第一章 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 ϕ. 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，或写出刻划全部元素的条件. 例如，$S=\{1,2,\cdots,100\}$ 也可表成

$$S=\{x \mid x \text{ 是不超过 } 100 \text{ 的自然数}\}.$$

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

$\mathbb{R}=\{\text{实数}\}$ 表示所有实数的集合

$\mathbb{C}=\{\text{复数}\}$，表示所有复数的集合. 也可写成 $\mathbb{C}=\{a+bi \mid a,b\in\mathbb{R}\}$

定义1.1.2 两个集合称为相等，如果它们的元素都一样. 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 Φ 是任一集合的子集。所以，任意集合 S 都有平凡子集：Φ，S。S 的其它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。

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

$X\cap Y$ 表示由 X 与 Y 中相同元素组成的集合，称为 X 与 Y 的交。

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

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 $X\cap Y=\Phi$，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：$$\bigcup_{i\in I}X_i=\{x \mid \text{存在 } i\in I \text{ 使 } x\in X_i\}.$$

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

集合的差集与补集：设 X, Y 是集合，则集合

$$X \setminus Y = \{x \in X \mid x \notin Y\}$$ 称为差集集。

① 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 Y^c）。

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 $(x_1, x_2, \cdots x_n) = (y_1, \cdots y_n)$ 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

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

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映射 $f: X \to Y$，我们也可用记号：$f: X \to Y$，$x \longmapsto f(x)$ 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 $f = \sin: \mathbb{R} \to \mathbb{R}$. ($x \longmapsto f(x) = \sin x$).

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

$$f: B^n \longrightarrow B, \quad (x_1, \cdots, x_n) \longmapsto f(x_1, \cdots, x_n)$$

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

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

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 1.5：子集合 $f(X) = \{ f(x) \mid \forall x \in X \} \subseteq Y$ 称为映射 $f: X \to Y$ 的像（也记为 $\operatorname{Im} f$）。而对任意 $y \in Y$，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 1.5：考虑映射 $f: \mathbb{R}^n \to \mathbb{R}^m$

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

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

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

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

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

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？

~~则~~ 为了方便，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"：

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

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

则对于任意 $x=\begin{pmatrix}x_1\\x_2\\\vdots\\x_n\end{pmatrix}\in\mathbb{R}^n$，$f(x)=x_1A^{(1)}+x_2A^{(2)}+\cdots+x_nA^{(n)}$

（$x_1A^{(1)}+x_2A^{(2)}+\cdots+x_nA^{(n)}$ 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 $y=\begin{pmatrix}y_1\\\vdots\\y_m\end{pmatrix}\in\mathbb{R}^m$ 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：

$$\begin{cases}a_{11}x_1+a_{12}x_2+\cdots+a_{1n}x_n=y_1\\a_{21}x_1+a_{22}x_2+\cdots+a_{2n}x_n=y_2\\\qquad\vdots\\a_{m1}x_1+a_{m2}x_2+\cdots+a_{mn}x_n=y_m\end{cases}$$ 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1.6.6. 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 $\Longleftrightarrow \forall y \in Y$，$f^{-1}(y)$ 最多只有一个元素.

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

f 是双射 $\Longleftrightarrow \forall y \in Y$，$f^{-1}(y)$ 恰有一个元素.

对任意集合 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

$$g \cdot f: X \longrightarrow W, \quad x \longmapsto f(x) \longmapsto g(f(x)).$$

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. 所以映射的合成在 $F(X)$ 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 $g \cdot f: \mathbb{R} \to \mathbb{R}$，$x \longmapsto g(\sin x) = 3 + \sin^2 x$

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

可见，一般来说 $g \cdot f \neq f \cdot g$（既使它们都有定义）. 另外，我们中学还学过函数的乘法：$fg: \mathbb{R} \to \mathbb{R}$，$x \longmapsto f(x) \cdot g(x)$（即函数值相乘）. 显然，函数的乘法满足 $fg = gf$。注意：映射的乘积不是函数的乘积.

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 $\Longleftrightarrow$ 存在映射 $g: Y \to X$ 使 $g \cdot f = 1_X$

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

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

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

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

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

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

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，并称为 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的证明：(1) 首先证明：如果存在 $g: Y \to X$ 使 $g f = 1_X$ 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 $Y = \overline{f(X)} \cup f(X)$，且 $\overline{f(X)} \cap f(X) = \phi$。固定任一元素 $x_0 \in X$，定义 $g_1: \overline{f(X)} \longrightarrow X$，$y \longmapsto x_0$（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：

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

可得 $g \cdot f = 1_X$（显然，当 $f(X) \subsetneqq Y$ 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 $y \in Y$，子集 $f^{-1}(y) \subseteq X$ 是非空子集。所以，$\forall y \in Y$，任取 $x \in f^{-1}(y)$ 并固定，记为 $x = h(y)$。则 $h: Y \xrightarrow{} X$ 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 $|X| = n \Leftrightarrow$ 存在双射 $f: X \to \{1, 2, \cdots, n\}$。即有限集合 X 与 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 $f(n) = \begin{cases} n & \text{如果 } n < a \\ n+1 & \text{如果 } n \geq a \end{cases}$。

所以对任意可数集 X，$\forall x \in X$，其 $\{x\}$ 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）。

但而更令人惊奇的是 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 $f: \mathbb{R}^2 \longrightarrow \mathbb{R}$。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。他在信中还提及说，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"保持结构的双映射"）意义下的分类。例如，带有"拓扑"结构的集合称为拓扑空间，带有"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:

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

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

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

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

4. 设映射 $f:\mathbb{R}\to\mathbb{R}$ 由 $f(x)=x^2$ 定义. 试确定
(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 $g\circ f: X\to Z$ 也是单射.
(b) 如果 f, g 都是满射, 则 $g\circ f: X\to Z$ 也是满射.
(c) 如果 g, $g\circ f$ 都是单射, f 是单射吗? 证明你的答案.
(d) 如果 g, $g\circ f$ 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 X 的子集. 证明: $f(A\cup B) = f(A)\cup f(B)$, $f(A\cap B)\subseteq f(A)\cap f(B)$.

§1.2 数学归纳法原理

数学中有大量涉及自然数的命题 $P(n)$，它们通常可以采用<u>数学归纳法</u>证明。数学归纳法原理的最简单形式可以表述如下：

<u>数学归纳原理</u>：设 $P(n)$ 是一个关于整数 $n \geq k_0$ 的命题。如果：(1) $P(k_0)$ 正确，(2) $\forall m \geq k_0$，有

$$\text{“}P(m)\text{正确”} \Longrightarrow \text{“}P(m+1)\text{正确”}.$$

则 $P(n)$ 对任意 $n \geqslant k_0$ 正确.

证明：设 $F(P) = \{ n \geq k_0 \mid P(n)\text{不正确} \} \subset \mathbb{N}$

只需证明：如果条件 (1), (2) 成立，则 $F(P) = \phi$。

如果 <u>$F(P)$ 非空</u>，则存在<u>最小自然数</u> $m_0 \in F(P)$。

由条件 (1)，$k_0 \notin F(P)$，所以 $m_0 > k_0$，从而 $m_0 - 1 \geq k_0$

但 $m_0 - 1 \notin F(P)$（因为 m_0 是 $F(P)$ 中的最小自然数），所以

$P(m_0-1)$ 正确。由条件 (2)：$P(m_0-1)$ 正确 $\Rightarrow P(m_0)$ 正确.

从而 $m_0 \notin F(P)$。与 $m_0 \in F(P)$ 矛盾. □

<u>例 1.2.1</u>：令 $P(n)$ 表示命题：任意 n 元集合 S 有 2^n 个子集合。试用数学归纳证明 $P(n)$ 对任意 $n \geqslant 0$ 成立。

证明：(1) $P(0)$ 是正确：此时 S 是一个空集，它的子集合只

S自己，所以 S 有 $2^0=1$ 个子集合.

(2) 如果 $P(m)$ 正确，则 $P(m+1)$ 正确.

设 S 是一个 $m+1$ 元集合，$\forall a\in S$，令 $S=T\cup\{a\}$

其中 $T=S\setminus\{a\}$. 令 2^T 表示所有 T 中子集的集合，

则 S 中的任何子集 $A\subseteq S$ 要么不包含 a（此时 $A\in 2^T$），要么包含 a（此时 $A=\bar{A}\cup\{a\}$，$\bar{A}\in 2^T$）.

形如 $\bar{A}\cup\{a\}$，($\bar{A}\in 2^T$) 的子集合共 $|2^T|=2^m$ 个，所以 S 共有 $|2^T|+|2^T|=2\cdot 2^m=2^{m+1}$ 个子集. 因此，$P(m+1)$ 是正确的.

□

注意：要证明命题 $P(n)$ 正确，归纳原理中的条件 (1), (2) 缺一不可！

例 1.2.2：令 $P(n)$ 表示命题：对任意 $n\geq 1$，有 $2^n=0$.

这个命题 $P(n)$ 显然对任意 $n\geq 1$ 不正确！但是是满足归纳原理中的条件 (2)：如果 $P(m)$ 正确，则 $P(m+1)$ 正确.（$P(m)$ 正确 $\Rightarrow 2^m=0\Rightarrow 2^{m+1}=0\Rightarrow P(m+1)$ 正确）.

例 1.2.3：令 $P(n)$ 表示命题：对任意 $n\geq 1$，有 $2^n=2$.

这个命题 $P(n)$ 满足归纳原理中的条件 (1)：$P(1)$ 正确，但不满足条件 (2)（$P(1)$ 正确，但 $P(2)$ 不正确）.

有些命题的证明需要应用下述形式的归纳原理(称为完全归纳原理):

完全归纳原理: 设 $P(n)$ 是一个关于自然数 $n \geq k_0$ 的命题. 如果 $P(n)$ 满足: (1) $P(k_0)$ 正确, (2) $\forall\ m > k_0$,

~~假设 $P(k)$ 对任意 $k<m$ 正确, 则 $P(m)$ 正确.~~

$P(k)$ 对任意 $k<m$ 正确 $\Rightarrow$ $P(m)$ 正确.

则 $P(n)$ 对任意 $n \geq k_0$ 正确.

例 1.2.4: 自然数 $p>1$ 称为素数, 如果 p 不能写成两个大于1的自然数的乘积。令 $P(n)$ 表示命题"任意大于1的自然数 n 可以写成有限个素数的乘积".

证明: (1) $P(2)$ 是正确的(因为2是素数).

(2) $\forall\ m>2$, 设 $P(k)$ 对任意 $k<m$ 正确.

需要证明 $P(m)$ 正确。

如果 m 是素数, 显然 $P(m)$ 正确。否则, $m = m_1 \cdot m_2$, 其中 $m_1 < m$, $m_2 < m$。由归纳假设, $P(m_1)$, $P(m_2)$ 正确. 即 $m_1 = p_1 \cdots p_s$, $m_2 = q_1 \cdots q_t$ (p_i, q_j 是素数)。所以

$$m = p_1 \cdots p_s q_1 \cdots q_t. \quad q_i, p_j \text{ 是素数}.$$

即 $P(m)$ 正确. □

数学归纳原理在更广泛的意义下成立, 但基本原理本质相同, 需要在实际应用中逐步掌握.

习题 1.2

1. 证明完全归纳原理.

2. 证明数学归纳原理与完全归纳原理等价.

3. 设 $P(m,n)$ 是一个关于 $m\geq a$, $n\geq b$ 的命题. 如果

 (1) 命题 $P(m,b)$, $P(a,n)$ 对所有 $m\geq a$, $n\geq b$ 成立.

 (2) 命题 $P(k,n)$ 和 $P(m,l)$ 对所有 $k<m$ 和 $l<n$ 正确 $\Longrightarrow P(m,n)$ 正确.

 证明命题 $P(m,n)$ 对所有 $m\geq a$, $n\geq b$ 正确.

4. 设 $P(m,n)=\#\{(x_1,\cdots,x_n)\in\mathbb{N}^n \mid x_1+\cdots+x_n=m\}$. 证明

$$P(m,n)=C_{n+m-1}^{m}=\frac{(n+m-1)(n+m-2)\cdots(n+1)n}{m!}$$

(约定 $0!=1$). (提示: 利用下题恒等式)

5. 证明: 对任意 $m\geq 0$, $n\geq 2$, 有

$$\sum_{k=0}^{m} C_{n-2+k}^{k} = C_{n+m-1}^{m}.$$

(提示: 利用下题中的 pascal 恒等式)

6. 对任意正整数 n,k, 如果 $n\geq k$, 则

$$C_{n+1}^{k} = C_{n}^{k} + C_{n}^{k-1}.$$

(提示: 利用 $(1+x)^n=\sum_{k=0}^{n} C_n^k x^{n-k}$, $(1+x)^{n+1}=(1+x)\sum_{k=0}^{n} C_n^k x^{n-k}$).

§1.3 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"但毕竟真实存在，所以命名实数（it is real）。但人们在解方程的时候发现实数并不够用。负数的平方根首先出现在1545年。但当时的人们（包括很多大数学家）认为它是想象中的数（imaginary number），是虚无飘渺的。要消除虚数的虚无感，最好莫过于给出它及它们的运算的几何解释，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 $a+bi$ 的数称为复数（complex number），a, b 分别称为 $a+bi$ 的实部和虚部，bi 称为纯虚数。令

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

表示所有复数的全体，其中两个复数 $x+yi$ 与 $a+bi$ 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 $z=a+bi,\ w=c+di \in \mathbb{C}$，则定义

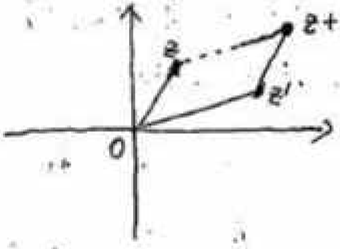
$$z+w=(a+c)+(b+d)i$$
$$z\cdot w=(ac-bd)+(bc+ad)i,$$

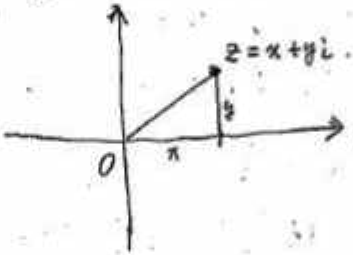
其中 a,b,c,d 的加法和乘法就是实数的加法和乘法

我们已经习惯将实数 a 与 $a+0i$ 等同，所以 $\mathbb{R}$ 可以看成 $\mathbb{C}$ 的子集合，且 $\mathbb{C}$ 上的运算在 $\mathbb{R}$ 上与实数的运算重合。

在平面Ω上，选定坐标系后，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 $z=x+yi$ 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) $|z|$.

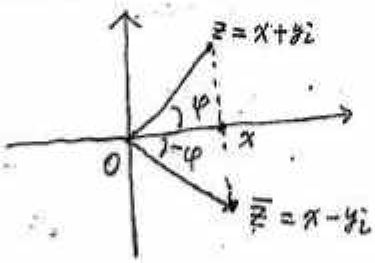




$$|z|=\sqrt{x^2+y^2}=|Oz|.$$

(II) 复数的辐角和复数的乘法.

对任意复数 $z=x+yi$，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，顺时针方向旋转所得角度为负值）. 如图.



其中 $\bar{z}=x-yi$ 称为 $z=x+yi$ 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述形式（三角形式）.

$$z=|z|(\cos\varphi+i\sin\varphi).$$

不难证明：

定理1.3.1：设 $z=|z|(\cos\varphi+i\sin\varphi)$，$z'=|z'|(\cos\varphi'+i\sin\varphi')$.

则 $$zz'=|z|\cdot|z'|(\cos(\varphi+\varphi')+i\sin(\varphi+\varphi'))$$

$$\frac{z}{z'}=\frac{|z|}{|z'|}(\cos(\varphi-\varphi')+i\sin(\varphi-\varphi'))$$ 如果 $|z'|\neq 0$.

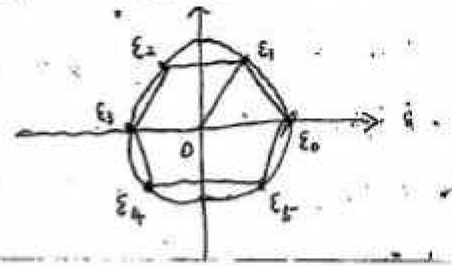
特别，我们有如下的棣莫弗公式

$$z^n = |z|^n(\cos n\varphi + i\sin n\varphi).$$ □

由棣莫弗公式，我们知道任意复数 z 都有 n 次方根：设 $x=|x|(\cos\theta+i\sin\theta)$ 是 $z=|z|(\cos\varphi+i\sin\varphi)$ 的一个 n 次方根，即 $x^n=z$。所以 $x^n=|x|^n(\cos n\theta+i\sin n\theta)=|z|(\cos\varphi+i\sin\varphi)$。从而 $|x|=\sqrt[n]{|z|}$，$n\theta=\varphi+2\pi k$。所以，如果 $z\neq0$ (i.e. $|z|\neq0$)，则 z 有 n 个 n 次方根…

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

特别，$\varepsilon_k=\cos\frac{2\pi k}{n}+i\sin\frac{2\pi k}{n}$，$(0\le k\le n-1)$ 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）



<u>定义1.3.1</u>：设 $R\subset\mathbb{C}$ 是一个至少包含一个非零数的子集合。如果 R 满足：

(1) $\forall\ a,b\in R$，必有 $a+b\in R$.

(2) $\forall\ a,b\in R$，必有 $a-b\in R$

(3) $\forall\ a,b\in R$，必有 $ab\in R$

（此时也称"R 关于 $\mathbb{C}$ 中的"加法"，"减法"，"乘法"是封闭"）。则称 R 是 $\mathbb{C}$ 的一个子环。如果一个子环 R 还满足：$\forall\ a,b\in R$, $a\neq0$ ⇒ $\frac{b}{a}\in R$.

则称子环 $R\subset\mathbb{C}$ 是 $\mathbb{C}$ 的一个子域.

$\square$

<u>例1.3.1</u>: $\mathbb{Q}$, $\mathbb{R}$ 都是 $\mathbb{C}$ 的子域, 分别称为有理数域、实数域. $\mathbb{Z}\subset\mathbb{C}$ 是 $\mathbb{C}$ 的子环, 但不是子域.

<u>例1.3.2</u>: $\mathbb{Z}[i]=\{a+bi\mid a,b\in\mathbb{Z}\}\subset\mathbb{C}$ 是 $\mathbb{C}$ 的子环 (称为高斯整数环). $\mathbb{Q}[i]=\{a+bi\mid a,b\in\mathbb{Q}\}\subset\mathbb{C}$ 是 $\mathbb{C}$ 的子域.

<u>例1.3.3</u>: 如果 $K\subset\mathbb{C}$, $L\subset\mathbb{C}$ 是两个子域. 则 $K\cap L\subset\mathbb{C}$ 也是一个子域.

<u>例1.3.4</u>: 设 $R\subset\mathbb{C}$ 是一个子环, $z\in\mathbb{C}$. 则

$$R[z]=\{a_0+a_1z+\cdots+a_nz^n\mid a_i\in R,\ n\ge 0\}\subset\mathbb{C}$$

是 $\mathbb{C}$ 的一个子环 (称为<u>向R添加z生成的子环</u>).

<u>例1.3.5</u>: 设 α 是 $x^2+bx+c=0$ 的根. (其中 $b,c\in\mathbb{Q}$, $b^2-4c<0$) 则子环 $\mathbb{Q}[\alpha]\subset\mathbb{C}$ 成为一个子域. 证明

$$\mathbb{Q}[\alpha]=\mathbb{Q}[\sqrt{\Delta}],\ \text{其中 } \Delta=b^2-4c.$$

习题 1.3

1. 计算下列表达式.

(a) $\dfrac{(1+3i)(8-i)}{(2+i)^2}$; (b) $\left(-\frac{1}{2}\pm\frac{\sqrt{3}}{2}i\right)^3$; (c) $\left(\dfrac{\sqrt{3}+i}{1-i}\right)^{30}$.

2. 解下述方程

(a) $|z|+z=8+4i$; (b) $z^2=3-4i$.

(c) 求 $(2+i)x+(1+2i)y=1-4i$ 的实数解.

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

3. 设 $R\subset\mathbb{C}$ 是一个子环, $z\in\mathbb{C}$, 证明: $R[z]$ 是 $\mathbb{C}$ 中包含 R 和 z 的最小子环.

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

证明: (1) $\varepsilon_k=\varepsilon_1^k$ $(0\le k<n)$.

(2) $\varepsilon_k\cdot\varepsilon_l=\begin{cases}\varepsilon_{k+l}, & \text{如果 } k+l<n\\\varepsilon_{k+l-n}, & \text{如果 } k+l\geq n\end{cases}$.

(3) 设 $a\in\mathbb{R}$, $a>0$, 则 $x^n-a=0$ 的全部根为

$\sqrt[n]{a}$, $\sqrt[n]{a}\varepsilon_1$, $\sqrt[n]{a}\varepsilon_1^2$, $\dots$, $\sqrt[n]{a}\varepsilon_1^{n-1}$.

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

6. 设 α 是 $x^2+bx+c=0$ 的一个根, $b,c\in\mathbb{R}$.

如果 $\Delta=b^2-4c<0$, 证明: $\mathbb{R}[\alpha]=\mathbb{C}$.

7. 求下述方程的全部解.

(1) $x^6-i=0$; (2) $x^6-64=0$, (3) $x^{10}-512(1-\sqrt{3}i)=0$.

§1.4 变换群

设 X 是一个非空集合，$S_X=\{X\xrightarrow{\sigma}X \mid \sigma\text{是双射}\}$ 是所有双射 $X\to X$ 的集合. 则映射的合成定义了 S_X 上的一个乘法： $\forall\ \sigma,\tau\in S_X$

$$\sigma\cdot\tau:\ X\xrightarrow{\tau}X\xrightarrow{\sigma}X.$$

该乘法满足：(1) $\forall\ \pi,\sigma,\tau\in S_X$，$(\pi\cdot\sigma)\cdot\tau=\pi\cdot(\sigma\cdot\tau)$

(2) 存在 $1\in S_X$ (X上的恒等映射)，$1\cdot\sigma=\sigma\cdot 1=\sigma$ $\forall\sigma\in S_X$

(3) $\forall\ \sigma\in S_X$，存在 $\tau\in S_X$ 使 $\sigma\cdot\tau=\tau\cdot\sigma=1$ ($\tau=\sigma^{-1}$ 是 σ 的逆映射.

<u>定义1.4.1</u>：设 $G\subset S_X$ 是非空集合. 如果 G 满足条件

(1) $\forall\ \sigma,\tau\in G$，有 $\sigma\cdot\tau\in G$，

(2) $\forall\ \sigma\in G$，有 $\sigma^{-1}\in G$.

则称 G 是一个<u>变换群</u>.

~~定义1.4.2~~. 如果 X 是有限集，$|X|=n$，则 $S_n:=S_X$ 称为<u>对称群</u>. □

变换群是一个极其重要的概念。在具体应用中，X 通常是带有某些"结构"的集合. 比如，X 是通常的<u>几何空间</u>，~~[illegible]~~ G 是 S_X 中保持两点距离不变的双射集合. 这样的 G 通常称为保距变换群. 是研究几何的重要工具。下面我们详细讨论对称群 S_n.

设 $X=\{1,2,\cdots,n\}$，则 S_n 中的元素

$$\sigma: X \longrightarrow X$$

由 $\sigma(1),\sigma(2),\cdots,\sigma(n)$ 唯一确定，所以我们通常称 S_n 中元素为一个置换（S_n 也称 n 元置换群），记为

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

引理1.4.1：$|S_n|=n!=1\cdot 2\cdot 3\cdots n$

证明：S_n 中的元素与 $1,2,\cdots,n$ 的所有"排列" 1-1 对应，而 $1,2,\cdots,n$ 共有 $n!$ 个不同的排列，所以 $|S_n|=n!$

□

引入记号：$\forall\ \sigma\in S_n$，$m\in\mathbb{Z}$，定义：

当 $m>0$，$\sigma^m=\underbrace{\sigma\cdot\sigma\cdots\sigma}_{m}$

当 $m<0$，$\sigma^m=(\sigma^{-1})^{-m}$

当 $m=0$，$\sigma^m=1$.

引理1.4.2：$\forall\ \sigma\in S_n$，$\sigma\neq 1$，则存在 $m>0$ 使

$$\sigma^m=1.$$

证明：考虑 S_n 中的元素：$\sigma,\sigma^2,\sigma^3,\cdots$，由 S_n 是有限集合，所以有 $m_1>0$，$m_2>0$ 且 $m_1\neq m_2$ 使 $\sigma^{m_1}=\sigma^{m_2}$，不妨设 $m_1>m_2$，则 $\sigma^{m_1-m_2}=\sigma^{m_2-m_2}=1$.

□

定义1.4.2：$\forall\ \sigma\in S_n$，$\sigma\neq 1$，令 $m>0$ 是使 $\sigma^m=1$ 的最小正整数，则称 m 是 σ 的阶。

例1.4.1（对换）：~~对~~ $\forall\ i,j\in X=\{1,2,\cdots,n\}$

定义双射 $\sigma=(ij):X\longrightarrow X$ 如下：

$\forall\ k\in X$，如果 $k\neq i,j$，则 $\sigma(k)=k$，

而 $\sigma(i)=j$，$\sigma(j)=i$，则 (ij) 的阶是 2.

例1.4.2（r-循环）：设 $i_1,i_2,\cdots,i_r\in X=\{1,2,\cdots,n\}$，

定义双射 $\sigma=(i_1i_2\cdots i_r):X\longrightarrow X$ 如下：

$\sigma(i_1)=i_2$，$\sigma(i_2)=i_3$，$\cdots$，$\sigma(i_{r-1})=i_r$，$\sigma(i_r)=i_1$，而

$\sigma(k)=k$，$\forall\ k\in X\setminus\{i_1,i_2,\cdots,i_r\}$.

则 $\sigma=(i_1i_2\cdots i_r)$ 的阶是 r；

例1.4.3：$\sigma=(1\,2)(3\,4)\in S_n$ 的阶是 2，但 σ 不是对换.

定义1.4.3：集合 $\{i_1,i_2,\cdots,i_r\}$ 称为 r-循环 $\sigma=(i_1i_2\cdots i_r)$ 的支撑集。$\sigma=(i_1i_2\cdots i_r)$，$\tau=(j_1j_2\cdots j_s)$ 称为不相交循环，如果 $\{i_1,i_2,\cdots,i_r\}\cap\{j_1,j_2,\cdots,j_s\}=\phi$.

引理1.4.3：如果 $\sigma=(i_1i_2\cdots i_r)$，$\tau=(j_1j_2\cdots j_s)$ 是不相交循环，则 $\sigma\cdot\tau=\tau\cdot\sigma$.

证明: $\forall\, k \in X = \{1, 2, \cdots, n\}$, 如果 $k \notin \{i_1, i_2, \cdots, i_r\}$

则 $\sigma(k) = k$, $\tau(k) \notin \{i_1, i_2, \cdots, i_r\}$ (否则, $\tau(k) \in \{i_1, \cdots, i_r\}$

从而 $\tau^2(k) = \tau(k)$ 得 $k = \tau(k) \in \{i_1, i_2, \cdots, i_r\}$ 矛盾).

从而 $\sigma\cdot\tau(k) = \tau(k) = \tau\cdot\sigma(k)$.

如果 $k \in \{i_1, i_2, \cdots, i_r\}$, 则 $\sigma(k) \in \{i_1, \cdots, i_r\}$. 所以

$$\sigma\cdot\tau(k) = \sigma(k) = \tau\cdot\sigma(k).$$

$\square$

定理1.4.1: S_n 中每个元素可以唯一地写成不相交循环的乘积.

证明: 对 n 用数学归纳法. 当 $n=1$ 时, $S_n = \{(1)\}$ 结论显然成立. ~~设对任意 $m<n$, S_m 中元素可以~~

设结论对任意 $m<n$ 成立.

设 $\sigma \in S_n$, $\sigma \neq 1$, 则存在 $i_1 \in X$ 使 $\sigma(i_1) \neq i_1$.

令 r 是 σ 的阶, $X_1 = X \setminus \{i_1, \sigma(i_1), \sigma^2(i_1), \cdots, \sigma^{r-1}(i_1)\}$.

~~由归纳假设, 存在不相交循环~~

则 $\forall\, k \in X_1$, $\sigma(k) \in X_1$. 所以, $\sigma|_{X_1}: X_1 \to X_1$ 是 S_{X_1} 中元素. 由归纳假设, 存在不相交循环 $\tau_1, \cdots, \tau_s \in S_{X_1}$ 使 $\sigma|_{X_1} = \tau_1 \cdots \tau_s$. 定义双射

$$\sigma_i: X \to X, \quad \sigma_i(k) = \begin{cases} \tau_i(k), & \text{如果 } k \in X_1 \\ k, & \text{否则} \end{cases}$$

则 $\sigma_1, \cdots, \sigma_s \in S_n$ 是不相交循环. 且

$$\sigma = \sigma_1 \cdots \sigma_s \cdot (i_1\ \sigma(i_1)\ \sigma^2(i_1) \cdots \sigma^{r-1}(i_1)).$$

是不相交循环的乘积. 下面证明分解唯一性:

如果 $\sigma = \sigma_1 \cdots \sigma_s = \pi_1 \cdots \pi_t$ 是两个这样的分解(假设 σ_i, π_i 非单位), 则 $s=t$, 且(适当排序后) $\sigma_i = \pi_i$. 只需证明: 存在 π_j 使 $\sigma_1 = \pi_j$.

设 $\sigma_1 = (i_1, \sigma_1(i_1), \cdots, \sigma_1^{r-1}(i_1))$, 由于 $\sigma(i_1) \neq i_1$, 故存在 π_j 使 $\pi_j(i_1) \neq i_1$. 所以, $\forall\, k>0$, 有

$$\sigma_1^k(i_1) = \sigma_1^k \cdot (\sigma_2^k \cdots \sigma_s^k(i_1)) = \sigma^k(i_1) = \pi_j^k(i_1)$$

从而 $\sigma_1 = \pi_j$.

$\square$

<u>推论1.4.1</u>: S_n 中每个元素都可写成<u>对换</u>的乘积

证明: $(i_1\ i_2 \cdots i_r) = (i_1\ i_2)\cdot(i_2\ i_3)\cdot \cdots \cdot(i_{r-1}\ i_r)$.

$\square$

这一推论直观上是容易见: $1, 2, \cdots, n$ 的任何一个排列 $j_1, j_2, \cdots, j_n$ 都可以通过<u>有限次</u>对换而得到. 下面的定理直观上不难接受, 但证明较繁琐, 所以省略.

<u>定理1.4.2</u>: $\forall\ \pi \in S_n$, 如果

$$\pi = \pi_1 \cdot \pi_2 \cdots \pi_s = \sigma_1 \cdots \sigma_t$$

是两个<u>对换分解</u>. 则 $s+t$ 是偶数 (s, t 的奇偶性相同).

定义1.4.4：$\forall\ \pi\in S_n$，如果 $\pi=\pi_1\cdot\pi_2\cdots\pi_s$

是一个对换分解（$\pi_1,\pi_2,\cdots,\pi_s$ 中可以有相同的对换），

则 $\varepsilon_\pi=(-1)^s$ 称为 π 的符号.

如果 $\varepsilon_\pi=1$，则称 π 为偶置换，否则称 π 为奇置换（即 $\varepsilon_\pi=-1$）.

推论1.4.2：(1) $\forall\ \pi,\sigma\in S_n$，有 $\varepsilon_{\pi\cdot\sigma}=\varepsilon_\pi\cdot\varepsilon_\sigma$.

(2) 如果 $\pi=\pi_1\cdot\pi_2\cdots\pi_s$ 是 π 的不相交循环分解.

则 $\varepsilon_\pi=(-1)^{\sum_{i=1}^{s}(l(\pi_i)-1)}$，

其中 $l(\pi_i)$ 是 π_i 的长度（即，π_i 是 $l(\pi_i)$-循环）.

证明：(1) $\varepsilon_{\pi\cdot\sigma}=\varepsilon_\pi\cdot\varepsilon_\sigma$ 由定义可得. □

(2) $\varepsilon_\pi=\varepsilon_{\pi_1}\cdots\varepsilon_{\pi_s}$. 而 $\varepsilon_{\pi_i}=(-1)^{l(\pi_i)-1}$ 可由

$$(i_1\,i_2\cdots i_r)=(i_1\,i_2)\cdot(i_2\,i_3)\cdots(i_{r-1}\,i_r)$$

得到. □

例1.4.4：将 $\pi=\begin{pmatrix}1&2&3&4&5&6&7\\5&4&1&7&3&6&2\end{pmatrix}\in S_7$ 分解成

不相交循环的乘积.

解：由于 $\pi(1)=5$，$\pi^2(1)=\pi(5)=3$，$\pi^3(1)=\pi(3)=1$. 所以

$\pi=(1\ 5\ 3)\cdot\pi_1$，其中 $\pi_1=\begin{pmatrix}1&2&3&4&5&6&7\\1&4&3&7&5&6&2\end{pmatrix}$.

又 $\pi_1(2)=4$，$\pi_1^2(2)=\pi_1(4)=7$，$\pi_1^3(2)=\pi_1(7)=2$ 所以

$$\pi_1=(2\ 4\ 7)\cdot\pi_2,\ \text{其中}\ \pi_2=\begin{pmatrix}1&2&3&4&5&6&7\\1&2&3&4&5&6&7\end{pmatrix}=1$$

因此，$\pi=(1\ 5\ 3)(2\ 4\ 7)$. □

定义1.4.5：设 V 是一个集合，函数

$$f:\ V\times\cdots\times V\longrightarrow\mathbb{R},$$
$$(x_1,x_2,\cdots,x_n)\longmapsto f(x_1,x_2,\cdots,x_n)$$

称为对称函数，如果 $\forall\pi\in S_n$，有

$$f(x_{\pi(1)},x_{\pi(2)},\cdots,x_{\pi(n)})=f(x_1,x_2,\cdots,x_n).$$

称为反对称函数，如果，$\forall\pi\in S_n$，有

$$f(x_{\pi(1)},x_{\pi(2)},\cdots,x_{\pi(n)})=\varepsilon_\pi f(x_1,x_2,\cdots,x_n).$$

□

例1.4.5：设 $f:V\times V\times\cdots\times V\to\mathbb{R}$ 是任意函数.

则 (1) $S(f):V\times V\times\cdots\times V\longrightarrow\mathbb{R}$

$$(x_1,x_2,\cdots,x_n)\longmapsto\sum_{\sigma\in S_n}f(x_{\sigma(1)},x_{\sigma(2)},\cdots,x_{\sigma(n)})$$

是对称函数

(2) $A(f):V\times V\times\cdots\times V\longrightarrow\mathbb{R}$,

$$(x_1,x_2,\cdots,x_n)\longmapsto\sum_{\sigma\in S_n}\varepsilon_\sigma f(x_{\sigma(1)},x_{\sigma(2)},\cdots,x_{\sigma(n)})$$

是反对称函数.

习题 1.4

1. 将 $\pi=\begin{pmatrix}1&2&3&4&5&6&7\\3&7&6&5&1&2&4\end{pmatrix}$ 分解成不相交循环的乘积，并求它的符号 ε_π.

2. 设
$$\sigma=\begin{pmatrix}1&2&3&4&5&6\\5&6&3&1&4&2\end{pmatrix},\quad \tau=\begin{pmatrix}1&2&3&4&5&6\\3&4&6&2&5&1\end{pmatrix}$$
是 S_6 中的元素。试计算：$\sigma\tau$，$\tau\sigma$，σ^{-1}，σ^2，σ^3，$\tau^{-1}\tau$.
并确定 σ 和 τ 的阶.

3. 将下列置换表成对换的乘积.
(1) $\sigma=\begin{pmatrix}1&2&3&4&5&6&7&8\\2&5&7&3&8&6&4&1\end{pmatrix}$，(2) $\sigma=\begin{pmatrix}1&2&3&4&5\\2&3&1&5&4\end{pmatrix}$

4. 证明：S_n 中任意置换都可写成形如
$$(1\,2),\ (1\,3),\ \cdots,\ (1\,n)$$
的对换的乘积.

5. 证明：任意偶置换都是形如 $(1\,2\,3), (1\,2\,4), \cdots, (1\,2\,n)$ 的 3-循环的乘积.

6. 设 $f(x_1,x_2,x_3,x_4,x_5)=x_1+2x_2^2+3x_3^3+4x_4^4+5x_5^5$，
$$\sigma=\begin{pmatrix}1&2&3&4&5\\2&3&1&5&4\end{pmatrix}\in S_5.$$
试求 $f(x_{\sigma(1)},x_{\sigma(2)},x_{\sigma(3)},x_{\sigma(4)},x_{\sigma(5)})$.

§1.5 多项式

设 $R\subset\mathbb{C}$ 是 $\mathbb{C}$ 中的子环，$x_1, x_2, \cdots, x_n$ 是不定元。形如 $a_{i_1i_2\cdots i_n}x_1^{i_1}x_2^{i_2}\cdots x_n^{i_n}$（$a_{i_1i_2\cdots i_n}\in R$）称为单项式，$a_{i_1i_2\cdots i_n}$ 称为它的系数，$i_1+i_2+\cdots+i_n$ 称为单项式

$$a_{i_1i_2\cdots i_n}x_1^{i_1}x_2^{i_2}\cdots x_n^{i_n},\quad (i_1, i_2, \cdots, i_n\in\mathbb{N})$$

的次数。两个单项式 $a_{i_1i_2\cdots i_n}x_1^{i_1}x_2^{i_2}\cdots x_n^{i_n}$ 与 $b_{j_1j_2\cdots j_n}x_1^{j_1}x_2^{j_2}\cdots x_n^{j_n}$ 称为同类项，如果

$$(i_1, i_2, \cdots, i_n)=(j_1, j_2, \cdots, j_n).$$

$$a_{i_1i_2\cdots i_n}x_1^{i_1}x_2^{i_2}\cdots x_n^{i_n}+b_{i_1i_2\cdots i_n}x_1^{i_1}x_2^{i_2}\cdots x_n^{i_n}=(a_{i_1i_2\cdots i_n}+b_{i_1i_2\cdots i_n})x_1^{i_1}x_2^{i_2}\cdots x_n^{i_n}$$

称为合并同类项。有限个单项式的形式和

$$f(x_1, x_2, \cdots, x_n)=\sum_{i_1i_2\cdots i_n}a_{i_1i_2\cdots i_n}x_1^{i_1}x_2^{i_2}\cdots x_n^{i_n},\quad a_{i_1i_2\cdots i_n}\in R$$

称为系数在 R 中的多项式。约定：多项式中的单项式"相加"与次序无关且满足结合律：

$$a_{i_1i_2\cdots i_n}x_1^{i_1}x_2^{i_2}\cdots x_n^{i_n}+a_{j_1j_2\cdots j_n}x_1^{j_1}x_2^{j_2}\cdots x_n^{j_n}$$
$$=a_{j_1j_2\cdots j_n}x_1^{j_1}x_2^{j_2}\cdots x_n^{j_n}+a_{i_1i_2\cdots i_n}x_1^{i_1}x_2^{i_2}\cdots x_n^{i_n} \qquad \text{（交换律）}$$

$$\left(a_{i_1i_2\cdots i_n}x_1^{i_1}x_2^{i_2}\cdots x_n^{i_n}+a_{j_1j_2\cdots j_n}x_1^{j_1}x_2^{j_2}\cdots x_n^{j_n}\right)+a_{k_1k_2\cdots k_n}x_1^{k_1}x_2^{k_2}\cdots x_n^{k_n}$$
$$=a_{i_1i_2\cdots i_n}x_1^{i_1}x_2^{i_2}\cdots x_n^{i_n}+\left(a_{j_1j_2\cdots j_n}x_1^{j_1}x_2^{j_2}\cdots x_n^{j_n}+a_{k_1k_2\cdots k_n}x_1^{k_1}x_2^{k_2}\cdots x_n^{k_n}\right)$$

（结合律）.

定义1.5.1：多项式 $f(x_1,x_2,\cdots,x_n)=\sum\limits_{i_1\cdots i_n} a_{i_1i_2\cdots i_n}x_1^{i_1}x_2^{i_2}\cdots x_n^{i_n}$ ($a_{i_1i_2\cdots i_n}\in R$) 称为零多项式 如果所有系数 $a_{i_1i_2\cdots i_n}=0$. 记为 0。如果

$$f(x_1,x_2,\cdots,x_n)=\sum_{i_1i_2\cdots i_n} a_{i_1i_2\cdots i_n}x_1^{i_1}x_2^{i_2}\cdots x_n^{i_n}$$

$$g(x_1,x_2,\cdots x_n)=\sum_{i_1i_2\cdots i_n} b_{i_1i_2\cdots i_n}x_1^{i_1}x_2^{i_2}\cdots x_n^{i_n}$$

是两个多项式，则

$$f(x_1,x_2,\cdots,x_n)+g(x_1,x_2,\cdots,x_n):=\sum_{i_1i_2\cdots i_n}(a_{i_1i_2\cdots i_n}+b_{i_1i_2\cdots i_n})x_1^{i_1}x_2^{i_2}\cdots x_n^{i_n}$$

$$f(x_1,x_2,\cdots,x_n)\,g(x_1,x_2,\cdots,x_n):=\sum_{\substack{i_1i_2\cdots i_n\\ j_1j_2\cdots j_n}} a_{i_1i_2\cdots i_n}b_{j_1j_2\cdots j_n}x_1^{i_1+j_1}x_2^{i_2+j_2}\cdots x_n^{i_n+j_n}$$

(右边表示经合并同类项后所得多项式). □

令 $R[x_1,x_2,\cdots,x_n]$ 表示所有系数在R中的多项式的集合，将R中的元素与 $R[x_1,x_2,\cdots,x_n]$ 中的零次多项式等同，即 $a x_1^0x_2^0\cdots x_n^0:=a$. 则 $R\subset R[x_1,\cdots,x_n]$.

对任意 $f(x_1,\cdots,x_n)=\sum\limits_{i_1\cdots i_n} a_{i_1\cdots i_n}x_1^{i_1}\cdots x_n^{i_n}\in R[x_1,\cdots,x_n]$ 的负多项式

$$-f(x_1,\cdots,x_n):=\sum_{i_1\cdots i_n}(-a_{i_1\cdots i_n})x_1^{i_1}\cdots x_n^{i_n},$$

$$f(x_1,\cdots,x_n)-g(x_1,\cdots,x_n):=f(x_1,\cdots,x_n)+(-g(x_1,\cdots,x_n))$$

$$1\cdot x_1^{i_1}\cdots x_n^{i_n}:=x_1^{i_1}\cdots x_n^{i_n}$$ 等等.

定义 1.5.2：多项式 $f(x_1,\cdots,x_n)\in R[x_1,\cdots,x_n]$ 称为 d 次齐次多项式，如果 $f(x_1,\cdots,x_n)$ 中的每个单项式都是 d 次的。

即：
$$f(x_1,\cdots,x_n)=\sum_{i_1+\cdots+i_n=d} a_{i_1\cdots i_n}\, x_1^{i_1}\cdots x_n^{i_n}$$
□

例 1.5.1：(1) $\overbrace{f(x_1,\cdots,x_n)}^{\in R[x_1\cdots x_n]}$ 是 1 次齐次多项式 $\Longleftrightarrow$
$$f(x_1,\cdots,x_n)=a_1x_1+\cdots+a_nx_n,\quad a_i\in R.$$

(2) $f(x_1,\cdots,x_n)\in R[x_1,\cdots,x_n]$ 是 2 次齐次多项式 $\Longleftrightarrow$
$$f(x_1,\cdots,x_n)=a_1x_1^2+\cdots+a_nx^n+\sum_{i<j}a_{ij}x_ix_j$$
□

设 $R[x_1,\cdots,x_n]_d\subset R[x_1,\cdots,x_n]$ 是 $R[x_1,\cdots,x_n]$ 中所有 d 次齐次多项式（包括零多项式）的集合，则 $\forall f(x_1,\cdots,x_n)\in R[x_1,\cdots,x_n]$，$f:=f(x_1,\cdots,x_n)$ 可唯一地写成 $f=f_0+f_1+\cdots+f_d$，$f_i\in R[x_1,\cdots,x_n]_i$

定义 1.5.3：$\forall f=f(x_1,\cdots,x_n)\in R[x_1,\cdots,x_n]$，如果 $f=f_0+f_1+\cdots+f_d$，其中 $f_d\neq0$，则定义 $\deg(f)=d$，f 称为 d 次多项式。□

例 1.5.2：如果 $\deg(f)=2$，则
$$f=a_1x_1^2+\cdots+a_nx_n^2+\sum_{i<j}a_{ij}x_ix_j+b_1x_1+\cdots+b_nx_n+c$$
□

引理1.5.1. $\forall\ f, g \in R[x_1, \cdots, x_n]$ 则

$$\deg(f\cdot g) = \deg(f) + \deg(g); \quad \deg(f+g) \le \max\{\deg(f), \deg(g)\}$$

证明：设 $f = f_0 + f_1 + \cdots + f_d$，$f_d \ne 0$，

$g = g_0 + g_1 + \cdots + g_{d'}$，$g_{d'} \ne 0$。则 $\deg(f+g) \le \max\{d, d'\}$。又

$$f\cdot g = f_0g_0 + (f_1g_0 + f_0g_1) + \cdots + f_dg_{d'}\,.$$

所以只需证明：两个非零齐次多项式 f_d，$g_{d'}$ 的乘积 $f_dg_{d'} \ne 0$。为此我们在所有 d次单项式中引入排序（字典排序法）：$\forall\ (i_1, i_2, \cdots, i_n), (j_1, \cdots, j_n) \in \mathbb{N}^n$

$$(i_1, \cdots, i_n) > (j_1, \cdots, j_n) \Longleftrightarrow (i_1, \cdots, i_n) - (j_1, \cdots, j_n) = (0, \cdots, 0, t, *, \cdots, *)$$

其中 $t > 0$。

令 $FT(f_d)$ 表示 f_d 的首项（即：在字典排序下，f_d 中"最大"的单项式），则 $FT(f_dg_{d'}) = FT(f_d)\cdot FT(g_{d'}) \ne 0$

所以 $f_dg_{d'} \ne 0$。 $\square$

当 $n=1$ 时，$R[x_1, \cdots, x_n]$ 中的多项式称为单变元多项式。用 $R[x]$ 表示所有单变元多项式（包括零多项式）的集合。$f(x) \in R[x]$ 通常记为

$$f(x) = a_0 + a_1x + \cdots + a_dx^d$$

或 $\quad f(x) = a_dx^d + \cdots + a_1x + a_0$. $\quad a_i \in R$.

如果 $a_d \ne 0$，则 $\deg(f) = d$，a_dx^d 称为 $f(x)$ 的首项。

引理 1.5.2（带余除法）：$\forall f(x), g(x) \in R[x]$，如果 $g(x)$ 的首项系数是 1，则存在唯一的 $q(x) \in R[x]$，$r(x) \in R[x]$ 使 $f(x) = q(x)g(x) + r(x)$，其中或者 $r(x)=0$，或者

$$\deg r(x) < \deg g(x).$$

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

$$g(x) = x^m + b_{m-1}x^{m-1} + \cdots + b_1 x + b_0$$

如果 $m>n$，则令 $q(x)=0$，$r(x)=f(x)$ 即得

$$f(x) = q(x)g(x) + r(x)$$

其中或者 $r(x)=0$，或者 $\deg r(x) = \deg f(x) < m = \deg g(x)$。

如果 $m \le n$，则可"消去 $f(x)$ 的首项"：

$$f_1(x) := f(x) - a_n x^{n-m} g(x) = f(x) - (a_n x^{n-m} x^m + a_n b_{m-1} x^{n-1} + \cdots + a_n b_0 x^{n-m})$$

$$= (a_{n-1} - a_n b_{m-1})x^{n-1} + \cdots + (a_{n-m} - a_n b_0)x^{n-m} + a_{n-m-1}x^{n-m-1} + \cdots + a_1 x + a_0.$$

$\deg(f_1) < \deg(f)$。对于 $\deg(f)$ 作归纳法，可设引理对 $f_1(x)$ 成立。即存在 $q_1(x), r_1(x)$ 使 $f_1(x) = q_1(x)g(x) + r_1(x)$，其中 $r_1(x)=0$ 或 $\deg r_1(x) < \deg g(x)$。由 $f_1(x) = f(x) - a_n x^{n-m} g(x)$ 得

$$f(x) = (a_n x^{n-m} + q_1(x))g(x) + r_1(x).$$

令 $q(x) := a_n x^{n-m} + q_1(x)$，$r(x) := r_1(x)$，即得满足引理条件的 $q(x), r(x)$。下面证明 $q(x), r(x)$ 的唯一性。

如果存在 $q(x), r(x)$ 及 $\bar{q}(x), \bar{r}(x)$ 满足：$r(x), \bar{r}(x)$ 或者是零多项式或者次数小于 $\deg(g)$，且 $f(x) = q(x)g(x) + r(x) = \bar{q}(x)g(x) + \bar{r}(x)$。

即 $(q(x) - \bar{q}(x))g(x) = \bar{r}(x) - r(x)$，$\deg[(q(x)-\bar{q}(x))g(x)] = \deg(\bar{r}(x) - r(x)) < \deg(g)$。

所以 $q(x) - \bar{q}(x) = 0$，从而 $\bar{r}(x) - r(x) = 0$。即 $q(x) = \bar{q}(x)$，$r(x) = \bar{r}(x)$。 $\square$

$\mathbb{R}[x]$ 中的任一多项式 $f(x)$ 都可看成 $\mathbb{C}$ 上的函数

$$f: \mathbb{C} \longrightarrow \mathbb{C}, \quad z \longmapsto f(z).$$

引理 1.5.3：设 $f(x), g(x) \in \mathbb{R}[x]$．则

$$f(x) = g(x) \iff \forall z \in \mathbb{C}, \text{有 } f(z) = g(z).$$

□

上述引理实际上由下述结论推出，$z \in \mathbb{C}$ 称为多项式 $f(x) = a_n x^n + a_{n-1} x^{n-1} + \cdots + a_1 x + a_0$ 的一个根．如果

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

命题 1.5.1：设 $f(x) \in \mathbb{R}[x]$ 是一个 n 次多项式．则 $f(x)$ 最多有 n 个根．

证明：注意 $\mathbb{R}[x] \subset \mathbb{C}[x]$；只需证明：$\forall f(x) \in \mathbb{C}[x]$．如果 $f(x)$ 非零，则 $\#\{z \in \mathbb{C} \mid f(z) = 0\} \leq \deg(f)$．

可对多项式次数应用归纳法：如果 $\deg(f) = 0$，则 $f(x)$ 没有根（因为 $f(x)$ 非零）。设结论对 $n-1$ 次多项式成立。如果 $\deg f(x) = n$，且 $\alpha_1 \in \mathbb{C}$ 是 $f(x)$ 的一个根．则由引理 1.5.2，存在 $q(x), r(x) \in \mathbb{C}[x]$，使 $f(x) = q(x)\cdot(x - \alpha_1) + r(x)$．$r(x) = 0$ 或 $\deg r(x) < \deg(x - \alpha_1) = 1$．由 $f(\alpha_1) = 0$，得知 $r(x)$ 必为零多项式即 $f(x) = q(x)(x - \alpha_1)$，$\deg q(x) = n-1$．所以

$$\{z \in \mathbb{C} \mid f(z) = 0\} = \{\alpha_1\} \cup \{z \in \mathbb{C} \mid q(z) = 0\}$$

最多只有 n 个元素．

□

习题 1.5

1. 设 $f(x_1,x_2,x_3)=x_1^3-2x_3^3+2x_2^2-x_2x_3-x_1+2$，$g(x_1,x_2,x_3)=x_2^4+2x_1x_2+3x_3^2+4x_1-2x_2+3$. 请计算

   (1) $f(x_1,x_2,x_3)+g(x_1,x_2,x_3)$，$f(x_1,x_2,x_3)g(x_1,x_2,x_3)$

   (2) f_d $(0\le d\le 3)$，g_d $(0\le d\le 4)$.

   (3) $\deg(f)$，$\deg(g)$，$\deg(f+g)$，$\deg(f\cdot g)$.

2. 对下列多项式 $f(x),g(x)\in\mathbb{Q}[x]$ 求带余除法中的 $q(x),r(x)\in\mathbb{Q}[x]$ 使 $f(x)=q(x)g(x)+r(x)$。

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

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

3. 设 $\mathbb{R}^n=\{(a_1,\cdots,a_n)\mid a_i\in\mathbb{R}\}$. $\forall$ $(a_1,\cdots,a_n),(b_1,\cdots,b_n)\in\mathbb{R}^n$，$\forall\lambda\in\mathbb{R}^n$，定义：(1) $(a_1,\cdots,a_n)+(b_1,\cdots,b_n)=(a_1+b_1,\cdots,a_n+b_n)$ (2) $\lambda\cdot(a_1,\cdots,a_n)=(\lambda a_1,\cdots,\lambda a_n)$. 设

$$M(\mathbb{R}^n)=\{\varphi:\mathbb{R}^n\longrightarrow\mathbb{R}^n\mid \varphi\text{是映射}\}.$$

$\forall\lambda\in\mathbb{R}$，$\varphi,\psi\in M(\mathbb{R}^n)$，$\lambda\varphi$，$\varphi+\psi$，$\varphi\psi\in M(\mathbb{R}^n)$ 分别定义为：$(\lambda\varphi)(\alpha)=\lambda\varphi(\alpha)$，$(\varphi+\psi)(\alpha)=\varphi(\alpha)+\psi(\alpha)$，$\forall\alpha\in\mathbb{R}^n$，$(\varphi\psi)(\alpha)=\varphi(\psi(\alpha))$. $\forall f(x)=a_nx^n+\cdots+a_1x+a_0\in\mathbb{R}[x]$，$\varphi\in M(\mathbb{R}^n)$，$f(\varphi):\mathbb{R}^n\longrightarrow\mathbb{R}^n$，定义为：$\forall\alpha\in\mathbb{R}^n$，

$$f(\varphi)(\alpha):=a_n\varphi^n(\alpha)+\cdots+a_1\varphi(\alpha)+a_0\alpha.$$

证明：(1) 如果 $h(x)=f(x)g(x)$，则 $h(\varphi)=f(\varphi)g(\varphi)$.

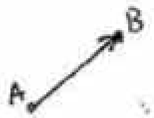
(2) 如果 $h(x)=f(x)+g(x)$，则 $h(\varphi)=f(\varphi)+g(\varphi)$.

第二章. 空间几何的线性化.

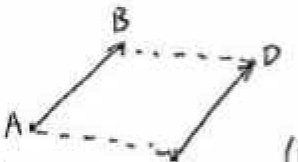
物理中的矢量是指既有大小,也有方向的物理量. 比如: 位移, 力, 速度等. 牛顿用空间中的有向线段来表示这些量, 比如作用于一点的两个力的合力可用平行四边形的对角线表示等。以此为背景, 我们可以定义有向线段的"加法"和"数乘运算", 它们满足8条规则, 从而使得"有向线段"关于上述运算成为一个向量空间。本章主要介绍如何利用上述向量空间来描述和解决空间几何问题, 为后面线性代数的学习建立几何背景。

§2.1: 向量空间.

不考虑矢量的物理意义, 数学中的向量是指: 既有大小, 又有方向的向量。可以用有向线段 $\overrightarrow{AB}$ (见图1.1) 来表示向量, 其中点A叫做向量 $\overrightarrow{AB}$ 的起点, 点B叫做向量 $\overrightarrow{AB}$ 的终点, 箭头表示向量的方向。如果一个有向线段 $\overrightarrow{CD}$ 可由 $\overrightarrow{AB}$ 通过平行移动得到(图1.2), 则认为它们代表相同的向量.



(图1.1)



(图1.2).

向量的长度定义为该有向线段 $\overrightarrow{AB}$ 的长度。

我们也可以在所有"有向线段"的集合中定义等价关系: $\overrightarrow{AB} \sim \overrightarrow{CD}$ 当且仅当它们通过平移可以重合, 从而可以将向量定义为有向线段的等价类, 通常用 α, β, γ 表示: $\alpha = \overrightarrow{AB} = \overrightarrow{CD}$. (如果 $\overrightarrow{AB} \sim \overrightarrow{CD}$).

定义2.1.1：表述向量α的有向线段$\overrightarrow{AB}$的长度称为α的模，记为$|\alpha|$。模为零的向量称为零向量，记为0. □

注记 (1) 表达零向量0的有向线段的"始点"与"终点"必重合；所以零向量0是唯一方向不确定的向量。

(2) 对于给定的非零向量α，我们可以用不同（但等价）的有向线段$\overrightarrow{AB}$, $\overrightarrow{CD}$来表达。

定义2.1.2：设α是一个向量，与α方向相反，但模长相等的向量称为α的负向量，记为$-\alpha$. □

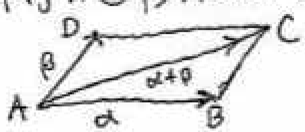
例2.1.1：如果α可由有向线段$\overrightarrow{AB}$表示，则$-\alpha$可由$\overrightarrow{BA}$表示。如果$|\alpha|=0$，则$|-\alpha|=0$. 所以，零向量的负向量还是零向量。□

例2.1.2：如果在空间中固定一点O，则所有"有向线段"可以等价于以O为"始点"的有向线段。如果用V表示所有向量的集合，Ω表示我们现实空间的点集合，则

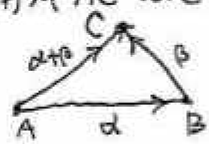
$$\Omega \longrightarrow V, \quad A \longmapsto \overrightarrow{OA}$$

是一个双射。□

定义2.1.3（向量的加法）. $\forall\ \alpha, \beta \in V$，我们定义向量$\alpha+\beta \in V$如下：通过平行移动可设表示$\alpha, \beta$的有向线段分别是：$\alpha = \overrightarrow{AB}$, $\beta = \overrightarrow{AD}$. 则$\alpha+\beta$定义为由平行四边形ABCD（见图1.3）的对角线$\overrightarrow{AC}$表达的向量。



(图1.3).



(图1.4). □

上述求 $\alpha+\beta$ 的方法称为<u>平行四边形法则</u>，另一个等价的法则称为<u>三角形法则</u>：将 α,β 分别用有向线段 $\overrightarrow{AB}$，$\overrightarrow{BC}$ 表述，则 $\alpha+\beta$ 可由三角形 $\triangle ABC$ 的第三边 $\overrightarrow{AC}$ 表述（见图1.4）。

<u>定义2.1.4</u>（向量的数乘法）。$\forall\ \alpha\in V$，$\lambda\in\mathbb{R}$．则

$\lambda\alpha\in V$ 的模定义为 $|\lambda||\alpha|$（即：$|\lambda\alpha|=|\lambda||\alpha|$），而 $\lambda\alpha$ 的方向定义如下：(1) 若 λ 或 α 有一个为零，则 $\lambda\alpha=0$；(2) 若 λ,α 均不为零，则 [illegible] 当 $\lambda>0$ 时，$\lambda\alpha$ 与 α 方向相同，否则 $\lambda\alpha$ 与 α 方向相反。

<u>例2.1.3</u>：设 $\alpha=\overrightarrow{AB}\neq 0$，$\lambda\in\mathbb{R}$，则当 $\lambda>0$ 时，$\lambda\alpha=\lambda\overrightarrow{AB}$．而当 $\lambda<0$ 时，$\lambda\alpha=(-\lambda)\overrightarrow{BA}$．

<u>定理2.1.1</u>：设 V 是全体向量的集合．则上述定义中的"向量加法"和"数乘运算"满足如下8个条件：

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

(2) 存在向量 $0\in V$，满足：$\forall\ \alpha\in V$，有 $\alpha+0=0+\alpha=\alpha$（零元存在性）

(3) $\forall\ \alpha\in V$，存在 $\beta\in V$，使 $\alpha+\beta=\beta+\alpha=0$．（负元存在性）．

(4) $\forall\alpha,\beta\in V$，有 $\alpha+\beta=\beta+\alpha$．（交换律）．

(5) $\forall\ \alpha\in V$，有 $1\alpha=\alpha$．

(6) $\forall\ \lambda,\mu\in\mathbb{R},\alpha\in V$，有 $\lambda(\mu\alpha)=(\lambda\mu)\alpha$．（数乘结合律）．

(7) $\forall\ \lambda,\mu\in\mathbb{R},\alpha\in V$，有 $(\lambda+\mu)\alpha=\lambda\alpha+\mu\alpha$

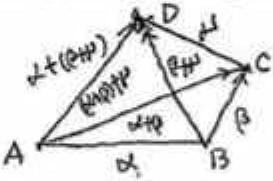
(8) $\forall\ \lambda\in\mathbb{R},\alpha,\beta\in V$，有 $\lambda(\alpha+\beta)=\lambda\alpha+\lambda\beta$．

（分配律）

证明：(1) 可由加法定义的三角形法则验证．（如图1.5）

零向量显然满足(2)的要求，而 $-\alpha$ 满足(3)中对 β 的要求.

(4) 可由加法定义的平行四边形法则验证.(图1.6)



(图1.5)

(图1.6)

~~(5) 可由定义直接验证~~.

(5), (6) 和 (7) 可由定义直接验证. 例如(7), 可设 α,λ,μ 均非零(否则显然). 如果 $\lambda>0,\mu>0$. 则 $(\lambda+\mu)\alpha$ 与 $\lambda\alpha+\mu\alpha$ 的方向与 α 一致. 且 $|\lambda\alpha+\mu\alpha| = |\lambda\alpha| + |\mu\alpha| = |\lambda||\alpha|+|\mu||\alpha|$

$$|(\lambda+\mu)\alpha| = |\lambda+\mu||\alpha| = |\lambda||\alpha|+|\mu||\alpha| = |\lambda\alpha+\mu\alpha|.$$

最后我们证明(8). 不妨设 λ,α,β 均非零. 如果 α,β 平行, 则存在 $\mu\in\mathbb{R}$ 使 $\beta=\mu\alpha$. 从而 $\lambda(\alpha+\beta)=\lambda(\alpha+\mu\alpha)=\lambda(1+\mu)\alpha$

$$\lambda(1+\mu)\alpha=(\lambda+\lambda\mu)\alpha=\lambda\alpha+\lambda\mu\alpha=\lambda\alpha+\lambda\beta.$$

如果 α,β 不平行, 则(8)可由下图证明: $\lambda\alpha+\lambda\beta=\overrightarrow{AB}=\lambda\overrightarrow{AC}=\lambda(\alpha+\beta)$ (不妨设 $\lambda>0$).

(图1.7)

□.

<u>定义2.1.5</u>: $\forall\ \alpha,\beta\in V$, 定义 $\alpha-\beta=\alpha+(-\beta)$.

$\forall\ \alpha_1,\alpha_2,\cdots,\alpha_n\in V$, $\lambda_1,\cdots,\lambda_n\in\mathbb{R}$, 则向量 $\lambda_1\alpha_1+\lambda_2\alpha_2+\cdots+\lambda_n\alpha_n\in V$ 称为 $\alpha_1,\cdots,\alpha_n$ 的一个线性组合. $\lambda_1,\cdots,\lambda_n$ 称为该线性组合的系数.

称向量 $\beta\in V$ 可由 $\alpha_1,\cdots,\alpha_n$ 线性表出, 如果存在 $\lambda_1,\cdots,\lambda_n\in\mathbb{R}$, 使 $\beta=\lambda_1\alpha_1+\cdots+\lambda_n\alpha_n$

□.

<u>注记</u>: 零向量可由任意一组向量线性表出.

定义2.1.6：设 $\alpha_1,\cdots,\alpha_n\in V$ 是一组向量，则

(1) $\alpha_1,\cdots,\alpha_n$ 称为共线，如果它们可由一条固定直线上的有向线段表示；

(2) $\alpha_1,\cdots,\alpha_n$ 称为共面，如果它们可由一个固定平面上的有向线段表示；

(3) $\alpha_1,\cdots,\alpha_n$ 称为线性相关，如果存在不全为零的数 $\lambda_1,\cdots,\lambda_n\in\mathbb{R}$ 使 $\lambda_1\alpha_1+\cdots+\lambda_n\alpha_n=0$。否则，$\alpha_1,\cdots,\alpha_n$ 称为线性无关。

□

练习：证明：(1) $\alpha_1,\cdots,\alpha_n$ 线性相关 $\Leftrightarrow$ 其中的某个 α_i 可由其它向量线性表出；(2) 设 $\alpha_1,\cdots,\alpha_n$ 线性无关，$\beta\in V$。则 $\alpha_1,\cdots,\alpha_n,\beta$ 线性相关 $\Leftrightarrow$ 存在唯一的一组数 $\lambda_1,\cdots,\lambda_n\in\mathbb{R}$ 使得 $\beta=\lambda_1\alpha_1+\cdots+\lambda_n\alpha_n$（此时称 β 可由 $\alpha_1,\cdots,\alpha_n$ 唯一线性表出）.

定理2.12：设 α,β,γ 是 V 中的三个向量。则

(1) α,β 共线（的充要条件是 $\Leftrightarrow$）α,β 线性相关；

(2) α,β,γ 共面，但 α,β 不共线 $\Leftrightarrow$ α,β,γ 线性相关，但 α,β 线性无关

(3) α,β,γ 不共面 $\Leftrightarrow$ α,β,γ 线性无关.

(4) 如果 α,β,γ 不共面，则对任意 $\delta\in V$，$\alpha,\beta,\gamma,\delta$ 线性相关（因此，δ 可由 α,β,γ 唯一线性表出）.

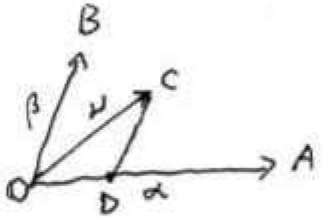
证明：(1) 当 α,β 中有一个为零，则 α,β 总是共线，α,β 也总是线性相关，因此结论成立。如果 α,β 均不为零，则 α,β 共线当且仅当存在 $\lambda\in\mathbb{R}$ 使 $\beta=\lambda\alpha$。而当 α,β 均不为零时，不难证明：

$$\alpha,\beta \text{ 线性相关} \Leftrightarrow \beta=\lambda\alpha,\ \lambda\in\mathbb{R}.$$

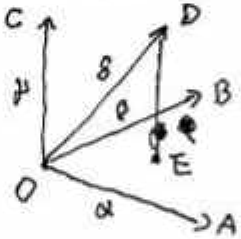
(2) 由(1)，α,β 不共线 $\Leftrightarrow$ α,β 线性无关。下面证明：

α,β,γ 共面 $\Leftrightarrow$ α,β,γ 线性相关.

令 $\alpha=\overrightarrow{OA}$, $\beta=\overrightarrow{OB}$, $\gamma=\overrightarrow{OC}$ (如图1.8), 则点 O, A, B 不共线.



(图1.8)



(图1.9)

但 O, A, B, C 四点共面. 过 C 点作平行于 $\overrightarrow{OB}$ 的直线交直线 OA 于 D 点 (如图1.8). 则存在 $\lambda,\mu\in\mathbb{R}$ 使 $\overrightarrow{DC}=\lambda\overrightarrow{OB}=\lambda\beta$, $\overrightarrow{OD}=\mu\alpha$.

所以, $\gamma=\overrightarrow{OC}=\overrightarrow{OD}+\overrightarrow{DC}=\mu\alpha+\lambda\beta$. 即 α,β,γ 线性相关.

反之, 如果 α,β,γ 线性相关, 则存在 $\mu,\lambda\in\mathbb{R}$ 使 $\gamma=\mu\alpha+\lambda\beta$ (因为 α,β 线性无关). 从而 $\overrightarrow{OC}$ 的落在由 O, A, B 确定的平面内. 因此 α,β,γ 共面.

(3) 可由 (1), (2) 推出: α,β,γ 不共面 $\Rightarrow$ α,β 不共线 $\overset{(1)}{\Rightarrow}$ α,β 线性无关 $\Rightarrow$ α,β,γ 线性无关 (否则, 由(2), α,β,γ 共面). 反之, α,β,γ 线性无关 $\Rightarrow$ α,β 也线性无关 $\Rightarrow$ α,β,γ 不共面 (否则, 由(2), α,β,γ 线性相关).

(4) 如果 α,β,γ 不共面, 则对任意 $\delta\in V$, 我们证明存在唯一一组数 $\lambda_1,\lambda_2,\lambda_3\in\mathbb{R}$ 使 $\delta=\lambda_1\alpha+\lambda_2\beta+\lambda_3\gamma$. 令

$$\alpha=\overrightarrow{OA},\quad \beta=\overrightarrow{OB},\quad \gamma=\overrightarrow{OC},\quad \delta=\overrightarrow{OD}\quad (\text{如图}1.9).$$

因为四点 O, A, B, C 不共面, 过点 D 平行于 $\overrightarrow{OC}$ 的直线与平面 OAB 相交于唯一一点 E (见图1.9). 由于 ED 平行于 $\overrightarrow{OC}$, 所以存在 $\lambda_3\in\mathbb{R}$ 使 $\overrightarrow{ED}=\lambda_3\gamma$. 又 $\alpha,\beta,\overrightarrow{OE}$ 共面 (α,β 不共线), 由(2) 存在 $\lambda_1,\lambda_2\in\mathbb{R}$ 使 $\overrightarrow{OE}=\lambda_1\alpha+\lambda_2\beta$. 因此,

$$\delta=\overrightarrow{OD}=\overrightarrow{OE}+\overrightarrow{ED}=\lambda_1\alpha+\lambda_2\beta+\lambda_3\gamma.$$

且 α,β,γ 线性无关推出这样的 $\lambda_1,\lambda_2,\lambda_3$ 必由 δ 唯一确定

□

推论 2.1.1：固定一组不共面的向量 $e_1, e_2, e_3 \in V$，则对任意 $\alpha \in V$，存在一组唯一由 α 确定的数 $l_1(\alpha), l_2(\alpha), l_3(\alpha) \in \mathbb{R}$ 使

$$\alpha = l_1(\alpha)e_1 + l_2(\alpha)e_2 + l_3(\alpha)e_3.$$

且 α 与 e_1, e_2 共面的充要条件是 $l_3(\alpha)=0$。将 $l_i(\alpha)$ 看成 V 上的函数 $l_i : V \to \mathbb{R}$ $(i=1,2,3)$ 满足下述条件：

$$l_i(\alpha+\beta) = l_i(\alpha) + l_i(\beta), \quad l_i(\lambda\alpha) = \lambda l_i(\alpha), \quad \forall \alpha, \beta \in V, \lambda \in \mathbb{R}.$$

□

例 2.1.4：证明：A_1, A_2, A_3 三点共线 $\Longleftrightarrow$ 存在不全为零的实数 $\lambda_1, \lambda_2, \lambda_3$ 使得

$$\lambda_1 \overrightarrow{OA_1} + \lambda_2 \overrightarrow{OA_2} + \lambda_3 \overrightarrow{OA_3} = 0, \quad \lambda_1 + \lambda_2 + \lambda_3 = 0$$

其中 O 是任意取定的点。

证明："$\Rightarrow$"。如果 A_1, A_2, A_3 共线，则存在 $\lambda, \mu \in \mathbb{R}$ 不全为零使 $\lambda \overrightarrow{A_1A_2} + \mu \overrightarrow{A_1A_3} = 0$。因此，对任意一点 O，有 $\lambda(\overrightarrow{OA_2} - \overrightarrow{OA_1}) + \mu(\overrightarrow{OA_3} - \overrightarrow{OA_1}) = 0$，整理可得

$$(-\lambda-\mu)\overrightarrow{OA_1} + \lambda \overrightarrow{OA_2} + \mu \overrightarrow{OA_3} = 0,$$ 令 $\lambda_1 = -\lambda-\mu, \lambda_2 = \lambda, \lambda_3 = \mu$

即得结论。

"$\Leftarrow$"。若存在不全为零的 $\lambda_1, \lambda_2, \lambda_3 \in \mathbb{R}$ 使得

$$\lambda_1 \overrightarrow{OA_1} + \lambda_2 \overrightarrow{OA_2} + \lambda_3 \overrightarrow{OA_3} = 0, \text{ 且 } \lambda_1+\lambda_2+\lambda_3 = 0.$$

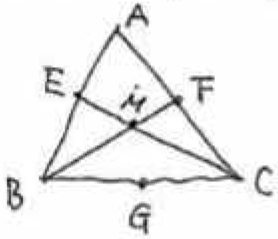
则

$$\begin{aligned} 0 &= \lambda_1 \overrightarrow{OA_1} + \lambda_2 \overrightarrow{OA_2} + \lambda_3 \overrightarrow{OA_3} = (-\lambda_2-\lambda_3)\overrightarrow{OA_1} + \lambda_2 \overrightarrow{OA_2} + \lambda_3 \overrightarrow{OA_3} \\ &= \lambda_2(\overrightarrow{OA_2} - \overrightarrow{OA_1}) + \lambda_3(\overrightarrow{OA_3} - \overrightarrow{OA_1}) = \lambda_2 \overrightarrow{A_1A_2} + \lambda_3 \overrightarrow{A_1A_3} \end{aligned}$$

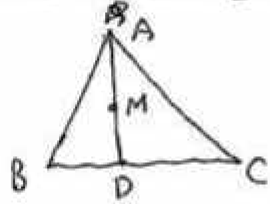
所以，$\overrightarrow{A_1A_2}, \overrightarrow{A_1A_3}$ 共线。从而 A_1, A_2, A_3 共线。 □

例 2.1.5：证明：三角形 $\triangle ABC$ 的三条中线交于一点 M，且 $\overrightarrow{MA} + \overrightarrow{MB} + \overrightarrow{MC} = 0$。

证明：设 E，F，G 分别是边 AB，AC，BC 上的中点。M 是中线 BF 与 CE 的交点（如图1.20）. 要证中线 AG 通过M，只需证明 $\overrightarrow{AM}$，$\overrightarrow{AG}$ 共线。显然，$\overrightarrow{AG}=\overrightarrow{AB}+\overrightarrow{BG}=\overrightarrow{AB}+\frac{1}{2}\overrightarrow{BC}=\frac{1}{2}\overrightarrow{AB}+\frac{1}{2}\overrightarrow{AC}$.



（图1.20）



（图1.21）

为了将 $\overrightarrow{AM}$ 也用 $\overrightarrow{AB}$，$\overrightarrow{AC}$ 线性表出，令 $\overrightarrow{BM}=\lambda\overrightarrow{BF}$，$\overrightarrow{CM}=\mu\overrightarrow{CE}$，

则 $\overrightarrow{AM}=\overrightarrow{AB}+\overrightarrow{BM}=\overrightarrow{AB}+\lambda\overrightarrow{BF}=\overrightarrow{AB}+\lambda(\overrightarrow{BA}+\frac{1}{2}\overrightarrow{AC})=(1-\lambda)\overrightarrow{AB}+\frac{\lambda}{2}\overrightarrow{AC}$.

同理，$\overrightarrow{AM}=\overrightarrow{AC}+\overrightarrow{CM}=\overrightarrow{AC}+\mu\overrightarrow{CE}=\overrightarrow{AC}+\mu(\overrightarrow{CA}+\frac{1}{2}\overrightarrow{AB})=(1-\mu)\overrightarrow{AC}+\frac{\mu}{2}\overrightarrow{AB}$.

所以，$(1-\lambda-\frac{\mu}{2})\overrightarrow{AB}=(1-\mu-\frac{\lambda}{2})\overrightarrow{AC}$. 但 $\overrightarrow{AB}$，$\overrightarrow{AC}$ 不共线，所以

$1-\lambda-\frac{\mu}{2}=0$，$1-\mu-\frac{\lambda}{2}=0$，可得 $\lambda=\mu=\frac{2}{3}$，从而

$\overrightarrow{AM}=\frac{1}{3}\overrightarrow{AB}+\frac{1}{3}\overrightarrow{AC}=\frac{2}{3}(\frac{1}{2}\overrightarrow{AB}+\frac{1}{2}\overrightarrow{AC})=\frac{2}{3}\overrightarrow{AG}$，即AG通过M.

又 $\overrightarrow{MA}+\overrightarrow{MB}+\overrightarrow{MC}=-\frac{2}{3}(\overrightarrow{AG}+\overrightarrow{BF}+\overrightarrow{CE})=-\frac{1}{3}(\overrightarrow{AB}+\overrightarrow{BC}-\overrightarrow{AC})=0$.

□.

<u>例2.1.6</u>：证明：点 M 在三角形 $\triangle ABC$ 内部(包括三条边)的充要条件是存在<u>非负实数</u> λ,μ 使

$$\overrightarrow{AM}=\lambda\overrightarrow{AB}+\mu\overrightarrow{AC} \quad 且 \quad \lambda+\mu\le 1.$$

证明：M 在 $\triangle ABC$ 内部 $\Leftrightarrow$ 存在BC上的一个点 D（如图1.21）使 $\overrightarrow{AM}=\lambda_1\overrightarrow{AD}$，$(0\le\lambda_1\le 1)$

令 $\overrightarrow{BD}=\mu_1\overrightarrow{BC}$ $(0\le\mu_1\le 1)$。可得 $\overrightarrow{AD}=\overrightarrow{AB}+\overrightarrow{BD}=\overrightarrow{AB}+\mu_1(\overrightarrow{AC}-\overrightarrow{AB})$，

$\overrightarrow{AM}=(\lambda_1-\lambda_1\mu_1)\overrightarrow{AB}+\lambda_1\mu_1\overrightarrow{AC}$。反之，如果存在 $0\le\lambda\le 1$，$0\le\mu\le 1$ 使

$$\overrightarrow{AM}=\lambda\overrightarrow{AB}+\mu\overrightarrow{AC}，且 \lambda+\mu\le 1.$$

则 AM 与 BC 不平行（否则，$\overrightarrow{AM}=\lambda'\overrightarrow{BC}=-\lambda'\overrightarrow{AB}+\lambda'\overrightarrow{AC}$，$\overrightarrow{AM}$ 必为零向量）.

令 D 是直线 AM 与 BC 的交点，$\overrightarrow{BD}=\mu_1\overrightarrow{BC}$，$\overrightarrow{AM}=\lambda_1\overrightarrow{AD}$.

则只需证明：$0\le\mu_1\le1$（说明 D 在 B, C 之间），$0\le\lambda_1\le1$（说明 M 在 A, D 之间）

由 $\overrightarrow{AM}=\lambda_1\overrightarrow{AD}=\lambda_1(\overrightarrow{AB}+\overrightarrow{BD})=\lambda_1(\overrightarrow{AB}+\mu_1\overrightarrow{BC})=(\lambda_1-\lambda_1\mu_1)\overrightarrow{AB}+\lambda_1\mu_1\overrightarrow{AC}$

因为 $\overrightarrow{AB}$, $\overrightarrow{AC}$ 不共线，所以 $\lambda=\lambda_1-\lambda_1\mu_1$，$\mu=\lambda_1\mu_1$. 不难证明，

$0\le\lambda\le1$ $0\le\mu\le1$ $\Rightarrow$ $\lambda_1\ge0$, $\mu_1\ge0$，而 $\lambda+\mu\le1\Rightarrow\lambda_1\le1$, $\mu_1\le1$。

□

习题 2.1

1.1. 设 AC, BD 是平行四边形 ABCD 的两条对角线. 已知向量 $\overrightarrow{AC}=\alpha$, $\overrightarrow{BD}=\beta$，求向量 $\overrightarrow{AB}$，和 $\overrightarrow{BC}$.

1.2. 试证：四点 A, B, C, D 共面的充要条件是：对任意点 O，存在不全为零的 $\lambda,\mu,\upsilon,\omega\in\mathbb{R}$ 使得 $\lambda+\mu+\upsilon+\omega=0$，且.

$$\lambda\overrightarrow{OA}+\mu\overrightarrow{OB}+\upsilon\overrightarrow{OC}+\omega\overrightarrow{OD}=0.$$

1.3. 设 O, A, B 三点不共线，则点 C 与 A, B 共线的充要条件是：存在 $s\in\mathbb{R}$ 使 $\overrightarrow{OC}=(1-s)\overrightarrow{OA}+s\overrightarrow{OB}$，且 s 与 O 无关，由 A, B, C 唯一确定。

1.4. 设 A, B, C 不共线，试证：点 M 与 A, B, C 共面的充要条件是：存在 $\lambda,\mu,\upsilon\in\mathbb{R}$ 使 $\lambda+\mu+\upsilon=1$，且

$$\overrightarrow{OM}=\lambda\overrightarrow{OA}+\mu\overrightarrow{OB}+\upsilon\overrightarrow{OC}.$$

其中 O 是任意取定的点.

1.5. 证明：点 M 在三角形 ΔABC 内部（包括三条边）的充要条件是：对任意给定点 O，存在非负实数 λ,μ,υ 使得

$$\overrightarrow{OM}=\lambda\overrightarrow{OA}+\mu\overrightarrow{OB}+\upsilon\overrightarrow{OC}，且\ \lambda+\mu+\upsilon=1.$$

1.6. 如果向量 α,β,γ 共面，证明其中至少有一个向量可以表示成其余向量的线性组合。是否每个向量均可由其余两个向量线性表出？

1.7. 设 $A_1, A_2, \cdots, A_n$ 是空间的任意一组点. 证明:

(1) 存在唯一的点 M 使得 $\overrightarrow{MA_1}+\cdots+\overrightarrow{MA_n}=0$;

(2) 对任意点 O, 有 $\overrightarrow{OA_1}+\cdots+\overrightarrow{OA_n}=n\overrightarrow{OM}$.

(点 M 称为点组 $A_1, A_2, \cdots, A_n$ 的重心).

1.8. 设 A, B, C, D 是空间中任意4个点, P, Q 分别是线段 AB, CD 的中点, 证明: 线段PQ的中点就是 A, B, C, D 的重心.

1.9. 证明: 四面体的3对对棱中点的连线(共三条)相交于一点, 且此点就是四面体4个顶点的重心.

1.10. 设 $A_1, A_2, \cdots, A_n$ 是圆周上的n个等分点. O是该圆的圆心. 证明: $\overrightarrow{OA_1}+\cdots+\overrightarrow{OA_n}=0$. (即O是$A_1, \cdots, A_n$的重心).

1.11. 设 $X=\{A_1, A_2, \cdots, A_n\}$, $Y=\{B_1, B_2, \cdots, B_n\}$ 是空间中给定的两组点. 证明: 对任意双射 $f: X\to Y$, 向量

$$\overrightarrow{A_1f(A_1)}+\overrightarrow{A_2f(A_2)}+\cdots+\overrightarrow{A_nf(A_n)}$$

与 f 的选择无关。

1.12. (Ceva定理). 设 D, E, F 依次是三角形ΔABC的边 AB, BC, CA 的内点, 且 $\overrightarrow{AD}=\lambda\overrightarrow{DB}$, $\overrightarrow{BE}=\mu\overrightarrow{EC}$, $\overrightarrow{CF}=\nu\overrightarrow{FA}$ ($\lambda,\mu,\nu\in\mathbb{R}$). 证明: 三条线段 AE, BF, CD 相交于一点 $\Longleftrightarrow$ $\lambda\mu\nu=1$.

1.13. (Menelaus定理). 设 A, B, C 不共线, 点 P, Q, R 依次在直线 AB, BC, CA 上(不与A, B, C重合). 令 $\overrightarrow{AP}=\lambda\overrightarrow{PB}$, $\overrightarrow{BQ}=\mu\overrightarrow{QC}$, $\overrightarrow{CR}=\nu\overrightarrow{RA}$. ($\lambda,\mu,\nu\in\mathbb{R}$). 证明: P, Q, R三点共线 $\Longleftrightarrow$ $\lambda\mu\nu=-1$.

1.14. 通过向量运算证明: 平行四边形的对角线互相平分.

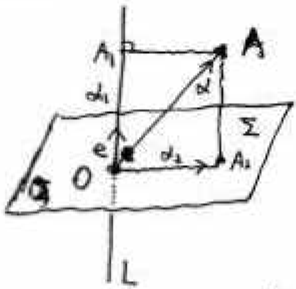
1.15. 证明: $\forall\ \alpha,\beta\in V$, 有 $|\alpha+\beta|\le|\alpha|+|\beta|$. 请说明等号成立的充要条件是什么?

1.16. 证明: $\forall\ \alpha,\beta\in V$, 有 $2|\alpha|^2+2|\beta|^2=|\alpha+\beta|^2+|\alpha-\beta|^2$.
如果α,β是不平行的非零向量, 上述恒等式对应的几何定理是什么?

§2.2. 内积、外积与混合积

上一节引入的向量空间是线性代数中线性空间的重要例子，它的线性运算反映了现实空间中三点共线，四点共面，直线相交等几何性质。为了进一步反映现实空间的度量性质，例如距离，夹角，面积，体积等，我们在本节中介绍向量的内积，外积和混合积等多重线性映射的例子，它们具有很强的物理和几何背景。为了讨论它们，我们需要定义向量α在一条给定直线和平面上的投影，它正好提供了研究一般线性映射（之间的）的理由和特例。

固定一个方向向量e和与它垂直的平面Σ，令L是以e为方向的直线，L与Σ的交点（垂足）为O.（如图2.1）：



（图 2.1）

对任意向量$\alpha\in V$，令 $\alpha=\overrightarrow{OA}$，$A_1$是$A$在直线$L$上的垂足，$A_2$是$A$在平面$\Sigma$上的垂足。则 $\alpha=\overrightarrow{OA_1}+\overrightarrow{OA_2}=\alpha_1+\alpha_2$，且

$$\alpha_1=\frac{|\alpha|}{|e|}\cos\langle\alpha,e\rangle e,\quad |\alpha_2|=|\alpha|\sin\langle e,\alpha\rangle.$$

其中α_1的方向由e给定，α_2的方向垂直于e但由α的方向而确定。

例2.2.1. α_1,α_2由α唯一确定，分别称为α在L上的投影和在Σ上的投影，记为：$\alpha_1=P_e(\alpha)$，$\alpha_2=P_\Sigma(\alpha)$.

证明：如果 $\alpha=\alpha_1+\alpha_2=\alpha_1'+\alpha_2'$，其中$\alpha_1'$平行$e$，$\alpha_2'$垂直于$e$。则 $\alpha_1-\alpha_1'=\alpha_2'-\alpha_2$ 既平行于e，又垂直于e，所以均为零 □

引理 2.2.1：上述投影映射 $P_e: V\to V$ 和 $P_\Sigma: V\to V$ 满足条件：

(1) $P_e(\lambda\alpha)=\lambda P_e(\alpha)$，$P_e(\alpha+\beta)=P_e(\alpha)+P_e(\beta)$，$\forall\lambda\in\mathbb{R},\alpha,\beta\in V$.

(2) $P_\Sigma(\lambda\alpha)=\lambda P_\Sigma(\alpha)$，$P_\Sigma(\alpha+\beta)=P_\Sigma(\alpha)+P_\Sigma(\beta)$，$\forall\lambda\in\mathbb{R},\alpha,\beta\in V$.

证明：在 Σ 上固定两个不共线的向量 e_1,e_2，则 e_1,e_2,e 不共面.
由推论 2.1.1，对任意 $\alpha\in V$，存在唯一由 α 确定的数 $l_1(\alpha),l_2(\alpha),l_3(\alpha)$ 使得 $\alpha=l_1(\alpha)e_1+l_2(\alpha)e_2+l_3(\alpha)e$. 由分解唯一性得

$$P_e(\alpha)=l_3(\alpha)e\ ,\quad P_\Sigma(\alpha)=l_1(\alpha)e_1+l_2(\alpha)e_2\ .$$

由于推论 2.1.1 中的函数 $l_i: V\to\mathbb{R}$ 满足：~~$l_i(\lambda\alpha)=\lambda l_i(\alpha)$~~

$$l_i(\lambda\alpha)=\lambda l_i(\alpha)\ ,\quad l_i(\alpha+\beta)=l_i(\alpha)+l_i(\beta)\ .\ \forall\lambda\in\mathbb{R},\alpha,\beta\in V.$$

所以 P_e，P_Σ 满足引理中的条件. □

定义 2.2.1：映射 $f: V\to V$ 称为线性映射，如果

$$f(\lambda\alpha)=\lambda f(\alpha)\ ,\quad f(\alpha+\beta)=f(\alpha)+f(\beta)\ ,\ \forall\lambda\in\mathbb{R},\alpha,\beta\in V.$$

特别，映射 $l: V\to\mathbb{R}$ 称为<u>线性函数</u>，如果

$$l(\lambda\alpha)=\lambda l(\alpha)\ ,\quad l(\alpha+\beta)=l(\alpha)+l(\beta)\ ,\ \forall\lambda\in\mathbb{R},\alpha,\beta\in V.$$

□

例 2.2.2：$P_e: V\to V$，$P_\Sigma: V\to V$ 是线性映射，且 $P_e(V)$ 中的向量共线（即平行于同一条直线），$P_\Sigma(V)$ 中的向量共面（即平行于同一个平面）。固定任意三个不共面向量 e_1,e_2,e_3，令

$$\alpha=l_1(\alpha)e_1+l_2(\alpha)e_2+l_3(\alpha)e_3\ ,\quad\forall\alpha\in V.$$

则 $l_i: V\to\mathbb{R}$，$(1\le i\le 3)$ 都是线性函数.

证明：仅证 $l_i(\alpha+\beta)=l_i(\alpha)+l_i(\beta)$. 由定义，$\alpha=l_1(\alpha)e_1+l_2(\alpha)e_2+l_3(\alpha)e_3$，$\beta=l_1(\beta)e_1+l_2(\beta)e_2+l_3(\beta)e_3$. $\alpha+\beta=l_1(\alpha+\beta)e_1+l_2(\alpha+\beta)e_2+l_3(\alpha+\beta)e_3$。但 $\alpha+\beta=(l_1(\alpha)+l_1(\beta))e_1+(l_2(\alpha)+l_2(\beta))e_2+(l_3(\alpha)+l_3(\beta))e_3$，由于 e_1,e_2,e_3 线性无关，得 $l_1(\alpha+\beta)=l_1(\alpha)+l_1(\beta)$，$l_2(\alpha+\beta)=l_2(\alpha)+l_2(\beta)$，$l_3(\alpha+\beta)=l_3(\alpha)+l_3(\beta)$. □

定义 2.2.2：设 V 是向量空间，函数

$$f: V\times V \longrightarrow \mathbb{R}, \quad (\alpha,\beta)\longmapsto f(\alpha,\beta)$$

称为双线性函数，如果，$\forall\ \alpha,\beta,\nu\in V$，$\lambda\in\mathbb{R}$，有

(1) $f(\alpha+\beta,\nu)=f(\alpha,\nu)+f(\beta,\nu)$，$f(\lambda\alpha,\beta)=\lambda f(\alpha,\beta)$

(2) $f(\alpha,\beta+\nu)=f(\alpha,\beta)+f(\alpha,\nu)$，$f(\alpha,\lambda\beta)=\lambda f(\alpha,\beta)$.

如果双线性函数 f 还满足：$f(\alpha,\beta)=f(\beta,\alpha)$.

则称 f 是对称双线性函数.

$\Box$

物理中，如果一个力 α 使物体产生的位移是 β，则

$$W:=|\alpha||\beta|\cos\theta$$

是力 α 作的功，其中 θ 是 α 和 β 的夹角。

定义 2.2.3：对任意两个向量 $\alpha,\beta\in V$，实数

$$\alpha\cdot\beta:=|\alpha|\cdot|\beta|\cos\langle\alpha,\beta\rangle$$

称为 α,β 的内积，其中 $\langle\alpha,\beta\rangle$ 表示向量 α 与 β 之间的夹角（当 α,β 中有一个零向量时，直接定义 $\alpha\cdot\beta=0$）.

$\Box$

定理 2.2.1：内积 $f: V\times V\to\mathbb{R}$，$f(\alpha,\beta)=\alpha\cdot\beta$ 是一个对称，双线性函数，且 $f(\alpha,\alpha)=\alpha\cdot\alpha\geq 0$，等号成立的充要条件是 $\alpha=0$.

证明：只需证明 f 是双线性的. 其它结论由定义是显然的.
由于 f 是对称的，故只需证明：对任意固定的 $\beta\in V$，函数 $f(\cdot,\beta): V\longrightarrow\mathbb{R}$，$\alpha\longmapsto f(\alpha,\beta)$ 是线性函数.

如果 $\beta=0$，函数 $f(\cdot,\beta): V\to\mathbb{R}$ 是零函数，显然是线性函数。
如果 $\beta\neq0$，设 L 是由 β 确定的直线（见例2.2.1），则任意非零向量 α 在 L 上的投影 $\alpha_1=|\alpha|\cos\langle\alpha,\beta\rangle\cdot\frac{\beta}{|\beta|}$，即

$$P_\beta(\alpha)=\frac{|\alpha|\cos\langle\alpha,\beta\rangle}{|\beta|}\cdot\beta=\frac{f(\alpha,\beta)}{|\beta|}\cdot\beta.$$

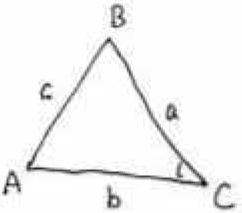
由于 $P_\beta: V\to V$ 是线性映射，所以

$$P_\beta(\alpha+\mu)=\frac{f(\alpha+\mu,\beta)}{|\beta|}\cdot\beta=P_\beta(\alpha)+P_\beta(\mu)=\left(\frac{f(\alpha,\beta)}{|\beta|}+\frac{f(\mu,\beta)}{|\beta|}\right)\beta.$$

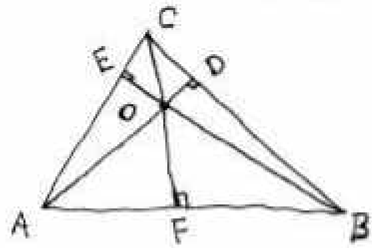
从而 $f(\alpha+\mu,\beta)=f(\alpha,\beta)+f(\mu,\beta)$。同样 $f(\lambda\alpha,\beta)=\lambda f(\alpha,\beta)$。 □

注记：由定义可知，$|\alpha|=\sqrt{\alpha\cdot\alpha}$，$\cos\langle\alpha,\beta\rangle=\frac{\alpha\cdot\beta}{|\alpha|\cdot|\beta|}$。
所以内积可用计算长度，角度等。所以上述的双线性性质对它的应用尤其重要。

例2.2.3（余弦定理）设三角形 $\triangle ABC$ 的三边长分别是 a,b,c（见图2.2）



（图2.2）



（图2.3）

则 $c^2=a^2+b^2-2ab\cos\angle C$。

证明：由 $\overrightarrow{AB}=\overrightarrow{AC}+\overrightarrow{CB}$（用 $\overrightarrow{AB}^2$ 表示内积 $\overrightarrow{AB}\cdot\overrightarrow{AB}$）

得

$$\begin{aligned}c^2=\overrightarrow{AB}\cdot\overrightarrow{AB}&=(\overrightarrow{AC}+\overrightarrow{CB})\cdot(\overrightarrow{AC}+\overrightarrow{CB})\\&=\overrightarrow{AC}\cdot\overrightarrow{AC}+2\overrightarrow{AC}\cdot\overrightarrow{CB}+\overrightarrow{CB}\cdot\overrightarrow{CB}=b^2+a^2+2ab\cos\langle\overrightarrow{AC},\overrightarrow{CB}\rangle\\&=a^2+b^2-2ab\cos\angle C.\end{aligned}$$

□

例2.2.4 证明三角形的三条高线交于一点。

证明：设 AD, BE 分别是三角形 $\triangle ABC$ 的两条高线(见图2.3)。设 O 是 AD 与 BE 的交点，F 是过 C, O 直线与 AB 的交点。则只需证明 $CF\perp AB$，即 $\overrightarrow{CO}\cdot\overrightarrow{AB}=0$。由 $\overrightarrow{CO}=\overrightarrow{CA}+\overrightarrow{AO}$，计算

$$\begin{aligned}\overrightarrow{CO}\cdot\overrightarrow{AB}&=\overrightarrow{CA}\cdot\overrightarrow{AB}+\overrightarrow{AO}\cdot\overrightarrow{AB}=\overrightarrow{CA}\cdot(\overrightarrow{AO}+\overrightarrow{OB})+\overrightarrow{AO}\cdot(\overrightarrow{AC}+\overrightarrow{CB})\\&=\overrightarrow{CA}\cdot\overrightarrow{AO}+\overrightarrow{CA}\cdot\overrightarrow{OB}+\overrightarrow{AO}\cdot\overrightarrow{AC}+\overrightarrow{AO}\cdot\overrightarrow{CB}.\end{aligned}$$

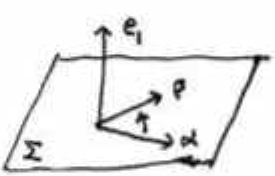
由于 $\overrightarrow{CA}\cdot\overrightarrow{OB}=0$，$\overrightarrow{AO}\cdot\overrightarrow{CB}=0$ ($CA\perp BE$, $CB\perp AD$)，得

$$\overrightarrow{CO}\cdot\overrightarrow{AB}=\overrightarrow{CA}\cdot\overrightarrow{AO}+\overrightarrow{AO}\cdot\overrightarrow{AC}=\overrightarrow{CA}\cdot\overrightarrow{AO}-\overrightarrow{AO}\cdot\overrightarrow{CA}=0.$$

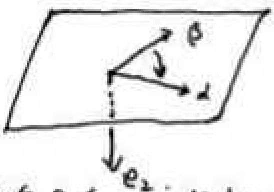
□

对于空间中的一个平面(或曲面Σ)往往需要区分平面(或曲面)的两个侧面，即需要给它定向。我们知道不是所有的曲面都可以定向，下面仅介绍平面的一种定向方法。

在平面上取定两个不共线的向量 α,β。如果规定了它们的先后顺序，则从第一个向量到第二个向量的转角小于 π 的旋转方向伸开右手四指时，拇指的指向就确定了平面的一个方向，它与平面垂直。如图.



(图2.4)



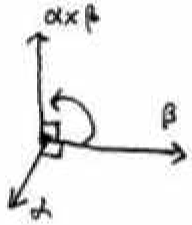
(图2.5)

通常称(图2.4) (α,β,e_1) 成右手系。例如，在空间取直角坐标系 e_1, e_2, e_3 时，我们通常要求它们成右手系。

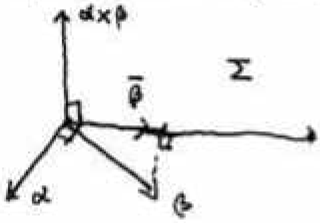
定义2.2.4. 设 $\alpha,\beta\in V$ 是两个向量，则向量 $\alpha\times\beta\in V$ 定义如下：(1) 它的模规定为 $|\alpha\times\beta|:=|\alpha|\cdot|\beta|\sin\langle\alpha,\beta\rangle$.

(2) 当 $|\alpha\times\beta|\neq0$ 时，它的方向规定为：与 α,β 均垂直，且使 $(\alpha,\beta,\alpha\times\beta)$ 成右手系。$\alpha\times\beta$ 称为 α 与 β 的外积。 □

由外积 $\alpha\times\beta$ 的定义可知：当 $|\alpha|=1$，$\langle\alpha,\beta\rangle=90^\circ$（即 α,β 垂直）时，$\alpha\times\beta$ 等于 β 绕 α 旋转 90°.（如图2.6）：



（图2.6）.



（图2.7）

如果 β 与 α 不垂直，令 $\bar\beta$ 表示 β 在与 α 垂直平面 Σ 上的投影，则 $|\bar\beta|=|\beta|\cos\langle\beta,\bar\beta\rangle=|\beta|\sin\langle\alpha,\beta\rangle=|\alpha||\beta|\sin\langle\alpha,\beta\rangle=|\alpha\times\beta|$.
所以 $\alpha\times\beta$ 等于 $\bar\beta$ 绕 α 旋转 90°。

<u>引理2.2.2</u>：设 $\alpha,\beta\in V$，$\bar\beta$ 表示 β 在与 α 垂直平面 Σ 上的投影。则

(1) $\alpha\times\beta=-\beta\times\alpha$；

(2) $|\alpha\times\beta|=0$ 当且仅当 α 与 β 平行。如果 $|\alpha\times\beta|\neq0$，则它是 α,β 所夹平行四边形的面积；

(3) $\alpha\times\beta=\alpha\times\bar\beta$，且 $\alpha\times\beta$ 就是 <u>$\bar\beta$ 绕 α 旋转 90°</u> 所得到向量的 $|\alpha|$ 倍。

证明：(1)，(2) 是定义的直接推论。

(3). 因为 $\alpha\times\beta$ 与 $\alpha\times\bar\beta$，以及"$\bar\beta$ 绕 α 旋转 90° 所得向量"的方向均相同，故只需比较它们的长度即得结论(3)。

$$|\alpha\times\bar\beta|=|\alpha||\bar\beta|\sin\langle\alpha,\bar\beta\rangle=|\alpha|\cdot|\bar\beta|=|\alpha|\cdot|\beta|\sin\langle\alpha,\beta\rangle=|\alpha\times\beta|.$$

□.

外积最重要的性质是它满足下述的双线性性质。

定理2.2.2：向量叉积定义的映射

$$f: V\times V \longrightarrow V,\quad f(\alpha,\beta):=\alpha\times\beta$$

是V上的一个反对称、双线性映射，即

(1) $f(\alpha,\beta)=-f(\beta,\alpha),\ \forall\ \alpha,\beta\in V.$

(2) $f(\lambda\alpha,\beta)=f(\alpha,\lambda\beta)=\lambda f(\alpha,\beta),\ \forall\ \lambda\in\mathbb{R}.$

(3) $f(\alpha,\beta_1+\beta_2)=f(\alpha,\beta_1)+f(\alpha,\beta_2).\ \forall\ \beta_1,\beta_2\in V.$

$f(\alpha_1+\alpha_2,\beta)=f(\alpha_1,\beta)+f(\alpha_1,\beta),\ \forall\alpha_1,\alpha_2\in V.$

证明：(1) 由定义可得，只需证明 (2) 和 (3)。由于 f 的反对称性，我们只需证明：$\forall\ \lambda\in\mathbb{R},\ \alpha,\beta,\beta_1,\beta_2\in V$ 有 $f(\alpha,\lambda\beta)=\lambda f(\alpha,\beta)$，及 $f(\alpha,\beta_1+\beta_2)=f(\alpha,\beta_1)+f(\alpha,\beta_2)$。

(2) 为证 $f(\alpha,\lambda\beta)=\lambda f(\alpha,\beta)$，可设 $\alpha\neq 0$。则

$$f(\alpha,\lambda\beta)=\alpha\times(\lambda\beta)=\alpha\times\overline{\lambda\beta}=\alpha\times\lambda\bar{\beta}$$

是"$\lambda\bar\beta$ 绕 α 旋转 90° 所得向量"的 $|\alpha|$ 倍.

而"$\lambda\bar\beta$ 绕 α 旋转 90° 所得向量"等于"$\bar\beta$ 绕 α 旋转 90° 所得向量"的 λ 倍，即 $\frac{\lambda}{|\alpha|}(\alpha\times\beta)=\frac{\lambda}{|\alpha|}f(\alpha,\beta)$.

所以，$f(\alpha,\lambda\beta)=|\alpha|\cdot\frac{\lambda}{|\alpha|}f(\alpha,\beta)=\lambda f(\alpha,\beta)$.

(3) 为证 $f(\alpha,\beta_1+\beta_2)=f(\alpha,\beta_1)+f(\alpha,\beta_2)$，可设 $\alpha\neq 0$.

则 $f(\alpha,\beta_1+\beta_2)=\alpha\times(\beta_1+\beta_2)=\alpha\times\overline{(\beta_1+\beta_2)}=\alpha\times(\bar\beta_1+\bar\beta_2)$.

因此，$f(\alpha,\beta_1+\beta_2)=|\alpha|\cdot(\bar{\beta}_1+\bar{\beta}_2$ 绕 α 旋转 90° 所得向量).

但 $\bar{\beta}_1+\bar{\beta}_2$ 是以 $\bar{\beta}_1,\bar{\beta}_2$ 为夹边平行四边形的对角线，所以 $\bar{\beta}_1+\bar{\beta}_2$ 绕 α 旋转 90° 所得向量 就是"$\bar{\beta}_1$ 绕 α 旋转 90° 所得向量"与"$\bar{\beta}_2$ 绕 α 旋转 90° 所得向量"之和，即：

$$f(\alpha,\beta_1+\beta_2)=\alpha\times\bar{\beta}_1+\alpha\times\bar{\beta}_2=f(\alpha,\beta_1)+f(\alpha,\beta_2)$$

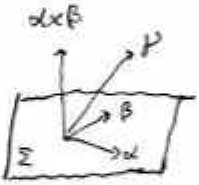
□

定义2.2.5：设 $\alpha,\beta,\gamma\in V$ 是任意三个向量，则 $(\alpha\times\beta)\cdot\gamma\in\mathbb{R}$ 称为 α,β,γ 的混合积，记为

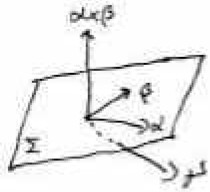
$$D(\alpha,\beta,\gamma):=(\alpha\times\beta)\cdot\gamma.$$

□.

如果 α,β,γ 不共面，设 Σ 是由 α,β 确定的平面.



(图2.8)



(图2.9).

如果 α,β,γ 构成右手系(即 γ 与 $\alpha\times\beta$ 位于同一侧)(见图2.8)，则角 $\langle\gamma,\alpha\times\beta\rangle$ 是锐角，于是

$$D(\alpha,\beta,\gamma)=(\alpha\times\beta)\cdot\gamma=|\alpha\times\beta|\,|\gamma|\cos\langle\gamma,\alpha\times\beta\rangle>0.$$

如果 α,β,γ 构成左手系(即 γ 与 $\alpha\times\beta$ 位于平面 Σ 不同侧面)，则角 $\langle\gamma,\alpha\times\beta\rangle$ 是钝角，于是

$$D(\alpha,\beta,\gamma)=|\alpha\times\beta|\cdot|\gamma|\cos\langle\gamma,\alpha\times\beta\rangle<0.$$

以 α,β,γ 为棱的平行六面体的体积为

$$V=|\alpha\times\beta|\cdot|\gamma|\,|\cos\langle\gamma,\alpha\times\beta\rangle|,$$

因此，当 μ 与 $\alpha\times\beta$ 位于Σ同一侧时，$D(\alpha,\beta,\mu)=V$，而当 μ 与 $\alpha\times\beta$ 位于Σ不同侧时，$D(\alpha,\beta,\mu)=-V$。

混合积的重要性除了上述的几何意义，它还是V上的一个反对称三重线性函数，而这样的函数在相差一个常数的意义下是唯一的。

定义 2.2.6：函数 $f: V\times V\times V\longrightarrow\mathbb{R}$ 称为一个反对称，三重线性函数，如果它满足：

(1) $\varepsilon_\pi f(X_1,X_2,X_3)=f(X_{\pi(1)},X_{\pi(2)},X_{\pi(3)})$，$\forall\pi\in S_3$.

(2) $f(X_1,X_2,X_3)$ 关于每个变量 X_i 都是线性的。

其中 ε_π 表示置换 π 的符号。 □

注记：条件(1)等价于 $f(X_2,X_1,X_3)=-f(X_1,X_2,X_3)$，

$f(X_1,X_3,X_2)=-f(X_1,X_2,X_3)$，$f(X_3,X_2,X_1)=-f(X_1,X_2,X_3)$.

定理 2.2.3：由混合积定义的函数

$$D: V\times V\times V\longrightarrow\mathbb{R},\quad D(\alpha,\beta,\mu):=(\alpha\times\beta)\cdot\mu$$

是一个反对称，三重线性函数。

证明：由于外积，内积关于每个变量都是线性的，所以 $D(\alpha,\beta,\mu)$ 关于每个变量都是线性的。因此，只需证明 $D(\alpha,\beta,\mu)$ 是反对称的。即：

(i) $D(\beta,\alpha,\mu)=-D(\alpha,\beta,\mu)$，(ii) $D(\mu,\beta,\alpha)=-D(\alpha,\beta,\mu)$，(iii) $D(\alpha,\mu,\beta)=-D(\alpha,\beta,\mu)$.

(i) 由定义 $D(\beta,\alpha,\mu)=(\beta\times\alpha)\cdot\mu=-(\alpha\times\beta)\cdot\mu=-D(\alpha,\beta,\mu)$ 可得。

(ii) 和(iii)的证明归结为证明：$(\alpha\times\beta)\cdot\mu=(\beta\times\mu)\cdot\alpha=(\mu\times\alpha)\cdot\beta$.

但是 $|(\alpha\times\beta)\cdot\gamma| = |(\beta\times\gamma)\cdot\alpha| = |(\gamma\times\alpha)\cdot\beta|$ (它们都等于以 α,β,γ 为同一顶点上三条棱的平行六面体的体积), 所以只需确定它们的符号相同。如果 α,β,γ 构成右手系, 则 β,γ,α 和 γ,α,β 也构成右手系 (见图2.8)。而当 α,β,γ 构成左手系时, β,γ,α 和 γ,α,β 也构成左手系. 因此, 它们有相同的符号 (见图2.9).

$\square$

<u>定理2.2.4</u>: 设 $f: V\times V\times V \longrightarrow \mathbb{R}$ 是一个<u>反对称</u>, 三重线性函数, $e_1,e_2,e_3\in V$ 是三个不共面的向量. 设 $\alpha_1,\alpha_2,\alpha_3\in V$ 是任意三个向量. 如果

$$\begin{aligned}\alpha_1 &= a_{11}e_1 + a_{12}e_2 + a_{13}e_3\\ \alpha_2 &= a_{21}e_1 + a_{22}e_2 + a_{23}e_3 \qquad (a_{ij}\in\mathbb{R}.)\\ \alpha_3 &= a_{31}e_1 + a_{32}e_2 + a_{33}e_3\end{aligned}$$

则 $$f(\alpha_1,\alpha_2,\alpha_3) = \Big(\sum_{\pi\in S_3}\varepsilon_\pi\, a_{1\pi(1)}a_{2\pi(2)}a_{3\pi(3)}\Big) f(e_1,e_2,e_3).$$

特别地, 如果 $f(e_1,e_2,e_3) = D(e_1,e_2,e_3)$, 则

$$f(\alpha_1,\alpha_2,\alpha_3) = D(\alpha_1,\alpha_2,\alpha_3), \quad \forall\, \alpha_1,\alpha_2,\alpha_3\in V.$$

证明: 只需根据<u>反对称</u>, 三重线性的定义, 将 $f(\alpha_1,\alpha_2,\alpha_3)$ 展开:

$$\begin{aligned}f(\alpha_1,\alpha_2,\alpha_3) &= f(a_{11}e_1+a_{12}e_2+a_{13}e_3,\ a_{21}e_1+a_{22}e_2+a_{23}e_3,\ a_{31}e_1+a_{32}e_2+a_{33}e_3)\\ &= \sum_{1\le i_1,i_2,i_3\le 3} f(a_{1i_1}e_{i_1},\ a_{2i_2}e_{i_2},\ a_{3i_3}e_{i_3}) = \sum_{1\le i_1,i_2,i_3\le 3} a_{1i_1}a_{2i_2}a_{3i_3} f(e_{i_1},e_{i_2},e_{i_3})\end{aligned}$$

上面仅用到了 f 是<u>三重线性</u>的条件. 当 i_1,i_2,i_3 中有相同的数时, 如果 f 是<u>反对称</u>, 则 $f(e_{i_1},e_{i_2},e_{i_3}) = 0$. 所以可设 i_1,i_2,i_3 是 $1,2,3$ 的一个排列. 且 $f(e_{i_1},e_{i_2},e_{i_3}) = \varepsilon_\pi f(e_1,e_2,e_3)$,

其中 $\pi(1)=i_1$，$\pi(2)=i_2$，$\pi(3)=i_3$. 所以

$$f(\alpha_1,\alpha_2,\alpha_3)=\left(\sum_{\pi\in S_3}\varepsilon_\pi a_{1\pi(1)}a_{2\pi(2)}a_{3\pi(3)}\right)f(e_1,e_2,e_3).$$

□.

<u>定义2.2.7</u>：称如下面的数表

$$A=\begin{pmatrix} a_{11} & a_{12} & a_{13}\\ a_{21} & a_{22} & a_{23}\\ a_{31} & a_{32} & a_{33}\end{pmatrix}\qquad a_{ij}\in\mathbb{R}.$$

称为<u>3阶实系数矩阵</u>. a_{ij} 称为<u>第 i 行</u>，<u>第 j 列</u>位置的元素，通常将它简记为 $A=(a_{ij})_{3\times3}$. 两个矩阵 $A=(a_{ij})_{3\times3}$，$B=(b_{ij})_{3\times3}$ 相等当且仅当 $a_{ij}=b_{ij}$ $(1\le i\le3, 1\le j\le3)$. 其中

$$A_{(i)}=(a_{i1},a_{i2},a_{i3}),\quad A^{(j)}=\begin{pmatrix}a_{1j}\\ a_{2j}\\ a_{3j}\end{pmatrix}:=[a_{1j},a_{2j},a_{3j}]$$

分别称为 A 的第 i 行（或第 i 个行向量），A 的第 j 列（或第 j 个列向量）. 而定理中的实数

$$|A|:=\sum_{\pi\in S_3}\varepsilon_\pi a_{1\pi(1)}a_{2\pi(2)}a_{3\pi(3)}\quad(\text{也用记号 } \det(A) \text{ 表示})$$

则称为 A 的<u>行列式</u>.

□

如果 e_1,e_2,e_3 是单位向量，且两两垂直，则 $D(e_1,e_2,e_3)=1$

<u>推论2.2.1</u>：若在定理2.2.4中设 e_1,e_2,e_3 是两两垂直的单位向量. 如果定理中的 $\alpha_1,\alpha_2,\alpha_3$ 构成右手系，则 $|A|$ 等于以 $\alpha_1,\alpha_2,\alpha_3$ 为三条共顶点棱的<u>平行六面体的体积</u>. 特别地，$|A|=0 \Leftrightarrow \alpha_1,\alpha_2,\alpha_3$ 共面.

□.

例2.2.5：考虑矩阵

$$A=\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix},\qquad {}^tA=\begin{pmatrix} a_{11} & a_{21} & a_{31} \\ a_{12} & a_{22} & a_{32} \\ a_{13} & a_{23} & a_{33} \end{pmatrix}.$$

其中 ${}^tA$ 称为 A 的转置矩阵. 则 $|{}^tA|=|A|$.

证明：令 ${}^tA=(b_{ij})_{3\times3}$. 则 $b_{ij}=a_{ji}$. 计算

$$|{}^tA|=\sum_{\pi\in S_3}\varepsilon_\pi\, b_{1\pi(1)}\, b_{2\pi(2)}\, b_{3\pi(3)}=\sum_{\pi\in S_3}\varepsilon_\pi\, a_{\pi(1)1}\, a_{\pi(2)2}\, a_{\pi(3)3}$$

$$=\sum_{\pi\in S_3}\varepsilon_\pi\, a_{1\pi^{-1}(1)}\, a_{2\pi^{-1}(2)}\, a_{3\pi^{-1}(3)}$$

$$=\sum_{\pi\in S_3}\varepsilon_{\pi^{-1}}\, a_{1\pi^{-1}(1)}\, a_{2\pi^{-1}(2)}\, a_{3\pi^{-1}(3)}=\sum_{\pi^{-1}\in S_3}\varepsilon_{\pi^{-1}}\, a_{1\pi^{-1}(1)}\, a_{2\pi^{-1}(2)}\, a_{3\pi^{-1}(3)}$$

$$=|A|.$$

□

例2.2.6：固定一组线性无关的向量 $e_1, e_2, e_3\in V$. 则存在三个线性函数 $l_i: V\to\mathbb{R}$ $(1\le i\le 3)$（由 e_1, e_2, e_3 唯一确定）使得：$\forall\,\alpha\in V$，有

$$\alpha=l_1(\alpha)e_1+l_2(\alpha)e_2+l_3(\alpha)e_3.$$

当选取 e_1, e_2, e_3 是一组两两垂直的单位向量时（即：$e_i\cdot e_j=0$, $|e_i|=1$）. 则 l_1, l_2, l_3 有如下的明确表达式：

$$l_i(\alpha)=e_i\cdot\alpha,\quad \forall\,\alpha\in V.$$

□

习题 2.2

2.1. 利用内积的对称双线性性质证明：$\forall\ \alpha,\beta\in V$，有

$$\alpha\cdot\beta=\frac{1}{2}\left(|\alpha-\beta|^2-|\alpha|^2-|\beta|^2\right).$$

2.2. 设四面体ABCD的棱 $|AB|=|AC|=2$，$|AD|=1$，$\angle BAC=\angle BAD=\frac{\pi}{3}$，$\angle CAD=\frac{\pi}{6}$，记BC的中点为P，三角形$\Delta BCD$的重心为Q，求：$|\overrightarrow{AP}|$，$|\overrightarrow{AQ}|$ 和 $\overrightarrow{AP}\cdot\overrightarrow{AQ}$.

2.3. 设α,β,γ不共面，δ是任意向量. 证明：

$$\delta=0\iff\delta\cdot\alpha=\delta\cdot\beta=\delta\cdot\beta=0.$$

2.4. 证明：三角形三条中线长度的平方和等于三边长度平方和的$\frac{3}{4}$.

2.5. 对空间中任意四点A，B，C，D. 证明：

(1) $\overrightarrow{AB}\cdot\overrightarrow{CD}+\overrightarrow{BC}\cdot\overrightarrow{AD}+\overrightarrow{CA}\cdot\overrightarrow{BD}=0$.

(2) $\overrightarrow{AB}^2+\overrightarrow{CD}^2=\overrightarrow{AC}^2+\overrightarrow{BD}^2\iff\overrightarrow{AD}\cdot\overrightarrow{BC}=0$，

并利用上述结论证明：如果一个四面体有两对对棱互相垂直，则第三对对棱也互相垂直，并且三对对棱长度的平方和相等.

2.6. 设$A_1,A_2,\cdots,A_n$是正n边形的顶点，证明：

$$\overrightarrow{A_1A_2}+\overrightarrow{A_2A_3}+\cdots+\overrightarrow{A_nA_1}=0.$$

并由此推导：

$$1+\cos\frac{2\pi}{n}+\cos\frac{4\pi}{n}+\cdots+\cos\frac{(2n-2)\pi}{n}=0$$

$$\sin\frac{2\pi}{n}+\sin\frac{4\pi}{n}+\cdots+\sin\frac{(2n-2)\pi}{n}=0$$

2.7. 证明：如果 α,β 与 γ 垂直，则 $(\alpha\times\beta)\times\gamma=0$.

2.8. 如果 β 垂直于 γ，而 α 平行于 γ。证明：$(\alpha\times\beta)\times\gamma=(\alpha\cdot\gamma)\beta$

2.9. 对任意三个向量 $\alpha,\beta,\gamma\in V$，证明：

$$(\alpha\times\beta)\times\gamma=(\alpha\cdot\gamma)\beta-(\beta\cdot\gamma)\alpha .$$

2.10. 证明：$(\alpha-\beta)\times(\alpha+\beta)=2(\alpha\times\beta)$. 并说明该等式的几何意义.

2.11. 证明：三角形的重心与三个顶点的连线分原三角形成三个等面积的三角形.

2.12. 证明：向量 α,β,γ 共面的充要条件是

$$\begin{vmatrix} \alpha\cdot\alpha, & \alpha\cdot\beta, & \alpha\cdot\gamma \\ \beta\cdot\alpha & \beta\cdot\beta & \beta\cdot\gamma \\ \gamma\cdot\alpha & \gamma\cdot\beta & \gamma\cdot\gamma \end{vmatrix}=0.$$

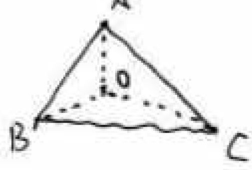
2.13. 证明拉格朗日（Lagrange）恒等式：任取向量 $\alpha_1,\alpha_2,\alpha_3,\alpha_4$，

有
$$(\alpha_1\times\alpha_2)\cdot(\alpha_3\times\alpha_4)=\underline{(\alpha_1\cdot\alpha_3)(\alpha_2\cdot\alpha_4)-(\alpha_2\cdot\alpha_3)(\alpha_2\cdot\alpha_4)}$$

$$:=\begin{vmatrix} \alpha_1\cdot\alpha_3 & \alpha_1\cdot\alpha_4 \\ \alpha_2\cdot\alpha_3 & \alpha_2\cdot\alpha_4 \end{vmatrix}$$ (←二阶行列式的定义).

2.14. 利用拉格朗日恒等式证明：

直角三棱锥（即：侧棱相互垂直）<u>斜面面积</u>的平方等于其它三个<u>直角面面积</u>的平方和.（=几何勾股定理）. 如图：



其中 $\angle AOB=\angle AOC=\angle BOC=90^\circ$，$\triangle ABC$ 是它的斜面.

§2.3. 基与坐标.

如果固定一组不共面的向量 $e_1, e_2, e_3 \in V$，则V中每一个向量可唯一地表示成 e_1, e_2, e_3 的线性组合。我们称 e_1, e_2, e_3 是 V 的一组基。由前面的讨论可知，e_1, e_2, e_3 确定了三个线性函数：$l_i : V \to \mathbb{R}$ $(i=1,2,3)$ 使得

$$\alpha = l_1(\alpha) e_1 + l_2(\alpha) e_2 + l_3(\alpha) e_3 \quad , \forall \alpha \in V.$$

为了更好地体现 l_1, l_2, l_3 由 e_1, e_2, e_3 确定，以后我们将用符号 e_1^*, e_2^*, e_3^* 表示线性函数 l_1, l_2, l_3. 即

$$\alpha = e_1^*(\alpha)\, e_1 + e_2^*(\alpha)\, e_2 + e_3^*(\alpha)\, e_3 \ , \ \forall \alpha \in V.$$

定义2.3.1：数组 $(e_1^*(\alpha), e_2^*(\alpha), e_3^*(\alpha)) \in \mathbb{R}^3$ 称为 α 在基 e_1, e_2, e_3 下的坐标向量，根据需要，有时也将坐标向量写成列向量的形式：

$$[e_1^*(\alpha), e_2^*(\alpha), e_3^*(\alpha)] := \begin{pmatrix} e_1^*(\alpha) \\ e_2^*(\alpha) \\ e_3^*(\alpha) \end{pmatrix} \in \mathbb{R}^3 .$$

□.

由于 $e_i^* : V \to \mathbb{R}$ 是线性函数，所以对任意 $\lambda \in \mathbb{R}, \alpha, \beta \in V$，有 $\lambda\alpha$，$\alpha+\beta$ 的坐标向量分别是：

$$(\lambda e_1^*(\alpha), \lambda e_2^*(\alpha), \lambda e_3^*(\alpha)), \quad (e_1^*(\alpha)+e_1^*(\beta), e_2^*(\alpha)+e_2^*(\beta), e_3^*(\alpha)+e_3^*(\beta)).$$

定义2.3.2：在 $\mathbb{R}^3$ 中定义：$\forall \lambda \in \mathbb{R}$，$x=(x_1, x_2, x_3)$，$y=(y_1, y_2, y_3) \in \mathbb{R}^3$，令

$$\lambda x = (\lambda x_1, \lambda x_2, \lambda x_3) , \quad x+y = (x_1+y_1, x_2+y_2, x_3+y_3).$$

我们称带有上述运算的 $\mathbb{R}^3$ 为坐标空间.

□.

推论2.3.1：映射 $e : V \longrightarrow \mathbb{R}^3$，$\alpha \longmapsto (e_1^*(\alpha), e_2^*(\alpha), e_3^*(\alpha))$ 是一个双射，且保持运算：$e(\lambda\alpha) = \lambda e(\alpha)$，$e(\alpha+\beta) = e(\alpha)+e(\beta)$

□.

通过选定 V 的一组基 e_1, e_2, e_3，不仅 V 中向量有坐标，甚至 V 上的线性函数，双线性函数，线性映射等都可以有"坐标"，而且它们的运算（比如，函数的加法，映射的合成）都可以通过"坐标的运算"来表述。

例 2.3.1：设 $l: V \to \mathbb{R}$ 是一个线性函数，e_1, e_2, e_3 是 V 的一组基。则 l 由它在 e_1, e_2, e_3 的取值 $a_1 = l(e_1)$, $a_2 = l(e_2)$, $a_3 = l(e_3)$ 唯一确定：$\forall\ x = x_1e_1 + x_2e_2 + x_3e_3 \in V$，有

$$l(x) = a_1x_1 + a_2x_2 + a_3x_3 .$$

所以，(a_1, a_2, a_3) 可以看成 l 在基 e_1, e_2, e_3 下的坐标。 □

例 2.3.2：设 $f: V \times V \to \mathbb{R}$ 是一个双线性函数，e_1, e_2, e_3 是 V 的一组基，则函数 f 由它在 e_1, e_2, e_3 的取值矩阵

$$B(f) := \begin{pmatrix} f(e_1,e_1), & f(e_1,e_2), & f(e_1,e_3) \\ f(e_2,e_1), & f(e_2,e_2), & f(e_2,e_3) \\ f(e_3,e_1), & f(e_3,e_2), & f(e_3,e_3) \end{pmatrix}$$

唯一确定：$\forall\ x = x_1e_1 + x_2e_2 + x_3e_3 \in V$，$y = y_1e_1 + y_2e_2 + y_3e_3 \in V$，有

$$f(x,y) = f(e_1,e_1)x_1y_1 + f(e_2,e_2)x_2y_2 + f(e_3,e_3)x_3y_3 + \sum_{i \neq j} f(e_i,e_j)x_iy_j .$$

所以，矩阵 $B(f)$ 可以看成 f 在基 e_1, e_2, e_3 下的坐标 □.

例 2.3.3：设 $A: V \to V$ 是一个线性映射，e_1, e_2, e_3 是 V 的一组基，则 A 由 e_1, e_2, e_3 在它之下的像 $Ae_1, Ae_2, Ae_3 \in V$ 唯一确定：$\forall\ x = x_1e_1 + x_2e_2 + x_3e_3 \in V$，$Ax = x_1Ae_1 + x_2Ae_2 + x_3Ae_3$。而 $Ae_j = e_1^*(Ae_j)\,e_1 + e_2^*(Ae_j)\,e_2 + e_3^*(Ae_j)\,e_3$ $(j=1,2,3)$ 由它的坐标向量 $(e_1^*(Ae_j), e_2^*(Ae_j), e_3^*(Ae_j))$ 唯一确定。所以，矩阵

$$e(A) := \begin{pmatrix} e_1^*(Ae_1) & e_1^*(Ae_2) & e_1^*(Ae_3) \\ e_2^*(Ae_1) & e_2^*(Ae_2) & e_2^*(Ae_3) \\ e_3^*(Ae_1) & e_3^*(Ae_2) & e_3^*(Ae_3) \end{pmatrix}$$

可以看成线性映射 $\mathcal{A}: V \to V$ 在基 $e:=(e_1,e_2,e_3)$ 下的坐标。习惯记为:

$$(\mathcal{A}e_1, \mathcal{A}e_2, \mathcal{A}e_3) = (e_1, e_2, e_3)\, e(\mathcal{A}) = (e_1,e_2,e_3)\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$$

其中 $a_{ij} = e_i^*(\mathcal{A}e_j)$。

$\square$.

设 $\mathbb{A}(V)$ 表示我们现实中的三维空间. 如果选取 $\mathbb{A}(V)$ 中的一个点 O; 则 $\mathbb{A}(V)$ 中的点与 V 中的向量有如下 1-1 对应:

$$\mathbb{A}(V) \xrightarrow{\Phi} V, \quad P \longmapsto \overrightarrow{OP}.$$

如果进一步固定 V 中的一组基 e_1, e_2, e_3. 则 $\forall\, P \in \mathbb{A}(V)$, 存在唯一的实数 $x_i(P) := e_i^*(\overrightarrow{OP})$ 使 $\overrightarrow{OP} = x_1(P)e_1 + x_2(P)e_2 + x_3(P)e_3$.

定义 2.3.3: (O, e_1, e_2, e_3) 称为 $\mathbb{A}(V)$ 的一个以 O 为原点的仿射坐标架, 上述的坐标向量 $(x_1(P), x_2(P), x_3(P))$ 称为点 P 在坐标架 (O, e_1, e_2, e_3) 下的坐标. 空间 $\mathbb{A}(V)$ 称为与 V 相伴的点空间。如果 e_1, e_2, e_3 是互相垂直的单位向量, 则 (O, e_1, e_2, e_3) 称为一个直角坐标架, P 点在它下的坐标称为 P 的直角坐标。

$\square$.

选定了一个坐标架 (O, e_1, e_2, e_3) 之后, 就在点空间 $\mathbb{A}(V)$, 向量空间 V 和坐标空间 $\mathbb{R}^3$ 之间建立了如下双射:

$$\mathbb{A}(V) \xrightarrow{\Phi} V \xrightarrow{e} \mathbb{R}^3.$$

$$\Phi(P) := \overrightarrow{OP}, \quad e(\overrightarrow{OP}) = (e_1^*(\overrightarrow{OP}), e_2^*(\overrightarrow{OP}), e_3^*(\overrightarrow{OP})) := (x_1(P), x_2(P), x_3(P)).$$

设 $H \subset \mathbb{A}(V)$ 是一个几何图形的点集, 它在 V 和 $\mathbb{R}^3$ 的像分别是:

$$V(H) := \{\overrightarrow{OP} \mid P \in H\} \subset V.$$

将在下节 $$C(H) := \{(x_1(P), x_2(P), x_3(P)) \in \mathbb{R}^3 \mid P \in H\} \subset \mathbb{R}^3.$$

我们讨论如下问题: (1) 刻划子集 $V(H)$ 和 $C(H)$ 的条件是什么?

(2) 如何利用 $V(H)$, $C(H)$ 满足的代数条件研究 $H\subset \mathbb{A}(V)$ 的几何？

在讨论上述问题时，我们通常选取直角坐标架 (O, e_1, e_2, e_3). 这样选取的唯一理由是可以简化问题的表述和计算. 例如，此时 V 中的内积可以用向量坐标表示成

$$\alpha\cdot\beta = e_1^*(\alpha)e_1^*(\beta)+e_2^*(\alpha)e_2^*(\beta)+e_3^*(\alpha)e_3^*(\beta).$$

而 $\mathbb{A}(V)$ 中任意两点 P,Q 的距离公式为:

$$\rho(P,Q) := |\overrightarrow{PQ}| = |\overrightarrow{OQ}-\overrightarrow{OP}| = \sqrt{(y_1-x_1)^2+(y_2-x_2)^2+(y_3-x_3)^2}$$

其中 $x_i = e_i^*(\overrightarrow{OP})$, $y_i = e_i^*(\overrightarrow{OQ})$ $(i=1,2,3)$ 是 P,Q 的坐标.

例2.3.4: 设 e_1, e_2, e_3 是互相垂直的单位向量，并形成右手系. (即: $e_3 = e_1\times e_2$). 则，$\forall\ \alpha,\beta\in V$，有

$$\alpha\times\beta = \begin{vmatrix} e_2^*(\alpha) & e_3^*(\alpha) \\ e_2^*(\beta) & e_3^*(\beta)\end{vmatrix} e_1 + \begin{vmatrix} e_3^*(\alpha) & e_1^*(\alpha) \\ e_3^*(\beta) & e_1^*(\beta)\end{vmatrix} e_2 + \begin{vmatrix} e_1^*(\alpha) & e_2^*(\alpha) \\ e_1^*(\beta) & e_2^*(\beta)\end{vmatrix} e_3$$

其中符号 $\begin{vmatrix} a & b \\ c & d\end{vmatrix}$ 称为二阶行列式，定义为

$$\begin{vmatrix} a & b \\ c & d\end{vmatrix} := ad - cb.$$

证明: 由叉积的反对称性, 双线性性质可得.

$$\alpha\times\beta = (e_1^*(\alpha)e_1 + e_2^*(\alpha)e_2 + e_3^*(\alpha)e_3)\times(e_1^*(\beta)e_1 + e_2^*(\beta)e_2 + e_3^*(\beta)e_3)$$

$$= e_2^*(\alpha)e_1^*(\beta)\, e_2\times e_1 + e_3^*(\alpha)e_1^*(\beta)\, e_3\times e_1 + e_1^*(\alpha)e_2^*(\beta)\, e_1\times e_2 + e_3^*(\alpha)e_2^*(\beta)\, e_3\times e_2 + e_1^*(\alpha)e_3^*(\beta)\, e_1\times e_3 + e_2^*(\alpha)e_3^*(\beta)\, e_2\times e_3$$

$$= (e_2^*(\alpha)e_3^*(\beta) - e_2^*(\beta)e_3^*(\alpha))\, e_2\times e_3 + (e_3^*(\alpha)e_1^*(\beta) - e_3^*(\beta)e_1^*(\alpha))\, e_3\times e_1 + \underbrace{(e_1^*(\alpha)e_2^*(\beta) - e_1^*(\beta)e_2^*(\alpha))}\, e_1\times e_2$$

$$= \begin{vmatrix} e_2^*(\alpha) & e_3^*(\alpha) \\ e_2^*(\beta) & e_3^*(\beta)\end{vmatrix} e_2\times e_3 + \begin{vmatrix} e_3^*(\alpha) & e_1^*(\alpha) \\ e_3^*(\beta) & e_1^*(\beta)\end{vmatrix} e_3\times e_1 + \begin{vmatrix} e_1^*(\alpha) & e_2^*(\alpha) \\ e_1^*(\beta) & e_2^*(\beta)\end{vmatrix} e_1\times e_2.$$

又 $e_2\times e_3=e_1$，$e_3\times e_1=e_2$，$e_1\times e_2=e_3$（因为 e_1,e_2,e_3 构成右手系,且是相互垂直的单位向量），可得向量 $\alpha\times\beta$ 的坐标为：

$$\left(\begin{vmatrix} e_2^*(\alpha) & e_3^*(\alpha) \\ e_2^*(\beta) & e_3^*(\beta)\end{vmatrix},\quad \begin{vmatrix} e_3^*(\alpha) & e_1^*(\alpha) \\ e_3^*(\beta) & e_1^*(\beta)\end{vmatrix},\quad \begin{vmatrix} e_1^*(\alpha) & e_2^*(\alpha) \\ e_1^*(\beta) & e_2^*(\beta)\end{vmatrix}\right).$$

□

例2.3.5：设 e_1,e_2,e_3 是相互垂直的单位向量，且构成右手系（即：$e_3=e_1\times e_2$，或等价地，$(e_1\times e_2)\cdot e_3=D(e_1,e_2,e_3)=1$）。则

$$\begin{vmatrix} e_1^*(\alpha) & e_2^*(\alpha) & e_3^*(\alpha) \\ e_1^*(\beta) & e_2^*(\beta) & e_3^*(\beta) \\ e_1^*(\gamma) & e_2^*(\gamma) & e_3^*(\gamma)\end{vmatrix}=D(\alpha,\beta,\gamma)=(\alpha\times\beta)\cdot\gamma=\begin{vmatrix} e_2^*(\alpha) & e_3^*(\alpha) \\ e_2^*(\beta) & e_3^*(\beta)\end{vmatrix}e_1^*(\gamma)+\begin{vmatrix} e_3^*(\alpha) & e_1^*(\alpha) \\ e_3^*(\beta) & e_1^*(\beta)\end{vmatrix}e_2^*(\gamma)$$
$$+\begin{vmatrix} e_1^*(\alpha) & e_2^*(\alpha) \\ e_1^*(\beta) & e_2^*(\beta)\end{vmatrix}e_3^*(\gamma).$$

□

例2.3.6：设点 A_1,A_2,A_3 在直角坐标架 (O,e_1,e_2,e_3) 的坐标分别是 (x_1,x_2,x_3)，(y_1,y_2,y_3)，(z_1,z_2,z_3)。则以 A_1,A_2,A_3 为顶点的三角形 $\Delta A_1A_2A_3$ 的面积公式为：

$$S=\frac{1}{2}\sqrt{\begin{vmatrix} y_2-x_2 & y_3-x_3 \\ z_2-x_2 & z_3-x_3\end{vmatrix}^2+\begin{vmatrix} y_3-x_3 & y_1-x_1 \\ z_3-x_3 & z_1-x_1\end{vmatrix}^2+\begin{vmatrix} y_1-x_1 & y_2-x_2 \\ z_1-x_1 & z_2-x_2\end{vmatrix}^2}$$

证明：$|\overrightarrow{A_1A_2}\times\overrightarrow{A_1A_3}|$ 是以 A_1A_2，A_1A_3 为夹边平行四边形的面积。

所以 $S=\frac{1}{2}|\overrightarrow{A_1A_2}\times\overrightarrow{A_1A_3}|$。而由条件知：

$$\overrightarrow{A_1A_2}=\overrightarrow{OA_2}-\overrightarrow{OA_1}=(y_1-x_1)e_1+(y_2-x_2)e_2+(y_3-x_3)e_3$$
$$\overrightarrow{A_1A_3}=\overrightarrow{OA_3}-\overrightarrow{OA_1}=(z_1-x_1)e_1+(z_2-x_2)e_2+(z_3-x_3)e_3.$$

利用例2.3.4中关于 $\overrightarrow{A_1A_2}\times\overrightarrow{A_1A_3}$ 的坐标公式，可得 $|\overrightarrow{A_1A_2}\times\overrightarrow{A_1A_3}|$。

□

例 2.3.7：设点 A_1, A_2, A_3, A_4 在直角坐标架 (O, e_1, e_2, e_3) 下的坐标分别是 (x_1, x_2, x_3)，(y_1, y_2, y_3)，(z_1, z_2, z_3)，(w_1, w_2, w_3)。则以 A_1, A_2, A_3, A_4 为顶点的四面体的体积公式为：

$$V^2 = \frac{1}{36}\begin{vmatrix} y_1-x_1 & y_2-x_2 & y_3-x_3 \\ z_1-x_1 & z_2-x_2 & z_3-x_3 \\ w_1-x_1 & w_2-x_2 & w_3-x_3 \end{vmatrix}^2$$

证明：$|(\overrightarrow{A_1A_2}\times\overrightarrow{A_1A_3})\cdot\overrightarrow{A_1A_4}|$ 等于以线段 A_1A_2，A_1A_3，A_1A_4 为共顶点棱构成的平行六面体的体积。所以

$$V = \frac{1}{6}|(\overrightarrow{A_1A_2}\times\overrightarrow{A_1A_3})\cdot\overrightarrow{A_1A_4}|.$$

由已知条件得：

$$\overrightarrow{A_1A_2} = \overrightarrow{OA_2}-\overrightarrow{OA_1} = (y_1-x_1)e_1+(y_2-x_2)e_2+(y_3-x_3)e_3$$

$$\overrightarrow{A_1A_3} = \overrightarrow{OA_3}-\overrightarrow{OA_1} = (z_1-x_1)e_1+(z_2-x_2)e_2+(z_3-x_3)e_3$$

$$\overrightarrow{A_1A_4} = \overrightarrow{OA_4}-\overrightarrow{OA_1} = (w_1-x_1)e_1+(w_2-x_2)e_2+(w_3-x_3)e_3.$$

并代入例 2.3.5 中的混合积公式即得

$$(\overrightarrow{A_1A_2}\times\overrightarrow{A_1A_3})\cdot\overrightarrow{A_1A_4} = \begin{vmatrix} y_1-x_1 & y_2-x_2 & y_3-x_3 \\ z_1-x_1 & z_2-x_2 & z_3-x_3 \\ w_1-x_1 & w_2-x_2 & w_3-x_3 \end{vmatrix}$$

□

在本节最后，我们讨论 $\mathbb{A}(V)$ 中的点在不同仿射坐标架 (O, e_1, e_2, e_3) 和 (O', e_1', e_2', e_3') 下坐标之间的关系。

设点 A 在 (O, e_1, e_2, e_3) 下的坐标向量是 (x_1, x_2, x_3)，而在 (O, e_1', e_2', e_3') 下的坐标向量是 (x_1', x_2', x_3')。即

$$\overrightarrow{OA} = x_1e_1+x_2e_2+x_3e_3, \quad \overrightarrow{O'A} = x_1'e_1'+x_2'e_2'+x_3'e_3'.$$

下面的定理给出了 (x_1', x_2', x_3') 与 (x_1, x_2, x_3) 的关系。

定理2.3.1：设 $(e_1',e_2',e_3')=(e_1,e_2,e_3)A$，$A=(a_{ij})_{3\times3}$，

$\overrightarrow{OO'}=b_1e_1+b_2e_2+b_3e_3$. 则

$$\begin{cases} x_1=a_{11}x_1'+a_{12}x_2'+a_{13}x_3'+b_1 \\ x_2=a_{21}x_1'+a_{22}x_2'+a_{23}x_3'+b_2 \\ x_3=a_{31}x_1'+a_{32}x_2'+a_{33}x_3'+b_3 \end{cases}$$ （坐标变换公式）.

证明：由 $\overrightarrow{OA}=\overrightarrow{OO'}+\overrightarrow{O'A}$，得

$$x_1e_1+x_2e_2+x_3e_3=b_1e_1+b_2e_2+b_3e_3+x_1'e_1'+x_2'e_2'+x_3'e_3'.$$

将 $(e_1',e_2',e_3')=(e_1,e_2,e_3)A$（即 $e_i'=a_{1i}e_1+a_{2i}e_2+a_{3i}e_3$）代入得

$$x_1e_1+x_2e_2+x_3e_3=(b_1+a_{11}x_1'+a_{12}x_2'+a_{13}x_3')e_1+(b_2+a_{21}x_1'+a_{22}x_2'+a_{23}x_3')e_2$$
$$+(b_3+a_{31}x_1'+a_{32}x_2'+a_{33}x_3')e_3.$$

比较系数，即得坐标变换公式.

□.

习题 2.3

3.1. 设 e_1,e_2,e_3 是 V 的一组基. 设 $\mathcal{A}:V\to V$，$\mathcal{B}:V\to V$ 是两个线性映射. 证明：$\mathcal{A},\mathcal{B}$ 的合成映射

$$\mathcal{C}:=\mathcal{A}\cdot\mathcal{B}:V\to V$$

也是线性映射。如果 $\mathcal{A},\mathcal{B}$ 在基 e_1,e_2,e_3 下的坐标矩阵分别是 A,B. 试求 $\mathcal{A}\mathcal{B}$ 在基 e_1,e_2,e_3 下的坐标矩阵。

3.2. 在定理2.3.1中，如果 (O,e_1,e_2,e_3)，(O',e_1',e_2',e_3') 都是直角坐标架，请问矩阵 $A=(a_{ij})_{3\times3}$ 必须满足什么条件？

3.3. 设 (O,e_1,e_2,e_3) 是 $\mathbb{A}(V)$ 的一个仿射坐标架，$A_1,A_2\in\mathbb{A}(V)$ 是坐标分别是 (x_1,x_2,x_3)，(y_1,y_2,y_3) 的两个点。设 A 是线段 A_1A_2 中的点使得 A_1A 与 AA_2 的长度之比为 λ_1/λ_2. 试求点 A 在 (O,e_1,e_2,e_3) 下的坐标.

3.4. 设 M 是点组 $A_1, A_2, \cdots, A_n$ 的重心。A_i 的坐标为 (x_i, y_i, z_i) $(1 \le i \le n)$. 试求点 M 的坐标.

3.5. 设点 A, B, C 在 (O, e_1, e_2, e_3) 下的坐标分别是 (x_1, x_2, x_3), (y_1, y_2, y_3), (z_1, z_2, z_3). 证明：如果 A, B, C 共线，则

$$\begin{vmatrix} x_1 & x_2 & x_3 \\ y_1 & y_2 & y_3 \\ z_1 & z_2 & z_3 \end{vmatrix} = 0.$$

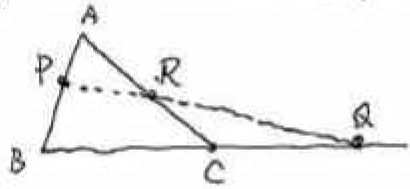
3.6. 证明定理2.3.1中坐标变换矩阵 A 的行列式不等于0.

3.7. 设 e_1, e_2, e_3 是V的一组基，$(e_1', e_2', e_3') = (e_1, e_2, e_3)A$. 证明：
e_1', e_2', e_3' 是一组基 $\Longleftrightarrow$ $|A| \ne 0$.

3.8. 设点 A, B, C 的坐标分别是 (x_1, x_2, x_3), (y_1, y_2, y_3), (z_1, z_2, z_3)
则 A, B, C 共线的充要条件是：

$$\begin{vmatrix} y_2-x_2 & y_3-x_3 \\ z_2-x_2 & z_3-x_3 \end{vmatrix} = \begin{vmatrix} y_3-x_3 & y_1-x_1 \\ z_3-x_3 & z_1-x_1 \end{vmatrix} = \begin{vmatrix} y_1-x_1 & y_2-x_2 \\ z_1-x_1 & z_2-x_2 \end{vmatrix} = 0.$$

3.9. (Menelaus 定理)：设点 P, Q, R 分别位于三角形 $\triangle ABC$ 的三条边所在的直线上（如图）：



如果 $\overrightarrow{AP} = \lambda \overrightarrow{PB}$, $\overrightarrow{BQ} = \mu \overrightarrow{QC}$, $\overrightarrow{CR} = \nu \overrightarrow{RA}$, 证明：

$$P, R, Q \text{ 共线} \Longleftrightarrow \lambda\mu\nu = -1.$$

（提示：适当选取坐标架 (O, e_1, e_2, e_3) 使 A, B, C 的坐标分别是 $(0,0,0)$, $(1,0,0)$, $(0,1,0)$, 并计算 P, R, Q 的坐标，然后利用习题3.8）.

§2.4. 平面，直线与子空间

$\mathbb{A}(V)$中的平面，直线是几何结构最简单的图形。如果固定$\mathbb{A}(V)$的一个坐标架 (O, e_1, e_2, e_3)，则在下述双射

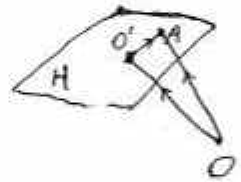
$$\mathbb{A}(V) \xrightarrow{\Phi} V \xrightarrow{e} \mathbb{R}^3, \quad P \longmapsto \overrightarrow{OP} \longmapsto (e_1^*(\overrightarrow{OP}), e_2^*(\overrightarrow{OP}), e_3^*(\overrightarrow{OP}))$$

下，$\mathbb{A}(V)$中的图形 H 分别对应了 V 和 $\mathbb{R}^3$ 中的子集：

$$V(H) = \{\overrightarrow{OA} \mid A\in H\} \subseteq V, \quad E(H) = \{(e_1^*(\overrightarrow{OA}), e_2^*(\overrightarrow{OA}), e_3^*(\overrightarrow{OA})) \mid A\in H\} \subseteq \mathbb{R}^3$$

我们将证明：H 是平面，或直线的充要条件是 $V(H)$, $E(H)$ 满足某些"线性条件"。

例 2.4.1：设 $H \subset \mathbb{A}(V)$ 是一个平面，过 O 作垂直于 H 的直线且交 H 于 O' 点（如图 2.4.1）。



(图2.4.1)

设 $W_H \subset V$ 是所有与 H 平行的向量的集合（包括零向量），$\alpha_0 = \overrightarrow{OO'}$。

则 $V(H) = \alpha_0 + W_H := \{\alpha_0 + \beta \mid \forall \beta \in W_H\}$。特别地，$W_H$ 满足条件：① $\forall \beta_1, \beta_2 \in W_H$，$\lambda_1, \lambda_2 \in \mathbb{R}$，必有 $\lambda_1\beta_1 + \lambda_2\beta_2 \in W_H$。

② 存在线性无关的向量 $\alpha_1, \alpha_2 \in W_H$ 使得 W_H 中任一向量都可由 α_1, α_2 线性表出。

证明：$\forall A \in H$，有 $\overrightarrow{OA} = \overrightarrow{OO'} + \overrightarrow{O'A} = \alpha_0 + \overrightarrow{O'A}$，其中 $\overrightarrow{O'A} \in W_H$。所以，$V(H) \subset \alpha_0 + W_H$。反之，$\forall A \in \mathbb{A}(V)$ 使得 $\overrightarrow{OA} \in \alpha_0 + W_H$，则必有 $A \in H$（从而可得 $V(H) = \alpha_0 + W_H$）。如果 $A \notin H$，则 $\overrightarrow{O'A}$ 不平行 H，但 $\overrightarrow{OA} = \overrightarrow{OO'} + \overrightarrow{O'A} = \alpha_0 + \overrightarrow{O'A} \in \alpha_0 + W_H$ 推得 $\overrightarrow{O'A} \in W_H$，此与 W_H 的定义矛盾。W_H 满足条件①②是显然的。 □.

定义2.4.1：~~设 $W \subset V$ 是一个非空子集，如果 W 满足条件：~~

(1) $\forall \beta_1, \beta_2 \in W$，$\lambda_1, \lambda_2 \in \mathbb{R}$，有 $\lambda_1\beta_1 + \lambda_2\beta_2 \in W$； (2) ~~存在线性无关的 $\alpha_1, \alpha_2 \in W$~~

定义 2.4.1：设 $W\subset V$ 是向量空间 V 的任一非空子集，如果 W 满足条件：$\forall\ \beta_1,\beta_2\in W$，$\forall\ \lambda_1,\lambda_2\in\mathbb{R}$，有 $\lambda_1\beta_1+\lambda_2\beta_2\in W$，则称 W 是 V 的一个子空间。如果存在线性无关的向量 $\alpha_1,\alpha_2\in W$ 使 W 中任意向量可由 α_1,α_2 线性表出，则称 W 是 V 的一个 2 维子空间。如果存在非零向量 $\alpha\in W$ 使 W 中任意向量可由 α 线性表出，则称 W 是 V 的一个 1 维子空间 $\square$.

例 2.4.2：设 $L\subset\mathbb{A}(V)$ 是一条直线，过 O 作垂直于 L 的平面交 L 于 O' 点。设 $W_L\subset V$ 是所有与 L 平行的向量的集合（包括零向量），$\alpha_0=\overrightarrow{OO'}$，则

$$V(L)=\alpha_0+W_L:=\{\alpha_0+\beta\mid\forall\ \beta\in W_L\}.$$

其中 W_L 是 1 维子空间。 $\square$

定理 2.4.1：设 $\mathbb{H}$ 表示 $\mathbb{A}(V)$ 中所有平面的集合，$\mathbb{L}$ 表示 $\mathbb{A}(V)$ 中所有直线的集合。令 $\Sigma_i=\{\alpha+W\mid \alpha\in V,\ W\subset V \text{是 } i \text{ 维子空间}\}$ $(i=1,2)$。则例 2.4.1 和例 2.4.2 中构造的映射

$$\mathbb{L}\longrightarrow\Sigma_1,\quad L\longmapsto\alpha_0+W_L,\qquad \mathbb{H}\longrightarrow\Sigma_2,\quad H\longmapsto\alpha_0+W_H.$$

都是双射。特别，$\mathbb{A}(V)$ 中过原点 O 的平面与 V 中的 2 维子空间 1-1 对应，过 O 点的直线与 V 中的 1 维子空间 1-1 对应。

证明：先证明它们是满射：$\forall\ \alpha+W\in\Sigma_1$，如果 $\alpha\in W$，则 $\alpha+W=W$。由于 W 是 1 维子空间，$W=\{\lambda\beta\mid\forall\lambda\in\mathbb{R}\}$，存在 $A_0\in\mathbb{A}(V)$ 使 $\overrightarrow{OA_0}=\beta$。令 L_0 是以 $\overrightarrow{OA_0}$ 为方向的直线，不难验证 $W=W_{L_0}$，即 W 是直线 L_0 对应的子空间。如果 $\alpha\notin W=\{\lambda\beta\mid\forall\lambda\in\mathbb{R}\}$，令 $H_0\subset\mathbb{A}(V)$ 是过 O 点垂直于 $\beta=\overrightarrow{OA_0}$ 的平面，$O'\in H_0$ 使得 $\alpha_0:=\overrightarrow{OO'}$ 是 α 在 H_0 上的投影，α_1 是 α 在 β 方向上的投影。则 $\alpha+W=\alpha_0+\alpha_1+W=\alpha_0+W$。令 L 是过 O' 点与 H_0 垂直的直线，则

$$\alpha+W=\alpha_0+W_L.$$

下面证明映射 $\mathbb{H} \to \Sigma_2$, $H \mapsto \alpha_0 + W_H$（其中 $\alpha_0 \perp W_H$）是满射：$\forall\ \alpha + W \in \Sigma_2$，令 $A_1, A_2 \in \mathbb{A}(V)$ 使得 $\overrightarrow{OA_1}, \overrightarrow{OA_2} \in W$ 线性无关且 $W = \{\lambda_1\overrightarrow{OA_1} + \lambda_2\overrightarrow{OA_2} \mid \forall \lambda_1, \lambda_2 \in \mathbb{R}\}$。设 H_0 是由 O, A_1, A_2 确定的平面，显然 $W_{H_0} = W$，所以，如果 $\alpha \in W$，则 $\alpha + W = W_{H_0}$。如果 $\alpha \notin W = W_{H_0}$，令 L_0 是 O 点与 H_0 垂直的直线，则存在 $O' \in L_0$ 使 $\alpha_0 := \overrightarrow{OO'}$ 是 α 在 L_0 上的投影，$\alpha_1 \in W_{H_0} = W$ 是 α 在 H_0 上的投影。令 $H \subset \mathbb{A}(V)$ 是过 O' 点与 H_0 平行的平面，则

$$\alpha + W = \alpha_0 + \alpha_1 + W = \alpha_0 + W = \alpha_0 + W_H.$$

最后我们证明映射 $\mathbb{L} \to \Sigma_1$，$\mathbb{H} \to \Sigma_2$ 是单射：设 $L_1, L_2 \subset \mathbb{A}(V)$ 是两条直线，$O_1 \in L_1$，$O_2 \in L_2$ 使 $\alpha_1 = \overrightarrow{OO_1}$，$\alpha_2 = \overrightarrow{OO_2}$ 分别垂直于 L_1, L_2。如果 $\alpha_1 + W_{L_1} = \alpha_2 + W_{L_2}$，则

$$\alpha_1 - \alpha_2 \in W_{L_1} \cap W_{L_2}$$

即：$\alpha_1 - \alpha_2$ 与 L_1, L_2 平行，所以 $\alpha_1 - \alpha_2$ 与 α_1, α_2 垂直，从而 $(\alpha_1-\alpha_2)\cdot(\alpha_1-\alpha_2) = 0$ 故 $\alpha_1 = \alpha_2$，$W_{L_1} = W_{L_2}$。由此可得 $L_1 = L_2$。同理可证 $\mathbb{H} \to \Sigma_2$ 是单射。 □

<u>定理 2.4.2</u>. 固定一个直角坐标架 (O, e_1, e_2, e_3)，则点集 $H \subset \mathbb{A}(V)$ 是一个平面的充要条件是：存在<u>不全为零</u>的 $a, b, c \in \mathbb{R}$ 及 $d \in \mathbb{R}$ 使得

$$H = \{A \in \mathbb{A}(V) \mid a e_1^*(\overrightarrow{OA}) + b e_2^*(\overrightarrow{OA}) + c e_3^*(\overrightarrow{OA}) + d = 0\}.$$

证明："$\Rightarrow$" 如果 $H \subset \mathbb{A}(V)$ 是一个平面，~~垂直于点 O'，则 $A \in H \Leftrightarrow \overrightarrow{OO'} \cdot \overrightarrow{O'A} = 0$~~

令 $\vec{n} = ae_1 + be_2 + ce_3$ 是一个垂直于 H 的<u>非零向量</u>.

过 O 点作垂直于 H 的直线交 H 于 O' 点，则

$$A \in H \Leftrightarrow \overrightarrow{O'A} \cdot \vec{n} = 0 \Leftrightarrow \vec{n} \cdot (\overrightarrow{OA} + \overrightarrow{O'O}) = 0.$$

所以，令 $d = \vec{n} \cdot \overrightarrow{O'O}$，我们有

$$A \in H \Leftrightarrow a e_1^*(\overrightarrow{OA}) + b e_2^*(\overrightarrow{OA}) + c e_3^*(\overrightarrow{OA}) + d = 0.$$

"$\Leftarrow$"：如果存在不全为零常数 a,b,c 及常数 d 使得

$$H=\{A\in \mathbb{A}(V)\mid ae_1^*(\overrightarrow{OA})+be_2^*(\overrightarrow{OA})+ce_3^*(\overrightarrow{OA})+d=0\}.$$

则 H 必为平面。事实上，设 $(x_0,y_0,z_0)\in\mathbb{R}^3$ 是方程 $ax+by+cz+d=0$ 的一组解，$\overrightarrow{OO'}=x_0e_1+y_0e_2+z_0e_3$，$H'$ 是过 O' 点与向量

$$\vec{n}=ae_1+be_2+ce_3$$

垂直的平面。则 $d=-(ax_0+by_0+cz_0)=-\vec{n}\cdot\overrightarrow{OO'}$，且

$$H'=\{A\in\mathbb{A}(V)\mid \vec{n}\cdot(\overrightarrow{OA}+\overrightarrow{O'O})=0\}=H$$

所以 H 就是平面 H'.

□.

上述定理表明：$\mathbb{A}(V)$ 中的一个点集 H 是一个平面当且仅当它~~的坐标~~在一个直角坐标架 (O,e_1,e_2,e_3) 下的坐标满足方程

$$ax+by+cz+d=0.$$

所以我们称它是平面 H 在直角坐标架下的方程。由坐标变换公式可知 H 在任意仿射坐标架下的方程也是一个形如 $a'x+b'y+c'z+d'=0$ 的一次多项式(称为线性方程)。但直角坐标架 (O,e_1,e_2,e_3) 下的方程系数有如下几何意义。

推论 2.4.1：设 $ax+by+cz+d=0$ 是平面 H 在直角坐标架 (O,e_1,e_2,e_3) 下的方程，则

(1) $\vec{n}=ae_1+be_2+ce_3$ 是垂直于 H 的向量(由它确定的直线称为 H 的法线)。

(2) 设 A 是坐标为 (x,y,z) 的点，令 $\rho(A,H)$ 表示点 A 到平面 H 的距离，则

$$\rho(A,H)=\frac{|ax+by+cz+d|}{|\vec{n}|}=\frac{|ax+by+cz+d|}{\sqrt{a^2+b^2+c^2}}.$$

证明：(1) 由定理 2.4.2 的证明可知。

(2) 过 A 点作垂直于 H 的直线交 H 于 A_0 点。如果 A_0 的坐标是 (x_0,y_0,z_0)，则 $ax_0+by_0+cz_0+d=0$.

所以 $ax+by+cz+d=a(x-x_0)+b(y-y_0)+c(z-z_0)=\vec{n}\cdot\overrightarrow{A_0A}$。由内积定义，$\vec{n}\cdot\overrightarrow{A_0A}=|\vec{n}||\overrightarrow{A_0A}|$（因为 $\vec{n}$ 与 $\overrightarrow{A_0A}$ 平行）。因此，

$$\rho(A,H):=|\overrightarrow{A_0A}|=\frac{|ax+by+cz+d|}{|\vec{n}|}=\frac{|ax+by+cz+d|}{\sqrt{a^2+b^2+c^2}}。$$

□

可以将平面 H 的方程 $ax+by+cz+d=0$ 改写为它的法式方程

$$\frac{a}{\sqrt{a^2+b^2+c^2}}x+\frac{b}{\sqrt{a^2+b^2+c^2}}y+\frac{c}{\sqrt{a^2+b^2+c^2}}z+\frac{d}{\sqrt{a^2+b^2+c^2}}=0$$

则它在任一点 A 的取值的绝对值等于 A 到 H 的距离。

推论2.4.2. 设 H_1,H_2 是两个不重合的平面，设它们在直角坐标架 (O,e_1,e_2,e_3) 下的方程分别是 $a_1x+b_1y+c_1z+d_1=0$ 和 $a_2x+b_2y+c_2z+d_2=0$，则

(1) H_1 与 H_2 平行 $\Leftrightarrow (a_2,b_2,c_2)=(\lambda a_1,\lambda b_1,\lambda c_1)$ 其中 $\lambda\in\mathbb{R}$；

(2) H_1 与 H_2 垂直 $\Leftrightarrow a_1a_2+b_1b_2+c_1c_2=0$；

(3) 如果 θ 是 H_1 与 H_2 的一个夹角，则

$$\cos\theta=\frac{a_1a_2+b_1b_2+c_1c_2}{\sqrt{a_1^2+b_1^2+c_1^2}\cdot\sqrt{a_2^2+b_2^2+c_2^2}}.$$

(4) 如果 H_1,H_2,H_3 是三个不同的平面，它们的定义方程是

$$\begin{aligned} a_1x+b_1y+c_1z+d_1&=0\\ a_2x+b_2y+c_2z+d_2&=0\\ a_3x+b_3y+c_3z+d_3&=0.\end{aligned}$$

则 H_1,H_2,H_3 相交于一个点的充要条件是

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

(5) 如果 $\Delta=0$，则 H_1,H_2,H_3 都与某条直线平行。特别，当 H_1,H_2,H_3 相交时，H_1,H_2,H_3 通过一条共同的直线.

证明：(1) 设 $\vec{n_1}=a_1e_1+b_1e_2+c_1e_3$，$\vec{n_2}=a_2e_1+b_2e_2+c_2e_3$，则 $\vec{n_1}$，$\vec{n_2}$ 分别与 H_1,H_2 垂直。所以，H_1 与 H_2 平行当且仅当 $\vec{n_1},\vec{n_2}$ 平行 $\Longleftrightarrow \vec{n_2}=\lambda\vec{n_1}$.

(2) H_1 与 H_2 垂直 $\Longleftrightarrow$ $\vec{n_1}$ 与 $\vec{n_2}$ 垂直 $\Longleftrightarrow \vec{n_1}\cdot\vec{n_2}=a_1a_2+b_1b_2+c_1c_2=0$.

(3) H_1 与 H_2 有两个夹角（如果它们不平行），$\vec{n_1}$ 与 $\vec{n_2}$ 的夹角 θ 等于其中之一。

$$\cos\theta=\frac{\vec{n_1}\cdot\vec{n_2}}{|\vec{n_1}|\,|\vec{n_2}|}=\frac{a_1a_2+b_1b_2+c_1c_2}{\sqrt{a_1^2+b_1^2+c_1^2}\sqrt{a_2^2+b_2^2+c_2^2}}.$$

(4) 由于 H_1,H_2,H_3 相交于一点 $\Longleftrightarrow$ $\vec{n_1},\vec{n_2},\vec{n_3}=a_3e_1+b_3e_2+c_3e_3$ 不共面. $\Longleftrightarrow$ $\Delta=(\vec{n_1}\times\vec{n_2})\cdot\vec{n_3}\neq 0$.

(5) $\Delta=0 \Longleftrightarrow \vec{n_1},\vec{n_2},\vec{n_3}$ 共面。所以存在一条直线与 $\vec{n_1},\vec{n_2},\vec{n_3}$ 垂直因此该直线与 H_1,H_2,H_3 平行。特别，可论 H_2,H_3 与 H_1 中的一条直线 L 平行。如果 H_2,H_3 与 H_1 相交，则 $L\subset H_2$，$L\subset H_3$

$\square$.

一条直线 $L\subset \mathbb{A}(V)$ 可以看成两个平面 H_1,H_2 的交，所以直线 L 在 (O,e_1,e_2,e_3) 下的方程是由两个独立的线性方程组成的相容方程组

$$\begin{cases}a_1x+b_1y+c_1z+d_1=0\\ a_2x+b_2y+c_2z+d_2=0.\end{cases}$$

反之，所有坐标满足 独立线性方程组成的相容线性方程组

$$\begin{cases}a_1x+b_1y+c_1z+d_1=0\\ a_2x+b_2y+c_2z+d_2=0\end{cases}$$

的点组成一条直线。对于给定的直线 $L\subset \mathbb{A}(V)$，定义 L 的相容线性方程组并不唯一，所以有时采用 L 的参数方程。

设 $\vec{v}=v_1e_1+v_2e_2+v_3e_3$ 是平行于 L 的一个非零向量，A_0 是 L 上坐标为 (x_0,y_0,z_0) 的给定一个点，则

$A\in L \Longleftrightarrow A$ 的坐标 (x,y,z) 满足如下"参数方程"：

$x=x_0+v_1t$，$y=y_0+v_2t$，$z=z_0+v_3t$ （t 是参数）.

如果已知 L 的参数方程，则消去参数 t（这是可能的，因为 v_1,v_2,v_3 不全为零）

可以得到一组定义L的相容独立的线性方程。反之，如果已知L是由一组相容，独立的线性方程

$$\begin{cases} a_1x+b_1y+c_1z+d_1=0 \\ a_2x+b_2y+c_2z+d_2=0 \end{cases}$$

定义，则不难得到L的一个参数方程。如下：

设(O,e_1,e_2,e_3)是直角坐标架，直线L由

$$\begin{cases} a_1x+b_1y+c_1z+d_1=0 \\ a_2x+b_2y+c_2z+d_2=0 \end{cases} \quad (*)$$

定义。如果H_i是由$a_ix+b_iy+c_iz+d_i=0$定义的平面$(i=1,2)$，则$\vec{n_i}:=a_ie_1+b_ie_2+c_ie_3$是与$H_i$垂直向量。如果$(v_1,v_2,v_3)\in\mathbb{R}^3$是方程组

$$\begin{cases} a_1x+b_1y+c_1z=0 \\ a_2x+b_2y+c_2z=0 \end{cases} \quad (**)$$

的一组非零解，则$\vec{v}:=v_1e_1+v_2e_2+v_3e_3$是与L平行的向量。所以，如果$(x_0,y_0,z_0)$是方程组$(*)$的一组解，则

$$x=x_0+v_1t\ ,\quad y=y_0+v_2t\ ,\quad z=z_0+v_3t$$

是L的一组参数方程。在求参数方程时，(v_1,v_2,v_3)可以取方程组$(**)$的任意一组非零解。例如：

$$v_1=\begin{vmatrix} b_1 & c_1 \\ b_2 & c_2 \end{vmatrix},\quad v_2=\begin{vmatrix} c_1 & a_1 \\ c_2 & a_2 \end{vmatrix},\quad v_3=\begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix}。$$

下面介绍如何利用定义平面，直线的方程来确定平面与直线，直线与直线的位置关系：平行，垂直，夹角等。

命题2.4.3（平面与直线的关系）：设平面$H\subset\mathbb{A}(V)$由

$$ax+by+cz+d=0$$

定义。如果直线L由参数方程：$x=x_0+v_1t$，$y=y_0+v_2t$，$z=z_0+v_3t$定义，则 (1) H与L平行$\Leftrightarrow$ $av_1+bv_2+cv_3=0$。

(2) H与L垂直$\Leftrightarrow$存在非零常数$\lambda\in\mathbb{R}$使$(a,b,c)=(\lambda v_1,\lambda v_2,\lambda v_3)$。 □

如果直线L的线性方程组

$$\begin{cases} a_1x+b_1y+c_1z+d_1=0 \\ a_2x+b_2y+c_2z+d_2=0 \end{cases}$$

定义，则由推论2.4.2中的结论(4)可得：

$$L\text{与}H\text{相交于一点} \Longleftrightarrow \begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a & b & c \end{vmatrix} \neq 0。$$

L与H不相交(平行)当且仅当线性方程组

$$\begin{cases} a_1x+b_1y+c_1z+d_1=0 \\ a_2x+b_2y+c_2z+d_2=0 \\ ax+by+cz+d=0 \end{cases}$$

无解。而L在平面H中的充要条件是线性方程组

$$\begin{cases} a_1x+b_1y+c_1z+d_1=0 \\ a_2x+b_2y+c_2z+d_2=0 \\ ax+by+cz+d=0 \end{cases}$$

有无穷多组解。

直线L与平面H的夹角(当它相交时)定义为L在H上的投影 l 与L的夹角，它们有两个不同的夹角θ_1,θ_2(但θ_1,θ_2互补). 一般将θ_1,θ_2中的锐角

$$\theta = \arccos(|\sin\langle\vec{v},\vec{n}\rangle|)$$

称为L与H的夹角，其中$\vec{v}$是与L平行的向量，$\vec{n}$是H的法向量(即$\vec{n}$与H垂直)。

推论2.4.4：设 $ax+by+cz+d=0$ 是平面H在直角坐标架(O,e_1,e_2,e_3)下的定义方程。如果L由参数方程

$$x=x_0+v_1t,\quad y=y_0+v_2t,\quad z=z_0+v_3t$$

定义，则L与H的夹角(当它们不平行时) θ 满足：

$$|\sin\theta| = \frac{|av_1+bv_2+cv_3|}{\sqrt{a^2+b^2+c^2}\cdot\sqrt{v_1^2+v_2^2+v_3^2}}.$$

证明：$\vec{u}=v_1e_1+v_2e_2+v_3e_3$ 与 L 平行，$\vec{n}=ae_1+be_2+ce_3$ 与 H 垂直

由 $\vec{n}\cdot\vec{u}=|\vec{n}|\cdot|\vec{u}|\cos\langle\vec{n},\vec{u}\rangle=|\vec{n}|\cdot|\vec{u}|\sin\theta$ 可得

$$|\sin\theta|=\frac{|av_1+bv_2+cv_3|}{\sqrt{a^2+b^2+c^2}\cdot\sqrt{v_1^2+v_2^2+v_3^2}}.$$

也可由外积公式 $|\vec{n}\times\vec{u}|=|\vec{u}|\cdot|\vec{n}|\cdot|\sin\langle\vec{u},\vec{n}\rangle|=|\vec{u}|\cdot|\vec{n}||\cos\theta|$

计算 $|\cos\theta|$：

$$|\cos\theta|=\frac{\sqrt{\begin{vmatrix}b&c\\v_2&v_3\end{vmatrix}^2+\begin{vmatrix}c&a\\v_3&v_1\end{vmatrix}^2+\begin{vmatrix}a&b\\v_1&v_2\end{vmatrix}^2}}{\sqrt{a^2+b^2+c^2}\cdot\sqrt{v_1^2+v_2^2+v_3^2}}.$$

□

直线与直线的关系有：异面[30]，共面（包括：相交，重合，平行）。

如果直线 L_1 与 L_2 由参数方程：

L_1： $x=x_0+v_1t$，$y=y_0+v_2t$，$z=z_0+v_3t$

L_2： $x=x_0'+v_1't$，$y=y_0'+v_2't$，$z=z_0'+v_3't$

定义，则 L_1, L_2 异面（即不共面）的充要条件是

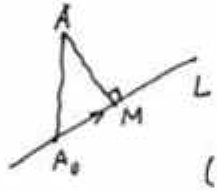
$$\begin{vmatrix}x_0-x_0' & y_0-y_0' & z_0-z_0'\\ v_1 & v_2 & v_3\\ v_1' & v_2' & v_3'\end{vmatrix}\neq 0.$$

如果 L_1, L_2 共面（即上述行列式等于0），则 L_1 与 L_2 平行的充要条件是 $\vec{v}=v_1e_1+v_2e_2+v_3e_3$ 与 $\vec{v}'=v_1'e_1+v_2'e_2+v_3'e_3$ 线性相关，即 L_1 与 L_2 重合的充要条件是 $\vec{v},\vec{v}'$，$(x_0-x_0')e_1+(y_0-y_0')e_2+(z_0-z_0')e_3$ 三个向量共线。

如果 L_1, L_2 由4个平面方程 $a_ix+b_iy+c_iz+d_i=0$ $(i=1,2,3,4)$ 给出，令 H_i 是由 $a_ix+b_iy+c_iz+d_i=0$ 定义的平面，则 L_1 与 L_2 的位置判断至少有两种方法，一种是从平面方程计算出 L_1, L_2 的参数方程，另一种是通过直线与平面，平面与平面之间的位置关系进行判断。但不管采用何种方法，都会涉及解线性方程组的问题。

设 L 是过点 A_0 的直线, A 是 L 外一点, 过 A 点作垂直于 L 的平面交 L 于点 M. 则点 A 到直线 L 的距离定义为

$$\rho(A,L) := |\overrightarrow{AM}|.$$



(图2.4.2)

推论 2.4.5: 设直线 L 的参数方程是

$$x = x_0 + v_1 t\ ,\quad y = y_0 + v_2 t\ ,\quad z = z_0 + v_2 t\ .$$

点 A 的坐标为 (a_1, a_2, a_3). 则

$$\rho(A,L) = \sqrt{\frac{\begin{vmatrix} a_2-y_0 & a_3-z_0 \\ v_2 & v_3 \end{vmatrix}^2 + \begin{vmatrix} a_3-z_0 & a_1-x_0 \\ v_3 & v_1 \end{vmatrix}^2 + \begin{vmatrix} a_1-x_0 & a_2-y_0 \\ v_1 & v_2 \end{vmatrix}^2}{v_1^2 + v_2^2 + v_3^2}}$$

证明: 如图2.4.2, $\overrightarrow{OA_0} = x_0 e_1 + y_0 e_2 + z_0 e_3$. $\overrightarrow{OA} = a_1 e_1 + a_2 e_2 + a_3 e_3$

$\vec{v} = v_1 e_1 + v_2 e_2 + v_3 e_3$ 与 $\overrightarrow{A_0M}$ 平行. 所以

$$\frac{|\overrightarrow{A_0A} \times \vec{v}|}{|\vec{v}|} = \frac{|\overrightarrow{A_0A} \times \overrightarrow{A_0M}|}{|\overrightarrow{A_0M}|} = \frac{|\overrightarrow{AM} \times \overrightarrow{A_0M}|}{|\overrightarrow{A_0M}|} = |\overrightarrow{AM}| := \rho(A,L).$$

将坐标代入即得公式。

□.

习题 2.4

4.1. 设 L_1, L_2 是两条直线, L_1 与 L_2 之间距离 $\rho(L_1, L_2)$ 定义为 L_1 上点与 L_2 上点之间的最短距离, 即

$$\rho(L_1, L_2) = \min\{\rho(A_1, A_2) \mid A_1 \in L_1,\ A_2 \in L_2\}.$$

证明: 如果 L_1, L_2 平行, 则 $\rho(L_1, L_2)$ 等于 L_1 上任一点到 L_2 的距离。

4.2. 设 L_1, L_2 是两条异面直线, 证明存在一条唯一的直线 L 分别与 L_1, L_2 垂直相交 (L 称为 L_1, L_2 的公垂线)。

4.3. 设 L_1, L_2 是两条异面直线，L 是 L_1, L_2 的公垂线分别交 L_1, L_2 于 A_1, A_2。证明：$\rho(L_1, L_2) = \rho(A_1, A_2)$（$A_1, A_2$ 之间的距离）。

4.4. 设两条异面直线 L_1, L_2 分别过点 M_1, M_2，且 L_1, L_2 的方向向量分别是 $\vec{v_1}, \vec{v_2}$。证明：

$$\rho(L_1, L_2) = \frac{|\overrightarrow{M_1M_2} \cdot (v_1 \times v_2)|}{|v_1 \times v_2|}$$

并说明上述公式的几何意义。

4.5. 设 L_1, L_2 分别由参数方程 $x = x_0 + v_1 t$, $y = y_0 + v_2 t$, $z = z_0 + v_3 t$ 和 $x = x_0' + v_1' t$, $y = y_0' + v_2' t$, $z = z_0' + v_3' t$ 定义。请用坐标 (x_0, y_0, z_0), (x_0', y_0', z_0'), (v_1, v_2, v_3), (v_1', v_2', v_3') 写出 $\rho(L_1, L_2)$ 的公式。

4.6. 设在直角坐标架 (O, e_1, e_2, e_3) 下，$H \subset \mathbb{A}(V)$ 由方程

$$ax + by + cz + d = 0$$

定义。证明存在线性函数 $l: V \to \mathbb{R}$ 使得

$$A \in H \iff l(\overrightarrow{OA}) + d = 0.$$

4.7. 设直线 L 在直角坐标架 (O, e_1, e_2, e_3) 下由

$$\begin{cases} a_1 x + b_1 y + c_1 z + d_1 = 0 \\ a_2 x + b_2 y + c_2 z + d_2 = 0 \end{cases}$$

定义。证明存在线性无关的两个线性函数 $l_i: V \to \mathbb{R}$ $(i=1,2)$ 使得 $A \in L \iff l_1(\overrightarrow{OA}) + d_1 = 0,\ l_2(\overrightarrow{OA}) + d_2 = 0$。

4.8. 设直线 L 由下列线性方程组

$$\begin{cases} 4x - y + 3z - 1 = 0 \\ x + 5y - z + 2 = 0 \end{cases}$$

定义。试求 L 的一组参数方程。

4.9. 证明：到三个两两不平行的平面等距离的点的几何轨迹是一条直线。（提示：利用三个平面的法式方程）。

4.10. 证明：到三角形三个顶点等距离的点的几何轨迹是一条直线（首先证明到两个给定点等距离的点的几何轨迹是一个平面）。

4.11. 证明：如果 $a_1x+b_1y+c_1z+d_1=0$ 和 $a_2x+b_2y+c_2z+d_2=0$ 定义同一个平面，则存在 $\lambda\in\mathbb{R}$ 使 $a_1x+b_1y+c_1z+d_1=\lambda(a_2x+b_2y+c_2z+d_2)$。

4.12. 如果两组相容，独立的线性方程组

$$\begin{cases} a_1x+b_1y+c_1z+d_1=0 \\ a_2x+b_2y+c_2z+d_2=0 \end{cases}, \quad \begin{cases} a_3x+b_3y+c_3z+d_3=0 \\ a_4x+b_4y+c_4z+d_4=0 \end{cases}$$

定义同一直线，令 $\alpha=(a_1,b_1,c_1,d_1)$，$\beta=(a_2,b_2,c_2,d_2)$，$\gamma=(a_3,b_3,c_3,d_3)$，$\mu=(a_4,b_4,c_4,d_4)$ 表示 $\mathbb{R}^4$ 中的元素。则存在 $\lambda_1,\lambda_2,\lambda_3,\lambda_4\in\mathbb{R}$，$\mu_1,\mu_2,\mu_3,\mu_4\in\mathbb{R}$ 使得

$$\begin{cases} \gamma=\lambda_1\alpha+\lambda_2\beta \\ \mu=\mu_1\alpha+\mu_2\beta \end{cases}, \quad \begin{cases} \alpha=\lambda_3\gamma+\lambda_4\mu \\ \beta=\mu_3\gamma+\mu_4\mu \end{cases}.$$

4.13. 设直线 $L\subset\mathbb{A}(V)$ 在直角坐标架 (O,e_1,e_2,e_3) 下由

$$\begin{cases} 2x+3y+z=0 \\ x+2y-z=0 \end{cases}$$

定义。试求向量 $\alpha\in V$ 使得

$$\{\overrightarrow{OA}\mid A\in L\}=\{\lambda\alpha\mid \lambda\in\mathbb{R}\}.$$

4.14. 设平面 $H\subset\mathbb{A}(V)$ 在直角坐标架 (O,e_1,e_2,e_3) 下由

$$2x+3y+z=0$$

定义。试求 $\alpha_1,\alpha_2\in V$ 使得

$$\{\overrightarrow{OA}\mid A\in H\}=\{\lambda_1\alpha_1+\lambda_2\alpha_2\mid \lambda_1,\lambda_2\in\mathbb{R}\}.$$

第三章 线性代数初步

上一章中的向量空间是有向线段等价类的集合V，通过几何方式在V上定义了"加法"和"数乘法"，它们满足8条规则。利用上述的线性运算解决几何问题时，我们可以忘记两个运算的具体定义，仅利用它们满足的"规则"。只有将计算结果应用于几何时，才需要利用"加法"和"数乘"的具体含义。在实际中有许多各种不同对象的集合，在这些集合上有各种不同应用背景的"加法"和"数乘"运算。带有这两种运算的集合虽然来自不同的应用场景，它们却满足与有向线段运算一样的8条规则。所以有必要忽略各种运算的具体含义，仅研究满足8条规则的"加法"和"数乘法"。

在第一节，我们将引入抽象向量空间的概念，建立抽象向量空间的维数理论，将抽象的"线性运算"变成具体的坐标运算。第二节是关于线性方程组的理论，它的解空间为上节的抽象向量空间提供了重要实例。在第三节，我们引入线性映射概念，主要强调矩阵是线性映射的坐标这一观点。第四节我们将引入多重线性函数的概念，本节将引入行列式的定义，建立二次型的标准型。

§3.1. 抽象向量空间

解析几何中的向量空间有两个方面的推广。一个方面是"加法"和"数乘运算"的推广。设V是任意非空集合，V上的"加法"和"数乘运算"是任意的运算。另一个方面是"数"的推广。"数乘运算"中的数不必是实数，它可以是任意抽象域里的元素。

定义3.1.1[域的定义]. 设K是至少有两个元素的集合，并带有两个映射

$$\psi_1: K\times K\longrightarrow K,\quad \psi_2: K\times K\longrightarrow K$$

(习惯上，$\forall\, a,b\in K$，将$(a,b)\in K\times K$在ψ_1，ψ_2下的像记为：$\psi_1(a,b):=a+b$，$\psi_2(a,b):=a\cdot b$. ψ_1,ψ_2分别称为K上的"加法"和"乘法")

如果上述两个映射(或称"加法运算"和"乘法运算")满足如下9条规则，则称$K=(K,\psi_1,\psi_2)$是一个域。

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

(2) 存在元素$0\in K$满足：$a+0=0+a=a$，$\forall\, a\in K$ (零元存在性)

(3) $\forall\, a\in K$，存在$b\in K$使$a+b=b+a=0$. (负元存在性)

(4) $\forall\, a,b\in K$，有 $a+b=b+a$ (交换律)

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

(6) 存在元素$1\in K$满足：$a\cdot 1=1\cdot a=a$，$\forall\, a\in K$ (单位元存在性).

(7) $\forall\, a\in K$，如果$a\neq 0$，则存在$b\in K$使得

$$a\cdot b=b\cdot a=1$$

(非零元可逆).

(8) $\forall\, a,b\in K$，有 $a\cdot b=b\cdot a$ (交换律).

(9) $\forall\, a,b,c\in K$，有

$$a\cdot(b+c)=a\cdot b+a\cdot c,\quad (a+b)\cdot c=a\cdot c+b\cdot c.$$ (分配律).

□

注记：(i) (2)中零元$0\in K$是唯一的：如果$0'\in K$也满足条件(2)，则 $0'=0'+0=0$.

(ii) 对给定的$a\in K$，(3)中的b由a唯一确定：如果$b,b'\in K$都满足：$a+b=b+a=0$，$a+b'=b'+a=0$. 则 (所以将b记为$b=-a$)

$$b'=b'+0=b'+(a+b)=(b'+a)+b=0+b=b.$$

(iii) 条件(6)中的1也是唯一的：如果$1'\in K$也满足(6). 则$1'=1\cdot 1'=1$.

(iv) 条件(7)中的 b 由 a 唯一确定：如果 $c\in K$ 也满足 $a\cdot c=c\cdot a=1$.

则 $b=1\cdot b=(c\cdot a)\cdot b=c\cdot(a\cdot b)=c\cdot 1=c$. (所以，将 b 记为 $b=a^{-1}$)

(v). $\forall a\in K$，有 $0\cdot a=a\cdot 0=0$： $0=0+0$，所以 $0\cdot a=0\cdot a+0\cdot a$

$$0=0\cdot a+(-0\cdot a)=(0\cdot a+0\cdot a)+(-0\cdot a)=0\cdot a+(0\cdot a+(-0\cdot a))=0\cdot a.$$

(vi). $1\neq 0$. 否则，$\forall a\in K$， $a=1\cdot a=0\cdot a=0$. (与K中至少有两个元素矛盾).

(vii) $\forall a,b\in K$， $(-a)\cdot b=a\cdot(-b)=-a\cdot b$. (只需证明：$(-a)\cdot b+a\cdot b=0$，$a\cdot(-b)+a\cdot b=0$)。

记号： $a-b:=a+(-b)$.

例3.1.1：$\mathbb{C}$ 关于复数的"加法"和"乘法"是一个域.(称为复数域)

设 $K\subset\mathbb{C}$ 是 $\mathbb{C}$ 的子域.(即：K 关于 $\mathbb{C}$ 的"四则运算和逆算封闭")

则 K 是一个域。($K=\mathbb{Q}$，$\mathbb{R}$ 等).

例3.1.2. 设 $p>0$ 是一个素数，$\forall a,b\in\mathbb{Z}$，定义：

$$a\sim b \iff p\mid(a-b),\quad \bar{a}=\{b\in\mathbb{Z}\mid b\sim a\}.$$

则 $\bar{a}_1=\bar{a}_2 \iff p\mid(a_1-a_2)$，特别 $\bar{0}=\{a\mid p\mid a\}=\{b\cdot p\mid \forall b\in\mathbb{Z}\}$.

令 $\mathbb{F}_p=\{\bar{0},\bar{1},\bar{2},\cdots,\overline{p-1}\}$，定义"加法"和"乘法"如下：

$$\forall \bar{a},\bar{b}\in\mathbb{F}_p,\qquad \bar{a}+\bar{b}:=\overline{a+b},\quad \bar{a}\cdot\bar{b}:=\overline{a\cdot b}.$$

则 $\mathbb{F}_p$ 关于上述的"加法"和"乘法"是一个域，$\mathbb{F}_p$ 中的零元和单位元分别是 $\bar{0}$ 和 $\bar{1}$。

证明. 仅证明 $\mathbb{F}_p$ 中每个非零元可逆：$\forall \bar{a}\in\mathbb{F}_p$ 则

$$\bar{a}\neq\bar{0} \iff p\nmid a \iff a \text{ 与 } p \text{ 互素}$$

所以，如果 $\bar{a}\neq\bar{0}$，则存在 $x,y\in\mathbb{Z}$ 使 $1=ax+p\cdot y$. 故

$$\bar{1}=\overline{ax+py}=\overline{a\cdot x}+\overline{py}=\overline{ax}=\bar{a}\cdot\bar{x}=\bar{x}\cdot\bar{a}.$$

□

对于初学者，在下面的讨论中(如果是关于抽象向量空间的定义)，可以将"一个域 K"想像为 $\mathbb{C}$ 的子域。

定义 3.1.2 [向量空间定义]. 设 K 是一个域，V 是一个非空集合，

$$\varphi_1 : V\times V \to V \quad , \quad \varphi_2 : K\times V \to V$$

是两个映射 ($\forall\ (\alpha,\beta)\in V$, $(\lambda,\alpha)\in K\times V$, 习惯上将它们在 φ_1,φ_2 下的像 $\varphi_1(\alpha,\beta)$, $\varphi_2(\lambda,\alpha)$ 分别记为：$\alpha+\beta$, $\lambda\cdot\alpha$。因此，φ_1 称为 V 上的"加法运算"，φ_2 称为 V 上的"数乘运算"). 如果 V 上的"加法运算"和"数乘运算"满足如下 8 条规则，则称

$$V =: (V, \varphi_1, \varphi_2) := (V, +, \cdot).$$

是一个 K-向量空间 (或称为 K-线性空间)。

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

(2) 存在零向量 $0\in V$ 满足：$\alpha+0=0+\alpha=\alpha$, $\forall\ \alpha\in V$. (零向量存在性)

(3) $\forall\ \alpha\in V$, 存在 $\beta\in V$ 使：$\alpha+\beta=\beta+\alpha=0$, (负向量存在性).

(4) $\forall\ \alpha,\beta\in V$ ，有 $\alpha+\beta=\beta+\alpha$ (加法交换律).

(5) $\forall\ \alpha\in V$ ，有 $1\cdot\alpha=\alpha$

(6) $\forall\ \lambda,\mu\in K$, $\alpha\in V$, 有 $\lambda\cdot(\mu\cdot\alpha)=(\lambda\mu)\cdot\alpha$ (数乘运算结合律)

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

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

(分配律).

$\square$

注记：(1) 条件 (2) 和 (3) 中的零向量 0 和 α 的负向量 β 都是唯一的：如果 β_1,β_2 满足：$\alpha+\beta_1=\beta_1+\alpha=0$, $\alpha+\beta_2=\beta_2+\alpha=0$. 则

$$\beta_1=\beta_1+0=\beta_1+(\alpha+\beta_2)=(\beta_1+\alpha)+\beta_2=0+\beta_2=\beta_2.$$

由于 (3) 中的 β 由 α 唯一确定，可将 β 记为 $\beta=-\alpha$.

(2). $\forall\, \alpha \in V$, 有 $0\cdot\alpha = 0$. ($0+0=0 \Rightarrow (0+0)\cdot\alpha = 0\cdot\alpha$. 由分配律得 $0\cdot\alpha + 0\cdot\alpha = 0\cdot\alpha$. 故 $0\cdot\alpha = (0\cdot\alpha+0\cdot\alpha)+(-0\cdot\alpha) = 0\cdot\alpha+(-0\cdot\alpha) = 0$.).

所以, $(-\lambda)\cdot\alpha = -\lambda\cdot\alpha$ (因为 $(-\lambda)\cdot\alpha + \lambda\cdot\alpha = (-\lambda+\lambda)\cdot\alpha = 0\cdot\alpha = 0$).

记号: $\forall\, \alpha, \beta \in V$, $\alpha - \beta := \alpha + (-\beta)$.

~~例~~

定义3.1.3 [子空间定义]. 设 V 是一个 K-向量空间, $W \subset V$ 是一个非空子集. 如果 W 满足如下条件:

(1) $\forall\, \alpha, \beta \in W$, 有 $\alpha + \beta \in W$.

(2) $\forall\, \lambda \in K, \alpha \in W$, 有 $\lambda\cdot\alpha \in W$.

则称 W 是 V 的一个子空间. □

引理3.1.1: 设 $W \subset V$ 是 K-向量空间 V 的子空间, 则 W 关于 V 的"加法"和"数乘法"是一个 K-向量空间.

证明: 由子空间的定义, V 的"加法"和"数乘法"诱导了 W 的"加法"和"数乘法". 所以, K-向量空间定义中的8个条件只需验证: W 中每个向量 α 的负向量 $-\alpha \in V$ 也在 W 中. 这里经由 $-\alpha = (-1)\cdot\alpha \in W$ 得到. 所以 W 是 K-向量空间 □

例3.1.3: $K^n = \{x = (x_1, \cdots, x_n) \mid x_i \in K\}$ 或 $K^n = \{x = [x_1, \cdots, x_n] \mid x_i \in K\}$ 关于"加法": $x+y := (x_1+y_1, \cdots, x_n+y_n)$. (或 $x+y = [x_1+y_1, \cdots, x_n+y_n]$) 和"数乘法": $\forall\, \lambda \in K$, $\lambda\cdot x := (\lambda x_1, \cdots, \lambda x_n)$. (或 $\lambda\cdot x = [\lambda x_1, \cdots, \lambda x_n]$) 成为 K-向量空间. (分别称为 n维行向量空间, 或 n维列向量空间). 特别, 当 $n=1$ 时, K 是1维 K-向量空间 □.

例3.1.4: $\mathbb{C}$ 关于通常的加法和乘法成为一个 $\mathbb{R}$-向量空间和 $\mathbb{Q}$-向量空间. □

例 3.1.5：~~[illegible]~~齐次线性方程组

$$(*)\quad \begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = 0 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = 0 \\ \cdots\cdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = 0 \end{cases}$$ $a_{ij}\in K$ 是给定常数。

则 (*) 的解空间 $V = \{x = (x_1, \cdots, x_n) \mid x \text{ 满足方程组 } (*)\} \subset K^n$ 是 K^n 的一个子空间（所以也是一个 K-向量空间）。 □

例 3.1.6：设 $V = K[x_1, \ldots, x_n]$ 关于多项式的加法和乘法（K 中元作为零次多项式与 V 中任意多项式相乘）成为一个 K-向量空间。其中 $V_d = \{d$ 次齐次多项式 $f_d(x_1, \cdots, x_n)\} \subset V$ 是一个子空间。 □

例 3.1.7：$V = \{$连续函数 $\mathbb{R} \xrightarrow{f} \mathbb{R}\}$ 关于函数的加法和乘法（$\mathbb{R}$ 中的数 $\lambda \in \mathbb{R}$ 确定一个常值函数与 V 中任意 f 相乘）成为一个 $\mathbb{R}$-向量空间。且 $W = \{\lambda_1 \sin t + \lambda_2 \cos t \mid \lambda_1, \lambda_2 \in \mathbb{R}\}$ 是 V 的一个子空间。 □

例 3.1.8：$V = \{$有向线段等价类$\}$ 关于平行四边形法则的加法和有向线段的数乘法成为一个 $\mathbb{R}$-向量空间。V 中与一个固定平面 H 平行的向量组成一个子空间 $W_H \subset V$.（二维子空间），与一个固定直线 L 平行的向量组成一个 1 维子空间 $W_L \subset V$。 □

定义 3.1.4：设 V 是一个 K-向量空间，$\alpha_1, \cdots, \alpha_m \in V$，$\lambda_1, \cdots, \lambda_m \in K$，则 $\lambda_1\alpha_1 + \cdots + \lambda_m\alpha_m \in V$ 称为 $\alpha_1, \cdots, \alpha_m$ 的一个线性组合，其中 $\lambda_1, \cdots, \lambda_m$ 称为该线性组合的系数。如果

$$\beta = \lambda_1\alpha_1 + \cdots + \lambda_m\alpha_m$$

则称 β 由 $\alpha_1, \cdots, \alpha_m$ 线性表出。 □

定义 3.1.5. 一个 K-向量空间 V 称为有限生成，如果存在有限个向量 $\alpha_1, \cdots, \alpha_m \in V$，使

$$V = \{ \lambda_1 \alpha_1 + \cdots + \lambda_m \alpha_m \mid \lambda_i \in K \}.$$

□

线性代数主要研究有限生成向量空间，无限生成的向量空间通常需要分析工具，属于泛函分析研究的对象。前述例 3.1.3，例 3.1.5，例 3.1.8 中的向量空间都是有限生成的。而例 3.1.6 中的 $K[x_1, \ldots, x_n]$ 和例 3.1.7 中的空间 V 都不是有限生成的，但它们的子空间 V_d 和 W 都是有限生成的。需要特别指出，K-向量空间 V 是否有限生成与 K 有密切关系：在例 3.1.4 中，$\mathbb{C}$ 作为 $\mathbb{R}$-向量空间是有限生成的（由 $1, i$ 生成），但 $\mathbb{C}$ 作 $\mathbb{Q}$-向量空间则不是有限生成的。

定义 3.1.6：设 V 是一个 K-向量空间，向量 $\alpha_1, \cdots, \alpha_m \in V$ 称为 K-线性相关（简称线性相关），如果存在不全为零的 $\lambda_1, \cdots, \lambda_m \in K$ 使 $\lambda_1 \alpha_1 + \cdots + \lambda_m \alpha_m = 0$。否则，$\alpha_1, \cdots, \alpha_m$ 称为线性无关。

□

引理 3.1.2：设 $\alpha_1, \ldots, \alpha_m \in V$ 是一组向量，则

(1) $\alpha_1, \ldots, \alpha_m$ 线性相关的充要条件是存在 α_i 可由其它向量线性表出。

(2) 如果 $\alpha_1, \cdots, \alpha_m$ 线性无关，$\beta \in V$ 是任意向量，则 $\alpha_1, \cdots, \alpha_m, \beta$ 线性相关当且仅当存在唯一的一组 $\lambda_1, \cdots, \lambda_m$ 使 $\beta = \lambda_1 \alpha_1 + \cdots + \lambda_m \alpha_m$

证明：留作练习。

□

下面重要引理是定义向量空间维数的基石，它的证明本质上依赖于下一节线性方程组理论的高斯消元法。

<u>引理3.1.3</u>. 设 V 是 K-向量空间，$\beta_1,\dots,\beta_s$, $\alpha_1,\dots,\alpha_r$ 是 V 中的两组向量。如果每个 β_i 可由 $\alpha_1,\dots,\alpha_r$ 线性表出，则

(1) 当 $s>r$ 时，$\beta_1,\dots,\beta_s$ 一定<u>线性相关</u>；

(2) 如果 $\beta_1,\dots,\beta_s$ <u>线性无关</u>，则 $s\le r$.

证明：(1) 与 (2) 逆否等价，所以只需证明 (1)。由条件得

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

由高斯消元法，<u>齐次线性方程组</u>

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

当 $s>r$ 时<u>有非零解</u> $(\lambda_1,\lambda_2,\dots,\lambda_s)$，计算得：

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

$= (a_{11}\lambda_1+\cdots+a_{s1}\lambda_s)\alpha_1 + (a_{12}\lambda_1+\cdots+a_{s2}\lambda_s)\alpha_2 + \cdots + (a_{1r}\lambda_1+\cdots+a_{sr}\lambda_r)\alpha_r = 0$

所以 $\beta_1,\dots,\beta_s$ 线性相关

□

定义3.1.7. 设 V 是一个 K-向量空间，$\alpha_1, \dots, \alpha_n \in V$ 称为 V 的一组基，如果它满足如下条件：

(1) $\alpha_1, \dots, \alpha_n$ 线性无关；

(2) V 中每个向量可由 $\alpha_1, \dots, \alpha_n$ 线性表出。

□

定理3.1.1. 设 V 是一个有限生成的 K-向量空间，则

(1) 在 V 中存在一组基；

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

(3) 如果 $\alpha_1, \dots, \alpha_t \in V$ 线性无关，$\alpha_1, \dots, \alpha_t$ 可扩充为 V 的一组基。

证明：(1) V 是有限生成的，令 $e_1, \dots, e_n$ 是 V 的一组其向量个数最少的生成元，则 $e_1, \dots, e_n$ 线性无关（由引理3.1.2中的(1)）。它又是 V 的生成元，所以 $e_1, \dots, e_n$ 是 V 的一组基。

(2) 由基的定义，$\alpha_1, \dots, \alpha_r$ 与 $\beta_1, \dots, \beta_s$ 互相线性表出，由引理3.1.3，得 $r=s$。

(3). 如果 $\alpha_1, \dots, \alpha_t \in V$ 生成 V，则 $\alpha_1, \dots, \alpha_t$ 已经是 V 的一组基。如果 $\alpha_1, \dots, \alpha_t$ 不生成 V，则存在 $\alpha_{t+1} \in V$，但 α_{t+1} 不能由 $\alpha_1, \dots, \alpha_t$ 线性表出，所以 $\alpha_1, \dots, \alpha_t, \alpha_{t+1}$ 线性无关（由引理3.1.2中的(2)）。对 $\alpha_1, \dots, \alpha_t, \alpha_{t+1}$ 继续上述操作可将 $\alpha_1, \dots, \alpha_t$ 扩充成 V 的一组基

□

定义3.1.8：设 V 是 K-向量空间，V 中任意一组基的向量个数称为 V 的维数，记为 $\dim_K(V)$。如果 $V=\{0\}$，则 $\dim_K(V)=0$。

□

例3.1.9：$\dim_K(K^n)=n$，$\dim_{\mathbb{R}}(\mathbb{C})=2$，$\dim_K(V_d)=C_{d+n-1}^{d}$.

□

例3.1.10. $\dim_{\mathbb{Q}}(\mathbb{C})=+\infty$, $\dim_K(K[x])=+\infty$, 但是 $K[x]$ 的子空间 $K[x]_d:=\{f(x)\in K[x] \mid \deg f(x)\le d\}$ 都是有限维 K-向量空间 $\dim_K(K[x]_d)=d+1$.

□

如果 V 是 n 维 K-向量空间，则存在一组基 $e_1,e_2,\cdots,e_n\in V$ 使得 V 中的每个向量 x 都可唯一地由 $e_1,\cdots,e_n$ 线性表出。即 $V=\{x=x_1e_1+x_2e_2+\cdots+x_ne_n \mid x_i\in K\}$, $x_1,x_2,\cdots,x_n$ 由 x 唯一确定。所以，$e_1,e_2,\cdots,e_n$ 确定了 n 个函数

$$e_i^*: V\longrightarrow K, \quad (1\le i\le n)$$

使得：如果 $x=x_1e_1+x_2e_2+\cdots+x_ne_n$，则 $e_i^*(x)=x_i$.

引理3.1.4: 由 V 的一组基 $e_1,e_2,\cdots,e_n$ 确定的函数 $e_i^*: V\to K$, $(1\le i\le n)$, 满足条件：$\forall\, x,y\in V$, $\lambda\in K$,

$$e_i^*(x+y)=e_i^*(x)+e_i^*(y), \quad e_i^*(\lambda x)=\lambda e_i^*(x).$$

证明：如果 $x=x_1e_1+\cdots+x_ie_i+\cdots+x_ne_n$, $y=y_1e_1+\cdots+y_ie_i+\cdots+y_ne_n$, 则 $x+y=(x_1+y_1)e_1+\cdots+(x_i+y_i)e_i+\cdots+(x_n+y_n)e_n$, $\lambda x=(\lambda x_1)e_1+\cdots+(\lambda x_i)e_i+\cdots+\lambda x_ne_n$.

所以 $e_i^*(x+y)=x_i+y_i=e_i^*(x)+e_i^*(y)$, $e_i^*(\lambda x)=\lambda x_i=\lambda e_i^*(x)$.

□

定义3.1.9. 函数 $l: V\to K$ 称为线性函数，如果 $\forall\, x,y\in V$, $\forall\,\lambda\in K$, 有 $l(x+y)=l(x)+l(y)$, $l(\lambda x)=\lambda l(x)$。

如果 l_1,l_2 是 V 上的两个线性函数，则

$$l_1+l_2: V\longrightarrow K, \quad x\mapsto l_1(x)+l_2(x),$$

$$\lambda l_1: V\longrightarrow K, \quad x\mapsto \lambda l_1(x)$$

也是线性函数 □

例3.1.11. 设V是n维K-向量空间，$e_1,\cdots,e_n\in V$是一组基.

则 (1) $e_i^*: V\longrightarrow K$，$(1\le i\le n)$，是线性函数，且$e_i^*(e_j)=0$ $(i\ne j)$，$e_i^*(e_i)=1$.

(2) 对任意给定的常数$a_1,a_2,\cdots,a_n\in K$，则函数

$$l_a: V\longrightarrow K,\quad x=x_1e_1+\cdots+x_ne_n\longmapsto l(x)=a_1x_1+\cdots+a_nx_n$$

是线性函数，且$l_a(e_i)=a_i$，$(1\le i\le n)$.

(3). $l_a=a_1e_1^*+\cdots+a_ne_n^*$ □.

定义3.1.10：设V^*表V上所有线性函数$l: V\longrightarrow K$的集合. V^*关于函数的"加法"和数乘法是一个K-向量空间，称为V的对偶空间.

□

引理3.1.5. 设V是n维K-向量空间，$e_1,\cdots,e_n\in V$是一组基，则$e_1^*,\cdots,e_n^*\in V^*$是$V^*$的一组基，即$\dim_K(V^*)=n$.

证明：(1) $e_1^*,\cdots,e_n^*$线性无关. 如果$\lambda_1e_1^*+\cdots+\lambda_ne_n^*=0$. 即$l:=\lambda_1e_1^*+\cdots+\lambda_ne_n^*: V\longrightarrow K$是零函数. 则

$$l(x)=\lambda_1e_1^*(x)+\cdots+\lambda_ne_n^*(x)=0,\quad \forall\ x\in V.$$

特别，$\lambda_i=l(e_i)=0$，$(1\le i\le n)$. 即 $\cancel{0=l(e_i)=\lambda_1e_1^*(e_i)+\cdots+\lambda_ie_i^*(e_i)+\cdots+\lambda_ne_n^*(e_i)}$ $=\lambda_i$

(2) $\forall\ l\in V^*$，即$l: V\longrightarrow K$是线性函数. 令$a_i=l(e_i)\in K$，则$l=a_1e_1^*+\cdots+a_ne_n^*$.

□.

定义3.1.11：设$e_1,e_2,\cdots,e_n\in V$是一组基，$\forall\ x\in V$，则数组

$$[e_1^*(x),e_2^*(x),\cdots,e_n^*(x)]:=\begin{pmatrix}e_1^*(x)\\ e_2^*(x)\\ \vdots\\ e_n^*(x)\end{pmatrix}\in K^n$$

称为x在基$e_1,\cdots,e_n$下的坐标. □

引入记号： $x = e_1^*(x)\, e_1 + \cdots + e_n^*(x)\, e_n$

$$\xlongequal{} (e_1, \cdots, e_n) \begin{pmatrix} e_1^*(x) \\ \vdots \\ e_n^*(x) \end{pmatrix} \xlongequal{} (e_1, \cdots, e_n)[e_1^*(x), \cdots, e_n^*(x)].$$

引理3.1.6： 设 V 是 n 维 K-向量空间，$e_1, \cdots, e_n \in V$ 是一组基.

则映射 $V \xrightarrow{\phi_e} K^n$， $x \longmapsto [e_1^*(x), \cdots, e_n^*(x)]$

是一个双射. 且满足： $\phi_e(x+y) = \phi_e(x) + \phi_e(y)$， $\phi_e(\lambda x) = \lambda \phi_e(x)$.

证明：(1) ϕ_e 是满射： $\forall\ [a_1, \cdots, a_n] \in K^n$， 令 $x = a_1 e_1 + \cdots + a_n e_n \in V$.

则 $e_i^*(x) = e_i^*(a_i e_i) = a_i$. 故 $\phi_e(x) = [e_1^*(x), \cdots, e_n^*(x)] = [a_1, \cdots, a_n]$.

(2) ϕ_e 是单射： $\forall\ x = (e_1, \cdots, e_n)[e_1^*(x), \cdots, e_n^*(x)] \in V$,

$y = (e_1, \cdots, e_n)[e_1^*(y), \cdots, e_n^*(y)] \in V$， 如果 $\phi_e(x) = \phi_e(y)$， 则

$[e_1^*(x), \cdots, e_n^*(x)] = [e_1^*(y), \cdots, e_n^*(y)]$。所以， $x = y$. 故 ϕ_e 是单射.

(3) 因为 $x+y = (e_1, \cdots, e_n)[e_1^*(x)+e_1^*(y), \cdots, e_n^*(x)+e_n^*(y)]$， 故

$$\phi_e(x+y) = [e_1^*(x)+e_1^*(y), \cdots, e_n^*(x)+e_n^*(y)] = \phi_e(x) + \phi_e(y).$$

同理可证 $\phi_e(\lambda x) = \lambda \phi_e(x)$， $\forall\ \lambda \in K$.

$\square$.

所以 V 上的所有线性函数和映射均可用向量的坐标来表达和计算. 例如，设 $l: V \to K$ 是一个线性函数，令 $a_i := l(e_i) \in K$， 则

$$l(x) = a_1 e_1^*(x) + a_2 e_2^*(x) + \cdots + a_n e_n^*(x) \xlongequal{} (a_1, \cdots, a_n)[e_1^*(x), \cdots, e_n^*(x)].$$

习题3.1

1. 设 V 是 K-向量空间，$x, y \in V$，$\lambda, \mu \in K$. 证明：

(1) $0\cdot x = 0$，(2) $\lambda\cdot 0 = 0$；(3) $(-1)\cdot x = -x$，(4) $\lambda\cdot(-x) = -\lambda\cdot x$

(5) $\lambda\cdot(x-y) = \lambda\cdot x - \lambda\cdot y$，(6) $(\lambda-\mu)\cdot x = \lambda\cdot x - \mu\cdot x$，(7) 若 $\lambda\cdot x = 0$，则 $\lambda = 0$ 或 $x = 0$.

2. 设 V 是 K-向量空间，$V_i \subset V$，$(i \in I)$，是若干个（可能无限个）子空间，证明它们的交集 $\bigcap_{i\in I} V_i$ 也是 V 的子空间.

3. 设 $V_1, V_2, \cdots, V_m$ 是 V 的子空间. 证明：它们"和"

$$V_1 + V_2 + \cdots + V_m := \{x_1 + x_2 + \cdots + x_m \mid \forall\, x_i \in V_i\}$$

也是 V 的子空间.

4. 设 V 是有限维向量空间，$W \subset V$ 是 V 的任意子空间. 证明：

$$\dim_K(W) \le \dim_K(V),$$

且等号成立的充要条件是 $W = V$.

5. 设 V_1, V_2 是 V 的子空间，证明：它们的并 $V_1 \cup V_2$ 不是 V 的子空间，除非 $V_1 \subseteq V_2$ 或 $V_2 \subseteq V_1$.

6. 设 $V = \mathbb{R}^2$ 是2维行向量空间，证明：(1) $V_1 = \{(x, 0) \mid \forall x \in \mathbb{R}\}$ 和 $V_2 = \{(0, y) \mid \forall y \in \mathbb{R}\}$ 是 $\mathbb{R}^2$ 的子空间；(2) $V = V_1 + V_2$.
(3) $\mathbb{R}^2 \neq V_1 \cup V_2$.

7. 设 $V = \{$连续函数 $\mathbb{R} \xrightarrow{f} \mathbb{R}\}$ 是 $\mathbb{R}$ 上所有连续函数关于函数加法和数乘运算构成的 $\mathbb{R}$-向量空间. 证明下列向量组在 V 中线性无关.

(1) $\sin x$，$\cos x$；(2) 1，$\sin x$，$\cos x$

(3) $\sin x$，$\sin 2x$，$\cdots$，$\sin nx$.（n 是任意正整数）.

8. 设 V 是 n 维 K-向量空间，$l: V\to K$ 是一非零线性函数。

试证明：(1) $V_l=\{x\in V \mid l(x)=0\}$ 是 V 的子空间；

(2) $\dim_K(V_l)=n-1$.

9. 设 $\dim_K(V)=n\geq 2$，$l_i: V\to K$ $(i=1,2)$ 是两个不全为零的线性函数。令 $W=\{x\in V \mid l_1(x)=0, l_2(x)=0\}$.

试证明：$\dim_K(W)\geq n-2$，等号成立当且仅当 l_1, l_2 在 V^* 中线性无关.

10. 设 $\dim_K(V)=n>0$，$x\in V$ 是非零向量，证明：

(1) $V_x^*=\{l\in V^* \mid l(x)=0\}$ 是 V^* 的子空间，(2) $\dim_K(V_x^*)=n-1$.

11. 设 V_1, V_2 是 V 的子空间，$\alpha_1,\cdots,\alpha_m\in V_1\cap V_2$ 是一组基。令 $\beta_1,\cdots,\beta_t\in V_1$，$\gamma_1,\cdots,\gamma_s\in V_2$ 使得 $\alpha_1,\cdots,\alpha_m,\beta_1,\cdots,\beta_t$ 是 V_1 的一组基，$\alpha_1,\cdots,\alpha_m,\gamma_1,\cdots,\gamma_s$ 是 V_2 的一组基。试证明：

(1) 向量组 $\alpha_1,\cdots,\alpha_m,\beta_1,\cdots,\beta_t,\gamma_1,\cdots,\gamma_s$ 是子空间 V_1+V_2 的一组基.

(2) $\dim_K(V_1+V_2)=\dim_K(V_1)+\dim_K(V_2)-\dim_K(V_1\cap V_2)$.

12. 设 $e_1,\cdots,e_n$ 是 V 的一组基，$\alpha_i=(e_1,\cdots,e_n)[e_1^*(\alpha_i),\cdots,e_n^*(\alpha_i)]$，$1\leq i\leq m$ 是 V 中一组坐标分别为 $A^{(i)}=[e_1^*(\alpha_i),\cdots,e_n^*(\alpha_i)]\in K^n$，$1\leq i\leq m$ 的向量。证明

(1) $\alpha_1,\cdots,\alpha_m$ 线性相关 $\Leftrightarrow$ $A^{(1)},\cdots,A^{(m)}$ 线性相关

(2) $\alpha_1,\cdots,\alpha_m$ 线性无关 $\Leftrightarrow$ $A^{(1)},\cdots,A^{(m)}$ 线性无关.

§3.2. 线性方程组

数学与实际应用中的许多问题都可以归结为求解下述的“线性方程组”：

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

其中 $x_1,x_2,\cdots,x_n$ 是未知元，$a_{ij}\in K$（K是一个域）称为方程组的系数，$a_{ij},b_i\in K$ 一般都是固定的数（或元素）。下述的线性方程组称为与方程组$(*)$相伴的“齐次线性方程组”.

$$
(*)^{齐}\quad\begin{cases}
a_{11}x_1+a_{12}x_2+\cdots+a_{1n}x_n=0\\
a_{21}x_1+a_{22}x_2+\cdots+a_{2n}x_n=0\\
\qquad\vdots\\
a_{m1}x_1+a_{m2}x_2+\cdots+a_{mn}x_n=0.
\end{cases}
$$

方程组$(*)$和$(*)^{齐}$也可写成列向量空间K^m中的线性组合形式.

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

$$
(*)^{齐}:\quad x_1\begin{pmatrix}a_{11}\\a_{21}\\\vdots\\a_{m1}\end{pmatrix}+x_2\begin{pmatrix}a_{12}\\a_{22}\\\vdots\\a_{m2}\end{pmatrix}+\cdots+x_n\begin{pmatrix}a_{1n}\\a_{2n}\\\vdots\\a_{mn}\end{pmatrix}=\begin{pmatrix}0\\0\\\vdots\\0\end{pmatrix}.
$$

为了更进一步简化方程组的表达，我们引入矩阵的概念.

定义3.2.1. 形如

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

的表格称为 m 行, n 列的矩阵（简称 $m\times n$ 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

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

称为 A 的第 j 列（或 A 的第 j 个列向量）. 矩阵 A 通常简记为

$$A=(a_{ij})_{m\times n}, \ (a_{ij} \text{表示位于第}i\text{行,第}j\text{列位置的元素})$$

如果 $a_{ij}\in K$, 则称 $A=(a_{ij})_{m\times n}$ 是系数在 K 中的矩阵。两个矩阵 $A=(a_{ij})_{m\times n}$, $B=(b_{ij})_{m'\times n'}$ 相等当且仅当 $m=m'$, $n=n'$, 且

$$a_{ij}=b_{ij}, \quad (1\le i\le m, 1\le j\le n).$$

□.

符号约定: 矩阵 $A=(a_{ij})_{m\times n}$ 也可写成

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

的形式（实际是矩阵的一种分块形式），即行向量形式和列向量形式.

定义3.2.2: 矩阵

$$A=\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} \text{，} \tilde{A}=(A,b)=\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} & b_1 \\ a_{21} & a_{22} & \cdots & a_{2n} & b_2 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} & b_m \end{pmatrix}$$

分别称为方程组的系数矩阵和增广矩阵. □

所以线性方程组 (*) 和 $(*)^{齐}$ 可以分别表示成：

$$(*):\quad x_1A^{(1)}+x_2A^{(2)}+\cdots+x_nA^{(n)}=b,\quad (b=[b_1,\cdots,b_m])$$

$$(*)^{齐}:\quad x_1A^{(1)}+x_2A^{(2)}+\cdots+x_nA^{(n)}=0$$

如果用 K^m 表示 m 维列向量空间，则方程组 (*) 是否有解可表述成：

(1) (*) 有解 $\Leftrightarrow$ b 在 K^m 可由 $A^{(1)},A^{(2)},\cdots,A^{(n)}$ 线性表出.

(2) $(*)^{齐}$ 有非零解 $\Leftrightarrow$ $A^{(1)},A^{(2)},\cdots,A^{(n)}$ 在 K^m 中线性相关.

(3) $(*)^{齐}$ 无非零解 $\Leftrightarrow$ $A^{(1)},A^{(2)},\cdots,A^{(n)}$ 在 K^m 中线性无关.

定理 3.2.1：令 K^n 表示 n 维列向量空间，则

$$W_{(*)^{齐}}=\{[\lambda_1,\lambda_2,\cdots,\lambda_n]\in K^n \mid \lambda_1A^{(1)}+\lambda_2A^{(2)}+\cdots+\lambda_nA^{(n)}=0\}$$

是 K^n 的一个子空间。如果 $\eta=[\eta_1,\eta_2,\cdots,\eta_n]\in K^n$ 是 (*) 的一个解，则

$$W_{(*)}=\{[\lambda_1,\lambda_2,\cdots,\lambda_n]\in K^n \mid \lambda_1A^{(1)}+\lambda_2A^{(2)}+\cdots+\lambda_nA^{(n)}=b\}=\eta+W_{(*)^{齐}}$$

证明：不难直接验证：$W_{(*)^{齐}}$ 是 K^n 的一个子空间.

如果 $\eta=[\eta_1,\eta_2,\cdots,\eta_n]$ 是 (*) 的一个解，$\forall\ \eta+a\in\eta+W_{(*)^{齐}}$，其中 $a=[a_1,a_2,\cdots,a_n]\in W_{(*)^{齐}}$，则

$$(\eta_1+a_1)A^{(1)}+(\eta_2+a_2)A^{(2)}+\cdots+(\eta_n+a_n)A^{(n)}=\eta_1A^{(1)}+\eta_2A^{(2)}+\cdots+\eta_nA^{(n)}=b$$

即 $\eta+a\in W_{(*)}$. 反之，$\forall\ \lambda=[\lambda_1,\lambda_2,\cdots,\lambda_n]\in W_{(*)}$，则 $\lambda-\eta$ 是 $(*)^{齐}$ 的一个解. 即 $\lambda-\eta\in W_{(*)^{齐}}$ 使 $\lambda=\eta+(\lambda-\eta)\in\eta+W_{(*)^{齐}}$

□

线性方程组理论通常需要回答下面几个问题：(1) 方程组 (*) 是否有解？(2) 如果 (*) 有解，何时有唯一解？(3) 如果 (*) 有唯一解，

该解的公式是什么？(4) 如果(*)有无限多个解，能否找出有限个解 $\alpha_1,\cdots,\alpha_k\in K^n$ 使其它解可由 $\alpha_1,\cdots,\alpha_k$ 唯一表出？
下面介绍一种算法（高斯消元法），它给出了问题(1),(2),(4)的完整答案。

该算法由对方程组的几个基本操作（称为初等变换）组成。初等变换(I)是指将方程组中的第i个方程与第j个方程交换位置；初等变换(II)是指将方程组中的第i个方程乘一个常数 $c\in K$ 加到第j个方程；初等变换(III)是指将方程组的第i个方程乘以非零常数 $c\in K$。

$\cancel{l_i = a_{i1}x_1 + a_{i2}x_2 + \cdots + a_{in}x_n}$

上述对线性方程组(*)的初等变换可以用对它的增广矩阵 $\tilde{A}=(A,b)$ 的行向量的初等变换来描述：令

$$\tilde{A}_{(i)}=(a_{i1},a_{i2},\cdots,a_{in},b_i).$$

则关于方程组(*)的三类初等变换可图示如下：

$$[\tilde{A}_{(1)},\cdots,\tilde{A}_{(i)},\cdots,\tilde{A}_{(j)},\cdots,\tilde{A}_{(m)}] \xleftrightarrow[(I)]{\text{初等变换}} [\tilde{A}_{(1)},\cdots,\tilde{A}_{(j)},\cdots,\tilde{A}_{(i)},\cdots,\tilde{A}_{(m)}]$$

$$[\tilde{A}_{(1)},\cdots,\tilde{A}_{(i)},\cdots,\tilde{A}_{(j)},\cdots,\tilde{A}_{(m)}] \xleftrightarrow[(II)]{\text{初等变换}} [\tilde{A}_{(1)},\cdots,\tilde{A}_{(i)},\cdots,\tilde{A}_{(j)}+c\tilde{A}_{(i)},\cdots,\tilde{A}_{(m)}]$$

$$[\tilde{A}_{(1)},\cdots,\tilde{A}_{(i)},\cdots,\tilde{A}_{(m)}] \xleftrightarrow[(III)]{\text{初等变换}} [\tilde{A}_{(1)},\cdots,c\tilde{A}_{(i)},\cdots,\tilde{A}_{(m)}].$$

引理3.2.1：如果方程组(*)'是由方程组(*)经有限次初等变换得到，则 $x=[x_1,x_2,\cdots,x_n]$ 是(*)'的解当且仅当 $x=[x_1,x_2,\cdots,x_n]$ 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 $x=[x_1,x_2,\cdots,x_n]$ 是 (*)' 的解当且仅当 $x=[x_1,x_2,\cdots,x_n]$ 是 (*) 的解。对第(I),(II)类初等变换，这是显然的。对第(III)类初等变换，令

$$l_k(x)=a_{k1}x_1+a_{k2}x_2+\cdots+a_{kn}x_n,$$

如果 $x=[x_1,x_2,\cdots,x_n]$ 是 (*)' 的解，则

$$l_k(x)=b_k,\ (k\neq j),\quad c\,l_i(x)+l_j(x)=cb_i+b_j.$$

故 $l_j(x)=b_j$，$x=[x_1,\cdots,x_n]$ 也是 (*) 的解。反之亦然

□.

定理 3.2.2（高斯消元法）。如果方程组 (*) 的系数矩阵 $A\neq 0$，则经过若干次第(I)类和第(III)类初等变换，(*) 可化简为下列的"阶梯型"方程组 $\overline{(*)}$：

$$\begin{array}{l} \bar{l}_1(x) \\ \bar{l}_2(x) \\ \vdots \\ \bar{l}_r(x) \\ \bar{l}_{r+1}(x) \\ \vdots \\ \bar{l}_m(x) \end{array}\left\{\begin{array}{l}
\bar{a}_{1i_1}x_{i_1}+\bar{a}_{1i_1+1}x_{i_1+1}+\bar{a}_{1i_1+2}x_{i_1+2}+\bar{a}_{1i_1+3}x_{i_1+3}+\cdots+\bar{a}_{1n}x_n=\bar{b}_1\\
\qquad\bar{a}_{2i_2}x_{i_2}+\bar{a}_{2i_2+1}x_{i_2+1}+\bar{a}_{2i_2+2}x_{i_2+2}+\bar{a}_{2i_2+3}x_{i_2+3}+\cdots+\bar{a}_{2n}x_n=\bar{b}_2\\
\qquad\qquad\ddots\qquad\qquad\vdots\\
\qquad\qquad\bar{a}_{ri_r}x_{i_r}+\bar{a}_{ri_r+1}x_{i_r+1}+\bar{a}_{ri_r+2}x_{i_r+2}+\cdots+\bar{a}_{rn}x_n=\bar{b}_r\\
0x_1+0x_2+\cdots+0x_r+0x_{r+1}+0x_{r+2}+0x_{r+3}+\cdots+0x_n=\bar{b}_{r+1}\\
\vdots\qquad\qquad\vdots\\
0x_1+0x_2+\cdots0x_r+0x_{r+1}+0x_{r+2}+0x_{r+3}+\cdots+0x_n=\bar{b}_m
\end{array}\right.$$

其中 $0<i_1<i_2<\cdots<i_r$，$\bar{a}_{1i_1}\neq 0,\ \bar{a}_{2i_2}\neq 0,\cdots,\bar{a}_{ri_r}\neq 0$（通过第(II)类初等变换可设 $x_{i_1},x_{i_2},\cdots,x_{i_r}$ 的系数均为1）。

证明：对不定元 $x_1,x_2,\cdots,x_n$ 的个数 n 做归纳法：当 $n=1$，结论显然（因为 $A\neq 0$）。设定理对不定元个数小于 n 的方程组成立。无妨设 $A^{(1)}\neq 0$（系数矩阵 A 的第1列），通过第(I)类初等变换，

可设 $a_{11}\neq 0$。为方便表述，令 $l_i(x)=a_{i1}x_1+a_{i2}x_2+\cdots+a_{in}x_n$.

则线性方程组 (*) 通过第(III)类初等变换可化为:

$$\begin{cases} l_1(x)=a_{11}x_1+a_{12}x_2+a_{13}x_3+\cdots+a_{1n}x_n=b_1 \\ l_2(x)-\frac{a_{21}}{a_{11}}l_1(x)=0x_1+(a_{22}-\frac{a_{21}}{a_{11}})x_2+\cdots+(a_{2n}-\frac{a_{21}}{a_{11}})x_n=b_2-\frac{a_{21}}{a_{11}}b_1 \\ \vdots \\ l_m(x)-\frac{a_{m1}}{a_{11}}l_1(x)=0x_1+(a_{m2}-\frac{a_{m1}}{a_{11}})x_2+\cdots+(a_{mn}-\frac{a_{m1}}{a_{11}})x_n=b_m-\frac{a_{m1}}{a_{11}}b_1 \end{cases}$$

令 $\bar{l}_1(x)=l_1(x)$, $l_2'(x)=l_2(x)-\frac{a_{21}}{a_{11}}l_1(x)$, $\cdots$, $l_m'(x)=l_m(x)-\frac{a_{m1}}{a_{11}}l_1(x)$

$\bar{b}_1=b_1$, $b_2'=b_2-\frac{a_{21}}{a_{11}}b_1$, $\cdots$, $b_m'=b_m-\frac{a_{m1}}{a_{11}}b_1$.

则上述方程组可写成

$$(*)'\quad \begin{cases} \bar{l}_1(x)=\bar{a}_{11}x_1+\bar{a}_{12}x_2+\cdots+\bar{a}_{1n}x_n=\bar{b}_1 \\ l_2'(x)=\qquad a_{22}'x_2+\cdots+a_{2n}'x_n=b_2' \\ \vdots \\ l_m'(x)=\qquad a_{m2}'x_2+\cdots+a_{mn}'x_n=b_m' \end{cases}$$

对方程组:
$$\begin{cases} l_2'(x)=a_{22}'x_2+\cdots+a_{2n}'x_n=b_2' \\ \vdots \\ l_m'(x)=a_{m2}'x_2+\cdots+a_{mn}'x_n=b_m' \end{cases}$$

应用归纳假设即得结论.

□

由证明过程，我们有下列观察:

(1) 方程组(*)'中的每个方程 $\bar{l}_i(x)=\bar{b}_i$ 都是(*)中方程 $l_1(x)=b_1$, $l_2(x)=b_2$, $\cdots$, $l_m(x)=b_m$ 的线性组合. 即存在常数 $\lambda_{i1},\cdots,\lambda_{im}\in K$ 使得

$$\bar{l}_i(x)=\lambda_{i1}l_1(x)+\lambda_{i2}l_2(x)+\cdots+\lambda_{im}l_m(x)$$
$$\bar{b}_i=\lambda_{i1}b_1+\lambda_{i2}b_2+\cdots+\lambda_{im}b_m$$

反之亦然.

(2) 在方程组(*)：$l_1(x)=b_1,\ l_2(x)=b_2,\ \cdots,\ l_m(x)=b_m$ 中存在 r 个方程（不妨设 $l_1(x)=b_1,\cdots,l_r(x)=b_r$）使得线性函数（或：线性无关）

$\{l_{r+1}(x),\cdots,l_m(x)\}$ 可由 $\{l_1(x),\cdots,l_r(x)\}$ 线性表出（但 $\{b_{r+1},\cdots,b_m\}$ 则未必由 $\{b_1,\cdots,b_r\}$ 线性表出）. 即：$\forall\, i>r$，存在常数 $\lambda_{i1},\cdots,\lambda_{ir}\in K$ 使得 $l_i(x)=\lambda_{i1}l_1(x)+\cdots+\lambda_{ir}l_r(x)$，$\bar{b}_i=b_i-(\lambda_{i1}b_1+\cdots+\lambda_{ir}b_r)$

(3) $r\le\min\{m,n\}$.

推论 3.2.1：(1) (*) 有解 $\Leftrightarrow$ $\overline{(*)}$ 有解 $\Leftrightarrow$ $\bar{b}_{r+1}=\cdots=\bar{b}_m=0$.

(2) 若 (*) 有解（即 $\bar{b}_{r+1}=\cdots=\bar{b}_m=0$），适当调整不定元 $x_1,\cdots,x_n$ 的次序（使阶梯型中的 $i_1,i_2,\cdots,i_r$ 为 $1,2,\cdots,r$），则 $x=[x_1,\cdots,x_n]\in K^n$ 是 (*) 一组解的充要条件是：

$$\begin{cases} x_1=d_1+c_{11}x_{r+1}+\cdots+c_{1\,n-r}x_n\\ x_2=d_2+c_{21}x_{r+1}+\cdots+c_{2\,n-r}x_n\\ \vdots\\ x_r=d_r+c_{r1}x_{r+1}+\cdots+c_{r\,n-r}x_n\end{cases}$$

其中 $d_1,d_2,\cdots,d_r$，$c_{ij}\in K$ 是（可计算的）常数。特别，

$r<n \Leftrightarrow$ (*) 有无穷多组解（当 K 是无限域时）.

(*) 有唯一解 $\Leftrightarrow r=n$.

(3) 齐次线性方程组 $(*)^h$ 有非零解 $\Leftrightarrow r<n$.

特别，当 $m<n$ 时，$(*)^h$ 有无穷多组解（当 K 是无限域时）.

□

例 3.2.1：线性方程组.
$$\begin{cases} l_1(x)=x_1+x_2+2x_3+2x_4=b_1\\ l_2(x)=3x_1+5x_2+6x_3-4x_4=b_2\\ l_3(x)=4x_1+5x_2-2x_3+3x_4=b_3\\ l_4(x)=3x_1+8x_2+24x_3-19x_4=b_4\end{cases}$$

何时有解？如有解，求出它的一般解.

解 消去后面三个方程中的 x_1 得

$$l_2'(x)=l_2(x)-3l_1(x)=2x_2+12x_3-10x_4=b_2-3b_1$$
$$l_3'(x)=l_3(x)-4l_1(x)=x_2+6x_3-5x_4=b_3-4b_1$$
$$l_4'(x)=l_4(x)-3l_1(x)=5x_2+30x_3-25x_4=b_4-3b_1$$

消去后两个方程中的 x_2，可得"阶梯型"方程组如下：

$$\begin{cases}\bar{l}_1(x)=l_1(x)=x_1+x_2-2x_3+2x_4=b_1=\bar{b}_1\\ \bar{l}_2(x)=\frac{1}{2}l_2'(x)=\quad x_2+6x_3-5x_4=\frac{1}{2}b_2-\frac{3}{2}b_1=\bar{b}_2\\ \bar{l}_3(x)=l_3'(x)-\bar{l}_2(x)=0x_1+0x_2+0x_3+0x_4=b_3-\frac{5}{2}b_1-\frac{1}{2}b_2=\bar{b}_3\\ \bar{l}_4(x)=l_4'(x)-5\bar{l}_2(x)=0x_1+0x_2+0x_3+0\cdot x_4=b_4+\frac{9}{2}b_1-\frac{5}{2}b_2=\bar{b}_4\end{cases}$$

故当 $b_3=\frac{5}{2}b_1+\frac{1}{2}b_2$，$b_4=-\frac{9}{2}b_1+\frac{5}{2}b_2$ 时，方程组有解。

如果 $x=[x_1,x_2,x_3,x_4]$ 是它的解，则 $\Longleftrightarrow$

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

令 $x_3=x_4=0$，可得方程组的一个特解 $\alpha=[\frac{5}{2}b_1-\frac{1}{2}b_2,-\frac{3}{2}b_1+\frac{1}{2}b_2,0,0]$.
与它相伴的齐次线性方程组的解空间是

$$W=\left\{x=[x_1,x_2,x_3,x_4]\in K^4 \;\middle|\; \begin{array}{l}x_1=8x_3-7x_4\\ x_2=-6x_3+5x_4\end{array}\right\}.$$

令 $x_3=1$，$x_4=0$ 得 $\beta_1=[8,-6,1,0]\in W$，令 $x_3=0$，$x_4=1$ 得 $\beta_2=[-7,5,0,1]\in W$. 可证明 β_1,β_2 线性无关，且 $\forall\ x=[x_1,x_2,x_3,x_4]\in W$，有 $x=x_3\beta_1+x_4\beta_2$. 即 β_1,β_2 是 W 的一组基. 即 $W=\{\lambda_1\beta_1+\lambda_2\beta_2\mid\lambda_1,\lambda_2\in K\}$，方程组的一般解（或称通解）可表示成

$$x=\alpha+\lambda_1\beta_1+\lambda_2\beta_2,\quad\forall\ \lambda_1,\lambda_2\in K.$$

□

注意：例3.2.1中的4个方程中，$l_1(x),l_2(x)$ 线性无关，$l_3(x)$，$l_4(x)$ 可由 $l_1(x),l_2(x)$ 线性表出：$l_3(x)=\frac{5}{2}l_1(x)+\frac{1}{2}l_2(x)$，$l_4(x)=-\frac{9}{2}l_1(x)+\frac{5}{2}l_2(x)$.

高斯消元法中出现的正整数 r 显然在解方程时意义重大. 一个自然的问题是: r 是否方程组的"不变量"? 即: r 是否与初等变换无关? 下面我们就来证明它确实是方程组系数矩阵的重要不变量.(方程组由系数矩阵 A 与 $\bar{A}=(A,b)$ 唯一确定, 所以它就是方程组的不变量).

定义 3.2.3: 对任意矩阵 $A=[A_{(1)},\cdots,A_{(m)}]=(A^{(1)},\cdots,A^{(n)})$

令 $\langle A_{(1)},\cdots,A_{(m)}\rangle:=\{\lambda_1A_{(1)}+\cdots+\lambda_mA_{(m)}\mid\forall\lambda_i\in K\}\subset K^n$

和 $\langle A^{(1)},\cdots,A^{(n)}\rangle=\{\lambda_1A^{(1)}+\cdots+\lambda_nA^{(n)}\mid\forall\lambda_i\in K\}\subset K^m$

分别表示由 A 的行向量生成的子空间, 和由 A 的列向量生成的子空间. 则 $r_r(A):=\dim_K\langle A_{(1)},\cdots,A_{(m)}\rangle$ 称为矩阵 A 的行秩, $r_c(A)=\dim_K\langle A^{(1)},\cdots,A^{(n)}\rangle$ 称为矩阵 A 的列秩. □

引理 3.2.2: 任意矩阵 $A=(a_{ij})_{m\times n}$ (系数在域 K 中) 通过对行向量施实若干次初等变换, 可化为如下"阶梯型"

$$\bar{A}=\begin{pmatrix}0\cdots0 & 1 & \bar{a}_{1,i_1+1} & \cdots & \bar{a}_{1i_2} & \cdots & 0 & \bar{a}_{1i_r} & \cdots & \bar{a}_{1n}\\ 0\cdots0 & 0 & 0 & \cdots & 1 & \cdots & 0 & \bar{a}_{2i_r} & \cdots & \bar{a}_{2n}\\ \vdots & \vdots & \vdots & & \vdots & & \vdots & \vdots & & \\ 0\cdots0 & 0 & 0 & \cdots & 0 & \cdots & 0 & 1 & \cdots & \bar{a}_{rn}\\ 0\cdots0 & 0 & 0 & \cdots & 0 & \cdots & 0 & 0 & \cdots & 0\\ \vdots & \vdots & \vdots & & \vdots & & & & & \\ 0\cdots0 & 0 & 0 & \cdots & 0 & \cdots & 0 & 0 & \cdots & 0\end{pmatrix}\qquad(i_1<i_2<\cdots<i_r).$$

其中 r 等于 A 的行秩 (即 r 与初等变换无关).

证明. A 可化为"阶梯型"是高斯消元法的推论. 仅需证明: r 是 A 的行秩 $r_r(A)$. 令 $\bar{A}=[\bar{A}_{(1)},\cdots,\bar{A}_{(m)}]$. 则

$$\langle A_{(1)},A_{(2)},\cdots,A_{(m)}\rangle=\langle\bar{A}_{(1)},\cdots,\bar{A}_{(m)}\rangle.$$

故 $r_r(A)=\dim_K\langle\bar{A}_{(1)},\cdots,\bar{A}_{(m)}\rangle=\dim_K\langle\bar{A}_{(1)},\cdots,\bar{A}_{(r)}\rangle$.

~~只需证明: $\bar{A}_{(1)},\cdots,\bar{A}_{(r)}$ 线性无关~~。

对 $[\bar{A}_{(1)},\cdots,\bar{A}_{(r)}]$ 施实进一步的初等行变换，对 $i_1<i_2<\cdots<i_r$ 可设 $\bar{a}_{j i_k}=0$ $(\forall j<k)$. 此时，行向量 $\lambda_1\bar{A}_{(1)}+\lambda_2\bar{A}_{(2)}+\cdots+\lambda_r\bar{A}_{(r)}$ 在第k行，第 i_k 列位置的元素必为 λ_k. 所以，如果

$$\lambda_1\bar{A}_{(1)}+\lambda_2\bar{A}_{(2)}+\cdots+\lambda_r\bar{A}_{(r)}=0$$

则 $\lambda_1=\lambda_2=\cdots=\lambda_r=0$. $\bar{A}_{(1)},\cdots,\bar{A}_{(r)}$ 线性无关. 故 $r=r_r(A)$. □

推论3.2.2. 设 $W\subset K^n$ 是齐次线性方程组

$$x_1A^{(1)}+x_2A^{(2)}+\cdots+x_nA^{(n)}=0$$

的解空间. 令 $A=(A^{(1)},A^{(2)},\cdots,A^{(n)})$. 则

$$\dim_K(W)=n-r_r(A)$$

证明: 对A的行向量作初等变换可将方程化成阶梯型(为简化符号可设 $i_1=1,i_2=2,\cdots,i_r=r$):

$$\begin{cases} x_1+\bar{a}_{12}x_2+\bar{a}_{13}x_3+\cdots+\bar{a}_{1r+1}x_{r+1}+\cdots+\bar{a}_{1n}x_n=0\\ \qquad x_2+\bar{a}_{23}x_3+\cdots+\bar{a}_{2r+1}x_{r+1}+\cdots+\bar{a}_{2n}x_n=0\\ \qquad\qquad\vdots\\ \qquad\qquad\qquad x_r+\bar{a}_{rr+1}x_{r+1}+\cdots+\bar{a}_{rn}x_n=0\end{cases}$$

解方程组得:

$$(2.1)\quad\begin{cases} x_1=c_{1r+1}x_{r+1}+c_{1r+2}x_{r+2}+\cdots+c_{1n}x_n\\ x_2=c_{2r+1}x_{r+1}+c_{2r+2}x_{r+2}+\cdots+c_{2n}x_n\\ \vdots\\ x_r=c_{rr+1}x_{r+1}+c_{rr+2}x_{r+2}+\cdots+c_{rn}x_n\end{cases}$$

其中 $c_{ij}\in K$ 是(可计算)常数 $x_{r+1},x_{r+2},\cdots,x_n$ 是自由变量. 所以，如果分别取 $[x_{r+1},\cdots,x_n]$ 等于: $[1,0\cdots0]$, $[0,1,0\cdots0]$, $\cdots$, $[0,0,\cdots0,1]$ 可得 $n-r$ 个解

$$\beta_1=\begin{pmatrix}c_{1r+1}\\ \vdots\\ c_{rr+1}\\ 1\\ 0\\ \vdots\\ 0\end{pmatrix},\quad \beta_2=\begin{pmatrix}c_{1r+2}\\ \vdots\\ c_{rr+2}\\ 0\\ 1\\ 0\\ \vdots\\ 0\end{pmatrix},\quad \beta_3=\begin{pmatrix}c_{1r+3}\\ \vdots\\ c_{rr+3}\\ 0\\ 0\\ 1\\ \vdots\\ 0\end{pmatrix},\quad\cdots\quad \beta_{n-r}=\begin{pmatrix}c_{1n}\\ \vdots\\ c_{rn}\\ 0\\ \vdots\\ 0\\ 1\end{pmatrix}.$$

它们是线性无关的. 且 $x_1A^{(1)}+x_2A^{(2)}+\cdots+x_nA^{(n)}=0$ 的任意一组解 $x=[x_1,x_2,\cdots,x_r,x_{r+1},\cdots,x_n]\in W$ 可以写成:

$$x=\left[\sum_{i=1}^{n-r}c_{1r+i}x_{r+i},\ \sum_{i=1}^{n-r}c_{2r+i}x_{r+i},\ \cdots,\ \sum_{i=1}^{n-r}c_{rr+i}x_{r+i},\ x_{r+1},\cdots,x_n\right]$$

$$=x_{r+1}\beta_1+x_{r+2}\beta_2+\cdots+x_n\beta_{n-r}\ .$$

故 $\dim_K(W)=n-r=n-r_r(A)$ □.

~~注~~. 上述~~过程~~证明中构造的 $x_1A^{(1)}+\cdots+x_nA^{(n)}=0$ 的线性无关解 $\beta_1,\beta_2,\cdots,\beta_{n-r}$ 称为基础解系, 方程 $x_1A^{(1)}+\cdots+x_nA^{(n)}=b$ 的一个固定解 $\alpha\in K^n$ 称为特解. 而形如 $\alpha+\lambda_1\beta_1+\cdots+\lambda_{n-r}\beta_{n-r}$ 的解称为

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

的一般解.

我们最后证明一个矩阵的行秩等于列秩. 证明之前先做一点一般的讨论.

引理3.2.3: 设 V 是一个 K-向量空间, $\alpha_1,\alpha_2,\cdots,\alpha_m\in V$ 是一组不全为零的向量. 则存在 $\alpha_{i_1},\cdots,\alpha_{i_r}\in\{\alpha_1,\alpha_2,\cdots,\alpha_m\}$ 满足: (1) $\alpha_{i_1},\alpha_{i_2},\cdots,\alpha_{i_r}$ 线性无关;

(2) $\{\alpha_1,\alpha_2,\cdots,\alpha_m\}$ 中每个向量可由 $\alpha_{i_1},\alpha_{i_2},\cdots,\alpha_{i_r}$ 线性表出.

(这样的一组向量 $\alpha_{i_1},\cdots,\alpha_{i_r}$ 称为 $\alpha_1,\cdots,\alpha_m$ 的一个极大线性无关组).

证明: 令 $S=\{\alpha_1,\alpha_2,\cdots,\alpha_m\}$. 子集合 $T\subset S$ 称为无关子集, 如果 T 中的向量线性无关. 令 $T\subset S$ 是一个向量个数 $|T|$ 最大的无关子集. 因为 $\alpha_1,\cdots,\alpha_m$ 不全为零, 故 T 非空. 令 $T=\{\alpha_{i_1},\cdots,\alpha_{i_r}\}$. 由定义, $\alpha_{i_1},\cdots,\alpha_{i_r}$ 线性无关. 另一方面, $\forall\alpha\in S$, $\{\alpha_{i_1},\cdots,\alpha_{i_r},\alpha\}=T'$ 必是线性相关(否则, $|T'|>|T|$ 与 T 的选取矛盾).

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 $\lambda_1, \cdots, \lambda_r$ 不全为零，且 $\lambda_1 \alpha_{i_1} + \cdots + \lambda_r \alpha_{i_r} = 0$ 与 $\alpha_{i_1}, \cdots, \alpha_{i_r}$ 线性无关矛盾。故 $\lambda_{r+1} \neq 0$。从而 α 可由 $\alpha_{i_1}, \cdots, \alpha_{i_r}$ 线性表出。

□

由引理 3.1.3，向量组 $\{\alpha_1, \cdots, \alpha_m\}$ 的任意两组极大线性无关组的向量个数是一样的。故可将极大线性无关组中向量的个数定义为向量组 $\{\alpha_1, \cdots, \alpha_m\}$ 的秩。

引理 3.2.4：设 $\alpha_1, \cdots, \alpha_m \in V$ 是任意 m 个不全为零的向量。

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

是 $\alpha_1, \cdots, \alpha_m$ 生成的子空间。则 $\alpha_1, \cdots, \alpha_m$ 中任意一组极大线性无关组 $\alpha_{i_1}, \cdots, \alpha_{i_r}$ 都是 $\langle \alpha_1, \cdots, \alpha_m \rangle$ 的一组基。特别，

$$\dim_K \langle \alpha_1, \cdots, \alpha_m \rangle = \text{向量组} \{\alpha_1, \cdots, \alpha_m\} \text{的秩}.$$

证明：设 $\alpha_{i_1}, \alpha_{i_2}, \cdots, \alpha_{i_r}$ 是向量组 $\{\alpha_1, \cdots, \alpha_m\}$ 的一个极大线性无关组。则 $\alpha_{i_1}, \cdots, \alpha_{i_r}$ 线性无关，且每个 α_i 可由它们线性表出。从而 $\langle \alpha_1, \cdots, \alpha_m \rangle$ 中每个向量可由 $\alpha_{i_1}, \cdots, \alpha_{i_r}$ 线性表出。由基的定义，$\alpha_{i_1}, \cdots, \alpha_{i_r}$ 是子空间 $\langle \alpha_1, \cdots, \alpha_m \rangle$ 的一组基。

□

定理 3.2.3：设 $A = (a_{ij})_{m \times n}$ 是系数在 K 中的矩阵。则

$$r_r(A) = r_c(A)$$

证明：为简化符号，可设经过初等行变换后，A 可化为"阶梯型"：

$$\bar{A} = \begin{pmatrix} 1 & \bar{a}_{12} & \cdots & \bar{a}_{1r} & \bar{a}_{1r+1} & \cdots & \bar{a}_{1n} \\ & 1 & & \vdots & \vdots & & \\ & & \ddots & & & & \\ & & & 1 & \bar{a}_{rr+1} & \cdots & \bar{a}_{rn} \\ 0 & 0 & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & & \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & 0 & 0 & \cdots & 0 \end{pmatrix} = [\bar{A}_{(1)}, \cdots, \bar{A}_{(m)}] = (\bar{A}^{(1)}, \cdots, \bar{A}^{(n)}).$$

则由引理3.2.2，$r_r(A)=r_r(\bar{A})=r$. 不妨设证明：$r_c(\bar{A})=r$.

事实上，对$\bar{A}$的列向量做初等变换可将$\bar{A}$化为

$$\bar{A}'=\begin{pmatrix} 1 & 0 & \cdots & 0 & 0 & \cdots & 0 \\ & \ddots & & \vdots & & & \\ * & & 1 & 0 & \cdots & 0 \\ 0 & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & & \vdots & \vdots & & \vdots \\ 0 & \cdots & 0 & 0 & \cdots & 0 \end{pmatrix}$$

从而 $r_c(\bar{A})=r_c(\bar{A}')=r$. 所以定理证明的关键是证明：$r_c(A)$在对A的行向量做初等变换时保持不变. 即：$r_c(A)=r_c(\bar{A})$. 它可由如下事实：$\forall\ x=[x_1,x_2,\cdots,x_n]\in K^n$，有

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

证明。由于$r_c(\bar{A})=r$，故存在$\{\bar{A}^{(1)},\cdots,\bar{A}^{(n)}\}$的极大线性无关组$\bar{A}^{(i_1)},\cdots,\bar{A}^{(i_r)}$. 则我们断言 $A^{(i_1)},\cdots,A^{(i_r)}$也线性无关，否则存在不全为零的 $\lambda_{i_1},\cdots,\lambda_{i_r}\in K$ 使得 $\lambda_{i_1}A^{(i_1)}+\cdots+\lambda_{i_r}A^{(i_r)}=0$，故 $\lambda_{i_1}\bar{A}^{(i_1)}+\cdots+\lambda_{i_r}\bar{A}^{(i_r)}=0$ 与 $\bar{A}^{(i_1)},\cdots,\bar{A}^{(i_r)}$线性无关矛盾。$\forall\ A^{(j)}\in\{A^{(1)},\cdots,A^{(n)}\}$，则 $\bar{A}^{(j)}$可由$\bar{A}^{(i_1)},\cdots,\bar{A}^{(i_r)}$线性表出，即存在 $x_{i_1},\cdots,x_{i_r}\in K$ 使

$$x_{i_1}\bar{A}^{(i_1)}+\cdots+x_{i_r}\bar{A}^{(i_r)}+\bar{A}^{(j)}=0 \qquad (\text{可设 } j\notin\{i_1,\cdots,i_r\})$$

从而 $x_{i_1}A^{(i_1)}+\cdots+x_{i_r}A^{(i_r)}+A^{(j)}=0$. 即 $A^{(j)}$可由$A^{(i_1)},\cdots,A^{(i_r)}$线性表出. 所以 $A^{(i_1)},\cdots,A^{(i_r)}$ 是 $\{A^{(1)},\cdots,A^{(n)}\}$ 的极大线性无关组. 由引理3.2.4，有 $r_c(A)=\dim_K\langle A^{(1)},\cdots,A^{(n)}\rangle=r$. □

定义3.2.4：$r(A):=r_r(A)=r_c(A)$ 称为矩阵A的秩

推论3.2.3：线性方程组 $x_1A^{(1)}+x_2A^{(2)}+\cdots+x_nA^{(n)}=b$ 有解的充要条件是。$r(A)=r(\tilde{A})$. 其中$\tilde{A}=(A,b)$是增广矩阵.

证明：如果有解，则 $r(\tilde{A})=r_c(\tilde{A})=r_c(A)=r(A)$. 反之，若 $r(A)=r(\tilde{A})$，则 $r_c(\tilde{A})=r_c(A)$，故b可由$A^{(1)},\cdots,A^{(n)}$线性表出 □.

习题3.2

1. 求解下列线性方程组

$$\begin{cases} x_1+2x_2+x_3-3x_4+2x_5=0 \\ 2x_1+x_2+x_3+x_4-3x_5=0 \\ x_1+x_2+3x_3+2x_4-2x_5=0 \\ 2x_1+3x_2-5x_3-17x_4+10x_5=0 \end{cases}, \qquad \begin{cases} x_1+2x_2+6x_3-2x_4=-7 \\ -2x_1-5x_2-10x_3+3x_4=10 \\ x_1+2x_2+4x_3=0 \\ x_2+3x_3-3x_4=-10 \end{cases}$$

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

3. 将方程组

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

写成 K^4 中的线性组合形式，并求系数矩阵和增广矩阵的秩.

4. 判定下列向量 β 能否用向量组 $\alpha_1,\alpha_2,\alpha_3$ 线性表出.

(1) $\beta=(5,4,-2,4)$，$\alpha_1=(1,0,-1,3)$，$\alpha_2=(2,1,0,1)$，$\alpha_3=(0,-2,1,1)$

(2) $\beta=(1,1,1,1)$，$\alpha_1=(-1,2,-1,1)$，$\alpha_2=(4,0,1,-1)$，$\alpha_3=(3,2,0,0)$.

5. 设矩阵 $A=(a_{ij})_{m\times n}$ 的秩为 r. 证明：通过初等行变换和初等列变换可将矩阵 A 化为

$$\bar{A}=\begin{pmatrix} 1 & 0 & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & 1 & & \vdots & \vdots & & \vdots \\ \vdots & & \ddots & 1 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & & \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & 0 & 0 & \cdots & 0 \end{pmatrix} \quad (r\geq 1).$$

6. 设 α 是方程组 $x_1A^{(1)}+\cdots+x_nA^{(n)}=b$ 的一个解，$\beta_1,\cdots,\beta_{n-r}$ 是齐次线性方程组 $x_1A^{(1)}+\cdots+x_nA^{(n)}=0$ 解空间的一组基. 证明：

$$\alpha,\ \alpha+\beta_1,\ \alpha+\beta_2,\ \cdots,\ \alpha+\beta_{n-r}$$

线性无关。

§3. 线性映射与矩阵

线性映射的概念可以说是随着K-向量空间V当然而自然出现的。例如，线性函数 $l: V \to K$ 就是一种特殊的线性映射。另一个自然出现的例子是引理3.1.6，一旦选定V的一组基 $e_1, \cdots, e_n \in V$，我们就得到了n个线性函数 $e_1^*, \cdots, e_n^* \in V^*$ 和一个双射

$$V \xrightarrow{\varphi_e} K^n, \quad x \longmapsto [e_1^*(x), \cdots, e_n^*(x)],$$

它满足：(1) $\varphi_e(x+y) = \varphi_e(x) + \varphi_e(y)$，(2) $\varphi_e(\lambda x) = \lambda\varphi_e(x)$。

定义3.3.1. 设V，W是两个K-向量空间。一个映射

$$\varphi: V \longrightarrow W$$

称为K-线性映射（简称线性映射），如果它满足：

$$(1)\ \varphi(x+y) = \varphi(x) + \varphi(y), \qquad (2)\ \varphi(\lambda x) = \lambda\varphi(x).$$

其中 $x, y \in V$ 是任意向量，$\lambda \in K$ 是任意常数。当 $W = V$ 时，我们称线性映射 $\varphi: V \to V$ 是V上的一个线性算子。 □

例3.3.1：设 $l_1, \cdots, l_m \in V^*$ 是V上任意m个线性函数。则映射 $\varphi: V \longrightarrow K^m$，$x \longmapsto [l_1(x), \cdots, l_m(x)]$，是一个线性映射。 □

例3.3.2：设 $M_{m\times n}(K)$ 表示所有系数在K中的 $m\times n$ 矩阵的集合。$\forall\ A = (A^{(1)}, \cdots, A^{(n)}) \in M_{m\times n}(K)$，映射

$$\varphi_A: K^n \longrightarrow K^m, \quad x = [x_1, \cdots, x_n] \longmapsto \varphi_A(x),$$

$$\varphi_A(x) := x_1 A^{(1)} + \cdots + x_n A^{(n)}$$

是一个K-线性映射。 □

例3.3.3：$\frac{d}{dx}$：$\mathbb{R}[x] \longrightarrow \mathbb{R}[x]$ 是一个$\mathbb{R}$-线性映射。另外，如果 $C(\mathbb{R}) = \{\mathbb{R} \xrightarrow{f} \mathbb{R}\}$ 表示$\mathbb{R}$上所有连续函数组成的$\mathbb{R}$-向量空间，则映射 $\varphi : C(\mathbb{R}) \longrightarrow C(\mathbb{R})$，$\forall f(x) \in C(\mathbb{R})$，

$$f(x) \longmapsto \int_0^t f(x)dx.$$

也是一个$\mathbb{R}$-线性映射.

□

例3.3.4：如果 $f : V \to W$，$g : V \to W$ 是两个K-线性映射，则映射 $f+g : V \to W$，$x \mapsto f(x)+g(x)$，和映射 $\lambda f : V \to W$，$x \mapsto \lambda f(x)$，都是K-线性映射（其中$\lambda \in K$是任意固定的常数）.

□

例3.3.5：令 $\mathcal{L}(V,W) = \{V \xrightarrow{f} W\}$ 表示所有V到W的K-线性映射的集合. 则 $\mathcal{L}(V,W)$ 关于例3.3.4中定义的"加法"和"数乘法"是一个K-向量空间.

例3.3.6：设 $e_1, \cdots, e_n \in V$，$w_1, \cdots, w_m \in W$ 分别是V和W的一组基，$A = [A^{(1)}, \cdots, A^{(n)}] \in M_{m\times n}(K)$，则映射

$$\varphi_A : V \longrightarrow W, \quad \varphi_A(x) = (w_1, w_2, \cdots, w_m)\sum_{i=1}^n e_i^*(x) A^{(i)},$$

是一个K-线性映射，使得 $\varphi_A(e_i) = (w_1, \cdots, w_m) A^{(i)}$.

证明：映射 $\varphi_A : V \to W$ 的定义分两步：首先指定φ_A在 $e_1, \cdots, e_n$ 上的像，$\varphi_A(e_i) = (w_1, \cdots, w_m) A^{(i)}$，然后对任意 $x = e_1^*(x) e_1 + \cdots + e_n^*(x) e_n$，指定.

$$\varphi_A(x) := e_1^*(x)\varphi_A(e_1) + \cdots + e_n^*(x)\varphi_A(e_n).$$

因此只须验证 $\varphi_A : V \to W$ 是一个K-线性映射：$\forall x, y \in V$，

$$\varphi_A(x+y) = e_1^*(x+y)\varphi_A(e_1) + \cdots + e_n^*(x+y)\varphi_A(e_n) = \varphi_A(x) + \varphi_A(y).$$

$$\varphi_A(\lambda x) = e_1^*(\lambda x)\varphi_A(e_1)+\cdots+e_n^*(\lambda x)\varphi_A(e_n) = \lambda\varphi_A(x).$$

通过计算，φ_A 的定义可以用如下公式表述：

$$(3.1)\qquad \varphi_A(x) = (w_1,\cdots,w_m)\sum_{i=1}^n e_i^*(x)A^{(i)}$$

□

引理3.3.1：分别固定 V 和 W 的一组基 $(v_1,\cdots,v_n)$ 和 $(w_1,\cdots,w_m)$. 则对任意 K-线性映射 $\varphi: V\to W$，存在唯一的矩阵 $A_\varphi = (A_\varphi^{(1)},\cdots,A_\varphi^{(n)})\in M_{m\times n}(K)$ 使得，$\forall x\in V$,

$$(3.2).\qquad \varphi(x) = (w_1,\cdots,w_m)\sum_{i=1}^n v_i^*(x)A_\varphi^{(i)}$$

证明：因为 $(w_1,\cdots,w_m)$ 是 W 的一组基，所以

$$\varphi(v_i) = a_{1i}w_1+\cdots+a_{mi}w_m = (w_1,\cdots,w_m)[a_{1i},\cdots,a_{mi}]$$

$(1\le i\le n)$.（事实上，$a_{ji} = w_j^*(\varphi(v_i))$）。令

$$A_\varphi := [A_\varphi^{(1)},\cdots,A_\varphi^{(n)}],\quad A_\varphi^{(i)} = [a_{1i},\cdots,a_{mi}].$$

则 A_φ 是由 $(\varphi(v_1),\cdots,\varphi(v_n))$ 和 $(w_1,\cdots,w_m)$ 唯一确定. 且

$$\varphi(x) = v_1^*(x)\varphi(v_1)+\cdots+v_n^*(x)\varphi(v_n) =: (\varphi(v_1),\cdots,\varphi(v_n))\begin{pmatrix}v_1^*(x)\\ \vdots\\ v_n^*(x)\end{pmatrix}$$

由 $\varphi(v_i) = (w_1,\cdots,w_m)A_\varphi^{(i)}$ 代入上式即得

$$\varphi(x) = (w_1,\cdots,w_m)\sum_{i=1}^n v_i^*(x)A_\varphi^{(i)}.$$

□.

符号约定：$(\varphi(v_1),\cdots,\varphi(v_n)) = (w_1,\cdots,w_m)A_\varphi$ 意指

$\varphi(v_1) = (w_1,\cdots,w_m)A_\varphi^{(1)},\cdots,\varphi(v_i) = (w_1,\cdots,w_m)A_\varphi^{(i)},\cdots,$ $\varphi(v_n) = (w_1,\cdots,w_m)A_\varphi^{(n)}$。因为 φ 由 $(\varphi(v_1),\cdots,\varphi(v_n))$ 唯一确定，所以我们通常用 $(\varphi(v_1),\cdots,\varphi(v_n)) = (w_1,\cdots,w_m)A_\varphi$ 来说明 φ 是如何被定义的。

定义 3.3.2：设 $(v_1,\cdots,v_n)$，$(w_1,\cdots,w_m)$ 分别是K-向量空间 V 和 W 的一组基。$\varphi: V\to W$ 是一个K-线性映射。令 $A_\varphi\in M_{m\times n}(K)$ 是~~使得~~满足等式

$$(\varphi(v_1),\cdots,\varphi(v_n))=(w_1,\cdots,w_m)A_\varphi$$

的矩阵。则称 A_φ 是 φ 在基 $(v_1,\cdots,v_n)$ 和 $(w_1,\cdots,w_m)$ 下的坐标矩阵。当 $V=W$ 时，通常取 $(w_1,\cdots,w_n)=(v_1,\cdots,v_n)$。 □

推论 3.3.1：设 $V\xrightarrow{\varphi}W$ 和 $V\xrightarrow{\psi}W$ 在基 $(v_1,\cdots,v_n)$ 和 $(w_1,\cdots,w_m)$ 下的坐标矩阵分别是 $A_\varphi=(a_{ij})_{m\times n}$，$A_\psi=(b_{ij})_{m\times n}$，则 $\varphi+\psi$ 和 $\lambda\varphi$ 在基 $(v_1,\cdots,v_n)$ 和 $(w_1,\cdots,w_m)$ 下的坐标矩阵分别是 $A_{\varphi+\psi}=(a_{ij}+b_{ij})_{m\times n}$ 和 $A_{\lambda\varphi}=(\lambda a_{ij})$。

证明：由定义，~~$(\varphi(v_1),\cdots,\varphi(v_n))=(w_1,\cdots,w_m)A_\varphi$~~，

$$\varphi(v_j)=(w_1,\cdots,w_m)A_\varphi^{(j)},\quad \psi(v_j)=(w_1,\cdots,w_m)A_\psi^{(j)}$$

故 $(\varphi+\psi)(v_j)=(w_1,\cdots,w_m)(A_\varphi^{(j)}+A_\psi^{(j)})$。所以 $A_{\varphi+\psi}^{(j)}=A_\varphi^{(j)}+A_\psi^{(j)}$。

即 $A_{\varphi+\psi}=(a_{ij}+b_{ij})_{m\times n}$，同理 $A_{\lambda\varphi}=(\lambda a_{ij})$。 □

定义 3.3.3：$\forall\ A=(a_{ij})_{m\times n}$，$B=(b_{ij})_{m\times n}\in M_{m\times n}(K)$，$\forall\lambda\in K$。

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

称为矩阵的加法和数乘法。 □

定义 3.3.4：一个K-线性映射 $V\xrightarrow{\varphi}W$ 称为同构映射，如果 φ 是一个双射。 □

引理 3.3.2：如果 $\varphi: V\to W$ 是一个同构映射，则 $\dim_K(V)=\dim_K(W)$ 且它的逆映射 $\varphi^{-1}: W\to V$ 也是同构映射。

证明：设 $(v_1,\cdots,v_n)$ 是 V 的一组基，只需证明
$(\varphi(v_1),\cdots,\varphi(v_n))$ 是 W 的一组基。首先，如果

$$\lambda_1\varphi(v_1)+\cdots+\lambda_n\varphi(v_n)=0.$$

则 $\varphi(\lambda_1v_1+\cdots+\lambda_nv_n)=\lambda\varphi(v_1)+\cdots+\lambda_n\varphi(v_n)=0$. 另外，
$\varphi(0)=0$，而 φ 是单射，故 $\lambda_1v_1+\cdots+\lambda_nv_n=0$。从而
$\lambda_1=\lambda_2=\cdots=\lambda_n=0$。即 $\varphi(v_1),\cdots,\varphi(v_n)$ 线性无关。其次，
$\forall\ \beta\in W$，存在 $\alpha\in V$ 使 $\varphi(\alpha)=\beta$.（φ 是满射），由于
$(v_1,\cdots,v_n)$ 是 V 的基，故存在 $\lambda_1,\cdots,\lambda_n\in K$ 使

$$\alpha=\lambda_1v_1+\cdots+\lambda_nv_n$$

从而 $\beta=\varphi(\alpha)=\lambda_1\varphi(v_1)+\cdots+\lambda_n\varphi(v_n)$. 所以 $(\varphi(v_1),\cdots,\varphi(v_n))$ 是
W 的一组基。而 φ 的逆映射 $\varphi^{-1}:W\longrightarrow V$ 是线性映射
可由定义直接验证：$\forall\ \alpha,\beta\in W$，有 $\varphi(\varphi^{-1}(\alpha+\beta))=\alpha+\beta=\varphi(\varphi^{-1}(\alpha)+\varphi^{-1}(\beta))$
故 $\varphi^{-1}(\alpha+\beta)=\varphi^{-1}(\alpha)+\varphi^{-1}(\beta)$. 同理可证 $\varphi^{-1}(\lambda\alpha)=\lambda\varphi^{-1}(\alpha)$.

□

例3.3.7：由例3.3.5知，所有线性映射 $V\xrightarrow{\varphi}W$ 的集合 $\mathcal{L}(V,W)$
关于线性映射的加法和数乘是一个 K-向量空间。容易
验证 $M_{m\times n}(K)$ 关于矩阵的加法和数乘成为一个 mn
维 K-向量空间。映射

$$\mathcal{L}(V,W)\xrightarrow{A}M_{m\times n}(K),\ \varphi\longmapsto A_\varphi$$

是两个 K-向量空间之间的同构映射。

□

引理3.3.3：设 $U\xrightarrow{\psi}V$，$V\xrightarrow{\varphi}W$ 是两个线性映射
则 (1) $\varphi\circ\psi:U\longrightarrow W$ 也是线性映射.

(2) 设 $(u_1,\cdots,u_s)$ 是 U 的一组基，$(v_1,\cdots,v_n)$ 是 V 的一组基，$(w_1,\cdots,w_m)$
是 W 的一组基。如果 ψ，φ 在上述基下的坐标矩阵分别
是 $A_\psi=(b_{ij})_{n\times s}$，$A_\varphi=(a_{ij})_{m\times n}$. 则 $\varphi\circ\psi$ 在上述基下的

坐标矩阵 $A_{\varphi\cdot\psi}=(c_{ij})_{m\times s}$，其中

$$c_{ij}=a_{i1}b_{1j}+a_{i2}b_{2j}+\cdots+a_{in}b_{nj}$$
$$:=(a_{i1},a_{i2},\cdots,a_{in})\begin{pmatrix}b_{1j}\\b_{2j}\\\vdots\\b_{nj}\end{pmatrix}:=(A_\varphi)_{(i)}\cdot A_\psi^{(j)}.$$

证明：(1) $\forall\ x,y\in U$，$\varphi\cdot\psi(x+y)=\varphi(\psi(x)+\psi(y))=\varphi\cdot\psi(x)+\varphi\psi(y)$

同理 $\varphi\cdot\psi(\lambda x)=\varphi(\lambda\psi(x))=\lambda\varphi(\psi(x))=\lambda\varphi\cdot\psi(x)$。故 $\varphi\cdot\psi:U\longrightarrow W$ 是线性映射。

(2) 由定义，$(\psi(u_1),\cdots,\psi(u_s))=(v_1,\cdots,v_n)A_\psi$，$(\varphi(v_1),\cdots,\varphi(v_n))=(w_1,\cdots,w_m)A_\varphi$

$(\varphi\cdot\psi(u_1),\cdots,\varphi\cdot\psi(u_s))=(w_1,\cdots,w_m)A_{\varphi\cdot\psi}$。令 $A_{\varphi\cdot\psi}=(c_{ij})_{m\times s}$，则

$\varphi\cdot\psi(u_j)=(w_1,\cdots,w_m)A_{\varphi\cdot\psi}^{(j)}$，$\psi(u_j)=(v_1,\cdots,v_n)A_\psi^{(j)}$，故

$$\varphi\cdot\psi(u_j)=\varphi(\psi(u_j))=(\varphi(v_1),\cdots,\varphi(v_n))A_\psi^{(j)}.$$

由 $\varphi(v_1)=(w_1,\cdots,w_m)A_\varphi^{(1)}$，$\cdots$，$\varphi(v_n)=(w_1,\cdots,w_m)A_\varphi^{(n)}$，计算得

$$(\varphi(v_1),\cdots,\varphi(v_n))A_\psi^{(j)}\underset{=(w_1,\cdots,w_m)(b_{1j}A_\varphi^{(1)}+\cdots+b_{nj}A_\varphi^{(n)})}{=}(w_1,\cdots,w_m)\left[(A_\varphi)_{(1)}A_\psi^{(j)},\cdots,(A_\varphi)_{(m)}A_\psi^{(j)}\right]$$

所以，$(w_1,\cdots,w_m)A_{\varphi\cdot\psi}^{(j)}=(w_1,\cdots,w_m)\left[(A_\varphi)_{(1)}A_\psi^{(j)},\cdots,(A_\varphi)_{(m)}A_\psi^{(j)}\right]$.

比较每个 w_i 的系数得：

$$c_{ij}=(A_\varphi)_{(i)}\cdot A_\psi^{(j)}=a_{i1}b_{1j}+a_{i2}b_{2j}+\cdots+a_{in}b_{nj}.$$

□

定义3.3.5 [矩阵乘法的定义]. $\forall\ A=(a_{ij})_{m\times n}$，$B=(b_{ij})_{n\times s}$，则矩阵

$$A\cdot B:=(c_{ij})_{m\times s}，\text{其中 } c_{ij}=A_{(i)}\cdot B^{(j)},$$

称为矩阵 A 与 B 的乘积。

□

推论3.3.2：若 $A\in M_{m\times n}(K)$，$B\in M_{n\times s}(K)$，$C\in M_{s\times t}(K)$，则

(1) $(A\cdot B)\cdot C=A\cdot(B\cdot C)$.　（结合律）.

(2) $(A_1+A_2)\cdot B=A_1\cdot B+A_2\cdot B$，$(A_1,A_2\in M_{m\times n}(K))$.（分配律）.

(3) $A\cdot(B_1+B_2)=A\cdot B_1+A\cdot B_2$，$(B_1,B_2\in M_{n\times s}(K))$.（分配律）.

证明：(1)：因为映射合成满足结合律. (2) 和 (3) 等价于验证线性映射的合成满足分配律（当然，也可由矩阵运算直接验证）.

$\forall$ ~~映射~~ 线性映射：$U\xrightarrow{\psi}V$，$V\xrightarrow{\varphi_1}W$，$V\xrightarrow{\varphi_2}W$，~~事~~ 证明 $(\varphi_1+\varphi_2)\cdot\psi=\varphi_1\cdot\psi+\varphi_2\cdot\psi$，事实上，$\forall x\in U$，有

$$(\varphi_1+\varphi_2)\cdot\psi(x)=(\varphi_1+\varphi_2)(\psi(x))=\varphi_1(\psi(x))+\varphi_2(\psi(x))$$

~~故~~ 同理，$\forall$ $U\xrightarrow{\psi_1}V$，$U\xrightarrow{\psi_2}V$，$V\xrightarrow{\varphi}W$，$\forall x\in U$，有

$\varphi\cdot(\psi_1+\psi_2)(x)=\varphi(\psi_1(x)+\psi_2(x))=\varphi\cdot\psi_1(x)+\varphi\cdot\psi_2(x)=(\varphi\cdot\psi_1+\varphi\cdot\psi_2)(x)$.

故 $\varphi\cdot(\psi_1+\psi_2)=\varphi\cdot\psi_1+\varphi\cdot\psi_2$.　　□.

推论3.3.3：设 $A_\varphi\in M_{m\times n}(K)$ 是线性映射 $V\xrightarrow{\varphi}W$ 在基 $(v_1,\cdots,v_n)$ 和 $(w_1,\cdots,w_m)$ 下的坐标矩阵. 则下列条件等价：

(1) φ 是一个同构映射；

(2) ~~[struck]~~ A_φ 是方阵（即 $n=m$）且存在方阵 $B\in M_n(K)$ 使得 $A_\varphi\cdot B=B\cdot A_\varphi=I_n$（$I_n$ 是 n 阶单位阵）；

(3) $\dim_K(V)=\dim_K(W)$，且 φ 是单射；

(4) $\dim_K(V)=\dim_K(W)$ 且 φ 是满射.

证明：我们将证明 (1) ⇒ (2) ⇒ (3) ⇒ (4) ⇒ (1).

(1) ⇒ (2)：由引理3.3.2 知 $\dim_K(V)=\dim_K(W)$. 即 $n=m$. 且逆映射 $\varphi^{-1}:W\longrightarrow V$ 也是线性映射. 令 $A_{\varphi^{-1}}$ 是 φ^{-1} 在基 $(w_1,\cdots,w_m)$ 和 $(v_1,\cdots,v_n)$ 下的坐标矩阵. 即

$$(\varphi^{-1}(w_1),\cdots,\varphi^{-1}(w_m))=(v_1,\cdots,v_n)A_{\varphi^{-1}}.$$

$(w_1,\cdots,w_m)=(\varphi(\varphi^{-1}(w_1)),\cdots,\varphi(\varphi^{-1}(w_m)))=(\varphi(v_1),\cdots,\varphi(v_n))A_{\varphi^{-1}}$

将 $(\varphi(v_1),\cdots,\varphi(v_n))=(w_1,\cdots,w_m)A_\varphi$ 代入上式可得

$$(w_1,\cdots,w_m)\cdot I_m=(w_1,\cdots,w_m)(A_\varphi\cdot A_{\varphi^{-1}}),$$

所以 $I_m=A_\varphi\cdot A_{\varphi^{-1}}$，同理 $A_{\varphi^{-1}}\cdot A_\varphi=I_n$。令 $B=A_{\varphi^{-1}}$，则 $A_\varphi\cdot B=BA_\varphi=I_n$。

(2) ⇒ (3)：A_φ 是方阵，故 $\dim_K(V)=\dim_K(W)$。另一方面，利用矩阵 B 定义线性映射 $\psi: W\to V$，使

$$(\psi(w_1),\cdots,\psi(w_m))=(v_1,\cdots,v_n)B.$$

则 $(\psi\cdot\varphi(v_1),\cdots,\psi\varphi(v_n))=(\psi(w_1),\cdots,\psi(w_m))A_\varphi$
$=(v_1,\cdots,v_n)B\cdot A_\varphi=(v_1,\cdots,v_n)\cdot I_n$。故 $\psi\cdot\varphi: V\to V$ 是恒等映射，从而 φ 是单射。

(3) ⇒ (4)．当 $\dim_K(V)=\dim_K(W)$，φ 是单射时，则 φ 必为满射。事实上，$\varphi(v_1),\cdots,\varphi(v_n)$ 在 W 中线性无关：若 $\lambda_1\varphi(v_1)+\cdots+\lambda_n\varphi(v_n)=0$，则

$$\varphi(\lambda_1 v_1+\cdots+\lambda_n v_n)=\lambda_1\varphi(v_1)+\cdots+\lambda_n\varphi(v_n)=0.$$

φ 是单射，故 $\lambda_1v_1+\cdots+\lambda_nv_n=0$，从而 $\lambda_1=\cdots=\lambda_n=0$。但 $\dim_K(W)=n$，故 $(\varphi(v_1),\cdots,\varphi(v_n))$ 是 W 的一组基，所以 $\forall\beta\in W$，$\beta=\mu_1\varphi(v_1)+\cdots+\mu_n\varphi(v_n)=\varphi(\mu_1v_1+\cdots+\mu_nv_n)$，即 φ 是满射。

(4) ⇒ (1)：只需证明 φ 必为单射。令 $\varphi(\alpha_i)=w_i$（因为 φ 是满射），则 $\alpha_1,\cdots,\alpha_m$ 在 V 中线性无关，但 $m=\dim_K(V)$，故 $(\alpha_1,\cdots,\alpha_m)$ 是 V 的一组基。如果 φ 不是单射，则存在 $\beta\in V$，$\beta\neq0$ 使得 $\varphi(\beta)=0$。令 $\beta=a_1\alpha_1+\cdots+a_m\alpha_m$，则

$$0=\varphi(\beta)=a_1\varphi(\alpha_1)+\cdots+a_m\varphi(\alpha_m)=a_1w_1+\cdots+a_mw_m$$

可得 $a_1=\cdots=a_m=0$，与 $\beta\neq0$ 矛盾。 □

<u>定义3.3.6</u> [可逆矩阵]. $A\in M_n(K)$ 称为一个<u>可逆矩阵</u>, 如果存在 $B\in M_n(K)$ 使 $AB=BA=I_n$. 可证明这样的 B 由 A 唯一确定, 故称 B 是 A 的<u>逆矩阵</u>, 记为 $B=A^{-1}$ □

<u>定义3.3.7</u>. 设 $\varphi: V\longrightarrow W$ 是一个线性映射. 则

$$\ker(\varphi)=\{v\in V \mid \varphi(v)=0\}$$

是 V 的一个子空间, 称为 φ 的核 (kernel), 而

$$\mathrm{Im}(\varphi)=\{\varphi(v) \mid \forall\, v\in V\}$$

是 W 的一个子空间, 称为 φ 的像 (Image). □

<u>定理3.3.1</u>. 设线性映射 $\varphi: V\longrightarrow W$ 定义如下:

$$(\varphi(v_1),\cdots,\varphi(v_n))=(w_1,\cdots,w_m)A_\varphi.$$

则 (1) $\dim_K(V)=\dim_K(\ker(\varphi))+\dim_K(\mathrm{Im}(\varphi))$.

(2) $\dim_K(\mathrm{Im}(\varphi))=r_c(A_\varphi)$.

(3) $\dim_K(\ker(\varphi))=\dim_K(V)-r_r(A_\varphi)$.

证明: (1). 设 $\varphi(\beta_1),\cdots,\varphi(\beta_s)\in \mathrm{Im}(\varphi)$ 是 $\mathrm{Im}(\varphi)$ 的一<u>组基</u>. $\alpha_1,\cdots,\alpha_t\in\ker(\varphi)$ 是 $\ker(\varphi)$ 的一组基. 则我们断言:

$$\alpha_1,\cdots,\alpha_t,\beta_1,\cdots,\beta_s$$

是 V 的<u>一组基</u>. 首先证明它们线性无关. 若

$$\lambda_1\alpha_1+\cdots+\lambda_t\alpha_t+\mu_1\beta_1+\cdots+\mu_s\beta_s=0.$$

则 $0=\varphi(\lambda_1\alpha_1+\cdots+\lambda_t\alpha_t+\mu_1\beta_1+\cdots+\mu_s\beta_s)=\mu_1\varphi(\beta_1)+\cdots+\mu_s\varphi(\beta_s)$

从而 $\mu_1=\cdots=\mu_s=0$, 故 $\lambda_1\alpha_1+\cdots+\lambda_t\alpha_t=0$, 从而 $\lambda_1=\cdots=\lambda_t=0$.

其次证明 V 中任意向量 ν 可由它们线性表出: $\varphi(\nu)$ 可由 $\varphi(\beta_1),\cdots,\varphi(\beta_s)$ 线性表出, 即存在 $b_1,\cdots,b_s\in K$ 使

$$\varphi(\nu)=b_1\varphi(\beta_1)+\cdots+b_s\varphi(\beta_s)$$

即 $\varphi(\nu-b_1\beta_1-\cdots-b_s\beta_s)=0$. 故 $\nu-b_1\beta_1-\cdots-b_s\beta_s\in\ker(\varphi)$

从而存在 $a_1,\cdots,a_t\in K$ 使 $\nu=a_1\alpha_1+\cdots+a_t\alpha_t+b_1\beta_1+\cdots+b_s\beta_s$.

(2) 回忆由基 $w=(w_1,\cdots,w_m)$ 确定的坐标映射

$$\Phi_w: W \longrightarrow K^m, \quad \alpha \longmapsto [w_1^*(\alpha),\cdots,w_m^*(\alpha)]$$

是一个同构映射。而 $\varphi(v_i)$ 在 $w_1,\cdots,w_m$ 下的坐标恰为 A_φ 的第 i 行列向量 $A_\varphi^{(i)}$。故 $\mathrm{Im}(\varphi)$ 在 Φ_w 下的像恰为

$$\langle A_\varphi^{(1)},\cdots,A_\varphi^{(n)}\rangle=\{x_1A_\varphi^{(1)}+\cdots+x_nA_\varphi^{(n)} \mid \forall x_i\in K\}.$$

所以 $\dim_K(\mathrm{Im}(\varphi))=\dim_K\langle A_\varphi^{(1)},\cdots,A_\varphi^{(n)}\rangle=r_c(A)$.

(3) 由基 $v=(v_1,\cdots,v_n)$ 确定的坐标映射

$$\Phi_v: V\longrightarrow K^n, \quad x=x_1v_1+\cdots+x_nv_n \longmapsto [x_1,\cdots,x_n],$$

是一个同构映射。而 $x\in\ker(\varphi) \Longleftrightarrow x_1\varphi(v_1)+\cdots+x_n\varphi(v_n)=0$

但 $x_1\varphi(v_1)+\cdots+x_n\varphi(v_n)=(w_1,\cdots,w_m)(x_1A_\varphi^{(1)}+\cdots+x_nA_\varphi^{(n)})$.

故 $x=x_1v_1+\cdots+x_nv_n\in\ker(\varphi) \Longleftrightarrow x_1A_\varphi^{(1)}+\cdots+x_nA_\varphi^{(n)}=0$

这说明子空间 $\ker(\varphi)\subset V$ 在 Φ_v 下的像恰好等于齐次线性方程组 $x_1A_\varphi^{(1)}+\cdots+x_nA_\varphi^{(n)}=0$ 的解空间：

$$\{[x_1,\cdots,x_n]\in K^n \mid x_1A_\varphi^{(1)}+\cdots+x_nA_\varphi^{(n)}=0\}.$$

故 $\dim_K\ker(\varphi)=n-r_r(A_\varphi)$. □

定理 3.3.2 [线性映射的坐标形式].

设 $A_\varphi\in M_{m\times n}(K)$ 是线性映射 $V\xrightarrow{\varphi}W$ 在基 $v=(v_1,\cdots,v_n)$ 和 $w=(w_1,\cdots,w_m)$ 下的坐标矩阵。即：

$$(\varphi(v_1),\cdots,\varphi(v_n))=(w_1,\cdots,w_m)A_\varphi$$

则对任意 $x\in V$ 有

$$[w_1^*(\varphi(x)),\cdots,w_m^*(\varphi(x))]=A_\varphi\cdot[v_1^*(x),\cdots,v_n^*(x)].$$

证明：由 $x=(v_1,\cdots,v_n)[v_1^*(x),\cdots,v_n^*(x)]$ 直接计算可得

$$\varphi(x)=(\varphi(v_1),\cdots,\varphi(v_n))\cdot[v_1^*(x),\cdots,v_n^*(x)]$$
$$=(w_1,\cdots,w_m)\underline{A_\varphi[v_1^*(x),\cdots,v_n^*(x)]}$$

故 $(w_1,\cdots,w_m)A_\varphi[v_1^*(x),\cdots,v_n^*(x)]=(w_1,\cdots,w_m)[w_1^*(\varphi(x)),\cdots,w_m^*(\varphi(x))]$.

比较系数可得：$[w_1^*(\varphi(x)),\cdots,w_m^*(\varphi(x))]=A_\varphi\cdot[v_1^*(x),\cdots,v_n^*(x)]$

□

设 $V\xrightarrow{\varphi}W$ 是任意映射，$l:W\to K$ 是任意函数，则函数 $\varphi^*(l):=l\cdot\varphi:V\xrightarrow{\varphi}W\xrightarrow{l}K$ 称为函数 l 的拉回。

$$\begin{tikzcd} V \arrow[r, "\varphi"] \arrow[rr, bend right, "\varphi^*(l)"'] & W \arrow[r, "l"] & K \end{tikzcd}$$

引理3.3.4.［线性映射~~的~~的函数刻划］．设 $V\xrightarrow{\varphi}W$ 是 K-向量空间 V 和 W 之间的任意映射，则

φ 是线性映射 $\Longleftrightarrow$ $\forall\, l\in W^*$，必有 $\varphi^*(l)\in V^*$.

证明："$\Rightarrow$"．若 φ 是线性映射，而 $l:W\to K$ 是线性函数，则不难验证：$\varphi^*(l):V\to K$ 必为线性函数．

"$\Leftarrow$" 如果 φ 将 W 上的所有线性函数 $W\xrightarrow{l}K$ 拉回成 V 上的线性函数，则 φ 必为线性映射。只需证明：

$$\alpha:=\varphi(\lambda x+\mu y)-\lambda\varphi(x)-\mu\varphi(y)\quad,\ (\forall\,\lambda,\mu\in K,\ x,y\in V)$$

是 W 中的零向量．另一方面，对任意 $\alpha\in W$，有如下事实．

(3.3)：$\qquad \alpha=0 \Longleftrightarrow \forall\, l\in W^*,\ l(\alpha)=0$．

故只需证明：$\forall\, l\in W^*$，$l(\alpha)=l\big(\varphi(\lambda x+\mu y)-\lambda\varphi(x)-\mu\varphi(y)\big)=0$.

计算：$l(\alpha)=l(\varphi(\lambda x+\mu y))-\lambda\, l(\varphi(x))-\mu\, l(\varphi(y))$ （因为 l 是线性函数）

$=\varphi^*(l)(\lambda x+\mu y)-\lambda\,\varphi^*(l)(x)-\mu\,\varphi^*(l)(y)=0$ （因为 $\varphi^*(l)$ 是线性函数）

□．

定义3.3.8［对偶映射］．设 $\varphi:V\to W$ 是一个线性映射，则

$$\varphi^*:W^*\longrightarrow V^*,\quad \varphi^*(l):=l\cdot\varphi$$

称为 φ 的对偶线性映射．

□

~~引理3.3.5．设 $V\xrightarrow{\varphi}W$ 在基 $(v_1,\cdots,v_n)$ 和基 $(w_1,\cdots,w_m)$ 下的坐标矩阵为 $A_\varphi\in M_{m\times n}(K)$，设 $A_{\varphi^*}\in M_{n\times m}(K)$ 是对偶映射 $W^*\xrightarrow{\varphi^*}V^*$ 在对偶基 $(w_1^*,\cdots,w_m^*)$ 和 $(v_1^*,\cdots,v_n^*)$ 下的坐标矩阵．~~

引理 3.3.5. 线性映射 $V \xrightarrow{\varphi} W$ 的对偶映射 $W^* \xrightarrow{\varphi^*} V^*$ 也为线性映射。如果 $A_\varphi=(a_{ij})_{m\times n}\in M_{m\times n}(K)$ 是 φ 在 $(v_1,\cdots,v_n)$ 和 $(w_1,\cdots,w_m)$ 下的坐标矩阵，$A_{\varphi^*}=(a^*_{ij})_{n\times m}\in M_{n\times m}(K)$ 是 φ^* 在对偶基 $(w_1^*,\cdots,w_m^*)$ 和 $(v_1^*,\cdots,v_n^*)$ 下的坐标矩阵。则

$$a^*_{ij}=a_{ji}\ ,\quad 1\le i\le n,\ 1\le j\le m.$$

证明：φ^* 是线性映射：$\forall\ \lambda_1,\lambda_2\in K$, $l_1,l_2\in W^*$, 则

$$\varphi^*(\lambda_1 l_1+\lambda_2 l_2)=(\lambda_1 l_1+\lambda_2 l_2)\cdot\varphi=\lambda_1\cdot l_1\cdot\varphi+\lambda_2 l_2\cdot\varphi=\lambda_1\varphi^*(l_1)+\lambda_2\varphi^*(l_2).$$

计算 A_{φ^*}：回忆 $(\varphi^*(w_1^*),\cdots,\varphi^*(w_m^*))=(v_1^*,\cdots,v_n^*)A_{\varphi^*}$. 由

$\varphi^*(w_j^*)=(v_1^*,\cdots,v_n^*)A^{(j)}_{\varphi^*}$，可知 a^*_{ij} 是 v_i^* 在前述线性组合中的系数。即 $a^*_{ij}=\varphi^*(w_j^*)(v_i)$. 由定义得：

$$\varphi^*(w_j^*)(v_i)=w_j^*(\varphi(v_i)).$$

将 $\varphi(v_i)=(w_1,\cdots,w_m)A^{(i)}_\varphi$ 代入上式得 $w_j^*(\varphi(v_i))=a_{ji}$

□.

定义 3.3.9【转置矩阵】. 设 $A=(a_{ij})_{m\times n}\in M_{m\times n}(K)$. 矩阵 ${}^tA=({}^ta_{ij})_{n\times m}\in M_{n\times m}(K)$ 称为 A 的转置矩阵，如果

$${}^ta_{ij}=a_{ji}\ ,\quad 1\le i\le n,\ 1\le j\le m.$$

□.

有了矩阵和线性映射的概念，我们可以从不同的视角看线性方程组理论。例如，设 $A\in M_{m\times n}(K)$ 是方程组 (*) 的系数矩阵，$b=[b_1,\cdots,b_m]\in M_{m\times 1}(K)$, $x=[x_1,\cdots,x_n]\in M_{n\times 1}(K)$, 则线性方程组 (*) 就是矩阵方程：$A\cdot x=b$.

从线性映射的视角看，由 A 可以定义线性映射

$$\varphi_A:\ K^n\longrightarrow K^m,\quad x=[x_1,\cdots,x_n]\longmapsto x_1A^{(1)}+\cdots+x_nA^{(n)}$$

则方程组 (*) 的解集合就是 φ_A 在 $b\in K^m$ 的纤维 $\varphi_A^{-1}(b)$. 而 $\varphi_A^{-1}(0)=\mathrm{Ker}(\varphi_A)$ 就是 (*) 的解空间. 如果 $\varphi_A^{-1}(b)$ 非空（即 (*) 有解），

$\forall\, y_0 \in \varphi_A^{-1}(b)$，(即 y_0 是 (*) 的一个特解)，则 $\varphi_A^{-1}(b) = y_0 + \varphi_A^{-1}(0)$.

习题 3.3

1. 设 $(\alpha_1, \cdots, \alpha_n)$ 是 V 的一组基，$\varphi: V \to V$ 是一个线性映射，

$$(\varphi(\alpha_1), \cdots, \varphi(\alpha_n)) = (\alpha_1, \cdots, \alpha_n) \cdot A_\varphi, \quad A_\varphi \in M_n(K)$$

此时称 A_φ 是 φ 在基 $\alpha_1, \cdots, \alpha_n$ 下的坐标矩阵. 试证明下列结论等价.

(1) $\varphi(\alpha_1), \cdots, \varphi(\alpha_n)$ 是 V 的一组基;

(2) A_φ 是可逆矩阵.

(3) φ 是同构映射.

2. 设 $\mathbb{R}[x]_n = \{ f(x) \in \mathbb{R}[x] \mid \deg f(x) < n \}$ 是多项式环 $\mathbb{R}[x]$ 的子空间. 试对下列的 $\varphi: \mathbb{R}[x]_n \to \mathbb{R}[x]_n$ 求 φ 在基 $1, x, \cdots, x^{n-1}$ 下的坐标矩阵 A_φ，并判断它们是否可逆矩阵.

(1) $\varphi: \mathbb{R}[x]_n \to \mathbb{R}[x]_n$ 定义为: $f(x) \longmapsto f(1-x)$;

(2) $\varphi: \mathbb{R}[x]_n \to \mathbb{R}[x]_n$ 定义为 $f(x) \longmapsto x f'(x)$.

(其中 $f'(x)$ 表示 $f(x)$ 的导数).

3. 设 V 是 n 维 K-向量空间，V^* 是它的对偶空间. 试证明:

(1) $\forall\, x \in V$，函数 $\varphi_x: V^* \to K$，$l \longmapsto l(x)$ 是 V^* 上的线性函数;

(2) 映射 $\varphi: V \to (V^*)^* := V^{**}$，$\varphi(x) = \varphi_x$，是一个同构映射.

~~4. 设 $(\alpha_1, \cdots, \alpha_n)$ 是 V 的一组基. 试求下列线性映射~~
$\varphi: V \to V$ 在 $(\alpha_1, \cdots, \alpha_n)$ 下的坐标矩阵 A_φ，和 A_φ^{-1}.

(1) $\varphi = \varphi_{ij}: V \to V$ 使得: $\varphi(\alpha_k) = \alpha_k$，$(k \neq i, j)$，且 $\varphi(\alpha_i) = \alpha_j$，$\varphi(\alpha_j) = \alpha_i$.

(2) $\varphi = \varphi_{ij}(\lambda): V \to V$ 使得.

4. 设 $e_i = [0,\cdots,0,\overset{i}{1},0,\cdots,0] \in K^n$ $(i=1,2,\cdots,n)$ 是 K^n 的标准基. 试求下列线性映射 $K^n \xrightarrow{\varphi} K^n$ 在基 $e_1,\cdots,e_n$ 下的坐标矩阵 A_φ.

(1) $[x_1,\cdots,x_i,\cdots,x_j,\cdots x_n] \xmapsto{\varphi_{ij}} [x_1,\cdots,x_j,\cdots,x_i,\cdots,x_n]$

(~~解~~即: 交换 x_i 与 x_j 的位置.

(2) 固定 $\lambda \in K$, $[x_1,\cdots,x_i,\cdots,x_j,\cdots,x_n] \xmapsto{\varphi_{ij}(\lambda)} [x_1,\cdots,x_i,\cdots,x_j+\lambda x_i,\cdots,x_n]$

(即: 将 x_i 乘以 λ 加到 x_j).

(3) 固定 $\lambda \in K$ 非零, $[x_1,\cdots,x_i,\cdots,x_n] \xmapsto{\varphi_i(\lambda)} [x_1,\cdots,\lambda x_i,\cdots,x_n]$.

5. 设 $A \in M_{m\times s}(K)$, $B \in M_{s\times n}(K)$. $K^s \xrightarrow{\varphi_A} K^m$ 和 $K^n \xrightarrow{\varphi_B} K^s$ 分别是以 A, B 为坐标矩阵的线性映射. 试证明:

(1) $\varphi_{AB} = \varphi_A \cdot \varphi_B$: $K^n \to K^m$ 是以 AB 为坐标矩阵的线性映射

(2) $Im(\varphi_{AB}) \subset Im(\varphi_A)$; $ker(\varphi_B) \subseteq ker(\varphi_{AB})$.

(3) $r(AB) \le min\{r(A), r(B)\}$, $r(A)$ 表示 A 的秩.

(4) 令 $U := Im(\varphi_B) \subset K^s$, $V = Im(\varphi_{AB}) \subset K^m$ 则 $f := \varphi_A|_U : U \to K^m$ 是线性映射, 且 $Im(f) = V$.

(5) $ker(f) \subset ker(\varphi_A)$

(6) $r(A) + r(B) - s \le r(AB)$. (提示: 维数定理 3.3.1 (3)).

6. 设 $A, B \in M_{m\times n}(K)$, 令 $\varphi_A, \varphi_B : K^n \to K^m$ 分别是以 A, B 为坐标矩阵的线性映射. 证明:

(1) $Im(\varphi_{A+B}) \subset Im(\varphi_A) + Im(\varphi_B)$

(2) $r(A+B) \le r(A) + r(B)$.

7. 设 $A, B, C \in M_n(K)$, 如果 $ABC = 0$, 证明:

$$r(A) + r(B) + r(C) \le 2n.$$

8. 证明: 线性映射 $\varphi: V \to W$ 是单射的充要条件是 $ker(\varphi) = \{0\}$

9. 设 $V=\mathbb{R}[x]_n=\{f(x)\in\mathbb{R}[x] \mid \deg f(x)<n\}$. 试证明:

(1) $l_k: V\longrightarrow\mathbb{R}$, $l_k(f(x))=\frac{f^{(k)}(0)}{k!}$。(其中 $f^{(k)}(0)$ 表示 $f(x)$ 在 $x=0$ 的 k 阶导数), 是线性函数.

(2) $l_0, l_1, \cdots, l_{n-1}\in V^*$ 是 $1, x, \cdots, x^{n-1}$ 的对偶基.

10. 设 V 是 n 维 K-向量空间, V^* 是它的对偶空间, $l_1,\cdots,l_m\in V^*$

$$\varphi: V\longrightarrow K^m,\quad x\longmapsto[l_1(x),\cdots,l_m(x)].$$

试证明: $\dim_K(\ker(\varphi))=n-\dim_K\langle l_1,\cdots,l_m\rangle$. 此处, $\langle l_1,\cdots,l_m\rangle=\{\lambda_1 l_1+\cdots+\lambda_m l_m \mid \forall \lambda_i\in K\}\subset V^*$ 是由 $l_1,\cdots,l_m$ 在 V^* 中生成的子空间.

§3.4. 多重线性函数.

n 维 K-向量空间 V 上的多重线性函数不仅数学意义重要, 也有极丰富的应用场景。本节我们仅讨论 2 重线性函数(亦称双线性函数) 和 n 重反对称线性函数。

<u>定义 3.4.1</u>. 设 $V_1,\cdots,V_m$ 是 K-向量空间. 函数

$$f: V_1\times\cdots\times V_m\longrightarrow K$$

称为 <u>m 重线性函数</u>. 如果 $f(X_1,\cdots,X_m)$ 关于每个变量都是线性的. 即: 对任意 $1\le i\le m$, 有

$$f(X_1,\cdots,X_{i-1},\lambda Y+\mu Z,X_{i+1},\cdots,X_m)=\lambda f(X_1,\cdots,X_{i-1},Y,X_{i+1},\cdots,X_m)+\mu f(X_1,\cdots,X_{i-1},Z,X_{i+1},\cdots,X_m).$$

当 $V_1=V_2=\cdots=V_m:=V$ 时, $f(X_1,\cdots,X_m)$ 称为 V 上的 m 重线性函数 ($m=1$ 时, 称 $f(X_1)$ 是线性函数, $m=2$ 称双线性函数). $\square$

例3.4.1：设 $l_i: V_i \to K$ 是线性函数，则

$$f = l_1 \otimes \cdots \otimes l_m : V_1 \times \cdots \times V_m \to K,$$

其中 $f(x_1, x_2, \cdots, x_m) := l_1(x_1) \cdot l_2(x_2) \cdots l_m(x_m)$，是一个 m-重线性函数。特别，当 $V_1 = V_2 = \cdots = V_m := V$ 时，$\forall\, l_1, \cdots, l_m \in V^*$，

$$f(x_1, x_2, \cdots, x_m) := l_1(x_1) \cdots l_m(x_m)$$

是 V 上的一个 m重线性函数。

例3.4.2：设 $f: V \times V \to K$ 是一个双线性函数，$e_1, e_2, \cdots, e_n \in V$ 是一组基。则，$\forall\, x, y \in V$，有

$$f(x,y) = (e_1^*(x), \cdots, e_n^*(x)) \begin{pmatrix} f(e_1,e_1), f(e_1,e_2), \cdots, f(e_1,e_n) \\ f(e_2,e_1), f(e_2,e_2), \cdots, f(e_2,e_n) \\ \vdots \\ f(e_n,e_1), f(e_n,e_2), \cdots, f(e_n,e_n) \end{pmatrix} \begin{pmatrix} e_1^*(y) \\ e_2^*(y) \\ \vdots \\ e_n^*(y) \end{pmatrix}$$

反之，$\forall\, A = (a_{ij})_{n\times n}$，存在唯一的双线性函数

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

使得 $f(e_i, e_j) = a_{ij}$。即 $f(x,y)$ 定义为

$$f(x,y) = (e_1^*(x), \cdots, e_n^*(x))\, A \begin{pmatrix} e_1^*(y) \\ \vdots \\ e_n^*(y) \end{pmatrix}.$$

□.

定义3.4.2：V 上的一个 m重线性函数

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

称为 m重对称线性函数。如果，$\forall\, \pi \in S_m$，$f(x_1, \cdots, x_m) = f(x_{\pi(1)}, \cdots, x_{\pi(m)})$。
称为 m重反对称线性函数，如果，$\forall\, \pi \in S_m$，有

$$f(x_1, \cdots, x_m) = \varepsilon_\pi f(x_{\pi(1)}, \cdots, x_{\pi(m)}).\quad \varepsilon_\pi \text{ 是 } \pi \text{ 的符号。}$$

□.

注记：在域 K 中，如果 $1=-1$，则称 K 是特征为2的域，记为 $\mathrm{char}(K)=2$。此时，$f(x_1,\cdots,x_m)$ 对称当且仅当 $f(x_1,\cdots,x_m)$ 是反对称的。

引理3.4.1：设 $f: V\times\cdots\times V \to K$ 是一个 m 重线性函数。则当 $\mathrm{char}(K)\neq 2$（即 $1\neq -1$）时，下述条件等价：

(1) $f(x_1,\cdots,x_m)$ 是反对称函数

(2) 如果存在 $i\neq j$ 使 $x_i=x_j$，则 $f(x_1,\cdots,x_m)=0$

(3) 如果 $x_1,\cdots,x_m$ 线性相关，则 $f(x_1,\cdots,x_m)=0$

证明：(1) ⇒ (2) ⇒ (3) ⇒ (1)，证明 (3) ⇒ (1)：

$\forall\ x_1,\cdots,x_m\in V$，$s<t$，令 $y_1,\cdots,y_s,\cdots,y_t,\cdots,y_m\in V$ 定义如下：$y_i=x_i$（当 $i\neq s,t$ 时），$y_s=x_s+x_t$，$y_t=x_s+x_t$。

则 $y_1,\cdots,y_m$ 线性相关，由 (3) 得 $f(y_1,\cdots,y_m)=0$。

但 $f(y_1,\cdots,y_s,\cdots,y_t\cdots y_m)=f(y_1,\cdots,x_s,\cdots,y_t,\cdots,y_m)+f(y_1,\cdots,x_t,\cdots,y_t,\cdots,y_m)$

$=\underline{f(x_1,\cdots,x_s,\cdots,x_s,\cdots,x_m)}+f(x_1,\cdots,x_s,\cdots,x_t,\cdots,x_m)+f(x_1,\cdots,x_t,\cdots,x_s,\cdots,x_m)$

$+\underline{f(x_1,\cdots,x_t,\cdots,x_t,\cdots,x_m)}=f(x_1,\cdots,x_s,\cdots,x_t,\cdots,x_m)+f(x_1,\cdots,x_t,\cdots,x_s,\cdots,x_m)$

故 $f(x_1,\cdots,x_s,\cdots,x_t,\cdots,x_m)=-f(x_1,\cdots,x_t,\cdots,x_s,\cdots,x_m)$。

(1) ⇒ (2)：无妨设 $i<j$，但 $x_i=x_j$。由 (1) 得

$$f(x_1,\cdots,x_i,\cdots,x_j\cdots,x_m)=-f(x_1,\cdots,x_j,\cdots,x_i\cdots,x_m).$$

由 $x_i=x_j$ 得 $f(x_1,\cdots,x_m)=-f(x_1,\cdots,x_m)$。但 $\mathrm{char}(K)\neq 2$。

故 $f(x_1,\cdots,x_m)=0$

(2) ⇒ (3)：$x_1,\cdots,x_m$ 线性相关，故存在 x_i 可由其它 x_j 线性表出。

即：$x_i=\sum_{j\neq i}\lambda_j x_j$。故 $f(x_1,\cdots,x_m)=0$（利用 (2)）

∎

推论3.4.1. 当 $\mathrm{Char}(K)\neq 2$，$m>\dim_K(V)$ 时，V 上没有非零的反对称 m 重线性函数。

$\square$

令 $L_m(V)$ 表示 V 上所有 m 重线性函数的集合（包括零函数），则 $L_m(V)$ 关于函数加法和数乘法是一个 K-向量空间。令

$S^m(V^*)=\{f\in L_m(V)\mid f 对称\}$，$\wedge^m(V^*)=\{f\in L_m(V)\mid f 反对称\}$，

则 $S^m(V^*)$，$\wedge^m(V^*)$ 是 $L_m(V)$ 的子空间，且当 $\mathrm{Char}(K)\neq 2$ 时，

$$S^m(V^*)\cap\wedge^m(V^*)=\{0\}.$$

定理3.4.1 [存在性]. 设 $\dim_K(V)=n$，$l_1,\cdots,l_m\in V^*$. 则

(1) $f(X_1,\cdots,X_m):=l_1(X_1)\cdot\cdots\cdot l_m(X_m)$ 是一个 m 重线性函数，且当 $l_1,\cdots,l_m$ 是非零函数时，$f(X_1,\cdots,X_m)$ 也是非零函数。

(2) 对任意 $\pi\in S_m$，任意 $f\in L_m(V)$，函数

$$f_\pi: V\times\cdots\times V\to K,\quad f_\pi(X_1,\cdots,X_m):=f(X_{\pi(1)},\cdots,X_{\pi(m)})$$

也是一个 m 重线性函数，且

$$S(f):=\sum_{\pi\in S_m}f_\pi,\qquad A(f):=\sum_{\pi\in S_m}\varepsilon_\pi f_\pi.$$

分别是对称和反对称 m 重线性函数。

(3) 令 $f(X_1,\cdots,X_n)=l_1(X_1)\cdots\cdot l_n(X_n)$，则当 $l_1,\cdots,l_n\in V^*$ 是一组基时，$D(l_1,\cdots,l_n):=A(f)$ 是非零反对称 n 重线性函数。

证明：(1) 显然 $f(X_1,\cdots,X_m):=l_1(X_1)\cdots l_m(X_m)$ 是 m 重线性函数，只需证明，若 $l_1,\cdots,l_m$ 非零，则 f 亦非零。因 $l_i\neq 0$，故存在 $X_i\in V$ 使 $l_i(X_i)\neq 0$，所以 $f(X_1,\cdots,X_m)=l_1(X_1)\cdots l_m(X_m)\neq 0$.

(2) f_π 显然是 m 重线性函数，从而 $S(f)$，$A(f)$ 亦然。$\forall\,\sigma\in S_m$，

$$S(f)(X_{\sigma(1)},\cdots,X_{\sigma(m)})=\sum_{\pi\in S_m}f_\pi(X_{\sigma(1)},\cdots,X_{\sigma(m)})=\sum_{\pi\in S_m}f(X_{\pi\sigma(1)},\cdots,X_{\pi\sigma(m)}).$$

根据定义，$f_{\pi\sigma}(X_1,\cdots,X_m)=f(X_{\pi\sigma(1)},\cdots,X_{\pi\sigma(m)})$，故

$$S(f)=\sum_{\pi\in S_m} f_{\pi\sigma}(X_1,\cdots,X_m)=\Big(\sum_{\pi\in S_m} f_{\pi\sigma}\Big)(X_1,\cdots,X_m)=S(f)(X_1,\cdots,X_m).$$

同理，$A(f)(X_{\sigma(1)},\cdots,X_{\sigma(m)})=\sum_{\pi\in S_m}\varepsilon_\pi f_{\pi\sigma}(X_1,\cdots,X_m)=\varepsilon_\sigma\cdot A(f)(X_1,\cdots,X_m).$

(3) 考虑线性映射 $\varphi: V\to K^n$，$\varphi(x)=[l_1(x),\cdots,l_n(x)]$，

$l_1,\cdots,l_n$ 线性无关当且仅当 φ 是同构映射。令 $e_i\in V$ 使

$$\varphi(e_i)=[0,\cdots,0,1,0\cdots,0]\in K^n.\ (i=1,\cdots,n).$$

则 $e_1,\cdots,e_n$ 是 V 的一组基使得 $e_i^*=l_i$。故

$$D(l_1,\cdots,l_n)(e_1,\cdots,e_n)=l_1(e_1)\cdots l_n(e_n)+\sum_{\pi\neq 1}\varepsilon_\pi l_1(e_{\pi(1)})\cdots l_n(e_{\pi(n)})=1.$$

□.

上述定理表明，在 n 维 K-向量空间 V 上存在非零的反对称 n 重线性函数。下面的定理告诉我们，当 $\mathrm{Char}(K)\neq 2$ 时，V 上的反对称 n 重线性函数本质上是唯一的。

定理3.4.2（唯一性）. 设 $f(X_1,\cdots,X_n)$ 是 n 维 K-向量空间 V 上的反对称 n 重线性函数，$\alpha_1,\cdots,\alpha_n$，$\beta_1,\cdots,\beta_n\in V$ 是两组向量，如果

$$\begin{cases}\beta_1=a_{11}\alpha_1+a_{12}\alpha_2+\cdots+a_{1n}\alpha_n\\ \beta_2=a_{21}\alpha_1+a_{22}\alpha_2+\cdots+a_{2n}\alpha_n\\ \quad\vdots\\ \beta_n=a_{n1}\alpha_1+a_{n2}\alpha_2+\cdots+a_{nn}\alpha_n\end{cases}\qquad a_{ij}\in K.$$

则，当 f 满足条件 (*)：如果 $X_i=X_j$ $(i\neq j)$，则 $f(X_1,\cdots,X_n)=0$ 时，有

$$f(\beta_1,\beta_2,\cdots,\beta_n)=\Big(\sum_{\pi\in S_n}\varepsilon_\pi a_{1i_1}a_{2i_2}\cdots a_{ni_n}\Big)f(\alpha_1,\cdots,\alpha_n).$$

特别，当 $\alpha_1,\cdots,\alpha_n\in V$ 是一组基时，f 由 $f(\alpha_1,\cdots,\alpha_n)$ 唯一确定。

证明：由 $f(X_1,\cdots,X_n)$ 是 n 重线性函数可知，

$$f(\beta_1,\cdots,\beta_n)=\sum_{1\le i_1,\cdots,i_n\le n} a_{1i_1}\cdot a_{2i_2}\cdots a_{ni_n}\, f(\alpha_{i_1},\alpha_{i_2},\cdots,\alpha_{i_n}).$$

其中 $(i_1,\cdots,i_n)$ 取遍所有 $1\le i_1\le n$，$1\le i_2\le n$，$\cdots$，$1\le i_n\le n$。

由于$f(X_1,\cdots,X_n)$是反对称的，且满足条件(2)，则~~...~~，

$$f(\alpha_{i_1},\cdots,\alpha_{i_n})=0 \quad (\text{当 } i_1,i_2,\cdots,i_n \text{ 中有两个相同整数时}).$$

故

$$f(\beta_1,\beta_2,\cdots,\beta_n)=\sum_{\pi\in S_n} a_{1\pi(1)}a_{2\pi(2)}\cdots a_{n\pi(n)} f(\alpha_{\pi(1)},\alpha_{\pi(2)},\cdots,\alpha_{\pi(n)})$$
$$=\Big(\sum_{\pi\in S_n}\varepsilon_\pi a_{1\pi(1)}a_{2\pi(2)}\cdots a_{n\pi(n)}\Big) f(\alpha_1,\alpha_2,\cdots,\alpha_n).$$

□

定义3.4.3：设 $A=[A_{(1)},A_{(2)},\cdots,A_{(n)}]=(a_{ij})_{n\times n}\in M_n(K)$，则

$$\det[A_{(1)},A_{(2)},\cdots,A_{(n)}]:=|A|:=\sum_{\pi\in S_n}\varepsilon_\pi a_{1\pi(1)}a_{2\pi(2)}\cdots a_{n\pi(n)}\in K$$

称为矩阵A的行列式。

□

定理3.4.3 [行列式基本性质]. 设K^n是n维行向量空间，

$$\det: K^n\times\cdots\times K^n=M_n(K)\longrightarrow K$$
$$A=[A_{(1)},A_{(2)},\cdots,A_{(n)}]\longmapsto \det[A_{(1)},\cdots,A_{(n)}]=|A|$$

看成矩阵A的行向量的函数。则

(D1). $\det[A_{(1)},\cdots,A_{(n)}]$是关于行向量的反对称函数。

(D2). $\det[A_{(1)},\cdots,A_{(n)}]$是关于行向量的n重线性函数。

(D3). $\det[I_{(1)},\cdots,I_{(n)}]=1$，其中$I_n=[I_{(1)},\cdots,I_{(n)}]$是单位矩阵。

(D4) 如果A中有两行相同，则$\det[A_{(1)},\cdots,A_{(n)}]=|A|=0$。

(D5) 对任意$\lambda\in K$，$\det[\lambda A_{(1)},\cdots,\lambda A_{(n)}]=\det(\lambda A)=\lambda|A|$.

(D6) 如果A的某行为零，则$\det(A)=|A|=0$.

(D7) 如果对A的行向量进行初等变换(III)，则$\det(A)$保持不变。

证明：设$V=K^n$是n维行向量空间，$e_1,\cdots,e_n\in V$是标准基，

$A=[A_{(1)},\cdots,A_{(n)}]=(a_{ij})_{n\times n}$，由~~[illegible]~~定理3.4.1(3)，存在反对称 n 重线性函数 $D(e_1^*,\cdots,e_n^*): V\times\cdots\times V\longrightarrow K$ 使得 $D(e_1^*,\cdots,e_n^*)(e_1,\cdots,e_n)=1$。由于

$$\begin{cases} A_{(1)}=a_{11}e_1+a_{12}e_2+\cdots+a_{1n}e_n\\ A_{(2)}=a_{21}e_1+a_{22}e_2+\cdots+a_{2n}e_n\\ \vdots\\ A_{(n)}=a_{n1}e_1+a_{n2}e_2+\cdots+a_{2n}e_n \end{cases}$$

则由定理3.4.2，当 $\mathrm{Char}(K)\neq 2$ 时，有

$$D(e_1^*,\cdots,e_n^*)(A_{(1)},\cdots,A_{(n)})=\det[A_{(1)},\cdots,A_{(n)}]=|A|.$$

故结论(D1)至(D7)得证。当 $\mathrm{Char}(K)=2$ 时，只需证明：函数

$$D(e_1^*,\cdots,e_n^*): V\times\cdots\times V\longrightarrow K$$

仍然满足引理3.4.1的(2)即可，~~[illegible]~~从而定理3.4.2仍然成立。在下面的引理中，我们证明引理3.4.1中的(2)对 $D(l_1,\cdots,l_n)$~~[illegible]~~均成立。 □

<u>引理3.4.2</u>：设 V 是任意 n 维~~[illegible]~~ K-向量空间，$l_1,\cdots,l_n\in V^*$，

$$D(l_1,\cdots,l_n)(X_1,X_2,\cdots,X_n):=\sum_{\pi\in S_n}\varepsilon_\pi\, l_1(X_{\pi(1)})\cdots l_n(X_{\pi(n)}).$$

则，当 $X_i=X_j\ (i\neq j)$ 时，$D(l_1,\cdots,l_n)(X_1,\cdots,X_n)=0$。

证明：设 $A_n\subset S_n$ 是所有<u>偶置换的集合</u>，$\bar{A}_n=S_n\setminus A_n$。则

$$D(l_1,\cdots,l_n)(X_1,\cdots,X_n)=\sum_{\pi\in A_n}l_1(X_{\pi(1)})\cdots l_n(X_{\pi(n)})-\sum_{\pi\in\bar{A}_n}l_1(X_{\pi(1)})\cdots l_n(X_{\pi(n)}).$$

若 $X_i=X_j$ 且 $i\neq j$，令 $\sigma=(ij)\in S_n$，则

$$l_1(X_{\pi(1)})\cdots l_n(X_{\pi(n)})=l_1(X_{\pi\sigma(1)})\cdots l_n(X_{\pi\sigma(n)}).$$

又 $\bar{A}_n\longrightarrow A_n$，$\pi\longmapsto\pi\sigma$，是双射，所以

$$\sum_{\pi\in\bar{A}_n}\cancel{\textstyle\sum_{\pi\in A_n}}\, l_1(X_{\pi(1)})\cdots l_n(X_{\pi(n)})=\sum_{\pi\in\bar{A}_n}l_1(X_{\pi\sigma(1)})\cdots l_n(X_{\pi\sigma(n)})$$

$$=\sum_{\pi\in A_n}l_1(X_{\pi(1)})\cdots l_n(X_{\pi(n)}).$$

故 $D(l_1,\cdots,l_n)(X_1,\cdots,X_n)=0$。 □

推论 3.4.2 【行列式乘法定理】. 设 $A=(a_{ij})_{n\times n}$, $B=(b_{ij})_{n\times n}$. 则

$$|A\cdot B| = |A|\cdot|B|.$$

证明：设 $V=K^n$ 是 n 维行向量空间, $e_1,\cdots,e_n\in V$ 是标准基. 令

$$\begin{pmatrix}\alpha_1\\ \vdots\\ \alpha_n\end{pmatrix} := B\begin{pmatrix}e_1\\ \vdots\\ e_n\end{pmatrix},\quad \begin{pmatrix}\beta_1\\ \vdots\\ \beta_n\end{pmatrix} := A\begin{pmatrix}\alpha_1\\ \vdots\\ \alpha_n\end{pmatrix},\quad \text{从而}\ \begin{pmatrix}\beta_1\\ \vdots\\ \beta_n\end{pmatrix} = AB\begin{pmatrix}e_1\\ \vdots\\ e_n\end{pmatrix}.$$

由定理 3.4.2, $D(e_1^*,\cdots,e_n^*)(\beta_1,\cdots,\beta_n) = |AB|\,D(e_1^*,\cdots,e_n^*)(e_1,\cdots,e_n) = |AB|$.
另一方面, $D(e_1^*,\cdots,e_n^*)(\beta_1,\cdots,\beta_n) = |A|\,D(e_1^*,\cdots,e_n^*)(\alpha_1,\cdots,\alpha_n) = |A|\cdot|B|$. 故

$$|A\cdot B| = |A|\cdot|B|.$$

□

对于固定的 $X=(X_1,\cdots,X_n)$, $D(l_1,\cdots,l_n)(X_1,\cdots,X_n)$ 也可看成 V^* 上的 n 重函数 $f_X: V^*\times\cdots\times V^*\longrightarrow K$, $f_X(l_1,\cdots,l_n) := D(l_1,\cdots,l_n)(X_1,\cdots,X_n)$.

命题 3.4.1：固定 $X=(X_1,\cdots,X_n)\in V\times\cdots\times V$, $D(l_1,\cdots,l_n)(X_1,\cdots,X_n)$ 是关于 $l_1,\cdots,l_n$ 的反对称 n 重线性函数, 且如果 $l_i=l_j\ (i\neq j)$ 则 $D(l_1,\cdots,l_n)(X_1,\cdots,X_n)=0$.

证明：$\forall\ \sigma\in S_n$, ~~$D(l_{\sigma(1)},\cdots,l_{\sigma(n)})(X_1,\cdots,X_n)=\sum_{\pi\in S_n}\varepsilon_\pi l_{\sigma(1)}(X_{\pi(1)})\cdots l_{\sigma(n)}(X_{\pi(n)})$~~

$$D(l_{\sigma(1)},\cdots,l_{\sigma(n)})(X_1,\cdots,X_n) = \sum_{\pi\in S_n}\varepsilon_\pi\, l_{\sigma(1)}(X_{\pi(1)})\cdots\cdots l_{\sigma(n)}(X_{\pi(n)})$$

$$= \sum_{\pi\in S_n}\varepsilon_\pi\, l_1(X_{\sigma^{-1}\pi(1)})\cdots l_n(X_{\sigma^{-1}\pi(n)}) = \varepsilon_\sigma\sum_{\pi\in S_n}\varepsilon_{\sigma^{-1}\pi}\, l_1(X_{\sigma^{-1}\pi(1)})\cdots l_n(X_{\sigma^{-1}\pi(n)})$$

$$= \varepsilon_\sigma\, D(l_1,\cdots,l_n)(X_1,\cdots,X_n)$$

所以 $D(l_1,\cdots,l_n)(X_1,\cdots,X_n)$ 关于 $l_1,\cdots,l_n$ 是反对称的. 只需证明它关于 l_1 是线性的, 令 $l_1=\lambda l_1'+\mu l_1''$, 则

$$D(\lambda l_1'+\mu l_1'', l_2,\cdots,l_n)(X_1\cdots X_n) = \sum_{\pi\in S_n}\varepsilon_\pi(\lambda l_1'+\mu l_1'')(X_{\pi(1)})\cdots l_n(X_{\pi(n)})$$

$$= \lambda\sum_{\pi\in S_n}\varepsilon_\pi\, l_1'(X_{\pi(1)})\, l_2(X_{\pi(2)})\cdots l_n(X_{\pi(n)}) + \mu\sum_{\pi\in S_n}\varepsilon_\pi\, l_1''(X_{\pi(1)})\cdots l_n(X_{\pi(n)})$$

$$= \lambda D(l_1', l_2,\cdots,l_n)(X_1,\cdots,X_n) + \mu D(l_1'', l_2,\cdots,l_n)(X_1,\cdots,X_n).$$

如果 $l_i=l_j\ (i\neq j)$, 令 $\sigma=(ij)\in S_n$, 则

$D(l_1,\cdots,l_n)(X_1,\cdots,X_n)=\sum_{\pi\in S_n}\varepsilon_\pi\, l_1(X_{\pi(1)})\cdots l_n(X_{\pi(n)})$. 由于 $l_i=l_j$，故

$l_i(X_{\pi(i)})=l_j(X_{\pi(i)})=l_j(X_{\pi\sigma(j)})$，$l_j(X_{\pi(j)})=l_i(X_{\pi(j)})=l_i(X_{\pi\sigma(i)})$

所以，$l_1(X_{\pi(1)})\cdot\cdots\cdot l_n(X_{\pi(n)})=l_1(X_{\pi\sigma(1)})\cdot\cdots\cdot l_n(X_{\pi\sigma(n)})$. 计算:

$$D(l_1,\cdots,l_n)(X_1,\cdots,X_n)=\sum_{\pi\in A_n}l_1(X_{\pi(1)})\cdots l_n(X_{\pi(n)})-\sum_{\pi\in\bar{A}_n}l_1(X_{\pi(1)})\cdots l_n(X_{\pi(n)})$$

但 $\sum_{\pi\in\bar{A}_n}l_1(X_{\pi(1)})\cdot\cdots\cdot l_n(X_{\pi(n)})=\sum_{\pi\in\bar{A}_n}l_1(X_{\pi\sigma(1)})\cdot\cdots\cdot l_n(X_{\pi\sigma(n)})$

$=\sum_{\pi\in A_n}l_1(X_{\pi(1)})\cdot\cdots\cdot l_n(X_{\pi(n)})$. 因此.

$$D(l_1,\cdots,l_n)(X_1,\cdots,X_n)=0.$$

□.

推论3.4.3：设 $\dim_K(V)=n$，$V^*\xrightarrow{\varphi^*}V^*$ 是线性算子 $V\xrightarrow{\varphi}V$ 的对偶算子. 则，$\forall\, l_1,\cdots,l_n\in V^*$，$X_1,\cdots,X_n\in V$，有

(1) $D(l_1,\cdots,l_n)(\varphi(X_1),\cdots,\varphi(X_n))=D(\varphi^*(l_1),\cdots,\varphi^*(l_n))(X_1,\cdots,X_n)$.

(2) 如果 $(l_1',\cdots,l_n')=(l_1,\cdots,l_n)A$，则

$$D(l_1',\cdots,l_n')=|{}^tA|\,D(l_1,\cdots,l_n).$$

(3). 对任意矩阵 $A=(a_{ij})_{n\times n}\in M_n(K)$，有 $|A|=|{}^tA|$.

证明: (1) 回忆: $\forall\,\alpha\in V$，$\varphi^*(l_i)(\alpha)=l_i(\varphi(\alpha))$. 故

$$D(l_1,\cdots,l_n)(\varphi(X_1),\cdots,\varphi(X_n))=\sum_{\pi\in S_n}\varepsilon_\pi\,\varphi^*(l_1)(X_{\pi(1)})\cdots\varphi^*(l_n)(X_{\pi(n)})$$
$$=D(\varphi^*(l_1),\cdots,\varphi^*(l_n))(X_1,\cdots,X_n).$$

(2) 定理3.4.2的直接应用. 只需注意到在定理3.4.2中，如果

$(\beta_1,\cdots,\beta_n)=(\alpha_1,\cdots,\alpha_n)A$，则 $f(\beta_1,\cdots,\beta_n)=|{}^tA|\,f(\alpha_1,\cdots,\alpha_n)$.

(3) 定义 $V\xrightarrow{\varphi}V$，$(\varphi(e_1),\cdots,\varphi(e_n))=(e_1,\cdots,e_n)A$，由引理3.3.5，有

$(\varphi^*(e_1^*),\cdots,\varphi^*(e_n^*))=(e_1^*,\cdots,e_n^*)\,{}^tA$.

故 $D(\varphi^*(e_1^*),\cdots,\varphi^*(e_n^*))=|A|\,D(e_1^*,\cdots,e_n^*)$. 又 $D(e_1^*,\cdots,e_n^*)(e_1,\cdots,e_n)=1$.

而 $D(\varphi^*(e_1^*),\cdots,\varphi^*(e_n^*))(e_1,\cdots,e_n)=D(e_1^*,\cdots,e_n^*)(\varphi(e_1),\cdots,\varphi(e_n))=|{}^tA|$. 所以，$|A|=|{}^tA|$. □

因为 $|A|=|{}^tA|$，所以 $\det[A_{(1)},\cdots,A_{(n)}]=\det(A^{(1)},\cdots,A^{(n)})$ 关于行向量成立的结论同样对列向量成立。下面是行列式的计算方法。

推论 3.4.4. 设 $A=(a_{ij})_{n\times n}\in M_n(K)$，则通过若次第(I)类和第(II)类单对行向量的初等变换，A 可化为如下的上三角型矩阵。

$$\bar{A}=\begin{pmatrix}\bar{a}_{11} & \bar{a}_{12},\cdots,\bar{a}_{1n}\\ 0 & \bar{a}_{22}\ \cdots\ \bar{a}_{2n}\\ \vdots & \ddots\ \ \vdots\\ 0 & 0\ \cdots\ \bar{a}_{nn}\end{pmatrix},$$

在此过程中，如果施行了 q 次第(I)类初等变换，则

$$|A|=(-1)^q\,\bar{a}_{11}\cdot\bar{a}_{22}\cdots\bar{a}_{nn}.$$

□.

最后讨论 V上的双线性函数 $f: V\times V\to K$，若固定一组基 $e_1,\cdots,e_n\in V$，则 $f(x,y)=(e_1^*(x),\cdots,e_n^*(x))M_f[e_1^*(y),\cdots,e_n^*(y)]$，其中

$$M_f=\begin{pmatrix}f(e_1,e_1), f(e_1,e_2),\cdots,f(e_1,e_n)\\ \vdots\\ f(e_n,e_1), f(e_n,e_2),\cdots f(e_n,e_n)\end{pmatrix}=(f(e_i,e_j))_{n\times n}$$

要解决的主要问题是：如何选择一组基 $e_1,\cdots,e_n$ 使 M_f 最简单，并能从中读出 f 的不变量？

令 $L_2(V)$ 表示双线性函数组成的 K-向量空间，$L_2^+(V)$，$L_2^-(V)$ 分别表示 V上对称，反对称双线性函数组成的子空间。则，当 $\mathrm{char}(K)\neq 2$，

$L_2(V)=L_2^+(V)+L_2^-(V)$，$f(x,y)=\frac{1}{2}(f(x,y)+f(y,x))+\frac{1}{2}(f(x,y)-f(y,x))$.

且 $L_2^+(V)\cap L_2^-(V)=\{0\}$。所以，任意双线性函数可以唯一表示成对称双线性函数 f^+ 和反对称双线性函数 f^- 之和：$f=f^++f^-$.

定理 3.4.4：设 $f(x,y)\in L_2^+(V)$，则当 $\mathrm{char}(K)\neq 2$ 时，存在一组基 $e_1,\cdots,e_n\in V$ 使得 $\forall\ x=x_1e_1+\cdots+x_ne_n$，$y=y_1e_1+\cdots+y_ne_n$，有

$$f(x,y)=\lambda_1x_1y_1+\lambda_2x_2y_2+\cdots+\lambda_nx_ny_n,\quad \lambda_i=f(e_i,e_i)\in K.$$

证明：若 $f=0$，结论显然。否则存在 $e_1\in V$ 使 $f(e_1,e_1)\neq 0$（需要 f 对称和 $\mathrm{char}(K)\neq 2$ 的条件）。令 $V_1=\{x\in V\mid f(e_1,x)=0\}\subset V$ 是一个 n-1 维子空间。对 $\dim_K(V)$ 作归纳法，可设存在 $e_2,\cdots,e_n\in V_1$ 使

故 $f(e_i,e_j)=0\ (\forall 2\le i,j\le n)$. 但由 V_1 的定义可知: $f(e_j,e_1)=f(e_1,e_j)=0$ ($j\ge2$). 所以 $M_f=(f(e_i,e_j))_{n\times n}$ 是对角矩阵, $\forall\ x,y\in V$,

$$f(x,y)=\lambda_1 e_1^*(x)e_1^*(y)+\cdots+\lambda_n e_n^*(x)e_n^*(y).\qquad \lambda_i:=f(e_i,e_i).$$

□

定义 3.4.4: $f\in L_2(V)$ 称为非退化, 如果存在一组基 $e_1,\cdots,e_n\in V$ 使得 $M_f=(f(e_i,e_j))_{n\times n}$ 是可逆矩阵.

□

命题 3.4.2. 设 $f\in L_2(V)$. 则 f 非退化当且仅当 ~~对任意~~ 对任意 $y\ne0$ 存在 $x\in V$ 使 $f(x,y)\ne0$. 若 $f\in L_2^-(V)$ 非退化, 则当 $char(K)\ne2$ 时, $dim_K(V)$ 必为偶数.

证明: 若 f 非退化, 则存在$\overset{\text{一组基}}{}$ $e_1,\cdots,e_n\in V$ 使 $M_f=(f(e_i,e_j))_{n\times n}$ 可逆, 从而, $\forall\ y=y_1e_1+\cdots+y_ne_n\ne0$, $M_f[y_1,\cdots,y_n]\ne0$. 故存在 $x=x_1e_1+\cdots+x_ne_n\ne0$ 使 $f(x,y)=(x_1,\cdots,x_n)M_f[y_1,\cdots,y_n]\ne0$. 反之, 任取一组基 $e_1,\cdots,e_n\in V$, 若 M_f 不可逆, 则存在 $[y_1,\cdots,y_n]\ne0$ 使 $M_f[y_1,\cdots,y_n]=0$. 故, 令 $y=y_1e_1+\cdots+y_ne_n$, $f(x,y)=(e_1^*(x),\cdots,e_n^*(x))M_f[y_1,\cdots,y_n]=0\ (\forall x\in V)$ 与条件矛盾. 若 $f\in L_2^-(V)$, 则对任一组基 $e_1,\cdots,e_n\in V$ 有

$${}^tM_f=(f(e_j,e_i))_{n\times n}=-(f(e_i,e_j))_{n\times n}=-M_f.$$

所以 $|M_f|=(-1)^n|M_f|$. 当 f 非退化且 $char(K)\ne2$ 时, 必有 $n=2m$

□.

若 $f\in L_2^-(V)$, 令 $V_0=\{y\in V\mid f(x,y)=0,\forall x\in V\}\subset V$, 则, $\forall\ x\in V_0, y\in V$, 有 $f(x,y)=-f(y,x)=0$. 所以, 当讨论反对称双线性函数 $f(x,y)$ 时, 我们无妨假设 f 非退化.

定理 3.4.5: 设 $f\in L_2^-(V)$ 非退化, $char(K)\ne2$, 则存在一组基 ($n=2m$) $\alpha_1,\beta_1,\alpha_2,\beta_2,\cdots,\alpha_m,\beta_m\in V$ 满足:

(1) $f(\alpha_i,\beta_i)=1$, $1\le i\le m$, (2) $\forall\ i\ne j$, $f(\alpha_i,\alpha_j)=f(\beta_i,\beta_j)=f(\alpha_i,\beta_j)=0$.

证明: 由命题 3.4.2, $dim_K(V)=n=2m$. 对于 m 应用数学归纳法;

当 $m=1$ 时, $\forall\ \alpha_1\in V$, $\alpha_1\ne0$. f 非退化, 故存在 $\beta\in V$ 使 $f(\alpha_1,\beta)\ne0$

令 $\lambda=f(\alpha_1,\beta)\in K$，$\beta_1=\frac{1}{\lambda}\beta\in V$. 则 $f(\alpha_1,\beta_1)=\frac{1}{\lambda}f(\alpha_1,\beta)=1$. 故 f 在基 α_1,β_1 下的坐标矩阵 $M_f=\begin{pmatrix}0&1\\-1&0\end{pmatrix}\in M_2(K)$.

假设定理对 $m-1$ 成立，令 $\langle\alpha_1,\beta_1\rangle$ 表示由 α_1,β_1 生成的子空间，考虑

$$V_1=\{x\in V \mid f(\alpha_1,x)=0,\ f(\beta_1,x)=0\}\subset V$$

因为 α_1,β_1 线性无关，故 $l_1:=f(\alpha_1,\cdot)$，$l_2:=f(\beta_1,\cdot)$ 在 V^* 中线性无关（引理1）从而 $\dim_K(V_1)=n-2=2(m-1)$. 令 $f_1=f|_{V_1\times V_1}: V_1\times V_1\to K$，则 $f_1\in L_2^-(V_1)$.
若 f_1 在 V_1 上非退化，则由归纳假设，存在 $\alpha_2,\beta_2,\cdots,\alpha_m,\beta_m\in V_1$ 满足：

(1) $f(\alpha_i,\beta_i)=1$，$(2\le i\le m)$，(2) $f(\alpha_i,\alpha_j)=f(\beta_i,\beta_j)=f(\alpha_i,\beta_j)=0$ $(2\le i\ne j\le m)$.

从而 $\alpha_1,\beta_1,\alpha_2,\beta_2,\cdots,\alpha_m,\beta_m\in V$ 满足定理要求。所以只需证明 f_1 在 V_1 上非退化。如果 f_1 在 V_1 上退化，则存在非零 $x\in V_1$ 使 $f_1(x,v)=0$ $(\forall v\in V_1)$.
下面证明：$\forall y\in V$，$f(x,y)=0$（从而与 f 在 V 上非退化矛盾）. 令

$$v=y-\lambda_1\beta_1-\lambda_2\alpha_1,\quad \lambda_1:=f(\alpha_1,y),\ \lambda_2=f(y,\beta_1)\in K.$$

则 $f(\alpha_1,v)=0$，$f(\beta_1,v)=0$. 所以 $v\in V_1$，从而 $f(x,v)=f_1(x,v)=0$.
但不难计算 $f(x,y)=f(x,v)$。 □

推论3.4.5：设 $f\in L_2^-(V)$ 非退化，$\mathrm{char}(K)\ne 2$，$\alpha_1,\beta_1,\cdots,\alpha_m,\beta_m$ 是定理3.4.5中的基.
则 f 在基 $\alpha_1,\beta_1,\alpha_2,\beta_2,\cdots,\alpha_m,\beta_m$ 及基 $\beta_1,\beta_2,\cdots,\beta_m,\alpha_1,\cdots,\alpha_m$ 下的坐标矩阵分别是

$$J=\begin{pmatrix}0&1&&&&\\-1&0&&&&\\&&0&1&&\\&&-1&0&&\\&&&&\ddots&&\\&&&&&0&1\\&&&&&-1&0\end{pmatrix}_{n\times n},\quad J_0=\begin{pmatrix}0&-I_m\\I_m&0\end{pmatrix}$$

其中 I_m 表示 m 阶单位阵. □

~~推论3.4.6：若 $f\in L_2^+(V)$，$\mathrm{char}(K)\ne 2$，则存在一组基 $e_1,\cdots,e_n\in V$ 使 f 在 $e_1,e_2,\cdots,e_n$ 下的坐标矩阵为~~

$$\cancel{M_f=\begin{pmatrix}\lambda_1&&&&&\\&\ddots&&&&\\&&\lambda_r&&&\\&&&0&&\\&&&&\ddots&\\&&&&&0\end{pmatrix}_{n\times n},\quad \lambda_i\ne 0\ (1\le i\le r)}.$$

删

引理3.4.3：设 $f\in L_2(V)$ 在基 $e_1,\cdots,e_n\in V$ 下的坐标矩阵为 M_f，在另一组基 $(e_1',\cdots,e_n')=(e_1,\cdots,e_n)T$ 下的坐标矩阵是 M_f'，则

$$M_f'={}^tT M_f T.$$

证明：已知 $M_f'=(f(e_i',e_j'))_{n\times n}$，$M_f=(f(e_i,e_j))_{n\times n}$，$e_i'=(e_1,\cdots,e_n)T^{(i)}$，$e_j'=(e_1,\cdots,e_n)T^{(j)}$。则 $f(e_i',e_j')={}^tT^{(i)}\cdot[f(e_1,e_j'),\cdots,f(e_n,e_j')]$

由 $f(e_i,e_j')=(f(e_i,e_1),\cdots,f(e_i,e_n))T^{(j)}$ 得

$$f(e_i',e_j')={}^tT^{(i)}M_fT^{(j)},\ \text{故}\ M_f'={}^tTM_fT.$$ □.

定义3.4.5：设 M_f 是 $f\in L_2(V)$ 在任意一组基下的坐标矩阵，则 M_f 的秩称为 f 的秩（由引理3.4.3，它与基的选取无关）。 □

推论3.4.6：如果 $f\in L_2^+(V)$ 的秩为 r，$\text{Char}(K)\neq 2$，则存在一组基 $e_1,\cdots,e_n$ 使

$$M_f=\begin{pmatrix}\lambda_1&&&&&\\&\ddots&&&&\\&&\lambda_r&&&\\&&&0&&\\&&&&\ddots&\\&&&&&0\end{pmatrix}_{n\times n}.\quad \lambda_i=f(e_i,e_i)\neq 0.$$

证明：只需证明：如果 $r(A)=r$，则对任意可逆矩阵 P,Q，有 $r(PA)=r$，$r(AQ)=r$。只需注意到：若 $V\xrightarrow{\varphi_A}W$，$V\xrightarrow{\varphi_Q}V$，$W\xrightarrow{\varphi_P}W$ 的坐标矩阵分别是 A,P,Q，则 $r(PA)=\dim_K \mathrm{Im}(\varphi_P\cdot\varphi_A)$，$r(AQ)=\dim_K\mathrm{Im}(\varphi_A\cdot\varphi_Q)$。而 P,Q 可逆 $\Leftrightarrow$ φ_P,φ_Q 是同构映射。故

$r(PA)=\dim_K\mathrm{Im}(\varphi_P\cdot\varphi_A)=\dim_K\mathrm{Im}(\varphi_A)=r(A)$。

同理 $r(AQ)=r(A)$。 □.

定理3.4.6：设 $f\in L_2^+(V)$，$\mathrm{rank}(f)=r$，则

(1) 当 $K=\mathbb{C}$ 时，存在一组基 $e_1,\cdots,e_n\in V$，使

$$f(x,y)=e_1^*(x)e_1^*(y)+\cdots+e_r^*(x)\cdot e_r^*(y),\quad \forall\ x,y\in V.$$

(2) 当 $K=\mathbb{R}$ 时，存在一组基 $e_1,\cdots,e_n$ 使

$$f(x,y)=e_1^*(x)e_1^*(y)+\cdots+e_s^*(x)e_s^*(y)-e_{s+1}^*(x)e_{s+1}^*(y)-\cdots-e_r^*(x)e_r^*(y).$$

(3) 当 $K=\mathbb{R}$ 时，(2) 中的 s 不依赖于基的选取。(惯性定理)。

证明：(1) 由推论 3.4.6，存在一组基 $\alpha_1,\cdots,\alpha_n \in V$ 使

$$f(x,y)=\lambda_1\alpha_1^*(x)\alpha_1^*(y)+\cdots+\lambda_r\alpha_r^*(x)\alpha_r^*(y). \quad \lambda_i=f(\alpha_i,\alpha_i)\neq 0$$

当 $K=\mathbb{C}$ 时，$\sqrt{\lambda_i}\in\mathbb{C}$，令 $e_i=\frac{1}{\sqrt{\lambda_i}}\alpha_i \ (1\le i\le r)$，$e_i=\alpha_i \ (i>r)$。则 $e_1,\cdots,e_r,e_{r+1},\cdots,e_n\in V$ 是一组基。且 $f(e_i,e_j)=0 \ (i\neq j)$，$f(e_i,e_i)=1 \ (1\le i\le r)$，$f(e_i,e_i)=0 \ (i>r)$。故，$\forall\, x,y\in V$，有

$$f(x,y)=e_1^*(x)e_1^*(y)+\cdots+e_r^*(x)\cdot e_r^*(y).$$

(2) 当 $K=\mathbb{R}$ 时，由推论 3.4.6，无妨设存在一组基 $\alpha_1,\cdots,\alpha_n\in V$ 使得

$$f(x,y)=\lambda_1\alpha_1^*(x)\alpha_1^*(y)+\cdots+\lambda_s\alpha_s^*(x)\alpha_s^*(y)-\lambda_{s+1}\alpha_{s+1}^*(x)\alpha_{s+1}^*(y)-\cdots-\lambda_r\alpha_r^*(x)\alpha_r^*(y).$$

其中，$\lambda_i=f(\alpha_i,\alpha_i)>0 \ (1\le i\le s)$，$-\lambda_i=f(\alpha_i,\alpha_i)<0 \ (s<i\le r)$。$f(\alpha_i,\alpha_i)=0 \ (i>r)$。

令 $e_i=\frac{1}{\sqrt{\lambda_i}}\alpha_i \ (1\le i\le r)$，$e_i=\alpha_i \ (r<i\le n)$。则 $e_1,\cdots,e_n\in V$ 是一组基。且 $f(e_i,e_i)=1 \ (1\le i\le s)$，$f(e_i,e_i)=-1 \ (s<i\le r)$，$f(e_i,e_i)=0 \ (i>r)$。$f(e_i,e_j)=0 \ (i\neq j)$，故，$\forall\, x,y\in V$ 有

$$f(x,y)=e_1^*(x)e_1^*(y)+\cdots+e_s^*(x)e_s^*(y)-e_{s+1}^*(x)e_{s+1}^*(y)-\cdots-e_r^*(x)e_r^*(y).$$

(3). 若有另一组基 $\alpha_1,\cdots,\alpha_n\in V$ 使得：$\forall\, x,y\in V$.

$$\begin{aligned}f(x,y)&=\alpha_1^*(x)\alpha_1^*(y)+\cdots+\alpha_t^*(x)\cdot\alpha_t^*(y)-\alpha_{t+1}^*(x)\alpha_{t+1}^*(y)-\cdots-\alpha_r^*(x)\alpha_r^*(y)\\&=e_1^*(x)e_1^*(y)+\cdots+e_s^*(x)e_s^*(y)-e_{s+1}^*(x)e_{s+1}^*(y)-\cdots-e_r^*(x)e_r^*(y).\end{aligned}$$

如果 $s\neq t$，无妨设 $s<t$。令 $W=\langle e_{s+1},\cdots,e_n\rangle\subset V$，$W'=\langle\alpha_1,\cdots,\alpha_t\rangle\subset V$。

则 $\dim_K(W)+\dim_K(W')-\dim_K(W\cap W')=n+t-s-\dim_K(W\cap W')$

若公式：$\dim_K(W+W')=\dim_K(W)+\dim_K(W')-\dim_K(W\cap W')$ 成立。(其中 $W+W':=\{\alpha+\beta \mid \forall\,\alpha\in W,\beta\in W'\}\subset V$ 是子空间)。则

$$\dim_K(W\cap W')\ge t-s>0.$$

故存在 $\beta\in W\cap W'$，$\beta\neq 0$，而从而.

$$\beta = \alpha_1^*(\beta)\alpha_1 + \cdots + \alpha_t^*(\beta)\alpha_t, \quad \alpha_i^*(\beta)=0 \ (i>t)$$
$$= e_{s+1}^*(\beta)e_{s+1} + \cdots + e_n^*(\beta)e_n, \quad e_i^*(\beta)=0 \ (i\le s).$$

故 $f(\beta,\beta)=\alpha_1^*(\beta)^2+\cdots+\alpha_t^*(\beta)^2>0$。但

$$f(\beta,\beta) = -\left(e_{s+1}^*(\beta)^2+\cdots+e_r^*(\beta)^2\right)\le 0. \quad 故矛盾.$$

□

<u>引理3.4.5</u>：设 W, W' 是 V 的子空，则 $W+W'=\{x+y \mid \forall x\in W, y\in W'\}$ 是 V 的子空间 且 $\dim_K(W+W')=\dim_K(W)+\dim_K(W')-\dim_K(W\cap W')$.

证明：直接验证 $W+W'\subset V$ 是子空间. 仅证维数公式。设 $\alpha_1,\cdots,\alpha_s\in W\cap W'$ 是一组基. 分别扩充成 W, W' 的基.

$$\alpha_1,\cdots,\alpha_s,\beta_{s+1},\cdots,\beta_m, \quad \alpha_1,\cdots,\alpha_s,\gamma_{s+1},\cdots,\gamma_{m'}$$

则 $\alpha_1,\cdots,\alpha_s,\beta_{s+1},\cdots,\beta_m,\gamma_{s+1},\cdots,\gamma_{m'}$ 是 $W+W'$ 的一组基. 故

$$\dim_K(W+W')=\dim_K(W)+\dim_K(W')-\dim_K(W\cap W')$$

□.

<u>推论3.4.7</u>. 设 $\mathrm{char}(K)\neq 2$, $A\in M_n(K)$. 则

(1) ${}^tA=A \iff$ 存在可逆矩阵 $T\in M_n(K)$ 使得

$$ {}^tTAT=\begin{pmatrix} \lambda_1 & & & & & \\ & \ddots & & & 0 & \\ & & \lambda_r & & & \\ & & & 0 & & \\ & 0 & & & \ddots & \\ & & & & & 0\end{pmatrix}_{n\times n}, \quad \lambda_i\neq 0,\ r=r(A).$$

(2) ${}^tA=-A \iff$ 存在可逆矩阵 $T\in M_n(K)$ 使得

$$ {}^tTAT=\begin{pmatrix} \begin{smallmatrix}0&1\\-1&0\end{smallmatrix} & & & & & & 0\\ & \begin{smallmatrix}0&1\\-1&0\end{smallmatrix} & & & & & \\ & & \ddots & & & & \\ & & & \begin{smallmatrix}0&1\\-1&0\end{smallmatrix} & & & \\ & & & & 0 & & \\ & 0 & & & & \ddots & \\ & & & & & & 0\end{pmatrix}_{n\times n}.$$

(3) 如果 $A\in M_n(\mathbb{R})$, ${}^tA=A$, 则存在可逆矩阵 $T\in M_n(\mathbb{R})$ 使

$$ {}^tTAT=\begin{pmatrix} \underbrace{\begin{smallmatrix}1&&\\&\ddots&\\&&1\end{smallmatrix}}_{s} & & & \\ & \underbrace{\begin{smallmatrix}-1&&\\&\ddots&\\&&-1\end{smallmatrix}}_{r-s} & & \\ & & 0 & \\ & & & \ddots \\ & & & & 0\end{pmatrix}_{n\times n}. \quad r=r(A).$$

s 称 A 的正惯性指数.

证明：设 V 是任意 n 维 K-向量空间，$e_1,\cdots,e_n\in V$ 是一组基，令

$$f_A(x,y)=(e_1^*(x),\cdots,e_n^*(x))A[e_1^*(y),\cdots,e_n^*(y)]$$

则 $f_A(x,y)\in L_2(V)$，使得 $A=(f_A(e_i,e_j))_{n\times n}$。

(1) 根据推论3.4.6，存在一组基 $(e_1',\cdots,e_n')=(e_1,\cdots,e_n)T$ 使得

$f_A(e_i',e_j')=0\ (i\neq j)$。$\lambda_i=f(e_i,e_i)\neq 0\ (1\le i\le r)$，$f(e_i,e_i)=0,\ i>r$。

故 $M_f'=(f(e_i',e_j'))_{n\times n}$ 是(1)中的对角矩阵。由引理3.4.3，

$${}^tTAT={}^tM_f'=\begin{pmatrix}\lambda_1&&&&\\&\ddots&&&\\&&\lambda_r&&\\&&&0&\\&&&&\ddots\ 0\end{pmatrix}_{n\times n}\quad \lambda_i\neq 0.$$

(2) 和 (3)：结合分别应用推论3.4.5，定理3.4.6 及引理3.4.3 即得证。 □

定义3.4.6. 任意二次齐次多项式 $q(x_1,\cdots,x_n)\in K[x_1,\cdots,x_n]$ 称为二次型。 □

引理3.4.6. [二次型的矩阵表示]. 设 $q(x_1,\cdots,x_n)$ 是一二次型，$char(K)\neq 2$，则存在唯一的对称矩阵 A（即：${}^tA=A$）使

$$q(x_1,\cdots,x_n)=(x_1,\cdots,x_n)A[x_1,\cdots,x_n]$$

证明：设 $q(x_1,\cdots,x_n)=\sum_{i\le j}q_{ij}x_ix_j$ 是一个二次型，令

$$A=(a_{ij})_{n\times n}，其\ a_{ij}=a_{ji}=\frac{1}{2}q_{ij}\ (i<j),\ a_{ii}=q_{ii}.$$

则 $q(x_1,\cdots,x_n)=(x_1,\cdots,x_n)A[x_1,\cdots,x_n]$。，但

~~$X=[x_1,\cdots,x_n]$~~

~~$f(X,Y)={}^tXAY$~~

若 $q(x_1,\cdots,x_n)=(x_1,\cdots,x_n)B[x_1,\cdots,x_n]$，${}^tB=B$，则比较系数即得 $A=B$。 □

推论3.4.8. 设 $q(x_1,\cdots,x_n)={}^tXAX$，$X=[x_1,\cdots,x_n]$。$char(K)\neq 2$ 则存在变量替换

$$\begin{pmatrix}x_1\\\vdots\\x_n\end{pmatrix}=T\begin{pmatrix}y_1\\\vdots\\y_n\end{pmatrix}，(T\in M_n(K)\text{ 是可逆矩阵})$$

使 $q(x_1,\cdots,x_n)=\lambda_1y_1^2+\cdots+\lambda_ry_r^2$，$r=r(A)$。$\lambda_i\neq 0$

证明：存在可逆矩阵 $T\in M_n(K)$ 使 ${}^tTAT=\begin{pmatrix}\lambda_1&&&\\&\ddots\ \lambda_r&&\\&&0&\\&&&\ddots\ 0\end{pmatrix}$，故

$q(x_1,\cdots,x_n)={}^tX\,{}^tT^{-1}\begin{pmatrix}\lambda_1&&&\\&\ddots\ \lambda_r&&\\&&0&\\&&&\ddots\ 0\end{pmatrix}T^{-1}X$，令 $\begin{pmatrix}y_1\\\vdots\\y_n\end{pmatrix}=T^{-1}X$ 即可 □

习题3.4

1. 计算下列行列式

(a) $\begin{vmatrix} 1 & -1 & 1 & -2 \\ 1 & 3 & -1 & 3 \\ -1 & -1 & 4 & 3 \\ -3 & 0 & -8 & -13 \end{vmatrix}$, (b) $\begin{vmatrix} 1 & n & n & \cdots & n \\ n & 2 & n & \cdots & n \\ n & n & 3 & \cdots & n \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ n & n & n & \cdots & n \end{vmatrix}$ (c) $\begin{vmatrix} a & b & \cdots & b & b \\ b & a & \cdots & b & b \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ b & b & \cdots & a & b \\ b & b & \cdots & b & a \end{vmatrix}$.

2. 不必计算行列式，证明下述4阶行列式

$$\begin{vmatrix} 1 & 7 & 9 & 8 \\ 2 & 1 & 3 & 9 \\ 3 & 2 & 5 & 5 \\ 4 & 8 & 6 & 7 \end{vmatrix}$$

可被1798，2139，3255，4867的最大公因数整除。

3. 下面的函数是否某个K-向量空间V上的对称双线性函数（若是，需指出相应的K-向量空间，若不是，需说明理由）。

(1) $f(x,y)={}^t x\cdot y$.（$x,y\in K^n$ 是列向量空间）.

(2) $f(A,B)=tr({}^t A\cdot B)$.（$A,B\in M_{m\times n}(K)$，$\forall C\in M_n(K)$，$tr(C)$表示对角线元素之和.）

(3) $f(A,B)=\det(A\cdot B)$.（$A,B\in M_n(K)$）.

(4) $f_{ij}(A,B)=A\cdot B$在(i,j)处的值.

(5) $f(u,v)=\int_a^b u\cdot v\, x^2\, dx$，（$u(x),v(x)$是$[a,b]$上连续函数）

(6) $f(u,v)=\int_a^b (u+v)^2 dx$，（$u(x),v(x)$是$[a,b]$上连续函数）

(7) $f(u,v)=(u\cdot v)(a)$，（$u,v\in K[x]$，$a\in K$固定）.

(8) $f(u,v)=\frac{d}{dx}(u\cdot v)(a)$，（$u,v\in K[x]$，$a\in K$固定）

4. 设$f: V\times V\to K$是一个双线性函数，$\forall\, x,\alpha\in V$，定义

$f_1(x): V\to K,\ \alpha\mapsto f(x,\alpha)$，$f_2(x): V\to K,\ \alpha\mapsto f(\alpha,x)$.

证明 (1) 对任意固定的x，$f_1(x)$，$f_2(x)$是V上的线性函数.

(2) $f_1: V\to V^*,\ x\mapsto f_1(x)$，$f_2: V\to V^*$是$V$到$V^*$的线性映射.

(3) f是对称的 $\Longleftrightarrow f_1=f_2$.

(4) f 非退化 $\Leftrightarrow$ $f_1: V\to V^*$, $f_2: V\to V^*$ 是同构映射.

5. 设 $A\in M_n(K)$ 是任意 n 阶方阵. 试证下面3个条件等价:

(1) A 是可逆矩阵; (2) $|A|\neq 0$, (3) $r(A)=n$.

6. 设 $e_1,\cdots,e_n\in V$ 是一组基. 对任意矩阵 $A\in M_n(K)$ 定义:

$$f_A(x,y)=(e_1^*(x),\cdots,e_n^*(x))A[e_1^*(y),\cdots,e_n^*(y)], \quad \forall x,y\in V.$$

证明: 映射 $M_n(K)\longrightarrow L_2(V)$, $A\mapsto f_A$, 是同构映射.
且 f 是对称(反对称) $\Leftrightarrow$ ${}^tA=A$. (${}^tA=-A$).

7. 将下述二次型表示成矩阵形式 ${}^tXAX$:

(a) $Q(x_1,x_2)=x_1^2+3x_1x_2+x_2^2$, (b) $Q(x_1,x_2,x_3)=x_1^2+x_1x_2-x_2^2-2\sqrt{2}x_1x_3+2x_3^2$

8. 设 $A\in M_n(\mathbb{R})$ 可逆, 且 ${}^tA=A$, 证明下述二次型:

$f_A(x_1,\cdots,x_n)=(x_1,\cdots,x_n)A[x_1,\cdots,x_n]$, 与 $f_{A^{-1}}(x_1,\cdots,x_n)=(x_1,\cdots,x_n)A^{-1}[x_1,\cdots,x_n]$
有相同的正惯性指数.

9. 设 $A=(a_{ij})_{n\times n}\in M_n(K)$, $tr(A):=a_{11}+a_{22}+\cdots+a_{nn}$ 称为 A 的迹.
易知 $tr: M_n(K)\to K$ ($A\mapsto tr(A)$) 是一个线性函数. 试证明:

(1) $f: M_n(K)\times M_n(K)\longrightarrow K$, $f(A,B):=tr(AB)$, 是一个非退化的对称双线性函数.

(2) 如果 $l: M_n(K)\longrightarrow K$ 是一个线性函数, 则存在 $A\in M_n(K)$ 使得 $l(X)=tr(A\cdot X)$, $\forall X\in M_n(K)$.

(提示: 利用(1)和前面的命题4).

10. 设 $A=(a_{ij})_{n\times n}\in M_n(\mathbb{R})$, ${}^tA=A$, 证明微分算子:

$$L: \mathbb{R}[x_1,\cdots,x_n]\longrightarrow\mathbb{R}[x_1,\cdots,x_n], \quad f\mapsto L(f)=\sum_{i,j}a_{ij}\frac{\partial^2 f}{\partial x_i\partial x_j}$$

是一个 $\mathbb{R}$-线性映射. 并且验证: 如果 $[y_1,\cdots,y_n]=C\cdot[x_1,\cdots,x_n]$, 其中 $C=(c_{ij})_{n\times n}\in M_n(\mathbb{R})$ 可逆, 则

$$L(f)=\sum_{i,j}b_{ij}\frac{\partial^2 f}{\partial y_i\partial y_j}, \quad 其中 (b_{ij})_{n\times n}={}^tCAC.$$

第四章. 矩阵与行列式

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，线性映射，双线性函数的工具。事实上，矩阵可以看成线性映射和双线性函数在给定基下的坐标，矩阵的运算也是由线性映射的运算运算定义的。但矩阵运算的引入，不仅仅起到了表述方便，简化计算的作用，矩阵本身也脱离线性映射坐标的角色成为一个内容丰富，有巨大的理论和应用价值的研究领域。

§4.1 矩阵的运算.

设 $M_{m\times n}(K)=\{A=(a_{ij})_{m\times n}\mid a_{ij}\in K\}$ 表示所有系数在 K 中的 m 行，n 列矩阵的集合，则 $M_{m\times n}(K)$ 在下列"加法"和数乘法下成为一个 K-向量空间：$\forall\ A=(a_{ij})_{m\times n},\ B=(b_{ij})_{m\times n},\ \lambda\in K$

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

它有一组标准基 $\{E_{ij}\}_{\substack{1\le i\le m\\1\le j\le n}}$，称为矩阵单位：$E_{ij}\in M_{m\times n}(K)$ 表示 (i,j) 位置元素为 1，其它位置元素为 0 的矩阵。它们在 $M_{m\times n}(K)$ 中线性无关，且 $\forall\ A=(a_{ij})_{m\times n}\in M_{m\times n}(K)$，有

$$A=\sum_{i,j}a_{ij}E_{ij}.$$

故 $\dim_K M_{m\times n}(K)=mn$。显然，取"转置"定义了下述 K-向量空间 $M_{m\times n}(K)$ 与 $M_{n\times m}(K)$ 之间的一个同构映射：

$${}^t(\cdot):\ M_{m\times n}(K)\longrightarrow M_{n\times m}(K),\quad A\longmapsto {}^tA.$$

回忆：对于任意 $A=[A_{(1)},\cdots,A_{(m)}]\in M_{m\times n}(K)$ 和 $B=(B^{(1)},\cdots,B^{(s)})\in M_{ns}(K)$，它们的乘积 $A\cdot B$ 定义如下：

$$A\cdot B=(A_{(i)}\cdot B^{(j)})_{m\times s}=\begin{pmatrix}A_{(1)}\cdot B^{(1)}, & A_{(1)}B^{(2)}, & \cdots, & A_{(1)}B^{(s)}\\ A_{(2)}B^{(1)}, & A_{(2)}B^{(2)}, & \cdots, & A_{(2)}B^{(s)}\\ \cdots & \cdots & & \cdots\\ A_{(m)}B^{(1)}, & A_{(m)}B^{(2)}, & \cdots & A_{(m)}B^{(s)}\end{pmatrix}_{m\times s},\quad A_{(i)}\cdot B^{(j)}=(a_{i1},\cdots,a_{in})\begin{pmatrix}b_{1j}\\b_{2j}\\\vdots\\b_{nj}\end{pmatrix}$$

命题4.1.1：若 $A\in M_{m\times n}(K)$, $B\in M_{n\times s}(K)$, $C\in M_{s\times t}(K)$，则

(1) $(A\cdot B)\cdot C=A\cdot(B\cdot C)$. (结合律)

(2) $(A_1+A_2)\cdot B=A_1\cdot B+A_2\cdot B$. ($\forall A_1,A_2\in M_{m\times n}(K)$). (右分配律)

(3) $A\cdot(B_1+B_2)=A\cdot B_1+A\cdot B_2$ ($\forall B_1,B_2\in M_{n\times s}(K)$) (左分配律).

(4), $\forall\lambda\in K$, $(\lambda A)\cdot B=A\cdot(\lambda B)=\lambda(A\cdot B)$. ($\forall A\in M_{m\times n}(K), B\in M_{n\times s}(K)$).

证明：既可通过运算定义直接验证，亦可利用它们运算定义的背景将其归结于下列等价形式的证明。

□.

命题4.1.2：设 T, U, V, W 分别是 t,s,n,m 维 K-向量空间。令 $\mathcal{L}(T,U)$, $\mathcal{L}(U,V)$, $\mathcal{L}(V,W)$ 分别表示 T 到 U, U 到 V, V 到 W 之间所有线性映射的集合。则

(1) $(\varphi\cdot\psi)\cdot\phi=\varphi\cdot(\psi\cdot\phi)$, $\forall\ \phi\in\mathcal{L}(T,U)$, $\psi\in\mathcal{L}(U,V)$, $\varphi\in\mathcal{L}(V,W)$.

(2) $(\varphi_1+\varphi_2)\cdot\psi=\varphi_1\cdot\psi+\varphi_2\cdot\psi$, $\forall\ \varphi_1,\varphi_2\in\mathcal{L}(V,W)$, $\forall\psi\in\mathcal{L}(U,V)$.

(3) $\varphi\cdot(\psi_1+\psi_2)=\varphi\cdot\psi_1+\varphi\cdot\psi_2$, $\forall\ \varphi\in\mathcal{L}(V,W)$, $\forall\psi_1,\psi_2\in\mathcal{L}(U,V)$.

(4) $\forall\ \lambda\in K$, $(\lambda\varphi)\cdot\psi=\varphi\cdot(\lambda\psi)=\lambda(\varphi\cdot\psi)$. $\forall\ \varphi\in\mathcal{L}(V,W)$, $\psi\in\mathcal{L}(U,V)$.

证明：根据定义直接验证（很好的练习）.

□

推论4.1.1：映射 $M_{m\times n}(K)\times M_{n\times s}(K)\xrightarrow{(\cdot,\cdot)}M_{m\times s}(K)$, $(A,B)\mapsto A\cdot B$ 和 $\mathcal{L}(V,W)\times\mathcal{L}(U,V)\longrightarrow\mathcal{L}(U,W)$, $(\varphi,\psi)\longmapsto\varphi\cdot\psi$, 是双线性映射。

证明：这是上述两个命题中 (2), (3), (4) 的统一表述. □.

命题4.1.3：(1). $\forall\ A,B\in M_{m\times n}(K)$, $\lambda\in K$, 有

$$^t({}^tA)=A\ ,\quad {}^t(A+B)={}^tA+{}^tB\ ,\quad {}^t(\lambda A)=\lambda\,{}^tA\ .$$

(2) $\forall\ A\in M_{m\times n}(K)$, $B\in M_{n\times s}(K)$, 则 ${}^t(A\cdot B)={}^tB\,{}^tA$.

证明：(1)是显然的. (2) 可以利用矩阵乘法的定义直接证明，也可利用转置矩阵和矩阵乘法的定义背景归结于线性映射的验证。

证明1：设 $A=[A_{(1)},\cdots,A_{(m)}]$，$B=(B^{(1)},\cdots,B^{(s)})$，故

$$A\cdot B=(A_{(i)}B^{(j)})_{m\times s}.$$

令 ${}^t(A\cdot B)=(C_{ij})_{s\times m}$，则 $C_{ij}=A_{(j)}B^{(i)}$，$1\le i\le s,\ 1\le j\le m$.

但 $A_{(j)}B^{(i)}={}^tB^{(i)}\cdot{}^tA_{(j)}=({}^tB)_{(i)}\cdot({}^tA)^{(j)}=({}^tB\cdot{}^tA)_{ij}$，故

$${}^t(A\cdot B)={}^tB\cdot{}^tA.$$

证明2：设 $U\xrightarrow{\varphi_B}V\xrightarrow{\varphi_A}W$ 分别是以 B,A 为坐标矩阵的线性映射，从而 $\varphi_A\cdot\varphi_B=\varphi_{AB}$. 若 $(\varphi_A\cdot\varphi_B)^*=\varphi_B^*\cdot\varphi_A^*$，由引理3.3.5知 $\varphi_{AB}^*,\varphi_A^*,\varphi_B^*$ 的坐标矩阵分别是 ${}^t(AB)$，${}^tA$，${}^tB$，从而

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

$(\varphi_A\cdot\varphi_B)^*=\varphi_B^*\cdot\varphi_A^*$ 可直接验证：$\forall\ l\in W^*$，$\varphi_B^*\cdot\varphi_A^*(l)=\varphi_B^*(l\cdot\varphi_A)$ $=(l\cdot\varphi_A)\cdot\varphi_B=l\cdot(\varphi_A\cdot\varphi_B)=(\varphi_A\cdot\varphi_B)^*(l)$. 故 $\varphi_B^*\cdot\varphi_A^*=(\varphi_A\cdot\varphi_B)^*$. □

我们接下来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4.1.1 [单位矩阵]：存在唯一的矩阵 $I_n\in M_n(K)$ 使得

$$A\cdot I_n=I_n\cdot A=A,\quad\forall\ A\in M_n(K),$$

称为 n 阶单位矩阵，它从而形如

$$I_n=\begin{pmatrix}1&0&0&\cdots&0\\0&1&0&\cdots&0\\0&0&1&\cdots&0\\\vdots&\vdots&\vdots&\ddots&\vdots\\0&0&0&\cdots&1\end{pmatrix}_{n\times n}$$

例4.1.2 [可逆矩阵]. $A\in M_n(K)$ 称为可逆，如果存在 $B\in M_n(K)$ 使得 $A\cdot B=BA=I_n$. 这样的 B 由 A 唯一确定，称为 A 的逆矩阵，记为 $B:=A^{-1}$.

~~命题4.1.4. 若 A,B 可逆，则 (1) AB 可逆，且 $(AB)^{-1}=B^{-1}A^{-1}$；(2) $({}^tA)^{-1}={}^t(A^{-1})$.~~

下面引入一类特殊的可逆矩阵，它们实际上是组成可逆矩阵最基本的要素，即任意可逆矩阵都是这些特殊矩阵的乘积。为讨论方便我们将单位矩阵 I_n 表示为

$$I_n=[e_1,e_2,\cdots,e_n]=({}^te_1,{}^te_2,\cdots,{}^te_n).$$

其中 $e_1=(1,0,\cdots,0)$，$e_2=(0,1,0,\cdots,0)$，$\cdots$，$e_n=(0,0,\cdots,1)\in K^n$.

定义 4.1.1 [初等矩阵].

(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：$F_{s,t}=[e_1,\cdots,e_{s-1},e_t,e_{s+1},\cdots,e_{t-1},e_s,e_{t+1},\cdots,e_n]$，或

$$F_{s,t}=I_n-E_{ss}-E_{tt}+E_{st}+E_{ts}\quad(E_{ij}\text{ 表示矩阵单位}).$$

(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：

$$F_{s,t}(\lambda)=[e_1,\cdots,e_{s-1},e_s,e_{s+1},\cdots,e_{t-1},e_t+\lambda e_s,e_{t+1},\cdots,e_n]=I_n+\lambda E_{t,s}$$

(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：

$$F_s(\lambda)=[e_1,\cdots,e_{s-1},\lambda e_s,e_{s+1},\cdots,e_n]=I_n+(\lambda-1)E_{ss}.$$

命题 4.1.4. (1) ${}^tF_{s,t}=F_{s,t}$，${}^tF_{s,t}(\lambda)=F_{t,s}(\lambda)$，${}^tF_s(\lambda)=F_s(\lambda)$.

(2) 设 $A=(a_{ij})_{m\times n}\in M_{m\times n}(K)$，$F_{s,t}$，$F_{s,t}(\lambda)$，$F_s(\lambda)$ 表示 m 阶初等矩阵。则

(I). $F_{s,t}\cdot A=[A_{(1)},\cdots,A_{(s-1)},A_{(t)},A_{(s+1)},\cdots,A_{(t-1)},A_{(s)},A_{(t+1)},\cdots,A_{(m)}]$

(II). $F_{s,t}(\lambda)\cdot A=[A_{(1)},\cdots,A_{(s)},\cdots,A_{(t-1)},A_{(t)}+\lambda A_{(s)},A_{(t+1)},\cdots,A_{(m)}]$

(III). $F_s(\lambda)\cdot A=[A_{(1)},\cdots,A_{(s-1)},\lambda A_{(s)},A_{(s+1)},\cdots,A_{(m)}]$.

(3). 若 $F_{s,t}$，$F_{s,t}(\lambda)$，$F_s(\lambda)$ 表示 n 阶初等矩阵，则

$$A\cdot F_{s,t}=(A^{(1)},\cdots,A^{(s-1)},A^{(t)},A^{(s+1)},\cdots,A^{(t-1)},A^{(s)},A^{(t+1)},\cdots,A^{(n)}).$$

$$A\cdot F_{s,t}(\lambda)=(A^{(1)},\cdots,A^{(s-1)},A^{(s)}+\lambda A^{(t)},\cdots,A^{(t)},\cdots,A^{(n)}).$$

$$A\cdot F_s(\lambda)=(A^{(1)},\cdots,A^{(s-1)},\lambda A^{(s)},A^{(s+1)},\cdots,A^{(n)})$$

证明：直接验证，留作练习 □.

推论 4.1.2：设 $A=(a_{ij})_{m\times n}\in M_{m\times n}(K)$ 的秩为 $r(A)=r$，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

$$P_sP_{s-1}\cdots P_1\,A\,Q_1\cdot Q_2\cdot\cdots\cdot Q_t=\begin{pmatrix} I_r & 0\\ 0 & 0\end{pmatrix}.$$

其中 I_r 是 r 阶单位矩阵. 3个 0 分别表示 $r\times(n-r)$, $(m-r)\times r$ 和 $(m-r)\times(n-r)$ 零矩阵.

证明：该结论等价于：秩为 r 的矩阵 A ~~[illegible]~~ 做初等行变换和初等列变换可将 A 化为

$$\begin{pmatrix} I_r & 0\\ 0 & 0\end{pmatrix}.$$

□

对任意 n 阶方阵 $A\in M_n(K)$，和任意 n 维 K-向量空间 V，我们可以定义一个线性映射 $\varphi_A: V\to V$，使得：如果 $v_1,\cdots,v_n\in V$ 是一组基，则 $(\varphi_A(v_1),\cdots,\varphi_A(v_n))=(v_1,\cdots,v_n)A$.

定理 4.4.1：设 $A\in M_n(K)$，$\varphi_A: V\to V$ 是在基 $v_1,\cdots,v_n$ 下以 A 为坐标矩阵的线性映射，则下述结论等价.

(1) 存在 $B\in M_n(K)$ 使 $AB=I_n$,

(2) $\varphi_A: V\to V$ 是满射.

(3) $r(A)=n$ (称 A 非退化);

(4) $\varphi_A: V\to V$ 是单射.

(5) $\varphi_A: V\to V$ 是双射.

(6) A 是可逆矩阵.

证明：(1) ⇒ (2) ⇒ (3) ⇒ (4) ⇒ (5) ⇒ (6) ⇒ (1).

(1) ⇒ (2). 令 $\varphi_B: V \to V$ 是以 B 为坐标矩阵的线性映射：

$$(\varphi_B(v_1), \cdots, \varphi_B(v_n)) = (v_1, \cdots, v_n)B, \quad (\varphi_A(v_1), \cdots, \varphi_A(v_n)) = (v_1, \cdots, v_n)A.$$

则 $\varphi_A \cdot \varphi_B = \varphi_{A\cdot B} = \varphi_{I_n}: V \to V$ 是恒等映射，从而 $\varphi_A: V \to V$ 是满射。

(2) ⇒ (3)：$\varphi_A: V \to V$ 是满射，故 $r(A) = \dim_K Im(\varphi_A) = \dim_K(V) = n$.

(3) ⇒ (4)：$Ker(\varphi_A)$ 的维数等于 $n - r(A) = 0$，故 $Ker(\varphi_A) = \{0\}$，从而 φ_A 是单射。

(4) ⇒ (5)：因 $\varphi_A: V \to V$ 是单射，故 $Ker(\varphi_A) = \{0\}$。应用定理3.3.1(1)

$\dim_K(Im(\varphi_A)) = \dim_K(V)$，所以 φ_A 也是满射，从而 φ_A 是双射。

(5) ⇒ (6)：$\varphi_A: V \to V$ 是双射，故存在逆映射 $\varphi_A^{-1}: V \to V$ 使得 $\varphi_A \cdot \varphi_A^{-1} = \varphi_A^{-1} \cdot \varphi_A = \varphi_{I_n}$。因 φ_A 是线性映射，故 φ_A^{-1} 也是线性映射。令 B 是 φ_A^{-1} 的坐标矩阵，即 $\varphi_A^{-1} = \varphi_B: V \to V$，则

$$AB = BA = I_n$$

(6) ⇒ (1)：显然。

推论4.1.3：(1) 初等矩阵都是可逆矩阵，且

$$F_{s,t}^{-1} = F_{s,t}, \quad F_{s,t}(\lambda)^{-1} = F_{s,t}(-\lambda), \quad F_s(\lambda)^{-1} = F_s(\lambda^{-1})$$

(2) 如果 $A_1, A_2, \ldots, A_m$ 是可逆矩阵，则它们的乘积 $A_1 \cdot A_2 \cdots A_m$ 也可逆，且 $(A_1A_2\cdots A_m)^{-1} = A_m^{-1} A_{m-1}^{-1} \cdots A_2^{-1} \cdot A_1^{-1}$.

(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

$$r(BAC) = r(A), \quad \forall A \in M_{m\times n}(K).$$

证明：(1) $F_{s,t}$, $F_{s,t}(\lambda)$, $F_s(\lambda)$ 分别是下述线性映射在基 $v_1, \cdots, v_n$ 下的坐标矩阵：

$\varphi_{F_{s,t}}: V \longrightarrow V$ 使得 $(v_1,\cdots v_s,\cdots,v_t,\cdots,v_n) \longmapsto (v_1,\cdots,v_t,\cdots,v_s\cdots v_n)$

$\varphi_{F_{s,t}(\lambda)}: V \longrightarrow V$ 使得 $(v_1,\cdots,v_s,\cdots,v_t,\cdots,v_n) \longmapsto (v_1,\cdots,v_s+\lambda v_t,\cdots,v_t,\cdots,v_n)$

$\varphi_{F_s(\lambda)}: V \longrightarrow V$ 使得 $(v_1,\cdots,v_s,\cdots,v_n) \longmapsto (v_1,\cdots,\lambda v_s,\cdots,v_n)$.

它们的逆映射分别是 $\varphi_{F_{s,t}}^{-1} = \varphi_{F_{s,t}}$, $\varphi_{F_{s,t}(\lambda)}^{-1} = \varphi_{F_{s,t}(-\lambda)}$, $\varphi_{F_s(\lambda)}^{-1} = \varphi_{F_s(\lambda^{-1})}$.

故 $F_{s,t}^{-1} = F_{s,t}$, $F_{s,t}(\lambda)^{-1} = F_{s,t}(-\lambda)$, $F_s(\lambda)^{-1} = F_s(\lambda^{-1})$.

(2) 因为 $(A_1\cdot A_2\cdots A_m)\cdot(A_m^{-1}\cdot A_{m-1}^{-1}\cdots A_2^{-1}\cdot A_1^{-1}) = I_n$, 故 $(A_1A_2\cdots A_m)^{-1} = A_m^{-1}\cdots A_1^{-1}$.

(3) 因为 $r(A)=n$, 由推论 4.1.2, 存在初等矩阵 $P_1,\cdots,P_s$, $Q_1,\cdots,Q_t$ 使

$$P_sP_{s-1}\cdots P_1\, A\, Q_1Q_2\cdots Q_t = I_n.$$

从而 $A = P_1^{-1}\cdots P_s^{-1}\cdot I_n Q_t^{-1}\cdots Q_1^{-1} = P_1^{-1}\cdots P_s^{-1}\cdot Q_t^{-1}\cdots Q_1^{-1}$. 初等矩阵的逆矩阵仍然是初等矩阵. 故 A 是初等矩阵的乘积.

(4) 由于 B, C 分别是初等矩阵的乘积. BAC 就是分别对 A 的行向量和 A 的列向量作若干次初等变换所得矩阵. 故 $r(BAC) = r(A)$.

□

上述结论(3)的一个显而易见的推论是: 对任一可逆矩阵 A, 存在初等矩阵 $P_1,\cdots,P_s$ 使 $P_sP_{s-1}\cdots P_2\cdot P_1\cdot A = I_n$. 而 $P_s\cdots P_1\cdot A$ 相当于对 A 的行向量做 s 次初等变换. 这一事实为下述求 A 的逆矩阵 A^{-1} 的算法提供了依据.

推论 4.1.4 [求逆矩阵算法]. 设 $A=(a_{ij})_{n\times n} \in M_n(K)$ 可逆. 令

$$\bar{A} := (A \mid I_n) = \left(\begin{array}{cccc|cccc} a_{11} & a_{12} & \cdots & a_{1n} & 1 & 0 & \cdots & 0 \\ a_{21} & a_{22} & \cdots & a_{2n} & 0 & 1 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} & 0 & 0 & \cdots & 1 \end{array}\right)_{n\times 2n}$$

对 $\bar{A}$ 的行向量依次做初等变换: $(A \mid I_n) \xrightarrow{P_1} (P_1A \mid P_1\cdot I_n) \xrightarrow{P_2}$ $(P_2P_1\cdot A \mid P_2P_1\cdot I_n) \longrightarrow \cdots \xrightarrow{P_s} (P_s\cdots P_2P_1A \mid P_s\cdots P_2\cdot P_1\cdot I_n)$.

直至 $P_s\cdots P_2\cdot P_1\cdot A = I_n$ 时停止. 则 $A^{-1} = P_s\cdots P_2\cdot P_1\cdot I_n$. □

例4.1.3. 如果 $A=\begin{pmatrix}a_{11}&a_{12}\\a_{21}&a_{22}\end{pmatrix}$ 可逆. 求 A^{-1}.

解: A 可逆 $\Leftrightarrow$ $r(A)=2$ $\Leftrightarrow$ $A^{(1)}=[a_{11},a_{21}]$, $A^{(2)}=[a_{12},a_{22}]$ 线性无关.
特别 a_{11},a_{21} 不全为零. 无妨设 $a_{11}\neq 0$. $|A|=a_{11}a_{22}-a_{12}a_{21}\neq 0$

$$(A|I_2)=\left(\begin{array}{cc|cc}a_{11}&a_{12}&1&0\\a_{21}&a_{22}&0&1\end{array}\right)\xrightarrow{F_{12}(-\frac{a_{21}}{a_{11}})}\left(\begin{array}{cc|cc}a_{11}&a_{12}&1&0\\0&\underbrace{a_{22}-\frac{a_{21}}{a_{11}}a_{12}}_{|A|/a_{11}}&-\frac{a_{21}}{a_{11}}&1\end{array}\right)$$

$$\xrightarrow{F_{21}(-\frac{a_{11}}{|A|}a_{12})}\left(\begin{array}{cc|cc}a_{11}&0&1+\frac{a_{12}a_{21}}{|A|}&-\frac{a_{11}}{|A|}a_{12}\\0&\frac{|A|}{a_{11}}&-\frac{a_{21}}{a_{11}}&1\end{array}\right),\quad \left(1+\frac{a_{12}a_{21}}{|A|}=\frac{a_{11}a_{22}}{|A|}\right).$$

$$\xrightarrow[F_2(\frac{a_{11}}{|A|})]{F_1(\frac{1}{a_{11}})}\left(\begin{array}{cc|cc}1&0&\frac{a_{22}}{|A|}&-\frac{a_{12}}{|A|}\\0&1&-\frac{a_{21}}{|A|}&\frac{a_{11}}{|A|}\end{array}\right).\quad 故\ A^{-1}=\frac{1}{|A|}\begin{pmatrix}a_{22}&-a_{12}\\-a_{21}&a_{11}\end{pmatrix}.$$

例4.1.4: 求下列矩阵 A 的逆矩阵

$$A=\begin{pmatrix}1&2&3&4\\2&3&1&2\\1&1&1&-1\\1&0&-2&-6\end{pmatrix}$$

解:

$$(A|I_4)=\left(\begin{array}{cccc|cccc}1&2&3&4&1&0&0&0\\2&3&1&2&0&1&0&0\\1&1&1&-1&0&0&1&0\\1&0&-2&-6&0&0&0&1\end{array}\right)\xrightarrow[\substack{F_{13}(-1)\\F_{14}(-1)}]{F_{12}(-2)}\left(\begin{array}{cccc|cccc}1&2&3&4&1&0&0&0\\0&-1&-5&-6&-2&1&0&0\\0&-1&-2&-5&-1&0&1&0\\0&-2&-5&-10&-1&0&0&1\end{array}\right)$$

$$\xrightarrow[F_{24}(-2)]{F_{23}(-1)}\left(\begin{array}{cccc|cccc}1&2&3&4&1&0&0&0\\0&-1&-5&-6&-2&1&0&0\\0&0&3&1&1&-1&1&0\\0&0&5&2&3&-2&0&1\end{array}\right)\xrightarrow{F_{34}(-\frac{5}{3})}\left(\begin{array}{cccc|cccc}1&2&3&4&1&0&0&0\\0&-1&-5&-6&-2&1&0&0\\0&0&3&1&1&-1&1&0\\0&0&0&\frac{1}{3}&\frac{4}{3}&-\frac{1}{3}&-\frac{5}{3}&1\end{array}\right)$$

$$\xrightarrow[F_4(3)]{F_2(-1),\ F_3(\frac{1}{5})}\left(\begin{array}{cccc|cccc} 1 & 2 & 3 & 4 & 1 & 0 & 0 & 0 \\ 0 & 1 & 5 & 6 & 2 & -1 & 0 & 0 \\ 0 & 0 & 1 & \frac{2}{3} & \frac{1}{3} & -\frac{1}{3} & \frac{1}{3} & 0 \\ 0 & 0 & 0 & 1 & 4 & -1 & -5 & 3 \end{array}\right) \xrightarrow[F_{41}(-4)]{F_{43}(-\frac{2}{3}),\ F_{42}(-6)} \left(\begin{array}{cccc|cccc} 1 & 2 & 3 & 0 & -15 & 4 & 20 & -12 \\ 0 & 1 & 5 & 0 & -22 & 5 & 30 & -18 \\ 0 & 0 & 1 & 0 & -1 & 0 & 2 & -1 \\ 0 & 0 & 0 & 1 & 4 & -1 & -5 & 3 \end{array}\right)$$

$$\xrightarrow[F_{31}(-3)]{F_{32}(-5)} \left(\begin{array}{cccc|cccc} 1 & 2 & 0 & 0 & -12 & 4 & 14 & -9 \\ 0 & 1 & 0 & 0 & -17 & 5 & 20 & -13 \\ 0 & 0 & 1 & 0 & -1 & 0 & 2 & -1 \\ 0 & 0 & 0 & 1 & 4 & -1 & -5 & 3 \end{array}\right) \xrightarrow{F_{21}(-2)} \left(\begin{array}{cccc|cccc} 1 & 0 & 0 & 0 & 22 & -6 & -26 & 17 \\ 0 & 1 & 0 & 0 & -17 & 5 & 20 & -13 \\ 0 & 0 & 1 & 0 & -1 & 0 & 2 & -1 \\ 0 & 0 & 0 & 1 & 4 & -1 & -5 & 3 \end{array}\right).$$

所以 $A^{-1} = \begin{pmatrix} 22 & -6 & -26 & 17 \\ -17 & 5 & 20 & -13 \\ -1 & 0 & 2 & -1 \\ 4 & -1 & -5 & 3 \end{pmatrix}$.

$\square$

~~假设~~在 n 阶方阵集合 $M_n(K)$ 上定义的矩阵加法和乘法满足一系列的规则，下面我们将其中最本质的几条规则抽象成一个重要代数~~对象~~对象的定义。

定义 4.1.2. [环的定义]. 设 R 是一个非空集合，如果在 R 上定义了两个运算：$R\times R \xrightarrow{\varphi_1} R$，$R\times R \xrightarrow{\varphi_2} R$，分别称为"加法"和"乘法"，满足下列规则（为了习惯，通常将 $\varphi_1(a,b)$ 记为 $a+b$，将 $\varphi_2(a,b)$ 记为 $a\cdot b$）：

(1) $\forall\ a,b,c\in R$，有 $(a+b)+c=a+(b+c)$

(2) 存在元素 $0\in K$，使 $a+0=0+a=a$，$\forall a\in R$.

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

(4) $\forall\ a,b\in R$，有 $a+b=b+a$.

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

(6) 存在元素 $1\in R$ 满足：$a\cdot 1=1\cdot a=a$，$\forall a\in R$.

(7) $\forall\ a,b,c\in R$，有

$$a\cdot(b+c)=a\cdot b+a\cdot c,\qquad (a+b)\cdot c=a\cdot c+b\cdot c.$$

则称 $R=(R,+,\cdot)$ 是一个环。~~如果 $a\cdot b=b\cdot a$ ($\forall a,b\in R$) 则称R是交换环~~

如果它的乘法还满足交换律：$a\cdot b=b\cdot a$，则称R是交换环 □

例4.1.5 [矩阵环]. $M_n(K)$ 关于矩阵加法和乘法是一个环(这是环的一个典型例子)，它不是交换环！

□

定理4.1.2 [矩阵环的中心]. 设 $A=(a_{ij})_{n\times n}\in M_n(K)$，则 $A=\lambda\cdot I_n$, $(\lambda\in K)$ 的充要条件是：$A\cdot B=B\cdot A$, $\forall B\in M_n(K)$.

证明：若 $A=\lambda\cdot I_n$，则显然 A 与 $M_n(K)$ 中每个矩阵可交换. 反之，若 A 与 $M_n(K)$ 中每个矩阵可交换，则 $A\cdot E_{ij}=E_{ij}\cdot A$，其中 $E_{ij}\in M_n(K)$ 是矩阵单位.(即：(i,j) 位置元素为1，其它位置元素为0的矩阵) 利用如下公式：

$$(4.1).\quad AE_{ij}=(0,\cdots,0,\underset{\text{第}j\text{列}}{A^{(i)}},0,\cdots,0)\ ;\quad E_{ij}\cdot A=[0,\cdots,0,\underset{\text{第}i\text{行}}{A_{(j)}},0,\cdots,0].$$

可得：$A\cdot E_{ij}=E_{ij}A \iff a_{ij}=0,\ (i\neq j)$，$a_{ii}=a_{jj}$. 故 $A=a_{11}\cdot I_n$ □.

例4.1.6 [多项式环]. 设 K 是一个域，则 $K[x]$ 关于多项式的加法和乘法是一个交换环.

□

例4.1.7：若 $f(x)=a_n x^{d}+a_{n-1}x^{d-1}+\cdots+a_1x+a_0\in K[x]$，$A\in M_n(K)$ 令 $f(A):=a_nA^d+a_{n-1}A^{d-1}+\cdots+a_1A+a_0I_n\in M_n(K)$. 则

$$K[A]:=\{g(A)\mid \forall g(x)\in K[x]\}\subset M_n(K)$$

关于矩阵的加法和乘法是一个环.

证明：$\forall\ g_1(A),g_2(A)\in K[A]$，其中

$$g_1(x)=a_mx^m+a_{m-1}x^{m-1}+\cdots+a_1x+a_0\in K[x],\quad g_2(x)=b_nx^n+b_{n-1}x^{n-1}+\cdots+b_1x+b_0\in K[x].$$

令 $f(x):=g_1(x)+g_2(x)$，$h(x):=g_1(x)\cdot g_2(x)\in K[x]$，则

$$g_1(A)+g_2(A)=f(A),\quad g_1(A)\cdot g_2(A)=h(A).$$

所以 $M_n(K)$ 中的加法和乘法诱导了 $K[A]$ 上的加法和乘法. 且 0（零矩阵）和 I_n（单位矩阵）在 $K[A]$ 中. 不难验证 $K[A]$ 是一个环 □

若 $A\in M_n(K)$, 则 $I_n, A, A^2, \cdots, A^{n^2}$ 在 $M_n(K)$ 中线性相关（因为 $\dim_K M_n(K)=n^2$）. 故存在不全为零的常数 $a_0,a_1,\cdots,a_{n^2}\in K$ 使得 $a_0I_n+a_1A+\cdots+a_{n^2}A^{n^2}=0$。即存在非零多项式 $f(x)=a_{n^2}x^{n^2}+\cdots+a_1A+a_0$ 使得 $f(A)=0$. 令

$$S_A=\{f(x)\in K[x]\mid f(A)=0\}\subset K[x].$$

则 S_A 中有非零多项式。令 $\mu_A(x)$ 是 S_A 中次数最小的首项系数为1的多项式。则有

命题4.1.5：设 $f(x)\in K[x]$，若 $f(A)=0$，则 $\mu_A(x)\mid f(x)$，即存在 $q(x)\in K[x]$ 使 $f(x)=\mu_A(x)\cdot q(x)$.

证明：由多项式的带余除法，存在 $q(x), r(x)\in K[x]$ 使

$$f(x)=\mu_A(x)\cdot q(x)+r(x),\ \text{其中 } r(x)=0 \text{ 或 } \deg r(x)<\deg\mu_A(x).$$

若 $r(x)\neq 0$，则 $\deg r(x)<\deg\mu_A(x)$. 故 $r(A)\neq 0$（否则与 $\mu_A(x)$ 选取矛盾）. 但 $r(x)=f(x)-\mu_A(x)q(x)$，故 $r(A)=f(A)-\mu_A(A)\cdot q(A)=0$. 矛盾，所以 $r(x)=0$，即 $f(x)=\mu_A(x)q(x)$. □

定义4.1.3 [矩阵的极小多项式]. 设 $A\in M_n(K)$，$\mu_A(x)\in K[x]$ 是满足 $\mu_A(A)=0$ 的最小次数的首系1多项式。$\mu_A(x)$ 由 A 唯一确定，故称 $\mu_A(x)$ 是 A 的极小多项式. □

推论4.1.5：$\dim_K(K[A])=\deg\mu_A(x)=d$.

证明：$\forall f(A)\in K[A]$，$f(x)=\mu_A(x)q(x)+r(x)$，其中 $r(x)=0$ 或 $\deg r(x)<\overset{d}{\overset{\|}{\deg\mu_A(x)}}$

$f(A)=r(A)\in\{\lambda_0I_n+\lambda_1A+\lambda_2A^2+\cdots+\lambda_{d-1}A^{d-1}\mid\lambda_i\in K\}=\langle I_n,A,A^2,\cdots,A^{d-1}\rangle$

故 $K[A]=\langle I_n,A,A^2,\cdots,A^{d-1}\rangle$. $I_n,A,A^2,\cdots,A^{d-1}$ 必线性无关，否则存在非零多项式 $f(x)=a_0+a_1x+\cdots+a_{d-1}x^{d-1}$ 使 $f(A)=0$. 与 $\mu_A(x)$ 选取矛盾 □

习题 4.1

1. 设 $f(x)\in K[x]$，$A\in M_n(K)$，若 $T\in M_n(K)$ 是可逆矩阵，证明：
$$f(TAT^{-1})=Tf(A)T^{-1}.$$

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 $A\in M_n(K)$，若 $E_{ii}A=AE_{ii}$ $(1\le i\le n)$ 成立，证明：A 必为对角矩阵 $\mathrm{diag}(a_{11},a_{22},\cdots,a_{nn}):=\begin{pmatrix}a_{11}&&0\\&a_{22}&\\0&&\ddots\ a_{nn}\end{pmatrix}_{n\times n}$.

4. 对任意 $A=(a_{ij})_{n\times n}\in M_n(K)$，$tr(A):=a_{11}+a_{22}+\cdots+a_{nn}\in K$ 称为 A 的迹 (trace). 试证明：(1) $tr(A\cdot B)=tr(BA)$，$\forall A,B\in M_n(K)$.

(2) $\forall$ 可逆矩阵 $T\in M_n(K)$，$tr(TAT^{-1})=tr(A)$.

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) $A=\begin{pmatrix}1&1&\cdots&1\\0&1&\cdots&1\\\cdots&\cdots&\cdots&\cdots\\0&0&\cdots&1\end{pmatrix}_{n\times n}$.

7. 证明：(1) 若 $\lambda\neq0$，A 可逆，则 $(\lambda A)^{-1}=\lambda^{-1}A^{-1}$.

(2) 若 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：
$$f(A)=0.$$

9. 矩阵 $A\in M_n(K)$ 称为幂零矩阵，如果存在 $m>0$ 使 $A^m=0$.
证明：若 A,B 是幂零矩阵，且 $A\cdot B=BA$，则 $A+B$ 也是幂零矩阵.

10. 如果 ${}^tA=A$，则称 A 是对称矩阵，若 ${}^tA=-A$，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。

11. 求矩阵
$$A=\begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & & & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \\ 0 & 0 & 0 & \cdots & 0 \end{pmatrix}_{n\times n}$$
的极小多项式。

12. 若 $A\in M_n(K)$ 是幂零矩阵，证明：I_n-A 可逆。

13. 若 $A\in M_n(K)$ 满足 $A^2=A$（称为幂等矩阵），证明：I_n+A 可逆

14. 设 $A\in M_{m\times n}(K)$，$B\in M_{m\times k}(K)$，考虑矩阵方程

$$(*)\qquad AX=B\ ,\quad X\in M_{n\times k}(K).$$

证明：$(*)$ 有解当且仅当 $r(A)=r(A|B)$，（$(A|B)$ 表示增广矩阵）

15. 设 $A,B\in M_n(K)$，若 I_n+AB 可逆，证明：I_n+BA 也可逆。

16. 求矩阵
$$A=\begin{pmatrix} x_1y_1 & x_1y_2 & \cdots & x_1y_n \\ x_2y_1 & x_2y_2 & \cdots & x_2y_n \\ \cdots & \cdots & \cdots & \cdots \\ x_ny_1 & x_ny_2 & \cdots & x_ny_n \end{pmatrix}_{n\times n}$$
的秩。

（提示：$A=[x_1,\cdots,x_n]\cdot(y_1,\cdots,y_n)$）。

17. 设 $A\in M_n(K)$，$f(x)=(x^3+2x^2+3x+1)(x^2-x+2)\in K[x]$。试求：$f(A)$。

18. 证明：$I_n+E_{ij}\ (i\neq j)$ 是可逆矩阵。

19. 证明：$A=\lambda I_n \Longleftrightarrow A\cdot B=BA$，$B$ 是任意可逆矩阵。

20. 求 $A\in M_n(K)$ 使得 $tr(A\cdot X)=0$ 对任意 $X\in M_n(K)$ 成立，证明你的结论。

21. 设 $A\in M_n(\mathbb{R})$ 是上三角型矩阵，若 $A\,{}^tA={}^tA\,A$，证明：A 是对角矩阵。

§4.2 行列式展开与应用.

在第三章定理3.4.3中，我们证明了矩阵行列式关于行向量的基本性质，根据 $|A| = |{}^tA|$ 不难看出所有对行向量成立的结论同样适用于列向量。本节将利用行列式基本性质将一个n阶行列式按行（或列）展开，从而将n阶行列式的研究转化为对n-1阶行列式的研究.

定义4.2.1. 对n阶矩阵 $A=(a_{ij})_{n\times n}\in M_n(K)$，去掉第i行与第j列，令 M_{ij} 表示所得n-1阶矩阵的行列式. 则称 M_{ij} 是A对应于元素 a_{ij} 的子式，记

$$A_{ij}=(-1)^{i+j}M_{ij}\quad\in K$$

称之为 a_{ij} 的代数余子式. □

引理4.2.1. 如果 $A=\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n}\\ 0 & a_{22} & \cdots & a_{2n}\\ \vdots & \vdots & & \\ 0 & a_{n2} & \cdots & a_{nn}\end{pmatrix}_{n\times n}$，则

$$|A|=a_{11}\begin{vmatrix} a_{22} & \cdots & a_{2n}\\ a_{32} & \cdots & a_{3n}\\ \vdots & & \\ a_{n2} & \cdots & a_{nn}\end{vmatrix}=a_{11}M_{11}=a_{11}A_{11}$$

证明：$\forall\ \pi\in S_n$，如果 $\pi(1)\neq 1$，则 $a_{\pi(1)1}=0$. 故

$$|A|=|{}^tA|=\sum_{\pi\in S_n}\varepsilon_\pi a_{\pi(1)1}\cdot a_{\pi(2)2}\cdots a_{\pi(n)n}=a_{11}\sum_{\substack{\pi\in S_n\\ \pi(1)=1}}\varepsilon_\pi a_{\pi(2)2}\cdots a_{\pi(n)n}.$$

$$=a_{11}M_{11}=a_{11}A_{11}$$

□.

定理4.2.1 [行列式展开]. 设 $A=(a_{ij})_{n\times n}\in M_n(K)$, 则

$$|A|=\sum_{i=1}^{n}(-1)^{i+j}a_{ij}M_{ij}=\sum_{i=1}^{n}a_{ij}A_{ij} \qquad (\text{按第}j\text{列展开})$$

$$|A|=\sum_{j=1}^{n}(-1)^{i+j}a_{ij}M_{ij}=\sum_{j=1}^{n}a_{ij}A_{ij} \qquad (\text{按第}i\text{行展开}).$$

证明: 仅证明按第j列展开. 对第j列应用基本性质(D2)(参见定理3.4.3):

$$|A|=\begin{vmatrix} a_{11} & \cdots & a_{1j} & \cdots & a_{1n}\\ a_{21} & \cdots & a_{2j} & \cdots & a_{2n}\\ \vdots & & \vdots & & \vdots\\ a_{n1} & \cdots & a_{nj} & \cdots & a_{nn}\end{vmatrix}=\sum_{i=1}^{n}\begin{vmatrix} a_{11} & \cdots & a_{1j-1} & 0 & a_{1j+1} & \cdots & a_{1n}\\ \vdots & & \vdots & \vdots & \vdots & & \vdots\\ a_{i1} & \cdots & a_{ij-1} & a_{ij} & a_{ij+1} & \cdots & a_{in}\\ \vdots & & \vdots & \vdots & \vdots & & \vdots\\ a_{n1} & \cdots & a_{nj-1} & 0 & a_{nj+1} & \cdots & a_{nn}\end{vmatrix}$$

此处利用了 $A^{(j)}=[a_{1j},a_{2j},\cdots,a_{nj}]=\sum_{i=1}^{n}[0,\cdots,a_{ij},0\cdots 0]$. 另外有:

$$\begin{vmatrix} a_{11} & \cdots & a_{1j-1} & 0 & a_{1j+1} & \cdots & a_{1n}\\ \vdots & & \vdots & \vdots & \vdots & & \vdots\\ a_{i1} & \cdots & a_{ij-1} & a_{ij} & a_{ij+1} & \cdots & a_{in}\\ \vdots & & \vdots & \vdots & \vdots & & \vdots\\ a_{n1} & \cdots & a_{nj-1} & 0 & a_{nj+1} & \cdots & a_{nn}\end{vmatrix}=(-1)^{j-1}\begin{vmatrix} 0 & a_{11} & \cdots & a_{1j-1} & a_{1j+1} & \cdots & a_{1n}\\ \vdots & \vdots & & \vdots & \vdots & & \vdots\\ a_{ij} & a_{i1} & \cdots & a_{ij-1} & a_{ij+1} & \cdots & a_{in}\\ \vdots & \vdots & & \vdots & \vdots & & \vdots\\ 0 & a_{n1} & \cdots & a_{nj-1} & a_{nj+1} & \cdots & a_{nn}\end{vmatrix}$$

$$=(-1)^{j-1}\cdot(-1)^{i-1}\begin{vmatrix} a_{ij} & a_{i1} & \cdots & a_{ij-1} & a_{ij+1} & \cdots & a_{in}\\ 0 & a_{11} & \cdots & a_{1j-1} & a_{1j+1} & \cdots & a_{1n}\\ 0 & a_{21} & \cdots & a_{2j-1} & a_{2j+1} & \cdots & a_{2n}\\ \vdots & \vdots & & \vdots & \vdots & & \vdots\\ 0 & a_{i-1\,1} & \cdots & a_{i-1\,j-1} & a_{i-1\,j+1} & \cdots & a_{i-1\,n}\\ 0 & a_{i+1\,1} & \cdots & a_{i+1\,j-1} & a_{i+1\,j+1} & \cdots & a_{i+1\,n}\\ \vdots & \vdots & & \vdots & \vdots & & \vdots\\ 0 & a_{n1} & \cdots & a_{nj-1} & a_{nj+1} & \cdots & a_{nn}\end{vmatrix}=(-1)^{i+j}a_{ij}M_{ij}=a_{ij}A_{ij}.$$

此处第一个等式应用了基本性质(D1)(将第j列依次与前j-1列交换位置). 第二个等式也应用了(D1)(将第i行依次与前i-1行交换位置), 第三个等式应用了引理4.2.1。因此得到|A|按第j列展开的公式

$$|A|=a_{1j}A_{1j}+a_{2j}A_{2j}+\cdots+a_{nj}A_{nj}=\sum_{i=1}^{n}a_{ij}A_{ij}.$$

同理可证: $|A|=a_{i1}A_{i1}+\cdots+a_{in}A_{in}$ (按第i行展开) □

例4.2.1. [范德蒙德(Vandermonde)行列式]. 设 $x_1,\cdots,x_n \in K$. 则

$$\Delta(x_1,\cdots,x_n)=\begin{vmatrix} 1 & 1 & \cdots & 1 \\ x_1 & x_2 & \cdots & x_n \\ x_1^2 & x_2^2 & \cdots & x_n^2 \\ \vdots & \vdots & & \vdots \\ x_1^{n-1} & x_2^{n-1} & \cdots & x_n^{n-1} \end{vmatrix} = \prod_{1\le i<j\le n}(x_j-x_i).$$

证明. [对n运用数学归纳法]: 当 $n=2$ 时, $\Delta(x_1,x_2)=\begin{vmatrix} 1 & 1 \\ x_1 & x_2 \end{vmatrix}=(x_2-x_1)$。
假设公式对 $n-1$ 成立, 则依次第$(i-1)$行乘$(-x_1)$加到第i行 (i从n开始) 可得

$$\Delta(x_1,\cdots,x_n)=\begin{vmatrix} 1 & 1 & 1 & \cdots & 1 \\ 0 & x_2-x_1 & x_3-x_1 & & x_n-x_1 \\ 0 & x_2^2-x_2x_1 & x_3^2-x_3x_1 & \cdots, & x_n^2-x_nx_1 \\ \vdots & & & & \\ 0 & x_2^i-x_2^{i-1}x_1 & x_3^i-x_3^{i-1}x_1 & \cdots & x_n^i-x_n^{i-1}x_1 \\ \vdots & \vdots & & & \\ 0 & x_2^{n-2}-x_2^{n-3}x_1 & x_3^{n-2}-x_3^{n-3}x_1 & \cdots & x_n^{n-2}-x_n^{n-3}x_1 \\ 0 & x_2^{n-1}-x_2^{n-2}x_1 & x_3^{n-1}-x_3^{n-2}x_1 & \cdots & x_n^{n-1}-x_n^{n-2}x_1 \end{vmatrix}$$

按第1列展开, 并从 M_{11} 的各列提出因子 x_2-x_1, $x_3-x_1,\cdots,x_n-x_1$ 可得

$$\Delta(x_1,\cdots,x_n)=(x_2-x_1)\cdots(x_n-x_1)\begin{vmatrix} 1 & 1 & \cdots & 1 \\ x_2 & x_3 & \cdots & x_n \\ x_2^2 & x_3^2 & \cdots & x_n^2 \\ \vdots & \vdots & & \vdots \\ x_2^{n-2} & x_3^{n-2} & \cdots & x_n^{n-2} \end{vmatrix}=(x_2-x_1)\cdots(x_n-x_1)\Delta(x_2,\cdots,x_n)$$

由归纳假设可得结论。

□

利用行列式展开定理, 可以得到关于逆矩阵的有趣公式.

定义4.2.2 [伴随矩阵] 设 $A=(a_{ij})_{n\times n}\in M_n(K)$. 则矩阵

$$A^*=\begin{pmatrix} A_{11} & A_{21} & \cdots & A_{n1} \\ A_{12} & A_{22} & \cdots & A_{n2} \\ \vdots & & & \\ A_{1n} & A_{2n} & \cdots & A_{nn} \end{pmatrix}\in M_n(K)$$

称为A的伴随矩阵.

□

引理4.2.2. $AA^* = A^*A = |A|\cdot I_n = \begin{pmatrix} |A| & 0 & \cdots & 0 \\ 0 & |A| & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & |A| \end{pmatrix}_{n\times n}$

证明：$A\cdot A^* = |A|\cdot I_n$ 等价于，$\forall 1\le i,j\le n$，有

$$(a_{i1}, a_{i2}, \cdots, a_{in})\begin{pmatrix} A_{j1} \\ A_{j2} \\ \vdots \\ A_{jn} \end{pmatrix} = \begin{cases} 0 & i\ne j \\ |A| & i=j. \end{cases}$$

当 $i=j$ 时，上式就是定理4.2.1中的按第 i 行展开。当 $i\ne j$ 时，将 A 的第 j 行 $A_{(j)}$ 用第 i 行 $A_{(i)}$ 代替，得：

$$A' = [A_{(1)}, \cdots, A_{(i)}, \cdots, A_{(i)}, \cdots, A_{(n)}] = (a'_{ij})_{n\times n}$$

第 j 行位置.

则 $|A'|=0$. 所以将 $|A'|$ 按第 j 行展开，得：

$$0 = |A'| = a'_{j1}A'_{j1} + a'_{j2}A'_{j2} + \cdots + a'_{jn}\cdot A'_{jn},$$

但 $a'_{jk} = a_{ik}$，$A'_{jk} = A_{jk}$，故 $a_{i1}A_{j1} + \cdots + a_{in}A_{jn} = 0$ $(i\ne j)$.

同理可证，$A^*A = |A|\cdot I_n$. □

定理4.2.2 [逆矩阵公式]. 设 $A = (a_{ij})_{n\times n} \in M_n(K)$. ($K$是一个域)

则 A 可逆 $\Longleftrightarrow$ $|A|$ 在 K 中可逆 (即 $|A|\ne 0$).

此时，$A^{-1} = |A|^{-1}A^* = \frac{1}{|A|}A^*$.

证明：如果 A 可逆，则存在 $B\in M_n(K)$ 使 $A\cdot B = B\cdot A = I_n$. 从而 $|A|\cdot|B| = |B|\cdot|A| = 1$，故 $|A|$ 在 K 中可逆 (即 $|A|\ne 0$). 反之，若 $|A|\in K$ 可逆，则 $A^{-1} = |A|^{-1}\cdot A^*$ 是 A 的逆矩阵 □

定理4.2.3 [克莱姆 (Cramer) 法则]. 线性方程组

$$x_1A^{(1)} + x_2A^{(2)} + \cdots + x_nA^{(n)} = b, \quad b = [b_1, \cdots, b_n]$$

有唯一解 $\Longleftrightarrow$ $\det(A^{(1)}, \cdots, A^{(n)}) = |A| \ne 0$，且它的解为

$$x_1 = \frac{D_1}{|A|},\ x_2 = \frac{D_2}{|A|},\ \cdots,\ x_n = \frac{D_n}{|A|},\quad D_k := \det(A^{(1)}, \cdots, A^{(k-1)}, b, A^{(k+1)}, \cdots, A^{(n)}).$$

证明: 方程组 $AX=b$ 有唯一解 $\Leftrightarrow r(A)=n \Leftrightarrow A$ 可逆 $\Leftrightarrow |A|\neq 0$

且 $X=A^{-1}\cdot b=\frac{1}{|A|}A^{*}\cdot b$，即 $x_k=\frac{1}{|A|}(b_1A_{1k}+\cdots+b_nA_{nk})$.

但 $b_1A_{1k}+b_2A_{2k}+\cdots+b_nA_{nk}=\det(A^{(1)},\cdots,A^{(k-1)},b,A^{(k+1)},\cdots,A^{(n)})$. 因此

$$x_1=\frac{D_1}{|A|},\ x_2=\frac{D_2}{|A|},\ \cdots,\ x_n=\frac{D_n}{|A|}.$$

□

设 $A=(a_{ij})_{m\times n}\in M_{m\times n}(K)$, A 的秩可用 A 的子矩阵的行列式来描述。设 $1\le p\le \min\{m,n\}$, 对任意 $1\le i_1<i_2<\cdots<i_p\le m$, $1\le j_1<j_2<\cdots<j_p\le n$.

定义4.2.3. [p阶子式]. 矩阵 A 的任意 p 行 $A_{(i_1)},\cdots,A_{(i_p)}$ 与任意 p 列 $A^{(j_1)},\cdots,A^{(j_p)}$ 交叉处的元素组成一个 p 阶方阵, 其行列式

$$M\begin{pmatrix} i_1 & i_2 & \cdots & i_p \\ j_1 & j_2 & \cdots & j_p \end{pmatrix}=\begin{vmatrix} a_{i_1j_1} & a_{i_1j_2} & \cdots & a_{i_1j_p} \\ a_{i_2j_1} & a_{i_2j_2} & \cdots & a_{i_2j_p} \\ \vdots & & & \\ a_{i_pj_1} & a_{i_pj_2} & \cdots & a_{i_pj_p} \end{vmatrix}$$

称为 A 的一个 p-阶子式.

□

定理4.2.4 [秩的行列式刻画]. 设 $A=(a_{ij})_{m\times n}\in M_{m\times n}(K)$. 则

$r(A)=r \Leftrightarrow$ A 有一个非零的 r 阶子式, 且 A 的任意 r+1 阶子式为零.

证明: 若 $r(A)=r$, 则存在 $A_{(i_1)},A_{(i_2)},\cdots,A_{(i_r)}$ 线性无关, 从而 $B=[A_{(i_1)},A_{(i_2)},\cdots,A_{(i_r)}]$ 的行秩等于 r. 因此, 存在 $B^{(j_1)},B^{(j_2)},\cdots,B^{(j_r)}$ 线性无关, 即 r 阶方阵 $(B^{(j_1)},B^{(j_2)},\cdots,B^{(j_r)})$ 的秩为 r, 故

$$M\begin{pmatrix} i_1 & \cdots & i_r \\ j_1 & \cdots & j_r \end{pmatrix}=\begin{vmatrix} a_{i_1j_1} & a_{i_1j_2} & \cdots & a_{i_1j_r} \\ a_{i_2j_1} & a_{i_2j_2} & \cdots & a_{i_2j_r} \\ \vdots & \vdots & & \vdots \\ a_{i_rj_1} & a_{i_rj_2} & \cdots & a_{i_rj_r} \end{vmatrix}\neq 0.$$

下面只需证明: 若 A 有一个非零的 p 阶子式, 则 $r(A)\ge p$.

无妨设 $M\begin{pmatrix} 1 & 2 & \cdots & p \\ 1 & 2 & \cdots & p \end{pmatrix}\neq 0$, 则 $A_{(1)},\cdots,A_{(p)}$ 线性无关. 否则存在 $A_{(k)}$ 可由 $A_{(1)},\cdots,A_{(k-1)},A_{(k+1)},\cdots,A_{(p)}$ 线性表出, 从而 $(a_{k1},a_{k2},\cdots,a_{kp})$

可由 $\{(a_{i1}, a_{i2}, \cdots, a_{ip}) \mid i \neq k\}_{1\le i\le p}$ 线性表出，故 $M\begin{pmatrix}1&2&\cdots&p\\1&2&\cdots&p\end{pmatrix}=0$.
此与 $M\begin{pmatrix}1&2&\cdots&p\\1&2&\cdots&p\end{pmatrix}\neq 0$ 矛盾。因此 $r(A)\ge p$. □

下面的 Laplace 定理是行列式按行(或列)展开的推广，它是将行列式按给定的若干行(或列)的展开。

定义 4.2.4 [(n-p)阶余子式]. 设 $A=(a_{ij})_{n\times n}$，$1\le i_1<i_2<\cdots<i_p\le n$，$1\le j_1<j_2<\cdots<j_p\le n$. 去掉 A 的第 $i_1, i_2, \cdots, i_p$ 行及第 $j_1, j_2, \cdots, j_p$ 列得一个 $(n-p)$ 阶子阵，它的行列式 $\overline{M}\begin{pmatrix}i_1&\cdots&i_p\\j_1&\cdots&j_p\end{pmatrix}$ 称为 $M\begin{pmatrix}i_1&\cdots&i_p\\j_1&\cdots&j_p\end{pmatrix}$ 的 $(n-p)$ 阶余子式. 而

$$A\begin{pmatrix}i_1&\cdots&i_p\\j_1&\cdots&j_p\end{pmatrix}:=(-1)^{i_1+\cdots+i_p+j_1+j_2+\cdots+j_p}\,\overline{M}\begin{pmatrix}i_1&\cdots&i_p\\j_1&\cdots&j_p\end{pmatrix}.$$

称为 $M\begin{pmatrix}i_1&\cdots&i_p\\j_1&\cdots&j_p\end{pmatrix}$ 的 $(n-p)$ 阶代数余子式. □

定理 4.2.5 [Laplace 定理]. 对任意 $1\le i_1<\cdots<i_p\le n$，有

$$|A|=\sum_{1\le j_1<\cdots<j_p\le n} M\begin{pmatrix}i_1&i_2&\cdots&i_p\\j_1&j_2&\cdots&j_p\end{pmatrix}A\begin{pmatrix}i_1&i_2&\cdots&i_p\\j_1&j_2&\cdots&j_p\end{pmatrix}.$$

(称为按第 $i_1, i_2, \cdots, i_p$ 行展开)，也可按第 $j_1, j_2, \cdots, j_p$ 列展开.

$$|A|=\sum_{1\le i_1<\cdots<i_p\le n} M\begin{pmatrix}i_1&i_2&\cdots&i_p\\j_1&j_2&\cdots&j_p\end{pmatrix}A\begin{pmatrix}i_1&\cdots&i_p\\j_1&\cdots&j_p\end{pmatrix}.$$

□

定理 4.2.6 [Cauchy-Binet 公式]. 设 $A=(a_{ij})_{n\times m}$，$B=(b_{ij})_{m\times n}$. 则 (1) 当 $m<n$ 时，有 $|A\cdot B|=0$；(2) 当 $m\ge n$ 时有

$$|A\cdot B|=\sum_{1\le j_1<\cdots<j_n\le m}\begin{vmatrix}a_{1j_1}&a_{1j_2}&\cdots&a_{1j_n}\\a_{2j_1}&a_{2j_2}&\cdots&a_{2j_n}\\\vdots&\vdots&&\vdots\\a_{nj_1}&a_{nj_2}&\cdots&a_{nj_n}\end{vmatrix}\cdot\begin{vmatrix}b_{j_11}&b_{j_12}&\cdots&b_{j_1n}\\b_{j_21}&b_{j_22}&\cdots&b_{j_2n}\\\vdots&\vdots&&\vdots\\b_{j_n1}&b_{j_n2}&\cdots&b_{j_nn}\end{vmatrix}.$$

证明：(1) $A\cdot B\in M_n(K)$，但 $r(A\cdot B)\le m<n$，故 $|A\cdot B|=0$.

(2) 的证明省略. □

习题 4.2

1. 请按要求计算下列行列式.

(a) $\begin{vmatrix} 2 & -3 & 4 & 1 \\ 4 & -2 & 3 & 2 \\ a & b & c & d \\ 3 & -1 & 4 & 3 \end{vmatrix}$ (按第三行展开).
(b) $\begin{vmatrix} 5 & a & 2 & -1 \\ 4 & b & 4 & -3 \\ 2 & c & 3 & -2 \\ 4 & d & 5 & -4 \end{vmatrix}$ (按第二列展开).

2. 计算下列行列式

(a) $\begin{vmatrix} x & y & 0 & \cdots & 0 & 0 \\ 0 & x & y & \cdots & 0 & 0 \\ \cdots & & & \cdots & & \cdots \\ 0 & 0 & 0 & \cdots & x & y \\ y & 0 & 0 & \cdots & 0 & x \end{vmatrix}$;
(b) $\begin{vmatrix} a_0 & 1 & 1 & 1 & \cdots & 1 \\ 1 & a_1 & 0 & 0 & \cdots & 0 \\ 1 & 0 & a_2 & 0 & \cdots & 0 \\ \cdots & & \cdots & & & \cdots \\ 1 & 0 & 0 & 0 & \cdots & a_n \end{vmatrix}$.

(c) $\begin{vmatrix} \cos(\alpha_1-\beta_1), & \cos(\alpha_1-\beta_2), & \cdots, & \cos(\alpha_1-\beta_n) \\ \cos(\alpha_2-\beta_1), & \cos(\alpha_2-\beta_2), & \cdots & \cos(\alpha_2-\beta_n) \\ \vdots & \vdots & \ddots & \vdots \\ \cos(\alpha_n-\beta_1), & \cos(\alpha_n-\beta_2) & \cdots & \cos(\alpha_n-\beta_n) \end{vmatrix}$ (提示: $\cos(\alpha_i-\beta_j) = (\cos\alpha_i, \sin\alpha_i)\begin{pmatrix}\cos\beta_j \\ \sin\beta_j\end{pmatrix}$.)

3. 利用 Cauchy-Binet 公式证明: 当 $n \geq 2$ 时, 有

$$\left(\sum_{i=1}^n a_i^2\right)\left(\sum_{i=1}^n b_i^2\right) - \left(\sum_{i=1}^n a_i b_i\right)^2 = \sum_{1\leq i<j\leq n} (a_i b_j - a_j b_i)^2.$$

(提示: 左边 $= \begin{vmatrix} \sum_{i=1}^n a_i^2 & \sum_{i=1}^n a_i b_i \\ \sum_{i=1}^n a_i b_i & \sum_{i=1}^n b_i^2 \end{vmatrix}$).

4. 试证明: (1) $|A^*| = |A|^{n-1}$; (2) $(A^*)^* = \begin{cases} |A|^{n-2}A, & 当 n>2 \\ A, & 当 n=2 \end{cases}$

(3) 如果 A^* 可逆, 则 A 也可逆. 且 $A^{-1} = \frac{1}{|A|}\cdot A^*$.

5. 设 $A \in M_{m\times n}(K)$, $m \geq n$, 证明: $|{}^tA \cdot A| = \sum_M M^2$, 其中 M 跑遍 A 的全部 n 阶子式.

6. 设 $A = (a_{ij})_{n\times n}$, 且 $A^* = {}^tA$. 证明: $|A| = \sum_{j=1}^n a_{1j}^2 = \sum_{j=1}^n a_{2j}^2 = \cdots = \sum_{j=1}^n a_{nj}^2$.

7. 设n阶方阵
$$A=\begin{pmatrix}1&1&1&\cdots&1\\0&1&1&\cdots&1\\0&0&1&\cdots&1\\\cdots&\cdots&\cdots&\cdots&\cdots\\0&0&0&\cdots&1\end{pmatrix}.$$

试求 $\sum_{i=1}^{n}\sum_{j=1}^{n}A_{ij}$.

8. 设 $A=(a_{ij})_{m\times n}\in M_{m\times n}(K)$, $(m<n)$. 已知 $AX=0$ 的一个基础解系为 $\beta_i=[b_{i1},b_{i2},\cdots,b_{in}]$ $(i=1,2,\cdots,n-m)$. 试求：线性方程组
$$\sum_{j=1}^{n}b_{ij}y_j=0\qquad(i=1,2,\cdots,n-m)$$
的一个基础解系（或解空间的一组基）.

9. 设 $\Delta(x_1,\cdots,x_n)$, $\Delta(y_1,\cdots,y_n)$ 表示范德蒙德行列式，且对任意 i,j 有 $x_iy_j\neq1$. 试证明：
$$\Delta(x_1,\cdots,x_n)\,\Delta(y_1,\cdots,y_n)=\det\left(\frac{1}{1-x_iy_j}\right)_{n\times n}\cdot\prod_{i,j}(1-x_iy_j).$$

10. 证明欧拉恒等式：
$$(x_1^2+x_2^2+x_3^2+x_4^2)(y_1^2+y_2^2+y_3^2+y_4^2)=(x_1y_1+x_2y_2+x_3y_3+x_4y_4)^2+$$
$$(x_1y_1-x_2y_1-x_3x_4+x_4y_3)^2+(x_1y_3+x_2y_4-x_3y_1-x_4y_2)^2+(x_1y_4-x_2y_3+x_3y_2-x_4y_1)^2.$$

提示：计算矩阵乘积：
$$\begin{pmatrix}x_1&x_2&x_3&x_4\\x_2&-x_1&-x_4&x_3\\x_3&x_4&-x_1&-x_2\\x_4&-x_3&x_2&-x_1\end{pmatrix}\cdot\begin{pmatrix}y_1&y_2&y_3&y_4\\y_2&-y_1&-y_4&y_3\\y_3&y_4&-y_1&-y_2\\y_4&-y_3&y_2&-y_1\end{pmatrix}.$$

§4.3. 矩阵的等价关系.

在矩阵空间 $M_{m\times n}(K)$ 中，两个不同的矩阵 $A=(a_{ij})_{m\times n}$ 和 $B=(b_{ij})_{m\times n}$ 可能是同一个线性映射 $V\xrightarrow{\varphi}W$ 在不同基下的坐标矩阵。在方阵空间 $M_n(K)$ 中，不同的方阵 A 与 B 也可能是同一个双线性函数 $f: V\times V\to K$ 在不同基下的坐标矩阵。所以我们有必要研究这些矩阵之间的关系。

引理4.3.1：设 $v_1,\cdots,v_n\in V$ 是一组基. 则

(1) $\forall$ 可逆矩阵 $T\in M_n(K)$，$(v_1,\cdots,v_n)T$ 是 V 的一组基.

(2) 如果 $(\alpha_1,\cdots,\alpha_n)$ 是 V 的一组基，则存在可逆矩阵 T 使得 $(\alpha_1,\cdots,\alpha_n)=(v_1,\cdots,v_n)T$.

证明：(1) 令 $(\alpha_1,\cdots,\alpha_n)=(v_1,\cdots,v_n)T$，则 $(v_1,\cdots,v_n)=(\alpha_1,\cdots,\alpha_n)T^{-1}$. 即 $v_1,\cdots,v_n$ 可由 $\alpha_1,\cdots,\alpha_n$ 线性表出，故 $V=\langle\alpha_1,\cdots,\alpha_n\rangle$. 只需证明 $\alpha_1,\cdots,\alpha_n$ 线性无关：若存在 $[\lambda_1,\cdots,\lambda_n]\neq 0$ 使 $(\alpha_1,\cdots,\alpha_n)\cdot[\lambda_1,\cdots,\lambda_n]=0$. 可得 $(v_1,\cdots,v_n)T\cdot[\lambda_1,\cdots,\lambda_n]=0$. 但 $v_1,\cdots,v_n$ 线性无关，故 $T\cdot[\lambda_1,\cdots,\lambda_n]=0$ 从而 $[\lambda_1,\cdots,\lambda_n]=T^{-1}\cdot T[\lambda_1,\cdots,\lambda_n]=0$. 矛盾.

(2) 设 $\alpha_1,\cdots,\alpha_n\in V$ 是一组基。由于 $v_1,\cdots,v_n$ 也是一组基，故 $\alpha_1,\cdots,\alpha_n$ 可由 $v_1,\cdots,v_n$ 线性表出. 即存在矩阵 $T\in M_n(K)$ 使 $(\alpha_1,\cdots,\alpha_n)=(v_1,\cdots,v_n)T$. 同理，$v_1,\cdots,v_n$ 也可由 $\alpha_1,\cdots,\alpha_n$ 线性表出，即存在 $T'\in M_n(K)$ 使

$$(v_1,\cdots,v_n)=(\alpha_1,\cdots,\alpha_n)T'.$$

从而 $(\alpha_1,\cdots,\alpha_n)=(v_1,\cdots,v_n)T=(\alpha_1,\cdots,\alpha_n)T'\cdot T$. 但 $\alpha_1,\cdots,\alpha_n$ 线性无关，故 $T'\cdot T=I_n$，从而 T 是可逆矩阵（定理4.4.1）. □

设 $A,B\in M_{m\times n}(V)$ 分别是线性映射 $V\xrightarrow{\varphi}W$ 在不同基 $v_1,\cdots,v_n\in V$, $w_1,\cdots,w_m\in W$ 和 $v_1',\cdots,v_n'\in V$, $w_1',\cdots,w_m'$ 下的坐标矩阵.

$(\varphi(v_1),\cdots,\varphi(v_n))=(w_1,\cdots,w_m)\cdot A$, $(\varphi(v_1'),\cdots,\varphi(v_n'))=(w_1',\cdots,w_m')\cdot B$.

令 $(v_1',\cdots,v_n')=(v_1,\cdots,v_n)T_1$, $(w_1',\cdots,w_m')=(w_1,\cdots,w_m)T_2$. $(T_1\in M_n(K), T_2\in M_m(K))$.

则 $(\varphi(v_1'),\cdots,\varphi(v_n'))=(\varphi(v_1),\cdots,\varphi(v_n))T_1=(w_1,\cdots,w_m)AT_1$. 另一方面,

$$(\varphi(v_1'),\cdots,\varphi(v_n'))=(w_1',\cdots,w_m')\cdot B=(w_1,\cdots,w_m)T_2B.$$

故 $AT_1=T_2B$, 即 $T_2^{-1}A\cdot T_1=B$ (或 $A=T_2BT_1^{-1}$).

<u>定理4.3.1</u>: 设 $A,B\in M_{m\times n}(K)$, 则 A,B 是<u>同一线性映射</u>在不同基下的坐标矩阵的充要条件是: 存在可逆矩阵 $P\in M_m(K)$, $Q\in M_n(K)$ 使 $B=PA\cdot Q$.

证明: 如果 A,B 是同一线性映射 $V\xrightarrow{\varphi}W$ 在不同基下的坐标矩阵, 则上述计算表明存在可逆矩阵 $P=T_2^{-1}$, $Q=T_1$ 使 $PAQ=B$.

反之, 若存在可逆矩阵 $P\in M_m(K)$, $Q\in M_n(K)$ 使 $PAQ=B$, 令 $V\xrightarrow{\varphi_A}W$ 是由 $(\varphi_A(v_1),\cdots,\varphi_A(v_n))=(w_1,\cdots,w_m)A$ 定义的线性映射(其中 $v_1,\cdots,v_n\in V$, $w_1,\cdots,w_m\in W$ 是任意选定的基). 则 φ_A 在基

$$(v_1',\cdots,v_n'):=(v_1,\cdots,v_n)\cdot Q,\quad (w_1',\cdots,w_m'):=(w_1,\cdots,w_m)\cdot P^{-1}$$

下的坐标矩阵是 $PAQ=B$. □

由于可逆矩阵可写成初等矩阵之积, $PAQ=B$ 表示 B 可以由 A 经<u>初等行变换</u>和<u>初等列变换</u>得到. 所以, 我们引入,

<u>定义4.3.1</u>: $A,B\in M_{m\times n}(K)$ 称为初等等价(或简称等价). 如果存在可逆矩阵 $P\in M_m(K)$, $Q\in M_n(K)$ 使

$$PAQ=B.$$ □

<u>推论4.3.1</u>: (1) 若 $r(A)=r$, 则 A (初等)等价于

$$\begin{pmatrix} I_r & 0\\ 0 & 0\end{pmatrix}.\quad I_r\text{是}r\text{阶单位阵}.$$

(2) A 与 ~~[illegible]~~ B 等价 $\Longleftrightarrow r(A)=r(B)$

$\square$

当 $V=W$ 时, 线性映射 $V\xrightarrow{\varphi}V$ 称为 V 上的线性算子(或称 V 上的线性变换), 它在基 $v_1,\cdots,v_n$ 下的坐标矩阵 A_φ 定义为

$$(\varphi(v_1),\cdots,\varphi(v_n))=(v_1,\cdots,v_n)A_\varphi,\ \text{即: } A_\varphi=(a_{ij})_{n\times n},\ a_{ij}=v_i^*(\varphi(v_j))\in K.$$

$M_n(K)$ 中每一个矩阵 A 均可实现为 V 上某个线性算子在基 $v_1,\cdots,v_n$ 下的坐标矩阵.

<u>引理4.3.2</u>: 设 $A,B\in M_n(K)$. 则 A,B 是同一线性算子 $V\xrightarrow{\varphi}V$ 在不同基下的坐标矩阵的充要条件是: 存在可逆矩阵 $T\in M_n(K)$ 使 $T^{-1}AT=B$.

证明: 若存在 $V\xrightarrow{\varphi}V$ 及基 $v_1,\cdots,v_n$ 和 $v_1',\cdots,v_n'$ 使

$$(\varphi(v_1),\cdots,\varphi(v_n))=(v_1,\cdots,v_n)A,\quad(\varphi(v_1'),\cdots,\varphi(v_n'))=(v_1',\cdots,v_n')B.$$

令 $(v_1',\cdots,v_n')=(v_1,\cdots,v_n)\cdot T$, 则 $(\varphi(v_1'),\cdots,\varphi(v_n'))=(\varphi(v_1),\cdots,\varphi(v_n))T$.
从而 $(v_1,\cdots,v_n)T\cdot B=(v_1,\cdots,v_n)A\cdot T$. 由于 $v_1,\cdots,v_n$ 线性无关, 故 $T\cdot B=A\cdot T$. 即 $B=T^{-1}AT$ (或 $A=TBT^{-1}$).

反之, 设 $V\xrightarrow{\varphi_A}V$ 由 $(\varphi_A(v_1),\cdots,\varphi_A(v_n))=(v_1,\cdots,v_n)A$ 定义, ($v_1,\cdots,v_n$ 是 V 的任一组基), 令 $(v_1',\cdots,v_n')=(v_1,\cdots,v_n)\cdot T$, 则 φ_A 在基 $v_1',\cdots,v_n'$ 下的坐标矩阵是 $T^{-1}AT=B$.

$\square$

<u>定义4.3.2</u>: 设 $A,B\in M_n(K)$. 如果存在可逆矩阵 $T\in M_n(K)$ 使 $T^{-1}A\cancel{B}T=B$ (或 $A=TBT^{-1}$), 则称 A 与 B 相似等价.

$\square$

一个自然的问题是: 对于给定的 $A\in M_n(K)$, 什么是 <u>与 A 相似等价</u> 的"最简单"矩阵 $J(A)$? 即存在可逆矩阵 $T\in M_n(K)$ 使 $T^{-1}AT=J(A)$. 这在理论上和应用中都是相当重要

的问题。理论上它回答了："线性算子最简单的表达式是什么？"另一方面，它也可以简化计算：$A^m = T\,J(A)^m\,T^{-1}$。例如，如果 $J(A)$ 是对角矩阵，则只需知道 T，人们就能计算出 $f(A)$. ($\forall f(x)\in K[x]$)，这在实際问题中有重要意义。在下一章中，我们将回答下列问题：(1) 对于给定的 $A\in M_n(K)$，明确在相似等价意义下"最简单"的含义。(2) "最简单矩阵 $J(A)$" 由 A 的哪些不变量唯一确定？(3) 如何求 T 使 $A = T\,J(A)\,T^{-1}$？

$M_n(K)$ 中的任意矩阵 $A=(a_{ij})_{n\times n}$ 均可实现为 n 维 K-向量空间 V 上的双线性函数 $f_A: V\times V\to K$ 在给定基 $v_1,\cdots,v_n$ 下的坐标矩阵：$f_A(X,Y) := (v_1^*(X),\cdots,v_n^*(X))\cdot A\cdot [v_1^*(Y),\cdots,v_n^*(Y)]$. ($\forall X,Y\in V$)。显然，$f_A$ 是对称双线性函数当且仅当 A 是对称矩阵，f_A 是反对称双线性函数当且仅当 A 是反对称矩阵。

<u>引理4.3.3</u>：设 $A, B\in M_n(K)$，则 A, B 是<u>同一双线性函数</u> $f: V\times V\to K$ 在 V 的不同基下的坐标矩阵的充要条件是：存在<u>可逆矩阵</u> $T\in M_n(K)$ 使 ${}^tTAT = B$.

证明：若 $A, B\in M_n(K)$ 是双线性函数 $f: V\times V\to K$ 分别在基 $v_1,\cdots,v_n\in V$ 和基 $\alpha_1,\cdots,\alpha_n\in V$ 下的<u>坐标矩阵</u>，则，$\forall x,y\in V$，

$$f(x,y) = (v_1^*(x),\cdots,v_n^*(x))A\begin{pmatrix}v_1^*(y)\\ \vdots\\ v_n^*(y)\end{pmatrix} = (\alpha_1^*(x),\cdots,\alpha_n^*(x))\,B\begin{pmatrix}\alpha_1^*(y)\\ \vdots\\ \alpha_n^*(y)\end{pmatrix}.$$

令 $(\alpha_1,\cdots,\alpha_n) = (v_1,\cdots,v_n)\cdot T$，由定理3.3.2知，$\forall x\in V$，有

$$T\cdot[\alpha_1^*(x),\cdots,\alpha_n^*(x)] = [v_1^*(x),\cdots,v_n^*(x)].$$

故，$\forall x,y\in V$，有 ~~$(v_1^*(x),\cdots,v_n^*(x))A[v_1^*(y),\cdots,v_n^*(y)] = (v_1^*(x),\cdots,v_n^*(x))BT$~~

$$(v_1^*(x),\cdots,v_n^*(x))A\begin{pmatrix}v_1^*(y)\\ \vdots\\ v_n^*(y)\end{pmatrix} = (\alpha_1^*(x),\cdots,\alpha_n^*(x))({}^tTAT)\begin{pmatrix}\alpha_1^*(y)\\ \vdots\\ \alpha_n^*(y)\end{pmatrix}.$$

所以 $B = {}^tTAT$.

反之，若 ${}^tTAT=B$，令 $f_A(x,y)=(v_1^*(x),\cdots,v_n^*(x))A[v_1^*(y),\cdots,v_n^*(y)]$.
则 $f_A: V\times V\to K$ 是双线性函数且 $f_A(v_i,v_j)=a_{ij}$. 不难验证，f_A 在基 $(\alpha_1,\cdots,\alpha_n):=(v_1,\cdots,v_n)\cdot T$ 下的坐标矩阵是 B. □

定义4.3.3：设 $A,B\in M_n(K)$，若存在可逆矩阵 $T\in M_n(K)$ 使 $A={}^tTBT$，则称 A 与 B 合同等价. □

定理4.3.2：设 $\mathrm{Char}(K)\neq 2$，$A=(a_{ij})_{n\times n}\in M_n(K)$. 则

(1) A 合同等价于一个对角矩阵 $\Longleftrightarrow$ A 是对称矩阵.

(2) A 是反对称矩阵 $\Longleftrightarrow$ A 合同等价于

$$\begin{pmatrix} 0 & 1 & & & & & & & \\ -1 & 0 & & & & & & & \\ & & 0 & 1 & & & & & \\ & & -1 & 0 & & & & & \\ & & & & \ddots & & & & \\ & & & & & 0 & 1 & & \\ & & & & & -1 & 0 & & \\ & & & & & & & 0 & \\ & & & & & & & & \ddots \\ & & & & & & & & & 0 \end{pmatrix}_{n\times n}.$$

(3) 如果 $A\in M_n(\mathbb{R})$ 是实对称矩阵，则存在可逆矩阵 $T\in M_n(\mathbb{R})$ 使

$${}^tTAT=\begin{pmatrix} \overbrace{1 \ \cdots \ 1}^{s} & & & \\ & \overbrace{-1 \ \cdots \ -1}^{r-s} & & \\ & & 0 & \\ & & & \ddots \ 0 \end{pmatrix}_{n\times n}$$

其中 s 不依赖于 T 的选取，是 A 的合同不变量，称为 A 的正惯性指数.

证明：这是推论3.4.7的重新叙述，证明已在第三章的第四节给出. □

定义4.3.4. 设 $A\in M_n(\mathbb{R})$ 是实对称矩阵，s 是 A 的正惯性指数，r 是 A 的秩，$2s-r$ 称为 A 的符号差. 则

(1) 如果 $r=s$，则称 A 是半正定矩阵 (semi-positive)

(2) 如果 $r=s=n$，则称 A 是正定矩阵 (positive).

(3) 如果 $s=0$，则称 A 是半负定矩阵.

(4) 如果 $s=0$，$r=n$，则称 A 是负定矩阵

(5) 如果 $0<s<r$，则称 A 是不定的.

一个二次型 $q(x_1,\cdots,x_n)=(x_1,\cdots,x_n)A\,{}^t[x_1,\cdots,x_n]$ 分别称为半正定，正定，半负定，负定，不定 如果它的坐标矩阵分别是半正定，正定，半负定，负定，不定的矩阵.

□

推论4.3.2: 设 $A\in M_n(\mathbb{R})$ 是对称实矩阵. 则

(1) A 是半正定矩阵 $\Leftrightarrow$ $\forall\ X=[x_1,\cdots,x_n]\in\mathbb{R}^n$，${}^tXAX\geq 0$.

(2) A 是正定矩阵 $\Leftrightarrow$ $\forall\ X=[x_1,\cdots,x_n]\neq 0$，有 ${}^tXAX>0 \Leftrightarrow A={}^tT\cdot T$ (T可逆).

(3) A 是半负定矩阵 $\Leftrightarrow$ $\forall\ X=[x_1,\cdots,x_n]$，有 ${}^tXAX\leq 0$

(4) A 是负定矩阵 $\Leftrightarrow$ $\forall\ X=[x_1,\cdots,x_n]\neq 0$，有 ${}^tXAX<0$

(5) A 是不定的 $\Leftrightarrow$ 存在 $X=[x_1,\cdots,x_n]$，$Y=[y_1,\cdots,y_n]$ 使 ${}^tXAX>0$，${}^tYAY<0$.

证明：由定理4.3.2 (3)，存在可逆矩阵 $T\in M_n(\mathbb{R})$ 使

$$A={}^tT\begin{pmatrix}1_s & 0 & 0\\ 0 & -1_{r-s} & 0\\ 0 & 0 & 0\end{pmatrix}_{n\times n}T={}^tT\cdot D(A)\cdot T.$$

~~若A半正定，则~~ $\forall\ X=[x_1,\cdots,x_n]$，令 $[t_1,\cdots,t_n]=T\cdot[x_1,\cdots,x_n]$，

则 ${}^tXAX={}^t(T\cdot X)D(A)(TX)=t_1^2+\cdots+t_s^2-t_{s+1}^2-\cdots-t_r^2$. 故

(1) A 半正定 $\Leftrightarrow r=s \Leftrightarrow {}^tXAX\geq 0$，$\forall\ X=[x_1,\cdots,x_n]\in\mathbb{R}^n$.

(2) A 正定 $\Leftrightarrow r=s=n \underbrace{\Leftrightarrow}_{A={}^tTT\Leftrightarrow} {}^tXAX>0$，$\forall\ X=[x_1,\cdots,x_n]\neq 0$

同理可证(3)和(4).

(5). 若 A 是不定的，即 $0<s<r$，令 $X=T^{-1}\cdot[1,0\cdots0]$，则 ${}^tXAX=1>0$，令 $Y=T^{-1}\cdot[\overbrace{0\cdots0}^{s},1,0\cdots0]$，则 ${}^tYAY=-1<0$。反之，若存在 $X=[x_1,\cdots,x_n]$，$Y=[y_1,\cdots,y_n]\in\mathbb{R}^n$ 使 ${}^tXAX>0$，${}^tYAY<0$。则由 (1) 知，A 不是半正定，即 $s<r$。因为 ${}^tXAX>0$，~~[illegible]~~ 由 (3) 知 A 不是半负定的，故 $s\neq0$。所以 $0<s<r$，即 A 是不定的。

□.

在将对称矩阵 A 合同等价地化为对角型：

$$
{}^tTAT=\begin{pmatrix}\lambda_1 & & 0\\ & \ddots & \\ 0 & & \lambda_n\end{pmatrix}
$$

时，对角线上的 $\lambda_1,\cdots,\lambda_n$ 依赖于 T 的选取。下面的定理表明，对某些特殊的对称矩阵 A，可以选取 T 使对角线上的数直接由 A 确定。

<u>定义4.3.5</u>. 设 $A=(a_{ij})_{n\times n}\in M_n(K)$，则

$$
\Delta_k=\begin{vmatrix}a_{11} & a_{12} & \cdots & a_{1k}\\ a_{21} & a_{22} & \cdots & a_{2k}\\ \vdots & \vdots & & \vdots\\ a_{k1} & a_{k2} & \cdots & a_{kk}\end{vmatrix}\qquad(1\leq k\leq n)
$$

称为 A 的 k-阶主子式。

<u>定理4.3.3</u>. 设 $A=(a_{ij})_{n\times n}\in M_n(K)$ 是对称矩阵且 A 的<u>所有主子式</u> $\Delta_1,\cdots,\Delta_n$ <u>都不为零</u>。则存在可逆矩阵 $T\in M_n(K)$ 使

$$
{}^tTAT=\begin{pmatrix}\Delta_1/\Delta_0 & & & 0\\ & \Delta_2/\Delta_1 & & \\ & & \ddots & \\ 0 & & & \Delta_n/\Delta_{n-1}\end{pmatrix}_{n\times n}\quad 其中\ \Delta_0=1.
$$

证明，设 V 是 n 维 K-向量空间，$e_1,\cdots,e_n\in V$ 是一组基。定义

$$
f(x,y):=(e_1^*(x),\cdots,e_n^*(x))A\begin{pmatrix}e_1^*(y)\\ \vdots\\ e_n^*(y)\end{pmatrix}.
$$

则 $f\in\mathcal{L}_2^+(V)$，$a_{ij}=f(e_i,e_j)$。

所以只需证明：存在一组基 $\alpha_1,\cdots,\alpha_n\in V$ 使 $f(\alpha_i,\alpha_j)=0\ (i\neq j)$，

$f(\alpha_i,\alpha_i)=\Delta_i/\Delta_{i-1}$ ~~$f(x,y)=\frac{\Delta_1}{\Delta_0}\alpha_1^*(x)\alpha_1^*(y)+\frac{\Delta_2}{\Delta_1}\alpha_2^*(x)\alpha_2^*(y)+\cdots+\frac{\Delta_n}{\Delta_{n-1}}\alpha_n^*(x)\alpha_n^*(y)$~~.

为证该结论，对 $\dim_K(V)$ 用归纳法：当 $\dim_K(V)=1$ 时，结论显然成立。设结论对 $n-1$ 成立，令 $V_1=\langle e_1,\cdots,e_{n-1}\rangle\subset V$，$f_1=f|_{V_1\times V_1}:V_1\times V_1\to K$。则 f_1 在 $e_1,\cdots,e_{n-1}$ 下的坐标矩阵是 $A_1=(a_{ij})_{(n-1)\times(n-1)}$，$A_1$ 的 $n-1$ 个主子式就是 A 的前 $n-1$ 个主子式 $\Delta_1,\Delta_2,\cdots,\Delta_{n-1}$。由归纳假设，存在 V_1 的一组基 $\alpha_1,\cdots,\alpha_{n-1}$ 使 $f(\alpha_i,\alpha_j)=0$，$f(\alpha_i,\alpha_i)=\Delta_i/\Delta_{i-1}\ (1\le i\le n-1)$。设 $\beta\in V$ 是 $f(\alpha_1,x)=\cdots=f(\alpha_{n-1},x)=0$ 的非零解。则对任意的 $0\neq\lambda\in K$，向量 $\alpha_1,\cdots,\alpha_{n-1},\lambda\beta$ 线性无关：若 $\mu_1\alpha_1+\cdots+\mu_{n-1}\alpha_{n-1}+\mu_n\lambda\beta=0$，则

$$\mu_i f(\alpha_i,\alpha_i)=f(\alpha_i,\mu_1\alpha_1+\cdots+\mu_{n-1}\alpha_{n-1}+\mu_n\lambda\beta)=0\quad(1\le i\le n-1)$$

从而 $\mu_i=0$，$\mu_n=0$。令 $(\alpha_1,\cdots,\alpha_{n-1},\beta)=(e_1,\cdots,e_n)B$，$\alpha_n=\frac{1}{|B|}\beta$。则 $\alpha_1,\cdots,\alpha_n\in V$ 是一组基。且 $f(\alpha_i,\alpha_j)=0\ (i\neq j)$，$f(\alpha_i,\alpha_i)=\Delta_i/\Delta_{i-1}\ (1\le i\le n-1)$。

为计算 $f(\alpha_n,\alpha_n)$，令 $(\alpha_1,\cdots,\alpha_n)=(e_1,\cdots,e_n)T$，则 $|T|=\frac{1}{|B|}\cdot|B|=1$。且

$$^tTAT=\begin{pmatrix} f(\alpha_1,\alpha_1) & & 0\\ & \ddots & \\ 0 & & f(\alpha_n,\alpha_n)\end{pmatrix}_{n\times n}=\begin{pmatrix}\Delta_1/\Delta_0 & & & \\ & \Delta_2/\Delta_1 & & \\ & & \ddots & \\ & & & \Delta_{n-1}/\Delta_{n-2} & \\ & & & & f(\alpha_n,\alpha_n)\end{pmatrix}.$$

所以，$\Delta_n=|A|=f(\alpha_1,\alpha_1)\cdots f(\alpha_{n-1},\alpha_{n-1})\cdot f(\alpha_n,\alpha_n)=\frac{\Delta_1}{\Delta_0}\cdot\frac{\Delta_2}{\Delta_1}\cdots\frac{\Delta_{n-1}}{\Delta_{n-2}}\cdot f(\alpha_n,\alpha_n)$

故 $f(\alpha_n,\alpha_n)=\frac{|A|}{\Delta_{n-1}}=\frac{\Delta_n}{\Delta_{n-1}}$。

□

推论4.3.3：设 $A=(a_{ij})_{n\times n}\in M_n(\mathbb{R})$ 是对称矩阵。则 A 是正定矩阵的充要条件是 $\Delta_i>0\ (1\le i\le n)$。

证明：若 $\Delta_i>0\ (1\le i\le n)$，由上述定理4.3.2 知 A 是正定矩阵。若 A 是正定矩阵。则对任意 $1\le k\le n$，设 $X=[x_1,\cdots,x_k,0\cdots0]\neq0$ 有

$$0<{}^tXAX=(x_1,\cdots,x_k)A_k[x_1,\cdots,x_k]，其中\ A_k=\begin{pmatrix}a_{11}&a_{12}&\cdots&a_{1k}\\a_{21}&a_{22}&\cdots&a_{2k}\\\vdots&\vdots&&\vdots\\a_{k1}&a_{k2}&\cdots&a_{kk}\end{pmatrix}.$$

由推论4.3.1，A_k 是正定矩阵，特别，故 $\Delta_k=|A_k|>0$。

□

最后介绍一种求可逆矩阵 T，使得

$$^{t}TAT=\begin{pmatrix}\lambda_1 & & & 0\\ & \lambda_2 & & \\ & & \ddots & \\ 0 & & & \lambda_n\end{pmatrix}$$

的算法，它的理论依据基于如下的观察：令 $^{t}T=P_s\cdots P_1$ 是初等矩阵之积，则 $T={}^{t}P_1\cdot{}^{t}P_2\cdots{}^{t}P_s$。同时注意到 $^{t}F_{s,t}=F_{s,t}$，$^{t}F_s(\lambda)=F_s(\lambda)$，$^{t}F_{s,t}(\lambda)=F_{t,s}(\lambda)$（见命题4.1.14）及 $A\cdot F_{t,s}(\lambda)$ 等于将 A 的第 s 列乘 λ 加到第 t 列（见命题4.1.14(2)）。所以，对称矩阵 A 可以通过对 A 的行与列同时做相同的初等变换化为对角型矩阵。要精确表述如下：

定义4.3.6：设 $A=(a_{ij})_{n\times n}\in M_n(K)$。$A$ 的下列变换称为合同变换

(i) 交换 A 的第 i 行与第 j 行，再交换第 i 列与第 j 列；

(ii) 将 A 的第 i 行乘以常数 λ 加到第 j 行，再将第 i 列乘 λ 加到第 j 列。

(iii) 将 A 的第 i 行乘非零常数 c，再将第 i 列乘相同的 c。 □

所以定理4.3.2中的结论(1)可以等价地叙述为：对称矩阵 A 可以通过若干次合同变换化为对角型矩阵。因此，如同求逆矩阵的算法，我们只需将合同变换中的行变换或列变换记录下来即得 ^{t}T。设 $T=P_s\cdots P_1$ 使 $TA{}^{t}T$ 成对角型，则

$$(A\,|\,I_n)\xrightarrow{①}(P_1A{}^{t}P_1\,|\,P_1I_n)\xrightarrow{②}\cdots\longrightarrow(P_s\cdots P_1A{}^{t}P_1\cdots{}^{t}P_s\,|\,P_s\cdots P_1\cdot I_n).$$

例4.3.1. 设 $A=\begin{pmatrix}1&-1&1\\-1&-3&-3\\1&-3&0\end{pmatrix}$，令 $(A\,|\,I_3)=\left(\begin{array}{ccc|ccc}1&-1&1&1&0&0\\-1&-3&-3&0&1&0\\1&-3&0&0&0&1\end{array}\right)$

$$\xrightarrow{F_{1,2}(1)}\left(\begin{array}{ccc|ccc}1&0&1&1&0&0\\0&-4&-2&1&1&0\\1&-2&0&0&0&1\end{array}\right)\xrightarrow{F_{1,3}(-1)}\left(\begin{array}{ccc|ccc}1&0&0&1&0&0\\0&-4&-2&1&1&0\\0&-2&-1&-1&0&1\end{array}\right)$$

$$\xrightarrow{F_{2,3}(-\frac{1}{2})}\left(\begin{array}{ccc|ccc}1&0&0&1&0&0\\0&-4&0&1&1&0\\0&0&0&-\frac{3}{2}&-\frac{1}{2}&1\end{array}\right).$$ 得 $T=\begin{pmatrix}1&0&0\\1&1&0\\-\frac{3}{2}&-\frac{1}{2}&1\end{pmatrix}$ 使 $TA{}^{t}T=\begin{pmatrix}1&&\\&-4&\\&&0\end{pmatrix}$.

矩阵A对应的二次型为 $q(x_1,x_2,x_3)=x_1^2-2x_1x_2+2x_1x_3-6x_2x_3$. 令

$$\begin{pmatrix}x_1\\x_2\\x_3\end{pmatrix}={}^tT\begin{pmatrix}y_1\\y_2\\y_3\end{pmatrix}=\begin{pmatrix}y_1+y_2-\frac{3}{2}y_3\\y_2-\frac{1}{2}y_3\\y_3\end{pmatrix},\quad 得\ q(x_1,x_2,x_3)=(y_1,y_2,y_3)\,TA{}^tT\begin{pmatrix}y_1\\y_2\\y_3\end{pmatrix}$$

$$=y_1^2-4y_2^2=(x_1-x_2+x_3)^2-(4x_2+2x_3)^2.$$

□

习题 4.3.

1. 求变量替换 $[x_1,x_2,x_3,x_4]=C\cdot[y_1,y_2,y_3,y_4]$, $C\in M_4(K)$ 可逆, 使下列二次型成为对角型:

(a) $q(x_1,x_2,x_3,x_4)=3(x_1^2+x_2^2+x_3^2+x_4^2)+2x_1x_2+2x_1x_4-2x_2x_3-2x_2x_4-2x_3x_4$

(b) $q(x_1,x_2,x_3,x_4)=x_1x_2+x_2x_3+x_1x_4-4x_2x_4+6x_3x_4$

2. 令 $A=\begin{pmatrix}0&2&-1&3\\-2&0&4&-2\\1&-4&0&1\\-3&2&-1&0\end{pmatrix}$, $J=\begin{pmatrix}0&1&0&0\\-1&0&0&0\\0&0&0&1\\0&0&-1&0\end{pmatrix}$, 求可逆矩阵C使 $J=CA{}^tC$.

3. 设A是半正定矩阵, 证明: 伴随矩阵 A^* 也是半正定的.

4 举例说明: (1) 存在正定矩阵 $A=(a_{ij})_{n\times n}$, 它有某个 $a_{ij}<0$. (2) 存在对称实矩阵 $A=(a_{ij})_{n\times n}$ 满足 $a_{ij}>0$ $(1\le i,j\le n)$, 但A不是正定矩阵.

5 设A是任意实对称矩阵, 证明: 对充分小的实数 ε, $B=I_n+\varepsilon A$ 是正定矩阵.

6. 证明下述结论等价: (1) A是正定矩阵; (2) 存在对角线元素全为1的上三角矩阵B使得 $A={}^tBDB$, 其中D是对角线元素全为正数的对角矩阵; (3) 存在对角线元素全为正数的上三角矩阵C使 $A={}^tC\cdot C$.

7. 设A是正定矩阵, 证明: 二次型

$$g(x_1,\cdots,x_n)=\begin{vmatrix}a_{11}&a_{12}&\cdots&a_{1n}&x_1\\a_{21}&a_{22}&\cdots&a_{2n}&x_2\\\vdots&\vdots&&\vdots&\vdots\\a_{n1}&a_{n2}&\cdots&a_{nn}&x_n\\x_1&x_2&\cdots&x_n&0\end{vmatrix}$$

是负定的.

8. 设 $A=(a_{ij})_{n\times n}$ 是正定矩阵，证明：$|A|\le a_{nn}\Delta_{n-1}$，$\Delta_{n-1}$是$(n-1)$阶主子式。

9. 如果 $A=(a_{ij})_{n\times n}$ 是正定矩阵，试证：$|A|\le a_{11}a_{22}\cdots a_{nn}$.

10. 设 A 是 n 阶实对称矩阵，证明：存在实数 $\lambda>0$ 使得对任意实数 $x_1,\dots,x_n$ 有

$$(x_1,\dots,x_n)A\begin{pmatrix}x_1\\ \vdots\\ x_n\end{pmatrix}\le\lambda(x_1^2+\dots+x_n^2).$$

11. 证明：对称矩阵 A 是半正定的充要条件是存在矩阵 $B\in M_n(\mathbb{R})$ 使 $A={}^tB\cdot B$，且 $r(B)=r(A)$.

12. 设 $T\in M_n(K)$ 是一个可逆矩阵，证明映射

$$f_T:\ M_n(K)\longrightarrow M_n(K),\quad A\longmapsto T^{-1}AT,$$

是一个K-线性映射且满足：$f_T(AB)=f_T(A)\cdot f_T(B)$.

13. 设 $A,B\in M_n(\mathbb{R})$ 是两个实矩阵，如果存在可逆复系数矩阵 $T\in M_n(\mathbb{C})$ 使 $B=T^{-1}AT$，试证明：存在实可逆矩阵 $C\in M_n(\mathbb{R})$ 使得 $B=C^{-1}AC$.

*14. 设 $f:\ M_n(k)\longrightarrow M_n(k)$ 是一个双射K-线性映射，且满足条件：

$$f(A\cdot B)=f(A)\cdot f(B),\quad \forall A,B\in M_n(k).$$

证明：存在可逆矩阵 $T\in M_n(K)$ 使

$$f(X)=T^{-1}XT,\quad \forall X\in M_n(K).$$

15. 设 $l:\ M_n(K)\longrightarrow K$ 是 K-向量空间 $M_n(K)$ 上的任一线性函数。试证明：存在 $A\in M_n(K)$，使得

$$f(X)=tr(A\cdot X),\quad \forall X\in M_n(K).$$

此处 $tr(B)=b_{11}+\dots+b_{nn}$ 表示矩阵 $B=(b_{ij})_{n\times n}$ 的迹（提示：证明双线性函数 $tr(\cdot,\cdot):\ M_n(K)\times M_n(K)\longrightarrow K,\ (A,B)\longmapsto tr(A\cdot B)$ 非退化）.

§4.4 矩阵分块与应用

本节我们介绍一种在矩阵研究中非常有用的技巧：矩阵分块。对矩阵 A 进行分块后的表述形式称为分块矩阵，而矩阵的运算，初等变换等可以表述为分块矩阵的运算和初等变换。需要特别指出：分块矩阵的运算及初等变换不是重新定义的，而是仅仅是原来矩阵运算和初等变换在分块矩阵语言下的重新表述。

一般地，对任意 $m\times n$ 矩阵 A，若先用若干条横虚线把它分成 r 块，再用若干条竖虚线把它分成 s 块，则我们得到一个 rs 块的分块矩阵，A 可记为：

$$A=\begin{pmatrix} A_{11} & A_{12} & \cdots & A_{1s} \\ A_{21} & A_{22} & \cdots & A_{2s} \\ \vdots & & & \\ A_{r1} & A_{r2} & \cdots & A_{rs} \end{pmatrix}=(A_{ij})_{r\times s} \qquad (A_{ij}\text{是矩阵})$$

例 4.4.1：

$$A=\left(\begin{array}{ccc|c} 1 & 2 & 1 & 1 \\ 2 & 5 & 3 & -2 \\ \hline 3 & 1 & -1 & 3 \end{array}\right),\quad \text{可记为 } A=\begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix}$$

其中 $A_{11}=\begin{pmatrix} 1 & 2 & 1 \\ 2 & 5 & 3 \end{pmatrix}_{2\times 3}$，$A_{12}=\begin{pmatrix} 1 \\ -2 \end{pmatrix}_{2\times 1}$，$A_{21}=(3\ \ 1\ \ -1)_{1\times 3}$，$A_{22}=(1)_{1\times 1}$

$\square$

另外，我们将一个矩阵 A 表示成列向量形式 $A=(A^{(1)},\cdots,A^{(n)})$ 或行向量形式 $A=[A_{(1)},\cdots,A_{(m)}]$ 也是一种矩阵分块。当然，怎么对一个矩阵分块？需要根据具体矩阵的形式和具体的问题来决定，是一个技巧性比较高的问题。例如，如果需要计算下列矩阵：

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

的逆矩阵，或A的幂A^m，则我们可用分块技巧将A表示成分块矩阵．

$$A=\begin{pmatrix} A_1 & 0 & 0 \\ 0 & A_2 & 0 \\ 0 & 0 & A_3 \end{pmatrix}，其中 A_1=\begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}，A_2=\begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}$$

$A_3=(1)$，0表示不同阶数的零矩阵。通过将矩阵运算总结成分块矩阵的运算，我们可以得到

$$A^{-1}=\begin{pmatrix} A_1^{-1} & 0 & 0 \\ 0 & A_2^{-1} & 0 \\ 0 & 0 & A_3^{-1} \end{pmatrix}，\quad A^m=\begin{pmatrix} A_1^m & 0 & 0 \\ 0 & A_2^m & 0 \\ 0 & 0 & A_3^m \end{pmatrix}。$$

定理4.4.1．设 $A,B\in M_{m\times n}(K)$，如果它们有相同的分块．
$A=(A_{ij})_{r\times s}$，$B=(B_{ij})_{r\times s}$，且A_{ij}与B_{ij}的行数，列数相等．则

(1) $A+B=(A_{ij}+B_{ij})_{r\times s}$，$A-B=(A_{ij}-B_{ij})_{r\times s}$

(2) $\forall\,\lambda\in K$，$A=(A_{ij})_{r\times s}$，则 $\lambda A=(\lambda A_{ij})_{r\times s}$

(3) 设 $A\in M_{n\times m}(K)$，$B\in M_{m\times a}(K)$，如果 $A=(A_{ij})_{r\times s}$，$B=(B_{ij})_{s\times t}$
且 $A_{ij}\in M_{m_i\times n_j}(K)$，$B_{ij}\in M_{n_i\times l_j}(K)$，则

$$A\cdot B=(C_{ij})_{r\times t}．$$

其中 $C_{ij}=A_{i1}B_{1j}+A_{i2}B_{2j}+\cdots+A_{is}\cdot B_{sj}$．

(4) $${}^t\begin{pmatrix} A_{11} & A_{12} & \cdots & A_{1s} \\ A_{21} & A_{22} & \cdots & A_{2s} \\ \vdots & & & \\ A_{r1} & A_{r2} & \cdots & A_{rs} \end{pmatrix}_{r\times s}=\begin{pmatrix} {}^tA_{11} & {}^tA_{21} & \cdots & {}^tA_{r1} \\ {}^tA_{12} & {}^tA_{22} & \cdots & {}^tA_{r2} \\ \vdots & \vdots & & \vdots \\ {}^tA_{1s} & {}^tA_{2s} & \cdots & {}^tA_{rs} \end{pmatrix}_{s\times r}．$$

$\square$

例 4.4.2：对任意 $A\in M_{m\times n}(K)$，$B\in M_{n\times r}(K)$．如果对 B 做列分块 $B=(B^{(1)},B^{(2)},\cdots,B^{(r)})$．则 $A\cdot B=(AB^{(1)},AB^{(2)},\cdots,AB^{(r)})$．如果对 A 做行分块 $A=[A_{(1)},A_{(2)},\cdots,A_{(m)}]$，则 $AB=[A_{(1)}\cdot B,A_{(2)}\cdot B,\cdots,A_{(m)}\cdot B]$．

□

例 4.4.3．设有两个分块矩阵

$$A=\begin{pmatrix}A_1&0&\cdots&0\\0&A_2&\cdots&0\\\vdots&\vdots&\ddots&\vdots\\0&0&\cdots&A_k\end{pmatrix},\quad B=\begin{pmatrix}B_1&0&\cdots&0\\0&B_2&\cdots&0\\\vdots&\vdots&\ddots&\vdots\\0&0&\cdots&B_k\end{pmatrix}.$$

如果 A_i,B_i 是阶数相同的方阵，则

$$A\cdot B=\begin{pmatrix}A_1B_1&0&\cdots&0\\0&A_2B_2&\cdots&0\\\vdots&\vdots&&\vdots\\0&0&\cdots&A_kB_k\end{pmatrix}.$$

如果每个 A_i 可逆，则 A 也可逆，且

$$A^{-1}=\begin{pmatrix}A_1^{-1}&0&\cdots&0\\0&A_2^{-1}&\cdots&0\\\vdots&\vdots&&\vdots\\0&0&\cdots&A_k^{-1}\end{pmatrix}.$$

□

矩阵 A 的初等变换可以用分块矩阵的形式表述如下：

初等变换(I)：交换分块矩阵的两块行，或两块列的次序．

初等变换(II)：以某个矩阵 C 左乘某分块行加到另一分块行，或以某个矩阵 B 右乘某分块列加到另一分块列．

初等变换(III)：将某可逆矩阵左乘某分块行，或右乘某分块列．

注记：初等变换(II)不改变行列式的值。

例4.4.4.

$$A=\begin{pmatrix} A_{11} & A_{12} & \cdots & A_{1s} \\ A_{21} & A_{22} & \cdots & A_{2s} \\ \vdots & \vdots & & \vdots \\ A_{r1} & A_{r2} & \cdots & A_{rs} \end{pmatrix} \xrightarrow{(\mathrm{II})} \begin{pmatrix} A_{11} & A_{12} & \cdots & A_{1s} \\ A_{21}+CA_{11}, & A_{22}+CA_{12} & \cdots & A_{2s}+CA_{1s} \\ \vdots & \vdots & & \vdots \\ A_{r1} & A_{r2} & & A_{rs} \end{pmatrix}$$

(以矩阵C左乘第一分块行加到第二分块行).

$$A=\begin{pmatrix} A_{11} & A_{12} & \cdots & A_{1s} \\ A_{21} & A_{22} & \cdots & A_{2s} \\ \vdots & \vdots & & \vdots \\ A_{r1} & A_{r2} & \cdots & A_{rs} \end{pmatrix} \xrightarrow{(\mathrm{II})} \begin{pmatrix} A_{11}+A_{12}B & A_{12} & \cdots & A_{1s} \\ A_{21}+A_{22}B & A_{22} & \cdots & A_{2s} \\ \vdots & \vdots & & \vdots \\ A_{r1}+A_{r2}B & A_{r2} & \cdots & A_{rs} \end{pmatrix}.$$

(以矩阵B右乘第二分块列加到第一分块列).

□

例4.4.5: 设 $A\in M_m(K)$ 可逆, $D\in M_n(K)$, $B\in M_{m\times n}(K)$, $C\in M_{n\times m}(K)$,

则
$$\begin{vmatrix} A & B \\ C & D \end{vmatrix} = |A|\cdot|D-CA^{-1}B|.$$

证明: 初等变换(II): 以 $-CA^{-1}$ 左乘第一分块行加到第二分块行.

得
$$\begin{pmatrix} A & B \\ C & D \end{pmatrix} \xrightarrow{(\mathrm{II})} \begin{pmatrix} A & B \\ 0 & D-CA^{-1}B \end{pmatrix}.$$

由下述引理可得

$$\begin{vmatrix} A & B \\ C & D \end{vmatrix} = \begin{vmatrix} A & B \\ 0 & D-CA^{-1}B \end{vmatrix} = |A|\cdot|D-CA^{-1}B|.$$

□

引理4.4.1. 设 $A_{11}\in M_m(K)$, $A_{12}\in M_{m\times n}(K)$, $A_{21}\in M_{n\times m}(K)$, $A_{22}\in M_n(K)$. 则

$$\begin{vmatrix} A_{11} & A_{12} \\ 0 & A_{22} \end{vmatrix} = |A_{11}|\cdot|A_{22}|, \qquad \begin{vmatrix} A_{11} & 0 \\ A_{21} & A_{22} \end{vmatrix} = |A_{11}|\cdot|A_{22}|.$$

证明: 只需证明第二个等式.

令
$$A=\begin{pmatrix} a_{11} & \cdots & a_{1m} & a_{1\,m+1} & \cdots & a_{1\,m+n}\\ \vdots & & \vdots & \vdots & & \vdots\\ a_{m1} & \cdots & a_{mm} & a_{m\,m+1} & \cdots & a_{m\,m+n}\\ a_{m+1\,1} & \cdots & a_{m+1\,m} & a_{m+1\,m+1} & \cdots & a_{m+1\,m+n}\\ \vdots & & \vdots & \vdots & & \vdots\\ a_{m+n\,1} & \cdots & a_{m+n\,m} & a_{m+n\,m+1} & \cdots & a_{m+n\,m+n}\end{pmatrix}=\begin{pmatrix} A_{11} & 0\\ A_{21} & A_{22}\end{pmatrix}.$$

则 $a_{ij}=0\ (1\le i\le m,\ m<j\le m+n)$. 即

$$|A|=\sum_{\pi\in S_{m+n}}\varepsilon_\pi a_{1\pi(1)}\cdots a_{m\pi(m)}\cdot a_{m+1\,\pi(m+1)}\cdots a_{m+n\,\pi(m+n)}.$$

若存在 $1\le i\le m$ 使 $\pi(i)>m$, 则 $a_{1\pi(1)}\cdots a_{m\pi(m)}\cdot a_{m+1\pi(m+1)}\cdots a_{m+n\,\pi(m+n)}=0$.
故只需考虑满足条件 $1\le\pi(i)\le m\ (\forall 1\le i\le m)$ 的置换 $\pi\in S_{m+n}$. ~~$\pi|_{\{1,2,\dots,m\}}$ 满足 $\{1,2,\dots,m\}$ 到一个置换~~. 这样的置换由 $\pi_1=\pi|_{\{1,2,\dots,m\}}\in S_m$ 和 $\pi_2=\pi|_{\{m+1,\dots,m+n\}}\in S_n$ 确定. 故

$$\begin{aligned}|A|&=\sum_{\substack{\pi_1\in S_m\\ \pi_2\in S_n}}\varepsilon_{\pi_1}\cdot\varepsilon_{\pi_2}\, a_{1\pi_1(1)}\cdots a_{m\pi_1(m)}\, a_{m+1\,\pi_2(m+1)}\cdots a_{m+n\,\pi_2(m+n)}\\ &=\Big(\sum_{\pi_1\in S_m}\varepsilon_{\pi_1}a_{1\pi_1(1)}\cdots a_{m\pi_1(m)}\Big)\cdot\Big(\sum_{\pi_2\in S_m}\varepsilon_{\pi_2}a_{m+1\,\pi_2(m+1)}\cdots a_{m+n\,\pi_2(m+n)}\Big)\\ &=|A_{11}|\cdot|A_{22}|.\end{aligned}$$

□.

<u>引理4.4.2</u>: 设 $A\in M_m(K)$, $B\in M_n(K)$, $C\in M_{n\times m}(K)$. 则
$$\begin{vmatrix}0 & A\\ B & C\end{vmatrix}=(-1)^{mn}|A|\cdot|B|.$$

证明: $\begin{pmatrix}0 & A\\ B & C\end{pmatrix}\xrightarrow{(\mathrm{I})}\begin{pmatrix}A & 0\\ C & B\end{pmatrix}$. (初等列变换(I)).

则 $\begin{vmatrix}0 & A\\ B & C\end{vmatrix}=(-1)^{mn}\begin{vmatrix}A & 0\\ C & B\end{vmatrix}=(-1)^{mn}|A|\cdot|B|$

□.

<u>例4.4.6</u>: 设 $A\in M_m(K)$, $D\in M_n(K)$ 可逆. 求矩阵 $\begin{pmatrix}A, & B\\ 0 & D\end{pmatrix}$ 的逆.

解：

$$\left(\begin{array}{cc|cc} A & B & I_m & 0 \\ 0 & D & 0 & I_n \end{array}\right) \xrightarrow[\text{加到第一分块行}]{\text{以}-BD^{-1}\text{乘第二分块行}} \left(\begin{array}{cc|cc} A & 0 & I_m & -BD^{-1} \\ 0 & D & 0 & I_n \end{array}\right)$$

$$\xrightarrow[\text{以}D^{-1}\text{左乘第二分块行}]{\text{以}A^{-1}\text{左乘第一分块行}} \left(\begin{array}{cc|cc} I_m & 0 & A^{-1}, & -A^{-1}BD^{-1} \\ 0 & I_n & 0 & D^{-1} \end{array}\right).\quad \text{故} \begin{pmatrix} A & B \\ 0 & D \end{pmatrix}^{-1} = \begin{pmatrix} A^{-1}, & -A^{-1}BD^{-1} \\ 0 & D^{-1} \end{pmatrix}$$

□.

下面介绍一种实系数矩阵研究中常用的方法，它通常可以将实矩阵的问题转化为可逆实矩阵的问题。

引理4.4.3. 设 $A = (a_{ij})_{n\times n} \in M_n(\mathbb{R})$. 则存在 $\delta>0$ 使得对任意 $0<t<\delta$，下述矩阵

$$tI_n + A = \begin{pmatrix} t+a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & t+a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \\ a_{n1} & a_{n2} & \cdots & t+a_{nn} \end{pmatrix}$$

是可逆矩阵.

证明：行列式 $|tI_n - A| = f(t)$ 是一个首项系数1的n次多项式. 所以 $f(t)$ 在 $\mathbb{R}$ 中最多只有 n 个根. 取 $\delta>0$ 充分小，可使 $f(t)$ 在 $(0,\delta)$ 中没有根. 故对任意 $0<t<\delta$，$|tI_n+A| \neq 0$，因此 tI_n+A 可逆.

□

例4.4.7. 设 $A, B \in M_n(\mathbb{R})$，则 $(AB)^* = B^*A^*$.

证明：若 A, B 可逆，则 $A^* = |A|\cdot A^{-1}$，$B^* = |B|A^{-1}$，$A\cdot B$ 也可逆。故 $(AB)^* = |A\cdot B|\,(AB)^{-1} = |A|\cdot|B|\,B^{-1}A^{-1} = B^*\cdot A^*$.

一般情形：存在 $\delta>0$ 使对任意 $0<t<\delta$，tI_n+A，tI_n-B 可逆.
故 $((tI_n+A)\cdot(tI_n+B))^* = (tI_n+B)^*\cdot(tI_n+A)^*$. 即

$$(t^2 I_n + t(A+B) + AB)^* = (tI_n+B)^*\cdot(tI_n+A)^*.$$

令 $t\to 0$，两边的极限分别是 $(AB)^*$ 和 B^*A^*. 所以 $(AB)^* = B^*A^*$

□

例4.4.8：证明：$|A^*|=|A|^{n-1}$，$(A^*)^*=|A|^{n-2}\cdot A$.

证明：若 A 可逆，则 $A^*=|A|A^{-1}$. 于是

$$|A^*|=|A|^n|A^{-1}|=|A|^{n-1},\quad (A^*)^*=|A^*|(A^*)^{-1}=|A^*|(|A|A^{-1})^{-1}=|A|^{n-2}A.$$

对一般矩阵 A，存在 $\delta>0$ 使得 tI_n+A 可逆 $(0<t<\delta)$. 所以

$$|(tI_n+A)^*|=|tI_n+A|^{n-1},\quad ((tI_n+A)^*)^*=|tI_n+A|^{n-2}(tI_n+A)$$

令 $t\to 0$ 得 $|A^*|=|A|^{n-1}$，$(A^*)^*=|A|^{n-2}A$ □

习题4.4

1. 设 $A,B\in M_n(\mathbb{C})$ 满足 $AB=BA$，证明

$$\begin{vmatrix} A & -B \\ B & A \end{vmatrix}=|A^2+B^2|.$$

2. 设 A,B,C,D 都是 n 阶方阵. 证明：

$$\begin{vmatrix} A & B & C & D \\ B & A & D & C \\ C & D & A & B \\ D & C & B & A \end{vmatrix}=|A+B+C+D|\cdot|A+B-C-D|\cdot|A-B+C-D|\cdot|A-B-C+D|.$$

3. 设 $A\in M_m(\mathbb{R})$，$B\in M_n(\mathbb{R})$，求 $\begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix}^*$.

4. 设 $A,B,C,D\in M_n(\mathbb{R})$，且 $AC=CA$，证明

$$\begin{vmatrix} A & B \\ C & D \end{vmatrix}=|AD-CB|.$$

(提示：先证 A 可逆的情形)

5. 设 $A\in M_{m\times s}(\mathbb{R})$，$B\in M_{m\times t}(\mathbb{R})$. 证明：

$$r\left(\begin{pmatrix} A & B \\ 2A & -5B \end{pmatrix}\right)=r(A)+r(B).$$

6. 设 $A, B \in M_n(\mathbb{R})$，证明：$r\left(\begin{pmatrix} A & AB \\ B & B+B^2 \end{pmatrix}\right) = r(A) + r(B)$.

7. 设 t 是不定元，

$$A(t) = \begin{pmatrix} a_{11}+t & a_{12}+t & \cdots & a_{1n}+t \\ a_{21}+t & a_{22}+t & \cdots & a_{2n}+t \\ \vdots & \vdots & & \vdots \\ a_{n1}+t & a_{n2}+t & \cdots & a_{nn}+t \end{pmatrix}_{n\times n}$$

证明：$|A(t)| = |A(0)| + \sum_{i,j}^{n} A_{ij}$，其中 A_{ij} 是 a_{ij} 在 $A(0)$ 中的代数余子式.

§4.5. 一般系数的矩阵和行列式

从前面的讨论可以看到，矩阵 $A=(a_{ij})_{m\times n}$ 中系数 a_{ij} 可以不是域 K 中的元素，它可以是多项式、甚至是矩阵。在本节我们将矩阵和行列式的理论推广至系数 a_{ij} 在一般的环 R 中，可以看到绝大部分定理和结果对一般系数的矩阵和行列式。如果不加说明，本节关于矩阵的结论对任意的环 R 成立，而关于行列式的定义和结果则仅对交换环成立。初学者可以将环 R 想象为下列具体例子：(1) $R=M_n(K)$ 是矩阵环，(2) $R=K[x]$ 是域 K 上的多项式环，(3) R 是域或整数环 $\mathbb{Z}$ 等。

定义 4.5.1：设 R 是一个环，$M_{m\times n}(R) = \{A=(a_{ij})_{m\times n} \mid a_{ij}\in R\}$ 是所有系数在 R 中的 $m\times n$ 矩阵的集合。$\forall\ A=(a_{ij})_{m\times n}, B=(b_{ij})_{m\times n}$，$\lambda\in R$，定义：$A+B := (a_{ij}+b_{ij})_{m\times n}$，$\lambda A = (\lambda a_{ij})_{m\times n}$. □

定义 4.5.2 [矩阵乘法]. 设 $A=(a_{ij})_{m\times n}\in M_{m\times n}(R)$，$B=(b_{ij})_{n\times s}\in M_{n\times s}(R)$，则 $A\cdot B = (c_{ij})_{m\times s}\in M_{m\times s}(R)$，其中

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

称为 A 与 B 的乘积. □

命题4.5.1 若 $A\in M_{m\times n}(R)$, $B\in M_{n\times s}(R)$, $C\in M_{s\times t}(R)$. 则

(1) $(A\cdot B)\cdot C = A\cdot(B\cdot C)$ (结合律).

(2) $(A_1+A_2)\cdot B = A_1\cdot B + A_2\cdot B$ $(\forall A_1,A_2\in M_{m\times n}(R))$.

(3) $A\cdot(B_1+B_2) = A\cdot B_1 + A\cdot B_2$. $(\forall B_1,B_2\in M_{n\times s}(R))$.

(4) $\forall\ \lambda\in R$, $A=(a_{ij})_{m\times n}\in M_{m\times n}(R)$,

$$\lambda A = \begin{pmatrix} \lambda & 0 & \cdots & 0 \\ 0 & \lambda & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda \end{pmatrix}_{m\times m}\cdot A \in M_{m\times n}(R).$$

证明: 直接验证即可. 我们仅验证(1). 设 $A=(a_{ij})_{m\times n}$, $B=(b_{ij})_{n\times s}$, $C=(c_{ij})_{s\times t}$. 则 $(A\cdot B)\cdot C$ 在 (i,j) 位置的元素为

$$\left[(A\cdot B)\cdot C\right]_{(i,j)} = \left(\sum_{k=1}^{n} a_{ik}b_{k1}\right)c_{1j} + \left(\sum_{k=1}^{n} a_{ik}b_{k2}\right)c_{2j} + \cdots + \left(\sum_{k=1}^{n} a_{ik}b_{ks}\right)c_{sj}$$

$$= a_{i1}\left(\sum_{k=1}^{s} b_{1k}c_{kj}\right) + a_{i2}\left(\sum_{k=1}^{s} b_{2k}c_{kj}\right) + \cdots + a_{in}\left(\sum_{k=1}^{s} b_{nk}c_{kj}\right).$$

$$= a_{i1}\left(B_{(1)}\cdot C^{(j)}\right) + a_{i2}\left(B_{(2)}\cdot C^{(j)}\right) + \cdots + a_{in}\left(B_{(n)}C^{(j)}\right)$$

$$= A_{(i)}\cdot(BC)^{(j)} = \left[A\cdot(BC)\right]_{(i,j)}.$$ 故 $(A\cdot B)\cdot C = A\cdot(BC)$. □

推论4.5.1: 设 R 是一个环, 则 $M_n(R)$ 关于矩阵的加法和乘法也是一个环 R.

证明: 定义4.1.2中的条件(1)至(4)显然由 R 的性质推出, 零元 $O_{n\times n}=(0)_{n\times n}$. 条件(5)和(7)由命题4.5.1的(1), (2)和(3)推出. 条件(6)中的1取单位矩阵 $I_n=\mathrm{diag}(1,\cdots,1)$ 即可. □

定义4.5.3 [可逆矩阵]. $A\in M_n(R)$ 称为可逆矩阵(或简称可逆), 如果存在 $B\in M_n(R)$ 使 $A\cdot B = B\cdot A = I_n$. □

注记: 满足 $A\cdot B = BA = I_n$ 的 B 唯一, 故称 B 是 A 的逆矩阵, 记为 $B = A^{-1}$.

定义4.5.4 [行列式]. 设 $A=(a_{ij})_{n\times n}\in M_n(R)$，$R$是交换环.

$$\det[A_{(1)},\cdots,A_{(n)}]=|A|=\sum_{\pi\in S_n}\varepsilon_\pi\, a_{1\pi(1)}\cdot\cdots\cdot a_{n\pi(n)}\quad\in R.$$

称为矩阵 A 的行列式. □

定理4.5.1 [行列式基本性质]. 设R是交换环. $A=(a_{ij})_{n\times n}\in M_n(R)$. 则

(D1) $\det[A_{(1)},\cdots,A_{(n)}]$ 是关于行向量的反对称函数：$\forall\pi\in S_n$

$$\det[A_{(\pi(1))},\cdots,A_{(\pi(n))}]=\varepsilon_\pi\det[A_{(1)},\cdots,A_{(n)}].$$

(D2) $\det[A_{(1)},\cdots,A_{(n)}]$ 是关于行向量的n重线性函数：若 $A_{(i)}=\lambda A'_{(i)}+\mu A''_{(i)}$，

$\det[A_{(1)},\cdots,\lambda A'_{(i)}+\mu A''_{(i)},\cdots,A_{(n)}]=\lambda\det[A_{(1)},\cdots,A'_{(i)},\cdots,A_{(n)}]+\mu\det[A_{(1)},\cdots,A''_{(i)},\cdots,A_{(n)}]$

(D3) $\det[I_{(1)},\cdots,I_{(n)}]=|I_n|=1$，其中 $I_n=[I_{(1)},\cdots,I_{(n)}]$ 是单位矩阵.

(D4) 如果A中两行相同，则 $\det[A_{(1)},\cdots,A_{(n)}]=|A|=0$.

(D5) 对任意 $\lambda\in R$，$\det[\lambda A_{(1)},\cdots,\lambda A_{(n)}]=\det(\lambda A)=\lambda^n|A|$.

(D6) 如果A的某行为零，则 $\det(A)=|A|=0$.

(D7). $\forall\ \lambda\in R$，有 $\det[A_{(1)},\cdots,A_{(j)},\cdots\lambda A_{(j)}+A_{(k)},\cdots,A_{(n)}]=\det[A_{(1)},\cdots,A_{(n)}]$.

(D8) $|{}^tA|=|A|$. 故上述所有对行向量成立的结论对列向量也成立.

证明：仅验证(D1)，(D2)，(D4)，(D8)，其它性质是它们的简单推论. (D1)，(D2)，(D4)，(D8)也均可由定义直接验证.

(D1)：$\det[A_{(\pi(1))},\cdots,A_{(\pi(n))}]=\sum_{\sigma\in S_n}\varepsilon_\sigma\, a_{\pi(1)\sigma(1)}\cdot a_{\pi(2)\sigma(2)}\cdot\cdots a_{\pi(n)\sigma(n)}$

$$=\sum_{\sigma\in S_n}\varepsilon_\sigma\, a_{1\sigma\pi^{-1}(1)}\cdot a_{2\sigma\pi^{-1}(2)}\cdots a_{n\sigma\pi^{-1}(n)}=\varepsilon_\pi\sum_{\sigma\in S_n}\varepsilon_{\sigma\pi^{-1}}a_{1\sigma\pi^{-1}(1)}\cdots a_{n\sigma\pi^{-1}(n)}$$

$$=\varepsilon_\pi\det[A_{(1)},\cdots,A_{(n)}]=\varepsilon_\pi\cdot|A|.$$

(D2)：为简化符号，仅验证它关于 $A_{(1)}$ 是线性的：若 $A_{(1)}=\lambda A'_{(1)}+\mu A''_{(1)}$

即 $a_{1j}=\lambda a'_{1j}+\mu a''_{1j}$，其中 $A'_{(1)}=(a'_{1j},\cdots a'_{nj})$，$A''_{(1)}=(a''_{1j},\cdots a''_{nj})\in R^n$，

则 $\det[\lambda A'_{(1)}+\mu A''_{(1)},A_{(2)},\cdots,A_{(n)}]=\sum\limits_{\pi\in S_n}\varepsilon_\pi\left(\lambda a'_{1\pi(1)}+\mu a''_{1\pi(1)}\right)a_{2\pi(2)}\cdots a_{n\pi(n)}$

$$=\lambda\sum_{\pi\in S_n}\varepsilon_\pi a'_{1\pi(1)}a_{2\pi(2)}\cdots a_{n\pi(n)}+\mu\sum_{\pi\in S_n}\varepsilon_\pi a''_{1\pi(1)}a_{2\pi(2)}\cdots a_{n\pi(n)}.$$

$$=\lambda\det[A'_{(1)},A_{(2)},\cdots,A_{(n)}]+\mu\det[A''_{(1)},A_{(2)},\cdots,A_{(n)}].$$

(D4)：若 $A_{(i)}=A_{(j)}$，$i\neq j$，令 $\sigma=(i,j)\in S_n$．则

$$a_{1\pi(1)}\cdots a_{n\pi(n)}=a_{1\pi\sigma(1)}\cdots a_{n\pi\sigma(n)}.$$

令 $A_n\subset S_n$ 是所有偶置换的集合，$\bar{A}_n=S_n\setminus A_n$ 是所有奇置换的集合，则映射 $A_n\longrightarrow\bar{A}_n$，$\pi\longmapsto\pi\sigma$，是双射．

$$|A|=\sum_{\pi\in A_n}a_{1\pi(1)}\cdots a_{n\pi(n)}-\sum_{\pi\in\bar{A}_n}a_{1\pi(1)}\cdots a_{n\pi(n)}=\sum_{\pi\in A_n}a_{1\pi(1)}\cdots a_{n\pi(n)}$$

$$-\sum_{\tau\in A_n}a_{1\tau\sigma(1)}\cdots a_{n\tau\sigma(n)}=\sum_{\pi\in A_n}a_{1\pi(1)}\cdots a_{n\pi(n)}-\sum_{\tau\in A_n}a_{1\tau(1)}\cdots a_{n\tau(n)}=0.$$

(D8)：令 ${}^tA=(a'_{ij})_{n\times n}$，则 $a'_{ij}=a_{ji}$．故

$$|{}^tA|=\sum_{\pi\in S_n}\varepsilon_\pi a'_{1\pi(1)}\cdots a'_{n\pi(n)}=\sum_{\pi\in S_n}\varepsilon_\pi a_{\pi(1)1}\cdots a_{\pi(n)n}$$

$$=\sum_{\pi\in S_n}\varepsilon_\pi a_{1\pi^{-1}(1)}\cdots a_{n\pi^{-1}(n)}=\sum_{\pi\in S_n}\varepsilon_{\pi^{-1}}a_{1\pi^{-1}(1)}\cdots a_{n\pi^{-1}(n)}=|A|.$$

□

所以，对矩阵 $A=(a_{ij})_{n\times n}\in M_n(R)$ 有与定理4.2.1完全一样的行列式展开定理，和伴随矩阵的定义．

定义4.5.5：[伴随矩阵]．设 $A=(a_{ij})_{n\times n}\in M_n(R)$，矩阵

$$A^*=\begin{pmatrix}A_{11}&A_{21}&\cdots&A_{n1}\\A_{12}&A_{22}&\cdots&A_{n2}\\\vdots&\vdots&\ddots&\vdots\\A_{1n}&A_{2n}&\cdots&A_{nn}\end{pmatrix}_{n\times n}\in M_n(R)$$

称为 A 的伴随矩阵．

□

推论4.5.2：设R是交换环，$A \in M_n(R)$，则 $A^*A = A\cdot A^* = |A|\cdot I_n$.

特别，若 $|A| \in R$ 在R中可逆，则 A 可逆，且 $A^{-1} = |A|^{-1}\cdot A^*$.

证明：展开定理4.2.1对 $M_n(R)$ 中矩阵也成立，故 $A\cdot A^* = A^*A = |A|\cdot I_n$。若 $|A| \in R$ 在R中可逆，即存在 $b \in R$ 使 $b\cdot|A| = |A|\cdot b = 1$，则 b 由 |A| 唯一确定，记为 $b := |A|^{-1} = \frac{1}{|A|} \in R$. 故 $A\cdot(|A|^{-1}\cdot A^*) = |A|^{-1}(A\cdot A^*) = |A|^{-1}\cdot(|A|\cdot I_n) = I_n$. 同理 $(|A|^{-1}\cdot A^*)A = I_n$，所以 $|A|^{-1}\cdot A^*$ 是 A 的逆矩阵．

□.

定义4.5.6［初等变换］ 设 $A = [A_{(1)}, \cdots, A_{(m)}] = \{A^{(1)}, \cdots, A^{(n)}\} \in M_{m\times n}(R)$.

初等行变换(I)：交换第i行与第j行的位置.

初等行变换(II)：将第i行乘以 $\lambda \in R$ 加到第j行：

$$[A_{(1)}, \cdots, A_{(i)}, \cdots, A_{(j)}, \cdots, A_{(m)}] \xrightarrow{F_{ij}(\lambda)} [A_{(1)}, \cdots, A_{(i)}, \cdots, A_{(j)} + \lambda A_{(i)}, \cdots, A_{(m)}]$$

初等行变换(III)：将第i行乘以R中的可逆元 $\lambda \in R$.

$$[A_{(1)}, \cdots, A_{(i)}, \cdots, A_{(m)}] \xrightarrow{F_i(\lambda)} [A_{(1)}, \cdots, \lambda A_{(i)}, \cdots, A_{(m)}].$$

对应的初等列变换可类似定义.

□

可同样定义初等矩阵（见定义4.1.1），命题4.1.4对 $M_{m\times n}(R)$ 中的矩阵仍然成立。即，对 A 的行（或列）做初等变换等价于对 A 左乘（或右乘）相应的初等矩阵。同时，推论4.1.3的(1)和(2)对系数在 R 中的矩阵仍然成立。特别，初等矩阵都是可逆矩阵，初等变换(II)不改变行列式|A|的值等仍然成立。矩阵分块的技巧对 $M_{m\times n}(R)$ 中矩阵仍然适用，特别，引理4.4.1与引理4.4.2对 $M_n(R)$ 中矩阵仍然成立（R是交换环）.

定理4.5.2［行列式乘法定理］：设 $A = (a_{ij})_{n\times n}$，$B = (b_{ij})_{n\times n} \in M_n(R)$.
则 $|A\cdot B| = |A|\cdot|B|$.

证明：由引理4.4.1 可得 $\begin{vmatrix} A & 0 \\ -I_n & B \end{vmatrix} = |A|\cdot|B|$. 另一方面，以 A左乘第二分块行加到第一分块行 可得

$$\begin{vmatrix} A & 0 \\ -I_n & B \end{vmatrix} = \begin{vmatrix} 0 & AB \\ -I_n, & B \end{vmatrix} = (-1)^{n^2} |AB|\cdot|-I_n| = (-1)^{n^2+n} |A\cdot B| = |A\cdot B|.$$

故 $|A\cdot B| = |A|\cdot|B|$ □.

推论4.5.3：设 R 是交换环，$A=(a_{ij})_{n\times n} \in M_n(R)$. 则

A 是可逆矩阵 $\Longleftrightarrow$ $|A| \in R$ 是可逆元.

证明：若 A 可逆，则存在 $B \in M_n(R)$ 使 $A\cdot B = B\cdot A = I_n$. 从而 $|A|\cdot|B| = |B|\cdot|A| = 1 \in R$. 所以 $|A| \in R$ 可逆. 若 $|A|$ 在 R 中可逆，则由推论4.5.2，A 可逆，且 $A^{-1} = |A|^{-1}\cdot A^*$. □.

例4.5.1. 设 $R=\mathbb{Z}$ 是整数环，则 $A \in M_n(\mathbb{Z})$ 可逆的充要条件是：$|A| = \pm 1$. □

例4.5.2. 设 $R=K[x]$ 是域K上的多项式环. 则 $A(x)=(a_{ij}(x))_{n\times n}$ 在 $M_n(K[x])$ 中可逆的充要条件是：$|A(x)|$ 是K中的非零常数. □

定义4.5.7：设 $A, B \in M_{m\times n}(R)$，A与B 称为初等等价，如果 B可由A经初等行变换和初等列变换得到，记为 $A \backsim B$. □

引理4.5.1. $M_{m\times n}(R)$ 中矩阵的"初等等价"是 $M_{m\times n}(R)$上的 等价关系，即：
(1) 若 $A \backsim B$，则 $B \backsim A$. (2) $\forall A \in M_n(R)$，则 $A \backsim A$.
(3) 若 $A \backsim B$, $B \backsim C$，则 $A \backsim C$.

证明：若初等变换是可逆变换 故结论显然. □

我们称 $M_{m\times n}(K[x])$ 中的矩阵 $A(x)=$ ~~(…)~~ $(a_{ij}(x))_{m\times n}$ 为多项式矩阵. 下面我们证明: 多项式矩阵总是初等等价于一个对角型矩阵.

引理4.5.2: 设 $A(x)=(a_{ij}(x))_{m\times n}$ 非零, 则 $A(x)$ 初等等价于

$$B(x)=(b_{ij}(x))_{m\times n}$$

满足: (1) $b_{11}(x)$ 是首项系数为1的多项式, (2) $b_{11}(x)$ 整除每个 $b_{ij}(x)$.

证明: 对任意非零多项式矩阵 $A(x)=(a_{ij}(x))_{m\times n}$ 定义:

$$d(A(x))=\min_{i,j}\{\deg a_{ij}(x)\mid a_{ij}(x)\neq 0\}.$$

当 $d(A(x))=0$ 时, 引理显然成立。若 $d(A(x))>0$, 可设引理对所有满足 $d(B(x))<d(A(x))$ 的 $B(x)$ 成立。

通过初等变换, 不妨设 $a_{11}(x)\neq 0$ (首项系数为1), 且

$$\deg a_{11}(x)=d(A(x)).$$

如果存在 $a_{1j_0}(x)$ 使 $a_{11}(x)\nmid a_{1j_0}(x)$, 则 $a_{1j_0}(x)=q(x)a_{11}(x)+\bar{a}_{1j_0}(x)$ 其中 $\bar{a}_{1j_0}(x)\neq 0$ 但 $\deg\bar{a}_{1j_0}(x)<\deg a_{11}(x)=d(A(x))$. 将 $A(x)$ 的第一列乘 $(-q(x))$ 加到第 j_0 列 得到 $\bar{A}(x)=(\bar{a}_{ij}(x))$, 且 $d(\bar{A}(x))\le\deg\bar{a}_{1j_0}(x)<d(A(x))$. 由归纳假设, $\bar{A}(x)$ 等价于引理中的 $B(x)$. 故引理对 $A(x)$ 成立。所以可设 $a_{11}(x)\mid a_{1j}(x)$ $(1\le j\le n)$. 同理可设 $a_{11}(x)\mid a_{i1}(x)$ $(1\le i\le m)$. 因此, 通过第四类初等变换, 可设 $A(x)$ 初等等价于

$$\bar{A}(x)=\begin{pmatrix} a_{11}(x) & 0 & \cdots & 0 \\ 0 & \bar{a}_{22}(x) & \cdots & \bar{a}_{2n}(x) \\ \vdots & \vdots & & \\ 0 & \bar{a}_{m2}(x) & \cdots & \bar{a}_{mn}(x) \end{pmatrix}_{m\times n}.$$

如果存在 $\bar{a}_{ij}(x)$ 使 $a_{11}(x)\nmid\bar{a}_{ij}(x)$, 则 $\bar{a}_{ij}(x)=\bar{q}(x)\cdot a_{11}(x)+\bar{b}_{ij}(x)$ 其中 $\bar{b}_{ij}(x)\neq 0$ 且 $\deg\bar{b}_{ij}(x)<\deg a_{11}(x)=d(A(x))$。将 $\bar{A}(x)$ 的第 i 行加

~~到第一行, 再将 $\bar{A}'(x)$ 的第一列乘 … 加到第 … 列, 得:~~

到第1行得 $\bar{A}'(x)$，再将 $\bar{A}'(x)$ 的第1列乘 $(-q(x))$ 加到第j列 得 $\bar{B}(x)=(\bar{b}_{ij}(x))$，则 $d(\bar{B}(x))\le \deg \bar{b}_{ij}(x) < \deg a_{11}(x) = d(A(x))$。由归纳假设，$\bar{B}(x)$ 初等等价于满足引理条件的 $B(x)$，从而 $A(x)$ 初等等价于 $B(x)$。

□

推论4.5.4. 设 $A(x)=(a_{ij}(x))_{m\times n}\ne 0$，则

$$A(x)\sim B(x)=\begin{pmatrix} b_{11}(x) & \cdots & & 0 & \cdots & 0 \\ \vdots & \ddots & & \vdots & & \vdots \\ 0 & & b_{rr}(x) & 0 & \cdots & 0 \\ \vdots & & & & & \\ 0 & \cdots & 0 & 0 & \cdots & 0 \end{pmatrix}_{m\times n}$$

其中 $b_{ii}(x)\mid b_{i+1\,i+1}(x)$ 是首项系数1的多项式。

□

推论4.5.5: $A(x)\in M_n(K[x])$ 可逆，则 $A(x)$ 可写成初等矩阵的乘积。

证明：由推论4.5.4，$A(x)\sim \mathrm{diag}(b_{11}(x),\cdots,b_{rr}(x),0\cdots0)$。但 $A(x)$ 可逆，$|A(x)|$ 在 $K[x]$ 中可逆，从而 $r=n$ 且 $b_{ii}(x)$ 是零次多项式。即 $\mathrm{diag}(b_{11}(x),\cdots,b_{nn}(x))=I_n$。

□

推论4.5.6，设 $A\in M_n(K)$，则存在可逆矩阵 $P(x),Q(x)\in M_n(K[x])$ 使

$$P(x)(x\cdot I_n-A)Q(x)=\begin{pmatrix} 1 & & & & 0 \\ & \ddots & & & \\ & & 1 & & \\ & & & d_1(x) & \\ 0 & & & & \ddots \\ & & & & & d_s(x) \end{pmatrix}_{n\times n}$$

其中 $d_i(x)\mid d_{i+1}(x)$，$d_i(x)$ 是次数大于零的首项系数为1的多项式。

证明：由推论4.5.4，存在可逆矩阵 $P(x),Q(x)\in M_n(K[x])$ 使

$$P(x)(x\cdot I_n-A)Q(x)=\begin{pmatrix} b_{11}(x) & & & & \\ & \ddots & & 0 & \\ & & b_{rr}(x) & & \\ & 0 & & 0 & \\ & & & & \ddots \\ & & & & & 0 \end{pmatrix}_{n\times n}$$

且 $b_{ii}(x)$ 是首项系数1的多项式 且 $b_{ii}(x)\mid b_{i+1\,i+1}(x)$。

但 $|x\cdot I_n-A|$ 是 n 次非零多项式，故 $r=n$. 存在 s 使 $b_{11}(x)=\cdots=b_{n-s,n-s}(x)=1$. 令 $d_i(x)=b_{n-s+i,n-s+i}(x)$ 即可. □

习题 4.5

1. 设 $A=(a_{ij})_{m\times n}\in M_{m\times n}(\mathbb{Z})$，$\mathbb{Z}$ 是整数环，证明：A 初等等价于 $B=(b_{ij})_{m\times n}$ 满足：$b_{11}\mid b_{ij}$.

2. 证明：$M_{m\times n}(\mathbb{Z})$ 任意非零矩阵都初等等价于

$$\begin{pmatrix} b_{11} & 0 & \cdots & 0 & 0 & \cdots & 0 \\ 0 & b_{22} & & & \vdots & & \vdots \\ \vdots & \vdots & \ddots & b_{rr} & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & & \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & 0 & 0 & \cdots & 0 \end{pmatrix}_{m\times n}$$

其中 $b_{ii}\mid b_{i+1,i+1}$ 且 $b_{ii}>0$.

3. $M_n(\mathbb{Z})$ 中可逆矩阵可以写成初等矩阵的乘积.

4. 通过初等变换将多项式矩阵

$$x\cdot I_n-A=\begin{pmatrix} x-3 & -2 & 3 \\ -4 & x-10 & 12 \\ -3 & -6 & x+7 \end{pmatrix}$$

化为推论4.5.6中的对角型矩阵.

5. 设

$$A=\begin{pmatrix} 1 & 1 & -1 \\ 3 & -2 & 0 \\ -1 & 1 & -1 \end{pmatrix}.$$

试计算行列式 $|x\cdot I_3-A|$.

6. 设 $A\in M_n(K)$. 证明：(1) $|xI_n-A|$ 是首项系数为1的 n 次多项式. $|xI_n-A|=x^n+a_1x^{n-1}+\cdots+a_{n-1}x+a_n$，(2) $a_1=-tr(A)$.

第五章 多项式

在第一章的第五节，我们已经介绍了系数在子环 $R\subset\mathbb{C}$ 中的多项式。这些基本概念可以推广至系数在任意域 K 中的多项式（甚至系数在任意环中的多项式），在此不再赘述。本章的内容包括：多项式的不可约分解，多项式函数，对称多项式，两个多项式的结式。

§5.1. 不可约分解

设 K 是一个域，$K[x_1,\dots,x_n]$ 表示系数在 K 中的所有多项式集合。$f(x_1,\cdots,x_n)\in K[x_1,\cdots,x_n]$ 称为 $K[x_1,\cdots,x_n]$ 中的可逆元，如果存在 $g(x_1,\cdots,x_n)\in K[x_1,\cdots,x_n]$ 使得 $f(x_1,\cdots,x_n)\cdot g(x_1,\cdots,x_n)=1$。显然，$f(x_1,\cdots,x_n)\in K[x_1,\cdots,x_n]$ 是可逆元当且仅当 $f(x_1,\cdots,x_n)$ 是非零的零次多项式（即 K 中的非零元）。

定义5.1.1. 设 $f(x_1,\cdots,x_n)$，$g(x_1,\cdots,x_n)\in K[x_1,\cdots,x_n]$。如果存在 $h(x_1,\cdots,x_n)\in K[x_1,\cdots,x_n]$ 使得 $g(x_1,\cdots,x_n)=f(x_1,\cdots,x_n)\cdot h(x_1,\cdots,x_n)$，则称 $f(x_1,\cdots,x_n)$ 整除 $g(x_1,\cdots,x_n)$，记为 $f(x_1,\cdots,x_n)\mid g(x_1,\cdots,x_n)$，其中 $f(x_1,\cdots,x_n)$，$h(x_1,\cdots,x_n)$ 称为 $g(x_1,\cdots,x_n)$ 的因子（或因式）。 □

对任意多项式 $f(x_1,\cdots,x_n)\in K[x_1,\cdots,x_n]$，显然 $a\cdot f(x_1,\cdots,x_n)$（其中 $a\in K$ 非零）和 K 中的非零元都是 $f(x_1,\cdots,x_n)$ 的因子。它们称为 $f(x_1,\cdots,x_n)$ 的平凡因子。

定义5.1.2. $f(x_1,\cdots,x_n)\in K[x_1,\cdots,x_n]$ 称为不可约多项式，如果 $f(x_1,\cdots,x_n)$ 的次数大于零，且没有非平凡的因子。 □

例5.1.1. $f(x)=x^2-2$ 是 $\mathbb{Q}[x]$ 中的不可约多项式，但不是 $\mathbb{R}[x]$ 中的不可约多项式；x^2+1 是 $\mathbb{R}[x]$ 中不可约多项式但不是 $\mathbb{C}[x]$ 中不可约多项式。 □

例5.1.2（代数基本定理）. 若 $f(x)\in\mathbb{C}[x]$ 是 $\mathbb{C}[x]$ 中的不可约多项式，则 $f(x)=ax+b$ （$a\in\mathbb{C}$ 非零）. □

例5.1.3：设 $f(x)\in\mathbb{R}[x]$ 是首项系数为1（简称首1）的实系数多项式. 则 $f(x)$ 在 $\mathbb{R}[x]$ 中不可约当且仅当 $f(x)=x-a$ （$a\in\mathbb{R}$）或

$$f(x)=x^2+bx+c\quad,\ b^2-4c<0.$$

□

例5.1.4：设 x_{ij} （$1\le i,j\le n$）是域 K 上的不定元. 则多项式 $\det(x_{ij})_{n\times n}=\sum_{\pi\in S_n}\varepsilon_\pi x_{1\pi(1)}\cdots x_{n\pi(n)}$ 在多项式环 $K[x_{ij}\mid 1\le i,j\le n]$ 中不可约.

证明：设 $X=(x_{ij})_{n\times n}$ 表 (i,j) 位置元素为 x_{ij} 的矩阵. 则行列式 $\det(X)$ 是 n 次齐多项式，且是关于任一行元素 $x_{i1},x_{i2},\cdots,x_{in}$（或任一列 $x_{1j},x_{2j},\cdots,x_{nj}$）的一次齐次多项式. 所以，若 $\det(X)=f\cdot g$, f 的某单项式包含 x_{ij}，则 f 必是关于 $x_{i1},x_{i2},\cdots,x_{in}$ 的一次齐次多项式且 g 与 $x_{i1},\cdots,x_{in}$ 无关.（对 $x_{1j},x_{2j},\cdots,x_{nj}$ 上述结论亦成立）. 即：若 f 包含变量 x_{ij}，则 g 与 $x_{i1},x_{i2},\cdots,x_{in}$, $x_{1j},x_{2j},\cdots,x_{nj}$ 无关，且 f 亦含 $x_{i1},x_{i2},\cdots x_{in}$. 分别应用上述结论于 $x_{11},x_{12},\cdots,x_{1n}$ 可得：g 与 $(x_{11},x_{21},\cdots,x_{n1})$, $(x_{12},x_{22},\cdots,x_{n2})$, $\cdots$ $(x_{1n},\cdots,x_{nn})$ 无关. 所以 g 为常数 □

引理 5.1.1：设 $f(x_1,\cdots,x_n)\in K[x_1,\cdots,x_n]$，则 $f(x_1,\cdots,x_n)$ 可以写成有限个不可约多项式的乘积，即：

$$f(x_1,\cdots,x_n)=f_1(x_1,\cdots,x_n)\cdots f_s(x_1,\cdots,x_n) \qquad (1.1)$$

其中 $f_i(x_1,\cdots,x_n)\in K[x_1,\cdots,x_n]$ $(1\le i\le s)$ 是不可约多项式。

证明：若 $f(x_1,\cdots,x_n)$ 不可约，则 $s=1$，$f_1(x_1,\cdots,x_n)=f(x_1,\cdots,x_n)$。若 $f(x_1,\cdots,x_n)$ 可约，则 $f(x_1,\cdots,x_n)=g(x_1,\cdots,x_n)\cdot h(x_1,\cdots,x_n)$ 其中 $\deg(g)>0$，$\deg(h)>0$。从而 $\deg(g)<\deg(f)$，$\deg(h)<\deg(f)$。对于 $\deg(f)$ 应用归纳法，可得引理。 □

定义 5.1.3. 上述引理中的 (1.1) 称为 $f(x_1,\cdots,x_n)$ 的一个不可约分解。如果对 $f(x_1,\cdots,x_n)$ 的任意两个不可约分解

$$f(x_1,\cdots,x_n)=f_1(x_1,\cdots,x_n)\cdots f_s(x_1,\cdots,x_n)=g_1(x_1,\cdots,x_n)\cdots g_t(x_1,\cdots,x_n)$$

必有 $s=t$，且经过适当调整不可约因子次序，必有

$$f_i(x_1,\cdots,x_n)=c_i\,g_i(x_1,\cdots,x_n),\quad c_i\in K^*:=K\setminus\{0\}.$$

则称 $f(x_1,\cdots,x_n)$ 有唯一不可约分解。 □

是否 $K[x_1,\cdots,x_n]$ 中的每个多项式都有唯一不可约分解？~~定义~~这问题实际上等价于最大公因子的存在性。

定义 5.1.4：设 $f,g\in K[x_1,\cdots,x_n]$，$d(x_1,\cdots,x_n)\in K[x_1,\cdots,x_n]$ 称为 f,g 的一个最大公因子，如果它满足：

(1) $d(x_1,\cdots,x_n)\mid f(x_1,\cdots,x_n)$，$d(x_1,\cdots,x_n)\mid g(x_1,\cdots,x_n)$。

(2) 若 $h(x_1,\cdots,x_n)$ 是 $f(x_1,\cdots,x_n)$ 与 $g(x_1,\cdots,x_n)$ 任一公因子，则 $h(x_1,\cdots,x_n)\mid d(x_1,\cdots,x_n)$。（即：$h\mid f$，$h\mid g \Rightarrow h\mid d$。） □

引理5.1.2：若 $K[x_1,\cdots,x_n]$ 中每个次数大于零的多项式均有唯一不可约分解，则 $K[x_1,\cdots,x_n]$ 中任意两个多项式均有最大公因子

证明：$\forall\ f,g\in K[x_1,\cdots,x_n]$ 无妨设 $f\neq 0, g\neq 0$，且 $\deg(f)>0$ $\deg(g)>0$（否则容易验证 f,g 有最大公因子）。令

$$f=p_1^{n_1}\cdots p_s^{n_s}\ ,\ g=p_1^{m_1}\cdots p_s^{m_s}\ ,\quad n_i\geq 0,\ m_i\geq 0\ .$$

其中 $p_i\neq p_j$（若 $i\neq j$）是 $K[x_1,\cdots,x_n]$ 中的不可约多项式。则

$$d=p_1^{l_1}\cdots p_s^{l_s}\ ,\quad l_i=\min\{n_i,m_i\}.$$

是 f 和 g 的最大公因子。事实上，$d|f, d|g$。故只需证明：若 $h|f, h|g$，则 $h|d$ ~~[illegible]~~。

设 $h=q_1^{n_1'}\cdots q_t^{n_t'}$ 是 h 的不可约分解，利用 f 不可约分解的唯一性，可知 $q_1,\cdots,q_t$ 必在 $p_1,\cdots,p_s$ 中出现，故可设

$$h=p_1^{k_1}\cdots p_s^{k_s}\ ,\quad k_i\geq 0$$

由于 $h|f$，且 f 的不可约分解唯一，故 $k_i\leq n_i$。同理 $k_i\leq m_i$。所以 $h|d$。 □.

练习：若 $d,d'\in K[x_1,\cdots,x_n]$ 都是 f 和 g 的最大公因子，则 $d'=cd$，其中 $c\in K$。

定理5.1.1：设 K 是一个域，则关于 $K[x_1,\cdots,x_n]$ 的下述断言等价。

(1) $K[x_1,\cdots,x_n]$ 中每个次数大于零的多项式均有唯一不可约分解。

(2) $K[x_1,\cdots,x_n]$ 中任意两个多项式均有最大公因子。

(3) 设 $p(x_1,\cdots,x_n)\in K[x_1,\cdots,x_n]$ 不可约，若 $p\,|\,f\cdot g$，则或者 $p|f$ 或者 $p|g$.

证明：我们将证明 (1) ⇒ (2) ⇒ (3) ⇒ (1)。其中 (1) ⇒ (2) 由引理 5.1.2 可得。下面先证 (2) ⇒ (3)。若 $p = p(x_1, \cdots, x_n)$ 不可约且 $p \mid f \cdot g$（其中 $f = f(x_1, \cdots, x_n)$，$g = g(x_1, \cdots, x_n)$）。令 d_1 是 p 和 f 的最大公因子，由 $d_1 \mid p$ 及 p 不可约可知 $d_1 \in K^*$ 或 $d_1 = cp$ $(c \in K^*)$。如果 $d_1 = c \cdot p$ $(c \in K^*)$，则 $p \mid f$。否则可设 $d_1 = 1$，此时可证 g 是 $p \cdot g$ 和 $f \cdot g$ 的一个最大公因子，由此可得 $p \mid g$（因为 p 是 $p \cdot g$ 和 $f \cdot g$ 的一个公因子）。设 d 是 $p \cdot g$ 和 $f \cdot g$ 的一个最大公因子，则 $g \mid d$。令 $d = g \cdot h$，$p \cdot g = d \cdot h_1$，$f \cdot g = d \cdot h_2$，可得（将 $d = g \cdot h$ 代入）

$$p \cdot g = g \cdot h \cdot h_1, \quad f \cdot g = g \cdot h \cdot h_2 \text{。}(\Rightarrow h \text{ 是 } p \text{ 和 } f \text{ 的公因子})$$

从而 $h \mid d_1$。但 $d_1 = 1$，故 $h \in K^*$，g 是 $p \cdot g$ 和 $f \cdot g$ 的一个最大公因子。最后证明 (3) ⇒ (1)。设 $f(x_1, \cdots, x_n) \in K[x_1, \cdots, x_n]$ 的次数大于 0。如果 $f = p_1 \cdots p_s = q_1 \cdots q_t$ 是 f 在 $K[x_1, \cdots, x_n]$ 中的两个不可约分解。由 (2)，p_1 必整除 $q_1, \cdots, q_t$ 中一个。无妨设 $p_1 \mid q_1$。则 $q_1 = c_1 p_1$ $(c_1 \in K^*)$。消去 p_1，q_1 可得

$$p_2 \cdots p_s = c_1 q_2 \cdots q_t .$$

则 p_2 必整除 $q_2, \cdots, q_t$ 中一个，无妨设 $p_2 \mid q_2$，得 $q_2 = c_2 p_2$ $(c_2 \in K^*)$。有限步后可得结论 (1)。（可对不可约因子个数采用归纳法）□.

<u>定义 5.1.5</u>：设 $f_1, \cdots, f_s \in K[x_1, \cdots, x_n]$，多项式 $d \in K[x_1, \cdots, x_n]$ 称为 $f_1, \cdots, f_s$ 的一个<u>最大公因子</u>，如果

(1) $d \mid f_i$ $(1 \le i \le s)$。(2) 若 $h \mid f_i$ $(1 \le i \le s)$，则 $h \mid d$。

如果 $f_1, \cdots, f_s$ 的最大公因子是<u>非零常数</u> $c \in K^*$（通常可设 $c=1$），则称 $f_1, \cdots, f_s$ 互素，记为 $(f_1, \cdots, f_s) = 1$。 □.

⑥

显然，如果 $k[x_1,\cdots,x_n]$ 中任意两个多项式存在最大公因子，则 $k[x_1,\cdots,x_n]$ 中任意有限个多项式存在最大公因子。本节的主要定理是

定理5.1.2：$k[x_1,\cdots,x_n]$ 中的每个正次数多项式均有唯一不可约分解。 □

仅证明 $n=1$ 的情形。$n>1$ 的情形可由归纳法得到，但证明需要高斯引理及分式域概念。

定理5.1.3：设 k 是一个域，$f(x), g(x) \in k[x]$。则存在 $d(x)\in k[x]$ 满足条件：(1) $d(x)|f(x)$，$d(x)|g(x)$.

(2) 存在 $u(x), v(x) \in k[x]$ 使 $d(x) = u(x)f(x) + v(x)g(x)$.

特别地，$d(x)$ 是 $f(x)$ 和 $g(x)$ 的一个最大公因子.

证明：令 $I = \{ u(x)f(x) + v(x)g(x) \mid \forall u(x), v(x) \in k[x]\} \subset k[x]$.

$d(x)\in I$ 是 I 中次数最小的非零多项式。则 $d(x)$ 整除 I 中的每一个多项式。事实上，$\forall$ $h(x) \in I$，由带余除法，存在 $q(x), r(x) \in k[x]$ 使得

$h(x) = q(x)d(x) + r(x)$，其中 $r(x)=0$ 或 $\deg r(x) < \deg d(x)$.

如果 $r(x)\neq 0$，则 $\deg(r(x)) < \deg d(x)$. 但 $r(x) = h(x) - q(x)\cdot d(x) \in I$. 与 $d(x)$ 的选取矛盾。故 $r(x)=0$ 即 $d(x)\,|\,h(x)$, $(\forall h(x) \in I)$.

特别地，$d(x)|f(x)$, $d(x)|g(x)$. 且 $d(x) = u(x)f(x) + v(x)g(x)$。

若 $d_1(x)$ 是 $f(x), g(x)$ 的公因子，由 $d(x) = u(x)f(x) + v(x)g(x)$ 显然可得 $d_1(x)\,|\,d(x)$. 故 $d(x)$ 是 $f(x)$ 和 $g(x)$ 的一个最大公因子. □

推论5.1.1. 若 $f(x), g(x) \in k[x]$ 互素，则存在 $u(x), v(x) \in k[x]$ 使 $u(x)f(x) + v(x)g(x) = 1$. □

对于给定的 $f(x), g(x) \in K[x]$，带余除法 $f(x)=q(x)g(x)+r(x)$ 中的 $q(x), r(x)$ 是可计算的。所以下面的算法告诉我们定理5.1.3中的 $d(x)$ 和 $u(x), v(x)$ 都是可计算的(计算机实现)。

例5.1.5 (辗转相除法求最大公因子)。设 $f(x), g(x) \in K[x]$ 非零。做带余除法 $f(x)=q_1(x)g(x)+r_1(x)$。若 $r_1(x)=0$，则 $f(x), g(x)$ 的最大公因子是 $d(x)=g(x)$。若 $r_1(x)\neq 0$，做带余除法

$$
\begin{aligned}
f(x) &= q_1(x)g(x)+r_1(x) & \deg r_1(x) &< \deg g(x) \\
g(x) &= q_2(x)r_1(x)+r_2(x) & \deg r_2(x) &< \deg r_1(x) \\
&\cdots & &\cdots \\
r_{n-2}(x) &= q_n(x)r_{n-1}(x)+r_n(x). & \deg r_n(x) &< \deg r_{n-1}(x) \\
r_{n-1}(x) &= q_{n+1}(x)r_n(x). & r_{n+1}(x) &= 0.
\end{aligned}
$$

则 $d(x)=r_n(x)$ 是 $f(x)$ 和 $g(x)$ 的一个最大公因子。令 $u_1(x)=1, v_1(x)=-q_1(x)$, $u_2(x)=-q_2(x)$, $v_2(x)=1+q_1(x)q_2(x)$, $\cdots$, $u_i(x)=u_{i-2}(x)-q_i(x)u_{i-1}(x)$ $(i\geq 3)$。$v_i(x)=v_{i-2}(x)-q_i(x)v_{i-1}(x)$ $(i\geq 3)$。则 $r_i(x)=u_i(x)f(x)+v_i(x)g(x)$。

特别地，$d(x)=r_n(x)=u_n(x)f(x)+v_n(x)g(x)$。

证明：从最后一个等式往前推可知：$r_n(x) \mid r_{n-1}(x), r_{n-2}(x), \cdots, r_2(x), r_1(x)$，从而 $r_n(x) \mid f(x), g(x)$。即 $r_n(x)$ 是 $f(x), g(x)$ 的公因子。若 $h(x) \mid f(x), g(x)$，则 $h(x) \mid r_1(x)$, $h(x) \mid r_2(x)$, $\cdots$, $h(x) \mid r_n(x)$。故 $r_n(x)$ 是 $f(x), g(x)$ 的最大公因子。$r_i(x)=u_i(x)f(x)+v_i(x)g(x)$ 可由第一个等式开始往后代入得到。 $\square$

上述算法亦称欧氏算法，它对将来抽象代数中的欧氏环也成立。特别，若 $f(x), g(x) \in K[x]$ 互素，上述算法可求出 $u(x), v(x) \in K[x]$ 使 $u(x)f(x)+v(x)g(x)=1$。

习题 5-1

1. 设 $f(x_1,\cdots,x_n)$, $g(x_1,\cdots,x_n)$, $h(x_1,\cdots,x_n)\in K[x_1,\cdots,x_n]$ 使
$$f(x_1,\cdots,x_n)=g(x_1,\cdots,x_n)\cdot h(x_1,\cdots,x_n)$$
证明：f 是齐次多项式当且仅当 $g(x_1,\cdots,x_n)$, $h(x_1,\cdots,x_n)$ 是齐次多项式.

2. 设 $f(x_1,\cdots,x_n)$, $g(x_1,\cdots,x_n)\in K[x_1,\cdots,x_n]$, 若 $f\mid g$, $g\mid f$, 试证明存在非零常数 $c\in K^*$ 使 $f(x_1,\cdots,x_n)=c\,g(x_1,\cdots,x_n)$.

3. 设 $f(x_1,\cdots,x_n)\in K[x_1,\cdots,x_n]$ 是齐次多项式. 证明：
   (1) 若 f 在 $K[x_1,\cdots,x_n]$ 不可约，则 $f_i=f(x_1,\cdots,x_{i-1},1,x_{i+1},\cdots,x_n)$ 在 $K[x_1,\cdots,x_{i-1},x_{i+1},\cdots,x_n]$ 中不可约对任意 $1\le i\le n$ 成立.
   (2) 若 $x_i\nmid f(x_1,\cdots,x_n)$ 且 $f_i(x_1,\cdots,x_{i-1},x_{i+1},\cdots,x_n)$ 不可约，则 $f(x_1,\cdots,x_n)$ 在 $K[x_1,\cdots,x_n]$ 中不可约.

4. 用辗转相除法求多项式 $f(x)=x^5+x^4+x^3+x^2+x+1$ 和 $g(x)=x^4+3x^3+2x^2+x+1$ 的最大公因子 $d(x)$, 并求 $u(x)$, $v(x)$ 使得
$$d(x)=u(x)f(x)+v(x)g(x)$$

5. 设 $p(x)\in K[x]$ 是不可约多项式，$f(x)\in K[x]$. 证明：若 $p(x)\nmid f(x)$, 则 $p(x)$, $f(x)$ 互素 即 $(p(x),f(x))=1$.

6. 对任意 $f(x),g(x)\in K[x]$, 令 $(f(x),g(x))$ 表示 $f(x),g(x)$ 的一个最大公因子. 两个多项式 $f(x),g(x)$ 称为相伴（记为 $f(x)\sim g(x)$）如果存在 $c\in K^*$ 使得 $f(x)=c\,g(x)$. 试证明：
   (1) $(f(x),0)\sim f(x)$, $(f(x),g(x))\sim f(x)\Leftrightarrow f(x)\mid g(x)$
   (2) $(h(x)f(x),h(x)g(x))\sim h(x)\cdot(f(x),g(x))$.
   (3) $((f(x),g(x)),h(x))\sim(f(x),(g(x),h(x)))$.

7. 设 $f_1(x),\cdots,f_s(x)\in K[x]$, $d(x)$ 是它们的最大公因子. 证明存在 $u_1(x),\cdots,u_s(x)\in K[x]$ 使得

$$d(x)=u_1(x)f_1(x)+\cdots+u_s(x)f_s(x).$$

特别地, 若 $f_1(x),\cdots,f_s(x)$ 互素, 则存在 $u_1(x),\cdots,u_s(x)\in K[x]$ 使 $u_1(x)f_1(x)+\cdots+u_s(x)f_s(x)=1$.

8. 设 $f(x_1,\cdots,x_n)\in K[x_1,\cdots,x_n]$. 证明: $f(x_1,\cdots,x_n)$ 是 d 次齐次多项式的充分必要条件是 $f(x_1,\cdots,x_n)$ 满足:

$$f(tx_1,\cdots,tx_n)=t^d f(x_1,\cdots,x_n)$$

其中 t 是一个新的独立变元.

9. 对下列多项式对 $f(x)$, $g(x)$, 求 $q(x)$, $r(x)$ 使得 $f(x)=q(x)g(x)+r(x)$.

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

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

10. 设 $d(x)$ 是 $f(x)$, $g(x)\in K[x]$ 的最大公因子. 证明: 若 $g(x)\nmid f(x)$, 则

(1) 存在 $u(x),v(x)\in K[x]$ 满足 $\deg u(x)<\deg g(x)-\deg d(x)$, $\deg v(x)<\deg f(x)-\deg d(x)$ 使得 $d(x)=u(x)f(x)+v(x)g(x)$.

(2) 上述多项式 $u(x)$, $v(x)$ 是唯一的.

(提示: 存在 $w(x),h(x)\in K[x]$ 使 $d(x)=f(x)w(x)+g(x)h(x)$, 令 $u(x)$ 是 $w(x)$ 除多项式 $g(x)/d(x)$ 的余项, 则 $u(x)\neq 0$).

§5.2 多项式函数

设 K 是一个域，R ⊃ K 是一个包含 K 作为子环的环。R 不必是交换环，但我们总是假设

$$a\cdot u = u\cdot a,\quad \forall\, a\in K,\ u\in R.$$

若 $f(x) = a_n x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0 \in K[x]$ 是一个多项式，对任意 $u\in R$，令 $f(u)$ 表示

$$a_n\cdot u^n + a_{n-1}\cdot u^{n-1} + \cdots + a_1\cdot u + a_0 \in R.$$

则多项式 $f(x)$ 定义了一个 R 上函数 $f_R: R\to R$，$u\longmapsto f(u)$。称为由 $f(x)$ 诱导的多项式函数。令

$$R_{pol} = \{ f_R : R\to R \mid f(x)\in K[x] \}$$

表示由 $K[x]$ 中所有多项式诱导的多项式函数集合。在 R_{pol} 上可定义函数的加法和 R 上的乘法：$\forall\, f_R, g_R \in R_{pol}$.

$$f_R + g_R : R\to R,\quad u\longmapsto f(u)+g(u).$$

$$f_R\cdot g_R : R\to R,\quad u\longmapsto f(u)\cdot g(u).$$

则 R_{pol} 关于上述运算是一个交换环。

<u>引理 5.2.1</u>：映射 $K[x]\longrightarrow R_{pol}$，$f(x)\longmapsto f_R$，满足条件：

$$f(x)+g(x)\longmapsto f_R+g_R,\quad f(x)g(x)\longmapsto f_R\cdot g_R.$$

（这样的映射称为环同态）。当 K 是无限域时，映射

$$K[x]\longrightarrow R_{pol},\quad f(x)\longmapsto f_R.$$

对任意 R（满足 $a u = u\cdot a,\ \forall\, a\in K, u\in R$）是双射.

证明：由定义，映射 $K[x] \longrightarrow R_{pol}$，$f(x) \longmapsto f_R$，是一个满射。在同态条件：$a \cdot u = u \cdot a$，$(\forall a \in K, u \in R)$ 可直接验证：$f(x)+g(x) \longmapsto f_R + g_R$，$f(x) \cdot g(x) \longmapsto f_R \cdot g_R$.

（注意：若存在 $a \in K$，$u \in R$ 使得 $a \cdot u \neq u \cdot a$，则令 $f(x)=x$，$g(x)=a$（零次多项式）。则 $f(x) \cdot g(x) = x \cdot a = ax$，但

$$(f \cdot g)_R(u) = a \cdot u \neq f_R \cdot g_R(u) = f_R(u) \cdot g_R(u) = u \cdot a).$$

下面证明，映射 $K[x] \longrightarrow R_{pol}$，$f(x) \longmapsto f_R$，当 K 是无限域时，是单射。设 $f(x), g(x) \in K[x]$ 满足：$f_R = g_R$。则 $f(u) = g(u)$ $(\forall u \in R)$。特别，$f(u) = g(u)$ $(\forall u \in K)$。令 $h(x) = f(x) - g(x)$，则 $h(u) = 0$ $(\forall u \in K)$。因此，$h(x)$ 在 K 中有无穷多个根。从而 $h(x)$ 必为零多项式。故

$$f(x) = g(x)$$

□.

因此，当 K 是无限域时，$K[x]$ 中的多项式 $f(x)$ 均可看成 K 上的函数 $K \longrightarrow K$，$u \longmapsto f(u)$。且 $f(x) = g(x)$ 当且仅当它们作为 K 上的函数相等！但当 K 是有限域时，这样的结论不成立！

<u>例 5.2.1</u>（费马小定理）。设 p 是一个素数，$\mathbb{F}_p = \{\bar{0}, \bar{1}, \cdots, \overline{p-1}\}$ 是模 p 剩余类域，$f(x) = x^p - x \in \mathbb{F}_p[x]$，则

$$f(u) = 0, \quad \forall u \in \mathbb{F}_p.$$

证明：$\forall\ u\in \mathbb{F}_p$，无妨设 $u\neq\bar{0}$，则

$$\{\bar{1}\cdot u,\ \bar{2}\cdot u,\ \cdots,\ \overline{p-1}\cdot u\}=\{\bar{1},\ \bar{2},\ \cdots,\ \overline{p-1}\}.$$

从而 $(\bar{1}\cdot u)\cdot(\bar{2}\cdot u)\cdots\cdot(\overline{p-1}\cdot u)=\bar{1}\cdot\bar{2}\cdot\cdots\cdot\overline{p-1}$。由此可得 $(\bar{1}\cdot\bar{2}\cdot\cdots\cdot\overline{p-1})\cdot u^{p-1}=\bar{1}\cdot\bar{2}\cdot\cdots\cdot\overline{p-1}$。故 $u^{p-1}=\bar{1}$ 即 $u^p=u$。所以 $f(u)=0\ (\forall\ u\in\mathbb{F}_p)$ □.

引理5.2.1 的一个重要应用是 $R=\mathcal{L}(V)$，其中 V 是一个 n 维 K-向量空间。$\mathcal{L}(V)=\{\mathcal{A}: V\to V\}$ 是 V 上所有线性算子（即 $\mathcal{A}: V\to V$ 是线性映射）的集合，它在运算

$$\mathcal{A}+\mathcal{B}: V\to V,\quad \alpha\mapsto \mathcal{A}\alpha+\mathcal{B}\alpha,\quad (\mathcal{A}\alpha:=\mathcal{A}(\alpha))$$
$$\mathcal{A}\cdot\mathcal{B}: V\to V,\quad \alpha\mapsto \mathcal{A}\cdot(\mathcal{B}\cdot\alpha).$$

下是一环。事实上，$\forall\ \lambda\in K$，$\mathcal{A}\in\mathcal{L}(V)$，还可定义"数乘"：

$$\lambda\mathcal{A}: V\to V,\quad \alpha\mapsto\lambda\mathcal{A}\alpha.$$

使得 $\mathcal{L}(V)$ 同时是一个 K-向量空间且 $\mathcal{A}\cdot(\lambda\mathcal{B})=(\lambda\mathcal{A})\cdot\mathcal{B}=\lambda(\mathcal{A}\cdot\mathcal{B})$。此时我们称 $\mathcal{L}(V)$ 是一个 K-代数。

对任意 $a\in K$，映射 $V\to V,\ \alpha\mapsto a\alpha$，是一个线性算子（称为数乘算子）。显然，$\forall\ a,b\in K$，则 $a=b$ 的充要条件是它们定义的数乘算子相等。另外注意，若 $a\in K$，$\mathcal{A}\in\mathcal{L}(V)$，则数乘 $a\mathcal{A}$ 与算子乘法 $a\cdot\mathcal{A}$ 相等。所以我们通常将 K 的元素 $a\in K$ 与它诱导的算子 $V\to V$，$\alpha\mapsto a\alpha$，等同，从而 $K\subset\mathcal{L}(V)$ 是 $\mathcal{L}(V)$ 的子环。

且 $a\cdot\mathcal{A}=\mathcal{A}\cdot a$ $(\forall a\in K,\ \mathcal{A}\in\mathcal{L}(V))$. 若

$$f(x)=a_nx^n+a_{n-1}x^{n-1}+\cdots+a_1x+a_0\in K[x],$$

$\mathcal{A}\in\mathcal{L}(V)$. 定义 $f(\mathcal{A})=a_n\mathcal{A}^n+a_{n-1}\mathcal{A}^{n-1}+\cdots+a_1\mathcal{A}+a_0\in\mathcal{L}(V)$,

$$\alpha\longmapsto f(\mathcal{A})\alpha:=a_n\mathcal{A}^n\alpha+a_{n-1}\mathcal{A}^{n-1}\alpha+\cdots+a_1\mathcal{A}\alpha+a_0\alpha\in V.$$

推论5.2.1. $\forall f(x),g(x)\in K[x]$, 有 $(f+g)_{\mathcal{L}(V)}=f_{\mathcal{L}(V)}+g_{\mathcal{L}(V)}$, $(f\cdot g)_{\mathcal{L}(V)}=f_{\mathcal{L}(V)}\cdot g_{\mathcal{L}(V)}$。即: 若 $f(x)+g(x)=h_1(x)$, $f(x)\cdot g(x)=h_2(x)$, 则 $\forall\mathcal{A}\in\mathcal{L}(V)$, 有

$$h_1(\mathcal{A})=f(\mathcal{A})+g(\mathcal{A}),$$
$$h_2(\mathcal{A})=f(\mathcal{A})\cdot g(\mathcal{A}).$$

□.

定义5.2.1 [取值映射]. 设 $\mathcal{A}\in\mathcal{L}(V)$, 则映射

$$K[x]\longrightarrow\mathcal{L}(V),\quad f(x)\longmapsto f(\mathcal{A}),$$

称为多项式在 $\mathcal{A}\in\mathcal{L}(V)$ 的取值映射。取值映射的像记为 $K[\mathcal{A}]:=\{f(\mathcal{A})\mid\forall f(x)\in K[x]\}$. □.

推论5.2.2. $\forall\mathcal{A}\in\mathcal{L}(V)$, 则取值映射

$$K[x]\longrightarrow\mathcal{L}(V),\quad f(x)\longmapsto f(\mathcal{A}),$$

是一个环同态, 即: $f(x)+g(x)\longmapsto f(\mathcal{A})+g(\mathcal{A})$, $f(x)\cdot g(x)\longmapsto f(\mathcal{A})\cdot g(\mathcal{A})$ 且它的像 $K[\mathcal{A}]\subset\mathcal{L}(V)$ 是 $\mathcal{L}(V)$ 的交换子环. □.

显然, $K[\mathcal{A}]\subset\mathcal{L}(V)$ 同时也是一个子空间, 称为交换子代数.

若 $K\subset\mathbb{C}$ 是复数域 $\mathbb{C}$ 的子域，由代数基本定理可知 $K[x]$ 中的每个非常数多项式 $f(x)\in K[x]$（即：$\deg f(x)>0$）在 $\mathbb{C}$ 均有一个根。对于任意域 K（比如，有限域 $\mathbb{F}_q$），我们将不予证明地应用下述定理。

定理 5.2.1（代数闭包的存在性）. 设 K 是一个域，则存在一个域 $\bar{K}$ 使得 $K\subset\bar{K}$ 是子域，且满足：

(1) 若 $f(x)\in\bar{K}[x]$ 是次数零的多项式，则存在 $\alpha\in\bar{K}$ 使 $f(\alpha)=0$

(2) 对任意 $\alpha\in\bar{K}$，存在非零 $f(x)\in K[x]$ 使 $f(\alpha)=0$.

上述域 $\bar{K}$ 称为 K 的代数闭包. □

例 5.2.2：设 $\mathbb{F}_q$ 是 q 元域，则它的代数闭包 $\bar{\mathbb{F}}_q$ 必为无限域.

证明：由定义，$\bar{\mathbb{F}}_q[x]$ 中的不可约多项式必为一次多项式，所以，若 $|\bar{\mathbb{F}}_q|<+\infty$（即 $\bar{\mathbb{F}}_q$ 是有限域），则 $\bar{\mathbb{F}}_q[x]$ 中仅有有限个不可约多项式。故我们只需证明：对任意域 K，$K[x]$ 中有无限多个不可约多项式（若 K 是无限域，这是显然的，因为 $x-a$ $(a\in K)$ 都是不可约多项式）。否则，设 $p_1(x),\dots,p_s(x)$ 是 $K[x]$ 中的全部不可约多项式，则 $f(x)=p_1(x)\cdots p_s(x)+1\in K[x]$ 必为 $K[x]$ 中的不可约多项式且 $f(x)\neq p_1(x),\dots,p_s(x)$，矛盾 □.

定理 5.2.2. 设 K 是一个域，$f(x)\in K[x]$ 是首项系数为 1 的多项式. 若 $\deg f(x)=n>0$，则存在 $\alpha_1,\dots,\alpha_n\in\bar{K}$（$\alpha_1,\dots,\alpha_n$ 可以有相同元）使得 $f(x)=(x-\alpha_1)\cdots(x-\alpha_n)$. 特别，$f(x)$ 在 $\bar{K}$ 中最多有 n 个根.

证明：将 $K[x]$ 中的所有多项式均看成 $\bar{K}[x]$ 中多项式。对 $\deg f(x)$ 作归纳法，证明若 $f(x)\in\bar{K}[x]$ 是 n 次首系1多项式，则 $f(x)=(x-\alpha_1)\cdots(x-\alpha_n)$，$\alpha_i\in\bar{K}$。 $n=1$ 时显然，若 $\deg f(x)=n>1$，由 $\bar{K}$ 的定义，存在 $\alpha_1\in\bar{K}$ 使 $f(\alpha_1)=0$。对 $f(x)$，$x-\alpha_1$ 在 $\bar{K}[x]$ 作用带余除法，可得 $f_1(x)\in\bar{K}[x]$ 使 $f(x)=(x-\alpha_1)f_1(x)$。且 $\deg f_1(x)=n-1$，由归纳假设，存在 $\alpha_2,\cdots,\alpha_n\in\bar{K}$ 使 $f_1(x)=(x-\alpha_2)\cdots(x-\alpha_n)$。故定理得证。

□.

对任意多项式 $f(x)\in K[x]$（K 可以是有限域），令

$$f_{\bar{K}}:\bar{K}\longrightarrow\bar{K},\quad u\longmapsto f(u),$$

表示由多项式 $f(x)$ 诱导的函数。 则我们有.

推论5.2.3： 设 $\bar{K}_{pol}=\{f_{\bar{K}}:\bar{K}\to\bar{K}\mid f(x)\in K[x]\}$ 表示由 $K[x]$ 中多项式诱导的 $\bar{K}$ 上的多项式函数环。 则映射

$$K[x]\longrightarrow\bar{K}_{pol},\quad f(x)\longmapsto f_{\bar{K}}$$

是一个环同构（即：它是双射且保持运算相容）。

证明：由引理5.2.1已知，映射 $K[x]\longrightarrow\bar{K}_{pol}$ 是满射且

$$f(x)+g(x)\longmapsto f_{\bar{K}}+g_{\bar{K}},\quad f(x)\cdot g(x)\longmapsto f_{\bar{K}}\cdot g_{\bar{K}}.$$

故只需证明：若 $f(x)\in K[x]$ 是非零多项式，则函数 $f_{\bar{K}}:\bar{K}\to\bar{K}$ 非零，它由定理5.2.2可证。

□

对任意 $f(x_1,\cdots,x_n)\in K[x_1,\cdots,x_n]$，可定义多项式函数

$$f_{\bar{K}}:\bar{K}^n\longrightarrow K,\quad (\alpha_1,\cdots,\alpha_n)\longmapsto f(\alpha_1,\cdots,\alpha_n).$$

不难证明（对 n 使用归纳法）： $f(x_1,\cdots,x_n)=g(x_1,\cdots,x_n)$ 当且仅当 $f_{\bar{K}}=g_{\bar{K}}$.

定义 5.2.2（初等对称多项式）对不定元 $x_1,\cdots,x_n$，下述 n 个对称多项式

$$S_k(x_1,\cdots,x_n)=\sum_{1\le i_1<i_2<\cdots<i_k\le n} x_{i_1}x_{i_2}\cdots x_{i_k}, \quad (1\le k\le n)$$

称为 n 元初等对称多项式. □

定理 5.2.3（韦达公式）. 设 $\alpha_1,\cdots,\alpha_n\in K$ 是

$$f(x)=x^n+a_1x^{n-1}+\cdots+a_kx^{n-k}+\cdots+a_{n-1}x+a_n\in K[x]$$

的根. 则 $a_1=-s_1(\alpha_1,\cdots,\alpha_n)=-(\alpha_1+\cdots+\alpha_n)$, $\cdots$

$$a_k=(-1)^k s_k(\alpha_1,\cdots,\alpha_n),\ \cdots,\ a_n=(-1)^n\alpha_1\cdots\alpha_n.$$

证明：由定理 5.2.2，在 $K[x]$ 中，有多项式等式

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

展开比较系数可得 $a_k=(-1)^k s_k(\alpha_1,\cdots,\alpha_n)$. $(1\le k\le n)$. □.

习题 5.2

1. 设 $p>2$ 是一个素数，证明：$s_k(1,2,\cdots,p-1)\equiv 0 \pmod p$, $1\le k\le p-2$ 和 $(p-1)!+1\equiv 0 \pmod p$.（提示：对 $f(x)=x^{p-1}-1\in\mathbb{F}_p[x]$ 利用韦达公式和费马小定理）.

2. 利用上题证明威尔逊定理：$p>2$ 是一个素数的充要条件是
$$(p-1)!+1\equiv 0 \pmod p.$$
（提示：若 $p=p_1p_2$, $p_1>1, p_2>1$, 则 p_1 整除 $(p-1)!$ 从而 $(p-1)!+1\not\equiv 0 \pmod p$）.

3. 若 $f(x_1,\cdots,x_n)\in K[x_1,\cdots,x_n]$ 是非零多项式，~~[illegible]~~ K 是无限域，证明存在 $\alpha_1,\cdots,\alpha_n\in K$ 使 $f(\alpha_1,\cdots,\alpha_n)\neq 0$.

4. 设 $f(x_1,\cdots,x_n)\in\mathbb{F}_p[x_1,\cdots,x_n]$ 非零。若 $f(x_1,\cdots,x_n)$ 关于每个 x_i 的次数皆小于 p（即 $\deg_{x_i}(f)<p$）。则存在 $\alpha_1,\cdots,\alpha_n\in\mathbb{F}_p$ 使 $f(\alpha_1,\cdots,\alpha_n)\neq 0$.

5. 证明：每个多项式 $f(x_1,\cdots,x_n)\in\mathbb{F}_p[x_1,\cdots,x_n]$ 皆可表示为

$$f(x_1,\cdots,x_n)=\sum_{i=1}^{n} g_i(x_1,\cdots,x_n)(x_i^p-x_i)\quad+f^*(x_1,\cdots,x_n)$$

其中 $f^*(x_1,\cdots,x_n)\in\mathbb{F}_p[x_1,\cdots,x_n]$ 或者是零多项式或者满足：

$$\deg_{x_i}(f^*)<p\ (1\le i\le n),\quad \deg f^*(x_1,\cdots,x_n)\le\deg f(x_1,\cdots,x_n).$$

6. 设 $f(x_1,\cdots,x_n)\in\mathbb{F}_p[x_1,\cdots,x_n]$. 证明：$f_{\mathbb{F}_p}:\mathbb{F}_p^n\longrightarrow\mathbb{F}_p$, $(\alpha_1,\cdots,\alpha_n)\mapsto f(\alpha_1,\cdots,\alpha_n)$ 是零函数的充分必要条件是存在 $g_i(x_1,\cdots,x_n)\in\mathbb{F}_p[x_1,\cdots,x_n]$ 使

$$f(x_1,\cdots,x_n)=\sum_{i=1}^{n} g_i(x_1,\cdots,x_n)(x_i^p-x_i).$$

7. 若 $f(x_1,\cdots,x_n)\in\mathbb{F}_p[x_1,\cdots,x_n]$ 满足条件：(1) $\deg_{x_i}(f)<p\ (1\le i\le n)$
(2) $f(0,\cdots,0)=1$, 但 $f(\alpha_1,\cdots,\alpha_n)=0$ 对所有非零 $(\alpha_1,\cdots,\alpha_n)\in\mathbb{F}_p^n$ 成立。
证明：$f(x_1,\cdots,x_n)=(1-x_1^{p-1})\cdots\cdots(1-x_n^{p-1})$.
（提示：利用多项式的唯一性和上述习题 6 ）.

8. 设 $f(x_1,\cdots,x_n)\in\mathbb{F}_p[x_1,\cdots,x_n]$ m 次齐次多项式. 若 $0<m<n$, 证明存在不全为零的 $\alpha_1,\cdots,\alpha_n\in\mathbb{F}_p$ 使

$$f(\alpha_1,\cdots,\alpha_n)=0$$

（提示：反证法，若 $f(\alpha_1,\cdots,\alpha_n)\neq0$, $\forall$ 非零 $(\alpha_1,\cdots,\alpha_n)\in\mathbb{F}_p^n$, 令

$$g(x_1,\cdots,x_n)=1-f(x_1,\cdots,x_n)^{p-1},$$

则 $g^*(x_1,\cdots,x_n)$ 是习题5中由 $g(x_1,\cdots,x_n)$ 而确定的多项式. 利用习题7 可证 $g^*(x_1,\cdots,x_n)=(1-x_1^{p-1})\cdots\cdots(1-x_n^{p-1})$. 从而

$$\deg(g)=(p-1)\deg(f)=m(p-1)<n(p-1)=\deg(g^*)$$ 与习题5中的要求 $\deg(g^*)\le\deg(g)$ 矛盾）.

§5.3. 对称多项式

在第一章（定义1.4.5）我们定义了对称函数和反对称函数。而多项式 $f(x_1,\cdots,x_n)\in K[x_1,\cdots,x_n]$ 可由多项式函数（完满导出）

$$f_{\bar K}: \bar K^n = \bar K\times\cdots\times\bar K \longrightarrow \bar K,\quad (\alpha_1,\cdots,\alpha_n)\longmapsto f(\alpha_1,\cdots,\alpha_n)$$

唯一确定义。如果 $f_{\bar K}$ 是对称函数（或反对称函数），则称 $f(x_1,\cdots,x_n)$ 是对称多项式（或反对称多项式）。它们的等价定义如下.

定义5.3.1. $f(x_1,\cdots,x_n)\in K[x_1,\cdots,x_n]$ 称为对称多项式，如果

$$f(x_{\pi(1)},\cdots,x_{\pi(n)}) = f(x_1,x_2,\cdots,x_n),\quad \forall \pi\in S_n$$

$f(x_1,\cdots,x_n)$ 称为反对称多项式，如果

$$f(x_{\pi(1)},\cdots,x_{\pi(n)}) = \varepsilon_\pi f(x_1,\cdots,x_n),\quad \forall \pi\in S_n$$

□.

例5.3.1：上节中定义的初等对称多项式

$$s_k(x_1,\cdots,x_n) = \sum_{1\le i_1<\cdots<i_k\le n} x_{i_1}x_{i_2}\cdots x_{i_k},$$

$(k=1,2,\cdots,n)$ 是对称多项式.

□

例5.3.2：$\Delta(x_1,\cdots,x_n) = \prod_{1\le i<j\le n}(x_j-x_i)$ 是反对称多项式

而 $\Delta(x_1,\cdots,x_n)^2$ 是对称多项式.

□

例5.3.3，$f(x_1,x_2,x_3) = x_1^2+x_2+x_3^3\in K[x_1,x_2,x_3]$ 不是对称多项式. 但 $f(s_1,s_2,s_3) = s_1^2+s_2+s_3^3 = (x_1+x_2+x_3)^2+(x_1x_2+x_2x_3+x_1x_3)+(x_1x_2x_3)^3\in K[x_1,x_2,x_3]$ 是对称多项式

□

事实上，对任意多项式 $g(y_1, y_2, \cdots, y_n) \in K[y_1, \cdots, y_n]$，将 $y_1, \cdots, y_n$ 以 $K[x_1, \cdots, x_n]$ 中的 n 个初等对称多项式

$$s_k = s_k(x_1, \cdots, x_n) = \sum_{1 \le i_1 < \cdots < i_k \le n} x_{i_1} x_{i_2} \cdots x_{i_k} \qquad (1 \le k \le n)$$

代入可得 $K[x_1, \cdots, x_n]$ 中的对称多项式

$$f(x_1, \cdots, x_n) := g(s_1, s_2, \cdots, s_n) \in K[x_1, \cdots, x_n].$$

本节的主要定理是：$K[x_1, \cdots, x_n]$ 中的所有对称多项式皆可如此得到！

定理 5.3.1（对称多项式基本定理）

设 $f(x_1, \cdots, x_n) \in K[x_1, \cdots, x_n]$ 是对称多项式。则存在唯一的 $g(y_1, \cdots, y_n) \in K[y_1, \cdots, y_n]$ 使得

$$f(x_1, \cdots, x_n) = g(s_1, \cdots, s_n) \in K[x_1, \cdots, x_n].$$

并且 g 的系数是 f 系数的整系数线性组合。

□

设 $f = f(x_1, \cdots, x_n) = f_0 + f_1 + \cdots + f_m$，$f_m \ne 0$，其中 $f_i = f_i(x_1, \cdots, x_n)$ 是 i 次齐次多项式。对任意 $\pi \in S_n$，令

$$f^\pi(x_1, \cdots, x_n) := f(x_{\pi(1)}, \cdots, x_{\pi(n)}) \in K[x_1, \cdots, x_n].$$

则 $f^\pi = f_0^\pi + f_1^\pi + \cdots + f_m^\pi$，其中 $f_i^\pi = f_i^\pi(x_1 \cdots x_n) := f_i(x_{\pi(1)}, \cdots, x_{\pi(n)})$。因此 $f^\pi = f \iff f_i = f_i^\pi \quad (i = 0, 1, \cdots, m)$。即 f 是对称多项式当且仅当它的每个齐次部分是对称多项式。所以只需对 $f(x_1, \cdots, x_n)$ 是齐次多项式的情形证明上述定理。

我们将采用"消去首项法"证明该定理。但齐次多项式 $f(x_1, \cdots, x_n)$ 中的每个单项式次数相同，所以我们不能按单

项式的次数排序（$n=1$时例外）. 下面我们在所有 m次单项式中引入"字典排序".

定义5.3.2（字典排序法）. 设 $u=ax_1^{i_1}x_2^{i_2}\cdots x_n^{i_n}$，$v=bx_1^{j_1}x_2^{j_2}\cdots x_n^{j_n}$ 是两个m次单项式。如果 $(i_1-j_1, i_2-j_2, \cdots, i_n-j_n)=(0,\cdots,0,t,\cdots)$ 其中 $t>0$，则称u位于v之前（记为 $u>v$，或 $(i_1,\cdots,i_n)>(j_1,\cdots,j_n)$）。对任意$m$次齐次多项式 $f=f(x_1,\cdots,x_n)$，令 $FT(f)$ 表示f中（按上述方式）排在最前面的单项式，称为$f(x_1,\cdots,x_n)$的首项. □

引理5.3.1. 若 $h(x_1,\cdots,x_n)=h_1(x_1,\cdots,x_n)\cdots h_r(x_1,\cdots,x_n)$，则

$$FT(h)=FT(h_1)\cdots FT(h_r).$$

证明: 只需证明，若 $h=h_1\cdot h_2$ 是齐次多项式，则 $FT(h)=FT(h_1)FT(h_2)$。它可由下述结论推出.
若 $(i_1,i_2,\cdots,i_n)>(k_1,k_2,\cdots,k_n)$，则对任意 $(j_1,\cdots,j_n)$ 有

$$(i_1,i_2,\cdots,i_n)+(j_1,j_2,\cdots,j_n)>(k_1,\cdots,k_n)+(j_1,j_2,\cdots,j_n).$$

□

引理5.3.2. 设 $f(x_1,\cdots,x_n)\in K[x_1,\cdots,x_n]$ 是m次齐次对称多项式，$FT(f)=ax_1^{i_1}x_2^{i_2}\cdots x_n^{i_n}$，则 $i_1\geq i_2\geq\cdots\geq i_n$.

证明:（反证法）. 若存在 $i_k<i_{k+1}$ $(1\leq k<n)$，置换 x_k 和 x_{k+1} 的位置，可得单项式 $u=ax_1^{i_1}\cdots x_k^{i_{k+1}}x_{k+1}^{i_k}\cdots x_n^{i_n}$。因为$f$是对称多项式，故$u$也是$f$中的一个单项式。但 $u>FT(f)$ 与 $FT(f)$是首项矛盾。 □

定理5.3.1的证明: 不妨设 $f=f(x_1,\cdots,x_n)\in K[x_1,\cdots,x_n]$ 是m次齐次对称多项式，$FT(f)=ax_1^{i_1}x_2^{i_2}\cdots x_n^{i_n}$，$i_1\geq i_2\geq\cdots\geq i_n$.

令 $f_{(1)}=f_{(1)}(x_1,x_2,\cdots,x_n):=f(x_1,x_2,\cdots,x_n)-as_1^{i_1-i_2}s_2^{i_2-i_3}\cdots s_n^{i_n}$.

由引理5.3.1, $FT(as_1^{i_1-i_2}s_2^{i_2-i_3}\cdots s_n^{i_n})=aFT(s_1)^{i_1-i_2}FT(s_2)^{i_2-i_3}\cdots FT(s_n)^{i_n}$

$$=a\,x_1^{i_1-i_2}(x_1x_2)^{i_2-i_3}\cdots(x_1x_2\cdots x_n)^{i_n}=a x_1^{i_1}x_2^{i_2}\cdots x_n^{i_n}=FT(f).$$

所以 $FT(f)>FT(f_{(1)})$. 且 $f_{(1)}$ 仍是齐次对称多项式.

令 $\bar{g}(y_1,\cdots,y_n)=a y_1^{i_1-i_2}y_2^{i_2-i_3}\cdots y_n^{i_n}$, 则

$$f_{(1)}(x_1,\cdots,x_n)=f(x_1,\cdots,x_n)-\bar{g}(s_1,s_2,\cdots,s_n).$$

且 $f_{(1)}(x_1,\cdots,x_n)$ 的系数形如 $c-qa$ $(q\in\mathbb{Z})$, c,a 是 $f(x_1,\cdots,x_n)$ 的系数. 重复上述消首项的过程, 可设存在 $g_{(1)}(y_1,\cdots,y_n)$ 使

$$f_{(1)}(x_1,\cdots,x_n)=g_{(1)}(s_1,\cdots,s_n)$$

其中 $g_{(1)}$ 的系数是 $f_{(1)}$ 系数的整线性组合. 令

$$g(y_1,\cdots,y_n)=g_{(1)}(y_1,\cdots,y_n)+\bar{g}(y_1,\cdots,y_n)$$

即得 $f(x_1,\cdots,x_n)=g(s_1,\cdots,s_n)$. 下面证明这样的 $g(y_1,\cdots,y_n)$ 是唯一的.

若 $g_1(y_1,\cdots,y_n)$, $g_2(y_1,\cdots,y_n)\in K[y_1,\cdots,y_n]$ 使得

$$g_1(s_1,\cdots,s_n)=f(x_1,\cdots,x_n)=g_2(s_1,\cdots,s_n).$$

则 $g_1(y_1,\cdots,y_n)=g_2(y_1,\cdots,y_n)$. 令 $g(y_1,\cdots,y_n):=g_1(y_1,\cdots,y_n)-g_2(y_1,\cdots,y_n)$

则 $g(s_1,\cdots,s_n)\in K[x_1,\cdots,x_n]$ 是零多项式. 故只需证明: 若 $g(y_1,\cdots,y_n)$ 是 $K[y_1,\cdots,y_n]$ 中的非零多项式, 则 $g(s_1,\cdots,s_n)$ 是 $K[x_1,\cdots,x_n]$ 中的非零多项式。事实上, 若 $ay_1^{k_1}y_2^{k_2}\cdots y_n^{k_n}$ 是 $g(y_1,\cdots,y_n)$ 的一个非零单项式, 则 $FT(as_1^{k_1}s_2^{k_2}\cdots s_n^{k_n})=a x_1^{k_1+k_2+\cdots+k_n}x_2^{k_2+k_3+\cdots+k_n}\cdots x_n^{k_n}$.

所以, 若 $ay_1^{k_1}y_2^{k_2}\cdots y_n^{k_n}$, $by_1^{l_1}y_2^{l_2}\cdots y_n^{l_n}$ 是 $g(y_1,\cdots,y_n)$ 中不同同类项 (即 $(k_1,\cdots,k_n)\neq(l_1,\cdots,l_n)$), 则 $FT(as_1^{k_1}s_2^{k_2}\cdots s_n^{k_n})$, $FT(bs_1^{l_1}s_2^{l_2}\cdots s_n^{l_n})$ 是 $g(s_1,\cdots,s_n)$ 中的不同类项. 故 $FT(g(s_1,\cdots,s_n))\neq 0$ 矛盾 □

注记：定理的一个推论是：$K[s_1, s_2, \cdots, s_n]$ 与 $K[x_1, x_2, \cdots, x_n]$ 同构（尽管 $K[s_1, s_2, \cdots, s_n] \subset K[x_1, \cdots, x_n]$ 是真子集合）.

推论5.3.1．设 $z_1, \cdots, z_n \in \mathbb{C}$ 是多项式

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

的根．则对任意对称多项式 $h(x_1, \cdots, x_n) \in \mathbb{Z}[x_1, \cdots, x_n]$ 有

$$h(z_1, \cdots, z_n) \in \mathbb{Z}.$$

证明：由定理5.3.1，存在 $g(y_1, \cdots, y_n) \in \mathbb{Z}[y_1, \cdots, y_n]$ 使

$$h(x_1, \cdots, x_n) = g(s_1, s_2, \cdots, s_n).$$

由韦达公式，$h(z_1, \cdots, z_n) = g(s_1(z_1, \cdots, z_n), \cdots, s_n(z_1, \cdots, z_n)) \in \mathbb{Z}$. □

例5.3.4：$f(x) = x^5 - 1 \in \mathbb{Z}[x]$ 在 $\mathbb{C}$ 中有5个根．

$$\varepsilon_k = \cos\frac{2k\pi}{5} + i\sin\frac{2k\pi}{5}, \quad 0 \le k < 5,$$

由韦达公式 $1+\varepsilon_1+\varepsilon_2+\varepsilon_3+\varepsilon_4 = 0$ 可得（实部）$2\cos\frac{4\pi}{5} + 2\cos\frac{2\pi}{5} + 1 = 0$．令 $y = \cos\frac{2\pi}{5}$，则 $\cos\frac{4\pi}{5} = 2y^2 - 1$．所以 $4y^2 + 2y - 1 = 0$．解二次方程可得 $\cos\frac{2\pi}{5} = \frac{\sqrt{5}-1}{4}$，$\cos\frac{4\pi}{5} = \frac{\sqrt{5}+1}{4}$ □.

例5.3.5．$P_2 = x_1^2 + \cdots + x_n^2$ 是对称多项式，$FT(P_2) = x_1^2$，

$$f_{(1)} = P_2 - s_1^2 = -2(x_1x_2 + x_1x_3 + \cdots + x_1x_n + \cdots + x_{n-1}x_n) = -2s_2.$$

故 $g(y_1, \cdots, y_n) = y_1^2 - 2y_2$，$P_2 = g(s_1, \cdots, s_n) = s_1^2 - 2s_2$ □

例5.3.6 $(n \ge 3)$．$P_3 = x_1^3 + x_2^3 + \cdots + x_n^3$ 是对称多项式，$FT(P_3) = x_1^3$，令

$f_{(1)} = P_3 - s_1^3$，$FT(f_{(1)}) = -3x_1^2x_2$．令 $f_{(2)} = f_{(1)} + 3s_1s_2$

则 $FT(f_{(2)}) = 3x_1x_2x_3$，$f_{(2)} - 3s_3 = 0$．故

$$P_3 = s_1^3 - 3s_1s_2 + 3s_3.$$ □

定理5.3.2 (牛顿公式). 设 $p_k = p_k(x_1,\cdots,x_n) = x_1^k + x_2^k+\cdots+x_n^k$. 我们有如下的递推关系(称为牛顿公式).

(1) 如果 $1\le k\le n$, 则有递推公式

(5.3.1) $$p_k - p_{k-1}s_1+\cdots+(-1)^{k-1}p_1 s_{k-1} + (-1)^k k s_k = 0$$

(2) 如果 $k>n$, 则有递推公式

(5.3.2) $$p_k - p_{k-1}s_1+\cdots+(-1)^{n-1}p_{k-n+1}s_{n-1}+(-1)^n p_{k-n}s_n=0.$$

证明: 在 $R:=k[x_1,x_2,\cdots,x_n]$ 上引入新的不定元 Y, 考虑多项式

$$f(Y)=(Y-x_1)(Y-x_2)\cdots(Y-x_n)\in R[Y].$$

则 $$(Y-x_1)(Y-x_2)\cdots(Y-x_n)=Y^n - s_1Y^{n-1}+s_2Y^{n-2}+\cdots+(-1)^n s_n$$

其中 $s_k = s_k(x_1,\cdots,x_n)\in k[x_1,\cdots,x_n]$ 是第k个初等对称多项式。令 $Y=x_i$, 可得恒等式

$$x_i^n - s_1x_i^{n-1}+\cdots+(-1)^{n-1}s_{n-1}x_i+(-1)^n s_n=0 \qquad (1\le i\le n).$$

若 $k\ge n$, 以 x_i^{k-n} 同乘上式两边得

$$x_i^k - s_1x_i^{k-1}+\cdots+(-1)^{n-1}s_{n-1}x_i^{k-n+1}+(-1)^n s_n x_i^{k-n}=0$$

对 $i=1,2,\cdots,n$ 求和即得 (5.3.2),

$$p_k - s_1p_{k-1}+\cdots+(-1)^{n-1}s_{n-1}p_{k-n+1}+(-1)^n s_n p_{k-n}=0.$$

当 $k=n$ 时, 上式与公式(5.3.1)重合 (注意: $p_0 = x_1^0+\cdots+x_n^0=n$).

为证公式(5.3.1), 只需证明: 当 $k\le n$ 时,

$$f_{k,n}(x_1,\cdots,x_n) := p_k - p_{k-1}s_1+\cdots+(-1)^{k-1}p_1s_{k-1}+(-1)^k k s_k$$

是零多项式. 对 $r=n-k$ 作归纳法, 已知 $r=0$ 时 $f_{k,n}=0$.

设结论对 $r-1$ 成立. 若 $r>0$（即 $n>k$），令 $x_n=0$，得

$$f_{k,n}(x_1,\cdots,x_{n-1},0)=P_k(x_1,\cdots,x_{n-1},0)-P_{k-1}(x_1,\cdots,x_{n-1},0)S_1(x_1,\cdots,x_{n-1},0)+\cdots+(-1)^{k-1}P_1(x_1,\cdots,x_{n-1},0)S_{k-1}(x_1,\cdots,x_{n-1},0)+(-1)^k kS_k(x_1,\cdots,x_{n-1},0)$$

$$=P_k(x_1,\cdots,x_{n-1})-P_{k-1}(x_1,\cdots,x_{n-1})S_1(x_1,\cdots,x_{n-1})+\cdots+(-1)^{k-1}P_1(x_1,\cdots,x_{n-1})S_{k-1}(x_1,\cdots,x_{n-1})+(-1)^k k\,S_k(x_1,\cdots,x_{n-1})=0$$

（因为 $n-1-k=(n-k)-1=r-1$）.

故 $x_n\mid f_{k,n}(x_1,\cdots,x_n)$. 因为 $f_{k,n}(x_1,\cdots,x_n)$ 是对称多项式，由 $x_n\mid f_{k,n}(x_1,\cdots,x_n)$ 可得 $x_i\mid f_{k,n}(x_1,\cdots,x_n)$（$1\le i\le n$）. 但 $x_1,x_2,\cdots,x_n$ 是 $K[x_1,\cdots,x_n]$ 中互不相同的不可约多项式，由于 $K[x_1,\cdots,x_n]$ 中每个多项式都有唯一不可约分解，故 $S_n=x_1x_2\cdots x_n$ 整除 $f_{k,n}(x_1,\cdots,x_n)$. 即存在 $t(x_1,\cdots,x_n)\in K[x_1,\cdots,x_n]$ 使

$$f_{k,n}(x_1,\cdots,x_n)=S_n(x_1,\cdots,x_n)\,t(x_1,\cdots,x_n).$$

若 $f_{k,n}$ 是非零多项式，则 $\deg f_{k,n}=k<n=\deg S_n$ 与 $r>0$ 矛盾. □

例5.3.7 多项式 $\Delta_n^2=\prod_{i<j}(x_j-x_i)^2\in K[x_1,\cdots,x_n]$ 是对称多项式. 由定理5.3.1，存在唯一多项式 $\mathrm{Dis}(y_1,y_2,\cdots,y_n)\in K[y_1,\cdots,y_n]$ 使得

$$\Delta_n^2=\prod_{i<j}(x_j-x_i)^2=\mathrm{Dis}(S_1,S_2,\cdots,S_n).$$

Δ_n 是下列矩阵

$$A=\begin{pmatrix}1&1&\cdots&1\\ x_1&x_2&\cdots&x_n\\ \vdots&\vdots&&\\ x_1^{n-1}&x_2^{n-1}&\cdots&x_n^{n-1}\end{pmatrix}$$

的行列式. 由 $\Delta_n^2=|A^tA|$ 可得

$$\mathrm{Dis}(S_1,S_2,\cdots,S_n)=\begin{vmatrix}n&P_1&P_2&\cdots&P_{n-1}\\ P_1&P_2&P_3&\cdots&P_n\\ P_2&P_3&P_4&\cdots&P_{n+1}\\ \cdots&&\cdots&\cdots&\cdots\\ P_{n+1}&P_{n+2}&P_{n+3}&\cdots&P_{2n-2}\end{vmatrix},\quad P_k=x_1^k+\cdots+x_n^k.$$

□

定义 5.3.3 (方程的判别式). 设

$$f=f(x)=x^n+a_1x^{n-1}+\cdots+a_{n-1}x+a_n\in K[x].$$

$D(f):=Dis\left(-a_1,a_2,\cdots,(-1)^ka_k,\cdots,(-1)^na_n\right)$ (即在多项式 $D(y_1,\cdots,y_n)$ 中分别令 $y_k=(-1)^ka_k$) 称为方程 $f=x^n+a_1x^{n-1}+\cdots+a_{n-1}x+a_n=0$ 的判别式. □

推论 5.3.2: 方程 $x^n+a_1x^{n-1}+\cdots+a_{n-1}x+a_n=0$ 有重根当且仅当 $D(f)=0$.

证明: 设 $\alpha_1,\cdots,\alpha_n\in\bar{K}$ 是 $x^n+a_1x^{n-1}+\cdots+a_{n-1}x+a_n=0$ 的根. 由韦达公式得 $s_k(\alpha_1,\cdots,\alpha_n)=(-1)^ka_k$. $(1\le k\le n)$.
所以 $\Delta_n^2(\alpha_1,\cdots,\alpha_n)=\prod_{i<j}(\alpha_j-\alpha_i)^2=Dis\left(-a_1,a_2,\cdots,(-1)^ka_k,\cdots,(-1)^na_n\right)=0$ 当且仅当 $\alpha_1,\cdots,\alpha_n$ 中有两个根相等. □

习题 5.3

1. 计算方程 $f(x)=x^2+bx+c=0$ 的判别式 $D(f)$.

2. 计算方程 $f(x)=x^3+ax+b=0$ 的判别式 $D(f)$.

3. 设 $p>2$ 是素数, 利用牛顿公式证明

$$\sum_{i=1}^{p-1}i^m=\begin{cases}-1\ (\mathrm{mod}\ p), & 若\ (p-1)\mid m\\ 0\ (\mathrm{mod}\ p), & 若\ (p-1)\nmid m.\end{cases}$$

4. 将 $f(x_1,x_2,x_3)=x_1^2x_2+x_1x_2^2+x_1^2x_3+x_1x_3^2+x_2^2x_3+x_2x_3^2\in K[x_1,x_2,x_3]$ 表示成初等对称多项式 s_1,s_2,s_3 的多项式.

5. 若 $f(x_1,\cdots,x_n)\in\mathbb{Q}[x_1,\cdots,x_n]$ 是反对称多项式. 证明存在对称多项式 $g(x_1,\cdots,x_n)\in\mathbb{Q}[x_1,\cdots,x_n]$ 使 $f(x_1,\cdots,x_n)=g(x_1,\cdots,x_n)\prod_{i<j}(x_j-x_i)$.

6. 利用牛顿公式和克莱姆法则证明下述公式

$$P_k = \begin{vmatrix} s_1 & 1 & 0 & 0 & \cdots & 0 \\ 2s_2 & s_1 & 1 & 0 & \cdots & 0 \\ 3s_3 & s_2 & s_1 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & & \vdots \\ (k-1)s_{k-1} & s_{k-2} & s_{k-3} & s_{k-4} & \cdots & 1 \\ ks_k & s_{k-1} & s_{k-2} & s_{k-3} & \cdots & s_1 \end{vmatrix}$$

$$s_k = \frac{1}{k!} \begin{vmatrix} P_1 & 1 & 0 & 0 & \cdots & 0 \\ P_2 & P_1 & 1 & 0 & \cdots & 0 \\ P_3 & P_2 & P_1 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & & \vdots \\ P_{k-1} & P_{k-2} & P_{k-3} & P_{k-4} & \cdots & 1 \\ P_k & P_{k-1} & P_{k-2} & P_{k-3} & \cdots & P_1 \end{vmatrix}$$

其中 $P_k = x_1^k + \cdots + x_n^k$，$s_k = s_k(x_1, \cdots, x_n) = \sum\limits_{1 \le i_1 < \cdots < i_k \le n} x_{i_1} x_{i_2} \cdots x_{i_k}$，

§5.4. 两个多项式的结式.

若 $K\subset\mathbb{C}$ 是一个子域，$f(x_1,\cdots,x_n)\in K[x_1,\cdots,x_n]$，无论是实际应用还是数学本身都需要研究由 f 定义的超曲面

$$V(f)=\{(\alpha_1,\cdots,\alpha_n)\in\mathbb{C}^n \mid f(\alpha_1,\cdots,\alpha_n)=0\}\subset\mathbb{C}^n.$$

对任意域 K，及 $f(x_1,\cdots,x_n)\in K[x_1,\cdots,x_n]$，我们亦称

$$V(f)=\{(\alpha_1,\cdots,\alpha_n)\in\bar{K}^n \mid f(\alpha_1,\cdots,\alpha_n)=0\}\subset\bar{K}^n$$

是空间 $\bar{K}^n$ 中的一个超曲面。下所以自然的问题是：对于给定的 $f(x_1,\cdots,x_n)$，$g(x_1,\cdots,x_n)\in K[x_1,\cdots,x_n]$，两个超曲面是否相交？或者可问：如何判断 f,g 是否有次数大于 0 的公因子？一个自然的想法是：通过方程组

$$\begin{cases} f(x_1,\cdots,x_n)=0 \\ g(x_1,\cdots,x_n)=0 \end{cases}$$

消元，将问题归结于 $n-1$ 个变元的问题。令 $R=K[x_1,\cdots,x_{n-1}]$，$x_n=x$，将上述多项式写成

$$(5.4.1).\quad \begin{cases} f(x)=a_0x^n+a_1x^{n-1}+\cdots+a_{n-1}x+a_n \\ g(x)=b_0x^m+b_1x^{m-1}+\cdots+b_{m-1}x+b_m. \end{cases}$$

$a_i,b_i\in R$，$n>0$，$m>0$。下面的讨论中，系数 a_i，b_i 可以看成 $x_1,\cdots,x_{n-1}$ 的函数，或者看成这些函数在 $(x_1,\cdots,x_n)=(t_1,\cdots,t_{n-1})$ 的取值 $a_i(t_1,\cdots,t_{n-1})$，$b_i(t_1,\cdots,t_{n-1})\in F$（其中 t_i 是某个域 F 中的一般元素）。将 a_i,b_i 看成在某个域中的变量，且不排除 $a_0=0$ 或 $b_0=0$ 的情形。所以下面定义中的结式 $\mathrm{Res}(f,g)$ 可以看成关于 $a_0,\cdots,a_n,b_0,\cdots,b_m$ 的多项式，或看成关于 $a_0,\cdots,a_n,b_0,\cdots,b_m$ 的多项式函数。

定义5.4.1. 多项式 f 和 g 的结式定义为

$$\mathrm{Res}(f,g)=\begin{vmatrix} a_0 & a_1 & \cdots & a_n & & & \\ & a_0 & a_1 & \cdots & a_n & & \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ & & & a_0 & a_1 & \cdots & a_n \\ b_0 & b_1 & \cdots & b_m & & & \\ & b_0 & b_1 & \cdots & b_m & & \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ & & & b_0 & b_1 & \cdots & b_m \end{vmatrix}\begin{matrix}\left.\vphantom{\begin{matrix}a\\a\\a\\a\end{matrix}}\right\}m\text{行}\\ \left.\vphantom{\begin{matrix}a\\a\\a\\a\end{matrix}}\right\}n\text{行}\end{matrix}$$

□

由行列式定义，$\mathrm{Res}(f,g)$ 是关于 $a_0,\cdots,a_n,b_0,\cdots,b_m$ 的 $m+n$ 次齐次多项式（关于 $a_0,\cdots,a_n$ 的次数是 m，关于 $b_0,\cdots,b_m$ 的次数是 n）。下面将介绍结式 $\mathrm{Res}(f,g)$ 的基本性质和应用。在证明中经常需要将 $a_0,a_1,\cdots,a_m,b_0,b_1,\cdots,b_m$ 看成某个域 F 中的元素，这总是成立的。比如，环 $K[x_1,\cdots,x_{n-1}]$ 可以看成有理函数域 $\mathbb{F}=\left\{\frac{h_2(x_1,\cdots,x_{n-1})}{h_1(x_1,\cdots,x_{n-1})}\,\middle|\, h_1(x_1,\cdots,x_{n-1})\neq 0,\ h_i(x_1,\cdots,x_{n-1})\in K[x_1,\cdots,x_{n-1}]\right\}$ 的子环。详细讨论将留待抽象代数课程。

定理5.4.1. $\mathrm{Res}(f,g)=0$ 当且仅当($\Leftrightarrow$) 或者 $a_0=b_0=0$，

或者 f 和 g 有非平凡的公因子（在 $F[x]$ 中）。

证明：首先证明条件"$a_0=b_0=0$ 或者 f 和 g 有非平凡公因子"

等价于：存在非零多项式 $f_1,g_1\in F[x]$，使

(5.4.2) $\quad f\cdot g_1+f_1 g=0,\ \deg f_1<n,\ \deg g_1<m$。

若"$a_0=b_0=0$，取 $f_1=f,\ g_1=-g$ 得条件 (5.4.2)。若 a_0,b_0 不全为零但 f,g 的最大公因子 h 的次数大于零，则 $f=hf_1,\ g=-hg_1$ 得条件(5.4.2)：$fg_1+gf_1=0$ 且 $\deg f_1<n$，$\deg g_1<m$.

下面证明：$\mathrm{Res}(f,g)=0 \Longleftrightarrow$ 条件 (5.4.2) 成立.

令 $f_1=c_0x^{n-1}+c_1x^{n-2}+\cdots+c_{n-1}$，$g_1=d_0x^{m-1}+d_1x^{m-2}+\cdots+d_{m-1}$

其中 $c_0,c_1,\cdots,c_{n-1}$，$d_0,d_1,\cdots,d_{m-1}$ 是 F 中待定的元（c_0,d_0 可能为零）.

则存在非零多项式 f_1,g_1 使得 (5.4.2) 成立的充分必要条件是下列关于 $c_0,c_1,\cdots,c_{n-1},d_0,d_1,\cdots,d_{m-1}$ 的线性方程组（它由恒等式 $fg_1+f_1g=0$ 导出）有非零解：

$$
\begin{cases}
a_0d_0 \quad \cdots + b_0c_0 \quad \cdots = 0\\
a_1d_0+a_0d_1 \quad \cdots + b_1c_0+b_0c_1 \quad \cdots = 0\\
a_2d_0+a_1d_1+a_0d_2 \quad \cdots + b_2c_0+b_1c_1+b_0c_2 = 0\\
\cdots\cdots\cdots
\end{cases}
$$

上述方程组的系数矩阵转置的行列式恰为 $\mathrm{Res}(f,g)$.

故 $\mathrm{Res}(f,g)=0$ 当且仅当存在不全为零的 $c_0,c_1,\cdots,c_{n-1}$ 及不全为零的 $d_0,d_1,\cdots,d_{m-1}$（注意：$(c_0,c_1,\cdots,c_{n-1})$，$(d_0,d_1,\cdots,d_{m-1})$ 中任一组全为零，则由上述方程组，另一组也必全为零）使 (5.4.2) 成立 $\square$.

定理5.4.2：设 K 是一个域，若 f,g 在 $K[x]$ 中可分解为

$$f(x)=a_0(x-\alpha_1)\cdots(x-\alpha_n)$$

$$g(x)=b_0(x-\beta_1)\cdots(x-\beta_m)$$

则 $$\mathrm{Res}(f,g)=a_0^m\prod_{i=1}^n g(\alpha_i)=(-1)^{mn}b_0^n\prod_{j=1}^m f(\beta_j)=a_0^m b_0^n\prod_{i,j}(\alpha_i-\beta_j).$$

证明：设 F 是包含多项式环 $K[x_1,\cdots,x_n,y]$ 的域. $K[x]\subset F[x]$.

考虑 $\Phi(x)=a_0(x-x_1)\cdots(x-x_n)\in F[x]$

我们先证明：在 $F[x]$ 中有 $$\mathrm{Res}(\Phi,g)=a_0^m\prod_{i=1}^n g(x_i)$$

由结式的定义可得：$\mathrm{Res}(\Phi, g-y) = (-1)^n a_0^m y^n + \cdots + \mathrm{Res}(\Phi, g)$

是关于 y 的 n 次多项式，系数在 $K[x_1, \cdots, x_n]$ 中，首项系数为 $(-1)^n a_0^m$，常数为 $\mathrm{Res}(\Phi, g)$。$\Phi(x) = a_0(x-x_1)\cdots(x-x_n)$ 与 $g(x)-g(x_i)$ 有 共同的根 x_i，故 $\Phi(x)$ 与 $g(x)-g(x_i)$ 在 $F[x]$ 中有非平凡公因子。由定理 5.4.1，$\mathrm{Res}(\Phi, g-g(x_i))=0$。即 $g(x_1), \cdots, g(x_n)$ 是多项式 $\mathrm{Res}(\Phi, g-y)$（关于 y）的 n 个根。 故

$$\mathrm{Res}(\Phi, g-y) = a_0^m \prod_{i=1}^{n} (g(x_i) - y) .$$

令 $y=0$ 得 $\mathrm{Res}(\Phi, g) = a_0^m \prod_{i=1}^{m} g(x_i)$。进一步取 $x_i = \alpha_i$ $(1 \le i \le n)$ 得 $\mathrm{Res}(f, g) = a_0^m \prod_{i=1}^{n} g(\alpha_i)$.

另一方面，$\mathrm{Res}(f,g) = (-1)^{mn} \mathrm{Res}(g,f)$（由定义），在上述证明中交换 f 与 g 的次序可得

$$\mathrm{Res}(g,f) = b_0^n \prod_{j=1}^{m} g(\beta_j) .$$

故 $\mathrm{Res}(f,g) = a_0^m \prod_{i=1}^{n} g(\alpha_i) = (-1)^{mn} b_0^n \prod_{j=1}^{m} f(\beta_j)$. □

<u>定义 5.4.2</u>（多项式的导数）. 设 $f(x) = a_0 x^n + a_1 x^{n-1} + \cdots + a_{n-1} x + a_n$ 则多项式 $f'(x) = n a_0 x^{n-1} + (n-1) a_1 x^{n-2} + \cdots + a_{n-1}$ 称为多项式 $f(x)$ 的导数. □

<u>引理 5.4.1</u>: (1) $(f+g)' = f' + g'$.

(2) $(fg)' = f'g + fg'$

(3) 零次多项式的导数是 0 多项式.

证明: 直接验证. □

定理5.4.3 设 $f(x)=a_0(x-\alpha_1)\cdots(x-\alpha_n)\in K[x]$

则 $\mathrm{Res}(f,f')=(-1)^{\frac{n(n-1)}{2}}a_0^{2n-1}\prod_{i>j}(\alpha_i-\alpha_j)^2$.

若定义 $D(f)=a_0^{2n-2}\prod_{i>j}(\alpha_i-\alpha_j)^2$ 为 $f(x)=0$ 的判别式. 则

$$D(f)=(-1)^{\frac{n(n-1)}{2}}a_0^{-1}\mathrm{Res}(f,f')$$

□

证明：由定理5.4.2，$\mathrm{Res}(f,f')=a_0^{n-1}\prod_{i=1}^{n}f'(\alpha_i)$.

计算 $f'(x)$ 在 $x=\alpha_i$ 的取值得 $f'(\alpha_i)=a_0\prod_{j\neq i}(\alpha_i-\alpha_j)$. 故

$$\mathrm{Res}(f,f')=a_0^{2n-1}\prod_{i=1}^{n}\prod_{j\neq i}(\alpha_i-\alpha_j)=(-1)^{\frac{n(n-1)}{2}}a_0^{2n-1}\prod_{i>j}(\alpha_i-\alpha_j)^2$$

□.

$R(f,g)$ 也可看成，通过方程组

$$f(x)=a_0x^n+a_1x^{n-1}+\cdots+a_{n-1}x+a_n$$

$$g(x)=b_0x^m+b_1x^{m-1}+\cdots+b_{m-1}x+b_m,$$

消去 x 后所得关于系数的多项式：存在多项式 $A(x),B(x)$

使 $\mathrm{Res}(f,g)=A(x)f(x)+B(x)g(x)$. (线性组合式).

考虑线性方程组 (关于 $x^{n+m-1},x^{n+m-2},\ldots,x^2,x,1$)：

$$\begin{cases}
x^{m-1}f=a_0x^{n+m-1}+a_1x^{n+m-2}+\cdots+a_nx^{m-1}\\
x^{m-2}f=\qquad a_0x^{n+m-2}+a_1x^{n+m-3}+\cdots+a_nx^{m-2}\\
\vdots\\
xf=\qquad a_0x^{n+1}+a_1x^{n}+\cdots+a_nx\\
f=\qquad a_0x^{n}+a_1x^{n-1}+\cdots+a_n\cdot 1\\
x^{n-1}g=b_0x^{n+m-1}+b_1x^{n+m-2}+\cdots+b_mx^{n-1}\\
\vdots\\
xg=\qquad b_0x^{m+1}+b_1x^{m}+\cdots+b_mx\\
g=\qquad b_0x^{m}+b_1x^{m-1}+\cdots+b_m\cdot 1
\end{cases}$$

写成矩阵形式

$$\begin{pmatrix} x^{m-1}f \\ \vdots \\ xf \\ f \\ x^{n-1}g \\ \vdots \\ xg \\ g \end{pmatrix} = \begin{pmatrix} a_0 & a_1 & \cdots & a_n & & & \\ & a_0 & a_1 & \cdots & a_n & & \\ & & \ddots & & & \ddots & \\ & & & a_0 & a_1 & \cdots & a_n \\ b_0 & b_1 & \cdots & b_m & & & \\ & b_0 & b_1 & \cdots & b_m & & \\ & & \ddots & & & \ddots & \\ & & & b_0 & b_1 & \cdots & b_m \end{pmatrix} \begin{pmatrix} x^{n+m-1} \\ x^{n+m-2} \\ \vdots \\ x^n \\ x^{n-1} \\ \vdots \\ x \\ 1 \end{pmatrix}$$

用 M 表示系数矩阵，~~Res(f,g)~~ 两边同乘伴随矩阵 M^* 得

$$\mathrm{Res}(f,g)\begin{pmatrix} x^{n+m-1} \\ x^{n+m-2} \\ \vdots \\ x \\ 1 \end{pmatrix} = M^* \begin{pmatrix} x^{m-1}f \\ \vdots \\ xf \\ f \\ x^{n-1}g \\ \vdots \\ xg \\ g \end{pmatrix}$$

故 $\mathrm{Res}(f,g) = M^*_{(n)} \begin{pmatrix} x^{m}f \\ \vdots \\ f \\ x^{n}g \\ \vdots \\ g \end{pmatrix} = A(x)f(x) + B(x)g(x)$.

习题 5.4

1. 设 K 是一个域，$f, g, h \in K[x]$. 证明

$$\mathrm{Res}(fg, h) = \mathrm{Res}(f,h)\cdot \mathrm{Res}(g,h).$$

~~在$\bar{K}[x]$中~~（提示：在 $\bar{K}[x]$ 中应用定理5.4.2）.

2. 证明：$D(fg) = D(f)D(g)[\mathrm{Res}(f,g)]^2$.

3. 计算 $\mathrm{Res}(f(x), x-a)$，和 $D(x^n+a)$.

4. 设 $f(x) = x^{n-1}+x^{n-2}+\cdots+x+1$. 求 $D(f)$.（提示：$x^n-1=(x-1)f(x)$）.

第六章. 线性算子初步

设 V 是一个 n 维 K-向量空间, $V \xrightarrow{\mathcal{A}} V$ 是一个线性映射(称为 V 上的线性算子). 对于 V 的任一组基 $e_1, \cdots, e_n \in V$, 线性算子 $V \xrightarrow{\mathcal{A}} V$ 的坐标矩阵 $A \in M_n(K)$ 由等式

$$(\mathcal{A}e_1, \cdots, \mathcal{A}e_n) = (e_1, \cdots, e_n)A, \qquad \mathcal{A}e_i := \mathcal{A}(e_i)$$

确定。一个自然的问题是: 如何选取基 $e_1, \cdots, e_n$ 使得上述坐标矩阵 A "最简单"? 该问题亦可用矩阵语言描述。由于 $V \xrightarrow{\mathcal{A}} V$ 在另一组基 $(e_1', \cdots, e_n') = (e_1, \cdots, e_n)T$ 下的坐标矩阵为 $T^{-1}AT$, 所以我们的问题变成如下的纯矩阵问题: 对给定的矩阵 $A \in M_n(K)$, 如何选取可逆矩阵 T 使 $T^{-1}AT$ "最简单"? 当然, 所谓的"最简单"并没有唯一的判断标准, 我们将从空间分解的角度定义何谓"最简单"。回答上述问题是本章的中心任务。

§6.1. 直和分解与线性算子.

设 V 是 n 维 K-向量空间, $V_1, \cdots, V_s$ 是 V 的子空间, 如果每一个向量 $\alpha \in V$ 皆可写成

$$\alpha = \alpha_1 + \cdots + \alpha_s, \quad \alpha_i \in V_i.$$

则称 V 可分解成 $V_1, \cdots, V_s$ 之和, 记为 $V = V_1 + \cdots + V_s$。

通常对给定的 $\alpha \in V$, 上述的分解 $\alpha = \alpha_1 + \cdots + \alpha_s$ 并不唯一。

定义 6.1.1. $V = V_1 + \cdots + V_s$ 称为直和, 如果每个 $\alpha \in V$ 可唯一地写成 $\alpha = \alpha_1 + \cdots + \alpha_s$ $(\alpha_i \in V_i)$. 记为 $V = V_1 \oplus \cdots \oplus V_s$. □

下述命题给出了直和 $V=V_1\oplus\cdots\oplus V_s$ 的各种等价定义.

命题6.1.1. 设 $V=V_1+\cdots+V_s$，则下述结论等价.

(1) $V=V_1\oplus\cdots\oplus V_s$（即 $V_1+\cdots+V_s$ 是直和）.

(2) 若 $\alpha_1+\cdots+\alpha_s=0$（其中 $\alpha_i\in V_i$），则必有 $\alpha_1=\cdots=\alpha_s=0$.

(4) $(V_1+\cdots+V_{i-1}+V_{i+1}+\cdots+V_s)\cap V_i=\{0\}$. $(1\le i\le s)$.

(3) $\dim_K(V_1+\cdots+V_s)=\dim_K(V_1)+\cdots+\dim_K(V_s)$.

证明：将证明 (1) $\Rightarrow$ (2) $\Rightarrow$ (3) $\Rightarrow$ (4)

(1) $\Rightarrow$ (2). 因为零向量的分解唯一，故由 $0=\alpha_1+\cdots+\alpha_s$ 可得 $\alpha_1=\cdots=\alpha_s=0$.

(2) $\Rightarrow$ (3). 设 $\dim_K(V_i)=n_i$，$\alpha_{i1},\cdots,\alpha_{in_i}\in V_i$ 是 V_i 的一组基，则 $V_1+\cdots+V_s$ 由 $\alpha_{11},\cdots,\alpha_{1n_1},\cdots,\alpha_{s1},\cdots,\alpha_{sn_s}$ 生成. 由(2)易证 $\alpha_{11},\cdots,\alpha_{1n_1},\cdots,\alpha_{s1},\cdots,\alpha_{sn_s}$ 线性无关 故 $\dim_K(V_1+\cdots+V_s)=n_1+\cdots+n_s$.

(3) $\Rightarrow$ (4)：令 $W=V_1+\cdots+V_{i-1}+V_{i+1}+\cdots+V_s$，则

$$\dim_K(W)\le\dim_K(V_1)+\cdots+\dim_K(V_{i-1})+\dim_K(V_{i+1})+\cdots+\dim_K(V_s)$$

故 $\dim_K(V_1)+\cdots+\dim_K(V_s)\ge\dim_K(W)+\dim_K(V_i)\ge\dim_K(W+V_i)=\dim_K(V_1+\cdots+V_s)$

从而 $\dim_K(W+V_i)=\dim_K(W)+\dim_K(V_i)$. 故 $W\cap V_i=\{0\}$.

(4) $\Rightarrow$ (1)：只需证明：若 $\alpha_1+\cdots+\alpha_s=\beta_1+\cdots+\beta_s$.（其中 $\alpha_i,\beta_i\in V_i$）. 则必有 $\alpha_i=\beta_i$ $(i=1,2,\cdots,s)$. ~~由~~ $\cancel{\alpha_i-\beta_i=-(\alpha_1-\beta_1+\cdots+\alpha_{i-1}-\beta_{i-1}+\alpha_{i+1}-\beta_{i+1}+\cdots+\alpha_s-\beta_s)}$ ~~可得~~ $\cancel{\alpha_i-\beta_i\in V_i}$

由 $\alpha_i-\beta_i=-(\alpha_1-\beta_1+\cdots+\alpha_{i-1}-\beta_{i-1}+\alpha_{i+1}-\beta_{i+1}+\cdots+\alpha_s-\beta_s)$ 可得由

$$\alpha_i-\beta_i\in V_i\cap(V_1+\cdots+V_{i-1}+V_{i+1}+\cdots+V_s)=\{0\}.$$

□

研究线性算子 $V \xrightarrow{\mathcal{A}} V$ 的方法之一是考虑 $\mathcal{A}$ 在某个子空间 $W \subset V$ 上诱导的线性算子 $\mathcal{A}|_W$. 但这要求子空间 $W \subset V$ 满足如下条件.

<u>定义 6.1.2</u>. 子空间 $W \subset V$ 称为 $\mathcal{A}$ 的不变子空间. 如果 ~~$\mathcal{A}$~~ $\forall\, \alpha \in W$, 有 $\mathcal{A}\alpha \in W$ (即: $\mathcal{A}(W) \subseteq W$).

$\square$

例 6.1.1 (循环子空间). 设 $V \xrightarrow{\mathcal{A}} V$ 是线性算子. $\alpha \in V$. 则包含 α 的 <u>最小不变子空间</u> 是

$$K[\mathcal{A}]\cdot\alpha = \{ f(\mathcal{A})\alpha \mid \forall f(x) \in K[x] \} \subset V.$$

(称为由 α 生成的循环子空间).

$\square$

<u>命题 6.1.2</u>. 设 $V \xrightarrow{\mathcal{A}} V$ 是 n 维 K-向量空间 V 上的线性算子. 则 V 可分解为 $\mathcal{A}$ 的<u>不变子空间</u> $V_1, \cdots, V_s$ 的直和 $V = V_1 \oplus \cdots \oplus V_s$, ($\mathcal{A}(V_i) \subset V_i$, $\dim_K(V_i) = n_i$) 的充要条件是存在一组基 $e_1, \cdots, e_n$ 使

$$(\mathcal{A}e_1, \cdots, \mathcal{A}e_n) = (e_1, \cdots, e_n) \begin{pmatrix} A_1 & & 0 \\ & \ddots & \\ 0 & & A_s \end{pmatrix}$$

其中 $A_i \in M_{n_i}(K)$.

证明: "$\Rightarrow$": 若存在 $\mathcal{A}$ 的不变子空间 $V_1, \cdots, V_s$ 使

$$V = V_1 \oplus \cdots \oplus V_s \quad , \quad \dim_K(V_i) = n_i .$$

任取 V_i 的一组基 $\alpha_{i1}, \cdots, \alpha_{in_i}$, 令

$$(\mathcal{A}\alpha_{i1}, \cdots, \mathcal{A}\alpha_{in_i}) = (\alpha_{i1}, \cdots, \alpha_{in_i}) A_i \ , \ A_i \in M_{n_i}(K).$$

则 $\mathcal{A}$ 在基 $\alpha_{11},\cdots\alpha_{1n_1},\alpha_{21},\cdots\alpha_{2n_2},\cdots,\alpha_{s1},\cdots,\alpha_{sn_s}$ 下的坐标矩阵为

$$\begin{pmatrix} A_1 & & & O \\ & A_2 & & \\ & & \ddots & \\ O & & & A_s \end{pmatrix}.$$

"$\Leftarrow$"。若存在一组基 $e_1,\cdots,e_n\in V$ 使

$$(\mathcal{A}e_1,\cdots,\mathcal{A}e_n)=(e_1,\cdots,e_n)\begin{pmatrix} A_1 & & & O \\ & A_2 & & \\ & & \ddots & \\ O & & & A_s \end{pmatrix},$$

其中 $A_i\in M_{n_i}(k)$。令

$$\alpha_{11}=e_1,\cdots,\alpha_{1n_1}=e_{n_1},\ \alpha_{21}=e_{n_1+1},\cdots,\ \alpha_{2n_2}=e_{n_1+n_2},\cdots,$$
$$\alpha_{s1}=e_{n-n_s+1},\cdots,\alpha_{sn_s}=e_n,$$ 则

$$(\mathcal{A}\alpha_{i1},\mathcal{A}\alpha_{i2},\cdots,\mathcal{A}\alpha_{in_i})=(\alpha_{i1},\alpha_{i2},\cdots,\alpha_{in_i})A_i.$$

故 $V_i=\langle\alpha_{i1},\cdots,\alpha_{in_i}\rangle\subset V$ 是 $\mathcal{A}$ 的不变子空间。且

$$V=V_1\oplus V_2\oplus\cdots\oplus V_s$$

□

定义 6.1.3 设 $V\xrightarrow{\mathcal{A}}V$ 是线性算子。如果 V 不能分解为两个非平凡不变子空间的直和，则称 $(V,\mathcal{A})$ 是不可分解空间。

□

推论 6.1.1. 对任意 $\mathcal{A}\in\mathcal{L}(V)$，存在 $\mathcal{A}$ 的不变子空间 $V_i\subset V$ $(1\le i\le s)$ 使得 $V=V_1\oplus\cdots\oplus V_s$，且 $(V_i,\mathcal{A}|_{V_i})$ 是不可分解空间。

□

所以我们需要确定不可分解空间 $(V_i,\mathcal{A}|_{V_i})$ 的结构。本节最后讨论一类特殊的线性算子。

1维不变子空间是显然的不可分解子空间，它引出了线性代数中一个非常重要的概念.

定义 6.1.4. $\lambda\in K$ 称为 $\mathcal{A}\in\mathcal{L}(V)$ 的一个特征值，如果存在非零向量 $\alpha\in V$ 使 $\mathcal{A}\alpha=\lambda\alpha$。这样的 α 称为 $\mathcal{A}$ 的一个特征向量，而 $V_\lambda=\{v\in V\mid \mathcal{A}v=\lambda v\}$ 称为（特征值 λ 对应的）特征子空间。 □

对于给定的线性算子 $\mathcal{A}\in\mathcal{L}(V)$，一个有趣的问题是：$V$ 是否可以分解成 1 维不变子空间的直和？它等价于：是否存在一组基使 $\mathcal{A}$ 在该组基下的坐标矩阵是对角阵？在讨论该问题之前，我们引入线性算子 $\mathcal{A}\in\mathcal{L}(V)$ 的两个重要多项式：极小多项式与特征多项式.

设 $e_1,\cdots,e_n\in V$ 是一组基，$(\mathcal{A}e_1,\cdots,\mathcal{A}e_n)=(e_1,\cdots,e_n)A$

令 $\chi_{\mathcal{A}}(x)=|x\cdot I_n-A|\in K[x]$，$\mu_{\mathcal{A}}(x)=\mu_A(x)$，则有

命题 6.1.3：(1) $\chi_{\mathcal{A}}(x)$，$\mu_{\mathcal{A}}(x)$ 与基 $e_1,\cdots,e_n$ 的选取无关.

(2) $\mu_{\mathcal{A}}(\mathcal{A})=0$，如果 $f(x)\in K[x]$ 满足 $f(\mathcal{A})=0$，则必有 $\mu_{\mathcal{A}}(x)\mid f(x)$.

证明：(1) 设 $\mathcal{A}$ 在另一组基 $(e_1',\cdots,e_n')=(e_1,\cdots,e_n)T$ 下的坐标[illegible] A'，则 $A'=T^{-1}AT$，[illegible] $|xI_n-A'|=|T^{-1}(xI_n-A)T|$，

$\mu_A(A')=T^{-1}\mu_A(A)T=0$，从而 $\mu_{A'}(x)\mid\mu_A(x)$. 同理可证 $\mu_A(x)\mid\mu_{A'}(x)$.

(2) 对任意多项式 $f(x)\in K[x]$，由 $(f(\mathcal{A})e_1,\cdots,f(\mathcal{A})e_n)=(e_1,\cdots,e_n)f(A)$. 可得结论。 □.

多项式 $\chi_A(x)$ 和 $\mu_A(x)$ 分别称为线性算子 A 的特征多项式和极小多项式，它们都是研究 V 与 V 的主要工具。我们将分别利用 $\mu_A(x)$ 和 $\chi_A(x)$ 给出 ~~[illegible]~~ (V, A) 的不变子空间分解。

引理 6.1.1. 设 $A \in \mathcal{L}(V)$ 是 K-向量空间 V 上的线性算子。则下述结论等价：

(1) $\lambda \in K$ 是 A 的特征值。

(2) λ 是特征多项式 $\chi_A(x)$ 在 K 中的根。

(3) λ 是极小多项式 $\mu_A(x)$ 在 K 中的根。

证明：我们分别证明：(1) $\Leftrightarrow$ (2)，(1) $\Leftrightarrow$ (3)（从而 (2) $\Leftrightarrow$ (3)）。

设 $e_1, \cdots, e_n \in V$ 是一组基，若 $\lambda \in K$ 是 A 的特征向量，则存在 $\alpha = x_1e_1 + \cdots + x_ne_n$ 使 $A\alpha = \lambda\alpha$ 且 $x_1, \cdots, x_n$ 不全为零。

令 $(Ae_1, \cdots, Ae_n) = (e_1, \cdots, e_n)A$，则 $A\alpha = \lambda\alpha$ 等价于

$$(e_1, \cdots, e_n)(\lambda I_n - A)\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = 0.$$

即方程组 $(\lambda I_n - A)\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = 0$ 有非零解。故 $\chi_A(\lambda) = |\lambda I_n - A| = 0$.

反之，若 λ 是 $\chi_A(x) = 0$ 的根（即 $|\lambda I_n - A| = 0$），则

$$(\lambda I_n - A)\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = 0$$

有非零解 $(x_1, \cdots, x_n) \in K^n$。令 $\alpha = x_1e_1 + \cdots + x_ne_n$，则

$$\lambda\alpha - A\alpha = (e_1, \cdots, e_n)(\lambda I_n - A)\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = 0.$$

故 $\lambda \in K$ 是 A 的特征值。从而 (1) $\Leftrightarrow$ (2)。

(1) $\Leftrightarrow$ (3)，若 $\lambda \in K$ 是特征值，则存在非零向量 $\alpha \in V$ 使 $A\alpha = \lambda\alpha$.
故 $\mu_A(\lambda)\alpha = \mu_A(A)\alpha = 0$. 从而 $\mu_A(\lambda) = 0$.

反之，若 $\lambda\in K$ 是 $\mu_{\mathcal{A}}(x)$ 的根，即 $\mu_{\mathcal{A}}(\lambda)=0$。则

$$\mu_{\mathcal{A}}(x)=(x-\lambda)\,g(x)\quad,\ g(x)\in K[x].$$

因 $\deg g(x)<\deg \mu_{\mathcal{A}}(x)$，故 $g(\mathcal{A})\neq 0$。从而存在 $\beta\in V$ 使 $g(\mathcal{A})\beta\neq 0$

令 $\alpha=g(\mathcal{A})\beta\in V$，则 $(\mathcal{A}-\lambda)\alpha=(\mathcal{A}-\lambda)g(\mathcal{A})\cdot\beta=\mu_{\mathcal{A}}(\mathcal{A})\beta=0$。

即 $\mathcal{A}\alpha=\lambda\alpha$。由定义，$\lambda\in K$ 是 $\mathcal{A}$ 的特征值。

□

注记：对任意矩阵 $A\in M_n(K)$，$\chi_A(x)=|xI_n-A|$ 也称为矩阵 A 的特征多项式，它的根 $\lambda\in K$ 也称为矩阵 A 的特征值，即线性方程组

$$(\lambda I_n-A)\begin{pmatrix}x_1\\ \vdots\\ x_n\end{pmatrix}=0$$

的非零解也称为矩阵 A 的特征向量。

<u>定义 6.1.5</u>. $\mathcal{A}\in\mathcal{L}(V)$ 称为<u>可对角化算子</u>，如果 V 可分解成<u>1维不变子空间</u>的直和。对应地，矩阵 $A\in M_n(K)$ 称为<u>可对角化</u>矩阵，如果 A 相似于一个对角矩阵。

□

<u>定理 6.1.1</u>：$\mathcal{A}\in\mathcal{L}(V)$ 是<u>可对角化算子</u>的充分必要条件是 $\mu_{\mathcal{A}}(x)=(x-\lambda_1)\cdots(x-\lambda_s)$，其中 $\lambda_i\in K$ 互不相同。

证明："⟹"：若 $\mathcal{A}$ 可对角化，则存在一组基 $e_1,\dots,e_n\in V$ 使 $\mathcal{A}$ 在该组基下的坐标矩阵为

$$A=\begin{pmatrix}\overbrace{\lambda_1\ \cdots\ \lambda_1}^{n_1} & & & O\\ & \overbrace{\lambda_2\ \cdots\ \lambda_2}^{n_2} & & \\ & & \ddots & \\ O & & & \overbrace{\lambda_s\ \cdots\ \lambda_s}^{n_s}\end{pmatrix}_{n\times n}\qquad \lambda_i\neq\lambda_j\ (i\neq j).$$

令 $f(x)=(x-\lambda_1)(x-\lambda_2)\cdots(x-\lambda_s)$，则 $f(A)=0$。从而 $\mu_{\mathcal{A}}(x)\,|\,f(x)$.

但由引理6.1.1，$\lambda_1,\cdots,\lambda_s$ 是 $\mu_A(x)$ 的根，故 $f(x)\mid\mu_A(x)$。因此

$$\mu_A(x)=f(x)=(x-\lambda_1)\cdots(x-\lambda_s).$$

"$\Leftarrow$"：若 $\mu_A(x)=(x-\lambda_1)\cdots(x-\lambda_s)$，$\lambda_i\in K$ 互不相同，令

$$f_i(x)=\frac{\mu_A(x)}{x-\lambda_i}=(x-\lambda_1)\cdots(x-\lambda_{i-1})(x-\lambda_{i+1})\cdots(x-\lambda_s),\quad(1\le i\le s).$$

则 $f_1(x),\cdots,f_s(x)$ 互素。故存在 $u_i(x)\in K[x]$ 使

$$u_1(x)f_1(x)+\cdots+u_s(x)f_s(x)=1.$$

可得 $\mathcal{L}(V)$ 中恒等式：$u_1(A)f_1(A)+\cdots+u_s(A)f_s(A)=1$。

令 $V_i=u_i(A)f_i(A)(V)\subset V$，则 $V_i\subset V_{\lambda_i}$。且

$$V=V_1+\cdots+V_s\subset V_{\lambda_1}+\cdots+V_{\lambda_s}\subset V.$$

故 $V=V_1+\cdots+V_s=V_{\lambda_1}+\cdots+V_{\lambda_s}$。下面证明：$V_{\lambda_1}+\cdots+V_{\lambda_s}$ 是直和。$\forall\ \beta\in V_{\lambda_i}\cap(V_{\lambda_1}+\cdots+V_{\lambda_{i-1}}+V_{\lambda_{i+1}}+\cdots+V_{\lambda_s})$

则 $(A-\lambda_i)\beta=0$，$f_i(A)\beta=0$。但 $x-\lambda_i$ 与 $f_i(x)$ 互素，故存在 $u(x),v(x)\in K[x]$ 使 $u(x)(x-\lambda_i)+v(x)f_i(x)=1$。可得

$$\beta=\big(u(A)(A-\lambda_i)+v(A)f_i(A)\big)\cdot\beta=0.$$

因此 $V=V_{\lambda_1}\oplus\cdots\oplus V_{\lambda_s}$，其中特征子空间 V_{λ_i} 显然可以写成1维不变子空间的直和。由定义，A 可对角化。 □

<u>定理6.1.2</u>. 设 V 是 n 维 K-向量空间，则 $A\in\mathcal{L}(V)$ <u>可对角化</u>的充要条件是存在 $\lambda_i\in K$ 使得

$$\chi_A(x)=(x-\lambda_1)^{n_1}\cdots(x-\lambda_s)^{n_s},\quad\lambda_i\ne\lambda_j\ (i\ne j)$$

且 $n_i=\dim_K(V_{\lambda_i})$。

证明："$\Rightarrow$"：若 A 可对角化，则存在一组基 $e_1,\cdots,e_n$ 使 A

在该组基下的坐标矩阵为

$$A=\begin{pmatrix} \overbrace{\lambda_1 \cdots \lambda_1}^{n_1} & & & & 0 \\ & \overbrace{\lambda_2 \cdots \lambda_2}^{n_2} & & & \\ & & \ddots & & \\ 0 & & & \overbrace{\lambda_s \cdots \lambda_s}^{n_s} \end{pmatrix}_{n\times n},$$ 其中 $\lambda_i \neq \lambda_j \ (i\neq j)$ 且 $n_1+\cdots+n_s=n$.

所以 $\chi_A(x)=|xI_n-A|=(x-\lambda_1)^{n_1}\cdot(x-\lambda_2)^{n_2}\cdots(x-\lambda_s)^{n_s}$. 显然,

$$n_i \le \dim_K(V_{\lambda_i}) \qquad i=1,2,\cdots,s.$$

下面说明: $n_i \ge \dim_K(V_{\lambda_i})$. 设 $\dim_K(V_{\lambda_i})=m$, $\alpha_1,\cdots,\alpha_m$ 是 V_{λ_i} 的一组基. 将 $\alpha_1,\cdots,\alpha_m$ 扩充成 V 的一组基 $\alpha_1,\cdots,\alpha_m,\alpha_{m+1},\cdots,\alpha_n$. 则 A 在 $\alpha_1,\cdots,\alpha_m,\alpha_{m+1},\cdots,\alpha_n$ 下的坐标矩阵为

$$A'=\begin{pmatrix}\lambda_i I_m & A_{12}\\ 0 & A_{22}\end{pmatrix}.$$

故 $\chi_A(x)=|xI_n-A'|=(x-\lambda_i)^m\,|xI_{n-m}-A_{22}|$. 即 $(x-\lambda_i)^m \mid \chi_A(x)$

因此, $m\le n_i$. 从而 $n_i=\dim_K(V_{\lambda_i})$.

"$\Leftarrow$": 若 $\chi_A(x)=(x-\lambda_1)^{n_1}\cdots(x-\lambda_s)^{n_s}$, 其中 $n_i=\dim_K(V_{\lambda_i})$ 且 $\lambda_1,\cdots,\lambda_s\in K$ 互不相同, 则 $V_{\lambda_1}+\cdots+V_{\lambda_s}$ 必是直和且

$$\dim_K(V_{\lambda_1}\oplus\cdots\oplus V_{\lambda_s})=\dim_K(V_{\lambda_1})+\cdots+\dim_K(V_{\lambda_s})=n_1+\cdots+n_s=n$$

所以 $V=V_{\lambda_1}\oplus\cdots\oplus V_{\lambda_s}$. A 可对角化.

$\square$

注记: 若 $\lambda_1,\cdots,\lambda_s$ 是 A 的互不相同的特征值. $\alpha_i\in V_{\lambda_i}$. 若 $\alpha_1+\cdots+\alpha_s=0$ 则 $\lambda_1^k\alpha_1+\cdots+\lambda_s^k\alpha_s=0 \ (k\ge 0)$. 即

$$\begin{pmatrix}1 & 1 & \cdots & 1\\ \lambda_1 & \lambda_2 & \cdots & \lambda_s\\ \vdots & \vdots & & \vdots\\ \lambda_1^{s-1} & \lambda_2^{s-1} & \cdots & \lambda_s^{s-1}\end{pmatrix}\begin{pmatrix}\alpha_1\\ \vdots\\ \alpha_s\end{pmatrix}=0$$

由此可得 $\alpha_1=\alpha_2=\cdots=\alpha_s=0$.

习题 6.1

1. 设 $\lambda_1, \cdots, \lambda_s$ 是 ~~线性变换~~ $\mathcal{A} \in \mathcal{L}(V)$ 的互不相同的特征值. ~~设~~ $\alpha_i \in V_{\lambda_i}$, $i=1,2,\cdots,s$, 是特征向量, 证明: $\alpha_1, \cdots, \alpha_s$ 线性无关.

2. 设 $\mathcal{A} \in \mathcal{L}(V)$. 证明: (1) $\mathrm{Ker}(\mathcal{A})$, $\mathrm{Im}(\mathcal{A})$ 皆为 $\mathcal{A}$ 的不变子空间, (2) 若 V_1, V_2 是不变子空间, 则 V_1+V_2, $V_1 \cap V_2$ 也是不变子空间.

3. 设 V 是 n 维 ~~复~~ $\mathbb{C}$-向量空间, 则任意 $\mathcal{A} \in \mathcal{L}(V)$ 必有 1 维不变子空间.

4. 设 V 是 n 维 $\mathbb{R}$-向量空间, 则任意 $\mathcal{A} \in \mathcal{L}(V)$ 必有 1 维或 2 维不变子空间.

5. 设 V 是 n 维 $\mathbb{C}$-向量空间, 证明任意 $\mathcal{A} \in \mathcal{L}(V)$ 必有 $n-1$ 维不变~~量~~子空间 (提示: 若 $l \in V^*$ 是 $V^* \xrightarrow{\mathcal{A}^*} V^*$ 的特征向量, ~~则~~ 则 $W = \{v \in V \mid l(v)=0\} \subset V$ 是 $\mathcal{A}$ 的不变子空间).

6. 设 $\mathcal{A} \in \mathcal{L}(V)$, 如果存在自然数 s 使 $\mathrm{Im}(\mathcal{A}^s) = \mathrm{Im}(\mathcal{A}^{s+1})$. 证明: $V = \mathrm{Ker}(\mathcal{A}^s) \oplus \mathrm{Im}(\mathcal{A}^s)$ 是不变子空间的直和.

7. 证明: 对任意 $A, B \in M_n(\mathbb{R})$, 矩阵 AB 与 BA 有相同的特征多项式.

8. 设 $\mathcal{A} \in \mathcal{L}(V)$, $\mathcal{A}^* \in \mathcal{L}(V^*)$ 是 $\mathcal{A}$ 的对偶算子. 证明:

$$\mu_{\mathcal{A}}(x) = \mu_{\mathcal{A}^*}(x), \quad \chi_{\mathcal{A}}(x) = \chi_{\mathcal{A}^*}(x).$$

9. 设 V 是 n 维 K-向量空间, $\mathcal{A} \in \mathcal{L}(V)$, $\mathcal{A}^* \in \mathcal{L}(V^*)$ 是 $\mathcal{A}$ 的对偶算子. 若 $W \subset V^*$ 是 r 维子空间, 试证明:

(1) $W^{\perp} := \{v \in V \mid l(v)=0, \forall l \in W\} \subset V$ 是 ~~不变子空间~~ $n-r$ 维子空间.

(2) 若 W 是 $\mathcal{A}^*$ 的不变子空间, 则 $W^{\perp}$ 是 $\mathcal{A}$ 的不变子空间.

§6.2. 不可分解空间

设 V 是 n 维 K-向量空间，$W\subset V$ 是任意子空间，则一定存在子空间 $W'\subset V$ 使 $V=W\oplus W'$。一个自然的问题是：对任意不变子空间 $W\subset V$，是否一定存在不变子空间 $W'\subset V$ 使得 $V=W\oplus W'$？如果回答是肯定的，则我们至少可得结论：复向量空间上的线性算子总可以对角化。但事实并非如此！

例 6.2.1. 设 V 是 m 维复向量空间，$A\in\mathcal{L}(V)$ 的坐标矩阵

$$J_m(\lambda)=\begin{pmatrix}\lambda & 1 & & 0\\ & \lambda & 1 & \\ & & \ddots & 1\\ 0 & & & \lambda\end{pmatrix}_{m\times m}\qquad (m>1)$$

则 A 不可对角化（等价于 $J_m(\lambda)$ 不相似于对角矩阵）。 □

本节的主要任务是：对任意给的线性算子 $A\in\mathcal{L}(V)$，寻找 A 的不变子空间 $W_A\subset V$ 使得存在 A 的不变子空间 $W'\subset V$ 满足 $V=W_A\oplus W'$。它对于确定不可分解空间的结构至关重要。我们的主要工具是极小多项式 $\mu_A(x)$ 和它的下述推广。

定义 6.2.1. 设 $A\in\mathcal{L}(V)$，$\alpha\in V$ 非零。$\mu_{A,\alpha}(x)\in K[x]$ 是满足下述条件的多项式。

(1) $\mu_{A,\alpha}(x)$ 首项系数为 1，且 $\mu_{A,\alpha}(A)\cdot\alpha=0$。

(2) 若 $f(x)\in K[x]$ 满足 $f(A)\cdot\alpha=0$，则 $\mu_{A,\alpha}(x)\mid f(x)$。 □

定义中的多项式 $\mu_{A,\alpha}(x)$ 存在且唯一。它的唯一性可由定义得到。而存在性可由下述构造得到：令

$$S_{A,\alpha}=\{f(x)\in K[x]\mid f(A)\cdot\alpha=0\},$$

则 $S_{A,\alpha}$ 是非空集合（$\mu_A(x)\in S_{A,\alpha}$），故存在一个首项系数为 1，

次数最小的多项式 $\mu_{A,\alpha}(x)\in S_{A,\alpha}$. 由带余除法可验证：$\mu_{A,\alpha}(x)$ 满足定义的要求。

引理 6.2.1. 设 $\alpha\in V$ 非零，$A\in \mathcal{L}(V)$.

$$V_\alpha := K[A]\cdot\alpha = \{ f(A)\cdot\alpha \mid \forall f(x)\in K[x] \}\subset V.$$

则 V_α 是 A 的包含 α 的最小不变子空间，且满足：

(1) $\dim_K(V_\alpha)=\deg \mu_{A,\alpha}(x)$.

(2) $\mu_{A,\alpha}(x)$ 是 $A_\alpha=A|_{V_\alpha}: V_\alpha\to V_\alpha$ 的极小多项式.

证明：V_α 显然是 A 的不变子空间，且任何包含 α 的不变子空间 $W\subset V$ 必包含 $K[A]\cdot\alpha$，故 V_α 是包含 α 的最小不变子空间。

(1) 设 $\mu_{A,\alpha}(x)=x^m+a_1x^{m-1}+\cdots+a_{m-1}x+a_m\in K[x]$，则

(6.2.1) $\quad e_1=A^{m-1}\alpha,\ e_2=A^{m-2}\alpha,\ \cdots,\ e_{m-1}=A\alpha,\ e_m=\alpha$

是 $V_\alpha=K[A]\alpha$ 的一组基。事实上，$\forall\ \beta=f(A)\alpha\in V_\alpha$，存在 $g(x), r(x)$ 使 $f(x)=g(x)\mu_{A,\alpha}(x)+r(x)$，其中

$$r(x)=\lambda_1x^{m-1}+\lambda_2x^{m-2}+\cdots+\lambda_{m-1}x+\lambda_m\in K[x].$$

从而 $\beta=f(A)\alpha=g(A)\mu_{A,\alpha}(A)\cdot\alpha+r(A)\cdot\alpha=r(A)\cdot\alpha=\lambda_1e_1+\cdots+\lambda_me_m$.

所以 $V_\alpha=\langle e_1,\cdots,e_m\rangle$. $e_1,\cdots,e_m$ 的线性无关性可证明如下。

若 $\lambda_1e_1+\cdots+\lambda_me_m=0$，令 $f(x)=\lambda_1x^{m-1}+\lambda_2x^{m-2}+\cdots+\lambda_{m-1}x+\lambda_m\in K[x]$

则 $f(A)\cdot\alpha=\lambda_1e_1+\cdots+\lambda_me_m=0$. 故 $\mu_{A,\alpha}(x)\mid f(x)$. 但 $\deg\mu_{A,\alpha}(x)=m>m-1$，因此 $f(x)$ 是零多项式，即 $\lambda_1=\lambda_2=\cdots=\lambda_m=0$.

(2) $\mu_{A,\alpha}(x)$ 是 A_α 的极小多项式 $\mu_{A_\alpha}(x)$ 证明如下：

首先，$\mu_{A,\alpha}(A_\alpha)=0$，即 $\forall\ \beta=f(A)\alpha\in V_\alpha$，有

$$\mu_{A,\alpha}(A_\alpha)\beta=\mu_{A,\alpha}(A)\cdot\beta=f(A)\mu_{A,\alpha}(A)\alpha=0.$$

所以 $\mu_{A_\alpha}(x) \mid \mu_{A,\alpha}(x)$，（$\mu_{A_\alpha}(x)$ 的定义）. 其次，由 $\mu_{A_\alpha}(A_\alpha)=0$ 可得 $\mu_{A_\alpha}(A)\cdot\alpha=0$. 故由 $\mu_{A,\alpha}(x)$ 的定义可得 $\mu_{A,\alpha}(x) \mid \mu_{A_\alpha}(x)$. 因此,

$$\mu_{A_\alpha}(x)=\mu_{A,\alpha}(x).$$

□.

注记：在证明中由 (6.2.1) 定义的基

$$e_1=A^{m-1}\alpha,\ e_2=A^{m-2}\alpha,\ \cdots,\ e_{m-1}=A\alpha,\ e_m=\alpha$$

在以后的讨论中会经常出现，$A_\alpha: V_\alpha\to V_\alpha$ 在它下的坐标矩阵可以由 $\mu_{A,\alpha}(x)=x^m-a_1x^{m-1}-a_2x^{m-2}-\cdots-a_{m-1}x-a_m\in K[x]$ 确定如下：

$$(6.2.2).\qquad \begin{pmatrix} a_1 & 1 & 0 & \cdots & 0 \\ a_2 & 0 & 1 & \cdots & 0 \\ a_3 & 0 & 0 & \cdots & 0 \\ \vdots & & & & \\ a_{m-1} & 0 & 0 & \cdots & 1 \\ a_m & 0 & 0 & \cdots & 0 \end{pmatrix}_{m\times m}.$$

引理 6.2.2：设 $\mu_A(x)=p_1(x)^{m_1}\cdot p_2(x)^{m_2}\cdots p_s(x)^{m_s}$ $(m_i>0)$ 是 A 的极小多项式的不可约分解. 则

$$V=V_1\oplus\cdots\oplus V_s\ ,\quad A(V_i)\subset V_i,$$

且 $A_i=A|_{V_i}\in\mathcal{L}(V_i)$ 的极小多项式 $\mu_{A_i}(x)=p_i(x)^{m_i}$.

证明：令 $f_i(x)=p_1(x)^{m_1}\cdots p_{i-1}(x)^{m_{i-1}}\,p_{i+1}(x)^{m_{i+1}}\cdots p_s(x)^{m_s}$（即 $\mu_A(x)=p_i(x)^{m_i}f_i(x)$）. 则 $f_1(x),\cdots,f_s(x)$ 互素，故存在 $u_1(x),\cdots,u_s(x)\in K[x]$ 使

$$1=f_1(x)u_1(x)+\cdots+f_s(x)u_s(x).$$

令 $V_i=f_i(A)u_i(A)(V)=\{f_i(A)u_i(A)\cdot v \mid \forall v\in V\}\subset V$.

则 $V_i\subset V$ 是 A 的不变子空间，且 $V=V_1+\cdots+V_s$.

首先证明 $\mu_{A_i}(x)=p_i(x)^{m_i}$. 事实上，$p_i(A)^{m_i}(V_i)=u_i(A)\mu_A(A)(V)=0$,

即 $P_i(A_i)^{m_i}=0$，故 $\mu_{A_i}(x)\mid P_i(x)^{m_i}$. 从而 $\mu_{A_i}(x)=P_i(x)^{m_i'}$，$m_i'\le m_i$.

若 $m_i'<m_i$，令 $g(x)=f_i(x)\cdot P_i(x)^{m_i'}$，则 $\deg g(x)<\deg \mu_A(x)$. 但由

$$g(A)(V)=g(A)(V_1)+\cdots+g(A)(V_s)=0 \quad (注意: f_i(A)f_j(A)=0)$$

可得 $g(A)=0$ 与 $\mu_A(x)$ 的定义矛盾. 故 $\mu_{A_i}(x)=P_i(x)^{m_i}$.

最后证明 $V_1+\cdots+V_s$ 是直和. 令 $W_i=V_1+\cdots+V_{i-1}+V_{i+1}+\cdots+V_s$，则 $f_i(A)(W_i)=0$ (因为 $f_i(A)f_j(A)=0$, $(i\neq j)$), $P_i(A)^{m_i}(V_i)=0$

由于 $f_i(x)$ 与 $P_i(x)^{m_i}$ 互素，存在 $u(x),v(x)\in K[x]$ 使

$$u(x)f_i(x)+v(x)P_i(x)^{m_i}=1.$$

所以，$\forall\beta\in W_i\cap V_i$，有 $\beta=u(A)f_i(A)\cdot\beta+v(A)P_i(A)^{m_i}\cdot\beta=0$. □

<u>定理 6.2.1</u>. 对任意 $A\in \mathcal{L}(V)$，存在 $\alpha\in V$ 使得

$$\mu_{A,\alpha}(x)=\mu_A(x).$$

证明：设 $\mu_A(x)=P_1(x)^{m_1}\cdots P_s(x)^{m_s}$ $(m_i>0)$ 是 $\mu_A(x)$ 的不可约分解. 由引理 6.2.2，$V=V_1\oplus\cdots\oplus V_s$，$A(V_i)\subset V_i$，且 $A_i=A|V_i$ 的极小多项式 $\mu_{A_i}(x)=P_i(x)^{m_i}$. 若 $m_i=1$，$\forall\alpha_i\in V_i$ 非零，有 $\mu_{A,\alpha_i}(x)\mid\mu_{A_i}(x)$，故 $\mu_{A,\alpha_i}(x)=P_i(x)=\mu_{A_i}(x)$ ($\alpha_i\neq0$，故 $\deg\mu_{A,\alpha_i}(x)>0$).
若 $m_i>1$，则 $P_i(A_i)^{m_i-1}\neq0$，因此存在 $\alpha_i\in V_i$ 使 $P_i(A)^{m_i-1}\alpha_i\neq0$. 从而 $\mu_{A,\alpha_i}(x)=P_i(x)^{m_i}$. 总之，存在 $\alpha_i\in V_i$ 使 $\mu_{A,\alpha_i}(x)=P_i(x)^{m_i}$.

令 $\alpha=\alpha_1+\cdots+\alpha_s$，由 $0=\mu_{A,\alpha}(A)\cdot\alpha=\mu_{A,\alpha}(A)\cdot\alpha_1+\cdots+\mu_{A,\alpha}(A)\cdot\alpha_s$
得 $\mu_{A,\alpha}(A)\cdot\alpha_i=0$ $(1\le i\le s)$，故 $P_i(x)^{m_i}\mid\mu_{A,\alpha}(x)$ $(1\le i\le s)$.
所以 $\mu_A(x)\mid\mu_{A,\alpha}(x)$，$\mu_{A,\alpha}(x)\mid\mu_A(x)$. 因此

$$\mu_{A,\alpha}(x)=\mu_A(x).$$

□

推论 6.2.1. 设 $\mathcal{A}\in\mathcal{L}(V)$，则 $\deg \mu_{\mathcal{A}}(x) \le \dim_K(V)$，且等号成立的充要条件是存在 $\alpha\in V$ 使得

$$V = K[\mathcal{A}]\cdot\alpha .$$

证明：由定理6.2.1，存在 $\alpha\in V$ 使 $\mu_{\mathcal{A},\alpha}(x)=\mu_{\mathcal{A}}(x)$. 故

$$\deg\mu_{\mathcal{A}}(x) = \deg\mu_{\mathcal{A},\alpha}(x) = \dim_K(K[\mathcal{A}]\cdot\alpha) \le \dim_K(V).$$

若等号成立，显然 $V_\alpha := K[\mathcal{A}]\cdot\alpha = V$. 反之，若存在 $\alpha\in V$ 使 $V=K[\mathcal{A}]\cdot\alpha$. 则 $\mu_{\mathcal{A}}(x)=\mu_{\mathcal{A},\alpha}(x)$. 因此

$$\deg\mu_{\mathcal{A}}(x) = \deg\mu_{\mathcal{A},\alpha}(x) = \dim_K K[\mathcal{A}]\cdot\alpha = \dim_K(V).$$

□

推论 6.2.2. 设 $\mathcal{A}\in\mathcal{L}(V)$，如果 $(V,\mathcal{A})$ 满足

(1) $\mu_{\mathcal{A}}(x)=p(x)^m$，$p(x)\in K[x]$ 不可约.

(2) $\dim_K V = \deg\mu_{\mathcal{A}}(x)$.

则 $(V,\mathcal{A})$ 是不可分解空间.

证明：若 $V=V_1\oplus V_2$，其中 V_1, V_2 是 $\mathcal{A}$ 的不变子空间. 令 $\mathcal{A}_i=\mathcal{A}|_{V_i}: V_i\to V_i$ $(i=1,2)$. 则 $\mu_{\mathcal{A}_1}(x)=p(x)^{m_1}$，$\mu_{\mathcal{A}_2}(x)=p(x)^{m_2}$ 其中 $m_i\le m$ $(i=1,2)$. 不妨设 $m_1\ge m_2$，则 $p(\mathcal{A})^{m_1}(V_1)=0$，$p(\mathcal{A})^{m_1}(V_2)=0$. 从而 $p(\mathcal{A})^{m_1}(V)=0$. 因此，$\mu_{\mathcal{A}_1}(x)=\mu_{\mathcal{A}}(x)$. 由条件(2)，

$$\dim_K(V_1)\ge\deg\mu_{\mathcal{A}_1}(x)=\deg\mu_{\mathcal{A}}(x)=\dim_K(V).$$

所以 $V_1=V$，即 $(V,\mathcal{A})$ 不可分解.

□

设 $\mathcal{A}\in\mathcal{L}(V)$，$\alpha\in V$ 使得 $\mu_{\mathcal{A},\alpha}(x)=\mu_{\mathcal{A}}(x)$. 我们将证明 $V_\alpha=K[\mathcal{A}]\cdot\alpha\subset V$ 就是我们要找的 $W_{\mathcal{A}}$. 即存在 $\mathcal{A}$ 的不变子空间 $W^*\subset V$ 使 $V=V_\alpha\oplus W^*$.

引理 6.2.3. 对任意非零向量 $\alpha\in V$，令 $V_\alpha=K[\mathcal{A}]\cdot\alpha$，我们有如下结论

(1) 存在 $l_\alpha\in V_\alpha^*$ 使 $V_\alpha^*=K[\mathcal{A}_\alpha^*]\cdot l_\alpha$，其中 $\mathcal{A}_\alpha^*: V_\alpha^*\to V_\alpha^*$ 是 $\mathcal{A}_\alpha=\mathcal{A}|_{V_\alpha}: V_\alpha\to V_\alpha$ 的对偶算子。

(2) 设 $\dim_K(V_\alpha)=m$，$l\in V^*$ 使 $l|_{V_\alpha}=l_\alpha$，令

$$W_*=\langle l, \mathcal{A}^*(l), \mathcal{A}^{*^2}(l), \cdots, \mathcal{A}^{*^{m-1}}(l)\rangle\subset V^*,$$

则 $V=V_\alpha\oplus W_*^\perp$，其中 $W_*^\perp=\{v\in V \mid f(v)=0, \forall f\in W_*\}$.

(3) 若 $W_*\subset V^*$ 是对偶算子 $\mathcal{A}^*: V^*\to V^*$ 的不变子空间，则 $W_*^\perp\subset V$ 是 $\mathcal{A}$ 的不变子空间。

(4) 若 $\alpha\in V$ 满足 $\mu_{\mathcal{A},\alpha}(x)=\mu_{\mathcal{A}}(x)$，则

$$V=V_\alpha\oplus W_*^\perp,\quad \mathcal{A}(W_*^\perp)\subset W_*^\perp.$$

证明. (1) 因为 $\dim_K(V_\alpha^*)=\dim_K(V_\alpha)=\deg\mu_{\mathcal{A}_\alpha}(x)=\deg\mu_{\mathcal{A}_\alpha^*}(x)$，由推论 6.2.1 的 可知 $V_\alpha^*=K[\mathcal{A}_\alpha^*]\cdot l_\alpha$，其中 $l_\alpha\in V_\alpha^*$.

(2) 若 $\dim_K(V_\alpha)=m$，则 $\dim_K(V_\alpha^*)=m$，且

$$l_\alpha,\ \mathcal{A}_\alpha^*(l_\alpha),\ \mathcal{A}_\alpha^{*^2}(l_\alpha),\cdots,\ \mathcal{A}_\alpha^{*^{m-1}}(l_\alpha)$$

是 V_α^* 的一组基。由 $l|_{V_\alpha}=l_\alpha$ 可得 $\mathcal{A}^{*^k}(l)|_{V_\alpha}=\mathcal{A}_\alpha^{*^k}(l_\alpha)$，故

$$W_*=\langle l, \mathcal{A}^*(l), \mathcal{A}^{*^2}(l), \cdots, \mathcal{A}^{*^{m-1}}(l)\rangle\subset V^*$$

是 m 维子空间。由线性方程组理论，$W_*^\perp\subset V$ 是 $\dim_K(V)-m$ 维子空间。 $\forall\ \beta\in V_\alpha\cap W_*^\perp$，由 $W_*^\perp$ 的定义可知

$$\mathcal{A}_\alpha^{*^k}(l_\alpha)(\beta)=\mathcal{A}^{*^k}(l)(\beta)=0,\quad 0\le k\le m-1.$$

因此 $\beta=0$，从而 $\dim_K(V_\alpha+W_*^\perp)=\dim_K(V_\alpha)+\dim_K(W_*^\perp)=\dim_K(V)$.
故 $V=V_\alpha\oplus W_*^\perp$.

(3) 若 $W_*\subset V^*$ 是 $V^*\xrightarrow{\mathcal{A}^*}V^*$ 的不变子空间，则对任意 $f\in W_*$ 有 $\mathcal{A}^*(f)\in W_*$。 $\forall\ v\in W_*^\perp$，有 $f(\mathcal{A}v)=\mathcal{A}^*(f)(v)=0$，$\forall f\in W_*$

故 $A\cdot v\in W_*^{\perp}$. 即 $W_*^{\perp}\subset V$ 是 A 的不变子空间.

(4). 若 $\mu_{A,\alpha}(x)=\mu_A(x)$. 设 $\deg\mu_{A,\alpha}(x)=\dim_K(V_\alpha)=m$. 由 $m=\dim_K(W_*)\le\dim_K K[A^*]\cdot l=\deg\mu_{A^*,l}(x)\le\deg\mu_{A^*}(x)=m$ 可得 $W_*=K[A^*]\cdot l\subset V^*$ 是 A^* 的不变子空间.

□

<u>定理 6.2.2</u>: 设 $A\in\mathcal{L}(V)$, 则 (V,A) 是不可分解空间的充要条件是 (1) $\mu_A(x)=p(x)^m$, $p(x)\in K[x]$ 不可约;

(2) $\dim_K(V)=\deg\mu_A(x)$.

证明: 由推论 6.2.2, 若条件 (1), (2) 成立, 则 (V,A) 是不可分解空间。若 (V,A) 不可分解, 由引理 6.2.3 中的结论 (4), 可得 $V=K[A]\cdot\alpha$, $\mu_A(x)=\mu_{A,\alpha}(x)$. 故 $\deg\mu_A(x)=\dim_K(V)$. 由引理 6.2.2 可知 $\mu_A(x)=p(x)^m$, $p(x)\in K[x]$.

□

<u>推论 6.2.3</u>: 设 $A\in\mathcal{L}(V)$, 则存在 $v_1,\cdots,v_m\in V$ 使得

$$V=K[A]\cdot v_1\oplus\cdots\oplus K[A]\cdot v_m \qquad (6.2.3)$$

其中 $\mu_{A,v_i}(x)=p_i(x)^{m_i}$, $p_i(x)\in K[x]$ 不可约。

证明: 推论 6.1.1, $V=V_1\oplus\cdots\oplus V_m$, 其中 V_i $(1\le i\le m)$ 是 A 的不变子空间. 且 $(V_i, A|V_i)$ 不可分解. 由定理 6.2.2, 存在 $v_i\in V_i$ 使

$$V_i=K[A|V_i]\cdot v_i=K[A]\cdot v_i$$

且 $\mu_{A,v_i}(x)=\mu_{A_i}(x)=p_i(x)^{m_i}$ (其中 $A_i=A|V_i$).

□

不可分解空间 (V,A) 的结构显然与 K 密切相关. K 越大, $K[x]$ 中的不可约多项式越简单. 例如, 当 K 是代数闭域(即 $K=\bar{K}$)时, $K[x]$ 中不可约多项式仅为一次多项式. 特别当 $K=\mathbb{C}$ 时, 有

推论 6.2.4. 设 V 是 n 维复向量空间, $\mathcal{A}\in\mathcal{L}(V)$. 则存在一组基 ~~……~~ 使 $\mathcal{A}$ 在该基下的坐标矩阵等于

$$J(\mathcal{A})=\begin{pmatrix} J_{m_1}(\lambda_1) & & & 0 \\ & J_{m_2}(\lambda_2) & & \\ & & \ddots & \\ 0 & & & J_{m_s}(\lambda_s)\end{pmatrix},\quad J_{m_i}(\lambda_i)=\begin{pmatrix}\lambda_i & 1 & & 0\\ & \lambda_i & 1 & \\ & & \ddots & 1 \\ 0 & & & \lambda_i\end{pmatrix}_{m_i\times m_i}.$$

($J_{m_i}(\lambda_i)$ 称为特征值为 λ_i 的 m_i 阶 Jordan 块, $J(\mathcal{A})$ 称为 Jordan 标准型).

证明: 由推论 6.2.3, $V=K[\mathcal{A}]\cdot v_1\oplus\cdots\oplus K[\mathcal{A}]\cdot v_s$, 其中 $\mu_{\mathcal{A},v_i}(x)=(x-\lambda_i)^{m_i}$. 在 $V_i=K[\mathcal{A}]\cdot v_i$ 中取基

$$e_1=(\mathcal{A}-\lambda_i)^{m_i-1}v_i,\ e_2=(\mathcal{A}-\lambda_i)^{m_i-2}v_i,\ \cdots,\ e_{m_i-1}=(\mathcal{A}-\lambda_i)v_i,\ e_{m_i}=v_i,$$

则 $\mathcal{A}e_j=\lambda_i e_j+e_{j-1}\ (1\le j\le m_i)$, $\mathcal{A}_i=\mathcal{A}|_{V_i}$ 在该组基下的坐标矩阵为 $J_{m_i}(\lambda_i)$. □

推论 6.2.5. 设 $(V,\mathcal{A})$ 不可分解, $\mu_{\mathcal{A}}(x)=p(x)^m$,

$$p(x)=x^p-a_{p-1}x^{p-1}-\cdots-a_1x-a_0\in K[x]$$

是不可约多项式. 则存在一组基使 $\mathcal{A}$ 在该基下的坐标矩阵为

$$J_m(\mathcal{A})=\begin{pmatrix}\boxed{A_1} & & & & 0\\ 1 & \boxed{A_2} & & & \\ & 1 & \boxed{A_3} & & \\ & & & \ddots & \\ 0 & & & 1 & \boxed{A_m}\end{pmatrix},\quad 其中\ A_i=\begin{pmatrix}0&0&0&\cdots&0&a_0\\1&0&0&\cdots&0&a_1\\0&1&0&\cdots&0&a_2\\\vdots&\vdots&\vdots&&\vdots&\vdots\\0&0&0&\cdots&1&a_{p-1}\end{pmatrix}$$

证明: $(V,\mathcal{A})$ 不可分解, 故存在 $\alpha\in V$ 使 $V=K[\mathcal{A}]\cdot\alpha$. 令

$$e_0=\alpha,\ e_1=\mathcal{A}\alpha,\ e_2=\mathcal{A}^2\alpha,\ \cdots,\ e_{p-1}=\mathcal{A}^{p-1}\alpha$$

则 $\mathcal{A}e_0=e_1$, $\mathcal{A}e_1=e_2,\cdots$, $\mathcal{A}e_{p-2}=e_{p-1}$, $\mathcal{A}e_{p-1}=a_0e_1+a_1e_2+\cdots+a_{p-1}e_{p-1}+p(\mathcal{A})e_0$

断言: $\underbrace{e_0,e_1,\cdots,e_{p-1}}$, $\underbrace{p(A)e_0,p(A)e_1,\cdots,p(A)e_{p-1}}$, $\underbrace{p(A)^2e_0,p(A)^2e_1,\cdots,p(A)^2e_{p-1}}$

$\cdots\cdots$ $\underbrace{p(A)^{m-1}e_0,p(A)^{m-1}e_1,\cdots,p(A)^{m-1}e_{p-1}}$ (共 mp 个向量) 是 V 的一组基.

由于 $\dim_K(V)=mp$, 故只需证明上述向量生成 V. 但 V 是包含 $e_0:=\alpha$ 的$\underbrace{\text{最小不变子空间}}$, 故只需证明

$$W:=\langle e_0,\cdots,e_{p-1},p(A)e_0,\cdots,p(A)e_{p-1},\cdots,p(A)^{m-1}e_0,\cdots,p(A)^{m-1}e_{p-1}\rangle$$

是不变子空间。令 $e_{kp}=p(A)^ke_0$, $e_{kp+1}=p(A)^ke_1,\cdots,e_{(k+1)p-1}=p(A)^ke_{p-1}$ $(0\le k\le m-1)$. 利用 $Ae_i=e_{i+1}$, $Ae_{p-1}=a_0e_0+a_1e_1+\cdots+a_{p-1}e_{p-1}+e_p$,

得: $Ae_{kp+i}=A(p(A)^ke_i)=p(A)^ke_{i+1}=e_{kp+i+1}$. $(0\le i<p-1)$

$$Ae_{kp+p-1}(=Ae_{(k+1)p-1})=p(A)^kAe_{p-1}=a_0p(A)^ke_0+a_1p(A)^ke_1+\cdots+a_{p-1}p(A)^ke_{p-1}+p(A)^ke_p$$

$$=a_0e_{kp}+a_1e_{kp+1}+\cdots+a_{p-1}e_{(k+1)p-1}+e_{(k+1)p}.$$

所以 $\langle e_0,e_1,\cdots,e_{p-1},e_p,e_{p+1},\cdots,e_{2p-1},e_{2p},e_{2p+1},\cdots e_{3p-1},\cdots,e_{(m-1)p},\cdots,e_{mp-1}\rangle$ 是 V 的一个包含 e_0 的不变子空间. 从而 $\{e_{kp+i}\}_{0\le k\le m-1,0\le i\le p-1}$ 是 V 的一组基. A 在该组基下的坐标矩阵是 $J_m(A)$。例如:

$$(Ae_0,Ae_1,\cdots,Ae_{p-2},Ae_{p-1})=(e_0,\cdots,e_{p-1},e_p)\begin{pmatrix}0&0&\cdots&0&a_0\\1&0&\cdots&0&a_1\\0&1&\cdots&0&a_2\\0&0&\cdots&0&a_3\\\vdots&\vdots&&\vdots&\vdots\\0&0&\cdots&1&a_{p-1}\\0&0&\cdots&0&1\end{pmatrix}_{(p+1)\times p}$$

即有

$$(Ae_0,\cdots,Ae_{p-1},Ae_p,\cdots,Ae_{2p-1})=(e_0,\cdots,e_{p-1},e_p,\cdots,e_{2p-2},e_{2p-1},e_{2p})A,\text{ 其中}$$

$$A=\left(\begin{array}{ccccc|ccccc}0&0&\cdots&0&a_0&&&&&\\1&0&\cdots&0&a_1&&&&&\\0&1&\cdots&0&a_2&&&\bigcirc&&\\0&0&\cdots&0&a_3&&&&&\\\vdots&\vdots&&\vdots&\vdots&&&&&\\0&0&\cdots&1&a_{p-1}&&&&&\\\hline0&0&\cdots&0&1&0&0&\cdots&0&a_0\\&&&&&1&0&\cdots&0&a_1\\&&&&&0&1&\cdots&0&a_2\\&&\bigcirc&&&\vdots&\vdots&&\vdots&\vdots\\&&&&&0&0&\cdots&1&a_{p-1}\\&&&&&0&0&\cdots&0&1\end{array}\right)$$

$\square$

对任意 $\mathcal{A}\in\mathcal{L}(V)$，$V$ 总可以写不可分解子空间 $(V_i,\mathcal{A}|_{V_i})$ 的直和，然后应用定理6.2.2中关于不可分解空间 $(V_i,\mathcal{A}|_{V_i})$ 的结构得到推论6.2.3中的分解(6.2.3)。另一方面，引理6.2.3中的结论(4)提供了满足推论6.2.3中条件的另一分解。事实上，引理6.2.3中的结论(4)有下述直接推论。

推论6.2.6. 设 $\mathcal{A}\in\mathcal{L}(V)$，则存在 $z_i\in V$ $(1\le i\le s)$ 使得

$$V=K[\mathcal{A}]\cdot z_1\oplus\cdots\oplus K[\mathcal{A}]\cdot z_s \tag{6.2.4}$$

其中 $d_i(x):=\mu_{\mathcal{A},z_i}(x)$ 满足 $d_i(x)\mid d_{i+1}(x)$。

证明：由定理6.2.1，存在 $z_s\in V$ 使 $\mu_{\mathcal{A},z_s}(x)=\mu_{\mathcal{A}}(x)$。由引理6.2.3中结论(4)，存在不变子空间 $V'\subset V$ 使得

$$V=V'\oplus K[\mathcal{A}]\cdot z_s\quad,\quad d_s(x):=\mu_{\mathcal{A},z_s}(x)=\mu_{\mathcal{A}}(x).$$

令 $\mathcal{A}'=\mathcal{A}|_{V'}\in\mathcal{L}(V')$，重复应用定理6.2.1和引理6.2.3中结论(4)于 $(V',\mathcal{A}')$ 可得 $V=K[\mathcal{A}]\cdot z_1\oplus\cdots\oplus K[\mathcal{A}]\cdot z_{s-1}\oplus K[\mathcal{A}]\cdot z_s$，其中 $d_i(x):=\mu_{\mathcal{A},z_i}(x)$ 满足条件 $d_i(x)\mid d_{i+1}(x)$.（注意：$\mu_{\mathcal{A}'}(x)\mid\mu_{\mathcal{A}}(x)=d_s(x)$）。 □

对推论6.2.6中的分解(6.2.4)应用下述引理可得到另一个满足推论6.2.3中条件的分解。

引理6.2.4. 设 $\mathcal{A}\in\mathcal{L}(W)$，$W=K[\mathcal{A}]\cdot z$。如果

$$\mu_{\mathcal{A},z}(x)=f(x)\cdot g(x)\ \text{且}\ (f(x),g(x))=1,$$

则存在 $\alpha,\beta\in W$ 使 $W=K[\mathcal{A}]\cdot\alpha\oplus K[\mathcal{A}]\cdot\beta$，且

$$\mu_{\mathcal{A},\alpha}(x)=f(x)\ ,\quad \mu_{\mathcal{A},\beta}(x)=g(x).$$

证明：由 $(f(x),g(x))=1$ 知存在 $u(x),v(x)\in K[x]$ 使 $f(x)u(x)+g(x)v(x)=1$.

故 $z=g(\mathcal{A})v(\mathcal{A})\cdot z+f(\mathcal{A})\cdot u(\mathcal{A})\cdot z=\alpha+\beta$，其中 $\alpha=g(\mathcal{A})\cdot v(\mathcal{A})\cdot z$，

$\beta = f(A)u(A)\cdot z$，因此，$W = k[A]\cdot z = k[A]\cdot\alpha + k[A]\cdot\beta$。由 $f(A)\cdot\alpha = 0$，$g(A)\cdot\beta = 0$。可得 $k[A]\alpha \cap k[A]\cdot\beta = \{0\}$，故

$$W = k[A]\cdot\alpha \oplus k[A]\cdot\beta,\ \text{且}\ \mu_{A,\alpha}(x) \mid f(x),\ \mu_{A,\beta}(x) \mid g(x).$$

另一方面，由 $\deg f(x) + \deg g(x) = \deg \mu_{A,z}(x) = \dim_k(W) = \deg \mu_{A,\alpha}(x) + \deg \mu_{A,\beta}(x)$
可得 $\deg f(x) = \deg \mu_{A,\alpha}(x)$，$\deg g(x) = \deg \mu_{A,\beta}(x)$。所以

$$\mu_{A,\alpha}(x) = f(x),\quad \mu_{A,\beta}(x) = g(x).$$

□.

下面我们证明满足定理6.2.3中条件的分解是唯一的。

定理6.2.3（不变性定理）。对任意 $A \in \mathcal{L}(V)$，如果分解

$$V = k[A]\cdot z_1 \oplus \cdots \oplus k[A]\cdot z_s = k[A]\cdot w_1 \oplus \cdots \oplus k[A]\cdot w_t$$

满足 $\mu_{A,z_i}(x) = p_i(x)^{n_i}$，$\mu_{A,w_j}(x) = q_j(x)^{m_j}$（$1\le i\le s$，$1\le j\le t$）其中 $p_i(x), q_j(x) \in k[x]$ 是不可约多项式，则 $s=t$，且适当调整次序后有 $p_i(x)^{n_i} = q_i(x)^{m_i}$。

证明：设 $p(x)\in k[x]$ 是不可约多项式，令 $M_{p(x)} := \bigoplus_{p_i(x)=p(x)} k[A]\cdot z_i \subset V$，
$N_{p(x)} = \bigoplus_{q_j(x)=p(x)} k[A]\cdot w_j \subset V$，则 $V = M_{p(x)} \oplus M'_{p(x)} = N_{p(x)} \oplus N'_{p(x)}$。其中

$$M'_{p(x)} = \bigoplus_{p_i(x)\neq p(x)} k[A]\cdot z_i,\quad N'_{p(x)} = \bigoplus_{q_j(x)\neq p(x)} k[A]\cdot w_j.$$

不难证明 $M_{p(x)} = N_{p(x)}$：$\forall \alpha \in M_{p(x)}$，存在 $\alpha_1 \in N_{p(x)}$，$\alpha_2 \in N'_{p(x)}$ 使 $\alpha = \alpha_1 + \alpha_2$。必存在 k 使 $p(A)^k\cdot\alpha_2 = p(A)^k\cdot(\alpha-\alpha_1) = 0$ 以及 $f(A)\cdot\alpha_2 = 0$ 其中 $f(x) = \prod_{q_j(x)\neq p(x)} q_j(x)^{m_j}$，且 $(p(x)^k, f(x)) = 1$，故 $\alpha_2 = 0$。所以 $M_{p(x)} \subseteq N_{p(x)}$。
同理 $N_{p(x)} \subseteq M_{p(x)}$。因此，在定理6.2.3中不妨假设 $p_1(x) = \cdots = p_s(x) = p(x)$，$q_1(x) = \cdots = q_t(x) = p(x)$，且 $n_1 \le n_2 \le \cdots \le n_s$，$m_1 \le m_2 \le \cdots \le m_t$。此时不难说

~~$E = \bigoplus_{n_i=n_s} k[A]\cdot z_i$，$W = \bigoplus_{m_j=m_t} k[A]\cdot w_j$，$E' = \bigoplus_{n_i<n_s} k[A]\cdot z_i$，$W' = \bigoplus_{m_j<m_t} k[A]\cdot w_j$。~~

~~则 $V = E' \oplus E = W' \oplus W$~~ 明 $\mu_A(x) = p(x)^{n_s}$，$\mu_A(x) = p(x)^{m_t}$。故 $n_s = m_t = m$。

对任意 $1\le k\le m$，令 $N_k=\#\{i \mid n_i=k\}$，$N_k'=\#\{j \mid m_j=k\}$.

若 $N_k=N_k'$ 对任意 $1\le k\le m$ 成立，则 $s=t$，$n_i=m_i$ $(1\le i\le s)$.

我们将证明 N_k 与分解无关，它们事实上由线性算子 $p(A)^j$ $(1\le j\le m)$ 的秩唯一确定！

对任意 $1\le k\le m$，令 $E_k=\bigoplus_{n_i=k}k[A]\cdot z_i$，则 $V=\bigoplus_{k=1}^{m}E_k$，考虑线性算子 $p(A): V=\bigoplus_{k=1}^{m}E_k \longrightarrow V=\bigoplus_{k=1}^{m}E_k$，$Im(p(A))=\bigoplus_{k=1}^{m}p(A)\cdot E_k$.

由于 $p(A)\cdot E_k=\bigoplus_{n_i=k}k[A]\,p(A)\cdot z_i$，$\dim_k k[A]\cdot p(A)\cdot z_i=(k-1)\deg p(x)$，故

$$r(p(A))=\deg p(x)\sum_{k=2}^{m}(k-1)N_k .$$

同理可得公式 ($j>m$，则 $N_j=0$，$p(A)^j=0$).

$$(6.2.5)\quad r(p(A)^j)=\deg p(x)\sum_{k=j+1}^{m}(k-j)N_k ,\qquad j=1,2,\cdots,m.$$

特别地，对任意 $1\le j\le m$，我们有 (注意 $p(A)^0=id$):

$$r(p(A)^{j+1})=\deg p(x)\sum_{k=j+2}^{m}(k-j-1)N_k,\quad r(p(A)^{j-1})=p\sum_{k=j}^{m}(k-j+1)N_k .$$

故 $r(p(A)^{j+1})+r(p(A)^{j-1})=N_j\deg p(x)+2\,r(p(A)^j)$. 因此，

$$(6.2.6)\qquad N_j=\frac{1}{\deg p(x)}\left[r(p(A)^{j+1})+r(p(A)^{j-1})-2\,r(p(A)^j)\right],\quad 1\le j\le m.$$

$\Box$

习题 6.2

1. 设 $\dim_K(V)=n$，$A\in \mathcal{L}(V)$. 试证明：

(1) V 是循环空间当且仅当存在一组基 $e_0,e_1,\cdots,e_{n-1}$ 使
$$Ae_i=e_{i+1},\quad 0\le i<n-1.$$

(2) 如果 $Ae_{n-1}=a_0e_0+a_1e_1+\cdots+a_{n-1}e_{n-1}$，求 A 的极小多项式 $\mu_A(x)$，并求 A 在上述基下的坐标矩阵.

2. 设 $(V, \mathcal{A})$ 是不可分解空间，$\mu_{\mathcal{A}}(x)=p(x)^m$，$p(x)\in K[x]$ 不可约。试证明：

(1) $\mathrm{Ker}(p(\mathcal{A}))$, $\mathrm{Ker}(p(\mathcal{A})^2)$, $\cdots$, $\mathrm{Ker}(p(\mathcal{A})^{m-1})$ 是 $\mathcal{A}$ 的全部非平凡 $\mathcal{A}$ 不变子空间；

(2) $\mathrm{Im}(p(\mathcal{A})^k)=\mathrm{Ker}(p(\mathcal{A})^{m-k})$，$(0\le k\le m)$.

3. 设 $(V, \mathcal{A})$ 是不可分解空间，$V=K[\mathcal{A}]\cdot e$，

$$\mu_{\mathcal{A}}(x)=(x^2-ax-b)^m,\quad p(x)=x^2-ax-b \text{ 不可约.}$$

令 $\alpha_i=(\mathcal{A}^2-a\mathcal{A}-b)^{i-1}e$，$\beta_i=(\mathcal{A}^2-a\mathcal{A}-b)^{i-1}\mathcal{A}e$，$(1\le i\le m)$. 试证：

(1) $\alpha_1, \beta_1, \alpha_2, \beta_2, \cdots, \alpha_m, \beta_m$ 是 V 的一组基.

(2) $\mathcal{A}$ 在基 $\alpha_1, \beta_1, \alpha_2, \beta_2, \cdots, \alpha_m, \beta_m$ 下的坐标矩阵是

$$A=\begin{pmatrix} 0 & b & & & & & & & \\ 1 & a & & & & & & \text{\Large 0} & \\ & 1 & 0 & b & & & & & \\ & & 1 & a & & & & & \\ & & & 1 & 0 & b & & & \\ & & & & 1 & a & & & \\ & & & & & & \ddots & & \\ & \text{\Large 0} & & & & & 1 & 0 & b \\ & & & & & & & 1 & a \end{pmatrix}.$$

4. 设 $\mathcal{A}\in\mathcal{L}(V)$，$\mathcal{A}^*\in\mathcal{L}(V^*)$ 是 $\mathcal{A}$ 的对偶算子，$W\subset V$ 是 r 维子空间，$n=\dim_K(V)$. 试证明：

(1) $W^{\perp}=\{l\in V^* \mid l|_W=0\}\subset V^*$ 是一个 $n-r$ 维子空间，

(2) 若 W 是 $\mathcal{A}$ 的不变子空间，则 $W^{\perp}$ 是 $\mathcal{A}^*$ 的不变子空间.

(3) 线性映射 $V^*\xrightarrow{\varphi}W^*$，$l\mapsto l|_W$，是满映射.

§6.3 多项式矩阵方法.

对于给定的线性算子 $\mathcal{A}\in\mathcal{L}(V)$，我们可以定义 $K[x]$ 对 V 的"数乘法"： $K[x]\times V\longrightarrow V$， $(f(x),v)\longmapsto f(x)\cdot v=f(\mathcal{A})(v)$. 当 $f(x)$ 是 K 中常数时，$f(x)\cdot v$ 就是 K-向量空间 V 上的"数乘法". 所以运算 $K[x]\times V\longrightarrow V$, $(f(x),v)\longmapsto f(x)\cdot v$, 可以看成 K-向量空间 V 的"数乘法" $K\times V\longrightarrow V$, $(\lambda,v)\longmapsto\lambda v$. 的扩展。

<u>引理 6.3.1</u>：给定 $\mathcal{A}\in\mathcal{L}(V)$，则下述"多项式数乘运算"

$$K[x]\times V\longrightarrow V\ ,\ (f(x),v)\longmapsto f(x)\cdot v:=f(\mathcal{A})\cdot v:=f(\mathcal{A})(v)$$

是 V 中"数乘运算"的扩展，即：若 $f(x)$ 是零多项式 $a\in K$，则

$$f(x)\cdot v=av.$$

且满足：(1) $f(x)\cdot(g(x)\cdot v)=(f(x)g(x))\cdot v$ ， $\forall f(x),g(x)\in K[x]$, $v\in V$,

(2) $(f(x)+g(x))\cdot v=f(x)\cdot v+g(x)\cdot v$ ， (3) $f(x)\cdot(v_1+v_2)=f(x)\cdot v_1+f(x)\cdot v_2$.

证明：由定义直接验证. □

对于任意多项式矩阵 $A(x)=(a_{ij}(x))_{m\times n}$ 及 $\alpha_1,\cdots,\alpha_m\in V$, $\beta_1,\cdots,\beta_n\in V$，我们可以定义运算

$$(\alpha_1,\cdots,\alpha_m)\cdot A(x)=\left(\sum_{i=1}^{m}a_{i1}(x)\cdot\alpha_i\ ,\sum_{i=1}^{m}a_{i2}(x)\cdot\alpha_i,\cdots,\sum_{i=1}^{m}a_{in}(x)\cdot\alpha_i\right)$$

$$A(x)\begin{pmatrix}\beta_1\\ \vdots\\ \beta_n\end{pmatrix}=\begin{pmatrix}\sum_{j=1}^{n}a_{1j}(x)\cdot\beta_j\\ \vdots\\ \sum_{j=1}^{n}a_{mj}(x)\cdot\beta_j\end{pmatrix},$$ 它们满足通常矩阵运算的

规则。虽然对初学者而言，对这些运算可能不习惯（甚至有些不习惯）但我们将会看到，通过这种抽象形式的运算，不仅无需任何技巧地重新证明了上一节的结果，而且对任意矩阵 $A\in M_n(K)$，提供了寻找 T 使 $T^{-1}AT=J(A)$ 的有效计算方法。

设 $e_1,\cdots,e_n\in V$ 是一组基，$\mathcal{A}\in\mathcal{L}(V)$ 由下述等式

(6.3.1) $\qquad (\mathcal{A}e_1,\cdots,\mathcal{A}e_n)=(e_1,\cdots,e_n)A\ ,\quad A=(a_{ij})_{n\times n}\in M_n(K),$

唯一确定。根据约定：$x\cdot e_i=\mathcal{A}e_i$，所以 (6.3.1) 可翻译为

(6.3.2) $\qquad (e_1,\cdots,e_n)\cdot(xI_n-A)=0 \qquad$ (矩阵等式).

这一翻译的好处立竿见影，它使我们很自然地得到

<u>定理 6.3.1</u> (凯莱-哈密尔顿，Cayley-Hamilton 定理). 设

$$\chi_A(x)=|x\cdot I_n-A|=x^n-a_1x^{n-1}+a_2x^{n-2}+\cdots+(-1)^{n-1}a_{n-1}x+(-1)^na_n\ (\in K[x])$$

是 A 的特征多项式. 则

$$\chi_A(\mathcal{A})=\mathcal{A}^n-a_1\mathcal{A}^{n-1}+a_2\mathcal{A}^{n-2}+\cdots+(-1)^{n-1}a_{n-1}\mathcal{A}+(-1)^na_n=0.$$

证明：在等式 (6.3.2) 两边同时乘 xI_n-A 的伴随矩阵 $(xI_n-A)^*$ 得

$$(e_1,\cdots,e_n)(x\cdot I_n-A)\cdot(x\cdot I_n-A)^*=(e_1,\cdots,e_n)\begin{pmatrix}\chi_A(x) & & 0\\ & \ddots & \\ 0 & & \chi_A(x)\end{pmatrix}=0.$$

所以，对所有 $1\le i\le n$，$\chi_A(x)\cdot e_i=0$. 将每个 $\chi_A(x)\cdot e_i$ 翻译过来就是 $\chi_A(\mathcal{A})e_i=\chi_A(\mathcal{A})(e_i)=0\ (1\le i\le n)$. 所以 $\chi_A(\mathcal{A})$ 是零算子 □.

等式 (6.3.2) 包含线性算子 $\mathcal{A}\in\mathcal{L}(V)$ 的全部信息，所以是研究 $\mathcal{A}$ 的源泉。为进一步通过 (6.3.2) 研究 $\mathcal{A}$，先介绍一个技术引理.

<u>引理 6.3.2</u> (多项式矩阵的带余除法). 设 $A(x)\in M_{n\times m}(K[x])$，则存在 $B(x)\in M_{n\times m}(K[x])$，$R\in M_{n\times m}(K)$ 使

$$A(x)=(xI_n-A)B(x)+R.$$

证明：任意 $A(x)\in M_{n\times m}(K[x])$ 可写成

$$A(x)=A_kx^k+A_{k-1}x^{k-1}+\cdots+A_1x+A_0,\qquad A_i\in M_{n\times m}(K).$$

对 $A(x)$ 的次数 k 做归纳法。消去首项 A_kx^k：令

$$\bar{A}(x) = A(x) - (xI_n - A)A_k x^{k-1} = A(x) - A_k x^k + AA_k x^{k-1}$$
$$= (A_{k-1} + AA_k)x^{k-1} + A_{k-2}x^{k-2} + \cdots + A_1 x + A_0$$

由归纳假设，存在 $\bar{B}(x) \in M_{n\times m}(K[x])$，$\bar{R} \in M_{n\times m}(K)$ 使

$$\bar{A}(x) = (xI_n - A)\bar{B}(x) + \bar{R}.$$

所以，$A(x) = \bar{A}(x) + (xI_n - A)A_k x^{k-1} = (xI_n - A)(A_k x^{k-1} + \bar{B}(x)) + \bar{R}$.
令 $B(x) = A_k x^{k-1} + \bar{B}(x)$，$R = \bar{R}$ 即得引理。 □.

定理 6.3.2. 存在可逆矩阵 $P(x), Q(x) \in M_n(K[x])$ 使

$$(6.3.3)\quad P(x)\cdot(xI_n - A)\cdot Q(x) = \begin{pmatrix} 1 & & & & & \\ & \ddots & & & & \\ & & 1 & & & \\ & & & d_1(x) & & \\ & & & & \ddots & \\ & & & & & d_s(x) \end{pmatrix}_{n\times n}$$

其中 $d_i(x) | d_{i+1}(x)$ 是首项系数为1的非常数多项式。令

$$(6.3.4)\quad (e_1, \cdots, e_n)P(x)^{-1} := (u_1, \cdots, u_{n-s}, v_1, \cdots, v_s),\quad u_i, v_i \in V.$$

则 $V = K[A]\cdot v_1 \oplus \cdots \oplus K[A]\cdot v_s$，且 $\mu_{A,v_i}(x) = d_i(x)$.

证明：由等式(6.3.2)，得 $(e_1, \cdots, e_n)\cdot P(x)^{-1}\cdot P(x)\cdot(xI_n - A)Q(x) = 0$，故

$$(u_1, \cdots, u_{n-s}, v_1, \cdots, v_s)\cdot \begin{pmatrix} 1 & & & & & \\ & \ddots & & & & \\ & & 1 & & & \\ & & & d_1(x) & & \\ & & & & \ddots & \\ & & & & & d_s(x) \end{pmatrix}_{n\times n} = 0.$$

因此，$u_1, \cdots, u_{n-s} = 0$，$d_1(x)\cdot v_1 = 0, \cdots, d_s(x)\cdot v_s = 0$. 由于

$$(e_1, \cdots, e_n) = (u_1, \cdots, u_{n-s}, v_1, \cdots, v_s)P(x) = (0, \cdots, 0, v_1, \cdots, v_s)P(x)$$

可得 $V = K[x]\cdot e_1 + \cdots + K[x]\cdot e_n = K[x]\cdot v_1 + \cdots + K[x]\cdot v_s = K[A]\cdot v_1 + \cdots + K[A]\cdot v_s$.
所以只需证明：(1) $V = K[A]\cdot v_1 + \cdots + K[A]\cdot v_s$ 是直和；(2) $\mu_{A,v_i}(x) = d_i(x)$.
下述引理是为了证明(1)和(2). □

引理 6.3.3. 若 $f_1(x)\cdot v_1+\cdots+f_s(x)\cdot v_s=0$，则 $d_i(x)\mid f_i(x)$. $(1\le i\le s)$.

证明：$f_1(x)\cdot v_1+\cdots+f_s(x)\cdot v_s=0$ 可以写成

$$0=(u_1,\cdots,u_{n-s},v_1,\cdots,v_s)\begin{pmatrix}0\\ \vdots\\ 0\\ f_1(x)\\ \vdots\\ f_s(x)\end{pmatrix}=(e_1,\cdots,e_n)P(x)^{-1}\begin{pmatrix}0\\ \vdots\\ 0\\ f_1(x)\\ \vdots\\ f_s(x)\end{pmatrix}.$$

由引理 6.3.2（带余除法），存在多项式矩阵 $h(x)\in M_{n\times1}(K[x])$ 使

$$P(x)^{-1}\begin{pmatrix}0\\ \vdots\\ 0\\ f_1(x)\\ \vdots\\ f_s(x)\end{pmatrix}=(xI_n-A)h(x)+\begin{pmatrix}b_1\\ \vdots\\ b_n\end{pmatrix},\quad h(x):=\begin{pmatrix}h_1(x)\\ \vdots\\ h_n(x)\end{pmatrix},\ b_i\in K.$$

由

$$0=(e_1,\cdots,e_n)P(x)^{-1}\begin{pmatrix}0\\ \vdots\\ 0\\ f_1(x)\\ \vdots\\ f_s(x)\end{pmatrix}=(e_1,\cdots,e_n)\cdot(xI_n-A)\begin{pmatrix}h_1(x)\\ \vdots\\ h_n(x)\end{pmatrix}+(e_1,\cdots,e_n)\begin{pmatrix}b_1\\ \vdots\\ b_n\end{pmatrix}.$$

得 $b_1e_1+\cdots+b_ne_n=0$. $(b_i\in K)$. 故 $b_1=\cdots=b_n=0$（因 $e_1,\cdots,e_n$ 是基）. 由因

$$\begin{pmatrix}0\\ \vdots\\ 0\\ f_1(x)\\ \vdots\\ f_s(x)\end{pmatrix}=P(x)\cdot(xI_n-A)\cdot Q(x)\cdot Q(x)^{-1}\cdot h(x)=P(x)\cdot(xI_n-A)Q(x)\cdot\begin{pmatrix}\tilde h_1(x)\\ \vdots\\ \tilde h_{n-s}(x)\\ \vdots\\ \tilde h_n(x)\end{pmatrix}.$$

所以 $f_i(x)=d_i(x)\tilde h_{n-s+i}(x)$，$1\le i\le s$. $\square$

下面引入一些关于多项式矩阵的术语用来证明 $d_1(x),\cdots,d_s(x)$ 与 $P(x),Q(x)$ 的选取无关（即与初等行变换和列变换无关）.

设 $A(x)\in M_{m\times n}(K[x])$，$1\le k\le \min\{m,n\}$，$D_k(A(x))$ 表示 $A(x)$ 中所有 k-阶非零子式的最大公因。如果 $A(x)$ 的所有 k 阶子式均为零，则令 $D_k(A(x))=0$.

引理 6.3.4. $D_i(A(x))\mid D_{i+1}(A(x))$. 如果 $A(x)$ 与 $B(x)$ 等价，则 $D_i(A(x))=a_iD_i(B(x))$，$a_i\in K^*=K\setminus\{0\}$.

证明：因为任意 $i+1$ 阶子式都是 i 阶子式的线性组合，所以 $D_i(A(x))\mid D_{i+1}(A(x))$.

$B(x)$ 的子式行列式由 $A(x)$ 的子式行列式通过初等变换得到，故 $B(x)$ 的任意 k 阶子式与 $A(x)$ 中对应的 k 阶子式相差一个非零常数。因此
$D_i(A(x)) = a_i D_i(B(x))$. □

定义 6.3.1. $D_k(A(x))$ 称为 $A(x)$ 的 k 阶行列式因子，而

$$\{ D_1(A(x)), \ \cdots, \ D_{\min\{m,n\}}(A(x)) \}$$

称为 $A(x)$ 的行列式因子集合。令 r 是使 $D_r(A(x)) \neq 0$ 的最大整数，则
$d_1(A(x)) = D_1(A(x))$, $d_2(A(x)) = \dfrac{D_2(A(x))}{D_1(A(x))}$, $\cdots$, $d_r(A(x)) = \dfrac{D_r(A(x))}{D_{r-1}(A(x))}$
称为 $A(x)$ 的不变因子（要求 $D_i(A(x))$, $d_i(A(x))$ 是首项系数为1的多项式，从而它们在"等价"关系下不变）。□

引理 6.3.5. 如果 $xI_n - A$ 初等等价于

$$B = \begin{pmatrix} 1 & & & & & 0 \\ & \ddots & & & & \\ & & 1 & & & \\ & & & d_1(x) & & \\ & & & & \ddots & \\ 0 & & & & & d_s(x) \end{pmatrix},$$

其中 $d_i(x) | d_{i+1}(x)$，$d_i(x)$首1.

则 $xI_n - A$ 的不变因子是：$\underbrace{1, 1, \cdots, 1}_{n-s}$, $d_1(x), d_2(x), \cdots, d_s(x)$. □

证明：直接计算 B 即可.

定义 6.3.2. $xI_n - A$ 的不变因子 $d_1(x), \cdots, d_s(x)$ 称为 A 的不变因子。

如果 $d_i(x) = P_{i1}(x)^{m_{i1}} \cdots P_{in_i}(x)^{m_{in_i}}$ $(1 \le i \le s)$, $m_{ij} > 0$,

是不可约分解，则 $P_{ij}(x)^{m_{ij}}$ 称为 A 的一个初等因子。称

A 的初等因子全体称为 A 的初等因子组。□

推论 6.3.1: 如果 $d_1(x), \cdots, d_s(x)$, $d_i(x) | d_{i+1}(x)$, 是 A 的不变因子，

则 $\mu_A(x) = d_s(x)$.

证明：由定理 6.3.2, $V = k[A]\cdot v_1 \oplus \cdots \oplus k[A]\cdot v_s$, $\mu_{A,v_i}(x) = d_i(x)$, 所以 $d_s(A)\cdot V = 0$, 从而 $\mu_A(x) | d_s(x)$. 由于 $\mu_{A,v_s}(x) = d_s(x)$,
可得 $d_s(x) | \mu_A(x)$. 因此 $\mu_A(x) = d_s(x) = \mu_{A,v_s}(x)$. □

推论6.3.2. 设 $A\in M_n(K)$，如果

$$d_i(x)=x^{m_i}-a^{(i)}_{m_i-1}x^{m_i-1}-\cdots-a_1^{(i)}x-a_0^{(i)}, \quad (i=1,2,\cdots,s).$$

是 A 的不变因子. 则 A 相似于

$$F(A)=\begin{pmatrix} A_1 & & & 0\\ & A_2 & & \\ & & \ddots & \\ 0 & & & A_s\end{pmatrix},\quad 其中\ A_i=\begin{pmatrix} 0 & 0 & \cdots & 0 & a_0^{(i)}\\ 1 & 0 & \cdots & 0 & a_1^{(i)}\\ 0 & 1 & \cdots & 0 & a_2^{(i)}\\ \vdots & \vdots & \ddots & 1 & \vdots\\ 0 & 0 & \cdots & 0 & a_{m_i-2}^{(i)}\\ 0 & 0 & \cdots & 1 & a_{m_i-1}^{(i)}\end{pmatrix}$$

证明：由定理6.3.2，$V=K[A]\cdot v_1\oplus\cdots\oplus K[A]\cdot v_s$，$\mu_{A,v_i}(x)=d_i(x)$.

在 $K[A]\cdot v_i$ 中取基 $v_i, Av_i, A^2v_i,\cdots,A^{m_i-1}v_i$，$A_i=A|_{K[A]\cdot v_i}$ 在该基下的坐标矩阵为

$$A_i=\begin{pmatrix} 0 & 0 & 0 & \cdots & 0 & a_0^{(i)}\\ 1 & 0 & 0 & \cdots & 0 & a_1^{(i)}\\ 0 & 1 & 0 & \cdots & 0 & a_2^{(i)}\\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots\\ 0 & 0 & 0 & \cdots & 0 & a_{m_i-2}^{(i)}\\ 0 & 0 & 0 & \cdots & 1 & a_{m_i-1}^{(i)}\end{pmatrix}.$$

每个 $K[A]\cdot v_i$ 中上述基一起可得 V 的一组基使 A 在该基下的坐标矩阵

$$F(A)=\begin{pmatrix} A_1 & & 0\\ & A_2 & \\ 0 & \ddots & A_s\end{pmatrix}.$$

所以 A 相似于 F(A). □

推论6.3.3. 设 $(Ae_1,\cdots,Ae_n)=(e_1,\cdots,e_n)A$，$d_1(x),\cdots,d_s(x)$ 是 A 的不变因子. 如果

$$d_i(x)=p_{i1}(x)^{m_{i1}}\cdots p_{in_i}(x)^{m_{in_i}},\quad m_{ij}>0,\quad (1\le i\le s).$$

是 $d_i(x)$ 的不可约分解，则存在 $\alpha_{ij}\in V$ $(1\le i\le s, 1\le j\le n_i)$ 使

$$V=\bigoplus_{i,j}K[A]\cdot\alpha_{ij}$$

其中 $\mu_{A,\alpha_{ij}}(x)=p_{ij}(x)^{m_{ij}}$.

证明：由定理6.3.2，$V=K[A]\cdot v_1\oplus\cdots\oplus K[A]\cdot v_s$ 且 $\mu_{A,v_i}(x)=d_i(x)$

应用引理6.3.4于每个 $K[A]\cdot v_i$ 即得推论。 □

推论6.3.4. 设 A, $B\in M_n(K)$. 则下述结论等价

(1) A 与 B 相似;

(2) xI_n-A 与 xI_n-B 初等等价.

(3) A 与 B 有相同的初等因子组

(4) A 与 B 有相同的不变因子组.

证明: (1)$\Rightarrow$(2): 存在 $T\in M_n(K)$ 使 $B=T^{-1}AT$, 从而

$$xI_n-B=T^{-1}(xI_n-A)T.$$

所以 xI_n-A 与 xI_n-B 等价.

(2)$\Rightarrow$(3): 显然.

(3)$\Rightarrow$(4): 由 A 的初等因子组可唯一确定 A 的不变因子.

(4)$\Rightarrow$(1): 由推论6.3.2, A 相似于 $F(A)$, $F(A)$ 由 A 的不变因子组确定. 所以,若 A,B 有相同不变因子组, 则 $F(A)$ 与 $F(B)$ 相似. 从而 A 与 B 相似. □

例6.3.1. 矩阵 (Jordan 块)

$$J_m(\lambda_0)=\begin{pmatrix}\lambda_0 & 1 & & 0\\ & \lambda_0 & 1 & \\ & & \ddots & 1\\ 0 & & & \lambda_0\end{pmatrix}_{m\times m}$$

的初等因子组是 $(x-\lambda_0)^m$, 而 Jordan 矩阵

$$J=\begin{pmatrix}J_{m_1}(\lambda_1) & & & 0\\ & J_{m_2}(\lambda_2) & & \\ & & \ddots & \\ 0 & & & J_{m_s}(\lambda_s)\end{pmatrix}$$

的初等因子组是: $(x-\lambda_1)^{m_1}, (x-\lambda_2)^{m_2}, \dots, (x-\lambda_s)^{m_s}$. □

例6.3.2. 设 $A\in M_7(K)$, A 的初等因子组是

$$x-1,\quad (x-1)^3,\quad (x+1)^2,\quad x-2.$$

则 A 有4个 Jordan 块: $J_1(1)=1$, $J_3(1)=\begin{pmatrix}1&1&0\\0&1&1\\0&0&1\end{pmatrix}$, $J_2(-1)=\begin{pmatrix}-1&1\\0&-1\end{pmatrix}$, $J_1(2)=2$. 所以 A 的 Jordan 标准型为

$$J(A)=\begin{pmatrix} 1 & & & & & & \\ & 1 & 1 & 0 & & & \\ & 0 & 1 & 1 & & & \\ & 0 & 0 & 1 & & & \\ & & & & -1 & 1 & \\ & & & & 0 & -1 & \\ & & & & & & 2 \end{pmatrix}_{7\times 7}$$

此时A的不变因子组可按下列方式求出：将初等因子组分类按降幂排列：

$$\begin{matrix} (x-1)^3, & (x-1) \\ (x+1)^2, & 1 \\ (x-2), & 1 \end{matrix} \quad \Rightarrow \quad \begin{aligned} d_2(x) &= (x-1)^3(x+1)^2(x-2) \\ d_1(x) &= x-1 \end{aligned}$$

是A的不变因子组. □

<u>例6.3.3</u>. 设 $A\in M_{10}(K)$ 的初等因子组是：$x-1,\ x-1,\ (x-1)^2,\ (x+1)^2,\ (x+1)^3,\ x-2$，求A的<u>不变因子组</u>：将它们分类按降幂排列

$$\begin{matrix} (x-1)^2, & x-1, & x-1 \\ (x+1)^3, & (x+1)^2, & 1 \\ x-2, & 1, & 1 \end{matrix} \quad \Rightarrow \text{A的}\underline{\text{不变因子组}}\text{是：} \quad \begin{aligned} d_3(x) &= (x-1)^2(x+1)^3(x-2) \\ d_2(x) &= (x-1)(x+1)^2 \\ d_1(x) &= x-1. \end{aligned}$$

<u>例6.3.4</u>. 设 $A\in M_9(K)$ 的不变因子组为：

$$\overbrace{1, 1, \cdots, 1}^{7},\ (x-1)(x^2+1),\ (x-1)^2(x^2+1)(x^2-2).$$

分别在 $K=\mathbb{Q},\ \mathbb{R},\ \mathbb{C}$ 时，求A的初等因子组。

当 $K=\mathbb{Q}$ 时，A的初等因子组是 $x-1,\ x^2+1,\ (x-1)^2,\ x^2+1,\ x^2-2$。当 $K=\mathbb{R}$ 时，A的初等因子组是：$x-1,\ x^2+1,\ (x-1)^2,\ x^2+1,\ x-\sqrt{2},\ x+\sqrt{2}$。而当 $K=\mathbb{C}$ 时，A的初等因子组是 $x-1,\ x-i,\ x+i,\ (x-1)^2,\ x-i,\ x+i,\ x-\sqrt{2},\ x+\sqrt{2}$。由此我们可分别写出标准型，当 $K=\mathbb{Q},\mathbb{R},\mathbb{C}$ 时，A分别相似于：

$$\begin{pmatrix} A_1 & & & & 0 \\ & A_2 & & & \\ & & A_3 & & \\ & & & A_4 & \\ 0 & & & & A_5 \end{pmatrix},\quad \begin{pmatrix} B_1 & & & & & 0 \\ & B_2 & & & & \\ & & B_3 & & & \\ & & & B_4 & & \\ & & & & B_5 & \\ 0 & & & & & B_6 \end{pmatrix},\quad \begin{pmatrix} 1 & & & & & & & & \\ & i & & & & & & & 0 \\ & & -i & & & & & & \\ & & & i & & & & & \\ & & & & -i & & & & \\ & & & & & \sqrt{2} & & & \\ & & & & & & -\sqrt{2} & & \\ & 0 & & & & & & 1 & 0 \\ & & & & & & & 1 & 1 \end{pmatrix}$$

其中 $A_1=(1),\ A_2=\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix},\ A_3=\begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix},\ A_4=\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix},\ A_5=\begin{pmatrix} 0 & 2 \\ 1 & 0 \end{pmatrix},\ B_1=(1),\ B_2=(\sqrt{2}),$

$B_3=(-\sqrt{2})$，$B_4=\begin{pmatrix}0&-1\\1&0\end{pmatrix}$，$B_5=\begin{pmatrix}1&0\\1&1\end{pmatrix}$，$B_6=\begin{pmatrix}0&-1\\1&0\end{pmatrix}$.

□

例6.3.5. 设

$$A=\begin{pmatrix}3&2&-3\\4&10&-12\\3&6&-7\end{pmatrix}.$$

试求：(1) $\mu_A(x)$；(2) A的Jordan标准型 $J(A)$；(3) 求T使 $T^{-1}AT=J(A)$.

解：设V是3维复空间，$e_1,e_2,e_3\in V$是一组基，$\mathcal{A}\in L(V)$ 由

$$(\mathcal{A}e_1,\mathcal{A}e_2,\mathcal{A}e_3)=(e_1,e_2,e_3)A.$$

定义. 则 $J(A)$ 和使 $T^{-1}AT=J(A)$ 的 $T\in M_3(\mathbb{C})$ 可按下列方式求得

第一步：求可逆矩阵 $P(x)\in M_3(\mathbb{C}[x])$ 使

$$P(x)(xI_3-A)Q(x)=\begin{pmatrix}1&&&0\\&\ddots1&&\\&&d_1(x)&\\0&&&\ddots d_s(x)\end{pmatrix}.$$

$$(xI_3-A\mid I_3)=\left(\begin{array}{ccc|ccc}x-3&-2&3&1&0&0\\-4&x-10&12&0&1&0\\-3&-6&x+7&0&0&1\end{array}\right)\longrightarrow\left(\begin{array}{ccc|ccc}-3,&-6,&x+7&0&0&1\\x-3&-2&3&1&0&0\\-4&x-10,&12&0&1&0\end{array}\right)$$

$$\longrightarrow\left(\begin{array}{ccc|ccc}1&2&-\frac{1}{3}x-\frac{7}{3}&0&0&-\frac{1}{3}\\x-3&-2&3&1&0&0\\-4&x-10&12&0&1&0\end{array}\right)\longrightarrow\left(\begin{array}{ccc|ccc}1&2&-\frac{1}{3}x-\frac{7}{3}&0&0&-\frac{1}{3}\\0&-2(x-2)&\frac{1}{3}x^2+\frac{8}{3}x-\cdots&1&0&\frac{1}{3}(x-3)\\0&x-2&-\frac{4}{3}x+\frac{8}{3}&0&1&-\frac{4}{3}\end{array}\right)$$

$$\longrightarrow\left(\begin{array}{ccc|ccc}1&2&-\frac{1}{3}(x+7)&0&0&-\frac{1}{3}\\0&-2(x-2)&\frac{1}{3}(x-2)(x+6)&1&0&\frac{1}{3}(x-3)\\0&0&\frac{1}{6}(x-2)^2&\frac{1}{2}&1&\frac{1}{6}(x-11)\end{array}\right)\xrightarrow[\text{变换}]{\text{列}}\left(\begin{array}{ccc|ccc}1&0&0&0&0&-\frac{1}{3}\\0&-2(x-2),&\frac{1}{3}(x-2)(x+6)&1&0&\frac{1}{3}(x-3)\\0&0&\frac{1}{6}(x-2)^2&\frac{1}{2}&1&\frac{1}{6}(x-11)\end{array}\right)$$

$$\longrightarrow\left(\begin{array}{ccc|ccc}1&0&0&0&0&-\frac{1}{3}\\0&(x-2)&0&\frac{-1}{2}&0&\frac{-1}{6}(x-3)\\0&0&(x-2)^2&\frac{6}{2}&6&(x-11)\end{array}\right),\quad P(x)=\begin{pmatrix}0&0&-\frac{1}{3}\\-\frac{1}{2}&0&-\frac{1}{6}(x-3)\\3&6&x-11\end{pmatrix},\quad \begin{array}{l}d_1(x)=x-2\\d_2(x)=(x-2)^2\end{array}$$

因此有：(1) $\mu_A(x)=(x-2)^2$；(2) $J(A)=\begin{pmatrix}2&0&0\\0&2&1\\0&0&2\end{pmatrix}$.

第二步：求 $P(x)^{-1}$ 得. $P(x)^{-1}=\begin{pmatrix}x-3,&-2,&0\\-4,&1,&\frac{1}{6}\\-3,&0,&0\end{pmatrix}$.

所以 $(e_1,e_2,e_3)P(x)^{-1}=\left((x-3)e_1-4e_2-3e_3=0,\ -2e_1+e_2,\ \frac{1}{6}e_2\right)$.

令 $v_1=-2e_1+e_2$, $v_2=\frac{1}{6}e_2$, 则 $V=\mathbb{C}[A]\cdot v_1\oplus\mathbb{C}[A]\cdot v_2$,

$\mu_{A,v_1}(x)=x-2$, $\mu_{A,v_2}(x)=(x-2)^2$. 在 $\mathbb{C}[A]\cdot v_1$ 取基: v_1, 在 $\mathbb{C}[A]\cdot v_2$ 中取基: $(A-2)v_2$, v_2. 令 $\alpha_1=v_1$, $\alpha_2=(A-2)\cdot v_2$, $\alpha_3=v_2$, 则

$$(A\alpha_1,A\alpha_2,A\alpha_3)=(\alpha_1,\alpha_2,\alpha_3)\begin{pmatrix}2&0&0\\0&2&1\\0&0&2\end{pmatrix}=(\alpha_1,\alpha_2,\alpha_3)J(A).$$

由 $\alpha_1=-2e_1+e_2$, $\alpha_2=\frac{1}{6}(A-2)e_2=\frac{1}{3}e_1+\frac{4}{3}e_2+e_3$, $\alpha_3=\frac{1}{6}e_2$, 得

$$(\alpha_1,\alpha_2,\alpha_3)=(e_1,e_2,e_3)\begin{pmatrix}-2&\frac{1}{3}&0\\1&\frac{4}{3}&\frac{1}{6}\\0&1&0\end{pmatrix}.$$

所以 $T=\begin{pmatrix}-2&\frac{1}{3}&0\\1&\frac{4}{3}&\frac{1}{6}\\0&1&0\end{pmatrix}$ 使得 $T^{-1}AT=J(A)$. □.

习题 6.3

1. 设

$$A=\begin{pmatrix}0&1&-1&1\\-1&2&-1&1\\-1&1&1&0\\-1&1&0&1\end{pmatrix}.$$

试求: (1) 极小多项式 $\mu_A(x)$; (2) A 的 Jordan 标准型 $J(A)$.
(3) $T\in M_4(\mathbb{C})$ 使 $T^{-1}AT=J(A)$.

2. 已知某矩阵 A 的初等因子组为:
$x-1$, $(x-1)^3$, $x+1$, $x+1$, $(x+1)^3$, $x-2$, $(x-2)^2$.
试求 A 的不变因子组和 $J(A)$.

3. 设 V 是 n 维 k-向量空间，$\mathcal{A}\in\mathcal{L}(V)$

如果 $$\chi_{\mathcal{A}}(x)=p_1(x)^{n_1}\cdot p_2(x)^{n_2}\cdots p_s(x)^{n_s}\qquad (n_i>0)$$
$$\mu_{\mathcal{A}}(x)=p_1(x)^{m_1}\cdot p_2(x)^{m_2}\cdots p_s(x)^{m_s}\qquad (m_i>0)$$

分别是 $\mathcal{A}$ 的特征多项式和极小多项式的不可约分解。令

$$V_i=\ker\left(p_i(\mathcal{A})^{n_i}\right),\quad W_i=\ker\left(p_i(\mathcal{A})^{m_i}\right).$$

证明：(1) $V=V_1\oplus\cdots\oplus V_s=W_1\oplus\cdots\oplus W_s$

(2) $\dim_k W_i = n_i\cdot\deg p_i(x)$. $(1\le i\le s)$.

4. 设 $\mathcal{A}\in\mathcal{L}(V)$ 满足 $\mu_{\mathcal{A}}(x)=\chi_{\mathcal{A}}(x)$. 试证明.

$$C(\mathcal{A}):=\{B\in\mathcal{L}(V)\mid B\cdot\mathcal{A}=\mathcal{A}\cdot B\}=k[\mathcal{A}].$$

5. 设
$$J_n(\lambda)=\begin{pmatrix}\lambda & 1 & & 0\\ & \lambda & 1 & \\ & & \ddots & 1\\ 0 & & & \lambda\end{pmatrix}_{n\times n}$$
是 n 阶 Jordan 块

证明：
$$J_n(\lambda)^m=\begin{pmatrix}\lambda^m, & C_m^1\lambda^{m-1}, & C_m^2\lambda^{m-2}, & \cdots\\ & \lambda^m, & C_m^1\lambda^{m-1}, & \cdots\\ & & \lambda^m, & C_m^1\lambda^{m-1}, \cdots\\ & & & \ddots\\ & & & & \lambda^m\end{pmatrix}.$$

其中 $C_m^k=\frac{m(m-1)\cdots(m-k+1)}{k!}=\frac{1}{k!}(\lambda^m)^{(k)}$ （$(\lambda^m)^{(k)}$ 表示对 λ 求 k 阶导数）

6. 设 $A\in M_n(\mathbb{Q})$，试证明：

A 是幂零矩阵 $\Longleftrightarrow$ $\mathrm{tr}(A^i)=0$ $(1\le i\le n)$.

§6.4. 内积空间.

为了更准确地描述现实世界，也为了数学，物理等领域的研究需要，我们需要在向量空间 V 上引入额外的结构。例如，~~为了~~ 在抽象实向量空间 V 上引入长度，角度等概念，我们需要在 V 上指定一个正定的二次型。如果在实向量空间 V 上指定了一个反对称双线性函数，则称 V 是辛空间。带有一个 Hermitian 形式的有限维复空间称为 Hermitian。这些都是数学中重要的研究对象。

由于时间关系，我们在本课程中仅介绍带有一个正定对称双线性函数的实向量空间（内积空间）.

定义 6.4.1. 设 V 是一个实向量空间（可能无限维），双线性函数 $(\cdot|\cdot): V\times V\to\mathbb{R}$ 称为 V 上的一个内积 (inner product)，如果它满足：

(1) $\forall\ x, y\in V$，有 $(x|y)=(y|x)$ （对称性）

(2) $\forall\ x\in V$，有 $(x|x)\geq 0$，且

$(x|x)=0 \iff x=0$ （正定性）.

带有指定内积的实向量空间 $V=(V,(\cdot|\cdot))$ 称为一个内积空间。有限维内积空间称为欧几里得空间（简称欧氏空间）。 □

例6.4.1 $V=\mathbb{R}^2$（标准向量空间）. $\forall\ A=(x_1,x_2),\ B=(y_1,y_2)\in V$.

$$(A|B)=(x_1,x_2)\begin{pmatrix}y_1\\y_2\end{pmatrix}=x_1y_1+x_2y_2,$$ 是$\mathbb{R}^2$上的一个内积

而 A, B 之间的距离 $d(A,B)=\sqrt{(x_1-y_1)^2+(x_2-y_2)^2}=\sqrt{(A-B|A-B)}$.

如果将$\mathbb{R}^2$中向量 $x=(x_1,x_2)$ 看成是始于原点的(有向)线段，则向量 $x=(x_1,x_2)$ 的长度为 $\|x\|=\sqrt{x_1^2+x_2^2}=\sqrt{(x|x)}$，两个向量 $x=(x_1,x_2)$, $y=(y_1,y_2)$ 的夹角 φ 满足：$\cos\varphi=\dfrac{(x|y)}{\|x\|\cdot\|y\|}$.

□

例6.4.2：$V=\mathbb{R}^n$（标准向量空间）. $\forall\ x=(x_1,x_2,\cdots,x_n),\ y=(y_1,\cdots,y_n)\in V$

$$(x|y)=x_1y_1+\cdots+x_ny_n$$ 是$\mathbb{R}^n$上的一个内积（称为标准内积）.

事实上，$\forall$ ~~[illegible]~~ 一个正定矩阵 $A\in M_n(\mathbb{R})$，双线性函数

$$(x|y)_A:=xA\,{}^ty$$ 是$\mathbb{R}^n$上的一个内积.

□

例6.4.3：$V=C([a,b])=\{[a,b]\text{上的连续实函数}\}$，则

$$(f|g):=\int_a^b f(x)g(x)\,dx,\quad \forall\ f,g\in C([a,b])=V$$

是V上的一个内积.

□

例6.4.4 设 V 是一个内积空间，$W\subset V$ 是一子空间，则 V 上内积 $(\cdot|\cdot):V\times V\to\mathbb{R}$ 限制在 $W\times W$ 也是 W 上的一个内积，所以 W（在V的内积下）也是一个内积空间。

□

定理6.4.1 设 V 是一个内积空间，则

$$(\alpha|\beta)^2\le(\alpha|\alpha)\cdot(\beta|\beta),\quad \forall\,\alpha,\beta\in V.$$

且等号成立当且仅当 α,β 线性相关。

证明：令 $W=\langle\alpha,\beta\rangle\subset V$. $\forall\ x=x_1\alpha+x_2\beta\in W$；有

$\langle x|x\rangle\ge0$，等号成立 $\Leftrightarrow x=0$

即 $(x_1, x_2)\begin{pmatrix}(\alpha|\alpha), (\alpha|\beta)\\ (\alpha|\beta), (\beta|\beta)\end{pmatrix}\begin{pmatrix}x_1\\ x_2\end{pmatrix} = \langle x|x\rangle \geq 0$ (*)

对任意 x_1, x_2 成立，知矩阵 $A=\begin{pmatrix}(\alpha|\alpha), (\alpha|\beta)\\ (\alpha|\beta), (\beta|\beta)\end{pmatrix}$ 半正定．所以

$$(\alpha|\alpha)\cdot(\beta|\beta)-(\alpha|\beta)^2=|A|\geqslant 0.$$

且 $|A|=0 \Longleftrightarrow$ 存在不全为零的 x_1, x_2 使 $\langle x_1\alpha+x_2\beta \mid x_1\alpha+x_2\beta\rangle=0$
$\Longleftrightarrow x_1\alpha+x_2\beta=0 \Longleftrightarrow \alpha, \beta$ 线性相关．

$\square$

对于不同内积的选取，上述不等式可推论为许多著名不等式．例如：

(1) 在 $\mathbb{R}^n$ 中选取内积：$\forall\ x=(x_1, \cdots, x_n),\ y=(y_1, \cdots, y_n)\in\mathbb{R}^n$

$$\langle x|y\rangle = x_1y_1+\cdots+x_ny_n.$$

则定理6.1中不等式变为 Cauchy 不等式：

$$\left|\sum_{i=1}^n x_iy_i\right| \leq \sqrt{\sum_{i=1}^n x_i^2}\cdot\sqrt{\sum_{i=1}^n y_i^2}.$$ 等号成立 $\Leftrightarrow y_i=\lambda x_i$．

(2) 在 $C([a,b])$ 上选取内积：$(f|g)=\int_a^b f(x)g(x)dx$．则

$$\left|\int_a^b f(x)g(x)dx\right| \leq \sqrt{\int_a^b f(x)^2dx}\cdot\sqrt{\int_a^b g(x)^2dx}$$ （Schwarz 不等式）．

推论6.4.1：$\forall\ \alpha, \beta\in V$，令 $\|\alpha\|=\sqrt{(\alpha|\alpha)}$，$\|\beta\|=\sqrt{(\beta|\beta)}$．则

$$-1 \leq \frac{\langle\alpha, \beta\rangle}{\|\alpha\|\cdot\|\beta\|} \leq 1$$

$\square$

推论6.4.2：$\forall\ \alpha, \beta\in V$，$\|\alpha\pm\beta\| \leq \|\alpha\|+\|\beta\|$．

证明：$\|\alpha\pm\beta\|^2=(\alpha\pm\beta\mid\alpha\pm\beta)=\|\alpha\|^2+\|\beta\|^2\pm 2\langle\alpha|\beta\rangle \leq \|\alpha\|^2+\|\beta\|^2+2|(\alpha|\beta)|$
$\leq \|\alpha\|^2+\|\beta\|^2+2\|\alpha\|\cdot\|\beta\| = (\|\alpha\|+\|\beta\|)^2$．

$\square$．

定义6.4.2: 设 V 是内积空间, $\forall\ x,y\in V$. 则.

(1) $\|x\|=\sqrt{(x|x)}$ 称为 x 的长度, $\|x-y\|$ 称为 x,y 之间的距离, $x\in V$ 称为单位向量, 如果 $\|x\|=1$.

(2) 存在唯一 $0\le\varphi\le\pi$ 使 $\cos\varphi=\dfrac{(x|y)}{\|x\|\cdot\|y\|}$. φ 称为 x 与 y 的夹角.

(3) x 与 y 称为正交 (记为 $x\perp y$) 如果它们的夹角 $\varphi=\frac{\pi}{2}$.

即: $x\perp y \Leftrightarrow (x|y)=0$.

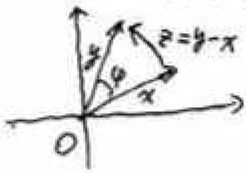
(4). $V=\mathbb{R}^n$ 与它的标准内积: $\forall\ x=(x_1,\cdots,x_n),\ y=(y_1,\cdots,y_n)$

$$(x|y)=x^t y=x_1y_1+x_2y_2+\cdots+x_ny_n$$

一起被称为 n-维标准欧氏空间. □

例6.4.5: 设 $\mathbb{R}^3$ 是 3-维标准欧氏空间, $\forall\ x=(x_1,x_2,x_3),\ y=(y_1,y_2,y_3)$

则 $\|x\|=\sqrt{x_1^2+x_2^2+x_3^2}$, $\|y\|=\sqrt{y_1^2+y_2^2+y_3^2}$.



$\|z\|^2=\|x\|^2+\|y\|^2-2\|x\|\cdot\|y\|\cos\varphi$ ← (余弦定理)

另一方面, $\|z\|^2=\|y-x\|^2=(y-x|y-x)=\|y\|^2+\|x\|^2-2(x|y)$

所以 $(x|y)=\|x\|\cdot\|y\|\cos\varphi$. 即 $\cos\varphi=\dfrac{(x|y)}{\|x\|\cdot\|y\|}$. □

定义6.4.3: 设 V 是 n-维欧氏空间. $e_1,\cdots,e_n\in V$ 称为标准正交基, 如果 $(e_i|e_j)=\begin{cases}1 & i=j\\ 0 & i\ne j\end{cases}$. □

例6.4.6 设 $\mathbb{R}^n$ 是标准欧氏空间. 则 $e_1=(1,0,\cdots,0,0),\cdots,$ $e_n=(0,\cdots,0,1)$ 是 $\mathbb{R}^n$ 中的标准正交基. □

例6.4.7：设 V 是内积空间，$\alpha_1,\cdots,\alpha_m\in V$ 是一组正交向量（i.e. $\alpha_i\perp\alpha_j$，$\forall\, i\neq j$），则 $\alpha_1,\cdots,\alpha_m$ 线性无关。 □

定理6.4.2：任何 n-维欧氏空间必有标准正交基。

证明：$(\cdot|\cdot): V\times V\to\mathbb{R}$ 是对称双线性函数，所以存在一组基 $e_1',\cdots,e_n'\in V$ 使 $e_i'\perp e_j'$ $(i\neq j)$。但 $(e_i',e_i')>0$ $(1\le i\le n)$。令 $e_i=\frac{1}{\sqrt{(e_i'|e_i')}}e_i'$，$(1\le i\le n)$，则 $(e_i|e_i)=1$，$(e_i|e_j)=0$ $(i\neq j)$。所以 $e_1,\cdots,e_n\in V$ 是一组标准正交基。 □

定理6.4.3（Gram-Schmidt 方法）。设 V 是内积空间，$u_1,\cdots,u_m\in V$ 线性无关，令 $v_1=u_1$，$v_2=u_2-\frac{(u_2|v_1)}{\|v_1\|^2}v_1$，$\cdots$

$$v_{k+1}=u_{k+1}-\sum_{i=1}^{k}\frac{(u_{k+1}|v_i)}{\|v_i\|^2}v_i,\quad (1\le k\le m-1).$$

则 $v_1,v_2,\cdots,v_m$ 互相正交 $(v_i\perp v_j)$，$\langle u_1,\cdots,u_m\rangle=\langle v_1,\cdots,v_m\rangle$，且 $(v_1,v_2,\cdots,v_m)=(u_1,u_2,\cdots,u_m)T$，其中 T 是对角线为1的上三角矩阵。

证明：直接验证（留作练习）。 □

定义6.4.4：两个内积空间 V，V' 称为同构，如果存在一个线性同构 $f: V\to V'$ 使得，$\forall\, x,y\in V$，

$$(x|y)_V=(f(x)|f(y))_{V'}$$ □

推论6.4.3：任意 n-维欧氏空间 V 都同构于标准欧氏空间 $(\mathbb{R}^n, (\cdot|\cdot)_{\mathbb{R}^n})$。

证明：存在标准正交基 $e_1,\cdots,e_n$。$\forall\, x=x_1e_1+\cdots+x_ne_n\in V$，定义 $f(x)=(x_1,\cdots,x_n)\in\mathbb{R}^n$，则 $f: V\to\mathbb{R}^n$ 是同构。 □

推论 6.4.4：欧氏空间 V 中的任一组标准正交向量 $\alpha_1,\cdots,\alpha_s$ 可以扩充成 V 的一组标准正交基。特别地，对任意子空间 $L\subset V$，有 $V=L\oplus L^{\perp}$，其中 $L^{\perp}=\{v\in V \mid \forall x\in L,\ (v|x)=0\}$ 且 $(L^{\perp})^{\perp}=L$。

习题 6.4

1. 设 V 是 n-维欧氏空间，$e_1,\cdots,e_n$ 和 $v_1,\cdots,v_n$ 是 V 的两组标准正交基，证明：$(v_1,\cdots,v_n)=(e_1,\cdots,e_n)T$，其中 ${}^tT\cdot T=I_n$（这样的 T 称为正交矩阵）。

2. 设 $P_2=\{f(t)\in\mathbb{R}[t] \mid \deg f(t)\leq 2\}$，内积由 $(f|g)=\int_{-1}^{1}f(t)g(t)dt$ 定义
(1) 求子空间 $\langle 1,t\rangle^{\perp}$；(2) 求 P_2 的一组标准正交基。

3. 在 $\mathbb{R}[t]$ 中定义 $(f|g)=\int_0^{\infty}e^{t}f(t)g(t)dt$。证明：
(1) 它是 $\mathbb{R}[t]$ 的一个内积
(2) $(t^n|t^m)=(m+n)!$

4. 设 $\langle\cdot|\cdot\rangle_1$ 与 $\langle\cdot|\cdot\rangle_2$ 是实向量空间 V（可能无穷维）上的两个内积，证明：$\langle\cdot|\cdot\rangle_1=\langle\cdot|\cdot\rangle_2 \Leftrightarrow \|x\|_1=\|x\|_2,\ \forall x\in V$.

5. 在 $\mathbb{R}^3$ 中定义内积使：$\forall\ x=[x_1,x_2,x_3]\in V$，
$$\|x\|^2=3x_1^2+2x_2^2+x_3^2-4x_1x_2-2x_1x_3+2x_2x_3,$$ 试求：
(1) $x=[1,1,1]$ 与 $y=[2,2,1]$ 之间的夹角；(2) 所有与 x 正交的向量。

6. 证明：所有可逆矩阵 $A=(a_{ij})_{n\times n}\in M_n(\mathbb{R})$ 都可写成 $A=B\cdot C$ 其中 B 是正交矩阵，C 是上三角形矩阵，且 $|C|=\pm|A|$。

7. 设 $A=[A_{(1)},\cdots,A_{(n)}]\in M_n(\mathbb{R})$，$A_{(1)},\cdots,A_{(n)}$ 是标准欧氏空间 $\mathbb{R}^n$ 的一组正交向量。证明：$|A|=\|A_{(1)}\|\cdot\|A_{(2)}\|\cdot\cdots\cdot\|A_{(n)}\|$.

73

8. 证明：$\forall\ X=[X_{(1)},\cdots,X_{(n)}]\in M_n(\mathbb{R})$，有

$$|\det(X)|\leq\|X_{(1)}\|\cdot\|X_{(2)}\|\cdot\cdots\|X_{(n)}\|$$，（其采用 $\mathbb{R}^n$ 中标准内积）.

9. 设 V 是 n-维欧氏空间，$e_1,\cdots,e_n\in V$ 是一组基. 如果 $(\alpha|e_i)=(\beta|e_i)$，$(1\leq i\leq n)$，证明：$\alpha=\beta$.

10. 在内积空间 V 中，证明：

$$\|\alpha+\beta\|^2+\|\alpha-\beta\|^2=2\|\alpha\|^2+2\|\beta\|^2.$$

（即：平行四边形中两对角线平方和等于 4 边平方和）.

~~设 $A=[A_{(1)},\cdots,A_{(n)}]\in M_n(\mathbb{R})$~~,

11. 设 $A=(A^{(1)},\cdots,A^{(n)})\in M_n(\mathbb{R})$，如果 $A^{(1)},\cdots,A^{(n)}\in\mathbb{R}^n$ 是标准欧氏空间的一组正交向量，试求 $A^{-1}=?$

12. 在标准欧氏空间 $\mathbb{R}^4$ 中构造子空间

$$\langle(1,2,2,-1),(1,1,-5,3),(3,2,8,-7)\rangle\subset\mathbb{R}^4$$

的一组标准正交基

13.（Bessel 不等式）. 设 V 是 n-维欧氏空间，$e_1,\cdots,e_k\in V$ 是一组正交向量. 证明，$\forall\ x\in V$，有

$$\sum_{i=1}^{k}|(x|e_i)^2\leq\|x\|^2.$$

且等号对任意 $x\in V$ 成立 $\Longleftrightarrow k=n$.

在研究欧氏空间上线性算子之前，我们先做一点一般性讨论。设 V 是 n 维 K-线性空间，$\mathcal{A}\in\mathcal{L}(V)$ 是一个线性算子，从§6.2中的讨论可以看出 $\mathcal{A}$ 的对偶算子 $\mathcal{A}^*: V^*\longrightarrow V^*$ 对研究 $\mathcal{A}$ 很有帮助。例如，对任意 $\mathcal{A}$ 的不变子空间 $W\subset V$，$W^\perp\subset V^*$ 也是 $\mathcal{A}^*$ 的不变子空间。如果在 V 与 V^* 之间有更直接的联系，则 $\mathcal{A}^*$ 可提供更多研究 $\mathcal{A}$ 的信息。

命题6.5.1：设 $\mathcal{L}(V,V^*)$ 表示所有 K-线性映射 $V\xrightarrow{\varphi}V^*$ 组成的 K-向量空间，$\mathcal{L}_2(V)$ 表示所有双线性映射

$$f: V\times V\longrightarrow K$$

组成的 K-向量空间。则映射

$$\Phi_1:\ \mathcal{L}(V,V^*)\longrightarrow\mathcal{L}_2(V)$$
$$\varphi\longmapsto\Phi_1(\varphi): V\times V\to K$$
$$(x,y)\longmapsto\varphi(x)(y)$$

是一个 K-线性同构，且满足：

(1) $\varphi: V\to V^*$ 是同构 $\Longleftrightarrow$ $\Phi_1(\varphi)$ 非退化。

(2) $\varphi=\varphi^*$ (此处 $\varphi^*: V\cong(V^*)^*\xrightarrow{\varphi^*}V^*$) $\Longleftrightarrow$ $\Phi_1(\varphi)$ 对称。

证明：不难验证 Φ_1 是 K-线性映射，且映射

$$\Psi_1:\ \mathcal{L}_2(V)\longrightarrow\mathcal{L}(V,V^*)$$
$$f\longmapsto\Psi_1(f): V\to V^*,\quad x\longmapsto \overset{V^*}{f(x,\cdot)}: V\to K$$

是 Φ_1 的逆映射。为证明(1)，(2)，可设 $e_1,\cdots,e_n\in V$ 是一组基，$e_1^*,\cdots,e_n^*\in V^*$ 是它的对偶基。令

$$(\varphi(e_1),\cdots,\varphi(e_n))=(e_1^*,\cdots,e_n^*)A_\varphi,\quad A_\varphi=(a_{ij})_{n\times n}.$$

则 $\Phi_1(\varphi)(e_i,e_j)=\varphi(e_i)(e_j)=(a_{1i}e_1^*+\cdots+a_{ni}e_n^*)(e_j)=a_{ji}$

所以，$\left(\Phi_1(\varphi)(e_i,e_j)\right)_{n\times n} = {}^tA_\varphi$．$\varphi$ 是同构 $\Leftrightarrow A_\varphi$ 可逆

$\Leftrightarrow \left(\Phi_1(\varphi)(e_i,e_j)\right)_{n\times n}$ 可逆 $\Leftrightarrow \Phi_1(\varphi): V\times V\to K$ 非退化。

又 $(\varphi^*(e_1),\cdots,\varphi^*(e_n)) = (e_1^*,\cdots,e_n^*)\,{}^tA_\varphi$．所以，$\varphi=\varphi^* \Leftrightarrow A_\varphi = {}^tA_\varphi$

$\Leftrightarrow \Phi_1(\varphi)(e_i,e_j) = \Phi_1(\varphi)(e_j,e_i) \Leftrightarrow \Phi_1(\varphi): V\times V\to K$ 对称

□．

<u>注记</u>：显然，还有同构 $\Phi_2: \mathcal{L}(V,V^*) \longrightarrow \mathcal{L}_2(V)$：

$$\Phi_2(\varphi): V\times V\longrightarrow K,\quad (x,y)\longmapsto \varphi(y)(x).$$

它的逆映射为 $\Psi_2: \mathcal{L}_2(V)\longrightarrow \mathcal{L}(V,V^*)$：对任意 $f\in\mathcal{L}_2(V)$

$$\Psi_2(f): V\longrightarrow V^*,\quad x\longmapsto f(\cdot,x)\in V^*;$$

练习：对任意 $\varphi\in\mathcal{L}(V,V^*)$，证明：

$$\Psi_2\Phi_1(\varphi) = \varphi^*,$$ 其中 $\varphi^*: V\to V^*$ 定义为：

$\varphi^*(x): V\to K$，$\varphi^*(x)(y) = \varphi(y)(x)$．$(\forall\, x\in V, y\in V)$，

□

<u>命题 6.5.2</u>：设 n 维 K-线性空间 V 带有一个指定的<u>非退化</u>，<u>对称</u>，<u>双线性函数</u>：$(\cdot|\cdot): V\times V\longrightarrow K$。则对任意 $A\in\mathcal{L}(V)$，存在唯一的<u>线性算子</u> $A^*\in\mathcal{L}(V)$ 使得

$$(Ax\mid y) = (x\mid A^*y)$$

对任意 $x,y\in V$ 成立。且对任意 $A,B\in\mathcal{L}(V)$ 有

(1) $\forall\, \lambda,\mu\in K$，$(\lambda A+\mu B)^* = \lambda A^* + \mu B^*$

(2) $(AB)^* = B^*\cdot A^*$，(3) $(A^*)^* = A$．

证明：该双线性函数定义了同构 $\tau: V\longrightarrow V^*$，$\tau(x)=(\cdot|x)\in V^*$

对任意 $\mathcal{A}\in\mathcal{L}(V)$，令 $\mathcal{A}^\vee: V^*\to V^*$ 表示 $\mathcal{A}$ 的对偶：

$\forall\, l\in V^*$, $\mathcal{A}^\vee(l): V\to K$, $x\mapsto l(\mathcal{A}x)$.

则 $\mathcal{A}^* := \tau^{-1}\mathcal{A}^\vee\tau : V\xrightarrow{\tau}V^*\xrightarrow{\mathcal{A}^\vee}V^*\xrightarrow{\tau^{-1}}V$

满足：$(\mathcal{A}x\,|\,y)=(x\,|\,\mathcal{A}^*y)$, $\forall\, x,y\in V$.

另外，由于双线性函数 $(\cdot|\cdot): V\times V\to K$ 非退化，命题6.5.2中的结论 (1),(2),(3) 等价于：$\forall\, x,y\in V$ 有

(1) $(x\,|\,(\lambda\mathcal{A}+\mu\mathcal{B})^*y)=(x\,|\,\lambda\mathcal{A}^*y+\mu\mathcal{B}^*y)$

(2) $(x\,|\,(\mathcal{A}\mathcal{B})^*y)=(x\,|\,\mathcal{B}^*(\mathcal{A}^*y))$

(3) $(x\,|\,(\mathcal{A}^*)^*y)=(x\,|\,\mathcal{A}y)$。

即上述等式可由 $\mathcal{A}^*$ 的定义直接验证：

(1) $(x\,|\,(\lambda\mathcal{A}+\mu\mathcal{B})^*y)=((\lambda\mathcal{A}+\mu\mathcal{B})x\,|\,y)=\lambda(\mathcal{A}x\,|\,y)+\mu(\mathcal{B}x\,|\,y)$

$=\lambda(x\,|\,\mathcal{A}^*y)+\mu(x\,|\,\mathcal{B}^*y)=(x\,|\,\lambda\mathcal{A}^*y+\mu\mathcal{B}^*y)$.

(2) $(x\,|\,(\mathcal{A}\mathcal{B})^*y)=(\mathcal{A}(\mathcal{B}x)\,|\,y)=(\mathcal{B}x\,|\,\mathcal{A}^*y)=(x\,|\,\mathcal{B}^*(\mathcal{A}^*y))$.

(3) $(x\,|\,(\mathcal{A}^*)^*y)=(\mathcal{A}^*x\,|\,y)=(y\,|\,\mathcal{A}^*x)=(\mathcal{A}y\,|\,x)=(x\,|\,\mathcal{A}y)$.

□

定义6.5.1：$\mathcal{A}^*\in\mathcal{L}(V)$ 称为 $\mathcal{A}\in\mathcal{L}(V)$（在 $(\cdot|\cdot):V\times V\to K$ 下）的 对偶算子（或称为 $\mathcal{A}$ 的伴随算子，相伴算子等）。

□

~~定义6.5.2 $\mathcal{A}$ 称为自伴算子，如果 $\mathcal{A}=\mathcal{A}^*$~~.

从现在开始，我们仅考虑欧氏空间 $V=(V,(\cdot|\cdot))$.
即 V 是 n 维实向量空间，$(\cdot|\cdot): V\times V\longrightarrow\mathbb{R}$ 是 V 上的内积.
从而有下面简单（但重要）的事实.

引理 6.5.1：设 V 是 n 维欧氏空间，$A\in\mathcal{L}(V)$，$A^*\in\mathcal{L}(V)$. 则

(1) 对任意子空间 $W\subset V$，有 $V=W\oplus W^\perp$，其中
$$W^\perp=\{x\in V \mid (x|y)=0,\ \forall y\in W\}$$
称为 W 的正交补.

(2) 如果 W 是 A 的不变子空间，则 $W^\perp$ 是 A^* 的不变子空间.

(3) 如果 $e_1,\cdots,e_n\in V$ 是标准正交基，令
$$(Ae_1,\cdots,Ae_n)=(e_1,\cdots,e_n)A$$
$$(A^*e_1,\cdots,A^*e_n)=(e_1,\cdots,e_n)A^*.$$
则 $A^*={}^tA$.

证明：(1) 设 $e_1,\cdots,e_m\in W$ 是 W 的标准正交基，则 $e_1,\cdots,e_m$ 可扩充成 V 的标准正交基 $e_1,\cdots,e_m,e_{m+1},\cdots,e_n\in V$. 则
$$W^\perp=\langle e_{m+1},\cdots,e_n\rangle,\quad V=W\oplus W^\perp.$$

(2) 只需证明：$\forall\ x\in W^\perp$，有 $A^*x\in W^\perp$，这等价于证明：$\forall y\in W$，有 $(A^*x|y)=0$。由 A^* 的定义：
$$(A^*x|y)=(y|A^*x)=(Ay|x)=(x|Ay)=0\quad(\text{由于 } Ay\in W).$$

(3) 设 $A=(a_{ij})_{n\times n}$，$A^*=(a^*_{ij})_{n\times n}$. 则
$$A^*e_j=a^*_{1j}e_1+\cdots+a^*_{ij}e_i+\cdots+a^*_{nj}e_n$$
$$Ae_i=a_{1i}e_1+\cdots+a_{ji}e_j+\cdots+a_{ni}e_n.$$
所以，$a^*_{ij}=(e_i|A^*e_j)=(Ae_i|e_j)=a_{ji}$. 即 $A^*={}^tA$.

$\square$.

定义 6.5.2：设 $\mathcal{A} \in \mathcal{L}(V)$ 是欧氏空间上的线性算子，则

(1) $\mathcal{A}$ 称为<u>对称算子</u>，如果 $\mathcal{A}^* = \mathcal{A}$；

(2) $\mathcal{A}$ 称为<u>反对称算子</u>，如果 $\mathcal{A}^* = -\mathcal{A}$；

(3) $\mathcal{A}$ 称为<u>正交算子</u>，如果 $\mathcal{A}^* = \mathcal{A}^{-1}$；

(4) $\mathcal{A}$ 称为<u>正规算子</u>，如果 $\mathcal{A}\mathcal{A}^* = \mathcal{A}^*\mathcal{A}$。 □

引理 6.5.2：设 $\mathcal{A} \in \mathcal{L}(V)$ 满足上述定义中条件 (1), (2), (3) 中之一，则

(1) $W \subset V$ 是 $\mathcal{A}$ 的不变子空间 $\Longleftrightarrow$ W 是 $\mathcal{A}^*$ 的不变子空间.

(2) $V = V_1 \oplus \cdots \oplus V_m$，其中 V_i 是 $\mathcal{A}$ 的不变子空间，且 $1 \le \dim_{\mathbb{R}} V_i \le 2$，$V_i \perp V_j$ $(i \neq j)$.

证明：(1) 如果 $\mathcal{A}$ 满足条件 (1) 或 (2)，结论是显然。因此，设 $\mathcal{A}$ 是正交算子。如果 W 是 $\mathcal{A}$ 的不变子空间，则 $\mathcal{A}|_W : W \to W$ 是双射，所以，$\forall\ x \in W$，存在 $y \in W$ 使 $\mathcal{A}y = x$. 从而 $\mathcal{A}^* x = \mathcal{A}^{-1} x = y \in W$，即 W 是 $\mathcal{A}^*$ 的不变子空间。反之，如果 W 是 $\mathcal{A}^*$ 的不变子空间，则 $\mathcal{A}^*|_W : W \to W$ 是双射，所以，$\forall\ x \in W$，存在 $y \in W$ 使

$$\mathcal{A}^* y = \mathcal{A}^{-1} y = x$$

即 $\mathcal{A}x = y \in W$. 因此，W 是 $\mathcal{A}$ 的不变子空间.

(2) 由引理 6.5.1 和上述结论 (1)，如果 $W \subset V$ 是 $\mathcal{A}$ 的不变子空间，则 $V = W \oplus W^{\perp}$，$W^{\perp}$ 也是 $\mathcal{A}$ 的不变子空间。所以，我们只需证明：实向量空间 V 上的任意线性算子 $\mathcal{A}$

都有1维或2维的不变子空间。事实上，设$\mu_A(x)$是A的极小多项式，$p(x)$是$\mu_A(x)$的一个不可约因子，即

$$\mu_A(x) = p(x) f(x), \quad p(x) \in \mathbb{R}[x] \text{ 不可约}.$$

则 $p(x) = x - \lambda$，或 $p(x) = x^2 + bx + c$ $(b^2 - 4c < 0)$。由$\mu_A(x)$是极小多项式，存在 $\alpha \in V$ 使 $f(A)\cdot\alpha \neq 0$。令 $\beta = f(A)\cdot\alpha$，则

$$p(A)\cdot\beta = 0, \quad \beta \neq 0.$$

当 $p(x) = x - \lambda$ 时，$A\beta = \lambda\beta$，即 $\langle\beta\rangle \subset V$ 是A的1维不变子空间。当 $p(x) = x^2 + bx + c$ 时，$A^2\beta = -c\beta - bA\beta$。所以 $\langle\beta, A\beta\rangle \subset V$ 是A的2维不变子空间。

□

定理6.5.1：A是对称算子（亦称自伴算子）$\Leftrightarrow$ 存在标准正交基 $e_1, e_2, \cdots, e_n \in V$ 使

$$(Ae_1, \cdots, Ae_n) = (e_1, \cdots, e_n)\begin{pmatrix} \lambda_1 & & & 0 \\ & \lambda_2 & & \\ & & \ddots & \\ 0 & & & \lambda_n \end{pmatrix}.$$

证明：设 $\mu_A(x) \in \mathbb{R}[x]$ 是 A 的极小多项式，$p(x) \in \mathbb{R}[x]$ 不可约，使得 $\mu_A(x) = p(x) f(x)$。只需证明：$p(x)$是一次多项式。否则，$p(x) = x^2 + bx + c$，$b^2 - 4c < 0$。存在 $\alpha \in V$ 使 $\beta := f(A)\cdot\alpha \neq 0$，令 $W = \langle\beta, A\beta\rangle$，$A_W = A|_W : W \to W$。则 $A_W^* = A_W$（事实上，$A_W^* = A^*|_W$）$\chi_{A_W}(x) = p(x)$。设 $e_1, e_2 \in W$ 是标准正交基，则

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

得 $p(x) = \chi_{A_W}(x) = \begin{vmatrix} x - a_{11} & -a_{12} \\ -a_{12} & x - a_{22} \end{vmatrix} = x^2 - (a_{11} + a_{22}) + a_{11}a_{22} - a_{12}^2$。

$b^2 - 4c = (a_{11} + a_{22})^2 - 4(a_{11}a_{22} - a_{12}^2) = (a_{11} - a_{22})^2 + 4a_{12}^2 \geq 0$ 矛盾。

□

定理 6.5.2：$\mathcal{A}$ 是反对称算子 $\Leftrightarrow$ 存在标准正交基 $e_1,\cdots,e_n\in V$ 使得

$$(\mathcal{A}e_1,\cdots,\mathcal{A}e_n)=(e_1,\cdots,e_n)\begin{pmatrix}\boxed{\begin{matrix}0&\lambda_1\\-\lambda_1&0\end{matrix}} & & & & & \\ & \ddots & & & 0 & \\ & & \boxed{\begin{matrix}0&\lambda_s\\-\lambda_s&0\end{matrix}} & & & \\ & 0 & & 0 & & \\ & & & & \ddots & \\ & & & & & 0\end{pmatrix}$$

证明：由引理 6.5.2，无妨设 $V=V_1\oplus\cdots\oplus V_s\oplus V_{s+1}\oplus\cdots\oplus V_m$，其中 $V_1,\cdots,V_s$ 是 $\mathcal{A}$ 的（不可分解）2维不变子空间，而 $V_{s+1},\cdots,V_m$ 是 $\mathcal{A}$ 的1维不变子空间，且 $V_i\perp V_j$（$\forall\ i\neq j,1\le i,j\le m$）。由于对于 $\mathcal{A}$ 的任意不变子空间 $W\subset V$，$\mathcal{A}_W=\mathcal{A}|_W:W\to W$ 也是 W 上的反对称算子，所以 $\mathcal{A}|_{V_i}=0$（$s<i\le m$），而 $\mathcal{A}|_{V_i}$（$1\le i\le s$）在 V_i 的标准正交基下的坐标矩阵必为

$$A_i=\begin{pmatrix}0&\lambda_i\\-\lambda_i&0\end{pmatrix}.\qquad(1\le i\le s).$$

令 $e_1,\cdots,e_n$ 是由 $V_1,V_2,\cdots,V_m$ 的标准正交基组成的 V 的一组基，则 $e_1,\cdots,e_n$ 是 V 的标准正交基（由于 $V_i\perp V_j$），且 $\mathcal{A}$ 在它下的坐标矩阵为

$$\begin{pmatrix}A_1&&&&&\\&\ddots&&&0&\\&&A_s&&&\\&&&0&&\\&0&&&\ddots&\\&&&&&0\end{pmatrix}$$

□.

在讨论正交算子标准型之前，我们先给一个正交算子的等价刻划。

命题 6.5.3：设 $\mathcal{A}:V\to V$ 是欧氏空间 V 上的线性算子，则下列结论等价：

(1) $\mathcal{A}$ 是正交算子；(2) $e_1,\cdots,e_n\in V$ 是标准正交基 $\Leftrightarrow$ $\mathcal{A}e_1,\cdots,\mathcal{A}e_n$ 是标准正交基；

(3) $\forall\ x\in V$，$\|\mathcal{A}x\|=\|x\|$；(4) $\forall\ x,y\in V$，$(\mathcal{A}x|\mathcal{A}y)=(x|y)$。

证明：(1)⇒(2)：由于 $A^*=A^{-1}$ 得 $(Ae_i|Ae_j)=(e_i|e_j)$，所以 $e_1,\cdots,e_n$ 是标准正交基 $\Longleftrightarrow$ $Ae_1,\cdots,Ae_n$ 是标准正交基.

(2)⇒(3)：设 $e_1,\cdots,e_n\in V$ 是标准正交基，$x=x_1e_1+\cdots+x_ne_n$ 则，由于 $Ae_1,\cdots,Ae_n$ 也是 V 的标准正交基，我们有：

$$\|Ax\|=\sqrt{x_1^2+\cdots+x_n^2}=\|x\|.$$

(3)⇒(4)：$\forall\, x,y\in V$，由 $(Ax-Ay|Ax-Ay)=\|A(x-y)\|^2$ 和

$$(Ax-Ay|Ax-Ay)=(Ax|Ax)+(Ay|Ay)-2(Ax|Ay).$$ 得

$$(Ax|Ay)=\frac{1}{2}\left(\|Ax\|^2+\|Ay\|^2-\|A(x-y)\|^2\right)$$
$$=\frac{1}{2}\left(\|x\|^2+\|y\|^2-\|x-y\|^2\right)=(x|y).$$

(4)⇒(1)：$\forall\, x,y\in V$，$(x|y)=(Ax|Ay)=(x|A^*Ay)$.

所以 $A^*Ay=y$ 对任意 $y\in V$ 成立。因此，$A^*=A^{-1}$

□.

例 6.5.1（平面的旋转）. $V=\{x+yi\mid x,y\in\mathbb{R}\}=\mathbb{C}$.

定义：$(\cdot|\cdot): V\times V\longrightarrow\mathbb{R},\ (\alpha,\beta)\longmapsto \mathrm{Re}(\alpha\bar{\beta})$.

则 $(\cdot|\cdot)$ 是 V 的一个内积. $e_1=1$，$e_2=i$ 是 V 的一组标准正交基.

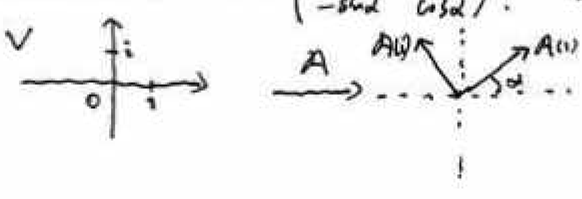
因为 $e^{i\alpha}=\cos\alpha+i\sin\alpha\in\mathbb{C}$，则

$$A: V\longrightarrow V,\quad z\longmapsto e^{i\alpha}\cdot z$$

是一个线性算子，它在 $e_1=1, e_2=i$ 下的坐标矩阵是

$$(A(1),A(i))=(1,i)\begin{pmatrix}\cos\alpha & -\sin\alpha\\ \sin\alpha & \cos\alpha\end{pmatrix}$$

$$(A^*(1),A^*(i))=(1,i)\begin{pmatrix}\cos\alpha & \sin\alpha\\ -\sin\alpha & \cos\alpha\end{pmatrix}.$$ A 是正交算子，表示将平面旋转变换.



定理 6.5.3: $\mathcal{A}$ 是正交算子 $\Longleftrightarrow$ 存在标准正交基 $e_1,\cdots,e_n\in V$

使
$$(\mathcal{A}e_1,\cdots,\mathcal{A}e_n)=(e_1,\cdots,e_n)\begin{pmatrix} \boxed{\begin{matrix}\cos\varphi_1 & -\sin\varphi_1\\ \sin\varphi_1 & \cos\varphi_1\end{matrix}} & & & & 0\\ & \ddots & & & \\ & & \boxed{\begin{matrix}\cos\varphi_r & -\sin\varphi_r\\ \sin\varphi_r & \cos\varphi_r\end{matrix}} & & \\ & & & -I_k & \\ 0 & & & & I_l \end{pmatrix}$$

其中 $k+l+2r=n$.

证明: 由引理 6.5.2 可设 $V=V_1\oplus\cdots\oplus V_r\oplus V_{r+1}\oplus\cdots\oplus V_m$ 其中 V_i $(1\le i\le r)$ 是 $\mathcal{A}$ 的 2 维不可分解不变子空间, $m=n-2r$, $V_{r+1},\cdots,V_m$ 是 $\mathcal{A}$ 的 1 维不变子空间, 且 $V_i\perp V_j$ $(i\neq j)$. 令

$$e_{2r+1}\in V_{r+1},\ e_{2r+2}\in V_{r+2},\ \cdots,\ e_n\in V_{n-2r}=V_m .$$

满足 $\|e_{2r+i}\|=1$, 则 $\mathcal{A}e_{2r+i}=\pm e_{2r+i}$. 所以无妨设

当 $1\le i\le k$ 时, $\mathcal{A}e_{2r+i}=-e_{2r+i}$. 在 2 维不变子空间 V_i $(1\le i\le r)$ 上, $\mathcal{A}_i=\mathcal{A}|_{V_i}: V_i\to V_i$ 是 2 维欧氏空间上的正交算子, 且极小多项式 $\mu_{\mathcal{A}_i}(x)\in\mathbb{R}[x]$ 是二次不可约多项式. 因此, 定理的证明由下列引理完成. □

引理 6.5.3: 设 $\mathcal{A}: V\to V$ 是 2 维欧氏空间上的正交算子. 如果 $\mu_{\mathcal{A}}(x)\in\mathbb{R}[x]$ 是二次不可约多项式, 则存在标准正交基 $e_1,e_2\in V$ 使

$$(\mathcal{A}e_1,\mathcal{A}e_2)=(e_1,e_2)\begin{pmatrix}\cos\varphi & -\sin\varphi\\ \sin\varphi & \cos\varphi\end{pmatrix}.$$

证明: 设 $e_1,e_2\in V$ 是一组标准正交基, $(\mathcal{A}e_1,\mathcal{A}e_2)=(e_1,e_2)A$.

其中 $A=\begin{pmatrix}a & b\\ c & d\end{pmatrix}\in M_2(\mathbb{R})$ 是正交矩阵.

由于 $\mu_A(x)=\chi_A(x)=x^2-(a+d)x+|A|$ 没有实根，$|A|=1$.

又 $A^{-1}=\begin{pmatrix} d, & -b \\ -c, & a \end{pmatrix}={}^tA=\begin{pmatrix} a & c \\ b & d \end{pmatrix}$，所以 $A=\begin{pmatrix} a & -c \\ c & a \end{pmatrix}$，$a^2+c^2=1$.

令 $a=\cos\varphi$，$c=\sin\varphi$. 即得

$$(\mathcal{A}e_1,\mathcal{A}e_2)=(e_1,e_2)\begin{pmatrix} \cos\varphi, & -\sin\varphi \\ \sin\varphi, & \cos\varphi \end{pmatrix}.$$

$\square$.

对称算子，反对称算子，正交算子都是正规算子的特例。但下面我们将会看到，正规算子与正交算子其实"相差不远"。

命题 6.5.4：设 $\mathcal{A}: V\to V$ 是正规算子．则

(1) 对任意整数 $k\geq 2$，$r(\mathcal{A}^k)=r(\mathcal{A})$。特别，任意幂零正规算子必为零算子。

(2) $\mathcal{A}$ 的极小多项式 $\mu_A(x)$ 没有重因子．即

$$\mu_A(x)=P_1(x)\cdots P_m(x),\quad P_i(x) \text{ 不可约 且 } P_i(x)\neq P_j(x).$$

证明：只需证明：$r(\mathcal{A}^k)=r(\mathcal{A}^{k-1})$，即 $\dim_K \mathrm{Im}(\mathcal{A}^k)=\dim_K \mathrm{Im}(\mathcal{A}^{k-1})$. 显然，$\mathrm{Im}(\mathcal{A}^k)\subseteq \mathrm{Im}(\mathcal{A}^{k-1})$（当 $k\geq 2$），只需证明：

$$\dim_K \mathrm{Im}(\mathcal{A}^{k-1})\leq \dim_K \mathrm{Im}(\mathcal{A}^k).$$

考虑：$V\xrightarrow{\mathcal{A}^{k-1}}V\xrightarrow{\mathcal{A}}V$，只需证明：$\mathrm{Im}(\mathcal{A}^{k-1})\xrightarrow{\mathcal{A}}\mathrm{Im}(\mathcal{A}^k)$ 是单射，即：$\mathrm{Im}(\mathcal{A}^{k-1})\cap \mathrm{Ker}(\mathcal{A})=\{0\}$．注意到：对任意线性算子 $\mathcal{A}: V\to V$ 有 $V=\mathrm{Im}(\mathcal{A})\oplus \mathrm{Im}(\mathcal{A})^{\perp}$，而且

$$\mathrm{Im}(\mathcal{A})^{\perp}=\mathrm{Ker}(\mathcal{A}^*).$$

（即 $\mathcal{A}$ 是正规算子 $\Longleftrightarrow \forall x,y\in V$，$(\mathcal{A}x|\mathcal{A}y)=(\mathcal{A}^*x|\mathcal{A}^*y)$，特别可得

$$\mathrm{Ker}(\mathcal{A}^*)=\mathrm{Ker}(\mathcal{A}).$$

所以, $\mathrm{Im}(A^{k-1})\cap \ker(A) = \mathrm{Im}(A^{k-1})\cap \mathrm{Im}(A)^{\perp} = \{0\}$.

(2) 设 $\mu_A(x)=p_1(x)^{k_1}\cdots p_m(x)^{k_m}$ 是 $\mu_A(x)$ 的不可约分解. 令 $f_i(x)=\frac{\mu_A(x)}{p_i(x)^{k_i}}$
则 $f_1(x),\cdots,f_m(x)$ 互素, 存在 $h_1(x),\cdots,h_m(x)\in\mathbb{R}[x]$ 使

$$f_1(x)h_1(x)+\cdots+f_m(x)h_m(x)=1.$$

令 $V_i=\mathrm{Im}(f_i(A)h_i(A))=f_i(A)\cdot h_i(A)\cdot V\subset V$, 则

$$V=V_1\oplus\cdots\oplus V_m,\qquad A(V_i)\subset V_i,\ A^*(V_i)\subset V_i.$$

且 $\mu_{A_i}(x)=p_i(x)^{k_i}$ (令 $A_i=A|_{V_i}$). 不难验证: $A_i^*=A^*|_{V_i}$, A_i 也是 V_i 上的正规算子. 所以, $p_i(A_i)$ 也是 V_i 上的正规算子. 但 $p_i(A_i)^{k_i}=\mu_{A_i}(A_i)=0$. 因此, $p_i(A_i)=0$
从而必有 $k_i=1$

□.

~~因此, 对任意正规算子 $A: V\to V$, $\mu_A(x)=p_1(x)\cdots p_m(x)$.~~

命题 6.5.5: 设 $A: V\to V$ 是正规算子, $\mu_A(x)=p_1(x)\cdots p_m(x)$ 是极小多项式的不可约分解. 令

$$V_i=\ker(p_i(A))\subset V\quad (1\le i\le m).$$

$A_i=A|_{V_i}$ $(1\le i\le m)$. 则

(1) $V=V_1\oplus\cdots\oplus V_m$ 是正交分解 ($V_i\perp V_j$).

(2) 存在 $\lambda_i\in\mathbb{R}$ 和正交算子 $\breve{A}_i: V_i\to V_i$ 使 $A_i=\lambda_i\breve{A}_i$.

证明: (1) 令 $f_i(x)=\frac{\mu_A(x)}{p_i(x)}$, 则存在 $h_i(x)\in\mathbb{R}[x]$ 使

$$f_1(x)h_1(x)+\cdots+f_m(x)h_m(x)=1$$

令 $W_i=f_i(A)h_i(A)\cdot V=\mathrm{Im}(f_i(A)h_i(A))\subset V$,

则 $W_i \subseteq \ker(P_i(A))$ （因为 $P_i(A)f_i(A)=\mu_A(A)=0$），

$$V = W_1\oplus\cdots\oplus W_m \subset V_1\oplus\cdots\oplus V_m \subset V.$$

所以，$V=V_1\oplus\cdots\oplus V_m$，$V_i=W_i=\mathrm{Im}(f_i(A)h_i(A))$.

下面证明：当 A 是正规算子时，$V=V_1\oplus\cdots\oplus V_m$ 是正交分解。由 $P_i(x)$，$P_j(x)$ 互素，存在 $g_1(x)$，$g_2(x)\in\mathbb{R}[x]$ 使 $g_1(x)P_i(x)+g_2(x)P_j(x)=1$，从而

$$g_1(A)P_i(A)+g_2(A)P_j(A)=1$$

是恒等算子。对任意 $x\in V_i$，$y\in V_j$，有

$$x=\big(g_1(A)P_i(A)+g_2(A)P_j(A)\big)\cdot x = P_j(A)g_2(A)\cdot x,$$

及 $(x\,|\,y)=(P_j(A)g_2(A)\cdot x\,|\,y)=(g_2(A)x\,|\,P_j(A)^*y)$。

当 A 是正规算子时，$P_j(A)$ 也是正规算，从而

$$\ker(P_j(A)^*)=\ker(P_j(A)).$$

由 $y\in V_j=\ker(P_j(A))$，得 $P_j(A)^*y=0$。所以 $(x|y)=0$。

(2) $\mu_{A_i}(x)=P_i(x)\in\mathbb{R}[x]$ 不可约。所以 $\deg P_i(x)\le 2$.

如果 $\mu_{A_i}(x)=x-\lambda_i$ ($\lambda_i\in\mathbb{R}$)。则 $A_i=\lambda_i\cdot 1$（1是正交算子）.

如果 $\mu_{A_i}(x)=x^2+b_ix+c_i$ 且 $b_i^2-4c_i<0$。则

$$A_i^2+b_iA_i+c_i=0\ ,\quad A_i^{*2}+b_iA_i^*+c_i=0\ .$$

注意 $V_i=W_i=\mathrm{Im}(f_i(A)h_i(A))$ 也是 A^* 的不变子空间，$A_i^*=A^*|_{V_i}$.
所以，A_i 也是 V_i 上的正规算子，从而

$$(A_iA_i^*-c_i\cdot 1)(A_i-A_i^*)=(A_i^2+b_iA_i+c_i)A_i^*-(A_i^{*2}+b_iA_i^*+c_i)A=0.$$

如果 $A_i-A_i^*:V_i\to V_i$ 不是同构，则 $K=\ker(A_i-A_i^*)\subset V_i$ 是 A_i 的非零不变子空间，在 K 上有 $A_i=A_i^*$。所以 $A_i|_K$ 是 K 上的对称算子，从而存在 $0\ne\alpha\in K$ 使 $A_i\alpha=\lambda\alpha$，$\mu_{A_i}(\lambda)=0$ 矛盾.

因此，由 $(A_iA_i^*-c_i\cdot 1)(A_i-A_i^*)=0$ 可得

$$A_iA_i^*=c_i1.$$

令 $\tilde{A}_i=\frac{1}{\sqrt{c_i}}A_i$ （因为 $4c_i>b_i^2\geq 0$），$\lambda_i=\sqrt{c_i}$，则 $A_i=\lambda_i\tilde{A}_i$，$\tilde{A}_i$ 是 V_i 上的正交算子.

□

推论 6.5.1：如果 $A:V\to V$ 是正规算子，$W\subset V$ 是 A 的不变子空间，则 $W^\perp\subset V$ 也是 A 的不变子空间.

证明：设 $V=V_1\oplus\cdots\oplus V_m$，$V_i=\ker(p_i(x))$，是命题 6.5.5 中的正交直和分解，$A_i=A|V_i\in\mathcal{L}(V_i)$。则，由命题 6.5.5 (2)，

$$A_i^*=\lambda_i\tilde{A}_i^*=\lambda_i\tilde{A}_i^{-1}=\lambda_i^2A_i^{-1}.$$

$\forall\ x\in W^\perp$，$y\in W$，令 $y=y_1+\cdots+y_m$，$y_i\in V_i$. 则

$$(Ax\mid y)=(x\mid A^*y)=\sum_{i=1}^m(x\mid A^*y_i)=\sum_{i=1}^m(x\mid A_i^*y_i)$$

如果 $A_i^*y_i\in W$，则 $(x\mid A_i^*y_i)=0$，从而 $(Ax\mid y)=0$. 即 $Ax\in W^\perp$.

所以，只需说明：在分解 $y=y_1+\cdots+y_m$ 中，$y_i\in V_i\cap W$.

而由命题 6.5.5 (1) 证明中的恒等算子分解

$$1=f_1(A)h_1(A)+\cdots+f_m(A)h_m(A)$$

可得 $y=f_1(A)h_1(A)y+\cdots+f_m(A)h_m(A)\cdot y$，其中

$$f_i(A)h_i(A)y\in V_i\cap W$$ （因为 $y\in W$，W 是 A 的不变子空间）.

由于 $V=V_1\oplus\cdots\oplus V_m$ 是直和，可得

$$y_i=f_i(A)h_i(A)y\in V_i\cap W$$

□

推论 6.5.2：$A\in\mathcal{L}(V)$ 是正规算子 $\Longleftrightarrow$ 存在一组标准正交基 $e_1,e_2,\cdots,e_n\in V$，使.

$$(Ae_1,Ae_2,\cdots,Ae_n)=(e_1,\cdots,e_n)A\cdot,$$ 其中,

$$A=\begin{pmatrix} \boxed{\begin{matrix} a_1 & -b_1 \\ b_1 & a_1 \end{matrix}} & & & & & 0 \\ & \ddots & & & & \\ & & \boxed{\begin{matrix} a_r & -b_r \\ b_r & a_r \end{matrix}} & & & \\ & & & \mu_1 & & \\ & & & & \ddots & \\ 0 & & & & & \mu_l \end{pmatrix}_{n\times n}, \quad n=2r+l.$$

证明：对任意线性算子 $\mathcal{A}\in\mathcal{L}(V)$，及 $\mathcal{A}$ 的不变子空间 $W\subset V$，有 $V=W\oplus W^\perp$，$W^\perp$ 是 $\mathcal{A}^*$ 的不变子空间。当 $\mathcal{A}$ 是<u>正规算子</u>时，由推论 6.5.1，$W^\perp$ 也是<u>$\mathcal{A}$ 的不变子空间</u>，从而 $\mathcal{A}|_{W^\perp}$ 是 $W^\perp$ 的<u>正规算子</u>且 $(\mathcal{A}|_{W^\perp})^*=\mathcal{A}^*|_{W^\perp}$。所以，$V=L_1\oplus\cdots\oplus L_r\oplus L_{r+1}\oplus\cdots\oplus L_{n-r}$，其中 L_i 是 $\mathcal{A}$ 的<u>不可分解</u>不变子空间，$L_i\perp L_j$ $(i\neq j)$，且

$$\dim_{\mathbb{R}} L_i=2,\ (1\le i\le r),\quad \dim_{\mathbb{R}} L_i=1.\ (r+1\le i\le n-r),$$

$\mathcal{A}_i=\mathcal{A}|_{L_i}$ 是 L_i 上<u>正规算子</u>。当 $1\le i\le r$ 时，$\mathcal{A}_i$ 在 L_i 的<u>标准正交基</u>下的矩阵必形如

$$\begin{pmatrix} a_i & -b_i \\ b_i & a_i \end{pmatrix} \quad (1\le i\le r).$$

□.

<u>推论 6.5.3</u>：设 $e_1,\cdots,e_n\in V$ 是标准正交基，

$$(\mathcal{A}e_1,\cdots,\mathcal{A}e_n)=(e_1,\cdots,e_n)A.$$

则 (1) $\mathcal{A}$ 是<u>对称算子</u> $\Longleftrightarrow$ ${}^tA=A$ （即对称矩阵）

(2) $\mathcal{A}$ 是<u>反对称算子</u> $\Longleftrightarrow$ ${}^tA=-A$ （反对称矩阵）

(3) $\mathcal{A}$ 是<u>正交算子</u> $\Longleftrightarrow$ ${}^tA=A^{-1}$ （正交矩阵）.

(4) $\mathcal{A}$ 是<u>正规算子</u> $\Longleftrightarrow$ ${}^tA\cdot A=A\cdot{}^tA$ （这样 A 称为正规矩阵）.

□

习题 6.5

1. 设 V 是欧氏空间，$\mathcal{A}\in\mathcal{L}(V)$ 是线性算子. 试证：

(1) $\mathrm{Im}(\mathcal{A})^{\perp}=\mathrm{Ker}(\mathcal{A}^*)$；

(2) 如果 $\mathcal{A}$ 是对称算子，则 $\mathcal{A}=0 \Leftrightarrow (\mathcal{A}x\mid x)=0,\ \forall x\in V$.

(3) 设 $\mathcal{A}$ 是反对称算子，则 $(\mathcal{A}x\mid x)=0,\ \forall x\in V$.

(4) $\mathcal{A}$ 是正规算子，则 $\mathrm{Ker}(\mathcal{A})=\mathrm{Ker}(\mathcal{A}^*)$.

2. 设 $\mathcal{A}$ 是欧氏空间 V 上的线性算子，$V=V_1\oplus\cdots\oplus V_s$ 是其不变子空间的正交分解（i.e. $V_i\perp V_j$）. 证明：

$\mathcal{A}$ 是正规算子 $\Leftrightarrow$ ① V_i 也是 $\mathcal{A}^*$ 的不变子空间；② $\mathcal{A}_i=\mathcal{A}|_{V_i}$ 是正规算子.

3. 设 $A\in M_n(\mathbb{R})$ 是对称矩阵，证明存在正交矩阵 $T\in M_n(\mathbb{R})$ 使 ${}^tTAT$ 是对角阵.

4. 设 $\alpha_1,\cdots,\alpha_n$ 是欧氏空间 V 的一组基（不必是标准正交），$\mathcal{A}\in\mathcal{L}(V)$。令 $B=(b_{ij})_{n\times n}\in M_n(\mathbb{R})$，$b_{ij}=(\alpha_i\mid\alpha_j)$. 及

$$(\mathcal{A}\alpha_1,\cdots,\mathcal{A}\alpha_n)=(\alpha_1,\cdots,\alpha_n)A,\quad (\mathcal{A}^*\alpha_1,\cdots,\mathcal{A}^*\alpha_n)=(\alpha_1,\cdots,\alpha_n)A^*.$$

(1) 求 A^* 与 A,B 的关系；(2) 求 $\mathcal{A}$ 是对称算子的充要条件.

5. 设 $\mathcal{A}$ 是 2 维欧氏空间上的正规算子，e_1,e_2 是标准正交基，

$$(\mathcal{A}e_1,\mathcal{A}e_2)=(e_1,e_2)A$$

证明：$A=\begin{pmatrix} a & b\\ -b & a\end{pmatrix}$ 或 $\begin{pmatrix} a & b\\ b & d\end{pmatrix}$.

6. 设 $\mathcal{A}\in\mathcal{L}(V)$ 是正规算子。如果 $\mathcal{A}^2=\mathcal{A}$，证明：$\mathcal{A}$ 必为对称算子（提示：$\forall x\in V,\ x-\mathcal{A}x\in\mathrm{Ker}(\mathcal{A})=\mathrm{Ker}(\mathcal{A}^*)$）.

7. 设 $A\in M_n(\mathbb{R})$ 满足 ${}^tAA=A\,{}^tA$。如果 $A^2=A$，试用矩阵方法证明：${}^tA=A$。

第七章. 仿射几何与欧氏几何.

对于我们生活的三维空间，我们观察到的不仅仅是点的集合，而且是包含点，直线，平面及它们之间相互位置的"几何直观"。在第二章我们通过"有向线段"的等价类(在平移下重合)来描述几何直观。例如，在"有向线段"等价类的集合V_3中定义了有几何意义的运算使得V_3成为一个3维实向量空间。同时在V_3中定义了有几何意义的函数：内积(对称双线性函数)，混合积(三重反对称线性函数)。通过这些运算和函数，可以解决现实三维空间的诸多几何问题。需要特别指出的是，在解决这些几何问题的计算过程中仅需利用这些"运算"和"函数"满足的规则，而无需关注它们的具体定义。

对于没有"几何直观"的空间(比如高维空间)，如何定义"有向线段的等价类"呢？在本章，我们将从一个给定的抽象实向量空间V开始，定义抽象集合$\mathbb{A}$中的有向线段等价类使得$\mathbb{A}$具有直线，平面，平行等几何概念(这样的$\mathbb{A}=(\mathbb{A},V)$称为与$V$相伴的仿射空间)。如果$V$是欧氏向量空间，这样的$\mathbb{A}=(\mathbb{A},V)$称为与$V$相伴的欧几里得空间。§7.1将引入仿射空间$\mathbb{A}=(\mathbb{A},V)$及其仿射子空间(即直线，平面)等概念。§7.2将在欧几里得空间$\mathbb{A}=(\mathbb{A},V)$中引入距离，直线之间的夹角，体积等概念。§7.3将讨论变换与几何的关系。§7.4将讨论仿射空间上二次函数的一般性质。§7.5将给出仿射空间和欧几里得空间中二次曲面的分类。§7.6将介绍射影空间的基本概念。

§7.1. 仿射空间.

在本节我们将通过一个给定的抽象向量空间 V 来定义集合 $\mathbb{A}$ 上的仿射空间结构(本质上是在集合 $\mathbb{A}$ 中定义有向线段). 为简单起见, 我们将假设 V 总是 n 维实向量空间.

<u>定义 7.1.1</u>. 设 $\mathbb{A}$ 是一非空集合, V 是 n 维实向量空间. 如果存在映射 $\mathbb{A}\times V \longrightarrow \mathbb{A}$, $(P,v)\longmapsto P\dotplus v$, 且满足条件

(1) $\forall\ P\in\mathbb{A}$, $u,v\in V$, 有 $P\dotplus 0 = P$ ($0\in V$ 是零向量) 及 $(P\dotplus u)\dotplus v = P\dotplus(u+v)$.

(2) $\forall\ P,Q\in\mathbb{A}$, 存在<u>唯一</u>向量 $v\in V$ 使得 $P\dotplus v = Q$. (此时称 v 是由 P 到 Q 的 ~~有向线段的等价类~~ 向量, 记为 $v=\overrightarrow{PQ}$).

则称集合 $\mathbb{A}$ 有一个与 V 相伴的仿射空间结构. $\mathbb{A}:=(\mathbb{A},V)$ 称为与 V 相伴的<u>仿射空间</u>. $\dim\mathbb{A}:=\dim(V)$.

□

注记 (1) 定义中的<u>仿射结构映射</u> $\mathbb{A}\times V\longrightarrow\mathbb{A}$, $(P,v)\longmapsto P\dotplus v$, 确定了一个满映射 $\mathbb{A}\times\mathbb{A}\longrightarrow V$, $(P,Q)\longmapsto\overrightarrow{PQ}$. 但 $\mathbb{A}\times\mathbb{A}\longrightarrow V$, $(P,Q)\longmapsto\overrightarrow{PQ}$, 一般不是单射. 所以我们将 $\overrightarrow{PQ}$ 视为从点 P 到点 Q 的有向线段. 而 v 是有向线段的等价类.

(2). 如果固定一个点 $O\in\mathbb{A}$, 则 $\mathbb{A}\longrightarrow V$, $P\longmapsto\overrightarrow{OP}$, 是双射, 它的逆映射是 $V\longrightarrow\mathbb{A}$, $v\longmapsto O\dotplus v$.

(3) 如果固定一个向量 $v\in V$, 则映射

$$T_v:\mathbb{A}\longrightarrow\mathbb{A},\quad P\longmapsto P\dotplus v,$$

是一个双射, 称为仿射空间 $\mathbb{A}$ 的<u>沿方向 v 的平移</u>.

□

引理7.1.1：所有平移的集合 $T(\mathbb{A}):=\{T_v \mid \forall v\in V\}$ 关于映射的合成是一个群.

证明：$\forall\, u, v\in V$，有 $T_u\cdot T_v = T_{u+v}$.（事实上，$\forall\, P\in\mathbb{A}$，有

$T_u\cdot T_v(P) = T_u(P\dotplus v) = (P\dotplus v)\dotplus u = P\dotplus(v+u) = T_{u+v}(P)$）.

故 $T(\mathbb{A})$ 关于映射的合成是封闭的. 另一方面，$T(\mathbb{A})$ 含有恒等映射（由零向量确定），且 $T_v\cdot T_{-v} = T_0 = e$. 故 $T(\mathbb{A})$ 是一个（交换）群

□

例7.1.1：设 $\mathbb{A}_3$ 表示我们所生活的三维空间中点的集合，V 表示 $\mathbb{A}_3$ 中有向线段等价类（在平移下等价）组成的向量空间.

~~$\mathbb{A}\times V \longrightarrow \mathbb{A},\ (P, v)\longmapsto P\dotplus v$~~

则对任意 $(P, v)\in\mathbb{A}\times V$，存在唯一的点 $Q\in\mathbb{A}$ 使得有向线段 $\overrightarrow{PQ}$ 的（平移）等价类是 v. 故

$$\mathbb{A}\times V \longrightarrow \mathbb{A},\quad (P, v)\longmapsto P\dotplus v := Q$$

定义了仿射空间 $\mathbb{A}_3$.

□

例7.1.2：设 V 是任意实向量空间. $U\subset V$ 是一个子空间. $\mathbb{A} = v_0 + U$（$v_0\in V$ 是固定向量）. 则映射

$$\mathbb{A}\times U \longrightarrow \mathbb{A},\quad (v_0+u, v)\longmapsto v_0+(u+v),$$

满足定义7.1.1中的条件(1)和(2). 故 v_0+U 是与 U 相伴的仿射空间. 特别地，V 也是与 V 相伴的仿射空间.

□

~~定义7.1.2. 设 $\mathbb{A}$, $\mathbb{A}'$ 分别是与 V, V' 相伴的仿射空间. 映射~~

~~$f: \mathbb{A} \longrightarrow \mathbb{A}'$~~

~~称为一个仿射映射，如果~~

定义7.1.2：设$\mathbb{A}$，$\mathbb{A}'$分别是与V，V'相伴的仿射空间。映射

$$f: \mathbb{A} \longrightarrow \mathbb{A}'$$

称为仿射映射，如果存在线性映射$f': V \to V'$使得

(1.1). $$f(P \dotplus v) = f(P) \dotplus f'(v)$$

对任意$(P, v) \in \mathbb{A} \times V$成立。

□.

定义7.1.2中的线性映射$f': V \to V'$由f唯一确定，称为仿射映射$f: \mathbb{A} \to \mathbb{A}'$的线性部分（或$f$的微分），此时

(1.2). $$f'(\overrightarrow{Pq}) = \overrightarrow{f(P)f(q)}, \quad \forall\ P, q \in \mathbb{A}.$$

引理7.1.2：设$f: \mathbb{A} \to \mathbb{A}'$是仿射映射。则$f$是双射当且仅当$f$的线性部分$f': V \to V'$是双射。此时$f^{-1}: \mathbb{A}' \to \mathbb{A}$也是仿射映射，且$(f^{-1})' = (f')^{-1}$。

证明：若f是双射，固定点$P \in \mathbb{A}$，则

$V = \{\overrightarrow{Pq} \mid \forall q \in \mathbb{A}\}$，$V' = \{\overrightarrow{f(P)q'} \mid \forall q' \in \mathbb{A}'\}$.

由公式$f'(\overrightarrow{Pq}) = \overrightarrow{f(P)f(q)}$可得$f': V \to V'$是双射。

（f'显然是满射，且$\overrightarrow{Pq_1} \neq \overrightarrow{Pq_2} \Rightarrow q_1 \neq q_2 \Rightarrow f(q_1) \neq f(q_2)$
$\Rightarrow \overrightarrow{f(P)f(q_1)} \neq \overrightarrow{f(P)f(q_2)}$）。令$f^{-1}: \mathbb{A}' \to \mathbb{A}$，则

$$f^{-1}(f(P) \dotplus f'(v)) = P \dotplus v, \quad \forall\ P \in \mathbb{A}, v \in V.$$

即 $f^{-1}(P' \dotplus v') = f^{-1}(P') \dotplus (f')^{-1}(v')$. $\forall\ P' \in \mathbb{A}', v' \in V'$.

所以f^{-1}是仿射映射，且$(f^{-1})' = (f')^{-1}$.

若$f': V \to V'$是双射，固定$O \in \mathbb{A}$，则

$$f'(\overrightarrow{OP}) = \overrightarrow{f(O)f(P)}, \quad \forall\ P \in \mathbb{A}.$$

故f是双射。

□.

定义7.1.3：若仿射映射 $f: \mathbb{A} \to \mathbb{A}'$ 是双射，则称 f 是仿射同构。此时称 $\mathbb{A}$ 与 $\mathbb{A}'$ 同构。 □

定理7.1.1：$\mathbb{A} = (\mathbb{A}, V)$ 与 $\mathbb{A}' = (\mathbb{A}', V')$ 同构当且仅当

$$\dim(V) = \dim(V').$$

证明：若 $\mathbb{A}$ 与 $\mathbb{A}'$ 同构，由引理7.1.2，$\dim(V) = \dim(V')$.

反之，若 $\dim(V) = \dim(V')$，任取向量空间同构 $F: V \to V'$，并固定点 $O \in \mathbb{A}$，$O' \in \mathbb{A}'$。由于 $\mathbb{A} = \{O \dotplus v \mid v \in V\}$（见注记(2)），定义 $f: \mathbb{A} \to \mathbb{A}'$，$f(O \dotplus v) = O' \dotplus F(v)$，则 f 是仿射映射（$\forall\, P = O \dotplus u \in \mathbb{A}$，$v \in V$，$f(P \dotplus v) = f(O \dotplus (u+v)) = O' \dotplus F(u+v)$ $= O' \dotplus (F(u) + F(v)) = (O' \dotplus F(u)) \dotplus F(v) = f(P) \dotplus F(v)$），且 $f' = F$. 由引理7.1.2，f 是同构. □

定义7.1.4：设 $\mathbb{A} = (\mathbb{A}, V)$ 是仿射空间，$O \in \mathbb{A}$，$e_1, \cdots, e_n$ 是 V 的一组基，则 $(O, e_1, \cdots, e_n)$ 称为 $\mathbb{A}$ 的一个仿射坐标系，（O 称为原点）。向量 $\overrightarrow{OP}$ 在 $e_1, \cdots, e_n$ 的坐标称为点 $P \in \mathbb{A}$ 在坐标系 $(O, e_1, \cdots, e_n)$ 下的坐标。 □

设 $(O, e_1, \cdots, e_n)$，$(O', e_1', \cdots, e_n')$ 是 $\mathbb{A} = (\mathbb{A}, V)$ 的两个坐标系，O' 在 $(O, e_1, \cdots, e_n)$ 下的坐标为 $(b_1, \cdots, b_n)$，且

$$(e_1', \cdots, e_n') = (e_1, \cdots, e_n) A, \quad A = (a_{ij})_{n \times n}.$$

定理7.1.2：设点 $P \in \mathbb{A}$ 在坐标 $(O, e_1, \cdots, e_n)$ 和 $(O', e_1', \cdots, e_n')$ 下的坐标分别是 $(x_1, \cdots, x_n)$ 和 $(y_1, \cdots, y_n)$，则

$$\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = A \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} + \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}.$$

□

证明：由定义，

$$\overrightarrow{OP}=(e_1\cdots,e_n)\begin{pmatrix}x_1\\ \vdots\\ x_n\end{pmatrix},\quad \overrightarrow{O'P}=(e_1',\cdots e_n')\begin{pmatrix}y_1\\ \vdots\\ y_n\end{pmatrix}=(e_1,\cdots,e_n)A\begin{pmatrix}y_1\\ \vdots\\ y_n\end{pmatrix},\ \overrightarrow{OO'}=(e_1\cdots e_n)\begin{pmatrix}b_1\\ \vdots\\ b_n\end{pmatrix}.$$

故

$$\overrightarrow{OP}=\overrightarrow{OO'}+\overrightarrow{O'P}=(e_1,\cdots,e_n)\left(A\begin{pmatrix}y_1\\ \vdots\\ y_n\end{pmatrix}+\begin{pmatrix}b_1\\ \vdots\\ b_n\end{pmatrix}\right).\ 即\ \begin{pmatrix}x_1\\ \vdots\\ x_n\end{pmatrix}=A\begin{pmatrix}y_1\\ \vdots\\ y_n\end{pmatrix}+\begin{pmatrix}b_1\\ \vdots\\ b_n\end{pmatrix}.$$

□.

定义7.1.5．设 $\mathbb{A}=(\mathbb{A},V)$ 是 n 维仿射空间，$U\subset V$ 是 m 维子空间，则

$$\pi=P\dotplus U=\{P\dotplus u\mid u\in U\}\subset\mathbb{A}$$

称为 $\mathbb{A}$ 的一个过点 $P\in\mathbb{A}$ 的 m 维平面（或仿射子空间）．当 $m=0$ 时，π 是一个点，当 $m=1$ 时，π 称为一条直线；当 $m=n-1$ 时，π 称为超平面．（U 通常称为平面 π 的方向子空间）．

□

引理7.1.5：$\pi\subset\mathbb{A}$ 是一条直线当且仅当存在不同的两点 $P,Q\in\mathbb{A}$ 使得 $\pi=\{P\dotplus\lambda\overrightarrow{PQ}\mid\lambda\in\mathbb{R}\}$．

证明：若 $\pi\subset\mathbb{A}$ 是一条直线，则存在点 $P\in\mathbb{A}$，1 维子空间 $U\subset V$ 使得 $\pi=P\dotplus U$．令 $U=\{\lambda u\mid\lambda\in\mathbb{R}\}$，$Q=P\dotplus u\in\pi$，则 $u=\overrightarrow{PQ}$，从而 $\pi=\{P\dotplus\lambda\overrightarrow{PQ}\mid\lambda\in\mathbb{R}\}$．

若 $\pi=\{P\dotplus\lambda\overrightarrow{PQ}\mid\lambda\in\mathbb{R}\}$，$P\neq Q$，则 $u=\overrightarrow{PQ}\neq0$．令 $U=\{\lambda u\mid\lambda\in\mathbb{R}\}$，则 $\dim U=1$．且 $\pi=P\dotplus\lambda\overrightarrow{PQ}$．由定义 π 是一条直线

□.

定理7.1.3．子集合 $\Omega\subset\mathbb{A}=(\mathbb{A},V)$ 是一个 $m>0$ 维的平面当且仅当通过 Ω 中任意两个点的直线必落在 Ω 中．

证明：若 $\Omega=P\dotplus U$，$\dim(U)>0$，$Q_1,Q_2\in\Omega$ 是两个不同的点．则过 Q_1,Q_2 的直线必可写如 $\overline{Q_1Q_2}:=\{Q_1\dotplus\lambda\overrightarrow{Q_1Q_2}\mid\lambda\in\mathbb{R}\}$．

设 $Q_1=P\dotplus u_1$，$Q_2=P\dotplus u_2$ $(u_1,u_2\in U)$．则 $\overrightarrow{Q_1Q_2}=u_2-u_1$，故

$$Q_1\dotplus\lambda\overrightarrow{Q_1Q_2}=(P\dotplus u_1)\dotplus\lambda(u_2-u_1)=P\dotplus(u_1+\lambda u_2-\lambda u_1)\in P\dotplus U.$$

反之，设 $P\in\Omega$，$U=\{\overrightarrow{Pq}\mid q\in\Omega\}\subset V$，则 $\Omega=P\dotplus U$。故只需证明 U 是子空间。设 $u_1=\overrightarrow{Pq_1}$，$u_2=\overrightarrow{Pq_2}\in U$，则过 q_1,q_2 的直线 $\overline{q_1q_2}=\{q_1\dotplus\lambda\overrightarrow{q_1q_2}\mid\lambda\in\mathbb{R}\}$ 落在 Ω 中。即

$$q_1\dotplus\lambda\overrightarrow{q_1q_2}=(P\dotplus u_1)\dotplus\lambda(u_2-u_1)=P\dotplus(u_1+\lambda u_2-\lambda u_1)\in\Omega.$$

换言之，$\forall\ u_1,u_2\in U$，$\lambda\in\mathbb{R}$，必有 $(1-\lambda)u_1+\lambda u_2\in U$。因 $P\in\Omega$，故 U 包含零向量。取 $u_1=0$ 可得 $\lambda u_2\in U$ 对任意 $u_2\in U$ 成立。当 $\lambda=\frac{1}{2}$ 时得 $\frac{1}{2}(u_1+u_2)\in U$ 对任意 $u_1,u_2\in U$ 成立。从而 $u_1+u_2\in U$。故 $U\subset V$ 是子空间。 □

<u>推论7.1.1</u>. 若 $\pi',\pi''\subset\mathbb{A}=(\mathbb{A},V)$ 是两个平面，则

$$\pi'\cap\pi''\subset\mathbb{A}$$

也是平面。若 π',π'' 对应的子空间分别是 $U',U''\subset V$，则 $\pi'\cap\pi''$ 对应的子空间是 $U'\cap U''$。

证明：若 $\pi'\cap\pi''$ 仅含一个点，结论显然成立（它对应零子空间）。

若 $q_1,q_2\in\pi'\cap\pi''$ 是不同的点，则 $\overline{q_1q_2}\subset\pi'$，$\overline{q_1q_2}\subset\pi''$，从而 $\overline{q_1q_2}\subset\pi'\cap\pi''$。故 $\pi'\cap\pi''\subset\mathbb{A}$ 是平面。亦可直接由定义证明：设 $P\in\pi'\cap\pi''$，则 $\pi'=P\dotplus U'$，$\pi''=P\dotplus U''$。不难证得

$$\pi'\cap\pi''=P\dotplus U'\cap U''.$$

□.

<u>定义7.1.6</u>：设 $\mathbb{A}=(\mathbb{A},V)$ 是仿射空间。映射

$$f:\ \mathbb{A}\longrightarrow\mathbb{R}$$

称为仿射(线性)函数，如果存在 $f'\in V^*$ 使

$$f(P\dotplus v)=f(P)+f'(v),\quad\forall\ P\in\mathbb{A},\ v\in V.$$

□

<u>引理7.1.6</u>：设 $(O,e_1,\cdots,e_n)$ 是 $\mathbb{A}=(\mathbb{A},V)$ 的坐标，$\overrightarrow{OP}=x_1e_1+\cdots+x_ne_n$。则映射 $f:\mathbb{A}\to\mathbb{R}$ 是仿射函数的充要条件为存在常数 $a_1,\cdots,a_n\in\mathbb{R}$，$c\in\mathbb{R}$ 使得

$$f(P)=a_1x_1+\cdots+a_nx_n+c.$$

证明：若 $f: \mathbb{A} \to \mathbb{R}$ 是仿射函数，则存在 $f' \in V^*$ 使得

$$f(P \dotplus v) = f(P) + f'(v), \quad \forall P \in \mathbb{A}, v \in V.$$

令 $a_i = f'(e_i)$，$c = f(O)$。则 $f(P) = f(O \dotplus \overrightarrow{OP}) = f(O) + f'(\overrightarrow{OP})$。即

$$f(P) = a_1 x_1(P) + \cdots + a_n x_n(P) + c.$$

其中 $(x_1(P), \cdots, x_n(P))$ 是点 P 在 $(O, e_1, \cdots, e_n)$ 下的坐标。

反之，若函数 $f: \mathbb{A} \to \mathbb{R}$ 满足 $f(P) = a_1 x_1(P) + \cdots + a_n x_n(P) + c$，令 $f': V \to \mathbb{R}$ 是由 $f'(e_i) := a_i$ 确定的线性函数。则对任意 $v = v_1 e_1 + \cdots + v_n e_n \in V$，及 $P \in \mathbb{A}$，有 $\overrightarrow{O(P \dotplus v)} = \overrightarrow{OP} + v$，故

$$\begin{aligned} f(P \dotplus v) &= a_1 x_1(P \dotplus v) + \cdots + a_n x_n(P \dotplus v) + c \\ &= a_1 x_1(P) + \cdots + a_n x_n(P) + c + a_1 v_1 + \cdots + a_n v_n \\ &= f(P) + f'(v). \end{aligned}$$

所以 f 是仿射函数。 $\square$

定理 7.1.4：设 $(O, e_1, \cdots, e_n)$ 是仿射空间 $\mathbb{A} = (\mathbb{A}, V)$ 的一个坐标系，$\Pi \subset \mathbb{A}$ 是非空子集。则 Π 是 $(n-r)$ 维平面的充要条件是

$$\Pi = \left\{ P \in \mathbb{A} \;\middle|\; \begin{matrix} a_{11} x_1(P) + \cdots + a_{1n} x_n(P) = b_1 \\ \vdots \\ a_{m1} x_1(P) + \cdots + a_{mn} x_n(P) = b_m \end{matrix} \right\}$$

其中 $(x_1(P), \cdots, x_n(P))$ 是点 P 在 $(O, e_1, \cdots, e_n)$ 下的坐标，

$$(7.3) \quad \begin{cases} a_{11} x_1 + \cdots + a_{1n} x_n = b_1 \\ \vdots \\ a_{m1} x_1 + \cdots + a_{mn} x_n = b_m \end{cases}$$

是秩为 r 的相容方程组。

证明：若 $\Pi = P_0 \dotplus U$，$\dim(U) = n-r$，则存在秩为 r 的齐次方程组

$$(*) \quad \begin{cases} a_{11} x_1 + \cdots + a_{1n} x_n = 0 \\ \vdots \\ a_{m1} x_1 + \cdots + a_{mn} x_n = 0 \end{cases}$$

使得 $U \subset V$ 恰由坐标满足 (*) 的向量组成。而 $P \in P_0 + U$ 的充要条件是 $\overrightarrow{P_0P} \in U$. 即 $\overrightarrow{OP} - \overrightarrow{OP_0}$ 的坐标 $(x_1(P)-x_1(P_0), \cdots, x_n(P)-x_n(P_0))$ 满足齐次线性方程组 (*). 令 $b_i = a_{i1}x_1(P_0) + \cdots + a_{in}x_n(P_0) \in \mathbb{R}$. 则 $P \in \pi = P_0 + U$ 的充要条件是 P 的坐标 $(x_1(P), \cdots, x_n(P))$ 满足

$$\begin{cases} a_{11}x_1(P) + \cdots + a_{1n}x_n(P) = b_1 \\ \vdots \\ a_{m1}x_1(P) + \cdots + a_{mn}x_n(P) = b_m \end{cases}$$

反之，若子集合 $\Omega \subset \mathbb{A}$ 由坐标 $(x_1(P), \cdots, x_n(P))$ 满足

$$(**) \quad \begin{cases} a_{11}x_1(P) + \cdots + a_{1n}x_n(P) = b_1 \\ \vdots \\ a_{m1}x_1(P) + \cdots + a_{mn}x_n(P) = b_m \end{cases}$$

的点组成（其中 (**) 有解）. 令

$$U = \{ v = v_1e_1 + \cdots + v_ne_n \mid (v_1, \cdots, v_n) 满足 (**)^{齐} \}.$$

$P_0 \in \Omega$. 则不难验证 $\Omega = P_0 + U$. □

<u>定义7.1.7</u> 设 $\pi' = P \dotplus U'$, $\pi'' = Q \dotplus U''$ 是 $\mathbb{A} = (\mathbb{A}, V)$ 中的两个平面，且 $\dim U' \le \dim U''$. 如果 $U' \subseteq U''$, 则称 π' 平行于 π''. □

<u>定理7.1.5</u>. 对任意平面 $\pi \subset \mathbb{A}$ 和任意点 $Q \in \mathbb{A}$, 存在唯一通过 Q 并平行于 π 的平面 $\pi' \subset \mathbb{A}$ 且 $\dim \pi' = \dim \pi$. 如果 $Q \in \pi$, 则 $\pi' = \pi$. 如果 $Q \notin \pi$, 则 π 与 π' 不相交.

证明：设 $\pi = P_0 \dotplus U$, $Q \in \mathbb{A}$. 则 $\pi' = Q \dotplus U$ 是平行于 π 的平面。若 $\pi' = Q \dotplus U'$ 是平行且 $\dim \pi' = \dim \pi$. 则 $U' \subseteq U$. 但 $\dim \pi' = \dim \pi$, 故 $U' = U$. 若 $Q \in \pi$, 显然 $\pi' = \pi$. 若 $Q \notin \pi$, 则 π 与 π' 不相交。否则存在 $u_1, u_2 \in U$ 使

$$P \dotplus u_1 = Q \dotplus u_2$$

从而 $Q = (Q \dotplus u_2) \dotplus (-u_2) = (P \dotplus u_1) \dotplus (-u_2) = P \dotplus (u_1 - u_2) \in \pi$. 矛盾 □

例7.1.3：设 π,π' 是 $\mathbb{A}$ 中分别由方程

$$a_1x_1+\cdots+a_nx_n=b \quad 和 \quad a_1'x_1+\cdots+a_n'x_n=b'$$

定义的超平面，则 π 与 π' 平行的充要条件是

$$a_1'=\lambda a_1,\ \ldots,\ a_n'=\lambda a_n.$$

而 π 与 π' 重合的充要条件是 $a_1'=\lambda a_1,\ldots,a_n'=\lambda a_n,\ b'=\lambda b$. □

习题 7.1

1. 设 $\pi=P\dotplus U$ 是 $\mathbb{A}=(A,V)$ 的平面（亦称仿射子空间），证明 π 是与 U 相伴的仿射空间.

2. 试证明：平面 $\pi'=P\dotplus U'$ 与平面 $\pi''=Q\dotplus U''$ 相交的充要条件是 $\overrightarrow{PQ}\in U'+U''$.

3. 设 $(P_0,P_1,\ldots,P_n)$ 是仿射空间 $\mathbb{A}=(A,V)$ 的 $n+1$ 个点使得 $e_1:=\overrightarrow{P_0P_1},\ldots,e_n=\overrightarrow{P_0P_n}$ 是 V 的一组基. 试证明：对仿射空间 $\mathbb{A}'=(A',V')$ 中任意 $n+1$ 个点 $(P_0',P_1',\ldots,P_n')$，存在唯一的仿射映射 $f:\mathbb{A}\to\mathbb{A}'$ 使得 $f(P_i)=P_i'$ $(i=0,1,\ldots,n)$.

4. 设 V 是 n 维实向量空间，$\mathbb{A}=V$，若定义

$$\mathbb{A}\times V\to\mathbb{A},\quad (P,v)\longmapsto P+v$$

其中 $P+v$ 是 V 中的加法. 试验证 V 是与 V 相伴的仿射空间（为了与 V 区别，记为 $\mathbb{A}(V):=(V,V)$）.

由于任意 n 维仿射空间均同构于 $\mathbb{A}(V)$，下面的习题仅针对 $\mathbb{A}(V)$.

5. 设 $X \subset \mathbb{A}(V)$ 的非空子集. 证明下列条件等价

(1) $\forall\ P, Q \in X$, 则 $\lambda P + (1-\lambda) Q \in X \quad (\forall \lambda \in \mathbb{R})$.

(2) $\forall\ P_1, \cdots, P_m \in X$, 若 $\lambda_1 + \cdots + \lambda_m = 1$, 则 $\lambda_1 P_1 + \cdots + \lambda_m P_m \in X$

(3) X 是 $\mathbb{A}(V)$ 中的平面(仿射子空间).

6. 设 $M \subset \mathbb{A}(V)$ 的非空间子集, $\mathbb{A}(M) \subset \mathbb{A}(V)$ 表示包含 M 的最小平面(或仿射子空间). 证明.

$$\mathbb{A}(M) = \left\{ \sum_{i=1}^{m} \lambda_i x_i \;\middle|\; m \geq 1,\ x_1, \cdots, x_m \in M,\ \sum_{i=1}^{m} \lambda_i = 1 \right\}.$$

7. 设 $f: \mathbb{A}(V) \longrightarrow \mathbb{A}(V')$ 是一个映射. 试证明:

$$f \text{ 是仿射映射} \Longleftrightarrow f(\lambda P + (1-\lambda) Q) = \lambda f(P) + (1-\lambda) f(Q).$$

8. 设 $f: \mathbb{A}(V) \longrightarrow \mathbb{A}(V')$ 是一个映射. 试证明: f 是仿射映射的充要条件是存在 $v' \in V'$ 及线性映射 $A: V \to V'$ 使得 $f = T_{v'} \cdot A$, 其中 $T_{v'}: \mathbb{A}(V') \to \mathbb{A}(V')$ 定义如下:

$$T_{v'}(P') = P' + v', \quad \forall P' \in \mathbb{A}(V').$$

9. 设 V 是一个实向量空间, $\mathbb{A}$ 是一个非空集合. 试证明: $\mathbb{A}$ 是与 V 相伴的充分必要条件是. 存在映射

$$\mathbb{A} \times \mathbb{A} \longrightarrow V, \quad (P, Q) \longmapsto \overrightarrow{PQ},$$

满足: (1) $\forall\ P \in \mathbb{A}$, $v \in V$, 存在唯一的 $Q \in \mathbb{A}$ 使得

$$\overrightarrow{PQ} = v.$$

(2) $\forall\ P, Q, R \in \mathbb{A}$, 有 $\overrightarrow{PQ} + \overrightarrow{QR} = \overrightarrow{PR}$.

§7.2. 欧几里得空间

定义7.2.1：设 V 是欧氏向量空间，则称与 V 相伴的仿射空间 $\mathbb{A}=(\mathbb{A},V)$ 为欧几里得空间。对任意两点 $P,Q\in\mathbb{A}=(\mathbb{A},V)$，$\rho(P,Q):=\|\overrightarrow{PQ}\|=\sqrt{\langle\overrightarrow{PQ}\mid\overrightarrow{PQ}\rangle}$，称为 P 与 Q 之间的距离。 $\square$

由内积的性质，我们不难证明距离的基本性质：

(1) $\rho(P,Q)=\rho(Q,P)$，　(2) $\rho(P,Q)=0 \Leftrightarrow P=Q$.

(3)（三角不等式）. $\rho(P,Q)+\rho(Q,R)\geq\rho(P,R)$.

定义7.2.2：设 $l_{P,Q}$ 是表示由点 $P,Q\in\mathbb{A}$ 确定的直线，则直线 $l_{P,Q}$ 与直线 $l_{R,S}$ 之间夹角 φ 由下式定义：

$$\cos\varphi=\frac{\langle\overrightarrow{PQ}\mid\overrightarrow{RS}\rangle}{\|\overrightarrow{PQ}\|\cdot\|\overrightarrow{RS}\|}.$$

$\square$

如果 $e_1,\cdots,e_n\in V$ 是一组标准正交基，则 $(O,e_1,\cdots,e_n)$ 称为 $\mathbb{A}=(\mathbb{A},V)$ 的一个直角坐标系。若 P,Q 在直角坐标系 $(O,e_1,\cdots,e_n)$ 的坐标分别是 $(x_1,\cdots,x_n)$，$(y_1,\cdots,y_n)$，则

$$\rho(P,Q)=\sqrt{(y_1-x_1)^2+\cdots+(y_n-x_n)^2}.$$

定理7.2.1. 设 V,V' 是欧氏向量空间，$\mathbb{A}=(\mathbb{A},V)$，$\mathbb{A}'=(\mathbb{A}',V')$ 分别是与 V,V' 相伴的欧几里得空间。若 $\dim(V)=\dim(V')$，则存在仿射同构 $f:\mathbb{A}\to\mathbb{A}'$ 使

$$\rho(P,Q)=\rho(f(P),f(Q)),\ \forall P,Q\in\mathbb{A}.$$

证明. 设 $(O,e_1,\cdots,e_n)$，$(O',e_1',\cdots,e_n')$ 分别是 $\mathbb{A},\mathbb{A}'$ 的直角坐标系，

则存在线性映射 $F: V \to V'$ 使 $F(e_i)=e_i'$ $(1\le i\le n)$.
由定理7.1.1的证明可知: 存在仿射同构 $f: \mathbb{A}\to\mathbb{A}'$ 使得

$$f(O)=O', \quad f'=F.$$

且 $\forall\ p,q\in\mathbb{A}$, 有 $f'(\overrightarrow{pq})=\overrightarrow{f(p)f(q)}$. 即 $F(\overrightarrow{pq})=\overrightarrow{f(p)f(q)}$.
由F的选取可得 $\|F(\overrightarrow{pq})\|=\|\overrightarrow{pq}\|$. 故

$$\rho(f(p),f(q))=\|\overrightarrow{f(p)f(q)}\|=\|\overrightarrow{pq}\|=\rho(p,q).$$

$\square$

<u>定义7.2.3</u>. $\overline{pq}=\{p\dotplus\lambda\overrightarrow{pq}\mid 0\le\lambda\le 1\}\subset\mathbb{A}$ 称为连接 $p,q\in\mathbb{A}$ 两点的<u>线段</u>. 称 $|\overline{pq}|:=\|\overrightarrow{pq}\|=\rho(p,q)$ 称为线段 $\overline{pq}$ 的长度.

$\square$

<u>定义7.2.4</u>(点到平面的距离). 设 $\Pi\subset\mathbb{A}=(\mathbb{A},V)$ 是欧几里得空间的m维平面. $p\in\mathbb{A}\setminus\Pi$, $q\in\Pi$. 若

$$\langle\overrightarrow{pq}\mid\overrightarrow{rs}\rangle=0 \qquad (\forall\ r,s\in\Pi)$$

则称直线 $L_{p,q}$(通过 p,q 的直线)垂直于 Π(记为 $L_{p,q}\perp\Pi$).
此时称 $\rho(p,q)$ 是点p到 Π 的距离, 线段 $\overline{pq}$ 称为点p到 Π 的垂直线段(记为 $\overline{pq}\perp\Pi$).

$\square$

<u>引理7.2.1</u>. 设 $\Pi=O\dotplus U\subset\mathbb{A}$ 是n维欧几里得空间 $\mathbb{A}=(\mathbb{A},V)$ 中的m维平面, $p\in\mathbb{A}\setminus\Pi$. 则存在唯一的 $q\in\Pi$ 使

$$\langle\overrightarrow{pq}\mid\overrightarrow{rs}\rangle=0 \qquad (\forall\ r,s\in\Pi).$$

证明: 设 $V=U\oplus U^\perp$, $\overrightarrow{pO}=u+v$, $(u\in U, v\in U^\perp)$
令 $q=O\dotplus(-u)$, 则 $\overrightarrow{pq}=\overrightarrow{pO}+\overrightarrow{Oq}=u+v+(-u)=v\in U^\perp$
故 $\langle\overrightarrow{pq}\mid\overrightarrow{rs}\rangle=0$, $\forall\ r,s\in\Pi$.
即 $\overline{pq}$ 是点p到平面 Π 的垂直线.

若在 $\mathbb{A}$ 中选取坐标系 $(O, e_1, \cdots, e_n)$ 使得 $e_1, \cdots, e_m$ 是 U 的一组基，($e_{m+1}, \cdots, e_n \in U^\perp$) 令 $\overrightarrow{Oq} = x_1 e_1 + \cdots + x_n e_n$，则 $\overrightarrow{pq} \perp U$ 等价于

$$0 = \langle \overrightarrow{pq} \mid e_i \rangle = \langle \overrightarrow{Oq} \mid e_i \rangle - \langle \overrightarrow{Op} \mid e_i \rangle \qquad (1 \le i \le m)$$

即点 q 的坐标 $(x_1, \cdots, x_n)$ 满足方程组

$$\begin{pmatrix} \langle e_1|e_1\rangle, & \cdots, & \langle e_1|e_m\rangle \\ \vdots & & \vdots \\ \langle e_m|e_1\rangle, & \cdots & \langle e_m|e_m\rangle \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_m \end{pmatrix} = \begin{pmatrix} \langle \overrightarrow{Op}|e_1\rangle \\ \vdots \\ \langle \overrightarrow{Op}|e_m\rangle \end{pmatrix} \qquad (2.1)$$

它有唯一解，故点 q 由 p 和 $\Pi = O + U$ 唯一确定。 □

引理 7.2.2. 设 $\Pi = p \dotplus U$，$\Pi' = p' \dotplus U'$ 是 $\mathbb{A} = (\mathbb{A}, V)$ 的两个平面，则存在一条直线同时垂直于 Π 和 Π'。(称为 Π 与 Π' 的公垂线)

证明：~~由于 p' 可以用 Π' 中任一点代替，由引理 7.2.1 可设 $\overrightarrow{pp'} \perp \Pi'$。即~~ 需要找到 $q = p \dotplus u$，$q' = p' \dotplus u'$ 使得 $\overrightarrow{qq'} \perp \Pi$，$\overrightarrow{qq'} \perp \Pi'$。即 $\overrightarrow{qq'} \perp (U + U')$。但 $\overrightarrow{qq'} = -u + u' + \overrightarrow{pp'}$

$V = (U+U') \oplus (U+U')^\perp$，令 $\overrightarrow{pp'} = b + c$ $(b \in U+U', c \in (U+U')^\perp)$，$b = v + v'$ $(v \in U, v' \in U')$。则 $\overrightarrow{qq'} = (-u+v) + (u'+v') + c$。由于 $c \in (U+U')^\perp$，所以只需选取 $u = v$，$u' = -v'$ 即得 $\overrightarrow{qq'} \in (U+U')^\perp$。 □

定理 7.2.2. 设 $q \in \Pi$，$q' \in \Pi'$ 使得 $\overrightarrow{qq'}$ 是 Π 与 Π' 的公垂线。则
$$\rho(q, q') = \min\{\rho(x, x') \mid x \in \Pi, x' \in \Pi'\}.$$
且公垂线唯一当且仅当 Π 与 Π' 的方向子空间 U 与 U' 满足 $U \cap U' = \{0\}$。

证明：不妨设 $\Pi = q \dotplus U$，$\Pi' = q' \dotplus U'$，$x = q \dotplus u$，$x' = q' \dotplus u'$

则 $\overrightarrow{xx'} = \overrightarrow{qq'} + u'-u$. 故由 $\overrightarrow{qq'} \perp (U+U')$ 可得

$$\|\overrightarrow{xx'}\|^2 = \|\overrightarrow{qq'}\|^2 + \|u'-u\|^2 \geq \|\overrightarrow{qq'}\|.$$

即 $\rho(x,x') \geq \rho(q,q')$.

若 $p = q+v$, $p' = q'+v'$ 使得 $\overline{pp'}$ 也是 π 与 π' 的公垂线,

则 $\rho(p,p') = \rho(q,q')$. 即 $v = v'$. 所以, 如果 $U \cap U' = \{0\}$, 则 $p = q$, $p' = q'$. 若 π 与 π' 的公垂线唯一, 则 $U \cap U' = \{0\}$. 否则, 取 $v \in U \cap U'$ 是非零向量, 令 $p = q+v$, $p' = q'+v$. 则 $\overline{pp'}$ 与 $\overline{qq'}$ 是不同的线段, 且 $\overline{pp'} \perp \pi$, $\overline{pp'} \perp \pi'$. □.

~~[illegible]~~ (平面之间的距离)

定义7.2.5: 设 $\pi, \pi' \subset \mathbb{A}$ 是两个平面, 它们公垂线的长度称为 π 与 π' 之间的距离. □.

习题7.2

1. 求由点 $p = (2,1,-3,4)$ 到平面

$$\pi: \quad 2x_1 - 4x_2 - 8x_3 + 13x_4 + 19 = 0, \quad x_1 + x_2 - x_3 + 2x_4 - 1 = 0$$

的距离.

2. 求平面 π_1: $x_1 + x_3 + x_4 - 2x_5 - 2 = 0$, $x_2 + x_3 - x_4 - x_5 - 3 = 0$, $x_1 - x_2 + 2x_3 - x_5 - 3 = 0$ 和平面

$$\pi_2 = (1,-2,5,8,2) + \langle (0,1,2,1,2), (2,1,2,-1,1) \rangle$$

之间的距离.

§7.3. 群与几何.

设 $\mathbb{A}=(A,V)$ 是一个 n 维仿射空间，若 $\mathbb{A}\xrightarrow{f}\mathbb{A}$ 和 $\mathbb{A}\xrightarrow{g}\mathbb{A}$ 是仿射同构，则 $g\cdot f:\mathbb{A}\to\mathbb{A}$ 也是仿射同构. 故映射的合成诱导了集合

$$\mathrm{Aff}(\mathbb{A})=\{\mathbb{A}\xrightarrow{f}\mathbb{A}\mid f\text{是仿射同构}\}.$$

上的一个运算. 不难验证，$\mathrm{Aff}(\mathbb{A})$ 关于该运算是一个群。

定义7.3.1. $\mathrm{Aff}(\mathbb{A})$ 称为 $\mathbb{A}$ 上的 n 维仿射群. □.

例7.3.1. 对任意 $v\in V$, 沿方向 v 的平移映射

$$T_v:\mathbb{A}\to\mathbb{A},\quad P\longmapsto P+v$$

是仿射同构(它的线性部分是恒等映射). 故

$$T(\mathbb{A})=\{T_v\mid v\in V\}\subset\mathrm{Aff}(\mathbb{A})$$

是一个子群，称为 $\mathrm{Aff}(\mathbb{A})$ 的平移子群. □

引理7.3.1. 对任意 $f\in\mathrm{Aff}(\mathbb{A})$, 令 $D(f)$ 表示 f 的线性部分，则映射 $D:\mathrm{Aff}(\mathbb{A})\longrightarrow GL(V)$ 是一个群同态. 且 $T(\mathbb{A})=\mathrm{Ker}(D):=\{f\in\mathrm{Aff}(\mathbb{A})\mid D(f)=\mathrm{id}.\}$.

证明: 对任意线性映射 $F:V\to V$, 由定理7.1.1的证明，可构造仿射映射 $f:\mathbb{A}\to\mathbb{A}$ 使得 $D(f)=F$, 且 f 是同构当且仅当 $F:V\to V$ 是线性同构. 故 D 是满射. 不难验证 $D(f\cdot g)=D(f)\cdot D(g)$, $\forall f,g\in\mathrm{Aff}(\mathbb{A})$, 故 D 是同态.

若 $D(f)=\mathrm{id}$, 则 $f(P+v)=f(P)+v$, $\forall P\in A, v\in V$.

)

故 $\overrightarrow{(p+v)f(p+v)} = \overrightarrow{pf(p)}$，即 $\overrightarrow{qf(q)} = \overrightarrow{pf(p)}$（因为 $v \mapsto p+v$ 诱导了双射 $V \to \mathbb{A}$）. 即 $v := \overrightarrow{pf(p)}$ 与 $p \in \mathbb{A}$ 无关. 且

$$f(p) = p + v, \quad \forall p \in \mathbb{A}.$$

所以 $f \in T(\mathbb{A})$，即 $\ker(D) \subset T(\mathbb{A})$. 而 $T(\mathbb{A}) \subset \ker(D)$ 是显然的，于是 $T(\mathbb{A}) = \ker(D)$. □

固定坐标系 $(O, e_1, \cdots, e_n)$，对 $f \in \mathrm{Aff}(\mathbb{A})$，令 $D(f)$ 在 $e_1, \cdots, e_n$ 下的坐标系矩阵为 $A = (a_{ij})_{n\times n}$，即

$$(D(f)e_1, \cdots, D(f)e_n) = (e_1, \cdots, e_n)A.$$

若 P，$f(P)$，$f(O)$ 在 $(O, e_1, \cdots, e_n)$ 下的坐标分别是 $(x_1, \cdots, x_n)$，$(y_1, \cdots, y_n)$，$(b_1, \cdots, b_n)$。则

$$(3.1).\qquad \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} = A \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} + \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix}.$$

__定义7.3.2__. 设 $\mathbb{A} = (\mathbb{A}, V)$ 是一个 n 维欧氏仿射空间，映射 $f: \mathbb{A} \to \mathbb{A}$ 是一个映射，若

$$\rho(f(p), f(q)) = \rho(p, q), \quad \forall p, q \in \mathbb{A}.$$

则称 f 是一个运动（或称保距映射）. □

__定理7.3.1__. 映射 $f: \mathbb{A} \to \mathbb{A}$ 是一个运动（保距映射）的充分必要条件是：f 是仿射映射且 f 的线性部分 $D(f): V \to V$ 是正交算子.

证明："$\Leftarrow$". 若 $f: \mathbb{A} \to \mathbb{A}$ 是仿射映射，且 $D(f): V \to V$ 是正交算子，则

$$\rho(f(p), f(q)) = \rho(f(p), f(p+v)) = \|\overrightarrow{f(p)f(p+v)}\| = \|D(f)(v)\| = \|v\| = \|\overrightarrow{pq}\| = \rho(p, q).$$

若 $f: \mathbb{A}\to\mathbb{A}$ 是一个运动，下面证明，f 是仿射映射且 $D(f): V\to V$ 是正交算子。我们分两步证明：

(1) 若存在 $P_0\in\mathbb{A}$，使 $f(P_0)=P_0$，则 f 是仿射映射，且 $D(f):V\to V$ 是正交算子。

(2)（一般情形）。固定一点 $P_0\in\mathbb{A}$，令 $Q_0=f(P_0)$，$v=\overrightarrow{P_0Q_0}\in V$，$g=T_{-v}\circ f:\mathbb{A}\to\mathbb{A}$，其中 $T_{-v}:\mathbb{A}\to\mathbb{A}$ 是沿 $-v$ 方向的平移。则 $g:\mathbb{A}\to\mathbb{A}$ 是一个运动，且 $g(P_0)=P_0$。由上述结论(1)可得 $g:\mathbb{A}\to\mathbb{A}$ 是仿射映射且 $D(g):V\to V$ 是正交算子。故 $f=T_v\circ g$ 是仿射映射，且 $D(f)=D(T_v)\cdot D(g)=D(g)$ 是正交算子。

所以定理的证明归结为下面的引理。 □

引理7.3.2：设 $f:\mathbb{A}\to\mathbb{A}$ 是一运动且存在 $P_0\in\mathbb{A}$ 使 $f(P_0)=P_0$，则 f 是仿射映射且 $D(f):V\to V$ 是正交算子。

证明：定义映射 $F:V\to V$ 如下：$F(v)=\overrightarrow{P_0 f(P_0\dotplus v)}\in V$，则

~~则 $f(P_0\dotplus v)=$~~

$$f(P_0\dotplus v)=P_0\dotplus F(v). \tag{3.2}$$

下面我们证明：$F:V\to V$ 是正交算子。事实上，我们有

$$F(0)=0,\quad \|F(u)-F(v)\|=\|u-v\|. \tag{3.3}$$

$F(0)=0$ 是显然的（因为 $F(0)=\overrightarrow{P_0f(P_0)}=0$）。由(3.2)，可得

$$\begin{aligned}\|F(u)-F(v)\| &= \|\overrightarrow{P_0f(P_0\dotplus u)}-\overrightarrow{P_0f(P_0\dotplus v)}\| = \|\overrightarrow{f(P_0\dotplus v)f(P_0\dotplus u)}\| \\ &= \rho(f(P_0\dotplus v),f(P_0\dotplus u))=\rho(P_0\dotplus v,P_0\dotplus u) \\ &= \|\overrightarrow{(P_0\dotplus v)(P_0\dotplus u)}\|=\|v-u\|.\end{aligned}$$

由(3.3)，$\|u\|^2-2\langle u|v\rangle+\|v\|^2=\|u-v\|=\|F(u)-F(v)\|$

$$=\|F(u)\|^2-2\langle F(u)|F(v)\rangle+\|F(v)\|^2=\|u\|^2-2\langle F(u)|F(v)\rangle+\|v\|^2$$

故 $\langle F(u)|F(v)\rangle=\langle u|v\rangle$，即 $F:V\to V$ 保持内积不变.

下面证明 $F:V\to V$ 是线性映射. 事实上，

$$\|F(u+v)-F(u)-F(v)\|^2=\|F(u+v)\|^2-2\langle F(u+v)|F(u)+F(v)\rangle+\|F(u)+F(v)\|^2$$

$$=\|F(u+v)\|^2-2\langle F(u+v)|F(u)\rangle-2\langle F(u+v)|F(v)\rangle+\|F(u)\|^2+2\langle F(u)|F(v)\rangle+\|F(v)\|^2$$

$$=\|u+v\|^2-2\langle u+v|u\rangle-2\langle u+v|v\rangle+\|u\|^2+2\langle u|v\rangle+\|v\|^2=0$$

故 $F(u+v)=F(u)+F(v)$. 同理

$$\|F(\lambda u)-\lambda F(u)\|^2=\|F(\lambda u)\|^2-2\langle F(\lambda u)|\lambda F(u)\rangle+\|\lambda F(u)\|^2$$

$$=\|\lambda u\|^2-2\langle \lambda u|\lambda u\rangle+\lambda^2\|u\|^2=0.$$

故 $F(\lambda u)=\lambda F(u)$. 从而 $F:V\to V$ 是一个正交算子，且

$$f(P\dotplus v)=f(P_0\dotplus(\overrightarrow{P_0P}+v))=P_0\dotplus F(\overrightarrow{P_0P}+v)$$

$$=(P_0\dotplus F(\overrightarrow{P_0P}))\dotplus F(v)$$

$$=f(P_0\dotplus\overrightarrow{P_0P})\dotplus F(v)=f(P)\dotplus F(v).$$

故 f 是仿射映射且 $D(f)=F$. □

定义 7.3.3 子群 $\mathrm{Iso}(\mathbb{A})=\{f\in\mathrm{Aff}(\mathbb{A})\mid D(f)\text{ 是正交算子}\}\subset\mathrm{Aff}(\mathbb{A})$

称为保距离群(运动群). □

根据克莱因在其"爱尔兰根纲领"中的观点，仿射几何学的目标就是研究 $\mathbb{A}=(\mathbb{A},V)$ 中在仿射变换群 $\mathrm{Aff}(\mathbb{A})$ 的作用之下不变的空间图形的性质. 而欧氏几何学的目标则是研究 $\mathbb{A}=(\mathbb{A},V)$ 中在 $\mathrm{Iso}(\mathbb{A})$ 作用下不变的空间图形性质.

为下面讨论方便，我们在 $\mathbb{A}=(\mathbb{A},V)$ 中引入下面的记号。对任意 $P_0,P_1,\cdots,P_m\in\mathbb{A}$ 及满足 $\lambda_0+\lambda_1+\cdots+\lambda_m=1$ 的实数 $\lambda_0,\lambda_1,\cdots,\lambda_m$，定义

$$\lambda_0P_0+\lambda_1P_1+\cdots+\lambda_mP_m := P\dotplus\sum_{i=0}^{m}\lambda_i\overrightarrow{PP_i}.$$

其中 $P\in\mathbb{A}$ 是任意的点。

引理7.3.3. 表达式 $P\dotplus(\lambda_0\overrightarrow{PP_0}+\lambda_1\overrightarrow{PP_1}+\cdots+\lambda_m\overrightarrow{PP_m})\in\mathbb{A}$ 与 P 的选取无关。

证明：若 $q=P\dotplus v\in\mathbb{A}$，则 $\overrightarrow{(P\dotplus v)P_i}=\overrightarrow{PP_i}-v$. 故

$$q\dotplus\sum_{i=0}^{m}\lambda_i\overrightarrow{qP_i}=(P\dotplus v)\dotplus\sum_{i=0}^{m}\lambda_i\overrightarrow{(P\dotplus v)P_i}=(P\dotplus v)\dotplus\sum_{i=0}^{m}\lambda_i(\overrightarrow{PP_i}-v)$$
$$=P\dotplus\Big(v-\sum_{i=1}^{m}\lambda_iv+\sum_{i=0}^{m}\lambda_i\overrightarrow{PP_i}\Big)=P\dotplus\sum_{i=0}^{m}\overrightarrow{PP_i}.$$

□.

在下面的讨论中，我们总是假设给定的 $m+1$ 个点满足条件：其中 $\overrightarrow{P_0P_1},\cdots,\overrightarrow{P_0P_m}$ 线性无关。此时表达式 $q=\lambda_0P_0+\cdots+\lambda_mP_m$ 中满足条件 $\lambda_0+\cdots+\lambda_m=1$ 的系数 $\lambda_0,\cdots,\lambda_m$ 由 q 唯一确定。

定义7.3.4：子集合 $\Delta(P_0,\cdots,P_m)=\{\lambda_0P_0+\cdots+\lambda_mP_m \mid \sum_{i=0}^{m}\lambda_i=1,\ \lambda_i>0\}$ 称为 $\mathbb{A}$ 中以 $P_0,\cdots,P_m$ 为顶点的 m维开单纯形。而

$$\overline{\Delta}(P_0,\cdots,P_m)=\{\lambda_0P_0+\cdots+\lambda_mP_m \mid \sum_{i=0}^{m}\lambda_i=1,\ \lambda_i\geq 0\}.$$

称为以 $P_0,P_1,\cdots,P_m$ 为顶点的闭单纯形。 □

例7.3.2：当 $m=1$ 时，$\overline{\Delta}(P_0,P_1)=\{\lambda_0P_0+\lambda_1P_1 \mid \lambda_0+\lambda_1=1,\ \lambda_0,\lambda_1\geq 0\}=\overline{P_0P_1}$.
当 $m=2$ 时，$\Delta(P_0,P_1,P_2)=\{\lambda_0P_0+\lambda_1P_1+\lambda_2P_2 \mid \lambda_i>0,\ \lambda_0+\lambda_1+\lambda_2=1\}$ 是以 P_0,P_1,P_2 的开三角形。 □

定理 7.3.2：对任意 $f \in \mathrm{Aff}(\mathbb{A})$，单纯形的像仍是单纯形。任意两个 m 维单纯形在仿射群 $\mathrm{Aff}(\mathbb{A})$ 作用下重合。

证明：对任意 $f \in \mathrm{Aff}(\mathbb{A})$，只需证明：若 $\lambda_0+\lambda_1+\cdots+\lambda_m=1$，则

$$f(\lambda_0 P_0+\lambda_1 P_1+\cdots+\lambda_m P_m)=\lambda_0 f(P_0)+\lambda_1 f(P_1)+\cdots+\lambda_m f(P_m).$$

且对任意两组点 $P_0, P_1, \cdots, P_m$，$q_0, q_1, \cdots, q_m \in \mathbb{A}$，如果 $\{\overrightarrow{P_0P_1}, \cdots, \overrightarrow{P_0P_m}\}$ 和 $\{\overrightarrow{q_0q_1}, \cdots, \overrightarrow{q_0q_m}\}$ 都是线性无关组，则存在 $f \in \mathrm{Aff}(\mathbb{A})$ 使 $f(P_i)=q_i$ $(0 \le i \le m)$。上述结论的证明留作练习。 □

定义 7.3.5. 子集合 $M \subset \mathbb{A}$ 称为凸集，如果它满足下列条件：对任意 $P, q \in M$，必有 $\overline{Pq} \subset M$。 □

例 7.3.3：单纯形必为凸集。任意多个凸集的交集必为凸集。 □

定义 7.3.6：设 $M \subset \mathbb{A}$ 是任意的非空集合。$C(M)$ 是 $\mathbb{A}$ 中所有包含 M 的凸集之交。$C(M)$ 称为 M 的凸闭包。 □

引理 7.3.4：设 $M \subset \mathbb{A}$ 是凸集，$P \in \mathbb{A}$。则

$$C(M \cup \{P\}) = \bigcup_{q \in M} \overline{Pq}.$$

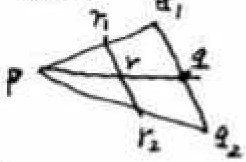
证明：显然 $\bigcup_{q \in M} \overline{Pq} \subset C(M \cup \{P\})$。而 $C(M \cup \{P\})$ 是包含 $M \cup \{P\}$ 的最小凸集。故只需证明集合 $\bigcup_{q \in M} \overline{Pq}$ 是凸集。

下面我们给出两个证明，一个采用"几何直观"，另一个是纯代数化证明。

设 $q_1,q_2\in M$，$r_1\in\overline{pq_1}$，$r_2\in\overline{pq_2}$，$r\in\overline{r_1r_2}$.（如图）

令 $q\in\overline{q_1q_2}$ 是直线 L_{pr} 与 $\overline{q_1q_2}$ 的交点.

则 $q\in M$，$r\in\overline{pq}\subset\bigcup\limits_{q\in M}\overline{pq}$.



亦可采用代数证明，$\overline{pq_1}=\{p\dotplus\lambda\overrightarrow{pq_1}\mid 0\le\lambda\le1\}$

$\overline{pq_2}=\{p\dotplus\lambda\overrightarrow{pq_2}\mid 0\le\lambda\le1\}$. 由于 $r_1\in\overline{pq_1}$，$r_2\in\overline{pq_2}$，可令

$$r_1=p\dotplus\lambda_1\overrightarrow{pq_1},\quad r_2=p\dotplus\lambda_2\overrightarrow{pq_2},\quad 0<\lambda_1\le1,\ 0\le\lambda_2\le1.$$

若 $r=r_1\dotplus\lambda_0\overrightarrow{r_1r_2}\in\overline{r_1r_2}$，（不妨设 $0<\lambda_0<1$，$0<\lambda_1<1$，$0<\lambda_2<1$）.

令 $t_0=\dfrac{\lambda_0\lambda_2}{\lambda_0\lambda_2+(1-\lambda_0)\lambda_1}$，$\lambda=\dfrac{1}{\lambda_0\lambda_2+(1-\lambda_0)\lambda_1}$，则

$$0<t_0<1,\qquad \lambda\ge1,\qquad p\dotplus\lambda\overrightarrow{pr}=q_1\dotplus t_0\overrightarrow{q_1q_2}.$$

从而 $q:=q_1\dotplus t_0\overrightarrow{q_1q_2}\in\overline{q_1q_2}\subset M$，$\overrightarrow{pq}=\lambda\overrightarrow{pr}$. 故

$$r=p\dotplus\overrightarrow{pr}=p\dotplus\frac{1}{\lambda}\overrightarrow{pq}\in\overline{pq}\subset\bigcup_{q\in M}\overline{pq}.$$

□

定理7.3.3：设 $S=C(p_0,p_1,\cdots,p_m)$，f 是 S 上的仿射线性函数. 则，$\forall p\in S$ 有

$$f(p)\le\max\{f(p_0),f(p_1),\cdots,f(p_m)\}.$$

证明：对 m 作归纳法. 设 $M=C(p_0,\cdots,p_{m-1})$ 且

$$f(q)\le\max\{f(p_0),\cdots,f(p_{m-1})\},\quad\forall q\in M.$$

由于 $S\subset C(M\cup\{p_m\})=\bigcup\limits_{q\in M}\overline{p_mq}$，$\forall s\in S$ 有

$$s=p_m\dotplus\lambda\overrightarrow{p_mq},\quad 0\le\lambda\le1.$$

即 $s=(1-\lambda)p_m+\lambda q$. 故 $f(s)=(1-\lambda)f(p_m)+\lambda f(q)\le\max\{f(p_m),f(q)\}$.

从而 $f(s) \le \max\{f(P_0), \cdots, f(P_m)\}$. $(\forall s \in S)$. □

上述定理在线性规划等实际问题中也有重要意义。

习题 7.3

1. 设 $f: \mathbb{A} \to \mathbb{A}'$ 是仿射映射, $P_0, \cdots, P_m \in \mathbb{A}$.

证明: 若 $\lambda_0 + \lambda_1 + \cdots + \lambda_m = 1$, 则

$$f(\lambda_0 P_0 + \cdots + \lambda_m P_m) = \lambda_0 f(P_0) + \cdots + \lambda_m f(P_m).$$

2. 设 $P_0, \cdots, P_m \in \mathbb{A}$, 若 $\overrightarrow{P_0P_1}, \cdots, \overrightarrow{P_0P_m}$ 线性无关, 证明: 对任意 $q_0, \cdots, q_m \in \mathbb{A}'$. $(n = \dim \mathbb{A})$, 存在唯一仿射映射 $f: \mathbb{A} \to \mathbb{A}'$ 使得 $f(P_i) = q_i$ $(0 \le i \le m)$.

3. 证明: 仿射映射将凸集映到凸集。

4. 设 $P_0, P_1, \cdots, P_m \in \mathbb{A}$. 令

$$S = \{\lambda_0 P_0 + \cdots + \lambda_m P_m \mid \lambda_0 + \cdots + \lambda_m = 1, \lambda_i \ge 0, 0 \le i \le m\}$$

证明: (1) S 是凸集. , (2) S 是有限个单纯形的并集.

(3) $S = C(P_0, \cdots, P_m)$.

5. 设 $f: \mathbb{A} \to \mathbb{A}$ 是仿射映射, $F: V \to V$ 是 f 的线性部分.

证明: $f = T_v \circ g$, 其中 T_v 是沿 $v = \overrightarrow{Of(O)}$ 方向的平移. g 是保持 O 不动且线性部分等于 F 的仿射映射(其中 $O \in \mathbb{A}$ 是预先固定的点).

§7.4 二次函数

设 $\mathbb{A}=(\mathbb{A},V)$ 是 n 维仿射空间. 本章第一节讨论过仿射函数 $f:\mathbb{A}\to\mathbb{R}$ 满足 ($O\in\mathbb{A}$ 固定).

$$f(O\dot{+}x)=l(x)+f(O)\ ,\quad 其中\ l\in V^*,\ x\in V.$$

亦可称为一次函数. 本节我们讨论下述二次函数.

<u>定义7.4.1</u>. 函数 $Q:\mathbb{A}\to\mathbb{R}$ 称为<u>二次函数</u>, 如果

$$(4.1)\qquad Q(O\dot{+}x)=q(x)+l(x)+c\ ,\quad \forall\, x\in V.$$

其中 $O\in\mathbb{A}$ 是选定的点, $q:V\to\mathbb{R}$ 是非零二次型, $l\in V^*$, $c\in\mathbb{R}$ 是常数 (即 $c=Q(O)$). □

下面的引理表明, 在公式 (4.1) 中的二次型 $q:V\to\mathbb{R}$ 与点 $O\in\mathbb{A}$ 的选取无关. (但一次式 $l(x)$ 和常数项 c 与 O 的选取有关). 所以, q 也称为由 Q 确定的二次型, 定义 $\mathrm{rank}(Q)=\mathrm{rank}(q)$.

<u>引理7.4.1</u>. 设 $f:V\times V\to\mathbb{R}$ 是由 $q:V\to\mathbb{R}$ 唯一确定的对称双线性函数, $O'\in\mathbb{A}$, $v_0=\overrightarrow{OO'}$. 则

$$\cancel{Q(O'\dot{+}x)=q(x)+[2f(\overrightarrow{OO'},x)+l(x)]+[q(\overrightarrow{OO'})+l(\overrightarrow{OO'})+Q(O)]}$$

$$Q(O'\dot{+}x)=q(x)+[2f(v_0,x)+l(x)]+[q(v_0)+l(v_0)+c].$$

证明: $Q(O'\dot{+}x)=Q(O\dot{+}(v_0+x))=q(v_0+x)+l(v_0+x)+c.$

但 $q(v_0+x)=f(v_0+x,v_0+x)=f(x,x)+2f(v_0,x)+f(v_0,v_0)$. 故

$$Q(O'\dot{+}x)=q(x)+[2f(v_0,x)+l(x)]+[q(v_0)+l(v_0)+c].$$

□

定义7.4.2. 设 $Q: \mathbb{A} \to \mathbb{R}$ 是一个二次函数，点 $P \in \mathbb{A}$ 称为Q的一个中心，如果

$$Q(P \dotplus x) = Q(P \dotplus (-x)), \quad \forall\, x \in V.$$

令 $C(Q)$ 表示Q的全部中心的集合。若 $C(Q)$ 非空，则称Q是中心的。 □

设 $(O, e_1, \cdots, e_n)$ 是仿射空间 $\mathbb{A} = (\mathbb{A}, V)$ 的一个坐标系。则二次函数 $Q(O \dotplus x) = q(x) + l(x) + c$ 的坐标表达式为

$$(4.2) \qquad Q\Big(O \dotplus \sum_{i=1}^{n} x_i e_i\Big) = \sum_{1 \le i,j \le n} a_{ij} x_i x_j + \sum_{i=1}^{n} b_i x_i + c, \quad a_{ij} = a_{ji}.$$

其中 $c = Q(O)$，$b_i = \frac{\partial Q}{\partial x_i}(O)$。

$P = O \dotplus \sum_{i=1}^{n} x_i e_i \in \mathbb{A}$ 是 Q 的中心当且仅当

$$(4.3) \qquad \begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n + \frac{1}{2}b_1 = 0 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n + \frac{1}{2}b_2 = 0 \\ \vdots \\ a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n + \frac{1}{2}b_n = 0 \end{cases}.$$

定理7.4.1. 设 $Q(O \dotplus x) = q(x) + l(x) + c$。如果 $C(Q) \neq \emptyset$，则 $C(Q) \subset \mathbb{A}$ 是一个 $(n-r)$ 维平面，该平面的方向子空间为

$$\operatorname{Ker} q = \{ x \in V \mid f(x,y) = 0, \forall y \in V \}$$

其中 $f: V \times V \to \mathbb{R}$ 是对称双线性函数使得 $q(x) = f(x,x)$，r 是 q 的秩。特别，Q仅有一个中心的充要条件是 $r = n$（即 $q(x)$ 非退化）。

证明：$P = O \dotplus x$ 是Q的中心 $\Leftrightarrow Q(O \dotplus (x+y)) = Q(O \dotplus (x-y))$，$\forall y \in V$。

令 $f: V\times V \to \mathbb{R}$ 是满足 $q(x)=f(x,x)$ 的对称双线性函数. 则 $Q(O\dot{+}(x+y)) = Q(O\dot{+}(x-y)) \Longleftrightarrow f(x+y,x+y)+l(x+y)=f(x-y,x-y)+l(x-y)$

即 $f(x,y)+\frac{1}{2}l(y)=0$, $\forall y\in V$, 它的坐标表达式正好是(4.3).

若 $C(Q)$ 非空, 则 $C(Q)$ 是中心的坐标恰为(4.3)的解空间. □

定义7.4.3. 二次函数 $Q_i: \mathbb{A}\to\mathbb{R}$ $(i=1,2)$ 称为仿射等价, 如果存在 $\varphi\in \mathrm{Aff}(\mathbb{A})$ 使得 $Q_2=\varphi^*(Q_1):=Q_1\circ\varphi$. 若 $\mathbb{A}$ 是欧氏几何空间, 且存在 $\varphi\in \mathrm{Isom}(\mathbb{A})$ 使得 $Q_2=\varphi^*(Q_1):=Q_1\circ\varphi$, 则称 Q_1, Q_2 是保距等价的. □

不难验证, 仿射等价和保距等价确实在所有二次函数的集合中定义了等价关系。

定理7.4.2. 设 $Q: \mathbb{A}\to\mathbb{R}$ 是秩为 r 的二次函数. 则

(1) 如果 Q 是中心的(即: $C(Q)$ 非空), 则存在以中心 $O\in\mathbb{A}$ 为原点的坐标系 $(O, e_1,\dots,e_n)$ 使得

$$Q\Big(O\dot{+}\sum_{i=1}^n x_ie_i\Big)=x_1^2+\cdots+x_s^2-x_{s+1}^2-\cdots-x_r^2+c.$$

其中 s, r, c 仅依赖于 Q, 不依赖于坐标系的选取.

(2) 若 Q 无中心(即 $C(Q)$ 是空集合), 则存在坐标系 $(O, e_1,\dots,e_n)$ 使得

$$Q\Big(O\dot{+}\sum_{i=1}^n x_ie_i\Big)=x_1^2+\cdots+x_s^2-x_{s+1}^2-\cdots-x_r^2+x_{r+1}.$$

(3) 有中心的二次函数与无中心的二次函数不是仿射等价.

证明: (1) 设 $Q(O\dot{+}x)=q(x)+l(x)+c$, 其中 $O\in C(Q)$. 由二次型理论, 存在 V 的一组基使

$$q\Big(\sum_{i=1}^n x_ie_i\Big)=x_1^2+\cdots+x_s^2-x_{s+1}^2-\cdots-x_r^2.$$

故 $Q(O\dotplus\sum_{i=1}^n x_ie_i)=x_1^2+\cdots+x_s^2-x_{s+1}^2-\cdots-x_r^2+b_1x_1+\cdots+b_nx_n+C$.

由于 $O\in\mathbb{A}$ 是 Q 的中心，O 的坐标 $(0,0,\cdots,0)$ 满足方程组 (4.3)，从而 $b_1=b_2=\cdots=b_n=0$. 即

$$Q(O\dotplus\sum_{i=1}^n x_ie_i)=x_1^2+\cdots+x_s^2-x_{s+1}^2-\cdots-x_r^2+C.$$

其中 $C=Q(O)$. 因 $C(Q)\subset\mathbb{A}$ 是一个平面，其方向子空间为

$$\mathrm{Ker}(q)=\{v\in V\mid f(v,y)=0,\forall y\in V\}$$

其中 $f: V\times V\to\mathbb{R}$ 是由 $q: V\to\mathbb{R}$ 唯一确定的对称双线性函数满足：$q(x)=f(x,x)$. 所以，$\forall\, O'\in C(Q)$，$q(\overrightarrow{OO'})=0$. 特别

$$Q(O')=Q(O\dotplus\overrightarrow{OO'})=C=Q(O).$$

(2) 任取 $O'\in\mathbb{A}$，令 $Q(O'\dotplus x)=q(x)+l(x)+C$. 存在一组基 $e_1',\cdots,e_n'\in V$ 使得

$$q\left(\sum_{i=1}^n x_i'e_i'\right)=x_1'^2+\cdots+x_s'^2-x_{s+1}'^2-\cdots-x_r'^2$$

令 $l(e_i')=2b_i$，则

$$\begin{aligned}Q(O'+\sum_{i=1}^n x_i'e_i')&=x_1'^2+\cdots+x_s'^2-x_{s+1}'^2-\cdots-x_r'^2+2b_1x_1'+\cdots+2b_nx_n'+C\\&=\sum_{i=1}^s(x_i'+b_i)^2-\sum_{i=s+1}^r(x_i'-b_i)^2+2b_{r+1}x_{r+1}'+\cdots+2b_n'x_n'+\mu\end{aligned}$$

其中 $\mu=C-(b_1^2+\cdots+b_s^2)+(b_{s+1}^2+\cdots+b_r^2)$. 令

$$P_0=O'\dotplus v,\quad v=-\sum_{i=1}^s b_ie_i'+\sum_{i=s+1}^r b_ie_i'$$

则 $\quad Q(P_0\dotplus\sum_{i=1}^n x_i'e_i')=x_1'^2+\cdots+x_s'^2-x_{s+1}'^2-\cdots-x_r'^2+2b_{r+1}x_{r+1}'+\cdots+2b_nx_n'+\mu.$

由于 $C(Q)=\phi$，故 $b_{r+1},\cdots,b_n$ 不全为零. 作坐标变换：

$x_i=x_i'\ (1\le i\le r)$，$x_{r+1}=2b_{r+1}x_{r+1}'+\cdots+2b_nx_n'+\mu$，及

$x_i=c_{i1}x_1'+\cdots+c_{in}x_n'\ (r+1<i\le n)$.（适当选取 $c_{ij}\in\mathbb{R}$）. 则存在坐标系 $(O,e_1,\cdots,e_n)$ 使.

$$Q(O\dotplus\sum_{i=1}^{n}x_ie_i)=x_1^2+\cdots+x_s^2-x_{s+1}^2-\cdots-x_r^2+x_{r+1}.$$

(3) 设 $P\in\mathbb{A}$ 是 Q 的中心，$\forall\,\varphi\in \mathrm{Aff}(\mathbb{A})$，令 $P'\in\mathbb{A}$ 使 $\varphi(P')=P$. 则，$\forall\,x\in V$，有

$$\varphi^*(Q)(P'\dotplus x)=Q(\varphi(P'\dotplus x))=Q(\varphi(P')\dotplus D(\varphi)(x))$$
$$=Q(P\dotplus D(\varphi)(x))=Q(P\dotplus D(\varphi)(-x))=\varphi^*Q(P'\dotplus(-x)).$$

故 P' 是 $\varphi^*(Q)$ 的中心. 所以，有中心的二次函数不可能仿射等价于无中心的二次函数.

□

定理7.4.3：设 $\mathbb{A}=(A,V)$ 是 n 维欧氏几何空间，$Q:\mathbb{A}\to\mathbb{R}$ 是秩为 r 的二次函数. 则

(1) 当 $O\in C(Q)$ 时，存在直角坐标系 $(O,e_1,\cdots,e_n)$ 使
$$Q(O\dotplus\sum_{i=1}^{n}x_ie_i)=\lambda_1x_1^2+\cdots+\lambda_rx_r^2+c,\quad \lambda_i\neq0\ (1\le i\le r).$$

(2) 当 $C(Q)=\varnothing$ 时，存在直角坐标系 $(O,e_1,\cdots,e_n)$ 使
$$Q(O\dotplus\sum_{i=1}^{n}x_ie_i)=\lambda_1x_1^2+\cdots+\lambda_rx_r^2+\mu x_{r+1},\ (\lambda_i\neq0,\ \mu>0).$$

证明：(1) 令 $Q(O\dotplus x)=q(x)+l(x)+c$，$e_1,\cdots,e_n$ 是 V 的标准正交基 使得 $q(x_1e_1+\cdots+x_ne_n)=\lambda_1x_1^2+\cdots+\lambda_rx_r^2$，$(\lambda_i\neq0)$.

$$Q(O\dotplus\sum_{i=1}^{n}x_ie_i)=\lambda_1x_1^2+\cdots+\lambda_rx_r^2+c.$$

且 $Q(O')=Q(O)=c\quad(\forall\,O'\in C(Q))$.

(2) 由定理7.4.2的证明可知，存在直角坐标系 $(P_0,e_1',\cdots,e_n')$ 使得
$$Q(P_0\dotplus\sum_{i=1}^{n}x_i'e_i')=\lambda_1x_1'^2+\cdots+\lambda_rx_r'^2+\sum_{i=r+1}^{n}b_ix_i'+d.$$

考虑

$$\begin{pmatrix} x_1\\ \vdots\\ x_r\\ x_{r+1}\\ x_{r+2}\\ \vdots\\ x_n \end{pmatrix}=\begin{pmatrix} \overbrace{1 \ddots 1}^{r} & & & \text{O} & \\ 0 & \cdots & 0 & \frac{b_{r+1}}{\mu},\cdots, & \frac{b_n}{\mu}\\ c_{r+2,1}, & \cdots & c_{r+2,r+1} & \cdots & c_{r+2,n}\\ \vdots & & \vdots & & \vdots\\ c_{n1}, & \cdots & c_{n,r+1}, & \cdots & c_{nn} \end{pmatrix}\begin{pmatrix} x_1'\\ \vdots\\ x_r'\\ x_{r+1}'\\ x_{r+2}'\\ \vdots\\ x_n' \end{pmatrix}+\begin{pmatrix} 0\\ \vdots\\ 0\\ \gamma/\mu\\ 0\\ \vdots\\ 0 \end{pmatrix}$$

适当选取 $c_i=(c_{i1},\cdots,c_{in})$ $(r+1<i\leq n)$ 使得

$$A=\begin{pmatrix} \overbrace{1 \ddots 1}^{r} & & & \text{O} \\ 0\cdots 0 & \frac{b_{r+1}}{\mu}, & \cdots, & \frac{b_n}{\mu}\\ c_{r+2,1}, & \cdots & \cdots & c_{r+2,n}\\ \vdots & & & \vdots\\ c_{n1}, & \cdots & \cdots & c_{nn} \end{pmatrix}_{n\times n}$$

，其中 $\mu=\sqrt{b_{r+1}^2+\cdots+b_n^2}$

是正交矩阵. 由此可得直角坐标系 $(O,e_1,\cdots,e_n)$ 使

$$Q\left(O\dotplus\sum_{i=1}^{n}x_ie_i\right)=\lambda_1x_1^2+\cdots+\lambda_rx_r^2+\mu x_{r+1}$$

□

习题 7.4

1. 设 $Q(O\dotplus x)=q(x)+l(x)+c$ 是二次函数. 证明: $P\in\mathbb{A}$ 是中心的充要条件是

$$Q(P\dotplus y)=Q(P)+q(y),\ \forall y\in V.$$

2. 设二次函数 $Q:\mathbb{A}\to\mathbb{R}$ 在坐标系 $(O,e_1,\cdots,e_n)$ 下分别由下述二次多项式定之.

(1) $Q(O\dotplus\sum_{i=1}^{n}x_ie_i)=2\sum_{1\leq i<j\leq n}x_ix_j+2\sum_{i=1}^{n}x_i+1$

(2) $Q(O\dotplus\sum_{i=1}^{n}x_ie_i)=\sum_{i=1}^{n}x_i^2+2\sum_{1\leq i<j\leq n}x_ix_j+x_1$

试分别由 Q 的中心.

§7.5. 二次曲面.

设 $\mathbb{A}=(\mathbb{A},V)$ 是一个 n 维仿射空间, $\mathbb{A}\xrightarrow{Q}\mathbb{R}$ 是一个二次函数. 则 Q 的零点集

$$S_Q := \{ P\in\mathbb{A} \mid Q(P)=0 \}.$$

称为 $\mathbb{A}$ 中的二次曲面. 在下面的讨论中, 我们总是假设 $n\geq 2$, S_Q 非空. 我们首先要解决 S_Q 与 Q 的关系问题.

任意固定一个坐标系 $(O, e_1,\cdots,e_n)$, 则 $P\in S_Q$ 的充要条件是 P 的坐标满足二次多项式方程.

$$(x_1,\cdots,x_n)\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn}\end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n\end{pmatrix} + 2(b_1,\cdots,b_n)\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n\end{pmatrix} + c = 0 .$$

因此, 若存在坐标系 $(O, e_1,\cdots,e_n)$ 使得

$$Q(O+\sum_{i=1}^{n} x_i e_i) = x_1^2+\cdots+x_r^2 .$$

则 S_Q 是 $\mathbb{A}$ 中的 $(n-r)$ 维平面.

定义7.5.1. 点 $P\in\mathbb{A}$ 称为二次曲面 S_Q 的一个中心(图5.1). 如果 P 满足条件: 若 $P+x\in S_Q$, 则 $P+(-x)\in S_Q$. (即 S_Q 中的点关于 P 点是对称的). 若 S_Q 的中心 P 同时在 S_Q 上, 则称 P 是 S_Q 的一个顶点. (见图5.2) □

引理7.5.1. (1) 设 $L\subset\mathbb{A}$ 是直线, 若 $L\not\subset S_Q$, 则 L 与 S_Q 最多相交两个点.

(2) 若 $P\in S_Q$ 是 S_Q 的顶点, $Q\in S_Q$ 且 $Q\neq P$, 则 $L_{PQ}\subset S_Q$.

证明：(1) 设 $L = O \dotplus \langle v \rangle = \{ O \dotplus tv \mid t \in \mathbb{R} \}$,

$$Q(O \dotplus x) = q(x) + l(x) + c.$$

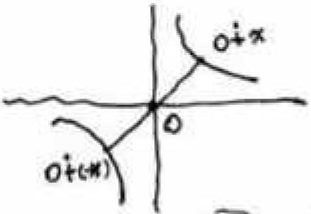
则 L 与 S_Q 的交点由方程 $Q(O \dotplus tv) = q(v)t^2 + l(v)t + c = 0$ 确定. 如它的系数全为零, 则 $L \subset S_Q$. 故条件 $L \not\subset S_Q$ 表明方程 $q(v)t^2 + l(v)t + c = 0$ 最多仅有二个实数解.

(2) 设 $q = p \dotplus x$, 则 $p,\ q = p \dotplus x,\ p \dotplus (-x) \in S_Q$, 且 $p \dotplus (-x) \in L_{pq}$. 所以 $L_{pq} \cap S_Q$ 包含三个点. 由(1), 得 $L_{pq} \subset S_Q$ □.

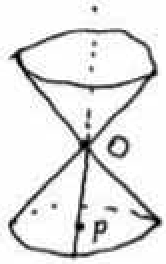
推论7.5.1. 如果二次曲面 S 上的每个点都是顶点, 则 $S \subset \mathbb{A}$ 是平面.

证明：由引理7.5.1中的结论(2), S 中任意两点连成的直线皆在 S 中. 由定理7.1.3知, S 是平面 □.

定义7.5.2：二次曲面 $S \subset \mathbb{A}$ 称为二次锥面, 如果 S 有顶点, 且 S 不是平面. □



(图5.1)



(图5.2)

定理7.5.1：设 Q_1, Q_2 是 $\mathbb{A}$ 上的二次函数, 且

$$S_{Q_1} = S_{Q_2} := S.$$

则或者 S 是平面, 或者存在 $\lambda \in \mathbb{R}$ 使 $Q_2 = \lambda Q_1$.

证明：无妨设 $S=S_{Q_1}=S_{Q_2}$ 不是平面。由推论(5.1)，存在 $O\in S$ 使得 O 不是 S 的中心（即 O 不是 S 的顶点）。设

$$Q_1(O\dotplus x)=g_1(x)+l_1(x), \quad Q_2(O\dotplus x)=g_2(x)+l_2(x).$$

其中 $l_1, l_2\in V^*$ 是非零线性函数。如果 $x\in V$ 满足条件 $g_1(x)\neq 0$，$l_1(x)\neq 0$，令 $t_x=-\frac{l_1(x)}{g_1(x)}$，则点 $O\dotplus t_x x$ 是直线 $L=\{O\dotplus tx \mid t\in\mathbb{R}\}$ 与 S_{Q_1} 的交点。由 $S_{Q_1}=S_{Q_2}$，它也是 L 与 S_{Q_2} 的交点。所以 $t_x=-\frac{l_1(x)}{g_1(x)}$ 是 $g_2(x)t^2+l_2(x)t=0$ 的解。从而 $g_1(x)l_2(x)=g_2(x)l_1(x)$ 对任意 $x\in\left\{x\in V \,\middle|\, \begin{matrix} g_1(x)\neq 0 \\ l_1(x)\neq 0\end{matrix}\right\}$ 成立。令 $f(x)=g_1(x)l_1(x)$，则

$$(5.1) \qquad g_1(x)l_2(x)f(x)=g_2(x)l_1(x)f(x), \quad \forall x\in V.$$

设 $(x_1,\cdots,x_n)$ 是 x 在某固定坐标系 $(O, e_1,\cdots,e_n)$ 下的坐标，则 $g_1(x), g_2(x), l_1(x), l_2(x), f(x)$ 是关于 $x_1,\cdots,x_n$ 的非零多项式。从而 (5.1) 在多项式环 $\mathbb{R}[x_1,\cdots,x_n]$ 中成立。由于 $f(x)\in\mathbb{R}[x_1,\cdots,x_n]$ 是非零多项式，由 (5.1) 可消去 $f(x)$ 得

$$(5.2) \qquad g_1(x)l_2(x)=g_2(x)l_1(x), \quad \forall x\in V.$$

若 l_1, l_2 在 V^* 线性相关，即存在 $\lambda\in\mathbb{R}$ 使 $l_2(x)=\lambda l_1(x)$，则由 (5.2) 可得 $g_2(x)=\lambda g_1(x)$。从而 $Q_2=\lambda Q_1$。

若 l_1, l_2 在 V^* 线性无关，令 $e_1, e_2, \cdots, e_n\in V$ 是一组基，满足条件：$l_1(e_1)=1$，$l_1(e_i)=0$ $(i\neq 1)$，$l_2(e_2)=1$，$l_2(e_i)=0$ $(i\neq 2)$。则

$$g_1(x)x_2=g_2(x)x_1, \quad \forall x=\sum_{i=1}^{n}x_ie_i\in V.$$

从而存在 $l\in V^*$ 使 $g_1(x)=l(x)x_1$，$g_2(x)=l(x)x_2$。即

$$Q_1(O\dotplus x)=(l(x)+1)x_1, \quad Q_2(O\dotplus x)=(l(x)+1)x_2.$$

特别，S_{Q_1} 包含超平面 $H_1=\{O\dotplus\sum_{i=1}^{n}x_ie_i \mid x_1=0\}$。由 $S_{Q_2}=S_{Q_1}$ 可知

$Q_2(O\dot{+}\sum_{i=1}^{n}x_ie_i)=(l(x)+1)x_2$ 在 H_1 上恒等于零. 但不难构造 $P=O\dot{+}x_2e_2\in H_1$ 使 $Q_2(P)=(x_2l(e_2)+1)x_2\neq 0$ (取 $0\neq x_2\in\mathbb{R}$ 使 $x_2l(e_2)+1\neq 0$ 即可). 此矛盾表明 l_1, l_2 在 V^* 中线性相关. □

推论7.5.2: (1) 设二次曲面 S_{Q_1} 不是平面, $O\in\mathbb{A}$. 则

O 是 S_{Q_1} 的中心 $\Leftrightarrow$ O 是 Q_1 的中心.

(2) 设二次曲面 S_Q 不是平面. 若 S_Q 在平移 T_v 下不变. 则二次函数在 T_v 的作用下也不变.

证明: (1) 若 O 是 S_{Q_1} 的中心, 令 $Q_2: \mathbb{A}\to\mathbb{R}$ 定义如下:

$$Q_2(O\dot{+}x):=Q_1(O\dot{+}(-x)).$$

则 $S_{Q_1}=S_{Q_2}$. 事实上, $\forall P=O\dot{+}x\in S_{Q_1}$, 则 $O\dot{+}(-x)\in S_{Q_1}$ (因为 O 是 S_{Q_1} 的中心). 从而 $Q_2(P)=Q_2(O\dot{+}x)=Q_1(O\dot{+}(-x))=0$. 即 $S_{Q_1}\subset S_{Q_2}$, 反之亦然. 由定理7.5.1, $Q_2=\lambda Q_1$. 但 Q_1, Q_2 有相同的二次项, 所以 $\lambda=1$. 从而 $Q_2=Q_1$ 即

$$Q_1(O\dot{+}x)=Q_2(O\dot{+}x)=Q_1(O\dot{+}(-x)),\ \forall x\in V.$$

故 O 是 Q_1 的中心. 反之显然.

(2) 令 $Q':=Q\cdot T_v$, 则 $S_Q=S_{Q'}$. 事实上, $\forall P\in S_Q$, 有 $T_v(P)\in S_Q$ 从而 $Q'(P)=Q(T_v(P))=0$. 即 $S_Q\subseteq S_{Q'}$. 反之, $\forall P\in S_{Q'}$ 有 $T_v(P)\in S_Q$, 从而 $P\in S_Q$ 即 $S_{Q'}\subseteq S_Q$. 由定理7.5.1, 存在 $\lambda\in\mathbb{R}$ 使 $Q'=\lambda Q$. 但 Q' 与 Q 有相同的二次项. 故 $\lambda=1$. 从而 $Q'=Q$ 即 $T_v^*(Q):=Q\cdot T_v=Q$. □

设 $Q(O\dotplus x)=q(x)+l(x)+c$ 是 $\mathbb{A}=(A,V)$ 上的二次函数。下面我们寻找 V 中的向量 v 使得 S_Q 在平移 T_v 下不变的条件。

由于二次型 $q(x)$ 与点 O 的选取无关，令 $f: V\times V\to\mathbb{R}$ 是对称双线性函数使得 $q(x)=f(x,x)$ $(\forall x\in V)$，则子空间

$$\mathrm{Ker}(q)\cap\mathrm{Ker}(l)\subset V$$

与点 O 的选取无关。其中 $\mathrm{Ker}(q)=\{x\in V\mid f(x,y)=0,\ \forall y\in V\}$。事实上，若 $O'\in\mathbb{A}$ 是另一个点，令 $Q(O'\dotplus x)=q(x)+l'(x)+c'$。则由引理7.4.1，$l'(x)=2f(v_0,x)+l(x)$（其中 $v_0=\overrightarrow{OO'}\in V$）。容易验证：

$$\mathrm{Ker}(q)\cap\mathrm{Ker}(l)=\mathrm{Ker}(q)\cap\mathrm{Ker}(l').$$

引理7.5.2. 设 $Q(O\dotplus x)=q(x)+l(x)+c$ 是二次函数。令

$$\mathrm{Ker}(Q):=\mathrm{Ker}(q)\cap\mathrm{Ker}(l).$$

则函数 Q 在 T_v 作用下不变的充要条件是 $v\in\mathrm{Ker}(Q)$。

证明：Q 在 T_v 作用不变 $\Leftrightarrow$ $Q(O\dotplus x)=Q(O\dotplus x+v)$，$\forall x\in V$。

即 $Q(O\dotplus x)=Q(O\dotplus x+v)$ 对任意 $x\in V$ 成立的充要条件是 $v\in\mathrm{Ker}(Q)$。 $\square$

若 $\mathrm{Ker}(Q)\subset V$ 是非零子空间，$\forall P\in S_Q$，则平面 $P\dotplus\mathrm{Ker}(Q)\subset S_Q$。这样的曲面 S_Q 称为柱形二次曲面。

引理7.5.3. 若 S_Q 是非柱形二次曲面（即 $\mathrm{Ker}(Q)=\{0\}$），则 S_Q 最多只有一个中心。

证明：设 $O,O'\in\mathbb{A}$ 是 S_Q 的两个中心，则

$$Q(O\dotplus x)=Q(O\dotplus(-x)),\quad Q(O'\dotplus x)=Q(O'\dotplus(-x)),\quad \forall x\in V.$$

从而 $Q(O'\dot{+}x)=Q(O\dot{+}(\overrightarrow{OO'}+x))=Q(O\dot{+}(-\overrightarrow{OO'}-x))$

$=Q(O'\dot{+}(\overrightarrow{O'O}-\overrightarrow{OO'}-x))=Q(O'\dot{+}(-2\overrightarrow{OO'}-x))$

$=Q((O'\dot{+}x)\dot{+}2\overrightarrow{OO'})$.

即 Q 在平移 $T_{2\overrightarrow{OO'}}$ 作用下不变. 由引理7.5.2, $2\overrightarrow{OO'}\in\mathrm{Ker}(Q)$.
但 S_Q 是非柱形二次曲面, 故 $\overrightarrow{OO'}=0$. 从而 $O=O'$. $\square$.

设 $\dim\mathrm{Ker}(Q)=m>0$ (即 S_Q 是柱形二次曲面). 选取 V 的一组基 $e_1,\cdots,e_n$ 使 $e_{n-m+1},\cdots,e_n\in\mathrm{Ker}(Q)$, $O\in S_Q$. 则 $Q(O\dot{+}\sum_{i=1}^{n}x_ie_i)=q(\sum_{i=1}^{n-m}x_ie_i)+l(\sum_{i=1}^{n-m}x_ie_i):=Q_0(O\dot{+}x)$
其中 $x\in V_0:=\langle e_1,\cdots,e_{n-m}\rangle$. Q_0 是 $(n-m)$ 维空间 $O\dot{+}V_0$ 上的二次函数, S_{Q_0} 是 $(n-m)$ 维仿射空间中的二次曲面, 称为柱形曲面 S_Q 的底曲面. 事实上,

$$S_Q=\{P\dot{+}u \mid P\in S_{Q_0},\ u\in\mathrm{Ker}(Q)\}.$$

所以二次曲面的研究可以归结为研究非柱形二次曲面.

定理7.5.2 (仿射分类). 设 $\mathbb{A}=(\mathbb{A},V)$ 是 n 维仿射空间. $S\subset\mathbb{A}$ 是非柱形二次曲面. 则在适当的仿射坐标系下 (或等价的说法: 经过适当仿射变换) S 可由下述标准形式的方程之一定义:

$\mathrm{I}_{s,n}$: $x_1^2+\cdots+x_s^2-x_{s+1}^2-\cdots-x_n^2=1$, $(0<s\le n)$.

$\mathrm{I}'_{s,n}$: $x_1^2+\cdots+x_s^2-x_{s+1}^2-\cdots-x_n^2=0$, $(\frac{n}{2}\le s<n)$.

$\mathrm{II}_{s,n-1}$: $x_1^2+\cdots+x_s^2-x_{s+1}^2-\cdots-x_{n-1}^2=x_n$, $(\frac{n-1}{2}\le s\le n-1)$.

证明: 若 S 有中心 $O\in\mathbb{A}$, 则 S 由形如 $Q(O\dot{+}x)=q(x)+c$ 的二次函数定义, 其中 $q(x)$ 非退化. $c=Q(O)$.

如果 $c\neq 0$ (即 $O\notin S$, 或者说 S 没有顶点), 无妨设 $Q(O)=-1$. 从而存在坐标系 $(O, e_1,\cdots,e_n)$ 使得 S 由

$$I_{s,n}:\quad x_1^2+\cdots+x_s^2-x_{s+1}^2-\cdots-x_n^2=1 \qquad (0<s\leq n)$$

定义。如果 $c=0$, (即 $O\in S$ 是顶点), 无妨设 $q(x)$ 的正惯性指数满足 $\frac{n}{2}\leq s<n$ (若 $s=n$, 则 S 是一个点). 则在适当坐标 $(O, e_1,\cdots,e_n)$ 下, S 由方程

$$I'_{s,n}:\quad x_1^2+\cdots+x_s^2-x_{s+1}^2-\cdots-x_n^2=0, \qquad (\tfrac{n}{2}\leq s<n)$$

若 S 无中心, 设 $O\in S$, 则 S 可由形如 $Q(O+x)=q(x)+l(x)$ 的二次函数定义, 其中 $l\in V^*$ 非零. 由 $\mathrm{Ker}(q)\cap\mathrm{Ker}(l)=\{0\}$ 可得 $\dim(\mathrm{Ker}(q))\leq 1$. 若 $\dim(\mathrm{Ker}(q))=\{0\}$, 则 q 非退化, 从而线性方程 (4.3) 有唯一解. 即 Q 有唯一中心, 此与 S 无中心矛盾. 所以二次型 $q(x)$ 的秩为 $n-1$. 无妨设 $q(x)$ 的正惯性指数满足 $\frac{n-1}{2}\leq s\leq n-1$. 则在适当坐标系下, S 可由下面标准方程定义.

$$II_{s,n-1}:\quad x_1^2+\cdots+x_s^2-x_{s+1}^2-\cdots-x_{n-1}^2=x_n, \qquad (\tfrac{n-1}{2}\leq s\leq n-1).$$

□

~~定理7.5.2: 椭球面 ($I_{n,n}$), 双曲面 ($I_{s,n}$, $0<s<n$), 二次锥面 ($I'_{s,n}$); 椭圆抛物面 ($II_{n-1,n-1}$) 双曲抛物面 ($II_{s,n}$, $\frac{n-1}{2}\leq s<n-1$).~~ □

例7.5.1 ($n=2$) 二次曲线仿射等价于下列曲线之一.

$I_{n,s}$: (1) $x_1^2+x_2^2=1$ (圆). (2) $x_1^2-x_2^2=1$ (双曲线)

$I'_{n,s}$: (3) $x_1^2-x_2^2=0$ (两相交于顶点的直线), $II_{s,n-1}$: $x_1^2=x_2$ (抛物线).

□

例 7.5.2（三维空间二次曲面的仿射分类）.

$I_{n,s}$：$x_1^2+x_2^2+x_3^2=1$（球面），$x_1^2+x_2^2-x_3^2=1$（单叶双曲面），

$x_1^2-x_2^2-x_3^2=1$（双叶双曲面）

$I'_{n,s}$：$x_1^2+x_2^2-x_3^2=0$（二次锥面），

$II_{n,s}$：$x_1^2+x_2^2=x_3$（椭圆抛物面），$x_1^2-x_2^2=x_3$（双曲抛物面）.

$\square$

定理 7.5.3（欧氏几何分类）. 设 $S\subset \mathbb{A}=(A,V)$ 是 n 维欧几里得空间 $\mathbb{A}$ 中的非柱形二次曲面. 则 S 等距等价于下列方程定义的曲面之一（其中 $a_i>0$）：

$$I_{s,n}: \frac{x_1^2}{a_1^2}+\cdots+\frac{x_s^2}{a_s^2}-\frac{x_{s+1}^2}{a_{s+1}^2}-\cdots-\frac{x_n^2}{a_n^2}=1,\quad 0<s\le n$$

$$I'_{s,n}: \frac{x_1^2}{a_1^2}+\cdots+\frac{x_s^2}{a_s^2}-\frac{x_{s+1}^2}{a_{s+1}^2}-\cdots-\frac{x_n^2}{a_n^2}=0,\quad \frac{n}{2}\le s<n$$

$$II_{s,n}: \frac{x_1^2}{a_1^2}+\cdots+\frac{x_s^2}{a_s^2}-\frac{x_{s+1}^2}{a_{s+1}^2}-\cdots-\frac{x_{n-1}^2}{a_{n-1}^2}=x_n,\quad \frac{n-1}{2}\le s\le n-1.$$

证明：若 S 有中心 $O\in\mathbb{A}$. 则 S 由形如 $Q(O+v)=q(v)+c$ 的二次函数定义，其中 $q(v)$ 非退化，$c=Q(O)$.

如果 $c\neq 0$，不妨设 $c=-1$. 从而存在直角坐标系 $(O,e_1,\cdots,e_n)$ 使得

$$Q(O+\sum_{i=1}^n x_ie_i)=\lambda_1x_1^2+\cdots+\lambda_sx_s^2-\lambda_{s+1}x_{s+1}^2-\cdots-\lambda_nx_n^2-1$$

其中 $\lambda_i>0$。令 $a_i=\frac{1}{\sqrt{\lambda_i}}$，可得 S 在 $(O,e_1,\cdots,e_n)$ 下的定义方程

$$I_{s,n}: \frac{x_1^2}{a_1^2}+\cdots+\frac{x_s^2}{a_s^2}-\frac{x_{s+1}^2}{a_{s+1}^2}-\cdots-\frac{x_n^2}{a_n^2}=1.$$

如果 $c=0$（即 O 是 S 的顶点），不妨设 $q(v)$ 的正惯性指数满足 $\frac{n}{2}\le s<n$（若 $s=n$，则 S 是一个点）. 故 S 在适当的直角坐标系 $(O,e_1,\cdots,e_n)$ 下由如下述方程定义：

$$I'_{s,n}: \frac{x_1^2}{a_1^2}+\cdots+\frac{x_s^2}{a_s^2}-\frac{x_{s+1}^2}{a_{s+1}^2}-\cdots-\frac{x_n^2}{a_n^2}=0,\quad \frac{n}{2}\le s<n$$

若 S 是无中心二次曲面，设 $O\in S$，则 S 由二次函数

$$Q(O+x)=q(x)+l(x)$$

定义，其中 $q(x)$ 是秩为 $n-1$ 的二次型，$l\in V^*$ 非零. 无妨设 $q(x)$ 的正惯性指数满足条件 $\frac{n-1}{2}\le s\le n-1$. 适当修改定理7.4.3的证明，可以证明存在直角坐标系 $(O,e_1,\cdots,e_n)$ 使

$$Q\left(O+\sum_{i=1}^{n}x_ie_i\right)=\lambda_1x_1^2+\cdots+\lambda_sx_s^2-\lambda_{s+1}x_{s+1}^2-\cdots-\lambda_{n-1}x_{n-1}^2-\mu x_n.$$

其中 $\mu>0$，$\lambda_i>0$. 令 $a_i=\sqrt{\frac{\mu}{\lambda_i}}$，$S$ 可由方程

$$\mathrm{II}_{s,n}:\ \frac{x_1^2}{a_1^2}+\cdots+\frac{x_s^2}{a_s^2}-\frac{x_{s+1}^2}{a_{s+1}^2}-\cdots-\frac{x_{n-1}^2}{a_{n-1}^2}=x_n,\quad \frac{n-1}{2}\le s\le n-1.$$

□

由 $\frac{x_1^2}{a_1^2}+\cdots+\frac{x_n^2}{a_n^2}=1$ 定义的二次曲面称为椭球面，由

$$\frac{x_1^2}{a_1^2}+\cdots+\frac{x_s^2}{a_s^2}-\frac{x_{s+1}^2}{a_{s+1}^2}-\cdots-\frac{x_n^2}{a_n^2}=1\qquad(0<s<n)$$

定义的二次曲面称为双曲面. 而由下述方程

$$\frac{x_1^2}{a_1^2}+\cdots+\frac{x_{n-1}^2}{a_{n-1}^2}=x_n,\quad \frac{x_1^2}{a_1^2}+\cdots+\frac{x_s^2}{a_s^2}-\frac{x_{s+1}^2}{a_{s+1}^2}-\cdots-\frac{x_{n-1}^2}{a_{n-1}^2}=x_n,\ (0<s<n-1)$$

定义的二次曲面分别称为椭圆抛物面，双曲抛物面. 其中的 a_i 称为半轴，它们是 $\mathrm{Iso}(\mathbb{A})$ 不变量. 由方程

$$\frac{x_1^2}{a_1^2}+\cdots+\frac{x_s^2}{a_s^2}-\frac{x_{s+1}^2}{a_{s+1}^2}-\cdots-\frac{x_n^2}{a_n^2}=0\qquad\left(\frac{n}{2}\le s<n\right)$$

称为二次锥面.

习题 7.5

1. 当参数 t 取怎样的值时，二次曲面

$$x_1^2+x_2^2+x_3^2+2tx_1x_2+2tx_1x_3+2tx_2x_3-4t=0$$

是一个椭圆面？

2. 证明：在 3 维仿射空间 $\mathbb{A}$ 中，任意二次曲面都可以在适当坐标系下，由下列方程之一给出：

(1) $x_1^2+x_2^2+x_3^2=1$；(2) $x_1^2+x_2^2-x_3^2=1$，(3) $x_1^2+x_2^2-x_3^2=-1$

(4) $x_1^2-x_2^2=2x_3$；(5) $x_1^2+x_2^2=2x_3$ (6) $x_1^2+x_2^2+x_3^2=-1$

(7) $x_1^2+x_2^2-x_3^2=0$；(8) $x_1^2+x_2^2+x_3^2=0$，(9) $x_1^2+x_2^2=-1$

(10) $x_1^2+x_2^2=1$ (11) $x_1^2=2x_2$ (12) $x_1^2-x_2^2=1$

(13) $x_1^2-x_2^2=0$ (14) $x_1^2-1=0$ (15) $x_1^2+x_2^2=0$

(16) $x_1^2+1=0$ (17) $x_1^2=0$.

并指出哪些是平面对？哪些是柱形二次曲面？

3. 非零向量 $v\in V$ 称为 S_Q 的渐近向量，如果 $q(v)=0$。设 $P\in S_Q$，$v\in V$ 是 S_Q 的一个渐近向量。证明：直线 $x=p+tv$ 或者整个落在曲面 S_Q 上，或者与 S_Q 只有一个交点。

4. 找出下列二次曲面与平面相交曲线的仿射类型。

(1) $x_1^2+5x_2^2+x_3^2+2x_1x_2+2x_2x_3+6x_1x_3-2x_1+6x_2+2x_3=0$

与平面 $2x_1-x_2+x_3=0$.