线性代数(I)

目 录

第一章　向量

向量是力学与物理等领域中诸多概念的数学抽象, 例如力、速度、加速度、动量等等. 向量具有很强的几何直观, 同时也可以进行代数运算. 利用向量可以很简洁地表示数学与物理中的结构, 在力学、物理和工程技术领域有着广泛的应用. 此外, 向量是解析几何的重要工具, 也是我们学习抽象线性空间理论的基础.

中学教材中介绍了平面向量和空间向量的加法、数乘、内积等运算, 以及向量法在求解平面几何和空间几何问题中的应用. 为保证本教材的自完备性, 我们还是从向量的定义出发, 完整地介绍向量的各种运算规则, 并着重强调它与后续章节(特别是第五章)的关系.

本章前五节我们讨论的向量均为三维空间中的向量.

§1.1　向量的线性运算

§1.1.1　向量及其表示

向量的概念来源于物理学. 很多物理量不仅有大小, 而且有方向, 例如速度、位移、力等等. 抛开它们的物理意义, 只保留大小与方向两个要素, 就抽象为数学中的向量概念: 既有大小, 又有方向的量称为**向量**.

一般用有向线段表示一个向量, 线段的长度表示它的大小, 线段的方向表示它的方向. 以空间中 A 为起点, B 为终点的有向线段所表示的向量记为 $\overrightarrow{AB}$. 常用黑斜体小写字母 $\mathbf{a}, \mathbf{b}, \mathbf{c}$ 等表示向量. 如果两个向量大小相等、方向相同, 就称这两个向量是**相等**的.

如果两个向量的大小相等而方向相反, 则称这两个向量互为**反向量**. 向量 $\mathbf{a}$ 的反向量记为 $-\mathbf{a}$, 也称为 $\mathbf{a}$ 的**负向量**.

向量的长度也称为向量的**模**, 向量 $\mathbf{a}$ 的模用 $|\mathbf{a}|$ 表示. 模为 1 的向量称为**单位向量**, 模为零的向量称为**零向量**, 记作 $\mathbf{0}$. 零向量的起点和终点是重合的, 因此它没有确定的方向.

如果向量 $\mathbf{a}$ 与向量 $\mathbf{b}$ 的方向相同或相反, 就称它们**平行**, 记作 $\mathbf{a}//\mathbf{b}$. 如果向量 $\mathbf{a}$ 与向量 $\mathbf{b}$ 的方向互相垂直, 就称它们**垂直**或**正交**, 记作 $\mathbf{a} \perp \mathbf{b}$. 规定零向量与任何向量都平行且正交.

向量 $\overrightarrow{OA}$与 $\overrightarrow{OB}$所夹的角$\angle AOB$称为它们之间的**夹角**. 向量夹角介于0与π之间. 向量 $\overrightarrow{OA}$与 $\overrightarrow{OB}$平行当且仅当它们之间的夹角为0或π, 而向量 $\overrightarrow{OA}$与 $\overrightarrow{OB}$垂直当且仅当它们之间夹角为$\pi/2$.

§1.1.2 向量的线性运算

将物理中速度、力的合成法加以抽象, 就得到向量加法的定义. 给定具有相同起点 O 的两个向量 $\mathbf{a}=\overrightarrow{OA}, \mathbf{b}=\overrightarrow{OB}$, 则以 $\overrightarrow{OA}, \overrightarrow{OB}$ 为邻边的平行四边形的对角线向量 $\mathbf{c}=\overrightarrow{OC}$(图 1.1)就称为这两个向量的**和**, 记作

$$\overrightarrow{OC}=\overrightarrow{OA}+\overrightarrow{OB} \quad \text{或者} \quad \mathbf{c}=\mathbf{a}+\mathbf{b}.$$

这种求和的方法称为**平行四边形法则**.

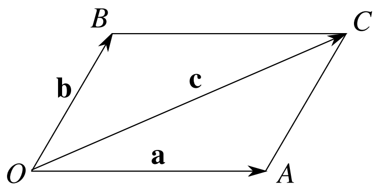


图 1.1 向量的加法

从图 1.1可知, $\overrightarrow{OB}=\overrightarrow{AC}$, 所以 $\overrightarrow{OC}=\overrightarrow{OA}+\overrightarrow{AC}$, 这称为两个向量的和的**三角形法则**. 由定义不难看出向量的加法满足以下的性质:

$$\mathbf{a}+\mathbf{b}=\mathbf{b}+\mathbf{a} \tag{1.1}$$

$$\mathbf{a}+(\mathbf{b}+\mathbf{c})=(\mathbf{a}+\mathbf{b})+\mathbf{c} \tag{1.2}$$

$$\mathbf{a}+\mathbf{0}=\mathbf{a} \tag{1.3}$$

$$\mathbf{a}+(-\mathbf{a})=\mathbf{0} \tag{1.4}$$

向量的**减法**为向量加法的逆运算. 对于向量 $\mathbf{a},\mathbf{b}$, 定义向量减法 $\mathbf{a}-\mathbf{b}=\mathbf{a}+(-\mathbf{b})$.

定义向量 $\mathbf{a}$ 与实数 λ 的乘积为一个向量, 记为 $\lambda\mathbf{a}$, 它的模为 $|\lambda||\mathbf{a}|$, 它的方向规定为: 当 $\lambda>0$ 时, 与 $\mathbf{a}$ 同向; 当 $\lambda<0$时, 与 $\mathbf{a}$ 反向. 这种运算称为向量的**数乘**. 由数乘的定义可知 $0\mathbf{a}=\mathbf{0}$, 并且对任意实数 λ,μ, 都有

$$1\mathbf{a}=\mathbf{a} \tag{1.5}$$

$$\lambda(\mu\mathbf{a})=(\lambda\mu)\mathbf{a} \tag{1.6}$$

$$(\lambda+\mu)\mathbf{a}=\lambda\mathbf{a}+\mu\mathbf{a} \tag{1.7}$$

$$\lambda(\mathbf{a}+\mathbf{b})=\lambda\mathbf{a}+\lambda\mathbf{b} \tag{1.8}$$

对于非零向量 $\mathbf{a}$, 用 $\mathbf{a}^0$ 表示与 $\mathbf{a}$ 同向的单位向量, 则由向量数乘的定义知

$$\mathbf{a}^0 = \frac{\mathbf{a}}{|\mathbf{a}|}.$$

向量的加法与数乘运算统称为向量的**线性运算**. 线性运算并非三维向量所特有. 事实上, 在很多非空集合中都可以定义所谓的加法与数乘运算, 并且满足相应于 (1.1)-(1.8) 的 8 条性质. 我们将这样的集合(附带加法与数乘运算)称为**线性空间**或**向量空间**, 线性空间中的元素称为(抽象的) 向量. 具体内容详见第五章.

§1.1.3 向量的共线与共面

一组向量称为是**共线的**, 如果它们都平行于某条直线. 一组向量称为是**共面的**, 如果它们都平行于某个平面.

命题1.1.1. *向量 $\boldsymbol{a},\boldsymbol{b}$ 共线的充分必要条件是存在不全为零的实数 λ,μ, 使得*

$$\lambda\boldsymbol{a}+\mu\boldsymbol{b}=\boldsymbol{0}.$$

证明. 必要性:设向量 $\mathbf{a},\mathbf{b}$ 共线, 不妨设 $\mathbf{a}$ 不是零向量. 若向量 $\mathbf{b}$ 与向量 $\mathbf{a}$ 同向, 则 $\mathbf{b}=\frac{|\mathbf{b}|}{|\mathbf{a}|}\mathbf{a}$, 因此 $\frac{|\mathbf{b}|}{|\mathbf{a}|}\mathbf{a}+(-1)\mathbf{b}=0$. 若向量 $\mathbf{b}$ 与向量 $\mathbf{a}$ 反向, 则 $\mathbf{b}=-\frac{|\mathbf{b}|}{|\mathbf{a}|}\mathbf{a}$, 因此 $\frac{|\mathbf{b}|}{|\mathbf{a}|}\mathbf{a}+1\cdot\mathbf{b}=0$.

充分性: 设 λ,μ 为不全为零的实数且 $\lambda\mathbf{a}+\mu\mathbf{b}=\mathbf{0}$. 不妨设 $\mu\neq 0$, 则 $\mathbf{b}=-\frac{\lambda}{\mu}\mathbf{a}$, 因此 $\mathbf{a},\mathbf{b}$ 共线. □

命题1.1.2. *向量 $\boldsymbol{a},\boldsymbol{b},\boldsymbol{c}$ 共面的充分必要条件是存在不全为零的实数 λ,μ,ν, 使得*

$$\lambda\boldsymbol{a}+\mu\boldsymbol{b}+\nu\boldsymbol{c}=\boldsymbol{0}.$$

证明. 必要性:若 $\mathbf{a},\mathbf{b},\mathbf{c}$ 中有两个向量, 例如 $\mathbf{a},\mathbf{b}$ 共线, 则存在不全为零的实数 λ,μ, 使得 $\lambda\mathbf{a}+\mu\mathbf{b}=\mathbf{0}$. 从而 $\lambda\mathbf{a}+\mu\mathbf{b}+0\cdot\mathbf{c}=\mathbf{0}$, 其中 $\lambda,\mu,0$ 不全为零.

设 $\mathbf{a},\mathbf{b},\mathbf{c}$ 中任意两个向量都不共线. 取定一点 O, 作 $\overrightarrow{OA}=\mathbf{a}$, $\overrightarrow{OB}=\mathbf{b}$, $\overrightarrow{OC}=\mathbf{c}$. 过 C 点作 OB 的平行线交直线 OA 于点 D(图 1.2), 则存在实数 λ,μ, 使得$\overrightarrow{OC}=\overrightarrow{OD}+\overrightarrow{DC}=\lambda\overrightarrow{OA}+\mu\overrightarrow{OB}$. 移项后得 $\lambda\mathbf{a}+\mu\mathbf{b}+(-1)\mathbf{c}=\mathbf{0}$, 其中 $\lambda,\mu,-1$ 不全为零.

充分性:设存在不全为零的实数 λ,μ,ν, 使得 $\lambda\mathbf{a}+\mu\mathbf{b}+\nu\mathbf{c}=\mathbf{0}$. 不妨设 $\nu\neq\mathbf{0}$, 于是 $\mathbf{c}=-\frac{\lambda}{\nu}\mathbf{a}-\frac{\mu}{\nu}\mathbf{b}$. 因此 $\mathbf{c}$ 是以 $-\frac{\lambda}{\nu}\mathbf{a},-\frac{\mu}{\nu}\mathbf{b}$ 为边的平行四边形的对角线, 从而 $\mathbf{a},\mathbf{b},\mathbf{c}$ 共面. □

定义1.1.1. 设 $\mathbf{a}_1,\mathbf{a}_2,\cdots,\mathbf{a}_n$ 为一组向量, $\lambda_1,\lambda_2,\cdots,\lambda_n$为实数. 称向量

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

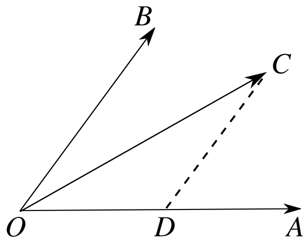


图 1.2 三向量共面条件

为向量 $\mathbf{a}_1,\mathbf{a}_2,\cdots,\mathbf{a}_n$ 的**线性组合**.

利用这个定义, 命题 1.1.1和命题 1.1.2也有如下的表述方式:两个向量共线当且仅当某一个向量为另一个向量的线性组合(倍数)；三个向量共面当且仅当某一个向量为另外两个向量的线性组合.

定义1.1.2. 一组向量 $\mathbf{a}_1,\mathbf{a}_2,\cdots,\mathbf{a}_n$ 称为**线性相关**, 如果存在不全为零的实数 $\lambda_1,\lambda_2,\cdots,\lambda_n$, 使得

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

反之, 不是线性相关的一组向量称为**线性无关**. 也就是, 如果上式成立, 则$\lambda_1=\lambda_2=\ldots=\lambda_n=0$.

利用命题 1.1.1与命题 1.1.2可以得到

- 一个向量 $\mathbf{a}$ 线性相关当且仅当 $\mathbf{a}=\mathbf{0}$;
- 两个向量线性相关当且仅当它们共线;
- 三个向量线性相关当且仅当它们共面.

类似可以得到线性无关的等价条件. 线性相关与线性无关是线性代数中最基本的概念之一, 在第五章中我们会作进一步的讨论.

例1.1.1. 对任意向量$\mathbf{a},\mathbf{b},\mathbf{c}$, 证明向量 $\mathbf{a}+\mathbf{b}+\mathbf{c}$, $\mathbf{a}-\mathbf{b}-\mathbf{c}$, $\mathbf{a}+2\mathbf{b}+2\mathbf{c}$ 线性相关.

证明. 我们只要证明, 存在不全为零的实数解 λ,μ,ν使得

$$\lambda(\mathbf{a}+\mathbf{b}+\mathbf{c})+\mu(\mathbf{a}-\mathbf{b}-\mathbf{c})+\nu(\mathbf{a}+2\mathbf{b}+2\mathbf{c})=\mathbf{0}.$$

化简上式得

$$(\lambda+\mu+\nu)\mathbf{a}+(\lambda-\mu+2\nu)\mathbf{b}+(\lambda-\mu+2\nu)\mathbf{c}=\mathbf{0}.$$

所以只要证明方程组

$$\begin{cases} \lambda+\mu+\nu=0 \\ \lambda-\mu+2\nu=0 \end{cases}$$

有不全为零的解. 易见 $\lambda=-3,\mu=1,\nu=2$ 为一组非零解, 因此三个向量线性相关. $\square$

例1.1.2. 证明:空间中任意三点 A,B,C 共线的充分必要条件是, 存在不全为零的实数 k_1,k_2,k_3 使得对任意点 O 都有

$$k_1\overrightarrow{OA}+k_2\overrightarrow{OB}+k_3\overrightarrow{OC}=\mathbf{0}, \quad 且 \quad k_1+k_2+k_3=0.$$

证明. 必要性:设 A,B,C 三点共线, 则向量 $\overrightarrow{AB}$ 与向量 $\overrightarrow{AC}$ 共线. 因此存在不全为零的实数 λ,μ, 使得 $\lambda\overrightarrow{AB}+\mu\overrightarrow{AC}=\mathbf{0}$, 即

$$\lambda(\overrightarrow{OB}-\overrightarrow{OA})+\mu(\overrightarrow{OC}-\overrightarrow{OA})=\mathbf{0}.$$

化简得 $(-\lambda-\mu)\overrightarrow{OA}+\lambda\overrightarrow{OB}+\mu\overrightarrow{OC}=\mathbf{0}$. 取 $k_1=-\lambda-\mu, k_2=\lambda, k_3=\mu$, 则

$$k_1\overrightarrow{OA}+k_2\overrightarrow{OB}+k_3\overrightarrow{OC}=\mathbf{0}, \quad 且 \quad k_1+k_2+k_3=0.$$

充分性:设存在不全为零的 k_1,k_2,k_3 满足条件, 不妨设 $k_1\neq 0$. 则 $k_3=-k_1-k_2$, 且 $k_1\overrightarrow{OA}+k_2\overrightarrow{OB}+(-k_1-k_2)\overrightarrow{OC}=\mathbf{0}$. 所以

$$k_1(\overrightarrow{OA}-\overrightarrow{OC})+k_2(\overrightarrow{OB}-\overrightarrow{OC})=\mathbf{0},$$

即 $k_1\overrightarrow{CA}+k_2\overrightarrow{CB}=\mathbf{0}$. 由于 k_1,k_2 不全为零, 向量 $\overrightarrow{CA}$ 与$\overrightarrow{CB}$ 共线, 即 A,B,C 三点共线. $\square$

利用向量运算可以解决许多几何问题, 其思想是将几何性质转化为向量的代数运算.

例1.1.3. $\triangle ABC$ 中, D,E 分别是边 BC,AC 的中点, AD,BE 相交于点 G(图 1.3). 证明: $AG=\dfrac{2}{3}AD$.

证明. 设 $\overrightarrow{AG}=x\overrightarrow{AD}$, 由于 D 为 BC 的中点, $\overrightarrow{AD}=\dfrac{1}{2}(\overrightarrow{AB}+\overrightarrow{AC})$. 所以

$$\overrightarrow{AG}=\frac{x}{2}(\overrightarrow{AB}+\overrightarrow{AC}).$$

由于 B,G,E 共线, 根据例 1.1.2, 可设 $\overrightarrow{AG}=y\overrightarrow{AB}+(1-y)\overrightarrow{AE}$. 由于 E 为 AC 的中点, 所以

$$\overrightarrow{AG}=y\overrightarrow{AB}+\frac{(1-y)}{2}\overrightarrow{AC}.$$

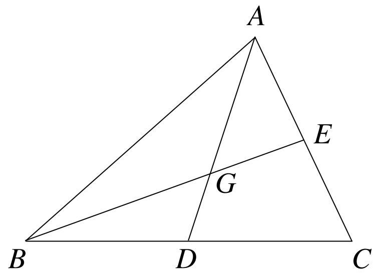


图 1.3

因此

$$(\frac{x}{2}-y)\overrightarrow{AB}+(\frac{x}{2}-\frac{1-y}{2})\overrightarrow{AC}=\mathbf{0}.$$

由于 $\overrightarrow{AB},\overrightarrow{AC}$ 不共线, 由命题 1.1.1得

$$\begin{cases}\dfrac{x}{2}-y=0\\ \dfrac{x}{2}-\dfrac{1-y}{2}=0\end{cases}$$

解得 $x=\dfrac{2}{3}$, 因此 $AG=\dfrac{2}{3}AD$. □

§1.2 坐标系

§1.2.1 仿射坐标系

在中学我们学习了直角坐标系. 在直角坐标系中, 三个坐标轴两两垂直. 本节我们将坐标系推广到坐标轴不相互垂直的情形. 我们先陈述向量的基本定理.

定理1.2.1. *设 $\boldsymbol{e}_1,\boldsymbol{e}_2,\boldsymbol{e}_3$ 为空间中三个不共面的向量, 则对每个向量 $\boldsymbol{a}$ 都存在唯一的三元有序实数组 (x_1,x_2,x_3), 使得*

$$\boldsymbol{a}=x_1\boldsymbol{e}_1+x_2\boldsymbol{e}_2+x_3\boldsymbol{e}_3. \tag{1.9}$$

证明. 在空间中任取一点 O, 作向量 $\overrightarrow{OA}=\mathbf{e}_1,\overrightarrow{OB}=\mathbf{e}_2,\overrightarrow{OC}=\mathbf{e}_3,\overrightarrow{OP}=\mathbf{a}$(图 1.4).

过 P 点作直线 OC 的平行线, 交 AOB 平面于 Q 点. 再过 Q 点作直线 OB 的平行线, 交 OA 直线于 R 点. 则 $\mathbf{a}=\overrightarrow{OP}=\overrightarrow{OR}+\overrightarrow{RQ}+\overrightarrow{QP}$. 由于 $\overrightarrow{OR}//\overrightarrow{OA}$, $\overrightarrow{RQ}//\overrightarrow{OB}$, $\overrightarrow{QP}//\overrightarrow{OC}$, 利用命题 1.1.1知, 存在实数 x_1,x_2,x_3 使得

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

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

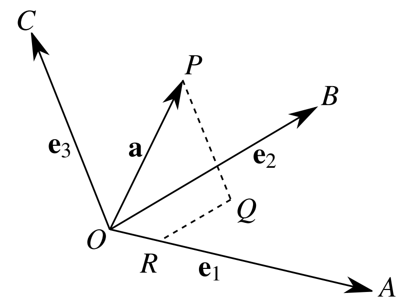


图 1.4 向量基本定理

下面证明唯一性. 如果向量 $\mathbf{a}$ 有两种表示方式:

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

则有

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

由于向量 $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$ 不共面, 由命题 1.1.2 知 $x_1 = y_1, x_2 = y_2, x_3 = y_3$, 因此表示方式是唯一的. □

定义1.2.1. 空间中任意三个有序的不共面的向量 $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$ 称为空间的一组**基**. 对于向量 $\mathbf{a}$, 若

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

则称 (x_1, x_2, x_3) 为向量 $\mathbf{a}$ 在基 $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$ 下的**仿射坐标**或简称**坐标**.

定义1.2.2. 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 $[O; \mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3]$. 点 O 称为**坐标原点**, $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$ 称为**坐标向量**. $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$ 所在直线分别称为 x **轴**, y **轴**和z **轴**, 统称为**坐标轴**. 三个坐标轴的任意两个决定了一个平面, 称为**坐标面**, 分别记为 Oxy, Oyz, Ozx.

给定仿射坐标系 $[O; \mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3]$, 对空间中一点 P, 向量 $\overrightarrow{OP}$ 在基 $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$ 下的坐标称为点 P 在仿射坐标系 $[O; \mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3]$ 下的**坐标**. 因此点 P 在 $[O; \mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3]$ 下的坐标为 (x_1, x_2, x_3) 当且仅当 $\overrightarrow{OP} = x_1\mathbf{e}_1 + x_2\mathbf{e}_2 + x_3\mathbf{e}_3$.

从上面的讨论知, 在引入仿射坐标系后, 在下列三者间存在一一对应的关系:

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

例1.2.1. 设 $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$ 为一组基,

(1) 证明:$\mathbf{a} = \mathbf{e}_1 + \mathbf{e}_2 + \mathbf{e}_3$, $\mathbf{b} = \mathbf{e}_1 - \mathbf{e}_2 + \mathbf{e}_3$, $\mathbf{c} = \mathbf{e}_1 + \mathbf{e}_2 - \mathbf{e}_3$ 也是一组基;

(2) 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.

解. (1) 为证明 $\mathbf{a},\mathbf{b},\mathbf{c}$ 为一组基, 只需证明 $\mathbf{a},\mathbf{b},\mathbf{c}$ 不共面. 为此设实数 λ,μ,ν 满足 $\lambda\mathbf{a}+\mu\mathbf{b}+\nu\mathbf{c}=\mathbf{0}$, 即

$$\lambda(\mathbf{e}_1+\mathbf{e}_2+\mathbf{e}_3)+\mu(\mathbf{e}_1-\mathbf{e}_2+\mathbf{e}_3)+\nu(\mathbf{e}_1+\mathbf{e}_2-\mathbf{e}_3)=\mathbf{0}.$$

化简得

$$(\lambda+\mu+\nu)\mathbf{e}_1+(\lambda-\mu+\nu)\mathbf{e}_2+(\lambda+\mu-\nu)\mathbf{e}_3=\mathbf{0}.$$

由于 $\mathbf{e}_1,\mathbf{e}_2,\mathbf{e}_3$ 不共面, 我们得到下面的方程组

$$\begin{cases}\lambda+\mu+\nu=0\\ \lambda-\mu+\nu=0\\ \lambda+\mu-\nu=0\end{cases}$$

不难解得 $\lambda=\mu=\nu=0$, 因此 $\mathbf{a},\mathbf{b},\mathbf{c}$ 不共面.

(2) 由于

$$\begin{aligned}\mathbf{d}&=\mathbf{a}+2\mathbf{b}-3\mathbf{c}\\&=(\mathbf{e}_1+\mathbf{e}_2+\mathbf{e}_3)+2(\mathbf{e}_1-\mathbf{e}_2+\mathbf{e}_3)-3(\mathbf{e}_1+\mathbf{e}_2-\mathbf{e}_3)\\&=-4\mathbf{e}_2+6\mathbf{e}_3\end{aligned}$$

所以向量 $\mathbf{d}$ 在基 $\mathbf{e}_1,\mathbf{e}_2,\mathbf{e}_3$ 下的坐标为 $(0,-4,6)$. □

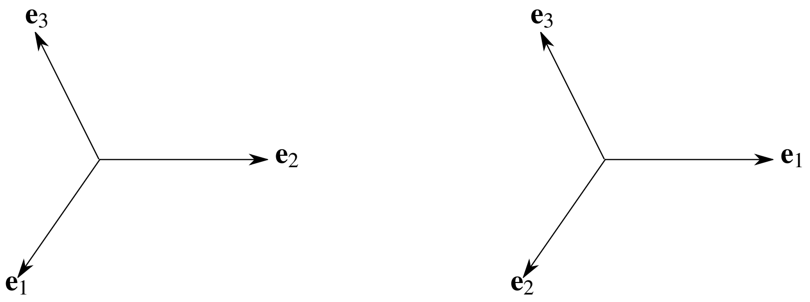
一个仿射坐标系的三个坐标面将空间分成八个部分, 称为八个**卦限**. 其中每个卦限内点的坐标的正负号规定为:

$$\begin{array}{llll}\mathrm{I}(+,+,+), & \mathrm{II}(-,+,+), & \mathrm{III}(-,-,+), & \mathrm{IV}(+,-,+)\\ \mathrm{V}(+,+,-), & \mathrm{VI}(-,+,-), & \mathrm{VII}(-,-,-), & \mathrm{VIII}(+,-,-)\end{array}$$

一组基 $\mathbf{e}_1,\mathbf{e}_2,\mathbf{e}_3$ 在空间中的位置关系有两种情形, 如图 1.5. 称图 1.5(a)所示的坐标系为**右手仿射坐标系**, 图 1.5(b) 所示的为**左手仿射坐标系**.

§1.2.2 向量的坐标运算

设在空间中取定了一个仿射坐标系 $[O;\mathbf{e}_1,\mathbf{e}_2,\mathbf{e}_3]$. 由于向量与其坐标之间存在一一对应的关系, 向量的运算可以转化为其坐标间的运算.



(a) 右手系 (b) 左手系

图 1.5

设 $\mathbf{a}=(x_1,x_2,x_3)$, $\mathbf{b}=(y_1,y_2,y_3)$, λ 为一个实数, 则

$$\begin{aligned}\mathbf{a}+\mathbf{b}&=(x_1\mathbf{e}_1+x_2\mathbf{e}_2+x_3\mathbf{e}_3)+(y_1\mathbf{e}_1+y_2\mathbf{e}_2+y_3\mathbf{e}_3)\\&=(x_1+y_1)\mathbf{e}_1+(x_2+y_2)\mathbf{e}_2+(x_3+y_3)\mathbf{e}_3,\\ \lambda\mathbf{a}&=\lambda(x_1\mathbf{e}_1+x_2\mathbf{e}_2+x_3\mathbf{e}_3)=\lambda x_1\mathbf{e}_1+\lambda x_2\mathbf{e}_2+\lambda x_3\mathbf{e}_3.\end{aligned}$$

所以我们有下面的坐标计算公式:

$$(x_1,x_2,x_3)+(y_1,y_2,y_3)=(x_1+y_1,x_2+y_2,x_3+y_3), \tag{1.10}$$

$$\lambda(x_1,x_2,x_3)=(\lambda x_1,\lambda x_2,\lambda x_3). \tag{1.11}$$

例1.2.2. 设 $A(x_1,y_1,z_1),B(x_2,y_2,z_2)$ 为空间中两点, 若点 P 将 $\overrightarrow{AB}$ 分割成定比 λ, 即 $\overrightarrow{AP}=\lambda\overrightarrow{PB}$, 求分点 P 的坐标.

解. 由 $\overrightarrow{AP}=\lambda\overrightarrow{PB}$, 知 $\overrightarrow{OP}-\overrightarrow{OA}=\lambda(\overrightarrow{OB}-\overrightarrow{OP})$. 因此

$$\overrightarrow{OP}=\frac{1}{1+\lambda}\overrightarrow{OA}+\frac{\lambda}{1+\lambda}\overrightarrow{OB}=\left(\frac{x_1+\lambda x_2}{1+\lambda},\frac{y_1+\lambda y_2}{1+\lambda},\frac{z_1+\lambda z_2}{1+\lambda}\right).$$

特别地, AB 中点的坐标为 $((x_1+x_2)/2,(y_1+y_2)/2,(z_1+z_2)/2)$. □

例1.2.3. 设 $A(x_1,y_1,z_1),B(x_2,y_2,z_2),C(x_3,y_3,z_3)$, 求 $\triangle ABC$ 重心 G 的坐标.

解. 设 BC 边的中点为 D, 则 D 点的坐标为

$$\left(\frac{x_2+x_3}{2},\frac{y_2+y_3}{2},\frac{z_2+z_3}{2}\right).$$

由例 1.1.3知, $\overrightarrow{AG}=2\overrightarrow{GD}$, 利用定比分点公式得

$$x=\frac{x_1+2\frac{x_2+x_3}{2}}{1+2}=\frac{x_1+x_2+x_3}{3},$$

同理

$$y=\frac{y_1+y_2+y_3}{3},\quad z=\frac{z_1+z_2+z_3}{3}.$$

所以, 重心 G 的坐标为

$$\left(\frac{x_1+x_2+x_3}{3},\frac{y_1+y_2+y_3}{3},\frac{z_1+z_2+z_3}{3}\right).$$

□

§1.2.3 直角坐标系

空间直角坐标系为一个特殊的仿射坐标系, 它要求三个坐标向量为两两垂直的单位向量. 一般用 $\mathbf{i},\mathbf{j},\mathbf{k}$ 表示这三个坐标向量, 相应的坐标轴为 x 轴, y 轴和 z 轴.

关于仿射坐标系的所有概念与结论都适用于直角坐标系, 而直角坐标系的特殊性使得利用向量的坐标来计算模长、夹角等变得容易.

设 $[O;\mathbf{i},\mathbf{j},\mathbf{k}]$ 为一个空间直角坐标系, 若向量 $\mathbf{a}=a_1\mathbf{i}+a_2\mathbf{j}+a_3\mathbf{k}$, 则

$$|\mathbf{a}|=\sqrt{a_1^2+a_2^2+a_3^2}. \tag{1.12}$$

设向量 $\mathbf{a}=(a_1,a_2,a_3)$ 与坐标向量$\mathbf{i}$, $\mathbf{j}$和$\mathbf{k}$的夹角分别为 α,β,γ, $\cos\alpha,\cos\beta,\cos\gamma$ 称为向量 $\mathbf{a}$ 的**方向余弦**.

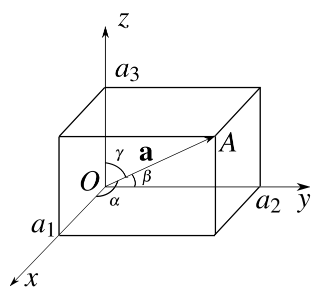


图 1.6 方向余弦

从图 1.6不难看出

$$(\cos\alpha,\cos\beta,\cos\gamma)=\left(\frac{a_1}{\sqrt{a_1^2+a_2^2+a_3^2}},\frac{a_2}{\sqrt{a_1^2+a_2^2+a_3^2}},\frac{a_3}{\sqrt{a_1^2+a_2^2+a_3^2}}\right). \tag{1.13}$$

从而

$$\cos^2\alpha+\cos^2\beta+\cos^2\gamma=1.$$

例1.2.4. 已知 $P(1,2,3), Q(2,4,-1)$, 求向量 $\overrightarrow{PQ}$ 的方向余弦.

解. 由于

$$\overrightarrow{PQ} = (1,2,-4)$$
$$|\overrightarrow{PQ}| = \sqrt{1^2+2^2+(-4)^2} = \sqrt{21},$$

所以 $\overrightarrow{PQ}$ 的方向余弦为 $(\frac{1}{\sqrt{21}}, \frac{2}{\sqrt{21}}, \frac{-4}{\sqrt{21}})$. □

§1.3 向量的数量积

§1.3.1 数量积的定义与性质

设一物体位于光滑水平面上, 力 $\mathbf{F}$ 作用于该物体上. 设 $\mathbf{F}$ 与水平面的夹角为 θ, 物体产生的位移是 $\mathbf{S}$, 则力 $\mathbf{F}$ 所作的功为

$$W = |\mathbf{F}||\mathbf{S}|\cos\theta.$$

功是由力 $\mathbf{F}$ 和位移 $\mathbf{S}$ 两个向量所唯一决定的一个数量. 除去其中的物理含义, 我们就得到了两个向量的数量积的概念.

定义1.3.1. 两个向量 $\mathbf{a}$ 与 $\mathbf{b}$ 的**数量积**为一个实数, 它等于两个向量的模长与两向量夹角的余弦的乘积, 记为 $\mathbf{a}\cdot\mathbf{b}$. 如果向量 $\mathbf{a},\mathbf{b}$ 的夹角为 θ, 则

$$\mathbf{a}\cdot\mathbf{b} = |\mathbf{a}||\mathbf{b}|\cos\theta. \tag{1.14}$$

数量积也常称为**内积**.

利用这个定义, 力所作的功为 $\mathbf{W} = \mathbf{F}\cdot\mathbf{S}$.

由数量积的定义知, $\mathbf{a}\cdot\mathbf{b} = 0$ 当且仅当两个向量是正交的(包含其中一个向量为零向量的情形).

数量积的运算具有如下的性质.

命题1.3.1. *对向量 $\boldsymbol{a},\boldsymbol{b},\boldsymbol{c}$ 及实数 λ, 我们有*

$$\boldsymbol{a}\cdot\boldsymbol{b} = \boldsymbol{b}\cdot\boldsymbol{a}, \tag{1.15}$$
$$(\boldsymbol{a}+\boldsymbol{b})\cdot\boldsymbol{c} = \boldsymbol{a}\cdot\boldsymbol{c}+\boldsymbol{b}\cdot\boldsymbol{c}, \tag{1.16}$$
$$(\lambda\boldsymbol{a})\cdot\boldsymbol{b} = \lambda(\boldsymbol{a}\cdot\boldsymbol{b}) = \boldsymbol{a}\cdot(\lambda\boldsymbol{b}), \tag{1.17}$$
$$\boldsymbol{a}^2 := \boldsymbol{a}\cdot\boldsymbol{a} \geq 0, \text{等号成立当且仅当 } \boldsymbol{a} = \boldsymbol{0}. \tag{1.18}$$

证明. 我们只证(1.16), 其余等式可利用数量积的定义直接验证.

不妨设 $\mathbf{c} \neq \mathbf{0}$. 如图 1.7, 作向量 $\overrightarrow{OA} = \mathbf{a}, \overrightarrow{AB} = \mathbf{b}, \overrightarrow{OC} = \mathbf{c}$, 则有 $\overrightarrow{OB} = \overrightarrow{OA} + \overrightarrow{AB} = \mathbf{a} + \mathbf{b}$. 过点 A,B 作直线 OC 的垂线, 垂足分别为 A',B'. 则存在实数 x,y 使得$\overrightarrow{OA'} = x\mathbf{c}$, $\overrightarrow{A'B'} = y\mathbf{c}$, $\overrightarrow{OB'} = (x+y)\mathbf{c}$.

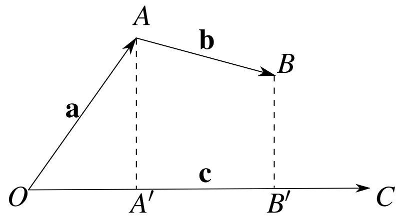


图 1.7

由数量积的定义可知

$$\mathbf{a} \cdot \mathbf{c} = x|\mathbf{c}|^2, \quad \mathbf{b} \cdot \mathbf{c} = y|\mathbf{c}|^2, \quad (\mathbf{a}+\mathbf{b}) \cdot \mathbf{c} = (x+y)|\mathbf{c}|^2.$$

所以 $(\mathbf{a}+\mathbf{b}) \cdot \mathbf{c} = \mathbf{a} \cdot \mathbf{c} + \mathbf{b} \cdot \mathbf{c}$. □

利用数量积的性质可以得到下面的等式

$$(\mathbf{a} \pm \mathbf{b})^2 = \mathbf{a}^2 \pm 2\mathbf{a} \cdot \mathbf{b} + \mathbf{b}^2. \tag{1.19}$$

由于

$$(\mathbf{a}+\mathbf{b})^2 = \mathbf{a}^2 + \mathbf{b}^2 + 2\mathbf{a} \cdot \mathbf{b} = |\mathbf{a}|^2 + |\mathbf{b}|^2 + 2|\mathbf{a}||\mathbf{b}|\cos\theta \leq (|\mathbf{a}| + |\mathbf{b}|)^2,$$

因此有向量模的三角不等式

$$|\mathbf{a}| + |\mathbf{b}| \geq |\mathbf{a}+\mathbf{b}|. \tag{1.20}$$

从(1.19)式还可以得到余弦定理的一个简单证明, 读者不妨一试.

下面我们通过一个例子来看向量的数量积在求解几何问题中的应用.

例1.3.1. 四面体 $OABC$ 中, $OA \perp BC, OB \perp AC$, 证明: $OC \perp AB$.

证明. 记 $\overrightarrow{OA} = \mathbf{a}, \overrightarrow{OB} = \mathbf{b}, \overrightarrow{OC} = \mathbf{c}$, 则

$$\overrightarrow{BC} = \mathbf{c} - \mathbf{b}, \quad \overrightarrow{CA} = \mathbf{a} - \mathbf{c}, \quad \overrightarrow{AB} = \mathbf{b} - \mathbf{a}.$$

由 $OA \perp BC$ 知, $\mathbf{a} \cdot (\mathbf{c} - \mathbf{b}) = 0$, 即 $\mathbf{a} \cdot \mathbf{c} = \mathbf{a} \cdot \mathbf{b}$. 同理, $OB \perp AC$ 蕴含了 $\mathbf{b} \cdot \mathbf{c} = \mathbf{b} \cdot \mathbf{a}$. 因此 $\mathbf{a} \cdot \mathbf{c} - \mathbf{b} \cdot \mathbf{c} = 0$, 即 $(\mathbf{a} - \mathbf{b}) \cdot \mathbf{c} = 0$, 所以 $OC \perp AB$.

□

§1.3.2 直角坐标系下数量积的计算

设 $[O;\mathbf{i},\mathbf{j},\mathbf{k}]$ 为一个直角坐标系. 由于向量 $\mathbf{i},\mathbf{j},\mathbf{k}$ 为两两垂直的单位向量, 所以

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

给定两个向量 $\mathbf{a}=a_1\mathbf{i}+a_2\mathbf{j}+a_3\mathbf{k},\quad \mathbf{b}=b_1\mathbf{i}+b_2\mathbf{j}+b_3\mathbf{k}$, 则有

$$\begin{aligned}\mathbf{a}\cdot\mathbf{b}&=(a_1\mathbf{i}+a_2\mathbf{j}+a_3\mathbf{k})\cdot(b_1\mathbf{i}+b_2\mathbf{j}+b_3\mathbf{k})\\&=a_1b_1(\mathbf{i}\cdot\mathbf{i})+a_1b_2(\mathbf{i}\cdot\mathbf{j})+a_1b_3(\mathbf{i}\cdot\mathbf{k})\\&\quad+a_2b_1(\mathbf{j}\cdot\mathbf{i})+a_2b_2(\mathbf{j}\cdot\mathbf{j})+a_2b_3(\mathbf{j}\cdot\mathbf{k})\\&\quad+a_3b_1(\mathbf{k}\cdot\mathbf{i})+a_3b_2(\mathbf{k}\cdot\mathbf{j})+a_3b_3(\mathbf{k}\cdot\mathbf{k})\\&=a_1b_1+a_2b_2+a_3b_3.\end{aligned}$$

所以有

$$(a_1,a_2,a_3)\cdot(b_1,b_2,b_3)=a_1b_1+a_2b_2+a_3b_3. \tag{1.21}$$

就是说, 两个向量的数量积等于它们对应坐标乘积之和.

设向量 $\mathbf{a},\mathbf{b}$ 之间的夹角为 θ, 则由数量积的定义知

$$\cos\theta=\frac{\mathbf{a}\cdot\mathbf{b}}{|\mathbf{a}||\mathbf{b}|}=\frac{a_1b_1+a_2b_2+a_3b_3}{\sqrt{a_1^2+a_2^2+a_3^2}\sqrt{b_1^2+b_2^2+b_3^2}}. \tag{1.22}$$

例1.3.2. 求一个单位向量, 使它与向量 $\mathbf{a}=-\mathbf{i}+2\mathbf{j}+\mathbf{k}$, $\mathbf{b}=\mathbf{i}+3\mathbf{k}$ 都垂直.

解. 设所求的单位向量为 $\mathbf{x}=x_1\mathbf{i}+x_2\mathbf{j}+x_3\mathbf{k}$, 由条件知

$$\mathbf{x}\cdot\mathbf{a}=0,\quad \mathbf{x}\cdot\mathbf{b}=0,\quad |\mathbf{x}|=1.$$

用坐标表示, 即得

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

解得

$$(x_1,x_2,x_3)=\pm(\frac{3}{\sqrt{14}},\frac{2}{\sqrt{14}},\frac{-1}{\sqrt{14}}).$$

所以

$$\mathbf{x}=\pm\frac{1}{\sqrt{14}}(3\mathbf{i}+2\mathbf{j}-\mathbf{k}).$$

□

例1.3.3. 证明 Cauchy 不等式

$$(a_1b_1+a_2b_2+a_3b_3)^2 \le (a_1^2+a_2^2+a_3^2)(b_1^2+b_2^2+b_3^2).$$

证明. 设 $\mathbf{a}=(a_1,a_2,a_3)$, $\mathbf{b}=(b_1,b_2,b_3)$, 则有

$$\mathbf{a}\cdot\mathbf{b}=a_1b_1+a_2b_2+a_3b_3,$$
$$|\mathbf{a}|^2=a_1^2+a_2^2+a_3^2,\quad |\mathbf{b}|^2=b_1^2+b_1^2+b_3^2.$$

由数量积的定义 $\mathbf{a}\cdot\mathbf{b}=|\mathbf{a}||\mathbf{b}|\cos\theta$ 及 $|\cos\theta|\le 1$ 推知 $(\mathbf{a}\cdot\mathbf{b})^2\le|\mathbf{a}|^2|\mathbf{b}|^2$. 把它写成坐标的形式, 即得 Cauchy 不等式. □

§1.4 向量的向量积

§1.4.1 向量积的定义与性质

设刚体以等角速度 ω 绕定轴转动, O 为转动轴上的一个定点, M 为刚体上的一个点, 到转动轴的距离为 R, 记向量 $\mathbf{r}=\overrightarrow{OM}$(图 1.8). 由物理学知道, 点 M 处的速度向量 $\mathbf{v}$ 的方向与定轴及点 M 决定的平面垂直, 大小等于 $\omega\cdot R$. 为了更好地描述这一特征, 在转动轴上引一向量 ω, 称为角速度向量, 它的模 $|\omega|=\omega$, 它的方向与转动的方向构成右手螺旋系统, 即向量 $\omega,\mathbf{r},\mathbf{v}$ 满足右手法则. 若记向量 ω 与 $\mathbf{r}$ 的夹角为 θ, 则

$$|\mathbf{v}|=\omega R=|\omega||\mathbf{r}|\sin\theta,$$

象这样由两个向量 $\omega,\mathbf{r}$ 决定第三个向量 $\mathbf{v}$ 的现象在物理学中十分常见. 将其抽象化, 就得到了两个向量的向量积的概念.

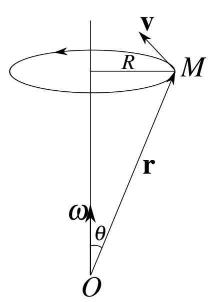


图 1.8 向量的外积

定义1.4.1. 两个向量 $\mathbf{a},\mathbf{b}$ 的**向量积** $\mathbf{a}\times\mathbf{b}$ 为一个向量, 它的方向与 $\mathbf{a},\mathbf{b}$ 都垂直, 且使 $\mathbf{a},\mathbf{b},\mathbf{a}\times\mathbf{b}$ 构成右手系；它的模等于以 $\mathbf{a},\mathbf{b}$ 为边的平行四边形的面积, 即 $|\mathbf{a}\times\mathbf{b}|=|\mathbf{a}||\mathbf{b}|\sin\theta$, 其中 θ 为 $\mathbf{a},\mathbf{b}$ 的夹角.

向量的向量积运算具有以下性质.

命题1.4.1. *设 $\boldsymbol{a},\boldsymbol{b},\boldsymbol{c}$ 为三个向量, λ 为实数, 则有*

$$\boldsymbol{a}\times\boldsymbol{b}=-\boldsymbol{b}\times\boldsymbol{a}, \tag{1.23}$$

$$(\lambda\boldsymbol{a})\times\boldsymbol{b}=\lambda(\boldsymbol{a}\times\boldsymbol{b})=\boldsymbol{a}\times(\lambda\boldsymbol{b}), \tag{1.24}$$

$$(\boldsymbol{a}+\boldsymbol{b})\times\boldsymbol{c}=\boldsymbol{a}\times\boldsymbol{c}+\boldsymbol{b}\times\boldsymbol{c}. \tag{1.25}$$

证明. 我们只证(1.25), 为此首先来看当 $\mathbf{c}$ 为一个单位向量时, 向量积 $\mathbf{a}\times\mathbf{c}$ 的几何作图法.

设 $\mathbf{a}$ 与 $\mathbf{c}$ 有共同的起点 O, $\mathbf{a}$ 与 $\mathbf{c}$ 的夹角为 θ. 把向量 $\mathbf{a}$ 投影到与 $\mathbf{c}$ 垂直的平面上得向量 $\mathbf{a}_1$, 再将 $\mathbf{a}_1$ 绕$\mathbf{c}$顺时针旋转 $\dfrac{\pi}{2}$ 得向量 $\mathbf{a}_2$, 使得 $\mathbf{a},\mathbf{c},\mathbf{a}_2$ 组成右手系(图1.9(a)).

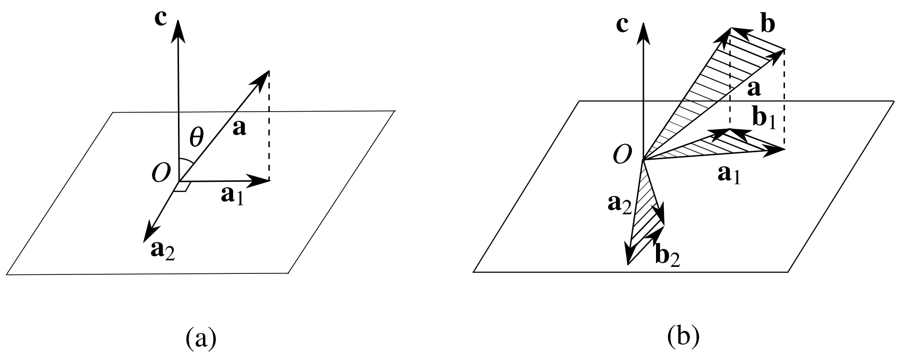


图 1.9

由于

$$|\mathbf{a}_2|=|\mathbf{a}_1|=|\mathbf{a}|\sin\theta=|\mathbf{a}\times\mathbf{c}|,$$

因此

$$\mathbf{a}_2=\mathbf{a}\times\mathbf{c}.$$

现在设 $\mathbf{c}$ 为单位向量, 将 $\mathbf{a}$ 平移使它与 $\mathbf{c}$ 有共同起点 O, 过 $\mathbf{a}$ 的终点引向量 $\mathbf{b}$, 按三角形法则得向量 $\mathbf{a}+\mathbf{b}$(图 1.9(b)).将向量 $\mathbf{a},\mathbf{b},\mathbf{a}+\mathbf{b}$ 组成的三角形投影到与 $\mathbf{c}$ 垂直的平面中, 得到一个由向量 $\mathbf{a}_1,\mathbf{b}_1,\mathbf{a}_1+\mathbf{b}_1$ 组成的三角形. 再将所得的三角形在

这个平面上绕 O 点顺时针方向旋转 $\dfrac{\pi}{2}$ 角度, 得到一个由向量 $\mathbf{a}_2, \mathbf{b}_2, \mathbf{a}_2+\mathbf{b}_2$ 组成的三角形.

根据上述向量积的几何作图法, 我们得到

$$\mathbf{a}\times\mathbf{c}=\mathbf{a}_2,\quad \mathbf{b}\times\mathbf{c}=\mathbf{b}_2,\quad (\mathbf{a}+\mathbf{b})\times\mathbf{c}=\mathbf{a}_2+\mathbf{b}_2,$$

所以 $(\mathbf{a}+\mathbf{b})\times\mathbf{c}=\mathbf{a}\times\mathbf{c}+\mathbf{b}\times\mathbf{c}$.

进而考察一般情形. 因为 $\mathbf{c}=|\mathbf{c}|\mathbf{c}^0$, 其中$\mathbf{c}^0$为与$\mathbf{c}$同向的单位向量, 利用上式及性质 (2), 得

$$\begin{aligned}(\mathbf{a}+\mathbf{b})\times(|\mathbf{c}|\mathbf{c}^0)&=|\mathbf{c}|\left((\mathbf{a}+\mathbf{b})\times\mathbf{c}^0\right)\\&=|\mathbf{c}|\left(\mathbf{a}\times\mathbf{c}^0+\mathbf{b}\times\mathbf{c}^0\right)\\&=\mathbf{a}\times(|\mathbf{c}|\mathbf{c}^0)+\mathbf{b}\times(|\mathbf{c}|\mathbf{c}^0)),\end{aligned}$$

即

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

$\square$

§1.4.2 直角坐标系下向量积的计算

设 $[O;\mathbf{i},\mathbf{j},\mathbf{k}]$ 为一个直角坐标系. 由于 $\mathbf{i},\mathbf{j},\mathbf{k}$ 为两两垂直的单位向量且满足右手法则, 由向量积的定义知

$$\begin{aligned}&\mathbf{i}\times\mathbf{i}=\mathbf{0},\quad \mathbf{j}\times\mathbf{j}=\mathbf{0},\quad \mathbf{k}\times\mathbf{k}=\mathbf{0}\\&\mathbf{i}\times\mathbf{j}=\mathbf{k},\quad \mathbf{j}\times\mathbf{k}=\mathbf{i},\quad \mathbf{k}\times\mathbf{i}=\mathbf{j}.\end{aligned}$$

给定两个向量

$$\mathbf{a}=a_1\mathbf{i}+a_2\mathbf{j}+a_3\mathbf{k},\quad \mathbf{b}=b_1\mathbf{i}+b_2\mathbf{j}+b_3\mathbf{k},$$

则有

$$\begin{aligned}\mathbf{a}\times\mathbf{b}=&(a_1\mathbf{i}+a_2\mathbf{j}+a_3\mathbf{k})\times(b_1\mathbf{i}+b_2\mathbf{j}+b_3\mathbf{k})\\=&a_1b_1(\mathbf{i}\times\mathbf{i})+a_1b_2(\mathbf{i}\times\mathbf{j})+a_1b_3(\mathbf{i}\times\mathbf{k})\\&+a_2b_1(\mathbf{j}\times\mathbf{i})+a_2b_2(\mathbf{i}\times\mathbf{j})+a_2b_3(\mathbf{j}\times\mathbf{k})\\&+a_3b_1(\mathbf{k}\times\mathbf{i})+a_3b_2(\mathbf{k}\times\mathbf{j})+a_3b_3(\mathbf{k}\times\mathbf{k})\\=&(a_2b_3-a_3b_2)\mathbf{i}+(a_3b_1-a_1b_3)\mathbf{j}+(a_1b_2-a_2b_1)\mathbf{k}\end{aligned}$$

为了便于记忆上式, 我们引入 2 阶和 3 阶行列式的概念.

把 4 个数排成两行两列, 在两边加上两条竖线, 就得到一个 2 阶行列式. 它代表一个数, 其运算规则是

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

类似地, 9 个数排成 3 行 3 列, 就得到一个 3 阶行列式. 它的运算规则是

$$\begin{aligned}\begin{vmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{vmatrix} &= a_1\begin{vmatrix} b_2 & b_3 \\ c_2 & c_3 \end{vmatrix} - a_2\begin{vmatrix} b_1 & b_3 \\ c_1 & c_3 \end{vmatrix} + a_3\begin{vmatrix} b_1 & b_3 \\ c_1 & c_3 \end{vmatrix} \\ &= a_1b_2c_3 + a_2b_3c_1 + a_3b_1c_2 - a_1b_3c_2 - a_2b_1c_3 - a_3b_2c_1.\end{aligned}$$

关于行列式的详细介绍, 请参看第三章.

利用行列式的语言, 前面关于向量积的运算公式可以改写为

$$\begin{aligned}\mathbf{a}\times\mathbf{b} &= \begin{vmatrix} a_2 & a_3 \\ b_2 & b_3 \end{vmatrix}\mathbf{i} - \begin{vmatrix} a_1 & a_3 \\ b_1 & b_3 \end{vmatrix}\mathbf{j} + \begin{vmatrix} a_1 & a_2 \\ b_1 & b_2 \end{vmatrix}\mathbf{k} \\ &= \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \end{vmatrix}.\end{aligned} \tag{1.26}$$

例1.4.1. 求垂直于向量 $\mathbf{a} = (-1,2,1)$ 和 $\mathbf{b} = (1,0,3)$ 的单位向量.

解. $\mathbf{a}\times\mathbf{b}$ 就是一个垂直于 $\mathbf{a}$ 和 $\mathbf{b}$ 的向量, 而

$$\begin{aligned}\mathbf{a}\times\mathbf{b} = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ -1 & 2 & 1 \\ 1 & 0 & 3 \end{vmatrix} &= \begin{vmatrix} 2 & 1 \\ 0 & 3 \end{vmatrix}\mathbf{i} - \begin{vmatrix} -1 & 1 \\ 1 & 3 \end{vmatrix}\mathbf{j} + \begin{vmatrix} -1 & 2 \\ 1 & 0 \end{vmatrix}\mathbf{k} \\ &= 6\mathbf{i} + 4\mathbf{j} - 2\mathbf{k}. \\ |\mathbf{a}\times\mathbf{b}| &= \sqrt{6^2+4^2+(-2)^2} = 2\sqrt{14}\end{aligned}$$

因此所求的单位向量为

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

与例 1.3.2相比, 这里的解法更为简单.

□

例1.4.2. 设三角形的三个定点为 $A(1,1,2), B(-2,0,3), C(2,4,5)$, 求 $\triangle ABC$ 的面积以及 BC 边上的高 h.

解. 设 $\triangle ABC$ 的面积为 S, 由向量积的几何意义知 $S=\dfrac{1}{2}|\overrightarrow{AB}\times\overrightarrow{AC}|$. 而 $\overrightarrow{AB}=(-3,-1,1)$, $\overrightarrow{AC}=(1,3,3)$, 所以

$$\overrightarrow{AB}\times\overrightarrow{AC}=\begin{vmatrix}\mathbf{i} & \mathbf{j} & \mathbf{k}\\ -3 & -1 & 1\\ 1 & 3 & 3\end{vmatrix}=-6\mathbf{i}+10\mathbf{j}-8\mathbf{k}.$$

因此

$$S=\frac{1}{2}\sqrt{(-6)^2+10^2+(-8)^2}=5\sqrt{2}.$$

另一方面, 由 $S=\dfrac{1}{2}|\overrightarrow{BC}|\cdot h$ 及 $\overrightarrow{BC}=(4,4,2)$ 可解得 $h=\dfrac{5\sqrt{2}}{3}$. □

§1.5 向量的混合积

§1.5.1 混合积的定义

给定三个向量 $\mathbf{a},\mathbf{b},\mathbf{c}$, 称 $(\mathbf{a}\times\mathbf{b})\cdot\mathbf{c}$ 为 $\mathbf{a},\mathbf{b},\mathbf{c}$的**混合积**.

以 $\mathbf{a},\mathbf{b},\mathbf{c}$ 为棱的平行六面体的体积 V 等于以 $\mathbf{a},\mathbf{b}$ 为边的平行四边形的面积 S 乘以高 h(图 1.10), 即$V=Sh$. 由向量积的定义知 $S=|\mathbf{a}\times\mathbf{b}|$, 另一方面, 设 $\mathbf{a}\times\mathbf{b}$ 与 $\mathbf{c}$ 的夹角为 φ, 则有 $|h|=|\mathbf{c}||\cos\varphi|$. 于是

$$V=|\mathbf{a}\times\mathbf{b}||\mathbf{c}||\cos\varphi|=|(\mathbf{a}\times\mathbf{b})\cdot\mathbf{c}|.$$

注意到 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$(\mathbf{a}\times\mathbf{b})\cdot\mathbf{c}$ 表示的是以 $\mathbf{a},\mathbf{b},\mathbf{c}$ 为棱的平行六面体的“有向体积”. 即:当 $\mathbf{a},\mathbf{b},\mathbf{c}$ 为右手系时, 就是六面体的体积; 当 $\mathbf{a},\mathbf{b},\mathbf{c}$ 为左手系时, 它是六面体体积的相反数.

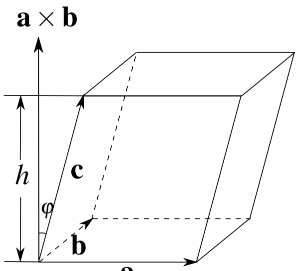


图 1.10 向量的混合积

由于轮换 $\mathbf{a},\mathbf{b},\mathbf{c}$ 的次序时, 不会改变左右手系, 因此混合积的值不变. 即

$$(\mathbf{a}\times\mathbf{b})\cdot\mathbf{c}=(\mathbf{b}\times\mathbf{c})\cdot\mathbf{a}=(\mathbf{c}\times\mathbf{a})\cdot\mathbf{b}. \tag{1.27}$$

另一方面, 交换 $\mathbf{a},\mathbf{b},\mathbf{c}$ 中任意两个的次序会改变左右手系, 因此混合积的值改变符号, 例如

$$(\mathbf{a}\times\mathbf{b})\cdot\mathbf{c}=-(\mathbf{b}\times\mathbf{a})\cdot\mathbf{c}. \tag{1.28}$$

此外, 当 $\mathbf{a},\mathbf{b},\mathbf{c}$ 中任意两个平行时, 混合积为零, 例如

$$(\mathbf{a}\times\mathbf{b})\cdot\mathbf{a}=(\mathbf{a}\times\mathbf{b})\cdot\mathbf{b}=0.$$

§1.5.2 直角坐标系下混合积的计算

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

$$\mathbf{a}\times\mathbf{b}=\begin{vmatrix}\mathbf{i}&\mathbf{j}&\mathbf{k}\\a_1&a_2&a_3\\b_1&b_2&b_3\end{vmatrix}=\begin{vmatrix}a_2&a_3\\b_2&b_3\end{vmatrix}\mathbf{i}-\begin{vmatrix}a_1&a_3\\b_1&b_3\end{vmatrix}\mathbf{j}+\begin{vmatrix}a_1&a_2\\b_1&b_2\end{vmatrix}\mathbf{k},$$

所以

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

即

$$(\mathbf{a}\times\mathbf{b})\cdot\mathbf{c}=\begin{vmatrix}a_1&a_2&a_3\\b_1&b_2&b_3\\c_1&c_2&c_3\end{vmatrix}. \tag{1.29}$$

命题1.5.1. 三个向量 $\boldsymbol{a}=(a_1,a_2,a_3),\boldsymbol{b}=(b_1,b_2,b_3),\boldsymbol{c}=(c_1,c_2,c_3)$ 共面当且仅当

$$\begin{vmatrix}a_1&a_2&a_3\\b_1&b_2&b_3\\c_1&c_2&c_3\end{vmatrix}=0.$$

证明. 由于混合积 $(\mathbf{a}\times\mathbf{b})\cdot\mathbf{c}$ 表示的是 $\mathbf{a},\mathbf{b},\mathbf{c}$ 构成的平行六面体的有向体积, $\mathbf{a},\mathbf{b},\mathbf{c}$ 共面当且仅当 $(\mathbf{a}\times\mathbf{b})\cdot\mathbf{c}=0$. 由上面的公式即得命题. □

例1.5.1. 设四面体的四个顶点为 $A(1,2,3),B(2,1,4),C(1,3,5),D(3,2,1)$, 求该四面体的体积.

解. 所求四面体的体积 V 是以 $\overrightarrow{AB},\overrightarrow{AC},\overrightarrow{AD}$ 为棱的平行六面体的体积的六分之一, 故

$$V=\frac{1}{6}|(\overrightarrow{AB}\times\overrightarrow{AC})\cdot\overrightarrow{AD}|.$$

由于 $\overrightarrow{AB}=(1,-1,1), \overrightarrow{AC}=(0,1,2), \overrightarrow{AD}=(2,0,-2)$, 所以

$$(\overrightarrow{AB}\times\overrightarrow{AC})\cdot\overrightarrow{AD}=\begin{vmatrix}1 & -1 & 1\\ 0 & 1 & 2\\ 2 & 0 & -2\end{vmatrix}=-8.$$

于是得到 $V=\dfrac{4}{3}$. □

§1.5.3 二重外积

给定三个向量 $\mathbf{a},\mathbf{b},\mathbf{c}$, 称 $(\mathbf{a}\times\mathbf{b})\times\mathbf{c}$ 为这三个向量的二重外积.

命题1.5.2. 对任意向量 $\boldsymbol{a},\boldsymbol{b},\boldsymbol{c}$, 有

$$(\boldsymbol{a}\times\boldsymbol{b})\times\boldsymbol{c}=(\boldsymbol{a}\cdot\boldsymbol{c})\boldsymbol{b}-(\boldsymbol{b}\cdot\boldsymbol{c})\boldsymbol{a}.$$

证明. 取一个右手直角坐标系, 设

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

只要验证等式两边的向量具有相同的坐标即可. 由于

$$\mathbf{a}\times\mathbf{b}=(a_2b_3-a_3b_2,\, a_3b_1-a_1b_3,\, a_1b_2-a_2b_1),$$

所以$(\mathbf{a}\times\mathbf{b})\times\mathbf{c}$ 的第一个坐标为

$$(a_3b_1-a_1b_3)c_3-(a_1b_2-a_2b_1)c_2.$$

另一方面, $(\mathbf{a}\cdot\mathbf{c})\mathbf{b}-(\mathbf{b}\cdot\mathbf{c})\mathbf{a}$ 的第一个坐标为

$$\begin{aligned}&(a_1c_1+a_2c_2+a_3c_3)b_1-(b_1c_1+b_2c_2+b_3c_3)a_1\\ &=(a_3b_1-a_1b_3)c_3-(a_1b_2-a_2b_1)c_2.\end{aligned}$$

因此等式两边向量的第一个坐标相同, 同理另两个坐标相同, 从而等式成立. □

上述公式通常称作二重外积展开式, 从这个公式可以看出, 外积不满足结合律, 就是说, 一般情况下,

$$(\mathbf{a}\times\mathbf{b})\times\mathbf{c}\neq\mathbf{a}\times(\mathbf{b}\times\mathbf{c}),$$

因为上式左边是 $\mathbf{a}$ 和 $\mathbf{b}$ 的线性组合, 而右边是 $\mathbf{b}$ 和 $\mathbf{c}$ 的线性组合.

例1.5.2. 证明: $(\mathbf{a}\times\mathbf{b})\cdot(\mathbf{c}\times\mathbf{d})=(\mathbf{a}\cdot\mathbf{c})(\mathbf{b}\cdot\mathbf{d})-(\mathbf{a}\cdot\mathbf{d})(\mathbf{b}\cdot\mathbf{c})$.

证明.

$$\begin{aligned}(\mathbf{a}\times\mathbf{b})\cdot(\mathbf{c}\times\mathbf{d}) &= ((\mathbf{a}\times\mathbf{b})\times\mathbf{c})\cdot\mathbf{d}\\ &= ((\mathbf{a}\cdot\mathbf{c})\mathbf{b}-(\mathbf{b}\cdot\mathbf{c})\mathbf{a})\cdot\mathbf{d}\\ &= (\mathbf{a}\cdot\mathbf{c})(\mathbf{b}\cdot\mathbf{d})-(\mathbf{a}\cdot\mathbf{d})(\mathbf{b}\cdot\mathbf{c}).\end{aligned}$$

□

§1.6 高维数组向量

在现实生活中, 很多量无法在三维空间中表示, 如物体在某点的温度、密度、压力等. 因为物体的每一点的位置要用一个三维坐标表示, 温度, 密度等量还要一个坐标表示. 因此, 表示物体在某一点的温度(或密度)要四个坐标. 而表示物体一点的速度则要六个坐标. 因此, 有必要将三维向量推广到高维向量.

取定一个坐标系, 一个空间向量可以用一个三元数组(坐标)来表示. 向量的线性运算(加法与数乘)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1.6.1. 一个n维**数组向量** $\mathbf{a}$是一个有序的n元数组

$$\mathbf{a}=(a_1,a_2,\ldots,a_n), \tag{1.30}$$

其中$a_i\in F(i=1,2,\ldots,n)$称为向量$\mathbf{a}$的第i个**分量**. 这里F表示实数集、复数集或其它数域(定义见1.8节).

n维数组向量有时需要写成行的形式, 如$\mathbf{a}=(a_1,a_2,\ldots,a_n)$, 称为**行向量**, 有时需要写成列的形式, 如

$$\mathbf{a}=\begin{pmatrix}a_1\\a_2\\\vdots\\a_n\end{pmatrix} \tag{1.31}$$

称为**列向量**. 虽然形式上看上去不同, 但我们将它们认为是一样的.

给定两个n维数组向量$\mathbf{a}=(a_1,a_2,\ldots,a_n)$与$\mathbf{b}=(b_1,b_2,\ldots,b_n)$, 规定$\mathbf{a}$与$\mathbf{b}$ **相等**, 当且仅当它们对应的分量分别相等, 即$a_i=b_i$, $i=1,2,\ldots,n$.

定义n维数组向量的加法与数乘运算如下

$$\begin{aligned}&\mathbf{a}+\mathbf{b}=(a_1+b_1,a_2+b_2,\ldots,a_n+b_n),\\ &\lambda\mathbf{a}=(\lambda a_1,\lambda a_2,\ldots,\lambda a_n).\end{aligned} \tag{1.32}$$

即按分量分别执行加法与数乘运算. 规定**零向量**$\mathbf{0}=(0,\ldots,0)$, **负向量** $-\mathbf{a}=(-a_1,\ldots,-a_n)$. 容易验证, 数组向量的加法与数乘运算满足几何向量的性质(1.1)–(1.8).

定义了向量的加法与数乘, 就可以引进向量的一种重要的运算-线性组合.

定义1.6.2. 给定一组n维数组向量$\mathbf{a}_1,\mathbf{a}_2,\ldots,\mathbf{a}_m$及一组数$\lambda_1,\lambda_2,\ldots,\lambda_m$, 称和式

$$\lambda_1\mathbf{a}_1+\lambda_2\mathbf{a}_2+\ldots+\lambda_m\mathbf{a}_m$$

为$\mathbf{a}_1,\mathbf{a}_2,\ldots,\mathbf{a}_m$的**线性组合**. $\lambda_1,\lambda_2,\ldots,\lambda_m$称为**组合系数**. 如果$\mathbf{a}$可以写成$\mathbf{a}_1,\mathbf{a}_2,\ldots,\mathbf{a}_m$的线性组合, 则称$\mathbf{a}$可以用$\mathbf{a}_1,\mathbf{a}_2,\ldots,\mathbf{a}_m$ **线性表示**.

用$\mathbf{e}_i=(0,\ldots,1,\ldots,0)$表示第$i$个分量为1, 其余分量为0的$n$维数组向量, $\mathbf{e}_1,\ldots,\mathbf{e}_n$称为**单位坐标向量**. 任何一个$n$维数组向量都可以表示为单位坐标向量的线性组合. 实际上, 设$\mathbf{a}=(a_1,a_2,\ldots,a_n)$, 则

$$\mathbf{a}=a_1\mathbf{e}_1+a_2\mathbf{e}_2+\ldots+a_n\mathbf{e}_n.$$

关于数组向量的一般理论我们将在第五章详细研究.

§1.7 集合与映射*

集合论是由数学家Cantor于十九世纪建立, 现已成为现代数学的基础. 中学阶段我们已经学习了集合与映射的基本概念及运算, 本节对其做一个简单的回顾.

§1.7.1 集合的定义*

将一些不同的对象放在一起, 即为**集合**, 其中的对象称为集合的**元素**. 在本书中, 我们将使用大写字母$A,B,C,\ldots$来表示集合, 用小写字母$a,b,c,\ldots$来表示集合中的元素. 记A为一个集合. 如果a是A中的元素, 则称a属于A, 记为$a\in A$, 否则记为$a\notin A$. 本书中我们总是假设集合中元素是不重复的.

集合的描述通常有两者方式. 一种是列出集合中的所有元素, 并用花括号将这些元素括起来, 例如$A=\{1,3,5,7,9\}$ 表示由1,3,5,7,9 五个整数构成的集合. 另一种方式是通过语言描述来定义集合, 如**偶数集合**=$\{a$ 为整数 $|$ a被2 整除$\}$. 我们通常分别用$\mathbb{N}$, $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$, $\mathbb{C}$ 表示自然数、整数、有理数、实数及复数集合. 平面上以原点为圆心、半径为1 的圆可以表示为满足以下条件的点的集合$C=\{(x,y)\in\mathbb{R}^2\,|\,x^2+y^2=1\}$.

如果集合 A 中的每一个元素均是集合 B 中的元素, 则称 A 是 B 的**子集**, 记为 $A \subseteq B$ 或 $B \supseteq A$. 图 1.11 表示了集合A包含于集合B, 即$A \subseteq B$. 如果 $A \subseteq B$ 且 $A \neq B$, 则称 A 为 B 的**真子集**, 记为 $A \subset B$ 或者$A \subsetneq B$. 如果集合A与B都包含相同的元素, 则称 A 与 B **相等**, 记为 $A = B$. 易知, 集合A与B相等当且仅当 $A \subseteq B$ 且 $B \subseteq A$.

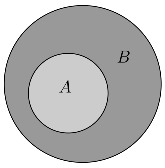


图 1.11 集合的包含

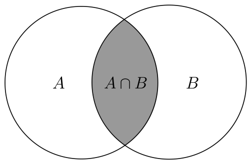


图 1.12 集合的交

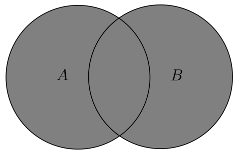


图 1.13 集合的并

我们可以用一个简单例子来理解集合.

- 班级 $\longleftrightarrow$ 集合,
- 班上的学生 $\longleftrightarrow$ 元素,
- 班上的一个学习小组 $\longleftrightarrow$ 子集合,
- 学校的所有班级 $\longleftrightarrow$ 集合构成的集族.

不含任何元素的集合称为**空集**, 记为 $\varnothing$. 由定义可知, 空集 $\varnothing$ 是任何集合的子集, 且是任何非空集合的真子集.

如果集合A的元素个数有限, 称A 为**有限集**, 其元素个数称为集合的**阶**或**基数**, 记为$|A|$. 元素个数无限的集合称为**无限集**, 它的阶定义为 ∞.

§1.7.2 集合的基本运算*

一般来说, 对于某固定集合U 的子集合, 有如下四种基本运算.

(1) **集合的交** 设 A,B 为U 的两个子集合, 则A **与** B **的交集**定义为

$$A \cap B := \{x \mid x \in A \text{ 且} x \in B\}.$$

如图 1.12所示. 在上式中, 记号$:=$ 表示的是将其右边的集合记作$A \cap B$.

更一般地, 设 I 为集合, I 中每个元素 i 对应U 的子集合 A_i, 则集合族 $A_i (i \in I)$ 的交定义为

$$\bigcap_{i \in I} A_i := \{x \mid x \in A_i, \text{对每个} i \in I \text{成立}\}.$$

(2) **集合的并** 设 A,B 是集合U的子集合. 则A **与** B **的并集**定义为

$$A \cup B := \{x \mid x \in A \text{ 或} x \in B\}.$$

如图 1.13所示. 更一般地, 集合族 $A_i(i\in I)$ 的并定义为

$$\bigcup_{i\in I}A_i := \{x \mid x\in A_i, \text{对某个}i\in I\text{ 成立}\}.$$

如果 A_i 两两不交(即交集为空集), 我们称 $\bigcup\limits_{i\in I}A_i$ 为**不交并**, 并记为 $\bigsqcup\limits_{i\in I}A_i$.

命题1.7.1. 设A,B,C为集合U的子集, 则

$$A\cap(B\cup C)=(A\cap B)\cup(A\cap C),\quad A\cup(B\cap C)=(A\cup B)\cap(A\cup C). \tag{1.33}$$

证明. 我们只证前一式, 后一式类似可证. 设$x\in A\cap(B\cup C)$, 则$x\in A$ 且$x\in B\cup C$. 若$x\in B$, 则$x\in A\cap B$, 于是$x\in(A\cap B)\cup(A\cap C)$. 若$x\in C$, 则$x\in A\cap C$, 同样有$x\in(A\cap B)\cup(A\cap C)$. 故$A\cap(B\cup C)\subset(A\cap B)\cup(A\cap C)$.

反之, 设$x\in(A\cap B)\cup(A\cap C)$, 则$x\in A\cap B$ 或$x\in A\cap C$. 无论哪种情形都有$x\in A$, $x\in B\cup C$. 故$x\in A\cap(B\cup C)$. 从而$(A\cap B)\cup(A\cap C)\subset A\cap(B\cup C)$. 证毕. □

(3) **集合的差集与补集** 设 A,B 为 U 的子集, 则A **对** B **的补集或差集** 定义为

$$A-B=A\backslash B:=\{x\mid x\in A\text{ 且}x\notin B\}.$$

如图 1.14 所示. 由补集定义, 我们有

$$A=(A\cap B)\sqcup(A-B).$$

A **在** U **中的补集**定义为

$$A^c:=\{x\in U\mid x\notin A\}.$$

如图 1.15 所示.

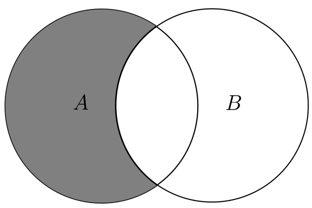


图 1.14 集合的差集$A-B$

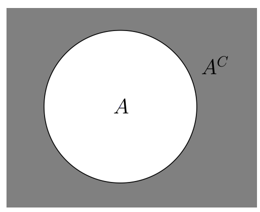


图 1.15 集合的补集A^c

关于补集运算, 有以下结论:

命题1.7.2. 设 A_i $(i \in I)$ 为某固定集合 U 的子集, 则

$$\bigcap_{i\in I} A_i^c = \Big(\bigcup_{i\in I} A_i\Big)^c. \tag{1.34}$$

通俗地说, 补集的交等于并集的补.

证明. 我们有

$$\begin{aligned} x \in \bigcap_{i\in I} A_i^c &\Longleftrightarrow x \in A_i^c \text{ 对任意} i \in I \text{ 成立} \\ &\Longleftrightarrow x \notin A_i \text{ 对任意} i \in I \text{ 成立} \\ &\Longleftrightarrow x \notin \bigcup_{i\in I} A_i,\ \text{即 } x \in \Big(\bigcup_{i\in I} A_i\Big)^c. \end{aligned}$$

□

关于两个集合并集的阶数有以下结论.

命题1.7.3. *(容斥原理)* 如果 A,B 为有限集合, 则 $A\cup B$ 和$A\cap B$ 均为有限集, 且

$$|A\cup B| = |A| + |B| - |A\cap B|. \tag{1.35}$$

证明. 易证

$$A\cup B = (A-B)\sqcup B, \qquad A = (A-B)\sqcup(A\cap B).$$

故

$$|A\cup B| = |A-B| + |B|, \qquad |A| = |A-B| + |A\cap B|.$$

由此即可得(1.35). □

(4) **集合的笛卡尔积** 集合 A 与 B 的**笛卡尔积** 是由所有元素对 (a,b) (其中 $a\in A, b\in B$) 构成的集合, 即

$$A\times B := \{(a,b) \mid a\in A, b\in B\}.$$

$A\times B$ 中的两个元素(a,b) 与(c,d)相等当且仅当$a=c, b=d$.

更进一步地, 集合族 A_i $(i\in I)$ 的笛卡尔积定义为

$$\prod_{i\in I} A_i := \{(a_i)_{i\in I} \mid a_i \in A_i\}.$$

如所有的A_i 均为A, 我们通常用A^I 表示其笛卡尔积. 特别地, 我们用A^n 表示n 个A 的笛卡尔积. 例如, $\mathbb{R}^n = \{(x_1,x_2,\ldots,x_n) \,|\, x_i\in\mathbb{R}, i=1,2,\ldots,n\}$. 再如, 设$A=\{1,2\}$, 则$A^2 = \{(1,1),(1,2),(2,1),(2,2)\}$. 容易证明, $|A\times B| = |A|\times|B|$.

§1.7.3 映射*

设 A,B 为两个非空集合. 如果对 A 中每个元素 a, 均有唯一元素 $b\in B$ 与之对应, 则称此对应为 A 到 B 的**映射**, 记之为

$$f:A\to B,\quad a\mapsto b=f(a).$$

有时候也记为

$$A\xrightarrow{f}B.$$

集合A 称为 f 的**定义域**, $Im(f):=f(A)=\{f(a)\mid a\in A\}\subseteq B$ 称为f 的**值域** 或**像集**. b 称为a 的**像**, a 称为b 的**原像**.

设f 与g 为集合A 到B的两个映射. 如果对于A 中任意元素a, 均有$f(a)=g(a)$, 则称映射f 与g **相等**, 记为$f=g$.

当集合B 是数(如有理数或者实数) 的集合时, 映射f 习惯上称为**函数**. 因此, 我们在中学阶段学习到的各种初等函数, 如多项式函数, 三角函数, 指数函数等都可以看成从$\mathbb{R}$ 到$\mathbb{R}$ 的映射.

如果映射f将不同的元素映射到不同的元素, 即当$a\neq b\in A$时, $f(a)\neq f(b)$, 则称映射 f 为**单射**. 容易验证, f为单射当且仅当f的每个像都有唯一的原像, 即对元素 $a,b\in A$, 当 $f(a)=f(b)$ 时, 必有 $a=b$. 如果对任意 $b\in B$, 存在 $a\in A$, 使得 $f(a)=b$, 则称 f 为**满射**. 如果 f 既是单射又是满射, 则称 f 为**一一映射** 或**双射**. 例如, **恒同映射** $I_A:A\mapsto A, a\mapsto a$ (即将每个元素映射到自身的映射) 是双射.

设 $f:A\mapsto B$ 和$g:B\mapsto C$ 为映射, 则映射

$$g\circ f:A\mapsto C,\quad a\mapsto g(f(a))$$

称为 f 与 g 的**复合映射**.

例如, 设f是集合A到集合B的映射, 则$f\circ I_A=f$, $I_B\circ f=f$. 映射复合的一个重要性质是结合律.

命题1.7.4. *(结合律)* 设 $f:A\mapsto B$, $g:B\mapsto C$ 与$h:C\mapsto D$ 为集合间的映射, 则

$$(h\circ g)\circ f=h\circ(g\circ f).$$

证明. 直接验证即可. □

设f是集合A到B的映射, $A_0\subset A$. 映射$f|_{A_0}:A_0\mapsto B, a\mapsto f(a)$ 称为映射f在集合A_0上的**限制**. 映射$i_{A_0}:A_0\mapsto A, a\mapsto a$ 称为**嵌入映射**. 显然, $f|_{A_0}=f\circ i_{A_0}$.

下面我们介绍逆映射. 设f是集合A到B的映射. 如果存在从集合B到集合A的映射g, 满足$g\circ f=I_A$, $f\circ g=I_B$, 则称g是f的**逆映射**, 并称f 是**可逆的**. f的逆映射记

为f^{-1}. 容易证明, 逆映射如果存在则是唯一的, 并且$(f^{-1})^{-1}=f$. 实际上, 设g,h均为f的逆, 则$g=g\circ I_B=g\circ(f\circ h)=(g\circ f)\circ h=I_A\circ h=h$.

例如, $f:\mathbb{R}\mapsto\mathbb{R}_+, x\mapsto e^x$ 是可逆的, f的逆为$f^{-1}:\mathbb{R}_+\mapsto\mathbb{R}, x\mapsto\log(x)$. 这里$\mathbb{R}_+$表示全体正实数的集合.

定理1.7.5. *映射$f:A\mapsto B$ 可逆当且仅当f是双射.*

证明. 设f是双射. 则对任意$b\in B$, 存在唯一$a\in A$ 使得$f(a)=b$. 于是存在映射$g:B\mapsto A$ 满足$g(b)=a$, 且$f(g(b))=b$, 或等价地, $f\circ g=I_B$. 另一方面, $g(f(a))=g(b)=a$, 亦即$g\circ f=I_A$. 从而g 是f 的逆.

反过来, 设f 可逆. 则存在映射$g:B\mapsto A$ 满足$f\circ g=I_B$, $g\circ f=I_A$. 对任何$a_1\neq a_2\in A$, $a_1=(g\circ f)(a_1)=g(f(a_1))\neq g(f(a_2))=(g\circ f)(a_2)=a_2$, 故$f(a_1)\neq f(a_2)$, 从而$f$是单射. 另一方面, 对任意$b\in B$, $f(g(b))=(f\circ g)(b)=b$, 故f是满射, 从而f是双射. □

对于有限集合, 我们有以下结论.

命题1.7.6. *设A为有限集, $f:A\mapsto A$为映射. 则f是单射当且仅当f是满射.*

证明. 设f为单射, 则$|Im(f)|\geq|A|$. 但$Im(f)\subset A$, 故$|Im(f)|=|A|$, 进而$Im(f)=A$, 即f是满射.

反之, 设f是满射, 即$Im(f)=A$, 于是$|Im(f)|=|A|$. 若f不是单射, 则$|Im(f)|<|A|$. 矛盾! 故f为单射.

□

定义1.7.1. 设 S 为集合. 我们称映射 $f:S\times S\mapsto S$, $(a,b)\mapsto p$ 为 S 上的一个**二元运算**.

注1. 在数学应用中, 记号 $p=f(a,b)$ 并不是一个很适宜的记号. 实际上, 我们经常使用 $+,\times,*,\cdot$ 等符号来表示二元运算, 即

$$\boxed{p=ab,\ a\times b,\ a+b,\ a*b,\ a\cdot b,\text{诸如此类.}}$$

例1.7.1. 加法, 减法和乘法是实数集$\mathbb{R}$上的二元运算, 除法是非零实数集$\mathbb{R}^\times:=\mathbb{R}\backslash\{0\}$ 上的二元运算.

例1.7.2. 记 Σ_A 为集合 A 到自身的所有映射的集合, 则映射的复合构成 Σ_A 上的二元运算. 记 S_A 为集合 A 到自身的所有双射构成的集合, 则映射的复合构成 S_A 上的二元运算.

定义1.7.2. 集合 S 上的二元运算如果满足条件: 对所有 $a,b,c\in S$,

$$(ab)c=a(bc), \tag{1.36}$$

则称该二元运算满足**结合律**. 如果对任意 $a,b\in S$,

$$ab=ba, \tag{1.37}$$

则称其满足**交换律**.

容易看出, 例1.7.1和1.7.2中的二元运算均满足结合律, 但映射的复合并不满足交换律. 因此, 结合律是更一般性的规律.

§1.7.4 置换与排列*

本小节, 我们简单介绍一个特殊的映射–置换. 置换在诸多代数问题, 如多项式方程求根以及行列式中都会用到.

设$A=\{1,2,\dots,n\}$. 则A到自身的一一映射称为一个**置换**, 也称为**排列**. 所有置换的全体构成的集合记为S_n. 设σ 是一个置换, 我们一般将置换表示成形式:

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

或者写成一个排列的形式$\sigma=(\sigma(1),\sigma(2),\dots,\sigma(n))$. 特别地, $\sigma=(1,2,\dots,n)$称为**恒同置换**, 记为e. 显然, $e\sigma=\sigma e=\sigma$.

例如, 置换$\sigma=\begin{pmatrix}1&2&3\\2&3&1\end{pmatrix}$ 表示$A=\{1,2,3\}$到自身的映射: $\sigma(1)=2$, $\sigma(2)=3$, $\sigma(3)=1$. 该置换也可以表示成一个排列$(2,3,1)$.

S_n中一共有$n!$个元素. 例如, S_3 由以下6个置换构成:

$$\begin{pmatrix}1&2&3\\1&2&3\end{pmatrix},\begin{pmatrix}1&2&3\\1&3&2\end{pmatrix},\begin{pmatrix}1&2&3\\2&1&3\end{pmatrix},\begin{pmatrix}1&2&3\\2&3&1\end{pmatrix},\begin{pmatrix}1&2&3\\3&1&2\end{pmatrix},\begin{pmatrix}1&2&3\\3&2&1\end{pmatrix}.$$

或者写成$S_3=\{(1,2,3),(1,3,2),(2,1,3),(2,3,1),(3,1,2),(3,2,1)\}$.

任意两个置换的复合仍然为一个置换, 我们称之为两个置换的**乘积**. 置换σ,τ的乘积记为$\sigma\tau$. 例如, $\sigma=\begin{pmatrix}1&2&3\\2&1&3\end{pmatrix}$, $\tau=\begin{pmatrix}1&2&3\\3&2&1\end{pmatrix}$. 则$\sigma\tau=\begin{pmatrix}1&2&3\\3&1&2\end{pmatrix}$, $\tau\sigma=\begin{pmatrix}1&2&3\\2&3&1\end{pmatrix}$.

称只改变两个位置的置换为**对换**. 若改变的元素为i,j, 则该置换记为(ij). 例如, $\begin{pmatrix}1&2&3&4&5\\1&2&5&4&3\end{pmatrix}$ 即是对换(35).

命题1.7.7. 任意置换都是有限个对换的乘积.

证明. 我们对n做归纳. 当$n=2$时, $S_2=\{(12),(21)\}$. (21)本身是对换, 而(12)是恒同置换, 它可以表示为$(12)(12)$. 现假设结论对$n-1$成立. 对n, 我们先证: 任意置换σ可以写成最多一个对换与一个置换τ的乘积, 这里$\tau(n)=n$.

设$\sigma(i)=n$, 如果$i=n$, 取$\tau=\sigma$即可. 下设$i<n$, 令

$$\tau=(\sigma(1),\ldots,\sigma(i-1),\sigma(n),\sigma(i+1),\ldots,\sigma(n-1),\sigma(i)),$$

则$\sigma=\tau\cdot(i,n)$, 且$\tau(n)=\sigma(i)=n$.

对$\tau'=(\tau(1),\tau(2),\ldots,\tau(n-1))$ 用归纳假设知, τ' 可以写成有限个对换的乘积, 于是τ 也可以写成有限个对换的乘积, 从而σ也是有限个对换的乘积. □

作为一个例子, 我们给出$\sigma=(31542)$的对换分解: $\sigma=(35)(13)(12)$. 当然, 上述分解并不唯一. 例如, σ 还可以分解为$\sigma=(12)(34)(45)(34)(23)$. 不过无论如何分解, σ分解出的对换个数具有相同的奇偶性.

定理1.7.8. 一个置换的所有对换分解中, 对换的个数具有相同的奇偶性.

证明. 对置换σ, 引进σ 对函数$f(x_1,x_2,\ldots,x_n)$的作用

$$\sigma(f)(x_1,x_2,\ldots,x_n):=f(x_{\sigma(1)},x_{\sigma(2)},\ldots,x_{\sigma(n)}).$$

令$\Delta(x_1,x_2,\ldots,x_n):=\Pi_{1\le i<j\le n}(x_j-x_i)$. 容易验证, 如果$\sigma$是$i$个对换的乘积, 则

$$\sigma(\Delta)(x_1,x_2,\ldots,x_n)=(-1)^i\Delta(x_1,x_2,\ldots,x_n).$$

如果σ又可以分解为j个对换的乘积, 则

$$\sigma(\Delta)(x_1,x_2,\ldots,x_n)=(-1)^j\Delta(x_1,x_2,\ldots,x_n).$$

于是i与j必须具有相同的奇偶性. □

由上面的定理, 我们可以定义一个置换或排列的奇偶性. 如果一个置换的对换分解中对换的个数为偶数(奇数), 则称该置换为为**偶置换**(**奇置换**). 对任意置换σ, 定义它的**符号**为$\varepsilon_\sigma=(-1)^k$, 这里$k$ 是σ的对换分解中对换的个数. 因此, 偶置换的符号为1, 奇置换的符号为-1. 容易验证, 对任意两个置换σ与τ, $\varepsilon_{\sigma\tau}=\varepsilon_\sigma\varepsilon_\tau$.

§1.7.5 等价关系, 等价类与分拆*

定义1.7.3. 集合 A 中的元素间的关系 $\sim$ 称为**等价关系**, 是指下述三条性质成立:

(1) (**自反性**) 对所有 $a \in A$, $a \sim a$.

(2) (**对称性**) 如果 $a \sim b$, 则 $b \sim a$.

(3) (**传递性**) 如果 $a \sim b$ 且 $b \sim c$, 则 $a \sim c$.

定义1.7.4. 集合 A 作为它的一些子集合的不交并, 称为 A 的一个**分拆**.

设 $\sim$ 是 A 上的一个等价关系. 对 $a \in A$, 记 $[a] = \{b \in A \mid b \sim a\}$, 即 $[a]$ 为 A 中所有与 a 等价的元素构成的子集合. 子集合$[a]$ 称为a 所在的**等价类**. 注意到

$$[a] \cap [b] = \begin{cases} [a] = [b], & \text{如果} a \sim b, \\ \varnothing, & \text{如果} a \nsim b. \end{cases}$$

记$A/\sim$ 为A 中所有等价类构成的集合, 即

$$A/\sim := \{[a] \mid a \in A\} \quad (\text{去掉重复项}).$$

则A 可以写为不交并

$$A = \bigsqcup_{[a] \in A/\sim} [a]. \tag{1.38}$$

由此我们得到 A 的一个分拆. 反过来, 如果 $A = \bigsqcup\limits_{i \in I} A_i$ 为A 的分拆, 则很容易在 A 上定义等价关系:

$$a \sim b \quad \text{当且仅当} \quad a, b \text{ 属于同一个 } A_i.$$

故我们有如下定理

定理1.7.9. 集合 A 的分拆与定义在A上的等价关系一一对应.

例1.7.3. 整数集合 $\mathbb{Z}$ 可以分拆为偶数集合和奇数集合的不交并. 另一方面, 在 $\mathbb{Z}$ 上可以定义等价关系: $a \sim b$ 如果 $a - b$ 是偶数. 则偶数集合是此等价关系中 0 所在的等价类, 奇数集合为 1 所在的等价类.

设 $f: A \mapsto B$ 为集合间的映射. 对于元素 $b \in B$, 令b 的原像集合 $f^{-1}(b) = \{a \in A \mid f(a) = b\}$, 则 $f^{-1}(b)$ 为 A 的子集. 对于B 中不同的元素b 和b', 有$f^{-1}(b) \cap f^{-1}(b') = \varnothing$. 并且, $f^{-1}(b) = \varnothing$ 当且仅当 $b \notin f(A)$. 故我们得到分拆

$$A = \bigsqcup_{b \in f(A)} f^{-1}(b). \tag{1.39}$$

我们称集合A 的这个分拆为**映射 f 决定的分拆**. 它决定的等价关系即

$$a \sim a' \iff f(a) = f(a').$$

例1.7.4. 如果$\sim$ 是集合A 上的等价关系, 对于自然映射

$$p : A \to A/\sim, \quad a \mapsto [a],$$

可以看出, p 所决定的分拆即等价关系$\sim$ 所决定的分拆.

例1.7.5. 定义映射 $f : \mathbb{Z} \mapsto \{0,1\}$, 其中 $f(2n) = 0$, $f(2n+1) = 1$. 则映射 f 决定的等价关系和分拆即与例1.7.3 给出的等价关系是同一等价关系.

例1.7.6. 设 $f : \mathbb{R}^2 = \mathbb{R} \times \mathbb{R} \mapsto \mathbb{R}$ 为实数减法映射 $(x,y) \mapsto x-y$, 则 $f^{-1}(a)$ 为直线 $y = x-a$. 实平面 $\mathbb{R}^2$ 由映射 f 决定的分拆即是平行直线束 $y = x-a$ $(a \in \mathbb{R})$ 的不交并.

§1.8 复数*

这一节中, 我们简单回顾一下有关复数的一些知识, 特别是复数的几何表示与运算.

§1.8.1 复数的四则运算*

复数就是形如 $z = x+iy$ 的数, 其中 i 为虚数单位 $\sqrt{-1}$, 而 x, y 为实数, 分别称为复数 z 的**实部**和**虚部**, 记作 $\mathrm{Re}\,z$ 和 $\mathrm{Im}\,z$.

设 $z_1 = x_1+iy_1, z_2 = x_2+iy_2$ 为两个复数, 复数的加法与减法定义为:

$$\begin{aligned} z_1+z_2 &= (x_1+x_2)+i(y_1+y_2), \\ z_1-z_2 &= (x_1-x_2)+i(y_1-y_2). \end{aligned} \tag{1.40}$$

复数的乘法与实数的乘法一样, 只要注意 $i^2 = -1$ 即可. 就是说

$$z_1z_2 = (x_1+iy_1)(x_2+iy_2) = (x_1x_2-y_1y_2)+i(x_1y_2+x_2y_1). \tag{1.41}$$

除法定义为

$$\frac{x_2+iy_2}{x_1+iy_1} = \frac{(x_2+iy_2)(x_1-iy_1)}{(x_1+iy_1)(x_1-iy_1)} = \frac{(x_1x_2+y_1y_2)+i(x_1y_2-x_2y_1)}{x_1^2+y_1^2}. \tag{1.42}$$

§1.8.2 复数的几何表示*

在平面中取一个直角坐标系 Oxy, 我们用坐标为 (x,y) 的点 P 表示复数 $z=x+iy$. 这样, 复数就与平面中的点一一对应. x 轴与实数对应, 也称为**实轴**; y 轴与纯虚数对应, 也称为**虚轴**. 与复数建立了这种对应关系的平面称为**复平面**.

我们也常用向量 $\overrightarrow{OP}$ 表示复数 $z=x+iy$. 这样复数 z, 复平面上的点 P 以及向量 $\overrightarrow{OP}$ 之间建立了一一对应关系. 如果 z_1,z_2 对应的向量分别为 $\overrightarrow{OP_1},\overrightarrow{OP_2}$, 则 z_1+z_2 对应的向量恰为 $\overrightarrow{OP_1}+\overrightarrow{OP_2}$. 因此复数的加法运算和向量的加法运算是一致的. 由于复数 z_1,z_2,z_1+z_2 对应的向量恰好构成了一个三角形, 由三角形两边之和大于第三边的性质可得到复数的三角不等式

$$|z_1+z_2|\le|z_1|+|z_2|. \tag{1.43}$$

复数 z 对应的向量 $\overrightarrow{OP}$ 的长度称为复数 z 的**模**, 记作 $|z|=r=\sqrt{x^2+y^2}$. 将x轴正向逆时针旋转至向量 $\overrightarrow{OP}$所得的角 θ 称为复数 z 的**幅角**, 记作 $arg\ z$(图 1.16). 复数 z 的幅角并不唯一, 它们彼此相差2π的整数倍. 在实际中, 我们一般规定 $0\le arg\ z<2\pi$, 称之为幅角的主值. 例如, $arg(-i)=\frac{3}{2}\pi$. 当然也可以规定幅角主值在其它的范围, 如$-\pi\le arg\ z<\pi$. 规定复数 $z=0$ 的幅角是不定的.

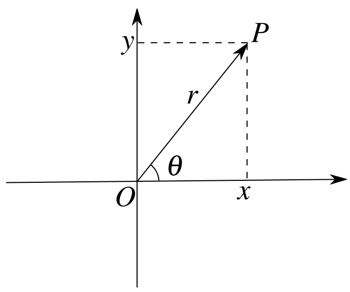


图 1.16 复数的几何定义

复数 $x-iy$ 称为复数 $z=x+iy$ 的**共轭复数**, 记为 $\bar{z}$. 从几何上看, z 和 $\bar{z}$ 关于实轴对称, 因此有 $|\bar{z}|=|z|$, $\arg\bar{z}=2\pi-\arg z$. 关于复数的共轭有如下简单的性质, 证明留作习题.

$$|z|^2=z\bar{z},\quad \overline{z_1+z_2}=\overline{z_1}+\overline{z_2},\quad \overline{z_1z_2}=\overline{z_1}\,\overline{z_2}. \tag{1.44}$$

知道了复数 z的模长 r与幅角 θ, 则 z 可以表示为

$$z=x+iy=r\cos\theta+ir\sin\theta=r(\cos\theta+i\sin\theta) \tag{1.45}$$

称上式为复数的**三角表示**. 利用复数的三角表示, 很容易实现复数的乘法运算, 并且可以给予简单的几何解释.

假设 $z_1 = r_1(\cos\theta_1 + i\sin\theta_1)$, $z_2 = r_2(\cos\theta_2 + i\sin\theta_2)$, 则由乘法定义知

$$\begin{aligned} z_1 z_2 &= r_1 r_2[(\cos\theta_1\cos\theta_2 - \sin\theta_1\sin\theta_2) + i(\sin\theta_1\cos\theta_2 + \sin\theta_2\cos\theta_1)] \\ &= r_1 r_2\left(\cos(\theta_1+\theta_2) + i\sin(\theta_1+\theta_2)\right). \end{aligned} \tag{1.46}$$

由此可得

$$|z_1 z_2| = |z_1||z_2|, \tag{1.47}$$

$$\arg(z_1 z_2) = \arg z_1 + \arg z_2. \tag{1.48}$$

即复数乘积的模等于复数模的乘积, 复数乘积的幅角等于各幅角的和(如超出主值范围, 则还要减去2π的整数倍). 由此可以得到乘法的一个几何解释. 设复数 $w = r(\cos\theta + i\sin\theta)$, 则 $w\cdot z$ 表示将复数 z 对应的向量模长伸缩 r 倍, 再逆时针旋转 θ 角. 因此, 利用复数乘法可以方便地表示旋转变换.

复数的三角形式还有一种更为常用的表示方式. 根据 Euler 公式

$$e^{i\theta} = \cos\theta + i\sin\theta$$

(只要给出复指数函数的适当定义, 上述公式是可以证明的, 这里我们可以将它看成是一个记号), 模长为 r 幅角为 θ 的复数 z 可记为 $z = re^{i\theta}$. 由 (1.46) 知

$$r_1 e^{i\theta_1}\cdot r_2 e^{i\theta_2} = r_1 r_2 e^{i(\theta_1+\theta_2)}.$$

因此 $e^{i\theta}$ 可以象普通的指数函数一样进行运算.例如:

$$(re^{i\theta})^n = r^n e^{in\theta}, \quad n = 0, \pm1, \pm2, \ldots \qquad \text{(de Moivre 公式)} \tag{1.49}$$

例1.8.1. 求复数 $z = 1 + \cos\theta + i\sin\theta\,(-\pi \le \theta < \pi)$ 的三角形式.

解.

$$|z| = \sqrt{(1+\cos\theta)^2 + \sin^2\theta} = 2\left|\cos\frac{\theta}{2}\right| = 2\cos\frac{\theta}{2},$$

$$\arg z = \arccos\frac{1+\cos\theta}{2\cos\frac{\theta}{2}} = \arccos\cos\frac{\theta}{2} = \frac{\theta}{2}.$$

从而 $z = 2\cos\dfrac{\theta}{2}\left(\cos\dfrac{\theta}{2} + i\sin\dfrac{\theta}{2}\right)$. □

例1.8.2. 在复数域范围内求方程 $z^n = a$ 的根, 其中 a 为复数, n 为正整数.

解. 设 $a = re^{i\theta}$, 其中 $r \geq 0, 0 \leq \theta < 2\pi$. 设 $z = se^{i\phi}$, 则

$$s^n e^{in\phi} = re^{i\theta}.$$

因此

$$s = r^{\frac{1}{n}}, \quad \phi = \frac{\theta + 2k\pi}{n} \quad (k = 0, 1, \cdots, n-1)$$

所求方程有 n 个不同的根, 它们是

$$r^{\frac{1}{n}} e^{i\frac{\theta+2k\pi}{n}} \quad (k = 0, 1, \cdots, n-1).$$

□

本节最后, 我们给出两个例子说明复数在几何中的应用.

例1.8.3. 证明:平行四边形的两条对角线的平方和等于四边的平方和.

证明. 假设平行四边形某相邻两边向量对应的复数分别为z_1, z_2, 则两对角线向量分别对应复数$z_1 + z_2$与$z_1 - z_2$. 于是我们只需证明

$$|z_1 + z_2|^2 + |z_1 - z_2|^2 = 2(|z_1|^2 + |z_2|^2).$$

实际上,

$$\begin{aligned} |z_1 + z_2|^2 &= (z_1 + z_2)(\overline{z_1 + z_2}) = (z_1 + z_2)(\bar{z}_1 + \bar{z}_2) \\ &= z_1\bar{z}_1 + z_1\bar{z}_2 + z_2\bar{z}_1 + z_2\bar{z}_2 \\ &= |z_1|^2 + z_1\bar{z}_2 + z_2\bar{z}_1 + |z_2|^2. \end{aligned}$$

同理可得

$$|z_1 - z_2|^2 = |z_1|^2 - z_1\bar{z}_2 - z_2\bar{z}_1 + |z_2|^2.$$

两式相加, 即得要证的等式. □

例1.8.4. 在平面直角坐标系Oxy中, 求将点$P = (x, y)$绕原点逆时针旋转θ角后所得点P'的坐标.

解. 设P'的坐标为(x', y'). 点P与P'对应的复数分别为$z = x + yi$及$z' = x' + y'i$. 根据复数乘法的意义, $z' = e^{i\theta}z$, 即

$$x' + y'i = (x + iy)(\cos\theta + i\sin\theta) = (x\cos\theta - y\sin\theta) + (x\sin\theta + y\cos\theta)i.$$

因此,

$$\begin{cases} x' &= x\cos\theta - y\sin\theta, \\ y' &= x\sin\theta + y\cos\theta. \end{cases} \tag{1.50}$$

□

上述公式即平面直角坐标系中, 绕原点逆时针旋转θ角的点变换公式.

§1.9 数域*

复数集的任何一个子集称为一个数集. 自然数集 $\mathbb{N}$, 整数集 $\mathbb{Z}$, 有理数集 $\mathbb{Q}$, 实数集 $\mathbb{R}$ 都是大家熟知的数集.

设 F 为一个数集, 在 F 中任取两数作某种运算, 如果其结果仍在 F 中, 则称数集 F 对这种运算是**封闭的**. 例如 $\mathbb{C},\mathbb{R},\mathbb{Q}$ 对数的加减乘除运算是封闭的, $\mathbb{Z}$ 对数的加减乘运算是封闭的, 但对除法运算不封闭, 而 $\mathbb{N}$只对加法乘法运算封闭.

定义1.9.1. 设数集 F 至少包含两个不同的元素, 称 F 为**数域**, 如果 F 对数的加减乘除运算是封闭的, 即当 $a,b\in F$ 时, $a\pm b,\ ab,\ \dfrac{a}{b}\,(b\neq 0)\in F$.

由定义知, $\mathbb{C},\mathbb{R},\mathbb{Q}$ 为数域, 分别称为复数域, 实数域和有理数域. $\mathbb{Z},\mathbb{N}$ 则不是数域.

由定义不难看出, 一个数域必含有 $0,1$ 两个元素.实际上可以证明, 在通常的四则运算下任何数域都包含有理数域(读者不妨证明).下面的例子表明, 除了上述三种数域, 还有其它的数域.

例1.9.1. 证明数集 $\mathbb{Q}(\sqrt{2})=\{a+b\sqrt{2}:a,b\in\mathbb{Q}\}$ 为一个数域.

证明. 首先 $\mathbb{Q}(\sqrt{2})$ 对加减运算封闭是显然的. 任取 $\mathbb{Q}(\sqrt{2})$ 中的两个元素 $a+b\sqrt{2},\ c+d\sqrt{2}$,

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

而当 $c+d\sqrt{2}\neq 0$时,

$$\frac{a+b\sqrt{2}}{c+d\sqrt{2}}=\frac{(a+b\sqrt{2})(c-d\sqrt{2})}{(c+d\sqrt{2})(c-d\sqrt{2})}=\frac{ac-2bd}{c^2-2d^2}+\frac{bc-ad}{c^2-2d^2}\sqrt{2}.$$

由于 $a,b,c,d\in\mathbb{Q}$, $\mathbb{Q}$ 为数域, 所以 $ac+2bd\in\mathbb{Q}$, $ad+bc\in\mathbb{Q}$, 即 $\mathbb{Q}(\sqrt{2})$ 对乘法运算是封闭的. 又由于 $c^2-2d^2\neq 0$, 所以

$$\frac{ac-2bd}{c^2-2d^2}\in\mathbb{Q},\quad \frac{bc-ad}{c^2-2d^2}\in\mathbb{Q},$$

即 $\mathbb{Q}(\sqrt{2})$ 对除法运算是封闭的.故 $\mathbb{Q}(\sqrt{2})$ 为一个数域. □

在本书的后续章节中, 我们考虑的数域 F 均指包含有理数域, 运算为通常四则运算的数域, 如实数域、复数域等.

§1.10 求和符号*

为了后面的需要, 我们在这里简单介绍一下求和符号的运用.

设 $a_1,a_2\cdots,a_n$ 为 n 个数, 通常我们用 $\sum\limits_{i=1}^{n}a_i$ 表示和式 $a_1+a_2+\cdots+a_n$. 其中 $\sum$ 为求和符号, i 为求和指标, $\sum$ 的上下标表示求和指标 i 的取值范围. 注意求和指标可以用其它的字母代替, 例如 $\sum\limits_{j=1}^{n}a_j$, $\sum\limits_{k=1}^{n}a_k$ 表示的都是同一个和式. 容易看出, 求和符号满足以下性质:

$$\sum_{i=1}^{n}(a_i+b_i)=\sum_{i=1}^{n}a_i+\sum_{i=1}^{n}b_i,$$
$$\sum_{i=1}^{n}\lambda a_i=\lambda\sum_{i=1}^{n}a_i. \qquad (\text{其中 } \lambda \text{ 为常数})$$

在线性代数课程中, 我们经常会遇到双重, 甚至多重求和. 多重求和的原则是从内向外对每个求和符号逐次求和. 例如 $\sum\limits_{j=1}^{m}\sum\limits_{i=1}^{n}a_{ij}$ 中, 就应该先对求和指标 i 求和, 再对求和指标 j 求和. 在对指标 i 求和时, 指标 j 是固定不变的, 因此有

$$\sum_{i=1}^{n}a_{ij}=a_{1j}+a_{2j}+\cdots+a_{nj},$$

所以

$$\begin{aligned}\sum_{j=1}^{m}\sum_{i=1}^{n}a_{ij}&=\sum_{j=1}^{m}(a_{1j}+a_{2j}+\cdots+a_{nj})\\&=\sum_{j=1}^{m}a_{1j}+\sum_{j=1}^{m}a_{2j}+\cdots+\sum_{j=1}^{m}a_{nj}.\end{aligned}$$

从上式可以看出 $\sum\limits_{j=1}^{m}\sum\limits_{i=1}^{n}a_{ij}$ 等于所有$a_{ij}(1\le i\le n,1\le j\le m)$ 的和. 同样的道理, $\sum\limits_{i=1}^{n}\sum\limits_{j=1}^{m}a_{ij}$ 表示的也是所有 a_{ij} 的和. 事实上, 如果我们将这些数 a_{ij} 排成如下的一个方块:

$$\begin{matrix}a_{11}&a_{12}&\cdots&a_{1m}\\a_{21}&a_{22}&\cdots&a_{2m}\\\cdots&\cdots&\cdots&\cdots\\a_{n1}&a_{n2}&\cdots&a_{nm}\end{matrix}$$

那么 $\sum\limits_{j=1}^{m}\sum\limits_{i=1}^{n}a_{ij}$ 就是先固定一个 j, 对上述方块的第 j 列求和, 再将得到的和相加; 而 $\sum\limits_{i=1}^{n}\sum\limits_{j=1}^{m}a_{ij}$ 则是先固定一个 i, 对第 i 行求和, 再将得到的和相加. 因此我们有

$$\sum_{j=1}^{m}\sum_{i=1}^{n}a_{ij}=\sum_{i=1}^{n}\sum_{j=1}^{m}a_{ij}.$$

上式表明在多重求和时, 可以交换求和符号的次序. 这一点非常重要, 今后会经常用到.

尽管是多重求和, 有时也可以只用一个求和号表示. 例如 $\sum\limits_{1\le i,j\le n} a_{ij}$ 就是 $\sum\limits_{j=1}^{n}\sum\limits_{i=1}^{n} a_{ij}$. 有些时候, 我们并不是对所有的项求和, 而只是对满足一定条件的项求和. 这时我们通常将这些条件写在求和符号的下方. 例如

$$\sum_{1\le i\le j\le n} a_{ij} = a_{11} + (a_{12}+a_{22}) + (a_{13}+a_{23}+a_{33}) \\ + \cdots + (a_{1n}+a_{2n}+\cdots+a_{nn})$$

表示的是满足 $i \le j$ 的所有项的和. 如果用二重和式表示, 可得

$$\sum_{1\le i\le j\le n} a_{ij} = \sum_{j=1}^{n}\sum_{i=1}^{j} a_{ij} = \sum_{i=1}^{n}\sum_{j=i}^{n} a_{ij}.$$

注意上式中的后一个等式, 形式上似乎与前面所说的求和可交换次序不符, 这是由于和式 $\sum\limits_{j=1}^{n}\sum\limits_{i=1}^{j} a_{ij}$ 中, 内层的求和 $\sum\limits_{i=1}^{j} a_{ij}$ 中, 指标 i 的取值范围涉及到了外层指标 j, 这时就不能轻易的直接交换求和次序了.

例1.10.1. 证明: $\sum\limits_{j=1}^{m}\sum\limits_{i=1}^{n} a_i b_j = (\sum\limits_{i=1}^{n} a_i)(\sum\limits_{j=1}^{m} b_j)$.

证明. 在左端的二重求和中, 先固定一个 j, 对指标 i 求和. 这时 b_j 与求和指标 i 无关, 可看成常数, 可以提到求和号的外面, 所以

$$\sum_{j=1}^{m}\sum_{i=1}^{n} a_i b_j = \sum_{j=1}^{m}\left(b_j \sum_{i=1}^{n} a_i\right).$$

再对 j 求和时, $\sum\limits_{i=1}^{n} a_i$ 与求和指标 j 无关, 可看成常数, 因此有

$$\sum_{j=1}^{m}\left(b_j \sum_{i=1}^{n} a_i\right) = \left(\sum_{i=1}^{n} a_i\right)\left(\sum_{j=1}^{m} b_j\right).$$

□

求和符号还有更一般的形式. 设 Λ 为一个有限集, Λ 中每个元素 λ 对应了一个数 a_λ, 我们将所有 $a_\lambda\ (\lambda \in \Lambda)$ 的和记为 $\sum\limits_{\lambda\in\Lambda} a_\lambda$. 上面介绍的 $\sum\limits_{i=1}^{n} a_i$ 可以看成一个特例, 即取 $\Lambda = \{1,2,\cdots,n\}$. 今后在定义 n 阶行列式时, 我们要用到这样的求和符号. 在那里, Λ 为 n 个元素的所有排列构成的集合.

最后, 介绍一下乘积符号 $\prod$. 设 $a_1,a_2,\cdots,a_n$ 为一列数, 我们用 $\prod\limits_{i=1}^{n} a_i$ 表示乘积 $a_1,a_2,\cdots a_n$. 乘积符号的上下标的含义与求和符号完全一样, 不再赘述.

习题一

注: 以下各题中涉及的坐标系均为直角坐标系.

1. 设 $\overrightarrow{AM}=\overrightarrow{MB}$, 证明: 对任意一点 O, $\overrightarrow{OM}=\dfrac{1}{2}(\overrightarrow{OA}+\overrightarrow{OB})$.

2. 设 O 为一定点, A,B,C 为不共线的三点. 证明点 M 位于平面 ABC 上的充分必要条件是存在实数 k_1,k_2,k_3, 使得

$$\overrightarrow{OM}=k_1\overrightarrow{OA}+k_2\overrightarrow{OB}+k_3\overrightarrow{OC},\quad \text{且}\quad k_1+k_2+k_3=1.$$

3. 证明向量 $\mathbf{a}-\mathbf{b}+\mathbf{c}$, $-2\mathbf{a}+3\mathbf{b}-2\mathbf{c}$, $2\mathbf{a}-\mathbf{b}+2\mathbf{c}$ 线性相关.

4. 证明: 三维空间中四个或四个以上的向量一定线性相关.

5. 设 $\mathbf{e}_1,\mathbf{e}_2,\mathbf{e}_3$ 为一组基.

   (1) 证明:$\mathbf{a}=\mathbf{e}_1+2\mathbf{e}_2-\mathbf{e}_3$, $\mathbf{b}=2\mathbf{e}_1+\mathbf{e}_2+\mathbf{e}_3$, $\mathbf{c}=3\mathbf{e}_1+2\mathbf{e}_3$ 为一组基;

   (2) 设 $\tilde{\mathbf{c}}=3\mathbf{e}_1+x\mathbf{e}_2+2\mathbf{e}_3$, 当 x 取何值时, $\mathbf{a},\mathbf{b},\tilde{\mathbf{c}}$ 共面.

6. 已知三点 $A(2,1,-1),B(3,5,1),C(1,-3,-3)$, 问 A,B,C 是否共线?

7. 已知线段 AB 被点 $C(1,2,3)$ 和 $D(2,-1,5)$ 三等分, 求端点 A,B 的坐标.

8. 设 $\mathbf{a}=(1,-2,4),\mathbf{b}=(2,2,1)$, 试计算 $\mathbf{a}\cdot\mathbf{b}$, $(\mathbf{a}+\mathbf{b})\cdot(\mathbf{a}-\mathbf{b})$, $(\mathbf{a}-\mathbf{b})^2$.

9. 设三个向量 $\mathbf{a},\mathbf{b},\mathbf{c}$ 两两间的夹角为 $45°$, 且 $|\mathbf{a}|=1,|\mathbf{b}|=2,|\mathbf{c}|=3$. 求向量 $\mathbf{a}+2\mathbf{b}-\mathbf{c}$ 的模.

10. 设 $\mathbf{a},\mathbf{b},\mathbf{c}$ 是满足 $\mathbf{a}+\mathbf{b}+\mathbf{c}=\mathbf{0}$ 的单位向量, 试求 $\mathbf{a}\cdot\mathbf{b}+\mathbf{b}\cdot\mathbf{c}+\mathbf{c}\cdot\mathbf{a}$ 的值.

11. 设向量 $\mathbf{a},\mathbf{b}$ 的夹角为 $60°$, 且 $|\mathbf{a}|=1,|\mathbf{b}|=2$, 试求 $(\mathbf{a}\times\mathbf{b})^2$, $|(\mathbf{a}+\mathbf{b})\times(\mathbf{a}-\mathbf{b})|$.

12. 设向量 $\mathbf{a}=(1,-1,2),\mathbf{b}=(2,3,-4)$, 求 $\mathbf{a}\times\mathbf{b}$, $(\mathbf{a}+\mathbf{b})\times(\mathbf{a}-\mathbf{b})$.

13. 一个四面体的顶点为 $A(1,2,3),B(-1,0,2),C(2,4,5),D(0,-3,4)$, 求它的体积.

14. 判断下列结论是否成立, 不成立时请举例说明.

    (1) 若 $\mathbf{a}\cdot\mathbf{b}=0$, 则 $\mathbf{a}=\mathbf{0}$ 或 $\mathbf{b}=\mathbf{0}$;

    (2) 若$\mathbf{a}\times\mathbf{b}=\mathbf{a}\times\mathbf{c}$, 则必有 $\mathbf{b}=\mathbf{c}$;

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

(4) $(\mathbf{a}\cdot\mathbf{b})^2=\mathbf{a}^2\cdot\mathbf{b}^2$;

(5) $(\mathbf{a}+\mathbf{b})\times(\mathbf{a}+\mathbf{b})=\mathbf{a}\times\mathbf{a}+2\mathbf{a}\times\mathbf{b}+\mathbf{b}\times\mathbf{b}$;

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

15. 证明下列等式:

(1) $(\mathbf{a}\times\mathbf{b})^2=\mathbf{a}^2\mathbf{b}^2-(\mathbf{a}\cdot\mathbf{b})^2$.

(2) $(\mathbf{a}\times\mathbf{b})\times\mathbf{c}+(\mathbf{b}\times\mathbf{c})\times\mathbf{a}+(\mathbf{c}\times\mathbf{a})\times\mathbf{b}=\mathbf{0}$.

16. 证明共轭复数的下列性质:

$$|z|^2=z\bar{z},\quad \overline{z_1+z_2}=\overline{z_1}+\overline{z_2},\quad \overline{z_1z_2}=\overline{z_1}\,\overline{z_2}.$$

17. 设 $z=\cos\theta+i\sin\theta\neq 1$, 求 $\dfrac{1+z}{1-z}$.

18. 求下列和式:

(1) $1+\cos\theta+\cos 2\theta+\cdots+\cos n\theta$;

(2) $\sin\theta+\sin 2\theta+\cdots+\sin n\theta$.

19. 证明: $|1+z_1\bar{z}_2|^2+|z_1-z_2|^2=(1+|z_2|^2)(1+|z_1|^2)$.

第二章　线性方程组

我们在中学学过了多元一次方程组的求解方法. 在那里变量(或变元)的个数通常为两个或三个, 并且方程的个数与变量的个数相同. 在本章中, 我们将一般地讨论由任意个变量及任意个方程构成的一次方程组的求解方法以及解的属性. 我们称关于变量都为一次的方程组成的方程组为**线性方程组**. 求解线性方程组的问题在科学与工程计算中有广泛的应用, 如天气预报、大坝与桥梁的设计、导弹的发射等.

一般地, 具有n个变量$x_1, \cdots, x_n$, m个方程组成的线性方程组具有形式:

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2 \\ \cdots\cdots\cdots\cdots\cdots\cdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = b_m \end{cases} \tag{2.1}$$

其中, a_{ij}是第i个方程中第j个变量x_j的**系数**, b_i是第i个方程的**常数项**. 本章中, 我们假设系数a_{ij}, 常数项b_i及方程组的解x_j均属于某个数域F (如有理数域$\mathbb{Q}$, 实数域$\mathbb{R}$或复数域$\mathbb{C}$). 如果常数项都为零, 则称相应的线性方程组为**齐次线性方程组**. 否则, 称为**非齐次线性方程组**. 若将$x_1 = c_1, \cdots, x_n = c_n$代入上述方程等式都成立, 则称$(c_1 \cdots c_n)$为该方程组的**一组解**. 线性方程组(2.1)的解的全体称为该方程组的**解集**. 如果解集非空, 则称线性方程组(2.1)是**相容**的; 否则, 称线性方程组**不相容**.

关于线性方程组的解, 有以下**几个基本问题**:

(1) 线性方程组是否存在解? 如果有解, 有多少个解?

(2) 如何求线性方程组的解?

(3) 解的公式表示.

(4) 解集的几何结构.

(5) 线性方程组的解是否符合实际的需要(可行解问题)?

本章及后面的第三章、第四章及第五章将围绕这些问题展开讨论, 并由此引入线性代数的基本工具–矩阵与行列式, 以及线性代数的基本研究对象–线性空间.

§2.1 Gauss 消元法

关于线性方程组的求解, 有十分悠久的历史. 我国古代名著《九章算术》就有求解线性方程组的一般方法, 现在一般称这种方法为**Gauss消元法**. Gauss消元法的基本思想是:将线性方程组化成所谓的三角形(或阶梯形)线性方程组, 然后再通过回代技术求解出所有变量. 下面我们通过几个实例来说明这种方法.

例2.1.1. 求解下列线性方程组

$$\begin{cases} 3x_1+2x_2-x_3=6 & (1) \\ x_1+3x_2+2x_3=9 & (2) \\ 2x_1-x_2+3x_3=3 & (3) \end{cases}$$

解. 我们在中学学过了求解这类方程的方法. 其基本思想是, 从三个方程中的某一个方程(如第一个方程)出发, 将另两个方程的某个变量(如x_1)消去. 则变换后的这两个方程都是二元一次方程. 即将一个三元一次方程组化为一个二元一次方程组求解. 类似地, 二元一次方程组又可以化为一元一次方程. 再用回代的方法求出所有变元的解. 对本例, 为了避免分数运算, 我们先将三个方程互换

$$\begin{cases} x_1+3x_2+2x_3=9 & (4) \\ 2x_1-x_2+3x_3=3 & (5) \\ 3x_1+2x_2-x_3=6 & (6) \end{cases}$$

将第(4)个方程分别乘-2与-3, 再分别加到第(5)与第(6)个方程即消去了这两个方程中的变量x_1:

$$\begin{cases} x_1 \quad +3x_2+2x_3=9 & (7) \\ \quad -7x_2-x_3=-15 & (8) \\ \quad -7x_2-7x_3=-21 & (9) \end{cases}$$

再将第(8)个方程乘-1加到第(9)个方程得

$$\begin{cases} x_1 \quad +3x_2+2x_3=9 & (10) \\ \quad -7x_2-x_3=-15 & (11) \\ \quad\quad 6x_3=6 & (12) \end{cases}$$

于是, 我们将原来的方程组化为一个三角型的方程组. 从方程(12)得$x_3=1$, 将$x_3=1$代入方程(11)得$x_2=2$. 再将$x_2=2, x_3=1$代入方程(10)得$x_1=1$. 这个过程称为**回代**. 由此, 原方程组的解为

$$x_1=1, \quad x_2=2, \quad x_3=1.$$

□

从上述求解线性方程组的过程可以看出, 我们主要对方程做了如下三种基本变换:

1. 交换两个方程;
2. 某个方程乘一个非零常数;
3. 某方程乘一非零常数加到另一个方程.

我们称这三种变换为方程的**初等变换**, 并分别用记号$(i) \leftrightarrow (j)$(第i个方程与第j个方程交换), $\lambda(i)$(第i个方程乘非零常数λ), $\lambda(i) \to (j)$(第i个方程乘λ加到第j个方程)表示. 经过上述三种变换, 我们可以将原线性方程组化为三角形(或阶梯形)的线性方程组, 最后利用回代法求出所有变量. 接下来, 我们再举两个例子.

例2.1.2. 解线性方程组

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

解. 我们以简记形式表示求解过程.

$$\begin{cases} x_1 & +2x_2+3x_3 & +4x_4=-3 & (1) \\ x_1 & +2x_2 & -5x_4=1 & (2) \\ 2x_1 & +4x_2-3x_3 & -19x_4=6 & (3) \\ 3x_1 & +6x_2-3x_3 & -24x_4=7 & (4) \end{cases} \xrightarrow[-3(1)\to(4)]{-1(1)\to(2),-2(1)\to(3)}$$

$$\begin{cases} x_1+2x_2+3x_3+4x_4=-3 & (5) \\ -3x_3-9x_4=4 & (6) \\ -9x_3-27x_4=12 & (7) \\ -12x_3-36x_4=16 & (8) \end{cases} \xrightarrow[-4(6)\to(8)]{-3(6)\to(7)}$$

$$\begin{cases} x_1+x_2+3x_3+4x_4=-3 & (9) \\ -3x_3-9x_4=4 & (10) \end{cases}$$

在最后一个方程组中, 由于后两个方程为$0=0$, 故予舍去. 因此, 变换后的方程组只由两个方程组成. 由后一个方程知道, x_3与x_4之一可以在数域F中任意取值. 不妨设$x_4=t_2$, 则$x_3=-\frac{4}{3}-3t_2$. 而由前一个方程知x_1与x_2之一可以在数域F中任意取值, 不妨设$x_2=t_1$. 将x_2,x_3,x_4代入前一个方程得$x_1=1-2t_1+5t_2$. 因此, 原方程组有多组解, 其解的一般表示形式为

$$\begin{cases} x_1=1-2t_1+5t_2 \\ x_2=t_1 \\ x_3=-\frac{4}{3}-3t_2 \\ x_4=t_2 \end{cases} \qquad t_1,t_2\in F \text{ 任意取值}$$

□

例2.1.3. 解线性方程组

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

解.

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

$$\begin{cases} x_1-x_2+x_3=1 & (4) \\ \qquad -2x_3=2 & (5) \\ \qquad -3x_3=1. & (6) \end{cases} \xrightarrow{-\frac{3}{2}(5)\to(6)}$$

$$\begin{cases} x_1-x_2+x_3=1 & (7) \\ \qquad -2x_3=2 & (8) \\ \qquad 0=-2. & (9) \end{cases}$$

最后一个方程组的最后一个方程$0=-2$是一个矛盾方程, 因此原方程组无解. □

关于Gauss消元求解线性方程组还有一个基本问题未解决, 即

原线性方程组与经过初等变换后的线性方程组是否有相同的解?

如果两个线性方程组有相同的解, 则我们称它们为**同解方程组**. 很显然, 原线性方程组的解一定是变换后的线性方程组的解. 但反过来结论是否成立? 经过变换后的线性方程组是否会产生“增根”? 要回答这个问题, 我们只需考察三个初等变换是否将原方程组化为同解方程组. 显然, 互换方程后方程组的解不变. 将某个方程乘一个非零常数λ也不会改变方程组的解, 因为将这个方程乘λ^{-1}后得到原方程. 对第三个初等变换$\lambda(i)\to(j)$, 设原方程组为

$$\begin{cases} l_1 = b_1 \\ \cdots\cdots \\ l_i = b_i \\ \cdots\cdots \\ l_m = b_m \end{cases}$$

其中$l_i(x_1,\ldots,x_n) := a_{i1}x_1 + a_{i2}x_2 + \ldots + a_{in}x_n$. 变换后的方程组为

$$\begin{cases} l_1 = b_1 \\ \cdots\cdots \\ \lambda l_i + l_j = \lambda b_i + b_j \\ \cdots\cdots \\ l_m = b_m \end{cases}$$

设$x_1 = c_1, x_2 = c_2, \ldots, x_n = c_n$是变换后方程组的解. 则$l_i(c_1,c_2,\ldots,c_n) = b_i$, $1 \le i \le m$, $i \ne j$, 且$\lambda l_i(c_1,c_2,\ldots,c_n) + l_j(c_1,c_2,\ldots,c_n) = \lambda b_i + b_j$. 于是也有$l_j(c_1,c_2,\ldots,c_n) = b_j$. 从而$x_1 = c_1, x_2 = c_2, \ldots, x_n = c_n$也是原方程组的解. 由此, 我们有以下结论.

命题2.1.1. 三种初等变换将线性方程组变为同解线性方程组, 因此不会产生增根.

§2.2 Gauss 消元的矩阵表示

观察Gauss消元法解线性方程组的过程, 其中的变元并不参与运算, 只是变元的系数与常数项参与运算, 因此可以省去变元与等号. 这样一个线性方程组可以表示成一个矩形表格的形式, 我们称之为**矩阵**. 一个**矩阵**是由若干行及若干列

的数构成的矩形阵列. 例如, 线性方程组(2.1)可以表示为矩阵形式:

$$\begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} & b_1 \\ a_{21} & a_{22} & \dots & a_{2n} & b_2 \\ \dots & \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & \dots & a_{mn} & b_m \end{pmatrix} \tag{2.2}$$

其中圆括号是为了表示矩阵是一个整体. 称矩阵(2.2)为线性方程组(2.1)的**增广矩阵**, 而矩阵(2.2)的m行及前n列构成的矩阵称为线性方程组(2.1)的**系数矩阵**. 容易观察到, 线性方程组(2.1)的第i个方程正好对应矩阵(2.2)的第i行. 这样对线性方程组做初等变换等价于对相应矩阵的行做相同的初等变换. 对应矩阵的三个**初等行变换**为:

1 互换矩阵的两行;

2 将某行乘一个非零常数;

3 将某行乘一个非零常数加到另一行.

我们分别用以下记号表示矩阵的三个初等行变换: $r_i \leftrightarrow r_j$(互换第i与第j两行), λr_i(第i行乘非零常数λ), $\lambda r_i \to r_j$(第i行乘λ加到第j行). 利用矩阵的初等行变换, 将矩阵化为三角或阶梯形式的矩阵, 就可以求出方程组的解了. 下面我们举几个例子加以说明.

例2.2.1. 重新考察例3.1.1. 其Gauss消元的矩阵形式如下:

解.

$$\begin{pmatrix} 3 & 2 & -1 & 6 \\ 1 & 3 & 2 & 9 \\ 2 & -1 & 3 & 3 \end{pmatrix} \xrightarrow[r_2 \leftrightarrow r_3]{r_1 \leftrightarrow r_2} \begin{pmatrix} 1 & 3 & 2 & 9 \\ 2 & -1 & 3 & 3 \\ 3 & 2 & -1 & 6 \end{pmatrix} \xrightarrow[-3r_1 \to r_3]{-2r_1 \to r_2}$$

$$\begin{pmatrix} 1 & 3 & 2 & 9 \\ 0 & -7 & -1 & -15 \\ 0 & -7 & -7 & -21 \end{pmatrix} \xrightarrow{-r_2 \to r_3} \begin{pmatrix} 1 & 3 & 2 & 9 \\ 0 & -7 & -1 & -15 \\ 0 & 0 & 6 & 6 \end{pmatrix} \xrightarrow{\frac{1}{6} r_3}$$

$$\begin{pmatrix} 1 & 3 & 2 & 9 \\ 0 & -7 & -1 & -15 \\ 0 & 0 & 1 & 1 \end{pmatrix} \xrightarrow[-2r_3 \to r_1]{r_3 \to r_2} \begin{pmatrix} 1 & 3 & 0 & 7 \\ 0 & -7 & 0 & -14 \\ 0 & 0 & 1 & 1 \end{pmatrix} \xrightarrow{-\frac{1}{7} r_2}$$

$$\begin{pmatrix} 1 & 3 & 0 & 7 \\ 0 & 1 & 0 & 2 \\ 0 & 0 & 1 & 1 \end{pmatrix} \xrightarrow{-3r_2 \to r_1} \begin{pmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 2 \\ 0 & 0 & 1 & 1 \end{pmatrix}$$

于是原线性方程组的解为$x_1=2, x_2=2, x_3=1$. □

例2.2.2. 解线性方程组

$$\begin{cases} x_1 + x_2 + x_3 + x_4 + x_5 = 0 \\ 3x_1 + 2x_2 + x_3 + x_4 - 3x_5 = 0 \\ x_2 + 2x_3 + 2x_4 + 6x_5 = 0 \\ 5x_1 + 4x_2 + 3x_3 + 3x_4 - x_5 = 0 \end{cases}$$

解.

$$\begin{pmatrix} 1 & 1 & 1 & 1 & 1 \\ 3 & 2 & 1 & 1 & -3 \\ 0 & 1 & 2 & 2 & 6 \\ 5 & 4 & 3 & 3 & -1 \end{pmatrix} \xrightarrow[-5r_1\to r_4]{-3r_1\to r_2} \begin{pmatrix} 1 & 1 & 1 & 1 & 1 \\ 0 & -1 & -2 & -2 & -6 \\ 0 & 1 & 2 & 2 & 6 \\ 0 & -1 & -2 & -2 & -6 \end{pmatrix}$$

$$\xrightarrow[-r_2\to r_4]{r_2\to r_3} \begin{pmatrix} 1 & 1 & 1 & 1 & 1 \\ 0 & -1 & -2 & -2 & -6 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

因此原方程组等价于

$$\begin{cases} x_1 + x_2 + x_3 + x_4 + x_5 = 0 \\ x_2 + 2x_3 + 2x_4 + 6x_5 = 0 \end{cases}$$

令$x_3=t_1, x_4=t_2, x_5=t_3$, 得$x_1=t_1+t_2+5t_3$, $x_2=-2t_1-2t_2-6t_3$. 即原方程组的解为:

$$\begin{cases} x_1 = t_1 + t_2 + 5t_3 \\ x_2 = -2t_1 - 2t_2 - 6t_3 \\ x_3 = t_1 \\ x_4 = t_2 \\ x_5 = t_3 \end{cases} \qquad t_1, t_2, t_3 \in F$$

□

例2.2.3. 解线性方程组

$$\begin{cases} 3x_1 - 2x_2 + 5x_3 + 4x_4 = 2 \\ 6x_1 - 7x_2 + 4x_3 + 3x_4 = 3 \\ 9x_1 - 9x_2 + 9x_3 + 7x_4 = -1 \end{cases}$$

解.

$$\begin{pmatrix} 3 & -2 & 5 & 4 & 2 \\ 6 & -7 & 4 & 3 & 3 \\ 9 & -9 & 9 & 7 & -1 \end{pmatrix} \xrightarrow[-3r_1\to r_3]{-2r_1\to r_2} \begin{pmatrix} 3 & -2 & 5 & 4 & 2 \\ 0 & -3 & -6 & -5 & -1 \\ 0 & -3 & -6 & -5 & -7 \end{pmatrix}$$

$$\xrightarrow{-r_2\to r_3} \begin{pmatrix} 3 & -2 & 5 & 4 & 2 \\ 0 & -3 & -6 & -5 & -1 \\ 0 & 0 & 0 & 0 & -6 \end{pmatrix}$$

从最后一个矩阵的最后一行看出, 原方程组无解. □

从上面的例子可以看到, 矩阵形式的Gauss消元本质上是对矩阵的行做两个基本运算: (1) 两行相加; (2) 某行乘一个非零常数. 而矩阵的每一行都可以视为一个$n+1$维数组向量, 因此Gauss消元就是对矩阵的行向量做线性运算, 变换后的矩阵的行向量都是原矩阵行向量的线性组合. 方程组的解也可以写成一些向量的线性组合, 本章中我们将方程组的解写成列向量形式. 例如, 例 (2.2.2)中方程组的解可以写成:

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{pmatrix} = \begin{pmatrix} 1 \\ -2 \\ 1 \\ 0 \\ 0 \end{pmatrix} t_1 + \begin{pmatrix} 1 \\ -2 \\ 0 \\ 1 \\ 0 \end{pmatrix} t_2 + \begin{pmatrix} 5 \\ -6 \\ 0 \\ 0 \\ 1 \end{pmatrix} t_3,$$

其中t_1,t_2,t_3为参数.

§2.3 一般线性方程组的Gauss 消元法

从前两节的例子看出, 线性方程组的解有三种情况: 唯一解, 多解(如数域F无限, 则有无穷多解) 及无解. 本节我们对一般的线性方程组给出Gauss消元的过程, 并由此判断线性方程组解的属性, 即有唯一解、多解还是无解.

§2.3.1 算法描述

给定线性方程组(2.1)或其矩阵形式(2.2), 我们先将它化为三角形(或阶梯形).

首先, $a_{i1}, i=1,2,\dots,m$中至少有一个非零元素, 否则原方程组不含变量x_1. 通过行的交换, 不妨设$a_{11}\neq 0$. 对矩阵(2.2)做如下初等行变换: $-a_{11}^{-1}a_{i1}r_1\to r_i$, $i=2,3,\dots,m$, 则矩阵(2.2)变为

$$\begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} & b_1 \\ 0 & a_{22}^{(1)} & \dots & a_{2n}^{(1)} & b_2^{(1)} \\ \dots & \dots & \dots & \dots & \dots \\ 0 & a_{m2}^{(1)} & \dots & a_{mn}^{(1)} & b_m^{(1)} \end{pmatrix}$$

如果$a_{i2}^{(1)}, i = 2,3,\dots,m$全为零, 则进一步考察$a_{i3}^{(1)}, i = 2,3,\dots,m$是否全为零. 假设$j_2$是满足如下条件的下标: 使得$a_{ij}^{(1)}, i = 2,3,\dots,m, j = 2,\dots, j_2 - 1$全为零, 但$a_{ij_2}^{(1)}$, $i = 2,3,\dots,m$中至少有一个非零. 通过行的互换, 不妨设$a_{2j_2}^{(1)} \neq 0$. 对上述矩阵做行变换$-(a_{2j_2}^{(1)})^{-1} a_{ij_2}^{(1)} r_2 \to r_i, i = 3,\dots,m$, 则矩阵进一步化为

$$\begin{pmatrix} a_{11} & \dots & \dots & \dots & \dots & \dots & \dots & \dots & a_{1n} & b_1 \\ 0 & \dots & 0 & a_{2j_2}^{(1)} & \dots & \dots & \dots & \dots & a_{2n}^{(1)} & b_2^{(1)} \\ 0 & \dots & \dots & 0 & \dots & 0 & a_{3j_3}^{(2)} & \dots & a_{3n}^{(2)} & b_3^{(2)} \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & \dots & \dots & 0 & \dots & 0 & a_{mj_3}^{(2)} & \dots & a_{mn}^{(2)} & b_m^{(2)} \end{pmatrix}$$

上述过程可以一直进行下去, 最后矩阵变为如下阶梯形矩阵

$$\begin{pmatrix} c_{11} & \dots & \dots & \dots & \dots & \dots & \dots & \dots & c_{1n} & d_1 \\ 0 & \dots & 0 & c_{2j_2} & \dots & \dots & \dots & \dots & c_{2n} & d_2 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & \dots & \dots & \dots & \dots & 0 & c_{rj_r} & \dots & c_{rn} & d_r \\ 0 & \dots & \dots & \dots & \dots & \dots & \dots & \dots & 0 & d_{r+1} \\ 0 & \dots & \dots & \dots & \dots & \dots & \dots & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & \dots & \dots & \dots & \dots & \dots & \dots & \dots & 0 & 0 \end{pmatrix} \tag{2.3}$$

其中$c_{11}, c_{2j_2}, \dots, c_{rj_r}$均非零. 我们称上述形式为矩阵的**最简形式**或**标准形式**. 当矩阵化为最简形式后, 我们就可以由此判断原线性方程组解的属性, 并给出其通解.

§2.3.2 线性方程组解的属性

利用矩阵的最简形式(2.3), 我们很方便得到线性方程组的解存在、唯一及有多个解的条件.

定理2.3.1. *线性方程组(2.1)的解的属性如下:*

1. 当$d_{r+1} \neq 0$时, 线性方程组(2.1)无解.

2. 当$d_{r+1}=0$且$r=n$时, 线性方程组(2.1)有唯一解.

3. 当$d_{r+1}=0$且$r<n$时, 线性方程组(2.1)有多解.

证明. (1) 当$d_{r+1}\neq 0$时, 矩阵(2.3)的第$r+1$行表示方程$0=d_{r+1}$. 此时, 原线性方程组无解.

(2) 当$d_{r+1}=0$且$r=n$时, 矩阵(2.3)具有形式

$$\begin{pmatrix} c_{11} & \dots & \dots & \dots & c_{1n} & d_1 \\ 0 & c_{22} & \dots & \dots & c_{2n} & d_2 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & \dots & \dots & 0 & c_{nn} & d_n \\ 0 & \dots & \dots & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & \dots & \dots & \dots & 0 & 0 \end{pmatrix}$$

对上述矩阵做行变换: $-c_{jj}^{-1}c_{ij}\,r_j\to r_i$, $i=1,2,\dots,j-1$, $j=n,n-1,\dots,2$. 则矩阵变为

$$\begin{pmatrix} c_{11} & 0 & \dots & \dots & 0 & \tilde{d}_1 \\ 0 & c_{22} & 0 & \dots & 0 & \tilde{d}_2 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & \dots & \dots & 0 & c_{nn} & \tilde{d}_n \\ 0 & \dots & \dots & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & \dots & \dots & \dots & 0 & 0 \end{pmatrix}$$

从而, 原线性方程组(2.1)有唯一解: $x_1=\tilde{d}_1/c_{11},\dots,x_n=\tilde{d}_n/c_{nn}$.

(3) 当$d_{r+1}=0$且$r<n$时, 矩阵(2.3)对应的线性方程组为

$$\begin{cases} c_{11}x_1+c_{12}x_2+\dots & \dots\dots & +c_{1n}x_n=d_1 \\ \qquad c_{2j_2}x_{j_2}+ & \dots\dots & +c_{2n}x_n=d_2 \\ & \dots\dots & \dots\dots \\ & c_{rj_r}x_{j_r}+\dots & +c_{rn}x_n=d_r \end{cases}$$

在上述方程组中, 将$x_{j_1},\dots,x_{j_r}(j_1=1)$视为变量, 而其余$x_j$(共$n-r$个)视为参数.

于是上述方程组可以改写为

$$\begin{cases} c_{11}x_1+c_{1j_2}x_{j_2}+ \quad \dots\dots+c_{1j_r}x_{j_r} = d_1 - c_{12}x_2 - \dots - c_{1n}x_n \\ \qquad c_{2j_2}x_{j_2}+ \quad \dots\dots+c_{2j_r}x_{j_r} = d_2 - c_{23}x_3 - \dots - c_{2n}x_n \\ \qquad\qquad \dots\dots\dots\dots\dots \\ \qquad\qquad\qquad c_{rj_r}x_{j_r} = d_r - c_{rj_r+1}x_{j_r+1} - \dots - c_{rn}x_n \end{cases}$$

用$t_1,\cdots,t_{n-r}$表示$n-r$个独立参数. 则由(2)知, 对于取定的一组参数值, 上述方程组的解唯一. 求解上述方程组, 可得线性方程组(2.1)的解具有如下形式

$$\begin{cases} x_1 = & \alpha_{11}t_1 + \dots + \alpha_{1,n-r}t_{n-r} + \beta_1 \\ x_2 = & \alpha_{21}t_1 + \dots + \alpha_{2,n-r}t_{n-r} + \beta_2 \\ \quad \cdots & \qquad\qquad \dots\dots \\ x_n = & \alpha_{n1}t_1 + \dots + \alpha_{n,n-r}t_{n-r} + \beta_n \end{cases} \tag{2.4}$$

或者写成向量形式

$$\mathbf{x} = \alpha_1 t_1 + \dots + \alpha_{n-r}t_{n-r} + \beta, \tag{2.5}$$

其中

$$\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}, \quad \alpha_i = \begin{pmatrix} \alpha_{1i} \\ \alpha_{2i} \\ \vdots \\ \alpha_{ni} \end{pmatrix}, \quad i = 1,2,\dots,n-r, \quad \beta = \begin{pmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_n \end{pmatrix}.$$

我们称(2.4)为线性方程组(2.1)的**通解**. 将参变量$t_1,\cdots,t_{n-r}$ 取定一组值代入(2.4)得到的解称为一个**特解**. 由于参变量可以任意取值, 此时线性方程组(2.1)有多解. □

对于齐次线性方程组, 由于总有$d_{r+1}=0$, 故线性方程组总有解, 特别地, $x_1 = 0,\dots,x_n = 0$ 就是齐次线性方程组的一组解, 我们称之为**零解**或**平凡解**. 齐次线性方程组的非零解称为**非平凡解**.

推论2.3.1. *齐次线性方程组有非零解的充要条件为$r<n$.*

证明. 当$r=n$时, 原齐次线性方程组有唯一解–零解. 当$r<n$时, 原齐次线性方程组有多解, 因而有非零解. □

推论2.3.2. *若$m<n$, 则齐次线性方程组一定有非零解.*

证明. 由于$r \leq m < n$, 由推论 2.3.1, 结论成立. □

从上述论证可以看出, 最简形式(2.3)中的整数r是一个非常重要的量. 它表示了线性方程组(2.1)独立方程的个数, 并且决定了解集的“大小”.

例2.3.1. 当λ为何值时, 下列线性方程组有解, 并求其通解.

$$\begin{cases} 2x_1 - x_2 + x_3 + x_4 = 1 \\ x_1 + 2x_2 - x_3 + 4x_4 = 2 \\ x_1 + 7x_2 - 4x_3 + 11x_4 = \lambda \end{cases}$$

解.

$$\begin{pmatrix} 2 & -1 & 1 & 1 & 1 \\ 1 & 2 & -1 & 4 & 2 \\ 1 & 7 & -4 & 11 & \lambda \end{pmatrix} \xrightarrow{r_1 \leftrightarrow r_2} \begin{pmatrix} 1 & 2 & -1 & 4 & 2 \\ 2 & -1 & 1 & 1 & 1 \\ 1 & 7 & -4 & 11 & \lambda \end{pmatrix} \xrightarrow[-r_1 \to r_3]{-2r_1 \to r_2}$$

$$\begin{pmatrix} 1 & 2 & -1 & 4 & 2 \\ 0 & -5 & 3 & -7 & -3 \\ 0 & 5 & -3 & 7 & \lambda - 2 \end{pmatrix} \xrightarrow{r_2 \to r_3} \begin{pmatrix} 1 & 2 & -1 & 4 & 2 \\ 0 & -5 & 3 & -7 & -3 \\ 0 & 0 & 0 & 0 & \lambda - 5 \end{pmatrix}$$

因此, 当$\lambda = 5$时原线性方程组有解. 此时, 原方程组等价于

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

令$x_3 = t_1, x_4 = t_2$, 解得

$$\begin{cases} x_1 = -\frac{1}{5}t_1 - \frac{6}{5}t_2 + \frac{4}{5} \\ x_2 = \frac{3}{5}t_1 - \frac{7}{5}t_2 + \frac{3}{5} \\ x_3 = t_1 \\ x_4 = t_2 \end{cases} \qquad t_1, t_2 \in F$$

□

或写成向量形式

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} -\frac{1}{5} \\ \frac{3}{5} \\ 1 \\ 0 \end{pmatrix} + \begin{pmatrix} -\frac{6}{5} \\ -\frac{7}{5} \\ 0 \\ 1 \end{pmatrix} + \begin{pmatrix} \frac{4}{5} \\ \frac{3}{5} \\ 0 \\ 0 \end{pmatrix}.$$

我们现在回头看看本章开头提出的几个基本问题:

1. 解的存在性与唯一性问题已解决;

2. 求解方法以及通解的表达形式已解决.

我们还需继续研究线性方程组的公式解、Gauss消元的矩阵表示以及解的几何结构. 这些问题将留待第三章、第四章与第五章解决.

关于线性方程组的解集的几何结构问题, 我们先对$n = 3$的情形做一点初步的分析. 对于$n = 3$, 由于每个方程表示三维空间中的一个平面, 因此方程组的解集将是一些平面的交集. 因此解集可以是一个平面, 一条直线, 一个点或空集. 这里, r(独立方程的个数)是决定解集"大小"(即是平面, 还是直线还是点)的一个非常重要的量!

虽然我们解决了线性方程组的解的属性(存在性, 唯一性等)问题以及求解方法, 但是我们是对原线性方程组做了一系列初等变换将它化成阶梯形方程组后再判断, 求解. 能否不做初等变换, 直接从原线性方程组出发判断解的属性, 并给出线性方程组的解? 为此, 我们重新提出以下**几个新的问题**:

1. 如何从原方程组直接判别解的存在性、唯一性及多解性?

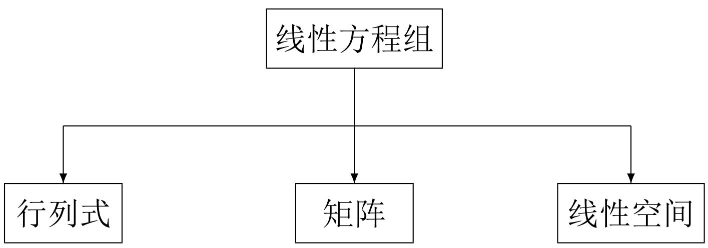
2. 如何从原方程组直接确定r?

3. r是否唯一?

4. 解集的大小与r有何关系?

5. 直接从原方程获得公式解.

为研究线性方程组的公式解, 我们将引入行列式的概念. 为研究线性方程组解的属性(存在性, 唯一性等) 及解集的结构, 我们将引入线性空间的概念与有关理论. 为研究线性方程组求解, 我们将引入矩阵的运算(特别是乘法运算). 以上问题将引出我们以下各章的内容, 我们用下列图形表示:



小知识:《九章算术》

《九章算术》是我国古代一部数学专著,成书于公元一世纪,作者不详.西汉张苍、耿寿昌做过增补与整理,魏晋时期刘徽(公元225年-295年)为《九章》做注,称为《九章算术注》.

《九章算术》共收集了246个应用问题和各种问题的解法,分别隶属于方田、粟米、衰分、少广、商功、均输、盈不足、方程、勾股九章.《九章算术》系统地总结了我国从先秦到东汉的数学成就.其中负数、分数计算,联立一次方程(线性方程组)解法等都是具有世界意义的数学成就.

在西方,线性方程组的研究是在17世纪后期由莱布尼茨开创的,18世纪Gauss提出了求解线性方程组的Gauss消元法,其本质与《九章算术》是相同的.中国在线性方程组的研究比西方早1500 余年.

习题一

1. 解下列线性方程组

$$(1)\quad \begin{cases} x_1+x_2-3x_3=-1 \\ 2x_1+x_2-2x_3=1 \\ x_1+x_2+x_3=3 \\ x_1+2x_2-3x_3=1 \end{cases} \qquad (2)\quad \begin{cases} x_1-2x_2+3x_3-4x_4=4 \\ x_2-x_3+x_4=-3 \\ x_1+3x_2-3x_4=1 \\ -7x_2+3x_3+x_4=-3 \end{cases}$$

$$(3)\quad \begin{cases} 2x_1+3x_2-x_3+x_4=1 \\ 8x_1+12x_2-9x_3+8x_4=3 \\ 4x_1+6x_2+3x_3-2x_4=3 \\ 2x_1+3x_2+9x_3-7x_4=3 \end{cases} \qquad (4)\quad \begin{cases} 2x_1+4x_2-6x_3+x_4=2 \\ x_1-x_2+4x_3+x_4=1 \\ -x_1+x_2-x_3+x_4=0 \end{cases}$$

$$(5)\quad \begin{cases} 2x_1-3x_2+7x_3+5x_4=7 \\ x_1-x_2+x_3-x_4=1 \\ x_1+x_2-8x_3+x_4=0 \\ 4x_1-3x_2+5x_4=0 \end{cases} \qquad (6)\quad \begin{cases} 3x_1-5x_2+x_3-2x_4=0 \\ 2x_1+3x_2-5x_3+x_4=0 \\ -x_1+7x_2-4x_3+3x_4=0 \\ 4x_1+15x_2-7x_3+9x_4=0 \end{cases}$$

$$(7)\quad \begin{cases} x_1+x_2-3x_4-x_5=0 \\ x_1-x_2-2x_3-x_4=0 \\ 4x_1-2x_2+6x_3+3x_4-4x_5=0 \\ 2x_1+4x_2-2x_3+4x_4-7x_5=0 \end{cases}$$

2. 当a为何值时, 下列线性方程组有解? 有解时求出它的通解.

$$(1)\quad \begin{cases} 3x_1+2x_2+x_3=2 \\ x_1-x_2-2x_3=-3 \\ ax_1-2x_2+2x_3=6 \end{cases} \qquad (2)\quad \begin{cases} x_1-4x_2+2x_3=-1 \\ -x_1+11x_2-x_3=3 \\ 3x_1-5x_2+7x_3=a \end{cases}$$

3. a为何值时, 下述线性方程组有唯一解? a为何值时, 此方程组无解?

$$\begin{cases} x_1+x_2+x_3=3 \\ x_1+2x_2-ax_3=9 \\ 2x_1-x_2+3x_3=6 \end{cases}$$

4. 求三次多项式$f(x)=ax^3+bx^2+cx+d$, 使$y=f(x)$的图象经过以下4个点:

   $A(1,2), B(-1,3), C(3,0), D(0,2)$.

5. 求三次多项式$f(x)=ax^3+bx^2+cx+d$满足

$$f(0)=1, f(1)=2, f'(0)=1, f'(1)=-1.$$

6. 给定线性方程组

$$\begin{cases} x_1+2x_2-3x_3+4x_4=2 \\ 2x_1+5x_2-2x_3+x_4=1 \\ 3x_1+8x_2-x_3-2x_4=0 \end{cases}$$

   将常数项改为零得到另一个方程组, 求解这两个方程组, 并研究这两个方程组解之间的关系. 对其它方程组做类似的讨论.

7. 兽医建议某宠物的食谱每天要包含100单位的蛋白质, 200单位的糖, 50单位的脂肪. 某宠物商店出售四种食品A,B,C,D. 这四种食物每千克含蛋白质、

糖、脂肪的含量(单位)如下

食物	蛋白质	糖	脂肪
A	5	20	2
B	4	25	2
C	7	10	10
D	10	5	6

问是否可以适量配备上述四种食品,满足兽医的建议.

第三章　行列式

行列式的概念最早出现于线性方程组的求解, 它是由日本数学家关孝和及德国数学家莱布尼茨于十七世纪末期提出, 现已成为线性代数的基本工具之一. 行列式在坐标变换、多重积分中的变量替换、二次型及多元多项式消元等有重要的应用. 本章将介绍行列式的定义、基本性质以及行列式的计算. 作为行列式的应用, 我们将给出求解线性方程组公式解的Cramer法则.

§3.0　行列式的引入

我们从求线性方程组的公式解入手. 考察n个变元与n个方程的线性方程组

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2 \\ \cdots\cdots\cdots\cdots\cdots\cdots \\ a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n = b_n \end{cases} \tag{3.1}$$

的公式解.

当$n = 2$时, 上述线性方程组化为

$$\begin{cases} a_{11}x_1 + a_{12}x_2 &= b_1 \\ a_{21}x_1 + a_{22}x_2 &= b_2. \end{cases}$$

将第一个方程乘a_{22}减去第二个方程乘a_{12}得

$$(a_{11}a_{22} - a_{12}a_{21})x_1 = b_1a_{22} - b_2a_{12}.$$

类似地, 将第二个方程乘a_{11}减去第一个方程乘a_{21}得

$$(a_{11}a_{22} - a_{12}a_{21})x_2 = b_2a_{11} - b_1a_{21}.$$

因此, 当$a_{11}a_{22} - a_{12}a_{21} \neq 0$时, 方程组有唯一解

$$\begin{cases} x_1 &= \dfrac{b_1a_{22} - b_2a_{12}}{a_{11}a_{22} - a_{12}a_{21}} \\ x_2 &= \dfrac{b_2a_{11} - b_1a_{21}}{a_{11}a_{22} - a_{12}a_{21}} \end{cases}$$

引入记号

$$\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}a_{22} - a_{12}a_{21}, \quad \begin{vmatrix} b_1 & a_{12} \\ b_2 & a_{22} \end{vmatrix} = b_1a_{22} - b_2a_{12}, \quad \begin{vmatrix} a_{11} & b_1 \\ a_{21} & b_2 \end{vmatrix} = b_2a_{11} - b_1a_{21}.$$

则二阶线性方程组的公式解可以简洁地表示为

$$x_1=\frac{\begin{vmatrix}b_1 & a_{12}\\ b_2 & a_{22}\end{vmatrix}}{\begin{vmatrix}a_{11} & a_{12}\\ a_{21} & a_{22}\end{vmatrix}},\qquad x_2=\frac{\begin{vmatrix}a_{11} & b_1\\ a_{21} & b_2\end{vmatrix}}{\begin{vmatrix}a_{11} & a_{12}\\ a_{21} & a_{22}\end{vmatrix}}.$$

我们称上式中的分子与分母为二阶行列式. 二阶行列式也可以视为2×2阶矩阵的行列式. 一般地, 设矩阵

$$A=\begin{pmatrix}a_1 & a_2\\ b_1 & b_2\end{pmatrix}$$

则A的行列式(记为$\det(A)$或$|A|$)为

$$\det(A)=\begin{vmatrix}a_1 & a_2\\ b_1 & b_2\end{vmatrix}=a_1b_2-a_2b_1.$$

下面我们研究三阶线性方程组的公式解. 考虑三阶线性方程组

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

将第一个方程乘A_1, 第二个方程乘A_2, 第三个方程乘A_3, 然后相加得

$$\begin{aligned}&(a_{11}A_1+a_{21}A_2+a_{31}A_3)x_1+(a_{12}A_1+a_{22}A_2+a_{32}A_3)x_2+\\ &(a_{13}A_1+a_{23}A_2+a_{33}A_3)x_3=b_1A_1+b_2A_2+b_3A_3.\end{aligned}$$

为了消去变元x_2与x_3, 令

$$a_{12}A_1+a_{22}A_2+a_{32}A_3=0,\qquad a_{13}A_1+a_{23}A_2+a_{33}A_3=0.$$

将上式改写成关于A_1,A_2的线性方程组

$$\begin{cases}a_{12}A_1+a_{22}A_2=-a_{32}A_3\\ a_{13}A_1+a_{23}A_2=-a_{33}A_3\end{cases}$$

解得

$$A_1=\frac{\begin{vmatrix}a_{22} & a_{23}\\ a_{32} & a_{33}\end{vmatrix}}{\begin{vmatrix}a_{12} & a_{13}\\ a_{22} & a_{23}\end{vmatrix}}A_3,\quad A_2=-\frac{\begin{vmatrix}a_{12} & a_{13}\\ a_{32} & a_{33}\end{vmatrix}}{\begin{vmatrix}a_{12} & a_{13}\\ a_{22} & a_{23}\end{vmatrix}}A_3.$$

因此, 令

$$A_1 = \begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix}, \qquad A_2 = -\begin{vmatrix} a_{12} & a_{13} \\ a_{32} & a_{33} \end{vmatrix}, \qquad A_3 = \begin{vmatrix} a_{12} & a_{13} \\ a_{22} & a_{23} \end{vmatrix}.$$

即得

$$\begin{aligned} &\left(a_{11}\begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} - a_{21}\begin{vmatrix} a_{12} & a_{13} \\ a_{32} & a_{33} \end{vmatrix} + a_{31}\begin{vmatrix} a_{12} & a_{13} \\ a_{22} & a_{23} \end{vmatrix}\right)x_1 \\ &= b_1\begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} - b_2\begin{vmatrix} a_{12} & a_{13} \\ a_{32} & a_{33} \end{vmatrix} + b_3\begin{vmatrix} a_{12} & a_{13} \\ a_{22} & a_{23} \end{vmatrix}. \end{aligned}$$

引进记号

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

$$\begin{vmatrix} b_1 & a_{12} & a_{13} \\ b_2 & a_{22} & a_{23} \\ b_3 & a_{32} & a_{33} \end{vmatrix} := b_1\begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} - b_2\begin{vmatrix} a_{12} & a_{13} \\ a_{32} & a_{33} \end{vmatrix} + b_3\begin{vmatrix} a_{12} & a_{13} \\ a_{22} & a_{23} \end{vmatrix}.$$

则得

$$x_1 = \frac{\begin{vmatrix} b_1 & a_{12} & a_{13} \\ b_2 & a_{22} & a_{23} \\ b_3 & a_{32} & a_{33} \end{vmatrix}}{\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}}$$

类似可得

$$x_2 = \frac{\begin{vmatrix} a_{11} & b_1 & a_{13} \\ a_{21} & b_2 & a_{23} \\ a_{31} & b_3 & a_{33} \end{vmatrix}}{\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}}, \qquad x_3 = \frac{\begin{vmatrix} a_{11} & a_{12} & b_1 \\ a_{21} & a_{22} & b_2 \\ a_{31} & a_{32} & b_3 \end{vmatrix}}{\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}}$$

其中

$$\begin{vmatrix} a_{11} & b_1 & a_{13} \\ a_{21} & b_2 & a_{23} \\ a_{31} & b_3 & a_{33} \end{vmatrix} := -b_1 \begin{vmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{vmatrix} + b_2 \begin{vmatrix} a_{11} & a_{13} \\ a_{31} & a_{33} \end{vmatrix} - b_3 \begin{vmatrix} a_{11} & a_{13} \\ a_{21} & a_{23} \end{vmatrix}.$$

$$\begin{vmatrix} a_{11} & a_{12} & b_1 \\ a_{21} & a_{22} & b_2 \\ a_{31} & a_{32} & b_3 \end{vmatrix} := b_1 \begin{vmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{vmatrix} - b_2 \begin{vmatrix} a_{11} & a_{12} \\ a_{31} & a_{32} \end{vmatrix} + b_3 \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}.$$

x_1,x_2,x_3的分子与分母均称为三阶行列式. 可以看到, 三阶行列式都可以表达为二阶行列式的求和形式, 并且是按某一列进行展开得到的.

利用归纳法, 一般地n元线性方程组(3.1)的公式解可以表示为两个n阶行列式的比, 而n阶行列式可以表示为$n-1$阶行列式求和的形式. 据此, 我们用递归的方式引入n阶行列式的定义.

§3.1 行列式的定义

一个行数与列数相等的矩阵称之为**方阵**, 由n行与n列组成的方程$A=(a_{ij})$称为n**阶**方阵, 其中a_{ij}表示位置(i,j)的**元素**. 由前一节的讨论, 我们可以用递归的方式定义方阵A的行列式.

定义3.1.1. n阶方阵$A=(a_{ij})$的**行列式**通常记为

$$\det(A) \quad 或 \quad \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \cdots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix}.$$

当$n=1$时, $\det(A)$定义为a_{11}. 当$n\geq 2$时, $\det(A)$递归地定义为

$$\det(A) := \sum_{i=1}^{n} (-1)^{i+1} a_{i1} M_{i1} = \sum_{i=1}^{n} a_{i1} A_{i1}, \tag{3.2}$$

其中

$$M_{ij} = \begin{vmatrix} a_{11} & \cdots & a_{1,j-1} & a_{1,j+1} & \cdots & a_{1n} \\ \vdots & \cdots & \vdots & \vdots & \ldots & \vdots \\ a_{i-1,1} & \cdots & a_{i-1,j-1} & a_{i-1,j+1} & \cdots & a_{i-1,n} \\ a_{i+1,1} & \cdots & a_{i+1,j-1} & a_{i+1,j+1} & \cdots & a_{i+1,n} \\ \vdots & \cdots & \vdots & \vdots & \ldots & \vdots \\ a_{n1} & \cdots & a_{n,j-1} & a_{n,j+1} & \cdots & a_{nn}, \end{vmatrix}, \tag{3.3}$$

即M_{ij}是删除A的第i行与第j列所得$n-1$阶方阵的行列式, 称为行列式$\det(A)$的元素a_{ij}的**余子式**, $A_{ij}=(-1)^{i+j}M_{ij}$称为a_{ij}的**代数余子式**. 定义 (3.2)表明, 一个n阶行列式可表示为第一列元素与之对应代数余子式($n-1$阶行列式)乘积的和. 显然, 2阶和3阶行列式的定义与上一节中引入的定义是一致的.

例3.1.1. 计算行列式$\begin{vmatrix} 2 & 1 & 0 \\ 1 & 2 & 1 \\ 0 & 1 & 2 \end{vmatrix}$.

解:

$$\begin{aligned}\begin{vmatrix} 2 & 1 & 0 \\ 1 & 2 & 1 \\ 0 & 1 & 2 \end{vmatrix} &= 2\times\begin{vmatrix} 2 & 1 \\ 1 & 2 \end{vmatrix} - 1\times\begin{vmatrix} 1 & 0 \\ 1 & 2 \end{vmatrix} + 0\times\begin{vmatrix} 1 & 0 \\ 2 & 1 \end{vmatrix} \\ &= 2\times(2\times 2-1\times 1)-1\times(1\times 2-0\times 1)=2\times 3-2=4.\end{aligned}$$

例3.1.2. 证明:

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

证明.

$$\begin{aligned}\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} &= a_{11}\begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} - a_{21}\begin{vmatrix} a_{12} & a_{13} \\ a_{32} & a_{33} \end{vmatrix} + a_{31}\begin{vmatrix} a_{12} & a_{13} \\ a_{22} & a_{23} \end{vmatrix} \\ &= a_{11}(a_{22}a_{33}-a_{23}a_{32})-a_{21}(a_{12}a_{33}-a_{13}a_{32})+a_{31}(a_{12}a_{23}-a_{13}a_{32}) \\ &= a_{11}a_{22}a_{33}-a_{11}a_{23}a_{32}-a_{21}a_{12}a_{33}+a_{21}a_{13}a_{32}+a_{31}a_{12}a_{23}-a_{31}a_{13}a_{32} \\ &= a_{11}a_{22}a_{33}+a_{12}a_{23}a_{31}+a_{13}a_{21}a_{32}-a_{11}a_{23}a_{32}-a_{12}a_{21}a_{33}-a_{13}a_{22}a_{31}.\end{aligned}$$

从上可以看到, 三阶行列的展开式中总共有6项, 每项元素具有形式$a_{1j_1}a_{2j_2}a_{3j_3}$, 其中j_1,j_2,j_3是$1,2,3$的一个排列. 当$(j_1j_2j_3)=(123),(231),(312)$时, 对应的项符号为正. 当$(j_1j_2j_3)=(132),(213),(321)$时, 对应的项符号为负. 实际上, 这不是偶然的. 后面我们将看到, 一般n阶行列式有类似的展开形式. □

例3.1.3. 证明:

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

特别地,

$$\begin{vmatrix} a_{11} & 0 & \cdots & 0 \\ 0 & a_{22} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & a_{nn} \end{vmatrix} = a_{11}a_{22}\cdots a_{nn}, \qquad \begin{vmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{vmatrix} = 1.$$

证明. 根据行列式的定义,

$$\begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ 0 & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & a_{nn} \end{vmatrix} = a_{11} \begin{vmatrix} a_{22} & a_{23} & \cdots & a_{2n} \\ 0 & a_{33} & \cdots & a_{3n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & a_{nn} \end{vmatrix} = \ldots\ldots = a_{11}a_{22}\cdots a_{nn}. \qquad \square$$

在行列式的定义中, n阶行列式是按照第一列元素展开成n个$n-1$阶行列式的和. 实际上, 行列式也可以按照任意一列展开, 我们有

定理3.1.1. 设$A=(a_{ij})_{n\times n}$为n阶方阵. 则对固定的k($1\leq k\leq n$)有

$$\det(A)=\sum_{i=1}^{n}a_{ik}A_{ik}=\sum_{i=1}^{n}(-1)^{i+k}a_{ik}M_{ik}, \tag{3.4}$$

即行列式可以按任意列第k列展开.

证明. 对行列式的阶数n用数学归纳法. 容易验证, 2阶行列式按第一列展开与按第二列展开结果相同. 现假设结论对$n-1$成立, 即$n-1$阶行列式按任意一列展开结果相同. 我们要证结论对n阶行列式也成立.

固定$2\leq k\leq n$, 用D_{ij}表示删除A的第$1,k$列与第i,j行所得$n-2$阶矩阵的行列式, 则显然有$D_{ij}=D_{ji}$. 为证明(3.4), 我们将等式两边都展开成$n-2$阶行列式D_{ij}的和的形式.

首先, 由行列式的定义,

$$\det(A)=\sum_{i=1}^{n}(-1)^{i+1}a_{i1}M_{i1}.$$

由归纳假设, 将$n-1$阶行列式M_{i1}按第$k-1$列展开

$$M_{i1}=\sum_{j=1}^{i-1}(-1)^{j+(k-1)}a_{jk}D_{ij}+\sum_{j=i+1}^{n}(-1)^{(j-1)+(k-1)}a_{jk}D_{ij}.$$

因此

$$
\begin{aligned}
\det(A) &= \sum_{i=1}^{n}(-1)^{i+1}a_{i1}M_{i1} \\
&= \sum_{i=1}^{n}\sum_{j=1}^{i-1}(-1)^{i+j+k}a_{i1}a_{jk}D_{ij} + \sum_{i=1}^{n}\sum_{j=i+1}^{n}(-1)^{i+j+k-1}a_{i1}a_{jk}D_{ij} \\
&= \sum_{i=1}^{n}\sum_{j=1}^{i-1}(-1)^{i+j+k}a_{i1}a_{jk}D_{ij} + \sum_{j=1}^{n}\sum_{i=1}^{j-1}(-1)^{i+j+k-1}a_{i1}a_{jk}D_{ij} \qquad (3.5)\\
&= \sum_{i=1}^{n}\sum_{j=1}^{i-1}(-1)^{i+j+k}(a_{i1}a_{jk}-a_{j1}a_{ik})D_{ij}.
\end{aligned}
$$

其中对第二行第二个式子中求和号$\sum_{i=1}^{n}\sum_{j=i+1}^{n}$进行交换变成了$\sum_{j=1}^{n}\sum_{i=1}^{j-1}$, 而对第三行第二个式子的指标$i,j$做对换并与第一个式子合并构成了第四行的式子.

另一方面, 将$n-1$阶行列式M_{ik}按第一列展开得

$$
M_{ik} = \sum_{j=1}^{i-1}(-1)^{j+1}a_{j1}D_{ij} + \sum_{j=i+1}^{n}(-1)^{j}a_{j1}D_{ij}.
$$

因此,

$$
\begin{aligned}
&\sum_{i=1}^{n}(-1)^{i+k}a_{ik}M_{ik} \\
&= \sum_{i=1}^{n}\sum_{j=1}^{i-1}(-1)^{i+j+k+1}a_{ik}a_{j1}D_{ij} + \sum_{i=1}^{n}\sum_{j=i+1}^{n}(-1)^{i+j+k}a_{ik}a_{j1}D_{ij} \\
&= \sum_{i=1}^{n}\sum_{j=1}^{i-1}(-1)^{i+j+k+1}a_{ik}a_{j1}D_{ij} + \sum_{j=1}^{n}\sum_{i=1}^{j-1}(-1)^{i+j+k}a_{ik}a_{j1}D_{ij} \qquad (3.6)\\
&= \sum_{i=1}^{n}\sum_{j=1}^{i-1}(-1)^{i+j+k}(a_{i1}a_{jk}-a_{j1}a_{ik})D_{ij}.
\end{aligned}
$$

比较(3.5)与(3.6)两式即知结论成立. 定理证毕. □

例3.1.4. 证明

$$
\begin{vmatrix}
a_{11} & a_{12} & \cdots & a_{1n} \\
\vdots & \vdots & \iddots & \\
a_{n-1,1} & a_{n-1,2} & & \\
a_{n1} & & &
\end{vmatrix}
= (-1)^{\frac{n(n-1)}{2}}a_{n1}a_{n-1,2}\cdots a_{1n}.
$$

证明. 将行列式逐次按第$n,n-1,\ldots,2$列展开得

$$\begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & \iddots & \\ a_{n-1,1} & a_{n-1,2} & & \\ a_{n1} & & & \end{vmatrix} = (-1)^{n+1}a_{1n}\begin{vmatrix} a_{21} & a_{22} & \cdots & a_{2,n-1} \\ \vdots & \vdots & \iddots & \\ a_{n-1,1} & a_{n-1,2} & & \\ a_{n1} & & & \end{vmatrix}$$

$$= \ldots\ldots = (-1)^{(1+n)+n+\ldots+3}a_{1n}a_{2,n-1}\cdots a_{n1} = (-1)^{\frac{n(n-1)}{2}}a_{n1}a_{n-1,2}\cdots a_{1n}.$$

□

行列式不仅可以按列展开, 还可以按行展开. 事实上, 我们有

定理3.1.2. 设$A=(a_{ij})$为n阶方阵. 则

$$\det(A) = \sum_{j=1}^{n} a_{kj}A_{kj} = \sum_{j=1}^{n}(-1)^{k+j}a_{kj}M_{kj}, \qquad 1 \le k \le n. \tag{3.7}$$

证明. 不妨设$k=1$. 我们需要证明:

$$\sum_{j=1}^{n}(-1)^{j+1}a_{1j}M_{1j} = \sum_{i=1}^{n}(-1)^{i+1}a_{i1}M_{i1}. \tag{3.8}$$

对n用数学归纳法. 当$n=1,2$时, 易验证结论成立. 现假设结论对$n-1$阶行列式成立, 我们要证结论对n阶行列式成立.

用D_{ij}表示删除第1行, 第i行, 第1列, 第j列所得到的$n-2$阶子行列式. 一方面, 将M_{1j}按第一列展开得

$$M_{1j} = \sum_{i=2}^{n}(-1)^{i}a_{i1}D_{ij}, \qquad j=2,\ldots,n.$$

因此,

$$\begin{aligned} &\sum_{j=1}^{n}(-1)^{j+1}a_{1j}M_{1j} = a_{11}M_{11} + \sum_{j=2}^{n}(-1)^{j+1}a_{1j}\sum_{i=2}^{n}(-1)^{i}a_{i1}D_{ij} \\ &= a_{11}M_{11} + \sum_{i=2}^{n}\sum_{j=2}^{n}(-1)^{i+j+1}a_{i1}a_{1j}D_{ij}. \end{aligned} \tag{3.9}$$

另一方面, 由归纳假设, 对M_{i1}按第一行展开

$$M_{i1} = \sum_{j=2}^{n}(-1)^{j}a_{1j}D_{ij}.$$

因此,

$$\begin{aligned} &\sum_{i=1}^{n}(-1)^{i+1}a_{i1}M_{i1} = a_{11}M_{11} + \sum_{i=2}^{n}(-1)^{i+1}a_{i1}\sum_{j=2}^{n}(-1)^{j}a_{1j}D_{ij} \\ &= a_{11}M_{11} + \sum_{i=2}^{n}\sum_{j=2}^{n}(-1)^{i+j+1}a_{i1}a_{1j}D_{ij}. \end{aligned} \tag{3.10}$$

由 (3.9)和 (3.10)知, (3.8)式的两边相等. 证毕. □

由定理 3.1.1与定理 3.1.2, 行列式即可以按任意一行展开也可以按任意一列展开. 实际上, (3.4)式与(3.7) 式都可以作为行列式的定义.

定义3.1.2. 设$A=(a_{ij})_{n\times n}$为n阶方阵, 固定$k(1\le k\le n)$. A的行列式 $\det(A)$ 可递归定义如下: 当$n=1$时, $\det(A)=a_{11}$; 当$n\ge 2$ 时,

$$\det(A)=\sum_{i=1}^{n}(-1)^{i+k}a_{ik}M_{ik} \tag{3.11}$$

其中M_{ik}是a_{ik}的余子式.

定义3.1.3. 设$A=(a_{ij})_{n\times n}$为n阶方阵, 固定$k(1\le k\le n)$. A的行列式 $\det(A)$ 可递归定义如下: 当$n=1$时, $\det(A)=a_{11}$; 当$n\ge 2$ 时,

$$\det(A)=\sum_{j=1}^{n}(-1)^{k+j}a_{kj}M_{kj}, \tag{3.12}$$

其中M_{kj}是a_{kj}的余子式.

易知, 上述定义与定义 3.1.1 是等价的.

最后, 我们来研究一下行列式的几何含义.

在平面直角坐标系oxy中, 令向量$\overrightarrow{OA}=(a_{11},a_{12})$, $\overrightarrow{OB}=(a_{21},a_{22})$. 以$OA,OB$为两邻边的平行四边形为$OABC$(如图 3.1 所示). 将$\overrightarrow{OB}$绕原点$O$顺时针旋转$90^\circ$得到向量$\overrightarrow{OB'}=(a_{22},-a_{21})$. 则

$$\begin{aligned}\begin{vmatrix}a_{11} & a_{12}\\ a_{21} & a_{22}\end{vmatrix} &= a_{11}a_{22}-a_{12}a_{21}=\overrightarrow{OA}\cdot\overrightarrow{OB'}\\ &=|OA|\,|OB|\cos\angle B'OA=|OA|\,|OB|\sin\angle AOB=S_{OABC}.\end{aligned}$$

这里S_{OABC}是平行四边形的有(带符号)向面积, 它的符号由$\sin\angle AOB$确定. 因此, 二阶行列式就是以行列式的两个行向量为边的平行四边形的有向面积.

接下来, 我们研究三阶行列式的几何意义. 在空间直角坐标系$oxyz$中, 设$\overrightarrow{OA}=(a_{11},a_{12},a_{13})$, $\overrightarrow{OB}=(a_{21},a_{22},a_{23})$, $\overrightarrow{OC}=(a_{31},a_{32},a_{33})$(如图 3.2所示). 则

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

$$(a_{11},a_{12},a_{13})\cdot\left(\begin{vmatrix}a_{22} & a_{23}\\ a_{32} & a_{33}\end{vmatrix},-\begin{vmatrix}a_{21} & a_{23}\\ a_{31} & a_{33}\end{vmatrix},\begin{vmatrix}a_{21} & a_{22}\\ a_{31} & a_{32}\end{vmatrix}\right)=\overrightarrow{OA}\cdot\overrightarrow{OB}\times\overrightarrow{OC}=V_{OABC}.$$

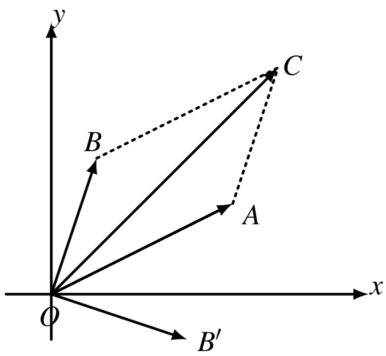


图 3.1 二阶行列式的几何意义

其中V_{OABC}是以OA, OB, OC为棱边的平行六面体的有向体积. 这就是三阶行列式的几何意义. 类似地可以想象, n阶行列式是n维空间中以行列式的行向量为棱边的平行多变体体积.

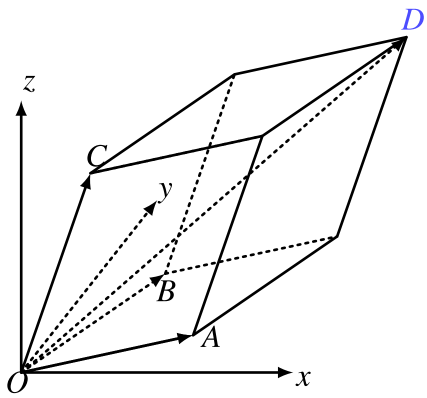


图 3.2 三阶行列式的几何意义

§3.2 行列式的性质

行列式有很多重要的性质, 这些性质对于理解行列式的概念、计算行列式, 以及行列式的应用起着至关重要的作用.

定理3.2.1. 行列式具有下列性质:

(1) 任意方阵与它的转置方阵(行列互换所得到的方阵)的行列式相等, 即

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

(2) 交换方阵A的某两行, 得到的方阵B满足$\det(B) = -\det(A)$, 即

$$\begin{vmatrix} \cdots & \cdots & \cdots & \cdots \\ a_{q1} & a_{q2} & \cdots & a_{qn} \\ \cdots & \cdots & \cdots & \cdots \\ a_{p1} & a_{p2} & \cdots & a_{pn} \\ \cdots & \cdots & \cdots & \cdots \end{vmatrix} = -\begin{vmatrix} \cdots & \cdots & \cdots & \cdots \\ a_{p1} & a_{p2} & \cdots & a_{pn} \\ \cdots & \cdots & \cdots & \cdots \\ a_{q1} & a_{q2} & \cdots & a_{qn} \\ \cdots & \cdots & \cdots & \cdots \end{vmatrix}.$$

(3) 将方阵A的某一行乘以常数λ, 得到的方阵B满足$\det(B) = \lambda\det(A)$, 即

$$\begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \cdots & \cdots & \cdots & \cdots \\ \lambda a_{k1} & \lambda a_{k2} & \cdots & \lambda a_{kn} \\ \cdots & \cdots & \cdots & \cdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix} = \lambda \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{k1} & a_{k2} & \cdots & a_{kn} \\ \cdots & \cdots & \cdots & \cdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix}.$$

(4) 若方阵A的某一行是两个向量之和, 则$\det(A)$可拆成两个行列式之和, 即

$$\begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \cdots & \cdots & \cdots & \cdots \\ b_1+c_1 & b_2+c_2 & \cdots & b_n+c_n \\ \cdots & \cdots & \cdots & \cdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix} = \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \cdots & \cdots & \cdots & \cdots \\ b_1 & b_2 & \cdots & b_n \\ \cdots & \cdots & \cdots & \cdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix} + \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \cdots & \cdots & \cdots & \cdots \\ c_1 & c_2 & \cdots & c_n \\ \cdots & \cdots & \cdots & \cdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix}.$$

(5) 若方阵A有某两行成比例, 则$\det(A) = 0$. 特别地, 若A有两行相同, 则其行列式为零.

(6) 将方阵A的某一行的常数倍加到另一行, 得到的方阵B满足$\det(B) = \det(A)$, 即

$$\begin{vmatrix} \cdots & \cdots & \cdots & \cdots \\ a_{p1}+\lambda a_{q1} & a_{p2}+\lambda a_{q2} & \cdots & a_{pn}+\lambda a_{qn} \\ \cdots & \cdots & \cdots & \cdots \\ a_{q1} & a_{q2} & \cdots & a_{qn} \\ \cdots & \cdots & \cdots & \cdots \end{vmatrix} = \begin{vmatrix} \cdots & \cdots & \cdots & \cdots \\ a_{p1} & a_{p2} & \cdots & a_{pn} \\ \cdots & \cdots & \cdots & \cdots \\ a_{q1} & a_{q2} & \cdots & a_{qn} \\ \cdots & \cdots & \cdots & \cdots \end{vmatrix}.$$

证明. (1) 用数学归纳法. 当$n=1$时, 结论显然成立. 设结论对$n-1$阶行列式成立, 下证结论对n阶行列式成立.

设$A=(a_{ij})$是n阶方阵, $A'=(a'_{ij})$是A的转置方阵, 即$a'_{ij}=a_{ji}$. 分别用M_{ij}与M'_{ij}表示A与A'的(i,j)元素的余子式. 则由归纳假设, $M'_{ij}=M_{ji}$. 将$\det(A')$按第一列展开得

$$\det(A')=\sum_{i=1}^{n}(-1)^{i+1}a'_{i1}M'_{i1}=\sum_{i=1}^{n}(-1)^{i+1}a_{1i}M_{1i}=\det(A).$$

故结论成立.

(2) 用D_{ij}表示删除 A 的第 p,q 行与i,j列所得$n-2$阶矩阵的行列式. 将行列式A按p,q两行先后展开, 类似定理 3.1.1的证明可得

$$\det(A)=\sum_{i=1}^{n}\sum_{j=1}^{i-1}(-1)^{p+q+i+j-1}(a_{pi}a_{qj}-a_{pj}a_{qi})D_{ij}.$$

从上式看出, 如果A的p,q两行互换, 则行列式变号.

(3) 设将方阵A的第k行乘常数λ得到矩阵B. 将$\det(B)$按第k行展开得

$$\det(B)=\sum_{i=1}^{n}(\lambda a_{ki}A_{ki})=\lambda\sum_{i=1}^{n}a_{ki}A_{ki}=\lambda\det(A).$$

即知结论成立.

(4) 设方阵A的第k行被拆成两向量之和, 将$\det(A)$按第k行展开得

$$\det(A)=\sum_{i=1}^{n}(b_i+c_i)A_{ki}=\sum_{i=1}^{n}b_iA_{ki}+\sum_{i=1}^{n}c_iA_{ki}.$$

从而知结论成立.

(5) 假设矩阵A的p,q两行相等. 将A的p,q两互换后, 方阵不变, 但由性质(1), 行列式变号, 即$\det(A)=-\det(A)$. 因此, $\det(A)=0$. 如果A的p,q两行成比例, 则由性质(2), 将比例因子提出后, 行列式的p,q两行相等, 从而行列式仍为零.

(6) 根据性质(3), $\det(B)$可拆成两个行列式之和, 其中一个行列式是$\det(A)$; 另一个行列式的p,q两行成比例, 根据性质(4), 其为零. 因此, $\det(B)=\det(A)$. □

由上述定理的第一条性质知道, 行列式的行与列的地位是对等的. 因此, 定理中的有关行变换的性质对列变换也成立. 例如, 交换行列式某两列, 行列式变

号; 行列式某列乘常数λ, 等于λ乘原行列式; 行列式某列乘常数加到另一列, 行列式不变; 等等.

上述性质可以用行列式的几何意义来解释. 我们以2阶行列式来说明. 如前所述, 2阶行列式$\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}$表示以向量$\overrightarrow{OA}=(a_{11},a_{12})$, $\overrightarrow{OB}=(a_{21},a_{22})$为边构成的平行四边形面积. 性质(2) 表明, 向量$\overrightarrow{OA}$,$\overrightarrow{OB}$ 互换, 则角$\angle AOB$ 变号, 从而平行四边形面积S_{OABC}也变号. 性质(3)表明, 如果行列式某边长放缩λ 倍, 则平行四边形面积相应地放缩λ 倍. 性质(4)可以用图 3.3来解释, 其中平面四边形$ACA''C''$ 的面积等于平行四边形$ABA'B'$ 与平行四边形$BCB'C'$ 面积之和(实际上$S_{ABC}=S_{A'B'C'}$, $S_{ACA'C'}=S_{ACA''C''}$). 如果平行四边形两相邻边平行, 则平行四边形退化, 其面积为零, 这就是性质(5). 请读者自己解释性质(1).

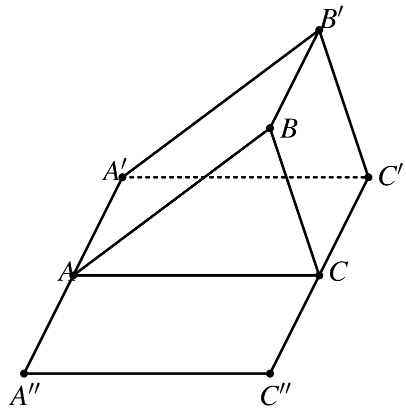


图 3.3 行列式性质4的几何解释.

例3.2.1. 证明:

$$\begin{vmatrix} a_{11} & 0 & \cdots & 0 \\ a_{21} & a_{22} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix} = a_{11}a_{22}\cdots a_{nn}.$$

证明: 由定理 3.2.1的第一条性质知, 行列式与它的转置行列式相等, 再由例3.1.3的结论立即可得. □

根据行列式的上述性质, 我们还可以从另一个角度理解行列式. 设$\mathbf{a}_i$是n阶方阵A的第i行构成的行向量, 则$\det(A)$可以看成是向量组$\mathbf{a}_1,\mathbf{a}_2,\cdots,\mathbf{a}_n$的函数, 记作$\det(\mathbf{a}_1,\mathbf{a}_2,\cdots,\mathbf{a}_n)$. 该函数具有以下性质

(1) **反对称性**: 互换两个变量的位置, 行列式变号, 即

$$\det(\mathbf{a}_1,\cdots,\mathbf{a}_i,\cdots,\mathbf{a}_j,\cdots,\mathbf{a}_n)=-\det(\mathbf{a}_1,\cdots,\mathbf{a}_j,\cdots,\mathbf{a}_i,\cdots,\mathbf{a}_n).$$

(2) **多重线性**: 行列式关于每行是线性的, 即

$$\det(\mathbf{a}_1,\ldots,\lambda\mathbf{a}_i+\mu\mathbf{b}_i,\ldots,\mathbf{a}_n)=\lambda\ \det(\mathbf{a}_1,\ldots,\mathbf{a}_i,\ldots)+\mu\ \det(\ldots,\mathbf{b}_i,\ldots,\mathbf{a}_n),$$

其中$\mathbf{a}_i,\mathbf{b}_i$是任意$n$维行向量, λ,μ是任意常数.

(3) **规范性**: 设$\mathbf{e}_1,\mathbf{e}_2,\ldots,\mathbf{e}_n$是单位坐标向量, 即$\mathbf{e}_i$为第$i$个分量为1, 其余分量为0的$n$维行向量. 则有

$$\det(\mathbf{e}_1,\mathbf{e}_2,\ldots,\mathbf{e}_n)=1.$$

实际上, 函数$\det(\mathbf{a}_1,\mathbf{a}_2,\cdots,\mathbf{a}_n)$由上述三条性质唯一确定, 一些教科书也将满足上述三条性质的函数$\det(\mathbf{a}_1,\mathbf{a}_2,\cdots,\mathbf{a}_n)$作为行列式的定义. 从下一节我们将会看到, 该定义与我们的递归定义是等价的.

例3.2.2. 设$\mathbf{a},\mathbf{b},\mathbf{c}$都是3维向量, 证明: $\det(\mathbf{a}+\mathbf{b},\mathbf{b}+\mathbf{c},\mathbf{c}+\mathbf{a})=2\det(\mathbf{a},\mathbf{b},\mathbf{c})$.

证明. 利用行列式的性质, 易得

$$\begin{aligned}&\det(\mathbf{a}+\mathbf{b},\mathbf{b}+\mathbf{c},\mathbf{c}+\mathbf{a})=\det(\mathbf{a},\mathbf{b}+\mathbf{c},\mathbf{c}+\mathbf{a})+\det(\mathbf{b},\mathbf{b}+\mathbf{c},\mathbf{c}+\mathbf{a})\\&=\det(\mathbf{a},\mathbf{b}+\mathbf{c},\mathbf{c})+\det(\mathbf{b},\mathbf{c},\mathbf{c}+\mathbf{a})=\det(\mathbf{a},\mathbf{b},\mathbf{c})+\det(\mathbf{b},\mathbf{c},\mathbf{a})=2\det(\mathbf{a},\mathbf{b},\mathbf{c}).\end{aligned}$$

□

§3.3 行列式的完全展开式

在本小中, 我们将给出行列式的完全展开式, 即将$\det(A)$展开为A的元素的多项式.

设$\mathbf{a}_i$是方阵$A=(a_{ij})_{n\times n}$的第i行构成的行向量, $\mathbf{e}_i$是第i个n维单位坐标向量, 则$\mathbf{a}_i$可以表示为$\mathbf{e}_1,\mathbf{e}_2,\ldots,\mathbf{e}_n$的线性组合, 即$\mathbf{a}_i=\sum\limits_{k=1}^{n}a_{ik}\mathbf{e}_k$. 根据行列式关于行的线

性性可得

$$\begin{aligned}\det(A) &= \det\left(\sum_{j_1=1}^{n} a_{1j_1}\mathbf{e}_{j_1}, \sum_{j_2=1}^{n} a_{2j_2}\mathbf{e}_{j_2}, \ldots, \sum_{j_n=1}^{n} a_{nj_n}\mathbf{e}_{j_n}\right)\\ &= \sum_{j_1=1}^{n} a_{1j_1}\det\left(\mathbf{e}_{j_1}, \sum_{j_2=1}^{n} a_{2j_2}\mathbf{e}_{j_2}, \ldots, \sum_{j_n=1}^{n} a_{nj_n}\mathbf{e}_{j_n}\right)\\ &= \sum_{j_1=1}^{n}\sum_{j_2=1}^{n} a_{1j_1}a_{2j_2}\det\left(\mathbf{e}_{j_1}, \mathbf{e}_{j_2}, \sum_{j_3=1}^{n} a_{3j_3}\mathbf{e}_{j_3}, \ldots, \sum_{j_n=1}^{n} a_{nj_n}\mathbf{e}_{j_n}\right)\\ &= \ldots = \sum_{j_1=1}^{n}\sum_{j_2=1}^{n}\cdots\sum_{j_n=1}^{n} a_{1j_1}a_{2j_2}\cdots a_{nj_n}\det(\mathbf{e}_{j_1}, \mathbf{e}_{j_2}, \ldots, \mathbf{e}_{j_n}).\end{aligned}$$

再由行列式的性质(5)可知, 当$j_1, j_2, \ldots, j_n$中有两个相等时, $\det(\mathbf{e}_{j_1}, \mathbf{e}_{j_2}, \ldots, \mathbf{e}_{j_n}) = 0$; 而当$j_1, j_2, \ldots, j_n$ 两两不等时, $(j_1, j_2, \ldots, j_n)$是$(1, 2, \ldots, n)$的一个排列. 因此,

$$\det(A) = \sum_{(j_1, j_2, \ldots, j_n)\in S_n} a_{1j_1}a_{2j_2}\cdots a_{nj_n}\det(\mathbf{e}_{j_1}, \mathbf{e}_{j_2}, \ldots, \mathbf{e}_{j_n}). \tag{3.13}$$

其中S_n是$(1, 2, \ldots, n)$的所有排列构成的集合. 为进一步展开上式, 我们引入逆序数的概念.

定义3.3.1. 将n个两两不同的正整数$s_1, s_2, \ldots, s_n$按顺序排成的一个有序数组称为一个**排列**, 记做$s = (s_1, s_2, \ldots, s_n)$. 满足$s_1 < s_2 < \ldots < s_n$的排列称为**顺序排列**. 特别地, 排列$(1, 2, \ldots, n)$称为**自然排列**. 互换$s$中$i, j$位置的两个数$s_i$与$s_j$称为一次**对换**. 满足$i < j$且$s_i > s_j$的一对数$(s_i, s_j)$称为$s$的一个**逆序**. s的逆序的个数称为s的**逆序数**, 记作$\tau(s)$. 逆序数为奇数的排列称为**奇排列**; 逆序数为偶数的排列称为**偶排列**.

例如, $(3, 1, 4, 2), (1, 2, 3, 4)$都是1, 2, 3, 4的排列, 其中$(1, 2, 3, 4)$是自然排列, 其逆序数为零, 因此是偶排列. 排列$(3, 1, 4, 2)$ 有逆序$(3, 1), (3, 2), (4, 2)$, 其逆序数$\tau(3, 1, 4, 2) = 3$, 因此是奇排列.

任何一个排列都可以通过对换将其变成一个顺序排列. 例如, $(3, 1, 4, 2) \to (1, 3, 4, 2) \to (1, 3, 2, 4) \to (1, 2, 3, 4)$, 即排列$(3, 1, 4, 2)$ 经过3次对换变成了顺序排列. 实际上, 我们有

引理3.3.1. *每个排列s可经过$\tau(s)$次相邻位置的对换变成从小到大的顺序排列.*

证明. 设s是正整数列$a_1 < a_2 < \ldots < a_n$的一个排列, s中形如$(*, a_i)$的逆序共有m_i 个, $i = 1, 2, \ldots, n$. 则可经过m_1次相邻位置间的对换, 将a_1 置换到位置1; 再经过m_2次相邻位置间的对换, 将a_2置换到位置2; 以此类推, 可经过总共$m_1 + m_2 + \ldots + m_{n-1} = \tau(s)$次相邻位置的对换, 把$s$ 变成顺序排列$(a_1, a_2, \cdots, a_n)$. □

有了排列的逆序数的概念, 下面我们就可以给出行列式的完全展开式.

定理3.3.1. 设$A=(a_{ij})_{n\times n}$为n阶方阵, 则

$$\det(A)=\sum_{(j_1,j_2,\cdots,j_n)\in S_n}(-1)^{\tau(j_1,j_2,\ldots,j_n)}a_{1j_1}a_{2j_2}\cdots a_{nj_n}. \tag{3.14}$$

证明. 根据引理3.3.1, 可经过$\tau(j_1,j_2,\ldots,j_n)$次对换, 将$(j_1,j_2,\ldots,j_n)$变成标准排列$(1,2,\cdots,n)$. 又根据行列式的性质(1), 每次对换都会使行列式变号. 因此,

$$\det(\mathbf{e}_{j_1},\mathbf{e}_{j_2},\ldots,\mathbf{e}_{j_n})=(-1)^{\tau(j_1,j_2,\ldots,j_n)}.$$

代入(3.13)得

$$\det(A)=\sum_{(j_1,j_2,\ldots,j_n)\in S_n}(-1)^{\tau(j_1,j_2,\ldots,j_n)}a_{1j_1}a_{2j_2}\cdots a_{nj_n}. \qquad \square$$

上式就是行列式的完全展开式, 许多教科书也直接将 (3.14)式作为行列式的定义.

特别地, 我们再次得到二阶与三阶行列式的完全展开式

$$\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}=a_{11}a_{22}-a_{21}a_{12},$$

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

例3.3.1. 证明:

$$\begin{vmatrix} a_{11} & \cdots & a_{1r} & a_{1,r+1} & \cdots & a_{1n} \\ \vdots & \cdots & \vdots & \vdots & \cdots & \vdots \\ a_{r1} & \cdots & a_{rr} & a_{r,r+1} & \cdots & a_{rn} \\ 0 & \cdots & 0 & a_{r+1,r+1} & \cdots & a_{r+1,n} \\ \vdots & \cdots & \vdots & \vdots & \cdots & \vdots \\ 0 & \cdots & 0 & a_{n,r+1} & \cdots & a_{nn} \end{vmatrix}=\begin{vmatrix} a_{11} & \cdots & a_{1r} \\ \vdots & \cdots & \vdots \\ a_{r1} & \cdots & a_{rr} \end{vmatrix}\cdot\begin{vmatrix} a_{r+1,r+1} & \cdots & a_{r+1,n} \\ \vdots & \cdots & \vdots \\ a_{n,r+1} & \cdots & a_{nn} \end{vmatrix}.$$

证明. 用Δ记左边的行列式. 考察Δ的完全展开式(3.14)中的非零项. 由于当$r+1\le i\le n$且$1\le j\le r$时, $a_{ij}=0$, 故若$a_{1j_1}a_{2j_2}\cdots a_{nj_n}\ne 0$, 则必然有$r+1\le j_{r+1},\ldots,j_n\le n$. 从而$(j_{r+1},\ldots,j_n)$ 是$(r+1,\ldots,n)$的一个排列, 进而$(j_1,\ldots,j_r)$也是$(1,\ldots,r)$的一个排列. 而且容易验证, 逆序数$\tau(j_1,\ldots,j_n)=\tau(j_1,\ldots,j_r)+\tau(j_{r+1},\cdots,j_n)$. 记$S_1$是$1,2,\ldots,r$的排列的集合, S_2是$r+1,r+2,\ldots,n$的排列的集合. 则

$$\begin{aligned}\Delta &= \sum_{(j_1,j_2,\ldots,j_n)\in S_n}(-1)^{\tau(j_1,j_2,\ldots,j_n)}a_{1j_1}a_{2j_2}\cdots a_{nj_n}\\ &= \sum_{\substack{(j_1,\cdots,j_r)\in S_1\\(j_{r+1},\cdots,j_n)\in S_2}}(-1)^{\tau(j_1\cdots j_r)+\tau(j_{r+1}\cdots j_n)}a_{1j_1}\cdots a_{rj_r}a_{r+1,j_{r+1}}\cdots a_{nj_n}\\ &= \sum_{(j_1,\cdots,j_r)\in S_1}(-1)^{\tau(j_1\cdots j_r)}a_{1j_1}\cdots a_{rj_r}\cdot\sum_{(j_{r+1},\cdots,j_n)\in S_2}(-1)^{\tau(j_{r+1}\cdots j_n)}a_{r+1,j_{r+1}}\cdots a_{nj_n}\\ &= \begin{vmatrix}a_{11}&\cdots&a_{1r}\\ \vdots&\cdots&\vdots\\ a_{r1}&\cdots&a_{rr}\end{vmatrix}\cdot\begin{vmatrix}a_{r+1,r+1}&\cdots&a_{r+1,n}\\ \vdots&\cdots&\vdots\\ a_{n,r+1}&\cdots&a_{nn}\end{vmatrix}.\end{aligned}$$

□

§3.4 Cramer法则

在第一章中, 我们提出了求线性方程组的公式解的问题. 在本章的第0节, 我们给给出了二元一次方程组及三元一次方程组的公式解. 本节中, 我们将给出求解具有n 个方程和n 个未知数的线性方程组的公式解的法则–Cramer 法则. 为此, 我们先给出一个有用的结论.

定理3.4.1. 设$A=(a_{ij})$为n阶方阵, 则

$$\sum_{k=1}^{n}a_{ik}A_{jk}=\det(A)\delta_{ij},\quad \sum_{k=1}^{n}a_{ki}A_{kj}=\det(A)\delta_{ij},\qquad i,j=1,2,\ldots,n, \tag{3.15}$$

其中δ_{ij}为*Kronecker*记号

$$\delta_{ij}=\begin{cases}1, & i=j\\ 0, & i\ne j\end{cases}$$

证明. 我们只证 $\sum_{k=1}^{n}a_{ik}A_{jk}=\det(A)\delta_{ij}$, 也就是

$$\sum_{k=1}^{n}a_{ik}A_{ik}=\det(A),\qquad \sum_{k=1}^{n}a_{ik}A_{jk}=0,\, i\ne j.$$

另一式同理可得.

$\det(A)$按第i行展开显然有$\sum_{k=1}^n a_{ik}A_{ik} = \det(A)$. 下设$i \neq j$, 记矩阵$B$为将$A$的第$j$行用第$i$行替换所得矩阵. 由于$B$的第$i,j$两行元素相同, 由行列式的性质(4), $\det(B)=0$. 另一方面, 将行列式$\det(B)$按第j行展开得

$$\det(B) = \sum_{k=1}^n a_{ik}A_{jk} = 0.$$

定理证毕. □

我们以$n=3$为例解释一下定理 3.4.1的结论. 由 (3.15)式得

$$a_{11}A_{11}+a_{12}A_{12}+a_{13}A_{13}=\det(A), \quad a_{21}A_{11}+a_{22}A_{12}+a_{23}A_{13}=0, \quad a_{31}A_{11}+a_{32}A_{12}+a_{33}A_{13}=0.$$

分别用$\mathbf{a}_1,\mathbf{a}_2,\mathbf{a}_3$表示三阶行列式$\det(A)$的三个行向量. 则$(A_{11},A_{12},A_{13}) = \mathbf{a}_2\times\mathbf{a}_3$, 于是上面几个式子等价于

$$\det(A)=\mathbf{a}_1\cdot\mathbf{a}_2\times\mathbf{a}_3, \quad \mathbf{a}_2\cdot\mathbf{a}_2\times\mathbf{a}_3=0, \quad \mathbf{a}_3\cdot\mathbf{a}_2\times\mathbf{a}_3=0.$$

利用上述定理的结论, 我们可以方便地证明Cramer法则.

定理3.4.2 (Cramer法则). 当系数矩阵$A=(a_{ij})_{n\times n}$的行列式不等于零时, 线性方程组

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

有唯一解

$$(x_1,x_2,\ldots,x_n)=\left(\frac{\Delta_1}{\Delta},\frac{\Delta_2}{\Delta},\ldots,\frac{\Delta_n}{\Delta}\right),$$

其中$\Delta=\det(A)$, Δ_i是将A的第i列换成常数项后所得方阵的行列式, $i=1,2,\ldots,n$.

证. 将线性方程组的每个方程分别乘$A_{1i},A_{2i},\ldots,A_{ni}$, 然后求和得

$$\sum_{j=1}^n A_{ji}\sum_{k=1}^n a_{jk}x_k = \sum_{j=1}^n A_{ji}b_j.$$

上式左边为

$$\sum_{j=1}^n A_{ji}\sum_{k=1}^n a_{jk}x_k = \sum_{k=1}^n\left(\sum_{j=1}^n a_{jk}A_{ji}\right)x_k = \sum_{k=1}^n \delta_{ki}\det(A)x_k = \Delta\cdot x_i.$$

而行列式Δ_i按第i列展开得到$\Delta_i=\sum_{j=1}^n A_{ji}b_j$. 因此当$\Delta\neq 0$时, $x_i=\Delta_i/\Delta$, $i=1,2,\ldots,n$.
定理证毕. □

我们以$n=3$来理解一下上面证明的思路. 当$n=3$时, 原方程组可以写成向量形式

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

其中

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

对方程两边点乘$\mathbf{a}_2\times\mathbf{a}_3$可得

$$\mathbf{a}_1\cdot\mathbf{a}_2\times\mathbf{a}_3x_1=\mathbf{b}\cdot\mathbf{a}_2\times\mathbf{a}_3.$$

从而

$$x_1=\frac{\mathbf{b}\cdot\mathbf{a}_2\times\mathbf{a}_3}{\mathbf{a}_1\cdot\mathbf{a}_2\times\mathbf{a}_3}=\frac{\Delta_1}{\Delta}.$$

同理, 可以得到x_2,x_3的表示式.

例3.4.1. 用Cramer法则解线性方程组

$$\begin{cases}\quad\ \ x_2+x_3+x_4=1\\ x_1\quad\ \ +x_3+x_4=2\\ x_1+x_2\quad\ \ +x_4=3\\ x_1+x_2+x_3\quad\ \ =4\end{cases}$$

解.

$$\Delta=\begin{vmatrix}0&1&1&1\\1&0&1&1\\1&1&0&1\\1&1&1&0\end{vmatrix}=-3,\quad \Delta_1=\begin{vmatrix}1&1&1&1\\2&0&1&1\\3&1&0&1\\4&1&1&0\end{vmatrix}=-7,\quad \Delta_2=\begin{vmatrix}0&1&1&1\\1&2&1&1\\1&3&0&1\\1&4&1&0\end{vmatrix}=-4,$$

$$\Delta_3=\begin{vmatrix}0&1&1&1\\1&0&2&1\\1&1&3&1\\1&1&4&0\end{vmatrix}=-1,\quad \Delta_4=\begin{vmatrix}0&1&1&1\\1&0&1&2\\1&1&0&3\\1&1&1&4\end{vmatrix}=2.$$

故方程组的解为

$$x_1=\frac{7}{3},\ x_2=\frac{4}{3},\ x_3=\frac{1}{3},\ x_4=-\frac{2}{3}.$$

□

一般来说, Cramer法则主要用于理论分析, 除非相关的行列式容易计算, 否则用Cramer法则求解方程的运算量非常巨大.

§3.5 行列式的计算

计算行列式的最基本的方法是: 通过行与列的初等变换, 化一般方阵的行列式为三角方阵的行列式. 根据行列式的性质,

- 交换方阵的两行(列)之后, 行列式变为原来的相反数.
- 将某行(列)乘以非零常数λ之后, 行列式变为原来的λ倍.
- 将某行(列)的常数倍加到另一行(列)之后, 行列式保持不变.

为方便起见, 我们采用以下记号来表示行列式的初等变换. 分别用$r_i \leftrightarrow r_j$、λr_i、$\lambda r_i \to r_j$ 表示第i,j两行互换、第i行乘常数λ、第i行乘λ加到第j行. 类似地, 分别用记号$c_i \leftrightarrow c_j$、λc_i、$\lambda c_i \to c_j$表示三类列变换. 例如, 把第2行的3倍加到第3行, 记为$3r_2 \to r_3$; 将第3列乘-2记为$-2c_3$.

例3.5.1. 计算以下行列式

$$\begin{vmatrix} 0 & 5 & -4 & 5 \\ -3 & -1 & -5 & 3 \\ 3 & 1 & -2 & -3 \\ -1 & 4 & -5 & -1 \end{vmatrix}.$$

解.

$$\begin{vmatrix} 0 & 5 & -4 & 5 \\ -3 & -1 & -5 & 3 \\ 3 & 1 & -2 & -3 \\ -1 & 4 & -5 & -1 \end{vmatrix} \xrightarrow{r_3 \to r_2} \begin{vmatrix} 0 & 5 & -4 & 5 \\ 0 & 0 & -7 & 0 \\ 3 & 1 & -2 & -3 \\ -1 & 4 & -5 & -1 \end{vmatrix} \xrightarrow{3r_4 \to r_3} \begin{vmatrix} 0 & 5 & -4 & 5 \\ 0 & 0 & -7 & 0 \\ 0 & 13 & -17 & -6 \\ -1 & 4 & -5 & -1 \end{vmatrix} \xrightarrow{-c_4 \to c_2}$$

$$\begin{vmatrix} 0 & 0 & -4 & 5 \\ 0 & 0 & -7 & 0 \\ 0 & 19 & -17 & -6 \\ -1 & 5 & -5 & -1 \end{vmatrix} \xrightarrow{r_1 \leftrightarrow r_4, r_2 \leftrightarrow r_3} \begin{vmatrix} -1 & 5 & -5 & -1 \\ 0 & 19 & -17 & -6 \\ 0 & 0 & -7 & 0 \\ 0 & 0 & -4 & 5 \end{vmatrix} = -19 \cdot \begin{vmatrix} -7 & 0 \\ -4 & 5 \end{vmatrix} = 665. \qquad \square$$

例3.5.2. 计算n阶行列式

$$\begin{vmatrix} x & 1 & \cdots & 1 \\ 1 & x & \ddots & \vdots \\ \vdots & \ddots & \ddots & 1 \\ 1 & \cdots & 1 & x \end{vmatrix}.$$

解.

$$\begin{vmatrix} x & 1 & \cdots & 1 \\ 1 & x & \ddots & \vdots \\ \vdots & \ddots & \ddots & 1 \\ 1 & \cdots & 1 & x \end{vmatrix} \xrightarrow[\text{其它各行}]{\text{第1行加上}} \begin{vmatrix} x+n-1 & x+n-1 & \cdots & x+n-1 \\ 1 & x & \ddots & \vdots \\ \vdots & \ddots & \ddots & 1 \\ 1 & \cdots & 1 & x \end{vmatrix} \xrightarrow[\text{提取因子}]{\text{第1行}}$$

$$(x+n-1)\begin{vmatrix} 1 & 1 & \cdots & 1 \\ 1 & x & \ddots & \vdots \\ \vdots & \ddots & \ddots & 1 \\ 1 & \cdots & 1 & x \end{vmatrix} \xrightarrow[\text{减去第1行}]{\text{其它各行}} (x+n-1)\begin{vmatrix} 1 & 1 & \cdots & 1 \\ & x-1 & & \\ & & \ddots & \\ & & & x-1 \end{vmatrix}$$

$$= (x+n-1)(x-1)^{n-1}.$$ □

对于一些结构性很高的行列式, 可以将行列式按某一行(列)展开, 建立行列式的递推公式求解.

例3.5.3. 计算Vandermonde行列式

$$\Delta_n(a_1,a_2,\ldots,a_n) = \begin{vmatrix} 1 & a_1 & \cdots & a_1^{n-1} \\ 1 & a_2 & \cdots & a_2^{n-1} \\ \vdots & \vdots & \cdots & \vdots \\ 1 & a_n & \cdots & a_n^{n-1} \end{vmatrix}.$$

解. 从行列式的第n列开始至第2列, 依次把各列减去前一列的a_n倍, 得

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

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

$$= \prod_{i=1}^{n-1}(a_n-a_i)\cdot\Delta_{n-1}(a_1,a_2,\ldots,a_{n-1}).$$

易知$\Delta_1 = 1, \Delta_2(a_1,a_2) = a_2 - a_1$. 利用数学归纳法可得

$$\Delta_n(a_1,a_2,\ldots,a_n) = \prod_{1\le i<j\le n}(a_j-a_i).$$

□

例3.5.4. 计算n阶行列式

$$\Delta_n = \begin{vmatrix} 2 & 1 & 0 & \cdots & 0 \\ 1 & 2 & 1 & \ddots & \vdots \\ 0 & 1 & \ddots & \ddots & 0 \\ \vdots & \ddots & \ddots & 2 & 1 \\ 0 & \cdots & 0 & 1 & 2 \end{vmatrix}$$

.

解. 容易算得$\Delta_1 = 2, \Delta_2 = 3$. 当$n \geq 3$时, 按第一行展开行列式得

$$\Delta_n = 2\Delta_{n-1} - \Delta_{n-2}.$$

由此

$$\Delta_n - \Delta_{n-1} = \Delta_{n-1} - \Delta_{n-2} = \cdots = \Delta_2 - \Delta_1 = 1.$$

从而

$$\Delta_n = n + 1.$$

□

小知识:行列式

行列式的概念作为一个速记工具最早出现于线性方程组的求解, 它是由日本数学家关孝和及德国数学家Leibnitz于十七世纪末期提出, 现已称为线性代数的基本工具之一.

1683年关孝和写了一部专著《解伏题之法》, 意思就是《解行列式问题的方法》. 1693年Leibnitz在写给洛必达的一封信中使用了行列式记号, 但没有提出行列式的概念. 1729年英国数学家Maclaurin以行列式为工具解含有2, 3, 4个未知量的线性方程组. 1750 年, 瑞士数学家Cramer更完整地叙述了行列式的展开法则并将它用于解线性方程组. 1772年法国数学家Vandermonde专门对行列式做了理论上的研究, 建立了行列式展开法则, 用子式和代数余子式表示一个行列式. 1772年法国数学家Laplace推广了Vandermonde展开行列式的方法, 得到我们熟知的拉普拉斯展开定理. 1813至1815年法国数学家Cauchy对行列式做了系统的代数处理. 1841年, 德国数学家Jacobi发表的《论行列式的形成与性质》一文, 总结了行列式的理论.

行列式在坐标变换、多重积分中的变量替换、二次型及多元多项式消元等有重要的应用.

习题二

1. 计算下列行列式:

$$(1)\quad \begin{vmatrix} 1 & 0 & 1 & -4 \\ -1 & -3 & 0 & -2 \\ 2 & -1 & 4 & 0 \\ 0 & 3 & -3 & 2 \end{vmatrix} \qquad (2)\quad \begin{vmatrix} 1 & 4 & -1 & -1 \\ 1 & 0 & 0 & 1 \\ -3 & 3 & -4 & 0 \\ 0 & 1 & -1 & -1 \end{vmatrix}$$

$$(3)\quad \begin{vmatrix} x+a & x+b & x+c \\ y+a & y+b & y+c \\ z+a & z+b & z+c \end{vmatrix} \qquad (4)\quad \begin{vmatrix} & & & a_{1n} \\ & & a_{2,n-1} & a_{2n} \\ & \iddots & \iddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix}$$

$$(5)\quad \begin{vmatrix} a^2 & (a+1)^2 & (a+2)^2 \\ b^2 & (b+1)^2 & (b+2)^2 \\ c^2 & (c+1)^2 & (c+2)^2 \end{vmatrix} \qquad (6)\quad \begin{vmatrix} a_1 & a_2 & 0 & 0 & 0 \\ b_1 & b_2 & 0 & 0 & 0 \\ c_1 & c_2 & c_3 & 0 & 0 \\ d_1 & d_2 & d_3 & d_4 & 0 \\ e_1 & e_2 & e_3 & e_4 & e_5 \end{vmatrix}$$

2. 在三维空间直角坐标系中, 已知点A,B,C,D的坐标分别是 $(1,1,0)$, $(3,1,2)$, $(0,1,3)$, $(2,2,4)$. 求四面体$ABCD$ 的体积及各个面的面积.

3. 将行列式

$$\begin{vmatrix} x-2 & 1 & 0 & 0 \\ 1 & x-2 & 1 & 0 \\ 0 & 1 & x-2 & 1 \\ 0 & 0 & 1 & x-2 \end{vmatrix}$$

展开为关于x的多项式.

4. A为n阶方阵, λ为常数. 证明: $\det(\lambda A) = \lambda^n \det(A)$.

5. 方阵A称为**反对称方阵**, 如果它的转置方阵等于$-A$. 证明: 奇数阶反对称方阵的行列式为零.

6. 证明:

$$\begin{vmatrix} a_{11}b_{11}+a_{12}b_{21} & a_{11}b_{12}+a_{12}b_{22} \\ a_{21}b_{11}+a_{22}b_{21} & a_{21}b_{12}+a_{22}b_{22} \end{vmatrix} = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} \cdot \begin{vmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{vmatrix}$$

7. 证明:

$$\begin{vmatrix} a_{11} & a_{12} & 1 & 0 \\ a_{21} & a_{22} & 0 & 1 \\ 1 & 0 & b_{11} & b_{12} \\ 0 & 1 & b_{21} & b_{22} \end{vmatrix} = \begin{vmatrix} a_{11}b_{11}+a_{12}b_{21}-1 & a_{11}b_{12}+a_{12}b_{22} \\ a_{21}b_{11}+a_{22}b_{21} & a_{21}b_{12}+a_{22}b_{22}-1 \end{vmatrix}$$

8. 设$\mathbf{a},\mathbf{b},\mathbf{c},\mathbf{d}$为4维数组向量. 证明: $\det(2\mathbf{a}-\mathbf{b},-\mathbf{a}+2\mathbf{b}-\mathbf{c},-\mathbf{b}+2\mathbf{c}-\mathbf{d},-\mathbf{c}+2\mathbf{d}) = 5\det(\mathbf{a},\mathbf{b},\mathbf{c},\mathbf{d})$.

9. 求以下排列的逆序数, 并指出其奇偶性.
(1) $(6,8,1,4,7,5,3,2,9)$ (2) $(6,4,2,1,9,7,3,5,8)$ (3) $(7,5,2,3,9,8,1,6,4)$

10. 证明: 任意一个排列经过一次对换后, 必改变其奇偶性.

11. 证明: 等式 (3.14)作为行列式的定义与行列式其它定义等价.

12. 写出四阶行列式的完全展开式.

13. 用Cramer法则求解下列线性方程组:

(1) $\begin{cases} x_1-x_2+x_3=3 \\ x_1+2x_2+4x_3=5 \\ x_1+3x_2+9x_3=7 \end{cases}$ (2) $\begin{cases} 2x_1+x_2-5x_3+x_4=8 \\ x_1-3x_2-6x_4=9 \\ 2x_2-x_3+2x_4=-5 \\ x_1+4x_2-7x_3+6x_4=0 \end{cases}$

14. 设$x_0,x_1,\dots,x_n$及$y_0,y_1,\dots,y_n$是任给实数, 其中$x_i(0\le i\le n)$两两互不相等. 证明: 存在唯一的次数不超过n 的多项式$p(x)$ 满足$p(x_i)=y_i$, $i=0,1,\dots,n$.

15. 证明: n个变元、n个方程的线性方程组有非零解当且仅当系数矩阵的行列式为零.

16. 计算下列n阶行列式

(1) $$\begin{vmatrix} a_1 & & & & & b_1 \\ & \ddots & & & \iddots & \\ & & a_n & b_n & & \\ & & c_n & d_n & & \\ & \iddots & & & \ddots & \\ c_1 & & & & & d_1 \end{vmatrix}$$

(2) $$\begin{vmatrix} 1+a_1 & 1 & \cdots & 1 \\ 1 & 1+a_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 1 \\ 1 & \cdots & 1 & 1+a_n \end{vmatrix}$$

(3) $$\begin{vmatrix} 2\cos(\theta) & 1 & 0 & \cdots & 0 \\ 1 & 2\cos(\theta) & 1 & \cdots & 0 \\ 0 & 1 & 2\cos(\theta) & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \cdots & 2\cos(\theta) \end{vmatrix}$$

第四章 矩阵

在第一章介绍求解线性方程组的Gauss消元法时, 我们引进了矩阵的概念. 但在那里, 矩阵仅仅是方程组的一个简略记号, 并没有定义矩阵之间的运算. 实际上, 通过引入矩阵之间的运算(特别是乘法), 我们可以将求解线性方程组的Gauss消元法用矩阵运算来表述, 从而理解Gauss消元法的本质. 经过几个世纪的发展, 矩阵现已称为线性代数的主要研究工具和基本研究对象之一. 科学与工程计算中的许多问题都可以转化成矩阵的计算问题. 本章将介绍矩阵的基本运算, 包括线性运算、乘法、求逆、分块运算、初等矩阵及其应用.

§4.1 矩阵的定义

矩阵的概念在现实生活中有着广泛的应用. 例如, 我们经常见到的各种统计表格就是矩阵的具体体现. 第一章中我们用矩阵来表示线性方程组. 后面我们将看到, 矩阵还可以表示更多的数学对象. 虽然第一章引入了矩阵的记号, 但是这里我们还是给出矩阵的正式定义.

定义4.1.1. 对任意正整数m和n, 由$m\times n$个数排成的m行n列的矩形列表

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \cdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} \tag{4.1}$$

称为一个$m\times n$阶**矩阵**, 记作$A=(a_{ij})_{m\times n}$. 称表中的每个数为矩阵A的**元素**; 排在第i行第j列的元素a_{ij}被称为A的**第**(i,j)**元素**; 当$i=j$时, a_{ii}被称为A的**对角元**. 两个矩阵A与B称为是**相等**的, 如果它们的行数和列数都相等并且每个位置上的元素也相等, 记作$A=B$.

矩阵的概念可以看成数组向量的一个自然推广. 两者之间既略有区别, 同时也有紧密的联系. 一个n 维数组行向量可以看成一个1行n列的矩阵; 同理, 一个n维数组列向量可视为一个n行1列的矩阵. 在向量的意义下, 行向量

$$(a_1,a_2,\cdots,a_n)$$

和列向量

$$\begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix}$$

表示同一个向量; 但是在矩阵的意义下, 它们表示不同的矩阵.

一个$m\times n$阶矩阵A 既可以被看作是m个行向量按列排在一起, 又可以被看作n个列向量按行排在一起, 还可以被看作是将一个mn维的数组向量. 除此之外, 矩阵还可以表示许多数学对象. 下面, 我们给出几个例子.

例4.1.1. 旋转变换矩阵 在平面直角坐标系Oxy中, 将点$P=(x,y)$绕原点逆时针旋转θ角得到点$P'=(x',y')$, 则有

$$\begin{cases} x'=\cos(\theta)x-\sin(\theta)y \\ y'=\sin(\theta)x+\cos(\theta)y \end{cases}$$

该坐标变换公式的系数排成的2×2阶矩阵

$$P=\begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$$

表示平面上绕原点逆时针旋转θ角的旋转变换. 类似地, 3×3阶矩阵

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

表示空间中点绕z轴正向逆时针旋转θ角的坐标变换公式.

例4.1.2. 线性映射 设$\mathbf{x}=(x_1,x_2,\cdots,x_m)$是一个$m$维数组向量, 其中每个$x_i$都可以表示为$n$维数组向量$\mathbf{y}=(y_1,y_2,\cdots,y_n)$的分量的齐次线性函数:

$$\begin{cases} x_1=a_{11}y_1+a_{12}y_2+\cdots+a_{1n}y_n, \\ x_2=a_{21}y_1+a_{22}y_2+\cdots+a_{2n}y_n, \\ \qquad\qquad\cdots\cdots\cdots \\ x_m=a_{m1}y_1+a_{m2}y_2+\cdots+a_{mn}y_n. \end{cases}$$

由上式确定的映射$\mathscr{A}:\mathbf{y}\mapsto\mathbf{x}$称为**线性映射**. 该线性映射可用矩阵$A=(a_{ij})_{m\times n}$来表示. 线性映射$\mathscr{A}$与矩阵$A$ 建立了一一对应关系, 即每个矩阵A对应了一个线性映射$\mathscr{A}$; 反过来, 每个线性映射$\mathscr{A}$都对应一个矩阵A. 显然, 例 4.1.1中的旋转变换是线性映射的一个特例.

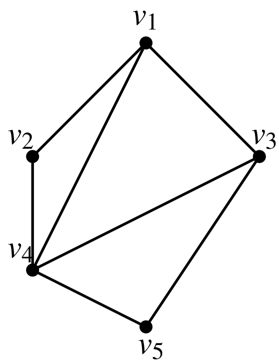


图 4.1 图的邻接矩阵

例4.1.3. 图的邻接矩阵 一个平面图G是由顶点与边构成. 每条边连接两个顶点. 设图G有n个顶点$v_1, v_2, \ldots, v_n$. 现构造一个矩阵$A = (a_{ij})_{n\times n}$, 如果顶点v_i与v_j有一条边相连, 则定义$a_{ij} = 1$; 否则, 定义$a_{ij} = 0$. 称矩阵A为图G的**邻接矩阵**. 例如, 图 4.1 的邻接矩阵

$$A = \begin{pmatrix} 0 & 1 & 1 & 1 & 0 \\ 1 & 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 & 1 \\ 1 & 1 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 & 0 \end{pmatrix}.$$

图的邻接矩阵完全反映了图的拓扑关系.

例4.1.4. 社团关系 某校有n名学生加入了m个社团. 引进一个$m\times n$阶矩阵$A = (a_{ij})$, 其中

$$a_{ij} = \begin{cases} 1 & \text{如果第}j\text{号学生属于第}i\text{个社团} \\ 0 & \text{其它} \end{cases}$$

则矩阵A描述了学生与社团之间的关系. 例如,

$$A = \begin{pmatrix} 1 & 1 & 1 & 1 & 0 \\ 1 & 0 & 1 & 0 & 1 \\ 0 & 1 & 1 & 0 & 1 \end{pmatrix}$$

表示有三个社团, 5名学生. 第一个社团有编号位1,2,3,4的学生, 第二个社团有编号位1,3,5的学生, 第三个社团有编号位2,3,5的学生.

例4.1.5. 数字图像 一幅灰度数字图像通过分割成$m\times n$的均匀小方格, 可以看成一个$m\times n$阶矩阵$A = (a_{ij})$. 每一个小方格称为一个 **像素**. 因此, 一幅灰度图像总

共有mn个像素. 每个像素对应的**灰度值**由矩阵A的元素$a_{ij}(0 \le a_{ij} \le 1)$ 给出. 例如, $a_{ij}=0$ 表示像素(i,j) 为黑色; $a_{ij}=1$表示像素(i,j) 为白色; 一般地, a_{ij}越小表示对应像素的颜色越暗; 反之, a_{ij} 越大表示对应像素的颜色越明亮.

一幅数字图像可以看成一幅真实图像的离散表示. m,n越大(即像素越多), 则逼近效果越好. 图 4.2显示了一幅256×256的灰度图像.

图 4.2 数字图像

下面介绍几种常见的矩阵名称及其记号.

- 元素都是0的矩阵称为**零矩阵**, 记作O或0.
- $n\times n$矩阵称为n阶**方阵**.
- 对角元是1其它元素都是0的n阶方阵称为**单位阵**, 记作I, I_n或$I^{(n)}$.
- 对角元是a其它元素都是0的方阵称为**数量阵**, 记作aI.
- 若方阵$A=(a_{ij})_{n\times n}$的元素除对角元外都为零, 则称A为**对角阵**, 记作

$$A=\begin{pmatrix} a_{11} & & \\ & \ddots & \\ & & a_{nn} \end{pmatrix} \quad 或 \quad A=\mathrm{diag}(a_{11},a_{22},\cdots,a_{nn}).$$

- 若矩阵$A=(a_{ij})_{m\times n}$满足$a_{ij}=0$对所有$i>j$成立, 则A称为**上三角阵**.
- 若矩阵$A=(a_{ij})_{m\times n}$满足$a_{ij}=0$对所有$i<j$成立, 则A称为**下三角阵**.
- 上三角阵和下三角阵统称为**三角阵**.
- 若方阵$A=(a_{ij})_{n\times n}$满足$a_{ij}=a_{ji}$对所有i,j成立, 则A称为**对称阵**.

- 若方阵$A=(a_{ij})_{n\times n}$满足$a_{ij}=-a_{ji}$对所有i,j成立, 则A称为**反对称阵**.
- 若矩阵A的元素都是整数、有理数、实数、复数、多项式等, 则A分别称为整数矩阵、有理数矩阵、实矩阵、复矩阵、多项式矩阵. 一般地, 若A的元素都取自某个数域F, 则A称为**数域F上的矩阵**.数域F上的$m\times n$矩阵的全体, 记作$F^{m\times n}$.

例4.1.6. 在下列2阶方阵

$$A=\begin{pmatrix}0&1\\1&0\end{pmatrix},B=\begin{pmatrix}1&0\\\frac{1}{2}&1\end{pmatrix},C=\begin{pmatrix}0&\pi\\-\pi&0\end{pmatrix},D=\begin{pmatrix}x&0\\0&x^2\end{pmatrix},E=\begin{pmatrix}e^{i\pi/3}&0\\0&e^{i\pi/3}\end{pmatrix}$$

中, A为整数方阵、对称阵, B为有理方阵、下三角阵, C为实方阵、反对称阵, D为多项式方阵、对角阵, E为复方阵、数量阵.

§4.2 矩阵的运算

在本节中, 我们介绍矩阵的最基本运算, 包括加法、数乘、乘积、求逆、转置、分块等. 其中最重要的矩阵乘法与逆. 初学者只有通过加深对基本概念的理解、提高矩阵运算的熟练程度, 才能逐渐地掌握矩阵运算技巧、并灵活地加以运用.

§4.2.1 加法与数乘

矩阵的加法与数乘又称为线性运算. 矩阵的线性运算可以视为数组向量的线性运算的简单推广.

定义4.2.1. 设矩阵$A=(a_{ij})_{m\times n}\in F^{m\times n}$, $B_{m\times n}=(b_{ij})\in F^{m\times n}$, $\lambda\in F$. 定义

$$A+B:=\begin{pmatrix}a_{11}+b_{11}&a_{12}+b_{12}&\cdots&a_{1n}+b_{1n}\\a_{21}+b_{21}&a_{22}+b_{22}&\cdots&a_{2n}+b_{2n}\\\vdots&\vdots&\cdots&\vdots\\a_{m1}+b_{m1}&a_{m2}+b_{m2}&\cdots&a_{mn}+b_{mn}\end{pmatrix}\tag{4.2}$$

和

$$\lambda A:=\begin{pmatrix}\lambda a_{11}&\lambda a_{12}&\cdots&\lambda a_{1n}\\\lambda a_{21}&\lambda a_{22}&\cdots&\lambda a_{2n}\\\vdots&\vdots&\cdots&\vdots\\\lambda a_{m1}&\lambda a_{m2}&\cdots&\lambda a_{mn}\end{pmatrix}\tag{4.3}$$

分别称为矩阵的**加法**运算和**数乘**运算, 记作

$$A+B=(a_{ij}+b_{ij})_{m\times n},\quad \lambda A=(\lambda a_{ij})_{m\times n}.$$

类似地, 可以定义矩阵的**减法**运算和**负矩阵**

$$A-B=(a_{ij}-b_{ij})_{m\times n},\quad -A=(-a_{ij})_{m\times n}. \tag{4.4}$$

例4.2.1. 设A,B,C,D,E为例4.1.6中的矩阵, 则有

$$A+B=\begin{pmatrix}1 & 1\\ \frac{3}{2} & 1\end{pmatrix},\quad C-D=\begin{pmatrix}-x & \pi\\ -\pi & -x^2\end{pmatrix},\quad 2E=\begin{pmatrix}1+\sqrt{3}i & 0\\ 0 & 1+\sqrt{3}i\end{pmatrix}$$

定理4.2.1. *矩阵的加法和数乘运算具有下列性质:*

(1) **加法交换律** $A+B=B+A$

(2) **加法结合律** $(A+B)+C=A+(B+C)$

(3) **有零矩阵** $A+O=O+A=A$

(4) **有负矩阵** $A+(-A)=(-A)+A=O$

(5) **左分配律** $(\lambda+\mu)A=\lambda A+\mu A$

(6) **右分配律** $\lambda(A+B)=\lambda A+\lambda B$

(7) **数乘结合律** $(\lambda\mu)A=\lambda(\mu A)$

(8) **数乘单位元** $1A=A$

其中A,B,C是使运算有意义的矩阵, λ,μ是数.

证明. 以上性质可通过矩阵运算的定义直接验证, 留作习题. □

同数组向量类似, 我们也可以定义一组矩阵的**线性组合**概念.

定义4.2.2. 设$A_1,A_2,\ldots,A_k\in F^{m\times n}$是一组矩阵, $\lambda_1,\lambda_2,\ldots,\lambda_k\in F$是一组数. 称和式

$$\lambda_1A_1+\lambda_2A_2+\ldots+\lambda_kA_k \tag{4.5}$$

为$A_1,A_2,\ldots,A_k$的线性组合.

第(i,j)元素是1, 其它元素全是0的$m\times n$矩阵

$$\begin{array}{cc} \begin{pmatrix} 0 & \cdots & 0 \\ \vdots & 1 & \vdots \\ 0 & \cdots & 0 \end{pmatrix} & \leftarrow \text{第}i\text{行} \\ \uparrow & \\ \text{第}j\text{列} & \end{array}$$

称为**基本矩阵**, 记作E_{ij}. 每个$m\times n$矩阵$A=(a_{ij})_{m\times n}$都可以表示为基本矩阵的线性组合:

$$A=\sum_{i=1}^{m}\sum_{j=1}^{n}a_{ij}E_{ij}.$$

§4.2.2 矩阵的乘法

在定义矩阵的乘法运算之前, 我们首先看一个例子.

例4.2.2. 给定线性映射$\mathscr{A}$: $\mathbf{y}\in F^n\mapsto \mathbf{x}\in F^m$

$$\begin{cases} x_1=a_{11}y_1+a_{12}y_2+\cdots+a_{1n}y_n, \\ x_2=a_{21}y_1+a_{22}y_2+\cdots+a_{2n}y_n, \\ \qquad\qquad\cdots\cdots\cdots \\ x_m=a_{m1}y_1+a_{m2}y_2+\cdots+a_{mn}y_n. \end{cases} \tag{4.6}$$

该映射可以用矩阵$A=(a_{ij})_{m\times n}$来表示. 再给定线性映射$\mathscr{B}$: $\mathbf{z}\in F^p\mapsto \mathbf{y}\in F^n$

$$\begin{cases} y_1=b_{11}z_1+b_{12}z_2+\cdots+b_{1p}z_p, \\ y_2=b_{21}z_1+b_{22}z_2+\cdots+b_{2p}z_p, \\ \qquad\qquad\cdots\cdots\cdots \\ y_n=b_{n1}z_1+b_{n2}z_2+\cdots+b_{np}z_p. \end{cases} \tag{4.7}$$

该线性映射可以用矩阵$B=(b_{ij})_{n\times p}$来表示.

下面我们来考虑$\mathscr{A}$与$\mathscr{B}$的复合映射. 由

$$x_i=\sum_{k=1}^{n}a_{ik}y_k=\sum_{k=1}^{n}a_{ik}\sum_{j=1}^{p}b_{kj}z_j=\sum_{j=1}^{p}\sum_{k=1}^{n}a_{ik}b_{kj}z_j$$

可知, 线性映射$\mathscr{A}$和$\mathscr{B}$的复合$\mathscr{C}=\mathscr{A}\circ\mathscr{B}:\mathbf{z}\mapsto\mathbf{x}$仍是一个线性映射:

$$\begin{cases} x_1 = c_{11}z_1 + c_{12}z_2 + \cdots + c_{1p}z_p, \\ x_2 = c_{21}z_1 + c_{22}z_2 + \cdots + c_{2p}z_p, \\ \qquad\qquad \cdots\cdots\cdots \\ x_m = c_{m1}z_1 + c_{m2}z_2 + \cdots + c_{mp}z_p. \end{cases} \tag{4.8}$$

其中

$$c_{ij} = \sum_{k=1}^{n} a_{ik}b_{kj}, \quad 1 \le i \le m, 1 \le j \le p.$$

$\mathscr{C}$对应的矩阵为$C=(c_{ij})_{m\times p}$.

由于线性映射与矩阵之间存在着一一对应的关系, 线性映射的复合对应于矩阵的某种运算, 我们定义它为矩阵的**乘积**.

定义4.2.3. 设矩阵$A=(a_{ij})_{m\times n}\in F^{n\times p}$与$B=(b_{ij})_{n\times p}\in F^{n\times p}$, 定义$A$与$B$的**乘积**为$m\times p$阶矩阵$C:=(c_{ij})_{m\times p}$, 其中元素

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

是A的第i行与B的第j列对应元素的乘积之和. 简记为$C=AB$.

我们看两个矩阵乘法的例子.

例4.2.3. 设

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix}, \quad B = \begin{pmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{pmatrix}, \quad C = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \quad D = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}.$$

则有

$$AB = \begin{pmatrix} 1\cdot1+2\cdot3+3\cdot5 & 1\cdot2+2\cdot4+3\cdot6 \\ 4\cdot1+5\cdot3+6\cdot5 & 4\cdot2+5\cdot4+6\cdot6 \end{pmatrix} = \begin{pmatrix} 22 & 28 \\ 49 & 64 \end{pmatrix},$$

$$BA = \begin{pmatrix} 1\cdot1+2\cdot4 & 1\cdot2+2\cdot5 & 1\cdot3+2\cdot6 \\ 3\cdot1+4\cdot4 & 3\cdot2+4\cdot5 & 3\cdot3+4\cdot6 \\ 5\cdot1+6\cdot4 & 5\cdot2+6\cdot5 & 5\cdot3+6\cdot6 \end{pmatrix} = \begin{pmatrix} 9 & 12 & 15 \\ 19 & 26 & 33 \\ 29 & 40 & 51 \end{pmatrix},$$

$$CD = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \qquad DC = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}, \qquad CC = DD = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}.$$

例4.2.4. 设

$$A=\begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \cdots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix},\quad B=\begin{pmatrix} b_1 & & \\ & \ddots & \\ & & b_m \end{pmatrix},\quad C=\begin{pmatrix} c_1 & & \\ & \ddots & \\ & & c_n \end{pmatrix}.$$

则有

$$BA=\begin{pmatrix} b_1a_{11} & \cdots & b_1a_{1n} \\ \vdots & \cdots & \vdots \\ b_ma_{m1} & \cdots & b_ma_{mn} \end{pmatrix},\quad AC=\begin{pmatrix} a_{11}c_1 & \cdots & a_{1n}c_n \\ \vdots & \cdots & \vdots \\ a_{m1}c_1 & \cdots & a_{mn}c_n \end{pmatrix}.$$

注2. 根据矩阵乘法的定义和上面的例子,我们有以下的观察:

(1) 并非任意两个矩阵A与B都可以相乘. 只有当A的列数等于B的行数时,A与B才可以相乘.

(2) 即使A与B是同阶方阵, AB与BA也不一定相等. 若$AB=BA$, 则称A,B乘法可交换.

(3) 两个非零矩阵的乘积可能是零矩阵, 即由$AB=O$不能推出$A=O$或$B=O$.

(4) 在A的左边乘上对角阵相当于将A的各行分别乘上一个数, 在A的右边乘上对角阵相当于将A的各列分别乘上一个数.

(5) 数量阵λI与矩阵A相乘的效果等于矩阵的数乘λA; 特别地, $IA=AI=A$, $OA=AO=O$.

定理4.2.2. *矩阵的乘法运算具有以下性质:*

(1) 乘法结合律 $(AB)C=A(BC)$

(2) 乘法单位元 $IA=AI=A$

(3) 左分配律 $(A+B)C=AC+BC$

(4) 右分配律 $A(B+C)=AB+AC$

(5) 数乘结合律 $\lambda(AB)=(\lambda A)B=A(\lambda B)$

其中A,B,C是使运算有意义的矩阵, λ是数.

证明. 我们仅证明性质(1), 其余性质证明留作习题.

设$A=(a_{ij})_{m\times n}, B=(b_{ij})_{n\times p}, C=(c_{ij})_{p\times q}$. 则有

$$\big((AB)C\big)_{ij}=\sum_{l=1}^{p}(AB)_{il}c_{lj}=\sum_{l=1}^{p}\sum_{k=1}^{n}a_{ik}b_{kl}c_{lj},$$

$$\big(A(BC)\big)_{ij}=\sum_{k=1}^{n}a_{ik}(BC)_{kj}=\sum_{k=1}^{n}\sum_{l=1}^{p}a_{ik}b_{kl}c_{lj}.$$

通过交换求和的次序知, 以上两式相等, 故$(AB)C=A(BC)$. □

矩阵的乘法运算在矩阵运算中有着特别重要的地位. 许多代数问题都可以利用矩阵乘法得到简洁的表示和求解. 例如, 线性方程组 (2.1) 可以被简单地写成矩阵乘积的形式

$$A\mathbf{x}=\mathbf{b},$$

其中$A=(a_{ij})_{m\times n}$称为线性方程组的**系数矩阵**, $\mathbf{x}$和$\mathbf{b}$为列向量

$$\mathbf{x}=\begin{pmatrix}x_1\\x_2\\\cdots\\x_n\end{pmatrix},\qquad \mathbf{b}=\begin{pmatrix}b_1\\b_2\\\cdots\\b_m\end{pmatrix}.$$

再如, 例 4.1.2中的线性映射可以表示为矩阵乘积形式

$$\begin{pmatrix}x_1\\\vdots\\x_m\end{pmatrix}=\begin{pmatrix}a_{11}&\cdots&a_{1n}\\\vdots&\vdots&\vdots\\a_{m1}&\cdots&a_{mn}\end{pmatrix}\begin{pmatrix}y_1\\\vdots\\y_n\end{pmatrix}$$

类似地, 例 4.1.1中的坐标变换公式可以写成

$$\begin{pmatrix}x'\\y'\end{pmatrix}=\begin{pmatrix}\cos(\theta)&-\sin(\theta)\\\sin(\theta)&\cos(\theta)\end{pmatrix}\begin{pmatrix}x\\y\end{pmatrix}$$

及

$$\begin{pmatrix}x'\\y'\\z'\end{pmatrix}=\begin{pmatrix}\cos(\theta)&-\sin(\theta)&0\\\sin(\theta)&\cos(\theta)&0\\0&0&1\end{pmatrix}\begin{pmatrix}x\\y\\z\end{pmatrix}$$

以后, 我们会看到更多的矩阵乘法的应用.

通过矩阵的乘法, 可以定义任意方阵A的**正整数次幂**

$$A^k=\underbrace{A\cdots A}_{k},\quad k=1,2,\cdots \tag{4.9}$$

另外, 对任意方阵A, 规定$A^0=I$. 有了方阵的各次幂, 就可以将方阵代入多项式求值. 设多项式$f(x)=c_0+c_1x+\cdots+c_kx^k$, 定义**矩阵多项式**

$$f(A)=c_0I+c_1A+\cdots+c_kA^k. \tag{4.10}$$

例4.2.5. 设$A=\begin{pmatrix}\cos(\theta) & -\sin(\theta)\\ \sin(\theta) & \cos(\theta)\end{pmatrix}$. 求$A^n$, 这里$n$为正整数.

解. 容易验证

$$A^2=\begin{pmatrix}\cos(\theta) & -\sin(\theta)\\ \sin(\theta) & \cos(\theta)\end{pmatrix}\begin{pmatrix}\cos(\theta) & -\sin(\theta)\\ \sin(\theta) & \cos(\theta)\end{pmatrix}=\begin{pmatrix}\cos(2\theta) & -\sin(2\theta)\\ \sin(2\theta) & \cos(2\theta)\end{pmatrix}.$$

用归纳法, 一般可证

$$A^n=\begin{pmatrix}\cos(n\theta) & -\sin(n\theta)\\ \sin(n\theta) & \cos(n\theta)\end{pmatrix}.$$

上式有明显的几何意义. 我们知道, A表示逆时针旋转θ角的旋转变换. 因此, A^2表示该变换与自身的复合变换, 从而是旋转2θ角的变换. 一般地, A^n表示n个旋转θ角的变换的复合, 从而是一个旋转$n\theta$角的变换. □

例4.2.6. 设A,B为同阶方阵, 且$AB=BA$. 证明:

$$(A+B)^n=C_n^0A^n+C_n^1A^{n-1}B+\ldots+C_n^{n-1}AB^{n-1}+C_n^nB^n.$$

证明. : 用数学归纳法, 请读者自己完成. □

例4.2.7. 设$A=\begin{pmatrix}a & b\\ 0 & a\end{pmatrix}$. 求$A^n$, 这里$n$为正整数.

解. 将A分解为$A=aI_2+N$, 这里$N=\begin{pmatrix}0 & b\\ 0 & 0\end{pmatrix}$. 易验证$N^k=0, k\geq 2$. 于是

$$A^n=(aI_2+N)^n=a^nI_2+C_n^1a^{n-1}N=\begin{pmatrix}a^n & na^{n-1}b\\ 0 & a^n\end{pmatrix}.$$

□

下面我们给出关于矩阵乘法的行列式的一个重要结论.

定理4.2.3. 设A,B均为同阶方阵. 则$\det(AB)=\det(A)\det(B)$.

证明. : 设$A=(a_{ij})_{n\times n}, B=(b_{ij})_{n\times n}$. 用$\beta_1,\ldots,\beta_n$表示$B$的$n$个行. 则$AB$的第$i$行可以写成

$$\left(\sum_{k=1}^{n} a_{ik}b_{k1},\ldots,\sum_{k=1}^{n} a_{ik}b_{kn}\right)=\sum_{k=1}^{n} a_{ik}\left(b_{k1},\ldots,b_{kn}\right)=\sum_{k=1}^{n} a_{ik}\beta_k.$$

这样

$$\det(AB)=\det\left(\sum_{j_1=1}^{n} a_{1j_1}\beta_{j_1},\ldots,\sum_{j_n=1}^{n} a_{nj_n}\beta_{j_n}\right).$$

利用行列式的多重线性性得

$$\det(AB)=\sum_{j_1=1}^{n}\cdots\sum_{j_n=1}^{n} a_{1j_1}\cdots a_{nj_n}\det(\beta_{j_1},\ldots,\beta_{j_n}).$$

上式中, 当下标$j_1,\ldots,j_n$中有两个相等时, 对应项为零. 因此

$$\begin{aligned}\det(AB)&=\sum_{(j_1\cdots j_n)\in S_n} a_{1j_1}\cdots a_{nj_n}\det(\beta_{j_1},\ldots,\beta_{j_n})\\&=\sum_{(j_1\cdots j_n)\in S_n} a_{1j_1}\cdots a_{nj_n}(-1)^{\tau(j_1\cdots j_n)}\det(\beta_1,\ldots,\beta_n)\\&=\det(A)\det(B).\end{aligned}$$

定理证毕. □

上述定理有明确的几何意义. 如图 §4.2.2所示. 设A_0为单位正方形, 其面积$S_0=1$. 经线性变换$\mathscr{B}:\mathbf{y}=B\mathbf{x}$, A_0 变为平行四边形A_1, 其面积$S_1=\det(B)$. 平行四边形A_1经线性变换$\mathscr{A}:\mathbf{z}=A\mathbf{y}$变为平行四边形$A_2$, 其面积$S_2=\det(A)S_1=\det(A)\det(B)$. 另一方面, A_2 是由A_0经复合线性变换$\mathscr{A}\circ\mathscr{B}$得到, 故$S_2=\det(AB)$. 因此, $\det(AB)=\det(A)\det(B)$.

例4.2.8. 计算行列式

$$\begin{vmatrix}3 & x+y+z & x^2+y^2+z^2\\ x+y+z & x^2+y^2+z^2 & x^3+y^3+z^3\\ x^2+y^2+z^2 & x^3+y^3+z^3 & x^4+y^4+z^4\end{vmatrix}$$

解. 由于

$$\begin{pmatrix}3 & x+y+z & x^2+y^2+z^2\\ x+y+z & x^2+y^2+z^2 & x^3+y^3+z^3\\ x^2+y^2+z^2 & x^3+y^3+z^3 & x^4+y^4+z^4\end{pmatrix}=\begin{pmatrix}1&1&1\\x&y&z\\x^2&y^2&z^2\end{pmatrix}\begin{pmatrix}1&x&x^2\\1&y&y^2\\1&z&z^2\end{pmatrix}$$

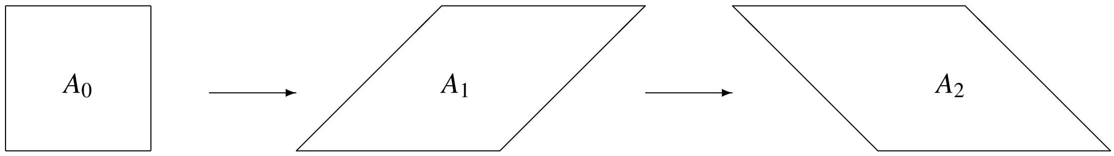


图 4.3 矩阵乘积行列式的几何意义

因此

$$\begin{aligned}&\begin{vmatrix}3 & x+y+z & x^2+y^2+z^2\\ x+y+z & x^2+y^2+z^2 & x^3+y^3+z^3\\ x^2+y^2+z^2 & x^3+y^3+z^3 & x^4+y^4+z^4\end{vmatrix}=\begin{vmatrix}1&1&1\\x&y&z\\x^2&y^2&z^2\end{vmatrix}\cdot\begin{vmatrix}1&x&x^2\\1&y&y^2\\1&z&z^2\end{vmatrix}\\&=(x-y)^2(y-z)^2(z-x)^2.\end{aligned}$$

□

§4.2.3 逆矩阵

我们定义了矩阵的乘法, 自然也希望能够定义乘法的逆运算"除法". 为此我们引入逆矩阵的概念.

定义4.2.4. 设A是一个n阶方阵, 如果存在n阶方阵X满足$XA=AX=I$, 则称A**可逆**, 并称X为A的**逆矩阵**, 记作A^{-1}. 可逆方阵也称为**非奇异**方阵, 称不可逆方阵为**奇异**方阵.

容易知道, 若A可逆, 其逆矩阵是唯一的. 这是因为: 若X和Y都是A的逆矩阵, 则有

$$X=XI=X(AY)=(XA)Y=IY=Y.$$

显然不是每个方阵都可逆, 例如零方阵就不可逆. 那么如何判断一个给定的方阵是否可逆呢?

定理4.2.4. 设$A=(a_{ij})_{n\times n}$为n阶方阵. 则A可逆当且仅当$\det(A)\neq 0$. 且当A可逆时,

$$A^{-1}=\frac{1}{\det(A)}A^*, \tag{4.11}$$

这里

$$A^* = \begin{pmatrix} A_{11} & A_{21} & \cdots & A_{n1} \\ A_{12} & A_{22} & \cdots & A_{n2} \\ \vdots & \vdots & \vdots & \vdots \\ A_{1n} & A_{2n} & \cdots & A_{nn} \end{pmatrix}, \tag{4.12}$$

A_{ij}是a_{ij}的关于行列式$\det(A)$的代数余子式. 称A^*为A的**伴随方阵**.

证明. 若A可逆, 则存在n阶方阵X使得$AX=I$. 上式两边取行列式得 $\det(A)\det(X)=1$. 从而$\det(A)\neq 0$.

反过来, 设$\det(A)\neq 0$. 则由第二章定理 3.4.1直接验证可得, $AA^*=A^*A=\det(A)I^{(n)}$. 因此, $A^{-1}=\frac{1}{\det(A)}A^*$. □

注3. 上述定理还表明, n阶方阵A可逆当且仅当存在n阶方阵X使得

$$AX=I \quad 或者 \quad XA=I.$$

注4. 设$A\in F^{n\times n}$, $\mathbf{b}\in F^n$. 则当$\det(A)\neq 0$时, 线性方程组$A\mathbf{x}=\mathbf{b}$的解

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

由此可以立即得到Cramer法则.

下面给出可逆矩阵的基本性质.

定理4.2.5. 对任意同阶可逆方阵A,B, 都有

(1) $(A^{-1})^{-1}=A$

(2) $(\lambda A)^{-1}=\lambda^{-1}A^{-1}$, $\lambda\neq 0$

(3) $(AB)^{-1}=B^{-1}A^{-1}$

证明. 可通过逆矩阵的定义直接验证, 留作习题. □

例4.2.9. 证明:

(1) 当$ad\neq bc$时,

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}^{-1} = \frac{1}{ad-bc}\begin{pmatrix} d & -b \\ -c & a \end{pmatrix}.$$

(2) 设向量$\mathbf{a}=(a_1,a_2,a_3),\mathbf{b}=(b_1,b_2,b_3),\mathbf{c}=(c_1,c_2,c_3)$, 则当$(\mathbf{a}\times\mathbf{b})\cdot\mathbf{c}\neq 0$ 时,

$$\begin{pmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{pmatrix}^{-1} = \frac{1}{(\mathbf{a}\times\mathbf{b})\cdot\mathbf{c}} \begin{pmatrix} u_1 & v_1 & w_1 \\ u_2 & v_2 & w_2 \\ u_3 & v_3 & w_3 \end{pmatrix},$$

其中$\mathbf{u}=(u_1,u_2,u_3)=\mathbf{b}\times\mathbf{c},\mathbf{v}=(v_1,v_2,v_3)=\mathbf{c}\times\mathbf{a},\mathbf{w}=(w_1,w_2,w_3)=\mathbf{a}\times\mathbf{b}$.

(3) 当$a_1\cdots a_n\neq 0$时,

$$\begin{pmatrix} a_1 & & \\ & \ddots & \\ & & a_n \end{pmatrix}^{-1} = \begin{pmatrix} a_1^{-1} & & \\ & \ddots & \\ & & a_n^{-1} \end{pmatrix}.$$

证明. 我们只证明(2), (1)与(3)留给读者完成.

$$\begin{pmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{pmatrix} \begin{pmatrix} u_1 & v_1 & w_1 \\ u_2 & v_2 & w_2 \\ u_3 & v_3 & w_3 \end{pmatrix} = \begin{pmatrix} \mathbf{a}\cdot\mathbf{u} & \mathbf{a}\cdot\mathbf{v} & \mathbf{a}\cdot\mathbf{w} \\ \mathbf{b}\cdot\mathbf{u} & \mathbf{b}\cdot\mathbf{v} & \mathbf{b}\cdot\mathbf{w} \\ \mathbf{c}\cdot\mathbf{u} & \mathbf{c}\cdot\mathbf{v} & \mathbf{c}\cdot\mathbf{w} \end{pmatrix}.$$

右边矩阵的对角元为$(\mathbf{b}\times\mathbf{c})\cdot\mathbf{a}=(\mathbf{c}\times\mathbf{a})\cdot\mathbf{b}=(\mathbf{a}\times\mathbf{b})\cdot\mathbf{c}$, 而非对角元全为零. 例如, $(1,2)$元素为$\mathbf{a}\cdot(\mathbf{c}\times\mathbf{a})=0$. 因此, 右边矩阵为数量矩阵$(\mathbf{a}\times\mathbf{b})\cdot\mathbf{c}I$, 从而结论正确. □

§4.2.4 转置、共轭与迹

定义4.2.5. 将矩阵$A=(a_{ij})_{m\times n}$的行列互换, 得到的矩阵

$$\begin{pmatrix} a_{11} & a_{21} & \cdots & a_{m1} \\ a_{12} & a_{22} & \cdots & a_{m2} \\ \vdots & \vdots & \cdots & \vdots \\ a_{1n} & a_{2n} & \cdots & a_{mn} \end{pmatrix}, \tag{4.13}$$

称为A的**转置矩阵**, 记作$A^T=(a_{ji})_{n\times m}$.

将复矩阵A的每个元素换成它的共轭复数, 得到的矩阵

$$\begin{pmatrix} \overline{a_{11}} & \overline{a_{12}} & \cdots & \overline{a_{1n}} \\ \overline{a_{21}} & \overline{a_{22}} & \cdots & a_{2n} \\ \vdots & \vdots & \cdots & \vdots \\ \overline{a_{n1}} & \overline{a_{n2}} & \cdots & \overline{a_{mn}} \end{pmatrix}, \tag{4.14}$$

称为A的**共轭矩阵**, 记作$\overline{A}=(\overline{a_{ij}})_{m\times n}$.

n阶方阵A的对角元之和

$$a_{11}+a_{22}+\cdots+a_{nn} \tag{4.15}$$

称为A的**迹**, 记作$\mathrm{tr}(A)$.

例4.2.10. 设$A=\begin{pmatrix}1+i & 2-i\\ 3-i & 4+i\end{pmatrix}$, 计算$A^T$、$\overline{A}$、$\mathrm{tr}(A)$.

解. $A^T=\begin{pmatrix}1+i & 3-i\\ 2-i & 4+i\end{pmatrix}, \overline{A}=\begin{pmatrix}1-i & 2+i\\ 3+i & 4-i\end{pmatrix}, \mathrm{tr}(A)=5+2i.$ □

定理4.2.6. 矩阵的转置运算具有以下性质:

(1) $(A+B)^T=A^T+B^T$

(2) $(\lambda A)^T=\lambda A^T$

(3) $(AB)^T=B^TA^T$

(4) $(A^{-1})^T=(A^T)^{-1}$

其中A,B是使运算有意义的矩阵, λ是数.

证明. 性质(1)与(2)可通过矩阵运算的定义直接验证, 留作习题. 下证性质(3)与(4). 设$A=(a_{ij})_{m\times n}, B=(b_{ij})_{n\times p}$, 则有

$$\left((AB)^T\right)_{ij}=(AB)_{ji}=\sum_{k=1}^{n}a_{jk}b_{ki},$$

$$\left(B^TA^T\right)_{ij}=\sum_{k=1}^{n}b_{ki}a_{jk}.$$

上述两和式相等, 故$(AB)^T=B^TA^T$. 特别地, 对可逆方阵A,

$$(A^{-1})^TA^T=(AA^{-1})^T=I,\quad A^T(A^{-1})^T=(A^{-1}A)^T=I.$$

根据定义4.2.4, $(A^{-1})^T=(A^T)^{-1}$. □

定理4.2.7. 矩阵的迹具有以下性质:

(1) $\mathrm{tr}(A+B)=\mathrm{tr}(A)+\mathrm{tr}(B)$

(2) $\mathrm{tr}(\lambda A) = \lambda\,\mathrm{tr}(A)$

(3) $\mathrm{tr}(A^T) = \mathrm{tr}(A)$, $\mathrm{tr}(\overline{A}) = \overline{\mathrm{tr}(A)}$

(4) $\mathrm{tr}(AB) = \mathrm{tr}(BA)$

其中A,B是使运算有意义的矩阵, λ是数.

证明. 性质(1)、(2)、(3)可通过矩阵运算的定义直接验证, 留作习题. 下证性质(4).

设$A=(a_{ij})_{m\times n}, B=(b_{ij})_{n\times m}$, 则有

$$\mathrm{tr}(AB) = \sum_{i=1}^{m}(AB)_{ii} = \sum_{i=1}^{m}\sum_{j=1}^{n} a_{ij}b_{ji},$$

$$\mathrm{tr}(BA) = \sum_{j=1}^{n}(BA)_{jj} = \sum_{j=1}^{n}\sum_{i=1}^{m} b_{ji}a_{ij}.$$

通过交换求和的次序, 以上两式相等, 故$\mathrm{tr}(AB) = \mathrm{tr}(BA)$. □

例4.2.11. 设复矩阵A满足$\mathrm{tr}(A\overline{A}^T) = 0$, 证明:$A = O$.

证明. 设$A=(a_{ij})_{m\times n}, A\overline{A}^T = (b_{ij})_{m\times m}$. 由

$$0 = \mathrm{tr}(A\overline{A}^T) = \sum_{i=1}^{m} b_{ii} = \sum_{i=1}^{m}\sum_{j=1}^{n} a_{ij}\overline{a_{ij}} = \sum_{i=1}^{m}\sum_{j=1}^{n} |a_{ij}|^2$$

可知所有$a_{ij} = 0$, 即$A = O$. □

§4.2.5 分块运算

在矩阵运算过程中, 有时候将某些子矩阵看成一个整体来参加运算会方便许多. 例如, 可以将矩阵看成一组行向量或列向量:

$$A = \begin{array}{cc} \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \cdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} & \begin{array}{c} \leftarrow \mathbf{b}_1 \\ \leftarrow \mathbf{b}_2 \\ \vdots \\ \leftarrow \mathbf{b}_m \end{array} \\ \begin{array}{cccc} \uparrow & \uparrow & & \uparrow \\ \mathbf{a}_1 & \mathbf{a}_2 & \cdots & \mathbf{a}_n \end{array} & \end{array}$$

记作

$$A = \begin{pmatrix} \mathbf{a}_1 & \mathbf{a}_2 & \cdots & \mathbf{a}_n \end{pmatrix} \quad 或 \quad A = \begin{pmatrix} \mathbf{b}_1 \\ \mathbf{b}_2 \\ \vdots \\ \mathbf{b}_m \end{pmatrix}$$

一般地, 可以将A同时按行按列分成若干块

$$A=\begin{pmatrix}A_{11} & A_{12} & \cdots & A_{1s}\\ A_{21} & A_{22} & \cdots & A_{2s}\\ \vdots & \vdots & \cdots & \vdots\\ A_{r1} & A_{r2} & \cdots & A_{rs}\end{pmatrix}$$

称为**分块矩阵**, 记作$A=(A_{ij})_{r\times s}$, 每个A_{ij}称为A的**子块**.

更一般地, 由A的第$i_1,i_2,\cdots,i_r$行和第$j_1,j_2,\cdots,j_s$列上的元素依次排列组成的$r\times s$矩阵称为A的**子矩阵**, 记作

$$A\begin{pmatrix}i_1 i_2\cdots i_r\\ j_1 j_2\cdots j_s\end{pmatrix}=\begin{pmatrix}a_{i_1j_1} & a_{i_1j_2} & \cdots & a_{i_1j_s}\\ a_{i_2j_1} & a_{i_2j_2} & \cdots & a_{i_2j_s}\\ \vdots & \vdots & \cdots & \vdots\\ a_{i_rj_1} & a_{i_rj_2} & \cdots & a_{i_rj_s}\end{pmatrix} \tag{4.16}$$

若$A_{ij}=O$对所有$i\neq j$成立, 则称A为**准对角阵**, 记作

$$A=\begin{pmatrix}A_{11} & & \\ & \ddots & \\ & & A_{rr}\end{pmatrix} \quad 或 \quad A=\mathrm{diag}(A_{11},A_{22},\cdots,A_{rr}).$$

若$A_{ij}=O$ 对所有$i>j$成立, 则A称为**准上三角阵**. 若$A_{ij}=O$对所有$i<j$成立, 则称A为**准下三角阵**. 准上三角阵和准下三角阵统称为**准三角阵**.

例4.2.12. 若将矩阵

$$A=\begin{pmatrix}1 & 1 & 0 & 0\\ 0 & 1 & 0 & 0\\ 0 & 0 & 1 & 0\\ 0 & 0 & 1 & 1\end{pmatrix}$$

按照下列方式分块

$$\left(\begin{array}{c|ccc}1 & 1 & 0 & 0\\ \hline 0 & 1 & 0 & 0\\ 0 & 0 & 1 & 1\\ 0 & 0 & 1 & 1\end{array}\right), \quad \left(\begin{array}{cc|cc}1 & 1 & 0 & 0\\ 0 & 1 & 0 & 0\\ \hline 0 & 0 & 1 & 0\\ 0 & 0 & 1 & 1\end{array}\right), \quad \left(\begin{array}{ccc|c}1 & 1 & 0 & 0\\ 0 & 1 & 0 & 0\\ 0 & 0 & 1 & 0\\ \hline 0 & 0 & 1 & 1\end{array}\right).$$

则A分别成为准上三角阵、准对角阵、准下三角阵.

由此可见,“准对角”、“准上三角”、“准下三角”的概念并不是孤立的,是和矩阵分块的方式联系在一起的. 在实际运用的时候,选择恰当的分块方式也许会有助于问题的解决.

下面我们考虑分块矩阵的运算. 先看一个例子. 假设矩阵

$$A=\left(\begin{array}{cc|cc} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ \hline a_{31} & a_{32} & a_{33} & a_{34} \\ a_{41} & a_{42} & a_{43} & a_{44} \end{array}\right), \quad B=\left(\begin{array}{cc|cc} b_{11} & b_{12} & b_{13} & b_{14} \\ b_{21} & b_{22} & b_{23} & b_{24} \\ \hline b_{31} & b_{32} & b_{33} & b_{34} \\ b_{41} & b_{42} & b_{43} & b_{44} \end{array}\right)$$

按照以上方式分块成

$$A=\begin{pmatrix} A_1 & A_2 \\ A_3 & A_4 \end{pmatrix}, \quad B=\begin{pmatrix} B_1 & B_2 \\ B_3 & B_4 \end{pmatrix},$$

则有

$$A+B=\left(\begin{array}{cc|cc} a_{11}+b_{11} & a_{12}+b_{12} & a_{13}+b_{13} & a_{14}+b_{14} \\ a_{21}+b_{21} & a_{22}+b_{22} & a_{23}+b_{23} & a_{24}+b_{24} \\ \hline a_{31}+b_{31} & a_{32}+b_{32} & a_{33}+b_{33} & a_{34}+b_{34} \\ a_{41}+b_{41} & a_{42}+b_{42} & a_{43}+b_{43} & a_{44}+b_{44} \end{array}\right)=\begin{pmatrix} A_1+B_1 & A_2+B_2 \\ A_3+B_3 & A_4+B_4 \end{pmatrix},$$

$$AB=\left(\begin{array}{cc|cc} \begin{gathered}a_{11}b_{11}+a_{12}b_{21}\\+a_{13}b_{31}+a_{14}b_{41}\end{gathered} & \begin{gathered}a_{11}b_{12}+a_{12}b_{22}\\+a_{13}b_{32}+a_{14}b_{42}\end{gathered} & \begin{gathered}a_{11}b_{13}+a_{12}b_{23}\\+a_{13}b_{33}+a_{14}b_{43}\end{gathered} & \begin{gathered}a_{11}b_{14}+a_{12}b_{24}\\+a_{13}b_{34}+a_{14}b_{44}\end{gathered} \\ \begin{gathered}a_{21}b_{11}+a_{22}b_{21}\\+a_{23}b_{31}+a_{24}b_{41}\end{gathered} & \begin{gathered}a_{21}b_{12}+a_{22}b_{22}\\+a_{23}b_{32}+a_{24}b_{42}\end{gathered} & \begin{gathered}a_{21}b_{13}+a_{22}b_{23}\\+a_{23}b_{33}+a_{24}b_{43}\end{gathered} & \begin{gathered}a_{21}b_{14}+a_{22}b_{24}\\+a_{23}b_{34}+a_{24}b_{44}\end{gathered} \\ \hline \begin{gathered}a_{31}b_{11}+a_{32}b_{21}\\+a_{33}b_{31}+a_{34}b_{41}\end{gathered} & \begin{gathered}a_{31}b_{12}+a_{32}b_{22}\\+a_{33}b_{32}+a_{34}b_{42}\end{gathered} & \begin{gathered}a_{31}b_{13}+a_{32}b_{23}\\+a_{33}b_{33}+a_{34}b_{43}\end{gathered} & \begin{gathered}a_{31}b_{14}+a_{32}b_{24}\\+a_{33}b_{34}+a_{34}b_{44}\end{gathered} \\ \begin{gathered}a_{41}b_{11}+a_{42}b_{21}\\+a_{43}b_{31}+a_{44}b_{41}\end{gathered} & \begin{gathered}a_{41}b_{12}+a_{42}b_{22}\\+a_{43}b_{32}+a_{44}b_{42}\end{gathered} & \begin{gathered}a_{41}b_{13}+a_{42}b_{23}\\+a_{43}b_{33}+a_{44}b_{43}\end{gathered} & \begin{gathered}a_{41}b_{14}+a_{42}b_{24}\\+a_{43}b_{34}+a_{44}b_{44}\end{gathered} \end{array}\right)$$

$$=\begin{pmatrix} A_1B_1+A_2B_3 & A_1B_2+A_2B_4 \\ A_3B_1+A_4B_2 & A_3B_2+A_4B_4 \end{pmatrix}.$$

类似地, 对于数乘、转置、共轭、迹运算, 还有

$$\lambda A=\begin{pmatrix} \lambda A_1 & \lambda A_2 \\ \lambda A_3 & \lambda A_4 \end{pmatrix},\ A^T=\begin{pmatrix} A_1^T & A_3^T \\ A_2^T & A_4^T \end{pmatrix},\ \overline{A}=\begin{pmatrix} \overline{A_1} & \overline{A_2} \\ \overline{A_3} & \overline{A_4} \end{pmatrix},\ \mathrm{tr}(A)=\mathrm{tr}(A_1)+\mathrm{tr}(A_4).$$

从以上例子可以看出, 在对分块矩阵施行矩阵的加法、数乘、乘法、转置、共轭、迹运算的时候, 可以把每个矩阵块看作是一个元素进行运算, 然后再对每个矩阵块施行同样的运算.

定理4.2.8. 矩阵的分块运算具有以下性质:

(1) 设$A=(A_{ij})_{r\times s}$, $B=(B_{ij})_{r\times s}$, 则$A+B=(A_{ij}+B_{ij})_{r\times s}$.

(2) 设$A=(A_{ij})_{r\times s}$, 则$\lambda A=(\lambda A_{ij})_{r\times s}$.

(3) 设$A=(A_{ij})_{r\times s}$, $B=(B_{ij})_{s\times t}$, 则$AB=(C_{ij})_{r\times t}$, 其中$C_{ij}=\sum\limits_{k=1}^{s}A_{ik}B_{kj}$.

(4) 设$A=(A_{ij})_{r\times s}$, 则$A^T=(A_{ji}^T)_{s\times r}$.

(5) 设$A=(A_{ij})_{r\times s}$是复方阵, 则$\overline{A}=(\overline{A_{ij}})_{r\times s}$.

(6) 设$A=(A_{ij})_{r\times r}$且每个A_{ii}都是方阵, 则$\mathrm{tr}(A)=\sum\limits_{i=1}^{r}\mathrm{tr}(A_{ii})$.

(7) 当$A_1,\cdots,A_r$都可逆时, $(\mathrm{diag}(A_1,\ldots,A_r))^{-1}=\mathrm{diag}(A_1^{-1},\ldots,A_r^{-1})$.

其中矩阵A,B的分块方式使运算有意义, λ是数.

证明. 可通过矩阵运算的定义直接验证, 留作习题. □

注5. 在上述定理(3)中, 若使矩阵的乘法运算有意义, 每个A_{ik}的列数必须等于B_{kj}的行数. 而且, 在公式$C_{ij}=\sum\limits_{k=1}^{s}A_{ik}B_{kj}$ 中, A_{ik}和B_{kj}的次序不可以交换.

下面, 我们说明分块在矩阵运算中的一些应用.

例4.2.13. 重新考虑例 4.2.4. 设矩阵A的行为$\mathbf{a}_1,\ldots,\mathbf{a}_m$, 列为$\mathbf{b}_1,\ldots,\mathbf{b}_n$. 则

$$BA=\begin{pmatrix}b_1&&\\&\ddots&\\&&b_m\end{pmatrix}\begin{pmatrix}\mathbf{a}_1\\ \vdots\\ \mathbf{a}_m\end{pmatrix}=\begin{pmatrix}b_1\mathbf{a}_1\\ \vdots\\ b_m\mathbf{a}_m\end{pmatrix}=\begin{pmatrix}b_1a_{11}&\cdots&b_1a_{1n}\\ \vdots&\vdots&\vdots\\ b_1a_{m1}&\cdots&b_ma_{mn}\end{pmatrix}.$$

同理

$$AC=\begin{pmatrix}\mathbf{b}_1&\cdots&\mathbf{b}_n\end{pmatrix}\begin{pmatrix}c_1&&\\&\ddots&\\&&c_n\end{pmatrix}=\begin{pmatrix}c_1\mathbf{b}_1&\cdots&c_n\mathbf{b}_n\end{pmatrix}=\begin{pmatrix}c_1a_{11}&\cdots&c_na_{1n}\\ \vdots&\vdots&\vdots\\ c_1a_{m1}&\cdots&c_na_{mn}\end{pmatrix}.$$

例4.2.14. 设A是一个$m\times n$矩阵, B是一个$n\times m$矩阵, $m>n$. 证明: $\det(AB)=0$.

证明. 在矩阵A的右端添加$m-n$列零向量使之成为方阵$\tilde{A}=(A,O)$. 在矩阵B的下端添加$m-n$行零向量使之成为方阵$\tilde{B}=\begin{pmatrix}B\\O\end{pmatrix}$. 则$\tilde{A}\tilde{B}=AB$. 根据定理 4.2.3,

$$\det(AB)=\det(\tilde{A}\tilde{B})=\det(\tilde{A})\det(\tilde{B})=0.$$

□

例4.2.15. 设准上三角方阵$A=(A_{ij})_{k\times k}$的每个对角块A_{ii}都是方阵, 则有

$$\det(A)=\begin{vmatrix}A_{11} & A_{12} & \cdots & A_{1k}\\ O & A_{22} & \cdots & A_{2k}\\ \vdots & \ddots & \ddots & \vdots\\ O & \cdots & O & A_{kk}\end{vmatrix}=\det(A_{11})\det(A_{22})\cdots\det(A_{kk}).$$

证明. 对k用归纳法. 当$k=1$时, 结论显然成立. 假设结论对$k-1$成立. 对$k\geq 2$, 将矩阵A可分块成

$$A=\begin{pmatrix}A_{11} & \bar{A}_{12}\\ O & \bar{A}_{22}\end{pmatrix},$$

其中

$$\bar{A}_{12}=\begin{pmatrix}A_{12} & \cdots & A_{1k}\end{pmatrix},\qquad \bar{A}_{22}=\begin{pmatrix}A_{22} & A_{23} & \cdots & A_{2k}\\ O & A_{33} & \cdots & A_{3k}\\ \vdots & \ddots & \ddots & \vdots\\ O & \cdots & O & A_{kk}\end{pmatrix}.$$

由第二章例 3.3.1知,

$$\det(A)=\det(A_{11})\det(\bar{A}_{22}).$$

而由归纳假设

$$\det(\bar{A}_{22})=\det(A_{22})\cdots\det(A_{kk}).$$

故

$$\det(A)=\det(A_{11})\det(A_{22})\cdots\det(A_{kk}).$$

□

例4.2.16. 给定矩阵

$$A=\begin{pmatrix}1&1&1&0\\0&1&0&1\\0&0&1&1\\0&0&0&1\end{pmatrix}.$$

计算A^n, 其中n为正整数.

解. 将A分块成

$$A=\begin{pmatrix} B & I \\ O & B \end{pmatrix}, \quad B=\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}.$$

用数学归纳法可以证明:

$$A^n=\begin{pmatrix} B^n & nB^{n-1} \\ O & B^n \end{pmatrix}, \qquad B^n=\begin{pmatrix} 1 & n \\ 0 & 1 \end{pmatrix}, \qquad n \geq 1.$$

于是

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

□

例4.2.17. 设$A \in F^{m\times m}$, $B \in F^{n\times n}$, $C \in F^{m\times n}$, $D \in F^{n\times m}$, 且A,B可逆. 证明: 矩阵$\begin{pmatrix} A & C \\ O & B \end{pmatrix}$与$\begin{pmatrix} A & O \\ D & B \end{pmatrix}$可逆, 并求它们的逆.

证明. 由于

$$\det\begin{pmatrix} A & C \\ O & B \end{pmatrix}=\det(A)\det(B)\neq 0,$$

因此矩阵$\begin{pmatrix} A & C \\ O & B \end{pmatrix}$可逆. 假设它的逆可以表示为分块矩阵形式$X=\begin{pmatrix} X_1 & X_2 \\ X_3 & X_4 \end{pmatrix}$, 其中$X_1 \in F^{m\times m}, X_2 \in F^{m\times n}, X_3 \in F^{n\times m}, X_4 \in F^{n\times n}$. 由

$$\begin{pmatrix} A & C \\ O & B \end{pmatrix}\begin{pmatrix} X_1 & X_2 \\ X_3 & X_4 \end{pmatrix}=\begin{pmatrix} I_m & O \\ O & I_n \end{pmatrix}$$

可得

$$AX_1+CX_3=I_m, \quad AX_2+CX_4=O, \quad BX_3=O, \quad BX_4=I.$$

由上式解得

$$X_3=O, \quad X_4=B^{-1}, \quad X_1=A^{-1}, \quad X_2=-A^{-1}CB^{-1}.$$

即

$$\begin{pmatrix} A & C \\ O & B \end{pmatrix}^{-1}=\begin{pmatrix} A^{-1} & -A^{-1}CB^{-1} \\ O & B^{-1} \end{pmatrix}.$$

类似可得
$$\begin{pmatrix} A & O \\ D & B \end{pmatrix}^{-1} = \begin{pmatrix} A^{-1} & O \\ -B^{-1}DA^{-1} & B^{-1} \end{pmatrix}.$$
特别地,
$$\begin{pmatrix} I & C \\ O & I \end{pmatrix}^{-1} = \begin{pmatrix} I & -C \\ O & I \end{pmatrix}, \quad \begin{pmatrix} I & O \\ D & I \end{pmatrix}^{-1} = \begin{pmatrix} I & O \\ -D & I \end{pmatrix}.$$

□

例4.2.18. 求例 4.2.16中矩阵A的逆.

解. 同例 4.2.16 类似, 将矩阵A分块为
$$A = \begin{pmatrix} B & I \\ O & B \end{pmatrix}, \quad B = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}.$$
由上例结论,
$$A^{-1} = \begin{pmatrix} B^{-1} & -B^{-2} \\ O & B^{-1} \end{pmatrix}.$$
由于
$$B^{-1} = \begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix}, \quad B^{-2} = \begin{pmatrix} 1 & -2 \\ 0 & 1 \end{pmatrix}$$
因此
$$A^{-1} = \begin{pmatrix} 1 & -1 & -1 & 2 \\ 0 & 1 & 0 & -1 \\ 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

□

§4.3 初等变换

在本书第一章中, 我们引入了求解线性方程组的三种初等变换. 上一章我们看到了初等变换在行列式计算中的重要作用. 本节我们将看到初等变换也是矩阵运算的重要工具. 利用矩阵乘法, 我们可以简洁地表示矩阵的初等变换. 因此, Gauss消元法可以用矩阵乘法表示, 进而可以帮助我们用矩阵的理解解决线性方程组的基本理论问题, 如解的存在性与唯一性等.

回忆一下三种初等变换

(1) 交换矩阵的两行;

(2) 将某行乘以一个非零常数;

(3) 将某行的常数倍加到另一行.

以上变换称为矩阵的**初等行变换**. 类似地, 对矩阵的列也可以实施三种对应的变换, 称为矩阵的**初等列变换**, 这六种变换统称为矩阵的**初等变换**. 我们仍然用行列式初等变换相同的记号表示矩阵的初等变换, 即用$r_i \leftrightarrow r_j$、λr_i、$\lambda r_i \to r_j$表示三个行变换, 用$c_i \leftrightarrow c_j$、λc_i、$\lambda c_i \to c_j$表示三个列变换. 每个初等变换对应一个初等矩阵. 下面分别予以介绍.

- 交换单位阵的第i, j行(或交换第i, j列), 得到

$$S_{ij} = \begin{array}{c} \begin{pmatrix} 1 & & & & & & \\ & \ddots & & & & & \\ & & 0 & & 1 & & \\ & & & \ddots & & & \\ & & 1 & & 0 & & \\ & & & & & \ddots & \\ & & & & & & 1 \end{pmatrix} \end{array} \begin{array}{l} \\ \\ \leftarrow \text{第}i\text{行} \\ \\ \leftarrow \text{第}j\text{行} \\ \\ \\ \end{array} \tag{4.17}$$

- 将单位阵的第i行(或第i列)乘以非零数λ, 得到

$$D_i(\lambda) = \begin{pmatrix} 1 & & & & & & \\ & \ddots & & & & & \\ & & 1 & & & & \\ & & & \lambda & & & \\ & & & & 1 & & \\ & & & & & \ddots & \\ & & & & & & 1 \end{pmatrix} \begin{array}{l} \\ \\ \\ \leftarrow \text{第}i\text{行} \\ \\ \\ \\ \end{array} \tag{4.18}$$

- 将单位阵的第j行的λ倍加到第i行(或将第i列的λ倍加到第j列), 得到

$$T_{ij}(\lambda)=\begin{array}{cl}\begin{pmatrix}1&&&&&&\\&\ddots&&&&&\\&&1&&\lambda&&\\&&&\ddots&&&\\&&&&1&&\\&&&&&\ddots&\\&&&&&&1\end{pmatrix}&\begin{array}{l}\\ \\ \leftarrow 第i行\\ \\ \leftarrow 第j行\\ \\ \\ \end{array}\end{array}\tag{4.19}$$

上述三类方阵称为**初等方阵**. 每一类初等方阵与一类初等变换相对应.

定理4.3.1. 对矩阵作初等行变换, 相当于在矩阵的左边乘上一个相应的初等方阵; 对矩阵作初等列变换, 相当于在矩阵的右边乘上一个相应的初等方阵.

证明. 我们仅证明行变换的情形. 设$m\times n$阶矩阵A的行为$\mathbf{a}_1,\mathbf{a}_2,\ldots\mathbf{a}_m$, 则有

$$S_{ij}A=\begin{pmatrix}1&&&&&&\\&\ddots&&&&&\\&&0&&1&&\\&&&\ddots&&&\\&&1&&0&&\\&&&&&\ddots&\\&&&&&&1\end{pmatrix}\begin{pmatrix}\mathbf{a}_1\\ \vdots\\ \mathbf{a}_i\\ \vdots\\ \mathbf{a}_j\\ \vdots\\ \mathbf{a}_m\end{pmatrix}=\begin{pmatrix}\mathbf{a}_1\\ \vdots\\ \mathbf{a}_j\\ \vdots\\ \mathbf{a}_i\\ \vdots\\ \mathbf{a}_m\end{pmatrix}$$

$$D_i(\lambda)A=\begin{pmatrix}1&&&&&&\\&\ddots&&&&&\\&&1&&&&\\&&&\lambda&&&\\&&&&1&&\\&&&&&\ddots&\\&&&&&&1\end{pmatrix}\begin{pmatrix}\mathbf{a}_1\\ \vdots\\ \vdots\\ \mathbf{a}_i\\ \vdots\\ \vdots\\ \mathbf{a}_m\end{pmatrix}=\begin{pmatrix}\mathbf{a}_1\\ \vdots\\ \vdots\\ \lambda\mathbf{a}_i\\ \vdots\\ \vdots\\ \mathbf{a}_m\end{pmatrix}$$

$$T_{ij}(\lambda)A=\begin{pmatrix}1&&&&&&\\&\ddots&&&&&\\&&1&&\lambda&&\\&&&\ddots&&&\\&&&&1&&\\&&&&&\ddots&\\&&&&&&1\end{pmatrix}\begin{pmatrix}\mathbf{a}_1\\ \vdots\\ \mathbf{a}_i\\ \vdots\\ \mathbf{a}_j\\ \vdots\\ \mathbf{a}_m\end{pmatrix}=\begin{pmatrix}\mathbf{a}_1\\ \vdots\\ \mathbf{a}_i+\lambda\mathbf{a}_j\\ \vdots\\ \mathbf{a}_j\\ \vdots\\ \mathbf{a}_m\end{pmatrix}.$$

列变换的情形可以类似证明,留作习题. □

定理4.3.2. *初等方阵具有下列性质:*

(1) S_{ij}*为对称方阵, 且*$S_{ij}^{-1}=S_{ij}$*;*

(2) $D_i(\lambda)$*为对角方阵, 且*$D_i(\lambda)^{-1}=D_i(\lambda^{-1})$*;*

(3) $T_{ij}(\lambda)$*为三角方阵, 且*$T_{ij}(\lambda)^{-1}=T_{ij}(-\lambda)$*.*

证明. 留作习题. □

第一章中求解线性方阵组的Gauss消元法本质上是对矩阵做行的初等变换,最终将矩阵化为阶梯形的标准形式. 利用矩阵语言表述就是,存在一系列初等方阵$P_1,P_2,\cdots,P_s$使得

$$P_s\cdots P_2P_1\,A=J,$$

其中J是阶梯形矩阵 (2.3).

特别地, 假设A是n阶可逆方阵, 则J是可逆上三角阵. 如果再假设我们做的初等变换只有第三类, 即用(k,k)位置的元素消去(j,k)位置的元素, $j>k$, $k=1,2,\ldots,n-1$, 则$P_s\cdots P_2P_1$是一个可逆下三角阵, 且对角元均为1, 记为L^{-1}. 同时记$J=U$, 则$A=LU$. 我们称$A=LU$为矩阵A的LU**分解**. 如果我们得到了矩阵A的LU分解, 解线性方程组$A\mathbf{x}=\mathbf{b}$就可以转化为求解两个简单的线性方程组: $L\mathbf{y}=\mathbf{b},U\mathbf{x}=\mathbf{y}$.

如果对矩阵同时实施行和列的初等变换,则矩阵可以化为更简单的形式. 我们有

定理4.3.3. *对任意矩阵*$A=(a_{ij})_{m\times n}$*, 存在一系列*m*阶初等方阵*$P_1,P_2,\ldots,P_s$*和*n*阶初等方阵*$Q_1,Q_2,\ldots,Q_t$*, 使得*

$$P_s\cdots P_2P_1\,A\,Q_1Q_2\cdots Q_t=\begin{pmatrix}I_r&O\\O&O\end{pmatrix}. \tag{4.20}$$

其中r为非负整数.

证明. 若$A=O$, 结论显然成立. 下设$A\neq O$, 不妨设$a_{pq}\neq 0$. 则将A的第一行与第p行互换, 第一列与第q列互换, 则可以将a_{pq}置换到$(1,1)$位置. 因此, 矩阵

$$B=(b_{ij})_{m\times n}=D_1(a_{pq}^{-1})\,S_{1p}\,A\,S_{1q}$$

满足$b_{11}=1$. 利用初等变换将B的第一行和第一列元素(除b_{11}外)都化为零, 即

$$T_{m1}(-b_{m1})\cdots T_{21}(-b_{21})\,B\,T_{12}(-b_{12})\cdots T_{1n}(-b_{1n})=\begin{pmatrix}1 & 0\\ 0 & C\end{pmatrix},$$

其中C为$(m-1)\times(n-1)$阶矩阵. 对矩阵C重复以上步骤, 最终可化为 (4.20)形式. □

由于初等方阵都是可逆方阵, 并且可逆方阵的乘积也是可逆方阵, 由上述定理有

定理4.3.4. 对任意矩阵$A=(a_{ij})_{m\times n}$, 存在m阶可逆方阵P和n阶可逆方阵Q, 使得

$$P\,A\,Q=\begin{pmatrix}I_r & O\\ O & O\end{pmatrix}, \tag{4.21}$$

其中非负整数r由A唯一决定.

证明. 只需证r的唯一性. 用反证法. 假设存在$r\neq s$ (不妨设$r<s$)及可逆方阵P_1,P_2,Q_1,Q_2满足

$$P_1AQ_1=\begin{pmatrix}I_r & O\\ O & O\end{pmatrix},\quad P_2AQ_2=\begin{pmatrix}I_s & O\\ O & O\end{pmatrix}.$$

令$P=P_2P_1^{-1}$, $Q=Q_1^{-1}Q_2$, 则

$$P\begin{pmatrix}I_r & O\\ O & O\end{pmatrix}Q=\begin{pmatrix}I_s & O\\ O & O\end{pmatrix}.$$

把P和Q分块

$$P=\begin{pmatrix}P_1 & P_2\\ P_3 & P_4\end{pmatrix},\quad Q=\begin{pmatrix}Q_1 & Q_2\\ Q_3 & Q_4\end{pmatrix},$$

其中P_1是$s\times r$矩阵, Q_1是$r\times s$矩阵. 于是$P_1Q_1=I_s$. 然而, 由于$r<s$, 根据例 4.2.14的结论, $\det(P_1Q_1)=0$. 这与$P_1Q_1=I_s$矛盾. □

上述定理中的非负整数r有特别的含义. 它对应于第一章中矩阵做初等变换后, (2.3)式中的梯形矩阵J的非零行数. 它是A的一个非常重要的量, 称为矩阵的秩. 我们将在下一节详细介绍它.

在(4.20)式中, 如果A为n阶可逆方阵, 则由于$P_s \cdots P_1\, A\, Q_1 \cdots Q_t$也可逆, 因此$r = n$. 从而

$$P_s \cdots P_1\, A\, Q_1 \cdots Q_t = I.$$

由此

$$A = P_1^{-1} \cdots P_s^{-1} Q_t^{-1} \cdots Q_1^{-1}. \tag{4.22}$$

由于初等矩阵的逆矩阵仍为初等矩阵, 因此有

定理4.3.5. *方阵A可逆的充要条件是, A可以分解为一系列(有限个)初等方阵的乘积.*

从(4.22)式还可以得到

$$\begin{aligned} Q_1 \cdots Q_t P_s \cdots P_1 A = I, \\ A Q_1 \cdots Q_t P_s \cdots P_1 = I. \end{aligned} \tag{4.23}$$

以上两式的含义是, 可以对可逆方阵A做一系列的初等行变换将A变为最简形式I, 也可以对可逆方阵A做一系列的初等列变换将A变为最简形式I. 利用这一结论, 可以给出计算可逆方阵的逆矩阵的有效方法–**初等变换法**.

设X是方阵的逆矩阵, 即$AX = I$. 再设$P_s \cdots P_1 A = I$, 其中P_i为初等矩阵. 则$X = P_s \cdots P_1$. 为计算矩阵A的逆, 我们将方阵A与I组合在一起形成一个n行、$2n$列的矩阵(A,I). 现对矩阵(A,I)实施一系列的初等行变换(对应的初等矩阵为$P_1,\ldots,P_s$). 则由于

$$P_s \cdots P_1 (A,I) = (P_s \cdots P_1 A, P_s \cdots P_1 I) = (I, A^{-1}), \tag{4.24}$$

这些初等变换将A简化为I, 与此同时将I化为A^{-1}, 由此通过一系列初等变换就可以求得矩阵的逆. 用类似的方法还可以求解矩阵方程$AX = B$. 下面我们给一个具体例子.

例4.3.1. 计算矩阵A的逆矩阵, 这里

$$A = \begin{pmatrix} 2 & 1 & 0 \\ 1 & 2 & 1 \\ 0 & 1 & 2 \end{pmatrix}.$$

解. 对矩阵(A,I)做一系列初等行变换, 使得A简化为I:

$$\begin{pmatrix}2&1&0&1&0&0\\1&2&1&0&1&0\\0&1&2&0&0&1\end{pmatrix}\xrightarrow{\frac{1}{2}r_1}\begin{pmatrix}1&\frac{1}{2}&0&\frac{1}{2}&0&0\\1&2&1&0&1&0\\0&1&2&0&0&1\end{pmatrix}\xrightarrow{-r_1\to r_2}\begin{pmatrix}1&\frac{1}{2}&0&\frac{1}{2}&0&0\\0&\frac{3}{2}&1&-\frac{1}{2}&1&0\\0&1&2&0&0&1\end{pmatrix}\xrightarrow{\frac{2}{3}r_2}$$

$$\begin{pmatrix}1&\frac{1}{2}&0&\frac{1}{2}&0&0\\0&1&\frac{3}{2}&-\frac{1}{3}&\frac{2}{3}&0\\0&1&2&0&0&1\end{pmatrix}\xrightarrow{-r_2\to r_3}\begin{pmatrix}1&\frac{1}{2}&0&\frac{1}{2}&0&0\\0&1&\frac{3}{2}&-\frac{1}{3}&\frac{2}{3}&0\\0&0&\frac{4}{3}&\frac{1}{3}&-\frac{2}{3}&1\end{pmatrix}\xrightarrow{\frac{3}{4}r_3}\begin{pmatrix}1&\frac{1}{2}&0&\frac{1}{2}&0&0\\0&1&\frac{3}{2}&-\frac{1}{3}&\frac{2}{3}&0\\0&0&1&\frac{1}{4}&-\frac{2}{4}&\frac{3}{4}\end{pmatrix}$$

$$\xrightarrow{-\frac{2}{3}r_3\to r_2}\begin{pmatrix}1&\frac{1}{2}&0&\frac{1}{2}&0&0\\0&1&0&-\frac{1}{2}&1&-\frac{1}{2}\\0&0&1&\frac{1}{4}&-\frac{2}{4}&\frac{3}{4}\end{pmatrix}\xrightarrow{-\frac{1}{2}r_2\to r_1}\begin{pmatrix}1&0&0&\frac{3}{4}&-\frac{2}{4}&\frac{1}{4}\\0&1&0&-\frac{1}{2}&1&-\frac{1}{2}\\0&0&1&\frac{1}{4}&-\frac{2}{4}&\frac{3}{4}\end{pmatrix}.$$

因此

$$A^{-1}=\begin{pmatrix}\frac{3}{4}&-\frac{2}{4}&\frac{1}{4}\\-\frac{1}{2}&1&-\frac{1}{2}\\\frac{1}{4}&-\frac{2}{4}&\frac{3}{4}\end{pmatrix}.$$ □

例4.3.2. 设$A=\begin{pmatrix}0&1&1&1\\1&0&1&1\\1&1&0&1\\1&1&1&0\end{pmatrix}$. 求$A^{-1}$.

解法一. 用初等变换.

$$\begin{pmatrix}0&1&1&1&1&0&0&0\\1&0&1&1&0&1&0&0\\1&1&0&1&0&0&1&0\\1&1&1&0&0&0&0&1\end{pmatrix}\xrightarrow{r_i\to r_1,i=2,3,4}\begin{pmatrix}3&3&3&3&1&1&1&1\\1&0&1&1&0&1&0&0\\1&1&0&1&0&0&1&0\\1&1&1&0&0&0&0&1\end{pmatrix}\xrightarrow{\frac{1}{3}r_1,-r_1\to r_i,i=2,3,4}$$

$$\begin{pmatrix}1&1&1&1&1/3&1/3&1/3&1/3\\0&-1&0&0&-1/3&2/3&-1/3&-1/3\\0&0&-1&0&-1/3&-1/3&2/3&-1/3\\0&0&0&-1&-1/3&-1/3&-1/3&2/3\end{pmatrix}\xrightarrow{r_i\to r_1,-r_i,i=2,3,4}$$

$$\begin{pmatrix}1&0&0&0&-2/3&1/3&1/3&1/3\\0&1&0&0&1/3&-2/3&1/3&1/3\\0&0&1&0&1/3&1/3&-2/3&1/3\\0&0&0&1&1/3&1/3&1/3&-2/3\end{pmatrix}.$$

因此 $A^{-1}=\begin{pmatrix}-2/3 & 1/3 & 1/3 & 1/3\\ 1/3 & -2/3 & 1/3 & 1/3\\ 1/3 & 1/3 & -2/3 & 1/3\\ 1/3 & 1/3 & 1/3 & -2/3\end{pmatrix}$. □

解法二. 由于

$$A+I_4=\begin{pmatrix}1&1&1&1\\1&1&1&1\\1&1&1&1\\1&1&1&1\end{pmatrix}$$

易验证, $(A+I_4)^2=4(A+I_4)$. 于是$A^2-2A-3I_4=0$, $A(A-2I_4)=3I_4$. 因此, $A^{-1}=\frac{1}{3}(A-2I_4)$. □

对于分块矩阵也可以施行与上面类似的“初等行(列)变换”, 这种变换也是通过对分块矩阵左(右)乘相应的“初等方阵”实现的.

例4.3.3. 给定分块矩阵

$$\begin{pmatrix}A & B\\ C & D\end{pmatrix},$$

其中A为可逆方阵. 可以直接验证

$$\begin{pmatrix}I & O\\ -CA^{-1} & I\end{pmatrix}\begin{pmatrix}A & B\\ C & D\end{pmatrix}=\begin{pmatrix}A & B\\ O & D-CA^{-1}B\end{pmatrix} \tag{4.25}$$

$$\begin{pmatrix}A & B\\ C & D\end{pmatrix}\begin{pmatrix}I & -A^{-1}B\\ O & I\end{pmatrix}=\begin{pmatrix}A & O\\ C & D-CA^{-1}B\end{pmatrix} \tag{4.26}$$

$$\begin{pmatrix}I & O\\ -CA^{-1} & I\end{pmatrix}\begin{pmatrix}A & B\\ C & D\end{pmatrix}\begin{pmatrix}I & -A^{-1}B\\ O & I\end{pmatrix}=\begin{pmatrix}A & O\\ O & D-CA^{-1}B\end{pmatrix} \tag{4.27}$$

其中准三角“初等方阵”

$$\begin{pmatrix}I & O\\ -CA^{-1} & I\end{pmatrix} \quad 和 \quad \begin{pmatrix}I & -A^{-1}B\\ O & I\end{pmatrix}$$

具有与初等方阵$T_{ij}(\lambda)$类似的性质, 并且可以验证

$$\begin{pmatrix}I & O\\ X & I\end{pmatrix}^{-1}=\begin{pmatrix}I & O\\ -X & I\end{pmatrix}, \quad \begin{pmatrix}I & Y\\ O & I\end{pmatrix}^{-1}=\begin{pmatrix}I & -Y\\ O & I\end{pmatrix}.$$

例4.3.4. 设A,D分别为m,n阶方阵, $B\in F^{n\times m}, C\in F^{m\times n}$, 且$A$可逆. 证明

$$\det\begin{pmatrix}A & C\\ B & D\end{pmatrix}=\det(A)\det(D-BA^{-1}C). \tag{4.28}$$

证明. 容易验证

$$\begin{pmatrix}I_m & 0\\ -BA^{-1} & I_n\end{pmatrix}\begin{pmatrix}A & C\\ B & D\end{pmatrix}=\begin{pmatrix}A & C\\ 0 & D-BA^{-1}C\end{pmatrix}$$

上式两边取行列式即得证.

特别地, 如果$m=n$且$AB=BA$, 则由于

$$A(D-BA^{-1}C)=AD-ABA^{-1}C=AD-BC,$$

故

$$\det\begin{pmatrix}A & C\\ B & D\end{pmatrix}=\det(AD-BC).$$

□

对于分块矩阵, 也可以使用初等变换法求逆.

例4.3.5. 设A,B,I为n阶方阵, 且$BA=0$. 计算分块矩阵

$$M=\begin{pmatrix}I & A\\ B & I\end{pmatrix}$$

的逆矩阵.

解.

$$\begin{pmatrix}I & A & I & O\\ B & I & O & I\end{pmatrix}\longrightarrow\begin{pmatrix}I & A & I & O\\ O & I & -B & I\end{pmatrix}\longrightarrow\begin{pmatrix}I & O & I+AB & -A\\ O & I & -B & I\end{pmatrix}.$$

因此

$$\begin{pmatrix}I & A\\ B & I\end{pmatrix}^{-1}=\begin{pmatrix}I+AB & -A\\ -B & I\end{pmatrix}.$$

□

利用定理 4.3.5, 我们还可以方便地证明行列式的基本性质–行列式的行列互换不改变行列式的值.

例4.3.6. 设A为n阶方阵. 则$\det(A^T)=\det(A)$.

证明. : 当A不可逆时, A^T也不可逆. 此时$\det(A^T)=\det(A)=0$. 下面假设A可逆. 由定理 4.3.5知, 存在初等方阵$P_1,P_2,\ldots,P_s$使得

$$A=P_1P_2\cdots P_s.$$

于是

$$A^T=P_s^T\cdots P_2^TP_1^T.$$

注意到$\det(P_i^T)=\det(P_i)$, $i=1,2,\ldots,s$, 故

$$\det(A^T)=\det(A)=\Pi_{i=1}^s\det(P_i).$$

总之, $\det(A^T)=\det(A)$. □

§4.4 矩阵的秩与相抵

在第一章我们讲到, 线性方程组及相应的矩阵可以通过初等行变换化成阶梯形式, 而阶梯形矩阵中非零行数是决定线性方程组的解的重要参数. 这个非零行数正是定理 4.3.4 中的非负整数r, 它是由矩阵A唯一确定的, 这就是矩阵A的秩. 矩阵秩的概念在矩阵运算及线性方程组的求解理论中有着重要的作用.

§4.4.1 矩阵秩与相抵的定义

回顾一下Gauss消元的过程, 我们实际上是对一个矩阵做一系列初等行变换. 变换后的矩阵对应的方程组与原矩阵对应的方程组是同解的. 此时, 我们说变换后的矩阵与原矩阵是等价(相抵)的. 据此, 我们引进以下概念.

定义4.4.1. 设A,B是$m\times n$阶矩阵, 如果存在一系列(有限个)初等变换将矩阵A化成矩阵B, 则称矩阵A和B **相抵**.

由于可逆方阵可以分解为一系列初等方阵的乘积, 因此有

定理4.4.1. *设A,B是$m\times n$阶矩阵. 则A和B相抵当且仅当, 存在m阶可逆方阵P和n阶可逆方阵Q使得*

$$B=PAQ.$$

容易验证, 矩阵的相抵关系满足性质

1. A与A本身相抵;

2. 若A与B相抵, 则B与A相抵;

3. 若A与B相抵, 且B与C相抵, 则A与C相抵.

满足上述性质的关系称为**等价关系**, 等价关系用$\sim$表示. 容易验证, 相抵是一种等价关系. 所有$m\times n$矩阵的全体依据相抵关系可以被分解为一些不相交的子集的并, 每个子集称为一个**相抵等价类**. 两个矩阵相抵当且仅当它们属于同一个相抵等价类. 关于矩阵相抵, 我们有以下基本问题:

1. 两个矩阵属于同一相抵等价类的条件是什么? 或者说, 设A,B均为$m\times n$矩阵, A与B相抵的充要条件是什么?

2. 在每个相抵等价类中, 最简单的代表元具有怎样的形式? 也就是说, 对每个$m\times n$矩阵, 与它相抵的最简矩阵(相抵标准形)是什么?

定理 4.3.4表明, 任意矩阵A相抵于准对角方阵$\begin{pmatrix} I_r & O \\ O & O \end{pmatrix}$, 这是与$A$相抵的最简矩阵.

定义4.4.2. 设A是$m\times n$矩阵, 公式 (4.21)中的矩阵$\begin{pmatrix} I_r & O \\ O & O \end{pmatrix}$称为$A$的**相抵标准形**. 整数$r$称为矩阵$A$的**秩**, 记为$\mathrm{rank}(A)$或$r(A)$. 若$r=m$, 则$A$称为是**行满秩**的; 若$r=n$, 则$A$称为是**列满秩**的. 特别地, 零矩阵的秩等于0.

接下来我们回答两个矩阵何时相抵的问题.

定理4.4.2. *设A,B是同阶矩阵, 则A与B相抵的充要条件是$\mathrm{rank}(A)=\mathrm{rank}(B)$.*

证明. 充分性. 设$\mathrm{rank}(A)=\mathrm{rank}(B)=r$, 则$A$与$B$都相抵于标准形$\begin{pmatrix} I_r & O \\ O & O \end{pmatrix}$, 从而$A$与$B$也相抵.

必要性. 设A与B相抵, 则A与B的标准形也相抵. 设$\mathrm{rank}(A)=r$, $\mathrm{rank}(B)=s$. 则$\begin{pmatrix} I_r & O \\ O & O \end{pmatrix}$与$\begin{pmatrix} I_s & O \\ O & O \end{pmatrix}$相抵. 由定理 4.3.4的证明知, $r=s$. □

上述定理表明, 矩阵根据它的秩分成不同的等价类. 对于$m\times n$阶矩阵全体$F^{m\times n}$, 由于$0\le r\le \min(m,n)$, 并且每个r对应唯一的一个相抵等价类, 因此$F^{m\times n}$总共有$\min(m,n)+1$个相抵等价类. 同一等价类的矩阵都相抵于$\begin{pmatrix} I_r & O \\ O & O \end{pmatrix}$.

定理4.4.3. *设A是$m\times n$矩阵, P,Q分别是m,n阶可逆方阵, 则$\mathrm{rank}(PAQ)=\mathrm{rank}(A)$.*

证明. 由于PAQ与A相抵, 故它们的秩相等. □

上述定理表明, 初等变换不改变矩阵的秩.

例4.4.1. 设A为n阶方阵. 证明: A可逆当且仅当$\mathrm{rank}(A)=n$.

证明. 设$\mathrm{rank}(A)=r$, 则存在n阶可逆方阵P,Q使得

$$PAQ=\begin{pmatrix} I_r & O \\ O & O \end{pmatrix}.$$

如果A可逆, 则PAQ可逆, 从而$r=n$. 反之, 如果$r=n$, 则$PAQ=I_n$, 从而$A=P^{-1}Q^{-1}$可逆. □

例4.4.2. 证明: 对任意矩阵A, $\mathrm{rank}(A^T)=\mathrm{rank}(A)$.

证明. 设$\mathrm{rank}(A)=r$, 由定理 4.3.4知, 存在可逆方阵P,Q使得$PAQ=\begin{pmatrix} I_r & O \\ O & O \end{pmatrix}$. 于是$Q^TA^TP^T=\begin{pmatrix} I_r & O \\ O & O \end{pmatrix}$. 从而, $\mathrm{rank}(A^T)=r=\mathrm{rank}(A)=r$. □

例4.4.3. 设A,B为任意矩阵. 证明

$$\mathrm{rank}\begin{pmatrix} A & O \\ O & B \end{pmatrix}=\mathrm{rank}(A)+\mathrm{rank}(B).$$

证明. 设$\mathrm{rank}(A)=r$, $\mathrm{rank}(B)=s$. 则存在可逆方阵P_1,Q_1与P_2,Q_2使得$P_1AQ_1=\begin{pmatrix} I_r & O \\ O & O \end{pmatrix}$, $P_2BQ_2=\begin{pmatrix} I_s & O \\ O & O \end{pmatrix}$. 于是

$$\begin{pmatrix} P_1 & O \\ O & P_2 \end{pmatrix}\begin{pmatrix} A & O \\ O & B \end{pmatrix}\begin{pmatrix} Q_1 & O \\ O & Q_2 \end{pmatrix}=\begin{pmatrix} I_r & O & O & O \\ O & O & O & O \\ O & O & I_s & O \\ O & O & O & O \end{pmatrix}.$$

通过简单的初等变换(行列互换)知, 上式等式右端的矩阵相抵于$\begin{pmatrix} I_{r+s} & O \\ O & O \end{pmatrix}$. 因此需证的等式成立. □

矩阵的秩与其非零子式有着紧密的联系. 为揭示这种联系, 我们先给出一个引理.

引理4.4.1. 设A是$m\times n$方阵, P,Q分别是m,n阶初等方阵. 若A的所有k阶子式都为零, 则PA与AQ的所有k阶子式也为零.

证明. 当$P=S_{ij}$或$P=D_i(\lambda)$时, 结论显然成立.当$P=T_{pq}(\lambda)$时, 设$B=PA$, 则

$$\det B\begin{pmatrix} i_1 i_2\cdots i_k \\ j_1 j_2\cdots j_k\end{pmatrix}=\begin{cases}\det A\binom{i_1 i_2\cdots i_k}{j_1 j_2\cdots j_k}, & p\notin\{i_1,i_2,\ldots,i_k\}\\ \det A\binom{i_1\cdots i_t\cdots i_k}{j_1\cdots j_t\cdots j_k}+\lambda\det A\binom{i_1\cdots q\cdots i_k}{j_1\cdots j_t\cdots j_k}, & p=i_s\ (1\le s\le k)\end{cases}.$$

因此若A的所有k阶子式都是为零, 则$B=PA$的所有k阶子式也都为零. 类似可证, QA的所有k阶子式也都为零. □

矩阵的秩与其非零子式之间的联系由下列定理给出.

定理4.4.4. 矩阵A的非零子式的最高阶数等于矩阵A的秩.

证明. 设矩阵A的秩为r, 它的非零子式的最大阶数为s. 将矩阵A 通过一系列初等变换变成其相抵标准形$B=\begin{pmatrix} I_r & O\\ O & O\end{pmatrix}$, 即存在可逆方阵$P,Q$使得$A=PBQ$, 则$A$至少有一个$r$阶子式非零. 否则, 由引理 4.4.1知, B 的所有r 阶子式都为零, 矛盾. 因此, $s\ge r$.

另一方面, 由于$B=\begin{pmatrix} I_r & O\\ O & O\end{pmatrix}$的任意$k(k\ge r+1)$阶子式为零, 由引理 4.4.1, $A=PBQ$的任意$k(k\ge r+1)$阶子式也为零, 从而$s\le r$. 因此, $r=s$. 定理证毕. □

由上述定理, 我们给出矩阵秩的一个等价定义.

定义4.4.3. 设矩阵A至少有一个r阶非零子式, 且A的所有$r+1$阶子式都为零, 则称A的秩为r.

利用上述定义可以方便地验证, 第一章中Gauss消元得到的标准型矩阵J ((2.3)式)的秩为

$$\operatorname{rank}(J)=\begin{cases} r & d_{r+1}=0\\ r+1 & d_{r+1}\neq 0.\end{cases}$$

例如, 当$d_{r+1}=0$时, J的第$1,2,\ldots,r$行, 第$1,j_2,\ldots,j_r$列的子式非零, 但J的任意$r+1$阶子式为零. 因此, $\operatorname{rank}(J)=r$. 同理, 当$d_{r+1}\neq 0$时, $\operatorname{rank}(J)=r+1$. 这正好表明, 标准型$J$中的非零行数恰好是$J$的秩, 也是矩阵$A$的秩.

这样, 我们实际上可以回答第一章中提出的如下问题:

1. (2.3)式中, r是由矩阵A唯一确定的, 它就是矩阵A的秩.

2. 线性方程组$A\mathbf{x}=\mathbf{b}$有解, 当且仅当$\operatorname{rank}(A)=\operatorname{rank}(J)=\operatorname{rank}(A,\mathbf{b})=r$.

§4.4.2 秩的计算

根据定理4.4.3, 初等变换不改变矩阵的秩. 因此, 对给定矩阵做行列初等变换将其化为简单方阵, 是计算矩阵秩的一般方法.

例4.4.4. 对于x的不同值, 讨论n阶方阵A的秩, 这里

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

解. 作初等变换

$$A \longrightarrow \begin{pmatrix} x+n-1 & 1 & \cdots & 1 \\ x+n-1 & x & \ddots & \vdots \\ \vdots & \ddots & \ddots & 1 \\ x+n-1 & \cdots & 1 & x \end{pmatrix} \longrightarrow \begin{pmatrix} x+n-1 & 1 & \cdots & 1 \\ & x-1 & & \\ & & \ddots & \\ & & & x-1 \end{pmatrix} = B.$$

当$x=1$时, B和$\mathrm{diag}(1,O_{n-1})$相抵. 当$x\neq 1$时, B和$(x+n-1,I_{n-1})$相抵.
综上可得

$$\mathrm{rank}(A)=\begin{cases} 1, & x=1; \\ n-1, & x=1-n; \\ n, & \text{其它情形}. \end{cases}$$ □

另一种计算秩的方法是计算矩阵的子式, 不过这种方法只对特殊矩阵(结构简单)有效.

例4.4.5. 计算n阶方阵A的秩, 这里

$$A=\begin{pmatrix} 1 & 1 & & \\ & 1 & \ddots & \\ & & \ddots & 1 \\ 1 & & & 1 \end{pmatrix}.$$

解. 由A的右上角的$n-1$阶子式非零知, $\mathrm{rank}(A)\geq n-1$. 又由$\det(A)=1-(-1)^n$, 故当n是奇数时, $\mathrm{rank}(A)=n$; 当n是偶数时, $\mathrm{rank}(A)=n-1$. □

§4.4.3 相抵标准形的应用*

利用矩阵的相抵标准形, 可以使一些矩阵问题得到简化, 把关于一般矩阵的问题转化为关于标准形矩阵的问题处理. 下面给出一些例子.

例4.4.6. 设$A \in F^{m\times n}$ 且rank$(A) = 1$. 证明:存在m维非零列向量**a**及n维非零行向量**b**使得$A = \mathbf{ab}$.

证明. : 由rank$(A) = 1$知存在m阶可逆方阵P及n阶可逆方阵Q使得

$$A = P\begin{pmatrix} 1 & O \\ O & O \end{pmatrix} Q = P\begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}\begin{pmatrix} 1 & 0 & \cdots & 0 \end{pmatrix} Q.$$

令

$$\mathbf{a} = P\begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \quad \mathbf{b} = \begin{pmatrix} 1 & 0 & \cdots & 0 \end{pmatrix} Q$$

即有$A = \mathbf{ab}$, 且**a**为m维非零列向量, **b**为n维非零行向量. □

例4.4.7. 每个秩为r的矩阵都可以写成r个秩是1的矩阵之和.

证明. 设rank$(A) = r$, 则存在可逆方阵P, Q满足$A = P\begin{pmatrix} I_r & O \\ O & O \end{pmatrix} Q$. 于是$A = PE_{11}Q + \ldots + PE_{rr}Q$, 其中$E_{ii}$是基本矩阵, 它的秩为1. 从而矩阵$PE_{ii}Q$的秩也都是1. □

例4.4.8. 证明: 对任意$m\times n$矩阵A, $n\times p$矩阵B, 都有

$$\operatorname{rank}(AB) \le \min\big(\operatorname{rank}(A), \operatorname{rank}(B)\big).$$

证明. 设rank$(A) = r$, 则存在可逆P, Q使得$A = P\begin{pmatrix} I_r & O \\ O & O \end{pmatrix} Q$. 于是$AB = P\begin{pmatrix} I_r & O \\ O & O \end{pmatrix} QB$. 将矩阵$QB$做如下分块$QB = \begin{pmatrix} B_1 \\ B_2 \end{pmatrix}$, 其中$B_1 \in F^{r\times p}$, $B_2 \in F^{(n-r)\times p}$. 则

$$\operatorname{rank}(AB) = \operatorname{rank}\left(\operatorname{diag}(I_r, O)\begin{pmatrix} B_1 \\ B_2 \end{pmatrix}\right) = \operatorname{rank}(B_1) \le \operatorname{rank}(B).$$

另一方面, $\operatorname{rank}(AB) = \operatorname{rank}((AB)^T) = \operatorname{rank}(B^TA^T) \le \operatorname{rank}(A^T) = \operatorname{rank}(A)$. 即$\operatorname{rank}(AB) \le \min\big(\operatorname{rank}(A), \operatorname{rank}(B)\big)$. □

例4.4.9. 设$A\in F^{m\times p}, B\in F^{n\times p}$. 证明: $\operatorname{rank}\begin{pmatrix}A\\B\end{pmatrix}\le\operatorname{rank}(A)+\operatorname{rank}(B)$.

证明. 由于
$$\begin{pmatrix}A\\B\end{pmatrix}=\begin{pmatrix}A&O\\O&B\end{pmatrix}\begin{pmatrix}I_p\\I_p\end{pmatrix},$$
因此
$$\operatorname{rank}\begin{pmatrix}A\\B\end{pmatrix}\le\operatorname{rank}\begin{pmatrix}A&O\\O&B\end{pmatrix}=\operatorname{rank}(A)+\operatorname{rank}(B).$$
□

例4.4.10. 设$A\in F^{m\times n}, B\in F^{n\times m}, m>n$. 证明: $\det(AB)=0$.

证明. 由$\operatorname{rank}(AB)\le\operatorname{rank}(A)\le n<m$, 即知$AB$不可逆, 从而 $\det(AB)=0$. □

例4.4.11. 设A为n阶方阵, I为同阶单位阵, 且$A^2=A$. 证明: $\operatorname{rank}(A)+\operatorname{rank}(I-A)=n$.

证明. 对以下分块矩阵做初等变换
$$\begin{pmatrix}A&O\\O&I-A\end{pmatrix}\longrightarrow\begin{pmatrix}A&O\\A&I-A\end{pmatrix}\longrightarrow\begin{pmatrix}A&A\\A&I\end{pmatrix}\longrightarrow\begin{pmatrix}A-A^2&O\\A&I\end{pmatrix}\longrightarrow\begin{pmatrix}O&O\\O&I\end{pmatrix}.$$
由于初等变换不改变矩阵的秩, 因此
$$\operatorname{rank}(A)+\operatorname{rank}(I-A)=\operatorname{rank}\begin{pmatrix}A&O\\O&I-A\end{pmatrix}=\operatorname{rank}(I)=n.$$
□

小知识: 矩阵发展史

从逻辑上, 矩阵应该早于行列式.但矩阵概念比行列式要晚. 1850年英国数学家Sylvester(1814–1897)在研究线性方程组的时候, 引入了矩阵的概念. 但这时矩阵只是线性方程组的一个速记形式.

1855年, 英国数学家凯莱(Caylay, 1821–1895)在研究线性变换下的不变量时, 引入了矩阵的概念. 1858年, 凯莱在《矩阵论的研究报告》中, 定义了两个矩阵相等、相加以及数与矩阵的数乘等运算和算律. 同时, 定义了零矩阵、单位阵等特殊矩阵, 更重要的是在该文中他给出了矩阵相乘、矩阵可逆等概念, 以及利用伴随阵求逆阵的方法, 证明了有关的性质, 如矩阵乘法有结合律, 没有交换律, 两个非零阵乘积可以为零矩阵等结论. 定义了转置阵、对称阵、反对称阵等概

念. 因此可以说, 凯莱是矩阵理论的奠基人. 凯莱还提出了矩阵的特征多项式的概念以及凯莱-哈密尔顿定理, 并验证了3×3阶矩阵的情况. 哈密尔顿证明了4×4阶矩阵的情况, 而一般情况下的证明是德国数学家弗罗贝尼乌斯(F.G.Frohenius, 1849一1917)于1898年给出的.

1854年时法国数学家埃尔米特(C.Hermite)使用了 “正交矩阵” 这一术语, 但他的正式定义直到1878年才由费罗贝尼乌斯发表. 1879年, 费罗贝尼乌斯(Frobenius)引入矩阵秩的概念, 他还引入了矩阵最小多项式. 迹的概念由Taber引入. 至此, 矩阵的体系基本上建立起来了.

1878年, 弗罗伯纽斯(Frobenius)在他的论文中引入了多项式矩阵的行列式因子、不变因子和初等因子等概念, 证明了两个多项式矩阵等价当且仅当它们有相同的不变因子和初等因子.

关于行列式的乘积定理($\det(AB)=\det(A)\det(B)$), Lagrange对三阶方阵给出了证明. 对一般情形, Binet叙述过但没有给出满意的证明, Cauchy给出了给定理的证明.

习题三

1. 完成定理4.2.1、定理4.2.2的证明.

2. 证明: 每个方阵都可以表示为一个对称阵与一个反对称阵之和的形式.

3. 设$A=\begin{pmatrix}-3&-1&-2\\1&3&4\end{pmatrix}$, $B=\begin{pmatrix}2&2&-2\\4&-1&-4\\4&3&-3\end{pmatrix}$, $C=\begin{pmatrix}1&1\\-4&1\\-1&-2\end{pmatrix}$, 计算$AB, BC, ABC, B^2, AC, CA$.

4. 计算$\begin{pmatrix}1&x&x^2&\cdots&x^n\\1&y&y^2&\cdots&y^n\\1&z&z^2&\cdots&z^n\end{pmatrix}\begin{pmatrix}a_0&b_0&c_0\\a_1&b_1&c_1\\a_2&b_2&c_2\\\vdots&\vdots&\vdots\\a_n&b_n&c_n\end{pmatrix}$.

5. 计算$\begin{pmatrix}x_1&x_2&\cdots&x_m\end{pmatrix}\begin{pmatrix}a_{11}&a_{12}&\cdots&a_{1n}\\a_{21}&a_{22}&\cdots&a_{2n}\\\vdots&\vdots&\cdots&\vdots\\a_{m1}&a_{m2}&\cdots&a_{mn}\end{pmatrix}\begin{pmatrix}y_1\\y_2\\\vdots\\y_n\end{pmatrix}$.

6. 举例求满足条件的2阶实方阵A.

(1) $A^2=\begin{pmatrix}0&1\\1&0\end{pmatrix}$ (2) $A^2=\begin{pmatrix}0&1\\-1&0\end{pmatrix}$ (3) $A^3=I$且$A\neq I$.

7. 计算下列方阵的k次方幂, k为正整数.

(1) $\begin{pmatrix}\cos\theta&\sin\theta\\-\sin\theta&\cos\theta\end{pmatrix}$ (2) $\begin{pmatrix}a&b\\-b&a\end{pmatrix}$ (3) $\begin{pmatrix}1&a&1&0\\0&1&0&1\\0&0&1&a\\0&0&0&1\end{pmatrix}$

(4) $\begin{pmatrix}1&1&&\\&1&\ddots&\\&&\ddots&1\\&&&1\end{pmatrix}_{n\times n}$ (5) $\begin{pmatrix}a_1b_1&a_1b_2&\cdots&a_1b_n\\a_2b_1&a_2b_2&\cdots&a_2b_n\\\vdots&\vdots&\cdots&\vdots\\a_nb_1&a_nb_2&\cdots&a_nb_n\end{pmatrix}$

8. 设A,B都是n阶对称方阵, 且$AB=BA$. 证明AB也是对称方阵.

9. 证明: 两个n阶上(下)三角方阵的乘积仍是上(下)三角.

10. 证明: 与任意n阶方阵都乘法可交换的方阵一定是数量阵.

11. 完成定理 4.2.5的证明.

12. 设$A_1,A_2,\ldots,A_k$都是n阶可逆方阵. 证明: $(A_1A_2\cdots A_k)^{-1}=A_k^{-1}\cdots A_2^{-1}A_1^{-1}$.

13. 设方阵A满足$A^k=O$, k为正整数. 证明: $I+A$可逆, 并求$(I+A)^{-1}$.

14. 设方阵A满足$I-2A-3A^2+4A^3+5A^4-6A^5=O$. 证明: $I-A$可逆. 并求$(I-A)^{-1}$.

15. 完成定理4.2.6, 定理4.2.7, 定理4.2.8的证明.

16. 证明: $(A_1A_2\cdots A_k)^T=A_k^T\cdots A_2^TA_1^T$(假设其中的矩阵乘法有意义).

17. 求所有满足$A^2=O, B^2=I, \overline{C}^TC=I$的2阶复方阵$A$、$B$、$C$.

18. 证明: 不存在n阶复方阵A,B满足$AB-BA=I_n$.

19. 证明: 可逆上(下)三角、准对角、对称、反对称方阵的逆矩阵仍然分别是上(下)三角、准对角、对称、反对称的.

20. 设$A=\begin{pmatrix}2&1&0\\1&2&1\\2&-1&1\end{pmatrix}$. 求$A$的LU分解, 并基于$A$的LU分解, 求解线性方程组$A\mathbf{x}=\mathbf{b}$, 这里$\mathbf{b}=(1,-2,1)^T$.

21. 求解下列矩阵方程

(1) $X\begin{pmatrix}-2 & -2 & 1\\ 1 & 4 & -3\\ 1 & -1 & 1\end{pmatrix}=\begin{pmatrix}2 & -1 & 3\\ -3 & 2 & 5\\ -1 & 5 & 4\end{pmatrix}$

(2) $\begin{pmatrix}0 & 1 & 0\\ 0 & 0 & 1\\ 1 & 0 & 0\end{pmatrix}X-X\begin{pmatrix}0 & 1 & 0\\ 0 & 0 & 1\\ 0 & 0 & 0\end{pmatrix}=\begin{pmatrix}1 & 2 & 3\\ 2 & 4 & 6\\ 3 & 6 & 9\end{pmatrix}$

22. 设A,B是n阶方阵, λ是数. 证明:

(1) $(\lambda A)^*=\lambda^{n-1}A^*$; (2) $(AB)^*=B^*A^*$; (3) $\det(A^*)=(\det(A))^{n-1}$.

23. 设方阵A的逆矩阵$A^{-1}=\begin{pmatrix}1 & 1 & 1\\ 1 & 2 & 1\\ 1 & 1 & 3\end{pmatrix}$, 求$A^*$.

24. 设方阵A的伴随矩阵$A^*=\begin{pmatrix}0 & 0 & 0 & 1\\ 0 & 0 & 2 & 0\\ 0 & -1 & 0 & 0\\ 4 & 0 & 0 & 0\end{pmatrix}$, 求$A$.

25. 设n阶方阵A的每行、每列元素之和都是0, 证明:A^*的所有元素都相等.

26. 设A是方阵, 证明: 线性方程组$A\mathbf{x}=\mathbf{0}$有非零解当且仅当$\det(A)=0$.

27. 完成定理4.3.1、定理4.3.2的证明.

28. 设$A=\begin{pmatrix}1 & -1 & 0 & 0\\ 0 & 1 & 0 & 0\\ 1 & 0 & 1 & -1\\ 0 & 1 & 0 & 1\end{pmatrix}$. 计算$A^n$及$A^{-1}$.

29. 设A,B为同阶方阵, 且满足$AB=BA$. 计算$\begin{pmatrix}A & B\\ O & A\end{pmatrix}^n$.

30. 设A为n阶方阵, 且满足$A^3=I_n$. 计算$\begin{pmatrix}O & I_n\\ A & O\end{pmatrix}^{2024}$.

31. 设$A\in F^{m\times m},B\in F^{n\times n},C\in F^{n\times m}$, 且$A,B$可逆. 求分块矩阵$\begin{pmatrix}A & O\\ C & B\end{pmatrix}$的逆.

32. 证明: 初等方阵具有以下性质

(1) $T_{ij}(\lambda)T_{ij}(\mu)=T_{ij}(\lambda+\mu)$;

(2) 当$i\neq q$且$j\neq p$时, $T_{ij}(\lambda)T_{pq}(\mu)=T_{pq}(\mu)T_{ij}(\lambda)$;

(3) $D_i(-1)S_{ij}=S_{ij}D_j(-1)=T_{ji}(1)T_{ij}(-1)T_{ji}(1)$.

33. 将单位矩阵做一系列(有限次)行与行的对换得到的矩阵称为**置换矩阵**. 证明:

(1) 将单位矩阵做一系列的列与列的对换得到的矩阵也为置换矩阵;

(2) 置换矩阵可以通过按某种次序排列单位阵的行(或列)得到;

(3) 置换矩阵的乘积仍为置换矩阵;

(4) 置换矩阵的逆、转置也是置换矩阵.

34. 计算下列矩阵的逆矩阵

$$(1)\quad \begin{pmatrix} 1 & 0 & 1 & -4 \\ -1 & -3 & -4 & -2 \\ 2 & -1 & 4 & 4 \\ 2 & 3 & -3 & 2 \end{pmatrix} \qquad (2)\quad \begin{pmatrix} 1 & 4 & -1 & -1 \\ 1 & -2 & -1 & 1 \\ -3 & 3 & -4 & -2 \\ 0 & 1 & -1 & -1 \end{pmatrix}$$

$$(3)\quad \begin{pmatrix} & & & 1 \\ & & 1 & 1 \\ & \iddots & \iddots & \vdots \\ 1 & 1 & \cdots & 1 \end{pmatrix} \qquad (4)\quad \begin{pmatrix} & & & A_1 \\ & & A_2 & \\ & \iddots & & \\ A_k & & & \end{pmatrix},\ A_i\text{可逆}$$

$$(5)\quad \begin{pmatrix} 1+a_1 & 1 & \cdots & 1 \\ 1 & 1+a_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 1 \\ 1 & \cdots & 1 & 1+a_n \end{pmatrix}$$

35. 本题给出定理 4.2.3的另一个证明. 设A,B均为n阶方阵.

(1) 若A或者B不可逆, 证明AB也不可逆.

(2) 设P为初等方阵, 证明$\det(PB)=\det(P)\det(B)$.

(3) 证明: $\det(AB)=\det(A)\det(B)$.

36. 设A是n阶可逆方阵, $\mathbf{a}$为n维列向量. 证明:

$$\det(A-\mathbf{a}\mathbf{a}^T)=(1-\mathbf{a}^TA^{-1}\mathbf{a})\det(A).$$

37. 设A是n阶可逆方阵, $\mathbf{a},\mathbf{b}$为n维列向量. 证明: $A+\mathbf{a}\mathbf{b}^T$可逆当且仅当$1+\mathbf{b}^TA^{-1}\mathbf{a}\neq 0$. 且$A+\mathbf{a}\mathbf{b}^T$可逆时,

$$(A+\mathbf{a}\mathbf{b}^T)^{-1}=A^{-1}-\frac{A^{-1}\mathbf{a}\mathbf{b}^TA^{-1}}{1+\mathbf{b}^TA^{-1}\mathbf{a}}.$$

38. 设$A\in F^{m\times m}$,$B\in F^{n\times n}$,$C\in F^{n\times m}$, 且A为对称可逆方阵. 证明: 存在可逆方阵P使得$P\begin{pmatrix}A & C^T\\ C & B\end{pmatrix}P^T$ 为准对角阵.

39. 设A是$m\times n$矩阵, B是$n\times m$矩阵. 证明:

$$\det(I_n-BA)=\det\begin{pmatrix}I_m & A\\ B & I_n\end{pmatrix}=\det(I_m-AB).$$

40. 计算下列矩阵的秩.

(1) $\begin{pmatrix}3 & 2 & -1 & 9\\ -2 & 1 & -4 & 2\\ -1 & -2 & 3 & -2\\ 3 & 2 & -1 & 9\end{pmatrix}$ (2) $\begin{pmatrix}0 & 4 & 8 & -5\\ -7 & 9 & -3 & 1\\ 1 & -7 & -11 & 7\\ -5 & 7 & -1 & 0\end{pmatrix}$ (3) $\begin{pmatrix}1 & 4 & 9 & 16\\ 4 & 9 & 16 & 25\\ 9 & 16 & 25 & 36\\ 16 & 25 & 36 & 49\end{pmatrix}$

41. 对于a,b的各种取值, 讨论实矩阵$\begin{pmatrix}1 & 2 & 3\\ 2 & 4 & a\\ 3 & b & 9\end{pmatrix}$的秩.

42. 设A为n阶方阵, 且$A^2=I$. 求方阵$\mathrm{diag}(I+A,I-A)$的相抵标准形.

43. 设A是n阶方阵, 证明: $\mathrm{rank}(A^*)=\begin{cases}n, & \mathrm{rank}(A)=n;\\ 1, & \mathrm{rank}(A)=n-1;\\ 0, & \mathrm{rank}(A)\leq n-2.\end{cases}$

44. 设$A\in F^{m\times n}$. 证明:线性方程组$A\mathbf{x}=\mathbf{0}$有非零解的充分必要条件是$\mathrm{rank}(A)<n$.

45. 证明下列关于秩的不等式.

(1) $\max\Big(\mathrm{rank}(A),\mathrm{rank}(B),\mathrm{rank}(A+B)\Big)\leq\mathrm{rank}\begin{pmatrix}A & B\end{pmatrix}$;

(2) $\mathrm{rank}\begin{pmatrix}A & B\end{pmatrix}\leq\mathrm{rank}(A)+\mathrm{rank}(B)$;

(3) $\begin{pmatrix}A & C\\ O & B\end{pmatrix}\geq\mathrm{rank}(A)+\mathrm{rank}(B)$.

其中A,B,C是使运算有意义的矩阵.

46* 设n阶方阵A满足$A^2=I$, 证明: $\operatorname{rank}(I+A)+\operatorname{rank}(I-A)=n$.

47* 设A是$m\times n$矩阵, B是$n\times p$矩阵. 证明: $\operatorname{rank}(AB)\geq\operatorname{rank}(A)+\operatorname{rank}(B)-n$.

48* 设A是$m\times n$矩阵, B是$n\times m$矩阵. 证明

$$m+\operatorname{rank}(I_n-BA)=\operatorname{rank}\begin{pmatrix}I_m & A\\ B & I_n\end{pmatrix}=n+\operatorname{rank}(I_m-AB).$$

第五章　线性空间

在第一章中, 我们熟悉了向量与三维几何空间. 本章中, 我们将向量与几何空间的概念加以推广到一般的抽象空间–线性空间, 并且要回答第一章提出的以下问题:

1. 确定线性方程组解存在、唯一的条件.
2. 确定线性方程组解的几何结构.

我们先研究一个具体的线性空间–n维数组空间的相关理论, 包括向量组的线性相关/无关性, 向量组的极大无关组与秩, 线性空间的基与维数等. 然后我们将有关结论推广到一般的抽象线性空间.

§5.1　数组空间

在第一章中, 我们引进了n-维数组向量及其加法与数乘运算. 下面我们做一个简单回顾.

一个n维**数组向量** $\mathbf{a}$是一个有序的n元数组

$$\mathbf{a}=(a_1,a_2,\ldots,a_n), \tag{5.1}$$

其中$a_i\in F(i=1,2,\ldots,n)$称为向量$\mathbf{a}$的第i个**分量**. 这里F是一个数域.

n维数组向量的加法与数乘运算定义如下

$$\begin{aligned}\mathbf{a}+\mathbf{b}&=(a_1+b_1,a_2+b_2,\ldots,a_n+b_n),\\ \lambda\mathbf{a}&=(\lambda a_1,\lambda a_2,\ldots,\lambda a_n).\end{aligned} \tag{5.2}$$

即按分量分别执行加法与数乘运算. 规定**零向量**$\mathbf{0}=(0,\ldots,0)$, **负向量** $-\mathbf{a}=(-a_1,\ldots,-a_n)$. 容易验证, 数组向量的加法与数乘运算满足三维几何向量相同的性质:

(1) **加法交换律**:　$\mathbf{a}+\mathbf{b}=\mathbf{b}+\mathbf{a}$,

(2) **加法结合律**:　$(\mathbf{a}+\mathbf{b})+\mathbf{c}=\mathbf{a}+(\mathbf{b}+\mathbf{c})$,

(3) **分配律**:　$\lambda(\mathbf{a}+\mathbf{b})=\lambda\mathbf{a}+\lambda\mathbf{b}$,

(4) **分配律**:　$(\lambda+\mu)\mathbf{a}=\lambda\mathbf{a}+\mu\mathbf{a}$.

(5) **零向量**: 称每个分量为零的向量$\mathbf{0}=(0,0,\ldots,0)$为零向量. 零向量满足: $\mathbf{a}=\mathbf{a}+\mathbf{0}=\mathbf{0}+\mathbf{a}$.

(6) **负向量**: 称满足$\mathbf{b}+\mathbf{a}=\mathbf{0}$的向量$\mathbf{b}$为$\mathbf{a}$的负向量, 简记为$-\mathbf{a}$. 显然, 负向量存在唯一. 实际上, 若$\mathbf{a}=(a_1,a_2,\ldots,a_n)$, 则$-\mathbf{a}=(-a_1,-a_2,\ldots,-a_n)$.

(7) **乘法结合律**: $\lambda(\mu\mathbf{a})=(\lambda\mu)\mathbf{a}$.

(8) $1\cdot\mathbf{a}=\mathbf{a}$.

定义了上述加法与数乘运算的n维数组向量全体就构成n维数组空间.

定义5.1.1. 设F是数域. 定义了线性运算 (5.2)的n维数组向量全体

$$F^n:=\{(a_1,a_2,\ldots,a_n)\,|\,a_i\in F\}$$

称为n维**数组空间**.

注6. 本书中, 通过建立单位直角坐标系我们将平面与三维几何空间分别与数组空间$\mathbb{R}^2$及$\mathbb{R}^3$等价起来. 在平面直角坐标系中, 点(x,y)与$\mathbb{R}^2$ 中的向量(x,y) 对应, 在空间直角坐标系中, 点(x,y,z)与$\mathbb{R}^3$中的向量(x,y,z)对应.

有了向量的加法与数乘, 就可以引进向量的一种重要的运算–线性组合.

定义5.1.2. 给定一组n维数组向量$\mathbf{a}_1,\mathbf{a}_2,\ldots,\mathbf{a}_m$及一组数$\lambda_1,\lambda_2,\ldots,\lambda_m$, 称和式

$$\lambda_1\mathbf{a}_1+\lambda_2\mathbf{a}_2+\ldots+\lambda_m\mathbf{a}_m$$

为$\mathbf{a}_1,\mathbf{a}_2,\ldots,\mathbf{a}_m$的**线性组合**. $\lambda_1,\lambda_2,\ldots,\lambda_m$称为**组合系数**. 如果$\mathbf{a}$可以写成$\mathbf{a}_1,\mathbf{a}_2,\ldots,\mathbf{a}_m$的线性组合, 则称$\mathbf{a}$可以用$\mathbf{a}_1,\mathbf{a}_2,\ldots,\mathbf{a}_m$ **线性表示**.

定义5.1.3. 给定两个向量组$S=\{\mathbf{a}_1,\ldots,\mathbf{a}_m\}$与$T=\{\mathbf{b}_1,\ldots,\mathbf{b}_l\}$, 如果$S$ 中的每一个向量$\mathbf{a}_i$都可以用向量组T线性表示, 则称向量组S可以由向量组T **线性表示**. 如果两个向量组S与T可以相互表示, 则称S与T **等价**. 记为$S\sim T$.

容易证明, 向量组等价具有下列性质:

1. **反身性**:$S\sim S$;
2. **对称性**: 若$S\sim T$, 则$T\sim S$;
3. **传递性**: 若$S\sim T$, $T\sim U$, 则$S\sim U$.

因此, 向量组的等价是一种等价关系. 向量组的等价有明确的几何意义, 我们稍后介绍.

例5.1.1. 设$\mathbf{a},\mathbf{b},\mathbf{c}$是三维几何空间中三个共面的向量, 且其中任意两个向量不平行. 则$\{\mathbf{a},\mathbf{b}\}\sim\{\mathbf{b},\mathbf{c}\}\sim\{\mathbf{a},\mathbf{c}\}$.

实际上, 由$\mathbf{a},\mathbf{b},\mathbf{c}$共面知, 存在不全为零的数$\lambda,\mu,\nu$ 使得

$$\lambda\mathbf{a}+\mu\mathbf{b}+\nu\mathbf{c}=\mathbf{0}.$$

下证λ,μ,ν均不为零. 若不然, 假设其中某个数为零, 比如$\lambda=0$, 则$\mu\mathbf{b}+\nu\mathbf{c}=\mathbf{0}$. 由于$\mu,\nu$不全为零, 则$\mathbf{b},\mathbf{c}$平行, 矛盾! 因此, λ,μ,ν均不为零. 于是

$$\mathbf{a}=-\frac{\mu}{\lambda}\mathbf{b}-\frac{\nu}{\lambda}\mathbf{c},\quad \mathbf{c}=-\frac{\lambda}{\nu}\mathbf{a}-\frac{\mu}{\nu}\mathbf{b}.$$

这样$\{\mathbf{a},\mathbf{b}\}\sim\{\mathbf{b},\mathbf{c}\}$. 同理, $\{\mathbf{a},\mathbf{b}\}\sim\{\mathbf{a},\mathbf{c}\}$.

本章将研究由一组给定向量的线性组合全体构成的集合性质与结构, 为此考察集合

$$V:=\{\lambda_1\mathbf{a}_1+\lambda_2\mathbf{a}_2+\ldots+\lambda_m\mathbf{a}_m\,|\,\lambda_i\in F, i=1,2,\ldots,m\}$$

容易验证, 集合V具有以下性质(见本章习题2): 对任意$\mathbf{b}_1,\ldots\mathbf{b}_l\in V$及$\mu_1,\ldots,\mu_l\in F$, $\sum_{i=1}^{l}\mu_i\mathbf{b}_i\in V$. 也就是说, 集合$V$对线性运算是封闭的. 我们称具有如此性质的集合为一个子空间.

定义5.1.4. 设$V\subset F^n$为非空向量集合, 它满足

1. 若$\mathbf{a},\mathbf{b}\in V$, 则$\mathbf{a}+\mathbf{b}\in V$;
2. 如$\mathbf{a}\in V$, $\lambda\in F$, 则$\lambda\mathbf{a}\in V$.

则称V为F^n的**子空间**.

设V为子空间, 则对任意 $\mathbf{a}_1,\ldots,\mathbf{a}_m\in V$ 及 $\lambda_1,\ldots,\lambda_m\in F$, 都有 $\sum\limits_{i=1}^{m}\lambda_i\mathbf{a}_i\in V$. 显然, $V=\{\mathbf{0}\}$以及$V=F^n$都是F^n的子空间, 我们称之为**平凡子空间**. 一个非平凡子空间的例子是由一组向量的线性组合全体构成的集合. 由前面的讨论我们有

定理5.1.1. 设$\mathbf{a}_1,\mathbf{a}_2,\ldots,\mathbf{a}_m\in F^n$是一组给定的$n$维数组向量. 则集合

$$\langle\mathbf{a}_1,\mathbf{a}_2,\ldots,\mathbf{a}_m\rangle:=\{\lambda_1\mathbf{a}_1+\lambda_2\mathbf{a}_2+\ldots+\lambda_m\mathbf{a}_m\,|\,\lambda_i\in F, i=1,2,\ldots,m\}\tag{5.3}$$

是F^n的子空间, 称为由向量组$\mathbf{a}_1,\mathbf{a}_2,\ldots,\mathbf{a}_m$ **生成的子空间**.

利用生成子空间, 向量组等价的概念可以表述为

定理5.1.2. 两组向量组S与T等价当且仅当$\langle S\rangle=\langle T\rangle$.

证明. 必要性. 由于向量组S可以由T线性表示, S中每个向量都可以用T线性表示. 因此, $S\subset\langle T\rangle$. 进而, S中任意若干元素的线性组合均属于$\langle T\rangle$, 即$\langle S\rangle\subset\langle T\rangle$. 类似可证, $\langle T\rangle\subset\langle S\rangle$. 因而, $\langle S\rangle=\langle T\rangle$.

充分性. 若$\langle S\rangle=\langle T\rangle$, 则$S$中每个元素都属于$\langle T\rangle$, 即$S$中每个元素都可以用$T$线性表示. 同理可证, T中每个元素可以用S线性表示. 因此, 向量组S与T等价. □

上述定理表明, 两组向量等价当且仅当它们生成的子空间相同, 这就是向量组等价的几何意义. 下面我们来看一些子空间的例子.

例5.1.2. 设$\mathbf{a},\mathbf{b},\mathbf{c}$是三维几何空间中三个不共面的向量. 由三维空间向量的基本定理, 任何向量可以唯一地表示成三个不共面的向量的线性组合, 因此生成子空间$\langle\mathbf{a},\mathbf{b},\mathbf{c}\rangle$正好是整个三维几何空间$\mathbb{R}^3$. 而$\langle\mathbf{a}\rangle$, $\langle\mathbf{b}\rangle$, $\langle\mathbf{c}\rangle$, $\langle\mathbf{a},\mathbf{b}\rangle$, $\langle\mathbf{a},\mathbf{c}\rangle$, $\langle\mathbf{b},\mathbf{c}\rangle$ 等都是$\mathbb{R}^3$的生成子空间. 从几何上, $\langle\mathbf{a}\rangle$表示由$\mathbf{a}$确定的直线, $\langle\mathbf{a},\mathbf{b}\rangle$ 表示$\mathbf{a},\mathbf{b}$ 确定的平面.

例5.1.3. 设矩阵$A\in F^{m\times n}$. 由A的行向量生成F^n的子空间, 称为A的**行空间**, 记为$R(A)$. 由A的列向量生成F^m的子空间, 称为A 的**列空间**, 记为$C(A)$. 齐次线性方程组$A\mathbf{x}=\mathbf{0}$的解集全体构成一个线性子空间, 称为A的**零空间**, 记为$N(A)$. 例如, n阶单位向量生成的行空间与列空间均是F^n.

例5.1.4. 用$\mathbf{a}_j$表示m维列向量$(a_{1j},a_{2j},\ldots,a_{mj})^T$, $\mathbf{b}$表示m维列向量$(b_1,b_2,\ldots,b_m)^T$, 则线性方程组(2.1) 可以表示为向量形式

$$x_1\mathbf{a}_1+x_2\mathbf{a}_2+\cdots+x_n\mathbf{a}_n=\mathbf{b}. \tag{5.4}$$

即$\mathbf{b}$是$\mathbf{a}_1,\mathbf{a}_2,\ldots,\mathbf{a}_n$的线性组合, 组合系数就是线性方程组的解. 因此, 线性方程组有解等价于$\mathbf{b}\in\langle\mathbf{a}_1,\mathbf{a}_2,\ldots,\mathbf{a}_n\rangle$.

例5.1.5. 由本书第一章的理论, 齐次线性方程组的通解可以写成一些向量的线性组合形式

$$\mathbf{x}=t_1\alpha_1+t_2\alpha_2+\cdots+t_{n-r}\alpha_{n-r},$$

其中$\alpha_1,\alpha_2,\ldots,\alpha_{n-r}\in F^n$, $t_1,t_2,\ldots,t_{n-r}$为参数. 因此, 齐次线性方程组解的全体构成一个子空间$\langle\alpha_1,\alpha_2,\ldots,\alpha_{n-r}\rangle$.

由上面的例子我们看到, 要研究线性方程组的解的属性(存在性、唯一性等)以及解的结构, 就要研究向量的线性组合及其性质. 为此, 先研究向量之间的线性关系.

§5.2 线性相关与线性无关

我们先回顾一下三维向量之间的线性关系. 三维几何空间的两个向量$\mathbf{a},\mathbf{b}$共线, 当且仅当其中一个向量是另一个向量的常数倍, 如$\mathbf{a}=\lambda\mathbf{b}$, 这里$\lambda$为常数. 此时, 我们称向量$\mathbf{a},\mathbf{b}$ **线性相关**. 反之, $\mathbf{a},\mathbf{b}$不共线, 当且仅当$\mathbf{a}$不是$\mathbf{b}$的常数倍, 且$\mathbf{b}$不是$\mathbf{a}$的常数倍. 此时, 我们称$\mathbf{a},\mathbf{b}$ **线性无关**. 类似地, 三个向量$\mathbf{a},\mathbf{b},\mathbf{c}$共面, 当且仅当其中一个向量是另两个向量的线性组合, 如$\mathbf{c}=\lambda\mathbf{a}+\mu\mathbf{b}$, 这里$\lambda,\mu$为常数, 此时我们称向量$\mathbf{a},\mathbf{b},\mathbf{c}$ **线性相关**. 反之, $\mathbf{a},\mathbf{b},\mathbf{c}$不共面, 当且仅当$\mathbf{a},\mathbf{b},\mathbf{c}$中任一向量都不是另两个向量的线性组合. 此时, 我们称向量$\mathbf{a},\mathbf{b},\mathbf{c}$ **线性无关**.

接下来, 我们讨论n维数组向量的线性相关性与线性无关性. 我们从考察线性方程组的方程之间的线性关系开始.

给定线性方程组

$$\begin{cases} l_1 := a_{11}x_1+a_{12}x_2+\cdots+a_{1n}x_n-b_1=0 \\ l_2 := a_{21}x_1+a_{22}x_2+\cdots+a_{2n}x_n-b_2=0 \\ \qquad\qquad\vdots \\ l_m := a_{m1}x_1+a_{m2}x_2+\cdots+a_{mn}x_n-b_m=0 \end{cases} \tag{5.5}$$

在本书第一章介绍求解线性方程组的Gauss消元法时, 我们知道了线性方程也可以进行线性运算(求和及数乘), 从而可以定义线性方程的线性组合. 例如,

$$\lambda_1 l_1+\lambda_2 l_2 := (\lambda_1 a_{11}+\lambda_2 a_{21})x_1+\ldots+(\lambda_1 a_{1n}+\lambda_2 a_{2n})x_n-(\lambda_1 b_1+\lambda_2 b_2)=0.$$

如果方程l_i是其余方程的线性组合, 即$l_i=\sum\limits_{j\neq i}\lambda_j l_j$($\lambda_j$为常数), 则去掉方程$l_i$后, 原方程组与去掉$l_i$后的方程组等价, 此时称方程组$l_1,l_2,\cdots,l_m$是**线性相关**的. 反之, 如果任何一个方程都不能写成其它方程的线性组合, 则称方程组$l_1,l_2,\cdots,l_m$**线性无关**.

例5.2.1. 线性方程组

$$\begin{cases} x+y+z=1 \\ 2x+y+5z=2 \\ x-3y+13z=1 \end{cases}$$

是否线性相关?

解. 令$l_1=x+y+z-1$, $l_2=2x+y+5z-2$, $l_3=x-3y+13z-1$. 假设$l_3=\lambda_1 l_1+\lambda_2 l_2$,

则

$$\begin{cases}\lambda_1+2\lambda_2=1\\ \lambda_1+\lambda_2=-3\\ \lambda_1+5\lambda_2=13\\ \lambda_1+2\lambda_2=1\end{cases}$$

解得$\lambda_1=-7,\lambda_2=4$. 于是$l_3=-7l_1+4l_2$, 即原方程组线性相关. □

线性方程与数组向量之间存在一一对应关系. 实际上, 方程组(5.5)中的方程l_i与一个$n+1$维的数组向量$\mathbf{a}_i=(a_{i1},a_{i2},\ldots,a_{in},-b_i)$一一对应, $i=1,2,\ldots,m$. 因此, 方程组(5.5)与一组数组向量$\mathbf{a}_1,\mathbf{a}_2,\ldots,\mathbf{a}_m$一一对应, 并且线性方程之间的线性运算也与向量组之间的线性运算一一对应. 例如, $\lambda_1l_1+\lambda_2l_2$与$\lambda_1\mathbf{a}_1+\lambda_2\mathbf{a}_2$对应. 由此, 可以类似于线性方程组的线性相关性与无关性, 定义数组向量组的线性相关性与无关性.

定义5.2.1. 设$\mathbf{a}_1,\mathbf{a}_2,\cdots,\mathbf{a}_m\in F^n$, $m\geq 2$. 如果其中某个向量能用其它向量线性表示, 即存在$\mathbf{a}_i$ 及$\lambda_j\in F(j\neq i)$使得$\mathbf{a}_i=\sum\limits_{j\neq i}\lambda_j\mathbf{a}_j$, 则称$\mathbf{a}_1,\cdots,\mathbf{a}_m$ **线性相关**. 否则, 称它们**线性无关**. 一个向量$\mathbf{a}_1$称为**线性相关**如果$\mathbf{a}_1=\mathbf{0}$, 否则称它**线性无关**.

本章中, 我们规定向量组的个数都是有限的, 以后不再说明.

例5.2.2. 重新考察例5.2.1 线性方程l_1, l_2与l_3分别与向量$\mathbf{a}_1=(1,1,1,1)$, $\mathbf{a}_2=(2,1,5,2)$, $\mathbf{a}_3=(1,-3,13,1)$ 对应. 由于$\mathbf{a}_3=-7\mathbf{a}_1+4\mathbf{a}_2$, $\mathbf{a}_1$, $\mathbf{a}_2$, $\mathbf{a}_3$ 线性相关, 因而l_1,l_2,l_3也线性相关.

利用生成子空间的概念, 向量组的线性相关性可以表述如下

定理5.2.1. *给定向量组$\mathbf{a}_1,\beta a_2,\cdots,\mathbf{a}_m\in F^n$, $m\geq 2$. 则$\mathbf{a}_1,\mathbf{a}_2,\cdots,\mathbf{a}_m$线性相关当且仅当, 存在$i\in\{1,2,\ldots,m\}$满足*

$$\langle\mathbf{a}_1,\mathbf{a}_2,\ldots,\mathbf{a}_m\rangle=\langle\mathbf{a}_1,\ldots,\mathbf{a}_{i-1},\mathbf{a}_{i+1},\ldots,\mathbf{a}_m\rangle. \tag{5.6}$$

证明. 由定义 5.2.1, $\mathbf{a}_1,\mathbf{a}_2,\cdots,\mathbf{a}_m$线性相关当且仅当存在$i\in\{1,2,\ldots,m\}$使得, $\mathbf{a}_i=\sum_{j=1,j\neq i}^m\lambda_j\mathbf{a}_j$, 或等价地

$$\mathbf{a}_i\in\langle\mathbf{a}_1,\ldots,\mathbf{a}_{i-1},\mathbf{a}_{i+1},\ldots,\mathbf{a}_m\rangle.$$

上式显然与(5.6)等价. □

上述定理表明, 向量组$\mathbf{a}_1,\mathbf{a}_2,\cdots,\mathbf{a}_m$线性相关等价于, 其中某个向量对生成子空间 $\langle \mathbf{a}_1,\mathbf{a}_2,\ldots,\mathbf{a}_m \rangle$ 而言是多余的.

用记号$\Leftrightarrow$表示“等价”或“当且仅当”. 由于

$$\begin{aligned}
&\text{存在常数}\lambda_j\text{使得} \quad \mathbf{a}_i = \sum_{j\neq i}\lambda_j\mathbf{a}_j \Leftrightarrow \\
&\text{存在常数}\lambda_j\text{使得} \quad \sum_{j=1}^{m}\lambda_j\mathbf{a}_j = \mathbf{0},\ \lambda_i = -1 \Leftrightarrow \\
&\text{存在常数}\mu_j\text{使得} \quad \sum_{j=1}^{m}\mu_j\mathbf{a}_j = \mathbf{0},\ \text{某个}\mu_i \neq 0
\end{aligned}$$

因此有

定理5.2.2. 设$\mathbf{a}_1,\mathbf{a}_2,\cdots,\mathbf{a}_m \in F^n$, 则$\mathbf{a}_1,\mathbf{a}_2,\cdots,\mathbf{a}_m$线性相关的充要条件是存在不全为零的常数$\lambda_1,\lambda_2,\cdots,\lambda_m$, 使得

$$\sum_{i=1}^{m}\lambda_i\mathbf{a}_i = \lambda_1\mathbf{a}_1+\lambda_2\mathbf{a}_2+\ldots+\lambda_m\mathbf{a}_m = \mathbf{0}. \tag{5.7}$$

例5.2.3. 包含零向量的任何向量组一定线性相关.

证明. 实际上, 如果向量组$\mathbf{a}_1,\mathbf{a}_2,\cdots,\mathbf{a}_m$含零向量, 不妨设$\mathbf{a}_1=\mathbf{0}$. 则

$$1\,\mathbf{a}_1+0\,\mathbf{a}_2+\ldots+0\,\mathbf{a}_m = \mathbf{0}.$$

因此, 向量组$\mathbf{a}_1,\mathbf{a}_2,\cdots,\mathbf{a}_m$线性相关. □

上述例子的一般地情形是如下结论.

定理5.2.3. 设向量组$\tilde{S}$是向量组S的一个子集. 那么, 如果$\tilde{S}$线性相关, 则S也线性相关; 反之, 如果S线性无关, 则$\tilde{S}$也线性无关.

证明. 不妨设, $\tilde{S}=\{\mathbf{a}_1,\mathbf{a}_2,\ldots,\mathbf{a}_k\}$, $S=\{\mathbf{a}_1,\mathbf{a}_2,\ldots,\mathbf{a}_m\}$ $(k<m)$. 设$\tilde{S}$线性相关, 则存在不全为零的常数$\lambda_1,\lambda_2,\ldots,\lambda_k$使得

$$\lambda_1\mathbf{a}_1+\lambda_2\mathbf{a}_2+\ldots+\lambda_k\mathbf{a}_k = \mathbf{0}.$$

则

$$\lambda_1\mathbf{a}_1+\ldots+\lambda_k\mathbf{a}_k+0\cdot\mathbf{a}_{k+1}+\ldots+0\cdot\mathbf{a}_m = \mathbf{0}.$$

这表明, S也线性相关. 反之, 如果S线性无关, 则$\tilde{S}$必线性无关.

如果用生成子空间的语言来说大致意思是, 如果$\tilde{S}$中有多余的向量, 则S中必有多余的向量.

□

下面, 我们给出向量组线性相关(无关)的进一步的判别方法.

定理5.2.4. 设$\mathbf{a}_i=(a_{i1},a_{i2},\ldots,a_{in})\in F^n$, $i=1,2,\ldots,m$. 用A表示以$\mathbf{a}_1,\mathbf{a}_2,\cdots,\mathbf{a}_m$为行构成的$m\times n$阶矩阵. 则$\mathbf{a}_1,\mathbf{a}_2,\cdots,\mathbf{a}_m$线性相关当且仅当, 关于$\lambda_1,\lambda_2,\ldots,\lambda_m$的齐次线性方程组

$$A^T\begin{pmatrix}\lambda_1\\ \vdots\\ \lambda_m\end{pmatrix}=\mathbf{0} \tag{5.8}$$

有非零解.

证明. 由定理 5.2.2, 向量组$\mathbf{a}_1,\mathbf{a}_2,\cdots,\mathbf{a}_m$线性相关, 当且仅当存在不全为零的常数$\lambda_1,\lambda_2,\ldots,\lambda_m$使得

$$\lambda_1\mathbf{a}_1+\lambda_2\mathbf{a}_2+\cdots+\lambda_m\mathbf{a}_m=\mathbf{0}$$

亦即线性方程组(5.8)有非零解. □

由上述定理, 下面的推论是显然的.

推论5.2.1. 设$\mathbf{a}_1,\mathbf{a}_2,\cdots,\mathbf{a}_m\in F^n$是一组数组向量, 则

1. 若$m>n$, 则 $\mathbf{a}_1,\mathbf{a}_2,\cdots,\mathbf{a}_m$必然线性相关.
2. 若$m=n$, 则 $\mathbf{a}_1,\mathbf{a}_2,\cdots,\mathbf{a}_m$线性相关当且仅当$\det(A)=0$.

总结起来, 下列说法等价

1. n维向量组$\mathbf{a}_1,\mathbf{a}_2,\cdots,\mathbf{a}_m$ 线性相关;
2. 某个 $\mathbf{a}_i$是其它向量的线性组合, 这里$m\geq 2$;
3. 关于$\lambda_1,\lambda_2,\ldots,\lambda_m$的线性方程组 $\lambda_1\mathbf{a}_1+\lambda_2\mathbf{a}_2+\cdots+\lambda_m\mathbf{a}_m=\mathbf{0}$ 有非零解;

对应地, 可以得到线性无关的判别条件, 在此略去. 下面, 我们给出一些实例.

例5.2.4. 判断下列向量组的线性相关(无关)性.

(1) $\mathbf{e}_1,\mathbf{e}_2,\cdots,\mathbf{e}_n$, 其中$\mathbf{e}_i\in F^n$是$n$维单位坐标向量, $i=1,2,\ldots,n$.

(2) $\mathbf{a}_1=(1,0,\cdots,0)$, $\mathbf{a}_2=(1,1,0,\cdots)$, $\cdots$, $\mathbf{a}_n=(1,\cdots,1)$.

(3) $\mathbf{a}_1+\mathbf{a}_2$, $\mathbf{a}_2+\mathbf{a}_3$, $\mathbf{a}_3+\mathbf{a}_1$, 其中$\mathbf{a}_1,\mathbf{a}_2,\mathbf{a}_3\in F^n$线性无关.

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

解. (1) 由于以$\mathbf{e}_1,\mathbf{e}_2,\cdots,\mathbf{e}_n$为行的矩阵为单位阵, 其行列式显然非零. 因此, $\mathbf{e}_1,\mathbf{e}_2,\cdots,\mathbf{e}_n$线性无关.

(2) 因为以$\mathbf{a}_1,\mathbf{a}_2,\ldots,\mathbf{a}_n$构成矩阵的行列式为1, 所以向量组$\mathbf{a}_1,\mathbf{a}_2,\cdots,\mathbf{a}_n$线性无关.

(3) 由

$$\lambda_1(\mathbf{a}_1+\mathbf{a}_2)+\lambda_2(\mathbf{a}_2+\mathbf{a}_3)+\lambda_3(\mathbf{a}_3+\mathbf{a}_1)=\mathbf{0}$$

得

$$(\lambda_1+\lambda_3)\mathbf{a}_1+(\lambda_1+\lambda_2)\mathbf{a}_2+(\lambda_2+\lambda_3)\mathbf{a}_3=\mathbf{0}.$$

由于$\mathbf{a}_1,\mathbf{a}_2,\mathbf{a}_3$线性无关, 因此

$$\lambda_1+\lambda_3=0,\quad \lambda_1+\lambda_2=0,\quad \lambda_2+\lambda_3=0.$$

故$\lambda_1=\lambda_2=\lambda_3=0$. 从而, $\mathbf{a}_1+\mathbf{a}_2$, $\mathbf{a}_2+\mathbf{a}_3$, $\mathbf{a}_3+\mathbf{a}_1$线性无关.

(4) 解线性方程组

$$\lambda_1\mathbf{a}_1+\lambda_2\mathbf{a}_2+\lambda_3\mathbf{a}_3+\lambda_4\mathbf{a}_4=\mathbf{0}$$

得一组解$(\lambda_1,\lambda_2,\lambda_3,\lambda_4)=(2,1,-1,1)$, 即$2\mathbf{a}_1+\mathbf{a}_2-\mathbf{a}_3-\mathbf{a}_4=\mathbf{0}$. 因此, $\mathbf{a}_1,\mathbf{a}_2,\mathbf{a}_3,\mathbf{a}_4$线性相关.

□

例5.2.5. 设$\mathbf{a}_1,\mathbf{a}_2\in\mathbf{R}^3$是三维空间向量, $\mathbf{b}_1,\mathbf{b}_2$ 是$\mathbf{a}_1,\mathbf{a}_2$在$oxy$平面上的投影向量. 如果$\mathbf{b}_1,\mathbf{b}_2$线性无关, 则$\mathbf{a}_1,\mathbf{a}_2$也线性无关. 反之不然.

证明. 请读者完成. □

上述例子中的结果可以推广到一般情形.

定理5.2.5. *设$\mathbf{a}_i=(a_{i1},a_{i2},\ldots,a_{ir})\in F^r$, $i=1,2,\ldots,m$. 它们的加长向量组为$\mathbf{b}_i=(a_{i1},\ldots,a_{ir},\ldots,a_{in})\in F^n$ $(n>r)$, $i=1,2,\ldots,m$. 则有*

1. *若$\mathbf{a}_1,\mathbf{a}_2,\cdots,\mathbf{a}_m$线性无关, 则$\mathbf{b}_1,\mathbf{b}_2,\cdots,\mathbf{b}_m$也线性无关;*

2. *若$\mathbf{b}_1,\mathbf{b}_2,\cdots,\mathbf{b}_m$线性相关, 则$\mathbf{a}_1,\mathbf{a}_2,\cdots,\mathbf{a}_m$也线性相关.*

证明. 只用证明第2条. 设$\mathbf{b}_1,\mathbf{b}_2,\cdots,\mathbf{b}_m$线性相关, 则存在不全为零的数$\lambda_i$, $i=1,\ldots,m$ 使得$\lambda_1\mathbf{b}_1+\lambda_2\mathbf{b}_2+\ldots+\lambda_m\mathbf{b}_m=\mathbf{0}$, 即

$$\begin{cases} a_{11}\lambda_1+a_{21}\lambda_2+\ldots+a_{m1}\lambda_m=0 \\ a_{12}\lambda_1+a_{22}\lambda_2+\ldots+a_{m2}\lambda_m=0 \\ \cdots\cdots \\ a_{1n}\lambda_1+a_{2n}\lambda_2+\ldots+a_{mn}\lambda_m=0 \end{cases}$$

上述方程组的前r个方程即$\lambda_1\mathbf{a}_1+\lambda_2\mathbf{a}_2+\ldots+\lambda_m\mathbf{a}_m=\mathbf{0}$. 从而$\mathbf{a}_1,\mathbf{a}_2,\cdots,\mathbf{a}_m$也线性相关. □

在上述定理中, 向量组$\mathbf{a}_1,\mathbf{a}_2,\cdots,\mathbf{a}_m$可以视为向量组$\mathbf{b}_1,\mathbf{b}_2,\cdots,\mathbf{b}_m$在低维空间中的投影向量.

§5.3 极大无关组与秩

我们先来考察线性方程组独立方程的个数问题, 也就是Gauss消元的标准型(2.3)中非零行数的问题. 独立方程的个数决定了解集的"大小". 例如, 下列线性方程组的解集分别为一个点、一条直线、一个平面, 它们独立方程的个数分别为$3,2,1$.

$$\begin{cases} x+y+z=1 \\ 2x+y-z=2 \\ 3x-y+z=3 \end{cases} \qquad \begin{cases} x+y+z=0 \\ 2x+y+z=0 \\ 3x+2y+2z=0 \end{cases} \qquad \begin{cases} x+y+z=1 \\ 2x+2y+2z=2 \\ 3x+3y+3z=3 \end{cases}$$

一般地, 给定一个线性方程组, 如何确定它的独立方程的个数?

注意到, 如果有一个方程l_i可以表示成其它方程的线性组合, 则去掉l_i后得到的线性方程组与原方程组同解. 因此, 如果线性方程组线性相关, 则其独立方程个数小于m. 如果去掉l_i后的线性方程组仍然线性相关, 则可以进一步去掉其中一个方程, 剩下方程构成的方程组仍与原方程组等价. 这个过程可以一直进行下去, 直到剩下的方程构成的方程组线性无关. 这时, 剩下方程构成的方程组仍与原方程组等价, 它的个数就是原方程组独立方程的个数. 我们称这个数为原方程组的**秩**, 剩下的方程称为原方程组的**极大无关组**.

上述概念可以平行推广到到n维向量组.

定义5.3.1. 设S是一组数组向量, S_1是S的子向量组. 若S_1线性无关, 且对任意向量$\mathbf{a}\in S\backslash S_1$, $S_1\cup\{\mathbf{a}\}$线性相关, 则称S_1是S的**极大无关组**.

由本章习题 13的结论, 向量组的每一向量都可以由它的极大无关组线性表示.

例5.3.1. 求$\mathbf{a}_1=(2,-1,3,1)$, $\mathbf{a}_2=(4,-2,5,4)$, $\mathbf{a}_3=(2,-1,4,-1)$的极大无关组.

解. 由于$3\mathbf{a}_1=\mathbf{a}_2+\mathbf{a}_3$, 且$\mathbf{a}_1$, $\mathbf{a}_2$, $\mathbf{a}_3$中任两个线性无关, 故其中任何两个为极大无关组. □

上述例子表明, 极大无关组不唯一. 一般地, 如何有效地求一组极大无关组? 在给出算法之前, 我们先叙述一个基本结论.

定理5.3.1. *设$\mathbf{a}_1,\mathbf{a}_2,\cdots,\mathbf{a}_m\in F^n$为一组列向量, $A=(\mathbf{a}_1,\mathbf{a}_2,\cdots,\mathbf{a}_m)$是以$\mathbf{a}_1,\mathbf{a}_2,\cdots,\mathbf{a}_m$为列构成的$n\times m$阶矩阵. A经一系列的初等行变换变为矩阵$B=(\mathbf{b}_1,\mathbf{b}_2,\cdots,\mathbf{b}_m)$. 则*

1. *$\mathbf{a}_1,\mathbf{a}_2,\cdots,\mathbf{a}_m$线性相关(无关), 当且仅当$\mathbf{b}_1,\mathbf{b}_2,\cdots,\mathbf{b}_m$线性相关(无关).*
2. *$\mathbf{a}_{i_1},\mathbf{a}_{i_2},\cdots,\mathbf{a}_{i_r}$为$\mathbf{a}_1,\mathbf{a}_2,\cdots,\mathbf{a}_m$的极大无关组, 当且仅当$\mathbf{b}_{i_1},\mathbf{b}_{i_2},\cdots,\mathbf{b}_{i_r}$为$\mathbf{b}_1,\mathbf{b}_2,\cdots,\mathbf{b}_m$的极大无关组. 这里$1\le i_1<\ldots<i_r\le m$.*

证明. 1. 记$\mathbf{x}=(\lambda_1,\lambda_2,\ldots,\lambda_m)^T$. 由于初等变换不改变线性方程组的解, 因此

$\mathbf{a}_1,\mathbf{a}_2,\cdots,\mathbf{a}_m$线性相关$\Leftrightarrow$
线性方程组$A\mathbf{x}=\mathbf{0}$有非零解$\Leftrightarrow$
线性方程组$B\mathbf{x}=\mathbf{0}$有非零解$\Leftrightarrow$
$\mathbf{b}_1,\mathbf{b}_2,\cdots,\mathbf{b}_m$线性相关.

2. 由1的结论, 容易验证,

$\mathbf{a}_{i_1},\mathbf{a}_{i_2},\cdots,\mathbf{a}_{i_r}$为$\mathbf{a}_1,\mathbf{a}_2,\cdots,\mathbf{a}_m$的极大无关组$\Leftrightarrow$
$\mathbf{a}_{i_1},\mathbf{a}_{i_2},\cdots,\mathbf{a}_{i_r}$线性无关, 而$\mathbf{a}_{i_1},\mathbf{a}_{i_2},\cdots,\mathbf{a}_{i_r},\mathbf{a}_{i_{r+1}}$ 线性相关, 这里$\mathbf{a}_{i_{r+1}}$为$\mathbf{a}_1,\mathbf{a}_2,\cdots,\mathbf{a}_m$中除去$\mathbf{a}_{i_1},\mathbf{a}_{i_2},\cdots,\mathbf{a}_{i_r}$中向量的任一向量$\Leftrightarrow$
$\mathbf{b}_{i_1},\mathbf{b}_{i_2},\cdots,\mathbf{b}_{i_r}$线性无关, 而$\mathbf{b}_{i_1},\mathbf{b}_{i_2},\cdots,\mathbf{b}_{i_r},\mathbf{b}_{r_{r+1}}$ 线性相关$\Leftrightarrow$
$\mathbf{b}_{i_1},\mathbf{b}_{i_2},\cdots,\mathbf{b}_{i_r}$为$\mathbf{b}_1,\mathbf{b}_2,\cdots,\mathbf{b}_m$的极大无关组

□

根据上述定理, 要求一组列向量$\mathbf{a}_1,\mathbf{a}_2,\cdots,\mathbf{a}_m$的极大无关组, 只需要对它做初等变换得到一组较简单的向量组, 其极大无关组就容易求得.

例5.3.2. 给定向量组$\mathbf{a}_1=(-1,5,3,-2)$, $\mathbf{a}_2=(4,1,-2,9)$, $\mathbf{a}_3=(2,0,-1,4)$, $\mathbf{a}_4=(0,3,4,-5)$, 求它的极大无关组.

解. 对$(\mathbf{a}_1^T,\mathbf{a}_2^T,\mathbf{a}_3^T,\mathbf{a}_4^T)$做初等行变换得

$$\begin{pmatrix} -1 & 4 & 2 & 0 \\ 5 & 1 & 0 & 3 \\ 3 & -2 & -1 & 4 \\ -2 & 9 & 4 & -5 \end{pmatrix} \longrightarrow \begin{pmatrix} -1 & 4 & 2 & 0 \\ 0 & 1 & 0 & -5 \\ 0 & 0 & 5 & 54 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

易知, $\{\mathbf{a}_1,\mathbf{a}_2,\mathbf{a}_3\}$, $\{\mathbf{a}_1,\mathbf{a}_2,\mathbf{a}_4\}$分别为极大无关组. 实际上, $\{\mathbf{a}_2,\mathbf{a}_3,\mathbf{a}_4\}$也为极大无关组. □

前面说过, 线性方程组的极大无关组与原方程组等价, 因此, 原方程组的任何两个极大无关组也等价. 这个结论对一般向量组也成立.

定理5.3.2. 一组向量组与它的任何一组极大无关组等价.

证明. 不妨设$\mathbf{a}_1,\mathbf{a}_2,\ldots,\mathbf{a}_r$是向量组$\mathbf{a}_1,\mathbf{a}_2,\ldots,\mathbf{a}_m$ $(m\geq r)$的极大无关组. 显然, 极大无关组$\mathbf{a}_1,\mathbf{a}_2,\ldots,\mathbf{a}_r$可以由向量组$\mathbf{a}_1,\mathbf{a}_2,\ldots,\mathbf{a}_m$线性表示, 因为

$$\mathbf{a}_i = 0\mathbf{a_1}+\ldots+0\mathbf{a}_{i-1}+1\mathbf{a}_i+0\mathbf{a}_{i+1}+\ldots+0\mathbf{a}_m.$$

$i=1,2,\ldots,r$.

反过来, 由极大无关组的定义, 对任意$j\in\{r+1,\ldots,m\}$, $\mathbf{a}_1,\mathbf{a}_2,\ldots,\mathbf{a}_r,\mathbf{a}_j$线性相关, 而$\mathbf{a}_1,\mathbf{a}_2,\ldots,\mathbf{a}_r$线性无关, 由本章习题 13的结论, $\mathbf{a}_j$可以表示成$\mathbf{a}_1,\mathbf{a}_2,\ldots,\mathbf{a}_r$的线性组合. 因此, 向量组$\mathbf{a}_1,\mathbf{a}_2,\ldots,\mathbf{a}_m$也可以由向量组$\mathbf{a}_1,\mathbf{a}_2,\ldots,\mathbf{a}_r$线性表示. 从而两者等价. □

由上述定理及向量组等价的传递性有

推论5.3.1. 向量组的任何两个极大无关组彼此等价.

对线性方程组, 极大无关组的个数即独立方程的个数. 然而, 这里有一点需要澄清. 对于给定的某个线性方程组, 由于Gauss消元过程并不唯一, 所得的独立方程组也不唯一. 那么, 不同的独立方程组个数是否一致? 如果不一致, 哪一个才是真正的独立方程的个数? 幸运的是, 不同方法获得的独立方程组的个数完全一样! 且这个结论对向量组也成立!

命题5.3.3. 设向量组$\{\mathbf{a}_1,\mathbf{a}_2,\cdots,\mathbf{a}_r\}$可以由向量组$\{\mathbf{b}_1,\mathbf{b}_2,\cdots,\mathbf{b}_s\}$ 线性表示, 且$\{\mathbf{a}_1,\mathbf{a}_2,\cdots,\mathbf{a}_r\}$ 线性无关, 则$r\leq s$.

证明. 由于$\{\mathbf{a}_1,\mathbf{a}_2,\cdots,\mathbf{a}_r\}$可以用$\{\mathbf{b}_1,\mathbf{b}_2,\cdots,\mathbf{b}_s\}$线性表示, 则存在常数$\mu_{ij}$, $i=1,\ldots,r$, $j=1,\ldots,s$使得$\mathbf{a}_i=\sum_{j=1}^{s}\mu_{ij}\mathbf{b}_j$. 考察关于$\lambda_1,\lambda_2,\ldots,\lambda_r$的线性方程组

$$\sum_{i=1}^{r}\lambda_i\mathbf{a}_i=\mathbf{0}.$$

由于$\{\mathbf{a}_1,\mathbf{a}_2,\cdots,\mathbf{a}_r\}$线性无关, 该方程组只有零解. 将它重新改写成

$$\sum_{j=1}^{s}\left(\sum_{i=1}^{r}\lambda_i\mu_{ij}\right)\mathbf{b}_j=\mathbf{0}.$$

由于$\{\mathbf{b}_1,\mathbf{b}_2,\cdots,\mathbf{b}_s\}$线性无关, 上式等价于

$$\sum_{i=1}^{r}\lambda_i\mu_{ij}=0, \quad ,j=1,2,\ldots,s$$

这是关于$\lambda_1,\ldots,\lambda_r$的齐次线性方程组, 且知道它只有零解. 由推论 2.3.2, $r\le s$.

□

由上述命题立即可得

定理5.3.4. 两个分别线性无关向量组$\{\mathbf{a}_1,\mathbf{a}_2,\cdots,\mathbf{a}_r\}$和$\{\mathbf{b}_1,\mathbf{b}_2,\cdots,\mathbf{b}_s\}$等价, 则$r=s$.

上述定理的一个直接推论是

推论5.3.2. 设$\mathbf{a}_{i_1},\mathbf{a}_{i_2},\cdots,\mathbf{a}_{i_r}$和$\mathbf{a}_{j_1},\mathbf{a}_{j_2},\cdots,\mathbf{a}_{j_s}$分别为$\mathbf{a}_1,\mathbf{a}_2,\cdots,\mathbf{a}_m$的两个极大无关组, 则$r=s$.

根据上述结论, 线性方程组的独立方程的个数是唯一确定的, 它就是方程组的秩.

定义5.3.2. 向量组$\mathbf{a}_1,\mathbf{a}_2,\cdots,\mathbf{a}_m$的极大无关组元素的个数称为向量组的**秩**, 记为$\mathrm{rank}(\mathbf{a}_1,\mathbf{a}_2,\cdots,\mathbf{a}_m)$, 或$r(\mathbf{a}_1,\mathbf{a}_2,\cdots,\mathbf{a}_m)$.

下面我们进一步从几何上考察秩的含义.

设$\mathbf{a},\mathbf{b},\mathbf{c}$是三维几何空间中的向量. 如果$\mathbf{a},\mathbf{b},\mathbf{c}$共线且至少有一个非零, 则生成子空间$\langle\mathbf{a},\mathbf{b},\mathbf{c}\rangle$表示一条直线, 此时$r(\mathbf{a},\mathbf{b},\mathbf{c})=1$. 如果$\mathbf{a},\mathbf{b},\mathbf{c}$共面, 且至少有两个向量不共线, 则$\langle\mathbf{a},\mathbf{b},\mathbf{c}\rangle$表示一个平面, 此时$r(\mathbf{a},\mathbf{b},\mathbf{c})=2$. 如果$\mathbf{a},\mathbf{b},\mathbf{c}$不共面, 则$\langle\mathbf{a},\mathbf{b},\mathbf{c}\rangle$就是整个几何空间, 此时$r(\mathbf{a},\mathbf{b},\mathbf{c})=3$. 因此, 向量组的秩表示了向量组生成子空间的维数. 秩越大, 则生成子空间也越大, 反之亦然. 这就是秩的几何意义.

关于秩, 我们有下列有用的结论.

定理5.3.5. 设向量$\mathbf{a}_1,\mathbf{a}_2,\cdots,\mathbf{a}_m\in F^n$, 则有

1. $\mathbf{a}_1,\mathbf{a}_2,\cdots,\mathbf{a}_m$线性无关, 当且仅当$r(\mathbf{a}_1,\mathbf{a}_2,\cdots,\mathbf{a}_m)=m$.
2. $\mathbf{a}_1,\mathbf{a}_2,\cdots,\mathbf{a}_m$线性相关, 当且仅当$r(\mathbf{a}_1,\mathbf{a}_2,\cdots,\mathbf{a}_m)<m$.
3. $\{\mathbf{b}_1,\mathbf{b}_2,\cdots,\mathbf{b}_s\}$可以用$\{\mathbf{a}_1,\mathbf{a}_2,\cdots,\mathbf{a}_r\}$线性表示, 则$r(\mathbf{b}_1,\mathbf{b}_2,\cdots,\mathbf{b}_s)\le r(\mathbf{a}_1,\mathbf{a}_2,\cdots,\mathbf{a}_r)$.
4. $\{\mathbf{b}_1,\mathbf{b}_2,\cdots,\mathbf{b}_s\}$与$\{\mathbf{a}_1,\mathbf{a}_2,\cdots,\mathbf{a}_r\}$等价, 则$r(\mathbf{b}_1,\mathbf{b}_2,\cdots,\mathbf{b}_s)=r(\mathbf{a}_1,\mathbf{a}_2,\cdots,\mathbf{a}_r)$.
5. 向量$\mathbf{b}$可以表示成$\mathbf{a}_1,\mathbf{a}_2,\cdots,\mathbf{a}_m$的线性组合, 当且仅当$r(\mathbf{a}_1,\mathbf{a}_2,\cdots,\mathbf{a}_m)=r(\mathbf{a}_1,\mathbf{a}_2,\cdots,\mathbf{a}_m,\mathbf{b})$.

证明. 只证3与5, 其余留给读者自己完成.

3. 不妨设$\{\mathbf{b}_1,\mathbf{b}_2,\cdots,\mathbf{b}_k\}$是$\{\mathbf{b}_1,\mathbf{b}_2,\cdots,\mathbf{b}_s\}$的极大无关组, $\{\mathbf{a}_1,\mathbf{a}_2,\cdots,\mathbf{a}_l\}$是$\{\mathbf{a}_1,\mathbf{a}_2,\cdots,\mathbf{a}_r\}$的极大无关组. 则$\{\mathbf{b}_1,\mathbf{b}_2,\cdots,\mathbf{b}_k\}$可以用$\{\mathbf{a}_1,\mathbf{a}_2,\cdots,\mathbf{a}_l\}$线性表示. 由命题 5.3.4 可知$k\le l$, 亦即$r(\mathbf{b}_1,\mathbf{b}_2,\cdots,\mathbf{b}_s)\le r(\mathbf{a}_1,\mathbf{a}_2,\cdots,\mathbf{a}_r)$.

5. 假设向量$\mathbf{b}$可以表示成向量组$\mathbf{a}_1,\mathbf{a}_2,\cdots,\mathbf{a}_m$的线性组合, 则$\mathbf{a}_1,\mathbf{a}_2,\cdots,\mathbf{a}_m$与$\mathbf{a}_1,\mathbf{a}_2,\cdots,\mathbf{a}_m,\mathbf{b}$等价, 从而$r(\mathbf{a}_1,\mathbf{a}_2,\cdots,\mathbf{a}_m)=r(\mathbf{a}_1,\mathbf{a}_2,\cdots,\mathbf{a}_m,\mathbf{b})$.

反之, 假设$r(\mathbf{a}_1,\mathbf{a}_2,\cdots,\mathbf{a}_m)=r(\mathbf{a}_1,\mathbf{a}_2,\cdots,\mathbf{a}_m,\mathbf{b})$, 且$\mathbf{a}_1,\mathbf{a}_2,\cdots,\mathbf{a}_k$是$\mathbf{a}_1,\mathbf{a}_2,\cdots,\mathbf{a}_m$的极大无关组, 则$\mathbf{a}_1,\mathbf{a}_2,\cdots,\mathbf{a}_k$也是$\mathbf{a}_1,\mathbf{a}_2,\cdots,\mathbf{a}_m,\mathbf{b}$的极大无关组, 从而$\mathbf{b}$是$\mathbf{a}_1,\mathbf{a}_2,\cdots,\mathbf{a}_k$的线性表示, 因而也是$\mathbf{a}_1,\mathbf{a}_2,\cdots,\mathbf{a}_m$的线性表示. □

上述定理表明, 利用秩可以判断向量组的线性相关(无关)性及等价性等. 这表明秩是很深刻的一个概念.

定理 5.3.5的第5条实际上再次给出了线性方程组(5.4)有解的充要条件. 这个条件直接由原方程的系数矩阵及常数项给出, 不像第一章那样需要经过初等变换后才能给出.

接下来, 我们研究向量组的秩与矩阵秩的关系.

给定$m\times n$阶方阵A, 它可以看成由m个n维行向量组成, 也可以看成由n个m维列向量组成, 即

$$A=\begin{pmatrix}\mathbf{a}_1\\ \mathbf{a}_2\\ \vdots\\ \mathbf{a}_m\end{pmatrix}=\begin{pmatrix}\mathbf{b}_1 & \mathbf{b}_2 & \cdots & \mathbf{b}_n\end{pmatrix}. \tag{5.9}$$

作为向量组$\mathbf{a}_1,\mathbf{a}_2,\ldots,\mathbf{a}_m$与$\mathbf{b}_1,\mathbf{b}_2,\cdots,\mathbf{b}_n$都有秩, 分别称为矩阵$A$的**行秩**与**列秩**. 问题是, 矩阵$A$的行秩、列秩与它本身的秩有什么关系? 我们先看几个具体实例.

例5.3.3. 计算下列矩阵的秩、行秩与列秩.

$$(1)\, A=\begin{pmatrix}1&2&-3\\-1&3&1\\1&7&-5\end{pmatrix}\qquad (2)\, A=\begin{pmatrix}1&1&1&1\\1&2&3&0\\1&4&9&3\end{pmatrix}$$

$$(3)\, A=\begin{pmatrix}c_{11}&c_{12}&c_{13}&c_{14}&c_{15}\\0&0&c_{23}&c_{24}&c_{25}\\0&0&0&c_{34}&c_{35}\\0&0&0&0&0\end{pmatrix}\ (c_{11}c_{23}c_{34}\neq 0)$$

解. 直接验算可知, 对矩阵(1)有, A的行秩$=A$的列秩$=A$的秩$=2$. 对矩阵(2)有, A的行秩$=A$的列秩$=A$的秩$=3$.

对矩阵(3), 由于$c_{11}c_{23}c_{34}\neq 0$, 矩阵A的前三行线性无关, 因此A的行秩为3. 同样, A的第一、三、四列线性无关, 且为A的所有列的极大无关组, 因此A的列秩也为3. 最后, 由于A的前三行、第一、三、四列构成的三阶子式非零, 因此$r(A)=3$. 总之, 仍然有A的行秩$=A$的列秩$=A$的秩. □

上述例子表明, 矩阵的行秩、列秩与秩是一致的. 这一结论对任何矩阵成立.

定理5.3.6. 任何矩阵的行秩等于它的列秩, 等于该矩阵的秩.

证明. 该定理依赖于以下事实

1. 初等变换不改变矩阵的列秩;
2. 初等变换不改变矩阵的行秩;
3. 初等变换不改变矩阵的秩.

我们只证初等行变换对秩的不变性, 初等列变换类似证明. 第1条从定理5.3.1 可以直接得到. 第3条由第三章定理 4.4.3也是显然的.下证第2条.

设矩阵A的行为$\mathbf{a}_1,\mathbf{a}_2,\ldots,\mathbf{a}_m\in F^n$. 只要说明经过三种初等行变换, 所得的向量组与$\mathbf{a}_1,\mathbf{a}_2,\ldots,\mathbf{a}_m$等价即可. 第一、二种初等变换显然将向量组变为等价向量组. 第三种初等变换将$\mathbf{a}_1,\mathbf{a}_2,\ldots,\mathbf{a}_m$变为

$$\mathbf{a}_1,\mathbf{a}_2,\ldots,\mathbf{a}_i+\lambda\mathbf{a}_j,\ldots,\mathbf{a}_m.$$

由于$\mathbf{a}_i=(\mathbf{a}_i+\lambda\mathbf{a}_j)-\lambda\mathbf{a}_j$, 变换前后的向量组可以相互表示, 亦即彼此等价. 因此, 经过初等行变换, 矩阵的行秩不变.

根据以上事实, 以及任何矩阵可以经过初等变换变为相抵标准形$\begin{pmatrix} I_r & O \\ O & O \end{pmatrix}$, 而相抵标准形的秩、行秩与列秩都相等, 即知结论成立.

□

上述定理给出了矩阵秩的一种新的解释, 即矩阵的秩就是矩阵行向量极大无关组的个数, 也是列向量极大无关组的个数. 秩越大, 则矩阵元素之间的相关性越小; 反之, 秩越小, 则矩阵元素之间的相关性越大.

下面两个推论比较显然, 请读者完成.

推论5.3.3. n阶方阵A可逆$\Leftrightarrow$ $\mathrm{rank}(A)=n$ $\Leftrightarrow$ A的行向量线性无关$\Leftrightarrow$ A的列向量线性无关.

推论5.3.4. 若$\mathrm{rank}(A)=r$, 则A的不等于零的r阶子式所在行(列)构成A的行(列)向量的极大无关组.

例5.3.4. 设$A\in F^{m\times n}, B\in F^{n\times l}$. 证明: $\mathrm{rank}(AB)\le \min(\mathrm{rank}(A),\mathrm{rank}(B))$.

证明. 记$C=AB$, $A=(a_{ij})_{m\times n}$, B的n个行向量为$\mathbf{b}_i$, $i=1,\dots,n$, C的m行向量为$\mathbf{c}_i$, $i=1,\dots,m$. 则由$C=AB$得

$$\mathbf{c}_i=a_{i1}\mathbf{b}_1+a_{i2}\mathbf{b}_2+\ldots+a_{in}\mathbf{b}_n,\quad i=1,\dots,m$$

这表明向量组$\mathbf{c}_1,\mathbf{c}_2,\dots,\mathbf{c}_m$可以由向量组$\mathbf{b}_1,\mathbf{b}_2,\dots,\mathbf{b}_n$线性表示. 因此, $\mathrm{rank}(\mathbf{c}_1,\mathbf{c}_2,\dots,\mathbf{c}_m)\le\mathrm{rank}(\mathbf{b}_1,\mathbf{b}_2,\dots,\mathbf{b}_n)$. 从而$\mathrm{rank}(AB)\le\mathrm{rank}(B)$. 另一方面, $\mathrm{rank}(AB)=\mathrm{rank}(B^TA^T)\le\mathrm{rank}(A^T)=\mathrm{rank}(A)$. 证毕. □

§5.4 基与维数

在本章第一节我们引进了数组空间的子空间的概念, 本节将研究子空间的结构. 首先要回答的一个基本的问题是, 是不是数组空间F^n的每个子空间都可以由一组向量生成? 回答是肯定的. 我们有

定理5.4.1. 设V是F^n的子空间. 则存在线性无关的向量组$\mathbf{a}_1,\mathbf{a}_2,\dots,\mathbf{a}_r$使得

$$V=\langle\mathbf{a}_1,\mathbf{a}_2,\dots,\mathbf{a}_r\rangle.$$

证明. 设$V\ne\mathbf{0}$. 在V中任取向量$\mathbf{a}_1\ne\mathbf{0}$, 若$V=\langle\mathbf{a}_1\rangle$, 则结论成立. 否则, 任取$\mathbf{a}_2\in V\backslash\langle\mathbf{a}_1\rangle$. 下证$\mathbf{a}_1,\mathbf{a}_2$线性无关. 实际上, 若$\mathbf{a}_1,\mathbf{a}_2$线性相关, 由于$\mathbf{a}_1\ne\mathbf{0}$, 则$\mathbf{a}_2$是$\mathbf{a}_1$的常数倍, 即$\mathbf{a}_2\in\langle\mathbf{a}_1\rangle$, 这与$\mathbf{a}_2$的取法矛盾. 因此, $\mathbf{a}_1,\mathbf{a}_2$线性无关.

接下来考察$V=\langle \mathbf{a}_1,\mathbf{a}_2\rangle$是否成立. 若成立, 则定理成立. 否则, 任取$\mathbf{a}_3\in V\backslash\langle \mathbf{a}_1,\mathbf{a}_2\rangle$. 下证$\mathbf{a}_1,\mathbf{a}_2,\mathbf{a}_3$线性无关. 若$\mathbf{a}_1,\mathbf{a}_2,\mathbf{a}_3$线性相关, 由于$\mathbf{a}_1,\mathbf{a}_2$线性无关, 则$\mathbf{a}_3\in\langle \mathbf{a}_1,\mathbf{a}_2\rangle$, 这与$\mathbf{a}_3$ 的取法矛盾. 因此, $\mathbf{a}_1,\mathbf{a}_2,\mathbf{a}_3$线性无关.

继续上述过程, 由于F^n的任何线性无关向量组中向量的个数最多为n, 因此, 存在线性无关向量组$\mathbf{a}_1,\mathbf{a}_2,\ldots,\mathbf{a}_r(r\le n)$使得$V=\langle \mathbf{a}_1,\mathbf{a}_2,\ldots,\mathbf{a}_r\rangle$. □

上述定理表明, 任意向量$\mathbf{a}\in V$都可以表示成$\mathbf{a}_1,\mathbf{a}_2,\ldots,\mathbf{a}_r$的线性组合

$$\mathbf{a}=x_1\mathbf{a}_1+x_2\mathbf{a}_2+\ldots+x_r\mathbf{a}_r,$$

并且上述表示是唯一的. 实际上, 如果

$$\mathbf{a}=y_1\mathbf{a}_1+y_2\mathbf{a}_2+\ldots+y_r\mathbf{a}_r$$

是$\mathbf{a}$的另一个表示, 则

$$(x_1-y_1)\mathbf{a}_1+(x_2-y_2)\mathbf{a}_2+\ldots+(x_r-y_r)\mathbf{a}_r=\mathbf{0}.$$

由$\mathbf{a}_1,\mathbf{a}_2,\ldots,\mathbf{a}_r$的线性无关性得, $x_i=y_i$, $i=1,2,\ldots,r$, 即$\mathbf{a}$的表示唯一. 因此, 从这个意义上, 一组线性无关的向量决定了子空间的结构. 这样一组线性无关向量称为子空间的一组基.

定义5.4.1. $V\subset F^n$是子空间. V中一组向量$\{\mathbf{a}_1,\mathbf{a}_2,\cdots,\mathbf{a}_r\}$称为$V$的一组**基**, 如果它满足

1. 对任意向量$\mathbf{a}\in V$, $\mathbf{a}$可表示成$\mathbf{a}_1,\mathbf{a}_2,\cdots,\mathbf{a}_r$ 的线性组合
$$\mathbf{a}=x_1\mathbf{a}_1+x_2\mathbf{a}_2+\ldots+\mathbf{x}_r\mathbf{a}_r. \tag{5.10}$$
2. $\mathbf{a}_1,\mathbf{a}_2,\cdots,\mathbf{a}_r$线性无关.

称$(x_1,x_2,\ldots,x_r)$为向量$\mathbf{a}$在基$\{\mathbf{a}_1,\mathbf{a}_2,\cdots,\mathbf{a}_r\}$下的**坐标**.

由前面的分析, 表达式 (5.10)是唯一的. 同时, 基可以看成是子空间作为一组向量组的极大无关组! 由于极大无关组不唯一, 因此基也不是唯一的. 但同一个子空间的基是彼此等价的, 因而它们所含向量的个数是一样的, 这个数称之为子空间的维数, 它可以看作是子空间作为一组向量组的秩.

定义5.4.2. 设$V\subset F^n$为子空间. 称V的一组基的向量个数为V的**维数**, 记为$\dim V=\mathrm{rank}(V)$.

例5.4.1. 在$\mathbf{R}^2$中, 设$\mathbf{a}_1,\mathbf{a}_2$是其中两个不共线的向量. 则对任意$\mathbf{a}\in\mathbf{R}^2$, $\mathbf{a}$可以唯一地表示成$x_1\mathbf{a}_1+x_2\mathbf{a}_2$. 因此, $\mathbf{a}_1,\mathbf{a}_2$为$\mathbf{R}^2$的一组基, (x_1,x_2)是$\mathbf{a}$在基$\mathbf{a}_1,\mathbf{a}_2$下的坐标.

类似地, 在$\mathbf{R}^3$中, 设$\mathbf{a}_1,\mathbf{a}_2,\mathbf{a}_3$是不共面的向量. 则对任意$\mathbf{a}\in\mathbf{R}^3$, $\mathbf{a}$可以唯一地表示成$x_1\mathbf{a}_1+x_2\mathbf{a}_2+x_3\mathbf{a}_3$. 因此, $\mathbf{a}_1,\mathbf{a}_2,\mathbf{a}_3$为$\mathbf{R}^3$的一组基, (x_1,x_2,x_3)为$\mathbf{a}$在基$\mathbf{a}_1,\mathbf{a}_2,\mathbf{a}_3$下的坐标.

例5.4.2. 设$\mathbf{e}_i$是F^n中的单位坐标向量. 则向量组$\{\mathbf{e}_1,\mathbf{e}_2,\ldots,\mathbf{e}_n\}$是$F^n$的一组基, 称为**自然基**. 任何向量$\mathbf{b}=(b_1,b_2,\ldots,b_n)\in F^n$可以唯一地表示成

$$\mathbf{b}=b_1\mathbf{e}_1+b_2\mathbf{e}_2+\ldots+b_n\mathbf{e}_n.$$

即$\mathbf{b}$在基$\{\mathbf{e}_1,\mathbf{e}_2,\ldots,\mathbf{e}_n\}$下的坐标是$(b_1,b_2,\ldots,b_n)$.

设$\mathbf{a}_i=(1,\ldots,1,0,\ldots,0)\in F^n$是前$i$个分量为1, 后$n-i$个分量为零的向量. 显然, $\mathbf{a}_1,\mathbf{a}_2,\ldots,\mathbf{a}_n$线性无关, 因此也是$F^n$的一组基, 并且有

$$\mathbf{a}_i=\mathbf{e}_1+\mathbf{e}_2+\ldots+\mathbf{e}_i,\qquad i=1,\ldots,n.$$

因此, $\mathbf{e}_1=\mathbf{a}_1$, $\mathbf{e}_i=\mathbf{a}_i-\mathbf{a}_{i-1}$, $i=2,\ldots,n$. 从而

$$\mathbf{b}=\sum_{i=1}^n b_i\mathbf{e}_i=b_1\mathbf{a}_1+\sum_{i=2}^n b_i(\mathbf{a}_i-\mathbf{a}_{i-1})=\sum_{i=1}^{n-1}(b_i-b_{i+1})\mathbf{a}_i+b_n\mathbf{a}_n.$$

即$\mathbf{b}$在$\mathbf{a}_1,\mathbf{a}_2,\ldots,\mathbf{a}_n$下的坐标为$(b_1-b_2,b_2-b_3,\ldots,b_n)$.

例5.4.3. 设$A\in F^{m\times n}$, 且$\operatorname{rank}(A)=r$. 则A的行向量的极大无关组是A的行空间$R(A)$的一组基, 行空间的维数为r. 同理, A的列向量的极大无关组是A的列空间$C(A)$的一组基, 且列空间的维数也为r.

一个子空间的不同基是彼此等价的, 即它们之间可以相互表示. 利用这种相互表示关系, 一个向量在一组基下的坐标与在另一组基下的坐标之间的关系也就清楚了. 这就是所谓的**坐标变换**. 具体地,

设F^n有两组基

$$\mathbf{a}_1,\mathbf{a}_2,\ldots,\mathbf{a}_n$$

$$\mathbf{b}_1,\mathbf{b}_2,\ldots,\mathbf{b}_n$$

设两组基之间的关系由下式确定

$$\begin{cases}\mathbf{b}_1=t_{11}\mathbf{a}_1+t_{21}\mathbf{a}_2+\ldots+t_{n1}\mathbf{a}_n\\ \mathbf{b}_2=t_{12}\mathbf{a}_1+t_{22}\mathbf{a}_2+\ldots+t_{n2}\mathbf{a}_n\\ \qquad\qquad\cdots\cdots\cdots\cdots\cdots\cdots\\ \mathbf{b}_n=t_{1n}\mathbf{a}_1+b_{2n}\mathbf{a}_2+\ldots+t_{nn}\mathbf{a}_n\end{cases}$$

上式可以形式上写成矩阵的形式

$$(\mathbf{b}_1,\mathbf{b}_2,\dots,\mathbf{b}_n)=(\mathbf{a}_1,\mathbf{a}_2,\dots,\mathbf{a}_n)T, \tag{5.11}$$

其中矩阵

$$T=\begin{pmatrix} t_{11} & \dots & t_{1n} \\ \vdots & \vdots & \vdots \\ t_{n1} & \dots & t_{nn} \end{pmatrix} \tag{5.12}$$

称矩阵T为从基$\mathbf{a}_1,\mathbf{a}_2,\dots,\mathbf{a}_n$到基$\mathbf{b}_1,\mathbf{b}_2,\dots,\mathbf{b}_n$的**过渡矩阵**. 由于基都是线性无关的向量组, 因此过渡矩阵T是可逆的, 并且

$$(\mathbf{a}_1,\mathbf{a}_2,\dots,\mathbf{a}_n)=(\mathbf{b}_1,\mathbf{b}_2,\dots,\mathbf{b}_n)T^{-1}, \tag{5.13}$$

即从基$\mathbf{b}_1,\mathbf{b}_2,\dots,\mathbf{b}_n$到基$\mathbf{a}_1,\mathbf{a}_2,\dots,\mathbf{a}_n$的过渡矩阵为$T^{-1}$ (请读者证明). 现设向量$\mathbf{v}\in F^n$在两组基$\mathbf{a}_1,\mathbf{a}_2,\dots,\mathbf{a}_n$及$\mathbf{b}_1,\mathbf{b}_2,\dots,\mathbf{b}_n$下的坐标分别为$X=(x_1,x_2,\dots,x_n)^T$ 与$Y=(y_1,y_2,\dots,y_n)^T$, 则由

$$\mathbf{v}=(\mathbf{a}_1,\mathbf{a}_2,\dots,\mathbf{a}_n)X=(\mathbf{b}_1,\mathbf{b}_2,\dots,\mathbf{b}_n)Y=(\mathbf{a}_1,\mathbf{a}_2,\dots,\mathbf{a}_n)TY$$

可得$X=TY$. 因此, 从原坐标X到新坐标Y的**坐标变换公式**为

$$Y=T^{-1}X. \tag{5.14}$$

例5.4.4. 重新考察例5.4.2. 从自然基$\mathbf{e}_1,\mathbf{e}_2,\dots,\mathbf{e}_n$到基$\mathbf{a}_1,\mathbf{a}_2,\dots,\mathbf{a}_n$的过渡矩阵为

$$T=\begin{pmatrix} 1 & 1 & \dots & 1 \\ 0 & 1 & \dots & 1 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 1 \end{pmatrix}$$

不难算得

$$T^{-1}=\begin{pmatrix} 1 & -1 & 0 & \dots & 0 \\ 0 & 1 & -1 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \dots & \dots & \dots & \dots & \\ 0 & 0 & 0 & \dots & 1 \end{pmatrix}$$

因此$\mathbf{v}$在自然基下的坐标$X=(x_1,x_2,\dots,x_n)^T$与它在基$\mathbf{a}_1,\mathbf{a}_2,\dots,\mathbf{a}_n$下的坐标$Y=(y_1,y_2,\dots,y_n)^T$之间有关系(5.14), 从而$(y_1,y_2,\dots,y_n)=(x_1-x_2,x_2-x_3,\dots,x_n)$. 这与例5.4.2中的结果一致.

例5.4.5. 考察坐标系的旋转. 将平面直角坐标系oxy逆时针旋转θ角得到新坐标系$ox'y'$. 记x轴正向的单位向量为$\mathbf{e}_1$, y轴正向的单位向量为$\mathbf{e}_2$. 记x'轴正向的单位向量为$\mathbf{e}_1'$, y'轴正向的单位向量为$\mathbf{e}_2'$. 则$\mathbf{R}^2$的两组基$\mathbf{e}_1', \mathbf{e}_2'$与$\mathbf{e}_1, \mathbf{e}_2$之间的关系为

$$(\mathbf{e}_1', \mathbf{e}_2') = (\mathbf{e}_1, \mathbf{e}_2)\begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}.$$

因此, 点$\mathbf{P}$在原坐标系下的坐标(x, y)与在新坐标系下的坐标(x', y')之间的关系为

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}^{-1}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix}.$$

类似地, 在三维几何空间中, 直角坐标系$[O; \mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3]$经过转轴得到新坐标系$\tilde{\mathbf{e}}_1, \tilde{\mathbf{e}}_2, \tilde{\mathbf{e}}_3$. 若$\tilde{\mathbf{e}}_1, \tilde{\mathbf{e}}_2, \tilde{\mathbf{e}}_3$在当前的坐标系中的坐标分别为$(a_1, a_2, a_3), (b_1, b_2, b_3), (c_1, c_2, c_3)$, 则新旧坐标系之间的过渡变换为

$$(\tilde{\mathbf{e}}_1, \tilde{\mathbf{e}}_2, \tilde{\mathbf{e}}_3) = (\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3)\begin{pmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_3 \\ a_3 & b_3 & c_3 \end{pmatrix},$$

相应的坐标变换为

$$\begin{pmatrix} \tilde{x} \\ \tilde{y} \\ \tilde{z} \end{pmatrix} = \begin{pmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{pmatrix}\begin{pmatrix} x \\ y \\ z \end{pmatrix}.$$

关于子空间的维数与基有下列有用的结论.

定理5.4.2. *n维数组空间F^n中的下列结论成立.*

1. *设$V \subset F^n$为r维子空间, 则V中任意$r+1$个向量线性相关.*
2. *设V为r维子空间, 则V中任意r个线性无关向量为V的一组基.*
3. *设U与V为F^n的子空间, 且$U \subseteq V$, 则$\dim U \le \dim V$.*
4. *设U与V为F^n的子空间, 且$U \subseteq V$, 若$\dim U = \dim V$, 则$U = V$.*

证明. 请读者自己完成. □

定理5.4.3. *设$V \subset F^n$为r维子空间. $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_s \in V$是$s(s < r)$个线性无关的向量. 则存在$V$中的向量$\mathbf{a}_{s+1}, \ldots, \mathbf{a}_r$使得$\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_r$构成$V$的一组基. 称$\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_r$为线性无关组$\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_s$ 的一组**扩充基**.*

证明. 类似定理5.4.1的证明, 请读者完成. □

例5.4.6. 在$\mathbb{R}^3$中, 令$V=\{(x_1,x_2,x_3)|\, x_1+x_2+x_3=0\}$. 证明: V是$\mathbb{R}^3$的子空间, 并求V的一组基与维数.

证明. 设$\mathbf{x}=(x_1,x_2,x_3),\mathbf{y}=(y_1,y_2,y_3)\in V$, 即$x_1+x_2+x_3=0, y_1+y_2+y_3=0$. 则对任意实数$\lambda,\mu$,

$$\lambda\mathbf{x}+\mu\mathbf{y}=(\lambda x_1+\mu y_1,\lambda x_2+\mu y_2,\lambda x_3+\mu y_3)\in V,$$

因为$(\lambda x_1+\mu y_1)+(\lambda x_2+\mu y_2)+(\lambda x_3+\mu y_3)=\lambda(x_1+x_2+x_3)+\mu(y_1+y_2+y_3)=0$, 故$V$是子空间.

容易验证, $(-1,1,0)$与$(-1,0,1)$是V的一组线性无关组, 且任意$\mathbf{x}=(x_1,x_2,x_3)\in V$, $\mathbf{x}=x_2(-1,1,0)+x_3(-1,0,1)$. 从而$(-1,1,0)$与$(-1,0,1)$是$V$的一组基, 故$\dim(V)=2$. □

例5.4.7. 在$\mathbb{R}^3$中, 令$V=\{(x_1,x_2,x_3)|\ x_1+x_2+x_3=0, 2x_1-x_2+x_3=3\}$. 证明: V不是$\mathbb{R}^3$的子空间.

证明. 设$\mathbf{x}=(x_1,x_2,x_3),\mathbf{y}=(y_1,y_2,y_3)\in V$. 显然,

$$\mathbf{x}+\mathbf{y}=(x_1+y_1,x_2+y_2,x_3+y_3)\notin V,$$

因为$2(x_1+y_1)-(x_2+y_2)+(x_3+y_3)=(2x_1-x_2+x_3)+(2y_1-y_2+y_3)=6$, 故$V$不是子空间. □

§5.5 线性方程组解集的结构

在第一章, 我们提出了关于线性方程组的下列问题:

1. 确定线性方程组解存在、唯一的条件;
2. 确定线性方程组解的几何结构;
3. 给出线性方程组的公式解.

本节我们将利用前几节的理论对上述问题给出一个完整的解答.

§5.5.1 线性方程组解的存在性与唯一性

根据前几节的结论, 我们可以给出线性方程组解的存在性与唯一性条件.

定理5.5.1. 设$A \in F^{m\times n}$为$m\times n$阶矩阵, $\mathbf{b}\in F^m$为m维列向量. 则线性方程组

$$A\mathbf{x} = \mathbf{b} \tag{5.15}$$

有解的充要条件是$\operatorname{rank}(A) = \operatorname{rank}(A,\mathbf{b})$. 线性方程组有唯一解的充要条件是$\operatorname{rank}(A) = \operatorname{rank}(A,\mathbf{b}) = n$.

证明. 设矩阵A的列为$\mathbf{a}_1,\mathbf{a}_2,\dots,\mathbf{a}_n$, $\mathbf{x} = (x_1,\dots,x_n)^T$. 则线性方程组$A\mathbf{x} = \mathbf{b}$ 等价于

$$x_1\mathbf{a}_1 + x_2\mathbf{a}_2 + \dots + x_n\mathbf{a}_n = \mathbf{b}.$$

因此, 线性方程组有解$\Leftrightarrow$ $\mathbf{b}\in\langle\mathbf{a}_1,\mathbf{a}_2,\dots,\mathbf{a}_n\rangle \Leftrightarrow \langle\mathbf{a}_1,\mathbf{a}_2,\dots,\mathbf{a}_n\rangle = \langle\mathbf{a}_1,\mathbf{a}_2,\dots,\mathbf{a}_n,\mathbf{b}\rangle \Leftrightarrow \operatorname{rank}(\mathbf{a}_1,\mathbf{a}_2,\dots,\mathbf{a}_n) = \operatorname{rank}(\mathbf{a}_1,\mathbf{a}_2,\dots,\mathbf{a}_n,\mathbf{b}) \Leftrightarrow \operatorname{rank}(A) = \operatorname{rank}(A,\mathbf{b})$.

线性方程组有唯一解$\Leftrightarrow$ $\mathbf{a}_1,\mathbf{a}_2,\dots,\mathbf{a}_n$线性无关, 且$\operatorname{rank}(A) = \operatorname{rank}(A,\mathbf{b}) \Leftrightarrow \operatorname{rank}(A) = \operatorname{rank}(A,\mathbf{b}) = n$. □

推论5.5.1. 齐次线性方程组

$$A\mathbf{x} = \mathbf{0} \tag{5.16}$$

有非零解的充要条件是$\operatorname{rank}(A) < n$. 特别地, 若$A$为$n$阶方阵, 则齐次线性方程组有非零解当且仅当$\det(A) = 0$.

§5.5.2 齐次线性方程组解集的结构

我们先研究三元线性方程组的解集的几何结构. 三元齐次线性方程组的解集表示每个方程定义的平面的公共交集, 它可以是一个点(原点), 过原点的一条直线或过原点的一个平面. 注意到, 这些解集有子空间的特性. 并且解集的“大小”是由独立的方程的个数决定. 如果独立方程是三个, 则解集是原点; 如果独立方程是两个, 则解集是过原点的直线; 如果独立方程是一个, 则解集是过原点的平面. 本节我们将指出, 一般的齐次线性方程组的解集有类似的特性.

考虑齐次线性方程组(5.16)的解集的全体

$$V := \{\mathbf{x}\in F^n \mid A\mathbf{x} = \mathbf{0}\}. \tag{5.17}$$

定理5.5.2. 齐次线性方程组*(5.16)*的解集V是F^n的子空间, 并且$\dim V = n - r(A)$. 称V为矩阵A的**零空间**, 记为$N(A)$.

证明. 设$\mathbf{x},\mathbf{y}\in V$, 即$A\mathbf{x}=\mathbf{0}, A\mathbf{y}=\mathbf{0}$. 对任意$\lambda,\mu\in F$, 由于$A(\lambda\mathbf{x}+\mu\mathbf{y})=\lambda A\mathbf{x}+\mu A\mathbf{y}=\mathbf{0}$, 因此, $\lambda\mathbf{x}+\mu\mathbf{y}\in V$. 所以$V$是$F^n$的子空间.

对线性方程组$A\mathbf{x}=\mathbf{0}$做初等行变换不改变方程组的解, 并且方程组可以化成简化形式$J\mathbf{x}=\mathbf{0}$. 其中J是阶梯形式的矩阵(2.3). 由于$\operatorname{rank}(J)=\operatorname{rank}(A)=r$, J有r个非零行. 由第一章的结论, 齐次方程组的通解具有形式

$$\begin{cases} x_1=\alpha_{11}t_1+\alpha_{12}t_2+\ldots+\alpha_{1,n-r}t_{n-r} \\ x_2=\alpha_{21}t_1+\alpha_{22}t_2+\ldots+\alpha_{2,n-r}t_{n-r} \\ \qquad\qquad\cdots\cdots\cdots\cdots\cdots\cdots \\ x_m=\alpha_{m1}t_1+\alpha_{m2}t_2+\ldots+\alpha_{m,n-r}t_{n-r} \end{cases}$$

其中$t_1,\ldots,t_{n-r}$为自由参数. 写成向量形式是

$$\mathbf{x}=t_1\alpha_1+t_1\alpha_2+\ldots+t_{n-r}\alpha_{n-r},$$

这里$\alpha_i=(\alpha_{1i},\alpha_{2i},\ldots,\alpha_{mi})^T\in F^m$, $i=1,2,\ldots,n-r$.

下面说明, $\alpha_1,\alpha_2,\ldots,\alpha_{n-r}$线性无关. 一个重要的观察是, 在齐次方程组的通解中, 除去$x_1,x_{j_2},\ldots,x_{j_r}$外, 其余变量的解为某个$t_j$, 即向量$t_1\alpha_1+t_2\alpha_2+\ldots+t_{n-r}\alpha_{n-r}$中有$n-r$个分量分别为$t_1,t_2,\ldots,t_{n-r}$. 因此, 如果

$$t_1\alpha_1+t_2\alpha_2+\ldots+t_{n-r}\alpha_{n-r}=\mathbf{0},$$

则$t_1=t_2=\ldots=t_{n-r}=0$. 这表明, $\alpha_1,\alpha_2,\ldots,\alpha_{n-r}$线性无关, 从而是$V$的一组基.

□

推论5.5.2.

$$\dim(C(A))+\dim(N(A))=n.$$

齐次线性方程组(5.16)的解集V也称为**解空间**. 解空间的一组基$\alpha_1,\alpha_2,\ldots,\alpha_{n-r}$称为齐次线性方程组的一个**基础解系**.

例5.5.1. 设A是n阶奇异方阵, 且A的(i,j)元素的代数余子式$A_{ij}\neq 0$. 证明: $\alpha=(A_{i1},A_{i2},\ldots,A_{in})^T$是齐次线性方程组$A\mathbf{x}=\mathbf{0}$的基础解系.

证明. 由于$A_{ij}\neq 0$及$\det(A)=0$, $\operatorname{rank}(A)=n-1$. 因此, 齐次线性方程组$A\mathbf{x}=\mathbf{0}$的解空间是1维的. 由行列式展开定理, $\sum_{j=1}^n a_{kj}A_{ij}=\delta_{ki}\det(A)=0$, $k=1,\ldots,n$, 所以$\alpha=(A_{i1},A_{i2},\ldots,A_{in})^T$ 是齐次方程组$A\mathbf{x}=\mathbf{0}$的解. 由于α为非零向量, 故它是基础解系.

□

例5.5.2. 求下列齐次线性方程组的一组基础解系, 并求它的通解

$$\begin{cases} x_1+x_2+x_3+x_4+x_5=0 \\ x_1+2x_2+3x_3-2x_4+2x_5=0 \\ x_1-x_3+4x_4=0 \end{cases}$$

解.

$$\begin{pmatrix} 1 & 1 & 1 & 1 & 1 \\ 1 & 2 & 3 & -2 & 2 \\ 1 & 0 & -1 & 4 & 0 \end{pmatrix} \xrightarrow[-r_1\to r_2]{-r_1\to r_3} \begin{pmatrix} 1 & 1 & 1 & 1 & 1 \\ 0 & 1 & 2 & -3 & 1 \\ 0 & -1 & -2 & 3 & -1 \end{pmatrix} \xrightarrow[-r_2\to r_1]{r_2\to r_3}$$

$$\begin{pmatrix} 1 & 0 & -1 & 4 & 0 \\ 0 & 1 & 2 & -3 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

因此原方程组等价于

$$\begin{cases} x_1-x_3+4x_4=0 \\ x_2+2x_3-3x_4+x_5=0 \end{cases}$$

取$x_3=1, x_4=x_5=0$得方程的一组解$\alpha_1=(1,-2,1,0,0)^T$. 类似地, 分别取$x_4=1, x_3=x_5=0$及$x_5=1, x_3=x_4=0$得方程组的另两组解$\alpha_2=(-4,3,0,1,0)^T$与$\alpha_3=(0,-1,0,0,1)^T$. 则$\alpha_1,\alpha_2,\alpha_3$构成原线性方程组的一组基础解系, 原方程组的通解为$t_1\alpha_1+t_2\alpha_2+t_3\alpha_3$, t_1,t_2,t_3为任意参数.

当然, 也可以直接令$x_3=t_1, x_4=t_2, x_5=t_3$求得$x_1=t_1-4t_2$, $x_2=-2t_1+3t_2-t_3$. 故

$$(x_1,x_2,x_3,x_4,x_5)^T=t_1\alpha_1+t_2\alpha_2+t_3\alpha_3.$$

由此可求方程组的基础解系与通解. □

§5.5.3 非齐次线性方程组解集的结构

本小节我们研究非齐次线性方程组的解集的几何结构. 从例 5.4.7看出, 非齐次线性方程组的解集不再是一个子空间, 但它与子空间有密切关系.

考察线性方程$ax+by+cz=d(d\neq 0)$所确定的点集W, 它表示$\mathbb{R}^3$中不过原点的平面, 因而$\{\overrightarrow{OP}\,|\,P\in W\}$不是一个子空间. 相应的齐次方程$ax+by+cz=0$所确定的点集$V$表示$\mathbb{R}^3$中过原点的平面, 它对应一个子空间. W是对V沿向量$(0,0,d/c)$平移的结果. 这个观察对一般非齐次线性方程组的解集也是成立的.

考虑非齐次线性方程组(5.15)的解集. 记

$$W=\{\mathbf{x}\in F^n\,|\,A\mathbf{x}=\mathbf{b}\}. \tag{5.18}$$

V如(5.17)所示. 我们知道, V是$n-r(A)$维子空间. 那么, W与V有什么关系?

容易验证, W具有性质

1. $\alpha,\beta \in W \Rightarrow \alpha-\beta \in V$.
2. $\alpha \in W,\ \gamma \in V \Rightarrow \alpha+\gamma \in W$.

事实上, 设$\alpha,\beta \in W$, 则$A\alpha=\mathbf{b}, A\beta=\mathbf{b}$. 于是, $A(\alpha-\beta)=\mathbf{0}$, 即$\alpha-\beta \in V$. 另一方面, 如果$\alpha \in W,\ \gamma \in V$, 则$A\alpha=\mathbf{b}, A\gamma=\mathbf{0}$. 因此, $A(\alpha+\gamma)=\mathbf{b}$, 即$\alpha+\gamma \in W$.

利用上述性质可以得到

定理5.5.3. *设V,W分别(5.17), (5.18)由定义. 则有*

$$W=\gamma_0+V:=\{\gamma_0+\alpha \mid \alpha \in V\}.$$

其中γ_0是$A\mathbf{x}=\mathbf{b}$的一个特解.

证明. 对任意$\alpha \in V$, $A(\gamma_0+\alpha)=A\gamma_0+A\alpha=\mathbf{b}$. 因此, $\gamma_0+\alpha \in W$.

反过来, 任取$\beta \in W$, 由$A\beta=\mathbf{b}$及$A\gamma_0=\mathbf{b}$有, $A(\beta-\gamma_0)=\mathbf{0}$. 所以, $\beta-\gamma_0 \in V$. 从而$\beta \in \gamma_0+V$. □

上述定理表明, 非齐次线性方程组的解集是相应齐次方程组的解集的一个平移. 假设$\alpha_1,\ldots,\alpha_{n-r}$是对应的齐次方程组的基础解系, γ_0是非齐次线性方程组的特解, 则非齐次线性方程组的通解为

$$\mathbf{x}=t_1\alpha_1+t_2\alpha_2+\ldots+t_{n-r}\alpha_{n-r}+\gamma_0.$$

其中$t_1,t_2,\ldots,t_{n-r}$为自由参数. 或者说, 非齐次线性方程组的解集为

$$W=\{t_1\alpha_1+t_2\alpha_2+\ldots+t_{n-r}\alpha_{n-r}+\gamma_0 \,|\, t_1,t_2,\ldots,t_{n-r} \in F\}. \tag{5.19}$$

例5.5.3. 给出线性方程组$A\mathbf{x}=\mathbf{b}$的公式解.

解. 设$A \in F^{m\times n}$, $\mathbf{b} \in F^m$, 且$\mathrm{rank}(A)=\mathrm{rank}(A,\mathbf{b})=r$. 不妨设$A$的前$r$行、前$r$列构成的子式非零. 则线性方程组$A\mathbf{x}=\mathbf{b}$等价于$\bar{A}\mathbf{x}=\bar{\mathbf{b}}$, 这里$\bar{A}$是由$A$的前$r$行组成的矩阵, $\bar{\mathbf{b}}$是$\mathbf{b}$的前r个分量组成的向量. 记$\bar{A}=(A_1,A_2)$, $\mathbf{x}=(\mathbf{x}_1,\mathbf{x}_2)^T$. 这里$A_1 \in F^{r\times r}, A_2 \in F^{r\times(n-r)}$, $\mathbf{x}_1 \in F^r, \mathbf{x}_2 \in F^{n-r}$. 于是, 线性方程组$\bar{A}\mathbf{x}=\bar{\mathbf{b}}$又可以改写成

$$A_1\mathbf{x}_1=\bar{b}-A_2\mathbf{x}_2.$$

由于A_1可逆, 则

$$\mathbf{x}_1 = A_1^{-1}(\bar{b} - A_2\mathbf{x}_2) = \frac{1}{\det(A_1)}A_1^*(\bar{b} - A_2\mathbf{x}_2) = -\frac{1}{\det(A_1)}A_1^*A_2\mathbf{x}_2 + \frac{1}{\det(A_1)}A_1^*\bar{b}.$$

这里A_1^*是A_1的伴随方阵, $\mathbf{x}_2$是可任意取值的$n-r$维向量. 于是

$$\mathbf{x} = \begin{pmatrix} \mathbf{x}_1 \\ \mathbf{x}_2 \end{pmatrix} = \begin{pmatrix} -A_1^*A_2/\det(A_1) \\ I_{n-r} \end{pmatrix}\mathbf{t} + \begin{pmatrix} A_1^*\bar{b}/\det(A_1) \\ O \end{pmatrix},$$

这里$\mathbf{t} \in F^{n-r}$为独立参数. □

§5.6 一般线性空间

本节中, 我们将数组空间的概念与理论推广到一般的线性空间.

§5.6.1 一般线性空间的定义

在前面几节中, 我们研究了n维数组空间的理论. 这些理论建立在数组向量的两个基本运算之上, 即向量的加法与数乘. 简而言之, n维数组空间就是赋予了加法与数乘运算的n数组向量的集合. 一个基本的问题是, 对一个集合, 是否可以恰当地定义其中元素的加法与数乘运算, 从而具有n维数组空间一样的性质与结构呢? 我们先看一些实例.

例5.6.1. 考察具有n个变元的线性方程的全体

$$E_n := \{a_1x_1 + a_2x_2 + \ldots + a_nx_n = b \,|\, a_1, a_2, \ldots, a_n, b \in F\}.$$

定义方程的加法(用$\oplus$表示)与数乘(用$\circ$表示)如下

$$\begin{aligned}&(a_1x_1 + a_2x_2 + \ldots + a_nx_n = b) \oplus (a_1'x_1 + a_2'x_2 + \ldots + a_n'x_n = b') \\ &:= (a_1 + a_1')x_1 + (a_2 + a_2') + \ldots + (a_n + a_n')x_n = b + b'. \\ &\lambda \circ (a_1x_1 + a_2x_2 + \ldots + a_nx_n = b) := \lambda a_1x_1 + \lambda a_2x_2 + \ldots + \lambda a_nx_n = \lambda b.\end{aligned}$$

有了加法与数乘运算, 就可以定义方程组的线性组合、线性相关、线性无关、极大无关组、秩等概念. 对线性方程组做初等变换就是做线性组合运算, 线性方程组的极大无关组就是等价的独立方程组, 秩就是独立方程的个数.

例5.6.2. 用$F_n[x]$表示数域F上次数不超过n的多项式全体

$$F_n[x] := \{a_0 + a_1x + \ldots + a_nx^n \,|\, a_0, a_1, \ldots, a_n \in F\}.$$

定义多项式的加法与数乘如下

$$
\begin{aligned}
&(a_0+a_1x+\ldots+a_nx^n)+(b_0+b_1x+\ldots+b_nx^n)\\
&=(a_0+b_0)+(a_1+b_1)x+\ldots+(a_n+b_n)x^n,\\
&\lambda(a_0+a_1x+\ldots+a_nx^n)=\lambda a_0+\lambda a_1x+\ldots+\lambda a_nx^n
\end{aligned}
$$

由此可以定义多项式的线性相关(无关)、极大无关组、秩、基等概念, 并且相应于数组空间的有关定理与结论仍然成立. 例如, 多项式$p_1(x)=x^2-2x+1$, $p_2(x)=x^2+2$, $p_3(x)=x^2+2x+3$线性相关, 因为$p_3(x)-2p_2(x)+p_1(x)=0$. 再如$S:=\{1,x,x^2\}$线性无关(因为不存在不全为零常数a_0,a_1,a_2使$a_0+a_1x+a_2x^2=0$), 且它们构成$F_2[x]$的一组基. 这表明, n次多项式全体在定义了加法与数乘运算后与数组空间有相似的结构与性质.

例5.6.3. 对所有$m\times n$阶矩阵的全体$F^{m\times n}$, 按照矩阵的加法与数乘, 也可以定义一组矩阵的线性相关(无关)、极大无关组、秩、基等, 并且数组空间的有关结论可以平移过来. 例如, 矩阵

$$
\begin{pmatrix}1&0\\0&0\end{pmatrix},\quad \begin{pmatrix}0&1\\0&0\end{pmatrix},\quad \begin{pmatrix}0&0\\1&0\end{pmatrix},\quad \begin{pmatrix}0&0\\0&1\end{pmatrix}
$$

线性无关, 且它们构成$F^{2\times 2}$的一组基.

上述实例表明, 对任何一个集合V, 在恰当定义了元素的加法与数乘运算, 并且运算满足一定的规律后, V就有了像数组空间一样的结构与性质及相关理论. 这样, 我们可以将数组空间推广到一般的线性空间, 从而大大扩充有关数学理论的应用范围. 在进入以后的学习中, 读者会见到各种各样的线性空间, 并会发现线性空间理论在许多工程领域有着广泛的应用. 下面, 我们给出一般线性空间的定义.

定义5.6.1. 设V是一个非空集合, F是一个数域. 对V中的元素定义两种运算

1. **加法**: 对V中的任意两个元素α,β组成的有序对(α,β), 存在V中唯一的一个元素γ与之对应, 简记为$\alpha+\beta=\gamma$.
2. **数乘**: 对任意常数$\lambda\in F$及向量$\alpha\in V$, 存在V中唯一的一个元素γ与之对应, 简记为$\lambda\alpha=\gamma$.

加法与数乘运算满足下列运算规律

(A1) $\alpha+\beta=\beta+\alpha$ 对任意$\alpha,\beta\in V$成立.

(A2) $(\alpha+\beta)+\gamma=\alpha+(\beta+\gamma)$ 对任意$\alpha,\beta,\gamma\in V$成立.

(A3) 存在元素$\theta\in V$, 使得$\alpha+\theta=\theta+\alpha=\alpha$对任意$\alpha\in V$成立. θ 称为零元素. 在不致混淆的情况下, 一般线性空间中的零元素也常简记为0.

(A4) 对任意$\alpha\in V$, 存在$\beta\in V$使得$\alpha+\beta=\beta+\alpha=0$. β称为α的负元素, 简记为$-\alpha$, 并且定义$\beta-\alpha=\beta+(-\alpha)$.

(D1) $\lambda(\alpha+\beta)=\lambda\alpha+\lambda\beta$对任意$\lambda\in F$及$\alpha,\beta\in V$成立.

(D2) $(\lambda+\mu)\alpha=\lambda\alpha+\mu\alpha$对任意$\lambda,\mu\in F$及$\alpha\in V$成立.

(M1) $\lambda(\mu\alpha)=(\lambda\mu)\alpha$对任意$\lambda,\mu\in F$及$\alpha\in V$成立.

(M2) $1\alpha=\alpha$对任意$\alpha\in V$成立.

则称V是数域F上的**线性空间**, 简记为$V(F)$或V. 线性空间V中的元素称为**向量**.

对上述定义我们做两点说明.

关于加法与数乘两种运算可以用映射的观点解释. 用$V\times V$表示所有有序向量对的集合, 即$V\times V:=\{(\alpha,\beta)\,|\,\alpha,\beta\in V\}$. 类似地, 记$F\times V:=\{(\lambda,\alpha)\,|\,\lambda\in F,\alpha\in V\}$. 所谓$V$中的加法实际上是从$V\times V$到$V$的一个映射. 而数乘是$F\times V$到$V$的一个映射.

接下来我们要说明为什么要求加法与数乘运算要满足上述八条规律.

容易知道, 数组向量满足上述八条规律. 这些运算规律保证了后续一系列定理与结论的正确推导. 我们当然希望一般线性空间也有数组空间一样的性质与结论, 因此对这两种运算就要求满足一定的规则. 下面, 我们从以下这样一个命题加以说明.

设$\alpha_1,\alpha_2,\ldots,\alpha_m$为线性空间$V$中的向量, 若存在不全为零的数$\lambda_1,\lambda_2,\ldots,\lambda_m$使得 $\lambda_1\alpha_1+\lambda_2\alpha_2+\cdots+\lambda_m\alpha_m=\theta$, 则存在$\alpha_i(1\le i\le m)$使得

$$\alpha_i=-\frac{\lambda_1}{\lambda_i}\alpha_1-\ldots-\frac{\lambda_{i-1}}{\lambda_i}\alpha_{i-1}-\frac{\lambda_{i+1}}{\lambda_i}\alpha_{i+1}-\ldots-\frac{\lambda_m}{\lambda_i}\alpha_m.$$

设$\lambda_i \neq 0$. 则

$$\lambda_1\alpha_1+\lambda_2\alpha_2+\cdots+\lambda_m\alpha_m=\theta \xrightarrow{(A3)}$$
$$(\lambda_1\alpha_1+\lambda_2\alpha_2+\cdots+\lambda_m\alpha_m)+(-\lambda_i)\alpha_i=(-\lambda_i)\alpha_i \xrightarrow{(A1,A2)}$$
$$\lambda_1\alpha_1+\lambda_2\alpha_2+\ldots+\lambda_{i-1}\alpha_{i-1}+\lambda_{i+1}\alpha_{i+1}+\ldots+\lambda_m\alpha_m+$$
$$(\lambda_i\alpha_i+(-\lambda_i)\alpha_i)=(-\lambda_i)\alpha_i \xrightarrow{(D2,A3,*)} \quad (\text{这里用到性质}*\text{: } 0\alpha=\theta)$$
$$\lambda_1\alpha_1+\ldots+\lambda_{i-1}\alpha_{i-1}+\lambda_{i+1}\alpha_{i+1}+\ldots+\lambda_m\alpha_m=(-\lambda_i)\alpha_i \to$$
$$(-\frac{1}{\lambda_i})(-\lambda_i)\alpha_i=(-\frac{1}{\lambda_i})(\lambda_1\alpha_1+\ldots+\lambda_{i-1}\alpha_{i-1}+\lambda_{i+1}\alpha_{i+1}+\ldots+\lambda_m\alpha_m)$$
$$\xrightarrow{(M1,M2,D1,A4)} \alpha_i=-\frac{\lambda_1}{\lambda_i}\alpha_1-\ldots-\frac{\lambda_{i-1}}{\lambda_i}\alpha_{i-1}-\frac{\lambda_{i+1}}{\lambda_i}\alpha_{i+1}-\ldots-\frac{\lambda_m}{\lambda_i}\alpha_m.$$

我们看到, 要证明上述命题, 八条运算规律都用到了. 实际上, 这八条规律也足够证明所有相关结论.

下面我们列举线性空间的一些**基本性质**.

1. 零向量唯一.

   设θ_1与θ_2都为零向量, 则有$\theta_1=\theta_1+\theta_2=\theta_2+\theta_1=\theta_2$.

2. 负向量唯一.

   设β_1,β_2均为α的负向量. 则$\beta_1=\beta_1+\theta=\beta_1+(\alpha+\beta_2)=\theta+\beta_2=\beta_2$.

3. $0\alpha=\theta$; $(-1)\alpha=-\alpha$; $\lambda\theta=\theta$.

   由$\alpha=(1+0)\alpha=\alpha+0\alpha$得, $0\alpha=\theta$.

   由$\theta=(1-1)\alpha=\alpha+(-1)\alpha$得, $(-1)\alpha=-\alpha$.

   最后, $\lambda\theta=\lambda(0\alpha)=(\lambda\cdot 0)\alpha=0\alpha=\theta$.

4. $\lambda\alpha=\theta$ 当且仅当$\lambda=0$或$\alpha=\theta$.

   由3, 若$\lambda=0$或$\alpha=\theta$, 则显然有$\lambda\alpha=\theta$. 现设$\lambda\neq 0$, $\lambda\alpha=\theta$, 则$\alpha=(1/\lambda)(\lambda\alpha)=(1/\lambda)\theta=\theta$.

接下来, 我们给出一些线性空间的实例.

例5.6.4. 一些线性空间的例子, 其中F为数域.

(1) 数组空间F^n按数组向量的加法与数乘构成线性空间.

(2) 所有n元的一次方程全体E_n按线性方程的加法与数乘构成线性空间.

(3) 所有次数不超过n的多项式全体$F_n[x]$按多项式的加法, 及数与多项式的乘法为数乘构成线性空间.

(4) 所有$m\times n$阶矩阵全体$F^{m\times n}$按矩阵的加法与数乘构成线性空间.

(5) $F=\mathbf{R}$, 所有复数全体$\mathbf{C}$按复数的加法、实数与复数的乘法为数乘构成线性空间.

(6) 用C_n表示如下三角多项式全体

$$C_n=\{a_0+a_1\cos\theta+b_1\sin\theta+\ldots+a_n\cos(n\theta)+b_n\sin(n\theta)\,|\,a_i,b_i\in F\}.$$

则在函数通常的加法、数与函数的乘法下构成线性空间. 特别地, $\cos^3\theta\in C_3$, 因为$\cos^3\theta=\frac{3}{4}\cos\theta+\frac{1}{4}\cos(3\theta)$.

(7) 用$C[a,b]$表示$[a,b]$上连续函数的全体. 则$C[a,b]$在函数通常的加法, 及数与函数乘法运算下构成线性空间.

下面我们引进一般线性空间的子空间的概念.

定义5.6.2. 设V是数域F上的线性空间, W是V的非空子集. 如果W对于线性空间V的加法与数乘运算保持封闭, 即

1. 对任意$\alpha,\beta\in W$, 有$\alpha+\beta\in W$;

2. 对任意$\lambda\in F,\alpha\in W$, 有$\lambda\alpha\in W$.

则称W是V的**子空间**.

容易验证, W在V中的加法与数乘下构成数域F上的线性空间. 特别地, $W=\{0\}$及$W=V$也是V 的子空间, 称为**平凡子空间**. 一类非平凡的子空间是生成子空间.

定义5.6.3. 设V是数域F上的线性空间, $S\subset V$是非空集合. 则集合

$$\langle S\rangle:=\{\lambda_1\alpha_1+\lambda_2\alpha_2+\ldots+\lambda_m\alpha_m\,|\,\lambda_1,\lambda_2,\ldots,\lambda_m\in F,\ \alpha_1,\alpha_2,\ldots,\alpha_m\in S\} \tag{5.20}$$

是一个子空间, 称为V的**生成子空间**, S称为生成子空间的**生成元**.

下面, 我们给几个子空间的实例.

例5.6.5. 子空间的例子.

(1) 在例 5.6.4中, $F_n[x]$及C_n都是$C[a,b]$的子空间.

(2) 设W 是$F_n[x]$中满足$p(-x)=p(x)$的多项式全体, 则W是$F_n[x]$的子空间, 且$W=\langle 1,x^2,\ldots,x^{2m}\rangle$, 其中$m=[n/2]$.

(3) 设$A\in F^{n\times n}$, 则集合$W:=\{X\in F^{n\times n}\,|\,AX=0\}$是$F^{n\times n}$ 的子空间.

证明. 我们以(2)为例给出证明. 设$p,q\in W$, 则$p(-x)=p(x),q(-x)=q(x)$. 于是

$$(p+q)(-x)=p(-x)+q(-x)=p(x)+q(x)=(p+q)(x),$$

即$p+q\in W$. 另一方面,

$$(\lambda p)(-x)=\lambda p(-x)=\lambda p(x)=(\lambda p)(x).$$

因此W在加法与数乘下封闭, 从而是$F_n[x]$的子空间.

另一方面, 由于W中多项式都是偶多项式, 故$W=\{a_0+a_2x^2+\ldots+a_{2m}x^{2m}\}$, 其中$m=[n/2]$. 因此, $W=\langle 1,x^2,\ldots,x^{2m}\rangle$. □

最后对线性空间再做一点说明. 同三维几何空间不同, 一般线性空间中只定义了向量的加法与数乘运算, 没有定义向量长度、夹角等几何概念.

§5.6.2 一般线性空间的理论

本小节中, 我们将线性组合、线性相关、线性无关、极大无关组与秩、基与维数等概念推广到一般线性空间. 并且关于数组空间的有关结论都可以平行推广过来, 因为这些定义与结论只涉及到向量的加法与数乘, 而不管向量本身是什么, 向量的运算如何进行. 我们将主要的概念与结论重新叙述一遍, 但省略有关证明.

定义5.6.4. 设V是数域F上的线性空间. 给定V中的一组向量$S=\{\alpha_1,\alpha_2,\ldots,\alpha_m\}$及一组数$\lambda_1,\lambda_2,\ldots,\lambda_m\in F$, 称和式

$$\lambda_1\alpha_1+\lambda_2\alpha_2+\ldots+\lambda_m\alpha_m$$

为向量组S的**线性组合**, $\lambda_1,\lambda_2,\ldots,\lambda_m$称为**组合系数**. 如果$\alpha$可以写成$S$的线性组合, 则称$\alpha$可以用$S$ **线性表示**.

显然, α可以用S 线性表示当且仅当$\alpha\in\langle S\rangle$.

定义5.6.5. 设V是数域F上的线性空间. 称向量组$T=\{\beta_1,\beta_2,\cdots,\beta_l\}$可以由向量组$S=\{\alpha_1,\alpha_2,\cdots,\alpha_m\}$ **线性表示**, 如果每一个$\beta_i(i=1,2,\ldots,l)$可以用向量组S线性表示. 如果向量组S与T可以相互线性表示, 则称S与T **等价**.

关于向量组的线性表示与等价有下列结论

定理5.6.1. 设S与T是线性空间V的两个向量组. 则

1. T是S的线性组合, 当且仅当$\langle T\rangle\subset\langle S\rangle$.
2. U是T的线性组合, T是S的线性组合, 则U是S的线性组合.
3. S与T等价, 当且仅当$\langle S\rangle=\langle T\rangle$.
4. U与T等价, T与S等价, 则U与S等价.

定义5.6.6. 设V是数域F上的线性空间, S 是V中一组向量. 当$\#S\geq 2$($\#S$表示S中元素的个数)时, 如果S中某个向量能用其它向量线性表示, 则称S **线性相关**. 否则, 称它们**线性无关**. 特别地, 一个向量组成的向量组**线性相关**, 当且仅当该向量为零向量.

关于向量组的线性相关性, 有下列等价的说法

定理5.6.2. 设$\alpha_1,\alpha_2,\cdots,\alpha_m$是线性空间$V$中的向量, 则下列说法等价

1. $\alpha_1,\alpha_2,\cdots,\alpha_m$线性相关;
2. 存在不全为零的常数$\lambda_1,\lambda_2,\cdots,\lambda_m\in F$, 使得$\sum\limits_{i=1}^{m}\lambda_i\alpha_i=\theta$;
3. 存在向量α_i使得$\alpha_i=\sum\limits_{j\neq i}\lambda_j\alpha_j$;
4. 存在向量α_i使得$\alpha_i\in\langle\alpha_1,\ldots,\alpha_{i-1},\alpha_{i+1},\ldots,\alpha_m\rangle$;
5. 存在α_i使得$\langle\alpha_1,\alpha_2,\ldots,\alpha_m\rangle=\langle\alpha_1,\ldots,\alpha_{i-1},\alpha_{i+1},\ldots,\alpha_m\rangle$.

定理5.6.3. 设向量组S_1是向量组S的一个子集. 那么, 如果S_1线性相关, 则S也线性相关; 如果S线性无关, 则S_1也线性无关.

定义5.6.7. 设S是线性空间V中的向量组. 若S的子集S_1线性无关, 且对任意$\alpha\in S\backslash S_1$, $S_1\cup\{\alpha\}$线性相关, 则称S_1为向量组S的**极大无关组**.

定理5.6.4. 向量组的极大无关组有下列等价的说法

1. 向量组S的子集S_1是S的极大无关组;

2. 向量组S可以由子集S_1线性表示, 且S_1线性无关;

3. 向量组S与它的子集S_1等价, 且S_1线性无关;

4. $\langle S\rangle = \langle S_1\rangle$, 且$S_1$线性无关.

推论5.6.1. 向量组的任何两个极大无关组彼此等价.

定理5.6.5. 两个等价向量组$\{\alpha_1, \alpha_2, \cdots, \alpha_r\}$和$\{\beta_1, \beta_2, \cdots, \beta_s\}$ 分别线性无关, 则$r = s$.

推论5.6.2. 设$\alpha_{i_1}, \alpha_{i_2}, \cdots, \alpha_{i_r}$和$\alpha_{j_1}, \alpha_{j_2}, \cdots, \alpha_{j_s}$分别为$\alpha_1, \alpha_2, \cdots, \alpha_m$的两个极大无关组, 则$r = s$.

定义5.6.8. 向量组S的极大无关组的向量的个数称为向量组的**秩**, 记为$\mathrm{rank}(S)$或$r(S)$.

关于秩, 我们有下列有用的结论.

定理5.6.6. 设S, T是线性空间V中的向量组. 则有下列结论

1. S线性无关当且仅当$\mathrm{rank}(S) = \#S$, 这里$\#S$表示S中向量的个数.

2. S线性相关当且仅当$\mathrm{rank}(S) < \#S$.

3. 若T可以用S线性表示, 则$\mathrm{rank}(T) \leq \mathrm{rank}(S)$.

4. 若T与S等价, 则$\mathrm{rank}(T) = \mathrm{rank}(S)$.

5. 若T可以用S线性表示, 且T线性无关, 则$\#T \leq \#S$.

例5.6.6. 设C_n是例5.6.4中由三角多项式构成的线性空间. 求向量组$S = \{1, \sin x, \cos x, \sin^2 x, \cos^2 x, \sin 2x, \cos 2x\}$的一组极大无关组.

解. 由于$\cos^2 x = \frac{1}{2} + \frac{1}{2}\cos 2x$, $\sin^2 x = \frac{1}{2} - \frac{1}{2}\cos 2x$, 因此, S等价于$S_1 = \{1, \cos x, \sin x, \cos 2x, \sin 2x\}$. 下面证明, S_1线性无关, 从而是S的极大无关组.

设$a_0 + a_1\cos x + b_1\sin x + a_2\cos 2x + b_2\sin 2x = 0$. 分别令$x = 0, \pi/4, \pi/2, 3\pi/4, \pi$得

$$\begin{cases} a_0 + a_1 + a_2 = 0 \\ a_0 + \dfrac{\sqrt{2}}{2}a_1 + \dfrac{\sqrt{2}}{2}b_1 + b_2 = 0 \\ a_0 + b_1 - a_2 = 0 \\ a_0 - \dfrac{\sqrt{2}}{2}a_1 + \dfrac{\sqrt{2}}{2}b_1 - b_2 = 0 \\ a_0 - a_1 + a_2 = 0 \end{cases}$$

解得$a_0 = a_1 = a_2 = b_1 = b_2 = 0$. 从而$S_1$线性无关(也可以用Fourier分析的理论证明$S_1$线性无关).

由于$S_2 = \{1, \cos x, \sin x, \sin^2 x, \sin 2x\}$与$S_1$等价，因此，$S_2$也是一个极大无关组. □

例5.6.7. 设线性空间V的向量组S含m个向量, 且其秩为r. 在S中任取s个向量组成向量组T. 证明: $\mathrm{rank}(T) \geq r+s-m$.

证明. 若$r+s-m \leq 0$, 结论显然成立. 下设$r+s-m > 0$. 设S_1为S的极大无关组, 则T至少含$s-(m-r) = r+s-m$个S_1中的向量, 这$r+s-m$个向量线性无关. 因此, $\mathrm{rank}(T) \geq r+s-m$. □

下面我们引进线性空间的基、坐标与维数的概念.

定义5.6.9. 设V是数域F上的线性空间. S是V中一组线性无关向量. 如果V中任何向量都能表示成S的线性组合, 则称S为V的一组**基**. 若S是有限的, 则称V为**有限维线性空间**, S中元素的个数称为线性空间V的**维数**, 记为$\dim V$. 不是有限维的线性空间称为**无限维线性空间**, 其维数为无穷大. 设基$S = \{\alpha_1, \alpha_2, \ldots, \alpha_n\}$是有限的, 则任意向量$\alpha \in V$可以唯一地表示成$S$的线性组合

$$\alpha = \lambda_1\alpha_1 + \lambda_2\alpha_2 + \ldots + \lambda_n\alpha_n. \tag{5.21}$$

称$(\lambda_1, \lambda_2, \ldots, \lambda_n)$为向量$\alpha$在基$S$下的**坐标**.

除非特别申明, 本书只考虑有限维线性空间. 有限维的线性空间一定可以由一组向量生成.

(5.21)可以写成向量形式

$$\alpha = (\alpha_1, \alpha_2, \ldots, \alpha_n)\begin{pmatrix}\lambda_1\\ \lambda_2\\ \vdots\\ \lambda_n\end{pmatrix}. \tag{5.22}$$

注意, 上式完全是一个形式上的记号, 因为$(\alpha_1, \alpha_2, \ldots, \alpha_n)$可能没有实质上的意义, 但运算规律跟矩阵运算相同.

定理5.6.7. 设V是数域F上的有限维线性空间. 则存在线性无关向量组$\alpha_1, \alpha_2, \ldots, \alpha_n \in V$使得$V = \langle\alpha_1, \alpha_2, \ldots, \alpha_n\rangle$.

证明. 同定理 5.4.1的证明一样. □

显然, 基实际上是空间作为向量组的极大无关组. 而空间V的维数实际上是V作为向量组的秩. 关于线性空间的维数与基有下列结论.

定理5.6.8. 设V是数域F上的n维线性空间. 则有

1. V中任意$n+1$个向量线性相关.
2. V中任意n个线性无关向量为一组基.
3. 设$\alpha_1,\alpha_2,\dots,\alpha_r \in V$是$r(r<n)$个线性无关的向量, 则存在$V$中的向量$\alpha_{r+1},\dots,\alpha_n$使得$\alpha_1,\alpha_2,\dots,\alpha_n$构成$V$的一组基. 称$\alpha_1,\alpha_2,\dots,\alpha_n$为线性无关组$\alpha_1,\alpha_2,\dots,\alpha_r$的一组**扩充基**.

线性空间的基与坐标的变换, 同数组空间完全类似, 不再赘述.

例5.6.8. 设$\mathbf{C}$是例5.6.4中定义的复数构成的线性空间. 求它的一组基与维数.

解. 首先证明$\{1,i\}$是$\mathbf{C}$的一组基. 实际上, 设有实数a,b使得$a+bi=0$, 则必有$a=b=0$, 即$\{1,i\}$在$\mathbf{R}$上线性无关. 另一方面, 任何复数都可以写成$\{1,i\}$在$\mathbf{R}$上的线性组合$a+bi$. 因此, $\{1,i\}$构成$\mathbf{C}$的一组基. 从而$\dim\mathbf{C}=2$.

如果数域$F=\mathbf{C}$, 则任何非零数复数都是$\mathbf{C}$的一个基, 从而其维数为1. □

例5.6.9. 考虑例5.6.4中$m\times n$阶矩阵构成的线性空间$F^{m\times n}$. 求$F^{m\times n}$的一组基与维数.

解. 用$E_{ij}\in F^{m\times n}$表示(i,j)元素为1, 其余元素为零的矩阵. 先证E_{ij}, $i=1,2,\dots,m$, $j=1,2,\dots,n$线性无关. 由

$$\sum_{i=1}^{m}\sum_{j=1}^{n}\lambda_{ij}E_{ij}=(\lambda_{ij})_{m\times n}=0$$

知$\lambda_{ij}=0$, $i=1,2,\dots,m$, $j=1,2,\dots,n$, 从而$\{E_{ij}\}$线性无关. 另一方面, 任何矩阵$A=(a_{ij})\in F^{m\times n}$可以写成$\{E_{ij}\}$的线性组合

$$A=\sum_{i=1}^{m}\sum_{j=1}^{n}a_{ij}E_{ij}.$$

因此, $\{E_{ij}\}$构成$F^{m\times n}$的一组基, 并且$\dim F^{m\times n}=mn$. □

§5.7 子空间运算*

本节我们讨论一般线性空间的子空间的运算, 如求交, 求和等.

§5.7.1 子空间的交*

设W_1,W_2是数域F上的线性空间V的子空间. 我们要问, W_1与W_2的交集还是不是子空间? 如果是的话, 如何求它的基与维数?

对于三维几何空间, 两个过原点的平面对应两个子空间, 它们的交集是过原点的直线或平面, 因而也对应子空间. 这个观察对一般线性空间也成立.

定理5.7.1. 设W_1,W_2是数域F上的线性空间V的子空间, 则W_1与W_2的**交** $W_1\cap W_2$也是V的子空间.

证明. 设$\alpha,\beta\in W_1\cap W_2$, 则$\alpha,\beta\in W_1$且$\alpha,\beta\in W_2$. 由于$W_1,W_2$是子空间, $\alpha+\beta\in W_1$且$\alpha+\beta\in W_2$. 因此, $\alpha+\beta\in W_1\cap W_2$.

另一方面, 若$\lambda\in F$, $\alpha\in W_1\cap W_2$, 则$\lambda\alpha\in W_1$且$\lambda\alpha\in W_2$. 因而$\lambda\alpha\in W_1\cap W_2$. 这表明$W_1\cap W_2$是$V$的子空间. □

上述定理的结论可以推广到任意多个子空间的情形, 即任意个子空间的交为子空间.

例5.7.1. 设$V=F^4$, $W_1=\langle\alpha_1,\alpha_2,\alpha_3\rangle$, $W_2=\langle\beta_1,\beta_2\rangle$. 其中$\alpha_1=(1,2,1,0)$, $\alpha_2=(-1,1,1,1)$, $\alpha_3=(0,3,2,1)$, $\beta_1=(2,-1,0,1)$, $\beta_2=(1,-1,3,7)$. 求$W_1\cap W_2$.

解. 任取$\alpha\in W_1\cap W_2$, 则

$$\alpha=x_1\alpha_1+x_2\alpha_2+x_3\alpha_3=y_1\beta_1+y_2\beta_2.$$

上式是关于x_1,x_2,x_3,y_1,y_2的线性方程组, 解得$(y_1,y_2)=(3,-1)t$, t为任意参数. 于是$\alpha=y_1\beta_1+y_2\beta_2=(5,-2,-3,-4)t$. 因此, $W_1\cap W_2=\langle\alpha\rangle$, 这里$\alpha=(5,-2,-3,-4)$. □

§5.7.2 子空间的和*

接下来, 我们讨论子空间的和. 设W_1,W_2是V的子空间, $W_1\cup W_2$是不是子空间? 从三维几何空间观察, 结论是否定的. 例如, 过原点的两条不同直线的并不能对应子空间, 包含这两条直线的子空间至少要包含这两条直线所在的平面. 一般地, 要想构造包含$W_1\cup W_2$的子空间W, W必须包含W_1中的任何向量α_1与W_2中的任何向量α_2的和. 恰巧, 由W_1中的任何向量α_1与W_2中的任何向量α_2的和构成的集合是子空间.

定理5.7.2. 设W_1,W_2是线性空间V的子空间. W_1与W_2的**和**

$$W_1+W_2:=\{\alpha_1+\alpha_2\,|\,\alpha_1\in W_1,\alpha_2\in W_2\} \tag{5.23}$$

构成V的子空间, 并且它是包含$W_1\cup W_2$的最小子空间.

证明. 设$\alpha := \alpha_1+\alpha_2 \in W_1+W_2$, $\beta := \beta_1+\beta_2 \in W_1+W_2$, 这里$\alpha_1,\beta_1 \in W_1$, $\alpha_2,\beta_2 \in W_2$. 由于$\alpha_1+\beta_1 \in W_1$, $\alpha_2+\beta_2 \in W_2$, 因此$\alpha+\beta = (\alpha_1+\beta_1)+(\alpha_2+\beta_2) \in W_1+W_2$. 另一方面, 对$\lambda \in F$, $\lambda\alpha = \lambda\alpha_1+\lambda\alpha_2 \in W_1+W_2$. 由于$W_1+W_2$在加法与数乘下封闭, 因此$W_1+W_2$是子空间.

接下来证明W_1+W_2是包含$W_1\cup W_2$的最小子空间. 对任意$\alpha_1 \in W_1$, $\alpha_1 = \alpha_1+0 \in W_1+W_2$, 因此$W_1 \subset W_1+W_2$. 同理, $W_2 \subset W_1+W_2$. 故$W_1\cup W_2 \subset W_1+W_2$. 下设$W$是包含$W_1\cup W_2$ 的任意子空间. 则任意$\alpha_1 \in W_1$, $\alpha_2 \in W_2$, $\alpha_1,\alpha_2 \in W$, 从而$\alpha_1+\alpha_2 \in W$. 这说明$W_1+W_2 \subset W$, 即W_1+W_2是包含$W_1\cup W_2$的最小子空间. □

下面, 我们研究W_1+W_2的基与维数.

定理5.7.3. 设$\alpha_1,\ldots,\alpha_r$与$\beta_1,\ldots,\beta_s$是线性空间V的两个向量组. 则有

$$\langle\alpha_1,\ldots,\alpha_r\rangle+\langle\beta_1,\ldots,\beta_s\rangle = \langle\alpha_1,\ldots,\alpha_r,\beta_1,\ldots,\beta_s\rangle.$$

证明. 由生成子空间及子空间和的定义, 结论显然(请读者完成). □

例5.7.2. 设V, W_1与W_2如例 5.7.1所示. 求W_1+W_2的一组基与维数.

解. 由定理 5.7.3, $W_1+W_2 = \langle\alpha_1,\alpha_2,\alpha_3,\beta_1,\beta_2\rangle$. 易算得$\alpha_1,\alpha_2,\alpha_3,\beta_1,\beta_2$的一个极大无关组$\alpha_1,\alpha_2,\beta_1$, 因而$W_1+W_2 = \langle\alpha_1,\alpha_2,\beta_1\rangle$, 且$\alpha_1,\alpha_2,\beta_1$为$W_1+W_2$的一组基. 因此, $\dim(W_1+W_2) = 3$. □

关于子空间的和的维数有下面的结论.

定理5.7.4. *(维数公式)* 设W_1,W_2是线性空间V的子空间. 则

$$\dim(W_1+W_2) = \dim W_1+\dim W_2-\dim(W_1\cap W_2). \tag{5.24}$$

证明. 设$\dim W_1 = r$, $\dim W_2 = s$, $\dim(W_1\cap W_2) = t$. 再设$\alpha_1,\alpha_2,\ldots,\alpha_t$是$W_1\cap W_2$的一组基, 则它可以分别扩充为$W_1$与$W_2$ 的一组基

$$\alpha_1,\ldots,\alpha_t,\beta_1,\ldots,\beta_{r-t},$$

$$\alpha_1,\ldots,\alpha_t,\gamma_1,\ldots,\gamma_{s-t}$$

于是

$$W_1+W_2 = \langle\alpha_1,\ldots,\alpha_t,\beta_1,\ldots,\beta_{r-t},\gamma_1,\ldots,\gamma_{s-t}\rangle.$$

如果$\alpha_1,\ldots,\alpha_t,\beta_1,\ldots,\beta_{r-t},\gamma_1,\ldots,\gamma_{s-t}$线性无关, 则它构成$W_1+W_2$的一组基, 则有$\dim(W_1+W_2) = t+(r-t)+(s-t) = r+s-t$, 从而定理得证.

下证$\alpha_1,\ldots,\alpha_t,\beta_1,\ldots,\beta_{r-t},\gamma_1,\ldots,\gamma_{s-t}$线性无关.

设$\lambda_1\alpha_1+\ldots+\lambda_t\alpha_t+\mu_1\beta_1+\ldots+\mu_{r-t}\beta_{r-t}+\nu_1\gamma_1+\ldots+\nu_{s-t}\gamma_{s-t}=0$. 则

$$\lambda_1\alpha_1+\ldots+\lambda_t\alpha_t+\mu_1\beta_1+\ldots+\mu_{r-t}\beta_{r-t}=-\nu_1\gamma_1-\ldots-\nu_{s-t}\gamma_{s-t}\in W_1\cap W_2.$$

故存在常数$\delta_1,\ldots,\delta_t$使得

$$-\nu_1\gamma_1-\ldots-\nu_{s-t}\gamma_{s-t}=\delta_1\alpha_1+\ldots+\delta_t\alpha_t.$$

移项得

$$\nu_1\gamma_1+\ldots+\nu_{s-t}\gamma_{s-t}+\delta_1\alpha_1+\ldots+\delta_t\alpha_t=0.$$

由于$\alpha_1,\ldots,\alpha_t,\gamma_1,\ldots,\gamma_{s-t}$线性无关, 必有$\nu_1=\ldots=\nu_{s-t}=0$, 进而有

$$\lambda_1\alpha_1+\ldots+\lambda_t\alpha_t+\mu_1\beta_1+\ldots+\mu_{r-t}\beta_{r-t}=0.$$

再由$\alpha_1,\ldots,\alpha_t,\beta_1,\ldots,\beta_{r-t}$的线性无关性得$\lambda_1=\ldots=\lambda_t=\mu_1=\ldots=\mu_{r-t}=0$. 因此, $\alpha_1,\ldots,\alpha_t,\beta_1,\ldots,\beta_{r-t},\gamma_1,\ldots,\gamma_{s-t}$线性无关. 定理证毕. □

推论5.7.1. 设W_1,W_2是线性空间V的子空间, 则

$$\dim(W_1+W_2)=\dim W_1+\dim W_2\Leftrightarrow W_1\cap W_2=0.$$

我们回头看例 5.7.2, 由于$\dim W_1=2$, $\dim W_2=2$, $\dim(W_1\cap W_2)=1$, $\dim(W_1+W_2)=3$, 因此符合上述定理的结论.

§5.7.3 子空间的直和*

设V是三维几何空间, W_1与W_2分别是过原点的平面与直线, 且直线不位于平面之内. 显然, $V=W_1+W_2$, 并且任何向量$\alpha\in V$, α可以唯一地分解为$\alpha=\alpha_1+\alpha_2$, 这里$\alpha_1\in W_1,\alpha_2\in W_2$. 此时, 称和式$W_1+W_2$为**直和**. 由此引出

定义5.7.1. 设W_1,W_2是数域F上的线性空间V的子空间. 如果任意$\alpha\in W_1+W_2$可以唯一地写成

$$\alpha=\alpha_1+\alpha_2,\quad \alpha_1\in W_1,\alpha_2\in W_2.$$

则称和式W_1+W_2为**直和**, 记为$W_1\oplus W_2$. 如果$V=W_1\oplus W_2$, 则称W_1是W_2的**补空间**. 此时, W_2也是W_1的补空间.

关于子空间的直和有以下等价表示

定理5.7.5. 设W_1,W_2是线性空间V的子空间. 则下列命题等价

1. W_1+W_2为直和;

2. $W_1 \cap W_2 = 0$;

3. $\dim(W_1+W_2) = \dim W_1 + \dim W_2$;

4. 设$\alpha_1,\dots,\alpha_r$是W_1的一组基，$\beta_1,\dots,\beta_s$是W_2的一组基，则$\alpha_1,\dots,\alpha_r,\beta_1,\dots,\beta_s$构成$W_1+W_2$的一组基.

证明.$\Rightarrow$ 3. 若$W_1\cap W_2\neq 0$, 任取非零向量$\beta\in W_1\cap W_2$, 则向量$\alpha\in W_1+W_2$有两种不同的表示$\alpha=\alpha_1+\alpha_2=(\alpha_1+\beta)+(\alpha_2-\beta)$, 矛盾. 因此$W_1\cap W_2=0$.

$3\Rightarrow 2$. 设$\alpha\in W_1+W_2$有两种不同的表示$\alpha=\alpha_1+\alpha_2=\beta_1+\beta_2$, 这里$\alpha_1,\beta_1\in W_1$, $\alpha_2,\beta_2\in W_2$, $\alpha_1\neq\beta_1$, $\alpha_2\neq\beta_2$. 于是$0\neq\alpha_1-\beta_1=\beta_2-\alpha_2\in W_1\cap W_2$. 故$W_1\cap W_2\neq 0$, 矛盾.

$3\Leftrightarrow 4$. 由推论 5.7.1立即得到.

$4\Rightarrow 5$. 设$\alpha_1,\dots,\alpha_r$是W_1的基, $\beta_1,\dots,\beta_s$是W_2的基. 则有

$$W_1+W_2=\langle\alpha_1,\dots,\alpha_r,\beta_1,\dots,\beta_s\rangle.$$

由$\dim(W_1+W_2)=\dim W_1+\dim W_2$有, $\alpha_1,\dots,\alpha_r,\beta_1,\dots,\beta_s$构成$W_1+W_2$的一组基.

$5\Rightarrow 4$. 结论显然.

定理证毕. □

例5.7.3. 设$V=F^{n\times n}$是数域F上所有n阶方阵按矩阵的加法与数乘构成的线性空间，W_1是所有n 阶对称矩阵的全体，W_2是所有n阶反对称方阵全体. 证明: W_1,W_2是V的子空间, 且$V=W_1\oplus W_2$.

证明. 由于对称方阵在矩阵加法与数乘运算下封闭, 因此W_1是V的子空间. 类似地, W_2也是V的子空间.

对任意$A\in V$, $A=\frac{1}{2}(A+A^T)+\frac{1}{2}(A-A^T)\in W_1+W_2$. 因此$V=W_1+W_2$. 另一方面, 对任意$A\in W_1\cap W_2$有$A=A^T=-A^T$, 从而$A=0$. 因此$W_1\cap W_2=0$.

综上所述, $V=W_1\oplus W_2$. □

习题四

1. 设$\mathbf{a}_1=(1,2,-1)$, $\mathbf{a}_2=(2,0,3)$, $\mathbf{a}_3=(2,1,0)$是三维几何空间中的三个向量. 能否将其中某一个向量写成其它两个向量的线性组合? 这三个向量是否共面?

2. 设$\mathbf{a}_1=(0,1,-2)$, $\mathbf{a}_2=(2,1,3)$, $\mathbf{a}_3=(4,5,0)$是三维几何空间中的三个向量. 能否将其中某一个向量写成其它两个向量的线性组合? 这三个向量是否共面?

3. 在F^4中, 判断向量$\mathbf{b}$能否写成$\mathbf{a}_1,\mathbf{a}_2,\mathbf{a}_3$的线性组合. 若能, 请写出一种表示方式.

$$
\begin{aligned}
(1)\quad & \mathbf{a}_1=(-1,3,0,-5), \mathbf{a}_2=(2,0,7,-3),\\
& \mathbf{a}_3=(-4,1,-2,6), \mathbf{b}=(8,3,-1,-25).
\end{aligned}
$$

$$
\begin{aligned}
(2)\quad & \mathbf{a}_1=(3,-5,2,-4)^T, \mathbf{a}_2=(-1,7,-3,6)^T,\\
& \mathbf{a}_3=(3,11,-5,10)^T, \mathbf{b}=(2,-30,13,-26)^T.
\end{aligned}
$$

4. 设$\mathbf{a}_1=(1,0,0,0)$, $\mathbf{a}_2=(1,1,0,0)$, $\mathbf{a}_3=(1,1,1,0)$, $\mathbf{a}_4=(1,1,1,1)$. 证明: F^4中任何向量可以写成$\mathbf{a}_1,\mathbf{a}_2,\mathbf{a}_3,\mathbf{a}_4$的线性组合, 且表示唯一.

5. 设$\mathbf{P}_i=(x_i,y_i,z_i)$, $i=1,2,3,4$是三维几何空间中的点. 证明: $\mathbf{P}_i$, $i=1,2,3,4$共面的条件是

$$
\begin{vmatrix}
x_1 & y_1 & z_1 & 1\\
x_2 & y_2 & z_2 & 1\\
x_3 & y_3 & z_3 & 1\\
x_4 & y_4 & z_4 & 1
\end{vmatrix}=0.
$$

6. 设$\mathbf{a}_1,\mathbf{a}_2,\mathbf{a}_3,\mathbf{a}_4$是三维几何空间中的四个向量. 证明它们必线性相关.

7. 设$\mathbf{b}_1,\mathbf{b}_2,\ldots,\mathbf{b}_s$中每一个向量是$n$维数组向量$\mathbf{a}_1,\mathbf{a}_2,\ldots,\mathbf{a}_r$的线性组合. 证明:$\mathbf{b}_1,\mathbf{b}_2,\ldots,\mathbf{b}_s$ 的任何线性组合都是$\mathbf{a}_1,\mathbf{a}_2,\ldots,\mathbf{a}_r$的线性组合.

8. 证明: 对线性方程组做初等变换后得到的线性方程组中的每一个方程都是原方程组的线性组合.

9. 判别下列线性方程组是否线性相关

$$
(1)\begin{cases}
-x_1+2x_2+3x_3=4\\
5x_1-x_3=-1\\
8x_1-6x_2-10x_3=-13
\end{cases}
\qquad
(2)\begin{cases}
2x_1+3x_2+x_3+5x_4=2\\
3x_1+2x_2+4x_3+2x_4=3\\
x_1+x_2+2x_3+4x_4=1
\end{cases}
$$

10. 判断下列向量组是否线性相关

(1) $\mathbf{a}_1=(1,1,1)$, $\mathbf{a}_2=(1,-2,3)$, $\mathbf{a}_3=(1,4,9)$;

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

(3) $\mathbf{a}_1=(-2,1,0,3)$, $\mathbf{a}_2=(1,-3,2,4)$, $\mathbf{a}_3=(3,0,2,-1)$, $\mathbf{a}_4=(2,-2,4,6)$;

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

11. 证明: 任何一个经过以下两个平面

$$A_1x+B_1y+C_1z+D_1=0, \qquad A_2x+B_2y+C_2z+D_2=0$$

交线的平面的方程能写成:

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

其中λ,μ为不全为零的常数.

12. 下列说法是否正确? 为什么?

(1) 若$\mathbf{a}_1,\mathbf{a}_2,\ldots,\mathbf{a}_s(s\geq 2)$线性相关, 则其中每一个向量都可以表示成其它向量的线性组合.

(2) 如果向量组的任何不是它本身的子向量组都线性无关, 则该向量组也线性无关.

(3) 若向量组线性无关, 则它的任何子向量组都线性无关.

(4) F^n的$n+1$个向量组成的向量组必线性相关.

(5) 设$\mathbf{a}_1,\mathbf{a}_2,\ldots,\mathbf{a}_s$线性无关, 则$\mathbf{a}_1+\mathbf{a}_2,\mathbf{a}_2+\mathbf{a}_3,\ldots,\mathbf{a}_s+\mathbf{a}_1$必线性无关.

(6) 设$\mathbf{a}_1,\mathbf{a}_2,\ldots,\mathbf{a}_s$线性相关, 则$\mathbf{a}_1+\mathbf{a}_2,\mathbf{a}_2+\mathbf{a}_3,\ldots,\mathbf{a}_s+\mathbf{a}_1$必线性相关.

(7) 设$\mathbf{a}_1,\mathbf{a}_2,\ldots,\mathbf{a}_s\in F^n$线性无关, 则它们的加长向量组也必线性无关.

(8) 设$\mathbf{a}_1,\mathbf{a}_2,\ldots,\mathbf{a}_s\in F^n$线性相关, 则它们的加长向量组也必线性相关.

13. 若向量组$\mathbf{a}_1,\mathbf{a}_2,\ldots,\mathbf{a}_n\in F^n$线性无关, 而$\mathbf{a}_1,\mathbf{a}_2,\ldots,\mathbf{a}_n,\mathbf{b}$线性相关, 则$\mathbf{b}$可以表示成$\mathbf{a}_1,\mathbf{a}_2,\ldots,\mathbf{a}_n$的线性组合, 且表示唯一.

14. 证明**向量表示基本定理**: 设$\mathbf{a}_1,\mathbf{a}_2,\ldots,\mathbf{a}_n\in F^n$线性无关, 则任意向量$\mathbf{b}\in F^n$可以表示为$\mathbf{a}_1,\mathbf{a}_2,\ldots,\mathbf{a}_n$的线性组合, 且表示唯一.

15. 证明: 非零向量组$\mathbf{a}_1,\mathbf{a}_2,\ldots,\mathbf{a}_s$线性无关的充要条件是, 每个$\mathbf{a}_i(1<i\leq s)$都不能用它前面的向量线性表示.

16. 设向量组$\mathbf{a}_1,\mathbf{a}_2,\ldots,\mathbf{a}_s$线性无关, $\mathbf{b}=\lambda_1\mathbf{a}_1+\ldots+\lambda_s\mathbf{a}_s$. 如$\lambda_i\neq 0$, 则用$\mathbf{b}$代替$\mathbf{a}_i$后, 向量组$\mathbf{a}_1,\ldots,\mathbf{a}_{i-1},\mathbf{b},\mathbf{a}_{i+1},\ldots,\mathbf{a}_s$线性无关.

17. 设向量组$\mathbf{a}_1,\mathbf{a}_2,\ldots,\mathbf{a}_r$线性无关, 且$\mathbf{a}_1,\mathbf{a}_2,\ldots,\mathbf{a}_r$可以由向量组$\mathbf{b}_1,\mathbf{b}_2,\ldots,\mathbf{b}_r$线性表示, 则$\mathbf{b}_1,\mathbf{b}_2,\ldots,\mathbf{b}_r$也线性无关.

18. 证明:向量组等价具有自身性、对称性与传递性.

19. 求下列向量组的极大无关组与秩

(1) $\mathbf{a}_1 = (3,-2,0), \mathbf{a}_2 = (27,-18,0), \mathbf{a}_3 = (-1,5,8)$.

(2) $\mathbf{a}_1 = (1,-1,2,4), \mathbf{a}_2 = (0,3,1,2), \mathbf{a}_3 = (3,0,7,14), \mathbf{a}_4 = (1,-1,2,0)$,
$\mathbf{a}_5 = (2,1,5,6)$.

(3) $\mathbf{a}_1 = (0,1,2,3), \mathbf{a}_2 = (1,2,3,4), \mathbf{a}_3 = (3,4,5,6), \mathbf{a}_4 = (4,3,2,1)$,
$\mathbf{a}_5 = (6,5,4,3)$.

20. 证明: $\mathbf{a}_{i_1},\mathbf{a}_{i_2},\cdots,\mathbf{a}_{i_r}$是$\mathbf{a}_1,\mathbf{a}_2,\cdots,\mathbf{a}_m$的极大无关组, 当且仅当$\langle\mathbf{a}_1,\mathbf{a}_2,\cdots,\mathbf{a}_m\rangle = \langle\mathbf{a}_{i_1},\mathbf{a}_{i_2},\cdots,\mathbf{a}_{i_r}\rangle$ 且$\mathbf{a}_{i_1},\mathbf{a}_{i_2},\cdots,\mathbf{a}_{i_r}$ 线性无关.

21. 证明: 若向量$\mathbf{b}$可以由向量组$\mathbf{a}_1,\mathbf{a}_2,\ldots,\mathbf{a}_m$线性表示, 则$\mathbf{b}$可以由$\mathbf{a}_1,\mathbf{a}_2,\ldots,\mathbf{a}_m$的极大无关组表示.

22. 设向量组$\mathbf{a}_1,\mathbf{a}_2,\ldots,\mathbf{a}_m$的秩为$r$, 则其中任何$r$个线性无关的向量构成$\mathbf{a}_1$, $\mathbf{a}_2,\ldots,\mathbf{a}_m$的极大无关组.

23. 设向量组$\mathbf{a}_1,\mathbf{a}_2,\ldots,\mathbf{a}_m$的秩为$r$, 如$\mathbf{a}_1,\mathbf{a}_2,\ldots,\mathbf{a}_m$可以由它的$r$个向量线性表示, 则这$r$个向量构成$\mathbf{a}_1,\mathbf{a}_2,\ldots,\mathbf{a}_m$的极大无关组.

24. 证明定理 5.3.5的1, 2, 4.

25. 求下列矩阵的秩, 并求出它的行空间、列空间及零空间的一组基.

$$(1)\quad \begin{pmatrix} 3 & -2 & 0 & 1 \\ -1 & -3 & 2 & 0 \\ 2 & 0 & -4 & 5 \\ 4 & 1 & -2 & 1 \end{pmatrix} \qquad (2)\quad \begin{pmatrix} 3 & 6 & 1 & 5 \\ 1 & 4 & -1 & 3 \\ -1 & -10 & 5 & -7 \\ 4 & -2 & 8 & 0 \end{pmatrix}$$

26. 证明: 线性方程组(5.4)有解, 当且仅当$\mathbf{b} \in \langle\mathbf{a}_1,\mathbf{a}_2,\cdots,\mathbf{a}_m\rangle$, 当且仅当$\langle\mathbf{a}_1,\mathbf{a}_2,\cdots,\mathbf{a}_m\rangle = \langle\mathbf{a}_1,\mathbf{a}_2,\cdots,\mathbf{a}_m,\mathbf{b}\rangle$.

27. 证明: $\mathrm{rank}(\mathbf{a}_1,\mathbf{a}_2,\ldots,\mathbf{a}_r,\mathbf{b}_1,\mathbf{b}_2,\ldots,\mathbf{b}_s) \le \mathrm{rank}(\mathbf{a}_1,\mathbf{a}_2,\ldots,\mathbf{a}_r)+\mathrm{rank}(\mathbf{b}_1,\mathbf{b}_2,\ldots,\mathbf{b}_s)$.

28. 证明: n阶方阵A可逆$\Leftrightarrow$ $\mathrm{rank}(A) = n$ $\Leftrightarrow$ A的行向量线性无关$\Leftrightarrow$ A的列向量线性无关.

29. 设矩阵$A \in F^{m\times n}$的秩为r, 则A的不等于零的r阶子式所在行(列)构成A的行(列)向量的极大无关组.

30. 设A,B是同阶矩阵. 证明: $\operatorname{rank}(A+B) \le \operatorname{rank}(A)+\operatorname{rank}(B)$.

31. 设$\mathbf{a}_1,\mathbf{a}_2,\ldots,\mathbf{a}_n$是$F^n$的基, 向量组$\mathbf{b}_1,\mathbf{b}_2,\ldots,\mathbf{b}_n$与$\mathbf{a}_1,\mathbf{a}_2,\ldots,\mathbf{a}_n$有关系式

$$(\mathbf{b}_1,\mathbf{b}_2,\ldots,\mathbf{b}_n)=(\mathbf{a}_1,\mathbf{a}_2,\ldots,\mathbf{a}_n)T.$$

证明: $\mathbf{b}_1,\mathbf{b}_2,\ldots,\mathbf{b}_n$为$F^n$的基当且仅当$T$为可逆方阵.

32. 证明定理 5.4.2.

33. 证明定理 5.4.3.

34. 以向量组$\mathbf{a}_1=(3,1,0)$, $\mathbf{a}_2=(6,3,2)$, $\mathbf{a}_3=(1,3,5)$为基, 求$\beta=(2,-1,2)$的坐标.

35. 设$\mathbf{a}_1=(3,2,-1,4)$, $\mathbf{a}_2=(2,3,0,-1)$.

(1) 将$\mathbf{a}_1,\mathbf{a}_2$扩充为$\mathbf{R}^4$的一组基;

(2) 给出标准基在该组基下的表示;

(3) 求$\beta=(1,3,4,-2)$在该组基下的坐标.

36. 将三维几何空间中的直角坐标系$[O;\mathbf{e}_1,\mathbf{e}_2,\mathbf{e}_3]$绕单位向量$\mathbf{e}=\dfrac{1}{\sqrt{3}}(1,1,1)$逆时针旋转$\theta$角, 求新坐标与原坐标之间的关系.

37. 设$\mathbf{a}_1,\mathbf{a}_2,\ldots,\mathbf{a}_s$为非齐次线性方程组$A\mathbf{x}=\mathbf{b}$的一组线性无关解, $\lambda_1,\lambda_2,\ldots,\lambda_s$为常数. 给出$\lambda_1\mathbf{a}_1+\lambda_2\mathbf{a}_2+\ldots+\lambda_s\mathbf{a}_s$为该线性方程组的解的充要条件.

38. 给出三个平面$a_ix+b_iy+c_iz=d_i$, $i=1,2,3$相交于一条直线的充要条件.

39. 求下列齐次线性方程组的基础解系与通解

$$(1)\quad \begin{cases} x_1-3x_2+x_3-2x_4=0\\ -5x_1+x_2-2x_3+3x_4=0\\ -x_1-11x_2+2x_3-5x_4=0\\ 3x_1+5x_2+x_4=0 \end{cases}$$

$$(2)\quad \begin{cases} x_1+x_2+x_3+x_4-4x_5=0\\ x_1-2x_2+3x_3-4x_4+2x_5=0\\ -x_1+3x_2-5x_3+7x_4-4x_5=0\\ x_1+2x_2-x_3+4x_4-6x_5=0 \end{cases}$$

40. 已知F^5中向量$\mathbf{a}_1=(1,2,3,2,1)^T$及$\mathbf{a}_2=(1,3,2,1,2)^T$, 求找一个齐次线性方程组使得$\mathbf{a}_1$与$\mathbf{a}_2$为该方程组的基础解系.

41. 判断下列集合关于规定的运算是否构成线性空间

(1) V是所有实数对(x,y)的集合, 数域$F=\mathbf{R}$. 定义

$$(x_1,y_1)+(x_2,y_2)=(x_1+x_2,y_1+y_2),\quad \lambda(x,y)=(x,y).$$

(2) V是所有满足$f(-1)=0$的实函数的集合, 数域$F=\mathbf{R}$. 定义加法为函数的加法, 数乘为数与函数的乘法.

(3) V是所有满足$f(0)\neq 0$的实函数的集合, 数域$F=\mathbf{R}$. 定义加法为函数的加法, 数乘为数与函数的乘法.

(4) V是数域F上所有n阶可逆方阵的全体, 加法为矩阵的加法, 数乘为矩阵的数乘.

42. 设V是所有实函数全体在实数域上构成的线性空间, 判断V中下列函数组是否线性相关.

(1) $1,x,\sin x$;

(2) $1,x,e^x$;

(3) $1,\cos 2x,\cos^2 x$;

(4) $1,x^2,(x-1)^3,(x+1)^3$;

(5) $\cos x,\cos 2x,\ldots,\cos(nx)$.

43. 证明定理 5.6.1.

44. 设$F_n[x]$是次数小于或等于n的多项式全体构成的线性空间.

(1) 证明: $S=\{1, x-1, (x-1)^2,\ldots,(x-1)^n\}$构成$F^n$的一组基;

(2) 求S到基$T=\{1,x,\ldots,x^n\}$的过渡矩阵;

(3) 求多项式$p(x)=a_0+a_1x+\ldots+a_nx^n\in F_n[x]$在基$S$下的坐标.

45. V是数域F上n阶对称方阵的全体, 定义加法为矩阵的加法, 数乘为矩阵的数乘. 证明: V是线性空间, 并求V的一组基及维数.

46. 给定矩阵

$$A = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 4 & -2 & 1 \end{pmatrix}$$

令V是与A乘法可交换的三阶实方阵的全体. 证明: V在矩阵加法与数乘下构成实数域上的线性空间, 并求V的一组基与维数.

47. $V = F^{n\times n}$是数域F上所有n阶矩阵构成的线性空间, 令W是数域F上所有满足$Tr(A) = 0$的n阶矩阵的全体. 证明: W是V的线性子空间, 并求W的一组基与维数.

48* 证明定理 5.7.3.

49* 证明:有限维线性空间的任何子空间都有补空间.

50* 令$F_n[x]$是数域F上次数不超过n的多项式全体按多项式的加法与数乘构成的线性空间, W_1是$F_n[x]$的偶多项式(满足$f(-x) = f(x)$)全体, W_2是$F_n[x]$奇多项式(满足$f(-x) = -f(x)$)全体. 证明: W_1, W_2是$F_n[x]$的子空间, 且$F_n[x] = W_1 \oplus W_2$.

第六章 线性变换

上一章介绍了线性空间的基本理论. 本章将研究线性空间上的线性变换. 我们首先从数组空间上的线性映射开始研究线性映射的性质, 接着重点介绍矩阵和线性变换的特征值和特征向量, 它们是矩阵与线性变换最重要的几何量, 在科学与工程计算中有广泛的应用. 随后, 我们将研究矩阵在相似关系下的标准形等问题. 特别地, 我们将给出矩阵相似于对角阵的充分必要条件. 最后, 我们将数组空间上的线性变换理论推广到一般线性空间. 作为一个选学内容, 我们还将对矩阵的若当(Jordan标准形理论与计算做一个简单的介绍. 本章将重点关注线性代数的几何理论(变换)与代数理论(矩阵)间的有机结合, 而用代数的方法研究几何问题是线性代数的一个基本思想.

§6.1 数组空间上的线性映射

在第二章, 我们介绍矩阵乘法时引进了线性映射的概念. 这里, 我们给出正式的定义.

定义6.1.1. 设$\mathscr{A}:\mathbf{x}\in F^n\to\mathbf{y}\in F^m$是数组空间$F^n$到$F^m$的映射. 如果

$$\mathbf{y}=A\mathbf{x} \tag{6.1}$$

其中$A\in F^{m\times n}$, 则称$\mathscr{A}$为**线性映射**. 特别地, F^n到F^n上的线性映射称为F^n上的**线性变换**.

实际上, 线性映射与线性变换的概念在解析几何中早已有之. 例如, 对坐标平面的单位圆$C:=x^2+y^2-1=0$做如下的伸缩变换

$$x'=3x,\qquad y'=2y$$

得到一个椭圆$C':=(x')^2/9+(y')^2/4-1=0$. 上述变换将单位圆沿$x$轴方向放大3倍, 沿$y$轴方向放大2倍, 从而得到一个椭圆. 这里, 伸缩变换就是平面上的线性变换.

再如, 将图形绕坐标原点逆时针旋转θ角的图形变换为

$$x'=x\cos\theta-y\sin\theta,\qquad y'=x\sin\theta+y\cos\theta.$$

旋转变换不改变图形的形状, 只改变它的位置, 它也是一种线性变换.

图形的线性变换不仅可沿坐标轴方向伸缩, 而且可以产生错切现象.

例6.1.1. 考虑坐标平面上的如下变换

$$x' = x + 0.1y, \qquad y' = 0.2x + y.$$

设C_1是由边平行于坐标轴的矩形网格, C_2是单位圆$x^2+y^2-1=0$, C_3是正弦曲线$y=\sin(x)$. 绘制变换前后的图形, 观察图形的变化.

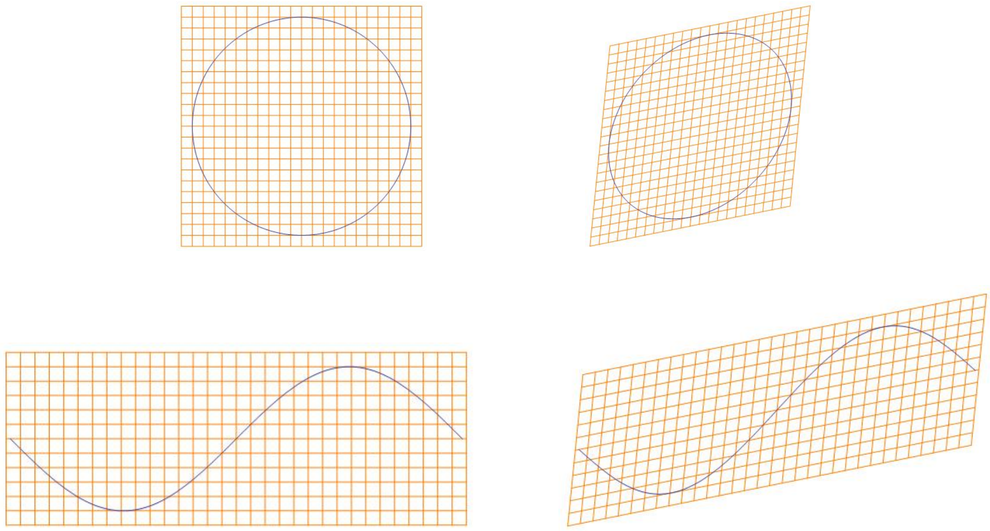


图 6.1 图形的线性变换

我们看到, 图形不仅沿斜线方向发生伸缩变化, 并且产生错切现象. 但上述变换仍保持图形的基本性状不变, 例如, 直线仍变为直线, 平行直线变为平行直线, 圆变为椭圆.

例6.1.2. 设A为2阶方阵, 则当A分别是

$$A = \begin{pmatrix} \lambda & \\ & \mu \end{pmatrix}, \quad \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}, \quad \begin{pmatrix} 1 & \\ & -1 \end{pmatrix}$$

时, 线性变换 (6.1)分别是平面上的**伸缩变换**, **旋转变换**, **反射变换**.

另一个基本的线性映射是投影变换. 例如, 三维空间中的点(x,y,z)到xoy平面的投影$(x,y,0)$就是投影变换.

例6.1.3. 在三维直角坐标系中, 取定一个向量$\mathbf{v}=(a,b,c)$, 其中$c\neq 0$. 则任意点$\mathbf{P}(x,y,z)$沿方向$\mathbf{v}$在oxy平面上的平行投影为$\mathbf{P}'$:

$$(x',y',z') = (x-az/c, y-bz/c, 0).$$

投影变换显然也是一个线性映射.

将直线变为直线,且将平行直线变为平行直线,是线性映射的基本特性. 我们以平面上的线性变换来说明.

命题6.1.1. *设$A \in F^{2\times 2}$为可逆方阵. 则线性变换(6.1)将平面上的直线变为直线, 平行直线变为平行直线.*

证明. 设l是由$\mathbf{P}_1,\mathbf{P}_2(\mathbf{P}_1 \neq \mathbf{P}_2)$决定的直线, $\mathbf{P}_1,\mathbf{P}_2$的坐标分别为$\mathbf{X}_1,\mathbf{X}_2$. 经过线性变换 (6.1), l变为l', $\mathbf{P}_1,\mathbf{P}_2$分别变为$\mathbf{P}'_1,\mathbf{P}'_2$. 则$\mathbf{P}'_1,\mathbf{P}'_2$的坐标分别为$\mathbf{X}'_1 = A\mathbf{X}_1$, $\mathbf{X}'_2 = A\mathbf{X}_2$. 由于$A$可逆, 显然$\mathbf{X}'_1 \neq \mathbf{X}'_2$, 即$\mathbf{P}'_1 \neq \mathbf{P}'_2$.

l上任一点$\mathbf{P}$可以表示为$\mathbf{P} = (1-t)\mathbf{P}_1 + t\mathbf{P}_2, t \in F$, 其坐标为$\mathbf{X} = (1-t)\mathbf{X}_1 + t\mathbf{X}_2$. 则$\mathbf{P}$在线性变换 (6.1) 下的像$\mathbf{P}'$的坐标为

$$\mathbf{X}' = A\mathbf{X} = A((1-t)\mathbf{X}_1 + t\mathbf{X}_2) = (1-t)A\mathbf{X}_1 + tA\mathbf{X}_2 = (1-t)\mathbf{X}'_1 + t\mathbf{X}'_2.$$

这表明, $\mathbf{P}' = (1-t)\mathbf{P}'_1 + t\mathbf{P}'_2$, 即$\mathbf{P}'$在直线$l'$上. 因此, 线性变换 (6.1)将直线变为直线.

设直线$\mathbf{P}_1\mathbf{P}_2 \parallel \mathbf{P}_3\mathbf{P}_4$, 其中$\mathbf{P}_i$的坐标为$\mathbf{X}_i$, $i = 1,2,3,4$. $\mathbf{P}_i$在线性变换 (6.1) 下的像为$\mathbf{P}'_i$, 其坐标$\mathbf{X}'_i = A\mathbf{X}_i$, $i = 1,2,3,4$. 直线$\mathbf{P_1P_2}$的像为直线$\mathbf{P}'_1\mathbf{P}'_2$, 直线$\mathbf{P_3P_4}$的像为直线$\mathbf{P}'_3\mathbf{P}'_4$. 由$\mathbf{P}_1\mathbf{P}_2 \parallel \mathbf{P}_3\mathbf{P}_4$, 可设$\overrightarrow{\mathbf{P_3P_4}} = \lambda\overrightarrow{\mathbf{P_1P_2}}$, $\lambda \neq 0$. 则$\mathbf{X}_4 - \mathbf{X}_3 = \lambda(\mathbf{X}_2 - \mathbf{X}_1)$. 于是$\overrightarrow{\mathbf{P}'_1\mathbf{P}'_2}$的坐标为$A\mathbf{X}_2 - A\mathbf{X}_1 = A(\mathbf{X}_2 - \mathbf{X}_1)$, $\overrightarrow{\mathbf{P}'_3\mathbf{P}'_4}$的坐标为$A\mathbf{X}_4 - A\mathbf{X}_3 = A(\mathbf{X}_4 - \mathbf{X}_3) = \lambda A(\mathbf{X}_2 - \mathbf{X}_1)$. 从而$\overrightarrow{\mathbf{P}'_1\mathbf{P}'_2} \parallel \overrightarrow{\mathbf{P}'_3\mathbf{P}'_4}$.

□

将直线变为直线, 平行直线变为平行直线这一特性可以由以下关键条件刻画: 设$\mathbf{x},\mathbf{y} \in F^n$, 则

$$\mathscr{A}(\lambda\mathbf{x} + \mu\mathbf{y}) = \lambda\mathscr{A}\mathbf{x} + \mu\mathscr{A}\mathbf{y}. \tag{6.2}$$

这里λ,μ是任意数. 这条性质为今后将线性映射推广到一般线性空间奠定了基础.

由性质 (6.2)可知, 线性映射保持线性组合运算不变, 即

$$\mathscr{A}(\lambda_1\mathbf{x}_1 + \lambda_2\mathbf{x}_2 + \ldots + \lambda_k\mathbf{x}_k) = \lambda_1\mathscr{A}\mathbf{x}_1 + \lambda_2\mathscr{A}\mathbf{x}_2 + \ldots + \lambda_k\mathscr{A}\mathbf{x}_k.$$

由此我们可以得到, 若$\mathbf{x}_1,\mathbf{x}_2,\ldots,\mathbf{x}_k$线性相关, 则$\mathscr{A}\mathbf{x}_1,\mathscr{A}\mathbf{x}_2,\ldots,\mathscr{A}\mathbf{x}_k$ 也线性相关. 反之, 若$\mathscr{A}\mathbf{x}_1,\mathscr{A}\mathbf{x}_2,\ldots,\mathscr{A}\mathbf{x}_k$线性无关, 则$\mathbf{x}_1,\mathbf{x}_2,\ldots,\mathbf{x}_k$ 也线性无关.

.

线性映射 (6.1)将空间F^n映射后的像的集合是一个子空间. 实际上, 我们有

定理6.1.2. 线性映射 (6.1)的像的集合$Im(\mathscr{A}) := \{A\mathbf{x} | ,\mathbf{x} \in F^n\}$是$A$的列空间, 称为$\mathscr{A}$的**像空间**. 线性映射 (6.1) 的核$Ker(\mathscr{A}) := \{\mathbf{x} \in F^n | A\mathbf{x} = \mathbf{0}\}$为$F^n$的零空间, 也称为$\mathscr{A}$的**核空间**. 且$\dim(Im(\mathscr{A})) = \operatorname{rank}(A)$, $\dim(Ker(\mathscr{A})) = n - \operatorname{rank}(A)$, $\dim(Im(\mathscr{A})) + \dim(Ker(\mathscr{A})) = n$.

证明. 设$\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n \in F^n$为单位坐标列向量, $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n$为$A$的$n$个列向量. 则$\mathscr{A}\mathbf{e}_i = \mathbf{a}_i$, $i = 1, 2, \ldots, n$. 于是, 对任意$\mathbf{x} = \sum_{i=1}^n x_i \mathbf{e}_i \in F^n$,

$$\mathscr{A}\mathbf{x} = \sum_{i=1}^n x_i \mathbf{a}_i.$$

故$Im(\mathscr{A})$是由A的列向量$\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n$生成的子空间, 即$A$的列空间. 因此, $\dim(Im(\mathscr{A})) = \operatorname{rank}(A)$.

另一方面, $Ker(\mathscr{A})$是齐次线性方程组$A\mathbf{x} = \mathbf{0}$的解集, 即A的零空间, 其维数为$n - \operatorname{rank}(A)$. 从而 $\dim(Im(\mathscr{A})) + \dim(Ker(\mathscr{A})) = n$.

□

上述定理刻画了矩阵A的秩的几何意义. 简单来说, 矩阵A的秩就是A的列(行)空间的维数.

推论6.1.1. 用$\mathscr{A}$表示线性映射 (6.1). 则有

1. $\mathscr{A}$是单射$\Leftrightarrow \operatorname{rank}(A) = n$.
2. $\mathscr{A}$是满射$\Leftrightarrow \operatorname{rank}(A) = m$.
3. $\mathscr{A}$是双射$\Leftrightarrow \operatorname{rank}(A) = m = n$, 即$A$为可逆方阵.

证明. 由定理 6.1.2, $\mathscr{A}$是单射当且仅当$Ker(\mathscr{A}) = \{0\}$, 当且仅当$\operatorname{rank}(A) = n$. $\mathscr{A}$是满射当且仅当$Im(\mathscr{A}) = F^m$, 当且仅当$\operatorname{rank}(A) = m$. 因此, $\mathscr{A}$是双射当且仅当$\operatorname{rank}(A) = m = n$. □

§6.2 线性变换的特征值与特征向量

§6.2.1 特征值与特征向量的定义

上一节介绍的伸缩变换沿两个数轴方向将图形进行拉伸或压缩变形, 这时数轴方向是该变换的两个特殊方向, 即变换前的向量与变换后的向量重合, 这种方向称为线性变换的特征方向或特征向量. 而沿特征方向伸缩的比例就是变换的特征值. 基于上述观察, 我们引进线性变换的特征向量与特征值的概念.

定义6.2.1. 设$\mathscr{A}$是数组空间F^n上的线性变换, 其变换公式为$\mathbf{y}=A\mathbf{x}$, 这里$A\in F^{n\times n}$, $\mathbf{x},\mathbf{y}\in F^n$. 如果存在 $\lambda\in F$ 及非零列向量 $\mathbf{x}\in F^n$, 使得

$$\mathscr{A}\mathbf{x}=\lambda\mathbf{x}\qquad(\text{或者}A\mathbf{x}=\lambda\mathbf{x}),\tag{6.3}$$

则称 λ 为线性变换$\mathscr{A}$(或者方阵 A) 的一个**特征值**, 而称 $\mathbf{x}$ 为属于特征值 λ 的一个**特征向量**.

由上述定义看出, 特征向量$\mathbf{x}$在线性变换$\mathscr{A}$ 下保持方向不变(相同或相反), 即$\mathscr{A}\mathbf{x}\parallel\mathbf{x}$, 而向量的模长伸缩$|\lambda|$ 倍. 也就是说, $\mathbf{x}$是线性变换$\mathscr{A}$下保持方向不变的一个特别方向, 而λ 是变换$\mathscr{A}$在该方向的伸缩比. 下面我们看一个具体的例子.

例6.2.1. 给定矩阵$A=\begin{pmatrix}1&1\\1&3\end{pmatrix}$, 它确定了坐标平面上的线性变换

$$\mathscr{A}:F^2\to F^2,\quad \mathscr{A}\begin{pmatrix}x\\y\end{pmatrix}=\begin{pmatrix}1&1\\1&3\end{pmatrix}\begin{pmatrix}x\\y\end{pmatrix}.$$

求线性变换$\mathscr{A}$的特征值与特征向量. 该线性变换将单位圆变为什么图形?

解. 设λ是矩阵$\mathscr{A}$的特征值, $(x,y)^T$是对应的特征向量. 则由$A(x,y)^T=\lambda(x,y)^T$有

$$\begin{pmatrix}1-\lambda&1\\1&3-\lambda\end{pmatrix}\begin{pmatrix}x\\y\end{pmatrix}=\mathbf{0}.$$

上述线性方程组关于x,y有非零解, 故

$$\begin{vmatrix}1-\lambda&1\\1&3-\lambda\end{vmatrix}=0.$$

解得$\lambda_1=2+\sqrt{2},\lambda_2=2-\sqrt{2}$, 它们是线性变换 $\mathscr{A}$的两个特征值. 求得对应的特征向量分别为$T_1=(1,1+\sqrt{2})^T,T_2=(1,1-\sqrt{2})^T$. 故$\mathscr{A}T_i=\lambda_iT_i$, $i=1,2$. 因此, 线性变换$\mathscr{A}$沿T_1,T_2保持方向不变, 且在T_1,T_2方向的伸缩比分别为$2+\sqrt{2},2-\sqrt{2}$. 于是, $\mathscr{A}$将单位圆C变为椭圆C'. 椭圆C'长轴方向为T_1, 长半轴长度为$2+\sqrt{2}$; C'短轴方向为T_2, 短半轴长度为$2-\sqrt{2}$.

□

显然, 数组空间上的线性变换$\mathscr{A}$与方阵A的特征值与特征向量是一样的, 因此下面我们只讨论矩阵A的特征值与特征向量. 对应于给特征值, 特征向量不唯一. 实际上, 如果 $\mathbf{x}$ 为属于 λ 的特征向量, 则对任意非零的 $k\in F$, $k\mathbf{x}$ 也是属于 λ 的特征向量.

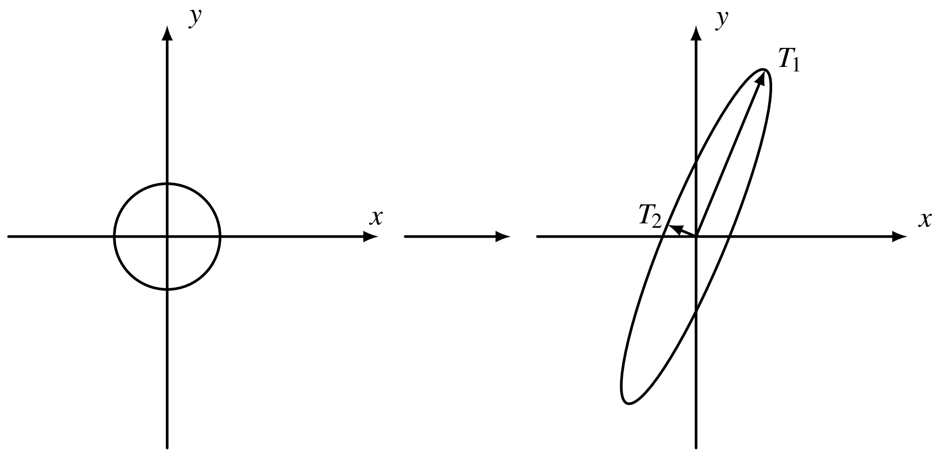


图 6.2 线性变换的特征值与特征向量的几何意义

定理6.2.1. 设$\lambda \in F$是矩阵$A \in F^{n\times n}$的特征值, 则

$$V_A(\lambda) = \{\mathbf{x} \in F^n \,|\, A\mathbf{x} = \lambda\mathbf{x}\}.$$

是F^n的子空间, 称为矩阵 A 的属于特征值 λ 的**特征子空间**, 即特征子空间$V_A(\lambda)$由λ的所有特征向量与零向量一起构成.

证明. 设$\mathbf{x}, \mathbf{y} \in V_A(\lambda)$, 即$A\mathbf{x} = \lambda\mathbf{x}$, $A\mathbf{y} = \lambda\mathbf{y}$. 则对任意$\mu_1, \mu_2$,

$$A(\mu_1\mathbf{x} + \mu_2\mathbf{y}) = \mu_1 A\mathbf{x} + \mu_2 A\mathbf{y} = \lambda(\mu_1\mathbf{x} + \mu_2\mathbf{y}).$$

因此, $\mu_1\mathbf{x} + \mu_2\mathbf{y} \in V_A(\lambda)$. 即$V_A(\lambda)$为$F^n$的子空间. □

§6.2.2 特征值与特征向量的计算

设 $A = (a_{ij})$ 为数域 F 上的一个 n 阶方阵, λ是A的特征值, $\mathbf{x}$ 是对应的特征向量. 则

$$(\lambda I - A)\mathbf{x} = \mathbf{0}. \tag{6.4}$$

上述关于$\mathbf{x}$的线性方程组有非零解, 因此

$$\det(\lambda I - A) = \begin{vmatrix} \lambda - a_{11} & -a_{12} & \cdots & -a_{1n} \\ -a_{21} & \lambda - a_{22} & \cdots & -a_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ -a_{n1} & -a_{n2} & \cdots & \lambda - a_{nn} \end{vmatrix} = 0. \tag{6.5}$$

对于给定的 n 阶方阵 A, 将行列式 $\det(\lambda I - A)$ 完全展开得到一个以 λ 为变量的 n 次首项系数为1的多项式.

定义6.2.2. 设A为n阶方阵, 称行列式 $\det(\lambda I-A)$为矩阵 A 的 **特征多项式**, 记为$p_A(\lambda)$.

由上面的分析, λ 为矩阵 A 的特征值当且仅当 λ 为A 的特征多项式的根. 但是, 数域 F 上的多项式在数域 F 中并不一定有根, 例如实系数多项式 x^2+1 在实数域中没有实根. 为了确保特征值的存在性, 在本章剩下的各节中, 除非特别声明, 我们总假设 $F=\mathbb{C}$. 由代数基本定理知, 任何一个非常数的复系数多项式都有复根, 因此复数域 $\mathbb{C}$ 上的任意一个 n 阶方阵恰有 n 个特征值 (可能有相同的). 求出特征值λ后, 求解线性方程组(6.4)即可求得对应λ的特征向量.

复数域 $\mathbb{C}$ 上的方阵 A 的特征值和特征向量的算法可归纳如下

(1) 计算特征多项式 $p_A(\lambda)=\det(\lambda I-A)$. 设

$$p_A(\lambda)=(\lambda-\lambda_1)^{n_1}(\lambda-\lambda_2)^{n_2}\cdots(\lambda-\lambda_s)^{n_s},$$

其中 $\lambda_i\in\mathbb{C}, n_i\geq 1\,(i=1,2,\cdots,s)$ 且 $n_1+n_2+\cdots+n_s=n$. 则 $\lambda_1,\lambda_2,\cdots,\lambda_s$ 为 A 的全部不同的特征值, 其重数分别为 $n_1,n_2,\cdots,n_s$.

(2) 对每个特征值 λ_i, 求解方程组

$$(\lambda_i I-A)\mathbf{x}=0.$$

设 $\mathbf{x}_{i1},\mathbf{x}_{i2},\cdots,\mathbf{x}_{im_i}$ 为它的一个基础解系, 则所有的非零线性组合

$$c_1\mathbf{x}_{i1}+c_2\mathbf{x}_{i2}+\cdots+c_{m_i}\mathbf{x}_{im_i}$$

为 A 的属于 λ_i 的所有特征向量.

例6.2.2. 求矩阵A的全部特征值和特征向量, 这里

$$A=\begin{pmatrix}3&-1&-2\\2&0&-2\\2&-1&-1\end{pmatrix}.$$

解. A 的特征多项式为

$$p_A(\lambda)=\det(\lambda I-A)=\begin{vmatrix}\lambda-3&1&2\\-2&\lambda&2\\-2&1&\lambda+1\end{vmatrix}=\lambda(\lambda-1)^2.$$

由 $p_A(\lambda)=\lambda(\lambda-1)^2=0$, 得到 A 的全部特征值为 $\lambda_1=0, \lambda_2=\lambda_3=1$. 下面求各个特征值对应的特征向量.

对于 $\lambda_1=0$, 解线性方程组 $(0I-A)X=0$, 即解

$$\begin{pmatrix}-3&1&2\\-2&0&2\\-2&1&1\end{pmatrix}\begin{pmatrix}x_1\\x_2\\x_3\end{pmatrix}=\begin{pmatrix}0\\0\\0\end{pmatrix}$$

得特征值 0 对应的特征向量为 $c_1(1,1,1)^T\ \ (c_1\neq 0)$.
对于 $\lambda_2=\lambda_3=1$, 解方程组

$$\begin{pmatrix}-2&1&2\\-2&1&2\\-2&1&2\end{pmatrix}\begin{pmatrix}x_1\\x_2\\x_3\end{pmatrix}=\begin{pmatrix}0\\0\\0\end{pmatrix}$$

得特征值 1 对应的特征向量为 $c_2(1,2,0)^T+c_3(0,-2,1)^T\ \ (c_2,c_3$不同时为零). □

例6.2.3. 设 λ 为 n 阶方阵 A 的一个特征值, 证明:

(1) λ^k 为 A^k 的特征值, 其中 k 为正整数;

(2) λ 为 A^T 的特征值;

(3) 若 $\lambda\neq 0$, 则 $\dfrac{1}{\lambda}\det(A)$ 为 A 的伴随方阵 A^* 的特征值.

证明. (1) 设 $A\mathbf{x}=\lambda\mathbf{x}$, 则

$$A^k\mathbf{x}=A^{k-1}(A\mathbf{x})=A^{k-1}(\lambda\mathbf{x})=\lambda A^{k-1}\mathbf{x}=\cdots=\lambda^k\mathbf{x}.$$

(2) 由于 $\det(\lambda I-A)=0$, 转置后得 $\det(\lambda I-A^T)=\det(\lambda I-A)^T=0$, 所以 λ 为 A^T 的特征值.

(3) 因为 $AA^*=\det(A)I$, 由 $\det(\lambda I-A)=0$ 知 $\det(A^*)\det(\lambda I-A)=0$, 即 $\det(\lambda A^*-A^*A)=0$. 所以 $\det(A^*-\frac{1}{\lambda}\det(A)I)=0$, 从而 $\dfrac{1}{\lambda}\det(A)$ 为 A^* 的特征值.

□

特征值是矩阵非常重要的量. 下面我们给出特征值的几个主要结论.

设 $A=(a_{ij})$ 为 $\mathbb{C}$ 上的一个 n 阶方阵, 记

$$p_A(\lambda)=\det(\lambda I-A)=\lambda^n+\sigma_1\lambda^{n-1}+\cdots+\sigma_{n-1}\lambda+\sigma_n.$$

取 $\lambda=0$, 可得 $\sigma_n=(-1)^n\det(A)$. 从行列式的完全展开式可以看出, 除了主对角线上 n 个元素的乘积, 其它乘积项中 λ 的次数不超过 $n-2$, 因此 λ^{n-1} 项只能出现在 $(\lambda-a_{11})(\lambda-a_{22})\cdots(\lambda-a_{nn})$ 中, 从而 $\sigma_1=\sum\limits_{i=1}^{n}a_{ii}$.

另一方面, 假设 A 的 n 个特征值为 $\lambda_1, \lambda_2, \cdots, \lambda_n$ (可能有相同的), 则

$$p_A(\lambda) = (\lambda - \lambda_1)(\lambda - \lambda_2)\cdots(\lambda - \lambda_n).$$

对比上面两式, 我们得到下面的命题.

命题6.2.2. 设 $A = (a_{ij})$ 为 $\mathbb{C}$ 上的一个 n 阶方阵, $\lambda_1, \lambda_2, \cdots, \lambda_n$ 为 A 的 n 个特征值. 则

(1) $Tr(A) = \sum_{i=1}^{n} \lambda_i$,

(2) $\det(A) = \lambda_1\lambda_2\cdots\lambda_n$.

推论6.2.1. n 阶方阵可逆当且仅当它的 n 个特征值都不为零.

我们给一个例子说明上述命题的应用.

例6.2.4. 设 n 阶方阵 A 的 n 个特征值分别为 $\lambda_1, \lambda_2, \cdots, \lambda_n$, 求 $I+A$ 的特征值及 $\det(I+A)$.

解. 由于 $\det(\lambda I - A) = \prod_{i=1}^{n}(\lambda - \lambda_i)$, 因此

$$\det(\lambda I - (I+A)) = \det((\lambda - 1)I - A) = \prod_{i=1}^{n}(\lambda - 1 - \lambda_i).$$

所以 $I+A$ 的 n 个特征值为 $1+\lambda_1, 1+\lambda_2, \cdots, 1+\lambda_n$, 从而 $\det(I+A) = \prod_{i=1}^{n}(1+\lambda_i)$. □

§6.3 矩阵的相似

§6.3.1 矩阵相似的定义

设A为n阶复方阵, 其n个特征值为$\lambda_1, \lambda_2, \ldots, \lambda_n$, 它们对于的特征向量分别为$\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n$. 假设这$n$ 个特征向量线性无关, 令矩阵$P = (\mathbf{x}_1, \ldots, \mathbf{x}_n)$, 则$P$为可逆矩阵. 由$A\mathbf{x}_i = \lambda_i\mathbf{x}_i, i = 1, 2, \ldots, n$可得

$$\begin{aligned} AP &= A(\mathbf{x}_1, \ldots, \mathbf{x}_n) = (A\mathbf{x}_1, \ldots, A\mathbf{x}_n) = (\lambda_1\mathbf{x}_1, \ldots, \lambda_n\mathbf{x}_n) \\ &= (\mathbf{x}_1, \ldots, \mathbf{x}_n)\,\mathrm{diag}(\lambda_1, \ldots, \lambda_n) = P\,\mathrm{diag}(\lambda_1, \ldots, \lambda_n), \end{aligned}$$

从而

$$A = P\,\mathrm{diag}(\lambda_1, \ldots, \lambda_n)P^{-1}. \tag{6.6}$$

此时我们称矩阵A与对角$\mathrm{diag}(\lambda_1, \ldots, \lambda_n)$阵相似.

定义6.3.1. 设$A, B \in F^{n\times n}$. 若存在可逆方阵$T \in F^{n\times n}$满足$B = T^{-1}AT$, 则称方阵A与B(在数域F上)相似.

矩阵相似是一种重要的等价关系. 容易验证, 相似矩阵满足以下性质

(1) (反身性) A与A相似;

(2) (对称性) 若 A与B相似, 则 B与A相似;

(3) (传递性) 若 A与B相似, B与C相似, 则 A与C相似.

由于相似关系为等价关系, 可以将 n 阶方阵按相似关系进行分类: 将相互之间相似的方阵归成一类. 两个类要么是一样的, 要么就不相交. 每个类称之为一个**相似类**, 该类中的每个元素都称为一个**代表元**.

类似于矩阵相抵关系, 研究矩阵相似关系需要解决以下两个根本问题

(1) 两个矩阵属于同一个相似类的条件是什么?

(2) 每个相似类中, 最简单的矩阵具有什么形式?

第一个问题即是要判定两个矩阵是否相似的问题, 而第二个问题就是所谓的**相似标准形**问题, 它等价于在空间中找到一组适当的基, 使得给定的线性变换在该基下的矩阵具有最简单的形式.

不同于矩阵相抵的情形, 回答矩阵相似的上述问题要困难许多, 我们将在本章以后各节将对上述问题做初步解答.

相似矩阵具有相同的重要基本量. 例如, 若A与B相似, 则A与B的行列式与秩均相等. 实际上, 它们的特征值(从而迹)也相同. 我们称这些量为 **相似不变量**.

命题6.3.1. 相似的矩阵具有相同的特征多项式和特征值.

证明. 设 $B = P^{-1}AP$, 其中 T 为可逆方阵. 则

$$\det(\lambda I - B) = \det(\lambda I - P^{-1}AP) = \det(P^{-1}(\lambda I - A)P) = \det(\lambda I - A).$$

因此 A 和 B 有相同的特征多项式, 从而有相同的特征值, 并且它们的迹也相同.

□

例6.3.1. 已知 矩阵A与B相似, 求 x 和 y. 这里

$$A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & x \end{pmatrix}, \qquad B = \begin{pmatrix} y & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

解. 相似的方阵有相同的特征多项式, 所以 $\det(\lambda I-A)=\det(\lambda I-B)$, 即

$$(\lambda-1)(\lambda(\lambda-x)-1)=(\lambda-y)(\lambda+1)(\lambda-1).$$

比较系数可得 $x=0, y=1$.

另一个解法是利用相似的矩阵具有相同的行列式和迹, 可以得到 $-1=-y$ 且 $1+x=y$, 从而 $x=0, y=1$. □

§6.3.2 矩阵相似于对角阵的条件

如果一个方阵相似于对角阵, 则称该方阵**可对角化**, 也称相应的线性变换**可对角化**. 下面的例子表明, 并不是每个方阵都可对角化.

例6.3.2. 证明: 2 阶方阵 $A=\begin{pmatrix}2&1\\0&2\end{pmatrix}$ 不能相似于对角阵.

证明. 假设 A 能够相似于对角阵 B. 由于 A 的两个特征值都是 2, 而特征值是相似不变量, 因此 B 的两个特征值也都是 2, 所以 $B=2I_2$. 由 A 相似于 B 知, 存在 2 阶可逆方阵 T, 使得

$$A=T^{-1}BT=T^{-1}(2I_2)T=2I_2.$$

这显然是矛盾的, 因此 A 不可能相似于对角阵. □

下面我们给出一个矩阵相似于对角阵的条件.

定理6.3.2. *数域 F 上的 n 阶方阵 A 相似于对角阵的充分必要条件是, A 有 n 个线性无关的特征向量.*

证明. 充分性. 由 (6.6), 若设 A 有 n 个线性无关的特征向量, 则A相似于对角阵.

必要性. 设存在可逆方阵 P, 使得

$$P^{-1}AP=\mathrm{diag}(\lambda_1,\lambda_2,\cdots,\lambda_n).$$

记 $P=(P_1,P_2,\cdots,P_n)$, 其中 $P_i\in F^n\,(i=1,2,\cdots,n)$ 为矩阵 P 的第 i 列, 则

$$\begin{aligned}&(AP_1,AP_2,\cdots,AP_n)=(P_1,P_2,\cdots,P_n)\,\mathrm{diag}(\lambda_1,\lambda_2,\cdots,\lambda_n)\\&=(\lambda_1P_1,\lambda_2P_2,\cdots,\lambda_nP_n).\end{aligned}$$

因此 $AP_i=\lambda_iP_i\ i=1,2,\cdots,n$. 所以 $P_1,P_2,\cdots,P_n$ 为 A 的 n 个特征向量. 由于这 n 个向量构成的矩阵 T 是可逆的, 它们是线性无关的. □

若矩阵 A 相似于对角阵, 则该对角阵的 n 个主对角线元素恰为 A 的 n 个特征值, 因此如果不计主对角线上元素的先后次序, 该对角阵是唯一的.

例6.3.3. 在例 6.2.2中, 矩阵 A 有三个线性无关特征向量:

$$T_1=\begin{pmatrix}1\\1\\1\end{pmatrix},\ T_2=\begin{pmatrix}1\\2\\0\end{pmatrix}\ T_3=\begin{pmatrix}0\\-2\\1\end{pmatrix},$$

且 $AT_1=0, AT_2=T_2, AT_3=T_3$. 因此, A相似于对角阵. 若令 $T=(T_1,T_2,T_3)$, 则有

$$T^{-1}AT=\begin{pmatrix}0&0&0\\0&1&0\\0&0&1\end{pmatrix}.$$

下面我们给出矩阵可对角化的进一步条件, 为此先介绍一个引理.

引理6.3.1. *设 A 为数域 F 上的 n 阶方阵, 则属于 A 的不同特征值的特征向量是线性无关的.*

证明. 设 $\lambda_1,\lambda_2,\cdots,\lambda_k$ 为 A 的互不相同的特征值, $\mathbf{x}_1,\mathbf{x}_2,\cdots,\mathbf{x}_k$ 为相应于它们的特征向量. 对 k 用数学归纳法证明. 当 $k=1$ 时, $\mathbf{x}_1\neq\mathbf{0}$, 它是线性无关的. 假设 $k-1$ 时命题成立, 下面证明命题对 k 成立.

假设

$$\mu_1\mathbf{x}_1+\mu_2\mathbf{x}_2+\cdots+\mu_k\mathbf{x}_k=\mathbf{0} \tag{6.7}$$

用 λ_k 乘 (6.7) 式两端, 得

$$\mu_1\lambda_k\mathbf{x}_1+\mu_2\lambda_k\mathbf{x}_2+\cdots+\mu_k\lambda_k\mathbf{x}_k=\mathbf{0}$$

用 A 左乘 (6.7) 式, 得

$$\mu_1\lambda_1\mathbf{x}_1+\mu_2\lambda_2\mathbf{x}_2+\cdots+\mu_k\lambda_k\mathbf{x}_k=\mathbf{0}$$

前两式相减得

$$\mu_1(\lambda_k-\lambda_1)\mathbf{x}_1+\mu_2(\lambda_k-\lambda_2)\mathbf{x}_2+\cdots+\mu_{k-1}(\lambda_k-\lambda_{k-1})\mathbf{x}_{k-1}=\mathbf{0}.$$

由归纳假设, $\mathbf{x}_1,\mathbf{x}_2,\cdots,\mathbf{x}_{k-1}$ 线性无关. 因此

$$\mu_j(\lambda_k-\lambda_j)=0,\quad j=1,2,\cdots,k-1,$$

由于 $\lambda_k-\lambda_j\neq0$, 我们得到 $\mu_1=\mu_2=\ldots=\mu_{k-1}=0$. 再由 (6.7) 知 $\mu_k\mathbf{x}_k=\mathbf{0}$. 因为 $\mathbf{x}_k\neq\mathbf{0}$, 所以 $\mu_k=0$. 这就证明了 $\mathbf{x}_1,\mathbf{x}_2,\cdots,\mathbf{x}_k$ 线性无关. □

由定理 6.3.2及引理 6.3.1, 立即可得

推论6.3.1. 如果矩阵 A 的 n 个特征值两两不同, 则 A 相似于对角阵.

证明. 由引理 6.3.1知, A 有 n 个线性无关的特征向量. 再由定理 6.3.2知推论成立. □

例6.3.4. 设方阵 A 相似于对角阵, 求 x 和 y 应满足的条件, 这里

$$A=\begin{pmatrix}0&0&x\\1&1&y\\1&0&0\end{pmatrix}.$$

解. 由于

$$|\lambda I-A|=\begin{vmatrix}\lambda&0&-x\\-1&\lambda-1&-y\\-1&0&\lambda\end{vmatrix}=(\lambda-1)(\lambda^2-x),$$

当 $x\neq 0,1$时, A 有三个不同的特征值, 由推论 6.3.1知, 此时 A 可对角化.

当 $x=1$ 时, A 有特征值 $\lambda_1=\lambda_2=1,\ \lambda_3=-1$. 对应$\lambda_1=\lambda_2=1$的特征向量为: 当$y\neq-1$时, $T_1=(0,1,0)^T$; 当$y=-1$时, $T_1=(1,0,1)^T$, $T_2=(1,1,1)^T$. 对应$\lambda_3=-1$的特征向量为: $T_3=(1,(y-1)/2,-1)^T$. 因此, 当$y=-1$时, 得到三个线性无关特征向量T_1,T_2,T_3, 从而 A 可以对角化.

当 $x=0$ 时, A 有特征值 $\lambda_1=\lambda_2=0,\ \lambda_3=1$. 对应特征值$\lambda_1=\lambda_2=0$, 特征向量$T_1=(0,-y,1)^T$. 对应特征值$\lambda_3=1$, 特征向量$T_2=(0,1,0)^T$. 由于只有两个线性无关特征向量, 故此时 A 不相似于对角阵.

综上所述, A 可对角化的条件是 $x\neq 0,1$ 或者 $x=1$ 且 $y=-1$. □

例6.3.5. 设 n 阶方阵 A 满足 $\mathrm{rank}(A+I)+\mathrm{rank}(A-I)=n$, 证明 A 可对角化.

证明. 设 $\mathrm{rank}(A+I)=r$, 则方程组 $(A+I)X=0$ 的解空间是 $n-r$ 维的, 它是 A 的特征值 -1 对应的特征子空间. 取 $X_{r+1},\cdots,X_n$ 为它的一组基.

由条件知 $\mathrm{rank}(A-I)=n-r$, 所以 $(A-I)X=0$ 的解空间是 r 维的, 它是 A 的特征值 1 对应的特征子空间. 取 $\mathbf{x}_1,\mathbf{x}_2,\cdots,\mathbf{x}_r$ 为它的一组基.

下证$\mathbf{x}_1,\mathbf{x}_2,\ldots,\mathbf{x}_n$ 线性无关. 实际上, 设

$$c_1\mathbf{x}_1+c_2\mathbf{x}_2+\ldots+c_n\mathbf{x}_n=\mathbf{0},$$

其中$c_1,c_2,\ldots,c_n$为常数. 上式两边作用 A 得

$$c_1\mathbf{x}_1+\ldots+c_r\mathbf{x}_r-c_{r+1}\mathbf{x}_{r+1}-\ldots-c_n\mathbf{x}_n=\mathbf{0}.$$

于是

$$c_1\mathbf{x}_1+\ldots+c_r\mathbf{x}_r=c_{r+1}\mathbf{x}_{r+1}+\ldots+c_n\mathbf{x}_n=\mathbf{0}.$$

由于$\mathbf{x}_1,\ldots,\mathbf{x}_r$, $\mathbf{x}_{r+1},\ldots,\mathbf{x}_n$分别线性无关, 因此, $c_1=\ldots=c_{r+1}=\ldots=c_n=0$, 即$\mathbf{x}_1,\mathbf{x}_2,\ldots,\mathbf{x}_n$ 线性无关. 于是A 可对角化.

实际上, 令 $P=(\mathbf{x}_1,\cdots,\mathbf{x}_r,\mathbf{x}_{r+1},\cdots,\mathbf{x}_n)$, 则 P 可逆, 且

$$AP=A(\mathbf{x}_1,\cdots,\mathbf{x}_r,\mathbf{x}_{r+1},\cdots,\mathbf{x}_n)=P\begin{pmatrix} I_r & 0\\ 0 & -I_{n-r}\end{pmatrix}.$$

所以$A=P\begin{pmatrix} I_r & 0\\ 0 & -I_{n-r}\end{pmatrix}P^{-1}$. □

§6.3.3 相似于上三角阵

虽然不是每个复方阵都可以相似于对角阵, 但我们可以证明它总可以相似于一个上三角阵.

定理6.3.3. 任何一个 n 阶复方阵 A 都可以相似于一个上三角阵, 且该上三角阵的主对角线上的元素都是 A 的特征值.

证明. 对方阵 A 的阶数 n 用数学归纳法. 当 $n=1$ 时, 命题显然成立. 假设命题对 $n-1$ 阶方阵成立, 现在考虑 n 阶方阵 A. 设 λ_1 为 A 的一个特征值, $\mathbf{x}_1$ 为属于 λ_1 的一个特征向量. 将 $\mathbf{x}_1$ 扩充为 $\mathbf{C}^n$ 的一组基$\mathbf{x}_1,\mathbf{x}_2,\cdots,\mathbf{x}_n$. 令$P_1=(\mathbf{x}_1,\mathbf{x}_2,\cdots,\mathbf{x}_n)$, 则 P_1 为 n 阶可逆方阵. 由 $A\mathbf{x}_1=\lambda_1\mathbf{x}_1$ 知

$$AP_1=A(\mathbf{x}_1,\mathbf{x}_2,\cdots,\mathbf{x}_n)=(\lambda_1\mathbf{x}_1,A\mathbf{x}_2,\cdots,A\mathbf{x}_n)=(\mathbf{x}_1,\mathbf{x}_2,\cdots,\mathbf{x}_n)\begin{pmatrix}\lambda_1 & *\\ 0 & A_1\end{pmatrix}.$$

其中 A_1 为一个 $(n-1)$ 阶方阵. 所以 $P_1^{-1}AP_1=\begin{pmatrix}\lambda_1 & *\\ 0 & A_1\end{pmatrix}$.

根据归纳假设, 存在 $n-1$ 阶可逆方阵 P_2, 使得 $P_2^{-1}A_1P_2$ 为上三角阵. 令 $P=P_1\begin{pmatrix}1 & 0\\ 0 & P_2\end{pmatrix}$, 则

$$\begin{aligned}P^{-1}AP&=\begin{pmatrix}1 & 0\\ 0 & P_2^{-1}\end{pmatrix}(P_1^{-1}AP_1)\begin{pmatrix}1 & 0\\ 0 & P_2\end{pmatrix}\\ &=\begin{pmatrix}1 & 0\\ 0 & P_2^{-1}\end{pmatrix}\begin{pmatrix}\lambda_1 & *\\ 0 & A_1\end{pmatrix}\begin{pmatrix}1 & 0\\ 0 & P_2\end{pmatrix}\\ &=\begin{pmatrix}\lambda_1 & *\\ 0 & P_2^{-1}A_1P_2\end{pmatrix}\end{aligned}$$

由于 $P_2^{-1}A_1P_2$ 为上三角阵, $P^{-1}AP$ 为上三角阵. 因为上三角阵的主对角线元素都是它的特征值, 而特征值是相似不变量, 所以 $P^{-1}AP$ 的主对角线元素都是 A 的特征值. □

在上述定理证明中, 矩阵 P_1, P_2 的选取方法不是唯一的, 最后得到的上三角阵也不是唯一的.

例6.3.6. 设 2 阶方阵 A 满足$A^2=0$, 证明 $A=0$ 或者 A 相似于$\begin{pmatrix}0&1\\0&0\end{pmatrix}$.

证明. 设 λ 为 A 的一个特征值, $A\mathbf{x}=\lambda\mathbf{x}$. 则 $0=A^2\mathbf{x}=\lambda^2\mathbf{x}$. 由于 $\mathbf{x}$ 是非零向量, 所以有 $\lambda=0$, 即 A 的两个特征值都是 0. 由定理 6.3.3知, A 相似于上三角矩阵 $\begin{pmatrix}0&a\\0&0\end{pmatrix}$. 若 $a=0$, 则 $A=0$. 否则利用初等变换很容易将其进一步相似于 $\begin{pmatrix}0&1\\0&0\end{pmatrix}$. 例如取 $P=\begin{pmatrix}a^{-1}&0\\0&1\end{pmatrix}$, 则 $P\begin{pmatrix}0&a\\0&0\end{pmatrix}P^{-1}=\begin{pmatrix}0&1\\0&0\end{pmatrix}$. □

例6.3.7. 设$A\in\mathbb{C}^{2\times 2}$. 证明: $A^2-Tr(A)A+\det(A)I_2=0$.

证明. 一种直接的证明方法是令$A=(a_{ij})$, 将A的表达式代入等式验证等式成立. 实际上, 我们可以利用定理 6.3.3 将证明简化. 由定理 6.3.3, 存在上三角矩阵$B=\begin{pmatrix}b_{11}&b_{12}\\0&b_{22}\end{pmatrix}$及二阶可逆方阵$P$使得$A=PBP^{-1}$. 于是

$$A^2-Tr(A)A+\det(A)I_2=PB^2P^{-1}-Tr(B)PBP^{-1}+\det(B)I_2=P(B^2-Tr(B)B+\det(B)I_2)P^{-1}.$$

这样只需要证: $B^2-Tr(B)B+\det(B)I_2=0$. 注意到,

$$p_B(\lambda)=\lambda^2-Tr(B)\lambda+\det(B)=(\lambda-b_{11})(\lambda-b_{22}),$$

于是我们需要证明:$p_B(B)=(B-b_{11}I_2)(B-b_{22}I_2)=0$. 由于

$$p_B(B)=(B-b_{11}I_2)(B-b_{22}I_2)=\begin{pmatrix}0&b_{12}\\0&b_{22}-b_{11}\end{pmatrix}\begin{pmatrix}b_{11}-b_{22}&b_{12}\\0&0\end{pmatrix}=0.$$

因此, $A^2-Tr(A)A+\det(A)I_2=0$. □

§6.4 一般空间上的线性变换

本节我们将数组空间上的线性变换理论推广到一般线性空间上.

§6.4.1 线性变换的定义

从第一节我们知道, 线性映射的关键特性是保持线性组合的不变性, 即(6.2)式成立. 据此, 我们给出一般线性空间上线性映射的定义.

定义6.4.1. 设 V, V' 为数域 F 上的两个线性空间, 若映射 $\mathscr{A}: V \to V'$ 满足: 对任意 $\alpha, \beta \in V, \lambda \in F$, 都有

$$\mathscr{A}(\lambda+\beta) = \mathscr{A}(\alpha) + \mathscr{A}(\beta) \tag{6.8}$$

$$\mathscr{A}(\lambda\alpha) = \lambda\mathscr{A}(\alpha) \tag{6.9}$$

则称 $\mathscr{A}$ 为从线性空间 V 到线性空间 V' 的**线性映射**. 特别地, 如果 $V' = V$, 则称 $\mathscr{A}$ 为线性空间 V 上的一个**线性变换**.

本书只讨论线性变换,但关于线性变换的很多结论对线性映射同样成立. 我们先给出几个一般线性空间上线性变换的例子.

例6.4.1. 把每个向量映为自身的变换

$$\mathscr{E}: \mathscr{E}(x) = x, \quad x \in V.$$

是线性变换, 称为**单位变换**或**恒等变换**.

把空间的每个向量都映为零向量的变换 $\mathscr{O}: \mathscr{O}(x) = 0, x \in V$ 也是一个线性变换, 称为**零变换**.

例6.4.2. 设$P_n[x]$是次数不超过n的多项式全体, $\mathscr{A}$ 为微分算子

$$\mathscr{A}(p(x)) = \frac{\mathrm{d}}{\mathrm{d}x}p(x).$$

由求导法则知 $\mathscr{A}$ 为线性变换.

例6.4.3. 用 $C[a,b]$ 表示闭区间 $[a,b]$ 上所有实值连续函数构成的集合. 映射 $\mathscr{A}: C[a,b] \to C[a,b]$ 定义为:对每个 $f \in C[a,b]$,

$$\mathscr{A}(f)(x) = \int_a^b K(x,t)f(t)\mathrm{d}t,$$

其中 $K(x,t)$ 是给定的 $[a,b]\times[a,b]$ 上的实值连续函数. 由积分的性质知 $\mathscr{A}$ 为线性变换.

例6.4.4. 设V是区间$[a,b]$上的函数构成的线性空间. 映射$\mathscr{A}: V \to V$定义为: 对任意$f(x) \in V$,

$$\mathscr{A}: \quad f(x) \mapsto (1-x)f(a) + xf(b).$$

则$\mathscr{A}$是线性变换.

下面介绍线性变换的简单性质.

定理6.4.1. 设V是数域F上的线性空间, $\mathscr{A}$是V上的线性变换. $\mathscr{A}$具有以下性质

(1) $\mathscr{A}(\theta)=\theta$;

(2) $\mathscr{A}(-\alpha)=-\mathscr{A}(\alpha)$, $\alpha\in V$.

(3) $\mathscr{A}(\lambda_1\alpha_1+\lambda_2\alpha_2+\cdots+\lambda_n\alpha_n)=\lambda_1\mathscr{A}(\alpha_1)+\lambda_2\mathscr{A}(\alpha_2)+\cdots+\lambda_n\mathscr{A}(\alpha_n)$.

(4) 设 $\alpha_1,\alpha_2,\ldots,\alpha_n$ 为线性空间 V 的一组基, 若 $\alpha=\lambda_1\alpha_1+\lambda_2\alpha_2+\cdots+\lambda_n\alpha_n$, 则

$$\mathscr{A}(\alpha)=\lambda_1\mathscr{A}(\alpha_1)+\lambda_2\mathscr{A}(\alpha_2)+\cdots+\lambda_n\mathscr{A}(\alpha_n).$$

(5) 若 $\alpha_1,\alpha_2,\cdots,\alpha_m$ 为 V 中线性相关的向量, 则 $\mathscr{A}(\alpha_1),\mathscr{A}(\alpha_2),\cdots,\mathscr{A}(\alpha_m)$ 也线性相关.

证明. (1)–(4)请读者自己完成. 下面我们证性质(5).

设 $\alpha_1,\alpha_2,\cdots,\alpha_m$ 线性相关, 则存在不全为零的 $\lambda_1,\lambda_2,\cdots,\lambda_m\in F$, 使得

$$\lambda_1\alpha_1+\lambda_2\alpha_2+\cdots+\lambda_m\alpha_m=0.$$

两边用线性变换 $\mathscr{A}$ 作用后得到

$$\begin{aligned}&\lambda_1\mathscr{A}(\alpha_1)+\lambda_2\mathscr{A}(\alpha_2)+\cdots+\lambda_m\mathscr{A}(\alpha_m)\\&=\mathscr{A}(\lambda_1\alpha_1+\lambda_2\alpha_2+\cdots+\lambda_m\alpha_m)=\mathscr{A}(0)=0.\end{aligned}$$

因此 $\mathscr{A}(\alpha_1),\mathscr{A}(\alpha_2),\cdots,\mathscr{A}(\alpha_m)$ 线性相关. □

性质(4)说明, 如果知道线性变换 $\mathscr{A}$ 在一组基下的像给定, 则每个向量的像也就唯一确定了. 性质(5)则表明, 线性相关的向量组经过线性变换后, 仍保持线性相关性. 特别地, 将它应用到三维几何空间就意味着线性变换把共线的向量映为共线的向量, 把共面的向量映为共面的向量.

定理6.4.2. 设$\mathscr{A}$是线性空间V上的线性变换. $\mathscr{A}$的像的全体$Im(\mathscr{A}):=\{A\alpha\,|,\alpha\in V\}$是$V$的子空间, 称为$\mathscr{A}$ 的**像空间**. $\mathscr{A}$的核$Ker(\mathscr{A}):=\{\alpha\in V\,|A\alpha=\theta\}$也为$V$的子空间, 称为$\mathscr{A}$的**核空间**.

证明. 设$\alpha,\beta\in Im(\mathscr{A})$, 则存在$\tilde\alpha,\tilde\beta\in V$使得$\alpha=\mathscr{A}\tilde\alpha,\beta=\mathscr{A}\tilde\beta$. 于是对$\lambda,\mu\in F$,

$$\lambda\alpha+\mu\beta=\mathscr{A}(\lambda\tilde\alpha+\mu\tilde\beta)\in Im(\mathscr{A}),$$

从而$Im(\mathscr{A})$是V的子空间.

另一方面, 设$\alpha,\beta\in Ker(\mathscr{A})$, 则$\mathscr{A}\alpha=\theta$, $\mathscr{A}\beta=\theta$. 于是

$$\mathscr{A}(\lambda\alpha+\mu\beta)=\lambda\mathscr{A}\alpha+\mu\mathscr{A}\beta=\theta,$$

即$\lambda\alpha+\mu\beta\in Ker(\mathscr{A})$. 故$Ker(\mathscr{A})$为$V$的子空间. □

§6.4.2 线性变换在一组基下的矩阵

设$\mathscr{A}$是线性空间V上的线性变换, $\alpha_1,\alpha_2,\ldots,\alpha_n$是$V$的一组基. 显然, $\mathscr{A}(\alpha_i)$ 可以唯一地表示为这组基的线性组合, 即存在 $a_{ij}\in F\ (i,j=1,2,\cdots,n)$满足

$$\begin{cases}\mathscr{A}(\alpha_1)=a_{11}\alpha_1+a_{21}\alpha_2+\cdots+a_{n1}\alpha_n\\ \mathscr{A}(\alpha_2)=a_{12}\alpha_1+a_{22}\alpha_2+\cdots+a_{n2}\alpha_n\\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\\ \mathscr{A}(\alpha_n)=a_{1n}\alpha_1+a_{2n}\alpha_2+\cdots+a_{nn}\alpha_n\end{cases}$$

记$\mathscr{A}(\alpha_1,\alpha_2,\cdots,\alpha_n)=(\mathscr{A}(\alpha_1),\mathscr{A}(\alpha_2),\cdots,\mathscr{A}(\alpha_n))$, 则上式可以改写成

$$\mathscr{A}(\alpha_1,\alpha_2,\cdots,\alpha_n)=(\alpha_1,\alpha_2,\cdots,\alpha_n)A, \tag{6.10}$$

其中矩阵$A=(a_{ij})_{n\times n}$称为**线性变换 $\mathscr{A}$ 在基 $\alpha_1,\alpha_2,\cdots,\alpha_n$下 的矩阵**. 注意 (6.10) 式右端的 $(\alpha_1,\alpha_2,\cdots,\alpha_n)$ 并非通常意义下由数构成的向量, 但将α_i视为向量元素, 右端的向量与矩阵的乘法仍然是有意义的. 特别地, 当$V=F^n$ 且$\alpha_1,\alpha_2,\cdots,\alpha_n$为列向量时, 可以将$(\alpha_1,\alpha_2,\cdots,\alpha_n)$及$\mathscr{A}(\alpha_1,\alpha_2,\ldots,\alpha_n)$ 视为n阶方阵. 另一方面, 注意矩阵 A 是由变换 $\mathscr{A}$ 及基 $\alpha_1,\alpha_2,\cdots,\alpha_n$ 唯一确定的, 它的第 j 列为 $\mathscr{A}(\alpha_j)$ 在基 $\alpha_1,\alpha_2,\cdots,\alpha_n$ 下的坐标.

例6.4.5. 线性变换 (6.1) 在自然基$\mathbf{e}_1,\mathbf{e}_2,\ldots,\mathbf{e}_n$下的矩阵正好是$A$, 这是因为$A\mathbf{e}_i$就是$A$的第$i$列. 也就是说, 线性变换在自然基下的矩阵就是定义(6.1)中的矩阵A.

例6.4.6. 设 F^3 的线性变换 $\mathscr{A}$ 将

$$\alpha_1=(2,3,5)^T,\quad \alpha_2=(0,1,2)^T,\quad \alpha_3=(1,0,0)^T$$

分别变换为

$$\beta_1=(1,2,0)^T,\quad \beta_2=(2,4,-1)^T,\quad \beta_3=(3,0,5)^T.$$

求

(1) $\mathscr{A}$ 在基 $\alpha_1,\alpha_2,\alpha_3$ 下的矩阵;

(2) $\mathscr{A}$ 在自然基下的矩阵.

解. (1) 设 $\mathscr{A}$ 在基 $\alpha_1,\alpha_2,\alpha_3$ 下的矩阵矩阵为 A, 由定义知

$$(\beta_1,\beta_2,\beta_3)=(\alpha_1,\alpha_2,\alpha_3)A.$$

所以

$$\begin{aligned}A&=(\alpha_1,\alpha_2,\alpha_3)^{-1}(\beta_1,\beta_2,\beta_3)\\&=\begin{pmatrix}2&0&1\\3&1&0\\5&2&0\end{pmatrix}^{-1}\begin{pmatrix}1&2&3\\2&4&0\\0&-1&5\end{pmatrix}=\begin{pmatrix}4&9&-5\\-10&-23&15\\-7&-16&13\end{pmatrix}.\end{aligned}$$

(2) 设变换 $\mathscr{A}$ 在自然基下的矩阵为 B, 由例 6.4.5知$\beta_i=B\alpha_i$, $i=1,2,3$. 因此 $(\beta_1,\beta_2,\beta_3)=B(\alpha_1,\alpha_2,\alpha_3)$, 从而

$$\begin{aligned}B&=(\beta_1,\beta_2,\beta_3)(\alpha_1,\alpha_2,\alpha_3)^{-1}\\&=\begin{pmatrix}1&2&3\\2&4&0\\0&-1&5\end{pmatrix}\begin{pmatrix}0&2&-1\\0&-5&3\\1&-4&2\end{pmatrix}=\begin{pmatrix}3&-20&11\\0&-16&10\\5&-15&7\end{pmatrix}.\end{aligned}$$

□

例6.4.7. 设线性空间$V=F^{2\times 2}$, $\mathscr{A}$是V上的线性变换

$$\mathscr{A}(M)=AM,\quad \forall\, M\in V$$

其中$A=\begin{pmatrix}1&2\\3&4\end{pmatrix}$. 求线性变换$\mathscr{A}$在基

$$e_1=\begin{pmatrix}1&0\\0&0\end{pmatrix}, e_2=\begin{pmatrix}0&1\\0&0\end{pmatrix}, e_3=\begin{pmatrix}0&0\\1&0\end{pmatrix}, e_4=\begin{pmatrix}0&0\\0&1\end{pmatrix}$$

下的矩阵 $\tilde{A}$.

解.

$$\mathscr{A}(e_1)=\begin{pmatrix}1&2\\3&4\end{pmatrix}\begin{pmatrix}1&0\\0&0\end{pmatrix}=\begin{pmatrix}1&0\\3&0\end{pmatrix}=1e_1+0e_2+3e_3+0e_4,$$

所以矩阵 $\tilde{A}$ 的第一列为 $(1,0,3,0)^T$. 类似可求得 $\tilde{A}$ 的其他列, 得到

$$\tilde{A}=\begin{pmatrix}1&0&2&0\\0&1&0&2\\3&0&4&0\\0&3&0&4\end{pmatrix}.$$

□

上例中的变换 $\mathscr{A}$ 虽然是通过左乘矩阵 A 得到, 但 $\mathscr{A}$ 在基本基下的矩阵 $\tilde{A}\neq A$, 它们的阶数甚至都不相同. 读者应注意它与例 6.4.5的区别.

固定线性空间V的一组基$(\alpha_1,\alpha_2,\cdots,\alpha_n)$, 则通过关系式(6.10), V上的线性变换$\mathscr{A}$与一个n阶方阵A建立了一一对应关系. 这样可以通过研究矩阵A来研究线性变换$\mathscr{A}$.

定理6.4.3. 设线性空间V上的线性变换$\mathscr{A}$在基$\{\alpha_1,\alpha_2,\cdots,\alpha_n\}$下的矩阵为$A$. $\beta=\mathscr{A}\alpha$, $\alpha,\beta\in V$.若α, β在基$\{\alpha_1,\alpha_2,\cdots,\alpha_n\}$下的坐标分别为$\mathbf{x},\mathbf{y}\in F^n$, 则

$$\mathbf{y}=A\mathbf{x}. \tag{6.11}$$

证明: 设$\mathbf{x}=(x_1,x_2,\ldots,x_n)^T$, $\mathbf{y}=(y_1,y_2,\ldots,y_n)^T$. 则

$$\begin{aligned}\beta=\mathscr{A}(\alpha)&=\mathscr{A}(x_1\alpha_1+x_2\alpha_2+\cdots+x_n\alpha_n)\\&=x_1\mathscr{A}(\alpha_1)+x_2\mathscr{A}(\alpha_2)+\cdots+x_n\mathscr{A}(\alpha_n)\\&=\big(\mathscr{A}(\alpha_1),\mathscr{A}(\alpha_2),\cdots,\mathscr{A}(\alpha_n)\big)\mathbf{x}\\&=\big((\alpha_1,\alpha_2,\cdots,\alpha_n)A\big)\mathbf{x}\\&=(\alpha_1,\alpha_2,\cdots,\alpha_n)(A\mathbf{x})\end{aligned}$$

由于一个向量在一组基下的坐标是唯一的, 我们立即得到(6.11)式.

□

上述定理表明, $\beta=\mathscr{A}\alpha$的矩阵形式就是$\mathbf{y}=A\mathbf{x}$!

例6.4.8. 设$\mathscr{A}$是$\mathbb{R}^2$上绕坐标原点逆时针旋转θ角的线性变换, $\mathbf{e}_1,\mathbf{e}_2$是$\mathbb{R}^2$的单位坐标向量. 则

$$\mathscr{A}(\mathbf{e}_1,\mathbf{e}_2)=(\mathbf{e}_1,\mathbf{e}_2)\begin{pmatrix}\cos(\theta)&-\sin(\theta)\\\sin(\theta)&\cos(\theta)\end{pmatrix}$$

于是由定理 6.4.3, 点$P=(x,y)^T$经$\mathscr{A}$变换后的点$P'=(x',y')^T$的坐标为

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}.$$

特别注意此例与第四章例 5.4.5的联系与区别.

在第四章中, 我们研究过基变换与坐标变换, 得到了坐标变换公式(5.14). 它与定理 6.4.3中的的坐标变换公式有何区别呢? 实际上, 坐标变换公式(5.14)表示的是坐标系做了改变, 同一个向量在新坐标系与旧坐标系下的坐标之间的关系. 而 (6.11) 表示的是, 坐标系不变, 但向量做了线性变换, 变换前后的向量的坐标之间的关系.

下面我们研究线性变换的像空间与核空间的结构.

定理6.4.4. 设线性空间V上的线性变换$\mathscr{A}$在基$\{\alpha_1,\alpha_2,\cdots,\alpha_n\}$下的矩阵为$A$. 则

1. $Im(\mathscr{A})=\{(\alpha_1,\alpha_2,\cdots,\alpha_n)\mathbf{y}\,|\,\mathbf{y}\in C(A)\}$, $Ker(\mathscr{A})=\{(\alpha_1,\alpha_2,\cdots,\alpha_n)\mathbf{x}\,|\,\mathbf{x}\in N(A)\}$. 其中$C(A)$,$N(A)$分别为矩阵$A$的列空间与零空间.

2. $\dim(Im\mathscr{A})=r(A)$, $\dim(Ker\mathscr{A})=n-r(A)$, $\dim(Im(\mathscr{A}))+\dim(KerIm(\mathscr{A}))=n$.

证明. 我们只对像空间证明有关结论, 核空间的结论类似证明.

1. 对任意$\alpha=(\alpha_1,\alpha_2,\ldots,\alpha_n)\mathbf{x}\in V$, 其中$\mathbf{x}\in F^n$, 由定理 6.4.3,

$$\mathscr{A}\alpha=(\alpha_1,\alpha_2,\ldots,\alpha_n)A\mathbf{x}.$$

由于$\mathbf{y}=A\mathbf{x}\in C(A)$, 所以

$$\mathscr{A}\alpha=(\alpha_1,\alpha_2,\cdots,\alpha_n)\mathbf{y}\in\{(\alpha_1,\alpha_2,\cdots,\alpha_n)\mathbf{y}\,|\,\mathbf{y}\in C(A)\},$$

故而$Im(\mathscr{A})\subset\{(\alpha_1,\alpha_2,\cdots,\alpha_n)\mathbf{y}\,|\,\mathbf{y}\in C(A)\}$.

另一方面, 对任意$\mathbf{y}\in C(A)$, 存在$\mathbf{x}\in F^n$使得$\mathbf{y}=A\mathbf{x}$, 故

$$(\alpha_1,\alpha_2,\cdots,\alpha_n)\mathbf{y}=(\alpha_1,\alpha_2,\cdots,\alpha_n)A\mathbf{x}=\mathscr{A}((\alpha_1,\alpha_2,\ldots,\alpha_n)\mathbf{x})\in Im(\mathscr{A}).$$

因此, $\{(\alpha_1,\alpha_2,\cdots,\alpha_n)\mathbf{y}\,|\,\mathbf{y}\in C(A)\}\subset Im(\mathscr{A})$, 进而$Im(\mathscr{A})=\{(\alpha_1,\alpha_2,\cdots,\alpha_n)\mathbf{y}\,|\,\mathbf{y}\in C(A)\}$.

2. 不妨设$A_1,A_2,\ldots,A_r$为$C(A)$的一组基, 下证$(\alpha_1,\alpha_2,\ldots,\alpha_n)A_i, i=1,2,\ldots,r$ 为$Im\mathscr{A}$的一组基, 从而$\dim(Im\mathscr{A})=\dim C(A)=r(A)$.

   首先, 由于$A_1,A_2,\ldots,A_r$为$C(A)$的一组基, 对任意$\mathbf{y}\in C(A)$, $(\alpha_1,\alpha_2,\ldots,\alpha_n)\mathbf{y}$都可以由$(\alpha_1,\alpha_2,\ldots,\alpha_n)A_i, i=1,2,\ldots,r$线性表示.

   其次, 我们要说明$(\alpha_1,\alpha_2,\ldots,\alpha_n)A_i$, $i=1,2,\ldots,r$线性无关. 实际上, 设$\sum_{i=1}^r x_i(\alpha_1,\alpha_2,\ldots,\alpha_n)A_i=\theta$, 则

$$\sum_{i=1}^r x_i(\alpha_1,\alpha_2,\ldots,\alpha_n)A_i=(\alpha_1,\alpha_2,\ldots,\alpha_n)\sum_{i=1}^r x_iA_i=(\alpha_1,\alpha_2,\ldots,\alpha_n)\tilde{A}\tilde{\mathbf{x}}=\theta,$$

   这里$\tilde{A}=(A_1,A_2,\ldots,A_r)$, $\tilde{\mathbf{x}}=(x_1,x_2,\ldots,x_r)^T$. 于是$A\tilde{\mathbf{x}}=\mathbf{0}$. 由于$A_1,A_2,\ldots,A_r$ 线性无关, 故$\tilde{\mathbf{x}}=\mathbf{0}$. 从而$(\alpha_1,\alpha_2,\ldots,\alpha_n)A_i, i=1,2,\ldots,r$ 线性无关, 进而构成$Im\mathscr{A}$的一组基.

定理证毕. □

上述定理表明了, 线性变换$\mathscr{A}$的像空间与矩阵A的列空间对应, $\mathscr{A}$的核空间与A的零空间对应.

推论6.4.1. 设线性空间V上的线性变换$\mathscr{A}$在基$\{\alpha_1,\alpha_2,\cdots,\alpha_n\}$下的矩阵为$A$. 则$\mathscr{A}$可逆当且仅当矩阵$A$可逆.

证明. 由定理 6.4.4, 线性变换$\mathscr{A}$可逆当且仅当$Ker(\mathscr{A})=\{\theta\}$且$Im(\mathscr{A})=V$, 当且仅当$\mathrm{rank}(A)=n$, 当且仅当$A$可逆. □

下面我们研究一个线性变换在不同基下的矩阵之间的关系.

定理6.4.5. 设线性空间V上的线性变换 $\mathscr{A}$ 在 V 的两组基 $\alpha_1,\alpha_2,\cdots,\alpha_n$ 和 $\beta_1,\beta_2,\cdots,\beta_n$ 下的矩阵分别为 A 和 B. 设基 $\alpha_1,\alpha_2,\cdots,\alpha_n$ 到基 $\beta_1,\beta_2,\cdots,\beta_n$ 的过渡矩阵为 T, 即 $(\beta_1,\beta_2,\cdots,\beta_n)=(\alpha_1,\alpha_2,\cdots,\alpha_n)T$. 则

$$B=T^{-1}AT.$$

即矩阵A与B相似.

证明. 已知

$$\begin{aligned}\mathscr{A}(\alpha_1,\alpha_2,\cdots,\alpha_n)&=(\alpha_1,\alpha_2,\cdots,\alpha_n)A\\ \mathscr{A}(\beta_1,\beta_2,\cdots,\beta_n)&=(\beta_1,\beta_2,\cdots,\beta_n)B\\ (\beta_1,\beta_2,\cdots,\beta_n)&=(\alpha_1,\alpha_2,\cdots,\alpha_n)T\end{aligned}$$

于是

$$\begin{aligned}\mathscr{A}(\beta_1,\beta_2,\cdots,\beta_n) &= \mathscr{A}\big((\alpha_1,\alpha_2,\cdots,\alpha_n)T\big) = \mathscr{A}(\alpha_1,\alpha_2,\cdots,\alpha_n)T \\ &= (\alpha_1,\alpha_2,\cdots,\alpha_n)AT = (\beta_1,\beta_2,\cdots,\beta_n)(T^{-1}AT)\end{aligned}$$

由于线性变换 $\mathscr{A}$ 在基 $\beta_1,\beta_2,\cdots,\beta_n$ 下的矩阵是唯一的, 所以有

$$B = T^{-1}AT.$$

□

定理 6.4.5 指出, 一个线性变换在不同基下的矩阵是彼此相似的. 那么反过来, 彼此相似的方阵, 是否是同一个线性变换在不同基下对应的矩阵呢? 回答是肯定的. 事实上, 设 A 为线性变换 $\mathscr{A}$ 在基 $\alpha_1,\alpha_2,\cdots,\alpha_n$ 下的矩阵. 若 B 与 A 相似, 即存在可逆方阵 T, 使得 $B = T^{-1}AT$. 令 $(\beta_1,\beta_2,\cdots,\beta_n) = (\alpha_1,\alpha_2,\cdots,\alpha_n)T$, 则 $\beta_1,\beta_2,\cdots,\beta_n$ 也是 V 的一组基, 且不难验证 $\mathscr{A}$ 在这组基下的矩阵为 B.

由于一个线性空间上的线性变换的性质与该空间的基的选取没有关系, 通过上面的分析我们看出, n维线性空间上的线性变换与n阶方阵的相似类之间建立了一一对应关系. 因此, 通过矩阵研究线性变换的性质时, 只有所有相似矩阵都具有的性质, 才能反映线性变换的性质. 例如, 矩阵的相似不变量对应线性变换的量. 由此我们可以定义线性变换$\mathscr{A}$的行列式$\det(\mathscr{A})$、秩$\mathrm{rank}(\mathscr{A})$、迹$Tr(\mathscr{A})$以及特征值, 等.

例6.4.9. 设 $\mathscr{A}: F^3 \to F^3$ 为例 6.4.6 中的线性变换. 求 $\mathscr{A}$ 在基

$$\gamma_1 = (1,0,1)^T, \quad \gamma_2 = (2,1,2)^T, \quad \gamma_3 = (-2,0,-1)^T$$

下的矩阵.

解: 在例 6.4.6 中, 我们已求得 $\mathscr{A}$ 在基 $\alpha_1,\alpha_2,\alpha_3$ 下的矩阵为

$$B = \begin{pmatrix} 4 & 9 & -5 \\ -10 & -23 & 15 \\ -7 & -16 & 13 \end{pmatrix}.$$

设基 $\alpha_1,\alpha_2,\alpha_3$ 到基 $\gamma_1,\gamma_2,\gamma_3$ 的过渡矩阵为 T, 则

$$(\alpha_1,\alpha_2,\alpha_3)T = (\gamma_1,\gamma_2,\gamma_3).$$

所以

$$T = (\alpha_1,\alpha_2,\alpha_3)^{-1}(\gamma_1,\gamma_2,\gamma_3) = \begin{pmatrix} -1 & 0 & 1 \\ 3 & 1 & -3 \\ 3 & 2 & -4 \end{pmatrix}.$$

由定理 6.4.5 知变换 $\mathscr{A}$ 在基 $\gamma_1,\gamma_2,\gamma_3$ 下的矩阵

$$C=T^{-1}BT=\begin{pmatrix}2&2&-1\\3&1&0\\3&2&-1\end{pmatrix}\begin{pmatrix}4&9&-5\\-10&-23&15\\-7&-16&13\end{pmatrix}\begin{pmatrix}-1&0&1\\3&1&-3\\3&2&-4\end{pmatrix}$$
$$=\begin{pmatrix}-10&2&3\\10&4&-10\\-2&1&0\end{pmatrix}.$$

如果利用例 6.4.6 中 $\mathscr{A}$ 在自然基下的矩阵 A 来求矩阵 C, 则计算更为简单, 因为从自然基到 $\gamma_1,\gamma_2,\gamma_3$ 的过渡矩阵就是 $(\gamma_1,\gamma_2,\gamma_3)$.

§6.4.3 线性变换的特征值与特征向量

在第§5.3节, 我们定义了矩阵的特征值与特征向量的概念. 这个概念可以推广到任意线性变换上.

定义6.4.2. 设 V 是数域 F 上 n 维线性空间, $\mathscr{A}$ 为V上的线性变换. 如果存在 $\lambda\in F$ 及非零向量 $\alpha\in V$ 满足 $\mathscr{A}\alpha=\lambda\alpha$, 则称 λ 为线性变换 $\mathscr{A}$ 的一个**特征值**, α 称为属于特征值 λ 的一个**特征向量**. 称所有对应特征值 λ的特征向量全体(包括零向量)

$$V_{\mathscr{A}}(\lambda)=\{\alpha\in V|\mathscr{A}\alpha=\lambda\alpha\},$$

为属于特征值 λ 的**特征子空间**.

下面的命题揭示了线性变换的特征向量与其对应的矩阵的特征向量之间的关系.

命题6.4.6. 设V是数域F上n维线性空间, $\alpha_1,\alpha_2,\cdots,\alpha_n$是$V$的一组基. 线性变换$\mathscr{A}:V\to V$ 在基 $\alpha_1,\alpha_2,\cdots,\alpha_n$ 下的矩阵为 A. 则

(1) 线性变换$\mathscr{A}$与矩阵A有相同的特征值;

(2) 设λ为$\mathscr{A}$的一个特征值, 则

$$V_{\mathscr{A}}(\lambda)=\{(\alpha_1,\alpha_2,\cdots,\alpha_n)\mathbf{x}:\mathbf{x}\in V_A(\lambda)\}.$$

证明. (1) 设λ为$\mathscr{A}$的特征值, α为$\mathscr{A}$的属于λ的特征向量. 令$\alpha=(\alpha_1,\alpha_2,\cdots,\alpha_n)\mathbf{x}$. 由于

$$\mathscr{A}\alpha=\mathscr{A}(\alpha_1,\alpha_2,\cdots,\alpha_n)\mathbf{x}=(\alpha_1,\alpha_2,\cdots,\alpha_n)A\mathbf{x},$$

所以$\mathscr{A}\alpha=\lambda\alpha$蕴含了$A\mathbf{x}=\lambda\mathbf{x}$.

反过来, 假设λ为矩阵A的特征值, $\mathbf{x}$为A的属于λ的特征向量. 令$\alpha=(\alpha_1,\alpha_2,\cdots,\alpha_n)\mathbf{x}$. 类似上面的讨论, 由$A\mathbf{x}=\lambda\mathbf{x}$可以推出$\mathscr{A}\alpha=\lambda\alpha$. 所以线性变换$\mathscr{A}$与矩阵$A$具有完全相同的特征值.

(2) 由(1)的证明过程可知结论正确. □

利用命题 6.4.6, 有关线性变换的特征值与特征向量计算问题可以转化为矩阵的特征值与特征向量计算问题. 这也是将几何问题转化为代数问题的实例之一.

例6.4.10. 设$A=\begin{pmatrix}1&1\\0&2\end{pmatrix}$, $V=F^{2\times 2}$为线性空间, 定义V上的线性变换$\mathscr{A}$:$\mathscr{A}M=AM$, $M\in V$. 求$\mathscr{A}$的特征值和特征向量.

解. 取V的一组基

$$e_1=\begin{pmatrix}1&0\\0&0\end{pmatrix},e_2=\begin{pmatrix}0&1\\0&0\end{pmatrix},e_3=\begin{pmatrix}0&0\\1&0\end{pmatrix},e_4=\begin{pmatrix}0&0\\0&1\end{pmatrix}.$$

类似于例 6.4.7, 线性变换$\mathscr{A}$在基e_1,e_2,e_3,e_4下的矩阵

$$\tilde{A}=\begin{pmatrix}1&0&1&0\\0&1&0&1\\0&0&2&0\\0&0&0&2\end{pmatrix}.$$

$\tilde{A}$的特征多项式$p_{\tilde{A}}(\lambda)=(\lambda-1)^2(\lambda-2)^2$, 所以$\mathscr{A}$的特征值为$\lambda_1=1,\lambda_2=2$, 重数均为 2 .

对于$\lambda=1$, 解$(I-\tilde{A})\mathbf{x}=0$得到两个线性无关解$\mathbf{x}_1=(1,0,0,0)^T$和$\mathbf{x}_2=(0,1,0,0)^T$. 由于$(e_1,e_2,e_3,e_4)\mathbf{x}_1=e_1$, $(e_1,e_2,e_3,e_4)\mathbf{x}_2=e_2$, $\mathscr{A}$的属于特征值 1 的所有特征向量为$c_1e_1+c_2e_2$, $(c_1,c_2$不全为 0).

对于$\lambda=2$, 解$(2I-\tilde{A})\mathbf{x}=0$得到两个线性无关解$\mathbf{x}_3=(1,0,1,0)^T$和$\mathbf{x}_4=(0,1,0,1)^T$. 由于$(e_1,e_2,e_3,e_4)\mathbf{x}_3=e_1+e_3$, $(e_1,e_2,e_3,e_4)\mathbf{x}_4=e_2+e_4$, $\mathscr{A}$的属于特征值 2 的所有特征向量为$c_3(e_1+e_3)+c_4(e_2+e_4)$, $(c_3,c_4$不全为 0). □

设线性空间V上的线性变换$\mathscr{A}$在一组基$\{\alpha_1,\ldots,\alpha_n\}$下的矩阵为$A$. 由于相似矩阵具有相同的特征多项式与特征向量, 因此可以定义线性变换$\mathscr{A}$的**特征多项式**$p_{\mathscr{A}}(\lambda)=p_A(\lambda)$, **行列式**$\det(\mathscr{A})=\det(A)$, **秩** $\mathrm{rank}(\mathscr{A})=\mathrm{rank}(A)$, **迹**$Tr(\mathscr{A})=Tr(A)$.

关于线性变换的下述结论都是显然的, 请读者自己证明.

命题6.4.7. 线性变换$\mathscr{A}$可逆当且仅当$\mathscr{A}$的特征值都不为零.

命题6.4.8. n维线性空间上的线性变换$\mathscr{A}$可对角化当且仅当$\mathscr{A}$有n个线性无关特征向量.

命题6.4.9. 若线性变换$\mathscr{A}$的特征值都不相等, 则$\mathscr{A}$可对角化.

§6.5 若当标准形简介*

在上节中我们给出了矩阵相似于对角阵的条件. 本节中, 我们将简单介绍任意矩阵在相似关系下的标准形–**若当标准形**. 我们将介绍计算若当标准形的方法及其应用.

在例 6.3.2中, 我们证明了矩阵 $\begin{pmatrix} 2 & 1 \\ 0 & 2 \end{pmatrix}$ 不能相似于对角阵.这个矩阵只有一个重数为 2 特征值 2, 它在相似关系下已经是最简单的形式了, 称为一个若当块.一般地, 我们有

定义6.5.1. 设 λ 是任意复数, m 是任意正整数, 形如

$$\begin{pmatrix} \lambda & 1 & 0 & \cdots & 0 \\ & \lambda & 1 & \ddots & \vdots \\ & & \ddots & \ddots & 0 \\ & & & \lambda & 1 \\ & & & & \lambda \end{pmatrix}$$

的 m 阶方阵称为**若当块**, 记作 $J_m(\lambda)$, 其中 m 表示它的阶数, λ 是它的对角线元, 也是它的特征值.因此 $J_m(\lambda)$ 也称为特征值为 λ 的 m 阶若当块.

如果一个方阵是准对角阵, 并且每个对角块都是若当块, 则称之为**若当形矩阵**.注意一个若当形矩阵的某些若当块可能具有相同的特征值.例如矩阵

$$\mathrm{diag}\,(J_4(2), J_3(2), J_1(2), J_3(5))$$

是一个若当形矩阵.

定理6.5.1. 任何一个复方阵 A 都相似于一个若当形矩阵 J, 它的每个若当块的特征值都是 A 的特征值.如果不计若当块的排列顺序, 则 J 是唯一的.

由上面的定理知, 每个复矩阵都相似于唯一的一个若当形矩阵, 若当形矩阵可以作为相似等价类的代表元, 所以若当形矩阵通常也称为**若当标准形**.

定理 6.5.1的证明超过了本书的范围, 我们在此略过.下面我们介绍如何计算一个矩阵的若当标准形.

例6.5.1. 设 5 阶复方阵 A 有一个重数为 5 的特征值 λ, 若 A 的若当标准形 J 有两个若当块, 那么 J 具有什么样的形式?

解. 由于 J 有两个若当块, 如果不计若当块在对角线上的排列顺序, 则有

$$J=J_1=\mathrm{diag}(J_4(\lambda),\lambda)\quad \text{或者}\quad J=J_2=\mathrm{diag}(J_3(\lambda),J_2(\lambda)).$$

注意, J_1 和 J_2 是不相似的.因为如果 J_1 与 J_2 相似, 则 $J_1-\lambda I$ 与 $J_2-\lambda I$ 相似, 从而 $(J_1-\lambda I)^2$ 与 $(J_2-\lambda I)^2$ 相似.但是, $\mathrm{rank}(J_1-\lambda I)^2=2$, 而 $\mathrm{rank}(J_2-\lambda I)^2=1$, 矛盾.

事实上, 如果考虑所有可能的若当块的个数, 方阵 A 的若当标准形共有 7 种形式, 而且这 7 个若当标准形是两两不相似的, 请读者自行验证. □

在上例中, 为说明两个若当形矩阵不相似, 我们需计算 $(J-\lambda I)^k$的秩.而 $J-\lambda I$ 的每个若当块的特征值都是零, 因此弄清楚 $(J_m(0))^k$的秩 十分必要.不难计算得

$$(J_m(0))^2=\begin{pmatrix}0&0&1&\cdots&0\\ &0&0&\ddots&\vdots\\ &&\ddots&\ddots&1\\ &&&0&0\\ &&&&0\end{pmatrix},\quad\cdots,$$

$$(J_m(0))^{m-1}=\begin{pmatrix}0&0&0&\cdots&1\\ &0&0&\ddots&\vdots\\ &&\ddots&\ddots&0\\ &&&0&0\\ &&&&0\end{pmatrix},\quad (J_m(0))^m=0$$

从计算结果可以看出, $k\,(k\le m)$ 每增加 1, $(J_m(0))^k$ 中 1 的列就向右上方移动一位, 矩阵的秩减少一.因此我们得到下面的等式

$$\mathrm{rank}(J_m(0)^k)=\begin{cases}m-k & k\le m\\ 0 & k>m\end{cases}$$

实际上, 我们一开始并不知道若当标准形J, 如何计算$(J-\lambda I)^k$的秩呢? 注意到 A 与 J 相似, 则对任意 $k\ge 1$, $(A-\lambda I)^k$ 与 $(J-\lambda I)^k$ 相似(请读者验证), 从而有$\mathrm{rank}(J-\lambda I)^k=\mathrm{rank}(A-\lambda I)^k$.因此, 我们只需要计算$(A-\lambda I)^k$的秩就可以了.

例6.5.2. 求方阵

$$A=\begin{pmatrix}3&-2&1\\2&-2&2\\3&-6&5\end{pmatrix}$$

的若当标准形.

解. 矩阵 A 的特征多项式为

$$p_A(\lambda) = |\lambda I - A| = (\lambda - 2)^3.$$

由于 A 只有一个 3 重的特征值 $\lambda_1 = 2$, A 的若当标准形只能是以下三者之一

$$J_1 = \begin{pmatrix} 2 & & \\ & 2 & \\ & & 2 \end{pmatrix}, \quad J_2 = \begin{pmatrix} 2 & 1 & \\ & 2 & \\ & & 2 \end{pmatrix}, \quad J_3 = \begin{pmatrix} 2 & 1 & \\ & 2 & 1 \\ & & 2 \end{pmatrix}.$$

但是

$$\operatorname{rank}(J - 2I) = \operatorname{rank}(A - 2I) = \operatorname{rank}\begin{pmatrix} 1 & -2 & 1 \\ 2 & -4 & 2 \\ 3 & -6 & 3 \end{pmatrix} = 1,$$

故 A 的若当标准形为J_2. □

例6.5.3. 求矩阵

$$A = \begin{pmatrix} -2 & 4 & 1 & 0 & 2 \\ -4 & 6 & 1 & -1 & 0 \\ 0 & 0 & 3 & 0 & 0 \\ 0 & 0 & 2 & 3 & 0 \\ 0 & 0 & 1 & 0 & 3 \end{pmatrix}$$

的若当标准形.

解. 矩阵 A 的特征多项式

$$p_A(\lambda) = |\lambda I - A| = (\lambda - 2)^2(\lambda - 3)^3.$$

对于特征值 $\lambda_1 = 2$, 对应的若当块可能为下列二者之一

$$J_{11} = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}, \quad J_{12} = \begin{pmatrix} 2 & 1 \\ 0 & 2 \end{pmatrix}.$$

对于特征值 $\lambda_2 = 3$, 对应的若当块可能为下列三者之一

$$J_{21} = \begin{pmatrix} 3 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 3 \end{pmatrix}, \quad J_{22} = \begin{pmatrix} 3 & 1 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 3 \end{pmatrix}, \quad J_{32} = \begin{pmatrix} 3 & 1 & 0 \\ 0 & 3 & 1 \\ 0 & 0 & 3 \end{pmatrix}.$$

方阵A的若当标准形为$J=\mathrm{diag}(J_{1i},J_{2j})$, $i=1,2$, $j=1,2,3$六种情形之一. 由$\mathrm{rank}(J-2I)=\mathrm{rank}(A-2I)=4$以及$\mathrm{rank}(J-3I)=\mathrm{rank}(A-3I)=3$知, A 的若当标准形为

$$J=\mathrm{diag}(J_{12},J_{22})=\begin{pmatrix}2&1&&&\\0&2&&&\\&&3&1&\\&&0&3&\\&&&&3\end{pmatrix}.$$

□

现在我们考察任意方阵若当标准形的计算问题.设n阶方阵A的特征多项式为

$$p_A(\lambda)=(\lambda-\lambda_1)^{n_1}(\lambda-\lambda_2)^{n_2}\cdots(\lambda-\lambda_s)^{n_s},\qquad \sum_{i=1}^{s}n_i=n,$$

其中$\lambda_1,\ldots,\lambda_s$为互不相等的特征值.记

$$r_k^i=\mathrm{rank}\,(A-\lambda_i I)^k,\quad k\geq 0,\quad \text{其中约定 } r_0=n.$$

记A的若当标准形为J.我们考察特征值为λ_i的若当块中, 各阶若当块的个数, 设p阶若当块个数为δ_p^i. 则$\sum_{p=1}^{n_i}p\delta_p^i=n_i$, 以及有

$$r_1^i=\mathrm{rank}(A-\lambda_i I)=\mathrm{rank}(J-\lambda_i I)=n-n_i+\sum_{p=2}^{n_i}(p-1)\delta_p^i.$$

一般地

$$r_k^i=\mathrm{rank}(A-\lambda_i I)^k=\mathrm{rank}(J-\lambda_i I)^k=n-n_i+\sum_{p=k+1}^{n_i}(p-k)\delta_p^i,\qquad k=1,2,\ldots,n_i.$$

由此

$$d_k^i:=r_{k-1}^i-r_k^i=\sum_{p=k}^{n_i}\delta_p^i.$$

进而J 中特征值为 λ_i 的 k 阶若当块的个数为

$$\delta_k^i=d_k^i-d_{k+1}^i.$$

例6.5.4. 重新考察例 6.5.3. 矩阵 A 有一个二重特征值2和一个三重特征值3.

对于特征值 $\lambda_1=2$, 计算可得

$$r_0^1=5,\ r_1^1=\mathrm{rank}(A-2I)=4,\ r_2^1=\mathrm{rank}(A-2I)^2=3,\ r_3^1=\mathrm{rank}(A-2I)^3=3.$$

因此

$$d_1^1 = r_0^1 - r_1^1 = 1, \quad d_2^1 = r_1^1 - r_2^1 = 1, \quad d_3^1 = r_2^1 - r_3^3 = 0.$$

所以

$$\delta_1^1 = d_1^1 - d_2^1 = 0, \quad \delta_2^1 = d_2^1 - d_3^1 = 1.$$

即 A 的特征值为 2的若当块只有如下一个 2 阶的若当块$\begin{pmatrix} 2 & 1 \\ 0 & 2 \end{pmatrix}$.

对于特征值 $\lambda_2 = 3$, 计算可得

$$r_0 = 5^2, r_1^2 = \operatorname{rank}(A-3I) = 3, r_2^2 = \operatorname{rank}(A-3I)^2 = 2, r_3^2 = \operatorname{rank}(A-2I)^3 = 2.$$

因此

$$d_1^2 = r_0^2 - r_1^2 = 2, \quad d_2^2 = r_1^2 - r_2^2 = 1, \quad d_3^2 = r_2^2 - r_3^2 = 0.$$

所以

$$\delta_1^2 = d_1^2 - d_2^2 = 1, \quad \delta_2^2 = d_2^2 - d_3^2 = 1.$$

即 A 的特征值为 3 的若当块有一个 2 阶的和一个 1 阶的

$$\begin{pmatrix} 3 & 1 \\ 0 & 3 \end{pmatrix} \qquad (3)$$

所以 A 的若当标准形为

$$\begin{pmatrix} 2 & 1 & & & \\ 0 & 2 & & & \\ & & 3 & 1 & \\ & & 0 & 3 & \\ & & & & 3 \end{pmatrix}.$$

例6.5.5. 设

$$A = \begin{pmatrix} 3 & -2 & 1 \\ 2 & -2 & 2 \\ 3 & -6 & 5 \end{pmatrix}$$

求可逆矩阵 T, 使得 $T^{-1}AT$ 为 A 的若当标准形.

解. 在例 6.5.2中, 我们已计算A的若当标准形为

$$J = \begin{pmatrix} 2 & 1 & \\ & 2 & \\ & & 2 \end{pmatrix}.$$

求过渡矩阵 T, 就是解矩阵方程 $AT=TJ$. 如果将 T 中元素看成未知数, 则上述矩阵方阵转化为 9 个未知数的线性方程组.从理论上讲, 任何一个 n 阶方阵与它的若当标准形间的过渡矩阵都可以这样求出:解含 n^2 个未知数的线性方程组.尽管当阶数 n 比较小时, 这个方法是可行的, 但是当 n 很大时, 这样做计算量很大, 不如我们下面将要采用的方法简单.

设 $T=(T_1,T_2,T_3)$, 其中 T_1,T_2,T_3 为列向量.则

$$A\,(T_1,T_2,T_3)=(T_1,T_2,T_3)\begin{pmatrix}2&1&\\&2&\\&&2\end{pmatrix}$$

所以

$$A\,T_1=2T_1,\ A\,T_2=T_1+2T_2,\ A\,T_3=2T_3.$$

即

$$\begin{cases}(A-2I)\,T_1=0\\(A-2I)\,T_2=T_1\\(A-2I)\,T_3=0\end{cases}$$

由前两式得

$$(A-2I)^2T_2=0.$$

由于 $(A-2I)^2=0$, 故 T_2 可取任何一个满足$(A-2I)T_2\neq 0$的向量, 不妨取 $T_2=(1,0,0)^T$. 代入第二个式得 $T_1=(1,2,3)^T$. 再解第三式, 注意 T_3 要与 T_1,T_2 线性无关, 例如取 $T_3=(-1,0,1)^T$. 于是得到过渡矩阵

$$T=\begin{pmatrix}1&1&-1\\2&0&0\\3&0&-1\end{pmatrix}.$$

注意, 我们是先通过第四个式子求出 T_2, 再代入第一个式子解得 T_1. 那么请读者思考是否可以先通过第一个式子解得 T_1, 再代入第二式解 T_2 呢?

□

下面我们来看若当标准形的两个应用.

例6.5.6. 证明:n阶方阵A满足$A^2=A$的充要条件是A相似于准对角阵$\mathrm{diag}(I_r,0)$, 这里$r=\mathrm{rank}(A)$.

证明. 充分性.假设A相似于$\mathrm{diag}(I_r,0)$,即存在可逆方阵T使得$A=T\,\mathrm{diag}(I_r,0)T^{-1}$.则易验证,$A^2=A$.

必要性.设方阵A满足$A^2=A$, A的若当标准形为J, 则$J^2=J$.设J_i是J的特征值为λ_i的若当块, 则$J_i^2=J_i$.由此, $\lambda_i^2=\lambda_i$, 即$\lambda_i=0$或$\lambda_i=1$, 且J_i必为1阶若当块.从而,$J=\mathrm{diag}(I_r,0)$. □

例6.5.7. 设 x,y,z 都是 t 的函数, 求解常微分方程组

$$\begin{cases}\dfrac{\mathrm{d}x}{\mathrm{d}t}=3x-2y+z\\ \dfrac{\mathrm{d}y}{\mathrm{d}t}=2x-2y+2z\\ \dfrac{\mathrm{d}z}{\mathrm{d}t}=3x-6y+5z\end{cases}$$

解. 原方程组可写成

$$\frac{\mathrm{d}}{\mathrm{d}t}\begin{pmatrix}x\\y\\z\end{pmatrix}=\begin{pmatrix}3&-2&1\\2&-2&2\\3&-6&5\end{pmatrix}\begin{pmatrix}x\\y\\z\end{pmatrix},$$

或

$$\frac{\mathrm{d}\mathbf{x}}{\mathrm{d}t}=A\mathbf{x},$$

其中 $\mathbf{x}=(x(t),y(t),z(t))$, 矩阵 A 恰为例 6.5.2中的矩阵. 设 T 为例 6.5.2中的过渡矩阵, 作线性代换 $\mathbf{x}=T\mathbf{x}^*$, 其中$\mathbf{x}^*=(x^*(t),y^*(t),z^*(t))$, 则 $T\dfrac{\mathrm{d}\mathbf{x}^*}{\mathrm{d}t}=AT\mathbf{x}^*$, 即

$$\frac{\mathrm{d}\mathbf{x}^*}{\mathrm{d}t}=T^{-1}AT\mathbf{x}^*.$$

写出相应的方程组, 就是

$$\begin{cases}\dfrac{\mathrm{d}x^*}{\mathrm{d}t}=2x^*+y^*\\ \dfrac{\mathrm{d}y^*}{\mathrm{d}t}=2y^*\\ \dfrac{\mathrm{d}z^*}{\mathrm{d}t}=2z^*\end{cases}$$

解此方程组得

$$z^*=c_1e^{2t},\ y^*=c_2e^{2t},\ x^*=(c_2t+c_3)e^{2t}.$$

再利用 $\mathbf{x}=T\mathbf{x}^*$, 即

$$\begin{pmatrix}x\\y\\z\end{pmatrix}=\begin{pmatrix}1&1&-1\\2&0&0\\3&0&-1\end{pmatrix}\begin{pmatrix}x^*\\y^*\\z^*\end{pmatrix},$$

得到通解是

$$\begin{cases} x(t) = (c_2t + c_2 + c_3 - c_1)e^{2t} \\ y(t) = (2c_2t + 2c_3)e^{2t} \\ z(t) = (3c_2t + 3c_3 + c_1)e^{2t} \end{cases}$$

其中 c_1, c_2, c_3 为任意常数. □

这个例子告诉我们, 利用若当标准形, 可完全解决常系数线性微分方程组的求解问题.

习题五

1. 判断下面所定义的变换$\mathscr{A}$, 哪些是线性的, 哪些不是线性的.

   (1) 在 $\mathbb{R}^2$ 中, $\mathscr{A}(x,y)^T = (x+y, x^2)^T$;

   (2) 在 $\mathbb{R}^3$ 中, $\mathscr{A}(x,y,z)^T = (3x-2y+z, 0, x+2y)^T$;

   (3) 在 $\mathbb{R}^3$ 中, $\mathscr{A}(x,y,z)^T = (x-y, z, x+1)^T$.

2. 给定 $\mathbb{R}^2$上的线性变换$\mathscr{A}(x,y)^T = (2x-y, x+y)^T$. 绘制图形, 观察$\mathscr{A}$将四边形网格变成什么图形? 将单位圆变为什么图形?

3. 在平面直角坐标系中, 求关于直线$\sin(\theta)x - \cos(\theta)y = 0$的对称变换.

4. 求 $\mathbb{R}^2$中如下线性变换$\mathscr{A}$: $\mathscr{A}$沿方向$(2,1)$拉伸2倍, 沿方向$(1,-2)$压缩2倍.

5. 证明: $\mathbb{R}^2$上的可逆线性变换可以分解为关于坐标轴的伸缩、反射及旋转变换的复合.

6. 在三维几何空间的直角坐标系中, 求关于平面$x+2y+3z=0$的对称变换.

7. 在三维几何空间的直角坐标系中, 求关于直线$z=2y=3x$的对称变换.

8. 在三维几何空间的直角坐标系中, 求绕单位向量$\mathbf{e} = (1,-1,1)^T$逆时针旋转30°角的变换.

9. 求 $\mathbb{R}^3$中如下线性变换$\mathscr{A}$: $\mathscr{A}$沿方向$(1,1,0)$拉伸2倍, 沿方向$(1,-1,1)$压缩2倍, 沿方向$(1,-1,-2)$拉伸3 倍.

10. 证明:三维几何空间中的可逆线性变换将平面变为平面, 平行平面变为平行平面.

11. 给定$\mathbb{R}^3$中的线性变换$\mathscr{A}(x,y,z)^T=(2x+y-z,x+2y+z,4x+5y+z)^T$. 求$\mathscr{A}$的像空间与核空间的维数及一组基.

12. 证明: 从高维数组空间到低维数组空间的投影变换不是可逆变换.

13. 求下列矩阵的全部特征值和特征向量

$$(1)\ \begin{pmatrix} 0 & a \\ -a & 0 \end{pmatrix};\quad (2)\ \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix};\quad (3)\ \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix};$$

$$(4)\ \begin{pmatrix} 3 & -2 & 1 \\ 2 & -2 & 2 \\ 3 & -6 & 5 \end{pmatrix};\quad (5)\ \begin{pmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{pmatrix};\quad (6)\ \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & 1 & -1 \\ 1 & 1 & -1 & 1 \end{pmatrix}.$$

14. 给定$\mathbb{R}^2$上的线性变换$\mathscr{A}(x,y)^T=(3x+2y,2x+y)^T$. $\mathscr{A}$将单位圆C映射为椭圆C'. 求椭圆C'的长短半轴方向与长度, 以及椭圆的面积.

15. 设 A 是可逆方阵. 证明: 若 λ 是 A 的一个特征值, 则 λ^{-1} 是 A^{-1} 的特征值, 且对应的特征向量相同.

16. 设 A 为方阵, $f(\lambda)$为多项式. 证明: 如果 λ_0 是 A 的特征值, 则 $f(\lambda_0)$ 是 $f(A)$ 的特征值; 且如果 $\mathbf{x}$ 是属于 λ_0 的特征向量, 则 $\mathbf{x}$ 也是矩阵 $f(A)$ 的属于特征值 $f(\lambda_0)$ 的特征向量.

17. 设三阶方阵A的特征多项式为$p_A(\lambda)=(\lambda-1)^2(\lambda-2)$. 求方阵$3A+2I$的特征值与行列式.

18. (1) 若 $A^2=I$, 证明:A 的特征值只能是 ±1;

(2) 设 n 阶实方阵 A 满足 $A^T=-A$, 证明 A 的特征值为零或纯虚数.

19. 设 λ_1,λ_2 是方阵 A 的两个不同的特征值, $\mathbf{x}_1,\mathbf{x}_2$ 是分别属于 λ_1,λ_2 的特征向量.证明:$\mathbf{x}_1+\mathbf{x}_2$ 不是 A 的特征向量.

20. 设 n 阶方阵 $A=(a_{ij})$ 的全部特征值为 $\lambda_i\,(i=1,2,\cdots,n)$, 证明

$$\sum_{i=1}^{n}\lambda_i^2=\sum_{i,j=1}^{n}a_{ij}a_{ji}.$$

21. 判断以下矩阵A,B是否相似? 请说明理由.

$$(1)\quad A=\begin{pmatrix}1&2&3\\0&1&0\\0&0&1\end{pmatrix},\qquad B=\begin{pmatrix}1&2&0\\0&1&3\\0&0&1\end{pmatrix};$$

$$(2)\quad A=\begin{pmatrix}1&2&0&0\\0&1&3&0\\0&0&1&0\\0&0&0&1\end{pmatrix},\qquad B=\begin{pmatrix}1&2&0&0\\0&1&0&0\\0&0&1&3\\0&0&0&1\end{pmatrix}.$$

22. 求三阶可逆方阵P使得$P\begin{pmatrix}1&0&0\\0&2&0\\0&0&3\end{pmatrix}P^{-1}=\begin{pmatrix}3&0&0\\0&1&0\\0&0&2\end{pmatrix}$.

23. 设A,B为同阶方阵, 且A可逆. 证明; A与B相似.

24. 设A与B相似, C与D相似. 证明: $\mathrm{diag}(A,C)$与$\mathrm{diag}(B,D)$相似. 反之是否成立?

25. 设方阵 A 与 B 相似, 证明:

(1) 对每个正整数 k, A^k 相似于 B^k;

(2) 对每个多项式 f, $f(A)$ 相似于$f(B)$.

26. 设矩阵 $A=\begin{pmatrix}0&0&1\\x&1&y\\1&0&0\end{pmatrix}$ 有三个线性无关的特征向量, 求 x 和 y 应满足的条件.

27. 设矩阵 A 与 B 相似, 其中

$$A=\begin{pmatrix}-2&0&0\\2&x&2\\3&1&1\end{pmatrix},\quad B=\begin{pmatrix}-1&0&0\\0&2&0\\0&0&y\end{pmatrix}.$$

(1) 求 x 和 y 的值;

(2) 求可逆矩阵 T, 使得$T^{-1}AT=B$.

28. 设 n 阶方阵 $A\neq 0$, 满足 $A^m=0$, 其中 $m\geq 2$ 为正整数.

(1) 求 A 的特征值;

(2) 证明:A 不能相似于对角阵;

(3) 证明:$|I+A|=1$.

29. 判断下列矩阵 A 是否可对角化?若可以, 试求变换矩阵 T, 使得 $T^{-1}AT$ 为对角阵.

$$(1)\quad A=\begin{pmatrix}-2&0&0\\3&1&1\\2&2&0\end{pmatrix};\qquad (2)\quad A=\begin{pmatrix}1&-1&1\\2&4&-2\\-3&-3&5\end{pmatrix};$$

$$(3)\quad A=\begin{pmatrix}3&0&1\\4&-2&-8\\-4&0&-1\end{pmatrix};\qquad (4)\quad A=\begin{pmatrix}2&1&1\\1&2&1\\1&1&2\end{pmatrix}.$$

30. 设 A,B 均为 n 阶方阵, A 有 n 个互异的特征值且 $AB=BA$.证明 B 相似于对角阵.

31. 设 n 阶方阵 A 满足 $A^2=I$, 证明 A 相似于 $\begin{pmatrix}I_r&0\\0&-I_{n-r}\end{pmatrix}$, 其中 $0\le r\le n$.

32. 设 A 为 3 阶实方阵, 若 A 不实相似于上三角阵, 问 A 是否一定复相似于对角阵?

33. 设n阶复方阵A的全部特征值为$\lambda_1,\lambda_2,\dots,\lambda_n$, $f(\lambda)$为多项式. 证明: $f(A)$的全部特征值为$f(\lambda_1),f(\lambda_2),\dots,f(\lambda_n)$, 由此$\det(f(A))=f(\lambda_1)f(\lambda_2)\cdots f(l_n)$.

34. 设三阶方阵A的特征多项式为$p_A(\lambda)=(\lambda-1)^2(\lambda-2)$. 求多项式矩阵$A^3-2A+3I$的特征值与行列式.

35. 判断下面所定义的变换, 哪些是线性的, 哪些不是线性的.

(3) 取定 $A,B\in M_n(F)$, 对每个 $\mathbf{x}\in M_n(F)$, $\mathscr{A}(\mathbf{x})=A\mathbf{x}-\mathbf{x}B$;

(4) 在线性空间 V 中, $\mathscr{A}(x)=\alpha$, 其中 α 为 V 中的一个固定的向量.

36. 求下列线性变换在所指定的基下的矩阵

(1) 在 $F_n[x]$ 中, $\mathscr{A}(P(x))=P'(x)$, 在基 $e_0=1,e_1=x,\cdots,e_{n-1}=\dfrac{x^{n-1}}{(n-1)!}$ 下;

(2) 以四个线性无关的函数

$$\alpha_1=e^{ax}\cos bx\qquad \alpha_2=e^{ax}\sin bx$$
$$\alpha_3=xe^{ax}\cos bx\qquad \alpha_4=xe^{ax}\sin bx$$

为基的四维空间中, 线性变换为微分变换;

(3) 给定 2 阶实方阵 A, 求 2 阶实方阵构成的线性空间上的线性变换 $\mathscr{A}(\mathbf{x}) = A\mathbf{x} - \mathbf{x}A$ 在基

$$e_1 = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, e_2 = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, e_3 = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, e_4 = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}.$$

下的矩阵.

37. 在 $\mathbb{R}^3$ 中定义线性变换

$$\mathscr{A}(x,y,z)^T = (x+2y, x-3z, 2y-z)^T.$$

求 $\mathscr{A}$ 在基 $\mathbf{e}_1 = (1,0,0)^T, \mathbf{e}_2 = (0,1,0)^T, \mathbf{e}_3 = (0,0,1)^T$ 下的矩阵.

38. 设 $\mathbb{R}^3$ 中的线性变换 $\mathscr{A}$ 将

$$\alpha_1 = (0,0,1)^T, \quad \alpha_2 = (0,1,1)^T, \quad \alpha_3 = (1,1,1)^T$$

变换到

$$\beta_1 = (2,3,5)^T, \quad \beta_2 = (1,0,0)^T, \quad \beta_3 = (0,1,-1)^T.$$

求 $\mathscr{A}$ 在自然基和 $\alpha_1, \alpha_2, \alpha_3$ 下的矩阵.

39. 设线性变换 $\mathscr{A}$ 在基 $\alpha_1 = (1,-1), \alpha_2 = (1,1)$ 下的矩阵为 $\begin{pmatrix} 2 & 3 \\ 0 & 1 \end{pmatrix}$, 求 $\mathscr{A}$ 在基 $\beta_1 = (2,0), \beta_2 = (-1,1)$ 下的矩阵.

40. 设 $\mathscr{A}$ 在基 $\alpha_1, \alpha_2, \alpha_3$ 下的矩阵为

$$A = \begin{pmatrix} 1 & 2 & 3 \\ -1 & 0 & 3 \\ 2 & 1 & 5 \end{pmatrix},$$

求 $\mathscr{A}$ 在基 $\beta_1, \beta_2, \beta_3$ 下的矩阵:

(1) $\beta_1 = \alpha_3, \beta_2 = \alpha_1, \beta_3 = \alpha_2$;

(2) $\beta_1 = \alpha_1, \beta_2 = \alpha_2 + \alpha_3, \beta_3 = \alpha_2 - \alpha_3$.

41. 在 $\mathbb{R}^3$ 中给定两组基:

$$\alpha_1 = (1,0,1)^T, \quad \alpha_2 = (2,1,0)^T, \quad \alpha_3 = (1,1,1)^T;$$

$$\beta_1 = (2,3,1)^T, \quad \beta_2 = (7,9,5)^T, \quad \beta_3 = (3,4,3)^T.$$

定义线性变换 $\mathscr{A}(\alpha_i) = \beta_i \, (i = 1,2,3)$.

(1) 求 $\mathscr{A}$ 在基 $\alpha_1,\alpha_2,\alpha_3$ 下的矩阵；

(2) 求 $\mathscr{A}$ 在基 β_1,β_2,β_3 下的矩阵.

42. 设 V 为 n 维线性空间, $\mathscr{A}:V\to V$ 为线性变换.若存在 $\alpha\in V$, 使得 $\mathscr{A}^{n-1}\alpha\neq 0$, 但是 $\mathscr{A}^n(\alpha)=0$, 证明: $\mathscr{A}$ 在某组基下的矩阵为

$$\begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ & 0 & 1 & \ddots & \vdots \\ & & \ddots & \ddots & 0 \\ & & & 0 & 1 \\ & & & & 0 \end{pmatrix}$$

43. 设 $\alpha_1,\alpha_2,\cdots,\alpha_n$ 为线性空间 V 的一组基, 证明:对于任意 $\beta_1,\beta_2,\cdots,\beta_n\in V$, 存在线性变换 $\mathscr{A}$, 使得 $\mathscr{A}(\alpha_i)=\beta_i\,(i=1,2,\cdots,n)$.

44. 设 V 为次数不超过 2 的多项式构成的线性空间, 线性变换 $\mathscr{A}:V\to V$ 满足

$$\mathscr{A}(1)=x^2+x+3,\quad \mathscr{A}(x)=2x+1,\quad \mathscr{A}(x^2)=2x^2+3.$$

求 $\mathscr{A}$ 的特征值和特征向量.

45. 设$V=F^{n\times n}$.

(1) 证明: V的变换$\mathscr{A}:\mathbf{x}\to\mathbf{x}^T$是线性变换;

(2) 求$\mathscr{A}$的特征值与特征向量. $\mathscr{A}$是否可对角化?

46* 求下列矩阵的若当标准形

(1) $\begin{pmatrix} 2 & -2 & 3 \\ 1 & 1 & 1 \\ 1 & 3 & -1 \end{pmatrix}$ (2) $\begin{pmatrix} 4 & -9 & -4 \\ 6 & -12 & -5 \\ -7 & 13 & 5 \end{pmatrix}$

(3) $\begin{pmatrix} 1 & 2 & 3 & 4 \\ 0 & 1 & 2 & 3 \\ 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 1 \end{pmatrix}$ (4) $\begin{pmatrix} 0 & 1 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 1 & \cdots & 0 & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \cdots & 0 & 1 \\ 1 & 0 & 0 & \cdots & 0 & 0 \end{pmatrix}_{n\times n}$

47* 设 $A=\begin{pmatrix} 2 & -2 & 3 \\ 10 & -4 & 5 \\ 5 & -4 & 6 \end{pmatrix}$, 求方阵 T, 使 $T^{-1}AT$ 为若当标准形.

48* 设方阵A满足$A^2 = A$.用矩阵的若当标准型证明:$Tr(A) = \mathrm{rank}(A)$.

49* 设 $x = x(t), y = y(t), z = z(t)$, 求解常微分方程组

$$\begin{cases} \dfrac{\mathrm{d}x}{\mathrm{d}t} = 2x - 2y + 3z \\ \dfrac{\mathrm{d}y}{\mathrm{d}t} = 10x - 4y + 5z \\ \dfrac{\mathrm{d}z}{\mathrm{d}t} = 5x - 4y + 6z \end{cases}$$

第七章　内积空间及变换

在第四章我们引进了线性空间的概念. 作为线性空间的一个具体模型–三维几何空间$\mathbb{R}^3$, 其中向量的长度、夹角等几何性质在线性空间中没有得到反映. 而这种反映向量之间的度量等几何性质的空间在物理和几何等许多问题中具有十分重要的作用. 因此, 本章的主要目的是在线性空间中引入内积形成具有度量的线性空间–内积空间. 我们主要考虑两个数域上的内积空间: 在实数域上的内积线性空间–**欧几里得(Euclid)空间**及复数域上的内积空间–**酉空间**. 我们先介绍欧几里得空间(简称欧氏空间), 特别地, 我们从数组空间上的内积空间开始介绍.

§7.1　数组内积空间的定义与性质

从解析几何中我们知道, 对于$\mathbb{R}^3$中任意两个向量$\mathbf{a}$和$\mathbf{b}$, 设$|\mathbf{a}|$和$|\mathbf{b}|$分别是它们的长度, θ是它们之间的夹角, 则其内积(有时又称为数量积或点乘)定义为

$$(\mathbf{a},\mathbf{b}) = |\mathbf{a}||\mathbf{b}|\cos\theta. \tag{7.1}$$

反之, 向量$\mathbf{a}$的长度以及向量$\mathbf{a}$和$\mathbf{b}$之间的夹角也可由内积表示

$$|\mathbf{a}| = \sqrt{(\mathbf{a},\mathbf{a})}, \tag{7.2}$$

$$\theta = \arccos\frac{(\mathbf{a},\mathbf{b})}{\sqrt{(\mathbf{a},\mathbf{a})(\mathbf{b},\mathbf{b})}}. \tag{7.3}$$

因此, 在一般线性空间中, 如果我们定义好了内积, 则向量之间的长度和夹角也就可以由内积来定义, 从而可以引进度量. 下面, 我们先在实数域$\mathbb{R}$上的n维数组空间$\mathbb{R}^n$中引入内积.

定义7.1.1. 设$\mathbb{R}^n$是实数域$\mathbb{R}$上的n维数组空间. 对任意(列)向量$\mathbf{a},\mathbf{b}\in\mathbb{R}^n$, 定义它们的内积为

$$(\mathbf{a},\mathbf{b}) := \mathbf{a}^T G\mathbf{b} \tag{7.4}$$

其中$G\in\mathbb{R}^{n\times n}$称为**度量矩阵**, 它是满足对任意$\mathbf{0}\neq\mathbf{a}\in\mathbb{R}^n$, $\mathbf{a}^T G\mathbf{a} > 0$的$n$阶实对称方阵, 称为**实对称正定方阵**. 定义了内积 (7.4) 的数组空间$\mathbb{R}^n$称为**欧几里得(Euclid)空间**, 简称**欧氏空间**.

我们看几个例子.

例7.1.1. 取度量矩阵G为n阶单位阵. 则对任意$\mathbf{0}\neq\mathbf{a}\in\mathbb{R}^n$, $\mathbf{a}^T G\mathbf{a} = \mathbf{a}^T\mathbf{a} > 0$, 故单位阵是实正定方阵. 这时内积定义为

$$(\mathbf{a},\mathbf{b}) = a_1b_1 + a_2b_2 + \ldots + a_nb_n = \mathbf{a}^T\mathbf{b}, \tag{7.5}$$

这里$\mathbf{a}=(a_1,a_2,\ldots,a_n)^T\in\mathbb{R}^n$, $\mathbf{b}=(b_1,b_2,\ldots,b_n)^T\in\mathbb{R}^n$. 我们称该内积为**标准内积**.

例7.1.2. 令$G=\mathrm{diag}(\lambda_1,\lambda_2,\ldots,\lambda_n)$, 这里$\lambda_i>0$, $i=1,2,\ldots,n$. 对于$\mathbf{a}=(a_1,a_2,\ldots,a_n)^T\neq\mathbf{0}$,

$$\mathbf{a}^TG\mathbf{a}=\lambda_1a_1^2+\lambda_2a_2^2+\ldots+\lambda_na_n^2>0.$$

因此, G是对称正定方阵. 这时内积为

$$(\mathbf{a},\mathbf{b})=\lambda_1a_1b_1+\lambda_2a_2b_2+\ldots+\lambda_na_nb_n.$$

例7.1.3. 在$\mathbb{R}^2$中, 取$G=\begin{pmatrix}1&1\\1&2\end{pmatrix}$. 由于对于$\mathbf{a}=(a_1,a_2)^T\neq\mathbf{0}$,

$$\mathbf{a}^TG\mathbf{a}=a_1^2+2a_1a_2+2a_2^2>0,$$

因此G是对称正定方阵. 这时内积为

$$(\mathbf{a},\mathbf{b})=a_1b_1+a_1b_2+a_2b_1+2a_2b_2.$$

本章中, 除非特别申明, 欧氏空间$\mathbb{R}^n$中的内积均指标准内积.

同三维几何空间向量的内积一样, 欧氏空间的内积满足以下性质:

(1) **对称性**: 即对任意两个向量$\mathbf{a},\mathbf{b}\in\mathbb{R}^n$, 有

$$(\mathbf{a},\mathbf{b})=(\mathbf{b},\mathbf{a}) \tag{7.6}$$

(2) **线性性**: 即对任意一个实数λ和任意三个向量$\mathbf{a},\mathbf{b},\mathbf{c}\in\mathbb{R}^n$, 有

$$(\lambda\mathbf{a},\mathbf{b})=\lambda(\mathbf{a},\mathbf{b}) \tag{7.7}$$

$$(\mathbf{a}+\mathbf{b},\mathbf{c})=(\mathbf{a},\mathbf{c})+(\mathbf{b},\mathbf{c}) \tag{7.8}$$

(3) **正定性**: 即对于任意一个向量$\mathbf{a}\in\mathbb{R}^n$, 都有$(\mathbf{a},\mathbf{a})\geq 0$, 且等号成立当且仅当$\mathbf{a}=\mathbf{0}$.

由上述性质, 很易容得到: 对任意实数λ_i,μ_j和任意向量$\mathbf{a}_i,\mathbf{b}_j\in\mathbb{R}^n$, 有

$$\left(\sum_{i=1}^{l}\lambda_i\mathbf{a}_i,\sum_{j=1}^{m}\mu_j\mathbf{b}_j\right)=\sum_{i=1}^{l}\sum_{j=1}^{m}\lambda_i\mu_j(\mathbf{a}_i,\mathbf{b}_j). \tag{7.9}$$

特别地

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

读者可以自己验证上述性质.

定义了内积, 我们就可以定义向量长度.

定义7.1.2. 设$\mathbf{a}\in\mathbb{R}^n$, 称
$$|\mathbf{a}|=\sqrt{(\mathbf{a},\mathbf{a})} \tag{7.10}$$
为向量$\mathbf{a}$的**长度**或**模**.

显然, $|\mathbf{a}|=0$的充分必要条件是$\mathbf{a}$为零向量. 当$|\mathbf{a}|=1$时, 称$\mathbf{a}$为**单位向量**. 对于任意一个非零向量$\mathbf{a}$, 向量$\frac{1}{|\mathbf{a}|}\mathbf{a}$为单位向量. 因此通过这样的方式可以把任何一个非零向量“压缩”(或“放大”)为一个单位向量, 这个过程称为向量的**单位化**.

为了定义向量之间的夹角, 我们先引入下面的

定理7.1.1. *[Cauchy-Schwarz不等式]* **对两个任意向量$\mathbf{a},\mathbf{b}\in\mathbb{R}^n$, 均有**
$$|(\mathbf{a},\mathbf{b})|\le|\mathbf{a}|\cdot|\mathbf{b}|. \tag{7.11}$$
等号成立当且仅当$\mathbf{a}$与$\mathbf{b}$线性相关.

证明. 如果$\mathbf{a}$或者$\mathbf{b}$有一个为零向量, 则结论显然成立. 否则对任意实数λ, 有
$$0\le(\lambda\mathbf{a}+\mathbf{b},\lambda\mathbf{a}+\mathbf{b})=(\mathbf{a},\mathbf{a})\lambda^2+2(\mathbf{a},\mathbf{b})\lambda+(\mathbf{b},\mathbf{b}).$$
上式右端是λ的二次多项式, 其值总是非负的, 故其判别式满足
$$4(\mathbf{a},\mathbf{b})^2-4(\mathbf{a},\mathbf{a})(\mathbf{b},\mathbf{b})\le 0.$$
由此知不等式 (7.11)成立.

当不等式 (7.11)中的等号成立时, 令$\lambda=-(\mathbf{a},\mathbf{b})/(\mathbf{a},\mathbf{a})$, 则
$$(\lambda\mathbf{a}+\mathbf{b},\lambda\mathbf{a}+\mathbf{b})=(\mathbf{a},\mathbf{a})\lambda^2+2(\mathbf{a},\mathbf{b})\lambda+(\mathbf{b},\mathbf{b})=0.$$
从而$\lambda\mathbf{a}+\mathbf{b}=0$, 即$\mathbf{a}$与$\mathbf{b}$线性相关. 反之, 从$\mathbf{a}$与$\mathbf{b}$线性相关, 易知等式成立. □

将Cauchy-Schwarz不等式应用于标准内积 (7.5)即得著名的Cauchy不等式:
$$(a_1b_1+a_2b_2+\ldots+a_nb_n)^2\le(a_1^2+a_2^2+\ldots+a_n^2)(b_1^2+b_2^2+\ldots+b_n^2).$$

由Cauchy-Schwarz不等式, 我们可进一步证明下列关于长度的三角不等式
$$|\mathbf{a}+\mathbf{b}|\le|\mathbf{a}|+|\mathbf{b}|. \tag{7.12}$$
实际上,
$$\begin{aligned}|\mathbf{a}+\mathbf{b}|^2&=(\mathbf{a}+\mathbf{b},\mathbf{a}+\mathbf{b})=(\mathbf{a},\mathbf{a})+2(\mathbf{a},\mathbf{b})+(\mathbf{b},\mathbf{b})\\&\le|\mathbf{a}|^2+2|\mathbf{a}||\mathbf{b}|+|\mathbf{b}|^2=(|\mathbf{a}|+|\mathbf{b}|)^2\end{aligned}$$

因此,

$$|\mathbf{a}+\mathbf{b}| \leq |\mathbf{a}|+|\mathbf{b}|.$$

利用长度可给出欧氏空间中任意两个向量$\mathbf{a}$和$\mathbf{b}$之间的“距离”:

$$d(\mathbf{a},\mathbf{b}) = |\mathbf{a}-\mathbf{b}|.$$

读者不难验证, 向量之间的距离满足

- **对称性**: $d(\mathbf{a},\mathbf{b}) = d(\mathbf{b},\mathbf{a})$;
- **正定性**: $d(\mathbf{a},\mathbf{b}) \geq 0$, 等号成立当且仅当$\mathbf{a}=\mathbf{b}$;
- **三角不等式**: $d(\mathbf{a},\mathbf{b}) \leq d(\mathbf{a},\mathbf{c})+d(\mathbf{c},\mathbf{b})$.

由Cauchy-Schwarz不等式还可以得到

$$-1 \leq \frac{(\mathbf{a},\mathbf{b})}{|\mathbf{a}||\mathbf{b}|} \leq 1.$$

由此可以定义两向量的夹角.

定义7.1.3. 对于$\mathbb{R}^n$中两个非零向量$\mathbf{a}$和$\mathbf{b}$, 定义它们之间的夹角为

$$\theta = \arccos\frac{(\mathbf{a},\mathbf{b})}{|\mathbf{a}||\mathbf{b}|}. \tag{7.13}$$

特别地, 当$(\mathbf{a},\mathbf{b})=0$时, 称向量$\mathbf{a}$和$\mathbf{b}$**正交**或**垂直**, 记作$\mathbf{a}\perp\mathbf{b}$.

对于标准内积, 向量$\mathbf{a}$的长度为

$$|\mathbf{a}| = \sqrt{a_1^2+a_2^2+\cdots+a_n^2}, \tag{7.14}$$

两个向量$\mathbf{a}$与$\mathbf{b}$的夹角为

$$\theta = \arccos\frac{a_1b_1+a_2b_2+\cdots+a_nb_n}{\sqrt{a_1^2+a_2^2+\cdots+a_n^2}\sqrt{b_1^2+b_2^2+\cdots+b_n^2}}. \tag{7.15}$$

由于通过内积可以引进向量的长度及向量之间的夹角, 因此, 欧氏空间不但具有线性运算结构, 同时也具有了几何结构.

对于相互垂直的向量, 我们有下面的勾股定理

命题7.1.2. 设$\mathbf{a},\mathbf{b}\in\mathbb{R}^n$, 则$\mathbf{a}\perp\mathbf{b}$当且仅当

$$|\mathbf{a}+\mathbf{b}|^2 = |\mathbf{a}|^2+|\mathbf{b}|^2.$$

证明. 由于

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

命题结论显然. □

最后, 我们提一下欧氏空间的子空间. 设$W\subset\mathbb{R}^n$是$\mathbb{R}^n$的线性子空间. 作为欧氏空间$\mathbb{R}^n$, 其内积自然也作用在W上. 因此, 赋予了内积的线性子空间W也成为一个欧氏空间, 称为欧氏空间$\mathbb{R}^n$的**子空间**.

§7.2 标准正交基

正如三维几何空间中有正交标架, 在欧氏空间中也可以定义类似的正交标架–正交基.

定义7.2.1. 在n维欧氏空间$\mathbb{R}^n$中, 一组两两正交的非零向量称为**正交向量组**. 由正交向量组构成的基称为**正交基**. 由单位向量组构成的正交基称为**标准正交基**.

由定义, 向量组$\mathbf{a}_1,\dots,\mathbf{a}_n$构成$n$维欧氏空间$\mathbb{R}^n$的一组标准正交基当且仅当

$$(\mathbf{a}_i,\mathbf{a}_j)=\delta_{ij},\quad i,j=1,2,\dots,n. \tag{7.16}$$

显然, $\mathbb{R}^n$中的单位坐标向量组$\mathbf{e}_i$, $i=1,2,\dots,n$在标准内积之下是标准正交基. 标准正交基是三维几何空间中单位直角坐标系的推广.

命题7.2.1. 设$\mathbf{a}_1,\mathbf{a}_2,\cdots,\mathbf{a}_r$为欧氏空间$\mathbb{R}^n$中的正交向量组, 则$\mathbf{a}_1,\mathbf{a}_2,\cdots,\mathbf{a}_r$线性无关.

证明. 设存在实数$\lambda_1,\lambda_2,\dots,\lambda_r$满足

$$\lambda_1\mathbf{a}_1+\lambda_2\mathbf{a}_2+\cdots+\lambda_r\mathbf{a}_r=\mathbf{0},$$

对任意$1\le i\le r$, 两边分别与$\mathbf{a}_i$作内积得

$$\lambda_i|\mathbf{a}_i|^2=0.$$

因为$\mathbf{a}_i$是非零向量, 所以$\lambda_i=0$, 即向量组$\mathbf{a}_1,\mathbf{a}_2,\dots,\mathbf{a}_r$线性无关. □

上述命题表明, $\mathbb{R}^n$中的n个正交向量组构成$\mathbb{R}^n$的一组基.

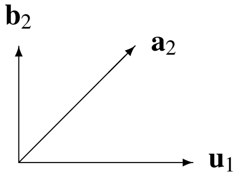
n维欧氏空间的标准正交基总是存在的. 实际上, 从任意一组基出发, 我们可以构造出一组标准正交基, 其过程称为**Schmidt正交化**.

定理7.2.2. *[Schmidt正交化]* **从n维欧氏空间$\mathbb{R}^n$的任意一组基出发, 可以构造一组标准正交基.**

证明. 设$\mathbf{a}_1,\mathbf{a}_2,\ldots,\mathbf{a}_n$是欧氏空间$\mathbb{R}^n$的一组基. 首先将$\mathbf{a}_1$归一化, 令

$$\mathbf{u}_1=\frac{\mathbf{a}_1}{|\mathbf{a}_1|},$$

则有$(\mathbf{u}_1,\mathbf{u}_1)=1$且$\langle\mathbf{a}_1\rangle=\langle\mathbf{u}_1\rangle$. 然后在$\mathbf{u}_1$与$\mathbf{a}_2$所张成的子空间中找出一个与$\mathbf{u}_1$ 正交的向量.



上图启发我们这样的向量为$\mathbf{a}_2$减去$\mathbf{a}_2$在$\mathbf{u}_1$上的投影, 其代数表示为

$$\mathbf{b}_2=\mathbf{a}_2-(\mathbf{a}_2,\mathbf{u}_1)\mathbf{u}_1$$

不难验证, $(\mathbf{b}_2,\mathbf{u}_1)=0$且$\mathbf{b}_2\neq 0$. 将$\mathbf{b}_2$归一化得

$$\mathbf{u}_2=\frac{\mathbf{b}_2}{|\mathbf{b}_2|}.$$

显然, $(\mathbf{u}_1,\mathbf{u}_2)=0$, $(\mathbf{u}_2,\mathbf{u}_2)=1$ 且$\langle\mathbf{a}_1,\mathbf{a}_2\rangle=\langle\mathbf{u}_1,\mathbf{u}_2\rangle$.

继续上面的过程, 假设从$\mathbf{a}_1,\mathbf{a}_2,\ldots,\mathbf{a}_{k-1}$已经构造出了单位化的正交向量组$\mathbf{u}_1,\mathbf{u}_2,\ldots,\mathbf{u}_{k-1}$, 且$\langle\mathbf{u}_1,\mathbf{u}_2,\ldots,\mathbf{u}_{k-1}\rangle=\langle\mathbf{a}_1,\mathbf{a}_2,\ldots,\mathbf{a}_{k-1}\rangle$. 我们再从$\mathbf{a}_k$和$\mathbf{u}_1,\mathbf{u}_2,\ldots,\mathbf{u}_{k-1}$所张成的子空间中挑选一个与$\mathbf{u}_1,\mathbf{u}_2,\ldots,\mathbf{u}_{k-1}$ 都正交的向量$\mathbf{b}_k$如下

$$\mathbf{b}_k=\mathbf{a}_k-\sum_{i=1}^{k-1}(\mathbf{a}_k,\mathbf{u}_i)\mathbf{u}_i.$$

显然它一定非零, 否则就有

$$\mathbf{a}_k=\sum_{i=1}^{k-1}(\mathbf{a}_k,\mathbf{u}_i)\mathbf{u}_i,$$

即$\mathbf{a}_k$是$\mathbf{u}_1,\mathbf{u}_2,\ldots,\mathbf{u}_{k-1}$的线性组合, 从而也是$\mathbf{a}_1,\mathbf{a}_2,\ldots,\mathbf{a}_{k-1}$的线性组合, 这与$\mathbf{a}_1,\mathbf{a}_2,\ldots,\mathbf{a}_k$线性无关矛盾.

另一方面

$$(\mathbf{b}_k,\mathbf{u}_j)=(\mathbf{a}_k,\mathbf{u}_j)-\sum_{i=1}^{k-1}(\mathbf{a}_k,\mathbf{u}_i)(\mathbf{u}_i,\mathbf{u}_j)=0,\quad j=1,2,\ldots,k-1,$$

故$\mathbf{b}_k$与所有$\mathbf{u}_1,\mathbf{u}_2,\dots,\mathbf{u}_{k-1}$正交. 将$\mathbf{b}_k$归一化得到一个与所有$\mathbf{u}_1,\mathbf{u}_2,\dots,\mathbf{u}_{k-1}$都正交的单位向量

$$\mathbf{u}_k = \frac{\mathbf{b}_k}{|\mathbf{b}_k|}.$$

容易验证,

$$\langle \mathbf{u}_1,\mathbf{u}_2,\dots,\mathbf{u}_k\rangle = \langle \mathbf{a}_1,\mathbf{a}_2,\dots,\mathbf{a}_k\rangle.$$

将上述过程一直进行到$k=n$, 这样就得到了一组标准正交基$\mathbf{u}_1,\mathbf{u}_2,\dots,\mathbf{u}_n$. □

下面我们举一例说明上述正交化过程.

例7.2.1. 把下列向量按标准内积化为标准正交基

$$\alpha_1 = (1,1,0,0)^T, \alpha_2 = (1,0,1,0)^T, \alpha_3 = (-1,0,0,1)^T, \alpha_4 = (1,-1,-1,1)^T.$$

解. 先把$\mathbf{a}_1$归一化得

$$\mathbf{u}_1 = \frac{1}{\sqrt{2}}(1,1,0,0)^T.$$

令

$$\mathbf{b}_2 = \mathbf{a}_2 - (\mathbf{a}_2,\mathbf{u}_1)\mathbf{u}_1 = \frac{1}{2}(1,-1,2,0)^T,$$

再归一化得

$$\mathbf{u}_2 = \frac{1}{\sqrt{6}}(1,-1,2,0)^T.$$

继续正交化$\mathbf{a}_3,\mathbf{a}_4$得

$$\mathbf{b}_3 = \mathbf{a}_3 - (\mathbf{a}_3,\mathbf{u}_1)\mathbf{u}_1 - (\mathbf{a}_3,\mathbf{u}_2)\mathbf{u}_2 = \frac{1}{3}(-1,1,1,3)^T, \quad \mathbf{u}_3 = \frac{1}{\sqrt{12}}(-1,1,1,3)^T,$$

$$\mathbf{b}_4 = \mathbf{a}_4 - \sum_{i=1}^{3}(\mathbf{a}_4,\mathbf{u}_i)\mathbf{u}_i = (1,-1,-1,1)^T, \quad \mathbf{u}_4 = \frac{1}{2}(1,-1,-1,1)^T.$$

则$\mathbf{u}_1,\mathbf{u}_2,\mathbf{u}_3,\mathbf{u}_4$即为一组标准正交基. □

在Schmidt正交化过程中, 注意到

$$\mathbf{a}_k = \sum_{i=1}^{k}(\mathbf{a}_k,\mathbf{u}_i)\mathbf{u}_i, \quad k=1,2,\dots,n.$$

因此

$$(\mathbf{a}_1,\mathbf{a}_2,\dots,\mathbf{a}_n) = (\mathbf{u}_1,\mathbf{u}_2,\dots,\mathbf{u}_n)\begin{pmatrix} (\mathbf{a}_1,\mathbf{u}_1) & (\mathbf{a}_2,\mathbf{u}_1) & \cdots & (\mathbf{a}_n,\mathbf{u}_1) \\ 0 & (\mathbf{a}_2,\mathbf{u}_2) & \cdots & (\mathbf{a}_n,\mathbf{u}_2) \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & (\mathbf{a}_n,\mathbf{u}_n) \end{pmatrix}. \tag{7.17}$$

其中假设$\mathbf{a}_i, \mathbf{u}_i$均为列向量形式, 则上式即为

$$A = QR, \tag{7.18}$$

其中A是一个n阶可逆实方阵, Q是一个任意两列都正交的单位向量构成的方阵, R是一个上三角阵, 其对角元为正. 称上式为矩阵A的**QR分解**.

容易验证

$$Q^TQ = \begin{pmatrix} \mathbf{u}_1^T \\ \vdots \\ \mathbf{u}_n^T \end{pmatrix} (\mathbf{u}_1, \dots, \mathbf{u}_n) = \begin{pmatrix} (\mathbf{u}_1, \mathbf{u}_1) & \cdots & (\mathbf{u}_1, \mathbf{u}_n) \\ \vdots & \vdots & \vdots \\ (\mathbf{u}_n, \mathbf{u}_1) & \cdots & (\mathbf{u}_n, \mathbf{u}_n) \end{pmatrix} = I. \tag{7.19}$$

因此, $Q^TQ = I = QQ^T$.

定义7.2.2. 称满足

$$QQ^T = Q^TQ = I \tag{7.20}$$

的实方阵Q为正交方阵.

由前面的讨论, 下面的结论显然.

定理7.2.3. *n阶实方阵Q为正交矩阵当且仅当Q的行(或列)向量构成$\mathbb{R}^n$的一组标准正交基.*

证明. 设Q的列为$\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_n$. 由 (7.19), $Q^TQ = I$当且仅当$(\mathbf{u}_i, \mathbf{u}_j) = \delta_{ij}, i = 1, 2, \dots, n$, 当且仅当 $\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_n$ 构成$\mathbb{R}^n$ 的标准正交基. 同理, $QQ^T = I$当且仅当Q的行向量构成$\mathbb{R}^n$的标准正交基. □

例如, 旋转矩阵就是正交矩阵. 再如

$$Q = \begin{pmatrix} \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{6}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{3}} & -\frac{2}{\sqrt{6}} & 0 \end{pmatrix}$$

是正交矩阵.

定理7.2.4. *正交矩阵有下列性质:*

(1) 单位阵是正交方阵.

(2) A, B为同阶正交方阵, 则AB也为正交方阵.

(3) 设A是正交方阵, 则$A^T = A^{-1}$也是正交方阵.

(4) 正交方阵的行列式的值为1或-1.

(5) 正交方阵的特征值的模长为1.

证明. 我们只证(5), 其它几条留给读者自己证明.

设 $A\mathbf{x}=\lambda\mathbf{x}$, 两边转置得 $\mathbf{x}^TA^T=\lambda\mathbf{x}^T$. 两边取复共轭, 由于 A 为实矩阵, 我们有 $\bar{\mathbf{x}}^TA^T=\bar{\lambda}\bar{\mathbf{x}}^T$. 因此 $\bar{\mathbf{x}}^TA^TA\mathbf{x}=\bar{\lambda}\bar{\mathbf{x}}^T\cdot\lambda\mathbf{x}=|\lambda|^2\bar{\mathbf{x}}^T\mathbf{x}$. 从而 $\bar{\mathbf{x}}^T\mathbf{x}=|\lambda|^2\bar{\mathbf{x}}^T\mathbf{x}$.

若记 $\mathbf{x}=(x_1,x_2,\cdots,x_n)^T$, 则 $\bar{\mathbf{x}}^T\mathbf{x}=\sum\limits_{i=1}^{n}|x_i|^2>0$, 因为特征向量不是零向量. 所以 $|\lambda|^2=1$, 即 $|\lambda|=1$.

□

有了标准正交基, 任意向量在其下的坐标可以通过内积方便获得.

命题7.2.5. 设$\mathbf{u}_1,\ldots,\mathbf{u}_n$为$\mathbb{R}^n$的一组标准正交基, 则任意$\mathbf{a}\in\mathbb{R}^n$,

$$\mathbf{a}=(\mathbf{a},\mathbf{u}_1)\mathbf{u}_1+(\mathbf{a},\mathbf{u}_2)\mathbf{u}_2+\ldots+(\mathbf{a},\mathbf{u}_n)\mathbf{u}_n. \tag{7.21}$$

证明. 设

$$\mathbf{a}=c_1\mathbf{u}_1+c_2\mathbf{u}_2+\ldots+c_n\mathbf{u}_n,$$

将上式两边与$\mathbf{u}_i$做内积, 由$\mathbf{u}_1,\ldots,\mathbf{u}_n$的单位正交性得,

$$(\mathbf{a},\mathbf{u}_i)=c_1(\mathbf{u}_1,\mathbf{u}_i)+\ldots+c_i(\mathbf{u}_i,\mathbf{u}_i)+\ldots+(\mathbf{u}_n,\mathbf{u}_i)=c_i.$$

命题得证. □

下面研究正交子空间.

定义7.2.3. 设W_1,W_2是欧氏空间$\mathbb{R}^n$的子空间. 对向量$\mathbf{a}\in\mathbb{R}^n$, 若$\mathbf{a}\perp\mathbf{b}$对任意$\mathbf{b}\in W_1$成立, 则称$\mathbf{a}$与子空间$W_1$正交, 记为$\mathbf{a}\perp W_1$. 若对任意$\mathbf{a}\in W_1$, $\mathbf{b}\in W_2$都有$\mathbf{a}\perp\mathbf{b}$, 则称$W_1$与$W_2$ 正交, 记为$W_1\perp W_2$.

定义7.2.4. 设W_1,W_2是欧氏空间$\mathbb{R}^n$的子空间. 则向量集合

$$W_1+W_2:=\{\mathbf{a}+\mathbf{b}\,|\,\mathbf{a}\in W_1,\mathbf{b}\in W_2\}$$

也是$\mathbb{R}^n$的子空间(请读者自行验证), 我们称之为子空间W_1与W_2的**和**. 如果W_1,W_2还满足$W_1\cap W_2=\{\mathbf{0}\}$, 则称和$W_1+W_2$ 为**直和**, 记为$W_1\oplus W_2$.

定理7.2.6. 设W是欧氏空间$\mathbb{R}^n$的子空间, 则向量集合

$$W^{\perp}:=\{\mathbf{a}\in\mathbb{R}^n\,|\,\mathbf{a}\perp W\}$$

是$\mathbb{R}^n$的子空间, 且有$\mathbb{R}^n=W\oplus W^{\perp}$. 称$W^{\perp}$为$W$的**正交补空间**.

证明. 先证$W^{\perp}$是子空间. 实际上, 设$\mathbf{a},\mathbf{b}\in W^{\perp}$及$\lambda,\mu\in\mathbb{R}$, 则对任意$\mathbf{c}\in W$, $(\mathbf{a},\mathbf{c})=0$, $(\mathbf{b},\mathbf{c})=0$. 于是

$$(\lambda\mathbf{a}+\mu\mathbf{b},\mathbf{c})=\lambda(\mathbf{a},\mathbf{c})+\mu(\mathbf{b},\mathbf{c})=0,$$

即$\lambda\mathbf{a}+\mu\mathbf{b}\in W^{\perp}$. 因此, $W^{\perp}$为子空间.

接下来证明, $W\cap W^{\perp}=0$. 实际上, 设$\mathbf{a}\in W\cap W^{\perp}$, 则$(\mathbf{a},\mathbf{a})=0$, 从而$\mathbf{a}=0$. 此即, $W\cap W^{\perp}=0$.

最后说明, $\mathbb{R}^n=W\oplus W^{\perp}$. 为此设$\mathbf{a}_1,\ldots,\mathbf{a}_r$为$W$的一组基. 对任意$\mathbf{x}\in W^{\perp}$, 则有$(\mathbf{a}_i,\mathbf{x})=\mathbf{a}_i^T\mathbf{x}=0$, $i=1,2,\ldots,r$. 即

$$\begin{pmatrix}\mathbf{a}_1^T\\ \vdots\\ \mathbf{a}_r^T\end{pmatrix}\mathbf{x}=0.$$

$W^{\perp}$即上述线性方程组的解集, 因此是一个$n-r$维的子空间. 设$\mathbf{a}_{r+1},\ldots,\mathbf{a}_n$为$W^{\perp}$的一组基, 下面我们证明$\mathbf{a}_1,\mathbf{a}_2,\ldots,\mathbf{a}_n$为$\mathbb{R}^n$的一组基, 从而 $\mathbb{R}^n=W\oplus W^{\perp}$. 实际上, 假设

$$\lambda_1\mathbf{a}_1+\ldots+\lambda_r\mathbf{a}_r+\ldots+\lambda_n\mathbf{a}_n=\mathbf{0}.$$

则

$$\lambda_1\mathbf{a}_1+\ldots+\lambda_r\mathbf{a}_r=-\lambda_{r+1}\mathbf{a}_{r+1}-\ldots-\lambda_n\mathbf{a}_n\in W\cap W^{\perp}=\{\mathbf{0}\},$$

故

$$\lambda_1\mathbf{a}_1+\ldots+\lambda_r\mathbf{a}_r=\mathbf{0},\qquad -\lambda_{r+1}\mathbf{a}_{r+1}-\ldots-\lambda_n\mathbf{a}_n=\mathbf{0}.$$

从而$\lambda_i=0$, $i=1,2,\ldots,n$, 即$\mathbf{a}_1,\mathbf{a}_2,\ldots,\mathbf{a}_n$线性无关. 因而$\mathbf{a}_1,\mathbf{a}_2,\ldots,\mathbf{a}_n$为$\mathbb{R}^n$ 的一组基. □

例7.2.2. 设$\mathbf{a}=(1,0,1,0)^T$, $\mathbf{b}=(0,1,0,1)^T$, $W=\langle\mathbf{a},\mathbf{b}\rangle\subset\mathbb{R}^4$. 求$W^{\perp}$的一组基.

解. 设$\mathbf{x}=(x_1,x_2,x_3,x_4)^T\in W^{\perp}$. 则由$\mathbf{a}\cdot\mathbf{x}=0$, $\mathbf{b}\cdot\mathbf{x}=0$得

$$\begin{cases}x_1+x_3=0\\ x_2+x_4=0\end{cases}$$

以求得上述线性方程组的两个线性无关解为$\mathbf{x}=(1,0,-1,0)^T$及$(0,1,0,-1)^T$, 它们构成$W^{\perp}$的一组基. □

接下来研究向量在子空间上的正交投影.

定义7.2.5. 设W是欧氏空间$\mathbb{R}^n$的子空间, $\mathbf{x}\in\mathbb{R}^n$. 如果向量$\mathbf{y}\in W$满足$\mathbf{x}-\mathbf{y}\perp W$, 则称$\mathbf{y}$为向量$\mathbf{x}$在$W$上的正交投影, 记为$\mathbf{y}=P\mathbf{x}$.

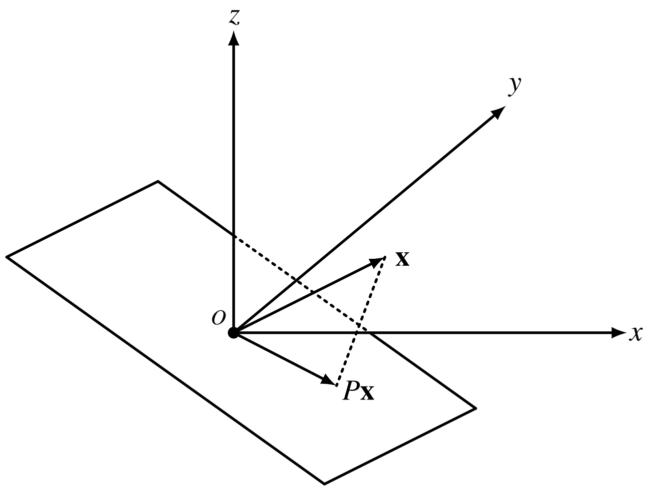


图 7.1 向量的正交投影

图 7.1给出了向量在子空间投影的例子. 显然, 若$\mathbf{x} \in W$, 则$P\mathbf{x} = \mathbf{x}$.

定理7.2.7. 设W是欧氏空间$\mathbb{R}^n$的子空间, $\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_r$是$W$的一组标准正交基. 则下列结论成立:

1. 对任意$\mathbf{x} \in \mathbb{R}^n$, $\mathbf{x}$在W上的正交投影存在唯一, 且

$$P\mathbf{x} = (\mathbf{x}, \mathbf{u}_1)\mathbf{u}_1 + (\mathbf{x}, \mathbf{u}_2)\mathbf{u}_2 + \ldots + (\mathbf{x}, \mathbf{u}_r)\mathbf{u}_r. \tag{7.22}$$

2. 对任意$\mathbf{y} \in W$,

$$|\mathbf{x} - \mathbf{y}| \geq |\mathbf{x} - P\mathbf{x}|.$$

   等号成立当且仅当$\mathbf{y} = P\mathbf{x}$.

证明. 1. 先证投影的唯一性. 设$\mathbf{y}_1, \mathbf{y}_2 \in W$是$\mathbf{x}$在$W$上的正交投影, 则$\mathbf{x} - \mathbf{y}_1 \perp W$, $\mathbf{x} - \mathbf{y}_2 \perp W$. 于是$\mathbf{y}_1 - \mathbf{y}_2 = (\mathbf{x} - \mathbf{y}_2) - (\mathbf{x} - \mathbf{y}_1) \perp W$. 而$\mathbf{y}_1 - \mathbf{y}_2 \in W$, 故$\mathbf{y}_1 = \mathbf{y}_2$.

设$P\mathbf{x} = c_1\mathbf{u}_1 + c_2\mathbf{u}_2 + \ldots + c_r\mathbf{u}_r$. 由$\mathbf{x} - P\mathbf{x} \perp W$, 故

$$(\mathbf{x} - P\mathbf{x}, \mathbf{u}_i) = 0, \quad i = 1, 2, \ldots, r.$$

即

$$(\mathbf{x}, \mathbf{u}_i) = (P\mathbf{x}, \mathbf{u}_i) = c_i.$$

因此

$$P\mathbf{x} = (\mathbf{x}, \mathbf{u}_1)\mathbf{u}_1 + (\mathbf{x}, \mathbf{u}_2)\mathbf{u}_2 + \ldots + (\mathbf{x}, \mathbf{u}_r)\mathbf{u}_r.$$

- 由勾股定理,

$$|\mathbf{x}-\mathbf{y}|^2=|(\mathbf{x}-P\mathbf{x})+(P\mathbf{x}-\mathbf{y})|^2=|\mathbf{x}-P\mathbf{x}|^2+|P\mathbf{x}-\mathbf{y}|^2\geq|\mathbf{x}-P\mathbf{x}|^2.$$

上式等号成立当且仅当$\mathbf{y}=P\mathbf{x}$.

□

上述定理表明, 对任意向量$\mathbf{x}\in\mathbb{R}^n$, $\mathbf{x}$在子空间W上的正交投影$P\mathbf{x}$就是$\mathbf{x}$与W中距离最小的向量.

下面我们给一个例子.

例7.2.3. 设$W=\langle(1,1,0),(1,0,-1)\rangle$是欧氏空间$\mathbb{R}^3$的子空间. 求向量$(2,1,3)$在$W$上的正交投影.

解法一. 将向量组$(1,1,0),(1,0,-1)$正交化得$\mathbf{u}_1=\frac{1}{\sqrt{2}}(1,1,0)$, $\mathbf{u}_2=\frac{1}{\sqrt{6}}(1,-1,-2)$. 它们构成$W$的一组标准正交基. 于是向量$\mathbf{x}=(2,1,3)$在$W$上的正交投影为

$$P\mathbf{x}=(\mathbf{x},\mathbf{u}_1)\mathbf{u}_1+(\mathbf{x},\mathbf{u}_2)\mathbf{u}_2=\frac{3}{2}(1,1,0)-\frac{5}{6}(1,-1,-2)=\frac{1}{3}(2,7,5).$$

□

解法二. 令$\mathbf{a}_1=(1,1,0)$, $\mathbf{a}_2=(1,0,-1)$. 设向量$\mathbf{x}=(2,1,3)$在W上的正交投影为$P\mathbf{x}=c_1\mathbf{a}_1+c_2\mathbf{a}_2$, 这里$c_1,c_2$ 为待定常数. 则由$\mathbf{x}-P\mathbf{x}\perp W$得

$$(\mathbf{x}-c_1\mathbf{a}_1+c_2\mathbf{a}_2,\mathbf{a}_1)=0,\quad(\mathbf{x}-c_1\mathbf{a}_1+c_2\mathbf{a}_2,\mathbf{a}_2)=0.$$

上式即

$$\begin{cases}(\mathbf{a}_1,\mathbf{a}_1)c_1+(\mathbf{a}_1,\mathbf{a}_2)c_2=(\mathbf{x},\mathbf{a}_1)\\(\mathbf{a}_2,\mathbf{a}_1)c_1+(\mathbf{a}_2,\mathbf{a}_2)c_2=(\mathbf{x},\mathbf{a}_2)\end{cases}$$

将$\mathbf{a}_1,\mathbf{a}_2,\mathbf{x}$代入上式解得$c_1=\frac{7}{3},c_2=-\frac{5}{3}$. 因此, 向量$\mathbf{x}=(2,1,3)$在$W$上的正交投影为

$$P\mathbf{x}=c_1\mathbf{a}_1+c_2\mathbf{a}_2=\frac{7}{3}(1,1,0)-\frac{5}{3}(1,0,-1)=\frac{1}{3}(2,7,5).$$

□

例7.2.4. 设$A\in\mathbb{R}^{m\times n}$, $\mathbf{b}\in\mathbb{R}^m$, 且$A$列满秩. 求线性方程组$A\mathbf{x}=\mathbf{b}$的**最小二乘解**, 即求$\mathbf{x}\in\mathbb{R}^n$使得$|A\mathbf{x}-\mathbf{b}|$达到最小.

解. 若线性方程组$A\mathbf{x}=\mathbf{b}$有解$\mathbf{x}$, 则$\mathbf{x}$显然是该线性方程组的最小二乘解. 下设线性方程组$A\mathbf{x}=\mathbf{b}$ 无解, 即$\mathbf{b}\notin C(A)$, 这里$C(A)$是矩阵A的列空间. 若$|A\mathbf{x}-\mathbf{b}|$达到最小, 由定理 7.2.7及$A\mathbf{x}\in C(A)$ 知, $A\mathbf{x}$ 正是$\mathbf{b}$在$C(A)$上的正交投影, 即$A\mathbf{x}-\mathbf{b}\perp C(A)$. 于是$A^T(A\mathbf{x}-\mathbf{b})=0$. 即

$$(A^TA)\mathbf{x}=A^T\mathbf{b}.$$

由于A列满秩, $r(A^TA)=r(A)=n$. 从而上述线性方程组有唯一解$\mathbf{x}=(A^TA)^{-1}A^T\mathbf{b}$. 它就是线性方程组$A\mathbf{x}=\mathbf{b}$的最小二乘解. □

§7.3 $\mathbb{R}^n$上的线性变换

由于有了内积, 欧氏空间既保留了线性空间的代数性质, 又具有几何性质. 本节我们感兴趣的是那些与度量密切相关的线性变换. 我们主要讨论两类变换–正交变换与对称变换.

§7.3.1 正交变换

定义了标准内积的n维欧氏空间$\mathbb{R}^n$上的正交变换就是由正交方阵定义的线性变换. 本节中, n维欧氏空间$\mathbb{R}^n$均指定义了标准内积的实数组空间.

定义7.3.1. n维欧氏空间$\mathbb{R}^n$上的正交变换定义为$\mathscr{A}:\mathbf{x}\to A\mathbf{x}$, 这里$\mathbf{x}\in\mathbb{R}^n$, A为n阶正交矩阵.

例如, 平面上的旋转变换就是由矩阵$A=\begin{pmatrix}\cos(\theta) & -\sin(\theta)\\ \sin(\theta) & \cos(\theta)\end{pmatrix}$定义的正交变换.

正交变换有以下等价说法.

定理7.3.1. *设$\mathscr{A}:\mathbf{x}\to A\mathbf{x}$是$n$维欧氏空间$\mathbb{R}^n$上的线性变换. 则下列命题等价:*

1. *$\mathscr{A}$是正交变换;*
2. *对任意$\mathbf{x}\in\mathbb{R}^n$, $|\mathscr{A}\mathbf{x}|=|\mathbf{x}|$;*
3. *对任意$\mathbf{x},\mathbf{y}\in\mathbb{R}^n$, $(\mathscr{A}\mathbf{x},\mathscr{A}\mathbf{y})=(\mathbf{x},\mathbf{y})$.*
4. *$\mathscr{A}$将一组标准正交基变为另一组标准正交基.*
5. *$\mathscr{A}$在任一组标准正交基下的矩阵为正交方阵.*

Proof. $1 \Rightarrow 2$. 由于

$$|\mathscr{A}\mathbf{x}|^2 = (\mathscr{A}\mathbf{x}, \mathscr{A}\mathbf{x}) = (A\mathbf{x})^T A\mathbf{x} = \mathbf{x}^T A^T A\mathbf{x} = \mathbf{x}^T\mathbf{x} = (\mathbf{x}, \mathbf{x}) = |\mathbf{x}|^2$$

故$|\mathscr{A}\mathbf{x}| = |\mathbf{x}|$.

$2 \Rightarrow 3$. 一方面,

$$\begin{aligned}
|\mathscr{A}(\mathbf{x}+\mathbf{y})|^2 &= (\mathscr{A}(\mathbf{x}+\mathbf{y}), \mathscr{A}(\mathbf{x}+\mathbf{y})) \\
&= (\mathscr{A}\mathbf{x}, \mathscr{A}\mathbf{x}) + 2(\mathscr{A}\mathbf{x}, \mathscr{A}\mathbf{y}) + (\mathscr{A}\mathbf{y}, \mathscr{A}\mathbf{y}) \\
&= |\mathscr{A}\mathbf{x}|^2 + 2(\mathscr{A}\mathbf{x}, \mathscr{A}\mathbf{y}) + |\mathscr{A}\mathbf{y}|^2,
\end{aligned}$$

另一方面,

$$\begin{aligned}
|\mathbf{x}+\mathbf{y}|^2 &= (\mathbf{x}+\mathbf{y}, \mathbf{x}+\mathbf{y}) = (\mathbf{x}, \mathbf{x}) + 2(\mathbf{x}, \mathbf{y}) + (\mathbf{y}, \mathbf{y}) \\
&= |\mathbf{x}|^2 + 2(\mathbf{x}, \mathbf{y}) + |\mathbf{y}|^2,
\end{aligned}$$

由$\mathscr{A}(\mathbf{x}+\mathbf{y})| = |\mathbf{x}+\mathbf{y}|$, $\mathscr{A}\mathbf{x}| = |\mathbf{x}|$, $\mathscr{A}\mathbf{y}| = |\mathbf{y}|$ 知,

$$(\mathscr{A}\mathbf{x}, \mathscr{A}\mathbf{y}) = (\mathbf{x}, \mathbf{y}).$$

$3 \Rightarrow 4$. 设$\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_n$是$\mathbb{R}^n$的一组标准正交基, 则由

$$(\mathscr{A}\mathbf{u}_i, \mathscr{A}\mathbf{u}_j) = (\mathbf{u}_i, \mathbf{u}_j) = \delta_{ij}$$

知, $\mathscr{A}\mathbf{u}_i$, $i = 1, 2, \ldots, n$也为$\mathbb{R}^n$的一组标准正交基.

$4 \Rightarrow 5$.

设$\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_n$是$\mathbb{R}^n$的任一组(列)标准正交基, $\mathscr{A}$在该组基下的矩阵为$\tilde{A}$. 则

$$(\mathscr{A}\mathbf{u}_1, \mathscr{A}\mathbf{u}_2, \ldots, \mathscr{A}\mathbf{u}_n) = \mathscr{A}(\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_n) = (\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_n)\tilde{A}.$$

由于$\mathscr{A}\mathbf{u}_1, \mathscr{A}\mathbf{u}_2, \ldots, \mathscr{A}\mathbf{u}_n$也是一组标准正交基, 故$(\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_n)$与$(\mathscr{A}\mathbf{u}_1, \mathscr{A}\mathbf{u}_2, \ldots, \mathscr{A}\mathbf{u}_n)$均为正交方阵. 从而

$$\tilde{A} = \begin{pmatrix} \mathscr{A}\mathbf{u}_1 & \mathscr{A}\mathbf{u}_2 & \ldots & \mathscr{A}\mathbf{u}_n \end{pmatrix} \begin{pmatrix} \mathbf{u}_1 & \mathbf{u}_2 & \ldots & \mathbf{u}_n \end{pmatrix}^{-1}$$

也为正交方阵.

$5 \Rightarrow 1$.

$\mathbb{R}^n$的单位坐标向量为一组标准正交基, 因此$\mathscr{A}$在单位坐标向量下的矩阵A也是正交方阵, 从而$\mathscr{A}$是正交变换. □

由上述定理知道, 正交变换就是保持度量(内积、长度与夹角)都不变的变换, 也就是一种刚体变换, 如旋转变换, 反射变换等.

定理7.3.2. *正交变换有下列性质:*

1. *恒同变换是正交变换;*
2. *两个正交变换的复合是正交变换;*
3. *正交变换的逆变换是正交变换;*
4. *正交变换的行列式的值为1或-1;*
5. *正交变换的特征值的模长为1.*

正交变换的性质由正交矩阵决定, 请读者自己完成上述定理的证明.

下面我们具体看看二维与三维欧氏空间中的正交变换.

例7.3.1. 确定$\mathbb{R}^2$上的正交变换.

解. 设$\mathscr{A}$是$\mathbb{R}^2$上的正交变换: $\mathbf{x} \to A\mathbf{x}$, 其中$A$为二阶正交方阵. 显然,

$$A = \begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix} \quad 或 \quad A = \begin{pmatrix} \cos(\theta) & \sin(\theta) \\ \sin(\theta) & -\cos(\theta) \end{pmatrix}$$

当$A = \begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix}$时, $\mathscr{A}$是平面旋转变换. 此时, $\det(\mathscr{A}) = 1$, 且$\mathscr{A}$的特征值为$\cos(\theta) \pm \sin(\theta)$.

当$A = \begin{pmatrix} \cos(\theta) & \sin(\theta) \\ \sin(\theta) & -\cos(\theta) \end{pmatrix}$时, $\det(\mathscr{A}) = -1$, $\mathscr{A}$的特征值为± 1. 容易验证, 对任意a, 变换$\mathscr{A}$将$(a,0)$变为$(a\cos(\theta), a\sin(\theta))$, 并将$(a\cos(\theta), a\sin(\theta))$ 变为$(a,0)$. 因此, $\mathscr{A}$是关于过原点且与x轴正向夹角为$\theta/2$的直线l的反射变换.

□

设$\mathscr{A}$为正交变换. 若$\det(\mathscr{A}) = 1$, 则称$\mathscr{A}$为**第一类**正交变换; 若$\det(\mathscr{A}) = -1$, 则称$\mathscr{A}$为**第二类**正交变换.

例7.3.2. 设$\mathscr{A}$为$\mathbb{R}^3$上的第一类正交变换. 证明: $\mathscr{A}$是绕过原点的某直线的旋转变换.

证明. 设线性变换 $\mathscr{A}$为$\mathbf{x} \to A\mathbf{x}$, 其中$A$为三阶正交方阵. 设$\mathscr{A}$的三个特征值分别为$\lambda_1, \lambda_2, \lambda_3$. 由于$\mathscr{A}$是第一类正交变换, 故$\det(\mathscr{A}) = \lambda_1\lambda_2\lambda_3 = 1$. 下面证明, $\lambda_1, \lambda_2, \lambda_3$至少有一个为1.

实际上, 由于三次实系数多项式必有实根, 因此λ_i, $i = 1,2,3$中至少有一个为实数. 我们考虑两种情况:

1. $\lambda_i, i=1,2,3$均为实数. 由于$|\lambda_i|=1, i=1,2,3$且$\lambda_1\lambda_2\lambda_3=1$, 因此至少有一个$\lambda_i$为1, 不妨设$\lambda_1=1$. 否则, $\lambda_1\lambda_2\lambda_3=-1$.

2. λ_1为实数, λ_2,λ_3为一对共轭复数, 则$\lambda_2\lambda_3=1$. 从而, $\lambda_1=1$.

设$\mathbf{x}_1$是对应特征值$\lambda_1=1$的特征向量, 即$\mathscr{A}\mathbf{x}_1=\mathbf{x}_1$. 于是$\mathscr{A}$就是绕$\mathbf{x}_1$的旋转变换.

□

§7.3.2 对称变换

接下来我们讨论对称变换. $\mathbb{R}^n$上的对称变换就是由对称方阵定义的变换.

定义7.3.2. n维欧氏空间$\mathbb{R}^n$上的对称变换定义为$\mathscr{A}:\mathbf{x}\to A\mathbf{x}$, 这里$\mathbf{x}\in\mathbb{R}^n$, A为n阶实对称方阵.

对称变换有下列等价说法.

定理7.3.3. *设$\mathscr{A}:\mathbf{x}\to A\mathbf{x}$是$n$维欧氏空间$\mathbb{R}^n$上的线性变换. 则下列命题等价:*

1. *$\mathscr{A}$是对称变换;*

2. *对任意$\mathbf{x},\mathbf{y}\in\mathbb{R}^n$, $(\mathscr{A}\mathbf{x},\mathbf{y})=(\mathbf{x},\mathscr{A}\mathbf{y})$;*

3. *$\mathscr{A}$在任一组标准正交基下的矩阵是实对称方阵.*

证明. : $1\Rightarrow 2$. 对任意$\mathbf{x},\mathbf{y}\in\mathbb{R}^n$,

$$(\mathscr{A}\mathbf{x},\mathbf{y})=(A\mathbf{x})^T\mathbf{y}=\mathbf{x}^TA^T\mathbf{y}=\mathbf{x}^TA\mathbf{y}=(\mathbf{x},\mathscr{A}\mathbf{y}).$$

$2\Rightarrow 3$. 设$\mathbf{u}_1,\mathbf{u}_2,\ldots,\mathbf{u}_n$为$\mathbb{R}^n$的一组(列)标准正交基, $\mathscr{A}$在该组基下的矩阵为$\tilde{A}=(\tilde{a}_{ij})_{n\times n}$. 则

$$(\mathscr{A}\mathbf{u}_1,\mathscr{A}\mathbf{u}_2,\ldots,\mathscr{A}\mathbf{u}_n)=\mathscr{A}(\mathbf{u}_1,\mathbf{u}_2,\ldots,\mathbf{u}_n)=(\mathbf{u}_1,\mathbf{u}_2,\ldots,\mathbf{u}_n)\tilde{A},$$

于是

$$\tilde{A}=\begin{pmatrix}\mathbf{u}_1^T\\ \vdots\\ \mathbf{u}_n^T\end{pmatrix}(\mathscr{A}\mathbf{u}_1,\mathscr{A}\mathbf{u}_2,\ldots,\mathscr{A}\mathbf{u}_n)=\begin{pmatrix}(\mathbf{u}_1,\mathscr{A}\mathbf{u}_1) & (\mathbf{u}_1,\mathscr{A}\mathbf{u}_2) & \cdots & (\mathbf{u}_1,\mathscr{A}\mathbf{u}_n)\\ (\mathbf{u}_2,\mathscr{A}\mathbf{u}_1) & (\mathbf{u}_2,\mathscr{A}\mathbf{u}_2) & \cdots & (\mathbf{u}_2,\mathscr{A}\mathbf{u}_n)\\ \vdots & \vdots & \vdots & \vdots\\ (\mathbf{u}_n,\mathscr{A}\mathbf{u}_1) & (\mathbf{u}_n,\mathscr{A}\mathbf{u}_2) & \cdots & (\mathbf{u}_n,\mathscr{A}\mathbf{u}_n)\end{pmatrix}$$

由$(\mathscr{A}\mathbf{u}_i,\mathbf{u}_j)=(\mathbf{u}_i,\mathscr{A}\mathbf{u}_j)$即知, $\tilde{A}$为对称方阵.

$3\Rightarrow 1$.

$\mathbb{R}^n$的单位坐标向量为一组标准正交基, 因此$\mathscr{A}$在单位坐标向量下的矩阵A也是实对称方阵, 从而$\mathscr{A}$是对称变换. □

接下来我们看看对称变换的特性. 为此, 先证明两个预备结论.

命题7.3.4. 实对称方阵的特征值均为实数.

证明. 设A为n阶实对称方阵, λ是A的特征值, $\mathbf{x}$是对应特征向量, 则$A\mathbf{x} = \lambda\mathbf{x}$. 对该式两边做共轭转置得$\bar{\mathbf{x}}^T A = \bar{\lambda}\bar{\mathbf{x}}^T$. 对前一式两边左乘$\bar{\mathbf{x}}^T$, 对后一式两边右乘$\mathbf{x}$得

$$\bar{\mathbf{x}}^T A\mathbf{x} = \lambda\bar{\mathbf{x}}^T\mathbf{x}, \qquad \bar{\mathbf{x}}^T A\mathbf{x} = \bar{\lambda}\bar{\mathbf{x}}^T\mathbf{x}.$$

于是$(\bar{\lambda} - \lambda)\bar{\mathbf{x}}^T\mathbf{x} = 0$. 设$\mathbf{x} = (x_1, x_2, \dots, x_n)^T \neq \mathbf{0}$, 则$\bar{\mathbf{x}}^T\mathbf{x} = \sum_{i=1}^n |x_i|^2 > 0$. 从而$\bar{\lambda} = \lambda$, 即$\lambda$为实数.

□

命题7.3.5. 实对称方阵的不同特征值对应的特征向量彼此正交.

证明. 设A为n阶实对称方阵, λ_1, λ_2是A的两个不相同的实特征值, $\mathbf{x}_1, \mathbf{x}_2$是对应的特征向量, 即

$$A\mathbf{x}_1 = \lambda_1\mathbf{x}_1, \quad A\mathbf{x}_2 = \lambda_2\mathbf{x}_2.$$

于是

$$\lambda_1(\mathbf{x}_1, \mathbf{x}_2) = (A\mathbf{x}_1, \mathbf{x}_2) = (\mathbf{x}_1, A\mathbf{x}_2) = \lambda_2(\mathbf{x}_1, \mathbf{x}_2).$$

因为$\lambda_1 \neq \lambda_2$, 所以$(\mathbf{x}_1, \mathbf{x}_2) = 0$, 即$\mathbf{x}_1$和$\mathbf{x}_2$相互正交. □

有了上面的预备知识, 我们可以证明实对称方阵的以下重要结论.

定理7.3.6. 对于任意n阶实对称阵A, 存在一个n阶正交阵T, 使得$T^{-1}AT = \mathrm{diag}(\lambda_1, \lambda_2, \dots, \lambda_n)$. 这里$\lambda_1, \lambda_2, \dots, \lambda_n$是$A$的特征值.

证明. 对矩阵A的阶数n用数学归纳法. 当$n = 1$时, 结论显然成立. 现假设$n - 1$时结论成立.现在看n的情形, 由命题 7.3.4知A的特征值都是实数, 设λ_1是A的一个特征值.$\mathbf{x}_1$是属于λ_1 的单位特征向量.则由Schmidt正交化可知, 可以将$\mathbf{x}_1$扩充成F^n的一组标准正交基$\mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_n$.

令$T_n = (\mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_n)$, 则$T_n$是一个正交矩阵.注意到$A$是对称阵, 则$T_n^{-1}AT_n$是对称的.利用$A\mathbf{x}_1 = \lambda_1\mathbf{x}_1$以及$\mathbf{x}_1$与$\mathbf{x}_i (2 \leq i \leq n)$的正交性有

$$T_n^{-1}AT_n = (\mathbf{x}_1, \dots, \mathbf{x}_n)^T(\lambda_1\mathbf{x}_1, A\mathbf{x}_2, \cdots, A\mathbf{x}_n) = \begin{pmatrix} \lambda_1 & 0 \\ 0 & A_{n-1} \end{pmatrix},$$

其中A_{n-1}是$n - 1$阶实对称阵, 由归纳假设, 存在$n - 1$阶正交阵T_{n-1}, 使得$T_{n-1}^{-1}A_{n-1}T_{n-1} = \mathrm{diag}(\lambda_2, \lambda_3, \cdots, \lambda_n)$.

令$T=T_n\begin{pmatrix}1 & 0\\ 0 & T_{n-1}\end{pmatrix}$, 显然$T$仍然是正交阵, 且有

$$
\begin{aligned}
T^{-1}AT &= \begin{pmatrix}1 & 0\\ 0 & T_{n-1}\end{pmatrix}^{-1} T_n^{-1}AT_n\begin{pmatrix}1 & 0\\ 0 & T_{n-1}\end{pmatrix} = \begin{pmatrix}1 & 0\\ 0 & T_{n-1}^{-1}\end{pmatrix}\begin{pmatrix}\lambda_1 & 0\\ 0 & A_{n-1}\end{pmatrix}\begin{pmatrix}1 & 0\\ 0 & T_{n-1}\end{pmatrix}\\
&= \begin{pmatrix}\lambda_1 & 0\\ 0 & T_{n-1}^{-1}A_{n-1}T_{n-1}\end{pmatrix} = \mathrm{diag}(\lambda_1,\lambda_2,\cdots,\lambda_n).
\end{aligned}
$$

定理证毕. □

下面我们以一个具体实例说明对称方阵对角化的方法.

例7.3.3. 设$A=\begin{pmatrix}1 & 2 & 2\\ 2 & 1 & 2\\ 2 & 2 & 1\end{pmatrix}$, 求正交阵$T$, 使$T^{-1}AT$为对角阵.

解. 矩阵A的特征多项式为

$$
p_A(\lambda)=\det(\lambda I-A)=\begin{vmatrix}\lambda-1 & -2 & -2\\ -2 & \lambda-1 & -2\\ -2 & -2 & \lambda-1\end{vmatrix}=(\lambda-5)(\lambda+1)^2.
$$

故A的特征值为$\lambda_1=5$, $\lambda_2=-1$(二重).

对于$\lambda_1=5$, 求解

$$(5I-A)\mathbf{x}=0$$

得特征向量$\mathbf{x}_1=(1,1,1)^T$, 把它单位化得$\mathbf{e}_1=\dfrac{1}{\sqrt{3}}(1,1,1)^T$.

对于$\lambda_2=-1$, 求解

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

得特征向量$\mathbf{x}_2=(-1,1,0)^T$, $\mathbf{x}_3=(-1,0,1)^T$. 把它们标准正交化得$\mathbf{e}_2=\frac{1}{\sqrt{2}}(-1,1,0)^T$, $\mathbf{e}_3=\frac{1}{\sqrt{6}}(-1,-1,2)^T$.

令

$$
T=(\mathbf{e}_1,\mathbf{e}_2,\mathbf{e}_3)=\begin{pmatrix}1/\sqrt{3} & -1/\sqrt{2} & -1/\sqrt{6}\\ 1/\sqrt{3} & 1/\sqrt{2} & -1/\sqrt{6}\\ 1/\sqrt{3} & 0 & 2/\sqrt{6}\end{pmatrix},
$$

则T为一个正交矩阵, 且有

$$
T^{-1}AT=\begin{pmatrix}5 & & \\ & -1 & \\ & & -1\end{pmatrix}.
$$

□

由定理 7.3.6, 我们即可得到对称变换的主要特性.

定理7.3.7. 设$\mathscr{A}$是欧氏空间$\mathbb{R}^n$上的对称变换. 则

1. $\mathscr{A}$的特征值均为实数;
2. $\mathscr{A}$的不同特征值对应的特征向量彼此正交;
3. 存在$\mathbb{R}^n$的一组标准正交基$\mathbf{u}_1,\mathbf{u}_2,\dots,\mathbf{u}_n$, 使得$\mathscr{A}$在这组基下的矩阵为$\mathrm{diag}(\lambda_1,\lambda_2,\dots,\lambda_n)$, 这里$\lambda_1,\lambda_2,\dots,\lambda_n$为$\mathscr{A}$的特征值.

证明. 结论1, 2是命题 7.3.4与7.3.5的显然推论. 下面我们证结论3.

设$\mathscr{A}$由实对称方阵A定义: $\mathbf{x}\to A\mathbf{x}$. 则$\mathscr{A}$在自然基$\mathbf{e}_1,\mathbf{e}_2,\dots,\mathbf{e}_n$下的矩阵为$A$, 即

$$\mathscr{A}(\mathbf{e}_1,\mathbf{e}_2,\dots,\mathbf{e}_n)=(\mathbf{e}_1,\mathbf{e}_2,\dots,\mathbf{e}_n)A.$$

由定理 7.3.6知, 存在正交方阵T使得$A=T\Lambda T^{-1}$, 其中$\Lambda=\mathrm{diag}(\lambda_1,\lambda_2,\dots,\lambda_n)$. 于是

$$\mathscr{A}((\mathbf{e}_1,\mathbf{e}_2,\dots,\mathbf{e}_n)T)=(\mathbf{e}_1,\mathbf{e}_2,\dots,\mathbf{e}_n)T\Lambda.$$

令$(\mathbf{u}_1,\mathbf{u}_2,\dots,\mathbf{u}_n)=(\mathbf{e}_1,\mathbf{e}_2,\dots,\mathbf{e}_n)T$. 由于$T$是正交方阵, 则$\mathbf{u}_1,\mathbf{u}_2,\dots,\mathbf{u}_n$ 是$\mathbb{R}^n$的一组标准正交基, 且有

$$\mathscr{A}(\mathbf{u}_1,\mathbf{u}_2,\dots,\mathbf{u}_n)=(\mathbf{u}_1,\mathbf{u}_2,\dots,\mathbf{u}_n)\Lambda.$$

□

上述定理表明, $\mathscr{A}\mathbf{u}_i=\lambda_i\mathbf{u}_i$, $i=1,2,\dots,n$. 也就是说, 对称变换就是沿n个互相正交的特征方向的伸缩(若$\lambda_i>0$)加反射(若$\lambda_i<0$).

§7.4 一般的欧氏空间

本节我们将数组内积空间的结果推广到一般的线性空间上.

§7.4.1 一般欧氏空间的定义与性质

设V是实数域$\mathbb{R}$上的线性空间, 如果我们在V上引进内积就可以得到一般的欧氏空间. 同数组空间$\mathbb{R}^n$不同, 我们无法直接在一般线性空间V上给出内积的公式表示, 我们只能根据内积所具有的性质来定义内积.

定义7.4.1. 设V是实数域$\mathbb{R}$上的线性空间, 如果V中任意两个向量α和β都按某一法则对应于一个实数, 记作(α,β), 且满足

(1) **对称性**:即对任意两个向量$\alpha,\beta\in V$, 有

$$(\alpha,\beta)=(\beta,\alpha) \tag{7.23}$$

(2) **线性性**:即对任意一个实数λ和任意三个向量$\alpha,\beta,\gamma\in V$, 有

$$(\lambda\alpha,\beta)=\lambda(\alpha,\beta) \tag{7.24}$$

$$(\alpha+\beta,\gamma)=(\alpha,\gamma)+(\beta,\gamma) \tag{7.25}$$

(3) **正定性**:即对于任意一个向量$\alpha\in V$, 有$(\alpha,\alpha)\geq 0$, 等号成立当且仅当$\alpha=\theta$.

则称(α,β)为向量α和β的**内积**, 定义了内积的实数域$\mathbb{R}$上的线性空间V称为**欧几里得(Euclid)空间**, 简称**欧氏空间**.

显然, 一般的欧氏空间是$\mathbb{R}^n$的推广. 但由于两者内积性质完全一样, 因此$\mathbb{R}^n$上的有关理论都可以平行推广到一般欧氏空间上. 下面, 我们直接叙述有关结论.

定理7.4.1. ***[Cauchy-Schwarz不等式] 设V是欧氏空间, $(\cdot,\cdot)$是V的内积, 则对V中的任意两个向量α和β, 有***

$$|(\alpha,\beta)|\leq\sqrt{(\alpha,\alpha)(\beta,\beta)}. \tag{7.26}$$

定义7.4.2. 设V是欧氏空间, $(\cdot,\cdot)$是V的内积, 对于任意的$\alpha\in V$, 称

$$|\alpha|=\sqrt{(\alpha,\alpha)} \tag{7.27}$$

称为α的**长度**或**模**.若$|\alpha|=1$, 则称α为单位向量. 对于欧氏空间V中两个非零向量α和β, 定义它们之间的夹角为

$$\theta=\arccos\frac{(\alpha,\beta)}{|\alpha||\beta|}. \tag{7.28}$$

特别, 当$(\alpha,\beta)=0$时, 称向量α和β**正交**或**垂直**, 记作$\alpha\perp\beta$.

一般欧氏空间中同样有勾股定理

$$|\alpha+\beta|^2=|\alpha|^2+|\beta|^2\Leftrightarrow(\alpha,\beta)=0. \tag{7.29}$$

此外, 利用Cauchy-Schwarz不等式可以证明下列三角不等式

$$|\alpha+\beta|\leq|\alpha|+|\beta|. \tag{7.30}$$

下面我们给出几个一般欧氏空间的具体例子.

例7.4.1. 对于实数域上任何一个n维线性空间V, 取定V的一组基$\alpha_1,\alpha_2,\ldots,\alpha_n$.对任意向量

$$\alpha=a_1\alpha_1+a_2\alpha_2+\cdots+a_n\alpha_n,\quad \beta=b_1\alpha_1+b_2\alpha_2+\cdots+b_n\alpha_n,$$

定义它们的内积如下

$$(\alpha,\beta)=a_1b_1+a_2b_2+\cdots+a_nb_n.$$

容易验证, 上述定义满足内积的三条基本性质. 在上述内积下, V成为一个欧氏空间.

上述例子说明, 实数域上n维线性空间V的内积是一定存在的, 从而可以构造相应的欧氏空间.

例7.4.2. 设$C[a,b]$是闭区间$[a,b]$上实的连续函数的全体, 它是实数域上的线性空间.对于任意的$f,g\in C[a,b]$, 定义

$$(f,g)=\int_a^b f(x)g(x)dx.$$

容易验证, 它满足内积定义中的(1)和(2).如果

$$(f,f)=\int_a^b f^2(x)dx=0,$$

则由连续函数的性质可知$f\equiv 0$, 所以(3)也满足.因此, (f,g) 定义了$C[a,b]$上的内积.对应的Cauchy-Schwarz不等式为

$$\left|\int_a^b f(x)g(x)dx\right|\le\sqrt{\int_a^b f^2(x)dx}\sqrt{\int_a^b g^2(x)dx}.$$

§7.4.2 标准正交基

一般欧氏空间的内积并没有一个具体公式. 但当我们给定空间的一组基$\alpha_1,\alpha_2,\ldots,\alpha_n$后, 内积可以用矩阵形式表示. 实际上, 设

$$\alpha=\sum_{i=1}^n x_i\alpha_i,\quad \beta=\sum_{i=1}^n y_i\alpha_i,$$

则

$$(\alpha,\beta)=\sum_{i=1}^n\sum_{j=1}^n x_iy_j(\alpha_i,\alpha_j)=\mathbf{x}^TG\mathbf{y},\tag{7.31}$$

其中$\mathbf{x}=(x_1,x_2,\ldots,x_n)^T$, $\mathbf{y}=(y_1,y_2,\ldots,y_n)^T$, $G=((\alpha_i,\alpha_j))_{n\times n}$是**度量矩阵**, 也称为**Gram 阵**. 由内积的正定性知, G 是实对称正定方阵. 注意到, 上式与欧氏空间$\mathbb{R}^n$的内积定义(7.4)一致.

特别地, 当$G=I$, 即$(\alpha_i,\alpha_j)=\delta_{ij}$时,

$$(\alpha,\beta)=\mathbf{x}^T\mathbf{y}. \tag{7.32}$$

定义7.4.3. 设$\alpha_1,\alpha_2,\ldots,\alpha_n$为欧氏空间$V$的一组基, 若$(\alpha_i,\alpha_j)=\delta_{ij}$, $i,j=1,2,\ldots,n$, 则称$\alpha_1,\alpha_2,\ldots,\alpha_n$为$V$的一组标准正交基.

定理7.4.2. *(**Schmidt正交化**) 从n维欧氏空间V的任意一组基$\alpha_1,\alpha_2,\ldots,\alpha_n$出发, 可以构造一组标准正交基$\beta_1,\beta_2,\ldots,\beta_n$使得$\{\alpha_1,\ldots,\alpha_k\}\sim\{\beta_1,\ldots,\beta_k\}$, $k=1,2,\ldots,n$.*

定义7.4.4. 设W_1,W_2是欧氏空间V的子空间. 对向量$\alpha\in V$, 若$\alpha\perp\beta$对任意$\beta\in W_1$成立, 则称α与子空间W_1正交, 记为$\alpha\perp W_1$. 若对任意$\alpha\in W_1$, $\beta\in W_2$都有$\alpha\perp\beta$, 则称W_1与W_2正交, 记为$W_1\perp W_2$.

定理7.4.3. *设W是欧氏空间V的子空间, 则向量集合*

$$W^{\perp}:=\{\alpha\in V\,|\,\alpha\perp W\}$$

*是V的子空间, 且有$V=W\oplus W^{\perp}$. 称$W^{\perp}$为W的*正交补空间.

定义7.4.5. 设W是欧氏空间V的子空间, $\alpha\in V$. 如果向量$\beta\in W$满足$\alpha-\beta\perp W$, 则称β为向量α在W上的正交投影, 记为$\beta=P\alpha$.

定理7.4.4. *设W是欧氏空间V的子空间, $\alpha_1,\alpha_2,\ldots,\alpha_r$是$W$的一组标准正交基. 则下列结论成立:*

1. 对任意$\alpha\in V$, α在W上的正交投影存在唯一, 且

$$P\alpha=(\alpha,\alpha_1)\alpha_1+(\alpha,\alpha_2)\alpha_2+\ldots+(\alpha,\alpha_r)\alpha_r.$$

2. 设$\alpha\in V$. 则对任意$\beta\in W$,

$$|\alpha-\beta|\geq|\alpha-P\alpha|.$$

等号成立当且仅当$\beta=P\alpha$.

例7.4.3. 设$C[-\pi,\pi]$是例 7.4.2中定义的欧氏空间. 令$W=\langle 1,\cos x,\sin x,\ldots,\cos(nx),\sin(nx)\rangle$ 为$C[-\pi,\pi]$ 的子空间. 容易验证,

$$\frac{1}{\sqrt{2\pi}},\frac{1}{\sqrt{\pi}}\cos x,\frac{1}{\sqrt{2\pi}}\sin x,\ldots,\frac{1}{\sqrt{2\pi}}\cos(nx),\frac{1}{\sqrt{2\pi}}\sin(nx)$$

为W的一组标准正交基. 因此, $f(x)\in C[-\pi,\pi]$在子空间W上的正交投影为

$$\frac{a_0}{2}+a_1\cos(x)+b_1\sin(x)+\ldots+a_n\cos(nx)+b_n\sin(nx),$$

其中

$$a_k=\frac{1}{\pi}(f(x),\cos(kx))=\frac{1}{\pi}\int_{-\pi}^{\pi}f(x)\cos(kx)\,dx,\quad b_k=\frac{1}{\pi}(f(x),\sin(kx))=\frac{1}{\pi}\int_{-\pi}^{\pi}f(x)\sin(kx)\,dx.$$

$k=1,2,\ldots,n$. 这正式$f(x)$的Fourier级数的前n项的和.

§7.4.3 正交变换与对称变换

接下来, 我们研究一般欧氏空间上的正交变换与对称变换. 这个时候, 我们是从变换的本质特征来定义.

定义7.4.6. 设V是一个n维的欧氏空间, $\mathscr{A}$是V上的一个线性变换, 如果$\mathscr{A}$保持V的内积不变, 即对于任意的两个向量$\alpha,\beta\in V$都有

$$(\mathscr{A}\alpha,\mathscr{A}\beta)=(\alpha,\beta). \tag{7.33}$$

则称$\mathscr{A}$是V内的**正交变换**.

定理7.4.5. *设V是一个n维的欧氏空间, $\mathscr{A}$是V上的一个线性变换, 则下列命题等价:*

1. *$\mathscr{A}$为正交变换;*
2. *$\mathscr{A}$保持任意向量的模长不变;*
3. *$\mathscr{A}$将标准正交基变换为标准正交基;*
4. *$\mathscr{A}$在任一组标准正交基下的矩阵为正交矩阵.*

证明. $1\Leftrightarrow 2$ 及$1\Rightarrow 3$ 与定理 7.3.1的证明一样. 我们仅证明$3\Rightarrow 4$及$4\Rightarrow 1$.

$3\Rightarrow 4$. 设$\alpha_1,\alpha_2,\ldots,\alpha_n$是欧氏空间$V$的标准正交基, $\mathscr{A}$在该组基下的矩阵为$A=(a_{ij})$, A的列分别为$A_1,A_2,\ldots,A_n$. 则

$$(\mathscr{A}\alpha_1,\mathscr{A}\alpha_2,\ldots,\mathscr{A}\alpha_n)=(\alpha_1,\alpha_2,\ldots,\alpha_n)A$$

或

$$\mathscr{A}\alpha_j=(\alpha_1,\alpha_2,\ldots,\alpha_n)A_j=\sum_{i=1}^{n}a_{ij}\alpha_i,\quad j=1,2,\ldots,n$$

于是

$$\delta_{ij} = (\mathscr{A}\alpha_i, \mathscr{A}\alpha_j) = \left(\sum_{l=1}^{n} a_{li}\alpha_l, \sum_{k=1}^{n} a_{kj}\alpha_k\right) = \sum_{l,k=1}^{n} a_{li}a_{kj}\delta_{lk} = \sum_{k=1}^{n} a_{ki}a_{kj} = A_i^T A_j,$$

即A是正交方阵.

$4 \Rightarrow 1$. 设$\alpha_1, \alpha_2, \ldots, \alpha_n$是$V$的一组标准正交基, $\mathscr{A}$在该组基下的矩阵A正交. 则

$$(\mathscr{A}\alpha_i, \mathscr{A}\alpha_j) = A_i^T A_j = \delta_{ij}, \qquad i, j = 1, 2, \ldots, n$$

于是对V中任意两个向量

$$\alpha = \sum_{i=1}^{n} a_i\alpha_i, \quad \beta = \sum_{i=1}^{n} b_i\alpha_i$$

有

$$\begin{aligned}(\mathscr{A}\alpha, \mathscr{A}\beta) &= \left(\mathscr{A}\left(\sum_{i=1}^{n} a_i\alpha_i\right), \mathscr{A}\left(\sum_{i=1}^{n} b_i\alpha_i\right)\right) = \sum_{i,j=1}^{n} a_i b_j (\mathscr{A}\alpha_i, \mathscr{A}\alpha_j) \\ &= \sum_{i,j=1}^{n} a_i b_j \delta_{ij} = \sum_{i=1}^{n} a_i b_i = (\alpha, \beta).\end{aligned}$$

即$\mathscr{A}$是正交变换. □

一般欧氏空间上的正交变换具有$\mathbb{R}^n$上的正交变换一样的性质, 详见定理7.3.2.

设$\mathscr{A}$是欧氏空间V上的正交变换, 则$\det(\mathscr{A}) = \pm 1$. 若$\det(\mathscr{A}) = 1$, 则称$\mathscr{A}$为**第一类正交变换**; 若$\det(\mathscr{A}) = -1$, 则称$\mathscr{A}$为**第二类正交变换**;

命题7.4.6. *设$\mathscr{A}$是欧氏空间V上的第一类正交变换, 且V的维数为奇数. 则$\mathscr{A}$一定有特征值1.*

证明. 由于$\mathscr{A}$是第一类正交变换, 所以它的所有特征值乘积为1. 因为V的维数为奇数, 故$p_{\mathscr{A}}(\lambda)$为奇数次多项式, 从而必有实根. 另一方面, 由于$p_{\mathscr{A}}(\lambda)$的复根成对共轭出现, 且任一对共轭复根$\mu, \bar{\mu}$的乘积为$\mu\bar{\mu} = |\mu|^2 = 1$, 故$p_{\mathscr{A}}(\lambda)$一定有奇数个实根, 且其乘积为1. 从而必有一个特征值为1. □

接下来介绍对称变换.

定义7.4.7. 设V是n维欧氏空间, $\mathscr{A}$是V上的线性变换.如果$\mathscr{A}$满足

$$(\alpha, \mathscr{A}\beta) = (\mathscr{A}\alpha, \beta) \tag{7.34}$$

对V中任意两个向量α和β成立, 则称$\mathscr{A}$是V上的**对称变换**.

定理7.4.7. 设$\mathscr{A}$是欧氏空间上的线性变换, 则$\mathscr{A}$是对称变换的充分必要条件是$\mathscr{A}$在任何一组标准正交基下的矩阵A是实对称方阵.

证明. 设$\alpha_1,\alpha_2,\dots,\alpha_n$是$V$中的标准正交基, $A=(a_{ij})$是对称变换$\mathscr{A}$在这组基下的矩阵, 即

$$(\mathscr{A}\alpha_1,\mathscr{A}\alpha_2,\dots,\mathscr{A}\alpha_n)=(\alpha_1,\alpha_2,\dots,\alpha_n)A$$

或写成

$$\mathscr{A}\alpha_j=a_{1j}\alpha_1+\cdots+a_{nj}\alpha_n=\sum_{i=1}^{n}a_{kj}\alpha_k.$$

因此,

$$(\alpha_i,\mathscr{A}\alpha_j)=(\alpha_i,\sum_{k=1}^{n}a_{kj}\alpha_k)=\sum_{k=1}^{n}a_{kj}(\alpha_i,\alpha_k)=a_{ij},$$

$$(\mathscr{A}\alpha_i,\alpha_j)=(\sum_{k=1}^{n}a_{ki}\alpha_k,\alpha_j)=\sum_{k=1}^{n}a_{ki}(\alpha_k,\alpha_j)=a_{ji}.$$

以上两式相等得$a_{ij}=a_{ji}$.

反之, 如果$\mathscr{A}$在标准正交基下的矩阵$A=(a_{ij})$满足$a_{ij}=a_{ji}$, 则

$$(\alpha_i,\mathscr{A}\alpha_j)=(\mathscr{A}\alpha_i,\alpha_j).$$

因此, 对于任意两个向量

$$\alpha=\sum_{i=1}^{n}a_i\alpha_i,\quad \beta=\sum_{j=1}^{n}b_j\alpha_j$$

有

$$(\alpha,\mathscr{A}\beta)=\sum_{i,j=1}^{n}a_ib_j(\alpha_i,\mathscr{A}\alpha_j)=\sum_{i,j=1}^{n}a_ib_j(\mathscr{A}\alpha_i,\alpha_j)=(\mathscr{A}\alpha,\beta).$$

即$\mathscr{A}$是对称变换. □

定理7.4.8. 设$\mathscr{A}$是欧氏空间V上的对称变换, 则$\mathscr{A}$的特征值均为实数, 且不同特征值对应的特征向量相互正交.

证明. 由定理 7.4.7, $\mathscr{A}$在一组标准正交基下的矩阵为实对称方阵, 从而其特征值为实数. 不同特征值对应特征向量的正交性的证明同命题 7.3.5. □

定理7.4.9. 设$\mathscr{A}$是n维欧氏空间V上的对称变换, 则存在V的一组标准正交基$\alpha_1,\alpha_2,\dots,\alpha_n$使得$\mathscr{A}\alpha_i=\lambda_i\alpha_i$, $i=1,2,\dots,n$. 这里$\lambda_1,\lambda_2,\dots,\lambda_n$为$\mathscr{A}$ 的特征值.

证明. 设$\beta_1,\beta_2,\ldots,\beta_n$是$V$的一组标准正交基, $\mathscr{A}$在该组基下的矩阵为实对称方阵, 记为A. 由定理, 存在可逆方阵T使得$A = T\,\mathrm{diag}(\lambda_1,\lambda_2,\ldots,\lambda_n)T^{-1}$. 记$(\alpha_1,\alpha_2,\ldots,\alpha_n)=(\beta_1,\beta_2,\ldots,\beta_n)T$, 则

$$\begin{aligned}
&\mathscr{A}(\alpha_1,\alpha_2,\ldots,\alpha_n)=\mathscr{A}((\beta_1,\beta_2,\ldots,\beta_n)T)\\
&=\mathscr{A}(\beta_1,\beta_2,\ldots,\beta_n)T=(\beta_1,\beta_2,\ldots,\beta_n)AT\\
&=(\beta_1,\beta_2,\ldots,\beta_n)T\,\mathrm{diag}(\lambda_1,\lambda_2,\ldots,\lambda_n)\\
&=(\alpha_1,\alpha_2,\ldots,\alpha_n)\,\mathrm{diag}(\lambda_1,\lambda_2,\ldots,\lambda_n).
\end{aligned}$$

因此, $\mathscr{A}\alpha_i=\lambda_i\alpha_i$, $i=1,2,\ldots,n$. □

§7.5 酉空间*

欧几里得空间是实数域上的线性空间装配上内积(或者说度量)后形成的.自然会问,可否将这些结果推广到复数域上的线性空间,形成一种复数域上的度量空间? 答案是可行的.

§7.5.1 酉空间的基本概念*

设$\mathbb{C}^n$为复数域$\mathbb{C}$上的n维数组空间. 如何在$\mathbb{C}^n$上引进内积使得其具有度量? 一个自然的想法是用$\mathbb{R}^n$一样的内积, 即对$\mathbf{a}=(a_1,a_2,\ldots,a_n)^T,\mathbf{b}=(b_1,b_2,\ldots,b_n)^T\in\mathbb{C}^n$, 定义

$$(\mathbf{a},\mathbf{b})=a_1b_1+a_2b_2+\ldots+a_nb_n.$$

则$(\mathbf{a},\mathbf{a})=\sum_{i=1}^n a_i^2$. 此时, 我们不能保证$(\mathbf{a},\mathbf{a})\geq 0$, 从而无法定义$\mathbf{a}$的范数. 因此, 欧氏空间的内积定义不适合推广到复数域上的线性空间. 但如果我们对上述内积公式做一点修改, 就可以保证$(\mathbf{a},\mathbf{a})\geq 0$. 事实上, 我们引入内积

$$(\mathbf{a},\mathbf{b})=\bar{a}_1b_1+\bar{a}_2b_2+\ldots+\bar{a}_nb_n=\mathbf{a}^*\mathbf{b}, \tag{7.35}$$

这里$\mathbf{a}^*=\bar{\mathbf{a}}^T=(\bar{a}_1,\bar{a}_2,\ldots,\bar{a}_n)$表示向量$\mathbf{a}$的共轭转置. 则$(\mathbf{a},\mathbf{a})=\sum_{i=1}^n|a_i|^2\geq 0$. 更一般地, 我们可以定义两个向量的内积

$$(\mathbf{a},\mathbf{b})=\mathbf{a}^*H\mathbf{b}, \tag{7.36}$$

其中方阵$H\in\mathbb{C}^{n\times n}$满足:

(1) H是Hermite方阵, 即$H^*=H$. 这里H^*表示H的共轭转置$\bar{H}^T$.

(2) H正定, 即对任意$\mathbf{x}\in\mathbb{C}^n$, $\mathbf{x}^*H\mathbf{x}\geq 0$, 且等号成立当且仅当$\mathbf{x}=0$.

定义7.5.1. 定义了内积 (7.36)的数组空间$\mathbb{C}^n$称为酉空间.

显然内积 (7.35)是内积 (7.36)的特殊情形($H=I$). 我们称 (7.35)为标准内积. 如无特别说明, 酉空间$\mathbb{C}^n$中的内积都指标准内积.

§7.5.2 酉空间的基本性质*

酉空间的内积性质与欧氏空间内积的性质有相似性, 但也有些不同.

定理7.5.1. 酉空间内积具有以下性质

(1) **共轭对称性**:即对任意两个向量$\mathbf{a},\mathbf{b}\in V$, 有

$$(\mathbf{a},\mathbf{b})=\overline{(\mathbf{b},\mathbf{a})}. \tag{7.37}$$

(2) **共轭线性性**:即对任意一个复数λ和任意三个向量$\mathbf{a},\mathbf{b},\mathbf{c}\in V$, 有

$$(\mathbf{a},\lambda\mathbf{b})=\lambda(\mathbf{a},\mathbf{b}),\quad (\mathbf{a},\mathbf{b}+\mathbf{c})=(\mathbf{a},\mathbf{b})+(\mathbf{a},\mathbf{c}) \tag{7.38}$$

$$(\lambda\mathbf{a},\mathbf{b})=\bar{\lambda}(\mathbf{a},\mathbf{b}),\quad (\mathbf{a}+\mathbf{b},\mathbf{c})=(\mathbf{a},\mathbf{c})+(\mathbf{b},\mathbf{c}). \tag{7.39}$$

(3) **正定性**:即对于任意一个向量$\mathbf{a}\in V$, 有$(\mathbf{a},\mathbf{a})\geq 0$, 等号成立当且仅当$\mathbf{a}=\mathbf{0}$.

证明. 请读者自己完成证明. □

内积的共轭线性表明, 内积关于第二个变量是线性的, 但关于第一个变量不是线性的, 而是所谓的共轭线性的, 即有

$$(\lambda_1\mathbf{a}+\lambda_2\mathbf{b},\mathbf{c})=\overline{\lambda_1}(\mathbf{a},\mathbf{c})+\overline{\lambda_2}(\mathbf{b},\mathbf{c}),$$
$$(\mathbf{a},\lambda_1\mathbf{b}+\lambda_2\mathbf{c})=\lambda_1(\mathbf{a},\mathbf{b})+\lambda_2(\mathbf{a},\mathbf{c}).$$

一般地, 我们有

$$\left(\sum_{i=1}^{k}\lambda_i\mathbf{a}_i,\sum_{j=1}^{m}\mu_j\mathbf{b}_j\right)=\sum_{i=1}^{k}\sum_{j=1}^{m}\overline{\lambda_i}\mu_j(\mathbf{a}_i,\mathbf{b}_j). \tag{7.40}$$

类似欧氏空间, 有了内积就可以定义酉空间中向量的长度或模.

定义7.5.2. 向量$\mathbf{a}$的**长度(模)**定义为

$$|\mathbf{a}|=\sqrt{(\mathbf{a},\mathbf{a})}. \tag{7.41}$$

然而对于酉空间, 由于任何两个向量之间的内积$(\mathbf{a}, \mathbf{b})$不再是一个实数, 所以无法定义两个向量之间的夹角. 这是酉空间和欧氏空间之间的一个重要区别! 尽管如此, 我们仍然可以定义酉空间中向量之间的“正交”或者“垂直”的概念.

定义7.5.3. 酉空间中两个向量$\mathbf{a}$和$\mathbf{b}$满足$(\mathbf{a}, \mathbf{b}) = 0$, 则称$\mathbf{a}$与$\mathbf{b}$相互**正交**或**垂直**, 记为$\mathbf{a} \perp \mathbf{b}$.

再次强调, 这里的“正交”或者“垂直”, 仅仅是指内积等于零, 而不是夹角等于$90°$, 因为这里没有夹角的概念.但欧氏空间中凡涉及正交性的结果均可移植到酉空间中. 例如, 可以考虑子空间的相互正交性以及子空间的正交补空间.

定理7.5.2. *设$\mathbb{C}^n$是n维酉空间, 则*

(1) *$\mathbb{C}^n$中两两正交的一组非零向量一定是线性无关的;*

(2) *$\mathbb{C}^n$中存在标准正交基, 即存在一组基$\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_n$, 满足*

$$(\mathbf{u}_i, \mathbf{u}_j) = \delta_{ij}, \quad i, j = 1, 2, \ldots, n$$

(3) *从酉空间中任何一组基出发, 可以通过Schmidt正交化过程, 得到一组标准正交基.*

读者可以自行完成上述定理的证明.下面, 我们用一个例子说明酉空间的Schmidt正交化过程. 该正交化方法与欧氏空间向量组的正交化完全一样, 只是在内积的计算时要注意用酉空间的内积而不是欧氏空间的内积.

例7.5.1. 将$\mathbb{C}^3$的一组基$\mathbf{a}_1 = (1, i, 1)$, $\mathbf{a}_2 = (-1, 0, 1)$与$\mathbf{a}_3 = (0, 1, i)$单位正交化.

解. 首先将$\mathbf{a}_1$单位化得

$$\mathbf{u}_1 = \frac{\mathbf{a}_1}{|\mathbf{a}_1|} = \frac{1}{\sqrt{3}}(1, i, 1).$$

令$\mathbf{b}_2 = \mathbf{a}_2 - (\mathbf{a}_2, \mathbf{u}_1)\mathbf{u}_1 = (-1, 0, 1)$, 则$(\mathbf{b}_2, \mathbf{u}_1) = 0$. 将$\mathbf{b}_2$单位化得

$$\mathbf{u}_2 = \frac{\mathbf{b}_2}{|\mathbf{b}_2|} = \frac{1}{\sqrt{2}}(-1, 0, 1).$$

令$\mathbf{b}_3 = \mathbf{a}_3 - (\mathbf{a}_3, \mathbf{u}_1)\mathbf{u}_1 - (\mathbf{a}_3, \mathbf{u}_2)\mathbf{u}_2 = \frac{1}{2}(i, 2, i)$, 则$(\mathbf{b}_3, \mathbf{u}_1) = (\mathbf{b}_3, \mathbf{u}_2) = 0$. 将$\mathbf{b}_3$正交化得

$$\mathbf{u}_3 = \frac{\mathbf{b}_3}{|\mathbf{b}_3|} = \frac{1}{\sqrt{6}}(i, 2, i).$$

则 $\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3$为$\mathbb{C}^3$的一组标准正交基.

□

根据 7.5.2定理, 设$\mathbf{u}_1,\mathbf{u}_2,\dots,\mathbf{u}_n$是酉空间$V$内的一组标准正交基, 对于$V$内任意两个向量$\mathbf{a},\mathbf{b}$

$$\mathbf{a}=a_1\mathbf{u}_1+a_2\mathbf{u}_2+\cdots+a_n\mathbf{u}_n,\quad \mathbf{b}=b_1\mathbf{u}_1+b_2\mathbf{u}_2+\cdots+b_n\mathbf{u}_n,$$

有

$$\begin{aligned}(\mathbf{a},\mathbf{b})&=\left(\sum_{i=1}^n a_i\mathbf{u}_i,\sum_{j=1}^n b_j\mathbf{u}_j\right)=\sum_{i=1}^n\sum_{j=1}^n\overline{a_i}b_j(\mathbf{u}_i,\mathbf{u}_j)\\&=\sum_{i=1}^n\sum_{j=1}^n\delta_{ij}\overline{a_i}b_j=\sum_{i=1}^n\overline{a_i}b_i.\end{aligned}$$

也就是说, 在酉空间的一组标准正交基下, 内积的表示与标准内积是一致的.

定义7.5.4. 设W_1,W_2是酉空间$\mathbb{C}^n$的子空间. 对向量$\mathbf{x}\in\mathbb{C}^n$, 若$\mathbf{x}\perp\mathbf{y}$对任意$\mathbf{y}\in W_1$成立, 则称$\mathbf{x}$与子空间$W_1$正交, 记为$\mathbf{x}\perp W_1$. 若对任意$\mathbf{x}\in W_1$, $\mathbf{y}\in W_2$都有$\mathbf{x}\perp\mathbf{y}$, 则称$W_1$与$W_2$正交, 记为$W_1\perp W_2$.

定理7.5.3. 设W是欧氏空间$\mathbb{C}^n$的子空间, 则向量集合

$$W^{\perp}:=\{\mathbf{a}\in\mathbb{C}^n\,|\,\mathbf{a}\perp W\}$$

是$\mathbb{C}^n$的子空间, 且有$\mathbb{C}^n=W\oplus W^{\perp}$. 称$W^{\perp}$为$W$的**正交补空间**.

定义7.5.5. 设W是欧氏空间$\mathbb{C}^n$的子空间, $\mathbf{x}\in\mathbb{C}^n$. 如果向量$\mathbf{y}\in W$满足$\mathbf{x}-\mathbf{y}\perp W$, 则称$\mathbf{y}$为向量$\mathbf{x}$在$W$上的正交投影, 记为$\mathbf{y}=P\mathbf{x}$.

定理7.5.4. 设W是欧氏空间$\mathbb{C}^n$的子空间, $\mathbf{u}_1,\mathbf{u}_2,\dots,\mathbf{u}_r$是$W$的一组标准正交基. 则下列结论成立:

1. 对任意$\mathbf{x}\in\mathbb{C}^n$, $\mathbf{x}$在W上的正交投影存在唯一, 且

$$P\mathbf{x}=(\mathbf{x},\mathbf{u}_1)\mathbf{u}_1+(\mathbf{x},\mathbf{u}_2)\mathbf{u}_2+\ldots+(\mathbf{x},\mathbf{u}_r)\mathbf{u}_r.$$

2. 对任意$\mathbf{y}\in W$,

$$|\mathbf{x}-\mathbf{y}|\geq|\mathbf{x}-P\mathbf{x}|.$$

等号成立当且仅当$\mathbf{y}=P\mathbf{x}$.

下表就一些重要性质给出了欧氏空间和酉空间之间的对比.

<table>
<tr><th>欧氏空间</th><th>酉空间</th></tr>
<tr><td>内积$(\mathbf{a},\mathbf{b})$为实数, 满足
$(\mathbf{a},\mathbf{b})=(\mathbf{b},\mathbf{a})$,
$(\lambda\mathbf{a},\mathbf{b})=(\mathbf{a},\lambda\mathbf{b})=\lambda(\mathbf{a},\mathbf{b})$</td><td>内积$(\mathbf{a},\mathbf{b})$为复数, 满足
$(\mathbf{a},\mathbf{b})=\overline{(\mathbf{b},\mathbf{a})}$,
$(\lambda\mathbf{a},\mathbf{b})=\overline{\lambda}(\mathbf{a},\mathbf{b}), (\mathbf{a},\lambda\mathbf{b})=\lambda(\mathbf{a},\mathbf{b})$</td></tr>
<tr><td>模$|\mathbf{a}|=\sqrt{(\mathbf{a},\mathbf{a})}\geq 0$
向量$\mathbf{a}$的单位化向量$\frac{\mathbf{a}}{|\mathbf{a}|}$</td><td>模$|\mathbf{a}|=\sqrt{(\mathbf{a},\mathbf{a})}\geq 0$
向量$\mathbf{a}$的单位化向量$\frac{\mathbf{a}}{|\mathbf{a}|}$</td></tr>
<tr><td>Cauchy-Schwarz不等式
$(\mathbf{a},\mathbf{b})^2\leq(\mathbf{a},\mathbf{a})(\mathbf{b},\mathbf{b})$
即$|(\mathbf{a},\mathbf{b})|\leq|\mathbf{a}||\mathbf{b}|$,
当且仅当$\mathbf{a},\mathbf{b}$线性相关时等号成立</td><td>Cauchy-Schwarz不等式
$(\mathbf{a},\mathbf{b})\overline{(\mathbf{a},\mathbf{b})}\leq(\mathbf{a},\mathbf{a})(\mathbf{b},\mathbf{b})$
即$|(\mathbf{a},\mathbf{b})|\leq|\mathbf{a}||\mathbf{b}|$,
当且仅当$\mathbf{a},\mathbf{b}$线性相关时等号成立</td></tr>
<tr><td>非零向量$\mathbf{a},\mathbf{b}$的夹角$\varphi=\arccos\frac{(\mathbf{a},\mathbf{b})}{|\mathbf{a}||\mathbf{b}|}$</td><td>无定义</td></tr>
<tr><td>勾股定理: 即$(\mathbf{a},\mathbf{b})=0\Leftrightarrow$
$|\mathbf{a}+\mathbf{b}|^2=|\mathbf{a}|^{\mathbf{2}}+|\mathbf{b}|^{\mathbf{2}}$</td><td>勾股定理: 即$(\mathbf{a},\mathbf{b})=0\Rightarrow$
$|\mathbf{a}+\mathbf{b}|^2=|\mathbf{a}|^{\mathbf{2}}+|\mathbf{b}|^{\mathbf{2}}$</td></tr>
<tr><td>三角不等式成立
$|\mathbf{a}+\mathbf{b}|\leq|\mathbf{a}|+|\mathbf{b}|$</td><td>三角不等式成立
$|\mathbf{a}+\mathbf{b}|\leq|\mathbf{a}|+|\mathbf{b}|$</td></tr>
<tr><td>度量矩阵G为实对称矩阵
设$\mathbf{a}_1,\mathbf{a}_2,\ldots,\mathbf{a}_n$为基,
$\sigma_{ij}=(\mathbf{a}_i,\mathbf{a}_j)=\sigma_{ji}$
度量矩阵$G=(\sigma_{ij})=G^T$</td><td>度量矩阵G为Hermite阵
设$\mathbf{a}_1,\mathbf{a}_2,\ldots,\mathbf{a}_n$为基,
$\sigma_{ij}=(\mathbf{a}_i,\mathbf{a}_j)=\overline{(\mathbf{a}_j,\mathbf{a}_i)}=\overline{\sigma_{ji}}$
度量矩阵$H=(\sigma_{ij})=H^*$</td></tr>
<tr><td>用Schmidt方法可将任一组基改造
为标准正交基$\mathbf{u}_1,\mathbf{u}_2,\ldots,\mathbf{u}_n$
$(\mathbf{u}_i,\mathbf{u}_j)=\delta_{ij},\ i,j=1,2,\ldots,n$</td><td>用Schmidt方法可将任一组基改造
为标准正交基$\mathbf{u}_1,\mathbf{u}_2,\ldots,\mathbf{u}_n$
$(\mathbf{u}_i,\mathbf{u}_j)=\delta_{ij},\ i,j=1,2,\ldots,n$</td></tr>
<tr><td>在标准正交基下, $(\mathbf{a},\mathbf{b})=\sum\limits_{i=1}^{n}a_ib_i$</td><td>在标准正交基下, $(\mathbf{a},\mathbf{b})=\sum\limits_{i=1}^{n}\overline{a_i}b_i$</td></tr>
</table>

类比于欧氏空间中的正交变换和对称变换, 下面将讨论酉空间中的两类重要变换, 即酉变换和Hermite变换.

§7.5.3 酉矩阵与酉变换*

酉矩阵就是正交方阵在复数域上的推广, 因此它的性质及证明与正交方阵类似.

定义7.5.6. 设U是一个n阶复矩阵, 如果它满足

$$U^*U=UU^*=I, \tag{7.42}$$

则称U为**酉矩阵**.

二阶酉矩阵的一般形式如下

$$A=\begin{pmatrix} a & b \\ -\bar{b}e^{i\theta} & \bar{a}e^{i\theta} \end{pmatrix},\quad |a|^2+|b|^2=1$$

定理7.5.5. *n阶复方阵U是酉矩阵当且仅当U的行(或列)向量构成$\mathbb{C}^n$的一组标准正交基.*

定理7.5.6. *酉矩阵有下列性质:*

(1) 单位阵是酉方阵.

(2) A,B为同阶酉方阵, 则AB也为酉方阵.

(3) 设A是酉方阵, 则$\bar{A}$, A^T与A^{-1}也是酉方阵.

(4) 酉方阵的行列式的模长为1.

(5) 酉方阵的特征值的模长为1.

有了酉矩阵就可以定义$\mathbb{C}^n$上的酉变换.

定义7.5.7. 设$\mathscr{U}$是酉空间$\mathbb{C}^n$上的一个线性变换: $\mathbf{x}\to U\mathbf{x}$, $\mathbf{x}\in\mathbb{C}^n$. 若$U$是酉矩阵, 则称$\mathscr{U}$为酉变换.

定理7.5.7. *设$\mathscr{U}$是酉空间$\mathbb{C}^n$上的线性变换, 则下列命题等价*

(1) $\mathscr{U}$是酉变换;

(2) $\mathscr{U}$保持内积不变, 对任意$\mathbf{x},\mathbf{y}\in\mathbb{C}^n$, $(\mathscr{U}\mathbf{x},\mathscr{U}\mathbf{y})=(\mathbf{x},\mathbf{y})$;

(3) $\mathscr{U}$保持向量的模长不变, 即对任意的$\mathbf{x}\in V$, 有$|\mathscr{U}\mathbf{x}|=|\mathbf{x}|$;

(4) $\mathscr{U}$把酉空间的标准正交基变为标准正交基;

(5) $\mathscr{U}$在标准正交基下的矩阵是酉矩阵.

定理7.5.8. *设$U(\mathbb{C}^n)$是n维酉空间$\mathbb{C}^n$中所有酉变换的全体所形成的集合, 则*

(1) 单位变换$\mathscr{E}\in U(\mathbb{C}^n)$;

(2) 如果$\mathscr{U}_1,\mathscr{U}_2\in U(\mathbb{C}^n)$, 则$\mathscr{U}_1\circ\mathscr{U}_2\in U(\mathbb{C}^n)$;

(3) 如果$\mathscr{U}\in U(\mathbb{C}^n)$, 则$\mathscr{U}^{-1}\in U(\mathbb{C}^n)$.

即$U(\mathbb{C}^n)$在变换的乘法和逆运算之下是封闭的.

定理7.5.9. 酉方阵均酉相似于对角阵. 即设$A\in\mathbb{C}^{n\times n}$为酉方阵, 则存在酉方阵$U$使得$U^{-1}AU=\mathrm{diag}(e^{i\theta_1},e^{i\theta_2},\ldots,e^{i\theta_n})$, 这里$e^{i\theta_1},e^{i\theta_2},\ldots,e^{i\theta_n}$为$A$的特征值.

定理7.5.10. 设$\mathscr{U}$是酉空间$\mathbb{C}^n$上的线性变换, 则存在$\mathbb{C}^n$的一组标准正交基$\mathbf{u}_1,\mathbf{u}_2,\ldots,\mathbf{u}_n$使得$\mathscr{U}(\mathbf{u}_j)=\lambda_j\mathbf{u}_j$, $j=12,\ldots,n$, 这里$\lambda_j=e^{i\theta_j}$, $j=1,2,\ldots n$是$\mathscr{U}$的特征值.

上述两个定理的证明超出本书范围, 感兴趣的读者可参阅[17].

§7.5.4 Hermite矩阵与Hermite变换*

Hermite矩阵是实对称方阵在复数域上的推广.

定义7.5.8. n阶复方阵H称为Hermite矩阵, 如果H满足$H^*=H$.

例如, $H=\begin{pmatrix}1 & 1+i\\ 1-i & -1\end{pmatrix}$就是Hermite阵. Hermite阵有以下性质.

命题7.5.11. *Hermite*方阵的特征值均为实数.

命题7.5.12. *Hermite*方阵的不同特征值对应的特征向量彼此正交.

上述两个命题的证明同命题 7.3.4, 7.3.5的证明一样. 利用上述命题即可以得到

定理7.5.13. 设H是一个*Hermite*矩阵, 则H酉相似于实对角阵, 即存在酉矩阵U使得

$$U^{-1}HU=\mathrm{diag}(\lambda_1,\lambda_2,\ldots,\lambda_n),$$

其中$\lambda_i(i=1,2,\ldots,n)$是$H$的特征值.

上述定理的证明同定理 7.3.6的证明类似.

例7.5.2. 将Hermite阵

$$H=\begin{pmatrix}0 & i & 1\\ -i & 0 & 0\\ 1 & 0 & 0\end{pmatrix}$$

酉相似对角化.

解. H的特征多项式为$P_H(\lambda)=\lambda(\lambda^2-2)$, 故特征值为$\lambda_1=0,\lambda_2=\sqrt{2},\lambda_2=-\sqrt{2}$.

直接计算, 对应$\lambda_1=0$的单位特征向量为$\mathbf{u}_1=\frac{1}{\sqrt{2}}(0,1,-i)^T$; 对应$\lambda_2=\sqrt{2}$的单位特征向量为$\mathbf{u}_2=\frac{1}{2}(\sqrt{2},-i,1)^T$; 对应$\lambda_3=-\sqrt{2}$的单位特征向量为$\mathbf{u}_3=\frac{1}{2}(\sqrt{2},i,-1)^T$.

令

$$U=(\mathbf{u}_1,\mathbf{u}_2,\mathbf{u}_3)=\frac{1}{2}\begin{pmatrix}0 & \sqrt{2} & \sqrt{2}\\ \sqrt{2} & -i & i\\ -\sqrt{2}i & 1 & -1\end{pmatrix},$$

则有

$$U^{-1}HU=\mathrm{diag}(0,\sqrt{2},-\sqrt{2}).$$

□

下面我们介绍Hermite变换.

定义7.5.9. 设$\mathscr{H}$是$\mathbb{C}^n$上的线性变换: $\mathbf{x}\to H\mathbf{x}$, $\mathbf{x}\in\mathbb{C}^n$. 若$H$是Hermite矩阵, 则称$\mathscr{H}$为Hermite变换.

定理7.5.14. *设$\mathscr{H}$是n维酉空间$\mathbb{C}^n$上的线性变换. 则下列命题等价:*

1. *$\mathscr{H}$是Hermite变换;*
2. *对任意$\mathbf{x},\mathbf{y}\in\mathbb{C}^n$, $(\mathscr{H}\mathbf{x},\mathbf{y})=(\mathbf{x},\mathscr{H}\mathbf{y})$;*
3. *$\mathscr{H}$在任一组标准正交基下的矩阵是Hermite方阵.*

证明. $1\Rightarrow 2$.

$$(\mathscr{H}\mathbf{x},\mathbf{y})=(H\mathbf{x})^*\mathbf{y}=\mathbf{x}^*H^*\mathbf{y}=\mathbf{x}H\mathbf{y}=(\mathbf{x},\mathscr{H}\mathbf{x}).$$

$2\Rightarrow 3$.

设Hermite变换$\mathscr{H}$在酉空间的一组标准正交基$\mathbf{u}_1,\mathbf{u}_2,\ldots,\mathbf{u}_n$下的矩阵为$H=(h_{ij})$, 即

$$\mathscr{H}\mathbf{u}_j=\sum_{k=1}^n h_{kj}\mathbf{u}_k,\quad j=1,2,\ldots,n$$

则一方面有

$$(\mathbf{u}_i,\mathscr{H}\mathbf{u}_j)=(\mathbf{u}_i,\sum_{k=1}^n h_{kj}\mathbf{u}_k)=\sum_{k=1}^n h_{kj}(\mathbf{u}_i,\mathbf{u}_k)=h_{ij},$$

另一方面有

$$(\mathscr{H}\mathbf{u}_i,\mathbf{u}_j)=(\sum_{k=1}^n h_{ki}\mathbf{u}_k,\mathbf{u}_j)=\sum_{k=1}^n \overline{h_{ki}}(\mathbf{u}_k,\mathbf{u}_j)=\overline{h_{ji}}.$$

由$(\mathbf{u}_i,\mathscr{H}\mathbf{u}_j)=(\mathscr{H}\mathbf{u}_i,\mathbf{u}_j)$得$h_{ij}=\overline{h_{ji}}, i,j=1,2,\ldots,n$. 于是$H=H^*$.

$3\Rightarrow 1$.

由于自然基为$\mathbb{C}^n$的一组标准正交基，变换$\mathscr{H}$在自然基下的矩阵H为Hermite阵，从而$\mathscr{H}$为Hermite变换.

□

定理7.5.15. *设$\mathscr{H}$是酉空间$\mathbb{C}^n$上的Hermite变换. 则*

1. *$\mathscr{H}$的特征值均为实数;*
2. *$\mathscr{H}$的不同特征值对应的特征向量相互正交;*
3. *存在$\mathbb{C}^n$的一组标准正交基$\mathbf{u}_1,\mathbf{u}_2,\ldots,\mathbf{u}_n$使得$\mathscr{H}\mathbf{u}_i=\lambda_i\mathbf{u}_i,\ i=1,2,\ldots,n,$ 这里$\lambda_1,\lambda_2,\ldots,\lambda_n$是$\mathscr{H}$的特征值.*

最后我们简单提一下如何将酉空间从$\mathbb{C}^n$推广到一般线性空间. 类似于一般欧氏空间, 一般酉空间中的内积需要用性质定义.

定义7.5.10. 设V是复数域上的线性空间, 如果V内任意两个向量α和β都按某一法则对应于一个复数, 记作(α,β), 且满足

(1) **共轭对称性**:即对任意两个向量$\alpha,\beta\in V$, 有

$$(\alpha,\beta)=\overline{(\beta,\alpha)}. \tag{7.43}$$

(2) **共轭线性性**:即对任意一个复数λ和任意三个向量$\alpha,\beta,\gamma\in V$, 有

$$(\alpha,\lambda\beta)=\lambda(\alpha,\beta),\quad (\alpha,\beta+\gamma)=(\alpha,\beta)+(\alpha,\gamma) \tag{7.44}$$

$$(\lambda\alpha,\beta)=\overline{\lambda}(\alpha,\beta),\quad (\alpha+\beta,\gamma)=(\alpha,\gamma)+(\beta,\gamma) \tag{7.45}$$

(3) **正定性**:即对于任意一个向量$\alpha\in V$, 有$(\alpha,\alpha)\ge 0$, 等号成立当且仅当$\alpha=0$.

则称(α,β)为α和β的**内积**, 定义了内积的复数域**C**上的线性空间V称为**酉空间**.

我们给出一个一般酉空间的例子.

例7.5.3. 用$C[a,b]$表示区间$[a,b]$上连续复函数全体构成的线性空间. 在$C[a,b]$上定义内积

$$(f(x),g(x)):=\int_a^b\overline{f(x)}g(x)\,dx$$

容易验证, 上述内积满足共轭对称性及共轭线性性. 另一方面,

$$(f(x),f(x))=\int_a^b |f(x)|^2\,dx \geq 0,$$

且等号成立当且仅当$f(x)=0$. 因此, 上述内积满足正定性. 从而$C[a,b]$构成一个酉空间.

设$\alpha_1,\alpha_2,\ldots,\alpha_n$是酉空间$V$的一组基, 对任意向量$\alpha=x_1\alpha_1+x_2\alpha_2+\ldots+x_n\alpha_n$, $\beta=y_1\alpha_1+y_2\alpha+\ldots+y_n\alpha_n$,

$$(\alpha,\beta)=(\sum_{i=1}^n x_i\alpha_i,\sum_{j=1}^n y_j\alpha_j)=\sum_{i=1}^n\sum_{j=1}^n \bar{x}_i y_j(\alpha_i,\alpha_j)=\mathbf{x}^*H\mathbf{y}, \tag{7.46}$$

其中$\mathbf{x}=(x_1,x_2,\ldots,x_n)^T$, $\mathbf{y}=(y_1,y_2,\ldots,y_n)^T$, $H=((\alpha_i,\alpha_j))$. 因此, 给定一组基后, 酉空间中的内积表示与 (7.36)式是一致的.

由于一般酉空间中的内积与$\mathbb{C}^n$的内积性质一样, 有关结论也基本一样, 在此不再赘叙.

习题六

1. 已知$\alpha_1=(1,2,-1,1)$, $\alpha_2=(2,3,1,-1)$, $\alpha_3=(-1,-1,-2,2)$.

   (1) 求$\alpha_1,\alpha_2,\alpha_3$的长度及彼此间的夹角.

   (2) 求与$\alpha_1,\alpha_2,\alpha_3$都正交的向量.

2. 设$\mathbf{x},\mathbf{y},\mathbf{z}$是欧氏空间$\mathbb{R}^n$中的向量, 证明以下不等式:

   (1) $|\mathbf{x}-\mathbf{y}|\geq|\mathbf{x}|-|\mathbf{y}|$;

   (2) $|\mathbf{x}-\mathbf{y}|+|\mathbf{y}-\mathbf{z}|\geq|\mathbf{x}-\mathbf{z}|$.

3. 设$\mathbf{x},\mathbf{y}$是欧氏空间$\mathbb{R}^n$的两个向量, 它们之间的夹角为θ. 证明:

   (1) (余弦定理) $|\mathbf{x}-\mathbf{y}|^\mathbf{2}=|\mathbf{x}|^\mathbf{2}+|\mathbf{y}|^\mathbf{2}-\mathbf{2}|\mathbf{x}||\mathbf{y}|\cos\theta$.

   (2) (平行四边形定理) $|\mathbf{x}+\mathbf{y}|^\mathbf{2}+|\mathbf{x}-\mathbf{y}|^\mathbf{2}=\mathbf{2}(|\mathbf{x}|^\mathbf{2}+|\mathbf{y}|^\mathbf{2})$.

   (3) (菱形对角线定理) 若$|\mathbf{x}|=|\mathbf{y}|$, 则$(\mathbf{x}+\mathbf{y})\perp(\mathbf{x}-\mathbf{y})$.

4. 设$\mathbf{a}_1,\mathbf{a}_2,\ldots,\mathbf{a}_n$为$\mathbb{R}^n$的一组基, $\mathbf{a}\in\mathbb{R}^\mathbf{n}$. 证明: $\mathbf{a}=\mathbf{0}$当且仅当$(\mathbf{a},\mathbf{a}_i)=0$, $i=1,2,\ldots,n$.

5. 用Schmidt正交化方法构造标准正交向量组:

(1) $(0,0,1),(0,1,1),(1,1,1)$;

(2) $(1,1,1,2),(1,1,-5,3),(3,2,8,-7)$.

6. 设在$\mathbb{R}^3$中, 基$\mathbf{a}_1,\mathbf{a}_2,\mathbf{a}_3$的度量矩阵是

$$\begin{pmatrix} 1 & 0 & -1 \\ 0 & 2 & 0 \\ -1 & 0 & 2 \end{pmatrix}$$

试求$\mathbb{R}^3$中由$\mathbf{a}_1,\mathbf{a}_2,\mathbf{a}_3$表示的一组标准正交基.

7. 证明:n维向量空间中任一个正交向量组都能扩充成一组正交基.

8. 验证下列各组向量是正交的, 并添加向量改造为标准正交基:

(1) $(2,1,2),(1,2,-2)$;

(2) $(1,1,1,2),(1,2,3,-3)$.

9. 如$\mathbf{b}$与$\mathbf{a}_1,\mathbf{a}_2,\ldots,\mathbf{a}_n$都正交, 证明:$\mathbf{b}$与$\mathbf{a}_1,\mathbf{a}_2,\ldots,\mathbf{a}_n$的任意线性组合也正交.

10. 设$\mathbf{e}_1,\mathbf{e}_2,\mathbf{e}_3$是$\mathbb{R}^3$的一组标准正交基. 令

$$\begin{aligned} \mathbf{a}_1 &= \frac{1}{3}(2\mathbf{e}_1+2\mathbf{e}_2-\mathbf{e}_3), \\ \mathbf{a}_2 &= \frac{1}{3}(2\mathbf{e}_1-\mathbf{e}_2+2\mathbf{e}_3), \\ \mathbf{a}_3 &= \frac{1}{3}(\mathbf{e}_1-2\mathbf{e}_2-2\mathbf{e}_3), \end{aligned}$$

证明:$\mathbf{a}_1,\mathbf{a}_2,\mathbf{a}_3$也是$\mathbb{R}^3$的一组标准正交基.

11. 设$\mathbf{e}_1,\mathbf{e}_2,\ldots,\mathbf{e}_n$是$\mathbb{R}^n$的标准正交基, $\mathbf{x}_1,\mathbf{x}_2,\ldots,\mathbf{x}_k$是$\mathbb{R}^n$中任意$k$个向量, 试证:$\mathbf{x}_1,\mathbf{x}_2,\ldots,\mathbf{x}_k$两两正交的充分必要条件是

$$\sum_{s=1}^{n}(\mathbf{x}_i,\mathbf{e}_s)(\mathbf{x}_j,\mathbf{e}_s)=0, \quad i,j=1,2,\ldots,k,\ i\neq j$$

12. 设$\mathbf{a}_1,\mathbf{a}_2,\ldots,\mathbf{a}_n$是$\mathbb{R}^n$的标准正交基. 证明:

(1) 对任意$\mathbf{a},\mathbf{b}\in\mathbb{R}^n$, $(\mathbf{a},\mathbf{b})=\sum_{i=1}^n(\mathbf{a},\mathbf{a}_i)(\mathbf{b},\mathbf{a}_i)$.

(2) 对任意$\mathbf{a}\in\mathbb{R}^n$, $|\mathbf{a}|^2=\sum_{i=1}^n(\mathbf{a},\mathbf{a}_i)^2$.

13. 证明二阶正交矩阵必取下面两种形式之一:

$$\begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix}, \quad \begin{pmatrix} \cos\theta & \sin\theta \\ \sin\theta & -\cos\theta \end{pmatrix}, \qquad -\pi \le \theta < \pi$$

14. 写出所有3阶正交矩阵, 它的元素是0或1.

15. 如果一个正交阵中每个元素都是$\frac{1}{4}$或$-\frac{1}{4}$, 这个正交阵是几阶的?

16. 若$\mathbf{a}$是$\mathbb{R}^n$的单位向量, 证明: $Q = I_n - 2\mathbf{a}\mathbf{a}^T$是一个正交阵. 当$\mathbf{a} = \frac{\mathbf{1}}{\sqrt{\mathbf{3}}}(\mathbf{1},\mathbf{1},\mathbf{1})^{\mathbf{T}}$时, 具体求出$Q$.

17. 在什么条件下, 对角矩阵是正交矩阵?

18. 设A,B都为n阶正交阵, 证明:

    (1) A的伴随矩阵A^*也是正交阵;

    (2) AB也为正交阵;

    (3) A^{-1}也为正交阵.

    (4) A的行列式为1或-1.

19. 给定三阶矩阵$A = \begin{pmatrix} 1 & 2 & 0 \\ 0 & 1 & 2 \\ 2 & 0 & 1 \end{pmatrix}$. 求$A$的QR分解.

20. 设W是由$(1,1,0)$, $(1,-2,1)$生成的$\mathbb{R}^3$的子空间.

    (1) 求W的一组标准正交基.

    (2) 求$W^{\perp}$的标准正交基.

    (3) 求向量$(0,0,2)$在W的正交投影.

21. 在三维空间直角坐标系中, 已知点A,B,C,D的坐标分别是 $(1,1,0)$, $(3,1,2)$, $(0,1,3)$, $(2,2,4)$. 求四面体的顶点A 到底面BCD的高.

22. 求线性方程组

$$\begin{cases} 2x - y = 1 \\ x + 2y = 1 \\ x + y = 1 \end{cases}$$

    的最小二乘解.

23. 证明: 三维空间中绕过原点的某条直线的旋转变换是正交变换.

24. 设$\mathbf{a}$为$\mathbb{R}^n$中的单位向量. 定义线性变换: $\mathscr{A}\mathbf{b}=\mathbf{b}-\mathbf{2}(\mathbf{a},\mathbf{b})\mathbf{a}$. 证明: $\mathscr{A}$为正交变换, 并找$\mathbb{R}^n$ 的一组标准正交基, 使得$\mathscr{A}$在该组基下的矩阵为对角阵.

25. 证明: 第五章习题4, 9给出的线性变换是对称变换.

26. 设

$$A=\begin{pmatrix}1 & -2 & 0\\ -2 & 2 & -2\\ 0 & -2 & 3\end{pmatrix},$$

求正交矩阵P, 使$P^{-1}AP$为对角阵.由此求A^k, k是自然数.

27. 证明:下列三个条件中只要有二个成立, 另一个也必成立.
(1)A是对称的；(2)A是正交的；(3)$A^2=I$.

28. 求正交阵T, 使$T^{-1}AT$为对角矩阵

(1)$A=\begin{pmatrix}2 & 1\\ 1 & 2\end{pmatrix}$ (2)$A=\begin{pmatrix}0 & -1 & -1\\ -1 & 0 & -1\\ -1 & -1 & 0\end{pmatrix}$

(3)$A=\begin{pmatrix}1 & 2 & 4\\ 2 & -2 & 2\\ 4 & 2 & 1\end{pmatrix}$ (4)$A=\begin{pmatrix}3 & -2 & 0\\ -2 & 2 & -2\\ 0 & -2 & 1\end{pmatrix}$

29. 设A为n阶实对称方阵, 且$A^2=A$. 证明: 存在正交方阵T使得$T^{-1}AT=\mathrm{diag}(I_r,O)$. 这里$r=\mathrm{rank}(A)$.

30. 设A为n阶实对称方阵. 证明: $\max_{\mathbf{0}\neq\mathbf{x}\in\mathbb{R}^n}\frac{\mathbf{x}^TA\mathbf{x}}{\mathbf{x}^T\mathbf{x}}=\lambda_{\max}$, 这里$\lambda_{\max}$是$A$的最大特征值.

31* 设V是n阶实方阵全体构成的线性空间. 对$A,B\in V$, 定义$(A,B)=Tr(AB^T)$.

(1) 证明: (A,B)构成内积, 从而V在该内积下成为欧氏空间.

(2) 设$A=(a_{ij})_{n\times n}\in V$, 求$A$的模长.

(3) 基本矩阵E_{ij}是否构成V的一组标准正交基? 请说明理由.

32 设$\alpha_1,\alpha_2,\ldots,\alpha_n$是$n$维欧氏空间$V$的一组向量. 定义其Gram矩阵$G=((\alpha_i,\alpha_j))\in\mathbb{R}^{n\times n}$.

(1) 证明: $\alpha_1,\alpha_2,\ldots,\alpha_n$构成$V$的一组基当且仅当$\det(G)\neq 0$.

(2) 设α_i在一组标准正交基下的坐标为$\mathbf{x}_i$, $i=1,2,\ldots,n$. 记$X=(\mathbf{x}_1,\ldots,\mathbf{x}_n)\in\mathbb{R}^{n\times n}$. 则$\det(G)=\det(X)^2$.

33 设$\mathbb{R}_n[x]$是次数不超过n次的实系数多项式全体构在内积$(f(x),g(x))=\int_{-1}^{1}f(x)g(x)\,dx$下的欧氏空间. 令

$$P_0(x)=1,\ P_1(x)=x,\ P_n(x)=\frac{2n+1}{n+1}xP_{n-1}(x)-\frac{n}{n+1}P_{n-2}(x),\quad n\geq 2.$$

证明: $P_0(x),P_1(x),\ldots,P_n(x)$是$\mathbb{R}_n[x]$的一组正交基. 称$P_n(x)$为**Legendre多项式**.

34 设$\mathbb{R}_n[x]$是次数不超过n次的实系数多项式全体构在内积$(f(x),g(x))=\int_{-1}^{1}\frac{f(x)g(x)}{\sqrt{1-x^2}}\,dx$下的欧氏空间. 令

$$T_0(x)=1,\ T_1(x)=x,\ T_n(x)=2xT_{n-1}(x)-T_{n-2}(x),\quad n\geq 2.$$

证明: $T_0(x),T_1(x),\ldots,T_n(x)$是$\mathbb{R}_n[x]$的一组标准正交基. 称$T_n(x)$为**Chebyshev多项式**.

35* 验证

(1) $A=\begin{pmatrix}-(\frac{1}{2}+\frac{\sqrt{2}}{4})i & \frac{1}{2}i & -(\frac{1}{2}-\frac{\sqrt{2}}{4})i\\ \frac{1}{2}i & \frac{\sqrt{2}}{2}i & -\frac{1}{2}i\\ (\frac{1}{2}-\frac{\sqrt{2}}{4})i & \frac{1}{2}i & (\frac{1}{2}+\frac{\sqrt{2}}{4})i\end{pmatrix}$ 为酉阵;

(2) $B=\begin{pmatrix}1 & 1+i & i\\ 1-i & 2 & -i\\ -i & i & 0\end{pmatrix}$为Hermite阵.

36* 证明:酉变换的特征值的模等于1.

37* 证明:

(1) 酉阵的不同特征值所对应的特征向量是两两正交的;

(2) Hermite阵的不同特征值所对应的特征向量是两两正交的.

38* 证明:

(1) 设A为可逆Hermite阵, 则A^{-1}也为Hermite阵;

(2) 设A,B都是n阶Hermite阵, 那么AB也是Hermite阵的充分必要条件是$AB=BA$.

39* 设A为Hermite阵, 证明:

(1) $I \pm iA$为可逆阵;

(2) $I+iA$与$I-iA$乘法可交换;

(3) $U=(I+iA)(I-iA)^{-1}$为酉阵.

40* 若$AA^H=0$, 证明:$A=0$.

41* 设U为酉阵, 且$U^{-1}AU=B$, 证明:

$$\operatorname{tr}(A^HA)=\operatorname{tr}(B^HB).$$

42* 已知Hermite阵

$$H=\begin{pmatrix} 0 & -i & 1 \\ i & 0 & i \\ 1 & -i & 0 \end{pmatrix}$$

求酉方阵U使得$U^{-1}HU$为对角阵.

第八章　实二次型

在各种工程与实际应用中,广泛地存在处理多元二次多项式的问题.例如,数据拟合的最小二乘法导致一个多元二次多项式的极小值问题.再如,二次曲线(或二次曲面)的分类问题就是要确定多元二次方程所决定的几何图形的类型与形状.还有其它许多问题,如多元二次多项式的正性(非负性)的判别,等等.幸运地,线性代数的工具对上述这些问题给出了一个简单有效的处理方法.

要处理多元二次多项式,关键是要处理其二次项.其二次项的全体组成一个二次齐次多项式,我们称之为**二次型**.本章的目的就是要利用线性代数的工具研究二次型的基本理论、性质、简化及其应用.本章只讨论系数为实数的二次型—**实二次型**.

§8.1　二次型的矩阵表示

定义8.1.1. 在实数域上,一个含n个变元$x_1,x_2,\ldots,x_n$的**二次型** $Q(x_1,x_2,\ldots,x_n)$是一个齐次的二次多项式

$$Q(x_1,x_2,\ldots,x_n)=\sum_{i=1}^{n}\sum_{j=1}^{n}a_{ij}x_ix_j. \tag{8.1}$$

这里$a_{ji}=a_{ij}$, $i,j=1,\ldots,n$为实系数.利用矩阵乘法运算,上式可以表示成

$$Q(x_1,x_2,\ldots,x_n)=\mathbf{x}^TA\mathbf{x}, \tag{8.2}$$

这里A为n阶实对称方阵

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

而$\mathbf{x}$为n维列向量

$$\mathbf{x}=(x_1,x_2,\ldots,x_n)^T.$$

称实对称矩阵A为二次型$Q(x_1,x_2,\ldots,x_n)$的**矩阵**, A的秩称为**二次型的秩**.

每一个二次型通过(8.2)与一个实对称矩阵建立了一一对应关系,而这正是利用矩阵研究二次型的基础.例如,

$$Q(x_1,x_2,x_3)=x_1x_2+x_2x_3+x_1x_3=(x_1,x_2,x_3)\begin{pmatrix} 0 & \frac12 & \frac12\\ \frac12 & 0 & \frac12\\ \frac12 & \frac12 & 0\end{pmatrix}\begin{pmatrix} x_1\\ x_2\\ x_3\end{pmatrix}$$

对应的实对称矩阵就是

$$A = \begin{pmatrix} 0 & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & 0 & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} & 0 \end{pmatrix},$$

而一个只含平方项的二次型

$$Q(x_1,\ldots,x_n) = \lambda_1 x_1^2 + \lambda_2 x_2^2 + \cdots + \lambda_n x_n^2$$

对应的实对称矩阵是对角矩阵

$$A = \mathrm{diag}(\lambda_1,\lambda_2,\ldots,\lambda_n) = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix}.$$

二次型的一个基本问题是, 如何对变量做可逆(非退化)的线性变换, 使得变换后的二次型尽量简单. 具体来说, 给定二次型$Q(\mathbf{x}) = \mathbf{x}^T A\mathbf{x}$, 对变量$\mathbf{x}$做可逆线性变换$\mathbf{x} = P\mathbf{y}$, 其中$P$是可逆矩阵, 则变换后的二次型为

$$\tilde{Q}(y_1,y_2,\ldots,y_n) := \tilde{Q}(\mathbf{y}) = Q(\mathbf{x})|_{\mathbf{x}=P\mathbf{y}} = (P\mathbf{y})^T A(P\mathbf{y}) = \mathbf{y}^T(P^T AP)\mathbf{y} := \mathbf{y}^T B\mathbf{y}, \tag{8.3}$$

这里$B = P^T AP$. 我们的目标是, 对给定的对称矩阵A, 找合适的可逆方阵P使得$B = P^T AP$尽量简单.

显然矩阵B也是对称的.而矩阵$B = P^T AP$给出了变换前后二次型对应矩阵之间的关系, 这种关系就是所谓的相合关系.

定义8.1.2. 对于实数域上两个n阶矩阵A, B, 如果存在一个可逆的实矩阵P使得

$$B = P^T AP, \tag{8.4}$$

则称A和B是**相合的**, 或者说B**相合于**A, 矩阵P称为**相合变换矩阵**.

注意到矩阵之间的相合关系是一种等价关系, 即它具有性质

(1) 反身性:$A = I^T AI$, 这里I是n阶单位矩阵, 即自身与自身相合;

(2) 对称性:如果$B = P^T AP$, 则$A = (P^{-1})^T BP^{-1}$, 即如果B相合于A, 则A相合于B;

(3) 传递性:如果$C = Q^T BQ, B = P^T AP$, 则$C = (PQ)^T A(PQ)$, 即如果C相合于B, B相合于A, 则C相合于A.

根据矩阵之间相合这种等价关系, 可以将所有n阶实对称矩阵进行分类, 即同一类中的矩阵是彼此相合的, 不同类的矩阵之间彼此不相合.关于矩阵相合, 我们有以下基本问题

1. 两个矩阵相合的充要条件是什么?
2. 在每一个相合类中, 找到一个最简单的矩阵作为代表元, 即所谓的相合标准形.

由于实对称矩阵与实二次型之间是一一对应的, 因此, 实对称矩阵在相合下的分类与化简, 对应了二次型的分类与化简.

§8.2 二次型的标准形

给定一个二次型, 我们希望做一个可逆线性变换使得变换使得变换后的二次型尽量简单. 这里尽量简单的意思是, 二次型中尽量不要出现交叉项, 只有平方项. 下面介绍二次型简化的*配平方法*. 我们以两个例子来说明.

例8.2.1. 化三元的二次型$Q(x_1,x_2,x_3)=2x_1^2+x_1x_2+x_3^2$为尽量简单的形式.

解. 对变量x_1配方得

$$Q(x_1,x_2,x_3)=2(x_1+\tfrac{1}{4}x_2)^2-\tfrac{1}{8}x_2^2+x_3^2.$$

令

$$\begin{cases} y_1=x_1+\frac{1}{4}x_2 \\ y_2=x_2 \\ y_3=x_3 \end{cases} \quad 或 \quad \begin{cases} x_1=y_1-\frac{1}{4}y_2 \\ x_2=y_2 \\ x_3=y_3 \end{cases}$$

便有

$$\tilde{Q}(y_1,y_2,y_3)=Q(x_1,x_2,x_3)\Big|_{\mathbf{x}=P\mathbf{y}}=2y_1^2-\tfrac{1}{8}y_2^2+y_3^2.$$

用矩阵语言表述就是, 矩阵

$$A=\begin{pmatrix} 2 & \frac{1}{2} & 0 \\ \frac{1}{2} & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

经过相合变换阵

$$P=\begin{pmatrix} 1 & -\frac{1}{4} & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

做相合变换得到对角阵

$$P^TAP=\begin{pmatrix}2&0&0\\0&-\frac{1}{8}&0\\0&0&1\end{pmatrix}.$$

□

上述例子中, 化简后的二次型$\tilde{Q}(y_1,y_2,y_3)$没有交叉项.

例8.2.2. 化三元的二次型$Q(x_1,x_2,x_3)=2x_1x_2-6x_2x_3+2x_1x_3$为尽量简单的形式.

解. 做非退化变换

$$\begin{cases}x_1=y_1+y_2\\x_2=y_1-y_2\\x_3=y_3\end{cases}$$

得

$$Q(x_1,x_2,x_3)=\tilde{Q}(y_1,y_2,y_3)=2y_1^2-2y_2^2-4y_1y_3+8y_2y_3.$$

从而消去x_1和x_2的交叉项.再利用配方消去其他的交叉项, 即令

$$\begin{cases}y_1=z_1+z_3\\y_2=z_2+2z_3\\y_3=z_3\end{cases}$$

则

$$Q(x_1,x_2,x_3)=2z_1^2-2z_2^2+6z_3^2.$$

综合两次变换得

$$\mathbf{x}=P\mathbf{z},\quad P=\begin{pmatrix}1&1&3\\1&-1&-1\\0&0&1\end{pmatrix}$$

满足

$$P^TAP=\begin{pmatrix}2&0&0\\0&-2&0\\0&0&6\end{pmatrix},$$

其中

$$A=\begin{pmatrix}0&1&1\\1&0&-3\\1&-3&0\end{pmatrix}.$$

□

上述例子并不是特例. 实际上我们有

定理8.2.1. 任何一个由 (8.1)式定义的实二次型$Q(x_1,x_2,\ldots,x_n)$ 均可通过配平方法找到可逆变换$\mathbf{x}=P\mathbf{y}$, 将它化为

$$\tilde{Q}(y_1,y_2,\ldots,y_n):=Q(x_1,x_2,\ldots,x_n)\Big|_{\mathbf{x}=P\mathbf{y}}=\mu_1y_1^2+\mu_2y_2^2+\cdots+\mu_ny_n^2. \tag{8.5}$$

称$\tilde{Q}(y_1,\ldots,y_n)$为二次型$Q(x_1,\ldots,x_n)$的**标准型**.

证明(配平方法). 对变量的个数作归纳法.对于$n=1$, $Q(x_1)=a_{11}x_1^2$, 结论是显然的. 假设定理对$n-1$元的二次型成立, 则对于n元的情形, 分三种情形讨论.

(1) $a_{11}\neq 0$.如果$a_{11}=0$, 但存在某个$a_{ii}\neq 0$, 只需将x_1和x_i对换, 使二次型成为上述形式即可.这时

$$\begin{aligned}
Q(x_1,\ldots,x_n)&=a_{11}x_1^2+2\sum_{j=2}^n a_{1j}x_1x_j+\sum_{i,j=2}^n a_{ij}x_ix_j\\
&=a_{11}\left(x_1+\sum_{j=2}^n a_{1j}a_{11}^{-1}x_j\right)^2-a_{11}^{-1}\left(\sum_{j=2}^n a_{1j}x_j\right)^2+\sum_{i,j=2}^n a_{ij}x_ix_j\\
&=a_{11}\left(x_1+\sum_{j=2}^n a_{1j}a_{11}^{-1}x_j\right)^2+\sum_{i,j=2}^n b_{ij}x_ix_j.
\end{aligned}$$

上式中

$$\sum_{i,j=2}^n b_{ij}x_ix_j=-a_{11}^{-1}\left(\sum_{j=2}^n a_{1j}x_j\right)^2+\sum_{i,j=2}^n a_{ij}x_ix_j$$

是一个含$n-1$个变元$x_2,x_3,\ldots,x_n$的实二次型.令

$$\begin{cases} y_1=x_1+\sum\limits_{j=2}^n a_{1j}a_{11}^{-1}x_j\\ y_2=x_2\\ \cdots\cdots\\ y_n=x_n\end{cases} \quad 或 \quad \begin{cases} x_1=y_1-\sum\limits_{j=2}^n a_{1j}a_{11}^{-1}y_j\\ x_2=y_2\\ \cdots\cdots\\ x_n=y_n\end{cases}$$

其对应的变换矩阵是

$$S=\begin{pmatrix}1 & -a_{12}a_{11}^{-1} & \cdots & -a_{1n}a_{11}^{-1}\\ 0 & 1 & \cdots & 0\\ \vdots & \vdots & \ddots & \vdots\\ 0 & 0 & \cdots & 1\end{pmatrix}.$$

经过变换$\mathbf{x}=S\mathbf{y}$, 原二次型化为

$$Q(x_1,\dots,x_n)=a_{11}y_1^2+\sum_{i,j=2}^{n}b_{ij}y_iy_j.$$

由归纳假设, 对于$n-1$个元的二次型$\sum_{i,j=2}^{n}b_{ij}y_iy_j$, 存在一个可逆的线性变换$\mathbf{y}=T\mathbf{z}$使得

$$\sum_{i,j=2}^{n}b_{ij}y_iy_j=\sum_{i=2}^{n}\mu_iz_i^2.$$

令$z_1=y_1$, $\mu_1=1$, 则原二次型化为标准形$\sum_{i=1}^{n}\mu_iz_i^2$.相应的变换为$\mathbf{x}=P\mathbf{z}$, 这里

$$P=S\begin{pmatrix}1&0\\0&T\end{pmatrix}.$$

(2) 如果所有的$a_{ii}=0$, $i=1,2,\dots,n$, 但至少有一个$a_{1j}\neq 0$, $j>1$.不妨设$a_{12}\neq 0$, 令

$$\begin{cases}x_1=y_1-y_2\\x_2=y_1+y_2\\x_3=y_3\\\cdots\cdots\\x_n=y_n\end{cases}$$

它是可逆的线性变换, 对应的变换矩阵为

$$P=\begin{pmatrix}1&-1&0&\cdots&0\\1&1&0&\cdots&0\\0&0&1&\cdots&0\\\vdots&\vdots&\vdots&\ddots&\vdots\\0&0&0&\cdots&1\end{pmatrix}$$

于是

$$\begin{aligned}Q(x_1,\dots,x_n)&=2a_{12}x_1x_2+\cdots=2a_{12}(y_1^2-y_2^2)+\cdots\\&=2a_{12}y_1^2-2a_{12}y_2^2+\cdots\end{aligned}$$

化为关于变元$y_1,y_2,\dots,y_n$的二次型, 且平方项y_1^2的系数不为零, 从而化成情形(1).

(3) 如果对所有的$a_{ii}=0, i=1,2,\ldots,n$ 及$a_{1j}=0, j=1,2,\ldots,n$, 则由对称性, 得$a_{j1}=0, j=1,2,\ldots,n$. 因此二次型简化为

$$Q(x_1,\ldots,x_n)=\sum_{i,j=2}^{n}a_{ij}x_ix_j.$$

这是一个含有$n-1$个变元的二次型, 由归纳假设, 它经过线性变换可以化为标准形.

总之, 对于任一种形式的实二次型都存在可逆的线性变换将其化为标准形. □

上述结论可以用矩阵语言表示如下.

定理8.2.2. *每一个实对称矩阵A都相合于对角阵. 具体来说, 存在初等矩阵$P_1,P_2,\ldots,P_r$, 使得*

$$P_r^TP_{r-1}^T\cdots P_1^TAP_1P_2\cdots P_r=\operatorname{diag}(\mu_1,\mu_2,\ldots,\mu_n). \tag{8.6}$$

这里$\mu_1,\mu_2,\ldots,\mu_n$为实数.

证明. 由定理8.2.1, 结论显然.下面我们用矩阵语言将证明重叙.

对矩阵A的阶数n进行归纳.当$n=1$时, 结论是显然的.假设对于$n-1$阶对称矩阵, 定理已成立, 则对n阶对称矩阵$A=(a_{ij})_{n\times n}$, 分如下情形讨论

(1) 如果$a_{11}\neq 0$, 做"把第1列乘$-a_{i1}a_{11}^{-1}$加到第i列"的初等变换, 同时做"把第1行乘$-a_{i1}a_{11}^{-1}$加到第i行"的初等变换($i=2,3,\ldots,n$), 由于A是对称矩阵, 经过上述变换, 矩阵A化为$\operatorname{diag}(a_{11},A_{n-1})$, 即存在可逆矩阵$P_1$使得

$$P_1^TAP_1=\begin{pmatrix}a_{11}&0\\0&A_{n-1}\end{pmatrix}.$$

这里A_{n-1}是一个$n-1$阶的实对称矩阵.根据归纳假设, A_{n-1}可用相合初等变换阵化成对角矩阵, 即存在可逆方阵P_2使得$P_2^TA_{n-1}P_2=\Lambda$, Λ为对角阵.令$P=P_1\operatorname{diag}(1,P_2)$, 则

$$P^TAP=\begin{pmatrix}a_{11}&0\\0&\Lambda\end{pmatrix}.$$

(2) 当$a_{11}=0$时, 如果有一个$a_{ii}\neq 0$, $i=2,3,\ldots,n$, 则只要把A的第一行与第i行互换, 再把第1列和第i列互换, 就成第一种情形.这个过程就是对A作相合变换$S_{1i}AS_{1i}$.

(3) 当$a_{ii}=0, i=1,2,\ldots,n$时, 即A的所有对角线上的元素都等于零, 则第一行的其他元素至少有一个不等于零.不妨设$a_{i1}\neq 0$, 则做初等相合变换

$$T_{1i}(1)AT_{1i}(1)^T=\begin{pmatrix}2a_{i1} & C\\ C^T & A_{n-1}\end{pmatrix},$$

于是问题又归结到第一种情形.这样就完成了定理的证明.

□

由定理 8.2.2可知, 对矩阵施行一个初等行变换, 同时要对矩阵作一次相应的列变换, 以保证每次变换后得到的矩阵和原矩阵相合. 为了求得变换矩阵P, 须将每次所作的变换保留下来.具体作法是, 将单位阵放在待变换矩阵下面, 构成$2n\times n$矩阵$(A,I)^T$, 对该矩阵每作一次列变换, 同时仅对A作一次相应的行变换, 即

$$\begin{pmatrix}A\\ I\end{pmatrix}\xrightarrow[\text{对}I\text{只作初等列变换}]{\text{对}A\text{作成对的初等行列变换}}\begin{pmatrix}P^TAP\\ P\end{pmatrix}$$

当P^TAP是对角阵时, I就成了P.当然, 也可对矩阵(A,I)实施类似的变换, 但对I只做行变换.

例8.2.3. 用初等变换法将二次型

$$Q(x_1,x_2,x_3)=x_1^2-x_2^2+x_3^2+4x_1x_3+2x_2x_3$$

化成标准形.

解. 二次型Q的矩阵是

$$A=\begin{pmatrix}1 & 0 & 2\\ 0 & -1 & 1\\ 2 & 1 & 1\end{pmatrix}.$$

对A做初等相合变换化为对角阵

$$\begin{pmatrix}A\\ I\end{pmatrix}=\begin{pmatrix}1 & 0 & 2\\ 0 & -1 & 1\\ 2 & 1 & 1\\ 1 & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & 1\end{pmatrix}\xrightarrow{-2c_1\to c_3}\begin{pmatrix}1 & 0 & 0\\ 0 & -1 & 1\\ 2 & 1 & -3\\ 1 & 0 & -2\\ 0 & 1 & 0\\ 0 & 0 & 1\end{pmatrix}\xrightarrow{-2r_1\to r_3}\begin{pmatrix}1 & 0 & 0\\ 0 & -1 & 1\\ 0 & 1 & -3\\ 1 & 0 & -2\\ 0 & 1 & 0\\ 0 & 0 & 1\end{pmatrix}$$

$$\xrightarrow{c_2\to c_3}\begin{pmatrix}1&0&0\\0&-1&0\\0&1&-2\\1&0&-2\\0&1&1\\0&0&1\end{pmatrix}\xrightarrow{r_2\to r_3}\begin{pmatrix}1&0&0\\0&-1&0\\0&0&-2\\1&0&-2\\0&1&1\\0&0&1\end{pmatrix}$$

于是原二次型的标准型为

$$\tilde{Q}(y_1,y_2,y_3)=y_1^2-y_2^2-2y_3^2.$$

相应的变换矩阵为

$$P=\begin{pmatrix}1&0&-2\\0&1&1\\0&0&1\end{pmatrix}.$$

□

例8.2.4. 用初等变换法将二次型

$$Q(x_1,x_2,x_3)=2x_1x_2+4x_1x_3$$

化成标准形.

解. 二次型的矩阵

$$A=\begin{pmatrix}0&1&2\\1&0&0\\2&0&0\end{pmatrix}.$$

将A的第二列加到第一列, 再将第二行加到第一行, 就可使矩阵的第一行第一列的元素变为非零. 我们用行的形式来记录P^T.

$$(A,I)=\begin{pmatrix}0&1&2&1&0&0\\1&0&0&0&1&0\\2&0&0&0&0&1\end{pmatrix}\longrightarrow\begin{pmatrix}1&1&2&1&1&0\\1&0&0&0&1&0\\2&0&0&0&0&1\end{pmatrix}\longrightarrow\begin{pmatrix}2&1&2&1&1&0\\1&0&0&0&1&0\\2&0&0&0&0&1\end{pmatrix}$$

$$\longrightarrow\begin{pmatrix}2&1&2&1&1&0\\0&-\frac{1}{2}&-1&-\frac{1}{2}&\frac{1}{2}&0\\0&-1&-2&-1&-1&1\end{pmatrix}\longrightarrow\begin{pmatrix}2&0&0&1&1&0\\0&-\frac{1}{2}&-1&-\frac{1}{2}&\frac{1}{2}&0\\0&-1&-2&-1&-1&1\end{pmatrix}$$

$$\longrightarrow\begin{pmatrix}2&0&0&1&1&0\\0&-\frac{1}{2}&-1&-\frac{1}{2}&\frac{1}{2}&0\\0&0&0&-&-2&1\end{pmatrix}\longrightarrow\begin{pmatrix}2&0&0&1&1&0\\0&-\frac{1}{2}&0&-\frac{1}{2}&\frac{1}{2}&0\\0&0&0&0&-2&1\end{pmatrix}=(D,P^T).$$

令

$$P=\begin{pmatrix}1 & -\frac{1}{2} & 0\\ 1 & \frac{1}{2} & -2\\ 0 & 0 & 1\end{pmatrix},$$

作可逆线性替换$\mathbf{x}=P\mathbf{y}$, 则

$$Q(x_1,x_2,x_3)=\mathbf{y}^T\operatorname{diag}(2,-\frac{1}{2},0)\mathbf{y}=2y_1^2-\frac{1}{2}y_2^2.$$

□

除了配方法与初等变换方法之外, 上一章的定理 7.3.6实际上给出了实对称方阵正交相合于对角阵的方法.

定理8.2.3. 设A为n阶实对称方阵, 则存在正交方阵P使得

$$P^TAP=\operatorname{diag}(\lambda_1,\lambda_2,\ldots,\lambda_n),$$

这里$\lambda_1,\lambda_2,\ldots,\lambda_n$是$A$的特征值.

证明. 由于P为正交阵, 故$P^{-1}=P^T$. 由定理 7.3.6, 结论显然. □

§8.3 相合不变量与分类

在上节中, 我们我们证明了实二次型可以化为标准形, 或者等价地, 实对称矩阵可以通过一系列初等变换使之相合于实对角矩阵. 但二次型的标准型不是唯一的. 实际上, 设

$$\tilde{Q}(y_1,y_2,\ldots,y_n)=\mu_1y_1^2+\mu_2y_2^2+\ldots+\mu_ny_n^2$$

是二次型$Q(x_1,x_2,\ldots,x_n)$的标准型. 做可逆变换

$$y_1=c_1z_1,\, y_2=c_2z_2,\, \ldots,\, y_n=c_nz_n,\quad c_i\neq 0,\, i=1,2,\ldots,n$$

得到新的二次型标准型

$$\tilde{\tilde{Q}}(z_1,z_2,\ldots,z_n)=\mu_1c_1^2z_1^2+\mu_2c_2^2z_2^2+\ldots+\mu_nc_n^2z_n^2.$$

由于$c_i, i=1,2,\ldots,n$的任意性, 因此二次型的标准型不唯一, 或者说一个实矩阵通过不同的相合变换可以化为不同的对角矩阵. 正是如此, 我们要进一步研究刻画二次型在可逆线性变换下本质的性质.

.

定理8.3.1. 设A是一个n阶实对称矩阵, 则存在可逆矩阵P, 使得

$$P^TAP=\begin{pmatrix} I_r & & \\ & -I_s & \\ & & 0 \end{pmatrix}, \quad \mathrm{rank}A=r+s\le n \tag{8.7}$$

其中, r是标准形中正项的项数, s是负项的项数.上式右边的对角矩阵称为矩阵A的**规范形**.等价地, 给定实二次型$Q(x_1,x_2,\dots,x_n)$, 存在可逆方阵P使得

$$Q(x_1,x_2,\dots,x_n)\Big|_{\mathbf{x}=P\mathbf{y}}=y_1^2+\dots+y_r^2-y_{r+1}^2-\dots-y_{r+s}^2. \tag{8.8}$$

上式称为二次型$Q(x_1,x_2,\dots,x_n)$的**规范形**.

证明. 从定理8.2.2知, 存在可逆矩阵P_1使得$P_1^TAP_1$是对角矩阵.将对角矩阵中的对角元分成三类:正的、负的和零.不妨设

$$P_1^TAP_1=\mathrm{diag}(\mu_1,\dots,\mu_r,-\mu_{r+1},\dots,-\mu_{r+s},0,\dots,0),$$

这里$\mu_i>0$, $i=1,2,\dots,r+s$, 否则可通过初等相合变换将对角线上的元素调整至上述情形.令

$$P_2=\mathrm{diag}\left(\frac{1}{\sqrt{\mu_1}},\dots,\frac{1}{\sqrt{\mu_r}},\frac{1}{\sqrt{\mu_{r+1}}},\dots,\frac{1}{\sqrt{\mu_{r+s}}},1,\dots,1\right),$$

则$P=P_1P_2$满足

$$P^TAP=(P_1P_2)^TAP_1P_2=P_2^T(P_1^TAP_1)P_2=\begin{pmatrix} I_r & & \\ & -I_s & \\ & & 0 \end{pmatrix}.$$

□

显然, 相合变换不改变矩阵A的秩, 即实对称矩阵(二次型)的秩是相合变换下的不变量, 因此$\mathrm{rank}A=r+s$.不仅如此, 我们还可以证明, 定理8.3.1中的量r和s, 即二次型的标准形中的正项数和负项数也是相合意义下的不变量.

定理8.3.2. 实二次型

$$Q(x_1,\dots,x_n)=\mathbf{x}^TA\mathbf{x}$$

的规范形中正项数r和负项数s是由二次型$Q(x_1,\dots,x_n)$唯一确定的, 或者说是由实对称矩阵A唯一确定的.

证明. 假如有两种不同的可逆变换$\mathbf{x}=P_1\mathbf{y}$和$\mathbf{x}=P_2\mathbf{z}$, 使得二次型化为下列两种规范形

$$Q(x_1,\ldots,x_n)\Big|_{\mathbf{x}=P_1\mathbf{y}}=y_1^2+\cdots+y_r^2-y_{r+1}^2-\cdots-y_{r+s}^2$$

和

$$Q(x_1,\ldots,x_n)\Big|_{\mathbf{x}=P_2\mathbf{z}}=z_1^2+\cdots+z_p^2-z_{p+1}^2-\cdots-z_{p+q}^2.$$

这里$r+s=p+q=\mathrm{rank}A$.假设$r\neq p$, 不妨设$r<p$. 记$P_1^{-1}=(b_{ij})_{n\times n}$, $P_2^{-1}=(c_{ij})_{n\times n}$. 令

$$\begin{aligned}
y_1&=b_{11}x_1+b_{12}x_2+\cdots+b_{1n}x_n=0\\
&\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\\
y_r&=b_{r1}x_1+b_{r2}x_2+\cdots+b_{rn}x_n=0\\
z_{p+1}&=c_{p+1,1}x_1+c_{p+1,2}x_2+\cdots+c_{p+1,n}x_n=0\\
&\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\\
z_n&=c_{n1}x_1+c_{n2}x_2+\cdots+c_{nn}x_n=0
\end{aligned}$$

则得到一个关于n个变量$x_1,\ldots,x_n$的齐次线性方程组. 方程组中方程的个数是$r+(n-p)=n-(p-r)<n$, 因此有非零解$\mathbf{x}=(x_1,x_2,\ldots,x_n)^T$.这个解对应的

$$\mathbf{y}=P_1^{-1}\mathbf{x}=(0,\ldots,0,y_{r+1},\ldots,y_n)^T,\quad \mathbf{z}=P_2^{-1}\mathbf{x}=(z_1,\ldots,z_p,0,\ldots,0)^T.$$

显然, $y_{r+1},\ldots,y_{r+s},\ldots,y_n$不全为零; $z_1,\ldots,z_p$也不全为零. 于是有

$$\begin{aligned}
Q(x_1,\ldots,x_n)\Big|_{X=P_1Y}&=-y_{r+1}^2-\cdots-y_{r+s}^2\leq 0,\\
Q(x_1,\ldots,x_n)\Big|_{X=P_2Z}&=z_1^2+\cdots+z_p^2>0.
\end{aligned}$$

上面两个式子是相互矛盾的.因此$r=p$, 从而有$s=q$.定理证毕. □

我们称定理8.3.2为**惯性定理**, r为**正惯性指数**, s为**负惯性指数**.它们的和就是二次型的秩(即对应的实对称矩阵A的秩$r+s=\mathrm{rank}A$).它们的差$r-s=r-(\mathrm{rank}A-r)=2r-\mathrm{rank}A$称为二次型(或矩阵$A$)的**符号差**.

定理8.3.2说明, 在相合意义下, 两个实二次型(实对称矩阵)是否属于同一个等价类取决于实二次型(实对称矩阵)的秩以及正惯性指数.只有秩与正惯性指数相等的实二次型(实对称矩阵)才是相互等价的, 它们等价于同一个规范形.这就是实二次型(实对称矩阵)的分类.

§8.4 二次曲线与曲面的分类

二次型的标准型及分类的一个直接应用是将二次曲线与曲面化为标准形, 并由此给出二次曲线与曲面的分类.

给定平面二次曲线的方程

$$a_{11}x^2+2a_{12}xy+a_{22}y^2+2b_1x+2b_2y+c=0. \tag{8.9}$$

为判别它表示的曲线类型, 我们先将它化为标准型.为此, 首先要消去方程中的交叉项xy. 考察方程的二次项

$$a_{11}x^2+2a_{12}xy+a_{22}y^2=(x,y)\begin{pmatrix}a_{11} & a_{12}\\ a_{12} & a_{22}\end{pmatrix}\begin{pmatrix}x\\ y\end{pmatrix}:=(x,y)A\begin{pmatrix}x\\ y\end{pmatrix}.$$

只要找到二阶正交矩阵P使P^TAP为对角阵$\mathrm{diag}(\lambda_1,\lambda_2)$, 则变换

$$\begin{pmatrix}x\\ y\end{pmatrix}=P\begin{pmatrix}x'\\ y'\end{pmatrix}$$

将原方程的交叉项消去, 而且可以保持曲线的形状.

注意到二阶正交矩阵可以具有形式

$$P=\begin{pmatrix}\cos\theta & -\sin\theta\\ \sin\theta & \cos\theta\end{pmatrix}$$

因此, 我们只需要做旋转变换

$$\begin{pmatrix}x\\ y\end{pmatrix}=\begin{pmatrix}\cos\theta & -\sin\theta\\ \sin\theta & \cos\theta\end{pmatrix}\begin{pmatrix}x'\\ y'\end{pmatrix}$$

就可以消去方程中的xy项. 将上式代入原方程得交叉项系数的一半为

$$(a_{22}-a_{11})\sin\theta\cos\theta+a_{12}(\cos^2\theta-\sin^2\theta).$$

令上式为零得

$$\cot 2\theta=\frac{a_{11}-a_{22}}{2a_{12}}.$$

这样原方程变为

$$\lambda_1(x')^2+\lambda_2(y')^2+2b_1'x'+2b_2'y'+c'=0.$$

其中λ_1,λ_2是方阵A的特征值.

λ_1,λ_2中至少有一个非零.不妨设$\lambda_1 \neq 0$. 作坐标轴的平移

$$\tilde{x} = x' + b_1/\lambda_1, \qquad \tilde{y} = y' + b_2/\lambda_2$$

(若$\lambda_2 = 0$, 则令$\tilde{y} = y'$).最后原方程化为标准形式

1. **椭圆型** $(\lambda_1\lambda_2 > 0)$

$$\lambda_1\tilde{x}^2 + \lambda_2\tilde{y}^2 + \lambda_3 = 0.$$

2. **双曲型** $(\lambda_1\lambda_2 < 0)$

$$\lambda_1\tilde{x}^2 + \lambda_2\tilde{y}^2 + \lambda_3 = 0.$$

3. **抛物形** $(\lambda_1 \neq 0, \lambda_2 = 0, \tilde{b}_2, \tilde{c}$中至少一个为零.)

$$\lambda_1\tilde{x}^2 + 2\tilde{b}_2\tilde{y} + \tilde{c} = 0.$$

对于二次曲面, 我们类似处理. 给定二次曲面的方程

$$a_{11}x^2 + a_{22}y^2 + a_{33}z^2 + 2a_{12}xy + 2a_{13}xz + 2a_{23}yz + 2b_1x + 2b_2y + 2b_3z + c = 0. \tag{8.10}$$

记

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{12} & a_{22} & a_{23} \\ a_{13} & a_{23} & a_{33} \end{pmatrix}$$

我们找正交阵P使P^TAP为对角阵$\mathrm{diag}(\lambda_1,\lambda_2,\lambda_3)$, 其中$\lambda_1,\lambda_2,\lambda_3$为$A$的特征值. 于是正交变换

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = P \begin{pmatrix} x' \\ y' \\ z' \end{pmatrix}$$

将原方程的所有交叉项消去. 即原方程变为

$$\lambda_1(x')^2 + \lambda_2(y')^2 + \lambda_3(z')^2 + 2b_1'x' + 2b_2'y' + 2b_3'z' + c' = 0.$$

在作坐标轴的平移后, 原方程可以化成标准形式

- **椭球面型** $(\lambda_1,\lambda_2,\lambda_3 > 0)$

$$\lambda_1\tilde{x}^2 + \lambda_2\tilde{y}^2 + \lambda_3\tilde{z}^2 + \lambda_4 = 0.$$

- **双曲面型** ($\lambda_1, \lambda_2, \lambda_3$不同号, $\lambda_4 \neq 0$).

$$\lambda_1 \tilde{x}^2 + \lambda_2 \tilde{y}^2 + \lambda_3 \tilde{z}^2 + \lambda_4 = 0.$$

- **抛物型** ($\lambda_1, \lambda_2, \lambda_3$中有两个非零, 一个为零, $\tilde{b}_3 \neq 0$).

$$\lambda_1 \tilde{x}^2 + \lambda_2 \tilde{y}^2 + 2\tilde{b}_3 \tilde{z} = 0.$$

- **锥面** ($\lambda_1, \lambda_2, \lambda_3$不同号).

$$\lambda_1 \tilde{x}^2 + \lambda_2 \tilde{y}^2 + \lambda_3 \tilde{z}^2 = 0.$$

- **柱面** ($\lambda_1, \lambda_2, \lambda_3$中至少一个为零, $\tilde{b}_2, \tilde{c}$至少有一为零).

$$\lambda_1 \tilde{x}^2 + \lambda_2 \tilde{y}^2 + 2\tilde{b}_2 \tilde{y} + \tilde{c} = 0.$$

下面我们给出两个实例.

例8.4.1. 确定以下二次曲线的类型、形状与位置.

$$4x^2 + 8xy + 4y^2 + 13x + 3y + 4 = 0.$$

解. 我们先做坐标系的旋转变换使得曲线方程中的交叉项xy消失.设将坐标系逆时针旋转θ角, 其中θ满足

$$\operatorname{ctg} 2\theta = (4-4)/16 = 0.$$

取$\theta = \frac{\pi}{4}$, 得坐标变换公式

$$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} \frac{\sqrt{2}}{2} & -\frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \end{pmatrix} \begin{pmatrix} x' \\ y' \end{pmatrix}$$

将上式代入原方程得

$$8(x')^2 + 8\sqrt{2}x' - 5\sqrt{2}y' + 4 = 0.$$

配方得

$$8\left(x' + \frac{\sqrt{2}}{2}\right)^2 - 5\sqrt{2}y' = 0.$$

作移轴

$$\tilde{x} = x' + \frac{\sqrt{2}}{2}, \quad \tilde{y} = y',$$

则原方程化为

$$\tilde{y} = \frac{4\sqrt{2}}{5}\tilde{x}^2.$$

因此, 所求二次曲线为抛物线, 开口朝y'轴, 焦参数$p = \frac{5\sqrt{2}}{16}$.

总的坐标变换公式为

$$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} \frac{\sqrt{2}}{2} & -\frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \end{pmatrix} \begin{pmatrix} \tilde{x} - \frac{\sqrt{2}}{2} \\ \tilde{y} \end{pmatrix}$$

□

例8.4.2. 给定二次曲面在直角坐标系下的方程

$$x^2 + 4y^2 + z^2 - 4xy - 8xz - 4yz - 1 = 0,$$

将它化为标准方程, 并指出这是什么二次曲面.

解. 将给定方程写成矩阵形式

$$(x, y, z) \begin{pmatrix} 1 & -2 & -4 \\ -2 & 4 & -2 \\ -4 & -2 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} - 1 = 0.$$

可以找到正交矩阵

$$P = \begin{pmatrix} \frac{\sqrt{5}}{5} & \frac{4\sqrt{5}}{15} & \frac{2}{3} \\ -\frac{2\sqrt{5}}{5} & \frac{2\sqrt{5}}{15} & \frac{1}{3} \\ 0 & -\frac{\sqrt{5}}{3} & \frac{2}{3} \end{pmatrix}$$

使得$P^TAP = \mathrm{diag}(5, 5, -4)$.作正交变换$(x, y, z) = (\tilde{x}, \tilde{y}, \tilde{z})P^T$, 则原方程化为标准形式

$$5\tilde{x}^2 + 5\tilde{y}^2 - 4\tilde{z}^2 = 1.$$

它表示一个单叶双曲面. □

§8.5 正定二次型

在第六章中, 我们指出由内积确定的度量矩阵G具有性质:对任意非零向量$\mathbf{x} \in \mathbb{R}^n$, $\mathbf{x}^T G\mathbf{x} > 0$.具有这种性质的矩阵称为正定方阵, 相应的二次型$\mathbf{x}^T G\mathbf{x}$称为正定的二次型.

定义8.5.1. n元实二次型

$$Q(x_1,x_2,\dots,x_n)=\sum_{i,j=1}^{n}a_{ij}x_ix_j=\mathbf{x}^TA\mathbf{x}$$

称为是**正定二次型**, 如果对任意非零向量$\mathbf{x}\in\mathbb{R}^n$,

$$Q(x_1,x_2,\dots,x_n)>0.$$

正定的二次型对应的矩阵A称为**正定矩阵**.矩阵A正定简记为$A>0$.

设正定二次型$Q(x_1,x_2,\dots,x_n)$的正负惯性指数分别为r与s, 即存在可逆变换$\mathbf{x}=P\mathbf{y}$使得

$$Q(x_1,x_2,\dots,x_n)\Big|_{\mathbf{x}=P\mathbf{y}}=y_1^2+\dots+y_r^2-y_{r+1}^2-\dots-y_{r+s}^2.$$

下面我们说明, $s=0$且$r=n$, 即

$$Q(x_1,x_2,\dots,x_n)\Big|_{\mathbf{x}=P\mathbf{y}}=y_1^2+y_2^2+\dots+y_n^2,$$

或者说A相合于单位阵.

实际上, 如果$r<n$, 取$\mathbf{x}=P(0,\dots,0,1)^T$, 则

$$Q(x_1,x_2,\dots,x_n)=y_1^2+\dots+y_r^2-y_{r+1}^2-\dots-y_{r+s}^2\le 0,$$

这与Q的正定性矛盾!

上面的结论反过来也成立, 即若$r=n$, 则二次型$Q(x_1,x_2,\dots,x_n)$正定. 这是因为对任意非零向量$\mathbf{x}$, $\mathbf{y}=P^{-1}\mathbf{x}$也非零, 从而$Q(x_1,x_2,\dots,x_n)=y_1^2+y_2^2+\dots+y_n^2>0$.因此, 我们有

定理8.5.1. n元实二次型

$$Q(x_1,x_2,\dots,x_n)=\mathbf{x}^TA\mathbf{x}$$

正定的充要条件是, Q的正惯性指数为n.等价地, 实对称方阵A正定的充要条件是A相合于单位阵.

由于二次型与实对称方阵一一对应, 下面我们仅研究实对称方阵的正定问题.

正定矩阵有以下性质

定理8.5.2. 设A为n阶实对称方阵.

1 设P为n阶实可逆方阵, $B = P^TAP$.则$B > 0$当且仅当$A > 0$.

2 设$A = \mathrm{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n)$. 则$A > 0$当且仅当$\lambda_i > 0$, $i = 1, 2, \ldots, n$.

3 若$A > 0$, 则$\det(A) > 0$.

证明. 1. 若$A > 0$, 则对任意实向量$\mathbf{x} \neq \mathbf{0}$, $\mathbf{x}^TA\mathbf{x} > 0$.由于$P\mathbf{x} \neq \mathbf{0}$, 于是$\mathbf{x}^T(P^TAP)\mathbf{x} = (P\mathbf{x})^TA(P\mathbf{x}) > 0$, 即$B > 0$.反之亦然.

2. 若$A > 0$, 取$\mathbf{x} = \mathbf{e}_i$, 这里$\mathbf{e}_i$为单位坐标向量, 则$\mathbf{x}^TA\mathbf{x} = \lambda_i > 0$. 反之, 设$\lambda_i > 0$, $i = 1, 2, \ldots, n$. 则对任意$\mathbf{x} = (x_1, x_2, \ldots, x_n) \neq \mathbf{0}$,

$$\mathbf{x}^TA\mathbf{x} = \lambda_1 x_1^2 + \lambda_2 x_2^2 + \ldots + \lambda_n x_n^2 > 0.$$

即A正定.

3. 由定理 8.5.1, 存在可逆方阵P使得$P^TAP = I$. 两边取行列式得$\det(P)^2\det(A) = 1$. 从而$\det(A) > 0$.

□

下面, 我们给出矩阵正定的若干等价条件.

定理8.5.3. *设A为n阶实对称矩阵, 则下列命题等价*

1. A正定.

2. A的特征值均为正.

3. 存在可逆方阵P使得$A = P^TP$.

4. A的各阶顺序主子式均为正, 即

$$a_{11} > 0, \quad \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} > 0, \quad \cdots \quad , \begin{vmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ldots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{vmatrix} = \det A > 0. \tag{8.11}$$

证明. $1 \Leftrightarrow 2$.

由定理 8.2.3, 存在正交阵P使得

$$P^TAP = \mathrm{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n),$$

其中$\lambda_1, \lambda_2, \ldots, \lambda_n$为$A$的特征值. 由定理 8.5.5,

$$A > 0 \Leftrightarrow P^TAP > 0 \Leftrightarrow \mathrm{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n) > 0 \Leftrightarrow \lambda_i > 0, i = 1, 2, \ldots, n.$$

$1 \Leftrightarrow 3$.

由定理 8.5.1, A正定当且仅当存在可逆方阵Q使得$Q^TAQ = I$, 当且仅当$A = (Q^{-1})^T(Q^{-1})$, 当且仅当存在可逆阵P使得$A = P^TP$.

$1 \Leftrightarrow 4$.

先证$1 \Rightarrow 4$. 设A正定, 则二次型$Q(x_1,x_2,\ldots,x_n) = \mathbf{x}^TA\mathbf{x}$正定. 令$x_{r+1} = x_{r+2} = \cdots = x_n = 0$, 则

$$Q_r(x_1,x_2,\ldots,x_r) = Q(x_1,x_2,\ldots,x_r,0,\ldots,0) = \sum_{i,j=1}^{r} a_{ij}x_ix_j$$

是一个r变元的二次型.对于任意不全为零的$x_1,x_2,\ldots,x_r$, 有

$$Q_1(x_1,x_2,\ldots,x_r) > 0.$$

所以$Q_1(x_1,x_2,\ldots,x_r)$是一个正定的二次型, 其对应的矩阵A_r也正定. 由定理 8.5.5, $\det(A_r) > 0$. 而$\det(A_r)$正是A的第r阶顺序主子式.

再证$4 \Rightarrow 1$. 对实对称矩阵的阶数n作归纳法.当$n = 1$时, 结论显然成立. 假设当$n-1$时充分性成立, 由于对称矩阵

$$A_{n-1} = \begin{pmatrix} a_{11} & \cdots & a_{1,n-1} \\ \vdots & \cdots & \vdots \\ a_{n-1,1} & \cdots & a_{n-1,n-1} \end{pmatrix}$$

的各阶顺序主子式均大于零, 根据归纳假设, 矩阵A_{n-1}是正定的.因此存在可逆矩阵P_{n-1}使得

$$P_{n-1}^TA_{n-1}P_{n-1} = I_{n-1}.$$

对于n阶对称矩阵A作如下分块

$$A = \begin{pmatrix} A_{n-1} & C \\ C^T & a_{nn} \end{pmatrix},$$

其中$C = (a_{1n},a_{2n},\ldots,a_{n-1,n})^T$.令

$$R = \begin{pmatrix} P_{n-1} & -A_{n-1}^{-1}C \\ 0 & 1 \end{pmatrix}$$

则R可逆, 且

$$\begin{aligned} R^TAR &= \begin{pmatrix} P_{n-1}^T & 0 \\ -C^TA_{n-1}^{-1} & 1 \end{pmatrix} \begin{pmatrix} A_{n-1} & C \\ C^T & a_{nn} \end{pmatrix} \begin{pmatrix} P_{n-1} & -A_{n-1}^{-1}C \\ 0 & 1 \end{pmatrix} \\ &= \begin{pmatrix} I_{n-1} & 0 \\ 0 & a_{nn} - C^TA_{n-1}^{-1}C \end{pmatrix} \end{aligned}$$

记$a=a_{nn}-C^TA_{n-1}^{-1}C$, 在上式两边取行列式得

$$\det(A)\det(R)^2=a.$$

由$\det(A)>0$知$a>0$. 令

$$P=R\begin{pmatrix} I_{n-1} & 0 \\ 0 & \frac{1}{\sqrt{a}} \end{pmatrix}$$

便有

$$P^TAP=I_n.$$

所以A是正定矩阵.定理证毕. □

下面我们给几个例子.

例8.5.1. 证明: 对任意实数x,y,

$$f(x,y)=x^2+2xy+2y^2+2x+6y+6>0.$$

证明. 考虑二次型

$$Q(x,y,z)=x^2+2y^2+6z^3+2xy+2xz+6yz,$$

其对应的矩阵为

$$A=\begin{pmatrix} 1 & 1 & 1 \\ 1 & 2 & 3 \\ 1 & 3 & 6 \end{pmatrix}$$

而顺序主子式满足

$$1>0,\quad \begin{vmatrix} 1 & 1 \\ 1 & 2 \end{vmatrix}=1>0,\quad \begin{vmatrix} 1 & 1 & 1 \\ 1 & 2 & 3 \\ 1 & 3 & 6 \end{vmatrix}=1>0,$$

所以$Q(x,y,z)$正定. 于是$f(x,y)=Q(x,y,1)>0$. □

例8.5.2. 三维欧氏空间$\mathbb{R}^3$中, 是否存在向量$\mathbf{a}_1,\mathbf{a}_2,\mathbf{a}_3$满足

$$(\mathbf{a}_1,\mathbf{a}_1)=4,\quad (\mathbf{a}_2,\mathbf{a}_2)=9,\quad (\mathbf{a}_3,\mathbf{a}_3)=4,$$
$$(\mathbf{a}_1,\mathbf{a}_2)=9/2,\quad (\mathbf{a}_1,\mathbf{a}_3)=-1/2,\quad (\mathbf{a}_2,\mathbf{a}_3)=9/2.$$

解. 假设存在向量$\mathbf{a}_1,\mathbf{a}_2,\mathbf{a}_3$满足上述条件, 令$\mathbf{a}=x\mathbf{a}_1+y\mathbf{a}_2+z\mathbf{a}_3$. 则对任意实数$x,y,z$,

$$\begin{aligned}(\mathbf{a},\mathbf{a}) &= (\mathbf{a}_1,\mathbf{a}_1)x^2+(\mathbf{a}_2,\mathbf{a}_2)y^2+(\mathbf{a}_3,\mathbf{a}_3)z^2+2(\mathbf{a}_1,\mathbf{a}_2)xy+2(\mathbf{a}_1,\mathbf{a}_3)xz+2(\mathbf{a}_2,\mathbf{a}_3)yz \\ &= 4x^2+9y^2+4z^2+9xy-xz+9yz \geq 0.\end{aligned}$$

记上述二次型为$Q(x,y,z)$, 其对应的矩阵为$A=\begin{pmatrix} 4 & 9/2 & -1/2 \\ 9/2 & 9 & 9/2 \\ -1/2 & 9/2 & 4 \end{pmatrix}$. 由于$\det(A)=-81/2<0$, 因此$A$不正定. 这与上述不等式矛盾! 因此, 满足条件的向量组$\mathbf{a}_1,\mathbf{a}_2,\mathbf{a}_3$不存在. □

例8.5.3. 给定实二次型$Q(x,y,z)=\lambda(x^2+y^2+z^2)-2xy-2yz-2zx$. λ为何值时, $Q(x,y,z)$正定?

解. $Q(x,y,z)$对应的矩阵为

$$A=\begin{pmatrix} \lambda & -1 & -1 \\ -1 & \lambda & -1 \\ -1 & -1 & \lambda \end{pmatrix}$$

则$Q(x,y,z)$正定当且仅当

$$\lambda>0, \quad \begin{pmatrix} \lambda & -1 \\ -1 & \lambda \end{pmatrix}=\lambda^2-1, \quad \det(A)=(\lambda-1)^2(\lambda+2)>0.$$

因此, $Q(x,y,z)$正定当且仅当$\lambda>1$. □

例8.5.4. 设向量组$\mathbf{a}_1,\mathbf{a}_2,\ldots,\mathbf{a}_m\in\mathbb{R}^n$. 则方阵$G=((\mathbf{a}_i,\mathbf{a}_j)_{m\times m}$正定当且仅当$\mathbf{a}_1,\mathbf{a}_2,\ldots,\mathbf{a}_m$ 线性无关.

证明. : 对任意$\mathbf{x}=(x_1,x_2,\ldots,x_m)\in\mathbb{R}^m$,

$$\mathbf{x}^TG\mathbf{x}=\left(\sum_{i=1}^m x_i\mathbf{a}_i,\sum_{i=1}^m x_i\mathbf{a}_i\right)\geq 0.$$

并且$\mathbf{x}^TG\mathbf{x}>0$当且仅当$\sum_{i=1}^m x_i\mathbf{a}_i\neq\mathbf{0}$. 因此$G$正定, 当且仅当向量组$\mathbf{a}_1,\mathbf{a}_2,\ldots,\mathbf{a}_m$ 线性无关. □

由于二次型的秩和正惯性指数是相合不变量, 我们可以根据它们对二次型进行分类. 正定二次型就是其中重要的一类.另一类重要的二次型是半正定二次型.

定义8.5.2. n元实二次型

$$Q(x_1,x_2,\ldots,x_n)=\sum_{i,j=1}^{n}a_{ij}x_ix_j=\mathbf{x}^TA\mathbf{x}$$

称为是**半正定二次型**, 如果对任意向量$\mathbf{x}\in\mathbb{R}^n$,

$$Q(x_1,x_2,\ldots,x_n)\geq 0.$$

半正定的二次型对应的矩阵A称为**半正定矩阵**.矩阵A半正定简记为$A\geq 0$.

同正定二次型(矩阵)类似, 容易证明下列结论.

定理8.5.4. n元实二次型

$$Q(x_1,x_2,\ldots,x_n)=\mathbf{x}^TA\mathbf{x}$$

半正定的充要条件是, Q的负惯性指数为0.等价地, 实对称方阵A半正定的充要条件是A相合于$\mathrm{diag}(I_r,0)$, 这里r是矩阵A的秩.

半正定矩阵有以下性质

定理8.5.5. *设A为n阶实对称方阵.*

1 设P为n阶实可逆方阵, $B=P^TAP$.则$B\geq 0$当且仅当$A\geq 0$.

2 设$A=\mathrm{diag}(\lambda_1,\lambda_2,\ldots,\lambda_n)$. 则$A\geq 0$当且仅当$\lambda_i\geq 0$, $i=1,2,\ldots,n$.

3 若$A\geq 0$, 则$\det(A)\geq 0$.

定理8.5.6. *设A为n阶实对称矩阵, 则下列命题等价*

1. A半正定.

2. A的特征值均为非负.

3. 存在实方阵P使得$A=P^TP$.

4. A的各阶主子式均为非负.

注意在上面的等价命题4中, 同正定矩阵不同, “A的各阶顺序主子式”变成了“A的各阶主子式”.

例8.5.5. 设A,B,C为三角形ABC的三个内角.证明:对任意实数x,y,z,

$$x^2+y^2+z^2\geq 2xy\cos(A)+2yz\cos(B)+2zx\cos(C).$$

证明. 只需证明二次型

$$Q(x,y,z)=x^2+y^2+z^2-2xy\cos(A)-2xz\cos(B)-2yz\cos(C)$$

半正定.实际上, 二次型$Q(x,y,z)$对应的矩阵为

$$A=\begin{pmatrix}1 & -\cos(A) & -\cos(B)\\ -\cos(A) & 1 & -\cos(C)\\ -\cos(B) & -\cos(C) & 1\end{pmatrix}.$$

A的各阶主子式为

$$\begin{aligned}
&1>0,\quad \begin{vmatrix}1 & -\cos(A)\\ -\cos(A) & 1\end{vmatrix}=\sin^2(A)\geq 0,\\
&\begin{vmatrix}1 & -\cos(B)\\ -\cos(B) & 1\end{vmatrix}=\sin^2(B)\geq 0,\quad \begin{vmatrix}1 & -\cos(C)\\ -\cos(C) & 1\end{vmatrix}=\sin^2(C)\geq 0,\\
&\det(A)=1-\cos^2 A-\cos^2 B-\cos^2 C-2\cos(A)\cos(B)\cos(C)=0.
\end{aligned}$$

因此, A半正定, 从而$Q(x,y,z)\geq 0$. □

除了正定与半正定二次型(矩阵), 还可以定义负定、半负定二次型(矩阵).简单来说, $Q(\mathbf{x})$**负定**当且仅当$-Q(\mathbf{x})$正定; $Q(\mathbf{x})$**半负定**当且仅当$-Q(\mathbf{x})$半正定.因此, 负定与半负定二次型(矩阵)都可以化为正定与半正定二次型(矩阵)处理. 总结起来, 对于n个变元的二次型$Q(x_1,x_2,\ldots,x_n)=\mathbf{x}^TA\mathbf{x}$, 它可以分为以下类别

(1) **正定二次型**:$r=n$.

(2) **半正定二次型**:$s=0, r<n$, 其规范形为

$$y_1^2+y_2^2+\cdots+y_r^2,\quad r\leq n$$

或者说对任意的变元$x_1,x_2,\ldots,x_n$, $Q(x_1,x_2,\ldots,x_n)\geq 0$. 半正定二次型的矩阵称为**半正定矩阵**, 简记为$A\geq 0$. 半正定矩阵相合于对角阵$J=\mathrm{diag}(I_r,0)$.

(3) **负定二次型**:$r=0$, $s=n$, 其规范形是

$$-y_1^2-y_2^2-\cdots-y_n^2.$$

或者说对任意的不全为零的$x_1,x_2,\ldots,x_n$, 有$Q(x_1,x_2,\ldots,x_n)<0$. 负定二次型的矩阵称为**负定矩阵**, 简记为$A<0$, 它相合于负的单位矩阵.

(4) **半负定二次型**:$r=0$, $s<n$, 其规范形为

$$-y_1^2-y_2^2-\cdots-y_s^2,\quad s<n,$$

或者说对任意的$x_1,x_2,\ldots,x_n$, 有$Q(x_1,x_2,\ldots,x_n)\le 0$.半负定二次型的矩阵称为**半负定矩阵**, 简记为$A\le 0$.它相合于对角阵$J=\operatorname{diag}(-I_s,0)$.

(5) 除上述4类之外, 其它实二次型都称为**不定型**.

作为正定与半正定矩阵的两个重要应用, 我们给出矩阵的奇异值分解与多元函数达到极小(极大)的一个充分条件.

定理8.5.7. *(**矩阵的奇异值分解**) 设$A\in\mathbb{R}^{m\times n}$, 则存在$m$阶正交阵$U$及$n$阶正交阵$V$使得*

$$A=U\begin{pmatrix}\Lambda & O\\ O & O\end{pmatrix}V, \tag{8.12}$$

*其中$\Lambda=\operatorname{diag}(\sigma_1,\sigma_2,\ldots,\sigma_r)$, $\sigma_1\ge\sigma_2\ge\ldots\ge\sigma_r>0$称为矩阵$A$的**奇异值**.*

证明. 显然$AA^T\ge 0$. 于是存在m阶正交阵U使得

$$AA^T=U\operatorname{diag}(\sigma_1^2,\ldots,\sigma_r^2,0,\ldots,0)U^T$$

其中$\sigma_1\ge\sigma_2\ge\ldots\ge\sigma_r>0$. 令$B=U^{-1}A$, 则

$$BB^T=\operatorname{diag}(\sigma_1^2,\ldots,\sigma_r^2,0,\ldots,0).$$

记$B=(\mathbf{b}_1^T,\mathbf{b}_2^T,\ldots,\mathbf{b}_m^T)^T$, 则有

$$(\mathbf{b}_i,\mathbf{b}_j)=0,\quad 1\le i\ne j\le r,\qquad |\mathbf{b}_i|=\sigma_i,\quad 1\le i\le r,\quad \mathbf{b}_i=\mathbf{0},\quad r+1\le i\le m.$$

令$\mathbf{c}_i=\frac{1}{\sigma_i}\mathbf{b}_i$, $i=1,2,\ldots,r$, 则$\mathbf{c}_i$, $i=1,2,\ldots,r$为$\mathbb{R}^n$的标准正交向量组, 将它们扩充为$\mathbb{R}^n$的一组标准正交基$\mathbf{c}_i$, $i=1,2,\ldots,n$. 令$V=(\mathbf{c}_1^T,\ldots,\mathbf{c}_n^T)^T$. 则

$$B=\begin{pmatrix}\Lambda & O\\ O & O\end{pmatrix}V.$$

因此

$$A=U\begin{pmatrix}\Lambda & O\\ O & O\end{pmatrix}V.$$

特别注意, $\sigma_1,\ldots,\sigma_r$是AA^T的非零特征值的平方根. □

如果用$\mathbf{u}_i$表示矩阵U的列向量, $i = 1,2,\ldots,m$, 用$\mathbf{v}_i^T$表示矩阵V的行向量, $i = 1,2,\ldots,n$. 则公式 (8.12) 可以改写为

$$A = \sigma_1 \mathbf{u}_1 \mathbf{v}_1^T + \sigma_2 \mathbf{u}_2 \mathbf{v}_2^T + \ldots + \sigma_r \mathbf{u}_r \mathbf{v}_r^T. \tag{8.13}$$

在实际应用中, 当$\sigma_{s+1},\ldots,\sigma_r (s < r)$比较小时, 我们可以用矩阵

$$B = \sigma_1 \mathbf{u}_1 \mathbf{v}_1^T + \sigma_2 \mathbf{u}_2 \mathbf{v}_2^T + \ldots + \sigma_s \mathbf{u}_s \mathbf{v}_s^T$$

给出矩阵A的一个良好的低秩逼近. 这一结论在实际中有广泛的应用.

例8.5.6. (**多元函数的极值**) 设$\Omega \subset \mathbb{R}^n$是有界区域, $\mathbf{x}^0 = (x_1^0,\ldots,x_n^0)$是$\Omega$ 内一点, $f(\mathbf{x}) \in C^2(\Omega)$, $\mathbf{x} = (x_1,\ldots,x_n)$. 且$\frac{\partial f}{\partial x_i}(\mathbf{x}^0) = 0$, $i = 1,2,\ldots,n$, $H(\mathbf{x}^0) = \left(\frac{\partial^2 f}{\partial x_i \partial x_j}(\mathbf{x}^0)\right)_{n\times n} > 0$. 则$f(\mathbf{x})$在$\mathbf{x}^0$取极小值.

证明. 由Taylor展开,

$$\begin{aligned} f(\mathbf{x}) &= f(\mathbf{x}^0) + \nabla f(\mathbf{x}^0)\cdot(\mathbf{x}-\mathbf{x}^0) + \frac{1}{2}\sum_{i=1}^n\sum_{j=1}^n \frac{\partial^2 f}{\partial x_i \partial x_j}(\mathbf{x}^0)(x_i - x_i^0)(x_j - x_j^0) + o(||\mathbf{x}-\mathbf{x}^0||^2) \\ &= f(\mathbf{x}^0) + \frac{1}{2}(\mathbf{x}-\mathbf{x}^0)^T H(\mathbf{x}^0)(\mathbf{x}-\mathbf{x}^0) + o(||\mathbf{x}-\mathbf{x}^0||^2) \end{aligned}$$

由$H(\mathbf{x}^0) > 0$知, 存在正交方阵P使得

$$P^T H(\mathbf{x}^0) P = \mathrm{diag}(\lambda_1,\lambda_2,\ldots,\lambda_n),$$

其中$\lambda_1 \ge \lambda_2 \ge \ldots \ge \lambda_n > 0$. 令$\mathbf{x}-\mathbf{x}^0 = P\mathbf{y}$, 则$||\mathbf{x}-\mathbf{x}^0|| = ||\mathbf{y}||$. 于是

$$\begin{aligned} (\mathbf{x}-\mathbf{x}^0)^T H(\mathbf{x}^0)(\mathbf{x}-\mathbf{x}^0) &= \mathbf{y}^T \mathrm{diag}(\lambda_1,\lambda_2,\ldots,\lambda_n)\mathbf{y} \\ &= \lambda_1 y_1^2 + \lambda_2 y_2^2 + \ldots + \lambda_n y_n^2 \ge \lambda_n (y_1^2 + y_2^2 + \ldots + y_n^2) \\ &= \lambda_n ||\mathbf{y}||^2 = \lambda_n ||\mathbf{x}-\mathbf{x}^0||^2. \end{aligned}$$

这样

$$f(\mathbf{x}) \ge f(\mathbf{x}^0) + \frac{\lambda_n}{2}||\mathbf{x}-\mathbf{x}^0||^2 + o(||\mathbf{x}-\mathbf{x}^0||^2).$$

因此, 存在$\delta > 0$使得$||\mathbf{x}-\mathbf{x}^0|| < \delta$时, $f(\mathbf{x}) \ge f(\mathbf{x}^0)$, 即$f(\mathbf{x})$在$\mathbf{x} = \mathbf{x}^0$达到极小.

类似可证, 当$H(\mathbf{x}^0) < 0$时, $f(\mathbf{x})$在$\mathbf{x} = \mathbf{x}^0$达到极大. □

习题七

1. 将下列二次型表示成矩阵形式

(1) $Q(x_1,x_2,x_3)=-x_1^2+x_2^2+2x_3^2+4x_1x_2+3x_1x_3+4x_2x_3$;

(2) $Q(x_1,x_2,x_3)=2x_1^2+2x_1x_2-x_2x_3$;

(3) $Q(x_1,x_2,x_3,x_4)=x_1x_2+3x_1x_3+5x_1x_4+6x_2x_3+7x_2x_4+10x_3x_4$;

(4) $Q(x_1,x_2,\cdots,x_n)=\sum\limits_{i=1}^{n-2}(x_i-x_{i+2})^2$.

2. 写出下列对称矩阵对应的二次型

(1)$A=\begin{pmatrix}1&3&5\\3&2&6\\5&6&4\end{pmatrix}$ (2)$A=\begin{pmatrix}a&b&0\\b&a&b\\0&b&a\end{pmatrix}$

(3)$A=\begin{pmatrix}0&-1&2&-3\\-1&0&3&-2\\2&3&0&2\\-3&-2&2&0\end{pmatrix}$

3. 用配方法将下列二次型化为标准形, 并求相应的可逆线性变换

(1)$Q=x_1^2+x_2^2+x_3^2+x_1x_2+x_1x_3+x_2x_3$

(2)$Q=x_1^2+x_2x_3$

(3)$Q=x_1x_2+x_2x_3+x_3x_4+x_4x_1$

(4)$Q=x_1^2+5x_2^2-4x_3^2+2x_1x_2-4x_2x_3$

4. 用初等变换法将下列二次型化成标准形, 并求相应的可逆线性变换

(1)$Q=\mathbf{x}^T\begin{pmatrix}2&1&1\\1&3&2\\1&2&1\end{pmatrix}\mathbf{x}$ (2)$Q=\mathbf{x}^T\begin{pmatrix}0&-1&0\\-1&1&1\\0&1&1\end{pmatrix}\mathbf{x}$

(3)$Q=\mathbf{x}^T\begin{pmatrix}0&1&1\\1&0&-1\\1&-1&0\end{pmatrix}\mathbf{x}$ (4)$Q=\mathbf{x}^T\begin{pmatrix}0&\frac12&0&\frac12\\\frac12&0&\frac12&0\\0&\frac12&0&\frac12\\\frac12&0&\frac12&0\end{pmatrix}\mathbf{x}$

5. 求正交变换化下列实二次型为标准形

(1)$Q=x_1^2+4x_2^2+4x_3^2-4x_1x_2+4x_1x_3-8x_2x_3$

(2)$Q=x_1^2+7x_2^2+x_3^2+8x_1x_2+8x_2x_3-16x_1x_3$

(3)$Q=6x_1x_2+6x_1x_3+6x_2x_3$

(4)$Q=2x_1x_2-2x_3x_4$

6. 试证:在实数域上, 对称矩阵$\begin{pmatrix}1&0\\0&1\end{pmatrix}$和$\begin{pmatrix}1&0\\0&-1\end{pmatrix}$不相合.

7. 证明:秩等于r的对称矩阵可以表成r个秩等于1的对称矩阵之和.

8. 设A是一个n阶方阵, 证明:
(1)A是反对称阵, 当且仅当对任一个n维向量X, 有$\mathbf{x}^T A\mathbf{x}=0$;
(2)如果A为对称阵, 且对任一个n维向量$\mathbf{x}$, 有$\mathbf{x}^T A\mathbf{x}=0$, 那么$A=0$.

9. 求习题5中二次型的规范形.

10. 设A是n阶实对称阵, 且$A^2=I$, 求A的相合规范形.

11. 确定以下二次曲线的类型、形状与位置.

$$\begin{aligned}
&(1)\quad 3x^2+4xy+2y^2+8x+4y+6=0;\\
&(2)\quad x^2+xy-2y^2-11x-y+28=0;\\
&(3)\quad 4x^2-4xy+y^2-2x-14y+7=0.
\end{aligned}$$

12. 将二次方程化为最简形式, 并判断曲面类型.
(1)$4x^2-6y^2-6z^2-4yz-4x+4y+4z-5=0$.
(2)$x^2-\frac{4}{3}y^2+\frac{4}{3}z^2-\frac{16}{3}\sqrt{2}yz+4x+\frac{8}{3}\sqrt{3}z-1=0$.

13. 在三维空间直角坐标系中, 给定曲面方程

$$2x^2+4y^2+5z^2-2xy+4xz+6yz-20=0.$$

利用正交变换将其化为标准型. 该方程表示什么曲面?

14. 判断下列二次型是否是正定二次型
(1)$Q(x_1,x_2,x_3)=x_1^2+2x_2^2+4x_3^2+2x_1x_2-4x_2x_3$;
(2)$Q(x_1,x_2,x_3)=-5x_1^2-6x_2^2-4x_3^2+4x_1x_2+4x_1x_3$;
(3)$Q(x_1,x_2,x_3)=7x_1^2+x_2^2+x_3^2-2x_1x_2-4x_1x_3$;
(4)$Q(x_1,x_2,x_3)=x_1^2+5x_2^2+9x_3^2+4x_1x_2-6x_2x_3+x_1x_3$.

15. 判断下列矩阵是否是正定矩阵
(1) $\begin{pmatrix}2&\frac{1}{2}&2\\ \frac{1}{2}&2&-2\\ 2&-2&4\end{pmatrix}$ (2) $\begin{pmatrix}1&-1&1\\ -1&2&0\\ 1&0&1\end{pmatrix}$

16. 问参数t满足什么条件时, 下列二次型正定?
(1)$Q(x_1,x_2,x_3)=2x_1^2+x_2^2+x_3^2+2x_1x_2+tx_1x_3$
(2)$Q(x_1,x_2,x_3)=x_1^2+2x_2^2+3x_3^2+tx_1x_2+tx_1x_3+x_2x_3$

17. 证明: 对任意实数x,y, 二元二次多项式$f(x,y)=2x^2+3y^2+4z^2-2xy+2xz-6yz$恒大于零.

18. 设$Q=ax_1^2+bx_2^2+ax_3^2+2cx_1x_3$, 问$a,b,c$满足什么条件时, Q为正定?

19. 试证:二次型$Q(x_1,x_2,\dots,x_n)=\mathbf{x}^TA\mathbf{x}$负定的充分必要条件是$A$的奇数阶顺序主子式全小于零, 偶数阶顺序主子式全大于零.

20. 设有n元实二次型

$$Q(x_1,x_2,\cdots,x_n)=(x_1+a_1x_2)^2+(x_2+a_2x_3)^2+\cdots+(x_{n-1}+a_{n-1}x_n)^2+(x_n+a_nx_1)^2$$

其中$a_i(i=1,\cdots,n)$为实数, 试问:当$a_1,\cdots,a_n$满足何种条件时, $Q(x_1,\cdots,x_n)$为正定二次型?

21. 在$Q=\lambda(x^2+y^2+z^2)+2xy+2xz-2yz$中, 问
(1)哪些λ的值使得Q为正定?
(2)哪些λ的值使得Q为负定?
(3)当$\lambda=2$和$\lambda=-1$时, Q为什么类型?

22. 设A为n阶实对称阵, 证明:若A正定, 则对任意正整数k, A^k为正定.

23. 设A为n阶可逆实对称阵, 证明:
(1) 若A正定, 则A^{-1}亦正定.
(2) 若A正定, 则A的伴随矩阵A^*也是正定矩阵.

24. 设A为n阶实对称矩阵, 若$A^2=I$, 证明$A+I$为正定矩阵或半正定矩阵.

25. 设A,B是两个n阶实对称正定矩阵, 证明:
(1)$A+B$亦正定； (2)AB正定的充分必要条件是$AB=BA$.

26. 设A,B均为n阶实对称阵, 其中A为正定阵, 证明:当实数t充分大后, $tA+B$亦正定.

27. A是实对称正定矩阵, 证明:存在下三角阵P, 使得$A=PP^T$. 进一步证明, 存在下三角阵L, 其对角线元素均为1, 以及对角阵D, 其对角元均为正使得$A=LDL^T$.

28. 设A为n阶实对称阵, 且$A>0$.求证:$\det(A)\le\left(\frac{1}{n}Tr(A)\right)^n$.

29. 证明:若$A=(a_{ij})$是n阶正定矩阵, 则$\det(A)\le a_{11}a_{22}\cdots a_{nn}$.

30. 设A为n阶实对称方阵, 证明下列命题等价

(1) A为半正定方阵.

(2) $A=P^TP$, 其中P为n阶实方阵.

(3) A的各阶主子式皆非负.

31. 在二次型$Q=\mathbf{x}^TA\mathbf{x}$中, 实对称阵$A=(\alpha_{ij})_{n\times n}$, 若

$$\begin{vmatrix} a_{11} & \cdots & a_{1k} \\ \cdots & \cdots & \cdots \\ a_{k1} & \cdots & a_{kk} \end{vmatrix} > 0, \quad k=1,2,\ldots,n-1, \quad \det A=0$$

证明:Q为半正定.

32. 设$A\in\mathbb{R}^{m\times n}$, $A=U\begin{pmatrix}\Lambda & O\\ O & O\end{pmatrix}V$是$A$的奇异值分解, 如定理 8.5.7 所述. 令$A^+=V^T\begin{pmatrix}\Lambda^{-1} & O\\ O & O\end{pmatrix}U^T$, 称$A^+$为矩阵$A$的**广义逆**. 证明:

$$AA^+A=A, \quad A^+AA^+=A^+, \quad (AA^+)^*=AA^+, \quad (A^+A)^*=A^+A.$$

33. 设$A\in\mathbb{R}^{m\times n}$. 证明: $|A|:=\max_{\mathbf{x}\neq\mathbf{0}}\frac{|A\mathbf{x}|}{|\mathbf{x}|}=\sigma_{\max}$, 这里$\sigma_{\max}$是矩阵$A$ 的最大奇异值. 称$|A|$为矩阵A的**谱范数**.

参考文献

[1] 代数学引论, 许以超编著, 上海科学技术出版社, 1966年5月第1版.

[2] 高等数学引论(余篇), 华罗庚著, 科学出版社, 1984年7月第1版, 2009年4月高等教育出版社重印(王元校).

[3] 线性代数, 李炯生、查建国编著, 中国科学技术大学出版社, 1989年4月第1版, 2010年1月第2版.

[4] 高等代数, 刘昌堃、叶世源、叶家琛编, 同济大学出版社, 1995年4月第1版.

[5] 解析几何, 丘维声编, 北京大学出版社, 1996年10月第2版.

[6] 高等代数教程(上、下), 王萼芳编著, 清华大学出版社, 1997年4月第1版.

[7] 线性代数简明教程, 陈龙玄、钟立敏编, 中国科学技术大学出版社, 1989年2月第1版, 1997年9月修订版.

[8] 高等代数与解析几何(上、下), 陈志杰主编, 高等教育出版社, 2000年6月第1版.

[9] 简明线性代数, 丘维声编著, 北京大学出版社, 2002年2月第1版.

[10] 高等代数(上、下), 丘维声编著, 高等教育出版社, 2002年7月第2版.

[11] 高等代数学, 张贤科、许甫华编著, 清华大学出版社, 2004年7月第2版.

[12] 微积分(上、下), 谢盛刚、李娟、陈秋桂编, 科学出版社, 2005年1月第1版.

[13] 线性代数及其应用(原书第2版), (美)David C.Lay著, 刘深泉、洪毅、马东魁、郭国雄、刘勇平译, 机械工业出版社, 2005年8月.

[14] 线性代数, 李尚志编著, 高等教育出版社, 2006年5月第1版.

[15] 高等代数与解析几何(上、下), 孟道骥著, 科学出版社, 2007年1月第2版.

[16] 高等代数简明教程(上、下), 蓝以中编著, 北京大学出版社, 2007年7月第2版.

[17] 解析几何简明教程, 吴光磊、田畴编, 高等教育出版社, 2008年3月第2版.