目　录

第1章 预备知识

1.1 集合与映射

定义 1.1.1 某些特定对象 (object) 的汇集 (collection) S 称为一个集合 (set), 其中的对象 x 称为集合 S 的元素, 记为 $x \in S$. 不含任何对象的集合称为空集, 记为 $\emptyset$. 如果一个元素不在集合 S 中, 则记为 $x \notin S$.

一般有两种方式表示集合, 列出全部元素或写出刻画全部元素的条件. 例如, $S = \{0, 1, 2, \ldots, 100\}$, 也可表成 $S = \{x \mid x$ 是不超过 100 的自然数$\}$.
下面我们固定几个常用集合的符号:

$\mathbb{N} = \{0, 1, 2, \cdots\}$ 表示所有自然数的集合;

$\mathbb{Z} = \{\cdots, -2, -1, 0, 1, 2, \cdots\}$ 表示所有整数的集合;

$\mathbb{Z}_+ = \{1, 2, 3, \cdots\}$ 表示所有正整数的集合;

$\mathbb{Q} = \{p/q \mid p, q \in \mathbb{Z}, q \neq 0\}$ 表示所有有理数的集合;

$\mathbb{R} = \{$实数$\}$ 表示所有实数的集合;

$\mathbb{C} = \{$复数$\}$ 表示所有复数的集合, 也可以写成 $\mathbb{C} = \{a + b\mathrm{i} \mid a, b \in \mathbb{R}\}$.

定义 1.1.2 两个集合称为相等, 如果他们元素都一样. 集合 X 称为 Y 的子集合, 如果 X 中的元素都在 Y 中, 记为 $X \subseteq Y$ (或 $X \subset Y$). Y 的子集合 X 称为 Y 的真子集, 如果存在 $x \in X$ 但 $x \notin Y$, 记为 $X \subsetneq Y$. 显然,

$$X = Y \Leftrightarrow X \subseteq Y,\ Y \subseteq X.$$

我们约定: 空集 $\emptyset$ 是任一集合的子集. 所以, 任意集合 S 都有平凡子集: $\emptyset, S$. S 的其他子集称为非平凡子集.

定义 1.1.3 设 S 是一个集合, $P(S)$ 表示 S 中所有子集合的集合, 则称 $P(S)$ 是集合 S 的幂集.

集合的并与交: 设 X, Y 是两个集合. $X \cup Y$ 是由 X 中所有元素与 Y 中所有元素合并而成的集合, 称为 X 与 Y 的并. 显然,

$$x \in X \cup Y \Leftrightarrow x \in X \text{ 或 } x \in Y.$$

$X \cap Y$ 是由 X 与 Y 中相同元素组成的集合, 称为 X 与 Y 的交. 显然

$$x \in X \cap Y \Leftrightarrow x \in X,\ x \in Y.$$

所以我们也可以用数学符号来定义集合 X 与 Y 的并与交:

$$X \cup Y = \{x \mid x \in X \text{ 或 } x \in Y\}.$$

$$X \cap Y = \{x \mid x \in X,\ x \in Y\}.$$

如果 $X \cap Y = \emptyset$, 则称 X 与 Y 不相交. 我们事实上有更一般的定义, 设 $\{X_i\}_{i \in I}$ 是一族集合(可以是无限多个), 定义:

$$\bigcup_{i \in I} X_i = \left\{x \mid \text{存在 } i \in I \text{ 使 } x \in X_i\right\},$$

$$\bigcap_{i \in I} X_i = \{x \mid \forall i \in I, x \in X_i\}.$$

集合的差集与补集: 设 X, Y 是集合, 则集合

$$X \setminus Y = \{x \in X \mid x \notin Y\}$$

称为差集. 此时不要求 Y 是 X 的子集, 如果 Y 是 X 的子集, 则 $X \setminus Y$ 称为 Y 在 X 中的补集 (也记为 $\overline{Y}$).

集合的笛卡儿积: 设 X, Y 是集合, 则所有有序对 (x, y) 形成的集合称为 X与Y 的笛卡儿积 (也称乘积), 记为 $X \times Y$, 即 $X \times Y = \{(x, y) \mid x \in X, y \in Y\}$. 其中两个元素 (x, y) 与 (z, w) 相等当且仅当 $x = z, y = w$. 所以一般 $X \times Y \neq Y \times X$. 我们有更一般的定义, 设 $X_1, X_2, \cdots, X_n$ 是有限个集合, 则

$$X_1 \times X_2 \times \cdots \times X_n = \{(x_1, x_2, \cdots, x_n) \mid x_1 \in X_1, x_2 \in X_2, \cdots, x_n \in X_n\}$$

称为 $X_1, X_2, \cdots, X_n$ 的乘积, 其中 $(x_1, x_2, \cdots, x_n) = (y_1, y_2, \cdots, y_n)$ 当且仅当 $x_1 = y_1, x_2 = y_2, \cdots, x_n = y_n$. 如果 $X_1 = X_2 = \cdots = X_n := X$, 则 $X_1 \times X_2 \times \cdots \times X_n$ 记为

$$X^n = \underbrace{X \times \cdots \times X}_{n}.$$

例 1.1.1 设 $X = \mathbb{R}$, 则 $\mathbb{R}^2 = \mathbb{R} \times \mathbb{R}$ (实平面). 一般地,

$$\mathbb{R}^n = \underbrace{\mathbb{R} \times \cdots \times \mathbb{R}}_{n} = \{(x_1, x_2, \cdots, x_n) \mid x_i \in \mathbb{R}\}$$

称为 n-维实空间.

下面我们讨论集合之间映射 (也有教科书称为函数), 它是数学中最重要的概念 (没有之一!).

定义 1.1.4 设 X, Y 是集合, 从 X 到 Y 的一个映射 (或函数) $f : X \to Y$ 是指一个规则 (用 f 表示), 它对 X 中任一元素 $x \in X$ 指定 (通过规则 f) Y 中的唯一一

个元素 $y \in Y$. (由于 y 是由 x (通过规则 f) 唯一确定的, 所以我们称 y 是 x 在 f 下的像, 记为 $y = f(x)$.

有时为了更形象地描述 $f : X \to Y$, 我们也使用记号: $f : X \to Y, x \mapsto f(x)$ 等. 两个映射 $f : X \to Y$, $g : Z \to W$ 称为相等, 如果 $X = Z, Y = W$ 且 $f(x) = g(x)$ 对任意 $x \in X$ 成立.

例 1.1.2 我们学习过的三角函数, 指数函数, 多项式函数等都可以看成映射 $f : \mathbb{R} \to \mathbb{R}$. 例如, $\sin x$, 它对应的“规则”就是 $f = \sin : \mathbb{R} \to \mathbb{R}$ ($x \mapsto f(x) = \sin x$).

例 1.1.3 设 $B = \{0, 1\}$ 是两个元素的集合, 任何映射

$$f : B^n \to B,\ (x_1, \cdots, x_n) \mapsto f(x_1, \cdots, x_n)$$

称为一个次数为 n 的布尔函数 (Boolean function).

例 1.1.4 设 $a_1, a_2, \cdots, a_n \in \mathbb{R}$ 是固定的实数, 则

$$L : \mathbb{R}^n \to \mathbb{R},\ (x_1, \cdots, x_n) \mapsto L(x_1, \cdots, x_n) = a_1x_1 + \cdots + a_nx_n$$

是一个映射, 称为 n 个变元的线性函数 (或线性映射).

对任意映射 $f : X \to Y$ 及 $x \in X$, $f(x) \in Y$ 称为 x 的像, 而 x 称为 $f(x)$ 的一个原像.

定义 1.1.5 子集合 $f(X) = \{f(x) \mid \forall x \in Y\} \subseteq Y$ 称为映射 $f : X \to Y$ 的像 (也记为 Imf). 而对任意的 $y \in Y$, 子集合 $f^{-1}(y) = \{x \in X \mid f(x) = y\} \subseteq X$ 称为 f 在 $y \in Y$ 的纤维. 更一般地, 对任意子集 $Y_0 \subseteq Y$, 子集合

$$f^{-1}(Y_0) = \{x \in X \mid f(x) \in Y_0\} \subseteq X$$

称为 Y_0 在 f 下的逆像 (也称 Y_0 的原像).

例 1.1.5 考虑映射 $f : \mathbb{R}^n \to \mathbb{R}^m$,

$$x = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \mapsto \begin{pmatrix} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots\cdots + a_{mn}x_n \end{pmatrix} = \begin{pmatrix} \sum\limits_{j=1}^{n} a_{1j}x_j \\ \sum\limits_{j=1}^{n} a_{2j}x_j \\ \vdots \\ \sum\limits_{j=1}^{n} a_{mj}x_j \end{pmatrix}$$

其中, $a_{ij}\ (1 \le i \le m, 1 \le j \le m)$ 是固定的实数. 则对任意的

$$b = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix} \in \mathbb{R}^m, f^{-1}(b) = \left\{ x = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \middle| \begin{pmatrix} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2 \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots\cdots + a_{mn}x_n = b_m \end{pmatrix} \right\}.$$

显然, $f^{-1}(b) \neq \emptyset \Leftrightarrow$ 方程组 (称为线性方程组)

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

有解. 所以, 映射 $f : \mathbb{R}^n \to \mathbb{R}^m$ 的纤维可以是空集, 也可以不止一个元素. 那么 f 的像 $f(\mathbb{R}^n) \subseteq \mathbb{R}^m$ 又是什么呢?

为了方便表述, 我们引入记号:

$$A^{(1)} = \begin{pmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{pmatrix}, A^{(2)} = \begin{pmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{m2} \end{pmatrix}, \cdots, A^{(n)} = \begin{pmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{mn} \end{pmatrix} \in \mathbb{R}^m,$$

在 $\mathbb{R}^m$ 中, 我们可以引入两个运算 “加法” 和 “数乘法”:

$$\forall \lambda \in \mathbb{R}, \alpha = \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_m \end{pmatrix}, \beta = \begin{pmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_m \end{pmatrix} \in \mathbb{R}^m, \text{定义:}$$

$$\alpha + \beta = \begin{pmatrix} \alpha_1 + \beta_1 \\ \alpha_2 + \beta_2 \\ \vdots \\ \alpha_m + \beta_m \end{pmatrix}, \ \lambda \cdot \alpha = \begin{pmatrix} \lambda\alpha_1 \\ \lambda\alpha_2 \\ \vdots \\ \lambda\alpha_m \end{pmatrix}.$$

则对任意 $x = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \in \mathbb{R}^n$, $f(x) = x_1A^{(1)}+x_2A^{(2)}+\cdots+x_nA^{(n)}$. $(x_1A^{(1)}+x_2A^{(2)}+\cdots+x_nA^{(n)}$ 称为 $A^{(1)}, A^{(2)}, \cdots, A^{(n)}$ 的一个"线性组合", $x_1, x_2, \cdots, x_n$ 称为该"线性组合"的系数).

所以, $f(\mathbb{R}^n) = \{x_1A^{(1)} + \cdots + x_nA^{(n)} \mid \forall x_1, \cdots, x_n \in \mathbb{R}\} \subseteq \mathbb{R}^m$, 即 $\mathbb{R}^m$ 中元素 $y = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{pmatrix} \in \mathbb{R}^m$ 在 $f(\mathbb{R}^n)$ (f 的像) 中的充分必要条件是: y 可以写成 $A^{(1)}, A^{(2)}, \cdots, A^{(n)}$ 的"线性组合", 亦即方程组:

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = y_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = y_2 \\ \qquad\qquad\vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots\cdots + a_{mn}x_n = y_m \end{cases}$$

有解!

显然, $f(\mathbb{R}^n)$ 可以是 $\mathbb{R}^m$ 的真子集, i.e. $f(\mathbb{R}^n) \subsetneq \mathbb{R}^m$.

定义 1.1.6 映射 $f: X \to Y$ 称为单射, 如果 f 将 X 中不同的元素映到不同的元素: $a \neq b \Rightarrow f(a) \neq f(b)$. 映射 f 称为满射, 如果对任意 $y \in Y$, 都至少存在一个 $x \in X$ 使得 $f(x) = y$. 如果 f 既是单射又是满射, 则称 f 是双射.

等价的定义是: 设 $f: X \to Y$ 是一个映射, 则

f 是单射 $\Leftrightarrow \forall y \in Y,\ f^{-1}(y)$ 最多只有一个元素.

f 是满射 $\Leftrightarrow \forall y \in Y,\ f^{-1}(y)$ 至少有一个元素.

f 是双射 $\Leftrightarrow \forall y \in Y,\ f^{-1}(y)$ 恰一个元素.

对任意的集合 X, 有一个双射 $1_X: X \to X, x \mapsto x$ 称为恒等映射.

定义 1.1.7 设 $f: X \to Y, g: Z \to W$ 是两个映射, 当 $Y = Z$ 时, 可定义映射

$$g \cdot f: X \to W, x \mapsto f(x) \mapsto g(f(x)).$$

称为 f 和 g 的合成 (或称 f 和 g 的乘积).

注意, f 和 g 可以合成的充分必要条件是 $Y = Z$. 如果考虑所有 X 到自身的映射的集合

$$F(X) = \{f : X \to X \mid f\text{是映射}\},$$

则对任意 $f, g \in F(X)$, $f \cdot g$, $g \cdot f$ 都可定义. 所以映射的合成 $F(X)$ 上定义了一个乘法 (非常重要的乘法!)

例 1.1.6 设 $f = \sin : \mathbb{R} \to \mathbb{R}, x \mapsto \sin x$, $g = \mathbb{R} \to \mathbb{R}, x \mapsto 3 + x^2$, 则

$g \cdot f : \mathbb{R} \to \mathbb{R}, x \mapsto g(\sin x) = 3 + \sin^2 x$,

$f \cdot g : \mathbb{R} \to \mathbb{R}, x \mapsto f(g(x)) = f(3 + x^2) = \sin(3 + x^2)$.

可见, 一般来说, $g \cdot f \neq f \cdot g$ (即使它们都有定义). 另外, 我们中学还学过函数的乘法: $fg : \mathbb{R} \to \mathbb{R}, x \mapsto f(x) \cdot g(x)$ (即函数值相等). 显然, 函数的乘法满足 $fg = gf$. 注意: 映射的乘法对应函数的复合.

定理 1.1.1 设 $f : X \to Y$ 是一个映射 (我们有时也用符号 $X \xrightarrow{f} Y$). 则我们有下述结论:

(1) f 是单射 $\Leftrightarrow$ 存在映射 $g : Y \to X$ 使 $g \cdot f = 1_X$.

(2) f 是满射 $\Leftrightarrow$ 存在映射 $h : Y \to X$ 使 $f \cdot h = 1_Y$.

(3) f 是双射 $\Leftrightarrow$ 存在映射 $g : Y \to X, h : Y \to X$ 使

$$g \cdot f = 1_X,\ f \cdot h = 1_Y.$$

在证明定理之前, 我们先介绍一个关于映射合成 (或乘法) 的重要性质: 映射合成满足结合律!

性质 1.1.1 设 $X \xrightarrow{f} Y, Y \xrightarrow{g} Z, Z \xrightarrow{h} W$ 是三个映射, 则 $h \cdot (g \cdot f) = (h \cdot g) \cdot f$.

证明: 首先注意到, $h \cdot (g \cdot f)$, $(h \cdot g) \cdot f$ 都是从 X 到 W 的映射, 所以我们只需证明: $\forall x \in X$, 有

$$h \cdot (g \cdot f)(x) = (h \cdot g) \cdot f(x).$$

验证: $h \cdot (g \cdot f)(x) = h(g \cdot f(x)) = h(g(f(x))) = h \cdot g(f(x)) = (h \cdot g) \cdot f(x)$. □

定理 1.1.1 (3) 的推论: 对于映射 $f : X \to Y$, 如果存在 $g : Y \to X, h : Y \to X$ 使 $g \cdot f = 1_X,\ f \cdot h = 1_Y$. 则 $g = h$ 且由 f 唯一确定. 所以我们统一用: $f^{-1} : Y \to X$

表示, 称为 f 的逆映射. 为证明 $g=h$, 对任意 $y\in Y$, 只需证明 $g(y)=h(y)$. 由 $f\cdot h=1_Y$, 得 $f(h(y))=y$. 所以 $g(y)=g(f(h(y)))=g\cdot f(h(y))=1_X(h(y))=h(y)$.

定理 1.1.1 的证明: (1) 首先证明: 如果存在 $g:Y\to X$ 使 $g\cdot f=1_X$, 则 f 必为单射: $\forall x_1,x_2\in X$, 只需证明, 如果 $f(x_1)=f(x_2)$, 则 $x_1=x_2$. 事实上, $f(x_1)=f(x_2)\Rightarrow g(f(x_1))=g(f(x_2))$ 即 $g\cdot f(x_1)=g\cdot f(x_2)$. 由 $g\cdot f=1_X$, 可得 $x_1=x_2$. 反之, 如果 f 是单射, 我们需要构造映射 $g:Y\to X$ 使 $g\cdot f=1_X$: 令 $\overline{f(X)}=Y\setminus f(X)$, 则 $Y=\overline{f(X)}\bigcup f(X)$ 且 $\overline{f(X)}\bigcap f(X)=\emptyset$. 固定任一元素 $x_0\in X$, 定义 $g_1:\overline{f(X)}\to X, y\mapsto x_0$ (即将 $\overline{f(X)}$ 中所有元素映射到同一元素 x_0). 另一方面, 对任意 $y\in f(X)$ 存在唯一 $x\in X$ 使 $f(x)=y$ (因为 f 是单射). 由于 x 是由 y 唯一确定, 可记 $x=g_2(y)$. 从而得到 (唯一) 映射 $g_2:f(X)\to X$ 使 $g_2(f(x))=x$ ($\forall x\in X$). 定义 $g:Y\to X$ 如下:

$$g(y)=\begin{cases} g_1(y), \text{ 如果 } y\in\overline{f(X)},\\ g_2(y), \text{ 如果 } y\in f(X),\end{cases}$$

可得 $g\cdot f=1_X$ (显然, 当 $f(X)\subsetneq Y$ 时, 映射 g 不是唯一的).

(2) 如果存在 $h:Y\to X$ 使 $f\cdot h=1_Y$, 则 $\forall y\in Y$, $f(h(y))=y$. 所以 f 是满射. 如果 $f:X\to Y$ 是满射, 则对任意 $y\in Y$, 纤维 $f^{-1}(y)\subset X$ 是非空子集. 所以, $\forall y\in Y$, 任取 $x\in f^{-1}(y)$ 并固定, 记为 $x=h(x)$. 则: $Y\xrightarrow{h}X$ 定义了一个映射, 使 $f(h(y))=y$, 即 $f\cdot h=1_Y$ (如果 f 不是单射, 这样的 h 显然不是唯一的).

(3) 由 (1), (2) 直接推出.

注记: (1) 中的 g 称为 f 的左逆, (2) 中的 h 称为 f 的右逆, 它们一般不唯一. 如果 f 既有左逆, 又有右逆, 则它们必唯一且相等. 所以, $f:X\to Y$ 是双射 $\Leftrightarrow$ 存在唯一映射 $g:Y\to X$ 使得 $g\cdot f=1_X$, $f\cdot g=1_Y$. 这样的 g 称为 f 的逆映射, 记为 $g=f^{-1}$.

定义 1.1.8 两个集合 X 和 Y 称为等价 (或同构), 如果存在双射 $X\xrightarrow{f}Y$ (双射可以记为 $X\xrightarrow{\sim}Y$).

集合论的任务之一是试图在同构意义下对集合进行分类, 对于有限集合, 这样的分类比较符合我们的直观. 如果用 $|X|$ 表示集合 X 中元素的个数, 则 $|X|=n\Leftrightarrow$ 存在 $f:X\to\{1,2,\cdots,n\}$, 即有限集合 X 和 Y 同构的充分必要条件是 $|X|=|Y|$. 但对于无限集合, 人们很容易找到例子 X 使得 X 同构于它自己的一个真子集.

定义 1.1.9 与正自然数集 $\mathbb{Z}=\{1,2,3,\cdots\}$ 同构的集合称为可数集.

例 1.1.7 $\forall a\in\mathbb{Z}_+$, 存在双射 $f:\mathbb{Z}_+\to\overline{\{a\}}=\mathbb{Z}_+\setminus\{a\}$, 其中

$$f(n)=\begin{cases} n, \text{ 如果} n<a,\\ n+1, \text{ 如果} n\geqslant a.\end{cases}$$

所以对任意可数集 X, $\forall x\in X$, $\{x\}$ 的补集 $\overline{\{x\}}=X\setminus\{x\}$ 与 X 同构.

显然, 这种 "整体" 与 "部分" 同构的现象对有限集合是不会出现的. Dedekind 在 19 世纪提出把这一现象作为无限集合的定义: 集合 X 是无限集的充分必要条件是存在 X 的真子集 $Y\subsetneq X$ 使得 X 与 Y 同构 (等价).

然而更令人惊奇的是 Cantor 发现: 平面上点的集合与直线上点的集合等价, 即存在双射 $f:\mathbb{R}^2\to\mathbb{R}$. 这样不同维数图形的点集之间居然存在 $1-1$ 对应, 令 Cantor 自己也感到不可思议! 他在写给 Dedekind 的信中甚至认为数学中关于 "维数" 的直观描述需要修改! 但 Dedekind 在回信中认为, 这一发现与我们关于 "维数是确定一个点所需坐标的个数" 并不矛盾. 因为, 当我们用两组坐标定义同一个点时, 我们要求一组坐标是另一组坐标的连续函数. 他在信中还提议, 在考虑几何图形之间的映射时应加上 "连续性" 的要求, 并且断言不同维数的几何图形之间不存在连续的双射! 该断言直到 1910 年才被证明.

数学中讨论的对象都是带有 "结构" 的集合, 它们之间的映射都要求 "保持结构" (这样的映射一般称为态射), 而数学的主要任务之一就是对 "带结构集合" 在同构 (即 "保持结构的双射") 意义下的分类. 例如, 带有 "拓扑结构" 的集合称为拓扑空间, 带 "线性结构" 的空间称为线性空间等.

我们这门课主要讨论线性空间 (亦称向量空间) 及它们之间的线性映射 (即保持线性结构的映射). 代数学主要讨论各类带有代数结构的集合, 而所谓带代数结构的集合, 就是带有各种 "运算" 的集合, 它们之间的映射一般要求 "保持运算", 这样的映射一般称为同态.

习题1.1

1. 设 $\{A_i\}_{i\in I}$ 是一组集合 (可以是无限个), B 是任意集合. 试证明:

$$\left(\bigcup_{i\in I}A_i\right)\cap B=\bigcup_{i\in I}(A_i\cap B),\ \left(\bigcap_{i\in I}A_i\right)\cup B=\bigcap_{i\in I}(A_i\cup B).$$

2. 设 $\{A_i\}_{i\in I}$ 是集合 X 的一组子集 (可以是无限个). 试证明:

$$\overline{\bigcup_{i\in I}A_i}=\bigcap_{i\in I}\overline{A_i},\ \overline{\bigcap_{i\in I}A_i}=\bigcup_{i\in I}\overline{A_i}.$$

3. 设 $A_1, A_2, \cdots, A_n$ 是有限集合 (i.e. $|A_i| < +\infty$), 试证明:

$$\left|\bigcup_{i \in I} A_i\right| = \sum_{k=1}^{n} (-1)^{k-1} \sum_{1 \le i_1 < \cdots < i_k \le n} \left|A_{i_1} \cap A_{i_2} \cap \cdots \cap A_{i_k}\right|.$$

(提示: 对 n 使用数学归纳法, 归结为 $|A \cup B| = |A| + |B| - |A \cap B|$).

4. 设映射 $f : \mathbb{R} \to \mathbb{R}$ 由 $f(x) = x^2$ 定义, 确定:

   (a) $f^{-1}(1)$, (b) $f^{-1}(\{x \mid 0 < x < 1\})$, (c) $f^{-1}(\{x \mid x > 4\})$.

5. 设 $f : X \to Y$ 是映射, A, B 是 Y 的子集. 试证明:

   (a) $f^{-1}(A \cup B) = f^{-1}(A) \cup f^{-1}(B)$, (b) $f^{-1}(A \cap B) = f^{-1}(A) \cap f^{-1}(B)$.

6. 设 $f : X \to Y, g : Y \to Z$ 是映射, 证明:

   (a) 如果 f, g 都是单射, 则 $g \cdot f : X \to Z$ 也是单射.

   (b) 如果 f, g 都是满射, 则 $g \cdot f : X \to Z$ 也是满射.

   (c) 如果 $g, g \cdot f$ 都是单射, f 是单射吗? 证明你的答案.

   (d) 如果 $g, g \cdot f$ 都是满射, f 是满射吗? 证明你的答案.

7. 设 $f : \mathbb{R} \to \mathbb{R}$ 由 $f(x) = 2x$ 定义, 分别求: $f(\mathbb{Z}), f(\mathbb{N}), f(\mathbb{R})$.

8. 设 $f : X \to Y$ 是映射, A, B 是 Y 的子集. 证明:

$$f(A \cup B) = f(A) \cup (B), \ f(A \cap B) \subseteq f(A) \cap f(B).$$

1.2 带运算的集合

定义 1.2.1 设 S 是一个非空集合, S 上的一个运算是指一个映射 $S\times S\xrightarrow{f} S,(x,y)\mapsto f(x,y)$.

显然, 一个集合上有太多的运算. 而数学中的运算都来自实际问题的抽象, 是为了描述问题的本质, 所以它们一般要求满足一定的条件. 我们先来看几个重要的例子.

例 1.2.1 用 K 表示集合 $\mathbb{Q}$ 或 $\mathbb{R}$ 或 $\mathbb{C}$. 在 K 中有两个运算:

$$K\times K\xrightarrow{f} K,\ f(a,b)=a+b.$$

$$K\times K\xrightarrow{g} K,\ g(a,b)=a\cdot b.$$

它们满足下述条件:

$$\text{关于“加法”}\begin{cases}(1)\ \forall a,b,c\in K,(a+b)+c=a+(b+c).\ (\text{结合律}).\\(2)\ \text{存在}\ 0\in K\ \text{使得}:\forall a\in K,a+0=0+a=a.\ (\text{零元的存在性}).\\(3)\ \forall a\in K,\text{存在}\ b\in K\ \text{使}\ a+b=b+a=0.\ (\text{负元的存在性}).\\(4)\ \forall a,b\in K,a+b=b+a.\ (\text{交换律}).\end{cases}$$

$$\text{关于“乘法”}\begin{cases}(5)\ \forall a,b,c\in K,(a\cdot b)\cdot c=a\cdot(b\cdot c).\ (\text{结合律}).\\(6)\ \text{存在}\ 1\in K\ \text{使得}:\forall a\in K,a\cdot 1=1\cdot a=a.\ (\text{单位元的存在性}).\\(7)\ \forall a\in K,\text{存在}\ b\in K\ \text{使}\ a\cdot b=b\cdot a=1.\ (\text{逆元的存在性}).\\(8)\ \forall a,b\in K,a\cdot b=b\cdot a.\ (\text{交换律}).\end{cases}$$

(9) $\forall a,b,c\in K,(a+b)\cdot c=a\cdot c+b\cdot c$. (分配律).

一般考虑集合 $\mathbb{Q},\mathbb{R},\mathbb{C}$ 时, 都和上述两种运算一并考虑在内通常称 $\mathbb{Q}$ 是有理数域, $\mathbb{R}$ 是实数域, $\mathbb{C}$ 是复数域.

例 1.2.2 所有整数的集合 $\mathbb{Z}$ 上也存在两个运算 “加法” 和 “乘法”. 满足上述例子中除条件 (7) 外的所有条件, 一般称 $\mathbb{Z}$ 是整数环.

例 1.2.3 设 X 是一个非空集合, 令 S_X 表示所有双射 $f:X\to X$ 的集合. 则在 S_X 有一个自然的运算 (亦称乘法): $S_X\times S_X\to S_X,(f,g)\mapsto f\cdot g$, 此处 $f\cdot g:X\to X$ 是映射 $X\xrightarrow{g} X$ 与 $X\xrightarrow{f} X$ 的合成.

容易验证该运算满足:

(a) $\forall f, g, h \in S_X, (f \cdot g) \cdot h = f \cdot (g \cdot h)$. (结合律).

(b) 恒等映射 $1_X \in S_X$ 满足 $1_X \cdot f = f \cdot 1_X = f$. (单位元的存在性).

(c) $\forall f \in S_X$, 存在 $g \in S_X$ 使 $f \cdot g = g \cdot f = 1_X$. (可逆元的存在性).

这样的 S_X 称为 X 的变换群. 特别, 当 $|X| = n$ 时, S_X 记为 S_n 称为 n 个元素的置换群 (或称 n 阶对称群).

上述的三个例子可以抽象出数学中三个非常重要的概念: 域, 环, 群.

定义 1.2.2 设 G 是一个带有运算 $G \times G \to G, (a, b) \mapsto a \cdot b$ 的非空集合, 如果该运算满足:

(1) $\forall a, b, c \in G, (a \cdot b) \cdot c = a \cdot (b \cdot c)$. (结合律).

(2) 存在元素 $1 \in G$, 使 $1 \cdot a = a \cdot 1 = a$ 对任意 $a \in G$ 成立. (单位元存在性).

(3) $\forall a \in G$, 存在 $b \in G$ 使 $a \cdot b = b \cdot a = 1$. (每个元素都可逆).

则称 G 是一个群. 如果该运算还满足交换律:

(4) $\forall a, b \in G, a \cdot b = b \cdot a$.

则称 G 是一个交换群 (或阿贝尔群).

注记: (1) 单位元 $1 \in G$ 是唯一的, 即如果有 $1 \in G$ 和 $1' \in G$ 满足条件 (2), 则 $1 = 1'$.

(2) 对任意给定的 $a \in G$, 条件 (3) 中的 b 是由 a 唯一确定的. 即, 如果存在 $b, b' \in G$, 使 $a \cdot b = b \cdot a = 1, a \cdot b' = b' \cdot a = 1$, 则 $b = b'$. 所以, 我们一般用 a^{-1} 表示条件 (3) 中的 b, 称为 a 的逆元.

(3) 对于交换群, 我们一般用加法符号 $+$ 表示运算, 并且用 0 表示 G 的单位元 (称为零元), 用 $-a$ 表示 a 的逆元 (称为 a 的负元).

定义 1.2.3 设 R 是一个带有两个运算的非空集合, 分别用 $(a, b) \mapsto a + b$ (加法), $(a, b) \mapsto a \cdot b$ (乘法) 表示它的两个运算. 如果它们满足:

① $(a + b) + c = a + (b + c), \forall a, b, c \in R$. (结合律).

② 存在 $0 \in R$ 使 $a + 0 = 0 + a = a, \forall a \in R$. (零元的存在性).

③ $\forall a \in R$, 存在 $-a \in R$ 使 $a + (-a) = (-a) + a = 0$.

④ $\forall a, b \in R, a + b = b + a$.

则称 R 关于运算 $(a, b) \mapsto a + b$ 成为一个交换群. 如果还满足:

⑤ $\forall a, b, c \in R, (a \cdot b) \cdot c = a \cdot (b \cdot c)$. (结合律).

⑥ 存在 $1 \in R$, 使 $1 \cdot a = a \cdot 1 = a, a \in R$ 成立. (单位元的存在性).

⑦ $\forall a,b,c \in R$, $(a+b)\cdot c = a\cdot c + b\cdot c$, $c\cdot(a+b) = c\cdot a + c\cdot b$. (分配律).

则称 R 是一个环. 如果它的乘法也满足交换律:

⑧ $\forall a,b,c \in R$, $a\cdot b = b\cdot a$.

则称 R 是一个交换环.

定义 1.2.4 设 K 是一个交换环, $K^* = K\setminus\{0\}$ 表示 K 中非零元的集合, 如果 K 中的乘法诱导出 K^* 上的运算 $K^*\times K^*\to K^*$, 使得 K^* 成为一个群, 则称 K 是一个域.

K 是一个域的定义隐含了下述条件: ① K^* 非空, 即 K 至少包含两个元素 0 和 1 且 $1\neq 0$. ② $\forall a,b\in K$, 如果 a,b 都不等于 0, 则 $a\cdot b\neq 0$. ③ $\forall a\in K$, 如果 $a\neq 0$, 则 a 可逆.

例 1.2.4 元素个数最小的域是 $K=\{0,1\}$ (二元域), 其中的运算定义为: $0+0=0$, $0+1=1+0=1$, $1+1=0$, $0\cdot 0=0\cdot 1=1\cdot 0=0$, $1\cdot 1=1$. 二元域一般用符号 $\mathbb{F}_2$ 表示, 在编码, 通讯和计算机领域有重要应用.

群, 环, 域不是本课程讨论的主要对象, 我们的主要研究对象是域 K 上的线性空间 (或称 K-线性空间, K-向量空间等). 如果对抽象的域不太习惯 (尤其是不太敢做加, 减, 乘, 除的运算), 在后面的学习中可以将 K 想象为 $\mathbb{Q}$, $\mathbb{R}$, $\mathbb{C}$ 或 $\mathbb{C}$ 中的一些子域 (后面将会介绍 $\mathbb{C}$ 中的子域). 下面的例子一般称为 n-维标准 K-向量空间 (或 K-线性空间).

例 1.2.5 设 K 是一个域, $V=K^n=\{(x_1,\cdots,x_n)\mid x_i\in K\}$. 则 K 上的 "加法" 和 "乘法" 诱导了 V 的两个运算:

(I) (加法) $\forall\alpha=(\alpha_1,\cdots,\alpha_n)$, $\beta=(\beta_1,\cdots,\beta_n)\in V$,

$$\alpha+\beta=(\alpha_1+\beta_1,\cdots,\alpha_n+\beta_n).$$

(II) (数乘) $\forall\lambda\in K$ (一般将 K 中元素称为"数"), $\forall\alpha=(\alpha_1,\cdots,\alpha_n)\in V$, 定义 "数乘运算"

$$\lambda\alpha=(\lambda\alpha_1,\lambda\alpha_2,\cdots,\lambda\alpha_n).$$

可以验证 V 上的这两个运算 (指 "加法" 和 "数乘运算") 满足下面 8 个条件 (简称 "八项规定").

① $\forall\alpha,\beta,\gamma\in V$, $(\alpha+\beta)+\gamma=\alpha+(\beta+\gamma)$. (结合律).

② 存在 $0\in V$, 使得对任意 $\alpha\in V$, 有 $\alpha+0=0+\alpha=\alpha$. (零元的存在性).

③ $\forall \alpha \in V$, 存在 $-\alpha \in V$ 使 $\alpha + (-\alpha) = (-\alpha) + \alpha = 0$. (负元的存在性)

④ $\forall \alpha, \beta \in V, \alpha + \beta = \beta + \alpha$.

⑤ $\lambda, \mu \in K, \alpha \in V$, 有 $(\lambda \cdot \mu)\alpha = \lambda \cdot (\mu\alpha)$.

⑥ $\forall \alpha \in V$, $1 \cdot \alpha = \alpha$.

⑦ $(\lambda + \mu) \cdot \alpha = \lambda \cdot \alpha + \lambda \cdot \alpha, \forall \lambda, \mu \in K, \alpha \in V$.

⑧ $\lambda \cdot (\alpha + \beta) = \lambda\alpha + \lambda\beta, \forall \lambda \in K, \alpha, \beta \in V$.

其中, 满足前四条称 V 关于 "加法" 是一个交换群, 满足后四条称 V 上保持 K-线性结构.

定义 1.2.5 设 K 是一个域, 加法群 V 的一个 K-向量空间结构 (或 K-线性空间结构) 是指一个映射

$$f : K \times V \to V$$

满足如下条件 ($\forall \lambda \in K, \alpha \in V$, 记 $f(\lambda, \alpha) = \lambda\alpha$):

① $(\lambda \cdot \mu)\alpha = \lambda \cdot (\mu\alpha)$, ② $1 \cdot \alpha = \alpha$,

③ $(\lambda + \mu) \cdot \alpha = \lambda \cdot \alpha + \mu \cdot \alpha$, ④ $\lambda \cdot (\alpha + \beta) = \lambda\alpha + \lambda\beta$.

加法群 V 称为一个 K-向量空间 (或 K-线性空间). 如果存在这样一个 "数乘" 映射 $K \times V \to V$.

注记: 不是所有的加法群都有 K-向量空间结构. 例如, 整数集合 $\mathbb{Z}$ 关于加法是一个交换群, 但是任何域 K 都不存在 K-向量空间结构. i.e. 不存在映射 $K \times \mathbb{Z} \to \mathbb{Z}$ 满足定义中的条件①-④.

定义 1.2.6 两个 K-向量空间 K^n 与 K^m 称为同构. 如果存在双射 $f : K^n \to K^m$ 满足:

(1) $\forall \alpha, \beta \in K^n,\ f(\alpha + \beta) = f(\alpha) + f(\beta)$,

(2) $\forall \lambda \in K, \alpha \in K^n,\ f(\lambda\alpha) = \lambda f(\alpha)$.

(这样的映射称为保运算的映射).

例 1.2.6 显然 $\mathbb{R}^2, \mathbb{R}$ 是两个 $\mathbb{R}$-向量空间, Cantor 证明了存在双射: $f : \mathbb{R}^2 \to \mathbb{R}$ (即: 作为集合 $\mathbb{R}^2$ 与 $\mathbb{R}$ 同构). 我们可以证明, 不存在保运算的双射 $f : \mathbb{R}^2 \to \mathbb{R}$. 否则, 令 $e_1 = (0, 1), e_2 = (0, 1) \in \mathbb{R}^2$, 得 $f(e_1) \neq f(e_2) \Rightarrow$ 存在 $\lambda_1, \lambda_2 \in \mathbb{R}$ 不全为 0 使 $\lambda_1 f(e_1) + \lambda_2 f(e_2) = 0 \Rightarrow f(\lambda_1 e_1 + \lambda_2 e_2) = \lambda_1 f(e_1) + \lambda_2 f(e_2) = 0$. 但 f 是单射, 所以 $\lambda_1 e_1 + \lambda_2 e_2 = 0 \Rightarrow \lambda_1 = \lambda_2 = 0$. 矛盾.

习题1.2

1. 证明两个 $\mathbb{R}$-向量空间 $\mathbb{R}^n$ 与 $\mathbb{R}^m$ 同构的充要条件是 $n = m$.

2. 逻辑中的命题 P 是指一段句子, 它要么是对的, 要么是错的. 令 P 表示所有命题 (proposition) 的集合, $B = \{0, 1\}$, 定义映射 $f : P \to B$ 如下:

$$p \in P,\ f(p) = \begin{cases} 1, \text{ 如果命题 } P \text{ 是对的}, \\ 0, \text{ 如果命题 } P \text{ 是错的}. \end{cases}$$

在逻辑中有下述运算: $\forall p, q \in \mathrm{P}$, 令 $p \vee q \in P$ 表示命题: “p 或者 q”, 而 $p \wedge q$ 表示命题: “p 和 q”. 所以集合 P 上有运算:

$$P \times P \to P, (p, q) \mapsto p \vee q,\ P \times P \to P, (p, q) \mapsto p \wedge q.$$

请确定集合 $B = \{0, 1\}$ 的两个运算 “加法” 和 “乘法” 使得

$$\forall p, q \in P,\ f(p \vee q) = f(p) + f(q),\ f(p \wedge q) = f(p) \cdot f(q).$$

3. 设 S 是一个集合, $P(S)$ 是 S 的幂集 (i.e. S 中所有子集合的集合). 定义集合 $P(S)$ 的 “加法” 和 “乘法” 如下: $\forall A, B \in P(S)$,

$$A + B = A \cup B \setminus A \cap B,\ A \cdot B = A \cap B.$$

证明: (a) 带有上述运算的 $P(S)$ 是一个环.

(b) $P(S)$ 中的零元, 单位元分别是什么?

(c) 什么时候 $P(S)$ 的零元与单位元相等?

4. 设 K 是一个域, $e_1 = (1, 0, \cdots, 0)$, $e_2 = (0, 1, 0, \cdots, 0), \cdots$, $e_n(0, \cdots, 0, 1)$ 是 K^n 中的元素. 证明:

(a) $\lambda_1 e_1 + \lambda_2 e_2 + \cdots + \lambda_n e_n = 0 \Leftrightarrow \lambda_1 = \lambda_2 = \cdots \lambda_n = 0$.

(b) $\forall x \in K^n$, 存在唯一的 $x_1, \cdots, x_n \in K$ 使

$$x = x_1 e_1 + x_2 e_2 + \cdots + x_n e_n.$$

5. 设 $f : K^n \to K$ 是保持运算的映射, 即: $\forall \alpha, \beta \in K^n$ 和 $\lambda \in K$, 有 $f(\alpha + \beta) = f(\alpha) + f(\beta)$, $f(\lambda\alpha) = \lambda f(\alpha)$.

证明: 存在常数 $a_1, a_2, \cdots, a_n \in K$, 使得对任意 $x = (x_1, x_2, \cdots, x_n) \in K^n$ 有 $f(x) = a_1 x_1 + a_2 x_2 + \cdots + a_n x_n$.

1.3 整数环与多项式环

我们首先复习有关整数环 $\mathbb{Z}$ 的基本算术性质. 对任意 $a, b \in \mathbb{Z}$, 我们称 a 整除 b (记为 $a|b$). 如果存在 $c \in \mathbb{Z}$ 使 $b = a \cdot c$, 此时 a, c 也称为 b 的因子, b 称为 a, c 的倍数. 如果 a 不整除 b, 我们用符号 $a \nmid b$ 表示.

定义 1.3.1 大于 1 的整数 $P \in \mathbb{Z}$ 称为素数, 如果 P 只能被 ± 1 和 $\pm p$ 整除.

在整数环 $\mathbb{Z}$ 中, 最重要的性质应该是下面的算术基本定理.

定理 1.3.1 每一个大于 1 的整数 $a \in \mathbb{Z}$ 都可以唯一写成

$$a = p_1 \cdot p_2 \cdots p_s,\ p_1, p_2, \cdots, p_s \text{是素数 (可以相同)}.$$

证明: 根据素数的定义, 很容易证明每一个大于 1 的整数可以分解成素数的乘积. 该定理的重要性是断言这样的分解是唯一的, 即如果 $a = p_1 p_2 \cdots p_s = q_1 q_2 \cdots q_t$, 其中 p_i, q_i 是素数, 则该定理断言: ① $s = t$, ② 适当交换 $q_1, q_2, \cdots, q_t$ 的次序, 可使 $q_1 = p_1, q_2 = p_2, \cdots, q_t = p_t$. 我们可以证明定理 1.3.1 与下面的定理等价. □

定理 1.3.2 设 p 是素数, 如果 $p|ab$, 则 $p|a$ 或 $p|b$ (i.e. 如果 $p|ab$, 则 p 一定整除 a, b 之一).

显然, 定理 1.3.2 推出定理 1.3.1. 事实上, 如果 $a = p_1 p_2 \cdots p_s = q_1 q_2 \cdots q_t$, 则 p_1 一定整除 $q_1, q_2, \cdots, q_t$ 中的一个. 重新调整 $q_1, q_2, \cdots, q_t$ 的次序, 可设 $p_1|q_1$, 但 p_1, q_1 是素数, $p_1|q_1$ 推出 $p_1 = q_1$. 从而 $p_2 \cdots p_s = q_2 \cdots q_t$. 继续上述论证, 可得 $s = t$, $p_2 = q_2, \cdots, p_s = q_t$. 如果定理 1.3.1 成立, 则不难证明定理 1.3.2. 设 $a = p_1 \cdots p_s$, $b = q_1 \cdots q_t$ 是 a, b 的不可约分解 (我们称定理 1.3.1 的分解 $a = p_1 \cdots p_s$ 为 a 的不可约分解). 则 $ab = p_1 \cdots p_s q_1 \cdots q_t$ 是 ab 的不可约分解. 有 $p|ab$, 得 $ab = p \cdot c$. 令 $c = h_1 h_2 \cdots h_m$ 是 c 的不可约分解, 则 $ab = p \cdot h_1 h_2 \cdots h_m$ 是 ab 的不可约分解. 由定理 1.3.1 分解的唯一性, p 一定是 $p_1, \cdots, p_s$ 或 $q_1, \cdots, q_t$ 中的一个, 所以 $p|a$ 或 $p|b$. 我们证明了定理 1.3.1 与定理 1.3.2 的等价性. 为了证明定理 1.3.2, 我们回忆有关 $\mathbb{Z}$ 的一些基本概念.

定义 1.3.2 设 $a, b \in \mathbb{Z}$, $d \in \mathbb{Z}$ 称为 a, b 的最大公因子, 如果 ① $d|a, d|b$, ② 对任意公因子 c (i.e. $c|a, c|b$), 有 $c|d$.

显然, 如果 d 是 a,b 的最大公因子, 则 $-d$ 也是 a,b 的最大公因子. 为了方便, 我们用 (a,b) 表示大于零的最大公因子, 称为最大公因数 (它确是公因子里最大的). 如果 $(a,b)=1$, 则称 a,b 互素. 我们发现, 下面显而易见的引理实际上反映了整数环 $\mathbb{Z}$ 的一个重要性质.

引理 1.3.1 (带余除法) $\forall a, b \neq 0 \in \mathbb{Z}$, 存在唯一的 $q, r \in \mathbb{Z}$ 使得 $a = qb + r$, 其中 $a \leq r < |b|$.

证明: 令 $S = \{a - xb \mid x \in \mathbb{Z}, a - xb \geq 0\}$ (验证: S 是一个非空集合!). 则 S 中有一个最小整数 $r \in S$, 即存在 $q \in \mathbb{Z}$ 使 $r = a - qb \geq 0$ 是 S 中最小整数, 且 $r < |b|$ (否则, $r - |b| = a - qb - |b| \in S$, 且 $r - |b| < r$ 与 r 的最小性矛盾).

如果 $a = qb + r = q_1 b + r_1$ 且 $a \leq r_1 < |b|$. 则 $(q - q_1)b = r_1 - r$. 所以必有 $q = q_1, r_1 = r$ (否则 $|r_1 - r| \geq |b|$ 矛盾). □

利用带余除法, 我们可以证明:

定理 1.3.3 设 $a, b \in \mathbb{Z}$ 不全为零, 则存在最大公因数 (a,b) 且存在 $x, y \in \mathbb{Z}$ 使 $(a,b) = ax + by$.

证明: 考虑集合 $S = \{ax + by \mid \forall x, y \in \mathbb{Z}\}$, 它显然非空且包含正整数 (例如 $a \cdot a + b \cdot b = a^2 + b^2 \in S$). 令 $d \in S$ 是最小正整数, 则 d 整除 S 中的每一个元素. 否则, 存在 $ax + by \in S$, $d \nmid (ax + by)$. 则由带余除法, 存在 $q, r \in \mathbb{Z}$ 使

$$ax + by = q \cdot d + r,\ 0 < r < d.$$

但 $r = ax + by - q \cdot d \in S$ 与 d 是 S 中最小正整数矛盾. 所以存在 $x, y \in \mathbb{Z}$ 使 $d = ax + by$, 且 $d|a$, $d|b$. 对任意 a, b 的公因数 c 必有 $c|d$. 所以 d 是 a, b 的最大公因数. □

注记: $\forall a, b \in \mathbb{Z}$, 如果 $(a,b) = 1$, 则称 a, b 互素. 对于任意素数 p 和任意整数 $a \in \mathbb{Z}$, 如果 $p \nmid a$, 则 $(p,a) = 1$.

定理 1.3.2 的证明: 设 p 是素数, 且 $p|ab$, 如果 $p \nmid a$, 则 $(p,a) = 1$. 由定理 1.3.3, 存在 $x, y \in \mathbb{Z}$, 使 $1 = px + ay$, 所以 $b = pxb + aby$. 但 $p|ab$, 从而 $p|(pxb + aby) = b$.

我们下面引入一个与整数环非常类似的环, 称为域上的多项式环. 事实上, 在数学上经常将它与 $\mathbb{Z}$ 进行比较. 设 K 是一个域 (对不习惯抽象运算的同学, 可以将 K 看成 $\mathbb{Q}$, 或 $\mathbb{R}$, 或 $\mathbb{C}$).

设 x 是一个不定元 (也可以用 t, λ, y, z 等任何符号表示), $ax^m (a \in K)$ 称为一个 m 次的单项式, a 称为它的系数, 而有限个单项式的和称为多项式, 其中相同次数的单项式可以合并同类项: $ax^m + bx^m$ 可写成 $(a+b)x^m$. 而且单项式相加与它们的次序无关. 所以, 任何一个多项式都可以写成 $f(x) = a_n x^n + a_{n-1} x^{n-1} + \cdots + a_1 x + a_0$, 其中 $a_n \neq 0$. $a_n x^n$ 称为 $f(x)$ 的首项, a_n 称为 $f(x)$ 的首项系数, n 称为 $f(x)$ 的次数, 记为 $deg(f(x))$ (或 $deg(f)$). 当 $m = 0$ 时, ax^m 就是 a. 所以, a_0 称为 $f(x)$ 的零次项, $a_0, a_1, \cdots, a_n$ 称为 $f(x)$ 的系数.

我们将用 $K[x] = \{a_n x^n + a_{n-1} x^{n-1} + \cdots + a_1 x + a_0 \mid n \in \mathbb{Z}, a_i \in K\}$ 表示所有多项式的集合 ($K[x]$ 中所有零次多项式的集合与 K 等同, 称为常值多项式). 有时我们将一个多项式 $f(x) = a_n x^n + a_{n-1} x^{n-1} + \cdots + a_1 x + a_0$ 写成 $f(x) = a_0 + a_1 x + \cdots + a_{n-1} x^{n-1} + a_n x^n$. 在 $K[x]$ 中有一个显然的 "加法": $f(x) + g(x)$ 是通过合并 $f(x)$ 和 $g(x)$ 的同类项而得的多项式. (注意: $f(x)$ 称为零多项式, 如果 $f(x)$ 的所有系数都为零, 记成 $f(x) \equiv 0$, 它与 K 中的零元等同). 多项式的乘法定义如下:

$\forall f(x) = a_n x^n + a_{n-1} x^{n-1} + \cdots + a_1 x + a_0, g(x) = b_m x^m + b_{m-1} x^{m-1} + \cdots + b_1 x + b_0 \in K[x]$, 定义: $f(x)g(x) = c_{m+n} x^{m+n} + c_{m+n-1} x^{m+n-1} + \cdots + c_1 x + c_0$, 其中 $c_k = \sum\limits_{i+j=k} a_i b_j = a_0 b_k + a_1 b_{k-1} + \cdots + a_k b_0$.

定理 1.3.4 $K[x]$ 关于上述的 "加法" 与 "乘法" 是一个交换环 (称为多项式环). 即:

(1) $\forall f(x), g(x), h(x) \in K[x]$, $(f(x) + g(x)) + h(x) = f(x) + (g(x) + h(x))$.

(2) 存在 $0 \in K[x]$, $f(x) + 0 = 0 + f(x)$.

(3) $\forall f(x) = a_n x^n + a_{n-1} x^{n-1} + \cdots + a_1 x + a_0 \in K[x]$, 则

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

是 $f(x)$ 的负元 (记为: $-f(x) = -a_n x^n - a_{n-1} x^{n-1} + \cdots - a_1 x - a_0$).

(4) $f(x) + g(x) = g(x) + f(x)$.

(5) $(f(x) \cdot g(x))h(x) = f(x)(g(x) \cdot h(x))$.

(6) $1 \in K$ 使 $1 \cdot f(x) = f(x) \cdot 1 = f(x)$.

(7) $f(x) \cdot g(x) = g(x) \cdot h(x)$.

(8) $(f(x) + g(x)) \cdot h(x) = f(x)h(x) + g(x)h(x)$,

$h(x) \cdot (f(x) + g(x)) = h(x)f(x) + h(x)g(x)$.

满足条件 (1)-(4) 称 $K[x]$ 关于 "加法" 是交换群.

证明是直接验证, 根据定义, 单项式 a_ix^i 与 b_jx^j 的乘积是 $(a_ix^i)\cdot(b_jx^j) = a_ib_jx^{i+j}$, 而 $f(x)\cdot g(x)$ 由 $f(x)$ 的单项式 a_ix^i 与 $g(x)$ 的单项式 b_jx^j 相乘, 然后合并同类项所得的多项式. 特别, $f(x)\cdot g(x)$ 的首项是 $a_nb_mx^{n+m}$. 所以我们有下面简单但重要的引理.

引理 1.3.2 任意非零多项式 $f(x), g(x) \in K[x]$, 则 $f(x)\cdot g(x)$ 是非零多项式, 且

$$deg(f(x)g(x)) = deg(f(x)) + deg(g(x)).$$

如果 $f(x)+g(x)$ 非零, 则 $deg(f(x)+g(x)) \le max\{deg(f), deg(g)\}$.

证明是显然的, 我们省略.

对任意 $f(x), g(x) \in K[x]$, 如果存在 $h(x) \in K[x]$ 使

$$f(x) = g(x)\cdot h(x),$$

则称 $g(x)$ 整除 $f(x)$, 记为 $g(x)|f(x)$ (否则, 记为 $g(x) \nmid f(x)$), 其中 $g(x), h(x)$ 称为 $f(x)$ 的因子. 任何多项式 $f(x)$ 都可以被自己和非零的零次多项式 (即: $f(x)$ 和 K 中非零元) 整除, 所以 $f(x)$ 和 K 中非零元称为 $f(x)$ 的平凡因子.

定义 1.3.3 $p(x) \in K[x]$ 称为不可约 (irreducible) 多项式, 如果它满足: ① $deg(p(x)) > 0$, ② $p(x)$ 没有非平凡因子 (即: $p(x)$ 不能写成两个次数大于零的多项式的乘积).

定理 1.3.5 (唯一分解定理) 设 $f(x) \in K[x]$ 是一首项系数为 1 (简称 "首系 1 ") 的多项式, 且 $deg(f(x)) > 0$. 则

(1) 存在首项系数为 1 的不可约多项式 $p_1(x), \cdots, p_s(x)$ 使

$$f(x) = p_1(x)\cdots p_s(x), \text{其中 } p_1(x), \cdots, p_s(x) \text{ 可以相同.}$$

(2) 如果 $f(x) = p_1(x)\cdots p_s(x) = q_1(x)\cdots q_t(x)$, 其中 $q_i(x)$ 也是首系 1 不可约多项式, 则 $s = t$, 且(适当调整次序) $p_i(x) = q_i(x)$.

证明: (1) 的证明很显然, 如果 $f(x)$ 不可约, 则 $s = 1$, $p_1(x) = f(x)$. 如果 $f(x)$ 可约, 则 $f(x) = f_1(x)\cdot f_2(x)$, $deg(f_1) > 0$, $deg(f_2) > 0$, 所以 $deg(f_1) < deg(f)$, $deg(f_2) < deg(f)$, 对 $f_i(x)$ 的次数 $deg(f_i)$ 做数学归纳法, 可得 (1).

(2) 与整数环 $\mathbb{Z}$ 类似, 分解的唯一性与下面的定理等价. □

定理 1.3.6 设 $p(x) \in K[x]$ 不可约, 如果 $p(x)|f(x)\cdot g(x)$, 则 $p(x)|f(x)$ 或者 $p(x)|g(x)$.

为了证明定理 1.3.6, 我们需要类似的带余除法.

引理 1.3.3 设 $f(x), g(x) \in K[x]$, 如果 $g(x) \not\equiv 0$, 则存在唯一 $q(x), r(x) \in K[x]$ 满足 (1) $f(x) = q(x)g(x) + r(x)$, (2) $r(x) \equiv 0$ 或 $deg(r(x)) < deg(g(x))$.

证明: 令 $S = \{f(x) - h(x)g(x) \mid \forall h(x) \in K[x]\}$. 如果 $0 \in S$ (i.e. $g(x)|f(x)$), 则取 $r(x) \equiv 0$ 即可, 否则对任意 $h(x) \in K[x]$, $f(x) - h(x)g(x) \not\equiv 0$, 它们的次数是非负整数, 令 $r(x) \in S$ 是 S 中次数最小的多项式, 则存在 $q(x) \in K[x]$ 使 $r(x) = f(x) - q(x)g(x)$. 下面我们证明: $deg(r(x)) < deg(g(x))$. 令 $g(x) = b_m x^n + b_{m-1}x^{m-1} + \cdots + b_1 x + b_0, r(x) = c_d x^d + c_{d-1}x^{d-1} + \cdots + c_1 x + c_0$, 其中 $b_m \neq 0$, $c_d \neq 0$. 如果 $deg(r(x)) = d \geq deg(g(x)) = m$, 则我们可消去 $r(x)$ 的首项, 所以 $r(x) - c_d \cdot b_m^{-1} x^{d-m} \cdot g(x)$ 的次数小于 $d = deg(r(x))$. 但 $r(x) - c_d \cdot b_m^{-1} x^{d-m} \cdot g(x) = f(x) - (q(x) + c_d \cdot b_m^{-1} x^{d-m})g(x)$ 在 S 中与 $r(x)$ 是 S 中次数最小的多项式矛盾. 所以, $deg(r(x)) < deg(g(x))$. 我们证明了 $q(x), r(x)$ 的存在性. 唯一性证明如下: 如果存在另外的 $q_1(x), r_1(x)$ 也满足: $f(x) = q_1(x)g(x) + r_1(x)$, 且 $r_1(x) \equiv 0$ 或 $deg(r_1(x)) < deg(g(x))$, 则得 $(q(x) - q_1(x))g(x) = r_1(x) - r(x)$. 如果 $q(x) - q_1(x) \not\equiv 0$ (i.e. $q(x) \neq q_1(x)$), 则 $deg(r_1(x) - r(x)) \geq deg(g(x))$. 但 $deg(r_1(x) - r(x)) \leq max\{deg(r_1(x)), deg(r(x))\} < deg(g(x))$, 得到矛盾. □

带余除法 (引理 1.3.3) 有另一个构造性证明, 对给定的

$$f(x) = a_n x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0 \ (a_n \neq 0),$$

$$g(x) = b_m x^n + b_{m-1}x^{m-1} + \cdots + b_1 x + b_0 \ (b_m \neq 0),$$

我们可以给一种"算法", 求出满足引理 1.3.3 条件的 $q_(x)$ 与 $r(x)$, 称为" 消去首项法":

① 如果 $deg(f(x)) < deg(g(x))$, 令 $q(x) = 0, r(x) = f(x)$.

② 如果 $deg(f(x)) \geq deg(g(x))$, 令 $f_1(x) = f(x) - a_n b_m^{-1} x^{n-m} g(x)$, 则 $deg(f_1(x)) < deg(f(x))$. 如果 $deg(f_1(x)) < deg(g(x))$, 可得 $r(x) = f_1(x), q(x) = a_n b_m^{-1} x^{n-m}$.

③ 如果 $deg(f(x)) \geq deg(g(x))$, 继续第 ② 步, 消耗 $f_1(x)$ 的首项.

定义 1.3.4 设 $f(x), g(x) \in K[x]$, $d(x) \in K[x]$ 称为 $f(x), g(x)$ 的最大公因子, 如果 $d(x)$ 满足:

(1) $d(x)|f(x), d(x)|g(x)$, (2) 如果 $c(x)$ 是 $f(x), g(x)$ 的公因子, 则 $c(x)|d(x)$.

注记: 如果 $d_1(x), d_2(x)$ 都是 $f(x), g(x)$ 的最大公因子, 则 $d_1(x)|d_2(x)$, $d_2(x)|d_1(x)$, 所以 $d_1(x) = d_2(x) \cdot a$ ($a \in K$ 非零).

我们将用 $(f(x), g(x))$ 表示 $f(x), g(x)$ 的首项系数为 1 的最大公因子. 特别, $f(x), g(x)$ 称为互素的, 如果 $(f(x), g(x)) = 1$.

定理 1.3.7 设 $f(x), g(x) \in K[x]$ 不全为零, 则存在最大公因子 $d(x) = (f(x), g(x))$, 且存在 $u(x), v(x) \in K[x]$ 使

$$d(x) = (f(x), g(x)) = u(x)f(x) + v(x)g(x).$$

证明: 令 $S = \{u(x)f(x) + v(x)g(x) \mid \forall u(x), v(x) \in K[x]\}$, 则 S 中含非零多项式. 令 $d(x) \in S$ 是 S 中次数最小的首项系数为 1 的多项式, 则存在 $u(x), v(x) \in K[x]$ 使

$$d(x) = u(x)f(x) + v(x)g(x).$$

我们只需证明 $d(x)$ 是 $f(x), g(x)$ 的最大公因子. 上述等式表明: 如果 $c(x)$ 是 $f(x), g(x)$ 公因子 (i.e. 同时整除 $f(x)$ 和 $g(x)$), 则 $c(x)|d(x)$. 所以, 我们只需证明 $d(x)$ 是 $f(x), g(x)$ 的公因子. 由带余除法 (引理 1.3.3), 存在 $q_1(x), q_2(x), r_1(x), r_2(x)$, 使

$$f(x) = q_1(x)d(x) + r_1(x), \ g(x) = q_2(x)d(x) + r_2(x).$$

如果 $r_1(x)$, $r_2(x)$ 不全为零, 不妨设 $r_1(x) \neq 0$, 则 $deg(r_1(x)) < deg(d(x))$. 但 $r_1(x) = f(x) - q_1(x)d(x) = f(x) - q_1(x)u(x)f(x) - v(x)g(x) \in S$ 与 $d(x)$ 的选取矛盾. □

定理 1.3.6 的证明: 设 $p(x)$ 不可约, 且 $p(x)|f(x)g(x)$. 如果 $p(x) \nmid f(x)$, 则 $(p(x), f(x)) = 1$. 所以, 存在 $u(x), v(x)$ 使 $1 = u(x)p(x) + v(x)f(x)$, 得 $g(x) = g(x)u(x)p(x) + v(x)f(x)g(x)$. 由 $p(x)|f(x)g(x)$, 得 $p(x)|g(x)$.

带余除法还有一个很重要的推论. 对任意 $a \in K$ 和 $f(x) \in K[x]$, $f(a) \in K$ 表示多项式 $f(x)$ 在 $x = a$ 时的取值. 如果 $f(a) = 0$, 则称 a 是方程 $f(x) = 0$ 的根 (或简称 $f(x)$ 的根).

定理 1.3.8 如果 $a \in K$ 是 $f(x)$ 的根, 则 $(x - a)|f(x)$.

证明: 存在 $q(x), r(x)$ 使 $f(x) = q(x) \cdot (x - a) + r(x)$, 其中 $r(x) = 0$ 或非零常数. 但 $0 = f(a) = q(a) \cdot (a - a) + r(a) = r(a)$, 所以 $r(x) = 0$, 即 $f(x) = (x - a) \cdot q(x)$. □

习题1.3

1. 设 $f(x) \in \mathbb{Q}[x], a \in \mathbb{R}$, $f(a) \in \mathbb{R}$ 表示 $f(x)$ 在 $x = a$ 的取值. 证明: $f(a) = 0 \Leftrightarrow$ 存在 $q(x) \in \mathbb{R}[x]$, 使 $f(x) = (x - a)q(x)$ (即在 $\mathbb{R}[x]$ 中, $(x - a)|f(x)$).

2. 证明: $f(x)=x^2-2$ 在 $\mathbb{Q}[x]$ 中不可约, 但在 $\mathbb{R}[x]$ 中可约; $f(x)=x^2+1$ 在 $\mathbb{R}[x]$ 中不可约, 但在 $\mathbb{C}[x]$ 中可约.

3. 对下列多项式 $f(x),g(x)$ 求带余除法中的 $q(x),r(x)$ 使 $f(x)=q(x)g(x)+r(x)$.

(1) $f(x)=2x^4-3x^2+4x^2-5x+6$, $g(x)=x^2-3x+1$.

(2) $f(x)=x^3-3x^2-x-1$, $g(x)=3x^2-2x+1$.

4. 本题给出求 $f(x),g(x)$ 最大公因子的算法 (称辗转相除法, 或欧几里得算法).

(1) 用 $g(x)$ 除 $f(x)=q_1(x)g(x)+r_1(x)$, 如果 $r_1(x)=0$, 证明 $g(x)$ 是 $f(x)$ 与 $g(x)$ 的最大公因子.

(2) 如果 $r_1(x)\neq 0$, 用 $r_1(x)$ 除 $g(x)$ 得 $g(x)=q_2(x)r_1(x)+r_2(x)$; 如果 $r_2(x)=0$, 证明 $r_1(x)$ 是 $f(x),g(x)$ 的最大公因子, 且 $r_1(x)=f(x)-q_1(x)g(x)$.

(3) 如果 $r_2(x)\neq 0$, 用 $r_2(x)$ 除 $r_1(x)$ 得 $r_1(x)=q_3(x)r_2(x)+r_3(x)$; 如果 $r_3(x)=0$, 终止计算. 否则, 继续做带余除法, 可得一系列等式:

$$\begin{aligned}
f(x)&=q_1(x)g(x)+r_1(x),\\
g(x)&=q_2(x)r_1(x)+r_2(x),\\
r_1(x)&=q_3(x)r_2(x)+r_3(x), &&\text{其中 } deg r_1(x)>deg r_2(x)>\cdots>deg r_n(x)\\
r_2(x)&=q_4(x)r_3(x)+r_4(x), &&r_{n+1}(x)=0.\\
&\vdots\\
r_{n-2}(x)&=q_n(x)r_{n-1}(x)+r_n(x),\\
r_{n-1}(x)&=q_{n+1}(x)r_n(x)+r_{n+1}(x),
\end{aligned}$$

证明: $r_n(x)$ 是 $f(x),g(x)$ 的最大公因子 (该算法实际提供了 $u(x),v(x)$ 使 $r_n(x)=u(x)f(x)+v(x)g(x)$).

5. 求下列 $f(x)$ 和 $g(x)$ 的最大公因子, 并求 $u(x),v(x)$ 使 $(f,g)=u(x)f(x)+v(x)g(x)$.

(1) $f(x)=x^4+x^3-3x^2-4x-1$, $g(x)=x^3+x^2-x-1$;

(2) $f(x)=x^5+3x^2-2x+2$, $g(x)=x^6+x^5+x^4-3x^2+2x-6$;

(3) $f(x)=x^4-10x^2+1$, $g(x)=x^4-4\sqrt{2}x^3+6x^2+4\sqrt{2}x+1$.

6. 令 $F(\mathbb{R})=\{\varphi:\mathbb{R}\to\mathbb{R}\mid \varphi\text{是映射}\}$ 表示所有从 $\mathbb{R}$ 到 $\mathbb{R}$ 映射的集合. 对任意 $\lambda\in\mathbb{R},\psi\in F(\mathbb{R})$ 定义:

(1) $\varphi+\psi:\mathbb{R}\to\mathbb{R}$, $a\mapsto\varphi(a)+\psi(a)$.

(2) $\varphi\cdot\psi:\mathbb{R}\to\mathbb{R}$ 表示映射合成.

(3) $\lambda\varphi : \mathbb{R} \to \mathbb{R}, a \mapsto \lambda\varphi(a)$, 其中 $\varphi(a) + \psi(a)$, $\lambda\varphi(a)$ 就是实数的加法和乘法. 对任意多项式

$$f(x) = a_n x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0 \in \mathbb{R}[x].$$

及 $\varphi \in F(\mathbb{R})$, $f(\varphi) = a_n\varphi^n + a_{n-1}\varphi^{n-1} + \cdots + a_1\varphi + a_0 \in F(\mathbb{R})$, 其中 $\varphi^k := \overbrace{\varphi \cdot \varphi \cdots \varphi}^{k}$ 表示 k 个映射 φ 的合成, a_0 表示常值映射. 对下列的 $f(x)$, 和 $\varphi : \mathbb{R} \to \mathbb{R}$, 求 $f(\varphi)$:

(1) $f(x) = 2x^3 + x^2 - 3x + 7$, $\varphi : \mathbb{R} \to \mathbb{R}, a \mapsto a + 1$.

(2) $f(x) = x^2 - 3x + 2$, $\varphi : \mathbb{R} \to \mathbb{R}, a \mapsto 2a$.

7. 设 $f(x) \in K[x]$, $\alpha, \beta \in K$ 是 $f(x)$ 的根, 且 $\alpha \neq \beta$. 证明存在 $h(x) \in K[x]$ 使 $f(x) = (x - \alpha)(x - \beta) \cdot h(x)$.

8. 设 $f(x) \in K[x]$ 中首项系数为 1 的 n 次多项式. 如果 $\alpha_1, \alpha_2, \cdots, \alpha_n \in K$ 是 $f(x)$ 的 n 个不同的根. 证明

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

9. 设 $f(x) \in K[x]$, $\alpha \in K$ 是 $f(x)$ 的根. 证明: 存在 m 和 $h(x) \in K[x]$ 使得 $f(x) = (x - \alpha)^m h(x)$ 且 $h(x) \neq 0$.

10. 证明习题 9 中的 m 是唯一的 (即: 如果 $f(x) = (x - \alpha)^{m_1} h_1(x)$ 且 $h_1(x) \neq 0$, 则 $m_1 = m$). 此时, m 称为根 α 的重数.

11. 习题 8 中的分解去掉条件“$\alpha_1, \alpha_2, \cdots, \alpha_n$ 两两不同”是否仍然成立? 证明你的结论.

1.4 复数域及其子域

人类对数的认识经历了漫长的历史, 从有理数到实数, 虽然 “无理” 但毕竟真是存在, 所以命名实数 (it is real) . 但人们在解方程的时发现了实数并不够用, 负数的平方数首先出现在 1545 年. 但当时人们 (包括很多大数学家) 认为它是想象中的数 (imaginary number) , 是虚无缥缈的. 要消除虚数的虚无感, 最好莫过于给它及它们的运算以几何解释, 用已知的实数 (已被广泛接受) 来描述虚数, 这一过程花了近二百年.

在中学我们已经知道, 用 i 表示 $\sqrt{-1}$, 且所有形如 $a + b\mathrm{i}$ 的数称为复数 (complex number), a, b 分别称为 $a + b\mathrm{i}$ 的实部和虚部, $b\mathrm{i}$ 称为纯虚数. 令

$$\mathbb{C} = \{x + y\mathrm{i} \mid x, y \in \mathbb{R}\}$$

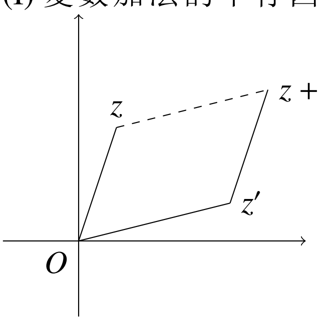
表示所有复数的全体, 其中两个复数 $x+y\mathrm{i}$ 与 $a+b\mathrm{i}$ 相等的定义是 $x = a, y = b$. 在中学课程里, 我们已经在集合 $\mathbb{C}$ 上定义了两种运算 (分别称为 “ 加法” 和 “乘法”), 如果 $z = a + b\mathrm{i}, \omega = c + d\mathrm{i} \in \mathbb{C}$, 则定义

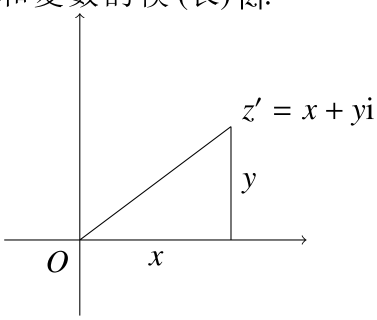
$$z + \omega = (a + c) + (b + d)\,\mathrm{i}$$

$$z \cdot \omega = (ac - bd) + (bc + ad)\,\mathrm{i},$$

不难验证, 这两个运算满足 “域” 定义中的公理, 其中 a, b, c, d 的加法和乘法就是实数的加法和乘法. 称为复数域. 我们已经习惯将实数 a 与 $a + 0\mathrm{i}$ 等同, 所以 $\mathbb{R}$ 可以看成 $\mathbb{C}$ 的子集合, 且 $\mathbb{C}$ 上的运算在 $\mathbb{R}$ 与实数的运算重合. 在平面 Ω 上, 选取坐标系后, Ω 上的点可由坐标 (x, y) 确定. 我们不妨认为 Ω 中的点由一个复数 $z = x + y\mathrm{i}$ 确定, 因此, 可以试图对复数的运算给出几何解释.

(I) 复数加法的平行四边形法则和复数的模 (长) $|z|$.



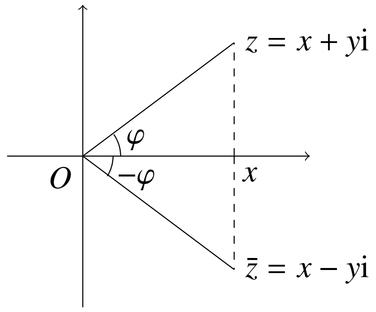


$$|z| = \sqrt{x^2 + y^2} = |Oz|$$

(II) 复数的辐角和复数的乘法.

对任意的复数 $z = x + y\mathrm{i}$, 从实轴到射线 $\overrightarrow{Oz}$ 的夹角称为 z 的辐角, 记为

arg z (按逆时针方向旋转得到的角度为正值, 顺时针方向所得角度为负值) . 如图.



其中 $\bar{z} = x - y\mathrm{i}$ 称为 $z = x + y\mathrm{i}$ 的共轭.

注意, 对任意整数 $m \in \mathbb{Z}$, $\arg z + 2\pi m$ 仍是 z 的辐角. 所以复数有另一个表达形式 (三角形式) .

$$z = |z|\,(\cos\varphi + \mathrm{i}\sin\varphi)\,.$$

不难证明:

定理 1.4.1 设 $z = |z|\,(\cos\varphi + \mathrm{i}\sin\varphi)$, $z' = |z'|\,(\cos\varphi' + \mathrm{i}\sin\varphi')$. 则:

$$z'z = |z'|\,|z|\,(\cos(\varphi + \varphi') + \mathrm{i}\sin(\varphi + \varphi'))$$

$$\frac{z}{z'} = \frac{|z|}{|z'|}\,(\cos(\varphi - \varphi') + \mathrm{i}\sin(\varphi - \varphi'))$$

特别, 我们有如下的棣莫弗公式

$$z^n = |z|^n\,(\cos n\varphi + \mathrm{i}\sin n\varphi)\,.$$

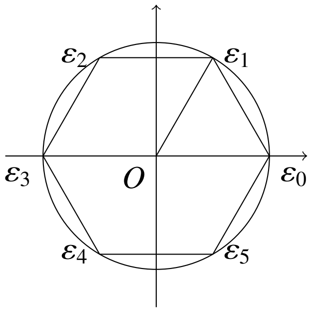
由棣莫弗公式, 我们知道任意复数 z 都有 n 次方根: 设 $x = |x|\,(\cos\theta + \mathrm{i}\sin\theta)$ 是 $z = |z|\,(\cos\theta + \mathrm{i}\sin\theta)$ 的一个 n 次方根, 即$x^n = z$. 所以

$$x^n = |x^n|\,(\cos\theta + i\sin\theta) = |z|\,(\cos\theta + \mathrm{i}\sin\theta)\,.$$

从而 $|x| = \sqrt[n]{|z|}, n\theta = \varphi + 2\pi k$. 所以, 如果 $z \neq 0$ (i.e. $|z| \neq 0$), 则 z 有 n 个 n 次方根

$$\sqrt[n]{|z|}\left(\cos\frac{\varphi + 2\pi k}{n} + \mathrm{i}\sin\frac{\varphi + 2\pi k}{n}\right), k = 0, 1, 2, \cdots, n - 1.$$

特别, $\varepsilon_k = \cos\frac{2\pi k}{n} + \mathrm{i}\sin\frac{2\pi k}{n}, (1 \leqslant k \leqslant n - 1)$ 是 1 的 n 个 n 次方根 (称为 n 次单位根), i.e. 多项式 $f(x) = x^n - 1 \in \mathbb{Q}[x]$ 在 $\mathbb{C}$ 中有 n 个根 $\varepsilon_0, \varepsilon_1, \cdots, \varepsilon_{n-1}$. 几何上, 它们位于单位圆的 n 个等分点上, 形成正 n-边形的 n 个顶点. 如图 ($n = 6$ 时)



下面我们讨论复数域 $\mathbb{C}$ 的代数模型, 仅仅实数域 $\mathbb{R}$ 出发构造一个与 $\mathbb{C}$ 同构的域. 由上面讨论的启发, 有一个集合间的双射. $f:\mathbb{R}^2\to\mathbb{C}$, $(x,y)\mapsto x+y\mathrm{i}$. 我们可以在 $\mathbb{R}^2$ 上直接定义运算:

$$(a,b)+(c,d)=(a+c,b+d)$$

$$(a,b)\cdot(c,d)=(ac-bd,ad+bc)$$

可以验证 $\mathbb{R}^2$ 关于上述运算是一个域, 而且双射 $f:\mathbb{R}^2\to\mathbb{C}$ 保持运算:

$$f((a,b)+(c,d))=f((a,b))+f((c,d))$$

$$f((a,b)\cdot(c,d))=f((a,b))\cdot f((c,d))$$

所以 $\mathbb{R}^2$ 在上述定义的运算下是一个同构于 $\mathbb{C}$ 的域. 当然, 这一构造显得太人工化. 显然乘法 $(a,b)\cdot(c,d)$ 的定义是通过双射 f 将 $\mathbb{C}$ 中的乘法 $(a+b\mathrm{i})\cdot(c+d\mathrm{i})=ac-bd+(ad+bc)\mathrm{i}$ 翻译而成, 所以 $\mathbb{R}^2$ 中乘法的定义很不自然. 下面我们构造出另一个与 $\mathbb{C}$ 同构的域, 它的“加法”和“乘法”的定义更加自然!

在下面的讨论中, 我们总是将 $\mathbb{R}^2$ 看出带有 $\mathbb{R}$-向量空间结构: $(a,b)+(c,d)=(a+c,b+d)$, $\lambda(a,b)=(\lambda a,\lambda b)\,(\forall\lambda\in\mathbb{R})$ 的 $\mathbb{R}$-向量空间.

定义 1.4.1 映射 $\mathscr{A}:\mathbb{R}^2\to\mathbb{R}^2$ 称为一个线性算子, 如果它满足条件: $\forall\alpha,\beta\in\mathbb{R}^2,\lambda\in\mathbb{R}$, 有

$$\mathscr{A}(\alpha+\beta)=\mathscr{A}(\alpha)+\mathscr{A}(\beta),\ \mathscr{A}(\lambda\alpha)=\lambda\mathscr{A}(\alpha).$$

练习: (1) 如果 $\mathscr{A}:\mathbb{R}^2\to\mathbb{R}^2$ 是线性算子, $0=(0,0)\in\mathbb{R}$, 证明 $\mathscr{A}(0)$, 如果 $\mathscr{A},\mathscr{B}$ 是线性算子, 证明 $\mathscr{A}\cdot\mathscr{B}$ 也是.

(2) 令 $e_1=(1,0),e_2=(0,1)$, 证明: 任一线性算子 $\mathscr{A}:\mathbb{R}^2\to\mathbb{R}^2$ 由 e_1,e_2 的像 $\mathscr{A}(e_1),\mathscr{A}(e_2)$ 唯一确定.

令 $\mathscr{L}\left(\mathbb{R}^2\right)=\left\{\mathscr{A}:\mathbb{R}^2\to\mathbb{R}^2\mid\mathscr{A}\text{是线性算子}\right\}$ 表示所有线性算子的集合, 在 $\mathscr{L}\left(\mathbb{R}^2\right)$ 定义运算: $\forall$ 线性算子 $\mathscr{A}:\mathbb{R}^2\to\mathbb{R}^2,\mathscr{B}:\mathbb{R}^2\to\mathbb{R}^2,\forall\lambda\in\mathbb{R}$,(1) $\mathscr{A}+\mathscr{B}$:

$\mathbb{R}^2 \to \mathbb{R}^2, \alpha \mapsto \mathscr{A}(\alpha) + \mathscr{B}(\alpha)$ (函数值相加), (2) $\mathscr{A} \cdot \mathscr{B} : \mathbb{R}^2 \to \mathbb{R}^2, \alpha \mapsto \mathscr{A}(\mathscr{B}(\alpha))$ (i.e. 映射合成), (3) $\lambda\mathscr{A} : \mathbb{R}^2 \to \mathbb{R}^2, \alpha \mapsto \lambda\mathscr{A}(\alpha)$.

定义 1.4.2 $\mathscr{L}\left(\mathbb{R}^2\right)$ 和上述定义的运算 ①$\mathscr{A} + \mathscr{B}$,②$\mathscr{A} \cdot \mathscr{B}$,③$\lambda\mathscr{A}$ 一起称为 $\mathbb{R}^2$ 上的线性算子代数.

由于 $\mathscr{A} \in \mathscr{L}\left(\mathbb{R}^2\right)$ 保持 $\mathbb{R}^2$ 中运算, 而 $\mathbb{R}^2$ 中每个元素 $x = (x_1, x_2)$ 都可写成 $x = x_1e_1 + x_2e_2$, 其中 $e_1 = (1, 0), e_2 = (0, 1) \in \mathbb{R}^2$, 所以任意线性算子 $\mathscr{A}$ 都由 $\mathscr{A}(e_1), \mathscr{A}(e_2)$ 唯一确定. 令 $\mathscr{A}(e_1) = (a_{11}, a_{21}), \mathscr{A}(e_2) = (a_{12}, a_{22})$, 则我们可以说任一线性算子 $\mathscr{A}$ 都由数组 $A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$ 唯一确定.

定义 1.4.3 形如 $\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$ (其中 $a_{ij} \in \mathbb{R}$) 的数组称为一个系数在 $\mathbb{R}$ 中的 2×2 矩阵 (或简称二阶实矩阵), 两个矩阵 $A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}, B = \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix}$ 称为相等, 如果 $a_{ij} = b_{ij}$. 其中 $\begin{pmatrix} a_{11} \\ a_{21} \end{pmatrix}, \begin{pmatrix} a_{12} \\ a_{22} \end{pmatrix}$ 分别为 A 的第一列, 第二列. 而 $(a_{11}, a_{12}), (a_{21}, a_{22})$ 则分别称为 A 的第一行, 第一行 (也分别称它们是列向量, 行向量).

有可上述定义, 我们可以简单地说: 对任一 $\mathscr{A} \in \mathscr{L}\left(\mathbb{R}^2\right)$, 如果 $\mathscr{A}(e_1) = a_{11}e_1 + a_{21}e_2, \mathscr{A}(e_2) = a_{12}e_1 + a_{22}e_2$, 则 $\mathscr{A}$ 由矩阵$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$ 唯一确定, 称 A 是$\mathscr{A}$ 对应的矩阵. 为了表示 $\mathscr{A}$ 与矩阵 A 的对应关系, 我们可以将 $\mathscr{A} \in \mathscr{L}\left(\mathbb{R}^2\right)$. 表示为

$$(\mathscr{A}(e_1), \mathscr{A}(e_2)) = (e_1, e_2)\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} = (e_1, e_2)A.$$

这样我们得到一个映射

$$f : \mathscr{L}\left(\mathbb{R}^2\right) \longrightarrow M_2(\mathbb{R}) = \left\{\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \mid a_{ij} \in \mathbb{R}\right\} \text{(所有二阶实系数的集合).}$$

引理 1.4.1 上述映射 $f:\mathscr{L}\left(\mathbb{R}^2\right)\longrightarrow M_2(\mathbb{R})$ (i.e.$\forall\mathscr{A}\in\mathscr{L}\left(\mathbb{R}^2\right), f(\mathscr{A})=\begin{pmatrix}a_{11}&a_{12}\\a_{21}&a_{22}\end{pmatrix}$ 使 $\mathscr{A}(e_1)=a_{11}e_1+a_{21}e_2$, $\mathscr{A}(e_2)=a_{12}e_1+a_{22}e_2$) 是一个双射.

证明: $\forall\mathscr{A},\mathscr{B}\in\mathscr{L}\left(\mathbb{R}^2\right)$. 如果 $f(\mathscr{A})=f(\mathscr{B})$, 则 $\mathscr{A}(e_1)=\mathscr{B}(e_1)$, $\mathscr{A}(e_2)=\mathscr{B}(e_2)$. 所以 $\mathscr{A}=\mathscr{B}$, f 是单射. 为了证明 f 是满射, $\forall A=\begin{pmatrix}a_{11}&a_{12}\\a_{21}&a_{22}\end{pmatrix}\in M_2(\mathbb{R})$, 定义 $\mathscr{A}:\mathbb{R}^2\to\mathbb{R}^2$ 如下: $\forall x=(x_1,x_2)=x_1e_1+x_2e_2$, 令

$$\mathscr{A}(x)=(a_{11}x_1+a_{12}x_2)\,e_1+(a_{21}x_1+a_{22}x_2)$$

则 $\mathscr{A}$ 是一个线性算子使得 $f(\mathscr{A})=A$. □

注意: 在 $\mathscr{L}\left(\mathbb{R}^2\right)$ 中有下述特殊算子: ① 零算子 $\mathscr{O}:\mathbb{R}^2\to\mathbb{R}^2$ 它将所有 $\alpha\in\mathbb{R}^2$ 映到 $0=(0,0)$, 所以 $f(0)=\begin{pmatrix}0&0\\0&0\end{pmatrix}:=0$ (零矩阵), ② 恒等算子 $1_{\mathbb{R}^2}:\mathbb{R}^2\to\mathbb{R}^2$ 它将每个 $\alpha\in\mathbb{R}^2$ 映到 α 自己, 所以 $f(1_{\mathbb{R}^2})=\begin{pmatrix}1&0\\0&1\end{pmatrix}:=I$ (单位矩阵). ③$\forall\lambda\in\mathbb{R}$, 有数乘算子 $m_\lambda:\mathbb{R}^2\to\mathbb{R}^2$, 它将 $\alpha\in\mathbb{R}^2$ 映到 $\lambda\alpha$, 所以 $f(m_\lambda)=\begin{pmatrix}\lambda&0\\0&\lambda\end{pmatrix}$(数乘矩阵).

通过引理1.4.1中的双射, $\mathscr{L}\left(\mathbb{R}^2\right)$ 中的三个运算: 加法, 乘法和数乘可以诱导 $M_2(\mathbb{R})$ 中的三个运算: $\forall\lambda\in\mathbb{R}, A, B\in M_2(\mathbb{R})$, 令 $\mathscr{A},\mathscr{B}\in\mathscr{L}\left(\mathbb{R}^2\right)$ 使 $f(\mathscr{A})=A, f(\mathscr{B})=B$, 定义: ① $A+B=f(\mathscr{A}+\mathscr{B})$, ② $\lambda A=f(\lambda\mathscr{A})$, ③ $A\cdot B=f(\mathscr{A}\cdot\mathscr{B})$. 则我们得到关于矩阵的加法, 矩阵的数乘法, 矩阵的乘法.

引理 1.4.2 $\forall\lambda\in\mathbb{R}, A=\begin{pmatrix}a_{11}&a_{12}\\a_{21}&a_{22}\end{pmatrix}, B=\begin{pmatrix}b_{11}&b_{12}\\b_{21}&b_{22}\end{pmatrix}\in M_2(\mathbb{R})$

① $A+B=\begin{pmatrix}a_{11}+b_{11}&a_{12}+b_{12}\\a_{21}+b_{21}&a_{22}+b_{22}\end{pmatrix}$, ② $\lambda A=\begin{pmatrix}\lambda a_{11}&\lambda a_{12}\\\lambda a_{21}&\lambda a_{22}\end{pmatrix}$,

③ $A\cdot B=\begin{pmatrix}a_{11}b_{11}+a_{12}b_{21}&a_{11}b_{12}+a_{12}b_{22}\\a_{21}b_{11}+a_{22}b_{21}&a_{21}b_{12}+a_{22}b_{22}\end{pmatrix}$.

证明: 令 $\mathscr{A}, \mathscr{B} \in \mathscr{L}\left(\mathbb{R}^2\right), f(\mathscr{A}) = A, f(\mathscr{B}) = B$, 则

$$\mathscr{A}(e_1) = (a_{11}, a_{21}), \mathscr{A}(e_2) = (a_{12}, a_{22}), \mathscr{B}(e_1) = (b_{11}, b_{21}), \mathscr{B}(e_2) = (b_{12}, b_{22}).$$

从而

$$(\mathscr{A} + \mathscr{B})(e_1) = \mathscr{A}(e_1) + \mathscr{B}(e_1) = (a_{11} + b_{11}, a_{21} + b_{21}),\ \lambda\mathscr{A}(e_1) = (\lambda a_{11}, \lambda a_{21})$$

$$(\mathscr{A} + \mathscr{B})(e_2) = \mathscr{A}(e_2) + \mathscr{B}(e_2) = (a_{12} + b_{12}, a_{22} + b_{22}),\ \lambda\mathscr{A}(e_2) = (\lambda a_{12}, \lambda a_{22})$$

所以

$$f(\mathscr{A} + \mathscr{B}) = \begin{pmatrix} a_{11} + b_{11} & a_{12} + b_{12} \\ a_{21} + b_{21} & a_{22} + b_{22} \end{pmatrix} = A + B,\ f(\lambda\mathscr{A}) = \begin{pmatrix} \lambda a_{11} & \lambda a_{12} \\ \lambda a_{21} & \lambda a_{22} \end{pmatrix} = \lambda A.$$

下面确定 $f(\mathscr{A} \cdot \mathscr{B})$: 由 $\mathscr{A}(e_1) = a_{11}e_1 + a_{21}e_2, \mathscr{A}(e_2) = a_{12}e_1 + a_{22}e_2$, 及 $\mathscr{B}(e_1) = b_{11}e_1 + b_{21}e_2, \mathscr{B}(e_2) = b_{12}e_1 + b_{22}e_2$. 得

$$\begin{aligned}\mathscr{A} \cdot \mathscr{B}(e_1) &= \mathscr{A}(\mathscr{B}(e_1)) = \mathscr{A}(b_{11}e_1 + b_{21}e_2) = b_{11}\mathscr{A}(e_1) + b_{21}\mathscr{A}(e_2) \\ &= (a_{11}b_{11} + a_{12}b_{21})\, e_1 + (a_{21}b_{11} + a_{22}b_{21})\, e_2\end{aligned}$$

$$\begin{aligned}\mathscr{A} \cdot \mathscr{B}(e_2) &= \mathscr{A}(\mathscr{B}(e_2)) = \mathscr{A}(b_{12}e_1 + b_{22}e_2) = b_{12}\mathscr{A}(e_1) + b_{22}\mathscr{A}(e_2) \\ &= (a_{11}b_{12} + a_{12}b_{22})\, e_1 + (a_{21}b_{12} + a_{22}b_{22})\, e_2\end{aligned}$$

□

引理 1.4.3 $\mathscr{L}\left(\mathbb{R}^2\right), M_2(\mathbb{R})$ 关于各自运算是一个环. 且双射 $f: \mathscr{L}\left(\mathbb{R}^2\right) \to M_2(\mathbb{R})$ 是一个同构, 即

$$f(\mathscr{A} + \mathscr{B}) = f(\mathscr{A}) + f(\mathscr{B}),\ f(\mathscr{A} \cdot \mathscr{B}) = f(\mathscr{A}) \cdot f(\mathscr{B})$$

且 $f(\mathscr{O}) = 0,\ f(1_{\mathbb{R}^2}) = I$.

证明: $\mathscr{L}\left(\mathbb{R}^2\right), M_2(\mathbb{R})$ 是环科直接验证. f 保持运算时显然的, 因为 $M_2(\mathbb{R})$ 上的运算是由 $\mathscr{L}\left(\mathbb{R}^2\right)$ 上的运算通过 f 定义的. □

推论 1.4.1 $\forall \lambda \in \mathbb{R}, A, B, C \in M_2(\mathbb{R})$. 我们有

(1) $(A + B) + C = A + (B + C)$, (2)$A + 0 = 0 + A = 0$

(3) $-A = \begin{pmatrix} -a_{11} & -a_{12} \\ -a_{21} & -a_{22} \end{pmatrix}$, (4) $(A \cdot B)\, C = A \cdot (B \cdot C)$

(5) $A \cdot (B + C) = A \cdot B + A \cdot C$, (6) $(B + C) \cdot A = B \cdot A + C \cdot A$

(7) $A \cdot I = I \cdot A = A$, 其中 $I = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$.

推论 1.4.2 $\forall A \in M_2(\mathbb{R})$. 令 $\mathscr{A}: \mathbb{R}^2 \to \mathbb{R}^2$ 是由 A 定义的线性算子. 则 $\mathscr{A}$ 是双射的充要条件是存在 $\mathscr{B} \in M_2\left(\mathbb{R}^2\right)$ 使 $A \cdot B = B \cdot A = I$.

证明: $\mathscr{A}: \mathbb{R}^2 \to \mathbb{R}^2$ 是双射 $\Leftrightarrow \exists \mathscr{B}: \mathbb{R}^2 \to \mathbb{R}^2$ 使 $\mathscr{A} \cdot \mathscr{B} = \mathscr{B} \cdot \mathscr{A}$. 且 $\mathscr{A}$ 是线性算子 $\Leftrightarrow \mathscr{B}$ 是线性算子. 令 $B = f(\mathscr{B})$. 所以 $\mathscr{A} \cdot \mathscr{B} = \mathscr{B} \cdot \mathscr{A} \Leftrightarrow A \cdot B = B \cdot A$. □

定义 1.4.4 $A \in M_2(\mathbb{R})$ 称为可逆矩阵, 如果存在 $B \in M_2(\mathbb{R})$ 使 $A \cdot B = B \cdot A = I$. B 称为 A 可逆矩阵.

在环 $M_2(\mathbb{R})$ 中的加法和乘法是非常自然的, 因为它们实质上是 $\mathbb{R}^2$ 到 $\mathbb{R}^2$ 函数的加法和复合, 而且不难发现.

推论 1.4.3 数量矩阵的集合 $\mathbb{R} \cdot I = \{\lambda \cdot I \mid \forall \lambda \in \mathbb{R}\}$ 关于 $M_2(\mathbb{R})$ 的加法和乘法是一个与实数域 $\mathbb{R}$ 同构的域.

证明: 子集合 $\mathbb{R} \cdot I = \left\{\begin{pmatrix} \lambda & 0 \\ 0 & \lambda \end{pmatrix} \mid \forall \lambda \in \mathbb{R}\right\}$ 矩阵 $\lambda I = \begin{pmatrix} \lambda & 0 \\ 0 & \lambda \end{pmatrix}$ 可逆. 所以 $\mathbb{R} \cdot I$ 是一个域. 且 $\mathbb{R} \to \mathbb{R} \cdot I$, $\lambda \mapsto \lambda I \begin{pmatrix} \lambda & 0 \\ 0 & \lambda \end{pmatrix}$ 是域同构. □

引理 1.4.4

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \in M_2(\mathbb{R}) \text{ 可逆} \Leftrightarrow a_{11}a_{22} - a_{12}a_{21} \neq 0 \tag{1-2}$$

证明: 如果 A 可逆, 则存在 $B = \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix} \in M_2(\mathbb{R})$ 使

$$A \cdot B = B \cdot A = I \tag{1-3}$$

即我们有下述方程组

$$\begin{cases} a_{11}b_{11} + a_{12}b_{21} = 1 \\ a_{21}b_{12} + a_{22}b_{22} = 1 \\ a_{11}b_{12} + a_{12}b_{22} = 0 \\ a_{21}b_{11} + a_{22}b_{21} = 0 \end{cases} \tag{1-4}$$

$$\Leftarrow \text{由 } A \cdot B = \begin{pmatrix} a_{11}b_{11}+a_{12}b_{21} & a_{11}b_{12}+a_{12}b_{22} \\ a_{21}b_{11}+a_{22}b_{21} & a_{21}b_{12}+a_{22}b_{22} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

由方程 (1-2) ,a_{11}, a_{12} 不同时为零. 如果 $a_{11} \neq 0$, 由方程组 (1-3) ,方程组 (1-4) 得

$$a_{11} = a_{11}(a_{21}b_{12}+a_{22}b_{22}) = a_{21}a_{11}b_{12}+a_{11}a_{22}b_{22} = (a_{11}a_{22}-a_{12}a_{21})\,b_{22}$$

如果 $a_{12} \neq 0$, 同理得

$$a_{12} = a_{12}(a_{21}b_{12}+a_{22}b_{22}) = a_{21}a_{12}b_{12}+a_{12}a_{22}b_{22} = -(a_{11}a_{22}-a_{12}a_{21})\,b_{12}$$

所以 $a_{11}a_{22}-a_{12}a_{21} \neq 0$. 如果 $a_{11}a_{22}-a_{12}a_{21} \neq 0$, 令 $|A| = a_{11}a_{22}-a_{12}a_{21}$, 则$B = \begin{pmatrix} \frac{a_{22}}{|A|} & -\frac{a_{12}}{|A|} \\ -\frac{a_{21}}{|A|} & \frac{a_{11}}{|A|} \end{pmatrix}$是 A 的逆矩阵. □

定义 1.4.5 $A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$, $|A| = a_{11}a_{22}-a_{12}a_{21}$ 称为 A 的行列式 (因此, A可逆 $\Leftrightarrow$ A的行列式 $|A| \neq 0$).

定理 1.4.2 子集合 $\mathbb{F} = \left\{ \begin{pmatrix} a & b \\ -b & a \end{pmatrix} \mid a, b \in \mathbb{R} \right\} \subset M_2(\mathbb{R})$ 关于 $M_2(\mathbb{R})$ 的加法和乘法是一个同构于 $\mathbb{C}$ 的域.

证明: (1) $\mathbb{F}$ 对加法封闭, 且 $\mathbb{F}$ 中元素的负元在 $\mathbb{F}$ 中.

$$\begin{pmatrix} a & b \\ -b & a \end{pmatrix} + \begin{pmatrix} c & d \\ -d & c \end{pmatrix} = \begin{pmatrix} a+c & b+d \\ -(b+d) & a+c \end{pmatrix} \in \mathbb{F}$$

$$-\begin{pmatrix} a & b \\ -b & a \end{pmatrix} = \begin{pmatrix} -a & -b \\ -(-b) & -a \end{pmatrix} \in \mathbb{F}$$

(2) $\mathbb{F}$ 对乘法封闭, 且非零矩阵可逆.

$$\begin{pmatrix} a & b \\ -b & a \end{pmatrix} \cdot \begin{pmatrix} c & d \\ -d & c \end{pmatrix} = \begin{pmatrix} ac-bd & ad+bc \\ -(ad+bc) & ac-bd \end{pmatrix}$$

且

$$A = \begin{pmatrix} a & b \\ -b & a \end{pmatrix} \neq 0 \text{ (零矩阵)} \Leftrightarrow a^2+b^2 \neq 0 \Leftrightarrow A \text{ 可逆 (引理1.4.4)}$$

(3) $f:\mathbb{C}\to\mathbb{F},\ a+b\mathrm{i}\mapsto\begin{pmatrix}a & b\\ -b & a\end{pmatrix}$是域同构.

$$(a+b\mathrm{i})\cdot(c+d\mathrm{i})=(ac-bd)+(ad+bc)$$

所以

$$f\left((a+b\mathrm{i})\cdot(c+d\mathrm{i})\right)=f\left((a+b\mathrm{i})\right)f\left((c+d\mathrm{i})\right).$$

保持加法是显然的. 且 $f(0)=0$ (零矩阵) $f(1)=I=\begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix}$. 且

$$f(\mathbb{R})=\left\{\begin{pmatrix}\lambda & 0\\ 0 & \lambda\end{pmatrix}\mid\forall\lambda\in\mathbb{R}\right\}.\ f(i)=\begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix}=J.$$

可验证: $J^2+I=0$ □

前面我们对复数域的运算给出了几何描述和代数描述. 但关于复数域 $\mathbb{C}$ 的最重要定理 (至少在代数方面) 也许是所谓的代数基本定理.

定理 1.4.3 (代数基本定理). 设 $f(x)\in\mathbb{C}[x]$. 如果 $degf(x)>0$, 则存在 $\alpha\in\mathbb{C}$ 使 $f(\alpha)=0$. 特别, 任何首项系数为 1 的非常值多项式 $f(x)\in\mathbb{C}[x]$. 都可分解成

$$f(x)=(x-\alpha_1)(x-\alpha_2)\cdots(x-\alpha_n),\ \text{其中}\ n=degf(x).$$

$\alpha_1,\cdots,\alpha_n$ 可以相同.

这里不给证明 (尽管它有近百种证明) , 因为易懂的纯代数证明还没有找到. 但我们可以用它讨论 $\mathbb{C}$ 的子域.

定义 1.4.6 设 $K\subset\mathbb{C}$ 是一个包含非零元的子集合. 如果 K 关于 $\mathbb{C}$ 中的加, 减, 乘, 除封闭, 则称 K 是 $\mathbb{C}$ 的子域. (即: $\forall x,y\in K$, 有 $x+y, x-y, xy, \frac{x}{y}$(如果$y\neq 0$) $\in K$)

显然, $\mathbb{C}$ 中的子域 K 是一个域. 下面是 $\mathbb{C}$ 中子域的一些例子.

例 1.4.1 显然, $\mathbb{Q},\mathbb{R}$ 是 $\mathbb{C}$ 的子域. 事实上, $\mathbb{Q}\subset\mathbb{C}$ 是 $\mathbb{C}$ 的最小子域. i.e. 如果 $K\subset\mathbb{C}$ 是一个子域, 则$\mathbb{Q}\subset K$.

设 $z\in\mathbb{C}$, 我们可以包含 z 和 $\mathbb{Q}$ 的最小子域是什么?

定义 1.4.7 $z \in \mathbb{C}$ 称为 $\mathbb{Q}$ 上的代数元 (或代数数). 如果存在首项系数为1的多项式 $f(x) \in \mathbb{Q}[x]$ 使 $f(z) = 0$. 否则称 z 是 $\mathbb{Q}$ 上的超越元 (或超越数).

定理 1.4.4 设 $z \in \mathbb{C}$, $\mathbb{Q}(x) \subset \mathbb{C}$ 表示包含 $\mathbb{Q}$ 和 z 的最小子域. 则我们有

(1) 当 $z \in \mathbb{C}$ 是 $\mathbb{Q}$ 上的超越元时,

$$\mathbb{Q}(z) = \left\{ \frac{f(z)}{g(z)} \mid \forall f(x), g(x) \in \mathbb{Q}[x], g(x) \neq 0 \right\}.$$

(2) 当 $z \in \mathbb{C}$ 是 $\mathbb{Q}$ 上的代数元时,

$$\mathbb{Q}(z) = \{ f(z) \mid \forall f(x) \in \mathbb{Q}[x] \}.$$

此时我们用 $\mathbb{Q}[z]$ 表示 $\mathbb{Q}(z)$ (因为, 此时 $\mathbb{Q}(x)$ 中的元素都可以表成 z 的多项式).

证明: 如果 $K \subset \mathbb{C}$ 是一个包含 $\mathbb{Q}$ 和 z 的子域, 则 $f(z) \in K$. 对任意多项式 $f(x) = a_n x^n + a_{n-1} x^{n-1} + \cdots + a_1 x + a_0$(其中$a_i \in \mathbb{Q}$) 成立. 所以, $\forall f(x), g(x) \in \mathbb{Q}[x]$ 如果 $g(z) \neq 0$, 则 $\frac{f(z)}{g(z)} \in K$. i.e.

$$\mathbb{F} = \left\{ \frac{f(z)}{g(z)} \mid \forall f(x), g(x) \in \mathbb{Q}[x], g(x) \neq 0 \right\} \subset K.$$

可以验证集合 $\mathbb{F} \subset \mathbb{C}$ 也是 $\mathbb{C}$ 的子域. 且包含 $\mathbb{Q}$ 和 z. 所以, $\mathbb{F}$ 就是 $\mathbb{C}$ 中包含 $\mathbb{Q}$ 和 z 的最小子域 $\mathbb{Q}(z)$.

(1) 当 z 是 $\mathbb{Q}$ 上超越元时, $\forall g(x) \in \mathbb{Q}[x]$ 非零, $g(z) \neq 0$ 所以,

$$\mathbb{Q}(z) = \mathbb{F} = \left\{ \frac{f(z)}{g(z)} \mid \forall f(x), g(x) \in \mathbb{Q}[x], g(x) \neq 0 \right\}.$$

(2) 当 z 是 $\mathbb{Q}$ 上代数元时, 令

$$S = \{ f(x) \in \mathbb{Q}[x] \mid f(z) = 0 \}$$

在 S 中存在非零多项式. 令 $m(x) \in S$ 表示 S 中次数最小的首项系数为 1 的多项式, 则 $m(x)$ 满足:

$$① \ m(z) = 0, \ ② \ \forall f(x) \in \mathbb{Q}[x], \text{如果} f(z) = 0, \text{则}\ m(x) \mid f(x) .$$

(如果 $f(z) = 0$, 令 $f(x) = g(x)m(x) + r(x)$, 则 $r(z) = 0$, $r(x)$ 必为零, 否则 $deg r(x) < deg m(x)$, 与 $m(x)$ 的选取矛盾)

显然, $m(x)$ 必为不可约多项式(否则, $m(x) = m_1(x) m_2(x)$, $deg m_1(x) < deg m(x)$, 则$0 = m(z) = m_1(z) \cdot m_2(z) \Rightarrow m_1(z) = 0$, 或 $m_2(z) = 0$. 与 $m(x)$ 选取矛盾).

$$\forall \frac{f(z)}{g(z)} \in \mathbb{F}, \ \text{i.e.}\ g(z) \neq 0. \text{则}\ m(x) \text{与} g(x) \text{互素}.$$

$$\Rightarrow \exists u(x), v(x) \in \mathbb{Q}[x],\ 使\ m(x)\cdot u(x) + g(x)\cdot v(x) = 1.$$

$$\Rightarrow 1 = m(z)\cdot u(z) + g(z)\cdot v(z) = g(z)\cdot v(z).$$

所以

$$\frac{f(z)}{g(z)} = \frac{f(z)\cdot v(z)}{g(z)\cdot v(z)} = f(z)\cdot v(z)$$

即 $\mathbb{F}$ 中每个元素都可以写成关于 z 的多项式. 所以

$$\mathbb{Q}(z) = \mathbb{Q}[z] = \{f(z) \mid \forall f(x) \in \mathbb{Q}[x]\}.$$

$\square$

例 1.4.2 设 z 是二次多项式 $x^2 + bx + c = 0\,(b, c \in \mathbb{Q})$ 的根. $\triangle = b^2 - 4c$, 则

$$\mathbb{Q}[z] = \left\{\lambda_0 + \lambda_1\sqrt{\triangle} \mid \lambda_0, \lambda_1 \in \mathbb{Q}\right\} = \mathbb{Q}[\sqrt{\triangle}].$$

证明: $z = -\frac{b}{2} \pm \frac{\sqrt{\triangle}}{2}$. 如果 $\sqrt{\triangle} \notin \mathbb{Q}$, 则 $p(x) = x^2 + bx + c$ 不可约. 且 $p(z) = 0$.

$$\forall f(x) \in \mathbb{Q}[x],\ f(x) = q(x)p(x) + a_1x + a_0 \Rightarrow$$

$$f(z) = a_1z + a_0 = \lambda_1\sqrt{\triangle} + \lambda_0,\ (\lambda_0, \lambda_1 \in \mathbb{Q}) \Rightarrow$$

$$\mathbb{Q}[z] = \mathbb{Q}[\triangle] = \left\{\lambda_0 + \lambda_1\sqrt{\triangle} \mid \lambda_0, \lambda_1 \in \mathbb{Q}\right\}.$$

$\square$

定理 1.4.4 中包含 $\mathbb{Q}$ 与 z 的最小域 $\mathbb{Q}(z)$ 称为由 $\mathbb{Q}$ 添加 z 生成的域. 我们可以对 $\mathbb{Q}$ 添加任意次多项式的根 z 而得到各种子域 $\mathbb{Q}[z] \subset \mathbb{C}$, 它们称为 $\mathbb{Q}$ 的代数扩张. 下面介绍一个由初等几何作用问题产生的子域.

在平面 Ω 上给定有限个点 $S = \{P_1, \cdots, P_m\}$, 历史上的尺规作图问题关心一个特定的点 P 能否由 $\{P_1, \cdots, P_m\}$ 仅用直尺, 圆规作出. 例如, 能否任意角度都可以仅用直尺-圆规三等分 ? 能否用直尺-圆规构造一个正七边形? 设

$$C(P_1, P_2, \cdots, P_m) \subset \Omega$$

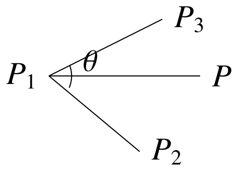
是由 $\{P_1, P_2, \cdots, P_m\}$ 经直尺-圆规构造的所有点集 (称为可构造集) . $C(P_1, P_2, \cdots, P_m)$ 可以更明确地刻画如下:

(1) $S = \{P_1, P_2, \cdots, P_m\} \subset C(P_1, P_2, \cdots, P_m)$.

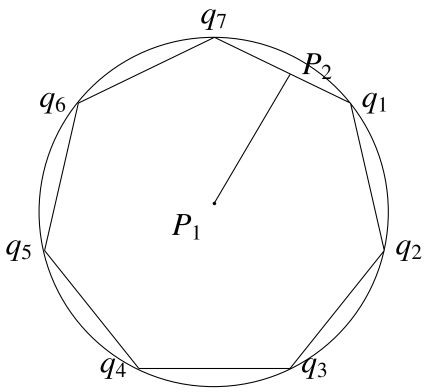
(2) $\forall A, B \in C(P_1, P_2, \cdots, P_m)$. 连接 A, B 的直线 $\overline{AB}$ 称为可构造直线, 以 A 为圆心, AB 为半径的圆周称为可构造圆周. 则它们交点 (i.e. 直线之间的交点, 直线与圆周的交点, 圆周之间的交点) 也在 $C(P_1, P_2, \cdots, P_m)$ 中.

(3) $C(P_1, P_2, \cdots, P_m)$ 是满足 (1), (2) 的最小点集.
平面 Ω 上的一个点 P 称为可由直尺-圆规构造, 如果 $P \in C(P_1, P_2, \cdots, P_m)$.

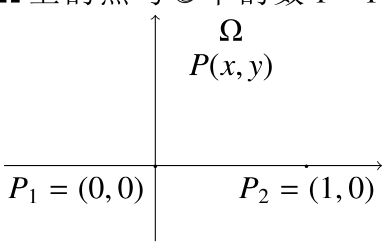
例 1.4.3 设角 θ 由三个点 (不在同一直线上) P_1, P_2, P_3 确定. 角 θ 是否可由直尺-圆规三等分? 等价于是否存在 $P \in C(P_1, P_2, \cdots, P_m)$ 使 $\angle PP_1P_2 = \frac{1}{3}\theta$?



例 1.4.4 正七变形是否可由直尺-圆规作出? 等价于给定 P_1, P_2 是否可在以 P_1 为圆心, P_1P_2 为半径的圆周上作出 7 等分点? i.e. 是否 $q_1, q_2, \cdot, q_7 \in C(P_1, P_2, \cdots, P_m)$?



如果在平面 Ω 上选取以 P_1 为原点的直角坐标系. 使 $P_2 = (1, 0)$. 则可将 Ω 上的点与 $\mathbb{C}$ 中的数 $1-1$ 对应.



$\longleftrightarrow$ $\quad z = x + y\mathrm{i} \in \mathbb{C}$

$P \leftrightarrow z = x + y\mathrm{i}$

令 $z_1, z_2, \cdots, z_m$ 是点 $P_1, P_2, \cdots, P_m$ 对应的复数 ($z_1 = 0, z_2 = 1$) 集合 $C(z_1, z_2, \cdots, z_m) \subset \mathbb{C}$ 是 $C(P_1, P_2, \cdots, P_m)$ 中点对应复数的集合. 则

定理 1.4.5 $C(z_1, z_2, \cdots, z_m) \subset \mathbb{C}$ 是 $\subset \mathbb{C}$ 的子域且满足

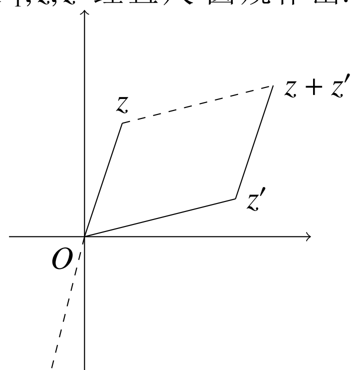
(I) $z \in (z_1, z_2, \cdots, z_n) \Rightarrow$ 共轭 $\bar{z} \in C(z_1, z_2, \cdots, z_n)$.

(II) $z \in (z_1, z_2, \cdots, z_n) \Rightarrow$ 它的平方根 $\sqrt{z} \in C(z_1, z_2, \cdots, z_n)$.

证明: 首先证明 $C(z_1, z_2, \cdots, z_m) \subset \mathbb{C}$ 是 $\subset \mathbb{C}$ 的子域. 只需证明:

$$\forall z, z' \in C(z_1, z_2, \cdots, z_m) \Rightarrow z + z', -z, z \cdot z', \frac{z}{z'} \in C(z_1, \cdots z_m).$$

它的证明是一系列的初等几何作图问题, 例如 (如图1) , $z + z'$ 是平行四边形的顶点, $-z$ 是以 P_1 为圆心 $|z|$ 为半径圆周与直线 $\overline{P_1 z}$ 的交点. 它们都可由 P_1, z, z' 经直尺-圆规作出.

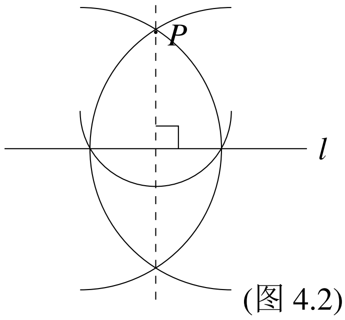


(图 4.1)

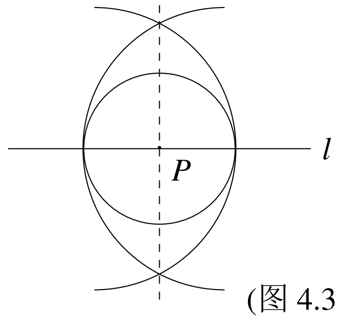
为证明 $z \cdot z', \frac{z}{z'}(z \neq 0)$ 也在 $C(z_1, z_2, \cdots, z_m)$ 中, 首先注意到一个事实

$$z = x + y\mathrm{i} \in C(z_1, z_2, \cdots; z_m) \Leftrightarrow x, y \in C(z_1, z_2, \cdots z_m)$$

(i.e. z 可构造 $\Leftrightarrow$ 它的实部与虚部可构造). 该事实有下述作图给出.



(图 4.2)



(图 4.3)

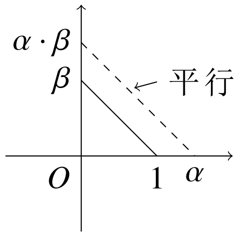
图 4.2: 给定直线 l 及 l 外一点 P , 可作一条过 P 点垂直 l 的直线.

图 4.3: 给定直线 l 及 l 上一点 P , 可作一条过 P 点垂直 l 的直线.

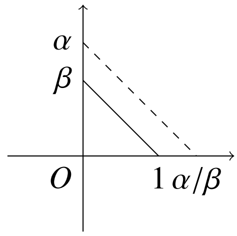
上述作图也推出: 给定直线 l 及 l 外一点 P , 可构造过 P 点且平行于 l 的直线 (所以平行四边形的顶点可构造) . 设 $z = x + y\mathrm{i}$, $z' = x' + y'\mathrm{i}$ 则

$$zz' = (xx' - yy') + (xy' + x'y)\,\mathrm{i},\ \frac{z}{z'} = \frac{xx' + yy'}{x'^2 + y'^2} + \frac{x'y - xy'}{x'^2 + y'^2}\mathrm{i}$$

所以, 只需证明: 如果 $\alpha, \beta \in C(z_1, z_2, \cdots, z_m)$ 是实数, 则 $\alpha \cdot \beta, \frac{\alpha}{\beta} \in C(z_1, z_2, \cdots, z_m)$, 它的证明由下图推出:



(图 4.4)

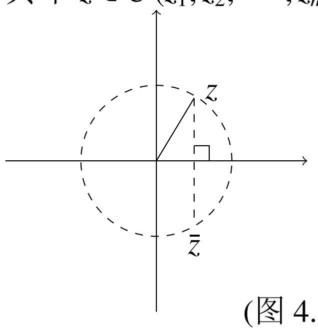


(图 4.5)

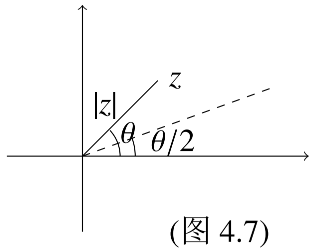
所以, 我们证明 $C(z_1, z_2, \cdots, z_m) \subset \mathbb{C}$ 是一个子域. 下面证明:

$$\forall z \in C(z_1, z_2, \cdots, z_m) \Rightarrow \bar{z},\ \sqrt{z} \in C(z_1, z_2, \cdots, z_m).$$

其中 $\bar{z} \in C(z_1, z_2, \cdots, z_m)$. 由图 4.6 推出.

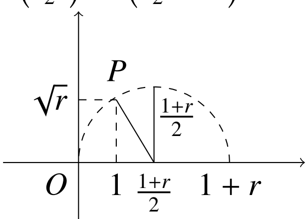


(图 4.6)



(图 4.7)

令 $z = |z|(\cos\theta + \mathrm{i}\sin\theta)$, 则 $\sqrt{z} = \sqrt{|z|}(\cos\theta + \mathrm{i}\sin\theta)$, 而角的平分线可由直尺- 圆规作出 (图4.7) . 所以只需证明: $C(z_1, z_2, \cdots, z_m)$ 中的实数 $|z|$ 的平方根 $\sqrt{|z|}$ 可由直尺-圆规作出; $r = |z|, r$ 可构造 $\Rightarrow 1 + r, \frac{1+r}{2}$ 可构造 . 即 $\overline{P1}^2 = \left(\frac{1+r}{2}\right)^2 - \left(\frac{1+r}{2} - 1\right)^2 = r$, 所以 $\overline{P1} = \sqrt{r}$ 可构造.



(图 4.8)

□

习题1.4

1. 计算下列表达式.

(a) $\frac{(1+3\mathrm{i})}{8-\mathrm{i}}$, (b) $(-\frac{1}{2} \pm \frac{\sqrt{3}}{2}\mathrm{i})^3$, (c) $(\frac{\sqrt{3}+\mathrm{i}}{1-\mathrm{i}})^{30}$.

2. 解下述方程.

(a) $|z| + z = 8 + 4\mathrm{i}$,

(b) 解方程组 $\begin{cases}(1+\mathrm{i})z_1+(1-\mathrm{i})z_2=1+\mathrm{i},\\ (1-\mathrm{i})z_1+(1+\mathrm{i})z_2=1+3\mathrm{i}.\end{cases}$

(c) $z^2=3-4\mathrm{i}$,

(d) 求 $(2+\mathrm{i})x+(1+2\mathrm{i})y=1-4\mathrm{i}$ 的实数解.

3. 求下述方程的全部解.

(a) $x^6-\mathrm{i}=0$, (b) $x^6-64=0$, (c) $x^{10}-512(1-\mathrm{i}\sqrt{3})=0$.

4. 设 $\varepsilon_k=\cos\frac{2\pi k}{n}+\mathrm{i}\sin\frac{2\pi k}{n}$ $(0\le k<n)$ 是 n 次单位根. 证明:

(a) $\varepsilon_k=\varepsilon_1^k$ $(0\le k<n)$, (b) $\varepsilon_k\varepsilon_l=\begin{cases}\varepsilon_{k+l}, & 如果k+l<n,\\ \varepsilon_{k+l-n}, & 如果k+l\geqslant n.\end{cases}$

(c) 设 $a\in\mathbb{R}$ 是大于零的实数. 则 $x^n-a=0$ 的全部根为

$$\sqrt{a},\ \sqrt{a}\varepsilon_1,\ \sqrt{a}\varepsilon_1^2,\cdots,\ \sqrt{a}\varepsilon_1^{n-1}.$$

5. 设 α 是 $x^3-2=0$ 的一个根, 求 $\mathbb{Q}[\alpha]=$?

6. 设 $\mathscr{A}:\mathbb{R}^2\to\mathbb{R}^2$ 是一个线性算子 ($\mathbb{R}^2$表示标准 $\mathbb{R}$-向量空间) . 证明:

$$\mathscr{A}是单射\Leftrightarrow\forall x=(x_1,x_2)\neq 0,\mathscr{A}(x)\neq 0\ (0=(0,0)\in\mathbb{R}).$$

7. 设 $\mathscr{A}:\mathbb{R}^2\to\mathbb{R}^2$ 是一个线性算子. $e_1=(1,0),e_2=(0,1)$. 证明:

(1) $\mathscr{A}$ 是单射 $\Leftrightarrow$ 对任意不全为零的数 $\lambda_1,\lambda_2\in\mathbb{R},\lambda_1\mathscr{A}(e_1)+\lambda_2\mathscr{A}(e_2)\neq 0$. (此时称向量 $\mathscr{A}(e_1),\mathscr{A}(e_2)$ 线性无关).

(2) 令 $\mathscr{A}(e_1)=a_{11}e_1+a_{21}e_2,\mathscr{A}(e_2)=a_{12}e_1+a_{22}e_2$. 则

$$\mathscr{A}(e_1),\mathscr{A}(e_2)\ 线性无关\Leftrightarrow\begin{vmatrix}a_{11}&a_{12}\\a_{21}&a_{22}\end{vmatrix}=a_{11}a_{22}-a_{12}a_{21}\neq 0.$$

(3) $\forall b=(b_1,b_2)\in\mathbb{R}$, 如果 $\begin{vmatrix}a_{11}&a_{12}\\a_{21}&a_{22}\end{vmatrix}\neq 0$, 则方程组 $\begin{cases}a_{11}x_1+a_{12}x_2=b_1\\a_{21}x_1+a_{22}x_2=b_1\end{cases}$ 有唯一解.

8. 设 $\mathscr{A}:\mathbb{R}^2\to\mathbb{R}^2$ 是一个线性算子, $\mathscr{A}(e_1)=(a_{11},a_{21}),\mathscr{A}(e_2)=(a_{12},a_{22})$ 证明下述等价条件: (a) $\mathscr{A}$ 是单射, (b) $\begin{vmatrix}a_{11}&a_{12}\\a_{21}&a_{22}\end{vmatrix}\neq 0$, (c) $\mathscr{A}$ 是满射. (提示: 利用习题7)

9. 设 $z\in\mathbb{C}$, 定义映射 $m_z:\mathbb{C}\to\mathbb{C},\ w\mapsto zw$,$\varphi:\mathbb{C}\to\mathbb{R}^2,\ x=x_1+x_2\mathrm{i}\mapsto(\mathrm{x}_1,\mathrm{x}_2)$, $f_z=\varphi\cdot m_z\cdot\varphi^{-1}:\mathbb{R}^2\to\mathbb{R}^2$ 试证明:

(1) $f_z:\mathbb{R}^2\to\mathbb{R}^2$ 是双射 (如果 $z\neq 0$).

(2) $f_z:\mathbb{R}^2\to\mathbb{R}^2$ 是线性算子 ($\mathbb{R}^2$表示标准 $\mathbb{R}$-向量空间).

(3) 如果 $z=a+b\mathrm{i}$, 求 $f_z(e_1)$ 和 $f_z(e_2)$. 并写出 f_z 对应的矩阵 $A(z)$.

(4) $\forall z\in\mathbb{C}$ ，令 $A(z)$ 表示 (3) 中的矩阵, 则

$$A(z+z')=A(z)+A(z'),A(zz')=A(z)\cdot A(z').$$

(5) 映射 $A:\mathbb{C}\to M_2(\mathbb{R})$ 是单射, 并求它的像集合 $K=Im(A)\subset M_2(\mathbb{R})$.

(6) 映射 $A:\mathbb{C}\to K,\ z\mapsto A(Z)$ 是域同构.

10. 设 $f(x)\in\mathbb{R}[x]$, 如果 $z\in\mathbb{C}$ 是 $f(x)$ 的根, 证明: z 的共轭 $\bar{z}$ 也是 $f(x)$ 的根. 利用此事实证明: 奇数次多项式 $f(x)\in\mathbb{R}[x]$ 必有实数根 (i.e. 存在 $\alpha\in\mathbb{R}$ 使 $f(\alpha)=0$).

11. 利用习题10 证明: $\mathbb{R}[x]$ 中的不可约多项式必为一次多项式. 或二次多项式 (提示: $(x-z)(x-\bar{z})$ 是实系数多项式).

12. 设 $f(x)=x^2+bx+c\in\mathbb{R}$. 证明:

(1) $f(x)$ 在 $\mathbb{R}[x]$ 中不可约 $\Leftrightarrow\triangle=b^2-4ac<0$.

(2) 如果 $z\in\mathbb{C}$ 是 $f(x)=x^2+bx+c\in\mathbb{R}$ 的一个根. 则当 $\triangle<0$ 时, $\mathbb{C}$ 中包含 $\mathbb{R}$ 和 z 的子域必为 $\mathbb{C}$.

第2章 线性代数初步

2.1 线性方程组与子空间

有许多应用和几何问题归结为求解下述的“线性方程组”:

$$\begin{cases} a_{11}x_1+a_{12}x_2+\cdots+a_{1n}x_n=b_1 \\ a_{21}x_1+a_{22}x_2+\cdots+a_{2n}x_n=b_n \\ \vdots \\ a_mx_1+a_mx_2+\cdots\cdots+a_{mn}x_n=b_m \end{cases} \tag{2-1}$$

其中 $x_1,x_2,\cdots,x_n$ 是不定元, $a_{ij}\in K$（K是一个域）称为 x_j 的系数. $a_{ij},b_i\in K$ 都是固定的数（或元素）. 我们称下述线性方程组

$$\begin{cases} a_{11}x_1+a_{12}x_2+\cdots+a_{1n}x_n=0 \\ a_{21}x_1+a_{22}x_2+\cdots+a_{2n}x_n=0 \\ \vdots \\ a_mx_1+a_mx_2+\cdots\cdots+a_{mn}x_n=0 \end{cases} \tag{2-2}$$

为与 (2-1) 相伴的“齐次线性方程组”. 齐次线性方程组有下述特别的性质:

定理 2.1.1 设 $V\subset K^n$ 是齐次线性方程组

$$\begin{cases} a_{11}x_1+a_{12}x_2+\cdots+a_{1n}x_n=0 \\ a_{21}x_1+a_{22}x_2+\cdots+a_{2n}x_n=0 \\ \vdots \\ a_{11}x_1+a_{m2}x_2+\cdots+a_{mn}x_n=0 \end{cases}$$

的解集合. 则 V 是 K^n 的子空间, 即 $\forall\alpha,\beta\in V,\lambda\in K$ 有 $\alpha+\beta\in V,\lambda\alpha\in V$.

证明: 设 $\alpha=(\alpha_1,\alpha_2,\cdots,\alpha_n)$, $\beta=(\beta_1,\beta_2,\cdots,\beta_n)\in V$. 即

$$\begin{cases} a_{11}\alpha_1+a_{12}\alpha_2+\cdots+a_{1n}\alpha_n=0 \\ a_{21}\alpha_1+a_{22}\alpha_2+\cdots+a_{2n}\alpha_n=0 \\ \vdots \\ a_{11}\alpha_1+a_{m2}\alpha_2+\cdots+a_{mn}\alpha_n=0 \end{cases} \quad \begin{cases} a_{11}\beta_1+a_{12}\beta_2+\cdots+a_{1n}\beta_n=0 \\ a_{21}\beta_1+a_{22}\beta_2+\cdots+a_{2n}\beta_n=0 \\ \vdots \\ a_{m1}\beta_1+a_{m2}\beta_2+\cdots+a_{mn}\beta_n=0 \end{cases}$$

所以不难验证 $\alpha+\beta=(\alpha_1+\beta_1,\alpha_2+\beta_2,\cdots,\alpha_n+\beta_n)$, $\lambda\alpha=(\lambda\alpha_1,\lambda\alpha_2,\cdots,\lambda\alpha_n)$ 也是齐次线性方程组:

$$\begin{cases} a_{11}x_1+a_{12}x_2+\cdots+a_{1n}x_n=0 \\ a_{21}x_1+a_{22}x_2+\cdots+a_{2n}x_n=0 \\ \vdots \\ a_{11}x_1+a_{m2}x_2+\cdots+a_{mn}x_n=0 \end{cases}$$

的解. 即: $\alpha+\beta\in V, \lambda\alpha\in V$. □

例 2.1.1 设 $L_1=\left\{(x_1,x_2)\in\mathbb{R}^2 \mid a_{11}x_1+a_{12}x_2=0\right\}\subset\mathbb{R}^2$ 是实平面上过原点的直线 (我们总假设 $a_{11},a_{12}\in\mathbb{R}$ 不全为0). 如果 $a_{11}\neq 0$, 则映射

$$\varphi: L_1\to\mathbb{R},\ (x_1,x_2)\mapsto x_2$$

是$\mathbb{R}$-向量空间的同构 $\left(\varphi^{-1}:\mathbb{R}\to L_1,\ x\mapsto\left(-\frac{a_{12}}{a_{11}}x,x\right)\right)$.

例 2.1.2 设 $L_1=\left\{(x_1,x_2,x_3)\in\mathbb{R}^3 \mid a_{11}x_1+a_{12}x_2+a_{13}x_3=0\right\}\subset\mathbb{R}^3$ 是三维空间中过原点的平面 (假设$(a_{11},a_{12},a_{13})\in\mathbb{R}^3$非零) 如果 $a_{1j}\neq 0$, 则 $L_1=\left\{\left(-\frac{a_{12}}{a_{11}}x_2-\frac{a_{13}}{a_{11}}x_3,x_2,x_3\right) \mid \forall x_2,x_3\in\mathbb{R}\right\}$ 所以

$$\varphi: L_1\to\mathbb{R}^2,\ (x_1,x_2,x_3)\mapsto(x_1,x_2)$$

是 $\mathbb{R}$-向量空间的同构$\left(\varphi^{-1}:\mathbb{R}^2\to L_1,\ (x,y)\mapsto\ \left(-\frac{a_{12}}{a_n}x-\frac{a_{13}}{a_{11}}y,x,y\right)\right)$.

例 2.1.3 设 $L_1=\{(x_1,x_2,\cdots,x_n)\in\mathbb{R}^n \mid a_{11}x_1+a_{12}x_2+\cdots+a_{1n}x_n=0\}\subset\mathbb{R}^n$ (假设$(a_{11},a_{12},\cdots,a_{1n})\neq 0$) 称为 n 维空间$\mathbb{R}^n$中的超平面. 如果$a_{1j}\neq 0$, 则 $L_1=\left\{\left(x_1,\cdots,x_{j-1}-\sum_{i\neq j}a_{1i}a_{ij}x_i,x_{j+1},\cdots,x_n\right) \mid \forall\left(x_1,\cdots,x_{j-1},x_{j+1},\cdots x_n\right)\in\mathbb{R}^{n-1}\right\}$ 所以

$$\varphi: L_1\to\mathbb{R}^{n-1},\ \left(x_1,\cdots x_{j+1},x_j,x_{j+1},\cdots x_n\right)\mapsto\left(x_1,\cdots x_{j-1},x_{j+1},\cdots,x_n\right)$$

是 $\mathbb{R}$-向量空间的同构

$$\left(\varphi^{-1}:\mathbb{R}^{n-1}\to L_1,\ (x_1,\cdots,x_{n-1})\mapsto\left(x_1,\cdots,x_{i-1},-\sum_{i+j}\frac{a_i}{a_j}x_i,x_{i+1},\cdots x_{n+1}\right)\right).$$

例 2.1.4

$$V=\left\{(x_1,x_2,\cdots,x_n)\in\mathbb{R}^n\;\middle|\;\begin{array}{c}a_{11}x_1+a_{12}x_2+\cdots+a_{1n}x_n=0\\ a_{21}x_1+a_{22}x_2+\cdots+a_{\mathrm{en}}x_n=0\\ \vdots\\ a_{m1}x_1+a_{m2}x_2,+\cdots+a_{mn}x_n=0\end{array}\right\}\subset\mathbb{R}^n$$

是 m 个超平面 $L_i=\{(x_1,x_2,\cdots,x_n)\in\mathbb{R}^n\mid a_{i1}x_1+a_{i2}x_2+\cdots+a_{in}x_n=0\}\subset\mathbb{R}^n$ $(1\leqslant i\leqslant m)$ 的交: $V=L_1\bigcap L_2\bigcap\cdots\bigcap L_m$, 一个自然的问题是: 求整数r和 $\mathbb{R}$-向量空间同构 $\varphi:V\to\mathbb{R}^n$, 这是我们将要讨论的问题之一.

对于一般的线性方程组 (2-1) ,有几个基本问题需要回答: (1) 方程组 (2-1) 是否有解 ? (2) 如果 (2-1) 有解, 何时有唯一解 ? 何时有无限多个解 ? (3) 如果 (2-1) 有唯一解, 是否可求出它的解公式 ? 如果 (2-1) 有无限多个解, 能否找出有限个解 $\eta_1,\cdots,\eta_k$ 使得其他所有解可由 $\eta_1,\cdots,\eta_k$ 表示 ? 在回答这些问题之前, 我们有下面重要的观察, 设 V 是 (2-2) 的解空间, 如果 $\alpha=(\alpha_1,\cdots,\alpha_n)$ 是 (2-1) 的一个解, 则

$$\alpha+V=\{\alpha+\beta\mid\forall\beta=(\beta_1,\cdots,\beta_n)\in V\}$$

是 (2-1) 的全部解. 事实上, 不难验证 $\alpha+\beta=(\alpha_1+\beta_1,\cdots,\alpha_n+\beta_n)$ 是 (2-1) 的解. 另一方面, 如果 $x=(x_1,\cdots,x_n)$ 是 (2-1) 的解, 则 $x-\alpha=(x_1-\alpha_1,\cdots,x_n-\alpha_n)$ 是 (2-2) 的解. 所以

$$x=\alpha+(x-\alpha)\in\alpha+V\quad(\text{令}\beta=x-\alpha\in V),$$

因此, (2-1) 有唯一解当且仅当 $V=\{0\}$. 即 (2-2) 只有零解! 反之, (2-1) 有无限多个解 (假设域 K 是无限集合).

上述基本问题的回答依赖于高斯消元法. 为简化记号, 令 $L_i=a_{i1}x_1+a_{i2}x_2+\cdots+a_{in}x_n$ 表示系数为 $a_{i1},a_{i2},\cdots,a_{in}$ 的线性函数 (或称一次齐次

多项式), 它由系数 $a_{i1}, a_{i2}, \cdots, a_{in}$ 唯一确定. 而方程组 (2-1) 可记为

$$\begin{cases} L_1 &= & b_1 \\ L_2 &= & b_2 \\ & \vdots & \\ L_m &= & b_m \end{cases}$$

所谓高斯消元法是一种算法, 它由几个基本的操作 (称为初等变换) 组成:

$$\begin{cases} L_1 &= & b_1 \\ & \vdots & \\ L_i &= & b_i \\ & \vdots & \\ L_j &= & b_j \\ & \vdots & \\ L_m &= & b_m \end{cases} \xleftrightarrow{\text{(I)}} \begin{cases} L_1 &= & b_1 \\ & \vdots & \\ L_j &= & b_j \\ & \vdots & \\ L_i &= & b_i \\ & \vdots & \\ L_m &= & b_m \end{cases}$$

初等变换 (I) : 交换第 i 个方程与第 j 个方程的位置.

$$\begin{cases} L_1 &= & b_1 \\ & \vdots & \\ L_i &= & b_i \\ & \vdots & \\ L_j &= & b_j \\ & \vdots & \\ L_m &= & b_m \end{cases} \xleftrightarrow{\text{(II)}} \begin{cases} L_1 &= & b_1 \\ & \vdots & \\ L_i &= & b_i \\ & \vdots & \\ cL_i + L_j &= & cb_i + b_j \\ & \vdots & \\ L_m &= & b_m \end{cases}$$

初等变换 (II) : 将第 i 个方程乘以常数 $c \in K$ 加到与第 j 个方程.

$$\begin{cases} L_1 & = & b_1 \\ & \vdots & \\ L_i & = & b_i \\ & \vdots & \\ L_m & = & b_m \end{cases} \overset{\text{(III)}}{\longleftrightarrow} \begin{cases} L_1 & = & b_1 \\ & \vdots & \\ cL_i & = & cb_i \\ & \vdots & \\ L_m & = & b_m \end{cases}$$

初等变换 (III) : 将第 i 个方程乘以非零常数 $c \in K$.

引理 2.1.1 如果方程组 $(*)'$ 是由方程组 $(*)$ 经初等变换得到, 则

(1) $(*)'$ 有解 $\Leftrightarrow$ $(*)$ 有解;

(2) $\alpha = (\alpha_1, \cdots, \alpha_n)$ 是 $(*)'$ 的解 $\Leftrightarrow$ $\alpha = (\alpha_1, \cdots, \alpha_n)$ 是 $(*)$ 的解.

证明: 由于初等变换都可逆, 所以 $(*)$ 也可由 $(*)'$ 经初等变换得到. 因此只需证明: 如果 $(*)'$ 是由 $(*)$ 经一次初等变换得到, 则 $\alpha = (\alpha_1, \cdots, \alpha_n)$ 是 $(*)'$ 的解 $\Leftrightarrow$ $\alpha = (\alpha_1, \cdots, \alpha_n)$ 是 $(*)$ 的解. 对于第 I 类, 第 III 类初等变换, 这是显然的. 对于第 II 类初等变换. 如果 $\alpha = (\alpha_1, \cdots, \alpha_n)$ 是 $(*)'$ 的解, 则

$$L_k(\alpha_1, \cdots, \alpha_n) = b_k \, (k \neq j), cL_i(\alpha_1, \cdots, \alpha_n) + L_j(\alpha_1, \cdots, \alpha_n) = cb_i + b_j$$

所以 $L_j(\alpha_1, \cdots, \alpha_n) = b_j$, 从而 $\alpha = (\alpha_1, \cdots, \alpha_n)$ 是 $(*)$ 的解, 反之亦然. □

定理 2.1.2 (高斯消元法). 经过有限次初等变换, 方程组

$$\begin{cases} L_1 = a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1 \\ L_2 = a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2 \\ \vdots \\ L_m = a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = b_m \end{cases} \tag{2-3}$$

(不妨设 x_1 的系数 $a_{11}, a_{21}, \cdots, a_{m1}$ 不全为零) 可化简为下列的 “阶梯型” 方程

组:

$$
\begin{cases}
\overline{L_1} = & x_1 + \overline{a_{12}}x_2 + \cdots + \overline{a_{1n}}x_n & = \overline{b_1} \\
\overline{L_2} = & x_{i_2} + \overline{a_{2i_2+1}}x_{i_2+1} + \cdots + \overline{a_{2n}}x_n & = \overline{b_2} \\
\quad\vdots & \ddots & \\
\overline{L_r} = & x_{i_r} + \overline{a_{ri_r+1}}x_{i_r+1} + \cdots + \overline{a_{rn}}x_n & = \overline{b_r} \\
\overline{L_{r+1}} = & 0x_1 + \cdots + 0x_{i_r} + 0x_{i_r+1} + \cdots + 0x_n & = \overline{b_{r+1}} \\
\quad\vdots & & \\
\overline{L_m} = & 0x_1 + \cdots + 0x_{i_r} + \cdots + 0x_n & = \overline{b_{r+1}}
\end{cases}
\tag{2-4}
$$

其中 $1 < i_2 < i_3 < \cdots < i_r$.

证明: 交换方程次序(初等变换 (I)), 可设 $a_{11} \neq 0$, 则通过若干次初等变换 (II) , (III) , 可化为

$$
\begin{cases}
\frac{1}{a_{11}}L_1 = x_1 + \frac{a_{12}}{a_{11}}x_2 + \cdots + \frac{a_{1n}}{a_{11}}x_n = \frac{1}{a_{11}}b_1 \\
L_2 - \frac{a_{21}}{a_{11}}L_1 = 0x_1 + \left(a_{22} - \frac{a_{21}}{a_{11}}a_{12}\right)x_2 + \cdots + \left(a_{2n} - \frac{a_{21}}{a_{11}}a_{1n}\right)x_n = b_2 - \frac{a_{21}}{a_{11}}b_1 \\
\vdots \\
L_m - \frac{a_{m1}}{a_{11}}L_1 = 0x_1 + \left(a_{m2} - \frac{a_{m1}}{a_{11}}a_{12}\right)x_2 + \cdots + \left(a_{mn} - \frac{a_{m1}}{a_{11}}a_{1n}\right)x_n = b_m - \frac{a_{m1}}{a_{11}}b_{11}
\end{cases}
\tag{2-5}
$$

令

$$
\begin{aligned}
\overline{L_1} &= \frac{1}{a_{11}}L_1, \quad L'_2 = L_2 - \frac{a_{21}}{a_{11}}L_1, \cdots, L'_m = L_m - \frac{a_{m1}}{a_{11}}L_1 \\
\overline{b_1} &= \frac{1}{a_{11}}b_1, \quad b'_2 = b_2 - \frac{a_{21}}{a_{11}}b_1, \cdots b'_m = b_m - \frac{a_{m1}}{a_{11}}b_1
\end{aligned}
$$

则上述方程组 (2-5) 可写成

$$
\begin{cases}
\overline{L_1} = x_1 + \overline{a_{12}}x_2 + \cdots + \overline{a_{1n}}x_n = b_1 \\
L'_2 = a'_{2i_2}x_{i_2} + \cdots + a'_{2n}x_n = b'_2 \\
\vdots \\
L'_m = a'_{mi_2}x_{i_2} + \cdots + a'_{mn}x_n = b'_m
\end{cases}
$$

(其中 x_{i_2} 的系数 $a'_{2i_2}, \cdots, a'_{mi_2}$ 不全为零), 对方程组 $\begin{cases} L'_2 = b'_2 \\ \vdots \\ L'_m = b'_m \end{cases}$ 重复上述操作步骤可得结论. □

我们有下述重要观察: (1) 方程组 (2-4) 中的每个方程 $\overline{L_i} = \overline{b_i}$ 都是方程 $L_1 = b_1$, $L_2 = b_2, \cdots, L_m = b_m$ 的"线性组合". 即存在常数 $\lambda_{i_1}, \lambda_{i_2}, \cdots, \lambda_{i_m} \in K$ 使得:

$$\overline{L_i} = \lambda_{i_1} L_1 + \lambda_{i_2} L_2 + \cdots + \lambda_{i_m} L_m$$

$$\overline{b_i} = \lambda_{i_1} b_1 + \lambda_{i_2} b_2 + \cdots + \lambda_{i_m} b_m$$

反之亦然: 每个方程 $L_i = b_i$ 都是 $L_1 = b_1$, $L_2 = b_2, \cdots, L_m = b_m$ 的"线性组合".

(2) 对任意线性方程组 $\begin{cases} L_1 = b_1 \\ \vdots \\ L_m = b_m \end{cases}$ 存在 r 个方程 (无妨设 $L_1 = b_1, \cdots, L_r = b_r$) 满足: 对任意 $i > r$ 存在常数 $\lambda_{i_1}, \lambda_{i_2}, \cdots, \lambda_{i_m} \in K$ 使

$$L_i = \lambda_{i_1} L_1 + \cdots + \lambda_{i_r} L_r, \quad \overline{b_i} = b_i - (\lambda_{i_1} b_1 + \cdots + \lambda_{i_r} b_r)$$

即 $\{L_{r+1}, \cdots, L_m\}$ 可由 $\{L_1, \cdots, L_r\}$ 线性表出. (3) $r \leqslant \min\{m, n\}$.

推论 2.1.1 (1) 方程组 (2-3) 有解 $\Leftrightarrow$ 方程组 (2-4) 有解 $\Leftrightarrow \overline{b_{r+1}} = \cdots = \overline{b_m} = 0$.
(2) 如果方程组 (2-3) 有解 (*i.e.* $\overline{b_{r+1}} = \cdots = \overline{b_m} = 0$), 则 $x = (x_1, \cdots, x_n)$ 是方程组 (2-3) 的解 $\Leftrightarrow$

$$\begin{cases} x_1 = \overline{b_1} - \overline{a_{12}} x_2 - \cdots - \overline{a_{1n}} x_n \\ x_{i_2} = \overline{b_2} - \overline{a_{2i_2+1}} x_{i_2+1} - \cdots - \overline{a_{2n}} x_n \\ \vdots \\ x_{i_r} = \bar{b}_r - \overline{a_{ri_{r+1}}} x_{i_{r+1}} - \cdots - \overline{a_{rn}} x_n \end{cases}$$

特别: 方程组 (2-3) 有无穷多组解 $\Leftrightarrow r < n$, 方程组 (2-3) 有唯一解 $\Leftrightarrow r = n$.
(3) 齐次线性方程组 (2-2) 有非零解 $\Leftrightarrow r < n$. 特别: 当 $m < n$ 时, 方程组 (2-2) 有无穷多组解.

例 2.1.5 线性方程组

$$\begin{cases} L_1 = x_1 + x_2 - 2x_3 + 2x_4 = b_1 \\ L_2 = 3x_1 + 5x_2 + 6x_3 - 4x_4 = b_2 \\ L_3 = 4x_1 + 5x_2 - 2x_3 + 3x_4 = b_3 \\ L_4 = 3x_1 + 8x_2 + 24x_3 - 19x_4 = b_4 \end{cases}$$

何时有解 ? 如果有解, 求出它的一般解.

解: 做初等变换可得

$$L_2' = L_2 - 3L_1 = 2x_2 + 12x_3 - 10x_4 = b_2 - 3b_1$$

$$L_3' = L_3 - 4L_1 = x_2 + 6x_3 - 5x_4 = b_3 - 4b_1$$

$$L_4' = L_4 - 3L_1 = 5x_2 + 30x_3 - 25x_4 = b_4 - 3b_1$$

化为“阶梯型”:

$$\begin{cases} \overline{L_1} = L_1 = x_1 + x_2 - 2x_3 + 2x_4 = b_1 = \overline{b_1} \\ \overline{L_2} = \frac{1}{2}L_2' = x_2 + 6x_3 - 5x_4 = \frac{1}{2}b_2 - \frac{3}{2}b_1 = \overline{b_2} \\ L_3 = L_3' - \overline{L}_2 = 0x_1 + 0x_2 + 0x_3 + 0x_4 = b_3 - 4b_1 - \frac{1}{2}b_2 + \frac{3}{2}b_1 = \overline{b_3} \\ \overline{L_4} = L_4' - 5\overline{L_2} = 0x_1 + 0x_2 + 0x_3 + 0x_4 = b_4 - 3b_1 - \frac{5}{2}b_2 + \frac{15}{2}b_1 = \overline{b_4} \end{cases}$$

不难看出, $\overline{L_1} = L_1$, $\overline{L_2} = -\frac{3}{2}L_1 + \frac{1}{2}L_2$, $\overline{L_3} = -\frac{5}{2}L_1 - \frac{1}{2}L_2 + L_3$, $\overline{L_4} = \frac{9}{2}L_1 - \frac{5}{2}L_2 + L_4$, $\overline{b_1} = b_1$, $\overline{b_2} = -\frac{3}{2}b_1 + \frac{1}{2}b_2$, $\overline{b_3} = -\frac{5}{2}b_1 - \frac{1}{2}b_2 + b_3$, $\overline{b_4} = \frac{9}{2}b_1 - \frac{5}{2}b_2 + b_4$, $r = 2$. 所以, 上述方程组有解 $\Leftrightarrow \overline{b_3} = 0$, $\overline{b_4} = 0$ $\Leftrightarrow \begin{cases} 5b_1 + b_2 - 2b_3 & = 0 \\ b_1 - 5b_2 + 2b_4 & = 0 \end{cases} \Leftrightarrow \begin{cases} b_3 & = \frac{5}{2}b_1 + \frac{1}{2}b_2 \\ b_4 & = -\frac{1}{2}b_1 + \frac{5}{2}b_2 \end{cases}$.

当 b_1, b_2, b_3, b_4 满足该条件时, $x = (x_1, x_2, x_3, x_4)$ 是一组解

$$\Leftrightarrow \begin{cases} x_2 = -\frac{3}{2}b_1 + \frac{1}{2}b_2 - 6x_3 + 5x_4 \\ x_1 = -x_2 + 2x_3 + 2x_4 + b_1 = \frac{5}{2}b_1 - \frac{1}{2}b_2 + 8x_3 - 7x_4 \end{cases}$$

令 $x_3 = x_4 = 0$,得方程组的一个特解 $\alpha = \left(\frac{5}{2}b_1 - \frac{1}{2}b_2, -\frac{3}{2}b_1 + \frac{1}{2}b_2, 0, 0\right)$.

方程组

$$\begin{cases} x_1 + x_2 - 2x_3 + 2x_4 = b_1 \\ 3x_1 + 5x_2 + 6x_3 - 4x_4 = b_2 \\ 4x_1 + 5x_2 - 2x_3 + 3x_4 = \frac{5}{2}b_1 + \frac{1}{2}b_2 \\ 3x_1 + 8x_2 + 24x_3 - 19x_4 = -\frac{1}{2}b_1 + \frac{5}{2}b_2 \end{cases}$$

对任意常数 b_1, b_2 都有解, 与它相伴的齐次线性方程组的解空间是

$$V = \left\{ (x_1, x_2, x_3, x_4) \;\middle|\; \begin{array}{l} x_1 = 8x_3 - 7x_4 \\ x_2 = -6x_3 + 5x_4 \end{array} \right\}$$

令 $x_3 = 1, x_4 = 0$ 得 $\eta_1 = (8, -6, 1, 0) \in V$, 令 $x_3 = 0, x_4 = 1$ 得 $\eta_= (-7, -5, 0, 1) \in V$ 且 V 中的任一组解 $x = (x_1, x_2, x_3, x_4)$ 唯一表示成:

$$x = (8x_3 - 7x_4, -6x_3 + 5x_4, x_3, x_4) = (8x_3, -6x_3, x_3, 0) + (-7x_4, 5x_4, 0, x_4)$$
$$= x_3(8, -6, 1, 0) + x_4(-7, 5, 0, 1) = x_3\eta_1 + x_3\eta_2$$

所以 V 可以写成 $V = \{\lambda_1\eta_1 + \lambda_2\eta_2 \mid \forall\lambda_1, \lambda_2 \in K\}$. 而方程组的一般解可写成

$$x = \lambda_1\eta_1 + \lambda_2\eta_2 \quad (\forall\lambda_1, \lambda_2 \in K)$$

其中 $\alpha = \left(\frac{5}{2}b_1 - \frac{7}{b_2}, -\frac{3}{2}b_1 + \frac{1}{2}b_2, 0, 0\right)$ 是方程组的一组特解.

注意: $L_3 = \frac{5}{2}L_1 + \frac{1}{2}L_2$, $L_4 = -\frac{9}{2}L_1 + \frac{5}{2}L_2$, 即: L_3, L_4 可由 L_1, L_2 "线性表出", 而 L_1, L_2 "线性无关". 即: 如果 $\lambda_1L_1 + \lambda_2L_2 = 0$, 则必有 $\lambda_1 = \lambda_2 = 0$.

为了简化符号, 我们下面引入线性方程组的矩阵表达:

定义 2.1.1 形如 $A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}$ 的表格称为 $m \times n$ 矩阵 (或称为一个 m 行, n 列的矩阵): 其中, $A_{(i)} = (a_{i1}, a_{i2}, \cdots, a_{in})$ 称为 A 的第 i 行 (或 A 的第 i 个 "行向量"), 而 $A^{(j)} = \begin{pmatrix} a_{1j} \\ a_{2j} \\ \cdots \\ a_{mj} \end{pmatrix}$ 称为 A 的第 j 列 (或 A 的第 j 个 "列向量").

有时也把矩阵 A 记为 $\left(a_{ij}\right)_{m\times n}$, 其中 $\left(a_{ij}\right)$ 表示位于第 i 行 , 第 j 列 的元素. 两个矩阵 $A = \left(a_{ij}\right)_{m\times n}, B = \left(b_{ij}\right)_{m'\times n'}$ 相等当且仅当 $m = m', n = n'$ 且 $a_{ij} = b_{ij}$ 对任意 $1 \leqslant i \leqslant m, 1 \leqslant j \leqslant n$ 成立. 由于 $L_i = a_{i1}x_1 + a_{i2}x_2 + \cdots + a_{in}x_n$ 由它系数向量 $(a_{i1}, a_{i2}, \cdots,)$ 唯一确定, $L_i \leftrightarrow (a_{i1}, a_{i2}, \cdots, a_{in})$ 且

$$L_1 + L_2 \leftrightarrow (a_{11}, a_{12}, \cdots, a_{1n}) + (a_{21}, a_{22}, \cdots, a_{2n}) = (a_{11} + a_{21}, a_{12} + a_{22}, \cdots, a_{1n} + a_{2n})$$

$$\lambda L_i \leftrightarrow \lambda\,(a_{i1}, a_{i2}, \cdots, a_{in}) = (\lambda a_{i1}, \lambda a_{i2}, \cdots, \lambda a_{in})\,;$$

线性方程组 (2-1) 由它的 "系数矩阵"

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

和 “增广矩阵”

$$
\widetilde{A} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} & b_1 \\ a_{21} & a_{22} & \cdots & a_{2n} & b_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} & b_m \end{pmatrix}
$$

确定, 对 (2-1) 施行 “初等变换” 等价于对 $\widetilde{A}$ 的行向量施行相应的 “初等变换”.

推论2.2 (高斯消元法). 如果 $\widetilde{A}$ 的第1 列不全为零, 则通过对 $\widetilde{A}$ 的行向量施行 “初等变换”, $\widetilde{A}$ 可化为阶梯型.

$$
\begin{pmatrix} 1 & \overline{a_{12}} & \cdots & \cdots & \cdots & \cdots & \overline{a_{1n}} & \overline{b_1} \\ 0 & \cdots & 1 & \overline{a_{2i_2+1}} & \cdots & \cdots & \overline{a_{2n}} & \overline{b_2} \\ \cdots & & \cdots & & \cdots & & \cdots & \\ 0 & 0 & \cdots & 1 & \overline{a_{ri_r+1}} & \cdots & \overline{a_{rn}} & \overline{b_r} \\ 0 & 0 & \cdots & \cdots & \cdots & \cdots & \cdots & \overline{b_{r+1}} \\ \cdots & & \cdots & & \cdots & & \cdots & \\ 0 & 0 & \cdots & 0 & \cdots & 0 & \cdots & \overline{b_n} \end{pmatrix}
$$

在之前的讨论中, 我们将 K^n 中的元素用 $x = (x_1, \cdots, x_n)$ 表示. 为了应用方便, 我们约定: K^n 即可以表示 “行向量” $x = (x_1, \cdots, x_n)$ 的集合, 也可以表示 “列向

量” $x = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$ 的集合:

$$K^n = \{(x_1, \cdots x_n) \mid \forall x_i \in k\} = \left\{\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \mid \forall x_i \in K\right\}$$

其中 $\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} \Leftrightarrow x_i = y_i$; $\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} + \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} = \begin{pmatrix} x_1 + y_1 \\ x_2 + y_2 \\ \vdots \\ x_n + y_n \end{pmatrix}$, $\lambda \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} \lambda x_1 \\ \lambda x_2 \\ \vdots \\ \lambda x_n \end{pmatrix}$. 这样方程组

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_n \\ \vdots \\ a_m x_1 + a_m x_2 + \cdots\cdots + a_{mn}x_n = b_m \end{cases} \tag{2-6}$$

可以写成

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

或

$$x_1 A^{(1)} + x_2 A^{(2)} + \cdots + x_n A^{(n)} = b, \text{其中} b = \begin{pmatrix} b_1 \\ b_2 \\ \cdots \\ b_m \end{pmatrix}$$

所以, 我们可以说: 方程组 (2-6) 有解 $\Leftrightarrow$ 增广矩阵 $\widetilde{A}=(A,b)$ 的最后一列 b 可由 A 的“列向量”线性表出.

习题2.1

1. 求下列线性方程组的一个特解, 并写完它们的一般解.

$$\begin{cases} 5x_1+3x_2+5x_3+12x_4=10 \\ 2x_1+2x_2+3x_3+5x_4=4 \\ x_1+7x_2+9x_3+4x_4=2 \end{cases} \qquad \begin{cases} -6x_1+8x_2-5x_3-x_4=9 \\ -2x_1+4x_2+7x_3+3x_4=1 \\ -3x_1+5x_2+4x_3+2x_4=3 \\ -3x_1+7x_2+17x_3+7x_4=\lambda \end{cases}$$

(确定 λ 的值使它有解)

2. 求二次实系数多项式 $f(x)\in\mathbb{R}[x]$ 使 $f(1)=8, f(-1)=2, f(2)=14$.

3. 求三次实系数多项式 $f(x)\in\mathbb{R}[x]$ 使 $f(-2)=1, f(-1)=3, f(1)=13, f(2)=33$.

4. 对行向量施行初等变换, 将下列矩阵化为阶梯型

$$A=\begin{pmatrix} 5 & 6 & -2 & 7 & 4 \\ 2 & 3 & -1 & 4 & 2 \\ 7 & 9 & -3 & 5 & 6 \\ 5 & 9 & -3 & 1 & 6 \end{pmatrix}, B=\begin{pmatrix} 1 & 1 & -2 & 2 \\ 3 & 5 & 6 & -4 \\ 4 & 5 & -2 & 3 \\ 3 & 8 & 24 & -19 \end{pmatrix}$$

5. 设 $V\subset\mathbb{R}^4$ 是齐次线性方程组

$$\begin{cases} x_1+x_2-2x_3+2x_4=0 \\ 3x_1+5x_2+6x_3-4x_4=0 \\ 4x_1+5x_2-2x_3+3x_4=0 \\ 3x_1+8x_2+24x_3-19x_4=0 \end{cases}$$

的解空间. 证明存在 $\alpha,\beta\in V$ 满足下列条件:

(1) α,β 线性无关 (即: 如果 $\lambda_1\alpha+\lambda_2\beta=0$, 则 $\lambda_1=\lambda_2=0$).

(2) V 中任一向量可由 α,β 线性表出.(即: $\forall x\in V$, 存在 $\lambda_1,\lambda_2\in\mathbb{R}$ 使 $x=\lambda_1\alpha+\lambda_2\beta$)

6. 设 $\alpha_1=\begin{pmatrix}5\\2\\1\end{pmatrix},\alpha_2=\begin{pmatrix}3\\2\\7\end{pmatrix},\alpha_3=\begin{pmatrix}5\\3\\9\end{pmatrix},\alpha_4=\begin{pmatrix}12\\5\\4\end{pmatrix}$,证明存在不全为零的数 x_1,x_2,x_3,x_4 使

$$x_1\alpha_1+x_2\alpha_2+x_3\alpha_3+x_4\alpha_4=0$$

(此时称 $\alpha_1,\alpha_2,\alpha_3,\alpha_4$ 线性无关).

2.2 线性映射与矩阵

设 K^n,K^m 表示由列向量组成的标准 K-向量空间. 其中 K 是一个域(例如, $K=\mathbb{Q},\mathbb{R},\mathbb{C}$,或它们的子域).

定义 2.2.1 映射 $f:K^n\to K^m$ 称为 K-线性映射:

如果 $\forall\lambda\in K,\alpha,\beta,\in K^n$,有

$$f(\alpha+\beta)=f(\alpha)+f(\beta)$$

$$f(\lambda\alpha)=\lambda f(\alpha).$$

例 2.2.1 令 $M_{m\times n}(K)=\left\{\left(a_{ij}\right)_{m\times n}\mid a_{ij}\in K\right\}$ 表示所有 K 上的 $m\times n$ 矩阵的集合, 则对任意 $A=\left(a_{ij}\right)_{m\times n}\in M_{m\times n}(K)$,定义映射 $\varphi_A:K^n\to K^m$ 如下: 设 $A^{(1)}=\begin{pmatrix}a_{11}\\a_{21}\\\cdots\\a_{m1}\end{pmatrix},A^{(2)}=\begin{pmatrix}a_{12}\\a_{22}\\\cdots\\a_{m2}\end{pmatrix},\cdots,A^{(n)}=\begin{pmatrix}a_{1n}\\a_{2n}\\\cdots\\a_{mn}\end{pmatrix}\in K^n$ 表示矩阵 A 的列向量. $\forall x=\begin{pmatrix}x_1\\x_2\\\vdots\\x_n\end{pmatrix}\in K^n$, 定义

$$\varphi_A(x)=x_1A^{(1)}+x_2A^{(2)}+\cdots+x_nA^{(n)}$$

则 $\varphi_A:K^n\to K^m$ 是一个 K-线性映射.

引理 2.2.1 设 $L(K^n,K^m)=\left\{f:K^n\to K^m\mid f\text{是}K-\text{线性映射}\right\}$ 表示所有从 K^n 到 K^m 的 K-线性映射的集合. 则映射

$$\varphi:M_{m\times n}(K)\to L(K^n,K^m),\ A\mapsto\varphi_A$$

是一个双射.

证明： 首先证明 φ 是单射: $\forall A,B\in M_{m\times n}(K)$, 如果 $\varphi_A=\varphi_B$. 则 $\forall x\in K^n,\varphi_A(x)=\varphi_B(x)$. 即: 对任意 $x=\begin{pmatrix}x_1\\ \vdots\\ x_n\end{pmatrix}\in K^n$, 有 $x_1A^{(1)}+\cdots+x_nA^{(n)}=x_1B^{(1)}+\cdots+x_nB^{(n)}$. 从而 $A^{(i)}=B^{(i)}\,(1\leqslant i\leqslant n)$. 所以 $A=B$.

下面证明 φ 是满射: $\forall f\in L(K^n,K^m)$, 只需证明存在 $A\in M_{m\times n}(K)$ 使 $\varphi_A=f$, 考虑 f 在 $e_1=\begin{pmatrix}1\\0\\0\\ \vdots\\0\end{pmatrix},e_2=\begin{pmatrix}0\\0\\1\\ \vdots\\0\end{pmatrix},\cdots,e_n=\begin{pmatrix}0\\0\\ \vdots\\0\\1\end{pmatrix}\in K^n$ 的像: $f(e_1)=\begin{pmatrix}a_{11}\\a_{21}\\ \vdots\\a_{m1}\end{pmatrix},f(e_2)=\begin{pmatrix}a_{12}\\a_{22}\\ \vdots\\a_{m2}\end{pmatrix},\cdots,f(e_n)=\begin{pmatrix}a_{1n}\\a_{2n}\\ \vdots\\a_{mn}\end{pmatrix}\in K^m$, 令 $A\in M_{m\times n}(K)$ 使 $A^{(1)}=f(e_1),A^{(2)}=f(e_2),\cdots,A^{(n)}=f(e_n)$. 则对任意 $x=\begin{pmatrix}x_1\\ \vdots\\ x_n\end{pmatrix}\in K^n$, 由于 $x=x_1e_1+x_2e_2+\cdots+x_ne_n$, 且 f 是 K-线性映射, 我们有

$$f(x)=x_1f(e_1)+x_2f(e_2)+\cdots+x_nf(e_n)=x_1A^{(1)}+x_2A^{(2)}+\cdots+x_nA^{(n)}=\varphi_A(x)$$

所以 $f=\varphi_A$. 因此, $\varphi:M_{m\times n}(K)\to L(K^n,K^m)$, $A\mapsto\varphi_A$ 是双射. 如果我们将矩阵 $A\in M_{m\times n}$ 写成列向量表达式: $A=\left(A^{(1)},A^{(2)},\cdots,A^{(n)}\right)$, $A^{(i)}\in K^m$, 则 $\varphi^{-1}:L(K^n,K^m)\to M_{m\times n}(K)$ 可表示为:

$$f\mapsto(f(e_1),f(e_2),\cdots,f(e_n)).$$

□

例 2.2.2 在 $L(K^n,K^m)$ 上定义 "加法" 和 "数乘法" :

$$\forall f,g\in L(K^n,K^m),\lambda\in K,$$

$$f+g: K^n \to K^m,\ x \mapsto f(x)+g(x)$$

$$\lambda f: K^n \to K^m,\ x \mapsto \lambda f(x)$$

则 $L(K^n, K^m)$ 关于它们是一个 K-向量空间, 通过双射

$$\varphi^{-1}:\ L(K^n, K^m) \to M_{m\times n}(K)$$

可以定义 $M_{m\times n}(K)$ 上的“加法”和“数乘法”使得 $\varphi^{-1}:\ L(K^n, K^m) \to M_{m\times n}(K)$ 是 K-向量空间的同构. 不难看出, $M_{m\times n}(K)$ 上通过 φ^{-1} 定义的“加法”和“数乘法”如下: $\forall\, A=\left(a_{ij}\right)_{m\times n}, B=\left(b_{ij}\right)_{m\times n}$, $\lambda\in K$, 则 $A+B=\left(a_{ij}+b_{ij}\right)_{m\times n}$, $\lambda A=\left(\lambda a_{ij}\right)_{m\times n}$. 因此我们引入矩阵的加法和数乘法的定义如下.

定义 2.2.2 $\forall\, A=\left(a_{ij}\right)_{m\times n}, B=\left(b_{ij}\right)_{m\times n}$, $\lambda\in M_{m\times n}(K)$, $\lambda\in K$, 则

$$A+B=\left(a_{ij}+b_{ij}\right)_{m\times n},\ \lambda A=\left(\lambda a_{ij}\right)_{m\times n}.$$

引理 2.2.2 $M_{m\times n}(K)$ 关于定义 2.3 中的运算成为一个 K- 向量空间, 且双射 $\varphi: M_{m\times n}(K) \to L(K^n, K^m)$, $A \mapsto \varphi_A$ 是 K-向量空间之间的同构.

证明: $M_{m\times n}(K)$ 关于矩阵的“加法”和“数乘法”是一个 K-向量空间可直接验证. (实际上, 它是一个标准 K-向量空间 K^{mn}).证明双射 φ 是同构, 只需验证(留作练习):

$$\varphi_{A+B}=\varphi_A+\varphi_B,\ \varphi_{\lambda A}=\lambda\varphi_A.$$

□

引理 2.2.3 设 $A=\left(a_{ij}\right)_{m\times n}\in M_{m\times n}(K), B=\left(b_{ij}\right)_{n\times s}\in M_{n\times s}(K)$ 则线性映射 $\varphi_B: K^s\to K^n$, $\varphi_A: K^n\to K^m$ 的合成映射 $f=\varphi_A\cdot\varphi_B: K^s\to K^m$ 也是 K- 线性映射, 且 f 由矩阵 $C=\left(c_{ij}\right)_{m\times s}\in M_{m\times s}(K)$ 确定. 其中

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

证明: 可以直接验证 $f=\varphi_A\cdot\varphi_B$ 是 K-线性映射:

$$\forall\, \alpha,\beta\in K^s,\ \lambda\in K,\ f(\alpha+\beta)=\varphi_A(\varphi_B(\alpha)+\varphi_B(\beta))=f(\alpha)+f(\beta)$$

$$f(\lambda\alpha)=\varphi_A(\lambda\varphi_B(\alpha))=\lambda\varphi_A(\varphi_B(\alpha))=\lambda f(\alpha)$$

所以, 由引理 2.2 , 存在唯一的矩阵 $C=\left(c_{ij}\right)_{m\times s}$ 使 $f=\varphi_c$. 即: $\forall\ x=\begin{pmatrix}x_1\\ \vdots\\ x_s\end{pmatrix}\in K^s$,

$$f(x)=\varphi_c(x)=x_1C^{(1)}+x_2C^{(2)}+\cdots+x_sC^{(s)}.$$

为了计算 c_{ij}, 我们来求矩阵 $C=\left(c_{ij}\right)_{m\times s}$ 的第 j 列 $C^{(j)}$. 显然, $C^{(j)}=f\left(e_j\right)$, 其中 $e_j=\begin{pmatrix}0\\ \vdots\\ 0\\ 1\\ 0\\ \vdots\\ 0\end{pmatrix}$ (1 位于第 j 个位置). 由定义, $\varphi_B\left(e_j\right)=B^{(j)}\in K^n$, 所以

$$C^{(j)}=f\left(e_j\right)=\varphi_A\left(B^{(j)}\right)=b_{1j}A^{(1)}+b_{2j}A^{(2)}+\cdots+b_{nj}A^{(n)}$$

即

$$\begin{pmatrix}c_{1j}\\ c_{2j}\\ \vdots\\ c_{mj}\end{pmatrix}=\begin{pmatrix}a_{11}b_{1j}\\ a_{21}b_{1j}\\ \vdots\\ a_{m1}b_{1j}\end{pmatrix}+\begin{pmatrix}a_{12}b_{2j}\\ a_{22}b_{2j}\\ \vdots\\ a_{m2}b_{2j}\end{pmatrix}+\cdots+\begin{pmatrix}a_{1n}b_{nj}\\ a_{2n}b_{nj}\\ \vdots\\ a_{mn}b_{nj}\end{pmatrix}$$

所以

$$c_{ij}=a_{i1}b_{1j}+a_{i2}b_{2j}+\cdots+a_{in}b_{nj}\ .$$

□

定义 2.2.3 将上述引理 2.4 中的矩阵 $C=\left(c_{ij}\right)_{m\times s}$ 记为 $A=B\cdot C$,称为 A,B 的乘积. 所以, 如果 $A=\left(a_{ij}\right)_{m\times n}$, $B=\left(b_{ij}\right)_{n\times s}$, 则 $A\cdot B=\left(c_{ij}\right)_{m\times s}$ 其中

$$c_{ij}=\sum_{k=1}^{n}a_{ik}b_{kj}\ .$$

注记 2.2.1 A,B 有意义 $\Leftrightarrow$ A 的 "列数' = B 的 "行数". 所以, 有可能 $A\cdot B$ 有意义, 但 $B\cdot A$ 没有意义. 特别地, 在 $M_{m\times n}(K)$ 中, $A\cdot B, B\cdot A$ 都有意义, 但不一定

相等.

定义 2.2.4 行数与列数相等的矩阵称为方阵. 用 $M_n(k) = \left\{\left(a_{ij}\right)_{n\times n} \mid \quad a_{ij} \in K\right\}$ 表示所有 $n \times n$ 矩阵的集合(也称 n 阶方阵的集合).

例 2.2.3 行向量 $(a_1, \cdots, a_n)$ 可以看成 $1 \times n$ 矩阵. 而列向量 $\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ 可以看成 $n \times 1$ 矩阵. 根据矩阵乘法,

$$(a_1, \cdots, a_n)\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = a_1x_1 + a_2x_2 + \cdots + a_nx_n .$$

由矩阵相等的定义, 线性方程组 (2.1) 可以写成矩阵形式

$$\begin{pmatrix} \sum_{j=1}^n a_{1j}x_j \\ \sum_{j=1}^n a_{2j}x_j \\ \vdots \\ \sum_{j=1}^n a_{mj}x_j \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix}$$

根据矩阵乘法, 即

$$\begin{pmatrix} a_{11} & a_{12}, & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix}$$

令

$$A = \begin{pmatrix} a_{11} & a_{12}, & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}, \ b = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix}, \ X = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_m \end{pmatrix}$$

则线性方程组 (2.1) 可以写成矩阵方程

$$A \cdot X = b .$$

而 A 确定的 K-线性映射 $\varphi_A : K^n \to K^m$ 就是

$$A : M_{n\times 1}(K) \to M_{m\times 1}(K),\ X \mapsto A\cdot X$$

即对每个 $X \in M_{n\times 1}(K)$, 它的像是用 A 乘 X 的结果.

习题2.2

1. 计算: (a) $\begin{pmatrix} 3 & 0 & 2 & 0 \\ 0 & 1 & 2 & 1 \\ 2 & 3 & 0 & 0 \end{pmatrix} \cdot \begin{pmatrix} 1 & -2 & 2 \\ 2 & -1 & 1 \\ -1 & 1 & -2 \\ 2 & 2 & -1 \end{pmatrix} + \begin{pmatrix} -2 & 0 & -3 \\ 0 & 6 & -3 \\ 5 & -2 & 8 \end{pmatrix}$,

(b) $\begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}^2$, $\begin{pmatrix} \cos\alpha & \sin\alpha \\ -\sin\alpha & \cos\alpha \end{pmatrix}^n$, $\begin{pmatrix} 2 & 1 \\ 5 & 3 \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} \cdot \begin{pmatrix} 3 & -1 \\ -5 & 2 \end{pmatrix}^n$.

2. 设 $\varphi_A : K^n \to K^m$ 是由 $A \in M_n(K)$ 确定的 K-线性映射. 证明:

(1) φ_A 是恒等映射 $\Leftrightarrow A = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix} = I_n$ (称为 n 阶单位矩阵);

(2) φ_A 是双射 $\Leftrightarrow$ 存在唯一 $B \in M_n(K)$ 使 $A\cdot B = B\cdot A = I_n$;

(3) φ_A 的像 $\varphi_A(k^n) = \left\{x_1A^{(1)} + x_2A^{(2)} + \cdots + x_nA^{(n)} \mid \forall x_i \in K\right\}$.

3. 下题中的 K^n ($n = 1, 2, \cdots$) 表示标准 K-向量空间. 试证明:

(1) 如果 $f : K \to K$ 是一个 K-线性映射, 则存在常数 $a \in K$ 使对任意 $x \in K, f(x) = ax$ (称 f 由常数 $a \in K$ 确定).

(2) 如果 $f : K \to K^n$ 是一个 K-线性映射, 则存在常数 $b_1, b_2, \cdots, b_n \in K$ 使对任意 $x \in K,\ f(x) = \begin{pmatrix} b_1x \\ b_2x \\ \vdots \\ b_nx \end{pmatrix} = x\begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix}$.

(3) 如果 $f : K^n \to K$ 是一个 K-线性映射, 则存在常数 $a_1, a_2, \cdots, a_n \in K$ 使

对任意 $x=\begin{pmatrix}x_1\\ \vdots\\ x_n\end{pmatrix}\in K^n$, $g(x)=a_1x_1+a_2x_2+\cdots+a_nx_n$.

(4) 如果 $g\circ f: K\to K$ 由常数 $c\in K$ 确定, 则

$$c=a_1b_1+\cdots+a_nb_n=(a_1,a_2,\cdots,a_n)\begin{pmatrix}b_1\\ b_2\\ \vdots\\ b_n\end{pmatrix}.$$

4. 将线性方程组

$$\begin{cases}x_1+x_2-2x_3+2x_4=1\\ 3x_1+5x_2+6x_3-4x_4=4\\ 4x_1+5x_2-2x_3+3x_4=3\\ 3x_1+8x_2+24x_3-19x_4=2\end{cases}$$

写出矩阵形式, 并指出它的系数矩阵和增广矩阵.

2.3 矩阵的秩

设 K 是一个域 (比如, $K=\mathbb{Q},\mathbb{R},\mathbb{C}$ 或 $\mathbb{C}$ 的子域). $V\subset K^n$ 表示 K^n 的子空间.

定义 2.3.1 非空子集 $V\subset K^n$ 称为 K^n 的子空间, 如果 $\forall\alpha,\beta\in V$. $\lambda\in K$ 有 $\alpha+\beta\in V$ 和 $\lambda\alpha\in V$.

定义 2.3.2 设 $\alpha_1,\alpha_2,\cdots,\alpha_m\in K^n$ 是一组向量, 则 $\lambda_1\alpha_1+\lambda_2\alpha_2+\cdots+\lambda_m\alpha_m$ 称为线性组合 (其中 $\lambda_1,\lambda_2,\cdots,\lambda_m$ 称为该线性组合的系数). $\alpha_1,\alpha_2,\cdots,\alpha_m$ 称为线性相关. 如果存在不全为零的 $\lambda_1,\lambda_2,\cdots,\lambda_m\in K$ 使 $\lambda_1\alpha_1+\lambda_2\alpha_2+\cdots+\lambda_m\alpha_m=0$. 否则 $\alpha_1,\alpha_2,\cdots,\alpha_m$ 称为线性无关. (i.e. 如果 $\lambda_1\alpha_1+\lambda_2\alpha_2+\cdots+\lambda_m\alpha_m=0$, 则 $\lambda_1=\lambda_2=\cdots=\lambda_m=0$).

引理 2.3.1 设 $\alpha_1,\alpha_2,\cdots,\alpha_m\in K^n$ 是一组向量, 则

$$\langle\alpha_1,\alpha_2,\cdots,\alpha_m\rangle=\{\lambda_1\alpha_1+\lambda_2\alpha_2+\cdots+\lambda_m\alpha_m\mid\forall\lambda_i\in K\}$$

是 K^n 中包含 $\alpha_1,\alpha_2,\cdots,\alpha_m$ 的最小子空间.

证明: 首先 $\langle\alpha_1,\alpha_2,\cdots,\alpha_m\rangle$ 是包含 $\alpha_1,\alpha_2,\cdots,\alpha_m$ 的子空间: 显然 $\langle\alpha_1,\alpha_2,\cdots,\alpha_m\rangle$ 包含 $\alpha_1,\alpha_2,\cdots,\alpha_m$. 另外, $\forall\alpha,\beta\in\langle\alpha_1,\alpha_2,\cdots,\alpha_m\rangle$ 则

$$\alpha=\lambda_1\alpha_1+\lambda_2\alpha_2+\cdots+\lambda_m\alpha_m,\beta=\mu_1\alpha_1+\mu_2\alpha_2+\cdots+\mu_m\alpha_m.$$

所以,

$$\alpha+\beta=(\mu_1+\lambda_1)\alpha_1+(\mu_2+\lambda_2)\alpha_2+\cdots+(\mu_m+\lambda_m)\alpha_m,$$

$$\lambda\alpha=\lambda\lambda_1\alpha_1+\lambda\lambda_2\alpha_2+\cdots+\lambda\lambda_m\alpha_m\in\langle\alpha_1,\alpha_2,\cdots,\alpha_m\rangle.$$

根据子空间的定义, $\langle\alpha_1,\alpha_2,\cdots,\alpha_m\rangle\subset K^n$ 是 K^n 的子空间.

其次 $\langle\alpha_1,\alpha_2,\cdots,\alpha_m\rangle$ 是包含 $\alpha_1,\alpha_2,\cdots,\alpha_m$ 的最小子空间; 如果 $W\subset K^n$ 是一个包含 $\alpha_1,\alpha_2,\cdots,\alpha_m$ 的最小子空间, 则 $\lambda_1\alpha_1+\lambda_2\alpha_2+\cdots+\lambda_m\alpha_m\in W$ (对任意 $\lambda_1,\lambda_2,\cdots,\lambda_m$), 所以 $\langle\alpha_1,\alpha_2,\cdots,\alpha_m\rangle\in W$. □

定义 2.3.3 $\langle\alpha_1,\alpha_2,\cdots,\alpha_m\rangle$ 称为由 $\alpha_1,\alpha_2,\cdots,\alpha_m$. 生成的子空间. 一个子空间 $V\subset K^n$ 称为有限生成的子空间, 如果存在有限个 $\alpha_1,\alpha_2,\cdots,\alpha_m\in V$ 使 $V=\langle\alpha_1,\alpha_2,\cdots,\alpha_m\rangle$.

定义 2.3.4 设 $\{\alpha_1,\alpha_2,\cdots,\alpha_m\}$ 是一组向量, $\alpha_{i_1},\alpha_{i_2},\cdots,\alpha_{i_r}\in\{\alpha_1,\alpha_2,\cdots,\alpha_m\}$ 称为 $\alpha_1,\alpha_2,\cdots,\alpha_m$ 的一个极大线性无关组. 如果它满足如下条件:

(1) $\alpha_{i_1},\alpha_{i_2},\cdots,\alpha_{i_r}$ 线性无关.

(2) $\forall\alpha\in\{\alpha_1,\alpha_2,\cdots,\alpha_m\}$ 存在 $\lambda_{i_1},\lambda_{i_2},\cdots,\lambda_{i_r}\in K$ 使 $\alpha=\lambda_{i_1}\alpha_{i_1}+\lambda_{i_2}\alpha_{i_2}+\cdots+\lambda_{i_r}\alpha_{i_r}$.

引理 2.3.2 如果 $\alpha_1,\alpha_2,\cdots,\alpha_m$ 不全为零, 则存在极大线性无关组 $\alpha_{i_1},\alpha_{i_2},\cdots,\alpha_{i_r}$.

证明: 令 $S=\{\alpha_1,\alpha_2,\cdots,\alpha_m\}$. 一个子集合 $T\subset S$ 称为线性无关子集. 如果 T 中的向量线性无关. 令 $T=\{\alpha_{i_1},\alpha_{i_2},\cdots,\alpha_{i_r}\}$ 是 $|T|$ 最大的线性无关子集, 可以证明: $T=\{\alpha_{i_1},\alpha_{i_2},\cdots,\alpha_{i_r}\}$ 是一个极大线性无关组. 由 T 的选取, $\alpha_{i_1},\alpha_{i_2},\cdots,\alpha_{i_r}$ 线性无关.

另一方面, 任取 $\alpha\in S,\alpha_1,\alpha_2,\cdots,\alpha_m,\alpha$ 一定是线性相关 (如果, $\alpha\in T=\{\alpha_{i_1},\alpha_{i_2},\cdots,\alpha_{i_r}\},\alpha\in S,\alpha_1,\alpha_2,\cdots,\alpha_m,\alpha$ 显然线性相关, 如果 $\alpha\notin T$, 令

$T' = \{\alpha_{i_1}, \alpha_{i_2}, \cdots, \alpha_{i_r}, \alpha\}$. 则 $|T'| > |T|$, 所以 T' 中的向量线性相关), 所以存在不全为零的 $\lambda_1, \cdots, \lambda_r, \lambda_{r+1} \in K$ 使

$$\lambda_1 \alpha_{i_1} + \cdots + \lambda_r \alpha_{i_r} + \lambda_{r+1} \alpha = 0$$

如果 $\lambda_{r+1} = 0$, 则 $\lambda_1 \alpha_{i_1} + \cdots + \lambda_r \alpha_{i_r} = 0$. 且 $\lambda_1, \cdots, \lambda_r$ 不全为零, 这与 $\alpha_{i_1}, \alpha_{i_2}, \cdots, \alpha_{i_r}$ 线性无关矛盾! 因此, $\lambda_{r+1} \neq 0, \alpha = -\frac{\lambda_1}{\lambda_{r+1}} \alpha_{i_1} - \cdots - \frac{\lambda_r}{\lambda_{r+1}} \alpha_{i_r}$(i.e. α 可由 $\alpha_{i_1}, \alpha_{i_2}, \cdots, \alpha_{i_r}$ 线性表出). 所以 $\alpha_{i_1}, \alpha_{i_2}, \cdots, \alpha_{i_r}$ 是一个极大线性无关组. □

下面重要的引理是定义向量维数的基石, 它的证明本质上应用了高斯消元法.

引理 2.3.3 设 $\beta_1, \cdots, \beta_s, \alpha_1, \cdots, \alpha_r \in K^n$, 且每个 β_i 可由 $\alpha_1, \cdots, \alpha_r$ 线性表出, 则

(1) 如果 $s > r$, $\beta_1, \cdots, \beta_s$ 一定线性相关.

(2) 如果 $\beta_1, \cdots, \beta_s$ 线性无关, 则 $s \leqslant r$.

证明: (1) 和 (2)显然是等价的, 所以, 只需要证明 (1): 根据条件,

$$\begin{cases} \beta_1 = a_{11}\alpha_1 + a_{12}\alpha_2 + \cdots + a_{1r}\alpha_r \\ \beta_2 = a_{21}\alpha_1 + a_{22}\alpha_2 + \cdots + a_{2r}\alpha_r \\ \vdots \\ \beta_s = a_{s1}\alpha_1 + a_{s2}\alpha_2 + \cdots + a_{sr}\alpha_r \end{cases}, \quad a_{ij} \in K$$

考虑线性组合 $x_1\beta_1 + x_2\beta_2 + \cdots + x_s\beta_s = (a_{11}x_1 + a_{21}x_2 + \cdots + a_{s1}x_s)\alpha_1 + (a_{12}x_1 + a_{22}x_2 + \cdots + a_{s2}x_s)\alpha_2 + \cdots + (a_{1r}x_1 + \cdots + a_{sr}x_s)\alpha_r$

由于 $s > r$, 由高斯消元法, 下述齐次线性方程组

$$\begin{cases} a_{11}x_1 + a_{21}x_2 + \cdots + a_{s1}x_s = 0 \\ a_{12}x_1 + a_{22}x_2 + \cdots + a_{s2}x_s = 0 \\ \vdots \\ a_{1r}x_1 + a_{2r}x_2 + \cdots + a_{sr}x_s = 0 \end{cases}$$

有非零解 $(\lambda_1, \lambda_2, \cdots, \lambda_s)$, 所以有不全为零的 $\lambda_1, \lambda_2, \cdots, \lambda_s$. 使

$$\lambda_1\beta_1 + \lambda_2\beta_2 + \cdots + \lambda_s\beta_s = 0$$

即 $\beta_1, \beta_2, \cdots, \beta_s$ 线性相关. □

定义 2.3.5 设 $V \subset K^n$ 是一个子空间. $\alpha_1, \cdots, \alpha_m \in V$ 称为 V 的一组基, 如果它满足条件:

(1) $\alpha_1,\cdots,\alpha_m\in V$ 线性无关.

(2) V 中每一个向量可由 $\alpha_1,\cdots,\alpha_m\in V$ 线性表出. (i.e. $\forall\beta\in V$ 存在 $\lambda_1,\cdots,\lambda_m\in K$ 使 $\beta=\lambda_1\alpha_1+\cdots+\lambda_m\alpha_m$).

定理 2.3.1 设 $V\subset K^n$ 是一个非零子空间. 则

(1) 在 V 中存在一组基.

(2) 如果 $\alpha_1,\cdots,\alpha_r$ 和 $\beta_1,\cdots,\beta_s$ 分别是 V 的一组基, 则 $r=s$.

(3) 如果 $\alpha_1,\cdots,\alpha_t\in V$ 是一组线性无关的向量, 则 $\alpha_1,\cdots,\alpha_t$ 可以扩充为 V 的一组基.

证明: 反复利用到引理2.7.

(1) 只需证明 V 是有限生成的 (i.e. 存在 $\alpha_1,\cdots,\alpha_m\in V$ 使 $V=\langle\alpha_1,\cdots,\alpha_m\rangle$), 则 $\{\alpha_1,\cdots,\alpha_m\}$ 的一个极大线性无关组 $\alpha_{i_1},\cdots,\alpha_{i_r}$ 就是 V 的一组基. 由引理2.7, 可以证明 V 使有限生成的. 事实上, 设 $e_1,e_2,\cdots,e_n\in K^n$, 其中 $e_i=(0,\cdots,0,1,0,\cdots,0)$, 则 K^n 中每个向量可由 $e_1,e_2,\cdots,e_n$ 线性表出. 所以, V 中任一组线性无关向量 $\alpha_1,\cdots,\alpha_m$ 可由 $e_1,e_2,\cdots,e_n$ 线性表出, 由引理2.7, $m\leqslant n$. 所以, 存在一组个数最多的线性无关向量 $\alpha_1,\cdots,\alpha_m$. 不难证明 $V=\langle\alpha_1,\cdots,\alpha_m\rangle$, 事实上, $\forall\beta\in V$, $\alpha_1,\cdots,\alpha_m,\beta$ 线性相关. 从而 β 可由 $\alpha_1,\alpha_2,\cdots,\alpha_m$ 线性表出.

(2) 因为 $\alpha_1,\cdots,\alpha_r$ 和 $\beta_1,\cdots,\beta_s$ 分别是 V 的一组基, 所以它们可以相互线性表出, 由引理2.7, $r=s$.

(3) 设 m 表示 V 中任一组的向量个数. $\alpha_1,\cdots,\alpha_t\in V$ 线性无关, 如果 $\langle\alpha_1,\cdots,\alpha_t\rangle=V$, 则 $\alpha_1,\cdots,\alpha_t$ 已经是一组基. 否则, 存在 $\alpha_{t+1}\in V\setminus\langle\alpha_1,\cdots,\alpha_t\rangle$, 则 α_{t+1} 不能由 $\alpha_1,\cdots,\alpha_t$ 线性表出. 所以 $\alpha_1,\cdots,\alpha_t,\alpha_{t+1}$ 线性无关. 如果 $V=\langle\alpha_1,\cdots,\alpha_t,\alpha_{t+1}\rangle$, 则 $\alpha_1,\cdots,\alpha_t,\alpha_{t+1}$ 使 V 的一组基. 否则可继续添加 $\alpha_{t+2}\in V\setminus\langle\alpha_1,\cdots,\alpha_t,\alpha_{t+1}\rangle$, 得到线性无关组 $\alpha_1,\cdots,\alpha_t,\alpha_{t+1},\alpha_{t+2}$. 有限步后, $\alpha_1,\cdots,\alpha_t$ 可扩充为 V 的一组基. □

定义 2.3.6 设 $V\subset K^n$ 是一个非零子空间, $\alpha_1,\cdots,\alpha_m\in V$ 是 V 的一组基, 则 m 称为 V 的维数. 记为

$$\dim_K(V)=m$$

如果 $V=0$, 则 $\dim_K(V)=0$.

例 2.3.1 $e_1, \cdots, e_n \in K^n$ 是 K^n 的一组基 (称为标准基). 所以

$$\dim_K(K^n) = n\,.$$

例 2.3.2 $M_{m\times n}(K) = \left\{A = \left(a_{ij}\right)_{m\times n} \mid a_{ij} \in K\right\}$ 关于矩阵加法和数乘法是一个 K-向量空间. 令 $e_{ij} \in M_{m\times n}(K)$ 表示矩阵

$$e_{ij} = \begin{pmatrix} 0 & \cdots & 0 & 0 & 0 & \cdots & 0 \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ 0 & \cdots & 0 & 0 & 0 & \cdots & 0 \\ 0 & \cdots & 0 & 1 & 0 & \cdots & 0 \\ 0 & \cdots & 0 & 0 & 0 & \cdots & 0 \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ 0 & \cdots & 0 & 0 & 0 & \cdots & 0 \end{pmatrix}$$

(即矩阵 e_{ij} 除了第 i 行第 j 行交叉处的元是 1 , 其他位置均为0). 则

$$A = \left(a_{ij}\right)_{m\times n} = \sum_{i,j} a_{ij}e_{ij}\,,$$

且 $\left\{e_{ij}\right\}_{1\leqslant i\leqslant m, 1\leqslant j\leqslant n}$ 线性无关. $\dim_K M_{m\times n}(K) = n$.

例 2.3.3 $\mathbb{C}$ 可以看成 $\mathbb{R}$-向量空间. $1, i \in \mathbb{C}$ 是一组基. $\dim_{\mathbb{R}}(\mathbb{C}) = 2$.

设 $A = \left(a_{ij}\right)_{m\times n}$ 是一个矩阵, 用 $A_{(1)}, \cdots, A_{(m)}$ 分别表示 A 的第1, 第2,$\cdots$, 第 m 个行向量. 而

$$A^{(j)} = \begin{pmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{mj} \end{pmatrix} = \left[a_{1j}, a_{2j}, \cdots, a_{mj}\right]$$

表示 A 的第 j 个列向量. 所以, 有时我们也将矩阵 A 写成

$$A = [A_{(1)}, A_{(2)}, \cdots, A_{(m)}]\,,\ A = \left(A^{(1)}, A^{(2)}, \cdots, A^{(n)}\right).$$

定义 2.3.7 令 $\langle A_{(1)}, A_{(2)}, \cdots, A_{(m)}\rangle \subset K^n$, $\left\langle A^{(1)}, A^{(2)}, \cdots, A^{(n)}\right\rangle \subset K^m$ 分别是 A 的 "行向量" 和 "列向量" 生成的子空间. 而

$$r_r(A) = \dim_K \langle A_{(1)}, A_{(2)}, \cdots, A_{(m)}\rangle$$

$$r_c(A) = \dim_K \left\langle A^{(1)}, A^{(2)}, \cdots, A^{(n)}\right\rangle$$

分别称为 A 的行秩和列秩.

例 2.3.4 设 $\varphi_A : K^n \to K^m$ 是由矩阵 $A = \left(a_{ij}\right)_{m\times n}$ 确定的 K-线性映射, 则 φ_A 的像

$$\varphi_A\left(K^n\right) = \left\{x_1A^{(1)} + x_2A^{(2)} + \cdots + x_nA^{(n)} \mid \forall x_i \in K\right\} = \left\langle A^{(1)}, \cdots, A^{(n)}\right\rangle$$

所以 $\dim_K \varphi_A\left(K^n\right) = r_c(A)$. 而

$$\ker\left(\varphi_A\right) := \{x = [x_1, \cdots, x_n] \in K^n \mid \varphi_A(x) = 0\} \in K^n$$

是齐次线性方程组 $x_1A^{(1)} + x_2A^{(2)} + \cdots + x_nA^{(n)} = 0$ 的解空间.

定理 2.3.2 $r_r(A) = r_c(A)$ (所以, $r(A) = r_r(A) = r_c(A)$ 称为 A 的秩).

证明: 对 A 的行向量做第 I, II 类初等变换可得矩阵

$$\overline{A} = \begin{pmatrix}
0 & \cdots & 0 & a_{1i_1} & \cdots & a_{1i_2} & \cdots & a_{1i_3} & \cdots & a_{1i_r} & \cdots & a_{1n} \\
0 & \cdots & 0 & 0 & \cdots & a_{2i_2} & \cdots & a_{2i_3} & \cdots & a_{2i_r} & \cdots & a_{2n} \\
0 & \cdots & 0 & 0 & \cdots & 0 & \cdots & a_{3_3} & \cdots & a_{3i_r} & \cdots & a_{3n} \\
\vdots & & \vdots & \vdots & & \vdots & & \vdots & & \vdots & & \vdots \\
0 & \cdots & 0 & 0 & \cdots & 0 & \cdots & 0 & \cdots & a_{ri_r} & \cdots & a_{rn} \\
0 & \cdots & 0 & 0 & \cdots & 0 & \cdots & 0 & \cdots & 0 & \cdots & 0 \\
\vdots & & \vdots & \vdots & & \vdots & & \vdots & & \vdots & & \vdots \\
0 & \cdots & 0 & 0 & \cdots & 0 & \cdots & 0 & \cdots & 0 & \cdots & 0
\end{pmatrix}$$

$$= \left[\bar{A}_{(1)}, \bar{A}_{(2)}, \cdots, \bar{A}_{(m)}\right] = \left(\bar{A}^{(1)}, \bar{A}^{(2)}, \cdots, \bar{A}^{(n)}\right), \ (a_{1i_1} \cdot a_{2i_2} \cdots a_{ri_r} \neq 0)$$

首先证明: $r_r(A) = r_r\left(\overline{A}\right)$, $r_c(A) = r_c\left(\overline{A}\right)$, 其中$r_r(A) = r_r\left(\overline{A}\right)$ 是显然的, 因为 $\langle A_{(1)}, A_{(2)}, \cdots, A_{(m)}\rangle = \left\langle \bar{A}_{(1)}, \bar{A}_{(2)}, \cdots, \bar{A}_{(m)}\right\rangle$, 所以只需证明: $r_c(A) = r_c\left(\overline{A}\right)$. 注意到:

$$x_1A^{(1)} + x_2A^{(2)} + \cdots + x_nA^{(n)} = 0 \Leftrightarrow x_1\bar{A}^{(1)} + x_2\bar{A}^{(2)} + \cdots + x_n\bar{A}^{(n)} = 0 .$$

设 $r_c(A) = s, A^{(i_1)}, A^{(i_2)}, \cdots, A^{(i_s)}$ 是 $\left\{A^{(1)}, A^{(2)}, \cdots, A^{(n)}\right\}$ 的极大线性无关组, 则 $x_{i_1}\bar{A}^{(i_1)} + x_{i_2}\bar{A}^{(i_2)} + \cdots + x_{i_s}\bar{A}^{(i_s)} = 0$ 只有零解 (否则 $x_{i_1}A^{(i_1)} + x_{i_2}A^{(i_2)} + \cdots + x_{i_s}A^{(i_s)} = 0$ 有非零解. 这与 $A^{(i_1)}, A^{(i_2)}, \cdots, A^{(i_s)}$ 线性无关矛盾), 所以

$$\bar{A}^{(i_1)}, \bar{A}^{(i_2)}, \cdots, \bar{A}^{(i_s)}.$$

线性无关. $\forall A^{(j)} \in \left\{A^{(1)}, A^{(2)}, \cdots, A^{(n)}\right\}$, $A^{(i_1)}, A^{(i_2)}, \cdots, A^{(j)}$ 必然线性相关, 即 $x_{i_1}A^{(i_1)} + \cdots + x_{i_s}A^{(i_s)} + x_jA^{(j)} = 0$ 有非零解. 从而 $x_{i_1}\bar{A}^{(i_1)} + \cdots + x_{i_s}\bar{A}^{(i_s)} + x_j\bar{A}^{(j)} = 0$

有非零解. 即 $\left\{\bar{A}^{(i_1)},\cdots,\bar{A}^{(i_s)},\bar{A}^{(j)}\right\}$ 线性相关, 所以 $\bar{A}^{(i_1)},\cdots,\bar{A}^{(i_s)}$ 是 $\left\{\bar{A}^{(1)},\cdots,\bar{A}^{(n)}\right\}$ 的极大线性无关组. 所以, $r_c\left(\bar{A}\right)=s=r_c(A)$.

最后我们证明: $r_r\left(\bar{A}\right)=r_c\left(\bar{A}\right)$. 由于 $\bar{A}_{(1)},\bar{A}_{(2)},\cdots,\bar{A}_{(r)}$ 线性无关, 所以 $r_r\left(\bar{A}\right)=r$. 下面我们证明: $\bar{A}^{(i_1)},\cdots,\bar{A}^{(i_r)}$ 是 $\left\{\bar{A}^{(1)},\bar{A}^{(2)},\cdots,\bar{A}^{(n)}\right\}$ 的极大线性无关组, 从而 $r_c\left(\bar{A}\right)=r$.

事实上, 如果

$$\lambda_1\bar{A}^{(i_1)}+\lambda_2\bar{A}^{(i_2)}+\cdots+\lambda_r\bar{A}^{(i_r)}=\lambda_1\begin{pmatrix}a_{1i_1}\\0\\0\\\vdots\\0\\0\end{pmatrix}+\lambda_2\begin{pmatrix}a_{1i_2}\\a_{2i_2}\\0\\0\\\vdots\\0\end{pmatrix}+\cdots+\lambda_r\begin{pmatrix}a_{1i_r}\\a_{2i_r}\\a_{3i_r}\\\vdots\\a_{ri_r}\\0\\\vdots\\0\end{pmatrix}=0$$

则 $\lambda_r a_{ri_r}=0\left(a_{ri_r}\neq 0\right)\Rightarrow\lambda_r=0$.从而

$$\lambda_1\begin{pmatrix}a_{1i_1}\\0\\0\\\vdots\\0\\0\end{pmatrix}+\lambda_2\begin{pmatrix}a_{1i_2}\\a_{2i_2}\\0\\\vdots\\0\\0\end{pmatrix}+\cdots+\lambda_{r-1}\begin{pmatrix}a_{1i_{r-1}}\\a_{2i_{r-1}}\\\vdots\\a_{r-1,i_{r-1}}\\0\\\vdots\\0\end{pmatrix}=0\Rightarrow\lambda_{r-1}a_{r-1,i_{r-1}}=0\,(a_{r-1,i_{r-1}}\neq 0)\Rightarrow\lambda_{r-1}=0$$

类推可得 $\lambda_1=\lambda_2=\cdots=\lambda_{r-1}=\lambda_r=0$. 所以 $\bar{A}^{(i_1)},\bar{A}^{(i_2)},\cdots,\bar{A}^{(i_r)}$ 线性无关. 另一方面, 对任意 $\bar{A}^{(j)}\in\left\{\bar{A}^{(1)},\cdots,\bar{A}^{(n)}\right\}$, 向量组 $\left\{\bar{A}^{(i_1)},\cdots,\bar{A}^{(i_r)},\bar{A}^{(j)}\right\}$ 线性相关. 因此, $\bar{A}^{(i_1)},\cdots,\bar{A}^{(i_r)}$ 是 $\bar{A}^{(1)},\cdots,\bar{A}^{(n)}$ 的极大线性无关组. □

定理 2.3.3 线性方程组 $A\cdot X=b$ 有解的充要条件是

$$r(A)=r\left(\widetilde{A}\right).$$

其中 $\widetilde{A}=(A,b)$ 是增广矩阵. 如果 $r_r(A)=r$, 则

$$V=\{X\in K^n \mid A\cdot X=0\}=\langle \eta_1,\cdots,\eta_{n-r}\rangle$$

$\eta_1,\eta_2,\cdots,\eta_{n-r}$ 线性无关. 如果又是 $A\cdot X=b$ 的一组解. 则

$$\widetilde{V}=\{X\in K^n \mid A\cdot X=b\}=\{\alpha+\lambda_1\eta_1+\cdots+\lambda_{n-r}\eta_{n-r} \mid \forall \lambda_i\in K\}.$$

证明: $A\cdot X=b$ 有解 $\Leftrightarrow$ 存在 $x_1,x_2,\cdots,x_n\in K$ 使

$$x_1A^{(1)}+x_2A^{(2)}+\cdots+x_nA^{(1)}=b$$

$\Leftrightarrow \left\{A^{(1)},A^{(2)},\cdots,A^{(n)}\right\}$ 的极大线性无关组也是 $\left\{A^{(1)},A^{(2)},\cdots,A^{(n)},b\right\}$ 的极大线性无关组 $\Leftrightarrow r_c(A)=r_c(\tilde{A})\Leftrightarrow r(A)=r(\tilde{A})$.

如果 $r_r(A)=r$, 则通过对矩阵 A 的行向量做初等变换, A 可化为阶梯型. 为了化简符号, 无妨设该阶梯型为

$$A=\begin{pmatrix} 1 & \overline{a_{12}} & \overline{a_{13}} & \overline{a_{14}} & \cdots & \overline{a_{1r}} & \cdots & \overline{a_{1n}} \\ 0 & 1 & \overline{a_{23}} & \overline{a_{24}} & \cdots & \overline{a_{2r}} & \cdots & \overline{a_{2n}} \\ 0 & 0 & 1 & \overline{a_{34}} & \cdots & \overline{a_{3r}} & \cdots & \overline{a_{3n}} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & 0 & \cdots & 1 & \cdots & \overline{a_{rn}} \\ 0 & 0 & 0 & 0 & 0 & 0 & \cdots & 0 \end{pmatrix},$$

即方程组 $AX=0$ 可化为

$$\begin{cases} x_1+\overline{a_{12}}x_2+\overline{a_{13}}x_3+\cdots+\overline{a_{1r}}x_r+\cdots+\overline{a_{1n}}x_n=0 \\ x_2+\overline{a_{23}}x_3+\cdots+\overline{a_{2r}}x_r+\cdots+\overline{a_{2n}}x_n=0 \\ x_3+\cdots+\overline{a_{3r}}x_r+\cdots+\overline{a_{3n}}x_n=0 \\ \cdots\cdots\cdots\cdots\cdots\cdots \\ x_r+\overline{a_{r,r+1}}x_{r+1}+\cdots+\overline{a_{rn}}x_n=0 \end{cases}$$

解该方程组得

$$\begin{cases} x_1=c_{1,r+1}x_{r+1}+c_{1,r+2}x_{r+2}+\cdots+c_{1n}x_n=\sum_{i=1}^{n-r}c_{1,r+i}x_{r+i} \\ x_2=c_{2,r+1}x_{r+1}+c_{2,r+2}x_{r+2}+\cdots+c_{2n}x_n=\sum_{i=1}^{n-r}c_{2,r+i}x_{r+i} \\ \vdots \\ x_r=c_{r,r+1}x_{r+1}+c_{r,r+2}x_{r+2}+\cdots+c_{rn}x_n=\sum_{i=1}^{n-r}c_{r,r+i}x_{r+i} \end{cases}$$

其中 $x_{r+1},x_{r+2},\cdots,x_n$ 是自由变量. 所以, 令 $x_{r+1}=1,x_{r+2}=\cdots=x_n=0$, 可得

$AX=0$ 的一组解.

$$x_1=c_{1,r+1},x_2=c_{2,r+1},\cdots,x_r=c_{r,r+1},x_{r+1}=1,x_{r+2}=0,\cdots,x_n=0$$

即

$$\eta_1=[c_{1,r+1},c_{2,r+1},\cdots,c_{r,r+1},1,0,0,0,\cdots,0,0]$$

同理, 令 $x_{r+1}=0,x_{r+2}=1,x_{r+3}=\cdots=x_n$, 得 $AX=0$ 的一个解

$$\eta_2=[c_{1,r+2},c_{2,r+2},\cdots,c_{r,r+2},0,1,0,0,\cdots,0,0]$$

这样可得到 $AX=0$ 的 $n-r$ 个解

$$\eta_1=\begin{pmatrix} c_{1,r+1}\\ c_{2,r+1}\\ \vdots\\ c_{r,r+1}\\ 1\\ 0\\ 0\\ \vdots\\ 0\end{pmatrix},\eta_2=\begin{pmatrix} c_{1,r+2}\\ c_{2,r+2}\\ \vdots\\ c_{r,r+2}\\ 0\\ 1\\ 0\\ \vdots\\ 0\end{pmatrix},\cdots,\eta_{n-r}=\begin{pmatrix} c_{1,r+2}\\ c_{2,r+2}\\ \vdots\\ c_{r,r+2}\\ 0\\ 0\\ \vdots\\ 0\\ 1\end{pmatrix}.$$

它们是线性无关的, 且 $AX=0$ 的任一解 $X=[x_1,\cdots,x_r,x_{r+1},\cdots,x_n]$ 都可以写成

$$X=\begin{pmatrix} \sum_{i=1}^{n-r}c_{1,r+i}x_{r+i}\\ \sum_{i=1}^{n-r}c_{2,r+i}x_{r+i}\\ \vdots\\ \sum_{i=1}^{n-r}c_{r,r+i}x_{r+i}\\ x_{r+1}\\ \vdots\\ x_n\end{pmatrix}=x_{r+1}\eta_1+x_{r+2}\eta_2+\cdots+x_n\eta_{n-r}$$

所以 $\eta_1,\eta_2,\cdots,\eta_{n-r}\in V=\{X\mid AX=0\}$ 是一组基. 如果 $\alpha=[\alpha_1,\cdots,\alpha_n]\in K^n$ 是

$AX = b$ 的一固定解, 则

$$\widetilde{V} = \{X \in K^n \mid AX = b\} = \{\alpha + \lambda_1\eta_1 + \lambda_2\eta_2 + \cdots + \lambda_{n-r}\eta_{n-r} \mid \forall \lambda_i \in K\}.$$

□

定义 2.3.8 上述定理证明中出现的 $\eta_1, \eta_2, \cdots, \eta_{n-r}$ 称为 $AX = 0$ 的基础解系. $\alpha = [\alpha_1, \cdots, \alpha_n]$ 称为 $AX = b$ 的一个特解. 而形如 $\alpha + \lambda_1\eta_1 + \lambda_2\eta_2 + \cdots + \lambda_{n-r}\eta_{n-r}$ 的向量称为

$$AX = b$$

的一般解.

例 2.3.5 求线性方程组

$$(*) \quad \begin{cases} x_1 + x_2 - 2x_3 + 2x_4 = 1 \\ 3x_1 + 5x_2 + 6x_3 - 4x_4 = 1 \\ 4x_1 + 5x_2 - 2x_3 + 3x_4 = 3 \\ 3x_1 + 8x_2 + 24x_3 - 19x_4 = 2 \end{cases}$$

的一个特解和一般解.

解: 通过高斯消元法 (或称初等变换), 方程组 $(*)$ 可化为

$$\begin{cases} x_1 + x_2 - 2x_3 + 2x_4 = 1 \\ \quad x_2 + 6x_3 - 5x_4 = -1 \end{cases},$$

令 $x_3 = x_4 = 0$, 得方程组 $(*)$ 的一个特解

$$\alpha = [2, -1, 0, 0],$$

$(*)$ 对应都齐次方程组 $(*)'$ 可化为

$$\begin{cases} x_1 + x_2 - 2x_3 + 2x_4 = 0 \\ \quad x_2 + 6x_3 - 5x_4 = 0 \end{cases},$$

解该方程组得:

$$\begin{cases} x_1 = -8x_3 = 7x_4 \\ x_2 = -6x_3 + 5x_4 \end{cases},$$

分别令 $x_3 = 1, x_4 = 0$ 和 $x_3 = 0, x_4 = 1$ 得 $(*)'$ 的基础解系

$$\eta_1 = [8, -6, 1, 0], \eta_2 = [-7, 5, 0, 1],$$

$(*)$ 的一般解为

$$[2+8\lambda_1-7\lambda_2,-1-6\lambda_1+5\lambda_2,\lambda_1,\lambda_2].$$

习题2.3

1. 证明下面关于“线性相关”的结论:

(a) 如果 $\alpha_1,\cdots,\alpha_m$ 中的一部分向量线性相关, 则 $\alpha_1,\cdots,\alpha_m$ 线性相关.

(b) 任何一组向量 $\alpha_1,\cdots,\alpha_m$, 如果包含零向量, 则 $\alpha_1,\cdots,\alpha_m$ 线性相关.

(c) $\alpha_1,\cdots,\alpha_m$ 线性相关 $\Leftrightarrow$ 其中有一个向量是其余向量的线性组合.

(d) 如果 $\alpha_1,\cdots,\alpha_m$ 线性无关, 则

$$\alpha_1,\alpha_2,\cdots\alpha_m,\beta \text{ 线性相关} \Leftrightarrow \beta \text{ 可由} \alpha_1,\alpha_2,\cdots\alpha_m \text{ 线性相关.}$$

2. 证明: $\alpha_1=(1,1,1),\alpha_2=(1,1,2),\alpha_3=(1,2,3)$ 是 $\mathbb{R}^3$ 的一组基, 并求 $\alpha_1,\alpha_2,\alpha_3$ 使 $\beta=(6,9,14)=\lambda_1\alpha_1+\lambda_2\alpha_2+\lambda_3\alpha_3$.

3. 设 $\varphi_A: K^n\to K^m$ 是由矩阵 $A=\left(a_{ij}\right)_{m\times n}$ 确定的 K-线性映射. 证明:

(1) φ_A 的像 $\varphi_A(K^n)=\{\varphi_A(x)\mid\forall x\in K^n\}$ 是 K^m 的子空间且$\dim_K(\varphi_A(K^n))=r_c(A)$.

(2) φ_A 的核 $\ker(\varphi_A)=\{x\in K^n\mid\varphi_A(x)=0\}$ 是 K^n 的子空间且 $\dim_K(\ker(\varphi_A))=n-r_r(A)$.

4. 设 $f:K^n\to K^m$ 是任意 K-线性映射, 令 $\dim_K(\ker(f))=s,\dim_K(f(K^n))=t$. 证明: $\dim_K(K^n)=s+t$ (i.e. $n=s+t$).

(提示: 令 $\alpha_1,\cdots,\alpha_s\in\ker(f)$ 是一组基, $\beta_1,\cdots,\beta_t\in K^n$ 使 $f(\beta_1),\cdots,f(\beta_t)\in f(K^n)$ 是一组基. 证明: $\alpha_1,\cdots,\alpha_s,\beta_1,\cdots,\beta_t\in K^n$ 是一组基.)

5. 利用习题3和习题4证明: $r_r(A)=r_c(A)$.(行秩 = 列秩的另一证明)

6.求线性方程组

$$\begin{cases}2x_1+5x_2+x_3+3x_4=0\\4x_1+6x_2+3x_3+5x_4=0\\4x_1+14x_2+x_3+7x_4=0\\2x_1-3x_2+3x_3+x_4=0\end{cases}$$

的基础解系.

7. 设 U,V 是两个子空间, 如果 $U\subseteq V$, 则 $\dim_K(U)\leqslant\dim_K(V)$ 且 $\dim_K(U)=$

$\dim_K(V) \Leftrightarrow U = V$.

8. 设 U, V 是两个子空间, $f : U \to V$ 是 K-线性映射. 证明:

(1) f 是单射 $\Leftrightarrow \ker(f) = \{0\}$.

(2) f 是满射 $\Leftrightarrow \dim_K f(U) = \dim_K(V)$.

(3)(推广习题4). $\dim_K(U) = \dim_K \ker(f) + \dim_K f(U)$.

9. 设 $A \in M_{m\times s}(K), B \in M_{\mathrm{sxn}}(K)$. $\varphi_A : K^s \to K^m, \varphi_B : K^n \to K^s$ 是它们定义的 K-线性映射. $\varphi_{A\cdot B} = \varphi_A \circ \varphi_B : K^n \to K^m$. 证明:

(1) $\varphi_{AB}(K^n) \subseteq \varphi_A(K^s)$.

(2) $\ker(\varphi_B) \subseteq \ker(\varphi_{AB})$.

(3) $r(AB) \leqslant \min\{r(A), r(B)\}$.

10. $A, B, \varphi_A, \varphi_B$ 如上题. 设$U = \varphi_B(K^n)$, $V = \varphi_{AB}(K^n)$, $f = \varphi_A|_U$: $U \to K^m$表示 φ_A 在 $U \subseteq K^s$ 的限制映射 (i.e. $\forall x \in U, f(x) = \varphi_A(x)$). 证明:

(1) $f : U \to K^m$ 是 K-线性映射. 且 $f(U) = V$.

(2) $\ker(f) \subseteq \ker(\varphi_A)$.

(3) 利用习题8(3), 证明 $r(A) + r(B) - s \leqslant r(AB)$.

11. 设 U, V 是 K^m 的子空间, 证明:

(1) 集合 $V + U = \{\alpha + \beta \mid \alpha \in V, \beta \in U\}$ 也是 K^m 的子空间.

(2) $V \bigcap U$ 也是 K^m 的子空间.

(3) $\dim_K(V + U) \leqslant \dim_K(V) + \dim_K(U)$.

12. 设 $A, B \in M_{m\times n}(K)$, 证明: $r(A + B) \leqslant r(A) + r(B)$.

(提示: 设 $\varphi_A, \varphi_B : K^n \to K^m$, 证明 φ_{A+B} 的像在 $\varphi_A(K^n) + \varphi_B(K^n)$ 中 i.e. $Im(\varphi_{A+B}) \subset Im(\varphi_A) + Im(\varphi_B)$.)

2.4 矩阵

通过前面的讨论, 我们有这样一种感觉, 矩阵是表示线性方程组, K-线性映射的工具. 事实上, 矩阵可以看成线性映射的坐标, 矩阵的运算是由线性映射的运算定义的, 但矩阵的引入不仅简化了计算, 矩阵论本身也是一个内容丰富的研究方向.

命题 2.4.1 矩阵的乘法满足结合律和分配律:

(1)$(AB)C = A(BC)$,

(2)$(A+B)C=AC+BC, D(A+B)=DA+DB$.

(在上面等式中, 我们假设乘法都是有意义的).

证明: 由于线性映射的合成满足结合律, 所以矩阵的乘法自然也满足. 不难验证分配律:

$(\varphi_A+\varphi_B)\cdot\varphi_C(x)=(\varphi_A+\varphi_B)(\varphi_c(x))=\varphi_A(\varphi_c(x))+\varphi_B(\varphi_c(x))=\varphi_A\cdot\varphi_c(x)+\varphi_B\cdot\varphi_C(x)=(\varphi_A\varphi_C+\varphi_B\varphi_C)(x)$. □

定义 2.4.1 设$A=(a_{ij})_{m\times n}$, 则矩阵$B=(b_{ij})_{n\times m}$, 其中$b_{ij}=a_{ji}$, 称为A的转置矩阵(或简称A的转置), 记为${}^tA$: 即

$$A=\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}_{m\times n}, \text{则}\ {}^tA=\begin{pmatrix} a_{11} & a_{21} & \cdots & a_{m1} \\ a_{12} & a_{22} & \cdots & a_{m2} \\ \vdots & \vdots & \vdots & \vdots \\ a_{1n} & a_{2n} & \cdots & a_{nm} \end{pmatrix}_{n\times m}$$

命题 2.4.2 (1)${}^t({}^tA)=A, {}^t(A+B)={}^tA+{}^tB, {}^t(\lambda A)=\lambda{}^tA$,

(2)${}^t(AB)={}^tB{}^tA$.

证明: (1) 中的等式是显然的.

(2) 我们分别用行向量和列向量来表示A,B.

$$A=[A_{(1)},A_{(2)},\cdots,A_{(m)}], B=(B^{(1)},B^{(2)},\cdots,B^{(n)}).$$

这样方便计算矩阵的乘法. 事实上,

$$AB=(A_{(i)}\cdot B^{(j)})=\begin{pmatrix} A_{(1)}B^{(1)} & A_{(1)}B^{(2)} & \cdots & A_{(1)}B^{(n)} \\ A_{(2)}B^{(1)} & A_{(2)}B^{(2)} & \cdots & A_{(2)}B^{(n)} \\ \vdots & \vdots & \vdots & \vdots \\ A_{(m)}B^{(1)} & A_{(m)}B^{(2)} & \cdots & A_{(m)}B^{(n)} \end{pmatrix}$$

其中$A_{(i)}B^{(j)}=(a_{i1},a_{i2},\cdots,a_{is})\begin{pmatrix} b_{1j} \\ b_{2j} \\ \vdots \\ b_{sj} \end{pmatrix}=a_{i1}b_{1j}+a_{i2}b_{2j}+\cdots+a_{is}b_{sj}={}^t(B^{(j)})\cdot{}^t(A_{(i)})$. 所以${}^t(AB)_{i,j}=(AB)_{ji}=A_{(j)}B^{(i)}={}^t(B^{(i)})\cdot{}^t(A_{(j)})={}^tB_{(i)}\cdot{}^tA^{(j)}=({}^tB{}^tA)_{ij}$, 因此, ${}^t(AB)={}^tB{}^tA$ □

下面我们主要讨论行数与列数相等的矩阵(称为方阵):

$$M_n(K) = \{A = (a_{ij})_{n\times n} \mid a_{ij} \in K\}.$$

例 2.4.1 $I_n = (a_{ij})_{n\times n}$, 其中$a_{ij} = \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases}$称为$n$阶单位矩阵. A是单位矩阵$\Leftrightarrow$ $\varphi_A : K^n \to K^n$是恒等映射.

例 2.4.2 $diag(\lambda_1, \lambda_2, \cdots, \lambda_n) = (a_{ij})_{n\times n}$, 其中$a_{ij} = \begin{cases} \lambda_i & i = j \\ 0 & i \neq j \end{cases}$称为对角矩阵. 如果$\lambda_1 = \lambda_2 = \cdots = \lambda_n = \lambda$, 则它称为纯量矩阵, 此时显然有$diag(\lambda_1, \lambda_2, \cdots, \lambda_n) = \lambda \cdot I_m$.

例 2.4.3 $E_{ij} \in M_n(K)$表示在(i, j)处元素是1, 而其他处都为0的矩阵, 称为矩阵单位. 因为, 任一矩阵$A = (a_{ij})_{n\times n}$均可唯一表成$A = \sum_{i,j} a_{ij}E_{ij}$.

定理 2.4.1 设$A = (a_{ij})_{n\times n} \in M_n(K)$. 则

$$A = \lambda I_n \Leftrightarrow \forall B \in M_n(K), A \cdot B = B \cdot A.$$

证明: $A = (a_{ij})_{n\times n} = \sum_{i,j} a_{ij}E_{ij}$. 所以$AB = BA$对任意$B \in M_n(K)$成立$\Leftrightarrow$ $AE_{ij} = E_{ij}A, (1 \leq i \leq n, 1 \leq j \leq n)$.令$A = [A_{(1)}, A_{(2)}, \cdots, A_{(n)}] = (A^{(1)}, A^{(2)}, \cdots, A^{(n)})$. 则$AE_{ij} = (0, \cdots, 0, A^{(i)}, 0, \cdots, 0), E_{ij}A = [0, \cdots, 0, A_{(j)}, 0, \cdots, 0]$.其中$E_{ij} = (0, \cdots, 0, e_i, 0, \cdots, 0), e_i = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$, 从而有$AE_{ij} = E_{ij}A \Leftrightarrow a_{ij} = 0 (i \neq j), a_{ii} = a_{jj} \Leftrightarrow$ $A = a_{11} \cdot I_n$. □

下面我们引入一类重要矩阵, 称为初等矩阵, 为了讨论方便, 我们仍将单位矩阵I_n表示为$I_n = [e_1, e_2, \cdots, e_n] = ({}^te_1, {}^te_2, \cdots, {}^te_n)$, 其中$e_1 = (1, 0, \cdots, 0), e_2 = (0, 1, 0, \cdots, 0), \cdots, e_n = (0, 0, \cdots, 0, 1)$.

定义 2.4.2 (I) 交换I_n的第s行与第t行所得矩阵$F_{s,t}$称为第一类初等矩阵.*i.e.* $F_{s,t} = [e_1, \cdots, e_{s-1}, e_t, e_{s+1}, \cdots, e_{t-1}, e_s, e_{t+1}, \cdots, e_n] = I_n - E_{ss} - E_{tt} + E_{st} + E_{ts}$.
(II) 将I_n的第s行乘λ加到第t行所得矩阵$F_{s,t}(\lambda)$称为第二类初等矩阵. *i.e.* $F_{s,t}(\lambda) = [e_1, \cdots, e_{s-1}, e_s, e_{s+1}, \cdots, e_{t-1}, e_t + \lambda e_s, e_{t+1}, \cdots, e_n] = I_n + \lambda E_{t,s}$.
(III) 将I_n的第s行乘以不为零的常数λ所得矩阵$F_s(\lambda)$称为第三类初等矩阵. *i.e.* $F_s(\lambda) = [e_1, \cdots, e_{s-1}, \lambda e_s, e_{s+1}, \cdots, e_n] = I_n + (\lambda - 1)E_{ss}$.

命题 2.4.3 (1)${}^tF_{s,t} = F_{s,t}, {}^tF_{s,t}(\lambda) = F_{t,s}(\lambda), {}^tF_s(\lambda) = F_s(\lambda)$,
(2)设$A = (a_{ij})_{m\times n}, F_{s,t}, F_{s,t}(\lambda), F_s(\lambda)$是$m$阶初等矩阵.
则(I) $F_{s,t} \cdot A$是交换A的第s行与第t行所得矩阵.
(II) $F_{s,t}(\lambda) \cdot A$是将A的第s行乘λ加到第t行所得矩阵.
(III) $F_s(\lambda) \cdot A$是将A的第s 行乘以非零常数所得矩阵.
即:

$$F_{s,t} \cdot A = [A_1, \cdots, A_{s-1}, A_t, A_{s+1}, \cdots, A_{t-1}, A_s, A_{t+1}, \cdots, A_n].$$

$$F_{s,t}(\lambda) \cdot A = [A_1, \cdots, A_s, \cdots, A_{t-1}, A_t + \lambda A_s, A_{t+1}, \cdots, A_n].$$

$$F_s(\lambda) \cdot A = [A_1, \cdots, A_{s-1}, \lambda A_s, A_{s+1}, \cdots, A_n].$$

同理对列的初等变换也成立(右乘n阶初等矩阵):

$$A \cdot F_{s,t} = [A^{(1)}, \cdots, A^{(s-1)}, A^{(t)}, A^{(s+1)}, \cdots, A^{(t-1)}, A^{(s)}, A^{(t+1)}, \cdots, A^{(n)}].$$

$$A \cdot F_{s,t}(\lambda) = [A^{(1)}, \cdots, A^{(s)}, \cdots, A_{(t-1)}, A_{(t)} + \lambda A_{(s)}, A_{(t+1)}, \cdots, A_{(n)}].$$

$$A \cdot F_s(\lambda) = [A^{(1)}, \cdots, A^{(s-1)}, \lambda A^{(s)}, A^{(s+1)}, \cdots, A^{(n)}].$$

证明: 直接验证, 留作练习. □

定理 2.4.2 设$A = (a_{ij})_{m\times n}$的秩$r(A) = r$, 则存在m阶初等矩阵$P_1, \cdots, P_s$和n阶初等矩阵$Q_1, \cdots, Q_t$, 使得

$$P_sP_{s-1}\cdots P_1AQ_1Q_2\cdots Q_t = \begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix},$$

其中I_r是r阶单位矩阵, 另外三个0分别表示$r\times(n-r), (m-r)\times r$和$(m-r)\times(n-r)$阶零矩阵.

证明： 上述定理等价于说通过对A的行向量和列向量做初等变换可得

$$\overline{A}=\begin{pmatrix}1&0&0&\cdots&0&\cdots&0\\0&1&0&\cdots&0&\cdots&0\\\vdots&\vdots&\vdots&\vdots&\vdots&\vdots&\vdots\\0&0&0&\cdots&1&0\cdots&0\\0&0&0&0&0&0&0\\\vdots&\vdots&\vdots&\vdots&\vdots&\vdots&\vdots\\0&0&0&0&0&0&0\end{pmatrix}=\begin{pmatrix}I_r&0\\0&0\end{pmatrix}.$$

如果$A\neq 0$, 通过交换行和列的次序可设$a_{11}\neq 0$, 无妨设$a_{11}=1$(将A的第一行乘$\frac{1}{a_{11}}$). 通过对行和列做第(II)类初等变换, 可得$\begin{pmatrix}1&0&\cdots&0\\0&\overline{a_{22}}&\cdots&\overline{a_{2n}}\\\vdots&\vdots&\vdots&\vdots\\0&\overline{a_{m2}}&\cdots&\overline{a_{mn}}\end{pmatrix}$,

对$\begin{pmatrix}\overline{a_{22}}&\cdots&\overline{a_{2n}}\\\vdots&\vdots&\vdots\\\overline{a_{m2}}&\cdots&\overline{a_{mn}}\end{pmatrix}$继续, 可得结论. □

对于方阵$A\in M_n(K)$, 如果它对应的K-线性映射

$$\varphi_A:K^n\to K^n$$

(也称线性算子)是双射, 则其逆映射$\varphi_A^{-1}:K^n\to K^n$也是$K$-线性映射. 所以存在$B\in M_n(K)$使$\varphi_A^{-1}=\varphi_B$ 也是K-线性映射, 所以存在$B\in M_n(K)$ 使$\varphi_A^{-1}=\varphi_B$, 此时我们称φ_A是可逆线性映射(或可逆线性算子), 而A也称为可逆矩阵, B由A唯一确定(因为φ_A^{-1}由φ_A唯一确定), 所以B称为A的逆矩阵, 记为$B=A^{-1}$.事实上, 我们有如下纯矩阵语言的定义.

定义 2.4.3 方阵$A\in M_n(K)$称为可逆(或可逆矩阵). 如果存在$B\in M_n(K)$使$AB=BA=I_n$, 其中B是满足上述等式的唯一矩阵. 所以B称为A的逆(或A的逆矩阵)记为$B=A^{-1}$.

定理 2.4.3 对任意$A\in M_n(K)$, 令$\varphi_A:K^n\to K^n$是A对应的K-线性映射, 则下述论断等价.

(1)存在$B \in M_n(K)$使$AB = I_n$;

(2)$\varphi_A : K^n \to K^n$是满射;

(3)$r(A) = n$(此时称A非退化);

(4)$\varphi_A : K^n \to K^n$是单射;

(5)$\varphi_A : K^n \to K^n$是双射;

(6)A是可逆矩阵.

证明: 我们将证明:(1) $\Rightarrow$ (2) $\Rightarrow$ (3) $\Rightarrow$ (4) $\Rightarrow$ (5) $\Rightarrow$ (6) $\Rightarrow$ (1).

证明:(1) $\Rightarrow$ (2):令$\varphi_B : K^n \to K^n$是矩阵B对应的K-线性映射, 则$\varphi_A \cdot \varphi_B = \varphi_{I_n} = 1_{K^n}$(i.e. 恒等映射). 所以$\varphi_A$有右逆, 从而$\varphi_A$是满射.

(2) $\Rightarrow$ (3):由于$\varphi_A : K^n \to K^n$的像$Im(\varphi_A) = \langle A^{(1)}, A^{(2)}, \cdots, A^{(n)} \rangle$, 所以它的维数等于$r_c(A) = r(A)$. 但$\varphi_A$是满射推出

$$Im(\varphi_A) = K^n.$$

从而$dimIm(\varphi_A) = n$. 因此, $r(A) = n$.

(3) $\Rightarrow$ (4):由于$\varphi_A : K^n \to K^n$是单射的充要条件是$Ker(\varphi_A) = \{\alpha \in K^n \mid \varphi_A(\alpha) = 0\} = \{0\}$, 而$Ker(\varphi_A)$正好是线性方程组$AX = 0$的解空间. 由定理 **??**, $AX = 0$的解空间的维数等于$n - r(A)$. 所以$r(A) = n \Rightarrow dim_K Ker(\varphi_A) = 0 \Rightarrow Ker(\varphi_A) = \{0\} \Rightarrow \varphi_A$是单射.

(4) $\Rightarrow$ (5):φ_A是单射$\Rightarrow Ker(\varphi_A) = \{0\} \Rightarrow r(A) = n \Rightarrow r_c(A) = n \Rightarrow dim_K Im(\phi_A) = n$, 但$Im(\varphi_A) \subset K^n$, 所以$K^n = Im(\varphi_A)$, 从而$\varphi_A : K^n \to K^n$是满射, 因此$\varphi_A$是双射.

(5) $\Rightarrow$ (6):$\varphi_A : K^n \to K^n$是双射$\Rightarrow$存在逆映射$\varphi_A^{-1} : K^n \to K^n$, 不难证明映射$\varphi_A^{-1} : K^n \to K^n$是$K$-线性映射, 即$\forall \alpha, \beta \in K^n$和$\lambda \in K$, 有$\varphi_A^{-1}(\alpha + \beta) = \varphi_A^{-1}(\alpha) + \varphi_A^{-1}(\beta), \varphi_A^{-1}(\lambda\alpha) = \lambda\varphi_A^{-1}(\alpha)$, 而要证明上述等式, 只需证明它们在$\varphi_A$下的像相等(因为$\varphi_A$是单射):

$$\varphi_A(\varphi_A^{-1}(\alpha) + \varphi_A^{-1}(\beta)) = \varphi_A(\varphi_A^{-1}(\alpha)) + \varphi_A(\varphi_A^{-1}(\beta)) = \alpha + \beta = \varphi_A(\phi_A^{-1}(\alpha + \beta))$$

$$\varphi_A(\lambda\varphi_A^{-1}(\alpha)) = \lambda\varphi(\varphi_A^{-1}(\alpha)) = \lambda\alpha = \varphi_A(\varphi_A^{-1}(\lambda\alpha))$$

(上述第二个等号成立是因为φ_A是K-线性映射).由于$\varphi_A^{-1} : K^n \to K^n$是$K$-线性映射, 所以存在$B \in M_n(K)$使$\varphi_A^{-1} = \varphi_B$. 因此$\varphi_A \cdot \varphi_B = \varphi_B \cdot \varphi_A = \varphi_{I_n}$, 从而$AB = BA = I_n$. 所以$A$是可逆矩阵. □

命题 2.4.4 (1)初等矩阵都是可逆矩阵, 且它们的逆也是初等矩阵:

$F_{s,t}^{-1}=F_{s,t},F_{s,t}(\lambda)^{-1}=F_{s,t}(-\lambda),F_s(\lambda)^{-1}=F_s(\lambda^{-1})$.

(2)如果$A_1,A_2,\cdots,A_m$是可逆矩阵, 则它的乘积$A=A_1A_2\cdots A_m$也是可逆矩阵, 且$A^{-1}=A_m^{-1}A_{m-1}^{-1}\cdots A_2^{-1}A_1^{-1}$.

(3)任何可逆矩阵A都可写成初等矩阵的乘积.

(4)如果$B\in M_m(K),C\in M_n(K)$分别是m阶和n阶可逆矩阵, 则对任意$A\in M_{m\times n}(K)$, 有$r(BAC)=r(A)$.

证明: (1)它们之间的K-线性映射如下:

$$\varphi_{F_{s,t}}:K^n\to K^n;$$

$$x=\begin{pmatrix}x_1\\ \vdots\\ x_{s-1}\\ x_s\\ x_{s+1}\\ \vdots\\ x_{t-1}\\ x_t\\ x_{t+1}\\ \vdots\\ x_n\end{pmatrix}\xrightarrow{\varphi_{F_{s,t}}}\begin{pmatrix}x_1\\ \vdots\\ x_{s-1}\\ x_t\\ x_{s+1}\\ \vdots\\ x_{t-1}\\ x_s\\ x_{t+1}\\ \vdots\\ x_n\end{pmatrix}\text{(交换}x_s\text{与}x_t\text{的位置), 所以}F_{s,t}^{-1}=F_{s,t}.$$

$$x=\begin{pmatrix}x_1\\ \vdots\\ x_s\\ \vdots\\ x_t\\ \vdots\\ x_n\end{pmatrix}\xrightarrow{\varphi_{F_{s,t}(\lambda)}}\begin{pmatrix}x_1\\ \vdots\\ x_s\\ \vdots\\ x_t+\lambda x_s\\ \vdots\\ x_n\end{pmatrix},x=\begin{pmatrix}x_1\\ \vdots\\ x_s\\ \vdots\\ x_n\end{pmatrix}\xrightarrow{\varphi_{F_s(\lambda)}}\begin{pmatrix}x_1\\ \vdots\\ \lambda x_s\\ \vdots\\ x_n\end{pmatrix}.$$ 所以,$F_{s,t}(\lambda)^{-1}=F_{s,t}(-\lambda),F_s(\lambda)^{-1}=F_s(\lambda^{-1})$.

(2)$A_1,\cdots,A_m$可逆, 所以$A_1^{-1},\cdots,A_m^{-1}$存在, 令$B=A_m^{-1}A_{m-1}^{-1}\cdots A_2^{-1}A_1^{-1}$, 则$AB=$

$(A_1\cdots A_m)(A_m^{-1}A_{m-1}^{-1}\cdots A_2^{-1}A_1^{-1})=I_n$, 所以$A$可逆, 且$A^{-1}=B=A_m^{-1}A_{m-1}^{-1}\cdots A_2^{-1}A_1^{-1}$.

(3)如果A可逆, 则$r(A)=n$. 由定理 2.4.2, 存在初等矩阵$P_1,\cdots,P_s,Q_1,\cdots,Q_t$使

$$P_sP_{s-1}\cdot P_1AQ_1Q_2\cdot Q_t=I_n.$$

所以$A=P_1^{-1}P_2^{-1}\cdots P_{s-1}^{-1}P_s^{-1}Q_t^{-1}\cdots Q_2^{-1}Q_1^{-1}$.

(4)由于B,C可逆, 所以B,C分别可写成m阶和n阶初等矩阵的乘积:$B=P_sP_{s-1}\cdots P_1,C=Q_1Q_2\cdots Q_t$. 而$BAC=P_sP_{s-1}\cdots P_1AQ_1Q_2\cdots Q_t$表示分别对$A$的行向量和列向量作若干次初等变换, 所以$r(BAC)=r_r(BAC)=r_r(AC)=r_c(AC)=r_c(A)=r(A)$. □

最后我们介绍一种求逆矩阵的算法, 它的原理建立在下述事实上: 对任一可逆矩阵A, 存在初等矩阵$P_1,P_2,\cdots,P_s$使$P_sP_{s-1}\cdots P_2P_1A=I_n$, 而$P_sP_{s-1}\cdots P_2P_1A$相当于对$A$的行向量做$s$次初等变换$P_1,P_2,\cdots,P_s$所得矩阵. 所以, 令

$$(A\mid I_n)=\left(\begin{array}{cccc|ccccc} a_{11} & a_{12} & \cdots & a_{1n} & 1 & 0 & 0 & \cdots & 0\\ a_{21} & a_{22} & \cdots & a_{2n} & 0 & 1 & 0 & \cdots & 0\\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots\\ a_{n1} & a_{n2} & \cdots & a_{nn} & 0 & 0 & 0 & \vdots & 1 \end{array}\right)$$

, 对$(A|I_n)$的行向量依次做初等变换$(A|I_n)\xrightarrow{P_1}(P_1A|P_1\cdot I_n)\xrightarrow{P_2}(P_2P_1A|P_2P_1I_n)\rightarrow\cdots\xrightarrow{P_s}(P_s\cdots P_2P_1A|P_s\cdots P_2P_1\cdot I_n)$.直到$P_s\cdots P_2P_1A=I_n$停止, 此时在右边得到的矩阵$P_s\cdots P_2P_1$就是$A$的逆矩阵. *i.e.* $A^{-1}=P_s\cdots P_2P_1$.

例 2.4.4 如果$A=\begin{pmatrix} a_{11} & a_{12}\\ a_{21} & a_{22}\end{pmatrix}$可逆 (*i.e.*$r(A)=2$), 求$A^{-1}$.

解:A^{-1}可逆$\Leftrightarrow r(A)=2\Leftrightarrow\begin{pmatrix} a_{11}\\ a_{21}\end{pmatrix},\begin{pmatrix} a_{12}\\ a_{22}\end{pmatrix}$线性无关$\Rightarrow\begin{pmatrix} a_{11}\\ a_{21}\end{pmatrix}\neq\begin{pmatrix} 0\\ 0\end{pmatrix}$. 无妨设$a_{11}\neq 0\Rightarrow$ $|A|=a_{11}a_{22}-a_{12}a_{21}\neq 0$(否则, $a_{22}=\frac{a_{12}}{a_{11}}a_{21},a_{12}=\frac{a_{12}}{a_{11}}a_{11}$, *i.e.* $\begin{pmatrix} a_{12}\\ a_{22}\end{pmatrix}=\frac{a_{12}}{a_{11}}\begin{pmatrix} a_{11}\\ a_{22}\end{pmatrix}$, 矛盾).

$$(A\mid I_2)=\left(\begin{array}{cc|cc} a_{11} & a_{12} & 1 & 0\\ a_{21} & a_{22} & 0 & 1\end{array}\right)\xrightarrow{F_{12}(-\frac{a_{21}}{a_{11}})}\left(\begin{array}{cc|cc} a_{11} & a_{12} & 1 & 0\\ 0 & a_{22}-\frac{a_{12}a_{21}}{a_{11}} & -\frac{a_{21}}{a_{11}} & 1\end{array}\right)\xrightarrow{F_{21}(-\frac{a_{11}a_{12}}{|A|})}=$$

$$\left(\begin{array}{cc|cc} a_{11} & 0 & 1+\frac{a_{12}a_{21}}{|A|} & -\frac{a_{11}a_{12}}{|A|} \\ 0 & \frac{|A|}{a_{11}} & -\frac{a_{21}}{a_{11}} & 1 \end{array}\right) = \left(\begin{array}{cc|cc} a_{11} & 0 & \frac{a_{11}a_{22}}{|A|} & -\frac{a_{11}a_{12}}{|A|} \\ 0 & \frac{|A|}{a_{11}} & -\frac{a_{21}}{a_{11}} & 1 \end{array}\right) \xrightarrow{F_1(\frac{1}{a_{11}}),F_2(\frac{a_{11}}{|A|})}$$

$$\left(\begin{array}{cc|cc} 1 & 0 & \frac{a_{22}}{|A|} & -\frac{a_{12}}{|A|} \\ 0 & 1 & -\frac{a_{21}}{|A|} & \frac{a_{11}}{|A|} \end{array}\right).$$ 所以$A^{-1} = \frac{1}{|A|}\begin{pmatrix} a_{22} & -a_{12} \\ -a_{21} & a_{11} \end{pmatrix}$.

例 2.4.5 求矩阵$A = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 3 & 1 & 2 \\ 1 & 1 & 1 & -1 \\ 1 & 0 & -2 & -6 \end{pmatrix}$的逆矩阵$A^{-1}$.

解:令$(A|I_4) = \left(\begin{array}{cccc|cccc} 1 & 2 & 3 & 4 & 1 & 0 & 0 & 0 \\ 2 & 3 & 1 & 2 & 0 & 1 & 0 & 0 \\ 1 & 1 & 1 & -1 & 0 & 0 & 1 & 0 \\ 1 & 0 & -2 & -6 & 0 & 0 & 0 & 1 \end{array}\right) \xrightarrow{F_{12}(-2),F_{13}(-1),F_1(-1)}$

$$\left(\begin{array}{cccc|cccc} 1 & 2 & 3 & 4 & 1 & 0 & 0 & 0 \\ 0 & -1 & -5 & -6 & -2 & 1 & 0 & 0 \\ 0 & -1 & -2 & -5 & -1 & 0 & 1 & 0 \\ 0 & -2 & -5 & -10 & -1 & 0 & 0 & 1 \end{array}\right) \xrightarrow{F_{23}(-1),F_{24}(-2)} \left(\begin{array}{cccc|cccc} 1 & 2 & 3 & 4 & 1 & 0 & 0 & 0 \\ 0 & -1 & -5 & -6 & -2 & 1 & 0 & 0 \\ 0 & 0 & 3 & 1 & 1 & -1 & 1 & 0 \\ 0 & 0 & 5 & 2 & 3 & -2 & 0 & 1 \end{array}\right) \xrightarrow{F_{34}(-\frac{5}{3})}$$

$$\left(\begin{array}{cccc|cccc} 1 & 2 & 3 & 4 & 1 & 0 & 0 & 0 \\ 0 & -1 & -5 & -6 & -2 & 1 & 0 & 0 \\ 0 & 0 & 3 & 1 & 1 & -1 & 1 & 0 \\ 0 & 0 & 0 & \frac{1}{3} & \frac{4}{3} & -\frac{1}{3} & -\frac{5}{3} & 1 \end{array}\right) \xrightarrow{F_2(-1),F_3(\frac{1}{3}),F_4(3)}$$

$$\left(\begin{array}{cccc|cccc} 1 & 2 & 3 & 4 & 1 & 0 & 0 & 0 \\ 0 & 1 & 5 & 6 & 2 & -1 & 0 & 0 \\ 0 & 0 & 1 & \frac{1}{3} & \frac{1}{3} & -\frac{1}{3} & \frac{1}{3} & 0 \\ 0 & 0 & 0 & 1 & 4 & -1 & -5 & 3 \end{array}\right) \xrightarrow{F_{43}(-\frac{1}{3}),F_{42}(-6),F_{41}(-4)} \left(\begin{array}{cccc|cccc} 1 & 2 & 3 & 4 & -15 & 4 & 20 & -12 \\ 0 & 1 & 5 & 0 & -22 & 5 & 30 & -18 \\ 0 & 0 & 1 & 0 & -1 & 0 & 2 & -1 \\ 0 & 0 & 0 & 1 & 4 & -1 & -5 & 3 \end{array}\right) \xrightarrow{F_{32}(-5),F_{31}(-3)}$$

$$\left(\begin{array}{cccc|cccc} 1 & 2 & 0 & 0 & -12 & 4 & 14 & -9 \\ 0 & 1 & 0 & 0 & -17 & 5 & 20 & -13 \\ 0 & 0 & 1 & 0 & -1 & 0 & 2 & -1 \\ 0 & 0 & 0 & 1 & 4 & -1 & -5 & 3 \end{array}\right) \xrightarrow{F_{21}(-2)} \left(\begin{array}{cccc|cccc} 1 & 0 & 0 & 0 & 22 & -6 & -26 & 17 \\ 0 & 1 & 0 & 0 & -17 & 5 & 20 & -13 \\ 0 & 0 & 1 & 0 & -1 & 0 & 2 & -1 \\ 0 & 0 & 0 & 1 & 4 & -1 & -5 & 3 \end{array}\right).$$

所以$A^{-1} = \begin{pmatrix} 22 & -6 & 26 & 17 \\ -17 & 5 & 20 & -13 \\ -1 & 0 & 2 & -1 \\ 4 & -1 & -5 & 3 \end{pmatrix}$.

习题2.4

1.对于多项式$f(x) = a_m x^m + a_{m-1}x^{m-1} + \cdots + a_1 x + a_0 \in K[x]$和矩阵$A \in M_n(K)$, 定义:$f(A) = a_m A^m + a_{m-1}A^{m-1} + \cdots + a_1 A + a_0 I_n$. 试证明:

$$f(BAB^{-1}) = Bf(A)B^{-1}.$$

2.计算下列多项式$f(x) \in K[x]$在矩阵A的值$f(A)$.

(1)$f(x) = x^3 - 2x^2 + 1, A = \begin{pmatrix} 2 & 1 & 0 \\ 0 & 2 & 0 \\ 1 & 1 & 1 \end{pmatrix}$;

(2)$f(x) = x^3 - 3x + 2, A = \begin{pmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{pmatrix}$.

3.设$A \in M_n(K)$, 如果$E_{ii}A = AE_{ii}$对所有的$1 \le i \le n$成立, 证明:A必为对角矩阵$diag(a_{11}, a_{22}, \cdots, a_{nn})$.

4.对任意矩阵$A \in M_n(K)$, 令$tr(A)$表示A的对角线元素之和($i.e.$ $tr(A) = a_{11} + a_{22} + \cdots + a_{nn}$)(称为$A$的迹). 试证明:对任意可逆矩阵$B \in M_n(K)$, 有$tr(BAB^{-1}) = tr(A)$.

5.求解下列矩阵方程.

(1)$\begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \\ 3 & 1 & 2 \end{pmatrix} X = \begin{pmatrix} 0 & 0 \\ 1 & 0 \\ 0 & 1 \end{pmatrix}$; (2)$\begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 2 \\ 0 & 1 & 1 \end{pmatrix} X = \begin{pmatrix} 5 \\ -6 \\ -2 \end{pmatrix}$; (3)$\begin{pmatrix} 3 & 1 \\ 2 & 1 \end{pmatrix} X \begin{pmatrix} 1 & 3 \\ 1 & 2 \end{pmatrix} = \begin{pmatrix} 3 & 3 \\ 2 & 2 \end{pmatrix}$.

6.求下列矩阵A的逆矩阵, 并将A表示成初等矩阵的乘积.

$$(1)A=\begin{pmatrix}1&2&0&0\\2&3&0&0\\1&-1&1&3\\0&1&0&2\end{pmatrix};\ (2)A=\begin{pmatrix}1&1&\cdots&1\\0&1&\cdots&1\\\vdots&\vdots&\vdots&\vdots\\0&0&\cdots&1\end{pmatrix}_{n\times n}$$

7.证明:(1)如果$\lambda\neq 0$, A可逆, 则$(\lambda A)^{-1}=\lambda^{-1}A^{-1}$.

(2)如果A可逆, 则${}^tA$可逆, 且$({}^tA)^{-1}={}^t(A^{-1})$.

8.设$A\in M_2(K)$, $f(x)=x^2-tr(A)x+|A|\in K[x]$, 证明:$f(A)=0$(零矩阵).

9.设$A\in M_n(K)$, 如果${}^tA=A$, 则称A为对称矩阵, 而如果${}^tA=-A$, 则称A是反对称矩阵. 证明: 对称(反对称)矩阵的逆矩阵还是对称(反对称)矩阵.

10.矩阵$A\in M_n(K)$称为幂零矩阵, 如果存在m使$A^m=0$. 证明: 如果A,B是幂零矩阵, 且$AB=BA$, 则$A+B$也是幂零矩阵.

11.证明:

$$A=\begin{pmatrix}0&1&0&\cdots&0\\0&0&1&\cdots&0\\\vdots&\vdots&\vdots&\vdots&\vdots\\0&0&0&\cdots&1\\0&0&0&\cdots&0\end{pmatrix}_{n\times n}$$

是幂零矩阵, 并求最小的m使$A^m=0$.

12.如果$A\in M_n(K)$是幂零矩阵, 证明:I_n-A可逆.

13.如果$A\in M_n(K)$满足$A^2=A$(特征不为2)(称为幂等矩阵). 证明:I_n+A可逆.

14.对任意$A,B,C\in M_n(K)$, 如果$ABC=0$, 证明:$r(A)+r(B)+r(c)\leq 2n$.

15.设V,U是K^n的子空间, 证明:$V\cap U$和$V+U=\{\alpha+\beta\mid\forall\alpha\in V,\forall\beta\in U\}$也是$K^n$的子空间, 且$dim_K(V+U)=dim_K(V)+dim_K(U)-dim_K(V\cap U)$.

16.如果$V\cap U=\{0\}$, 则称$V+U$是一个直和, 记为$V\bigoplus U$. 证明:$V+U$是直和$\Leftrightarrow$ $V+U$中的每个向量都可唯一写成$\alpha+\beta$($i.e.$ 如果$\alpha+\beta=\alpha'+\beta'$, 其中$\alpha,\alpha'\in V,\beta,\beta'\in U$, 则$\alpha=\alpha',\beta=\beta'$)$\Leftrightarrow$ 如果$\alpha+\beta=0(\alpha\in V,\beta\in U)$, 则$\alpha=0,\beta=0$.($i.e.$ 0向量可唯一表成$0=0+0$).

参考文献