序

陆少华教授嘱我为他和沈灏教授合著的《大学代数》作序，当时我正担任该书的复审编辑，虽然出于情面应允了此事，但心里还是十分的茫然。复审刚刚开始，尚未见《大学代数》之一斑，还真不知该如何去避免一个平淡的序言。随着编审过的内容越来越多，心底的迷雾越来越少，落笔也渐渐地变得有把握起来。

20世纪总被称为是一个知识爆炸的时代。作为一个科学工作者，我深深感受到知识发展的速度。如果有几年不看文献，就会产生恍入二世的感觉，在学术交流会上，对一些发言有感坠入五里云雾中，这时连解决问题的方法都听不懂。这种感觉常常来自知识的贫乏。研究的对象越来越深入，研究的手段越来越进步，越来越多的近代数学用到了应用问题的解决中。记得我在求学的时代就有人提出"用'新三高(泛函分析、拓扑和近世代数)'来代替'老三高(微积分、解析几何和高等代数)'"的说法，对这种观点的正确性的认识是在从事科研工作后才真正建立起来的，因为不懂泛函分析、拓扑学和近世代数简直无法与处于前沿的研究者沟通。产生这种"容易落伍"现象的原因不只是由于应用科学的发展，同时由于数学本身的发展，一批应用性的数学成果建立，例如控制界称为20世纪标志性成果的最大值原理，使得抽象数学与应用科学的距离越来越近。在这种情况下，再照以前的模式，大学数学从空间解析几何开始，讲述经典微积分，讲解一点微分方程和矩阵计算，必然地使得数学教学落后于研究工作的现代化了，从而有必要将大学数学教育的锋线推进到一个新的层次，有必要对大学生特别是一些前沿大学里的学生讲述一些现代数学知识。陆少华和沈灏两位教授的《大学代数》做了这方面的尝试，这本教材，从多项式讲起，涵盖了行列式、矩阵论、线性空间与线性变换、二次型，还加入了通常非数学专业不讲的群、环、域的基本内容。我以为这些知识已经能满足攻读工科的学生包

括研究生在一般研究工作中对数学工具的需要了。

在与过去一样多的教学时间内要讲解比过去多一倍以上的内容，如果还是以前一样的平铺直叙，也许没有一位教师会敢采用这样的一本教材，也没有一位学生会喜爱它。于是课程的整合是必要的。整合是近年来教育改革的一个重要概念。科学在发展，课目随之而增多；由于开展素质教育，让学员有更大的活动空间，课时又要缩减。常常出现一门课的时间内要上过去两三门课的内容。简单地将两三门课程的内容放在一起肯定是不行的。这时候，就要求对课程内容进行筛选，对课程的线索进行调整，讲述的方法要作更新，要融合进新鲜的内容，这就需要课程的整合。读了陆少华教授和沈灏教授的《大学代数》，感到他们在课程的整合上是下了功夫的。书中的讲述体系作了新的安排，内容的处理有了新的思路，例如，先讲具体的多项式，实数上的矩阵和线性空间，使得学员产生比较形象的印象，然后进入抽象的域中，这种升华也许更自然，更有利于学生的学习；又如采用矩阵将线性代数部分都串起来，使得结构紧凑，论述更方便；我特别对书中介绍 Jordan 形的方式很赞赏，作者采用现代的方法，一鼓作气给出了矩阵标准形的结论。

陆少华和沈灏是上海交通大学的知名数学教授，陆教授长期从事数学教学，讲课以思路清晰和细腻著称；沈灏教授在组合数学上颇有造诣，在国内外很有影响。两位携手共著教材，可谓珠联璧合。陆少华教授谦逊地要我提一点建议，我想也许内容可以作进一步的精选，多将一些结论留给学员，让他们有更自由的思想活动空间，这作为再版时考虑。

以上是我的一些感受，窃以为可以作为序言。

韩正之

2001 年 1 月

§2 数

一、数集的概念

由复数组成的集合常称为数集。常见数集有：

(1) 正整数全体所成的集合，用 $\mathbf{Z}^+$ 表示，即

$$\mathbf{Z}^+ = \{1,2,3,\cdots\};$$

(2) 整数全体所成的集合，用 $\mathbf{Z}$ 表示，即

$$\mathbf{Z} = \{0, \pm 1, \pm 2, \pm 3, \cdots\};$$

(3) 有理数全体所成的集合，用 $\mathbf{Q}$ 表示，即

$$\mathbf{Q} = \left\{ \frac{a}{b} \,\middle|\, a,b \in \mathbf{Z}, b \neq 0 \right\};$$

(4) 实数全体所成的集合，用 $\mathbf{R}$ 表示；

(5) 复数全体所成的集合，用 $\mathbf{C}$ 表示，即

$$\mathbf{C} = \{a + b\mathrm{i} \mid a,b \in \mathbf{R}\}。$$

这些例子中，我们用了集合的两种常用表示法，即列举法((1)与(2))和特性法((3)与(5))。

二、数环

定义 1.1 非空数集 R 叫数环，如果对任意 $a,b \in R$，必 $a+b$，$a-b$，$ab \in R$。即数环对数的加、减、乘三种运算封闭：数环包含了两个数的同时，包含了它们的和、差与积。

因为数环 R 非空，R 中有某个数 a，从而 $0 = a - a \in R$，即数环包含数零。

单个数零成一数环。数集 $\mathbf{Z},\mathbf{Q},\mathbf{R},\mathbf{C}$ 都是数环。

由于 $1 - 2 \overline{\in} \mathbf{Z}^+$，因而 $\mathbf{Z}^+$ 不是数环。

三、数域

定义 1.2 数环 P 称为数域，如果 $0,1\in P$，且对任意 $a,b\in P$，$b\neq 0$，必 $\frac{a}{b}\in P$，即数域对数的加、减、乘、除四种运算封闭。

例 1.1 $\mathbf{Q},\mathbf{R},\mathbf{C}$ 都是数域。

例 1.2 $Q(\sqrt{2})=\{a+b\sqrt{2}\mid a,b\in\mathbf{Q}\}$是数域。

证 若 $a+b\sqrt{2}$ 与 $c+d\sqrt{2}\in Q(\sqrt{2})$，则

$$
\begin{aligned}
&(a+b\sqrt{2})\pm(c+d\sqrt{2})\\
=&(a\pm c)+(b\pm d)\sqrt{2}\in Q(\sqrt{2}),\\
&(a+b\sqrt{2})(c+d\sqrt{2})\\
=&(ac+2bd)+(bc+ad)\sqrt{2}\in Q(\sqrt{2}),
\end{aligned}
$$

当 $c+d\sqrt{2}\neq 0$ 时，

$$
\begin{aligned}
\frac{a+b\sqrt{2}}{c+d\sqrt{2}}&=\frac{(a+b\sqrt{2})(c-d\sqrt{2})}{(c+d\sqrt{2})(c-d\sqrt{2})}\\
&=\frac{1}{c^2-2d^2}[(ac-2bd)+(bc-ad)\sqrt{2}]\in Q(\sqrt{2}),
\end{aligned}
$$

故 $Q(\sqrt{2})$ 是数域。

命题 1.1 若 P 是数域，则

$$P\supseteq\mathbf{Q},$$

即任意数域都包含了有理数域。

证 $0,1\in P$，从而 $n=\underbrace{1+1+\cdots+1}_{n个}\in P$。

$$-n=0-n\in P,$$

故

$$P\supseteq\mathbf{Z}。$$

任意 $a,b\in\mathbf{Z}\subseteq P$，$b\neq 0$，必 $\frac{a}{b}\in P$。由此，$P\supseteq\mathbf{Q}$。

§3 一元多项式

中学数学中常把一元多项式表达成：$f(x)=a_0+a_1x+\cdots+a_nx^n$，这里 $a_0,a_1,\cdots,a_n$ 是数，x 是字母。但究竟字母 x 代表什么，x 的一个多项式的含意是什么，都是不够明确的。

为了确切理解多项式，我们给出下面的定义。

一、多项式的定义

定义 1.3 称无穷序列 $(a_0,a_1,\cdots,a_i,\cdots)$ 为数环 R 上的一元多项式，如果

(1) $a_i\in R, i=0,1,2,\cdots$;

(2) 存在 $n\geqslant 0$，当 $i\geqslant n$ 时，$a_i=0$。

即数环 R 上的一元多项式是指数环 R 中无穷个数组成的序列，这无穷个数中仅有有限个可能不是零。

二、多项式的次数

我们采用传统的记号 $R[x]$ 表示数环 R 上一元多项式全体。用 $f(x),g(x)$ 等表示 $R[x]$ 中的多项式。

设 $f(x)=(a_0,a_1,\cdots,a_n,\cdots)$，称 $a_0,a_1,\cdots$ 为 $f(x)$ 的系数；如果 $a_n\neq 0$，而当 $i>n$ 时，$a_i=0$，则称 $f(x)$ 为 n 次多项式，记成 $\deg f(x)=n$，a_n 称为 $f(x)$ 的首项系数，一般地，a_i 称为 $f(x)$ 的 i 次项系数。

系数全为零的多项式 $(0,0,\cdots,0,\cdots)$ 称为零多项式，仍记成 0。为了方便，定义 $\deg 0=-\infty$，且规定，对任意整数 n 均有：$-\infty<n$，$n+(-\infty)=-\infty$，$(-\infty)+(-\infty)=-\infty$。

若多项式 $f(x)$ 与 $g(x)$ 的所有系数一一对应相等，则称两个多项式相等，记成 $f(x)=g(x)$。

三、多项式的运算及其性质

我们定义多项式的加法、数乘、乘法于下：

$$(a_0,a_1,\cdots)+(b_0,b_1,\cdots)=(a_0+b_0,a_1+b_1,\cdots),$$
$$k(a_0,a_1,\cdots)=(ka_0,ka_1,\cdots),$$
$$(a_0,a_1,\cdots)\times(b_0,b_1,\cdots)=(c_0,c_1,\cdots),$$

这里

$$c_t=\sum_{i=0}^{t}a_ib_{t-i}=\sum_{i+j=t}a_ib_j。$$

显然,多项式的和与积仍是多项式。

例 1.3 (1) $(1,2,4,1,0,\cdots)\times(2,0,3,0,\cdots)$

$$=(1\times 2,1\times 0+2\times 2,1\times 3+2\times 0+4\times 2,$$
$$1\times 0+2\times 3+4\times 0+1\times 2,\cdots)$$
$$=(2,4,11,8,12,3,0,\cdots);$$

(2) $2(1,-1,0,3,0\cdots)+(1,3,0\cdots)\times(2,0,4,0\cdots)$

$$=(2,-2,0,6,0\cdots)+(2,6,4,12,0,\cdots)$$
$$=(4,4,4,18,0\cdots)。$$

命题 1.2 (1) $\deg(f(x)+g(x))\leqslant\max(f(x),g(x))$,

(2) $\deg(f(x)g(x))=\deg f(x)+\deg g(x)$。

证 按定义,(1) 是显然的。证(2)于下:

若 $f(x)$与 $g(x)$中有一个为零多项式,则其积也是零多项式。此时,$\deg[f(x)\cdot g(x)]=-\infty=\deg f(x)+\deg g(x)$。

若 $f(x)\neq 0,g(x)\neq 0$。设 $\deg f(x)=n,\deg g(x)=s,f(x)$ 与 $g(x)$ 的 首项系数分别为 a_n 与 b_s。此时,积 $f(x)\cdot g(x)$ 的$(n+s)$ 次项系数为

$$\sum_{i+j=n+s}a_ib_j,$$

因为 $i>n$ 时,$a_i=0$;$j>s$ 时,$b_j=0$。故上述和式中的非零项下标必满足 $i\leqslant n,j\leqslant s$,欲使其和为 $n+s$,必 $i=n,j=s$。
于是

$$\sum_{i+j=n+s}a_ib_j=a_nb_s\neq 0。$$

当 $t>n+s$ 时,积 $f(x)\cdot g(x)$ 的 t 次项系数为

$$\sum_{i+j=t} a_i b_j。$$

欲使和式中单项 $a_i b_j \neq 0$，必 $i \leqslant n, j \leqslant s$，但此时 $i+j \leqslant n+s<t$，故当 $i+j=t$ 时，和式中单项全为零，于是积 $f(x) \cdot g(x)$ 的 $t(>n+s)$ 次项系数为 0。

由上分析，得

$$\deg (f(x) \cdot g(x))=n+s=\deg f(x)+\deg g(x)。$$

从证明过程中可知，若 $f(x)$ 与 $g(x)$ 都不是零多项式，它们的首项系数分别为 a_n 与 b_s，则积的首项系数为 $a_n \cdot b_s$。

推论 非零多项式积的首项系数等于首项系数之积。

与数的运算相同，多项式的运算满足下面规律：

(1) 加法交换律 $f(x)+g(x)=g(x)+f(x)$；

(2) 加法结合律 $[f(x)+g(x)]+h(x)$
$=f(x)+[g(x)+h(x)]$；

(3) 乘法交换律 $f(x) \cdot g(x)=g(x) \cdot f(x)$；

(4) 乘法结合律 $[f(x) \cdot g(x)] h(x)$
$=f(x)[g(x) \cdot h(x)]$；

(5) 数乘法则 $(k+l) f(x)=k f(x)+l f(x)$，
$(k l) f(x)=k[l f(x)]$，
$k[f(x)+g(x)]=k f(x)+k g(x)$，
$k[f(x) \cdot g(x)]=[k f(x)] \cdot g(x)$
$=f(x) \cdot[k g(x)]$；

(6) 乘法对于加法的分配律 $f(x)[g(x)+h(x)]$
$=f(x) g(x)+f(x) h(x)$；

(7) 乘法消去律 若 $f(x) g(x)=f(x) h(x)$，且 $f(x) \neq 0$，则

$$g(x)=h(x)。$$

(1)、(2)、(3)、(5)和(6)的证明可由数的运算满足相应规律立即推出。

我们证(4)和(7)于下：

(4) 的证明 设 $f(x)=(a_0,a_1,\cdots)$,

$$g(x)=(b_0,b_1,\cdots),h(x)=(c_0,c_1,\cdots)。$$

$f(x)[g(x)h(x)]$ 的 t 次项系数 $=\sum\limits_{i+l=t}a_i(\sum\limits_{j+k=l}b_jc_k)=\sum\limits_{i+j+k=t}a_i(b_jc_k)=\sum\limits_{i+j+k=t}(a_ib_j)c_k=\sum\limits_{u+k=t}(\sum\limits_{i+j=u}a_ib_j)c_k=[f(x)\cdot g(x)]h(x)$ 的 t 次项系数。

故 $f(x)[g(x)h(x)]=[(f(x)g(x)]h(x)$。

(7) 的证明 由 $f(x)g(x)=f(x)h(x)$ 得

$$f(x)[g(x)-h(x)]=0。$$

若 $g(x)\neq h(x)$,即 $g(x)-h(x)\neq 0$,由命题 1.2

$$\deg f(x)+\deg[g(x)-h(x)]=\deg 0=-\infty,$$

但上式左边是两个整数之和,不等于 $-\infty$,故必 $g(x)=h(x)$。

四、特殊多项式与多项式的传统表示

零多项式在多项式的运算中起着数零在数的运算中的作用:

$$0+f(x)=f(x),$$
$$0\cdot f(x)=0。$$

我们发现,若 $1\in R$,另有两个多项式在运算中起着特殊的作用:第一个是$(1,0,0,\cdots)$,称为单位多项式,记成 **1**,它在多项式的运算中起着数 1 在数的运算中的类似作用:

$$\mathbf{1}\cdot f(x)=f(x)。$$

第二个是$(0,1,0,0\cdots)$,记它为 x,对 n 用归纳法,可证得:

$$x^n=\underbrace{x\cdot x\cdot\cdots\cdot x}_{n个}=(\underbrace{0,\cdots,0}_{n个},1,0,0\cdots)。$$

利用单位多项式 **1** 与特殊多项式 x 可表多项式$(a_0,a_1,\cdots,a_n,0,\cdots)$为

$$(a_0,0,\cdots)+(0,a_1,0,\cdots)+\cdots+(0,\cdots,0,a_n,0\cdots)$$

$$= a_0(1,0,0,\cdots) + a_1(0,1,0,0,\cdots) + \cdots + a_n(\underbrace{0,\cdots,0}_{n\text{个}},1,0,\cdots)$$

$$= a_0\mathbf{1} + a_1x + a_2x^2 + \cdots + a_nx^n。$$

这使我们回到了熟悉的形式，并明确了字母 x 的含意：x 起着定位作用。

注 在多项式的概念中，其实 R 可以不是数环，仅要求它的元素可作加、乘运算且满足一定规律。例如我们可用 $R[x]$ 代替 R，就得到 R 上的二元多项式。

§4 最大公因式

下面设 P 是数域。

一、带余除法

定理 1.1 设 $f(x)$，$g(x)$ 都是数域 P 上的一元多项式，$g(x)\neq 0$，则存在 $P[x]$ 中惟一的 $q(x)$ 和 $r(x)$ 使

$$f(x) = q(x)g(x) + r(x),$$
$$\deg r(x) < \deg g(x)。$$

$q(x)$ 与 $r(x)$ 分别称为 $g(x)$ 除 $f(x)$ 所得的商和余式。

证 记 $T = \{f(x) - t(x)g(x) \mid t(x) \in P[x]\}$。设 $r(x)$ 是 T 中次数最小的多项式，$r(x) = f(x) - q(x)g(x)$。

若 $\deg r(x) \geqslant \deg g(x)$。设 $r(x)$ 的首项为 cx^m，$g(x)$ 的首项为 $bx^s (m \geqslant s)$，则

$r(x)$ 与 $\frac{c}{b}x^{m-s}g(x)$ 有相同首项 cx^m。

从而 $r_1(x) = r(x) - \frac{c}{b}x^{m-s}g(x)$ 的次数小于 $r(x)$ 的次数，又

$$r_1(x) = [f(x) - q(x)g(x)] - \frac{c}{b}x^{m-s}g(x)$$
$$= f(x) - [q(x) + \frac{c}{b}x^{m-s}]g(x) \in T,$$

于是 $r_1(x)$ 是 T 中次数小于 $r(x)$ 的多项式，这与假设矛盾（$r(x)$ 是 T

中次数最小的多项式），故必

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

于是

$$\begin{cases} f(x) = q(x)g(x) + r(x), \\ \deg r(x) < \deg g(x)。\end{cases}$$

如果另有

$$f(x) = q_1(x)g(x) + r_1(x),$$
$$\deg r_1(x) < \deg g(x),$$

则

$$[q(x) - q_1(x)]g(x) = r_1(x) - r(x)。$$

因为 $g(x) \neq 0$，故 $q(x) - q_1(x)$ 与 $r_1(x) - r(x)$ 同时为 0 或同时不为 0。若同时不为 0，由命题 1.2．$\deg[r_1(x) - r(x)] \geqslant \deg g(x)$，此不可能，故必同时为 0，即

$$q(x) = q_1(x), r_1(x) = r(x)。$$

本定理得证。

例如，非零多项式 $g(x) = x^2 + 3x + 1$ 除多项式 $f(x) = x^3 + 4x^2 + 5x - 3$ 的商与余分别为

$$\begin{aligned} q(x) &= x + 1, r(x) = x - 4, \\ f(x) &= x^3 + 4x^2 + 5x - 3 \\ &= (x + 1)(x^2 + 3x + 1) + (x - 4) \\ &= q(x)g(x) + r(x)。\end{aligned}$$

除法算式：

$$\begin{array}{r|l} & \quad (1, \quad 1) \\ \hline (1,3,1) & (1, 4, 5, -3) \\ & \underline{1, 3, 1} \\ & \quad 1, 4, -3 \\ & \quad \underline{1, 3, \quad 1} \\ & \quad (1, -4) \end{array} \qquad \begin{array}{r|l} & \quad q\text{ 的系数} \\ \hline g\text{ 的系数} & f\text{ 的系数} \\ & \cdot \; \cdot \; \cdot \\ & \underline{\cdot \; \cdot \; \cdot} \\ & \quad \cdot \; \cdot \; \cdot \\ & \quad \underline{\cdot \; \cdot \; \cdot} \\ & \quad r\text{ 的系数} \end{array}$$

二、整除性

定义 1.4 设 $g(x)\neq 0$，若 $g(x)$ 除 $f(x)$ 所得余式为 0，则称 $g(x)$ 能整除 $f(x)$，或称 $g(x)$ 是 $f(x)$ 的因式，记成 $g(x)\mid f(x)$。

推论 1 $g(x)\mid f(x)\Leftrightarrow$ 存在 $q(x)$ 使 $f(x)=q(x)g(x)$。

推论 2 若 $f(x)\neq 0$，且 $g(x)\mid f(x)$，则 $\deg g(x)\leqslant \deg f(x)$。

性质 (1) 若 $g(x)\mid f(x)$ 且 $f(x)\mid g(x)$，则存在非零常数 c 使 $f(x)=cg(x)$；

(2) 若 $h(x)\mid g(x)$，$g(x)\mid f(x)$，则 $h(x)\mid f(x)$；

(3) 若 $h(x)\mid f_i(x)$，$i=1,\cdots,n$，则对任意多项式 $g_1(x),\cdots,g_n(x)$ 有 $h(x)\mid \sum\limits_{i=1}^{n}g_i(x)f_i(x)$。

证 (1) 由假设，存在 $h(x)$ 与 $e(x)$ 使

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

从而

$$f(x)=h(x)e(x)f(x),$$

两边消去非零多项式 $f(x)$，得

$$h(x)e(x)=1。$$

由命题 1.2 导出 $\deg h(x)=0$，即 $h(x)$ 为非零常数 c，进而 $f(x)=cg(x)$。

(2) 由假设

$$g(x)=e(x)h(x),f(x)=q(x)g(x),$$

于是

$$f(x)=q(x)e(x)g(x),$$

故

$$g(x)\mid f(x)。$$

(3) 由假设，$f_i(x)=d_i(x)\cdot h(x)$，$i=1,\cdots,n$。于是

$$\sum_{i=1}^{n} g_i(x) f_i(x) = \sum_{i=1}^{n} g_i(x) d_i(x) h(x) = \left(\sum_{i=1}^{n} g_i(x) d_i(x) \right) h(x),$$

故

$$h(x) \Big| \sum_{i=1}^{n} g_i(x) f_i(x)。$$

例 1.4 求证：

$$(x^2 + x + 1) \Big| (x^{3m} + x^{3n+1} + x^{3p+2}),$$

这里 m,n,p 是三个正整数。

证 $x^{3m} + x^{3n+1} + x^{3p+2}$

$$\begin{aligned} &= (x^{3m} - 1) + (x^{3n+1} - x) + (x^{3p+2} - x^2) + (1 + x + x^2) \\ &= [(x^3)^m - 1] + x[(x^3)^n - 1] + x^2[(x^3)^p - 1] + \\ &\quad (1 + x + x^2), \end{aligned}$$

因为

$$(x^2 + x + 1) \Big| (x^3 - 1) = (x - 1)(x^2 + x + 1),$$

$$\begin{aligned} &(x^3 - 1) \Big| ((x^3)^t - 1) \\ &= (x^3 - 1)[(x^3)^{t-1} + (x^3)^{t-2} + \cdots + x^3 + 1], \end{aligned}$$

故

$$(x^2 + x + 1) \Big| (x^{3m} + x^{3n+1} + x^{3p+2})。$$

三、最大公因式

定义 1.5 多项式 $d(x)$ 称为多项式 $f(x)$ 与 $g(x)$ 的最大公因式，记成 $(f(x),g(x))$，如果

(1) $d(x) \mid f(x)$，$d(x) \mid g(x)$，即 $d(x)$ 是 $f(x)$ 与 $g(x)$ 的公因式；

(2) 若 $c(x) \mid f(x)$，且 $c(x) \mid g(x)$，则 $c(x) \mid d(x)$。即 $d(x)$ 是 $f(x)$ 与 $g(x)$ 的任一公因式的倍式；

(3) $d(x)$ 的首项系数为 1（首项系数为 1 的多项式简称首 1 多项式）。

定理 1.2 $P[x]$ 中任意两个不全为零的多项式 $f(x)$ 与 $g(x)$ 在

$P[x]$中必有且仅有一个最大公因式 $d(x)$，且存在 $u(x),v(x)\in P[x]$ 使

$$d(x)=u(x)f(x)+v(x)g(x)。$$

证 记 $T=\{k(x)f(x)+l(x)g(x)\mid k(x),l(x)\in P[x]\}$。因为

$$\begin{aligned}f(x)&=1\cdot f(x)+0\cdot g(x)\quad\in T,\\ g(x)&=0\cdot f(x)+1\cdot g(x)\quad\in T,\end{aligned}$$

且 $f(x),g(x)$不全为 0，故 T 中有非零多项式，进而 T 中有首 1 多项式。设 $d(x)$为 T 中次数最小的首 1 多项式，$d(x)\in T$，即存在 $u(x)$，$v(x)\in P[x]$ 使

$$d(x)=u(x)f(x)+v(x)g(x)。$$

下面证明 $d(x)$是 $f(x)$与 $g(x)$的最大公因式。

若用 $d(x)$除 $f(x)$得商 $q(x)$与余式 $r(x)$：

$$\begin{cases}f(x)=q(x)d(x)+r(x),\\ \deg r(x)<\deg d(x),\end{cases}$$

$$\begin{aligned}r(x)&=f(x)-q(x)d(x)\\ &=f(x)-q(x)[u(x)f(x)+v(x)g(x)]\\ &=[1-q(x)u(x)]f(x)+[-q(x)v(x)]g(x)\in T,\end{aligned}$$

若 $r(x)\neq 0$，则 T 中存在次数小于 $d(x)$ 的非零多项式。进而 T 中存在次数小于 $d(x)$ 的首 1 多项式，这与 $d(x)$ 是 T 中次数最小的首 1 多项式的假设矛盾，故必 $r(x)=0$，即

$$d(x)\mid f(x),$$

同理

$$d(x)\mid g(x)。$$

若 $c(x)\mid f(x),c(x)\mid g(x)$，则

$$c(x)\mid u(x)f(x)+v(x)g(x)=d(x)。$$

因而 $d(x)$ 是 $f(x)$ 与 $g(x)$ 的最大公因式。

又，若 $d_1(x)$ 也是 $f(x)$ 与 $g(x)$ 的最大公因式。把 $d(x)$ 看成是 $f(x)$ 与 $g(x)$ 的最大公因式，$d_1(x)$ 仅看成是 $f(x)$ 与 $g(x)$ 的公因式，则 $d_1(x)\mid d(x)$。同理，$d(x)\mid d_1(x)$。从而 $d_1(x)=cd(x)$，比较两边首项系数，得 $c=1$，即 $d_1(x)=d(x)$。

定理 1.2 告诉我们：任意两个不全为零的多项式的最大公因式存在且唯一。

性质 (1) $((f(x),g(x)),h(x))=(f(x),(g(x),h(x)))$；

(2) 若 $u(x)$ 为首 1 多项式，则 $(u(x)f(x),u(x)g(x))=u(x)\cdot(f(x),g(x))$；

(3) 若 $(f(x),g(x))=1$，$(f(x),h(x))=1$，则 $(f(x),g(x)h(x))=1$。

证 (1) 等式左边 $\mid(f(x),g(x)),h(x)$. 从而是 $f(x),g(x),h(x)$ 的公因式，于是也是 $f(x)$ 与 $(g(x),h(x))$ 的公因式。由最大公因式定义，等式左边能整除等式右边。同理，等式右边能整除等式左边。而两者都是首 1 多项式，故相等。

(2) 设 $(f(x),g(x))=d(x)$。于是 $u(x)d(x)\mid u(x)f(x),u(x)g(x)$。因而 $u(x)d(x)\mid(u(x)f(x),u(x)g(x))$，可设后者为 $u(x)d(x)e(x)$。此时 $u(x)d(x)e(x)\mid u(x)f(x)$，故 $d(x)e(x)\mid f(x)$。同理 $d(x)e(x)\mid g(x)$，进而 $d(x)e(x)\mid(f(x),g(x))=d(x)$。故得 $e(x)=1$，即 $(u(x)f(x),u(x)g(x))=u(x)d(x)e(x)=u(x)d(x)=u(x)\,(f(x),g(x))$。

(3) $1=(f(x),g(x))=(f(x),g(x)(f(x),h(x)))=(f(x),(g(x)f(x),g(x)h(x)))=((f(x),g(x)f(x)),g(x)h(x))=(f(x),g(x)h(x))$。

定义 1.6 若 $(f(x),g(x))=1$，则称多项式 $f(x)$ 与 $g(x)$ 互素。

推论 若存在多项式 $u(x)$ 与 $v(x)$ 使

$$u(x)f(x)+v(x)g(x)=1,$$

则 $f(x)$ 与 $g(x)$ 互素。

证 1 是 $f(x)$ 与 $g(x)$ 的当然首 1 公因式。又若 $c(x)|f(x)$，$g(x)$，则 $c(x)|u(x)f(x)+v(x)g(x)$，即 $c(x)|1$，故 1 是 $f(x)$ 与 $g(x)$ 的最大公因式，即 $(f(x),g(x))=1$。

例 1.5 求证：$(x^3-1,x^2-1)=x-1$。

证 因为

$$(x-1)|(x^2-1),$$
$$(x-1)|(x^3-1),$$

又 $x-1=-x(x^2-1)+(x^3-1)$ 是 (x^2-1) 与 (x^3-1) 的一个组合，故若 $c(x)$ 满足

$$c(x)|(x^2-1),c(x)|(x^3-1),$$

则必

$$c(x)|(x-1)。$$

所以 $(x-1)$ 是 (x^2-1) 与 (x^3-1) 的最大公因式。

例 1.6 若 $f(x),g(x)$ 是两个不全为零的多项式，求证：

$$(f(x),g(x))=(f(x)+g(x),f(x)-g(x))。$$

证 因为

$$(f(x),g(x))|f(x),g(x),$$

故

$$(f(x),g(x))|(f(x)+g(x)) 与 (f(x)-g(x)),$$

从而

$$(f(x),g(x))|(f(x)+g(x),f(x)-g(x))。$$

又

$$(f(x)+g(x),f(x)-g(x))$$
$$\left|\left(\frac{1}{2}[f(x)+g(x)]+\frac{1}{2}[f(x)-g(x)]\right)=f(x),\right.$$
$$(f(x)+g(x),f(x)-g(x))$$
$$\left|\left(\frac{1}{2}[f(x)+g(x)]-\frac{1}{2}[f(x)-g(x)]\right)=g(x),\right.$$

故

$$(f(x)+g(x),f(x)-g(x))\mid(f(x),g(x)),$$

于是必

$$(f(x)+g(x),f(x)-g(x))=(f(x),g(x))。$$

例 1.7 若 $a_0\neq 0$,求证:多项式 $f(x)$ 与 $g(x)$ 互素。这里

$$f(x)=a_0+a_1x+\cdots+a_nx^n,$$
$$g(x)=a_1+a_2x+\cdots+a_nx^{n-1}。$$

证 因为

$$\frac{1}{a_0}f(x)-\frac{1}{a_0}xg(x)=\frac{1}{a_0}[f(x)-xg(x)]=\frac{1}{a_0}\cdot a_0=1,$$

由定义 1.6 的推论,$f(x)$与 $g(x)$互素。

四、$n(\geqslant 2)$个多项式的最大公因式

定义 1.7 $d(x)$称为多项式 $f_1(x),f_2(x),\cdots,f_n(x)$的最大公因式,记成 $d(x)=(f_1(x),f_2(x),\cdots,f_n(x))$。如果

(1) $d(x)\mid f_i(x),i=1,2,\cdots,n$;

(2) 若 $c(x)\mid f_i(x),i=1,2,\cdots,n$,则 $c(x)\mid d(x)$;

(3) $d(x)$为首 1 多项式。

定理 1.3 $n(\geqslant 2)$ 个不全为零的多项式 $f_1(x),\cdots,f_n(x)$ 的最大公因式存在且惟一。

证 记 $T=\left\{\sum\limits_{i=1}^{n}k_i(x)f_i(x)\,\middle|\,k_i(x)\in P[x],i=1,\cdots,n\right\}$。因为 $f_i(x)\in T,i=1,\cdots,n$,故 T 中有非零多项式,从而有首 1 多项式。设 $d(x)$ 是 T 中次数最小的首 1 多项式,类似定理 1.2 的证明,可得 $d(x)$ 是 $f_1(x),\cdots,f_n(x)$ 惟一的最大公因式。

我们指出,$d(x)\in T$,因而存在多项式 $u_1(x),\cdots,u_n(x)$,使 $d(x)=(f_1(x),\cdots,f_n(x))=\sum\limits_{i=1}^{n}u_i(x)f_i(x)$。

§5 因式分解惟一性定理

一、不可约多项式

定义 1.8 数域 P 上一元多项式 $p(x)$称为 $P[x]$中不可约多项式，如果

(1) $\deg p(x) \geqslant 1$；

(2) 若 $d(x) \in P[x]$，$d(x) \mid p(x)$，必 $\deg d(x) = 0$ 或 $\deg d(x) = \deg p(x)$。

显然，一次多项式都是不可约多项式。若 $p(x)$是不可约多项式，$c \neq 0$，则 $cp(x)$ 也是不可约多项式。

多项式是否可约常依赖于它所在的数域，如多项式 $x^2 - 2$ 是 $Q[x]$ 中的不可约多项式，但它不是 $R[x]$中的不可约多项式。

性质 (1) 若 $d(x)$ 是不可约多项式 $p(x)$ 的因式，则 $d(x) = c$ 或 $d(x) = cp(x)$，这里 c 是非零常数；

(2) 若 $p(x)$ 是不可约多项式，且 $p(x) \mid f(x)g(x)$，则 $p(x) \mid f(x)$ 或 $p(x) \mid g(x)$。

证 (1) 由定义 1.8，若 $\deg d(x) = 0$，则 $d(x)$ 为非零常数；若 $\deg d(x) = \deg p(x)$，则 $d(x) = c \cdot p(x)(c \neq 0)$。

(2) 若 $p(x) \nmid f(x)$，且 $p(x) \nmid g(x)$，则 $(p(x), f(x)) = 1$，$(p(x), g(x)) = 1$。由最大公因式的性质(3)，$(p(x), f(x) \cdot g(x)) = 1$，但 $p(x)$ 是 $p(x)$ 与 $f(x) \cdot g(x)$ 的公因式，因而 $p(x) \mid 1$，这不可能。故必 $p(x) \mid f(x)$ 或 $p(x) \mid g(x)$。

对 n 用归纳法，即得下面推论。

推论 若不可约多项式 $p(x) \mid f_1(x)f_2(x)\cdots f_n(x)$，则存在 i，$1 \leqslant i \leqslant n$，使 $p(x) \mid f_i(x)$。

二、因式分解惟一性定理

定理 1.4 次数大于 0 的多项式 $f(x)$必可惟一表成有限个不可

约多项式 $p_1(x),\cdots,p_n(x)$ 的积：

$$f(x)=p_1(x)\cdots p_n(x)。$$

此处"惟一"是指，若 $f(x)$另有分解式 $f(x)=q_1(x)\cdots q_s(x)$，这里 $q_1(x),\cdots,q_s(x)$ 是不可约多项式，则 $s=n$，且适当调整 $q_1(x),\cdots,q_s(x)$ 的次序，可使 $p_i(x)=c_iq_i(x),i=1,\cdots,n(c_i$ 为非零常数)。

证 对 $f(x)$的次数 n 用归纳法：

$n=1$ 时，因为 1 次多项式都是不可约的，故结论成立。

若对次数小于 n 的多项式结论成立。我们观察 n 次多项式 $f(x)$：

如果 $f(x)$是不可约多项式，则结论成立。

如果 $f(x)$不是不可约的，即 $f(x)$有因式 $d(x)$，$0<\deg d(x)<\deg f(x)$。此时有多项式 $h(x)$使

$$f(x)=d(x)h(x),$$

由命题 1.2，$\deg h(x)=\deg f(x)-\deg d(x)$，从而

$$0<\deg h(x)<n。$$

由归纳法假设，$d(x)$与 $h(x)$都是有限个不可约多项式的积，故 $f(x)$也是有限个不可约多项式的积。

若 $f(x)=p_1(x)\cdots p_n(x)=q_1(x)\cdots q_s(x)$，这里 $p_i(x)$ 与 $q_j(x)$ $(i=1,\cdots,n;j=1,\cdots,s)$ 都是不可约多项式。此时，不可约多项式 $p_1(x)|q_1(x)\cdots q_s(x)$。由不可约多项式的性质可知，$p_1(x)$能整除 $q_1(x),\cdots,q_s(x)$中的某一个，适当调整次序可使

$$p_1(x)|q_1(x)。$$

由于 $q_1(x)$也是不可约多项式，必

$$p_1(x)=c_1q_1(x),c_1\neq 0,$$

由于 $q_1(x)\neq 0$，在 $f(x)=p_1(x)\cdots p_n(x)=q_1(x)\cdots q_s(x)$ 中消去 $q_1(x)$得

$$f_1(x)=(c_1p_2(x))\cdot p_3(x)\cdots p_n(x)=q_2(x)\cdots q_s(x),$$

因为 $0 < \deg f_1(x) < n$，由归纳法假设

$$n-1=s-1,$$
$$c_1p_2(x)=dq_2(x), d\neq 0,$$
$$p_i(x)=c_iq_i(x), c_i\neq 0, i=3,\cdots,n。$$

记 $\dfrac{d}{c_1}=c_2$，即得。对 n 次多项式 $f(x)$结论成立。

由归纳法，本定理得证。

若把不可约多项式首 1 化(除以首项系数)，且把相同的归在一起，可得标准分解式：

$$f(x)=cp_1^{\alpha_1}(x)\cdots p_t^{\alpha_t}(x)。$$

命题 1.3 如果 $f(x)=ap_1^{\alpha_1}(x)\cdots p_t^{\alpha_t}(x), g(x)=bp_1^{\beta_1}(x)\cdots p_t^{\beta_t}(x)$，这里 $a\neq 0, b\neq 0, \alpha_i\geqslant 0, \beta_i\geqslant 0, i=1,\cdots,t$。$p_1(x),\cdots,p_t(x)$ 是两两不同的首 1 不可约多项式，则

(1) $g(x)\mid f(x)\Leftrightarrow\beta_i\leqslant\alpha_i, i=1,2,\cdots,t$；

(2) $(g(x),f(x))=p_1^{r_1}(x)\cdots p_t^{r_t}(x)$。

这里 $r_i=\min(\alpha_i,\beta_i), i=1,2,\cdots,t$。

证 (1) $g(x)\mid f(x)\Leftrightarrow f(x)=g(x)h(x)\Leftrightarrow g(x)$ 标准分解式中所有首 1 不可约多项式是 $f(x)$ 标准分解式中首 1 不可约多项式的部分 $\Leftrightarrow\beta_i\leqslant\alpha_i, i=1,\cdots,t$。

(2) 记右边为 $d(x)$。由(1)得 $d(x)\mid f(x), g(x)$。

若 $c(x)\mid f(x), g(x)$，则 $c(x)=kp_1^{l_1}(x)\cdots p_t^{l_t}(x)$，且 $l_i\leqslant\alpha_i,\beta_i$，从而 $l_i\leqslant\min(\alpha_i,\beta_i)=r_i, i=1,\cdots,t$。故 $c(x)\mid d(x)$，即 $d(x)=(g(x),f(x))$。

例 1.8 求多项式

$$f(x)=(x-1)^{10}(x+2)^3(x+3)^5$$

的首 1 因式个数。

解 若 $d(x)$为 $f(x)$的一个首 1 因式，则

$$d(x)=(x-1)^{\alpha}(x+2)^{\beta}(x+3)^{r},$$

这里

$$0 \leqslant \alpha \leqslant 10,\ 0 \leqslant \beta \leqslant 3,\ 0 \leqslant r \leqslant 5,$$

α 有 $(10+1)$ 种取法，β 有 $(3+1)$ 种取法，r 有 $(5+1)$ 种取法，因而 $d(x)$ 的个数为

$$(10+1)(3+1)(5+1) = 11 \times 4 \times 6 = 264,$$

即 $f(x)$ 的首 1 因式个数为 264。

定义 1.9 多项式 $l(x)$ 称为多项式 $f_1(x),\cdots,f_n(x)$ 的最小公倍式，如果

(1) $f_i(x) \mid l(x), i=1,\cdots,n$;

(2) 若 $f_i(x) \mid h(x), i=1,\cdots,n$，则 $l(x) \mid h(x)$;

(3) $l(x)$ 为首 1 多项式。

定理 1.5 有限个非零多项式的最小公倍式存在且惟一。

代替证明，我们指出，若 $f_i(x) = a_i p_1^{\alpha_{i1}}(x)\cdots p_t^{\alpha_{it}}(x)(i=1,\cdots,n)$，则 $f_1(x),\cdots,f_n(x)$ 的最小公倍式为

$$l(x) = p_1^{\beta_1}(x)\cdots p_t^{\beta_t}(x),$$

这里 $\beta_j = \max\{\alpha_{1j},\alpha_{2j},\cdots,\alpha_{nj}\}$。

例 1.9 求多项式 $f(x)$ 与 $g(x)$ 的最大公因式与最小公倍式。这里：

$$f(x) = (x^4-1)^{10},$$
$$g(x) = (x^2-3x+2)^7(x^2+3x+2)^{15}。$$

解

$$f(x) = (x-1)^{10}(x+1)^{10}(x^2+1)^{10},$$
$$g(x) = (x-1)^7(x-2)^7(x+1)^{15}(x+2)^{15}。$$

于是

$$(f(x),g(x)) = (x-1)^7(x+1)^{10}。$$

$f(x)$ 与 $g(x)$ 的最小公倍式为

$$(x-1)^{10}(x+1)^{15}(x-2)^7(x+2)^{15}(x^2+1)^{10}。$$

三、重因式

定义 1.10 不可约多项式 $p(x)$ 称为 $f(x)$ 的 k 重因式，如果 $p^k(x)\mid f(x)$，而 $p^{k+1}(x)\nmid f(x)$。

若 $k=0$，$p(x)$ 不是 $f(x)$ 的因式。若 $k=1$，称 $p(x)$ 为 $f(x)$ 的单因式；若 $k>1$，称 $p(x)$ 为 $f(x)$ 的重因式。

下面我们以导数为工具研究多项式的重因式。

定义 1.11 若多项式 $f(x)=a_0+a_1x+a_2x^2+\cdots+a_nx^n$，称多项式 $a_1+2a_2x+\cdots+na_nx^{n-1}$ 为 $f(x)$ 的导数，且记成 $f'(x)$。

直接验证，可得下列求导法则：

求导法则 (1) $(f(x)+g(x))'=f'(x)+g'(x)$；

(2) $(kf(x))'=kf'(x)$；

(3) $(f(x)\cdot g(x))'=f'(x)g(x)+f(x)g'(x)$；

(4) $(f^s(x))'=sf^{s-1}(x)\cdot f'(x)$。

引理 若不可约多项式 $p(x)$ 是 $f(x)$ 的因式，则 $p(x)$ 是 $f(x)$ 的重因式 $\Leftrightarrow p(x)\mid f'(x)$。

证 设 $p(x)$ 是 $f(x)$ 的 k 重因式 $(k\geqslant 1)$：

$f(x)=p^k(x)g(x)$，$p(x)\nmid g(x)$。

$$
\begin{aligned}
f'(x)&=kp^{k-1}(x)p'(x)g(x)+p^k(x)g'(x)\\
&=p^{k-1}(x)\cdot[kp'(x)g(x)+p(x)g'(x)].
\end{aligned}
$$

因为 $p(x)\nmid p'(x)$，$p(x)\nmid g(x)$，所以 $p(x)$ 不是 $kp'(x)g(x)$ 的因子，但它是 $p(x)g'(x)$ 的因子。于是它不是 $kp'(x)g(x)+p(x)g'(x)$ 的因子。因而 $p(x)$ 是 $f'(x)$ 的 $(k-1)$ 重因式。

因而 $p(x)\mid f'(x)\Leftrightarrow k-1>0\Leftrightarrow k>1\Leftrightarrow p(x)$ 是 $f(x)$ 的重因式。

推论 不可约多项式 $p(x)$ 是 $f(x)$ 重因式 $\Leftrightarrow p(x)\mid(f(x),f'(x))$。

定理 1.6 多项式 $f(x)$ 没有重因式 $\Leftrightarrow (f(x),f'(x))=1$。

证 $f(x)$ 有重因式 $\Leftrightarrow \deg(f(x),f'(x))>0$，故

$f(x)$ 没有重因式 $\Leftrightarrow \deg(f(x),f'(x))=0\Leftrightarrow(f(x),f'(x))=1$。

例 1.10 求证：多项式 $f(x)=x^n-1$ 无重因式。

证 因为 $f'(x) = nx^{n-1}$，
于是

$$\frac{x}{n}f'(x) - f(x) = x^n - (x^n - 1) = 1,$$

由定义 1.6 的推论，知

$$(f(x), f'(x)) = 1,$$

故 $f(x)$没有重因式。

§6 复数域与实数域上的一元多项式

一、多项式赋值

定义 1.12 在多项式 $f(x) = a_0 + a_1x + \cdots + a_nx^n$ 中，用数 c 代替字母 x，得到了一个数的算式，运算结果仍是一个数，称它为多项式 $f(x)$在 c 处的值，记成 $f(c)$，即

$$f(c) = a_0 + a_1c + \cdots + a_nc^n。$$

命题 1.4 若 $h(x) = f(x) + g(x), l(x) = f(x)g(x)$，则 $h(c) = f(c) + g(c); l(c) = f(c)g(c)$。

证 若 $f(x) = a_0 + a_1x + \cdots, g(x) = b_0 + b_1x + \cdots$，则

$$h(x) = (a_0 + b_0) + (a_1 + b_1)x + \cdots,$$

从而

$$\begin{aligned} h(c) &= (a_0 + b_0) + (a_1 + b_1)c + \cdots \\ &= a_0 + a_1c + \cdots + b_0 + b_1c + \cdots \\ &= f(c) + g(c), \end{aligned}$$

$$l(x) = d_0 + d_1x + \cdots, d_t = \sum_{i+j=t} a_ib_j。$$

而

$$f(c)g(c) = (\sum_i a_ic^i)(\sum_j b_jc^j) = \sum_t (\sum_{i+j=t} a_ib_j)c^t$$

$$= \sum_t d_t c^t = l(c)。$$

二、多项式的零点

定义 1.13 c 是一个数，$f(x)$为一多项式，若 $f(c) = 0$，则称 c 为 $f(x)$的一个零点或根。

命题 1.5 $(x-c) \mid f(x) \Leftrightarrow f(c) = 0$。

证 由带余除法，存在 $q(x)$和 b 使

$$f(x) = q(x)(x-c) + b。$$

由命题 1.4 得，$f(c) = q(c)(c-c) + b = b$，即 $x-c$ 除 $f(x)$ 所得的余式为 $f(c)$，从而

$$(x-c) \mid f(x) \Leftrightarrow 余式\ f(c) = 0。$$

推论 若 $a_1, a_2, \cdots, a_n$ 是多项式 $f(x)$的互异零点，则

$$(x-a_1)(x-a_2)\cdots(x-a_n) \mid f(x)。$$

证 对 n 用归纳法（$n = 1$ 时，即命题 1.5）。

若 $n-1$ 时结论成立。设 $a_1, a_2, \cdots, a_n$ 是 $f(x)$的 n 个互异零点，特别 $a_1, \cdots, a_{n-1}$是 $f(x)$的$(n-1)$个互异零点，由归纳法假设，$(x-a_1)\cdots(x-a_{n-1}) \mid f(x)$，即

$$f(x) = (x-a_1)\cdots(x-a_{n-1})f_1(x),$$

从而

$$f(a_n) = (a_n - a_1)\cdots(a_n - a_{n-1})f_1(a_n),$$

由 $f(a_n) = 0$ 得 $f_1(a_n) = 0$。由命题 1.5 知，$(x-a_n) \mid f_1(x)$，即 $f_1(x) = f_2(x)(x-a_n)$，故 $f(x) = (x-a_1)\cdots(x-a_n)f_2(x)$。于是 n 时结论仍成立。由归纳法，本推论得证。

代数基本定理 次数大于零的复系数多项式至少有一个复数根（证略）。

例 1.11 若 $(x+1)^2 \mid (x^3 + Ax^2 + B)$，求 A 与 B。

解 设$f(x) = x^3 + Ax^2 + B$，$(x+1)$是$f(x)$的重因式，故$(x+$

1) 是 $f(x)$ 与 $f'(x)$ 的公因式，从而

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

即

$$\begin{cases}-1+A+B=0,\\3-2A=0。\end{cases}$$

解之，得

$$A=\frac{3}{2},B=-\frac{1}{2}。$$

例 1.12 若二次多项式 $f(x)=ax^2+bx+c(a\neq 0)$ 有重根。(数 α 称为多项式 $f(x)$ 的重根，如果 $(x-\alpha)^2|f(x)$)，求系数 a,b,c 满足的关系式。

解 因为 $(x-\alpha)$ 是 $f(x)$ 的重因式，所以 $(x-\alpha)$ 是 $f(x)$ 与 $f'(x)=2ax+b$ 的公因式，因而

$$\begin{cases}f(\alpha)=a\alpha^2+b\alpha+c=0,\\f'(\alpha)=2a\alpha+b=0,\end{cases}$$

消去 α，得 $b^2-4ac=0$。

例 1.13 求二次多项式 $f(x)=ax^2+bx+c$ 使

$$f(n)=\frac{1}{n+1},\ n=1,2,3。$$

解 1 由 $\begin{cases}f(1)=a+b+c=\dfrac{1}{2},\\f(2)=4a+2b+c=\dfrac{1}{3},\\f(3)=9a+3b+c=\dfrac{1}{4},\end{cases}$

解之，得

$$a=\frac{1}{24},\ b=\frac{-7}{24},\ c=\frac{3}{4},$$

故所求多项式为

$$f(x)=\frac{1}{24}x^2-\frac{7}{24}x+\frac{3}{4}。$$

解 2 设 $F(x)=(x+1)f(x)-1$ 为三次多项式。按题意，$F(x)$ 有 3 个零点：1，2 与 3。

故

$$F(x)=k(x-1)(x-2)(x-3),$$

但

$$F(-1)=-1, 又 F(-1)=k(-1-1)(-1-2)(-1-3),$$

得

$$k=\frac{1}{24}。$$

故所求多项式为

$$f(x)=\frac{F(x)+1}{x+1}=\frac{1}{x+1}\left[1+\frac{1}{24}(x-1)(x-2)(x-3)\right]$$
$$=\frac{1}{24}(x^2-7x+18)。$$

三、复系数多项式的标准分解式

命题 1.6 $\mathbf{C}[x]$ 中多项式 $p(x)$ 不可约 $\Leftrightarrow \deg p(x)=1$。

证 若 $p(x)$ 是 $\mathbf{C}[x]$ 的一个不可约多项式，由代数基本定理，$p(x)$ 有一复数根 a，从而 $(x-a)\mid p(x)$。因 $p(x)$ 是不可约多项式，故 $\deg p(x)=\deg(x-a)=1$。反之，一次多项式当然是不可约的。本命题得证。

定理 1.7 在 $\mathbf{C}[x]$ 中，任一次数大于零的多项式 $f(x)$ 都可惟一表成如下标准形：

$$f(x)=a(x-c_1)^{\alpha_1}\cdots(x-c_t)^{\alpha_t}。$$

这里 $a\neq0$ 为 $f(x)$ 的首项系数；$c_1,\cdots,c_t$ 是 $f(x)$ 的全部两两不同的零点。

证 因为 $\mathbf{C}[x]$ 中不可约多项式全是一次的，由因式分解惟一性定

理(标准形)即得本定理。

四、实系数多项式的标准分解式

命题 1.7 若 $p(x)$是 $\mathbf{R}[x]$中首 1 不可约多项式,则

(1) $p(x) = x - c$;

或

(2) $p(x) = (x - c)^2 + d^2, (d \neq 0)$。

证 若 $p(x)$有一实数根 c,则 $x - c \mid p(x)$,因 $p(x)$ 为首 1 不可约多项式,故 $p(x) = x - c$。

若 $p(x)$没有实数根,则 $p(x)$有一复数根 $c + d\mathrm{i}(d \neq 0)$,$p(c + d\mathrm{i}) = 0$。此时 $p(c - d\mathrm{i}) = p(\overline{c + d\mathrm{i}}) = \overline{p(c + d\mathrm{i})} = 0$,即 $c - d\mathrm{i}$ 也是 $p(x)$ 根,从而

$$(x - c)^2 + d^2 = [x - (c + d\mathrm{i})][x - (c - d\mathrm{i})] \text{ 整除 } p(x)。$$

在复数域内 $(x - c)^2 + d^2$ 能整除 $p(x)$,由带余除法余数惟一性,$\mathbf{C} \supseteq \mathbf{R}$,故在实数域内$(x - c)^2 + d^2$ 仍能整除 $p(x)$。

因 $p(x)$是首 1 不可约多项式,故 $p(x) = (x - c)^2 + d^2$。

把命题 1.7 用到因式分解惟一性定理上,得

定理 1.8 次数大于零的实系数多项式有如下标准分解式

$$f(x) = a(x - c_1)^{\alpha_1}\cdots(x - c_t)^{\alpha_t}[(x - p_1)^2 + q_1^2]^{\beta_1}\cdots \\ [(x - p_s)^2 + q_s^2]^{\beta_s}。$$

这里 $a \neq 0$,是 $f(x)$ 的首项系数;$c_1, \cdots, c_t$ 是 $f(x)$的两两不同的实数根,$p_1 + q_1\mathrm{i}, \cdots, p_s + q_s\mathrm{i}$ 是 $f(x)$的两两不同且不共轭的复数根。

习 题 1

1. 计算:

(1) $\sum\limits_{k=0}^{n} \frac{1}{2^k}$;

(2) $\sum\limits_{1\leqslant j\leqslant i\leqslant 5}(i-j)$；

(3) $\sum\limits_{k=1}^{n}(-1)^k C_n^k$；

(4) $\sum\limits_{k=1}^{n}\dfrac{1}{k(k+1)}$。

2. 化简：

(1) $\sum\limits_{i=1}^{n}\sum\limits_{j=1}^{n}(x_i-x_j)$；

(2) $\sum\limits_{i=1}^{n}\sum\limits_{j=1}^{n}(x_i+x_j)^2+\sum\limits_{i=1}^{n}\sum\limits_{j=1}^{n}(x_i-x_j)^2$。

3. 若 $\sum\limits_{1\leqslant j\leqslant i\leqslant n}a_{ij}=\sum\limits_{i=1}^{n}\sum\limits_{j}a_{ij}=\sum\limits_{j=1}^{n}\sum\limits_{i}a_{ij}$，试确定第二个和式中 j 与第三个和式中 i 的上下限。

4. 验证：

(1) n 为一固定正整数，$n\mathbf{Z}=\{nk\mid k\in\mathbf{Z}\}$ 是数环；

(2) $Z[i]=\{a+b\mathrm{i}\mid a,b\in\mathbf{Z}\}$ 是数环。

5. 设 ω 是 $x^2+x+1=0$ 的一个复数根，试验证：$K=Q(\omega)=\{a\omega+b\mid a,b\in\mathbf{Q}\}$ 是数域。

6. 设 b 是正整数，$a\in\mathbf{Z}$，求证：存在惟一一对整数 q,r，使 $a=qb+r$ 且 $0\leqslant r<b$。

7. 求证：任意两个不全为零的整数 a,b 必有最大公因子 d，即 d 满足：

(1) $d\in\mathbf{Z}^+$；

(2) d 是 a 与 b 的公因子；

(3) 若 c 也是 a 与 b 的公因子，则 c 是 d 的因子。

8. 称大于 1 的正整数 p 为素数，如果由 $d\in\mathbf{Z}^+$ 且 $d\mid p$，必有 $d=1$ 或 $d=p$，求证：若 $p\mid ab$，则必 $p\mid a$ 或 $p\mid b$，这里 $a,b\in\mathbf{Z}$。

9. 求 $g(x)$ 除 $f(x)$ 的商与余式：

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

(2) $f(x)=6x^4+3x^3-7$，$g(x)=x+3$。

10. m,p,q 适合什么条件时，有

$$x^2 + mx - 1 \mid x^3 + px + q。$$

11. 若正整数 d 是正整数 n 的因子，求证：

$$x^{p^d-1} - 1 \mid x^{p^n-1} - 1。$$

12. 若 $(f(x),g(x)) = 1$，m 是正整数，求证：$(f(x^m),g(x^m)) = 1$。

13. 若 $(f(x),g(x))=1$，求证：$(f(x)g(x),f(x)+g(x)) = 1$。

14. 若 $f_1(x) = af(x) + bg(x)$，$g_1(x) = cf(x) + dg(x)$，且 $ad - bc \neq 0$，求证：$(f(x),g(x)) = (f_1(x),g_1(x))$。

15. 求 t 的值使 $f(x) = x^3 + 3x^2 + tx + 1$ 有重根(a 称为 $f(x)$的重根，如果 $(x-a)^2 \mid f(x)$)。

16. 求多项式 $x^3 + px + q$ 有重根的条件。

17. 若 $(x-1)^2 \mid Ax^4 + Bx^2 + 1$，求 A,B。

18. 求证：多项式 $f(x) = 1 + x + \dfrac{x^2}{2!} + \cdots + \dfrac{x^n}{n!}$ 无重根。

19. (1) 若 $x^2 + x + 1 \mid f_1(x^3) + xf_2(x^3)$，求证：

$$(x-1) \mid (f_1(x),f_2(x));$$

(2) 若 $x^2 - x + 1 \mid f_1(x^3) + xf_2(x^3)$，求证：

$$(x+1) \mid (f_1(x),f_2(x))。$$

20. 若 n 是大于 1 的正整数，且 $f(x) \mid f(x^n)$，求证：$f(x)$ 的根只能是零或单位根(即其某个正整数次幂为 1)。

21. 求次数不大于 2 的多项式 $f(x)$使 $f(0) = -1$，$f(1) = 0$，$f(2) = 7$。

22. 若实系数首 1 多项式 $f(x)$无实根，求证：存在实系数多项式 $u(x)$与 $v(x)$使 $f(x) = u^2(x) + v^2(x)$。

23. 求次数不大于 n 的多项式 $f(x)$使

$$f(k) = \frac{k}{k+1},\ k = 0,1,2,\cdots,n。$$

第二章　行　列　式

本章主要介绍行列式的定义、性质与计算方法。

§1　二阶与三阶行列式

一、二元线性方程组与二阶行列式

用消去法解二元线性方程组

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

为消去未知数 x_2，以 a_{22} 与 a_{12} 分别乘上列两方程的两端，然后两个方程相减，得

$$(a_{11}a_{22} - a_{12}a_{21})x_1 = b_1a_{22} - a_{12}b_2.$$

类似地，消去 x_1，得

$$(a_{11}a_{22} - a_{12}a_{21})x_2 = a_{11}b_2 - b_1a_{21}.$$

当 $a_{11}a_{22} - a_{12}a_{21} \neq 0$ 时，求得方程组的解为

$$x_1 = \frac{b_1a_{22} - a_{12}b_2}{a_{11}a_{22} - a_{12}a_{21}},\quad x_2 = \frac{a_{11}b_2 - b_1a_{21}}{a_{11}a_{22} - a_{12}a_{21}}.$$

未知量 x_1，x_2 表达式中的分子、分母都是四个数分两对相乘再相减而得。

为了方便，我们把由四个数 a,b,c,d 组成的算式 $(ad - bc)$ 记成

$$ad - bc = \begin{vmatrix} a & b \\ c & d \end{vmatrix},$$

称为二阶行列式,运算的结果称为行列式的值。若记

$$D=\begin{vmatrix}a_{11} & a_{12}\\ a_{21} & a_{22}\end{vmatrix},\ D_1=\begin{vmatrix}b_1 & a_{12}\\ b_2 & a_{22}\end{vmatrix},\ D_2=\begin{vmatrix}a_{11} & b_1\\ a_{21} & b_2\end{vmatrix},$$

则当 $D\neq 0$ 时,二元线性方程组

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

的解可表为

$$x_1=\frac{D_1}{D},\ x_2=\frac{D_2}{D}。$$

注意:这里的分母 D 是由方程组未知量的系数所确定的二阶行列式;x_1的分子 D_1是用常数项 b_1,b_2代替 D 中 x_1的系数 a_{11},a_{21}所得的二阶行列式;x_2的分子 D_2是用常数项 b_1,b_2代替 D 中 x_2的系数 a_{12},a_{22}所得的二阶行列式。

例 2.1 求解二元线性方程组

$$\begin{cases}3x_1+2x_2=8\\ 5x_1-4x_2=6\end{cases}$$

的解。

解 由于

$$D=\begin{vmatrix}3 & 2\\ 5 & -4\end{vmatrix}=-12-10=-22,$$

$$D_1=\begin{vmatrix}8 & 2\\ 6 & -4\end{vmatrix}=-32-12=-44,$$

$$D_2=\begin{vmatrix}3 & 8\\ 5 & 6\end{vmatrix}=18-40=-22,$$

故 $$x_1=\frac{-44}{-22}=2,\ x_2=\frac{-22}{-22}=1。$$

二、三阶行列式

为方便计，我们称由 9 个数$\{a_{ij} | i,j=1,2,3\}$组成的运算式子

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

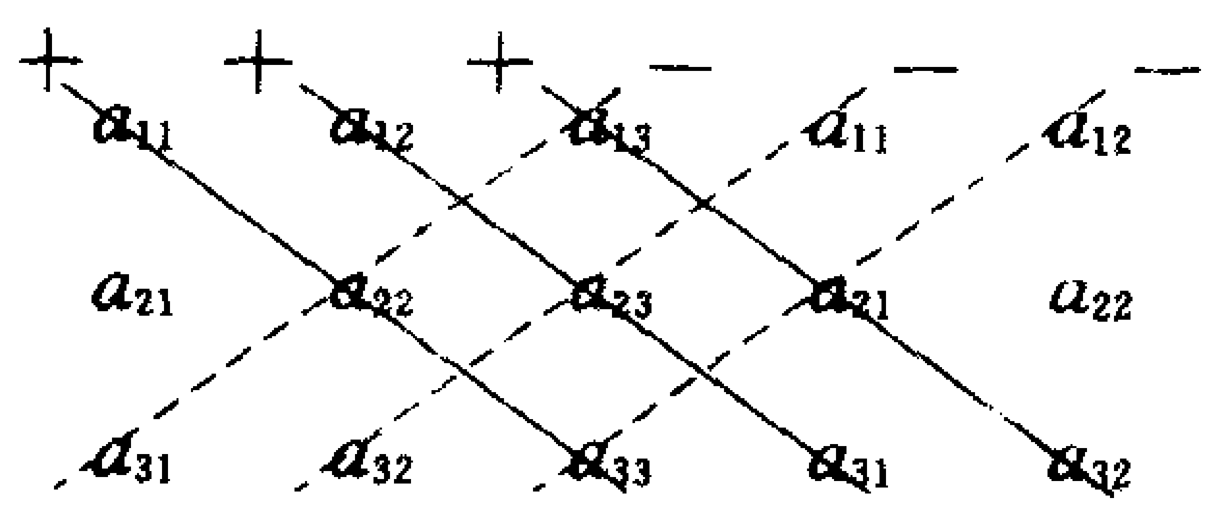
为三阶行列式，记成$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}$，

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

$$= a_{11}(a_{22}a_{33} - a_{23}a_{32}) - a_{12}(a_{21}a_{33} - a_{23}a_{31}) + a_{13}(a_{21}a_{32} - a_{22}a_{31})$$

$$= (a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32}) - (a_{11}a_{23}a_{32} + a_{12}a_{21}a_{33} + a_{13}a_{21}a_{32})。$$

三阶行列式含 6 个单项，每项均为不同行不同列的三个元素的乘积再冠以正负号，其规律遵循下图所示的对角线法则：图中有三条实线看作是平行于主对角线的连线，三条虚线看作是平行于副对角线的连线，实线上三元素的乘积冠以正号，虚线上三元素的乘积冠以负号。



例 2.2　计算三阶行列式

$$D = \begin{vmatrix} 1 & 2 & -4 \\ -2 & 2 & 1 \\ -3 & 4 & 2 \end{vmatrix},$$

解　按对角线法则，有

$$
\begin{aligned}
D &= 1\times 2\times(-2)+2\times 1\times(-3)+(-4)\times(-2)\times 4\\
&\quad -1\times 1\times 4-2\times(-2)\times(-2)-(-4)\times 2\times(-3)\\
&=-4-6+32-4-8-24\\
&=-14。
\end{aligned}
$$

例 2.3 求解方程

$$\begin{vmatrix} 1 & 1 & 2 \\ 2 & 3 & x \\ 7 & 12 & x^2 \end{vmatrix} = 0。$$

解 方程左端的三阶行列式

$$
\begin{aligned}
D &= 3x^2+7x+48-42-12x-2x^2\\
&= x^2-5x+6,
\end{aligned}
$$

由 $x^2-5x+6=0$，解得 $x=2$ 或 $x=3$。

类似于二阶行列式与二元线性方程组的关系，我们有：当系数行列式

$$D=\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} \neq 0$$

时，三元线性方程组

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

的解为：

$$x_1=\frac{D_1}{D},\ x_2=\frac{D_2}{D},\ x_3=\frac{D_3}{D}。$$

这里 D_1 是用常数项 b_1, b_2, b_3 替换系数行列式 D 中的 x_1 系数 a_{11}，a_{21}，a_{31} 而得的三阶行列式：

$$D_1 = \begin{vmatrix} b_1 & a_{12} & a_{13} \\ b_2 & a_{22} & a_{23} \\ b_3 & a_{32} & a_{33} \end{vmatrix},$$

类似有：

$$D_2 = \begin{vmatrix} a_{11} & b_1 & a_{13} \\ a_{21} & b_2 & a_{23} \\ a_{31} & b_3 & a_{33} \end{vmatrix},\ D_3 = \begin{vmatrix} a_{11} & a_{12} & b_1 \\ a_{21} & a_{22} & b_2 \\ a_{31} & a_{33} & b_3 \end{vmatrix}。$$

例 2.4 求解三元线性方程组

$$\begin{cases} x + y - 2z = -3, \\ 5x - 2y + 7z = 22, \\ 2x - 5y + 4z = 4。 \end{cases}$$

解 系数行列式

$$D = \begin{vmatrix} 1 & 1 & -2 \\ 5 & -2 & 7 \\ 2 & -5 & 4 \end{vmatrix} = -8 + 50 + 14 - 8 - 20 + 35 = 63,$$

又

$$D_1 = \begin{vmatrix} -3 & 1 & -2 \\ 22 & -2 & 7 \\ 4 & -5 & 4 \end{vmatrix} = 24 + 220 + 28 - 16 - 88 - 105 = 63,$$

$$D_2 = \begin{vmatrix} 1 & -3 & -2 \\ 5 & 22 & 7 \\ 2 & 4 & 4 \end{vmatrix} = 88 - 40 - 42 + 88 + 60 - 28 = 126,$$

$$D_3 = \begin{vmatrix} 1 & 1 & -3 \\ 5 & -2 & 22 \\ 2 & -5 & 4 \end{vmatrix} = -8 + 75 + 44 - 12 - 20 + 110 = 189,$$

故

$$\begin{cases} x = \dfrac{D_1}{D} = \dfrac{63}{63} = 1, \\ y = \dfrac{D_2}{D} = \dfrac{126}{63} = 2, \\ z = \dfrac{D_3}{D} = \dfrac{189}{63} = 3。 \end{cases}$$

§2 n 阶行列式的归纳定义

为了把二元线性方程组与二阶行列式的关系推广到一般情形，我们引进 n 阶行列式的概念。

n 阶行列式是指由 n^2 个数 $\{a_{ij} | i, j=1,2,\cdots,n\}$ 组成的算式。它的运算结果为一个数。为方便计，常将此运算式记成

$$\begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix},$$

运算的结果称为行列式的值。

定义 2.1 若 T 是数环，$a_{ij} \in T$，归纳定义 T 上行列式如下：

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

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

$$\vdots \qquad\qquad \vdots \qquad\qquad \vdots$$

$$\begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix} = \sum_{i=1}^{n} a_{1j}(-1)^{1+j} M\begin{pmatrix} 1 \\ j \end{pmatrix},$$

$$= a_{11}M\begin{pmatrix} 1 \\ 1 \end{pmatrix} - a_{12}M\begin{pmatrix} 1 \\ 2 \end{pmatrix} + a_{13}M\begin{pmatrix} 1 \\ 3 \end{pmatrix} - \cdots + (-1)^{1+n}M\begin{pmatrix} 1 \\ n \end{pmatrix}。$$

这里 $M\begin{pmatrix}1\\j\end{pmatrix}$ 表示原来的行列式中划去第 1 行和第 j 列后得到的 $(n-1)$ 阶行列式。按定义，n 阶行列式是它的第一行的元素乘上相应的 $(n-1)$ 阶行列式加减而得。按定义，计算行列式的值是对它的元素作加、减、乘三种运算，故 T 上的行列式的值仍在 T 中。

为了方便，常用 $M_A\begin{pmatrix}i_1 & \cdots & i_k\\ j_1 & \cdots & j_k\end{pmatrix}$ 表示阵列 $A=\begin{pmatrix}a_{11} & \cdots & a_{1n}\\ \vdots & & \vdots\\ a_{n1} & \cdots & a_{nn}\end{pmatrix}$ 中划去第 $i_1,i_2,\cdots,i_k$ 行与第 $j_1,j_2,\cdots,j_k$ 列后所得 $(n-k)$ 阶行列式，在 A 明确的前提下，简记为 $M\begin{pmatrix}i_1 & \cdots & i_k\\ j_1 & \cdots & j_k\end{pmatrix}$。

显然 $M\begin{pmatrix}i_1\\ j_1\end{pmatrix}\begin{pmatrix}i_2\\ j_2\end{pmatrix}=M\begin{pmatrix}i_1 & i_2\\ j_1 & j_2\end{pmatrix}$。

例如，设 $A=\begin{pmatrix}1 & 2 & 3 & 4\\ 5 & 6 & 7 & 8\\ -1 & 0 & 1 & 2\\ -4 & 3 & 2 & -1\end{pmatrix}$，则

$$M\begin{pmatrix}1 & 4\\ 2 & 3\end{pmatrix}=\begin{vmatrix}5 & 8\\ -1 & 2\end{vmatrix}=18,$$

$$M\begin{pmatrix}2\\ 4\end{pmatrix}=\begin{vmatrix}1 & 2 & 3\\ -1 & 0 & 1\\ -4 & 3 & 2\end{vmatrix}=1\times\begin{vmatrix}0 & 1\\ 3 & 2\end{vmatrix}-2\times\begin{vmatrix}-1 & 1\\ -4 & 2\end{vmatrix}$$

$$+3\times\begin{vmatrix}-1 & 0\\ -4 & 3\end{vmatrix}=1\times(-3)-2\times2+3\times(-3)$$

$$=-16。$$

n 阶行列式 $\begin{vmatrix}a_{n} & \cdots & a_{1n}\\ \vdots & & \vdots\\ a_{n1} & \cdots & a_{nn}\end{vmatrix}$ 可简记成 $|a_{ij}|_n$，

a_{ij} 表示行列式中 i 行、j 列的元素，右下标 n 表示行列式的阶数。

例如

$$|i \cdot j|_3 = \begin{vmatrix} 1\cdot 1 & 1\cdot 2 & 1\cdot 3 \\ 2\cdot 1 & 2\cdot 2 & 2\cdot 3 \\ 3\cdot 1 & 3\cdot 2 & 3\cdot 3 \end{vmatrix} = \begin{vmatrix} 1 & 2 & 3 \\ 2 & 4 & 6 \\ 3 & 6 & 9 \end{vmatrix}。$$

§3 行列式的性质

引理 2.1 n 阶行列式的值可由它的第一列的元素分别乘上相应 $(n-1)$ 阶行列式加减而得，即

$$\begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & & \vdots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix} = a_{11}M\begin{pmatrix} 1 \\ 1 \end{pmatrix} - a_{21}M\begin{pmatrix} 2 \\ 1 \end{pmatrix} + \cdots + (-1)^{n+1}a_{n1}M\begin{pmatrix} n \\ 1 \end{pmatrix}$$

$$= \sum_{i=1}^{n}(-1)^{i+1}a_{i1}M\begin{pmatrix} i \\ 1 \end{pmatrix}。$$

证 对行列式的阶数 n 用归纳法。

$n=2$ 时，等式两端都等于 $a_{11}a_{22} - a_{12}a_{21}$，结论成立。

若行列式的阶数为 $(n-1)$ 时等式成立，考虑 n 阶行列式的情形。

$i \geqslant 2$ 时，$M\begin{pmatrix} i \\ 1 \end{pmatrix}$ 是原行列式中划去第 i 行、第 1 列后得 $(n-1)$ 阶行列式，它的 j 列是原行列式的 $(j+1)$ 列。于是，按行列式的定义：

$$M\begin{pmatrix} i \\ 1 \end{pmatrix} = \begin{vmatrix} \cancel{a_{11}} & a_{12} & a_{13} & \cdots & a_{1n} \\ \cancel{a_{21}} & a_{22} & a_{23} & \cdots & a_{2n} \\ \vdots & & & & \vdots \\ \cancel{a_{i1}} & \cancel{a_{i2}} & \cancel{a_{i3}} & \cancel{\cdots} & \cancel{a_{in}} \\ \vdots & & & & \vdots \\ \cancel{a_{n1}} & a_{n2} & a_{n3} & \cdots & a_{nn} \end{vmatrix}$$

$$= a_{12}M\begin{pmatrix} i \\ 1 \end{pmatrix}\begin{pmatrix} 1 \\ 2 \end{pmatrix} - a_{13}M\begin{pmatrix} i \\ 1 \end{pmatrix}\begin{pmatrix} 1 \\ 3 \end{pmatrix} + \cdots$$

$$= \sum_{j=2}^{n}(-1)^{j}a_{1j}M\begin{pmatrix} i \\ 1 \end{pmatrix}\begin{pmatrix} 1 \\ j \end{pmatrix}。$$

把 $M\begin{pmatrix} i \\ 1 \end{pmatrix}$ $(i \geqslant 2)$ 的表达式代入求证等式的右端得

$$
\begin{aligned}
\text{右端} &= a_{11}M\begin{pmatrix} 1 \\ 1 \end{pmatrix} + \sum_{i=2}^{n}(-1)^{i+1}a_{i1} \cdot \sum_{j=2}^{n}(-1)^{j}a_{1j}M\begin{pmatrix} i & 1 \\ 1 & j \end{pmatrix} \\
&= a_{11}M\begin{pmatrix} 1 \\ 1 \end{pmatrix} + \sum_{i=2}^{n}\sum_{j=2}^{n}(-1)^{i+j+1}a_{i1}a_{1j}M\begin{pmatrix} 1 \\ j \end{pmatrix}\begin{pmatrix} i \\ 1 \end{pmatrix}。
\end{aligned}
$$

交换求和号得

$$
\text{右端} = a_{11}M\begin{pmatrix} 1 \\ 1 \end{pmatrix} + \sum_{j=2}^{n}(-1)^{j+1}a_{1j}\sum_{i=2}^{n}(-1)^{i}a_{i1}M\begin{pmatrix} 1 \\ j \end{pmatrix}\begin{pmatrix} i \\ 1 \end{pmatrix}。
$$

由归纳法假设，$(n-1)$ 阶行列式 $M\begin{pmatrix} 1 \\ j \end{pmatrix}$ 的值可由它的第一列元素乘上相应 $(n-2)$ 阶行列式加减而得，即$\left(\text{注意：} M\begin{pmatrix} 1 \\ j \end{pmatrix} \text{的第 } i \text{ 行为原行列式的第 } (i+1) \text{ 行}\right)$

$$
\begin{aligned}
M\begin{pmatrix} 1 \\ i \end{pmatrix} &= \begin{vmatrix} a_{11}a_{12} & \cdots & a_{1j} & \cdots & a_{1n} \\ a_{21}a_{22} & \cdots & a_{2j} & \cdots & a_{2n} \\ a_{31}a_{32} & \cdots & a_{3j} & \cdots & a_{3n} \\ \vdots & \vdots & & & \vdots \\ a_{n1}a_{n2} & \cdots & a_{nj} & \cdots & a_{nn} \end{vmatrix} \\
&= a_{21}M\begin{pmatrix} 1 \\ j \end{pmatrix}\begin{pmatrix} 2 \\ 1 \end{pmatrix} - a_{31}M\begin{pmatrix} 1 \\ j \end{pmatrix}\begin{pmatrix} 3 \\ 1 \end{pmatrix} + \cdots \\
&\quad - (-1)^{i}a_{i1}M\begin{pmatrix} 1 \\ j \end{pmatrix}\begin{pmatrix} i \\ 1 \end{pmatrix} + \cdots + (-1)^{n}a_{n1}M\begin{pmatrix} 1 \\ j \end{pmatrix}\begin{pmatrix} n \\ 1 \end{pmatrix} \\
&= \sum_{i=2}^{n}(-1)^{i}a_{i1}M\begin{pmatrix} 1 \\ j \end{pmatrix}\begin{pmatrix} i \\ j \end{pmatrix},
\end{aligned}
$$

代入右端表达式得

$$
\text{右端} = a_{11}M\begin{pmatrix} 1 \\ 1 \end{pmatrix} + \sum_{j=2}^{n}(-1)^{j+1}a_{1j}M\begin{pmatrix} 1 \\ j \end{pmatrix}
$$

$$= \sum_{j=1}^{n} (-1)^{j+1} a_{1j} M\begin{pmatrix} 1 \\ j \end{pmatrix} = \text{左端}。$$

由归纳法，本引理得证。

引理 2.2 行列式相邻两列互换，其值改号。即若互换行列式

$$|A| = \begin{array}{c} \begin{vmatrix} a_{11} & \cdots & a_{1i} & a_{1i+1} & \cdots & a_{1n} \\ a_{21} & \cdots & a_{2i} & a_{2i+1} & \cdots & a_{2n} \\ \vdots & & \vdots & \vdots & & \vdots \\ a_{n1} & \cdots & a_{ni} & a_{ni+1} & \cdots & a_{nn} \end{vmatrix} \\ \begin{array}{cc} i & i+1 \\ \text{列} & \text{列} \end{array} \end{array}$$

的相邻两列，如第 i 列与第 $(i+1)$ 列，得行列式

$$|B| = \begin{array}{c} \begin{vmatrix} a_{11} & \cdots & a_{1i+1} & a_{1i} & \cdots & a_{1n} \\ a_{21} & \cdots & a_{2i+1} & a_{2i} & \cdots & a_{2n} \\ \vdots & & \vdots & \vdots & & \vdots \\ a_{n1} & \cdots & a_{ni+1} & a_{ni} & \cdots & a_{nn} \end{vmatrix} \\ \begin{array}{cc} i & i+1 \\ \text{列} & \text{列} \end{array} \end{array},$$

则

$$|B| = -|A|。$$

证 对行列式的阶数 n 用归纳法。

$n=2$ 时，$|B| = \begin{vmatrix} a_{12} & a_{11} \\ a_{22} & a_{21} \end{vmatrix} = a_{12}a_{21} - a_{11}a_{22} = -\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = -|A|$，结论成立。

若对 $(n-1)$ 阶行列式引理正确，我们考虑 n 阶行列式的情形。由行列式定义：

$$|B| = a_{11}M_B\begin{pmatrix} 1 \\ 1 \end{pmatrix} + \cdots + (-1)^{1+i} a_{1i+1} M_B\begin{pmatrix} 1 \\ i \end{pmatrix}$$

$$+ (-1)^{1+(i+1)} a_{1i} M_B\begin{pmatrix} 1 \\ i+1 \end{pmatrix} + \cdots + (-1)^{n+1} a_{1n} M_B\begin{pmatrix} 1 \\ n \end{pmatrix}。$$

当 $i \neq i, i+1$ 时，$M_B\begin{pmatrix}1\\j\end{pmatrix}$ 是行列式 $|B|$ 中划去第 1 行、第 j 列（保留 i 列、$(i+1)$ 列）后留下的 $(n-1)$ 阶行列式：

$$M_B\begin{pmatrix}1\\j\end{pmatrix} = \begin{vmatrix} a_{11} & \cdots & a_{1i+1} & a_{1i} & \cdots & a_{1j} & \cdots & a_{1n} \\ a_{21} & \cdots & a_{2i+1} & a_{2i} & \cdots & a_{2j} & \cdots & a_{2n} \\ a_{31} & \cdots & a_{3i+1} & a_{3i} & \cdots & a_{3j} & \cdots & a_{3n} \\ \vdots & & \vdots & \vdots & & \vdots & & \vdots \\ a_{n1} & \cdots & a_{ni+1} & a_{ni} & \cdots & a_{nj} & \cdots & a_{nn} \end{vmatrix},$$

$$\begin{matrix} i & & i+1 \\ \text{列} & & \text{列} \end{matrix}$$

它可由 $M_A\begin{pmatrix}1\\j\end{pmatrix}$ 互换第 i 列与第 $(i+1)$ 列得到：

$$M_A\begin{pmatrix}1\\j\end{pmatrix} = \begin{vmatrix} a_{11} & \cdots & a_{1i} & a_{1i+1} & \cdots & a_{1j} & \cdots & a_{1n} \\ a_{21} & \cdots & a_{2i} & a_{2i+1} & \cdots & a_{2j} & \cdots & a_{2n} \\ \vdots & & \vdots & \vdots & & & & \vdots \\ a_{n1} & \cdots & a_{ni} & a_{ni+1} & \cdots & a_{nj} & \cdots & a_{nn} \end{vmatrix}。$$

$$\begin{matrix} i & & i+1 \\ \text{列} & & \text{列} \end{matrix}$$

由于 $M_B\begin{pmatrix}1\\j\end{pmatrix}$，$M_A\begin{pmatrix}1\\j\end{pmatrix}$ 为 $(n-1)$ 阶行列式，由归纳法假设：

$$M_B\begin{pmatrix}1\\j\end{pmatrix} = -M_A\begin{pmatrix}1\\j\end{pmatrix}。$$

而

$$M_B\begin{pmatrix}1\\i\end{pmatrix} = \begin{vmatrix} a_{11} & \cdots & a_{1i+1} & a_{1i} & \cdots & a_{1n} \\ a_{21} & \cdots & a_{2i+1} & a_{2i} & \cdots & a_{2n} \\ \vdots & & \vdots & \vdots & & \vdots \\ a_{n1} & \cdots & a_{ni+1} & a_{ni} & \cdots & a_{nn} \end{vmatrix}$$

$$\begin{matrix} i \\ \text{列} \end{matrix}$$

$$
= \begin{array}{c}
\begin{vmatrix}
a_{11} & \cdots & a_{1i} & a_{1i+1} & \cdots & a_{1n} \\
a_{21} & \cdots & a_{2i} & a_{2i+1} & \cdots & a_{2n} \\
\vdots & & \vdots & \vdots & & \vdots \\
a_{n1} & \cdots & a_{ni} & a_{ni+1} & \cdots & a_{nn}
\end{vmatrix} \\
i+1 \\
\text{列}
\end{array}
= M_A\begin{pmatrix} 1 \\ i+1 \end{pmatrix},
$$

$$
M_B\begin{pmatrix} 1 \\ i+1 \end{pmatrix} = \begin{array}{c}
\begin{vmatrix}
a_{11} & \cdots & a_{1i+1} & a_{1i} & \cdots & a_{1n} \\
a_{21} & \cdots & a_{2i+1} & a_{2i} & \cdots & a_{2n} \\
\vdots & & \vdots & \vdots & & \vdots \\
a_{n1} & \cdots & a_{ni+1} & a_{ni} & \cdots & a_{nn}
\end{vmatrix} \\
i+1 \\
\text{列}
\end{array}
$$

$$
= \begin{array}{c}
\begin{vmatrix}
a_{11} & \cdots & a_{1i} & a_{1i+1} & \cdots & a_{1n} \\
a_{21} & \cdots & a_{2i} & a_{2i+1} & \cdots & a_{2n} \\
\vdots & & \vdots & \vdots & & \vdots \\
a_{n1} & \cdots & a_{ni} & a_{ni+1} & \cdots & a_{nn}
\end{vmatrix} \\
i \\
\text{列}
\end{array}
= M_A\begin{pmatrix} 1 \\ i \end{pmatrix},
$$

于是

$$
\begin{aligned}
|B| = & -a_{11}M_A\begin{pmatrix} 1 \\ 1 \end{pmatrix} + \cdots + (-1)^{i+1}a_{1i+1}M_A\begin{pmatrix} 1 \\ i+1 \end{pmatrix} + \\
& (-1)^{i+2}a_{1i}M_A\begin{pmatrix} 1 \\ i \end{pmatrix} + \cdots + (-1)^{j+1}a_{1j}\left[-M_A\begin{pmatrix} 1 \\ j \end{pmatrix}\right] + \cdots \\
& + (-1)^{n+1}a_{1n}\left[-M_A\begin{pmatrix} 1 \\ n \end{pmatrix}\right],
\end{aligned}
$$

交换和式中的第 i 项与第 $(i+1)$ 项，注意到 $(-1)^i = (-1)^{i+2}$，提取公因子(-1)得

$$
|B| = -\left[a_{11}M_A\begin{pmatrix} 1 \\ 1 \end{pmatrix} - \cdots + (-1)^{1+i}a_{1i}M_A\begin{pmatrix} 1 \\ i \end{pmatrix}\right.
$$

$$+(-1)^{i+2}a_{1i+1}M_A\begin{pmatrix}1\\ i+1\end{pmatrix}+\cdots+(-1)^{1+n}a_{1n}M_A\begin{pmatrix}1\\ n\end{pmatrix}\Big]$$
$$=-|A|。$$

即对 n 阶行列式，引理正确。由归纳法，本引理得证。

定理 2.1 任意两列互换，行列值的值改号，即 $|\alpha_1,\cdots,\alpha_i,\beta_1,\cdots,\beta_t,\alpha_j,\cdots,\alpha_n| = -|\alpha_1,\cdots,\alpha_j,\beta_1,\cdots,\beta_t,\alpha_i,\cdots,\alpha_n|$。这里，用 $\alpha_1,\cdots,\alpha_i,\beta_1,\cdots,\beta_t,\alpha_j,\cdots,\alpha_n$ 表示行列式的各个列。

证 左端行列式的 α_i 列可经 $(t+1)$ 次相邻两列互换到 α_j 的位置，即 α_i 列依次与 β_1 列，β_2 列，$\cdots$，β_t 列及 α_j 列互换，由引理 2 得

$$\begin{aligned}左端 &= -|\alpha_1,\cdots,\beta_1,\alpha_i,\beta_2,\cdots,\beta_t,\alpha_j,\cdots,\alpha_n|\\ &= \cdots = (-1)^{t+1}|\alpha_1,\cdots,\beta_1,\cdots,\beta_t,\alpha_j,\alpha_i,\cdots,\alpha_n|,\end{aligned}$$

α_j 列继续与 β_t 列，$\cdots$，β_1 列互换，即再经 t 次相邻两列互换得

$$\begin{aligned}左端 &= (-1)^{t+1}(-1)|\alpha_1,\cdots,\beta_1,\cdots,\alpha_j,\beta_t,\alpha_i,\cdots,\alpha_n|\\ &= \cdots = (-1)^{t+1}(-1)^t|\alpha_1,\cdots,\alpha_j,\beta_1,\cdots,\beta_t,\alpha_i,\cdots,\alpha_n|\\ &= (-1)^{2t+1}|\alpha_1,\cdots,\alpha_j,\beta_1,\cdots,\beta_t,\alpha_i,\cdots,\alpha_n| = 右端。\end{aligned}$$

推论 两列相等，行列式值为零。

证 设行列式的第 i 列与第 j 列相等，其值为 d。互换 i 列与 j 列后新的行列式值为 $-d$；但因为第 i 列与第 j 列相等，互换该两列后，行列式依旧，其值仍为 d。故 $-d=d$，于是必 $d=0$。

若 $\boldsymbol{A}$ 表示由 $s\times n$ 个数组成的 s 行、n 列的阵列 $\begin{pmatrix}a_{11} & \cdots & a_{1n}\\ \vdots & & \vdots\\ a_{s1} & \cdots & a_{sn}\end{pmatrix}$，常称 $\boldsymbol{A}$ 为 s 行、n 列矩阵，且用 $\boldsymbol{A}'$ 表示由这 $s\times n$ 个数组成的 n 行、s 列矩阵 $\begin{pmatrix}a_{11} & a_{21} & \cdots & a_{s1}\\ \vdots & \vdots & & \vdots\\ a_{1n} & a_{2n} & \cdots & a_{sn}\end{pmatrix}$，称它为 $\boldsymbol{A}$ 的转置。

如 $\boldsymbol{A}=\begin{pmatrix}1 & 2 & 3\\ 4 & 5 & 6\end{pmatrix}$，则 $\boldsymbol{A}'=\begin{pmatrix}1 & 4\\ 2 & 5\\ 3 & 6\end{pmatrix}$。一般规则是：$\boldsymbol{A}$ 的第 i 行、j

列的元素恰是 A' 的第 j 行、i 列元素。

性质 1

$$|A| = \begin{vmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{vmatrix} = |A'|\text{。即行列式转置，其值不变。}$$

证 对行列式的阶数 n 用归纳法。

$n=2$ 时，$|A| = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}a_{22} - a_{12}a_{21} = \begin{vmatrix} a_{11} & a_{21} \\ a_{12} & a_{22} \end{vmatrix} = |A'|$。

若行列式的阶数为 $(n-1)$ 时，性质成立。考虑行列式阶数为 n 时情形。此时

$$|A'| = \begin{vmatrix} a_{11} & a_{21} & \cdots & a_{i1} & \cdots & a_{n1} \\ \vdots & & & & & \vdots \\ a_{1n} & a_{2n} & \cdots & a_{in} & \cdots & a_{nn} \end{vmatrix} = a_{11}M_{A'}\begin{pmatrix} 1 \\ 1 \end{pmatrix} - a_{21}M_{A'}\begin{pmatrix} 2 \\ 1 \end{pmatrix} + \cdots$$

$$= \sum_{i=1}^{n}(-1)^{i+1}a_{i1}M_{A'}\begin{pmatrix} 1 \\ i \end{pmatrix}\text{。}$$

$(n-1)$ 阶行列式 $M_{A'}\begin{pmatrix} 1 \\ i \end{pmatrix}$ 是 n 阶行列式 $|A'|$ 中划去第 1 行、第 i 列而得到的，恰是 n 阶行列式 $|A|$ 中划去第 i 行、第 1 列而得到的 $(n-1)$ 阶行列式的转置，即 $M_A\begin{pmatrix} i \\ 1 \end{pmatrix}$ 的转置：

$$M_{A'}\begin{pmatrix} 1 \\ i \end{pmatrix} = \begin{vmatrix} a_{11} & a_{21} & \cdots & a_{i1} & \cdots & a_{n1} \\ a_{12} & a_{22} & \cdots & a_{i2} & \cdots & a_{n2} \\ \vdots & & & \vdots & & \vdots \\ a_{1n} & a_{2n} & \cdots & a_{in} & \cdots & a_{nn} \end{vmatrix}$$

$$= \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{i1} & a_{i2} & \cdots & a_{in} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix}' = \left[M_A\begin{pmatrix} i \\ 1 \end{pmatrix}\right]'$$

按归纳法假设，$(n-1)$ 阶行列式 $M_{A'}\begin{pmatrix}1\\i\end{pmatrix}$ 的值等于其转置 $M_A\begin{pmatrix}i\\1\end{pmatrix}$ 的值，即

$$M_{A'}\begin{pmatrix}1\\i\end{pmatrix}=M_A\begin{pmatrix}i\\1\end{pmatrix},$$

于是

$$|A'|=\sum_{i=1}^{n}(-1)^{i+1}a_{i1}M_A\begin{pmatrix}i\\1\end{pmatrix},$$

由引理 2.1 得

$$|A'|=|A|,$$

由归纳法，结论成立。

推论 两行互换，行列式值改号。

证 设行列式 $|A|$ 经第 i 行与第 j 行互换得 $|B|$，则 $|A'|$ 经 i 列与 j 列互换得 $|B'|$，于是

$$|A|=|A'|=-|B'|=-|B|。$$

性质 2 $|A|=|a_{ij}|_n=\sum_{j=1}^{n}(-1)^{k+j}a_{kj}M\begin{pmatrix}k\\j\end{pmatrix},\quad(1)$

$$|A|=\sum_{i=1}^{n}(-1)^{i+l}a_{il}M\begin{pmatrix}i\\l\end{pmatrix}。\quad(2)$$

式(1)称行列式按第 k 行的展开式，式(2)称按第 l 列的展开式。

证 若用 α_i 表行列式 $|A|$ 的第 i 行。把行列式 $|A|$ 的第 k 行 α_k 依次与 $\alpha_{k-1},\cdots,\alpha_1$ 互换，即经 $(k-1)$ 次相邻两行互换，利用性质 1 的推论得

$$|A|=\begin{vmatrix}\alpha_1\\\vdots\\\alpha_{k-1}\\\alpha_k\\\vdots\end{vmatrix}=-\begin{vmatrix}\alpha_1\\\vdots\\\alpha_k\\\alpha_{k-1}\\\vdots\end{vmatrix}=\cdots=(-1)^{k-1}\begin{vmatrix}\alpha_k\\\alpha_1\\\alpha_2\\\vdots\\\alpha_{k-1}\\\vdots\end{vmatrix}$$

$$= (-1)^{k-1}\begin{vmatrix} a_{k1} & a_{k2} & \cdots & a_{kn} \\ a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & & \vdots \\ a_{k-1,1} & a_{k-1,2} & \cdots & a_{k-1,n} \\ a_{k+1,1} & a_{k-1,2} & \cdots & a_{k+1,n} \\ \vdots & \vdots & & \vdots \end{vmatrix}$$

$$= (-1)^{k-1}\sum_{j=1}^{n}(-1)^{1+j}a_{kj}M\begin{pmatrix} k \\ j \end{pmatrix}$$

$$= \sum_{j=1}^{n}(-1)^{k+j}a_{kj}M\begin{pmatrix} k \\ j \end{pmatrix}。$$

又：$|A| = |A'| = \begin{vmatrix} a_{11} & a_{21} & \cdots & a_{n1} \\ \vdots & & & \vdots \\ a_{1l} & a_{2l} & \cdots & a_{nl} \\ \vdots & & & \vdots \\ a_{1n} & a_{2n} & \cdots & a_{nn} \end{vmatrix}$（按 l 行展开）

$$= \sum_{i=1}^{n}(-1)^{l+i}a_{il}M_{A'}\begin{pmatrix} l \\ i \end{pmatrix}$$

$$= \sum_{i=1}^{n}(-1)^{l+i}a_{il}M\begin{pmatrix} i \\ l \end{pmatrix}。$$

例 2.5 若 $\begin{vmatrix} 1 & 3 & 5 \\ 0 & 2 & k \\ 2 & 1 & 7 \end{vmatrix} = 0$，求 k 的值。

解 按第二行展开行列式，得

$$(-1)^{2+1}0\cdot\begin{vmatrix} 3 & 5 \\ 1 & 7 \end{vmatrix} + (-1)^{2+2}2\cdot\begin{vmatrix} 1 & 5 \\ 2 & 7 \end{vmatrix} + (-1)^{2+3}\cdot k\begin{vmatrix} 1 & 3 \\ 2 & 1 \end{vmatrix}$$

$= 2(7-10) - k(1-6) = -6 + 5k = 0$，

故

$$k = \frac{6}{5}。$$

性质 3 若 n 阶行列式

$$|\boldsymbol{A}| = |a_{ij}|_n = |\alpha_1, \cdots, \alpha_k, \cdots, \alpha_n|$$

的第 k 列 α_k为两列之和：

$$\alpha_k = \begin{pmatrix} a_{1k} \\ a_{2k} \\ \vdots \\ a_{nk} \end{pmatrix} = \begin{pmatrix} b_1 + c_1 \\ b_2 + c_2 \\ \vdots \\ b_n + c_n \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix} + \begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{pmatrix} = \beta + \gamma,$$

则行列式$|\boldsymbol{A}|$为两个行列式之和

$$\begin{aligned} |\boldsymbol{A}| &= |\alpha_1, \cdots, \beta + \gamma, \cdots, \alpha_n| \\ &= |\alpha_1, \cdots, \beta, \cdots, \alpha_n| + |\alpha_1, \cdots, \gamma, \cdots, \alpha_n| 。\end{aligned}$$

行的情形也正确，即若行列式的一行为两行之和，则行列式为两个行列式之和。

证 按第 k 列展开行列式$|\boldsymbol{A}|$：

$$\begin{aligned} |\boldsymbol{A}| &= \sum_{i=1}^{n} (-1)^{i+k} a_{ik} M\begin{pmatrix} i \\ k \end{pmatrix} = \sum_{i=1}^{n} (-1)^{i+k} [b_i + c_i] M\begin{pmatrix} i \\ k \end{pmatrix} \\ &= \sum_{i=1}^{n} (-1)^{i+k} b_i M\begin{pmatrix} i \\ k \end{pmatrix} + \sum_{i=1}^{n} (-1)^{i+k} c_i M\begin{pmatrix} i \\ k \end{pmatrix} \\ &= |\alpha_1, \cdots, \beta, \cdots, \alpha_n| + |\alpha_1, \cdots, \gamma, \cdots, \alpha_n| 。\end{aligned}$$

行的情形：若用 $\alpha_1, \cdots, \alpha_k, \cdots, \alpha_n$分别表示行列式的各个行，$\alpha_k = \beta_k + \gamma_k$，则

$$\begin{aligned} \begin{vmatrix} \alpha_1 \\ \vdots \\ \alpha_k \\ \vdots \\ \alpha_n \end{vmatrix} &= \begin{vmatrix} \alpha_1 \\ \vdots \\ \beta + \gamma \\ \vdots \\ \alpha_n \end{vmatrix} = \begin{vmatrix} \alpha_1 \\ \vdots \\ \beta + \gamma \\ \vdots \\ \alpha_n \end{vmatrix}' = |\alpha_1', \cdots, \beta' + \gamma', \cdots, \alpha_n'| \\ &= |\alpha_1', \cdots, \beta', \cdots, \alpha_n'| + |\alpha_1', \cdots, \gamma', \cdots, \alpha_n'| \\ &= |\alpha_1', \cdots, \beta', \cdots, \alpha_n'|' + |\alpha_1', \cdots, \gamma', \cdots, \alpha_n'|' \end{aligned}$$

$$=\begin{vmatrix}\alpha_1\\ \vdots\\ \beta\\ \vdots\\ \alpha_n\end{vmatrix}+\begin{vmatrix}\alpha_1\\ \vdots\\ \gamma\\ \vdots\\ \alpha_n\end{vmatrix}。$$

性质 3 得证。

性质 4 (1) 两列(或行)互换,行列式值改号;

(2) 一列(或行)的每个数乘以 k,则行列式的值是原行列式的值与 k 之积;

(3) 一列(或行)的每个数的 k 倍加到另一列(行)上,行列式的值不变。

证 我们仅证列的情形,行的情形可仿性质 3 的证明同样处理。

(1) 即定理 2.1。

(2) $$\begin{vmatrix}a_{11}, & \cdots, & ka_{1j}, & \cdots a_{1n}\\ \vdots & & & \vdots\\ \vdots & & & \vdots\\ a_{n1}, & \cdots, & ka_{nj}, & \cdots a_{nn}\end{vmatrix}\xlongequal[\text{按 } j \text{ 列展开}]{}\sum_{i=1}^{n}(-1)^{i+j}ka_{ij}M\begin{pmatrix}i\\ j\end{pmatrix}$$

$$=k\sum_{i=1}^{n}(-1)^{i+j}a_{ij}M\begin{pmatrix}i\\ j\end{pmatrix}=k|a_{ij}|_n。$$

(3) $$|\alpha_1,\cdots,\alpha_i+k\alpha_j,\cdots,\alpha_j,\cdots,\alpha_n|$$

$$=|\alpha_1,\cdots,\alpha_i,\cdots,\alpha_j,\cdots,\alpha_n|+|\alpha_1,\cdots,k\alpha_j,\cdots,\alpha_j,\cdots,\alpha_n|$$

$$=|\alpha_1,\cdots,\alpha_i,\cdots,\alpha_j,\cdots,\alpha_n|+k|\alpha_1,\cdots,\alpha_j,\cdots,\alpha_j,\cdots\alpha_n|$$

$$=|\alpha_1,\cdots,\alpha_i,\cdots,\alpha_j,\cdots,\alpha_n|。$$

性质 4 的三种情形对行列式所作的变动常称为初等变换。我们将在矩阵这一章,详细研究初等变换。

§4 行列式的计算

一、三角行列式

若 $1\leqslant j<i\leqslant n$ 时,$a_{ij}=0$,则称 n 阶行列式

$$
|a_{ij}|_n=\begin{vmatrix} a_{11} & \cdots & \cdots & a_{1n} \\ 0 & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & a_{nn} \end{vmatrix}
$$

为上三角行列式，即其对角线以下的元素全为 0。这里我们把连结 a_{11}，a_{22}，…，a_{nn}的线称为行列式的主对角线或对角线。

若 $1 \leqslant i < j \leqslant n$ 时，$a_{ij} = 0$，则称 n 阶行列式

$$
|a_{ij}|_n=\begin{vmatrix} a_{11} & 0 & \cdots & \cdots & 0 \\ \vdots & \ddots & \ddots & & \vdots \\ \vdots & & \ddots & \ddots & 0 \\ a_{n1} & \cdots & & \cdots & a_{nn} \end{vmatrix}
$$

为下三角行列式，其对角线以上的元素全为 0。

命题 2.1 上、下三角行列式 $|a_{ij}|_n$ 的值等于其对角线上元素的积，即 $|a_{ij}|_n = a_{11}\cdots a_{nn}$，后者常记成 $\prod\limits_{i=1}^{n} a_{ij}$。

证

$$
\begin{vmatrix} a_{11} & 0 & \cdots & \cdots & 0 \\ \vdots & \ddots & \ddots & & \vdots \\ \vdots & & \ddots & \ddots & \vdots \\ \vdots & & & \ddots & 0 \\ \vdots & & & & \ddots \\ a_{n1} & \cdots & \cdots & \cdots & a_{nn} \end{vmatrix} \xlongequal{\text{按第一行展开}}
$$

$$
a_{11}\begin{vmatrix} a_{22} & 0 & 0 \\ \vdots & \ddots & \vdots \\ \vdots & \ddots & 0 \\ \vdots & & \ddots \\ a_{12} & & a_{nn} \end{vmatrix} \xlongequal{\text{继续按第一行展开}} a_{11}a_{22}\begin{vmatrix} a_{33} & 0 & \cdots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & 0 \\ a_{3n} & \cdots & \cdots & a_{nn} \end{vmatrix}
$$

$$= \cdots = a_{11}a_{22}a_{33}\cdots a_{nn} = \prod_{i=1}^{n} a_{ii}。$$

又
$$\begin{vmatrix} a_{11} & \cdots & \cdots & \cdots & a_{1n} \\ & \ddots & & & \vdots \\ 0 & & \ddots & & \vdots \\ \vdots & \ddots & & \ddots & \vdots \\ 0 & \cdots & 0 & & a_{nn} \end{vmatrix} \xlongequal{\text{性质 1}} \begin{vmatrix} a_{11} & 0 & \cdots & \cdots & 0 \\ \vdots & \ddots & \ddots & & \vdots \\ \vdots & & \ddots & \ddots & 0 \\ \vdots & & & \ddots & \\ a_{1n} & \cdots & \cdots & \cdots & a_{nn} \end{vmatrix}$$

$$= \prod_{i=1}^{n} a_{ii}。$$

命题 2.2

$$\begin{vmatrix} 0 & \cdots & \cdots & 0 & a_{1n} \\ \vdots & & \cdot^{\cdot^{\cdot}} & \cdot^{\cdot^{\cdot}} & \vdots \\ 0 & \cdot^{\cdot^{\cdot}} & \cdot^{\cdot^{\cdot}} & & \vdots \\ & \cdot^{\cdot^{\cdot}} & & & \vdots \\ a_{n1} & \cdots & \cdots & \cdots & a_{nn} \end{vmatrix} = \begin{vmatrix} a_{11} & \cdots & \cdots & \cdots & a_{1n} \\ \vdots & & & \cdot^{\cdot^{\cdot}} & 0 \\ \vdots & & \cdot^{\cdot^{\cdot}} & \cdot^{\cdot^{\cdot}} & \vdots \\ \vdots & \cdot^{\cdot^{\cdot}} & \cdot^{\cdot^{\cdot}} & & \vdots \\ a_{n1} & 0 & \cdots & \cdots & 0 \end{vmatrix}$$

$$= (-1)^{\left[\frac{n}{2}\right]} \prod_{i=1}^{n} a_{i,n+1-i},$$

这里$\left[\dfrac{n}{2}\right]$表示不大于$\dfrac{n}{2}$的最大整数，如$\left[\dfrac{7}{2}\right]=3$，$\left[\dfrac{8}{2}\right]=4$ 等。

证　把行列式
$$\begin{vmatrix} 0 & \cdots & \cdots & 0 & a_{1n} \\ \vdots & & \cdot^{\cdot^{\cdot}} & \cdot^{\cdot^{\cdot}} & \vdots \\ \vdots & \cdot^{\cdot^{\cdot}} & \cdot^{\cdot^{\cdot}} & & \vdots \\ 0 & \cdot^{\cdot^{\cdot}} & & & \vdots \\ a_{n1} & \cdots & \cdots & \cdots & a_{nn} \end{vmatrix}$$
的第 1 列与第 n 列互换，第 2 列与第 $(n-1)$ 列互换，……，得

$$\begin{vmatrix} 0 & \cdots & \cdots & 0 & \cdots & a_{1n} \\ \vdots & & & & \mathinner{\kern1mu\raise1pt{.}\kern2mu\raise4pt{.}\kern2mu\raise7pt{.}} & \vdots \\ \vdots & & \mathinner{\kern1mu\raise1pt{.}\kern2mu\raise4pt{.}\kern2mu\raise7pt{.}} & & & \vdots \\ 0 & \mathinner{\kern1mu\raise1pt{.}\kern2mu\raise4pt{.}\kern2mu\raise7pt{.}} & & & & \vdots \\ & & & & & \vdots \\ a_{n1} & \cdots & \cdots & \cdots & \cdots & a_{nn} \end{vmatrix} \xlongequal[i=1,2,\cdots,\left[\frac{n}{2}\right]]{(i)\longleftrightarrow(n+1-i)}$$

$$(-1)^{\left[\frac{n}{2}\right]}\begin{vmatrix} a_{1n} & & 0 & \cdots & \cdots & 0 \\ \vdots & \ddots & & \ddots & & \vdots \\ \vdots & & \ddots & & \ddots & \vdots \\ \vdots & & & \ddots & & 0 \\ \vdots & & & & \ddots & \\ a_{nn} & \cdots & \cdots & \cdots & \cdots & a_{n1} \end{vmatrix} = (1)^{\left[\frac{n}{2}\right]}\prod_{i=1}^{n} a_{i,n+1-i}.$$

类似可得另一等式。

注 为了方便与明了，对行列式的变换作如下记号上的规定：等号上方表示行变换，下方表示列变换。如$\xlongequal{(2)\longleftrightarrow(3)}$表示第 2 行与第 3 行互换；$\xlongequal[\frac{(2)}{4}]{}$表示第 2 列除以 4；$\xlongequal{(2)+4(3)}$表示第 3 行的 4 倍加到第 2 行，$\xlongequal[\frac{(2)+(3)}{4}]{}$表示第 3 列加到第 2 列后再除以 4，等等。

二、二线行列式

两条线以外的元素全为零的行列式常称二线行列式，有如下结果。

命题 2.3

$$\begin{vmatrix} a_1 & & & & & \\ & a_2 & & & & \\ & & \ddots & & & \\ & & & \ddots & & \\ b_1 & \cdots & \cdots & a_i & \cdots & b_n \\ & & & & \ddots & \\ & & & & & \ddots \\ & & & & & a_n \end{vmatrix} = \begin{vmatrix} a_1 & & & b_1 & & \\ & \ddots & & \vdots & & \\ & & \ddots & & & \\ & & & a_i & & \\ & & & \vdots & \ddots & \\ & & & b_n & & \ddots \;\; a_n \end{vmatrix} = \prod_{j=1}^{n} a_j,$$

两个行列式中，线外的元素全为零。

证

$$\begin{vmatrix} a_1 & & & & \\ & \ddots & & & \\ b_1 & \cdots & a_i & \cdots & b_n \\ & & & \ddots & \\ & & & & a_n \end{vmatrix} \xlongequal[\text{按 } i \text{ 列展开}]{} a_i(-1)^{i+i} \begin{vmatrix} a_1 & & & & & 0 \\ & \ddots & & & & \\ & & a_{i-1} & & & \\ & & & a_{i+1} & & \\ & & & & \ddots & \\ 0 & & & & & a_n \end{vmatrix}$$

$$= \prod_{j=1}^{n} a_j。$$

类似可得另一等式及下述命题 2.4。

命题 2.4

$$\begin{vmatrix} & & & & a_1 \\ & & & \cdot^{\cdot^{\cdot}} & \\ b_1 & \cdots & a_i & \cdots & b_n \\ & \cdot^{\cdot^{\cdot}} & & & \\ a_n & & & & \end{vmatrix} = \begin{vmatrix} & & b_1 & & a_1 \\ & & \vdots & \cdot^{\cdot^{\cdot}} & \\ & & a_i & & \\ & \cdot^{\cdot^{\cdot}} & \vdots & & \\ a_n & & b_n & & \end{vmatrix} = (-1)^{\left[\frac{n}{2}\right]} \prod_{i=1}^{n} a_i。$$

三、例

例 2.6 $\begin{vmatrix} 3 & 36 & 79 \\ 6 & 73 & 99 \\ 9 & 108 & 238 \end{vmatrix} \xlongequal[(3)-3(1)]{(2)-2(1)} \begin{vmatrix} 3 & 36 & 79 \\ 0 & 1 & * \\ 0 & 0 & 1 \end{vmatrix} = 3。$

例 2.7 $\begin{vmatrix} 1+x & 1 & 1 & 1 \\ 1 & 1-x & 1 & 1 \\ 1 & 1 & 1+y & 1 \\ 1 & 1 & 1 & 1-y \end{vmatrix}$

$$= \begin{vmatrix} 1+x & 1+0 & 1+0 & 1+0 \\ 1+0 & 1-x & 1+0 & 1+0 \\ 1+0 & 1+0 & 1+y & 1+0 \\ 1+0 & 1+0 & 1+0 & 1-y \end{vmatrix}$$

$$
=\begin{vmatrix} x & 0 & 0 & 0 \\ 0 & -x & 0 & 0 \\ 0 & 0 & y & 0 \\ 0 & 0 & 0 & -y \end{vmatrix}+\begin{vmatrix} 1 & 0 & 0 & 0 \\ 1 & -x & 0 & 0 \\ 1 & 0 & y & 0 \\ 1 & 0 & 0 & -y \end{vmatrix}
$$

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

$$
+\begin{vmatrix} x & 0 & 0 & 1 \\ 0 & -x & 0 & 1 \\ 0 & 0 & y & 1 \\ 0 & 0 & 0 & 1 \end{vmatrix}+11\text{ 个至少有两列全是 1 的行列式}
$$

$$
=x^2y^2-xy^2-xy^2+x^2y-x^2y+\sum 0=x^2y^2。
$$

例 2.8 计算 n 阶行列式 $D=\begin{vmatrix} x & & & \\ & x & a & \\ & a & \ddots & \\ & & & x \end{vmatrix}$，行列式中对角线上元素全为 x，对角线外元素全为 a。

解法 1 $D \xlongequal{(1)+(2)+\cdots+(n)}$

$$
\begin{vmatrix} x+(n-1)a & x+(n-1)a & \cdots & \cdots & x+(n-1)a \\ a & x & \cdots & \cdots & a \\ a & a & & & a \\ \vdots & \vdots & & & \vdots \\ \vdots & \vdots & & & \vdots \\ a & a & & & a \\ a & a & \cdots & \cdots & x \end{vmatrix}
$$

$$= [x + (n-1)a]\begin{vmatrix} 1 & 1 & \cdots\cdots & & 1 \\ a & x & a & \cdots\cdots & a \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \ddots & a \\ a & \cdots\cdots & & a & x \end{vmatrix}$$

$$\xlongequal[i=2,\cdots,n]{(i)-a(1)} [x + (n-1)a]\begin{vmatrix} 1 & 1 & \cdots\cdots & 1 \\ & x-a & & \\ & & \ddots & \\ & & & x-a \end{vmatrix}$$

$$\xlongequal{\text{三线行列式}} [x + (n-1)a](x-a)^{n-1}。$$

解法 2 把 D 的第 i 列化成两列之和 $\begin{pmatrix} a \\ \vdots \\ a \\ x \\ a \\ \vdots \\ a \end{pmatrix} = \begin{pmatrix} a \\ \vdots \\ a \\ \vdots \\ a \end{pmatrix} + \begin{pmatrix} 0 \\ \vdots \\ 0 \\ x-a \\ 0 \\ \vdots \\ 0 \end{pmatrix}$，由性质 3，$D$ 是 2^n 个行列式之和，其中有一个行列式不含 $\begin{pmatrix} a \\ \vdots \\ a \end{pmatrix}$ 列，有 n 个行列式仅含一列 $\begin{pmatrix} a \\ \vdots \\ a \end{pmatrix}$，余者至少含两列 $\begin{pmatrix} a \\ \vdots \\ a \end{pmatrix}$，故全为零。从而 D 的值为

$$\begin{vmatrix} x-a & 0 & \cdots & \cdots & 0 \\ 0 & \ddots & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \ddots & 0 \\ 0 & \cdots & \cdots & 0 & x-a \end{vmatrix} + \sum_{i=1}^{n} \begin{array}{c} \begin{vmatrix} x-a & & a & & \\ & \ddots & \vdots & & \\ & & a & & \\ & & \vdots & \ddots & \\ & & a & & x-a \end{vmatrix} \\ i\text{列} \end{array}$$

$$= (x-a)^n + na(x-a)^{n-1}$$

$$= (x-a)^{n-1}[x+(n-1)a]。$$

例 2.9 $\begin{vmatrix} a_1 & b_1 & & 0 \\ & \ddots & \ddots & \\ & & \ddots & b_{n-1} \\ b_n & 0 & & a_n \end{vmatrix} \xlongequal{\text{按第 1 列展开}}$

$$a_1 \begin{vmatrix} a_2 & b_2 & & \\ & \ddots & \ddots & \\ & & \ddots & b_{n-1} \\ & & & a_n \end{vmatrix} + (-1)^{n+1} b_n \begin{vmatrix} b_1 & & & \\ a_2 & \ddots & & \\ & \ddots & \ddots & \\ & & a_{n-1} & b_{n-1} \end{vmatrix}$$

$$= \prod_{i=1}^{n} a_i + (-1)^{n+1} \prod_{i=1}^{n} b_i。$$

例 2.10 求证：Vandermonde 行列式

$$D_n = \begin{vmatrix} 1 & \cdots & 1 \\ a_1 & \cdots & a_n \\ a_1^2 & \cdots & a_n^2 \\ \vdots & & \vdots \\ a_1^{n-1} & \cdots & a_n^{n-1} \end{vmatrix} = \prod_{1 \leqslant j < i \leqslant n} (a_i - a_j)。$$

证 对 n 用归纳法。

$n=2$ 时，$D_2=\begin{vmatrix}1 & 1\\ a_1 & a_2\end{vmatrix}=a_2-a_1$，等式成立。

若 $n-1$ 时等式成立。考虑 n 时情形：

(1) 如果 $a_1,\cdots,a_n$ 中有两个相等，则 D_n 中有两列相等，其值为零。又 $\prod(a_i-a_j)$ 有因子为零，其值也为零。

(2) 如果 $a_1,\cdots,a_n$ 两两不等，设

$$f(x)=\begin{vmatrix}1 & \cdots & 1 & 1\\ a_1 & \cdots & a_{n-1} & x\\ \vdots & & \vdots & \vdots\\ a_1^{n-1} & \cdots & a_{n-1}^{n-1} & x^{n-1}\end{vmatrix},$$

按第 n 列展开，可知 $f(x)$ 是首项系数为 D_{n-1} 的字母 x 的 $(n-1)$ 次多项式。由行列式的性质 4 可知，$f(x)$ 有 $(n-1)$ 个不同零点 $a_1,a_2,\cdots,a_{n-1}$，故 $f(x)=D_{n-1}(x-a_1)\cdots(x-a_{n-1})$，从而

$$\begin{aligned}D_n=f(a_n)&=\Big(\prod_{1\leqslant j<i\leqslant n-1}(a_i-a_j)\Big)(a_n-a_1)\cdots(a_n-a_{n-1})\\&=\prod_{1\leqslant j<i\leqslant n}(a_i-a_j),\end{aligned}$$

即 n 时等式成立，由归纳法，本例得证。

例 2.11 若 $a_{ij}=1$ 或 -1，$i,j=1,2,\cdots,n$。令 $D=|a_{ij}|_n$，求证：

$$2^{n-1}\mid D。$$

证 $D\xlongequal[i=2,\cdots,n]{(i)+(1)}\begin{vmatrix}a_{11} & \cdots & a_{1n}\\ b_{21} & \cdots & b_{2n}\\ \vdots & & \vdots\\ b_{n1} & \cdots & b_{nn}\end{vmatrix}$，

$b_{ij}=0,2$ 或 -2，故 $2\mid b_{ij}$，$i=2,\cdots,n$；$j=1,\cdots,n$。从而

$$D=2^{n-1}\begin{vmatrix}a_{11} & \cdots & a_{1n}\\ & \dfrac{b_{ij}}{2} & \end{vmatrix}，即\ 2^{n-1}\Big|D。$$

例 2.12 求证：

$$\begin{vmatrix} A & C \\ O & B \end{vmatrix} = \begin{cases} O & \text{若矩阵 } A \text{ 的行数 } k < \text{列数 } t, \\ |A| \cdot |B| & \text{若 } A \text{ 的行数 } k = \text{列数 } t, \end{cases}$$

这里 A 是 k 行、t 列矩阵，C 是 k 行、s 列矩阵，B 是 $(n-k)$ 行、s 列矩阵，O 是 $(n-k)$ 行、t 列的零矩阵，$n = t + s$。

证 对 k 用归纳法。

$k = 1$ 时，按第一列展开即得结论。

若 $k-1$ 时结论成立。k 时，按第一列展开行列式：

$$\begin{vmatrix} A & C \\ O & B \end{vmatrix} = \sum_{j=1}^{k} (-1)^{j+1} a_{j1} M\begin{pmatrix} j \\ 1 \end{pmatrix},$$

$a_{11}, a_{21}, \cdots, a_{k1}$ 为 A 的第一列元素。$M\begin{pmatrix} j \\ 1 \end{pmatrix}$ 是 $\begin{vmatrix} A & C \\ O & B \end{vmatrix}$ 中划掉第 j 行、第 1 列后的行列式，仍有原行列式的形状，由归纳法假设：

$$M\begin{pmatrix} j \\ 1 \end{pmatrix} = \begin{cases} O, & \text{当 } k-1 < t-1 \text{ 时} \\ M_A\begin{pmatrix} j \\ 1 \end{pmatrix} \cdot |B|, & \text{当 } k-1 = t-1 \text{ 时} \end{cases}$$

从而

$$\begin{vmatrix} A & C \\ O & B \end{vmatrix} = \begin{cases} O, & \text{当 } k < t \text{ 时} \\ \sum (-1)^{j+1} a_{j1} M_A\begin{pmatrix} j \\ 1 \end{pmatrix} \cdot |B| = |A| \cdot |B|, & \text{当 } k = t \text{ 时} \end{cases}$$

即 k 时，结论也成立。

由归纳法，本例得证。

例 2.13 求行列式值：$D = \begin{vmatrix} t_1 & a & a & \cdots & a \\ b & t_2 & a & \cdots & a \\ \vdots & \vdots & \vdots & & \vdots \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ b & b & b & \cdots & t_n \end{vmatrix}$。$(a \neq b)$

解 构造多项式 $f(x)=\begin{vmatrix} t_1+x & a+x & \cdots & a+x \\ b+x & t_2+x & \cdots & a+x \\ \vdots & \vdots & \ddots & \vdots \\ b+x & b+x & \cdots & t_n+x \end{vmatrix}$,

它的第 i 列为 $\begin{pmatrix} x \\ \vdots \\ \vdots \\ \vdots \\ \vdots \\ \vdots \\ \vdots \\ \vdots \\ x \end{pmatrix}+\begin{pmatrix} a \\ \vdots \\ a \\ \\ t_i \\ \\ b \\ \vdots \\ b \end{pmatrix}$。

由性质 3:

$$f(x)=D+\sum_{i=1}^{n}\begin{array}{c} \begin{vmatrix} t_1 & \cdots\cdots & x & \cdots\cdots & a \\ b & & x & & a \\ \vdots & & \vdots & & \vdots \\ b & & x & & a \\ b & \cdots\cdots & x & \cdots\cdots & t_n \end{vmatrix} \\ i\text{ 列} \end{array}$$

$$=D+x\sum_{i=1}^{n}\begin{array}{c} \begin{vmatrix} t_1 & \cdots\cdots & 1 & \cdots\cdots & a \\ b & \cdots\cdots & 1 & \cdots\cdots & a \\ \vdots & & \vdots & & \vdots \\ b & \cdots\cdots & 1 & \cdots\cdots & a \\ b & \cdots\cdots & 1 & \cdots\cdots & t_n \end{vmatrix} \\ i\text{ 列} \end{array},$$

即 $f(x)$是 x 的一次多项式。而

$$
f(-a)=\begin{vmatrix}
t_1-a & & & \\
 & \ddots & & 0 \\
 & & \ddots & \\
 & b-a & & \ddots \\
 & & & t_n-a
\end{vmatrix}=\prod_{i=1}^{n}(t_i-a),
$$

同理

$$
f(-b)=\prod_{i=1}^{n}(t_i-b),
$$

故

$$
\begin{aligned}
D=f(0)&=\frac{f(-a)-f(-b)}{(-a)-(-b)}[0-(-b)]+f(-b)\\
&=\frac{bf(-a)-af(-b)}{b-a}。
\end{aligned}
$$

四、三线行列式

三线行列式常可展开成两个低阶行列式的组合，由递推关系求得其值。在一些特殊情形，更有简便方法。详见下面例子。

例 2.14

$$
\begin{vmatrix}
a_1 & a_2 & \cdots & \cdots & \cdots & a_n \\
-1 & b & 0 & \cdots & \cdots & 0 \\
0 & \ddots & \ddots & \ddots & & \vdots \\
\vdots & & \ddots & \ddots & \ddots & \vdots \\
\vdots & & & \ddots & \ddots & \vdots \\
\vdots & & & & \ddots & 0 \\
0 & \cdots & \cdots & 0 & -1 & b
\end{vmatrix}
\xlongequal[(n)+\sum_{i=1}^{n-1} b^{n-i}(i)]{}
$$

$$
\begin{vmatrix}
a_1 & a_2 & \cdots & a_{n-1} & c \\
-1 & b & & & 0 \\
 & \ddots & \ddots & 0 & \vdots \\
 & & \ddots & \ddots & \vdots \\
 & 0 & & \ddots\ \ b & 0 \\
 & & & -1 & 0
\end{vmatrix},
$$

这里 $c = \sum_{i=1}^{n} a_i b^{n-i}$。

按最后一列展开得

$$原式 = (-1)^{n+1} c \begin{vmatrix} -1 & b & & \\ & \ddots & \ddots & \\ & & \ddots & b \\ & & & -1 \end{vmatrix} = (-1)^{n+1} c \cdot (-1)^{n-1}$$

$$= c = \sum_{i=1}^{n} a_i b^{n-i}。$$

例 2.15 求证：$D = \begin{vmatrix} a_1 & a_2 & \cdots & a_n \\ c_2 & b_2 & & \\ \vdots & & \ddots & \\ \vdots & & & \ddots \\ c_n & & & b_n \end{vmatrix}$

$$= a_1 b_2 \cdots b_n - \sum_{i=2}^{n} b_2 \cdots b_{i-1} (a_i c_i) b_{i+1} \cdots b_n。$$

证 当 $b_2, \cdots, b_n$ 全不为 0 时

$$D \xlongequal{(1) - \sum_{i=2}^{n} \frac{a_i}{b_i}(i)} \begin{vmatrix} a_1 - \sum_{i=2}^{n} \frac{a_i}{b_i} c_i & 0 & \cdots & 0 \\ c_2 & b_2 & & \\ \vdots & & \ddots & \\ \vdots & & & \ddots \\ c_n & \cdots & \cdots & b_n \end{vmatrix}$$

$$= \left(a_1 - \sum_{i=2}^{n} \frac{a_i}{b_i} c_i \right) b_2 \cdots b_n = 右边。$$

当 $b_2, \cdots, b_n$ 中有一个如 b_i 为 0 时

$$D \xlongequal{(1)\leftrightarrow(i)} -\begin{vmatrix} a_i & a_2\cdots a_{i-1} & & a_1 & a_{i+1}\cdots a_n \\ & & & \vdots & \\ 0 & b_2 & & \vdots & \\ \vdots & & \ddots & \vdots & \\ 0 & & b_{i-1} & \vdots & \\ \vdots & & & c_i & \\ 0 & & & \vdots & \ddots \\ \vdots & & & \vdots & \quad\ddots \\ 0 & & & \vdots & \qquad b_n \end{vmatrix} \xlongequal{\text{按第 1 列展开}}$$

$$= -a_i\begin{vmatrix} b_2 & & & \vdots & \\ & \ddots & & \vdots & \\ & & b_{i-1} & \vdots & \\ & & & c_i & \\ & & & \vdots & \ddots \\ & & & \vdots & \quad b_n \end{vmatrix} = -\, b_2\cdots b_{i-1}(a_ic_i)\cdots b_n = \text{右端。}$$

例 2.16 （1）若 $D_n=\begin{vmatrix} a_1 & b_1 & & & 0 \\ c_2 & a_2 & b_2 & & \\ & \ddots & \ddots & \ddots & \\ & & \ddots & \ddots & b_{n-1} \\ 0 & & & c_n & a_n \end{vmatrix}$，求证：

数列 $D_1,D_2,D_3,\cdots$ 满足递推关系（$D_1=a_1$）：

$$D_n = a_nD_{n-1} - c_nb_{n-1}D_{n-2}\text{。}(n \geqslant 3)$$

（2）求证：

$$E_n = \begin{vmatrix} 7 & 4 & & & 0 \\ 3 & 7 & \ddots & & \\ & \ddots & \ddots & \ddots & \\ & & \ddots & 7 & 4 \\ 0 & & & 3 & 7 \end{vmatrix} = 4^{n+1} - 3^{n+1}\text{。}(E_1 = 7)$$

证 (1) 按最后一行展开行列式：

$$D_n = (-1)^{n+n} a_n D_{n-1} + (-1)^{n+(n-1)} c_n \begin{vmatrix} a_1 & b_1 & & & 0 \\ & \ddots & \ddots & & \vdots \\ c_2 & a_2 & \ddots & & \vdots \\ & \ddots & \ddots & b_{n-3}, & 0 \\ & & \ddots & \ddots & \\ & & \ddots & a_{n-2}, & 0 \\ & & & \ddots & \\ & & & c_{n-1}, & b_{n-1} \end{vmatrix},$$

按最后一列展开第二个行列式，得

$$D_n = a_n D_{n-1} - c_n b_{n-1} D_{n-1}。$$

(2) 对 n 用归纳法：

$$E_1 = 7 = 4^{1+1} - 3^{1+1},$$

若 $E_k = 4^{k+1} - 3^{k+1}$，$(k = 1,2,\cdots,n-1)$，我们考虑 n 时的情形，由(1)

$$E_n = 7E_{n-1} - 3 \times 4E_{n-2} = 7E_{n-1} - 12E_{n-2},$$

用 $E_{n-1} = 4^n - 3^n$，$E_{n-2} = 4^{n-1} - 3^{n-1}$ 代入，得

$$E_n = 7(4^n - 3^n) - 12(4^{n-1} - 3^{n-1}) = 4^{n+1} - 3^{n+1}。$$

由归纳法，本题得证。

例 2.17

求 $D_n = \begin{vmatrix} a_1 & a_2 & \cdots & a_n \\ c_2 & b_2 & & \\ & \ddots & \ddots & \\ & & \ddots & \ddots \\ & & c_n & b_n \end{vmatrix}$ 的递推关系。

解 按最后一列展开行列式 D_n，即得递推关系：

$$D_n = b_n D_{n-1} + (-1)^{n+1} a_n \begin{vmatrix} c_2 & b_2 & & \\ & c_3 & \ddots & \\ & & \ddots & b_{n-1} \\ & & & c_n \end{vmatrix}$$

$$= b_n D_{n-1} + (-1)^{n+1} a_n c_2 c_3 \cdots c_n。$$

五、行列式计算方法小结

方法 1 用初等变换化行列式为三角行列式。

方法 2 展开成低阶行列式的组合，或得其值或得递推关系。

方法 3 化一行(或列)为两行(列)和，行列式化为同阶简易行列式(常为二线行列式)的和。

常见形式为

$$|\beta_1 + k_1\alpha, \beta_2 + k_2\alpha, \cdots, \beta_n + k_n\alpha|$$

$$= |\beta_1, \cdots, \beta_n| + \sum_{i=1}^{n} k_i |\beta_1, \cdots, \beta_{i-1}, \alpha, \beta_{i+1}, \cdots, \beta_n|$$

习　题　2

1. 若 $|A| = \begin{vmatrix} 1 & 2 & 3 & 4 \\ 0 & 6 & 1 & 7 \\ -4 & 3 & -2 & 5 \\ 1 & 0 & 3 & 3 \end{vmatrix}$，求 $M\begin{pmatrix} 1 & 4 \\ 2 & 3 \end{pmatrix}$，$M\begin{pmatrix} 1 \\ 2 \end{pmatrix}$。

2. 按定义证明：$\begin{vmatrix} x & z & y \\ y & x & z \\ z & y & x \end{vmatrix} = x^3 + y^3 + z^3 - 3xyz$。

3. 若 $\begin{vmatrix} 1 & 2 & k \\ 4 & 5 & 2 \\ 3 & 4 & 6 \end{vmatrix} = 0$，求 k 的值。

4. 求下列行列式的值：

(1) $\begin{vmatrix} a & a-1 \\ a+1 & a \end{vmatrix}$；

(2) $\begin{vmatrix} x-1 & x^2 \\ x & x^2+x+1 \end{vmatrix}$；

(3) $\begin{vmatrix} 2 & 46 & 75 \\ 4 & 93 & 134 \\ 6 & 138 & 226 \end{vmatrix}$；

(4) $\begin{vmatrix} & & & & & -1 \\ & 0 & & & & \\ & & & -1 & & \\ & -1 & & & & \\ & & & 0 & & \\ - & & & & & \end{vmatrix}$；

(5) $\begin{vmatrix} 1 & & & \\ & 2 & 2 & \\ & 2 & 3 & \\ & & & 4 \end{vmatrix}_4$；

(6) $\begin{vmatrix} 1 & 2 & 3 & 4 \\ 2 & 3 & 4 & 1 \\ 3 & 4 & 1 & 2 \\ 4 & 1 & 2 & 3 \end{vmatrix}_4$；

(7) $\begin{vmatrix} a^2 & (a+1)^2 & (a+2)^2 & (a+3)^2 \\ b^2 & (b+1)^2 & (b+2)^2 & (b+3)^2 \\ c^2 & (c+1)^2 & (c+2)^2 & (c+3)^2 \\ d^2 & (d+1)^2 & (d+2)^2 & (d+3)^2 \end{vmatrix}$；

(8) $|i+j|_4$。

5. **求证：**

$$\begin{vmatrix} a+x & x+u & u+a \\ b+y & y+v & v+b \\ c+z & z+w & w+c \end{vmatrix} = 2\begin{vmatrix} a & x & u \\ b & y & v \\ c & z & w \end{vmatrix}。$$

6. 若 α,β,γ 为方程 $x^3+px+q=0$ 的三个根，求行列式

$$\begin{vmatrix} \alpha & \beta & \gamma \\ \gamma & \alpha & \beta \\ \beta & \gamma & \alpha \end{vmatrix}。$$

7. 计算下列 n 阶行列式的值：

(1) $$\begin{vmatrix} 0 & & & & n \\ & \ddots & & & \\ 1 & & \ddots & & \\ & 2 & & \ddots & \\ & & \ddots & & \\ & & & n-1 & 0 \end{vmatrix};$$

(2) $$\begin{vmatrix} & & & a_1 & 0 \\ & & \iddots & \iddots & \\ & a_{n-1} & \iddots & & \\ 0 & \iddots & & & \\ & & & & a_n \end{vmatrix}_n;$$

(3) $$\begin{vmatrix} & & & b & a \\ & & \iddots & \iddots & \\ & \iddots & \iddots & & \\ b & \iddots & & & \\ a & & & & \\ & & & & b \end{vmatrix}_n;$$

(4) $\begin{vmatrix} a_1 & a_2 \cdots a_n & \\ b_2 & 1 & \\ \vdots & \ddots & \\ b_n & & 1 \end{vmatrix}$;

(5) $|a_i-b_j|_n$，$(n\geqslant 3)$；

(6) $\begin{vmatrix} & & & 1 \\ & 3 & 2 & \\ & & 3 & \\ & \iddots & 3 & \\ n & & & \end{vmatrix}$；$(n\geqslant 3)$；

(7) $\begin{vmatrix} a_0 & -1 & & & \\ a_1 & x & \ddots & & \\ \vdots & & \ddots & \ddots & \\ \vdots & & & \ddots & -1 \\ \vdots & & & & \\ a_{n-1} & & & & x \end{vmatrix}$;

(8) $\begin{vmatrix} x & & & \\ & \ddots & a & \\ & b & \ddots & \\ & & & x \end{vmatrix}$;

(9) $\begin{vmatrix} x_1 & a_2 & & a_n \\ a_1 & x_2 & & a_n \\ \vdots & \vdots & \cdots & \vdots \\ \vdots & \vdots & & a_n \\ a_1 & a_2 & & x_n \end{vmatrix}$;

(10) $|a_i^{n-j}\cdot b_i^{j-1}|_n$;

(11) $|1+x_i^j|_n$;

(12) $\begin{vmatrix} 1 & 2 & 3 & 4 & \cdots & n \\ 2 & 3 & 4 & 5 & \cdots n & 1 \\ 3 & 4 & 5 & \cdots\cdots & n\ 1 & 2 \\ \vdots & & & & \vdots & \\ n & 1 & 2 & \cdots\cdots & \cdots\cdots n- & 1 \end{vmatrix}$;

(13) $\begin{vmatrix} 2 & 1 & & & \\ 1 & \ddots & \ddots & & \\ & \ddots & \ddots & \ddots & \\ & & \ddots & \ddots & 1 \\ & & & 1 & 2 \end{vmatrix}_n$;

(14) $|\sin(i\alpha_j)|_n$。

8. **求证：**

(1) $\begin{vmatrix} 5 & 3 & & & \\ 2 & \ddots & \ddots & & \\ & \ddots & \ddots & \ddots & \\ & & \ddots & \ddots & 3 \\ & & & 2 & 5 \end{vmatrix}_n = 3^{n+1} - 2^{n+1}$;

(2) $\begin{vmatrix} \alpha+\beta & \alpha\beta & & & \\ 1 & \ddots & \ddots & & \\ & \ddots & \ddots & \ddots & \\ & & \ddots & \ddots & \alpha\beta \\ & & & 1 & \alpha+\beta \end{vmatrix}_n = \dfrac{\alpha^{n+1} - \beta^{n+1}}{\alpha - \beta}$;

(3) $\begin{vmatrix} \cos\alpha & & & & \\ & 2\cos\alpha & 1 & & \\ 1 & \ddots & \ddots & \ddots & \\ & \ddots & \ddots & \ddots & 1 \\ & & \ddots & 1 & 2\cos\alpha \end{vmatrix}_n = \cos n\alpha$。

9. 若 $|a_{ij}|_n = d, c \neq 0$，令 $b_{ij} = c^{i-j} a_{ij}$，求证：$|b_{ij}|_n = d$。

10. 若 A 是奇数阶行列式，且 $A' = -A$，求证：$|\boldsymbol{A}| = 0$。

11. 求多项式 $f(x) = \begin{vmatrix} x - a_1^2 & -a_1a_2 & \cdots & -a_1a_n \\ -a_2a_1 & x - a_2^2 & \cdots & -a_2a_n \\ \vdots & \vdots & & \vdots \\ -a_na_1 & -a_na_2 & \cdots & x - a_n^2 \end{vmatrix}$ 的根。

12. 若 $a_1, a_2, \cdots, a_n$，是非零整数，求证：存在 n 阶行列式 $|\boldsymbol{A}|$，它的 n^2 个元素全是整数，且

(1) $|\boldsymbol{A}|$ 的第一行元素分别为 $a_1, \cdots, a_n$；

(2) $|\boldsymbol{A}| = (a_1, \cdots, a_n)$。

第三章 矩　　阵

第二章已提及矩阵的概念。

数域 P 中 $s \times n$ 个元素 $\{a_{ij} | i = 1, \cdots, s; j = 1, \cdots, n\}$ 排成的 s 行、n 列的阵式

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1j} & \cdots & a_{1n} \\ \vdots & \vdots & & \vdots & & \vdots \\ a_{i1} & a_{i2} & \cdots & a_{ij} & \cdots & a_{in} \\ \vdots & \vdots & & \vdots & & \\ a_{s1} & a_{s2} & \cdots & a_{sj} & \cdots & a_{sn} \end{pmatrix},$$

称为数域 P 上 s 行 n 列矩阵，简称 $s \times n$ 阵。常记成 $(a_{ij})_{s \times n}$。将 1×1 的矩阵就看成 P 的元素。

通常，我们用英文大写字母表示矩阵。如

$$\boldsymbol{A} = (a_{ij})_{s \times n}$$

用 $\boldsymbol{A}_{(i)}$ 表示矩阵 $\boldsymbol{A}$ 的第 i 行，用 $\boldsymbol{A}^{(j)}$ 表示矩阵 $\boldsymbol{A}$ 的第 j 列，用 $\boldsymbol{A}(i, j)$ 表示矩阵 $\boldsymbol{A}$ 的第 i 行第 j 列的元素，即

$$\boldsymbol{A}_{(i)} = (a_{i1}, a_{i2}, \cdots, a_{in}),$$

$$\boldsymbol{A}^{(j)} = \begin{pmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{sj} \end{pmatrix},$$

$$\boldsymbol{A}(i, j) = a_{ij}。$$

例如，若 $\boldsymbol{A} = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \\ 7 & 4 & 1 & 3 \end{pmatrix}$，则

$$A_{(2)} = (5,6,7,8),$$

$$A^{(3)} = \begin{pmatrix} 3 \\ 7 \\ 1 \end{pmatrix},$$

$$A(2,4) = 8。$$

两个矩阵的行数相同，列数也相同时，就称它们是同型矩阵。如果 $A = (a_{ij})_{s\times n}$ 与 $B = (b_{ij})_{s\times n}$ 是同型矩阵，并且它们的对应元素相等，即

$$A(i,j) = a_{ij} = b_{ij} = B(i,j),$$
$$(i = 1,\cdots,s;\ j = 1,\cdots,n)$$

则称矩阵 A 与 B 相等，记作

$$A = B。$$

若矩阵 A 的行数等于列数，如同为 n，则称 A 为 n 阶方阵。

§1 矩阵的运算

一、加法

若 A 与 B 为同型矩阵，$A = (a_{ij})_{s\times n}$ $B = (b_{ij})_{s\times n}$，则称 $(a_{ij} + b_{ij})_{s\times n}$ 为 A 与 B 的和，记成 $A + B$。

易验证，矩阵的加法满足如下规律：

(1) 交换律 $A + B = B + A$;

(2) 结合律 $(A + B) + C = A + (B + C)$;

(3) 若用 O 表示所有元素为 0 的矩阵，则

$$A + O = O + A = A;$$

(4) 若 $A = (a_{ij})_{s\times n}$，记 $-A = (-a_{ij})_{s\times n}$，则

$$A + (-A) = O。$$

二、数乘

若 $A=(a_{ij})_{s\times n}$，$k\in P$，则称 $(ka_{ij})_{s\times n}$ 为数 k 与矩阵 A 的数乘，记成 kA。

易验证，矩阵的数乘运算有如下性质：

(1) $(kl)A=k(lA)$， $k,l\in P$；

(2) $(k+l)A=kA+lA$；

(3) $k(A+B)=kA+kB$；

(4) $1\cdot A=A$，$(-1)A=-A$。

例 3.1

$$3\begin{pmatrix}1&2&3\\4&5&6\end{pmatrix}-2\begin{pmatrix}-1&0&1\\2&-3&4\end{pmatrix}$$
$$=\begin{pmatrix}3&6&9\\12&15&18\end{pmatrix}-\begin{pmatrix}-2&0&2\\4&-6&8\end{pmatrix}=\begin{pmatrix}5&6&7\\8&21&10\end{pmatrix}。$$

三、乘法

若 $A=(a_{ij})_{s\times n}$，$B=(b_{ij})_{n\times t}$，则称 $C=(c_{ij})_{s\times t}$ 为 A 与 B 的积，记成 $C=AB$。这里：

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

注意：(1) 当且仅当 A 的列数等于 B 的行数时，A 与 B 才能相乘。

(2) 按乘法定义，积 $C=AB$ 的第 i 行第 j 列的元素本身是两个矩阵的积，即 $C(i,j)=c_{ij}=(a_{i1},\cdots,a_{in})\begin{pmatrix}b_{1j}\\\vdots\\b_{nj}\end{pmatrix}=A_{(i)}\cdot B^{(j)}$。

例 3.2 $(1\ 2\ 3)\begin{pmatrix}4\\5\\6\end{pmatrix}=(1\times4+2\times5+3\times6)=32$。

例 3.3 $\begin{pmatrix}4\\5\\6\end{pmatrix}(1\ 2\ 3)=\begin{pmatrix}4&8&12\\5&10&15\\6&12&18\end{pmatrix}$。

例 3.4 积 $\boldsymbol{C}=\begin{pmatrix}1&2&3\\4&5&6\end{pmatrix}\begin{pmatrix}1&2\\0&1\\5&1\end{pmatrix}$ 是 2×2 阵，

$$\boldsymbol{C}(1\ 1)=(1\ 2\ 3)\begin{pmatrix}1\\0\\5\end{pmatrix}=16,$$

$$\boldsymbol{C}(1\ 2)=(1\ 2\ 3)\begin{pmatrix}2\\1\\1\end{pmatrix}=7,$$

$$\boldsymbol{C}(2\ 1)=(4\ 5\ 6)\begin{pmatrix}1\\0\\5\end{pmatrix}=34,$$

$$\boldsymbol{C}(2\ 2)=(4\ 5\ 6)\begin{pmatrix}2\\1\\1\end{pmatrix}=19。$$

例 3.5 积 $\boldsymbol{B}=\begin{pmatrix}a_{11}&\cdots&a_{1n}\\\vdots&&\vdots\\a_{s1}&\cdots&a_{sn}\end{pmatrix}\begin{pmatrix}x_1\\\vdots\\x_n\end{pmatrix}$ 是 $s\times 1$ 阵，

$$\boldsymbol{B}(i1)=(a_{i1}\cdots a_{in})\begin{pmatrix}x_1\\\vdots\\x_n\end{pmatrix}=a_{i1}x_1+a_{i2}x_2+\cdots+a_{in}x_n,$$

$$i=1,\cdots,s$$

例 3.6 用 $\boldsymbol{E}_n$ 表示对角线上元素全为 1，对角线外元素全为 0 的 n 阶方阵，即 $\boldsymbol{E}_n=\begin{pmatrix}1&&0\\&\ddots&\\0&&1\end{pmatrix}_{n\times n}$。

若 $\boldsymbol{A}=(a_{ij})_{s\times n}$，则 $\boldsymbol{A}\boldsymbol{E}_n$ 为 $s\times n$ 阵，且它的第 i 行第 j 列元素为

$$A_{(i)}E_n^{(j)} = (a_{i1}\cdots a_{is}\cdots a_{in})\begin{pmatrix}0\\ \vdots\\ 0\\ 1\\ 0\\ \vdots\\ 0\end{pmatrix} j\text{行} = a_{ij}$$,即矩阵 AE_n 与 A 同为$s \times n$,且对应元素一一相等,故 $AE_n = A$。

同理可得 $E_sA = A$。

E_n, E_s 在乘法中起了单位元的作用,故常称它们为单位矩阵。

如果引进记号 δ_{ij}: $\delta_{ij} = \begin{cases} 1 & \text{当 } i = j \text{ 时,} \\ 0 & \text{当 } i \neq j \text{ 时,} \end{cases}$ 则 $E_n = (\delta_{ij})_{n\times n}$。

例 3.7 在 n 阶方阵中,若用 E_{ij}表示第 i 行第 j 列元素为 1,其余元素全为 0 的矩阵。则 E_{ij}中除 i 行外元素全为 0,E_{kl}中除了 l 列外元素全为 0,从而 $E_{ij}\cdot E_{kl}$ 除了第 i 行第 l 列的元素外全为 0,而它的第 i 行第 l 列元素为

$$(E_{ij}\text{ 的 } i \text{ 行}) \times \begin{pmatrix} E_{kl} \\ \text{的} \\ l \\ \text{列} \end{pmatrix} = (\underset{j\text{位}}{(0\cdots 0\ 1\ 0\cdots 0)}) \begin{pmatrix}0\\ \vdots\\ 0\\ 1\\ 0\\ \vdots\\ 0\end{pmatrix} k\text{ 位} = \delta_{jk},$$

故

$$E_{ij}E_{kl} = \delta_{jk}E_{il}。$$

例 3.8 kE_n常称数量矩阵,若 A 是$s\times n$阵,易验证

$$A(kE_n) = kA,\ (kE_s)A = kA。$$

例 3.9 对角线外元素全为 0 的矩阵称为对角阵。

$$\begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_s \end{pmatrix}(a_{ij})_{s\times n}\text{的第 } i \text{ 行第 } j \text{ 列元素为}$$

$$(0\cdots 0\ \lambda_i\ 0\cdots 0)\begin{pmatrix} a_{1j} \\ \vdots \\ a_{ij} \\ \vdots \\ a_{sj} \end{pmatrix}=\lambda_i a_{ij}\text{。即对角阵}\begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_s \end{pmatrix}$$

左乘矩阵 $\boldsymbol{A}$，则 $\boldsymbol{A}$ 的第 i 行元素均扩大了 λ_i 倍。

若对角阵 $\begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_s \end{pmatrix}$ 右乘 $\boldsymbol{A}$，则 $\boldsymbol{A}$ 的第 j 列元素均扩大了 λ_j 倍，变成了 $\lambda_j a_{ij}$。

例 3.10 若 n 阶方阵

$$\boldsymbol{H}=\begin{pmatrix} 0 & & & & \\ 1 & \ddots & & 0 & \\ 0 & \ddots & \ddots & & \\ \vdots & \ddots & \ddots & \ddots & \\ 0 & \cdots & 0 & 1 & 0 \end{pmatrix}=(\delta_{i-1,j})_{n\times n},$$

则 $\boldsymbol{H}^2=\boldsymbol{H}\boldsymbol{H}$ 的第 i 行第 j 列元素为

$$\underset{i-1\text{位}}{(0\cdots 0\ 1\ \cdots 0)}\begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} j+1\text{ 位}=\delta_{i-2,j},$$

故

$$
\boldsymbol{H}^2=\begin{bmatrix}
0 & & & & & \\
0 & \ddots & & & 0 & \\
1 & \ddots & \ddots & & & \\
0 & \ddots & \ddots & \ddots & & \\
\vdots & \ddots & \ddots & \ddots & \ddots & \\
0 & \cdots & 0 & 1 & 0 & 0
\end{bmatrix}。
$$

依次类推，可得

$$
\boldsymbol{H}^k=\begin{array}{r}
k\text{个}\left\{\begin{matrix} \\ \\ \end{matrix}\right. \\ \\ \\ \\ \\ \\
\end{array}
\begin{bmatrix}
0 & & & & & & \\
\vdots & \ddots & & & 0 & & \\
0 & & \ddots & & & & \\
1 & \ddots & & \ddots & & & \\
0 & \ddots & \ddots & & \ddots & & \\
\vdots & \ddots & \ddots & \ddots & & \ddots & \\
0 & \cdots & 0 & 1 & \underbrace{0 & \cdots & 0}_{k\text{个}}
\end{bmatrix}。
$$

例 3.11 n 阶方阵 $\boldsymbol{A}=(a_{ij})_{n\times n}$，若满足 $i>j$ 时，$a_{ij}=0$，则称 $\boldsymbol{A}$ 为上三角阵；若满足 $i<j$ 时，$a_{ij}=0$，则称 $\boldsymbol{A}$ 为下三角阵。

若 $\boldsymbol{A}$ 和 $\boldsymbol{B}=(b_{ij})_{n\times n}$ 都是上（下）三角阵，则 $\boldsymbol{AB}$ 也是上（下）三角阵。这是因为 $i>j$ 时

$$
\begin{aligned}
(\boldsymbol{AB})(i,j) &= \sum_{k=1}^{n} a_{ik}b_{kj}=\sum_{i>k} a_{ik}b_{kj}+\sum_{k\geqslant i>j} a_{ik}b_{kj} \\
&= \sum_{i>k} 0\cdot b_{kj}+\sum_{k\geqslant i>j} a_{ik}\cdot 0=0。
\end{aligned}
$$

性质 (1) $k(\boldsymbol{AB})=(k\boldsymbol{A})\boldsymbol{B}=\boldsymbol{A}(k\boldsymbol{B})$；

(2) $\boldsymbol{A}(\boldsymbol{B}+\boldsymbol{C})=\boldsymbol{AB}+\boldsymbol{AC}$，

$(\boldsymbol{A}+\boldsymbol{B})\boldsymbol{C}=\boldsymbol{AC}+\boldsymbol{BC}$（加，乘分配律）；

(3) $(\boldsymbol{AB})\boldsymbol{C}=\boldsymbol{A}(\boldsymbol{BC})$（乘法结合律）。

我们仅证(3)于下：

设 $\boldsymbol{A}=(a_{ij})_{s\times n}$，$\boldsymbol{B}=(b_{ij})_{n\times t}$，$\boldsymbol{C}=(c_{ij})_{t\times l}$。则 $(\boldsymbol{AB})\boldsymbol{C}$ 与 $\boldsymbol{A}(\boldsymbol{BC})$ 都是

$s \times l$ 阵，且

$$
\begin{aligned}
[(\boldsymbol{AB})\boldsymbol{C}](i,j) &= (\boldsymbol{AB})_{(i)}\boldsymbol{C}^{(j)} \\
&= \sum_{u=1}^{t}\left(\sum_{v=1}^{n} a_{iv}b_{vu}\right)c_{uj} \\
&= \sum_{v=1}^{n} a_{iv}\cdot\left(\sum_{u=1}^{t} b_{vu}c_{uj}\right) \\
&= \boldsymbol{A}_{(i)}(\boldsymbol{BC})^{(j)} \\
&= [\boldsymbol{A}(\boldsymbol{BC})](i,j)。
\end{aligned}
$$

故(3)得证。

从(3)的证明，不难得出下面的推论。

推论 若矩阵 $\boldsymbol{A},\boldsymbol{B},\boldsymbol{C}$ 的元素 a,b,c 满足

$$(ab)c = a(bc) \in D,$$

且 D 中元素加法满足交换律与结合律，则

$$(\boldsymbol{AB})\boldsymbol{C} = \boldsymbol{A}(\boldsymbol{BC})。$$

由
$$\begin{pmatrix} 0 & 3 \\ 0 & 7 \end{pmatrix}\begin{pmatrix} 1 & 2 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} 1 & 2 \\ 0 & 0 \end{pmatrix},$$
$$\begin{pmatrix} 1 & 2 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} 0 & 3 \\ 0 & 7 \end{pmatrix} = \begin{pmatrix} 0 & 17 \\ 0 & 0 \end{pmatrix}。$$

可知，矩阵乘法不满足交换律与消去律，且非零矩阵的积可以是零阵。

四、方阵的幂与多项式

若 $\boldsymbol{A}$ 是 n 阶方阵，归纳定义 $\boldsymbol{A}$ 的各次幂为

$$
\begin{aligned}
&\boldsymbol{A}^1 = \boldsymbol{A}, \\
&\boldsymbol{A}^2 = \boldsymbol{A}\cdot\boldsymbol{A} \\
&\boldsymbol{A}^3 = \boldsymbol{A}^2\cdot\boldsymbol{A} = (\boldsymbol{A}\cdot\boldsymbol{A})\boldsymbol{A} \\
&\qquad\vdots \qquad\qquad \vdots \\
&\boldsymbol{A}^{t+1} = \boldsymbol{A}^t\cdot\boldsymbol{A} = (\boldsymbol{A}^{t-1}\cdot\boldsymbol{A})\boldsymbol{A} = \cdots = \underbrace{\boldsymbol{A}\cdot\boldsymbol{A}\cdots\boldsymbol{A}}_{(t+1)\text{个}}。
\end{aligned}
$$

由于乘法满足结合律，我们有

$$A^t \cdot A^s = A^{t+s},$$
$$(A^t)^s = A^{ts},$$

这里 t 与 s 为正整数。若

$$f(x) = a_0 + a_1x + \cdots + a_tx^t \in P[x],$$

记

$$f(A) = a_0E_n + a_1A + \cdots + a_tA^t,$$

称它为方阵 A 的一个多项式，且如果

$$f(x) + g(x) = h(x),\ f(x)g(x) = l(x),$$

则

$$f(A) + g(A) = h(A),\ f(A) \cdot g(A) = l(A)。$$

例 3.12 若 $f(x) = x^2 - x + 2$，则

$$f\left(\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}\right) = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}^2 - \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} + 2\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}。$$

例 3.13 $(E - A)(E + A + A^2 + \cdots + A^{t-1}) = E - A^t$。这是因为 $(1 - x)(1 + x + x^2 + \cdots + x^{t-1}) = 1 - x^t$，两边用 A 代 x，用 E 代 1，即得结果。

例 3.14 若 $A = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$，$B = \begin{pmatrix} 2 & 0 \\ 1 & 2 \end{pmatrix}$，求证：

(1) $A^n = 2^{n-1}A$；

(2) $B^n = \begin{pmatrix} 2^n & 0 \\ n2^{n-1} & 2^n \end{pmatrix}$。

证 对 n 用归纳法。

$n = 1$ 时

$$A^1 = A = 2^{1-1}A,$$

$$B^1 = B = \begin{pmatrix} 2^1 & 0 \\ 1 \cdot 2^0 & 2^1 \end{pmatrix},$$

结论正确。若 $n=k-1$ 时结论正确，即

$$A^{k-1}=2^{k-2}A,\ B^{k-1}=\begin{pmatrix}2^{k-1} & 0\\ (k-1)2^{k-2} & 2^{k-1}\end{pmatrix},$$

$n=k$ 时

$$A^k=A^{k-1}A=2^{k-2}AA=2^{k-2}\begin{pmatrix}2 & 2\\ 2 & 2\end{pmatrix}=2^{k-1}\begin{pmatrix}1 & 1\\ 1 & 1\end{pmatrix}=2^{k-1}A,$$

$$B^k=B^{k-1}B=\begin{pmatrix}2^{k-1} & 0\\ (k-1)2^{k-2} & 2^{k-1}\end{pmatrix}\begin{pmatrix}2 & 0\\ 1 & 2\end{pmatrix}=\begin{pmatrix}2^k & 0\\ k2^{k-1} & 2^k\end{pmatrix},$$

结论正确。

故由归纳法，本例得证。

五、转置

定义 3.1 若 $A=(a_{ij})_{s\times n}$，记 $b_{ij}=a_{ji}$，则称 $(b_{ij})_{n\times s}$ 为 A 的转置，记成 A'。

如 $A=\begin{pmatrix}1 & 2 & 3\\ 4 & 5 & 6\end{pmatrix}$，则 $A'=\begin{pmatrix}1 & 4\\ 2 & 5\\ 3 & 6\end{pmatrix}$。

转置矩阵的运算有如下性质：

(1) $(A')'=A$；

(2) $(kA)'=kA'$；

(3) $(A+B)'=A'+B'$；

(4) $(AB)'=B'A'$。

我们仅证(4)于下：

设 $A=(a_{ij})_{s\times n}$，$B=(b_{ij})_{n\times t}$，则 $(AB)'$ 与 $B'A'$ 都是 $t\times s$ 矩阵，且

$$[(AB)'](i,j)=(AB)(j,i)=\sum_{k=1}^{n}a_{jk}b_{ki}$$

$$=\sum_{k=1}^{n}b_{ki}a_{jk}=(B')_{(i)}(A')^{(j)}=(B'A')(i,j)。$$

定义 3.2 若矩阵 A 满足 $A'=A$，则称 A 为对称阵。

若 $A=(a_{ij})_{s\times n}$ 为对称阵，则由 $A'=(a_{ji})_{n\times s}=A$ 得 $s=n$，且 $a_{ij}=a_{ji}$。

定义 3.3 如果 $A'=-A$，矩阵 A 称为反对称阵。

若 $A=(a_{ij})_{s\times n}$ 为反对称阵，由 $A'=-A$ 可得 $s=n$，且 $a_{ji}=-a_{ij}$，特别当 $i=j$ 时有 $a_{ii}=0$。

例如，矩阵 $\begin{pmatrix}1 & -4 & 5\\ -4 & 2 & 6\\ 5 & 6 & 3\end{pmatrix}$ 为对称阵；矩阵 $\begin{pmatrix}0 & 1 & -2\\ -1 & 0 & 3\\ 2 & -3 & 0\end{pmatrix}$ 为反对称阵。

§2 矩阵的初等变换、初等矩阵与矩阵的标准形

一、初等变换

定义 3.4 两行(或列)互换，非零常数乘某行(列)，一行(列)的若干倍加到另一行(列)分别称矩阵的第一，二，三类行(列)初等变换。

例如，互换矩阵 $A=\begin{pmatrix}1 & 2 & 3 & 4\\ 5 & 6 & 7 & 8\\ 9 & 7 & 5 & 3\end{pmatrix}$ 的第一行与第三行得到矩阵

$A_1=\begin{pmatrix}9 & 7 & 5 & 3\\ 5 & 6 & 7 & 8\\ 1 & 2 & 3 & 4\end{pmatrix}$，用数 (-1) 乘 A 的第二列得到矩阵

$A_2=\begin{pmatrix}1 & -2 & 3 & 4\\ 5 & -6 & 7 & 8\\ 9 & -7 & 5 & 3\end{pmatrix}$，把 A 的第一行的 2 倍加到第二行得到矩阵

$A_3=\begin{pmatrix}1 & 2 & 3 & 4\\ 7 & 10 & 13 & 16\\ 9 & 7 & 5 & 3\end{pmatrix}$。袭用计算行列式时采用的记号，$A\xrightarrow{(1)\leftrightarrow(3)}A_1$，

$A\xrightarrow[-1(2)]{}A_2$，$A\xrightarrow{(2)+2(1)}A_3$。

性质 (1) 初等变换是可逆的。即若矩阵 $\boldsymbol{A}$ 可经初等变换化为 $\boldsymbol{B}$，则 $\boldsymbol{B}$ 也可经初等变换化为 $\boldsymbol{A}$；

(2) 初等变换不改变方阵行列式的零性。即方阵 $\boldsymbol{A}$ 经初等变换化为 $\boldsymbol{B}$，则 $|\boldsymbol{A}|$ 与 $|\boldsymbol{B}|$ 同时为 0 或同时不为 0。

证 因为若 $\boldsymbol{A} \xrightarrow{(i)\leftrightarrow(j)} \boldsymbol{B}$，则 $\boldsymbol{B} \xrightarrow{(i)\leftrightarrow(j)} \boldsymbol{A}$，且 $|\boldsymbol{B}| = -|\boldsymbol{A}|$，

若 $\boldsymbol{A} \xrightarrow{c(i)} \boldsymbol{B}$，则 $\boldsymbol{B} \xrightarrow{\frac{(i)}{c}} \boldsymbol{A}$，且 $|\boldsymbol{B}| = c|\boldsymbol{A}|$ $(c \neq 0)$，

若 $\boldsymbol{A} \xrightarrow{(i)+k(j)} \boldsymbol{B}$，则 $\boldsymbol{B} \xrightarrow{(i)-k(j)} \boldsymbol{A}$，且 $|\boldsymbol{B}| = |\boldsymbol{A}|$。

二、初等矩阵

定义 3.5 对单位矩阵 $\boldsymbol{E}$ 分别作第一、二、三类初等变换所得的矩阵，分别称为第一、二、三类初等矩阵(我们用 $\boldsymbol{E}$ 泛指各阶单位阵)。

用 $\boldsymbol{I}_{ij}$ 表示互换 $\boldsymbol{E}$ 的第 i 行与第 j 行所得的第一类初等矩阵

$$\boldsymbol{I}_{ij} = \begin{array}{c} \begin{pmatrix} 1 & & & & & & \\ & \ddots & & & & & \\ & & 0 & & 1 & & \\ & & \vdots & \ddots & & & \\ & & 1 & \cdots & 0 & & \\ & & & & & \ddots & \\ & & & & & & 1 \end{pmatrix} \end{array} \begin{array}{l} \\ \\ i\text{ 行} \\ \\ j\text{ 行} \\ \\ \\ \end{array}$$

，$\boldsymbol{I}_{ij}$ 也是互换 $\boldsymbol{E}$ 的第 i 列与第 j 列所得的矩阵，且 $|\boldsymbol{I}_{ij}| = -1$。

用 $\boldsymbol{I}_i(c)$ 表示用非零数 c 乘 $\boldsymbol{E}$ 的第 i 行所得的第二类初等矩阵

$$\boldsymbol{I}_i(c) = \begin{pmatrix} 1 & & & & \\ & \ddots & & & \\ & & c & & \\ & & & \ddots & \\ & & & & 1 \end{pmatrix} \begin{array}{l} \\ \\ i\text{ 行} \\ \\ \\ \end{array}$$

，$\boldsymbol{I}_i(c)$ 也是用 c 乘 $\boldsymbol{E}$ 的 i 列所得的矩阵，且 $|\boldsymbol{I}_i(c)| = c \neq 0$。

用 $\boldsymbol{I}_{ij}(k)$ 表示把 $\boldsymbol{E}$ 的第 j 行的 k 倍加到第 i 行所得的第三类初等矩阵

$$
I_{ij}(k)=\begin{array}{cl}
\left(\begin{array}{cccccc}
1 & & & & & \\
 & \ddots & & & & \\
 & & 1 & & & \\
 & & \vdots & \ddots & & \\
 & & k & \cdots & 1 & \\
 & & & & \ddots & \\
 & & & & & 1
\end{array}\right)
&
\begin{array}{l}
\\ \\ j\text{ 行} \\ \\ i\text{ 行} \\ \\ \\
\end{array}
\\
\begin{array}{cccccc}
 & & j & & i & \\
 & & \text{列} & & \text{列} &
\end{array}
&
\end{array}
$$

，$I_{ij}(k)$ 也是把 E 的第 i 列的 k 倍加到第 j 列所得的矩阵，且 $|I_{ij}(k)|=1$。

直接验证可得：

若 $A\xrightarrow{(i)\leftrightarrow(j)}A_1$，则 $A_1=I_{ij}A$，

若 $A\xrightarrow{c(i)}A_2$，则 $A_2=I_i(c)A$，

若 $A\xrightarrow{(i)+k(j)}A_3$，则 $A_3=I_{ij}(k)A$。

故有：

命题 3.1 矩阵左乘初等阵相当于矩阵作相应的行初等变换。

类似可得：

命题 3.2 矩阵右乘初等阵相当于矩阵作相应的列初等变换。

如
$$
\begin{pmatrix}0&0&1\\0&1&0\\1&0&0\end{pmatrix}\begin{pmatrix}a_{11}&a_{12}&a_{13}&a_{14}\\a_{21}&a_{22}&a_{23}&a_{24}\\a_{31}&a_{32}&a_{33}&a_{34}\end{pmatrix}=\begin{pmatrix}a_{31}&a_{32}&a_{33}&a_{34}\\a_{21}&a_{22}&a_{23}&a_{24}\\a_{11}&a_{12}&a_{13}&a_{14}\end{pmatrix};
$$
$$
\begin{pmatrix}a_{11}&a_{12}&a_{13}&a_{14}\\a_{21}&a_{22}&a_{23}&a_{24}\\a_{31}&a_{32}&a_{33}&a_{34}\end{pmatrix}\begin{pmatrix}1&2&&\\&1&&\\&&1&\\&&&1\end{pmatrix}=\begin{pmatrix}a_{11}&2a_{11}+a_{12}&a_{13}&a_{14}\\a_{21}&2a_{21}+a_{22}&a_{23}&a_{24}\\a_{31}&2a_{31}+a_{32}&a_{33}&a_{34}\end{pmatrix}。
$$

三、矩阵的各种标准形

1. 阶梯形

若 $A=(a_{ij})_{s\times n}$ 是非零矩阵，设 A 的第一个非零列是第 j_1 列，经两

行互换可把第 j_1 列的非 0 元素换到第一行，从而 A 变成

$$\begin{array}{c}\begin{pmatrix} 0 & \cdots\cdots & 0 & b_1 & \\ \vdots & & \vdots & b_2 & * \\ \vdots & & \vdots & \vdots & \\ 0 & \cdots\cdots & 0 & b_s & \end{pmatrix} = \boldsymbol{B}, b_1 \neq 0; \\ \qquad\qquad\quad j_1 \\ \qquad\qquad\quad \text{列} \end{array}$$

$$\boldsymbol{B} \xrightarrow[\frac{1}{b_1}(1)]{(i)-\frac{b_i}{b_1}(1),\, i=2,\cdots,s} \begin{pmatrix} 0 & \cdots\cdots & 0 & 1 & * \cdots * \\ \vdots & & \vdots & 0 & \\ \vdots & & \vdots & \vdots & \boldsymbol{A}_1 \\ \vdots & & \vdots & \vdots & \\ 0 & \cdots\cdots & 0 & 0 & \end{pmatrix} = \boldsymbol{C},$$

即 $\boldsymbol{A}$ 经过若干次初等行变换化为 $\boldsymbol{C}$。若 $\boldsymbol{A}_1$ 不是零矩阵，对 $\boldsymbol{A}_1$ 作类似的行初等变换（$\boldsymbol{A}_1$ 的行数 $\leqslant s-1$）使

$$\boldsymbol{A}_1 \longrightarrow \begin{pmatrix} 0 & \cdots\cdots & 0 & 1 & * \cdots * \\ \vdots & & \vdots & 0 & \\ \vdots & & \vdots & \vdots & \boldsymbol{A}_2 \\ \vdots & & \vdots & \vdots & \\ 0 & \cdots\cdots & 0 & 0 & \end{pmatrix}。$$

若 $\boldsymbol{A}_2$ 不是零阵，对 $\boldsymbol{A}_2$ 作类似的行初等变换，……。$\boldsymbol{A}_i$ 的行数 $\leqslant s-i$，故上述过程必经有限步后终止，即存在 $i<s$ 且 $\boldsymbol{A}_i=0$ 或者 $\boldsymbol{A}_{s-1}$ 是仅有一行的非零矩阵。于是我们得到了下面的定理 3.1：

定理 3.1 任意非零矩阵可经有限次行初等变换化为阶梯形：

$$\begin{pmatrix} 0 \cdots\cdots 0 & 1 & * & * \cdots\cdots\cdots\cdots\cdots & * \\ 0 \cdots\cdots\cdots\cdots\cdots\cdots & & 0 & 1 \; * \quad \cdots\cdots & * \\ & & & \vdots & \vdots \\ 0 \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots & & & 0 \quad 1 \; * \cdot * & \\ 0 \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots & & & 0 \cdots\cdots\cdots & 0 \\ \vdots & & & \vdots & \vdots \\ 0 \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots & & & 0 \cdots\cdots & 0 \end{pmatrix} \begin{array}{l} \left.\begin{array}{l} \\ \\ \\ \\ \end{array}\right\} r \text{ 个阶梯}(r \leqslant \text{矩阵行数}) \\ \\ \\ \\ \end{array}$$

注 有时不要求阶梯形中每个阶梯的第 1 个元素为 1，而仅要求它是非零数即可。

2. 标准阶梯形

若可作列互换，则阶梯形可进一步化为

$$\boldsymbol{D}=\begin{pmatrix} 1 & c_{12} & \cdots & c_{1r} & & & \\ & \ddots & \ddots & \vdots & & & \\ & 0 & \ddots & c_{r-1,r} & & \boldsymbol{B} & \\ & & & 1 & & & \\ 0 & \cdots & & 0 & 0 & \cdots & 0 \\ \vdots & & & \vdots & \vdots & & \vdots \\ 0 & \cdots & & 0 & 0 & \cdots & 0 \end{pmatrix}\begin{matrix}\left.\begin{matrix}\\ \\ \\ \end{matrix}\right\} r\text{ 个阶梯} \\ \\ \\ \\ \\ \end{matrix},$$

把 $\boldsymbol{D}$ 的 r 行的 $(-c_{ir})$ 倍加到 i 行 $(i=1,\cdots,r-1)$，得

$$\boldsymbol{D}_1=\begin{pmatrix} 1 & & & 0 & & & \\ & 1 & * & \vdots & & & \\ & & \ddots & \vdots & & \boldsymbol{B}_1 & \\ & 0 & & 0 & & & \\ & & & 1 & & & \\ 0 & \cdots & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & & & \vdots & \vdots & & \vdots \\ 0 & \cdots & \cdots & 0 & 0 & \cdots & 0 \end{pmatrix}\begin{matrix}\left.\begin{matrix}\\ \\ \\ \end{matrix}\right\} r\text{ 个阶梯} \\ \\ \\ \\ \\ \end{matrix}。$$

仿上把 $\boldsymbol{D}_1$ 的 $(r-1)$ 行的若干倍分别加到第 $1,2,\cdots,(r-2)$ 行，得

$$\boldsymbol{D}_2=\begin{pmatrix} 1 & * & 0 & 0 & & & \\ & \ddots & \vdots & \vdots & & & \\ 0 & \ddots & 0 & \vdots & & \boldsymbol{B}_2 & \\ & & 1 & 0 & & & \\ & & & 1 & & & \\ 0 & \cdots & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & & & \vdots & \vdots & & \vdots \\ 0 & \cdots & \cdots & 0 & 0 & \cdots & 0 \end{pmatrix}\begin{matrix}\left.\begin{matrix}\\ \\ \\ \end{matrix}\right\} r\text{ 个阶梯} \\ \\ \\ \\ \\ \end{matrix}。$$

继续这个过程，即得下面定理 3.2。

定理 3.2 任意非零矩阵 $\boldsymbol{A}=(a_{ij})_{s\times n}$ 可经有限次行初等变换与列互换化为标准阶梯形

$$\begin{pmatrix} \boldsymbol{E}_r & \boldsymbol{X} \\ \boldsymbol{0} & \boldsymbol{0} \end{pmatrix},$$

$\boldsymbol{X}$ 是 r 行、$(n-r)$ 列矩阵，$r\leqslant s$。

3. 标准形

若允许作列初等变换，则标准阶梯形

$$\begin{pmatrix} \boldsymbol{E}_r & \begin{matrix} x_{11}, & \cdots & x_{1t} \\ \vdots & & \vdots \\ x_{r1}, & \cdots & x_{rt} \end{matrix} \\ 0 & 0 \end{pmatrix} \xrightarrow[\substack{j=1,\cdots,t}]{(r+j)-\sum\limits_{i=1}^{r} x_{ij}(i)} \begin{pmatrix} \boldsymbol{E}_r & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{0} \end{pmatrix}。$$

定理 3.3 任意非零矩阵可经有限次行、列初等变换化为标准形 $\begin{pmatrix} \boldsymbol{E}_r & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{0} \end{pmatrix}$。

例 $$\boldsymbol{A}=\begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 5 & 8 & 7 \\ 3 & 7 & 11 & 14 \\ 5 & 12 & 19 & 21 \end{pmatrix} \xrightarrow{\substack{(1)-(2)-(3) \\ [(3)-(1)-(2)]\times\frac{1}{3} \\ (2)-2(1)}} \begin{pmatrix} 1 & 2 & 3 & 4 \\ 0 & 1 & 2 & -1 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix} = \boldsymbol{B},$$

$\boldsymbol{B}$ 为阶梯形。

$$\boldsymbol{B} \xrightarrow[(4)\leftrightarrow(3)]{} \begin{pmatrix} 1 & 2 & 4 & 3 \\ 0 & 1 & -1 & 2 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \xrightarrow{\substack{(2)+(3) \\ (1)-4(3)}}$$

$$\begin{pmatrix}1&2&0&3\\0&1&0&2\\0&0&1&0\\0&0&0&0\end{pmatrix}\xrightarrow{(1)-2(2)}\begin{pmatrix}1&0&0&3\\0&1&0&2\\0&0&1&0\\0&0&0&0\end{pmatrix}=\boldsymbol{C},$$

$\boldsymbol{C}$ 为标准阶梯形。

$$\boldsymbol{C}\xrightarrow[(4)-3(1)-2(2)]{}\begin{pmatrix}1&0&0&0\\0&1&0&0\\0&0&1&0\\0&0&0&0\end{pmatrix}=\boldsymbol{D},\boldsymbol{D}\text{ 即为标准形。}$$

利用“初等变换相当于乘初等阵”，我们得到上述定理的乘积形式。

定理 3.4 若 $\boldsymbol{A}$ 为 $s\times n$ 阵，则存在 s 阶初等阵 $\boldsymbol{P}_1,\cdots,\boldsymbol{P}_t,\boldsymbol{Q}_1,\cdots\boldsymbol{Q}_l$ 与 n 阶初等阵 $\boldsymbol{R}_1,\cdots,\boldsymbol{R}_u$ 使

$$\boldsymbol{P}_1\cdots\boldsymbol{P}_t\boldsymbol{A}\text{ 为阶梯形，}$$
$$\boldsymbol{Q}_1\cdots\boldsymbol{Q}_l\boldsymbol{A}\boldsymbol{R}_1\cdots\boldsymbol{R}_u\text{ 为标准形。}$$

注 矩阵作初等变换，仅在它的元素间作加、减、乘、除运算，故各类标准形的元素与原矩阵的元素均在同一数域中。

§3 矩阵的秩

一、矩阵的子阵与子式

若 $\boldsymbol{B}$ 是 s 行、n 列矩阵，$\boldsymbol{B}=(b_{ij})_{s\times n}$。用 $\boldsymbol{B}\begin{pmatrix}i_1&\cdots&i_k\\j_1&\cdots&j_t\end{pmatrix}$ 表示，由 $\boldsymbol{B}$ 的第 i_1 行、…、第 i_k 行与第 j_1 列、…、第 j_t 列交叉点上 $k\cdot t$ 个元素组成的 $k\times t$ 矩阵，称为 $\boldsymbol{B}$ 的一个 $k\times t$ 子阵，$(k\leqslant s,t\leqslant n)$。用 $\boldsymbol{D}_{\boldsymbol{B}}\begin{pmatrix}i_1&\cdots&i_k\\j_1&\cdots&j_k\end{pmatrix}$ 表示 k 阶子阵 $\boldsymbol{B}\begin{pmatrix}i_1&\cdots&i_k\\j_1&\cdots&j_k\end{pmatrix}$ 的行列式，称为 $\boldsymbol{B}$ 的一个 k 阶子式。$k=1$ 时，$\boldsymbol{D}_{\boldsymbol{B}}\begin{pmatrix}i\\j\end{pmatrix}$ 表示数 b_{ij}。

如 $\boldsymbol{B}=\begin{pmatrix}1 & 2 & 3 & 4\\ 5 & 6 & 7 & 8\\ 9 & 7 & 5 & 3\end{pmatrix}$，则 $\boldsymbol{B}\begin{pmatrix} & 1 & 2\\ 2 & 3 & 4\end{pmatrix}=\begin{pmatrix}2 & 3 & 4\\ 6 & 7 & 8\end{pmatrix}$，$\boldsymbol{D}_{\boldsymbol{B}}\begin{pmatrix}1 & 2\\ 3 & 4\end{pmatrix}=$ $\begin{vmatrix}3 & 4\\ 7 & 8\end{vmatrix}=-4$。$\boldsymbol{B}$ 共有 C_4^3 个 3 阶子式，$C_4^2\times C_3^2$ 个 2 阶子式。

二、矩阵的秩

定义 3.6 非零矩阵 $\boldsymbol{A}$ 的非零子式的最大阶数称为矩阵的秩，记作 $\mathrm{r}(\boldsymbol{A})$。零矩阵的秩为 0。

例如，矩阵 $\boldsymbol{A}=\begin{pmatrix}1 & 2 & 3 & 4\\ 2 & 4 & 6 & 8\end{pmatrix}$ 的所有 2 阶子式全为 0，而 $\boldsymbol{D}_{\boldsymbol{A}}\begin{pmatrix}1\\ 1\end{pmatrix}=1\neq 0$，所以 $\boldsymbol{A}$ 的秩为 1。矩阵 $\boldsymbol{B}=\begin{pmatrix}1 & 2 & 3\\ 2 & 4 & 5\\ 3 & 6 & 8\end{pmatrix}$ 仅有一个 3 阶子式 $\begin{vmatrix}1 & 2 & 3\\ 2 & 4 & 5\\ 3 & 6 & 8\end{vmatrix}$，其值为 0，而 $\boldsymbol{D}_{\boldsymbol{B}}\begin{pmatrix}1 & 2\\ 1 & 3\end{pmatrix}=\begin{vmatrix}1 & 3\\ 2 & 5\end{vmatrix}=-1\neq 0$，故 $\boldsymbol{B}$ 的秩为 2。

定义 3.7 若 n 阶方阵 $\boldsymbol{A}$ 的秩为 n，即其唯一 n 阶子式 $|\boldsymbol{A}|\neq 0$，则称 $\boldsymbol{A}$ 为满秩矩阵。

若矩阵的秩等于其行数，则称为行满秩矩阵。若矩阵的秩等于其列数，则称为列满秩矩阵。

性质 (1) $s\times n$ 阵 $\boldsymbol{A}$ 的秩 $\leqslant \min(s,n)$；

(2) $\mathrm{r}(\boldsymbol{A})=\mathrm{r}(\boldsymbol{A}')$；

(3) 阶梯形 $\boldsymbol{A}=\begin{pmatrix}0 & \cdots & 0 & a_{1i_1} & * & \cdots\cdots & * \\ 0 & \cdots\cdots & 0 & & a_{2i_2} & *\cdots & * \\ & & & \vdots & & & \\ 0 & \cdots\cdots & 0 & & a_{ri_r} & *\cdots & * \\ 0 & \cdots\cdots & & & & & 0\end{pmatrix}$ 的秩等于阶梯个数 r，这里 $a_{ti_t}\neq 0,t=1,\cdots,r$；

(4) 矩阵经初等变换其秩不变。

证 按秩的定义,性质(1)与(2)是显然的,今证(3)与(4)于下:

(3) $\boldsymbol{A}$ 仅有 r 个非零行,$s>r$ 时,$\boldsymbol{A}$ 的任意 s 阶子式必有零行,故其值为 0。但

$$\boldsymbol{D}_A\begin{pmatrix} 1 & 2 & \cdots & r \\ i_1 & i_2 & \cdots & i_r \end{pmatrix} = \begin{vmatrix} a_{1i_1} & & \\ & \ddots & * \\ & 0 & \ddots & \\ & & & a_{ri_r} \end{vmatrix} = a_{1i_1}\cdots a_{ri_r} \neq 0$$

故 $\boldsymbol{A}$ 的秩为 r。

(4) 设 $\boldsymbol{A}$ 经一次初等变换化为 $\boldsymbol{B}$,又设 $s>\mathrm{r}(\boldsymbol{A})$。若 $\boldsymbol{A}$ 是经第一或第二类初等变换化为 $\boldsymbol{B}$,则 $\boldsymbol{B}$ 的任意一个 s 阶子式等于某个非零数乘以 $\boldsymbol{A}$ 的某个 s 阶子式,因为 $s>\mathrm{r}(\boldsymbol{A})$,所以 $\boldsymbol{A}$ 的 s 阶子式全为 0,从而 $\boldsymbol{B}$ 的 s 阶子式也全为 0,于是 $\mathrm{r}(\boldsymbol{B})\leqslant\mathrm{r}(\boldsymbol{A})$。

若 $\boldsymbol{A}$ 经第三类初等变换化为 $\boldsymbol{B}$,如 $\boldsymbol{A}\xrightarrow{(i)+k(j)}\boldsymbol{B}$,有两种可能:① $\boldsymbol{B}$ 的一个 s 阶子式不含第 i 行或同时含第 i 行与第 j 行,则它等于 $\boldsymbol{A}$ 的一个 s 阶子式,即为 0;② $\boldsymbol{B}$ 的一个 s 阶子式含 i 行,但不含 j 行,则它等于 $\boldsymbol{D}_1\pm k\boldsymbol{D}_2$,$\boldsymbol{D}_1$ 是 $\boldsymbol{A}$ 的一个含第 i 行的 s 阶子式,$\boldsymbol{D}_2$ 是 $\boldsymbol{A}$ 的含第 j 行的一个 s 阶子式,故仍为 0。由此 $\mathrm{r}(\boldsymbol{B})\leqslant\mathrm{r}(\boldsymbol{A})$。

综上所述,$\boldsymbol{A}$ 经一次初等变换得到 $\boldsymbol{B}$ 时,$\mathrm{r}(\boldsymbol{B})\leqslant\mathrm{r}(\boldsymbol{A})$。由初等变换的可逆性,$\boldsymbol{B}$ 也可经一次初等变换得到 $\boldsymbol{A}$,于是 $\mathrm{r}(\boldsymbol{A})\leqslant\mathrm{r}(\boldsymbol{B})$。从而 $\mathrm{r}(\boldsymbol{A})=\mathrm{r}(\boldsymbol{B})$。

矩阵经一次初等变换其秩不变,因而经有限次初等变换,其秩仍不变。

由性质(3)、(4)可得求秩方法:

用初等变换化矩阵为阶梯形,则阶梯个数就是矩阵的秩。

命题 3.3 矩阵的秩等于它的各类标准形中阶梯个数。

三、相抵矩阵

定义 3.8 若 $s\times n$ 阵 $\boldsymbol{A}$ 可经有限次初等变换化为 $s\times n$ 阵 $\boldsymbol{B}$,则

称 A 与 B 相抵，记成 $A \frown B$。

由于作初等变换相当于乘以初等阵，故 $A \frown B \Leftrightarrow$ 存在初等阵 $P_1, \cdots, P_t, Q_1, \cdots, Q_l$ 使 $B = P_t, \cdots, P_1 A Q_1, \cdots, Q_l$。

由初等变换性质易得：

(1) $A \frown A$；

(2) 若 $A \frown B$，则 $B \frown A$；

(3) 若 $A \frown B, B \frown C$，则 $A \frown C$。

即矩阵的相抵关系是一个等价关系。

命题 3.4 $s \times n$ 阵 A 与 B 相抵 $\Leftrightarrow \mathrm{r}(A) = \mathrm{r}(B)$。

证 $\Rightarrow A \frown B$，A 可经有限次初等变换化为 B，由于初等变换不改变矩阵的秩，故 $\mathrm{r}(A) = \mathrm{r}(B)$；

$\Leftarrow$ 若 $\mathrm{r}(A) = \mathrm{r}(B)$，则 A 与 B 有相同标准形 D，$A \frown D, B \frown D$，从而 $A \frown B$。

命题 3.5 满秩阵是初等阵的积。

证 若 A 是满秩阵，则 A 的标准形为 E，$E \frown A$，从而存在初等阵 $P_1, \cdots, P_t, Q_1, \cdots, Q_l$ 使

$$A = P_1 \cdots P_t E Q_1 \cdots Q_l = P_1 \cdots P_t \cdot Q_1 \cdots Q_l。$$

推论 满秩阵可只经行初等变换或只经列初等变换化为单位阵。

证 满秩阵 A 是初等阵 $P_1, \cdots, P_s$ 的积：

$$A = P_1 \cdots P_s = P_1 \cdots P_s E,$$

即 E 可经行初等变换化为 A，从而 A 可只经行初等变换化为 E。

类似可证，A 可只经列初等变换化为 E。

命题 3.6 列满秩阵可经行初等变换化为标准形。

证 设 $s \times n$ 阵 A 为列满秩阵，即其秩为 n，A 可经行初等变换化为阶梯形 B(共有 n 个阶梯)：

$$
\boldsymbol{B}=\begin{pmatrix}
0 & \cdots\cdots & 0 & b_{1i_1} & * & \cdots\cdots\cdots\cdots\cdots\cdots & * \\
0 & \cdots\cdots & 0 & \cdots\cdots & 0 & b_{2i_2} & * \cdots\cdots\cdots\cdots * \\
\vdots & & \vdots & & \vdots & & \vdots \\
0 & \cdots\cdots & 0 & \cdots\cdots & 0 & \cdots\cdots\ 0\ b_{ni_n} & * \cdots * \\
0 & \cdots\cdots & 0 & \cdots\cdots & 0 & \cdots\cdots\ 0 & \cdots\cdots\cdots\ 0 \\
\vdots & & \vdots & & \vdots & \vdots & \vdots \\
0 & \cdots\cdots & 0 & \cdots\cdots & 0 & \cdots\cdots\ 0 & \cdots\cdots\cdots\ 0
\end{pmatrix},
$$

$\boldsymbol{B}$ 是 $s \times n$ 阵，其列数 $= n =$ 阶梯数，于是

$$1 \leqslant i_1 < i_2 \cdots < i_n \leqslant n,$$

故必　$i_1 = 1, i_2 = 2, \cdots, i_n = n$，即

$$
\boldsymbol{B}=\begin{pmatrix}
b_{11} & & * & \\
 & b_{22} & & \\
 & 0 & \ddots & \\
 & & & b_{nn} \\
0 & & \cdots\cdots & 0 \\
\vdots & & & \vdots \\
0 & & \cdots\cdots & 0
\end{pmatrix}。
$$

仿照标准阶梯形化标准形的方法，可用行初等变换化 $\boldsymbol{B}$ 为标准形。本命题得证。

类似地，可得命题 3.7 的证明：

命题 3.7　行满秩矩阵可经列初等变换化为标准形。

四、矩阵积的秩与行列式

命题 3.8　$\mathrm{r}(\boldsymbol{AB}) \leqslant \mathrm{r}(\boldsymbol{A})$。

证　设 $\mathrm{r}(A) = s$，由定理 3.4，存在初等阵 $\boldsymbol{P}_1, \cdots, \boldsymbol{P}_t, \boldsymbol{Q}_1, \cdots, \boldsymbol{Q}_l$ 使

$$A = P_1 \cdots P_t \begin{pmatrix} E_s & 0 \\ 0 & 0 \end{pmatrix} Q_1 \cdots Q_l,$$

$$AB = P_1 \cdots P_t \begin{pmatrix} E_s & 0 \\ 0 & 0 \end{pmatrix} Q_1 \cdots Q_l B = P_1 \cdots P_t \begin{pmatrix} E_s & 0 \\ 0 & 0 \end{pmatrix} C,$$

这里 $C = Q_1 \cdots Q_l B$。从而 $\begin{pmatrix} E_s & 0 \\ 0 & 0 \end{pmatrix} C$ 可经初等变换化为 AB，所以 AB 的秩等于 $\begin{pmatrix} E_s & 0 \\ 0 & 0 \end{pmatrix} C$ 的秩，而后者的非零行最多有 s 行，于是其秩 $\leqslant s$。故 $\mathrm{r}(AB) \leqslant s = \mathrm{r}(A)$。

推论 1 $\mathrm{r}(AB) \leqslant \mathrm{r}(B)$

证 $\mathrm{r}(AB) = \mathrm{r}((AB)') = \mathrm{r}(B'A') \leqslant \mathrm{r}(B') = \mathrm{r}(B)$。

推论 2 若 A,B 均为 n 阶方阵，且 $|A| = 0$ 或 $|B| = 0$，则 $|AB| = 0$。

证 由 $|A| = 0$ 或 $|B| = 0$ 得 $\mathrm{r}(A) < n$ 或 $\mathrm{r}(B) < n$，由命题 3.6 或上述推论 1 可知 $\mathrm{r}(AB) < n$，从而 AB 的 n 阶子式 $|AB| = 0$。

引理 3.1 若 B 为 n 阶方阵，P 是 n 阶初等阵，则 $|PB| = |P||B|$。

证 (1) 若 $P = I_{ij}$，$|P| = -1$，$B \xrightarrow{(i)\leftrightarrow(j)} PB$，于是 $|PB| = -|B| = |P||B|$。

(2) 若 $P = I_i(c)$，$|P| = c$，$B \xrightarrow{c(i)} PB$，于是 $|PB| = c|B| = |P||B|$。

(3) 若 $P = I_{ij}(k)$，$|P| = 1$，$B \xrightarrow{(i)+k(j)} PB$，于是 $|PB| = |B| = |P||B|$。

定理 3.5 若 A,B 同为 n 阶方阵，则

$$|AB| = |A||B|。$$

证 (1) 若 $\mathrm{r}(A) < n$，$|A| = 0$，由命题 3.6 的推论 2 得 $|AB| = 0$，故结论成立。

(2) 若 $\mathrm{r}(A) = n$，则 A 为初等阵 $P_1, \cdots, P_t$ 的积，设为

$$A = P_1 \cdots P_t,$$

则

$$
\begin{aligned}
|A| &= |P_1| \cdot |P_2 \cdots P_t| = |P_1| \cdot |P_2| \cdot |P_3 \cdots P_t| \\
&= \cdots = |P_1| \cdots |P_t|。\\
|AB| &= |P_1 \cdots P_t B| = |P_1| \cdot |P_2 \cdots P_t B| \\
&= |P_1| \cdots |P_t| \cdot |B| = |A| \cdot |B|。
\end{aligned}
$$

由(1)、(2),定理得证。

例 3.15 若 n 阶方阵 A 的第 i 行第 j 列元素为 $\dfrac{a_i^n - b_j^n}{a_i - b_j}$,这里:$a_1, \cdots, a_n, b_1, \cdots, b_n$ 是两两不同的数,求 A 的行列式值。

解
$$
\begin{aligned}
\frac{a_i^n - b_j^n}{a_i - b_j} &= a_i^{n-1} b_j^0 + a_i^{n-2} b_j^1 + \cdots + a_i^0 b_j^{n-1} \\
&= (a_i^{n-1}, a_i^{n-2}, \cdots, a_i^1, 1) \begin{pmatrix} 1 \\ b_j \\ b_j^2 \\ \vdots \\ b_j^{n-1} \end{pmatrix}。
\end{aligned}
$$

故 $A = BC$,这里 B 的 i 行为$(a_i^{n-1}, a_i^{n-2}, \cdots, a_i, 1)$,$C$ 的 j 列为$\begin{pmatrix} 1 \\ b_j \\ \vdots \\ b_j^{n-1} \end{pmatrix}$。于是

$$
\begin{aligned}
|A| &= |B| \cdot |C| \\
&= \begin{vmatrix} a_1^{n-1} & a_1^{n-2} & \cdots & a_1 & 1 \\ a_2^{n-1} & a_2^{n-2} & \cdots & a_2 & 1 \\ \vdots & & & \vdots & \vdots \\ a_n^{n-1} & a_n^{n-2} & \cdots & a_n & 1 \end{vmatrix} \begin{vmatrix} 1 & 1 & \cdots & 1 \\ b_1 & b_2 & \cdots & b_n \\ \vdots & \vdots & & \vdots \\ b_1^{n-1} & b_2^{n-1} & \cdots & b_n^{n-1} \end{vmatrix} \\
&= (-1)^{\left[\frac{n}{2}\right]} \prod_{1 \leqslant j < i \leqslant n} (a_i - a_j) \cdot \prod_{1 \leqslant j < i \leqslant n} (b_i - b_j)。
\end{aligned}
$$

§4 可逆矩阵

一、可逆矩阵

定义 3.8 n 阶方阵 $\boldsymbol{A}$ 称为可逆阵，如果存在 n 阶方阵 $\boldsymbol{B}$ 使 $\boldsymbol{AB}=\boldsymbol{BA}=\boldsymbol{E}_n$。此时，$\boldsymbol{B}$ 称为 $\boldsymbol{A}$ 的逆。

例如 $\boldsymbol{A}=\begin{pmatrix}2&3\\1&2\end{pmatrix}$，取 $\boldsymbol{B}=\begin{pmatrix}2&-3\\-1&2\end{pmatrix}$，则 $\boldsymbol{AB}=\boldsymbol{BA}=\boldsymbol{E}_2$，故 $\boldsymbol{A}$ 可逆，$\boldsymbol{B}$ 是 $\boldsymbol{A}$ 的逆。

性质 (1) 若 $\boldsymbol{A}$ 可逆，则 $\boldsymbol{A}$ 的逆惟一，记为 $\boldsymbol{A}^{-1}$，且 $\boldsymbol{A}^{-1}$ 也可逆，$(\boldsymbol{A}^{-1})^{-1}=\boldsymbol{A}$；

(2) 若 $\boldsymbol{A}$ 可逆，$k\neq 0$，则 $k\boldsymbol{A}$ 也可逆，且 $(k\boldsymbol{A})^{-1}=\frac{1}{k}\cdot\boldsymbol{A}^{-1}$；

(3) 若 $\boldsymbol{A}$ 可逆，则 $\boldsymbol{A}'$ 也可逆，且 $(\boldsymbol{A}')^{-1}=(\boldsymbol{A}^{-1})'$；

(4) 若 $\boldsymbol{A}_1,\cdots,\boldsymbol{A}_s$ 皆可逆，则 $\boldsymbol{A}_1\cdots\boldsymbol{A}_s$ 也可逆，且 $(\boldsymbol{A}_1\cdots\boldsymbol{A}_s)^{-1}=\boldsymbol{A}_s^{-1}\cdots\boldsymbol{A}_1^{-1}$。

证 (1) 若 $\boldsymbol{B}$ 与 $\boldsymbol{C}$ 都是 $\boldsymbol{A}$ 的逆，则

$$\boldsymbol{B}=\boldsymbol{BE}=\boldsymbol{BAC}=(\boldsymbol{BA})\boldsymbol{C}=\boldsymbol{EC}=\boldsymbol{C},\text{ 即逆惟一。}$$

由 $\boldsymbol{A}\cdot\boldsymbol{A}^{-1}=\boldsymbol{A}^{-1}\boldsymbol{A}=\boldsymbol{E}$ 可知，$(\boldsymbol{A}^{-1})^{-1}=\boldsymbol{A}$。

(2) 由 $(k\boldsymbol{A})\cdot\left(\frac{1}{k}\boldsymbol{A}^{-1}\right)=\left(\frac{1}{k}\boldsymbol{A}^{-1}\right)\cdot k\boldsymbol{A}=\boldsymbol{E}$，知

$$(k\boldsymbol{A})^{-1}=\frac{1}{k}\boldsymbol{A}^{-1}。$$

(3) $(\boldsymbol{A}\cdot\boldsymbol{A}^{-1})'=\boldsymbol{E}'=\boldsymbol{E}$，故 $(\boldsymbol{A}^{-1})'\cdot\boldsymbol{A}'=\boldsymbol{E}$，

又

$$\boldsymbol{A}'\cdot(\boldsymbol{A}^{-1})'=(\boldsymbol{A}^{-1}\cdot\boldsymbol{A})'=\boldsymbol{E}'=\boldsymbol{E},$$

所以

$$(\boldsymbol{A}')^{-1}=(\boldsymbol{A}^{-1})'。$$

(4) 因为 $(\boldsymbol{A}_1\cdots\boldsymbol{A}_s)\cdot(\boldsymbol{A}_s^{-1}\cdots\boldsymbol{A}_1^{-1})$

$$= (A_s^{-1}\cdots A_1^{-1})(A_1\cdots A_s) = E,$$

所以

$$(A_1\cdots A_s)^{-1} = A_s^{-1}\cdots A_1^{-1}。$$

命题 3.9 初等阵是可逆阵且

$$(I_{ij})^{-1} = I_{ij}, (I_i(c))^{-1} = I_i\left(\frac{1}{c}\right), (I_{ij}(k))^{-1} = I_{ij}(-k)。$$

定理 3.6 若 A 是 n 阶方阵，则下列命题是彼此等价的：

(1) A 可逆；

(2) 存在 n 阶方阵 B 使 $AB = E$ 或 $BA = E$；

(3) $|A| \neq 0$；

(4) A 满秩；

(5) A 是有限个初等阵的积。

证 (1) $\Rightarrow$(2) 显然。

(2) $\Rightarrow$(3)由 $AB = E$ 或 $BA = E$，两边取行列式，由定理 3.5 得 $|A| \cdot |B| = 1$，从而 $|A| \neq 0$。

(3) $\Rightarrow$(4)由 $|A| \neq 0$ 得 A 的秩为 n，即 A 是满秩阵。

(4) $\Rightarrow$(5) 见命题 3.5。

(5) $\Rightarrow$(1) A 是有限个初等阵的积，而初等阵是可逆的，由性质(4)，A 可逆。

例 3.16 若方阵 A 满足：

$$A^2 - A = 6E,$$

求证：方阵 $A - 4E$ 可逆。

证 因为

$$\begin{aligned}(A - 4E)(A + 3E) &= A^2 - A - 12E \\ &= 6E - 12E = -6E,\end{aligned}$$

于是

$$(A - 4E)\left(\frac{A + 3E}{-6}\right) = E,$$

故

$$(A - 4E)$$

可逆。

二、伴随矩阵

设 $A = (a_{ij})_{n\times n}$，由行列式的性质(2)得

$$|A| = \sum_{j=1}^{n} a_{ij}(-1)^{i+j}M\begin{pmatrix} i \\ j \end{pmatrix} = \sum_{i=1}^{n} a_{ij}(-1)^{i+j}M\begin{pmatrix} i \\ j \end{pmatrix}$$

若记 $A_{ij} = (-1)^{i+j}M\begin{pmatrix} i \\ j \end{pmatrix}$，则

$$|A| = \sum_{j=1}^{n} a_{ij}A_{ij} = \sum_{i=1}^{n} a_{ij}A_{ij}。$$

若 $i \neq k$，令 $|B| = \begin{array}{c} \begin{vmatrix} a_{11} & \cdots\cdots & a_{1n} \\ \vdots & & \vdots \\ a_{k1} & \cdots\cdots & a_{kn} \\ \vdots & & \vdots \\ a_{k1} & \cdots\cdots & a_{kn} \\ \vdots & & \vdots \\ a_{n1} & \cdots\cdots & a_{nn} \end{vmatrix} \end{array}\begin{array}{l} \\ \\ k\text{行} \\ \\ i\text{行} \\ \\ \\ \end{array}$，即把 $|A|$ 中的 i 行用 k 行代替，$|A|$ 与 $|B|$ 除了第 i 行外全部相同，因而 $M_A\begin{pmatrix} i \\ j \end{pmatrix} = M_B\begin{pmatrix} i \\ j \end{pmatrix}$，故可记成 $M\begin{pmatrix} i \\ j \end{pmatrix}$。由于 $|B|$ 中第 k 行与第 i 行相等，因而其值为 0，又 $|B|$ 按 i 行展开得

$$\sum_{j=1}^{n} a_{kj}A_{ij} = 0, \quad (i \neq k)。$$

如果把 $i = k$ 的情形一起考虑，则得

$$\sum_{j=1}^{n} a_{kj}A_{ij} = \delta_{ki}|A|。 \tag{3.1}$$

同理

$$\sum_{i=1}^{n} a_{ik} A_{ij} = \delta_{kj} |A|。\tag{3.2}$$

定义 3.9 设 A 为 n 阶方阵，称 n 阶方阵

$$A^* = \begin{pmatrix} A_{11} & A_{12} & \cdots & A_{1n} \\ A_{21} & A_{22} & \cdots & A_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ A_{n1} & A_{n2} & \cdots & A_{nn} \end{pmatrix}'$$

为 A 的伴随矩阵。

由式(3.1)和式(3.2)可得

命题 3.10 $A \cdot A^* = A^* A = (\delta_{ij} |A|)_{n \times n} = |A| E_n$。

推论 若 A 可逆，则 $A^{-1} = \dfrac{A^*}{|A|}$。

例 $A = \begin{pmatrix} 2 & 5 \\ 3 & 7 \end{pmatrix}$，$|A| = -1$，$A_{11} = 7$，$A_{22} = 2$，$A_{12} = -3$，$A_{21} = -5$，$A^* = \begin{pmatrix} 7 & -5 \\ -3 & 2 \end{pmatrix}$，从而 $A^{-1} = -A^* = \begin{pmatrix} -7 & 5 \\ 3 & -2 \end{pmatrix}$。

三、实用求逆法

用求伴随阵的方法来求逆，计算量太大，不实用。下面我们利用初等变换与初等矩阵导出一个实用的求逆方法。

设 A 是可逆矩阵，$A^{-1} \cdot A = E$。A^{-1} 也是可逆阵，从而它可表初等矩阵的积：$A^{-1} = P_t \cdots P_1$。于是 $P_t \cdots P_1 A = E$，即 A 经过 t 次行初等变换化为 E，则在同样的行初等变换下，E 化为 $P_t \cdots P_1 E = P_t \cdots P_1 = A^{-1}$。由此，得实用求逆法。

求逆方法 用行初等变换化 $n \times 2n$ 矩阵(AE)为(EX)(即根据定理 3.4，用行初等变换化 A 为 E)，则 X 为所求 A 的逆阵。

例 3.17 求 $A = \begin{pmatrix} 2 & 3 \\ 1 & 2 \end{pmatrix}$ 的逆。

解 $(A\ E) = \begin{pmatrix} 2 & 3 & 1 & 0 \\ 1 & 2 & 0 & 1 \end{pmatrix} \xrightarrow{(1)\leftrightarrow(2)} \begin{pmatrix} 1 & 2 & 0 & 1 \\ 2 & 3 & 1 & 0 \end{pmatrix} \xrightarrow{(2)-2(1)}$

$$\begin{pmatrix} 1 & 2 & 0 & 1 \\ 0 & -1 & 1 & -2 \end{pmatrix} \xrightarrow[-1\cdot(2)]{(1)+2(2)} \begin{pmatrix} 1 & 0 & 2 & -3 \\ 0 & 1 & -1 & 2 \end{pmatrix}。$$

故 $A^{-1}=\begin{pmatrix} 2 & -3 \\ -1 & 2 \end{pmatrix}$。

例 3.18 $A=\begin{pmatrix} 0 & 1 & & & \\ & \ddots & 2 & & \\ & & \ddots & \ddots & \\ & & & \ddots & n-1 \\ n & & & & 0 \end{pmatrix}$，求 A^{-1}。

解 $(AE)=\left(\begin{array}{ccccc|cccc} 0 & 1 & & & & 1 & & & \\ & \ddots & 2 & & & & \ddots & & \\ & & \ddots & \ddots & & & & \ddots & \\ & & & \ddots & n-1 & & & & \ddots \\ n & & & & 0 & & & & 1 \end{array}\right) \xrightarrow{\substack{(n)\to(1) \\ (1)\to(2) \\ \vdots \\ (n-1)\to(n)}}$

$$\left(\begin{array}{ccccc|cccc} n & & & & & 0 & \cdots & \cdots & 0 \ \ 1 \\ & 1 & & & & 1 & 0 & & \\ & & 2 & & & & \ddots & \ddots & \\ & & & \ddots & & & & \ddots & \ddots \\ & & & & n-1 & & & 1 & 0 \end{array}\right) \xrightarrow{\substack{\frac{1}{n}(1) \\ \frac{1}{i-1}(i) \\ i=2,\cdots,n-1}}$$

$$\left(\begin{array}{ccccc|ccccc} 1 & & & & & 0 & & & & \frac{1}{n} \\ & 1 & & & & 1 & 0 & & & \\ & & 1 & & & & \frac{1}{2} & 0 & & \\ & & & \ddots & & & & \ddots & \ddots & \\ & & & & 1 & & & & \frac{1}{n-1} & 0 \end{array}\right),$$

故 $A^{-1}=\begin{pmatrix} 0 & & & & \frac{1}{n} \\ 1 & \ddots & & & \\ & \frac{1}{2} & \ddots & & \\ & & \ddots & \ddots & \\ & & & \frac{1}{n-1} & 0 \end{pmatrix}$。

方法推广 1 若$(\boldsymbol{A},\boldsymbol{B})$经行初等变换化为$(\boldsymbol{E},\boldsymbol{X})$，则

$$\boldsymbol{X}=\boldsymbol{A}^{-1}\boldsymbol{B}(\boldsymbol{A}\text{ 与 }\boldsymbol{B}\text{ 有相同行数，且 }\boldsymbol{A}\text{ 可逆})。$$

方法推广 2 若$\begin{pmatrix}\boldsymbol{A}\\ \boldsymbol{B}\end{pmatrix}$经列初等变换化为$\begin{pmatrix}\boldsymbol{E}\\ \boldsymbol{Y}\end{pmatrix}$，则

$$\boldsymbol{Y}=\boldsymbol{B}\boldsymbol{A}^{-1}(\boldsymbol{A}\text{ 与 }\boldsymbol{B}\text{ 有相同列数，且 }\boldsymbol{A}\text{ 可逆})。$$

证 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 作相同的行初等变换，若 $\boldsymbol{A}$ 化为 $\boldsymbol{E}$(即 $\boldsymbol{P}_1\cdots\boldsymbol{P}_t\boldsymbol{A}=\boldsymbol{E}$，从而 $\boldsymbol{P}_1\cdots\boldsymbol{P}_t=\boldsymbol{A}^{-1}$)，则 $\boldsymbol{B}$ 化为 $\boldsymbol{P}_1\cdots\boldsymbol{P}_t\boldsymbol{B}=\boldsymbol{A}^{-1}\boldsymbol{B}$。

若 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 作相同的列初等变换，$\boldsymbol{A}$ 化为 $\boldsymbol{E}$，即 $\boldsymbol{A}\boldsymbol{P}_1\cdots\boldsymbol{P}_t=\boldsymbol{E}$，$\boldsymbol{B}$ 化为 $\boldsymbol{B}\boldsymbol{P}_1\cdots\boldsymbol{P}_t=\boldsymbol{B}\boldsymbol{A}^{-1}$。

例 3.19 求 2 阶方阵 $\boldsymbol{X}$ 与 $\boldsymbol{Y}$ 使

$$\begin{pmatrix}2&3\\1&2\end{pmatrix}\boldsymbol{X}=\begin{pmatrix}1&4\\2&3\end{pmatrix},\boldsymbol{Y}\begin{pmatrix}2&3\\1&2\end{pmatrix}=\begin{pmatrix}1&4\\2&3\end{pmatrix}。$$

解 $\boldsymbol{X}=\begin{pmatrix}2&3\\1&2\end{pmatrix}^{-1}\begin{pmatrix}1&4\\2&3\end{pmatrix},\boldsymbol{Y}=\begin{pmatrix}1&4\\2&3\end{pmatrix}\begin{pmatrix}2&3\\1&2\end{pmatrix}^{-1}$，

$$\begin{pmatrix}2&3&1&4\\1&2&2&3\end{pmatrix}\xrightarrow{(1)\leftrightarrow(2)}\begin{pmatrix}1&2&2&3\\2&3&1&4\end{pmatrix}\xrightarrow{(2)-2(1)}$$

$$\begin{pmatrix}1&2&2&3\\0&-1&-3&-2\end{pmatrix}\xrightarrow[-1\times(2)]{(1)+2(2),}\begin{pmatrix}1&0&-4&-1\\0&1&3&2\end{pmatrix},$$

故

$$\boldsymbol{X}=\begin{pmatrix}-4&-1\\3&2\end{pmatrix},$$

$$\begin{pmatrix}2&3\\1&2\\1&4\\2&3\end{pmatrix}\xrightarrow[(2)-(1)]{}\begin{pmatrix}2&1\\1&1\\1&3\\2&1\end{pmatrix}\xrightarrow[(2)-2(1)]{(1)\leftrightarrow(2)}\begin{pmatrix}1&0\\1&-1\\3&-5\\1&0\end{pmatrix}\xrightarrow[-1\cdot(2)]{(1)+(2)}\begin{pmatrix}1&0\\0&1\\-2&5\\1&0\end{pmatrix},$$

故 $\boldsymbol{Y}=\begin{pmatrix}-2&5\\1&0\end{pmatrix}$。

§5 分块矩阵

一、矩阵的分块

为了表示和运算的方便，可把矩阵分块，从而把原矩阵看成由许多小矩阵排列而成的大矩阵。

设 $\boldsymbol{A}$ 为 $s\times n$ 阵，$\boldsymbol{A}=(a_{ij})_{s\times n}$。给 s 与 n 一个划分，即

$$s=i_1+\cdots+i_t,\quad \text{这里 } i_1,\cdots,i_t\in\mathbf{Z}^+,$$
$$n=j_1+\cdots+j_l,\quad \text{这里 } j_1,\cdots,j_l\in\mathbf{Z}^+。$$

用 $\boldsymbol{A}_{u,v}$ 表示由 $\boldsymbol{A}$ 的第 $s_{u-1}+1,\cdots$，第 $s_{u-1}+i_u$ 行与第 $n_{v-1}+1,\cdots$，第 $n_{v-1}+j_v$ 列所成 $i_u\times j_v$ 阵。这里 $s_{u-1}=i_1+\cdots+i_{u-1}$，$n_{v-1}=j_1+\cdots+j_{v-1}$。则得 $\boldsymbol{A}$ 的一个分块形式 $(\boldsymbol{A}_{uv})_{t\times l}=\begin{pmatrix}\boldsymbol{A}_{11}, & \cdots, & \boldsymbol{A}_{1l}\\ \vdots & & \\ \boldsymbol{A}_{t1}, & \cdots, & \boldsymbol{A}_{tl}\end{pmatrix}$。

如 $\boldsymbol{A}=\left(\begin{array}{cc:ccc:c}1&2&0&0&0&1\\3&4&0&0&0&2\\ \hdashline 0&0&1&2&7&4\\0&0&2&1&0&3\\ \hdashline 2&4&3&1&1&0\end{array}\right)$ 是 5×6 阵。$5=2+2+1$，$6=2+3+1$，从而对 $\boldsymbol{A}$ 作了分块，$\boldsymbol{A}$ 可看成由 9 个小矩阵为元素的大矩阵。

二、准对角矩阵

定义 3.10 若 $\boldsymbol{A}$ 为 n 阶方阵，给 n 一个划分，$n = i_1 + \cdots + i_t$，如果 $\boldsymbol{A}$ 的分块形式为$\begin{pmatrix} \boldsymbol{A}_1 & & \boldsymbol{0} \\ & \ddots & \\ \boldsymbol{0} & & \boldsymbol{A}_t \end{pmatrix}$，即对角线外的矩阵全是零矩阵，则称 $\boldsymbol{A}$ 为准对角阵。这里 $\boldsymbol{A}_k$ 为 i_k 阶方阵($k = 1,\cdots,t$)。

若 $\boldsymbol{A}$、$\boldsymbol{B}$ 都是 n 阶方阵，给 n 一个划分 $n = i_1 + \cdots + i_t$，$\boldsymbol{A}$ 与 $\boldsymbol{B}$ 的相应分块形式为$\begin{pmatrix} \boldsymbol{A}_1 & & \boldsymbol{0} \\ & \ddots & \\ \boldsymbol{0} & & \boldsymbol{A}_t \end{pmatrix}$与$\begin{pmatrix} \boldsymbol{B}_1 & & \boldsymbol{0} \\ & \ddots & \\ \boldsymbol{0} & & \boldsymbol{B}_t \end{pmatrix}$。准对角矩阵 $\boldsymbol{A}$,$\boldsymbol{B}$ 有如下运算法则：

(1) $\boldsymbol{A} + \boldsymbol{B} = \begin{pmatrix} \boldsymbol{A}_1 + \boldsymbol{B}_1 & & \boldsymbol{0} \\ & \ddots & \\ \boldsymbol{0} & & \boldsymbol{A}_t + \boldsymbol{B}_t \end{pmatrix}$；

(2) $\boldsymbol{AB} = \begin{pmatrix} \boldsymbol{A}_1\boldsymbol{B}_1 & & \boldsymbol{0} \\ & \ddots & \\ \boldsymbol{0} & & \boldsymbol{A}_t\boldsymbol{B}_t \end{pmatrix}$；

(3) $|\boldsymbol{A}| = |\boldsymbol{A}_1|\cdots|\boldsymbol{A}_t|$；

(4) 若 $\boldsymbol{A}$ 可逆，则 $\boldsymbol{A}^{-1} = \begin{pmatrix} \boldsymbol{A}_1^{-1} & & \boldsymbol{0} \\ & \ddots & \\ \boldsymbol{0} & & \boldsymbol{A}_t^{-1} \end{pmatrix}$；

(5) 若 $f(x)$ 为数域 P 上的多项式，则

$$f(\boldsymbol{A}) = \begin{pmatrix} f(\boldsymbol{A}_1) & & \boldsymbol{0} \\ & \ddots & \\ \boldsymbol{0} & & f(\boldsymbol{A}_t) \end{pmatrix}。$$

运算法则(1)可直接由加法定义得到，(3)可由§2的例2.7推出，(2)、(4)可由分块矩阵的乘法法则(见后)推出，(5)可由(1)、(2)直接推出。

三、分块矩阵的乘法

设 $A=(a_{ij})_{s\times n}$，$B=(b_{ij})_{n\times m}$，分别给 s,n,m 一个划分：

$$s=i_1+\cdots+i_t,$$
$$n=j_1+\cdots+j_l,$$
$$m=k_1+\cdots+k_r。$$

A 与 B 相应的分块形式为

$A=(A_{uv})_{t\times l}$，$B=(B_{uv})_{l\times r}$。则有

分块乘法法则 $AB=C=(C_{uv})_{t\times r}$，此处

$C_{uv}=\sum\limits_{w=1}^{l}A_{uw}B_{wv}$。即与把 A_{uw}，B_{wv} 看成数时的乘法法则相同。

证 直接验证可知，矩阵 C 和 AB 都是 $s\times m$ 阵，又设

$$i=i_1+\cdots+i_{u-1}+i',\ 1\leqslant i'\leqslant i_u,$$
$$j=j_1+\cdots+j_{v-1}+j',\ 1\leqslant j'\leqslant i_v,$$

则 $C(i,j)$ 是 $C_{u,v}$ 中第 i' 行第 j' 列的元素，即 $\sum\limits_{w=1}^{l}A_{uw}B_{wv}$ 中第 i' 行第 j' 列的元素。

$$C_{uv}(i',j')=\Big[\sum_{w=1}^{l}A_{uw}B_{wv}\Big](i',j')$$

$$=\underset{i'\text{行}}{(A_{u1},A_{u2},\cdots,A_{ul})}\underset{j'\text{列}}{\begin{pmatrix}B_{1v}\\B_{2v}\\\vdots\\B_{lv}\end{pmatrix}}$$

$$=A_{(i)}B^{(j)}=(AB)(i,j),$$

故

$$C=AB。$$

推论 1 $AB=A(B^{(1)},\cdots,B^{(m)})=(AB^{(1)},\cdots,AB^{(m)})$。

证 给 s,n,m 如下划分：

$s=s, n=n, m=1+1+\cdots+1$，由分块矩阵乘法即得本推论。

推论 2 $AB=\begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{s1} & \cdots & a_{sn} \end{pmatrix}\begin{pmatrix} \boldsymbol{B}_{(1)} \\ \vdots \\ \boldsymbol{B}_{(n)} \end{pmatrix}=\begin{pmatrix} \sum\limits_{i=1}^{n} a_{1i}\boldsymbol{B}_{(i)} \\ \vdots \\ \sum\limits_{i=1}^{n} a_{si}\boldsymbol{B}_{(i)} \end{pmatrix}$。

证 给 s,n,m 如下划分：

$$s=1+1+\cdots+1, n=1+\cdots+1, m=m$$

即得。

推论 3 $$AB=(\boldsymbol{A}^{(1)}\boldsymbol{A}^{(2)}\cdots\boldsymbol{A}^{(n)})\begin{pmatrix} b_{11} & \cdots & b_{1m} \\ \vdots & & \vdots \\ b_{n1} & \cdots & b_{nm} \end{pmatrix}$$
$$=\left(\sum_{i=1}^{n} b_{i1}\boldsymbol{A}^{(i)}, \cdots, \sum_{i=1}^{n} b_{im}\boldsymbol{A}^{(i)}\right)。$$

证 给 s,n,m 如下划分：

$$s=s, n=1+\cdots+1, m=1+1+\cdots+1$$

即得。

四、准初等变换与准初等矩阵

设 $\boldsymbol{A}$ 是一个 n 阶方阵，给 n 一个划分 $n=i_1+\cdots+i_t$，$\boldsymbol{A}$ 的相应分块形式为 $\boldsymbol{A}=\begin{pmatrix} \boldsymbol{A}_{11} & \cdots & \boldsymbol{A}_{1t} \\ \vdots & & \vdots \\ \boldsymbol{A}_{u1} & \cdots & \boldsymbol{A}_{ut} \\ \vdots & & \vdots \\ \boldsymbol{A}_{k1} & \cdots & \boldsymbol{A}_{kt} \\ \vdots & & \vdots \\ \boldsymbol{A}_{t1} & \cdots & \boldsymbol{A}_{tt} \end{pmatrix}$，$\boldsymbol{A}_{uk}$ 为 $i_n \times i_k$ 阵，特别 $\boldsymbol{A}_{uk}$ 为 i_u 阶方阵。

下列变换称行准初等变换（类似定义列准初等变换）。

(1) $A \xrightarrow{\text{准}(u)\leftrightarrow(k)} \begin{pmatrix} A_{11} & \cdots & A_{1t} \\ \vdots & & \vdots \\ A_{k1} & \cdots & A_{kt} \\ \vdots & & \vdots \\ A_{u1} & \cdots & A_{ut} \\ \vdots & & \vdots \\ A_{t1} & \cdots & A_{tt} \end{pmatrix}$;

(2) $A \xrightarrow{\text{准}C(u)} \begin{pmatrix} A_{11} & \cdots & A_{1t} \\ \vdots & & \vdots \\ CA_{u1}, & \cdots, & CA_{ut} \\ \vdots & & \vdots \\ A_{k1} & \cdots & A_{kt} \\ \vdots & & \vdots \\ A_{t1} & \cdots & A_{tt} \end{pmatrix}$,$C$ 是 i_u 阶可逆阵;

(3) $A \xrightarrow{\text{准}(u)+K(k)} \begin{pmatrix} A_{11}, \cdots\cdots \\ \vdots \\ A_{u1} + KA_{k1}, \cdots \\ \vdots \\ A_{k1}, \cdots\cdots \\ \vdots \\ A_{t1}, \cdots\cdots \end{pmatrix}$,$K$ 是 $i_u \times i_k$ 阵。

易验证准初等变换是可逆的。

定义 3.11 分块形式的单位矩阵,作一次准初等变换所得的矩阵称为准初等矩阵。

三类准初等矩阵分别为:

$$
\begin{pmatrix}
E &&&&&&\\
& \ddots &&&&&\\
&& 0 & \cdots & E &&\\
&& \vdots & E & \vdots &&\\
&& \vdots & \ddots & \vdots &&\\
&& \vdots & E & \vdots &&\\
&& E & \cdots & 0 &&\\
&&&&& \ddots &\\
&&&&&& E
\end{pmatrix};
$$

(这里为简便计,用 $\boldsymbol{E}$ 表示各阶单位阵)。

$$
\begin{pmatrix}
E &&&&&&\\
& \ddots &&&&&\\
&& E &&&&\\
&&& C &&&\\
&&&& E &&\\
&&&&& \ddots &\\
&&&&&& E
\end{pmatrix}
$$
($\boldsymbol{C}$ 是可逆阵);

$$
\begin{pmatrix}
E &&&&&&\\
& \ddots &&&&&\\
&& \ddots && K &&\\
&&& \ddots &&&\\
&&&& \ddots &&\\
&&&&& \ddots &\\
&&&&&& E
\end{pmatrix}。
$$

由分块矩阵的乘法法则,可得

命题 3.10 矩阵作行(列)准初等变换相当于左(右)乘准初等矩阵。

易见,准初等矩阵的行列式不为 0,故可逆,从而准初等变换不改变矩阵的秩。

例 3.20 若 A 是 n 阶可逆阵，D 是 s 阶方阵，则 $n+s$ 阶行列式 $\begin{vmatrix} A & B \\ C & D \end{vmatrix} = |A||D-CA^{-1}B|$。

证 $\begin{pmatrix} A & B \\ C & D \end{pmatrix} \xrightarrow{\text{准}(2)-CA^{-1}(1)} \begin{pmatrix} A & B \\ 0 & D-CA^{-1}B \end{pmatrix}$.

故

$$\begin{pmatrix} E & 0 \\ -CA^{-1} & E \end{pmatrix}\begin{pmatrix} A & B \\ C & D \end{pmatrix} = \begin{pmatrix} A & B \\ 0 & D-CA^{-1}B \end{pmatrix}。$$

两边取行列式即得结论。类似可得例 3.21。

例 3.21 若 A 是 n 阶方阵，D 是 s 阶可逆阵，则

$$\begin{vmatrix} A & B \\ C & D \end{vmatrix} = |D||A-BD^{-1}C|。$$

由例 3.20 与 3.21 可得下述命题：

命题 3.11 若 A 与 D 分别是 n 阶与 s 阶可逆矩阵，B 与 C 分别是 $n\times s$ 与 $s\times n$ 矩阵，则

$$|A-BD^{-1}C| = \frac{|A|}{|D|}|D-CA^{-1}B|。$$

在命题 3.11 中，令 $D=-1(s=1)$，$B=\beta$，$C=\gamma$，这里 β,γ 都是 $n\times 1$ 阵，则得：

命题 3.12 若 A 是 n 阶方阵；β 与 γ 都是 $n\times 1$ 阵，则

$$|A+\beta\cdot\gamma| = |A|(1+\gamma' A\beta)。$$

在命题 3.12 中，令 $A=\lambda E$ $(\lambda\neq 0)$，则得下面推论。

推论 若 $\lambda\neq 0$，β,γ 为 $n\times 1$ 阵，则

$$|\lambda E+\beta\gamma'| = \lambda^{n-1}(\lambda+\gamma'\beta).$$

因为等式两端都是 λ 的 n 次多项式，对非 0 的 λ 都成立，故必是恒等式(否则最多对 n 个不同的 λ 值等式成立——n 次多项式最多只有 n 个不同零点)，即 $\lambda=0$ 时，结论依然成立。

例 3.22 若 n 阶方阵

$$A=\begin{pmatrix} 2+1\cdot 1 & 1\cdot 2 & 1\cdot 3 & \cdots\cdots & 1\cdot n \\ 2\cdot 1 & 2+2\cdot 2 & 2\cdot 3 & \cdots\cdots & 2\cdot n \\ \vdots & & & \vdots & \vdots \\ n\cdot 1 & n\cdot 2 & n\cdot 3 & \cdots\cdots & 2+n\cdot n \end{pmatrix},$$

求 A 的行列式的值。

解 $A=2E+\begin{pmatrix}1\\2\\\vdots\\n\end{pmatrix}(1\ 2\cdots n)$，由命题 3.12 的推论得

$$|A|=2^{n-1}\left[2+(1\ 2\cdots n)\begin{pmatrix}1\\2\\\vdots\\n\end{pmatrix}\right]$$

$$=2^{n-1}\left[2+\sum_{k=1}^{n}k^2\right]。$$

例 3.23 求 $2n$ 阶行列式

$$\begin{vmatrix} 2E_n & E_n \\ E_n & 2E_n \end{vmatrix}。$$

解 由命题 3.10 得

$$\begin{vmatrix} 2E_n & E_n \\ E_n & 2E_n \end{vmatrix}=|2E_n|\left|2E_n-E_n\frac{E_n}{2}E_n\right|$$

$$=2^n\left|\frac{3}{2}E_n\right|=2^n\left(\frac{3}{2}\right)^n=3^n。$$

例 3.24 若 A 与 B 都是 n 阶方阵，求证：

$$\begin{vmatrix} A+B & A-B \\ A-B & A+B \end{vmatrix}=4^n|A||B|。$$

证 用准初等变换化 $2n$ 阶矩阵 $\begin{pmatrix} A+B & A-B \\ A-B & A+B \end{pmatrix}$ 为准对角阵：

$$\begin{pmatrix} A+B & A-B \\ A-B & A+B \end{pmatrix} \xrightarrow{\text{准}(1)+(2)} \begin{pmatrix} 2A & 2A \\ A-B & A+B \end{pmatrix}$$

$$\xrightarrow[\text{准}(2)-(1)]{} \begin{pmatrix} 2A & \boldsymbol{0} \\ A-B & 2B \end{pmatrix}$$

相应的矩阵等式为

$$\begin{pmatrix} E_n & E_n \\ \boldsymbol{0} & E_n \end{pmatrix} \begin{pmatrix} A+B & A-B \\ A-B & A+B \end{pmatrix} \begin{pmatrix} E_n & -E_n \\ \boldsymbol{0} & E_n \end{pmatrix} = \begin{pmatrix} 2A & \boldsymbol{0} \\ A-B & 2B \end{pmatrix},$$

两边取行列式，得

$$\begin{vmatrix} A+B, & A-B \\ A-B, & A+B \end{vmatrix} = \begin{vmatrix} 2A & \boldsymbol{0} \\ A-B & 2B \end{vmatrix} = |2A|\,|2B|$$
$$= 2^n|A| \cdot 2^n|B| = 4^n|A|\,|B|。$$

习　题　3

1. 设 $A = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \\ 0 & 7 & 4 \end{pmatrix}$，$B = \begin{pmatrix} 2 & -1 & 1 \\ 1 & 0 & 1 \\ 3 & 5 & 0 \end{pmatrix}$，求 $AB - BA$。

2. 若 n 阶方阵 A 的行列式值 $|A| = d$，求 $|kA|$。

3. 计算：

(1) $(1\ 2\ 3)\begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}$；

(2) $\begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}(1\ 2\ 3)$；

(3) $\begin{pmatrix} 2 & 1 & 1 \\ 3 & 1 & 0 \\ 0 & 1 & 2 \end{pmatrix}^2$；

(4) $(x\ y\ 1)\begin{pmatrix} a & b & d \\ b & c & e \\ d & e & f \end{pmatrix}\begin{pmatrix} x \\ y \\ 1 \end{pmatrix}$；

(5) $\begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix}^n$；

(6) $\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}^n$；

(7) $\begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix}^n$；

(8) $\begin{pmatrix} \cos\varphi, & -\sin\varphi \\ \sin\varphi, & \cos\varphi \end{pmatrix}^n$；

(9) $\begin{pmatrix} 1 & -1 & -1 & -1 \\ -1 & 1 & -1 & -1 \\ -1 & -1 & 1 & -1 \\ -1 & -1 & -1 & 1 \end{pmatrix}^n$

(10) $\begin{pmatrix} a & -b \\ b & a \end{pmatrix}^n$。

4. 若 $\boldsymbol{A},\boldsymbol{B}$ 为同阶方阵，且 $\boldsymbol{AB}=\boldsymbol{BA}$，试用归纳法证明：

$$(\boldsymbol{A}+\boldsymbol{B})^n=\sum_{k=0}^{n}C_n^k\boldsymbol{A}^{n-k}\boldsymbol{B}^k,$$

由此求 $\begin{pmatrix} \lambda & 0 & 0 \\ 1 & \lambda & 0 \\ 0 & 1 & \lambda \end{pmatrix}^n$ 与 $\begin{pmatrix} a & b \\ b & a \end{pmatrix}^n$。

5. 若 $f(x)=x^2-4x+1$，求 $f\left(\begin{pmatrix} 2 & 1 \\ 3 & 2 \end{pmatrix}\right)$。

6. 设 $A = \begin{pmatrix} a & & & 0 \\ & a_2 & & \\ & & \ddots & \\ 0 & & & a_n \end{pmatrix}$, $a_1, a_2, \cdots, a_n$ 两两不等。若 n 阶方阵 B 与 A 可交换，即 $AB = BA$，求证：B 也是对角矩阵。

7. 若 n 阶方阵 A 与所有 n 阶方阵可交换，求证：A 是数量矩阵，即存在数 k 使 $A = kE_n$。

8. 若 $A = \dfrac{1}{2}(B - E)$，求证：$A^2 = A \Leftrightarrow B^2 = E$。

9. 求证：

(1) (反)对称矩阵的和、差与数乘仍是(反)对称矩阵；

(2) 对称矩阵的幂与多项式仍是对称矩阵；

(3) 若 A 为方阵，则 $A + A'$, $A'A$ 为对称矩阵，$A - A'$ 为反对称矩阵。

10. 若 a 是 $n \times 1$ 阵，$A = E_n - 2aa'$, $a'a = 1$，求证：A 是对称阵，且 $A^2 = E$。

11. 若实对称 A 满足 $A^2 = 0$，求证：$A = 0$。

12. 求证：奇数阶反对称阵的行列式为 0。

13. 求证：任意矩阵可表示为对称阵与反对称阵之和。

14. 求 n 阶行列式 $|(a_i + b_j)^k|_n$, $k = 1, 2, \cdots, n-1$。

15. 若 $A = (a_{ij})_{2n \times 2n}$, $B = (b_{ij})_{2n \times 2n}$，且

$$b_{ij} = \sum_{k=1}^{n} \begin{vmatrix} a_{i,2k-1} & a_{j,2k-1} \\ a_{i,2k} & a_{j,2k} \end{vmatrix},$$

求证：$|B| = |A^2|$。

16. 设 $s_k = \sum_{i=1}^{n} x_i^k$, $a_{ij} = s_{i+j-2}$，求证：

$$|a_{ij}|_n = \prod_{1 \leqslant j < i \leqslant n} (x_i - x_j)^2。$$

17. 求 n 阶行列式 $\begin{vmatrix} a_1 & a_2 & \cdots\cdots & & a_n \\ a_n & a_1 & a_2, & \cdots & a_{n-1} \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \ddots & a_2 \\ a_2 & \cdots & a_n & & a_1 \end{vmatrix}$ 的值。

18. 求下列矩阵的秩：

(1) $\begin{pmatrix} 1 & 1 & 1 & 1 & 1 \\ 1 & 2 & 2 & 2 & 2 \\ 2 & 3 & 3 & 3 & 3 \\ 3 & 6 & 6 & 6 & 6 \end{pmatrix}$；

(2) $\begin{pmatrix} 1 & -1 & 2 & 1 & 0 \\ 3 & -2 & 4 & -2 & 0 \\ 3 & 0 & 6 & -1 & 1 \\ 0 & 3 & 0 & 0 & 1 \end{pmatrix}$；

(3) $\begin{pmatrix} 1 & -1 & & & \\ & 1 & \ddots & & \\ & & \ddots & \ddots & \\ & & & \ddots & -1 \\ -1 & & & & 1 \end{pmatrix}_n$。

19. 若 $\boldsymbol{A}$ 为方阵，多项式 $f(x)$ 满足 $f(0)=0$，求证：

$$\mathrm{r}(f(\boldsymbol{A})) \leqslant \mathrm{r}(\boldsymbol{A})。$$

20. 分别用行初等变换和行、列初等变换化下列矩阵为阶梯形与标准形：

(1) $\begin{pmatrix} 2 & 5 & 1 \\ 1 & 3 & 4 \\ 3 & 8 & 5 \end{pmatrix}$；

(2) $\begin{pmatrix} 1 & 1 & -1 & 1 \\ 0 & 2 & 2 & 2 \\ 1 & -1 & 0 & 3 \end{pmatrix}$；

(3) $\begin{pmatrix} 1 & -1 & & & & a_1 \\ & \ddots & \ddots & & & \vdots \\ & & \ddots & \ddots & & \vdots \\ & & & \ddots & -1 & \vdots \\ -1 & & & & 1 & a_n \end{pmatrix}$。

21. 求证：$\mathrm{r}\begin{pmatrix} \boldsymbol{A} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{B} \end{pmatrix} = \mathrm{r}(\boldsymbol{A}) + \mathrm{r}(\boldsymbol{B})$。

22. 求证：$\mathrm{r}\begin{pmatrix} \boldsymbol{A} & \boldsymbol{C} \\ \boldsymbol{0} & \boldsymbol{B} \end{pmatrix} \geqslant \mathrm{r}(\boldsymbol{A}) + \mathrm{r}(\boldsymbol{B})$。

23. 若 $\boldsymbol{A}$ 是 $s \times n$ 阵，$\boldsymbol{B}$ 是 $n \times s$ 阵，且 $\mathrm{r}(\boldsymbol{B}) = n$，$\boldsymbol{AB} = 0$，求证：$\boldsymbol{A} = 0$。

24. 若 $s \times n$ 阵 $\boldsymbol{A}$ 的秩为 r，求证：存在 $s \times r$ 阵 $\boldsymbol{L}$ 与 $r \times n$ 阵 $\boldsymbol{R}$，两者的秩均为 r. 且 $\boldsymbol{A} = \boldsymbol{LR}$。

25. 若方阵 $\boldsymbol{A}$ 的秩为 1，求证：存在数 k 使 $\boldsymbol{A}^2 = k\boldsymbol{A}$。

26. 求下列矩阵的逆：

(1) $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$，$ad - bc = 1$；

(2) $\begin{pmatrix} 1 & -1 & 0 \\ -1 & 2 & 1 \\ 2 & 2 & 3 \end{pmatrix}$；

(3) $\begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & 1 & -1 \\ 1 & -1 & -1 & 1 \end{pmatrix}$；

(4) $\begin{pmatrix} 1 & 3 & 1 & -4 \\ 0 & 1 & 2 & -3 \\ 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 1 \end{pmatrix}$；

(5) $\begin{pmatrix} 0 & & & & & n \\ 1 & \ddots & & & & \\ & 2 & \ddots & & & \\ & & 3 & \ddots & & \\ & & & \ddots & \ddots & \\ & & & & n-1 & 0 \end{pmatrix}$；

(6) $\begin{pmatrix} 1 & & & \\ & \ddots & \boldsymbol{0} & \\ & & \ddots & \\ & 1 & & \ddots & \\ & & & & 1 \end{pmatrix}_{n\times n}$。

27. 若方阵 $\boldsymbol{A}$ 满足 $\boldsymbol{A}^{100}=0$，求证：$\boldsymbol{E}-\boldsymbol{A}$ 可逆。

28. 若方阵 $\boldsymbol{A}$ 满足 $\boldsymbol{A}^2-5\boldsymbol{A}+6\boldsymbol{E}=0$，且 $\boldsymbol{A}+k\boldsymbol{E}$ 可逆，求数 k 的范围。

29. 求矩阵 $\boldsymbol{X}$，若

(1) $\begin{pmatrix} 2 & 3 \\ 3 & 5 \end{pmatrix}\boldsymbol{X}=\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$；

(2) $\boldsymbol{X}\begin{pmatrix} 2 & 3 \\ 3 & 5 \end{pmatrix}=\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$；

(3) $\begin{pmatrix} 1 & 1 & -1 \\ 0 & 2 & 2 \\ 1 & -1 & 0 \end{pmatrix}\boldsymbol{X}=\begin{pmatrix} 1 & -1 & 1 \\ 1 & 1 & 0 \\ 2 & 1 & 1 \end{pmatrix}$；

(4) $\begin{pmatrix} 1 & & & \\ & \ddots & & 1 \\ & & \ddots & \\ & 0 & & \ddots & \\ & & & & 1 \end{pmatrix}_{n\times n}\boldsymbol{X}=\begin{pmatrix} 2 & 1 & & & \\ 1 & \ddots & \ddots & & \\ & \ddots & \ddots & \ddots & \\ & & \ddots & \ddots & 1 \\ & & & 1 & 2 \end{pmatrix}_{n\times n}$。

30. 已知 $\boldsymbol{A}^{-1}$，$\boldsymbol{C}^{-1}$，求矩阵 $\begin{pmatrix} \boldsymbol{0} & \boldsymbol{A} \\ \boldsymbol{C} & \boldsymbol{0} \end{pmatrix}$ 的逆。

31. 求证：(1) 对称阵的逆仍是对称阵；

(2) 上(下)三角阵的逆仍是上(下)三角阵。

32. 若 $\boldsymbol{A}$ 是 n 阶方阵，$n \geqslant 2$，求证：

$$|\boldsymbol{A}^*| = |\boldsymbol{A}|^{n-1}。$$

33. 若 $\boldsymbol{A}$ 是 n 阶方阵，$n \geqslant 2$，求证：

$$\mathrm{r}(\boldsymbol{A}^*) = \begin{cases} n, \text{当 } \mathrm{r}(\boldsymbol{A}) = n \text{ 时}, \\ 1, \text{当 } \mathrm{r}(\boldsymbol{A}) = n-1 \text{ 时}, \\ 0, \text{当 } \mathrm{r}(\boldsymbol{A}) < n-1 \text{ 时}。 \end{cases}$$

34. 设 $\boldsymbol{A},\boldsymbol{B},\boldsymbol{C},\boldsymbol{D}$ 都是 n 阶方阵，$\boldsymbol{A}$ 可逆，$|\boldsymbol{A}| = d$，$\boldsymbol{AC} = \boldsymbol{CA}$，求证：$\begin{vmatrix} \boldsymbol{A} & \boldsymbol{B} \\ \boldsymbol{C} & \boldsymbol{D} \end{vmatrix} = |\boldsymbol{AD} - \boldsymbol{CB}|$，并求 $2n$ 阶行列式 $\begin{vmatrix} 3\boldsymbol{A} & \boldsymbol{A} \\ \boldsymbol{A} & 3\boldsymbol{A} \end{vmatrix}$ 的值。

35. 若 $\boldsymbol{A},\boldsymbol{B}$ 都是 n 阶方阵，求证：

$$\begin{vmatrix} \boldsymbol{A} & \boldsymbol{B} \\ \boldsymbol{B} & \boldsymbol{A} \end{vmatrix} = |\boldsymbol{A} + \boldsymbol{B}||\boldsymbol{A} - \boldsymbol{B}|。$$

36. 设 $\boldsymbol{A}$ 是 $s \times n$ 阵，$\boldsymbol{B}$ 是 $n \times s$ 阵，$s \geqslant n$，$\lambda \neq 0$，求证：$|\lambda \boldsymbol{E}_s - \boldsymbol{AB}| = \lambda^{s-n}|\lambda \boldsymbol{E}_n - \boldsymbol{BA}|$。

37. 若 $\boldsymbol{A}$ 是 n 阶可逆阵，$\boldsymbol{\alpha},\boldsymbol{\beta}$ 都是 $n \times 1$ 阵，且 $\boldsymbol{\beta}'\boldsymbol{A}^{-1}\boldsymbol{\alpha} + 1 \neq 0$，求证：$\boldsymbol{A} + \boldsymbol{\alpha\beta}'$ 可逆。

38. 若 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 都是 n 阶方阵，求证：

$$\mathrm{r}(\boldsymbol{AB}) \geqslant \mathrm{r}(\boldsymbol{A}) + \mathrm{r}(\boldsymbol{B}) - n。$$

第四章　线性方程组

线性方程组

$$\begin{cases} a_{11}x_1 + \cdots + a_{1j}x_j + \cdots + a_{1n}x_n = b_1 \\ \vdots \qquad\qquad \vdots \qquad\qquad \vdots \qquad \vdots \\ a_{i1}x_1 + \cdots + a_{ij}x_j + \cdots + a_{in}x_n = b_i \\ \vdots \qquad\qquad \vdots \qquad\qquad \vdots \qquad \vdots \\ a_{s1}x_1 + \cdots + a_{sj}x_j + \cdots + a_{sn}x_n = b_s \end{cases} \tag{4.1}$$

称为 s 式 n 元线性方程组。这里 $x_1,\cdots,x_n$ 是未知量，$\{a_{ij}|i=1,\cdots,s;j=1,\cdots,n\}$ 是未知量的系数，a_{ij} 是第 i 式中未知量 x_j 的系数，$b_1,\cdots,b_s$ 称为常数项。

记 $\boldsymbol{A}=(a_{ij})_{s\times n}$，$\boldsymbol{A}$ 称为式(4.1)的系数矩阵。$\boldsymbol{X}=\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$，$\boldsymbol{\beta}=\begin{pmatrix} b_1 \\ \vdots \\ b_s \end{pmatrix}$，则式(4.1)的矩阵形式为

$$\boldsymbol{AX}=\boldsymbol{\beta}。\tag{4.2}$$

若记 $\boldsymbol{\alpha}_i=\begin{pmatrix} a_{1i} \\ a_{2i} \\ \vdots \\ a_{si} \end{pmatrix}$，$i=1,\cdots,n$，则式(4.1)又可表示成

$$x_1\boldsymbol{\alpha}_1+x_2\boldsymbol{\alpha}_2+\cdots+x_n\boldsymbol{\alpha}_n=\boldsymbol{\beta}。\tag{4.3}$$

常称式(4.3)为线性方程组(4.1)的向量形式。

一组数 $c_1,\cdots,c_n$ 称为上述线性方程组的解，如果 $\boldsymbol{A}\begin{pmatrix} c_1 \\ \vdots \\ c_n \end{pmatrix}=\boldsymbol{\beta}$。

例如，三式四元线性方程组

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

它的矩阵形式为

$$\begin{pmatrix} 1 & -1 & 2 & -3 \\ 2 & 3 & -1 & 4 \\ 3 & -5 & 2 & -7 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} 2 \\ 5 \\ 1 \end{pmatrix},$$

它的向量形式为

$$x_1 \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} + x_2 \begin{pmatrix} -1 \\ 3 \\ -5 \end{pmatrix} + x_3 \begin{pmatrix} 2 \\ -1 \\ 2 \end{pmatrix} + x_4 \begin{pmatrix} -3 \\ 4 \\ -7 \end{pmatrix} = \begin{pmatrix} 2 \\ 5 \\ 1 \end{pmatrix}。$$

容易验证$\begin{pmatrix} 1 \\ 2 \\ 3 \\ 0 \end{pmatrix}$为它的一组解。

本章中，我们以矩阵为工具导出线性方程组有解的条件，求解的方法；更以向量为工具探讨解的结构，完善求解的理论与方法。

§1 矩阵消元法

一、同解线性方程组

定义 4.1 称两个线性方程组 $AX = \beta$ 与 $CX = \delta$ 为同解的，如果 $AX = \beta$ 的解都是 $CX = \delta$ 的解，并且 $CX = \delta$ 的解也都是 $AX = \beta$ 的解。这里 A 是 $s \times n$ 阵，C 是 $t \times n$ 阵，β 是 $s \times 1$ 阵，δ 是 $t \times 1$ 阵，X 是由未知量组成的 $n \times 1$ 阵。

定理 4.1 若 $s \times (n+1)$ 阵 $(A\ \ \beta)$ 经行初等变换化为 $(C\ \ \delta)$，这里

A 与 C 是 $s\times n$ 阵，β 与 δ 是 $s\times 1$ 阵，则线性方程组 $AX=\beta$ 与 $CX=\delta$ 同解。

证 $CX=\delta$ 的各个方程式或者是 $AX=\beta$ 的一个方程式，或者是 $AX=\beta$ 的方程式经乘数相加而得，所以 $AX=\beta$ 的解一定是 $CX=\delta$ 的解。又行初等变换是可逆的，故 $CX=\delta$ 的解也是 $AX=\beta$ 的解。因而两个线性方程组同解。

由于变更未知量的次序不改变线性方程组的解，而变更未知量的次序意味着系数矩阵作列互换，故有

推论 若 $s\times(n+1)$ 阵 $(A\ \ \beta)$ 经行初等变换与前 n 列互换化为 $(C\ \ \delta)$，则线性方程组 $AX=\beta$ 与 $CY=\delta$ 同解。这里 A,C,β,δ 同定理 4.1。$Y=\begin{pmatrix}y_1\\ \vdots\\ y_n\end{pmatrix}$，$y_1,\cdots,y_n$ 是 $x_1,\cdots,x_n$ 的一个排列。

例如，线性方程组

$$\begin{cases}x+2y+3z=6\\4x+5y+6z=15\\7x+8y+9z=24\end{cases}\quad\text{或}\quad\begin{pmatrix}1&2&3\\4&5&6\\7&8&9\end{pmatrix}\begin{pmatrix}x\\y\\z\end{pmatrix}=\begin{pmatrix}6\\15\\24\end{pmatrix}$$

中，更换未知量 y 与 z 的次序得

$$\begin{cases}x+3z+2y=6,\\4x+6z+5y=15,\\7x+9z+8y=24。\end{cases}\quad\text{或}\quad\begin{pmatrix}1&3&2\\4&6&5\\7&9&8\end{pmatrix}\begin{pmatrix}x\\z\\y\end{pmatrix}=\begin{pmatrix}6\\15\\24\end{pmatrix}。$$

二、矩阵消元法

设 A 为线性方程组(4.2)的系数矩阵，我们称 $s\times(n+1)$ 阵 $(A\ \ \beta)$ 为式(4.2)的增广矩阵。由 §3 定理 3.2，我们对 $(A\ \ \beta)$ 作行初等变换及前 n 列互换使 A 化为标准阶梯形，此时 $(A\ \ \beta)$ 化为

$$
\left(\begin{array}{cc|c}
 & & d_1 \\
\boldsymbol{E}_r & \boldsymbol{C} & \vdots \\
 & & d_r \\
 & & d_{r+1} \\
\boldsymbol{0} & \boldsymbol{0} & \vdots \\
 & & d_s
\end{array}\right) = \boldsymbol{B}, \boldsymbol{C} = \begin{pmatrix} c_{11} \cdots c_{1t} \\ \vdots \quad \vdots \\ c_{r1} \cdots c_{rt} \end{pmatrix}, r \text{ 是 } \boldsymbol{A} \text{ 的秩}, r + t = n。
$$

$\boldsymbol{AX} = \boldsymbol{\beta}$ 与下述方程组

$$
\begin{pmatrix} \boldsymbol{E}_r & \boldsymbol{C} \\ \boldsymbol{0} & \boldsymbol{0} \end{pmatrix} \begin{pmatrix} y_1 \\ \vdots \\ \vdots \\ \vdots \\ y_n \end{pmatrix} = \begin{pmatrix} d_1 \\ \vdots \\ d_r \\ d_{r+1} \\ \vdots \\ d_s \end{pmatrix} \tag{4.4}
$$

同解。式(4.4)有两种可能:

(1) 存在 $i > r, d_i \neq 0$;

此时式(4.4)的第 i 个方程为 $0 = d_i (\neq 0)$,这是不可能的,故原方程组无解。

(2) $d_{r+1} = \cdots = d_s = 0$。

此时原方程组同解于方程组

$$
\begin{cases} y_1 + c_{11} y_{r+1} + \cdots + c_{1t} y_n = d_1, \\ \vdots \qquad\qquad\qquad\qquad\qquad \vdots \\ y_r + c_{r1} y_{r+1} + \cdots + c_{rt} y_n = d_r, \end{cases}
$$

即

$$
\begin{cases} y_1 = -c_{11} y_{r+1} - c_{12} y_{r+2} - \cdots - c_{1t} y_n + d_1, \\ \vdots \qquad\qquad\qquad\qquad\qquad\qquad\qquad \vdots \\ y_r = -c_{r1} y_{r+1} - c_{r2} y_{r+2} - \cdots - c_{rt} y_n + d_r, \end{cases} \tag{4.5}
$$

这里 $y_{r+1}, \cdots, y_n$ 可取任意数值,称为自由未知量。一旦 $y_{r+1}, \cdots, y_n$ 取定一组数值后,由式(4.5),$y_1, \cdots, y_r$ 的数值也确定,于是得到了原方程组

的一组解。

当 $d_{r+1}=\cdots=d_s=0$ 时，$\boldsymbol{B}$ 是 $(\boldsymbol{A}\ \ \boldsymbol{\beta})$ 的阶梯形，其前 r 行为

$$\begin{pmatrix} \boldsymbol{E}_r & \boldsymbol{C} & \begin{matrix} d_1 \\ \vdots \\ d_r \end{matrix} \end{pmatrix} = \begin{pmatrix} 1 & & 0 & & \\ & \ddots & & & * \\ & 0 & \ddots & & \\ & & & 1 & \end{pmatrix},$$

共有 r 个阶梯，后 $s-r$ 行全为 0，故

$$\mathrm{r}(\boldsymbol{A}\ \ \boldsymbol{\beta})=\mathrm{r}(\boldsymbol{B})=r=\mathrm{r}(\boldsymbol{A})。$$

当 $d_{r+1},d_{r+2},\cdots,d_s$ 中有一个不为 0 时，$\boldsymbol{B}$ 的阶梯数 $\geqslant r+1$，于是

$$\mathrm{r}(\boldsymbol{A}\ \ \boldsymbol{\beta})=\mathrm{r}(\boldsymbol{B})\geqslant r+1>r=\mathrm{r}(\boldsymbol{A})。$$

在方程组(4.1)有解时，若 $r<n$，则存在自由未知量 $y_{r+1},\cdots,y_n$，它们可取任意值，从而线性方程组有无穷组解。若 $r=n$，则没有自由未知量，式(4.5)变为 $y_i=d_i(i=1,\cdots,n)$，线性方程组有惟一解。

综上分析，我们得到下面的结论：

定理 4.2 (1) 线性方程组 $\boldsymbol{AX}=\boldsymbol{\beta}$ 有解 $\Leftrightarrow \mathrm{r}(\boldsymbol{A})=\mathrm{r}(\boldsymbol{A}\ \ \boldsymbol{\beta})$；

(2) 线性方程组 $\boldsymbol{AX}=\boldsymbol{\beta}$ 有惟一解 $\Leftrightarrow \mathrm{r}(\boldsymbol{A})=\mathrm{r}(\boldsymbol{A}\ \ \boldsymbol{\beta})=$ 未知量个数。

由于 $\begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}$ 是齐次线性方程组 $\boldsymbol{AX}=\boldsymbol{0}$ 的当然解。故有

定理 4.3 齐次线性方程组 $\boldsymbol{AX}=\boldsymbol{0}$ 有非零解(即解不惟一)，$\Leftrightarrow \mathrm{r}(\boldsymbol{A})<$ 未知量个数。

推论 若 $\boldsymbol{A}$ 是 $s\times n$ 阵，$s<n$，则 $\boldsymbol{AX}=\boldsymbol{0}$ 必有非零解。

证 因为 $\mathrm{r}(\boldsymbol{A})\leqslant \boldsymbol{A}$ 的行数 $s<$ 未知量个数 n。

例 4.1 设线性方程组 $\begin{pmatrix} 1 & 1 & 1 & 1 & 1 \\ 1 & 2 & 2 & 2 & 2 \\ 2 & 3 & 3 & 3 & 3 \\ 3 & 6 & 6 & 6 & 6 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{pmatrix} = \begin{pmatrix} 5 \\ 9 \\ 14 \\ a \end{pmatrix}$ 有解，求 a 的值，并求所有解。

解 增广矩阵 $= \left(\begin{array}{ccccc|c} 1 & 1 & 1 & 1 & 1 & 5 \\ 1 & 2 & 2 & 2 & 2 & 9 \\ 2 & 3 & 3 & 3 & 3 & 14 \\ 3 & 6 & 6 & 6 & 6 & a \end{array}\right) \xrightarrow[(2)-(1)]{\substack{(4)-3(2) \\ (3)-(1)-(2)}}$

$$\begin{pmatrix} 1 & 1 & 1 & 1 & 1 & 5 \\ 0 & 1 & 1 & 1 & 1 & 4 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & a-27 \end{pmatrix} \xrightarrow{(1)-(2)} \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 1 \\ 0 & 1 & 1 & 1 & 1 & 4 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & a-27 \end{pmatrix}$$ 有解，

必 $a-27=0$，故 $a=27$，且解为

$$\begin{cases} x_1 = 1, \\ x_2 = 4 - x_3 - x_4 - x_5. \end{cases}$$

x_3, x_4, x_5 为自由未知量，可取任意数值。

例 4.2 讨论线性方程 $\begin{pmatrix} 1 & \cdots & \cdots & \cdots & 1 \\ & a & & 1 & \\ & & \ddots & & \\ & & & \ddots & \\ & 1 & & & a \end{pmatrix}\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix}$ 的解。

解 (1) 若 $a \neq 1$，则

增广矩阵 $= \begin{pmatrix} 1 & \cdots & \cdots & \cdots & 1 & b_1 \\ & a & & 1 & & \vdots \\ & & \ddots & & & \vdots \\ & & & \ddots & & \vdots \\ & 1 & & & a & b_n \end{pmatrix}$

$$\xrightarrow[i=2,\cdots,n]{(i)-a(1)}\begin{pmatrix} 1 & \cdots & \cdots & \cdots & 1 & b_1 \\ & a-1 & & & & b_2-b_1 \\ & & \ddots & 0 & & \vdots \\ & 0 & & \ddots & & \vdots \\ & & & & a-1 & b_n-b_1 \end{pmatrix}$$

$$\xrightarrow[i=2,\cdots,n]{\frac{1}{a-1}\cdot(i)}\begin{pmatrix} 1 & \cdots & \cdots & \cdots & \cdots & 1 & b_1 \\ & \ddots & & & & & \\ & & \ddots & & 0 & & \frac{b_2-b_1}{a-1} \\ & & & \ddots & & & \vdots \\ & & 0 & & \ddots & & \\ & & & & & 1 & \frac{b_n-b_1}{a-1} \end{pmatrix},$$

原方程组有惟一解：

$$x_1 = b_1 - \sum_{j=2}^{n} \frac{b_j - b_1}{a-1},$$

$$x_i = \frac{b_i - b_1}{a-1}, i = 2, \cdots, n。$$

(2) 若 $a = 1$，则

$$\text{增广矩阵} = \begin{pmatrix} 1 & \cdots & 1 & b_1 \\ 1 & \cdots & 1 & b_2 \\ & \vdots & & \vdots \\ 1 & \cdots & 1 & b_n \end{pmatrix} \xrightarrow[i=2,\cdots,n]{(i)-(1)} \begin{pmatrix} 1 & \cdots & 1 & b_1 \\ 0 & \cdots & 0 & b_2 - b_1 \\ \vdots & & \vdots & \\ 0 & \cdots & 0 & b_n - b_1 \end{pmatrix}$$

于是

原方程组有解 $\Leftrightarrow b_i - b_1 = 0, i = 2, \cdots, n, \Leftrightarrow b_1 = b_2 = \cdots = b_n$。有解时，解为 $x_1 = b - x_2 - \cdots - x_n$，这里 $x_2, \cdots, x_n$ 为自由未知量。

§2 Cramer 法则

若 $\boldsymbol{A}$ 是 n 阶可逆阵，则 $\boldsymbol{AX} = \boldsymbol{\beta} \Leftrightarrow \boldsymbol{X} = \boldsymbol{A}^{-1}\boldsymbol{\beta}$。

而 $A^{-1}\beta=\dfrac{A^*\beta}{|A|}$，它的第 i 行第 1 列元素为

$$\frac{1}{|A|}(A_{1i},A_{2i},\cdots,A_{ni})\begin{pmatrix}b_1\\ \vdots\\ b_n\end{pmatrix}=\frac{1}{|A|}\cdot\begin{array}{c}\begin{vmatrix}a_{11},&\cdots,&b_1,&\cdots,&a_{1n}\\ \vdots&&\vdots&&\vdots\\ a_{n1},&\cdots,&b_n,&\cdots,&a_{nn}\end{vmatrix}\\ \uparrow\\ i\text{列}\end{array}。$$

这样，我们得到了 Cramer 法则。

Cramer 法则　若矩阵 $A=(a_{ij})_{n\times n}$ 可逆，则线性方程组 $AX=\beta$ 有惟一解

$$x_i=\frac{|A^{(1)}\cdots A^{(i-1)}\beta A^{(i+1)}\cdots A^{(n)}|}{|A|},i=1,\cdots,n。$$

例 4.3　求线性方程组 $\begin{pmatrix}1&1&1\\1&2&3\\1&4&5\end{pmatrix}\begin{pmatrix}x_1\\x_2\\x_3\end{pmatrix}=\begin{pmatrix}2\\1\\1\end{pmatrix}$ 的解。

解　$\begin{vmatrix}1&1&1\\1&2&3\\1&4&5\end{vmatrix}=\begin{vmatrix}1&1&1\\0&1&2\\0&2&2\end{vmatrix}=-2\neq0$，方程组有惟一解：

$$x_1=\frac{\begin{vmatrix}2&1&1\\1&2&3\\1&4&5\end{vmatrix}}{-2}=\frac{-4}{-2}=2,$$

$$x_2=\frac{\begin{vmatrix}1&2&1\\1&1&3\\1&1&5\end{vmatrix}}{-2}=\frac{-2}{-2}=1,$$

$$x_3=\frac{\begin{vmatrix}1&1&2\\1&2&1\\1&4&1\end{vmatrix}}{-2}=\frac{2}{-2}=-1。$$

例 4.4 设 a,b,c 两两不等，求 $\begin{pmatrix} 1 & 1 & 1 \\ a & b & c \\ a^2 & b^2 & c^2 \end{pmatrix}\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} a+b+c \\ a^2+b^2+c^2 \\ a^3+b^3+c^3 \end{pmatrix}$ 的解。

解 因为系数矩阵的行列式为 Vandermonde 行列式，其值不为零，故系数矩阵可逆。按 Cramer 法则方程组有惟一解。略加观察得解 $x=a, y=b, z=c$。

例 4.5 讨论线性方程组 $\begin{pmatrix} a & 1 & 1 \\ 1 & a & 1 \\ 1 & 1 & a \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix}$ 的解。

解 $\begin{vmatrix} a & 1 & 1 \\ 1 & a & 1 \\ 1 & 1 & a \end{vmatrix} \xlongequal{(1)+(2)+(3)} \begin{vmatrix} a+2 & a+2 & a+2 \\ 1 & a & 1 \\ 1 & 1 & a \end{vmatrix}$

$$= (a+2)\begin{vmatrix} 1 & 1 & 1 \\ 1 & a & 1 \\ 1 & 1 & a \end{vmatrix} \xlongequal[(3)-(1)]{(2)-(1)} (a+2)\begin{vmatrix} 1 & 1 & 1 \\ 0 & a-1 & 0 \\ 0 & 0 & a-1 \end{vmatrix}$$

$= (a+2)(a-1)^2$。

(1) 若 $a \neq 1, -2$。系数矩阵可逆，方程组有惟一解。

(2) 若 $a=1$

增广矩阵 $= \begin{pmatrix} 1 & 1 & 1 & b_1 \\ 1 & 1 & 1 & b_2 \\ 1 & 1 & 1 & b_3 \end{pmatrix} \xrightarrow[(3)-(1)]{(2)-(1)} \begin{pmatrix} 1 & 1 & 1 & b_1 \\ 0 & 0 & 0 & b_2-b_1 \\ 0 & 0 & 0 & b_3-b_1 \end{pmatrix}$，方程组有解 $\Leftrightarrow b_1=b_2=b_3$，有解时有无穷组解：$x_1=b_1-x_2-x_3$。

(3) 若 $a=-2$

增广矩阵 $= \begin{pmatrix} -2 & 1 & 1 & b_1 \\ 1 & -2 & 1 & b_2 \\ 1 & 1 & -2 & b_2 \end{pmatrix} \xrightarrow{(1)\leftrightarrow(3)}$

$$\begin{pmatrix} 1 & 1 & -2 & b_3 \\ 1 & -2 & 1 & b_2 \\ -2 & 1 & 1 & b_1 \end{pmatrix} \xrightarrow[(2)-(1)]{(3)+(2)+(1)} \begin{pmatrix} 1 & 1 & -2 & b_3 \\ 0 & -3 & 3 & b_2-b_3 \\ 0 & 0 & 0 & b_1+b_2+b_3 \end{pmatrix}$$

$$\xrightarrow[(1)-(2)]{\frac{1}{-3}(2)}\begin{pmatrix}1 & 0 & -1 & \dfrac{2b_3+b_2}{3}\\ 0 & 1 & -1 & \dfrac{b_3-b_2}{3}\\ 0 & 0 & 0 & b_1+b_2+b_3\end{pmatrix},$$

故方程组有解 $\Leftrightarrow b_1+b_2+b_3=0$。

有解时，有无穷组解

$$\begin{cases}x_1=\dfrac{2b_3+b_2}{3}+x_3\\ x_2=\dfrac{b_3-b_2}{3}+x_3,\end{cases}\quad x_3\text{ 是自由未知量。}$$

§3 n 维向量及其线性关系

一、向量的概念

定义 4.2 称 $n\times 1$ 阵为 n 维列向量，$1\times n$ 阵为 n 维行向量。

下面若不作特别声明，我们研究的向量都是数域 P 上的 $n\times 1$ 阵（即列向量），用字母 $\alpha,\beta,\gamma,\cdots$ 来表示，它的元素，即 P 中的数用小写英文字母表示。$\alpha=\begin{pmatrix}a_1\\ \vdots\\ a_n\end{pmatrix}$，$a_1,\cdots,a_n$ 叫 α 的分量。

矩阵有加法与数乘运算，向量作为特殊类型的矩阵，有同样的运算，满足同样的运算规律。

二、向量的线性组合

定义 4.3 若 $\alpha_1,\cdots,\alpha_s,\beta$ 是 $(s+1)$ 个向量，$\kappa_1,\cdots,\kappa_s\in P$，且 $\beta=\kappa_1\alpha_1+\cdots+\kappa_s\alpha_s$，则称 β 是向量 $\alpha_1,\cdots,\alpha_s$ 的一个线性组合，也称向量 $\alpha_1,\cdots,\alpha_s$ 能线性表出 β，记成 $\alpha_1,\cdots,\alpha_s\Rightarrow\beta$。

$\alpha_1,\cdots,\alpha_s\Rightarrow\beta$ 意味着线性方程组 $\sum\limits_{i=1}^{s}x_i\alpha_i=\beta$ 有解。由定理 4.2

即得：

定理 4.4 $\alpha_1,\cdots,\alpha_s \Rightarrow \beta \Leftrightarrow r(\alpha_1,\cdots,\alpha_s)=r(\alpha_1,\cdots,\alpha_s,\beta)$。

例 4.6 因为 $\begin{pmatrix} a \\ b \end{pmatrix} = b\begin{pmatrix} 1 \\ 1 \end{pmatrix} + (a-b)\begin{pmatrix} 1 \\ 0 \end{pmatrix}$，故

$$\begin{pmatrix} 1 \\ 1 \end{pmatrix},\begin{pmatrix} 1 \\ 0 \end{pmatrix} \Rightarrow \begin{pmatrix} a \\ b \end{pmatrix}。$$

例 4.7 因为 $\begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ \vdots \\ a_n \end{pmatrix} = \alpha_1\begin{pmatrix} 1 \\ 0 \\ \vdots \\ \vdots \\ 0 \end{pmatrix} + \alpha_2\begin{pmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} + \cdots + \alpha_n\begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix}$，

若记

$$\varepsilon_1 = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ \vdots \\ 0 \end{pmatrix},\varepsilon_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix},\cdots,\varepsilon_n = \begin{pmatrix} 0 \\ \vdots \\ \vdots \\ 0 \\ 1 \end{pmatrix},$$

称它们为单位向量，则 $\varepsilon_1,\cdots,\varepsilon_n$ 能线性表出任意一个 n 维向量。

例 4.8 因为 $0=0\cdot\alpha_1+\cdots+0\cdot\alpha_s$，所以 $\alpha_1,\cdots,\alpha_s \Rightarrow 0$。

例 4.9 若 $\begin{pmatrix} 1 \\ 2 \\ 4 \end{pmatrix},\begin{pmatrix} 2 \\ 3 \\ 1 \end{pmatrix} \Rightarrow \begin{pmatrix} 2 \\ 0 \\ x \end{pmatrix}$，求 x。

解 矩阵 $\begin{pmatrix} 1 & 2 \\ 2 & 3 \\ 4 & 1 \end{pmatrix}$ 的秩为 2，由定理 4.4，矩阵 $\begin{pmatrix} 1 & 2 & 2 \\ 2 & 3 & 0 \\ 4 & 1 & x \end{pmatrix}$ 的秩也为 2，故 $\begin{vmatrix} 1 & 2 & 2 \\ 2 & 3 & 0 \\ 4 & 1 & x \end{vmatrix} = 0$，得 $x=-10$。

三、线性相关与线性无关

定义 4.4 向量组 $\alpha_1,\cdots,\alpha_s$ 称为线性相关，如果存在 P 中 s 个不全为零的数 $k_1,\cdots,k_s$ 使

$$k_1\alpha_1+k_2\alpha_2+\cdots+k_s\alpha_s=0,$$

即线性方程组

$$x_1\alpha_1+x_2\alpha_2+\cdots+k_s\alpha_s=0$$

在 P 中有非零解 $x_i=k_i(i=1,2,\cdots,s)$。

命题 4.1 向量组 $\alpha_1,\cdots,\alpha_s$ 线性相关

$\Leftrightarrow$矩阵$(\alpha_1,\cdots,\alpha_s)$的秩$<s$。

证 向量组 $\alpha_1,\cdots,\alpha_s$ 线性相关

$\Leftrightarrow$线性方程组

$$\sum_{i=1}^{s}x_i\alpha_i=(\alpha_1,\cdots,\alpha_s)\begin{pmatrix}x_1\\ \vdots\\ x_s\end{pmatrix}=0$$

有非零解$\Leftrightarrow$系数矩阵$(\alpha_1,\cdots,\alpha_s)$的秩$<$未知量个数 s。

推论 一个向量 α 线性相关 $\Leftrightarrow\alpha=0$。

例 4.10 因为 $1\cdot0+0\alpha_1+\cdots+0\alpha_s=0$，所以向量组 $0,\alpha_1,\alpha_2,\cdots\alpha_s$ 线性相关，即含零向量的向量组必线性相关。

例 4.11 因为 $1\cdot\alpha+1\cdot\beta+(-1)(\alpha+\beta)=0$，所以向量组 $\alpha,\beta,(\alpha+\beta)$ 线性相关。

例 4.12 若 $\alpha_1,\cdots,\alpha_s$ 线性相关，即存在不全为零的数 $k_1,\cdots,k_s$ 使

$$k_1\alpha_1+\cdots+k_s\alpha_s=0,$$

此时，$k_1,\cdots,k_s,0,\cdots,0$ 也不全为零，且

$$k_1\alpha_1+\cdots+k_s\alpha_s+0\beta_1+\cdots+0\beta_t=0,$$

所以 $\alpha_1,\cdots,\alpha_s,\beta_1,\cdots,\beta_t$ 也线性相关。即若部分线性相关，则整个向量组也线性相关。

例 4.13 若 $s>n$，则 s 个 n 维向量必线性相关。

三、线性相关与线性无关

定义 4.4 向量组 $\alpha_1,\cdots,\alpha_s$ 称为线性相关，如果存在 P 中 s 个不全为零的数 $k_1,\cdots,k_s$ 使

$$k_1\alpha_1+k_2\alpha_2+\cdots+k_s\alpha_s=0,$$

即线性方程组

$$x_1\alpha_1+x_2\alpha_2+\cdots+k_s\alpha_s=0$$

在 P 中有非零解 $x_i=k_i(i=1,2,\cdots,s)$。

命题 4.1 向量组 $\alpha_1,\cdots,\alpha_s$ 线性相关

$\Longleftrightarrow$矩阵$(\alpha_1,\cdots,\alpha_s)$的秩$<s$。

证 向量组 $\alpha_1,\cdots,\alpha_s$ 线性相关

$\Longleftrightarrow$线性方程组

$$\sum_{i=1}^{s}x_i\alpha_i=(\alpha_1,\cdots,\alpha_s)\begin{pmatrix}x_1\\ \vdots\\ x_s\end{pmatrix}=0$$

有非零解$\Longleftrightarrow$系数矩阵$(\alpha_1,\cdots,\alpha_s)$的秩$<$未知量个数 s。

推论 一个向量 α 线性相关 $\Longleftrightarrow\alpha=0$。

例 4.10 因为 $1\cdot0+0\alpha_1+\cdots+0\alpha_s=0$，所以向量组 $0,\alpha_1,\alpha_2,\cdots\alpha_s$ 线性相关，即含零向量的向量组必线性相关。

例 4.11 因为 $1\cdot\alpha+1\cdot\beta+(-1)(\alpha+\beta)=0$，所以向量组 $\alpha,\beta,(\alpha+\beta)$ 线性相关。

例 4.12 若 $\alpha_1,\cdots,\alpha_s$ 线性相关，即存在不全为零的数 $k_1,\cdots,k_s$ 使

$$k_1\alpha_1+\cdots+k_s\alpha_s=0,$$

此时，$k_1,\cdots,k_s,0,\cdots,0$ 也不全为零，且

$$k_1\alpha_1+\cdots+k_s\alpha_s+0\beta_1+\cdots+0\beta_t=0,$$

所以 $\alpha_1,\cdots,\alpha_s,\beta_1,\cdots,\beta_t$ 也线性相关。即若部分线性相关，则整个向量组也线性相关。

例 4.13 若 $s>n$，则 s 个 n 维向量必线性相关。

性相关，则 $\alpha_1,\cdots,\alpha_s \Rightarrow \beta$。

证 因为向量组 $\alpha_1,\cdots,\alpha_s$ 线性无关，故

$$r(\alpha_1,\cdots,\alpha_s)=s,$$

又 $\alpha_1,\cdots,\alpha_s,\beta$ 线性相关，故

$$r(\alpha_1,\cdots,\alpha_s,\beta)<s+1。$$

但

$$r(\alpha_1,\cdots,\alpha_s)\leqslant r(\alpha_1,\cdots,\alpha_s,\beta),$$

于是

$$s\leqslant r(\alpha_1,\cdots,\alpha_s,\beta)<s+1,$$

因而必有

$$r(\alpha_1,\cdots,\alpha_s,\beta)=s=r(\alpha_1,\cdots,\alpha_s),$$

由定理 4.4

$$\alpha_1,\cdots,\alpha_s \Rightarrow \beta。$$

命题 4.4 若 $\alpha_1,\cdots,\alpha_s \Rightarrow \beta_i, i=1,\cdots,t$，且 $t>s$，则向量组 $\beta_1,\cdots,\beta_t$ 线性相关。

证 由假设，存在数 $b_{1i},\cdots,b_{si}(i=1,\cdots,t)$ 使

$$\beta_i=\sum_{k=1}^{s}b_{ki}\alpha_k,\quad i=1,\cdots,t。$$

利用矩阵的分块乘法，有

$$(\beta_1,\beta_2,\cdots,\beta_t)=(\alpha_1,\cdots,\alpha_s)\cdot\begin{pmatrix}b_{11} & \cdots\cdots & b_{1t}\\ \vdots & & \vdots\\ \vdots & & \vdots\\ b_{s1} & \cdots\cdots & b_{st}\end{pmatrix}。$$

积矩阵$(\beta_1,\cdots,\beta_t)$秩 $\leqslant$ 矩阵$(\alpha_1,\cdots,\alpha_s)$ 秩 $\leqslant s<t$，故向量组 $\beta_1,\cdots,\beta_t$ 线性相关。

若 $\alpha_1,\cdots,\alpha_s$ 为 s 个 n 维向量，$s>n$。因为 $\varepsilon_1,\cdots,\varepsilon_n \Rightarrow \alpha_1,\cdots,\alpha_s$。由命题 4.4，向量组 $\alpha_1,\cdots,\alpha_s$ 线性相关，我们再次验证了例 4.13。

四、等价向量组

定义 4.6 设 Σ_1 与 Σ_2 都是向量组，若对任意 $\beta \in \Sigma_2$，均存在 Σ_1 中有限个向量 $\alpha_1, \cdots, \alpha_s$ 使 $\alpha_1, \cdots. \alpha_s \Rightarrow \beta$，则称向量组 Σ_1 能线性表出向量组 Σ_2，记成 $\Sigma_1 \Rightarrow \Sigma_2$。

推论 若 $\Sigma_1 \Rightarrow \Sigma_2, \Sigma_2 \Rightarrow \Sigma_3$，则 $\Sigma_1 \Rightarrow \Sigma_3$。

证 对任意 $\gamma \in \Sigma_3$，存在 $\beta_1, \cdots, \beta_t \in \Sigma_2$ 使

$$\gamma = \sum_{i=1}^{t} k_i \beta_i = (\beta_1, \cdots, \beta_t) \begin{pmatrix} k_1 \\ \vdots \\ k_t \end{pmatrix},$$

又 $\Sigma_1 \Rightarrow \Sigma_2$，故存在 $\alpha_1, \cdots, \alpha_s \in \Sigma_1$ 使

$$(\beta_1, \cdots, \beta_t) = (\alpha_1, \cdots, \alpha_s) \begin{pmatrix} b_{11} & \cdots & b_{1t} \\ \vdots & & \vdots \\ b_{s1} & \cdots & b_{st} \end{pmatrix},$$

这里 $\beta_t = \sum_{j=1}^{s} b_{ji} \alpha_j, i = 1, \cdots, t$。

于是

$$\gamma = (\alpha_1, \cdots, \alpha_s) \begin{pmatrix} b_{11} & \cdots & b_{1t} \\ \vdots & & \vdots \\ b_{s1} & \cdots & b_{st} \end{pmatrix} \begin{pmatrix} k_1 \\ \vdots \\ k_t \end{pmatrix},$$

若记 $s \times 1$ 阵

$$\begin{pmatrix} b_{11} & \cdots & b_{1t} \\ \vdots & & \vdots \\ b_{s1} & \cdots & b_{st} \end{pmatrix} \begin{pmatrix} k_1 \\ \vdots \\ k_t \end{pmatrix} \text{为} \begin{pmatrix} d_1 \\ \vdots \\ d_s \end{pmatrix},$$

则 $\gamma = d_1 \alpha_1 + \cdots + d_s \alpha_s$，即 $\Sigma_1 \Rightarrow \Sigma_s$。

定义 4.7 设 Σ_1 与 Σ_2 是两个向量组，若 $\Sigma_1 \Rightarrow \Sigma_2$ 且 $\Sigma_2 \Rightarrow \Sigma_1$，则称两个向量组等价，记成 $\Sigma_1 \sim \Sigma_2$。

例 4.19 若记 $\varepsilon_i = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$ 第 i 行，$\eta_i = \begin{pmatrix} 1 \\ \vdots \\ \vdots \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$ i 个 1，$(i = 1,\cdots,n)$

则 $\{\varepsilon_1,\cdots,\varepsilon_n\} \sim \{\eta_1,\cdots,\eta_n\}$。

证 这是因为 $\varepsilon_1 = \eta_1$；$\eta_i = \varepsilon_1 + \cdots + \varepsilon_i$，$(i = 2,\cdots,n)$

$$\varepsilon_i = \eta_i - \eta_{i-1}。$$

§4 向量组的秩

一、极大线性无关组

定义 4.8 设 Σ 是一个向量组，$\alpha_1,\cdots,\alpha_r \in \Sigma$，$\alpha_1,\cdots,\alpha_r$ 称为 Σ 的一个极大线性无关组，如果

(1) $\alpha_1,\cdots,\alpha_r$ 线性无关；

(2) 对任意 $\alpha \in \Sigma$，$\alpha_1,\cdots,\alpha_r,\alpha$ 线性相关。

例 4.20 设 $\Sigma = \left\{\begin{pmatrix}1\\0\end{pmatrix},\begin{pmatrix}0\\1\end{pmatrix},\begin{pmatrix}1\\1\end{pmatrix},\begin{pmatrix}1\\-1\end{pmatrix}\right\}$，则 $\left\{\begin{pmatrix}1\\0\end{pmatrix},\begin{pmatrix}0\\1\end{pmatrix}\right\}$，$\left\{\begin{pmatrix}1\\0\end{pmatrix},\begin{pmatrix}1\\1\end{pmatrix}\right\}$ 与 $\left\{\begin{pmatrix}1\\1\end{pmatrix},\begin{pmatrix}1\\-1\end{pmatrix}\right\}$ 都是 Σ 的极大线性无关组。

例 4.21 若 Σ 是数域 P 上 n 维向量的全体（记成 P^n），则 $\varepsilon_1,\cdots,\varepsilon_n$ 是 Σ 的一个极大线性无关组。

证 (1) 矩阵 $E_n = (\varepsilon_1,\cdots,\varepsilon_n)$ 秩为 n，故 $\varepsilon_1,\cdots,\varepsilon_n$ 线性无关；

(2) 对任意 $\begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} \in \Sigma$，有 $(-a_1)\varepsilon_1 + \cdots + (-a_n)\varepsilon_n + 1 \cdot \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} =$

0，即 $\varepsilon_1, \cdots, \varepsilon_n, \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix}$ 线性相关。

由定义，即得结论。

命题 4.5 设 $\alpha_1, \cdots, \alpha_r \in \Sigma$，则 $\alpha_1, \cdots, \alpha_r$ 是 Σ 的一个极大线性无关组 $\Leftrightarrow$(1) $\alpha_1, \cdots, \alpha_r$ 线性无关且(2′) $\alpha_1, \cdots, \alpha_r \Rightarrow \Sigma$，即在(1)成立的前提下，定义 4.8 中的(2)$\Leftrightarrow$(2′)。

证 (2)$\Rightarrow$(2′)：对任意 $\alpha \in \Sigma, \alpha_1, \cdots, \alpha_r$ 线性无关，$\alpha_1, \cdots, \alpha_r, \alpha$ 线性相关，由命题 4.3，$\alpha_1, \cdots, \alpha_r \Rightarrow \alpha$，即(2′)成立。

(2′)$\Rightarrow$(2)：对任意 $\alpha \in \Sigma$，若 $\alpha_1, \cdots, \alpha_r \Rightarrow \alpha$，即存在数 $k_1, \cdots, k_r$ 使

$$\alpha = k_1\alpha_1 + \cdots + k_r\alpha_r,$$

故

$$k_1\alpha_1 + \cdots + k_r\alpha_r + (-1)\alpha = 0,$$

即 $\alpha_1, \cdots, \alpha_r, \alpha$ 线性相关。

命题 4.6 若 $\alpha_1, \cdots, \alpha_r$ 与 $\beta_1, \cdots, \beta_t$ 是向量组 Σ 的两个极大线性无关组，则 $r = t$。

证 若 $r \neq t$，不妨一般地设 $r > t$，$\beta_1, \cdots, \beta_t$ 是 Σ 的一个极大线性无关组，由命题 4.5 得

$$\beta_1, \cdots, \beta_t \Rightarrow \Sigma,$$

因为

$$\alpha_1, \cdots, \alpha_r \in \Sigma,$$

故

$$\beta_1, \cdots, \beta_t \Rightarrow \alpha_i, i = 1, 2, \cdots, r。$$

由命题 4.4 得 $\alpha_1, \cdots, \alpha_r$ 线性相关，这与 $\alpha_1, \cdots, \alpha_r$ 是 Σ 的极大线性无关组的假设矛盾，故必 $r = t$。

二、向量组的秩

定义 4.9 由命题 4.6，向量组 Σ 的任意两个极大线性无关向量组

都含相同个数的向量，称这个个数为向量组 Σ 的秩，记作 $\mathrm{r}(\Sigma)$。

例 4.22 向量组 $\left\{\begin{pmatrix}1\\0\end{pmatrix},\begin{pmatrix}0\\1\end{pmatrix},\begin{pmatrix}1\\1\end{pmatrix},\begin{pmatrix}1\\-1\end{pmatrix}\right\}$ 的秩为 2。

例 4.23 因为向量组 $\boldsymbol{P}^n$ 有极大线性无关组 $\varepsilon_1,\cdots,\varepsilon_n$，故 $\boldsymbol{P}^n$ 的秩为 n。

性质 (1) 向量组 $\{\alpha_1,\cdots,\alpha_r\}$ 的秩为 $r \Leftrightarrow \alpha_1,\cdots,\alpha_r$ 线性无关；

(2) 若 $\boldsymbol{\Sigma}_1 \Rightarrow \boldsymbol{\Sigma}_2$，则 $\mathrm{r}(\boldsymbol{\Sigma}_1) \geqslant \mathrm{r}(\boldsymbol{\Sigma}_2)$；

特别由例 4.23，如果 $\boldsymbol{\Sigma}$ 是 n 维向量组，即 $\boldsymbol{\Sigma} \subseteq \boldsymbol{P}^n$，则 $\mathrm{r}(\boldsymbol{\Sigma}) \leqslant n$；

(3) 若 $\boldsymbol{\Sigma}_1 \sim \boldsymbol{\Sigma}_2$，则 $\mathrm{r}(\boldsymbol{\Sigma}_1) = \mathrm{r}(\boldsymbol{\Sigma}_2)$。

证 (1) 是显然的。

(2) 若 $\mathrm{r}(\boldsymbol{\Sigma}_2) = t > \mathrm{r}(\boldsymbol{\Sigma}_1) = r$，设 $\alpha_1,\cdots,\alpha_r$ 与 $\beta_1,\cdots,\beta_t$ 分别是 $\boldsymbol{\Sigma}_1$ 与 $\boldsymbol{\Sigma}_2$ 的极大线性无关组，$\alpha_1,\cdots,\alpha_r \Rightarrow \boldsymbol{\Sigma}_1 \Rightarrow \boldsymbol{\Sigma}_2$，从而 $\alpha_1,\cdots,\alpha_r \Rightarrow \beta_i$，$i = 1,\cdots,t$，由命题 4.4 得 $\beta_1,\cdots,\beta_t$ 线性相关，与 $\beta_1,\cdots,\beta_t$ 是 $\boldsymbol{\Sigma}_2$ 的极大线性无关组的假设矛盾，故必 $t \leqslant r$。

(3) 由(2)，$\mathrm{r}(\boldsymbol{\Sigma}_1) \geqslant \mathrm{r}(\boldsymbol{\Sigma}_2)$，且 $\mathrm{r}(\boldsymbol{\Sigma}_2) \geqslant \mathrm{r}(\boldsymbol{\Sigma}_1)$，故 $\mathrm{r}(\boldsymbol{\Sigma}_1) = \mathrm{r}(\boldsymbol{\Sigma}_2)$。

推论 若

$$\alpha_1,\cdots,\alpha_s \Rightarrow \beta_i, i=1,\cdots,s,$$

即 $(\beta_1,\cdots,\beta_s) = (\alpha_1,\cdots,\alpha_s)\boldsymbol{B}$，这里 $\boldsymbol{B}$ 是 $s \times s$ 可逆矩阵，则

$$\mathrm{r}(\alpha_1,\cdots,\alpha_s) = \mathrm{r}(\beta_1,\cdots,\beta_s)。$$

证 设 $\boldsymbol{B}^{-1}$ 为 $\boldsymbol{B}$ 的逆，则

$$(\alpha_1,\cdots,\alpha_s) = (\alpha_1,\cdots,\alpha_s)\boldsymbol{B}\boldsymbol{B}^{-1} = (\beta_1,\cdots,\beta_s)\boldsymbol{B}^{-1},$$

即

$$\beta_1,\cdots,\beta_s \Rightarrow \alpha_i, i = 1,\cdots,s.$$

于是

$$\{\alpha_1,\cdots,\alpha_s\} \sim \{\beta_1,\cdots,\beta_s\},$$

由性质(3)，两向量组秩相等。

例 4.24 若向量组 $\alpha_1,\cdots,\alpha_s$ 线性无关：

$$(\beta_1,\cdots,\beta_s)=(\alpha_1,\cdots,\alpha_s)\begin{pmatrix}1 & & & & 2\\ 2 & \ddots & & 0 & \\ & \ddots & \ddots & & \\ & 0 & \ddots & & \\ & & & 2 & 1\end{pmatrix},$$

则 $\beta_1,\cdots,\beta_s$ 线性无关。

证 因为

$$\begin{vmatrix}1 & & & 2\\ 2 & \ddots & & \\ & \ddots & \ddots & \\ & & 2 & 1\end{vmatrix}=1+(-1)^{s+1}2^s\neq 0,$$

故矩阵

$$\begin{pmatrix}1 & & & 2\\ 2 & \ddots & & \\ & \ddots & \ddots & \\ & & 2 & 1\end{pmatrix}$$

可逆，由推论得

$$\mathrm{r}(\beta_1,\cdots,\beta_s)=\mathrm{r}(\alpha_1,\cdots,\alpha_s)=s,$$

故 $\beta_1,\cdots,\beta_s$ 线性无关。

命题 4.7 若 $\mathrm{r}(\Sigma)=r$，$\alpha_1,\cdots,\alpha_r$ 是 Σ 的一个线性无关组，则 $\alpha_1,\cdots,\alpha_r$ 是 Σ 的一个极大线性无关组。

证 由 $\mathrm{r}(\Sigma)=r$，可设 $\beta_1,\cdots,\beta_r$ 是 Σ 的一个极大线性无关组。对任意 $\alpha\in\Sigma$，$\beta_1,\cdots,\beta_r\Rightarrow\alpha_1,\cdots,\alpha_r,\alpha$，$r+1>r$，由命题 4.4，$\alpha_1,\cdots,\alpha_r,\alpha$ 线性相关。按定义，$\alpha_1,\cdots,\alpha_r$ 是 Σ 的一个极大线性无关组。

命题 4.8 若 $\alpha_1,\cdots,\alpha_t$ 是向量组 Σ 的一个线性无关向量组，$\mathrm{r}(\Sigma)>t$，则存在 $\alpha_{t+1},\cdots,\alpha_r\in\Sigma$ 使 $\alpha_1,\cdots,\alpha_t,\cdots,\alpha_r$ 是 Σ 的一个极大线性无关组。

证 因 $\mathrm{r}(\Sigma)=r$，Σ 的极大线性无关组含 r 个向量，故 $\alpha_1,\cdots,\alpha_t$ 不是 Σ 的极大线性无关组，即存在 $\alpha_{t+1}\in\Sigma$ 使 $\alpha_1,\cdots,\alpha_t,\alpha_{t+1}$ 线性无关，若

$t+1<r$，仍存在 $\alpha_{t+1},\cdots$，于是存在 $\alpha_{t+1},\cdots,\alpha_r$ 使 $\alpha_1,\cdots,\alpha_t,\alpha_{t+1},\cdots,\alpha_r$ 线性无关，由命题 4.7，它们是 $\boldsymbol{\Sigma}$ 的一个极大线性无关组。

三、矩阵的秩

我们在第三章给出了矩阵的秩的定义。下面讨论矩阵的秩与它的行向量组或列向量组的秩之间的关系。

定理 4.5 矩阵的秩 = 它的列向量组的秩（称为列秩）= 它的行向量组的秩（称为行秩）。

证 设矩阵 A 的秩为 r，它有一个 r 阶子式 $\boldsymbol{D}_A\begin{pmatrix} i_1\cdots i_r \\ j_1\cdots j_r \end{pmatrix}\neq 0$，而所有 $s(>r)$ 阶子式全为 0。

此时 r 列矩阵 $(\boldsymbol{A}^{(j_1)}\cdots\boldsymbol{A}^{(j_r)})=\boldsymbol{B}$ 也有一个 r 阶子式 $\boldsymbol{D}_B\begin{pmatrix} i_1\cdots i_r \\ 1\ 2\cdots r \end{pmatrix}=\boldsymbol{D}_A\begin{pmatrix} i_1\cdots i_r \\ j_1\cdots j_r \end{pmatrix}\neq 0$，故其秩为 r，由命题 4.2 列向量组 $\boldsymbol{A}^{(j_1)},\cdots,\boldsymbol{A}^{(j_r)}$ 线性无关。

若 $\boldsymbol{\alpha}$ 是 $\boldsymbol{A}$ 的任意一个列向量，则 $(r+1)$ 列矩阵 $(\boldsymbol{A}^{(j_1)}\cdots\boldsymbol{A}^{(j_r)}\boldsymbol{\alpha})$ 的所有 $(r+1)$ 阶子式或者两列相等值为零，或者它是 $\boldsymbol{A}$ 的一个 $(r+1)$ 阶子式乘 (± 1)，值也为零，即 $(r+1)$ 列矩阵 $(\boldsymbol{A}^{(j_1)}\cdots\boldsymbol{A}^{(j_r)}\boldsymbol{\alpha})$ 的秩小于 $r+1$，由命题 4.1，列向量 $\boldsymbol{A}^{(j_1)},\cdots,\boldsymbol{A}^{(j_r)},\boldsymbol{\alpha}$ 线性相关。

于是，$\boldsymbol{A}^{(j_1)},\cdots,\boldsymbol{A}^{(j_r)}$ 是 $\boldsymbol{A}$ 的列向量组的一个极大线性无关组，从而 $\boldsymbol{A}$ 的列秩 $=r$。

又 $\boldsymbol{A}$ 的行秩 $=\boldsymbol{A}'$ 的列秩 $=\boldsymbol{A}'$ 的秩 $=\boldsymbol{A}$ 的秩 r。

例 4.25 矩阵 $\begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 2 & 4 & 6 & 8 & 10 \end{pmatrix}$ 的秩为 1。$\begin{pmatrix} 1 \\ 2 \end{pmatrix}$ 是列向量组的一个极大线性无关组。(1,2,3,4,5)是行向量组的一个极大线性无关组。列秩、行秩都为 1。

§5 线性方程组解的结构

一、齐次线性方程组的解空间

我们用 $\boldsymbol{P}^{s\times n}$ 表示数域 P 上 $s\times n$ 矩阵全体，即

$$\boldsymbol{P}^{s\times n}=\{(a_{ij})_{s\times n}\mid a_{ij}\in P, i=1,\cdots,s; j=1,\cdots,n\},$$

特别用 $\boldsymbol{P}^n$ 表示 $\boldsymbol{P}^{n\times 1}$，即

$$\boldsymbol{P}^n=\left\{\begin{pmatrix}a_1\\ \vdots\\ a_n\end{pmatrix}\middle| a_i\in P, i=1,\cdots,n\right\},$$

也就是用 $\boldsymbol{P}^n$ 表示 P 上 n 维向量全体所成集合。

若 $\boldsymbol{A}\in\boldsymbol{P}^{s\times n}$，记 $\boldsymbol{V}=\{\boldsymbol{X}\in\boldsymbol{P}^n\mid\boldsymbol{AX}=\boldsymbol{0}\}$，即 $\boldsymbol{AX}=\boldsymbol{0}$ 的解向量全体，称 $\boldsymbol{V}$ 为齐次线性方程组 $\boldsymbol{AX}=\boldsymbol{0}$ 的解空间。

性质 （1）若 $\alpha_1,\cdots,\alpha_s\in\boldsymbol{V}$，$k_1,\cdots,k_s\in P$，则

$$\sum_{i=1}^{s}k_i\alpha_i\in\boldsymbol{V};$$

（2）若 $\eta_1,\cdots,\eta_t$ 是 $\boldsymbol{V}$ 的一个极大线性无关组，则

$$\boldsymbol{V}=\left\{\sum_{i=1}^{t}k_i\eta_i\mid k_i\in P, i=1,\cdots,t\right\}。$$

证 （1）因为 $\boldsymbol{A}(\sum_{i=1}^{s}k_i\alpha_i)=\sum_{i=1}^{s}k_i\boldsymbol{A}\alpha_i=0$。

（2）由(1)，$\sum_{i=1}^{t}k_i\eta_i\in\boldsymbol{V}$。反之，对任意 $\alpha\in\boldsymbol{V}$，因为 $\eta_1,\cdots,\eta_t$ 是 $\boldsymbol{V}$ 的一个极大线性无关组，故 $\eta_1,\cdots,\eta_t\Rightarrow\alpha$，即存在 $k_1,\cdots,k_t\in P$，使 $\alpha=\sum_{i=1}^{t}k_i\eta_i$。

定义 4.10 $\boldsymbol{V}$ 的一个极大线性无关组称为齐次线性方程组 $\boldsymbol{AX}=\boldsymbol{0}$ 的一个基础解系。

由性质(2)可知，若 $\eta_1,\cdots,\eta_t$ 是 $\boldsymbol{AX}=\boldsymbol{0}$ 的一个基础解系，则 $\boldsymbol{AX}=\boldsymbol{0}$ 的全部解(称通解)为

$$\sum_{i=1}^{t} k_i\eta_t,\ k_i \in P, i=1,\cdots,t。$$

二、基础解系的求法

一式二元齐次线性方程组 $x+y=0$ 的解为 $x=-y$，向量形式为 $\begin{pmatrix} x \\ y \end{pmatrix}=\begin{pmatrix} -y \\ y \end{pmatrix}=k\begin{pmatrix} -1 \\ 1 \end{pmatrix}$，这里 $k=y$ 为自由未知量，可取任意数值。

三式五元齐次线性方程组

$$\begin{cases} x_1 \qquad\quad + a_1x_4+b_1x_5=0 \\ \quad x_2 \qquad + a_2x_4+b_2x_5=0 \\ \qquad\quad x_3+a_3x_4+b_3x_5=0 \end{cases}$$

的解为

$$\boldsymbol{X}=\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{pmatrix}=\begin{pmatrix} -a_1x_4-b_1x_5 \\ -a_2x_4-b_2x_5 \\ -a_3x_4-b_3x_5 \\ x_4+0x_5 \\ 0x_4+x_5 \end{pmatrix}$$

$$=(-x_4)\begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \\ -1 \\ 0 \end{pmatrix}+(-x_5)\begin{pmatrix} b_1 \\ b_2 \\ b_3 \\ 0 \\ -1 \end{pmatrix}=k_1\begin{pmatrix} a_1 \\ a_2 \\ a_3 \\ -1 \\ 0 \end{pmatrix}+k_2\begin{pmatrix} b_1 \\ b_2 \\ b_3 \\ 0 \\ -1 \end{pmatrix},$$

这里 $k_1=-x_4, k_2=-x_5$ 可取任意数值。

上述例子启发了我们用类似的方法来处理一般的线性方程组。

若给了 s 式 n 元齐次线性方程组

$$\boldsymbol{AX}=\boldsymbol{0},$$

我们用行初等变换(必要时改变未知量次序)化 $s\times n$ 阵 $\boldsymbol{A}$ 为标准阶梯形

$$\begin{pmatrix} \boldsymbol{E}_r & \boldsymbol{T} \\ \boldsymbol{0} & \boldsymbol{0} \end{pmatrix},$$

这里 r 是系数矩阵 $\boldsymbol{A}$ 的秩,$\boldsymbol{E}_r$ 是 r 阶单位阵,$\boldsymbol{T}$ 是 $r\times t$ 阵,$t=n-r$。

齐次线性方程组 $\boldsymbol{AX}=\boldsymbol{0}$ 同解于 $\begin{pmatrix} \boldsymbol{E}_r & \boldsymbol{T} \\ \boldsymbol{0} & \boldsymbol{0} \end{pmatrix}\boldsymbol{X}=\boldsymbol{0}$,或 $\left[\begin{pmatrix} \boldsymbol{E}_r & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{E}_t \end{pmatrix}+\begin{pmatrix} \boldsymbol{0} & \boldsymbol{T} \\ \boldsymbol{0} & -\boldsymbol{E}_t \end{pmatrix}\right]\boldsymbol{X}=\boldsymbol{0}$,从而

$$\boldsymbol{X}=\begin{pmatrix} \underbrace{\boldsymbol{0}}_{r列} & \underbrace{\boldsymbol{T}}_{t列} \\ \boldsymbol{0} & -\boldsymbol{E}_t \end{pmatrix}(-\boldsymbol{X})。$$

若记 $n\times t$ 阵 $\begin{pmatrix} \boldsymbol{T} \\ -\boldsymbol{E}_t \end{pmatrix}$ 的第 i 列为 $\boldsymbol{\alpha}_i$,$i=1,\cdots,t$。利用矩阵的分块记法与乘法得

$$\begin{aligned}\boldsymbol{X}&=(\overbrace{\boldsymbol{0},\cdots,\boldsymbol{0}}^{r列},\boldsymbol{\alpha}_1,\cdots,\boldsymbol{\alpha}_t)\begin{pmatrix} -x_1 \\ \vdots \\ -x_n \end{pmatrix}\\ &=(-x_{r+1})\boldsymbol{\alpha}_1+\cdots+(-x_n)\boldsymbol{\alpha}_t\\ &=k_1\boldsymbol{\alpha}_1+\cdots+k_t\boldsymbol{\alpha}_t,\end{aligned}$$

这里,$k_1=-x_{r+1},\cdots,k_t=-x_n$,可取任意数值.即通解为 $\boldsymbol{\alpha}_1,\cdots,\boldsymbol{\alpha}_t$ 的线性组合。矩阵 $(\boldsymbol{\alpha}_1,\cdots,\boldsymbol{\alpha}_t)$ 的后 t 行组成的 t 阶子式 $\begin{vmatrix} -1 & & 0 \\ & \ddots & \\ 0 & & -1 \end{vmatrix}\neq 0$,其秩为 t,所以解向量 $\boldsymbol{\alpha}_1,\cdots,\boldsymbol{\alpha}_t$ 线性无关,它们能线性表出任意一个解向量,由命题 4.5,它们是 $\boldsymbol{AX}=\boldsymbol{0}$ 的一个基础解系。

基础解系的求法 用行初等变换化 $s\times n$ 阵 $\boldsymbol{A}$ 为标准阶梯形

$$\begin{pmatrix} \boldsymbol{E}_r & \boldsymbol{T} \\ \boldsymbol{0} & \boldsymbol{0} \end{pmatrix},$$

则 $\boldsymbol{AX}=\boldsymbol{0}$ 有一个基础解系 $\boldsymbol{\alpha}_1,\cdots,\boldsymbol{\alpha}_t$。这里，$t=n-r$，$r$ 为 $\boldsymbol{A}$ 的秩，$\boldsymbol{\alpha}_i$ 为 $\begin{pmatrix} \boldsymbol{T} \\ -\boldsymbol{E}_t \end{pmatrix}$ 的 i 列，$i=1,\cdots,t$。$\boldsymbol{AX}=\boldsymbol{0}$ 的通解为

$$\boldsymbol{X}=\sum_{i=1}^{t}k_i\boldsymbol{\alpha}_i。$$

推论 $\boldsymbol{AX}=\boldsymbol{0}$ 的解空间的秩为未知量个数减去 $\boldsymbol{A}$ 的秩。

例 4.26 求齐次线性方程组

$$\begin{pmatrix} 1 & 1 & 1 & 1 & 1 \\ 1 & 2 & 2 & 2 & 2 \\ 2 & 3 & 3 & 3 & 3 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{pmatrix}=\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

的基础解系与通解。

解 $$\begin{pmatrix} 1 & 1 & 1 & 1 & 1 \\ 1 & 2 & 2 & 2 & 2 \\ 2 & 3 & 3 & 3 & 3 \end{pmatrix}\xrightarrow[(2)-(1)]{(3)-(1)-(2)}\begin{pmatrix} 1 & 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}\xrightarrow{(1)-(2)}$$

$$\begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}=\left(\begin{array}{cc|ccc} \multicolumn{2}{c|}{\multirow{2}{*}{$\boldsymbol{E}_2$}} & 0 & 0 & 0 \\ & & 1 & 1 & 1 \\ \hline 0 & 0 & 0 & 0 & 0 \end{array}\right),$$

故 $\boldsymbol{AX}=\boldsymbol{0}$ 有基础解系

$$\begin{pmatrix} 0 \\ 1 \\ -1 \\ 0 \\ 0 \end{pmatrix},\begin{pmatrix} 0 \\ 1 \\ 0 \\ -1 \\ 0 \end{pmatrix},\begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \\ -1 \end{pmatrix},$$

通解为

$$X = k_1\begin{pmatrix}0\\1\\-1\\0\\0\end{pmatrix} - k_2\begin{pmatrix}0\\1\\0\\-1\\0\end{pmatrix} + k_3\begin{pmatrix}0\\1\\0\\0\\-1\end{pmatrix}。$$

例 4.27 若 n 阶方阵 A 的行列式 $|A| = 0$,A 有一代数余子式 $A_{st} \neq 0$。求证：齐次线性方程组 $AX = 0$ 的通解为 $X = k\alpha$,其中

$$\alpha = \begin{pmatrix}A_{s1}\\A_{s2}\\\vdots\\A_{sn}\end{pmatrix}。$$

证 A 的唯一 n 阶子式 $|A| = 0$,而 A 有一个 $(n-1)$ 阶子式 $(-1)^{s+t}A_{st} \neq 0$,故 A 的秩为 $n-1$。从而齐次线性方程组 $AX = 0$ 基础解系个数为

$$n - (n-1) = 1。$$

所以只须证 α 为 $AX = 0$ 的非零解,因为 $A_{st} \neq 0$,故 $\alpha \neq 0$。又

$$A\alpha = \begin{pmatrix}a_{11} & \cdots & a_{1n}\\ \vdots & \vdots & \vdots\\ a_{n1} & \cdots & a_{nn}\end{pmatrix}\begin{pmatrix}A_{s1}\\ \vdots\\ A_{sn}\end{pmatrix} = \begin{pmatrix}\sum_{k=1}^{n} a_{1k}A_{sk}\\ \vdots\\ \sum_{k=1}^{n} a_{sk}A_{sk}\\ \vdots\\ \sum_{k=1}^{n} a_{nk}A_{sk}\end{pmatrix} = \begin{pmatrix}0\\ \vdots\\ |A|\\ \vdots\\ 0\end{pmatrix} = \begin{pmatrix}0\\ \vdots\\ 0\\ \vdots\\ 0\end{pmatrix},$$

即 α 是 $AX = 0$ 的非零解。从而是基础解系,于是通解为 $X = k\alpha$。

三、非齐次线性方程组解的结构

设 $A \in P^{s\times n}$,$\beta \in P^s$,$\beta \neq 0$。记 $W = \{X \in P^n | AX = \beta\}$,即 W 为 $AX = \beta$ 的解向量全体组成的集合。设 V 为 $AX = 0$ 的解空间,则有：

性质 (1) 若 $X_1,X_2\in W$,则 $X_2-X_1\in V$;

(2) 若 $X_0\in W$,则 $X\in W\Leftrightarrow X-X_0\in V$;

(3) 若 $X_0\in W,\alpha_1,\cdots,\alpha_t$ 是 $AX=0$ 的一个基础解系,则 $W=\{X_0+\sum\limits_{i=1}^{t}k_i\alpha_i\,|\,k_i\in P,i=1,\cdots,t\}$,即 $AX=\beta$ 的通解为 $X_0+\sum\limits_{i=1}^{t}k_i\alpha_i$。

证 (1) $AX_1=\beta,AX_2=\beta$,故 $A(X_2-X_1)=AX_2-AX_1=\beta-\beta=0$,即 $X_2-X_1\in V$。

(2) $X\in W$,则 $A(X-X_0)=AX-AX_0=\beta-\beta=0$,所以 $X-X_0\in V$。反之 $X-X_0\in V$,则 $AX=A(X-X_0+X_0)=A(X-X_0)+AX_0=0+\beta=\beta$,即 $X\in W$。

(3) 由(2)即得。

$AX=\beta$ 的求解方法 用行初等变换化增广矩阵 $(A\ \ \beta)$ 为标准阶梯形 $\begin{array}{c}\left(\begin{array}{ccc}E_r & T & \delta\\ 0 & 0 & 0\end{array}\right)\begin{array}{l}\}r\text{行}\\ \}s-r\text{行}\end{array}\\ \begin{array}{ccc}\underbrace{}_{r\text{列}} & \underbrace{}_{t\text{列}} & \underset{\text{列}}{1}\end{array}\end{array}$,则 $\gamma=\begin{pmatrix}\delta\\0\\\vdots\\0\end{pmatrix}$ 为 $AX=\beta$ 的一个解向量,$\alpha_1,\cdots,\alpha_t$ 是 $AX=0$ 的基础解系。这里 α_i 是 $\begin{pmatrix}T\\-E_t\end{pmatrix}$ 的 i 列,$i=1,\cdots,t$。由性质(3),$AX=\beta$ 的通解为

$$X=\gamma+\sum_{i=1}^{t}k_i\alpha_i,$$

$\gamma=\begin{pmatrix}\delta\\0\\\vdots\\0\end{pmatrix}$ 是解向量,这是因为在与 $AX=\beta$ 同解的线性方程组

$\begin{pmatrix} E_r & T \\ 0 & 0 \end{pmatrix}\begin{pmatrix} x_1 \\ \vdots \\ x_r \\ x_{r+1} \\ \vdots \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} \delta \\ 0 \\ \vdots \\ 0 \end{pmatrix}$中令自由未知量 $x_{r+1} = \cdots = x_n = 0$，得 $\begin{pmatrix} x_1 \\ \vdots \\ x_r \end{pmatrix} =$ δ，从而 $\gamma = \begin{pmatrix} \delta \\ 0 \\ \vdots \\ 0 \end{pmatrix}$ 是 $AX = \beta$ 的一个解向量。

例 4.28 求 $\begin{pmatrix} 1 & 1 & 2 & 1 & 2 \\ 1 & 2 & 3 & 4 & 5 \\ 2 & 3 & 5 & 5 & 7 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{pmatrix} = \begin{pmatrix} 7 \\ 15 \\ 22 \end{pmatrix}$ 的通解。

解 增广矩阵 $= \begin{pmatrix} 1 & 1 & 2 & 1 & 2 & 7 \\ 1 & 2 & 3 & 4 & 5 & 15 \\ 2 & 3 & 5 & 5 & 7 & 22 \end{pmatrix} \xrightarrow[(2)-(1)]{(3)-(1)-(2)}$

$$\begin{pmatrix} 1 & 1 & 2 & 1 & 2 & 7 \\ 0 & 1 & 1 & 3 & 3 & 8 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} \xrightarrow{(1)-(2)} \begin{pmatrix} 1 & 0 & 1 & -2 & -1 & -1 \\ 0 & 1 & 1 & 3 & 3 & 8 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

$$= \begin{pmatrix} E_2 & \begin{matrix} 1 \\ 1 \end{matrix} & \begin{matrix} -2 \\ 3 \end{matrix} & \begin{matrix} -1 \\ 3 \end{matrix} & \begin{matrix} -1 \\ 8 \end{matrix} \\ 0 \quad 0 & 0 & 0 & 0 & 0 \end{pmatrix},$$

通解为

$$X = k_1\begin{pmatrix} 1 \\ 1 \\ -1 \\ 0 \\ 0 \end{pmatrix} + k_2\begin{pmatrix} -2 \\ 3 \\ 0 \\ -1 \\ 0 \end{pmatrix} + k_3\begin{pmatrix} -1 \\ 3 \\ 0 \\ 0 \\ -1 \end{pmatrix} + \begin{pmatrix} -1 \\ 8 \\ 0 \\ 0 \\ 0 \end{pmatrix}。$$

例 4.29 若 $\alpha_1,\cdots,\alpha_m$ 都是 $AX=\beta$ 的解向量。求证：$\frac{1}{m}\left(\sum_{i=1}^{m}\alpha_i\right)$ 也是 $AX=\beta$ 的解向量。

证 因为 $A\left(\frac{1}{m}\sum_{i=1}^{m}\alpha_i\right)=\frac{1}{m}\sum_{i=1}^{m}A\alpha_i=\frac{1}{m}(\underbrace{\beta+\cdots+\beta}_{m\text{个}})=\beta$。

例 4.30 设 A 为 $s\times n$ 阵，$\beta=(b_1,\cdots,b_n)$，若线性方程组 $AX=\mathbf{0}$ 的解全是 $\sum_{i=1}^{n}b_ix_i=0$ 的解，求证：$A_{(1)},\cdots,A_{(s)}\Rightarrow\beta$。

证 由假设可知，线性方程组 $AX=\mathbf{0}$ 与 $\begin{pmatrix}A\\ \beta\end{pmatrix}X=\mathbf{0}$ 同解，从而解空间的秩相同，即 $n-\mathrm{r}(A)=n-\mathrm{r}\begin{pmatrix}A\\ \beta\end{pmatrix}$，所以 $\mathrm{r}(A)=\mathrm{r}\begin{pmatrix}A\\ \beta\end{pmatrix}=t$。设 $\alpha_1,\cdots,\alpha_t$ 是 A 的行向量组的一个极大线性无关组，它也是 $\begin{pmatrix}A\\ \beta\end{pmatrix}$ 的行向量组的一个线性无关组，由命题 4.7，它是 $\begin{pmatrix}A\\ \beta\end{pmatrix}$ 的行向量组的一个极大线性无关组，由命题 4.5，$\alpha_1,\cdots,\alpha_t\Rightarrow\beta$，从而

$$A_{(1)},\cdots,A_{(s)}\Rightarrow\beta。$$

习 题 4

1. 用矩阵消元法解下列线性方程组：

$$(1)\begin{cases}2x_1-x_2+3x_3=3,\\3x_1+x_2-5x_3=0,\\4x_1-x_2+x_3=3,\\x_1+3x_2-13x_3=-6。\end{cases}$$

$$(2)\begin{cases}x_1-2x_2+x_3+x_4=1,\\x_1-2x_2+x_3-x_4=-1,\\x_1-2x_2+x_3-5x_4=5。\end{cases}$$

(3) $\begin{cases} x_1 - x_2 + x_3 = 0, \\ 3x_1 - 2x_2 - x_3 = 0, \\ 3x_1 - x_2 + 5x_3 = 0, \\ -2x_1 + 2x_2 + 3x_3 = 0。\end{cases}$

(4) $\begin{cases} 2x_1 + x_2 - x_3 + x_4 = 1, \\ 3x_1 - 2x_2 + 2x_3 - 3x_4 = 2, \\ 5x_1 + x_2 - x_3 + 2x_4 = -1, \\ 2x_1 - x_2 + x_3 - 3x_4 = 4。\end{cases}$

2. 将向量 $\boldsymbol{\beta}$ 表示为向量 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3,\boldsymbol{\alpha}_4$ 的线性组合：

(1) $\boldsymbol{\beta} = \begin{pmatrix} b_1 \\ b_2 \\ b_3 \\ b_4 \end{pmatrix}, \boldsymbol{\alpha}_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \boldsymbol{\alpha}_2 = \begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \boldsymbol{\alpha}_3 = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 0 \end{pmatrix}, \boldsymbol{\alpha}_4 = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}$；

(2) $\boldsymbol{\beta} = \begin{pmatrix} 1 \\ 2 \\ 1 \\ 1 \end{pmatrix}, \boldsymbol{\alpha}_1 = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}, \boldsymbol{\alpha}_2 = \begin{pmatrix} 1 \\ 1 \\ -1 \\ -1 \end{pmatrix}, \boldsymbol{\alpha}_3 = \begin{pmatrix} 1 \\ -1 \\ 1 \\ -1 \end{pmatrix}$；

$\boldsymbol{\alpha}_4 = \begin{pmatrix} 1 \\ -1 \\ -1 \\ 1 \end{pmatrix}$。

3. 求线性方程组 $x_1 - x_2 = a_1, x_2 - x_3 = a_2, x_3 - x_4 = a_3, x_4 - x_1 = a_4$ 有解的条件，在有解时求所有解。

4. 讨论 a,b 为何值时下列线性方程组有解，并求解：

(1) $\begin{cases} 2x_1 - x_2 + x_3 + x_4 = 1, \\ x_1 + 2x_2 - x_3 + 4x_4 = 2, \\ x_1 + 7x_2 - 4x_3 + 11x_4 = a。\end{cases}$

(2) $\begin{cases} x_1 + x_2 + x_3 + x_4 + x_5 = 1, \\ x_2 + 2x_3 + 2x_4 + 6x_5 = 3, \\ 3x_1 + 2x_2 + x_3 + x_4 - 3x_5 = a, \\ 5x_1 + 4x_2 + 3x_3 + 3x_4 - x_5 = b。\end{cases}$

(3) $\begin{cases} ax - y + z = 4, \\ x - by + z = 3, \\ x - 2by + z = 4。 \end{cases}$

(4) $\begin{cases} ax_1 + x_2 + x_3 = 1, \\ x_1 + ax_2 + x_3 = a, \\ x_1 + x_2 + ax_3 = a^2。 \end{cases}$

5. 讨论 k 取何值时下列齐次线性方程组有非零解，并求解：

(1) $\begin{cases} kx + y + z = 0, \\ x + ky - z = 0, \\ 2x - y + z = 0。 \end{cases}$

(2) $\begin{cases} kx + 3y + 3z = 0, \\ 3x + ky + 3z = 0, \\ 3x + 3y + kz = 0。 \end{cases}$

6. 用 Cramer 法则解下列线性方程组：

(1) $\begin{cases} 2x + 5y = 1, \\ 3x + 7y = 2。 \end{cases}$

(2) $\begin{cases} x + y - 2z = -3, \\ 5x - 2y + 7z = 22, \\ 2x - 5y + 4z = 4。 \end{cases}$

(3) $\begin{cases} -bx + ay = 2ab, \\ -2cy + 3bz = bc, \\ cx + az = 0。 \end{cases}$

7. 若 $\begin{pmatrix} c_1 \\ \vdots \\ c_n \end{pmatrix}$ 是齐次线性方程组

$$\begin{pmatrix} 1 & 1 & \cdots & 1 \\ 1 & 2 & \cdots & n \\ 1^2 & 2^2 & \cdots & n^2 \\ \vdots & \vdots & & \vdots \\ 1^{n-2} & 2^{n-2} & \cdots & n^{n-2} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

的一个非零解，求 $c_1 : c_n$。

8. 若向量组 $\begin{pmatrix}3\\3\\4\end{pmatrix}, \begin{pmatrix}1\\2\\1\end{pmatrix}, \begin{pmatrix}3\\a\\0\end{pmatrix}$ 线性相关，求 a 的值。

9. 若向量组 $\alpha_1,\cdots,\alpha_s$ 线性相关，求证：其中必有一个向量是其余向量的线性组合。

10. 设 $t_1,\cdots,t_r$ 是两两不同的数 $(r \leqslant n)$

$$\alpha_i = \begin{pmatrix}1\\t_i\\t_i^2\\\vdots\\\vdots\\t_i^n\end{pmatrix}, i = 1,\cdots,r$$

求证：向量组 $\alpha_1,\cdots,\alpha_r$ 线性无关。

11. 若向量组 $\alpha_1,\alpha_2,\alpha_3$ 线性无关，则向量组 $\alpha_1 + k\alpha_2,\alpha_2 + k\alpha_3,\alpha_3 + k\alpha_1$ 也线性无关。这里 $k \in \mathbf{Z}^+$。

12. s 是奇数，若向量组 $\alpha_1,\cdots,\alpha_s$ 线性无关，则向量组 $\alpha_1 + \alpha_2,\alpha_2 + \alpha_3,\cdots,\alpha_{s-1} + \alpha_s,\alpha_s + \alpha_1$ 也线性无关。

13. 求证：向量组 $\{\alpha_1,\cdots,\alpha_s,\beta_1,\cdots,\beta_t\}$ 的秩 $\leqslant \{\alpha_1,\cdots,\alpha_s\}$ 的秩 $+ \{\beta_1,\cdots,\beta_t\}$ 的秩。

14. 若向量组 Σ 的秩为 r，$\alpha_1,\cdots,\alpha_r \in \Sigma$，且 $\alpha_1,\cdots,\alpha_r \Rightarrow \Sigma$，求证：$\alpha_1,\cdots,\alpha_r$ 是 Σ 的一个极大线性无关组。

15. 设 $\boldsymbol{A}$ 为 n 阶方阵，若对任意 n 维向量 β，线性方程组 $\boldsymbol{AX} = \beta$ 都有解，求证：$|\boldsymbol{A}| \neq 0$。

16. 若向量组 $\{\alpha_1,\cdots,\alpha_s\}$ 与 $\{\alpha,\cdots,\alpha_s,\beta,\cdots,\beta_t\}$ 有相同的秩，求证：$\alpha_1,\cdots,\alpha_s \Rightarrow \beta_i, i = 1,\cdots,t$。

17. 若线性方程组 $\boldsymbol{AX} = \boldsymbol{0}$ 与 $\boldsymbol{BX} = \boldsymbol{0}$ 同解，求证：$\boldsymbol{A}$ 的行向量组 $\sim \boldsymbol{B}$ 的行向量组。

18. 若 $\boldsymbol{\beta}_i = \sum_{j=1}^{s}\boldsymbol{\alpha}_j - \boldsymbol{\alpha}_i, i = 1,\cdots,s$，求证：秩$\{\boldsymbol{\alpha}_1,\cdots,\boldsymbol{\alpha}_s\}$ = 秩$\{\boldsymbol{\beta}_1,\cdots,\boldsymbol{\beta}_s\}$。

19. 若向量组 $\boldsymbol{\alpha}_1,\cdots,\boldsymbol{\alpha}_s$ 线性无关

$$\boldsymbol{\beta}_i = \sum_{k=1}^{s} a_{ik}\boldsymbol{\alpha}_k, i = 1,\cdots,s$$

求证：$\boldsymbol{\beta}_1,\cdots,\boldsymbol{\beta}_s$ 线性无关 $\Longleftrightarrow |a_{ij}|_n \neq 0$。

20. 若向量组 $\boldsymbol{\alpha}_1,\cdots,\boldsymbol{\alpha}_s$ 的秩为 r，则在其中任取 m 个向量组成的向量组的秩 $\geqslant r + m - s$。

21. 求下列齐次线性方程组的一个基础解系：

(1) $$\begin{pmatrix} 1 & -2 & 4 & -7 \\ 2 & 1 & -2 & 1 \\ 3 & -1 & 2 & -4 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix};$$

(2) $$\begin{pmatrix} 1 & 1 & 1 & 1 & 1 \\ 3 & 2 & 1 & 1 & -3 \\ 0 & 1 & 2 & 2 & 6 \\ 5 & 4 & 3 & 3 & -1 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix};$$

(3) $$\begin{pmatrix} 2 & -1 & 1 & -1 \\ 2 & -1 & 0 & -3 \\ 0 & 1 & 3 & -6 \\ 2 & -2 & -2 & 5 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}。$$

22. 求下列线性方程组的通解：

(1) $$\begin{pmatrix} 1 & -2 & 4 & -7 \\ 2 & 1 & -2 & 1 \\ 3 & -1 & 2 & -4 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} -4 \\ 2 \\ 0 \end{pmatrix};$$

(2) $\begin{pmatrix} 1 & -2 & 1 & 1 & -1 \\ 2 & 1 & -1 & -1 & -1 \\ 1 & 7 & -5 & -5 & 5 \\ 3 & -1 & -2 & 1 & -1 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}$;

(3) $\begin{pmatrix} 1 & 1 & 1 & 1 & 1 \\ 3 & 2 & 1 & 1 & -3 \\ 0 & 1 & 2 & 2 & 6 \\ 5 & 4 & 3 & 3 & -1 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{pmatrix} = \begin{pmatrix} 5 \\ 4 \\ 11 \\ 14 \end{pmatrix}$。

23. 若$\boldsymbol{\alpha}_1,\cdots,\boldsymbol{\alpha}_s$是线性方程组$\boldsymbol{AX}=\boldsymbol{\beta}$的$s$个解向量，$\sum\limits_{i=1}^{s}k_i=1$，求证：$\sum\limits_{i=1}^{s}k_i\boldsymbol{\alpha}_i$也是$\boldsymbol{AX}=\boldsymbol{\beta}$的解向量。

24. 若$\boldsymbol{A},\boldsymbol{B}\in \boldsymbol{P}^{n\times n}$，且$\boldsymbol{AB}=\boldsymbol{0}$，求证：

$$\mathrm{r}(\boldsymbol{A})+\mathrm{r}(\boldsymbol{B})\leqslant n。$$

25. 若$\boldsymbol{A}\in \boldsymbol{P}^{n\times n}$，且$\boldsymbol{A}^2=\boldsymbol{E}$，求证：

$$\mathrm{r}(\boldsymbol{A}+\boldsymbol{E})+\mathrm{r}(\boldsymbol{A}-\boldsymbol{E})=n。$$

26. 若$\boldsymbol{\alpha}_1,\cdots,\boldsymbol{\alpha}_t$是$\boldsymbol{AX}=\boldsymbol{0}$的一个基础解系，$\boldsymbol{A\beta}\neq\boldsymbol{0}$，求证：向量组$\boldsymbol{\beta},\boldsymbol{\beta}+\boldsymbol{\alpha}_1,\cdots,\boldsymbol{\beta}+\boldsymbol{\alpha}_t$线性无关。

27. $\boldsymbol{A}\in \boldsymbol{P}^{(n-1)\times n}$，$\mathrm{r}(\boldsymbol{A})=n-1$，求证：$\boldsymbol{AX}=\boldsymbol{0}$的通解为

$$\boldsymbol{X}=k\begin{pmatrix} \boldsymbol{M}_1 \\ -\boldsymbol{M}_2 \\ \vdots \\ (-1)^{n-1}\boldsymbol{M}_n \end{pmatrix},$$

这里$\boldsymbol{M}_i$是划去$\boldsymbol{A}$的第i列后余下的$(n-1)$阶行列式的值。

28. $\boldsymbol{A}$为方阵，s为一正整数，若$\mathrm{r}(\boldsymbol{A}^s)=\mathrm{r}(\boldsymbol{A}^{s+1})$，求证：对任意$n\in\mathbf{Z}^+$，有$\mathrm{r}(\boldsymbol{A}^{s+n})=\mathrm{r}(\boldsymbol{A}^s)$。

第五章 相似矩阵

如果两个 n 阶方阵 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 存在这样的关系：$\boldsymbol{B}=\boldsymbol{T}^{-1}\boldsymbol{A}\boldsymbol{T}$，这里 $\boldsymbol{T}$ 是 n 阶可逆阵，则 $\boldsymbol{B}^k=(\boldsymbol{T}^{-1}\boldsymbol{A}\boldsymbol{T})(\boldsymbol{T}^{-1}\boldsymbol{A}\boldsymbol{T})\cdots(\boldsymbol{T}^{-1}\boldsymbol{A}\boldsymbol{T})=\boldsymbol{T}^{-1}\boldsymbol{A}^k\boldsymbol{T}$。更一般地，若 $f(x)$ 是一个多项式，那么 $f(\boldsymbol{B})=\boldsymbol{T}^{-1}f(\boldsymbol{A})\boldsymbol{T}$。如果 $\boldsymbol{A}$ 很简单，比如为对角阵 $\begin{pmatrix}\lambda_1 & & \\ & \ddots & \\ & & \lambda_n\end{pmatrix}$，则 $f(\boldsymbol{B})=\boldsymbol{T}^{-1}\begin{pmatrix}f(\lambda_1) & & \\ & \ddots & \\ & & f(\lambda_n)\end{pmatrix}\boldsymbol{T}$，这样使 $\boldsymbol{B}$ 的多项式的计算大大简化。本章研究矩阵的这种关系（称相似关系），并寻求一个方阵相似于对角阵的条件。

§1 相似的概念

定义 5.1 两个 n 阶方阵 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 称为相似的，记成 $\boldsymbol{A}\sim\boldsymbol{B}$，如果存在 n 阶可逆阵 $\boldsymbol{T}$ 使 $\boldsymbol{B}=\boldsymbol{T}^{-1}\boldsymbol{A}\boldsymbol{T}$。例如

$$\begin{pmatrix}1 & 1\\ 1 & -5\end{pmatrix}^{-1}\begin{pmatrix}3 & 1\\ 5 & -1\end{pmatrix}\begin{pmatrix}1 & 1\\ 1 & -5\end{pmatrix}$$
$$=\begin{pmatrix}\frac{5}{6} & \frac{1}{6}\\ \frac{1}{6} & -\frac{1}{6}\end{pmatrix}\begin{pmatrix}3 & 1\\ 5 & -1\end{pmatrix}\begin{pmatrix}1 & 1\\ 1 & -5\end{pmatrix}=\begin{pmatrix}4 & 0\\ 0 & -2\end{pmatrix},$$

故

$$\begin{pmatrix}3 & 1\\ 5 & -1\end{pmatrix}\sim\begin{pmatrix}4 & 0\\ 0 & -2\end{pmatrix}。$$

性质 (1) $\boldsymbol{A}\sim\boldsymbol{A}$（自反性）；

(2) 若 $\boldsymbol{A}\sim\boldsymbol{B}$，则 $\boldsymbol{B}\sim\boldsymbol{A}$（对称性）；

(3) 若 $A\sim B$，$B\sim C$，则 $A\sim C$（传递性）；

(4) 若 $A_i\sim B_i$，$i=1,\cdots,t$，则 $\begin{pmatrix} A_1 & & \\ & \ddots & \\ & & A_t \end{pmatrix}\sim\begin{pmatrix} B_1 & & \\ & \ddots & \\ & & B_t \end{pmatrix}$。

证 (1) 因为 $A=E^{-1}AE$，故 $A\sim A$。

(2) 若 $A\sim B$，即存在可逆阵 T 使 $B=T^{-1}AT$，此时，$A=(T^{-1})^{-1}BT^{-1}=T_1^{-1}BT_1$，这里 $T_1=T^{-1}$ 是可逆阵，故 $B\sim A$。

(3) 若 $A\sim B$，$B\sim C$，即存在可逆阵 T_1 与 T_2 使 $B=T_1^{-1}AT_1$，$C=T_2^{-1}BT_2$，此时 $C=T_2^{-1}T_1^{-1}AT_1T_2=T^{-1}AT$。这里 $T=T_1T_2$ 是可逆阵的积，仍可逆，故 $A\sim C$。

(4) $B_i=T_i^{-1}AT_i$，$i=1,\cdots,t$。于是

$$\begin{pmatrix} B & & \\ & \ddots & \\ & & B_t \end{pmatrix}=\begin{pmatrix} T_1^{-1}A_1T_1 & & \\ & \ddots & \\ & & T_t^{-1}A_tT_t \end{pmatrix}$$

$$=\begin{pmatrix} T_1 & & \\ & \ddots & \\ & & T_t \end{pmatrix}^{-1}\begin{pmatrix} A_1 & & \\ & \ddots & \\ & & A_t \end{pmatrix}\begin{pmatrix} T_1 & & \\ & \ddots & \\ & & T_t \end{pmatrix},$$

故

$$\begin{pmatrix} A_1 & & \\ & \ddots & \\ & & A_t \end{pmatrix}\sim\begin{pmatrix} B_1 & & \\ & \ddots & \\ & & B \end{pmatrix}。$$

如果 $A\sim$ 对角阵 $\begin{pmatrix} \lambda_1 & & & \\ & \ddots & & \\ & & \ddots & \\ & & & \lambda_n \end{pmatrix}$，即存在可逆阵 T 使

$$T^{-1}AT=\begin{pmatrix} \lambda_1 & & & \\ & \ddots & & \\ & & \ddots & \\ & & & \lambda_n \end{pmatrix},$$

从而

$$AT = T\begin{pmatrix} \lambda_1 & & & \\ & \ddots & & \\ & & \ddots & \\ & & & \lambda_n \end{pmatrix},$$

或

$$A(T^{(1)},T^{(2)},\cdots,T^{(n)}) = (T^{(1)},T^{(2)},\cdots,T^{(n)})\begin{pmatrix} \lambda_1 & & & \\ & \ddots & & \\ & & \ddots & \\ & & & \lambda_n \end{pmatrix}。$$

由分块乘法的法则得

$$(AT^{(1)},\cdots,AT^{(n)}) = (\lambda_1 T^{(1)},\cdots,\lambda_n T^{(n)}),$$

从而

$$AT^{(i)} = \lambda_i T^{(i)},i = 1,\cdots,n。$$

反之,若存在 n 个线性无关的列向量 $\alpha_1,\cdots,\alpha_n$ 满足

$$A\alpha_i = \lambda_i\alpha_i,i = 1,\cdots,n,$$

则 $T = (\alpha_1,\cdots,\alpha_n)$ 可逆且

$$\begin{aligned} AT &= (A\alpha_1,\cdots,A\alpha_n) = (\lambda_1\alpha_1,\cdots,\lambda_n\alpha_n) \\ &= (\alpha_1,\alpha_2,\cdots,\alpha_n)\begin{pmatrix} \lambda_1 & & & \\ & \ddots & & \\ & & \ddots & \\ & & & \lambda_n \end{pmatrix} = T\begin{pmatrix} \lambda_1 & & & \\ & \ddots & & \\ & & \ddots & \\ & & & \lambda_n \end{pmatrix} \end{aligned}$$

从而 $T^{-1}AT = \begin{pmatrix} \lambda_1 & & & \\ & \ddots & & \\ & & \ddots & \\ & & & \lambda_n \end{pmatrix}$,即 $A\sim$对角阵。故有

定理 5.1 n 阶方阵 A 相似于对角阵$\Longleftrightarrow$存在 n 个线性无关的向量

$\alpha_1,\cdots,\alpha_n$，满足 $A\alpha_i=\lambda_i\alpha_i, i=1,\cdots,n$，且此时与 A 相似的对角阵的对角线上元素为 $\lambda_1,\cdots,\lambda_n$。

例 5.1 设 $\boldsymbol{A}=\begin{pmatrix} a & b \\ b & a \end{pmatrix}$，$\boldsymbol{A}\begin{pmatrix} 1 \\ 1 \end{pmatrix}=(a+b)\begin{pmatrix} 1 \\ 1 \end{pmatrix}$，

$\boldsymbol{A}\begin{pmatrix} 1 \\ -1 \end{pmatrix}=(a-b)\begin{pmatrix} 1 \\ -1 \end{pmatrix}$；由于 $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$，$\begin{pmatrix} 1 \\ -1 \end{pmatrix}$ 线性无关，

$\boldsymbol{T}=\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$ 有逆 $\boldsymbol{T}^{-1}=\frac{1}{2}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$，

从而

$$\begin{aligned}\boldsymbol{T}^{-1}\boldsymbol{A}\boldsymbol{T}&=\frac{1}{2}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}\begin{pmatrix} a & b \\ b & a \end{pmatrix}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}\\&=\frac{1}{2}\begin{pmatrix} a+b & a+b \\ a-b & b-a \end{pmatrix}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}=\begin{pmatrix} a+b & 0 \\ 0 & a-b \end{pmatrix},\end{aligned}$$

故有

$$\boldsymbol{A}=\boldsymbol{T}\begin{pmatrix} a+b & 0 \\ 0 & a-b \end{pmatrix}\boldsymbol{T}^{-1},$$

更有

$$\begin{aligned}\boldsymbol{A}^n&=\boldsymbol{T}\begin{pmatrix} a+b & 0 \\ 0 & a-b \end{pmatrix}^n\boldsymbol{T}^{-1}\\&=\frac{1}{2}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}\begin{pmatrix} (a+b)^n & 0 \\ 0 & (a-b)^n \end{pmatrix}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}\\&=\begin{pmatrix} x & y \\ y & x \end{pmatrix}。\end{aligned}$$

这里，$x=\dfrac{(a+b)^n+(a-b)^n}{2}$，$y=\dfrac{(a+b)^n-(a-b)^n}{2}$。

§2 特征值与特征向量

一、定义

定义 5.2 设 A 是数域 P 上 n 阶方阵，若存在 $\lambda\in P$，$\alpha\in P^n$，且

$\alpha \neq 0$，使 $A\alpha = \lambda\alpha$，则称 λ 为 A 的特征值，α 为 A 的对应于特征值 λ 的特征向量。

例 5.2 $\begin{pmatrix} 3 & 1 \\ 5 & -1 \end{pmatrix}\begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 4 \\ 4 \end{pmatrix} = 4\begin{pmatrix} 1 \\ 1 \end{pmatrix}$，故 4 是方阵 $\begin{pmatrix} 3 & 1 \\ 5 & -1 \end{pmatrix}$ 的特征值，$\begin{pmatrix} 1 \\ 1 \end{pmatrix}$ 是对应于特征值 4 的特征向量。又 $\begin{pmatrix} 3 & 1 \\ 5 & -1 \end{pmatrix}\begin{pmatrix} 1 \\ 2 \end{pmatrix} = \begin{pmatrix} 5 \\ 3 \end{pmatrix} \neq \lambda\begin{pmatrix} 1 \\ 2 \end{pmatrix}$，所以 $\begin{pmatrix} 1 \\ 2 \end{pmatrix}$ 不是 $\begin{pmatrix} 3 & 1 \\ 5 & -1 \end{pmatrix}$ 的特征向量。

利用特征值和特征向量的概念，定理 5.1 可表达为：

定理 5.2 n 阶方阵 A 相似于对角阵$\Longleftrightarrow A$ 有 n 个线性无关的特征向量。此时，相似于 A 的对角阵中对角线上元素都是 A 的特征值。

二、求法

设 A 为 n 阶方阵，λ 是 A 的特征值，α 是对应于 λ 的特征向量。我们来寻找求 λ 与 α 的方法。

由 $A\alpha = \lambda\alpha$ 得 $(\lambda E - A)\alpha = 0$，此时 $\alpha = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ 是系数矩阵为 $(\lambda E - A)$ 的齐次线性方程组 $(\lambda E - A)X = 0$ 的非零解。由于齐次线性方程组 $(\lambda E - A)X = 0$ 有非零解，故必

$$|\lambda E - A| = 0。$$

这是以 λ 为未知数的代数方程，由此求出特征值 $\lambda = \lambda_1, \cdots, \lambda_s$。

齐次线性方程组 $(\lambda_i E - A)X = 0$ 的非零解即 A 的对应于特征值 λ_i 的特征向量 $(i = 1, \cdots, s)$。

由此得到特征值与特征向量的求法：

(1) 由 $|\lambda E - A| = 0$ 求出特征值 $\lambda_1, \cdots, \lambda_s$；

(2) 求出齐次线性方程组 $(\lambda_i E - A)X = 0$ 的非零解，即求得对应于 λ_i 的特征向量 $(i = 1, \cdots, s)$。

例 5.3 求 $A=\begin{pmatrix}5&4&4\\4&5&4\\4&4&5\end{pmatrix}$ 的特征值与特征向量。

解 (1) $$|\lambda \boldsymbol{E}-\boldsymbol{A}|=\begin{vmatrix}\lambda-5&-4&-4\\-4&\lambda-5&-4\\-4&-4&\lambda-5\end{vmatrix}$$

$$=(\lambda-13)\begin{vmatrix}1&1&1\\-4&\lambda-5&-4\\-4&-4&\lambda-5\end{vmatrix}$$

$$=(\lambda-13)\begin{vmatrix}1&1&1\\0&\lambda-1&0\\0&0&\lambda-1\end{vmatrix}$$

$$=(\lambda-13)(\lambda-1)^2,$$

由 $(\lambda-13)(\lambda-1)^2=0$ 得特征值 $\lambda=13,1$。

(2) $\lambda=13$, $(13\boldsymbol{E}-\boldsymbol{A})=\begin{vmatrix}8&-4&-4\\-4&8&-4\\-4&-4&8\end{vmatrix}$

$$\xrightarrow[\frac{-(2)}{4}]{\frac{-(1)}{4}\leftrightarrow\frac{-(3)}{4}}\begin{pmatrix}1&1&-2\\1&-2&1\\-2&1&1\end{pmatrix}\xrightarrow[\frac{(2)-(1)}{3}]{(3)+(2)+(1)}$$

$$\begin{pmatrix}1&1&-2\\0&-1&1\\0&0&0\end{pmatrix}\xrightarrow[-(2)]{(1)+(2)}\begin{pmatrix}1&0&-1\\0&1&-1\\0&0&0\end{pmatrix},$$

得非零解为 $k\begin{pmatrix}-1\\-1\\-1\end{pmatrix}$ $(k\neq 0)$，即对应于特征值 13 的特征向量为

$l\begin{pmatrix}1\\1\\1\end{pmatrix}$, $(l=-k\neq 0)$。

(3) $\lambda = 1, (1\boldsymbol{E} - \boldsymbol{A}) = \begin{pmatrix} -4 & -4 & -4 \\ -4 & -4 & -4 \\ -4 & -4 & -4 \end{pmatrix} \to \begin{pmatrix} 1 & 1 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$,

非零解为 $k_1 \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix} + k_2 \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}$,($k_1, k_2$ 不全为 0),此即对应于特征值 1 的特征向量。

例 5.4 求 $\boldsymbol{A} = \begin{pmatrix} a & b \\ b & a \end{pmatrix}$ 的特征值与特征向量 ($b \neq 0$)。

解 (1) 由 $|\lambda \boldsymbol{E} - \boldsymbol{A}| = \begin{vmatrix} \lambda - a & -b \\ -b & \lambda - a \end{vmatrix} = (\lambda - a)^2 - b^2 = 0$, 求得特征值 $\lambda = a + b, a - b$。

(2) $\lambda = a - b, (a + b)\boldsymbol{E} - \boldsymbol{A} = \begin{pmatrix} b & -b \\ -b & b \end{pmatrix} \longrightarrow \begin{pmatrix} 1 & -1 \\ 0 & 0 \end{pmatrix}$ 非零解(即对应于 $a + b$ 的特征向量)为 $k \begin{pmatrix} -1 \\ -1 \end{pmatrix} = l \begin{pmatrix} 1 \\ 1 \end{pmatrix}$,($l = -k \neq 0$)。

(3) $\lambda = a - b, (a - b)\boldsymbol{E} - \boldsymbol{A} = \begin{pmatrix} -b & -b \\ -b & -b \end{pmatrix} \longrightarrow \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix}$, 对应于 $a - b$ 的特征向量为 $k \begin{pmatrix} 1 \\ -1 \end{pmatrix}$,($k \neq 0$)。

三、特征多项式

定义 5.3 称未知量 λ 的多项式 $f(\lambda) = |\lambda \boldsymbol{E} - \boldsymbol{A}|$ 为方阵 $\boldsymbol{A}$ 的特征多项式。

若 $\boldsymbol{A} = (a_{ij})_{n \times n}$, 则

$$|\lambda \boldsymbol{E} - \boldsymbol{A}| = \begin{vmatrix} \lambda - a_{11} & -a_{12} & \cdots\cdots & -a_{1n} \\ -a_{21} & \lambda - a_{22} & \cdots\cdots & -a_{2n} \\ \vdots & \vdots & & \vdots \\ \vdots & \vdots & & \vdots \\ -a_{n1} & -a_{n2} & \cdots\cdots & \lambda - a_{nn} \end{vmatrix}$$
$$= (\lambda - a_{11})(\lambda - a_{22})\cdots(\lambda - a_{nn})$$

$$+（次数不大于(n-2)的单项）$$

$$=\lambda^n-\left(\sum_{i=1}^{n}a_{ii}\right)\lambda^{n-1}+\cdots+f(0),$$

$f(\lambda)$的常数项为 $f(0)=|-\boldsymbol{A}|=(-1)^n|\boldsymbol{A}|$。

推论 方阵 $\boldsymbol{A}$ 的特征值是它的特征多项式的零点。

定理 5.3 相似矩阵有相同的特征多项式，从而有相同的特征值。

证 若 $\boldsymbol{B}=\boldsymbol{T}^{-1}\boldsymbol{A}\boldsymbol{T}$，则

$$\begin{aligned}|\lambda\boldsymbol{E}-\boldsymbol{B}|&=|\lambda\boldsymbol{E}-\boldsymbol{T}^{-1}\boldsymbol{A}\boldsymbol{T}|=|\boldsymbol{T}^{-1}(\lambda\boldsymbol{E}-\boldsymbol{A})\boldsymbol{T}|\\&=|\boldsymbol{T}^{-1}|\cdot|\lambda\boldsymbol{E}-\boldsymbol{A}|\cdot|\boldsymbol{T}|=|\boldsymbol{T}^{-1}|\cdot|\boldsymbol{T}|\cdot|\lambda\boldsymbol{E}-\boldsymbol{A}|\\&=|\boldsymbol{E}|\cdot|\lambda\boldsymbol{E}-\boldsymbol{A}|=|\lambda\boldsymbol{E}-\boldsymbol{A}|。\end{aligned}$$

定理得证。

命题 5.1 相似矩阵的秩相同。

证 $\boldsymbol{B}=\boldsymbol{T}^{-1}\boldsymbol{A}\boldsymbol{T}$，故 $\mathrm{r}(\boldsymbol{B})\leqslant\mathrm{r}(\boldsymbol{A})$。又 $\boldsymbol{A}=\boldsymbol{T}\boldsymbol{B}\boldsymbol{T}^{-1}$，所以 $\mathrm{r}(\boldsymbol{A})\leqslant\mathrm{r}(\boldsymbol{B})$。于是 $\mathrm{r}(\boldsymbol{A})=\mathrm{r}(\boldsymbol{B})$。

例 5.5 若 $\boldsymbol{A}=(a_{ij})_{n\times n}\sim\boldsymbol{B}=\begin{pmatrix}1&&&&\\&\ddots&&&\\&&1&&\\&&&0&\\&&&&\ddots\\&&&&&0\end{pmatrix}$，则

$$\mathrm{r}(\boldsymbol{A})=\sum_{i=1}^{n}a_{ii}。$$

证 由 $|\lambda\boldsymbol{E}-\boldsymbol{A}|=|\lambda\boldsymbol{E}-\boldsymbol{B}|$，比较两边 x^{n-1} 的系数，可知 $\sum_{i=1}^{n}a_{ii}=\boldsymbol{B}$ 的对角线上元素之和 $=\mathrm{r}(\boldsymbol{B})=\mathrm{r}(\boldsymbol{A})$。

命题 5.2 若 $\boldsymbol{A}\in\boldsymbol{P}^{n\times n}$，$f(x)\in\boldsymbol{P}[x]$，$\lambda$ 是 $\boldsymbol{A}$ 的特征值，则 $f(\lambda)$是 $f(\boldsymbol{A})$的特征值。

证 设 $\boldsymbol{A}\alpha=\lambda\alpha$，$\alpha\neq0$，即 α 是 $\boldsymbol{A}$ 的对应于特征值 λ 的特征向量。设

$$f(x)=a_0+a_1x+\cdots+a_lx^l,$$

则

$$\begin{aligned}f(\boldsymbol{A})\boldsymbol{\alpha} &= [a_0\boldsymbol{E}+a_1\boldsymbol{A}+\cdots+a_t\boldsymbol{A}^t]\boldsymbol{\alpha}\\ &= a_0\boldsymbol{\alpha}+a_1\boldsymbol{A}\boldsymbol{\alpha}+\cdots+a_t\boldsymbol{A}^t\boldsymbol{\alpha}\\ &= a_0\boldsymbol{\alpha}+a_1\lambda\boldsymbol{\alpha}+\cdots+a_t\lambda^t\boldsymbol{\alpha}\\ &= [a_0+a_1\lambda+\cdots+a_t\lambda^t]\boldsymbol{\alpha}\\ &= f(\lambda)\boldsymbol{\alpha},\end{aligned}$$

故 $f(\lambda)$是 $f(\boldsymbol{A})$的特征值。

例 5.6 求矩阵

$$\boldsymbol{B}=\begin{vmatrix}a & b & c\\ c & a & b\\ b & c & a\end{vmatrix}$$

的特征值。

解 令 $\boldsymbol{A}=\begin{pmatrix}0&1&0\\0&0&1\\1&0&0\end{pmatrix}$，则

$$\boldsymbol{A}^2=\begin{pmatrix}0&0&1\\1&0&0\\0&1&0\end{pmatrix},\boldsymbol{A}^3=\boldsymbol{E},$$

于是

$$\boldsymbol{B}=a\boldsymbol{E}+b\boldsymbol{A}+c\boldsymbol{A}^2=f(\boldsymbol{A}),f(x)=a+bx+cx^2,$$

由

$$|\lambda\boldsymbol{E}-\boldsymbol{A}|=\begin{vmatrix}\lambda & -1 & 0\\ 0 & \lambda & -1\\ -1 & 0 & \lambda\end{vmatrix}=\lambda^3-1=0$$

得 $\boldsymbol{A}$ 的三个特征值：$1,\omega,\omega^2$；这里 $\omega=\dfrac{-1+\sqrt{3}\mathrm{i}}{2}$。所以，$\boldsymbol{B}=f(\boldsymbol{A})$ 的三个特征值为：

$$\begin{aligned}f(1)&=a+b+c,\\ f(\omega)&=a+b\omega+c\omega^2,\\ f(\omega^2)&=a+b\omega^2+c\omega。\end{aligned}$$

四、特征向量的性质

命题 5.3 设 $A\in P^{n\times n}$，记 $V_\lambda=\{\alpha\in P^n\mid A\alpha=\lambda\alpha\}$，若 $\alpha_1,\cdots,\alpha_t\in V_\lambda,k_1,\cdots,k_t\in P$，则 $\sum\limits_{i=1}^{t}k_i\alpha_i\in V_\lambda$。

证 因为 $A(\sum k_i\alpha_i)=\sum k_iA\alpha_i=\sum k_i\lambda\alpha_i=\lambda(\sum k_i\alpha_i)$。所以 $\sum k_i\alpha_i\in V_\lambda$。

命题 5.4 方阵 A 的对应于不同特征值的特征向量线性无关。

证 设 $\lambda_1,\cdots,\lambda_s$ 是 A 的两两不同的特征值，对应的特征向量分别为 $\alpha_1,\cdots,\alpha_s$。

如果 $x_1\alpha_1+\cdots+x_s\alpha_s=0$，两边逐次左乘 A，利用 $A\alpha_i=\lambda_i\alpha_i$，可得

$$\begin{aligned}
\lambda_1x_1\alpha_1+\cdots\cdots+\lambda_sx_s\alpha_s&=0,\\
\lambda_1^2x_1\alpha_1+\cdots\cdots+\lambda_s^2x_s\alpha_s&=0,\\
\vdots\qquad\qquad&\vdots\\
\lambda_1^{s-1}x_1\alpha_1+\cdots\cdots+\lambda_s^{s-1}x_s\alpha_s&=0,
\end{aligned}$$

写成矩阵形式为

$$(x_1\alpha_1,x_2\alpha_2,\cdots,x_s\alpha_s)\begin{pmatrix}1&\lambda_1&\lambda_1^2&\cdots&\lambda_1^{s-1}\\1&\lambda_2&\lambda_2^2&\cdots&\lambda_2^{s-1}\\\vdots&\vdots&\vdots&&\vdots\\1&\lambda_s&\lambda_s^2&\cdots&\lambda_s^{s-1}\end{pmatrix}=0。$$

s 阶方阵 $B=\begin{pmatrix}1&\lambda_1&\cdots&\lambda_1^{s-1}\\1&\lambda_2&\cdots&\lambda_2^{s-1}\\\vdots&\vdots&&\vdots\\1&\lambda_s&\cdots&\lambda_s^{s-1}\end{pmatrix}$，它的行列式为 Vandermonde 行列式且其值不为零，故 B 可逆。上式两边右乘 B^{-1} 得

$$(x_1\alpha_1,\cdots,x_s\alpha_s)=0,$$

即

$$x_i\alpha_i = 0, \quad (i = 1,\cdots,s)。$$

因为 α_i 是特征向量，$\alpha_i \neq 0$，必 $x_i = 0\ (i = 1,\cdots,s)$。故向量组 $\alpha_1,\cdots,\alpha_s$ 线性无关。

推论 若 n 阶方阵有 n 个不同特征值，则相似于对角阵。

证 这是因为对应于每个特征值至少有一个特征向量，从而有 n 个特征向量，它们对应于不同特征值，因此线性无关，由定理 5.2，方阵相似于对角阵。

例 5.7 在例 5.3 中，$A = \begin{bmatrix} 5 & 4 & 4 \\ 4 & 5 & 4 \\ 4 & 4 & 5 \end{bmatrix}$，对应于特征值 13，$A$ 有特征向量 $\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$，对应于特征值 1，$A$ 有线性无关的特征向量 $\begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}$，$\begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}$。根据命题 5.3 与 5.4，$\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$，$\begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}$，$\begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}$ 是 A 的三个线性无关的特征向量。

§3 Jordan 标准形

给了方阵 A，不一定存在与 A 相似的对角阵。我们要问，与 A 相似的矩阵中哪一个最简单。在复矩阵的情形，答案是 Jordan 标准形。

定义 5.4 形如 $J_i = \begin{bmatrix} \lambda_i & & & & \\ 1 & \ddots & & \boldsymbol{0} & \\ & \ddots & \ddots & & \\ & \boldsymbol{0} & \ddots & \ddots & \\ & & & 1 & \lambda_i \end{bmatrix}_{n_i \times n_i}$ 的方阵称为 n_i

阶若 Jordan 块，其中 λ_i 为复数。由 Jordan 块组成的准对角矩阵

$$J=\begin{pmatrix} J_1 & & & \\ & \ddots & & \\ & & \ddots & \\ & & & \ddots & \\ & & & & J_s \end{pmatrix}$$

称为 Jordan 标准形。

例如，$\begin{pmatrix} 2 & 0 & 0 \\ 1 & 2 & 0 \\ 0 & 1 & 2 \end{pmatrix}$，$\begin{pmatrix} 0 & & & & \\ 1 & \ddots & & \boldsymbol{0} & \\ & \ddots & \ddots & & \\ & \boldsymbol{0} & \ddots & \ddots & \\ & & & 1 & 0 \end{pmatrix}$，$\begin{pmatrix} i & 0 \\ 1 & i \end{pmatrix}$，$(\lambda)$都是 Jordan 块。特别，一阶方阵是当然的 Jordan 块。方阵

$$\begin{pmatrix} 1 & 0 & & & & \\ 1 & 1 & & & & \\ & & 3 & & & \\ & & & i & 0 & \\ & & & 1 & i & \\ & & & & & i \end{pmatrix}$$

是由 4 个 Jordan 块组成的 Jordan 标准形。任意 n 阶对角阵是由 n 个 1 阶 Jordan 块组成的 Jordan 标准形。

下面叙述复数域上方阵相似标准形理论的基本定理，其证明将在第十章中给出。

定理 5.4 复方阵 A 必相似于一 Jordan 标准形。如果不考虑 Jordan 块的排列次序，Jordan 标准形由 A 惟一确定。

例 5.8 求下列方阵的 Jordan 标准形。

$$(1)\ \boldsymbol{A}=\begin{pmatrix}5&3\\2&4\end{pmatrix},\ (2)\ \boldsymbol{B}=\begin{pmatrix}3&0&0&&\\1&3&0&&\\0&1&3&&\\&&&5&3\\&&&2&4\end{pmatrix}$$

解 (1) 由 $|\lambda\boldsymbol{E}-\boldsymbol{A}|=\begin{vmatrix}\lambda-5&-3\\-2&\lambda-4\end{vmatrix}=\lambda^2-9\lambda+14=(\lambda-2)(\lambda-7)=0$ 得 $\boldsymbol{A}$ 的特征值为 2,7。因而 2 阶方阵 $\boldsymbol{A}$ 相似于对角阵 $\begin{pmatrix}2&0\\0&7\end{pmatrix}$,此即 $\boldsymbol{A}$ 的 Jordan 标准形。

(2) 因为 $\begin{pmatrix}5&3\\2&4\end{pmatrix}\sim\begin{pmatrix}2&0\\0&7\end{pmatrix}$,由矩阵相似的性质(4)(见§1)得

$$\boldsymbol{B}\sim\begin{pmatrix}3&0&0&&\\1&3&0&&\\0&1&3&&\\&&&2&0\\&&&0&7\end{pmatrix}$$

后者即为 $\boldsymbol{B}$ 的 Jordan 标准形。

§4 方阵的最小多项式

一、方阵的化零多项式

定义 5.5 $\boldsymbol{A}$ 是数域 P 上 n 阶方阵,数域 P 上多项式 $f(x)$ 叫 $\boldsymbol{A}$ 的化零多项式,如果 $f(\boldsymbol{A})=0$。

例如,因为 $\begin{pmatrix}3&1\\5&-1\end{pmatrix}^2-2\begin{pmatrix}3&1\\5&-1\end{pmatrix}-8\boldsymbol{E}_2=0$,所以 $f(x)=x^2-2x-8$ 是矩阵 $\begin{pmatrix}3&1\\5&-1\end{pmatrix}$ 的化零多项式。

命题 5.5 任意方阵都有非零化零多项式。

证 设 $\boldsymbol{A}$ 为 n 阶方阵。若把 n 阶方阵的 n^2 个元素排成一列,则可

把 n 阶方阵看成 n^2 维向量，加法与数乘运算如旧。于是，n^2+1 个 n^2 维向量 $\boldsymbol{E}_n, \boldsymbol{A}, \boldsymbol{A}^2, \cdots, \boldsymbol{A}^{n^2}$ 必线性相关，即存在 n^2+1 个不全为 0 的数 k_0，$k_1, \cdots, k_{n^2}$ 使

$$k_0\boldsymbol{E} + k_1\boldsymbol{A} + \cdots + k_{n^2}\boldsymbol{A}^{n^2} = 0。$$

从而非零多项式 $k_0 + k_1x + \cdots + k_{n^2}x^{n^2}$ 是方阵 $\boldsymbol{A}$ 的非零化零多项式。

二、方阵的最小多项式

定义 5.6 方阵 $\boldsymbol{A}$ 的次数最低的首 1 化零多项式称为 $\boldsymbol{A}$ 的最小多项，记成 $m_{\boldsymbol{A}}(x)$。

定理 5.5 任意方阵 $\boldsymbol{A}$ 均有最小多项式。

证 由命题 5.5，$\boldsymbol{A}$ 有非零化零多项式，从而有次数最小的非零化零多项式，首一化即得结论。

命题 5.6 $f(x)$是方阵 $\boldsymbol{A}$ 的化零多项式 $\Leftrightarrow m_{\boldsymbol{A}}(x) \mid f(x)$。

证 $\Leftarrow$若 $f(x) = m_{\boldsymbol{A}}(x)g(x)$，则 $f(\boldsymbol{A}) = m_{\boldsymbol{A}}(\boldsymbol{A})g(\boldsymbol{A}) = 0$。

$\Rightarrow$作带余除法，$f(x) = q(x)m_{\boldsymbol{A}}(x) + r(x)$，

$\deg r(x) < \deg m_{\boldsymbol{A}}(x)$。两边用 $\boldsymbol{A}$ 代 x，得 $r(\boldsymbol{A}) = 0$，由 $m_{\boldsymbol{A}}(x)$的定义，必 $r(x) = 0$，即 $m_{\boldsymbol{A}}(x) \mid f(x)$。

例 5.9 Jordan 块 $\boldsymbol{J}_0 = \begin{pmatrix} \lambda_0 & & & & \\ 1 & \ddots & & 0 & \\ & \ddots & \ddots & & \\ & & \ddots & \ddots & \\ & 0 & \ddots & \ddots & \\ & & & 1 & \lambda_0 \end{pmatrix}_{n_0 \times n_0}$ 的最小多项式为 $(x-\lambda_0)^{n_0}$，即 $\boldsymbol{J}_0$ 的特征多项式 $|x\boldsymbol{E} - \boldsymbol{J}_0|$。

证 $(\boldsymbol{J}_0 - \lambda_0\boldsymbol{E})^{n_0} = \begin{pmatrix} 0 & & & & \\ 1 & \ddots & & \boldsymbol{0} & \\ & \ddots & \ddots & & \\ & & \ddots & \ddots & \\ & \boldsymbol{0} & & \ddots & \ddots \\ & & & 1 & 0 \end{pmatrix}_{n_0 \times n_0}^{n_0} = \boldsymbol{0}$，

由命题 5.4 $m_{J_0}(x)\mid(x-\lambda_0)^{n_0}$，故 $m_{J_0}(x)=(x-\lambda_0)^k$，$k\leqslant n_0$，但 $k<n$ 时，$(\boldsymbol{J}_0-\lambda_0\boldsymbol{E})^k\neq 0$，故

$$m_{J_0}(x)=(x-\lambda_0)^{n_0}=|x\boldsymbol{E}-\boldsymbol{J}_0|。$$

命题 5.7 相似方阵有相同的最小多项式。

证 若 $\boldsymbol{B}=\boldsymbol{T}^{-1}\boldsymbol{A}\boldsymbol{T}$，则 $m_A(\boldsymbol{B})=\boldsymbol{T}^{-1}m_A(\boldsymbol{A})\boldsymbol{T}=0$，故 $m_B(x)\mid m_A(x)$。同理 $m_A(x)\mid m_B(x)$。故两个首 1 多项式 $m_A(x)$ 与 $m_B(x)$ 相等。

推论 方阵的最小多项式是惟一的。

命题 5.8 若 A 为准对角阵 $\quad A=\begin{pmatrix}\boldsymbol{A}_1 & & \\ & \ddots & \\ & & \boldsymbol{A}_s\end{pmatrix}$，则

$$m_A(x)=[m_{A_1}(x),\cdots,m_{A_s}(x)]。$$

这里 $[m_{A_1}(x),\cdots,m_{A_s}(x)]$，表 $m_{A_1}(x),\cdots,m_{A_s}(x)$ 的最小公倍式。

证 记 $[m_{A_1}(x),\cdots,m_{A_s}(x)]=f(x)$。

$$f(\boldsymbol{A})=\begin{pmatrix}f(\boldsymbol{A}_1) & & \\ & \ddots & \\ & & f(\boldsymbol{A}_s)\end{pmatrix}。$$

因为 $m_{A_i}(x)\mid f(x)$，所以 $f(\boldsymbol{A}_i)=0$，$i=1,\cdots,s$，故 $f(\boldsymbol{A})=0$，于是 $m_A(x)\mid f(x)$。

又

$$\boldsymbol{0}=m_A(\boldsymbol{A})=\begin{pmatrix}m_A(\boldsymbol{A}_1) & & \\ & \ddots & \\ & & m_A(\boldsymbol{A}_s)\end{pmatrix},$$

即 $m_A(\boldsymbol{A}_i)=\boldsymbol{0}$，故 $m_{A_i}(x)\mid m_A(x)$，$i=1,\cdots,s$，从而

$$f(x)\mid m_A(x)。$$

所以

$$f(x)=m_A(x)。$$

命题 5.9 Jordan 标准形 $J = \begin{pmatrix} J_1 & & \\ & \ddots & \\ & & J_s \end{pmatrix}$ 的最小多项式为 $[(x-\lambda_1)^{n_1},\cdots,(x-\lambda_s)^{n_s}]$，这里 $J_i = \begin{pmatrix} \lambda_i & & & & \boldsymbol{0} \\ 1 & \ddots & & & \\ & \ddots & \ddots & & \\ & & \ddots & \ddots & \\ \boldsymbol{0} & & & 1 & \lambda_i \end{pmatrix}_{n_i \times n_i}$，$i = 1,\cdots,s$。

命题 5.10 方阵 A 的特征多项式 $f(x) = |xE - A|$ 是 A 的化零多项式。

证 在复数域内

$$A \sim J = \begin{pmatrix} J_1 & & \\ & \ddots & \\ & & J_s \end{pmatrix},$$

$$m_A(x) = m_J(x) = [(x-\lambda_1)^{n_1},\cdots,(x-\lambda_s)^{n_s}],$$

而

$$|xE - A| = |xE - J| = (x-\lambda_1)^{n_1}\cdot\cdots\cdot(x-\lambda_s)^{n_s},$$

故得证明。

因为 $m_A(x)$是 $(x-\lambda_1)^{n_1},\cdots,(x-\lambda_s)^{n_s}$ 的最小公倍式，必

$$(x-\lambda_i)^{n_i} \mid m_A(x),$$

从而

$$m_A(\lambda_i) = 0,$$

即特征值是最小多项式的零点。故得下面推论。

推论 方阵 A 的特征值必是 $m_A(x)$的根。

三、相似于对角阵的矩阵

定理 5.6 数域 P 上方阵 A 相似于对角阵$\Leftrightarrow A$ 的最小多项式在 $P[x]$中能分解成不同一次因子的积。

证 $\Rightarrow$设 $A\sim\begin{pmatrix}\lambda_1 & & \\ & \ddots & \\ & & \lambda_n\end{pmatrix}$，则 A 的最小多项式 $=[\lambda-\lambda_1,\cdots,\lambda-\lambda_n]$，故是不同的一次因子之积。

$\Leftarrow$由命题 5.6 推论 2，A 的 Jordan 标准形 $J\in P^{n\times n}$：

$$J=\begin{pmatrix}J_1 & & & \\ & \ddots & & \\ & & \ddots & \\ & & & J_s\end{pmatrix},J_i=\begin{pmatrix}\lambda_i & & & & \\ 1 & \ddots & & \boldsymbol{0} & \\ & \ddots & \ddots & & \\ & & \ddots & \ddots & \\ & \boldsymbol{0} & & & \\ & & & 1 & \lambda_i\end{pmatrix}_{n_i\times n_i}。$$

由命题 5.8，J_i 的最小多项式 $(x-\lambda_i)^{n_i}$ 是 A 的最小多项式的因子，而后者是不同的一次因子之积，故必 $n_i=1$，即 Jordan 块全是一阶的，从而 Jordan 标准形是 P 上的对角阵。因此 A 与对角阵 J 相似。

注 若 $B\sim A,B,A\in P^{n\times n}$，$B$ 为对角阵，则存在 $T\in P^{n\times n}$ 使 $B=T^{-1}AT$。这是因为 T 的列向量是 A 的特征向量，是 P 上线性方程组 $(\lambda E-A)X=\boldsymbol{0}$ 的非零解。

例 5.10 若方阵 A 满足 $A^2-5A+6E=\boldsymbol{0}$，则

$$A\sim\begin{pmatrix}2E_r & \boldsymbol{0}\\ \boldsymbol{0} & 3E_t\end{pmatrix},r,t\geqslant 0。$$

证 由假设，多项式 $f(x)=x^2-5x+6=(x-2)(x-3)$ 是 A 的化零多项式。由于 $f(x)$ 是不同一次因子的积，A 的最小多项式是 $f(x)$的因子，因此也是不同一次因子的积，所以 A 相似于某个对角阵。此对角阵的对角线上元素是 A 的特征值，故必是最小多项式的根，从而也是化零多项式 $f(x)$的根，即 2 或 3。于是有非负整数 r 与 t 使

$$A \sim \begin{pmatrix} 2\boldsymbol{E}_r & \boldsymbol{0} \\ \boldsymbol{0} & 3\boldsymbol{E}_t \end{pmatrix}。$$

例 5.11 若方阵 $\boldsymbol{A}$ 满足

$$\boldsymbol{A}^2 - 5\boldsymbol{A} + 6\boldsymbol{E} = \boldsymbol{0}\ ,$$

求 $\boldsymbol{A}^{100}$。

解 用 $x^2 - 5x + 6$ 除 x^{100}，得商 $q(x)$ 及余式 $ax + b$：

$$x^{100} = q(x)(x^2 - 5x + 6) + ax + b。$$

令 $x = 2$，得

$$2^{100} = 2a + b,$$

令 $x = 3$，得

$$3^{100} = 3a + b,$$

解之，得

$$a = 3^{100} - 2^{100}, b = -2 \cdot 3^{100} + 3 \cdot 2^{100}.$$

于是

$$\begin{aligned} \boldsymbol{A}^{100} &= q(\boldsymbol{A})(\boldsymbol{A}^2 - 5\boldsymbol{A} + 6\boldsymbol{E}) + a\boldsymbol{A} + b\boldsymbol{E} \\ &= (3^{100} - 2^{100})\boldsymbol{A} + (3 \cdot 2^{100} - 2 \cdot 3^{100})\boldsymbol{E}。\end{aligned}$$

例 5.12 若方阵 $\boldsymbol{A}$ 满足 $\boldsymbol{A}^2 = \boldsymbol{A}$，求证：

$$\mathrm{r}(\boldsymbol{A}) = \sum_i \boldsymbol{A}(i,i)。$$

证 $\boldsymbol{A}$ 有化零多项式 $x^2 - x = x(x-1)$。由于 $x^2 - x$ 无重根，因而 $\boldsymbol{A}$ 相似于对角阵。因为 $x^2 - x$ 的根为 0 或 1，于是 $\boldsymbol{A}$ 的特征值为 0 或 1，从而 $\boldsymbol{A}$ 相似对角阵 $\begin{pmatrix} \boldsymbol{E}_r & \\ & \boldsymbol{0} \end{pmatrix}$，这里 r 是 $\boldsymbol{A}$ 的秩。由例题 5.5，得

$$\mathrm{r}(\boldsymbol{A}) = \sum_i \boldsymbol{A}(i,i)。$$

例 5.13 设 $\boldsymbol{A} = \begin{pmatrix} 1 & 4 & 5 \\ 0 & -2 & 3 \\ 0 & -1 & 2 \end{pmatrix}$，求 $\boldsymbol{A}^{100}$。

解 $|\lambda E-A|=\begin{vmatrix}\lambda-1 & -4 & -5\\ 0 & \lambda+2 & -3\\ 0 & 1 & \lambda-2\end{vmatrix}=(\lambda-1)[\lambda^2-4+3]=(\lambda-1)^2(\lambda+1)$，故 A 有化零多项式 $(x-1)^2(x+1)=f(x)$。

设用 $(x-1)^2(x+1)$ 除 x^{100}得商 $q(x)$与余式 ax^2+bx+c：

$$x^{100}=q(x)(x-1)^2(x+1)+ax^2+bx+c,$$

两边求导，得

$$100x^{99}=(x-1)p(x)+2ax+b,$$

此处 $p(x)$为某个多项式。

令 $x=1$，得

$$1=a+b+c,$$
$$100=2a+b,$$

令 $x=-1$，得

$$(-1)^{100}=a-b+c.$$

由

$$\begin{cases}a+b+c=1,\\ 2a+b\quad\quad=100,\\ a-b+c=1,\end{cases}$$

得

$$b=0, a=50, c=-49,$$

于是

$$\begin{aligned}\boldsymbol{A}^{100}&=q(\boldsymbol{A})f(\boldsymbol{A})+a\boldsymbol{A}^2+b\boldsymbol{A}+c\boldsymbol{E}\\&=50\boldsymbol{A}^2-49\boldsymbol{E}\\&=50\begin{pmatrix}1&-9&27\\0&1&0\\0&0&1\end{pmatrix}-49\begin{pmatrix}1&0&0\\0&1&0\\0&0&1\end{pmatrix}\end{aligned}$$

$$
= \begin{pmatrix} 1 & -450 & 1350 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}。
$$

§5　向量的内积

仿照平面几何与几何空间中向量的内积，本节引进 n 维向量的内积，从而导出向量的长度与正交化的概念。

一、共轭矩阵

定义 5.7　设 $A=(a_{ij})_{s\times n}$ 是复数域上的一个 $s\times n$ 阵，若用 $\overline{a}_{ij}$ 表示数 a_{ij} 的共轭复数，则称 $s\times n$ 阵 $(\overline{a}_{ij})_{s\times n}$ 为 A 的共轭矩阵，记成 $\overline{A}$。例如：

$$
\overline{\begin{pmatrix} 3 & 3+\mathrm{i} \\ i & 2-\mathrm{i} \end{pmatrix}} = \begin{pmatrix} 3 & 3-\mathrm{i} \\ -i & 2+\mathrm{i} \end{pmatrix}。
$$

由复数运算的性质，易得：

性质　(1) $\overline{(\overline{A})} = A$；

(2) $\overline{kA} = \overline{k}\,\overline{A}$；

(3) $\overline{A'} = (\overline{A})'$；

(4) $\overline{A+B} = \overline{A}+\overline{B}$；

(5) $\overline{AB} = \overline{A}\,\overline{B}$；

(6) 若 A 可逆，则 $\overline{A^{-1}} = (\overline{A})^{-1}$。

二、向量的内积

定义 5.8　设 $\alpha=\begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix}$，$\beta=\begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix}$ 是两个 n 维列向量，称数 $\overline{\alpha'}\cdot\beta=(\overline{a_1},\cdots,\overline{a_n})\begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix}=\sum\limits_{i=1}^{n}\overline{a_i}b_i$ 为向量 α 与 β 的内积，记成 $[\alpha,\beta]$。例如：

$$\left[\begin{pmatrix}1+\mathrm{i}\\2\\2-\mathrm{i}\end{pmatrix},\begin{pmatrix}\mathrm{i}\\-2\\2-\mathrm{i}\end{pmatrix}\right]=(\overline{1+\mathrm{i}},2,\overline{2-\mathrm{i}})\begin{pmatrix}\mathrm{i}\\-2\\2-\mathrm{i}\end{pmatrix}$$

$$=(1-\mathrm{i})\mathrm{i}+2(-2)+(2+\mathrm{i})(2-\mathrm{i})=2+\mathrm{i}。$$

性质 (1) $[\boldsymbol{\alpha},0]=[0,\boldsymbol{\beta}]=0$，

(2) $[\boldsymbol{\beta},\boldsymbol{\alpha}]=\overline{[\boldsymbol{\alpha},\boldsymbol{\beta}]}$，

(3) $[\boldsymbol{\alpha},l_1\boldsymbol{\beta}_1+l_2\boldsymbol{\beta}_2]=l_1[\boldsymbol{\alpha},\boldsymbol{\beta}_1]+l_2[\boldsymbol{\alpha},\boldsymbol{\beta}_2]$，

$[k_1\boldsymbol{\alpha}_1+k_2\boldsymbol{\alpha}_2,\boldsymbol{\beta}]=\overline{k_1}[\boldsymbol{\alpha}_1,\boldsymbol{\beta}]+\overline{k_2}[\boldsymbol{\alpha}_2,\boldsymbol{\beta}]$。

(4) $[\boldsymbol{\alpha},\boldsymbol{\alpha}]$是非负实数，且$[\boldsymbol{\alpha},\boldsymbol{\alpha}]=0\Leftrightarrow\boldsymbol{\alpha}=0$。

上述性质可直接由矩阵运算与复数的性质得出。如$[\boldsymbol{\alpha},\boldsymbol{\alpha}]=\sum\limits_{i=1}^{n}\overline{a_i}a_i=\sum\limits_{i=1}^{n}|a_i|^2\geqslant0$，且$[\boldsymbol{\alpha},\boldsymbol{\alpha}]=0\Leftrightarrow a_i$的模$|a_i|=0,i=1,\cdots,n\Leftrightarrow a_i=0,i=1,\cdots,n\Leftrightarrow\boldsymbol{\alpha}=0$。

三、向量的长度

定义 5.9 称非负实数$\sqrt{[\boldsymbol{\alpha},\boldsymbol{\alpha}]}$为向量$\boldsymbol{\alpha}$的长度，记成$\|\boldsymbol{\alpha}\|$。

例如，设$\boldsymbol{\alpha}=\begin{pmatrix}2+\mathrm{i}\\3\\-2\mathrm{i}\end{pmatrix}$，则$\|\boldsymbol{\alpha}\|=\sqrt{5+9+4}=\sqrt{18}=2\sqrt{3}$。

由内积的性质，易得向量长度的性质。

性质 (1) $\|\boldsymbol{\alpha}\|\geqslant0$，且$\|\boldsymbol{\alpha}\|=0\Leftrightarrow\boldsymbol{\alpha}=0$；

(2) $\|k\boldsymbol{\alpha}\|=|k|\,\|\boldsymbol{\alpha}\|$。

定义 5.10 长度为 1 的向量称为单位向量。若$\boldsymbol{\alpha}\neq0$，用$\dfrac{1}{\|\boldsymbol{\alpha}\|}$乘非零向量$\boldsymbol{\alpha}$的运算称单位化运算。

例如，把向量$\begin{pmatrix}1\\-2\\2\end{pmatrix}$单位化，得单位向量$\begin{pmatrix}\frac{1}{3}\\-\frac{2}{3}\\\frac{2}{3}\end{pmatrix}$。

四、正交

定义 5.11 若 $[\alpha,\beta]=0$，则称向量 α 与 β 正交。显然，零向量与任何向量正交。

定义 5.12 两两正交的非零向量组称为正交组，由单位向量组成的正交组称为标准正交组。

例 5.14 $\begin{pmatrix} \frac{1}{3} \\ \frac{2}{3} \\ \frac{2}{3} \end{pmatrix}, \begin{pmatrix} +\frac{2}{3} \\ +\frac{1}{3} \\ -\frac{2}{3} \end{pmatrix}, \begin{pmatrix} +\frac{2}{3} \\ -\frac{2}{3} \\ +\frac{1}{3} \end{pmatrix}$ 是标准正交组。

例 5.15 $\varepsilon_1=\begin{pmatrix} 1 \\ 0 \\ \vdots \\ \vdots \\ 0 \end{pmatrix},\cdots,\varepsilon_n=\begin{pmatrix} 0 \\ \vdots \\ \vdots \\ 0 \\ 1 \end{pmatrix}$ 是标准正交组。

命题 5.11 向量组 $\alpha_1,\cdots,\alpha_s$ 是标准正交组 $\Leftrightarrow[\alpha_i,\alpha_j]=\delta_{ij},i,j=1,\cdots,s\Leftrightarrow([\alpha_i,\alpha_j])_{s\times s}=\boldsymbol{E}_s$。

命题 5.12 正交组单位化则成标准正交组。

命题 5.13 正交组必线性无关。

证 设 $\alpha_1,\cdots,\alpha_s$ 为正交组，若

$$k_1\alpha_1+\cdots+k_s\alpha_s=0,$$

则

$$0=[\alpha_i,0]=[\alpha_i,\sum_{j=1}^{s}k_j\alpha_j]=k_i[\alpha_i,\alpha_i],$$

因为 $\alpha_i\neq 0$，故 $[\alpha_i,\alpha_i]\neq 0$，必 $k_i=0,i=1,\cdots,s$。所以 $\alpha_1,\cdots,\alpha_s$ 线性无关。

五、正交化方法

设 $\alpha_1,\cdots,\alpha_s$ 为线性无关向量组，我们寻求与向量组 $\alpha_1,\cdots,\alpha_s$ 等价

的正交组(称这个过程为正交化过程)。

令 $\gamma_1=\alpha_1,\{\gamma_1\}\sim\{\alpha_1\}$。若正交组 $\gamma_1,\cdots,\gamma_t$ 已求得：

$$\{\gamma_1,\cdots,\gamma_t\}\sim\{\alpha_1,\cdots,\alpha_t\},$$

可设

$$\gamma_{t+1}=\alpha_{t+1}+k_1\gamma_1+\cdots+k_t\gamma_t,$$

由 $[\gamma_i,\gamma_{t+1}]=0,i=1,\cdots,t$，得

$$[\gamma_i,\alpha_{t+1}]+k_i[\gamma_i,\gamma_i]=0,$$

于是

$$k_i=\frac{-[\gamma_i,\alpha_{t+1}]}{[\gamma_i,\gamma_i]},\qquad i=1,2,\cdots,t。$$

反之设

$$\gamma_{t+1}=\alpha_{t+1}-\sum_{i=1}^{t}\frac{[\gamma_i,\alpha_{t+1}]}{[\gamma_i,\gamma_i]}\gamma_i,$$

则满足

$$[\gamma_i,\gamma_{t+1}]=0,\quad i=1,\cdots,t。$$

从而 $\gamma_1,\cdots,\gamma_t,\gamma_{t+1}$ 也是正交组。

$$\{\gamma_1,\cdots,\gamma_t,\gamma_{t+1}\}\sim\{\gamma_1,\cdots,\gamma_t,\alpha_{t+1}\}\sim\{\alpha_1,\cdots,\alpha_t,\alpha_{t+1}\}。$$

这样，我们得到了正交化的方法。

正交化方法 若 $\alpha_1,\cdots,\alpha_s$ 是线性无关的向量组，令 $\gamma_1=\alpha_1,\gamma_i=\alpha_i-\sum\limits_{j=1}^{i-1}\frac{[\gamma_j,\alpha_i]}{[\gamma_j,\gamma_j]}\gamma_j,i=2,\cdots,s$。则 $\{\gamma_1,\cdots,\gamma_t\}$ 是与 $\{\alpha_1,\cdots,\alpha_t\}$ 等价的正交组，$t=1,\cdots,s$。

注意 (1) 因 $\{\gamma_1,\cdots,\gamma_t\}\sim\{\alpha_1,\cdots,\alpha_t\}$，故两者的秩皆为 t，于是 $\gamma_1,\cdots,\gamma_t$ 线性无关，其中必无零向量。

(2) 若 $\alpha_1,\cdots,\alpha_s$ 线性相关，则 $\{\gamma_1,\cdots,\gamma_s\}$ 仍等价于 $\{\alpha_1,\cdots,\alpha_s\}$，且 $\gamma_1,\cdots,\gamma_s$ 两两正交，但其中有零向量。

若记 $\varepsilon_i=\frac{\gamma_i}{\|\gamma_i\|},i=1,\cdots,s$。则 $\{\varepsilon_1,\cdots,\varepsilon_t\}$ 是与 $\{\alpha_1,\cdots,\alpha_t\}$ 等价的标准正交组，$t=1,\cdots,s$。

命题 5.14 若 n 维向量 $\alpha_1,\cdots,\alpha_t(t<n)$ 是正交组(或标准正交组),则存在 n 维向量 $\alpha_{t+1},\cdots,\alpha_n$ 使 $\alpha_1,\cdots,\alpha_t,\cdots,\alpha_n$ 为正交组(或标准正交组)。

证 因为 P^n 的秩为 n,由命题 4.8 与命题 5.13,存在 n 维向量 $\beta_{t+1},\cdots,\beta_n$ 使 $\alpha_1,\cdots,\alpha_t,\beta_{t+1},\cdots,\beta_n$ 线性无关,令

$$\alpha_i=\beta_i-\sum_{j=1}^{i-1}\frac{[\alpha_j,\beta_i]}{[\alpha_j,\alpha_j]}\alpha_j.\quad i=t+1,\cdots,n。$$

则$\alpha_1,\cdots,\alpha_t,\cdots,\alpha_n$ 是正交组。

若 $\alpha_1,\cdots,\alpha_t$ 是标准正交组,令 $\gamma_i=\dfrac{\alpha_i}{\|\alpha_i\|}$, $i=t+1,\cdots,n$,则 $\alpha_1,\cdots,\alpha_t,\gamma_{t+1},\cdots,\gamma_n$ 是标准正交组。

例 5.16 求与向量组 $\alpha_1=\begin{pmatrix}0\\1\\1\end{pmatrix},\alpha_2=\begin{pmatrix}1\\1\\1\end{pmatrix},\alpha_3=\begin{pmatrix}1\\2\\1\end{pmatrix}$ 等价的标准正交组。

解 正交化,得

$$\gamma_1=\alpha_1=\begin{pmatrix}0\\1\\1\end{pmatrix},$$

$$\gamma_2=\alpha_2-\frac{[\gamma_1,\alpha_2]}{[\gamma_1,\gamma_1]}\gamma_1=\begin{pmatrix}1\\1\\1\end{pmatrix}-\begin{pmatrix}0\\1\\1\end{pmatrix}=\begin{pmatrix}1\\0\\0\end{pmatrix},$$

$$\gamma_3=\alpha_3-\frac{[\gamma_1,\alpha_3]}{[\gamma_1,\gamma_1]}\gamma_1-\frac{[\gamma_2,\alpha_3]}{[\gamma_2,\gamma_2]}\gamma_2$$

$$=\begin{pmatrix}1\\2\\1\end{pmatrix}-\frac{3}{2}\begin{pmatrix}0\\1\\1\end{pmatrix}-\frac{1}{1}\begin{pmatrix}1\\0\\0\end{pmatrix}=\begin{pmatrix}0\\\frac{1}{2}\\-\frac{1}{2}\end{pmatrix},$$

单位化,得

$$\boldsymbol{\varepsilon}_1=\frac{\sqrt{2}}{2}\begin{pmatrix}0\\1\\1\end{pmatrix},\boldsymbol{\varepsilon}_2=\begin{pmatrix}1\\0\\0\end{pmatrix},\boldsymbol{\varepsilon}_3=\frac{\sqrt{2}}{2}\begin{pmatrix}0\\1\\-1\end{pmatrix},$$

则 $\boldsymbol{\varepsilon}_1,\boldsymbol{\varepsilon}_2,\boldsymbol{\varepsilon}_3$ 即为所求标准正交组。

§6 酉 相 似

一、关联阵

定义 5.13 设 A 为 n 阶方阵，称$\overline{A}'$为 A 的关联阵，记成 $A^{\star}$。由矩阵的转置与共轭运算的性质，即得

性质 (1) $(\boldsymbol{A}^{\star})^{\star}=\boldsymbol{A}$;

(2) $(\boldsymbol{A}+\boldsymbol{B})^{\star}=\boldsymbol{A}^{\star}+\boldsymbol{B}^{\star}$;

(3) $(k\boldsymbol{A})^{\star}=\bar{k}\boldsymbol{A}^{\star}$;

(4) $(\boldsymbol{A}')^{\star}=(\boldsymbol{A}^{\star})'$;

(5) $(\boldsymbol{A}\boldsymbol{B})^{\star}=\boldsymbol{B}^{\star}\boldsymbol{A}^{\star}$;

(6) 若 $\boldsymbol{A}$ 可逆，则 $(\boldsymbol{A}^{-1})^{\star}=(\boldsymbol{A}^{\star})^{-1}$。

命题 5.15 设 A 为 n 阶方阵，则对任意 n 维(列)向量 $\boldsymbol{\alpha},\boldsymbol{\beta}$,均有

$$[A\boldsymbol{\alpha},\boldsymbol{\beta}]=[\boldsymbol{\alpha},A^{\star}\boldsymbol{\beta}].$$

证 因为 $[A\boldsymbol{\alpha},\boldsymbol{\beta}]=(A\boldsymbol{\alpha})^{\star}\boldsymbol{\beta}=\overline{\boldsymbol{\alpha}}'A^{\star}\boldsymbol{\beta}=[\boldsymbol{\alpha},A^{\star}\boldsymbol{\beta}]$。

二、酉阵

定义 5.14 若 n 阶方阵 A 满足 $AA^{\star}=A^{\star}A=\boldsymbol{E}$,即 $A^{-1}=A^{\star}$，则称 A 为酉阵。

性质 (1) 若 A 是酉阵，则 $\boldsymbol{A}',\overline{\boldsymbol{A}}$ 与 $\boldsymbol{A}^{-1}=\boldsymbol{A}^{\star}$ 都是酉阵；

(2) 若 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 是两个 n 阶酉阵，则 $\boldsymbol{AB}$ 也是酉阵。

证 (1) 因为 $(\boldsymbol{A}')^{-1}=(\boldsymbol{A}^{-1})'=(\boldsymbol{A}^{\star})'=(\boldsymbol{A}')^{\star}$,

$$(\overline{\boldsymbol{A}})^{-1}=(\overline{\boldsymbol{A}^{-1}})=\overline{\boldsymbol{A}^{\star}}=(\overline{\boldsymbol{A}})^{\star},$$
$$(\boldsymbol{A}^{\star})^{-1}=(\boldsymbol{A}^{-1})^{\star}=(\boldsymbol{A}^{\star})^{\star}。$$

(2) $(\boldsymbol{AB})^{-1} = \boldsymbol{B}^{-1}\boldsymbol{A}^{-1} = \boldsymbol{B}^{\star}\boldsymbol{A}^{\star} = (\boldsymbol{AB})^{\star}$。

命题 5.16 n 阶方阵 $\boldsymbol{A}$ 是酉阵$\Leftrightarrow\boldsymbol{A}$ 的列向量是标准正交组$\Leftrightarrow\boldsymbol{A}$ 的行向量是标准正交组。

证 $\boldsymbol{A}$ 是酉阵 $\Leftrightarrow\boldsymbol{A}^{-1} = \boldsymbol{A}^{\star}\Leftrightarrow\boldsymbol{A}^{\star}\boldsymbol{A} = \boldsymbol{E}\Leftrightarrow\boldsymbol{A}^{\star}\boldsymbol{A}\ (i,j) = \delta_{ij}$，即 $(\overline{\boldsymbol{A}^{(i)}})'\cdot\boldsymbol{A}^{(j)} = \delta_{ij}$，即$[\boldsymbol{A}^{(i)},\boldsymbol{A}^{(j)}] = \delta_{ij}$，即 $\boldsymbol{A}$ 的列向量为标准正交组。

又 $\boldsymbol{A}$ 是酉阵$\Leftrightarrow\boldsymbol{A}'$是酉阵$\Leftrightarrow\boldsymbol{A}'$的列向量为标准正交组$\Leftrightarrow\boldsymbol{A}$ 的行向量为标准正交组。

由命题 5.14 与 5.16 即得

命题 5.17 若 $\alpha_1,\cdots,\alpha_t$ 是 $\boldsymbol{P}^n$ 中的标准正交组 $(t < n)$，则存在 n 维向量 $\alpha_{t+1},\cdots,\alpha_n$，使 n 阶方阵$(\alpha_1,\cdots,\alpha_t,\cdots,\alpha_n)$为酉阵。

定义 5.15 实酉阵称为正交阵，即实矩阵 $A(\overline{A} = A)$ 为正交阵 $\Leftrightarrow\boldsymbol{A}^{-1} = \boldsymbol{A}'$。

三、酉相似

定义 5.16 设 $\boldsymbol{A},\boldsymbol{B}$ 为两个 n 阶方阵，若存在酉阵 $\boldsymbol{T}$ 使 $\boldsymbol{B} = \boldsymbol{T}^{-1}\boldsymbol{AT} = \boldsymbol{T}^{\star}\boldsymbol{AT}$，则称 A 与 $\boldsymbol{B}$ 酉相似。记成 $\boldsymbol{A}\overset{U}{\sim}\boldsymbol{B}$。

利用单位阵是酉阵、酉阵的逆、两个酉阵的积仍是酉阵，可得

性质 (1) $\boldsymbol{A}\overset{U}{\sim}\boldsymbol{A}$；

(2) 若 $\boldsymbol{A}\overset{U}{\sim}\boldsymbol{B}$，则 $\boldsymbol{B}\overset{U}{\sim}\boldsymbol{A}$；

(3) 若 $\boldsymbol{A}\overset{U}{\sim}\boldsymbol{B},\boldsymbol{B}\overset{U}{\sim}\boldsymbol{C}$，则 $\boldsymbol{A}\overset{U}{\sim}\boldsymbol{C}$。

定理 5.7 任意方阵 $\boldsymbol{A}$ 必酉相似于上三角阵。

证 对 $\boldsymbol{A}$ 的阶 n 用归纳法。

$n = 1$ 时，$\boldsymbol{A}$ 已是上三角阵。

若 $(n-1)$ 时结论成立。设 $\boldsymbol{A}$ 为一 n 阶方阵，取 $\boldsymbol{A}$ 的一个特征值 λ_1，设 α_1 是对应的单位特征向量，由命题 5.13，存在 n 维向量 $\alpha_2,\cdots,\alpha_n$ 使 $\boldsymbol{T} = (\alpha_1,\cdots,\alpha_n)$ 为酉阵，于是

$$\begin{aligned}\boldsymbol{AT} &= \boldsymbol{A}(\alpha_1,\cdots,\alpha_n) = (\boldsymbol{A}\alpha_1,\boldsymbol{A}\alpha_2,\cdots,\boldsymbol{A}\alpha_n)\\ &= (\lambda_1\alpha_1,\boldsymbol{A}\alpha_2,\cdots,\boldsymbol{A}\alpha_n)\end{aligned}$$

$$= (\alpha_1, \cdots, \alpha_n) \begin{pmatrix} \lambda_1 & b_{12}, \cdots b_{1n} \\ 0 & \\ \vdots & A_1 \\ 0 & \end{pmatrix}$$

$$= T \begin{pmatrix} \lambda_1 & b_{12} \cdots b_{1n} \\ 0 & \\ \vdots & A_1 \\ 0 & \end{pmatrix} = TB,$$

即

$$T^{-1}AT = B。$$

$$A \overset{U}{\sim} B。$$

A_1 为 $(n-1)$ 阶方阵，由归纳法假设，存在 $(n-1)$ 阶酉阵 V 使

$V^{-1}A_1V$ 为上三角阵，$\begin{pmatrix} \lambda_2 & & & * \\ & \ddots & & \\ & & \ddots & \\ & & & \ddots \\ 0 & & & \lambda_n \end{pmatrix}$。

令 $Q = \begin{pmatrix} 1 & \\ & V \end{pmatrix}$ 为 n 阶准对角阵，则

$$Q^{-1} = \begin{pmatrix} 1 & \\ & V^{-1} \end{pmatrix} = \begin{pmatrix} 1 & \\ & V^{\star} \end{pmatrix} = \begin{pmatrix} 1 & \\ & V \end{pmatrix}^{\star} = Q^{\star},$$

即 Q 为酉阵，且

$$Q^{-1}BQ = \begin{pmatrix} 1 & \\ & V^{\star} \end{pmatrix} \begin{pmatrix} \lambda_1 & b_{12} \cdots b_{1n} \\ 0 & \\ \vdots & A_1 \\ 0 & \end{pmatrix} \begin{pmatrix} 1 & \\ & V \end{pmatrix}$$

$$= \begin{pmatrix} \lambda_1 & b_{12} \cdots b_{1n} \\ 0 & \\ \vdots & V^{-1}A_1 \\ 0 & \end{pmatrix} \begin{pmatrix} 1 & \\ & V \end{pmatrix}$$

$$
=\begin{pmatrix} \lambda_1 & * \cdots\cdots * \\ 0 & \\ \vdots & V^{-1}A_1V \\ 0 & \end{pmatrix}=\begin{pmatrix} \lambda_1 & & & * \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \ddots & \\ 0 & & & & \lambda_n \end{pmatrix},
$$

即

$$
B \overset{U}{\sim} \text{上三角阵}\begin{pmatrix} \lambda_1 & & & * \\ & \ddots & & \\ & & \ddots & \\ & & & \ddots & \\ 0 & & & & \lambda_n \end{pmatrix},
$$

从而 $A \overset{U}{\sim}$ 上三角阵。

四、正规阵

定义 5.17 n 阶方阵 A 称为 Hermite 阵，如果 $A^\star = A$。

推论 实对称阵是 Hermite 阵。

定义 5.18 方阵 A 称为正规阵，如果 $A^\star A = AA^\star$。

推论 酉阵、Hermite 阵都是正规阵。从而正交阵、实对称阵都是正规阵。

引理 5.1 设 A 是正规阵，α 是 A 的对应于特征值 λ 的特征向量，则 α 也是 $A^\star$ 的对应于特征值 $\bar{\lambda}$ 的特征向量。

证 因为 A 是正规阵，$A^\star A = AA^\star$，从而

$$
\begin{aligned}
(A-\lambda E)^\star(A-\lambda E) &= (A^\star - \bar{\lambda}E)(A-\lambda E) \\
&= A^\star A - (\lambda+\bar{\lambda})A + \bar{\lambda}\lambda E \\
&= AA^\star - (\lambda+\bar{\lambda})A + \lambda\bar{\lambda}E \\
&= (A-\lambda E)(A-\lambda E)^\star.
\end{aligned}
$$

由于 $A\alpha = \lambda\alpha$，因而 $(A-\lambda E)\alpha = 0$，由命题 5.16，我们有

$$
[(A-\lambda E)^\star\alpha,(A-\lambda E)^\star\alpha] = [\alpha,(A-\lambda E)(A-\lambda E)^\star\alpha]
$$

$$= [\alpha,(A-\lambda E)^{\star}(A-\lambda E)\alpha]$$
$$= [(A-\lambda E)\alpha,(A-\lambda E)\alpha] = 0,$$

于是

$$A^{\star}\alpha - \bar{\lambda}\alpha = (A-\lambda E)^{\star}\alpha = 0,$$

即

$$A^{\star}\alpha = \bar{\lambda}\alpha。$$

引理得证。

定理 5.8 正规阵 A 的对应于不同特征值的特征向量是正交的。

证 设 λ_1、λ_2 是 A 的不同特征值，对应的特征向量分别是 α,β，则

$$A\alpha = \lambda_1\alpha, A^{\star}\beta = \overline{\lambda_2}\beta,$$

又

$$[A\alpha,\beta] = [\alpha,A^{\star}\beta], \text{（命题 5.16）}$$

即

$$[\lambda_1\alpha,\beta] = [\alpha,\overline{\lambda_2}\beta],$$

于是

$$\overline{\lambda_1}[\alpha,\beta] = \overline{\lambda_2}[\alpha,\beta]。$$

因 $\lambda_1 \neq \lambda_2, \overline{\lambda_1} \neq \overline{\lambda_2}$，故必 $[\alpha,\beta] = 0$。

定理 5.9 方阵 $A \overset{U}{\sim}$ 对角阵 $\Leftrightarrow A$ 是正规阵。

证 $\Rightarrow A = T\begin{pmatrix}\lambda_1 & & \\ & \ddots & \\ & & \lambda_n\end{pmatrix}T^{-1}$，$T$ 是酉阵。

于是

$$A^{\star} = (T^{-1})^{\star}\begin{pmatrix}\overline{\lambda_1} & & \\ & \ddots & \\ & & \overline{\lambda_n}\end{pmatrix}T^{\star} = T\begin{pmatrix}\overline{\lambda_1} & & \\ & \ddots & \\ & & \overline{\lambda_n}\end{pmatrix}T^{-1},$$

从而

$$AA^{\star} = T\begin{pmatrix}\lambda_1\overline{\lambda_1} & & \\ & \ddots & \\ & & \lambda_n\overline{\lambda_n}\end{pmatrix}T^{-1} = A^{\star}A。$$

$\Leftarrow$由定理 5.7,存在酉阵 $\boldsymbol{T}$ 使

$$\boldsymbol{T}^{-1}\boldsymbol{A}\boldsymbol{T}=\boldsymbol{B}=\begin{pmatrix} b_{11} & \cdots & \cdots & \cdots & b_{1n} \\ & \ddots & & & \vdots \\ & & \ddots & & \vdots \\ & 0 & & \ddots & \vdots \\ & & & & b_{nn} \end{pmatrix}。$$

由 $\boldsymbol{A}^{\star}\boldsymbol{A}=\boldsymbol{A}\boldsymbol{A}^{\star}$ 可得 $\boldsymbol{B}^{\star}\boldsymbol{B}=\boldsymbol{B}\boldsymbol{B}^{\star}$,即

$$\begin{pmatrix} \overline{b_{11}} & & & & 0 \\ \vdots & \ddots & & & \\ \vdots & & \ddots & & \\ \vdots & & & \ddots & \\ \overline{b_{1n}} & \cdots & \cdots & \cdots & \overline{b_{nn}} \end{pmatrix}\begin{pmatrix} b_{11} & \cdots & \cdots & \cdots & b_{1n} \\ & \ddots & & & \vdots \\ & & \ddots & & \vdots \\ & & & \ddots & \vdots \\ 0 & & & & b_{nn} \end{pmatrix}$$

$$=\begin{pmatrix} b_{11} & \cdots & \cdots & \cdots & b_{1n} \\ & \ddots & & & \vdots \\ & & \ddots & & \vdots \\ & & & \ddots & \vdots \\ 0 & & & & b_{nn} \end{pmatrix}\begin{pmatrix} \overline{b_{11}} & & & & 0 \\ \vdots & \ddots & & & \\ \vdots & & \ddots & & \\ \vdots & & & \ddots & \\ \overline{b_{1n}} & \cdots & \cdots & \cdots & \overline{b_{nn}} \end{pmatrix}。$$

依次比较两边对角线上元素,得

$$|b_{11}|^2=|b_{11}|^2+|b_{12}|^2+\cdots+|b_{1n}|^2,$$

从而

$$b_{12}=b_{13}=\cdots=b_{1n}=0.$$

$$|b_{22}|^2=|b_{22}|^2+|b_{23}|^2+\cdots+|b_{2n}|^2,$$

$$b_{23}=\cdots b_{2n}=0。$$

……

故有 $i<j$ 时,$b_{ij}=0$,即 $\boldsymbol{B}$ 为对角阵。$\boldsymbol{A}$ 酉相似于对角阵 $\boldsymbol{B}$。

推论 酉阵、Hermite 阵必酉相似于对角阵,

五、实对称阵

设 A 是实对称阵，由定理 5.9，存在酉阵 T 使

$$T^{-1}AT=\begin{pmatrix}\lambda_1 & & \\ & \ddots & \\ & & \lambda_n\end{pmatrix}.$$

下面我们证明实对称阵 A 的特征值 $\lambda_1,\cdots,\lambda_n$ 是实数，上式中的 T 是实数矩阵。

设 λ 是 A 的一个特征值，于是存在 $\alpha\neq 0$ 使

$$A\alpha=\lambda\alpha,$$

由命题 5.16

$$[A\alpha,\alpha]=[\alpha,A^{\star}\alpha]=[\alpha,A\alpha],$$

于是

$$[\lambda\alpha,\alpha]=[\alpha,\lambda\alpha],$$

即

$$\bar{\lambda}[\alpha,\alpha]=\lambda[\alpha,\alpha],$$

因为 $[\alpha,\alpha]\neq 0$，故必 $\bar{\lambda}=\lambda$，即 A 的特征值为实数。

实对称阵 A 的特征值全是实数，特征向量是实系数线性方程组 $(\lambda_iE-A)X=0$ 的非零解向量，因而可取实向量。这样，在酉相似于实对角阵时可取实酉阵，即正交阵作为过渡矩阵。

定理 5.10 实对称阵 A 必正交相似于实对角阵，即存在正交阵 T 使 $T'AT=T^{-1}AT$ 为实对角阵 $\begin{pmatrix}\lambda_1 & & \\ & \ddots & \\ & & \lambda_n\end{pmatrix}$。

正交阵 T 是由 A 的线性无关特征向量经正交化、标准化而得。由于实对称阵是正规阵，它的对应于不同特征值的特征向量彼此正交。于是在求特征向量组成的正交组时，仅须将同一特征值的特征向量正交化。

例 5.17 求正交阵 T 使 $T'AT = T'\begin{pmatrix} 5 & 4 & 4 \\ 4 & 5 & 4 \\ 4 & 4 & 5 \end{pmatrix}T$ 为对角阵。

解 由 $\begin{vmatrix} \lambda - 5 & -4 & -4 \\ -4 & \lambda - 5 & -4 \\ -4 & -4 & \lambda - 5 \end{vmatrix} = 0$

得特征值 $\lambda = 1, 13$:

(1) $\lambda = 1, 1E - A = \begin{pmatrix} -4 & -4 & -4 \\ -4 & -4 & -4 \\ -4 & -4 & -4 \end{pmatrix} \xrightarrow{\text{初等行变换}}$

$\begin{pmatrix} 1 & 1 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$,得线性无关特征向量 $\begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}$。

正交化,得

$$\begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ -2 \end{pmatrix},$$

单位化,得

$$\frac{\sqrt{2}}{2}\begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}, \frac{\sqrt{6}}{6}\begin{pmatrix} 1 \\ 1 \\ -2 \end{pmatrix}.$$

(2) $\lambda = 13, 13E - A = \begin{pmatrix} 8 & -4 & -4 \\ -4 & 8 & -4 \\ -4 & -4 & 8 \end{pmatrix} \xrightarrow{\text{行初等变换}}$

$\begin{pmatrix} 1 & 0 & -1 \\ 0 & 1 & -1 \\ 0 & 0 & 0 \end{pmatrix}$,得线性无关特征向量 $\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$,单位化,得

$$\frac{\sqrt{3}}{3}\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix},$$

令

$$T=\begin{pmatrix}\frac{\sqrt{2}}{2} & \frac{\sqrt{6}}{6} & \frac{\sqrt{3}}{3}\\ -\frac{\sqrt{2}}{2} & \frac{\sqrt{6}}{6} & \frac{\sqrt{3}}{3}\\ 0 & -\frac{\sqrt{6}}{3} & \frac{\sqrt{3}}{3}\end{pmatrix}$$

则

$$\boldsymbol{T}'\boldsymbol{A}\boldsymbol{T}=\boldsymbol{T}^{-1}\boldsymbol{A}\boldsymbol{T}=\begin{pmatrix}1 & & \\ & 1 & \\ & & 13\end{pmatrix}。$$

习　题　5

1. 求下列矩阵的特征值与特征向量：

(1) $\begin{pmatrix}5 & 6 & -3\\ -1 & 0 & 1\\ 1 & 2 & 1\end{pmatrix}$；　　(2) $\begin{pmatrix}0 & 0 & 1\\ 0 & 1 & 0\\ 1 & 0 & 0\end{pmatrix}$；

(3) $\begin{pmatrix}-2 & 1 & 0\\ -4 & 3 & 0\\ 4 & -1 & 2\end{pmatrix}$；　　(4) $\begin{pmatrix}1 & 1 & 1 & 1\\ 1 & 1 & -1 & -1\\ 1 & -1 & 1 & -1\\ 1 & -1 & -1 & 1\end{pmatrix}$。

2. 若 n 阶方阵 $\boldsymbol{A}$ 的 i 行、j 列元素为 $i\cdot j$ $(n\geqslant 2)$，求 $\boldsymbol{A}$ 的所有特征值。

3. 若 $\boldsymbol{A}$ 为 n 阶可逆阵，$\boldsymbol{B}$ 为 n 阶方阵，求证：

$$\boldsymbol{AB}\sim\boldsymbol{BA}。$$

4. 若 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 为同阶方阵，求证：$\boldsymbol{AB}$ 与 $\boldsymbol{BA}$ 有相同的特征多项式。

5. 求下列矩阵的 n 次幂：

(1) $\begin{pmatrix}1 & 4 & 2\\ 0 & -3 & 4\\ 0 & 4 & 3\end{pmatrix}$；　　(2) $\begin{pmatrix}0 & 2 & -1\\ -2 & 5 & -2\\ -4 & 8 & -3\end{pmatrix}$。

6. 若 $\boldsymbol{A}$ 为 n 阶方阵，且 $\boldsymbol{A}^k = 0, k \in \mathbf{Z}^+$，求证：$\boldsymbol{A}$ 的特征值为 0。

7. 求证：0 不是可逆阵的特征值。

8. 若

$$\boldsymbol{A} = \begin{pmatrix} a & b & c & d \\ d & a & b & c \\ c & d & a & b \\ b & c & d & a \end{pmatrix},$$

求 $\boldsymbol{A}$ 的特征值。

9. 若 $\lambda_1, \cdots, \lambda_n$ 是 n 阶方阵 $\boldsymbol{A}$ 的全部特征值，求证：

$$\sum_{i=1}^{n} \lambda_i^2 = \sum_{i=1}^{n} \sum_{j=1}^{n} \boldsymbol{A}(ij)\boldsymbol{A}(ji)。$$

10. 若 $\boldsymbol{A}, \boldsymbol{B}$ 均为 n 阶方阵，$\boldsymbol{P} = \begin{pmatrix} \boldsymbol{A} & \boldsymbol{B} \\ \boldsymbol{B} & \boldsymbol{A} \end{pmatrix}$，求证：$\boldsymbol{P}$ 的特征值集为 $(\boldsymbol{A} + \boldsymbol{B})$ 与 $(\boldsymbol{A} - \boldsymbol{B})$ 特征值集的并集。

11. 若 λ_1, λ_2 是方阵 $\boldsymbol{A}$ 的两个不同特征值，对应的特征向量分别是 α_1 与 α_2，求证：$\alpha_1 + \alpha_2$ 不是 $\boldsymbol{A}$ 的特征向量。

12. 若 $\boldsymbol{A}$ 为对角线上元素两两不等的上三角阵，则 $\boldsymbol{A}$ 相似于对角阵。

13. 若 2 阶实方阵 $\boldsymbol{A}$ 的行列式值 $|\boldsymbol{A}| < 0$，求证：$\boldsymbol{A}$ 相似于对角阵。

14. 求下列矩阵的特征多项式与最小多项式：

(1) $\begin{pmatrix} 2 & & & \\ & 2 & & \\ & & 2 & \\ & & & 2 \end{pmatrix}$； (2) $\begin{pmatrix} 2 & 0 & & & \\ 1 & 2 & & & \\ & & 4 & 2 & \\ & & 3 & 5 & \\ & & & & 7 \end{pmatrix}$；

(3) $\begin{pmatrix} 3 & 0 & 0 & & & & \\ 1 & 3 & 0 & & & & \\ 0 & 1 & 3 & & & & \\ & & & 5 & 2 & & \\ & & & -4 & -1 & & \\ & & & & & 1 & 0 \\ & & & & & 1 & 1 \end{pmatrix}$；(4) $\begin{pmatrix} 0 & 2 & -1 \\ -2 & 5 & -2 \\ -4 & 8 & -3 \end{pmatrix}$。

15. 求下列矩阵的 100 次幂：

(1) $\begin{pmatrix} 4 & -2 \\ 3 & -1 \end{pmatrix}$；　　(2) $\begin{pmatrix} -1 & 1 & 0 \\ -4 & 3 & 0 \\ 1 & 0 & 2 \end{pmatrix}$；

(3) $\begin{pmatrix} -2 & 1 & 0 \\ -4 & 3 & 0 \\ 4 & -1 & 2 \end{pmatrix}$；(4) $\begin{pmatrix} 2 & -1 & -1 \\ 2 & -1 & -2 \\ -1 & 1 & 2 \end{pmatrix}$。

16. $n \geqslant 2$. 若 n 阶方阵 $\boldsymbol{A}$ 的元素全为 1，求 $\boldsymbol{A}$ 的特征值与最小多项式。

17. 若 $x^m - 1(m \in \mathbf{Z}^+)$ 是 $\boldsymbol{A}$ 的一个化零多项式，求证：$\boldsymbol{A}$ 相似于对角阵。

18. 若方阵 $\boldsymbol{A}$ 满足 $\boldsymbol{E} + \boldsymbol{A} + \dfrac{\boldsymbol{A}^2}{2!} + \cdots + \dfrac{\boldsymbol{A}^t}{t!} = 0$，则 $\boldsymbol{A}$ 相似于对角阵。

19. 若方阵 $\boldsymbol{A}$ 满足 $\boldsymbol{A}^2 = \boldsymbol{A}$，且 $\boldsymbol{A}$ 的秩为 r，则 $|\boldsymbol{A} + \boldsymbol{E}| = 2^r$。

20. 若 n 阶方阵 $\boldsymbol{A}$ 有 n 个不同特征值，且 $\boldsymbol{AB} = \boldsymbol{BA}$，求证：$\boldsymbol{B}$ 相似于对角阵。

21. 若 n 阶方阵 $\boldsymbol{A}$ 相似于对角阵，则 $2n$ 阶方阵 $\begin{pmatrix} \boldsymbol{A} & \boldsymbol{A}^2 \\ \boldsymbol{A}^2 & \boldsymbol{A} \end{pmatrix}$ 也相似于对角阵。

22. 设 $\boldsymbol{\alpha} = \begin{pmatrix} 2 \\ -\mathrm{i} \\ 3+\mathrm{i} \\ 4-\mathrm{i} \end{pmatrix}$，$\boldsymbol{\beta} = \begin{pmatrix} \mathrm{i} \\ 1 \\ 1-\mathrm{i} \\ 4+\mathrm{i} \end{pmatrix}$，求$[\boldsymbol{\alpha}, \boldsymbol{\beta}]$，$\|\boldsymbol{\alpha}\|$、$\|\boldsymbol{\beta}\|$。

23. 求单位向量 α 使与

$$\begin{pmatrix}1\\1\\-1\\1\end{pmatrix},\begin{pmatrix}1\\-1\\-1\\1\end{pmatrix},\begin{pmatrix}2\\1\\1\\3\end{pmatrix}$$

正交。

24. 求与下列向量组等价的正交组：

(1) $\begin{pmatrix}2\\3\\-4\\-6\end{pmatrix},\begin{pmatrix}1\\8\\-2\\-16\end{pmatrix},\begin{pmatrix}2\\4\\10\\-29\end{pmatrix},\begin{pmatrix}3\\11\\4\\-7\end{pmatrix}$；

(2) $\begin{pmatrix}1\\1\\-1\\-2\end{pmatrix},\begin{pmatrix}-2\\1\\5\\11\end{pmatrix},\begin{pmatrix}0\\3\\3\\7\end{pmatrix},\begin{pmatrix}3\\-3\\-3\\-9\end{pmatrix}$。

25. 若向量 $\boldsymbol{\beta}$ 是向量 $\boldsymbol{\alpha}_1,\cdots,\boldsymbol{\alpha}_s$ 的线性组合，且 $[\boldsymbol{\beta},\boldsymbol{\alpha}_i]=0,i=1,\cdots,s$，求证：$\boldsymbol{\beta}=0$。

26. 求证：

(1) 实反对称阵的特征值是零或纯虚数；

(2) 酉阵的特征值的模为 1。

27. 求正交阵 $\boldsymbol{T}$ 使 $\boldsymbol{T}'\boldsymbol{A}\boldsymbol{T}$ 为对角阵：

(1) $\begin{pmatrix}1&-2&0\\-2&2&-2\\0&-2&3\end{pmatrix}$； (2) $\begin{pmatrix}1&-2&2\\-2&-2&4\\2&4&-2\end{pmatrix}$；

(3) $\begin{pmatrix}a&b&b\\b&a&b\\b&b&a\end{pmatrix}$；$(a\neq b)$； (4) $\begin{pmatrix}1&1&1&1\\1&1&1&1\\1&1&1&1\\1&1&1&1\end{pmatrix}$。

28. 设 $\boldsymbol{\alpha}_1,\cdots,\boldsymbol{\alpha}_s$ 为 s 个 n 维向量，$s\leqslant n$，记 $\boldsymbol{A}=([\boldsymbol{\alpha}_i,\boldsymbol{\alpha}_j])_{s\times s}$，求证：$\boldsymbol{\alpha}_1,\cdots,\boldsymbol{\alpha}_s$ 线性无关 $\Leftrightarrow|\boldsymbol{A}|\neq 0$。

29. 求证：可逆阵 $\boldsymbol{A}$ 必可表示为酉阵 $\boldsymbol{Q}$ 与上三角阵 $\boldsymbol{T}$ 之积。

30. 求证：正交阵 $\boldsymbol{A}$ 必正交相似于实准对角阵 $\begin{pmatrix} \boldsymbol{A}_1 & & \\ & \ddots & \\ & & \boldsymbol{A}_s \end{pmatrix}$，这里 $\boldsymbol{A}_i$ 或是一阶方阵或是形如 $\begin{pmatrix} \cos\theta_i, & -\sin\theta_i \\ \sin\theta_i, & \cos\theta_i \end{pmatrix}^2$ 的二阶实方阵（$i = 1,\cdots,s$）。

第六章　二　次　型

在平面解析几何中，为了便于研究二次曲线

$$ax^2 + bxy + cy^2 = 1$$

的几何性质，我们选择适当的（坐标）旋转变换

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

把方程化为标准形。

$ax^2 + bxy + cy^2$ 可看成一个二次齐次多项式。从代数学的观点看，化标准形意味着通过变量的线性替换化简二次齐次多项式，使它仅含平方项。

§1　二次型的矩阵形式

一、二次型与它的矩阵

定义 6.1　设 P 是数域。一个系数在 P 中的字母 $x_1, x_2, \cdots, x_n$ 的二次齐次多项式

$$f(x_1, \cdots, x_n) = \sum_{i=1}^{n} a_{ii}x_i^2 + 2\sum_{1 \leqslant i < j \leqslant n} a_{ij}x_ix_j,$$

称为数域 P 上的 n 元二次型，简称二次型。

例如，$x_1^2 + 3x_2^2 - 7x_3^2 + 4x_1x_2 - 6x_1x_3 + 12x_2x_3$ 是有理数域 $\mathbf{Q}$ 上的一个三元二次型。

若 $i > j$ 时，令 $a_{ij} = a_{ji}$，注意到 $x_ix_j = x_jx_i$，$f(x_1, \cdots, x_n)$ 可写成

$\sum_{i=1}^{n}\sum_{j=1}^{n}a_{ij}x_ix_j$。

若记 $\boldsymbol{A}=(a_{ij})_{n\times n}$，称它为二次型 f 的矩阵，由于 $a_{ij}=a_{ji}$，所以 $\boldsymbol{A}'=\boldsymbol{A}$，即二次型的矩阵 $\boldsymbol{A}$ 为对称阵。

令 $\boldsymbol{X}=\begin{pmatrix}x_1\\ \vdots\\ x_n\end{pmatrix}$，则

$$\boldsymbol{X}'\boldsymbol{A}\boldsymbol{X}=(x_1,\cdots,x_n)(a_{ij})_{n\times n}\begin{pmatrix}x_1\\ \vdots\\ x_n\end{pmatrix}$$

$$=(x_1,\cdots,x_n)\begin{pmatrix}\sum_{j=1}^{n}a_{1j}x_j\\ \vdots\\ \sum_{j=1}^{n}a_{nj}x_j\end{pmatrix}=\sum_{i=1}^{n}\sum_{j=1}^{n}a_{ij}x_ix_j,$$

即二次型可用矩阵的乘积来表示：

$$f(x_1,\cdots,x_n)=\boldsymbol{X}'\boldsymbol{A}\boldsymbol{X}。$$

二次型 $\boldsymbol{X}'\boldsymbol{A}\boldsymbol{X}$ 的矩阵 $\boldsymbol{A}$ 的元素，当 $i\neq j$ 时，$a_{ij}=a_{ji}$ 是它的 x_ix_j 项的系数的一半，而 a_{ii}是它的 x_i^2 项的系数，因此二次型的矩阵是惟一的，即若二次型

$$f(x_1,\cdots,x_n)=\boldsymbol{X}'\boldsymbol{A}\boldsymbol{X}=\boldsymbol{X}'\boldsymbol{B}\boldsymbol{X},$$

且 $\boldsymbol{A}'=\boldsymbol{A},\boldsymbol{B}'=\boldsymbol{B}$，则 $\boldsymbol{A}=\boldsymbol{B}$。

例如，$3x_1^2-4x_1x_2+6x_1x_3+x_2^2+8x_2x_3+7x_3^2$

$$=(x_1,x_2,x_3)\begin{pmatrix}3&-2&3\\-2&1&4\\3&4&7\end{pmatrix}\begin{pmatrix}x_1\\x_2\\x_3\end{pmatrix}。$$

二、二次型的满秩线性替换

定义 6.2 设 $x_1,\cdots,x_n$ 与 $y_1,\cdots,y_n$ 是两组字母，系数在数域 P 中的一组关系式

$$\begin{cases} x_1 = t_{11}y_1 + \cdots + t_{1n}y_n \\ \vdots \qquad\qquad \vdots \\ \vdots \qquad\qquad \vdots \\ x_n = t_{n}y_1 + \cdots + t_{nn}y_n \end{cases}$$

即 $\boldsymbol{X} = \boldsymbol{TY}$。这里

$$\boldsymbol{X} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, \boldsymbol{T} = (t_{ij})_{n\times n}, \boldsymbol{Y} = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$$

称为由 $x_1,\cdots,x_n$ 到 $y_1,\cdots,y_n$ 的一个线性替换，简称线性替换。如果系数行列式 $|\boldsymbol{T}| \neq 0$，则称为满秩线性替换。此时矩阵 $\boldsymbol{T}$ 可逆，$\boldsymbol{Y} = \boldsymbol{T}^{-1}\boldsymbol{X}$；$x_1,\cdots,x_n$ 是 $y_1,\cdots,y_n$ 的一次齐式，$y_1,\cdots,y_n$ 也是 $x_1,\cdots,x_n$ 的一次齐式。

若把 $\boldsymbol{X} = \boldsymbol{TY}$ 代入 $f = \boldsymbol{X}'\boldsymbol{AX}$，得到 $y_1,\cdots,y_n$ 的二次型 $f = (\boldsymbol{TY})'\boldsymbol{ATY} = \boldsymbol{Y}'(\boldsymbol{T}'\boldsymbol{AT})\boldsymbol{Y}$，即线性替换 $\boldsymbol{X} = \boldsymbol{TY}$，把 $\boldsymbol{X}$ 的二次型变成了 $\boldsymbol{Y}$ 的二次型，由于$(\boldsymbol{T}'\boldsymbol{AT})' = \boldsymbol{T}'\boldsymbol{A}'(\boldsymbol{T}')' = \boldsymbol{T}'\boldsymbol{AT}$，即 $\boldsymbol{B} = \boldsymbol{T}'\boldsymbol{AT}$ 是对称阵，故 $\boldsymbol{B}$ 是 $\boldsymbol{Y}$ 的二次型 f 的矩阵。

$\boldsymbol{B} = \boldsymbol{T}'\boldsymbol{AT}$ 是二次型线性替换前后的矩阵关系。

例如，二次型

$$f = (x_1,x_2,x_3)\begin{pmatrix} 1 & 1 & 2 \\ 1 & 2 & -3 \\ 2 & -3 & 6 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$$

在满秩线性替换

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 1 & -1 & -7 \\ 0 & 1 & 5 \\ 0 & 0 & 1 \end{pmatrix}\begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix}$$

下，化为 y_1, y_2, y_3 的二次型

$$f=(y_1,y_2,y_3)\begin{pmatrix}1 & & \\ & 1 & \\ & & -23\end{pmatrix}\begin{pmatrix}y_1\\ y_2\\ y_3\end{pmatrix},$$

这里

$$\begin{pmatrix}1 & & \\ & 1 & \\ & & -23\end{pmatrix}=\begin{pmatrix}1 & -1 & -7\\ 0 & 1 & 5\\ 0 & 0 & 1\end{pmatrix}'\begin{pmatrix}1 & 1 & 2\\ 1 & 2 & -3\\ 2 & -3 & 6\end{pmatrix}\begin{pmatrix}1 & -1 & -7\\ 0 & 1 & 5\\ 0 & 0 & 1\end{pmatrix}。$$

定义 6.3 数域 P 上两个 n 阶方阵 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 称为合同的，如果存在 P 上 n 阶可逆阵 $\boldsymbol{T}$ 使 $\boldsymbol{B}=\boldsymbol{T}'\boldsymbol{A}\boldsymbol{T}$。

推论 二次型在满秩线性替换前后的矩阵是合同的。

性质 (1) $\boldsymbol{A}$ 与 $\boldsymbol{A}$ 合同；

(2) 若 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 合同，则 $\boldsymbol{B}$ 与 $\boldsymbol{A}$ 合同；

(3) 若 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 合同，$\boldsymbol{B}$ 与 $\boldsymbol{C}$ 合同，则 $\boldsymbol{A}$ 与 $\boldsymbol{C}$ 合同。

证 (1) 因为 $\boldsymbol{A}=\boldsymbol{E}'\boldsymbol{A}\boldsymbol{E}$，$|\boldsymbol{E}|=1\neq 0$。

(2) $\boldsymbol{B}=\boldsymbol{T}'\boldsymbol{A}\boldsymbol{T}$，则 $\boldsymbol{A}=(\boldsymbol{T}^{-1})'\boldsymbol{B}\boldsymbol{T}^{-1}$。

(3) $\boldsymbol{B}=\boldsymbol{T}_1'\boldsymbol{A}\boldsymbol{T}_1$，$\boldsymbol{C}=\boldsymbol{T}_2'\boldsymbol{A}\boldsymbol{T}_2$，$\boldsymbol{T}_1$，$\boldsymbol{T}_2$ 可逆，则

$$\boldsymbol{C}=\boldsymbol{T}_2'\boldsymbol{T}_1'\boldsymbol{A}\boldsymbol{T}_1\boldsymbol{T}_2=(\boldsymbol{T}_1\boldsymbol{T}_2)'\boldsymbol{A}(\boldsymbol{T}_1\boldsymbol{T}_2),\boldsymbol{T}=\boldsymbol{T}_1\boldsymbol{T}_2$$

是可逆阵的积，仍可逆。

定义 6.4 称二次型矩阵的秩为二次型的秩。

引理 6.1 合同矩阵有相同的秩。

证 若 $\boldsymbol{A}$ 合同于 $\boldsymbol{B}$，即存在可逆阵 $\boldsymbol{T}$ 使 $\boldsymbol{B}=\boldsymbol{T}'\boldsymbol{A}\boldsymbol{T}$，于是 $\mathrm{r}(\boldsymbol{B})\leqslant\mathrm{r}(\boldsymbol{A})$。由性质(2)，$\boldsymbol{B}$ 合同于 $\boldsymbol{A}$，故 $\mathrm{r}(\boldsymbol{A})\leqslant\mathrm{r}(\boldsymbol{B})$。这样 $\mathrm{r}(\boldsymbol{A})=\mathrm{r}(\boldsymbol{B})$。

由于二次型在满秩线性替换前后的矩阵是合同的，故其秩相等。

命题 6.1 满秩线性替换不改变二次型的秩。

§2 二次型的标准形

定义 6.5 矩阵为对角阵的二次型，即 $\sum_{i=1}^{n} a_i x_i^2$，称为标准二次型。

本节中，我们指出：任何二次型都可经满秩线性替换化为标准二次型，并提供具体方法。

例 6.1 用满秩线性替换化二次型

$$f = x_1^2 + 2x_2^2 + 6x_3^2 + 2x_1x_2 + 4x_1x_3 - 6x_2x_3$$

为标准形。

解 f 中含 x_1 的平方项，把 f 中含 x_1 的项归并在一起，配方得

$$\begin{aligned} f &= (x_1^2 + 2x_1x_2 + 4x_1x_3) + 2x_2^2 + 6x_3^2 - 6x_2x_3 \\ &= (x_1 + x_2 + 2x_3)^2 - x_2^2 - 4x_3^2 - 4x_2x_3 \\ &\quad + 2x_2^2 + 6x_3^2 - 6x_2x_3 \\ &= (x_1 + x_2 + 2x_3)^2 + x_2^2 + 2x_3^2 - 10x_2x_3。 \end{aligned}$$

继续把余下的项（已经没有含 x_1 的项）中含 x_2 的项归并在一起，配方得

$$f = (x_1 + x_2 + 2x_3)^2 + (x_2 - 5x_3)^2 - 23x_3^2。$$

作满秩线性替换

$$\begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} = \begin{pmatrix} 1 & 1 & 2 \\ 0 & 1 & -5 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix},$$

则 f 为 y_1, y_2, y_3 的标准二次型

$$f = y_1^2 + y_2^2 - 23y_3^2。$$

例 6.2 用满秩线性替换化二次型

$$f = 2x_1x_2 + 2x_1x_3 - 6x_2x_3$$

为标准二次型。

解 f 不含平方项，因而不能用配方法。

作满秩线性替换：

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 1 & 1 & 0 \\ 1 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix},$$

从而

$$\begin{aligned} f &= 2(y_1^2 - y_2^2) + 2y_3[(y_1 + y_2) - 3(y_1 - y_2)] \\ &= 2y_1^2 - 2y_2^2 - 4y_1y_3 + 8y_2y_3, \end{aligned}$$

此时，f 含平方项，配方得

$$f = 2(y_1 - y_3)^2 - 2(y_2 - 2y_3)^2 + 6y_3^2,$$

令

$$\begin{pmatrix} z_1 \\ z_2 \\ z_3 \end{pmatrix} = \begin{pmatrix} 1 & -1 & 0 \\ 0 & 1 & -2 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix},$$

$$\text{或} \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} z_1 \\ z_2 \\ z_3 \end{pmatrix}.$$

则 $f = 2z_1^2 - 2z_2^2 + 6z_3^2$。

从而 f 经过满秩线性替换：

$$\begin{aligned} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} &= \begin{pmatrix} 1 & 1 & 0 \\ 1 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} \\ &= \begin{pmatrix} 1 & 1 & 0 \\ 1 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} z_1 \\ z_2 \\ z_3 \end{pmatrix} \\ &= \begin{pmatrix} 1 & 1 & 2 \\ 1 & -1 & -1 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} z_1 \\ z_2 \\ z_3 \end{pmatrix} \end{aligned}$$

化为标准二次型 $f = 2z_1^2 - 2z_2^2 - 6z_3^2$。

注 由于可逆阵的积仍是可逆阵，故对二次型连续作多次满秩线性替换，相当于对二次型作一次满秩线性替换。

定理 6.1 二次型可经满秩线性替换化为标准形。

证 设二次型 $f = \boldsymbol{X}'\boldsymbol{AX}, \boldsymbol{A} = (a_{ij})_{k\times k}$。

我们对二次型 f 的元数 k 用归纳法。

$k = 1$ 时，$f = a_{11}x_1^2$ 已是标准形了。

若 $k = (n-1)$ 时结论成立。$f = \boldsymbol{X}'\boldsymbol{AX}$ 为 n 元二次型，分三种情形：

(1) $a_{11}, a_{22}, \cdots, a_{nn}$ 中至少有一个不为 0，如 $a_{11} \neq 0$。此时

$$f = a_{11}x_1^2 + 2a_{12}x_1x_2 + \cdots + 2a_{1n}x_1x_n + f_1, f_1 = \sum_{i=2}^{n}\sum_{j=2}^{n} a_{ij}x_ix_j$$

是 $x_2, \cdots, x_n$ 的 $(n-1)$ 元二次型。配方得

$$f = a_{11}\left(x_1 + \sum_{i=2}^{n} \frac{a_{1i}}{a_{11}}x_i\right)^2 + f_2。$$

$f_2 = f_1 - a_{11}\left(\sum_{i=2}^{n} \frac{a_{1i}}{a_{11}}x_i\right)^2$ 是 $x_2, \cdots, x_n$ 的 $(n-1)$ 元二次型。

由归纳法假设，存在满秩线性替换：

$$\begin{pmatrix} y_2 \\ \vdots \\ y_n \end{pmatrix} = \boldsymbol{T}_1 \begin{pmatrix} x_2 \\ \vdots \\ x_n \end{pmatrix},$$

使

$$f_2 = d_2y_2^2 + \cdots + d_ny_n^2。$$

作满秩线性替换：

$$\boldsymbol{Y} = \begin{pmatrix} y_1 \\ \vdots \\ \vdots \\ y_n \end{pmatrix} = \left(\begin{array}{c|c} 1 & \dfrac{a_{12}}{a_{11}} \cdots \dfrac{a_{1n}}{a_{11}} \\ \hline 0 & \\ \vdots & \boldsymbol{T}_1 \\ \vdots & \\ 0 & \end{array}\right) \begin{pmatrix} x_1 \\ \vdots \\ \vdots \\ x_n \end{pmatrix} = \boldsymbol{TX}。$$

（由于 $|T| = |T_1| \neq 0$，故是满秩线性替换。）

此时，$f = a_{11}y_1^2 + d_2y_2^2 + \cdots + d_ny_n^2$，已是标准形了。

(2) $a_{11}, a_{22}, \cdots, a_{nn}$全为0.但有一个 $a_{1j} \neq 0$ $(j > 1)$,可设 $a_{12} \neq 0$,作满秩线性替换:

$$\begin{cases} x_1 = y_1 + y_2, \\ x_2 = y_1 - y_2, \\ x_3 = y_3, \\ \vdots \quad\quad \vdots \\ x_n = y_n, \end{cases}$$

此时

$$\begin{aligned} f &= 2a_{11}(y_1 + y_2)(y_1 - y_2) + \cdots \\ &= 2a_{12}y_1^2 - 2a_{12}y_2^2 + \cdots, \end{aligned}$$

y_1^2 的系数不为0,化归为(1)的情形。

(3) 若 $a_{1i} = 0$, $(i = 1, \cdots, n)$。此时,f 是 $x_2, \cdots, x_n$ 的 $(n-1)$ 元二次型,由归纳法假设,结论成立。

由(1)、(2)、(3)可知,n 时结论成立。由归纳法,本定理得证。

推论 对称阵必合同于对角阵。即若 $\boldsymbol{A}' = \boldsymbol{A}$，则存在可逆阵 $\boldsymbol{T}$ 使 $\boldsymbol{T}'\boldsymbol{A}\boldsymbol{T}$ 为对角阵。

证 构造二次型 $f = \boldsymbol{X}'\boldsymbol{A}\boldsymbol{X}$，由定理6.1,存在满秩线性替换 $\boldsymbol{X} = \boldsymbol{T}\boldsymbol{Y}$ 使

$$f = \boldsymbol{Y}' \begin{pmatrix} d_1 & & \\ & \ddots & \\ & & d_n \end{pmatrix} \boldsymbol{Y},$$

此时

$$\boldsymbol{T}'\boldsymbol{A}\boldsymbol{T} = \begin{pmatrix} d_1 & & \\ & \ddots & \\ & & d_n \end{pmatrix}。$$

§3 实二次型

本节中，我们研究实数域上的二次型，即实二次型。

一、规范形

实二次型 $f = X'AX$ 可经满秩线性替换化为标准形：

$$f = d_1y_1^2 + \cdots + d_ny_n^2,$$

适当调整字母的次序（即作满秩线性替换）可使

$$d_i = \begin{cases} 正， & 0 \leqslant i \leqslant p \\ 负， & p < i \leqslant r \\ 0， & r < i \leqslant n \end{cases}$$

这里 $1 \leqslant p \leqslant r \leqslant n$，$r$ 是二次型 f 的秩。

进一步作满秩线性替换：

$$\begin{aligned} z_i &= \sqrt{d_i}\,y_i, & 1 \leqslant i \leqslant p \\ z_i &= \sqrt{-d_i}\,y_i, & p < i \leqslant r \\ z_i &= y_i, & r < i \leqslant n \end{aligned}$$

此时，$f = z_1^2 + \cdots + z_p^2 - z_{p+1}^2 - \cdots - z_r^2$，称它为 f 的规范形。r 是二次型 f 的秩，是一个从属于 f 的不变量。下面指出，正平方项的个数 p 也是 f 的一个不变量。

定理 6.2（惯性定理） 若 $y_1^2 + \cdots + y_p^2 - y_{p+1}^2 - \cdots - y_r^2$ 与 $z_1^2 + \cdots + z_q^2 - z_{q+1}^2 - \cdots - z_r^2$ 都是 n 元实二次型 $f = X'AX$ 的规范形，则 $p = q$。

证 若 $p \neq q$，不妨设 $q < p$。

由假设 $X = T_1Y = T_2Z$，从而 $Z = TY$，这里 $T = T_2^{-1}T_1$ 是两个可逆阵的积，也是可逆阵。设 $T = (t_{ij})_{n\times n}$，于是 $z_1, \cdots, z_n$ 是 $y_1, \cdots, y_n$ 的一次齐式。

n 个未知量 $y_1,\cdots,y_n$ 的齐次线性方程组

$$\begin{cases} z_1 = t_{11}y_1 + \cdots + t_{1n}y_n = 0 \\ \vdots \\ z_q = t_{q1}y_1 + \cdots + t_{qn}y_n = 0 \\ y_{p+1} = 0 \\ \vdots \qquad \vdots \\ y_r \quad = 0 \\ \vdots \qquad \vdots \\ y_n \quad = 0 \end{cases}$$

的式子个数为 $q + (n - p) = n - (p - q) <$ 未知量个数 n，故有非零解 $Y = \begin{pmatrix} c_1 \\ \vdots \\ c_n \end{pmatrix}$，$\begin{pmatrix} c_{p+1} = \cdots = c_n = 0 \\ c_1,\cdots,c_p \text{ 不全为 } 0 \end{pmatrix}$。

用 $\begin{pmatrix} c_1 \\ \vdots \\ c_n \end{pmatrix}$ 代 Y，得 $Z = T\begin{pmatrix} c_1 \\ \vdots \\ c_n \end{pmatrix}$，分别用 $Y = \begin{pmatrix} c_1 \\ \vdots \\ c_n \end{pmatrix}$，$Z = T\begin{pmatrix} c_1 \\ \vdots \\ c_n \end{pmatrix}$ 代入 f 得

$$f = c_1^2 + \cdots + c_p^2 > 0, \quad (\text{因为 } c_1,\cdots,c_p \text{ 不全为 } 0)$$

$$f = 0^2 + \cdots + 0^2 + 0^2 - \left(\sum_{i=1}^{n} t_{q+1,i}c_i\right)^2 - \cdots - \left(\sum_{i=1}^{n} t_{ri}c_i\right)^2 \leqslant 0,$$

这是不可能的。故必 $p = q$。

定义 6.6 实二次型 f 的规范形中正平方项的个数称为 f 的正惯性指数，负平方项的个数称为 f 的负惯性指数，两者的差称为符号差。

由惯性定理可知，实二次型的秩与正、负惯性指数，在满秩线性替换下保持不变，是从属于 f 的不变量。

二、正交替换

设 A 是实二次型 $f = X'AX$ 的矩阵，则 A 是实对称阵，由定理 5.10，存在正交阵 T 使

$$T'AT = T^{-1}AT = \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix},$$

这里 $\lambda_1,\cdots,\lambda_n$ 是 A 的特征值。

当 T 为正交阵时的满秩线性替换 $X = TY$ 叫正交替换。从而在满秩线性替换(称正交替换) $X = TY$ 下:

$$f = Y'\begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix}Y = \sum_{i=1}^{n}\lambda_i y_i^2,$$

即为标准形。于是得下述定理:

定理 6.3 实二次型 $f = X'AX$ 可经正交替换 $X = TY$ 化为标准形 $\sum_{i=1}^{n}\lambda_i y_i^2$,这里 $\lambda_1,\cdots,\lambda_n$ 是 A 的特征值。

例 6.3 用正交替换化实二次型 $f = 2x_1x_2 + 2x_1x_3 - 2x_1x_4 - 2x_2x_3 + 2x_2x_4 + 2x_3x_4$ 为标准形。

解 f 的矩阵为

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

特征多项式 $|\lambda E - A| = (\lambda + 3)(\lambda - 1)^3$,得特征值 $\lambda = -3,1$。

$\lambda = -3$ 时:

$$(-3E - A) = \begin{pmatrix} -3 & -1 & -1 & 1 \\ -1 & -3 & 1 & -1 \\ -1 & 1 & -3 & -1 \\ 1 & -1 & -1 & -3 \end{pmatrix} \xrightarrow{\frac{(1)+(2)+(3)+(4)}{-4}}$$

$$
\begin{bmatrix} 1 & 1 & 1 & 1 \\ -1 & -3 & 1 & -1 \\ -1 & 1 & -3 & -1 \\ 1 & -1 & -1 & -3 \end{bmatrix}
\xrightarrow{\substack{(2)+(1)\\(3)+(1)\\(4)-(1)}}
\begin{bmatrix} 1 & 1 & 1 & 1 \\ 0 & -2 & 2 & 0 \\ 0 & 2 & -2 & 0 \\ 0 & -2 & -2 & -4 \end{bmatrix}
\xrightarrow{\substack{(3)+(2)\\(4)-(2)\\-\frac{1}{2}(2)}}
$$

$$
\begin{bmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & -4 & -4 \end{bmatrix}
\xrightarrow{\substack{(1)-(2)\\ -\frac{(4)}{4}\leftrightarrow(3)}}
\begin{bmatrix} 1 & 0 & 2 & 1 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix}
$$

$$
\xrightarrow{\substack{(2)+(3)\\(1)-2(3)}}
\begin{bmatrix} 1 & 0 & 0 & -1 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix}
=
\left[\begin{array}{ccc|c} & \boldsymbol{E}_3 & & \begin{matrix}-1\\1\\1\end{matrix} \\ \hline 0 & 0 & 0 & 0 \end{array}\right],
$$

得其基础解系：$\begin{bmatrix} -1 \\ 1 \\ 1 \\ -1 \end{bmatrix}$，单位化，得 $\frac{1}{2}\begin{bmatrix} -1 \\ 1 \\ 1 \\ -1 \end{bmatrix}$.

$\lambda = 1$ 时：

$$
1\boldsymbol{E} - \boldsymbol{A} = \begin{bmatrix} -1 & 1 & 1 & -1 \\ 1 & -1 & -1 & 1 \\ 1 & -1 & -1 & 1 \\ -1 & 1 & 1 & -1 \end{bmatrix}
\xrightarrow{\substack{(2)+(1)\\(3)+(1)\\(4)-(1)}}
$$

$$
\begin{bmatrix} -1 & 1 & 1 & -1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}
\xrightarrow{-(1)}
\left[\begin{array}{c|ccc} 1 & -1 & -1 & 1 \\ \hline 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right],
$$

得其基础解系：$\begin{bmatrix} -1 \\ -1 \\ 0 \\ 0 \end{bmatrix}$，$\begin{bmatrix} -1 \\ 0 \\ -1 \\ 0 \end{bmatrix}$，$\begin{bmatrix} 1 \\ 0 \\ 0 \\ -1 \end{bmatrix}$ 或 $\begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \end{bmatrix}$，$\begin{bmatrix} 1 \\ 0 \\ 1 \\ 0 \end{bmatrix}$，$\begin{bmatrix} 1 \\ 0 \\ 0 \\ -1 \end{bmatrix}$，

正交化，得 $\begin{pmatrix}1\\1\\0\\0\end{pmatrix}$，$\begin{pmatrix}\frac{1}{2}\\-\frac{1}{2}\\1\\0\end{pmatrix}$，$\begin{pmatrix}\frac{1}{3}\\-\frac{1}{3}\\-\frac{1}{3}\\-1\end{pmatrix}$，

单位化，得 $\frac{\sqrt{2}}{2}\begin{pmatrix}1\\1\\0\\0\end{pmatrix}$，$\frac{\sqrt{6}}{6}\begin{pmatrix}1\\-1\\2\\0\end{pmatrix}$，$\frac{\sqrt{3}}{6}\begin{pmatrix}-1\\1\\1\\3\end{pmatrix}$。

于是经正交替换：$\begin{pmatrix}x_1\\x_2\\x_3\\x_4\end{pmatrix}=\begin{pmatrix}-\frac{1}{2} & \frac{\sqrt{2}}{2} & \frac{\sqrt{6}}{6} & -\frac{\sqrt{3}}{6}\\ \frac{1}{2} & \frac{\sqrt{2}}{2} & -\frac{\sqrt{6}}{6} & \frac{\sqrt{3}}{6}\\ \frac{1}{2} & 0 & \frac{\sqrt{6}}{3} & \frac{\sqrt{3}}{6}\\ -\frac{1}{2} & 0 & 0 & \frac{\sqrt{3}}{2}\end{pmatrix}\begin{pmatrix}y_1\\y_2\\y_3\\y_4\end{pmatrix}$，

f 化为标准形：

$$f=-3y_1^2+y_2^2+y_3^2+y_4^2。$$

三、正定二次型

定义 6.7 实二次型 $f(x_1,\cdots,x_n)$称为正定的，如果对任意一组不全为零的实数 $c_1,\cdots,c_n$，恒有 $f(c_1,\cdots,c_n)>0$。

推论 实二次型 $f(x_1,\cdots,x_n)$是正定的$\Leftrightarrow$对任意一组实数 $c_1,\cdots,c_n$，有 $f(c_1,\cdots,c_n)\geqslant 0$，且 $f(c_1,\cdots,c_n)=0$ 时必

$$c_1=c_2=\cdots=c_n=0。$$

例 6.4 n 元实二次型

$f(x_1,\cdots,x_n)=x_1^2+(x_1+x_2)^2+\cdots+(x_{n-1}+x_n)^2+x_n^2$

是正定的。

证 对任意实数 $c_1,\cdots,c_n$，$f(c_1,\cdots,c_n)\geqslant 0$，且 $f(c_1,\cdots,c_n)=0$，必

$$c_1=c_1+c_2=c_2+c_3=\cdots=c_{n-1}+c_n=c_n=0,$$

即

$$c_1=c_2=\cdots=c_n=0,$$

故 f 是正定的。

例 6.5 n 元实二次型

$$f(x_1,\cdots,x_n)=\sum_{i=1}^{n}x_i^2+\sum_{1\leqslant i<j\leqslant n}x_ix_j=\frac{1}{2}\sum_{i=1}^{n}x_i^2+\frac{1}{2}\left(\sum_{i=1}^{n}x_i\right)^2$$

是正定的。

命题 6.2 n 元实二次型 $f(\boldsymbol{X})=\boldsymbol{X}'\boldsymbol{A}\boldsymbol{X}$ 是正定的$\Leftrightarrow f$ 的正惯性指数 p 等于元数 n。

证 $\Rightarrow$若 $p<n$，f 的规范形为

$$f=y_1^2+\cdots+y_p^2-y_{p+1}^2-\cdots,$$

$$\boldsymbol{X}=\boldsymbol{T}\boldsymbol{Y},\ 令\ \boldsymbol{Y}_0=\begin{pmatrix}0\\ \vdots\\ 0\\ 1\\ \vdots\\ 1\end{pmatrix}\begin{matrix}\left.\begin{matrix}\\ \\ \\ \end{matrix}\right\}p\ 个\\ \\ \\ \\ \end{matrix},$$

因为 $p<n$，故 $\boldsymbol{Y}_0\neq 0$，此时 $\boldsymbol{X}_0=\boldsymbol{T}\boldsymbol{Y}_0$ 也不等于 0(否则 $\boldsymbol{Y}_0=\boldsymbol{T}^{-1}\boldsymbol{X}_0=0$)，但 $f\leqslant 0$，与 f 正定的假设矛盾，所以$p=n$。

$\Leftarrow$设 f 的正惯性指数等于 n，则 f 的规范形为

$$f=y_1^2+\cdots+y_n^2.$$

因而用实数代 $y_1,\cdots,y_n$ 时 $f\geqslant 0$，且 $f=0$ 时，$\boldsymbol{Y}=0$，从而 $\boldsymbol{X}=\boldsymbol{T}\boldsymbol{Y}=0$，即 f 为正定二次型。

推论 1 实二次型 f 是正定的$\Leftrightarrow f$ 的规范形的矩阵为单位阵$\Leftrightarrow f$ 的矩阵相合于 $\boldsymbol{E}$。

推论 2 满秩线性替换不改变二次型的正定性。

证 因为满秩线性替换不改变二次型的正惯性指数，从而替换前后的二次型其正惯性指数或同时等于元数或同时小于元数，所以替换前后的二次型或同时是正定的或同时不是正定的。

n 元实二次型可经正交替换化为标准形 $\lambda_1 y_1^2+\cdots+\lambda_n y_n^2$，这里 $\lambda_1,\cdots,\lambda_n$ 是二次型矩阵的全部特征值。而 $\lambda_1 y_1^2+\cdots+\lambda_n y_n^2$ 作为 $y_1,\cdots,y_n$ 的二次型是正定的 $\Leftrightarrow \lambda_i>0, i=1,\cdots,n$。由推论 2 即得下述命题：

命题 6.3 实二次型 f 是正定的$\Leftrightarrow f$ 的矩阵的特征值全大于零。

引理 6.2 若 $f=\boldsymbol{X}'\boldsymbol{A}\boldsymbol{X}$ 是正定的，则 f 的矩阵 $\boldsymbol{A}$ 的行列式 $|\boldsymbol{A}|>0$。

证 由推论 1，$\boldsymbol{A}$ 合同于 $\boldsymbol{E}$，即存在可逆阵 $\boldsymbol{T}$ 使

$$\boldsymbol{A}=\boldsymbol{T}'\boldsymbol{E}\boldsymbol{T}=\boldsymbol{T}'\boldsymbol{T},$$

从而

$$|\boldsymbol{A}|=|\boldsymbol{T}|^2>0。$$

定义 6.8 实对称矩阵 $\boldsymbol{A}$ 称为正定的，如果实二次型 $\boldsymbol{X}'\boldsymbol{A}\boldsymbol{X}$ 是正定二次型。

由命题 2 的推论 1 与命题 3 即得下述命题：

命题 6.4 实对称矩阵 $\boldsymbol{A}$ 是正定的$\Leftrightarrow\boldsymbol{A}$ 合同于 $\boldsymbol{E}$（或一已知正定矩阵）$\Leftrightarrow\boldsymbol{A}$ 的特征值全大于 0。

命题 6.5 n 元实二次型 $f(x_1,\cdots,x_n)=\boldsymbol{X}'\boldsymbol{A}\boldsymbol{X}$ 是正定的 $\Leftrightarrow f$ 的矩阵 $\boldsymbol{A}=(a_{ij})_{n\times n}$ 的 n 个顺序主子式 a_{11}，$\begin{vmatrix} a_{11} & a_{12}\\ a_{21} & a_{22}\end{vmatrix}$，$\cdots$，$\begin{vmatrix} a_{11} & \cdots & a_{1k}\\ \vdots & & \vdots\\ a_{k1} & \cdots & a_{kk}\end{vmatrix}$，$\cdots$，$|\boldsymbol{A}|$ 全大于零。

证 $\Rightarrow$ 令 $g(x_1,\cdots,x_k)=f(x_1,\cdots,x_k,0,\cdots,0)$，$k\leqslant n$，则

$$g(x_1,\cdots,x_k)=(x_1,\cdots,x_k)\begin{pmatrix} a_{11}, & \cdots & a_{1k}\\ \vdots & & \vdots\\ a_{k1}, & \cdots & a_{kk}\end{pmatrix}\begin{pmatrix} x_1\\ \vdots\\ x_k\end{pmatrix}$$ 即 g 的矩阵为 $(a_{ij})_{k\times k}$。

因为 f 正定，故对任意一组不全为 0 的实数 $c_1,\cdots,c_k$ 有

$$g(c_1,\cdots,c_k)=f(c_1,\cdots,c_k,0,\cdots 0)>0,$$

故 g 为 k 元正定二次型。由引理 6.2 有

$$\begin{vmatrix} a_{11} & \cdots & a_{1k} \\ \vdots & & \vdots \\ a_{k1} & \cdots & a_{kk} \end{vmatrix} > 0, k=1,\cdots,n。$$

$\Leftarrow$ 对元数 n 用归纳法证矩阵 $\boldsymbol{A}$ 是正定的(从而 f 是正定的)。$n=1$ 时，$f=a_{11}x_1^2$，因为 $a_{11}>0$，故 f 是正定的，于是 $\boldsymbol{A}$ 是正定的。若 $n-1$ 时结论成立。记

$$\boldsymbol{A}_1=(a_{ij})_{(n-1)\times(n-1)},$$

因为 $\boldsymbol{A}_1$ 的 $(n-1)$ 个顺序主子式也是 $\boldsymbol{A}$ 的顺序主子式，因此全大于 0，由归纳法假设，$(n-1)$ 阶方阵 $\boldsymbol{A}_1$ 是正定的。于是 $\boldsymbol{A}_1$ 合同于 $\boldsymbol{E}$，即存在可逆阵 $\boldsymbol{P}$ 使

$$\boldsymbol{P}'\boldsymbol{A}_1\boldsymbol{P}=\boldsymbol{E}。$$

若记 $\boldsymbol{\alpha}=\begin{pmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{n-1,n} \end{pmatrix}$，则

$$\boldsymbol{A}=\begin{pmatrix} \boldsymbol{A}_1 & \boldsymbol{\alpha} \\ \boldsymbol{\alpha}' & a_{nn} \end{pmatrix}。$$

令 $\boldsymbol{T}_1=\begin{pmatrix} \boldsymbol{P} & 0 \\ 0 & 1 \end{pmatrix}$，则

$$\begin{aligned} \boldsymbol{T}'_1\boldsymbol{A}\boldsymbol{T}_1 &= \begin{pmatrix} \boldsymbol{P}' & 0 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} \boldsymbol{A}_1 & \boldsymbol{\alpha} \\ \boldsymbol{\alpha}' & a_{nn} \end{pmatrix}\begin{pmatrix} \boldsymbol{P} & 0 \\ 0 & 1 \end{pmatrix} \\ &= \begin{pmatrix} \boldsymbol{P}'\boldsymbol{A}_1\boldsymbol{P} & \boldsymbol{P}'\boldsymbol{\alpha} \\ \boldsymbol{\alpha}'\boldsymbol{P} & a_{nn} \end{pmatrix}, \end{aligned}$$

若记 $\boldsymbol{P}'\boldsymbol{\alpha}=\boldsymbol{\beta}$，则

$$\boldsymbol{T}'_1\boldsymbol{A}\boldsymbol{T}_1=\begin{pmatrix} \boldsymbol{E} & \boldsymbol{\beta} \\ \boldsymbol{\beta}' & a_{nn} \end{pmatrix}。$$

令 $T_2 = \begin{pmatrix} E & -\beta \\ 0 & 1 \end{pmatrix}$，则

$$\begin{aligned}(T_1T_2)'A(T_1T_2) &= T_2'(T_1'AT_1)T_2 \\ &= \begin{pmatrix} E & 0 \\ -\beta' & 1 \end{pmatrix}\begin{pmatrix} E & \beta \\ \beta' & a_{nn} \end{pmatrix}\begin{pmatrix} E & -\beta \\ 0 & 1 \end{pmatrix} \\ &= \begin{pmatrix} E & 0 \\ 0 & a_{nn}-\beta'\beta \end{pmatrix}\end{aligned}$$

因为 $|T_1| = |P| \neq 0$，$|T_2| = 1 \neq 0$，故 T_1 与 T_2 都是可逆阵，从而 $T = T_1T_2$ 也是可逆阵。于是 A 合同于对角阵 $\begin{pmatrix} E & 0 \\ 0 & a_{nn}-\beta'\beta \end{pmatrix}$。

又

$$a_{nn}-\beta'\beta = \begin{vmatrix} E & 0 \\ 0 & a_{nn}-\beta'\beta \end{vmatrix} = |T'AT| = |T|^2 \cdot |A| > 0,$$

所以对角阵 $\begin{pmatrix} E & 0 \\ 0 & a_{nn}-\beta'\beta \end{pmatrix}$ 的对角线上元素全大于 0，故为正定阵，从而与之合同的矩阵 A 也正定。

由归纳法，充分性得证。

推论 实对称矩阵 A 是正定的$\Leftrightarrow A$ 的顺序主子式全大于 0。

例 6.6 判别二次型 $f = 4x_1^2 + 5x_2^2 - 7x_3^2 - 8x_1x_2 - 2x_1x_3 + 6x_2x_3$ 是否正定。

解 f 的矩阵为 $\begin{pmatrix} 4 & -4 & -1 \\ -4 & 5 & 3 \\ -1 & 3 & 7 \end{pmatrix}$，

因为 $4 > 0$，$\begin{vmatrix} 4 & -4 \\ -4 & 5 \end{vmatrix} = 4 > 0$，$\begin{vmatrix} 4 & -4 & -1 \\ -4 & 5 & 3 \\ -1 & 3 & 7 \end{vmatrix} = 11 > 0$，故 f 是正定的。

例 6.7 若 $a_i > b > 0\ (i = 1,\cdots,n)$，求证：$n$ 元实二次型 $f = X'\begin{pmatrix} a_1 & & b \\ & \ddots & \\ b & & a_n \end{pmatrix} X$ 是正定的。

证（第一种方法） $\begin{pmatrix} a_1 & & b \\ & \ddots & \\ b & & a_n \end{pmatrix}$ 的 k 阶顺序主子式

$$= \begin{vmatrix} a_1 & & & b \\ & \ddots & & \\ & & \ddots & \\ b & & & a_k \end{vmatrix} = \begin{vmatrix} b + (a_1 - b) & \cdots b + 0 \\ \vdots & \vdots \\ b + 0 & \cdots b + (a_n - b) \end{vmatrix}$$

$$= \begin{vmatrix} a_1 - b & & & \\ & \ddots & & \\ & & \ddots & \\ & & & a_k - b \end{vmatrix} + \sum_{i=1}^{k} \begin{array}{c} \begin{vmatrix} a_1 - b & b & & \\ \ddots & \vdots & & \\ & b & & \\ & \vdots & \ddots & \\ & b & & a_k - b \end{vmatrix} \\ i\text{列} \end{array}$$

$$= (a_1 - b)\cdots(a_k - b) + b\sum_{i=1}^{k} \frac{(a_1 - b)\cdots(a_k - b)}{(a_i - b)} > 0,\ k = 1,\cdots,n$$

故 f 是正定的，

证（第二种方法） $f = b\left(\sum_{i=1}^{n} x_i\right)^2 + \sum_{i=1}^{n}(a_i - b)x_i^2$，因为 $b > 0$，$a_i - b > 0$，$(i = 1,2,\cdots,n)$，故 f 为正定二次型。

四、半正定二次型

定义 6.9 n 元实二次型 $f(x_1,\cdots,x_n)$ 称为半正定的，如果对于任意一组实数 $c_1,\cdots,c_n$，恒有 $f(c_1,\cdots,c_n) \geqslant 0$。

定义 6.10 实对称阵称为半正定的，如果相应的实二次型 $X'AX$ 是半正定的。

例如，n 元实二次型 $(x_1+\cdots+x_n)^2$ 和 $\sum\limits_{i=1}^{n-1}(x_i-x_{i+1})^2$ 都是半正定二次型。

由惯性定理不难得到下面的命题：

命题 6.6 实二次型 $f=\boldsymbol{X}'\boldsymbol{AX}$ 是半正定的

$\Leftrightarrow$ f 的负惯性指数为 0（正惯性指数 $=f$ 的秩）。

$\Leftrightarrow$ f 的规范形为 $f=y_1^2+\cdots+y_r^2$。

$\Leftrightarrow$ f 的标准形为 $f=d_1z_1^2+\cdots+d_rz_r^2$, $d_i\geqslant 0$ $(i=1,\cdots,r)$。

$\Leftrightarrow$ $f(\boldsymbol{A})$的特征值全大于、等于 0。

推论 半正定矩阵 $\boldsymbol{A}$ 的行列式 $|\boldsymbol{A}|\geqslant 0$。

证 $\boldsymbol{A}$ 合同于 $f=\boldsymbol{X}'\boldsymbol{AX}$ 的规范形的矩阵 $\begin{pmatrix}\boldsymbol{E}_r & \boldsymbol{0}\\ \boldsymbol{0} & \boldsymbol{0}\end{pmatrix}$，即存在可逆阵 $\boldsymbol{T}$ 使

$$\boldsymbol{A}=\boldsymbol{T}'\begin{pmatrix}\boldsymbol{E}_r & \boldsymbol{0}\\ \boldsymbol{0} & \boldsymbol{0}\end{pmatrix}\boldsymbol{T},$$

两边取行列式，得

$$|\boldsymbol{A}|=|\boldsymbol{T}|^2\begin{vmatrix}\boldsymbol{E}_r & \boldsymbol{0}\\ \boldsymbol{0} & \boldsymbol{0}\end{vmatrix}\geqslant 0。$$

命题 6.7 n 元实二次型 $f=\boldsymbol{X}'\boldsymbol{AX}$ 是半正定的

$\Leftrightarrow$ $\boldsymbol{A}$ 的所有主子式：

$$\boldsymbol{D}_A\begin{pmatrix}i_1\cdots i_k\\ i_1\cdots i_k\end{pmatrix}=\begin{vmatrix}a_{i_1i_1} & a_{i_1i_2} & \cdots\cdots & a_{i_1i_k}\\ \vdots & \vdots & & \vdots\\ \vdots & \vdots & & \vdots\\ a_{i_ki_1} & a_{i_ki_2} & \cdots\cdots & a_{i_ki_k}\end{vmatrix},(1\leqslant i_1<i_2\cdots<i_k\leqslant n)$$

全大于、等于 0。

证 $\Rightarrow$若在 f 的表达式中令 $x_{i_t}=y_t(t=1,\cdots,k)$，其余的 $x_j=0$，得 k 元实二次型

$$g(y_1,\cdots,y_k)=f(0,\cdots,0,\underset{\substack{i_1\\ \text{位}}}{y_1},0,\cdots,0,\underset{\substack{i_2\\ \text{位}}}{y_2},0,\cdots,0,\cdots,0,\underset{\substack{i_k\\ \text{位}}}{y_k},0,\cdots,0)$$

$$=\sum_{s,t=1}^{n}a_{i_si_t}y_sy_t$$

$$=(y_1,\cdots,y_k)\begin{pmatrix}a_{i_1i_1} & \cdots\cdots & a_{i_1i_k}\\ \vdots & & \vdots\\ \vdots & & \vdots\\ a_{i_ki_1} & \cdots\cdots & a_{i_ki_k}\end{pmatrix}\begin{pmatrix}y_1\\ \vdots\\ \vdots\\ y_k\end{pmatrix},$$

因为 f 是半正定的，按定义 g 也是半正定的，故其矩阵的行列式 $D_A\begin{pmatrix}i_1\cdots i_k\\ i_1\cdots i_k\end{pmatrix}\geqslant 0$。

$$\Leftarrow |\lambda E-A|=\lambda^n-b_1\lambda^{n-1}+b_2\lambda^{n-2}-\cdots+(-1)^nb_n,$$

由多项式与行列式性质，得

$$b_k=A\text{ 的所有 }k\text{ 阶主子式之和}$$

$$=\sum_{1\leqslant i_1<i_2\cdots<i_k\leqslant n}D\begin{pmatrix}i_1\cdots i_k\\ i_1\cdots i_k\end{pmatrix}\geqslant 0。$$

若 A 有负的特征值 $-a,(a>0)$，代入特征多项式，得

$$(-a)^n-b_1(-a)^{n-1}+\cdots-(-1)^nb_n$$
$$=(-1)^n[a^n+b_1a^{n-1}+\cdots+b_n]=0,$$

即

$$a^n+b_1a^{n-1}+\cdots+b_n=0。$$

但因为 $a>0,b_k\geqslant 0\ (k=1,\cdots,n)$，上式不可能成立。故实对称阵 A 无负的特征值，即 A 的特征值是非负实数。于是由命题 6.6，$f=X'AX$ 为半正定二次型。

例 6.8 若 A 为 n 阶半正定矩阵，求证：

$$|A+E_n|\geqslant 1。$$

证 $\boldsymbol{A}$ 为半正定矩阵，其特征值 $\lambda_i \geqslant 0\ (i=1,\cdots,n)$。由定理6.3，存在正交阵 $\boldsymbol{T}$ 使

$$\boldsymbol{T}'\boldsymbol{A}\boldsymbol{T}=\boldsymbol{T}^{-1}\boldsymbol{A}\boldsymbol{T}=\begin{pmatrix}\lambda_1 & & \\ & \ddots & \\ & & \lambda_n\end{pmatrix},$$

或

$$\boldsymbol{A}=\boldsymbol{T}\begin{pmatrix}\lambda_1 & & \\ & \ddots & \\ & & \lambda_n\end{pmatrix}\boldsymbol{T}^{-1}。$$

于是

$$\begin{aligned}|\boldsymbol{A}+\boldsymbol{E}| &= \left|\boldsymbol{T}\begin{pmatrix}\lambda_1 & & \\ & \ddots & \\ & & \lambda_n\end{pmatrix}\boldsymbol{T}^{-1}+\boldsymbol{T}\begin{pmatrix}1 & & \\ & \ddots & \\ & & 1\end{pmatrix}\boldsymbol{T}^{-1}\right| \\ &= \left|\boldsymbol{T}\begin{pmatrix}\lambda_1+1 & & \\ & \ddots & \\ & & \lambda_n+1\end{pmatrix}\boldsymbol{T}^{-1}\right| \\ &= |\boldsymbol{T}|\begin{vmatrix}\lambda_1+1 & & \\ & \ddots & \\ & & \lambda_n+1\end{vmatrix}|\boldsymbol{T}^{-1}| \\ &= \begin{vmatrix}\lambda_1+1 & & \\ & \ddots & \\ & & \lambda_n+1\end{vmatrix}=(\lambda_1+1)\cdots(\lambda_n+1)\geqslant 1。\end{aligned}$$

习　题　6

1. 写出下列二次型的矩阵：

   (1) $4x_1^2+5x_2^2-7x_3^2+6x_1x_2+x_2x_3-x_1x_3$；

   (2) $x_1x_2-x_3x_4$。

2. 用满秩线性替换化下列二次型为标准形：

(1) $x_1^2-2x_2^2+5x_3^2-2x_1x_2+4x_1x_3-6x_2x_3$；

(2) $2x_1x_2-2x_1x_3-4x_2x_3$；

(3) $x_1^2+x_2^2+x_3^2+x_4^2+2x_1x_2+2x_2x_3+2x_3x_4$；

(4) $\sum\limits_{i=1}^{n}x_ix_{2n-i+1}$；

(5) $\sum\limits_{i=1}^{n}x_i^2+\sum\limits_{1\leqslant i<j\leqslant n}x_ix_j$。

3. 求证：秩为 r 的对称阵可表示为 r 个秩为 1 的对称阵之和。

4. 求证：$\begin{pmatrix}1&&&&\\&2&&&\\&&\ddots&&\\&&&\ddots&\\&&&&n\end{pmatrix}$ 与 $\begin{pmatrix}i_1&&&\\&\ddots&&\\&&\ddots&\\&&&i_n\end{pmatrix}$ 合同。这里 $i_1,\cdots,i_n$ 是 $1,\cdots,n$ 的一个排列。

5. 若 $\boldsymbol{A}$ 为 n 阶对称阵，且对任意 n 维向量 α 都有 $\alpha'A\alpha=0$，求证：$\boldsymbol{A}$ 为零矩阵。

6. 设 n 元实二次型 $f=\sum\limits_{i=1}^{s}\left(\sum\limits_{j=1}^{n}a_{ij}x_j\right)^2$，求证：$f$ 的秩等于矩阵 $\boldsymbol{A}=(a_{ij})_{s\times n}$ 的秩。

7. 如果把实 n 阶对称阵按合同分类，即两个 n 阶实对称阵属于同一类，当且仅当它们合同，问共有多少类？

8. 若 n 元实二次型 $f(x_1,\cdots,x_n)=l_1^2+\cdots+l_p^2-l_{p+1}^2-\cdots-l_{p+q}^2$，这里 l_i 是 $x_1,\cdots,x_n$ 的一次齐式($i=1,\cdots,p+q$)。求证：f 的正惯性指数 $\leqslant p$，f 的负惯性指数 $\leqslant q$。

9. 用正交替换化下列实二次型为标准形：

(1) $\boldsymbol{X}'\begin{pmatrix}7&3&3\\3&7&3\\3&3&7\end{pmatrix}\boldsymbol{X}$；

(2) $\boldsymbol{X}'\begin{pmatrix}0&1&-2\\1&0&-1\\-2&-1&0\end{pmatrix}\boldsymbol{X}$。

10. 判别下列实二次型是否正定：

(1) $2x_1^2+6x_2^2+4x_3^2-4x_1x_2+8x_1x_3-6x_2x_3$；

(2) $\sum\limits_{i=1}^{n}x_i^2+\sum\limits_{1\leqslant i<j\leqslant n}x_ix_j$；

(3) $\sum\limits_{i=1}^{n}x_i^2+\sum\limits_{i=1}^{n-1}x_ix_{i+1}$；

(4) $(n+1)\sum\limits_{i=1}^{n}x_i^2-\left(\sum\limits_{i=1}^{n}x_i\right)^2$。

11. 求 t 的值，使下列实二次型为正定：

(1) $3x^2+2txy+5y^2$；

(2) $5x_1^2+3x_2^2+tx_3^2+6x_1x_2-8x_1x_3+4x_2x_3$。

12. 若实 $s\times n$ 阵 $\boldsymbol{A}$ 的秩为 n，则实二次型 $\boldsymbol{X}'\boldsymbol{A}'\boldsymbol{A}\boldsymbol{X}$ 是正定的。

13. 若 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 为同阶正定阵，则 $\boldsymbol{A}+\boldsymbol{B}$ 也是正定阵。

14. 若 $\boldsymbol{A}$ 是正定阵，则 $\boldsymbol{A}^{-1}$ 也是正定阵。

15. (1) 设 $\boldsymbol{A}$ 为实对称阵，求证：存在正实数 a，使得对于任意实向量 $\boldsymbol{\alpha}$，恒有

$$|\boldsymbol{\alpha}'\boldsymbol{A}\boldsymbol{\alpha}|\leqslant a\boldsymbol{\alpha}'\boldsymbol{\alpha};$$

(2) 设 $\boldsymbol{A}$ 为实对称阵，求证：存在实数 $m>0,\varepsilon>0$，使得 $m\boldsymbol{E}+\boldsymbol{A}$ 与 $\boldsymbol{E}+\varepsilon\boldsymbol{A}$ 为正定矩阵。

16. 若 $\boldsymbol{A}$ 是 n 阶正定阵，求证：$f=-\begin{vmatrix} & & x_1\\ & \boldsymbol{A} & \vdots\\ & & x_n\\ x_1 & \cdots\ x_n & 0\end{vmatrix}$ 是正定二次型。

17. 若 $\boldsymbol{A}$ 为 n 阶正定阵，$\boldsymbol{B}$ 为 n 阶实对称阵，求证：存在可逆阵 $\boldsymbol{T}$ 使 $\boldsymbol{T}'\boldsymbol{A}\boldsymbol{T}$ 与 $\boldsymbol{T}'\boldsymbol{B}\boldsymbol{T}$ 同时为对角阵。

18. 若 $\boldsymbol{A},\boldsymbol{B}$ 为同阶正定阵，且 $\boldsymbol{AB}=\boldsymbol{BA}$，求证：$\boldsymbol{AB}$ 也是正定阵。

19. 若 $\boldsymbol{A},\boldsymbol{B}$ 为同阶正定阵，且多项式 $|x\boldsymbol{A}-\boldsymbol{B}|$ 的根全为 1，求证：

$$\boldsymbol{A}=\boldsymbol{B}。$$

20. 若 $\boldsymbol{A}$ 为正定矩阵，$\boldsymbol{B}$ 为半正定矩阵，求证：

$$|\boldsymbol{A}+\boldsymbol{B}| \geqslant |\boldsymbol{B}|。$$

21. 若 $\boldsymbol{A}$ 为正定矩阵，$\boldsymbol{B}$ 为反对称矩阵，求证：

$$|\boldsymbol{A}+\boldsymbol{B}| > 0。$$

22. 若矩阵方程

$$\boldsymbol{AX}+\boldsymbol{XA}=\boldsymbol{C}$$

有惟一解 $\boldsymbol{B}$，这里 $\boldsymbol{A}$，$\boldsymbol{C}$ 是同阶正定阵，求证：$\boldsymbol{B}$ 也是正定阵。

第七章　集合，映射，关系

本章介绍数学中的最基本概念“集合”、“映射”与“关系”。

§1　集　　合

一、集合

在数学中经常讨论的不是孤立的个体，也不是包罗万象的宇宙，而往往是对具有某些特征的个体的联合体进行研讨。这样，就产生了集合的概念。集合指“在一定范围内具有某些特性的个体组成的联合体”。组成一个集合的各个个体，称为这个集合的元素。

通常用大写英文字母 A，B，C…来表示集合，用小写英文字母 a，b，c…来表示集合的元素。当 a 是集合 A 的元素时，就说“a 属于 A”，或记为“$a \in A$”。当 a 不是集合 A 的元素时，记为“$a \notin A$”。

设有两个集合 A，B，若对一切 $a \in A$ 均有 $a \in B$，则称 A 是 B 的子集，记为 $A \subseteq B$。

若 A 是 B 的子集，B 也是 A 的子集，则称集合 A 和 B 相等，记为 $A = B$。

若 A 是 B 的子集，B 不是 A 的子集，则称 A 是 B 的真子集，记为 $A \subset B$。

为了方便，称不含任何元素的“客体”为空集合，记成$\varnothing$，并把空集合看成任何集合的子集。

数集是经常讨论的对象，为方便起见，继续用 $\mathbf{P}$，$\mathbf{Z}$，$\mathbf{Q}$，$\mathbf{R}$ 和 $\mathbf{C}$ 分别表示自然数、整数、有理数、实数和复数全体的集合。

二、集合的运算及其性质

设 A,B 是集合 U 的子集，定义 A 和 B 的交集为一切既属于 A 又属于 B 的元素组成的集合，记成 $A\cap B$；定义 A 和 B 的并集为一切属于 A 或者属于 B 的元素组成的集合，记成 $A\cup B$；A 和 B 的差集为一切属于 A 但不属于 B 的元素组成的集合，记成 $A-B$。即

$$A\cap B=\{x\mid x\in A \text{ 且 } x\in B\},$$
$$A\cup B=\{x\mid x\in A \text{ 或者 } x\in B\},$$
$$A-B=\{x\mid x\in A \text{ 且 } x\notin B\}。$$

集合的上述运算满足如下常用规律：

(1) 交换律　$A\cap B=B\cap A, A\cup B=B\cup A$;

(2) 结合律　$A\cap(B\cap C)=(A\cap B)\cap C$,
$A\cup(B\cup C)=(A\cup B)\cup C$;

(3) 分配律　$A\cap(B\cup C)=(A\cap B)\cup(A\cap C)$,
$A\cup(B\cap C)=(A\cup B)\cap(A\cup C)$;

(4) 对偶律　$A-(B\cup C)=(A-B)\cap(A-C)$,
$A-(B\cap C)=(A-B)\cup(A-C)$。

集合的交与并的概念可推广到任意多个集合上去，设

$$A_i(i\in I)$$

是集合 U 的子集，定义集合族 $\{A_i\mid i\in I\}$ 的交集为

$$\bigcap_{i\in I}A_i=\{x\mid x\in U,\forall i\in I, x\in A_i\},$$

这里 I 叫集合族 $\{A_i\}$ 的标集。

集合族 $\{A_i\mid i\in I\}$ 的并集为

$$\bigcup_{i\in I}A_i=\{x\mid x\in U, \text{存在一个 } i\in I \text{ 使 } x\in A_i\}。$$

特别当标集 $I=\varnothing$ 时，规定

$$\bigcap_{i\in I}A_i=U,\quad \bigcup_{i\in I}A_i=\varnothing。$$

对多个集合的交、并运算，分配律与对偶律依然成立：

$$A \cup (\bigcap_{i\in I} B_i) = \bigcap_{i\in I}(A \cup B_i),$$

$$A \cap (\bigcup_{i\in I} B_i) = \bigcup_{i\in I}(A \cap B_i),$$

$$A - (\bigcap_{i\in I} B_i) = \bigcup_{i\in I}(A - B_i),$$

$$A - (\bigcup_{i\in I} B_i) = \bigcap_{i\in I}(A - B_i)。$$

在此只证明最后一式，其余作为练习留给读者。

证　此题要证两点：(1) 左边集合$\subseteq$右边集合，即左边集合的元素必属于右边集合；(2) 右边集合$\subseteq$左边集合，即右边集合的元素必属于左边。

(1) $\forall x \in$ 左边集合，则 $x \in A$，且 $\forall i \in I, x \bar{\in} B_i$，即 $x \in A - B_i$，故 $x \in$ 右边集合。

(2) $\forall x \in$ 右边集合，$\forall i \in I, x \in A - B_i$，即 $x \in A, x \bar{\in} B_i$ $(\forall i \in I)$，则 $x \in A$ 且 $x \bar{\in} \bigcup_{i\in I} B_i$，故 $x \in$ 左边集合。

三、幂集合

定义 7.1　设 A 是给定的一个集合，则 A 的所有子集所组成的集合称为 A 的幂集合，用 2^A 来表示，即

$$2^A = \{B | B \subseteq A\}。$$

例如，$A = \{1,2,3\}$，则

$$2^A = \{\varnothing, \{1\}, \{2\}, \{3\}, \{1,2\}, \{1,3\}, \{2,3\}, \{1,2,3\}\}。$$

一般若 A 含有 n 个不同的元素，则 2^A 含有 2^n 个不同元素。理由是：2^A 的元素全是 A 的子集，而 A 的不含元素的子集有 $C_n^0 = 1$ 个($\varnothing$)，A 的含 1 个元素的子集有 C_n^1 个，含两个元素的子集有 C_n^2 个，…，含 n 个元素的子集有 $C_n^n = 1$ 个(A 本身)；故 2^A 的元素个数，即 A 的子集总数为

$$C_n^0 + C_n^1 + \cdots + C_n^n = (1+1)^n = 2^n。$$

§2 映　射

一、对应与映射

定义 7.2　设 A 与 B 为两个非空集合，f 为一规则：

(1) 若根据 f，A 中任意一个元素都能在 B 中找到至少一个元素与之对应，则称 f 确定了集合 A 到集合 B 的一个对应，记成 $A \xrightarrow{f} B$。

(2) 若根据 f，A 中任意一个元素 x 都有 B 中惟一确定的元素 y（记成 $f(x)$）与之对应，则称 f 确定了集合 A 到集合 B 的一个映射，记成 $A \mapsto\!\xrightarrow{f} B$（有时，在 A,B 都确定的情形下，也简记成 f）。

显然，映射是对应，但对应不一定是映射。当且仅当 A 中一个元素仅与 B 中一个元素对应时，对应才是映射。于是有下面推论：

推论　对应 $A \xrightarrow{f} B$ 是映射$\Leftrightarrow$，对任意 $x_1,x_2 \in A$，由 $x_1 = x_2$ 必有 $f(x_1) = f(x_2)$。

从映射的定义中可知，A、B、f 为映射的 3 个要素，分别称为映射 $A \mapsto\!\xrightarrow{f} B$ 的定义域、值域和对应规律。

两个映射称为相同，是指它们的 3 个要素对应相同。即映射“$A \mapsto\!\xrightarrow{f} B$”＝映射“$C \mapsto\!\xrightarrow{g} D$”意味着：

$$A = C, B = D,$$

且对任意 $x \in A = C$，恒有 $f(x) = g(x)$。

例 7.1　(1) $A = \mathbf{R}, B = \{x \mid x \in \mathbf{R}, x > 0\}, f(x) = |x|$，则 f 不是 A 到 B 的对应，更不是 A 到 B 的映射。因为根据 f，A 中的元素 0 在 B 中无对应元素。

(2) $A = B = P$，f：把 A 中的数对应于它的正因子。则 f 是 A 到 B 的对应，但不是 A 到 B 的映射。因为

$$f(12) = 1,2,3,4,6,12$$

不是 B 中惟一确定的元素。

(3) 设 $\mathbf{A} = M_n(\mathbf{R})$ 即实 n 阶方阵全体，$B = \mathbf{R}$，$f(x)$ 等于 x 的行列式，则 f 是 $\mathbf{A}$ 到 B 的映射。

(4) $A = B$，$I_A : x \longmapsto x$，是 A 到 A 的映射，称为 A 的恒等映射。

(5) 设 $A = C[a,b]$，即区间 $[a,b]$ 上连续函数全体，$B = \mathbf{R}$，则 $f(x(t)) = \int_a^b x(t)\mathrm{d}t$ 是 A 到 B 的映射。

给了映射 $A \stackrel{f}{\longmapsto} B$，若 $f(x) = y$，则称 y 为 x 在 f 下的像，x 为 y 的一个原像。

若 $C \subseteq A$，$D \subseteq B$，则定义 $f(C) = \{f(x) \mid x \in C\}$ 为 C 的像，定义 $f^{-1}(D) = \{x \mid x \in A, f(x) \in D\}$ 为 D 的原像。

给了映射 $A \xrightarrow{f} B$，可以导出两个相关的映射 $2^A \stackrel{f_1}{\longmapsto} 2^B$，$2^B \stackrel{f_2}{\longmapsto} 2^A$，这里 $f_1(C) = f(C)$，$f_2(D) = f^{-1}(D)$。

二、映射的积

定义 7.3 给了两个映射 $A \stackrel{f}{\longmapsto} B$，$B \stackrel{g}{\longmapsto} C$，定义映射 $A \stackrel{h}{\longmapsto} C$，这里 $h(x) = g(f(x))$ 为上述两个映射的积，记为 $h = gf$ 或 $h = g \cdot f$。

按定义，显然 $I_B \cdot f = f$，$fI_A = f$（这里 I 为恒等映射）。

下面证明映射的积满足结合律。

定理 7.1 若给出了三个映射 $A \stackrel{f}{\longmapsto} B \stackrel{g}{\longmapsto} C \stackrel{h}{\longmapsto} D$，则

$$(hg)f = h(gf)。$$

证 显然，上述两个映射的定义域都是 A，值域都是 D。而

$$\begin{aligned}[(hg)f](x) &= (hg)[f(x)] = h\{g[f(x)]\} = h[(gf)(x)] \\ &= [h(gf)](x)。\end{aligned}$$

三、单射，满射，双射

下面给出一些常见类型映射的定义。

定义 7.4 映射 $A|\xrightarrow{f}B$ 叫单射，如果 f 把 A 中不同元素对应到 B 中不同元素。即 $\forall\, x_1,x_2\in A, x_1\neq x_2$，则

$$f(x_1)\neq f(x_2)。$$

判别 (1) 映射 $A|\xrightarrow{f}B$ 是单射 $\Leftrightarrow$ 若 $f(x_1)=f(x_2)$，必 $x_1=x_2$。

(2) 对应 $A|\xrightarrow{f}B$ 是单射 $\Leftrightarrow\forall\, x_1,x_2\in A, f(x_1)=f(x_2)\Leftrightarrow x_1=x_2$。

定义 7.5 映射 $A|\xrightarrow{f}B$ 叫 A 到 B 上的一个满射，如果 $f(A)=B$，即对 $\forall\, y\in B$ 必存在 $x\in A$ 使 $f(x)=y$。

定义 7.6 若映射 $A|\xrightarrow{f}B$ 既是单射又是满射，则称为 A 到 B 上的一个双射。

例 映射 $\left[-\frac{\pi}{2},\frac{\pi}{2}\right]|\xrightarrow{\sin}(-\infty,+\infty)$ 是单射不是满射；映射 $(-\infty,+\infty)\xrightarrow{\sin}[-1,+1]$ 是满射不是单射；映射 $\left[-\frac{\pi}{2},\frac{\pi}{2}\right]|\xrightarrow{\sin}[-1,+1]$ 是双射。

性质 单射的积仍是单射，满射的积仍是满射，双射的积仍是双射。

证明作为读者的练习。

四、可逆映射

定义 7.7 映射 $A|\xrightarrow{f}B$ 称为左可逆的，若存在映射 $B|\xrightarrow{g}A$ 使 $gf=I_A$。这里 g 称为 f 的左逆。

定义 7.8 映射 $A|\xrightarrow{f}B$ 称为右可逆的，若存在映射 $B|\xrightarrow{h}A$ 使 $fh=I_B$。这里 h 称为 f 的右逆。

性质 若映射 f 有左逆 g，右逆 h，则 $g=h$。

证 因 $(gf)h=g(fh)$，即 $h=g$。

判别 映射 $A|\xrightarrow{f}B$ 是左可逆的 $\Leftrightarrow A|\xrightarrow{f}B$ 是单射。

证 $\Rightarrow$若 $f(x_1)=f(x_2)$，g 是 f 的左逆，则 $x_1=I_A(x_1)=gf(x_1)=gf(x_2)=I_A(x_2)=x_2$，故 f 是单射。

$\Leftarrow$若 $A|\xrightarrow{f}B$ 是单射，则可定义 $B|\xrightarrow{g}A$：

$$g(y)=\begin{cases}x,\text{若 } y=f(x)\in f(A)\text{，因 } f \text{ 是单射，故 } x \text{ 惟一，}\\ x_0,\text{若 } y\overline{\in} f(A)\text{，} x_0 \text{ 是 } A \text{ 中任意一个固定元。}\end{cases}$$

显然 $gf=I_A$，故 f 左可逆。

判别 映射 $A|\xrightarrow{f}B$ 是右可逆的$\Leftrightarrow A|\xrightarrow{f}B$ 是满射。证明留作读者练习。

定义 7.9 映射 $A|\xrightarrow{f}B$ 为可逆的，如果存在映射 $B|\xrightarrow{g}A$ 使 $gf=I_A, fg=I_B$。这里 g 称为 f 的逆映射。

判别 映射 $A|\xrightarrow{f}B$ 可逆$\Leftrightarrow A|\xrightarrow{f}B$ 是双射。

证 $\Rightarrow$ 设映射 $A|\xrightarrow{f}B$ 可逆，从而左、右可逆，故是单射又是满射，即为双射。

$\Leftarrow$ 设映射 $A|\xrightarrow{f}B$ 是双射，即是单射又是满射，因而有左逆 g，右逆 h，且 $g=h$ 为 f 的逆映射，即 f 可逆。

若映射 $A|\xrightarrow{f}B$ 可逆，它有逆映射 g 和 h，则可把它们分别看成 f 的左逆与右逆，因而 $f=g$。即映射 f 可逆时，其逆映射惟一，常记成 f^{-1}。

性质 (1) 若 f 可逆，则 $(f^{-1})^{-1}=f$；

(2) 若 f 与 g 可逆，则 $f\cdot g$ 也可逆，且 $(f\cdot g)^{-1}=g^{-1}\cdot f^{-1}$。

上述性质读者可直接按定义自行验证。

§3 等势集合

一、集合等势的概念

定义 7.10 给了两个集合 A 和 B，若存在 A 到 B 上的双射，则称集合 A 和 B 等势，记为 $A\sim B$。

因为恒等映射是双射，又双射是可逆映射，双射的积仍是双射，故有：

(1) $A \sim A$；

(2) 若 $A \sim B$，则 $B \sim A$；

(3) 若 $A \sim B$，$B \sim C$，则 $A \sim C$。

推论 (1) 若存在 A 到 B 的单射，则 A 等势于 B 的一个子集(因为若映射 $A \xrightarrow{f} B$ 是单射，则 $A \sim f(A) \subseteq B$)。

(2) 若存在 A 到 B 上的满射(于是存在 B 到 A 的单射)，则 B 等势于 A 的一个子集。

例 7.2 求证：

(1) $\mathbf{Z} \sim 2\mathbf{Z} = \{2k \mid k \in \mathbf{Z}\}$；

(2) $[a,b] \sim [0,1]$，$(a < b)$；

(3) $(-1,1) \sim (-\infty, +\infty)$；

证 (1) $f(x) = 2x$ 是 $\mathbf{Z}$ 到 $2\mathbf{Z}$ 上的双射。

(2) $f(x) = \dfrac{x-a}{b-a}$ 是$[a,b]$到$[0,1]$上的双射。

(3) $f(x) = \tan \dfrac{\pi}{2} x$ 是$(-1,1)$到$(-\infty,+\infty)$上的双射。

二、有限集

若集合 A 和自然数的一个前段 $N = \{1,2,\cdots,n\} = \{x \mid x \in P, 1 \leqslant x \leqslant n\}$ 等势(这里 n 是一固定正整数)，从而集合 A 可表示为 $A = \{a_1,\cdots,a_n\}$，则称 A 为有限集。n 即通常所说 A 的元素个数，称为 A 的阶，记为$|A|$。有限集 A 的阶也称作为 A 的势。

命题 (1) 两个有限集 A 和 B 等势$\Leftrightarrow |A| = |B|$；

(2) 有限集不能和它的真子集等势。

三、可数集

定义 7.11 和自然数集 P 等势的集合称为可数集。

例 7.3 (1) $\mathbf{Z} = \{0,1,-1,2,-2,\cdots,n,-n,\cdots\}$ 是可数集。

(2) **R**不是可数集。其证明如下：设 **R** 是可数集（反证法），$\mathbf{R}=\{a_1,a_2,\cdots,a_n,\cdots\}$。作闭区间套$\Delta_1,\Delta_2,\cdots,\Delta_n,\cdots$，满足：①$\Delta_1\supseteq\Delta_2\supseteq\cdots\Delta_{n-1}\supseteq\Delta_n\supseteq\cdots$；②$a_1\overline{\in}\Delta_1,\cdots,a_n\overline{\in}\Delta_n,\cdots$。由数学分析的知识知，至少存在一点$c\in\Delta_n,n=1,2,\cdots$。$c\in\mathbf{R}$，则存在$k$使$c=a_k\overline{\in}\Delta_k$，矛盾。故 **R** 不可数。

性质 (1) 可数集与有限集的并与差仍可数；

(2) 可数集的无限子集仍可数；

(3) 有限个可数集的并集仍可数；

(4) 可数个可数集的并集仍可数。

证 (1) 设 A 为可数集，B 为有限集，从而 B 的子集 $C=B-A$ 也为有限集。设

$$A=\{a_1,a_2,\cdots,a_n,\cdots\},$$
$$C=B-A=\{b_1,\cdots,b_t\},$$

则

$$A\cap C=\varnothing,$$
$$A\cup B=A\cup C=\{b_1,\cdots,b_t,a_1,\cdots a_n,\cdots\},$$

易知映射 f：

$$\begin{cases}f(b_t)=t\\ f(a_n)=n+t\end{cases}$$

是 $A\cup B$ 到 P 上的双射，故 $A\cup B$ 为可数集。

适当调整 A 中元素的次序可使

$$A\cap B=\{a_1,\cdots,a_s\},$$

于是

$$A-B=\{a_{s+1},\cdots,a_{s+n},\cdots\},$$

从而映射 f：

$$f(a_{s+n})=n$$

是 $A-B$ 到 P 上的双射，故 $A-B$ 为可数集。

(2) 设 A 为可数集，B 为 A 的无限子集：

$$A = \{a_1, \cdots, a_n, \cdots\},$$
$$B = \{a_{i1}, a_{i2}, \cdots, a_{it}, \cdots\}, 1 \leqslant i_1 < i_2 \cdots < i_t \cdots,$$
于是映射 f:
$$f(a_{i_t}) = t$$
是 B 到 P 上的双射，故 B 为可数集。

(3)的证明类似于(4)。

(4) 设 A_i 为可数集：
$$A_i = \{a_{i1}, a_{i2}, \cdots, a_{in}, \cdots\}, i = 1, 2, \cdots,$$
记
$$A = \bigcup_{i=1}^{\infty} A_i,$$
则映射 f:
$$f(a_{ij}) = (2i - 1)2^{j-1}$$
是 A 到 P 的单射，又因为 $A \supseteq A_i$ 为无限集，所以 A 等势于 P 的一个无限子集，从而是可数集。

推论 有理数集 **Q** 是可数集。

证 **Q** 可看成可数个可数集 A_n 的并集。这里 A_n 是分母为 n 的有理数全体。

例 7.4 求证：$[0,1] \sim (0,1)$。

证 记 $A = \left\{0, 1, \frac{1}{2}, \frac{1}{3}, \cdots, \frac{1}{n}, \cdots\right\}$,
$$A_1 = \left\{\frac{1}{2}, \frac{1}{3}, \cdots, \frac{1}{n}, \cdots\right\},$$
$$B = [0,1] - A = (0,1) - A_1,$$
于是
$$[0,1] = A \cup B; (0,1) = A_1 \cup B。$$
映射 f:
$$f(x) = \begin{cases} \frac{1}{2} & x = 0, \\ \frac{1}{2+n} & x = \frac{1}{n} \in A - \{0\}, \\ x & x \in B, \end{cases}$$

是$[0,1]$到$(0,1)$上的双射。

四、无限集

定义 7.12 不是有限集的集合统称无限集。

性质 (1) 无限集必有可数子集;

(2) 无限集有等势的真子集。

证 性质(1)显然。在性质(1)的基础上,证明性质(2):设 B 是无限集 A 的可数子集,$B=\{b_1,b_2,\cdots,b_n,\cdots\}$,令 $B_1=\{b_2,b_3,\cdots\}$,则 $A_1=B_1\cup(A-B)$ 是 $A=B\cup(A-B)$ 的真子集。映射 f:$f(x)=\begin{cases}b_{i-1},x=b_i\in B\\ x,x\overline{\in}B_1\end{cases}$ 是 A_1 到 A 上的双射。故 $A_1\sim A$。

命题 $A\nsim 2^A$。

证 采用反证法。若存在着 A 到 2^A 上的双射 φ,记 $S=\{x\mid x\in A,x\overline{\in}\varphi(x)\}\in 2^A$。因 φ 是双射,故存在 c 使 $\varphi(c)=S$。

(1) $c\in S=\varphi(c)$,按 S 定义,$c\overline{\in}\varphi(c)$,不可能;

(2) $c\overline{\in}S=\varphi(c)$,按 S 定义,$c\in S$,不可能。

由(1)、(2)知,这样的 c 是不存在的,故不存在 A 到 2^A 上的双射。本命题得证。

§4 等价关系与分类

一、集合的直积

若 $A_1,A_2,\cdots,A_n$ 为 n 个集合,定义集合

$$\{(a_1,a_2,\cdots,a_n)\mid a_i\in A_i,i=1,2,\cdots,n\}$$

为 $A_1,A_2,\cdots,A_n$ 的直积,记为

$$A_1\times A_2\times\cdots\times A_n,$$

这里 $$(a_1,a_2,\cdots,a_n)=(b_1,b_2,\cdots,b_n),$$

$$\Longleftrightarrow a_i = b_i, \quad i = 1, \cdots, n。$$

当 $A_1, \cdots, A_n$ 都是有限集时，有

$$|A_1 \times A_2 \times \cdots \times A_n| = |A_1| \cdot |A_2| \cdots |A_n|。$$

平面 R^2 可看成 R 与 R 的直积；$R^2 = R \times R$；空间 R^3 可看成 R 的三次直积；$R^3 = R \times R \times R$。

二、集合的二元关系

称 $A \times A$ 的子集 R 为集合 A 上的一个二元关系。A 的 2 个元素 a 和 b 称为有关系 R，当且仅当 $(a,b) \in R$，$(a,b) \in R$ 记为 aRb；$(a,b) \overline{\in} R$ 记为 $a\overline{R}b$。

例 7.5 $A = \{a,b,c\}$，$R = \{(a,b),(a,a),(c,c)\}$，则 R 是 A 的一个二元关系，aRb，$b\overline{R}a$，aRa，$b\overline{R}b$ 等。

例 7.6 令 $A = Z$。

设 $R_1 = \{(a,a) | a \in Z\}$，则 R_1 是 Z 上的“相等”关系。

设 $R_2 = \{(a,b) \in Z \times Z | a - b \geqslant 0\}$，则 R_2 是 Z 的“大于、等于”关系。

设 $R_3 = \{(a,b) \in Z \times Z |$ 存在 $c \in Z$ 使 $b = ac\}$，则 R_3 是 Z 上的整除关系。

若 n 是一固定正整数，令 $R_4 = \{(a,b) \in Z \times Z$ 存在 $c \in Z$ 使 $b - a = nc\}$，则 R_4 是 Z 关于 n 的同余关系。

三、等价关系

定义 7.13 集合 A 上的一个二元关系 R 称为等价关系，如果它满足：

(1) 反身性 $\forall a \in A$ 有 aRa；

(2) 对称性 $\forall a,b \in A$，若 aRb，则 bRa；

(3) 传递性 $\forall a,b,c \in A$，若 aRb，且 bRc，则 aRc。

等价关系“R”一般用记号“$\sim$”代替。

Z 的相等关系和同余关系是等价关系。三角形集合的相似关系是

等价关系。A 是集合，2^A 中的等势关系是等价关系。

若给了集合 A 上的一个等价关系"$\sim$"。设 $a\in A$，记

$$\overline{a} = \{x \mid x \in A, x \sim a\},$$

则$\overline{a}$是 A 的非空子集（因 $a\sim a$），称$\overline{a}$为 A 的一个等价元素类，或 a 所在的等价类。

易验证，$a \in \overline{a}$；若 $b \in \overline{a}$，则$\overline{b} \subseteq \overline{a}$。

命题 $\forall a,b \in A, \overline{a} \cap \overline{b} = \varnothing$ 或$\overline{a} = \overline{b}$。

证 若 $\overline{a} \cap \overline{b} \neq \varnothing$，则存在 $x \in \overline{a} \cap \overline{b}$，故 $a \sim x \sim b$，因而 $a \in \overline{b}$，$\overline{a} \subseteq \overline{b}$，同理$\overline{b} \subseteq \overline{a}$. 所以$\overline{a} = \overline{b} = \overline{a} \cap \overline{b}$。

定义 7.14 A 为一集合，$A_i \in 2^A (i \in I)$。若 $i \neq j$，$A_i \cap A_j = \varnothing$，且 $\bigcup\limits_{i\in I} A_i = A$，则称 $\{A_i \mid i \in I\}$ 是 A 的一个分类。此时，A 是两两不交（无公共元素）的子集的并。

定理 7.2 若给了集合 A 上的一个等价关系，（由上命题）则集合 A 是两两不交的等价类的并，从而给出了 A 的一个分类。

反之，若 $\{A_i \mid i\in I\}$ 是 A 的一个分类，在 A 中定义关系 R：$aRb \Leftrightarrow$ 存在 $i\in I$ 使 $a,b\in A_i$。则 R 是 A 上的一个等价关系。

四、商集合

给了集合 A 上的一个等价关系$\sim$，记 $A/\sim = \{\overline{a} \mid a \in A\}$，称为 A 关于$\sim$的商集合。

例 7.7 在 Z 中定义同余关系$\sim$，$a \sim b \Leftrightarrow 5 \mid a - b$。则 $Z/\sim = \{\overline{0},\overline{1},\overline{2},\overline{3},\overline{4}\}$。$Z$ 是可数集，而 $Z/\sim$是仅含 5 个元素的有限集。

定义 7.15 A 到 $A/\sim$的映射 φ：$\varphi(a) = \overline{a}$ 称为 A 到 $A/\sim$的自然映射。

设给了集合 A 到集合 B 上的一个满射 f。在 A 中定义关系 R，$aRb \Leftrightarrow f(a) = f(b)$。显然，$R$ 是等价关系。若 $\overline{a} = \overline{b} \Leftrightarrow f(a) = f(b)$，则 $\varphi(\overline{a}) = f(a)$ 是 $A/\sim$到 B 的单射。又因 f 是满射，故 φ 是 $A/\sim$到 B 上的双射。

§5 偏序关系与 Zorn 公理

一、偏序关系

定义 7.16 集合 A 上的一个关系 R 为偏序关系，如果满足：

(1) 反身性 $\forall a\in A, aRa$；

(2) 反对称性 $\forall a,b\in A$，若 aRb 且 bRa，则 $a=b$；

(3) 传递性 $\forall a,b,c\in A$，若 aRb, bRc，则 aRc。

偏序关系常用"$\leqslant$"来表示。定义了偏序关系的集合称为偏序集。

例 7.8 (1) 实数集 R 关于"小于或等于"关系成一偏序集。

(2) P 关于"整除"关系成一偏序集。

(3) A 为一集合，则 2^A 关于"$\subseteq$"关系成一偏序集。

定义 7.17 若$(A,\leqslant)$为一偏序集。

(1) A 的元素 m 为最大元，如果$\forall x\in A$ 有 $x\leqslant m$；

(2) A 的元素 l 为最小元，如果$\forall x\in A$ 有 $l\leqslant x$；

(3) A 的元素 u 为极大元，如果$\forall x\in A$，$u\leqslant x$，必 $u=x$；

(4) A 的元素 v 为极小元，如果$\forall x\in A$，$x\leqslant v$，必 $v=x$；

(5) 若 $T\subseteq A, b\in A$，b 为 T 的上界，如果$\forall x\in T$，有 $x\leqslant b$。而且如果对$\forall T$ 的上界 c 有 $b\leqslant c$，则称 b 为最小上界；

(6) 若 $T\subseteq A, d\in A$，d 为 T 的下界，如果$\forall x\in T$，有 $d\leqslant x$。而且若对$\forall T$ 的下界 e 有 $e\leqslant d$，则称 d 为最大下界。

偏序集 A 不一定有最大(小)元，极大(小)元。A 的子集 T 也不一定有上(下)界。

例如，P 关于整除关系成一偏序集，P 无最大元和极大元。P 的子集 T

$$T=\{x\in P \mid 1\leqslant x\leqslant 10\}$$

关于整除关系也成偏序集，T 无最大元，但 T 有极大元：6，7，8，9，10。T 作为 P 的子集，有上界，其最小上界为 $1,\cdots,10$ 的最小公倍数2 520。

二、Zorn 引理

定义 7.18 集合 A 的一个偏序关系"$\leqslant$"称为"序"关系，如果 $\forall x, y \in A$，均有 $x \leqslant y$ 或者 $y \leqslant x$。此时 $(A, \leqslant)$ 称为有序集。

实数关于"小于或等于"关系是有序集；而 Z 关于"整除"关系，2^A 关于"包含"关系都不是有序集。

Zorn 引理 若偏序集的任意非空有序子集有上界，则该偏序集有极大元。

§6 势

一、有限集的势

我们称有限集 A 所含元素的个数，即其阶 $|A|$ 为 A 的势，仍记成 $|A|$。

若存在有限集 B 到 A 的单射，则 B 的阶小于或等于 A 的阶，即 B 的势小于或等于 A 的势。若存在 B 到 A 上的满射，则 B 的阶大于或等于 A 的阶，即 A 的势小于或等于 B 的势。

我们把这一概念推广到无限集的情形。

二、势的大小

定义 7.19 S 为一集合，在 2^S 中称 S 的子集 B 的势小于或等于 S 的子集 A 的势，如果存在 B 到 A 的单射，记为 $|B| \leqslant |A|$。

例，若 A 为无限集，则

$$|P| \leqslant |A|。$$

推论 若存在 A 到 B 的满射，则

$$|B| \leqslant |A|。$$

证 设 $A \mapsto\xrightarrow{f} B$ 为满射，作映射 $B \mapsto\xrightarrow{g} A$：

$$g(b) = a, a \in f^{-1}(b)$$

若 $g(b_1) = g(b_2) = a$，则 $b_1 = f(a) = b_2$，即 g 为单射，于是

$$|B| \leqslant |A|。$$

引理 若 A,B 为集合 S 的 2 个子集，$|A| \leqslant |B|$，且 $|B| \leqslant |A|$，则 $A \sim B$。记成 $|A| = |B|$。

证 因 $|A| \leqslant |B|$，故存在 A 到 B 的单射 f。同理存在着 B 到 A 的单射 g。

$\forall a \in A$，由于 g 为单射，故 $g^{-1}(a)$或为空集，或恰由一个元素 $b \in B$ 组成。前者称 a 是无前辈的，后者称 b 为 a 的前辈。类似地，$\forall b \in B$，或者 $f^{-1}(b) = \varnothing$ (b 无前辈)，或者 $f^{-1}(b) = a' \in A$。(a' 为 b 的前辈)。如果继续用这种方法追寻元素 $x(\in A$ 或 $B)$的家谱，便会产生 3 种情况：(1) 追寻到 A 中一个无前辈的元素而停止，此元素称为 x 的祖宗；(2) 追寻到 B 中一个无前辈的元素而停止，此元素为 x 的祖宗；(3) 无休止地追寻下去(无祖宗)。

设 $A_1 = \{a \in A \mid a$ 在 A 中有祖宗$\}$

$A_2 = \{a \in A \mid a$ 在 B 中有祖宗$\}$

$A_3 = \{a \in A \mid a$ 无祖宗$\}$

$B_1 = \{b \in B \mid b$ 在 A 中有祖宗$\}$

$B_2 = \{b \in B \mid b$ 在 B 中有祖宗$\}$

$B_3 = \{b \in B \mid b$ 无祖宗$\}$

其中 A_1, A_2, A_3 是 A 的一个分类；B_1, B_2, B_3 是 B 的一个分类。

易得，$A_1 \cup A_3 \mid \xrightarrow{f} B_1 \cup B_3$ 和 $B_2 \mid \xrightarrow{g} A_2$ 是双射。作 A 到 B 的映射 h：

$$h(a) = \begin{cases} f(a), & 若 a \in A_1 \cup A_3 \\ g^{-1}(a), & 若 a \in A_2 \end{cases}$$

则 h 是双射。故 $A \sim B$。

定理 7.3 若 A,B 为集合 S 的两个子集，则必

$$|A| \leqslant |B| \text{ 或 } |B| \leqslant |A|。$$

证 作 $\sum = \{(X,f) | X \subseteq A, f \text{ 是 } X \text{ 到 } B \text{ 的单射}\}$，$\sum$ 中定义"$\leqslant$"：$(X_1,f_1) \leqslant (X_2,f_2) \Leftrightarrow X_1 \subseteq X_2, \forall x \in X_1, f_1(x) = f_2(x)$，则"$\leqslant$"是偏序关系。显然，$\sum$ 满足 Zorn 引理的条件，故有极大元(X_0, f_0)。作为极大元，仅有两种可能：

(1) $X_0 = A$，此时 $|A| \leqslant |B|$；

(2) $f_0(X_0) = B$，此时 $|B| \leqslant |A|$。

否则，取

$$x_1 \in A - X_0, y_1 \in B - f_0(X_0), \text{令 } X_1 = X_0 \cup \{X_1\},$$

$$f_1(x) = \begin{cases} f_0(x), & \text{当 } x \in X_0 \\ y_1, & \text{当 } x = x_1 \end{cases} \quad \text{则}(X_1, f_1) \in \Sigma,$$

$(X_0,f_0) < (X_1,f_1)$。这和(X_0,f_0)是极大元矛盾，不可能。

若用记号 $|A| < |B|$ 表示：存在 A 到 B 的单射，不存在 B 到 A 的单射，从而不存在 A 到 B 的双射，则有下面的推论：

推论 若 A 与 B 是集合 S 的两个子集，则

$$|A| < |B|, |A| = |B|, |B| < |A|$$

三者中有且仅有一个成立。

三、势的运算

引理 1 $|P \times \{0,1\}| = |P|$，即 $P \times \{0,1\} \sim P$。

证 映射 $f((n,0)) = 2n, f((n,1)) = 2n - 1$ 是集合 $P \times \{0,1\}$ 到集合 P 上的双射，故引理得证。

引理 2 若 A 为无限集，则 $|A \times \{0,1\}| = |A|$。

证 作 $\Sigma = \{(X,f) | X \subseteq A, f \text{ 是 } X \text{ 到 } X \times \{0,1\} \text{ 上的双射}\}$。因为 A 有可数子集，由引理 1，Σ 非空。类似定理 1.4 的证明，定义 Σ 的偏序关系，Σ 满足 Zorn 引理的条件，故有极大元(X_0,f_0)。

若 $|X_0| \neq |A|$，因 $A = X_0 \cup (A - X_0)$，故 $A - X_0$ 不是有限集，否则 $A \sim X_0$。设 X_1 是 $A - X_0$ 的可数子集，因为

$(X_0 \cup X_1) \times \{0,1\} = (X_0 \times \{0,1\}) \cup (X_1 \times \{0,1\}) \sim X_0 \cup X_1$，

这和(X_0, f_0)是极大元矛盾，故必 $|X_0| = |A|$，从而

$$|A| = |X_0| = |X_0 \times \{0,1\}| = |A \times \{0,1\}|。$$

引理得证。

定理 7.4 若 A, B 为两个集合，$|B| \leqslant |A|$，A 为无限集，则

$$|A \cup B| = |A|。$$

证 因为 $|B| \leqslant |A|$，设 g 为 B 到 A 的单射。显然

$$|A| \leqslant |A \cup B|,$$

又，映射：

$$f(x) = \begin{cases} (x,0), & x \in A \\ (g(x),1), & x \in B - A \end{cases}$$

是 $A \cup B$ 到 $A \times \{0,1\}$ 的单射，于是

$$|A \cup B| \leqslant |A \times \{0,1\}| = |A|,$$

故

$$|A| = |A \cup B|。$$

引理 3 $|P \times P| = |P|$。

证 因为 $f((a,b)) = 2^{a-1}(2b-1)$ 是 $P \times P$ 到 P 的双射。

下述引理的证明是显然的。

引理 4 若 $|A| = |C|$，$|B| = |D|$，则 $|A \times B| = |C \times D|$。

引理 5 若 A 为无限集，则 $|A \times A| = |A|$。

证 作 $\Sigma = \{(X,f) \mid X \subseteq A,\ f$ 为 X 到 $X \times X$ 的双射$\}$。由引理 3，Σ 非空，仿定理 1.4 的证明方法使 Σ 偏序化，Σ 满足 Zorn 引理的条件，有极大元(X_0, f_0)。

若 $|X_0| \neq |A|$，由定理 7.4，$|A - X_0| > |X_0|$。故存在 $B \subseteq A - X_0$，且 $|B| = |X_0|$。$B \cap X_0 = \varnothing$，从而

$$(X_0 \cup B) \times (X_0 \cup B) = (X_0 \times X_0) \cup [(X_0 \times B) \cup (B \times X_0) \cup (B \times B)]。$$

但

$$|(X_0 \times B) \cup (B \times X_0) \cup (B \times B)| = |B \times B| = |B|,$$

故存在 B 到 $(X_0 \times B) \cup (B \times X_0) \cup (B \times B)$ 的双射，从而存在保持 f_0 的$(X_0 \cup B)$到$(X_0 \cup B) \times (X_0 \cup B)$的双射，这和$(X_0, f_0)$是 Σ 极大元矛盾。故必 $|X_0| = |A|$，从而

$$|A \times A| = |X_0 \times X_0| = |X_0| = |A|。$$

定理 7.5 若 A 和 B 为两个非空集合，$|B| \leqslant |A|$，A 为无限集，则 $|A \times B| = |A|$。

证 因为 $|A| \leqslant |A \times B| \leqslant |A \times A| = |A|$。

推论 (1) 若 A 为无限集，则 $|P \times A| = |A|$；

(2) 若 B 为非空有限集，则 $|B \times P| = |P|$；

(3) 若 A 为无限集，则 $|A^n| = |\underbrace{A \times A \times \cdots \times A}_{n\text{个}}| = |A|$；

(4) 若 A 为无限集，则 $|\bigcup\limits_{n \in P} A^n| = |A|$。

证 推论(1)，(2)，(3)是显然的。在此只证明推论(4)：

由推论(3)，设 f_n 是 A^n 到 A 上的双射，则映射 $\sigma : \sigma(x) = (n, f_n(x))$。若 $x \in A^n$，则映射 σ 是 $\bigcup\limits_{n \in P} A^n$ 到 $P \times A$ 上的双射。故

$$\left|\bigcup_{n \in P} A^n\right| = |P \times A| = |A|。$$

例 7.9 设 B 为二元集，$B = \{b_1, b_2\}$；A 为无限集；Σ 为 A 到 B 的映射全体所成的集合，求证：

$$\Sigma \sim 2^A。$$

证 (1) 设 C 为 A 的子集，即 $C \in 2^A$，定义 A 到 B 的映射 f_C：

$$f_C(x) = \begin{cases} b_1, & x \in C \\ b_2, & x \in A - C \end{cases}$$

显然，映射

$$2^A \xrightarrow{\sigma} \Sigma$$
$$\sigma(C) = f_C$$

是单射，故

$$|2^A| \leqslant |\Sigma|。$$

(2) 作映射 $\Sigma \xrightarrow{\eta} 2^A \times 2^A$：

$$\eta(f) = (f^{-1}(b_1), f^{-1}(b_2))。$$

若 f_1 与 f_2 都是 A 到 B 的映射，且 $f_1 \neq f_2$，则

$$(f_1^{-1}(b_1), f_1^{-1}(b_2)) \neq (f_2^{-1}(b_1), f_2^{-1}(b_2)),$$

即 η 是单射，故

$$|\Sigma| \leqslant |2^A \times 2^A| = |2^A|。$$

由(1)与(2)得

$$|\Sigma| = |2^A|，即\ \Sigma \sim 2^A。$$

代替(2)，我们可以利用映射 $\Sigma \xrightarrow{\delta} 2^{A\times B}$：

$$\delta(f) = \{(x, f(x)) \mid x \in A\}$$

是单射这一点可知：

$$|\Sigma| \leqslant |2^{A\times B}| = |2^A|。$$

习　题　7

1. 设 $A = \{x \mid x \in \mathbf{R}, 1 \leqslant x \leqslant 4\}$，
   $B = \{x \mid x \in \mathbf{R}, 0 \leqslant x < 3\}$，
   试求 $A \cap B, A \cup B, A - B, B - A$。
2. 求证：$A = B \Leftrightarrow A \cup B = A \cap B$。
3. 设 $A_n = \{x \mid x \in \mathbf{R}, x \geqslant n\}, n = 1, 2, 3\cdots$。求
   $$\bigcup_{i=1}^{\infty} A_i \text{ 和 } \bigcap_{i=1}^{\infty} A_i。$$
4. 设 $A = \{x \mid x \in \mathbf{R}, x^4 - 3x^2 = 4\}$，写出 2^A 的所有元素。
5. 求证：单射(满射)的积仍是单射(满射)。

6. 给了映射 $A \xrightarrow{f} B, C \subseteq A, D \subseteq B$，求证：

(1) $f^{-1}[f(C)] \supseteq C$，且 f 是单射时取等号；

(2) $f[f^{-1}(D)] \subseteq D$，且 f 是满射时取等号。

7. 给了映射 $A \xrightarrow{f} B, A_i(i \in I)$ 是 A 的子集，求证：

(1) $f(\bigcup_{i \in I} A_i) = \bigcup_{i \in I} f(A_i)$；

(2) $f(\bigcap_{i \in I} A_i) \subseteq \bigcap_{i \in I} f(A_i)$。

举例说明，等号不一定成立。

8. 设集合 A 有 n 个元素，集合 B 有 s 个元素，试确定 A 到 B 的单射和满射个数。

9. 若 n 为一固定正整数，试求 n 个互不相同的 $\mathbf{Z}$ 到 $\mathbf{Z}$ 的映射使它们有共同的左逆。

10. 求证：坐标平面上所有格子点（即坐标为整数的点）的集合是可数的。

11. 求证：无理数的集合即 $\mathbf{R}-\mathbf{Q}$ 是不可数的。

12. 求证：若 $a < b$，则

(1) 闭区间 $[a,b] \sim [0,1]$；

(2) 开区间 $(a,b) \sim$ 闭区间 $[a,b]$。

13. 若 A 是无限集，B 是有限集，则 $(A \cup B) \sim A$。

14. 若 A 为有限集，则映射 $A \xrightarrow{f} A$ 是单射的充要条件是：f 是满射。

15. 在 $\mathbf{Z}$ 中，定义关系 R：$aRb \Leftrightarrow 4 | a + 3b$，求证：$R$ 是 Z 上的一个等价关系。

16. 在实数集 $\mathbf{R}$ 中定义关系 T：$aTb \Leftrightarrow a - b \in Z$，求证：$T$ 是 $\mathbf{R}$ 的一个等价关系，并确定商集合 $\mathbf{R}/T$。

17. 设 R_1, R_2 是集合 A 上的两个等价关系（即 $A \times A$ 的子集），试问 $R_1 \cap R_2$ 和 $R_1 \cup R_2$ 是不是 A 上的等价关系？试说明理由。

18. 若偏序集有最大元，则仅有一个。

19. 求偏序集 $(2^A, \subseteq)$ 的最大元和最小元。

20. P 关于整除关系成一偏序集，求子集 $T = \{2,3,4,5\}$ 的最小

第八章　线 性 空 间

本章把具体的、直观的平面与几何空间推广到抽象的线性空间。

§1　线性空间的概念

一、定义

我们从下面的例子导出线性空间的定义。

例 8.1　在几何空间，两个向量按平行四边形法则相加，一个向量可与实数作数量乘法。我们常用这两种运算来描述几何与力学对象的一些性质。

例 8.2　为了解线性方程组，我们引入了 n 维向量，它们有加法与数乘两种运算。

例 8.3　数域 P 上一元多项式同样有加法与数乘运算。

这些例子中，研究的对象各不相同，但它们有一个共同点：都有加法与数乘两种运算。为了在它们共同点的基础上，导出普遍适用的理论，我们引进抽象的线性空间的概念。

定义 8.1　非空集合 V 称为数域 P 上的线性空间，如果：

(1) 在集合 V 中定义了“加法”运算，即给出一个法则，对 V 中任意两个元素 α 与 β，都存在 V 中惟一的一个元素 γ 与之对应，γ 称为 α 与 β 的和，记成 $\gamma = \alpha + \beta$。“加法”运算满足如下规则：

i) 交换律，即任意 $\alpha,\beta \in V$，有 $\alpha + \beta = \beta + \alpha$；

ii) 结合律，即任意 $\alpha,\beta,\gamma \in V$，有 $(\alpha + \beta) + \gamma = \alpha + (\beta + \gamma)$；

iii) V 中存在零元素 $\boldsymbol{0}$，对 V 中任意向量 α 都有

$$\alpha + \boldsymbol{0} = \boldsymbol{0} + \alpha = \alpha;$$

iv) 对 V 中任意元素 α，都存在 V 中元素 β 使

$$\alpha+\beta=\boldsymbol{0}\ ,\beta\text{称为}\alpha\text{的负元素。}$$

(2) 定义了"数乘"运算，即给出了一个法则，对 P 中任意数 k 与 V 中任意元素 α，都有 V 中惟一的元素 δ 与之对应，δ 称为 k 与 α 的数量积，记成 $\delta=k\alpha$。数乘运算满足：

v) 任意 $\alpha\in V,1\cdot\alpha=\alpha$；

vi) 任意 $k,l\in P,\alpha\in V,(kl)\alpha=k(l\alpha)$；

vii) 任意 $k,l\in P,\alpha\in V,(k+l)\alpha=k\alpha+l\alpha$；

viii) 任意 $k\in P,\alpha,\beta\in V,k(\alpha+\beta)=k\alpha+k\beta$。

例 8.4 按 n 维向量的加法与数乘：

$$P^n=\left\{\begin{pmatrix}a_1\\ \vdots\\ a_n\end{pmatrix}\middle|\, a_i\in P,i=1,\cdots,n\right\}$$

是数域 P 上的线性空间。

例 8.5 按矩阵的加法与数乘：

$$P^{s\times n}=\{(a_{ij})_{s\times n}|a_{ij}\in P,i=1,\cdots s;j=1,\cdots,n\}$$

是数域 P 上的线性空间。

例 8.6 按多项式的加法与数乘，$P[x]$ 是数域 P 上的线性空间。

按多项式的加法与数乘

$$P[x]_n=\{f(x)\in P[x]|\deg f(x)<n\}$$

是数域 P 上的线性空间。但集合

$$\{f(x)\in P[x]|\deg f(x)=n\}$$

不是线性空间，因为 $(1-x^n)+x^n$ 不在其内。

例 8.7 按数的加法与乘法，数域 P 本身是 P 上的线性空间。

例 8.8 按数的加法与乘法，复数域 $\mathbf{C}$ 是实数域 $\mathbf{R}$ 上的线性空间。

线性空间的元素也称向量，当然这里的向量比以前引进的 n 维向

量要广泛得多。向量常用小写希腊字母 α,β,γ 等表示，数域 P 中的数常用小写英文字母表示。

二、简单性质

(1) 零向量(即零元素)是惟一的；

(2) 一个元素的负元素是惟一的，α 的负元素记成“$-\alpha$”；

(3) $0\alpha=k\mathbf{0}=\mathbf{0}$；

(4) $(-k)\alpha=k(-\alpha)=-k\alpha$，特别 $(-1)\alpha=-\alpha$；

(5) 若 $k\alpha=\mathbf{0}$，则 $k=0$ 或 $\alpha=\mathbf{0}$。

证 (1) 设 $\mathbf{0}_1$ 与 $\mathbf{0}_2$ 都是零向量。

考虑和 $\mathbf{0}_1+\mathbf{0}_2$，由于 $\mathbf{0}_1$ 是零向量，故 $\mathbf{0}_1+\mathbf{0}_2=\mathbf{0}_2$；又由于 $\mathbf{0}_2$ 也是零向量，故 $\mathbf{0}_1+\mathbf{0}_2=\mathbf{0}_1$。于是 $\mathbf{0}_2=\mathbf{0}_1$，即零向量惟一。

(2) 若向量 β,γ 都是向量 α 的负元素，则

$$\beta=\beta+\mathbf{0}=\beta+(\alpha+\gamma)=(\beta+\alpha)+\gamma=\mathbf{0}+\gamma=\gamma,$$

即负元素惟一。

记 α 的负元素为 $-\alpha$，由此可定义向量的减法：

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

$$\begin{aligned}(3)\ 0\alpha&=0\alpha+\mathbf{0}=0\alpha+\alpha-\alpha=0\alpha+1\alpha-\alpha\\&=(0+1)\alpha-\alpha=\alpha-\alpha=\mathbf{0}\ 。\end{aligned}$$

$$k\cdot\mathbf{0}=k(0\alpha)=(k\cdot 0)\alpha=0\alpha=\mathbf{0}$$

(4) $-k\alpha$ 表示 $k\alpha$ 的负元素，又

$$k\alpha+(-k)\alpha=[k+(-k)]\alpha=0\alpha=\mathbf{0}\ ,$$

$$k\alpha+k(-\alpha)=k[\alpha+(-\alpha)]=k\mathbf{0}=\mathbf{0}\ ,$$

即 $(-k)\alpha,k(-\alpha)$ 也是 $k\alpha$ 的负元素，由负元素的惟一性得

$$(-k)\alpha=k(-\alpha)=-k\alpha。$$

(5) $k\alpha=\mathbf{0}$，若 $k\neq 0$，则 $\alpha=1\cdot\alpha=\frac{1}{k}k\alpha=\frac{1}{k}\cdot\mathbf{0}=\mathbf{0}$ 。

三、向量的线性关系

定义 8.2 称向量 β 是向量 $\alpha_1,\cdots,\alpha_s$ 的线性组合，或称向量 $\alpha_1,\cdots,\alpha_s$ 能线性表出 β，记成 $\alpha_1,\cdots,\alpha_s \Rightarrow \beta$，如果存在数 $k_1,\cdots,k_s \in \boldsymbol{P}$ 使

$$\beta = \sum_{i=1}^{s} k_i\alpha_i。$$

定义 8.3 称向量组 Σ_1 能线性表出向量组 Σ_2，记成 $\Sigma_1 \Rightarrow \Sigma_2$，如果任意 $\beta \in \Sigma_2$，必存在 $\alpha_1,\cdots,\alpha_s \in \Sigma_1$ 使

$$\alpha_1,\cdots,\alpha_s \Rightarrow \beta,$$

即 Σ_2 中的任意一个向量可用 Σ_1 中有限个向量线性表出。

定义 8.4 $r \geqslant 1$，向量组 $\alpha_1,\cdots,\alpha_r$ 称为线性相关，如果存在 $\boldsymbol{P}$ 中 r 个不全为 0 的数 $k_1,\cdots,k_r$ 使

$$\sum_{i=1}^{r} k_i\alpha_i = \boldsymbol{0},$$

即向量组 $\alpha_1,\cdots,\alpha_r$ 线性相关$\Leftrightarrow$方程式 $\sum\limits_{i=1}^{r} x_i\alpha_i = 0$ 在 $\boldsymbol{P}$ 中有非零解。

若向量组 $\alpha_1,\cdots,\alpha_r$ 不线性相关，则称为线性无关，即方程式 $\sum\limits_{i=1}^{r} x_i\alpha_i = \boldsymbol{0}$ 在 $\boldsymbol{P}$ 中仅有零解。

引理 8.1 若 $\{\alpha_1,\cdots,\alpha_s\} \Rightarrow \{\beta_1,\cdots,\beta_t\}$，$t > s$，则向量组 $\beta_1,\cdots,\beta_t$ 线性相关。

证 按假设，采用矩阵记号，可设

$$(\beta_1,\cdots,\beta_t) = (\alpha_1,\cdots,\alpha_s)A, A \in \boldsymbol{P}^{s\times t}。$$

因为 $s < t$，齐次线性方程组 $AX = \boldsymbol{0}$ 有非零解 $\begin{pmatrix} k_1 \\ \vdots \\ k_t \end{pmatrix}$。此时

$$\sum_{i=1}^{t}k_i\beta_i=(\beta_1,\cdots,\beta_t)\cdot\begin{pmatrix}k_1\\ \vdots\\ k_t\end{pmatrix}=(\alpha_1,\cdots,\alpha_s)A\begin{pmatrix}k_1\\ \vdots\\ k_t\end{pmatrix}$$（由矩阵乘法性质(3)的推论）

$$=(\alpha_1,\cdots,\alpha_s)\cdot\begin{pmatrix}0\\ \vdots\\ 0\end{pmatrix}=\boldsymbol{0}\text{。}$$

故 $\beta_1,\beta_2,\cdots,\beta_t$ 线性相关。

定义 8.5 设 Σ 为一向量组，$\alpha_1,\cdots,\alpha_r\in\Sigma$。$\alpha_1,\cdots,\alpha_r$ 称为 Σ 的一个极大线性无关组，如果

(1) $\alpha_1,\cdots,\alpha_r$ 线性无关；

(2) 任意 $\alpha\in\Sigma$, $\alpha_1,\cdots,\alpha_r,\alpha$ 线性相关。

命题 8.1 $\alpha_1,\cdots,\alpha_r\in\Sigma$，则 $\alpha_1,\cdots,\alpha_r$ 是 Σ 的一个极大线性无关组$\Leftrightarrow$(1) $\alpha_1,\cdots,\alpha_r$ 线性无关；(2) $\alpha_1,\cdots,\alpha_r\Rightarrow\Sigma$。

证 $\Rightarrow$任意 $\alpha\in\Sigma,\alpha_1,\cdots,\alpha_r,\alpha$ 线性相关，即存在不全为 0 的数 $k_1,\cdots,k_r,k$ 使

$$k_1\alpha_1+\cdots+k_r\alpha_r+k\alpha=\boldsymbol{0}\text{.}$$

若 $k=0$，则 $k_1,\cdots,k_r$ 不全为 0，且 $\sum\limits_{i=1}^{r}k_i\alpha_i=\boldsymbol{0}$，与 $\alpha_1,\cdots,\alpha_r$ 线性无关的假设矛盾，故 $k\neq0$，从而

$$k=\left(-\frac{k_1}{k}\right)\alpha_1\cdots+\left(-\frac{k_r}{k}\right)\alpha_r,$$

即$\alpha_1,\cdots,\alpha_r\Rightarrow\alpha$，于是 $\alpha_1,\cdots,\alpha_r\Rightarrow\Sigma$。

$\Leftarrow$任意 $\alpha\in\Sigma,\alpha_1,\cdots,\alpha_r\Rightarrow\alpha$，即存在数 $k_1,\cdots,k_r$ 使 $\alpha=\Sigma k_i\alpha_i$，因而

$$\sum_{i=1}^{r}k_i\alpha_i+(-1)\alpha=\boldsymbol{0}\text{。}$$

由于 $k_1,\cdots,k_r,(-1)$，已经不全为 0，故 $\alpha_1,\cdots,\alpha_r,\alpha$ 线性相关。于是 $\alpha_1,\cdots,\alpha_r$ 是 Σ 的一个极大线性无关组。

定理 8.1 向量组的任意两个极大线性无关组所含向量个数相等，即若$\{\alpha_1,\cdots,\alpha_r\}$与$\{\beta_1,\cdots,\beta_t\}$都是向量组$\Sigma$的极大线性无关组，则$r=t$。

证 若$r\neq t$，可设$t>r$。

因为$\alpha_1,\cdots,\alpha_r$是Σ的一个极大线性无关组，由命题 8.1，$\alpha_1,\cdots,\alpha_r\Rightarrow\Sigma$。而$\beta_1,\cdots,\beta_t\in\Sigma$，故

$$\{\alpha_1,\cdots,\alpha_r\}\Rightarrow\{\beta_1,\cdots,\beta_t\},$$

且$t>r$，由引理 8.1，$\beta_1,\cdots,\beta_t$线性相关，于是导出了矛盾，故必$r=t$。

四、向量组的秩

定义 8.6 向量组Σ的极大线性无关组(如果存在的话)所含向量个数称为向量组Σ的秩，记成$\mathrm{r}(\Sigma)$。

性质 (1) 向量组$\alpha_1,\cdots,\alpha_r$线性无关$\Leftrightarrow\{\alpha_1,\cdots,\alpha_r\}$秩为$r$；

(2) 若向量组Σ的秩为r，$s>r$，$\alpha_1,\cdots,\alpha_s\in\Sigma$，则$\alpha_1,\cdots,\alpha_s$线性相关；

(3) 若Σ_1与Σ_2是秩有限的向量组，且$\Sigma_1\Rightarrow\Sigma_2$，则$\mathrm{r}(\Sigma_1)\geqslant\mathrm{r}(\Sigma_2)$。特别，若$\Sigma_1\Rightarrow\Sigma_2$，$\Sigma_2\Rightarrow\Sigma_1$，则$\mathrm{r}(\Sigma_1)=\mathrm{r}(\Sigma_2)$。

证 (1) 显然。

(2) 因为$\mathrm{r}(\Sigma)=r$，可设$\beta_1,\cdots,\beta_r$是Σ的一个极大线性无关组，由命题 8.1

$$\{\beta_1,\cdots,\beta_r\}\Rightarrow\{\alpha_1,\cdots,\alpha_s\},s>r,$$

由引理 8.1，$\alpha_1,\cdots,\alpha_s$线性相关。

(3) 设$\alpha_1,\cdots,\alpha_r$与$\beta_1,\cdots,\beta_s$分别是Σ_1与Σ_2的极大线性无关组，于是$\mathrm{r}(\Sigma_1)=r$，$\mathrm{r}(\Sigma_2)=s$。

若$s>r$，因为$\alpha_1,\cdots,\alpha_r\Rightarrow\Sigma_1\Rightarrow\Sigma_2$，于是

$$\{\alpha_1,\cdots,\alpha_r\}\Rightarrow\{\beta_1,\cdots,\beta_s\},$$

由引理 8.1，$\beta_1,\cdots,\beta_s$线性相关，与$\beta_1,\cdots,\beta_s$是Σ_2的极大线性无关组的假设矛盾，故必$s\leqslant r$，即$\mathrm{r}(\Sigma_1)\geqslant\mathrm{r}(\Sigma_2)$。

命题 8.2 若 $\mathrm{r}(\boldsymbol{\Sigma})=r$:

(1) 如果 $\alpha_1,\cdots,\alpha_r\in\boldsymbol{\Sigma}$,且 $\alpha_1,\cdots,\alpha_r$ 线性无关,则 $\alpha_1,\cdots,\alpha_r$ 是 $\boldsymbol{\Sigma}$ 的一个极大线性无关组;

(2) 如果 $\alpha_1,\cdots,\alpha_r\in\boldsymbol{\Sigma}$,且 $\alpha_1,\cdots,\alpha_r\Rightarrow\boldsymbol{\Sigma}$,则 $\alpha_1,\cdots,\alpha_r$ 是 $\boldsymbol{\Sigma}$ 的一个极大线性无关组。

证 $\mathrm{r}(\boldsymbol{\Sigma})=r$,设 $\beta_1,\cdots,\beta_r$ 是 $\boldsymbol{\Sigma}$ 的一个极大线性无关组:

(1) 任意 $\alpha\in\Sigma,\beta_1,\cdots,\beta_r\Rightarrow\alpha_1,\cdots,\alpha_r,\alpha$,由引理 8.1, $\alpha_1,\cdots,\alpha_r,\alpha$ 线性相关。按定义 8.5,$\alpha_1,\cdots,\alpha_r$ 是 $\boldsymbol{\Sigma}$ 的一个极大线性无关组。

(2) $\{\alpha_1,\cdots,\alpha_r\}\Rightarrow\boldsymbol{\Sigma}$, 由性质(3),$\{\alpha_1,\cdots,\alpha_r\}$秩$\geqslant r$.由性质(1), $\alpha_1,\cdots,\alpha_r$ 线性无关。根据命题 8.1,$\alpha_1,\cdots,\alpha_r$ 是 $\boldsymbol{\Sigma}$ 的一个极大线性无关组。

§2 有限维线性空间

定义 8.7 数域 $\boldsymbol{P}$ 上线性空间 $\boldsymbol{V}$,如果存在由有限个向量 $\boldsymbol{\varepsilon}_1,\cdots,\boldsymbol{\varepsilon}_n$ 组成 $\boldsymbol{V}$ 的极大线性无关组,则称 $\boldsymbol{V}$ 是 n 维线性空间,记成 $n=\dim\boldsymbol{V}$。$\boldsymbol{V}$ 的极大线性无关组称为 $\boldsymbol{V}$ 的基。

例 8.9 $\boldsymbol{P}^n$ 是数域 $\boldsymbol{P}$ 上 n 维线性空间,因为它有一组基

$$\boldsymbol{\varepsilon}_1=\begin{pmatrix}1\\0\\\vdots\\0\end{pmatrix},\cdots,\boldsymbol{\varepsilon}_n=\begin{pmatrix}0\\\vdots\\0\\1\end{pmatrix}。$$

例 8.10 $\boldsymbol{P}^{s\times n}$是 $\boldsymbol{P}$ 上 $s\times n$ 维线性空间,它有基

$$\{\boldsymbol{E}_{ij}\mid i=1,\cdots,s;j=1,\cdots,n\},$$

这里 $\boldsymbol{E}_{ij}$表示 i 行、j 列的元素为 1,其余元素为 0 的 $s\times n$ 阵。

例 8.11 $\boldsymbol{C}$ 是 $\mathbf{R}$ 上的二维线性空间,它有一组基: $1,\mathbf{i}$。

例 8.12 $\boldsymbol{P}[x]_n$ 是 $\boldsymbol{P}$ 上 n 维线性空间,它有一组基: $1,\boldsymbol{x},\boldsymbol{x}^2,\cdots,\boldsymbol{x}^{n-1}$。

一、基，维数，坐标

利用上一节的结果，把 Σ 换成 V，即得下列各命题。

命题 8.3 若 $s>n$，则 n 维线性空间中任意 s 个向量必线性相关，(见本章定义 8.6 后的性质(2))。

命题 8.4 n 维线性空间中任意 n 个线性无关的向量必是线性空间的一组基(见命题 8.2)。

命题 8.5 n 维线性空间 V 中，若 $\alpha_1,\cdots,\alpha_n \Rightarrow V$，则 $\alpha_1,\cdots,\alpha_n$ 是 V 的一组基(见命题 8.2)。

定义 8.8 若 $\varepsilon_1,\cdots,\varepsilon_n$ 是数域 P 上线性空间 V 的一组基，则对任意 $\alpha\in V$，$\varepsilon_1,\cdots,\varepsilon_n \Rightarrow \alpha$，即存在数 $a_1,\cdots,a_n\in P$，使

$$\alpha=\sum_{i=1}^{n}a_i\varepsilon_i=(\varepsilon_1,\cdots,\varepsilon_n)\begin{pmatrix}a_1\\ \vdots\\ a_n\end{pmatrix}。$$

称 $\begin{pmatrix}a_1\\ \vdots\\ a_n\end{pmatrix}$ 为向量 α 在基 $\varepsilon_1,\cdots,\varepsilon_n$ 下的坐标。

性质 (1) 任意向量 α 在基 $\varepsilon_1,\cdots,\varepsilon_n$ 下的坐标是惟一的；

(2) 若 α,β 在基 $\varepsilon_1,\cdots,\varepsilon_n$ 下的坐标分别为 $\begin{pmatrix}a_1\\ \vdots\\ a_n\end{pmatrix}$ 与 $\begin{pmatrix}b_1\\ \vdots\\ b_n\end{pmatrix}$，则向量 $(\alpha+\beta)$，$k\alpha$ 在基 $\varepsilon_1,\cdots,\varepsilon_n$ 下的坐标分别为

$$\begin{pmatrix}a_1+b_1\\ \vdots\\ a_n+b_n\end{pmatrix} 与 \begin{pmatrix}ka_1\\ \vdots\\ ka_n\end{pmatrix}。$$

证 (1) 若 α 在基 $\varepsilon_1,\cdots,\varepsilon_n$ 有两组坐标 $\begin{pmatrix}a_1\\ \vdots\\ a_n\end{pmatrix}$，$\begin{pmatrix}c_1\\ \vdots\\ c_n\end{pmatrix}$，即

$$\boldsymbol{\alpha}=(\boldsymbol{\varepsilon}_1,\cdots,\boldsymbol{\varepsilon}_n)\begin{pmatrix}a_1\\ \vdots\\ a_n\end{pmatrix}=(\boldsymbol{\varepsilon}_1,\cdots,\boldsymbol{\varepsilon}_n)\begin{pmatrix}c_1\\ \vdots\\ c_n\end{pmatrix},$$

从而

$$\sum_{i=1}^{n}a_i\boldsymbol{\varepsilon}_i=\sum_{i=1}^{n}c_i\boldsymbol{\varepsilon}_i,$$

于是

$$\sum_{i=1}^{n}(a_i-c_i)\boldsymbol{\varepsilon}_i=0。$$

由于 $\boldsymbol{\varepsilon}_1,\cdots,\boldsymbol{\varepsilon}_n$ 是基，因而线性无关，故 $a_i-c_i=0$，或 $a_i=c_i,i=1,\cdots,n$。即 $\begin{pmatrix}a_1\\ \vdots\\ a_n\end{pmatrix}=\begin{pmatrix}c_1\\ \vdots\\ c_n\end{pmatrix}$，故得坐标惟一性的证明。

(2) $$\boldsymbol{\alpha}+\boldsymbol{\beta}=\sum_{i=1}^{n}(a_i+b_i)\boldsymbol{\varepsilon}_i=(\boldsymbol{\varepsilon}_1,\cdots,\boldsymbol{\varepsilon}_n)\begin{pmatrix}a_1+b_1\\ \vdots\\ a_n+b_n\end{pmatrix},$$

$$k\boldsymbol{\alpha}=\sum_{i=1}^{n}ka_i\boldsymbol{\varepsilon}_i=(\boldsymbol{\varepsilon}_1,\cdots,\boldsymbol{\varepsilon}_n)\begin{pmatrix}ka_1\\ \vdots\\ ka_n\end{pmatrix}。$$

二、基变换与坐标变换

引理 8.2 若线性无关向量组 $\boldsymbol{\alpha}_1,\cdots,\boldsymbol{\alpha}_s$ 能线性表出向量 $\boldsymbol{\beta}_1,\cdots,\boldsymbol{\beta}_s$，即 $\boldsymbol{\beta}_i=\sum_{j=1}^{s}b_{ij}\boldsymbol{\alpha}_j,i=1,\cdots,s,b_{ij}\in\boldsymbol{P}$，写成矩阵形式：

$$(\boldsymbol{\beta}_1,\cdots,\boldsymbol{\beta}_s)=(\boldsymbol{\alpha}_1,\cdots,\boldsymbol{\alpha}_s)\begin{pmatrix}b_{11}, & \cdots, & b_{s1}\\ \vdots & & \vdots\\ b_{1s}, & \cdots, & b_{ss}\end{pmatrix},$$

记 $B=\begin{pmatrix}b_{11} & \cdots & b_{s1}\\ \vdots & & \vdots\\ b_{1s} & \cdots & b_{ss}\end{pmatrix}$，则 $\boldsymbol{\beta}_1,\cdots\boldsymbol{\beta}_s$ 线性无关 $\Longleftrightarrow|\boldsymbol{B}|\neq0$，即矩阵 $\boldsymbol{B}$

可逆。

证 由于 $\alpha_1,\cdots,\alpha_s$ 线性无关，故

$$(\alpha_1,\cdots,\alpha_s)\begin{pmatrix}y_1\\ \vdots\\ y_s\end{pmatrix}=0\Longleftrightarrow\begin{pmatrix}y_1\\ \vdots\\ y_s\end{pmatrix}=\begin{pmatrix}0\\ \vdots\\ 0\end{pmatrix},$$

$$\beta_1,\cdots,\beta_s\text{ 线性无关 }\Longleftrightarrow\sum_{i=1}^{s}x_i\beta_i=(\beta_1,\cdots,\beta_s)\begin{pmatrix}x_1\\ \vdots\\ x_s\end{pmatrix}=(\alpha_1,\cdots,\alpha_s)$$

$$B\begin{pmatrix}x_1\\ \vdots\\ x_s\end{pmatrix}=0\text{ 仅有零解},\Longleftrightarrow B\begin{pmatrix}x_1\\ \vdots\\ x_s\end{pmatrix}=0\text{ 仅有零解 }\Longleftrightarrow|B|\neq 0。$$

于是，若 $\varepsilon_1,\cdots,\varepsilon_n$ 与 $\eta_1,\cdots,\eta_n$ 是线性空间 V 的两组基，且 $(\eta_1,\cdots,\eta_n)=(\varepsilon_1,\cdots,\varepsilon_n)T$，由引理 8.2，$T$ 为可逆阵，称 T 为由基 $\varepsilon_1,\cdots,\varepsilon_n$ 到基 $\eta_1,\cdots,\eta_n$ 的过渡矩阵。T 的 i 列 $T^{(i)}$ 是向量 η_i 在基 $\varepsilon_1,\cdots,\varepsilon_n$ 下的坐标。

例 8.13 给了 $V=P^3$ 的两组基 $\varepsilon_1=\begin{pmatrix}1\\0\\0\end{pmatrix},\varepsilon_2=\begin{pmatrix}1\\1\\0\end{pmatrix},\varepsilon_3=\begin{pmatrix}1\\1\\1\end{pmatrix}$；

$\eta_1=\begin{pmatrix}1\\0\\1\end{pmatrix},\eta_2=\begin{pmatrix}2\\1\\1\end{pmatrix},\eta_3=\begin{pmatrix}1\\3\\4\end{pmatrix}$，求 $\varepsilon_1,\varepsilon_2,\varepsilon_3$ 到 η_1,η_2,η_3 的过渡矩阵 T。

解 由 $(\eta_1,\eta_2,\eta_3)=(\varepsilon_1,\varepsilon_2,\varepsilon_3)T$ 得

$$\begin{pmatrix}1&1&1\\0&1&1\\0&0&1\end{pmatrix}T=\begin{pmatrix}1&2&1\\0&1&3\\1&1&4\end{pmatrix},$$

于是

$$T=\begin{pmatrix}1&1&1\\0&1&1\\0&0&1\end{pmatrix}^{-1}\begin{pmatrix}1&2&1\\0&1&3\\1&1&4\end{pmatrix}。$$

计算如下：

$$\begin{pmatrix}1&1&1&1&2&1\\0&1&1&0&1&3\\0&0&1&1&1&4\end{pmatrix}\xrightarrow[(2)-(3)]{(1)-(2)}\begin{pmatrix}1&0&0&1&1&-2\\0&1&0&-1&0&-1\\0&0&1&1&1&4\end{pmatrix},$$

故

$$T=\begin{pmatrix}1&1&-2\\-1&0&-1\\1&1&4\end{pmatrix}。$$

下面我们寻求同一个向量 α 在两组基下的坐标关系。

设

$$\alpha=(\varepsilon_1,\cdots,\varepsilon_n)\begin{pmatrix}x_1\\\vdots\\x_n\end{pmatrix}=(\eta_1,\cdots,\eta_n)\begin{pmatrix}y_1\\\vdots\\y_n\end{pmatrix},$$

由于

$$(\eta_1,\cdots,\eta_n)=(\varepsilon_1,\cdots,\varepsilon_n)T,$$

代入上式得

$$\alpha=(\varepsilon_1,\cdots,\varepsilon_n)T\begin{pmatrix}y_1\\\vdots\\y_n\end{pmatrix}。$$

这样，$\begin{pmatrix}x_1\\\vdots\\x_n\end{pmatrix}$ 与 $T\begin{pmatrix}y_1\\\vdots\\y_n\end{pmatrix}$ 都是向量 α 在基 $\varepsilon_1,\cdots,\varepsilon_n$ 下的坐标，由坐标的惟一性得

$$\begin{pmatrix}x_1\\\vdots\\x_n\end{pmatrix}=T\begin{pmatrix}y_1\\\vdots\\y_n\end{pmatrix}。$$

例 8.14 P^n 中，向量 $\alpha=\begin{pmatrix}a_1\\\vdots\\a_n\end{pmatrix}$ 在基

$$\varepsilon_1 = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \cdots, \varepsilon_n = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix} \text{下的坐标为} \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix}\text{；另选一组基}$$

$$\eta_1 = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \\ 0 \end{pmatrix}, \eta_2 = \begin{pmatrix} 1 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \cdots, \eta_n = \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \\ 1 \end{pmatrix},$$

即

$$(\eta_1, \cdots, \eta_n) = (\varepsilon_1, \cdots, \varepsilon_n) \begin{pmatrix} 1 & & 1 \\ & \ddots & \\ 0 & & \ddots \\ & & & 1 \end{pmatrix}。$$

则 α 在基 $\eta_1, \cdots, \eta_n$ 下的坐标为

$$\begin{pmatrix} 1 & \cdots & 1 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & 1 \end{pmatrix}^{-1} \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} = \begin{pmatrix} 1 & 1 & & 0 \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ & 0 & & 1 \end{pmatrix} \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} = \begin{pmatrix} a_1 - a_2 \\ a_2 - a_3 \\ \vdots \\ a_{n-1} - a_n \\ a_n \end{pmatrix}。$$

§3 子 空 间

一、定义

几何空间看成三维线性空间，易验证，通过原点的平面与直线分别成二维与一维的线性空间。现在我们研究原线性空间的这些特殊“子集合”及它们的相互关系。

定义 8.9 数域 P 上线性空间 V 的一个非空子集 W，如果对于 V

的两种运算(即加法与数乘)仍是数域 P 上的线性空间,则称为 V 的子空间。

下面我们分析 V 的非空子集 W 满足什么条件才能成为子空间。

设 $W \subseteq V$, V 是线性空间,故 W 中的向量(也是 V 的向量)对 V 的两种运算满足线性空间定义中 i)、ii)、v)、vi)、vii) 与 viii)。为了使 W 是线性空间,仅需 W 对两种运算封闭及 iii)、iv) 成立,即

(1) 若 $\alpha \in W, k \in P$ 必 $k\alpha \in W$;

(2) 若 $\alpha, \beta \in W$,必 $\alpha + \beta \in W$;

(3) $\boldsymbol{0} = 0\alpha \in W$;

(4) $\forall \alpha \in W$,必 $-\alpha = (-1)\alpha \in W$。

其中要求(3)、(4)已经包含在(1)中(分别令 $k = 0$ 与 -1 即得),故有

定理 8.2 如果线性空间 V 的非空子集 W 对于 V 的两种运算封闭(即满足上述条件(1)与(2)),则 W 是 V 的子空间。

推论 若 W 是线性空间 V 的非空子集,对任意 $k, l \in P, \alpha, \beta \in W$,均有 $k\alpha + l\beta \in W$,则 W 是 V 的子空间。

证 分别令 $k = l = 1$ 和 $l = 0$ 即得条件(2)与(1)。

例 8.15 $\{\boldsymbol{0}\}$ 与 V 是 V 的当然子空间,称为 V 的平凡子空间。$\{\boldsymbol{0}\}$ 称为零子空间。

例 8.16 $P[x]_n$ 是 $P[x]$ 的子空间。

例 8.17 若 A 是数域 P 上 $s \times n$ 阵,记 $W = \{\alpha \in P^n \mid A\alpha = 0\}$,则 W 是 $V = P^n$ 的子空间。W 称为齐次线性方程组 $AX = \boldsymbol{0}$ 的解空间。且 $\dim W = n - r(A)$。

对 t 用归纳法,易得下面的性质:

性质 若 W 是 V 的子空间,$\alpha_1, \cdots, \alpha_t \in W, k_1, \cdots, k_t \in P$,则

$$\sum_{i=1}^{t} k_i \alpha_i \in W。$$

二、交与和

定理 8.3 任意多个子空间的交仍是子空间。

即若 $W_i(i\in I)$ 是 V 的子空间，则 $W=\bigcap\limits_{i\in I}W_i$ 也是 V 的子空间。

证 若 $\alpha,\beta\in W,k\in P$，则 $\alpha,\beta\in W_i(i\in I)$，因为 W_i 是子空间，所以 $\alpha+\beta,k\alpha\in W_i(i\in I)$，从而 $\alpha+\beta,k\alpha\in W$。

由本章定理 8.2，W 是子空间。

定义 8.10 若 $W_1,\cdots,W_s$ 是 V 的子空间，记

$$W=\{\alpha_1+\cdots+\alpha_i+\cdots+\alpha_s\,|\,\alpha_i\in W_i,i=1,\cdots,s\},$$

则 W 是 V 的子空间，称它为 $W_1,\cdots,W_s$ 的和空间，记成 $W_1+\cdots+W_s$。

证 设 $\alpha=\alpha_1+\cdots+\alpha_s,\beta=\beta_1+\cdots+\beta_s,\alpha_i,\beta_i\in W_i,i=1,\cdots,s$；$\alpha,\beta\in W$。$k\in P$。则 $\alpha+\beta=(\alpha_1+\beta_1)+\cdots+(\alpha_s+\beta_s)$，

$$k\alpha=k\alpha_1+\cdots+k\alpha_s。$$

因为 W_i 是子空间，所以 $\alpha_i+\beta_i,k\alpha_i\in W_i,i=1,\cdots,s$，于是 $\alpha+\beta$，$k\alpha\in W$，由本章定理 8.2，W 是子空间。

定理 8.4 若 $\alpha_1,\cdots,\alpha_s$ 是 n 维线性空间 V 的一个线性无关向量组 $(s<n)$，则存在 V 中向量 $\alpha_{s+1},\cdots,\alpha_n$ 使 $\alpha_1,\cdots,\alpha_s,\cdots,\alpha_n$ 是 V 的一组基。

证 由于 $\dim V=n,s<n,\alpha_1,\cdots,\alpha_s$ 不是 V 的基，故存在 V 中向量 α_{s+1}，使 $\alpha_1,\cdots,\alpha_s,\alpha_{s+1}$ 线性无关，若 $s+1<n$，可继续这个过程。于是找到 V 中向量 $\alpha_{s+1},\cdots,\alpha_n$ 使 $\alpha_1,\cdots\alpha_s,\alpha_{s+1},\cdots,\alpha_n$ 线性无关，由命题 8.4，$\alpha_1,\cdots,\alpha_s,\cdots,\alpha_n$ 是 V 的一组基。

定理 8.5 若 W_1,W_2 是 V 的两个子空间，则

$$\dim W_1+\dim W_2=\dim(W_1\cap W_2)+\dim(W_1+W_2)。$$

证 $W_1\cap W_2\subseteq W_1,W_2\subseteq W_1+W_2$。设 $\varepsilon_1,\cdots,\varepsilon_t$ 为 $W_1\cap W_2$ 的一组基。由定理 8.4，可设 $\varepsilon_1,\cdots,\varepsilon_t,\alpha_1,\cdots,\alpha_s$ 为 W_1 的一组基，$\varepsilon_1,\cdots,\varepsilon_t,\beta_1,\cdots,\beta_l$ 为 W_2 的一组基。

此时 $\dim(W_1\cap W_2)=t,\dim W_1=t+s,\dim W_2=t+l$，证明本定理意味着证明 $\dim(W_1+W_2)=t+s+l$。

若能证明 $\varepsilon_1,\cdots,\varepsilon_t,\alpha_1,\cdots,\alpha_s,\beta_1,\cdots,\beta_l$ 是 W_1+W_2 的一组基，则结论成立。其证如下：

若 $\boldsymbol{\gamma}=\boldsymbol{\gamma}_1+\boldsymbol{\gamma}_2\in \boldsymbol{W}_1+\boldsymbol{W}_2, \boldsymbol{\gamma}_1\in \boldsymbol{W}_1, \boldsymbol{\gamma}_2\in \boldsymbol{W}$。于是

$$\boldsymbol{\varepsilon}_1,\cdots,\boldsymbol{\varepsilon}_t,\boldsymbol{\alpha}_1,\cdots,\boldsymbol{\alpha}_s\Rightarrow\boldsymbol{\gamma}_1,$$

即

$$\boldsymbol{\gamma}_1=\sum_{i=1}^{t}k_i\boldsymbol{\varepsilon}_i+\sum_{i=1}^{s}m_i\boldsymbol{\alpha}_i,$$

$$\boldsymbol{\varepsilon}_1,\cdots,\boldsymbol{\varepsilon}_t,\boldsymbol{\beta}_1,\cdots,\boldsymbol{\beta}_l\Rightarrow\boldsymbol{\gamma}_2,$$

即

$$\boldsymbol{\gamma}_2=\sum_{i=1}^{t}u_i\boldsymbol{\varepsilon}_i+\sum_{i=1}^{l}v_i\boldsymbol{\beta}_i,$$

因而

$$\boldsymbol{\gamma}=\boldsymbol{\gamma}_1+\boldsymbol{\gamma}_2=\sum_{i=1}^{t}(k_i+u_i)\boldsymbol{\varepsilon}_i+\sum_{i=1}^{s}m_i\boldsymbol{\alpha}_i+\sum_{i=1}^{l}v_i\boldsymbol{\beta}_i,$$

即

$$\boldsymbol{\varepsilon}_1,\cdots,\boldsymbol{\varepsilon}_t,\boldsymbol{\alpha}_1,\cdots,\boldsymbol{\alpha}_s,\boldsymbol{\beta}_1,\cdots,\boldsymbol{\beta}_l\Rightarrow\boldsymbol{\gamma},$$

所以

$$\boldsymbol{\varepsilon}_1,\cdots,\boldsymbol{\varepsilon}_t,\boldsymbol{\alpha}_1,\cdots,\boldsymbol{\alpha}_s,\boldsymbol{\beta}_1,\cdots,\boldsymbol{\beta}_l,\Rightarrow\boldsymbol{W}_1+\boldsymbol{W}_2。$$

再证 $\boldsymbol{\varepsilon}_1,\cdots,\boldsymbol{\varepsilon}_t,\boldsymbol{\alpha}_1,\cdots,\boldsymbol{\alpha}_s,\boldsymbol{\beta}_1,\cdots,\boldsymbol{\beta}_l$ 线性无关。若

$$x_1\boldsymbol{\varepsilon}_1+\cdots+x_t\boldsymbol{\varepsilon}_t+y_1\boldsymbol{\alpha}_1+\cdots+y_s\boldsymbol{\alpha}_s+z_1\boldsymbol{\beta}_1+\cdots+z_l\boldsymbol{\beta}_l=0, \tag{8.1}$$

令 $\delta=x_1\boldsymbol{\varepsilon}_1+\cdots+x_t\boldsymbol{\varepsilon}_t+y_1\boldsymbol{\alpha}_1+\cdots+y_s\boldsymbol{\alpha}_s$，则 $\delta\in W_1$，由式(8.1)得 $\delta=-z_1\boldsymbol{\beta}_1-z_2\boldsymbol{\beta}_2-\cdots-z_l\boldsymbol{\beta}_l$，则 $\boldsymbol{\delta}\in\boldsymbol{W}_2$。因而 $\boldsymbol{\delta}\in\boldsymbol{W}_1\cap\boldsymbol{W}_2$，故$\boldsymbol{W}_1\cap\boldsymbol{W}_2$ 的基 $\boldsymbol{\varepsilon}_1,\cdots,\boldsymbol{\varepsilon}_t\Rightarrow\boldsymbol{\delta}$，即存在数 $a_1,\cdots,a_t$ 使

$$\boldsymbol{\delta}=a_1\boldsymbol{\varepsilon}_1+\cdots+a_t\boldsymbol{\varepsilon}_t,$$

于是$(a_1\boldsymbol{\varepsilon}_1+\cdots+a_t\boldsymbol{\varepsilon}_t)+(z_1\boldsymbol{\beta}_1+\cdots+z_l\boldsymbol{\beta}_l)=\boldsymbol{\delta}-\boldsymbol{\delta}=0$，因为 $\boldsymbol{\varepsilon}_1,\cdots,\boldsymbol{\varepsilon}_t,\boldsymbol{\beta}_1,\cdots,\boldsymbol{\beta}_l$ 是 $\boldsymbol{W}_2$ 的一组线性无关基，故 $a_1=\cdots=a_t=z_1=\cdots=z_l=0$。

把 $z_1=\cdots=z_l=0$ 代入式(8.1)得

$$x_1\varepsilon_1+\cdots+x_t\varepsilon_t+y_1\alpha_1+\cdots+y_s\alpha_s=\boldsymbol{0}\ ,$$

因为 $\varepsilon_1,\cdots,\varepsilon_t,\alpha_1,\cdots,\alpha_s$ 是 W_1 的一组基，故 $x_1=\cdots=x_t=y_1=\cdots=y_l=0$。

这就证明了 $\varepsilon_1,\cdots,\varepsilon_t,\alpha_1,\cdots,\alpha_s,\beta_1,\cdots,\beta_l$ 线性无关，从而它们是 $\boldsymbol{W}_1+\boldsymbol{W}_2$ 的一组基，定理得证。

例 8.18 在线性空间 $\boldsymbol{P}^3$ 中，令

$$\boldsymbol{W}_1=\left\{\begin{pmatrix}x\\y\\0\end{pmatrix}\middle|\,x,y\in\boldsymbol{P}\right\},\boldsymbol{W}_2=\left\{\begin{pmatrix}0\\y\\z\end{pmatrix}\middle|\,y,z\in\boldsymbol{P}\right\},$$

则

$$\boldsymbol{W}_1\cap\boldsymbol{W}_2=\left\{\begin{pmatrix}0\\y\\0\end{pmatrix}\middle|\,y\in\boldsymbol{P}\right\},$$

$$\boldsymbol{W}_1+\boldsymbol{W}_2=\boldsymbol{P}^3。$$

此时：

$$\dim\boldsymbol{W}_1=\dim\boldsymbol{W}_2=2,\dim\boldsymbol{W}_1\cap\boldsymbol{W}_2=1,$$

$$\dim(\boldsymbol{W}_1+\boldsymbol{W}_2)=\dim\boldsymbol{P}^3=3\text{，符合定理的结论。}$$

三、子集生成的子空间

设 $\boldsymbol{S}$ 是线性空间 $\boldsymbol{V}$ 的一个子集。记

$$L(\boldsymbol{S})=\bigcap_{\substack{\boldsymbol{W}\supseteq\boldsymbol{S}\\ \boldsymbol{W}\text{ 是子空间}}}\boldsymbol{W},$$

则 (1) $L(\boldsymbol{S})\supseteq\boldsymbol{S}$；

(2) $L(\boldsymbol{S})$ 是子空间的交，故仍是子空间；

(3) 若 $\boldsymbol{W}\supseteq\boldsymbol{S},\boldsymbol{W}$ 是子空间，则 $\boldsymbol{W}\supseteq L(\boldsymbol{S})$。

若 $\boldsymbol{S}$ 是有限个向量组成的集合，$\boldsymbol{S}=\{\alpha_1,\cdots,\alpha_t\}$，简记 $L(\boldsymbol{S})$ 为 $L(\alpha_1,\cdots,\alpha_t)$，更有

(4) $L(\alpha_1,\cdots,\alpha_t)=\left\{\sum_{i=1}^{t}k_i\alpha_i \mid k_i\in P,i=1,\cdots,t\right\}$。

证(4)于下：

$\alpha_1,\cdots,\alpha_t\in$ 左边，左边是子空间，故 $\sum_{i=1}^{t}k_i\alpha_i\in$ 左边，即左边 $\supseteq$ 右边。

又 $\alpha_i=o\alpha_1+\cdots+1\cdot\alpha_i+\cdots+o\alpha_t\in$ 右边，易验证，右边对两种运算封闭，因而是 V 的包含 $\{\alpha_1,\cdots,\alpha_t\}$ 的子空间，由(3)，右边 $\supseteq$ 左边，故右边＝左边。

由上面的分析，$L(S)$是 V 的包含 S 的最小子空间，它包含在所有包含 S 的子空间内，称它为集合 S 生成的子空间。

命题 8.6 若 $W_1,\cdots,W_s$ 是 V 的子空间，则

$$L(W_1\cup W_2\cdots\cup W_s)=W_1+\cdots+W_s。$$

证 左边$\supseteq W_1,\cdots,W_s$，任意 $\alpha_i\in W_i$，必 $\alpha_i\in$ 左边，因为左边是子空间，故 $\alpha_1+\cdots+\alpha_s\in$ 左边，即左边 $\supseteq$ 右边。

右边$\supseteq W_1,\cdots,W_s$，右边$\supseteq W_1\cup\cdots\cup W_s$，又右边是和空间，即子空间，由(3)，右边$\supseteq$左边，故右边＝左边。

命题 8.7 $\dim L(\alpha_1,\cdots,\alpha_t)$＝向量组$\{\alpha_1,\cdots,\alpha_t\}$的秩。

证 由(4)可知，$\alpha_1,\cdots,\alpha_t$ 的一个极大线性无关组能线性表出 $L(\alpha_1,\cdots,\alpha_t)$中的任意向量，因而是它的一组基，即得本命题的证明。

命题 8.8 $L(\alpha_1,\cdots,\alpha_t)=L(\beta_1,\cdots,\beta_s)\Leftrightarrow\{\alpha_1,\cdots,\alpha_t\}\sim\{\beta_1,\cdots,\beta_s\}$，即$\{\alpha_1,\cdots,\alpha_t\}\Rightarrow\{\beta_1,\cdots,\beta_s\}$且$\{\beta_1,\cdots,\beta_s\}\Rightarrow\{\alpha_1,\cdots,\alpha_t\}$。

证 由(4)，$L(\alpha_1,\cdots,\alpha_t)=L(\beta_1,\cdots,\beta_s)\Leftrightarrow\alpha_1,\cdots,\alpha_t$ 的线性组合必是 $\beta_1,\cdots,\beta_s$ 的线性组合，反之 $\beta_1,\cdots,\beta_s$ 的线性组合也是 $\alpha_1,\cdots,\alpha_t$ 的线性组合$\Leftrightarrow\{\alpha_1,\cdots,\alpha_t\}\sim\{\beta_1,\cdots,\beta_s\}$。

例 8.19 若$(\beta_1,\cdots,\beta_s)=(\alpha_1,\cdots,\alpha_s)A$，这里 A 是 s 阶可逆阵，则

$$L(\alpha_1,\cdots,\alpha_s)=L(\beta_1,\cdots,\beta_s)。$$

证 因 A 有逆 A^{-1}，从而

$$(\alpha_1,\cdots,\alpha_s)=(\beta_1,\cdots,\beta_s)A^{-1},$$

即

$$\{\alpha_1,\cdots,\alpha_s\}\sim\{\beta_1,\cdots,\beta_s\},$$

由命题 8.8

$$L(\alpha_1,\cdots,\alpha_s)=L(\beta_1,\cdots,\beta_s)。$$

四、直和

定义 8.11 设 $\boldsymbol{W}_1,\boldsymbol{W}_2$ 是 V 的两个子空间，如果和空间 $\boldsymbol{W}_1+\boldsymbol{W}_2$ 中每个向量 α 的分解式

$$\alpha=\alpha_1+\alpha_2,\alpha_1\in\boldsymbol{W}_1,\alpha_2\in\boldsymbol{W}_2.$$

是惟一的(即若又有分解式 $\alpha=\beta_1+\beta_2,\beta_1\in\boldsymbol{W}_1,\beta_2\in\boldsymbol{W}_2$，必 $\alpha_1=\beta_1$，$\alpha_2=\beta_2$)，则称和 $\boldsymbol{W}_1+\boldsymbol{W}_2$ 为直和，记成 $W_1\oplus W_2$。

定理 8.6 若 $\boldsymbol{W}_1,\boldsymbol{W}_2$ 是 $\boldsymbol{V}$ 的两个子空间，则下面 4 个命题彼此等价：

(1) $\boldsymbol{W}_1+\boldsymbol{W}_2$ 是直和；

(2) 零向量分解式惟一。即若 $\boldsymbol{0}=\alpha_1+\alpha_2,\alpha_1\in\boldsymbol{W}_1,\alpha_2\in\boldsymbol{W}_2$，则 $\alpha_1=\alpha_2=\boldsymbol{0}$；

(3) $\boldsymbol{W}_1\cap\boldsymbol{W}_2=\{\boldsymbol{0}\}$；

(4) $\dim(\boldsymbol{W}_1+\boldsymbol{W}_2)=\dim\boldsymbol{W}_1+\dim\boldsymbol{W}_2$。

证 (1) $\Rightarrow$(2)：因为 $\boldsymbol{0}=\boldsymbol{0}+\boldsymbol{0}\in\boldsymbol{W}_1+\boldsymbol{W}_2$。

(2)$\Rightarrow$(3)：$0\in\boldsymbol{W}_1,\boldsymbol{W}_2$，故 $\{\boldsymbol{0}\}\subseteq\boldsymbol{W}_1\cap\boldsymbol{W}_2$。

若 $\alpha\in\boldsymbol{W}_1\cap\boldsymbol{W}_2,\alpha\in\boldsymbol{W}_1,\alpha\in\boldsymbol{W}_2$ 因为 $\boldsymbol{W}_2$ 是子空间，所以 $-\alpha\in\boldsymbol{W}_2$。$\boldsymbol{0}=\alpha+(-\alpha)$，由零向量分解式惟一，故 $\alpha=\boldsymbol{0}$，即 $\boldsymbol{W}_1\cap\boldsymbol{W}_2=\{\boldsymbol{0}\}$。

(3) $\Rightarrow$(1)：若 α 有两种分解式 $\alpha=\alpha_1+\alpha_2=\beta_1+\beta_2$；$\alpha_1,\beta_1\in\boldsymbol{W}_1,\alpha_2,\beta_2\in\boldsymbol{W}_2$，则

$$\alpha_1-\beta_1=\beta_2-\alpha_2\in\boldsymbol{W}_1\cap\boldsymbol{W}_2=\{0\},$$

故 $\alpha_1-\beta_1=\beta_2-\alpha_2=0$，即 $\alpha_1=\beta_1,\alpha_2=\beta_2,\boldsymbol{W}_1+\boldsymbol{W}_2$ 是直和。

(3)$\Leftrightarrow$(4)：$W_1 \cap W_2 = \{0\} \Leftrightarrow \dim(W_1 \cap W_2) = 0 \xLeftrightarrow{\text{由定理8.5}} \dim(W_1 + W_2) = \dim W_1 + \dim W_2$。

例 8.20 在 $V = P^{n\times n}$ 中，记

$$W_1 = \{kE_n \mid k \in P\},$$

$$W_2 = \left\{A = (a_{ij})_{n\times n} \in V \mid \sum_{i=1}^{n} a_{ii} = 0\right\},$$

求证：$V = W_1 \oplus W_2$。

证 易验证 W_1 与 W_2 是 V 的两个子空间。

(1) 若 $A \in W_1 \cap W_2$，因 $A \in W_1$，故 $A = kE_n, k \in P$；又 $A \in W_2$，A 对角线上元素和 $k + \cdots + k = nk = 0$，必 $k = 0$，从而 $A = 0$，$W_1 \cap W_2 = \{0\}$，由定理 8.6，$W_1 + W_2$ 是直和。

(2) 若 $A = (a_{ij})_{n\times n} \in V$，记 $k = \dfrac{1}{n}\sum_{i=1}^{n} a_{ii}$，则

$$A = kE + (A - kE),$$

$kE \in W_1$，$A - kE$ 对角线上元素和为 $\sum_{i=1}^{n} a_{ii} - nk = 0$，即

$$A - kE \in W_2,$$

于是

$$A \in W_1 \oplus W_2,$$

故

$$V \subseteq W_1 \oplus W_2,\text{从而 } V = W_1 \oplus W_2。$$

定理 8.7 若 W_1 是 V 的子空间，则存在 V 的子空间 W_2 使 $V = W_1 \oplus W_2$。

证 取 W_1 一组基 $\alpha_1, \cdots, \alpha_s$，则存在 $\alpha_{s+1}, \cdots, \alpha_n$ 使 $\alpha_1, \cdots, \alpha_s, \alpha_{s+1}, \cdots, \alpha_n$ 为 V 的一组基。令 $W_2 = L(\alpha_{s+1}, \cdots, \alpha_n)$ 即满足要求。

子空间直和的概念可以推广到多个子空间的情形。

定义 8.12 设 $W_1, \cdots, W_s$ 都是 V 的子空间，如果和空间 $W_1 + \cdots + W_s$ 中每个向量 α 的分解式

$$\alpha = \alpha_1 + \cdots + \alpha_s, \alpha_i \in \boldsymbol{W}, i = 1, \cdots, s$$

是惟一的，则称和 $\boldsymbol{W}_1 + \cdots + \boldsymbol{W}_s$ 为直和，记成 $\boldsymbol{W}_1 \oplus \cdots \oplus \boldsymbol{W}_s$。

与两个子空间直和的情形类似，我们有下面的定理（证明略）。

定理 8.8 若 $\boldsymbol{W}_1, \cdots, \boldsymbol{W}_s$ 是 $\boldsymbol{V}$ 的子空间，则下面 4 个命题是等价的：

(1) $\boldsymbol{W}_1 + \cdots + \boldsymbol{W}_s$ 是直和；

(2) 零向量分解惟一；

(3) $\boldsymbol{W}_i \cap \sum_{j \neq i} \boldsymbol{W}_j = \{\boldsymbol{0}\}, i = 1, \cdots, s$；

(4) $\sum_{i=1} \dim \boldsymbol{W}_i = \dim(\boldsymbol{W}_1 + \cdots + \boldsymbol{W}_s)$。

§4 内积空间

一、向量的内积与长度

定义 8.13 数域 P 上线性空间 $\boldsymbol{V}$ 叫内积空间，如果 $\boldsymbol{V}$ 中任意两个向量 α, β 都对应到 P 中一个确定的数（记成 $[\alpha, \beta]$，称为 α 与 β 的内积），且满足

(1) $[\beta, \alpha] = \overline{[\alpha, \beta]}, \alpha, \beta \in \boldsymbol{V}$；

(2) $[\alpha, k_1\beta_1 + k_2\beta_2] = k_1[\alpha, \beta_1] + k_2[\alpha, \beta_2], k_1, k_2 \in P, \alpha, \beta_1, \beta_2 \in V$；

(3) $[\alpha, \alpha]$ 为非负实数，且 $[\alpha, \alpha] = 0 \Longleftrightarrow \alpha = 0, \alpha \in \boldsymbol{V}$。

复数域 $\mathbf{C}$ 上的内积空间常称酉空间，实数域 $\mathbf{R}$ 上的内积空间常称 Euclid 空间或欧氏空间，此时(1)改为，$[\beta, \alpha] = [\alpha, \beta]$。

例 8.21 $\mathbf{R}$ 上线性空间 $\boldsymbol{R}^n$ 关于向量的内积

$$[\alpha, \beta] = \alpha' \cdot \beta$$

成一内积空间。

例 8.22 $\mathbf{C}$ 上线性空间 $\boldsymbol{C}^n$ 关于向量的内积

$$[\alpha, \beta] = \bar{\alpha}' \beta$$

成一内积空间。

例 8.23 R 上线性空间 $\boldsymbol{R}[x]_n$ 关于内积

$$\left[\sum_{i=0}^{n-1}\boldsymbol{a}_i x^i,\ \sum_{i=0}^{n-1}\boldsymbol{b}_i x^i\right]=\sum_{i=0}^{n-1}\boldsymbol{a}_i\boldsymbol{b}_i$$

成一内积空间。

更有 $\boldsymbol{R}[x]_n$ 关于内积

$$[\boldsymbol{f}(x),\boldsymbol{g}(x)]=\sum_{i=1}^{n}\boldsymbol{f}(i)\boldsymbol{g}(i)$$

也成一内积空间。

性质 (1) $[0,\boldsymbol{\beta}]=[\boldsymbol{\alpha},0]=0$；

(2) $\left[\boldsymbol{\alpha},\sum_{i=1}^{s}k_i\boldsymbol{\beta}_i\right]=\sum_{i=1}^{s}k_i[\boldsymbol{\alpha},\boldsymbol{\beta}_i]$；

(3) $\left[\sum_{i=1}^{s}l_i\boldsymbol{\alpha}_i,\boldsymbol{\beta}\right]=\sum_{i=1}^{s}\bar{l}_i[\boldsymbol{\alpha},\boldsymbol{\beta}]$；

(4) $$\left[\sum_{i=1}^{s}l_i\boldsymbol{\alpha}_i,\sum_{j=1}^{t}k_j\boldsymbol{\beta}_j\right]=\sum_{i=1}^{s}\sum_{j=1}^{t}\bar{l}_i k_j[\boldsymbol{\alpha}_i,\boldsymbol{\beta}_j]$$

$$=(\bar{l}_1,\cdots,\bar{l}_s)([\boldsymbol{\alpha}_i,\boldsymbol{\beta}_j])_{s\times t}\begin{pmatrix}k_1\\ \vdots\\ k_t\end{pmatrix}。$$

证 (1) $[0,\boldsymbol{\beta}]=[0\boldsymbol{\alpha},\boldsymbol{\beta}]=0[\boldsymbol{\alpha},\boldsymbol{\beta}]=0$，同理 $[\boldsymbol{\alpha},0]=0$。

(2) 对 s 用归纳法即得。

(3) $\left[\sum_{i=1}^{s}l_i\boldsymbol{\alpha}_i,\boldsymbol{\beta}\right]=\overline{\left[\boldsymbol{\beta},\sum_{i=1}^{s}l_i\boldsymbol{\alpha}_i\right]}\xlongequal{\text{由(2)}}\overline{\sum_{i=1}^{s}l_i[\boldsymbol{\beta},\boldsymbol{\alpha}_i]}=\sum_{i=1}^{s}\bar{l}_i\overline{[\boldsymbol{\beta},\boldsymbol{\alpha}_i]}=\sum_{i=1}^{s}\bar{l}_i[\boldsymbol{\alpha}_i,\boldsymbol{\beta}]$。

(4) $\left[\sum_{i=1}^{s}l_i\boldsymbol{\alpha}_i,\sum_{j=1}^{t}k_j\boldsymbol{\beta}_j\right]\xlongequal{\text{由(3)}}\sum_{i=1}^{s}\bar{l}_i\left[\boldsymbol{\alpha}_i,\sum_{j=1}^{t}k_j\boldsymbol{\beta}_j\right]\xlongequal{\text{由(2)}}\sum_{i=1}^{s}\bar{l}_i\left(\sum_{j=1}^{t}k_j[\boldsymbol{\alpha}_i,\boldsymbol{\beta}_j]\right)=\sum_{i=1}^{s}\sum_{j=1}^{t}\bar{l}_i k_j[\boldsymbol{\alpha}_i,\boldsymbol{\beta}_j]$。

定义 8.14 称 $\sqrt{[\boldsymbol{\alpha},\boldsymbol{\alpha}]}$ 为内积空间的向量 $\boldsymbol{\alpha}$ 的长度，记为 $\|\boldsymbol{\alpha}\|$。

性质 (1) $\|\alpha\|$ 为非负实数，且 $\|\alpha\| = 0 \Leftrightarrow \alpha = 0$；

(2) $\|k\alpha\| = |k| \cdot \|\alpha\|$。

给了内积空间的一组基 $\varepsilon_1,\cdots,\varepsilon_n$。若

$$\alpha = (\varepsilon_1,\cdots,\varepsilon_n)\begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix}, \beta = (\varepsilon_1,\cdots,\varepsilon_n)\begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix},$$

由内积性质(4)

$$[\alpha,\beta] = (\overline{a_1},\cdots,\overline{a_n})([\varepsilon_i,\varepsilon_j])_{n\times n}\begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix}。 \tag{8.2}$$

称 n 阶方阵 $([\varepsilon_i,\varepsilon_j])_{n\times n} = A$ 为基 $\varepsilon_1,\cdots,\varepsilon_n$ 的度量矩阵。由上面推理可知，给了一组基的度量矩阵后，任意两个向量的内积可以把坐标代入式(8.2)算得，即内积由度量矩阵确定。

设 $\eta_1,\cdots,\eta_n$ 是 V 的另一组基，过渡矩阵为 T：

$$(\eta_1,\cdots,\eta_n) = (\varepsilon_1,\cdots,\varepsilon_n)T。$$

记 $([\eta_i,\eta_j])_{n\times n} = B$。我们来寻求两组不同基下的度量矩阵 A 与 B 的关系。

因为 η_i,η_j 在基 $\varepsilon_1,\cdots,\varepsilon_n$ 下的坐标分别为 $T^{(i)}$ 与 $T^{(j)}$，于是

$$B(i,j) = [\eta_i,\eta_j] = (\overline{T^{(i)}})'AT^{(j)},$$

而

$$(T^{\star}AT)(i,j) = T^{\star}_{(i)}AT^{(j)} = (\overline{T}^{(i)})'AT^{(j)} = B(i,j),$$

故

$$B = T^{\star}AT。$$

这是不同基下度量矩阵的关系。

二、标准正交基

定义 8.15 若 $[\alpha,\beta] = 0$，则称向量 α 与 β 正交。

定义 8.16 两两正交的非零向量组称为正交组。

定理 8.9 正交组线性无关。

证明见命题 5.9 的证明。

定义 8.17 由单位向量(长度为 1 的向量)组成的正交组称为标准正交组。若一个标准正交组是基,则称标准正交基。

推论 $\varepsilon_1,\cdots,\varepsilon_n$ 是标准正交基 $\Leftrightarrow[\varepsilon_i,\varepsilon_j]=\delta_{ij},i,j=1,\cdots,n$。$\Leftrightarrow$ 度量矩阵$([\varepsilon_i,\varepsilon_j])_{n\times n}=\boldsymbol{E}_n$。

若 $\varepsilon_1,\cdots,\varepsilon_n$ 是标准正交基,α 与 β 在 $\varepsilon_1,\cdots,\varepsilon_n$ 下的坐标分别为 $\begin{pmatrix}a_1\\ \vdots\\ a_n\end{pmatrix}$ 与 $\begin{pmatrix}b_1\\ \vdots\\ b_n\end{pmatrix}$,则 $[\alpha,\beta]=(\overline{a_1},\cdots,\overline{a_n})\cdot\boldsymbol{E}\begin{pmatrix}b_1\\ \vdots\\ b_n\end{pmatrix}=\sum\limits_{i=1}^{n}\overline{a_i}b_i$,恰是 n 维向量的内积形式。

正交化方法 若 $\alpha_1,\cdots,\alpha_s$ 为线性无关向量组,令 $\varepsilon_1=\alpha_1$, $\varepsilon_i=\alpha_i-\sum\limits_{j=1}^{i-1}\dfrac{[\varepsilon_j,\alpha_i]}{[\varepsilon_j,\varepsilon_j]}\varepsilon_j,i=2,\cdots,s$, 则,$\varepsilon_1,\cdots,\varepsilon_s$ 是与 $\alpha_1,\cdots,\alpha_s$ 等价的正交组。

推论 1 正交组必可扩充为正交基。

推论 2 标准正交基是存在的。

若 $\varepsilon_1,\cdots,\varepsilon_n$ 是一组标准正交基,向量 α 在这组基下的坐标为 $\begin{pmatrix}a_1\\ \vdots\\ a_n\end{pmatrix}$,即 $\alpha=\sum\limits_{i=1}^{n}a_i\varepsilon_i$, 则 $[\varepsilon_i,\alpha]=a_i[\varepsilon_i,\varepsilon_i]=a_i$。

若 $\varepsilon_1,\cdots,\varepsilon_n$ 与 $\eta_1,\cdots,\eta_n$ 是两组标准正交基,且

$$(\eta_1,\cdots,\eta_n)=(\varepsilon_1,\cdots,\varepsilon_n)\boldsymbol{T},$$

则

$$([\eta_i,\eta_j])_{n\times n}=\boldsymbol{T}^{\star}([\varepsilon,\varepsilon_j])_{n\times n}\boldsymbol{T}。$$

从而 $\boldsymbol{T}^{\star}\boldsymbol{T}=\boldsymbol{E}$, 即过渡矩阵为酉阵,反之也成立。

定理 8.10 标准正交基到标准正交基的过渡矩阵是酉阵。反之,若第一组基是标准正交基,过渡矩阵是酉阵,则第二组基也是标准正交基。

三、正交基下向量的坐标与内积

若给了内积空间 V 的一组正交基 $\varepsilon_1,\cdots,\varepsilon_n$，向量 α 在基 $\varepsilon_1,\cdots,\varepsilon_n$ 下的坐标为 $\begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix}$，即

$$\alpha = \sum_{i=1}^{n} a_i \varepsilon_i,$$

由 $[\varepsilon_i,\varepsilon_j]=\delta_{ij}$ 得

$$[\varepsilon_i,\alpha] = a_i[\varepsilon_i,\varepsilon_i],$$

从而坐标

$$a_i = \frac{[\varepsilon_i,\alpha]}{[\varepsilon_i,\varepsilon_i]}。$$

若向量 β 在正交基 $\varepsilon_1,\cdots,\varepsilon_n$ 下坐标为 $\begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix}$，因为正交基 $\varepsilon_1,\cdots,\varepsilon_n$ 的度量矩阵为对角阵 $\begin{pmatrix} [\varepsilon_1,\varepsilon_1] & & \\ & \ddots & \\ & & [\varepsilon_n,\varepsilon_n] \end{pmatrix}$，于是

$$\begin{aligned}[\alpha,\beta] &= (\bar{a}_1,\cdots,\bar{a}_n)\begin{pmatrix} [\varepsilon_1,\varepsilon_1] & & \\ & \ddots & \\ & & [\varepsilon_n,\varepsilon_n] \end{pmatrix}\begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix} \\ &= \sum_{i=1}^{n} \bar{a}_i b_i [\varepsilon_i,\varepsilon_i]。\end{aligned}$$

若 $\varepsilon_1,\cdots,\varepsilon_n$ 为标准正交基，则向量 α 在基 $\varepsilon_1,\cdots,\varepsilon_n$ 下的坐标为

$$\begin{pmatrix} [\varepsilon_1,\alpha] \\ \vdots \\ [\varepsilon_n,\alpha] \end{pmatrix},$$

向量 α 与 β 的内积为

$$[\alpha,\beta] = (\bar{a}_1,\cdots,\bar{a}_n)\begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix} = \sum_{i=1}^{n} \bar{a}_i b_i = \sum_{i=1}^{n} \overline{[\varepsilon_i,\alpha]}[\varepsilon_i,\beta]。$$

例 8.24 给了欧氏空间 $V=R^3$ 的一组正交基 $\varepsilon_1,\varepsilon_2,\varepsilon_3$:

$$\varepsilon_1 = \begin{pmatrix} 1 \\ 2 \\ 2 \end{pmatrix},\varepsilon_2 = \begin{pmatrix} 2 \\ 1 \\ -2 \end{pmatrix},\varepsilon_3 = \begin{pmatrix} 2 \\ -2 \\ 1 \end{pmatrix},$$

求向量 $\alpha = \begin{pmatrix} 2 \\ 5 \\ 6 \end{pmatrix}$ 在这组基下的坐标。

解 $[\varepsilon_1,\alpha] = \left[\begin{pmatrix} 2 \\ 5 \\ 6 \end{pmatrix},\begin{pmatrix} 1 \\ 2 \\ 2 \end{pmatrix}\right] = 2+10+12=24,$

$$[\varepsilon_2,\alpha] = \left[\begin{pmatrix} 2 \\ 5 \\ 6 \end{pmatrix},\begin{pmatrix} 2 \\ 1 \\ -2 \end{pmatrix}\right] = 4+5-12=-3,$$

$$[\varepsilon_3,\alpha] = \left[\begin{pmatrix} 2 \\ 5 \\ 6 \end{pmatrix},\begin{pmatrix} 2 \\ -2 \\ 1 \end{pmatrix}\right] = 4-10+6=0,$$

又

$$[\varepsilon_1,\varepsilon_1] = [\varepsilon_2,\varepsilon_2] = [\varepsilon_3,\varepsilon_3] = 9,$$

故向量 α 在基 $\varepsilon_1,\varepsilon_2,\varepsilon_3$ 下的坐标为 $\begin{pmatrix} \frac{24}{9} \\ -\frac{3}{9} \\ \frac{0}{9} \end{pmatrix} = \frac{1}{3}\begin{pmatrix} 8 \\ -1 \\ 0 \end{pmatrix}$。

四、正交补

若 W 为内积空间 V 的一个子空间。记

$$W^{\perp}=\{\alpha\in V \mid \text{任意 }\beta\in W,[\alpha,\beta]=0\},$$

称为 W 的正交补。

显然

$$\{\mathbf{0}\}^{\perp}=V, V^{\perp}=\{\mathbf{0}\}, (W^{\perp})^{\perp}=W。$$

易验证，$W^{\perp}$是 V 的子空间。

欧氏空间 $V=R^3$ 中，若

$$W=\left\{\begin{pmatrix}x\\y\\0\end{pmatrix} \middle| x,y\in\mathbf{R}\right\},$$

则

$$W^{\perp}=\left\{\begin{pmatrix}0\\0\\z\end{pmatrix} \middle| z\in\mathbf{R}\right\}。$$

定理 8.11 设 W 为 V 的任一子空间，则 $V=W\oplus W^{\perp}$。

证 (1) 若 $\alpha\in W\cap W^{\perp}$，则 $\alpha\in W$ 且 $\alpha\in W^{\perp}$，从而 $[\alpha,\alpha]=0$，$\alpha=0$，$W\cap W^{\perp}=\{\mathbf{0}\}$，故和 $W+W^{\perp}$ 是直和。

(2) 设 $\alpha_1,\cdots,\alpha_s$ 为 W 的正交基，扩充为 V 的正交基 $\alpha_1,\cdots,\alpha_s,\alpha_{s+1},\cdots,\alpha_n$。

$$W=L(\alpha_1,\cdots,\alpha_s), L(\alpha_{s+1},\cdots,\alpha_n)\subseteq W^{\perp},$$

于是

$$\begin{aligned}W+W^{\perp}&\supseteq L(\alpha_1,\cdots,\alpha_s)+L(\alpha_{s+1},\cdots,\alpha_n)\\&=L(\alpha_1,\cdots,\alpha_s,\cdots,\alpha_n)=V,\end{aligned}$$

所以

$$W\oplus W^{\perp}=V。$$

命题 8.9 若 $A\in R^{s\times n}$，则齐次线性方程组 $AX=\mathbf{0}$ 的解空间 $W=\{\alpha\in R^n \mid A\alpha=\mathbf{0}\}$ 作为欧氏空间 $V=R^n$ 的子空间，其正交补为系数矩阵 A 的行空间——由 A 的行向量生成的空间，即

$$W^{\perp} = L(A'_{(1)}, \cdots, A'_{(s)}),$$
$$V = R^n = W \oplus W^{\perp}。$$

证 $[L(A'_{(1)}, \cdots, A'_{(s)})]^{\perp} = \{\alpha \in R^n \mid [A'_{(i)}, \alpha] = 0, i = 1, \cdots, s\}$
$= \{\alpha \in R^n \mid A_{(i)} \cdot \alpha = 0, i = 1, \cdots, s\}$
$= \{\alpha \in R^n \mid A\alpha = 0\} = W,$

从而

$$W^{\perp} = L(A'_{(1)}, \cdots, A'_{(s)})。$$

命题 8.10 内积空间的子空间(在某种意义下,详见证明)必是某齐次线性方程组的解空间。

证 设 W 是内积空间 V 的一个子空间,$\varepsilon_1, \cdots, \varepsilon_t$ 为 $W^{\perp}$ 的一组基,则

$$W = \{\alpha \in V \mid [\varepsilon_i, \alpha] = 0, i = 1, \cdots, t\},$$

于是,W 是方程组

$$\begin{cases} [\varepsilon_1, \alpha] = 0 \\ \quad\vdots \qquad\quad \vdots \\ \quad\vdots \qquad\quad \vdots \\ [\varepsilon_t, \alpha] = 0 \end{cases}$$

的解空间。

若 $\eta_1, \cdots, \eta_n$ 为 V 的一组基,$\alpha = \sum\limits_{i=1}^{n} x_i \eta_i$,则上述方程组化为同解的齐次线性方程组:

$$\begin{cases} [\varepsilon_1, \eta_1]x_1 + \cdots + [\varepsilon_1, \eta_n]x_n = 0, \\ \quad\vdots \qquad\qquad\qquad\quad \vdots \qquad\quad \vdots \\ \quad\vdots \qquad\qquad\qquad\quad \vdots \qquad\quad \vdots \\ [\varepsilon_t, \eta_1]x_1 + \cdots + [\varepsilon_t, \eta_n]x_n = 0。 \end{cases}$$

而

$$W_1 = \left\{ \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \middle| (\eta_1, \cdots, \eta_n) \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \alpha \in W \right\}$$

为其解空间，即 W 中向量的坐标所生成的线性空间 W_1（下面将指出它与 W 同构）是上述齐次线性方程组的解空间。

若

$$V = W \oplus W^{\perp},$$
$$\alpha=\alpha_1+\alpha_2, \quad \alpha_1\in W, \alpha_2\in W^{\perp},$$

则称 α_1 为 α 在子空间 W 上的正交投影。

如果 W 有一组标准正交基 $\varepsilon_1,\cdots,\varepsilon_t$，扩充为 V 的一组标准正交基

$$\varepsilon_1,\ \cdots,\ \varepsilon_t,\ \varepsilon_{t+1},\cdots,\ \varepsilon_n,$$

则

$$W = L(\varepsilon_1,\ \cdots,\ \varepsilon)_t;\ W^{\perp}= L(\varepsilon_{t+1},\ \cdots,\ \varepsilon_n)。$$

向量

$$\alpha=\frac{[\varepsilon_1,\alpha]}{[\varepsilon_1,\varepsilon_1]}\varepsilon_1+\cdots+\frac{[\varepsilon_t,\alpha]}{[\varepsilon_t,\varepsilon_t]}\varepsilon_t+\cdots+\frac{[\varepsilon_n,\alpha]}{[\varepsilon_n,\varepsilon_n]}\varepsilon_n,$$

于是 α 在 W 上的正交投影为

$$\alpha_1 = \frac{[\varepsilon_1,\alpha]}{[\varepsilon_1,\varepsilon_1]}\varepsilon_1 + \cdots + \frac{[\varepsilon_t,\alpha]}{[\varepsilon_t,\varepsilon_t]}\varepsilon_t。$$

例 8.25 欧氏空间 $V = \boldsymbol{R}^4$，求向量 $\alpha = \begin{pmatrix} 7\\ -4\\ -1\\ 2\end{pmatrix}$ 在 V 的子空间 W 上的正交投影 α_1。这里 W 是下列齐次线性方程组的解空间：

$$\begin{cases} 2x_1+x_2+x_3+3x_4=0, \\ 3x_1+2x_2+2x_3+x_4=0, \\ x_1+2x_2+2x_3-9x_4=0。\end{cases}$$

解 $W^{\perp}=$ 系数矩阵的行空间 $=L\left(\begin{pmatrix}2\\1\\1\\3\end{pmatrix},\begin{pmatrix}3\\2\\2\\1\end{pmatrix},\begin{pmatrix}1\\2\\2\\-9\end{pmatrix}\right)$。用正交化

分法，得到与$\begin{pmatrix}2\\1\\1\\3\end{pmatrix}$，$\begin{pmatrix}3\\2\\2\\1\end{pmatrix}$，$\begin{pmatrix}1\\2\\2\\-9\end{pmatrix}$等价的正交组：

$$\varepsilon_1=\begin{pmatrix}2\\1\\1\\3\end{pmatrix},\ \varepsilon_2=\begin{pmatrix}19\\17\\17\\-24\end{pmatrix},$$

因而 α 在 $W^{\perp}$ 上的正交投影为

$$\begin{aligned}\alpha_2&=\frac{[\varepsilon_1,\alpha]}{[\varepsilon_1,\varepsilon_1]}\varepsilon_1+\frac{[\varepsilon_2,\alpha]}{[\varepsilon_2,\varepsilon_2]}\varepsilon_2\\&=\frac{15}{15}\begin{pmatrix}2\\1\\1\\3\end{pmatrix}+\frac{0}{1\,515}\begin{pmatrix}19\\17\\17\\-24\end{pmatrix}=\begin{pmatrix}2\\1\\1\\3\end{pmatrix}。\end{aligned}$$

于是 α 在 W 上正交投影为

$$\alpha_1=\alpha-\alpha_2=\begin{pmatrix}7\\-4\\-1\\2\end{pmatrix}-\begin{pmatrix}2\\1\\1\\3\end{pmatrix}=\begin{pmatrix}5\\-5\\-2\\-1\end{pmatrix}。$$

§5 同态与同构

一、定义

定义 8.18 设 V_1 与 V_2 是数域 P 上两个线性空间，映射 $V_1\xrightarrow{\sigma}V_2$ 称为线性空间的同态映射，如果

(1) 对任意 $\alpha,\beta\in V_1$，必 $\sigma(\alpha+\beta)=\sigma(\alpha)+\sigma(\beta)$；

(2) 对任意 $k\in P$，$\alpha\in V_1$，必 $\sigma(k\alpha)=k\sigma(\alpha)$。

若 σ 是 V_1 到 V_2 上的是双射，则称为同构映射。

常用"$\backsim$"表示同态，用"$\cong$"表示同构。

推论 若 σ 是数域 P 上线性空间 V_1 到数域 P 上线性空间 V_2 的映射，满足：

任意 $k,l \in P,\alpha,\beta \in V$，必有

$$\sigma(k\alpha + l\beta) = k\sigma(\alpha) + l\sigma(\beta),$$

则 σ 是同态映射，即 $V_1 \overset{\sigma}{\backsim} V_2$。

证 令 $k=l=1$ 得

$$\sigma(\alpha + \beta) = \sigma(\alpha) + \sigma(\beta),$$

令 $l = 0$ 得

$$\sigma(k\alpha) = k\sigma(\alpha),$$

故 σ 是同态映射。

定义 8.19 若存在线性空间V_1 到线性空间V_2 上的同构映射 σ，则称线性空间V_1 与线性空间V_2 同构，记成 $V_1 \overset{\sigma}{\cong} V_2$。

二、性质

同态的性质 若 $V_1 \overset{\sigma}{\backsim} V_2$，则

(1) $\sigma(O_1) = O_2$. 这里 O_1 与 O_2 分别表示V_1与V_2中的零向量；

(2) 任意 $\alpha \in V_1$，必 $\sigma(-\alpha) = -\sigma(\alpha)$；

(3) 对任意 $k_1,\cdots,k_t \in P,\alpha_1,\cdots,\alpha_t \in V$，必

$$\sigma\left(\sum_{i=1}^{t} k_i\alpha_i\right) = \sum_{i=1}^{t} k_i\sigma(\alpha_i);$$

(4) 若V_1 中向量$\alpha_1,\cdots,\alpha_t$ 线性相关，则V_2中向量 $\sigma(\alpha_1),\cdots,\sigma(\alpha_t)$ 也线性相关；

(5) 若W_1 是V_1 的子空间，则

$$\sigma(W_1) = \{\sigma(\alpha) \mid \alpha \in W_1\}$$

是V_2 的子空间。

证 (1) $\sigma(\boldsymbol{O}_1)=\sigma(\boldsymbol{O}\cdot\boldsymbol{\alpha})=O\cdot\sigma(\boldsymbol{\alpha})=\boldsymbol{O}_2$。

(2) $\sigma(-\boldsymbol{\alpha})=\sigma[(-1)\boldsymbol{\alpha}]=(-1)\sigma(\boldsymbol{\alpha})=-\sigma(\boldsymbol{\alpha})$。

(3) 对 t 用归纳法即得。

(4) 因为 $\boldsymbol{\alpha}_1,\cdots,\boldsymbol{\alpha}_t$ 线性相关，所以存在不全为 0 的 $k_1,\cdots,k_t\in\boldsymbol{P}$，使

$$\sum_{i=1}^{t}k_i\boldsymbol{\alpha}_i=\boldsymbol{O}_1,$$

于是

$$\boldsymbol{O}_2=\sigma(\boldsymbol{O}_1)=\sigma\left(\sum_{i=1}^{t}k_i\boldsymbol{\alpha}_i\right)=\sum_{i=1}^{t}k_i\sigma(\boldsymbol{\alpha}_i),$$

即

$\sigma(\boldsymbol{\alpha}_1),\cdots,\sigma(\boldsymbol{\alpha}_t)$ 线性相关。

(5) $\sigma(\boldsymbol{\alpha}),\sigma(\boldsymbol{\beta})\in\sigma(\boldsymbol{W}_1)$，即 $\boldsymbol{\alpha},\boldsymbol{\beta}\in\boldsymbol{W}_1$，因为$\boldsymbol{W}_1$是$\boldsymbol{V}_1$子空间，故对任意 $k_1l\in\boldsymbol{P}$，必 $k\boldsymbol{\alpha}+l\boldsymbol{\beta}\in\boldsymbol{W}_1$，于是

$$k\sigma(\boldsymbol{\alpha})+l\sigma(\boldsymbol{\beta})=\sigma(k\boldsymbol{\alpha}+l\boldsymbol{\beta})\in\sigma(\boldsymbol{W}_1),$$

所以 $\sigma(\boldsymbol{W}_1)$是 $\boldsymbol{V}_2$ 的子空间。

同构的性质 同构的线性空间，除了它们的元素记号不同，就两种运算（加法与数乘）而言是完全相同的。特别，若 $\boldsymbol{V}_1\cong\boldsymbol{V}_2$，则

(1) $\boldsymbol{V}_1$ 中向量 $\boldsymbol{\alpha}_1,\cdots,\boldsymbol{\alpha}_t$ 线性相关$\Leftrightarrow\boldsymbol{V}_2$ 中向量 $\sigma(\boldsymbol{\alpha}_1),\cdots,\sigma(\boldsymbol{\alpha}_t)$ 线性相关；

(2) $\boldsymbol{V}_1$ 中的向量 $\boldsymbol{\alpha}_1,\cdots,\boldsymbol{\alpha}_t$ 线性无关$\Leftrightarrow\boldsymbol{V}_2$ 中的向量 $\sigma(\boldsymbol{\alpha}_1),\cdots,\sigma(\boldsymbol{\alpha}_t)$线性无关；

(3) 若 $\boldsymbol{E}_1\subseteq\boldsymbol{V}_1$，$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_t$ 是 $\boldsymbol{E}_1$ 的一个极大线性无关组$\Leftrightarrow$ $\sigma(\boldsymbol{\alpha}_1),\sigma(\boldsymbol{\alpha}_2),\cdots,\sigma(\boldsymbol{\alpha}_t)$是 $\boldsymbol{E}_2=\{\sigma(\boldsymbol{\alpha})\mid\alpha\in\boldsymbol{E}_1\}$ 的一个极大线性无关组。

证(1) $\Rightarrow$：即同态性质(4)。

$\Leftarrow$：$\sigma(\boldsymbol{\alpha}_1),\cdots,\sigma(\boldsymbol{\alpha}_t)$线性相关，故存在不全为 0 的 $k_1,\cdots,k_t\in\boldsymbol{P}$，使

$$\sum_{i=1}^{t}k_i\sigma(\boldsymbol{\alpha}_i)=\boldsymbol{O}_2,$$

于是

$$\sigma\left(\sum_{i=1}^{t} k_i\alpha_i\right) = O_2,$$

因为 σ 是同构映射，从而是单射，故由同态性质(1)得

$$\sum_{i=1}^{t} k_i\alpha_i = O_1,$$

即 α_1，…，α_t 线性相关。

(2) 即(1)的逆否命题。

(3) 由(2)

$\alpha_1,\cdots,\alpha_t$ 线性无关$\Longleftrightarrow\sigma(\alpha_1)$，…，$\sigma(\alpha_t)$线性无关。由(1)，任意 $\alpha\in \boldsymbol{E}_1$，$\alpha_1,\cdots,\alpha_t,\alpha$ 线性相关$\Longleftrightarrow\sigma(\alpha_1)$，…，$\sigma(\alpha_t)$，$\sigma(\alpha)$线性相关，故 $\alpha_1,\cdots,\alpha_t$ 是 $\boldsymbol{E}_1$ 的一个极大线性无关组$\Longleftrightarrow\sigma(\alpha_1)$，…，$\sigma(\alpha_t)$是 $\boldsymbol{E}_2$ 的一个极大线性无关组。

三、运算

命题 8.11 同态的积仍是同态。

即若 $\boldsymbol{V}_1 \overset{\sigma_1}{\backsim} \boldsymbol{V}_2$，$\boldsymbol{V}_2 \overset{\sigma_2}{\backsim} \boldsymbol{V}_3$，则映射 $\boldsymbol{V}_1 \xmapsto{\sigma_2\sigma_1} \boldsymbol{V}_3$ 也是同态映射。

证 任意 $k,l\in P,\boldsymbol{\alpha},\boldsymbol{\beta}\in \boldsymbol{V}_1$，$(\sigma_2\sigma_1)(k\boldsymbol{\alpha}+l\boldsymbol{\beta})=\sigma_2[\sigma_1(k\boldsymbol{\alpha}+l\boldsymbol{\beta})]=\sigma_2[k\sigma_1(\boldsymbol{\alpha})+l\sigma_1(\boldsymbol{\beta})]=k\sigma_2\sigma_1(\boldsymbol{\alpha})+l\sigma_2\sigma_1(\boldsymbol{\beta})$，

故 $\sigma_2\sigma_1$ 是同态映射。

命题 8.12 同构的逆仍是同构。即若 $\boldsymbol{V}_1 \overset{\sigma}{\cong} \boldsymbol{V}_2$，则 $\boldsymbol{V}_2 \xmapsto{\sigma^{-1}} \boldsymbol{V}_1$ 是同构映射。

证 σ 是可逆映射，σ^{-1} 也可逆，故是双射。又

$$\sigma^{-1}(k\boldsymbol{\alpha}+l\boldsymbol{\beta})=\sigma^{-1}[k\sigma\sigma^{-1}(\boldsymbol{\alpha})+l\sigma\sigma^{-1}(\boldsymbol{\beta})]=\sigma^{-1}\sigma[k\sigma^{-1}(\boldsymbol{\alpha}) \\ +l\sigma^{-1}(\boldsymbol{\beta})]=k\sigma^{-1}(\boldsymbol{\alpha})+l\sigma^{-1}(\boldsymbol{\beta}),$$

故 σ^{-1} 是同构映射。

四、同构的判别

定理 8.14 数域 P 上两个线性空间 V_1 与 V_2 同构 $\Leftrightarrow \dim V_1 = \dim V_2$。

证 $\Rightarrow$ 设 $V_1 \overset{\sigma}{\cong} V_2$，由同构性质(3) 若 $\alpha_1,\cdots,\alpha_n$ 是 V_1 的一组基，则 $\sigma(\alpha_1),\cdots,\sigma(\alpha_n)$ 是 V_2 的一组基，故 $\dim V_1 = n = \dim V_2$。

$\Leftarrow$ 若 $\dim V_1 = \dim V_2 = n$，设 $\varepsilon_1,\cdots,\varepsilon_n$ 与 $\eta_1,\cdots,\eta_n$ 分别是 V_1 与 V_2 的基。作映射 $V_1 \xrightarrow{\sigma} V_2$：

$$\sigma\left((\varepsilon_1,\cdots,\varepsilon_n)\begin{pmatrix}a_1\\ \vdots\\ a_n\end{pmatrix}\right) = (\eta_1,\cdots,\eta_n)\begin{pmatrix}a_1\\ \vdots\\ a_n\end{pmatrix}。$$

由坐标的性质(存在与惟一)，易得 σ 是双射。又

$$\begin{aligned}
&\sigma\left[k(\varepsilon_1,\cdots,\varepsilon_n)\begin{pmatrix}a_1\\ \vdots\\ a_n\end{pmatrix} + l(\varepsilon_1,\cdots,\varepsilon_n)\begin{pmatrix}b_1\\ \vdots\\ b_n\end{pmatrix}\right]\\
&= \sigma\left((\varepsilon_1,\cdots,\varepsilon_n)\left[k\begin{pmatrix}a_1\\ \vdots\\ a_n\end{pmatrix} + l\begin{pmatrix}b_1\\ \vdots\\ b_n\end{pmatrix}\right]\right)\\
&= (\eta_1,\cdots,\eta_n)\left[k\begin{pmatrix}a_1\\ \vdots\\ a_n\end{pmatrix} + l\begin{pmatrix}b_1\\ \vdots\\ b_n\end{pmatrix}\right]\\
&= k(\eta_1,\cdots,\eta_n)\begin{pmatrix}a_1\\ \vdots\\ a_n\end{pmatrix} + l(\eta_1,\cdots,\eta_n)\begin{pmatrix}b_1\\ \vdots\\ b_n\end{pmatrix}\\
&= k\sigma\left[(\varepsilon_1,\cdots,\varepsilon_n)\begin{pmatrix}a_1\\ \vdots\\ a_n\end{pmatrix}\right] + l\sigma\left[(\varepsilon_1,\cdots,\varepsilon_n)\begin{pmatrix}b_1\\ \vdots\\ b_n\end{pmatrix}\right]
\end{aligned}$$

故 σ 是同构映射，从而 $V_1 \overset{\sigma}{\cong} V_2$。

定义 8.20 数域 P 上两个内积空间 V_1 与 V_2 称为同构，如果存在线性空间 V_1 到 V_2 上的同构映射 σ，且满足：

对任意 $\alpha,\beta \in V_1$，$[\alpha,\beta] = [\sigma(\alpha),\sigma(\beta)]$。

定理 8.15 数域 P 上两个内积空间 V_1 与 V_2 同构 $\Leftrightarrow \dim V_1 = \dim V_2$。

证 $\Rightarrow$ 内积空间 V_1 同构于 V_2，必线性空间 V_1 同构于 V_2，由定理 8.14，$\dim V_1 = \dim V_2$。

$\Leftarrow$ 若 $\dim V_1 = \dim V_2$。设 $\varepsilon_1, \cdots, \varepsilon_n$ 与 $\eta_1, \cdots, \eta_n$ 分别是 V_1 与 V_2 的标准正交基，由定理 8.14 的证明可知：

$$\sigma\left((\varepsilon_1, \cdots, \varepsilon_n)\begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix}\right) = (\eta_1, \cdots, \eta_n)\begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix}$$

是线性空间 V_1 到 V_2 上的同构映射。又若

$$\alpha = (\varepsilon_1, \cdots, \varepsilon_n)\begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix}, \beta = (\varepsilon_1, \cdots, \varepsilon_n)\begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix},$$

则

$$\sigma(\alpha) = (\eta_1, \cdots, \eta_n)\begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix}, \sigma(\beta) = (\eta_1, \cdots, \eta_n)\begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix}。$$

由于 $\varepsilon_1, \cdots, \varepsilon_n$ 是 V_1 的标准正交基，$\eta_1, \cdots, \eta_n$ 是 V_2 的标准正交基，故

$$[\alpha,\beta] = (\bar{a}_1, \cdots, \bar{a}_n)\begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix} = [\sigma(\alpha), \sigma(\beta)],$$

即内积空间 V_1 同构于 V_2。

习　题　8

1. 检验以下集合对于所指运算是否构成实数域上的线性空间：

(1) 次数等于 $n(n\geqslant 1)$ 的实系数多项式全体，对于多项式的加法和数乘运算；

(2) 全体实对称(反对称，上下三角)矩阵，对于矩阵的加法和数乘运算；

(3) 设 $\boldsymbol{A}$ 是 n 阶实方阵，$\boldsymbol{A}$ 的实系数多项式全体，对于矩阵的加法和数乘运算；

(4) $\boldsymbol{P}$ 为数域，集合 $\boldsymbol{P}^n$ 对于如下定义的加法与数乘：

$$\begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} + \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix} = \begin{pmatrix} a_1 + b_1 \\ \vdots \\ a_n + b_n \end{pmatrix},$$

$$k\begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \qquad k \in \boldsymbol{P};$$

(5) 集合与加法同(4)，数乘定义为

$$k\boldsymbol{\alpha} = \boldsymbol{\alpha}。\quad \boldsymbol{\alpha} \in \boldsymbol{P}^n, k \in \boldsymbol{P};$$

(6) 全体正实数 $\mathbf{R}^+$，加法和数乘定义为：

$$a \oplus b = ab,$$
$$k \otimes a = a^k。$$

2. 求下列线性空间的维数与一组基：

(1) 数域 $\boldsymbol{P}$ 上 n 阶上三角矩阵全体组成的线性空间；

(2) 数域 $\boldsymbol{P}$ 上线性空间 $\boldsymbol{V} = \{\boldsymbol{A} \in \boldsymbol{P}^{n\times n} \mid \boldsymbol{A}' = \boldsymbol{A}\}$；

(3) 数域 $\boldsymbol{P}$ 上线性空间 $\boldsymbol{V} = \{\boldsymbol{A} \in \boldsymbol{P}^{n\times n} \mid \boldsymbol{A}' = -\boldsymbol{A}\}$；

(4) 数域 $\boldsymbol{P}$ 上线性空间 $\boldsymbol{V}=\left\{\begin{pmatrix}a_1\\ \vdots\\ a_n\end{pmatrix}\in \boldsymbol{P}^n\ \middle|\ \sum_{i=1}^{n}a_i=0\right\}$;

(5) 习题 1 的 6;

(6) 实数域 $\mathbf{R}$ 上线性空间 $\boldsymbol{V}=\{f(\boldsymbol{A})\mid f(x)\in\mathbf{R}[x]\}$,这里

$$\boldsymbol{A}=\begin{pmatrix}0 & 1 & & 0\\ & \ddots & \ddots & \\ & & \ddots & 1\\ 0 & & & 0\end{pmatrix}_{n\times n}.$$

3. 设 $a\in\boldsymbol{P}$,求证:$1, x-a, \cdots, (x-a)^{n-1}$ 是 $\boldsymbol{P}[x]_n$ 的一组基。

4. 若 $a_1,\cdots,a_n$ 是 $\boldsymbol{P}$ 中 n 个两两不同的数,记

$$f(x)=\prod_{i=1}^{n}(x-a_i),\ f_i(x)=\frac{f(x)}{x-a_i},\ i=1,\cdots,n$$

求证:$f_1,\cdots,f_n$ 是 $\boldsymbol{P}[x]_n$ 的一组基。

5. 设 a_1,a_2,a_3,a_4 是 $\boldsymbol{P}$ 中 4 个两两不同的数,记

$$\boldsymbol{A}_i=\begin{pmatrix}1, & a_i\\ a_i^2, & a_i^3\end{pmatrix},i=1,2,3,4$$

求证:$\boldsymbol{A}_1,\boldsymbol{A}_2,\boldsymbol{A}_3,\boldsymbol{A}_4$ 是 $\boldsymbol{P}^{2\times 2}$ 的一组基。

6. 求向量 $\boldsymbol{\alpha}$ 在基 $\boldsymbol{\varepsilon}_1,\boldsymbol{\varepsilon}_2,\boldsymbol{\varepsilon}_3,\boldsymbol{\varepsilon}_4$ 下的坐标:

(1) $\boldsymbol{\varepsilon}_1=\begin{pmatrix}1\\1\\1\\1\end{pmatrix}$, $\boldsymbol{\varepsilon}_2=\begin{pmatrix}1\\1\\-1\\-1\end{pmatrix}$, $\boldsymbol{\varepsilon}_3=\begin{pmatrix}1\\-1\\1\\-1\end{pmatrix}$, $\boldsymbol{\varepsilon}_4=\begin{pmatrix}1\\-1\\-1\\1\end{pmatrix}$,

$\boldsymbol{\alpha}=\begin{pmatrix}1\\2\\1\\1\end{pmatrix}$;

(2) $\boldsymbol{\varepsilon}_1 = \begin{pmatrix} 1 \\ 2 \\ -1 \\ 0 \end{pmatrix}$, $\boldsymbol{\varepsilon}_2 = \begin{pmatrix} 1 \\ -1 \\ 1 \\ 1 \end{pmatrix}$, $\boldsymbol{\varepsilon}_3 = \begin{pmatrix} -1 \\ 2 \\ 1 \\ 1 \end{pmatrix}$, $\boldsymbol{\varepsilon}_4 = \begin{pmatrix} 4 \\ -1 \\ 0 \\ 1 \end{pmatrix}$,

$\boldsymbol{\alpha} = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}$。

7. 求由基 $\boldsymbol{\varepsilon}_1, \boldsymbol{\varepsilon}_2, \boldsymbol{\varepsilon}_3, \boldsymbol{\varepsilon}_4$ 到基 $\boldsymbol{\eta}_1, \boldsymbol{\eta}_2, \boldsymbol{\eta}_3, \boldsymbol{\eta}_4$ 的过渡矩阵：

(1) $\boldsymbol{\varepsilon}_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}$, $\boldsymbol{\varepsilon}_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}$, $\boldsymbol{\varepsilon}_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}$, $\boldsymbol{\varepsilon}_4 = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}$;

$\boldsymbol{\eta}_1 = \begin{pmatrix} 1 \\ 5 \\ 7 \\ 0 \end{pmatrix}$, $\boldsymbol{\eta}_2 = \begin{pmatrix} 3 \\ 1 \\ 4 \\ 2 \end{pmatrix}$, $\boldsymbol{\eta}_3 = \begin{pmatrix} 2 \\ 7 \\ 3 \\ 8 \end{pmatrix}$, $\boldsymbol{\eta}_4 = \begin{pmatrix} 3 \\ 1 \\ 2 \\ 1 \end{pmatrix}$;

(2) $\boldsymbol{\varepsilon}_1 = \begin{pmatrix} 1 \\ 2 \\ -1 \\ 0 \end{pmatrix}$, $\boldsymbol{\varepsilon}_2 = \begin{pmatrix} 1 \\ -1 \\ 1 \\ 1 \end{pmatrix}$, $\boldsymbol{\varepsilon}_3 = \begin{pmatrix} -1 \\ 2 \\ 1 \\ 1 \end{pmatrix}$, $\boldsymbol{\varepsilon}_4 = \begin{pmatrix} -1 \\ -1 \\ 0 \\ 1 \end{pmatrix}$;

$\boldsymbol{\eta}_1 = \begin{pmatrix} 2 \\ 1 \\ 0 \\ 1 \end{pmatrix}$, $\boldsymbol{\eta}_2 = \begin{pmatrix} 0 \\ 1 \\ 2 \\ 2 \end{pmatrix}$, $\boldsymbol{\eta}_3 = \begin{pmatrix} -2 \\ 1 \\ 1 \\ 2 \end{pmatrix}$, $\boldsymbol{\eta}_4 = \begin{pmatrix} 1 \\ 3 \\ 1 \\ 2 \end{pmatrix}$。

8. 在习题 7 的(2)中，求向量 α 使它在两组基下的坐标相同。

9. 若 $\boldsymbol{W}_1$ 与 $\boldsymbol{W}_2$ 都是线性空间 $\boldsymbol{V}$ 的子空间，$\boldsymbol{W}_1 \subseteq \boldsymbol{W}_2$，且 $\dim \boldsymbol{W}_1 = \dim \boldsymbol{W}_2$，求证：$\boldsymbol{W}_1 = \boldsymbol{W}_2$。

10. 在 $\boldsymbol{R}^4$ 中，求下列齐次线性方程组的解空间的维数与一组基：

(1) $x_1 + 2x_2 + 3x_3 + 4x_4 = 0$,

(2) $\begin{cases} x_1 + 2x_2 + 3x_3 - 7x_4 = 0, \\ x_1 + 3x_2 - x_3 - x_4 = 0, \\ x_1 + 5x_2 - 8x_3 + 11x_4 = 0, \\ 2x_1 + x_2 + 7x_3 - 13x_4 = 0。 \end{cases}$

11. 求$\begin{cases} x_1 + 2x_2 - 3x_3 + 2x_4 = 0, \\ x_1 + 4x_2 + x_3 + 4x_4 = 0, \\ 2x_1 + 5x_2 - 4x_3 - 5x_4 = 0, \end{cases}$

与$\begin{cases} x_1 + 2x_2 + 5x_4 = 0, \\ 2x_1 + 4x_2 - 5x_3 + 5x_4 = 0, \end{cases}$

的解空间的交的一组基。

12. 若 $a\alpha + b\beta + c\gamma = 0$，且 $ac \neq 0$，求证：

$$L(\alpha,\beta) = L(\beta,\gamma)。$$

13. 若 $\beta_1 = \alpha_1 + \alpha_2$，$\beta_2 = \alpha_2 + \alpha_3$，…，$\beta_{s-1} = \alpha_{s-1} + \alpha_s$，$\beta_s = \alpha_s + \alpha_1$，$s$ 是奇数，求证：$L(\alpha_1, \cdots, \alpha_s) = L(\beta_1, \cdots, \beta_s)$。

14. 若 W_1 与 W_2 分别是实系数齐次线性方程组 $x_1 + \cdots + x_n = 0$ 与 $x_1 = x_2 = \cdots = x_n$ 的解空间，求证：$R^n = W_1 \oplus W_2$。

15. 记 $W_1 = \{A \in P^{n\times n} \mid A' = A\}$，$W_2 = \{A \in P^{n\times n} \mid A' = -A\}$，求证：$P^{n\times n} = W_1 \oplus W_2$。

16. 求证：n 维线性空间是 n 个一维子空间的直和。

17. 设 W_1，…，W_s 是线性空间 V 的真子空间，即 $W_i \neq V$，$i = 1, \cdots, s$，求证：存在向量 $\alpha \in V$，但 $\alpha \notin (W_1 \cup W_2 \cup \cdots \cup W_s)$。

18. 若 A 是 n 阶正定阵，在 R^n 中定义内积：$[\alpha,\beta] = \alpha' A\beta$，求证：$R^n$ 是 R 上的内积空间。

19. 若 ε_1，ε_2、ε_3 是内积空间的一组标准正交基。求证：$\alpha_i = \frac{2}{3}(\varepsilon_1 + \varepsilon_2 + \varepsilon_3) - \varepsilon_i$，$(i = 1, 2, 3)$ 也是标准正交基。

20. 求由下列向量组生成的子空间 $L(\alpha_1, \alpha_2, \alpha_3, \alpha_4)$ 的正交基：

(1) $\boldsymbol{\alpha}_1=\begin{pmatrix}2\\3\\-4\\-6\end{pmatrix}$, $\boldsymbol{\alpha}_2=\begin{pmatrix}1\\8\\-2\\-16\end{pmatrix}$, $\boldsymbol{\alpha}_3=\begin{pmatrix}2\\4\\10\\-29\end{pmatrix}$, $\boldsymbol{\alpha}_4=\begin{pmatrix}3\\11\\4\\-7\end{pmatrix}$;

(2) $\boldsymbol{\alpha}_1=\begin{pmatrix}1\\1\\-1\\-2\end{pmatrix}$, $\boldsymbol{\alpha}_2=\begin{pmatrix}-2\\1\\5\\11\end{pmatrix}$, $\boldsymbol{\alpha}_3=\begin{pmatrix}0\\3\\3\\7\end{pmatrix}$, $\boldsymbol{\alpha}_4=\begin{pmatrix}3\\-3\\-3\\-9\end{pmatrix}$。

21. 求齐次线性方程组

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

的解空间(作为 $\boldsymbol{C}^5$ 的子空间)的一组标准正交基。

22. 设 $\boldsymbol{\varepsilon}_1, \cdots, \boldsymbol{\varepsilon}_n$ 是内积空间 $\boldsymbol{V}$ 的一组基，$\boldsymbol{\alpha}\in\boldsymbol{V}$，且 $[\boldsymbol{\alpha},\boldsymbol{\varepsilon}_i]=0$，$i=1,\cdots,n$，求证：$\boldsymbol{\alpha}=0$。

23. 设 $\boldsymbol{\alpha}_1, \cdots, \boldsymbol{\alpha}_s$ 是内积空间 $\boldsymbol{V}$ 的一组向量，记 $\Delta=([\boldsymbol{\alpha}_i,\boldsymbol{\alpha}_j])_{s\times s}$，求证：$\boldsymbol{\alpha}_1, \cdots, \boldsymbol{\alpha}_s$ 线性相关 $\Leftrightarrow |\Delta|=0$。

24. 求向量 $\boldsymbol{\alpha}$ 在一维子空间 $L(\boldsymbol{\beta})$ 上的正交投影。

(1) $\boldsymbol{\alpha}=\begin{pmatrix}2\\1\\3\end{pmatrix}$, $\boldsymbol{\beta}=\begin{pmatrix}6\\3\\9\end{pmatrix}$;

(2) $\boldsymbol{\alpha}=\begin{pmatrix}1\\2\\1\end{pmatrix}$, $\boldsymbol{\beta}=\begin{pmatrix}3\\4\\5\end{pmatrix}$。

25. 求向量 $\boldsymbol{\alpha}=\begin{pmatrix}-3\\4\\2\end{pmatrix}$ 在子空间 $\boldsymbol{W}=L\left[\begin{pmatrix}2\\1\\2\end{pmatrix},\begin{pmatrix}3\\3\\0\end{pmatrix},\begin{pmatrix}5\\4\\2\end{pmatrix}\right]$ 上的正交投影。

26. 求向量 $\alpha = \begin{pmatrix} 5 \\ -3 \\ 2 \\ 1 \end{pmatrix}$ 在子空间 W 上的正交投影。这里 W 为下列齐次线性方程组

$$\begin{cases} x_1 + 2x_2 + 3x_3 + 4x_4 = 0 \\ 2x_1 + x_2 + 2x_3 + 5x_4 = 0 \\ 3x_1 + 3x_2 + 5x_3 + 9x_4 = 0 \end{cases}$$

的解空间。

27. 在 $\boldsymbol{R}$ 上内积空间 $\boldsymbol{R}^5$ 中，设 $\boldsymbol{W} = L\left(\begin{pmatrix} 1 \\ 2 \\ 3 \\ -1 \\ 2 \end{pmatrix}, \begin{pmatrix} 2 \\ 4 \\ 7 \\ 2 \\ -1 \end{pmatrix} \right)$，求 $\boldsymbol{W}^{\perp}$ 的一组基。

28. 设 $\boldsymbol{W}_1$，$\boldsymbol{W}_2$ 是内积空间的两个子空间，求证：

$$(\boldsymbol{W}_1 + \boldsymbol{W}_2)^{\perp} = \boldsymbol{W}_1^{\perp} \cap \boldsymbol{W}_2^{\perp};$$

$$(\boldsymbol{W}_1 \cap \boldsymbol{W}_2)^{\perp} = \boldsymbol{W}_1^{\perp} + \boldsymbol{W}_2^{\perp}。$$

29. 若 $\boldsymbol{V}$ 为实数域上 n 维内积空间，求证：$\boldsymbol{V}$ 中不存在 $(n+2)$ 个向量，其中任意两个不同向量的内积都小于零。

第九章　线性变换

本章以矩阵为工具，研究线性空间到自身的同态映射，即线性变换。

§1　线性变换的概念

一、定义

定义 9.1　数域 P 上线性空间 V 到自身的同态映射

$$V \overset{\mathscr{A}}{\simeq} V$$

称为 V 的线性变换。即 V 到 V 的映射 $\mathscr{A}$ 称为 V 的一个线性变换，若下列条件满足：

(1) 任意 $\alpha,\beta \in V$，有

$$\mathscr{A}(\alpha+\beta)=\mathscr{A}(\alpha)+\mathscr{A}(\beta);$$

(2) 任意 $k \in P,\alpha \in V$，有

$$\mathscr{A}(k\alpha)=k\mathscr{A}(\alpha)。$$

或(代替(1)与(2))，对任意 $k,l \in P;\alpha,\beta \in V$，有

$$\mathscr{A}(k\alpha+l\beta)=k\mathscr{A}(\alpha)+l\mathscr{A}\beta)。$$

例 9.1　线性空间 V 到自身的零映射 O：任意 $\alpha \in V, O(\alpha)=0$ 是 V 的线性变换。

例 9.2　线性空间 V 到自身的恒等映射 I：任意 $\alpha \in V, I(\alpha)=\alpha$，是 V 的线性变换。

例 9.3　若 A 是数域 P 上 n 阶方阵，令 $V=P^n$，则映射 $V \xmapsto{\sigma} V$，

$\sigma(\alpha) = A\alpha$ 是 V 的线性变换。

证 因为 $\sigma(k\alpha + l\beta) = A(k\alpha + l\beta)$

$= A(k\alpha) + A(l\beta)$ （矩阵加乘分配律）

$= kA\alpha + lA(\beta)$ （矩阵数乘性质）

$= k\sigma(\alpha) + l\sigma(\beta)$。

例 9.4 若 A 与 B 分别是数域 P 上 n 阶与 s 阶方阵，$V = P^{n\times s}$ 是数域 P 上 $n \times s$ 阵全体组成的数域 P 上线性空间，则映射 $V \xrightarrow{\sigma} V$：

$$\sigma(\alpha) = A\alpha B$$

是 V 的线性变换。

证 任意 $k,l \in P,\alpha,\beta \in V$

$$\begin{aligned}\sigma(k\alpha + l\beta) &= A(k\alpha + l\beta)B \\ &= A[(k\alpha)B + (l\beta)B] = A(k\alpha)B + A(l\beta)B \\ &= k(A\alpha B) + l(A\beta B) \\ &= k\sigma(\alpha) + l\sigma(\beta),\end{aligned}$$

故 σ 是 V 的线性变换。

例 9.5 XOY 平面作为实数域 $\mathbf{R}$ 上的二维线性空间：

$$V = R^2 = \left\{ \begin{pmatrix} x \\ y \end{pmatrix} \middle| x,y \in \mathbf{R} \right\},$$

则绕原点 O 的旋转 σ 是 V 的线性变换。

证 设 XOY 平面绕原点 O 反时针方向旋转角度 φ，用 σ 表示这个旋转变换。σ 把具有极坐标(r,θ)的点 M 变成了具有极坐标$(r,\theta+\varphi)$的点 M_1。若 M 的直角坐标为$\begin{pmatrix} x \\ y \end{pmatrix}$，则

$$\begin{cases} x = r\cos\theta, \\ y = r\sin\theta。\end{cases}$$

点 M_1 的横、纵坐标分别为

$$\begin{cases} r\cos(\theta+\varphi) = r(\cos\theta\cos\varphi - \sin\theta\sin\varphi), \\ r\sin(\theta+\varphi) = r(\sin\varphi\cos\theta + \cos\varphi\sin\theta)。\end{cases}$$

于是

$$\sigma\left(\begin{pmatrix} x \\ y \end{pmatrix}\right) = \begin{pmatrix} r\cos\theta\cos\varphi - r\sin\theta\sin\varphi \\ r\sin\varphi\cos\theta + r\cos\varphi\sin\theta \end{pmatrix} = \begin{pmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{pmatrix}\begin{pmatrix} r\cos\theta \\ r\sin\theta \end{pmatrix}$$

$$= \begin{pmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix}。$$

由例 9.3，σ 是 V 的线性变换。

例 9.6 线性空间 $V = P[x]$，V 的微分变换 D：

$$D(f(x)) = f'(x),$$

由求导运算法则，可知 D 是 V 的线性变换。

例 9.7 设 V 是内积空间，W 是 V 的子空间，则 $V = W \oplus W^{\perp}$。定义 V 在 W 上的投影变换 Π 为：

若 $\alpha \in V, \alpha = \alpha_1 + \alpha^2, \alpha_1 \in W, \alpha_2 \in W^{\perp}, \Pi(\alpha) = \alpha_1$。

则 Π 为 V 的线性变换，$\Pi(\alpha)$ 为 α 在 W 上正交投影。

证 若 $\alpha = \alpha_1 + \alpha_2, \beta = \beta_1 + \beta_2, \alpha_1, \alpha_2 \in W; \beta_1, \beta_2 \in W^{\perp}$。则

$$k\alpha = k\alpha_1 + k\alpha_2, l\beta = l\beta_1 + l\beta_2。$$

由于 $W, W^{\perp}$ 是子空间，从而

$$k\alpha_1 + l\beta_1 \in W, k\alpha_2 + l\beta_2 \in W^{\perp},$$

则

$$k\alpha + l\beta = (k\alpha_1 + l\beta_1) + (k\alpha_2 + l\beta_2)。$$

所以

$$\Pi(k\alpha + l\beta) = k\alpha_1 + l\beta_1 = k\Pi(\alpha) + l\Pi(\beta),$$

即 Π 为 V 的线性变换。

二、性质

若 σ 是线性空间 V 的线性变换，即 V 到 V 的同态映射，故有

(1) $\sigma(0) = 0$；

(2) $\alpha\left(\sum_{i=1}^{t} k_i\alpha_i\right) = \sum_{i=1}^{t} k_i\sigma(\alpha_i)$；

(3) 若 $\alpha_1,\cdots,\alpha_t$ 线性相关，则 $\sigma(\alpha_1),\cdots,\sigma(\alpha_t)$ 也线性相关；

(4) 若 W 是 V 的子空间，则 $\sigma(W) = \{\sigma(\alpha)\,|\,\alpha \in V\}$ 也是 V 的子空间。

三、线性变换的运算

定义 9.2 设 $\mathscr{A}$，$\mathscr{B}$ 是数域 P 上线性空间 V 的两个线性变换，$k \in P$，定义

$$V \xmapsto{\mathscr{A}+\mathscr{B}} V,\ V \xmapsto{k\mathscr{A}} V,\ V \xmapsto{\mathscr{A}\cdot\mathscr{B}} V$$

如下：

$$(\mathscr{A}+\mathscr{B})(\alpha) = \mathscr{A}(\alpha) + \mathscr{B}(\alpha),$$
$$(k\mathscr{A})(\alpha) = k\mathscr{A}(\alpha),$$
$$(\mathscr{A}\mathscr{B})(\alpha) = \mathscr{A}[\mathscr{B}(\alpha)]。$$

分别称它们为线性变换的和、数乘与积。

直接验证易得，线性变换的和、数量积与乘积仍是线性变换。

定义 9.3 线性空间 V 的线性变换 σ 称为可逆线性变换，如果 σ 作为 V 到自身的映射是可逆的。此时，σ 是 V 到 V 的同构映射，其逆 σ^{-1} 也是 V 到 V 的同构映射，因而仍是 V 的线性变换。

定义 9.4 若 $\mathscr{A}$ 是线性空间 V 的线性变换，归纳定义 $\mathscr{A}$ 的幂为 $\mathscr{A}^k = \mathscr{A}^{k-1}\cdot\mathscr{A}$，则 $\mathscr{A}^k$ 仍是 V 的线性变换。

若 $f(x) = \sum_{i=0}^{t} a_i x^i \in P[x]$，称线性变换

$$a_0 I + a_1\mathscr{A} + \cdots + a_t\mathscr{A}^t$$

为 $\mathscr{A}$ 的多项式，记成 $f(\mathscr{A})$。这里 I 是恒等映射。

§2 线性变换与矩阵

一、线性变换在一组给定基下的矩阵

设 V 是数域 P 上 n 维线性空间，$\varepsilon_1,\cdots,\varepsilon_n$ 是 V 的一组基。给了 V 中一个向量 α，即给了它的坐标 $\begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix}$，$\alpha = \sum\limits_{i=1}^{n} a_i\varepsilon_i$。若 $\mathscr{A}$ 是 V 的一个线性变换，则 $\mathscr{A}(\alpha) = \sum\limits_{i=1}^{n} a_i\mathscr{A}(\varepsilon_i)$，因而 $\mathscr{A}$ 完全由 $\mathscr{A}(\varepsilon_1),\cdots,\mathscr{A}(\varepsilon_n)$ 确定，或由它们的坐标 $\begin{pmatrix} a_{11} \\ \vdots \\ a_{n1} \end{pmatrix},\cdots,\begin{pmatrix} a_{1n} \\ \vdots \\ a_{nn} \end{pmatrix}$ 确定。这里

$$\mathscr{A}(\varepsilon_i) = (\varepsilon_1,\cdots,\varepsilon_n)\begin{pmatrix} a_{1i} \\ a_{2i} \\ \vdots \\ a_{ni} \end{pmatrix}, i = 1,\cdots,n。$$

记 $A = (a_{ij})_{n\times n}$，即矩阵 A 的第 k 列为向量 $\mathscr{A}(\varepsilon_k)$ 的坐标，称矩阵 A 为线性变换 $\mathscr{A}$ 在基 $\varepsilon_1,\cdots,\varepsilon_n$ 下的矩阵。它们的关系是：

$$(\mathscr{A}(\varepsilon_1),\cdots,\mathscr{A}(\varepsilon_n)) = (\varepsilon_1,\cdots,\varepsilon_n)A。$$

为了书写和运算方便，我们记

$$(\mathscr{A}(\varepsilon_1),\cdots,\mathscr{A}(\varepsilon_n)) \text{ 为 } \mathscr{A}(\varepsilon_1,\cdots,\varepsilon_n)。$$

例 9.8 V 的恒等映射 I 在任何一组基下的矩阵都是单位阵 E。

例 9.9 $A \in P^{n\times n}, V = P^n$ 的线性变换 σ：$\sigma(\alpha) = A\alpha$。则 σ 在自

然基

$$\varepsilon_1=\begin{pmatrix}1\\0\\\vdots\\0\\0\end{pmatrix},\cdots,\varepsilon_i=\begin{pmatrix}0\\\vdots\\0\\1\\0\\\vdots\\0\end{pmatrix}\left.\begin{matrix}\\ \\ \\ \end{matrix}\right\}i,\cdots,\varepsilon_n=\begin{pmatrix}0\\0\\\vdots\\0\\1\end{pmatrix}$$

下的矩阵为 A。这是因为 $\sigma(\varepsilon_i)=A\varepsilon_i=A^{(i)}=(\varepsilon_1,\cdots,\varepsilon_n)A^{(i)}$，其坐标为矩阵 A 的第 i 列，因而 σ 在基 $\varepsilon_1,\cdots,\varepsilon_n$ 下的矩阵为 $(A^{(1)}\cdots A^{(n)})$，即矩阵 A。

例 9.10 在 $P^{2\times 2}$ 中定义线性变换 σ：$\sigma(\alpha)=\begin{pmatrix}1&1\\1&1\end{pmatrix}\alpha$。则 σ 在基

$$\varepsilon_1=\begin{pmatrix}1&0\\0&0\end{pmatrix},\varepsilon_2=\begin{pmatrix}0&1\\0&0\end{pmatrix},\varepsilon_3=\begin{pmatrix}0&0\\1&0\end{pmatrix},\varepsilon_4=\begin{pmatrix}0&0\\0&1\end{pmatrix}$$

下的矩阵为

$$\begin{pmatrix}1&0&1&0\\0&1&0&1\\1&0&1&0\\0&1&0&1\end{pmatrix}。$$

这是因为

$$\sigma(\varepsilon_1)=\begin{pmatrix}1&1\\1&1\end{pmatrix}\begin{pmatrix}1&0\\0&0\end{pmatrix}=\begin{pmatrix}1&0\\1&0\end{pmatrix}=\varepsilon_1+\varepsilon_3,$$

其坐标为 $\begin{pmatrix}1\\0\\1\\0\end{pmatrix}$。类似地，$\sigma(\varepsilon_2)$ 与 $\sigma(\varepsilon_4)$ 的坐标为 $\begin{pmatrix}0\\1\\0\\1\end{pmatrix}$，$\sigma(\varepsilon_3)$ 的坐标为 $\begin{pmatrix}1\\0\\1\\0\end{pmatrix}$。

4 个坐标列组成 σ 在基 $\varepsilon_1,\varepsilon_2,\varepsilon_3,\varepsilon_4$ 下的矩阵

$$\begin{pmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \end{pmatrix}。$$

例 9.11 $\boldsymbol{P}[x]_n$ 的微分变换 $\boldsymbol{D}$：$\boldsymbol{D}(f(x)) = f'(x)$。$\boldsymbol{D}$ 在基 $1,\boldsymbol{x},\boldsymbol{x}^2,\cdots,\boldsymbol{x}^{n-1}$ 下的矩阵为

$$\begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 2 & \cdots & \vdots \\ 0 & 0 & 0 & \cdots & \vdots \\ \vdots & \vdots & \vdots & & \vdots \\ \vdots & \vdots & \vdots & & 0 \\ 0 & 0 & 0 & \cdots & n-1 \\ 0 & 0 & 0 & \cdots & 0 \end{pmatrix},$$

这是因为 $\boldsymbol{D}(\boldsymbol{x}^i)=i\boldsymbol{x}^{i-1}$，$i=0,\cdots,n-1$。$\boldsymbol{D}(1)$的坐标为 $\begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}$，$\boldsymbol{D}(x^i)$的坐标为 $\begin{pmatrix} \left.\begin{matrix} 0 \\ \vdots \\ 0 \end{matrix}\right\} i-1 \\ i \\ 0 \\ \vdots \\ 0 \end{pmatrix}$，$i=1,\cdots,n-1$。$n$ 个坐标列组成的矩阵为

$$\begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & & \cdots & 0 \\ 0 & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \cdots & n-1 \\ 0 & 0 & 0 & \cdots & 0 \end{pmatrix}。$$

二、$L(V)$与$P^{n\times n}$

引理 9.1 设$\mathscr{A}$是数域P上n维线性空间V的一个线性变换，$\boldsymbol{B}=(b_{ij})\in \boldsymbol{P}^{n\times n}$，$\boldsymbol{\varepsilon}_1,\cdots,\boldsymbol{\varepsilon}_n\in \boldsymbol{V}$，则$\mathscr{A}[(\boldsymbol{\varepsilon}_1,\cdots,\boldsymbol{\varepsilon}_n)\boldsymbol{B}]=[\mathscr{A}(\boldsymbol{\varepsilon}_1,\cdots,\boldsymbol{\varepsilon}_n)]\boldsymbol{B}$。

证 等式左、右两边都是n个向量的集合：

$$\begin{aligned}\text{左边的第}\ i\ \text{个向量} &= \mathscr{A}[(\boldsymbol{\varepsilon}_1,\cdots,\boldsymbol{\varepsilon}_n)\boldsymbol{B}^{(i)}] = \mathscr{A}\left(\sum_{j=1}^{n} b_{ji}\boldsymbol{\varepsilon}_j\right)\\ &= \sum_{j=1}^{n} b_{ji}\mathscr{A}(\boldsymbol{\varepsilon}_j)\\ &= (\mathscr{A}(\boldsymbol{\varepsilon}_1),\cdots,\mathscr{A}(\boldsymbol{\varepsilon}_n))\begin{pmatrix} b_{1i}\\ \vdots\\ b_{ni}\end{pmatrix}\\ &= [\mathscr{A}(\boldsymbol{\varepsilon}_1,\cdots,\boldsymbol{\varepsilon}_n)]\boldsymbol{B}^{(i)} = \text{右边的第}\ i\ \text{个向量},\end{aligned}$$

故左边＝右边。

其实，因为$\mathscr{A}(\alpha k)=(\mathscr{A}(\alpha))k$，又$\mathscr{A}(\alpha k)$是$\boldsymbol{V}$中向量，加法满足交换律、结合律，故由矩阵乘法的广义结合律，即得本引理。

定理 9.1 设$\boldsymbol{V}$是数域$\boldsymbol{P}$上n维线性空间，$\varepsilon_1,\cdots,\varepsilon_n$是$\boldsymbol{V}$的一组基，用$L(\boldsymbol{V})$表示$\boldsymbol{V}$的线性变换全体组成的集合。作映射

$$L(\boldsymbol{V})\overset{\sigma}{\longmapsto}\boldsymbol{P}^{n\times n}:\ \sigma(\mathscr{A})=\boldsymbol{A},$$

这里，$\mathscr{A}(\boldsymbol{\varepsilon}_1,\cdots,\boldsymbol{\varepsilon}_n)=(\boldsymbol{\varepsilon}_1,\cdots,\boldsymbol{\varepsilon}_n)\boldsymbol{A}$，则映射$\sigma$是双射，且$\sigma$保持各类运算不变，即

(1) $\sigma(0)=0$，$\sigma(\boldsymbol{I})=\boldsymbol{E}$；

(2) $\sigma(\mathscr{A}+\mathscr{B})=\sigma(\mathscr{A})+\sigma(\mathscr{B})$；

(3) $\sigma(k\mathscr{A})=k\sigma(\mathscr{A})$；

(4) $\sigma(\mathscr{A}\mathscr{B})=\sigma(\mathscr{A})\cdot\sigma(\mathscr{B})$；

(5) 若$f(x)\in \boldsymbol{P}[x]$，则$\sigma(f(\mathscr{A}))=f(\sigma(\mathscr{A}))$；

(6) $\mathscr{A}$可逆$\Leftrightarrow\sigma(\mathscr{A})$可逆。且可逆时，$\sigma(\mathscr{A}^{-1})=[\sigma(\mathscr{A})]^{-1}$。

证 设

$\sigma(\mathscr{A}) = A = (a_{ij})_{n\times n}, \sigma(\mathscr{B}) = \boldsymbol{B} = (b_{ij})_{n\times n}$；$\alpha = \sum_{i=1}^{n} a_i\varepsilon_i$。

(1) 若 $\boldsymbol{A} = \boldsymbol{B}$，则

$$\begin{aligned}
\mathscr{A}(\alpha) &= \sum_{i=1}^{n} a_i \mathscr{A}(\varepsilon_i) = (\mathscr{A}(\varepsilon_1), \cdots, \mathscr{A}(\varepsilon_n))\begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} \\
&= (\varepsilon_1, \cdots, \varepsilon_n)\boldsymbol{A}\begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} = (\varepsilon_1, \cdots, \varepsilon_n)\boldsymbol{B}\begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} \\
&= (\mathscr{B}(\varepsilon_1), \cdots, \mathscr{B}(\varepsilon_n))\begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} = \sum_{i=1}^{n} a_i\mathscr{B}(\varepsilon_i) \\
&= \mathscr{B}(\alpha),
\end{aligned}$$

即 σ 为单射。

任意 $\boldsymbol{A} \in \boldsymbol{P}^{n\times n}$，定义映射 $\boldsymbol{V} \xrightarrow{\mathscr{A}} \boldsymbol{V}$：

$$\mathscr{A}(\alpha) = \mathscr{A}\left(\sum_{i=1}^{n} a_i\varepsilon_i\right) = (\varepsilon_1, \cdots, \varepsilon_n)\boldsymbol{A}\begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix},$$

即映射$\mathscr{A}$把坐标为

$\begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix}$的向量变成坐标为 $\boldsymbol{A}\begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix}$的向量。易验证，$\mathscr{A}$ 是线性变换，

即$\mathscr{A} \in L(\boldsymbol{V})$，$\varepsilon_i$ 的坐标为$\begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}\}i$，$\mathscr{A}(\varepsilon_i)$的坐标为 $\boldsymbol{A}\begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}\}i = \boldsymbol{A}^{(i)}$，故

$\mathscr{A}$ 在基 $\varepsilon_1, \cdots, \varepsilon_n$ 下的矩阵为

$$(A^{(1)}A^{(2)}\cdots A^{(n)})=A,$$

即 $\sigma(\mathscr{A})=A$，所以 σ 是满射，因而是双射。

(2) $(\mathscr{A}+\mathscr{B})(\varepsilon_i)=\mathscr{A}(\varepsilon_i)+\mathscr{B}(\varepsilon_i)$，其坐标为 $\mathscr{A}(\varepsilon_i)$ 与 $\mathscr{B}(\varepsilon_i)$ 的坐标和，即 $A^{(i)}+B^{(i)}=(A+B)^{(i)}$。故 $(\mathscr{A}+\mathscr{B})$ 在基 $\varepsilon_1,\cdots,\varepsilon_n$ 下的矩阵为

$$A+B,\text{即 } \sigma(\mathscr{A}+\mathscr{B})=A+B=\sigma(\mathscr{A})+\sigma(\mathscr{B})。$$

(3) $(k\mathscr{A})(\varepsilon_i)=k\cdot\mathscr{A}(\varepsilon_i)$，其坐标为 $k\cdot A^{(i)}=(kA)^{(i)}$。故 $k\mathscr{A}$ 在基 $\varepsilon_1,\cdots,\varepsilon_n$ 下矩阵为 kA，即

$$\sigma(k\mathscr{A})=k\sigma(\mathscr{A})。$$

(4) $(\mathscr{A}\mathscr{B})(\varepsilon_1,\cdots,\varepsilon_n)=\mathscr{A}[\mathscr{B}(\varepsilon_1,\cdots,\varepsilon_n)]=\mathscr{A}[(\varepsilon_1,\cdots,\varepsilon_n)B]$（由引理 9.1）$=[\mathscr{A}(\varepsilon_1,\cdots,\varepsilon_n)]B=(\varepsilon_1,\cdots,\varepsilon_n)AB$。于是

$$\sigma(\mathscr{A}\mathscr{B})=AB=\sigma(\mathscr{A})\sigma(\mathscr{B})。$$

(5) 因为由(2)，(3)，(4) 可知

$$\sigma(\Sigma a_i\mathscr{A}^i)=\Sigma a_i\sigma(\mathscr{A}^i)=\Sigma a_i[\sigma(\mathscr{A})]^i。$$

(6) $\mathscr{A}\mathscr{B}=I\overset{\sigma\text{是单射}}{\Longleftrightarrow}\sigma(\mathscr{A}\mathscr{B})=\sigma(I)\overset{\text{性质(3)、(4)}}{\Longleftrightarrow}\sigma(\mathscr{A})\sigma(\mathscr{B})=E$。故 $\mathscr{A}$ 可逆 $\Leftrightarrow\sigma(\mathscr{A})$ 可逆，且

$$\sigma(\mathscr{A}^{-1})=[\sigma(\mathscr{A})]^{-1}。$$

推论 1 矩阵关于加法、数乘、乘法所具有的各种性质，如加法交换律，结合律，乘法结合律，加乘分配律等，线性变换都具有。

推论 2 $f(x)$ 是线性变换 $\mathscr{A}$ 的化零多项式，即 $f(\mathscr{A})=0\Leftrightarrow f(x)$ 是矩阵 $\sigma(\mathscr{A})$ 的化零多项式。

$f(x)$ 是线性变换 $\mathscr{A}$ 的最小多项式（次数最低的首 1 化零多项式）$\Leftrightarrow f(x)$ 是矩阵 $\sigma(\mathscr{A})$ 的最小多项式。

三、同一线性变换在不同基下的矩阵

设 $\mathscr{A}$ 是线性空间 V 的一个线性变换；$\varepsilon_1,\cdots,\varepsilon_n$ 与 $\eta_1,\cdots,\eta_n$ 是 V 的

两组基。若

$$(\eta_1,\cdots,\eta_n)=(\varepsilon_1,\cdots,\varepsilon_n)\boldsymbol{T},$$
$$\mathscr{A}(\varepsilon_1,\cdots,\varepsilon_n)=(\varepsilon_1,\cdots,\varepsilon_n)\boldsymbol{A},$$
$$\mathscr{A}(\eta_1,\cdots,\eta_n)=(\eta_1,\cdots,\eta_n)\boldsymbol{B},$$

则

$$\mathscr{A}(\eta_1,\cdots,\eta_n)=\mathscr{A}[(\varepsilon_1,\cdots,\varepsilon_n)\boldsymbol{T}]=[\mathscr{A}(\varepsilon_1,\cdots,\varepsilon_n)]\boldsymbol{T},$$

而

$$(\eta_1,\cdots,\eta_n)\boldsymbol{B}=(\varepsilon_1,\cdots,\varepsilon_n)\boldsymbol{TB},$$

于是

$$\mathscr{A}(\varepsilon_1,\cdots,\varepsilon_n)=(\varepsilon_1,\cdots,\varepsilon_n)\boldsymbol{TBT}^{-1},$$

由定理 9.1

$$\boldsymbol{A}=\boldsymbol{TBT}^{-1},\text{或 }\boldsymbol{B}=\boldsymbol{T}^{-1}\boldsymbol{AT},$$

这是同一线性变换$\mathscr{A}$在不同基下的矩阵关系，此即下述定理：

定理 9.2 线性变换在不同基下的矩阵是相似的。

由于复矩阵都相似于 Jordan 标准形。故有：

命题 9.1 设 $\boldsymbol{V}$ 是复数域 **C** 上 n 维线性空间，$\mathscr{A}\in L(\boldsymbol{V})$，则存在 $\boldsymbol{V}$ 的基，使$\mathscr{A}$在这组基下矩阵为 Jordan 标准形，且存在 $\boldsymbol{V}$ 的基使$\mathscr{A}$在这组基下矩阵为对角阵$\Leftrightarrow\mathscr{A}$的最小多项式无重根。

证 任取 $\boldsymbol{V}$ 的一组基 $\varepsilon_1,\cdots,\varepsilon_n$，设$\mathscr{A}$在该基下矩阵为 $\boldsymbol{A}$。$\boldsymbol{A}$ 相似于 Jordan 标准形 $\boldsymbol{J}$，$\boldsymbol{J}=\boldsymbol{T}^{-1}\boldsymbol{AT}$。令$(\eta_1,\cdots,\eta_n)=(\varepsilon_1,\cdots,\varepsilon_n)\boldsymbol{T}$。因为 $\boldsymbol{T}$ 可逆，$\eta_1,\cdots,\eta_n$ 为 $\boldsymbol{V}$ 的基。$\mathscr{A}$在基 $\eta_1,\cdots,\eta_n$ 下的矩阵为 $\boldsymbol{T}^{-1}\boldsymbol{AT}=\boldsymbol{J}$。

又 $\boldsymbol{A}\sim$ 对角阵 $\Leftrightarrow\boldsymbol{A}$ 的最小多项式为不同的一次因子积 $\Leftrightarrow$(在 **C** 内)$\boldsymbol{A}$ 的最小多项式无重根 $\Leftrightarrow\mathscr{A}$ 的最小多项式无重根。

四、线性变换的特征值与特征向量

定义 9.5 设$\mathscr{A}$为数域 P 上线性空间 $\boldsymbol{V}$ 的一个线性变换，若 $\lambda\in P$，$\alpha\in\boldsymbol{V}$，$\alpha\neq0$，使$\mathscr{A}(\alpha)=\lambda\alpha$，则称 λ 为线性变换$\mathscr{A}$的特征值，α 为$\mathscr{A}$的对应于特征值 λ 的特征向量。

例 9.12 任意非零向量 α 都是恒等映射 I 的对应于 1 的特征向

量。因为$I(\alpha)=\alpha=1\cdot\alpha$。

例 9.13 设Π为内积空间V到子空间W的投影变换，即$\alpha=\alpha_1+\alpha_2,\alpha_1\in W,\alpha_2\in W^{\perp}$，则$\Pi(\alpha)=\alpha_1$。由此，$W$中的非零向量都是$\Pi$的对应于1的特征向量，而$W^{\perp}$中的非零向量是对应于0的特征向量。

任取V的一组基$\varepsilon_1,\cdots,\varepsilon_n$。设$\mathscr{A}$在该基下的矩阵为$A$。$\alpha\in V$，$\alpha$在该基下的坐标为$\begin{pmatrix}a_1\\ \vdots\\ a_n\end{pmatrix}$，则$\mathscr{A}(\alpha)$的坐标为$A\begin{pmatrix}a_1\\ \vdots\\ a_n\end{pmatrix}$。故$\alpha$是$\mathscr{A}$(对应于特征值$\lambda$)的特征向量

$$\Leftrightarrow\alpha\neq 0,\mathscr{A}(\alpha)=\lambda\alpha\Leftrightarrow\begin{pmatrix}a_1\\ \vdots\\ a_n\end{pmatrix}\neq 0,A\begin{pmatrix}a_1\\ \vdots\\ a_n\end{pmatrix}=\lambda\begin{pmatrix}a_1\\ \vdots\\ a_n\end{pmatrix}\Leftrightarrow\alpha$$

的坐标是矩阵A的特征向量(对应于特征值λ)。

因为相似矩阵有相同的特征多项式，而同一线性变换在不同基下的矩阵是相似的，故我们称线性变换(在任一组基下)的矩阵的特征多项式为线性变换的特征多项式。

因为矩阵的特征多项式是矩阵的化零多项式，从而线性变换的特征多项式也是线性变换的化零多项式。

命题 9.2 线性变换的对应于不同特征值的特征向量线性无关。

证 设$\lambda_1,\cdots,\lambda_s$是线性变换$\mathscr{A}$的两两不同特征值，对应的特征向量分别是$\alpha_1,\cdots,\alpha_s$，在一组基$\varepsilon_1,\cdots,\varepsilon_n$下，它们的坐标分别是$X_1,\cdots,X_s$。$\mathscr{A}$在基$\varepsilon_1,\cdots,\varepsilon_n$下矩阵为$A$，则$X_1,\cdots,X_s$，分别是矩阵$A$的对应于特征值$\lambda_1,\cdots,\lambda_s$的特征向量，故线性无关。若$\sum\limits_{i=1}^{s}k_i\alpha_s=0$，则坐标$\sum\limits_{i=1}^{s}k_iX_i=0$，因$X_1,\cdots,X_s$线性无关，必$k_1=\cdots=k_s=0$，即$\alpha_1,\cdots,\alpha_s$线性无关。

命题 9.3 设$\mathscr{A}$是n维线性空间V的一个线性变换。则存在V的

一组基 $\varepsilon_1,\cdots,\varepsilon_n$，$\mathscr{A}$ 在该基下的矩阵为对角阵$\Leftrightarrow\mathscr{A}$ 有 n 个线性无关的特征向量。

证 $\mathscr{A}$ 有 n 个线性无关的特征向量 $\varepsilon_1,\cdots,\varepsilon_n$，$\mathscr{A}(\varepsilon_i)=\lambda_i\varepsilon_i$，$i=1,\cdots,n$。$\Leftrightarrow\mathscr{A}$ 在基 $\varepsilon_1,\cdots,\varepsilon_n$ 下的矩阵为对角阵

$$\begin{pmatrix}\lambda_1 & & \\ & \ddots & \\ & & \lambda_n\end{pmatrix}。$$

推论 n 维线性空间 V 的线性变换 $\mathscr{A}$ 若有 n 个不同特征值，则存在 V 的基使 $\mathscr{A}$ 在该基下的矩阵为对角阵。

定义 9.6 设 λ 是线性变换 $\mathscr{A}$ 的一个特征值，记 $V_\lambda=\{\alpha\in V|\mathscr{A}(\alpha)=\lambda\alpha\}$。易验证 V_λ 是 V 的子空间，称 V_λ 为 $\mathscr{A}$ 的对应于特征值 λ 的特征子空间。

例 9.14 若线性变换 $\mathscr{A}$ 满足 $\mathscr{A}^2=\mathscr{A}$，求证：

$$V=V_0\oplus V_1$$

证 若 $\alpha\in V_0\cap V_1$，则 $\mathscr{A}(\alpha)=0\cdot\alpha$，$\mathscr{A}(\alpha)=1\cdot\alpha$，故 $\alpha=1\cdot\alpha=0\cdot\alpha=0$，即 $V_0\cap V_1=\{0\}$，所以和 V_0+V_1 是直和。

任意 $\alpha\in V$，$\alpha=\mathscr{A}(\alpha)+[\alpha-\mathscr{A}(\alpha)]$，$\mathscr{A}(\mathscr{A}(\alpha))=\mathscr{A}^2(\alpha)=\mathscr{A}(\alpha)$，所以 $\mathscr{A}(\alpha)\in V_1$。又 $\mathscr{A}[\alpha-\mathscr{A}(\alpha)]=\mathscr{A}(\alpha)-\mathscr{A}^2(\alpha)=0$，即 $\alpha-\mathscr{A}(\alpha)\in V_0$。故 $V\subseteq V_1+V_0$，因而 $V=V_0\oplus V_1$。

§3 不变子空间

一、不变子空间

定义 9.7 设 $\mathscr{A}$ 是线性空间 V 的一个线性变换，V 的子空间 W 称为 $\mathscr{A}$ 的不变子空间，如果任意 $\alpha\in W$，必 $\mathscr{A}(\alpha)\in W$。

例 9.15 $\{0\}$，V 是 $\mathscr{A}$ 的不变子空间。

例 9.16 $\mathscr{A}(V)=\{\mathscr{A}(\alpha)|\alpha\in V\}$ 是 $\mathscr{A}$ 的不变子空间。

例 9.17 若 λ 是$\mathscr{A}$ 的一个特征值，则 V_λ 是$\mathscr{A}$ 的不变子空间。

设$\mathscr{A}$ 是线性空间 V 的一个线性变换，$W_1,\cdots,W_s$ 是$\mathscr{A}$ 的不变子空间，$V=W_1\oplus\cdots\oplus W_s$，把 $W_1,\cdots,W_s$ 的基合并成 V 的一组基，则在该基下$\mathscr{A}$ 的矩阵为准对角阵 $\begin{pmatrix} A_1 & & \\ & \ddots & \\ & & A_s \end{pmatrix}$，$A_i$ 为 n_i 阶方阵，$n_i=\dim W_i$，$(i=1,\cdots,s)$。

二、线性变换的值域与核

定义 9.8 设$\mathscr{A}$ 是线性空间 V 的一个线性变换，记

$$\mathscr{A}(V)=\{\mathscr{A}(\alpha)\,|\,\alpha\in V\},$$

它是$\mathscr{A}$ 的不变子空间，称为$\mathscr{A}$ 的值域，且称 $\dim[\mathscr{A}(V)]$为$\mathscr{A}$ 的秩。

推论 线性变换$\mathscr{A}$ 的秩等于它在任一组基下矩阵的秩。

证 设 $\varepsilon_1,\cdots,\varepsilon_n$ 为一组基，则 $\mathscr{A}(V)=L(\mathscr{A}(\varepsilon_1),\cdots,\mathscr{A}(\varepsilon_n))$。于是 $\dim[\mathscr{A}(V)]=$ 向量组$\mathscr{A}(\varepsilon_1),\cdots,\mathscr{A}(\varepsilon_n)$ 的秩。

又 $(\mathscr{A}(\varepsilon_1),\cdots,\mathscr{A}(\varepsilon_n))=(\varepsilon_1,\cdots,\varepsilon_n)A$，向量组 $\mathscr{A}(\varepsilon_1),\cdots,\mathscr{A}(\varepsilon_n)$ 的秩等于它们坐标列的秩，即矩阵 A 的列秩，故得证明。

定义 9.9 设$\mathscr{A}$ 是线性空间 V 的线性变换，记

$$\operatorname{Ker}\mathscr{A}=\{\alpha\in V\,|\,\mathscr{A}(\alpha)=0\},$$

它是$\mathscr{A}$ 的不变子空间，称为$\mathscr{A}$ 的核，且称 $\dim(\operatorname{Ker}\mathscr{A})$为$\mathscr{A}$ 的零度。

推论 $\mathscr{A}$ 为单射$\Leftrightarrow\mathscr{A}$ 的零度为 0。

证 $\Rightarrow\mathscr{A}$ 为单射，$\mathscr{A}(0)=0$，$\alpha\neq0$，$\mathscr{A}(\alpha)\neq0$，即 $\operatorname{Ker}\mathscr{A}=\{0\}$，所以$\mathscr{A}$ 的零度为 0。

$\Leftarrow\operatorname{Ker}\mathscr{A}=\{0\}$，若 $\mathscr{A}(\alpha)=\mathscr{A}(\beta)$，则 $\mathscr{A}(\alpha-\beta)=0$，$\alpha-\beta\in\operatorname{Ker}\mathscr{A}=\{0\}$，所以 $\alpha-\beta=0$，$\alpha=\beta$，即$\mathscr{A}$ 为单射。

例 9.18 若 $V=P[X]_n$，则 $D(V)=P[X]_{n-1}$，$\operatorname{Ker}D=P$。

例 9.19 若 $V=P^n$，$A\in P^{n\times n}$，定义 V 的线性变换 σ：

$$\sigma(\alpha)=A\alpha,$$

则

$$\mathrm{Ker}\sigma = \{\alpha \in V \mid A\alpha = 0\}$$

即是齐次线性方程组 $AX=0$ 的解空间。

又

$$\sigma(V) = \left\{ A\alpha \,\middle|\, \alpha = \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} \in V \right\}$$

$$= \left\{ \sum_{i=1}^{n} a_i A^{(i)} \,\middle|\, a_i \in P, i = 1, \cdots, n \right\}$$

$$= L(A^{(1)}, \cdots, A^{(n)})$$

即矩阵 A 的列向量生成的 P^n 的子空间。

例 9.20 设 W 是内积空间 V 的子空间，Π 为 V 到 W 的投影变换，则 $\Pi(V) = W$，$\mathrm{Ker}\Pi = W^{\perp}$。

定理 9.3 设 $\mathscr{A}$ 是线性空间 V 的一个线性变换。则

$$\mathscr{A} \text{ 的秩} + \mathscr{A} \text{ 的零度} = \dim V \text{。}$$

证 设 $\varepsilon_1, \cdots, \varepsilon_t$ 为 $\mathrm{Ker}\mathscr{A}$ 的一组基，于是 $\mathscr{A}$ 的零度 $=t$。扩充 $\varepsilon_1, \cdots, \varepsilon_t$ 使 $\varepsilon_1, \cdots, \varepsilon_t, \varepsilon_{t+1}, \cdots, \varepsilon_n$ 为 V 的一组基，于是 $\dim V = n$。

$\mathscr{A}(V) = L(\mathscr{A}(\varepsilon_1), \cdots, \mathscr{A}(\varepsilon_t), \mathscr{A}(\varepsilon_{t+1}), \cdots, \mathscr{A}(\varepsilon_n)) = L(\mathscr{A}(\varepsilon_{t+1}), \cdots, \mathscr{A}(\varepsilon_n))$。从而 $\mathscr{A}$ 的秩 $=$ 向量组 $\mathscr{A}(\varepsilon_{t+1}), \cdots, \mathscr{A}(\varepsilon_n)$ 的秩。若能证明 $\mathscr{A}(\varepsilon_{t+1}), \cdots, \mathscr{A}(\varepsilon_n)$ 线性无关，则 $\mathscr{A}$ 的秩 $=$ 向量个数 $(n-t) = \dim V - \dim \mathrm{Ker}\,\mathscr{A}$，即得定理的证明。向量组 $\mathscr{A}(\varepsilon_{t+1}), \cdots, \mathscr{A}(\varepsilon_n)$ 线性无关的证明如下：

若 $\sum\limits_{i=t+1}^{n} k_i \mathscr{A}(\varepsilon_i) = 0$，则 $\mathscr{A}(\sum\limits_{i=t+1}^{n} k_i \varepsilon_i) = 0$，于是 $\alpha = \sum\limits_{i=t+1}^{n} k_i \varepsilon_i \in \mathrm{Ker}\,\mathscr{A}$。$\varepsilon_1, \cdots, \varepsilon_t$ 是 $\mathrm{Ker}\mathscr{A}$ 的一组基，因而存在数 $a_1, \cdots, a_t$ 使 $\alpha = \sum\limits_{i=1}^{t} a_i \varepsilon_i$。此时

$$0 = \alpha - \alpha = \sum_{i=t+1}^{n} k_i \varepsilon_i = \sum_{i=1}^{t} a_i \varepsilon_i$$

$$= \sum_{i=1}^{t}(-a_i)\varepsilon_i + \sum_{i=t+1}^{n} k_i\varepsilon_i,$$

因为 $\varepsilon_1,\cdots,\varepsilon_t,\varepsilon_{t+1},\cdots,\varepsilon_n$ 是 V 的一组基，它们线性无关，故必 $k_{t+1}=\cdots=k_n=0$。所以 $\mathscr{A}(\varepsilon_{t+1}),\cdots,\mathscr{A}(\varepsilon_n)$ 线性无关。

推论 线性变换 $\mathscr{A}$ 是单射 $\Leftrightarrow$ $\mathscr{A}$ 是满射 $\Leftrightarrow$ $\mathscr{A}$ 是双射。

证 $\mathscr{A}$ 是单射 $\Leftrightarrow$ $\mathscr{A}$ 的零度为 $0 \Leftrightarrow$ $\mathscr{A}$ 的秩 $=\dim V \Leftrightarrow \dim\mathscr{A}(V)=\dim V \Leftrightarrow \mathscr{A}(V)=V \Leftrightarrow$ $\mathscr{A}$ 是满射。

三、根子空间

定义 9.10 若 λ 是线性空间 V 的线性变换 $\mathscr{A}$ 的一个特征值，记 $\boldsymbol{R}_\lambda=\{\alpha\in V \mid 存在\ m\in \mathbf{Z}^+\ 使(\lambda\boldsymbol{I}-\mathscr{A})^m\alpha=\boldsymbol{0}\}$，则 $\boldsymbol{R}_\lambda$ 是 $\mathscr{A}$ 的不变子空间，称为 $\mathscr{A}$ 的对应于特征值 λ 的根子空间。

证 $\alpha,\beta\in\boldsymbol{R}_\lambda$，则存在 $m,s\in\mathbf{Z}^+$ 使 $(\lambda\boldsymbol{I}-\mathscr{A})^m(\alpha)=\boldsymbol{0}$，$(\lambda\boldsymbol{I}-\mathscr{A})^s(\beta)=\boldsymbol{0}$，记 $t=\max(m,s)$，此时

$$\begin{aligned}(\lambda\boldsymbol{I}-\mathscr{A})^t(k\alpha+l\beta)&=k(\lambda\boldsymbol{I}-\mathscr{A})^t\alpha+l(\lambda\boldsymbol{I}-\mathscr{A})^t\beta\\&=\boldsymbol{0}+\boldsymbol{0}=\boldsymbol{0}\end{aligned}$$

即

$$k\alpha+l\beta\in\boldsymbol{R}_\lambda,$$

故 $\boldsymbol{R}_\lambda$ 是子空间。

又 $\mathscr{A}$ 与 $(\lambda\boldsymbol{I}-\mathscr{A})$ 可交换，故

$$(\lambda\boldsymbol{I}-\mathscr{A})^m[\mathscr{A}(\alpha)]=\mathscr{A}[(\lambda\boldsymbol{I}-\mathscr{A})^m(\alpha)]=\mathscr{A}(\boldsymbol{0})=\boldsymbol{0},$$

即 $\mathscr{A}(\alpha)\in\boldsymbol{R}_\lambda$，所以 $\boldsymbol{R}_\lambda$ 是 $\mathscr{A}$ 的不变子空间。

例 9.21 设 $\boldsymbol{D}$ 为 $V=\boldsymbol{P}[X]_n$ 中的微分变换：$\boldsymbol{D}(f(x))=f'(x)$。$\boldsymbol{D}$ 有特征值 0，对应的根子空间为 V 本身。因为任意 $f(x)\in\boldsymbol{P}[x]_n$，都有 $f^{(n)}(x)=0$。即

$$(0\boldsymbol{I}-\boldsymbol{D})^n(f(x))=-\boldsymbol{D}^n(f(x))=-f^{(n)}(x)=0。$$

定理 9.4 若线性空间 V 的线性变换 $\mathscr{A}$ 的特征多项式为

$$f(\lambda) = (\lambda - \lambda_1)^{t_1} \cdots (\lambda - \lambda_s)^{t_s},$$

这里 $\lambda_1, \cdots, \lambda_s$ 两两不同,则

$$\boldsymbol{V} = \boldsymbol{R}_{\lambda_1} \oplus \cdots \oplus \boldsymbol{R}_{\lambda_s},$$

即 $\boldsymbol{V}$ 为根子空间的直和。

证 记 $f_i(\lambda) = \dfrac{f(\lambda)}{(\lambda - \lambda_i)^{t_i}}$, $W_i = f_i(\mathscr{A})(V), i = 1, \cdots, s$。

因为特征多项式是$\mathscr{A}$ 的化零多项式,所以 $f(\mathscr{A}) = \boldsymbol{0}$ 。于是

$$(\mathscr{A} - \lambda_i \boldsymbol{I})^{t_i} \boldsymbol{W}_i = f(\mathscr{A})(\boldsymbol{V}) = \boldsymbol{0} ,$$

故

$$\boldsymbol{W}_i \subseteq \boldsymbol{R}_{\lambda_i}。$$

由于 $(f_1(\lambda), \cdots, f_s(\lambda)) = 1$, 所以存在多项式 $u_1(\lambda), \cdots, u_s(\lambda)$使

$$u_1 f_1 + \cdots + u_s f_s = 1。$$

从而,对任意 $\alpha \in V$ 都有

$$\begin{aligned} \alpha = I(\alpha) &= [f_1(\mathscr{A}) u_1(\mathscr{A}) + \cdots + f_s(\mathscr{A}) u_s(\mathscr{A})](\alpha) \\ &= f_1(\mathscr{A})[u_1(\mathscr{A})(\alpha)] + \cdots + f_s(\mathscr{A})[u_s(\mathscr{A})(\alpha)]。 \end{aligned}$$

因为

$$f_i(\mathscr{A})[u_i(\mathscr{A})(\alpha)] \in f_i(\mathscr{A})(\boldsymbol{V}) = \boldsymbol{W}_i, i = 1, \cdots, s$$

故

$$\boldsymbol{V} \subseteq \boldsymbol{W}_1 + \cdots + \boldsymbol{W}_s \subseteq \boldsymbol{R}_{\lambda_1} + \cdots + \boldsymbol{R}_{\lambda_s},$$

从而

$$\boldsymbol{V} = \boldsymbol{R}_{\lambda_1} + \cdots + \boldsymbol{R}_{\lambda_s}。$$

又若

$$\boldsymbol{\alpha} = \boldsymbol{\alpha}_1 + \cdots + \boldsymbol{\alpha}_s = \boldsymbol{\beta}_1 + \cdots + \boldsymbol{\beta}_s, \boldsymbol{\alpha}_i, \boldsymbol{\beta}_1 \in \boldsymbol{R}_{\lambda_i}, i = 1, \cdots, s$$

由于 $\boldsymbol{\alpha}_j - \boldsymbol{\beta}_j \in \boldsymbol{R}_{\lambda_j}$,故存在 $m_j \in \mathbf{Z}^+$,且

$$(\mathscr{A} - \lambda_j \boldsymbol{I})^{m_j} (\boldsymbol{\alpha}_j - \boldsymbol{\beta}_j) = \boldsymbol{0} , j = 1, \cdots, n$$

令
$$\gamma = \alpha_i - \beta_i, 则 (\mathscr{A} - \lambda_i I)^{m_i}(\gamma) = \boldsymbol{0},$$
又
$$\gamma = \sum_{j \neq i} (\beta_j - \alpha_j),$$
故
$$\prod_{j \neq i} (\mathscr{A} - \lambda_j I)^{m_j}(\gamma) = \boldsymbol{0}。$$
因为
$$\left((x - \lambda_i)^{m_i}, \prod_{j \neq i} (x - \lambda_j)^{m_j}\right) = 1,$$
故存在多项式 $a(x)$ 与 $b(x)$ 使
$$a(x)(x - \lambda_i)^{m_i} + b(x) \prod_{j \neq i} (x - \lambda_j)^{m_j} = 1,$$
于是
$$\begin{aligned}\gamma = I(\gamma) &= \left[a(\mathscr{A})(\mathscr{A} - \lambda_i I)^{m_i} + b(\mathscr{A}) \prod_{j \neq i} (\mathscr{A} - \lambda_j I)^{m_j}\right](\gamma) \\ &= a(\mathscr{A})(\mathscr{A} - \lambda_i I)^{m_i}(\gamma) + b(\mathscr{A}) \cdot \prod_{j \neq i} (\mathscr{A} - \lambda_j I)^{m_j}(\gamma) \\ &= \boldsymbol{0} + \boldsymbol{0} = \boldsymbol{0},\end{aligned}$$
即
$$\alpha_i - \beta_i = \boldsymbol{0}, 或 \alpha_i = \beta_i, i = 1, \cdots, s$$
所以和 $\boldsymbol{R}_{\lambda_1} + \cdots + \boldsymbol{R}_{\lambda_s}$ 是直和。从而
$$\boldsymbol{V} = \boldsymbol{R}_{\lambda_1} \oplus \cdots \oplus \boldsymbol{R}_{\lambda_s}。$$

推论 1 $\boldsymbol{V} = \boldsymbol{W}_1 \oplus \cdots \oplus \boldsymbol{W}_s$。

证 因为 $\boldsymbol{V} = \boldsymbol{R}_{\lambda_1} \oplus \cdots \oplus \boldsymbol{R}_{\lambda_s}$，故
$$\dim \boldsymbol{V} = \sum_{i=1}^{s} \dim \boldsymbol{R}_{\lambda_i} \geqslant \sum_{i=1}^{s} \dim \boldsymbol{W}_i \geqslant \dim\left(\sum_{i=1}^{s} \boldsymbol{W}_i\right) = \dim \boldsymbol{V},$$
由于 $\dim \boldsymbol{V} = \dim \boldsymbol{V}$，故上列不等式中大于等于号全取等号，于是
$$\sum_{i=1}^{s} \dim \boldsymbol{W}_i = \dim\left(\sum_{i=1}^{s} \boldsymbol{W}_i\right),$$
从而和 $V = \sum_{i=1}^{s} \boldsymbol{W}_i$ 是直和，且

$$W_i = R_{\lambda_i}。$$

推论 2 $(\mathscr{A} - \lambda_i I)^{t_i} R_{\lambda_i} = (\mathscr{A} - \lambda_i I)^{t_i} W_i = f(\mathscr{A})(V) = \mathbf{0}$ 。

例 9.22 在 $V = R^4$ 中定义线性变换 σ:

$$\sigma(\alpha) = \begin{pmatrix} 1 & 1 & 2 & 1 \\ 0 & 1 & 3 & -1 \\ 0 & 0 & -1 & 2 \\ 0 & 0 & 0 & -1 \end{pmatrix} \alpha = A\alpha,$$

求 σ 的所有根子空间。

解 由例 9.9,线性变换 σ 在自然基 $\varepsilon_1,\varepsilon_2,\varepsilon_3,\varepsilon_4$ 下矩阵为

$$A = \begin{pmatrix} 1 & 1 & 2 & 1 \\ 0 & 1 & 3 & -1 \\ 0 & 0 & -1 & 2 \\ 0 & 0 & 0 & -1 \end{pmatrix},$$

σ 的特征多项式

$$f(\lambda) = |\lambda E - A| = (\lambda - 1)^2(\lambda + 1)^2,$$

故 σ 的特征值为 1,-1,有 2 个根子空间 R_1, R_{-1}:

$$\begin{aligned} R_1 &= \left(\left. \frac{f(\lambda)}{(\lambda - 1)^2} \right|_{\lambda=\sigma} \right) V = (\sigma + I)^2 V \\ &= L[(\sigma + I)^2\varepsilon_1, (\sigma + I)^2\varepsilon_2, (\sigma + I)^2\varepsilon_3, (\sigma + I)^2\varepsilon_4] \\ &= L(A + E)^2\varepsilon_1, (A + E)^2\varepsilon_2, (A + E)^2\varepsilon_3, (A + E)^2\varepsilon_4] \\ &= (A + E)^2 \text{ 的列空间}, \end{aligned}$$

而

$$(A + E)^2 = \begin{pmatrix} 2 & 1 & 2 & 1 \\ 0 & 2 & 3 & -1 \\ 0 & 0 & 0 & 2 \\ 0 & 0 & 0 & 0 \end{pmatrix}^2 = \begin{pmatrix} 4 & 6 & 7 & 5 \\ 0 & 4 & 6 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix},$$

于是

$$\boldsymbol{R}_1 = L\left[\begin{pmatrix}4\\0\\0\\0\end{pmatrix},\begin{pmatrix}6\\4\\0\\0\end{pmatrix},\begin{pmatrix}7\\6\\0\\0\end{pmatrix},\begin{pmatrix}5\\1\\0\\0\end{pmatrix}\right] = L(\boldsymbol{\varepsilon}_1,\boldsymbol{\varepsilon}_2);$$

类似地

$$\boldsymbol{R}_{-1} = (\boldsymbol{A}-\boldsymbol{E})^2 \text{ 的列空间},$$

由于

$$(\boldsymbol{A}-\boldsymbol{E})^2 = \begin{pmatrix}0&0&-1&1\\0&0&-6&8\\0&0&4&-8\\0&0&0&4\end{pmatrix},$$

故

$$\boldsymbol{R}_{-1} = L\left[\begin{pmatrix}-1\\-6\\4\\0\end{pmatrix},\begin{pmatrix}1\\8\\-8\\4\end{pmatrix}\right]。$$

§4 正 规 变 换

本节讨论内积空间的特殊线性变换，即正规变换。我们将证明，内积空间一定存在标准正交基，使某给定的正规变换在该基下的矩阵为对角阵，并将就两类常见正规变换（酉变换与 Hermite 变换）展开较详细的讨论（本节的内容是线性变换与矩阵间一一对应的一个具体例子）。

本节中出现的线性空间 $\boldsymbol{V}$ 都假定是数域 $\boldsymbol{P}$ 上的内积空间。

一、关联变换

定义 9.11 若$\mathscr{A}$是内积空间 $\boldsymbol{V}$ 的一个线性变换，$\boldsymbol{V}$ 的线性变换$\mathscr{B}$称为$\mathscr{A}$的关联变换，记成$\mathscr{B}=\mathscr{A}^*$，如果对于任意 $\alpha,\beta\in \boldsymbol{V}$，恒有

$$[\mathscr{A}(\alpha),\beta]=[\alpha,\mathscr{B}(\beta)]。$$

判别 若线性变换$\mathscr{A}$与$\mathscr{B}$在内积空间V的一组标准正交基$\varepsilon_1,\cdots,\varepsilon_n$下的矩阵分别为$A$与$B$,则

$$\mathscr{B}=\mathscr{A}^*,$$

即$\mathscr{B}$是$\mathscr{A}$的关联变换

$$\Leftrightarrow B=A^*,$$

即B是A的关联阵。

证 任意$\alpha,\beta\in V$,设α,β在基$\varepsilon_1,\cdots,\varepsilon_n$下的坐标分别为$X$与$Y$,即$\alpha=(\varepsilon_1,\cdots,\varepsilon_n)X,\beta=(\varepsilon_1,\cdots,\varepsilon_n)Y$,则$\mathscr{A}(\alpha)$与$\mathscr{B}(\beta)$的坐标分别为$AX$与$BY$。

$$\Rightarrow \mathscr{B}=\mathscr{A}^*,则[\mathscr{A}(\alpha),\beta]=[\alpha,\mathscr{B}(\beta)],$$

由在标准正交基下向量的内积公式得

$$(AX)^*Y=X^*(BY),$$

或

$$X^*A^*Y=X^*BY。$$

令$\alpha=\varepsilon_i,\beta=\varepsilon_j$,得

$$A^*(i,j)=B(i,j),\quad i,j=1,\cdots,n$$

故

$$A^*=B,$$

$$\Leftarrow B=A^*,$$

此时

$$\begin{aligned}[\mathscr{A}(\alpha),\mathscr{B}]&=(AX)^*Y=X^*A^*Y\\&=X^*(\mathscr{B}Y)=[\alpha,\mathscr{B}(\beta)]。\end{aligned}$$

推论 内积空间的任意线性变换的关联变换都存在且惟一。

证 设$\mathscr{A}$为内积空间V的一个线性变换,$\varepsilon_1,\cdots,\varepsilon_n$为$V$的一组标准正交基。$\mathscr{A}$在基$\varepsilon_1,\cdots,\varepsilon_n$下的矩阵为$A$。定义映射

$$V \xrightarrow{\mathscr{B}} V:$$

$$\mathscr{B}\left((\varepsilon_1,\cdots,\varepsilon_n)\begin{pmatrix}a_1\\ \vdots\\ a_n\end{pmatrix}\right)=(\varepsilon_1,\cdots,\varepsilon_n)A^*\begin{pmatrix}a_1\\ \vdots\\ a_n\end{pmatrix}$$

由定理 9.1 及其证明可知，$\mathscr{B}$ 是 V 的线性变换，$\mathscr{B}$ 在基 $\varepsilon_1,\cdots,\varepsilon_n$ 下的矩阵为 A^*，故由上述判别法，$\mathscr{B}$ 为 $\mathscr{A}$ 的关联变换。

若 $\mathscr{A}$ 另有关联变换 $\mathscr{B}_1$，则 $\mathscr{B}_1$ 在基 $\varepsilon_1,\cdots,\varepsilon_n$ 下的矩阵也是 A^*，由定理 9.1，$\mathscr{B}=\mathscr{B}_1$，故关联变换惟一。

二、正规变换

定义 9.12 内积空间 V 的线性变换 $\mathscr{A}$ 称为正规的，如果

$$\mathscr{A}\mathscr{A}^*=\mathscr{A}^*\mathscr{A}$$

判别 内接空间 V 的线性变换 $\mathscr{A}$ 是正规的；$\Longleftrightarrow$ $\mathscr{A}$ 在 V 的标准正交基下矩阵 A 是正规的。

证 $\mathscr{A}^*$ 在同一标准正交基下矩阵为 A^*。由定理 9.1，$\mathscr{A}\mathscr{A}^*$ 与 $\mathscr{A}^*\mathscr{A}$ 所对应矩阵分别为 AA^* 与 A^*A。由定理 9.1，这个对应是双射，故

$$\mathscr{A}\mathscr{A}^*=\mathscr{A}^*\mathscr{A}\Longleftrightarrow AA^*=A^*A,$$

即线性变换 $\mathscr{A}$ 是正规变换的充要条件是对应矩阵 A（在标准正交基下）是正规阵。

定理 9.5 若 $\mathscr{A}$ 是内积空间 V 的正规变换，则存在 V 的标准正交基，使 $\mathscr{A}$ 在该基下的矩阵为对角阵

$$\begin{pmatrix}\lambda_1 & & \\ & \ddots & \\ & & \lambda_n\end{pmatrix},$$

这里 $\lambda_i(i=1,\cdots,n)$ 为 $\mathscr{A}$ 的特征值。

证 设 $\varepsilon_1,\cdots,\varepsilon_n$ 为 V 的一组标准正交基，$\mathscr{A}$ 在该基下矩阵为 A。

因为$\mathscr{A}$为正规变换，故矩阵$\boldsymbol{A}$为正规阵。由定理5.9

$$\boldsymbol{A} \overset{U}{\sim} \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix},$$

即存在酉阵$\boldsymbol{T}$使

$$\boldsymbol{T}^{-1}\boldsymbol{A}\boldsymbol{T} = \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix}。$$

作基变换

$$(\eta_1, \cdots, \eta_n) = (\varepsilon_1, \cdots, \varepsilon_n)\boldsymbol{T},$$

由定理8.10，$\eta_1, \cdots, \eta_n$为$\boldsymbol{V}$的标准正交基。由定理9.2的证明可知，线性变换$\mathscr{A}$在标准正交基$\eta_1, \cdots, \eta_n$下的矩阵为

$$\boldsymbol{T}^{-1}\boldsymbol{A}\boldsymbol{T} = \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix}。$$

三、酉变换

定义 9.13 内积空间$\boldsymbol{V}$的线性变换$\mathscr{A}$称为酉变换，如果$\mathscr{A}$保持向量的内积不变，即任意$\alpha, \beta \in \boldsymbol{V}$有

$$[\mathscr{A}(\alpha), \mathscr{A}(\beta)] = [\alpha, \beta]。$$

欧氏空间，即实数域上的内积空间的酉变换常称为正交变换。

例 9.23 平面$\boldsymbol{V} = \boldsymbol{R}^2$的旋转变换$\sigma$：

$$\sigma\left(\begin{pmatrix} x \\ y \end{pmatrix}\right) = \begin{pmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = \boldsymbol{A}\begin{pmatrix} x \\ y \end{pmatrix}$$

是正交变换。

证 因为

$$\left[\boldsymbol{\sigma}\left(\begin{pmatrix}a_1\\a_2\end{pmatrix}\right),\boldsymbol{\sigma}\left(\begin{pmatrix}b_1\\b_2\end{pmatrix}\right)\right]=\left[\boldsymbol{A}\begin{pmatrix}a_1\\a_2\end{pmatrix},\boldsymbol{A}\begin{pmatrix}b_1\\b_2\end{pmatrix}\right]$$

$$=(a_1,a_2)\boldsymbol{A}'\boldsymbol{A}\begin{pmatrix}b_1\\b_2\end{pmatrix}=(a_1,a_2)\boldsymbol{E}\begin{pmatrix}b_1\\b_2\end{pmatrix}$$

$$=(a_1,a_2)\begin{pmatrix}b_1\\b_2\end{pmatrix}=\left[\begin{pmatrix}a_1\\a_2\end{pmatrix},\begin{pmatrix}b_1\\b_2\end{pmatrix}\right],$$

这里 $\boldsymbol{A}=\begin{pmatrix}\cos\varphi & -\sin\varphi\\ \sin\varphi & \cos\varphi\end{pmatrix}$ 是正交阵，$\boldsymbol{A}'\boldsymbol{A}=\boldsymbol{E}$。

例 9.24 平面 $\boldsymbol{R}^2$ 关于 X 轴的镜面反射 σ：$\boldsymbol{\sigma}\left(\begin{pmatrix}x\\y\end{pmatrix}\right)=\begin{pmatrix}x\\-y\end{pmatrix}$ 是正交变换。

定理 9.5 若 $\mathscr{A}$ 是数域 $\boldsymbol{P}$ 上内积空间 $\boldsymbol{V}$ 的一个线性变换，则下列命题彼此等价：

(1) $\mathscr{A}$ 是酉变换；

(2) 对任意 $\alpha\in\boldsymbol{V}$，$\|\mathscr{A}(\alpha)\|=\|\alpha\|$，即 $\mathscr{A}$ 保持向量长度不变；

(3) 若 $\varepsilon_1,\cdots,\varepsilon_n$ 是 $\boldsymbol{V}$ 的标准正交基，则 $\mathscr{A}(\varepsilon_1),\cdots,\mathscr{A}(\varepsilon_n)$ 也是 $\boldsymbol{V}$ 的标准正交基，即 $\mathscr{A}$ 把标准正交基变成标准正交基；

(4) $\mathscr{A}$ 在标准正交基下的矩阵为酉阵，从而是正规变换。

证 (1) $\Rightarrow$ (2)：因为

$$\|\mathscr{A}(\alpha)\|=\sqrt{[\mathscr{A}(\alpha),\mathscr{A}(\alpha)]}=\sqrt{[\alpha,\alpha]}=\|\alpha\|。$$

(2) $\Rightarrow$ (3)：$\|\mathscr{A}(\varepsilon_k)\|=\|\varepsilon_k\|=1$。$k\neq j$ 时，由

$$[\mathscr{A}(\varepsilon_k+\varepsilon_j),\mathscr{A}(\varepsilon_k+\varepsilon_j)]=[(\varepsilon_k+\varepsilon_j),(\varepsilon_k+\varepsilon_j)]=2,$$

易得

$$[\mathscr{A}(\varepsilon_k),\mathscr{A}(\varepsilon_j)]+[\mathscr{A}(\varepsilon_j),\mathscr{A}(\varepsilon_k)]=0。\qquad(9.1)$$

若 $P\subseteq R$，由式(9.1)及 $[\mathscr{A}(\varepsilon_k),\mathscr{A}(\varepsilon_j)]=[\mathscr{A}(\varepsilon_j),\mathscr{A}(\varepsilon_k)]$ 即得 $[\mathscr{A}(\varepsilon_k),\mathscr{A}(\varepsilon_j)]=0$。若存在 $a\in P$，$a\overline{\in}\mathbf{R}$，此时 $\bar{a}\neq a$。由

$$[\mathscr{A}(\varepsilon_k+a\varepsilon_j),\mathscr{A}(\varepsilon_k+a\varepsilon_j)]=[\varepsilon_k+a\varepsilon_j,\varepsilon_k+a\varepsilon_j],$$

可得

$$a[\mathscr{A}(\varepsilon_k),\mathscr{A}(\varepsilon_j)]+\bar{a}[\mathscr{A}(\varepsilon_j),\mathscr{A}(\varepsilon_k)]=0, \tag{9.2}$$

由 $\bar{a}\times$式(9.1)－式(9.2)得

$$(\bar{a}-a)[\mathscr{A}(\varepsilon_k),\mathscr{A}(\varepsilon_j)]=0。$$

因为 $\bar{a}-a\neq0$,必 $[\mathscr{A}(\varepsilon_k),\mathscr{A}(\varepsilon_j)]=0$, 于是,$\mathscr{A}(\varepsilon_1),\cdots,\mathscr{A}(\varepsilon_n)$ 是标准正交基。

(3) $\Rightarrow$(1): 设 $\alpha=(\varepsilon_1,\cdots,\varepsilon_n)\begin{pmatrix}a_1\\ \vdots\\ a_n\end{pmatrix},\beta=(\varepsilon_1,\cdots,\varepsilon_n)\begin{pmatrix}b_1\\ \vdots\\ b_n\end{pmatrix}$,则

$$\mathscr{A}(\alpha)=(\mathscr{A}(\varepsilon_1),\cdots,\mathscr{A}(\varepsilon_n))\begin{pmatrix}a_1\\ \vdots\\ a_n\end{pmatrix},$$

$$\mathscr{A}(\beta)=(\mathscr{A}(\varepsilon_1),\cdots,\mathscr{A}(\varepsilon_n))\begin{pmatrix}b_1\\ \vdots\\ b_n\end{pmatrix}。$$

由于 $\varepsilon_1,\cdots,\varepsilon_n$ 与 $\mathscr{A}(\varepsilon_1),\cdots,\mathscr{A}(\varepsilon_n)$ 都是标准正交基,故

$$[\mathscr{A}(\alpha),\mathscr{A}(\beta)]=\sum_{i=1}^{n}\bar{a}_ib_i=[\alpha,\beta]。$$

(3)$\Leftrightarrow$(4): 若 $\mathscr{A}$ 在基 $\varepsilon_1,\cdots,\varepsilon_n$ 下的矩阵为 A, 由定理 8.10, $\mathscr{A}(\varepsilon_1),\cdots,\mathscr{A}(\varepsilon_n)$ 是标准正交基 $\Leftrightarrow A$ 是酉阵。

命题 9.4 酉变换 $\mathscr{A}$ 的特征值的模全为 1。

证 设 λ 为 $\mathscr{A}$ 的特征值,α 为对应的特征向量,则

$$[\alpha,\alpha]=[\mathscr{A}(\alpha),\mathscr{A}(\alpha)]=[\lambda\alpha,\lambda\alpha]=\bar{\lambda}\lambda[\alpha,\alpha]。$$

因为 $\alpha\neq0,[\alpha,\alpha]\neq0$,故 $|\lambda|^2=\bar{\lambda}\lambda=1$。

由于酉变换是正规变换,即得下面定理:

定理 9.6 设 $\mathscr{A}$ 是内积空间 V 的酉变换,则存在 V 的标准正交基,使 $\mathscr{A}$ 在该基下的矩阵为对角阵:

$$\begin{pmatrix} \lambda & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix},$$

这里 $|\lambda_i| = 1, i = 1, \cdots, n$。

四、Hermite 变换

定义 9.14 内积空间 V 的线性变换 $\mathscr{A}$ 称为 Hermite 变换，如果任意 $\alpha,\beta \in V$，恒有 $[\mathscr{A}(\alpha),\beta]=[\alpha,\mathscr{A}(\beta)]$，即 $\mathscr{A}^* = \mathscr{A}$。

Euclid 空间（欧氏空间）的 Hermite 变换常称对称变换。

命题 9.5 Hermite 变换 $\mathscr{A}$ 的特征值全为实数。

证 设 λ 为 $\mathscr{A}$ 的特征值，α 为对应的特征向量，由

$$[\mathscr{A}(\alpha),\alpha] = [\alpha,\mathscr{A}(\alpha)],$$

得

$$[\lambda\alpha,\alpha] = [\alpha,\lambda\alpha],$$

或

$$\bar{\lambda}[\alpha,\alpha] = \lambda[\alpha,\alpha],$$

因为 $[\alpha,\alpha] \neq 0$，故 $\bar{\lambda} = \lambda$，即 λ 为实数。

定理 9.7 内积空间 V 的线性变换 $\mathscr{A}$ 是 Hermite 变换 $\Leftrightarrow \mathscr{A}$ 在标准正交基下矩阵为 Hermite 阵。

证 设 $\mathscr{A}$ 在标准正交基 $\varepsilon_1,\cdots,\varepsilon_n$ 下矩阵为 A，则 $\mathscr{A}$ 的关联变换 $\mathscr{A}^*$ 在基 $\varepsilon_1,\cdots,\varepsilon_n$ 下的矩阵为 A^*，因而 $\mathscr{A}$ 为 Hermite 变换，即

$$\mathscr{A}^* = \mathscr{A}$$

$\Leftrightarrow$（定理 9.1）$A^* = A$，即 A 为 Hermite 阵。

因为 Hermite 变换的关联就是自身，因而是正规变换，且其特征值全为实数。故有下面的定理：

定理 9.8 若 $\mathscr{A}$ 是内积空间 V 的 Hermite 变换，则存在 V 的标准正交基使 $\mathscr{A}$ 在该基下矩阵为实对角阵。

推论 若 $\mathscr{A}$ 是欧氏空间 V 的对称变换，则存在 V 的标准正交基

使$\mathscr{A}$在该基下的矩阵为实对角阵。

习　题　9

1．若 V 是线性空间，判断下面定义的 V 到 V 的映射，哪些是线性的，哪些不是？

(1) $\mathscr{A}(\alpha)=\alpha+\alpha_0$，$\alpha_0$ 是 V 的一个固定非零向量；

(2) $\mathscr{A}(\alpha)=\alpha_0$，$\alpha_0$ 是 V 的一个固定非零向量；

(3) $V=P^3,\mathscr{A}\left(\begin{pmatrix}x_1\\x_2\\x_3\end{pmatrix}\right)=\begin{pmatrix}x_1+x_2\\x_1-x_2\\3x_3+x_1\end{pmatrix}$；

(4) $V=P^3,\mathscr{A}\left(\begin{pmatrix}x_1\\x_2\\x_3\end{pmatrix}\right)=\begin{pmatrix}x_1^2\\x_2-x_3^2\\x_3\end{pmatrix}$；

(5) $V=P^{n\times n},\mathscr{A}(X)=BXC,B,C\in P^{n\times n}$；

(6) $V=P^{n\times n},\mathscr{A}(X)=X'$；

(7) $V=P[x],\mathscr{A}(f(x))=f(x+1)$；

(8) $V=P[x]_n,\mathscr{A}(f(x))=f(x+1)-f(x)$。

2．若 σ 是 P^n 的一个线性变换，$k,l\in P$，求证：$\eta\begin{pmatrix}\alpha\\\beta\end{pmatrix}=\begin{pmatrix}k\sigma(\alpha)\\l\sigma(\beta)\end{pmatrix}$ 是 P^{2n} 的线性变换。$\alpha,\beta\in P^n$。

3．若线性空间 V 的线性变换 σ 满足 $\sigma^s=0,(s\in \mathbf{Z}^+)$，求证：$I+\sigma$ 是 V 到自身的同构映射。

4．在线性空间 $P[x]$ 中，定义 $\mathscr{A}(f(x))=f'(x)$，$\mathscr{B}(f(x))=xf(x)$，求证：$\mathscr{A}\mathscr{B}-\mathscr{B}\mathscr{A}=I$。

5．设 $\mathscr{A},\mathscr{B}$ 是线性变换，如果 $\mathscr{A}\mathscr{B}-\mathscr{B}\mathscr{A}=I$，求证：对任意大于 1 的正整数 k，恒有

$$\mathscr{A}^k\mathscr{B}-\mathscr{B}\mathscr{A}^k=k\mathscr{A}^{k-1}。$$

6．设 $\varepsilon_1,\cdots,\varepsilon_n$ 是线性空间 V 的一组基，σ 是 V 的线性变换，求证：

σ 是双射$\Longleftrightarrow\sigma(\varepsilon_1),\cdots,\sigma(\varepsilon_n)$线性无关。

7. 求下列线性变换在指定基下的矩阵：

(1) 1 题(3)中线性变换$\mathscr{A}$在基$\begin{pmatrix}1\\0\\0\end{pmatrix},\begin{pmatrix}0\\1\\0\end{pmatrix},\begin{pmatrix}0\\0\\1\end{pmatrix}$下的矩阵；

(2) 1 题(8)中，线性变换$\mathscr{A}$在基

$$\varepsilon_0=1,\varepsilon_i=\frac{x(x-1)\cdots(x-i+1)}{i!},i=1,\cdots,n-1$$

下的矩阵；

(3) $P^{2\times 2}$ 中线性变换 σ：$\sigma(X)=\begin{pmatrix}1&2\\3&4\end{pmatrix}X-X\begin{pmatrix}1&2\\3&4\end{pmatrix}$，在基 $\varepsilon_1=\begin{pmatrix}1&0\\0&0\end{pmatrix},\varepsilon_2=\begin{pmatrix}0&1\\0&0\end{pmatrix},\varepsilon_3=\begin{pmatrix}0&0\\0&1\end{pmatrix},\varepsilon_4=\begin{pmatrix}0&0\\0&1\end{pmatrix}$ 下的矩阵；

(4) $P^{2\times 2}$ 中线性变换 σ：$\sigma(X)=\begin{pmatrix}a&b\\c&d\end{pmatrix}X$，在基 $\varepsilon_1=\begin{pmatrix}1&0\\0&0\end{pmatrix},\varepsilon_2=\begin{pmatrix}0&1\\0&0\end{pmatrix},\varepsilon_3=\begin{pmatrix}0&0\\1&0\end{pmatrix},\varepsilon_4=\begin{pmatrix}0&0\\0&1\end{pmatrix}$ 下的矩阵。

8. 设 $\varepsilon_1,\varepsilon_2$ 是线性空间 V 的一组基；$\mathscr{A},\mathscr{B}$ 是 V 的两个线性变换。若 $\mathscr{A}(\varepsilon_1+\varepsilon_2)=\mathscr{B}(\varepsilon_1-\varepsilon_2)$、$\mathscr{A}(\varepsilon_1-\varepsilon_2)=\mathscr{B}(\varepsilon_1-\varepsilon_2)$，求证：$\mathscr{A}=\mathscr{B}$。

9. 若 σ 是 P^n 的一个线性变换，求证：存在 P 上 n 阶方阵 A 使 $\sigma(\alpha)=A\cdot\alpha$。

10. 若线性空间 V 的线性变换 σ 满足 $2\sigma^2+\sigma=6I$。记

$$V_1=\{\alpha\in V|\sigma(\alpha)=\frac{3}{2}\alpha\},V_2=\{\alpha\in V|\sigma(\alpha)=-2\alpha\},$$

求证：

$$V=V_1\oplus V_2。$$

11. 设 3 维线性空间 σ 在基 $\varepsilon_1,\varepsilon_2,\varepsilon_3$ 下的矩阵为 $A=(\alpha_{ij})_{3\times 3}$：

(1) 求 σ 在基 $\varepsilon_3,\varepsilon_1,\varepsilon_2$ 下的矩阵；

(2) 求 σ 在基 $\varepsilon_1,k\varepsilon_2,\varepsilon_2$ 下的矩阵($k\neq 0$)；

(3) 求 σ 在基 $\varepsilon_1+\varepsilon_2,\varepsilon_1-\varepsilon_2,2\varepsilon_3$ 下的矩阵。

12. $\boldsymbol{P}^3$ 中线性变换 σ 在基 $\eta_1=\begin{pmatrix}-1\\1\\1\end{pmatrix},\eta_2=\begin{pmatrix}1\\0\\-1\end{pmatrix},\eta_3=\begin{pmatrix}0\\1\\1\end{pmatrix}$ 下的矩阵为 $\begin{pmatrix}1&0&1\\1&1&0\\-1&2&1\end{pmatrix}$。求 σ 在基 $\varepsilon_1=\begin{pmatrix}1\\0\\0\end{pmatrix},\varepsilon_2=\begin{pmatrix}0\\1\\0\end{pmatrix},\varepsilon_3=\begin{pmatrix}0\\0\\1\end{pmatrix}$ 下的矩阵。

13. 若 $\boldsymbol{V}=\boldsymbol{R}^2$ 中线性变换 σ 满足：

$$\sigma\left(\begin{pmatrix}1\\2\end{pmatrix}\right)=\begin{pmatrix}-2\\3\end{pmatrix},\sigma\left(\begin{pmatrix}3\\1\end{pmatrix}\right)=\begin{pmatrix}3\\-4\end{pmatrix},$$

求 $\sigma\left(\begin{pmatrix}x\\y\end{pmatrix}\right)$。

14. $\boldsymbol{P}^3$ 中线性变换 σ 满足

$$\sigma\left(\begin{pmatrix}-1\\0\\2\end{pmatrix}\right)=\begin{pmatrix}-5\\0\\3\end{pmatrix},\sigma\left(\begin{pmatrix}0\\1\\1\end{pmatrix}\right)=\begin{pmatrix}0\\-1\\6\end{pmatrix},$$

$$\sigma\left(\begin{pmatrix}3\\-1\\0\end{pmatrix}\right)=\begin{pmatrix}-5\\-1\\9\end{pmatrix},$$

求 σ 在基 $\begin{pmatrix}1\\0\\0\end{pmatrix},\begin{pmatrix}0\\1\\0\end{pmatrix},\begin{pmatrix}0\\0\\1\end{pmatrix}$ 下的矩阵。

15. 求 $\boldsymbol{P}^{2\times2}$ 中线性变换 σ：$\sigma(\boldsymbol{X})=\begin{pmatrix}1&1\\1&1\end{pmatrix}\boldsymbol{X}$ 的特征值与特征向量。

16. 若线性变换 σ 在基 $\varepsilon_1,\varepsilon_2,\varepsilon_3$ 下的矩阵为

(1) $\begin{pmatrix}1&2&2\\1&-1&1\\4&-12&1\end{pmatrix}$，

$$(2)\begin{pmatrix} 3 & 1 & 0 \\ -4 & -1 & 0 \\ 4 & -8 & -2 \end{pmatrix},$$

求 σ 的特征值与特征向量。

17. 若 σ 是可逆线性变换，求证：

(1) σ 的特征值一定不为 0；

(2) 如果 λ 是 σ 的特征值，则 $\dfrac{1}{\lambda}$ 是 σ^{-1} 的特征值。

18. 若线性变换 σ 满足 $\sigma^s = 0, s \in \mathbf{Z}^+$，$\lambda$ 是 σ 的特征值，求证：$\lambda=0$。

19. 设 σ 是线性空间 V 的线性变换，如果 $\sigma^{k-1}(\alpha) \neq 0, \sigma^k(\alpha)=0$，求证：$\alpha, \sigma(\alpha), \cdots, \sigma^{k-1}(\alpha)$ 线性无关 $(k > 0)$。

20. 若线性变换 σ 在任何基下的矩阵都相同，求证：存在数 k 使 $\sigma = k\boldsymbol{I}$。

21. 求下列线性变换 σ 的值域与核的基：

$$(1)\ \boldsymbol{V}=\boldsymbol{P}^3, \sigma(\boldsymbol{X})=\begin{pmatrix} 1 & -1 & 0 \\ 1 & 0 & 2 \\ -1 & 2 & 2 \end{pmatrix}\boldsymbol{X};$$

$$(2)\ \boldsymbol{V}=\boldsymbol{P}^3, \sigma\left(\begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix}\right)=\begin{pmatrix} 2a_1+a_2+3a_3 \\ 3a_1-a_2+a_3 \\ a_1+2a_2-3a_3 \end{pmatrix};$$

$$(3)\ \boldsymbol{V}=\boldsymbol{P}^n, \sigma\left(\begin{pmatrix} a_1 \\ \vdots \\ a_k \\ \vdots \\ a_n \end{pmatrix}\right)=\begin{pmatrix} a_1 \\ \vdots \\ a_k \\ 0 \\ \vdots \\ 0 \end{pmatrix} \quad (k<n);$$

$$(4)\ \boldsymbol{V}=\boldsymbol{P}^{2\times 2}, \sigma(\boldsymbol{X})=\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}\boldsymbol{X}。$$

22. 若线性变换 $\mathscr{A}$ 与 $\mathscr{B}$ 可交换，即 $\mathscr{A}\mathscr{B}=\mathscr{B}\mathscr{A}$，求证：$\mathscr{A}$ 的特征子空间是 $\mathscr{B}$ 的不变子空间。

23. 若 n 维线性空间 V 的线性变换 σ 有 n 个不同特征值，求 σ 的不变子空间的个数，并说明理由。

24. 设 $\sigma_1,\cdots,\sigma_s$ 是线性空间 V 的 s 个两两不同的线性变换，求证：V 中存在向量 α 使 $\sigma_1(\alpha),\cdots,\sigma_s(\alpha)$ 也两两不同。

25. 若线性变换 $\mathscr{A}$，$\mathscr{B}$ 满足 $\mathscr{A}^2=\mathscr{A}$，$\mathscr{B}^2=\mathscr{B}$，求证：

(1) $\mathscr{A}$ 与 $\mathscr{B}$ 有相同值域 $\Leftrightarrow \mathscr{A}\mathscr{B}=\mathscr{B}$，$\mathscr{B}\mathscr{A}=\mathscr{A}$；

(2) $\mathscr{A}$ 与 $\mathscr{B}$ 有相同的核 $\Leftrightarrow \mathscr{A}\mathscr{B}=\mathscr{A}$，$\mathscr{B}\mathscr{A}=\mathscr{B}$。

26. 设 σ 是线性空间 V 的线性变换，W 是 V 的子空间，求证：$\dim W=\dim(\sigma(W))-\dim[\mathrm{Ker}\sigma\cap W]$。

27. 设 $\mathscr{A}$，$\mathscr{B}$ 是 n 维线性空间的两个线性变换，求证：($\mathscr{A}\mathscr{B}$)的秩 $\geqslant \mathscr{A}$ 的秩 $+\ \mathscr{B}$ 的秩 $-\ n$。

28. σ 为 $V=\boldsymbol{R}^4$ 的一个线性变换：

$$(1)\ \sigma(\alpha)=\begin{pmatrix}1 & 1 & 3 & 4\\ -1 & 3 & 4 & 1\\ 0 & 0 & 3 & 5\\ 0 & 0 & 0 & 3\end{pmatrix}\alpha,$$

$$(2)\ \sigma(\alpha)=\begin{pmatrix}1 & 3 & 3 & 4\\ 4 & 2 & 7 & 8\\ 0 & 0 & 0 & 2\\ 0 & 0 & 1 & -1\end{pmatrix}\alpha,$$

求 σ 的所有根子空间。

29. 设 $\mathscr{A}$ 为线性空间 V 的一个线性变换，$\mathscr{A}$ 的特征多项式为 $f(x)\cdot g(x)$，$(f(x),g(x))=1$。记

$$W_1=f(\mathscr{A})V=\{f(\mathscr{A})(\alpha)\,|\,\alpha\in V\},$$
$$W_2=g(\mathscr{A})V=\{g(\mathscr{A})(\alpha)\,|\,\alpha\in V\},$$

求证：

$$V=W_1\oplus W_2。$$

30. 设 σ 是内积空间 V 到 V 的映射，若对任意 $\alpha,\beta\in V$，恒有 $[\sigma(\alpha),\sigma(\beta)]=[\alpha,\beta]$，求证：$\sigma$ 是 V 的线性变换，从而是酉变换。

性质 若$\mathscr{B}$是 n 维线性空间 V 的幂零线性变换，则

(1) $\mathscr{B}$ 的特征值全为 0，$\mathscr{B}$ 的特征多项式为 x^n；

(2) 若$\mathscr{B}$的幂零指数为 t_0，则 $t_0 \leqslant n$，$\mathscr{B}$ 的最小多项式为 x^{t_0}。

证 因为$\mathscr{B}$是幂零线性变换，故存在正整数 m 使

$$\mathscr{B}^m = 0,$$

于是 x^m 为$\mathscr{B}$的一个化零多项式，从而$\mathscr{B}$的特征值(必是最小多项式，从而也是化零多项式的根)全为 0。所以特征多项式的根(特征值)全为 0，又它是 n 次首 1 多项式，故 x^n 为$\mathscr{B}$的特征多项式。

因为最小多项式是特征多项式 x^n 的因式，故其形为

$$x^t \quad (t \leqslant n)。$$

由幂零指数的定义，可知 x^{t_0}为$\mathscr{B}$的最小多项式。

定义 10.2 数域 $\boldsymbol{P}$ 上 n 阶方阵 $\boldsymbol{B}$ 称为幂零阵，如果存在正整数 m 使

$$\boldsymbol{B}^m = \boldsymbol{0}。$$

若 $\boldsymbol{B}$ 为幂零阵，称非空正整数集合

$$\{m \in \mathbf{Z}^+ \mid \boldsymbol{B}^m = \boldsymbol{0}\}$$

中的最小正整数 t_0 为 $\boldsymbol{B}$ 的幂零指数。例如，幂零阵

$$\begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}, \quad \left(\begin{array}{cc|cc} 0 & 0 & & \\ 1 & 0 & & \\ \hline & & 0 & 0 \\ & & 1 & 0 \end{array}\right), \quad \begin{pmatrix} 4 & 2 \\ -8 & -4 \end{pmatrix}$$

的幂零指数分别为 3，2，2。

一般，n 阶幂零阵(特征多项式为 x^n)的幂零指数小于等于 n。

在本节余下部分，若不加说明，都假定$\mathscr{B}$是线性空间 V 的幂零变换，其幂零指数为 t_0。

二、向量的幂零指数与 t-关系

定义 10.3　$\alpha \in V$ 为线性空间 V 的非零向量，正整数 t 称为 α 关于幂零变换 $\boldsymbol{B}$ 的幂零指数，如果

$$\mathscr{B}^{t}(\alpha)=\boldsymbol{0}\ ,\text{而}\mathscr{B}^{t-1}(\alpha)\neq\boldsymbol{0}\ ,$$

即 t 是使 $\mathscr{B}^{t}(\alpha)=\boldsymbol{0}$ 的最小正整数。

因为 $\mathscr{B}^{0}(\alpha)=I(\alpha)=\alpha\neq\boldsymbol{0}$，而$\mathscr{B}^{t_0}$ 为零变换，$\mathscr{B}^{t_0}(\alpha)=\boldsymbol{0}$，故任意非零向量 α 的幂零指数 t 必满足：

$$1\leqslant t\leqslant t_0。$$

又因为 $\mathscr{B}^{0}(\boldsymbol{O})=\boldsymbol{I}(\boldsymbol{O})=\boldsymbol{O}$，故定义零向量的幂零指数为 0。

在 $\boldsymbol{V}=\boldsymbol{P}_n[x]$ 中，令$\mathscr{B}$ 为微分变换，则向量 1 的幂零指数为 1，向量 x 的幂零指数为 2，…，向量 x^i 的幂零指数为 $(i+1)$，…，向量 x^{n-1}的幂零指数为 n。

性质　(1) 若向量 α 的幂零指数为 $t(\geqslant 1)$，则向量$\mathscr{B}(\alpha)$的幂零指数为 $t-1$。

(2) 特征向量 α 的幂零指数为 1，反之也对。这是因为：$\mathscr{B}(\alpha)=\boldsymbol{O}\cdot\alpha=\boldsymbol{O}$，而$\mathscr{B}^{0}(\alpha)=\alpha\neq\boldsymbol{O}$。

(3) 若向量 $\alpha_1,\cdots,\alpha_s$ 中幂零指数最大者为 t，则其线性组合 $\sum\limits_{i=1}^{s}k_i\alpha_i$ 的幂零指数小于等于 t。这是因为$\mathscr{B}^{t}(\alpha_i)=\boldsymbol{0}$，$i=1,\cdots,s$，从而 $\mathscr{B}^{t}\left(\sum\limits_{i}k_i\alpha_i\right)=\sum k_i\mathscr{B}^{t}(\alpha_i)=\boldsymbol{0}$。

命题 10.1　若向量 α 的幂零指数为 t $(t\geqslant 1)$，则向量组

$$\bar{\alpha}=\{\alpha,\mathscr{B}(\alpha),\cdots,\mathscr{B}^{t-1}(\alpha)\}$$

线性无关。

证　用反证法。若存在不全为 0 的 $k_0,k_1,\cdots,k_{t-1}\in\boldsymbol{P}$ 使

$$k_0\alpha+k_1\mathscr{B}(\alpha)+\cdots+k_{t-1}\mathscr{B}^{t-1}(\alpha)=\boldsymbol{0}\ ,$$

设 $k_i(0\leqslant i\leqslant t-1)$ 为下标最小的非零数，此时

$$k_i \mathscr{B}^i(\alpha) + \cdots + k_{t-1}\mathscr{B}^{t-1}(\alpha) = \boldsymbol{0},$$

于是

$$\begin{aligned}\boldsymbol{0} = \mathscr{B}^{t-1-i}(\boldsymbol{0}) &= k_i \mathscr{B}^{t-1}(\alpha) + k_{i+1}\mathscr{B}^t(\alpha) + \cdots \\ &= k_i \mathscr{B}^{t-1}(\alpha),\end{aligned}$$

但 $k_i \neq 0$，α 的幂零指数为 t，$\mathscr{B}^{t-1}(\alpha) \neq \boldsymbol{0}$，于是

$$k_i \mathscr{B}^{t-1}(\alpha) \neq \boldsymbol{0},$$

这个矛盾说明了假设是错误的.从而证明了向量组

$$\bar{\alpha} = \{\alpha, \mathscr{B}(\alpha), \cdots, \mathscr{B}^{t-1}(\alpha)\}$$

是线性无关向量组。

例如，在 $\boldsymbol{V} = \boldsymbol{P}_n[x]$ 中，令 $\mathscr{B}$ 为微分变换，则

$$\overline{x^{n-1}} = \{x^{n-1}, (n-1)x^{n-2}, \cdots, (n-1)!\}$$

线性无关，且为 $\boldsymbol{V}$ 的一组基。

定义 10.3 向量组 $\alpha_1, \cdots, \alpha_s$ 称为 t-相关的，如果

(1) $\alpha_1, \cdots, \alpha_s$ 的幂零指数均为 t；

(2) 存在不全为零的 $k_1, \cdots, k_s \in \boldsymbol{P}$ 使

$$\sum_{i=1}^{s} k_i \alpha_i$$

的幂零指数小于 t。

不是 t-相关的，称为 t-无关。即对任意不全为零的 $k_1, \cdots, k_s \in \boldsymbol{P}$，向量

$$\sum_{i=1}^{s} k_i \alpha_i$$

的幂零指数仍为 t。

在 $\boldsymbol{V} = \boldsymbol{P}_4[x]$ 中，令 $\mathscr{B}$ 为微分变换，向量 $x^3, 2x^3 - x^2$ 的幂零指数均为 4，其非零线性组合

$$2 \cdot x^3 + (-1)(2x^3 - x^2) = x^2$$

的幂零指数为 3(<4)，故向量组 $x^3, 2x^3-x^2$ 为 4 -相关。

按定义，$t \geqslant 1$ 时，t -无关向量组的非零线性组合的幂零指数仍为 $t(\geqslant 1)$，即不可能是零向量(零向量的幂零指数为 0)。故得下面的命题：

命题 10.2 $t \geqslant 1$ 时，t -无关向量组必线性无关。

由定义易得，设 $t \geqslant 2$，若向量组 $\alpha_1, \cdots, \alpha_s$ 为 t -无关组，则向量组 $\mathscr{B}(\alpha_1), \cdots, \mathscr{B}(\alpha_s)$ 为 $(t-1)$ 无关组。

定义 10.4 $1 \leqslant t \leqslant t_0$，若向量 $\alpha_1, \cdots, \alpha_s, \beta$ 的幂零指数均为 t，且存在 $k_1, \cdots, k_s \in \boldsymbol{P}$ 使

$$\beta - k_1\alpha_1 - \cdots - k_s\alpha_s$$

的幂零指数小于 t，则称 $\alpha_1, \cdots, \alpha_s$ 能 t -表示 β，记成

$$\beta \equiv \sum_{i=1}^{s} k_i\alpha_i, \quad (\bmod t)$$

或

$$\alpha_1, \cdots, \alpha_s \overset{t}{\Longrightarrow} \beta。$$

命题 10.3 $1 \leqslant t \leqslant t_0$，若向量组 $\alpha_1, \cdots, \alpha_s, \beta$ 为 t -相关，而向量组 $\alpha_1, \cdots, \alpha_s$ 为 t -无关，则

$$\alpha_1, \cdots, \alpha_s \overset{t}{\Longrightarrow} \beta。$$

证 因为 $\alpha_1, \cdots, \alpha_s, \beta$ 为 t -相关，所以存在不全为零的 $k, k_1, \cdots, k_s \in \boldsymbol{P}$ 使

$$k\beta + k_1\alpha_1 + \cdots + k_s\alpha_s$$

的幂零指数小于 t，因为 $\alpha_1, \cdots, \alpha_s$ 为 t -无关，故必 $k \neq 0$。于是

$$\beta - \left(\frac{-k_1}{k}\right)\alpha_1 - \cdots - \left(\frac{-k_s}{k}\right)\alpha_s = \frac{1}{k}[k\beta + k_1\alpha_1 + \cdots + k_s\alpha_s]$$

的幂零指数小于 t。

三、极大 t-无关组与 t-秩

定义 10.5 $1 \leqslant t \leqslant t_0$，向量组 $\alpha_1, \cdots, \alpha_s$ 称为极大 t-无关组，如果

(1) $\alpha_1, \cdots, \alpha_s$ 为 t-无关向量组；

(2) 若 β 为任意一个幂零指数为 t 的向量，必 $\alpha_1, \cdots, \alpha_3, \beta$ 为 t-相关（由命题 10.3，必 $\alpha_1, \cdots, \alpha_3 \overset{t}{\Longrightarrow} \beta$，反之也对）。

引理 10.1 $1 \leqslant t \leqslant t_0$，若

$$\alpha_1, \cdots, \alpha_s \overset{t}{\Longrightarrow} \beta_i, i = 1, 2, \cdots, l. l > s,$$

则 $\beta_1, \cdots, \beta_l$ 必 t-相关。

证 由假设

$$(\beta_1, \cdots, \beta_l) = (\alpha_1, \cdots, \alpha_s)\boldsymbol{B} + (\gamma_1, \cdots, \gamma_l),$$

这里 $\boldsymbol{B}$ 为数域 $\boldsymbol{P}$ 上 s 行 l 列矩阵，$\gamma_1, \cdots, \gamma_l$ 为幂零指数小于 t 的向量。

因为 $l > s$，故齐次线性方程组 $\boldsymbol{BX} = 0$ 有非 0 解 $\begin{pmatrix} k_1 \\ \vdots \\ k_l \end{pmatrix}$，此时

$$\sum_{i=1}^{l} k_i \beta_i = (\beta_1, \cdots, \beta_l) \begin{pmatrix} k_1 \\ \vdots \\ k_l \end{pmatrix}$$

$$= (\alpha_1, \cdots, \alpha_s)\boldsymbol{B} \begin{pmatrix} k_1 \\ \vdots \\ k_l \end{pmatrix} + (\gamma_1, \cdots, \gamma_l) \begin{pmatrix} k_1 \\ \vdots \\ k_l \end{pmatrix}$$

$$= \sum_{i=1}^{l} k_i \gamma_i,$$

其幂零指数小于 t，即 $\beta_1, \cdots, \beta_l$ 为 t-相关。

$1 \leqslant t \leqslant t_0$，极大 t-无关组必线性无关，因为 $\boldsymbol{V}$ 为有限维线性空间，故线性无关向量组所含向量个数有限，从而极大 t-无关组所含向量个数有限。我们进而有下面定理：

定理 10.1 $1 \leqslant t \leqslant t_0$，任意两个极大 t-无关组所含向量个数相同。

证 设$\{\alpha_1,\cdots,\alpha_s\}$与$\{\beta_1,\cdots,\beta_l\}$为两个极大 t-无关组。如果 $s \neq l$，一般地，不妨可设 $l > s$，由引理 10.1，$\beta_1,\cdots,\beta_l$ 为 t-相关，这与假设"$\beta_1,\cdots,\beta_l$ 为极大 t-无关组"矛盾，故必 $s = l$。

定义 10.6 设 $1 \leqslant t \leqslant t_0$，称极大 t-无关组所含向量个数为 t-秩。

在 $\boldsymbol{V} = \boldsymbol{P}_n[x]$ 中，令$\mathscr{B}$为微分变换，当 $1 \leqslant t \leqslant n$ 时，$\boldsymbol{V}$ 中 $f(x)$的幂零指数为 $t \Leftrightarrow f(x)$为$(t-1)$次多项式。从而

(1) x^{t-1}的幂零指数为 t，且为 t-无关向量组；

(2) $f(x)$幂零指数为 t，$f(x)$为$(t-1)$次多项式：

$$f(x) = a_0 + a_1x + \cdots + a_{t-1}x^{t-1} \quad (a_{t-1} \neq 0)。$$

于是

$$f(x) - a_{t-1}x^{t-1}$$

的次数小于 $(t-1)$，其幂零指数小于 t，即 x^{t-1}能 t 表示任意幂零指数为 t 的向量。

由(1)，(2)，$\{x^{t-1}\}$为极大 t-无关组。从而当 $1 \leqslant t \leqslant n$ 时，$\boldsymbol{P}_n[x]$中微分变换的 t-秩均为 1。

§2 幂零线性变换的 Jordan 基

本节中，均假定$\mathscr{B}$是数域 $\boldsymbol{P}$ 上有限维线性空间 $\boldsymbol{V}$ 的幂零变换，其幂零指数为 t_0。

一、Jordan 链

定义 10.7 若向量 α 的幂零指数为 $t(\geqslant 1)$，则称向量组 α，$\mathscr{B}(\alpha),\cdots,\mathscr{B}^{t-1}(\alpha)$为由向量 α 生成的 Jordan 链，记成

$$\bar{\alpha} = \{\alpha, \mathscr{B}(\alpha), \cdots, \mathscr{B}^{t-1}(\alpha)\},$$

$$\langle\alpha\rangle = L(\bar{\alpha}) = L[\alpha, \mathscr{B}(\alpha), \cdots, \mathscr{B}^{t-1}(\alpha)],$$

这里$\langle\alpha\rangle$为由 Jordan 链$\bar{\alpha}$生成的子空间。

性质 (1) $\alpha,\mathscr{B}(\alpha),\cdots,\mathscr{B}^{t-1}(\alpha)$线性无关；

(2) $\langle\alpha\rangle$为$\mathscr{B}$不变子空间，$\dim\langle\alpha\rangle=t$，$\alpha,\mathscr{B}(\alpha),\cdots,\mathscr{B}^{t-1}(\alpha)$为$\langle\alpha\rangle$的一组基，该基下$\mathscr{B}$的矩阵为 Jordan 块：

$$\begin{pmatrix} 0 & & & \\ 1 & 0 & & \\ & \ddots & \ddots & \\ & & 1 & 0 \end{pmatrix}_{t\times t};$$

(3) $\mathscr{B}$作为$\langle\alpha\rangle$上幂零线性变换，$\langle\alpha\rangle$的 l 秩为：

$$\begin{cases} 0, & l>t \\ 1, & 1\leqslant l\leqslant t \end{cases}$$

证 (1) 见命题 10.1。

(2) 因为$\langle\alpha\rangle$的生成组 $\alpha,\mathscr{B}(\alpha),\cdots,\mathscr{B}^{t-1}(\alpha)$线性无关，故是一组基。又

$$\mathscr{B}(\alpha,\mathscr{B}(\alpha),\cdots,\mathscr{B}^{t-1}(\alpha))=(\mathscr{B}(\alpha),\mathscr{B}^2(\alpha),\cdots,\mathscr{B}^{t-1}(\alpha),0)$$

$$=(\alpha,\mathscr{B}(\alpha),\cdots,\mathscr{B}^{t-1}(\alpha))\begin{pmatrix} 0 & & & & \\ 1 & \ddots & & 0 & \\ & \ddots & \ddots & & \\ & 0 & \ddots & \ddots & \\ & & & 1 & 0 \end{pmatrix},$$

于是$\langle\alpha\rangle$是$\mathscr{B}$的不变子空间，在基 $\alpha,\mathscr{B}(\alpha),\cdots,\mathscr{B}^{t-1}(\alpha)$下的矩阵为 Jordan 块。

(3) $\boldsymbol{\beta}\in\langle\alpha\rangle$,

$$\boldsymbol{\beta}=a_k\mathscr{B}^k(\alpha)+a_{k+1}\mathscr{B}^{k+1}(\alpha)+\cdots+a_{k-1}\mathscr{B}^{t-1}(\alpha),$$

这里，$0\leqslant k\leqslant t-1$，$a_k\neq 0$，因为

$$\mathscr{B}^{t-k}(\alpha)=a_k\mathscr{B}^t(\alpha)+a_{k+1}\mathscr{B}^{t+1}(\alpha)+\cdots=\boldsymbol{0},$$

$$\mathscr{B}^{t-k-1}(\alpha)=a_k\mathscr{B}^{t-1}(\alpha)+a_{k+1}\mathscr{B}^t(\alpha)+\cdots=a_k\mathscr{B}^{t-1}(\alpha)\neq\boldsymbol{0},$$

故 β 的幂零指数为 $l=t-k, 1\leqslant l\leqslant t$。更由于

i) $\mathscr{B}^k(\alpha)$ 的幂零指数为 $l=t-k$，是 l 无关组，

ii) 任意 $\beta\in\langle\alpha\rangle$，$\beta$ 的幂零指数为 $l=t-k$，

则

$$\beta=a_k\mathscr{B}^k(\alpha)+\cdots+a_1\mathscr{B}^{t-1}(\alpha),$$

$\beta-a_k\mathscr{B}^k(\alpha)$ 的幂零指数小于 $(t-k)=l$，即 $\{\mathscr{B}^k(\alpha)\}$ 为极大 $l(=t-k)$ 无关组，故 $1\leqslant l\leqslant t$ 时，$\langle\alpha\rangle$ 的 l-秩为 1。

因为 $\mathscr{B}$ 在 $\langle\alpha\rangle$ 上的幂零指数为 t，从而 $\langle\alpha\rangle$ 中任何向量的幂零指数不大于 t。于是 $l>t$ 时，$\langle\alpha\rangle$ 的 l-秩为 0。

二、Jordan 基的定义与性质

定义 10.8 若向量组 $\overline{\alpha_1},\overline{\alpha_2},\cdots,\overline{\alpha_s}$ 为 V 的一组基，这里 α_i 的幂零指数为 n_i，$(i=1,\cdots,s)$，且 $n_1\geqslant n_2\geqslant\cdots\geqslant n_s$，则称这组基为 V 的 Jordan 基。

性质 若 $\overline{\alpha_1},\overline{\alpha_2},\cdots,\overline{\alpha_s}$ 为线性空间 V（关于幂零线性变换 $\mathscr{B}$）的一组 Jordan 基，这里 α_i 的幂零指数为 n_i，$n_1\geqslant n_2\geqslant\cdots\geqslant n_s$，则

(1) V 是 $\mathscr{B}$ 不变子空间 $\langle\alpha_1\rangle,\cdots,\langle\alpha_s\rangle$ 的直和

$$V=\langle\alpha_1\rangle\oplus\cdots\oplus\langle\alpha_s\rangle,$$

这里 $\dim\langle\alpha_i\rangle=n_i, i=1,\cdots,s$；

(2) V 在这组 Jordan 基下的矩阵为 Jordan 标准形：

$$J=\begin{pmatrix}J_1 & & \\ & \ddots & \\ & & J_s\end{pmatrix},$$

$$J_i=\begin{pmatrix}0 & & & & \\ 1 & 0 & & & \\ & \ddots & \ddots & & \\ & & \ddots & \ddots & \\ & & & 1 & 0\end{pmatrix}_{n_i\times n_i},$$

一般地，我们称线性空间的基为（关于线性变换$\mathscr{A}$）的 Jordan 基，如果$\mathscr{A}$在该基下的矩阵为 Jordan 标准形；

(3) $\boldsymbol{V}$ 的 t-秩 $=\sum\limits_{i=1}^{s}\langle\alpha_i\rangle$ 的 t-秩 $=\sum\limits_{n_i\geqslant t}1=$ Jordan 基中幂零指数为 t 的向量个数。

证 (1) 由 Jordan 链的性质，知$\langle\alpha_1\rangle,\cdots,\langle\alpha_s\rangle$为 $\boldsymbol{V}$ 的$\mathscr{B}$不变子空间：

$$\dim\langle\alpha_i\rangle=n_i, i=1,\cdots,s$$

$\overline{\alpha_i}$为$\langle\alpha_i\rangle$的基，$\overline{\alpha_1},\cdots,\overline{\alpha_s}$为 $\boldsymbol{V}$ 的基，故

$$\boldsymbol{V}=\langle\alpha_1\rangle\oplus\cdots\oplus\langle\alpha_s\rangle。$$

(2) 由(1)与 Jordan 链的性质直接推得。

(3) 由 Jordan 链的性质，仅须证：$\boldsymbol{V}$ 的 t-秩为 Jordan 基中幂零指数为 t 的向量个数 u。

设 $\beta_1,\cdots,\beta_u$ 为 Jordan 基中幂零指数为 t 的元素，则

i) $\beta_1,\cdots,\beta_u$ 为 t-无关，这是因为若存在不全为零的 $k_1,\cdots,k_u\in\boldsymbol{P}$ 使

$$\beta=\sum_{i=1}^{u}k_i\beta_i$$

的幂零指数小于 t，则

$$\mathscr{B}^{t-1}(\beta)=\sum_{i=1}^{u}k_i\,\mathscr{B}^{t-1}(\beta_i)=\boldsymbol{0},$$

但$\mathscr{B}^{t-1}(\beta_1),\cdots,\mathscr{B}^{t-1}(\beta_u)$是 Jordan 基的组成部分且线性无关，必 $k_1,\cdots,k_u$ 全为 0，与假设矛盾，故 $\beta_1,\cdots,\beta_u$ 为 t-无关。

ii) 若 γ 为任意一个幂零指数为 t 的向量。Jordan 基能线性表示 γ：

$$\gamma=\gamma_1+\cdots+\gamma_l,$$

这里 γ_i 是幂零指数等于 i 的基向量的线性组合（$i=1,2,\cdots,l$）。若

$\gamma_l \neq 0$，且 $l > t$，由 i)，γ_l 中基向量为 l-无关，从而 γ 的幂零指数为 l，大于 t，与假设矛盾。故 $l = t$，即 γ 可由基向量中幂零指数为 t 的向量（即 $\beta_1,\cdots,\beta_u$）t-表示。

由 i)与 ii)，$\beta_1,\cdots,\beta_u$ 为 V 的极大 t-无关组，故 u 为 V 的 t-秩。

三、存在与惟一性

定理 10.2 Jordan 基存在且惟一。

所谓"惟一"是指，若另有 Jordan 基 $\overline{\beta}_1,\cdots,\overline{\beta}_l$，必

(1) $l = s$；

且

(2) $\overline{\beta}_i$ 的幂零指数 $=$ α_i 的幂零指数 $= n_i (i = 1,\cdots,s)$。

证 用 $\Pi_{t_0} = \{\alpha_1,\cdots,\alpha_u\}$ 表示一个极大 t_0-无关组，这里 u 为 t_0-秩，固定不变。

因为 $\alpha_1,\cdots,\alpha_u$ 为 t_0-无关，故 $\mathscr{B}(\alpha_1),\cdots,\mathscr{B}(\alpha_u)$ 为 (t_0-1)-无关组，可扩充为一个极大 (t_0-1)-无关组：

$$\Pi_{t_0-1} = \{\mathscr{B}(\alpha_1),\cdots,\mathscr{B}(\alpha_u),\beta_1,\cdots,\beta_v\},$$

这里 $u+v = (t_0-1)$-秩，固定不变（从而 v 固定不变）。

归纳起来，由极大 t-无关组扩充为极大 $(t-1)$-无关组 Π_{t-1}：

$$\Pi_{t-1} \supseteq \mathscr{B}\Pi_t = \{\mathscr{B}(\alpha) \mid \alpha \in \Pi_t\}。$$

我们证明 $\Pi = \bigcup\limits_{t=1}^{t_0}\Pi_t$ 是一组基，由于

$$\Pi = \bigcup_{i=1}^{u}\overline{\alpha}_i \bigcup_{j=1}^{v}\overline{\beta}_j \cup \cdots,$$

从而是 Jordan 基。更由于 Jordan 基向量中幂零指数为 t $(1 \leqslant t \leqslant t_0)$ 的基向量个数为 t-秩，固定不变，从而 $u,v,\cdots$ 也固定不变．即得惟一性。

下面，证明 Π 是 V 的一组基。

(1) Π 线性无关。

若 Π 的一个非零线性组合为 $\boldsymbol{0}$：

$$\gamma_{i_1} + \cdots + \gamma_{i_l} = \boldsymbol{0} \ ,$$

这里,γ_{i_j}是 Π_{i_j}中向量的非零线性组合($1 \leqslant i_1 < i_2 \cdots < i_l \leqslant t_0$):

$$\gamma_{i_l} \equiv 0 \pmod{i_l},$$

因为 Π_{i_l}是 i_l -无关组,故 γ_{i_l}作为 Π_{i_l}中向量的线性组合,其系数全为 0,与假设矛盾,故 Π 为线性无关向量组。

(2) Π 能表示任意 V 中向量 α。对 α 的幂零指数 t 用归纳法。$t = 0$ 时,$\alpha = 0$ 能用 Π 线性表示。若 Π 能表示幂零指数小于 t 的向量,设 α 的幂零指数为 t。因为 Π_t 为极大 t -无关组,故

$$\Pi_t \overset{t}{\Longrightarrow} \alpha,$$

即存在 Π_t 的一个线性组合 β(Π_t 能线性表示 β)使 $(\alpha - \beta)$ 的幂零指数小于 t,由归纳法假设,Π 能线性表示 $(\alpha - \beta)$,从而 Π 能线性表示 $\beta + (\alpha - \beta) = \alpha$。由(1),(2),$\Pi$ 是 V 的一组基。

四、求法

引理 10.2 σ 为 V 的一个线性变换,$\varepsilon_1,\cdots,\varepsilon_t$ 为 $\operatorname{Ker}\sigma$ 的一组基,若 $\varepsilon_1,\cdots,\varepsilon_t,\varepsilon_{t+1},\cdots,\varepsilon_n$ 线性无关,则

$$\sigma(\varepsilon_{t+1}),\cdots,\sigma(\varepsilon_n)$$

也线性无关。

证 若 $\sum\limits_{i=t+1}^{n} k_i\sigma(\varepsilon_i) = 0$,则

$$\sigma\Big(\sum_{i=t+1}^{n} k_i\varepsilon_i\Big) = 0,$$

即

$$\sum_{i=t+1}^{n} k_i\varepsilon_i \in \operatorname{Ker}\sigma,$$

因而 $\varepsilon_1,\cdots,\varepsilon_t$ 能线性表示 $\sum\limits_{i=t+1}^{n} k_i\varepsilon_i$,即存在

$$k_1, \cdots, k_l \in \boldsymbol{P},$$

$$\sum_{i=t+1}^{n} k_i \boldsymbol{\varepsilon}_i = \sum_{i=1}^{t} k_i \boldsymbol{\varepsilon}_i,$$

因为 $\boldsymbol{\varepsilon}_1, \cdots, \boldsymbol{\varepsilon}_t, \boldsymbol{\varepsilon}_{t+1}, \cdots, \boldsymbol{\varepsilon}_n$ 线性无关，故 $k_i = 0 (i = 1, \cdots, t, t+1, \cdots, n)$。所以 $\sigma(\boldsymbol{\varepsilon}_{t+1}), \cdots, \sigma(\boldsymbol{\varepsilon}_n)$ 线性无关。

命题 10.4 若 $\boldsymbol{\varepsilon}_1, \cdots, \boldsymbol{\varepsilon}_t$ 为子空间

$$\boldsymbol{W}_1 = \{\alpha \in \boldsymbol{V} \mid \mathscr{B}^{l-1}(\alpha) = \boldsymbol{0}\} \quad (l \geqslant 1)$$

的一组基，扩充为

$$\boldsymbol{W}_2 = \{\alpha \in \boldsymbol{V} \mid \mathscr{B}^{l}(\alpha) = \boldsymbol{0}\} (\supseteq \boldsymbol{W}_1)$$

的一组基 $\boldsymbol{\varepsilon}_1, \cdots, \boldsymbol{\varepsilon}_t, \boldsymbol{\varepsilon}_{t+1}, \cdots, \boldsymbol{\varepsilon}_n$，则 $\boldsymbol{\varepsilon}_{t+1}, \cdots, \boldsymbol{\varepsilon}_n$ 为极大 l-无关组。

证 因为 $\boldsymbol{\varepsilon}_{t+1}, \cdots, \boldsymbol{\varepsilon}_n \in \boldsymbol{W}_2 - \boldsymbol{W}_1$，故它们的幂零指数都是 l。若 $\sum\limits_{i=t+1}^{n} k_i \boldsymbol{\varepsilon}_i$ 的幂零指数小于 l，则

$$\sum_{i=t+1}^{n} k_i \mathscr{B}^{l-1}(\boldsymbol{\varepsilon}_i) = 0,$$

由引理 10.2，$\mathscr{B}^{l-1}(\boldsymbol{\varepsilon}_{t+1}), \cdots, \mathscr{B}^{l-1}(\boldsymbol{\varepsilon}_n)$ 线性无关，故必

$$k_i = 0 \quad (i = t+1, \cdots, n),$$

所以 $\boldsymbol{\varepsilon}_{t+1}, \cdots, \boldsymbol{\varepsilon}_n$ 为 l-无关组。

若 α 为任意一个幂零指数为 l 的向量，$\alpha \in \boldsymbol{W}_2$，故 $\boldsymbol{\varepsilon}_1, \cdots, \boldsymbol{\varepsilon}_t, \cdots, \boldsymbol{\varepsilon}_n$ 能线性表示 α，但 $\boldsymbol{\varepsilon}_1, \cdots, \boldsymbol{\varepsilon}_t$ 的幂零指数小于 l，于是

$$\boldsymbol{\varepsilon}_{t+1}, \cdots, \boldsymbol{\varepsilon}_n \xRightarrow{l} \alpha。$$

由上分析，$\boldsymbol{\varepsilon}_{t+1}, \cdots, \boldsymbol{\varepsilon}_n$ 为极大 l-无关组。

若记

$$\boldsymbol{W}_i = \{\alpha \in \boldsymbol{V} \mid \mathscr{B}^{i}(\alpha) = \boldsymbol{0}\},$$

则

$$\boldsymbol{W}_1 \subseteq \boldsymbol{W}_2 \subseteq \boldsymbol{W}_3 \cdots \subseteq \boldsymbol{W}_{t_0}。$$

Jordan 基求法

(1) 将 $\boldsymbol{W}_{t_0-1}$的基扩充为 $\boldsymbol{W}_{t_0}$的基，扩充部分为极大 t_0-无关组

$$\Pi_{t_0} = \{\alpha_1, \alpha_2, \cdots\};$$

(2) 将 $\boldsymbol{W}_{t_0-2}$的基$\cup \mathscr{B}\Pi_{t_0}$扩充为 $\boldsymbol{W}_{t_0-1}$的基，$\mathscr{B}\Pi_{t_0}\cup$扩充部分为极大 $(t_0 - 1)$ 无关组 Π_{t_0-1}。设扩充部分为

$$\{\beta_1, \beta_2, \cdots\};$$

(3) 归纳起来，由 $\boldsymbol{W}_{t-2}$的基$\cup \mathscr{B}\Pi_t$ 扩充为 $\boldsymbol{W}_{t-1}$的基，$\mathscr{B}\Pi_t\cup$扩充部分为极大 $(t-1)$-无关组 Π_{t-1}。设扩充部分为

$$\{\delta_1, \delta_2, \cdots\} \quad (t = t_0, t_0 - 1, \cdots, 2);$$

(4) $\overline{\alpha_1} \cup \overline{\alpha_2} \cdots \cup \overline{\beta_1} \cup \overline{\beta_2} \cdots \cup \overline{\delta_1} \cup \overline{\delta_2} \cdots$ 即为所求 Jordan 基。

例 10.1 在 $\boldsymbol{V} = \boldsymbol{R}^3$ 中，求幂零线性变换$\mathscr{B}$ 的 Jordan 基，这里

$$\mathscr{B}(\alpha) = \boldsymbol{B} \cdot \alpha, \ \boldsymbol{B} = \begin{pmatrix} 1 & -1 & -1 \\ 2 & -2 & -2 \\ -1 & 1 & 1 \end{pmatrix}。$$

解 $\boldsymbol{B}^2 = \boldsymbol{0}$，故$\mathscr{B}$ 的幂零指数 $t_0 = 2$。$\boldsymbol{W}_1$ 的基为 $\boldsymbol{BX} = \boldsymbol{0}$ 的基础解系$\begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}$，$\begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}$，扩充 $\alpha = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$ 成 $\boldsymbol{V}$ 的一组基。于是 α 为极大 2-无关组。

$\boldsymbol{W}_0$ 为零子空间，$\mathscr{B}(\alpha) = \boldsymbol{B}\alpha = \begin{pmatrix} 1 & -1 & -1 \\ 2 & -2 & -2 \\ -1 & 1 & 1 \end{pmatrix}\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \\ -1 \end{pmatrix}$，求 β 使 $L(\mathscr{B}(\alpha), \beta) = \boldsymbol{W}_1 = L\left(\begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}\right)$，由

$$\begin{pmatrix} 1 & 1 & 1 \\ 2 & 1 & 0 \\ -1 & 0 & 1 \end{pmatrix} \xrightarrow[(3)+(1)-2(2)]{} \begin{pmatrix} 1 & 1 & 0 \\ 2 & 1 & 0 \\ -1 & 0 & 0 \end{pmatrix},$$

可取 $\boldsymbol{\beta} = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}$，于是得 Jordan 基：

$$\boldsymbol{\beta} = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \boldsymbol{\alpha} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \mathscr{B}(\alpha) = \begin{pmatrix} 1 \\ 2 \\ -1 \end{pmatrix}。$$

例 10.2 在 $\boldsymbol{V} = \boldsymbol{R}^4$ 中，求幂零线性变换$\mathscr{B}$的 Jordan 基。这里：

$$\mathscr{B}(\alpha) = \boldsymbol{B}\alpha,\ \boldsymbol{B} = \begin{pmatrix} 0 & -8 & 0 & 3 \\ 0 & -10 & 0 & 4 \\ 0 & -15 & 0 & 6 \\ 0 & -25 & 0 & 10 \end{pmatrix}。$$

解 $\boldsymbol{B}^2 = \begin{pmatrix} 0 & 5 & 0 & -2 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$，$\boldsymbol{B}^3 = 0$，$\mathscr{B}$的幂零指数 $t_0 = 3$。

$\boldsymbol{W}_2$ 的基为 $\boldsymbol{B}^2\boldsymbol{X} = \boldsymbol{0}$ 的基础解系 $\begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 2 \\ 0 \\ 5 \end{pmatrix}$，扩充 $\alpha = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}$ 为 $\boldsymbol{V}$ 的一组基。扩充部分，即 α 为极大 3-无关组。

$\boldsymbol{W}_1$ 的维数为 2，它的基与$\mathscr{B}(\alpha) = \boldsymbol{B}\alpha = \begin{pmatrix} 3 \\ 4 \\ 6 \\ 10 \end{pmatrix}$已成为 $\boldsymbol{W}_2$ 基($\boldsymbol{W}_2$ 维

数为 3)。$\mathscr{B}(\alpha)=\begin{pmatrix}3\\4\\6\\10\end{pmatrix}$为极大 2 -无关组。

W_0 为零子空间,求 β 使 $L(\mathscr{B}^2(\alpha),\beta)=W_1$。因为

$$\mathscr{B}^2(\alpha)=\mathscr{B}(\mathscr{B}(\alpha))=B\begin{pmatrix}3\\4\\6\\10\end{pmatrix}=\begin{pmatrix}-2\\0\\0\\0\end{pmatrix},$$

$$W\ =(BX=0\text{ 的解空间})=L\left[\begin{pmatrix}1\\0\\0\\0\end{pmatrix},\begin{pmatrix}0\\0\\1\\0\end{pmatrix}\right],$$

即求 β 使 $L\left[\begin{pmatrix}-2\\0\\0\\0\end{pmatrix},\beta\right]=L\left[\begin{pmatrix}1\\0\\0\\0\end{pmatrix},\begin{pmatrix}0\\0\\1\\0\end{pmatrix}\right]$,故可取 $\beta=\begin{pmatrix}0\\0\\1\\0\end{pmatrix}$。于是得 Jordan 基

$$\beta=\begin{pmatrix}0\\0\\1\\0\end{pmatrix},\alpha=\begin{pmatrix}0\\0\\0\\1\end{pmatrix},\mathscr{B}(\alpha)=\begin{pmatrix}3\\4\\6\\10\end{pmatrix},\mathscr{B}^2(\alpha)=\begin{pmatrix}-2\\0\\0\\0\end{pmatrix}。$$

幂零线性变换$\mathscr{B}$ 在这组 Jordan 基下的矩阵为

$$\left(\begin{array}{c|ccc}0&&&\\\hline&0&0&0\\&1&0&0\\&0&1&0\end{array}\right),$$

即由两个 Jordan 块组成的 Jordan 标准形。这里:

$$\mathscr{B}(\beta)=\mathbf{0}\ ,\mathscr{B}^3(\alpha)=\mathbf{0}\ ,$$

$$\mathscr{B}(\beta,\alpha,\mathscr{B}(\alpha),\mathscr{B}^2(\alpha))=(\beta,\alpha,\mathscr{B}(\alpha),\mathscr{B}^2(\alpha))\left(\begin{array}{c|ccc} 0 & & & \\ \hline & 0 & 0 & 0 \\ & 1 & 0 & 0 \\ & 0 & 1 & 0 \end{array}\right)。$$

§3 Jordan 标准形

基本定理 任意复方阵 $\boldsymbol{A}$ 都相似于 Jordan 标准形

$$\boldsymbol{J}=\begin{pmatrix} \boldsymbol{J}_1 & & & \\ & \ddots & & \\ & & \ddots & \\ & & & \boldsymbol{J}_s \end{pmatrix},$$

这里 $\boldsymbol{J}_i$ 为 Jordan 块

$$\boldsymbol{J}_i=\begin{pmatrix} \lambda_i & \ddots & & & \\ 1 & \ddots & \ddots & & \\ & \ddots & & \ddots & \\ & & \ddots & 1 & \lambda_i \end{pmatrix}。$$

若不考虑 Jordan 块的次序，Jordan 标准形 $\boldsymbol{J}$ 由 $\boldsymbol{A}$ 惟一确定。

基本定理的线性变换形式为：

基本定理 复数域上有限维线性空间都存在一组基，使给定线性变换在该基下的矩阵为 Jordan 标准形，若不考虑 Jordan 块的次序，则 Jordan 标准形由线性变换本身惟一决定。

证 (1) 设 $\mathscr{A}$ 为复数域上有限维线性空间 V 的一个线性变换，则 V 是根子空间的直和。根子空间 $\boldsymbol{R}_\lambda$ 是 $\mathscr{A}$ 不变子空间，$\mathscr{A}$ 在其内仅有一个特征值 λ，若令

$$\mathscr{B}=\mathscr{A}-\lambda\mathscr{I},$$

则$\mathscr{B}$可看成$\boldsymbol{R}_\lambda$上的线性变换，其特征值全为0；特征多项式，从而最小多项式为x^m型，因而是幂零线性变换，由定理9.4的推论2，$\mathscr{B}$在$\boldsymbol{R}_\lambda$上的幂零指数为λ作为$\mathscr{A}$的特征多项式的根的重数。$\mathscr{B}$在$\boldsymbol{R}_\lambda$上有Jordan基$\bar{\alpha}\cup\cdots$（也是$\mathscr{A}$的Jordan基）。

$\boldsymbol{W}=\langle\alpha\rangle=L(\bar{\alpha})$为$\mathscr{B}$（从而$\mathscr{A}$）不变子空间，在$\boldsymbol{W}$的基$\bar{\alpha}$下，$\mathscr{B}$的矩阵为Jordan块$\begin{pmatrix}0&&&\\1&0&&\\&\ddots&\ddots&\\&&\ddots&\ddots&\\&&&1&0\end{pmatrix}$，从而在基$\bar{\alpha}$下$\mathscr{A}$的矩阵为

$$\begin{pmatrix}0&&&&\\1&0&&&\\&\ddots&\ddots&&\\&&\ddots&\ddots&\\&&&1&0\end{pmatrix}+\begin{pmatrix}\lambda&&\\&\ddots&\\&&\lambda\end{pmatrix}=\begin{pmatrix}\lambda&&&&\\1&\lambda&&&\\&\ddots&\ddots&&\\&&\ddots&\ddots&\\&&&1&\lambda\end{pmatrix}$$，即也是Jordan块。

合并根子空间的Jordan基，成为$\boldsymbol{V}$的Jordan基$\bar{\alpha}_1\cup\cdots\cup\bar{\alpha}_s$，$\mathscr{A}$在这组基下的矩阵为Jordan标准形$\boldsymbol{J}$，Jordan基的每一个组成部分$\bar{\alpha}_i$对应于一个Jordan块

$$\boldsymbol{J}_i=\begin{pmatrix}\lambda_i&&&&\\1&\lambda_i&&&\\&\ddots&\ddots&&\\&&\ddots&\ddots&\\&&&1&\lambda_i\end{pmatrix},i=1,\cdots,s,$$

$$\boldsymbol{J}=\begin{pmatrix}\boldsymbol{J}_1&&\\&\ddots&\\&&\boldsymbol{J}_s\end{pmatrix}。$$

(2) 惟一性。若线性变换$\mathscr{A}$在另一组Jordan基下的矩阵也是Jordan标准形。此时每一个Jordan块对应到一个$\mathscr{A}$不变子空间，把同

一特征值 λ 的 Jordan 块所对应的不变子空间并在一起作直和，得不变子空间 $\boldsymbol{U}$，则

$$\boldsymbol{U} \subseteq \text{根子空间 } \boldsymbol{R}_\lambda,$$

由

$$\boldsymbol{V}=\oplus\Sigma\boldsymbol{U}\subseteq\oplus\Sigma\boldsymbol{R}_\lambda=\boldsymbol{V},$$

得

$$\dim\boldsymbol{V}=\Sigma\dim\boldsymbol{U}\leqslant\Sigma\dim\boldsymbol{R}_\lambda=\dim\boldsymbol{V},$$

必

$$\dim\boldsymbol{U}=\dim\boldsymbol{R}_\lambda,$$

于是

$$\boldsymbol{U}=\boldsymbol{R}_\lambda。$$

因为 $\boldsymbol{R}_\lambda$（作为(1)中幂零变换 $\mathscr{B}$）的 Jordan 基惟一，Jordan 基中每一个 $\bar{\alpha}$ 确定 $\mathscr{A}$ 的一个 Jordan 块，反之也对，故特征值为 λ 的 Jordan 块必惟一。即得 Jordan 标准形惟一性的证明。

例 10.3 在 $\boldsymbol{V}=\boldsymbol{R}^4$ 中，求线性变换 σ 的 Jordan 基。这里：

$$\sigma(\alpha) = \boldsymbol{A}\cdot\alpha, \boldsymbol{A} = \begin{pmatrix} 1 & 3 & -1 & 4 \\ 0 & 1 & -2 & 3 \\ 0 & 0 & 2 & 1 \\ 0 & 0 & 0 & 2 \end{pmatrix}。$$

解 σ，即 $\boldsymbol{A}$ 的特征多项式为 $|\lambda\boldsymbol{E}-\boldsymbol{A}|=(\lambda-1)^2(\lambda-2)^2$，从而 σ 的特征值为 1，2。

$$(1)\ \lambda = 2,\ 令\ \boldsymbol{B} = \boldsymbol{A} - 2\boldsymbol{E} = \begin{pmatrix} -1 & 3 & -1 & 4 \\ 0 & -1 & -2 & 3 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}，则$$

$$B_1^2=\begin{pmatrix}1&6&-5&4\\0&1&2&-5\\0&0&0&0\\0&0&0&0\end{pmatrix},$$

根子空间 R_2 为 $B_1^2X=0$ 的解空间 $W_2=L\left(\begin{pmatrix}7\\2\\-1\\0\end{pmatrix},\begin{pmatrix}26\\5\\0\\1\end{pmatrix}\right)$。又 $B_1X=0$ 有基础解系 $\begin{pmatrix}7\\2\\-1\\0\end{pmatrix}$，扩充 $\alpha=\begin{pmatrix}26\\5\\0\\1\end{pmatrix}$，成为 W_2 的一组基。于是得根子空间 R_2 的 Jordan 基 $\alpha=\begin{pmatrix}26\\5\\0\\1\end{pmatrix}$，$B_1\alpha=\begin{pmatrix}-7\\-2\\1\\0\end{pmatrix}$。

(2) $\lambda=1$，令 $B_2=A-E=\begin{pmatrix}0&3&-1&4\\0&0&-2&3\\0&0&1&1\\0&0&0&1\end{pmatrix}$，则

$$B_2^2=\begin{pmatrix}0&0&-7&12\\0&0&-2&1\\0&0&1&2\\0&0&0&1\end{pmatrix}.$$

根子空间 R_1 为 $B_2^2X=0$ 的解空间 $W_1'=L\left(\begin{pmatrix}1\\0\\0\\0\end{pmatrix},\begin{pmatrix}0\\1\\0\\0\end{pmatrix}\right)$。又 $B_2X=0$ 有基

础解系$\begin{pmatrix}1\\0\\0\\0\end{pmatrix}$，扩充$\beta=\begin{pmatrix}0\\1\\0\\0\end{pmatrix}$，成为$W_1'$的一组基。于是得根子空间$R_1$的Jordan基：

$$\beta=\begin{pmatrix}0\\1\\0\\0\end{pmatrix},\boldsymbol{B}_2\beta=\begin{pmatrix}3\\0\\0\\0\end{pmatrix},$$

从而得线性变换σ的Jordan基：

$$\alpha_1=\begin{pmatrix}26\\5\\0\\1\end{pmatrix},\alpha_2=\begin{pmatrix}-7\\-2\\1\\0\end{pmatrix},\alpha_3=\begin{pmatrix}0\\1\\0\\0\end{pmatrix},\alpha_4=\begin{pmatrix}3\\0\\0\\0\end{pmatrix},$$

$$\begin{aligned}
&\sigma(\alpha_1)=2\alpha_1+\alpha_2,\\
&\sigma(\alpha_2)=2\alpha_2,\\
&\sigma(\alpha_3)=\alpha_3+\alpha_4,\\
&\sigma(\alpha_4)=\alpha_4,\\
&\sigma(\alpha_1,\alpha_2,\alpha_3,\alpha_4)=(\alpha_1,\alpha_2,\alpha_3,\alpha_4)\left(\begin{array}{cc|cc}2&0&&\\1&2&&\\\hline &&1&0\\&&1&1\end{array}\right),
\end{aligned}$$

即幂零线性变换在这组Jordan基$\alpha_1,\alpha_2,\alpha_3,\alpha_4$下的矩阵为Jordan标准形

$$J=\left(\begin{array}{cc|cc}2&0&&\\1&2&&\\\hline &&1&0\\&&1&1\end{array}\right)。$$

若令

$$T=(\alpha_1,\alpha_2,\alpha_3,\alpha_4)=\begin{pmatrix}26 & -7 & 0 & 3\\ 5 & -2 & 1 & 0\\ 0 & 1 & 0 & 0\\ 1 & 0 & 0 & 0\end{pmatrix},$$

则

$$T^{-1}AT=J。$$

利用矩阵与线性变换的关系，下面我们导出求复方阵 A 的 Jordan 标准形 J 与过渡阵 T 的方法。这里 T 满足

$$T^{-1}AT=J。$$

设 A 为 n 阶复方阵，$V=\mathbf{C}^n$ 为复数域 $\mathbf{C}$ 上的 n 维线性空间，在 V 上定义线性变换

$$\sigma(\alpha)=A\alpha,$$

称 V 为关于 σ 的 Jordan 基 $\varepsilon_1,\cdots,\varepsilon_n$ 为 A 的 Jordan 基，此时

$$\begin{aligned}A(\varepsilon_1,\cdots,\varepsilon_n)&=(A\varepsilon_1,\cdots,A\varepsilon_n)=(\sigma(\varepsilon_1),\cdots,\sigma(\varepsilon_n))\\&=(\varepsilon_1,\cdots,\varepsilon_n)J。\end{aligned}$$

若令 $T=(\varepsilon_1,\cdots,\varepsilon_n)$，则

$$AT=TJ,$$

即

$$T^{-1}AT=J,$$

故求 J 与 T 归结为求 A 的 Jordan 基。

例 10.4 求方阵

$$A=\begin{pmatrix}1 & 3 & -1 & 4\\ 0 & 1 & -2 & 3\\ 0 & 0 & 2 & 1\\ 0 & 0 & 0 & 2\end{pmatrix}$$

的 Jordan 标准形 $\boldsymbol{J}$ 与过渡矩阵 $\boldsymbol{T}$。

解 $|\lambda \boldsymbol{E} - A| = (\lambda - 1)^2(\lambda - 2)^2$，故 $\boldsymbol{A}$ 的特征值为 1 与 2。

(1) $\lambda = 2$，令 $\boldsymbol{B}_1 = \boldsymbol{A} - 2\boldsymbol{E} = \begin{pmatrix} -1 & 3 & -1 & 4 \\ 0 & -1 & -2 & 3 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}$，

由定理 9.4，根子空间 $\boldsymbol{R}_2$ 为 $(\boldsymbol{A} - \boldsymbol{E})^2 = \begin{pmatrix} 0 & 0 & -7 & 12 \\ 0 & 0 & -2 & 1 \\ 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 1 \end{pmatrix}$ 的列空间

$$\boldsymbol{R}_2 = \boldsymbol{W}_2 = L\left(\begin{pmatrix} -7 \\ -2 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 12 \\ 1 \\ 2 \\ 1 \end{pmatrix}\right),$$

又 $\boldsymbol{B}_1\boldsymbol{X} = 0$ 有基础解系 $\begin{pmatrix} 7 \\ 2 \\ -1 \\ 0 \end{pmatrix}$，扩充 $\alpha = \begin{pmatrix} 12 \\ 1 \\ 2 \\ 1 \end{pmatrix}$ 成为 $\boldsymbol{W}_2$ 的一组基。于是得根子空间 $\boldsymbol{R}_2$ 的 Jordan 基：

$$\alpha = \begin{pmatrix} 12 \\ 1 \\ 2 \\ 1 \end{pmatrix}, \boldsymbol{B}_1\alpha = \begin{pmatrix} -7 \\ -2 \\ 1 \\ 0 \end{pmatrix}.$$

(2) $\lambda = 1$，令 $B_2 = \boldsymbol{A} - \boldsymbol{E} = \begin{pmatrix} 0 & 3 & -1 & 4 \\ 0 & 0 & -2 & 3 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{pmatrix}$，根子空间 $\boldsymbol{R}_1$ 为

$$(\boldsymbol{A}-2\boldsymbol{E})^2=\begin{pmatrix}1 & -6 & -5 & 4\\0 & 1 & 2 & -5\\0 & 0 & 0 & 0\\0 & 0 & 0 & 0\end{pmatrix}\text{ 的列空间，}$$

$$\boldsymbol{W}_1'=L\left(\begin{pmatrix}1\\0\\0\\0\end{pmatrix},\begin{pmatrix}0\\1\\0\\0\end{pmatrix}\right),$$

又 $\boldsymbol{B}_2\boldsymbol{X}=0$ 有基础解系 $\begin{pmatrix}1\\0\\0\\0\end{pmatrix}$，扩充 $\boldsymbol{\beta}=\begin{pmatrix}0\\1\\0\\0\end{pmatrix}$ 成为 $\boldsymbol{W}_1'$ 的一组基。于是得根子空间 $\boldsymbol{R}_1$ 的 Jordan 基：

$$\boldsymbol{\beta}=\begin{pmatrix}0\\1\\0\\0\end{pmatrix},\ \boldsymbol{B}_2\boldsymbol{\beta}=\begin{pmatrix}3\\0\\0\\0\end{pmatrix},$$

此时

$$\boldsymbol{T}^{-1}\boldsymbol{A}\boldsymbol{T}=\left(\begin{array}{cc|cc}2 & 0 & & \\1 & 2 & & \\\hline & & 1 & 0\\ & & 1 & 1\end{array}\right)=\boldsymbol{J},$$

这里

$$\boldsymbol{T}=(\boldsymbol{\alpha},\boldsymbol{B}_1\boldsymbol{\alpha},\boldsymbol{\beta},\boldsymbol{B}_2\boldsymbol{\beta})=\begin{pmatrix}12 & -7 & 0 & 3\\1 & -2 & 1 & 0\\2 & 1 & 0 & 0\\1 & 0 & 0 & 0\end{pmatrix}$$

为所求过渡矩阵。

例 10.5 求方阵

$$A=\begin{pmatrix}5 & 0 & 0 & 1\\ 1 & 4 & 0 & 1\\ 0 & 1 & 3 & 0\\ 0 & -1 & 1 & 4\end{pmatrix}$$

的 Jordan 标准形 $\boldsymbol{J}$ 与过渡矩阵 $\boldsymbol{T}$：

$$\boldsymbol{T}^{-1}\boldsymbol{A}\boldsymbol{T}=\boldsymbol{J}。$$

解 $|\lambda\boldsymbol{E}-\boldsymbol{A}|=(\lambda-4)^4$，$\boldsymbol{A}$ 的特征值为 4。令

$$\boldsymbol{B}=\boldsymbol{A}-4\boldsymbol{E}=\begin{pmatrix}1 & 0 & 0 & 1\\ 1 & 0 & 0 & 1\\ 0 & 1 & -1 & 0\\ 0 & 1 & -1 & 0\end{pmatrix},$$

则

$$\boldsymbol{B}^2=\begin{pmatrix}1 & -1 & 1 & 1\\ 1 & -1 & 1 & 1\\ 1 & -1 & 1 & 1\\ -1 & 1 & -1 & -1\end{pmatrix},\boldsymbol{B}^3=0。$$

$\boldsymbol{B}^2\boldsymbol{X}=0$ 的解空间 $\boldsymbol{W}_2=L\left(\begin{pmatrix}1\\1\\0\\0\end{pmatrix},\begin{pmatrix}1\\0\\-1\\0\end{pmatrix},\begin{pmatrix}1\\0\\0\\-1\end{pmatrix}\right)$ 的基，扩充 $\alpha=$ $\begin{pmatrix}1\\0\\0\\0\end{pmatrix}$ 为 $\boldsymbol{W}_3=\boldsymbol{V}=\boldsymbol{C}^4$ 的一组基。扩充部分 α 为极大 3-无关组。

$\boldsymbol{B}\boldsymbol{X}=0$ 的解空间 $\boldsymbol{W}_1=L\left(\begin{pmatrix}0\\1\\1\\0\end{pmatrix},\begin{pmatrix}1\\0\\0\\-1\end{pmatrix}\right)$ 的基，并同 $\boldsymbol{B}\alpha=\begin{pmatrix}1\\1\\0\\0\end{pmatrix}$ 已

成为 $\boldsymbol{W}_2$ 的一组基。故 $\boldsymbol{B}\alpha$ 为极大 2-无关组。

$\boldsymbol{W}_0=\{\boldsymbol{0}\}$ 并同 $\boldsymbol{B}^2\alpha=\begin{pmatrix}1\\1\\1\\-1\end{pmatrix}$ 扩充 $\beta=\begin{pmatrix}0\\1\\1\\0\end{pmatrix}$ 成为 $\boldsymbol{W}_1$ 的一组基，$\boldsymbol{B}^2\alpha$ 与 β 为极大 1-无关组。于是 $\alpha,\boldsymbol{B}\alpha,\boldsymbol{B}^2\alpha,\beta$ 为 Jordan 基。

令 $\boldsymbol{T}=(\alpha,\boldsymbol{B}\alpha,\boldsymbol{B}^2\alpha,\beta)=\begin{pmatrix}1&1&1&0\\0&1&1&1\\0&0&1&1\\0&0&-1&0\end{pmatrix}$，则

$$\boldsymbol{T}^{-1}\boldsymbol{A}\boldsymbol{T}=\boldsymbol{J}=\left(\begin{array}{ccc|c}4&0&0&\\1&4&0&\\0&1&4&\\\hline &&&4\end{array}\right)。$$

习　题　10

1. 在 $\boldsymbol{V}=\boldsymbol{C}^4$ 中，定义线性变换 σ：

$$\sigma(\alpha)=\boldsymbol{A}\alpha,$$

这里：$\boldsymbol{A}\in\boldsymbol{C}^{4\times4}$。

(1) $\boldsymbol{A}=\begin{pmatrix}2&1&0&0\\-4&-2&0&0\\6&3&0&0\\14&7&0&0\end{pmatrix}$;

(2) $\boldsymbol{A}=\begin{pmatrix}0&1&2&3\\0&0&1&2\\0&0&0&1\\0&0&0&0\end{pmatrix}$;

(3) $\boldsymbol{A} = \begin{pmatrix} 1 & 0 & -1 & -1 \\ 1 & 1 & 0 & 1 \\ 1 & 1 & 0 & 1 \\ 0 & -1 & 1 & 0 \end{pmatrix}$。

验证 σ 是幂零线性变换，并求 σ 的幂零指数。

2. 求证：非零幂零线性变换的 Jordan 标准形一定不是对角阵。

3. 若 $\alpha = k_1\beta_1 + \cdots + k_t\beta_t$，这里，$\beta_1,\cdots,\beta_t$ 是 Jordan 基的两两不同基向量，$k_i \neq 0, i = 1,\cdots,t$，求证：$\alpha$ 的幂零指数等于 $\beta_1,\cdots,\beta_t$ 的幂零指数中最大者。

4. 求本习题 1 中幂零线性变换 σ 的 Jordan 基。

5. 求证：任何一个复方阵 $\boldsymbol{A}$ 必可表为

$$\boldsymbol{A} = \boldsymbol{D} + \boldsymbol{N},$$

这里 $\boldsymbol{D}$ 相似于对角阵，$\boldsymbol{N}$ 是幂零阵。

6. 求矩阵 $\boldsymbol{A}$ 的 Jordan 标准形 $\boldsymbol{J}$ 及过渡矩阵 $\boldsymbol{T}$：$\boldsymbol{T}^{-1}\boldsymbol{A}\boldsymbol{T} = \boldsymbol{J}$。

(1) $\boldsymbol{A} = \begin{pmatrix} 1 & 2 & 0 \\ 0 & 2 & 0 \\ -2 & -1 & 1 \end{pmatrix}$；

(2) $\boldsymbol{A} = \begin{pmatrix} 3 & 1 & -3 \\ -7 & -2 & 9 \\ -2 & -1 & 4 \end{pmatrix}$；

(3) $\boldsymbol{A} = \begin{pmatrix} 3 & 1 & 0 & 0 \\ -4 & -1 & 0 & 0 \\ 7 & 1 & 2 & 1 \\ -17 & -6 & -1 & 0 \end{pmatrix}$；

(4) $\boldsymbol{A} = \begin{pmatrix} 3 & 1 & 0 & 0 \\ -4 & -1 & 0 & 0 \\ 3 & 1 & 2 & 1 \\ -3 & -1 & -1 & 0 \end{pmatrix}$。

第十一章　矩阵函数

本章将介绍矩阵函数的概念，导出矩阵函数的计算方法，并以矩阵函数为工具解决一些具体问题。

§1　矩阵函数的概念

矩阵运算中，我们定义了矩阵的加法、数乘与乘法三种基本运算。在这三种基本运算的基础上，我们引进矩阵多项式的概念。若 $\boldsymbol{A}$ 为方阵，$f(x)$为多项式：

$$f(x) = a_0 + a_1x + \cdots + a_tx^t,$$

我们定义矩阵多项式 $f(A)$：

$$f(A) = a_0E + a_1A + \cdots + a_tA^t。$$

但当 $f(x)$不是多项式而是一般函数时（如 e^x、$\sin x$，$\arctan x$ 等），我们尚未定义 $f(A)$。

在本节中，我们将就 $f(x)$是一般（具有各阶导数）的函数时，给出 $f(A)$的定义（当 $f(x)$是多项式时，与原定义一致）。

一、多项式矩阵赋值

在矩阵多项式中，我们证明：如果 $\boldsymbol{P}$ 是数域，$f(x)$，$g(x)$，$h(x)$，$l(x) \in \boldsymbol{P}[x]$，$\boldsymbol{A} \in \boldsymbol{P}^{n\times n}$，且

$$f(x) + g(x) = h(x),$$

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

则

$$f(\boldsymbol{A})+g(\boldsymbol{A})=h(\boldsymbol{A}),$$

$$f(\boldsymbol{A})\cdot g(\boldsymbol{A})=l(\boldsymbol{A})。$$

更进一步，我们有下面的命题：

命题 11.1 二元多项式 $f(x,y)\in \boldsymbol{P}[x,y]$，$f(x,y)=\sum_{i=0}^{t}a_i(x)y^i$，$a_i(x)\in \boldsymbol{P}[x]$，$i=0,\cdots,t$。$\boldsymbol{A},\boldsymbol{B}\in \boldsymbol{P}^{n\times n}$，$\boldsymbol{AB}=\boldsymbol{BA}$，则

$$f(\boldsymbol{A},\boldsymbol{B})=\sum_{i=0}^{t}a_i(\boldsymbol{A})\boldsymbol{B}^i。$$

证 因为 $f(x,y)$ 经字母 x,y 与 $\boldsymbol{P}$ 中数作加、乘运算（加、乘满足交换律、结合律与分配律）化成

$$\sum_{i=0}^{t}a_i(x)y^i。$$

在整个过程中，用 $\boldsymbol{A}$ 代 x，$\boldsymbol{B}$ 代 y，用 $k\boldsymbol{E}$ 代 $\boldsymbol{P}$ 中的数 k，由于 $\boldsymbol{A},\boldsymbol{B},k\boldsymbol{E}$ 间的加、乘同样满足交换律、结合律与分配律，因而不影响结果，故

$$f(\boldsymbol{A},\boldsymbol{B})=\sum_{i=0}^{t}a_i(\boldsymbol{A})\boldsymbol{B}^i。$$

例如

$$\begin{aligned}&(x+y)^2-(x+1)^2-(y-2)^2+2x-3y-7\\=&(x^2+2xy+y^2)-(x^2+2x+1)-(y^2-4y+4)\\&+2x-3y-7\\=&2xy+y-3,\end{aligned}$$

于是

$$\begin{aligned}&(\boldsymbol{A}+\boldsymbol{B})^2-(\boldsymbol{A}+\boldsymbol{E})^2-(\boldsymbol{B}-2\boldsymbol{E})^2+2\boldsymbol{A}-3\boldsymbol{B}-7\boldsymbol{E}\\=&(\boldsymbol{A}^2+2\boldsymbol{AB}+\boldsymbol{B}^2)-(\boldsymbol{A}^2+2\boldsymbol{A}+\boldsymbol{E})-(\boldsymbol{B}^2-4\boldsymbol{B}+4\boldsymbol{E})\\&+2\boldsymbol{A}-3\boldsymbol{B}-7\boldsymbol{E}\\=&2\boldsymbol{AB}+\boldsymbol{B}-3\boldsymbol{E}。\end{aligned}$$

二、Jordan 块的多项式

若 $\boldsymbol{A}$ 为 n 阶 Jordan 块

$$\boldsymbol{A}=\begin{pmatrix}\lambda & & & & \\ 1 & \lambda & & & \\ & \ddots & \ddots & & \\ & & \ddots & \ddots & \\ & & & 1 & \lambda\end{pmatrix},\ \lambda\in\mathbf{C},$$

即

$$\boldsymbol{A}=\lambda\boldsymbol{E}+\boldsymbol{H},$$

这里

$$\boldsymbol{H}=\begin{pmatrix}0 & & & & \\ 1 & \ddots & & & \\ & \ddots & \ddots & & \\ & & \ddots & \ddots & \\ & & & 1 & 0\end{pmatrix},\boldsymbol{H}^2=\begin{pmatrix}0 & & & & \\ 0 & \ddots & & & \\ 1 & \ddots & \ddots & & \\ & \ddots & \ddots & \ddots & \\ & & 1 & 0 & 0\end{pmatrix},\cdots$$

$$\boldsymbol{H}^{n-1}=\begin{pmatrix}0 & & & \\ 0 & \ddots & & \\ \vdots & & \ddots & \\ 1 & \cdots & 0 & 0\end{pmatrix},\boldsymbol{H}^n=0。$$

设 $p(x)$为一复系数 t 次多项式,我们来计算 Jordan 块 $\boldsymbol{A}$ 的多项式 $p(\boldsymbol{A})$。

引理 11.1 若 $p(x)$为一复系数 t 次多项式,则

$$p(x+h)=p(x)+p'(x)h+\frac{p''(x)}{2!}h^2+\cdots+\frac{p^{(t)}(x)}{t!}h^t。$$

证 (1) 当 $p(x)$为实系数多项式时,由 Taylor 公式得

$$p(x+h)=p(x)+p'(x)h+\cdots+\frac{p^{(t)}(x)}{t!}h^t+\frac{p^{(t+1)}(\xi)}{(t+1)!}h^{t+1},$$

因为 $p(x)$为 t 次多项式,故 $p^{(t+1)}(\xi)=0$,于是

$$p(x+h)=p(x)+p'(x)h+\cdots+\frac{p^{(t)}(x)}{t!}h^t。$$

(2) 若 $p(x)$为复系数多项式:

$$p(x)=u(x)+\mathrm{i}v(x),$$

这里 $u(x)$,$v(x)$为次数不大于 t 的实系数多项式。由(1)

$$u(x+h)=u(x)+u'(x)h+\cdots+\frac{u^{(t)}(x)}{t!}h,$$

$$v(x+h)=v(x)+v'(x)h+\cdots+\frac{v^{(t)}(x)}{t!}h^t,$$

于是

$$\begin{aligned}
&p(x+h)=u(x+h)+\mathrm{i}v(x+h)\\
=&[u(x)+\mathrm{i}z(x)]+[u'(x)+\mathrm{i}v'(x)]h+\cdots\\
&+\frac{1}{t!}[u^{(t)}(x)+\mathrm{i}v^{(t)}(x)]h^t\\
=&p(x)+p'(x)h+\cdots+\frac{p^t(x)}{t!}h^t。
\end{aligned}$$

定理 11.1 若 $p(x)$为复系数多项式,A 为 n 阶 Jordan 块:

$$\boldsymbol{A}=\begin{pmatrix}\lambda&&&\\1&\ddots&&\\&\ddots&\ddots&\\&&1&\lambda\end{pmatrix},$$

则

$$p(\boldsymbol{A})=p(\lambda\boldsymbol{E}+\boldsymbol{H})=p(\lambda)\boldsymbol{E}+p'(\lambda)\boldsymbol{H}+\cdots+\frac{p^{(n-1)}(\lambda)}{(n-1)!}\boldsymbol{H}^{n-1}。$$

证 在引理 11.1 中,令 $\boldsymbol{x}=\lambda\boldsymbol{E}$,$h=\boldsymbol{H}$,由命题 11.1 得:

$$p(A)=p(\lambda E+H)=p(\lambda)E+p'(\lambda)H+\cdots+\frac{p^{(t)}(\lambda)}{t!}H^t。$$

注意到 $H^n=H^{n+1}=\cdots=0$，即得证明。

三、矩阵多项式

设 A 为 n 阶复方阵，A 相似于 Jordan 标准形 J，即存在可逆阵 T 使

$$A=TJT^{-1}=T\begin{pmatrix}J_1&&\\&\ddots&\\&&J_s\end{pmatrix}T^{-1},$$

J_i 为 Jordan 块

$$J_i=\begin{pmatrix}\lambda_i&&&&\\1&\ddots&&&\\&\ddots&\ddots&&\\&&\ddots&\ddots&\\&&&1&\lambda_i\end{pmatrix},\quad i=1,\cdots,s。$$

若 $p(x)$ 为一复系数多项式，则

$$\begin{aligned}p(A)&=p(TJT^{-1})=Tp(J)T^{-1}\\&=T\begin{pmatrix}p(J_1)&&&\\&\ddots&&\\&&\ddots&\\&&&p(J_s)\end{pmatrix}T^{-1},\end{aligned}$$

即矩阵 A 的多项式 $p(A)$ 完全由 A 的 Jordan 块 J_i 的多项式 $p(J_i)$ $(i=1,\cdots,s)$ 确定。因而我们希望通过定义 Jordan 块的函数 $f(J_i)$ 来定义矩阵 A 的函数 $f(A)$。

四、矩阵函数的定义

定义 11.1 若 A 为 n 阶 Jordan 块：

$$
A=\begin{pmatrix}\lambda & & & \\ 1 & \ddots & & \\ & \ddots & \ddots & \\ & & \ddots & \ddots \\ & & & 1 & \lambda\end{pmatrix}_{n\times n},
$$

函数 $f(x)$ 在 $x=\lambda$ 处有直到 $(n-1)$ 阶的导数，则定义矩阵 $\boldsymbol{A}$ 的函数为

$$
f(\boldsymbol{A})=f(\lambda)\boldsymbol{E}+f'(\lambda)\boldsymbol{H}+\frac{f''(\lambda)}{2!}\boldsymbol{H}^2+\cdots+\frac{f^{(n-1)}(\lambda)}{(n-1)!}\boldsymbol{H}^{n-1}
$$

$$
=\begin{pmatrix} f(\lambda) & & & & \\ f'(\lambda) & \ddots & & & \\ \vdots & \ddots & \ddots & & \\ \vdots & & \ddots & \ddots & \\ \frac{f^{(n-1)}(\lambda)}{(n-1)!} & \cdots & \cdots & f'(\lambda) & f(\lambda)\end{pmatrix}。
$$

例如，设 $\boldsymbol{A}=\begin{pmatrix}3&0&0\\1&3&0\\0&1&3\end{pmatrix}$，则

(1) 若 $f(x)=\mathrm{e}^x$，于是 $f'(x)=f''(x)=\mathrm{e}^x$，故

$$
\mathrm{e}^{\boldsymbol{A}}=\begin{pmatrix} f(3) & 0 & 0\\ f'(3) & f(3) & 0\\ \frac{f''(3)}{2!} & f'(3) & f(3)\end{pmatrix}=\begin{pmatrix}\mathrm{e}^3 & 0 & 0\\ \mathrm{e}^3 & \mathrm{e}^3 & 0\\ \frac{\mathrm{e}^3}{2} & \mathrm{e}^3 & \mathrm{e}^3\end{pmatrix}。
$$

(2) 若 $f(x)=\sin x$，则 $f'(x)=\cos x$，$f''(x)=-\sin x$，故

$$
\sin\boldsymbol{A}=\begin{pmatrix} f(3) & 0 & 0\\ f'(3) & f(3) & 0\\ \frac{f''(3)}{2!} & f'(3) & f(3)\end{pmatrix}=\begin{pmatrix}\sin 3 & 0 & 0\\ \cos 3 & \sin 3 & 0\\ \frac{-\sin 3}{2} & \cos 3 & \sin 3\end{pmatrix}。
$$

定义 11.2 若复方阵 $\boldsymbol{A}$ 相似于 Jordan 标准形 $\boldsymbol{J}$，即

$$A=\boldsymbol{TJT}^{-1}=\boldsymbol{T}\begin{pmatrix}\boldsymbol{J}_1 & & \\ & \ddots & \\ & & \boldsymbol{J}_s\end{pmatrix}\boldsymbol{T}^{-1},$$

这里 $\boldsymbol{J}_1,\cdots,\boldsymbol{J}_s$ 为 Jordan 块。如果 Jordan 块函数 $f(\boldsymbol{J}_1),\cdots,f(\boldsymbol{J}_s)$有定义(定义 11.1),则我们定义方阵 A 的函数

$$f(\boldsymbol{A})=\boldsymbol{T}\begin{pmatrix}f(\boldsymbol{J}_1) & & \\ & \ddots & \\ & & f(\boldsymbol{J}_s)\end{pmatrix}\boldsymbol{T}^{-1}。$$

例如，
$$\boldsymbol{A}=\boldsymbol{T}\left(\begin{array}{cc|ccc}1 & 0 & & & \\ 1 & 1 & & & \\ \hline & & 2 & 0 & 0\\ & & 1 & 2 & 0\\ & & 0 & 1 & 2\end{array}\right)\boldsymbol{T}^{-1},\text{则}$$

$$\mathrm{e}^{\boldsymbol{A}}=\boldsymbol{T}\left(\begin{array}{cc|ccc}\mathrm{e} & 0 & & & \\ \mathrm{e} & \mathrm{e} & & & \\ \hline & & \mathrm{e}^2 & 0 & 0\\ & & \mathrm{e}^2 & \mathrm{e}^2 & 0\\ & & \frac{\mathrm{e}^2}{2} & \mathrm{e}^2 & \mathrm{e}^2\end{array}\right)\boldsymbol{T}^{-1},$$

$$\sin\boldsymbol{A}=\boldsymbol{T}\left(\begin{array}{cc|ccc}\sin 1 & 0 & & & \\ \cos 1 & \sin 1 & & & \\ \hline & & \sin 2 & 0 & 0\\ & & \cos 2 & \sin 2 & 0\\ & & \frac{-\sin 2}{2} & \cos 2 & \sin 2\end{array}\right)\boldsymbol{T}^{-1}。$$

推论 $f(x)$为多项式时,定义 11.2 所得 $f(\boldsymbol{A})$与按原矩阵多项式定义所得 $f(\boldsymbol{A})$是相同的。

§2 矩阵函数的幂级数展开

一、矩阵序列的极限

定义 11.3 P 为数域，$A_n \in P^{s\times s}$，$n = 1,2,\cdots$。若

$$\lim_{n\to\infty} A_n(i,j) = a_{ij},\ i,j = 1,\cdots,s,$$

记

$$A = (a_{ij})_{s\times s},$$

则称 $n\to\infty$ 时，矩阵序列$\{A_n\}$收敛于极限 A，记成

$$\lim_{n\to\infty} A_n = A,$$

也称极限 $\lim\limits_{n\to\infty} A_n$ 存在且为 A。

例如，$n\to\infty$时

$$A_n = \begin{pmatrix} 1-\dfrac{1}{n} & \dfrac{1}{2^n} \\ \left(1+\dfrac{1}{n}\right)^n & 3+\dfrac{2}{n} \end{pmatrix} \longrightarrow \begin{pmatrix} 1 & 0 \\ e & 3 \end{pmatrix},$$

或写成

$$\lim_{n\to\infty} \begin{pmatrix} 1-\dfrac{1}{n} & \dfrac{1}{2^n} \\ \left(1+\dfrac{1}{n}\right)^n & 3+\dfrac{2}{n} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ e & 3 \end{pmatrix}。$$

命题 11.2 若 $A_n, B_n \in P^{s\times s}$，$n = 1,2,3,\cdots$，且

$$\lim_{n\to\infty} A_n = A,\ \lim_{n\to\infty} B_n = B,$$

则

$$\lim_{n\to\infty}\left(A_n \begin{matrix}\pm\\ \times\end{matrix} B_n\right)$$

存在，且

$$\lim_{n\to\infty}\left(A_n \begin{matrix}\pm\\ \times\end{matrix} B_n\right) = A \begin{matrix}\pm\\ \times\end{matrix} B = \left(\lim_{n\to\infty} A_n\right) \begin{matrix}\pm\\ \times\end{matrix} \left(\lim_{n\to\infty} B_n\right),$$

即矩阵和、差、积序列的极限等于矩阵序列极限的和、差、积。

证 因为 $n\to\infty$ 时

$$(A_n \pm B_n)(i,j) = A_n(i,j) \pm B_n(i,j)$$

$$\longrightarrow A(i,j) \pm B(i,j) = (A \pm B)(i,j),$$

故

$$\lim_{n\to\infty}(A_n \pm B_n) = A \pm B。$$

又，$n \to \infty$ 时

$$(A_nB_n)(i,j) = \sum_{k=1}^{s} A_n(i,k)B_n(k,j)$$

$$\longrightarrow \sum_{k=1}^{s} A(i,k)B(k,j) = (AB)(i,j),$$

所以

$$\lim_{n\to\infty}(A_nB_n) = AB。$$

推论 (1) $\lim\limits_{n\to\infty}(kA_n) = k\lim\limits_{n\to\infty}A_n$；

(2) $\lim\limits_{n\to\infty}(BA_nC) = B\left(\lim\limits_{n\to\infty}A_n\right)C$；

(3) 若 $f(x)$是字母 x 的多项式，则

$$\lim_{n\to\infty}f(A_n) = f\left(\lim_{n\to\infty}A_n\right)。$$

推论(1),(2),(3)中都假定$\lim\limits_{n\to\infty}A_n$ 存在。

二、矩阵级数

定义 11.4 设 A_n 为 s 阶方阵 $(n=1,2,3,\cdots)$，称矩阵级数

$$A_1+A_2+A_3+\cdots+A_n+\cdots$$

收敛且记成

$$\sum_{n=1}^{+\infty} A_n = T,$$

如果 $\lim\limits_{n\to\infty}(A_1+\cdots+A_n)$ 存在且为 T，即

$$\lim_{n\to\infty}\left(\sum_{k=1}^{n} A_k\right)=T,$$

按定义

$$T(i,j)=\sum_{n=1}^{+\infty} A_n(i,j),\ i,j=1,2,\cdots,s。$$

例如

$$\sum_{n=1}^{+\infty}\begin{pmatrix}\dfrac{1}{2^n} & \dfrac{1}{3^n}\\ \dfrac{1}{n!} & \dfrac{(-1)^{n-1}}{n}\end{pmatrix}=\begin{pmatrix}\sum\limits_{n=1}^{+\infty}\dfrac{1}{2^n} & \sum\limits_{n=1}^{+\infty}\dfrac{1}{3^n}\\ \sum\limits_{n=1}^{+\infty}\dfrac{1}{n!} & \sum\limits_{n=1}^{+\infty}\dfrac{(-1)^{n-1}}{n}\end{pmatrix}$$

$$=\begin{pmatrix}\dfrac{\frac{1}{2}}{1-\frac{1}{2}} & \dfrac{\frac{1}{3}}{1-\frac{1}{3}}\\ e-1 & \ln 2\end{pmatrix}=\begin{pmatrix}1 & \dfrac{1}{2}\\ e-1 & \ln 2\end{pmatrix}。$$

由数项级数的收敛性质，即得下面矩阵级数的简单性质。

性质 (1) 若矩阵级数 $\sum\limits_{n=1}^{+\infty} A_n$ 收敛，则 $\lim\limits_{n\to\infty} A_n=0$；

(2) 若 A_n, B_n 均为同阶方阵，且

$$\sum_{n=1}^{+\infty} \boldsymbol{A}_n \text{ 与 } \sum_{n=1}^{+\infty} \boldsymbol{B}_n$$

收敛，则

$$\sum_{n=1}^{+\infty} (k\boldsymbol{A}_n + l\boldsymbol{B}_n)$$

也收敛，且

$$\sum_{n=1}^{+\infty} (k\boldsymbol{A}_n + l\boldsymbol{B}_n) = k\sum_{n=1}^{+\infty} \boldsymbol{A}_n + l\sum_{n=1}^{+\infty} \boldsymbol{B}_n。$$

三、矩阵幂级数

定义 11.5 若 A 为 s 阶方阵，称级数

$$a_0\boldsymbol{E} + a_1\boldsymbol{A} + \cdots + a_n\boldsymbol{A}^n + \cdots$$

为方阵 A 的幂级数。这里 $\boldsymbol{E}$ 为 s 阶单位阵。

定理 11.2 若 R 为幂级数 $\sum\limits_{n=0}^{+\infty} a_n x^n$ 的收敛半径，即当 $|x| < R$ 时，$\sum\limits_{n=0}^{+\infty} a_n x^n = f(x)$。

如果 s 阶方阵 A 的特征值的模全小于 R，则

$$f(\boldsymbol{A}) = \sum_{n=0}^{+\infty} a_n \boldsymbol{A}^n。$$

证 记 $S_n(x) = \sum\limits_{k=0}^{n} a_k x^k$，则 $\lim\limits_{n\to\infty} S_n(x) = f(x)$，$(|x| < R)$，由幂级数的性质（在收敛区间内可逐项求导）得

$$|x| < R \text{ 时}，\lim_{n\to\infty} S_n^{(k)}(x) = f^{(k)}(x)，k = 1,2,\cdots。$$

特别地，当 λ 为 A 的特征值时，$|\lambda| < R$，故有

$$\lim_{n\to\infty} S_n^{(k)}(\lambda) = f^{(k)}(\lambda)。$$

$$A \sim \text{Jordan 标准形 } \boldsymbol{J} = \begin{pmatrix} \boldsymbol{J}_1 & & \\ & \ddots & \\ & & \boldsymbol{J}_t \end{pmatrix},$$

$$A = \boldsymbol{T}\boldsymbol{J}\boldsymbol{T}^{-1} = \boldsymbol{T}\begin{pmatrix} \boldsymbol{J}_1 & & \\ & \ddots & \\ & & \boldsymbol{J}_t \end{pmatrix}\boldsymbol{T}^{-1},$$

$$f(\boldsymbol{A}) = \boldsymbol{T}\begin{pmatrix} f(\boldsymbol{J}_1) & & \\ & \ddots & \\ & & f(\boldsymbol{J}_t) \end{pmatrix}\boldsymbol{T}^{-1},$$

$$S_n(\boldsymbol{A}) = \boldsymbol{T}\begin{pmatrix} S_n(\boldsymbol{J}_1) & & \\ & \ddots & \\ & & S_n(\boldsymbol{J}_t) \end{pmatrix}\boldsymbol{T}^{-1},$$

而 $n \to \infty$ 时，$S_n'(\boldsymbol{J}_i) = S_n(\lambda_i \boldsymbol{E} + \boldsymbol{H}) = S_n'(\lambda_i)\boldsymbol{E} + S_n'(\lambda_i)\boldsymbol{H} + \cdots$

$$\longrightarrow f(\lambda_i)\boldsymbol{E} + f'(\lambda_i)\boldsymbol{H} + \cdots = f(\boldsymbol{J}_i),\ i = 1, \cdots, t,$$

于是，$n \to \infty$ 时

$$\begin{pmatrix} S_n(\boldsymbol{J}_1) & & \\ & \ddots & \\ & & S_n(\boldsymbol{J}_t) \end{pmatrix} \longrightarrow \begin{pmatrix} f(\boldsymbol{J}_1) & & \\ & \ddots & \\ & & f(\boldsymbol{J}_t) \end{pmatrix},$$

由命题 11.2 的推论(3)得

$$\begin{aligned} f(\boldsymbol{A}) &= \lim_{n \to \infty} S_n(\boldsymbol{A}) = \lim_{n \to \infty}(a_0\boldsymbol{E} + \cdots + a_n\boldsymbol{A}^n) \\ &= \sum_{n=0}^{+\infty} a_n \boldsymbol{A}^n 。\end{aligned}$$

四、常见矩阵函数的幂级数展开式

(1) 幂级数

$$e^x = 1 + x + \frac{x^2}{2!} + \cdots + \frac{x^n}{n!} + \cdots$$

的收敛半径为$+\infty$。故对任意方阵 A 均有

$$e^{A}=E+A+\frac{A^{2}}{2!}+\cdots+\frac{A^{n}}{n!}+\cdots。$$

类似地，有 $\sin A$，$\cos A$ 的幂级数展开式。

(2) 对任意方阵 A 均有

$$\sin A=\sum_{n=1}^{+\infty}\frac{(-1)^{n-1}A^{2n-1}}{(2n-1)!}=A-\frac{A^{3}}{3!}+\frac{A^{5}}{5!}+\cdots。$$

(3) 对任意方阵 A 均有

$$\cos A=\sum_{n=0}^{+\infty}\frac{(-1)^{n}A^{2n}}{(2n)!}=E-\frac{A^{2}}{2!}+\frac{A^{4}}{4!}-\frac{A^{6}}{6!}+\cdots。$$

(4) 当方阵 A 的特征值的模全小于 1 时有

$$(E+A)^{\alpha}=\sum_{n=0}^{+\infty}\frac{\alpha(\alpha-1)\cdots(\alpha-n+1)}{n!}A^{n},$$

特别地

$$(E-A)^{-1}=E+A+A^{2}+\cdots+A^{n}+\cdots。$$

(5) 当方阵 A 的特征值的模全小于 1 时有

$$\ln(E+A)=A-\frac{A^{2}}{2}+\frac{A^{3}}{3}-\cdots=\sum_{n=1}^{+\infty}\frac{(-1)^{n-1}}{n}A^{n},$$

$$\arctan A=A-\frac{A^{3}}{3}+\frac{A^{5}}{5}-\cdots=\sum_{n=1}^{+\infty}\frac{(-1)^{n-1}}{2n-1}A^{2n-1}。$$

五、矩阵函数的导函数

定义 11.6 若 s 阶方阵 $B(t)$ 的每个元素都是自变量 t 的函数

$$B(t)=(b_{ij}(t))_{s\times s},$$

则称 s 阶方阵

$$(b'_{ij}(t))_{s\times s} = \left(\frac{\mathrm{d}b_{ij}(t)}{\mathrm{d}t}\right)_{s\times s}$$

为 $\boldsymbol{B}(t)$ 的导函数，记成

$$\boldsymbol{B}'(t)，或 \frac{\mathrm{d}}{\mathrm{d}t}\boldsymbol{B}(t)。$$

例 11.1 若 A 为 s 阶方阵，$\boldsymbol{A} = (a_{ij})_{s\times s}$，记

$$\boldsymbol{A}t = (a_{ij}t)_{s\times s},$$

则

$$(\mathrm{e}^{\boldsymbol{A}t})' = \boldsymbol{A}\mathrm{e}^{\boldsymbol{A}t},$$
$$(\sin \boldsymbol{A}t)' = \boldsymbol{A}\cos \boldsymbol{A}t,$$
$$(\cos \boldsymbol{A}t)' = -\boldsymbol{A}\sin \boldsymbol{A}t。$$

证 $\mathrm{e}^{\boldsymbol{A}t} = \boldsymbol{E} + \boldsymbol{A}t + \cdots + \dfrac{\boldsymbol{A}^n}{n!}t^n + \cdots,$

$$\mathrm{e}^{\boldsymbol{A}t}(i,j) = \boldsymbol{E}(i,j) + \boldsymbol{A}(i,j)t + \cdots + \frac{\boldsymbol{A}^n(i,j)}{n!}\cdot t^n + \cdots,$$

两边求导，得

$$\begin{aligned}(\mathrm{e}^{\boldsymbol{A}t})'(i,j) &= \boldsymbol{A}(i,j) + \cdots + \frac{\boldsymbol{A}^n(i,j)}{(n-1)!}t^{n-1} + \cdots \\ &= \left[\boldsymbol{A} + \boldsymbol{A}^2t + \cdots + \frac{\boldsymbol{A}^n}{(n-1)!}t^{n-1} + \cdots\right](i,j),\end{aligned}$$

于是

$$\begin{aligned}(\mathrm{e}^{\boldsymbol{A}t})' &= \boldsymbol{A} + \boldsymbol{A}^2t + \cdots + \frac{\boldsymbol{A}^n}{(n-1)!}t^{n-1} + \cdots \\ &= \boldsymbol{A}\left[\boldsymbol{E} + \boldsymbol{A}t + \cdots + \frac{(\boldsymbol{A}t)^{n-1}}{(n-1)!} + \cdots\right] \\ &= \boldsymbol{A}\mathrm{e}^{\boldsymbol{A}t}。\end{aligned}$$

类似地可得其余两式。

§3 矩阵函数的计算

一、矩阵的谱与最小多项式

命题 11.3 若方阵 A 的最小多项式为 $m(x)$，$f(x)$ 与 $g(x)$ 为两个多项式，则

$$f(A)=g(A)\Longleftrightarrow m(x)\mid f(x)-g(x)。$$

证 $f(A)=g(A)$

$\Longleftrightarrow f(x)-g(x)$ 是 A 的化零多项式

$\Longleftrightarrow m(x)\mid f(x)-g(x)$。

定义 11.7 若

$$A=TJT^{-1},$$

$$J=\begin{pmatrix}J_1 & & \\ & \ddots & \\ & & J_s\end{pmatrix},\ J_i=\begin{pmatrix}\lambda_i & & & & \\ 1 & \ddots & & & \\ & \ddots & \ddots & & \\ & & \ddots & \ddots & \\ & & & 1 & \lambda_i\end{pmatrix}_{n_i\times n_i},(i=1,\cdots,s)$$

令

$$f(A)=T\begin{pmatrix}f(J_1) & & \\ & \ddots & \\ & & f(J_s)\end{pmatrix}T^{-1},$$

$$f(J_i)=f(\lambda_i)E+f'(\lambda_i)H+\cdots+\frac{f^{(n_i-1)}(\lambda_i)}{(n_i-1)!}H^{n_i-1},\ i=1,\cdots,s$$

$f(A)$ 完全由 $f(\lambda_i),\cdots,f^{(n_i-1)}(\lambda_i)$ $(i=1,\cdots,s)$ 确定。我们称数集 $\{f(\lambda_i),\cdots,f^{(n_i-1)}(\lambda_i)\mid i=1,\cdots,s\}$ 为函数 f 在 A 的谱上的值。

按定义，矩阵函数 $f(A)$ 完全由 f 在 A 的谱上的值确定。故可能

$f(\boldsymbol{A}) = g(\boldsymbol{A})$，但 $f(x) \neq g(x)$。

推论 方阵 $\boldsymbol{A}$ 的两个函数 $f(\boldsymbol{A})$ 与 $g(\boldsymbol{A})$ 相等：

$$f(\boldsymbol{A}) = g(\boldsymbol{A})$$

$\Longleftrightarrow f(x)$ 与 $g(x)$ 在 $\boldsymbol{A}$ 的谱上的值相同。

二、同值多项式

定义 11.8 若多项式 $p(x)$ 与函数 $f(x)$ 在方阵 $\boldsymbol{A}$ 的谱上的值相同（从而 $p(\boldsymbol{A}) = f(\boldsymbol{A})$），称 $p(x)$ 为 $f(x)$ 关于 $\boldsymbol{A}$ 的同值多项式。即若

$$\boldsymbol{A} \sim \boldsymbol{J} = \begin{pmatrix} \boldsymbol{J}_1 & & \\ & \ddots & \\ & & \boldsymbol{J}_s \end{pmatrix}, \boldsymbol{J}_i = \begin{pmatrix} \lambda_i & & & & \\ 1 & \ddots & & & \\ & \ddots & \ddots & & \\ & & \ddots & \ddots & \\ & & & 1 & \lambda_i \end{pmatrix}_{n_i \times n_i}, i = 1, \cdots, s$$

$p(x)$ 是 $f(x)$ 关于 $\boldsymbol{A}$ 的同值多项式，则 $p(x)$ 满足

$$p^{(k)}(\lambda_i) = f^{(k)}(\lambda_i), k = 0, 1, \cdots, n_i - 1,\ i = 1, \cdots, s$$

如果去掉重复的等式，适当调整次序，得

$$p^{(k)}(\lambda_i) = f^{(k)}(\lambda_i),\ k = 0, \cdots, n_i - 1,\ i = 1, \cdots, t$$

$\lambda_1, \cdots, \lambda_t$ 两两不等，此时

$$(x - \lambda_1)^{n_1} \cdots (x - \lambda_t)^{n_t}$$

为 $\boldsymbol{A}$ 的最小多项式 $m(x)$。

引理 11.2 若 $m(x)$ 为一已知多项式，λ 为一常数，$m(\lambda) \neq 0$。a_0，$a_1, \cdots, a_{n-1}$ 为一组常数，则存在次数不大于 $(n-1)$ 的多项式 $q(x)$ 使

$$p^{(k)}(\lambda) = a_k, k = 0, 1, \cdots, n - 1,$$

这里

$$p(x) = m(x)q(x)。$$

证　在乘积求导的 Leibniz 公式

$$p^{(k)}(x) = [m(x)q(x)]^{(k)} = \sum_{i=0}^{k} C_k^i m^{(i)}(x) \cdot q^{(k-i)}(x)$$

中，依次令 $k = 0,1,\cdots,n-1$ 得

$$\begin{cases} p(\lambda) = m(\lambda)q(\lambda), \\ p'(\lambda) = m(\lambda)q'(\lambda) + m'(\lambda)q(\lambda), \\ \qquad\vdots \\ p^{(n-1)}(\lambda) = m(\lambda)q^{(n-1)}(\lambda) + C_{n-1}^1 m'(\lambda) \cdot q^{(n-2)}(\lambda) + \cdots, \end{cases}$$

于是，由　$p^{(k)}(\lambda) = a_k \quad (k = 0,\cdots,n-1)$ 可解出

$$\begin{cases} q(\lambda) = \dfrac{a_0}{m(\lambda)} = b_0, \\ q'(\lambda) = \dfrac{1}{m(\lambda)}[a_1 - m'(\lambda) \cdot b_0] = b_1, \\ \qquad\vdots \\ q^{(n-1)}(\lambda) = \dfrac{1}{m(\lambda)}\left[a_{n-1} - \sum_{i=1}^{n-1} C_{n-1}^i m^{(i)}(\lambda) \cdot a_{n-1-i}\right] = b_{n-1}。 \end{cases}$$

反之，由　$q^{(k)}(\lambda) = b_k \quad (k = 0,\cdots,n-1)$ 可推得

$$p^{(k)}(\lambda) = a_k \quad (k = 0,\cdots,n-1)。$$

于是，由 Taylor 公式，多项式

$$b_0 + b_1(x-\lambda) + \cdots + \frac{b_{n-1}}{(n-1)!}(x-\lambda)^{n-1}$$

就是所求多项式 $q(x)$。

定理 11.3　若方阵 $\boldsymbol{A}$ 的最小多项式

$$m(x) = (x-\lambda_1)^{\alpha_1}\cdots(x-\lambda_t)^{\alpha_t},$$

$\lambda_1,\cdots,\lambda_t$ 两两不同，$\alpha_1 + \cdots + \alpha_t = n$，则任意 $\boldsymbol{A}$ 的矩阵函数 $f(\boldsymbol{A})$ 有且仅有一个次数小于 n 的同值多项式 $p(\boldsymbol{A})$。

证 记 $m_i(x)=\dfrac{m(x)}{(x-\lambda_i)^{\alpha_i}}\quad(i=1,\cdots,t)$,则

$m_i(\lambda_i)\neq 0$,由引理 11.2,存在多项式

$$p_i(x)=m_i(x)q_i(x),$$

满足

$$p_i^{(k)}(\lambda_i)=f^{(k)}(\lambda_i),\ k=0,\cdots,\alpha_i-1,$$

且

$$\begin{aligned}\deg p_i(x)&=\deg m_i(x)+\deg q_i(x)\\&\leqslant \deg m_i(x)+\alpha_i-1\\&\leqslant \alpha_1+\cdots+\alpha_t-1=n-1,\end{aligned}$$

又因为 $j\neq i$ 时,$(x-\lambda_j)^{\alpha_j}\mid p_i(x)$,故

$$p_i^{(k)}(\lambda_j)=0,\ k=0,\cdots,\alpha_j-1,$$

因而多项式

$$p(x)=p_1(x)+\cdots+p_t(x)$$

与 $f(x)$在 $\boldsymbol{A}$ 的谱上的值相同,且其次数小于 n。

若另有次数小于 n 的 $f(x)$的同值多项式 $q(x)$,则

$$p(\boldsymbol{A})=q(\boldsymbol{A})。$$

由命题 11.3 得

$$m(x)\mid p(x)-q(x),$$

但 $p(x)-q(x)$ 的次数小于 $m(x)$ 的次数,故必 $p(x)-q(x)=0$,即 $p(x)=q(x)$,定理得证。

计算矩阵函数的常用方法是寻求低次同值多项式,计算低次同值多项式的值即得到矩阵函数的值。本书中,我们提供三种方法:待定系数法、插值法与待定矩阵法。

三、待定系数法

若 A 为方阵，A 的最小多项式

$$m(x) = (x - \lambda_1)^{\alpha_1} \cdots (x - \lambda_t)^{\alpha_t},$$

$\lambda_1, \cdots, \lambda_t$ 两两不同；$\alpha_1 + \cdots + \alpha_t = n$。则矩阵函数 $f(\boldsymbol{A})$有次数小于 n 的同值多项式 $p(x)$，设

$$p(x) = a_0 + a_1 x + \cdots + a_{n-1} x^{n-1} \quad (a_0, \cdots, a_{n-1} \text{ 待定}),$$

由 $p(x)$与 $f(x)$在 A 的谱上取值相同，建立含 $a_0, \cdots, a_{n-1}$ 的 n 个线性方程式(n 个未知量 $a_0, \cdots, a_{n-1}$)，解之，得 $a_0, \cdots, a_{n-1}$ 的值，从而得

$$f(\boldsymbol{A}) = a_0 \boldsymbol{E} + a_1 \boldsymbol{A} + \cdots + a_{n-1} \boldsymbol{A}^{n-1}。$$

例 11.2 若 $\boldsymbol{A} = \begin{pmatrix} 1 & 4 \\ 3 & 2 \end{pmatrix}$，求 $e^{\boldsymbol{A}}$。

解 $\boldsymbol{A}$ 的特征多项式

$$|\lambda \boldsymbol{E} - \boldsymbol{A}| = \begin{vmatrix} \lambda - 1 & -4 \\ -3 & \lambda - 2 \end{vmatrix} = \lambda^2 - 3\lambda - 10 = (\lambda - 5)(\lambda + 2),$$

这也是 $\boldsymbol{A}$ 的最小多项式。于是，$f(x) = e^x$ 关于 $\boldsymbol{A}$ 的同值多项式 $p(x)$ 为一次多项式

$$p(x) = ax + b。$$

由

$$p(5) = f(5) = e^5, p(-2) = f(-2) = e^{-2},$$

得

$$\begin{cases} 5a + b = e^5, \\ -2a + b = e^{-2}, \end{cases}$$

解之，得

$$p(x)=\frac{1}{7}(e^{5}-e^{-2})x+\frac{1}{7}(2e^{5}+5e^{-2}),$$

于是

$$e^{A}=p(A)=\frac{1}{7}(e^{5}-e^{-2})A+\frac{1}{7}(2e^{5}+5e^{-2})E。$$

例 11.3 已知 $A=\begin{pmatrix}-2&2&-2&4\\-1&2&-1&1\\0&0&1&0\\-2&1&-1&4\end{pmatrix}$，求 $\cos\pi A$。

解 A 的最小多项式为

$$m(x)=(x-1)^{2}(x-2),$$

于是，$f(A)$有二次同值多项式 $p(A)$

$$p(x)=ax^{2}+bx+c,$$

由

$$\begin{cases}p(2)=f(2)=\cos 2\pi=1,\\ p(1)=f(1)=\cos\pi=-1,\\ p'(1)=f'(1)=-\pi\sin\pi=0,\end{cases}$$

得

$$\begin{cases}4a+2b+c=1,\\ a+b+c=-1,\\ 2a+b=0,\end{cases}$$

解之，得

$$p(x)=2x^{2}-4x+1,$$

于是

$$\cos \pi \boldsymbol{A} = p(\boldsymbol{A}) = 2\boldsymbol{A}^2 - 4\boldsymbol{A} + \boldsymbol{E} = \begin{pmatrix} -3 & 0 & 0 & 4 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ -2 & 0 & 0 & 3 \end{pmatrix}。$$

四、插值法

插值法即用求插值多项式的方法求同值多项式。

(1) Lagrange 插值

设方阵 $\boldsymbol{A}$ 的最小多项式 $m(x)$ 无重根：

$$m(x) = (x - \lambda_1)\cdots(x - \lambda_t),$$

记

$$m_i(x) = \frac{m(x)}{(x - \lambda_i)},\ l_i(x) = \frac{m_i(x)}{m_i(\lambda_i)} = \prod_{j \neq i} \frac{x - \lambda_j}{\lambda_i - \lambda_j},$$

则

$$l_i(x_j) = \delta_{ij},\ i, j = 1, \cdots, t,$$

于是多项式

$$p(x) = \sum_{i=1}^{t} f(\lambda_i) l_i(x)$$

与函数 $f(x)$ 在 $\boldsymbol{A}$ 的谱上取相同值，即

$$p(\lambda_i) = f(\lambda_i),\ i = 1, \cdots, t,$$

故

$$f(\boldsymbol{A}) = p(\boldsymbol{A}) = \sum_{i=1}^{t} f(\lambda_i) l_i(\boldsymbol{A})。$$

例 11.4 已知 $A = \begin{pmatrix} 2 & -3 & 0 \\ 1 & -2 & 0 \\ 2 & 3 & \sqrt{3} \end{pmatrix}$，求 $\operatorname{arc\,tan}\boldsymbol{A}$。

解 A 的特征多项式

$$m(x)=|xE-A|=(x-1)(x+1)(x-\sqrt{3})$$

也是 A 的最小多项式，无重根。

$$l_1(x)=\frac{(x+1)(x-\sqrt{3})}{(1+1)(1-\sqrt{3})}=\frac{(x+1)(x-\sqrt{3})}{2(1-\sqrt{3})},$$

$$l_2(x)=\frac{(x-1)(x-\sqrt{3})}{(-1-1)(-1-\sqrt{3})}=\frac{(x-1)(x-\sqrt{3})}{2(1+\sqrt{3})},$$

$$l_3(x)=\frac{(x-1)(x+1)}{(\sqrt{3}-1)(\sqrt{3}+1)}=\frac{x^2-1}{2},$$

于是

$$\begin{aligned}\arctan \boldsymbol{A} &= \arctan 1\cdot l_1(\boldsymbol{A})+\arctan(-1)\cdot l_2(\boldsymbol{A})+\\ &\quad \arctan\sqrt{3}\cdot l_3(\boldsymbol{A})\\ &=\frac{\pi}{4}\cdot\frac{(\boldsymbol{A}+\boldsymbol{E})(\boldsymbol{A}-\sqrt{3}\boldsymbol{E})}{2(1-\sqrt{3})}-\\ &\quad \frac{\pi}{4}\,\frac{(\boldsymbol{A}-\boldsymbol{E})(\boldsymbol{A}-\sqrt{3}\boldsymbol{E})}{2(1+\sqrt{3})}+\frac{\pi}{3}\,\frac{(\boldsymbol{A}^2-\boldsymbol{E})}{2}\\ &=\frac{\pi}{24}[5\boldsymbol{A}^2+3\sqrt{3}\boldsymbol{A}-8\boldsymbol{E}]。\end{aligned}$$

(2) Taylor 插值

设方阵 A 的最小多项式 $m(x)$ 仅有单根：

$$m(x)=(x-\lambda_1)^{\alpha_1}。$$

多项式

$$p(x)=f(\lambda_1)+f'(\lambda_1)(x-\lambda_1)+\cdots+\frac{f^{(\alpha_1-1)}(\lambda_1)}{(\alpha_1-1)!}(x-\lambda_1)^{\alpha_1-1}$$

与函数 $f(\lambda)$在 A 的谱上取相同值，即

$$p^{(k)}(\lambda_1)=f^{(k)}(\lambda_1),\ k=0,1,\cdots,\alpha_1-1,$$

故

$$f(\boldsymbol{A}) = f(\lambda_1)\boldsymbol{E} + f'(\lambda_1)(\boldsymbol{A} - \lambda_1\boldsymbol{E}) + \cdots + \frac{f^{(\alpha_1-1)}(\lambda_1)}{(\alpha_1 - 1)!}(\boldsymbol{A} - \lambda_1\boldsymbol{E})^{\alpha_1-1}。$$

例 11.5 已知 $\boldsymbol{A} = \begin{pmatrix} 0 & -1 \\ 4 & 4 \end{pmatrix}$，求 $f(\boldsymbol{A}) = \arcsin \dfrac{\boldsymbol{A}}{4}$。

解 $\boldsymbol{A}$ 的特征多项式

$$m(x) = |x\boldsymbol{E} - \boldsymbol{A}| = \begin{vmatrix} x & 1 \\ -4 & x-4 \end{vmatrix} = x^2 - 4x + 4 = (x-2)^2,$$

$$f(2) = \arcsin\frac{1}{2} = \frac{\pi}{6},\ f'(2) = \frac{1}{\sqrt{4^2 - x^2}}\bigg|_{x=2} = \frac{1}{2\sqrt{3}},$$

于是

$$\arcsin\frac{\boldsymbol{A}}{4} = f(\boldsymbol{A}) = f(2)\boldsymbol{E} + f'(2)[\boldsymbol{A} - 2\boldsymbol{E}]$$

$$= \begin{pmatrix} \dfrac{\pi}{6} - \dfrac{1}{\sqrt{3}} & -\dfrac{1}{2\sqrt{3}} \\ \dfrac{2}{\sqrt{3}} & \dfrac{\pi}{6} + \sqrt{3} \end{pmatrix}。$$

例 11.6 已知

$$\boldsymbol{A} = \begin{pmatrix} 3 & 1 & 0 & 0 \\ -4 & -1 & 0 & 0 \\ 3 & 1 & 2 & 1 \\ -3 & -1 & -1 & 0 \end{pmatrix},$$

求 $\arctan \boldsymbol{A}$。

解 $\boldsymbol{A}$ 的特征多项式为

$$|x\boldsymbol{E} - \boldsymbol{A}| = \begin{vmatrix} x-3 & -1 & 0 & 0 \\ 4 & x+1 & 0 & 0 \\ -3 & -1 & x-2 & -1 \\ 3 & 1 & 1 & x \end{vmatrix}$$

$$= \begin{vmatrix} x-3 & -1 \\ 4 & x+1 \end{vmatrix} \cdot \begin{vmatrix} x-2 & -1 \\ 1 & x \end{vmatrix}$$
$$= (x-1)^2(x-1)^2 = (x-1)^4,$$

A 的最小多项式(因为 $(A-E)^3=0$) 为

$$m(x) = (x-1)^3。$$

设 $f(x) = \arctan x$,则 $f'(x) = \dfrac{1}{1+x^2}$, $f''(x) = \dfrac{-2x}{(1+x^2)^2}$,于是

$$f(1) = \frac{\pi}{4},\ f'(1) = \frac{1}{2},\ f''(1) = -\frac{1}{2},$$

故

$$\begin{aligned}
\arctan A &= f(A) \\
&= f(1)E + f'(1)(A-E) + \frac{f''(1)}{2!}(A-E)^2 \\
&= \frac{\pi}{4}\begin{pmatrix} 1 & & & \\ & 1 & & \\ & & 1 & \\ & & & 1 \end{pmatrix} + \frac{1}{2}\begin{pmatrix} 2 & 1 & 0 & 0 \\ -4 & -2 & 0 & 0 \\ 3 & 1 & 1 & 1 \\ -3 & -1 & -1 & -1 \end{pmatrix} \\
&\quad - \frac{1}{4}\begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 2 & 1 & 0 & 0 \\ -2 & -1 & 0 & 0 \end{pmatrix} \\
&= \frac{1}{4}\begin{pmatrix} \pi+4 & 2 & 0 & 0 \\ -8 & \pi-4 & 0 & 0 \\ 4 & 1 & \pi+2 & 2 \\ -4 & -1 & -2 & \pi-2 \end{pmatrix}。
\end{aligned}$$

(3) 广义 Taylor 插值

若方阵 A 的最小多项式为

$$m(x) = (x-\lambda_1)^{a_1}(x-\lambda_2),$$

其中 $\lambda_1 \neq \lambda_2$，函数 $f(x)$关于 $\boldsymbol{A}$ 有次数不大于 α_1 的同值多项式 $p(x)$，可设

$$p(x) = f(\lambda_1) + f'(\lambda_1)(x - \lambda_1) + \cdots + \frac{f^{(\alpha_1-1)}(\lambda_1)}{(\alpha_1 - 1)!}(x - \lambda_1)^{\alpha_1 - 1} + a(x - \lambda_1)^{\alpha_1},$$

系数 a 待定，由

$$p(\lambda_2) = f(\lambda_2)$$

解出，即

$$a = \frac{1}{(\lambda_2 - \lambda_1)^{\alpha_1}}\Big[f(\lambda_2) - f(\lambda_1) - f'(\lambda_1)(\lambda_2 - \lambda_1) - \cdots - \frac{f^{(\alpha_1-1)}(\lambda_1)}{(\alpha_1 - 1)!}(\lambda_2 - \lambda_1)^{\alpha_1 - 1}\Big]。$$

易验证，$p(x)$ 与 $f(x)$在 $\boldsymbol{A}$ 的谱上取相同值，即 $p(\lambda_2) = f(\lambda_2)$，$p^{(k)}(\lambda_1) = f^{(k)}(\lambda_1)(k = 0, \cdots, \alpha_1 - 1)$，故 $p(x)$为 $f(x)$关于 $\boldsymbol{A}$ 的同值多项式，于是

$$\begin{aligned} f(\boldsymbol{A}) &= p(\boldsymbol{A}) \\ &= f(\lambda_1)\boldsymbol{E} + f'(\lambda_1)(\boldsymbol{A} - \lambda_1\boldsymbol{E}) + \cdots + \\ &\quad \frac{f^{(\alpha_1-1)}(\lambda_1)}{(\alpha_1 - 1)!}(\boldsymbol{A} - \lambda_1\boldsymbol{E})^{\alpha_1 - 1} + a(\boldsymbol{A} - \lambda_1\boldsymbol{E})^{\alpha_1}。\end{aligned}$$

例 11.7 已知

$$\boldsymbol{A} = \begin{pmatrix} 3 & 1 & -3 \\ -7 & -2 & 9 \\ -2 & -1 & 4 \end{pmatrix},$$

求 $e^{\boldsymbol{A}}$。

解 $\boldsymbol{A}$ 的特征多项式(也是最小多项式)为

$$m(x) = (x - 2)^2(x - 1)。$$

设 $f(x)=\mathrm{e}^x$，$f'(x)=\mathrm{e}^x$，$f(2)=f'(2)=\mathrm{e}^2$，令

$$p(x)=\mathrm{e}^2+\mathrm{e}^2(x-2)+a(x-2)^2,$$

由

$$p(1)=f(1)=\mathrm{e},$$

得

$$\mathrm{e}^2-\mathrm{e}^2+a=\mathrm{e},$$

即

$$a=\mathrm{e},$$

于是

$$\begin{aligned}\mathrm{e}^{\boldsymbol{A}}&=f(\boldsymbol{A})=p(\boldsymbol{A})=\mathrm{e}^2\boldsymbol{E}+\mathrm{e}^2(\boldsymbol{A}-2\boldsymbol{E})+\mathrm{e}(\boldsymbol{A}-2\boldsymbol{E})^2\\&=\mathrm{e}^2\begin{pmatrix}1&&\\&1&\\&&1\end{pmatrix}+\mathrm{e}^2\begin{pmatrix}1&1&-3\\-7&-4&9\\-2&-1&2\end{pmatrix}+\mathrm{e}\begin{pmatrix}0&0&0\\3&0&3\\1&0&1\end{pmatrix}\\&=\mathrm{e}\begin{pmatrix}2\mathrm{e}&\mathrm{e}&-3\mathrm{e}\\-7\mathrm{e}+3&-3\mathrm{e}&9\mathrm{e}+3\\-2\mathrm{e}+1&-\mathrm{e}&3\mathrm{e}+1\end{pmatrix}。\end{aligned}$$

五、待定矩阵法

若 $\boldsymbol{A}\sim\boldsymbol{J}=\begin{pmatrix}\boldsymbol{J}_1&&\\&\ddots&\\&&\boldsymbol{J}_s\end{pmatrix}$，$\boldsymbol{J}_i=\lambda_i\boldsymbol{E}+\boldsymbol{H}$ 为 Jordan 块

$$(i=1,\cdots,s)$$
$$\boldsymbol{A}=\boldsymbol{T}\boldsymbol{J}\boldsymbol{T}^{-1},$$

于是

$$f(\boldsymbol{A}) = \boldsymbol{T}\begin{pmatrix} f(\boldsymbol{J}_1) & & & \\ & \ddots & & \\ & & \ddots & \\ & & & f(\boldsymbol{J}_s) \end{pmatrix}\boldsymbol{T}^{-1}$$

$$= \sum_{i=1}^{s}\boldsymbol{T}\begin{pmatrix} 0 & & & & \\ & \ddots & & & \\ & & f(\boldsymbol{J}_i) & & \\ & & & \ddots & \\ & & & & 0 \end{pmatrix}\boldsymbol{T}^{-1},$$

而

$$f(\boldsymbol{J}_i) = f(\lambda_i\boldsymbol{E} + \boldsymbol{H}) = f(\lambda_i)\boldsymbol{E} + f'(\lambda_i)\boldsymbol{H} + \frac{f''(\lambda_i)}{2!}\boldsymbol{H}^2 + \cdots,$$

于是

$$f(\boldsymbol{A}) = \sum_{i=1}^{s}[f(\lambda_i)\boldsymbol{B}_{i0} + f'(\lambda_i)\boldsymbol{B}_{i1} + f''(\lambda_i)\boldsymbol{B}_{i2} + \cdots],$$

这里：

$$\boldsymbol{B}_{ij} = \frac{1}{j!}\boldsymbol{T}\begin{pmatrix} 0 & & & & \\ & \ddots & & & \\ & & \boldsymbol{H}^{j-1} & & \\ & & & \ddots & \\ & & & & 0 \end{pmatrix}\boldsymbol{T}^{-1}, i = 1,\cdots,s; j = 0,\cdots$$

由 $\boldsymbol{A}$ 本身唯一确定与 f 无关。若把相同的含 $f^{(k)}(\lambda_i)$的项合并在一起，得下列待定矩阵法。

待定矩阵法 若方阵 $\boldsymbol{A}$ 的最小多项式

$$m(x) = (x - \lambda_1)^{\alpha_1}\cdots(x - \lambda_t)^{\alpha_t},$$

这里 $\lambda_1,\cdots,\lambda_t$ 两两不等。可设矩阵函数（$\{f^{(\cdot)}(\lambda_i)\}$为 f 在 $\boldsymbol{A}$ 的谱上的值）：

$$f(A) = \sum_{i=1}^{t} \cdot \sum_{j=0}^{a_i-1} f^{(j)}(\lambda_i)\boldsymbol{B}_{ij},$$

这里 $\boldsymbol{B}_{ij}$ 完全由 $\boldsymbol{A}$ 本身确定，与 f 无关，为待定矩阵。可适当选择一些特殊函数代替 $f(x)$，以求出 $\boldsymbol{B}_{ij}$。从而求得矩阵函数 $f(\boldsymbol{A})$。

例 11.8 若方阵 $\boldsymbol{A}$ 的最小多项式为

$$m(x) = (x-1)^2(x+1),$$

试将 $e^{\boldsymbol{A}}$，$\sin \boldsymbol{A}$，$\arctan \boldsymbol{A}$ 表示为 $\boldsymbol{A}$ 的多项式。

解 $f(\boldsymbol{A})$ 由 $\boldsymbol{A}$ 谱上值 $f(1)$，$f'(1)$，$f(-1)$ 确定，设

$$f(A) = f(1)\boldsymbol{B} + f'(1)\boldsymbol{C} + f(-1)\boldsymbol{D},$$

令

$$f(x) = (x-1)^2,$$

得

$$f(1) = f'(1) = 0,\ f(-1) = 4,$$

$$(\boldsymbol{A} - \boldsymbol{E})^2 = 4\boldsymbol{D},$$

即

$$\boldsymbol{D} = \frac{1}{4}(\boldsymbol{A} - \boldsymbol{E})^2 = \frac{1}{4}(\boldsymbol{A}^2 - 2\boldsymbol{A} + \boldsymbol{E})。$$

令

$$f(x) = x - 1,$$

得

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

$$\boldsymbol{A} - \boldsymbol{E} = \boldsymbol{C} - 2\boldsymbol{D},$$

故

$$C = 2D + A - E = (A - E)\left[\frac{A - E}{2} + E\right] = \frac{A^2 - E}{2}。$$

令 $f(x) = x + 1$，得

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

$$A + E = 2B + C,$$

$$B = \frac{1}{2}(A + E - C) = \frac{1}{4}(2A + 3E - A^2),$$

于是

$$\begin{aligned} f(A) = &\frac{1}{4}[-f(1) + 2f'(1) + f(-1)]A^2 \\ &+ \frac{1}{2}[f(1) - f(-1)]A \\ &+ \frac{1}{4}[3f(1) - 2f'(1) + f(-1)]E。\end{aligned}$$

故

$$e^A = \frac{1}{4}\left(e + \frac{1}{e}\right)A^2 + \frac{1}{2}\left(e - \frac{1}{e}\right)A + \frac{1}{4}\left(e + \frac{1}{e}\right)E,$$

$$\sin A = \frac{1}{2}(-\sin 1 + \cos 1)A^2 + \sin 1 \cdot A + \frac{1}{2}(\sin 1 - \cos 1)E,$$

$$\arctan A = \frac{1}{4}\left(-\frac{\pi}{2} + 1\right)A^2 + \frac{\pi}{4}A + \frac{1}{4}\left(\frac{\pi}{2} - 1\right)E。$$

例 11.9 已知

$$A = \begin{pmatrix} 3 & 1 & -3 \\ -7 & -2 & 9 \\ -2 & -1 & 4 \end{pmatrix},$$

求 e^A，$\ln A$。

解 A 的特征多项式，也是最小多项式为

$$m(x) = (x - 2)^2(x - 1)。$$

$f(A)$由 f 在 A 谱上的值 $f(2)$,$f'(2)$,$f(1)$确定,设

$$f(\boldsymbol{A}) = f(2)\boldsymbol{B} + f'(2)\boldsymbol{C} + f(1)\boldsymbol{D}。$$

令

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

得

$$(\boldsymbol{A} - 2\boldsymbol{E})^2 = \boldsymbol{D},\text{或 } \boldsymbol{D} = \boldsymbol{A}^2 - 4\boldsymbol{A} + 4\boldsymbol{E}。$$

令 $f(x) = x - 2$, $f(2) = 0$, $f'(2) = 1$, $f(1) = -1$,得

$$\boldsymbol{A} - 2\boldsymbol{E} = \boldsymbol{C} - \boldsymbol{D},$$

$$\boldsymbol{C} = \boldsymbol{A} - 2\boldsymbol{E} + \boldsymbol{D} = \boldsymbol{A}^2 - 3\boldsymbol{A} + 2\boldsymbol{E}。$$

令 $f(x) = x - 1$, $f(2) = f'(2) = 1$, $f(1) = 0$,得

$$\boldsymbol{A} - \boldsymbol{E} = \boldsymbol{B} + \boldsymbol{C},$$

$$\boldsymbol{B} = \boldsymbol{A} - \boldsymbol{E} - \boldsymbol{C} = -\boldsymbol{A}^2 + 4\boldsymbol{A} - 3\boldsymbol{E},$$

于是

$$\begin{aligned} f(\boldsymbol{A}) = &[-f(2) + f'(2) + f(1)]\boldsymbol{A}^2 \\ &+ [4f(2) - 3f'(2) - 4f(1)]\boldsymbol{A} \\ &+ [-3f(2) + 2f'(2) + 4f(1)]\boldsymbol{E}。\end{aligned}$$

若 $f(x) = \mathrm{e}^x$,则 $f(2) = f'(2) = \mathrm{e}^2$, $f(1) = \mathrm{e}$,于是

$$\begin{aligned} \mathrm{e}^{\boldsymbol{A}} &= \mathrm{e}\boldsymbol{A}^2 + (\mathrm{e}^2 - 4\mathrm{e})\boldsymbol{A} + (-\mathrm{e}^2 + 4\mathrm{e})\boldsymbol{E} \\ &= \mathrm{e}\boldsymbol{A}^2 + (\mathrm{e}^2 - 4\mathrm{e})(\boldsymbol{A} - \boldsymbol{E}) \\ &= \mathrm{e}\begin{pmatrix} 8 & 4 & -12 \\ -25 & -12 & 39 \\ -7 & -4 & 13 \end{pmatrix} + (\mathrm{e}^2 - 4\mathrm{e})\begin{pmatrix} 2 & 1 & -3 \\ -7 & -3 & 9 \\ -2 & -1 & 3 \end{pmatrix} \\ &= \mathrm{e}\begin{pmatrix} 2\mathrm{e} & \mathrm{e} & -3\mathrm{e} \\ -7\mathrm{e} + 3 & -3\mathrm{e} & 9\mathrm{e} + 3 \\ -2\mathrm{e} + 1 & -\mathrm{e} & 3\mathrm{e} + 1 \end{pmatrix}。\end{aligned}$$

若 $f(x)=\ln x,\ f'(x)=\dfrac{1}{x}$，则 $f(2)=\ln 2, f'(2)=\dfrac{1}{2},\ f(1)=0$，于是

$$
\begin{aligned}
\ln \boldsymbol{A} &= \left(-\ln 2+\frac{1}{2}\right)\boldsymbol{A}^2+\left(4\ln 2-\frac{3}{2}\right)\boldsymbol{A}+(-3\ln 2+1)\boldsymbol{E}\\
&=\left(-\ln 2+\frac{1}{2}\right)\begin{pmatrix} 8 & 4 & -12\\ -25 & -12 & 39\\ -7 & -4 & 13\end{pmatrix}\\
&\quad+\left(4\ln 2-\frac{3}{2}\right)\begin{pmatrix} 3 & 1 & -3\\ -7 & -2 & 9\\ -2 & -1 & 4\end{pmatrix}\\
&\quad+(-3\ln 2+1)\begin{pmatrix} 1 & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & 1\end{pmatrix}\\
&=\begin{pmatrix} \ln 2+\frac{1}{2} & \frac{1}{2} & \frac{3}{2}\\ -3\ln 2-2 & \ln 2-2 & -3\ln 2+6\\ -\ln 2-\frac{1}{2} & -\frac{1}{2} & \frac{3}{2}\end{pmatrix}。
\end{aligned}
$$

§4　矩阵函数的应用

一、矩阵函数的两个性质

性质 1　若 $f(\boldsymbol{A}),g(\boldsymbol{A})$ 都是方阵 $\boldsymbol{A}$ 的函数，则

$$f(\boldsymbol{A})g(\boldsymbol{A})=g(\boldsymbol{A})f(\boldsymbol{A})。$$

证　设 $p(x),q(x)$ 分别为 $f(x),g(x)$ 关于方阵 $\boldsymbol{A}$ 的同值多项式，则

$$f(\boldsymbol{A})g(\boldsymbol{A})=p(\boldsymbol{A})q(\boldsymbol{A})=q(\boldsymbol{A})\cdot p(\boldsymbol{A})=g(\boldsymbol{A})f(\boldsymbol{A})。$$

性质 2　若 $f(\boldsymbol{A}),g(\boldsymbol{A}),h(\boldsymbol{A})$ 都是方阵 $\boldsymbol{A}$ 的函数，且

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

则

$$f(\boldsymbol{A})g(\boldsymbol{A}) = h(\boldsymbol{A})。$$

证 设 $p(x)$, $q(x)$, $r(x)$分别是 $f(x)$, $g(x)$, $h(x)$关于 $\boldsymbol{A}$ 的同值多项式。λ 是 $\boldsymbol{A}$ 的特征值。于是 $r(x)$在 $\boldsymbol{A}$ 的谱上的任一个值

$$\begin{aligned} r^{(i)}(\lambda) &= h^{(i)}(\lambda) = (f(x)g(x))^{(i)}\Big|_{x=\lambda} \\ &= \sum_{k=0}^{i} C_i^k f^{(k)}(\lambda)g^{(i-k)}(\lambda) \\ &= \sum_{k=0}^{i} C_i^k p^{(k)}(\lambda)q^{(i-k)}(\lambda) \\ &= (p(x)q(x))^{(i)}\Big|_{x=\lambda}, \end{aligned}$$

即 $r(x)$与 $l(x) = p(x)q(x)$ 在 $\boldsymbol{A}$ 的谱上的值相同,故

$$r(\boldsymbol{A}) = l(\boldsymbol{A}),$$

由多项式矩阵赋值的性质得

$$l(\boldsymbol{A}) = p(\boldsymbol{A})q(\boldsymbol{A}),$$

因此

$$h(\boldsymbol{A}) = r(\boldsymbol{A}) = l(\boldsymbol{A}) = p(\boldsymbol{A})q(\boldsymbol{A}) = f(\boldsymbol{A})g(\boldsymbol{A})。$$

推论 1 若方阵 $\boldsymbol{A}$ 的特征值的模全小于 1,则 $\boldsymbol{E}-\boldsymbol{A}$ 可逆,且

$$(\boldsymbol{E} - \boldsymbol{A})^{-1} = \boldsymbol{E} + \boldsymbol{A} + \boldsymbol{A}^2 + \cdots + \boldsymbol{A}^n + \cdots = \sum_{n=0}^{+\infty} \boldsymbol{A}^n。$$

证 因为 $(1-x)\cdot\dfrac{1}{1-x} = (1-x)\cdot(1-x)^{-1} = 1, (|x| < 1)$

故

$$(\boldsymbol{E} - \boldsymbol{A})\cdot(\boldsymbol{E} - \boldsymbol{A})^{-1} = \boldsymbol{E},$$

即

$(\boldsymbol{E}-\boldsymbol{A})^{-1}$ 为 $(\boldsymbol{E}-\boldsymbol{A})$ 的逆，而 $\boldsymbol{A}$ 的特征值的模全小于 1 时

$$(\boldsymbol{E}-\boldsymbol{A})^{-1}=\boldsymbol{E}+\boldsymbol{A}+\boldsymbol{A}^2+\cdots。$$

推论 2 若 A 为方阵，则 e^A 可逆，且其逆为 e^{-A}，即

$$(e^A)^{-1}=e^{-A}。$$

证 因为

$$e^x\cdot e^{-x}=e^0=1,$$

故

$$e^A\cdot e^{-A}=\boldsymbol{E},$$

即 e^{-A}为 e^A 的逆。

例 11.10 已知 $A=\begin{pmatrix}3&0\\-1&3\end{pmatrix}$，求 e^A 的逆。

解 $-A=\begin{pmatrix}-3&0\\1&-3\end{pmatrix}$ 为 Jordan 块，于是

$$(e^A)^{-1}=e^{-A}=\begin{pmatrix}e^{-3}&0\\e^{-3}&e^{-3}\end{pmatrix}。$$

二、解线性方程组的迭代法的原理

定理 11.4 给定 s 元、s 式线性方程组

$$\boldsymbol{AX}=\boldsymbol{\beta}。\tag{1}$$

令 $\boldsymbol{A}=\boldsymbol{E}-\boldsymbol{B}$，化式(1) 为同解的线性方程组

$$\boldsymbol{X}=\boldsymbol{BX}+\boldsymbol{\beta}。\tag{2}$$

如果 s 阶方阵 $\boldsymbol{B}$ 的特征值的模全小于 1，任取 s 维向量 $\boldsymbol{X}_0$，令

$$\boldsymbol{X}_{n+1}=\boldsymbol{BX}_n+\boldsymbol{\beta},$$

则 $n \to \infty$ 时，$\boldsymbol{X}_n$ 趋于式(2)，即式(1)的解。

证
$$\boldsymbol{X}_n = \boldsymbol{B}\boldsymbol{X}_{n-1} + \boldsymbol{\beta}, \qquad ①$$
$$\boldsymbol{X}_{n-1} = \boldsymbol{B}\boldsymbol{X}_{n-2} + \boldsymbol{\beta}, \qquad ②$$
$$\vdots \qquad \vdots$$
$$\vdots \qquad \vdots$$
$$\boldsymbol{X}_1 = \boldsymbol{B}\boldsymbol{X}_0 + \boldsymbol{\beta}, \qquad ⓝ$$

① + $\boldsymbol{B}$② + $\boldsymbol{B}^2$③ + … + $\boldsymbol{B}^{n-1}$ⓝ，得

$$\boldsymbol{X}_n = \boldsymbol{B}^n\boldsymbol{X}_0 + (\boldsymbol{E} + \boldsymbol{B} + \cdots + \boldsymbol{B}^{n-1})\boldsymbol{\beta}。$$

因为 $\boldsymbol{B}$ 的特征值的模全小于 1，故级数 $\sum\limits_{n=0}^{+\infty}\boldsymbol{B}^n$ 收敛，其和为 $(\boldsymbol{E} - \boldsymbol{B})^{-1}$，且 $n \to \infty$ 时，其通项 $\boldsymbol{B}^n \to 0$，即

$$\sum_{n=0}^{+\infty}\boldsymbol{B}^n = (\boldsymbol{E} - \boldsymbol{B})^{-1},$$

$$\lim_{n \to \infty}\boldsymbol{B}^n = 0,$$

所以

$$\lim_{n \to \infty}\boldsymbol{X}_n = \boldsymbol{O} \cdot \boldsymbol{X}_0 + (\boldsymbol{E} - \boldsymbol{B})^{-1}\boldsymbol{B} = \boldsymbol{A}^{-1}\boldsymbol{\beta},$$

这恰好是式(1)的解。

例如，线性方程组

$$\begin{cases} 5x_1 - 2x_2 + 2x_3 = 6, \\ -x_1 + 4x_2 + x_3 = 1, \\ 4x_1 - 5x_2 + 10x_3 = -7, \end{cases}$$

可化为同解方程组

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 & 0.4 & -0.4 \\ 0.25 & 0 & -0.25 \\ -0.4 & 0.5 & 0 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} + \begin{pmatrix} 1.2 \\ 0.25 \\ -0.7 \end{pmatrix},$$

矩阵 $\boldsymbol{B}=\begin{pmatrix} 0 & 0.4 & -0.4 \\ 0.25 & 0 & -0.25 \\ -0.4 & 0.5 & 0 \end{pmatrix}$ 的特征值模全小于 1。令

$$\boldsymbol{X}_0=\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix},\boldsymbol{X}_{n+1}=\boldsymbol{B}\boldsymbol{X}_n+\boldsymbol{\beta},\boldsymbol{\beta}=\begin{pmatrix} 1.2 \\ 0.25 \\ -0.7 \end{pmatrix},$$

当 $n=7$ 时

$$\boldsymbol{X}_7=\begin{pmatrix} 1.9995 \\ 0.9990 \\ -0.9993 \end{pmatrix}$$

与方程组的精确解 $\begin{pmatrix} 2 \\ 1 \\ -1 \end{pmatrix}$ 非常接近。

三、一阶常系数齐次线性微分方程组的一个解法

定理 11.5 若 A 为 s 阶方阵，则满足

$$\begin{cases} \dfrac{\mathrm{d}}{\mathrm{d}t}\boldsymbol{X}(t)=\boldsymbol{A}\boldsymbol{X}(t), \\ \boldsymbol{X}(0)=\boldsymbol{X}_0=\begin{pmatrix} c_1 \\ \vdots \\ c_s \end{pmatrix}, \end{cases}$$

的惟一解为

$$\boldsymbol{X}(t)=\mathrm{e}^{\boldsymbol{A}t}\boldsymbol{X}_0。$$

证
$$\boldsymbol{X}(0)=\mathrm{e}^{\boldsymbol{A}0}\cdot\boldsymbol{X}_0=\boldsymbol{X}_0,$$

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}\boldsymbol{X}(t)&=\frac{\mathrm{d}}{\mathrm{d}t}\mathrm{e}^{\boldsymbol{A}t}\boldsymbol{X}_0=\left(\frac{\mathrm{d}}{\mathrm{d}t}\mathrm{e}^{\boldsymbol{A}t}\right)\boldsymbol{X}_0,\\ &=\boldsymbol{A}\mathrm{e}^{\boldsymbol{A}t}\boldsymbol{X}_0=\boldsymbol{A}\boldsymbol{X}(t),\end{aligned}$$

故

$$X(t)=e^{At}X_0$$

为满足条件的解。又若另有解 $Y(t)$满足：

$$\begin{cases}\dfrac{d}{dt}Y(t)=AY(t),\\ Y(0)=X_0。\end{cases}$$

令　$Z(t)=e^{-At}\cdot Y(t)$，则

$$\begin{aligned}\frac{d}{dt}Z(t)&=\left(\frac{d}{dt}e^{-At}\cdot Y(t)+e^{-At}\frac{d}{dt}Y(t)\right)\\&=-A\cdot e^{-At}Y(t)+e^{-At}AY(t)\\&=-e^{-At}AY(t)+e^{-At}AY(t)=0,\end{aligned}$$

故

$$Z(t)=e^{-At}Y(t)$$

为常向量 α。令 $t=0$，得 $\alpha=Z(0)=Y(0)=X_0$，于是

$$Y(t)=e^{At}\alpha=e^{At}X_0=X(t),$$

所以，$X(t)=e^{At}X_0$ 为惟一解。

例 11.11　求常系数线性齐次方程组

$$\left\{\frac{dX(t)}{dt}=\begin{pmatrix}-7&-7&5\\-8&-8&-5\\0&-5&0\end{pmatrix}X(t)\right.$$

满足初始条件

$$X(0)=\begin{pmatrix}3\\-2\\1\end{pmatrix}$$

的特解。

解　设

$$A=\begin{pmatrix}-7 & -7 & 5\\ -8 & -8 & -5\\ 0 & -5 & 0\end{pmatrix}.$$

则 A 的特征多项式也是最小多项式，为

$$m(x)=|xE-A|=\begin{vmatrix}\lambda+7 & 7 & -5\\ 8 & \lambda+8 & 5\\ 0 & 5 & \lambda\end{vmatrix}$$

$$=(\lambda-5)(\lambda+5)(\lambda+15),$$

利用 Lagrange 插值得

$$\begin{aligned}
e^{At}&=(e^t)^A=\frac{(A+5E)(A+15E)}{(5+5)(5+15)}e^{5t}+\frac{(A-5E)(A+15E)}{(-5-5)(-5+15)}e^{-5t}\\
&\quad+\frac{(A-5E)(A+5E)}{(-15-5)(-15+5)}e^{-15t}\\
&=\frac{e^{5t}}{200}(A^2+20A+75E)+\frac{(-e^{-5t})}{100}\cdot(A^2+10A-75E)\\
&\quad+\frac{e^{-15t}}{200}(A^2-25E)\\
&=\frac{1}{200}(e^{5t}-2e^{-5t}+e^{-15t})A^2+\frac{1}{10}(e^{5t}-e^{-5t})A\\
&\quad+\frac{1}{8}(3e^{5t}+6e^{-5t}+e^{-15t})E,
\end{aligned}$$

用 $A=\begin{pmatrix}-7 & -7 & 5\\ -8 & -8 & -5\\ 0 & -5 & 0\end{pmatrix}$ 代入得

$$e^{At}=\frac{1}{10}\begin{pmatrix}2e^{5t}-4e^{-5t}+4e^{-15t} & -3e^{5t}-e^{-5t}+4e^{-15t} & 5e^{5t}-5e^{-5t}\\ -2e^{5t}-4e^{-5t}+6e^{-15t} & 3e^{5t}+e^{-5t}+6e^{-15t} & -5e^{5t}+5e^{-5t}\\ 2e^{5t}-4e^{-5t}+2e^{-15t} & -3e^{5t}+e^{-5t}+2e^{-15t} & 5e^{5t}+5e^{-5t}\end{pmatrix},$$

所以特解

$$X(t)=e^{At}X(0)=\frac{1}{10}\begin{bmatrix}17e^{5t}+9e^{-5t}+4e^{-15t}\\-17e^{5t}-9e^{-5t}+6e^{-15t}\\17e^{5t}-9e^{-5t}+2e^{-15t}\end{bmatrix}。$$

习　题　11

1. 已知 $A=\begin{bmatrix}1&0&3\\0&2&1\\0&0&1\end{bmatrix}$，$B=\begin{bmatrix}1&0&0\\0&2&1\\3&0&1\end{bmatrix}$，求

(1) $(A+B)(A-B)$；

(2) A^2-B^2。

2. 已知 $A=\begin{bmatrix}1&-2\\3&4\end{bmatrix}$，$B=\begin{bmatrix}7&6\\-9&-2\end{bmatrix}$，验证：$AB=BA$，求

(1) $(A+B)(A-B)$；

(2) A^2-B^2。

3. 试用 Taylor 公式表多项式 x^4+x+1 为 $(x-1)$ 与 $(x+1)$ 的多项式。

4. 已知 $A=\begin{bmatrix}2&0&0\\1&2&0\\0&1&2\end{bmatrix}$，求 e^A，$\sin A$，$\arctan(A-E)$。

5. 已知

$$A=\begin{bmatrix}-1&0&0&&\\1&-1&0&&\\0&1&-1&&\\&&&1&0\\&&&1&1\end{bmatrix},$$

求 e^A，$\cos A$，$\arctan A$。

6. 已知

$$A = \begin{pmatrix} 1 & & & \\ & \frac{1}{2} & & \\ & & \frac{1}{2} & 0 \\ & & 1 & \frac{1}{2} \end{pmatrix},$$

求 $\arcsin A, \ln A, (E+A)^{-1}$。

7. 设 $A_n = \begin{pmatrix} \frac{1}{n} & \frac{2n^2+n}{n^2-n} \\ 1 & \left(1-\frac{2}{n}\right)^n \end{pmatrix}$，求 $\lim\limits_{n\to\infty} A_n$。

8. 若 $\lim\limits_{n\to\infty} A_n = A$，求证：

$$\lim_{n\to\infty} |A_n| = |A|。$$

9. 若 $\lim\limits_{n\to\infty} A^n = B$，求证：

(1) $B^2 = B$；

(2) $(A-B)^n = A^n - B$。

10. 若方阵 A 的特征值的模全小于 1，求证：

$$\lim_{n\to\infty} A^n = 0。$$

11. 求和 $\sum\limits_{n=1}^{+\infty} \begin{pmatrix} \frac{1}{n^1} & \frac{1}{5^n} \\ \frac{(-1)^n}{n} & \frac{n}{2^{n-1}} \end{pmatrix}$。

12. 求和 $\sum\limits_{n=0}^{\infty} \frac{1}{3^n} \begin{pmatrix} 1 & 2 \\ 1 & 1 \end{pmatrix}^n$。

13. 已知 4 阶方阵 A 的特征多项式为 $x^2(x^2-\pi^2)$，求证：

$$\cos A = E - \frac{2A^2}{\pi^2},$$

$$\sin A = A - \frac{A^3}{\pi^2}。$$

14. 若 $\boldsymbol{A}(t)=\begin{pmatrix}\cos t, & \sin t\\ -\sin t, & \cos t\end{pmatrix}$，求

$$\frac{\mathrm{d}}{\mathrm{d}t}\boldsymbol{A}(t),\frac{\mathrm{d}}{\mathrm{d}t}\boldsymbol{A}^{-1}(t),\frac{\mathrm{d}}{\mathrm{d}t}|\boldsymbol{A}(t)|,$$

这里 $|\boldsymbol{A}(t)|$ 表矩阵 $\boldsymbol{A}(t)$ 的行列式。

15. 求证：$\frac{\mathrm{d}}{\mathrm{d}t}\cos \boldsymbol{A}t=-\boldsymbol{A}\sin \boldsymbol{A}t$。

16. 求项数不大于 2 的多项式 $q(x)$ 使

$$p(0)=1,p'(0)=0,p''(0)=2,$$

这里

$$p(x)=q(x)(x^2+x+1)。$$

17. 已知

$$\boldsymbol{A}=\begin{pmatrix}1 & 2 & 0\\ 0 & 2 & 0\\ -2 & -1 & 1\end{pmatrix},$$

用待定系数法求 $\ln \boldsymbol{A}$，$\mathrm{e}^{\boldsymbol{A}}$。

18. 已知

$$\boldsymbol{A}=\begin{pmatrix}1 & 2 & 1\\ 0 & 1 & 3\\ 0 & 0 & -1\end{pmatrix},$$

用待定系数法求 $\arctan \boldsymbol{A}$。

19. 已知方阵 $\boldsymbol{A}$ 的最小多项式为 $m(x)$，试用插值法表 $\mathrm{e}^{\boldsymbol{A}}$，$\arctan \boldsymbol{A}$ 为 $\boldsymbol{A}$ 的多项式：

(1) $m(x)=x^2-1$；

(2) $m(x)=x(x^2-3)$；

(3) $m(x)=(x-1)^3$；

(4) $m(x)=(x-1)^2(x+1)$。

20. 已知方阵 $\boldsymbol{A}$ 的最小多项式为

$$m(x) = (x-1)^2(x-3),$$

试用待定矩阵法将 $\ln \boldsymbol{A}$,$\arctan \boldsymbol{A}$,$\sin \boldsymbol{A}$ 表示为 $\boldsymbol{A}$ 的多项式。

21. 已知

$$\boldsymbol{A} = \begin{pmatrix} 0 & 1 & -1 \\ -4 & 4 & -2 \\ -2 & 1 & 1 \end{pmatrix},$$

求 $e^{\boldsymbol{A}}$,$\ln \boldsymbol{A}$。

22. 已知

$$\boldsymbol{A} = \begin{pmatrix} 2 & -1 & -1 \\ 2 & -1 & -2 \\ -1 & 1 & 2 \end{pmatrix},$$

求 $\arctan \boldsymbol{A}$。

23. 若 λ 为 n 阶方阵 $\boldsymbol{A}$ 的特征值,且 $\boldsymbol{A}$ 满足

$$\sum_{j=1}^{n} |\boldsymbol{A}(i,j)| < 1, \quad i = 1,\cdots,n$$

则

$$|\lambda| < 1。$$

24. 用迭代法求下列线性方程组的近似解:

$$\begin{cases} 5x_1 - 2x_2 + 2x_3 = 6, \\ -x_1 + 4x_2 + x_3 = 1, \\ 4x_1 - 3x_2 + 10x_3 = -6。 \end{cases}$$

25. 求 $\boldsymbol{X}(t)$使

$$\begin{cases} \dfrac{\mathrm{d}}{\mathrm{d}t}\boldsymbol{X}(t) = \begin{pmatrix} 1 & 12 \\ 3 & 1 \end{pmatrix}\boldsymbol{X}, \\ \boldsymbol{X}(0) = \begin{pmatrix} 0 \\ 1 \end{pmatrix}。 \end{cases}$$

26. 求方程组

$$\frac{\mathrm{d}}{\mathrm{d}t}\boldsymbol{X}(t) = \begin{pmatrix} 3 & 1 & -3 \\ -7 & -2 & 9 \\ -2 & -1 & 4 \end{pmatrix}\boldsymbol{X}(t)$$

的通解。

27. 若 $\boldsymbol{A}$ 是实反对称阵，求证：e^{A} 是正交阵。

第十二章　群

本章引进群的基本概念，介绍了“循环群”与“置换群”，导出了群的同态基本定理。

§1　群的概念

一、半群

1. 定义

定义 12.1　定义了满足结合律的二元运算“$\cdot$”的非空集合 A 称为半群，即

$\forall a,b\in A$，有 $a\cdot b\in A$；

$\forall a,b,c\in A$，有

$$(a\cdot b)\cdot c=a\cdot(b\cdot c)。$$

为区别于集合 A，常把上述半群记成 $(A,\cdot)$。

例 12.1　设 A 为一非空集合，由于 A 的子集的交与并仍是 A 的子集，且满足结合律：

$\forall B,C,D\subseteq A$，有

$$(B\cap C)\cap D=B\cap(C\cap D),$$
$$(B\cup C)\cup D=B\cup(C\cup D),$$

故 A 的幂集合，即 A 的全体子集所成的集合，2^A 关于交与并的运算均成半群，分别记成 $(2^A,\cap)$ 与 $(2^A,\cup)$。

例 12.2　设 A 为一非空集合，由于映射的积运算满足结合律，故 A 到 A 的映射全体 A^A 关于映射的积运算成一半群。

例 12.3 正整数集 $\mathbf{Z}^+$ 关于数的加法运算成一半群($\mathbf{Z}^+$,+)。$\mathbf{Z}^+$ 关于数的乘法运算成一半群($\mathbf{Z}^+$,×)。更若在 $\mathbf{Z}^+$ 中定义运算"∘":

$$a \circ b = a + b + ab,$$

则由于

$$\begin{aligned}(a \circ b) \circ c &= (a + b + ab) + c + (a + b + ab)c \\ &= a + b + c + ab + ac + bc + abc,\\ a \circ (b \circ c) &= a + (b + c + bc) + a(b + c + bc) \\ &= a + b + c + ab + ac + bc + abc \\ &= (a \circ b) \circ c,\end{aligned}$$

故($\mathbf{Z}^+$,∘)也成一半群。

例 12.4 由于矩阵的加、乘运算均满足结合律,故数环 T 上的 n 阶方阵全体 $\boldsymbol{T}^{n\times n}$ 关于矩阵的加、乘运算都成半群。

2. 给了半群(A,·),

可以归纳地定义

$$\prod_{i=1}^{n} a_i = a_1 \cdot a_2 \cdot \cdots \cdot a_n = \left(\prod_{i=1}^{n-1} a_i\right) \cdot a_n,$$

这里 $a_i \in A, i = 1,2,\cdots,n$。

半群 A 的运算"·"通常称为乘法。A 中 n 个元的乘积 $a_1 \cdot a_2 \cdot \cdots \cdot a_n$只要次序不变,不管结合的方式,其结果是相同的,即

命题 $\left(\prod_{i=1}^{n} a_i\right)\left(\prod_{j=n+1}^{n+s} a_j\right) = \prod_{i=1}^{n+s} a_i$。

证 对 $n+s$ 用归纳法。$n+s=3$ 时,由于 A 的运算"·"满足结合律,故命题成立。若 $n+s-1$ 时正确,有

$$\begin{aligned}\left(\prod_{i=1}^{n} a_i\right)\left(\prod_{j=n+1}^{n+s} a_j\right) &= \left(\prod_{i=1}^{n} a_i\right) \cdot \left[\left(\prod_{j=n+1}^{n+s-1} a_j\right) \cdot a_{n+s}\right] \\ &= \left[\left(\prod_{i=1}^{n} a_i\right) \cdot \left(\prod_{j=n+1}^{n+s-1} a_j\right)\right] a_{n+s}\end{aligned}$$

$$= \left(\prod_{i=1}^{n+s-1} a_i\right) a_{n+s} = \prod_{i=1}^{n+s} a_i。$$

故命题得证。

定义 12.2 a 为半群$(A,\cdot)$的一个元素,$n\in \mathbf{Z}^+$,则称$\underbrace{a,a\cdot\cdots\cdot a}_{n个}$为 a 的 n 次幂,记成 a^n。

$$a\in A, n,s\in \mathbf{Z}^+,$$

有
$$a^n\cdot a^s = a^{n+s};(a^n)^s = a^{n\cdot s}。$$

若$(A,\cdot)$是可换半群,即$\forall a,b\in A$ 有 $a\cdot b = b\cdot a$;这时,常用"+"代替"·",记 $a\cdot b$ 为 $a+b$。并用 na 表示 $\underbrace{a+\cdots+a}_{n个}$。此时,有

$$na + sa = (n+s)a,$$
$$n(sa) = (ns)a,$$
$$n(a+b) = na + nb。$$

这里,$a,b\in A$;$n,s\in \mathbf{Z}^+$。

3. 单位元

设$(A,\cdot)$为一半群。如果存在元素 $e\in A$,对$\forall a\in A$ 有 $ea = a$,则称 e 为 A 的一个左单位元。如果存在元素 $u\in A$,对$\forall a\in A$ 有 $au = a$,则称 u 为 A 的一个右单位元。若 A 的一个元素既是左单位元,又是右单位元,则称其为 A 的单位元。

一个半群可以没有单位元,也可以没有左(右)单位元。但若一个半群既有左单位元,又有右单位元,则必有单位元。即

命题 12.1 若半群$(A,\cdot)$有左单位元 e,右单位元 u,则 $e=u$,且 e 是 A 的惟一单位元。

证 因 e 是左单位元,故 $eu = u$,u 是右单位元,故 $u = e$,即 $u = e$ 是 A 的单位元。

若 A 有两个单位元 e,e',则 $e\cdot e' = e = e'$,故 A 的单位元若存在必惟一。

例 12.5 半群$(2^A,\cup)$有单位元ϕ;半群$(2^A,\cap)$有单位元 A。

例 12.6 半群$(\mathbf{Z},+)$有单位元 0；半群$(\mathbf{Z},\times)$有单位元 1。但半群$(\mathbf{Z}^+,+)$无单位元(因 $0\overline{\in}\mathbf{Z}^+$)。

例 12.7 A 为一非空集合，$\forall a,b\in A$，定义 $a\cdot b=b$，则$(A,\cdot)$为一半群。A 中任意元都是 A 的左单位元。当 A 含的元素多于 1 个时，A 无右单位元(因为若有右单位元 e，则右单位元等于左单位元，所有元素均等于同一个右单位元 e)。

4. 逆元

若$(A,\cdot)$是含单位元 e 的半群，A 中元素 a 称为左可逆的，如果存在 $b\in A$ 使 $ba=e$，b 为 a 的左逆元。a 称为右可逆的，如果存在$c\in A$ 使 $ac=e$，c 为 a 的右逆元。若 a 既是左可逆的又是右可逆的，则称 a 可逆。

在 a 可逆的情形下，a 的左逆 b 与右逆 c 必相等。因为 $b=be=b(ac)=(ba)c=ec=c$，$b=c$ 称为 a 的逆元，记为 a^{-1}。同理可证，a^{-1} 是惟一的。

若 a 可逆，则 $a^{-1}\cdot a=a\cdot a^{-1}=e$，于是 a^{-1}的逆为 a，即

$$(a^{-1})^{-1}=a。$$

又若 a,b 皆可逆，因为

$$(ab)(b^{-1}a^{-1})=e,(b^{-1}a^{-1})(ab)=e,$$

故 ab 可逆，其逆为 $b^{-1}a^{-1}$，即

$$(ab)^{-1}=b^{-1}a^{-1}。$$

5. 乘法表

为研究有限半群，通常把半群中元素的所有积按一定规则列成表格，称这样的表格为半群的乘法表。一般，若半群$(A,\cdot)$含 n 个元素，$A=\{a_1,\cdots,a_n\}$. 则作出 n 行 n 列矩阵，这矩阵第 i 行 j 列元为 $a_i\cdot a_j$，即

	a_1	a_2	$\cdots$	a_j	$\cdots$	a_n
a_1	$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$
a_2						
$\vdots$						
a_i	$\cdots$	$\cdots$	$\cdots$	a_ia_j	$\cdots$	$\cdots$
$\vdots$						
a_n	$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$

二、群

1. 群的定义

定义 12.3 半群$(G,\cdot)$称为群，如果G有左单位元e，且对G中任意元a都存在$b\in G$使$ba=e$。

在运算确切的情形下，为方便计，常把群$(G,\cdot)$简记成G。

例 12.8 半群$(Q,+)$是群，0是单位元。$\forall a\in Q$，有$(-a)\in Q$，且$a+(-a)=0$。

例 12.9 记$Q_0=Q-\{0\}$，则$(Q_0,\times)$是群，单位元为1。$\forall\ \dfrac{b}{a}\in Q_0$，有$\dfrac{a}{b}\in Q_0$，且$\dfrac{b}{a}\times\dfrac{a}{b}=1$。

例 12.10 记$Q_1=Q-\{-1\}$，则$(Q_1,\circ)$是群，这里

$$a\circ b=a+b+ab,$$

单位元是0：

$$0\circ b=0+b+0\cdot b=b,$$
$$0\circ 0=0,$$

又对任意$\dfrac{q}{p}\in Q_1$，有

$$\frac{q}{p}\circ\left(\frac{-q}{p+q}\right)=\frac{q}{p}-\frac{q}{p+q}-\frac{q^2}{p(p+q)}=0,$$

即$\dfrac{q}{p}$的逆元为$\dfrac{-q}{p+q}\left(\dfrac{q}{p}\neq-1,p+q\neq0\right)$。

例 12.11 设 A 为一非空集合，A 到 A 的双射全体所成的集合关于映射的乘法构成群。单位元是恒等映射 I_A，双射 f 为可逆映射，有逆映射 f^{-1}：

$$f^{-1} \cdot f = I_A。$$

2. 性质

若$(G, \cdot)$为一群，则有以下性质：

(1) G 的左单位元 e 是 G 的单位元，且是惟一的单位元；

(2) $\forall a \in G$，a 可逆，且 a 的左逆 b 就是它的逆元；

(3) $a, b, x \in G$，若 $ax = bx$，则 $a = b$，即满足消去律（同样，若 $ya = yb$，则 $a = b$）；

(4) $\forall a, b \in G$，必存在 $x, y \in G$ 满足 $ax = b$ 和 $ya = b$。

性质(1)～(4)有的已给了推导（如惟一性），有的显然（如性质(3)和(4)可直接由性质(2)得出）。需要证明的是：a 的左逆 b 也是 a 的右逆；左单位元 e 也是右单位元。其证明如下：

按群的定义，对 b，存在 $c \in G$ 使 $cb = e$。此时

$$ab = e(ab) = cbab = c(ba)b = c(eb) = cb = e,$$

又

$$ae = a(ba) = (ab)a = ea = a,$$

故 e 也是右单位元，从而是单位元。而

$$ab = ba = e,$$

即 a 可逆，b 是 a 的逆元。

3. 判别

定理 12.1 设$(G, \cdot)$是一半群，若$\forall a, b \in G$，方程 $ax = b$ 和 $ya = b$ 在 G 中都有解，则$(G, \cdot)$是群。

证 任取 $a \in G$，记 $xa = a$ 的解为 e_a，$\forall b \in G$，记 $ay = b$ 的解为 c，则 $e_a b = e_a ac = ac = b$，即 G 有左单位元 e_a，记成 e。

$\forall d \in G$，$xd = e$ 有解，所以 G 是群。

定理 12.2 若$(G,\cdot)$是有限半群，且满足左、右消去律，则$(G,\cdot)$是群。

证 设$G=\{a_1,a_2,\cdots,a_n\}$。$\forall a\in G$，由消去律知，$aa_1,aa_2,\cdots,aa_n$两两不等，故$G=aG=\{aa_1,\cdots,aa_n\}$。$\forall b\in G=aG$，即存在$a_i\in G$使$a_i$满足方程$ax=b$。同理，方程$ya=b$在$G$中有解。由定理12.1，即得本定理的证明。

三、子群

1. 定义

定义 12.4 群$(G,\cdot)$的非空子集H为$(G,\cdot)$的子群，如果$(H,\cdot)$仍是群，记成$H\leqslant G$。若$H\leqslant G$且$H\neq G$，则记成$H<G$。

例 12.12 $\mathbf{Z},\mathbf{Q},\mathbf{R},\mathbf{C}$关于数的加法运算都是群，前者是后者的子群。

例 12.13 设n为一固定正整数，则$H=\{nk\mid k\in\mathbf{Z}\}$是群$(\mathbf{Z},+)$的子群。特别，偶数全体$2\mathbf{Z}=\{2k\mid k\in\mathbf{Z}\}$是$(\mathbf{Z},+)$的子群。

证 $H\subseteq\mathbf{Z}$，加法满足结合律，又

$$nk+nl=n(k+l)\in\mathbf{Z}.$$

加法单位元：

$$0=n\cdot 0\in\mathbf{Z},$$
$$n(-k)\in H,n(-k)+nk=0,$$

故

$$H\leqslant\mathbf{Z}。$$

例 11.14 $\{-1,1\}$与$\{1,w,w^2\}=\{x\in\mathbf{C}\mid x^3=1\}$都是非零复数乘法群的子群。

每个群都有2个平凡的子群：单位元群和自身。

2. 判别

定理 12.3 群$(G,\cdot)$的非空子集H若满足：① $\forall a,b\in H$，都有$a\cdot b\in H$；② $\forall a\in H,a^{-1}\in H$，则$H\leqslant G$。

证 由①H 对"$\cdot$"封闭,因 $H \subseteq G$, G 中"$\cdot$"运算有结合律,故 $(H, \cdot)$是半群。$H \neq \varphi$,存在 $a \in H$。由②,$a^{-1} \in H$,得 $e = aa^{-1} \in H$,故 H 是群,因而 $H \leqslant G$。

定理 12.4 若 H 是群$(G, \cdot)$的非空子集,且$\forall a, b \in H$,则 $a^{-1}b \in H$,那么 $H \leqslant G$。

证 由于 $H \neq \varphi$,有 $a \in H$。又 $a, a \in H$,故 $e = a^{-1}a \in H$。$a, e \in H$,则 $a^{-1} = a^{-1}e \in H$。若 $a, b \in H$,则 $a^{-1}, b \in H$,故 $ab = (a^{-1})^{-1}b \in H$,由定理 12.3 得 $H \leqslant G$。

定理 12.5 若 H 是群$(G, \cdot)$的非空有限子集,$\forall a, b \in H$,都有 $a \cdot b \in H$,则 $H \leqslant G$。

证 由定理 12.3 的证明可知$(H, \cdot)$是有限半群,$H \subseteq G$, G 中"$\cdot$"满足消去律,故$(H, \cdot)$是群。所以 $H \leqslant G$。

例 12.15 设$(G, \cdot)$为一交换群,n 为一固定正整数,记

$$H = \{x \in G \mid x^n = e\},$$

则

$$H \leqslant G。$$

证 ① 因 $e^n = e$,故 $e \in H$, $H \neq \varphi$;

② 因 $a, b \in H$, $a^n = b^n = e$,故$(ab)^n = a^n b^n = e$,因而 $ab \in H$;

③ 因 $a \in H$, $a^n = e$, $(aa^{-1})^n = e^n = e$,则 $a^n(a^{-1})^n = e$,即 $(a^{-1})^n = e$,所以 $a^{-1} \in H$。

由①,②,③得 $H \leqslant G$。

由例 12.15 可知,$H = \{x \in \mathbf{C} \mid x^n = 1\}$ 是非 0 复数乘法群的子群。若用定理 12.5 可得更简易的证明:

因为 $1 \in H$, $|H| \leqslant n$,故 H 是 $\mathbf{C} - \{0\}$ 的非空有限子集。任意 $a, b \in H$, $a^n = b^n = 1$,于是$(ab)^n = a^n b^n = 1 \cdot 1 = 1$,即 $ab \in H$,故

$$H \leqslant \mathbf{C} - \{0\}。$$

例 12.16 设$(G, \cdot)$为一交换群。记 $H = \{x \mid x \in G$,存在 $n \in \mathbf{Z}^+$ 使 $x^n = e\}$,则 $H \leqslant G$。

证 ① 因 $e^1=e$，故 $e\in H$；

② $a,b\in H$，存在 $n,s\in P$ 使 $a^n=b^s=e$，此时 $(ab)^{n\cdot s}=(a^n)^s\cdot(b^s)^n=e$，故 $a\cdot b\in H$；

③ 同例 12.15 的③，故

$$H\leqslant G。$$

例 12.17 设 $(G,\cdot)$ 为一群，$a\in G$，记 $N_a=\{x\mid x\in G,ax=xa\}$，则 $N_a\leqslant G$。

证 $e\in N_a$，故 $N_a\neq\varphi$。若 $d,b\in N_a$，则

$$(db)a=d(ba)=d(ab)=(da)b=a(db),$$

即 $db\in N_a$。若 $b\in N_a,ba=ab$，故 $b^{-1}(ba)b^{-1}=b^{-1}(ab)b^{-1}$，因而 $ab^{-1}=b^{-1}a$，故 $b^{-1}\in N_a$。所以 $N_a\leqslant G$。

3. 子群的运算

定理 12.6 群的任意个子群的交仍是子群。即若 $H_i(i\in I)$ 是群 G 的子群，则 $\bigcap\limits_{i\in I}H_i\leqslant G$。

证 ① $e\in H_i,i\in I$，故 $e\in\bigcap\limits_{i\in I}H_i\neq\varphi$；

② $a,b\in\bigcap\limits_{i\in I}H_i$，则 $a,b\in H_i,i\in I$。因为 $H_i\leqslant G$，所以 $a^{-1}b\in H_i(i\in I)$，从而 $a^{-1}b\in\bigcap\limits_{i\in I}H_i$。

由上可知，$\bigcap\limits_{i\in I}H_i\leqslant G$。

子群的并一般不是子群，但有：

定理 12.7 若 $H_i(i\in I)$ 是群 G 的子群，$\forall i,j\in I$，必 $H_i\leqslant H_j$，或者 $H_j\leqslant H_i$，那么 $\bigcup\limits_{i\in I}H_i\leqslant G$。

证 ① $e\in\bigcup\limits_{i\in I}H_i\neq\varphi$；

② $\forall a,b\in\bigcup\limits_{i\in I}H_i$，存在 $k,j\in I$ 使 $a\in H_k,b\in H_j$，可设 $H_k\leqslant H_j$，则 $a,b\in H_j$，因而 $a^{-1}b\in H_j\subseteq\bigcup\limits_{i\in I}H_i$；

由上可得 $\bigcup\limits_{i\in I}H_i\leqslant G$。

4. 子集生成的子群

定义 12.5 设 S 为群 G 的一个子集，称 $\bigcap\limits_{\substack{H \supseteq S \\ H \leqslant G}} H$ 为 S 生成的子群，记成(S)。(S)有下面性质：

① $(S) \leqslant G$；

② $(S) \supseteq S$；

③ 若 $H \supseteq S$，且 $H \leqslant G$，则$(S) \leqslant H$。

例 12.18 在群$(\mathbf{Z},+)$中，令 $S_1 = \{2,3\}$，$S_2 = \{4,6\}$，求证：

(1) $(S_1) = \mathbf{Z}$；

(2) $(S_2) = \{2k \mid k \in \mathbf{Z}\} = (\{2\})$。

证 (1) 因为 $2,3 \in S_1$，(S_1)是群，所以$\forall n \in \mathbf{Z}^+$：

$$n = 3n - 2n = \underbrace{3 + \cdots + 3}_{n\text{个}} - \underbrace{2 - 2 - \cdots - 2}_{n\text{个}} \in (S_1),$$

$$-n = 2n - 3n = \underbrace{2 + 2 + \cdots + 2}_{n\text{个}} - \underbrace{3 - 3 - \cdots - 3}_{n\text{个}} \in (S_1),$$

故

$$(S_1) \supseteq \mathbf{Z},$$

从而

$$(S_1) = \mathbf{Z}。$$

(2) 因为 $4,6 \in S_2$，(S_2)是群，任意 $k \in \mathbf{Z}^+$：

$$2k = 6k - 4k = \underbrace{6 + \cdots + 6}_{k\text{个}} - \underbrace{4 - \cdots - 4}_{k\text{个}} \in (S_2),$$

$$-2k = 4k - 6k = \underbrace{4 + \cdots + 4}_{k\text{个}} - \underbrace{6 - \cdots - 6}_{k\text{个}} \in (S_2),$$

故

$$(S_2) \supseteq \{2k \mid k \in \mathbf{Z}\}。$$

由例 12.13

$$2\mathbf{Z} = \{2k \mid k \in \mathbf{Z}\} \leqslant (\mathbf{Z}, +),$$

且

$$4 = 2 \cdot 2 \in 2\mathbf{Z},$$
$$6 = 2 \cdot 3 \in 2\mathbf{Z},$$

由子集生成的子群的性质③得

$$2\mathbf{Z} \supseteq (S_2),$$

所以

$$(S_2) = \{2k \mid k \in \mathbf{Z}\}。$$

四、群元素的阶

定义 12.6 记 a 为群 G 的一个元素，记

$$T_a = \{n \in \mathbf{Z}^+ \mid a^n = e\},$$

若

(1) $T_a = \phi$，此时，对任意 $n \in \mathbf{Z}^+$，必 $a^n \neq e$，称元素 a 的阶为无穷大，记成 $|a| = +\infty$；

(2) $T_a \neq \phi$，T_a 为正整数集 $\mathbf{Z}^+$ 的非空子集，有最小元 k，即 $a^k = e$，且对任意 $n \in \mathbf{Z}^+$，$n < k$，必 $a^n \neq e$。称 k 为元素 a 的阶，记成 $|a| = k$。

若 $|a| = k$，$n \in \mathbf{Z}$，$k \mid n$，即 $n = kt$，则 $a^n = (a^k)^t = e^t = e$。

例如，非零复数乘法群中，元素(-1)的阶为 2，ω 的阶为 3，i 的阶为 4；而 2 的阶为无穷大。

整数加法群中，任意非零元的阶都为无穷大。

定理 12.8 设 a 为群 G 的一个元素，则 $|a| = k \in \mathbf{Z}^+$ 的充要条件是：① $a^k = e$；② $\forall d \in \mathbf{Z}$，$a^d = e$，必 $k \mid d$。

证 必要性。按阶定义：$a^k = e$。若 $d \in \mathbf{Z}$，$a^d = e$。设 d 被 k 去除余数为 r，即 $d = qk + r$，$0 \leqslant r < k$。此时，$e = a^d = a^{qk+r} = a^r$。按阶定义，必 $r = 0$，故 $k \mid d$。

充分性：略。

例 12.19 若群 G 的元素个数是偶数，则 G 中有阶为 2 的元素。

证 由于 $G=\bigcup_{a\in G}\{a,a^{-1}\}$ 是 G 的一个分类，则 $H=\{a|a\in G, a^{-1}=a\}$ 的元素个数 $\neq 1$。否则 G 含奇数个元素。故存在 $a\in H, a\neq e, a^{-1}=a, a^2=e, |a|=2$。

例 12.20 若群 G 中元素的阶都是 2 的因子，则 G 是交换群。

证 $\forall a,b\in G$，则 $|a|,|b|,|ab|$ 都是 2 的因子。故 $(ab)^2=(ab)(ab)=e=(a^2)(b^2)$，即 $abab=aabb$，由消去律得 $ba=ab$，所以 G 是交换群。

定理 12.9 若群 G 的元素 a 的阶为 $n\in\mathbf{Z}^+$，则 a^k 的阶为 $\dfrac{[n,k]}{k}=\dfrac{n}{(n,k)}$。这里 $k\in\mathbf{Z}^+$。

证 ① 因 $n|[n,k]$，所以

$$(a^k)^{\frac{[n,k]}{k}}=(a)^{[n,k]}=e;$$

② $\forall d\in\mathbf{Z}$，$(a^k)^d=e, a^{kd}=e$，故 $n|kd$。又 $k|kd$，因而 $[n,k]|kd$，所以 $[n,k]/k|d$。

根据定理 11.8，由①，②可知 a^k 的阶为 $\dfrac{[n,k]}{k}$。

§2 循环群与置换群

一、循环群

1. 群元素的幂

若 a 为群 G 的一个元素，对 $k\in\mathbf{Z}$ 定义 a 的幂 a^k 如下：

$a^k=\underbrace{a\cdots a}_{k个}$，若 $k\in\mathbf{Z}^+$，

$a^0=e$，

$a^k=(a^{-1})^n$，若 $k=-n, n\in\mathbf{Z}^+$，显然有

$$(a)=(\{a\})=\{a^0,a,a^{-1},a^2,a^{-2},\cdots\}。$$

2. 循环群的定义

定义 12.7 群 G 为循环群，如果它是由一个元素生成的，即 $G=(a)=\{a^k|k\in\mathbf{Z}\}$。

例 12.21 $(\mathbf{Z},+)$是一循环群，因为 $\mathbf{Z}=(1)$。它显然是无限群。

例 12.22 用 U_n 表示复数 $\mathbf{C}$ 中 n 个 n 次单位根所成的集合，即 $U_n=\left\{e^{\frac{2k\pi i}{n}}|k=1,2,\cdots,n\right\}$，这里 $n\in\mathbf{Z}^+$。则 U_n 关于数的乘法成一群，且 $U_n=\left(e^{\frac{2\pi i}{n}}\right)$。$e^{\frac{2\pi i}{n}}$ 的阶为 n，U_n 为 n 阶循环群。

3. 循环群的类型

下面指出，循环群仅有两种类型。

定义 12.8 若 G 为循环群，即 $G=(a)$。如果 $|a|=+\infty$，则称 G 为无限循环群。如果 $|a|=n\in\mathbf{Z}^+$，则称 G 为 n 阶循环群。

定理 12.10 无限循环群(a)含无限个元素。当 $i,j\in\mathbf{Z}$，$i\neq j$ 时，$a^i\neq a^j$。

证 若有 $i<j$，使 $a^i=a^j$，则 $a^{j-i}=e$，$j-i\in\mathbf{Z}^+$，和 a 的阶是 $+\infty$ 矛盾。故定理得证。

定理 12.11 n 阶循环群(a)有 n 个元素 $a^0,a^1,\cdots,a^{n-1}$，即

$$|(a)|=|a|。$$

证 $0\leqslant i<j<n$ 时，若 $a^i=a^j\Rightarrow a^{j-i}=e$，$0<j-i<n\in\mathbf{Z}^+$，和 a 的阶为 n 矛盾，故 $a^i\neq a^j$，即 $a^0,a^1,\cdots,a^{n-1}$ 两两不同。又$\forall k\in\mathbf{Z}$，$k=qn+r$，$0\leqslant r<n$，则 $a^k=a^r\in\{a^0,a^1,\cdots,a^{n-1}\}$。故得定理之证明。

推论 (a)为 n 阶循环群，$k\in\mathbf{Z}$，$a^k=e\Leftrightarrow n|k$。

4. 循环群的子群

设 $H\leqslant(a)$。若 $H=\{e\}$，则 $H=(e)$；若 $H\neq\{e\}$，则存在 $k\in\mathbf{Z}$，$k\neq0$，$a^k\in H$。因 $a^k\in H$，必 $a^{-k}\in H$，故集合 $T=\{k|k\in\mathbf{Z}^+,a^k\in H\}$ 非空，记 d 为 T 的最小元。$a^d\in H$，故$(a^d)\subseteq H$；又$\forall a^i\in H$，$i=qd+r$，$0\leqslant r<d$，则 $a^r=a^i(a^d)^{-q}\in H$。因 d 是 T 的最小元，故 $r=0$，所以 $a^i\in(a^d)$。即 $H=(a^d)$。故有

2. 循环群的定义

定义 12.7 群 G 为循环群，如果它是由一个元素生成的，即 $G=(a)=\{a^k|k\in\mathbf{Z}\}$。

例 12.21 $(\mathbf{Z},+)$是一循环群，因为 $\mathbf{Z}=(1)$。它显然是无限群。

例 12.22 用 U_n 表示复数 $\mathbf{C}$ 中 n 个 n 次单位根所成的集合，即 $U_n=\left\{e^{\frac{2k\pi i}{n}}\,|\,k=1,2,\cdots,n\right\}$，这里 $n\in\mathbf{Z}^+$。则 U_n 关于数的乘法成一群，且 $U_n=\left(e^{\frac{2\pi i}{n}}\right)$。$e^{\frac{2\pi i}{n}}$ 的阶为 n，U_n 为 n 阶循环群。

3. 循环群的类型

下面指出，循环群仅有两种类型。

定义 12.8 若 G 为循环群，即 $G=(a)$。如果 $|a|=+\infty$，则称 G 为无限循环群。如果 $|a|=n\in\mathbf{Z}^+$，则称 G 为 n 阶循环群。

定理 12.10 无限循环群(a)含无限个元素。当 $i,j\in\mathbf{Z}$，$i\neq j$ 时，$a^i\neq a^j$。

证 若有 $i<j$，使 $a^i=a^j$，则 $a^{j-i}=e$，$j-i\in\mathbf{Z}^+$，和 a 的阶是 $+\infty$ 矛盾。故定理得证。

定理 12.11 n 阶循环群(a)有 n 个元素 $a^0,a^1,\cdots,a^{n-1}$，即

$$|(a)|=|a|。$$

证 $0\leqslant i<j<n$ 时，若 $a^i=a^j\Rightarrow a^{j-i}=e$，$0<j-i<n\in\mathbf{Z}^+$，和 a 的阶为 n 矛盾，故 $a^i\neq a^j$，即 $a^0,a^1,\cdots,a^{n-1}$ 两两不同。又$\forall k\in\mathbf{Z}$，$k=qn+r$，$0\leqslant r<n$，则 $a^k=a^r\in\{a^0,a^1,\cdots,a^{n-1}\}$。故得定理之证明。

推论 (a)为 n 阶循环群，$k\in\mathbf{Z}$，$a^k=e\Leftrightarrow n|k$。

4. 循环群的子群

设 $H\leqslant(a)$。若 $H=\{e\}$，则 $H=(e)$；若 $H\neq\{e\}$，则存在 $k\in\mathbf{Z}$，$k\neq0$，$a^k\in H$。因 $a^k\in H$，必 $a^{-k}\in H$，故集合 $T=\{k|k\in\mathbf{Z}^+,a^k\in H\}$ 非空，记 d 为 T 的最小元。$a^d\in H$，故$(a^d)\subseteq H$；又$\forall a^i\in H$，$i=qd+r$，$0\leqslant r<d$，则 $a^r=a^i(a^d)^{-q}\in H$。因 d 是 T 的最小元，故 $r=0$，所以 $a^i\in(a^d)$。即 $H=(a^d)$。故有

$\{1,2,\cdots,n\}$，记 $S_n = \{\sigma|\sigma$ 是 A 的置换$\}$。

$\sigma\in S_n$，可把 σ 表成

$$\sigma=\begin{pmatrix}1 & 2 & 3 & \cdots & n\\ \sigma(1) & \sigma(2) & \sigma(3) & \cdots & \sigma(n)\end{pmatrix},$$

这里 $\sigma(i)$是 i 在 σ 映射下的像。由于 σ 是双射，所以 $\sigma(1),\sigma(2),\cdots,\sigma(n)$是 $1,2,\cdots,n$ 的一个排列。从而 $|S_n|=n!$。$\sigma\cdot\tau$ 即通常映射的积。如 $\sigma=\begin{pmatrix}1&2&3\\2&1&3\end{pmatrix}$，$\tau=\begin{pmatrix}1&2&3\\2&3&1\end{pmatrix}$。则 $\sigma\tau=\begin{pmatrix}1&2&3\\1&3&2\end{pmatrix}$。

2. 循环及其性质

设 $A=\{1,2,\cdots,n\}$，$\sigma\in S_n$，若 $\sigma(i_1)=i_2$，$\sigma(i_2)=i_3$，$\cdots$，$\sigma(i_{k-1})=\sigma(i_k)$，$\sigma(i_k)=i_1$，而 σ 把 A 中其他字母保持不变，则称 σ 为 S_n 的一个 k 项循环，记为$(i_1\,i_2\cdots\,i_k)$（这里 $i_1,i_2,\cdots,i_k$ 为 $1,\cdots,n$ 中 k 个不同的数）。2 项循环常称为对换。

若 $\{i_1\cdots i_k\}\cap\{j_1\cdots j_s\}=\varphi$，则称循环$(i_1\cdots i_k)$ 和$(j_1\cdots j_s)$ 是不相交的。

循环有如下的性质：

(1) 若$(i_1\cdots i_k)$和$(j_1\cdots j_s)$是不相交循环，则
$(i_1\cdots i_k)\cdot(j_1\cdots j_s)=(j_1\cdots j_s)\cdot(i_1\cdots i_k)$；

(2) $(i_1\,i_2\,i_3\,\cdots\,i_k)=(i_2\,i_3\,\cdots\,i_k\,i_1)=\cdots$
$=(i_k\,i_1\,i_2\,\cdots\,i_{k-1})$；

(3) $(i_1\,i_2\,\cdots\,i_{k-1}\,i_k)^{-1}=(i_k\,i_{k-1}\,\cdots\,i_2\,i_1)$；

(4) $(i_1\,i_2\cdots\,i_k)^k=$恒等映射 e，且 k 是具有这种性质的最小正整数；

(5) $(i_1\,i_2\,\cdots\,i_k)=(i_1i_k)\cdots(i_1i_4)(i_1i_3)(i_1i_2)$；

(6) $(ij)(i\cdots j\,\cdots)=(i\,\cdots)\,(j\,\cdots)$，
$(ij)(i\,\cdots)(j\,\cdots)=(i\,\cdots j\cdots)$；

(7) 若 $\sigma\in S_n$，则

$$\sigma(i_1\,\cdots\,i_k)\sigma^{-1}=(\sigma(i_1)\,\cdots\,\sigma(i_k))。$$

这些性质的验证留给读者作为练习。

3. 置换和循环

$\sigma \in S_n$,在 $A = \{1,2,\cdots,n\}$ 中定义关系 R: $aRb \Leftrightarrow$ 存在 $k \in \mathbf{Z}$ 使 $b = \sigma^k(a)$。定义 σ^0 为恒等映射,易证 R 是 A 的一个等价关系,从而得到 A 的一个分类 $A = \bar{a} \cup \bar{b} \cup \cdots$。

任取一类 $\bar{a} = \{\sigma^k(a) \mid k \in \mathbf{Z}\}$,因为 A 仅有有限个数,故存在 $i < j$ 使 $\sigma^i(a) = \sigma^j(a)$,即 $\sigma^{j-i}(a) = a, j - i \in \mathbf{Z}^+$,记 k 为使 $\sigma^k(a) = a$ 最小正整数,则 $\sigma(a), \sigma^2(a), \cdots, \sigma^k(a) = a$ 两两不同。$\forall l \in \mathbf{Z}$, $l = qk + r$, $0 \leqslant r < k$,则 $\sigma^l(a) = \sigma^r(a)$。故 $\bar{a} = \{a\ \sigma(a) \cdots \sigma^{k-1}(a)\}$。在 $\bar{a}$ 上的作用和循环 $(a\ \sigma(a) \cdots \sigma^{k-1}(a))$ 完全相同。由循环的性质有:

$$\sigma = (a\ \sigma(a)\ \cdots\ \sigma^{k-1}(a))\ (b\ \sigma(b) \cdots) \cdots。$$

即有下述定理:

定理 12.17 置换 σ 可表成不相交循环之积: $\sigma = (i_1\ \cdots\ i_k)(j_1 \cdots j_l) \cdots$,且 $A = \{i_1 \cdots j_k\} \cup \{j_1 \cdots j_l\} \cup \cdots$ 是 A 的一个分类。

如果不考虑循环次序,则这种表法是惟一的,即若 $\sigma \in S_n$,能按定理表成循环 $\tau_1, \cdots, \tau_s$ 之积,又能表成循环 $\eta_1, \cdots, \eta_u$ 之积,则必 $s = u$,且 $\tau_1 = \eta_{i_1}, \cdots, \tau_s = \eta_{i_s}$,这里 $i_1, \cdots, i_s$ 是 $1, \cdots, u$ 的一个排列。其证如下:

用 $\bar{\tau}_i$ 表示循环 τ_i 中字母组成的集合,则 $A = \bigcup\limits_i \bar{\tau}_i$ 是 A 的一个分类,$A = \bigcup\limits_i \bar{\eta}_i$ 也是 A 的一个分类。$\forall x \in A$, $x \in$ 某一 $\bar{\tau}_i$, $x \in$ 某 $\bar{\eta}_j$ 而得 $\bar{\tau}_i = \bar{\eta}_j = \{\sigma^k(x) \mid k \in \mathbf{Z}\}$,故 $\tau_i = \eta_j$,依次进行,即得所证。

例如,$\begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \\ 3 & 6 & 5 & 7 & 1 & 4 & 2 & 8 \end{pmatrix} = (1\ 3\ 5)\ (2\ 6\ 4\ 7)\ (8)$。

4. 奇置换与偶置换

定义 12.9 $\sigma \in S_n$, σ 可惟一表成不相交循环之积 $\sigma = \tau_1 \cdots \tau_s$,则称 $n - s$ 为 σ 的逆序数,记成 $\mu(\sigma)$。

例如,置换 $(135)\ (2647)\ (8)$ 的逆序数为 $8 - 3 = 5$,恒等映射 e 的逆序数为 $n - n = 0$,对换 (ij) 的逆序数为 $n - (n - 1) = 1$,这是由于恒等映射保持所有字母不变,因此它是 n 个不相交的 1 项循环之积,而对换保持 $n - 2$ 个字母不变,所以它是 $n - 1 = (n - 2) + 1$ 个不相交

循环之积。

定义 12.10 设$\sigma\in S_n$,若$\mu(\sigma)$是偶(奇)数,则称σ是偶(奇)置换。

引理 12.1 对换改变置换的奇偶性,即对$\sigma\in S_n$及$1\leqslant i,j\leqslant n$,$i\neq j$,都有$\mu[(ij)\sigma]\equiv\mu(\sigma)+1\pmod 2$。

证 表σ为不相交循环之积,有两种可能:

① i,j出现在同一循环中,即$\sigma=(i\cdots j\cdots)\cdots$;

② i,j出现在不同循环中,即$\sigma=(i\cdots)(j\cdots)\cdots$;

① 的情形:$(ij)\sigma=(i\cdots)(j\cdots)\cdots$仍为不相交循环之积,循环个数比$\sigma$的多1个;

② 的情形:$(ij)\sigma=(i\cdots j\cdots)\cdots$仍是不相交循环之积,循环个数比$\sigma$的少一个。

不论①,②都有:$\mu[(ij)\sigma]$比$\mu(\sigma)$大1或小1。即$\mu[(ij)\sigma]-\mu(\sigma)\equiv1\pmod 2$。

定理 12.18 偶(奇)置换能且仅能表成偶(奇)数个对换之积。

证 由循环的性质(5)和定理12.17,知置换能表成对换之积。设$\sigma\in S_n$,$\sigma=\underbrace{(pq)\cdots(ij)}_{x\text{个}}$。今证$\mu(\sigma)$和$x$有相同的奇偶性。

$\underbrace{(ij)\cdots(pq)}_{x\text{个}}\sigma=(ij)\cdots(pq)(pq)\cdots(ij)=e$,即$\sigma$乘上$x$个对换成了恒等映射,每乘一个对换逆序数模2加1,而恒等映射的逆序数为0,故有

$$\mu(\sigma)+x\equiv0\pmod 2,$$

即$\mu(\sigma)$和x有相同奇偶性,故得定理的证明。

记A_n为S_n中偶置换全体,即$A_n=\{\sigma|\sigma\in S_n,\mu(\sigma)\text{是偶数}\}$,则$S_n-A_n$为奇置换全体。

因为$f(\sigma)=(12)\sigma$既是A_n到S_n-A_n的单射,也是S_n-A_n到A_n的单射。故

$$|A_n|=|S_n-A_n|=|S_n|/2=\frac{1}{2}\times n!(n\geqslant2)。$$

5. 对称群与交错群

S_n 关于映射的乘法运算成一群，称为 n 次对称群。群 S_n 的单位元 e 是恒等映射。因为 $e \in A_n$，所以 A_n 是 S_n 的非空有限子集。又两个偶置换的积仍是偶数个对换之积，即仍是偶置换，根据定理 11.5，$A_n \leqslant S_n$，称 A_n 为 n 次交错群。

对称群 S_n 的任一子群，包括 S_n 与 A_n，都称为置换群。

例 12.23 试用不相交循环之积表示 S_3 与 A_4 中的所有置换。

解 S_3 中：

(1) 若置换 σ 是 3 个不交循环之积：

$$\sigma = (i_1)(i_2)(i_3) = (1)(2)(3),$$

则 a 为恒等映射 e。

(2) 若置换 σ 是两个不交循环之积，则

$$\sigma = (i_1)(i_2 i_3) = (i_2 i_3)(i_1) = (i_2 i_3),$$

此时

$$\sigma = (12), (13) \text{ 或} (23)。$$

(3) 置换 σ 是(一个)三项循环：

$$\sigma = (1\ i_1\ i_2),$$

此时

$$\sigma = (1\ 2\ 3) \text{ 或} (1\ 3\ 2),$$

从而

$$\begin{aligned} S_3 &= \{e, (1)(23), (2)(13), (3)(12), (123), (132)\} \\ &= \{e, (23), (13), (12), (123), (132)\}。\end{aligned}$$

A_4 中置换 σ 全是偶置换，因而或是四个不交循环之积或是两个不交循环之积，有下面三种情形：

(1) σ 是四个不交循环之积：

$$\sigma = (1)(2)(3)(4),$$

即 σ 为恒等映射。

(2) σ 是一个 1 项循环与一个 3 项循环之积：

$$\sigma = (i_1)(i_2 i_3 i_4),$$

此时

$$\sigma = (1)(234),(1)(243),(2)(134),(2)(143),$$
$$(3)(124),(3)(142),(4)(123) \text{ 或 } (4)(132)。$$

(3) σ 是两个 2 项循环之积：

$$\sigma = (1\ i_1)(i_2 i_3),$$

此时

$$\sigma = (12)(34),(13)(24) \text{ 或 } (14)(23),$$

从而

$$A_4 = \{e,(123),(132),(134),(143),(124),(142),(234),(243),$$
$$(12)(34),(13)(24),(14)(23)\}。$$

例 12.24 求证：

$$B_4 = \{e,(12)(34),(13)(24),(14)(23)\} \leqslant S_4。$$

证 若记 $a = (12)(34)$, $b = (13)(24)$, $c = (14)(23)$，直接验证 B_4 中元素乘法封闭(见乘法表)：

x	e	a	b	c
e	e	a	b	c
a	a	e	c	b
b	b	c	e	a
c	c	b	a	e

由定理 12.5

$$B_4 \leqslant S_4。$$

§3 陪集与指数

一、半群子集的运算

设$(A,\cdot)$是一个半群，在幂集合 2^A 中定义二元运算：

$$B\cdot C=\{b\cdot c\,|\,b\in B\subset A,c\in C\subset A\},$$

则$(2^A,\circ)$成半群，且有

$$(\bigcup_{i\in I}A_i)\cdot(\bigcup_{j\in J}B_j)=\bigcup_{i,j}(A_i\cdot B_j)。$$

这个结论可以证明如下：

若 $a\cdot b\in$ 左边，这里 $a\in\bigcup_i A_i,b\in\bigcup_j B_j$，于是，存在 $k\in I,t\in J$ 使

$$a\in A_k,b\in B_t,$$

从而

$$ab\in A_kB_t\subseteq 右边,$$

即

$$左边\subseteq右边。$$

反之，右边的元素可表为 ab，这里 $a\in A_{i_1}\subseteq\bigcup A_i,b\in B_{j_1}\subseteq\bigcup B_j$，于是 $ab\in$ 左边，即右边 $\subseteq$ 左边，故

$$左边=右边。$$

若 G 是一个群，则 2^G 有单位元$\{e\}$。如果 H 是 G 的子群，则

$$H=H\cdot\{e\}\subseteq H\cdot H=H^2=\{h_1\cdot h_2\,|\,h_1,h_2\in H\}\subseteq H,$$

于是

$$H^2=H,$$

对 $n(\in\mathbf{Z}^+)$用归纳法，即得

$$H^n = H。$$

当 B 仅含一个元素 b，即 $B = \{b\}$ 时，常记 BH 为 bH。

二、陪集

若 G 是一个群，$H \leqslant G$。利用 H，可以在 G 中定义一个等价关系，从而得到 G 的一个分类。

定义 12.11 若 H 是群 G 的子群。在 G 中，定义左陪集关系 R：$aRb \Leftrightarrow a \in b \cdot H$。

推论 群 G 关于子群 H 的左陪集关系是等价关系，且 $a \in G$，a 所在等价类为 aH，也称作 a 所在的左陪集。

证 (1) $a = ae$，$e \in H$，所以 $a \in aH$，故 aRa；

(2) 因 aRb，故 $a \in bH$，即存在 $h \in H$ 使 $a = bh$，$b = ah^{-1}$，$h^{-1} \in H$，即 $b \in aH$，故 bRa；

(3) aRb，bRc，即 $a \in bH$，$b \in cH$，于是存在 $h_1, h_2 \in H$ 使 $a = bh_1$，$b = ch_2$，$a = c(h_2h_1)$，$h_2h_1 \in H$，故 $a \in cH$，因而 aRc。

由(1)，(2)，(3)知，左陪集关系 R 是等价关系。

因为 $bRa \Leftrightarrow b \in aH$。故 a 所在等价类为 aH。

又映射 $\sigma(h) = ah$ 是集合 H 到 aH 的双射，故

$$H \sim aH。$$

类似地，定义群 G 关于子群 H 的右陪集关系 R：

$$aRb \Leftrightarrow a \in Hb,$$

右陪集关系也是等价关系，元素 a 所在的等价类为 Ha。同样，$H \sim Ha$。

陪集关系是等价关系，整个集合 G 是两两不交的等价类(陪集)的并：

$$G = \bigcup aH = \bigcup Ha。$$

因为陪集关系是等价关系，等价关系有对称性，从而有：

$$b \in aH \Leftrightarrow bH = aH \Leftrightarrow a \in bH,$$

$$b \in Ha \Leftrightarrow Hb = Ha \Leftrightarrow a \in Hb。$$

命题 12.2 $H \leqslant G$,则

$$aH = bH \Leftrightarrow a^{-1}b \in H,$$
$$Ha = Hb \Leftrightarrow ba^{-1} \in H。$$

证 若 $aH = bH$. 于是

$$be \in bH = aH,$$

即存在 $h \in H$,使

$$b = be = ah,$$

从而

$$a^{-1}b = h \in H。$$

反之,若 $a^{-1}b \in H$, 即存在 $h \in H$,使

$$a^{-1}b = h,$$

于是

$$b = ah \in aH,$$

故

$$aH = bH。$$

类似可得

$$Ha = Hb \Leftrightarrow ba^{-1} \in H。$$

例 给了 S_3 的子群 $H = \{e,(12)\}$,则

$$eH = H = (12)H,$$
$$(13)H = \{(13),(123)\} = (123)H,$$
$$(23)H = \{(23),(132)\} = (132)H,$$
$$S_3 = H \cup (13)H \cup (23)H。$$

又

$$He = H = H(12),$$
$$H(13) = \{(13),(132)\} = H(132),$$
$$H(23) = \{(23),(123)\} = H(123),$$
$$S_3 = H \cup H(13) \cup H(23).$$

命题 12.3 S_e 表群 G 关于子群 H 的所有左陪集所成集合，S_r 表群 G 关于子群 H 的所有右陪集所成集合，则 $S_e \sim S_r$。

证 因 $aH = bH \Leftrightarrow a^{-1}b \in H \Leftrightarrow a^{-1} \in Hb^{-1} \Leftrightarrow Ha^{-1} = Hb^{-1}$，故对应 $aH \xrightarrow{f} Ha^{-1}$ 是 S_e 到 S_r 的单射。又 $\forall a \in G$，存在 $a^{-1} \in G$，$f(a^{-1}H) = Ha$，即 f 是满射，从而 f 是 S_e 到 S_r 的双射。

定义 12.12 群 G 关于其子群 H 的左（右）陪集所成集合 S_l（$\sim S_r$）的势称作 H 在 G 中的指数，写成$[G : H]$（S_e 为有限集时，指数即陪集个数）。

三、陪集与指数

定理 12.19 若 G 是一个有限群，$K \leqslant H \leqslant G$。则

$$[G : K] = [G : H] \cdot [H : K]。$$

证 $G = \bigcup\limits_{i \in I_1} a_i H$ 是 G 的一个分类，$|I_1| = [G : H]$，$H = \bigcup\limits_{j \in I_2} b_j K$ 是 H 的一个分类，$|I_2| = [H : K]$，故 $G = \bigcup\limits_{i \in I_1} \bigcup\limits_{j \in I_2} a_i b_j K$ 是 G 的一个分类，即得

$$[G : K] = |I_1| \times |I_2| = [G : H] \cdot [H : K]。$$

推论 若 H 是有限群 G 的子群，则

(1) $|H| \mid |G|$；

(2) $\forall a \in G$，有 $|a| \mid |G|$，从而 $a^{|G|} = e$。

证 (1) 因为 $|G| = [G : H] \cdot |H|$。

(2) $|a| = |(a)| \mid |G|$。

定理 12.20 若 A,B 是群 G 的有限子群，则

$$|A \cdot B| = |A||B|/|A \cap B|。$$

证 记 $A \cap B = C$，设 C 在 A 中的指数为 n，$A = \bigcup_{i=1}^{n} a_i C$，$i \neq j$，$a_i C \neq a_j C$，即 $a_i^{-1} a_j \notin C$，从而 $a_i^{-1} a_j \notin B$。因为否则 $a_i^{-1} a_j \in A \cap B = C$，故 $a_i B$ 和 $a_j B$ 也是两个不交的 B 左陪集。又

$$AB = \left(\bigcup_{i=1}^{n} a_i C\right) B = \bigcup_{i=1}^{n} a_i CB = \bigcup_{i=1}^{n} a_i B,$$

故

$$|AB| = n \cdot |B| = \frac{|A| \cdot |B|}{|C|}.$$

即得定理证明。

定理 12.21 若 A 和 B 是群 G 的子群，记 $K = A \cap B$，则

$$[A : K] \leqslant [G : B],$$

且若 $[G : B]$ 有限，则

$$[A : K] = [G : B] \Leftrightarrow G = A \cdot B。$$

证 记 $S = \{aK \mid a \in A\}$，$T = \{xB \mid x \in G\}$。$\forall a_1 a_2 \in A$，$a_1 K = a_2 K \Leftrightarrow a_1^{-1} a_2 \in K \Leftrightarrow a_1^{-1} a_2 \in B \Leftrightarrow a_1 B = a_2 B$，所以

$$f(aK) = aB$$

是 S 到 T 的单射，故

$$[A : K] = |S| \leqslant |T| = [G : B]。$$

若 $[G : B] = |T|$ 有限，f 是双射 $\Leftrightarrow f$ 是满射 $\Leftrightarrow f(S) = T \Leftrightarrow AB = G$，即得所证。

例 12.25 阶为素数 p 的群 G 是循环群。

证 $p > 1$，即存在 $a \neq e$，$a \in G$。$|a| \neq 1$，$|a| \mid p$ 必 $|a| = p$，故 $G = (a)$。

例 12.26 阶为 4 的群一定是交换群。

证 设 G 为 4 阶群：

(1) G 有一个 4 阶元 a，则 $G=(a)$，是交换群；

(2) G 无 4 阶元。$\forall a\in G$，$|a|\mid 4 \Rightarrow |a|\mid 2$，故 G 是交换群（见例 12.21）。

例 12.27 阶为 6 的群 G 有且仅有一个 3 阶子群。

证 $\forall a\in G$，$|a|\mid |G|=6$，所以 $|a|=1,2,3,6$。

若 G 无 3 阶元，从而无 6 阶元，则 G 中元的阶都是 2 的因子，故 G 是交换群。（例 11.21）任选 G 中 e 外的 2 个元 a,b。此时集合 $\{e,a,b,ab\}$ 关于乘法封闭，成 G 的一个 4 阶子群，因为 4 不是 6 的因子，此不可能。故必 G 有一个 3 阶元 a，从而有 3 阶子群 (a)。

若 G 有 2 个 3 阶子群 A 和 B。$A\cap B\leqslant A$，所以 $|A\cap B|\,\Big|\,3$。若 $|A\cap B|=1$，则 $|A\cdot B|=\dfrac{|A|\cdot|B|}{|A\cap B|}=9>|G|$，此不可能。故必 $|A\cap B|=3$，即 $A=A\cap B=B$。故 G 仅有一个 3 阶子群。

例 12.28 若 A,B 为有限群 G 的子群，$[G:A]$和$[G:B]$为互素整数，则 $G=A\cdot B$。

证 设 $[G\cdot A]=s$，$[G:B]=t$，$(s,t)=1$。$t=[G:B]\mid[G:A\cap B]=[G:A][A:A\cap B]=s[A:A\cap B]$。由于 $(s,t)=1$，故 $t\mid[A:A\cap B]\leqslant[G:B]=t$，则

$$[A:A\cap B]=[G:B],$$

由定理 12.21 得

$$G=A\cdot B。$$

§4 正规子群、同态和商群

一、正规子群

定义 12.13 群 G 的子群 N 称为 G 的正规子群或不变子群，记成

$H \lhd G$，如果对$\forall a \in G$有$aH = Ha$。

判别 若H是G的子群，则$H \lhd G \Leftrightarrow \forall a \in G$，有$aHa^{-1} = H \Leftrightarrow \forall a \in G$有$aHa^{-1} \subseteq H \Leftrightarrow \forall a \in G, h \in H$有$aha^{-1} \in H$。

证 仅须证：由"$\forall a \in G$有$aHa^{-1} \subseteq H$"能推出"$\forall a \in G$有$aHa^{-1} = H$"，其证明如下：

$\forall a \in G, aHa^{-1} \subseteq H$。取$b = a^{-1}$，则$bHb^{-1} \subseteq H$，从而$a^{-1}Ha \subseteq H$，所以$H \subseteq aHa^{-1}$，故$H = aHa^{-1}$。

例 12.29 交换群的任意子群都是正规子群。

例 12.30 群G的指数为2的子群H必是G的正规子群。因为$\forall a \in H, aH = H = Ha$；$a \notin H, G = H \cup aH = H \cup Ha$是$G$的2个而实质上是一个的分类，故$aH = G - H = Ha$。所以$H \lhd G$。由此，$A_n \lhd S_n$。

例 12.31 若H是G的子群，且H的任意两个左陪集的积仍是左陪集，则$H \lhd G$。

证 设$aH \cdot bH = cH$，故$aebe \in aHbH = cH$。因$ab \in cH$，故$abH = cH$。$\forall a \in G$，$aHa^{-1}H = aa^{-1}H = H$，所以$aHa^{-1}e \subseteq aHa^{-1}H = H$，即$aHa^{-1} \subseteq H$，故$H \lhd G$。

例 12.32 求证$B_4 = \{e, (12)(34), (13)(24), (14)(23)\} \lhd S_4$

证 $\forall \sigma \in S_4$，$\sigma(e)\sigma^{-1} = e \in B_4$。又

$$\begin{aligned} &\sigma(i_1 i_2)(i_3 i_4)\sigma^{-1} = \sigma(i_1 i_2)\sigma^{-1}\sigma(i_3 i_4)\sigma^{-1} \\ &= (\sigma(i_1)\sigma(i_2))(\sigma(i_3)\sigma(i_4)) \in B_4。\end{aligned}$$

这是由于$i_1 i_2 i_3 i_4$是1,2,3,4的一个排列，σ是双射，故$\sigma(i_1), \sigma(i_2), \sigma(i_3), \sigma(i_4)$仍是1,2,3,4的一个排列的缘故。

二、商群

若N是群G的正规子群，则$G/N = \{aN \mid a \in N\}$关于群子集的积运算成一群，称为$G$关于$N$的商群。其证明如下：

(1) $\forall a, b \in G, (aN)(bN) = abNN = abN$，故$G/N$关于积运算封闭。$2^G$是半群，结合律成立，所以$G/N$是半群；

(2) $\forall a \in G$, $N \cdot aN = aN \cdot N = aN$，故 N 是 G/N 的单位元；

(3) $\forall a \in G$, $a^{-1}N \cdot aN = a^{-1}aN = N$，即 aN 可逆。故 G/N 是群。$|G/N| = [G \cdot N]$。

三、同态与同构

1. 定义

定义 12.14 群 G 到群 G' 的映射 f 称为同态映射。如果对 $\forall a, b \in G$ 有 $f(ab) = f(a) \cdot f(b)$。若 f 是单射，则称 f 为单同态；若 f 是满射，则称 f 为满同态；若 f 是双射，则称 f 为同构。如果存在群 G 到群 G' 的同态映射，则称 G 同态于 G'，记为 $G \sim G'$。如果存在群 G 到群 G' 上的同构映射，则称 G 同构于 G'，记为 $G \cong G'$。

例 12.33 任意无限循环群 (a) 同构于 $(\mathbf{Z}, +)$。因为 $f(a^k) = k$ 是 (a) 到 $(\mathbf{Z}, +)$ 的同构映射：$k \neq i \Leftrightarrow a^k \neq a^i$，故 f 是单射；显然 f 是满射；又 $f(a^k \cdot a^i) = f(a^{k+i}) = k + i = f(a^k) + f(a^i)$，所以 f 是同构。

例 12.34 若 (a) 和 (b) 为两个 n 阶循环群，则 $(a) \cong (b)$。

因为 $f(a^i) = b^i$ 是 $(a) = \{a^0, a^1, \cdots, a^{n-1}\}$ 到 $(b) = \{b^0, b^1, \cdots, b^{n-1}\}$ 的双射。又

$$f(a^i \circ a^j) = f(a^{i+j}) = f(a^r) = b^r = b^{i+j} = b^i \cdot b^j = f(a^i) \cdot f(a^j)。$$

这里 r 是 $i + j$ 被 n 除所得余数，故 f 是同构。

例 12.35 映射 $f(n) = a^n$ 是群 $(\mathbf{Z}, +)$ 到循环群 (a) 的同态映射。

例 12.36 映射 $f(x) = 2^x$ 是 $(\mathbf{R}, +)$ 到正实数乘法群 $(\mathbf{R}^+, \times)$ 的同构映射。

2. 同态的性质

若给了群 G 到群 G' 的同态 f，则有

(1) e 是 G 单位元，则 $f(e)$ 是 G' 单位元；

(2) $a \in G$，则 $f(a)$ 的逆元是 $f(a^{-1})$；

(3) $H \leqslant G$，则 $f(H) \leqslant G'$；特别 $f(G) \leqslant G'$；

(4) $H' \leqslant G'$，则 $f^{-1}(H') = \{a \mid a \in G, f(a) \in H'\} \leqslant G$。

证 (1) $a \in G$, $f(a) = f(ea) = f(e)f(a)$，由消去律得 $f(e)$ 是 G'

的单位元。

(2) 因 $f(a)\cdot f(a^{-1})=f(a\cdot a^{-1})=f(e)$，故 $f(a)$ 的逆元是 $f(a^{-1})$。

(3) $f(a),f(b)\in f(H)$，即 $a,b\in H$。有 $[f(a)]^{-1}\cdot f(b)=f(a^{-1}b)\in f(H)$，故 $f(H)\leqslant G'$。

(4) $a,b\in f^{-1}(H')$，$f(a),f(b)\in H'$，因为 $H'\leqslant G'$，所以

$$f(a^{-1}b)=[f(a)]^{-1}f(b)\in H'。$$

即 $a^{-1}b\in f^{-1}(H')$，故 $f^{-1}(H')\leqslant G$。

3. **同态的运算**

命题 12.4 群同态的积仍是群同态，群同构的逆仍是群同构。

证 若 $G_1\overset{f}{\sim}G_2$，$G_2\overset{g}{\sim}G_3$ 是群同态，则 $(gf)(xy)=g[f(xy)]=g[f(x)\cdot f(y)]=gf(x)\cdot gf(y)$，故积 gf 是群 G_1 到 G_3 的同态映射。

若 $G_1\overset{f}{\cong}G_2$ 是群同构，则 $f^{-1}(y_1\cdot y_2)=f^{-1}[(f\cdot f^{-1})(y_1)\cdot(ff^{-1})(y_2)]=f^{-1}\cdot f[f^{-1}(y_1)\cdot f^{-1}(y_2)]=f^{-1}(y_1)\cdot f^{-1}(y_2)$，故 f^{-1} 是 G_2 到 G_1 的同构。

§5 群的同态基本定理

一、自然同态

若 N 是群 G 的正规子群，则映射 φ：$\varphi(x)=xN$ 是群 G 到商群 G/N 的同态映射，称为自然同态。

二、同态的核

若给了群同态：$G\overset{f}{\sim}G'$。记 $\mathrm{Ker}f=\{x\mid x\in G,f(x)=f(e)\}$ 称为同态 f 的核。

因 $xN=N\Leftrightarrow x\in N$，故自然同态 φ 的核 $\mathrm{Ker}\varphi=N$。

命题 12.5 若给了群同态 $G\overset{f}{\sim}G'$，则 $\mathrm{Ker}f\lhd G$。

证 (1) $e \in \mathrm{Ker}f$，故 $\mathrm{Ker}f \neq \phi$；

(2) $a, b \in \mathrm{Ker}f, f(a) = f(b) = f(e)$，故

$$f(a^{-1}b) = f(a^{-1}) \cdot f(b) = [f(a)]^{-1} \cdot f(b)$$
$$= f(e) \cdot f(e) = f(e),$$

即 $a^{-1}b \in \mathrm{Ker}f$，所以 $\mathrm{Ker}f \leqslant G$。

(3) $\forall x \in G, a \in \mathrm{Ker}f$，则

$$f(xax^{-1}) = f(x) \cdot f(a) \cdot [f(x)]^{-1} = f(x) \cdot [f(x)]^{-1} = f(e),$$

所以

$$xax^{-1} \in \mathrm{Ker}f，故\ \mathrm{Ker}f \lhd G。$$

命题 12.6 若给了群同态 $G \overset{f}{\sim} G'$。则 f 是单射$\Leftrightarrow \mathrm{Ker}f = \{e\}$。

证 若 f 是单射，则 $x \neq e, f(x) \neq f(e)$，故 $\mathrm{Ker}f = \{e\}$。反之，若 $\mathrm{Ker}f = \{e\}, f(x) = f(y)$，则 $f(x^{-1}y) = f(e), x^{-1}y \in \mathrm{Ker}f = \{e\}$，所以 $x^{-1}y = e$，故 $x = y$，即 f 为单射。

推论 若满同态 $G \overset{f}{\sim} G'$ 的核为 $\{e\}$，则 f 是同构。特别 $G/\{e\} \cong G$。

三、同态基本定理

基本定理 若给了群的同态 $G \overset{f}{\sim} G'$，令 $N = \mathrm{Ker}f$，则

$$G/N \cong f(G)。$$

证 我们证明 $\sigma(aN) = f(a)$ 是 G/N 到 $f(G)$ 的一个同构映射。

(1) $aN = bN \Leftrightarrow a^{-1}b \in N \Leftrightarrow f(a^{-1}b) = f(e) \Leftrightarrow f(a) = f(b)$，故 σ 是单射。

(2) 按定义，σ 是满射。

(3) $\sigma(aN \cdot bN) = \sigma(abN) = f(ab) = f(a) \cdot f(b) = \sigma(aN) \cdot \sigma(bN)$，故 σ 是同态。

由(1)，(2)，(3)知 σ 是同构。定理获证。

例 12.37 (a)为 n 阶循环群，$f(k) = a^k$ 是$(\mathbf{Z}, +)$到(a)的同态

映射。$\mathrm{Ker}f = \{k \mid n \mid k\}$（$k$ 是 n 的倍数）记成 (n)。由基本定理得 $(a) \cong \mathbf{Z}/(n)$，后者常记为 $\mathbf{Z}_n$。

例 12.38 映射 $f(\sigma) = (-1)^{\mu(\sigma)}$ 是 S_n 到 $(-1, +1)$ 的乘法群的一个同态映射，$\mathrm{Ker}f = A_n$；由基本定理得

$$S_n/A_n \cong (\{-1,1\},x) \cong Z_2。$$

引理 12.2 若 A 与 B 都是群 G 的子群，且 $AB = BA$，则

$$AB \leqslant G。$$

证 $\forall a_ib_i \in AB, a_i \in A, b_i \in B, i = 1,2$，

$$(a_1b_1)^{-1}(a_2b_2) = b_1^{-1}a_1^{-1} \cdot a_2b_2 \in BAAB = BAB = AB^2 = AB,$$

故

$$AB \leqslant G。$$

第一同构定理 若 A 和 B 都是群 G 的子群，$B \lhd G$。则 $AB/B \cong A/(A \cap B)$。

证 因 $B \lhd G$，故 $AB = \bigcup\limits_{a\in A} aB = \bigcup\limits_{a\in A} Ba = BA$，由引理 12.2，$AB \leqslant G$，从而 $B \lhd AB$。

作群 A 到商群 AB/B 的映射 f：$f(a) = aB$。易验证，f 是满同态，$\mathrm{Ker}f = \{a \in A \mid a \in B\} = A \cap B$。由基本定理得 $A/(A \cap B) \cong AB/B$。

第二同构定理 若给了满同态 $G \overset{f}{\sim} G'$，$H' \lhd G'$，$H = f^{-1}(H')$，则 $H \lhd G$，且

$$G/H \cong G'/H'。$$

证 f 是 G 到 G' 满同态，自然同态 φ 是 G' 到 G'/H' 的满同态，故 φf 是 G 到 G'/H' 的满同态。$\mathrm{Ker}(\varphi f) = \{x \in G \mid f(x)H' = H'\} = \{x \in G \mid f(x) \in H'\} = H$。由基本定理得

$$G/H \cong G'/H'。$$

例 12.39 求证：$S_4/B_4 \cong S_3$。

证 记 $S_3' = \{\sigma \in S_4 | \sigma(4) = 4\}$，显然 $S_3' \cong S_3$。因为 B_4 中仅恒等映射保持 4 不变，故 $S'_3 \cap B_4 = \{e\}$，于是

$$|S_3' \cdot B_4| = \frac{|S_3'| \cdot |B_4|}{|S_3' \cap B_4|} = \frac{6 \cdot 4}{1} = 24 = |S_4|,$$

从而

$$S_3' \cdot B_4 = S_4。$$

由第一同构定理得

$$S_4/B_4 = \frac{S'_3 \cdot B_4}{B_4} \cong \frac{S'_3}{S_3' \cap B_4} = \frac{S'_3}{\{e\}} \cong S_3' \cong S_3。$$

§6 群的直积

本节中将讨论由给定的群构造新的、较为复杂的群——直积的方法。

一、两个群的直积

若 G_1,G_2 为两个群，在积集合 $G_1 \times G_2$ 中定义二元运算"·"：$(a_1, b_1) \cdot (a_2,b_2) = (a_1 \cdot a_2, b_1 \cdot b_2)$，则 $G_1 \times G_2$ 关于运算"·"成一群，称为群 G_1 和 G_2 的直积，仍记为 $G_1 \times G_2$。

若 G_1 的单位元为 e_1，G_2 的单位元为 e_2，则 $G_1 \times G_2$ 的单位元为(e_1, e_2)；元素(a,b)的逆元为(a^{-1},b^{-1})。如果 G_1 和 G_2 都有限，则 $G_1 \times G_2$ 也有限，且 $|G_1 \times G_2| = |G_1| \cdot |G_2|$。又 G_1 和 G_2 都是交换群时，$G_1 \times G_2$ 也是交换群。

例 12.40 若 G 表示 2 阶循环群，即由数(-1)和 1 所成乘法群。则 $G \times G = \{(1,1),(1,-1),(-1,1),(-1,-1)\} \cong B_4$。

例 12.41 若 $G_1 = \{-1,1\}$，$G_2 = \{1,\omega,\omega^2\}$。则 G_1 和 G_2 关于数的乘法分别是 2 阶和 3 阶循环群。且 $G_1 \times G_2$ 是 6 阶循环群。更一般地，有：

例 12.42 若用 C_n 表 n 阶循环群，$(n,s) = 1$。则 $C_n \times C_s \cong C_{ns}$。

证 设 $C_n = (a), C_s = (b)$，则

$$|(a,e)| = |a| = n, |(e,b)| = |b| = s, (n,s) = 1。$$

又 $(a,e)(e,b) = (e,b)(a,e) = (a,b)$，所以

$$|(a,b)| = |(a,e)| \cdot |(e,b)| = n \cdot s = |C_n \times C_s|$$

故

$$C_n \times C_s = ((a,b)) \cong C_{ns}。$$

定义 12.15 给了群 G 的 n 个子集 $(n \geqslant 2) G_1, G_2, \cdots, G_n$；如果 $\forall x \in G$，都存在 $a_i \in G_i (i = 1, \cdots, n)$ 使 $x = a_1 a_2 \cdots a_n$，则称 G 中元素可表成 $G_1, \cdots, G_n$ 中元素之积。又若由 $\forall x \in G$，$x = a_1 \cdots a_n = b_1 \cdots b_n$，这里 $a_i, b_i \in G_i (i = 1, \cdots, n)$，必能推出 $a_i = b_i (i = 1, \cdots, n)$，则称 G 中元素可惟一表成 $G_1, \cdots, G_n$ 中元素之积。

引理 12.3 若 G_1 和 G_2 是群 G 的子群，G 中元素可表成 G_1, G_2 中元素之积，则下面三个命题彼此等价：

(1) G 中元素可惟一表成 G_1, G_2 中元素之积；

(2) e 可惟一表成 G_1, G_2 中元素之积，即若 $e (= e \cdot e) = a_1 \cdot a_2, a_1 \in G_1, a_2 \in G_2$，则 $a_1 = a_2 = e$；

(3) $G_1 \cap G_2 = \{e\}$。

证 (1)⇒(2)：显然。

(2)⇒(3)：若 $x \in G_1 \cap G_2$，则 $x^{-1} \in G_1 \cap G_2$。$e = e \cdot e = x \cdot x^{-1} = x^{-1} \cdot x$，由 e 表法惟一，得 $x = e$，故 $G_1 \cap G_2 = \{e\}$。

(3)⇒(1)，$\forall x \in G$，$x = a_1 a_2 = b_1 b_2, a_1, b_1 \in G_1; a_2, b_2 \in G_2$。则 $b_1^{-1} a_1 = b_2 a_2^{-1} \in G_1 \cap G_2 = \{e\}$，所以 $b_1^{-1} a_1 = b_2 a_2^{-1} = e \Rightarrow b_1 = a_1, a_2 = b_2$，故得引理的证明。

引理 12.4 若 G_1 和 G_2 是群的子群，G 中元素可惟一表成 G_1 和 G_2 中元素之积，则

$$G_1, G_2 \lhd G \Leftrightarrow \forall a \in G_1, b \in G_2 \text{ 有 } ab = ba。$$

证 ⇒：由引理 12.3，$G_1 \cap G_2 = \{e\}$。$\forall a \in G_1$，$b \in G_2$，

$a(ba^{-1}b^{-1}) \in G_1$，同理 $\in G_2$，所以 $aba^{-1}b^{-1} \in G_1 \cap G_2$，故 $aba^{-1}b^{-1} = e$，从而 $ab = ba$。

$\Leftarrow$：$\forall a \in G_1$，$x \in G$，则 $x = a_1b_1, a_1 \in G_1, b_1 \in G_2$。此时 $xax^{-1} = a_1b_1ab_1^{-1}a_1^{-1} = a_1aa_1^{-1} \in G_1$。故 $G_1 \lhd G$，类此有 $G_2 \lhd G$。

定理 12.22 若 G_1 和 G_2 是群 G 的正规子群，G 中元素可惟一表成 G_1 和 G_2 中元素之积，则 $G \cong G_1 \times G_2$，称 G 为 G_1 和 G_2 的内部直积，也记成 $G = G_1 \times G_2$。

证 作 $G_1 \times G_2$ 到 G 的映射 f：$f((a,b)) = ab$ 由"G 中元素可惟一表成 G_1 和 G_2 中元素之积"可得，f 是双射(可表成意味着 f 是满射，惟一表成意味着 f 是单射)。

又 $f((a_1,b_1) \cdot (a_2,b_2)) = f((a_1a_2,b_1b_2)) = a_1a_2b_1b_2$。由引理 12.2，$a_1a_2b_1b_2 = (a_1b_1)(a_2b_2) = f((a_1,b_1)) \cdot f((a_2,b_2))$。

由上可得 f 是同构映射，本定理得证。

利用引理 12.3 和 12.4，即得

定理 12.23 若 G_1 和 G_2 是群 G 的子群，且满足：

(1) $G = G_1 \cdot G_2$(即 G 中元素可表成 G_1 和 G_2 中元素之积)；

(2) $G_1 \cap G_2 = \{e\}$；

(3) $\forall a \in G_1$，$b \in G_2$，$ab = ba$，则

$$G \cong G_1 \times G_2。$$

例 12.43 若 $(a) = C_{pq}$，$(p,q) = 1$，则 $(a) \cong (a^p) \times (a^q)$。

证 (1) 因 $(p,q) = 1$，所以存在 $s,t \in \mathbf{Z}$ 使 $sp + tq = 1$，故 $a = (a^p)^s \cdot (a^q)^t \in (a^p) \cdot (a^q) \Rightarrow (a) = (a^p) \cdot (a^q)$。

(2) $|a^p| = q$，$|a^q| = p$。$\forall x \in (a^p) \cap (a^q) \Rightarrow |x| \Big| p$ 和 $q \Rightarrow |x| \mid (p,q) = 1$，所以 $x = e$，故 $(a^p) \cap (a^q) = \{e\}$。

由(1)和(2)，加上"循环群是交换群"，利用定理 12.22，即得证明。

例 12.44 若群 G 是其子群 A 和 B 的内部直积，$N \lhd A$，则 N 是 G 的正规子群，且 $G/N \cong A/N \times B$。

证 作 G 到 $A/N \times B$ 的映射 f：$\forall x \in G$，$x = ab, a \in A, b \in B$，$f(x) = (aN, b)$。易得 f 是满同态。$\mathrm{Ker}f = \{x = ab \mid (aN,b) = (N,e)\}$

$= \{x = ab \mid a \in N, b = e\} = N$。故得证明。

二、有限个群的直积

类似于两个群直积的情形，有以下的定义和定理。

定义 12.16 若 $G_1, G_2, \cdots, G_n$ 是 n 个群，则积集合 $G_1 \times \cdots \times G_n$ 关于运算（乘法）：$(a_1, \cdots, a_n) \cdot (b_1, \cdots, b_n) = (a_1b_1, \cdots, a_nb_n)$ 成一群，称为群 $G_1, \cdots, G_n$ 的直积，仍记成 $G_1 \times G_2 \times \cdots \times G_n$。

若 G_i 的单位元素为 $e_i (i=1, \cdots, n)$，则 $G_1 \times \cdots \times G_n$ 的单位元素为 $(e_1, \cdots, e_n)$；元素 $(a_1, \cdots, a_n)$ 的逆元素为 $(a_1^{-1}, \cdots, a_n^{-1})$。当 $G_1, \cdots, G_n$ 都是交换群时，$G = G_1 \times \cdots \times G_n$ 也是交换群。

映射 φ_i：$\varphi_i(a) = (e_1, \cdots, e_{i-1}, a, e_{i+1}, \cdots, e_n)$ 是群 G_i 到 $G_1 \times \cdots \times G_n$ 的单同态，记 $\varphi_i(G_i) = G_i'$，则

$$G_i \cong G_i' \text{ 且 } G_i' \lhd G_1 \times \cdots \times G_n。$$

定义 12.17 若 $G_1, \cdots, G_n$ 是群 G 的 n 个正规子群，G 中元素可惟一表成 $G_1, \cdots, G_n$ 中元素之积，则称 G 为 $G_1, \cdots, G_n$ 的内部直积，记成 $G = G_1 \otimes \cdots \otimes G_n$。

定理 12.24 $G_1, \cdots, G_n$ 是群，G_i' 如上定义，则

$$G_1 \times \cdots \times G_n = G_1' \otimes \cdots \otimes G_n'。$$

定理 12.25 若群 G 为其正规子群 $G_1, \cdots, G_n$ 的内部直积，则 $G \cong G_1 \times \cdots \times G_n$。

定理 12.26 若 $G_1, \cdots, G_n$ 是群 G 的子群，满足：

(1) $G = G_1 \cdot G_2 \cdot \cdots \cdot G_n$；

(2) $G_i \cap \prod\limits_{j \neq i} G_j = \{e\}, i = 1, \cdots, n$；

(3) $i \neq j$，$\forall a_i \in G_i, a_j \in G_j$ 有 $a_ia_j = a_ja_i$，则

$$G \cong G_1 \times \cdots \times G_n。$$

定理 12.24 是显然的。定理 12.25 和 12.26 的证明类似于 $n=2$ 的情形，留给读者作为练习。

例 12.45 若 $G = (a)$ 是 n 阶循环群，$n = p_1^{\alpha_1} \cdots p_t^{\alpha_t}$，这里 $p_1, \cdots,$

p_t 是互异素数。记 $n_i = \frac{n}{p_i^{a_i}}$，$G_i = (a^{n_i})$，则

$$G \cong G_1 \times \cdots \times G_t。$$

证 (1) G 是循环群，故乘法可交换，所以 $G_i \lhd G \quad (i = 1,\cdots,t)$。

(2) 因 $(n_1,\cdots,n_t)=1$，所以存在 $k_1,\cdots,k_t$ 使 $\sum_{i=1}^{t} k_i n_i = 1$。故 $a = a^1 = (a^{n_1})^{k_1}\cdots(a^{n_t})^{k_t}$，即 a 可表成 $G_1,\cdots,G_t$ 中元素之积，故 G 中任意元素可由 $G_1,\cdots,G_t$ 中元素表出。

(3) 若 $a_1\cdots a_t = b_1\cdots b_t$，这里 $a_i,b_i \in G_i(i = 1,\cdots,t)$。记 $a_i \cdot b_i^{-1} = x_i$，则 $x_1\cdots x_t = e$。因 $x_i \in G_i$，故 $|x_i| \mid |G_i| = p_i^{a_i}$。又

$$x_i = \prod_{j\neq i} x_j^{-1} \in \prod_{j\neq i} G_j,$$

从而

$$|x_i| \mid n_i,$$

于是

$$|x_i| \mid (p_i^{a_i},n_i) = 1,$$

即

$$x_i = e，故\ a_i = b_i(i = 1,\cdots,t)。$$

由(1)，(2)，(3)可知，G 是 $G_1,\cdots,G_t$ 的内部直积，因而

$$G \cong G_1 \times \cdots \times G_t。$$

定理 12.27 若 N_i 是群 $G_i(i=1,\cdots,t)$的正规子群，则

$$N_1 \times \cdots \times N_t \lhd G_1 \times \cdots \times G_t,$$
$$\frac{G_1 \times \cdots \times G_t}{N_1 \times \cdots \times N_t} \cong \frac{G_1}{N_1} \times \cdots \times \frac{G_t}{N_t}。$$

证 映射 σ：

$$\sigma((a_1,\cdots,a_t)) = (a_1N_1,\cdots,a_tN_t)$$

是群 $G_1\times\cdots\times G_t$ 到群 $\dfrac{G_1}{N_1}\times\cdots\times\dfrac{G_t}{N_t}$ 的满同态，其核为

$$N_1\times\cdots\times N_t,$$

由同态基本定理得

$$\frac{G_1\times\cdots\times G_t}{N_1\times\cdots\times N_t}\cong\frac{G_1}{N_1}\times\cdots\times\frac{G_t}{N_t}。$$

习　题　12

1. 若$(\boldsymbol{A},\cdot)$为一半群，求证：$\boldsymbol{A}\times\boldsymbol{A}$ 关于下面的二元运算"$\circ$"也成一个半群：

$$(a_1,a_2)\circ(b_1,b_2)=(a_1\cdot b_1,a_2\cdot b_2)。$$

2. 若 n 为一固定正整数，在 $\mathbf{Z}$ 中定义运算"$\circ$"：

$$a\circ b=a+b-n,$$

求证：$(\mathbf{Z},\circ)$为一群，并求它的单位元和任意元的逆元。

3. 设 G 为有限群，$|G|=n$，在 $G\times G$ 中定义关系 R：$(a,b)R(c,d)\Leftrightarrow ab=cd$，求证：$R$ 是 $G\times G$ 的一个等价关系且商集合的阶 $|G\times G/R|=|G|=n$。

4. 设 $\boldsymbol{P}$ 为数域，$\boldsymbol{G}$ 表 $\boldsymbol{P}$ 上行列式不为零的 n 阶方阵全体：

$$\boldsymbol{G}=\{\boldsymbol{A}\in\boldsymbol{P}^{n\times n}\mid|\boldsymbol{A}|\neq0\},$$

求证：(1) $\boldsymbol{G}$ 关于矩阵的乘法运算构成群。

(2) 若记 $\boldsymbol{H}=\{\boldsymbol{A}\in\boldsymbol{G}\mid|\boldsymbol{A}|=1\}$，则 $\boldsymbol{H}\leqslant\boldsymbol{G}$。

5. 设 A 为一集合，求证 2^A 关于下面的二元运算"$\circ$"成一群：

$$B\circ C=(B\cup C)-(B\cap C)。$$

6. 若 A,B 为群 G 的两个子群，且 $A\not\subset B,B\not\subset A$，求证：$A\cup B$ 不是 G 的子群。

7. 在群$(\mathbf{Z},+)$中，设$S_1=\{7,9\}$，$S_2=\{8,12\}$，试确定(S_1)和(S_2)。

8. 若a,b是有限群G的两个元素，求证：$|ab|=|ba|$。

9. 求证：群中元和它的逆元有相同的阶，即a是群G的一个元素，则$|a|=|a^{-1}|$。

10. 若a,b为群G的两个元素，$ab=ba$，且$(|a|,|b|)=1$，求证：$|ab|=|a|\cdot|b|$。

11. 若G为有限群，求证：

(1) G的乘法表中任意一行(列)没有相同元素；

(2) G是交换群$\Leftrightarrow$乘法表作为矩阵是对称的；

(3) 若G的乘法表中i行、s列的元素为e；i行、t列的元素为b；j行、s列的元素为a；则j行、t列的元素必为$a\cdot b$。

12. 若$|a|=12$，求循环群(a)的所有子群和生成元。

13. 若$G=(a)$为循环群，$A=(a^s)$，$B=(a^t)$，求$A\cap B$和$(A\cup B)$的一个生成元。

14. 若a是群G的一个元素，a的阶为素数p，$H\leqslant G$，$a\overline{\in}H$，求证：$(a)\cap H=\{e\}$。

15. 若H是有限群G的子群，$a\in G$，记$k=\min\{n\in\mathbf{Z}^+\mid a^n\in H\}$，则

$$k\mid |a|。$$

16. 在S_8中，表下列置换为不相交循环之积，并求其阶与逆：

(1) $\begin{pmatrix}1&2&3&4&5&6&7&8\\5&8&6&1&4&3&2&7\end{pmatrix}$；

(2) $\begin{pmatrix}1&2&3&4&5&6&7&8\\8&5&2&7&3&6&4&1\end{pmatrix}$。

17. 作出S_3的乘法表。

18. 表S_4中所有奇置换为不相交循环之积。

19. 求证：S_n中置换必能表成下列$(n-1)$个对换之积(12)，(13)，$\cdots$，$(1n)$。

20. 求证：$S_n = (\{(12),(12\cdots n)\})$。

21. 求证：$A_n = (\{(123),(124),\cdots,(12n)\})$。

22. 在 S_n 中，设 $a = (12\cdots n)$，$b = \prod_{i=2}^{\left[\frac{n+1}{2}\right]} (i, n + 2 - i)$，求证：$\{a, b\}$ 生成的群的阶为 $2n$。

23. 作出 S_3 关于 $H = \{e,(13)\}$ 的左陪集和右倍集分解。

24. 作出 S_4 关于 B_4 的陪集分解。

25. A,B 是群 G 的子群，且 $A \cdot B \leqslant G$，求证：$AB = BA$。

26. 若 A,B 是群 G 的子群，C 表示 $A \cup B$ 生成的子群，求证：$[A : A \cap B] \leqslant [C : B]$。

27. 若 A,B 是群 G 的子群，$(|A|,|B|) = 1$，求证：

$$|AB| = |A| \cdot |B|。$$

28. 若 p,q 都是素数，$p > q$。则 $p \cdot q$ 阶群至多有一个 p 阶子群。

29. 若 A,B 是群 G 的指数有限的子群，$[G : A] = s$，$[G : B] = t$，则 $[G : A \cap B] \leqslant st$。

30. 若 H,K,N 是群 G 的子群，$H \leqslant K$，$H \cap N = K \cap N$，$H \cdot N = K \cdot N$，求证：$H = K$。

31. 若 A 和 B 为有限群 G 的非空子集，且 $|A| + |B| > |G|$，求证：$G = A \cdot B$。

32. 求证：一个群的正规子群的交仍是正规子群。

33. 若 H 是群 G 的子群 $\forall a \in G$，求证：

$$aHa^{-1} \leqslant G \text{ 且 } aHa^{-1} \sim H。$$

34. 若 H 是有限群 G 的惟一 n 阶子群，则 $H \lhd G$。

35. 若 A,B 都是群 G 的正规子群，且 $A \cap B = \{e\}$，求证：$\forall a \in A, b \in B$ 有 $ab = ba$。

36. 若 $N \lhd G$，$[G : N] = n$，求证：$\forall x \in G$，必 $x^n \in N$。

37. 求证：映射 $f(x) = x^{-1}$ 是群 G 的自同构 $\Leftrightarrow G$ 是交换群。

38. 求证：阶为 4 的群或者是循环群或者同构于 B_4。

39. 若 G 为有限交换群，整数 $p \mid |G|$，求证：G 中有阶为 p 的元。

40. 若 A 和 B 都是群 G 的正规子群，且商群 G/A 和 G/B 都是交换群，求证：商群 $G/A\cap B$ 也是交换群。

41. 给了群同态 $G \overset{f}{\sim} G'$，$\mathrm{Ker}f = N$，$H \leqslant G$，求证：$f^{-1}(f(H)) = HN$。

42. 若 N 是群 G 的正规子群，$H \leqslant G$；$[G:N]$ 和 $|H|$ 有限且互素，求证：$H \subseteq N$。

43. 若 N 是群 G 的正规子群，$H\leqslant G$；$[G:H]$和$|N|$有限且互素，求证：$N\subseteq H$。

44. 若 G_1,G_2 是两个群，求证：$G_1\times G_2\cong G_2\times G_1$。

45. 若群 G 是其子群 G_1 和 G_2 的内部直积，求证：

$$G/G_1\cong G_2 \quad G/G_2\cong G_1.$$

46. 若群 G 是其子群 G_1 和 G_2 的内部直积，且 $N\lhd G_1$，则 $N\lhd G$。

47. A 和 B 为有限群，$A\times B$ 为循环群，则 $(|A|,|B|)=1$。

48. 若 N 和 K 都是 G 的正规子群，且 $N\cap K=\{e\}$，求证：G 同构于 $G/N\times G/K$ 的一个子群。

49. 若 N 和 K 都是 G 的正规子群，且 $G=N\cdot K$。记 $H=N\cap K$，求证：$G/H\cong N/H\times K/H$。

50. $G_1,G_2,\cdots,G_s$ 为 s 个有限循环群，其阶分别为 $n_1,n_2,\cdots,n_s$，则

$$G_1\times\cdots\times G_s \text{ 为循环群} \Leftrightarrow [n_1,\cdots,n_s]=\prod_{i=1}^{s}n_i。$$

第十三章　环

环是比群略为复杂的一种常见的代数体系。环是一个集合连同两种运算所构成的。本章的内容大致分三部分：第一部分是群论中相应内容的推广，第二部分是环的嵌入，第三部分是因式分解理论。

§1　环的概念

一、环的定义

定义 13.1　带有两个二元运算（通常表示成加法和乘法）的非空集合 R 称为环，如果：

(1) $(R,+)$是交换群；

(2) $(R,\cdot)$是半群；

(3) 乘法对加法满足左、右分配律，即$\forall a,b,c\in R$，有

$$a\cdot(b+c)=ab+ac,$$
$$(a+b)\cdot c=ac+bc。$$

若$(R,\cdot)$是交换半群，则称 R 是交换环。若$(R,\cdot)$有单位元，则称 R 为含单位元的环。

环中加法群的单位元称作环的零元，表成 0。对 $a\in R$，$\underbrace{a+a+\cdots+a}_{n个}$常记成 na，$\underbrace{a\cdot a\cdots\cdot a}_{n个}$常记成 a^n。

本章中，我们习惯地用英文大写字母 R 表示环，读者不要与实数集合混淆。

常用数环　数集 $\mathbf{Z}$，$\mathbf{Q}$，$\mathbf{R}$，$\mathbf{C}$ 对通常数的加法与乘法均构成环，统称数环。

全矩阵环 设 R 是环，用$(\boldsymbol{R})_n$ 表示元素在 R 中的 n 阶方阵全体：

$$(\boldsymbol{R})_n = \{(a_{ij})_{n\times n} | a_{ij} \in R,\ i,\ j = 1,\ \cdots,\ n\},$$

在$(\boldsymbol{R})_n$ 中定义加法与乘法：

$$(a_{ij})_{n\times n} + (b_{ij})_{n\times n} = (a_{ij} + b_{ij})_{n\times n},$$

$$(a_{ij})_{n\times n} \times (b_{ij})_{n\times n} = \left(\sum_{k=1}^{n} a_{ik}b_{kj}\right)_{n\times n},$$

则$(\boldsymbol{R})_n$ 成环，称为 R 上的全矩阵环

多项式环 若 A 为数环，则多项式集 $A[x]$关于多项式的加法与乘法成环，称为 A 上的多项式环。一般地，若 R 是含单位元 e 的环，记

$R[x] = \{(a_0,\ a_1,\ \cdots) | a_i \in R$，存在非负整数 n，使 $k \geqslant n$ 时，$a_k = 0\}$

则 $R[x]$ 对于下面的加法与乘法成环，称为 R 上的多项式环：

$$(a_0,\ a_1,\ \cdots) + (b_0,\ b_1,\ \cdots) = (a_0 + b_0,\ a_1 + b_1,\ \cdots),$$

$$(a_0,\ a_1,\ \cdots) \cdot (b_0,\ b_1,\ \cdots) = (c_0,\ c_1,\ \cdots),$$

这里

$$c_k = \sum_{i=0}^{k} a_i b_{k-i} = \sum_{i+j=k} a_i b_j.$$

如果用 1 表示 $(e,\ 0,\ 0,\ \cdots)$，x 表示 $(0,\ e,\ 0,\ \cdots)$，则

$$x^n = (\underbrace{0,\ \cdots,\ 0}_{n\text{个}},\ e,\ 0\cdots),$$

于是

$$(a_0,\ a_1,\ \cdots,\ a_n,\ 0,\ 0,\ \cdots) = a_0 1 + a_1 x + \cdots + a_n x^n = \sum_{i=0}^{n} a_i x^i,$$

这里

$$x^0 = 1,\ a_i x^i = (\underbrace{0,\ \cdots,\ 0}_{i\text{个}},\ a_i,\ 0\cdots).$$

同余类环 设 n 为一正整数，在 $Z_n = \{\bar{0},\ \bar{1},\ \cdots,\ \overline{n-1}\}$ (Z 关于

n 的同余类所成的集合)中定义加法与乘法：

$$\bar{a} + \bar{b} = \overline{a + b},$$
$$\bar{a} \cdot \bar{b} = \overline{a \cdot b}。$$

当 $\bar{a}_1 = \bar{a}$，$\bar{b}_1 = \bar{b}$，即 $n|a - a_1$ 与 $b - b_1$ 时

$$n|(a - a_1) + (b - b_1) = (a + b) - (a_1 + b_1),$$

即$\overline{a+b}=\overline{a_1+b_1}$，故上面定义的加法是 Z_n 的一个二元运算。又

$$n|a(b-b_1)+(a-a_1)b_1=ab-a_1b_1,$$

即$\overline{ab}=\overline{a_1b_1}$，故上面定义的乘法也是 Z_n 的一个二元运算。

易验证，Z_n 成环，称为模 n 的同余类环。

群环 设

$$G = \{e_1, e_2, \cdots, e_n\}$$

为有限群，R 为环，在集合

$$R(G) = \{a_1e_1 + \cdots + a_ne_n | a_1, \cdots, a_n \in R\}$$

中定义加法与乘法：

$$\sum_{i=1}^{n} a_ie_i + \sum_{i=1}^{n} b_ie_i = \sum_{i=1}^{n} (a_i + b_i)e_i,$$
$$\left(\sum_{i=1}^{n} a_ie_i\right) \cdot \left(\sum_{i=1}^{n} b_ie_i\right) = \sum_{i,j=1}^{n} (a_ib_j)(e_ie_j)。$$

则 $R(G)$成环，称为群 G 上的 R 环，简称群环。

例如，设 G 为乘法群$\{1, -1, i, -i\}$，R 为实数环，则 G 上的 R 环是复数环 C。

二、性质

设 R 是一个环，则有

(1) $\forall a \in R$，有 $a \cdot 0 = 0 \cdot a = 0$；

(2) $\forall a, b \in R$，有$(-a)b = a(-b) = -(ab)$；

$$(-a)(-b)=ab;$$

(3) $\forall\, a, b \in R, n \in Z$，有$(na)b = a(nb) = n(ab)$；

(4) $\forall\, a_i, b_j \in R, (i = 1, \cdots, n, j = 1, \cdots, s)$ 有

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

证 (1) $0a = (0+0)a = 0a + 0a$，故 $0a = 0$。类似地有 $a0 = 0$。

(2) 因为 $ab + (-a)b = [a + (-a)]b = 0b = 0$，

$$ab + a(-b) = a[b + (-b)] = a0 = 0,$$

故$(-a)b$，$a(-b)$ 与 $-(ab)$ 都是 ab 的负元，由负元的惟一性，得

$$(-a)b = a(-b) = -(ab)。$$

类似地，$(-a)(-b)$ 与 ab 都是$(-a)b$ 的负元，由负元的惟一性，得

$$(-a)(-b) = ab。$$

(3) $n = 0$ 时，等式成立。$n \in \mathbf{Z}^+$ 时

$$(na)b = (\underbrace{a + \cdots + a}_{n\text{个}})b = \underbrace{ab + \cdots + ab}_{n\text{个}} \quad (= n(ab))$$
$$= a(\underbrace{b + \cdots + b}_{n\text{个}}) = a(nb),$$

$n = -k$ 而 $k \in \mathbf{Z}^+$ 时

$$(na)b = (-ka)b = -[(ka)\cdot b],$$
$$a(nb) = a(-kb) = -[a(kb)],$$
$$n(ab) = -[k(ab)],$$

由

$$(ka)b = a(kb) = k(ab),$$

即得

$$(na)b = a(nb) = n(ab)。$$

(4) 对 n, s 用归纳法即得证明。

三、整环、体和域

定义 13.2 若 a,b 是环 R 的两个非零元，且 $ab=0$，则称 a 为 R 的左零因子，b 为 R 的右零因子。既是左零因子又是右零因子的元素称为零因子。

定义 13.3 环 R 的元 a 称为幂零元，如果存在 $n\in\mathbf{Z}^+$ 使 $a^n=0$。

例如 在 $(Z)_2$ 中，$\begin{pmatrix}1&2\\2&4\end{pmatrix}\begin{pmatrix}-2&6\\1&-3\end{pmatrix}=\begin{pmatrix}0&0\\0&0\end{pmatrix}$，又 $\begin{pmatrix}0&1\\0&0\end{pmatrix}^2=\begin{pmatrix}0&0\\0&0\end{pmatrix}$，所以 $\begin{pmatrix}1&2\\2&4\end{pmatrix}$ 是左零因子，$\begin{pmatrix}-2&6\\1&-3\end{pmatrix}$ 是右零因子，$\begin{pmatrix}0&1\\0&0\end{pmatrix}$ 是幂零元。

又如，在交换环 Z_6 中，$\bar{2}\cdot\bar{3}=\bar{3}\cdot\bar{2}=0$，所以 $\bar{2}$，$\bar{3}$ 是零因子。

定义 13.4 含单位元 $(\neq 0)$ 且不含零因子的交换环称为整环。含单位元 $(\neq 0)$，且非零元都可逆的环称为体，交换体称为域。

域既是整环又是体。

命题 13.1 若环 R 中没有左、右零因子，则环中乘法消去律成立，即

$$a\neq 0,\ ax=ay\Rightarrow x=y;$$

$$a\neq 0,\ xa=ya\Rightarrow x=y.$$

反之，若 R 中乘法消去律成立，则 R 中没有左、右零因子。

证 $a\neq 0$，$ax=ay$，则 $ax-ay=a(x-y)=0$。由于 R 没有左、右零因子，故 $x=y$。同理由 $a\neq 0$，$xa=ya$ 可得 $x=y$，即 R 中乘法消去律成立。

反之，R 中乘法消去律成立，$a\neq 0$，$ab=0$，则 $ab=a\cdot 0\Rightarrow b=0$，即 R 中 $\forall$ 非 0 元都不是左零因子，故环 R 没有左零因子，同理没有右零因子。

例 13.1 用 $Z[\mathrm{i}]$ 表示复数集 $\mathbf{C}$ 的子集：

$$Z[\mathrm{i}]=\{a+b\mathrm{i}\mid a,b\in\mathbf{Z}\}$$

则 $Z[i]$是整环，称为 Gauss 整环。

例 13.2 若 p 是素数，则同余类环 Z_p 是域。

证 Z_p 是交换环，含单位元 $\bar{1} \neq \bar{0}$。又$\bar{a} \neq \bar{0}$，则 $p \nmid a$，$(p,a)=1$，所以存在 $s,t \in Z$ 使 $sa + pt = 1$，即$\bar{s}\bar{a} = \bar{1}$，故$\bar{a}$可逆，因而 Z_p 是域。

例 13.3 有限整环是域。因为有限整环的非零元乘法半群满足消去律，故是乘法群，从而整个环是域。

例 13.4 整环上的多项式环仍是整环，因为两个非零多项式积的首项等于首项之积，故非零。

四元数体 令 G 为 8 元群：$G = \{\pm e, \pm i, \pm j, \pm k\}$，这里 e 为单位元，$(-e)^2 = e$，$-i = (-e)i$，$-j = (-e)j$，$-k = (-e)k$。且

$$(\pm i)^2 = (\pm j)^2 = (\pm k)^2 = -e,$$
$$ij = (-j)i = k,$$
$$jk = (-k)j = i,$$
$$ki = (-i)k = j.$$

则群环 $Q[G]$和 $R[G]$都是体，称为四元数体（**R** 是实数集）。

四元数体中，单位元为

$$e = 1e + 0i + 0j + 0k,$$

非零元

$$\alpha = ae + bi + cj + dk, \ ((a,b,c,d) \neq (0,0,0,0)),$$

有逆元

$$\frac{ae - bi - cj - dk}{a^2 + b^2 + c^2 + d^2}。$$

8 元群也可取 S_8 的子群，它的 8 个元素分别为：

e 为恒等映射，

$-e = (13)(24)(57)(68)$，$(-e)^2 = e$，

$i = (1\,234)(6\,587)$，$i^2 = -e$，$-i = (-e)i$，

$j = (1\,537)(2\,648)$，$j^2 = -e$，$-j = (-e)j$，

$k = (1\,836)(2\,547)$，$k^2 = -e$，$-k = (-e)k$。

四、环的特征

设 R 是一个环，记 $S = \{n \in \mathbf{Z}^+ \mid \forall a \in R,\ na = 0\}$。

若 $S = \varphi$，称环 R 的特征为 0，记成 $\mathrm{ch}\,R = 0$。

若 $S \neq \varphi$，则 s 中有最小元 k，称 k 为环 R 的特征，记成 $\mathrm{ch}\,R = k$。

定理 13.1 若 R 是含单位元 e 的环，则

(1) 如果 e 作为加法群的元，其阶为 $|e|_+ = +\infty$，则 $\mathrm{ch}\,R = 0$；

(2) 如果 $|e|_+ = n$，则 $\mathrm{ch}\,R = n$。

证 (1) 是显然的。

(2) $|e|_+ = n$，仅需证 $\forall a \in R$，$na = 0$ 即可。因为 $na = n(ea) = (ne)a = 0 \cdot a = 0$，故得定理的证明。

定理 13.2 若环 R 不含零因子，且 $\mathrm{ch}\,R \neq 0$，则 $\mathrm{ch}\,R$ 必是一素数，且 R 中每个非零元，作为加法群的元，其阶都是 $\mathrm{ch}\,R$。

证 先证非零元的阶（加群）为素数，然后证这个素数就是环 R 的特征。

设 a 为 R 的一个非零元，若 $|a|_+ = n$ 非素数，则 $n = kl$ $(k, l < n)$。此时 $0 = (na) \cdot a = (kla) \cdot a = (la)(ka)$。因为 R 不含 0 因子，必 $la = 0$ 或 $ka = 0$，这都矛盾于 $|a|_+ = n$，故必 n 为某个素数 p：$|a|_+ = p$。$\forall b \in R$，$b \neq 0$，$(pb)a = b(pa) = 0$，因为 $a \neq 0$，必 $pb = 0$，所以 $\mathrm{ch}\,R = p$。

例 $\mathrm{ch}\,Z = 0$，$\mathrm{ch}\,Z_n = n$。

§2 理想与同余类环

一、子环

定义 13.5 如果环 R 的非空子集关于环的两种运算仍是环，则称它为 R 的子环。

若 N 是环 R 的子环，则 N 作为加法群 R 的子群是正规的，故有商群 $R/N=\{a+N|a\in R\}$。R/N 是加法群：$(a+N)+(b+N)=(a+b)+N$。要想在 R/N 中定义乘法，使它成一个环，很自然地类似同余类环，期望 $(a+N)\cdot(b+N)$ 对应 $ab+N$ 是 R/N 的二元运算。这要求 N 满足一定的条件，从而产生了理想的概念。理想在环中的地位相当于正规子群在群中的地位。

二、理想

1. 定义

定义 13.6 环 R 的非空子集 I 称为 R 的理想，如果满足：

(1) $\forall a,b\in I$，必 $a-b\in I$；

(2) $\forall a\in I$，$r\in R$，必 ar，$ra\in I$。

由(1)，I 是 R 的加法子群，(2)表明 I 对乘法封闭。故有下面的推论：

推论 环的理想一定是子环。

例 13.5 I 是整数环 Z 的理想$\Leftrightarrow$存在非负整数 n，使 $I=(n)=\{nk|k\in Z\}$。

证 $\Rightarrow$：设 I 是 Z 的理想，因而是 Z 的加法子群，Z 是由 1 生成的加法(无限)循环群，故存在非负整数 n 使

$$I=(n)=\{nk|k\in Z\}。$$

$\Leftarrow$：$I=\{nk|k\in Z\}$，因为

(1) $nk_1+nk_2=n(k_1-k_2)\in I$，

(2) $r(nk)=(nk)r=n(kr)\in I$，

故 I 是 Z 的理想。

例 13.6 R 是一个环，n 为一固定正整数，记

$$I=\{a\in R|na=0\},$$

则 I 是 R 的理想。

证 $0\in I$，故 $I\neq\varnothing$。

(1) $a,b\in I$，则 $na=0$，$nb=0$，从而

$$n(a-b)=na-nb=0,$$

即

$$a-b\in I。$$

(2) $r\in R$, $a\in I$, 则 $na=0$,从而

$$n(ra)=r(na)=r\cdot 0=0,$$
$$n(ar)=(na)\cdot r=0\cdot r=0,$$

即 ar, $ra\in I$。故 I 是 R 的理想。

2. 运算

定理 13.3 环的任意个理想之交仍是理想,即 $I_i(i\in T)$ 是环 R 的理想,则

$$I=\bigcap_{i\in T} I_i$$

也是 R 的理想。

证 由定理 12.6,I 是 R 的加法子群,又,$\forall r\in R$, $a\in I$,则 $a\in I_i(i\in T)$,因为 I_i 是 R 的理想,故

$$ra,\ ar\in I_i,\quad i\in T$$

于是

$$ra,\ ar\in \bigcap_{i\in T} I_i=I,$$

所以 I 是 R 的理想。

定理 13.4 环的有限个理想之和仍是理想。

证 就两个理想之和予以证明。若 I_1 和 I_2 是环 R 的两个理想,则

$$I=I_1+I_2=\{a_1+a_2\,|\,a_i\in I_i,\ i=1,\ 2\},\ I\supseteq I_i\neq\varnothing;$$

若

$$a_1+a_2,\ b_1-b_2\in I,$$

(这里 a_1, $b_1\in I_1$; a_2, $b_2\in I_2$),则

$$(a_1+a_2)-(b_1+b_2)=(a_1-b_1)+(a_2-b_2)\in I_1+I_2=I,$$

若 $r\in R$，则

$$r(a_1+a_2)=ra_1+ra_2\in I_1+I_2=I。$$

同理

$$(a_1+a_2)r\in I。$$

故 I 是 R 的理想。

理想的并与积一般不是理想，但有：

定理 13.5 若 $I_i(i\in T)$ 是环 R 的理想簇，且 $\forall i,\ j\in T$ 必 $I_i\subseteq I_j$ 或 $I_j\subseteq I_i$，则 $\bigcup_{i\in T}I_i$ 是 R 的理想。

证 记 $I=\bigcup_{i\in T}I_i$，则 $I\supseteq I_i$，$I\neq\varnothing$。$\forall r\in R$，$a,b\in I$，则有 $k,j\in T$ 使 $a\in I_k$，$b\in I_j$，一般地，可设

$$I_k\subseteq I_j,$$

于是 $a,b\in I_j$，因为 I_j 是 R 的理想，故

$$a-b\in I_j\subseteq I,$$

$$ra,ar\in I_j\subseteq I,$$

所以 I 是 R 的理想。

3. 子集生成的理想

设 S 是环 R 的一个子集，记

$$\langle S\rangle=\bigcap_{\substack{I\supseteq S\\ I\text{是}R\text{理想}}}I,$$

称它为子集 S 生成的理想。若 S 为有限集时，设

$$S=\{a_1,\cdots,a_n\},$$

可记 $\langle S\rangle$ 为 $\langle a_1,\cdots,a_n\rangle$。

性质 (1) $\langle S\rangle$ 是理想的交，从而本身也是 R 的理想；

(2) $\langle S\rangle\supseteq S$；

(3) 若 I 是 R 的理想，且 $I\supseteq S$，则 $I\supseteq\langle S\rangle$。

命题 13.2 若 $I_1, \cdots, I_n$ 是环 R 的理想，则

$$\langle I_1\cup\cdots\cup I_n\rangle=I_1+\cdots+I_n。$$

证 $I_i=0+\cdots+I_i+\cdots+0\subseteq I_1+\cdots+I_i+\cdots+I_n$，$i=1, \cdots, n$，所以

$$I_1\cup\cdots\cup I_n\subseteq I_1+\cdots+I_n。$$

因为 $I_1+\cdots+I_n$ 是理想的和，故是 R 的理想。由性质(3)：

$$I_1+\cdots+I_n\supseteq\langle I_1\cup\cdots\cup I_n\rangle，$$

又

$$I_i\subseteq\langle I_1\cup\cdots\cup I_n\rangle, (i=1, \cdots, n),$$

因 $\langle I_1\cup\cdots\cup I_n\rangle$ 是理想，从而是加法子群，它在包含一些元素的同时也包含它们的和，故

$$I_1+\cdots+I_n\subseteq\langle I_1\cup\cdots\cup I_n\rangle,$$

所以

$$I_1+\cdots+I_n=\langle I_1\cup\cdots\cup I_n\rangle。$$

命题 13.3 设 $a_1, \cdots, a_n$ 是环 R 的元素，则

$$\langle a_1, \cdots, a_n\rangle=\langle a_1\rangle+\cdots+\langle a_n\rangle。$$

证 $a_i\in\langle a_1, \cdots, a_n\rangle$，$i=1, \cdots, n$。因 $\langle a_1, \cdots, a_n\rangle$ 是理想，由性质(3)有

$$\langle a_1, \cdots, a_n\rangle\supseteq\langle a_i\rangle, i=1, \cdots, n$$

又由于 $\langle a_1, \cdots, a_n\rangle$ 是理想，从而是加法子群，它在包含一些元素同时也包含它们的和，故

$$\langle a_1, \cdots, a_n\rangle\supseteq\langle a_1\rangle+\cdots+\langle a_n\rangle,$$

又

$$\langle a \rangle = \{na + ra \mid n \in Z, r \in R\}。$$

证 $a \in \langle a \rangle$, $r \in R$,因为$\langle a \rangle$是 R 的理想,故 $ra \in \langle a \rangle$,理想是加法子群,所以 $na \in \langle a \rangle$,从而

$$na + ra \in \langle a \rangle,$$

即

$$\langle a \rangle \supseteq \{na + ra \mid n \in Z, r \in R\} = A。$$

又

$$a = 1 \cdot a + 0 \cdot a \in A,$$
$$(na + ra) - (n_1 a + r_1 a) = (n - n_1)a + (r - r_1)a \in A,$$
$$r'(na + ra) = (nr' + r'r)a \in A,$$

所以 A 是包含$\{a\}$ 的 R 的理想。由性质(3)

$$A \supseteq \langle a \rangle,$$

于是

$$\langle a \rangle = A = \{na + ra \mid n \in Z, r \in R\}。$$

推论 若 R 是含单位元 e 的交换环,$a \in R$,则

$$\langle a \rangle = \{ra \mid r \in R\}。$$

证 因为此时

$$na + ra = (ne) \cdot a + ra = (ne + r)a = r'a。$$

三、同余类环

若 I 是环 R 的一个理想,则加法商群 $R/I = \{a + I \mid a \in R\}$ 关于乘法$(a + I)(b + I) = ab + I$ 成一个环,称为 R 关于 I 的商环,也可称 R 关于 I 的同余类环,仍用记号 R/I 表示。

证 (1) R/I 是加法群。

(2) 乘法和代表元选择无关,即乘法满足运算的要求。因为若

$$a_1 + I = a + I,\ b_1 + I = b + I,$$

则

$$a_1 b_1 \in (a + I)(b + I) \subseteq ab + I,$$

故

$$a_1 b_1 + I = ab + I。$$

又因

$$\begin{aligned}
&[(a + I)(b + I)](c + I) \\
&= (ab + I)(c + I) = (ab)c + I \\
&= a(bc) + I = (a + I)(bc + I) \\
&= (a + I)[(b + I)(c + I)],
\end{aligned}$$

即乘法满足结合律，从而 R/I 是乘法半群。

(3) $(a + I)[(b + I) + (c + I)] = (a + I)[(b + c) + I] = a(b + c) + I = (ab + ac) + I = (ab + I) + (ac + I) = (a + I)(b + I) + (a + I)(c + I)$,

即加法与乘法满足左分配律。同理，可证明其满足右分配律。

由(1),(2),(3)可知，R/I 是一个环。

例 13.8 在整数环 Z 中，理想 $\langle n\rangle = \{kn \mid k \in Z\}$，商环 $Z/\langle n\rangle = \{\bar{0}, \bar{1}, \cdots, \overline{n-1}\}$，即模 n 的同余类环 Z_n。

§3 同态与直和

一、同态的概念

定义 13.8 环 A 到环 A' 的映射 f 称为环同态，如果对$\forall a, b \in R$，均有

$$f(a + b) = f(a) + f(b),$$
$$f(ab) = f(a) \cdot f(b).$$

此时，称环 A 同态于 A'，记成 $A \overset{f}{\sim} A'$。

如果 f 是单射，则称为单同态；如果 f 是满射，则称为满同态；如果 f 是双射，则称为同构，记成 $A \overset{f}{\cong} A'$。

例 13.9 映射 f；$f(n)=\bar{n}$ 是整数环 Z 到同余类环 Z_k 的满同态。

自然同态 若 I 是环 R 的一个理想，则映射 φ:

$$\varphi(a)=\bar{a}=a+I$$

是环 R 到同余类环 R/I 的满同态，称为自然同态。

赋值同态 若 R 是含单位元的环，$c \in R$，且 $\forall a \in R$ 有 $ac=ca$，则映射

$$\sigma(f(x))=f(c)$$

是多项式环 $R[x]$ 到环 R 的同态映射，称为赋值同态。

证 设 $f(x)=\sum\limits_i a_i x^i$, $g(x)=\sum\limits_j b_j x_j$,

$$\begin{aligned}
\sigma(f(x)+g(x)) &= \sigma\left(\sum_i (a_i+b_i)x^i\right) \\
&= \sum_i (a_i+b_i)c^i = \sum_i a_i c^i + \sum_i b_i c^i \\
&= f(c)+g(c)=\sigma(f(x))+\sigma(g(x)), \\
\sigma(f(x)\cdot g(x)) &= \sigma\left(\sum_t \left(\sum_{i+j=t} a_i b_j\right) x^t\right) \\
&= \sum_t \left(\sum_{i+j=t} a_i b_j\right) c^t \\
&= \sum_t \left(\sum_{i+j=t} a_i b_j c^{i+j}\right) \\
&= \sum_{i,j} (a_i c^i)(b_j c^j) \\
&= \sum_i a_i c^i \sum_j b_j c^j \\
&= f(c)\cdot g(c)=\sigma(f(x))\cdot\sigma(g(x))。
\end{aligned}$$

若 f 是环 A 到 A' 的一个同态映射，则 f 也是加法群 A 到 A' 的一个同态映射，核

$$\frac{A+B}{B} \cong \frac{A}{A \cap B}。$$

证 映射 $f: f(a) = a + B$ 是环 A 到环 $\frac{A+B}{B}$ 的满同态，其核为 $A \cap B$，由定理 13.6 即得证明。

定理 13.8 若 σ 是环 R 到 R' 的满同态，A' 是 R' 的理想，记 $A = \sigma^{-1}(A')$，则 A 是 R 的理想，且 $\frac{R}{A} \cong \frac{R'}{A'}$。

证 设 φ 为 R' 到 R'/A' 的自然(满)同态。易得 $\varphi\sigma$ 是环 R 到 R'/A' 的满同态，其核为 $\{x \in R \mid \sigma(x) \in A'\} = A$，由定理 13.6 即得证明。

例 13.11 $Z[\mathrm{i}]/\langle 1+\mathrm{i}\rangle \cong Z_2$。

证 Z_2 的特征为 2，易证映射 $f: f(a+b\mathrm{i}) = \overline{a+b}$ 是 $Z[\mathrm{i}]$ 到 Z_2 的满同态。

若整数 $c \in \langle 1+\mathrm{i}\rangle$，则 $(1+\mathrm{i})$ 的模的平方能整除 c^2，故 $2|c$。反之，若整数 c 是 $2 = (1+\mathrm{i})(1-\mathrm{i})$ 的倍数，显然 $c \in \langle 1+\mathrm{i}\rangle$。即整数 $c \in \langle 1+\mathrm{i}\rangle \Leftrightarrow c$ 是偶数。故有

$$\begin{aligned}
&a+b\mathrm{i} \in \operatorname{Ker} f \Leftrightarrow 2|a+b \Leftrightarrow a+b \in \langle 1+\mathrm{i}\rangle \Leftrightarrow a+b\mathrm{i} \\
&= a+b+b(1+\mathrm{i})\mathrm{i} \in \langle 1+\mathrm{i}\rangle,
\end{aligned}$$

即

$$\operatorname{Ker} f = \langle 1+\mathrm{i}\rangle。$$

由定理 13.6 即得证明。

三、直积与直和

设 $R_1, R_2, \cdots, R_n$ 为 n 个环，$R = R_1 \times R_2 \times \cdots \times R_n$ 为加法群的直积，在其内定义乘法：

$$(a_1, \cdots, a_n) \cdot (b_1, \cdots, b_n) = (a_1b_1, \cdots, a_nb_n),$$

则 R 成一个环，称为环 $R_1, \cdots, R_n$ 的直积，仍记为 $R_1 \times R_2 \times \cdots \times R_n$。

映射

$$\varphi_k: \varphi_k(a_k) = (0, \cdots, 0, a_k, 0, \cdots, 0)$$

是环 R_k 到环 R 的单同态。记 $\bar{R}_k = \varphi_k(R_k)$，从而 $\bar{R}_k \cong R_k$，$\bar{R}_k$ 是 R 的理想，因

$$(a_1, \cdots, a_n) = (a_1, 0, \cdots, 0) + \cdots + (0, \cdots, 0, a_n),$$

即 R 中元可以惟一表成 $\bar{R}_1, \cdots, \bar{R}_n$ 中元之和。

定义 13.9 环 R 称为是其理想 $I_1, \cdots, I_n$ 的内直和，如果 R 中任意元素都可以惟一表成 $I_1, \cdots, I_n$ 中元素之和，此时记为 $R = I_1 \oplus \cdots \oplus I_n$。

上段中，$R = \bar{R}_1 \oplus \cdots \oplus \bar{R}_n$。反之有如下定理：

定理 13.9 若环 R 是其理想 $I_1, \cdots, I_n$ 的内直和，则 $R \cong I_1 \times I_2 \times \cdots \times I_n$。

说明 在群论中，已证明了：表法惟一$\Leftrightarrow 0$ 表法惟一$\Leftrightarrow I_i \cap \left(\sum_{j \neq i} I_j\right) = \{0\}$。从而定理中的“表法惟一”用等价的条件代替即得相应的推论。

证 $i \neq j$，$I_i \cdot I_j \subseteq I_i \cap I_j = \{0\}$。在群论中，已经证明了映射

$$f: f[(a_1, \cdots, a_n)] = a_1 + \cdots + a_n$$

是加群 $I_1 \times \cdots \times I_n$ 到 R 的同构。又

$$\begin{aligned}
& f[(a_1, \cdots, a_n)] \cdot f[(b_1, \cdots, b_n)] \\
&= (a_1 + \cdots + a_n) \cdot (b_1 + \cdots + b_n) \\
&= \sum_{i=1}^{n} a_i b_i + \sum_{i \neq j} a_i b_j \\
&= \sum_{i=1}^{n} a_i b_i = f[(a_1 b_1, \cdots, a_n b_n)] \\
&= f[(a_1, \cdots, a_n)(b_1, \cdots, b_n)],
\end{aligned}$$

故映射 f 是环同构。

例 13.12 在 Z_6 中，记 $I_1 = \langle \bar{2} \rangle$，$I_2 = \langle \bar{3} \rangle$，则 $Z_6 = I_1 \oplus I_2$。

证 $I_1 \cap I_2 = \{0\}$，

$$|I_1+I_2|=\frac{|I_1||I_2|}{|I_1\cap I_2|}=6=|Z_6|,$$

故 $Z_6=I_1+I_2=I_1\oplus I_2$。

例 13.13 若 I_1 是交换环 R 的理想，I_1 有单位元 e，即

$$\forall x\in I_1,\text{有 } xe=x,$$

则存在 R 的一个理想 I_2 使

$$R=I_1\oplus I_2。$$

证 $\forall a\in R$，$a=ae+(a-ae)$。记 $I_2=\{a-ae|a\in R\}$。因为

$$(a-ae)-(b-be)=(a-b)-(a-b)e\in I_2,$$

又因

$$r(a-ae)=ra-rae\in I_2,$$

所以 I_2 是 R 的理想，显然，$R=I_1+I_2$。

对任意 $x\in I_1\cap I_2$，有 $x=xe$，且存在 $a\in R$ 使 $x=a-ae$，此时

$$x=xe=(a-ae)e=0.$$

从而

$$I_1\cap I_2=\{0\},$$

故

$$R=I_1\oplus I_2。$$

例 13.14 若环 R 的特征为 $n=p_1^{\alpha_1},\cdots,p_s^{\alpha_s}$，这里 $p_1,\cdots,p_s$ 是互异素数，则存在 R 的理想 $I_1,\cdots,I_s$ 使 I_i 的特征为 $p_i^{\alpha_i}(i=1,\cdots,s)$，且 $R=I_1\oplus I_2\oplus\cdots\oplus I_s$。

证 记 $n_i=\dfrac{n}{p_i^{\alpha_i}}$，$i=1,\cdots,s$。设

$$I_i=n_iR=\{n_ix|x\in R\},$$

易得，I_i 是 R 的理想，且 I_i 的特征为

$$\frac{n}{n_i} = p_i^{\alpha_i},\ (i = 1, \cdots, s)。$$

因为$(n_1, \cdots, n_s) = 1$，所以存在 $k_1, \cdots, k_s \in Z$ 使

$$\sum_{i=1}^{s} k_i n_i = 1。$$

故任意 $a \in R$，有

$$a = \left(\sum_{i=1}^{s} k_i n_i\right) a = \sum_{i=1}^{s} n_i (k_i a) \in I_1 + \cdots + I_s。$$

又，对任意 $x \in I_i \cap \left(\sum_{j \neq i} I_j\right)$，有 $p_i^{\alpha_i} x = 0$ 及 $n_i x = 0$，因为$(p_i^{\alpha_i}, n_i) = 1$，不难推出

$$x = 1x = (p_i^{\alpha_i}, n_i)x = 0,$$

即

$$I_i \cap \left(\sum_{j \neq i} I_j\right) = \{0\}\ (i = 1, \cdots, s).$$

故

$$R = I_1 \oplus \cdots \oplus I_s。$$

§4 商域与分式环

体内非零元都是可逆元，这给体内元素的运算带来了极大的方便（如满足乘法消去律，$a \neq 0$ 时，方程 $ax = b$ 与 $ya = b$ 可解等）。这就会产生一个想法：给了一个环，能否增加一些元素，使它成为一个体？例如，整数环中添加了所有两个非零整数的商，就成了体（有理数域）。这个方法是否带有普遍性？由于非交换环的情形异常复杂，本节中，仅就交换环的情形进行讨论。下面，先给出“嵌入”的概念。

定义 13.10 称环 R 可以嵌入到环 A 中去，如果存在 R 到 A 的（环）单同态。此时，R 同构于其同态像，A 有和 R 同构的子环。

一、体内子环的商体（域）

命题 体的任意多个子体的交仍是子体。

证 子体是加法子群，故其交仍是加法子群；子体的非零元集是乘法子群，故其交仍是乘法子群。加、乘法分配律是满足的，所以子体的交仍是子体。

定义 13.11 若 K 是体，R 是 K 的非零交换子环。记

$$P = \bigcap_{\substack{F\text{是}K\text{的子体} \\ F \supseteq R}} F,$$

称为环 R 的商体。

环 R 的商体 P 有如下性质：

(1) $P \supseteq R$；

(2) 对任意 F，F 是 K 的子体，若 $F \supseteq R$，则 $F \supseteq P$；

(3) P 是域，且 $P = \{ab^{-1} \mid a, b \in R, b \neq 0\}$，也称 P 为 R 的商域。

证 P 是 K 的子体的交，故仍是体，(1)，(2) 显然成立。$a, b \in R$，$ab = ba$，两边同时左、右乘 b^{-1} 得 $b^{-1}a = ab^{-1}$，故可记 $b^{-1}a = ab^{-1}$ 为 $\frac{a}{b}$。令

$$P_1 = \left\{ \frac{a}{b} \,\middle|\, a, b \in R, b \neq 0 \right\},$$

显然 $P \supseteq P_1$。又

$$\frac{a_1}{b_1}, \frac{a_2}{b_2} \in P_1 \Rightarrow \frac{a_1}{b_1} - \frac{a_2}{b_2} = \frac{a_1 b_2}{b_1 b_2} - \frac{a_2 b_1}{b_1 b_2} = \frac{a_1 b_2 - a_2 b_1}{b_1 b_2} \in P_1;$$

又 $a_1 \neq 0$ 时

$$\left(\frac{a_1}{b_1}\right)^{-1}\left(\frac{a_2}{b_2}\right) = \frac{b_1 a_2}{a_1 b_2} \in P_1,$$

所以 P_1 是包含 R 的 K 的子体，故 $P_1 \supseteq P$，从而 $P = P_1$。

$\forall \frac{a_1}{b_1}, \frac{a_2}{b_2} \in P$，有

$$\frac{a_1}{b_1} \cdot \frac{a_2}{b_2} = \frac{a_1 a_2}{b_1 b_2} = \frac{a_2 a_1}{b_2 b_1} = \frac{a_2}{b_2} \cdot \frac{a_1}{b_1},$$

故 P 是域，(3) 得证。

由以上证明可知，交换环 R 的商域 P 为

$$\left\{ \frac{a}{b} \middle| a, b \in R, b \neq 0 \right\}.$$

P 中元素相等和加、乘法运算如下：

$$\frac{a}{b} = \frac{c}{d} \Leftrightarrow ad = bc;$$

$$\frac{a}{b} + \frac{c}{d} = \frac{ad + bc}{bd};$$

$$\frac{a}{b} \cdot \frac{c}{d} = \frac{ac}{bd}。$$

即 P 中元素及其加、乘法运算完全由 R 中元素及其运算确定，即商域完全由环 R 本身确定。若交换环 R 能嵌入两个体中，那么得到的两个商域必同构。

二、可嵌入体的交换环

若非零交换环 R 能嵌入体 K，因为 K 没有零因子，故 R 也没有零因子，下面证明这个条件也是充分的。

定理 13.10 非零交换环 R 能嵌入体的充要条件是 R 没有零因子。

证 充分性的证明 若 R 没有零因子。在 $R \times R$ 的子集

$$S = \{(a, b) \in R \times R \mid b \neq 0\}$$

中定义二元关系"$\sim$"：

$$(a,b) \sim (c,d) \Leftrightarrow ad = bc,$$

这个关系显然满足反身性和对称性。又

$$(a,b) \sim (c,d), (c,d) \sim (e,f),$$

得

$$ad = bc,\ cf = de,$$

故

$$adf = bcf = bde,$$

由于 R 无零因子，可交换，从而 $af = be$，即 $(a,b) \sim (e,f)$，即上述关系也满足传递性，因而这个关系是等价关系。

把 (a,b) 所在的类记成 $\frac{a}{b}$，用 F 表示商集合 $S/\sim$，即

$$F = \left\{ \frac{a}{b} \middle| a,b \in R,\ b \neq 0 \right\},$$

$$\frac{a}{b} = \frac{c}{d} \Leftrightarrow ad = bc。$$

用 R 中元素的运算，来定义 F 中元素的运算：

$$\frac{a}{b} + \frac{c}{d} = \frac{ad + bc}{bd},\ \frac{a}{b} \cdot \frac{c}{d} = \frac{ac}{bd}。$$

因 R 不含零因子，故 $bd \neq 0$，从而和与积都 $\in F$。下面证明这样的定义符合运算规定的要求，即运算的结果是惟一的，和代表元的选择无关。

若

$$\frac{a_1}{b_1} = \frac{a}{b},\ \frac{c_1}{d_1} = \frac{c}{d},$$

即

$$a_1 b = a b_1,\ c_1 d = c d_1。$$

于是

$$a_1c_1bd = acb_1d_1,$$

故

$$\frac{a_1c_1}{b_1d_1} = \frac{ac}{bd}。$$

又

$$ab_1dd_1 = a_1bdd_1,$$

及

$$cd_1bb_1 = c_1dbb_1,$$

所以

$$(ad + bc)b_1d_1 = (a_1d_1 + b_1c_1)bd,$$

从而

$$\frac{ad + bc}{bd} = \frac{a_1d_1 + b_1c_1}{b_1d_1}。$$

易验证，F 对上述运算成一域，零元素为 $\frac{0}{b}$，$\frac{a}{b}$ 的负元为 $\frac{-a}{b}$，单位元为 $\frac{b}{b}$，$\frac{a}{b}$（$a \neq 0$）的逆元为 $\frac{b}{a}$。

易见，映射 f：$f(a) = \frac{ab_0}{b_0}$（这里 b_0 是 R 的一个固定非零元）是环 R 到域 F 的单同态，即 R 已嵌入到域 F 中，故定理得证。

例如，整数环 Z 的商域为有理数域 Q，高斯整环 $Z[\mathrm{i}]$ 的商域为

$$Q[\mathrm{i}] = \{a + b\mathrm{i} \mid a, b \in Q\}。$$

三、分式环

若设 R 是交换环，U 是 R 中所有非零因子所成集合。若 $U \neq \varnothing$，在 $R \times U$ 中定义关系：

$$(a,u) \sim (b,v) \Leftrightarrow av = bu,$$

这是一个等价关系，把等价类(a,u)记成$\frac{a}{u}$，在商集合

$$A = \left\{ \frac{a}{u} \,\middle|\, a \in R,\ u \in U \right\}$$

中仿定理3.14证明中的方法定义加法和乘法（注意U是乘法半群），则A成一环，称为R的分式环。R可嵌入A中，R中非零因子在A中皆可逆。A的单位元为$\frac{u}{u}$；v可看成A中的元$\frac{vu}{u}$，其逆为$\frac{u}{vu}$ $(v\in U)$，故得下面的定理：

定理 13.11 交换环R可嵌入到它的分式环A中，R中非零因子在A中有逆元。若R有非零因子，则A是含单位元的环。

例 13.15 求证：同余类环Z_n的分式环为自身。

证 设$\bar{i}$为Z_n中的元素，$(i, n) = d$，因为

$$\bar{i} \cdot \left(\overline{\frac{n}{d}} \right) = \bar{0},$$

故$(i, n) \neq 1$，则$\bar{i}$是零因子。$(i, n) = 1$，存在$a, b \in Z$，使$ai + bn = 1$，于是$\bar{a}\,\bar{i} = \bar{1}$，$\bar{i}$可逆，不是零因子。所以

$$\begin{aligned}(Z)_n\ \text{分式环} &= \left\{ \frac{\bar{j}}{\bar{i}} \,\middle|\, \bar{i}, \bar{j} \in Z_n,\ \begin{matrix}\bar{i}\ \text{不是零因子}\\ \text{因而可逆}\end{matrix} \right\} \\ &= \{ \bar{j} \cdot (\bar{i})^{-1} \mid \bar{j},\ (\bar{i})^{-1} \in Z_n \} = Z_n.\end{aligned}$$

四、可添加“单位元”的环

定理 13.12 任意环R都可嵌入到某个含单位元的环S中。环S（不是惟一的）的特征可以取成是零或者等于R的特征。

证 在积集合$Z \times R$中定义加、乘法如下：

$$(k_1, a_1) + (k_2, a_2) = (k_1 + k_2,\ a_1 + a_2),$$

$$(k_1, a_1) \cdot (k_2, a_2) = (k_1 \cdot k_2,\ k_1 a_2 + k_2 a_1 + a_1 a_2),$$

则 $S = Z \times R$ 成环，S 含单位元 $e = (1,0)$。S 的特征是零。映射 f：$f(a) = (0, a)$ 是环 R 到 S 的单同态，即 R 嵌入到含单位元 e 的特征为 0 的环 S 中。

若环 R 的特征为 n，在积集合 $S_1 = Z_n \times R$ 中定义加法和乘法：

$$(\bar{k}_1, a_1) + (\bar{k}_2, a_2) = (\overline{k_1 + k_2}, a_1 + a_2),$$
$$(\bar{k}_1, a_1)(\bar{k}_2, a_1) = (\overline{k_1 k_2}, k_1 a_2 + k_2 a_1 + a_1 a_2),$$

则 S_1 成环，S_1 的单位元为 $(\bar{1}, 0)$，其特征为 n。同样，映射 f：$f(a) = (\bar{0}, a)$ 是 R 到 S_1 的单同态，即 R 嵌入到含单位元且和它有相同特征的环 S_1 中。

§5　惟一因子分解整环

众所周知，在整数环 Z 中，不等于 1 的正整数都能分解成素数的乘积，而且，除了因子次序外，分解是惟一的。同样域 F 上多项式环 $F[x]$中，每个次数大于等于 1 的多项式都能分解成不可约多项式的乘积，而且，除了因子次序和零次因子的差别外，分解也是惟一的。本节中，将在整环中引进因子分解的概念，找出整环中因子分解惟一定理成立的一些条件，并给出具体例子。

一、交换环的素理想与极大理想

定义 13.12　交换环 R 的理想 P 称为素理想，如果$\forall a, b \in R$，$a \cdot b \in P$必 $a \in P$ 或 $b \in P$。

例 13.16　偶数全体 $2Z = \{2k \mid k \in Z\}$ 是整数环 Z 的素理想，这是因为两整数的积是偶数必有一个是偶数。

定理 13.13　设 P 是交换环 R 的理想，则 P 是素理想$\Leftrightarrow$同余类环 R/P 没有零因子。

证　P 是素理想 $\Leftrightarrow a, b \in R, a \cdot b \in P$　必 $a \in P$ 或 $b \in P$，即在 R/P 中，$\bar{a} \cdot \bar{b} = \bar{0}$　必$\bar{a} = \bar{0}$ 或$\bar{b} = \bar{0}$，即同余类环 R/P 无零因子。

例 13.17　证明：主理想$\langle x \rangle = \{x f(x) \mid f(x) \in Z[x]\}$ 是整系数

多项式环 $Z[x]$的素理想。

证 映射 σ: $\sigma(f(x)) = f(0)$ 是 $Z[x]$ 到 Z 的赋值满同态，同态核

$$\mathrm{Ker}\,\sigma = \{f(x) \in Z[x] \mid f(0) = 0\} = \langle x \rangle,$$

由同态基本定理

$$\frac{Z[x]}{\langle x \rangle} \cong Z,$$

而 Z 是整环，无零因子，故$\langle x \rangle$是 $Z[x]$的素理想。

定义 13.13 交换环 R 的真理想 M 称为极大理想，如果 H 是 R 的理想且 $H \supset M$，则 $H = R$，即 R 没有包含 M 且不等于 M 的真理想。

例 13.16 的理想$\langle Z \rangle$是 Z 的极大理想，但例 13.17 的理想$\langle x \rangle$不是 $Z[x]$ 的极大理想，因为 $\langle x \rangle \subset \langle x, 2 \rangle \subset Z[x]$。

命题 13.6 含单位元的交换环一定有极大理想。

证 设 R 是含单位元 e 的环，令 $S = \{H \mid H$ 是 R 的理想，$e \overline{\in} H\}$。因$\{0\} \in S$，故 $S \neq \varnothing$。S 对集合的包含关系成偏序集，S 的任一非空有序子集在 S 中有上界(并集)。由 Zorn 引理，S 有极大元，即 R 的极大理想。

定理 13.14 若 R 是含单位元的交换环，M 是 R 的理想，则 M 是 R 的极大理想 $\Leftrightarrow$ 商环 R/M 是域。

证 $\Rightarrow$：若 M 是 R 的极大理想，$\bar{a}$ 是同余类环 R/M 的非零元，即

$$a \overline{\in} M,$$

因此 R 的理想

$$H = M + \langle a \rangle \supset M,$$

由于 M 是 R 的极大理想，必 $H = R$，故 $1 \in H$，即存在 $m \in M$，$b \in R$ 使 $1 = m + ba$，从而$\bar{b}\bar{a} = \bar{1}$，这表明交换环 R/M 中非零元都可逆，从而是域。

$\Leftarrow$：若同余类环 R/M 是域，H 是 R 的理想，$H \supset M$，取 $a \in H - M$，在同余类域 R/M 中，$\bar{a} \neq \bar{0}$，因而$\forall c \in R$，必存在$\bar{b}((\bar{a})^{-1}\bar{c})$使

$$\bar{b}\ \bar{a}=\bar{c},$$

于是

$$c-ba\in M,$$

$$c\in H,$$

故 $H=R$,M 为极大理想。

在充分性的证明中,并未用到 R 有单位元的假设。故得下面推论:

推论 若 M 是交换环 R 的理想,且同余类环 R/M 是域,则 M 是 R 的极大理想。

例 13.18 求证:$\langle x,y\rangle$ 是 $Q[x,y]$ 的极大理想。

证 映射

$$\sigma(f(x,y))=f(0,0)$$

是环 $Q[x,y]$到 Q 的赋值满同态。同态核是常数项为零的多项式全体:

$$\mathrm{Ker}\ \sigma=\langle x,y\rangle。$$

由同态基本定理

$$\frac{Q[x,y]}{\langle x,y\rangle}\cong Q,$$

因 Q 是域,故$\langle x,y\rangle$ 是多项式环 $Q[x,y]$ 的极大理想。

例 13.19 求证:$\langle x,\ p\rangle$ 是 $Z[x]$ 的极大理想,这里 p 是素数。

证 $Z[x]$ 到 Z_p 的映射

$$\sigma[f(x)]=\overline{f(0)}$$

是赋值(满)同态 ε 与自然(满)同态 φ 的积:

$$\varepsilon(f(x))=f(0),\ \varphi(n)=\bar{n},$$

故 σ 是满同态。又

$$f(x)=q(x)\cdot x+f(0),$$

因而

$$\bar{f}(\bar{0})=\bar{0} \Leftrightarrow p \mid f(0) \Leftrightarrow f(x) \in \langle x, p\rangle,$$

即

$$\operatorname{Ker} \sigma=\langle x, p\rangle,$$

于是

$$\frac{Z[x]}{\langle x, p\rangle} \cong Z_p,$$

因为 Z_p 是域，故 $\langle x,p\rangle$ 是 $Z[x]$ 的极大理想。

二、整环内因子分解的概念

定义 13.14 设 R 是整环，$a,b \in R$，称 a 是 b 的因子，或 a 整除 b 记成 $a|b$，如果存在 $c \in R$ 使 $b=ac$。

推论 a 是整环 R 的元素，则 $a|1 \Leftrightarrow a$ 可逆。

定义 13.15 R 是整环，$a,b \in R$，称 a 和 b 等价，如果 $a|b$ 且 $b|a$，记成 $a \sim b$。

推论 a 和 b 是整环 R 的非零元素，则 $a \sim b \Leftrightarrow$ 存在 R 的可逆元 c 使 $a=bc$。

证 $\Leftarrow$ 显然。

$\Rightarrow a \sim b$，存在 $c \in R$ 使 $a=bc$。又 $a|b$，故存在 $d \in R$ 使 $b=ad$ 代入得 $a=adc$，消去得 $1=dc$，即 c 可逆。

推论 设 a 是整环 R 的元素，则 $a \sim 1 \Leftrightarrow a|1 \Leftrightarrow a$ 可逆。

定义 13.16 整环 R 的非零不可逆元 a 称为素元，如果由 $a|bc$ $(b, c \in R)$，必有 $a|b$ 或 $a|c$。

定义 13.17 整环 R 的非零不可逆元 a 称为既约元，如果 $\forall b \in R$，$b|a$，则 $b \sim 1$ 或 $b \sim a$。

命题 13.7 若素元 $p|a_1 \cdots a_n$，则存在 i $(1 \leqslant i \leqslant n)$ 使 $p|a_i$（证略）。

命题 13.8 整环的素元必是既约元。

证 设 p 是素元，$a|p$，即 $p=ab$，故 $p|ab$。因 p 是素元，故 $p|a$ 或 $p|b$。

若 $p|a$ 则 $a\sim p$。若 $p|b$，$b=pc$，$p=ab=apc$，消去 p，得 $ac=1$，故 $a\sim 1$。即得 p 为既约元。

定义 13.18 若 X 是整环 R 的非空子集，元素 $d\in R$ 为 X 的最大公因子。如果

(1) $d|a$，$\forall a\in X$；

(2) 若 $c|a$ $(\forall a\in X)\Rightarrow c|d$。

最大公因子并不一定存在。例如在偶数环(即偶数全体对通常数的加、乘作成的环)中 2 没有因子，从而 2 和 4 没有最大公因子，即使有最大公因子，它也不一定是惟一的，但由(2)可知，任意两个最大公因子是等价的。

三、惟一因子分解整环(F 环)

定义 13.19 整环 R 称为惟一因子分解整环，简称 F 环，如果：

(1) R 中每个非零不可逆元 a 均可写成 R 中有限个既约元 p_1，p_2，…，p_n 的乘积，即 $a=p_1\cdots p_n$；

(2) 若 $a=p_1\cdots p_n=q_1\cdots q_s$；这里 q_1，…，q_s 也是既约元。则 $n=s$，且存在 n 阶置换 σ 使 $p_i\sim q_{\sigma(i)}(i=1,\cdots,n)$。

如果 R 是 F 环，a 和 b 是 R 的非零不可逆元，若把等价的既约元并在一起，则可表 a 和 b 为两两不等价的既约元的幂的积，即 $a\sim\prod\limits_{i=1}^{n}p_i^{\alpha_i}$，$b\sim\prod\limits_{i=1}^{n}p_i^{\beta_i}$，这里 α_i 和 β_i 均为大于等于 0 的整数。则有

(1) $a|b\Leftrightarrow\alpha_i\leqslant\beta_i(i=1,\cdots,n)$，从而既约元是素元；

(2) (a,b) 存在且 $(a,b)\sim\prod\limits_{i=1}^{n}p_i^{\min(\alpha_i,\beta_i)}$。类似可求得 n 个元素的最大公因子。

不是所有的整环都是 F 环，如 $Z[\sqrt{-5}]$。

$$Z[\sqrt{-5}]=\{a+b\sqrt{-5}\,|\,a,b\in Z\}$$

它的可逆元为 ±1。

$$9=3\times3=(2+\sqrt{-5})(2-\sqrt{-5}),$$

易验证，3，$2\pm\sqrt{-5}$都是 $Z[\sqrt{-5}]$的既约元，于是 $Z[\sqrt{-5}]$中分解不惟一，因而不是 F 环。

四、主理想整环

定义 13.20 整环 R 称为主理想整环，如果 R 的任一理想都是主理想。

例 13.20 整数环 Z 是主理想整环，这是由于：Z 的理想是加法子群，而 Z 是加法循环群，其子群也是循环群。

例 13.21 若 K 是域，则 $K[x]$ 是主理想整环。

证 设 I 是 $K[x]$ 的理想，$I=\{0\}$是主理想。若 $I\neq\{0\}$，设 $p(x)$ 是 I 中次数最低的一个非零多项式，则

$$I\supseteq\langle p(x)\rangle,\quad 又\ \forall f(x)\in I,$$

作带余除法：

$$f(x)=q(x)p(x)+r(x),$$

其中 $r(x)=0$ 或 $r(x)$的次数$<p(x)$的次数。因

$$r(x)=f(x)-q(x)p(x)\in I,$$

必

$$r(x)=0,$$

即

$$I=\langle p(x)\rangle,$$

故 $K[x]$是主理想整环。

例 13.22 $Z[\mathrm{i}]$ 是主理想整环。

证 设 I 是 $Z[\mathrm{i}]$ 的非零理想。设 $a+b\mathrm{i}$ 是 I 中模平方最小的非零数。

$\forall\, c+d\mathrm{i}\in I$，在复数内作除法：

$$\frac{c+d\mathrm{i}}{a+b\mathrm{i}}=u+v\mathrm{i}\ (u,v\text{ 是实数}),$$

则存在 $u_1, v_1 \in Z$,使 $|u-u_1| \leqslant \frac{1}{2}$, $|v_1-v| \leqslant \frac{1}{2}$。此时,$I$ 中元

$$\begin{aligned}&(c+d\mathrm{i})-(a+b\mathrm{i})(u_1+v_1\mathrm{i})\\&=(a+b\mathrm{i})\left[\frac{c+d\mathrm{i}}{a+b\mathrm{i}}-(u_1+v_1\mathrm{i})\right]\\&=(a+b\mathrm{i})[(u-u_1)+(v-v_1)\mathrm{i}],\end{aligned}$$

其模平方显然小于 $a+b\mathrm{i}$ 的模的平方,故必为 0,即 $c+d\in\langle a+b\mathrm{i}\rangle$。从而

$$I=\langle a+b\mathrm{i}\rangle,$$

结论得证。

若 R 是主理想整环,则有

(1) $a|b \Leftrightarrow \langle a\rangle \supseteq \langle b\rangle$;

(2) $a\sim 1 \Leftrightarrow \langle a\rangle = R$;

(3) $a\sim b \Leftrightarrow \langle a\rangle = \langle b\rangle$;

(4) 非零不可逆元 a 是素元$\Leftrightarrow\langle a\rangle$是素理想;

(5) 非零不可逆元 a 是既约元$\Leftrightarrow\langle a\rangle$是极大理想。

如果$\langle a\rangle$是极大理想,则 $R/\langle a\rangle$是域,域是整环,从而$\langle a\rangle$是素理想。故有(前已证素元是既约元):

命题 13.9 主理想整环中,元素 a 是素元$\Leftrightarrow a$ 是既约元。

引理 13.1 若 R 是主理想整环,$\langle a_1\rangle\subseteq\langle a_2\rangle\cdots$ 是 R 中的理想链,则存在 $n\in\mathbf{Z}^+$ 使 $i\geqslant n$ 时,$\langle a_i\rangle=\langle a_n\rangle$。

证 令

$$A=\bigcup_{i\geqslant 1}\langle a_i\rangle。$$

易验证 A 是 R 的理想,故存在 $a\in R$ 使

$$A=\langle a\rangle。a\in A,$$

从而存在 $n\in\mathbf{Z}^+$ 使

$$a \in \langle a_n \rangle,$$

故

$$A = \langle a_n \rangle_\circ\ \forall\, i \geqslant n,\ \langle a_n \rangle \subseteq \langle a_i \rangle \subseteq A = \langle a_n \rangle,$$

故

$$\langle a_n \rangle = \langle a_i \rangle_\circ$$

定理 13.15 主理想整环是 F 环。

证 设 R 是主理想整环。令 $T = \{a \in R \mid a$ 非零,不可逆,且 a 不能表成 R 中有限个既约元的乘积。$\}$ 要证 $T = \varnothing$。若 $T \neq \varnothing$,令

$$S = \{\langle a \rangle \mid a \in T\},$$

则 S 关于集合的包含关系成一偏序集。由引理,S 的非空有序子集有上界。根据 Zorn 引理,S 有极大元 $\langle m \rangle$, $m \in T$。因 m 不是既约元,m 有因子 a, $a \not\sim 1$, $a \not\sim m$。此时 $m = ab$, 则 $b \not\sim m$, $b \not\sim 1$。故

$$\langle m \rangle \subset \langle a \rangle\ ,\ \langle m \rangle \subset \langle b \rangle_\circ$$

由 $\langle m \rangle$ 的极大性,知 a, $b \notin T$,即 a,b 可表成有限个既约元的乘积,从而 $m = a \cdot b$ 也能表成有限个既约元的乘积,矛盾。故 $T = \varnothing$,即 R 中每个非零不可逆元都能表成有限个既约元的乘积。

若 $a = p_1 \cdots p_n = q_1 \cdots q_s$。这里 p_i, $q_j (i = 1, \cdots, n, j = 1, \cdots, s)$ 都是既约元,p_1 是既约元。故是素元, $p_1 \mid a = q_1 \cdots q_s \Rightarrow$ 存在 i_1 使 $p_1 \mid q_{i_1}$,但 q_{i_1}是既约元,从而 $p_1 \sim q_{i_1}$,对 n 用归纳法,即得惟一性的证明。

推论 Z, $K[x]$, $Z[i]$是 F 环。这里 K 是域。

命题 13.10 若 R 是主理想整环,$a_1, \cdots, a_n \in R$,则 $a_1, \cdots, a_n$ 的最大公因子存在且能表成

$$\sum_{i=1}^{n} k_i a_i (k_i \in R,\ i = 1, \cdots, n)_\circ$$

证 R 的理想$\langle a_1, \cdots, a_n \rangle$是主理想,故存在 $d \in R$ 使

$$\langle d \rangle = \langle a_1, \cdots, a_n \rangle_\circ$$

因 $a_i \in \langle d \rangle$，故

$$d \mid a_i (i = 1, \cdots, n)。$$

又

$$d \in \langle a_1, \cdots, a_n \rangle = \langle a_1 \rangle + \cdots + \langle a_n \rangle,$$

即存在 $k_1, \cdots, k_n \in R$ 使

$$d = k_1 a_1 + \cdots + k_n a_n.$$

$$\forall c \in R, c \mid a_i (i = 1, \cdots, n)。$$

则

$$c \mid \sum_{i=1}^{n} k_i a_i = d,$$

故 d 是 $a_1, \cdots, a_n$ 的最大公因子。

五、Gauss 整环 Z[i]

Gauss 整环是主理想整环，因而是 F 环。我们往往通过模平方的关系把 Z[i]中的整除问题化为熟悉的整数环 Z 中的整除问题。

引理 13.2 $Z[i]$中，$a + bi \sim 1 \Leftrightarrow a^2 + b^2 = 1$，从而 $Z[i]$的可逆元为 ± 1，$\pm i$。

证 $a + bi \sim 1 \Leftrightarrow a + bi \mid 1 \Leftrightarrow$ 存在 $c + di \in Z[i]$ 使

$$(a + bi)(c + di) = 1 \Leftrightarrow a^2 + b^2 \mid 1 \Leftrightarrow a^2 + b^2 = 1。$$

引理 13.3 若 $Z[i]$中有 $x + yi \mid a + bi$，则必 $x^2 + y^2 \mid a^2 + b^2$ 在 Z 中成立。

证 Z[i]中，$x + yi \mid a + bi$，即存在 $c + di \in Z[i]$ 使

$$(a + bi) = (x + yi)(c + di),$$

两边取模平方，得

$$a^2 + b^2 = (x^2 + y^2)(c^2 + d^2),$$

即得结论。

推论 若 $Z[\mathrm{i}]$中，$(x+y\mathrm{i})\ (a+b\mathrm{i})$ 且 $x^2+y^2=a^2+b^2$，则

$$(x+y\mathrm{i})\sim(a+b\mathrm{i})$$

证 由引理 13.3 的证明，得 $c^2+d^2=1$，于是 $c+d\mathrm{i}\sim 1$，故

$$(x+y\mathrm{i})\sim(a+b\mathrm{i})。$$

定理 13.16 若 $a+b\mathrm{i}$ 为 $Z[\mathrm{i}]$ 的素元(也是既约元)，必

(1) a^2+b^2 为素数 p；或

(2) $a+b\mathrm{i}=\pm p,\ \pm p\mathrm{i}$，此时 $x^2+y^2=p$ 在 Z 中无解，p 为素数。反之亦对。

证 在 $Z[\mathrm{i}]$中，素元

$$(a+b\mathrm{i})\,|\,(a+b\mathrm{i})(a-b\mathrm{i})=a^2+b^2,$$

正整数 a^2+b^2 可表为素数的积，于是 $Z[\mathrm{i}]$ 的素元$(a+b\mathrm{i})$ 必能整除某个素数 p：

$$(a+b\mathrm{i})\,|\,p,$$

由引理 13.3 得

$$(a^2+b^2)\,|\,p^2,$$

因 $a+b\mathrm{i}$ 是素元，不可逆，故必 $a^2+b^2=p$ 或 p^2。

(1) 若 $a^2+b^2=p$，则 $a+b\mathrm{i}$ 是 $Z[\mathrm{i}]$ 的既约元，从而是素元。这是因为：若

$$x+y\mathrm{i}\,|\,a+b\mathrm{i},$$

必

$$(x^2+y^2)\,|\,(a^2+b^2)=p,$$

于是

1) $x^2+y^2=1,\ x+y\mathrm{i}\sim 1$,

2) $x^2+y^2=p$,

由引理 13.3 的推论，$x+yi\sim a+bi$，由 1)，2) 得，$a+bi$ 是 $Z[i]$ 的素元。

(2) $a^2+b^2=p^2$

1) $x^2+y^2=p$ 在 Z 中有解 x,y，此时若 $a+bi$ 为 $Z[i]$ 素元，由素元

$$a+bi\mid(a+bi)(a-bi)=a^2+b^2=p^2=(x+yi)^2(x-yi)^2$$

得

$$(a+bi)\mid(x+yi)\text{ 或 }(x-yi),$$

因而

$$p^2=(a^2+b^2)\mid(x^2+y^2)=p,$$

这不可能，故 $a+bi$ 不是 $Z[i]$ 的素元。

2) 若 $x^2+y^2=p$ 在 Z 中无解。此时，$a-bi$ 为 $Z[i]$ 的既约元，这是因为：若

$$(x+yi)\mid(a+bi),$$

必

$$(x^2+y^2)\mid(a^2+b^2)=p^2,$$

因而

$$x^2+y^2=1\text{ 或 }p^2,$$

于是必

$$x+yi\sim 1\text{ 或 }\sim a+bi。$$

又因为 $Z[i]$ 中

$$(a+bi)\mid(a+bi)(a-bi)=a^2+b^2=p\cdot p,$$

故必 $(a+bi)\mid p$，但 $a^2+b^2=p^2$，由引理 13.3 的推论：

$$a+bi\sim p,$$

所以

$$a + b\mathrm{i} = \pm p, \pm p\mathrm{i}。$$

例如，$\pm 1 \pm \mathrm{i}, \pm 3, \pm 3\mathrm{i}, \pm 2 \pm \mathrm{i}, \pm 1 \pm 2\mathrm{i}, \cdots$，都是 $Z[\mathrm{i}]$ 的素元。在 $Z[\mathrm{i}]$ 中，表 3 900 为既约元的积：

$$\begin{aligned}3\,900 &= 3 \times 13 \times 2^2 \times 5^2 \\ &= 3(3+2\mathrm{i})(3-2\mathrm{i})(1+\mathrm{i})^2(1-\mathrm{i})^2(2+\mathrm{i})^2(2-\mathrm{i})^2\end{aligned}$$

§6 多项式环

本节中，我们将证明 F 环上的多项式环仍是 F 环，从而 $Z[x]$ 及 F 环上的多元多项式环是 F 环。

一、带余除法

定义 13.21 $f(x) = \sum_{i=0}^{n} a_i x^i \in R[x]$，则元素 $a_i \in$ 环 R $(i = 0, 1, \cdots, n)$ 称为 $f(x)$ 系数，a_0 为常数项；若 $a_n \neq 0$，则称 a_n 为 $f(x)$ 的首项系数，n 为 $f(x)$ 的次数，记为 $\deg f = n$。$a \in R \subseteq R[x]$ 为常数多项式。若 a 为非零常数，规定 $\deg a = 0$，$0 \in R$ 为零多项式，规定 $\deg 0 = -\infty$。且规定：$-\infty < n$，$(-\infty) + n = n + (-\infty) = -\infty$，$(-\infty) + (-\infty) = -\infty$，这里 $n \in Z$。

关于多项式的次数，显然有以下结果：

定理 13.17 R 是环，$f(x), g(x) \in R[x]$，则

(1) $\deg[f(x) + g(x)] \leqslant \max[f(x), g(x)]$；

(2) $\deg[f(x) \cdot g(x)] \leqslant \deg f(x) + \deg g(x)$；

(3) 若 R 没有零因子，则 $\deg[f(x) \cdot g(x)] = \deg f(x) + \deg g(x)$。

定理 13.18 （带余除法）设 R 是含单位元 1 的环，$f(x), g(x)$ 为 $R[x]$ 的非零多项式，且 $g(x)$ 的首项系数为 R 中的可逆元，则存在惟一决定的多项式 $g(x), r(x) \in R[x]$ 使得

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

并且

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

证 若 $f(x)$次数$<g(x)$次数，取 $q(x)=0$，$r(x)=f(x)$ 即可。若 $\deg f(x) \geqslant \deg g(x)$，设 $f(x)$ 首项系数为 a_n，$g(x)$ 首项系数为 $b_s(n = \deg f(x) \geqslant \deg g(x) = s)$，则 $f_1(x) = f(x) - a_n b_s^{-1} g(x) \cdot x^{n-s}$ 的次数 $< f(x)$ 次数，对 $\deg f(x)$ 用归纳法，即得存在性的证明。

若 $f(x) = q_1(x)g(x) + r_1(x) = q_2(x)g(x) + r_2(x)$，这里 $\deg r_i(x) < \deg g(x)$ $(i=1,2)$，则$[q_1(x) - q_2(x)]g(x) = r_2(x) - r_1(x)$。若 $q_1(x) \neq q_2(x)$，得左边次数 $>$ 右边次数，不可能，故必 $q_1(x) = q_2(x)$，从而 $r_1(x) = r_2(x)$，得惟一性的证明。

二、多项式的零点

定义 13.22 若 R 是环，$f(x) \in R[x]$，$a \in R$，称为 $f(x)$的零点（或根），如果 $f(a) = 0$。

下面讨论整环上多项式的零点。

定理 13.19 若 R 是整环，$f(x) \in R[x]$，则 $a \in R$ 是 $f(x)$的零点 $\Leftrightarrow (x-a) \mid f(x)$。

证 由带余除法得

$$f(x) = q(x)(x-a) + c,$$

由多项式赋值同态得

$$f(a) = c,$$

于是

$$(x-a) \mid f(x) \Leftrightarrow c = f(a) = 0。$$

定理 13.20 设 R 是整环，$f(x) \in R[x]$，$a_1, \cdots, a_s$ 是 R 中两两不同的 s 个元素，则 $a_1, \cdots, a_s$ 是 $f(x)$的零点 $\Leftrightarrow (x-a_1)\cdots(x-a_s) \mid f(x)$。

证 充分性可由定理 13.19 立即推出。下证必要性：

因 a_1 是 $f(x)$的零点，故

$$f(x) = (x - a_1)f_1(x),$$

即

$$0 = f(a_i) = (a_i - a_1)f_1(a_i),\ i = 2, \cdots, s.$$

因为$a_i - a_1 \neq 0$，环R是整环，无零因子，必$f_1(a_i) = 0$，即$f_1(x)$以$a_2, \cdots, a_s$为零点。对s用归纳法，即得证明。

推论 整环上的n (>0)次多项式最多有n个不同零点。

为研究多项式的重根，现引进多项式的导数。

定义 13.23 R是整环，$f(x) \in R[x]$，$a \in R$，称为$f(x)$的k重根($k \in \mathbf{Z}^+$)，如果$(x-a)^k \mid f(x)$，而$(x-a)^{k+1} \nmid f(x)$；$k = 1$，a称为$f(x)$的单根；$k > 1$，a称为$f(x)$的重根。

定义 13.24 R是整环

$$R[x] \in f(x) = \sum_{i=0}^{n} a_i x^i.$$

记

$$f'(x) = \sum_{i=1}^{n} i a_i x^{i-1},$$

称它为多项式$f(x)$的导数。

直接验证，导数有下面性质：

(1) $[a \cdot f(x)]' = a \cdot f'(x)$；

(2) $[f(x) + g(x)]' = f'(x) + g'(x)$；

(3) $[f(x) \cdot g(x)]' = f'(x)g(x) + f(x)g'(x)$。

这里，$a \in R$，$f(x)$，$g(x) \in R[x]$。

定理 13.21 R是整环，$f(x) \in R[x]$，$a \in R$.则

(1) a是$f(x)$的重根$\Leftrightarrow f(a) = f'(a) = 0 \Leftrightarrow (x-a) \mid (f(x), f'(x))$，如果后者存在的话。

(2) 如果$(f(x), f'(x))$存在且~ 1，则$f(x)$无重根。

证 (2)是(1)的直接结果。只要证明(1)。

(1) 相当于说：a是$f(x)$的零点，则a是$f(x)$的重零点$\Leftrightarrow$ $f'(a)=0$。其证如下：

$$f(x) = (x-a)^k g(x),\ k \geqslant 1,\ g(a) \neq 0,$$
$$f'(x) = k(x-a)^{k-1}g(x) + (x-a)^k g'(x),$$

显然$f'(a)=0 \Leftrightarrow k>1$即$a$是$f(x)$的重根。

三、域上的多项式环

域K上的多项式环$K[x]$是主理想环，从而是F环。映射

$$\sigma(a) = a \cdot 1$$

是K到$K[x]$的单同态，因而可把K与$\{a1 \mid a \in K\}$等同起来，可把$K[x]$看成K的扩环。常称K中元为$K[x]$的常数项。

$K[x]$中元$f(x)$是可逆元$\Leftrightarrow f(x) \mid 1$，从而必$\deg f(x) \leqslant 0$，即$f(x) \in K$；反之$f(x)$为$K$中非零元$a$，则$a$有逆元$a^{-1}$。从而得下面的命题：

命题 13.11 设K是域，则$f(x)$是$K[x]$中的可逆元$\Leftrightarrow f(x) \in K-\{0\}$，即$K[x]$中的可逆元为$K$中非零元。

推论 A是整环，则$A[x]$中可逆元必是A中可逆元。

定义 13.25 设A是任意环，$f(x) \in A[x]$称为不可约多项式，如果(1) $\deg f(x) \geqslant 1$；(2) 若$g(x) \in A[x]$且$g(x) \mid f(x)$，则$\deg g(x) = 0$或$\deg f(x)$。

由于$K[x]$中可逆元是非零常数。故有：

推论 F环$K[x]$中次数大于等于1的多项式$f(x)$是既约元$\Leftrightarrow$ $f(x)$是$K[x]$的不可约多项式。

证 必要性显然。充分性证明如下：

若$g(x) \mid f(x)$，因$f(x)$是不可约多项式，故$\deg g(x) = 0$或$\deg f(x)$。$\deg g(x) = 0$，$g(x)$是非零常数，故可逆，即$g(x) \sim 1$。

$\deg g(x) = \deg f(x)$，则 $f(x) = c \cdot g(x)$，c 是非零常数，可逆，故 $f(x) \sim g(x)$，所以 $f(x)$ 是 $K[x]$ 的既约元。

复数域上多项式环 $C[x]$ 的既约元全是一次多项式；实数域上多项式环 $R[x]$ 的既约元是一次多项式或二次多项式 $ax^2 + bx + c$，这里判别式 $\Delta = b^2 - 4ac < 0$（即无实数根）。有理数域上多项式环 $Q[x]$ 的因式分解我们放在下一段与整数环上多项式环 $Z[x]$ 一起讨论。

四、F 环上的多项式环

设 R 是 F 环，P 是 R 的商域，则 $P[x]$ 也是 F 环。下面研究环 R 上多项式的因子分解。

引理 13.3 在 F 环中：

$$(ca_1, \cdots, ca_n) \sim c(a_1, \cdots, a_n)。$$

证 记 $d = (a_1, \cdots, a_n)$，则 $cd \mid ca_i$，$i = 1, \cdots, n$，于是

$$cd \mid (ca_1, \cdots, ca_n),$$

设

$$(ca_1, \cdots, ca_n) = cde,$$
$$cde \mid ca_i,$$

故

$$de \mid a_i, \quad (i = 1, \cdots, n),$$

从而

$$de \mid (a_1, \cdots, a_n) = d,$$

必

$$e \sim 1,$$

即

$$(ca_1, \cdots, ca_n) = cde \sim cd = c(a_1, \cdots, a_n),$$

定义 13.26 $f(x)\in R[x]$，若 $f(x)$的系数的最大公因子等价于1，则称 $f(x)$为本原多项式。

命题 13.12 任意 $f(x)\in R[x]$ $(f(x)\neq 0)$，存在 $d\in R$ 和本原多项式$\bar{f}(x)$使 $f(x)=d\cdot\bar{f}(x)$，且 $d\sim f(x)$系数的最大公因子，从而 d 在等价意义下是惟一的，称为多项式 $f(x)$的容度，记成 $c(f)$。

证 设

$$f(x)=a_0+a_1x+\cdots+a_nx^n,$$

记

$$d_1=(a_0,a_1,\cdots,a_n),\ b_i=\frac{a_i}{d_1},\ (i=0,1,\cdots,n),$$
$$f_1(x)=b_0+b_1x+\cdots+b_nx^n,$$

因

$$d_1=(a_0,\cdots,a_n)=(d_1b_0,\cdots,d_1b_n)\sim d_1(b_0,\cdots,b_n),$$

故

$$(b_0,b_1,\cdots,b_n)\sim 1.$$

即 $f_1(x)$ 为本原多项式。

又若

$$f(x)=d\bar{f}(x),\ \bar{f}(x)=c_0+c_1x+\cdots+c_nx^n$$

为本原多项式，则

$$d_1=(dc_0,\cdots,dc_n)\sim d(c_0,\cdots,c_n)\sim d。$$

推论 对任意 $f(x)\in P[x]$，必存在 $r,s\in R$ $(s\neq 0)$及 $R[x]$的本原多项式$\bar{f}(x)$使 $f(x)=\frac{r}{s}\bar{f}(x)$。

证 $f(x)\in P[x]$, $f(x)=\frac{a_0}{b_0}+\frac{a_1}{b_1}x+\cdots+\frac{a_n}{b_n}x^n=\frac{f_1(x)}{b_0\cdots b_n}$，这里：$a_i,b_i\in R$, $f_1(x)\in R[x]$，故

$$f(x)=\frac{r}{b}\overline{f_1}(x),$$

这里：

$$b=b_0\cdots b_n,\ r=c(f_1),\ f_1(x)=r\overline{f_1}(x)。$$

命题 13.13 R 中的素元 p 必是 $R[x]$ 中素元，从而是 $R[x]$ 的既约元。

证 因 p 是 R 的素元，故理想 $\langle p\rangle$ 是素理想，同余类环 $R/\langle p\rangle$ 是整环，映射

$$\sigma\left(\sum_i a_i x^i\right)=\sum_i \overline{a_i}x^i$$

是环 $R[x]$ 到环 $R/\langle p\rangle[x]$ 的满同态，核为 $\langle p\rangle$，由同态基本定理，$\frac{R[x]}{\langle p\rangle}$ 同构于整环 $R/\langle p\rangle[x]$，所以 $\langle p\rangle$ 为环 $R[x]$ 的素理想，从而 p 为 $R[x]$ 的素元。

推论 1 $R[x]$ 中本原多项式与本原多项式之积仍是本原多项式。

证 设 $g(x)$ 和 $h(x)$ 是本原多项式，$f(x)=g(x)h(x)$，若 $f(x)$ 不是本原多项式，则 $c(f)$ 是 R 的非零不可逆元，至少有一个素元因子 p，$p\mid f(x)=g(x)\cdot h(x)$。因 p 也是 $R[x]$ 的素元，所以 $p\mid g(x)$ 或 $p\mid h(x)$，这矛盾于 $g(x)$ 与 $h(x)$ 是本原多项式的假设，故必 $f(x)$ 是本原多项式。

推论 2 $f(x)$，$g(x)\in R[x]$，则 $c(f\cdot g)\sim c(f)\cdot c(g)$。

证 $f(x)\cdot g(x)=c(f)\cdot\overline{f}(x)\cdot c(g)\cdot\overline{g}(x)=c(f)\cdot c(g)\cdot\overline{f}(x)\cdot\overline{g}(x)$。由推论 1，$\overline{f}(x)\cdot\overline{g}(x)$ 是本原多项式，故 $c(f\cdot g)\sim c(f)\cdot c(g)$。

推论 3 $R[x]$ 中本原多项式的因子仍是本原多项式。

证 设 $f(x)$ 为本原多项式，$g(x)\mid f(x)$，则 $c(g)\mid c(f)$。因为 $c(f)\sim 1$，所以 $c(g)\sim 1$，故 $g(x)$ 是本原多项式。

若对 $f(x)$ 的次数用归纳法，则有

推论 4 $R[x]$ 中次数大于等于 1 的本原多项式必能表成有限个不

可约本原多项式的乘积。

命题 13.14 $R[x]$中本原不可约多项式$p(x)$是$P[x]$，也是$R[x]$的素元。

证 (1) 证$p(x)$是$P[x]$的既约元，从而是$P[x]$的素元。若$p(x)$在$P[x]$中不是既约元，即存在$P[x]$中次数小于$p(x)$的多项式$f(x)$与$g(x)$使

$$p(x)=f(x)g(x)=\frac{r_1}{s_1}\overline{f}(x)\cdot\frac{r_2}{s_2}\overline{g}(x),$$

于是在$R[x]$中：

$$s_1s_2p(x)=r_1r_2\overline{f}(x)\overline{g}(x),$$

必

$$s_1s_2\sim r_1r_2,\ p(x)\sim\overline{f}(x)\overline{g}(x),$$

这与$p(x)$是$R[x]$中不可约多项式的假设矛盾，故必$p(x)$是$P[x]$的既约元。

(2) 证$p(x)$是$R[x]$的素元。若在$R[x]$中，$p(x)\mid f(x)g(x)$，因为$p(x)$是$P[x]$的素元，又$R[x]\subseteq P[x]$，故在$P[x]$中：

$$p(x)\mid f(x)\ \text{或}\ p(x)\mid g(x)。$$

一般地，设$p(x)\mid f(x)$，即存在$P[x]$中多项式$h(x)$使

$$f(x)=p(x)h(x)=p(x)\cdot\frac{r}{s}\overline{h}(x),$$

于是在$R[x]$中：

$$sf(x)=rp(x)\overline{h}(x),$$

$$s\cdot c(f)\cdot\overline{\overline{f}}(x)=rp(x)\overline{h}(x),$$

$$s\cdot c(f)\sim r,\ \overline{\overline{f}}(x)\sim p(x)\overline{h}(x),$$

故

$$p(x) \mid \bar{f}(x) \mid c(f)\bar{f}(x) = f(x)。$$

所以 $p(x)$为 $R[x]$的素元。

命题 13.15 $R[x]$中素元仅两类：

(1) R 中素元；

(2) 本原不可约多项式。

证 设 $f(x)$是 $R[x]$的素元，从而是既约元：

$$f(x) = c(f) \cdot \bar{f}(x)。$$

(1) $c(f) \not\sim 1$，则 $f(x) \sim c(f)$为 R 中的既约元，从而是 R 中的素元。

(2) $c(f) \sim 1$，则 $f(x) \sim \bar{f}(x)$为本原多项式，因 $f(x)$是既约元，故是本原不可约多项式。

定理 13.22 $R[x]$是 F 环，即 F 环上的多项式环仍是 F 环。

证 设 $f(x)$是 $R[x]$的一个非零不可逆元，有

$$f(x) = c(f) \cdot \bar{f}(x)。$$

$c(f) \in R$，R 是 F 环，所以 $c(f)$可表成 R 中的从而也是 $R[x]$中的既约元的乘积 $c(f) = p_1 \cdots p_s$。$\bar{f}(x)$次数若$\geqslant 1$，由命题 13.13 的推论 4，可表成有限个本原不可约多项式的乘积，即有限个 $R[x]$中既约元的乘积，$\bar{f}(x) = f_1(x) \cdots f_t(x)$。因此，$f(x)$能表成 $R[x]$中有限个既约元的乘积。

若 $f(x) = p_1 \cdots p_s f_1(x) \cdots f_t(x) = q_1 \cdots q_l \cdot g_1(x) \cdots g_n(x)$，这里 p_i，q_j 是 R 的既约元，$f_i(x)$，$g_j(x)$是本原不可约多项式。因为 $p_1 \cdots p_s \sim c(f) \sim q_1 \cdots q_l$，$R$ 是 F 环，所以 $l = s$，且存在$\{1, \cdots, s\}$的置换 σ_1 使 $p_i \sim q_{\sigma_1(i)}$。又 $f_1(x) \cdots f_t(x) \sim g_1(x) \cdots g_n(x)$，素元 $f_i(x) \mid g_1(x) \cdots g_n(x)$，(适当调整次序)得

$$f_i(x) \mid \text{既约元 } g_i(x),$$

于是

$$f_i(x) \sim g_i(x),$$

定理得证。

定义 13.27 R 是环，归纳地定义 R 上多元多项式环为：

$$R[x_1, x_2] = (R[x_1])[x_2], \cdots, R[x_1, x_2, \cdots, x_n] = (R[x_1, \cdots, x_{n-1}])[x_n]。$$

推论 若 R 是 F 环，则 R 上多元多项式环 $R[x_1, \cdots, x_n]$ 也是 F 环。

特别，$Z[x]$，$Q[x_1, \cdots, x_n]$是 F 环。

下面给出多项式不可约的一个著名判别法。

Eisenstein 判别法 若 R 是 F 环，p 是 R 的一个素元。$f(x) = a_0 + a_1 + \cdots + a_n x^n \in R[x]$ $(n \geqslant 1)$，满足：

(1) $p \nmid a_n$；

(2) $p \mid a_i (i=0, 1, \cdots, n-1)$；

(3) $p^2 \nmid a_0$，

则 $f(x)$是不可约多项式。

证 若 $f(x)$可约，即存在

$g(x), h(x) \in R[x], \deg g(x), \deg h(x) > 0$.

且

$$f(x) = g(x) \cdot h(x)。$$

设

$$g(x) = b_0 + \cdots - b_m x^m, \quad (m \geqslant 1)$$

$$h(x) = c_0 + \cdots + c_t x^t, \quad (t \geqslant 1)$$

因

$$p \mid a_0 = b_0 c_0, \ p^2 \nmid a_0 = b_0 c_0,$$

故可设

$$p \mid b_0, \ p \nmid c_0。$$

又，p 不能整除所有 $g(x)$ 系数（否则 $p|a_n$），故存在 $1\leqslant j\leqslant m$ 使

$$p|b_0,\ \cdots,\ b_{j-1},$$

但

$$p\nmid b_j。$$

此时

$$a_j = b_jc_0 + b_{j-1}c_1 + \cdots + b_0c_j,$$

第一项不是 p 的倍数，其余的项都是 p 的倍数，因而

$$p\nmid a_j,\ 0\leqslant j\leqslant m<n,$$

这与假设矛盾，故 $f(x)$是不可约多项式。

推论 在判别法的条件下，若 P 是 R 的商域，则 $f(x)$是 $P[x]$的既约元。

例 13.23 $2x^3-6x^2+9x-3$ 是 $Z[x]$的既约元。

证 取 $p=3$，则由 Eisenstein 判别法知多项式是不可约多项式。显然，多项式又是本原的，故是既约元。

例 13.24 $f(x,y)=4x^2+5(y-1)x+3(y-1)$ 是 $Z[x,y]$的既约元。

证 $y-1$ 是 $z[y]$的本原不可约多项式，即既约元。因 $Z[y]$是 F 环，故 $y-1$ 是它的素元。用 Eisenstein 判别法：$Z[x,y]=Z[y][x]$ 中多项式 $f(x,y)$，作为系数在 $Z[y]$中的 x 的多项式是不可约的，又显然 $f(x,y)$是本原的，所以它是 $Z[x,y]$的既约元。

引理 13.4 $f(x)\in R[x]$，a 是 R 的可逆元，则 $f(x)$ 与 $f(ax+b)$ 同时可约或不可约。

证 若 $f(x)$可约，即存在次数小于 $f(x)$的多项式 $g(x)$与 $h(x)$使

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

于是

$$f(ax+b)=g(ax+b)h(ax+b),$$

即 $f(ax+b)$ 可约。

反之，若 $f(ax+b)=t(x)$ 可约，即存在次数小于 $t(x)$ 的多项式 $g(x)$，$h(x)$ 使

$$f(ax+b)=t(x)=g(x)h(x),$$

于是

$$f(x)=f\left[a\left(\frac{x-b}{a}\right)+b\right]=t\left(\frac{x-b}{a}\right)=g\left(\frac{x-b}{a}\right)\cdot h\left(\frac{x-b}{a}\right),$$

即可约。

命题 13.16 若 p 是素数，则多项式

$$f(x)=1+x+\cdots+x^{p-1}$$

是 $Z[x]$的从而也是 $Q[x]$的不可约多项式。

证 $f(x)=\dfrac{x^p-1}{x-1}$，于是

$$f(x+1)=\frac{(x+1)^p-1}{x}=x^{p-1}+C_p^1x^{p-2}+\cdots+C_p^{p-1},$$

因为

(1) p 不能整除 x^{p-1} 的系数 1，

(2) $p|C_p^k$，$k=1,\cdots,p-1$，

(3) $p^2\nmid p=C_p^{p-1}$，

故 $f(x+1)$，从而 $f(x)$ 是 $Z[x]$ 的不可约多项式。

习　题　13

1. 求出 Z_5 中所有非零元的逆元。
2. 在 Z_{15} 中求出方程 $x^2=1$ 的所有根。
3. 若 R 是交换环，证明二项式定理，即 $\forall a,b\in R$ 有

$$(a+b)^n=\sum_{k=0}^{n}C_n^ka^kb^{n-k}。$$

4. 求出高斯整环 Z[i]中的所有可逆元。

5. 在四元数体中，求出方程 $x^2 + e = 0$ 的所有根。

6. 设 $R = \{0, 1, a\}$ 是三元环，这里 0，1 分别是加法与乘法单位元，求证：$a^2 = 1$。

7. R 是一个环，在多项式环 $R[x]$ 中定义运算：

$$f(x) \cdot g(x) = f[g(x)].$$

问 $R[x]$，关于多项式的加法和"·"是否成一环？

8. 若 R 是含单位元的环，$a \in R$，a 有右逆，则下列命题是等价的：

(1) a 不可逆；

(2) a 是左零因子；

(3) a 有多于一个的右逆。

9. 若 R 是含单位元 e 的环，a 是 R 的幂零元，则 $e - a$ 是 R 的可逆元。

10. 设 R 是一个环，$\forall a \in R$，$a^2 = a$。则 R 是交换环，且 $\forall a \in R$ 有 $a + a = 0$。

11. 若 R 是阶大于 1 的环，$\forall a \in R$，$a \neq 0$，都有惟一 $b \in R$ 使 $aba = a$，求证：R 是体。

12. 若 R 是含单位元 e 的环，$a, b \in R$；若 $e - ab$ 有逆元 c，则 $e - ba$ 的逆元是 $e + bca$。

13. 若 R 是含单位元的环，$a \in R$，a 有多于一个的右逆，则 a 有无限多个右逆。

14. 若 Z[i]为 Gauss 整环，试求商环 $Z[i]/\langle 2 + i\rangle$ 的阶和特征。

15. 若 I 表示交换环 R 的幂零元全体所成的理想，求证：商环没有非零幂零元。

16. $R = (Z)_2$ 为整系数二阶方阵环，I 为元素是 3 的倍数所有二阶方阵所成集合，求证：I 是 R 的理想，且求商环 R/I 的元素个数。

17. 求证：Gauss 整环 Z[i]同构于 $Z[x]/\langle x^2 + 1\rangle$。

18. 求证：若 $n \in \mathbf{Z}^+$，则 $Z[x]/\langle x - 2, n\rangle \cong Z_n$。

19. 找出 Z 到自身的一切同态映射。

20. 求出 $Z[\mathrm{i}]$的一切自同态。

21. 若 R 是特征为 0 的整环，求同余类环 Z_n 到 R 的一切同态映射。

22. 若 R 是环 R_1，…，R_n 的直积 $R = R_1 \times R_2 \times \cdots \times R_n$，且 R_i 均有单位元($i = 1, \cdots, n$)，则 I 是 R 的理想 $\Leftrightarrow$ 存在 R_i 的理想 I_i ($i = 1, \cdots, n$) 使得 $I = I_1 \times I_2 \times \cdots \times I_n$。

23. 若 I_1 和 I_2 是环 R 的两个理想，$\langle I_1 \cup I_2 \rangle = R$，则

$$R/I_1 \cap I_2 \cong R/I_1 \times R/I_2。$$

24. 若 f, g 都是有理数域 $\mathbf{Q}$ 到特征为 0 的环 A 的环同态，且 $\forall\, n \in \mathbf{Z}$，$f(n) = g(n)$，求证：$f = g$。

25. 若 $a, b \in \mathbf{Z}$，且 $a^2 + b^2$ 为素数 p，求证：

$$\frac{Z[\mathrm{i}]}{\langle a + b\mathrm{i} \rangle} \cong Z_p。$$

26. 若 p 为一固定素数，记

$$R = \left\{ \frac{t}{s} \,\middle|\, t, s \in \mathbf{Z}, s \neq 0, (p, s) = 1 \right\},$$

求证：R 是整环，并求 R 的商域。

27. 记 $Z[\sqrt{-5}] = \{a + b\sqrt{-5} \mid a, b \in \mathbf{Z}\}$，求证：$Z[\sqrt{-5}]$是整环，并求它的商域。

28. 若 R 是整环，$a, b \in R$。s, t 是两个互素正整数，$a^s = b^s$，$a^t = b^t$，求证：$a = b$。

29. 若 S 是交换环 R 的乘法半群，在积集合 $R \times S$ 上定义二元关系：$(r_1, s_1) \sim (r_2, s_2) \Leftrightarrow$ 存在 $s \in S$ 使 $s(r_1 s_2 - r_2 s_1) = 0$，求证：这是等价关系，试在商集合 $R \times S/\sim$ 中定义加法和乘法，使之成为一个含单位元的交换环。

30. 求出同余类环 Z_{72}的所有素理想与极大理想。

31. $a, b \in \mathbf{Z}$，$b \neq 0$，求证：

$$\left|\frac{Z[x]}{\langle x-a,b\rangle}\right| = |b|。$$

32. 求证：$\langle x,y\rangle$和$\langle x,y,p\rangle$是环$Z[x,y]$的素理想，这里p是素数。

33. $Z[x]$中，$\langle x,n\rangle$是极大理想$\Leftrightarrow n$为素数。

34. 若R是含单位元的交换环，I是R的真理想，求证：存在R的极大理想M，且$M\supset I$。

35. 证明：$\langle 1+i\rangle$是$Z[i]$的素理想，但$\langle 2\rangle$不是。

36. 若R是交换环，$R^2=R$，则R中极大理想一定是素理想。

37. R是有限交换环，则R的真素理想必是R的极大理想。

38. 若R是F环，求证：

(1) $(da_1,\cdots,da_n)\sim d(a_1,\cdots,a_n)$。这里$d,a_1,\cdots,a_n\in R$；

(2) $a,b,c\in R$，且$(a,b)\sim 1$，$(a,c)\sim 1$，则$(a,bc)\sim 1$。

39. 若R是主理想整环，P是R的商域，求证：P的包含R的子环也是主理想整环。

40. 在$Z[i]$中，分解下列元素为既约元之积：

(1) 6 800；

(2) 1 430 + 2 860 i；

(3) 13 − i。

41. 若R是交换环，N为R的幂零元全体所成集合，则R中所有素理想的交恰好是N。

42. $s,t\in Z^+$，$s|t$，求证：在环$Z[x]$中，$x^s-1|x^t-1$。

43. 若R是含单位元1的环，$n\in\mathbf{Z}^+$，$f(x)\in R[x]$，若$(x-1)|f(x^n)$，则$(x-1)|f(x)$。

44. 若K是域，$f(x)$，$g(x)\in K[x]$，且$(f(x),g(x))\sim 1$，求证：$(f(x)\cdot g(x),f(x)+g(x))\sim 1$.

45. $Q[x]\ni f(x)=1+x+\dfrac{x^2}{2!}+\cdots+\dfrac{x^n}{n!}\ (n\geqslant 1)$，求证：$f(x)$没有重根。

46. 若K是域，$f(x)=x^3+bx+c\in K[x]$，求$f(x)$有重根的

条件。

47. 若 R 是 F 环，有素元 a，求证：$R[x]$不是主理想整环，由此推出 $Z[x]$和域上多元多项式环不是主理想整环。

48. 若 p 是奇素数，求证：

$$f(x) = x^p + px + 1$$

是 $Z[x]$的既约元。

49. 求证：$f(x,y) = 2x^2 - 3xy + 5y^2 - 3x - 5y - 10$ 是$Z[x,y]$的既约元。

50. K 是域，若 $x^n - a$ 是 $K[x]$中的不可约多项式，$m|n$，则 $x^m - a$ 也是 $K[x]$中的不可约多项式。

第十四章　域

域是环的一种类型，是最常见和最常用的代数结构。域的加法群是交换群，非零元集合成一乘法交换群。由于域内，方程 $a+x=b$ 和 $ax=b\ (a\neq 0)$ 均有惟一解，故加法和乘法有逆运算，其逆运算分别为减法和除法，因而域是其内元素可作四则运算的代数结构。

本章将详细讨论域的结构。因为任意域都可看成是它的子域的扩域，由子域添加若干元扩张而成。所以研究需从域的"扩张"入手。"扩张"的出发点，就是"素域"。

§1　素域和域的扩张

一、素域的概念

定义 14.1　域的子集若对域的两种运算仍是域，则称为子域。除了自身外，不再含其他子域的域叫素域。例如，有理数域 Q 和同余类域 Z_p（这里 p 是素数）。

同子群和理想的情形完全类似，有以下命题。

命题 14.1　域的任意个子域的交仍是子域。

证　因为子域的加法群是原来子域的加法子群，而加法子群的交仍是加法子群，故子域的交是加法子群。

又，子域的非零元组成乘法子群，而乘法子群的交仍是乘法子群，故子域的交中非零元成乘法子群。

子域的交是原来域的加法子群，非零元成乘法子群。又，分配律（作为原域的子集）是成立的，故子域的交仍是子域。

定理 14.1　任意域包含一个且仅一个素域。

证　设 F 是域，记 Π 为 F 的所有子域的交，由命题 14.1，Π 是 F

的子域，且 Π 不含真子域，故 Π 是 F 所含的素域。若 F 含两个素域 Π_1 和 Π_2，则 $\Pi_1 \cap \Pi_2$ 仍是域，必 $\Pi_1 = \Pi_1 \cap \Pi_2 = \Pi_2$，即 F 含且仅含一个素域。

二、素域的类型

下面证明，素域仅有两种类型，即 Q 和 Z_p。

设 Π 是素域，e 是 Π 的单位元。作为域，Π 必不含零因子；作为环，Π 的特征 $\mathrm{ch}\,\Pi$ 由单位元 e 确定，或者为 0（此时 e 作为加法群的元，其阶为 $+\infty$），或者为素数 p（此时，e 和所有非零元作为加法群的元，其阶均为同一素数 p）。

（1）$\mathrm{ch}\,\Pi = 0$，e 的阶为 $+\infty$，于是 Π 的（整）子环 $\{ne \mid n \in Z\}$ 同构于整数环 Z，由于同构的整环有同构的商域，从而 $\Pi \cong Q$。

（2）$\mathrm{ch}\,\Pi =$ 素数 p，e 的阶为 p，此时，Π 的子集 $\{0, e, 2e, \cdots, (p-1)e\}$ 已构成一个同构于 Z_p 的域，因为 Π 是素域，不含真子域，故必

$$\Pi = \{0, e, 2e, \cdots, (p-1)e\} \cong Z_p。$$

于是，证明了下面的定理 14.2：

定理 14.2 素域仅有两类，即 Q 与 Z_p（这里 p 为素数）。

作为环，域的特征由单位元惟一确定。又，域与其子域有相同单位元，故其特征相同。而 $\mathrm{ch}\,Q = 0$，$\mathrm{ch}\,Z_p = p$。故有下面的定理 14.3：

定理 14.3 F 是域，则

（1）$\mathrm{ch}\,F = 0 \Leftrightarrow F$ 含同构于 Q 的素域；

（2）$\mathrm{ch}\,F = p \Leftrightarrow F$ 含同构于 Z_p 的素域。

因为域不含零因子，故其非零元有相同阶（加法），都等于单位元的阶。故有：

性质 （1）F 是域，若 $\mathrm{ch}\,F = 0$，则 $n \in Z$，$a \in F$，$na = 0 \Leftrightarrow n = 0$ 或 $a = 0$（因为 F 的非零元阶都是 $+\infty$）。

（2）F 是域，若 $\mathrm{ch}\,F = p$，则 $n \in Z$，$a \in F$，$na = 0$，$\Leftrightarrow p \mid n$ 或 $a = 0$（因为 F 的非零元阶都 $= p = \mathrm{ch}\,F$）；

(3) F 是域，若 $\mathrm{ch}\, F = p$，a、$b \in F$，则

$$(a + b)^p = a^p + b^p \text{；} (a - b)^p = a^p - b^p\text{。}$$

证 因为 F 是域，$ab = ba$，所以

$$(a + b)^p = a^p + C_p^1 a^{p-1} b + \cdots + C_p^{p-1} ab^{p-1} + b^p\text{。}$$

由于

$$p \mid C_p^1,\ C_p^2,\ \cdots,\ C_p^{p-1},$$

故

$$(a + b)^p = a^p + b^p\text{。}$$

又

$$a^p = [(a - b) + b]^p = (a - b)^p + b^p,$$

所以

$$(a - b)^p = a^p - b^p\text{。}$$

三、素域 Z_p 及其特征方程

设 p 为素数，素域 Z_p 中共有 $(p - 1)$ 个非零元，它们构成一个 $(p - 1)$ 阶乘法群，因而非零元的 $(p - 1)$ 次幂均等于单位元 $e = \bar{1}$，即非零元都满足下列方程：

$$x^{p-1} = e = \bar{1},$$

由于零元满足方程

$$x^p = x \cdot x^{p-1} = x \cdot e = x,$$

或

$$x^p - x = 0,$$

故 Z_p 中的 p 个元恰是 p 次多项式

$$x^p - x = 0$$

的 p 个(全部)零点。

我们称方程

$$x^p - x = 0$$

为素域 Z_p 的特征方程。

若素域 Z_p 的 p 个元为 $\bar{0}, \bar{1}, \cdots, \overline{p-1}$,则

$$x^p - x = x(x - \bar{1})(x - \bar{2})\cdots(x - \overline{p-1})。$$

四、域的扩张

若 F 是域 K 的子域,M 是 K 的子集,记

$$F(M) = \bigcap_{\substack{L \supseteq F \cup M \\ L\text{是}K\text{子域}}} L,$$

称它为 F 添加 M 而得到的扩域。

$F(M)$有如下性质:

(1) $F(M)$是域 K 的子域的交,故仍是子域;

(2) $F(M) \supseteq F \cup M$;

(3) 若 K 的子域 $L \supseteq F \cup M$, 则 $L \supseteq F(M)$;

(4) $F(M)$是系数在 F 中的有限个 M 中元素的有理组合(加,减,乘,除)的全体,这是因为(若记后者为 Σ):

$F(M) \supseteq F \cup M$, $F(M)$ 是域,故 $F(M) \supseteq \Sigma$,又 $\Sigma \supseteq F$, M 且是域,由(3), $\Sigma \supseteq F(M)$。于是

$$F(M) = \bigcup_{\substack{N\text{是}M\text{的} \\ \text{有限子集}}} F(N)。$$

若 M_1, M_2 是 K 的两个子集,则 $F(M_1 \cup M_2) = F(M_1)(M_2)$,这是因为 $F(M_1 \cup M_2) \supseteq F \cup M_1$ 及 M_2,于是 $\supseteq F(M_1) \cup M_2$,故 $\supseteq F(M_1)(M_2)$。反过来,$F(M_1)(M_2) \supseteq F(M_1) \cup M_2$,于是 $\supseteq F \cup M_1 \cup M_2$,故 $F(M_1)(M_2) \supseteq F(M_1 \cup M_2)$。从而 $F(M_1 \cup M_2) = F(M_1)(M_2)$。

同理，$F(M_2)(M_1) = F(M_2 \cup M_1) = F(M_1 \cup M_2) = F(M_1)(M_2)$。

若 N 是有限集，$N = \{\alpha_1, \cdots, \alpha_s\}$，则记 $F(N) = F(\alpha_1, \cdots, \alpha_s)$。由 $F(M_1 \cup M_2) = F(M_1)(M_2)$ 可知：

$$F(\alpha_1, \cdots, \alpha_s) = F(\alpha_1)(\alpha_2)\cdots(\alpha_s)。$$

综上所述，添加无限集 M 到域 F 中去，相当于添加有限集 N 到 F 中去然后作并等。而添加有限集 N 到 F 中去，相当于有限次添加一个元到 F 中去。因此，添加一个元的扩张是最基本的。下面，研究添加单个元的扩张。

§2 单纯代数扩域

若 F 是域 K 的子域，$\alpha \in K$，记 $F(\alpha)$ 是域 F 添加单个元素 α 得到的扩域。现在来研究中间域 $F(\alpha)$ 的结构。我们从熟悉的多项式环出发，导出相应结论。

映射

$$\sigma[f(x)] = f(\alpha)$$

是域 F 上多项式环 $F[x]$ 到环 $F(\alpha)$ 的一个同态映射。由于 σ 保持 F 中元素不变，故不是零同态，即核 $\operatorname{Ker}\sigma$ 是 $F[x]$ 的真理想。由环同态基本定理得

$$F[x]/\operatorname{Ker}\sigma \cong \sigma(F[x]) = F[\alpha] = \{f(\alpha) \mid f(x) \in F[x]\}。$$

分两种情形：

(1) $\operatorname{Ker}\sigma = \{0\}$，此时 σ 是单同态，从而

$$F[x] \cong F[\alpha]。$$

同构的整环其商域也同构，故

$$F(\alpha) \cong F(x) = \left\{\frac{f(x)}{g(x)} \middle| f(x),\ g(x) \in F[x],\ g(x) \neq 0\right\}。$$

此时，由于 $\operatorname{Ker}\sigma = \{0\}$，得

$$f(\alpha)=\sigma[f(x)]=0\Leftrightarrow f(x)=0,$$

即 α 不是系数在 F 中的非零多项式的根；

(2) $\operatorname{Ker}\sigma$ 是主理想整环 $F[x]$的非零真理想：

$$\operatorname{Ker}\sigma=\langle p(x)\rangle,$$

由于

$$F[x]/\langle p(x)\rangle\cong F[\alpha],$$

$F[\alpha]\subseteq$ 域 $F(\alpha)$，无零因子，故 $\langle p(x)\rangle$ 为 $F[x]$ 的非零真素理想，从而 $p(x)$ 为 $F[x]$ 的一个不可约多项式。此时，同余类环 $F[x]/\langle p(x)\rangle$ 已经是域，故

$$F(\alpha)=F[\alpha],$$

且

$$\begin{aligned}&f(\alpha)=\sigma[f(x)]=0\Leftrightarrow f(x)\in\langle p(x)\rangle\\&\Leftrightarrow p(x)\mid f(x),\end{aligned}$$

即 α 是系数在 F 中多项式 $f(x)$的根（$f(x)$是 $p(x)$的倍式），特别是不可约（非零）多项式 $p(x)$的根。

综上所述，即得下面定理 14.4：

定理 14.4 F 是域 K 的子域，$\alpha\in K$，则

$$F(\alpha)\text{ 或 }\cong F(x)\text{ 或 }\cong F[x]/\langle p(x)\rangle。$$

这里，$p(x)$是系数在 F 中的一个不可约多项式。

定义 14.2 设 F 是域 K 的子域，$\alpha\in K$，则有两种情形：

(1) α 不是系数在 F 中的任意一个非零多项式的根，称 α 为域 F 的一个超越元，此时 $F(\alpha)\cong F(x)$。称 $F(\alpha)$为 F 的单纯超越扩域；

(2) α 是系数在 F 中的某个非零多项式的根（由上分析，α 必是系数在 F 中的某个不可约多项式 $p(x)$的根），称 α 为 F 的代数元。F 的扩域 $F(\alpha)$称为 F 的单纯代数扩域。此时

$$F(\alpha)\cong F[x]/\langle p(x)\rangle。$$

若 $p(x)$ 为 n 次多项式，则 $F(\alpha)$ 由系数在 F 中的次数不超过 $(n-1)$ 的 α 的多项式组成，$f(\alpha)=r(\alpha)$，这里 $r(x)$ 是 $f(x)$ 被 $p(x)$ 除后所得的余式，即

$$f(x)=q(x)p(x)+r(x),$$
$$\deg r(x)<\deg p(x)=n。$$

两个 α 的多项式相加即对应系数相加，两个 α 的多项式相乘按普通多项式乘法相乘并对 $p(x)$ 取余式。求非零元 $f(\alpha)$ 的逆的方法如下：

因为 $f(\alpha)\neq 0$，故 $p(x)\nmid f(x)$，于是 $(p(x),\ f(x))=1$，故存在多项式 $u(x)$，$v(x)$ 使

$$u(x)f(x)+v(x)p(x)=1,$$

于是

$$u(\alpha)\cdot f(\alpha)=1,$$

即 $u(\alpha)$ 为 $f(\alpha)$ 的逆（$u(x)$ 可经 $f(x)$ 与 $p(x)$ 转辗相除或待定系数法求得）。

例 14.1 设 ω 是 $Q[x]$ 中不可约多项式 $x^2+x+1=0$ 的根，于是 $Q(\omega)=\{a\omega+b \mid a,\ b\in Q\}$，试在 $Q(\omega)$ 中求

(1) $(\omega+1)+(2\omega+1)$；

(2) $(\omega+1)\cdot(2\omega+1)$；

(3) $3\omega+1$ 的逆元。

解 (1) $(\omega+1)+(2\omega+1)=3\omega+2$。

(2) $(\omega+1)\cdot(2\omega+1)=2\omega^2+3\omega+1=2(\omega^2+\omega+1)+\omega-1=\omega-1$。

(3) 因为 $x^2+x+1=\left(\frac{1}{3}x+\frac{2}{9}\right)(3x+1)+\frac{7}{9}$，故

$$\left(\frac{1}{3}\omega+\frac{2}{9}\right)(3\omega+1)=-\frac{7}{9},$$

即

$$\left[-\frac{1}{7}(3\omega+2)\right](3\omega+1)=1,$$

于是 $3\omega+1$ 的逆为

$$-\frac{1}{7}(3\omega+2)。$$

例 14.2 由 Eisenstein 不可约判别法知

$$p(x)=x^3+4x+2$$

是 $Q[x]$的三次不可约多项式，设 α 是 $p(x)$的一个根，于是 $Q(\alpha)$是系数在 Q 中的 α 的次数不大于 $2(=3-1)$的多项式全体，即

$$Q(\alpha)=\{a\alpha^2+b\alpha+c\,|\,a,b,c\in Q\},$$

试在 $Q(\alpha)$ 中表

$$\frac{2\alpha^2+\alpha+7}{\alpha^2+1}$$

为 α 的二次多项式。

解 用待定系数法，设

$$\frac{2\alpha^2+\alpha+7}{\alpha^2+1}=a\alpha^2+b\alpha+c,$$

于是

$$2\alpha^2+\alpha+7=(\alpha^2+1)(a\alpha^2+b\alpha+c),$$

或

$$2\alpha^2+\alpha+7-\alpha^3(a\alpha+b)-(a+c)\alpha^2-b\alpha-c=0。$$

用 $\alpha^3=-4\alpha-2$ 代入上式，化简，得

$$(3a-c+2)\alpha^2+(2a+3b+1)\alpha+(2b-c+7)=0。$$

因为以 α 为根的次数最低的非零多项式是 $p(x)$（三次多项式），故上式必为零多项式，即

$$\begin{cases}3a - c + 2 = 0, \\ 2a + 3b + 1 = 0, \\ 2b - c + 7 = 0。\end{cases}$$

解之,得

$$a = 1,\ b = -1,\ c = 5,$$

即

$$\frac{2\alpha^2 + \alpha + 7}{\alpha^2 + 1} = \alpha^2 - \alpha + 5。$$

定理 14.5 F 是域,若 $p(x)$是 $F[x]$的一个不可约多项式,则存在 F 的扩域 K 使 $p(x)$在 K 中有根。

证 $F[x]$是主理想整环,$p(x)$是其既约元,所以同余类环

$$F[x]/\langle p(x)\rangle$$

是域,记它为 K。映射

$$\sigma(a) = \bar{a}$$

是 F 到 K 的单同态,故 F 同构于 K 的子域

$$\{\bar{a} \mid a \in F\},$$

即 K 可看成 F 的扩域。在 K 中:

$$p(\bar{x}) = \overline{p(x)} = \bar{0},$$

即 $\alpha = \bar{x}$ 是 $p(x)$ 的根,定理得证。

例 14.3 求 Q 的扩域 K,使不可约多项式

$$p(x) = x^2 - 2$$

在 K 中有根。

解 令 $K = \{a\alpha + b \mid a, b \in Q\}$
$$\cong \{\overline{ax + b} \mid a, b \in Q\} = Q[x]/\langle x^2 - 2\rangle,$$
这里 $\alpha^2 = 2$,$\pm\alpha$ 都是 $p(x)$ 的根。

§3 有限扩域与代数扩域

一、域上线性空间

定义 14.3 若 K 是域，它的元用 $a,b,c,\cdots$ 来表示，V 是加法群，它的元用 $\alpha,\beta,\gamma,\cdots$ 来表示；若映射 $f: K\times V \xrightarrow{f} V$ 满足下列性质，则称 $\boldsymbol{V}$ 为域 K 上的线性空间（可简单地记 $f((a,\alpha))$ 为 $a\alpha$）。

(1) $a(\alpha+\beta)=a\alpha+a\beta$;

(2) $(a+b)\alpha=a\alpha+b\alpha$;

(3) $(ab)\alpha=a(b\alpha)$;

(4) $1\alpha=\alpha$，1 是 K 的单位元。

上述 α,β 是 $\boldsymbol{V}$ 中任意元；a,b 是 K 中任意元。

由定义，易得性质：$0\alpha=a0=0$；$(-a)\alpha=a(-\alpha)=-(a\alpha)$；$(-a)(-\alpha)=a\alpha$。$a\alpha=0$ 必 $a=0$ 或 $\alpha=0$。

$\boldsymbol{V}$ 的子群 $\boldsymbol{W}$ 若仍是 K 上线性空间，常称为 $\boldsymbol{V}$ 的子空间。任意线性空间 $\boldsymbol{V}$ 有两个平凡子空间 $\{\boldsymbol{0}\}$ 和 $\boldsymbol{V}$。

定义 14.4 $\boldsymbol{V}$ 中元 $\alpha_1,\cdots,\alpha_s$，称为关于 K 线性相关的，如果存在 K 中 s 个不全为 0 的元 $k_1,\cdots,k_s$ 使 $\sum\limits_{i=1}^{s}k_i\alpha_i=0$。如果不存在这样的元 $k_1,\cdots,k_s$，即上述关系仅当 $k_1,\cdots,k_s$ 全是零时成立，则称 $\alpha_1,\cdots,\alpha_s$ 关于 K 线性无关（简称线性相关和线性无关）。

若 $\beta=\sum\limits_{i=1}^{s}a_i\alpha_i$，称 β 为 $\alpha_1,\cdots,\alpha_s$ 的线性组合，记成

$$\alpha_1,\cdots,\alpha_s\Rightarrow\beta。$$

引理 若 n 个未定元 $x_1,\cdots,x_n$ 的齐次线性方程组

$$\sum_{j=1}^{n}a_{ij}x_j=0,\ (i=1,\cdots,s)$$

中系数 a_{ij} 都是域 K 中元，且 $n>s$，则在 K 中此方程组有非零解。

证 对 s 用归纳法。当 $s=1$ 时,引理显然成立。若 $s-1$ 时引理成立。设

$$l_i=\sum_{j=1}^{n}a_{ij}x_j(i=1,\cdots,s)。$$

如果 $a_{i1}=0\ (i=1,\cdots,s)$,显然有非零解:

$$x_1=1,x_2=\cdots=x_n=0。$$

如果 $a_{11},a_{21},\cdots,a_{s1}$,不全为 0,可设 $a_{11}\neq 0$。于是方程组 $l_i=0\ (i=1,\cdots,s)$ 同解于方程组

$$\begin{cases} l_1=0, & (1)\\ l_i-a_{11}^{-1}a_{i1}l_1=0\quad(i=2,\cdots,s)。 & (2)\end{cases}$$

但方程组(2) 有 $(n-1)$ 个未定元,$(s-1)$ 个方程式,由归纳法假设,在 K 中有非零解:$x_i=c_i(i=2,\cdots,n)$。从而

$$x_1=-a_{11}^{-1}\left(\sum_{i=2}^{n}a_{1i}c_i\right),x_i=c_i(i=2,\cdots,n)$$

是方程组 $l_i=0\ (i=1,\cdots,s)$ 在 K 中的非零解,引理得证。

由引理,方程组 $A\begin{pmatrix}x_1\\ \vdots\\ x_s\end{pmatrix}=\begin{pmatrix}0\\ \vdots\\ 0\end{pmatrix}$ 有非零解 $\begin{pmatrix}c_1\\ \vdots\\ c_s\end{pmatrix}$,此时

$$\sum_{i=1}^{s}c_i\boldsymbol{\beta}_i=(\boldsymbol{\beta}_1,\cdots,\boldsymbol{\beta}_s)\begin{pmatrix}c_1\\ \vdots\\ c_s\end{pmatrix}=(\boldsymbol{\alpha}_1,\cdots,\boldsymbol{\alpha}_n)A\begin{pmatrix}c_1\\ \vdots\\ c_n\end{pmatrix}=0。$$

故 $\boldsymbol{\beta}_1,\cdots,\boldsymbol{\beta}_s$ 线性无关。

定义 14.5 线性空间 V 的一个子集 Π 称为 V 的一个基底,如果

(1) Π 中任意有限个元均线性无关;

(2) V 中任意元都能表成 Π 中有限个元的线性组合。

定理 14.6 若线性空间 V 有一个由有限个元组成的基底 $\{\boldsymbol{\alpha}_1,\cdots,\boldsymbol{\alpha}_n\}$,则

(1) 任意 $\boldsymbol{\alpha}\in V$,存在 $a_1,\cdots,a_n\in K$ 使

$$\alpha=(\alpha_1,\cdots,\alpha_n)\begin{pmatrix}a_1\\ \vdots\\ a_n\end{pmatrix},$$

$\begin{pmatrix}a_1\\ \vdots\\ a_n\end{pmatrix}$ 称为 α 在基 $\alpha_1,\cdots,\alpha_n$ 下的坐标,坐标是惟一的;

(2) V 中任意 $(n+1)$ 个元必线性相关;

(3) 若 Π 是 V 的一个基底,则 Π 中元素个数也为 n。此时,称 V 为 n 维线性空间,记为 $[V:K]=n$;

(4) 若 $\beta_1,\cdots,\beta_n$ 线性无关,则 $\beta_1,\cdots,\beta_n$ 是 V 的基底。

证 (1) 由定义,知存在 $a_1,\cdots,a_n\in K$,使 $\alpha=\sum_{i=1}^{n}a_i\alpha_i$,若 α 又 $=\sum_{i=1}^{n}b_i\alpha_i$,则 $\sum_{i=1}^{n}(a_i-b_i)\alpha_i=0$。因为 $\alpha_1,\cdots,\alpha_n$ 线性无关,所以 $a_i-b_i=0$, $a_i=b_i(i=1,\cdots,n)$。

(2) 直接利用命题 1 (因为 $n+1>n$)。

(3) 由(2)Π 中元素个数必$\leqslant n$。反之,把 Π 看成基底:

$$\Pi\Rightarrow\{\alpha_1,\cdots,\alpha_n\},$$

必 $n\leqslant\Pi$ 的阶。故 Π 中元素个数为 n。

(4) 任意 $\alpha\in V$,由(2),$\beta_1,\cdots,\beta_n,\alpha$ 线性相关,故

$$\beta_1,\cdots,\beta_n\Rightarrow\alpha,$$

即 $\beta_1,\cdots,\beta_n$ 是 V 的基底。

定义 14.6 若线性空间没有阶为有限的基底,则称 V 为无限维线性空间。否则称有限维线性空间。

例如,$\boldsymbol{C}$ 是 R 的二维线性空间,$\{1,\boldsymbol{i}\}$是它的一个基底。四元数体是 $\boldsymbol{Q}$(或 R)上的四维线性空间,$\{\boldsymbol{e},\boldsymbol{i},\boldsymbol{j},\boldsymbol{k}\}$是它的基底。全矩阵环 $(\boldsymbol{K})_n$ 是 K 上的 n^2 维线性空间,$\{\boldsymbol{E}_{ij}\mid i,j=1,\cdots,n\}$ 是它的一个基底。多项式环 $\boldsymbol{K}[x]$是 K 上的无限维线性空间,$\{\boldsymbol{x}^i\mid i$ 是非负整数$\}$是它

的一个基底。

定理 14.7 线性空间 V 必有基底，且任意两组基底等势。

证 设 $\Sigma=\{A \mid A\subseteq V,\ A$ 的任意有限个元素线性无关$\}$，则 Σ 关于集合的包含关系成一偏序集。Σ 中非空有序子集有上界(并集)。按 Zorn 引理，Σ 有极大元 M。$M\in\Sigma$，M 中任意有限个元线性无关。又，任意 $\beta\in V$，$M\cup\{\beta\}\overline{\in}\Sigma$，因而存在 M 中有限个元 $\alpha_1, \cdots, \alpha_t$ 使 $\alpha_1, \cdots, \alpha_t, \beta$ 线性相关。因 $\alpha_1, \cdots, \alpha_t$ 线性无关，所以 $\alpha_1, \cdots, \alpha_t\Rightarrow\beta$，故 M 是 V 的基底。

若 X 和 Y 都是 V 的基底。如果 X 和 Y 中有一个是有限集，由定理 14.6，$|X| = |Y|$，故可设两者都是无限集，任意 $\beta\in Y$。β 是 X 中有限个元 $\alpha_1, \cdots, \alpha_n$ 的线性组合。若 $\alpha_1, \cdots, \alpha_n$ 能线性表示的 Y 中元为 $\beta_1, \cdots, \beta_t$(由命题，$t\leqslant n$)，其中 $\beta_j=\beta$，则映射 f: $f(\beta)=((\alpha_1, \cdots, \alpha_n), j)$ 是 Y 到 $A\times Z^+$ 的单射。这里 $A=\bigcup\limits_{n\in P}X^n$。从而 $|Y|\leqslant|A\times Z^+|=|A|=|X|$。同理 $|X|\leqslant|Y|$，故 $|X|=|Y|$。

二、有限扩域

若 F 是域 K 的子域，K 是 F 的扩域，K 看成 F 上的线性空间，若是有限维的，则称 K 是 F 的有限扩域。$n=[K:F]$ 称为 K 关于 F 的扩张次数。K 称为 F 的 n 次扩域。

$[K:F]=1$，$\{e\}$是 K 关于 F 的基底，故 $K=F$，反之也对。所以 $[K:F]=1\Leftrightarrow K=F$。

若 $K=F(\alpha)$，α 是 F 的代数元，它满足的不可约多项式 $p(x)$ 次数为 n，则 $[K:F]=n$。因为此时，α 不满足任何 $F[x]$中次数小于 n 的多项式，故 $1, \alpha, \cdots, \alpha^{n-1}$线性无关，且 K 中元都能表成它们的线性组合，所以 $1, \alpha, \cdots, \alpha^{n-1}$是 K 关于 F 的基底。

定理 14.8 α 是域 F 的代数元，则单纯代数扩域 $F(\alpha)$是 F 的有限扩域，扩张次数$[K:F]$等于 α 所满足的 $F[x]$中不可约多项式的次数。

例 14.4 C 是 R 的二次扩域。$Q(\sqrt{2})$也是 Q 的二次扩域。

定理 14.9 若 $\boldsymbol{K}$ 是域 $\boldsymbol{F}$ 的有限扩域，$\boldsymbol{\Sigma}$ 是域 $\boldsymbol{K}$ 的有限扩域，则 $\boldsymbol{\Sigma}$ 是 $\boldsymbol{F}$ 的有限扩域，且

$$[\boldsymbol{\Sigma}:\boldsymbol{F}]=[\boldsymbol{\Sigma}:\boldsymbol{K}]\cdot[\boldsymbol{K}:\boldsymbol{F}]。$$

证 设 $[\boldsymbol{\Sigma}:\boldsymbol{K}]=s$，$\boldsymbol{\Sigma}$ 关于 $\boldsymbol{K}$ 有一基底 $\beta_1, \cdots, \beta_s$，$[\boldsymbol{K}:\boldsymbol{F}]=t$，$\boldsymbol{K}$ 关于 $\boldsymbol{F}$ 有一基底 $\alpha_1, \cdots, \alpha_t$。下面证明

$$\{\alpha_i\beta_j \mid i=1, \cdots, t;\ j=1, \cdots, s\}$$

是 $\boldsymbol{\Sigma}$ 关于 $\boldsymbol{F}$ 的一组基底。

(1) 任意 $\alpha\in\boldsymbol{\Sigma}$，存在 $k_1, \cdots, k_s\in\boldsymbol{K}$，使

$$\alpha=\sum_{j=1}^{s}k_j\beta_j,$$

因为 $k_j\in\boldsymbol{K}$，故存在 $l_{1j}, \cdots, l_{tj}\in\boldsymbol{F}$，使

$$k_j=\sum_{i=1}^{t}l_{ij}\alpha_i,\ j=1, \cdots, s$$

于是

$$\alpha=\sum_{i=1}^{t}\sum_{j=1}^{s}l_{ij}\alpha_i\beta_j,$$

即 α 是系数在 $\boldsymbol{F}$ 中的 $\boldsymbol{\Sigma}$ 的 (ts) 个元素 $\{\alpha_i\beta_j \mid i=1, \cdots, t;\ j=1, \cdots, s\}$ 的线性组合，或

$$\{\alpha_i\beta_j \mid i=1, \cdots, t;\ j=1, \cdots, s\}\Rightarrow\alpha。$$

(2) 若 $x_{ij}\in F$，$i=1, \cdots, t$；$j=1, \cdots, s$，使

$$\sum_i\sum_j x_{ij}\alpha_i\beta_j=\boldsymbol{0},$$

于是

$$\sum_{j=1}^{s}k_j\beta_j=0,$$

这里

$$k_j = \sum_{i=1}^{t} x_{ij}\alpha_i \in \boldsymbol{K}。$$

因为 $\beta_1, \cdots, \beta_s$ 是 $\boldsymbol{\Sigma}$ 关于 $\boldsymbol{K}$ 一组基，故（关于 $\boldsymbol{K}$）线性无关，从而

$$k_j = \sum_{i=1}^{t} x_{ij}\alpha_i = 0,\ (j = 1, \cdots, s),$$

又，$\alpha_1, \cdots, \alpha_t$ 是 $\boldsymbol{K}$ 关于 $\boldsymbol{F}$ 一组基（关于 $\boldsymbol{F}$）线性无关，故

$$x_{ij} = 0,\ (i = 1, \cdots, t;\ j = 1, \cdots, s)。$$

即 $\{\alpha_i\beta_j \mid i = 1, \cdots, t;\ j = 1, \cdots, s\}$ 线性无关。

由(1)，(2) 可知

$$\{\alpha_i\beta_j \mid i = 1, \cdots, t;\ j = 1, \cdots, s\}$$

是 $\boldsymbol{\Sigma}$ 关于 $\boldsymbol{F}$ 一组基，故

$$[\boldsymbol{\Sigma} : \boldsymbol{F}] = ts = [\boldsymbol{\Sigma} : \boldsymbol{K}][\boldsymbol{K} : \boldsymbol{F}]。$$

例 14.5 若 $\boldsymbol{F}$，$\boldsymbol{K}$ 都是域 $\boldsymbol{\Sigma}$ 的子域，$\boldsymbol{F} \subseteq \boldsymbol{K}$，$\alpha \in \boldsymbol{\Sigma}$，且 $[\boldsymbol{F}(\alpha) : \boldsymbol{F}] = n$，$[\boldsymbol{K} : \boldsymbol{F}] = s$，$(n, s) = 1$，求证：

$$[\boldsymbol{K}(\alpha) : \boldsymbol{K}] = n。$$

证 因为

$$\boldsymbol{F} \subseteq \begin{matrix} \boldsymbol{K} \\ \boldsymbol{F}(\alpha) \end{matrix} \subseteq \boldsymbol{K}(\alpha),$$

所以，由定理 14.9：

$$[\boldsymbol{K}(\alpha) : \boldsymbol{K}][\boldsymbol{K} : \boldsymbol{F}] = [\boldsymbol{K}(\alpha) : \boldsymbol{F}(\alpha)][\boldsymbol{F}(\alpha) : \boldsymbol{F}],$$

故

$$n = [\boldsymbol{F}(\alpha) : \boldsymbol{F}] \mid [\boldsymbol{K}(\alpha) : \boldsymbol{K}] \cdot s,$$

因 $(n, s) = 1$，所以

$$n \mid [\boldsymbol{K}(\alpha):\boldsymbol{K}] \leqslant [\boldsymbol{F}(\alpha):\boldsymbol{F}] = n。$$

因而$[\boldsymbol{K}(\alpha):\boldsymbol{K}] = n$。

三、代数扩域

定义 14.7 K 是域 F 的扩域，若 K 中的元都是 F 的代数元，则称 K 为域 F 的代数扩域。

若 $\boldsymbol{\Sigma}$ 是域 F 的有限扩域，$[\Sigma:F] = n$，任意 $\alpha \in \Sigma$，则

$$1, \alpha, \alpha^2, \cdots, \alpha^{n-1}, \alpha^n$$

关于 F 线性相关，即存在 $a_0, a_1, \cdots, a_n \in F$，且不全为 0 使

$$a_0 + a_1\alpha + \cdots + a_n\alpha^n = 0。$$

故 α 满足 $F[x]$ 中非零多项式：

$$f(x) = a_0 + a_1x + \cdots + a_nx^n,$$

从而 α 是 F 的代数元，因此 $\boldsymbol{\Sigma}$ 是 F 的代数扩域。故有：

定理 14.10 有限扩域是代数扩域。

定理 14.11 代数扩域的代数扩域仍是代数扩域。

证 若 K 是 F 的代数扩域，$\boldsymbol{\Sigma}$ 是 K 的代数扩域。现要证 $\boldsymbol{\Sigma}$ 是 F 的代数扩域，即任意 $\alpha \in \boldsymbol{\Sigma}$，$\alpha$ 是 F 的代数元。其证如下：

因 $\boldsymbol{\Sigma}$ 是 K 的代数扩域，故 α 是 K 的代数元。设 α 是 $K[x]$中不可约多项式

$$p(x) = a_0 + a_1x + \cdots + a_nx^n$$

的零点。记

$$M = F(a_0, a_1, \cdots, a_n)(\alpha) = F(a_0)(a_1)\cdots(a_n)(\alpha)。$$

因 α 是 $F(a_0)(a_1)\cdots(a_n)$的代数元，故 M 是 $F(a_0)\cdots(a_n)$的单纯代数扩域，因而是有限扩域。

$a_n \in K$，a_n 是 F 的代数元，a_n 是 $F[x]$中（因而也是 $F(a_0, \cdots, a_{n-1})[x]$中）某非零多项式零点，所以 $F(a_0)\cdots(a_n)$是 $F(a_0)\cdots(a_{n-1})$的

单纯代数扩域，因而是有限扩域。类似地，$F(a_0)\cdots(a_{i-1})(a_i)$是$F(a_0)\cdots(a_{i-1})$的有限扩域 $(i=1,\cdots,n)$，$F(a_0)$是 F 的有限扩域。

由定理 14.9 可知，有限扩域的有限扩域仍是有限扩域，故 M 是 F 的有限扩域。由定理 14.10，M 是 F 的代数扩域，因此 M 的元素 α 是 F 的代数元。

推论 域 F 的代数元 α,β 的和，差，积，商 $\alpha+\beta$，$\alpha-\beta$，$\alpha\times\beta$，$\alpha\times\beta^{-1}$ $(\beta\neq 0)$ 仍是 F 的代数元。

证 因为 $F(\alpha,\beta)=F(\alpha)(\beta)$ 是 F 的代数扩域。

§4 代数闭包与分裂域

一、代数闭包

定义 14.8 没有真代数扩域的域称为代数闭域。

定理 14.12 域 Σ 是代数闭域 $\Leftrightarrow \Sigma[x]$ 中不可约多项式都是一次多项式。

证 $\Rightarrow$ 若 $p(x)$ 是 $\Sigma(x)$ 中次数 >1 的不可约多项式，则 $\Sigma[x]/\langle p(x)\rangle$ 是 Σ 的真代数扩域，故不可能。

$\Leftarrow$ 若 K 是 Σ 的代数扩域，任意 $\alpha\in K$，α 满足 $\Sigma[x]$ 中不可约多项式 $x-c$，故 $\alpha-c=0$，$\alpha=c\in\Sigma$，即 $K=\Sigma$，从而 Σ 没有真代数扩域。

例 14.6 因为 $C[x]$中不可约多项式都是一次的，故复数域 C 是代数闭域。

例 14.7 记 $\Sigma=\{\alpha\in C\mid \alpha$ 是 Q 的代数元$\}$，则 Σ 是代数闭域，即 Q 上代数数全体所成集合是代数闭域。

证 因 Q 上代数元的和，差，积，商是代数元，故 Σ 是 C 的子域。若 Σ 有代数扩域 K，由于 Σ 是 Q 的代数扩域，所以 K 是 Q 的代数扩域，即 K 中元是 Q 的代数元，因而$\in\Sigma$，故 $K\subseteq\Sigma$，即 Σ 没有真代数扩域，Σ 是代数闭域。

定理 14.13 设 F 是任意域，则存在 F 的代数扩域 Σ，Σ 是代数闭域，Σ 称为 F 的代数闭包。

证 记 $N=\{E|E$ 是 F 的代数扩域$\}$，因为 $F\in N$，故 N 不是空集。

在 N 中定义关系"$\leqslant$"：$E_1\leqslant E_2\Leftrightarrow E_1$ 是 E_2 的子域。易验证，这个关系是偏序关系。由于 N 中非空有序子集的并仍 $\in N$，而这个并是这个非空有序子集的上界，从而偏序集 N 满足 Zorn 公理的条件，故有极大元 Σ。

$\Sigma\in N$，Σ 是 F 的代数扩域．Σ 是极大元，所以 Σ 无真代数扩域，即 Σ 是 F 的代数闭包。

若 σ 是域$\triangle$到域 $\sigma(\triangle)$的同构映射。如果在$\triangle[x]$中有等式

$$\sum_{i=0}^{s+t}a_ix^i=\left(\sum_{j=0}^{s}b_jx^j\right)\left(\sum_{k=0}^{t}c_kx^k\right),$$

因为 $a_i=\sum\limits_{j+k=i}b_jc_k$，故 $\sigma(a_i)=\sum\limits_{j+k=i}\sigma(b_j)\sigma(c_k)$。从而在 $\sigma(\triangle)[x]$ 中有相应等式：

$$\sum_{i=0}^{s+t}\sigma(a_i)x^i=\left(\sum_{j=0}^{s}\sigma(b_j)x^j\right)\left(\sum_{k=0}^{t}\sigma(c_k)x^k\right)。$$

由此，可得以下几个命题：

命题 14.2 若 σ 是域$\triangle$到域 F 的同构映射。如果在$\triangle[x]$中：

$$\sum_{j=0}^{s}b_jx^j\Bigm|\sum_{i=0}^{n}a_ix^i,$$

则在 $F[x]$ 中：$\sum\limits_{j=0}^{s}\sigma(b_j)x^j\Bigm|\sum\limits_{i=0}^{n}\sigma(a_i)x^i$。

命题 14.3 若 σ 是域$\triangle$到域 F 的同构映射，则 $\sum\limits_{i=0}^{n}a_ix^i$ 是$\triangle[x]$的不可约多项式 $\Leftrightarrow\sum\limits_{i=0}^{n}\sigma(a_i)x^i$ 是 $F[x]$的不可约多项式。

命题 14.4 代数闭域的同构像仍是代数闭域。

证 若 σ 是代数闭域 Σ 到域 F 的同构映射。$F[x]$中多项式

$\sum_{i=0}^{n}\sigma(a_i)x^i$ 不可约 $\Leftrightarrow\Sigma[x]$ 中多项式 $\sum_{i=0}^{n}a_ix^i$ 不可约 $\Leftrightarrow n=1$，$a_1\neq 0$（因为 Σ 是代数闭域）$\Leftrightarrow n=1$，$\sigma(a_1)\neq 0\Leftrightarrow\sum_{i=0}^{n}\sigma(a_i)x^i$ 是一次多项式，故 F 是代数闭域。

定理 14.14 若给了域 F 到 F_1 的同构映射 σ，Σ 和 Σ_1 分别是 F 和 F_1 的代数闭包，则存在 Σ 到 Σ_1 的同构映射 η，对任意 $a\in F$ 均有

$$\eta(a)=\sigma(a)。$$

证 记 $T=\{(E,f)\mid E$ 是 Σ 的子域，$E\supseteq F$，f 是 E 到 Σ_1 的单同态，且任意 $a\in F$，有 $f(a)=\sigma(a)\}$，因为 $(F,\sigma)\in T$，故 T 不是空集。易验证，T 对二元关系"$\leqslant$"成一满足 Zorn 引理条件的偏序集，这里 $(E,f)\leqslant(E_1,f_1)\Leftrightarrow E$ 是 E_1 的子域，且任意 $a\in E$ 有 $f_1(a)=f(a)$。设 (M,f) 是 T 的极大元，若 $M\neq\Sigma$，则存在 $\alpha\in\Sigma$，$\alpha\overline{\in}M$，$M(\alpha)$ 是 M 的真代数扩域。设 α 是 $M[x]$ 中不可约多项式 $p(x)=\sum_{i=0}^{n}a_ix^i$ 的根，由命题 14.3，$p^*(x)=\sum_{i=0}^{n}f(a_i)x^i$ 是 $f(M)[x]$ 中不可约多项式。设 β 是其在 Σ_1 中一个根，作对应

$$M(\alpha)\xrightarrow{f_1}\Sigma_1:\ f_1(\Sigma a_ix^i)=\Sigma f(a_i)\beta^i,\ a_i\in M,$$

因为 $\Sigma a_i\alpha^i=\Sigma b_i\alpha^i\Leftrightarrow\Sigma(a_i-b_i)\alpha^i=0\Leftrightarrow$（因为 $p(x)$ 是 α 的最小多项式）$p(x)\mid\Sigma(a_i-b_i)x^i\Leftrightarrow$（命题 14.2）$p^*(x)\mid\Sigma f(a_i-b_i)x^i\Leftrightarrow$（因为 $p^*(x)$ 是 β 的最小多项式）$\Sigma f(a_i-b_i)\beta^i=0\Leftrightarrow\Sigma f(a_i)\beta^i=\Sigma f(b_i)\beta^i$，故 f_1 是单射，显然 f_1 是同态，从而 f_1 是 $M(\alpha)$ 到 Σ_1 的单同态，且任意 $a\in M$，$f_1(a)=f(a)$，即 $(M,f)<(M(\alpha),f_1)$，这矛盾于 (M,f) 是 T 的极大元的假设，故必 $M=\Sigma$。

由命题 14.4，$f(\Sigma)$ 是代数闭域，无真代数扩域，必 $f(\Sigma)=\Sigma_1$。

由定理 14.12 和 14.13 即得：任意域都有代数闭包，同构的域的代数闭包同构。

二、分裂域

定义 14.9 F 是域，$f(x)$是 $F[x]$中次数$\geqslant 1$ 的多项式。F 的扩域 K 称为 $f(x)$关于 F 的分裂域，如果：

(1) $K = F(\alpha_1, \cdots, \alpha_n)$；

(2) $f(x) = a_0(x - \alpha_1)\cdots(x - \alpha_n)$。

设 Σ 是 F 的代数闭包，$f(x) \in F[x] \subseteq \Sigma[x]$。$\Sigma[x]$ 中不可约多项式全是一次的，由因式分解惟一定理

$$f(x) = a_0(x - \alpha_1)\cdots(x - \alpha_n),$$

这里 $\alpha_1, \cdots, \alpha_n \in \Sigma$，则

$$K = F(\alpha_1, \cdots, \alpha_n)$$

为 $f(x)$关于 F 的分裂域。由于代数闭包在同构意义下是惟一的，因式分解是惟一的，从而 $\alpha_1, \cdots, \alpha_n$ 由 $f(x)$惟一确定，故分裂域也是惟一的。

定理 14.15 分裂域存在且(在同构意义下)惟一。

例 14.8 $f(x) = x^2 - 2$ 关于 Q 的分裂域是 $Q(\pm\sqrt{2}) = Q(\sqrt{2}) = \{a + b\sqrt{2} \mid a, b \in Q\}$。

例 14.9 求 $x^p - 1$ 关于 Q 的分裂域(p 是素数)。

解 因为 $f(x) = x^p - 1$ 和 $f'(x) = px^{p-1}$ 互素，故 $f(x)$无重根。$f(x)$的 p 个根成一乘法群。因阶为 p 的群是循环群，非单位元都是群的生成元。$f(x) = (x - 1)(x^{p-1} + \cdots + 1)$，记 α 为不可约多项式 $x^{p-1} + \cdots + 1$ 的一个根，则 $f(x)$关于 Q 的分裂域为

$$F(\alpha, \alpha^2, \cdots, \alpha^p) = F(\alpha) \cong F[x]/\langle x^{p-1} + \cdots + 1\rangle。$$

例 14.10 若域 F 的特征为 p，$F[x]$中多项式 $f(x) = x^p - x - a$ 有一个根为 α，则 $f(x)$关于 F 的分裂域为 $F(\alpha)$。

证 $f(\alpha) = 0 \Rightarrow \alpha^p - \alpha - a = 0$，故

$$f(x) = x^p - x - (\alpha^p - \alpha) = (x - \alpha)^p - (x - \alpha)。$$

因为 Z_p 中非零元成一个阶为 $p-1$ 的乘法群，故非零元满足 $y^{p-1}=e$，从而 Z_p 中元都满足 $y^p=y$，即 $y^p-y=0$。Z_p 中 p 个元为 $0, e, 2e, \cdots, (p-1)e$，所以

$$y^p-y=y(y-e)(y-2e)\cdots(y-\overline{p-1}e),$$

故

$$f(x)=(x-\alpha)(x-\alpha-e)\cdots(x-\alpha-\overline{p-1}e),$$

即 $f(x)$ 的所有零点为 $\alpha, \alpha+e, \alpha+2e, \cdots, \alpha+(p-1)e$。因此，$f(x)$ 关于 F 的分裂域为

$$F(\alpha, \alpha+e, \cdots, \alpha+(p-1)e)=F(\alpha)。$$

例 14.11 若 F, K 都是域 Σ 的子域，$F\subseteq K$，$\alpha\in\Sigma$。$[F(\alpha):F]=n$，$[K:F]=s$，$(n, s)=1$，求证：$[K(\alpha):K]=n$。

证 因为 $F\subseteq \begin{matrix}K\\F(\alpha)\end{matrix}\subseteq K(\alpha)$。所以，由定理 14.9：

$$[K(\alpha):K][K:F]=[K(\alpha):F(\alpha)][F(\alpha):F],$$

故

$$n=[F(\alpha):F]\mid[K(\alpha):K]\cdot s,$$

因 $(n, s)=1$，所以

$$n\mid[K(\alpha):K]\leqslant[F(\alpha):F]=n。$$

因而

$$[K(\alpha):K]=n。$$

例 14.12 q, p 是素数，若 K 是 $f(x)=x^p-q$ 关于 Q 的分裂域，求 $[K:Q]$。

解 若 α 是 $f(x)$ 的一个根，β 是 x^p-1 的一个根，$\beta\neq 1$

$$f(x)=x^p-\alpha^p=\alpha^p\left[\left(\frac{x}{\alpha}\right)^p-1\right],$$

故 $f(x)$ 的 p 个根为

$$\alpha,\ \alpha\beta,\ \alpha\beta^2,\ \cdots,\ \alpha\beta^{p-1},$$

所以

$$K = Q(\beta,\ \alpha),\ F = Q(\beta) \cong Q[x]/\langle x^{p-1} + \cdots + 1\rangle。$$

又，$[Q(\alpha):Q] = f(x)$ 次数 $= p$，$[F:Q] = p-1$，$(p,\ p-1) = 1$，由例 14.6，$[K:F] = p$，故

$$[K:Q] = [K:F][F:Q] = p(p-1)。$$

§5 有 限 域

一、有限域的阶

定义 14.10 含有限个元素的域称为有限域。

若 $\boldsymbol{F}$ 是有限域，$\mathrm{ch}\ F = p$，因而 $\boldsymbol{F}$ 是 Z_p 的有限扩域。可设 $[\boldsymbol{F}:Z_p] = n$，即 $\boldsymbol{F}$ 是 Z_p 上的 n 维线性空间，若 $\alpha_1,\ \cdots,\ \alpha_n$ 是 $\boldsymbol{F}$ 关于 Z_p 的一组基，则

$$F = \left\{\sum_{i=1}^{n} k_i\boldsymbol{\alpha}_i \,\middle|\, k_i \in Z_p,\ i = 1,\ \cdots,\ n\right\},$$

于是 $|\boldsymbol{F}| = p^n$，即有限域 F 的阶必是素数 p 的一个幂。

二、有限域的结构与特征方程

给了素数 p 的一个幂 p^n（n 为正整数），是否存在阶为 p^n 的有限域？若有，是否惟一？两个回答都是肯定的。下面我们给出具体推导。

设 Σ 是

$$f(x) = x^{p^n} - x$$

关于 Z_p 的分裂域。若 α,β 是 $f(x)$ 的两个零点，则

$$f(\alpha-\beta)=(\alpha-\beta)^{p^n}-(\alpha-\beta)=(\alpha^{p^n}-\alpha)-(\beta^{p^n}-\beta)=0,$$

当 $\beta\neq 0$ 时，则

$$\begin{aligned}f(\alpha\beta^{-1})&=(\alpha\beta^{-1})^{p^n}-\alpha\beta^{-1}=\alpha^{p^n}\cdot(\beta^{p^n})^{-1}-\alpha\beta^{-1}\\&=\alpha\beta^{-1}-\alpha\beta^{-1}=0,\end{aligned}$$

即 $\alpha-\beta,\alpha\beta^{-1}$ 也是 $f(x)$ 的零点，因而 $F=\{\alpha\in\Sigma\mid f(x)=0\}$ 是 Σ 子域。但 $F\supseteq Z_p$ 和 $f(x)$ 的所有零点，故 $\Sigma=F$，F 的元素个数等于 $f(x)$ 的零点个数，但$(f(x),\ f'(x))=(f(x),\ -1)=1$，所以 $f(x)$ 没有重零点，因而 $|F|=\deg f(x)=p^n$，即阶为 p^n 的域 F 是存在的。又，F 是阶为 p^n 的域，F 中非零元成一阶为 p^n-1 的乘法群，任意非零元 a 满足 $a^{p^n-1}=1$，从而 F 中元都是 $x^{p^n}-x$ 的零点，即 F 是 $x^{p^n}-x$ 关于 Z_p 的分裂域，而分裂域在同构意义下是惟一的。故有：

定理 14.16 有限域的阶是素数的幂 p^n。反之，任给素数的一个幂 p^n，存在一个且（在同构意义下）仅存在一个阶为 p^n 的域。F 是阶为 p^n 的域，则任意 $\alpha\in F$，$\alpha^{p^n}=\alpha$。称方程

$$x^{p^n}-x=0$$

为有 p^n 阶的有限域 F 的特征方程。

三、有限域的子域

若 F 是有限域，$|F|=p^n$。讨论 F 的子域 M 的结构。M 也是 Z_p 的扩域，如果 $[M:Z_p]=d$，则 $|M|=p^d$。又，$[F:Z_p]=[F:M][M:Z_p]$，故 $d\mid n$。反之，任给 n 的一个正因子 d，因 $p^d-1\mid p^n-1$，所以 $x(x^{p^d-1}-1)\mid x(x^{p^n-1}-1)$，因而 $x(x^{p^d-1}-1)=x^{p^d}-x$ 的零点 $\in F$，$x^{p^d}-x$ 的零点全体成 F 的一个阶为 p^d 的子域。由于阶为 p^d 的域的元素全是 $x^{p^d}-x$ 的零点，故 F 有且仅有一个阶为 p^d 的子域。

定理 14.17 若 F 是有限域，$|F|=p^n$（p 是素数），则 F 的子域阶为 p^d，这里 d 是 n 的正因子。反之任给 n 的一个正因子 d，F 有且仅有

一个阶为 p^d 的子域。

四、有限域 F 的非零元乘法群

设 $|F| = q$，表 $q-1=n$ 为素数幂的积：

$$n = p_1^{\alpha_1}\cdots p_s^{\alpha_s},$$

这里：p_1，…，p_s 为两两不同的素数。

记 $G = F - \{0\}$ 为 n 阶乘法群。在 G 中：

$$x^{\frac{n}{p_i}} - 1$$

的零点个数 $\leqslant \dfrac{n}{p_i} < n$，故 G 中存在 β_i 使

$$\beta_i^{\frac{n}{p_i}} - 1 \neq 0,$$

即

$$\beta_i^{\frac{n}{p_i^{\alpha_i}}} \neq 1。$$

记

$$u_i = \beta_i^{\frac{n}{p_i^{\alpha_i}}},\quad u_i^{p_i^{\alpha_i}} = \beta_i^n = 1$$

即 u_i 的阶是 $p_i^{\alpha_i}$ 的因子，但

$$u_i^{p_i^{\alpha_i-1}} = \beta_i^{\frac{n}{p_i}} \neq 1,$$

故 u_i 的阶为 $p_i^{\alpha_i}(i=1,\cdots,s)$。$u_1$，…，$u_s$ 的阶两两互素，且其乘法可交换，故

$$u = u_1\cdots u_s$$

的阶为 $p_1^{\alpha_1}\cdots p_s^{\alpha_s} = n$，即 $G = \langle u\rangle$ 为循环群。

定理 14.18 有限域的非零元素法群是循环群。

例 14.13 构造 9 个元素的域 F。

解　F 含素域 Z_3。$[F:Z_3]=2$。二次多项式 $p(x)=x^2-x-1$ 在 Z_3 中没有零点，所以在 $Z_3[x]$中无一次因子，从而 $p(x)$是 $Z_3[x]$的不可约多项式。故 $Z_3[x]/\langle p(x)\rangle$是含 9 个元素的域即为所求 F。

设 α 是 F 中 $p(x)$的一个零点，则 F 中 8 个非零元分别为：α，$\alpha^2=1+\alpha$，$\alpha^3=1+2\alpha$，$\alpha^4=\alpha$，$\alpha^5=2\alpha$，$\alpha^6=2+2\alpha$，$\alpha^7=2+\alpha$，$\alpha^8=1$。

例 14.14　在 $Z_2[x]$中分解多项式 x^8-x。

解　设 F 为 8 元域，则 F 是 x^8-x 关于 Z_2 的分裂域。任意 $\alpha\in F$，$Z_2\subseteq Z_2(\alpha)\subseteq F$，易得$[Z_2(\alpha):Z_2]=1$ 或 3，即 α 是 $Z_2[x]$ 中一次或三次不可约多项式的零点，从而 x^8-x 是 $Z_2[x]$ 中一次和三次不可约多项式的积。直接验证，可得：$Z_2[x]$ 中一次不可约多项式有两个：x 和 $x-1$；三次不可约多项式也有两个：x^3+x+1 和 x^3+x^2+1。从而

$$x^8-x=x(x-1)(x^3+x+1)(x^3+x^2+1)。$$

例 14.15　若 $f(x)$是 $Z_p[x]$的一个 n 次不可约多项式，α 为 $f(x)$ 的一个零点（$\alpha\in Z_p$ 的代数闭包），求证：$Z_p(\alpha)$ 为 $f(x)$ 关于 Z_p 的分裂域。

证　设 K 为 $f(x)$ 关于 Z_p 的分裂域，显然 $K\supseteq Z_p(\alpha)$。又，$|Z_p(\alpha)|=p^n$，$Z_p(\alpha)$ 的特征方程为

$$x^{p^n}-x=0,\alpha\in Z_p(\alpha)。$$

故

$$\alpha^{p^n}-\alpha=0,$$

但以 α 为零点的不可约多项式是 $f(x)$，必

$$f(x)\mid x^{p^n}-x,$$

即 $f(x)$ 的零点都是 $x^{p^n}-x$ 的零点，从而属于 $Z_p(\alpha)$，于是

$$K\subseteq Z_p(\alpha),$$

由

$$Z_p(\alpha)\subseteq K \text{ 及 } K\subseteq Z_p(\alpha),$$

得

$$K = Z_p(\alpha)。$$

例 14.16　若正整数 n 不能整除 2^n，求证多项式

$$f(x) = x^{2^n} + x + 1$$

在 $Z_2[x]$ 中可约。

证　用反证法。若 $f(x)$在 $Z_2[x]$中不可约，则

$$|Z_2(\alpha)| = 2^{2^n} = q,$$

这里 α 是 $f(x)$ 的一个零点，即 $f(\alpha) = 0$，$\alpha^{2^n} + \alpha = -1$，此时

$$\begin{aligned} f(x) &= x^{2^n} + x + 1 = x^{2^n} + x - (-1) \\ &= x^{2^n} + x - \alpha^{2^n} - \alpha \\ &= (x-\alpha)^{2^n} + (x-\alpha) \\ &= (x-\alpha)^{2^n} - (x-\alpha), \end{aligned}$$

因 $\alpha^q = \alpha$，即 α 是 $x^q - x$ 的零点，从而以 α 为零点的不可约多项式 $f(x)$ 能整除 $x^q - x$，即

$$f(x) = (x-\alpha)^{2^n} - (x-\alpha) \mid x^q - x,$$

于是 $f(x)$ 的根集 $\Pi \subseteq (x^q - x)$ 根集 $= Z_p(\alpha)$，进而 $y^{2^n} - y$ 根集

$$\Pi_1 = \{x - \alpha \mid x \in \Pi\} \subseteq Z_p(\alpha),$$

但 Π_1 为阶是 2^n 的有限域，即 $Z_p(\alpha)$ 有阶为 2^n 的子域，故必

$$n \mid 2^n,$$

这与假设矛盾，故必 $f(x)$ 在 $Z_2[x]$ 中可约。

习　题　14

1. 若域 K 的特征为 2，求证：K 中元素的平方和必是完全平方。

2. 若域 K 的特征为 p,求证:

(1) $\left(\sum_{i=1}^{s} a_i\right)^{p^n} = \sum_{i=1}^{s} a_i^{p^n}$, $a_i \in K$;

(2) $(a - b)^{p-1} = \sum_{i=0}^{p-1} a^i b^{p-1-i}$, $a,b \in K$。

3. 若域 K 的特征为 p, $\forall a \in K$ 有 $a^p = a$,求证: $K \cong Z_p$。

4. 验证: Q 和 Z_p 是素域。

5. 试表出 $Q(\sqrt{2}, \sqrt{3})$ 中元素。

6. 求证:映射 σ: $\sigma(a + b\mathrm{i}) = (a - b\mathrm{i})$ 是复数域 C 的自同构。记 $K = \{x \in C | \sigma(x) = x\}$,试确定 K。

7. 若 F,K,L 都是域 Σ 的子域,$F \subseteq K,L$, S 是 Σ 的子集,$K = F(S)$,求证: $F(K \cup L) = L(S)$。

8. 若域 K 的特征为 p,试在域 K 中解方程 $x^p - x = 0$。

9. ω 为 $x^2 + x + 1 = 0$ 的根,试在 $Q(w)$ 中表 $\frac{1}{\omega^3 + \omega}$ 为 ω 的一次多项式。

10. 若 α 为 $Q[x]$ 中不可约多项式

$$p(x) = x^3 - 6x + 2$$

的零点,试在 $Q(\alpha)$ 中表 $\frac{\alpha^2 - \alpha + 7}{2\alpha^2 + 3\alpha + 1}$ 为 α 的二次多项式。

11. 求证: $f(x) = x^3 + x + 1$ 是 $Z_2[x]$ 的不可约多项式。若它在 Z_2 的扩域中有根 α,则它的全部根为 α,α^2,α^4。

12. 若域 F 的特征为 p, $f(x)$ 是 $F[x]$ 中不可约多项式,求证: $f(x)$ 在 F 的扩域中有重根 $\Leftrightarrow f(x)$ 可表成 x^p 的多项式。

13. 若 $a + b\mathrm{i}$ 是 $Z[\mathrm{i}]$ 的素元,求证:存在素数 p 使 $Z[\mathrm{i}]/\langle a + b\mathrm{i}\rangle Z_p(\mathrm{i})$。这里 $Z_p(\mathrm{i})$ 是多项式 $x^2 + \bar{1} = \bar{0}$ 关于 Z_p 的分裂域,即添加 $x^2 + \bar{1} = \bar{0}$ 的根所得 Z_p 的扩域。

14. 若域 F 的特征为 p, $F[x]$ 中多项式 $x^p - a$ 和 $x^p - x - a$ 均在 F 中无根,求证: $x^p - a$ 和 $x^p - x - a$ 都是 $F[x]$ 的不可约多项式。

15. 若线性空间 V 中元 $\alpha_1,\cdots,\alpha_n$ 线性无关,而 $\alpha_1,\cdots,\alpha_n,\beta$ 线性相

关，则 $\alpha_1,\cdots,\alpha_n \Rightarrow \beta$。

16. 若线性空间 V 的维数为 n。$\alpha_1,\cdots,\alpha_n$ 是 V 的元素，任意 $\alpha\in V$，$\alpha_1,\cdots,\alpha_n\Rightarrow\alpha$，则 $\alpha_1,\cdots,\alpha_n$ 是 V 的基底。

17. 若 K 是有理数域 Q 的二次扩域，则存在 $a\in Z$ 使 $K=Q(\sqrt{a})$。

18. 若 p 为素数，$\alpha\neq1,\alpha^p=1$，求 $[Q(\alpha):Q]$。

19. 若 K 是 Z_p 的 n 次扩域，求 K 的元数 $|K|$。

20. 若 $F(\alpha)$ 是域 F 的奇数次单纯代数扩域，求证：$F(\alpha)=F(\alpha^2)$。

21. 若 F 是代数闭域 Σ 的子域，记 $K=\{\alpha\in\Sigma\mid\alpha$ 是 F 的代数元$\}$，则 K 是 F 的代数闭包。

22. 有限域不是代数闭域。

23. 若域 F 的特征为素数 p，$f(x)=x^p-x-a\in F[x]$，如果 α 为 $f(x)$ 在 F 的扩域中的一个零点，求证：$F(\alpha)$ 为 $f(x)$ 关于 F 的分裂域。

24. 设 $\alpha^p=1$，$\alpha\neq1$，这里 p 是素数，记 $K=Q(\alpha)$。若 β 为 $f(x)=x^p-2$ 的一个零点，求证：$K(\beta)$ 为 $f(x)$ 关于 K 的分裂域。

25. 若 K 是多项式 x^7-15 关于 Q 的分裂域，求 $[K:Q]$。

26. 求证：若 F 是域，则 $F[x]$ 中 n 次多项式的分裂域关于 F 的次数 $\leqslant n!$。

27. 若 p 是素数，试证：任意 $a\in Z$，$p\mid a^p-a$。

28. 若 $f(x)$ 是 $Z_p[x]$ 的一个 n 次不可约多项式。α 为 $f(x)$ 在其分裂域 K 中的一个零点，则 $f(x)$ 的全部零点为 $\alpha^p,\alpha^{p^2},\cdots,\alpha^{p^n}=\alpha$。

29. 分别构造 4 个元素和 8 个元素的域，并给出它们的加法表和乘法表。

30. 试在 $Z_3[x]$ 中分解多项式 x^9-x。在 $Z_2[x]$ 中分解多项式 $x^{16}-x$。

31. 有限域中每个元素均可表成两个元素的平方和。

32. 若域 F 的特征为 p，$a\in F$，x^p-a 在 F 中无根，$n\geqslant1$，求证：$x^{p^n}-a$ 是 $F[x]$ 的不可约多项式。

［Ｇｅｎｅｒａｌ　Ｉｎｆｏｒｍａｔｉｏｎ］
书名＝大学代数　陆少华，沈灏著，２００１
作者＝陆少华　　沈灏著
页数＝４８７
ＳＳ号＝１０４４９１５９
出版日期＝２００１年０４月第１版