线性代数（未完稿）

胡越　编著

第一版草稿说明

此讲义是在笔者 2023 年秋季学期所讲授的北京邮电大学理工类线性代数课程内容的基础上整理而成. 目前除了缺少习题, 部分内容仍待完善, 将在笔者后续课程内容基础上添加线性空间, 若当标准型以及部分矩阵论的内容. 希望读者能及时批评指正, 以让该讲义将来能以完善的面目问世.

在此感谢参与这门课程的北京邮电大学 2023 级数字媒体与设计艺术学院, 国际学院以及网络空间安全学院的学生.

胡越

2023 年 12 月 7 日

于北京邮电大学

目录

第 1 章　复数与向量

1.1　复数与平面向量

数系简单来说就是一些数构成的集合, 并且数之间带有一些运算. 我们已经了解的数系有:

$$\begin{aligned}\text{自然数}: &\quad \mathbb{N} = \{0, 1, 2, \ldots\}, \quad \text{加法, 乘法.}\\ \text{整数}: &\quad \mathbb{Z} = \{\ldots, -2, -1, 0, 1, 2, \ldots\}. \quad \text{加法, 减法, 乘法.}\\ \text{有理数}: &\quad \mathbb{Q} = \left\{ \frac{m}{n} \;\middle|\; m, n \in \mathbb{Z},\ n \neq 0 \right\}. \quad \text{加, 减, 乘, 除.}\end{aligned}$$

数系的发展, 一方面源于人类生产生活的需要, 例如自然数源于计数, 负数源于亏欠, 分数源于分割、分配等, 另一方面也源于数学自身理论发展的需要. 将有理数看成一条无限延伸的直线上的点, 也称为 "有理点", 这条直线称为实直线, 如图 1.1 所示. 在实直线上并不全是有理点. 记 $\sqrt{2}$ 为一个数 x, 满足 $x^2 = 2$, 则 x 在实直线上.

命题 1.1. $\sqrt{2}$ 不是有理数.

证明. 若 $\sqrt{2} = \dfrac{p}{q}$, 其中 $p, q \in \mathbb{Z}$, 且 p, q 互质, 即最大公因子为 1. 则 $2q^2 = p^2$, 于是 $2 \mid p$. 设 $p = 2p_1$, 则 $q^2 = 2p_1^2$, 于是 $2 \mid q$, 与 p, q 互质矛盾. □

我们当然认为直线上的所有点都对应一个数, 而那些填满有理点之间空隙的数就是无理数. 称实直线上的所有点, 或者所有数为实数, 记为 $\mathbb{R}$. 利用十进制表示, $\forall\, a \in \mathbb{R}$ 可以表

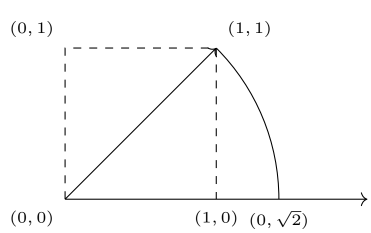


图 1.1: 实直线

示为小数的形式:

$$a = a_0.a_1a_2\cdots a_k\cdots = a_0 + \frac{a_1}{10} + \frac{a_2}{100} + \cdots + \frac{a_k}{10^k} + \cdots, \tag{1.1}$$

其中 $a_0 \in \mathbb{Z}$, $0 \le a_1, a_2, \cdots \le 9$.

命题 1.2. 设 $a \in \mathbb{R}$, 则 a 是有理数 $\Longleftrightarrow$ a 是有限小数或无限循环小数. 从而无理数是所有无限不循环小数.

证明. "$\Longleftarrow$": 若 a 是有限小数或无限循环小数, 可以设

$$a = a_0.a_1\cdots a_{s-1}a_s\cdots a_{s+\ell}a_s\cdots a_{s+\ell}\cdots,$$

其中若 a 是有限小数, 则 $a_s = \cdots = a_{s+\ell} = 0$. 于是

$$10^{s-1}a = a_0\cdots a_{s-1}.a_s\cdots a_{s+\ell}\cdots, \quad 10^{s+\ell-1}a = a_0\cdots a_{s+\ell}.a_s\cdots a_{s+\ell}\cdots,$$

于是 $(10^{s+\ell-1} - 10^{s-1})a$ 是一个整数, a 是有理数.
"$\Longrightarrow$": 设 $\dfrac{p}{q}$ 是有理数, 不妨设 $q > 0$. 反复利用带余除法, 有

$$\begin{aligned}
p &= a_0q + r_0, \quad a_0 \in \mathbb{Z},\ 0 \le r_0 < q,\\
10r_0 &= a_1q + r_1, \quad 0 \le a_1 \le 9,\ 0 \le r_1 < q,\\
&\ \ \vdots\\
10r_k &= a_{k+1}q + r_{k+1}, \quad 0 \le a_{k+1} \le 9,\ 0 \le r_{k+1} < q.
\end{aligned}$$

如果某个余数 $r_{k+1} = 0$, 则

$$\frac{p}{q} = a_0 + \frac{r_0}{q} = a_0 + \frac{a_1}{10} + \frac{r_1}{10q} = \cdots = a_0 + \frac{a_1}{10} + \frac{a_2}{100} + \cdots + \frac{a_k}{10^k},$$

即 $\dfrac{p}{q} = a_0.a_1a_2\cdots a_k$ 是有限小数. 如果 $\forall\ r_k \ne 0$, 则每个 r_k 都是 $1, \ldots, q$ 中某个数字, 故在 $\{r_0, \ldots, r_q\}$ 这 $q+1$ 个数中, 必定有两个数相同. 设 $r_{s-1} = r_{t-1}$, $t - s = \ell > 0$, 则由带余除法的过程可知对 $\forall\ k \ge s$, $a_{k+\ell} = a_k$, 从而

$$\frac{p}{q} = a_0.a_1a_2\cdots a_s\cdots a_{s+\ell}a_s\cdots a_{s+\ell}\cdots$$

是无限循环小数. □

代数的一个基本问题是 “解方程”. 但是很多方程在实数中没有解, 例如二次方程 $x^2+1=0$. 引入记号 i, 它是一个新的 “数”, 满足 $\mathrm{i}^2=-1$.

定义 1.3. 包含 $\mathbb{R}$ 与 i 并在加, 减, 乘, 除运算下封闭的最小的数系称为复数, 记为 $\mathbb{C}$.

设 $b\in\mathbb{R}$, 记 $\mathrm{i}b$ 为 i 和 b 的乘积, 则 $\mathrm{i}b$ 是方程 $x^2+b^2=0$ 的根. 具体的, 复数 $\mathbb{C}$ 中包含如下形式的数:

命题 1.4. $\mathbb{C}=\{a+\mathrm{i}b\mid a,b\in\mathbb{R}\}$.

证明. 一方面由复数的定义, $a+\mathrm{i}b$ 都在 $\mathbb{C}$ 中. 另一方面, 若 $z_1=a_1+\mathrm{i}b_1$, $z_2=a_2+\mathrm{i}b_2$, 则

$$
\begin{aligned}
z_1\pm z_2&=(a_1\pm a_2)+\mathrm{i}(b_1\pm b_2),\\
z_1\cdot z_2&=(a_1a_2-b_1b_2)+\mathrm{i}(a_1b_2+a_2b_1),\\
\frac{z_1}{z_2}&=\frac{a_1a_2+b_1b_2}{a_2^2+b_2^2}+\mathrm{i}\frac{a_2b_1-a_1b_2}{a_2^2+b_2^2},\quad(z_2\neq 0).
\end{aligned}
$$

从而 $\mathbb{C}$ 在四则运算下封闭. □

i 称为虚数单位. 若 $z=a+\mathrm{i}b$, a 称为 z 的实部, 记为 $\mathrm{Re}(z)$, b 称为 z 的虚部, 记为 $\mathrm{Im}(z)$. $\bar{z}=a-\mathrm{i}b$ 称为 z 的共轭, 则

$$z\bar{z}=(a+\mathrm{i}b)(a-\mathrm{i}b)=a^2+b^2. \tag{1.2}$$

定义 $z=$ 的模为 $|z|=\sqrt{a^2+b^2}$, 于是 $z=0\iff|z|=0$.

实数一一对应于直线上的点，而复数由它的实部和虚部决定, 自然的复数可以一一对应于平面上的点. 在平面上取直角坐标系 Oxy, 每个复数 $z=a+\mathrm{i}b$ 一一对应于平面上的坐标为 (a,b) 的点 P, 在这样的对应下平面称为复平面, 可以记作 $\mathbb{R}^2$. x 轴一一对应于实数, y 轴一一对应于纯虚数.

记 $\vec{OP}$ 为以原点 O 为起点, P 为终点的向量, 则复数 $z=a+\mathrm{i}b$ 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, 如图 1.2 所示.

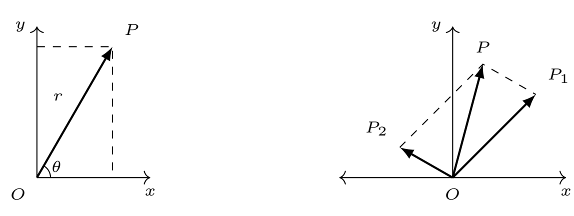


图 1.2: 向量的长度, 辐角和加法

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, 也称为 $\vec{OP}$ 的辐角, 为 θ. 则 z 可以写作如下三角形式

$$z = r(\cos\theta + \mathrm{i}\sin\theta). \tag{1.3}$$

在三角形式下, 可以看出复数乘法的几何含义:

$$\begin{aligned}
&z_1 = r_1(\cos\theta_1 + \mathrm{i}\sin\theta_1), \quad z_2 = r_2(\cos\theta_2 + \mathrm{i}\sin\theta_2),\\
&z_1 \cdot z_2 = r_1 r_2(\cos(\theta_1+\theta_2) + \mathrm{i}\sin(\theta_1+\theta_2)),\\
&\frac{z_1}{z_2} = \frac{r_1}{r_2}(\cos(\theta_1-\theta_2) + \mathrm{i}\sin(\theta_1-\theta_2)), \quad (z_2 \neq 0).
\end{aligned}$$

从而 $r(\cos\theta + \mathrm{i}\sin\theta)$ 乘以一个复数 z 就是把 z 对应的向量 $\vec{OP}$ 绕原点逆时针旋转 θ 角度, 再将长度拉伸 r 倍.

推论 1.5 (棣莫弗公式)**.** 对任意 $n \in \mathbb{Z}$, $z = r(\cos\theta + \mathrm{i}\sin\theta)$,

$$z^n = r^n(\cos n\theta + \mathrm{i}\sin n\theta). \tag{1.4}$$

例题 1.6. 求复数 $z = 1 + \cos\theta + \mathrm{i}\sin\theta$ $(-\pi \le \theta < \pi)$ 的三角形式.

解. 设 $z = r(\cos\alpha + \mathrm{i}\sin\alpha)$, 则

$$r^2 = |z|^2 = (1+\cos\theta)^2 + \sin^2\theta = 2 + 2\cos\theta = 4\cos^2\frac{\theta}{2}.$$

因为 $-\pi \le \theta < \pi$, $\cos\frac{\theta}{2} \ge 0$, 所以 $r = 2\cos\frac{\theta}{2}$.

进一步的, $\sin\alpha = \dfrac{\sin\theta}{r} = \sin\dfrac{\theta}{2}$. 因为 $\sin x$ 在 $[-\frac{\pi}{2}, \frac{\pi}{2}]$ 上单调递增, $\alpha = \frac{\theta}{2}$. 故 z 的三角形式为

$$z = 2\cos\frac{\theta}{2}(\cos\frac{\theta}{2} + \mathrm{i}\sin\frac{\theta}{2}). \qquad \square$$

在实数 $\mathbb{R}$ 中添加了 $x^2 + 1 = 0$ 的根 i, 我们得到了更大的复数系. 下面的定理说明, 对解决代数方程来说复数就足够了.

定理 1.7 (代数基本定理)**.** 复系数的 n 次代数方程

$$x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0 = 0, \quad a_0, \ldots, a_{n-1} \in \mathbb{C} \tag{1.5}$$

在 $\mathbb{C}$ 中存在一个根.

评注 1.8. 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》(柯斯特利金著) 的第六章, 3.3 节.

推论 1.9. 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 次代数方程恰好有 n 个根.

证明. 由代数基本定理, 设 $\alpha_1 \in \mathbb{C}$ 是 $f(x)$ 的一个根, 即 $f(\alpha_1) = 0$. 利用多项式的带余除法, 存在 $g(x)$ 以及常数 r 使得

$$f(x) = (x-\alpha_1)g(x) + r.$$

由 $f(\alpha_1) = 0$ 推出 $r = 0$, 即 $f(x) = (x-\alpha_1)g(x)$. 再对 $g(x)$ 进行分解即可. □

对任意 $n > 0$, 记

$$\varepsilon_k \coloneqq \cos\frac{2k\pi}{n} + \mathrm{i}\sin\frac{2k\pi}{n}, \quad k = 0, 1, \ldots, n-1, \tag{1.6}$$

则由棣莫弗公式有 $\varepsilon_k^n = 1$, 从而 $\varepsilon_0, \ldots, \varepsilon_{n-1}$ 是方程 $x^n - 1 = 0$ 的 n 个不同的根, 称为 *n 次单位根*. 由推论 1.9,

推论 1.10. (1.6) 中定义的 $\varepsilon_0, \ldots, \varepsilon_{n-1}$ 是方程 $x^n - 1 = 0$ 的所有根.

我们对数的认知已经扩大到了复数, 也了解到所研究的问题和限定的数的范围是密切相关的, 例如 $x^2 + 1 = 0$ 在 $\mathbb{R}$ 中无解而在 $\mathbb{C}$ 中有解. 为了后面叙述的方便和严谨, 需要对这门课程所限定的数的范围给一个精确的定义.

定义 1.11. 设 $\mathbb{F}$ 是由一些复数构成的集合, $0, 1 \in \mathbb{F}$, 并且 $\mathbb{F}$ 在加, 减, 乘, 除四则运算下封闭, 则称 $\mathbb{F}$ 是一个*数域*.

显然, 所有有理数, 实数, 复数构成的集合都是数域. 除此之外, 还有很多其他数域. 事实上, 任何数域 $\mathbb{F}$ 都包含 $\mathbb{Q}$.

例题 1.12. $\mathbb{Q}(\sqrt{2})=\{a+b\sqrt{2},a,b\in\mathbb{Q}\}$ 构成一个数域.

证明. 可以验证 $0,1\in\mathbb{Q}(\sqrt{2})$, 且 $\mathbb{Q}(\sqrt{2})$ 在加, 减, 乘, 除下封闭:

$$\begin{aligned}&(a+b\sqrt{2})\pm(c+d\sqrt{2})=(a\pm c)+(b\pm d)\sqrt{2},\\&(a+b\sqrt{2})(c+d\sqrt{2})=(ac+2bd)+(ad+bc)\sqrt{2},\\&\frac{a+b\sqrt{2}}{c+d\sqrt{2}}=\frac{ac-2bd}{c^2-2d^2}+\frac{-ad+bc}{c^2-2d^2}\cdot\sqrt{2},\ c+d\sqrt{2}\neq 0.\end{aligned}$$ □

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.

例题 1.13. 在平面直角坐标系 Oxy 中, 求将点 $P(a,b)$ 绕原点逆时针旋转 θ 角后得到的点 P' 的坐标.

解. 设 P' 的坐标为 (a',b'), 则 $\vec{OP'}$ 由 $\vec{OP}$ 逆时针旋转 θ 得到. 由复数乘法的几何意义,

$$a'+\mathrm{i}b'=(\cos\theta+\mathrm{i}\sin\theta)(a+\mathrm{i}b)=(\cos\theta a-\sin\theta b)+\mathrm{i}(\sin\theta a+\cos\theta b).$$

从而 P' 的坐标为 $(\cos\theta a-\sin\theta b,\sin\theta a+\cos\theta b)$. □

例题 1.14. 在平面直角坐标系 Oxy 中任取两点 $P_1(a_1,b_1)$, $P_2(a_2,b_2)$, 证明以 $\vec{OP_1}$, $\vec{OP_2}$ 为邻边的平行四边形的面积 $S=|a_1b_2-a_2b_1|$.

证明. 设 $\vec{OP_1}=(r_1\cos\theta_1,r_1\sin\theta_1)$, $\vec{OP_2}=(r_2\cos\theta_2,r_2\sin\theta_2)$, 则

$$\begin{aligned}S&=\left|\vec{OP_1}\right|\left|\vec{OP_2}\right||\sin(\theta_2-\theta_1)|=r_1r_2|\sin(\theta_2-\theta_1)|\\&=|r_1\cos\theta_1r_2\sin\theta_2-r_1\sin\theta_1r_2\cos\theta_2|=|a_1b_2-a_2b_1|.\end{aligned}$$ □

1.2 空间向量

我们已经看到了平面向量在解决几何问题中的应用, 这一节将介绍空间中的三维向量. 三维空间也记作 $\mathbb{R}^3$.

定义 1.15. 空间 $\mathbb{R}^3$ 中既有大小又有方向的量称为向量.

向量最初来源于物理, 也称作矢量, 物理学中研究的力, 加速度等就是向量. 一般我们用 $\vec{OA}$ 来表示从 O 到 A 的向量, 或简单用 $\mathbf{a}$ 来表示.

两个向量 $\mathbf{a}$, $\mathbf{b}$ 相等, 或者 $\mathbf{a}=\mathbf{b}$, 当且仅当它们的方向和大小相同. 所以如果两个向量经过平移重合, 则两个向量相同. 大小为 0 的向量称为零向量, 记为 $\mathbf{0}$, 零向量的方向是不确定的, 可以为任何方向. 任何向量 $\mathbf{a}$ 都有其负向量, 即大小相同但方向相反的向量, 记为 $-\mathbf{a}$. $\mathbf{0}$ 的负向量还是 $\mathbf{0}$. 向量的大小也称为其长度或模长, 记为 $|\mathbf{a}|$. 长度为 1 的向量称为单位向量.

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 $\mathbf{a}\perp\mathbf{b}$. 一般的, 向量 $\vec{OA}$ 与 $\vec{OB}$ 相交所形成的最小正角 (取值为 $[0,\pi]$) 称为它们的夹角.

空间中两个向量的加法和平面上一样, 按照平行四边形法则定义, 它满足如下性质:

$$\text{加法交换律}: \quad \mathbf{a}+\mathbf{b}=\mathbf{b}+\mathbf{a}, \tag{1.7}$$

$$\text{加法结合律}: \quad (\mathbf{a}+\mathbf{b})+\mathbf{c}=\mathbf{a}+(\mathbf{b}+\mathbf{c}), \tag{1.8}$$

$$\text{存在零向量}: \quad \mathbf{a}+\mathbf{0}=\mathbf{0}+\mathbf{a}=\mathbf{a}, \tag{1.9}$$

$$\text{存在负向量}: \quad \mathbf{a}+(-\mathbf{a})=(-\mathbf{a})+\mathbf{a}=\mathbf{0}. \tag{1.10}$$

向量的另一个基本运算是数与向量的乘法.

定义 1.16. 设 $\lambda\in\mathbb{R}$ 是一个实数, λ 乘以 $\mathbf{a}$ 是一个向量, 记为 $\lambda\mathbf{a}$, 称为 λ 与 $\mathbf{a}$ 的数乘. 其长度为 $|\lambda||\mathbf{a}|$, 其方向规定为:

- 当 $\lambda>0$ 时, $\lambda\mathbf{a}$ 与 $\mathbf{a}$ 同方向.
- 当 $\lambda<0$ 时, $\lambda\mathbf{a}$ 与 $\mathbf{a}$ 反方向.
- 当 $\lambda=0$ 时, $\lambda\mathbf{a}=\mathbf{0}$.

命题 1.17. 数乘满足如下性质:

$$\text{数乘对向量加法的分配律}: \quad \lambda(\mathbf{a}+\mathbf{b})=\lambda\mathbf{a}+\lambda\mathbf{b}, \tag{1.11}$$

$$\text{数乘对纯量加法的分配律}: \quad (\lambda+\mu)\mathbf{a}=\lambda\mathbf{a}+\mu\mathbf{a}, \tag{1.12}$$

$$\text{数乘结合律}: \quad (\lambda\mu)\mathbf{a}=\lambda(\mu\mathbf{a}), \tag{1.13}$$

$$\text{数乘单位元}: \quad 1\mathbf{a}=\mathbf{a}. \tag{1.14}$$

证明. 按照图 1.3 所示即可证明 (1.11). 其他性质按照定义即可. □

定义 1.18. 空间 $\mathbb{R}^3$ 中一组向量, 如果通过平移可以使它们同处一条直线, 称它们共线. 如果通过平移可以使它们同处一个平面, 称它们共面.

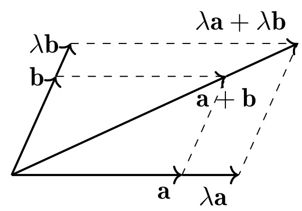


图 1.3: (1.11) 的证明

显然, 任何两个向量一定共面.

命题 1.19. (1) 向量 $\mathbf{a}$, $\mathbf{b}$ 共线 $\Longleftrightarrow$ 存在不全为 0 的实数 λ, μ 使得 $\lambda\mathbf{a}+\mu\mathbf{b}=\mathbf{0}$.

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

证明. (1) "$\Longrightarrow$": 若 $\mathbf{a}$, $\mathbf{b}$ 共线, 则 $\mathbf{a}$ 与 $\mathbf{b}$ 方向相同或相反. 若 $\mathbf{a}$ 和 $\mathbf{b}$ 同向, 则

$$\mathbf{b}=\frac{|\mathbf{b}|}{|\mathbf{a}|}\mathbf{a}\Longrightarrow\frac{|\mathbf{b}|}{|\mathbf{a}|}\mathbf{a}-\mathbf{b}=\mathbf{0}.$$

若 $\mathbf{a}$ 和 $\mathbf{b}$ 反向, 则

$$\mathbf{b}=-\frac{|\mathbf{b}|}{|\mathbf{a}|}\mathbf{a}\Longrightarrow\frac{|\mathbf{b}|}{|\mathbf{a}|}\mathbf{a}+\mathbf{b}=\mathbf{0}.$$

"$\Longleftarrow$": 若 $\lambda\mathbf{a}+\mu\mathbf{b}=\mathbf{0}$, 且不妨设 $\mu\neq 0$, 则 $\mathbf{b}=-\dfrac{\lambda}{\mu}\mathbf{a}$, 故 $\mathbf{b}$ 和 $\mathbf{a}$ 共线.

(2) "$\Longrightarrow$": 设 $\mathbf{a}$, $\mathbf{b}$, $\mathbf{c}$ 共面. 若其中两个向量共线, 不妨设为 $\mathbf{a}$ 和 $\mathbf{b}$, 则由 (1) 知存在不全为 0 的 λ, μ 使得 $\lambda\mathbf{a}+\mu\mathbf{b}=\mathbf{0}$, 于是 $\lambda\mathbf{a}+\mu\mathbf{b}+0\mathbf{c}=\mathbf{0}$, 即完成了证明. 现设三个向量两两不共线, 则 $\mathbf{a}$, $\mathbf{b}$, $\mathbf{c}$ 均不是零向量, 否则零向量和任何向量共线. 设 $\mathbf{a}=\vec{OA}$, $\mathbf{b}=\vec{OB}$, $\mathbf{c}=\vec{OC}$, 如图 1.4 所示.

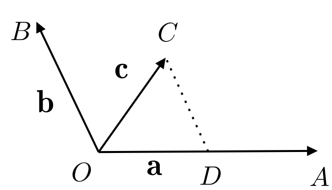


图 1.4: 三向量共面

因为 $\mathbf{a}$, $\mathbf{b}$ 不共线, 过 C 做 OB 的平行线与 OA 所在直线交于一点 D. 则 $\vec{OC}=\vec{OD}+\vec{DC}$. 因为 $\vec{OD}$ 和 $\mathbf{a}$ 共线, $\vec{DC}$ 和 $\mathbf{b}$ 共线, 故存在 $\lambda,\mu\in\mathbb{R}$ 使得 $\vec{OD}=\lambda\mathbf{a}$, $\vec{DC}=\mu\mathbf{b}$. 从而 $\mathbf{c}=\lambda\mathbf{a}+\mu\mathbf{b}$, 即 $\lambda\mathbf{a}+\mu\mathbf{b}-\mathbf{c}=\mathbf{0}$.

"$\Longleftarrow$": 若 $\lambda\mathbf{a}+\mu\mathbf{b}+\gamma\mathbf{c}=\mathbf{0}$, 且不妨设 $\gamma\neq 0$. 则 $\mathbf{c}=-\dfrac{\lambda}{\gamma}\mathbf{a}-\dfrac{\mu}{\gamma}\mathbf{b}$, 从而 $\mathbf{c}$ 在 $\mathbf{a}$, $\mathbf{b}$ 决定的平面上, 故 $\mathbf{a}$, $\mathbf{b}$, $\mathbf{c}$ 共面. □

定义 1.20. 设 $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n$ 为空间 $\mathbb{R}^3$ 中一组向量, $\lambda_1, \lambda_2, \ldots, \lambda_n$ 为一组实数, 称向量

$$\lambda_1\mathbf{a}_1 + \lambda_2\mathbf{a}_2 + \cdots + \lambda_n\mathbf{a}_n \tag{1.15}$$

为向量 $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n$ 的一个线性组合.

定义 1.21. 设 $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n$ 是空间 $\mathbb{R}^3$ 中一组向量, 如果存在不全为 0 的实数 $\lambda_1, \lambda_2, \ldots, \lambda_n$, 使得

$$\lambda_1\mathbf{a}_1 + \lambda_2\mathbf{a}_2 + \cdots + \lambda_n\mathbf{a}_n = \mathbf{0}, \tag{1.16}$$

则称 $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n$ 线性相关. 否则称它们线性无关.

结合命题 1.19, 我们有

- 一个向量线性相关 $\Longleftrightarrow$ 它是零向量.
- 两个向量线性相关 $\Longleftrightarrow$ 它们共线.
- 三个向量线性相关 $\Longleftrightarrow$ 它们共面.
- 那么四个向量呢?

回顾空间中的直角坐标系 $Oxyz$. 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 $\vec{OP} = x_1\mathbf{i} + x_2\mathbf{j} + x_3\mathbf{k}$. 于是直角坐标系 $Oxyz$ 的本质是, 在空间中选取了两两互相垂直的单位向量 $\mathbf{i}$, $\mathbf{j}$, $\mathbf{k}$, 空间中任何向量 $\mathbf{a}$ 可以唯一的表示成这三个向量的线性组合. 在这个观点下, 我们可以将坐标系推广到更一般的情形.

定理 1.22. 设 $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$ 为空间中三个不共面的向量, 则对任何向量 $\mathbf{a}$, 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. 设 $\mathbf{e}_1 = \vec{OA}$, $\mathbf{e}_2 = \vec{OB}$, $\mathbf{e}_3 = \vec{OC}$, $\mathbf{a} = \vec{OP}$, 如图 1.5 所示. 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 1.19 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

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

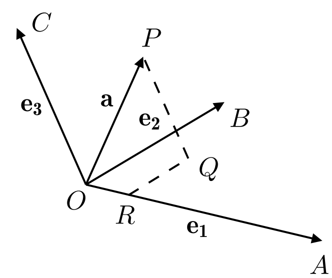


图 1.5: 仿射坐标系

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 1.19 知 $x_1=y_1$, $x_2=y_2$, $x_3=y_3$, 因此表示唯一. □

推论 1.23. 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 1.22 知存在 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 $x_1\mathbf{a}_1+x_2\mathbf{a}_2+x_3\mathbf{a}_3-\mathbf{a}_4=\mathbf{0}$, 从而 $\mathbf{a}_1$, $\mathbf{a}_2$, $\mathbf{a}_3$, $\mathbf{a}_4$ 线性相关. □

定义 1.24. 空间 $\mathbb{R}^3$ 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. 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 $[O;\mathbf{e}_1,\mathbf{e}_2,\mathbf{e}_3]$. 对任何向量 $\mathbf{a}$, 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.

引入仿射坐标系后, 向量可以用其坐标表示, 向量的运算可以化为在坐标上的运算:

$$\begin{aligned}\text{向量加法：}\quad &(x_1,x_2,x_3)+(y_1,y_2,y_3)=(x_1+y_1,x_2+y_2,x_3+y_3),\\ \text{数乘：}\quad &\lambda(x_1,x_2,x_3)=(\lambda x_1,\lambda x_2,\lambda x_3).\end{aligned}$$

例题 1.25. 设 $\mathbf{e}_1,\mathbf{e}_2,\mathbf{e}_3$ 为空间 $\mathbb{R}^3$ 的一组基.

$$\mathbf{a}=\mathbf{e}_1+\mathbf{e}_2+\mathbf{e}_3,\quad \mathbf{b}=\mathbf{e}_1-\mathbf{e}_2+\mathbf{e}_3,\quad \mathbf{c}=\mathbf{e}_1+\mathbf{e}_2-\mathbf{e}_3.$$

(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.

解. (1) $\mathbf{d} = -4\mathbf{e}_2 + 6\mathbf{e}_3$, 所以 $\mathbf{d}$ 在基 $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$ 下的坐标为 $(0, -4, 6)$.
(2) 需证明 $\mathbf{a}, \mathbf{b}, \mathbf{c}$ 线性无关, 即若存在 $x_1, x_2, x_3 \in \mathbb{R}$ 使得

$$x_1\mathbf{a} + x_2\mathbf{b} + x_3\mathbf{c} = \mathbf{0},$$

只能 $x_1 = x_2 = x_3 = 0$. 利用 $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$ 线性无关, 我们有

$$\begin{aligned}
x_1\mathbf{a} + x_2\mathbf{b} + x_3\mathbf{c} = \mathbf{0} &\Longleftrightarrow x_1(\mathbf{e}_1 + \mathbf{e}_2 + \mathbf{e}_3) + x_2(\mathbf{e}_1 - \mathbf{e}_2 + \mathbf{e}_3) + x_3(\mathbf{e}_1 + \mathbf{e}_2 - \mathbf{e}_3) = \mathbf{0} \\
&\Longleftrightarrow (x_1 + x_2 + x_3)\mathbf{e}_1 + (x_1 - x_2 + x_3)\mathbf{e}_2 + (x_1 + x_2 - x_3)\mathbf{e}_3 = \mathbf{0} \\
(\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3 \text{ 线性无关}) &\Longleftrightarrow \begin{cases} x_1 + x_2 + x_3 = 0 \\ x_1 - x_2 + x_3 = 0 \\ x_1 + x_2 - x_3 = 0 \end{cases} \Longleftrightarrow x_1 = x_2 = x_3 = 0.
\end{aligned}$$

这样就证明了 $\mathbf{a}, \mathbf{b}, \mathbf{c}$ 线性无关, 构成 $\mathbb{R}^3$ 的一组基. □

1.3 内积, 外积和混合积

这一节将介绍向量间的三种运算, 这些运算都有着几何或物理意义.

定义 1.26. 两个向量 $\mathbf{a}$, $\mathbf{b}$ 的内积 (或点积, 数量积) 为一个实数, 记为 $\mathbf{a} \cdot \mathbf{b}$, 等于两个向量的模长与两个向量夹角的余弦的乘积:

$$\mathbf{a} \cdot \mathbf{b} = |\mathbf{a}||\mathbf{b}| \cos\theta. \tag{1.17}$$

设 $\mathbf{a} = \vec{OA}$, $\mathbf{b} = \vec{OB}$, $\vec{OA}$ 在 $\vec{OB}$ 所在直线上的投影为向量 $\vec{OA'}$, 可以看出 $\vec{OA} \cdot \vec{OB} = \vec{OA'} \cdot \vec{OB}$, 如图 1.6 所示. 于是

命题 1.27. 一个向量 $\mathbf{a}$ 在另一个向量 $\mathbf{b}$ 上的投影向量为 $\dfrac{\mathbf{a} \cdot \mathbf{b}}{|\mathbf{b}|}\mathbf{b}$.

命题 1.28. 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 $\mathbf{a} = \mathbf{0}$.

证明. 按照图 1.7 所示即可证明性质 (2), 其他性质按照定义即可. □

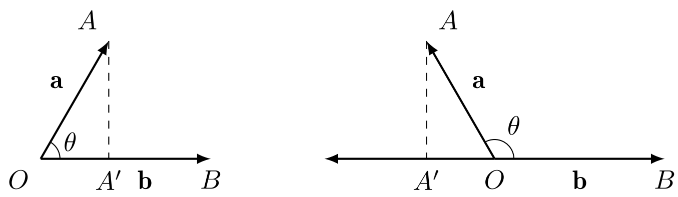


图 1.6: 内积

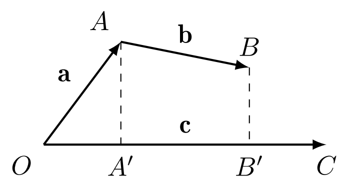


图 1.7: 内积性质 (2) 的证明

设 $[O;\mathbf{i},\mathbf{j},\mathbf{k}]$ 是一个直角坐标系, 因为 $\mathbf{i},\mathbf{j},\mathbf{k}$ 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 $\mathbf{a}=a_1\mathbf{i}+a_2\mathbf{j}+a_3\mathbf{k}$, $\mathbf{b}=b_1\mathbf{i}+b_2\mathbf{j}+b_3\mathbf{k}$, 则

$$\mathbf{a}\cdot\mathbf{b}=(a_1\mathbf{i}+a_2\mathbf{j}+a_3\mathbf{k})\cdot(b_1\mathbf{i}+b_2\mathbf{j}+b_3\mathbf{k})=a_1b_1+a_2b_2+a_3b_3.$$

于是两个向量的内积及夹角的余弦在直角坐标系下的公式为

$$(a_1,a_2,a_3)\cdot(b_1,b_2,b_3)=a_1b_1+a_2b_2+a_3b_3, \tag{1.18}$$
$$\cos\theta=\frac{\mathbf{a}\cdot\mathbf{b}}{|\mathbf{a}||\mathbf{b}|}=\frac{a_1b_1+a_2b_2+a_3b_3}{\sqrt{a_1^2+a_2^2+a_3^2}\sqrt{b_1^2+b_2^2+b_3^2}}. \tag{1.19}$$

定义 1.29. 向量 $\mathbf{a}$ 和 $\mathbf{b}$ 的外积 (或叉积, 向量积) 是一个向量, 记为 $\mathbf{a}\times\mathbf{b}$, 其模长为以 $\mathbf{a},\mathbf{b}$ 为邻边的平行四边形的面积:

$$|\mathbf{a}\times\mathbf{b}|=|a||b|\sin\theta. \tag{1.20}$$

方向与 $\mathbf{a}$, $\mathbf{b}$ 均正交, 且 $\mathbf{a},\mathbf{b},\mathbf{a}\times\mathbf{b}$ 构成右手系, 如图 1.8 所示.

设和 $\mathbf{b}$ 同方向的单位向量为 $\mathbf{b}_e$, 则 $\mathbf{a}\times\mathbf{b}=|\mathbf{b}|\mathbf{a}\times\mathbf{b}_e$. 求 $\mathbf{a}\times\mathbf{b}_e$ 的方法为: 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 $\mathbf{a}'$ 绕它的起点顺时针旋转 $\dfrac{\pi}{2}$ 就得到 $\mathbf{a}\times\mathbf{b}_e$, 如图 1.9 所示.

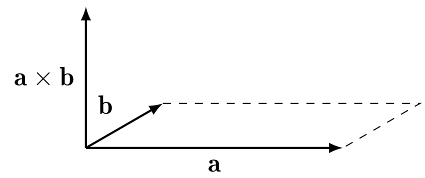


图 1.8: 外积

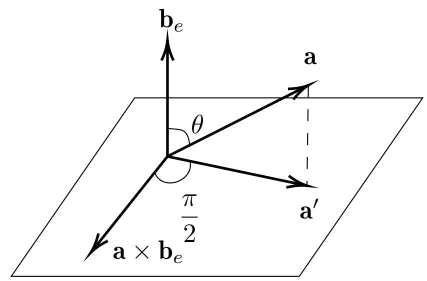


图 1.9: 求外积的方法

命题 1.30. 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) $(\lambda\mathbf{a})\times\mathbf{b}=\lambda(\mathbf{a}\times\mathbf{b})=\mathbf{a}\times(\lambda\mathbf{b})$.

证明. (1) 和 (3) 按照定义即可. 对 (2), 只需证明 $\mathbf{c}$ 是单位向量的情形, 按照图 1.10 所示即可. □

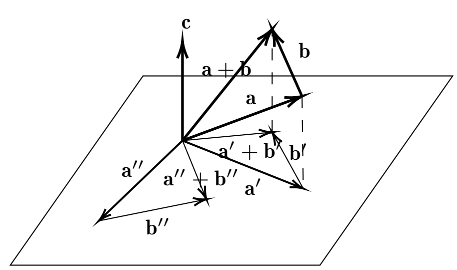


图 1.10: 外积性质 (2) 的证明

设 $[O;\mathbf{i},\mathbf{j},\mathbf{k}]$ 是一个直角坐标系, 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

$$\begin{aligned}\mathbf{a}\times\mathbf{b} &= (a_1\mathbf{i} + a_2\mathbf{j} + a_3)\times(b_1\mathbf{i} + b_2\mathbf{j} + b_3\mathbf{k})\\ &= (a_2b_3 - a_3b_2)\mathbf{i} + (a_3b_1 - a_1b_3)\mathbf{j} + (a_1b_2 - a_2b_1)\mathbf{k}.\end{aligned}\tag{1.21}$$

于是两个向量的外积在直角坐标系下的公式为:

$$(a_1, a_2, a_3)\times(b_1, b_2, b_3) = (a_2b_3 - a_3b_2, a_3b_1 - a_1b_3, a_1b_2 - a_2b_1).\tag{1.22}$$

例题 1.31. 求垂直于向量 $\mathbf{a} = (1, -1, 1)$ 和 $\mathbf{b} = (1, 2, -1)$ 的单位向量.

解. $\mathbf{a}\times\mathbf{b}$ 就是和 $\mathbf{a}, \mathbf{b}$ 均垂直的向量, 由 (1.22) 知 $\mathbf{a}\times\mathbf{b} = (-1, 2, 3)$. 于是所求为

$$\frac{\mathbf{a}\times\mathbf{b}}{|\mathbf{a}\times\mathbf{b}|} = \left(\frac{-1}{\sqrt{14}}, \frac{2}{\sqrt{14}}, \frac{3}{\sqrt{14}}\right).$$ □

定义 1.32. 对任何三个向量 $\mathbf{a}, \mathbf{b}, \mathbf{c}$, $(\mathbf{a}\times\mathbf{b})\cdot\mathbf{c}$ 称为它们的混合积 (或三重积).

设 V 是以 $\mathbf{a}, \mathbf{b}, \mathbf{c}$ 为棱的平行六面体的体积, 如图 1.11 所示. 当 φ 为锐角时, $\mathbf{a}, \mathbf{b}, \mathbf{c}$ 构成右手系, $V = (\mathbf{a}\times\mathbf{b})\cdot\mathbf{c}$; 当 φ 为钝角时, $\mathbf{a}, \mathbf{b}, \mathbf{c}$ 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".

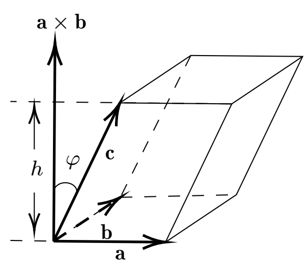


图 1.11: 混合积的几何意义

由内积和外积在直角坐标系下的公式 (1.18)(1.22) 可以得到混合积在直角坐标系下的公式. 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

$$(\mathbf{a}\times\mathbf{b})\cdot\mathbf{c} = (a_2b_3 - a_3b_2)c_1 + (a_3b_1 - a_1b_3)c_2 + (a_1b_2 - a_2b_1)c_3.\tag{1.23}$$

回忆例 1.14, 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

$$\begin{vmatrix} a_1 & a_2\\ b_1 & b_2\end{vmatrix} = a_1b_2 - a_2b_1.\tag{1.24}$$

于是混合积的公式可以写为

$$(\mathbf{a}\times\mathbf{b})\cdot\mathbf{c}=\begin{vmatrix}a_2 & a_3\\ b_2 & b_3\end{vmatrix}c_1-\begin{vmatrix}a_1 & a_3\\ b_1 & b_3\end{vmatrix}c_2+\begin{vmatrix}a_1 & a_2\\ b_1 & b_2\end{vmatrix}c_3. \tag{1.25}$$

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

$$\begin{vmatrix}a_1 & a_2 & a_3\\ b_1 & b_2 & b_3\\ c_1 & c_2 & c_3\end{vmatrix}=\begin{vmatrix}a_2 & a_3\\ b_2 & b_3\end{vmatrix}c_1-\begin{vmatrix}a_1 & a_3\\ b_1 & b_3\end{vmatrix}c_2+\begin{vmatrix}a_1 & a_2\\ b_1 & b_2\end{vmatrix}c_3. \tag{1.26}$$

比较 (1.21), 向量外积的公式可以形式的写为

$$\mathbf{a}\times\mathbf{b}=\begin{vmatrix}a_1 & a_2 & a_3\\ b_1 & b_2 & b_3\\ \mathbf{i} & \mathbf{j} & \mathbf{k}\end{vmatrix}. \tag{1.27}$$

推论 1.33. (1) $\mathbf{a},\mathbf{b},\mathbf{c}$ 共面 $\iff$ $\mathbf{a},\mathbf{b},\mathbf{c}$ 的行列式为 0.

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

证明. (1) 由混合积的几何意义及公式, $\mathbf{a},\mathbf{b},\mathbf{c}$ 共面 $\iff$ 以 $\mathbf{a},\mathbf{b},\mathbf{c}$ 为棱的平行六面体的体积为 0 $\iff$ $\mathbf{a},\mathbf{b},\mathbf{c}$ 的行列式为 0.

(2) 在轮换下不会改变 $\mathbf{a},\mathbf{b},\mathbf{c}$ 是左手系还是右手系, 于是三者均为以 $\mathbf{a},\mathbf{b},\mathbf{c}$ 为棱的平行六面体的有向体积. □

例题 1.34. 求以 $A(1,2,3)$, $B(2,1,4)$, $C(1,5,9)$, $D(2,2,2)$ 为顶点的四面体的体积.

证明. 设 $\mathbf{a}=\vec{AB}=(1,-1,1)$, $\mathbf{b}=\vec{AC}=(0,3,6)$, $\mathbf{c}=\vec{AD}=(1,0,-1)$. 所求四面体体积 V 是以 $\mathbf{a},\mathbf{b},\mathbf{c}$ 为棱的平行六面体体积的六分之一, 所以

$$V=\frac{1}{6}|(\mathbf{a}\times\mathbf{b})\cdot\mathbf{c}|=\frac{1}{6}\left|\begin{vmatrix}1 & -1 & 1\\ 0 & 3 & 6\\ 1 & 0 & -1\end{vmatrix}\right|=2.$$ □

1.4 直线与平面

本节默认在空间中取定一组直角坐标系, 并利用向量研究空间中的直线和平面.

设 L 是空间中一条直线, 在 L 上任取不同的点 A,B, 则非零向量 $\vec{AB}$ 称为 L 的一个方向向量. 若已知 L 上一点 M 的坐标 $M(x_0,y_0,z_0)$ 和它的一个方向向量 $\mathbf{u}=(u_1,u_2,u_3)$. 则

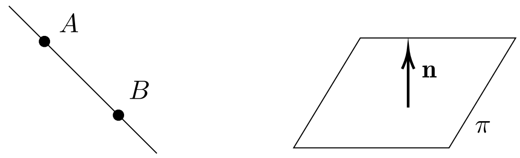


图 1.12: 直线和平面

对 L 上任意点 $P(x,y,z)$, 存在 $t \in \mathbb{R}$ 使得 $\vec{MP} = t\mathbf{u}$, 于是

$$x = x_0 + tu_1, \quad y = y_0 + tu_2, \quad z = z_0 + tu_3, \tag{1.28}$$

称为 L 的参数方程. 在 (1.28) 中消去 t 得到

$$\frac{x - x_0}{u_1} = \frac{y - y_0}{u_2} = \frac{z - z_0}{u_3}, \tag{1.29}$$

称为 L 的标准方程 (或点向式方程).

给定空间中任何一点 $M(x_0, y_0, z_0)$, 以及一个非零向量 $\mathbf{n} = (A, B, C)$, 存在唯一的平面 π 过 M, 且与 $\mathbf{n}$ 所在直线垂直. $\mathbf{n}$ 称为平面 π 的一个法向量. 对空间中任意点 $P(x,y,z)$, $P \in \pi \iff \vec{MP} \perp \mathbf{n}$, 从而

$$A(x - x_0) + B(y - y_0) + C(z - z_0) = 0, \tag{1.30}$$

称为平面的点法式方程. (1.30) 可简化为

$$Ax + By + Cz + D = 0, \tag{1.31}$$

称为平面的一般方程.

空间中不共线的三个点 $M_1(x_1, y_1, z_1)$, $M_2(x_2, y_2, z_2)$, $M_3(x_3, y_3, z_3)$ 也决定一个平面 π, 且 $P(x,y,z) \in \pi \iff \vec{M_1M_2}, \vec{M_1M_3}, \vec{M_1P}$ 共面. 由推论 1.33, 等价于

$$\begin{vmatrix} x_2 - x_1 & y_2 - y_1 & z_2 - z_1 \\ x_3 - x_1 & y_3 - y_1 & z_3 - z_1 \\ x - x_1 & y - y_1 & z - z_1 \end{vmatrix} = 0, \tag{1.32}$$

称为平面的三点式方程.

例题 1.35. (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 L: $x + 1 = 2y + 3 = 3z - 5$ 的平面方程.

解. (1) 由平面的三点式方程 (1.32), 以及

$$\begin{vmatrix} 1 & 3 & -1 \\ 2 & 3 & 4 \\ x-1 & y+2 & z-1 \end{vmatrix} = 15x - 6y - 3z - 24,$$

所求平面的一般方程为 $5x - 2y - z - 8 = 0$.
(2) L 的一个方向向量为 $\mathbf{u} = (6,3,2)$, 且 L 上一点为 $P(0,-1,2)$. 设所求平面的一个法向量为 $\mathbf{n} = (A,B,C)$, 则 $\mathbf{n}$ 和 $\mathbf{u}, \vec{MP}$ 均垂直, 故可取

$$\mathbf{n} = \mathbf{u} \times \vec{MP} = (6,3,2) \times (-1,-2,1) = (7,-8,-9).$$

设平面的一般方程为 $7x - 8y - 9z + D = 0$, 将 $M(1,1,1)$ 代入得到 $D = 10$, 故所求平面方程为 $7x - 8y - 9z + 10 = 0$. □

任何一条直线 L 可以看作两相交平面的交线, 故两个平面方程的联立方程组

$$\begin{cases} A_1x + B_1y + C_1z + D_1 = 0, \\ A_2x + B_2y + C_2z + D_2 = 0. \end{cases} \tag{1.33}$$

就给出了 L 的方程, 称为 L 的*一般方程*.

例题 1.36. 已知直线 L 的一般方程为 $\begin{cases} 2x - 3y - z - 3 = 0, \\ 4x - 6y + 5z + 1 = 0. \end{cases}$ 求它的标准方程和参数方程.

解. 设 L 的一个方向向量为 $\mathbf{u}$, 则 $\mathbf{u}$ 和两个平面的法向量 $\mathbf{n}_1 = (2,-3,-1)$, $\mathbf{n}_2 = (4,-6,5)$ 均垂直. 因为 $\mathbf{n}_1 \times \mathbf{n}_2 = (-21,-14,0)$, 可取 $\mathbf{u} = (3,2,0)$. 在一般方程中令 $y = 0$, 得到直线 L 上的一个点 $M(1,0,-1)$. 故 L 的标准方程和参数方程分别为

$$\frac{x-1}{3} = \frac{y}{2} = \frac{z+1}{0},$$
$$x = 3t+1, \quad y = 2t, \quad z = -1.$$ □

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, 如图 1.13 所示. 则点 P 到 L 的距离为

$$d = \left|\vec{BP}\right| = \left|\vec{AP}\right| \sin\theta = \frac{\left|\vec{AP} \times \mathbf{u}\right|}{|\mathbf{u}|}. \tag{1.34}$$

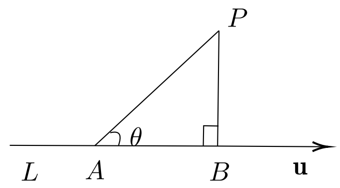


图 1.13: 点到直线的距离

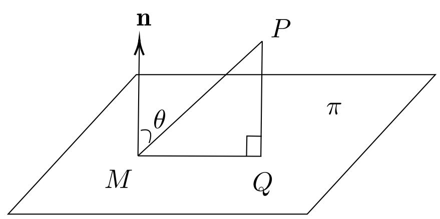


图 1.14: 点到平面的距离

例题 1.37. 求 $P(1,1,2)$ 到直线 $L\colon \dfrac{x-2}{2}=\dfrac{y-3}{1}=\dfrac{z}{2}$ 的距离.

解. L 上一点为 $M(2,3,0)$, 方向向量为 $\mathbf{u}=(2,1,2)$, 故 $d=\dfrac{\left|\vec{MP}\times\mathbf{u}\right|}{|\mathbf{u}|}=3$. □

设平面 π 的一般方程为 $Ax+By+Cz+D=0$, 法向量为 $\mathbf{n}=(A,B,C)$, 并任取平面上一点 $M(x_0,y_0,z_0)$. 设 P 为空间中任意点, 过 P 作平面 π 的垂线, 垂足为 Q, 如图 1.14 所示. 则 P 到 π 的距离为

$$\begin{aligned} d=\left|\vec{QP}\right|=\frac{\left|\mathbf{n}\cdot\vec{MP}\right|}{|\mathbf{n}|}&=\frac{|A(x-x_0)+B(y-y_0)+C(z-z_0)|}{\sqrt{A^2+B^2+C^2}}\\ &=\frac{|Ax+By+Cz+D|}{\sqrt{A^2+B^2+C^2}}. \end{aligned} \tag{1.35}$$

设 L_1,L_2 是空间中的两条直线, 方向向量分别为 $\mathbf{u}_1,\mathbf{u}_2$, 且分别取两条直线上的一个点 M_1,M_2. 则 L_1,L_2 的位置关系的判别方法为:

1. L_1 和 L_2 共面 $\Longleftrightarrow$ 向量 $\vec{M_1M_2},\mathbf{u}_1,\mathbf{u}_2$ 的行列式为 0.

2. L_1 和 L_2 异面 $\Longleftrightarrow$ 向量 $\vec{M_1M_2},\mathbf{u}_1,\mathbf{u}_2$ 的行列式非 0.

3. 两直线的夹角的余弦为 $\cos\theta=\dfrac{|\mathbf{u}_1\cdot\mathbf{u}_2|}{|\mathbf{u}_1||\mathbf{u}_2|}$.

例题 1.38. 求 a 使得直线 $L_1\colon x-a=\dfrac{y-3}{2}=z-2$ 和 $L_2\colon x=2y=2z$ 相交.

证明. 直线 L_1 的方向向量 $\mathbf{u}_1 = (1,2,1)$, 过点 $M_1 = (a,3,2)$; 直线 L_2 的方向向量 $\mathbf{u}_2 = (2,1,1)$, 过点 $M_2 = (0,0,0)$. 因为 L_1 和 L_2 不平行, 故它们相交 $\Longleftrightarrow$ 它们共面, 所以

$$\begin{vmatrix} a & 3 & 2 \\ 1 & 2 & 1 \\ 2 & 1 & 1 \end{vmatrix} = 0 \Longrightarrow a = 3.$$ □

设空间中两个平面 π_1, π_2 的一般方程分别为 $A_ix + B_iy + C_iz + D_i = 0$ $(i = 1,2)$, 法向量为 $\mathbf{n}_1, \mathbf{n}_2$, 则 π_1, π_2 的位置关系的判别方法为:

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 $\cos\theta = \dfrac{|\mathbf{n}_1 \cdot \mathbf{n}_2|}{|\mathbf{n}_1||\mathbf{n}_2|}$.

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

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

例题 1.39. 求直线 $L\colon x = 2y = 3z$ 和平面 $\pi\colon x + 2y + 3z = 4$ 的夹角和交点.

解. L 的方向向量 $\mathbf{u} = (1, \frac{1}{2}, \frac{1}{3})$, π 的法向量 $\mathbf{n} = (1,2,3)$, 故

$$\sin\theta = \frac{|\mathbf{u} \cdot \mathbf{n}|}{|\mathbf{u}||\mathbf{n}|} = \frac{18}{7\sqrt{14}}, \quad \theta = \arcsin\frac{18}{7\sqrt{14}}.$$

L 的参数方程为 $x = t$, $y = \dfrac{t}{2}$, $z = \dfrac{t}{3}$, 代入平面方程得到 $t = \dfrac{4}{3}$, 故交点为 $(\dfrac{4}{3}, \dfrac{2}{3}, \dfrac{4}{9})$. □

设 L 是两个平面

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

的相交直线, 则过 L 的任意平面的方程可以写为如下形式:

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

其中 λ, μ 不同时为 0, 称为平面束方程.

例题 1.40. 求直线 L: $\begin{cases} x-y+z+2=0, \\ x+2y-z+5=0. \end{cases}$ 在平面 π: $2x+y-3z+1=0$ 上的投影直线的方程.

解. 设投影直线为 L_1, 可以验证 L 不在 π 上, 故 L 和 L_1 不重合, 则它们决定一个平面 π_1. 一方面 π_1 是过 L 的平面, 设其方程为

$$\lambda(x-y+z+2)+\mu(x+2y-z+5)=0.$$

另一方面 π_1 和 π 垂直, 推出 $2\lambda=7\mu$, 得到 π_1 的方程 $9x-3y+5z+24=0$. 因为 L_1 是 π 和 π_1 的相交直线, 故 L_1 的一般方程为

$$\begin{cases} 2x+y-3z+1=0, \\ 9x-3y+5z+24=0. \end{cases}$$ □

1.5 高维向量

现实生活中, 很多量无法用三维向量中表示, 例如质点在一段时间内的运动轨迹, 物体某一点的速度等, 有必要将三维向量推广到高维向量.

在三维空间中, 取定一个坐标系, 一个向量可以用其坐标, 或者一个三元有序数组来表示, 且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 1.41. 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

$$\mathbf{a}=(a_1,a_2,\ldots,a_n), \quad a_1,a_2,\ldots,a_n\in\mathbb{F}. \tag{1.36}$$

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 $\mathbb{F}^n$ 上可以定义向量加法, 数乘, 零向量以及负向量:

$$\text{向量加法}: \quad (a_1,\ldots,a_n)+(b_1,\ldots,b_n)=(a_1+b_1,\ldots,a_n+b_n). \tag{1.37}$$

$$\text{数乘}: \quad \lambda(a_1,\ldots,a_n)=(\lambda a_1,\ldots,\lambda a_n), \quad \lambda\in\mathbb{F}. \tag{1.38}$$

$$\text{零向量}: \quad \mathbf{0}=(0,\ldots,0). \tag{1.39}$$

$$\text{负向量}: \quad -(a_1,\ldots,a_n)=(-a_1,\ldots,-a_n). \tag{1.40}$$

容易验证性质 (1.7)–(1.14) 仍然成立.

一个 n 维向量也可以写成列向量的形式:

$$\mathbf{a} = \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix}, \quad a_1, a_2, \ldots, a_n \in \mathbb{F}. \tag{1.41}$$

为了不占用空间, 我们也用 $[a_1, \ldots, a_n]$ 表示列向量.

定义 1.42. 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.

定义 1.43. 给定一组 n 维向量 $\mathbf{a}_1, \ldots, \mathbf{a}_m$, 如果存在 $\mathbb{F}$ 中一组不全为 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. 否则称它们线性无关.

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, 则任何量

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

可以唯一的表示为 $\boldsymbol{\epsilon}_1, \ldots, \boldsymbol{\epsilon}_m$ 的线性组合 $\mathbf{a} = a_1 \boldsymbol{\epsilon}_1 + \cdots + a_m \boldsymbol{\epsilon}_m$. 我们称

定义 1.44. $\boldsymbol{\epsilon}_1, \ldots, \boldsymbol{\epsilon}_m$ 为 $\mathbb{F}^m$ 的标准基.

我们已经研究了平面 $\mathbb{R}^2$, 以及空间 $\mathbb{R}^3$ 中的一些几何问题, 包含判断向量的共线和共面, 计算向量的长度和向量间的夹角, 计算平行四边形的面积和平行六面体的体积, 确定直线和平面的方程等. 除此之外, 还有一类重要的几何问题是研究平面或空间中的几何变换, 例如我们熟知的伸缩变换, 旋转变换以及反射变换, 我们将在第 4 章研究这些几何变换. 这些几何概念和问题都和“线性结构”相关, 即一组向量的线性组合和线性相关性, 称为线性几何. 研究线性几何的基本方法是取平面或空间的一组坐标系, 或者一组基向量, 将向量用坐标表示, 然后将问题转化为坐标上的计算或解一个代数方程.

线性代数这门课程的研究内容就是 n 维向量空间中的线性几何，其核心方法和思想就是用代数工具研究几何问题. 作为共线和共面, 以及直线和平面在 n 维向量空间 $\mathbb{F}^n$ 中的推广, 我们将在第 2.2 节研究 $\mathbb{F}^n$ 中向量间的线性相关性以及 $\mathbb{F}^n$ 的子空间; 作为平行六面体的有向体积, 或混合积的推广, 我们将在第 3.3 节研究 $\mathbb{F}^n$ 中 n 个向量的行列式; 作为二维及三维几何变换的推广, 我们将在第 4 章研究 n 维向量空间之间的线性变换; 最后, 我们在第 5 中将内积推广到 n 维向量空间, 以定义 n 维向量的长度和向量间的夹角, 或统称为度量.

我们已经看到线性几何的很多问题在坐标下都转化为解一个线性方程组. 给定一个三元一次方程组

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 = b_1 \\ a_{21}x_1 + a_{22}x_2 + a_{23}x_3 = b_2 \\ a_{31}x_1 + a_{32}x_2 + a_{33}x_3 = b_3 \end{cases} \tag{1.42}$$

一方面方程组 (1.42) 中的每个方程是一个平面的方程, 故方程组的所有解就是三个平面的所有交点. 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 (1.42) 可以写成

$$x_1\mathbf{a}_1 + x_2\mathbf{a}_2 + x_3\mathbf{a}_3 = \mathbf{b}. \tag{1.43}$$

故方程组的所有解就是将 $\mathbf{b}$ 用 $\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3$ 线性表示的所有方式.

如何求解一般的线性方程组就是第 2 章的内容. 我们在第 2.3 节引入了矩阵, 并在第 3 章介绍了矩阵的各种运算. 我们将看到, 矩阵就是研究线性几何问题的主要代数工具, 无论是解线性方程组, 计算行列式, 还是化简线性变换, 都可以转化为矩阵运算!

第 2 章　线性方程组

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 $x_1, \ldots, x_m$ 称为线性方程组的未知元 (或变量), a_{ij} $(1 \le i \le n, 1 \le j \le m)$, b_k $(1 \le k \le n)$ 均属于一个固定的数域 $\mathbb{F}$, 分别称为线性方程组的系数和常数项. 一组 $\mathbb{F}$ 中的数 $\{s_1, \ldots, s_m\}$ 称为线性方程组的一组解, 如果在 (2.1) 中令 x_i $(1 \le i \le m)$ 取值为 s_i $(1 \le i \le m)$ 时各个方程成为恒等式. 线性方程组的所有解构成的集合称为它的解集.

我们对 (2.1) 添加如下假设: 对任意 $1 \le j \le m$, 存在 $1 \le i \le n$ 使得 $a_{ij} \ne 0$. 换句话说, 我们假定未知元 $x_1, \ldots, x_m$ 在线性方程组中均真实出现.

对二元一次或三元一次线性方程组, 我们已经知道可以用消元法将方程组化为简单的形式来求解. 对一般的线性方程组是一样的, 先看一些例子.

例题 2.1. 求解线性方程组

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

解. 将第一个方程的 -2 倍和 -3 倍分别加到第二、三个方程, 可以将后两个方程中的未知元 x_1 消去, 化为一个二元一次方程组求解出 x_2, x_3, 然后再将 x_2, x_3 的值代入第一个方程

得到 x_1 的值. 将三个分别方程标记为 (1),(2),(3), 则具体求解过程为:

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

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

从最后一个线性方程组可以看出解为 $x_1=1$, $x_2=x_3=0$. □

例题 2.2. 求解线性方程组

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

证明.

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

从而原线性方程组无解. □

例题 2.3.

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

证明.

$$\begin{cases} x_1+2x_2+3x_3+4x_4=-3 \\ x_1+2x_2-5x_4=1 \\ 2x_1+4x_2-3x_3-19x_4=6 \\ 3x_1+6x_2-3x_3-24x_4=7 \end{cases} \xrightarrow[(1)\times-3+(4)]{\substack{(1)\times-1+(2) \\ (1)\times-2+(3)}} \begin{cases} x_1+2x_2+3x_3+4x_4=-3 \\ -3x_3-9x_4=4 \\ -9x_3-27x_4=12 \\ -12x_3-36x_4=16 \end{cases}$$

$$\xrightarrow[(2)\times-4+(4)]{(2)\times-3+(3)}\begin{cases}x_1+2x_2+3x_3+4x_4=-3\\ \qquad\qquad -3x_3-9x_4=4\end{cases}\xrightarrow{(2)\times1+(1)}\begin{cases}x_1+2x_2\qquad\qquad -5x_4=1\\ \qquad\qquad -3x_3-9x_4=4\end{cases}$$

此时 x_2,x_4 可以在 $\mathbb{F}$ 中任意取值, 且每一组 x_2,x_4 的取值就决定了 x_1,x_3 的值. 所以原方程组有无穷多组解. 为了写出所有的解, 设 $x_2=\lambda$, $x_4=\mu$, 则所有解可以表示为

$$x_1=-2\lambda+5\mu+1,\ x_2=\lambda,\ x_3=-3\mu-\frac{4}{3},\ x_4=\mu,\qquad \lambda,\mu\in\mathbb{F}.$$

进一步的, 将一组解写成列向量 $[x_1,x_2,x_3,x_4]$, 则

$$\begin{pmatrix}x_1\\x_2\\x_3\\x_4\end{pmatrix}=\begin{pmatrix}-2\lambda+5\mu+1\\ \lambda\\ -3\mu-\dfrac{4}{3}\\ \mu\end{pmatrix}=\lambda\begin{pmatrix}-2\\1\\0\\0\end{pmatrix}+\mu\begin{pmatrix}5\\0\\-3\\1\end{pmatrix}+\begin{pmatrix}1\\0\\-\dfrac{4}{3}\\0\end{pmatrix},\qquad \lambda,\mu\in\mathbb{F}.$$

这样将每一组解表示为了一些向量的线性组合的形式. □

以后我们都将用列向量 $\mathbf{x}=[x_1,\dots,x_n]$ 来表示线性方程组的一组解, 也称为解向量. 回顾消元法的过程, 我们对线性方程组的操作可以归纳为如下三类初等变换:

(1) 交换两个方程的位置.

(2) 将某个方程乘以一个非零常数.

(3) 将某个方程的一个倍数加到另一个方程.

并且我们默认了如下事实:

命题 2.4. 经过初等变换后的线性方程组和原方程组有相同的解.

证明. 记原方程组为 $\mathbf{LS}$, 经过某类初等变换后的方程为 $\mathbf{LS}'$. 显然, 若 $\mathbf{x}$ 是 $\mathbf{LS}$ 的解, 则 $\mathbf{x}$ 也是 $\mathbf{LS}'$ 的解. 另一方面, 注意到初等变换的操作是可逆的, 即

(1) 若 $\mathbf{LS}'$ 是将 $\mathbf{LS}$ 的第 i 和第 j 个方程交换得到的, 则将 $\mathbf{LS}'$ 的第 i 和第 j 个方程交换即回到原方程组 $\mathbf{LS}$.

(2) 若 $\mathbf{LS}'$ 是将 $\mathbf{LS}$ 的第 i 个方程乘以 λ $(\lambda\neq0)$ 得到的, 则将 $\mathbf{LS}'$ 的第 i 个方程乘以 $\dfrac{1}{\lambda}$ 即回到原方程组 $\mathbf{LS}$.

(3) 若 $\mathbf{LS}'$ 是将 $\mathbf{LS}$ 的第 j 个方程的 λ 倍加到第 i 个方程得到的, 则将 $\mathbf{LS}'$ 的第 j 个方程的 $-\lambda$ 倍加到第 i 个方程即回到原方程组 $\mathbf{LS}$.

就证明了 $\mathbf{LS}$ 和 $\mathbf{LS}'$ 有相同的解. □

接下来给出一般线性方程组的消元算法.

定理 2.5. 可以通过初等变换将方程组 (2.1) 化为如下阶梯形:

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

其中 $1 = s_1 < s_2 < \cdots < s_r \leq m$, $a'_{11}, a'_{2s_2}, \ldots, a'_{rs_r}$ 均不为 0, 且对 $\forall\ 1 \leq i \leq r$, 未知元 x_{s_i} 只在第 i 个方程出现.

证明. 结论对含有两个未知元的线性方程组显然成立, 接下来用归纳法, 对 $\forall\ k < n$, 假设含有 k 个未知元的线性方程组都可以通过初等变换化为阶梯形. 对线性方程组 (2.1), 按照我们的假定, 存在 i 使得 $a_{i1} \neq 0$, 通过交换两个方程, 总可以假定在第 1 个方程中 x_1 出现, 或者 $a_{11} \neq 0$. 此时可以通过将第 1 个方程的 $-\dfrac{a_{21}}{a_{11}}, \ldots, -\dfrac{a_{n1}}{a_{11}}$ 倍分别加到第 $2, \ldots, n$ 个方程, 将后 $n-1$ 个方程中的未知元 x_1 消去, 得到:

$$\begin{cases} a'_{11}x_1 + a'_{12}x_2 + \cdots + a'_{1m}x_m = b'_1 \\ \qquad a'_{22}x_2 + \cdots + a'_{2m}x_m = b'_2 \\ \qquad\quad \vdots \qquad\qquad \vdots \qquad \vdots \\ \qquad a'_{n2}x_2 + \cdots + a'_{nm}x_m = b'_n \end{cases} \tag{2.3}$$

利用归纳假设, 通过初等变换可以将 (2.3) 的后 $n-1$ 个方程构成的线性方程组化为阶梯形:

$$\begin{cases} a'_{11}x_1 + \cdots + a'_{1s_2}x_{s_2} + \cdots + a'_{1s_r}x_{s_r} + \cdots + a'_{1m}x_m = b'_1 \\ \qquad a'_{2s_2}x_{s_2} + \cdots + 0 + \cdots + a'_{2m}x_m = b'_2 \\ \qquad\qquad \ddots \quad \vdots \quad \vdots \quad \vdots \\ \qquad\qquad a'_{rs_r}x_{s_r} + \cdots + a'_{rm}x_m = b'_r \\ \qquad\qquad\qquad 0 = b'_{r+1} \\ \qquad\qquad\qquad \vdots \quad \vdots \\ \qquad\qquad\qquad 0 = b'_n \end{cases} \tag{2.4}$$

最后在 (2.4) 中, 分别将第 2 个方程的 $-\frac{a'_{1s_2}}{a'_{2s_2}}$ 倍, ..., 第 r 个方程的 $-\frac{a'_{1s_r}}{a'_{rs_r}}$ 倍加到第 1 个方程, 就将第 1 个方程中的未知元 $x_{s_2}, \dots, x_{s_r}$ 消去, 化为阶梯形 (2.2). □

化为阶梯形 (2.2) 后, 可以给出方程组是否有解的判定条件:

命题 2.6. (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) 有无穷多解当且仅当 $r < m$ 并且 $b'_{r+1} = \cdots = b'_n = 0$.

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

证明. 在 (2.2) 中, 若存在 $r+1 \le j \le n$ 使得 $b'_j \ne 0$, 则在方程组中有矛盾的等式 $0 = b'_j$, 故方程组无解.

若 $r < m$ 且 $b'_{r+1} = \cdots = b'_n = 0$. 记 $x_1, x_{s_2}, \dots, x_{s_r}$ 以外的所有未知元为 $x_{\ell_1}, \dots, x_{\ell_{m-r}}$. 当 $x_{\ell_1}, \dots, x_{\ell_{m-r}}$ 任意取一组值 $x_{\ell_1} = t_1, \dots, x_{\ell_{m-r}} = t_{m-r}$, 并代入前 r 个方程, 则方程组化为

$$\begin{cases} a'_{11}x_1 &= h_1 \\ a'_{2s_2}x_{s_2} &= h_2 \\ & \vdots \\ a'_{rs_r}x_{s_r} &= h_r \end{cases} \tag{2.5}$$

$x_1, x_{s_2}, \dots, x_{s_r}$ 的值就唯一确定, 得到了方程组的一组解. 因为 $x_{\ell_1}, \dots, x_{\ell_{m-r}}$ 可以在 $\mathbb{F}$ 中任意取值, 此时方程组有无穷多组解.

若 $r = m$ 且 $b'_{r+1} = \cdots = b'_n = 0$, 因为 $1 = s_1 < s_2 < \cdots < s_m \le m$, 只能 $s_2 = 2, \dots, s_m = m$, 即阶梯形为

$$\begin{cases} a'_{11}x_1 & & & &= b'_1 \\ & a'_{22}x_2 & & &= b'_2 \\ & & \ddots & & \vdots \\ & & & a'_{mm}x_m &= b'_m \end{cases} \tag{2.6}$$

其中 $a'_{11}, \dots, a'_{mm}$ 均不为 0. 此时方程组有唯一解. □

推论 2.7. 若线性方程组 (2.1) 有解, 且未知元的个数大于方程的个数, 即 $m > n$, 则方程组有无穷多组解.

证明. 因为线性方程组 (2.1) 有解, 化为阶梯形 (2.2) 后还剩下 r 个方程, 于是 $r \le n < m$, 由命题 2.6, 必有无穷多组解. □

综上, 用初等变换将一个线性方程组化为阶梯形求解的方法称为消元法. 当线性方程组 (2.1) 化为阶梯形 (2.2) 后, $x_1, x_{s_2}, \dots, x_{s_r}$ 称为方程组 (2.1) 的主未知元, 其余未知元 $x_{\ell_1}, \dots, x_{\ell_{m-r}}$ 称为方程组 (2.1) 的自由变量. 如果方程组 (2.1) 有解, 称它是相容的 (或可解的), 否则称它是不相容的.

若 (2.1) 是相容的, 由命题 2.6, 自由变量 $x_{\ell_1}, \dots, x_{\ell_{m-r}}$ 可以在 $\mathbb{F}$ 中任意取值, 并且主未知元 $x_1, x_{s_2}, \dots, x_{s_r}$ 随之唯一确定. 当自由变量均取 0 时, 得到一组解

$$x_1 = \frac{b_1'}{a_{11}'},\ x_{s_2} = \frac{b_2'}{a_{2s_2}'},\ \dots,\ x_{s_r} = \frac{b_r'}{a_{rs_r}'},\ x_{\ell_1} = \cdots = x_{\ell_{m-r}} = 0, \tag{2.7}$$

记这组解对应的解向量为 $\mathbf{d}$, 也称为方程组 (2.1) 的一个特解.

与方程组 (2.1) 相关联的是如下齐次线性方程组:

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1m}x_m = 0 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2m}x_m = 0 \\ \qquad\qquad\vdots \\ a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nm}x_m = 0 \end{cases} \tag{2.8}$$

即系数 a_{ij} $(1 \le i \le n, 1 \le j \le m)$ 和 (2.1) 相同, 而常数项 $b_1, \dots, b_n$ 均为 0 的线性方程组, 也称为与 (2.1) 关联的齐次线性方程组. 相应的, (2.1) 称为非齐次线性方程组. 齐次线性方程组是一定有解的, 比如零解.

命题 2.8. 若线性方程组 (2.1) 相容, 解集为 Ω, 且 $\mathbf{d}$ 是 (2.7) 中定义的 (2.1) 的特解. 设 (2.8) 是与 (2.1) 关联的齐次线性方程组, 解集为 Ω_0, 则对 $\forall\ \mathbf{x} \in \Omega_0$, $\mathbf{x} + \mathbf{d} \in \Omega$, 且有一一映射 $f\colon \Omega_0 \longrightarrow \Omega$, 其中 $f(\mathbf{x}) = \mathbf{x} + \mathbf{d}$.

证明. 设 $\mathbf{d} = [d_1, \dots, d_m] \in \Omega$, $\mathbf{x} = [x_1, \dots, x_m] \in \Omega_0$, 则对任意 $1 \le i \le n$,

$$\begin{cases} a_{i1}d_1 + \cdots + a_{im}d_m = b_i \\ a_{i1}x_1 + \cdots + a_{im}x_m = 0 \end{cases} \Longrightarrow a_{i1}(x_1 + d_1) + \cdots + a_{im}(x_m + d_m) = b_i, \tag{2.9}$$

从而 $\mathbf{x} + \mathbf{d} \in \Omega$, 故可以定义映射 $f\colon \Omega_0 \longrightarrow \Omega$, 其中 $f(\mathbf{x}) = \mathbf{x} + \mathbf{d}$. 反过来, 对 $\forall\ \mathbf{y} \in \Omega$, 可以验证 $\mathbf{y} - \mathbf{d} \in \Omega_0$, 故可以定义映射 $g\colon \Omega \longrightarrow \Omega_0$, 其中 $g(\mathbf{y}) = \mathbf{y} - \mathbf{d}$. 显然 g 是 f 的逆映射, 故 f 是一一映射. □

接着我们研究齐次线性方程组 (2.8) 的解集 Ω_0 的结构. 由定理 2.5, 当然也可以利用初等变将 (2.8) 换化为阶梯形, 此时等式右边的常数项全为 0, 形式更为简单:

$$
\begin{cases}
a'_{11}x_1 + \cdots + 0 + \cdots + 0 + \cdots + a'_{1m}x_m = 0 \\
\qquad a'_{2s_2}x_{s_2} + \cdots + 0 + \cdots + a'_{2m}x_m = 0 \\
\qquad\qquad \ddots \quad \vdots \qquad \vdots \quad \vdots \\
\qquad\qquad a'_{rs_r}x_{s_r} + \cdots + a'_{rm}x_m = 0
\end{cases}
\tag{2.10}
$$

设自由变量为 $x_{\ell_1}, \ldots, x_{\ell_{m-r}}$, 其中 $\ell_1 < \ell_2 < \cdots < \ell_{m-r}$, 考虑如下特殊的解:

引理 2.9. 对 $\forall\, 1 \le j \le m-r$, 设 $\mathbf{c}_j$ 是自由变量 $x_{\ell_j} = 1$, 其他自由变量均为 0 得到的 (2.8) 的解向量 $\mathbf{c}_j = [c_{1j}, c_{2j}, \ldots, c_{mj}]$, 则 $c_{\ell_j j} = 1$, 且对 $\forall\, \ell_j < k \le m$, $c_{kj} = 0$, 即

$$\mathbf{c}_j = [c_{1j}, c_{2j}, \ldots, c_{\ell_j - 1 j}, 1, 0, \ldots, 0].$$

证明. 先举一个例子. 假如 $s_2 > 2$, 则 x_2 是一个自由变量, 它只可能出现在阶梯形 (2.10) 中的第一个方程, 若 $x_2 = 1$, 其他自由变量均为 0, 则 (2.10) 化为

$$
\begin{cases}
x_1 = -a'_{12}/a'_{11} \\
x_{s_2} = 0 \\
\quad\vdots \\
x_{s_r} = 0
\end{cases}
\tag{2.11}
$$

得到一组解 $x_1 = -a'_{12}/a'_{11}$, $x_2 = 1$, $x_3 = \cdots = x_m = 0$, 解向量为 $\mathbf{x} = [-a'_{12}/a'_{11}, 1, 0, \ldots, 0]$.

一般的, 设 x_{ℓ_j} 是一个自由变量, 则存在 $1 \le t \le r$, 使得 $s_t < \ell_j < s_{t+1}$, 这里我们令 $s_1 = 1$, $s_{r+1} = m+1$. 则 x_{ℓ_j} 只可能出现在阶梯形 (2.10) 中的前 r 个方程. 因为除 x_{ℓ_j} 以外的其他自由变量均为 0, 对任意 $t < q \le r$, 主未知元 $x_{s_q} = 0$, 于是对 $\forall\, k > \ell_j$, $x_k = 0$, 即 $\mathbf{c}_j = [c_{1j}, c_{2j}, \ldots, c_{\ell_j - 1 j}, 1, 0, \ldots, 0]$. □

引理 2.10. 设 $\mathbf{c}_1, \ldots, \mathbf{c}_{m-r}$ 是引理 2.9 中定义的 (2.8) 的 $m-r$ 个解向量, 则 $\mathbf{c}_1, \ldots, \mathbf{c}_{m-r}$ 线性无关.

证明. 设 $\lambda_1 \mathbf{c}_1 + \cdots + \lambda_{m-r}\mathbf{c}_{m-r} = \mathbf{0}$, 即

$$
\begin{cases}
c_{11}\lambda_1 + \cdots + c_{1m-r}\lambda_{m-r} = 0 \\
\quad\vdots \\
c_{m1}\lambda_1 + \cdots + c_{mm-r}\lambda_{m-r} = 0
\end{cases}
\tag{2.12}
$$

考察第 ℓ_{m-r} 个方程, 因为 $c_{\ell_{m-r}m-r}=1$, 而对 $\forall\ 1\le j<m-r$, $c_{\ell_{m-r}j}=0$, 故第 ℓ_{m-r} 个方程为 $\lambda_{m-r}=0$. 类似的可以得到 $\lambda_1,\dots,\lambda_{m-r-1}$ 均为 0, 即 $\mathbf{c}_1,\dots,\mathbf{c}_{m-r}$ 线性无关. $\square$

一般的, 当自由变量 $x_{\ell_1}=t_1,\dots,x_{\ell_{m-r}}=t_{m-r}$ 时, 代入 (2.10) 得到的解可以表示为

$$\begin{cases} x_1=c'_{11}t_1+\cdots+c'_{1m-r}t_{m-r},\\ x_2=c'_{21}t_1+\cdots+c'_{2m-r}t_{m-r},\\ \quad\vdots\\ x_m=c'_{m1}t_1+\cdots+c'_{mm-r}t_{m-r}. \end{cases} \tag{2.13}$$

对 $\forall\ 1\le j\le m-r$, 记 $\mathbf{c}'_j=[c'_{1j},\dots,c'_{mj}]$, 则解向量 $\mathbf{x}$ 可以表示为

$$\mathbf{x}=\begin{pmatrix}x_1\\ \vdots\\ x_m\end{pmatrix}=t_1\begin{pmatrix}c'_{11}\\ \vdots\\ c'_{m1}\end{pmatrix}+\cdots+t_{m-r}\begin{pmatrix}c'_{1m-r}\\ \vdots\\ c'_{mm-r}\end{pmatrix}=t_1\mathbf{c}'_1+\cdots+t_{m-r}\mathbf{c}'_{m-r}. \tag{2.14}$$

对 $\forall\ 1\le j\le m-r$, 令 $t_j=1$, $t_{k\ne j}=0$, 就得到 $\mathbf{c}'_j=\mathbf{c}_j$. 于是

$$\mathbf{x}=\lambda_1\mathbf{c}_1+\cdots+\lambda_{m-r}\mathbf{c}_{m-r},\quad \lambda_1,\dots,\lambda_{m-r}\in\mathbb{F}. \tag{2.15}$$

定理 2.11. 设 Ω_0 是齐次线性方程组 (2.8) 的解集, 则可以在 Ω_0 中找到一组线性无关的解向量 $\mathbf{c}_1,\dots,\mathbf{c}_{m-r}$, 使得

$$\Omega_0=\{\lambda_1\mathbf{c}_1+\cdots+\lambda_{m-r}\mathbf{c}_{m-r}\ \mid\ \lambda_1,\dots,\lambda_{m-r}\in\mathbb{F}\}. \tag{2.16}$$

$\mathbf{c}_1,\dots,\mathbf{c}_{m-r}$ 称为齐次线性方程组 (2.8) 的一组*基础解系*.

由命题 2.8, 也可以给出非齐次线性方程组 (2.1) 的解集的结构:

定理 2.12. 若非齐次线性方程组 (2.1) 是相容的, Ω 是它的解集, $\mathbf{d}$ 是一个特解, $\mathbf{c}_1,\dots,\mathbf{c}_{m-r}$ 是与 (2.1) 关联的齐次线性方程组的一组基础解系, 则

$$\Omega=\{\lambda_1\mathbf{c}_1+\cdots+\lambda_{m-r}\mathbf{c}_{m-r}+\mathbf{d}\ \mid\ \lambda_1,\dots,\lambda_{m-r}\in\mathbb{F}\}. \tag{2.17}$$

$\mathbf{x}=\lambda_1\mathbf{c}_1+\cdots+\lambda_{m-r}\mathbf{c}_{m-r}+\mathbf{d}$ 称为 (2.1) 的*通解*.

以上我们给出了一般的线性方程组的消元算法, 并说明了线性方程组的通解可以表示成一组解向量的线性组合. 但这里对解集的描述依赖于我们做消元法 (初等变换) 的过程. 我们期待能够直接从原方程组给出线性方程组解集, 并回答如下两个问题:

(1) 在阶梯形中的 r 是否依赖于消元法的过程? 如何直接确定 r?

(2) 齐次线性方程组的解集 Ω_0 的大小和 r 有何关系?

要回答这两个问题, 需要进一步对解集 Ω_0 本身的结构加以研究.

2.2 子空间

考虑关于 3 个变量的齐次线性方程组

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 = 0 \\ a_{21}x_1 + a_{22}x_2 + a_{23}x_3 = 0 \end{cases}$$

它的每个方程表示三维空间 $\mathbb{F}^3$ 中过原点 O 的一个平面, 方程组的解就是两个平面的交点. 若两个平面不重合, 则所有交点构成一条过原点的直线 L. 设 $\mathbf{u}$ 是 L 的方向向量, 则 L 可以表示为

$$\vec{OP} = \lambda\mathbf{u}, \quad \lambda \in \mathbb{F}$$

若两个平面重合, 则所有交点构成一个过原点的平面 π. 任取 π 上两个线性无关的向量 $\mathbf{v}$, $\mathbf{w}$, 则 π 可以表示为

$$\vec{OP} = \lambda\mathbf{v} + \mu\mathbf{w}, \quad \lambda, \mu \in \mathbb{F}$$

直线和平面是三维空间 $\mathbb{F}^3$ 的一维和二维几何对象, 这说明有 m 个未知元的齐次线性方程组的解空间应是 $\mathbb{F}^m$ 中一类作为直线和平面的推广的几何对象.

命题 2.13. 设 Ω_0 是齐次线性方程组 (2.8) 的解空间. 则对 $\forall\ \mathbf{v}, \mathbf{w} \in \Omega_0$ 及 $\forall\ \lambda, \mu \in \mathbb{F}$,

$$\lambda\mathbf{v} + \mu\mathbf{w} \in \Omega_0.$$

证明. 设 $\mathbf{v} = [v_1, \dots, v_m]$, $\mathbf{w} = [w_1, \dots, w_m]$, 则对 $\forall\ 1 \le i \le n$, 代入 (2.8) 的第 i 个方程得到

$$a_{i1}v_1 + \cdots + a_{im}v_m = 0, \quad a_{i1}w_1 + \cdots + a_{im}w_m = 0.$$

从而对 $\forall\ \lambda, \mu \in \mathbb{F}$, 有

$$\begin{aligned} &\lambda(a_{i1}v_1 + \cdots + a_{im}v_m) + \mu(a_{i1}w_1 + \cdots + a_{im}w_m) = 0 \\ \Longrightarrow\ & a_{i1}(\lambda v_1 + \mu w_1) + \cdots + a_{im}(\lambda v_m + \mu w_m) = 0 \Longrightarrow \lambda\mathbf{v} + \mu\mathbf{w} \in \Omega_0 \end{aligned}$$ □

定义 2.14. 设 $V \subset \mathbb{F}^m$ 是由一些 m 维向量构成的非空集合, 如果 V 满足

$$\mathbf{v}, \mathbf{w} \in V,\ \lambda, \mu \in \mathbb{F} \Longrightarrow \lambda\mathbf{v} + \mu\mathbf{w} \in V, \tag{2.18}$$

则称 V 是 $\mathbb{F}^m$ 的子空间.

设 V 是 $\mathbb{F}^m$ 的子空间, 在 (2.18) 中取 $\lambda = \mu = 0$, $\mathbf{v}, \mathbf{w}$ 是 V 中任意向量, 则得到 $\mathbf{0} \in V$. 从而 $\mathbb{F}^m$ 的子空间中必有零向量. 另一方面, 若 $V = \{\mathbf{0}\}$, 即 V 中只有零向量, 则 V 也是 $\mathbb{F}^m$ 的子空间, 称为零空间.

齐次线性方程组 (2.8) 的解集 Ω_0 为 $\mathbb{F}^m$ 的子空间, 也称为解空间. 我们知道 $\mathbb{F}^m$ 中的向量都是用坐标表示的, 等价的说, 都可以唯一的表示成标准基 $\boldsymbol{\epsilon}_1, \ldots, \boldsymbol{\epsilon}_m$ 的线性组合. 但是对 $\mathbb{F}^m$ 的子空间 V, $\boldsymbol{\epsilon}_1, \ldots, \boldsymbol{\epsilon}_m$ 不一定在 V 中. 我们希望在 V 中也能找到一组 "基".

定义 2.15. 设 V 是 $\mathbb{F}^m$ 的子空间, 且 $V \neq \{\mathbf{0}\}$, $\boldsymbol{\alpha} \subset V$ 是一个有限子集. 若 V 中任意向量都可以唯一表示为 $\boldsymbol{\alpha}$ 中向量的线性组合, 则称 $\boldsymbol{\alpha}$ 是 V 的一组基.

引理 2.16. 设 $\mathbf{a}_1, \ldots, \mathbf{a}_k \in \mathbb{F}^m$, 则 $\mathbf{a}_1, \ldots, \mathbf{a}_k$ 线性无关 $\Longleftrightarrow$ $\mathbf{a}_1 \neq \mathbf{0}$, 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.

证明. "$\Longrightarrow$": 若 $\mathbf{a}_1 = \mathbf{0}$, 则 $\mathbf{a}_1 + 0\mathbf{a}_2 + \cdots + 0\mathbf{a}_k = \mathbf{0}$, 即 $\mathbf{a}_1, \ldots, \mathbf{a}_k$ 线性相关, 矛盾. 若存在某个 $2 \leq \ell \leq k$, 使得 $\mathbf{a}_\ell = \lambda_1\mathbf{a}_1 + \cdots + \lambda_{\ell-1}\mathbf{a}_{\ell-1}$, 则 $\lambda_1\mathbf{a}_1 + \cdots + \lambda_{\ell-1}\mathbf{a}_{\ell-1} - \mathbf{a}_\ell = \mathbf{0}$, 也矛盾.
"$\Longleftarrow$": 若存在不全为 0 的 $\lambda_1, \ldots, \lambda_k$, 使得 $\lambda_1\mathbf{a}_1 + \cdots + \lambda_k\mathbf{a}_k = \mathbf{0}$. 取 $\lambda_1, \ldots, \lambda_k$ 中下标最大的非零数, 即取 $1 \leq \ell \leq k$, 使得 $\lambda_\ell \neq 0$, 但是 $\lambda_{\ell+1} = \cdots = \lambda_k = 0$. 则

$$\lambda_1\mathbf{a}_1 + \cdots + \lambda_{\ell-1}\mathbf{a}_{\ell-1} + \lambda_\ell\mathbf{a}_\ell = \mathbf{0} \Longrightarrow \mathbf{a}_\ell = -\frac{\lambda_1}{\lambda_\ell}\mathbf{a}_1 - \cdots - \frac{\lambda_{\ell-1}}{\lambda_\ell}\mathbf{a}_{\ell-1},$$

于是 $\mathbf{a}_\ell$ 是 $\mathbf{a}_1, \ldots, \mathbf{a}_{\ell-1}$ 的线性组合, 矛盾. □

引理 2.17. $\mathbb{F}^m$ 中任何 $m+1$ 个向量一定线性相关.

证明. 设 $\mathbf{a}_1, \ldots, \mathbf{a}_{m+1}$ 是 $\mathbb{F}^m$ 中的向量, 且 $\mathbf{a}_i = [a_{1i}, \ldots, a_{mi}]$ $(1 \leq i \leq m+1)$, 则

$$x_1\mathbf{a}_1 + \cdots + x_{m+1}\mathbf{a}_{m+1} = \mathbf{0} \Longleftrightarrow x_1\begin{pmatrix} a_{11} \\ \vdots \\ a_{m1} \end{pmatrix} + \cdots + x_{m+1}\begin{pmatrix} a_{1m+1} \\ \vdots \\ a_{mm+1} \end{pmatrix} = \mathbf{0}$$

$$\Longleftrightarrow \left\{\begin{array}{rcrl} a_{11}x_1 + & \cdots & + \quad a_{1m+1}x_{m+1} & = 0 \\ & \vdots & & \\ a_{m1}x_1 + & \cdots & + \quad a_{mm+1}x_{m+1} & = 0 \end{array}\right. \tag{2.19}$$

(2.19) 是由 m 个方程组成的关于 $m+1$ 个未知元的齐次线性方程组, 由推论 2.7, 该方程组必有非零解, 于是 $\mathbf{a}_1,\ldots,\mathbf{a}_{m+1}$ 线性相关. □

可以看出引理 2.17 是推论 1.23 的推广. 接着我们来证明本节的主要定理:

定理 2.18. 设 V 是 $\mathbb{F}^m$ 的非零子空间, 则 V 中存在一组基.

证明. 因为 $V \neq \{\mathbf{0}\}$, 可以在 V 中找到一个非零向量 $\mathbf{a}_1$. 接下来递归的在 V 中进行如下操作: 若已得到 V 中向量 $\mathbf{a}_1,\ldots,\mathbf{a}_k$, 使得对任意 $1\leq \ell \leq k$, $\mathbf{a}_\ell$ 不是 $\mathbf{a}_1,\ldots,\mathbf{a}_{\ell-1}$ 的线性组合, 在 V 中继续寻找向量, 它不是 $\mathbf{a}_1,\ldots,\mathbf{a}_k$ 的线性组合, 若存在这样的向量, 任取其中一个 $\mathbf{a}_{k+1}$, 并继续操作, 否则操作停止.

一方面该操作必在 m 步内停止. 否则我们就在 V 中, 也在 $\mathbb{F}^m$ 中, 找到了 $\mathbf{a}_1,\ldots,\mathbf{a}_{m+1}$, 使得 $\mathbf{a}_1\neq\mathbf{0}$, 且对任意 $1\leq\ell\leq m+1$, $\mathbf{a}_\ell$ 不是 $\mathbf{a}_1,\ldots,\mathbf{a}_{\ell-1}$ 的线性组合. 由引理 2.16, 就得到 $\mathbf{a}_1,\ldots,\mathbf{a}_{m+1}$ 线性无关, 与引理 2.17 矛盾.

另一方面, 设该操作在第 r 步停止, 其中 $1\leq r\leq m$, 则我们在 V 中找到了向量 $\mathbf{a}_1,\ldots,\mathbf{a}_r$, 它们线性无关, 且 V 中任意向量都可以用 $\mathbf{a}_1,\ldots,\mathbf{a}_r$ 的线性表示. 若 V 中某个向量 $\mathbf{v}$ 可以表示成 $\mathbf{a}_1,\ldots,\mathbf{a}_r$ 的两个不同的线性组合:

$$\mathbf{v}=\lambda_1\mathbf{a}_1+\cdots+\lambda_r\mathbf{a}_r=\mu_1\mathbf{a}_1+\cdots+\mu_r\mathbf{a}_r,\quad \lambda_1,\ldots,\lambda_r,\mu_1,\ldots,\mu_r\in\mathbb{F}$$

则 $(\lambda_1-\mu_1)\mathbf{a}_1+\cdots+(\lambda_r-\mu_r)\mathbf{a}_r=\mathbf{0}$, 且存在 $1\leq i\leq r$ 使得 $\lambda_i-\mu_i\neq 0$, 与 $\mathbf{a}_1,\ldots,\mathbf{a}_r$ 线性无关矛盾.

综上我们证明了可以在 $\mathbb{F}^m$ 的子空间 V 中找到不超过 m 个向量 $\mathbf{a}_1,\ldots,\mathbf{a}_r$, 使得 V 中任意向量都可以唯一表示为 $\mathbf{a}_1,\ldots,\mathbf{a}_r$ 的线性组合, 即找到了 V 的一组基. □

回顾定理 2.18 的证明过程, 我们所构造的 V 的一组基 $\mathbf{a}_1,\ldots,\mathbf{a}_r$ 实际上满足如下条件:

定义 2.19. 设 V 是 $\mathbb{F}^m$ 的子空间, $\boldsymbol{\alpha}=\{\mathbf{a}_1,\ldots,\mathbf{a}_r\}$ 是 V 中线性无关的向量组. 如果在 $\boldsymbol{\alpha}$ 中添加任意 V 中向量就变为线性相关向量组, 就称 $\boldsymbol{\alpha}$ 是 V 的一个极大线性无关组.

于是找 V 的一组基就等价于找 V 的一组极大线性无关组, 极大线性无关组给出了子空间 V 的一组基的可构造性的等价定义.

推论 2.20. V 中任何一组线性无关的向量组均可扩充为 V 的一组基.

证明. 设 $\mathbf{a}_1,\ldots,\mathbf{a}_s$ 是 V 中给定的线性无关的向量组, 这等价于我们从定理 2.18 中递归操作的第 s 步出发, 继续该操作就可以得到由 $\mathbf{a}_1,\ldots,\mathbf{a}_s$ 扩充而来的 V 的一组基

$$\mathbf{a}_1,\ldots,\mathbf{a}_s,\mathbf{a}_{s+1},\ldots,\mathbf{a}_r.$$ □

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, 则

命题 2.21. 设 $\boldsymbol{\alpha}, \boldsymbol{\beta}$ 是 V 的两个极大线性无关组, 则 $|\boldsymbol{\alpha}| = |\boldsymbol{\beta}|$.

证明. 设 $\boldsymbol{\alpha} = \{\mathbf{a}_1, \ldots, \mathbf{a}_r\}$, $\boldsymbol{\beta} = \{\mathbf{b}_1, \ldots, \mathbf{b}_s\}$ 是 V 的两个极大线性无关组. 则一方面 $\boldsymbol{\alpha}$ 中向量线性无关, 且 $\boldsymbol{\alpha}$ 中向量都可以由 $\boldsymbol{\beta}$ 中向量线性表示. 设

$$\begin{aligned} \mathbf{a}_1 &= a_{11}\mathbf{b}_1 + a_{21}\mathbf{b}_2 + \cdots + a_{s1}\mathbf{b}_s, \\ \mathbf{a}_2 &= a_{12}\mathbf{b}_1 + a_{22}\mathbf{b}_2 + \cdots + a_{s2}\mathbf{b}_s, \\ &\vdots \\ \mathbf{a}_r &= a_{1r}\mathbf{b}_1 + a_{2r}\mathbf{b}_2 + \cdots + a_{sr}\mathbf{b}_s. \end{aligned}$$

则

$$\begin{aligned} x_1\mathbf{a}_1 + \cdots + x_r\mathbf{a}_r &= x_1(a_{11}\mathbf{b}_1 + \cdots + a_{s1}\mathbf{b}_s) + \cdots + x_r(a_{1r}\mathbf{b}_1 + \cdots + a_{sr}\mathbf{b}_s) \\ &= (a_{11}x_1 + \cdots a_{1r}x_r)\mathbf{b}_1 + \cdots + (a_{s1}x_1 + \cdots + a_{sr}x_r)\mathbf{b}_s. \end{aligned}$$

考虑齐次线性方程组

$$\begin{cases} a_{11}x_1 + \cdots + a_{1r}x_r = 0 \\ \quad\vdots \\ a_{s1}x_1 + \cdots + a_{sr}x_r = 0 \end{cases} \tag{2.20}$$

若 (2.20) 有非零解 $\mathbf{x}$, 则得到一组非平凡的线性组合 $x_1\mathbf{a}_1 + \cdots + x_r\mathbf{a}_r = \mathbf{0}$, 与 $\boldsymbol{\alpha}$ 是线性无关向量组矛盾. 所以 (2.20) 只有零解, 由命题 2.7, 必须方程的个数不少于未知元的个数, 即 $s \geq r$. 另一方面, $\boldsymbol{\beta}$ 也是线性无关向量组且 $\boldsymbol{\beta}$ 中向量都可以由 $\boldsymbol{\alpha}$ 中向量线性表示, 同理可知 $r \geq s$. 于是 $r = s$, 即两个极大线性无关组 $\boldsymbol{\alpha}$, $\boldsymbol{\beta}$ 中元素个数相同. □

定义 2.22. 设 V 是 $\mathbb{F}^m$ 的子空间, V 的所有极大线性无关组都含有同样多的元素个数. 这个共同的数称为是 V 的维数, 记为 $\dim V$.

定义 2.23. 设 V 是 $\mathbb{F}^m$ 的子空间, $\boldsymbol{\alpha} = \{\mathbf{v}_1, \ldots, \mathbf{v}_r\}$ 是 V 的极大线性无关组, 或一组基. 则 V 中任意向量 $\mathbf{v}$ 都可以唯一的表示为 $\boldsymbol{\alpha}$ 中向量的线性组合:

$$\mathbf{a} = x_1\mathbf{v}_1 + \cdots + x_r\mathbf{v}_r, \quad x_1, \ldots, x_r \in \mathbb{F}$$

$(x_1, \ldots, x_r)$ 称为 $\mathbf{v}$ 在基 $\mathbf{v}_1, \ldots, \mathbf{v}_r$ 下的坐标表示.

例题 2.24. 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.

解. V 是 $\mathbb{F}^3$ 中过原点的一个平面, 所以是 $\mathbb{F}^3$ 的子空间. 接下来找 V 的一组极大线性无关组. 一方面 $\mathbf{a}_1=(1,-1,0)$, $\mathbf{a}_2=(0,-1,1)$ 都在 V 中, 且 $\mathbf{a}_1,\mathbf{a}_2$ 线性无关. 另一方面对 V 中任意向量 $\mathbf{x}=(x_1,x_2,x_3)$, 因为 $x_1+x_2+x_3=0$, 所以 $x_2=-x_1-x_3$, 于是

$$\mathbf{x}=(x_1,-x_1-x_3,x_3)=x_1(1,-1,0)+x_2(0,-1,1)=x_1\mathbf{a}_1+x_3\mathbf{a}_2.$$

故 V 中向量都可以用 $\mathbf{a}_1,\mathbf{a}_2$ 线性表示. 从而 $\mathbf{a}_1,\mathbf{a}_2$ 是 V 的一组基, 且 $\dim V=2$. □

下面的命题说明了可以用维数刻画子空间的大小:

命题 2.25. (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 $V\subset W$. 若 $\dim V=\dim W$, 则 $V=W$.

证明. (1) 设 $\boldsymbol{\alpha}$ 是 V 的一组基, 因为 $V\subset W$, 则 $\boldsymbol{\alpha}$ 是 W 中的线性无关向量组, 由推论 2.20 可以扩充为 W 的一组基 $\boldsymbol{\beta}$, 故 $|\boldsymbol{\alpha}|\leq|\boldsymbol{\beta}|$, 即 $\dim V\leq\dim W$.
(2) 和 (1) 相同, 设 $\boldsymbol{\alpha}$ 是 V 的一组基, $\boldsymbol{\beta}$ 是由 $\boldsymbol{\alpha}$ 扩充得到的 W 的一组基, 因为 $\dim V=\dim W$, 故 $|\boldsymbol{\alpha}|=|\boldsymbol{\beta}|$, 只能 $\boldsymbol{\alpha}=\boldsymbol{\beta}$, 于是 $V=W$. □

回到齐次线性方程组 (2.8), 我们来回答第 2.1 节最后提出的关于齐次线性方程组的解空间 Ω_0 的问题. 2.1 的内容是, 我们总可以利用一系列初等变换将 (2.8) 化为某个阶梯形 (2.10). 设 r 是阶梯形方程中出现的非平凡的方程个数, 则我们可以通过阶梯形方程找到 Ω_0 中 $m-r$ 个线性无关的解向量 $\mathbf{c}_1,\ldots,\mathbf{c}_{m-r}$, 使得

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

因为 Ω_0 是 $\mathbb{F}^m$ 的子空间, $\mathbf{c}_1,\ldots,\mathbf{c}_{m-r}$ 实质上是 Ω_0 的一组基, 故

$$\dim\Omega_0=m-r,\tag{2.21}$$

即 $r=m-\dim\Omega_0$, 是一个由方程组 (2.8) 本身所决定的数, 和消元算法过程无关. 同时 (2.21) 也给出了 r 和解空间 Ω_0 的大小的关系.

极大线性无关组的概念还可以推广到 $\mathbb{F}^m$ 中任意一组向量, 且 $\mathbb{F}^m$ 中任意一组向量可以生成 $\mathbb{F}^m$ 的一个子空间:

定义 2.26. 设 $\boldsymbol{\alpha}=\{\mathbf{a}_1,\ldots,\mathbf{a}_k\}$ 是 $\mathbb{F}^m$ 中的向量组. 若 $\mathbf{a}_{i_1},\ldots,\mathbf{a}_{i_r}$ 是 $\boldsymbol{\alpha}$ 中线性无关的向量组, 且任取 $\boldsymbol{\alpha}$ 中向量 $\mathbf{a}_{i_{r+1}}$, $\mathbf{a}_{i_1},\ldots,\mathbf{a}_{i_r},\mathbf{a}_{i_{r+1}}$ 成为线性相关向量组, 则称 $\mathbf{a}_{i_1},\ldots,\mathbf{a}_{i_r}$ 是 $\boldsymbol{\alpha}$ 的极大线性无关组.

定义 2.27. 设 $\boldsymbol{\alpha}=\{\mathbf{a}_1,\dots,\mathbf{a}_k\}$ 是 $\mathbb{F}^m$ 中的向量组, 则如下 $\mathbb{F}^m$ 的子空间

$$\operatorname{span}\boldsymbol{\alpha}=\{\lambda_1\mathbf{a}_1+\cdots+\lambda_k\mathbf{a}_k \mid \lambda_1,\dots,\lambda_k\in\mathbb{F}\}$$

称为由 $\boldsymbol{\alpha}$ 生成的子空间. $\operatorname{span}\boldsymbol{\alpha}$ 的维数也称为向量组 $\boldsymbol{\alpha}$ 的秩, 记为 $\operatorname{rank}\boldsymbol{\alpha}$.

命题 2.28. 设 $\boldsymbol{\alpha}=\{\mathbf{a}_1,\dots,\mathbf{a}_k\}$ 是 $\mathbb{F}^m$ 中向量组, 则 $\boldsymbol{\alpha}$ 的任一极大线性无关组也是 $\operatorname{span}\boldsymbol{\alpha}$ 的一组极大线性无关组. 从而 $\boldsymbol{\alpha}$ 的任一极大线性无关组中向量的个数等于 $\operatorname{rank}\boldsymbol{\alpha}$.

证明. 设 $\boldsymbol{\beta}=\{\mathbf{a}_{i_1},\dots,\mathbf{a}_{i_r}\}$ 是 $\boldsymbol{\alpha}$ 的极大线性无关组, 则 $\boldsymbol{\beta}$ 是 $\operatorname{span}\boldsymbol{\alpha}$ 中一组线性无关的向量. 对 $\forall\,\mathbf{v}\in\operatorname{span}\boldsymbol{\alpha}$, $\mathbf{v}$ 是 $\mathbf{a}_1,\dots,\mathbf{a}_k$ 的线性组合 $\mathbf{v}=\lambda_1\mathbf{a}_1+\cdots+\lambda_k\mathbf{a}_k$. 又因为 $\boldsymbol{\beta}$ 是 $\boldsymbol{\alpha}$ 的极大线性无关组, $\boldsymbol{\alpha}$ 中向量都可以由 $\boldsymbol{\beta}$ 中向量线性表示, 所以 $\mathbf{v}$ 也可以由 $\boldsymbol{\beta}$ 中向量线性表示, 将 $\mathbf{v}$ 添加进 $\boldsymbol{\beta}$ 得到的向量组 $\mathbf{a}_{i_1},\dots,\mathbf{a}_{i_r},\mathbf{v}$ 必然线性相关, 所以 $\boldsymbol{\beta}$ 也是 $\operatorname{span}\boldsymbol{\alpha}$ 的极大线性无关组, 有 $|\boldsymbol{\beta}|=\operatorname{rank}\boldsymbol{\alpha}$. □

例题 2.29. 设 $\boldsymbol{\alpha}=\{\mathbf{a}_1,\dots,\mathbf{a}_k\}$, $\boldsymbol{\beta}=\{\mathbf{b}_1,\dots,\mathbf{b}_\ell\}$, 是 $\mathbb{F}^m$ 的向量组, 且 $\boldsymbol{\alpha}$ 中向量都可以由 $\boldsymbol{\beta}$ 中向量线性表示, 证明 $\operatorname{rank}\boldsymbol{\alpha}\le\operatorname{rank}\boldsymbol{\beta}$.

证明. 设 $\boldsymbol{\alpha}'=\{\mathbf{a}_{i_1},\dots,\mathbf{a}_{i_r}\}$ 是 $\boldsymbol{\alpha}$ 的极大线性无关组, $\boldsymbol{\beta}'=\{\mathbf{b}_{j_1},\dots,\mathbf{b}_{j_s}\}$ 是 $\boldsymbol{\beta}$ 的极大线性无关组. 则 $\boldsymbol{\alpha}'$ 中向量都可以用 $\boldsymbol{\beta}$ 中向量线性表示, 进一步的也可以用 $\boldsymbol{\beta}'$ 中向量线性表示. 和命题 2.21 的证明一样, 可以推出 $r\le s$, 即 $\operatorname{rank}\boldsymbol{\alpha}\le\operatorname{rank}\boldsymbol{\beta}$. □

例题 2.30. 求如下向量组的秩及一组极大线性无关组:

$$\mathbf{a}_1=(3,4,-2,5),\ \mathbf{a}_2=(2,-5,0,-3),\ \mathbf{a}_3=(5,0,-1,2),\ \mathbf{a}_4=(3,3,-3,5).$$

解. $\mathbf{a}_1,\mathbf{a}_2$ 显然线性无关. 考察 $\mathbf{a}_3$ 是否是 $\mathbf{a}_1,\mathbf{a}_2$ 的线性组合:

$$\mathbf{a}_3=x_1\mathbf{a}_1+x_2\mathbf{a}_2\Longrightarrow\begin{cases}3x_1+2x_2=5\\4x_1-5x_2=0\\-2x_1=-1\\5x_1-3x_2=2\end{cases}$$

该线性方程组无解, 故 $\mathbf{a}_1,\mathbf{a}_2,\mathbf{a}_3$ 线性无关. 考察 $\mathbf{a}_4$ 是否是 $\mathbf{a}_1,\mathbf{a}_2,\mathbf{a}_3$ 的线性组合:

$$\mathbf{a}_4=x_1\mathbf{a}_1+x_2\mathbf{a}_2+x_3\mathbf{a}_3\Longrightarrow\begin{cases}3x_1+2x_2+5x_3=3\\4x_1-5x_2=3\\-2x_1-x_3=-3\\5x_1-3x_2+2x_3=5\end{cases}\Longrightarrow\begin{cases}x_1=2\\x_2=1\\x_3=-1\end{cases}$$

故 $\mathbf{a}_4$ 是 $\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3$ 的线性组合. 从而 $\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3$ 是极大线性无关组, 原向量组的秩为 3. □

例题 2.31. 设 W_1, W_2 是 $\mathbb{F}^m$ 的子空间.

(1) 举例说明 $W_1 \cup W_2$ 不一定是 $\mathbb{F}^m$ 的子空间.

(2) 证明 $W_1 + W_2 \coloneqq \{\mathbf{w}_1 + \mathbf{w}_2 \mid \mathbf{w}_1 \in W_1, \mathbf{w}_2 \in W_2\}$ 是 $\mathbb{F}^m$ 中包含 W_1 及 W_2 的最小的子空间, 称为 W_1 与 W_2 的和.

(3) 证明 $\dim\,(W_1 + W_2) + \dim\,(W_1 \cap W_2) = \dim\, W_1 + \dim\, W_2$, 称为维数公式.

证明. (1) 设 $V = \mathbb{F}^2$, $W_1 = \{\lambda\boldsymbol{\epsilon}_1,\ \lambda \in \mathbb{F}\}$, $W_2 = \{\lambda\boldsymbol{\epsilon}_2,\ \lambda \in \mathbb{F}\}$, 则 $\boldsymbol{\epsilon}_1, \boldsymbol{\epsilon}_2 \in W_1 \cup W_2$, 但是 $\boldsymbol{\epsilon}_1 + \boldsymbol{\epsilon}_2 \notin W_1 \cup W_2$. 故 $W_1 \cup W_2$ 不是 V 的子空间.
(2) 首先, 对任意 $\mathbf{w}_1 + \mathbf{w}_2, \mathbf{w}_1' + \mathbf{w}_2' \in W_1 + W_2$,

$$\lambda(\mathbf{w}_1 + \mathbf{w}_2) + \mu(\mathbf{w}_1' + \mathbf{w}_2') = (\lambda\mathbf{w}_1 + \mu\mathbf{w}_1') + (\lambda\mathbf{w}_2 + \mu\mathbf{w}_2') \in W_1 + W_2,$$

从而 $W_1 + W_2$ 是 $\mathbb{F}^m$ 的子空间. 另一方面, 设 W' 是子空间, 且 $W_1, W_2 \subset W'$, 则对 $\forall\, w_1 \in W_1, w_2 \in W_2$, 由 W' 是子空间推出 $w_1 + w_2 \in W'$, 故 $W_1 + W_2 \subset W'$. 即 $W_1 + W_2$ 是包含 W_1 及 W_2 的最小的子空间.
(3) 设 $\{\mathbf{x}_1, \ldots, \mathbf{x}_r\}$ 是 $W_1 \cap W_2$ 的一组基, 并将其分别扩充为 W_1, W_2 的一组基

$$\boldsymbol{\alpha} = \{\mathbf{x}_1, \ldots, \mathbf{x}_r, \mathbf{y}_1, \ldots, \mathbf{y}_s\}, \quad \boldsymbol{\beta} = \{\mathbf{w}_1, \ldots, \mathbf{w}_r, \mathbf{z}_1, \ldots, \mathbf{z}_t\}.$$

我们证明向量组

$$\boldsymbol{\eta} = \{\mathbf{x}_1, \ldots, \mathbf{x}_r, \mathbf{y}_1, \ldots, \mathbf{y}_s, \mathbf{z}_1, \ldots, \mathbf{z}_t\}$$

是 $W_1 + W_2$ 的一组基, 从而得到维数公式.

一方面, 对 $\forall\, \mathbf{w}_1 \in W_1, \mathbf{w}_2 \in W_2$, $\mathbf{w}_1$ 是 $\boldsymbol{\alpha}$ 中向量的线性组合, $\mathbf{w}_2$ 是 $\boldsymbol{\beta}$ 中向量的线性组合, 从而 $\mathbf{w}_1 + \mathbf{w}_2$ 是 $\boldsymbol{\eta}$ 中向量的线性组合. 另一方面, 若

$$\lambda_1\mathbf{x}_1 + \cdots + \lambda_r\mathbf{x}_r + \mu_1\mathbf{y}_1 + \cdots + \mu_s\mathbf{y}_s + \gamma_1\mathbf{z}_1 + \cdots + \gamma_t\mathbf{z}_t = \mathbf{0},$$

记 $\mathbf{w} = \lambda_1\mathbf{x}_1 + \cdots + \lambda_r\mathbf{x}_r$, $\mathbf{v} = \mu_1\mathbf{y}_1 + \cdots + \mu_s\mathbf{y}_s$, $\mathbf{u} = \gamma_1\mathbf{z}_1 + \cdots + \gamma_t\mathbf{z}_t$, 则 $\mathbf{u} = -\mathbf{w} - \mathbf{v} \in W_1$, 推出 $\mathbf{u} \in W_1 \cap W_2$. 从而存在 $\xi_1, \ldots, \xi_r \in \mathbb{F}$ 使得

$$\mathbf{u} = \xi_1\mathbf{x}_1 + \cdots + \xi_r\mathbf{x}_r \iff \xi_1\mathbf{x}_1 + \cdots + \xi_r\mathbf{x}_r - \gamma_1\mathbf{z}_1 - \cdots - \gamma_t\mathbf{z}_t = \mathbf{0},$$

由 $\boldsymbol{\beta}$ 是 W_2 一组基知 $\xi_1 = \cdots = \xi_r = \gamma_1 = \cdots = \gamma_t = 0$, 即 $\mathbf{u} = \mathbf{0}$. 从而 $\mathbf{w} + \mathbf{v} = \mathbf{0}$, 再由 $\boldsymbol{\alpha}$ 是 W_1 一组基可知 $\lambda_1 = \cdots = \lambda_r = \mu_1 = \cdots = \mu_s = 0$, 即向量组 $\boldsymbol{\eta}$ 是线性无关的. □

例题 2.32. 若 W 是 $\mathbb{F}^m$ 的子空间且 $W \neq V$, 称 W 为 $\mathbb{F}^m$ 的真子空间. 证明 $\mathbb{F}^m$ 不能被有限个真子空间覆盖.

证明. 显然 $\mathbb{F}^m$ 不能被它的一个真子空间覆盖, 接下来用归纳法, 假设 $\mathbb{F}^m$ 不能被任何 k 个真子空间覆盖. 给定任何 $k+1$ 个真子空间 $V_1, \ldots, V_{k+1}$, 可以设 $V_{k+1} \not\subseteq V_1 \cup \cdots \cup V_k$, 否则由归纳假设即得. 于是存在 $\mathbf{v} \in V_{k+1}$ 但 $\mathbf{v} \notin V_1 \cup \cdots \cup V_k$. 因为 V_{k+1} 是真子空间, 存在非零向量 $\mathbf{u}$, 使得 $\mathbf{u} \notin V_{k+1}$, 于是对 $\forall\ \lambda \neq 0$, $\mathbf{v} + \lambda\mathbf{u} \notin V_{k+1}$. 考虑集合

$$T = \{\mathbf{v} + \lambda\mathbf{u} \mid \lambda \neq 0\}.$$

若 T 中有一个向量不在 $V_1 \cup \cdots \cup V_k$ 中, 则得到 $V_1, \ldots, V_{k+1}$ 不覆盖 $\mathbb{F}^m$. 否则存在 $\lambda_1 \neq \lambda_2$ 且 $\mathbf{v} + \lambda_1\mathbf{u}$, $\mathbf{v} + \lambda_2\mathbf{u}$ 都在某个 V_i 中. 于是

$$\frac{1}{\lambda_1}(\mathbf{v} + \lambda_1\mathbf{u}) - \frac{1}{\lambda_2}(\mathbf{v} + \lambda_2\mathbf{u}) = (\frac{1}{\lambda_1} - \frac{1}{\lambda_2})\mathbf{v}$$

在 V_i 中, 与 $\mathbf{v} \notin V_1 \cup \cdots \cup V_k$ 矛盾. □

2.3 矩阵的秩

对于齐次线性方程组 (2.8), 可以看出消元法的过程仅和方程组的系数 a_{ij} 相关, 而与未知元无关, 为了方便, 我们将 (2.8) 的系数提取出来写成如下矩阵的形式:

$$\mathbf{A} = (a_{ij})_{n\times m} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1m} \\ a_{21} & a_{22} & \cdots & a_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nm} \end{pmatrix}. \tag{2.22}$$

$\mathbf{A}$ 称为方程组 (2.8) 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 (2.8) 的所有信息.

矩阵 $\mathbf{A}$ 的每一行构成 $\mathbb{F}^m$ 中的一个行向量 $\mathbf{a}_i = (a_{i1}, a_{i2}, \ldots, a_{im})$. $\mathbf{A}$ 可以写为一个列向量, 称为 $\mathbf{A}$ 的列向量表示:

$$\mathbf{A} = \begin{bmatrix} \mathbf{a}_1 & \cdots & \mathbf{a}_n \end{bmatrix} = \begin{pmatrix} \mathbf{a}_1 \\ \vdots \\ \mathbf{a}_n \end{pmatrix}, \tag{2.23}$$

这个列向量中的元素是 $\mathbf{A}$ 的各个行向量.

利用三类初等变换对方程组 (2.8) 做消元法的过程等价于对矩阵 $\mathbf{A}$ 的所有行向量做类似的三类初等行变换:

(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.

定义 2.33. 矩阵 $\mathbf{A}$ 的行向量组 $\boldsymbol{\alpha}=\{\mathbf{a}_1,\ldots,\mathbf{a}_n\}$ 的秩称为矩阵的行秩, 记为 $\operatorname{rrank}\mathbf{A}$.

命题 2.34. 初等行变换不改变矩阵的行秩.

证明. 设 $\mathbf{A}$ 是 $n\times m$ 阶矩阵, 行向量组为 $\boldsymbol{\alpha}=\{\mathbf{a}_1,\ldots,\mathbf{a}_n\}$. 则
(1) 设 $\mathbf{A}'$ 是交换 $\mathbf{A}$ 的第 i 和第 j 行得到的矩阵, 则 $\mathbf{A}'$ 的行向量组和 $\mathbf{A}$ 的行向量组相同, 故行秩相同.
(2) 设 $\mathbf{A}'$ 是将 $\mathbf{A}$ 的第 i 行乘以 $\lambda\neq 0$ 得到的矩阵, 则 $\mathbf{A}'$ 的行向量组为

$$\boldsymbol{\alpha}'=\{\mathbf{a}_1,\ldots,\lambda\mathbf{a}_i,\ldots,\mathbf{a}_n\},$$

显然 $\operatorname{span}\boldsymbol{\alpha}'=\operatorname{span}\boldsymbol{\alpha}$, 故 $\mathbf{A}'$ 和 $\mathbf{A}$ 的行秩相同.
(3) 设 $\mathbf{A}'$ 是将 $\mathbf{A}$ 的第 j 行乘以 λ 加到第 i 行得到的矩阵, 则 $\mathbf{A}'$ 的行向量组为

$$\boldsymbol{\alpha}'=\{\mathbf{a}_1,\ldots,\mathbf{a}_i+\lambda\mathbf{a}_j,\ldots,\mathbf{a}_n\}.$$

显然 $\boldsymbol{\alpha}'$ 中向量都可以用 $\boldsymbol{\alpha}$ 中向量线性表示. 另一方面, $\mathbf{a}_i=(\mathbf{a}_i+\lambda\mathbf{a}_j)-\lambda\mathbf{a}_j$, 故 $\boldsymbol{\alpha}$ 中向量也可以用 $\boldsymbol{\alpha}'$ 中向量线性表示, 由例题 2.29, 知 $\operatorname{rank}\boldsymbol{\alpha}=\operatorname{rank}\boldsymbol{\alpha}'$. □

由定理 2.5, 每个线性方程组都可以利用初等变换化为阶梯形, 等价的, 每个矩阵 $\mathbf{A}$ 也可以通过初等行变换化为如下阶梯形

$$\mathbf{A}'=\begin{pmatrix} a'_{11} & \cdots & 0 & \cdots & 0 & \cdots & a'_{1m} \\ & & a'_{2s_2} & \cdots & 0 & \cdots & a'_{2m} \\ & & & \ddots & \vdots & & \vdots \\ & & & & a'_{rs_r} & \cdots & a'_{rm} \\ & & & & & & \end{pmatrix} \tag{2.24}$$

因为 $a'_{11},a'_{2s_2},\ldots,a'_{rs_r}$ 均不为 0, 且 $1<s_2<\cdots<s_r$, 容易看出 $\operatorname{rrank}\mathbf{A}'=r$. 于是由命题 2.34 知 $\operatorname{rrank}\mathbf{A}=\operatorname{rrank}\mathbf{A}'=r$. 由 (2.21) 可知

推论 2.35. 设 $\mathbf{A}$ 是 $n\times m$ 阶矩阵, Ω_0 是系数矩阵为 $\mathbf{A}$ 的齐次线性方程组的解空间, 则

$$\operatorname{rrank}\ \mathbf{A}+\dim\ \Omega_0=m. \tag{2.25}$$

于是 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. (2.25) 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 $n\times m$ 阶矩阵 $\mathbf{A}$ 的所有列向量为

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

则 $\mathbf{A}$ 也可以写成一个行向量, 称为 $\mathbf{A}$ 的行向量表示:

$$\mathbf{A}=\begin{pmatrix}\mathbf{b}_1 & \cdots & \mathbf{b}_m\end{pmatrix}, \tag{2.26}$$

这个行向量中的元素是 $\mathbf{A}$ 的列向量.

定义 2.36. 矩阵 $\mathbf{A}$ 的列向量组 $\boldsymbol{\beta}=\{\mathbf{b}_1,\ldots,\mathbf{b}_m\}$ 的秩称为 $\mathbf{A}$ 的列秩, 记为 crank $\mathbf{A}$.

命题 2.37. 设 $\mathbf{A}$ 是 $n\times m$ 阶矩阵, Ω_0 是系数矩阵为 $\mathbf{A}$ 的齐次线性方程组的解空间, 则

$$\operatorname{crank}\ \mathbf{A}+\dim\ \Omega_0=m. \tag{2.27}$$

证明. 设 $\mathbf{A}=\begin{pmatrix}\mathbf{b}_1 & \cdots & \mathbf{b}_m\end{pmatrix}$, $\mathbf{x}=[x_1,\ldots,x_m]\in\mathbb{F}^m$, 则

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

设 crank $\mathbf{A}=s$, 且不妨设 $\{\mathbf{b}_1,\ldots,\mathbf{b}_s\}$ 是列向量组 $\{\mathbf{b}_1,\ldots,\mathbf{b}_m\}$ 的极大线性无关向量组. 则 $\{\mathbf{b}_{s+1},\ldots,\mathbf{b}_m\}$ 可以由 $\{\mathbf{b}_1,\ldots,\mathbf{b}_s\}$ 线性表出:

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

于是

$$\begin{aligned} x_1\mathbf{b}_1+\cdots+x_m\mathbf{b}_m=\mathbf{0} &\Longleftrightarrow x_1\mathbf{b}_1+\cdots+x_s\mathbf{b}_s+x_{s+1}(a_{1s+1}\mathbf{b}_1+\cdots+a_{ss+1}\mathbf{b}_s)\\ &\qquad +\cdots+x_m(a_{1m}\mathbf{b}_1+\cdots+a_{sm}\mathbf{b}_s)=\mathbf{0}\\ &\Longleftrightarrow (x_1+a_{1s+1}x_{s+1}+\cdots+a_{1m}x_m)\mathbf{b}_1+\cdots+\\ &\qquad (x_s+a_{ss+1}x_{s+1}+\cdots+a_{sm}x_m)\mathbf{b}_s=\mathbf{0}\\ (\mathbf{b}_1,\ldots,\mathbf{b}_s \text{ 线性无关}) &\Longleftrightarrow \begin{cases} x_1+a_{1s+1}x_{s+1}+\cdots+a_{1m}x_m=0\\ x_2+a_{2s+1}x_{s+1}+\cdots+a_{2m}x_m=0\\ \qquad\vdots\\ x_s+a_{ss+1}x_{s+1}+\cdots+a_{sm}x_m=0 \end{cases} \end{aligned} \tag{2.28}$$

于是通过一系列的等价条件, 我们得到 $\mathbf{x}\in\Omega_0 \Longleftrightarrow \mathbf{x}$ 是 (2.28) 的解, 即 (2.28) 的解空间也为 Ω_0. 但是 (2.28) 本身就是一个阶梯形方程组, 其系数矩阵的行秩为 s, 由推论 2.35 知

$$s+\dim\,\Omega_0=m. \qquad \square$$

结合推论 2.35 和命题 2.37, 我们得到本节的主要定理:

定理 2.38. 任何矩阵 $\mathbf{A}$ 的行秩等于列秩, 统称为矩阵的秩, 记为 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

$$\mathrm{rank}\ \mathbf{A}+\dim\,\Omega_0=m. \tag{2.29}$$

定义 2.39. 设 $\mathbf{A}=(a_{ij})_{n\times m}$ 是 $n\times m$ 阶矩阵, 称如下 $m\times n$ 阶矩阵

$$\begin{pmatrix} a_{11} & a_{21} & \cdots & a_{n1}\\ a_{12} & a_{22} & \cdots & a_{n2}\\ \vdots & \vdots & \ddots & \vdots\\ a_{1m} & a_{2m} & \cdots & a_{nm} \end{pmatrix} \tag{2.30}$$

为 $\mathbf{A}$ 的转置, 记为 $\mathbf{A}^{\mathrm{T}}$.

一个行向量的转置就是含有同样元素的列向量. 若 $\mathbf{A}=\begin{bmatrix}\mathbf{a}_1 & \cdots & \mathbf{a}_n\end{bmatrix}$, 则 $\mathbf{A}^{\mathrm{T}}=\begin{pmatrix}\mathbf{a}_1^{\mathrm{T}} & \cdots & \mathbf{a}_n^{\mathrm{T}}\end{pmatrix}$. 由定理 2.38 得到

命题 2.40. 对任意矩阵 $\mathbf{A}$, rank $\mathbf{A}=$ rank $\mathbf{A}^{\mathrm{T}}$.

证明 rrank $\mathbf{A}$ = crank $\mathbf{A}$ 的另一种方法是证明初等行变换也不改变矩阵的列秩. 等价的, 我们证明如下更容易操作的命题:

命题 2.41. 设 $\mathbf{A}=(a_{ij})_{n\times m}$ 的列向量组为 $\mathbf{b}_1,\dots,\mathbf{b}_m$. $\mathbf{A}'$ 是对 $\mathbf{A}$ 做某个初等行变换后得到的矩阵, 列向量组为 $\mathbf{b}'_1,\dots,\mathbf{b}'_m$. 则对任意 $1\le j_1<j_2<\cdots<j_s\le m$, $\mathbf{b}_{j_1},\dots,\mathbf{b}_{j_s}$ 线性无关 $\iff$ $\mathbf{b}'_{j_1},\dots,\mathbf{b}'_{j_s}$ 线性无关.

证明. $\mathbf{b}_{j_1},\dots,\mathbf{b}_{j_s}$ 线性无关 $\iff$ 不存在 $\mathbf{x}=[x_1,\dots,x_s]\neq\mathbf{0}$ 使得 $x_1\mathbf{b}_{j_1}+\cdots+x_s\mathbf{b}_{j_s}=\mathbf{0}$, 这等价于如下齐次线性方程组只有零解:

$$\begin{cases} a_{1j_1}x_1+\cdots+a_{1j_s}x_s=0\\ \qquad\vdots\\ a_{nj_1}x_1+\cdots+a_{nj_s}x_s=0\end{cases}\tag{2.31}$$

同样的, $\mathbf{b}'_{j_1},\dots,\mathbf{b}'_{j_s}$ 线性无关也等价于某个齐次线性方程组只有零解, 而这个齐次线性方程组就是对 (2.31) 做和从 $\mathbf{A}$ 的 $\mathbf{A}'$ 所做的初等行变的同类初等变换所得到. □

将矩阵 $\mathbf{A}$ 通过行变换化为阶梯形 (2.24), 可以看出 $\mathbf{A}'$ 的第 $1,s_2,\dots,s_r$ 列即为 $\mathbf{A}'$ 的列向量组的极大线性无关组. 特别的, 阶梯形矩阵的行秩等于列秩. 再由初等行变换不改变行秩与列秩就得到定理 2.38 的另一个证明.

同时, 命题 (2.41) 还给出了求一组向量 $\boldsymbol{\alpha}=\{\mathbf{a}_1,\dots,\mathbf{a}_k\}$ 的秩及其极大线性无关组的矩阵算法:

(1) 将 $\mathbf{a}_1,\dots,\mathbf{a}_k$ 作为列向量排列成矩阵 $\mathbf{A}$.

(2) 通过行变换将 $\mathbf{A}$ 化为阶梯形 $\mathbf{A}'$. 设 $\mathbf{A}'$ 中非零行为第 $1,2,\dots,r$ 行, 且第 ℓ 行中的第一个非零元素在第 s_ℓ 列.

(3) 则 rank $\boldsymbol{\alpha}=r$, $\{\mathbf{a}_{s_1},\dots,\mathbf{a}_{s_r}\}$ 是 $\boldsymbol{\alpha}$ 的一个极大线性无关组.

(4) 设 $\{\mathbf{a}'_{j_1},\dots,\mathbf{a}'_{j_r}\}$ 是梯型矩阵 $\mathbf{A}'$ 中任意 r 个线性无关的列向量组, 则 $\{\mathbf{a}_{j_1},\dots,\mathbf{a}_{j_r}\}$ 也是 $\boldsymbol{\alpha}$ 的极大线性无关组.

例题 2.42. 利用矩阵算法求例题 2.30 中的向量组 $\boldsymbol{\alpha}=\{\mathbf{a}_1,\mathbf{a}_2,\mathbf{a}_3,\mathbf{a}_4\}$ 的秩及所有极大线性无关向量组.

解. 将 $\mathbf{a}_1,\mathbf{a}_2,\mathbf{a}_3,\mathbf{a}_4$ 按照列向量排成矩阵并通过行变换化为简单的形式:

$$\mathbf{A}=\begin{pmatrix}3&2&5&3\\4&-5&0&3\\-2&0&-1&-3\\5&-3&2&5\end{pmatrix}\longrightarrow\begin{pmatrix}-1&7&5&0\\0&23&20&3\\0&-14&-11&-3\\0&32&27&5\end{pmatrix}\longrightarrow$$

$$\begin{pmatrix} -1 & 7 & 5 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 3 & -3 \\ 0 & 0 & -5 & 5 \end{pmatrix} \longrightarrow \mathbf{A}' = \begin{pmatrix} -1 & 7 & 5 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

可以看出 rank $\mathbf{A}' = 3$, 且任意三列都线性无关. 故 rank $\boldsymbol{\alpha} = r$, $\boldsymbol{\alpha}$ 中任意 3 个向量都构成极大线性无关组. □

评注 2.43. 在例题 2.42 中我们并没有将 $\mathbf{A}$ 最终化为阶梯形, 而是停留在一个足够简单的矩阵 $\mathbf{A}'$.

最后我们给出求解非齐次线性方程组 (2.1) 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:

$$\mathbf{B} = (\mathbf{A}|\mathbf{b}) = \left(\begin{array}{cccc|c} a_{11} & a_{12} & \cdots & a_{1m} & b_1 \\ a_{21} & a_{22} & \cdots & a_{2m} & b_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nm} & b_n \end{array}\right). \tag{2.32}$$

用消元法对线性方程组 (2.1) 求解的过程可以简化为在增广矩阵上做行变换化为阶梯形矩阵.

命题 2.44. 设非齐次线性方程组的增广矩阵为 $\mathbf{B} = (\mathbf{A}|\mathbf{b})$, 则

(1) 若 rank $\mathbf{A} <$ rank $\mathbf{B}$, 则方程组不相容.

(2) 若 rank $\mathbf{A} =$ rank $\mathbf{B} < m$, 则方程组相容且有无穷多组解.

(3) 若 rank $\mathbf{A} =$ rank $\mathbf{B} = m$, 则方程组相容且有唯一解.

证明. 设 $\mathbf{A}$ 的列向量组 $\{\mathbf{a}_1, \ldots, \mathbf{a}_m\}$, 则非齐次线性方程组可以写为

$$x_1\mathbf{a}_1 + \cdots + x_m\mathbf{a}_m = \mathbf{b}.$$

若 rank $\mathbf{A} <$ rank $\mathbf{B}$, 即将 $\mathbf{b}$ 添加到 $\mathbf{A}$ 的列向量组后秩加 1, 故 $\mathbf{b}$ 不是 $\mathbf{a}_1, \ldots, \mathbf{a}_m$ 的线性组合, 方程组无解.

若 rank $\mathbf{A} =$ rank $\mathbf{B}$, 则 $\mathbf{b}$ 是 $\mathbf{a}_1, \ldots, \mathbf{a}_m$ 的线性组合, 方程组有解. 进一步的, 若 rank $\mathbf{A} = m$, 则 $\mathbf{a}_1, \ldots, \mathbf{a}_m$ 线性无关, 故 $\mathbf{b}$ 用 $\mathbf{a}_1, \ldots, \mathbf{a}_m$ 线性表示的方式唯一, 即有唯一解. 否则有无穷多组解. □

例题 2.45. 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:

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

解. 对增广矩阵做行变换:

$$\left(\begin{array}{ccc|c}\lambda & 1 & 1 & 1\\ 1 & \lambda & 1 & \lambda\\ 1 & 1 & \lambda & \lambda^2\end{array}\right) \longrightarrow \left(\begin{array}{ccc|c}1 & 1 & \lambda & \lambda^2\\ 0 & \lambda-1 & 1-\lambda & \lambda-\lambda^2\\ 0 & 1-\lambda & 1-\lambda^2 & 1-\lambda^3\end{array}\right),$$

(1) 若 $\lambda = 1$, 则方程组化为 $x_1 + x_2 + x_3 = 1$, 是相容的且有无穷多组解, 通解为

$$\mathbf{x} = \begin{pmatrix}x_1\\ x_2\\ x_3\end{pmatrix} = \begin{pmatrix}-x_2 - x_3 + 1\\ x_2\\ x_3\end{pmatrix} = x_2\begin{pmatrix}-1\\ 1\\ 0\end{pmatrix} + x_3\begin{pmatrix}-1\\ 0\\ 1\end{pmatrix} + \begin{pmatrix}1\\ 0\\ 0\end{pmatrix}, \quad x_2, x_3 \in \mathbb{F}.$$

(2) 若 $\lambda \neq 1$, 则增广矩阵可进一步化简为

$$\left(\begin{array}{ccc|c}1 & 1 & \lambda & \lambda^2\\ 0 & -1 & 1 & \lambda\\ 0 & 1 & 1+\lambda & 1+\lambda+\lambda^2\end{array}\right) \longrightarrow \left(\begin{array}{ccc|c}1 & 1 & \lambda & \lambda^2\\ 0 & -1 & 1 & \lambda\\ 0 & 0 & 2+\lambda & (\lambda+1)^2\end{array}\right)$$

由命题 2.6, 若 $\lambda = -2$, 则方程组无解; 若 $\lambda \neq -2, 1$, 则方程组有唯一解

$$x_1 = -\frac{\lambda+1}{\lambda+2}, \quad x_2 = \frac{1}{\lambda+2}, \quad x_3 = \frac{(\lambda+1)^2}{\lambda+2}.$$ □

例题 2.46. 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.

解. 设 $\mathbf{A} = \begin{pmatrix}\mathbf{a}_1 & \mathbf{a}_2 & \mathbf{a}_3 & \mathbf{a}_4\end{pmatrix}$, $\mathbf{B} = (\mathbf{A}|\mathbf{b})$. 则问题转化判断增广矩阵为 $\mathbf{B}$ 的非齐次线性方

程组何时无解, 以及何时有唯一解. 通过行变换将 $\mathbf{B}$ 化简:

$$\begin{pmatrix}1&1&1&1&1\\0&1&-1&2&1\\2&3&\lambda+2&4&\mu+3\\3&5&1&\lambda+8&5\end{pmatrix}\longrightarrow\begin{pmatrix}1&1&1&1&1\\0&1&-1&2&1\\0&1&\lambda&2&\mu+1\\0&2&-2&\lambda+5&2\end{pmatrix}$$
$$\longrightarrow\begin{pmatrix}1&1&1&1&1\\0&1&-1&2&1\\0&0&\lambda+1&0&\mu\\0&0&0&\lambda+1&0\end{pmatrix}.$$

(1) 若 $\lambda=-1$, $\mu\neq 0$, 则 $\Longleftrightarrow$ $\operatorname{rank}\mathbf{A}<\operatorname{rank}\mathbf{B}$, 方程组无解. 即 $\mathbf{b}$ 不能由 $\mathbf{a}_1$, $\mathbf{a}_2$, $\mathbf{a}_3$, $\mathbf{a}_4$ 线性表出.

(2) 若 $\lambda\neq -1$, 则 $\operatorname{rank}\mathbf{A}=4$, 又 $\mathbf{B}$ 的行秩至多为 4, 故只能 $\operatorname{rank}\mathbf{A}=\operatorname{rank}\mathbf{B}=4$, 方程组有唯一解. 即 $\mathbf{b}$ 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.

(3) 若 $\lambda=-1$, $\mu=0$, 则 $\operatorname{rank}\mathbf{A}=\operatorname{rank}\mathbf{B}=2$, 方程组有无穷多组解. 即 $\mathbf{b}$ 可以由 $\mathbf{a}_1$, $\mathbf{a}_2$, $\mathbf{a}_3$, $\mathbf{a}_4$ 线性表出, 且表达方式不唯一. □

例题 2.47. 当 λ 取不同值时, 求如下 $n\times n$ 阶矩阵的秩:

$$\mathbf{A}=\begin{pmatrix}\lambda&1&\cdots&1\\1&\lambda&\ddots&\vdots\\\vdots&\ddots&\ddots&1\\1&\cdots&1&\lambda\end{pmatrix}.$$

证明. 对 $\mathbf{A}$ 做如下行、列变换:

$$\mathbf{A}=\begin{pmatrix}\lambda&1&\cdots&1\\1&\lambda&\ddots&\vdots\\\vdots&\ddots&\ddots&1\\1&\cdots&1&\lambda\end{pmatrix}\xrightarrow[\text{加到第 1 行}]{\text{将第 }2,\ldots,n\text{ 行}}\begin{pmatrix}\lambda+n-1&\lambda+n-1&\cdots&\lambda+n-1\\1&\lambda&\ddots&\vdots\\\vdots&\ddots&\ddots&1\\1&\cdots&1&\lambda\end{pmatrix}$$
$$\xrightarrow[\text{加到第 }2,\ldots,n\text{ 列}]{\text{将第 1 列的 }(-1)\text{ 倍}}\mathbf{B}=\begin{pmatrix}\lambda+n-1&&&\\1&\lambda-1&&\\\vdots&&\ddots&\\1&&&\lambda-1\end{pmatrix}.$$

(1) 若 $\lambda = 1$, 则 $\mathrm{rank}\mathbf{A} = 1$.

(2) 若 $\lambda = 1 - n$, 则 $\mathrm{rank}\ \mathbf{A} = n - 1$.

(3) 若 $\lambda \neq 1, 1 - n$, 则 $\mathrm{rank}\ \mathbf{A} = n$. □

第 3 章 矩阵运算

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:

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

回忆在 $\mathbb{R}^3$ 中, 两个向量 $\mathbf{a}, \mathbf{b}$ 的内积在直角坐标系下的公式:

$$\mathbf{a} \cdot \mathbf{b} = (a_1, a_2, a_3) \cdot (b_1, b_2, b_3) = a_1b_1 + a_2b_2 + a_3b_3. \tag{3.2}$$

类似的, 在 $\mathbb{R}^m$ 中可以定义两个向量 $\mathbf{a}, \mathbf{b}$ 的 “内积” 运算:

$$\mathbf{a} \cdot \mathbf{b} = (a_1, \ldots, a_m) \cdot (b_1, \ldots, b_m) = a_1b_1 + \cdots + a_mb_m, \tag{3.3}$$

这样方程的第 i 行可以表示为

$$(a_{i1}, \ldots, a_{im}) \cdot (x_1, \ldots, x_m) = b_i. \tag{3.4}$$

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

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1m} \\ a_{21} & a_{22} & \cdots & a_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nm} \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_m \end{pmatrix} = \begin{pmatrix} a_{11}x_1 + \cdots + a_{1m}x_m \\ a_{21}x_1 + \cdots + a_{2m}x_m \\ \vdots \\ a_{n1}x_1 + \cdots + a_{nm}x_m \end{pmatrix}. \tag{3.5}$$

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

$$\mathbf{A} \cdot \mathbf{x} = \mathbf{b}. \tag{3.6}$$

定义 3.2. 设 $\mathbf{A}$ 是 $n \times m$ 阶矩阵, $\mathbf{B}$ 是 $m \times \ell$ 阶矩阵, 且 $\mathbf{B}$ 的行向量表示为 $\mathbf{B} = \begin{pmatrix} \mathbf{b}_1 & \cdots & \mathbf{b}_\ell \end{pmatrix}$. $\mathbf{A}$ 和 $\mathbf{B}$ 的矩阵乘积 $\mathbf{A} \cdot \mathbf{B}$ 是一个 $n \times \ell$ 阶矩阵, 且

$$\mathbf{A} \cdot \mathbf{B} = \begin{pmatrix} \mathbf{A}\mathbf{b}_1 & \cdots & \mathbf{A}\mathbf{b}_\ell \end{pmatrix}. \tag{3.7}$$

其中对 $\forall\, 1 \leq j \leq \ell$, $\mathbf{A}\mathbf{b}_i$ 是定义 3.1 中定义的 $n \times m$ 阶矩阵和 m 维列向量的乘积.

可以验证若 $\mathbf{A} = \begin{bmatrix} \mathbf{a}_1 & \cdots & \mathbf{a}_n \end{bmatrix}$, $\mathbf{B} = \begin{pmatrix} \mathbf{b}_1 & \cdots & \mathbf{b}_\ell \end{pmatrix}$, 则

$$\mathbf{AB} = \begin{pmatrix} \mathbf{a}_1\mathbf{B} \\ \vdots \\ \mathbf{a}_n\mathbf{B} \end{pmatrix} = \begin{pmatrix} \mathbf{a}_1 \cdot \mathbf{b}_1 & \mathbf{a}_1 \cdot \mathbf{b}_2 & \cdots & \mathbf{a}_1 \cdot \mathbf{b}_\ell \\ \mathbf{a}_2 \cdot \mathbf{b}_1 & \mathbf{a}_2 \cdot \mathbf{b}_2 & \cdots & \mathbf{a}_2 \cdot \mathbf{b}_\ell \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{a}_n \cdot \mathbf{b}_1 & \mathbf{a}_n \cdot \mathbf{b}_2 & \cdots & \mathbf{a}_n \cdot \mathbf{b}_\ell \end{pmatrix}. \tag{3.8}$$

以后我们都直接用 $\mathbf{AB}$ 表示两个矩阵的乘积.

例题 3.3. 计算 (1) $\begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} \begin{pmatrix} b_1 & \cdots & b_m \end{pmatrix}$. (2) $\begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix} \begin{pmatrix} \cos\beta & -\sin\beta \\ \sin\beta & \cos\beta \end{pmatrix}$.

解. (1) $\begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} \begin{pmatrix} b_1 & \cdots & b_m \end{pmatrix} = \begin{pmatrix} a_1b_1 & \cdots & a_1b_m \\ \vdots & \ddots & \vdots \\ a_nb_1 & \cdots & a_nb_m \end{pmatrix}$.

(2) $\begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix} \begin{pmatrix} \cos\beta & -\sin\beta \\ \sin\beta & \cos\beta \end{pmatrix} =$

$$\begin{pmatrix} \cos\alpha\cos\beta - \sin\alpha\sin\beta & -\sin\beta\cos\alpha - \sin\alpha\cos\beta \\ \sin\alpha\cos\beta + \sin\beta\cos\alpha & -\sin\alpha\sin\beta + \cos\alpha\cos\beta \end{pmatrix} = \begin{pmatrix} \cos(\alpha+\beta) & -\sin(\alpha+\beta) \\ \sin(\alpha+\beta) & \cos(\alpha+\beta) \end{pmatrix}.$$

□

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

$$\cdot : \quad \mathsf{M}_{n\times m}(\mathbb{F})\times \mathsf{M}_{m\times \ell}(\mathbb{F}) \longrightarrow \mathsf{M}_{n\times \ell}(\mathbb{F}). \tag{3.9}$$

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:

$$\mathbf{A}+\mathbf{B} = (a_{ij}+b_{ij})_{n\times m}, \tag{3.10}$$

$$\lambda\mathbf{A} = (\lambda a_{ij})_{n\times m}. \tag{3.11}$$

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. 记 $\mathbf{E}_{ij}\in \mathsf{M}_{n\times m}(\mathbb{F})$ 是 (i,j) 位置为 1, 其他位置全为 0 的矩阵:

$$\mathbf{E}_{ij} := \begin{array}{c} \begin{array}{ccccc} & & j & & \end{array} \\ \left(\begin{array}{ccccc} & & \vdots & & \\ & & 0 & & \\ \cdots & 0 & 1 & 0 & \cdots \\ & & 0 & & \\ & & \vdots & & \end{array}\right)_{i} \end{array}, \quad 1\le i\le n, 1\le j\le m. \tag{3.12}$$

则 $\{\mathbf{E}_{ij}\}_{1\le i\le n}^{1\le j\le m}$ 构成 $\mathsf{M}_{n\times m}(\mathbb{F})$ 的一组基, 称为 $n\times m$ 阶矩阵空间的标准基, 并且有公式:

$$\mathbf{E}_{ij}\mathbf{E}_{k\ell} = \delta_{jk}\mathbf{E}_{i\ell}, \quad \delta_{jk} = \begin{cases} 1 & j=k \\ 0 & j\neq k \end{cases}. \tag{3.13}$$

$n\times n$ 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. 矩阵乘法满足如下性质:

(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: $\lambda(\mathbf{AB}) = (\lambda\mathbf{A})\mathbf{B} = \mathbf{A}(\lambda\mathbf{B})$.

证明. 我们只验证乘法结合律, 其他性质的验证是容易的. 设 $\mathbf{A}=(a_{ij})_{n\times m}$, $\mathbf{B}=(b_{jk})_{m\times s}$, $\mathbf{C}=(c_{k\ell})_{s\times t}$, $\mathbf{D}=\mathbf{AB}=(d_{ik})_{n\times s}$, 则 $d_{ik}=\sum_{j=1}^m a_{ij}b_{jk}$, 从而

$$(\mathbf{AB})\mathbf{C}=(d_{ik})_{n\times s}(c_{k\ell})_{s\times t}=(\sum_{k=1}^s d_{ik}c_{k\ell})_{n\times t}=(\sum_{k=1}^s\sum_{j=1}^m a_{ij}b_{jk}c_{k\ell})_{n\times t}. \tag{3.14}$$

类似的, $\mathbf{A}(\mathbf{BC})$ 的 (i,ℓ) 位置元素也为 $\sum_{k=1}^s\sum_{j=1}^m a_{ij}b_{jk}c_{k\ell}$. □

特别的, $\mathsf{M}_n(\mathbb{F})$ 在矩阵加法, 乘法下构成一个环, 称为矩阵环. 只要 $n>1$, 则 $\mathsf{M}_n(\mathbb{F})$ 是非交换环, 并且两个非零矩阵的乘积可能是零矩阵. 例如, 在 $\mathsf{M}_2(\mathbb{F})$ 中, 由 (3.13), 有

$$\mathbf{E}_{12}\mathbf{E}_{21}=\mathbf{E}_{11},\quad \mathbf{E}_{21}\mathbf{E}_{12}=\mathbf{E}_{22},\quad \mathbf{E}_{12}\mathbf{E}_{12}=0. \tag{3.15}$$

例题 3.5. 设 $\mathbf{A}\in\mathsf{M}_{n\times m}(\mathbb{F})$, $\mathbf{B}\in\mathsf{M}_{m\times\ell}(\mathbb{F})$, 证明 $(\mathbf{AB})^{\mathrm{T}}=\mathbf{B}^{\mathrm{T}}\mathbf{A}^{\mathrm{T}}$.

证明. 设 $\mathbf{A}=(a_{ij})_{n\times m}$, $\mathbf{b}=[b_1,\ldots,b_m]\in\mathbb{F}^m$, 首先我们有

$$(\mathbf{Ab})^{\mathrm{T}}=\begin{pmatrix}\sum_{j=1}^m a_{1j}b_j & \cdots & \sum_{j=1}^n a_{nj}b_j\end{pmatrix}=\mathbf{b}^{\mathrm{T}}\mathbf{A}^{\mathrm{T}}.$$

接下来设 $\mathbf{B}=\begin{pmatrix}\mathbf{b}_1 & \cdots & \mathbf{b}_\ell\end{pmatrix}$, 于是

$$(\mathbf{AB})^{\mathrm{T}}=\begin{pmatrix}\mathbf{Ab}_1 & \cdots & \mathbf{Ab}_\ell\end{pmatrix}^{\mathrm{T}}=\begin{pmatrix}\mathbf{b}_1^{\mathrm{T}}\mathbf{A}^{\mathrm{T}}\\ \vdots\\ \mathbf{b}_\ell^{\mathrm{T}}\mathbf{A}^{\mathrm{T}}\end{pmatrix}=\begin{pmatrix}\mathbf{b}_1^{\mathrm{T}}\\ \vdots\\ \mathbf{b}_\ell^{\mathrm{T}}\end{pmatrix}\mathbf{A}^{\mathrm{T}}=\mathbf{B}^{\mathrm{T}}\mathbf{A}^{\mathrm{T}}.$$ □

例题 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.

证明. 显然纯量阵 $\lambda\mathbf{I}_n$ 和任意 n 阶方阵可交换. 另一方面, 设 $\mathbf{A}=\sum_{i=1}^n\sum_{j=1}^n a_{ij}\mathbf{E}_{ij}$ 与任意 n 阶方阵乘法交换, 特别的, 对任意 $1\le s,t\le n$, $\mathbf{AE}_{st}=\mathbf{E}_{st}\mathbf{A}$. 而

$$\mathbf{AE}_{st}=\sum_{i=1}^n\sum_{j=1}^n a_{ij}\mathbf{E}_{ij}\mathbf{E}_{st}=\sum_{i=1}^n a_{is}\mathbf{E}_{it}, \tag{3.16}$$

$$\mathbf{E}_{st}\mathbf{A}=\sum_{i=1}^n\sum_{j=1}^n a_{ij}\mathbf{E}_{st}\mathbf{E}_{ij}=\sum_{j=1}^n a_{tj}\mathbf{E}_{sj}. \tag{3.17}$$

比较两式得到对 $\forall\ i\neq j$, $a_{ij}=0$, 且 $a_{ii}=a_{jj}$, 即 $\mathbf{A}=\lambda\mathbf{I}_n$. □

定义 3.7. 对任意矩阵 $\mathbf{A} = (a_{ij})_{n\times m}$, $\mathbf{A}$ 的零空间为

$$\text{Null } \mathbf{A} = \left\{\mathbf{a} \in \mathbb{F}^m \mid \mathbf{A}\cdot\mathbf{a} = \mathbf{0}\right\}. \tag{3.18}$$

由 (3.6), Null $\mathbf{A}$ 也是齐次线性方程组 $\mathbf{A}\cdot\mathbf{x} = \mathbf{0}$ 的解空间 Ω_0, 从而

$$\dim \text{Null } \mathbf{A} = \dim \Omega_0. \tag{3.19}$$

定义 3.8. 对任意矩阵 $\mathbf{A} = (a_{ij})_{n\times m}$, $\mathbf{A}$ 的列空间为

$$\text{Col } \mathbf{A} = \left\{\mathbf{b} \in \mathbb{F}^n \;\middle|\; \exists\, \mathbf{a} \in \mathbb{F}^m,\ \mathbf{A}\cdot\mathbf{a} = \mathbf{b}\right\}. \tag{3.20}$$

引理 3.9. 矩阵 $\mathbf{A}$ 的列空间是 $\mathbb{F}^n$ 的子空间.

证明. 对 $\forall\, \mathbf{b}_1 = \mathbf{A}\mathbf{a}_1, \mathbf{b}_2 = \mathbf{A}\mathbf{a}_2 \in \text{Col } \mathbf{A}$,

$$\lambda\mathbf{b}_1 + \mu\mathbf{b}_2 = \lambda\mathbf{A}\mathbf{a}_1 + \mu\mathbf{A}\mathbf{a}_2 = \mathbf{A}(\lambda\mathbf{a}_1 + \mu\mathbf{a}_2) \in \text{Col } \mathbf{A}.$$ □

命题 3.10. 设 $\mathbf{A} = \begin{pmatrix}\mathbf{b}_1 & \cdots & \mathbf{b}_m\end{pmatrix}$, 即 $\mathbf{b}_1, \ldots, \mathbf{b}_m$ 是 $\mathbf{A}$ 的列向量组, 则

$$\text{Col } \mathbf{A} = \text{span}\{\mathbf{b}_1, \ldots, \mathbf{b}_m\}. \tag{3.21}$$

从而 $\dim \text{Col } \mathbf{A} = \text{rank } \mathbf{A}$.

证明. 因为 $\mathbf{A}\boldsymbol{\epsilon}_i = \mathbf{b}_i$, 从而 $\text{span}\{\mathbf{b}_1, \ldots, \mathbf{b}_m\} \subset \text{Col } \mathbf{A}$. 反过来, 对 $\forall\, \mathbf{a} \in \mathbb{F}^m$,

$$\mathbf{A}\mathbf{a} = \mathbf{A}(a_1\boldsymbol{\epsilon}_1 + \cdots + a_m\boldsymbol{\epsilon}_m) = a_1\mathbf{b}_1 + \cdots + a_m\mathbf{b}_m \in \text{span}\{\mathbf{b}_1, \ldots, \mathbf{b}_m\},$$

从而 $\text{Col } \mathbf{A} \subset \text{span}\{\mathbf{b}_1, \ldots, \mathbf{b}_m\}$. □

由定理 2.38, (3.19)以及命题 3.10, 我们得到

定理 3.11. 设 $\mathbf{A} \in \mathsf{M}_{n\times m}(\mathbb{F})$, 则 $\dim \text{Null } \mathbf{A} + \dim \text{Col } \mathbf{A} = m$.

例题 3.12. 设 $\mathbf{A} = \begin{pmatrix} 1 & 1 & & \\ & 1 & \ddots & \\ & & \ddots & 1 \\ & & & 1 \end{pmatrix}_{n\times n}$, 计算 $\mathbf{A}^k,\ k \in \mathbb{N}$.

解. 设 $\mathbf{B} = \begin{pmatrix} 0 & 1 & & \\ & 0 & \ddots & \\ & & \ddots & 1 \\ & & & 0 \end{pmatrix}_{n\times n}$, 则 $\mathbf{A} = \mathbf{I}_n + \mathbf{B}$, 并且

$$\mathbf{B}^\ell = \begin{array}{cc} \begin{matrix} & & & \ell+1 & & \end{matrix} & \\ \begin{pmatrix} 0 & 0 & \cdots & 1 & & \\ & 0 & 0 & & \ddots & \\ & & \ddots & \ddots & & 1 \\ & & & \ddots & \ddots & \vdots \\ & & & & 0 & 0 \\ & & & & & 0 \end{pmatrix} & \begin{matrix} \\ \\ \ell+1 \\ \\ \\ \\ \end{matrix} \end{array}, \quad 1 \le \ell \le n-1; \quad \mathbf{B}^\ell = 0, \quad \ell \ge n.$$

$\mathbf{A}^k = (\mathbf{I}_n + \mathbf{B})^k = \mathbf{I}_n + \binom{k}{1}\mathbf{B} + \cdots + \binom{k}{k}\mathbf{B}^k$, 其中 $\binom{k}{i} = \dfrac{k!}{i!(k-i)!}$ 是组合数. 从而

$$\mathbf{A}^k = \begin{pmatrix} 1 & \binom{k}{1} & \cdots & \binom{k}{n-1} \\ & 1 & \ddots & \vdots \\ & & \ddots & \binom{k}{1} \\ & & & 1 \end{pmatrix}. \quad (\text{若 } \ell > k,\ \text{则 } \tbinom{k}{\ell} = 0.)$$ □

例题 3.13. (1) 设 $\mathbf{A}, \mathbf{B} \in \mathsf{M}_{n\times m}(\mathbb{F})$, 则

$$\operatorname{rank}(\mathbf{A} + \mathbf{B}) \le \operatorname{rank} \mathbf{A} + \operatorname{rank} \mathbf{B}. \tag{3.22}$$

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

$$\operatorname{rank} \mathbf{AB} \le \min\{\operatorname{rank} \mathbf{A}, \operatorname{rank} \mathbf{B}\}. \tag{3.23}$$

证明. (1) 设 $\mathbf{A} = \begin{pmatrix} \mathbf{a}_1 & \cdots & \mathbf{a}_m \end{pmatrix}$, $\mathbf{B} = \begin{pmatrix} \mathbf{b}_1 & \cdots & \mathbf{b}_m \end{pmatrix}$, 则

$$\mathbf{A} + \mathbf{B} = \begin{pmatrix} \mathbf{a}_1 + \mathbf{b}_1 & \cdots & \mathbf{a}_m + \mathbf{b}_m \end{pmatrix},$$

且 $\operatorname{rank}(\mathbf{A} + \mathbf{B}) = \operatorname{rank} T_1 = \operatorname{rank}\{\mathbf{a}_1 + \mathbf{b}_1, \ldots, \mathbf{a}_m + \mathbf{b}_m\}$.

设 $\{\mathbf{a}_{i_1}, \ldots, \mathbf{a}_{i_r}\}$ 是 $\mathbf{A}$ 的列向量组的极大线性无关组, $\{\mathbf{b}_{j_1}, \ldots, \mathbf{b}_{j_s}\}$ 是 $\mathbf{B}$ 的列向量组的

极大线性无关组, 将它们拼起来得到向量组

$$T_2 = \{\mathbf{a}_{i_1}, \dots, \mathbf{a}_{i_r}, \mathbf{b}_{j_1}, \dots, \mathbf{b}_{j_s}\}.$$

对 $\forall\ 1 \le i \le m$, $\mathbf{a}_i$ 是 $\mathbf{a}_{i_1}, \dots, \mathbf{a}_{i_r}$ 的线性组合, $\mathbf{b}_i$ 是 $\mathbf{b}_{j_1}, \dots, \mathbf{b}_{j_s}$ 的线性组合, 所以 $\mathbf{a}_i + \mathbf{b}_i$ 是 T_2 中向量的线性组合, 由例题 2.29, 得到

$$\operatorname{rank}(\mathbf{A}+\mathbf{B}) = \operatorname{rank}\ T_1 \le \operatorname{rank}\ T_2 \le r+s = \operatorname{rank}\ \mathbf{A} + \operatorname{rank}\ \mathbf{B}.$$

(2) 设 $\mathbf{A} = (a_{ij})_{n\times m}$, $\mathbf{B} = \begin{bmatrix} \mathbf{b}_1 & \cdots & \mathbf{b}_m \end{bmatrix}$, 写为 $\mathbf{B}$ 的列向量形式, 则

$$\mathbf{AB} = \begin{pmatrix} a_{11} & \cdots & a_{1m} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nm} \end{pmatrix} \begin{pmatrix} \mathbf{b}_1 \\ \vdots \\ \mathbf{b}_n \end{pmatrix} = \begin{pmatrix} a_{11}\mathbf{b}_1 + \cdots a_{1m}\mathbf{b}_m \\ \vdots \\ a_{n1}\mathbf{b}_1 + \cdots + a_{nm}\mathbf{b}_m \end{pmatrix}, \tag{3.24}$$

即 $\mathbf{AB}$ 的每个行向量都是 $\mathbf{B}$ 的行向量组的线性组合, 从而由例题 2.29,

$$\operatorname{rank}\ \mathbf{AB} \le \operatorname{rank}\ \mathbf{B}.$$

另一方面, 设 $\mathbf{A} = \begin{pmatrix} \mathbf{a}_1 & \cdots & \mathbf{a}_m \end{pmatrix}$, 写为 $\mathbf{A}$ 的行向量形式, $\mathbf{B} = (b_{jk})_{m\times\ell}$, 则

$$\begin{aligned} \mathbf{AB} &= \begin{pmatrix} \mathbf{a}_1 & \cdots & \mathbf{a}_m \end{pmatrix} \begin{pmatrix} b_{11} & \cdots & b_{1\ell} \\ \vdots & \cdots & \vdots \\ b_{m1} & \cdots & b_{m\ell} \end{pmatrix} \\ &= \begin{pmatrix} b_{11}\mathbf{a}_1 + \cdots + b_{m1}\mathbf{a}_m & \cdots & b_{1\ell}\mathbf{a}_1 + \cdots + b_{m\ell}\mathbf{a}_m \end{pmatrix}, \end{aligned} \tag{3.25}$$

即 $\mathbf{AB}$ 的每个列向量都是 $\mathbf{A}$ 的列向量组的线性组合, 从而由例题 2.29,

$$\operatorname{rank}\ \mathbf{AB} \le \operatorname{rank}\ \mathbf{A}.$$

□

评注 3.14. 事实上, 由 $\operatorname{rank}\ \mathbf{AB} \le \operatorname{rank}\ \mathbf{B}$, 以及 $\operatorname{rank}\ \mathbf{A} = \operatorname{rank}\ \mathbf{A}^{\mathrm{T}}$, 有

$$\operatorname{rank}\ \mathbf{AB} = \operatorname{rank}\ (\mathbf{AB})^{\mathrm{T}} = \operatorname{rank}\ \mathbf{B}^{\mathrm{T}}\mathbf{A}^{\mathrm{T}} \le \operatorname{rank}\ \mathbf{A}^{\mathrm{T}} = \operatorname{rank}\ \mathbf{A}.$$

即得到 $\operatorname{rank}\ \mathbf{AB} \le \operatorname{rank}\ \mathbf{A}$ 的另一种证法.

例题 3.15. 设 $\mathbf{A} \in \mathsf{M}_{n\times m}(\mathbb{R})$, 证明 $\operatorname{rank}\ \mathbf{A}^{\mathrm{T}}\mathbf{A} = \operatorname{rank}\ \mathbf{A}$.

证明. 由 (3.23), $\text{rank } \mathbf{A}^{\mathrm{T}}\mathbf{A} \le \text{rank } \mathbf{A}$. 另一方面, 对 $\forall\ \mathbf{x} \in \text{Null } \mathbf{A}^{\mathrm{T}}\mathbf{A}$, 有

$$\mathbf{A}^{\mathrm{T}}\mathbf{A}\mathbf{x} = 0 \Longrightarrow \mathbf{x}^{\mathrm{T}}\mathbf{A}^{\mathrm{T}}\mathbf{A}\mathbf{x} = 0 \Longrightarrow (\mathbf{A}\mathbf{x})^{\mathrm{T}}\mathbf{A}\mathbf{x} = 0.$$

设 $\mathbf{A}\mathbf{x} = [y_1, \ldots, y_n] \in \mathbb{R}^n$, 则

$$(\mathbf{A}\mathbf{x})^{\mathrm{T}}\mathbf{A}\mathbf{x} = 0 \Longrightarrow y_1^2 + \cdots + y_n^2 = 0 \Longrightarrow y_i = 0,\ 1 \le i \le n$$

即 $\mathbf{A}\mathbf{x} = 0$. 从而 $\text{Null } \mathbf{A}^{\mathrm{T}}\mathbf{A} \subset \text{Null } \mathbf{A}$, 则 $\dim \text{Null } \mathbf{A}^{\mathrm{T}}\mathbf{A} \le \dim \text{Null } \mathbf{A}$. 由定理 3.11,

$$\text{rank } \mathbf{A}^{\mathrm{T}}\mathbf{A} = m - \dim \text{Null } \mathbf{A}^{\mathrm{T}}\mathbf{A} \ge m - \dim \text{Null } \mathbf{A} = \text{rank } \mathbf{A}.$$

□

例题 3.16. 设 $\mathbf{A} \in \mathsf{M}_n(\mathbb{F})$, $k \in \mathbb{N}$. 如果 $\text{rank } \mathbf{A}^k = \text{rank } \mathbf{A}^{k+1}$, 证明 $\text{rank } \mathbf{A}^k = \text{rank } \mathbf{A}^{k+1} = \text{rank } \mathbf{A}^{k+2} = \cdots$.

证明. 由 $\text{rank } \mathbf{A}^k = \text{rank } \mathbf{A}^{k+1}$ 有 $\dim \text{Col } \mathbf{A}^k = \dim \text{Col } \mathbf{A}^{k+1}$. 又因为 $\text{Col } \mathbf{A}^{k+1} \subset \text{Col } \mathbf{A}^k$, 故 $\text{Col } \mathbf{A}^{k+1} = \text{Col } \mathbf{A}^k$. 接下来证明 $\text{Col } \mathbf{A}^{k+2} = \text{Col } \mathbf{A}^{k+1}$, 或 $\text{Col } \mathbf{A}^{k+2} \supset \text{Col } \mathbf{A}^{k+1}$.

对 $\forall\ \mathbf{x} \in \text{Col } \mathbf{A}^{k+1}$, 存在 $\mathbf{y} \in \mathbb{F}^n$ 使得 $\mathbf{x} = \mathbf{A}^{k+1}\mathbf{y} = \mathbf{A}(\mathbf{A}^k\mathbf{y})$. 因为 $\mathbf{A}^k\mathbf{y} \in \text{Col } \mathbf{A}^k = \text{Col } \mathbf{A}^{k+1}$, 存在 $\mathbf{z} \in \mathbb{F}^n$ 使得 $\mathbf{A}^k\mathbf{y} = \mathbf{A}^{k+1}\mathbf{z}$. 从而 $\mathbf{x} = \mathbf{A}(\mathbf{A}^{k+1}\mathbf{z}) = \mathbf{A}^{k+2}\mathbf{z} \in \text{Col } \mathbf{A}^{k+2}$. □

3.2　可逆矩阵

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

定义 3.17. 设 $\mathbf{A} \in \mathsf{M}_n(\mathbb{F})$, 如果存在 $\mathbf{B} \in \mathsf{M}_n(\mathbb{F})$ 使得

$$\mathbf{A}\mathbf{B} = \mathbf{B}\mathbf{A} = \mathbf{I}_n, \tag{3.26}$$

称 $\mathbf{A}$ 可逆, $\mathbf{B}$ 是 $\mathbf{A}$ 的逆矩阵.

引理 3.18. (1) 若 $\mathbf{A}$ 可逆, 则其逆矩阵唯一, 记作 $\mathbf{A}^{-1}$.

(2) $(\mathbf{A}^{-1})^{-1} = \mathbf{A}$, $\quad (\lambda\mathbf{A})^{-1} = \lambda^{-1}\mathbf{A}^{-1}$ $(\lambda \ne 0)$.

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

证明. (1) 设 $\mathbf{B}, \mathbf{C}$ 都是 $\mathbf{A}$ 的逆矩阵, 则

$$\mathbf{C} = \mathbf{C}\mathbf{I}_n = \mathbf{C}(\mathbf{A}\mathbf{B}) = (\mathbf{C}\mathbf{A})\mathbf{B} = \mathbf{I}_n\mathbf{B} = \mathbf{B}.$$

(2),(3) 直接验证即可. □

定理 3.19. 设 $\mathbf{A} \in \mathsf{M}_n(\mathbb{F})$, 则 $\mathbf{A}$ 可逆 $\Longleftrightarrow \operatorname{rank} \mathbf{A} = n$.

证明. "$\Longrightarrow$": 设 $\mathbf{B}$ 是 $\mathbf{A}$ 的逆矩阵, 则 $\mathbf{AB} = \mathbf{I}_n$, 由 (3.23) 得到

$$n = \operatorname{rank} \mathbf{I}_n = \operatorname{rank} \mathbf{AB} \le \operatorname{rank} \mathbf{A} \le n,$$

从而 $\operatorname{rank} \mathbf{A} = n$.

"$\Longleftarrow$": 设 $\operatorname{rank} \mathbf{A} = n$, 则对 $\forall\, \mathbf{b} \in \mathbb{F}^n$, 有

$$n = \operatorname{rank} \mathbf{A} \le \operatorname{rank} (\mathbf{A}|\mathbf{b}) \le n,$$

从而 $\operatorname{rank} \mathbf{A} = \operatorname{rank} (\mathbf{A}|\mathbf{b}) = n$. 由命题 2.44, 对 $\forall\, \mathbf{b} \in \mathbb{F}^n$, 线性方程组 $\mathbf{Ax} = \mathbf{b}$ 有唯一解. 特别的, 对每个 $\boldsymbol{\epsilon}_i$ $(1 \le i \le n)$, 设 $\mathbf{b}_i$ 是 $\mathbf{Ax} = \boldsymbol{\epsilon}_i$ 的唯一解, 且 $\mathbf{B} = \begin{pmatrix} \mathbf{b}_1 & \cdots & \mathbf{b}_n \end{pmatrix}$, 则

$$\mathbf{AB} = \mathbf{A}\begin{pmatrix} \mathbf{b}_1 & \cdots & \mathbf{b}_n \end{pmatrix} = \begin{pmatrix} \boldsymbol{\epsilon}_1 & \cdots & \boldsymbol{\epsilon}_n \end{pmatrix} = \mathbf{I}_n.$$

接下来需要证明 $\mathbf{BA} = \mathbf{I}_n$. 由 $\mathbf{AB} = \mathbf{I}_n$ 可知 $\operatorname{rank} \mathbf{B} = n$, 用同样的方法可知存在 $\mathbf{C} \in \mathsf{M}_n(\mathbb{F})$ 使得 $\mathbf{BC} = \mathbf{I}_n$, 从而

$$\mathbf{A} = \mathbf{AI}_n = \mathbf{A}(\mathbf{BC}) = (\mathbf{AB})\mathbf{C} = \mathbf{I}_n\mathbf{C} = \mathbf{C},$$

即 $\mathbf{BA} = \mathbf{I}_n$. □

回顾定理 3.19 的证明过程, 我们实际证明了:

推论 3.20. 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 $\mathbf{CA} = \mathbf{I}_n$.

例题 3.21. 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 $1 \le i \le n$ 成立, 称 $\mathbf{A}$ 是对角占优的. 证明对角占优方阵可逆.

证明. 只需证明 $\operatorname{rank} \mathbf{A} = n$, 或等价的, $\operatorname{Null} \mathbf{A} = \{\mathbf{0}\}$. 用反证法, 设 $\mathbf{x} = [x_1, \dots, x_n] \neq \mathbf{0}$, 且 $\mathbf{Ax} = \mathbf{0}$. 设 $|x_i| = \max\{|x_1, \dots, |x_n||\}$, 则 $|x_i| \neq 0$. 考察方程组 $\mathbf{Ax} = \mathbf{0}$ 的第 i 个方程

$$a_{i1}x_1 + \cdots + a_{ii}x_i + \cdots + a_{in}x_n = 0,$$

可以推出

$$|a_{ii}||x_i| = \left|\sum_{j\neq i} a_{ij}x_j\right| \le \sum_{j\neq i} |a_{ij}||x_j| \le (\sum_{j\neq i} |a_{ij}|)|x_i|,$$

两边消去 $|x_i|$ 得到矛盾. $\square$

对任意 $1 \le i, j \le n$, 定义如下三类 $\mathsf{M}_n(\mathbb{F})$ 中的初等矩阵:

(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:

$$\mathbf{P}_{ij} \coloneqq \begin{array}{c@{}c} \begin{pmatrix} 1 & & & & & & \\ & \ddots & & & & & \\ & & 0 & & 1 & & \\ & & & \ddots & & & \\ & & 1 & & 0 & & \\ & & & & & \ddots & \\ & & & & & & 1 \end{pmatrix} & \begin{matrix} \\ \\ i \\ \\ j \\ \\ \\ \end{matrix} \end{array}, \quad 1 \le i, j \le n. \tag{3.27}$$

易知 $\mathbf{P}_{ij}^{-1} = \mathbf{P}_{ij}$.

(2) $\mathbf{D}_i(\lambda) = \mathbf{I}_n + (\lambda - 1)\mathbf{E}_{ii}, \lambda \ne 0$, 即将单位矩阵 $\mathbf{I}_n$ 的第 i 行乘以非零数 λ 得到的矩阵:

$$\mathbf{D}_i(\lambda) \coloneqq \begin{array}{c@{}c} \begin{pmatrix} 1 & & & & \\ & \ddots & & & \\ & & \lambda & & \\ & & & \ddots & \\ & & & & 1 \end{pmatrix} & \begin{matrix} \\ \\ i \\ \\ \\ \end{matrix} \end{array}, \quad 1 \le i, j \le n. \tag{3.28}$$

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:

$$\mathbf{P}_{ij} \coloneqq \begin{array}{c@{}c} \begin{pmatrix} 1 & & & & & & \\ & \ddots & & & & & \\ & & 1 & & \lambda & & \\ & & & \ddots & & & \\ & & & & 1 & & \\ & & & & & \ddots & \\ & & & & & & 1 \end{pmatrix} & \begin{matrix} \\ \\ i \\ \\ j \\ \\ \\ \end{matrix} \end{array}, \quad 1 \le i, j \le n. \tag{3.29}$$

易知 $\mathbf{T}_{ij}(\lambda)^{-1}=\mathbf{T}_{ij}(-\lambda)$.

从而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 3.22. 设 $\mathbf{A}\in\mathsf{M}_{n\times p}(\mathbb{F}),\mathbf{B}\in\mathsf{M}_{q\times n}(\mathbb{F})$, 则

(1) $\mathbf{P}_{ij}\mathbf{A}$ 是将 $\mathbf{A}$ 的 i 行和 j 行交换得到的矩阵. $\mathbf{BP}_{ij}$ 是将 $\mathbf{B}$ 的 i 列和 j 列交换得到的矩阵.

(2) $\mathbf{D}_i(\lambda)\mathbf{A}$ 是将 $\mathbf{A}$ 的第 i 行乘以非零数 λ 得到的矩阵. $\mathbf{BD}_i(\lambda)$ 是将 $\mathbf{B}$ 的第 i 列乘以非零数 λ 得到的矩阵.

(3) $\mathbf{T}_{ij}(\lambda)\mathbf{A}$ 是将 $\mathbf{A}$ 的第 j 行的 λ 倍加到第 i 行得到的矩阵. $\mathbf{BT}_{ij}(\lambda)$ 是将 $\mathbf{B}$ 的第 i 列的 λ 倍加到第 j 列得到的矩阵.

证明. 我们说明 $\mathbf{P}_{ij}$ 的情形. 设 $\mathbf{A}=\begin{bmatrix}\mathbf{a}_1 & \cdots & \mathbf{a}_n\end{bmatrix}$, $\mathbf{B}=\begin{pmatrix}\mathbf{b}_1 & \cdots & \mathbf{b}_n\end{pmatrix}$, 则

$$\mathbf{P}_{ij}\begin{bmatrix}\mathbf{a}_1 & \cdots & \mathbf{a}_i & \cdots & \mathbf{a}_j & \cdots & \mathbf{a}_n\end{bmatrix}=\begin{bmatrix}\mathbf{a}_1 & \cdots & \mathbf{a}_j & \cdots & \mathbf{a}_i & \cdots & \mathbf{a}_n\end{bmatrix}. \tag{3.30}$$

$$\begin{pmatrix}\mathbf{b}_1 & \cdots & \mathbf{b}_i & \cdots & \mathbf{b}_j & \cdots & \mathbf{b}_n\end{pmatrix}\mathbf{P}_{ij}=\begin{pmatrix}\mathbf{b}_1 & \cdots & \mathbf{b}_j & \cdots & \mathbf{b}_i & \cdots & \mathbf{b}_n\end{pmatrix}. \tag{3.31}$$

其他情形的证明是类似的. □

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. 从而

定理 3.23. 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 $\mathbf{P}_1,\dots,\mathbf{P}_k$, 以及 m 阶初等矩阵 $\mathbf{Q}_1,\dots,\mathbf{Q}_\ell$, 使得

$$\mathbf{P}_1\cdots\mathbf{P}_k\cdot\mathbf{A}\cdot\mathbf{Q}_1\cdots\mathbf{Q}_\ell=\begin{pmatrix}\mathbf{I}_r & 0\\ 0 & 0\end{pmatrix}. \tag{3.32}$$

推论 3.24. 可逆矩阵是一些初等矩阵的乘积. 特别的, 若 $\mathbf{A}\in\mathsf{M}_n(\mathbb{F})$ 可逆, 则存在初等矩阵 $\mathbf{P}_1,\dots,\mathbf{P}_k\in\mathsf{M}_n(\mathbb{F})$ 使得

$$\mathbf{P}_1\cdots\mathbf{P}_k\cdot\mathbf{A}=\mathbf{I}_n. \tag{3.33}$$

证明. 若 $\mathbf{A}$ 可逆, 由定理 3.19 知 $\operatorname{rank}\mathbf{A}=n$, 再由定理 3.23 知存在初等矩阵 $\mathbf{P}_1,\dots,\mathbf{P}_k$, $\mathbf{Q}_1,\dots,\mathbf{Q}_\ell$ 使得 $\mathbf{P}_1\cdots\mathbf{P}_k\cdot\mathbf{A}\cdot\mathbf{Q}_1\cdots\mathbf{Q}_\ell=\mathbf{I}_n$. 即得到

$$\mathbf{A}=\mathbf{P}_1^{-1}\cdots\mathbf{P}_k^{-1}\mathbf{Q}_\ell^{-1}\cdots\mathbf{Q}_1^{-1},$$

是一些初等方阵的乘积. 进一步的, 有

$$\mathbf{Q}_1\cdots\mathbf{Q}_\ell\mathbf{P}_1\cdots\mathbf{P}_k\cdot\mathbf{A}=\mathbf{I}_n,$$

即可通过一系列初等行变换将 $\mathbf{A}$ 化为单位阵 $\mathbf{I}_n$. □

另一方面, 若 $\mathbf{A}$ 可逆且 $\mathbf{P}_1\cdots\mathbf{P}_k\cdot\mathbf{A}=\mathbf{I}_n$, 则 $\mathbf{A}^{-1}=\mathbf{P}_1\cdots\mathbf{P}_k\cdot\mathbf{I}_n$, 从而 $\mathbf{A}^{-1}$ 可以看作将单位矩阵 $\mathbf{I}_n$ 依次作行变换 $\mathbf{P}_k,\mathbf{P}_{k-1},\dots,\mathbf{P}_1$ 得到的矩阵. 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:

$$(\mathbf{A}\mid\mathbf{I}_n)\xrightarrow{\mathbf{P}_k}(\mathbf{P}_k\mathbf{A}\mid\mathbf{P}_k)\xrightarrow{\mathbf{P}_{k-1}}\cdots\xrightarrow{\mathbf{P}_1}(\mathbf{I}_n\mid\mathbf{P}_1\cdots\mathbf{P}_k)=(\mathbf{I}_n\mid\mathbf{A}^{-1}).\tag{3.34}$$

例题 3.25. 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.

证明. 设 $\mathbf{A}=\mathbf{P}\begin{pmatrix}\mathbf{I}_r & 0\\ 0 & 0\end{pmatrix}\mathbf{Q}$, 则

$$\mathbf{A}=\mathbf{P}\begin{pmatrix}\mathbf{I}_r & 0\\ 0 & 0\end{pmatrix}\mathbf{Q}=\mathbf{P}\mathbf{E}_{11}\mathbf{Q}+\cdots+\mathbf{P}\mathbf{E}_{rr}\mathbf{Q}.$$ □

例题 3.26. 求矩阵 $\mathbf{A}=\begin{pmatrix}1&2&3&4\\2&3&1&2\\1&1&1&-1\\1&0&-2&-6\end{pmatrix}$ 的逆矩阵.

证明.

$$\left(\begin{array}{cccc|cccc}1&2&3&4&1&0&0&0\\2&3&1&2&0&1&0&0\\1&1&1&-1&0&0&1&0\\1&0&-2&6&0&0&0&1\end{array}\right)\longrightarrow\left(\begin{array}{cccc|cccc}1&1&1&-1&0&0&1&0\\2&3&1&2&0&1&0&0\\1&2&3&4&1&0&0&0\\1&0&-2&6&0&0&0&1\end{array}\right)\longrightarrow$$

$$\left(\begin{array}{cccc|cccc} 1 & 1 & 1 & -1 & 0 & 0 & 1 & 0 \\ 0 & 1 & -1 & 4 & 0 & 1 & -2 & 0 \\ 0 & 1 & 2 & 5 & 1 & 0 & -1 & 0 \\ 0 & -1 & -3 & -5 & 0 & 0 & -1 & 1 \end{array}\right) \longrightarrow \left(\begin{array}{cccc|cccc} 1 & 1 & 1 & -1 & 0 & 0 & 1 & 0 \\ 0 & 1 & -1 & 4 & 0 & 1 & -2 & 0 \\ 0 & 0 & 3 & 1 & 1 & -1 & 1 & 0 \\ 0 & 0 & -4 & -1 & 0 & 1 & -3 & 1 \end{array}\right) \longrightarrow$$

$$\left(\begin{array}{cccc|cccc} 1 & 1 & 1 & -1 & 0 & 0 & 1 & 0 \\ 0 & 1 & -1 & 4 & 0 & 1 & -2 & 0 \\ 0 & 0 & -1 & 0 & 1 & 0 & -2 & 1 \\ 0 & 0 & -4 & -1 & 0 & 1 & -3 & 1 \end{array}\right) \longrightarrow \left(\begin{array}{cccc|cccc} 1 & 1 & 1 & -1 & 0 & 0 & 1 & 0 \\ 0 & 1 & -1 & 4 & 0 & 1 & -2 & 0 \\ 0 & 0 & 1 & 0 & -1 & 0 & 2 & -1 \\ 0 & 0 & 0 & 1 & 4 & -1 & -5 & 3 \end{array}\right) \longrightarrow$$

$$\left(\begin{array}{cccc|cccc} 1 & 1 & 1 & 0 & 4 & -1 & -4 & 3 \\ 0 & 1 & -1 & 0 & -16 & 5 & 18 & -12 \\ 0 & 0 & 1 & 0 & -1 & 0 & 2 & -1 \\ 0 & 0 & 0 & 1 & 4 & -1 & -5 & 3 \end{array}\right) \longrightarrow \left(\begin{array}{cccc|cccc} 1 & 1 & 0 & 0 & 5 & -1 & -6 & 4 \\ 0 & 1 & 0 & 0 & -17 & 5 & 20 & -13 \\ 0 & 0 & 1 & 0 & -1 & 0 & 2 & -1 \\ 0 & 0 & 0 & 1 & 4 & -1 & -5 & 3 \end{array}\right) \longrightarrow$$

$$\left(\begin{array}{cccc|cccc} 1 & 0 & 0 & 0 & 22 & -6 & -26 & 17 \\ 0 & 1 & 0 & 0 & -17 & 5 & 20 & -13 \\ 0 & 0 & 1 & 0 & -1 & 0 & 2 & -1 \\ 0 & 0 & 0 & 1 & 4 & -1 & -5 & 3 \end{array}\right). \qquad \square$$

在之前表示矩阵 $\begin{pmatrix} \mathbf{I}_r & 0 \\ 0 & 0 \end{pmatrix}$, 以及将 $\mathbf{A}$ 写成它的列向量形式 $\mathbf{A} = \begin{bmatrix} \mathbf{a}_1, & \cdots & , \mathbf{a}_n \end{bmatrix}$ 以及行向量形式 $\mathbf{A} = \begin{pmatrix} \mathbf{b}_1, & \cdots & , \mathbf{b}_m \end{pmatrix}$ 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:

$$\begin{array}{c} \\ n_1 \\ n_2 \end{array} \overset{\begin{array}{cc} m_1 & m_2 \end{array}}{\begin{pmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{pmatrix}} \overset{\begin{array}{cc} \ell_1 & \ell_2 \end{array}}{\begin{pmatrix} \mathbf{A}' & \mathbf{B}' \\ \mathbf{C}' & \mathbf{D}' \end{pmatrix}} \begin{array}{c} m_1 \\ m_2 \end{array} = \begin{array}{c} n_1 \\ n_2 \end{array} \overset{\begin{array}{cc} \ell_1 & \ell_2 \end{array}}{\begin{pmatrix} \mathbf{AA}' + \mathbf{BC}' & \mathbf{AB}' + \mathbf{BD}' \\ \mathbf{CA}' + \mathbf{DC}' & \mathbf{CB}' + \mathbf{DD}' \end{pmatrix}}. \tag{3.35}$$

有了分块矩阵乘积, 就可以在分块矩阵中做行、列变换:

定理 3.27. 设 $\mathbf{A} \in \mathsf{M}_r(\mathbb{F})$, $\mathbf{B} \in \mathsf{M}_{r\times(n-r)}(\mathbb{F})$, $\mathbf{C} \in \mathsf{M}_{(n-r)\times r}(\mathbb{F})$, $\mathbf{D} \in \mathsf{M}_{(n-r)\times(n-r)}(\mathbb{F})$, 且 $\mathbf{A}$ 可逆, 则有

(1) $\begin{pmatrix} \mathbf{I}_r & \mathbf{0} \\ -\mathbf{CA}^{-1} & \mathbf{I}_{n-r} \end{pmatrix} \begin{pmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{pmatrix} = \begin{pmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{0} & \mathbf{D} - \mathbf{CA}^{-1}\mathbf{B} \end{pmatrix}$.

(2) $\begin{pmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{pmatrix} \begin{pmatrix} \mathbf{I}_r & -\mathbf{A}^{-1}\mathbf{B} \\ \mathbf{0} & \mathbf{I}_{n-r} \end{pmatrix} = \begin{pmatrix} \mathbf{A} & \mathbf{0} \\ \mathbf{C} & \mathbf{D}-\mathbf{C}\mathbf{A}^{-1}\mathbf{B} \end{pmatrix}$.

(3) $\begin{pmatrix} \mathbf{I}_r & \mathbf{0} \\ -\mathbf{C}\mathbf{A}^{-1} & \mathbf{I}_{n-r} \end{pmatrix} \begin{pmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{pmatrix} \begin{pmatrix} \mathbf{I}_r & -\mathbf{A}^{-1}\mathbf{B} \\ \mathbf{0} & \mathbf{I}_{n-r} \end{pmatrix} = \begin{pmatrix} \mathbf{A} & \mathbf{0} \\ \mathbf{0} & \mathbf{D}-\mathbf{C}\mathbf{A}^{-1}\mathbf{B} \end{pmatrix}$.

例题 3.28. 设 $\mathbf{A} = \begin{pmatrix} 1 & 1 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{pmatrix}$, 计算 $\mathbf{A}^n$.

证明. 设 $\mathbf{B} = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$, 则 $\mathbf{A} = \begin{pmatrix} \mathbf{B} & \mathbf{I}_2 \\ \mathbf{0} & \mathbf{B} \end{pmatrix}$. 因为 $\mathbf{A}^n = \begin{pmatrix} \mathbf{B}^n & n\mathbf{B}^{n-1} \\ \mathbf{0} & \mathbf{B}^n \end{pmatrix}$, $\mathbf{B}^n = \begin{pmatrix} 1 & n \\ 0 & 1 \end{pmatrix}$, 有

$$\mathbf{A}^n = \begin{pmatrix} 1 & n & n & n(n-1) \\ 0 & 1 & 0 & n \\ 0 & 0 & 1 & n \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

□

例题 3.29. 设 $\mathbf{A}, \mathbf{B} \in \mathsf{M}_n(\mathbb{F})$, 且 $\mathbf{AB} = 0$, 求 $\begin{pmatrix} \mathbf{I}_n & \mathbf{A} \\ \mathbf{B} & \mathbf{I}_n \end{pmatrix}^{-1}$.

证明. 做分块矩阵行变换:

$$\left(\begin{array}{cc|cc} \mathbf{I}_n & \mathbf{A} & \mathbf{I}_n & \mathbf{0} \\ \mathbf{B} & \mathbf{I}_n & \mathbf{0} & \mathbf{I}_n \end{array}\right) \longrightarrow \left(\begin{array}{cc|cc} \mathbf{I}_n & \mathbf{0} & \mathbf{I}_n & -\mathbf{A} \\ \mathbf{B} & \mathbf{I}_n & \mathbf{0} & \mathbf{I}_n \end{array}\right) \longrightarrow \left(\begin{array}{cc|cc} \mathbf{I}_n & \mathbf{0} & \mathbf{I}_n & -\mathbf{A} \\ \mathbf{0} & \mathbf{I}_n & -\mathbf{B} & \mathbf{I}_n + \mathbf{BA} \end{array}\right),$$

从而 $\begin{pmatrix} \mathbf{I}_n & \mathbf{A} \\ \mathbf{B} & \mathbf{I}_n \end{pmatrix}^{-1} = \begin{pmatrix} \mathbf{I}_n & -\mathbf{A} \\ -\mathbf{B} & \mathbf{I}_n + \mathbf{BA} \end{pmatrix}$. □

例题 3.30. 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 $\mathbf{A} \in \mathsf{M}_{n_1\times m_1}(\mathbb{F})$, $\mathbf{B} \in \mathsf{M}_{n_2\times m_2}(\mathbb{F})$, $\mathbf{C} \in \mathsf{M}_{n_1\times m_2}(\mathbb{F})$.

证明. 设 $\mathbf{A}$ 的行向量组的极大线性无关组为 $\mathbf{a}_{i_1}, \ldots, \mathbf{a}_{i_r}$, $\mathbf{B}$ 的行向量组的极大线性无关组为 $\mathbf{b}_{j_1}, \ldots, \mathbf{b}_{j_s}$. 在矩阵 $\begin{pmatrix} \mathbf{A} & \mathbf{C} \\ \mathbf{0} & \mathbf{B} \end{pmatrix}$ 中, $\mathbf{a}_{i_1}, \ldots, \mathbf{a}_{i_r}$ 所在行向量为

$$\mathbf{d}_{i_1} = (\mathbf{a}_{i_1}, \mathbf{c}_{i_1}), \ldots, \mathbf{d}_{i_r} = (\mathbf{a}_{i_r}, \mathbf{c}_{i_r}).$$

$\mathbf{b}_{j_1}, \ldots, \mathbf{b}_{j_s}$ 所在行向量为

$$\mathbf{e}_{j_1} = (\mathbf{0}, \mathbf{b}_{j_1}), \ldots, \mathbf{e}_{j_s} = (\mathbf{0}, \mathbf{b}_{j_s}).$$

显然 $\{\mathbf{d}_{i_1}, \ldots, \mathbf{d}_{i_r}, \mathbf{e}_{j_1}, \ldots, \mathbf{e}_{j_s}\}$ 线性无关, 故 $\operatorname{rank}\begin{pmatrix} \mathbf{A} & \mathbf{C} \\ \mathbf{0} & \mathbf{B} \end{pmatrix} \geq r + s = \operatorname{rank}\mathbf{A} + \operatorname{rank}\mathbf{B}$. □

3.3　行列式

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

$$|\mathbf{A}| = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}a_{22} - a_{21}a_{12}, \tag{3.36}$$

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

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = a_{31}\begin{vmatrix} a_{12} & a_{13} \\ a_{22} & a_{23} \end{vmatrix} - a_{32}\begin{vmatrix} a_{11} & a_{13} \\ a_{21} & a_{23} \end{vmatrix} + a_{33}\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}, \tag{3.37}$$

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 $\det_n \mathbf{A}$, 或简记为 $\det \mathbf{A}$, $|\mathbf{A}|$. 行列式是从 $\mathsf{M}_n(\mathbb{F})$ 到 $\mathbb{F}$ 的一个函数 $\det\colon \mathsf{M}_n(\mathbb{F}) \to \mathbb{F}$, 也可以看作关于 n 个 n 维行向量 $\mathbf{a}_1, \ldots, \mathbf{a}_n \in \mathbb{F}^n$ 的函数, 记作 $\det_n(\mathbf{a}_1, \ldots, \mathbf{a}_n)$.

要给出 n 阶行列式的定义, 我们先从二阶、三阶行列式出发, 收集行列式应满足的性质. 首先, 二阶行列式满足如下性质:

(1) $\begin{vmatrix} \lambda a + \mu a' & \lambda b + \mu b' \\ c & d \end{vmatrix} = \lambda\begin{vmatrix} a & b \\ c & d \end{vmatrix} + \mu\begin{vmatrix} a' & b' \\ c & d \end{vmatrix}$.

(2) $\begin{vmatrix} a & b \\ \lambda c + \mu c' & \lambda d + \mu d' \end{vmatrix} = \lambda\begin{vmatrix} a & b \\ c & d \end{vmatrix} + \mu\begin{vmatrix} a & b \\ c' & d' \end{vmatrix}$.

(3) $\begin{vmatrix} a & b \\ c & d \end{vmatrix} = -\begin{vmatrix} c & d \\ a & b \end{vmatrix}$.

(4) $\begin{vmatrix} 1 & 0 \\ 0 & 1 \end{vmatrix} = 1.$

将二阶行列式看成关于行向量的函数, 等价的有

(1) $\det_2(\lambda\mathbf{a}_1 + \mu\mathbf{a}_2, \mathbf{b}) = \lambda\det_2(\mathbf{a}_1, \mathbf{b}) + \mu\det_2(\mathbf{a}_2, \mathbf{b})$.

(2) $\det_2(\mathbf{a}, \lambda\mathbf{b}_1 + \mu\mathbf{b}_2) = \lambda\det_2(\mathbf{a}, \mathbf{b}_1) + \mu\det_2(\mathbf{a}, \mathbf{b}_2)$.

(3) $\det_2(\mathbf{a}, \mathbf{b}) = -\det_2(\mathbf{b}, \mathbf{a})$.

(4) $\det_2(\boldsymbol{\epsilon}_1, \boldsymbol{\epsilon}_2) = 1, \quad \boldsymbol{\epsilon}_1 = (1, 0), \boldsymbol{\epsilon}_2 = (0, 1)$.

其中 (1), (2) 称为线性, (3) 称为交错性, (4) 称为规范性.

引理 3.31. 设 $D\colon \mathbb{F}^2 \times \mathbb{F}^2 \to \mathbb{F}$ 满足线性, 则 D 是交错的 $\Longleftrightarrow$ 对 $\forall\, \mathbf{a} \in \mathbb{F}^2$, $D(\mathbf{a}, \mathbf{a}) = 0$.

证明. "$\Longrightarrow$": 由交错性, 对 $\forall\, \mathbf{a} \in \mathbb{F}^2$, $D(\mathbf{a}, \mathbf{a}) = -D(\mathbf{a}, \mathbf{a})$, 故 $D(\mathbf{a}, \mathbf{a}) = 0$.
"$\Longleftarrow$": 若对 $\forall\, \mathbf{a} \in \mathbb{F}^2$, $D(\mathbf{a}, \mathbf{a}) = 0$, 则特别的,

$$0 = D(\mathbf{a}+\mathbf{b}, \mathbf{a}+\mathbf{b}) = D(\mathbf{a}, \mathbf{a}) + D(\mathbf{b}, \mathbf{b}) + D(\mathbf{a}, \mathbf{b}) + D(\mathbf{b}, \mathbf{a}) = D(\mathbf{a}, \mathbf{b}) + D(\mathbf{b}, \mathbf{a}).$$

从而 $D(\mathbf{a}, \mathbf{b}) = -D(\mathbf{b}, \mathbf{a})$ 对 $\forall\, \mathbf{a}, \mathbf{b} \in \mathbb{F}^2$ 成立. □

命题 3.32. 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, 则对任意 $\mathbf{a}_1 = (a_{11}, a_{12})$, $\mathbf{a}_2 = (a_{21}, a_{22})$, $D(\mathbf{a}_1, \mathbf{a}_2) = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}$.

证明. 设 $\boldsymbol{\epsilon}_1, \boldsymbol{\epsilon}_2$ 是 $\mathbb{F}^2$ 的标准基, 则

$$\begin{aligned} D(\mathbf{a}_1, \mathbf{a}_2) &= D(a_{11}\boldsymbol{\epsilon}_1 + a_{12}\boldsymbol{\epsilon}_2, a_{21}\boldsymbol{\epsilon}_1 + a_{22}\boldsymbol{\epsilon}_2) \\ &= a_{11}a_{22}D(\boldsymbol{\epsilon}_1, \boldsymbol{\epsilon}_2) + a_{12}a_{21}D(\boldsymbol{\epsilon}_2, \boldsymbol{\epsilon}_1) \\ &= \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}. \end{aligned}$$ □

对 $\mathbb{F}^n$ 上的 n 元函数 $D_n\colon \overbrace{\mathbb{F}^n \times \cdots \times \mathbb{F}^n}^{n} \to \mathbb{F}$, 设 $\boldsymbol{\epsilon}_1, \ldots, \boldsymbol{\epsilon}_n$ 是 $\mathbb{F}^n$ 的标准基, 仍然可以定义线性, 交错性及规范性:

1. **线性:** $D_n(\ldots, \lambda\mathbf{a}_i + \mu\mathbf{a}'_i, \ldots) = \lambda D_n(\ldots, \mathbf{a}_i, \ldots) + \mu D_n(\ldots, \mathbf{a}'_i, \ldots), \quad \forall\, 1 \le i \le n.$

2. **交错性:** $D_n(\ldots, \mathbf{a}_i, \ldots, \mathbf{a}_j, \ldots) = -D_n(\ldots, \mathbf{a}_j, \ldots, \mathbf{a}_i, \ldots), \quad \forall\, 1 \le i < j \le n.$

3. **规范性:** $D_n(\boldsymbol{\epsilon}_1, \ldots, \boldsymbol{\epsilon}_n) = 1.$

定义 3.33. 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 $|\mathbf{A}|$.

练习 3.34. 设 $D \colon \mathbb{F}^3 \times \mathbb{F}^3 \times \mathbb{F}^3 \to \mathbb{F}$ 满足线性, 交错性及规范性, 则对 $\forall\, \mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3 \in \mathbb{F}^3$,

$$D(\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3) = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}.$$

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

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

所有 n 元置换构成的集合记为 $\mathfrak{S}_n$.

一个 n 元置换通常记为

$$\sigma = \begin{pmatrix} 1 & 2 & \cdots & n \\ \sigma(1) & \sigma(2) & \cdots & \sigma(n) \end{pmatrix}, \tag{3.38}$$

也可以简记为 $\sigma = (k_1 k_2 \cdots k_n)$, 其中 $k_i = \sigma(i)$ $(1 \le i \le n)$. 从而 n 元置换也可以看作 $\{1, 2, \ldots, n\}$ 的一个排列. 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 s_{ij}. 因为 $k_1, \ldots, k_n$ 是 $1, \ldots, n$ 的一个排列, 可以通过一系列对换将 σ 变为单位置换 $(12 \cdots n)$:

定理 3.36 (冒泡排序)**.** 按照如下排序算法可将 $\sigma = (k_1 k_2 \cdots k_n)$ 重新排列为单位置换:

1. 比较 k_1, k_2, 若 $k_1 > k_2$, 则做对换 s_{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 s_{23}, 否则不做操作. 依此类推, 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 $(12 \cdots n)$.

例题 3.37. 用冒泡排序算法将 $\sigma = (4231)$ 重排为 (1234).

证明. $(4231) \to (2431) \to (2341) \to (2314) \to (2134) \to (1234)$. □

定义 3.38. 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 # 表示集合中元素的个数. $(-1)^{\mathrm{inv}(\sigma)}$ 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.

例题 3.39. 设 $\sigma = (4231)$, 计算 σ 的逆序数, 并确定奇偶性.

证明. $\mathrm{inv}(\sigma) = \#\{(4,2),(4,3),(4,1),(2,1),(3,1)\} = 5$, $\mathrm{sign}(\sigma) = (-1)^5 = 1$, σ 是奇置换. □

引理 3.40. 利用冒泡排序将 $\sigma = (k_1 k_2 \cdots k_n)$ 重排为单位置换 $(12 \cdots n)$ 所需的相邻对换的操作次数为 $\mathrm{inv}(\sigma)$.

证明. 设 $\sigma = (k_1 \cdots k_i k_{i+1} \cdots k_n)$ 且 $k_i > k_{i+1}$. 对 σ 做相邻对换 s_{ii+1} 操作, 得到的置换为 $\sigma' = (k_1 \cdots k_{i+1} k_i \cdots k_n)$. 比较 σ 和 σ' 的逆序数, 可知 σ' 比 σ 恰好少一个逆序对 (k_i, k_{i+1}). 故 $\mathrm{inv}(\sigma') = \mathrm{inv}(\sigma) - 1$. 即在冒泡排序算法中, 每做一个相邻对换, 逆序数减 1. 因为单位置换的逆序数为 0, 故所做相邻对换的操作次数为 $\mathrm{inv}(\sigma)$. □

命题 3.41. 设 D 是满足交错性的 n 元函数, $\sigma = (k_1 k_2 \cdots k_n) \in \mathfrak{S}_n$, 则

$$D(x_{k_1}, x_{k_2}, \ldots, x_{k_n}) = \mathrm{sign}(\sigma) D(x_1, x_2, \ldots, x_n).$$

证明. 由引理 3.40, 利用冒泡排序将 $x_{k_1}, x_{k_2}, \ldots, x_{k_n}$ 重排为 $x_1, x_2, \ldots, x_n$, 对换操作次数为 $\mathrm{inv}(\sigma)$. 由交错性, 每次对换两变元函数改变符号, 故

$$D(x_{k_1}, x_{k_2}, \ldots, x_{k_n}) = (-1)^{\mathrm{inv}(\sigma)} D(x_1, x_2, \ldots, x_n) = \mathrm{sign}(\sigma) D(x_1, x_2, \ldots, x_n). \qquad \square$$

定理 3.42. 行列式函数 det 存在且唯一. 设 $\mathbf{A} = (a_{ij})_{n\times n} \in \mathsf{M}_n(\mathbb{F})$, 则

$$\det \mathbf{A} = \sum_{\sigma \in \mathfrak{S}_n} \mathrm{sign}(\sigma) a_{1\sigma(1)} a_{2\sigma(2)} \cdots a_{n\sigma(n)}. \tag{3.39}$$

证明. 设 $\mathbf{A} = (a_{ij})_{n\times n} \in \mathsf{M}_n(\mathbb{F})$, $\mathbf{a}_1, \ldots, \mathbf{a}_n$ 是 $\mathbf{A}$ 的行向量, 即 $\mathbf{a}_i = \sum_{j=1}^n a_{ij} \boldsymbol{\epsilon}_j$. 假设行列式函数 det 存在, 则由线性, 交错性及规范性, 有

$$\begin{aligned}\det \mathbf{A} &= \det(\mathbf{a}_1, \ldots, \mathbf{a}_n) \\ &= \det\Big(\sum_{k_1=1}^n a_{1k_1} \boldsymbol{\epsilon}_{k_1}, \ldots, \sum_{k_n=1}^n a_{nk_n} \boldsymbol{\epsilon}_{k_n}\Big)\end{aligned}$$

$$
\begin{aligned}
(\text{线性}) &= \sum_{1\le k_1,\dots,k_n\le n} a_{1k_1}\cdots a_{nk_n}\det(\boldsymbol{\epsilon}_{k_1},\dots,\boldsymbol{\epsilon}_{k_n})\\
(\text{交错性}) &= \sum_{(k_1\cdots k_n)\in\mathfrak{S}_n} a_{1k_1}\cdots a_{nk_n}\det(\boldsymbol{\epsilon}_{k_1},\dots,\boldsymbol{\epsilon}_{k_n})\\
(\text{命题 } 3.41\ ) &= \sum_{(k_1\cdots k_n)\in\mathfrak{S}_n} \operatorname{sign}(k_1\cdots k_n)a_{1k_1}\cdots a_{nk_n}\det(\boldsymbol{\epsilon}_1,\dots,\boldsymbol{\epsilon}_n)\\
(\text{规范性}) &= \sum_{\sigma\in\mathfrak{S}_n}\operatorname{sign}(\sigma)a_{1\sigma(1)}a_{2\sigma(2)}\cdots a_{n\sigma(n)}.
\end{aligned}
$$

这说明了若行列式函数存在, 则唯一.

另一方面, 按照公式 (3.39) 定义的函数 det: $\mathsf{M}_n(\mathbb{F})\to\mathbb{F}$ 显然满足线性, 交错性及规范性, 故公式 (3.39) 给出了行列式函数的存在性. □

例题 3.43. 按照公式 (3.39) 计算如下行列式:

$$
(1)\begin{vmatrix} & & & a_1\\ & & a_2 & \\ & \iddots & & \\ a_n & & & \end{vmatrix}.\qquad (2)\begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n}\\ 0 & a_{22} & \cdots & a_{2n}\\ \vdots & \vdots & \ddots & \vdots\\ 0 & \cdots & 0 & a_{nn}\end{vmatrix}.
$$

证明. (1) 在 (3.39) 只有 $\sigma=(nn-1\cdots1)$ 对应的项不为 0, 而

$$
\operatorname{inv}((nn-1\cdots1))=(n-1)+(n-2)+\cdots+1=n(n-1)/2,
$$

所以

$$
\det\mathbf{A}=\operatorname{sign}(nn-1\cdots1)a_1a_2\cdots a_n=(-1)^{n(n-1)/2}a_1a_2\cdots a_n.
$$

(2) 在 (3.39) 只有 $\sigma=(12\cdots n)$ 对应的项不为 0 故

$$
\det\mathbf{A}=a_{11}a_{22}\cdots a_{nn}.
$$

□

定义 3.44. 方阵 $\mathbf{A}\in\mathsf{M}_n(\mathbb{F})$ 称为是上三角矩阵, 如果它的主对角线下方的元素都为 0:

$$
\mathbf{A}=\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n}\\ & a_{22} & \cdots & a_{2n}\\ & & \ddots & \vdots\\ & & & a_{nn}\end{pmatrix}.
$$

由例题 3.43, 每个上三角矩阵的行列式是可以直接计算的. 另一方面, 每个方阵可以通过一系列初等行变换化为上三角矩阵, 说清楚初等行变换对行列式的影响, 就可以通过初等行变换计算行列式.

命题 3.45. 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 $\det \mathbf{PA} = \det \mathbf{A}$.

证明. 设 $\mathbf{A} = \begin{bmatrix} \mathbf{a}_1 & \cdots & \mathbf{a}_n \end{bmatrix}$ 是 $\mathbf{A}$ 的列向量形式, 则

$$\det \mathbf{P}_{ij}\mathbf{A} = \det(\mathbf{a}_1, \ldots, \mathbf{a}_j, \ldots, \mathbf{a}_i, \ldots, \mathbf{a}_n) = -\det(\mathbf{a}_1, \ldots, \mathbf{a}_i, \ldots, \mathbf{a}_j, \ldots, \mathbf{a}_n) = -\det \mathbf{A}.$$

$$\det \mathbf{D}_i(\lambda)\mathbf{A} = \det(\mathbf{a}_1, \ldots, \lambda\mathbf{a}_i, \ldots, \mathbf{a}_n) = \lambda\det(\mathbf{a}_1, \ldots, \mathbf{a}_i, \ldots, \mathbf{a}_n) = \lambda\det \mathbf{A}.$$

$$\det \mathbf{T}_{ij}(\lambda)\mathbf{A} = \det(\mathbf{a}_1, \ldots, \mathbf{a}_i + \lambda\mathbf{a}_j, \ldots, \mathbf{a}_j, \ldots, \mathbf{a}_n)$$

$$= \det(\mathbf{a}_1, \ldots, \mathbf{a}_i, \ldots, \mathbf{a}_j, \ldots, \mathbf{a}_n) + \lambda\det(\mathbf{a}_1, \ldots, \mathbf{a}_j, \ldots, \mathbf{a}_j, \ldots, \mathbf{a}_n) = \det \mathbf{A}.$$

□

例题 3.46. 计算如下行列式: (1) $\begin{vmatrix} 1 & 1 & 2 \\ 2 & 3 & 4 \\ 3 & 4 & 7 \end{vmatrix}$. (2) $\begin{vmatrix} 0 & 0 & 1 \\ 0 & 2 & 3 \\ 3 & 4 & 5 \end{vmatrix}$.

解. (1) $\begin{vmatrix} 1 & 1 & 2 \\ 2 & 3 & 4 \\ 3 & 4 & 7 \end{vmatrix} = \begin{vmatrix} 1 & 1 & 2 \\ 0 & 1 & 0 \\ 0 & 1 & 1 \end{vmatrix} = \begin{vmatrix} 1 & 1 & 2 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{vmatrix} = 1.$

(2) $\begin{vmatrix} 0 & 0 & 1 \\ 0 & 2 & 3 \\ 3 & 4 & 5 \end{vmatrix} = -\begin{vmatrix} 3 & 4 & 5 \\ 0 & 2 & 3 \\ 0 & 0 & 1 \end{vmatrix} = -6.$ □

例题 3.47. 计算行列式 $\begin{vmatrix} x & & & a_n \\ -1 & \ddots & & \vdots \\ & \ddots & x & a_2 \\ & & -1 & x + a_1 \end{vmatrix}$.

解. 依次将行列式的第 $n, n-1, \ldots, 2$ 行乘以 x 加到第 $n-1, n-2, \ldots, 1$ 行, 得到

$$\begin{vmatrix} x & & & a_n \\ -1 & \ddots & & \vdots \\ & \ddots & x & a_2 \\ & & -1 & x+a_1 \end{vmatrix} = \begin{vmatrix} 0 & & & x^n + a_1 x^{n-1} + \cdots + a_n \\ -1 & \ddots & & \vdots \\ & \ddots & 0 & x^2 + a_1 x + a_2 \\ & & -1 & x + a_1 \end{vmatrix}$$

$$= x^n + a_1 x^{n-1} + \cdots + a_n. \qquad \square$$

推论 3.48. 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

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

证明. 结合命题 3.45 以及 $\det \mathbf{P}_{ij} = -1$, $\det \mathbf{D}_i(\lambda) = \lambda\ (\lambda \neq 0)$, $\det \mathbf{T}_{ij}(\lambda) = 1$ 即证. $\square$

定理 3.49. 设 $\mathbf{A} \in \mathsf{M}_n(\mathbb{F})$, 则 $\mathbf{A}$ 可逆 $\iff \det \mathbf{A} \neq 0$.

证明. 对 $\mathbf{A} \in \mathsf{M}_n(\mathbb{F})$, 可以用初等行变换将其化为阶梯形, 因为阶梯形是上三角矩阵, 从而存在初等矩阵 $\mathbf{P}_1, \ldots, \mathbf{P}_k$ 使得

$$\mathbf{P}_1 \cdots \mathbf{P}_k \mathbf{A} = \mathbf{A}' = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ 0 & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & a_{nn} \end{pmatrix}.$$

从而 $\mathbf{A}$ 可逆 $\iff \operatorname{rank} \mathbf{A} = n \iff \operatorname{rank} \mathbf{A}' = n \iff a_{11}, \ldots, a_{nn} \neq 0 \iff \det \mathbf{A}' \neq 0 \iff \det \mathbf{P}_1 \cdots \mathbf{P}_k \mathbf{A} \neq 0 \iff \det \mathbf{P}_1 \cdots \det \mathbf{P}_k \det \mathbf{A} \neq 0 \iff \det \mathbf{A} \neq 0$. $\square$

例题 3.50. 设 $\mathbf{A} \in \mathsf{M}_n(\mathbb{R})$, 是对角占优矩阵, 证明 $\det \mathbf{A} > 0$.

证明. 在例题 3.21 中, 我们已经证明对角占优矩阵都可逆, 即 $\det \mathbf{A} \neq 0$. 设 $\mathbf{A}(\lambda) = \mathbf{A} + \lambda \mathbf{I}_n$, 且 $f(\lambda) = \det \mathbf{A}(\lambda)$ 是关于 λ 的首一 n 次多项式. 对 $\forall\, \lambda \geq 0$, $\mathbf{A}(\lambda)$ 仍是对角占优矩阵, 故 $f(\lambda)$ 在 $[0, +\infty]$ 中恒不为零. 若 $\det \mathbf{A} < 0$, 即 $f(0) < 0$, 则因为对充分大的 λ, $f(\lambda) > 0$, 由介值定理可知存在 $\mu \in [0, +\infty]$ 使得 $f(\mu) = 0$, 矛盾. 故 $\det \mathbf{A} > 0$. $\square$

定理 3.51. 设 $\mathbf{A},\mathbf{B}\in \mathsf{M}_n$, 则 $\det \mathbf{AB}=\det \mathbf{A}\det \mathbf{B}$.

证明一. 若 $\operatorname{rank}\mathbf{A}<n$, 则 $\det \mathbf{A}=0$. 又 $\operatorname{rank}\mathbf{AB}\le \operatorname{rank}\mathbf{A}<n$, 所以 $\det \mathbf{AB}=0$, 此时等式两边都为 0, 定理成立.

设 $\operatorname{rank}\mathbf{A}=n$, 则 $\mathbf{A}$ 可逆, 从而 $\mathbf{A}=\mathbf{P}_1\cdots\mathbf{P}_k$, 是一些初等矩阵的乘积. 于是由推论 3.48 知 $\det \mathbf{AB}=\det \mathbf{P}_1\mathbf{P}_2\cdots\mathbf{P}_k\mathbf{B}=\det \mathbf{P}_1\cdots\det \mathbf{P}_k\det \mathbf{B}=\det \mathbf{P}_1\cdots\mathbf{P}_k\det \mathbf{B}=\det \mathbf{A}\det \mathbf{B}$. □

证明二. 设 $\mathbf{A}=(a_{ij})_{n\times n}$, $\mathbf{B}=\begin{bmatrix}\mathbf{b}_1 & \cdots & \mathbf{b}_n\end{bmatrix}$ 写为列向量形式, 由 (3.24),

$$\mathbf{AB}=\begin{bmatrix}a_{11}\mathbf{b}_1+\cdots a_{1n}\mathbf{b}_n & \cdots & a_{n1}\mathbf{b}_1+\cdots+a_{nn}\mathbf{b}_n\end{bmatrix}.$$

于是

$$\begin{aligned}\det \mathbf{AB}&=\det(a_{11}\mathbf{b}_1+\cdots a_{1n}\mathbf{b}_n,\ldots,a_{n1}\mathbf{b}_1+\cdots+a_{nn}\mathbf{b}_n)\\&=\sum_{\sigma\in\mathfrak{S}_n}a_{1\sigma(1)}\cdots a_{n\sigma(n)}\det(\mathbf{b}_{\sigma(1)},\ldots,\mathbf{b}_{\sigma(n)})\\&=\sum_{\sigma\in\mathfrak{S}_n}\operatorname{sign}(\sigma)a_{1\sigma(1)}\cdots a_{n\sigma(n)}\det(\mathbf{b}_1,\ldots,\mathbf{b}_n)\\&=\det \mathbf{A}\det \mathbf{B}.\end{aligned}$$ □

命题 3.52. 设 $\mathbf{A}_1\in\mathsf{M}_{n_1}(\mathbb{F}),\mathbf{A}_2\in\mathsf{M}_{n_2}(\mathbb{F}),\mathbf{A}_3\in\mathsf{M}_{n_1\times n_2}(\mathbb{F})$, $n_1+n_2=n$, 则

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

证明. 设 $\mathbf{B}=(b_{ij})_{n\times n}=\begin{pmatrix}\mathbf{A}_1 & \mathbf{A}_3\\ 0 & \mathbf{A}_2\end{pmatrix}$. 在 $\det \mathbf{B}=\sum_{\sigma\in\mathfrak{S}_n}\operatorname{sign}(\sigma)b_{1\sigma(1)}b_{2\sigma(2)}\cdots b_{n\sigma(n)}$ 中, 对 $\forall\ n_1+1\le k\le n$, 若 $\sigma(k)\le n_1$, 则 $b_{k\sigma(k)}=0$. 从而在行列式公式中, 只需针对那些满足

$$\sigma\in\mathfrak{S}_n,\quad \sigma\{n_1+1,\ldots,n\}=\{n_1+1,\ldots,n\} \tag{3.40}$$

的 σ 计算. 设 σ 满足 (3.40), 则限制在 $\{1,\ldots,n_1\}$ 上唯一对应于置换 $\sigma_1\in\mathfrak{S}_{n_1}$, 限制在 $\{n_1+1,\ldots,n\}$ 上唯一对应于置换 $\sigma_2\in\mathfrak{S}_{n_2}$, 且 $\operatorname{sign}(\sigma)=\operatorname{sign}(\sigma_1)\operatorname{sign}(\sigma_2)$. 故

$$\det \mathbf{B}=\sum_{\sigma\in\mathfrak{S}_n}\operatorname{sign}(\sigma)b_{1\sigma(1)}\cdots b_{n\sigma(n)}$$

$$= (\sum_{\sigma_1 \in \mathfrak{S}_{n_1}} \operatorname{sign}(\sigma_1) b_{1\sigma_1(1)} \cdots b_{n_1\sigma_1(n_1)})(\sum_{\sigma_2 \in \mathfrak{S}_{n_2}} \operatorname{sign}(\sigma_2) b_{n_1+1\sigma_2(n_1+1)} \cdots b_{n\sigma_2(n)})$$
$$= \det \mathbf{A}_1 \det \mathbf{A}_2. \qquad \square$$

推论 3.53. 设 $\mathbf{A}_1 \in \mathsf{M}_{n_1}(\mathbb{F}), \mathbf{A}_2 \in \mathsf{M}_{n_2}(\mathbb{F}), \mathbf{A}_3 \in \mathsf{M}_{n_1 \times n_2}(\mathbb{F})$, $n_1 + n_2 = n$, 则

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

证明. 设 $\mathbf{P} = \begin{pmatrix} \mathbf{0} & \mathbf{I}_{n_1} \\ \mathbf{I}_{n_2} & \mathbf{0} \end{pmatrix}$, $\sigma = (n_2 + 1 \cdots n 1 \cdots n_2)$, 则 $\det \mathbf{P} = \operatorname{sign}(\sigma) = (-1)^{n_1 n_2}$. 又

$$\begin{pmatrix} \mathbf{A}_1 & \mathbf{A}_3 \\ \mathbf{0} & \mathbf{A}_2 \end{pmatrix} \begin{pmatrix} \mathbf{0} & \mathbf{I}_{n_1} \\ \mathbf{I}_{n_2} & \mathbf{0} \end{pmatrix} = \begin{pmatrix} \mathbf{A}_3 & \mathbf{A}_1 \\ \mathbf{A}_2 & \mathbf{0} \end{pmatrix},$$

于是

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

引理 3.54. 设 $\sigma \in \mathfrak{S}_n$, σ^{-1} 是 σ 的逆, 则 $\operatorname{inv}(\sigma) = \operatorname{inv}(\sigma^{-1})$. 特别的, $\operatorname{sign}(\sigma) = \operatorname{sign}(\sigma^{-1})$.

证明. 设 $\sigma = (k_1 k_2 \cdots k_n)$, 则 $\sigma^{-1}(k_i) = i$, 从而

$$\begin{aligned} \operatorname{inv}(\sigma) &= \#\{(k_i, k_j) \mid i < j,\ k_i > k_j\} \\ &= \#\{(j, i) \mid j > i,\ k_j < k_i\} \\ &= \#\{(\sigma^{-1}(k_j), \sigma^{-1}(k_i)) \mid k_j < k_i, \sigma^{-1}(k_j) > \sigma^{-1}(k_i)\} \\ &= \operatorname{inv}(\sigma^{-1}). \end{aligned} \qquad \square$$

命题 3.55. 设 $\mathbf{A} \in \mathsf{M}_n(\mathbb{F})$, 则 $\det \mathbf{A} = \det \mathbf{A}^{\mathrm{T}}$.

证明. 设 $\mathbf{A} = (a_{ij})_{n \times n}$, $\mathbf{A}^{\mathrm{T}} = (b_{ij})_{n \times n}$, 则 $a_{ij} = b_{ji}$. 从而

$$\begin{aligned} \det \mathbf{A} &= \sum_{\sigma \in \mathfrak{S}_n} \operatorname{sign}(\sigma) a_{1\sigma(1)} a_{2\sigma(2)} \cdots a_{n\sigma(n)} \\ &= \sum_{\sigma \in \mathfrak{S}_n} \operatorname{sign}(\sigma) a_{\sigma^{-1}(1)1} a_{\sigma^{-1}(2)2} \cdots a_{\sigma^{-1}(n)n} \\ &= \sum_{\sigma \in \mathfrak{S}_n} \operatorname{sign}(\sigma^{-1}) b_{1\sigma^{-1}(1)} b_{2\sigma^{-1}(2)} \cdots b_{n\sigma^{-1}(n)} \end{aligned}$$

$$= \sum_{\sigma \in \mathfrak{S}_n} \operatorname{sign}(\sigma) b_{1\sigma(1)} b_{2\sigma(2)} \cdots b_{n\sigma(n)} = \det \mathbf{A}^{\mathrm{T}}. \qquad \square$$

评注 3.56. 命题 3.55 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. 从而也可以通过列变换来简化计算.

例题 3.57. 设 $\mathbf{A} \in \mathsf{M}_n(\mathbb{F})$, 且 $\mathbf{A}^{\mathrm{T}}\mathbf{A} = \mathbf{I}_n$, $|\mathbf{A}| < 0$. 证明 $|\mathbf{A} + \mathbf{I}_n| = 0$.

证明. 设 $|\mathbf{A} + \mathbf{I}_n| = \left|\mathbf{A}^{\mathrm{T}} + \mathbf{I}_n\right| = a$. 则

$$a \cdot |\mathbf{A}| = \left|\mathbf{A}^{\mathrm{T}} + \mathbf{I}_n\right| |\mathbf{A}| = \left|\mathbf{A}^{\mathrm{T}}\mathbf{A} + \mathbf{A}\right| = |\mathbf{I}_n + \mathbf{A}| = a.$$

若 $a \neq 0$, 则 $|\mathbf{A}| = 1$, 矛盾. 故 $|\mathbf{A} + \mathbf{I}_n| = 0$. $\square$

例题 3.58. 计算行列式: $\begin{vmatrix} c_1 & a_2 & a_3 & \cdots & a_n \\ a_1 & c_2 & a_3 & \cdots & a_n \\ a_1 & a_2 & c_3 & \cdots & a_n \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_1 & a_2 & a_3 & \cdots & c_n \end{vmatrix}$.

证明. 利用 "加边" 的技巧, 可做如下行变换将行列式化为简单形式:

$$\begin{vmatrix} c_1 & a_2 & a_3 & \cdots & a_n \\ a_1 & c_2 & a_3 & \cdots & a_n \\ a_1 & a_2 & c_3 & \cdots & a_n \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_1 & a_2 & a_3 & \cdots & c_n \end{vmatrix} = \begin{vmatrix} 1 & -a_1 & -a_2 & -a_3 & \cdots & -a_n \\ 0 & c_1 & a_2 & a_3 & \cdots & a_n \\ 0 & a_1 & c_2 & a_3 & \cdots & a_n \\ 0 & a_1 & a_2 & c_3 & \cdots & a_n \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & a_1 & a_2 & a_3 & \cdots & c_n \end{vmatrix}$$

$$= \begin{vmatrix} 1 & -a_1 & -a_2 & -a_3 & \cdots & -a_n \\ 0 & c_1 & a_2 & a_3 & \cdots & a_n \\ 0 & a_1 & c_2 & a_3 & \cdots & a_n \\ 0 & a_1 & a_2 & c_3 & \cdots & a_n \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & a_1 & a_2 & a_3 & \cdots & c_n \end{vmatrix}$$

$$
= \begin{vmatrix}
1 & -a_1 & -a_2 & -a_3 & \cdots & -a_n \\
1 & c_1 - a_1 & 0 & 0 & \cdots & 0 \\
1 & 0 & c_2 - a_2 & 0 & \cdots & 0 \\
1 & 0 & 0 & c_3 - a_3 & \cdots & 0 \\
\vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\
1 & 0 & 0 & 0 & \cdots & c_n - a_n
\end{vmatrix}
$$

$$
= \prod_{i=1}^{n}(c_i - a_i) + \sum_{i=1}^{n} a_i \prod_{j \neq i}(c_j - a_j). \qquad \square
$$

例题 3.59. 计算范德蒙德行列式: $\begin{vmatrix} 1 & a_1 & \cdots & a_1^{n-1} \\ 1 & a_2 & \cdots & a_2^{n-1} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & a_n & \cdots & a_n^{n-1} \end{vmatrix}$.

证明. 分别将第 $n-1, n-2, \ldots, 1$ 列乘以 $-a_1$ 加到第 $n, n-1, \ldots, 2$ 列, 得到

$$
\begin{vmatrix} 1 & a_1 & \cdots & a_1^{n-1} \\ 1 & a_2 & \cdots & a_2^{n-1} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & a_n & \cdots & a_n^{n-1} \end{vmatrix}
= \begin{vmatrix} 1 & 0 & \cdots & 0 \\ 1 & a_2 - a_1 & \cdots & a_2^{n-2}(a_2 - a_1) \\ \vdots & \vdots & \ddots & \vdots \\ 1 & a_n - a_1 & \cdots & a_n^{n-2}(a_n - a_1) \end{vmatrix}
$$

$$
= (a_2 - a_1) \cdots (a_n - a_1) \begin{vmatrix} 1 & a_2 & \cdots & a_2^{n-2} \\ 1 & a_3 & \cdots & a_3^{n-2} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & a_n & \cdots & a_n^{n-2} \end{vmatrix}.
$$

由归纳法即知

$$
\begin{vmatrix} 1 & a_1 & \cdots & a_1^{n-1} \\ 1 & a_2 & \cdots & a_2^{n-1} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & a_n & \cdots & a_n^{n-1} \end{vmatrix} = \prod_{1 \leq i < j \leq n} (a_j - a_i). \qquad \square
$$

命题 3.60. 设 $\mathbf{A} \in \mathsf{M}_n(\mathbb{F})$, 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

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

证明. 不妨设 $\mathbf{A}$ 的第 j 列除了 a_{ij} 均为 0, 第 i 行除了 a_{ij} 均为 0 的情形的证明是类似的. 依次对 $\mathbf{A}$ 做如下行变换: 交换第 i 和第 $i-1$ 行, 交换第 $i-1$ 和第 $i-2$ 行, $\dots$, 交换第 2 行和第 1 行, 得到矩阵 $\mathbf{A}'$. 再依次对 $\mathbf{A}'$ 做如下列变换: 交换第 j 和第 $j-1$ 列, 交换第 $j-1$ 和第 $j-2$ 列, $\dots$, 交换第 2 列和第 1 列, 得到矩阵 $\mathbf{A}''$.

因为交换两行或交换两列改变行列式符号, 而从 $\mathbf{A}$ 到 $\mathbf{A}''$ 总共经历 $i+j-2$ 次交换两行或两列的操作, 故 $\det \mathbf{A} = (-1)^{i+j}\det \mathbf{A}''$. 又因为

$$\mathbf{A}'' = \begin{pmatrix} a_{ij} & * \\ 0 & \mathbf{M}_{ij} \end{pmatrix},$$

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

定义 3.61. 设 $\mathbf{A} \in \mathsf{M}_n(\mathbb{F})$, $\mathbf{M}_{ij}$ 是去掉 $\mathbf{A}$ 的第 i 行和第 j 列得到的 $n-1$ 阶方阵, 则

$$A_{ij} = (-1)^{i+j}\det \mathbf{M}_{ij} \tag{3.41}$$

称为 $\mathbf{A}$ 关于 a_{ij} 的代数余子式.

定理 3.62. 设 $\mathbf{A} \in \mathsf{M}_n(\mathbb{F})$, 对任意 $1 \le i \le n$,

$$\det \mathbf{A} = a_{i1}A_{i1} + a_{i2}A_{i2} + \cdots + a_{in}A_{in}, \tag{3.42}$$

$$\det \mathbf{A} = a_{1i}A_{1i} + a_{2i}A_{2i} + \cdots + a_{ni}A_{ni}. \tag{3.43}$$

分别称为行列式关于第 i 行和第 i 列的展开.

证明. 由行列式的线性, 有

$$\begin{aligned}\det \mathbf{A} &= \begin{vmatrix} a_{11} & \cdots & a_{1i} & \cdots & a_{1n} \\ a_{21} & \cdots & a_{2i} & \cdots & a_{2n} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ a_{n-11} & \cdots & a_{n-1i} & \cdots & a_{n-1n} \\ a_{n1} & \cdots & a_{ni} & \cdots & a_{nn} \end{vmatrix} \\ &= \begin{vmatrix} a_{11} & \cdots & a_{1i} & \cdots & a_{1n} \\ a_{21} & \cdots & 0 & \cdots & a_{2n} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ a_{n-11} & \cdots & 0 & \cdots & a_{n-1n} \\ a_{n1} & \cdots & 0 & \cdots & a_{nn} \end{vmatrix} + \cdots + \begin{vmatrix} a_{11} & \cdots & 0 & \cdots & a_{1n} \\ a_{21} & \cdots & 0 & \cdots & a_{2n} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ a_{n-11} & \cdots & 0 & \cdots & a_{n-1n} \\ a_{n1} & \cdots & a_{ni} & \cdots & a_{nn} \end{vmatrix}\end{aligned}$$

$$(\text{命题 } 3.60) = a_{1i}A_{1i} + \cdots a_{ni}A_{ni}.$$

得到关于第 i 列的展开. 关于第 i 行的展开公式的证明是类似的. □

推论 3.63. 设 $\mathbf{A} \in \mathsf{M}_n(\mathbb{F})$, $1 \le i, j \le n$, 则

$$a_{i1}A_{j1} + a_{i2}A_{j2} + \cdots + a_{in}A_{jn} = \delta_{ij}\det \mathbf{A}, \tag{3.44}$$

$$a_{1i}A_{1j} + a_{2i}A_{2j} + \cdots + a_{ni}A_{nj} = \delta_{ij}\det \mathbf{A}. \tag{3.45}$$

证明. 若 $i = j$, 则 (3.44) 和 (3.45) 即为关于第 i 行, 第 i 列的展开公式. 若 $i \neq j$ (不妨设 $i < j$), 设 $\mathbf{A}$ 的行向量组为 $\mathbf{a}_1, \ldots, \mathbf{a}_n$. 考虑方阵 $\mathbf{B}_{ij}$, 它是将 $\mathbf{A}$ 的第 j 个列向量替换为第 i 个列向量的矩阵, 即

$$\mathbf{B}_{ij} = \begin{pmatrix} \mathbf{a}_1 & \cdots & \mathbf{a}_i & \cdots & \mathbf{a}_i & \cdots & \mathbf{a}_n \end{pmatrix}.$$

一方面, 因为 $\mathbf{B}_{ij}$ 的两个列向量相同, 由行列式的交错性有 $\det \mathbf{B}_{ij} = 0$. 另一方面将 $\det \mathbf{B}_{ij}$ 关于第 j 列展开即得到 (3.45). (3.44) 的证明是类似的. □

定义 3.64. 设 $\mathbf{A} \in \mathsf{M}_n(\mathbb{F})$, $\mathbf{A}$ 的伴随矩阵为

$$\mathbf{A}^* = (A_{ij})^{\mathrm{T}}_{n\times n} = \begin{pmatrix} A_{11} & A_{21} & \cdots & A_{n1} \\ A_{12} & A_{22} & \cdots & A_{n2} \\ \vdots & \vdots & \ddots & \vdots \\ A_{1n} & A_{2n} & \cdots & A_{nn} \end{pmatrix}. \tag{3.46}$$

由推论 3.63 可知:

推论 3.65. 设 $\mathbf{A} \in \mathsf{M}_n(\mathbb{F})$, $\mathbf{A}^*$ 是其伴随矩阵, 则 $\mathbf{A}\mathbf{A}^* = (\det \mathbf{A})\mathbf{I}_n$. 特别的, 若 $\mathbf{A}$ 可逆, 则

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

例题 3.66. 设 $\mathbf{A} = \begin{pmatrix} 2 & 2 & 3 \\ 1 & -1 & 0 \\ -1 & 2 & 1 \end{pmatrix}$, 求 $\mathbf{A}^*$ 以及 $\mathbf{A}^{-1}$.

证明. 依次计算各个余子式:

$$A_{11} = (-1)^{1+1}\begin{vmatrix} -1 & 0 \\ 2 & 1 \end{vmatrix} = -1,\ A_{12} = (-1)^{1+2}\begin{vmatrix} 1 & 0 \\ -1 & 1 \end{vmatrix} = -1,\ A_{13} = (-1)^{1+3}\begin{vmatrix} 1 & -1 \\ -1 & 2 \end{vmatrix} = 1,$$

$$A_{21}=(-1)^{2+1}\begin{vmatrix}2&3\\2&1\end{vmatrix}=4,\ A_{22}=(-1)^{2+2}\begin{vmatrix}2&3\\-1&1\end{vmatrix}=5,\ A_{23}=(-1)^{2+3}\begin{vmatrix}2&2\\-1&2\end{vmatrix}=-6,$$

$$A_{31}=(-1)^{3+1}\begin{vmatrix}2&3\\-1&0\end{vmatrix}=3,\ A_{32}=(-1)^{3+2}\begin{vmatrix}2&3\\1&0\end{vmatrix}=3,\ A_{33}=(-1)^{3+3}\begin{vmatrix}2&2\\1&-1\end{vmatrix}=-4.$$

从而 $\mathbf{A}^*=\begin{pmatrix}-1&4&3\\-1&5&3\\1&-6&-4\end{pmatrix}$. 因为

$$\det\ \mathbf{A}=\begin{vmatrix}2&2&3\\1&-1&0\\-1&2&1\end{vmatrix}=\begin{vmatrix}0&6&5\\0&1&1\\-1&2&1\end{vmatrix}=-1,$$

从而 $\mathbf{A}^{-1}=\dfrac{\mathbf{A}^*}{\det\ \mathbf{A}}=-\mathbf{A}^*=\begin{pmatrix}1&-4&-3\\1&-5&-3\\-1&6&4\end{pmatrix}$. □

定理 3.67 (克拉默法则)**.** 如果线性方程组 $\mathbf{Ax}=\mathbf{b}$ 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:

$$x_k=\frac{\det(\mathbf{a}_1,\ldots,\mathbf{b},\ldots,\mathbf{a}_n)}{\det(\mathbf{a}_1,\ldots,\mathbf{a}_k,\ldots,\mathbf{a}_n)},\quad k=1,2,\ldots,n \tag{3.47}$$

证明. 若 $\mathbf{A}$ 可逆, 我们已知对 $\forall\ \mathbf{b}\in\mathbb{F}^n$, $\mathbf{Ax}=\mathbf{b}$ 有唯一解, 且

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

就是它的解. 从而

$$x_k=\frac{b_1A_{1k}+\cdots+b_nA_{nk}}{\det\ \mathbf{A}}=\frac{\det(\mathbf{a}_1,\ldots,\mathbf{b},\ldots,\mathbf{a}_n)}{\det(\mathbf{a}_1,\ldots,\mathbf{a}_k,\ldots,\mathbf{a}_n)},\quad k=1,2,\ldots,n$$ □

例题 3.68. 利用克拉默法则解线性方程组 $\begin{cases}x_1+2x_2+3x_3=1,\\2x_1+2x_2+5x_3=2,\\3x_1+5x_2+x_3=3.\end{cases}$

解. 线性方程组可写为

$$\begin{pmatrix}1&2&3\\2&2&5\\3&5&1\end{pmatrix}\begin{pmatrix}x_1\\x_2\\x_3\end{pmatrix}=\begin{pmatrix}1\\2\\3\end{pmatrix}.$$

因系数矩阵的第一列 $\mathbf{a}_1$ 和常数项 $\mathbf{b}$ 相同, 由 (3.47) 知 $x_1=1$, $x_2=x_3=0$. □

下面介绍行列式与矩阵的秩的关系:

定义 3.69. 设 $\mathbf{A}\in\mathsf{M}_{n\times m}(\mathbb{F})$, $\mathbf{A}$ 的第 $i_1,\dots,i_k$ 行与 $j_1,\dots,j_k$ 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

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

定理 3.70. 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.

证明. "$\Longrightarrow$": 设 $\operatorname{rank}\mathbf{A}=r$, 则任取 $\mathbf{A}$ 的 $i_1,\dots,i_{r+1}$ 行, 这 $r+1$ 个行向量线性相关. 特别的, 任取 $j_1,\dots,j_{r+1}$ 列, 向量组

$$\tilde{\mathbf{a}}_1=(a_{i_1j_1},\dots,a_{i_1j_{r+1}}),\dots,\tilde{\mathbf{a}}_{r+1}=(a_{i_{r+1}j_1},\dots,a_{i_{r+1}j_{r+1}})$$

也线性相关. 故任意 $r+1$ 阶子式 $A\begin{pmatrix}i_1&\cdots&i_{r+1}\\j_1&\cdots&j_{r+1}\end{pmatrix}=0$.

下面证明存在 $\mathbf{A}$ 的一个 r 阶子式不为 0. 因 $\operatorname{rank}\mathbf{A}=r$, 设 $\mathbf{A}$ 的第 $i_1,\dots,i_r$ 行为极大线性无关组. 考察这 r 个行向量构成的矩阵

$$\mathbf{B}=\begin{pmatrix}a_{i_11}&\cdots&a_{i_1m}\\\vdots&\ddots&\vdots\\a_{i_r1}&\cdots&a_{i_rm}\end{pmatrix},$$

则 $\mathbf{B}$ 的列向量组的秩也为 r. 设 $\mathbf{B}$ 的第 $j_1,\dots,j_r$ 列为极大线性无关组, 得到

$$A\begin{pmatrix}i_1&\cdots&i_r\\j_1&\cdots&j_r\end{pmatrix}\neq0.$$

“$\Longleftarrow$”: 设子式 $A\begin{pmatrix} i_1 & \cdots & i_r \\ j_1 & \cdots & j_r \end{pmatrix} \neq 0$, 则 $\mathbf{A}$ 的第 $i_1, \dots, i_r$ 行线性无关, 得到 rank $\mathbf{A} \geq r$. 若 rank $\mathbf{A} \geq r+1$, 由 “$\Longrightarrow$” 的证明可知在 $\mathbf{A}$ 的 $r+1$ 个线性无关的行向量中可以找到一个非零的 $r+1$ 阶子式, 矛盾. 从而 rank $\mathbf{A} = r$. □

例题 3.71. 设 $\mathbf{A} \in \mathsf{M}_n(\mathbb{F})$, 则 $\text{rank } \mathbf{A}^* = \begin{cases} n, & \text{rank } \mathbf{A} = n. \\ 1, & \text{rank } \mathbf{A} = n-1. \\ 0, & \text{rank } \mathbf{A} < n-1. \end{cases}$

证明. (1) 若 rank $\mathbf{A} = n$, 则 $\mathbf{A}$ 可逆, det $\mathbf{A} \neq 0$. 由 $\mathbf{A}\mathbf{A}^* = (\det \mathbf{A})\mathbf{I}_n$ 可知 $\mathbf{A}^*$ 也可逆, 故 rank $\mathbf{A}^* = n$.
(2) 若 rank $\mathbf{A} = n-1$, 则 det $\mathbf{A} = 0$, $\mathbf{A}\mathbf{A}^* = (\det \mathbf{A})\mathbf{I}_n = \mathbf{0}$. 将 $\mathbf{A}^*$ 写为行向量形式

$$\mathbf{A}^* = \begin{pmatrix} \mathbf{b}_1 & \cdots & \mathbf{b}_n \end{pmatrix},$$

由 $\mathbf{A}\mathbf{A}^* = \mathbf{0}$ 可知 $\mathbf{A}\mathbf{b}_i = \mathbf{0}$ 对 $\forall\, 1 \leq i \leq n$ 成立, 即 $\mathbf{b}_1, \dots, \mathbf{b}_n$ 都属于线性方程组 $\mathbf{A}\mathbf{x} = \mathbf{0}$ 的解空间 Ω_0. 由定理 2.38 知 $\dim \Omega_0 = n - \text{rank } \mathbf{A} = 1$, 于是

$$\text{rank } \mathbf{A}^* = \text{rank}\{\mathbf{b}_1, \dots, \mathbf{b}_n\} \leq \dim \Omega_0 = 1.$$

另一方面, 由定理 3.70, $\mathbf{A}$ 的某个 $n-1$ 阶子式非零, 即某个代数余子式 $A_{ij} \neq 0$, 于是 rank $\mathbf{A}^* = 1$.
(3) 若 rank $\mathbf{A} \leq n-2$, 则 $\mathbf{A}$ 的所有 $n-1$ 阶子式为零, 即所有代数余子式 $A_{ij} = 0$, 于是 $\mathbf{A}^* = \mathbf{0}$, rank $\mathbf{A}^* = 0$. □

例题 3.72. 设 $\mathbf{A}, \mathbf{B} \in \mathsf{M}_n(\mathbb{F})$, $\lambda \in \mathbb{F}$, 则伴随矩阵满足如下性质:
(1) $(\lambda\mathbf{A})^* = \lambda^{n-1}\mathbf{A}^*$.
(2) $|\mathbf{A}^*| = |\mathbf{A}|^{n-1}$.
(3) $(\mathbf{A}^*)^* = |\mathbf{A}|^{n-2}\mathbf{A},\ n \geq 2$.
(4) $(\mathbf{A}\mathbf{B})^* = \mathbf{B}^*\mathbf{A}^*$.

证明. (1) 设 B_{ij} 是 $\lambda\mathbf{A}$ 关于 (i,j) 位置的代数余子式. 因为 $\lambda\mathbf{A}$ 中去掉第 i 行, 第 j 列后的 $n-1$ 阶矩阵是 $\lambda\mathbf{M}_{ij}$, 故 $B_{ij} = (-1)^{i+j}\det(\lambda\mathbf{M}_{ij}) = \lambda^{n-1}A_{ij}$, 于是 $(\lambda\mathbf{A})^* = \lambda^{n-1}\mathbf{A}^*$.
(2) 由例题 3.71, 若 $\mathbf{A}$ 不可逆, 则 $|\mathbf{A}| = |\mathbf{A}^*| = 0$, 等式成立. 若 $\mathbf{A}$ 可逆, 则

$$\mathbf{A}\mathbf{A}^* = |\mathbf{A}|\mathbf{I}_n \Longrightarrow |\mathbf{A}||\mathbf{A}^*| = |\mathbf{A}|^n \Longrightarrow |\mathbf{A}^*| = |\mathbf{A}|^{n-1}.$$

(3) 若 $n=2$, 设 $\mathbf{A}=\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$, 则 $\mathbf{A}^*=\begin{pmatrix} a_{22} & -a_{12} \\ -a_{21} & a_{11} \end{pmatrix}$, $(\mathbf{A}^*)^*=\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}=\mathbf{A}$, 等式成立.

下设 $n>2$. 若 $\mathbf{A}$ 不可逆, 由例题 3.71, $\operatorname{rank}\mathbf{A}^*\le 1<n-1$, 故 $(\mathbf{A}^*)^*=\mathbf{0}$, 等式成立. 若 $\mathbf{A}$ 可逆, 则 $\mathbf{A}^*$ 也可逆, 从而

$$(\mathbf{A}^*)^*=|\mathbf{A}^*|(\mathbf{A}^*)^{-1}=|\mathbf{A}|^{n-1}(|\mathbf{A}|\mathbf{A}^{-1})^{-1}=|\mathbf{A}|^{n-2}\mathbf{A}.$$

(4) 若 $\mathbf{A}$, $\mathbf{B}$ 可逆, 则 $\mathbf{AB}$ 可逆, 于是

$$(\mathbf{AB})^*=|\mathbf{AB}|(\mathbf{AB})^{-1}=|\mathbf{B}|\mathbf{B}^{-1}|\mathbf{A}|\mathbf{A}^{-1}=\mathbf{B}^*\mathbf{A}^*.$$

接下来考虑一般情形. 设 $\mathbf{A}(\lambda)=\mathbf{A}+\lambda\mathbf{I}_n$, $\mathbf{B}(\lambda)=\mathbf{B}+\lambda\mathbf{I}_n$. 可以看出 $\det\mathbf{A}(\lambda)=f(\lambda)$, $\det\mathbf{B}(\lambda)=g(\lambda)$ 均是关于 λ 的首一 n 次多项式. 由代数基本定理, $f(\lambda),g(\lambda)$ 均有有限多个根, 故存在 $\mathbb{F}$ 的一个无穷多子集 S, 使得对 $\forall\,\lambda\in S$, $\mathbf{A}(\lambda),\mathbf{B}(\lambda)$ 均是可逆矩阵.

考虑矩阵 $\mathbf{C}(\lambda)=(\mathbf{A}(\lambda)\mathbf{B}(\lambda))^*-\mathbf{B}(\lambda)^*\mathbf{A}(\lambda)^*$, $\mathbf{C}(\lambda)$ 的每个位置的元素是关于 λ 的多项式, 故可设 $\mathbf{C}(\lambda)=(c_{ij}(\lambda))_{n\times n}$. 于是对 $\forall\,\lambda\in S$, $c_{ij}(\lambda)=0$, 即 $c_{ij}(\lambda)$ 有无穷多个根, 故 $c_{ij}(\lambda)\equiv 0$. 特别的, 令 $\lambda=0$, 得到 $(\mathbf{AB})^*=\mathbf{B}^*\mathbf{A}^*$. □

行列式还有一些几何应用:

命题 3.73. 过平面 $\mathbb{R}^2$ 上不同两点 (x_1,y_1), (x_2,y_2) 的直线的方程为

$$\begin{vmatrix} x & y & 1 \\ x_1 & y_1 & 1 \\ x_2 & y_2 & 1 \end{vmatrix}=0. \tag{3.48}$$

证明. 一方面将 (3.48) 中的行列式关于第 1 行展开, 得到方程为 $Ax+By+C=0$, 是平面上某条直线的方程. 另一方面, 令 (x,y) 分别取值 (x_1,y_1), (x_2,y_2) 可知这两点都在这条直线上. □

命题 3.74. 过平面 $\mathbb{R}^2$ 上不共线三点 (x_1,y_1), (x_2,y_2), (x_3,y_3) 的圆的方程为

$$\begin{vmatrix} x^2+y^2 & x & y & 1 \\ x_1^2+y_1^2 & x_1 & y_1 & 1 \\ x_2^2+y_2^2 & x_2 & y_2 & 1 \\ x_3^2+y_3^2 & x_3 & y_3 & 1 \end{vmatrix}=0. \tag{3.49}$$

证明. 一方面将 (3.49) 中的行列式关于第 1 行展开, 得到方程为

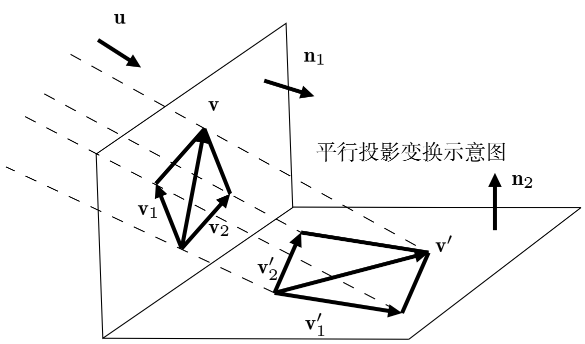
$$A(x^2+y^2)+Bx+Cy+D=0,$$

是平面上某个圆的方程. 另一方面, 令 (x,y) 分别取值 (x_1,y_1), (x_2,y_2), (x_3,y_3) 可知这三点都在这个圆上. $\square$

第 4 章　线性变换

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. 这些几何变换都将直线变为直线, 平行直线变为平行直线.



定义 4.1. 如果映射 $\mathcal{A}\colon \mathbb{F}^n \longrightarrow \mathbb{F}^n$ 满足

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

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

换句话说, 线性变换是从 $\mathbb{F}^n$ 到自身的保持加法与数乘运算的映射.

例题 4.2. 设 $\mathbf{A}\in\mathsf{M}_n(\mathbb{F})$, 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.

例题 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, 由例题 1.13,

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

(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

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

综上, 平面上的伸缩, 旋转, 反射都是某个二阶矩阵的线性变换.

例题 4.4. 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 $\mathbf{a}\neq\mathbf{0}$, 则一般的, 对 $\forall\,\lambda,\mu\in\mathbb{R}$ 且 $\lambda+\mu\neq 1$,

$$\mathcal{T}_{\mathbf{a}}(\lambda\mathbf{x}+\mu\mathbf{y})=\lambda\mathbf{x}+\mu\mathbf{y}+\mathbf{a}\neq\lambda\mathcal{T}_{\mathbf{a}}(\mathbf{x})+\mu\mathcal{T}_{\mathbf{a}}(\mathbf{y}).$$

故平移不是线性变换.

命题 4.5. 设 $\mathcal{A}\colon\mathbb{F}^n\longrightarrow\mathbb{F}^n$ 是线性变换, 则 $\mathcal{A}(\mathbf{0})=\mathbf{0}$.

证明. 在 (4.1) 中取 $\lambda=\mu=1$, $\mathbf{u}=\mathbf{v}=\mathbf{0}$, 则 $\mathcal{A}(\mathbf{0})=\mathcal{A}(\mathbf{0})+\mathcal{A}(\mathbf{0})$, 只能 $\mathcal{A}(\mathbf{0})=\mathbf{0}$. □

例题 4.6. 判断下列映射是否是线性变换:

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

解. $\mathcal{A}(0,0,0)=(0,0,1)$ 不是零向量, 故 $\mathcal{A}$ 不是线性变换. □

研究向量空间及其子空间的手段是找到它的一组基来确定维数和坐标, 研究线性变换的手段就是看它在一组基下的作用.

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

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

从而我们证明了

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 $\mathsf{M}_n(\mathbb{F})$ 中的矩阵.

例题 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

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

解. $\mathcal{A}(\boldsymbol{\epsilon}_1) = [2, 1, -1]$, $\mathcal{A}(\boldsymbol{\epsilon}_2) = [1, 3, -1]$, $\mathcal{A}(\boldsymbol{\epsilon}_3) = [1, 2, 3]$, 故 $\mathcal{A}$ 在标准基下的矩阵为

$$\begin{pmatrix} 2 & 1 & 1 \\ 1 & 3 & 2 \\ -1 & -1 & 3 \end{pmatrix}.$$

□

进一步的, 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 $\boldsymbol{\alpha}=\{\mathbf{v}_1,\dots,\mathbf{v}_n\}$ 下的矩阵. (4.5) 可以写为如下更为简单的形式:

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

对 $\forall\ \mathbf{v}\in\mathbb{F}^n$, 若 $\mathbf{v}$ 在基 $\mathbf{v}_1,\dots,\mathbf{v}_n$ 下的坐标是 $\mathbf{x}=[x_1,\dots,x_n]\in\mathbb{F}^n$, 则

$$\begin{aligned}\mathcal{A}(\mathbf{v})=\mathcal{A}(x_1\mathbf{v}_1+\cdots+x_n\mathbf{v}_n)&=\sum_{j=1}^n x_j\mathcal{A}(\mathbf{v}_j)\\ &=\sum_{j=1}^n x_j\sum_{i=1}^n a_{ij}\mathbf{v}_i=\sum_{i=1}^n(\sum_{j=1}^n a_{ij}x_j)\mathbf{v}_i.\end{aligned}$$

于是 $\mathcal{A}(\mathbf{v})$ 在基 $\mathbf{v}_1,\dots,\mathbf{v}_n$ 下的坐标为

$$\begin{pmatrix}a_{11}x_1+\cdots+a_{1n}x_n\\ \vdots\\ a_{n1}x_1+\cdots+a_{nn}x_n\end{pmatrix}=\begin{pmatrix}a_{11} & \cdots & a_{1n}\\ \vdots & \ddots & \vdots\\ a_{n1} & \cdots & a_{nn}\end{pmatrix}\begin{pmatrix}x_1\\ \vdots\\ x_n\end{pmatrix}=\mathbf{A}\mathbf{x}. \tag{4.7}$$

于是我们证明了

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 $\boldsymbol{\alpha}=\{\mathbf{v}_1,\dots,\mathbf{v}_n\}$ 下的矩阵. 若 $\mathbf{v}$ 在基 $\boldsymbol{\alpha}$ 下的坐标是 $\mathbf{x}$, 则 $\mathcal{A}(\mathbf{v})$ 在基 $\boldsymbol{\alpha}$ 下的坐标是 $\mathbf{A}\mathbf{x}$.

评注 4.10. 利用 (4.6), 定理 4.9 的证明可以简化为:

$$\begin{aligned}\mathcal{A}(\mathbf{v})&=\mathcal{A}(x_1\mathbf{v}_1+\cdots+x_n\mathbf{v}_n)=x_1\mathcal{A}(\mathbf{v}_1)+\cdots+x_n\mathcal{A}(\mathbf{v}_n)\\ &=\begin{pmatrix}\mathcal{A}(\mathbf{v}_1) & \cdots & \mathcal{A}(\mathbf{v}_n)\end{pmatrix}\begin{pmatrix}x_1\\ \vdots\\ x_n\end{pmatrix}=\begin{pmatrix}\mathbf{v}_1 & \cdots & \mathbf{v}_n\end{pmatrix}\cdot\mathbf{A}\mathbf{x}.\end{aligned}$$

映射间可以做复合运算, 容易验证两个 $\mathbb{F}^n$ 上线性变换 $\mathcal{A}$, $\mathcal{B}$ 的复合映射 $\mathcal{A}\circ\mathcal{B}$ (简记为 $\mathcal{AB}$) 仍然是 $\mathbb{F}^n$ 上线性变换. 自然的问题是研究 $\mathcal{AB}$ 在一组基下的矩阵.

命题 4.11. 设 $\mathcal{A}$, $\mathcal{B}$ 是 $\mathbb{F}^n$ 上线性变换, 且在一组基 $\boldsymbol{\alpha}=\{\mathbf{v}_1,\dots,\mathbf{v}_n\}$ 下的矩阵分别是 $\mathbf{A}$, $\mathbf{B}$. 则复合映射 $\mathcal{AB}\colon\mathbb{F}^n\longrightarrow\mathbb{F}^n$ 也是线性变换, 且在基 $\boldsymbol{\alpha}$ 下的矩阵为 $\mathbf{AB}$.

证明. 对 $\forall\ \mathbf{v}_1,\mathbf{v}_2\in\mathbb{F}^n$ 以及 $\forall\ \lambda,\mu\in\mathbb{F}$,

$$\mathcal{AB}(\lambda\mathbf{v}_1+\mu\mathbf{v}_2)=\mathcal{A}(\lambda\mathcal{B}(\mathbf{v}_1)+\mu\mathcal{B}(\mathbf{v}_2))=\lambda\mathcal{AB}(\mathbf{v}_1)+\mu\mathcal{AB}(\mathbf{v}_2),$$

故 $\mathcal{AB}$ 是 $\mathbb{F}^n$ 上的线性变换. 另一方面, 对 $\forall\ 1 \leq k \leq n$,

$$\mathcal{AB}(\mathbf{v}_k) = \mathcal{A}(\sum_{j=1}^{m} b_{jk}\mathbf{v}_j) = \sum_{j=1}^{n} b_{jk}\mathcal{A}(\mathbf{v}_j) = \sum_{j=1}^{n} b_{jk}\sum_{i=1}^{n} a_{ij}\mathbf{v}_i = \sum_{i=1}^{n}(\sum_{j=1}^{n} a_{ij}b_{jk})\mathbf{v}_i,$$

从而 $\mathcal{AB}$ 在基 $\boldsymbol{\alpha}$ 下矩阵为 $\mathbf{AB}$. □

定义 4.12. 设 $\mathcal{A}$ 是 $\mathbb{F}^n$ 上的线性变换,

$$\text{Ker } \mathcal{A} = \{\mathbf{v} \in \mathbb{F}^n \mid \mathcal{A}(\mathbf{v}) = \mathbf{0}\}, \tag{4.8}$$

$$\text{Im } \mathcal{A} = \{\mathbf{w} \in \mathbb{F}^n \mid \exists\ \mathbf{v} \in \mathbb{F}^n \text{ 使得 } \mathcal{A}(\mathbf{v}) = \mathbf{w}\}, \tag{4.9}$$

分别称为 $\mathcal{A}$ 的核空间和像空间.

例题 4.13. 设 $\varphi_{\mathbf{A}}$ 是 $\mathbf{A}$ 的线性变换, 则 $\varphi_{\mathbf{A}}$ 的核空间 Ker $\varphi_{\mathbf{A}}$ 就是 $\mathbf{A}$ 的零空间 Null $\mathbf{A}$, $\varphi_{\mathbf{A}}$ 的像空间 Im $\varphi_{\mathbf{A}}$ 就是 $\mathbf{A}$ 的列空间 Col $\mathbf{A}$.

由例题 4.13 及定理 3.11 就可以得到下面的像核维数定理. 这个定理是线性代数课程的核心定理, 将线性方程组, 矩阵, 线性变换统一的联系起来. 我们再给出一个新的证明, 帮助读者感受线性变换的美妙与用处.

定理 4.14 (像核维数定理)**.** 设 $\mathcal{A}$ 是 $\mathbb{F}^n$ 上的线性变换, 则

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

证明. 设 $\{\mathbf{v}_1, \ldots, \mathbf{v}_s\}$ 是 Ker $\mathcal{A}$ 的一组基, 并将其扩充为 $\mathbb{F}^n$ 的一组基

$$\{\mathbf{v}_1, \ldots, \mathbf{v}_s, \mathbf{v}_{s+1}, \ldots, \mathbf{v}_n\}.$$

我们证明 $\{\mathcal{A}(\mathbf{v}_{s+1}), \ldots, \mathcal{A}(\mathbf{v}_n)\}$ 是 Im $\mathcal{A}$ 的一组基, 从而得到定理的证明.

一方面, 对 $\forall\ \mathbf{v} \in \mathbb{F}^n$, 存在 $\lambda_1, \ldots, \lambda_m$ 使得 $\mathbf{v} = \lambda_1\mathbf{v}_1 + \cdots + \lambda_n\mathbf{v}_n$, 从而

$$\mathcal{A}(\mathbf{v}) = \lambda_1\mathcal{A}(\mathbf{v}_1) + \cdots + \lambda_n\mathcal{A}(\mathbf{v}_n) = \lambda_{s+1}\mathcal{A}(\mathbf{v}_{s+1}) + \cdots + \lambda_n\mathcal{A}(\mathbf{v}_n),$$

于是 Im $\mathcal{A} = \text{span}\{\mathcal{A}(\mathbf{v}_{s+1}), \ldots, \mathcal{A}(\mathbf{v}_n)\}$. 另一方面,

$$\begin{aligned} \lambda_{s+1}\mathcal{A}(\mathbf{v}_{s+1}) + \cdots + \lambda_n\mathcal{A}(\mathbf{v}_n) = \mathbf{0} &\Longrightarrow \mathcal{A}(\lambda_{s+1}\mathbf{v}_{s+1} + \cdots + \lambda_n\mathbf{v}_n) = \mathbf{0} \\ &\Longrightarrow \lambda_{s+1}\mathbf{v}_{s+1} + \cdots + \lambda_n\mathbf{v}_n \in \text{Ker } \mathcal{A}. \end{aligned}$$

于是存在 $-\mu_1, \ldots, -\mu_s \in \mathbb{F}$ 使得

$$\begin{aligned} &\lambda_{s+1}\mathbf{v}_{s+1} + \cdots + \lambda_n\mathbf{v}_n = -\mu_1\mathbf{v}_1 - \cdots - \mu_s\mathbf{v}_s \\ \Longrightarrow\ &\mu_1\mathbf{v}_1 + \cdots + \mu_s\mathbf{v}_s + \lambda_{s+1}\mathbf{v}_{s+1} + \cdots + \lambda_n\mathbf{v}_n = \mathbf{0}. \end{aligned}$$

由 $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ 是一组基就推出 $\lambda_{s+1} = \cdots = \lambda_n = 0$, 从而 $\mathcal{A}(\mathbf{v}_{s+1}), \ldots, \mathcal{A}(\mathbf{v}_n)$ 线性无关. 这样就证明了 $\{\mathcal{A}(\mathbf{v}_{s+1}), \ldots, \mathcal{A}(\mathbf{v}_n)\}$ 是 $\mathrm{Im}\,\mathcal{A}$ 的一组基, 从而 $\dim \mathrm{Ker}\,\mathcal{A} + \dim \mathrm{Im}\,\mathcal{A} = n$. □

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:

$$\begin{aligned} \begin{pmatrix} \mathbf{v}_1' & \mathbf{v}_2' & \cdots & \mathbf{v}_n' \end{pmatrix} &= \begin{pmatrix} \mathbf{v}_1 & \mathbf{v}_2 & \cdots & \mathbf{v}_n \end{pmatrix} \begin{pmatrix} p_{11} & p_{12} & \cdots & p_{1n} \\ p_{21} & p_{22} & \cdots & p_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ p_{n1} & p_{n2} & \cdots & p_{nn} \end{pmatrix} \\ &= \begin{pmatrix} \mathbf{v}_1 & \mathbf{v}_2 & \cdots & \mathbf{v}_n \end{pmatrix} \cdot \mathbf{P}. \end{aligned} \tag{4.10}$$

$\mathbf{P}$ 称为从 $\boldsymbol{\alpha}$ 到 $\boldsymbol{\alpha}'$ 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.

证明. 设 $\begin{pmatrix} \mathbf{v}_1' & \cdots & \mathbf{v}_n' \end{pmatrix} = \begin{pmatrix} \mathbf{v}_1 & \cdots & \mathbf{v}_n \end{pmatrix} \cdot \mathbf{P}$, $\begin{pmatrix} \mathbf{v}_1 & \cdots & \mathbf{v}_n \end{pmatrix} = \begin{pmatrix} \mathbf{v}_1' & \cdots & \mathbf{v}_n' \end{pmatrix} \cdot \mathbf{Q}$, 则

$$\begin{pmatrix} \mathbf{v}_1' & \cdots & \mathbf{v}_n' \end{pmatrix} = \begin{pmatrix} \mathbf{v}_1 & \cdots & \mathbf{v}_n \end{pmatrix} \cdot \mathbf{P} = \begin{pmatrix} \mathbf{v}_1' & \cdots & \mathbf{v}_n' \end{pmatrix} \cdot \mathbf{Q} \cdot \mathbf{P},$$

从而 $\mathbf{QP} = \mathbf{I}_n$, 即 $\mathbf{P}$ 可逆, 且 $\mathbf{Q} = \mathbf{P}^{-1}$. □

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

$$\begin{pmatrix} \mathcal{A}(\mathbf{v}_1) & \cdots & \mathcal{A}(\mathbf{v}_n) \end{pmatrix} = \begin{pmatrix} \mathbf{v}_1 & \cdots & \mathbf{v}_n \end{pmatrix} \cdot \mathbf{A}. \tag{4.11}$$

结合 (4.10) 以及 $\begin{pmatrix} \mathbf{v}_1 & \cdots & \mathbf{v}_n \end{pmatrix} = \begin{pmatrix} \mathbf{v}_1' & \cdots & \mathbf{v}_n' \end{pmatrix} \cdot \mathbf{P}^{-1}$ 就有

$$\begin{pmatrix} \mathcal{A}(\mathbf{v}_1') & \cdots & \mathcal{A}(\mathbf{v}_n') \end{pmatrix} = \begin{pmatrix} \mathcal{A}(\sum_{i=1}^n p_{i1}\mathbf{v}_i) & \cdots & \mathcal{A}(\sum_{i=1}^n p_{in}\mathbf{v}_i) \end{pmatrix} \tag{4.12}$$

$$= \begin{pmatrix} \sum_{i=1}^n p_{i1}\mathcal{A}(\mathbf{v}_i) & \cdots & \sum_{i=1}^n p_{in}\mathcal{A}(\mathbf{v}_i) \end{pmatrix} \tag{4.13}$$

$$= \begin{pmatrix} \mathcal{A}(\mathbf{v}_1) & \cdots & \mathcal{A}(\mathbf{v}_n) \end{pmatrix} \cdot \mathbf{P}. \tag{4.14}$$

$$= \begin{pmatrix} \mathbf{v}_1 & \cdots & \mathbf{v}_n \end{pmatrix} \cdot \mathbf{AP} = \begin{pmatrix} \mathbf{v}_1' & \cdots & \mathbf{v}_n' \end{pmatrix} \cdot \mathbf{P}^{-1}\mathbf{AP}. \tag{4.15}$$

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

定理 4.16. 设 $\mathcal{A}$ 是 $\mathbb{F}^n$ 上的线性变换, $\boldsymbol{\alpha}, \boldsymbol{\alpha}'$ 是 $\mathbb{F}^n$ 的两组基, 且从 $\boldsymbol{\alpha}$ 到 $\boldsymbol{\alpha}'$ 的过渡矩阵为 $\mathbf{P}$. 若 $\mathcal{A}$ 在基 $\boldsymbol{\alpha}$ 下的矩阵是 $\mathbf{A}$, 则 $\mathcal{A}$ 在基 $\boldsymbol{\alpha}'$ 下的矩阵是 $\mathbf{P}^{-1}\mathbf{AP}$.

定义 4.17. 设 $\mathbf{A}, \mathbf{B} \in \mathsf{M}_n(\mathbb{F})$, 若存在可逆矩阵 $\mathbf{P}$, 使得 $\mathbf{P}^{-1}\mathbf{AP} = \mathbf{B}$, 则称 $\mathbf{A}$ 与 $\mathbf{B}$ 相似.

综上, 相似的方阵可以看作同一个线性变换在不同基下的矩阵.

例题 4.18. 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.

解. 将 $\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3$ 按照列向量排成矩阵 $\mathbf{P}$, 则

$$\begin{pmatrix} \mathbf{a}_1 & \mathbf{a}_2 & \mathbf{a}_3 \end{pmatrix} = \begin{pmatrix} \boldsymbol{\epsilon}_1 & \boldsymbol{\epsilon}_2 & \boldsymbol{\epsilon}_3 \end{pmatrix} \cdot \mathbf{P} = \begin{pmatrix} \boldsymbol{\epsilon}_1 & \boldsymbol{\epsilon}_2 & \boldsymbol{\epsilon}_3 \end{pmatrix} \begin{pmatrix} 2 & 0 & 1 \\ 3 & 1 & 0 \\ 5 & 2 & 0 \end{pmatrix},$$

即 $\mathbf{P}$ 是从标准基到 $\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3$ 的过渡矩阵. 需要求 $\mathbf{P}^{-1}\mathbf{AP}$.

因为 $\det \mathbf{P} = 1$, 可知 $\mathbf{P}^{-1} = \mathbf{P}^* = \begin{pmatrix} 0 & 2 & -1 \\ 0 & -5 & 3 \\ 1 & -4 & 2 \end{pmatrix}$. 从而 $\mathcal{A}$ 在基 $\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3$ 下的矩阵是

$$\mathbf{P}^{-1}\mathbf{AP} = \begin{pmatrix} 0 & 2 & -1 \\ 0 & -5 & 3 \\ 1 & -4 & 2 \end{pmatrix} \begin{pmatrix} 3 & -20 & 11 \\ 0 & -16 & 10 \\ 5 & -15 & 7 \end{pmatrix} \begin{pmatrix} 2 & 0 & 1 \\ 3 & 1 & 0 \\ 5 & 2 & 0 \end{pmatrix} = \begin{pmatrix} 4 & 9 & -5 \\ -10 & -23 & 15 \\ -7 & -16 & 13 \end{pmatrix}. \qquad \square$$

例题 4.19. 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

$$\mathbf{P}^{-1}\mathbf{AP} = \begin{pmatrix} \mathbf{I}_r & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix}.$$

证明. 设 $\varphi_{\mathbf{A}}$ 是 $\mathbf{A}$ 的线性变换. 由条件 $\dim \operatorname{Ker} \varphi_{\mathbf{A}} = n-r$, 且对 $\forall\ \mathbf{x} \in \mathbb{F}^n$,

$$\varphi_{\mathbf{A}}^2(\mathbf{x}) = \varphi_{\mathbf{A}}(\varphi_{\mathbf{A}}(\mathbf{x})) = \mathbf{A}^2\mathbf{x} = \mathbf{A}\mathbf{x} = \varphi_{\mathbf{A}}(\mathbf{x}).$$

问题等价于证明 $\varphi_{\mathbf{A}}$ 在 $\mathbb{F}^n$ 的一组基下的矩阵为 $\begin{pmatrix} \mathbf{I}_r & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix}$.

任取 $\operatorname{Ker} \varphi_{\mathbf{A}}$ 的一组基 $\{\mathbf{v}_{r+1}, \ldots, \mathbf{v}_n\}$, 并将其扩充为 $\mathbb{F}^n$ 的一组基

$$\{\mathbf{v}_1, \ldots, \mathbf{v}_r, \mathbf{v}_{r+1}, \ldots, \mathbf{v}_n\}.$$

我们证明 $\boldsymbol{\alpha} = \{\varphi_{\mathbf{A}}(\mathbf{v}_1), \ldots, \varphi_{\mathbf{A}}(\mathbf{v}_r), \mathbf{v}_{r+1}, \ldots, \mathbf{v}_n\}$ 也是 $\mathbb{F}^n$ 的一组基, 只需证明它们线性无关. 若

$$\lambda_1\varphi_{\mathbf{A}}(\mathbf{v}_1) + \cdots + \lambda_r\varphi_{\mathbf{A}}(\mathbf{v}_r) + \lambda_{r+1}\mathbf{v}_{r+1} + \cdots + \lambda_n\mathbf{v}_n = \mathbf{0},$$

等式两边用 $\varphi_{\mathbf{A}}$ 作用得到

$$\begin{aligned}\mathbf{0} &= \varphi_{\mathbf{A}}(\lambda_1\varphi_{\mathbf{A}}(\mathbf{v}_1) + \cdots + \lambda_r\varphi_{\mathbf{A}}(\mathbf{v}_r) + \lambda_{r+1}\mathbf{v}_{r+1} + \cdots + \lambda_n\mathbf{v}_n) \\ &= \lambda_1\varphi_{\mathbf{A}}^2(\mathbf{v}_1) + \cdots + \lambda_r\varphi_{\mathbf{A}}^2(\mathbf{v}_r) = \varphi_{\mathbf{A}}(\lambda_1\mathbf{v}_1 + \cdots + \lambda_r\mathbf{v}_r).\end{aligned}$$

从而 $\lambda_1\mathbf{v}_1 + \cdots + \lambda_r\mathbf{v}_r \in \operatorname{Ker} \varphi_{\mathbf{A}}$, 存在 $\mu_{r+1}, \ldots, \mu_n \in \mathbb{F}$ 使得

$$\lambda_1\mathbf{v}_1 + \cdots + \lambda_r\mathbf{v}_r = \mu_{r+1}\mathbf{v}_{r+1} + \cdots + \mu_n\mathbf{v}_n.$$

因为 $\mathbf{v}_1, \ldots, \mathbf{v}_n$ 是一组基, 只能 $\lambda_1 = \cdots = \lambda_r = 0$, 进一步的也有 $\lambda_{r+1} = \cdots = \lambda_n = 0$. 从而 $\boldsymbol{\alpha}$ 中向量线性无关, 是 $\mathbb{F}^n$ 的一组基. 可以看出 $\varphi_{\mathbf{A}}$ 在 $\boldsymbol{\alpha}$ 下的矩阵为 $\begin{pmatrix} \mathbf{I}_r & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix}$. □

例题 4.20. 设 $\mathbf{A}, \mathbf{B} \in \mathsf{M}_n(\mathbb{F})$, $\mathbf{AB} = \mathbf{BA} = 0$, $\operatorname{rank} \mathbf{A}^2 = \operatorname{rank} \mathbf{A}$. 证明 $\operatorname{rank}(\mathbf{A}+\mathbf{B}) = \operatorname{rank} \mathbf{A} + \operatorname{rank} \mathbf{B}$.

证明. 设 $\varphi_{\mathbf{A}}, \varphi_{\mathbf{B}}$ 是 $\mathbf{A}, \mathbf{B}$ 的线性变换. 条件等价于 $\varphi_{\mathbf{A}}\varphi_{\mathbf{B}} = \varphi_{\mathbf{B}}\varphi_{\mathbf{A}} = 0$, $\operatorname{Im} \varphi_{\mathbf{A}}^2 = \operatorname{Im} \varphi_{\mathbf{A}}$. 设 $\{\mathbf{v}_1, \ldots, \mathbf{v}_s\}$ 是 $\operatorname{Ker} \varphi_{\mathbf{A}}$ 的一组基, 并将其扩充为 $\mathbb{F}^n$ 的一组基 $\{\mathbf{v}_1, \ldots, \mathbf{v}_s, \mathbf{v}_{s+1}, \ldots, \mathbf{v}_n\}$. 和例题 4.19 中类似, 可以证明 $\boldsymbol{\alpha} = \{\mathbf{v}_1, \ldots, \mathbf{v}_s, \varphi_{\mathbf{A}}(\mathbf{v}_{s+1}), \ldots, \varphi_{\mathbf{A}}(\mathbf{v}_n)\}$ 也是 $\mathbb{F}^n$ 的一组基. 在基 $\boldsymbol{\alpha}$ 下 $\varphi_{\mathbf{A}}, \varphi_{\mathbf{B}}$ 的矩阵分别为

$$\tilde{\mathbf{A}} = \begin{array}{cc} \begin{array}{cc} s & n-s \end{array} & \\ \begin{pmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{A}' \end{pmatrix} & \begin{array}{c} s \\ n-s \end{array} \end{array}, \quad \tilde{\mathbf{B}} = \begin{array}{cc} \begin{array}{cc} s & n-s \end{array} & \\ \begin{pmatrix} \mathbf{B}' & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix} & \begin{array}{c} s \\ n-s \end{array} \end{array}.$$

换句话说, 设 $\mathbf{P}$ 是从标准基到 $\boldsymbol{\alpha}$ 的过渡矩阵, 则 $\mathbf{P}^{-1}\mathbf{A}\mathbf{P} = \tilde{\mathbf{A}}$, $\mathbf{P}^{-1}\mathbf{B}\mathbf{P} = \tilde{\mathbf{B}}$. 因为

$$\operatorname{rank}\,(\tilde{\mathbf{A}}+\tilde{\mathbf{B}}) = \operatorname{rank}\begin{pmatrix}\mathbf{B}' & \mathbf{0}\\ \mathbf{0} & \mathbf{A}'\end{pmatrix} = \operatorname{rank}\,\mathbf{A}' + \operatorname{rank}\,\mathbf{B}' = \operatorname{rank}\,\tilde{\mathbf{A}} + \operatorname{rank}\,\tilde{\mathbf{B}}.$$

所以 $\operatorname{rank}\,(\mathbf{A}+\mathbf{B}) = \operatorname{rank}\,\mathbf{P}^{-1}(\mathbf{A}+\mathbf{B})\mathbf{P} = \operatorname{rank}\,(\tilde{\mathbf{A}}+\tilde{\mathbf{B}}) = \operatorname{rank}\,\tilde{\mathbf{A}} + \operatorname{rank}\,\tilde{\mathbf{B}} = \operatorname{rank}\,\mathbf{P}\tilde{\mathbf{A}}\mathbf{P}^{-1} + \operatorname{rank}\,\mathbf{P}\tilde{\mathbf{B}}\mathbf{P}^{-1} = \operatorname{rank}\,\mathbf{A} + \operatorname{rank}\,\mathbf{B}$. □

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:

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

但并不是每一个方阵都相似于对角矩阵, 例如:

例题 4.21. 证明 $\mathbf{A} = \begin{pmatrix}0 & 1\\ 0 & 0\end{pmatrix}$ 不可能相似于一个对角矩阵.

证明. 假设 $\mathbf{P} = \begin{pmatrix}a & b\\ c & d\end{pmatrix}$ 是可逆矩阵, 且 $\mathbf{P}^{-1}\mathbf{A}\mathbf{P} = \begin{pmatrix}\lambda_1 & 0\\ 0 & \lambda_2\end{pmatrix}$ 是对角矩阵. 因为 $\mathbf{P}$ 可逆, $\operatorname{rank}\,\mathbf{P}^{-1}\mathbf{A}\mathbf{P} = \operatorname{rank}\,\mathbf{A} = 1$. 从而 λ_1, λ_2 中恰好有一个为 0, 不妨设 $\lambda_1 \neq 0, \lambda_2 = 0$. 于是

$$\mathbf{P}^{-1}\mathbf{A}\mathbf{P} = \begin{pmatrix}\lambda_1 & 0\\ 0 & 0\end{pmatrix} \Longrightarrow \mathbf{A}\mathbf{P} = \mathbf{P}\begin{pmatrix}\lambda_1 & 0\\ 0 & 0\end{pmatrix} \Longrightarrow c = d = 0,$$

与 $\mathbf{P}$ 可逆矛盾. □

定义 4.22. 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

$$\mathbf{P}^{-1}\mathbf{A}\mathbf{P} = \operatorname{diag}(\lambda_1, \ldots, \lambda_n), \tag{4.16}$$

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

定义 4.23. 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

$$\mathcal{A}(\mathbf{v}_i) = \lambda_i \mathbf{v}_i, \quad 1 \le i \le n, \tag{4.17}$$

就称 $\mathcal{A}$ 是可对角化的.

定义 4.24. 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 4.25. 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 $\mathbf{A} \in \mathsf{M}_n(\mathbb{F})$, 当 λ 看作变量时, $\det\,(\lambda \mathbf{I}_n - \mathbf{A})$ 是关于 λ 的 $\mathbb{F}$ 上的首项系数为 1 的一元 n 次多项式, 记为 $p_{\mathbf{A}}(\lambda)$, 称为 $\mathbf{A}$ 的特征多项式. 从而

命题 4.26. 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 $\mathbb{C}$. $\mathbf{A} \in \mathsf{M}_n(\mathbb{C})$ 的特征值和特征向量的计算方法为:

(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.

例题 4.27. 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.

解. $\mathbf{A}$ 的特征多项式为

$$p_{\mathbf{A}}(\lambda)=\begin{vmatrix}\lambda-3&1&2\\-2&\lambda&2\\-2&1&\lambda+1\end{vmatrix}=\lambda(\lambda-1)^2.$$

故 $\mathbf{A}$ 的特征值为 $\lambda_1=0,\lambda_2=1$.

线性方程组 $(\mathbf{A}-\lambda_1\mathbf{I}_3)\mathbf{x}=\mathbf{0}$ 的解空间为

$$V_{\lambda_1}=\left\{\mu\cdot\begin{pmatrix}1\\1\\1\end{pmatrix},\ \mu\in\mathbb{C}\right\}.$$

线性方程组 $(\mathbf{A}-\lambda_2\mathbf{I}_3)\mathbf{x}=\mathbf{0}$ 的解空间为

$$V_{\lambda_2}=\left\{\mu\cdot\begin{pmatrix}1\\2\\0\end{pmatrix}+\gamma\cdot\begin{pmatrix}0\\-2\\1\end{pmatrix},\ \mu,\gamma\in\mathbb{C}\right\}.$$

$V_{\lambda_1},V_{\lambda_2}$ 中的向量是 $\mathbf{A}$ 的全部特征向量. □

引理 4.28. 设 $\mathbf{A}\in\mathsf{M}_n(\mathbb{F})$, 则 $|\mathbf{A}|=0\Longleftrightarrow\exists\,\mathbf{B}\neq\mathbf{0}\in\mathsf{M}_n(\mathbb{F})$ 使得 $\mathbf{AB}=\mathbf{0}$.

证明. "$\Longrightarrow$": $|\mathbf{A}|=0\Longleftrightarrow$ $\mathbf{Ax}=\mathbf{0}$ 有非零解 $\mathbf{b}$, 令 $\mathbf{B}=\begin{pmatrix}\mathbf{b}&\cdots&\mathbf{b}\end{pmatrix}$, 即 $\mathbf{B}$ 是列向量全为 $\mathbf{b}$ 的方阵, 则 $\mathbf{AB}=\mathbf{0}$ 且 $\mathbf{B}\neq\mathbf{0}$.
"$\Longleftarrow$": 设 $\mathbf{B}\neq\mathbf{0}$ 且 $\mathbf{AB}=\mathbf{0}$, 则存在 $\mathbf{B}$ 的一个列向量 $\mathbf{b}\neq\mathbf{0}$ 且 $\mathbf{Ab}=\mathbf{0}$, 故 $\mathbf{Ax}=\mathbf{0}$ 有非零解, 从而 $|\mathbf{A}|=0$. □

例题 4.29. 设 $\mathbf{A}=\begin{pmatrix}1&1\\0&2\end{pmatrix}$, 定义线性变换 $\mathcal{A}\colon\mathsf{M}_2(\mathbb{C})\longrightarrow\mathsf{M}_2(\mathbb{C})$ 为

$$\mathcal{A}(\mathbf{X})=\mathbf{AX},\quad \mathbf{X}\in\mathsf{M}_2(\mathbb{C}).$$

求 $\mathcal{A}$ 的所有特征值和特征向量.

解. 利用引理 4.28, 可知 λ 是 $\mathcal{A}$ 的特征值 $\iff \exists\, \mathbf{X} \neq \mathbf{0}$ 使得 $\mathbf{AX} = \lambda\mathbf{X} \iff \exists\, \mathbf{X} \neq \mathbf{0}$ 使得 $(\mathbf{A} - \lambda\mathbf{I}_2)\mathbf{X} = \mathbf{0} \iff |\mathbf{A} - \lambda\mathbf{I}_2| = 0 \iff \lambda$ 是 $\mathbf{A}$ 的特征值.

$\mathbf{A}$ 的特征值为 $1, 2$, 故 $\mathcal{A}$ 的全部特征值为 $\lambda_1 = 1, \lambda_2 = 2$. 设 $\mathbf{X} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$, 则

$$\mathbf{AX} = \lambda_1\mathbf{X} \iff \begin{pmatrix} 1 & 1 \\ 0 & 2 \end{pmatrix}\begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \iff \mathbf{X} = \begin{pmatrix} a & b \\ 0 & 0 \end{pmatrix}, \quad a, b \in \mathbb{C}.$$

$$\mathbf{AX} = \lambda_2\mathbf{X} \iff \begin{pmatrix} 1 & 1 \\ 0 & 2 \end{pmatrix}\begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} 2a & 2b \\ 2c & 2d \end{pmatrix} \iff \mathbf{X} = \begin{pmatrix} a & b \\ a & b \end{pmatrix}, \quad a, b \in \mathbb{C}.$$

故 $\mathcal{A}$ 的全部特征向量是所有形如 $\begin{pmatrix} a & b \\ 0 & 0 \end{pmatrix}$ 或 $\begin{pmatrix} a & b \\ a & b \end{pmatrix}$ 的二阶矩阵. □

虽然方阵不一定可对角化, 但可以证明 $\mathsf{M}_n(\mathbb{C})$ 中的方阵都可以上三角化, 即相似于一个上三角矩阵. 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. 一般来说, 我们只关心上三角矩阵中对角线上的元素, 记对角线上元素为 $\lambda_1, \ldots, \lambda_n$ 的上三角矩阵为 $\mathrm{utri}(\lambda_1, \ldots, \lambda_n)$.

引理 4.30. 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.

证明. 设 $\mathbf{B} = \mathbf{P}^{-1}\mathbf{AP}$, 其中 $\mathbf{P}$ 是可逆矩阵, 则

$$p_{\mathbf{B}}(\lambda) = |\lambda\mathbf{I}_n - \mathbf{B}| = \left|\lambda\mathbf{I}_n - \mathbf{P}^{-1}\mathbf{AP}\right| = \left|\mathbf{P}^{-1}(\lambda\mathbf{I}_n - \mathbf{A})\mathbf{P}\right| = |\lambda\mathbf{I}_n - \mathbf{A}| = p_{\mathbf{A}}(\lambda).$$ □

定理 4.31. 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.

证明. 用归纳法, $n = 1$ 时任何方阵都是上三角矩阵. 假设任何 $n-1$ 阶方阵都相似于上三角矩阵. 对 $\forall\, \mathbf{A} \in \mathsf{M}_n(\mathbb{C})$, 设 λ_1 是 $\mathbf{A}$ 的一个特征值, 且 $\mathbf{v}_1 \neq \mathbf{0}$ 是关于 λ_1 的一个特征向量. 将 $\mathbf{v}_1$ 扩充为 $\mathbf{C}^n$ 的一组基 $\boldsymbol{\alpha} = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$, 则 $\mathbf{A}$ 的线性变换 $\varphi_{\mathbf{A}}$ 在基 $\boldsymbol{\alpha}$ 下的矩阵为

$$\begin{pmatrix} \lambda_1 & \mathbf{A}_1 \\ \mathbf{0} & \mathbf{A}_2 \end{pmatrix}, \quad \mathbf{A}_2 \in \mathsf{M}_{n-1}(\mathbb{C}).$$

换句话说, 存在可逆矩阵 $\mathbf{P}_1 \in \mathsf{M}_n(\mathbb{C})$, 使得

$$\mathbf{P}_1^{-1}\mathbf{A}\mathbf{P}_1 = \begin{pmatrix} \lambda_1 & \mathbf{A}_1 \\ \mathbf{0} & \mathbf{A}_2 \end{pmatrix}.$$

由归纳假设, $\mathbf{A}_2$ 可上三角化, 即存在可逆矩阵 $\mathbf{Q} \in \mathsf{M}_{n-1}(\mathbb{C})$ 使得

$$\mathbf{Q}^{-1}\mathbf{A}_2\mathbf{Q} = \begin{pmatrix} \lambda_2 & \cdots & * \\ 0 & \ddots & \vdots \\ 0 & 0 & \lambda_n \end{pmatrix}.$$

考虑 $\mathbf{P}_2 = \begin{pmatrix} 1 & \mathbf{0} \\ \mathbf{0} & \mathbf{Q} \end{pmatrix} \in \mathsf{M}_n(\mathbb{C})$, 则 $\mathbf{P}_2$ 可逆且 $\mathbf{P}_2^{-1} = \begin{pmatrix} 1 & \mathbf{0} \\ \mathbf{0} & \mathbf{Q}^{-1} \end{pmatrix} \in \mathsf{M}_n(\mathbb{C})$. 令 $\mathbf{P} = \mathbf{P}_1\mathbf{P}_2$, 则

$$\mathbf{P}^{-1}\mathbf{A}\mathbf{P} = \begin{pmatrix} 1 & \mathbf{0} \\ \mathbf{0} & \mathbf{Q}^{-1} \end{pmatrix}\begin{pmatrix} \lambda_1 & \mathbf{A}_1 \\ \mathbf{0} & \mathbf{A}_2 \end{pmatrix}\begin{pmatrix} 1 & \mathbf{0} \\ \mathbf{0} & \mathbf{Q} \end{pmatrix} = \begin{pmatrix} \lambda_1 & \mathbf{A}_1\mathbf{Q} \\ \mathbf{0} & \mathbf{Q}^{-1}\mathbf{A}_2\mathbf{Q} \end{pmatrix} = \begin{pmatrix} \lambda_1 & \cdots & * \\ 0 & \ddots & \vdots \\ 0 & 0 & \lambda_n \end{pmatrix}$$

是上三角矩阵.

上三角矩阵 $\mathbf{P}^{-1}\mathbf{A}\mathbf{P}$ 的特征多项式为 $(\lambda-\lambda_1)\cdots(\lambda-\lambda_n)$, 由引理 4.30, 这也是 $\mathbf{A}$ 的特征多项式, 所以 $\lambda_1, \ldots, \lambda_n$ 就是 $\mathbf{A}$ 的全部特征值. □

推论 4.32. 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 $\lambda_i \neq 0$.

证明. 由定理 4.31, 存在可逆矩阵 $\mathbf{P}$ 使得 $\mathbf{P}^{-1}\mathbf{A}\mathbf{P} = \mathrm{utri}(\lambda_1, \ldots, \lambda_n)$, 从而

$$\det \mathbf{A} = \det\,(\mathbf{P}^{-1}\mathbf{A}\mathbf{P}) = \det \mathrm{utri}(\lambda_1, \ldots, \lambda_n) = \lambda_1 \cdots \lambda_n.$$ □

命题 4.33. 设 $\mathbf{A} = (a_{ij})_{n\times n} \in \mathsf{M}_n(\mathbb{F})$, 记 $\operatorname{tr}\mathbf{A} = a_{11} + a_{22} + \cdots + a_{nn}$, 称为 $\mathbf{A}$ 的迹. 证明 $\operatorname{tr}\mathbf{A}\mathbf{B} = \operatorname{tr}\mathbf{B}\mathbf{A}$.

证明. 设 $\mathbf{B} = (b_{ij})_{n\times n}$, $\mathbf{A}\mathbf{B} = (d_{ij})_{n\times n}$, $\mathbf{B}\mathbf{A} = (f_{ij})_{n\times n}$, 则

$$d_{ii} = \sum_{k=1}^{n} a_{ik}b_{ki}, \quad f_{ii} = \sum_{k=1}^{n} b_{ik}a_{ki}, \quad 1 \le i \le n.$$

从而

$$\mathrm{tr}\ \mathbf{AB}=\sum_{i=1}^{n} d_{ii}=\sum_{i,k=1}^{n} a_{ik}b_{ki}=\sum_{i,k=1}^{n} b_{ik}a_{ki}=\sum_{i=1}^{n} f_{ii}=\mathrm{tr}\ \mathbf{BA}. \qquad \square$$

推论 4.34. 设 $\mathbf{A}\in \mathsf{M}_n(\mathbb{C})$, $\lambda_1,\dots,\lambda_n$ 是 $\mathbf{A}$ 的全部特征值, 则 $\mathrm{tr}\,\mathbf{A}=\lambda_1+\cdots+\lambda_n$.

证明. 由定理 4.31, 存在可逆矩阵 $\mathbf{P}$ 使得 $\mathbf{P}^{-1}\mathbf{AP}=\mathrm{utri}(\lambda_1,\dots,\lambda_n)$, 再用命题 4.33,

$$\mathrm{tr}\ \mathbf{A}=\mathrm{tr}\ \mathbf{PP}^{-1}\mathbf{A}=\mathrm{tr}\ \mathbf{P}^{-1}\mathbf{AP}=\lambda_1+\cdots+\lambda_n. \qquad \square$$

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.

例题 4.35. 证明 $\mathbf{A}=\begin{pmatrix}1&1&0\\0&1&1\\0&0&1\end{pmatrix}$ 不可对角化.

证明. $p_{\mathbf{A}}(\lambda)=(\lambda-1)^3$, $\mathbf{A}$ 只有一个特征值 $\lambda_1=1$. 另一方面, $\mathrm{rank}\ (\lambda_1\mathbf{I}_3-\mathbf{A})=2$, 从而 $(\lambda_1\mathbf{I}_3-\mathbf{A})\mathbf{x}=\mathbf{0}$ 的解空间 V_{λ_1} 是 1 维子空间, 不可能找到 3 个线性无关的特征向量. 故 $\mathbf{A}$ 不可对角化. $\square$

定义 4.36. 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 $\forall\ 1\le i\le s$,

(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 4.37. 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.

证明. 设 λ 的代数重数为 s. 存在可逆矩阵 $\mathbf{P}$ 使得 $\mathbf{P}^{-1}\mathbf{AP}=\mathrm{utri}(\lambda_1,\dots,\lambda_n)$, 且不妨设对 $\forall\ 1\le i\le s$, $\lambda_i=\lambda$, 而对 $\forall\ s+1\le i\le n$, $\lambda_i\ne\lambda$. 从而

$$\mathbf{P}^{-1}(\mathbf{A}-\lambda\mathbf{I}_n)\mathbf{P}=\mathbf{P}^{-1}\mathbf{AP}-\lambda\mathbf{I}_n=\mathrm{utri}(0,\dots,0,\lambda_{s+1}-\lambda,\dots,\lambda_n-\lambda).$$

因为 $\lambda_{s+1}-\lambda,\dots,\lambda_n-\lambda$ 均不为 0, 故 $\mathrm{rank}\ \mathbf{P}^{-1}(\mathbf{A}-\lambda\mathbf{I}_n)\mathbf{P}\ge n-s$. 从而

$$\dim V_\lambda=n-\mathrm{rank}\ (\mathbf{A}-\lambda\mathbf{I}_n)=n-\mathrm{rank}\ \mathbf{P}^{-1}(\mathbf{A}-\lambda\mathbf{I}_n)\mathbf{P}\le s.$$

即 λ 的几何重数不超过代数重数. $\square$

命题 4.38. 设 $\mathbf{v}_1,\dots,\mathbf{v}_s$ 是 $\mathbf{A}$ 的特征值分别为 $\lambda_1,\dots,\lambda_s$ 的特征向量, 且 $\lambda_1,\dots,\lambda_s$ 互不相同, 则 $\mathbf{v}_1,\dots,\mathbf{v}_s$ 线性无关.

证明. 用归纳法. $s=1$ 时, 因为特征向量不是零向量, 命题成立. 现假设命题对 $s-1$ 成立, 即任意 $s-1$ 个特征值两两不同的特征向量都线性无关. 设 $\mathbf{v}_1,\ldots,\mathbf{v}_s$ 是 s 个特征值两两不同的特征向量, 且 $\mu_1\mathbf{v}_1+\cdots+\mu_{s-1}\mathbf{v}_{s-1}+\mu_s\mathbf{v}_s=\mathbf{0}$. 则

$$\lambda_1\mu_1\mathbf{v}_1+\cdots+\lambda_{s-1}\mu_{s-1}\mathbf{v}_{s-1}+\lambda_s\mu_s\mathbf{v}_s=\mathbf{A}(\mu_1\mathbf{v}_1+\cdots+\mu_{s-1}\mathbf{v}_{s-1}+\mu_s\mathbf{v}_s)=\mathbf{0},$$

$$\lambda_s\mu_1\mathbf{v}_1+\cdots+\lambda_s\mu_{s-1}\mathbf{v}_{s-1}+\lambda_s\mu_s\mathbf{v}_s=\lambda_s(\mu_1\mathbf{v}_1+\cdots+\mu_{s-1}\mathbf{v}_{s-1}+\mu_s\mathbf{v}_s)=\mathbf{0}.$$

从而 $(\lambda_s-\lambda_1)\mu_1\mathbf{v}_1+\cdots+(\lambda_s-\lambda_{s-1})\mu_{s-1}\mathbf{v}_{s-1}=\mathbf{0}$. 对任意 $1\le i\le s-1$, 由归纳假设知 $(\lambda_s-\lambda_i)\mu_i=0$, 因为 $\lambda_s\ne\lambda_i$, 故 $\mu_i=0$. 进而 $\mu_s=0$, 即 $\mathbf{v}_1,\ldots,\mathbf{v}_s$ 线性无关. □

定理 4.39. $\mathbf{A}\in\mathsf{M}_n(\mathbb{C})$ 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.

证明. "$\Longrightarrow$": 设 λ 是 $\mathbf{A}$ 的特征值, 代数重数为 s. 若存在 $\mathbf{P}$ 使得 $\mathbf{P}^{-1}\mathbf{AP}=\mathrm{diag}(\lambda_1,\ldots,\lambda_n)$, 和命题 4.37 的证明类似, 此时

$$\dim V_\lambda=n-\mathrm{rank}\ (\mathbf{A}-\lambda\mathbf{I}_n)=n-(n-s)=s.$$

故几何重数等于代数重数.
"$\Longleftarrow$": 设 $\lambda_1,\ldots,\lambda_s$ 是 $\mathbf{A}$ 的全部两两不同的特征值, 代数重数为 $n_1,\ldots,n_s$. 对任意 $1\le i\le s$, 因为 $\dim V_{\lambda_i}=n_i$, 故可取 V_{λ_i} 的一组基

$$\mathbf{v}_{i1},\ldots,\mathbf{v}_{in_i},\quad 1\le i\le s.$$

我们证明

$$\boldsymbol{\alpha}=\{\mathbf{v}_{11},\ldots,\mathbf{v}_{1n_1},\mathbf{v}_{21},\ldots,\mathbf{v}_{2n_2},\ldots,\mathbf{v}_{s1},\ldots,\mathbf{v}_{sn_s}\}$$

是 $\mathbf{F}^n$ 的一组基. 因为 $n_1+\cdots+n_s=n$, 只需证它们线性无关. 若

$$\lambda_{11}\mathbf{v}_{11}+\cdots+\lambda_{1n_1}\mathbf{v}_{1n_1}+\lambda_{21}\mathbf{v}_{21}+\cdots+\lambda_{2n_2}\mathbf{v}_{2n_2}+\cdots+\lambda_{s1}\mathbf{v}_{s1}+\cdots+\lambda_{sn_s}\mathbf{v}_{sn_s}=\mathbf{0},$$

设 $\mathbf{u}_i=\lambda_{i1}\mathbf{v}_{i1}+\cdots+\lambda_{in_i}\mathbf{v}_{in_i}$, 则由命题 4.38 及 $\mathbf{u}_1+\cdots+\mathbf{u}_t=\mathbf{0}$ 知每个 $\mathbf{u}_i=\mathbf{0}$. 再由 $\mathbf{v}_{i1},\ldots,\mathbf{v}_{in_i}$ 是一组基得到

$$\lambda_{i1}=\cdots=\lambda_{in_i}=0,\quad 1\le i\le s.$$

从而 $\boldsymbol{\alpha}$ 是 $\mathbb{F}^n$ 的一组基, 在这组基下 $\varphi_{\mathbf{A}}$ 的矩阵是对角矩阵, 故 $\mathbf{A}$ 可对角化. □

若 $\mathbf{A}$ 可对角化, 即存在可逆矩阵 $\mathbf{P}$ 使得 $\mathbf{P}^{-1}\mathbf{AP}=\mathrm{diag}(\lambda_1,\ldots,\lambda_n)$. 设 $\mathbf{P}$ 的行向量形

式为 $\mathbf{P}=\begin{pmatrix}\mathbf{p}_1 & \cdots & \mathbf{p}_n\end{pmatrix}$, 则

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

即 $\mathbf{p}_i$ 是 $\mathbf{A}$ 的关于特征值 λ_i 的特征向量.

例题 4.40. $\mathbf{A}=\begin{pmatrix}3 & -1 & -2\\ 2 & 0 & -2\\ 2 & -1 & -1\end{pmatrix}$ 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.

解. 在例题 4.27 中已算出 $p_{\mathbf{A}}(\lambda)=\lambda(\lambda-1)^2$, 特征值为 $\lambda_1=0$, $\lambda_2=1$, 且 $\dim V_{\lambda_1}=1$, $\dim V_{\lambda_2}=2$. 特征值的代数重数等于几何重数, 故 $\mathbf{A}$ 可对角化.

选取 V_{λ_1} 的一组基 $[1,1,1]$, V_{λ_2} 的一组基 $[1,2,0],[0,-2,1]$, 将它们按照列向量排成矩阵

$$\mathbf{P}=\begin{pmatrix}1 & 1 & 0\\ 1 & 2 & -2\\ 1 & 0 & 1\end{pmatrix},$$

就有 $\mathbf{P}^{-1}\mathbf{A}\mathbf{P}=\operatorname{diag}(0,1,1)$ 是对角矩阵. □

例题 4.41. 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?

证明. 由例题 4.3, 已经知道 $\mathcal{S}_{\lambda,\mu}$, $\mathcal{R}_\theta$ 以及 r_θ 在 $\mathbb{R}^2$ 的一组基下的矩阵.

(1) $\mathcal{S}_{\lambda,\mu}$ 在一组基下的矩阵是 $\begin{pmatrix}\lambda & 0\\ 0 & \mu\end{pmatrix}$, 故可对角化.

(2) $\mathcal{R}_\theta$ 在一组基下的矩阵是 $\begin{pmatrix}\cos\theta & -\sin\theta\\ \sin\theta & \cos\theta\end{pmatrix}$, $p(\lambda)=\lambda^2-2\cos\theta\lambda+1$. 若 $\cos\theta=\pm1$, 则已经是对角矩阵, 可对角化; 若 $\cos\theta\neq\pm1$, 则 $p(\lambda)$ 的根不是实数, 故在 $\mathbb{R}$ 上不可对角化.

(3) r_θ 在一组基下的矩阵是 $\begin{pmatrix}\cos 2\theta & \sin 2\theta\\ \sin 2\theta & -\cos 2\theta\end{pmatrix}$, $p(\lambda)=\lambda^2-1$, 有两不同特征值 1, -1, 故在 $\mathbb{R}$ 上可对角化, 且在一组基下矩阵为 $\begin{pmatrix}1 & 0\\ 0 & -1\end{pmatrix}$. □

例题 4.42. 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.

解. $\mathbf{A}$ 有三个线性无关的特征向量, 等价于 $\mathbf{A}$ 可对角化.

$$p_{\mathbf{A}}(\lambda)=\begin{vmatrix}\lambda & 0 & -1\\ -x & \lambda-1 & -y\\ -1 & 0 & \lambda\end{vmatrix}=(\lambda-1)^2(\lambda+1),$$

故特征值为 $\lambda_1=1$, $\lambda_2=-1$. 由代数重数等于几何重数知

$$\operatorname{rank}\,(\mathbf{A}-\mathbf{I}_3)=3-\dim V_{\lambda_1}=1,\quad \operatorname{rank}\,(\mathbf{A}+\mathbf{I}_3)=3-\dim V_{\lambda_2}=2.$$

对 $(\mathbf{A}-\mathbf{I}_3)$ 做行变换并由 $\operatorname{rank}\,(\mathbf{A}-\mathbf{I}_3)=1$ 可得

$$\begin{pmatrix}-1 & 0 & 1\\ x & 0 & y\\ 1 & 0 & -1\end{pmatrix}\longrightarrow\begin{pmatrix}1 & 0 & -1\\ x & 0 & y\\ 0 & 0 & 0\end{pmatrix}\longrightarrow\begin{pmatrix}1 & 0 & -1\\ 0 & 0 & x+y\\ 0 & 0 & 0\end{pmatrix}\Longrightarrow x+y=0.$$

可以验证对任何 $x+y=0$, $\operatorname{rank}\,(\mathbf{A}+\mathbf{I}_3)=2$. 故所求即为 $x+y=0$. □

例题 4.43. 求 $\mathbf{A}\in\mathsf{M}_3(\mathbb{C})$, 使得 $\mathbf{A}$ 的特征值为 $\lambda_1=1$, $\lambda_2=1$, $\lambda_3=2$, 对应的特征向量为

$$\mathbf{p}_1=\begin{pmatrix}1\\2\\1\end{pmatrix},\quad \mathbf{p}_2=\begin{pmatrix}1\\1\\0\end{pmatrix},\quad \mathbf{p}_3=\begin{pmatrix}2\\0\\-1\end{pmatrix}.$$

解. 设 $\mathbf{P}=\begin{pmatrix}\mathbf{p}_1 & \mathbf{p}_2 & \mathbf{p}_3\end{pmatrix}$, 由条件有 $\mathbf{AP}=\mathbf{P}\operatorname{diag}(1,1,2)$. 而

$$\left(\begin{array}{ccc|ccc}1 & 1 & 2 & 1 & 0 & 0\\ 2 & 1 & 0 & 0 & 1 & 0\\ 1 & 0 & -1 & 0 & 0 & 1\end{array}\right)\longrightarrow\left(\begin{array}{ccc|ccc}1 & 1 & 2 & 1 & 0 & 0\\ 0 & -1 & -4 & -2 & 1 & 0\\ 0 & 0 & 1 & 1 & -1 & 1\end{array}\right)\longrightarrow\left(\begin{array}{ccc|ccc}1 & 0 & 0 & 1 & -1 & 2\\ 0 & 1 & 0 & -2 & 3 & -4\\ 0 & 0 & 1 & 1 & -1 & 1\end{array}\right),$$

从而

$$\mathbf{A}=\begin{pmatrix}1 & 1 & 2\\ 2 & 1 & 0\\ 1 & 0 & -1\end{pmatrix}\begin{pmatrix}1 & & \\ & 1 & \\ & & 2\end{pmatrix}\begin{pmatrix}1 & -1 & 2\\ -2 & 3 & -4\\ 1 & -1 & 1\end{pmatrix}=\begin{pmatrix}3 & -2 & 2\\ 0 & 1 & 0\\ -1 & 1 & 0\end{pmatrix}.$$ □

例题 4.44. 设 $\mathbf{A}\in\mathsf{M}_n(\mathbb{C})$, 且 $\mathbf{A}^2=\mathbf{I}_n$, 证明 $\mathbf{A}$ 可对角化.

证明. 若 $\mathbf{v}$ 是特征值为 λ 的特征向量, 即 $\mathbf{Av}=\lambda\mathbf{v}$, 则 $\mathbf{v}=\mathbf{A}^2\mathbf{v}=\lambda^2\mathbf{v}$. 因为 $\mathbf{v}\neq\mathbf{0}$, 故 $\lambda^2=1$, 即 $\mathbf{A}$ 的特征值只能为 $\lambda_1=1$ 或 $\lambda_2=-1$. 要证明 $\mathbf{A}$ 可对角化, 只需要证明

$\dim V_{\lambda_1} + \dim V_{\lambda_2} = n$.

一方面, 对 $\forall \mathbf{v} \in \mathbb{C}^n$, $\mathbf{v} = \dfrac{\mathbf{I}_n + \mathbf{A}}{2}\mathbf{v} + \dfrac{\mathbf{I}_n - \mathbf{A}}{2}\mathbf{v} = \mathbf{v}_1 + \mathbf{v}_2$. 而 $\mathbf{A}\mathbf{v}_1 = \mathbf{v}_1$, $\mathbf{A}\mathbf{v}_2 = -\mathbf{v}_2$, 从而 $\mathbb{C}^n = V_{\lambda_1} + V_{\lambda_2}$. 另一方面, 若 $\mathbf{v} \in V_{\lambda_1} \cap V_{\lambda_2}$, 则 $\mathbf{v} = \mathbf{A}\mathbf{v} = -\mathbf{v}$, 只能 $\mathbf{v} = \mathbf{0}$, 即 $V_{\lambda_1} \cap V_{\lambda_2} = \{\mathbf{0}\}$. 从而由例题 2.31 中的维数公式知

$$\dim V_{\lambda_1} + \dim V_{\lambda_2} = \dim (V_{\lambda_1} + V_{\lambda_2}) + \dim V_{\lambda_1} \cap V_{\lambda_2} = n.$$

□

例题 4.45. 计算矩阵 $\mathbf{A} = \begin{pmatrix} a_1 & a_2 & \cdots & a_n \\ a_n & a_1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & a_2 \\ a_2 & \cdots & a_n & a_1 \end{pmatrix}$ 的行列式.

证明. 设 $\mathbf{Q} = \begin{pmatrix} 0 & 1 & \cdots & 0 \\ 0 & 0 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 1 \\ 1 & \cdots & 0 & 0 \end{pmatrix}$, 则 $\mathbf{A} = a_1\mathbf{I}_n + a_2\mathbf{Q} + \cdots + a_n\mathbf{Q}^{n-1}$. 因为

$$p_{\mathbf{Q}}(\lambda) = \begin{vmatrix} \lambda & -1 & \cdots & 0 \\ 0 & \lambda & \ddots & \vdots \\ \vdots & \ddots & \ddots & -1 \\ -1 & \cdots & 0 & \lambda \end{vmatrix} = \lambda^n - 1 = (\lambda - \epsilon_0)(\lambda - \epsilon_1)\cdots(\lambda - \epsilon_{n-1}),$$

其中 $\epsilon_0, \ldots, \epsilon_{n-1}$ 是所有 n 次单位根, 它们两两不同. 故 $\mathbf{Q}$ 可对角化, 即存在可逆矩阵 $\mathbf{P}$ 使得 $\mathbf{P}^{-1}\mathbf{Q}\mathbf{P} = \mathrm{diag}(\epsilon_0, \ldots, \epsilon_{n-1})$. 设 $f(x) = a_n x^{n-1} + \cdots + a_2 x + a_1$, 则

$$\det \mathbf{A} = \det \mathbf{P}^{-1}\mathbf{A}\mathbf{P} = \det \mathrm{diag}(f(\epsilon_0), \ldots, f(\epsilon_{n-1})) = f(\epsilon_0)\cdots f(\epsilon_{n-1}).$$

□

第 5 章　欧几里得空间

5.1　内积

基于向量的加法及数乘运算, 我们研究了向量空间 $\mathbb{F}^n$ 中的向量组的线性相关性, $\mathbb{F}^n$ 的子空间, 以及 $\mathbb{F}^n$ 上的线性变换. 为了研究 $\mathbb{F}^n$ 中的几何, 还需要定义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 $\mathbb{R}^n$ 中的几何问题.

在 $\mathbb{R}^3$ 中选取直角坐标系 $[O;\mathbf{i},\mathbf{j},\mathbf{k}]$, 若 $\mathbf{a}=[a_1,a_2,a_3]$, $\mathbf{b}=[b_1,b_2,b_3]$, 则

$$\mathbf{a}\cdot\mathbf{b} = a_1b_1+a_2b_2+a_3b_3.$$

在 $\mathbb{R}^n$ 中, 向量已经表示成坐标形式, 或者说已经固定了一组标准基 $\boldsymbol{\epsilon}_1,\ldots,\boldsymbol{\epsilon}_n$. 可以先给出内积在标准基下的公式, 然后通过内积定义向量的长度及向量间的夹角.

定义 5.1. 对 $\mathbb{R}^n$ 中两个向量 $\mathbf{a}=[a_1,\ldots,a_n]$, $\mathbf{b}=[b_1,\ldots,b_n]$,

$$\mathbf{a}\cdot\mathbf{b} = a_1b_1+\cdots+a_nb_n \tag{5.1}$$

称为 $\mathbf{a}$ 和 $\mathbf{b}$ 的点积, 也称为 $\mathbb{R}^n$ 上的标准内积.

在标准内积下, 因为 $\mathbf{a}\cdot\mathbf{a}\geq 0$ 对任意 $\mathbf{a}\in\mathbb{R}^n$ 成立, 可以定义向量 $\mathbf{a}$ 的长度为

$$|\mathbf{a}| = \sqrt{\mathbf{a}\cdot\mathbf{a}} = \sqrt{a_1^2+\cdots+a_n^2}. \tag{5.2}$$

从而 $\mathbf{a}\cdot\mathbf{a}=0 \iff \mathbf{a}=\mathbf{0}$, 故在这样的定义下只有零向量的长度为 0.

由柯西-施瓦茨不等式, $|\mathbf{a}\cdot\mathbf{b}|\leq|\mathbf{a}||\mathbf{b}|$ 对 $\forall\ \mathbf{a},\mathbf{b}\in\mathbb{R}^n$ 成立, 可以定义向量 $\mathbf{a}$, $\mathbf{b}$ 的夹角

的余弦为
$$\cos\theta = \frac{\mathbf{a}\cdot\mathbf{b}}{|\mathbf{a}||\mathbf{b}|} = \frac{a_1b_1+\cdots+a_nb_n}{\sqrt{a_1^2+\cdots+a_n^2}\sqrt{b_1^2+\cdots+b_n^2}}. \tag{5.3}$$
因为 $\mathbf{a}\cdot\mathbf{b}=\mathbf{b}\cdot\mathbf{a}$, 故 $\mathbf{a}$, $\mathbf{b}$ 的夹角和 $\mathbf{b}$, $\mathbf{a}$ 的夹角相同.

注意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.

评注 5.3. 由对称性, 内积对第二个分量也满足线性:

$$(\mathbf{a},\lambda\mathbf{b}_1+\mu\mathbf{b}_2)=(\lambda\mathbf{b}_1+\mu\mathbf{b}_2,\mathbf{a})=\lambda(\mathbf{b}_1,\mathbf{a})+\mu(\mathbf{b}_2,\mathbf{a})=\lambda(\mathbf{a},\mathbf{b}_1)+\mu(\mathbf{a},\mathbf{b}_2).$$

故内积满足 “双线性”.

设 $(\cdot,\cdot)$ 是欧式空间 $\mathbb{R}^n$ 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 5.4 (柯西-施瓦茨不等式)**.** $|(\mathbf{a},\mathbf{b})|^2\leq(\mathbf{a},\mathbf{a})(\mathbf{b},\mathbf{b})$, 等号成立 $\iff$ $\mathbf{a},\mathbf{b}$ 线性相关.

证明. 对 $\forall\ \lambda\in\mathbb{R}$,

$$0\leq(\mathbf{a}+\lambda\mathbf{b},\mathbf{a}+\lambda\mathbf{b})=(\mathbf{b},\mathbf{b})\lambda^2+2(\mathbf{a},\mathbf{b})\lambda+(\mathbf{a},\mathbf{a}),$$

故二次多项式 $(\mathbf{b},\mathbf{b})\lambda^2+2(\mathbf{a},\mathbf{b})\lambda+(\mathbf{a},\mathbf{a})$ 的判别式

$$\Delta=4|(\mathbf{a},\mathbf{b})|^2-4(\mathbf{b},\mathbf{b})(\mathbf{a},\mathbf{a})\leq 0,$$

等号成立当且仅当存在 $\lambda\in\mathbb{R}$ 使得 $(\mathbf{a}+\lambda\mathbf{b},\mathbf{a}+\lambda\mathbf{b})=0$, 即 $\mathbf{a},\mathbf{b}$ 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 5.5 (三角不等式)**.** 在欧式空间 $\mathbb{R}^n$ 中, $|\mathbf{a}+\mathbf{b}| \le |\mathbf{a}| + |\mathbf{b}|$.

证明. $|\mathbf{a}+\mathbf{b}|^2 = (\mathbf{a}+\mathbf{b}, \mathbf{a}+\mathbf{b}) = |\mathbf{a}|^2 + |\mathbf{b}|^2 + 2(\mathbf{a}, \mathbf{b}) \le (|\mathbf{a}| + |\mathbf{b}|)^2$. □

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

$$(\mathbf{a}, \mathbf{b}) = (\sum_{i=1}^n x_i \mathbf{v}_i, \sum_{j=1}^n y_j \mathbf{v}_j) = \sum_{1 \le i,j \le n} x_i y_j (\mathbf{v}_i, \mathbf{v}_j). \tag{5.4}$$

矩阵 $\mathbf{G} = ((\mathbf{v}_i, \mathbf{v}_j))_{n\times n}$ 称为 $(\cdot,\cdot)$ 在基 $\boldsymbol{\alpha}$ 下的度量矩阵. 且 (5.4) 可以写成

$$(\mathbf{a}, \mathbf{b}) = \mathbf{x}^{\mathrm{T}} \mathbf{G} \mathbf{y} = \begin{pmatrix} x_1 & \cdots & x_n \end{pmatrix} \begin{pmatrix} (\mathbf{v}_1, \mathbf{v}_1) & \cdots & (\mathbf{v}_1, \mathbf{v}_n) \\ \vdots & \ddots & \vdots \\ (\mathbf{v}_n, \mathbf{v}_1) & \cdots & (\mathbf{v}_n, \mathbf{v}_n) \end{pmatrix} \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}. \tag{5.5}$$

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

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

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 5.7. 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.

证明. 若 $\lambda_1 \mathbf{a}_1 + \cdots + \lambda_r \mathbf{a}_r = \mathbf{0}$, 则对 $\forall\ 1 \le i \le r$,

$$0 = (\lambda_1 \mathbf{a}_1 + \cdots + \lambda_i \mathbf{a}_i + \cdots + \lambda_r \mathbf{a}_r, \mathbf{a}_i) = \lambda_i |\mathbf{a}_i|^2 \Longrightarrow \lambda_i = 0,$$

从而 $\mathbf{a}_1, \ldots, \mathbf{a}_r$ 线性无关. □

定理 5.8 (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.

证明. (1) 取 $\mathbf{u}_1 = \mathbf{v}_1 / |\mathbf{v}_1|$, 则 $(\mathbf{u}_1, \mathbf{u}_1) = 1$.
(2) 令 $\mathbf{w}_2 = \mathbf{v}_2 - (\mathbf{v}_2, \mathbf{u}_1)\mathbf{u}_1$, 一方面

$$\mathbf{w}_2 = \mathbf{v}_2 - \frac{(\mathbf{v}_1, \mathbf{v}_1)\mathbf{v}_1}{|\mathbf{v}_1|^2},$$

由 $\mathbf{v}_1, \mathbf{v}_2$ 线性无关知 $\mathbf{w}_2 \neq 0$. 另一方面

$$(\mathbf{w}_2, \mathbf{u}_1) = (\mathbf{v}_2 - (\mathbf{v}_2, \mathbf{u}_1)\mathbf{u}_1, \mathbf{u}_1) = (\mathbf{v}_2, \mathbf{u}_1) - (\mathbf{v}_2, \mathbf{u}_1)(\mathbf{u}_1, \mathbf{u}_1) = 0.$$

取 $\mathbf{u}_2 = \mathbf{w}_2/|\mathbf{w}_2|$, 则 $(\mathbf{u}_2, \mathbf{u}_2) = 1$, $(\mathbf{u}_2, \mathbf{u}_1) = 0$, 且 $\mathbf{u}_2$ 是 $\mathbf{v}_1, \mathbf{v}_2$ 的线性组合.
(3) 假设已从 $\mathbf{v}_1, \ldots, \mathbf{v}_k$ 出发构造了 $\mathbf{u}_1, \ldots, \mathbf{u}_k$, 满足对 $\forall\ 1 \leq i \leq j \leq k$, $(\mathbf{u}_i, \mathbf{u}_j) = \delta_{ij}$, 且对 $\forall\ 1 \leq i \leq k$, $\mathbf{u}_i$ 是 $\mathbf{v}_1, \ldots, \mathbf{v}_i$ 的线性组合. 令

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

则一方面由 $\mathbf{v}_1, \ldots, \mathbf{v}_{k+1}$ 线性无关知 $\mathbf{w}_{k+1} \neq \mathbf{0}$. 另一方面对 $\forall\ 1 \leq i \leq k$, $(\mathbf{w}_{k+1}, \mathbf{u}_i) = 0$. 取 $\mathbf{u}_{k+1} = \mathbf{w}_{k+1}/|\mathbf{w}_{k+1}|$, 则对 $\forall\ 1 \leq i \leq j \leq k+1$, $(\mathbf{u}_i, \mathbf{u}_j) = \delta_{ij}$, 且对 $\forall\ 1 \leq i \leq k+1$, $\mathbf{u}_i$ 是 $\mathbf{v}_1, \ldots, \mathbf{v}_i$ 的线性组合.

由此就构造了 $\mathbf{u}_1, \ldots, \mathbf{u}_n$, 满足 $(\mathbf{u}_i, \mathbf{u}_j) = \delta_{ij}$ $(\forall\ 1 \leq i \leq j \leq n)$. 由命题 5.7 知 $\mathbf{u}_1, \ldots, \mathbf{u}_n$ 线性无关, 是 $\mathbb{R}^n$ 的一组标准正交基. □

推论 5.9. 欧式空间 $\mathbb{R}^n$ 中的一组两两正交的单位向量 $\mathbf{u}_1, \ldots, \mathbf{u}_k$ 可以扩充为 $\mathbb{R}^n$ 的一组标准正交基 $\mathbf{u}_1, \ldots, \mathbf{u}_k, \mathbf{u}_{k+1}, \ldots, \mathbf{u}_n$.

证明. 由命题 5.7 知 $\mathbf{u}_1, \ldots, \mathbf{u}_k$ 线性无关, 故可以扩充为 $\mathbb{R}^n$ 的一组基

$$\mathbf{u}_1, \ldots, \mathbf{u}_k, \mathbf{v}_{k+1}, \ldots, \mathbf{v}_n,$$

将这组基正交化就得到一组标准正交基 $\mathbf{u}_1, \ldots, \mathbf{u}_k, \mathbf{u}_{k+1}, \ldots, \mathbf{u}_n$. □

例题 5.10. 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.

解. 令 $\mathbf{u}_1 = \mathbf{v}_1/|\mathbf{v}_1| = (\frac{1}{\sqrt{2}}, \frac{-1}{\sqrt{2}}, 0)$, 以及

$$\mathbf{w}_2 = \mathbf{v}_2 - (\mathbf{v}_2, \mathbf{u}_1)\mathbf{u}_1 = (\frac{1}{2}, \frac{1}{2}, -1), \quad \mathbf{u}_2 = \mathbf{w}_2/|\mathbf{w}_2| = (\frac{1}{\sqrt{6}}, \frac{1}{\sqrt{6}}, \frac{-2}{\sqrt{6}}).$$

$$\mathbf{w}_3 = \mathbf{v}_3 - (\mathbf{v}_3, \mathbf{u}_1)\mathbf{u}_1 - (\mathbf{v}_3, \mathbf{u}_2)\mathbf{u}_2 = (\frac{2}{3}, \frac{2}{3}, \frac{2}{3}), \quad \mathbf{u}_3 = \mathbf{w}_3/|\mathbf{w}_3| = (\frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}).$$

则 $\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3$ 构成 $\mathbb{R}^3$ 在标准内积下的一组标准正交基. □

命题 5.11. 设 $\mathbf{u}_1, \ldots, \mathbf{u}_n$ 是欧式空间 $\mathbb{R}^n$ 的一组标准正交基, 则对 $\forall\ \mathbf{v} \in V$,
(1) $\mathbf{v} = (\mathbf{v}, \mathbf{u}_1)\mathbf{u}_1 + \cdots + (\mathbf{v}, \mathbf{u}_n)\mathbf{u}_n$.
(2) $|\mathbf{v}|^2 = (\mathbf{v}, \mathbf{u}_1)^2 + \cdots + (\mathbf{v}, \mathbf{u}_n)^2$.

证明. (1) 因为 $\mathbf{u}_1,\ldots,\mathbf{u}_i$ 是一组基, 故存在 $\lambda_1,\ldots,\lambda_n\in\mathbb{R}$ 使得 $\mathbf{v}=\lambda_1\mathbf{u}_1+\cdots+\lambda_n\mathbf{u}_n$. 对 $\forall\ 1\le i\le n$,

$$(\mathbf{v},\mathbf{u}_i)=(\lambda_1\mathbf{u}_1+\cdots+\lambda_i\mathbf{u}_i+\cdots+\lambda_n\mathbf{u}_n,\mathbf{u}_i)=\lambda_i.$$

(2) 由 (1) 知 $(\mathbf{v},\mathbf{v})=(\mathbf{v},(\mathbf{v},\mathbf{u}_1)\mathbf{u}_1+\cdots+(\mathbf{v},\mathbf{u}_n)\mathbf{u}_n)=(\mathbf{v},\mathbf{u}_1)^2+\cdots+(\mathbf{v},\mathbf{u}_n)^2$. □

命题 5.12. 设欧式空间 $\mathbb{R}^n$ 的内积在一组基 $\boldsymbol{\alpha}=\{\mathbf{v}_1,\ldots,\mathbf{v}_n\}$ 下的度量矩阵为 $\mathbf{G}$, 则存在可逆矩阵 $\mathbf{P}$ 使得 $\mathbf{G}=\mathbf{P}^{\mathrm{T}}\mathbf{P}$.

证明. 取 $\mathbb{R}^n$ 的一组标准正交基 $\boldsymbol{\eta}=\{\mathbf{u}_1,\ldots,\mathbf{u}_n\}$, $\mathbf{P}=(p_{ij})_{n\times n}$ 是从 $\boldsymbol{\eta}$ 到 $\boldsymbol{\alpha}$ 的过渡矩阵:

$$\begin{pmatrix}\mathbf{v}_1 & \cdots & \mathbf{v}_n\end{pmatrix}=\begin{pmatrix}\mathbf{u}_1 & \cdots & \mathbf{u}_n\end{pmatrix}\begin{pmatrix}p_{11} & \cdots & p_{1n}\\ \vdots & \ddots & \vdots\\ p_{n1} & \cdots & p_{nn}\end{pmatrix}.$$

则 $(\mathbf{v}_i,\mathbf{v}_j)=(\sum_{k=1}^n p_{ki}\mathbf{u}_k,\sum_{\ell=1}^n p_{\ell j}\mathbf{u}_\ell)=\sum_{k=1}^n p_{ki}p_{kj}$, 即 $\mathbf{G}=((\mathbf{v}_i,\mathbf{v}_j))_{n\times n}=\mathbf{P}^{\mathrm{T}}\mathbf{P}$. □

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

$$(\mathcal{A}(\mathbf{v}),\mathcal{A}(\mathbf{w}))=(\mathbf{v},\mathbf{w}),\quad\forall\ \mathbf{v},\mathbf{w}\in\mathbb{R}^n \tag{5.6}$$

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

我们考察正交变换所满足的性质. 首先在 (5.6) 中令 $\mathbf{v}=\mathbf{w}$, 得到

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

即正交变换保持向量长度. 因此正交变换也称为保距变换. 另一方面, 取 $\mathbb{R}^n$ 的一组标准正交基 $\mathbf{u}_1,\ldots,\mathbf{u}_n$, 由 (5.6) 知

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

从而 $\mathcal{A}(\mathbf{u}_1),\ldots,\mathcal{A}(\mathbf{u}_n)$ 也是 $\mathbb{R}^n$ 的一组标准正交基, 即正交变换将一组标准正交基映为一组标准正交基.

命题 5.14. 正交变换 $\mathcal{A}$ 在一组标准正交基 $\boldsymbol{\eta} = \{\mathbf{u}_1, \dots, \mathbf{u}_n\}$ 下的矩阵为 $\mathbf{A}$:

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

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

证明. 设 $\mathbf{A} = (a_{ij})_{n\times n}$, 则

$$\delta_{ij} = (\mathcal{A}(\mathbf{u}_i), \mathcal{A}(\mathbf{u}_j)) = (\sum_{k=1}^{n} a_{ki}\mathbf{u}_k, \sum_{\ell=1}^{n} a_{\ell j}\mathbf{u}_\ell) = \sum_{k=1}^{n} a_{ki}a_{kj},$$

即 $\mathbf{A}^{\mathrm{T}}\mathbf{A} = \mathbf{I}_n$. □

定义 5.15. $\mathsf{M}_n(\mathbb{R})$ 中满足 $\mathbf{A}^{\mathrm{T}}\mathbf{A} = \mathbf{I}_n$ 的方阵称为正交方阵.

设 $\boldsymbol{\eta}_1 = \{\mathbf{v}_1, \dots, \mathbf{v}_n\}$, $\boldsymbol{\eta}_2 = \{\mathbf{w}_1, \dots, \mathbf{w}_n\}$ 是欧式空间 $\mathbb{R}^n$ 中的两组标准正交基, 和命题 5.14 的证明类似, 可知从 $\boldsymbol{\eta}_1$ 到 $\boldsymbol{\eta}_2$ 的过渡矩阵是正交方阵. 从而若 $\mathcal{A}$ 是欧式空间 $\mathbb{R}^n$ 上的线性变换, 且在一组标准正交基下的矩阵为 $\mathbf{A}$, 则 $\mathcal{A}$ 在另一组标准正交基下的矩阵为 $\mathbf{O}^{-1}\mathbf{A}\mathbf{O}$, 其中 $\mathbf{O}$ 是正交方阵.

定义 5.16. 设 $\mathbf{A}, \mathbf{B} \in \mathsf{M}_n(\mathbb{R})$, 若存在正交方阵 $\mathbf{O} \in \mathsf{M}_n(\mathbb{R})$ 使得 $\mathbf{O}^{-1}\mathbf{A}\mathbf{O} = \mathbf{B}$, 则称 $\mathbf{A}$ 和 $\mathbf{B}$ 正交相似.

定理 5.17. 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) $|\mathcal{A}(\mathbf{v})| = |\mathbf{v}|$ 对 $\forall\ \mathbf{v} \in \mathbb{R}^n$ 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.

证明. 我们已经证明了 (1) $\Longrightarrow$ (2), (3) $\Longrightarrow$ (4).
(2) $\Longrightarrow$ (3): 一方面, 对 $\forall\ 1 \le i \le n$, $|\mathcal{A}(\mathbf{u}_i)| = |\mathbf{u}_i| = 1$. 另一方面, 对 $\forall\ 1 \le i < j \le n$,

$$|\mathcal{A}(\mathbf{u}_i + \mathbf{u}_j)| = |\mathbf{u}_i + \mathbf{u}_j| \Longrightarrow (\mathcal{A}(\mathbf{u}_i), \mathcal{A}(\mathbf{u}_j)) = 0.$$

故 $\mathcal{A}(\mathbf{u}_1), \dots, \mathcal{A}(\mathbf{u}_n)$ 构成一组标准正交基.
(4) $\Longrightarrow$ (1): 设 $\mathcal{A}$ 在标准正交基 $\mathbf{u}_1, \dots, \mathbf{u}_n$ 下的矩阵为正交方阵, 则

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

于是对 $\mathbf{v}=\sum_{k=1}^n a_k\mathbf{u}_k$, $\mathbf{w}=\sum_{\ell=1}^n b_\ell\mathbf{u}_\ell$,

$$\begin{aligned}(\mathcal{A}(\mathbf{v}),\mathcal{A}(\mathbf{w})) &= (\sum_{k=1}^n a_k\mathcal{A}(\mathbf{u}_k),\sum_{\ell=1}^n b_\ell\mathcal{A}(\mathbf{u}_\ell)) = \sum_{k,\ell=1}^n a_kb_\ell(\mathcal{A}(\mathbf{u}_i),\mathcal{A}(\mathbf{u}_j)) \\ &= \sum_{k,\ell=1}^n a_kb_\ell(\mathbf{u}_i,\mathbf{u}_j) = (\sum_{k=1}^n a_k\mathbf{u}_k,\sum_{\ell=1}^n b_\ell\mathbf{u}_\ell) = (\mathbf{v},\mathbf{w}).\end{aligned}$$

这样就完成了证明. □

例题 5.18. 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.

证明. 设 $\mathbf{A}=\begin{pmatrix}a & b\\ c & d\end{pmatrix}$, 则由 $\mathbf{A}^{\mathrm{T}}\mathbf{A}=\mathbf{I}_2$ 得到

$$a^2+c^2=1,\quad ,b^2+d^2=1,\quad ab+cd=0.$$

设 $a=\cos\alpha$, $c=\sin\alpha$, $b=\cos\beta$, $d=\cos\beta$, 则

$$ab+cd=0\Longrightarrow\sin(\alpha+\beta)=0\Longrightarrow\alpha+\beta=2n\pi \text{ 或 } (2n+1)\pi,\quad n\in\mathbb{Z}.$$

若 $\alpha+\beta=2n\pi$, 则 $b=-\sin\alpha$, $d=\cos\alpha$, $\mathbf{A}=\begin{pmatrix}\cos\alpha & -\sin\alpha\\ \sin\alpha & \cos\alpha\end{pmatrix}$.

若 $\alpha+\beta=(2n+1)\pi$, 则 $b=\sin\alpha$, $d=-\cos\alpha$, $\mathbf{A}=\begin{pmatrix}\cos\alpha & \sin\alpha\\ \sin\alpha & -\cos\alpha\end{pmatrix}$. □

命题 5.19. 设 $\mathbf{A}\in\mathsf{M}_n(\mathbb{R})$ 是正交方阵, 则

(1) $\det\mathbf{A}=\pm1$.

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

证明. (1) $\mathbf{A}^{\mathrm{T}}\mathbf{A}=\mathbf{I}_n\Longrightarrow|\mathbf{A}|^2=1\Longrightarrow|\mathbf{A}|=\pm1$.

(2) 设 $\mathbf{A}\mathbf{x}=\lambda\mathbf{x}$, 其中 $\mathbf{x}\in\mathbb{C}^n$ 是特征向量. 则取共轭转置得到 $\bar{\mathbf{x}}^{\mathrm{T}}\mathbf{A}^{\mathrm{T}}=\bar{\lambda}\bar{\mathbf{x}}^{\mathrm{T}}$, 从而

$$\bar{\mathbf{x}}^{\mathrm{T}}\mathbf{x}=\bar{\mathbf{x}}\mathbf{A}^{\mathrm{T}}\mathbf{A}\mathbf{x}=|\lambda|^2\bar{\mathbf{x}}^{\mathrm{T}}\mathbf{x}.$$

因为 $\mathbf{x}$ 是非零向量, $\bar{\mathbf{x}}^{\mathrm{T}}\mathbf{x}\neq0$, 故 $|\lambda|=1$. □

评注 5.20. 这里需要注意正交方阵 $\mathbf{A} \in \mathsf{M}_n(\mathbb{R})$, 但在求特征值及特征向量时需要将 $\mathbf{A}$ 看成 $\mathsf{M}_n(\mathbb{C})$ 中方阵. 换句话说, $\mathsf{M}_n(\mathbb{R})$ 中方阵的特征值不一定是实数, 特征向量不一定是 $\mathbb{R}^n$ 中向量.

例题 5.21. 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 $\det \mathbf{A} = 1$. 证明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.

证明. 设 $p_{\mathbf{A}}(\lambda) = |\lambda \mathbf{I}_3 - \mathbf{A}| = \lambda^3 + a_2\lambda^2 + a_1\lambda + a_0$, $a_0, a_1, a_2 \in \mathbb{R}$. 因为实系数 3 次方程必有实根, 故可设 $\mathbf{A}$ 的特征值为 $\lambda_1, \lambda_2, \lambda_3$, 其中 $\lambda_1 \in \mathbb{R}$.

若 $\lambda_2 \in \mathbb{R}$, 则 $\lambda_3 \in \mathbb{R}$. 由命题 5.19 知 $\lambda_i = \pm 1$. 又 $\det \mathbf{A} = 1$, 故 $\mathbf{A}$ 的特征值为 $1, 1, 1$ 或 $1, -1, -1$. 若 $\lambda_2 \notin \mathbb{R}$, 则 $\lambda_3 = \bar{\lambda}_2$, 于是

$$1 = \det \mathbf{A} = \lambda_1\lambda_2\lambda_3 = \lambda_1|\lambda_2|^2 = \lambda_1.$$

于是 $\mathcal{A}$ 的特征值必有 1.

设 $\mathbf{v}_1$ 是 $\mathcal{A}$ 的特征值为 1 的特征向量, 且 $|\mathbf{v}_1| = 1$, 由推论 5.9 可以扩充为 $\mathbb{R}^n$ 的一组标准正交基 $\boldsymbol{\eta} = \{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$. 则 $\mathcal{A}$ 在 $\boldsymbol{\eta}$ 下的矩阵为 $\mathbf{B} = \begin{pmatrix} 1 & \mathbf{A}_1 \\ \mathbf{0} & \mathbf{A}_2 \end{pmatrix}$, 其中 $\mathbf{A}_1 \in \mathsf{M}_{1\times 2}(\mathbb{R})$, $\mathbf{A}_2 \in \mathsf{M}_{2\times 2}(\mathbb{R})$. 因为 $\mathcal{A}$ 是正交变换, 故 $\mathbf{B}$ 是正交方阵, 于是

$$\begin{aligned} \mathbf{B}^{\mathrm{T}}\mathbf{B} = \mathbf{I}_3 &\Longrightarrow \begin{pmatrix} 1 & \mathbf{0} \\ \mathbf{A}_1^{\mathrm{T}} & \mathbf{A}_2^{\mathrm{T}} \end{pmatrix} \begin{pmatrix} 1 & \mathbf{A}_1 \\ \mathbf{0} & \mathbf{A}_2 \end{pmatrix} = \begin{pmatrix} 1 & \mathbf{0} \\ \mathbf{0} & \mathbf{I}_2 \end{pmatrix} \\ &\Longrightarrow \begin{pmatrix} 1 & \mathbf{A}_1 \\ \mathbf{A}_1^{\mathrm{T}} & \mathbf{A}_1^{\mathrm{T}}\mathbf{A}_1 + \mathbf{A}_2^{\mathrm{T}}\mathbf{A}_2 \end{pmatrix} = \begin{pmatrix} 1 & \mathbf{0} \\ \mathbf{0} & \mathbf{I}_2 \end{pmatrix} \Longrightarrow \mathbf{A}_1 = \mathbf{0},\ \mathbf{A}_2^{\mathrm{T}}\mathbf{A}_2 = \mathbf{I}_2. \end{aligned}$$

从而 $\mathbf{A}_2$ 是二阶正交方阵. 因为 $\mathbf{A}$ 和 $\mathbf{B}$ 都是 $\mathcal{A}$ 在某组基下的矩阵, $\mathbf{A}$ 和 $\mathbf{B}$ 相似, 故 $\det \mathbf{B} = \det \mathbf{A} = 1$, 于是 $\det \mathbf{A}_2 = 1$. 由例题 5.18 知 $\mathbf{A}_2 = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$, 从而 $\mathcal{A}$ 在 $\boldsymbol{\eta}$ 这组标准正交基下的矩阵为

$$\mathbf{B} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta \\ 0 & \sin\theta & \cos\theta \end{pmatrix}.$$

可以看出 $\mathcal{A}$ 固定以 $\mathbf{v}_1$ 为方向向量的直线 L, 且在与 L 垂直的平面 π 上的作用是以 L 与 π 交点为原点的旋转, 故是空间中以 L 为轴的旋转变换. □

定义 5.22. 设 $\mathcal{A}$ 是欧式空间 $\mathbb{R}^n$ 上的线性变换, 且满足

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

则称 $\mathcal{A}$ 为对称变换.

定理 5.23. 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.

证明. "$\Longrightarrow$": 设 $\mathbf{v}_1, \ldots, \mathbf{v}_n$ 是 $\mathbb{R}^n$ 的一组标准正交基, 对称变换 $\mathcal{A}$ 在这组基下的矩阵是 $\mathbf{A} = (a_{ij})_{n\times n}$:

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

则对 $\forall\ 1 \le i, j \le n$,

$$a_{ji} = (\sum_{k=1}^{n} a_{ki}\mathbf{v}_k, \mathbf{v}_j) = (\mathcal{A}(\mathbf{v}_i), \mathbf{v}_j) = (\mathbf{v}_i, \mathcal{A}(\mathbf{v}_j)) = (\mathbf{v}_i, \sum_{\ell=1}^{n} a_{\ell j}\mathbf{v}_\ell) = a_{ij}, \tag{5.7}$$

从而 $\mathbf{A}$ 是实对称矩阵.

"$\Longleftarrow$": 由 (5.7) 知若 $\mathcal{A}$ 在一组标准正交基下的矩阵是实对称矩阵, 则对 $\forall\ 1 \le i, j \le n$, $(\mathcal{A}(\mathbf{v}_i), \mathbf{v}_j) = (\mathbf{v}_i, \mathcal{A}(\mathbf{v}_j))$. 从而对 $\mathbf{v} = \sum_{k=1}^n a_k \mathbf{v}_k$, $\mathbf{w} = \sum_{\ell=1}^n b_\ell \mathbf{v}_\ell$,

$$\begin{aligned}(\mathcal{A}(\mathbf{v}), \mathbf{w}) =& (\sum_{k=1}^{n} a_k \mathcal{A}(\mathbf{v}_k), \sum_{\ell=1}^{n} b_\ell \mathbf{v}_\ell) = \sum_{k,\ell=1}^{n} a_k b_\ell (\mathcal{A}(\mathbf{v}_k), \mathbf{v}_\ell) \\ &= \sum_{k,\ell=1}^{n} a_k b_\ell (\mathbf{v}_k, \mathcal{A}(\mathbf{v}_\ell)) = (\sum_{k=1}^{n} a_k \mathbf{v}_k, \sum_{\ell=1}^{n} b_\ell \mathcal{A}(\mathbf{v}_\ell)) = (\mathbf{v}, \mathcal{A}(\mathbf{w})),\end{aligned}$$

即 $\mathcal{A}$ 是对称变换. □

命题 5.24. 设 $\mathcal{A}$ 是欧式空间 $\mathbb{R}^n$ 上的对称变换, 则 $\mathcal{A}$ 的特征值都是实数.

证明. 对矩阵 $\mathbf{A} = (a_{ij})_{n\times n} \in \mathsf{M}_n(\mathbb{C})$, 记 $\bar{\mathbf{A}} = (\bar{a}_{ij})_{n\times n}$, 其中 $\bar{a}_{ij}$ 是 a_{ij} 的共轭. 设对称变换 $\mathcal{A}$ 在一组标准正交基下的矩阵为 $\mathbf{A}$, 则 $\mathbf{A}$ 是实对称矩阵, 只需证明 $\mathbf{A}$ 的特征值都是实数:

$$\begin{gathered}\mathbf{A}\mathbf{v} = \lambda\mathbf{v} \Longrightarrow \mathbf{v}^{\mathrm{T}}\mathbf{A} = \lambda\mathbf{v}^{\mathrm{T}} \Longrightarrow \bar{\mathbf{v}}^{\mathrm{T}}\mathbf{A} = \bar{\lambda}\bar{\mathbf{v}}^{\mathrm{T}} \Longrightarrow \\ \bar{\mathbf{v}}^{\mathrm{T}}\mathbf{A}\mathbf{v} = \bar{\lambda}\bar{\mathbf{v}}^{\mathrm{T}}\mathbf{v} \Longrightarrow \lambda\bar{\mathbf{v}}^{\mathrm{T}}\mathbf{v} = \bar{\lambda}\bar{\mathbf{v}}^{\mathrm{T}}\mathbf{v} \Longrightarrow \lambda = \bar{\lambda}.\end{gathered}$$

于是 λ 是实数. □

命题 5.25. 设 $\mathcal{A}$ 是欧式空间 $\mathbb{R}^n$ 上的对称变换, $\mathbf{v}_1, \mathbf{v}_2$ 分别是特征值为 λ_1, λ_2 的特征向量, 且 $\lambda_1 \neq \lambda_2$, 则 $(\mathbf{v}_1, \mathbf{v}_2) = 0$.

证明. $\lambda_1(\mathbf{v}_1, \mathbf{v}_2) = (\mathcal{A}(\mathbf{v}_1), \mathbf{v}_2) = (\mathbf{v}_1, \mathcal{A}(\mathbf{v}_2)) = \lambda_2(\mathbf{v}_1, \mathbf{v}_2)$, 由 $\lambda_1 \neq \lambda_2$ 知 $(\mathbf{v}_1, \mathbf{v}_2) = 0$. □

定理 5.26. 设 $\mathcal{A}$ 是欧式空间 $\mathbb{R}^n$ 上对称变换, 则存在一组标准正交基 $\boldsymbol{\eta}$, $\mathcal{A}$ 在 $\boldsymbol{\eta}$ 下的矩阵是对角矩阵 $\mathrm{diag}(\lambda_1, \dots, \lambda_n)$. 即实对称方阵正交相似于对角矩阵.

证明. 对维数 n 用归纳法, 假设欧式空间 $\mathbb{R}^{n-1}$ 上的对称变换都可对角化，或者等价的，任何 $n-1$ 阶实对称矩阵可正交相似于对角矩阵.

设 $\mathcal{A}$ 是欧式空间 $\mathbb{R}^n$ 上的对称变换, 取 $\mathcal{A}$ 的一个特征值 λ_1, 以及特征值为 λ_1 特征向量 $\mathbf{v}_1$, 且 $|\mathbf{v}_1| = 1$. 将 $\mathbf{v}_1$ 扩充为 $\mathbb{R}^n$ 的一组标准正交基 $\boldsymbol{\eta} = \{\mathbf{v}_1, \dots, \mathbf{v}_n\}$, 则 $\mathcal{A}$ 在 $\boldsymbol{\eta}$ 下的矩阵为

$$\mathbf{B} = \begin{pmatrix} \lambda_1 & \mathbf{A}_2 \\ \mathbf{0} & \mathbf{A}_1 \end{pmatrix}, \quad \mathbf{A}_1 \in \mathsf{M}_{n-1}(\mathbb{R}), \mathbf{A}_2 \in \mathsf{M}_{1\times(n-1)}(\mathbb{R}).$$

因为 $\mathcal{A}$ 是对称变换, $\boldsymbol{\eta}$ 是标准正交基, 则由定理 5.23 知 $\mathbf{B}$ 是实对称矩阵, 只能 $\mathbf{A}_2 = \mathbf{0}$, $\mathbf{A}_1^{\mathrm{T}} = \mathbf{A}_1$. 由归纳假设, 存在正交方阵 $\mathbf{O}_1 \in \mathsf{M}_{n-1}(\mathbb{R})$ 使得

$$\mathbf{O}_1^{-1}\mathbf{A}_1\mathbf{O} = \mathrm{diag}(\lambda_2, \dots, \lambda_n), \quad \lambda_2, \dots, \lambda_n \in \mathbb{R}.$$

设 $\mathbf{O} = \begin{pmatrix} 1 & \\ & \mathbf{O}_1 \end{pmatrix}$, 则 $\mathbf{O}$ 是正交方阵, 且 $\mathbf{O}^{-1}\mathbf{B}\mathbf{O} = \mathrm{diag}(\lambda_1, \dots, \lambda_n)$, 即 $\mathcal{A}$ 在标准正交基

$$\begin{pmatrix} \mathbf{u}_1 & \cdots & \mathbf{u}_n \end{pmatrix} = \begin{pmatrix} \mathbf{v}_1 & \cdots & \mathbf{v}_n \end{pmatrix} \cdot \mathbf{O}$$

下的矩阵是对角矩阵. □

设 $\mathbf{A}$ 是实对称矩阵, 且正交方阵 $\mathbf{O}$ 满足 $\mathbf{O}^{-1}\mathbf{A}\mathbf{O} = \mathrm{diag}(\lambda_1, \dots, \lambda_n)$. 设 $\mathbf{O}$ 的行向量形式为 $\mathbf{O} = \begin{pmatrix} \mathbf{u}_1 & \cdots & \mathbf{u}_n \end{pmatrix}$, 则

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

即 $\mathbf{u}_i$ 是 $\mathbf{A}$ 的关于特征值 λ_i 的特征向量, 且 $\mathbf{u}_1, \dots, \mathbf{u}_n$ 构成 $\mathbb{R}^n$ 的一组标准正交基.

例题 5.27. 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.

证明. $p_{\mathbf{A}}(\lambda) = |\lambda \mathbf{I}_3 - \mathbf{A}| = (\lambda+1)^2(\lambda-5)$, 故 $\mathbf{A}$ 的特征值为 $\lambda_1 = -1$, $\lambda_2 = 5$. 可知其特征子空间为

$$V_{\lambda_1} = \left\{ \lambda \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix} + \mu \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix},\ \lambda, \mu \in \mathbb{R} \right\}, \quad V_{\lambda_2} = \left\{ \lambda \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix},\ \lambda \in \mathbb{R} \right\}.$$

利用格拉姆-施密特正交化过程可以求出 V_{λ_1} 的一组标准正交基 $\mathbf{u}_1 = [\frac{1}{\sqrt{2}}, \frac{-1}{\sqrt{2}}, 0]$, $\mathbf{u}_2 = [\frac{1}{\sqrt{6}}, \frac{1}{\sqrt{6}}, \frac{-2}{\sqrt{6}}]$, 以及 V_{λ_2} 的一组标准正交基 $\mathbf{u}_3 = [\frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}]$. 令

$$\mathbf{O} = \begin{pmatrix} \mathbf{u}_1 & \mathbf{u}_2 & \mathbf{u}_3 \end{pmatrix} = \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{3}} \\ \frac{-1}{\sqrt{2}} & \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{3}} \\ 0 & \frac{-2}{\sqrt{6}} & \frac{1}{\sqrt{3}} \end{pmatrix},$$

则 $\mathbf{O}^{-1}\mathbf{A}\mathbf{O} = \operatorname{diag}(-1, -1, 5)$. □

5.3 二次型

实对称矩阵的相似对角化的一个应用是研究二次型的分类.

定义 5.28. 实数域 $\mathbb{R}$ 上含有 n 个变元的二次齐次多项式

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

称为一个 n 阶（实）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 $Q(\mathbf{x})$, 则二次型可以写成

$$\mathbf{x}^{\mathrm{T}}\mathbf{A}\mathbf{x} = \begin{pmatrix} x_1 & \cdots & x_n \end{pmatrix} \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix},$$

其中 $\mathbf{A}=(a_{ij})_{n\times n}$ 是实对称矩阵, 称为二次型 $\mathbf{Q}(\mathbf{x})$ 的矩阵, 满足

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

例题 5.29. 写出二次型 $Q_1(\mathbf{x})=x_1x_2+x_2x_3+x_1x_3$, 及 $Q_2(\mathbf{x})=x_1^2+2x_2^2+3x_3^2$ 的矩阵.

证明. $Q_1(\mathbf{x})$ 的矩阵是 $\begin{pmatrix}0&\frac12&\frac12\\ \frac12&0&\frac12\\ \frac12&\frac12&0\end{pmatrix}$, $Q_2(\mathbf{x})$ 的矩阵是 $\begin{pmatrix}1&0&0\\0&2&0\\0&0&3\end{pmatrix}$. □

设 $(\cdot,\cdot)$ 是 $\mathbb{R}^n$ 上的内积, $\mathbf{G}$ 是内积在标准基下的度量矩阵. 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 5.30. 设 $Q(\mathbf{x})$ 是 n 阶二次型.

1. 若 $Q(\mathbf{x})>0$ ($Q(\mathbf{x})\ge 0$) 对 $\forall\ \mathbf{x}\neq\mathbf{0}$ 成立, 称 $Q(\mathbf{x})$ 是正定的 (半正定的).

2. 若 $Q(\mathbf{x})<0$ ($Q(\mathbf{x})\le 0$) 对 $\forall\ \mathbf{x}\neq\mathbf{0}$ 成立, 称 $Q(\mathbf{x})$ 是负定的 (半负定的).

3. 若既存在 $\mathbf{x}_1$ 使得 $Q(\mathbf{x}_1)>0$, 又存在 $\mathbf{x}_2$ 使得 $Q(\mathbf{x}_2)<0$, 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 $\mathbf{A}$ 是实对称矩阵, 由定理 5.26, 存在正交方阵 $\mathbf{O}$ 使得 $\mathbf{O}^{\mathrm{T}}\mathbf{A}\mathbf{O}=\mathrm{diag}(\lambda_1,\dots,\lambda_n)$, 其中 $\lambda_1,\dots,\lambda_n\in\mathbb{R}$, 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

$$Q(\mathbf{x})=\mathbf{x}^{\mathrm{T}}\mathbf{A}\mathbf{x}=\mathbf{y}^{\mathrm{T}}(\mathbf{O}^{\mathrm{T}}\mathbf{A}\mathbf{O})\mathbf{y}=\lambda_1y_1^2+\cdots+\lambda_ny_n^2.$$

因为 $\mathbf{O}$ 是可逆矩阵, 故

$$\varphi_{\mathbf{O}}\colon\mathbb{R}^n\longrightarrow\mathbb{R}^n,\quad \varphi_{\mathbf{O}}(\mathbf{y})=\mathbf{O}\mathbf{y}=\mathbf{x},$$

是从 $\mathbb{R}^n$ 到自身的一一映射. 从而 $Q(\mathbf{x})$ 的值域和二次型

$$\tilde{Q}(\mathbf{y})=\lambda_1y_1^2+\cdots+\lambda_ny_n^2$$

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. 求正交方阵 $\mathbf{O}$ 使得二次型 $\mathbf{x}^{\mathrm{T}}\mathbf{A}\mathbf{x}$ 化为标准形的算法和求正交方阵 $\mathbf{O}$ 使得实对称矩阵 $\mathbf{A}$ 对角化的算法相同.

例题 5.31. 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.

证明. 二次型 $Q(\mathbf{x})$ 的矩阵是

$$\mathbf{A} = \begin{pmatrix} 1 & -2 & 2 \\ -2 & 4 & -4 \\ 2 & -4 & 4 \end{pmatrix},$$

且 $p_{\mathbf{A}}(\lambda) = \lambda^2(\lambda - 9)$. 可以求出一组构成标准正交基的特征向量为

$$\mathbf{u}_1 = [\frac{1}{3}, \frac{-2}{3}, \frac{2}{3}], \quad \mathbf{u}_2 = [\frac{2}{\sqrt{5}}, \frac{1}{\sqrt{5}}, 0], \quad \mathbf{u}_3 = [\frac{2}{3\sqrt{5}}, \frac{-4}{3\sqrt{5}}, \frac{5}{3\sqrt{5}}].$$

令 $\mathbf{O} = \begin{pmatrix} \mathbf{u}_1 & \mathbf{u}_2 & \mathbf{u}_3 \end{pmatrix}$, 则在坐标变换 $\mathbf{x} = \mathbf{O}\mathbf{y}$ 下二次型化为标准形 $9y_1^2$. □

将二次型化为标准形后就可以得到二次型正定, 负定及不定的判定条件:

命题 5.32. 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) $\iff \lambda_i > 0$ $(\lambda_i \geq 0)$, $1 \leq i \leq n$.

(2) $Q(\mathbf{x})$ 是负定的 (半负定的) $\iff \lambda_i < 0$ $(\lambda_i \leq 0)$, $1 \leq i \leq n$.

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

评注 5.33. 事实上, 若目标是将二次型 $\mathbf{x}^{\mathrm{T}}\mathbf{A}\mathbf{x}$ 写成平方和 $\lambda_1 y_1^2 + \cdots + \lambda_n y_n^2$ 的形式, 而不要求 $\lambda_1, \ldots, \lambda_n$ 为 $\mathbf{A}$ 的特征值, 直接用 "配方法" 是更便捷的. 例如在例题 5.31 中, 可以直接看出 $Q(\mathbf{x}) = (x_1 - 2x_2 + 2x_3)^2$, 即在坐标变换

$$\begin{pmatrix} y_1 & y_2 & y_3 \end{pmatrix} = \begin{pmatrix} x_1 & x_2 & x_3 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 2 & 0 & 1 \end{pmatrix}$$

下 $Q(\mathbf{x})$ 化为标准形 y_1^2.

命题 5.34. 设 $\mathbf{A}$ 是实对称方阵, 则存在可逆矩阵 $\mathbf{P}$, 使得

$$\mathbf{P}^{\mathrm{T}}\mathbf{A}\mathbf{P} = \begin{pmatrix} \mathbf{I}_r & & \\ & -\mathbf{I}_s & \\ & & \mathbf{0} \end{pmatrix} = \mathrm{diag}(\mathbf{I}_r, -\mathbf{I}_s, \mathbf{0}), \tag{5.8}$$

等式右边只出现 1, −1, 0 的对角阵称为 $\mathbf{A}$ 的规范形. 相应的, 实二次型 $Q(\mathbf{x})$ 可以经过可逆坐标变换化为如下规范形:

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

证明. 已知存在正交矩阵 $\mathbf{O}$ 使得 $\mathbf{O}^{\mathrm{T}}\mathbf{A}\mathbf{O} = \mathrm{diag}(\lambda_1, \ldots, \lambda_n)$. 不妨设

$$\lambda_1, \ldots, \lambda_r > 0, \quad \lambda_{r+1}, \ldots, \lambda_{r+s} < 0, \quad \lambda_{r+s+1}, \ldots, \lambda_n = 0,$$

否则可以改变 $\mathbf{O}$ 的列向量的位置以达到此结果. 令

$$\mathbf{D} = \mathrm{diag}(\frac{1}{\sqrt{\lambda_1}}, \ldots, \frac{1}{\sqrt{\lambda_s}}, \frac{1}{\sqrt{-\lambda_{r+1}}}, \ldots, \frac{1}{\sqrt{-\lambda_{r+s}}}, 0, \ldots, 0), \quad \mathbf{P} = \mathbf{O}\mathbf{D},$$

则 $\mathbf{P}^{\mathrm{T}}\mathbf{A}\mathbf{P} = \mathbf{D}^{\mathrm{T}}\mathbf{O}^{\mathrm{T}}\mathbf{A}\mathbf{O}\mathbf{D} = \mathrm{diag}(\mathbf{I}_r, -\mathbf{I}_s, \mathbf{0})$. □

定理 5.35. 实对称矩阵 $\mathbf{A}$ 的规范形 $\mathrm{diag}(\mathbf{I}_r, -\mathbf{I}_s, \mathbf{0})$ 唯一.

证明. 假设存在可逆矩阵 $\mathbf{P}_1$, $\mathbf{P}_2$ 使得

$$\mathbf{P}_1^{\mathrm{T}}\mathbf{A}\mathbf{P}_1 = \begin{pmatrix} \mathbf{I}_r & & \\ & -\mathbf{I}_s & \\ & & \mathbf{0} \end{pmatrix}, \quad \mathbf{P}_2^{\mathrm{T}}\mathbf{A}\mathbf{P}_2 = \begin{pmatrix} \mathbf{I}_{r'} & & \\ & -\mathbf{I}_{s'} & \\ & & \mathbf{0} \end{pmatrix},$$

且 $\mathrm{diag}(\mathbf{I}_r, -\mathbf{I}_s, \mathbf{0})$, $\mathrm{diag}(\mathbf{I}_{r'}, -\mathbf{I}_{s'}, \mathbf{0})$ 是不同的规范形, 则不妨设 $r > r'$. 设 $\mathbf{P}_1$, $\mathbf{P}_2$ 的行向量形式为

$$\mathbf{P}_1 = \begin{pmatrix} \mathbf{v}_1 & \cdots & \mathbf{v}_n \end{pmatrix}, \quad \mathbf{P}_2 = \begin{pmatrix} \mathbf{w}_1 & \cdots & \mathbf{w}_n \end{pmatrix}.$$

对任意 $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$, $\mathbf{x}^{\mathrm{T}}\mathbf{A}\mathbf{y}$ 是一个实数, 从而可以定义函数 $f\colon \mathbb{R}^n \times \mathbb{R}^n \longrightarrow \mathbb{R}$, 满足

$$f(\mathbf{x}, \mathbf{y}) = \mathbf{x}^{\mathrm{T}}\mathbf{A}\mathbf{y}, \quad \mathbf{x}, \mathbf{y} \in \mathbb{R}^n.$$

利用分块矩阵乘法, 有

$$\begin{pmatrix} \mathbf{I}_r & & \\ & -\mathbf{I}_s & \\ & & \mathbf{0} \end{pmatrix} = \mathbf{P}_1^{\mathrm{T}}\mathbf{A}\mathbf{P}_1 = \begin{pmatrix} \mathbf{v}_1^{\mathrm{T}} \\ \vdots \\ \mathbf{v}_n^{\mathrm{T}} \end{pmatrix} \mathbf{A} \begin{pmatrix} \mathbf{v}_1 & \cdots & \mathbf{v}_n \end{pmatrix} = \begin{pmatrix} \mathbf{v}_1^{\mathrm{T}}\mathbf{A}\mathbf{v}_1 & \cdots & \mathbf{v}_1^{\mathrm{T}}\mathbf{A}\mathbf{v}_n \\ \vdots & \ddots & \vdots \\ \mathbf{v}_n^{\mathrm{T}}\mathbf{A}\mathbf{v}_1 & \cdots & \mathbf{v}_n^{\mathrm{T}}\mathbf{A}\mathbf{v}_n \end{pmatrix},$$

于是

$$f(\mathbf{v}_i, \mathbf{v}_j) = \begin{cases} 1 & 1 \le i = j \le r \\ -1 & r+1 \le i = j \le r+s \\ 0 & \text{其他情形} \end{cases} \tag{5.9}$$

类似的,

$$f(\mathbf{w}_i,\mathbf{w}_j)=\begin{cases}1 & 1\le i=j\le r'\\ -1 & r'+1\le i=j\le r'+s'\\ 0 & \text{其他情形}\end{cases} \tag{5.10}$$

因为 $r>r'$, 故 $\mathbf{v}_1,\dots,\mathbf{v}_r,\mathbf{w}_{r'+1},\dots,\mathbf{w}_n$ 线性相关, 存在 $\lambda_1,\dots,\lambda_r,\mu_{r'+1},\dots,\mu_n\in\mathbb{R}$ 不全为 0 使得

$$\lambda_1\mathbf{v}_1+\cdots+\lambda_r\mathbf{v}_r+\mu_{r'+1}\mathbf{w}_{r'+1}+\cdots+\mu_n\mathbf{w}_n=0.$$

设 $\mathbf{x}=\lambda_1\mathbf{v}_1+\cdots+\lambda_r\mathbf{v}_r=-\mu_{r'+1}\mathbf{w}_{r'+1}-\cdots-\mu_n\mathbf{w}_n$. 则一方面由 (5.9) 有

$$f(\mathbf{x},\mathbf{x})=f(\sum_{k=1}^r\lambda_k\mathbf{v}_k,\sum_{k=1}^r\lambda_k\mathbf{v}_k)=\lambda_1^2+\cdots+\lambda_r^2.$$

另一方面由 (5.10) 有

$$f(\mathbf{x},\mathbf{x})=f(-\sum_{\ell=r'+1}^n\mu_\ell\mathbf{w}_\ell,-\sum_{\ell=r'+1}^n\mu_\ell\mathbf{w}_\ell)=-\mu_{r'+1}^2-\cdots-\mu_n^2.$$

只能 $\lambda_1,\dots,\lambda_r,\mu_{r'+1},\dots,\mu_n$ 全为 0, 矛盾. □

推论 5.36. 二次型 $Q(\mathbf{x})$ 的规范形

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

唯一. r,s 分别称为 $Q(\mathbf{x})$ 的正惯性指数和负惯性指数.

定义 5.37. 设 $\mathbf{A}\in\mathrm{M}_n(\mathbb{R})$ 是实对称矩阵, $\mathbf{A}$ 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 5.38. $\mathbf{A}$ 是正定的 $\iff$ 存在可逆矩阵 $\mathbf{P}$ 使得 $\mathbf{A}=\mathbf{P}^\mathrm{T}\mathbf{P}$.

证明. "$\Longrightarrow$": 由命题 5.32 以及命题 5.34, 若 $\mathbf{A}$ 正定, 则存在可逆矩阵 $\mathbf{Q}$ 使得 $\mathbf{Q}^\mathrm{T}\mathbf{A}\mathbf{Q}=\mathbf{I}_n$. 取 $\mathbf{P}=\mathbf{Q}^{-1}$, 则 $\mathbf{A}=\mathbf{P}^\mathrm{T}\mathbf{P}$.
"$\Longleftarrow$": 设 $\mathbf{P}$ 是可逆矩阵, $\mathbf{A}=\mathbf{P}^\mathrm{T}\mathbf{P}$, 则对 $\forall\ \mathbf{x}\in\mathbb{R}^n$,

$$\mathbf{x}^\mathrm{T}\mathbf{A}\mathbf{x}=\mathbf{x}^\mathrm{T}\mathbf{P}^\mathrm{T}\mathbf{P}\mathbf{x}=(\mathbf{P}\mathbf{x})^\mathrm{T}(\mathbf{P}\mathbf{x})\ge 0,$$

且等式成立 $\iff\mathbf{P}\mathbf{x}=\mathbf{0}\iff\mathbf{x}=\mathbf{0}$. □

推论 5.39. 设 $\mathbf{A}$ 是正定矩阵, 则 $\det \mathbf{A} > 0$.

证明. 由推论 5.38, 存在可逆矩阵 $\mathbf{P}$ 使得 $\mathbf{A} = \mathbf{P}^{\mathrm{T}}\mathbf{P}$, 则 $\det \mathbf{A} = \det \mathbf{P}^{\mathrm{T}}\mathbf{P} = (\det \mathbf{P})^2 > 0$. □

下面给出判别矩阵正定 (负定) 的实用方法:

定义 5.40. 设 $\mathbf{A} = (a_{ij})_{n\times n} \in \mathsf{M}_n(\mathbb{R})$, 对每个 $1 \le k \le n$,

$$A_k = \begin{vmatrix} a_{11} & \cdots & a_{1k} \\ \vdots & \ddots & \vdots \\ a_{k1} & \cdots & a_{kk} \end{vmatrix}$$

称为 $\mathbf{A}$ 的 *k 阶顺序主子式*.

定理 5.41. 实对称方阵 $\mathbf{A}$ 是正定的 $\Longleftrightarrow$ $\mathbf{A}$ 的各阶顺序主子式 > 0.

证明. "$\Longrightarrow$": 若 $\mathbf{A}$ 正定, 则对 $\forall\, 1 \le k \le n$, 取 $\mathbf{x} \in \mathbb{R}^k$ 以及 $\tilde{\mathbf{x}} = \begin{bmatrix} \mathbf{x} & \mathbf{0} \end{bmatrix} \in \mathbb{R}^n$, 则

$$\tilde{\mathbf{x}}^{\mathrm{T}}\mathbf{A}\tilde{\mathbf{x}} = \begin{pmatrix} \mathbf{x}^{\mathrm{T}} & \mathbf{0} \end{pmatrix} \begin{pmatrix} \mathbf{A}_1 & \mathbf{A}_2 \\ \mathbf{A}_2^{\mathrm{T}} & \mathbf{A}_3 \end{pmatrix} \begin{pmatrix} \mathbf{x} \\ \mathbf{0} \end{pmatrix} = \mathbf{x}^{\mathrm{T}}\mathbf{A}_1\mathbf{x},$$

其中 $\mathbf{A}_1$ 是 $\mathbf{A}$ 的 $1, \ldots, k$ 行, $1, \ldots, k$ 列构成的子方阵, 也是实对称矩阵. 由 $\mathbf{A}$ 正定, 对 $\forall\, \tilde{\mathbf{x}} \neq \mathbf{0}$, $\tilde{\mathbf{x}}^{\mathrm{T}}\mathbf{A}\tilde{\mathbf{x}} > 0$, 从而对 $\forall\, \mathbf{x} \neq \mathbf{0}$, $\mathbf{x}^{\mathrm{T}}\mathbf{A}_1\mathbf{x} > 0$, 即 $\mathbf{A}_1$ 也是正定的. 于是 $\det \mathbf{A}_1 > 0$, 即 $\mathbf{A}$ 的 k 阶顺序主子式 > 0.
"$\Longleftarrow$": 用归纳法, 假设对 $\forall\, \mathbf{A} \in \mathsf{M}_{n-1}(\mathbb{R})$, $\mathbf{A}$ 的各阶顺序主子式 > 0 可以推出 $\mathbf{A}$ 正定. 任取 $\mathbf{A} \in \mathsf{M}_n(\mathbb{R})$, 且各阶顺序主子式 > 0. 设 $\mathbf{A}_1$ 是 $\mathbf{A}$ 的 $1, \ldots, n-1$ 行, $1, \ldots, n-1$ 列构成的子矩阵, 且

$$\mathbf{A} = \begin{pmatrix} \mathbf{A}_1 & \mathbf{A}_2 \\ \mathbf{A}_2^{\mathrm{T}} & a_{nn} \end{pmatrix}.$$

由条件及归纳假设, $\mathbf{A}_1$ 是正定矩阵, 故 $\mathbf{A}_1$ 可逆且 $\det \mathbf{A}_1 > 0$. 设

$$\mathbf{P}_1 = \begin{pmatrix} \mathbf{I}_{n-1} & -\mathbf{A}_1^{-1}\mathbf{A}_2 \\ \mathbf{0} & 1 \end{pmatrix},$$

则 $\mathbf{P}_1$ 是可逆矩阵且

$$\mathbf{P}_1^{\mathrm{T}}\mathbf{A}\mathbf{P}_1 = \begin{pmatrix} \mathbf{I}_{n-1} & \mathbf{0} \\ -\mathbf{A}_2^{\mathrm{T}}\mathbf{A}_1^{-1} & 1 \end{pmatrix} \begin{pmatrix} \mathbf{A}_1 & \mathbf{A}_2 \\ \mathbf{A}_2^{\mathrm{T}} & a_{nn} \end{pmatrix} \begin{pmatrix} \mathbf{I}_{n-1} & -\mathbf{A}_1^{-1}\mathbf{A}_2 \\ \mathbf{0} & 1 \end{pmatrix}$$

$$= \begin{pmatrix} \mathbf{I}_{n-1} & \mathbf{0} \\ -\mathbf{A}_2^{\mathrm{T}}\mathbf{A}_1^{-1} & 1 \end{pmatrix} \begin{pmatrix} \mathbf{A}_1 & \mathbf{0} \\ \mathbf{A}_2^{\mathrm{T}} & a_{nn} - \mathbf{A}_2^{\mathrm{T}}\mathbf{A}_1^{-1}\mathbf{A}_2 \end{pmatrix} = \begin{pmatrix} \mathbf{A}_1 & \mathbf{0} \\ \mathbf{0} & a_{nn} - \mathbf{A}_2^{\mathrm{T}}\mathbf{A}_1^{-1}\mathbf{A}_2 \end{pmatrix}.$$

设 $b_{nn} = a_{nn} - \mathbf{A}_2^{\mathrm{T}}\mathbf{A}_1^{-1}\mathbf{A}_2$, 则 $|\mathbf{A}_1|b_{nn} = \left|\mathbf{P}_1^{\mathrm{T}}\mathbf{A}\mathbf{P}_1\right| = |\mathbf{A}||\mathbf{P}_1|^2$, 由条件知 $b_{nn} > 0$. 因为 $\mathbf{A}_1$ 正定, 存在可逆矩阵 $\mathbf{Q}$ 使得 $\mathbf{Q}^{\mathrm{T}}\mathbf{A}_1\mathbf{Q} = \mathbf{I}_{n-1}$. 设

$$\mathbf{P}_2 = \begin{pmatrix} \mathbf{Q} & \\ & \dfrac{1}{\sqrt{b_{nn}}} \end{pmatrix}, \quad \mathbf{P} = \mathbf{P}_1\mathbf{P}_2,$$

则 $\mathbf{P}^{\mathrm{T}}\mathbf{A}\mathbf{P} = \mathbf{I}_n$, 从而 $\mathbf{A}$ 正定. □

推论 5.42. 实对称方阵 $\mathbf{A}$ 负定 $\iff$ $\mathbf{A}$ 的奇数阶顺序主子式 < 0, 偶数阶顺序主子式 > 0.

证明. $\mathbf{A}$ 负定 $\iff -\mathbf{A}$ 正定, 对 $-\mathbf{A}$ 应用定理 5.41 即可. □

例题 5.43. 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) $Q(\mathbf{x}) = x_1^2 - x_2^2 + x_3^2 + 4x_1x_3 + 2x_2x_3$.

解. (1) $Q(\mathbf{x})$ 的矩阵是

$$\mathbf{A} = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 2 & 3 \\ 1 & 3 & 6 \end{pmatrix},$$

可知其各阶顺序主子式 > 0, 故 $\mathbf{A}$ 正定, $Q(\mathbf{x})$ 的规范形为 $y_1^2 + y_2^2 + y_3^2$.
(2) $Q(\mathbf{x})$ 的矩阵是

$$\mathbf{A} = \begin{pmatrix} 1 & 0 & 2 \\ 0 & -1 & 1 \\ 2 & 1 & 1 \end{pmatrix},$$

各阶顺序主子式为 $A_1 = 1 > 0$, $A_2 = -1 < 0$, $A_3 = 2 > 0$, 故 $\mathbf{A}$ 既不正定, 也不负定. 设 $\mathbf{A}$ 的规范形为

$$\mathbf{P}^{\mathrm{T}}\mathbf{A}\mathbf{P} = \operatorname{diag}(\mathbf{I}_r, -\mathbf{I}_s, \mathbf{0}).$$

由 $\left|\mathbf{P}^{\mathrm{T}}\mathbf{A}\mathbf{P}\right| = |\mathbf{P}|^2|\mathbf{A}| > 0$ 知只能 $r = 1$, $s = 2$. 故 $Q(\mathbf{x})$ 的规范形为 $y_1^2 - y_2^2 - y_3^2$. □

参考文献

[1] 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, 2015.

[2] 刘吉佑, 莫骄. 线性代数与几何（第二版）, 北京邮电大学出版社, 2018.

[3] 李尚志. 线性代数（数学专业用）, 高等教育出版社, 2006.

[4] 席南华. 基础代数（第一卷）, 科学出版社, 2016.