本书分上、下两册，内容包括四部分：线性方程组，$n$ 元有序数组的向量空间和矩阵理论；一元多项式环和多元多项式环的理论；任意域上的线性空间和线性映射的理论；具有度量的线性空间包括欧几里得空间、酉空间、正交空间和辛空间的理论。

本书注意渗透现代的观点，力图体现代数与几何、分析的联系，注重联系实际，努力培养学生的数学素质。本书写得深入浅出，特别注意阐述清楚所讨论问题的想法。每一节后面都配有一定数量的习题。

本书可作为大学数学系、概率统计系和应用数学系的高等代数课程的教材，它也是一部内容丰富的教学参考书，可供有关师生参考。

# 下 册 前 言

本书上册讲了线性代数研究的具体对象：线性方程组和矩阵.上册内容包括：线性方程组的解法、行列式、$n$ 元有序数组的向量空间、线性方程组的理论、矩阵的运算、矩阵的相抵分类与相似分类、二次型与矩阵的合同分类.下册则讲多项式理论和线性代数研究的抽象对象：线性空间与线性映射.下册内容包括：一元多项式环与多元多项式环、线性空间、线性映射和线性变换、线性变换的Jordan 标准形、线性函数和双线性函数、欧几里得空间、酉空间、正交空间、辛空间.下册还结合高等代数研究的对象，水到渠成地引进了抽象代数的一些基本概念：群、环、域、域上的代数等.

面向 21 世纪，既要培养一批高水平的数学人才，又要培养一大批具有良好数学素质并在各个领域从事工作的人才.因此本书的内容分成三个层次，以适应培养各种类型人才的需要：

第一个层次是正文中用宋体字排印的不加“*”号的内容以及习题中不加“*”号的题.这些内容是作为教学的基本要求.

第二个层次是用楷体字排印的内容，它们不作为教学的基本要求.如果教学时数较充裕，可以讲；否则可以不讲.这些内容包括第七章 §9，§12；第八章 §7 的后半部分，§8；第十三章 §3，§4；第十四章 §1，§4 等.

第三个层次是正文中加“*”号的内容和阅读材料，它们不作为教学要求，是供学有余力的学生自学的，以便培养他们的阅读理解能力和开阔眼界.这些内容都是紧密配合正文的，为想看课外参考书的学生提供了方便.正文中加“*”号的内容包括第七章 §13，第十一章 §5，第十三章 §5，第十四章 §2，§3 等.

习题中加“*”号的题以及补充题不作为教学基本要求，供学

生选做.

在使用本书时，各个学校的教师还可根据具体情况作一些变通. 例如，本书讲的是任意域上的线性空间和线性映射，但是也可以只讲数域上的线性空间和线性映射，这时第七章 §14 可以不讲. 又如，本书讲线性变换的 Jordan 标准形时，采用空间分解的方法，但是也可以采用 $\lambda$- 矩阵的理论，这时第十章 §4 不用讲，而讲阅读材料三、五、七.

本书作为大学数学系、概率统计系和应用数学系的高等代数课程的教材，大体上，上册供第一学期使用，下册供第二学期使用. 如果第一学期教学时数较充裕，则下册第七章的前几节也可以放在第一学期讲.

本书的责任编辑胡乃冏同志为本书的编辑出版付出了辛勤劳动，作者向他表示衷心感谢.

热诚欢迎同行和读者对书中的疏漏之处提出批评指正.

丘维声

1996 年 8 月于北京大学燕北园

# 目　录

# 第七章　一元多项式环与多元多项式环

代数方程曾经是代数学研究的中心问题，它包括线性方程组与高次方程. 研究高次方程也就是研究多项式. 在前几章我们详细阐述了线性方程组的理论. 这一章我们将介绍多项式的基本理论. 多项式不仅是代数学的一个传统分支，而且在以研究代数结构为中心的近世代数学里占有重要位置. 在数学的其他分支以及工程技术中也常常用到多项式. 譬如，在数学分析中用多项式函数逼近一般的函数；在组合数学中，用多项式作为计数的生成函数；在代数几何中，多项式方程组的解集是基本研究对象；在计算机科学中，有多项式算法这一概念；等等.

## §1　一元多项式的概念及性质 · 环的基本概念

### 1.1　一元多项式的概念及其运算

读者在中学数学里已经学过多项式，那个时候是把多项式定义成几个单项式的代数和，而单项式被定义成用乘法(包括乘方)符号把数字或表示数字的字母连结而成的式子. 读者在数学分析课程中已学过多项式函数，它是由幂函数(幂指数为正整数)和常数函数经过有限次加、减、乘法运算所得的函数，即形如

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

的函数，其中 $a_i (i = 0, 1, \cdots, n)$ 是实常数，$x$ 是变量. 因此，迄今为止，读者对多项式的认识是形如(1) 那样的表达式，其中 $a_i (i = 0, 1, \cdots, n)$ 是实(或复) 常数，$x$ 是变量，$x$ 可取任意实数(或复数).

现在我们要把多项式的概念加以推广，使其有更广泛的应用.除了用任意数域 $K$ 代替实数域(或复数域)以外，最主要的推广是：$x$ 不仅能取数域 $K$ 中任意数，而且还能用其他一些满足一定要求的元素代入(详见本节后半部分的"一元多项式环的通用性质").

**定义 1** 设 $K$ 是一个数域.用一个不属于 $K$ 的符号 $x$ 作表达式

$$a_nx^n + a_{n-1}x^{n-1} + \cdots + a_1x + a_0 \tag{2}$$

其中 $n$ 是任意非负整数，系数 $a_i(i = 0,1,\cdots,n)$ 属于 $K$；两个这样的表达式相等规定为它们除去系数为零的项外含有完全相同的项，而系数为零的项是允许任意删去和添进来的；这种表达式称为**系数在数域 $K$ 中的一元多项式**，或者称为**数域 $K$ 上的一元多项式**，符号 $x$ 称为**不定元**.

系数全为零的多项式称为**零多项式**，记为 0.

在多项式(2)中，$a_ix^i$ 称为 $i$ **次项**，$a_i$ 称为 $i$ **次项的系数**，$i = 0,1,\cdots,n$.零次项 $a_0x^0$ 简记作 $a_0$，也称为**常数项**.

从定义 1 知道，数域 $K$ 上的两个一元多项式相等当且仅当它们的同次项的系数都相等.换句话说，一个一元多项式的表达法是唯一的(除了系数为 0 的项以外).

今后，我们用 $f(x),g(x),\cdots$ 或 $f,g,\cdots$ 等来代表一元多项式.

设 $f(x)$ 代表多项式(2).如果 $a_n \neq 0$.那么 $a_nx^n$ 称为多项式 $f(x)$ 的**首项**，$a_n$ 称为**首项系数**，$n$ 称为多项式 $f(x)$ 的**次数**，记作 $\deg f$.

零多项式的次数定义为 $-\infty$，并且规定：

$$(-\infty) + (-\infty) = -\infty$$

对一切非负整数 $n$，$(-\infty) + n = -\infty$，$-\infty < n$.

零次多项式是 $a_0x^0$，其中 $a_0 \neq 0$.我们可以把 $a_0x^0$ 与数域 $K$ 中的元素 $a_0$ 等同起来.因此数域 $K$ 中任一非零元素 $a$ 可看成是零次

多项式.

数域 $K$ 上的所有一元多项式组成的集合记作 $K[x]$. 现在我们在集合 $K[x]$ 中定义加法和乘法两种运算，其背景来自读者在中学数学里所熟悉的多项式的加法和乘法. 即按通常字母运算的规则把两个多项式 $f(x),g(x)$ 相加或相乘，$x$ 被认为是与数域 $K$ 中元素可交换的，并且把 $x$ 方次相同的项全合并起来，那么我们就得到一个多项式. 精确写出加法和乘法的定义如下：

**定义 2** 设 $f(x),g(x)\in K[x]$，其中

$$f(x)=a_nx^n+a_{n-1}x^{n-1}+\cdots+a_0=\sum_{i=0}^{n}a_ix^i$$

$$g(x)=b_mx^m+b_{m-1}x^{m-1}+\cdots+b_0=\sum_{j=0}^{m}b_jx^j$$

不妨设 $n\geqslant m$. $f(x)$ 与 $g(x)$ 的**和**是一个多项式

$$h(x)=\sum_{i=0}^{n}c_ix^i$$

其中 $h(x)$ 的 $i$ 次项的系数为

$$c_i=a_i+b_i,\quad i=0,1,\cdots,n \tag{3}$$

记作 $h(x)=f(x)+g(x)$.

$f(x)$ 与 $g(x)$ 的**乘积**是一个多项式

$$p(x)=\sum_{s=0}^{n+m}d_sx^s$$

其中 $p(x)$ 的 $s$ 次项的系数为

$$d_s=\sum_{i+j=s}a_ib_j,\quad s=0,1,\cdots,n+m \tag{4}$$

记作 $p(x)=f(x)g(x)$.

不难验证上面所定义的多项式的加法与乘法满足下列运算法则：$\forall f(x),g(x),h(x)\in K[x]$，有

1° 加法交换律，即 $f+g=g+f$；

2° 加法结合律，即 $(f+g)+h=f+(g+h)$；

3° 零多项式具有性质：$0+f=f+0=f$；

4° 设 $f(x)=\sum_{i=0}^{n}a_ix^i$，定义 $-f(x)=\sum_{i=0}^{n}(-a_i)x^i$，则

$$f+(-f)=(-f)+f=0$$

称 $-f$ 是 $f$ 的负元素；

5° 乘法交换律，即 $fg=gf$；

6° 乘法结合律，即 $(fg)h=f(gh)$；

7° 零次多项式 1 具有性质：$1f=f1=f$；

8° 乘法对于加法的分配律：

$$f(g+h)=fg+fh$$

$$(g+h)f=gf+hf$$

加法的四条运算法则是显然的，因为多项式的加法归结为同次项的系数相加.

从公式(4)易看出，多项式的乘法适合交换律，因为数域 $K$ 中乘法有交换律.

关于乘法的法则 7°也是显然的.

现在来验证乘法的结合律：设

$$f(x)=\sum_{i=0}^{n}a_ix^i,\quad g(x)=\sum_{j=0}^{m}b_jx^j,\quad h(x)=\sum_{k=0}^{l}c_kx^k \tag{5}$$

因为 $fg$ 的 $s$ 次项的系数为 $\sum\limits_{i+j=s}a_ib_j$，所以 $(fg)h$ 的 $t$ 次项的系数为

$$\sum_{s+k=t}\Big(\sum_{i+j=s}a_ib_j\Big)c_k=\sum_{i+j+k=t}a_ib_jc_k \tag{6}$$

因为 $gh$ 的 $r$ 次项系数为 $\sum\limits_{j+k=r}b_jc_k$，所以 $f(gh)$ 的 $t$ 次项系数为

$$\sum_{i+r=t}a_i\Big(\sum_{j+k=r}b_jc_k\Big)=\sum_{i+j+k=t}a_ib_jc_k \tag{7}$$

从(6)和(7)式得出，$(fg)h=f(gh)$.

下面来验证第一个分配律：设 $f,g,h$ 同(5). $f(g+h)$ 的 $s$ 次项的系数为

$$\sum_{i+j=s}a_i(b_j+c_j)=\sum_{i+j=s}a_ib_j+\sum_{i+j=s}a_ic_j \tag{8}$$

(8)式右端正好是 $fg+fh$ 的 $s$ 次项的系数，所以

$$f(g+h)=fg+fh$$

类似地可验证第二个分配律.

多项式的减法定义成：

$$f(x)-g(x):=f(x)+(-g(x))$$

**命题** 7.1.1 设 $f(x),g(x)\in K[x]$，则

$$\deg(f\pm g)\leqslant\max\{\deg f,\deg g\} \tag{9}$$

**证明** 设 $f(x)=\sum_{i=0}^{n}a_ix^i, g(x)=\sum_{i=0}^{m}b_ix^i$. 若 $f=g=0$，则 (9) 式显然成立. 下面设 $f\neq0$，并且 $a_n\neq0$，设 $n\geqslant m$. 因为

$$f(x)\pm g(x)=\sum_{i=0}^{n}(a_i\pm b_i)x^i$$

所以 $$\deg(f\pm g)\leqslant n=\max\{\deg f,\deg g\}$$ ▌

**命题** 7.1.2 设 $f(x),g(x)\in K[x]$. 如果 $f\neq0,g\neq0$，则 $fg\neq0$；并且有

$$\deg(fg)=\deg f+\deg g \tag{10}$$

**证明** 设

$$f(x)=\sum_{i=0}^{n}a_ix^i,\quad g(x)=\sum_{j=0}^{m}b_jx^j$$

如果 $f\neq0,g\neq0$，设 $a_n\neq0,b_m\neq0$，则 $a_nb_m\neq0$. 于是 $a_nb_mx^{n+m}$ 是 $f(x)g(x)$ 的首项，因此 $fg\neq0$，并且 $n+m$ 是 $fg$ 的次数，即

$$\deg(fg)=\deg f+\deg g$$

显然，当 $f$ 与 $g$ 有一个为 0 或者两个均为 0 时，(10) 式仍然成立. ▌

从命题 7.1.2 的证明中还可看出，多项式乘积的首项系数等于因子首项系数的乘积.

从命题 7.1.2 还可得出多项式乘法适合

9° 乘法消去律，即如果 $fg=fh$，且 $f\neq0$，则 $g=h$.

**证明** 从 $fg=fh$ 得

$$f(g-h)=0$$

由于 $f \neq 0$,所以

$$g - h = 0, \quad 即 \, g = h$$ ▌

1.2　环的基本概念

从上一小节知道，数域 $K$ 上所有一元多项式组成的集合 $K[x]$ 有两种运算：加法与乘法，并且满足一些运算法则. 整数集 $Z$ 也有加法与乘法两种运算，并且满足与上述同样的运算法则. 数域 $K$ 上所有 $n$ 级矩阵组成的集合 $M_n(K)$ 也有加法与乘法两种运算，并且满足加法交换律，加法结合律，乘法结合律，乘法对于加法的两个分配律. 从这些不同的对象中抽象出其关于运算的共同性质，便得到"环"的概念：

**定义** 3　设 $S$ 是一个非空集合，从 $S \times S$ 到 $S$ 的一个映射称为定义在 $S$ 上的一个**代数运算**.

**定义** 4　设 $R$ 是一个非空集合，在它上面定义了两个代数运算，一个叫做**加法**，记作 $a + b$；另一个叫做**乘法**，记作 $ab$，它们适合

1. 关于加法的以下规则：

1° 加法结合律，即 $(a + b) + c = a + (b + c)$；

2° 加法交换律，即 $a + b = b + a$；

3° 在 $R$ 中有元素 0，使得，$a + 0 = a, \forall\, a \in R$；

4° 对于 $R$ 中每个元素 $a$，在 $R$ 中有一个元素 $b$，使得 $a + b = 0$.

2. 关于乘法的以下规则：

5° 乘法结合律，即 $(ab)c = a(bc)$.

3. 关于加法和乘法的分配律：

6° 左分配律，即 $a(b + c) = ab + ac$；

　右分配律，即 $(b + c)a = ba + ca$.

这样的非空集合 $R$ 称为一个**环**.

容易证明，具有 3° 中性质的元素 0 是唯一的，称它是 $R$ 的零元素. 对于 $a \in R$，具有 4° 中性质的元素 $b$ 是唯一的，称它是 $a$ 的负元素，记作 $-a$.

**例** 1　整数集 $Z$ 对于数的加法与乘法形成一个环，称之为整数环.

**例** 2　数域 $K$ 上所有一元多项式组成的集合 $K[x]$ 对于节 1.1 中定义的加法与乘法形成一个环，称为数域 $K$ 上的一元多项式环.

**例** 3　数域 $K$ 上所有 $n$ 级矩阵组成的集合 $M_n(K)$ 对于矩阵的加法与乘法形成一个环，称为数域 $K$ 上的 $n$ 级矩阵环.

**例** 4　任意一个数域 $K$ 对于数的加法与乘法形成一个环.

*** 例** 5　设 $R$ 是一个环，$x$ 是一个不属于 $R$ 的符号，所有形式为

$$a_nx^n + a_{n-1}x^{n-1} + \cdots + a_1x + a_0$$
$$a_i \in R,\quad i = 0,1,\cdots,n,\quad n \geqslant 0$$

的表达式组成的集合记作 $R[x]$. $R[x]$ 中两个元素相等规定为它们除去系数为零的项外含有完全相同的项. $R[x]$ 中的元素称为**系数在环 $R$ 中的一元多项式**，符号 $x$ 称为不定元. 在 $R[x]$ 中按照 $K[x]$ 中的那样定义加法与乘法，同样可验证加法满足 4 条规则，乘法满足结合律，乘法对于加法的两个分配律也满足，因此 $R[x]$ 形成一个环，称它为**系数在环 $R$ 中的一元多项式环**.

*** 例** 6　$Z[x]$ 是系数在整数环 $Z$ 中的一元多项式环.

环 $R$ 中的减法定义成：

$$a - b := a + (-b)$$

如果环 $R$ 中的乘法还适合交换律，则称 $R$ 是交换环.

如果环 $R$ 中有一个元素 $e$ 具有性质：对于一切 $a \in R$，$ea = ae = a$，则称 $e$ 是 $R$ 的**单位元素**，此时称 $R$ 是有单位元素的环，环 $R$ 的单位元素通常就记作 1.

环 $R$ 中的元素 $a$ 称为一个**左零因子**，如果 $R$ 中存在 $b \neq 0$，使得 $ab = 0$. 类似地，$a$ 称为一个**右零因子**，如果 $R$ 中存在 $c \neq 0$，使得 $ca = 0$. 左零因子和右零因子都简称为零因子. 据阅读材料一的性质 1，0 既是左零因子，又是右零因子，称它是平凡零因子.

如果环 $R$ 没有非平凡的零因子，则 $R$ 称为**无零因子环**.

**定义** 5　如果环 $R$ 是有单位元素 $1(\neq 0)$、无零因子的交换环，则 $R$ 称为**整环**.

$Z, K[x], K$ 都是整环. $M_n(K)$ 不是整环，因为它不是交换环，并且有非平凡的零因子.

**定义** 6　如果环 $R$ 中一个非空子集 $R_1$ 对于 $R$ 的运算也组成一个环，则 $R_1$ 称为 $R$ 的一个**子环**.

**命题** 7.1.3　环 $R$ 的一个非空子集 $R_1$ 为一个子环的充分必要条件是 $R_1$ 对于 $R$ 的减法与乘法都封闭，即 $\forall\, a, b \in R_1$，有

$$a - b \in R_1, \qquad ab \in R_1$$

**证明**　必要性：设 $R_1$ 是环 $R$ 的一个子环. 由定义，$R_1$ 是一个环，因此 $\forall\, a, b \in R_1$，有 $a - b \in R_1, ab \in R_1$.

充分性. 要证 $R_1$ 对于 $R$ 的运算形成一个环. 任给 $a, b \in R_1$，由已知条件得 $a - a \in R_1$，即得 $0 \in R_1$. 从而 $0 - b \in R_1$，即 $-b \in R_1$. 于是 $a - (-b) \in R_1$，即 $a + b \in R_1$. 这证明了 $R$ 的加法也是 $R_1$ 的加法(即 $R_1$ 对于 $R$ 的加法封闭). 加法交换律 、加法结合律既然在 $R$ 中成立，自然在 $R_1$ 中也成立. 上面已证 $R$ 的零元素 $0 \in R_1$，从而 $R_1$ 也有了零元素 0. 上面已证对于 $b \in R_1$，有 $-b \in R_1$，从而 $R_1$ 中每个元素都有负元素. 从已知条件知，$R_1$ 对于 $R$ 的乘法封闭，因此 $R$ 的乘法也是 $R_1$ 的乘法. $R_1$ 中自然也成立乘法的结合律，乘法对于加法的两个分配律. 综上述，$R_1$ 成为一个环. 所以 $R_1$ 是 $R$ 的一个子环. ▌

**例** 7　在整数环 $Z$ 中，全体偶数组成的集合记作 $2Z$. 由于两个偶数的差、积都是偶数，所以 $2Z$ 是 $Z$ 的一个子环. 易看出，$2Z$ 没有单位元素. 显然，全体奇数组成的集合就不是 $Z$ 的子环.

**例** 8　在数域 $K$ 上的一元多项式环 $K[x]$ 中，所有零次多项式添上零多项式组成的集合 $S$ 是 $K[x]$ 的一个子环，这是因为两个零次多项式的差或者是零次多项式，或者是零多项式；两个零次多项式的积是零次多项式. 显然 $ax^0 \longmapsto a$ 是 $S$ 到 $K$ 的双射，并且这个映射保持加法与乘法. 这从下式可以看出：

$$ax^0 + bx^0 = (a+b)x^0, \quad (ax^0)(bx^0) = (ab)x^0$$

因此我们可以把 $S$ 与 $K$ 等同起来，从而可以把 $K$ 看成 $K[x]$ 的一个子环.

如果 $R_1$ 是环 $R$ 的一个子环，则称 $R$ 是环 $R_1$ 的一个扩环. 从例 8 知道，$K[x]$ 可以看成是 $K$ 的一个扩环.

**例** 9　设 $A \in M_n(K)$，令

$$K[A]: = \{a_m A^m + a_{m-1}A^{m-1} + \cdots + a_1 A + a_0 I \\ | a_i \in K, \quad i = 0,1,\cdots,m, m \geqslant 0\}$$

$K[A]$ 中的元素称为**矩阵 $A$ 的多项式**. 通常用 $f(A), g(A), \cdots$ 来代表矩阵 $A$ 的多项式. 从矩阵的加法、数乘、乘法知道，矩阵 $A$ 的多项式仍然是一个矩阵，因此 $K[A]$ 是 $M_n(K)$ 的一个非空子集. 设

$$f(A) = \sum_{i=0}^{m} a_i A^i, \quad g(A) = \sum_{i=0}^{l} b_i A^i$$

不妨设 $m \geqslant l$. 从矩阵的运算法则可得出

$$f(A) \pm g(A) = \sum_{i=0}^{m} (a_i \pm b_i) A^i \tag{11}$$

$$f(A)g(A) = \sum_{s=0}^{m+l} \Big( \sum_{i+j=s} a_i b_j \Big) A^s \tag{12}$$

因此 $K[A]$ 对于矩阵的加法、减法、乘法都封闭. 从而 $K[A]$ 是 $M_n(K)$ 的一个子环. 从公式(12) 得出

$$f(A)g(A) = g(A)f(A)$$

因此 $K[A]$ 是交换环. 称 $K[A]$ 是**矩阵 $A$ 的多项式环**.

**例** 10　设 $A \in M_n(K)$，在环 $K[A]$ 中考虑子集

$$W = \{aI \mid a \in K\}$$

即 $W$ 是所有 $n$ 级数量矩阵组成的集合. 显然 $W$ 对于矩阵的减法与乘法都封闭，因此 $W$ 是 $K[A]$ 的一个子环. 显然 $aI \longmapsto a$ 是 $W$ 到 $K$ 的一个双射，并且从

$$aI + bI = (a+b)I, \quad (aI)(bI) = (ab)I$$

知道，上述映射保持加法与乘法. 因此可以把 $W$ 与 $K$ 等同起来，从

而可以把 $K$ 看成环 $K[A]$ 的一个子环. 换句话说，$K[A]$ 可以看成是 $K$ 的一个扩环.

从例 8 与例 10 知道，$K[x]$ 与 $K[A]$ 都可以看成是 $K$ 的扩环. 但是 $K[x]$ 与 $K[A]$ 有本质的区别：$K[x]$ 中的元素写成

$$a_nx^n + a_{n-1}x^{n-1} + \cdots + a_1x + a_0$$

时，表法是唯一的(除了系数为 0 的项以外)，这是不定元 $x$ 的一个本质特点. 然而，$K[A]$ 中的元素写成

$$a_mA^m + a_{m-1}A^{m-1} + \cdots + a_1A + a_0I$$

时，表法是不唯一的. 例如，设

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

经直接计算得，$A^2 = I$. 于是此式左右两边是 $K[A]$ 中同一个元素，$A^2$ 与 $I$ 是这个元素的两种不同表示方式.

一元多项式环 $K[x]$ 中不定元 $x$ 的上述本质特点以及 $K[x]$ 中加法与乘运算使得 $K[x]$ 在交换环中占有重要地位，这点从下一小节可以看得很清楚.

### 1.3 一元多项式环的通用性质

一元多项式环具有通用性，这从下面的定理可以看出来：

**定理** 7.1.1　设 $K$ 是一个数域，设 $R$ 是一个交换环，并且 $R$ 可以看成是 $K$ 的一个扩环(意思是，$R$ 有一个子环 $R_1$ 与 $K$ 存在一一对应，它保持加法与乘法. 从而可以把 $R_1$ 与 $K$ 等同起来). 对于 $R$ 中每一个元素 $t$，令

$$\sigma_t : K[x] \longrightarrow R$$

$$f(x) = \sum_{i=0}^{n} a_i x^i \longmapsto \sum_{i=0}^{n} a_i t^i =: f(t) \tag{13}$$

则 $\sigma_t$ 是 $K[x]$ 到 $R$ 的一个映射，它使得

$$\sigma_t(a) = a, \quad \forall\, a \in K; \quad \sigma_t(x) = t \tag{14}$$

并且 $\sigma_t$ 保持加法与乘法运算，即 $\forall\, f(x), g(x) \in K[x]$，有

$$\sigma_t(f+g)=\sigma_t(f)+\sigma_t(g),\quad \sigma_t(fg)=\sigma_t(f)\cdot\sigma_t(g)$$

换句话说,如果

$$f(x)+g(x)=h(x),\quad f(x)g(x)=p(x)$$

则有 $$f(t)+g(t)=h(t),\quad f(t)g(t)=p(t) \tag{15}$$

由(13)式定义的映射 $\sigma_t$ 称为$x$ 用 $t$ 代入.

**证明** 因为 $K[x]$ 中每个元素写成 $\sum_{i=0}^{n} a_i x^i$ 的表法唯一(除了系数为 0 的项以外),并且 $R$ 可以看成是 $K$ 的一个扩环,所以由(13)式定义的对应法则是 $K[x]$ 到 $R$ 的一个映射,由(13)式立即得出,$\sigma_t(a)=a$,$\forall\ a\in K$;$\sigma_t(x)=t$. 现在来证映射 $\sigma_t$ 保持加法与乘法运算:设

$$f(x)=\sum_{i=0}^{n} a_i x^i,\quad g(x)=\sum_{i=0}^{m} b_i x^i$$

不妨设 $n\geqslant m$,并且设

$$h(x)=\sum_{i=0}^{n} c_i x^i,\quad p(x)=\sum_{s=0}^{n+m} d_s x^s$$

因为 $$f(x)+g(x)=h(x),\quad f(x)g(x)=p(x)$$

所以 $$c_i=a_i+b_i,\quad d_s=\sum_{i+j=s} a_i b_j$$

根据 $\sigma_t$ 的定义,得

$$f(t)=\sum_{i=0}^{n} a_i t^i,\quad g(t)=\sum_{i=0}^{m} b_i t^i$$

$$h(t)=\sum_{i=0}^{n} c_i t^i,\quad p(t)=\sum_{s=0}^{n+m} d_s t^s$$

于是由 $R$ 是交换环(实际上只需用到 $t$ 与 $K$ 中所有元素可交换)得到

$$f(t)+g(t)=\sum_{i=0}^{n} a_i t^i+\sum_{i=0}^{m} b_i t^i$$

$$=\sum_{i=0}^{n}(a_i+b_i)t^i=\sum_{i=0}^{n} c_i t^i=h(t)$$

$$f(t)g(t) = (\sum_{i=0}^{n} a_i t^i)(\sum_{j=0}^{m} b_j t^j)$$

$$= \sum_{i=0}^{n}\sum_{j=0}^{m} a_i b_j t^{i+j} = \sum_{s=0}^{n+m}(\sum_{i+j=s} a_i b_j)t^s = \sum_{s=0}^{n+m} d_s t^s = p(t)$$

这样我们证明了 $\sigma_t$ 保持加法与乘法运算. ▍

定理 7.1.1 的实质是：一元多项式 $f(x),g(x),\cdots$ 之间所有通过加法与乘法表示的关系在不定元 $x$ 用任意一个与 $K$ 中所有元素可交换的环元素 $t$ 代入之后仍然保持.

下面通过几个例子来理解定理 7.1.1 的实质并且看出定理 7.1.1 的用处.

**例** 11　取 $K$ 是有理数域 $Q$，取 $R$ 为复数域 $C$. 显然 $C$ 是 $Q$ 的扩环. 据定理 7.1.1，任一有理系数多项式 $f(x)$ 中的不定元 $x$ 可用任一复数 $t$ 代入得到 $f(t)$，并且这种代入保持加法与乘法运算. 譬如，据二项式定理(注意：一元多项式的加法与乘法定义与中学数学中多项式的加法与乘法定义一致，而中学数学里已证过二项式定理)，有

$$(1+x)^n = 1 + \mathrm{C}_n^1 x + \mathrm{C}_n^2 x^2 + \cdots + \mathrm{C}_n^i x^i + \cdots + \mathrm{C}_n^n x^n \tag{16}$$

$x$ 用 1 代入得

$$2^n = (1+1)^n = 1 + \mathrm{C}_n^1 + \mathrm{C}_n^2 + \cdots + \mathrm{C}_n^n \tag{17}$$

(17)式是读者在中学里已熟悉的关于组合数的一条性质. 这一性质的证明正是利用了定理 7.1.1：$x$ 用 1 代入保持乘法运算. 实际上，读者在中学时就早已屡次使用了“$x$ 用数 $t$ 代入保持加法与乘法运算”这一结论，而这一结论是定理 7.1.1 的一个特殊情形.

**例** 12　设 $K$ 是一个数域，取 $R$ 为 $K[x]$. 例 8 已指出 $K[x]$ 可以看成是 $K$ 的一个扩环. 据定理 7.1.1，数域 $K$ 上任意一个一元多项式 $f(x)$ 中的不定元 $x$ 可用环 $K[x]$ 中的元素 $x^m$ 代入，其中 $m$ 是任一正整数，并且这种代入保持加法与乘法运算. 于是如果 $f(x)g(x) = p(x)$，则有 $f(x^m)g(x^m) = p(x^m)$. 不定元 $x$ 也可用环 $K[x]$ 中元素 $1+x$ 代入. 于是如果 $f(x)g(x) = p(x)$，则有

$f(1+x)g(1+x)=p(1+x)$，等等.

**例** 13　在组合数学中，一元多项式常常被用来作为不同类型的数的生成函数. 譬如，从二项式定理得

$$(1+x)^n=1+C_n^1x+C_n^2x^2+\cdots+C_n^ix^i+\cdots+C_n^nx^n$$

因此，一元多项式 $(1+x)^n$ 可以作为组合数 $C_n^i$ 的**生成函数**，意思是，从多项式 $(1+x)^n$ 的各项系数可分别得到 $C_n^0,C_n^1,\cdots,C_n^n$. 利用组合数的生成函数 $(1+x)^n$ 可以证明组合数的一些性质. 例 11 中我们已证明了关于组合数的一条性质，现在再证明组合数的另一条性质：

$$\sum_{i=0}^{k}C_m^iC_n^{k-i}=C_{m+n}^k \tag{18}$$

**证明**　因为 $x^mx^n=x^{m+n}$，据例 12 知，$x$ 用 $1+x$ 代入是保持乘法运算的，所以我们得到

$$(1+x)^m(1+x)^n=(1+x)^{m+n} \tag{19}$$

(19)式左右两边的多项式相等，从而它们的 $k$ 次项系数应当相等，于是得到

$$\sum_{i+j=k}C_m^iC_n^j=C_{m+n}^k$$

即

$$\sum_{i=0}^{k}C_m^iC_n^{k-i}=C_{m+n}^k$$ ▌

**例** 14　取 $R$ 为 $K[A]$，其中 $A\in M_n(A)$. 例 10 已指出，$K[A]$ 可以看成是 $K$ 的一个扩环. 并且 $K[A]$ 是交换环. 于是据定理 7.1.1，任意一元多项式 $f(x)$ 中的 $x$ 可用矩阵 $A$ 代入得到 $f(A)$，且这种代入保持加法与乘法运算. 从而如果

$$f(x)g(x)=p(x),\quad f(x)+g(x)=h(x)$$

则有

$$f(A)g(A)=p(A),\quad f(A)+g(A)=h(A)$$

譬如，因为

$$(1+x)^n=1+C_n^1x+C_n^2x^2+\cdots+C_n^ix^i+\cdots+C_n^nx^n$$

所以我们有

$$(I+A)^n = I + C_n^1 A + C_n^2 A^2 + \cdots + C_n^i A^i + \cdots + C_n^n A^n$$

这里要注意当 $x$ 用 $A$ 代入时，常数项 $a_0$ 应当用与它等同的数量矩阵 $a_0 I$ 代替.

**例 15** 设矩阵 $B$ 是数域 $K$ 上的 $n$ 级幂零矩阵，其幂零指数为 $l$. 令

$$A = aI + bB,\quad a \neq 0,\quad a, b \in K$$

证明 $A$ 可逆，并且求 $A^{-1}$.

**证明** 在 $K[x]$ 中直接计算可得

$$(1-x)(1+x+x^2+\cdots+x^{l-1}) = 1 - x^l \tag{20}$$

因为 $K[B]$ 可看成是 $K$ 的交换扩环，所以在(20)中 $x$ 用 $-\dfrac{b}{a}B$ 代入得

$$\begin{aligned}&\left(I+\frac{b}{a}B\right)\left(I-\frac{b}{a}B+\frac{b^2}{a^2}B^2+\cdots+(-1)^{l-1}\frac{b^{l-1}}{a^{l-1}}B^{l-1}\right)\\ =&I-\left(-\frac{b}{a}\right)^l B^l = I\end{aligned} \tag{21}$$

从(21)得到

$$\begin{aligned}&(aI+bB)(a^{-1}I-a^{-2}bB+a^{-3}b^2B^2\\&+\cdots+(-1)^{l-1}a^{-l}b^{l-1}B^{l-1}) = I\end{aligned}$$

所以 $A = aI + bB$ 是可逆矩阵，并且

$$A^{-1} = a^{-1}I - a^{-2}bB + a^{-3}b^2B^2 + \cdots + (-1)^{l-1}a^{-l}b^{l-1}B^{l-1}$$ ▌

从以上例子可以看出，一元多项式环 $K[x]$ 中的不定元 $x$ 可以用任意环元素 $t$ 代入(只要 $t$ 与 $K$ 中所有元素可交换)，并且这种代入保持加法与乘法运算. 这就是一元多项式环的通用性质，正是由于 $x$ 可以用任意环元素代入(条件同前)，所以把 $x$ 叫做“不定元”. 一元多项式环的这个通用性质的意义在于：只要我们把一元多项式环的有关运算的性质研究清楚了，通过不定元 $x$ 用环元素代入，就可得到其他一些交换环的有关运算的性质，因为这种代入是保持运算的.

**定义 7** 设 $R$ 和 $R'$ 是两个环，从 $R$ 到 $R'$ 的一个映射 $\sigma$，如果保

持加法和乘法运算，即 $\forall\ a,b \in R$，有

$$\sigma(a+b)=\sigma(a)+\sigma(b) \tag{22}$$

$$\sigma(ab)=\sigma(a)\sigma(b) \tag{23}$$

则称 $\sigma$ 是环 $R$ 到 $R'$ 的一个**同态映射**（简称为**同态**），或者称 $\sigma$ 是环 $R$ 到 $R'$ 的一个**态射**.

从定理 7.1.1 知道，设 $R$ 可看成是 $K$ 的一个交换扩环，不定元 $x$ 用 $R$ 的元素 $t$ 代入，则 $\sigma_t$ 就是环 $K[x]$ 到 $R$ 的一个同态映射.

**定义** 8 设 $R$ 和 $R'$ 是两个环，如果存在从 $R$ 到 $R'$ 的一个双射 $\sigma$，并且 $\sigma$ 保持加法和乘法运算，那么称环 $R$ 与 $R'$ 是**同构的**，称 $\sigma$ 是 $R$ 到 $R'$ 的一个**同构映射**（简称为**同构**）. 此时记 $R \cong R'$.

注意：在(22)式和(23)式的左端中的加法、乘法是环 $R$ 的运算，而右端的加法、乘法是环 $R'$ 的运算.

从下一节开始，我们将用几节来讨论一元多项式环有关乘法的性质.

# 阅读材料一

一个非空集合 $S$，如果定义了一个或几个代数运算并且这些运算满足一定的法则，则称 $S$ 是一个**代数结构**.

环是一个代数结构，它有两个代数运算：加法和乘法（减法是通过加法定义的），并且满足六条运算法则. 现在我们要从环的定义出发经过逻辑推理得出环的进一步的性质（有关运算的性质），这种研究方法称为**公理化方法**. 下面设 $R$ 是一个环.

**性质 1** 环 $R$ 的零元素具有性质：$0a=a0=0, \forall\ a \in R$.

**证明** 在 $R$ 中取一个元素 $b$，我们有

$$ab=a(b+0)=ab+a0$$

上式两边加上 $(-ab)$，得

$$\text{左边}=(-ab)+ab=0$$

$$\text{右边}=[(-ab)+ab]+a0=0+a0=a0$$

由此得出 $a0=0$. 同理可证 $0a=0$. ▌

**性质 2** $\forall a,b\in R$,有

$$(-a)b=a(-b)=-ab$$
$$(-a)(-b)=ab$$

**证明** 因为

$$(-a)b+ab=[(-a)+a]b=0b=0$$

所以 $(-a)b=-ab$. 同理可证

$$a(-b)=-ab$$

从而
$$(-a)(-b)=-[a(-b)]=-[-(ab)]=ab$$
▌

由于环 $R$ 的加法满足结合律,所以可以定义环元素 $a$ 的正整数 $n$ 倍,即

$$na:=\underbrace{a+\cdots+a}_{n\text{个}}$$

还可以定义 $a$ 的零倍以及 $a$ 的负整数倍:

$$0a:=0$$
$$(-n)a:=n(-a)$$

注意:上面第一式左边的 0 是整数,而右边的 0 是环 $R$ 的元素.

**性质 3** 对任意 $a,b\in R$,任意整数 $m,n$,有

$$(m+n)a=ma+na$$
$$n(ma)=(nm)a$$
$$n(a+b)=na+nb$$
$$(na)b=a(nb)=n(ab)$$

**证明** 分情形讨论,细节从略. ▌

由于环 $R$ 的乘法满足结合律,因此可以定义环元素 $a$ 的正整数 $n$ 次幂:

$$a^n:=\underbrace{aa\cdots a}_{n\text{个}}$$

注意:对于一般的环,不能定义 $a$ 的零次幂与 $a$ 的负整数次幂. 如果环 $R$ 有单位元 1,那么可以定义 $a^0=1$.

**性质 4** 设 $a\in R$,对任意正整数 $m,n$,有

$$a^na^m=a^{n+m}$$
$$(a^n)^m=a^{nm}$$

**证明** 从环元素的正整数次幂的定义以及环的乘法结合律立即得到. ▌

注意:在环中,一般地,$(ab)^n$ 不等于 $a^nb^n$. 如果 $ab=ba$,那么 $(ab)^n=a^nb^n$ 成立. 同样,一般地,$(a+b)^n$ 的展开式不是二项式定理那种形式,然而如果

$ab = ba$，那么

$$(a+b)^n = a^n + C_n^1 a^{n-1}b + \cdots + C_n^i a^{n-i}b^i + \cdots + b^n$$

现在设 $R$ 是有单位元的环，并且 $R$ 的单位元 1 与 $R$ 的零元 0 不是同一个元素(这意味着 $R$ 至少有两个元素).

**定义 1** 设 $a \in R$，如果存在 $b \in R$，使得

$$ab = ba = 1$$

则称 $a$ 是 $R$ 的一个**可逆元**，也称 $a$ 是 $R$ 的一个**单位**.

注意环 $R$ 的单位与单位元这两个术语的区别. $R$ 的单位元 1 是 $R$ 的单位，但是 $R$ 的单位(即可逆元) 不一定是 $R$ 的单位元.

容易证明：如果 $a$ 是 $R$ 的可逆元，那么满足 $ab = ba = 1$ 的元素 $b$ 是唯一的，称 $b$ 是 $a$ 的**逆**，记作 $a^{-1}$.

从习题 7.1 第 2 题知道，数域 $K$ 上的一元多项式环 $K[x]$ 中，$f(x)$ 是可逆元当且仅当它是零次多项式，即它是 $K$ 中非零数.

容易看出，整数环 $Z$ 的可逆元只有两个：1 和 $-1$.

# 习 题 7.1

1. 证明在整数环 $Z$ 中，3 的倍数组成的子集 $3Z$ 是 $Z$ 的一个子环；对任一整数 $n$，$n$ 的倍数组成的子集 $nZ$ 是 $Z$ 的一个子环.

2. 设 $R$ 是一个有单位元素的环，$R$ 中一个元素 $a$ 称为可逆元素(或称为单位，注意不要与单位元素 1 混淆)，如果有 $b \in R$，使得

$$ab = ba = 1$$

证明：(1) $K[x]$ 中一个元素 $f(x)$ 是可逆元当且仅当 $f(x)$ 是零次多项式；(2) 整数环 $Z$ 的可逆元只有两个：1 和 $-1$.

3. 设 $R$ 是有单位元素的环，说明 $R$ 中的可逆元不可能是零因子.

4. 设 $n$ 级矩阵

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

令 $A = aI + B$，其中 $a \neq 0$. 证明：$A$ 可逆，并且求 $A^{-1}$.

5. 设

$$A = \begin{pmatrix} 1 & b & b^2 & \cdots & b^{n-2} & b^{n-1} \\ 0 & 1 & b & \cdots & b^{n-3} & b^{n-2} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 & b \\ 0 & 0 & 0 & \cdots & 0 & 1 \end{pmatrix}$$

求 $A^{-1}$.

6. 设 $A \in M_n(A)$，并且设 $A$ 的特征多项式为

$$|\lambda I - A| = (\lambda - \lambda_1)^{l_1}(\lambda - \lambda_2)^{l_2}\cdots(\lambda - \lambda_m)^{l_m}$$

其中 $\lambda_1, \lambda_2, \cdots, \lambda_m$ 是互不相同的复数，$l_1 + l_2 + \cdots + l_m = n$. 证明：对于 $K$ 中任一非零数 $k$，矩阵 $kA$ 的特征多项式为

$$|\lambda I - kA| = (\lambda - k\lambda_1)^{l_1}(\lambda - k\lambda_2)^{l_2}\cdots(\lambda - k\lambda_m)^{l_m}$$

由此得出：如果 $\lambda_1$ 是 $A$ 的 $l_1$ 重特征值，则 $k\lambda_1$ 是 $kA$ 的 $l_1$ 重特征值.

*7. 设 $A$ 及 $A$ 的特征多项式同第 6 题. 证明：$A^2$ 的特征多项式为

$$|\lambda I - A^2| = (\lambda - \lambda_1^2)^{l_1}(\lambda - \lambda_2^2)^{l_2}\cdots(\lambda - \lambda_m^2)^{l_m}$$

由此得出：如果 $\lambda_1$ 是 $A$ 的 $l_1$ 重特征值，则 $\lambda_1^2$ 是 $A^2$ 的 $l_1$ 重特征值.

*8. 设 $A$ 及 $A$ 的特征多项式同第 6 题. 证明：$A^3$ 的特征多项式为

$$|\lambda I - A^3| = (\lambda - \lambda_1^3)^{l_1}(\lambda - \lambda_2^3)^{l_2}\cdots(\lambda - \lambda_m^3)^{l_m}$$

由此得出：如果 $\lambda_1$ 是 $A$ 的 $l_1$ 重特征值，则 $\lambda_1^3$ 是 $A^3$ 的 $l_1$ 重特征值.

## §2 带余除法·整除性质初步

在数域 $K$ 上的一元多项式环 $K[x]$ 中，可以做加法与乘法运算，并且可以做加法的逆运算：减法. 但是乘法的逆运算：除法却不是 $K[x]$ 的运算，因为两个多项式相除所得的结果不一定是多项式. 什么样的两个多项式相除仍为多项式？这是我们关心的问题. 为此首先讨论两个多项式相除的一般情形，在中学数学里，讲过多项式的长除法，由于我们现在讨论的一元多项式，其加法、乘法的

定义与中学数学讲的多项式的加法、乘法一致，因此对于现在的一元多项式也可以做长除法. 例如，设

$$f(x) = 2x^3 + 3x^2 + 5, \quad g(x) = x^2 + 2x - 1$$

用 $g(x)$ 去除 $f(x)$ 可以按照下面的格式来作：

$$\begin{array}{r|rrrr|l}
x^2+2x-1 & 2x^3 & +3x^2 & & +5 & 2x-1 \\
 & 2x^3 & +4x^2 & -2x & & \\
\hline
 & & -x^2 & +2x & +5 & \\
 & & -x^2 & -2x & +1 & \\
\hline
 & & & 4x & +4 &
\end{array}$$

于是我们得到

$$2x^3 + 3x^2 + 5 = (2x - 1)(x^2 + 2x - 1) + (4x + 4)$$

其中 $2x - 1$ 称为 $g(x)$ 去除 $f(x)$ 所得的**商**，$4x + 4$ 称为**余式**. 这个结果具有一般性，即一元多项式环 $K[x]$ 具有下述重要性质：

**定理** 7.2.1(**带余除法**) 对于 $K[x]$ 中任意两个多项式 $f(x)$ 与 $g(x)$，其中 $g(x) \neq 0$，在 $K[x]$ 中存在唯一的一对多项式 $h(x)$，$r(x)$，使得

$$f(x) = h(x)g(x) + r(x), \quad \deg r(x) < \deg g(x) \tag{1}$$

**证明** 存在性. 已知 $g(x) \neq 0$，设 $\deg g(x) = m$. 如果 $m = 0$，则 $g(x) = b \in K$ 且 $b \neq 0$. 这时有 $f(x) = (b^{-1}f(x))b + 0$，(1) 式成立. 下面设 $m > 0$.

如果 $\deg f(x) < m$，则有 $f(x) = 0g(x) + f(x)$，(1) 式成立.

下面讨论 $\deg f(x) \geqslant m$ 的情形. 设 $\deg f(x) = n$，对 $n$ 作数学归纳法.

假设对于次数小于 $n$ 的每一个一元多项式，命题的存在性部分成立. 现在看 $n$ 次多项式 $f(x)$. 设 $f(x)$ 的首项是 $a_nx^n$，$g(x)$ 的首项是 $b_mx^m$. 于是 $a_nb_m^{-1}x^{n-m}g(x)$ 的首项是 $a_nx^n$. 从而

$$f_1(x) := f(x) - a_nb_m^{-1}x^{n-m}g(x) \tag{2}$$

的次数小于 $n$. 据归纳假设，$K[x]$ 中存在一对多项式 $h_1(x)$，

$r_1(x)$，使得

$$f_1(x) = h_1(x)g(x) + r_1(x), \quad \deg r_1(x) < \deg g(x) \qquad (3)$$

代入(2)得

$$f(x) = (h_1(x) + a_n b_m^{-1} x^{n-m})g(x) + r_1(x)$$

$$\deg r_1(x) < \deg g(x)$$

取 $h(x) = h_1(x) + a_n b_m^{-1} x^{n-m} \in K[x]$，$r(x) = r_1(x)$，则得(1)式.

据数学归纳法原理，定理 7.2.1 的存在性得证.

唯一性. 设另有多项式 $h'(x), r'(x) \in K[x]$，使得

$$f(x) = h'(x)g(x) + r'(x)$$

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

则得 $\qquad h(x)g(x) + r(x) = h'(x)g(x) + r'(x)$

即 $\qquad [h(x) - h'(x)]g(x) = r'(x) - r(x) \qquad (4)$

于是

$$\deg[h(x) - h'(x)] + \deg g(x) = \deg[r'(x) - r(x)]$$

$$\leqslant \max\{\deg r'(x), \deg r(x)\} < \deg g(x)$$

假如 $h(x) - h'(x) \neq 0$，则从上式得 $\deg[h(x) - h'(x)] < 0$，矛盾. 因此 $h(x) - h'(x) = 0$，即 $h(x) = h'(x)$. 从而得到

$$r(x) = r'(x)$$ ▌

(1)式中的 $h(x)$ 称为 $g(x)$ 除 $f(x)$ 的**商**，$r(x)$ 称为 $g(x)$ 除 $f(x)$ 的**余式**. 余式为零的情形是我们特别感兴趣的，为此引进下述概念：

**定义 1** 设 $f(x), g(x) \in K[x]$，若有 $h(x) \in K[x]$，使得

$$f(x) = h(x)g(x)$$

则称 $g(x)$ **整除** $f(x)$，记作 $g(x) \mid f(x)$. 当 $g(x)$ 整除 $f(x)$ 时，$g(x)$ 称为 $f(x)$ 的**因式(或因子)**，$f(x)$ 称为 $g(x)$ 的**倍式**.

当 $g(x) \neq 0$ 时，带余除法给出了整除性的一个判别法.

**定理 7.2.2** 设 $f(x), g(x) \in K[x]$，且 $g(x) \neq 0$，则 $g(x) \mid f(x)$ 的充分必要条件是 $g(x)$ 除 $f(x)$ 的余式为零.

**证明** 如果 $g(x)$ 除 $f(x)$ 的余式为零，则 $f(x) = h(x)g(x)$，

即 $g(x)|f(x)$.

如果 $g(x)|f(x)$,则有 $h(x)\in K[x]$,使得

$$f(x)=h(x)g(x)=h(x)g(x)+0$$

据带余除法的唯一性得,$g(x)$ 除 $f(x)$ 的余式为零. ▌

如果 $g(x)=0$,且 $g(x)|f(x)$,则

$$f(x)=h(x)g(x)=h(x)0=0$$

任取 $f(x)\in K[x]$,因为 $f(x)=1\cdot f(x)$,所以 $f(x)|f(x)$.因为 $0=0f(x)$, 所以 $f(x)|0$. 对于 $b\in K$, 且 $b\neq 0$, 有 $f(x)=[b^{-1}f(x)]b$, 所以 $b|f(x)$.

下面列举数域 $K$ 上一元多项式环 $K[x]$ 中,整除性的基本性质:

1. 如果 $f(x)|g(x)$,且 $g(x)|f(x)$,则 $f(x)=cg(x)$,其中 $c\in K$ 且 $c\neq 0$.

**证明** 因为 $f(x)|g(x)$,所以有 $h_1(x)\in K[x]$,使

$$g(x)=h_1(x)f(x)$$

因为 $g(x)|f(x)$,所以有 $h_2(x)\in K[x]$,使

$$f(x)=h_2(x)g(x)$$

于是

$$f(x)=h_2(x)h_1(x)f(x) \tag{5}$$

如果 $f(x)=0$,则 $g(x)=0$,从而结论显然成立.如果 $f(x)\neq 0$,则从(5)式得

$$1=h_2(x)h_1(x)$$

由此得

$$\deg h_2(x)+\deg h_1(x)=0$$

因此

$$\deg h_2(x)=\deg h_1(x)=0$$

从而 $h_2(x)=c\in K$,且 $c\neq 0$.于是 $f(x)=cg(x)$. ▌

如果 $f(x)|g(x)$,且 $g(x)|f(x)$,则称 $f(x)$ 与 $g(x)$ 是**相伴元素**,记作 $f(x)\sim g(x)$.从性质 1 以及定义 1 知道,$f(x)\sim g(x)$ 的充分必要条件是 $f(x)$ 与 $g(x)$ 相差一个非零数因子.

2. 如果 $f(x)|g(x)$,且 $g(x)|h(x)$,则 $f(x)|h(x)$(整除的传递性).

**证明** 由已知条件得

$$g(x)=g_1(x)f(x),\quad h(x)=h_1(x)g(x)$$

从而 $h(x)=[h_1(x)g_1(x)]f(x)$. 所以 $f(x)\mid h(x)$. ▌

3. 如果 $f(x)\mid g_i(x),i=1,2,\cdots,r$,则对于任意

$$u_i(x)\in K[x],\quad i=1,\cdots,r$$

有 $f(x)\mid(u_1(x)g_1(x)+u_2(x)g_2(x)+\cdots+u_r(x)g_r(x))$

**证明** 由 $g_i(x)=h_i(x)f(x),i=1,\cdots,r$,得

$$\begin{aligned}&u_1(x)g_1(x)+\cdots+u_r(x)g_r(x)\\=&[u_1(x)h_1(x)+\cdots+u_r(x)h_r(x)]f(x)\end{aligned}$$ ▌

通常,$u_1(x)g_1(x)+\cdots+u_r(x)g_r(x)$ 称为多项式 $g_1(x),\cdots,g_r(x)$ 的一个组合.

注意:两个多项式之间的整除关系不因系数域的扩大而改变. 即设 $f(x),g(x)\in K[x]$,数域 $F\supset K$,当然 $f(x),g(x)$ 也可看成是数域 $F$ 上的多项式. 如果在 $K[x]$ 中,$g(x)$ 不能整除 $f(x)$,那么在 $F[x]$ 中,$g(x)$ 也不能整除 $f(x)$. 证明如下:

假如在 $F[x]$ 中,$g(x)\mid f(x)$. 如果 $g(x)=0$,则 $f(x)=0$. 从而在 $K[x]$ 中有 $g(x)\mid f(x)$,矛盾. 下设 $g(x)\neq 0$,则存在 $h(x)\in F[x]$,使得

$$f(x)=h(x)g(x)+0\tag{6}$$

现在在 $K[x]$ 中作带余除法,有 $h_1(x),r_1(x)\in K[x]$,使得

$$f(x)=h_1(x)g(x)+r_1(x)\tag{7}$$

由于 $f(x),g(x),h_1(x),r_1(x)$ 都可看成是 $F[x]$ 中元素,因此(7)式可看成是在 $F[x]$ 中的带余除法. 比较(7)式和(6)式,由带余除法的唯一性得,$r_1(x)=0$. 因此 $f(x)=h_1(x)g(x)$. 从而在 $K[x]$ 中,$g(x)\mid f(x)$,矛盾. ▌

本节开头指出,用长除法可求得 $g(x)(\neq 0)$ 除 $f(x)$ 的商和余式. 如果 $g(x)=x-c$,其中 $c\in K$,且 $c\neq 0$,则我们用下述的综合除法求 $g(x)$ 除 $f(x)$ 的商与余式更为简便.

**例** 设 $f(x)=2x^4-6x^3+3x^2-2x+5,g(x)=x-2$,求

$g(x)$ 除 $f(x)$ 的商与余式.

**解**

$$\begin{array}{rrrrr|l} 2 & -6 & 3 & -2 & 5 & 2 \\ & 4 & -4 & -2 & -8 & \\ \hline 2 & -2 & -1 & -4 & -3 & \end{array}$$

所以 $f(x) = (2x^3 - 2x^2 - x - 4)(x - 2) - 3$

即 $g(x)$ 除 $f(x)$ 的商为 $2x^3 - 2x^2 - x - 4$,余式为 $-3$.

综合除法的理论根据如下:

设 $f(x) = \sum_{i=0}^{n} a_i x^i, a_n \neq 0, n \geqslant 1, g(x) = x - c$. 根据带余除法,我们有

$$f(x) = h(x)(x - c) + r, \quad r \in K \tag{8}$$

因为 $n \geqslant 1$,所以 $\deg f(x) = \deg h(x) + \deg(x - c)$. 由此得 $\deg h(x) = n - 1$. 设 $h(x) = \sum_{i=0}^{n-1} b_i x^i$. 比较(8)式两边首项系数,得 $a_n = b_{n-1}$. 比较 $s$ 次项的系数($s = 1, \cdots, n-1$),得

$$a_s = -cb_s + b_{s-1}$$

从而得 $b_{s-1} = a_s + cb_s, \quad s = n-1, \cdots, 1$

比较常数项得, $a_0 = -cb_0 + r$,从而得 $r = a_0 + cb_0$. 因此我们有

$$\begin{array}{ccccc|c} a_n & a_{n-1} & \cdots & a_1 & a_0 & c \\ & cb_{n-1} & \cdots & cb_1 & cb_0 & \\ \hline a_n & a_{n-1}+cb_{n-1} & \cdots & a_1+cb_1 & a_0+cb_0 & \\ \| & \| & & \| & \| & \\ b_{n-1} & b_{n-2} & \cdots & b_0 & r & \end{array}$$

## 阅读材料二

**$Z$ 中的带余除法.**

任给 $a, b \in Z, b \neq 0$,则存在唯一的一对整数 $q, r$,使得

$$a = bq + r, \quad 0 \leqslant r < |b|$$

**证明** 考虑集合

$$S = \{a - bs \mid s \in Z, \quad a - bs \geqslant 0\}$$

若 $b > 0$,则 $a - b(-a^2) = a + a^2b \geqslant 0$;若 $b < 0$,则 $a - ba^2 \geqslant 0$. 因此 $S$ 非空,从而$S$ 至少包含一个最小元素,设为 $a - bq$. 令 $r = a - bq$,根据 $S$ 的定义,$r \geqslant 0$. 如果 $r \geqslant |b|$,且 $b > 0$,则 $r - b \geqslant 0$,从而 $a - bq - b \geqslant 0$,即 $a - b(q+1) \geqslant 0$. 于是 $a - b(q+1) \in S$,即 $r - b \in S$,但 $r - b < r$,这与 $r$ 的定义矛盾. 如果 $r \geqslant |b|$,且 $b < 0$,则 $0 \leqslant r + b = a - b(q-1)$. 于是 $r + b \in S$,但 $r + b < r$,矛盾. 所以 $r < |b|$. 存在性得证.

唯一性. 假如还有整数 $q'$,$r'$,使得

$$a = bq' + r', \quad 0 \leqslant r' < |b|$$

则得 $bq + r = bq' + r'$. 不妨设 $r \leqslant r'$. 于是 $b(q - q') = r' - r$. 因为 $0 \leqslant r < |b|$,$0 \leqslant r' < |b|$,$r \leqslant r'$,所以 $0 \leqslant r' - r < |b|$. 于是

$$|q - q'| = \frac{r' - r}{|b|} < 1$$

由此得 $q - q' = 0$,即 $q = q'$. 从而得 $r = r'$. ▮

## 习 题 7.2

1. 用 $g(x)$ 除 $f(x)$,求商与余式.

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

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

2. 设 $f(x) = x^4 - 3x^2 + a_1x + a_0$,$g(x) = x^2 - 3x + 1$,求 $g(x)$ 整除 $f(x)$ 的充分必要条件.

3. 用综合除法求一次多项式 $g(x)$ 除 $f(x)$ 所得的商与余式:

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

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

4. 在第 3 题中,用 $g(x)$ 除 $f(x)$ 所得的商记作 $f_1(x)$. 接着用 $g(x)$ 除 $f_1(x)$,所得的商记作 $f_2(x)$. 然后用 $g(x)$ 除 $f_2(x)$,…,如此进行下去,将得到 $f(x)$ 的一个表达式. 对每一小题求出 $f(x)$ 的这种表达式.

*5. 设 $a,b \in Z$,如果有 $q \in Z$,使得

$$a = bq$$

则称 $b$ **整除** $a$，记作 $b|a$，此时 $b$ 叫做 $a$ 的**因数**（或**因子**），$a$ 叫做 $b$ 的**倍数**. 如果 $b$ 不能整除 $a$，则记作 $b\nmid a$. 证明下列有关整除的性质：

(1) 如果 $a|b$，并且 $b|a$（此时称 $a$ 与 $b$ 是**相伴的**），则 $a = \pm b$.

(2) 如果 $a|b$，并且 $b|c$，则 $a|c$（整除的**传递性**）.

(3) 如果 $a|b_i, i = 1,2,\cdots,r$，则对任意 $u_i \in Z, i = 1,2,\cdots,r$，有

$$a|(u_1b_1 + u_2b_2 + \cdots + u_rb_r)$$

(4) 如果 $a|b$，则对任意 $c \in Z$ 且 $c \neq 0$，有 $ca|cb$.

(5) 如果 $ca|cb$，且 $c \neq 0$，则 $a|b$.

(6) 如果 $b|a$，且 $a \neq 0$，则 $\frac{a}{b}|a$.

(7) 如果 $b|a$，且 $a \neq 0$，则 $|b| \leqslant |a|$.

# 阅读材料三

上册我们讲了数域 $K$ 上矩阵的运算，初等变换以及方阵的行列式. 它们的 定义都只用到数域 $K$ 中元素的加法和乘法运算. 因此我们可以把数域 $K$ 上的矩阵推广成整环 $R$ 上的矩阵：

**定义** 1　设 $R$ 是一个整环，由 $R$ 中的 $s \times n$ 个元素排成的 $s$ 行 $n$ 列的一张表称为 **$R$ 上的一个 $s \times n$ 矩阵**.

数域 $K$ 上矩阵的有关运算、初等变换、行列式的性质和概念对于整环 $R$ 上的矩阵仍然具有，但是要注意：数域 $K$ 中的非零数一定是可逆元，而整环 $R$ 中的非零元可能不是可逆元，因此凡是数域 $K$ 上矩阵的性质中用到"非零数"这一词，对于整环 $R$ 上的矩阵应当改成"可逆元". 譬如，对于整环 $R$ 上的矩阵，3°型初等行（列）变换应叙述成："用 $R$ 中一个可逆元乘某一行（列）". 又如，整环 $R$ 上 $n$ 级矩阵 $A$ 可逆的充分必要条件为 $|A|$ 是 $R$ 中的可逆元. 还要注意：数域 $K$ 中任一非零元的整数倍都不等于零，而对某些整环有可能发生非零元的某个正整数倍等于零，不过对于整数环 $Z$ 和数域 $K$ 上的一元多项式环 $K[x]$ 都不会发生这种情况. 需要注意的第三点是，整环 $R$ 上的矩阵的秩的概念通过子式来定义：

**定义** 2　设 $A$ 是整环 $R$ 上的一个非零矩阵，如果 $A$ 有一个 $r$ 级子式不为

零，而所有 $r+1$ 级子式（如果有的话）全为零，则称 $A$ 的**秩**为 $r$. 零矩阵的秩规定为零.

现在考虑整环 $K[\lambda]$ 上的矩阵，通常称它为 **$\lambda$- 矩阵**. $\lambda$- 矩阵的元素都是 $\lambda$ 的多项式. $\lambda$- 矩阵通常用 $A(\lambda), B(\lambda), \cdots$，表示. 由于 $K[\lambda]$ 中的可逆元都是 $K$ 中的非零数，因此 $\lambda$- 矩阵的 3° 型初等行（列）变换可叙述成：

3° 用 $K$ 中一个非零数乘某一行（列）.

同样地，$\lambda$- 矩阵 $A$ 可逆的充分必要条件可叙述成：$|A|$ 是 $K$ 中的非零数.

利用 $K[\lambda]$ 中的带余除法可以证明下面的关于 $\lambda$- 矩阵的相抵标准形的定理：

**定理 1** 任意一个非零的 $n$ 级 $\lambda$- 矩阵 $A(\lambda)$ 一定相抵于下述形式的 $\lambda$- 矩阵：

$$\operatorname{diag}\{d_1(\lambda), d_2(\lambda), \cdots, d_n(\lambda)\} \tag{1}$$

其中 $d_i(\lambda) \mid d_{i+1}(\lambda), i = 1,2,\cdots,n-1$；并且对于非零的 $d_i(\lambda)$，其首项系数为 1. 满足这些要求的 $\lambda$- 矩阵(1) 称为 $A(\lambda)$ 的一个相抵标准形.

**证明** 对 $\lambda$- 矩阵的级数 $n$ 作数学归纳法.

$n=1$ 时，$A(\lambda) = (a_{11}(\lambda))$，设 $a_{11}(\lambda)$ 的首项系数为 $b$，用 $b^{-1}$ 乘 $A(\lambda)$ 的第 1 行得 $A_1(\lambda) = (b^{-1}a_{11}(\lambda))$，于是 $A_1(\lambda)$ 就是 $A(\lambda)$ 的一个相抵标准形.

假设对于 $n-1$ 级 $\lambda$- 矩阵命题成立，来看 $n$ 级非零 $\lambda$- 矩阵 $A(\lambda) = (a_{ij}(\lambda))$. 不妨设 $a_{11}(\lambda) \neq 0$（否则对 $A(\lambda)$ 作两行互换或两列互换可使所得矩阵的 $(1,1)$ 元不为 0）. 令

$$S = \{G(\lambda) = (g_{ij}(\lambda)) \mid G(\lambda) \text{ 与 } A(\lambda) \text{ 相抵}, \quad g_{11}(\lambda) \neq 0\}$$

取 $B(\lambda) = (b_{ij}(\lambda)) \in S$，使得 $\deg b_{11}(\lambda) \leqslant \deg g_{11}(\lambda)$，对于一切的 $G(\lambda) = (g_{ij}(\lambda)) \in S$，并且 $b_{11}(\lambda)$ 的首项系数为 1. 我们来证 $b_{11}(\lambda)$ 能整除 $B(\lambda)$ 的第 1 行和第 1 列的所有元素.

假如 $b_{11}(\lambda) \nmid b_{i1}(\lambda)$，对于某个 $i \in \{2,3,\cdots,n\}$，作带余除法得

$$b_{i1}(\lambda) = h(\lambda)b_{11}(\lambda) + r(\lambda), \quad \deg r(\lambda) < \deg b_{11}(\lambda)$$

其中 $r(\lambda) \neq 0$. 把 $B(\lambda)$ 的第 1 行的 $-h(\lambda)$ 倍加到第 $i$ 行上得到 $B_1(\lambda)$，则 $B_1(\lambda)$ 的 $(i,1)$ 元为 $r(\lambda)$. 再把 $B_1(\lambda)$ 的第 $i$ 行与第 1 行互换得到 $B_2(\lambda)$，则 $B_2(\lambda)$ 的 $(1,1)$ 元为 $r(\lambda)$. 显然 $B_2(\lambda)$ 与 $A(\lambda)$ 相抵，且由于 $r(\lambda) \neq 0$，因此 $B_2(\lambda) \in S$. 由于 $\deg r(\lambda) < \deg b_{11}(\lambda)$，这与 $B(\lambda)$ 的取法矛盾. 所以 $b_{11}(\lambda)$ 能整除 $B(\lambda)$ 的第 1 列的所有元素. 同理可证 $b_{11}(\lambda)$ 能整除 $B(\lambda)$ 的第 1 行的所有元素. 于是对 $B(\lambda)$ 作一系列的 1° 型初等行变换和 1° 型初等列变换可将

$B(\lambda)$ 变成

$$C(\lambda)=\begin{pmatrix} b_{11}(\lambda) & 0 \\ 0 & D(\lambda) \end{pmatrix} \tag{2}$$

其中 $D(\lambda)$ 是 $n-1$ 级 $\lambda$-矩阵，显然 $C(\lambda)\in S$.

如果 $D(\lambda)=0$，则 $C(\lambda)$ 就是 $A(\lambda)$ 的一个相抵标准形.

如果 $D(\lambda)\neq 0$，则根据归纳假设，$D(\lambda)$ 相抵于下述形式的 $\lambda$-矩阵：

$$\mathrm{diag}\{d_2(\lambda),\cdots,d_n(\lambda)\} \tag{3}$$

其中 $d_i(\lambda)\mid d_{i+1}(\lambda)$，$i=2,\cdots,n-1$，并且非零的 $d_i(\lambda)$ 其首项系数为 1. 由于对 $D(\lambda)$ 作初等行(列)变换变成 $\lambda$-矩阵(3)时，对 $C(\lambda)$ 作相应的这些初等行(列)变换就变成

$$\mathrm{diag}\{b_{11}(\lambda),d_2(\lambda),\cdots,d_n(\lambda)\} \tag{4}$$

于是 $A(\lambda)$ 相抵于 $\lambda$-矩阵(4). 剩下只要证明 $b_{11}(\lambda)\mid d_2(\lambda)$. 假如 $b_{11}(\lambda)\nmid d_2(\lambda)$，作带余除法得

$$d_2(\lambda)=h_1(\lambda)b_{11}(\lambda)+r_1(\lambda),\quad \deg r_1(\lambda)<\deg b_{11}(\lambda)$$

其中 $r_1(\lambda)\neq 0$，对 $\lambda$-矩阵(4) 作下述初等变换：

$$\begin{pmatrix} b_{11}(\lambda) & & & \\ & d_2(\lambda) & & \\ & & \ddots & \\ & & & d_n(\lambda) \end{pmatrix} \xrightarrow{①+②} \begin{pmatrix} b_{11}(\lambda) & d_2(\lambda) & & \\ 0 & d_2(\lambda) & & \\ & & \ddots & \\ & & & d_n(\lambda) \end{pmatrix}$$

$$\xrightarrow[②+①\cdot(-h_1(\lambda))]{} \begin{pmatrix} b_{11}(\lambda) & r_1(\lambda) & & \\ 0 & d_2(\lambda) & & \\ & & \ddots & \\ & & & d_n(\lambda) \end{pmatrix} \xrightarrow[(①,②)]{} \begin{pmatrix} r_1(\lambda) & b_{11}(\lambda) & & \\ d_2(\lambda) & 0 & & \\ & & \ddots & \\ & & & d_n(\lambda) \end{pmatrix}$$

最后的 $\lambda$-矩阵记作 $L(\lambda)$，它属于 $S$. 但是 $\deg r_1(\lambda)<\deg b_{11}(\lambda)$，这与 $B(\lambda)$ 的取法矛盾. 因此 $b_{11}(\lambda)\mid d_2(\lambda)$.

根据数学归纳法原理，命题得证. ▌

**例 1** 求下述 $\lambda$-矩阵 $A(\lambda)$ 的一个相抵标准形

$$A(\lambda)=\begin{pmatrix} \lambda-1 & 3 & -4 \\ -4 & \lambda+7 & -8 \\ -6 & 7 & \lambda-7 \end{pmatrix}$$

**解** 首先要使 (1,1) 元变成能整除该矩阵的所有元素，从而可把第 1 行和第 1 列的其它元素变成 0. 然后把右下角的 2 级 $\lambda$-矩阵的(1,1) 元也变成能

整除该 2 级矩阵的所有元素，依次进行下去.

$$A(\lambda)\xrightarrow[③+②]{}\begin{pmatrix}\lambda-1 & 3 & -1\\ -4 & \lambda+7 & \lambda-1\\ -6 & 7 & \lambda\end{pmatrix}\xrightarrow[(①,③)]{}\begin{pmatrix}-1 & 3 & \lambda-1\\ \lambda-1 & \lambda+7 & -4\\ \lambda & 7 & -6\end{pmatrix}$$

$$\xrightarrow{\substack{②+①(\lambda-1)\\ ③+①\lambda}}\begin{pmatrix}-1 & 3 & \lambda-1\\ 0 & 4\lambda+4 & \lambda^2-2\lambda-3\\ 0 & 3\lambda+7 & \lambda^2-\lambda-6\end{pmatrix}$$

$$\longrightarrow\begin{pmatrix}1 & 0 & 0\\ 0 & \lambda+1 & \frac{1}{4}(\lambda-3)(\lambda+1)\\ 0 & 3\lambda+7 & \lambda^2-\lambda-6\end{pmatrix}$$

$$\xrightarrow[③+②[-\frac{1}{4}(\lambda-3)]]{}\begin{pmatrix}1 & 0 & 0\\ 0 & \lambda+1 & 0\\ 0 & 3\lambda+7 & \frac{1}{4}(\lambda-3)(\lambda+1)\end{pmatrix}$$

$$\xrightarrow{③+②\cdot(-3)}\begin{pmatrix}1 & 0 & 0\\ 0 & \lambda+1 & 0\\ 0 & 4 & \frac{1}{4}(\lambda-3)(\lambda+1)\end{pmatrix}$$

$$\longrightarrow\begin{pmatrix}1 & 0 & 0\\ 0 & 1 & (\lambda-3)(\lambda+1)\\ 0 & \lambda+1 & 0\end{pmatrix}$$

$$\longrightarrow\begin{pmatrix}1 & 0 & 0\\ 0 & 1 & (\lambda-3)(\lambda+1)\\ 0 & 0 & -(\lambda-3)(\lambda+1)^2\end{pmatrix}$$

$$\longrightarrow\begin{pmatrix}1 & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & (\lambda-3)(\lambda+1)^2\end{pmatrix}$$

最后一个 $\lambda$- 矩阵就是 $A(\lambda)$ 的一个相抵标准形.

练习：求下列 $\lambda$- 矩阵的相抵标准形：

(1) $\begin{pmatrix}\lambda & -1 & 0\\ 4 & \lambda-4 & 0\\ 2 & -1 & \lambda-2\end{pmatrix}$；　　(2) $\begin{pmatrix}\lambda-4 & 5 & -2\\ -5 & \lambda+7 & -3\\ -6 & 9 & \lambda-4\end{pmatrix}$.

## §3 最大公因式

从上一节知道，数域$K$上的一元多项式环$K[x]$有带余除法，这是$K[x]$的一个重要性质.这一节我们要由此出发推导出$K[x]$的另一个重要性质:$K[x]$中任何两个多项式都有最大公因式，并且$f(x)$与$g(x)$的最大公因式可以表成$f(x)$与$g(x)$的一个组合.

在$K[x]$中，如果$c(x)$既是$f(x)$的因式，又是$g(x)$的因式，则称$c(x)$是$f(x)$与$g(x)$的一个**公因式**.在$f(x)$与$g(x)$的所有公因式中，具有下述性质的公因式特别重要，即

**定义 1** $K[x]$中多项式$f(x)$与$g(x)$的一个公因式$d(x)$如果具有下述性质：对于$f(x)$与$g(x)$的任一公因式$c(x)$，都有$c(x)|d(x)$，则称$d(x)$是$f(x)$与$g(x)$的一个**最大公因式**.

从定义立即得出，两个零多项式的最大公因式就是0；对于任意多项式$f(x)$，$f(x)$是$f(x)$与0的一个最大公因式.

对于$K[x]$中任意两个多项式，是否存在它们的最大公因式？如果存在，如何找出它们的最大公因式？对于给定的两个多项式$f(x)$与$g(x)$，它们的最大公因式是否唯一？这些就是本节要讨论的问题.

我们先指出几个简单而有用的结论：

**命题 7.3.1** 设$f(x),g(x),p(x),q(x)$都是$K[x]$中的多项式，如果$f(x)$与$g(x)$的所有公因式组成的集合等于$p(x)$与$q(x)$的所有公因式组成的集合，那么$f(x)$与$g(x)$的最大公因式的集合等于$p(x)$与$q(x)$的最大公因式的集合.

**证明** 设$d(x)$是$f(x)$与$g(x)$的一个最大公因式，则$d(x)$是$p(x)$与$q(x)$的一个公因式.任取$p(x)$与$q(x)$的一个公因式$\varphi(x)$，则$\varphi(x)$也是$f(x)$与$g(x)$的公因式，从而$\varphi(x)|d(x)$.所以

$d(x)$是$p(x)$与$q(x)$的一个最大公因式.同理,$p(x)$与$q(x)$的任一最大公因式也是$f(x)$与$g(x)$的最大公因式. ▌

**推论** 7.3.1 设$f(x),g(x)\in K[x]$,$a$,$b$是$K$中非零数,则$f(x)$与$g(x)$的最大公因式的集合等于$af(x)$与$bg(x)$的最大公因式的集合.

**证明** 显然$f(x)$与$g(x)$的任一公因式是$af(x)$与$bg(x)$的公因式.容易说明$af(x)$与$bg(x)$的任一公因式也是$f(x)$与$g(x)$的公因式.于是由命题7.3.1立即得出结论. ▌

**引理** 7.3.1 在$K[x]$中,如果有等式

$$f(x)=h(x)g(x)+r(x) \tag{1}$$

成立,则$f(x)$与$g(x)$的最大公因式的集合等于$g(x)$与$r(x)$的最大公因式的集合.

**证明** 设$d(x)$是$f(x)$与$g(x)$的一个公因式,则$d(x)\mid f(x)$,并且$d(x)\mid g(x)$.因为从(1)式得

$$r(x)=f(x)-h(x)g(x)$$

所以$d(x)\mid r(x)$.即$d(x)$是$g(x)$与$r(x)$的一个公因式.现在任取$g(x)$与$r(x)$的一个公因式$c(x)$,由(1)式得,$c(x)\mid f(x)$.于是$c(x)$是$f(x)$与$g(x)$的公因式.由命题7.3.1立即得到所要求的结论. ▌

现在来证明这一节的主要结果:

**定理** 7.3.1 对于$K[x]$中任意两个多项式$f(x)$与$g(x)$,存在它们的一个最大公因式$d(x)$,并且$d(x)$可以表成$f(x)$与$g(x)$的一个组合,即有$K[x]$中多项式$u(x)$与$v(x)$,使

$$d(x)=u(x)f(x)+v(x)g(x) \tag{2}$$

**证明** 如果$g(x)=0$,则$f(x)$就是$f(x)$与$g(x)$的一个最大公因式,并且

$$f(x)=1\cdot f(x)+1\cdot 0$$

现在设$g(x)\neq 0$.据带余除法,有$h_1(x),r_1(x)\in K[x]$,使

$$f(x)=h_1(x)g(x)+r_1(x),\quad \deg r_1(x)<\deg g(x)$$

如果 $r_1(x) \neq 0$，则用 $r_1(x)$ 去除 $g(x)$，有 $h_2(x), r_2(x) \in K[x]$，使得

$$g(x) = h_2(x)r_1(x) + r_2(x)$$
$$\deg r_2(x) < \deg r_1(x)$$

又如果 $r_2(x) \neq 0$，则用 $r_2(x)$ 去除 $r_1(x)$，有 $h_3(x), r_3(x) \in K[x]$，使得

$$r_1(x) = h_3(x)r_2(x) + r_3(x)$$
$$\deg r_3(x) < \deg r_2(x)$$

如此辗转相除下去，显然，所得余式的次数不断降低，因此在有限次之后，必然有余式为零，即

$$
\begin{aligned}
& r_2(x) = h_4(x)r_3(x) + r_4(x) \\
& \quad \deg r_4(x) < \deg r_3(x) \\
& \cdots \quad \cdots \quad \cdots \quad \cdots \quad \cdots \quad \cdots \\
& r_{i-2}(x) = h_i(x)r_{i-1}(x) + r_i(x) \\
& \quad \deg r_i(x) < \deg r_{i-1}(x) \\
& \cdots \quad \cdots \quad \cdots \quad \cdots \quad \cdots \quad \cdots \\
& r_{s-3}(x) = h_{s-1}(x)r_{s-2}(x) + r_{s-1}(x) \\
& \quad \deg r_{s-1}(x) < \deg r_{s-2}(x) \\
& r_{s-2}(x) = h_s(x)r_{s-1}(x) + r_s(x) \\
& \quad \deg r_s(x) < \deg r_{s-1}(x) \\
& r_{s-1}(x) = h_{s+1}(x)r_s(x) + 0
\end{aligned}
$$

其中所有 $h_i(x), r_i(x) \in K[x]$. 从上述等式的最后一个式子得出，$r_s(x)$ 是 $r_{s-1}(x)$ 与 $r_s(x)$ 的一个最大公因式. 据引理 7.3.1 得，$r_s(x)$ 也是 $r_{s-2}(x)$ 与 $r_{s-1}(x)$ 的一个最大公因式；从而 $r_s(x)$ 也是 $r_{s-3}(x)$ 与 $r_{s-2}(x)$ 的一个最大公因式；依次往上推，$r_s(x)$ 也是 $f(x)$ 与 $g(x)$ 的一个最大公因式. 这证明了：**在对 $f(x)$ 与 $g(x)$ 作辗转相除时，最后一个不等于零的余式是 $f(x)$ 与 $g(x)$ 的一个最大公因式.**

从上述等式中倒数第二个得

$$r_s(x) = r_{s-2}(x) - h_s(x)r_{s-1}(x)$$

再由倒数第三式得

$$r_{s-1}(x) = r_{s-3}(x) - h_{s-1}(x)r_{s-2}(x)$$

代入上式得

$$r_s(x) = (1 + h_s(x)h_{s-1}(x))r_{s-2}(x) - h_s(x)r_{s-3}(x)$$

同理用更上面的等式逐个地消去 $r_{s-2}(x), r_{s-3}(x), \cdots, r_1(x)$，可得

$$r_s(x) = u(x)f(x) + v(x)g(x)$$

其中 $u(x), v(x) \in K[x]$. ▌

定理 7.3.1 的证明给出了求两个多项式的最大公因式的方法，称它为**辗转相除法**.

任意给定 $K[x]$ 中两个多项式 $f(x)$ 与 $g(x)$，它们的最大公因式是否唯一？设 $d_1(x), d_2(x)$ 都是 $f(x)$ 与 $g(x)$ 的最大公因式，据定义得，$d_1(x) \mid d_2(x)$ 且 $d_2(x) \mid d_1(x)$. 因此 $d_1(x)$ 与 $d_2(x)$ 相伴，即 $d_1(x)$ 与 $d_2(x)$ 仅相差一个非零数因子. 这说明：两个多项式的最大公因式在相伴的意义下是唯一确定的. 容易看出，两个不全为零的多项式的最大公因式一定是非零多项式，在这个情形，我们约定，用

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

来表示首项系数是 1 的那个最大公因式.

由推论 7.3.1 得出，设 $f(x), g(x)$ 不全为零，则对于 $a, b \in K$ 且 $a \neq 0, b \neq 0$，有

$$(f(x), g(x)) = (af(x), bg(x))$$

**例 1** 设

$$f(x) = x^3 + x^2 - 7x + 2, \quad g(x) = 3x^2 - 5x - 2$$

求 $(f(x), g(x))$，并且把它表示成 $f(x)$ 与 $g(x)$ 的一个组合.

**解** 根据上面所述，在作辗转相除法时，可以用适当的非零数去乘被除式或者除式，以便使计算简单一些.

| | $g(x)$ | $3f(x)$ | |
|---|---|---|---|
| $h_2(x)=3x+1$ | $3x^2-5x-2$ | $3x^3+3x^2-21x+6$ | $x+\frac{8}{3}=h_1(x)$ |
| | $3x^2-6x$ | $3x^3-5x^2-2x$ | |
| | $x-2$ | $8x^2-19x+6$ | |
| | $x-2$ | $8x^2-\frac{40}{3}x-\frac{16}{3}$ | |
| | $0$ | $r_1(x)=-\frac{17}{3}x+\frac{34}{3}$ | |
| | | $-\frac{3}{17}r_1(x)=x-2$ | |

因为最后一个不等于零的余式是 $r_1(x)$，所以

$$(f(x),g(x)) = x-2$$

把上述辗转相除过程写出来就是

$$3f(x) = \left(x+\frac{8}{3}\right)g(x) + r_1(x)$$

$$g(x) = (3x+1)\left(-\frac{3}{17}r_1(x)\right) + 0$$

于是

$$\begin{aligned}(f(x),g(x)) &= -\frac{3}{17}r_1(x)\\ &= -\frac{3}{17}\left[3f(x) - \left(x+\frac{8}{3}\right)g(x)\right]\\ &= -\frac{9}{17}f(x) + \frac{1}{17}(3x+8)g(x)\end{aligned}$$

一个极端情形是两个多项式的最大公因式是非零数(即零次多项式)，这个情形很重要，我们专门来讨论它.

**定义 2** $K[x]$ 中的两个多项式 $f(x),g(x)$，如果 $(f(x),g(x))=1$，则称 $f(x)$ 与 $g(x)$ **互素**.

显然，如果两个多项式互素，那么它们除去零次多项式外没有其他的公因式，反之亦然.

下面我们给出两个多项式互素的一个充分必要条件，它非常

有用.

**定理** 7.3.2　$K[x]$ 中两个多项式 $f(x)$ 与 $g(x)$ 互素的充分必要条件是有 $K[x]$ 中的多项式 $u(x),v(x)$,使

$$u(x)f(x)+v(x)g(x)=1 \tag{3}$$

**证明**　必要性从定理 7.3.1 立即得出.

现在证充分性. 设(3) 式成立. 因为$(f(x),g(x))\mid f(x)$,并且$(f(x),g(x))\mid g(x)$,所以从(3) 式得,$(f(x),g(x))\mid 1$. 于是

$$(f(x),g(x))=1$$ ▍

利用定理 7.3.2 可以证明下面关于整除性的两条重要性质:

**定理** 7.3.3　在 $K[x]$ 中,如果

$$f(x)\mid g(x)h(x),\quad 且(f(x),g(x))=1$$

则

$$f(x)\mid h(x)$$

**证明**　因为 $(f(x),g(x))=1$,所以有 $u(x),v(x)\in K[x]$,使得

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

等式两边乘以 $h(x)$,得

$$u(x)f(x)h(x)+v(x)g(x)h(x)=h(x) \tag{4}$$

因为 $f(x)\mid g(x)h(x)$,所以 $f(x)$ 整除(4) 式左端,从而

$$f(x)\mid h(x)$$ ▍

**推论** 7.3.2　在 $K[x]$ 中,如果

$$f(x)\mid h(x),\quad g(x)\mid h(x),\quad 且(f(x),g(x))=1$$

则

$$f(x)g(x)\mid h(x)$$

**证明**　因为 $f(x)\mid h(x)$,所以有 $p(x)\in K[x]$,使

$$h(x)=p(x)f(x) \tag{5}$$

因为 $g(x)\mid h(x)$,所以 $g(x)\mid p(x)f(x)$. 因为$(g(x),f(x))=1$,所以 $g(x)\mid p(x)$. 于是有 $q(x)\in K[x]$,使得 $p(x)=q(x)g(x)$. 代入(5) 式得

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

所以 $f(x)g(x)\mid h(x)$. ▍

利用定理 7.3.2 还可以证明关于多项式互素的一些性质，例如：

**定理** 7.3.4　在 $K[x]$ 中，如果

$$(f(x),h(x))=1,\quad (g(x),h(x))=1$$

则

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

**证明**　因为 $(f(x),h(x))=1$，$(g(x),h(x))=1$，所以有 $K[x]$ 中的多项式 $u_i(x),v_i(x),i=1,2$，使得

$$u_1(x)f(x)+v_1(x)h(x)=1$$

$$u_2(x)g(x)+v_2(x)h(x)=1$$

将上面两个等式相乘，得

$$u_1(x)u_2(x)f(x)g(x)+[u_1(x)f(x)v_2(x)$$

$$+v_1(x)u_2(x)g(x)+v_1(x)v_2(x)h(x)]h(x)=1$$

据定理 7.3.2 得，$(f(x)g(x),h(x))=1$. ▌

最大公因式和互素的概念可以推广到 $n(n>2)$ 个多项式的情形.

**定义** 3　在 $K[x]$ 中，如果多项式 $c(x)$ 能整除多项式 $f_1(x)$，$f_2(x),\cdots,f_n(x)$ 的每一个，那么 $c(x)$ 叫做这 $n$ 个多项式的一个**公因式**. 设 $d(x)$ 是 $f_1(x),\cdots,f_n(x)$ 的一个公因式，如果 $d(x)$ 具有下述性质：$f_1(x),\cdots,f_n(x)$ 的每一个公因式都能整除 $d(x)$，则 $d(x)$ 称为 $f_1(x),\cdots,f_n(x)$ 的一个**最大公因式**.

用数学归纳法可以证明，在 $K[x]$ 中，任意 $n$ 个 $(n\geqslant 2)$ 多项式 $f_1(x),\cdots,f_n(x)$ 的最大公因式存在，并且如果 $d_1(x)$ 是 $f_1(x),\cdots,f_{n-1}(x)$ 的一个最大公因式，则 $d_1(x)$ 与 $f_n(x)$ 的最大公因式是 $f_1(x),\cdots,f_{n-1}(x),f_n(x)$ 的最大公因式. 因此可以逐次用辗转相除法求出 $n$ 个多项式的一个最大公因式.

从定义 3 立即得出，$n$ 个多项式 $f_1(x),\cdots,f_n(x)$ 的最大公因式在相伴的意义下是唯一的. 对于 $n$ 个不全为零的多项式 $f_1(x),\cdots,f_n(x)$，我们约定，用

$$(f_1(x),f_2(x),\cdots,f_n(x))$$

表示首项系数是 1 的那个最大公因式. 由上面一段知

$$\begin{aligned}&(f_1(x),f_2(x),\cdots,f_n(x))\\&=((f_1(x),\cdots,f_{n-1}(x)),f_n(x))\end{aligned}\tag{6}$$

从(6)式以及定理 7. 3. 1 得，有 $K[x]$ 中多项式 $u_i(x)$，$i=1,\cdots,n$，使得

$$\begin{aligned}&u_1(x)f_1(x)+u_2(x)f_2(x)+\cdots+u_n(x)f_n(x)\\&=(f_1(x),f_2(x),\cdots,f_n(x))\end{aligned}\tag{7}$$

**定义** 4　$K[x]$ 中 $n(n\geqslant 2)$ 个多项式 $f_1(x),f_2(x),\cdots,f_n(x)$，如果 $(f_1(x),f_2(x),\cdots,f_n(x))=1$，则称 $f_1(x),f_2(x),\cdots,f_n(x)$ **互素**.

与定理 7. 3. 2 一样可以证明：在 $K[x]$ 中，$n$ 个多项式 $f_1(x)$，$f_2(x),\cdots,f_n(x)$ 互素的充分必要条件是有 $K[x]$ 中多项式 $u_1(x)$，$u_2(x),\cdots,u_n(x)$，使得

$$u_1(x)f_1(x)+u_2(x)f_2(x)+\cdots+u_n(x)f_n(x)=1\tag{8}$$

注意：$n(n>2)$ 个多项式互素时，它们不一定两两互素. 例如，设

$$f_1(x)=x+1,\quad f_2(x)=x^2+3x+2,\quad f_3(x)=x-1$$

容易算出

$$(f_1(x),f_2(x))=x+1,\quad (f_1(x),f_2(x),f_3(x))=1$$

因此，$f_1(x),f_2(x),f_3(x)$ 互素，但是 $f_1(x)$ 与 $f_2(x)$ 不互素.

我们还要指出一点，设 $K$ 与 $F$ 都是数域，并且 $K\subset F$. 设 $f(x),g(x)\in K[x]$，也可以把 $f(x)$ 与 $g(x)$ 看成 $F[x]$ 中的多项式. 注意 $f(x)$ 与 $g(x)$ 在 $K[x]$ 中的公因式与它们在 $F[x]$ 中的公因式不一定相同，例如设

$$f(x)=x^2+1,\quad g(x)=x^3+x^2+x+1$$

则 $f(x)$ 与 $g(x)$ 在 $R[x]$ 中没有一次公因式，但是 $f(x)$ 与 $g(x)$ 在 $C[x]$ 中有一次公因式 $x+\mathrm{i}$ 与 $x-\mathrm{i}$. 容易看出它们在 $R[x]$ 中的最大公因式是 $x^2+1$，在 $C[x]$ 中的最大公因式也是 $x^2+1$. 一般

地，对于 $K[x]$ 中任意两个多项式 $f(x)$ 与 $g(x)$，它们在 $K[x]$ 中的首项系数为 1 的最大公因式与它们在 $F[x]$ 中的首项系数为 1 的最大公因式相同. 也就是说，在数域扩大时，$f(x)$ 与 $g(x)$ 的最大公因式不改变. 证明如下：

若 $f(x)=g(x)=0$，则 $f(x)$ 与 $g(x)$ 在 $K[x]$ 中的最大公因式是零多项式，在 $F[x]$ 中的最大公因式也是零多项式.

下面设 $f(x)$ 与 $g(x)$ 不全为零. 设 $d_1(x)$ 是 $f(x)$ 与 $g(x)$ 在 $K[x]$ 中的最大公因式（首项系数为 1），设 $d_2(x)$ 是 $f(x)$ 与 $g(x)$ 在 $F[x]$ 中的最大公因式（首项系数为 1）. 在 $K[x]$ 中对 $f(x)$ 与 $g(x)$ 做辗转相除法，设最后一个余式为 $r_s(x)$，其首项系数为 $c$，则 $d_1(x)=\frac{1}{c}r_s(x)$. 由于每一步带余除法也可看成是在 $F[x]$ 中进行的（根据带余除法的唯一性），因此最后一个余式 $r_s(x)$ 也是 $f(x)$ 与 $g(x)$ 在 $F[x]$ 中的一个最大公因式，从而

$$d_2(x)=\frac{1}{c}r_s(x)=d_1(x)$$

我们再举一个例子以便提供一些解题方法.

* **例 2** 证明：在 $K[x]$ 中，如果 $(f(x),g(x))=1$，并且 $\deg f(x)>0$，$\deg g(x)>0$，则 $K[x]$ 中存在唯一的一对多项式 $u(x),v(x)$，使得

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

并且 $\deg u(x)<\deg g(x)$，$\deg v(x)<\deg f(x)$.

**证明** 因为 $(f(x),g(x))=1$，所以 $K[x]$ 中有多项式 $p(x)$，$q(x)$，使得

$$p(x)f(x)+q(x)g(x)=1 \tag{9}$$

用 $g(x)$ 去除 $p(x)$，据带余除法，有 $K[x]$ 中的多项式 $h(x)$，$r(x)$，使得

$$p(x)=h(x)g(x)+r(x),\quad \deg r(x)<\deg g(x) \tag{10}$$

因为 $\deg g(x)>0$，所以从 (10) 式和 (9) 式得出 $r(x)\neq 0$. 把 (10) 的第一式代入 (9) 式中得到

$$r(x)f(x)+(h(x)f(x)+q(x))g(x)=1 \tag{11}$$

令 $u(x)=r(x), v(x)=h(x)f(x)+q(x)$,则(11)式成为

$$u(x)f(x)+v(x)g(x)=1 \tag{12}$$

已指出 $\quad \deg u(x)=\deg r(x)<\deg g(x)$

假如 $\quad \deg v(x)\geqslant \deg f(x)$

则由于 $\quad \deg g(x)>0, \quad \deg f(x)>0, \quad u(x)\neq 0$

而得到
$$\begin{aligned}\deg(v(x)g(x)) &= \deg v(x)+\deg g(x)\\ &\geqslant \deg f(x)+\deg g(x)\\ &>\deg f(x)+\deg u(x)\\ &=\deg(u(x)f(x))\end{aligned}$$

从而
$$\begin{aligned}&\deg(u(x)f(x)+v(x)g(x))\\ &=\deg(v(x)g(x))\geqslant \deg f(x)+\deg g(x)>0\end{aligned}$$

这与(12)式矛盾. 所以 $\deg v(x)<\deg f(x)$. 存在性得证.

现在来证唯一性. 假设 $K[x]$ 中还有一对多项式 $u_1(x)$, $v_1(x)$,使得

$$u_1(x)f(x)+v_1(x)g(x)=1$$

并且 $\quad \deg u_1(x)<\deg g(x), \quad \deg v_1(x)<\deg f(x)$

则
$$(u_1(x)-u(x))f(x)=(v(x)-v_1(x))g(x) \tag{13}$$

由于 $(f(x),g(x))=1$,从(13)式得出

$$g(x)\mid(u_1(x)-u(x))$$

假如 $u_1(x)-u(x)\neq 0$,则

$$\deg(u_1(x)-u(x))\geqslant \deg g(x)$$

另一方面
$$\begin{aligned}&\deg(u_1(x)-u(x))\\ &\leqslant \max\{\deg u_1(x),\deg u(x)\}<\deg g(x)\end{aligned}$$

矛盾. 所以 $u_1(x)-u(x)=0$,即

$$u_1(x)=u(x)$$

从(13)式得 $v(x)-v_1(x)=0$,即

$$v(x)=v_1(x) \qquad \blacksquare$$

# 习 题 7.3

1. 求 $f(x)$ 与 $g(x)$ 的最大公因式，并且把 $(f(x),g(x))$ 表示成 $f(x)$ 与 $g(x)$ 的一个组合：

(1) $f(x)=x^4+9x^3+19x^2-9x-20$,

$g(x)=2x^3+3x^2-17x-30$;

(2) $f(x)=x^4+6x^3-6x^2+6x-7$,

$g(x)=x^3+x^2-7x+5$.

2. 证明：在 $K[x]$ 中，如果 $d(x)$ 是 $f(x)$ 与 $g(x)$ 的一个组合，并且 $d(x)$ 是 $f(x)$ 与 $g(x)$ 的一个公因式，则 $d(x)$ 是 $f(x)$ 与 $g(x)$ 的一个最大公因式.

3. 证明：在 $K[x]$ 中，$(f,g)h$ 是 $fh$ 与 $gh$ 的一个最大公因式；特别地，若 $h(x)$ 的首项系数为 1，则

$$(fh,gh)=(f,g)h$$

4. 证明：在 $K[x]$ 中，如果 $f(x),g(x)$ 不全为零，则

$$\left(\frac{f(x)}{(f(x),g(x))},\frac{g(x)}{(f(x),g(x))}\right)=1$$

5. 证明：在 $K[x]$ 中，如果 $f(x),g(x)$ 不全为零，并且

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

则 $(u(x),v(x))=1$.

6. 证明：在 $K[x]$ 中，如果 $(f,g)=1$，那么

$$(fg,f-g)=1$$

7. 设 $f(x),g(x)\in K[x]$，并且 $a,b,c,d\in K$，使得

$$ad-bc\neq 0$$

证明：$(af+bg,cf+dg)=(f,g)$.

8. 证明：在 $K[x]$ 中，如果 $(f(x),g(x))=1$，则对任意正整数 $m$，有

$$(f(x^m),g(x^m))=1$$

9. 证明：$K[x]$ 中两个非零多项式 $f(x)$ 与 $g(x)$ 不互素的充分必要条件是，存在两个非零多项式 $u(x),v(x)$，使

$$u(x)f(x)=v(x)g(x)$$

$$\deg u(x)<\deg g(x),\quad \deg v(x)<\deg f(x)$$

10. 设 $f(x), g(x) \in K[x]$，$K[x]$ 中的一个多项式 $m(x)$ 称为 $f(x)$ 与 $g(x)$ 的一个**最小公倍式**，如果

1) $f(x) \mid m(x), g(x) \mid m(x)$；

2) $f(x)$ 与 $g(x)$ 的任一公倍式(即 $K[x]$ 中既能被 $f(x)$ 整除，又能被 $g(x)$ 整除的多项式)都是 $m(x)$ 的倍式.

(1) 证明 $K[x]$ 任意两个多项式都有最小公倍式，并且在相伴的意义下是唯一的；

(2) 用 $[f(x), g(x)]$ 表示首项系数是1的那个最小公倍式. 证明：如果 $f(x), g(x)$ 的首项系数都是 1，则

$$[f(x), g(x)] = \frac{f(x)g(x)}{(f(x), g(x))}$$

*11. 设 $A \in M_n(K)$，$f(x), g(x) \in K[x]$. 证明：如果 $d(x)$ 是 $f(x)$ 与 $g(x)$ 的一个最大公因式，则齐次线性方程组 $d(A)X = 0$ 的解空间等于 $f(A)X = 0$ 的解空间与 $g(A)X = 0$ 的解空间的交.

*12. 设 $A \in M_n(K)$，$f_1(x), f_2(x) \in K[x]$，记 $f(x) = f_1(x)f_2(x)$. 证明：如果 $(f_1(x), f_2(x)) = 1$，则 $f(A)X = 0$ 的任一个解可以唯一地表示成 $f_1(A)X = 0$ 的一个解与 $f_2(A)X = 0$ 的一个解的和.

# 阅读材料四

在整数环 $Z$ 中也有带余除法，从而可以证明任意两个整数都有最大公因数. 先给出定义.

**定义 1** 设 $a, b$ 是整数，若整数 $c$ 既是 $a$ 的因数，又是 $b$ 的因数，则称 $c$ 是 $a$ 与 $b$ 的一个**公因数**(或**公因子**). 设 $d$ 是 $a$ 与 $b$ 的一个公因数，如果 $a$ 与 $b$ 的每一个公因数都能整除 $d$，则称 $d$ 是 $a$ 与 $b$ 的一个**最大公因数**(或**最大公因子**).

**引理 1** 在 $Z$ 中，如果有等式

$$a = bq + c$$

成立，则 $a$ 与 $b$ 的最大公因数是 $b$ 与 $c$ 的最大公因数，反之亦然.

**证明** 类似于引理 7.3.1 的证明. ▌

**定理 1** 任意给定两个整数 $a, b$，都存在它们的一个最大公因数 $d$，并且

$d$ 可以表示成 $a$ 与 $b$ 的一个组合，即存在整数 $u,v$，使得

$$d = ua + vb \tag{1}$$

**证明** 如果 $b = 0$，则 $a$ 就是 $a$ 与 $b$ 的一个最大公因数，并且

$$a = 1 \cdot a + 1 \cdot 0$$

现在设 $b \neq 0$. 容易看出 $a$ 与 $b$ 的最大公因数是 $a$ 与 $-b$ 的最大公因数，反之亦然. 因此我们不妨设 $b > 0$. 据带余除法，有整数 $q_1, r_1$，使得

$$a = bq_1 + r_1, \quad 0 \leqslant r_1 < b$$

如果 $r_1 \neq 0$，则用 $r_1$ 去除 $b$，有整数 $q_2, r_2$，使得

$$b = r_1 q_2 + r_2, \quad 0 \leqslant r_2 < r_1$$

如果 $r_2 \neq 0$，再用 $r_2$ 去除 $r_1$. 如此辗转相除下去，显然所得余数（正整数）不断减小，因此在有限次后，必然有余数为零. 以下类似于定理 7.3.1 的相应部分的证明，可证出最后一个不等于零的余数就是 $a$ 与 $b$ 的一个最大公因数，并且它可以表示成 $a$ 与 $b$ 的一个组合. ▍

从定义 1 可知，$a$ 与 $b$ 的最大公因数在相伴的意义下（即相差一个正负号）是唯一的. 对于不全为零的整数 $a,b$，我们约定用 $(a,b)$ 表示正的那个最大公因数.

**定义 2** 设 $a,b \in Z$，如果 $(a,b) = 1$，则称 $a$ 与 $b$ **互素**.

显然，如果两个整数互素，那么它们除去 $\pm 1$ 以外没有其他公因数，反之亦然.

**定理 2** 两个整数 $a$ 与 $b$ 互素的充分必要条件是有整数 $u,v$，使得

$$ua + vb = 1$$

**证明** 必要性从定理 1 立即得出. 充分性易证. ▍

用定理 2 可以得出

**定理 3** 在 $Z$ 中，如果 $a \mid bc$，并且 $(a,b) = 1$，则 $a \mid c$.

证明留给读者.

**定义 3** 在 $Z$ 中，如果 $c$ 是 $a_1, a_2, \cdots, a_n$ 中每一个的因数，则称 $c$ 是 $a_1, a_2, \cdots, a_n$ 的一个**公因数**. 设 $d$ 是 $a_1, a_2, \cdots, a_n$ 的一公因数，如果 $a_1, a_2, \cdots, a_n$ 的每一个公因数都能整除 $d$，则称 $d$ 是 $a_1, a_2, \cdots, a_n$ 的一个**最大公因数**.

用数学归纳法可以证明，任意 $n(n \geqslant 2)$ 个整数都有最大公因数，并且如果 $d_1$ 是 $a_1, a_2, \cdots, a_{n-1}$ 的一个最大公因数，则 $d_1$ 与 $a_n$ 的最大公因数是 $a_1, a_2, \cdots, a_{n-1}, a_n$ 的最大公因数. 因此可以用辗转相除法求出 $n$ 个整数的一个最大公因数.

从定义 3 得出，$n$ 个整数 $a_1,a_2,\cdots,a_n$ 的最大公因数在相伴的意义下是唯一的，对于 $n$ 个不全为零的整数 $a_1,a_2,\cdots,a_n$，我们约定用 $(a_1,a_2,\cdots,a_n)$ 表示正的那个最大公因数. 从上面一段知

$$(a_1,a_2,\cdots,a_n) = ((a_1,a_2,\cdots,a_{n-1}),a_n) \tag{2}$$

从(2)式以及定理 1 可得出，对于 $n$ 个不全为零的整数 $a_1,a_2,\cdots,a_n$，存在 $n$ 个整数 $u_1,u_2,\cdots,u_n$，使得

$$u_1a_1 + u_2a_2 + \cdots + u_na_n = (a_1,a_2,\cdots,a_n) \tag{3}$$

**定义** 4　对于 $n$ 个整数 $a_1,a_2,\cdots,a_n$，如果 $(a_1,a_2,\cdots,a_n) = 1$，则称 $a_1,a_2,\cdots,a_n$ **互素**.

$n$ 个整数 $a_1,a_2,\cdots,a_n$ 互素的充分必要条件是有整数 $u_1,u_2,\cdots,u_n$，使得

$$u_1a_1 + u_2a_2 + \cdots + u_na_n = 1$$

**定义** 5　设 $a,b$ 是整数，一个整数 $m$ 称为 $a$ 与 $b$ 的**最小公倍数**，如果

1)　$a|m,b|m$;

2)　从 $a|l,b|l$，可推出 $m|l$.

可以证明，任意两个整数都有最小公倍数，并且在相伴的意义下是唯一的. 对于不全为零的整数 $a,b$，我们用 $[a,b]$ 表示正的那个最小公倍数. 对于 $a > 0,b > 0$，有

$$[a,b](a,b) = ab \tag{4}$$

证明留给读者.

类似于定义 5，可给出 $n(n > 2)$ 个整数的最小公倍数的定义，可以证明它们存在，并且在相伴意义下唯一. 对于不全为零的整数 $a_1,a_2,\cdots,a_n$，用 $[a_1,a_2,\cdots,a_n]$ 表示正的那个最小公倍数，可证

$$[a_1,a_2,\cdots,a_n] = [[a_1,a_2,\cdots,a_{n-1}],a_n]$$

# 阅读材料五

下面介绍最大公因式理论在 $\lambda$- 矩阵(即整环 $K[\lambda]$ 上的矩阵) 中的应用.

**定义** 1　设 $A(\lambda)$ 是一个 $s \times n\lambda$- 矩阵，对于正整数 $k(1 \leqslant k \leqslant \min\{s,n\})$，$A(\lambda)$ 的所有 $k$ 级子式的首项系数为 1 的最大公因式 $D_k(\lambda)$ 称为 $A(\lambda)$ 的 $k$ 级行列式因子.

**例 1** 求下述 $\lambda$-矩阵 $A(\lambda)$ 的各级行列式因子：

$$A(\lambda)=\begin{pmatrix}\lambda-2 & 0 & 0\\ 0 & \lambda-2 & 1\\ 0 & 0 & \lambda-2\end{pmatrix}$$

**解** $A(\lambda)$ 有一个 1 级子式为 1，所以 $D_1(\lambda)=1$.

$A(\lambda)$ 的非零 2 级子式有：$(\lambda-2)^2,\lambda-2$. 因此 $D_2(\lambda)=\lambda-2$.

$A(\lambda)$ 的 3 级子式只有一个 $|A(\lambda)|$，因此 $D_3(\lambda)=(\lambda-2)^3$.

行列式因子有一个重要性质：它在初等变换下不变. 即我们有下述定理：

**定理 1** 相抵的 $\lambda$-矩阵有相同的秩与相同的各级行列式因子.

**证明** 把 $A(\lambda)$ 的第 $i$ 行的 $h(\lambda)$ 倍加到第 $j$ 行上得到 $B(\lambda)$. $B(\lambda)$ 中包含第 $j$ 行但不包含第 $i$ 行的 $k$ 级子式是 $A(\lambda)$ 的一个 $k$ 级子式与另一个 $k$ 级子式的 $h(\lambda)$ 倍或者 $-h(\lambda)$ 倍的和. $B(\lambda)$ 的其余 $k$ 级子式等于 $A(\lambda)$ 中对应的 $k$ 级子式. 因此 $A(\lambda)$ 的所有 $k$ 级子式的公因式能整除 $B(\lambda)$ 的所有 $k$ 级子式. 由于 $B(\lambda)$ 经过把第 $i$ 行的 $-h(\lambda)$ 倍加到第 $j$ 行上便得到 $A(\lambda)$，因此由刚才证得的结论知道，$B(\lambda)$ 的所有 $k$ 级子式的公因式能整除 $A(\lambda)$ 的所有 $k$ 级子式，所以 $A(\lambda)$ 的 $k$ 级行列式因子等于 $B(\lambda)$ 的 $k$ 级行列式因子.

把 $A(\lambda)$ 的第 $i$ 行与第 $j$ 行互换得到 $C(\lambda)$. $C(\lambda)$ 的任一 $k$ 级子式或者等于 $A(\lambda)$ 的某个 $k$ 级子式，或者与 $A(\lambda)$ 的某一 $k$ 级子式反号. 因此 $A(\lambda)$ 与 $C(\lambda)$ 的 $k$ 级行列式因子相等.

把 $A(\lambda)$ 的第 $i$ 行乘以 $K$ 中的一个非零数 $c$ 得到 $G(\lambda)$. $G(\lambda)$ 的每一个 $k$ 级子式或者等于 $A(\lambda)$ 的某个 $k$ 级子式，或者等于 $A(\lambda)$ 的某个 $k$ 级子式的 $c$ 倍. 因此 $G(\lambda)$ 的 $k$ 级行列式因子等于 $A(\lambda)$ 的 $k$ 级行列式因子.

以上表明 $\lambda$-矩阵经过一次初等行变换，$k$ 级行列式因子不变. 同理，$\lambda$-矩阵经过一次初等列变换，$k$ 级行列式因子也不变. 从而相抵的 $\lambda$-矩阵有相同的各级行列式因子.

设 $A(\lambda)$ 与 $B(\lambda)$ 相抵，如果 $A(\lambda)$ 的秩为 $r(>0)$，那么 $A(\lambda)$ 有一个 $r$ 级子式不为零，而所有 $r+1$ 级子式（如果有的话）全为零. 因此 $A(\lambda)$ 的 $r$ 级行列式因子不为零，而 $r+1$ 级行列式因子为零. 从而 $B(\lambda)$ 至少有一个 $r$ 级子式不为零，而所有 $r+1$ 级子式全为零. 所以 $B(\lambda)$ 的秩为 $r$. 如果 $A(\lambda)$ 的秩为零，则 $A(\lambda)=0$，从而 $B(\lambda)=0$. 于是 $B(\lambda)$ 的秩为零. ▎

阅读材料三的定理 1 指出，每个非零的 $n$ 级 $\lambda$-矩阵 $A(\lambda)$ 相抵于下述形式的 $\lambda$-矩阵：

$$\mathrm{diag}\{d_1(\lambda),d_2(\lambda),\cdots,d_n(\lambda)\} \tag{1}$$

其中 $d_i(\lambda)\mid d_{i+1}(\lambda), i=1,2,\cdots,n-1$，并且非零的 $d_i(\lambda)$ 其首项系数为 1. 现在来计算 $\lambda$- 矩阵(1) 各级行列式因子. 设

$$d_i(\lambda)\neq 0,\quad i=1,\cdots,r;\quad d_{r+1}(\lambda)=\cdots=d_n(\lambda)=0$$

则

$$\begin{aligned}
D_1(\lambda)&=(d_1(\lambda),d_2(\lambda),\cdots,d_r(\lambda),0)=d_1(\lambda)\\
D_2(\lambda)&=(d_1(\lambda)d_2(\lambda),d_1(\lambda)d_3(\lambda),\cdots,d_{r-1}(\lambda)d_r(\lambda),0)\\
&=d_1(\lambda)d_2(\lambda)\\
D_3(\lambda)&=d_1(\lambda)d_2(\lambda)d_3(\lambda)\\
&\cdots\ \cdots\ \cdots\ \cdots\ \cdots\ \cdots\\
D_r(\lambda)&=d_1(\lambda)d_2(\lambda)\cdots d_r(\lambda)\\
D_{r+1}(\lambda)&=\cdots=D_n(\lambda)=0
\end{aligned} \tag{2}$$

根据定理 1, $D_1(\lambda),D_2(\lambda),\cdots,D_n(\lambda)$ 也是 $A(\lambda)$ 的各级行列式因子. 由上述式子得到

$$\begin{aligned}
&d_1(\lambda)=D_1(\lambda),\quad d_2(\lambda)=\frac{D_2(\lambda)}{D_1(\lambda)}\\
&d_3(\lambda)=\frac{D_3(\lambda)}{D_2(\lambda)},\cdots,d_r(\lambda)=\frac{D_r(\lambda)}{D_{r-1}(\lambda)}
\end{aligned} \tag{3}$$

这表明 $A(\lambda)$ 的相抵标准形中主对角线上的非零元可以用 $A(\lambda)$ 的行列式因子计算出. 因此 $A(\lambda)$ 的相抵标准形中主对角线上的非零元是唯一确定的，其数目等于 $A(\lambda)$ 的秩. 这样我们证明了下面的定理 2：

**定理 2** $n$ 级 $\lambda$- 矩阵 $A(\lambda)$ 的相抵标准形是唯一的.

**定义 2** $n$ 级 $\lambda$- 矩阵 $A(\lambda)$ 的相抵标准形中主对角线上的非零元 $d_1(\lambda)$, $d_2(\lambda),\cdots,d_r(\lambda)$ 称为 $A(\lambda)$ 的**不变因子**.

**定理 3** 两个 $n$ 级 $\lambda$- 矩阵相抵的充分必要条件是它们有相同的不变因子，或者有相同的各级行列式因子.

**证明** 设 $A(\lambda)$ 与 $B(\lambda)$ 相抵，则它们有相同的各级行列式因子. 从而有相同的不变因子.

设 $A(\lambda)$ 与 $B(\lambda)$ 有相同的各级行列式因子，则它们有相同的不变因子，从而它们的相抵标准形一样，于是它们相抵. ▍

由(3)式和(2)式看出，$A(\lambda)$ 的各级行列式因子有下述关系：

$$D_i(\lambda)\mid D_{i+1}(\lambda),\quad i=1,2,\cdots,n-1 \tag{4}$$

从(4)式看出，在计算 $A(\lambda)$ 的各级行列式因子时，往往先求最高级的行列式

因子，较为简便.

**例** 2 求可逆的 $n$ 级 $\lambda$-矩阵 $A(\lambda)$ 的相抵标准形.

**解** 先计算 $A(\lambda)$ 的各级行列式因子. 因为 $A(\lambda)$ 可逆，所以 $|A(\lambda)|$ 是 $K$ 中非零数，于是 $D_n(\lambda)=1$. 由(4)式知道，$D_i(\lambda)=1, i=1,2,\cdots,n$. 从而

$$d_i(\lambda)=1,\quad i=1,2,\cdots,n$$

因此 $A(\lambda)$ 的相抵标准形是单位矩阵 $I$.

由单位矩阵 $I$ 经过一次 $\lambda$-矩阵的初等变换得到的矩阵称为初等 $\lambda$-矩阵 容易看出初等 $\lambda$-矩阵都是可逆的.

由例 2 的结论可得到

**定理** 4 $n$ 级 $\lambda$-矩阵 $A(\lambda)$ 可逆的充分必要条件是它可以表成一系列初等 $\lambda$-矩阵的乘积.

**证明** 设 $A(\lambda)$ 可逆，则 $A(\lambda)$ 相抵于 $I$. 从而有一系列初等 $\lambda$-矩阵 $P_1(\lambda),\cdots,P_s(\lambda),Q_1(\lambda),\cdots,Q_t(\lambda)$，使得

$$\begin{aligned}A(\lambda)&=P_s(\lambda)\cdots P_1(\lambda)IQ_1(\lambda)\cdots Q_t(\lambda)\\&=P_s(\lambda)\cdots P_1(\lambda)Q_1(\lambda)\cdots Q_t(\lambda)\end{aligned}$$

由于初等 $\lambda$-矩阵都是可逆的，因此充分性是显然的. ▌

**推论** 1 两个 $n$ 级 $\lambda$-矩阵 $A(\lambda)$ 与 $B(\lambda)$ 相抵的充分必要条件是存在可逆的 $\lambda$-矩阵 $P(\lambda)$ 和 $Q(\lambda)$，使得

$$B(\lambda)=P(\lambda)A(\lambda)Q(\lambda)$$

**定义** 3 设 $A$ 是数域 $K$ 上的 $n$ 级矩阵，$\lambda$-矩阵 $\lambda I-A$ 称为 $A$ 的**特征矩阵**.

由于 $|\lambda I-A|$ 是 $A$ 的特征多项式，它是 $n$ 次多项式，因此不等于零. 从而 $\lambda I-A$ 的秩等于 $n$，并且 $\lambda I-A$ 的 $n$ 级行列式因子 $D_n(\lambda)$ 等于 $A$ 的特征多项式. 于是由(3)式和(2)式知，$\lambda I-A$ 的不变因子有 $n$ 个，并且它们的乘积等于 $A$ 的特征多项式.

注意：可逆的 $\lambda$-矩阵一定满秩，但是满秩的 $\lambda$-矩阵不一定可逆. 例如，$A$ 的特征矩阵 $\lambda I-A$ 是满秩的，但 $\lambda I-A$ 不可逆，这是因为 $|\lambda I-A|$ 不是 $K[\lambda]$ 中的可逆元($|\lambda I-A|$ 是 $n$ 次多项式，$n\geqslant 1$).

## §4 不可约多项式·唯一因式分解定理

上一节我们证明了数域 $K$ 上的一元多项式环 $K[x]$ 中每一对

多项式都有最大公因式.这一节我们要利用这个性质证明 $K[x]$ 中多项式的因式分解的唯一性.

读者在中学代数里已经做过一些多项式的因式分解的练习题,知道多项式的因式分解是跟所考虑的数集有关系的,譬如多项式 $f(x)=x^4-9$,在有理数集 $Q$ 上可分解成

$$x^4-9=(x^2-3)(x^2+3)$$

在实数集 $R$ 上可分解成

$$x^4-9=(x+\sqrt{3})(x-\sqrt{3})(x^2+3)$$

而在复数集 $C$ 上可进一步分解成

$$x^4-9=(x+\sqrt{3})(x-\sqrt{3})(x+\sqrt{3}\mathrm{i})(x-\sqrt{3}\mathrm{i})$$

由此看出,在讨论多项式的因式分解时,首先要明确系数域 $K$,在这前提下,我们要问:$K[x]$ 中一个多项式 $f(x)$ 能分解成什么样的多项式的乘积,使得其中每一个多项式都不能进一步分解?容易想到这种不能分解的多项式应当含有最少的因式,为此我们给出不可约多项式的概念:

**定义 1** $K[x]$ 中一个次数大于零的多项式 $f(x)$ 如果在 $K[x]$ 中的因式只有 $K$ 中的非零数以及 $f(x)$ 的相伴元,则称 $f(x)$ 是域 $K$ 上的一个**不可约多项式**;否则叫做**可约多项式**.

从定义 1 立即得出,$K[x]$ 中一个不可约多项式 $p(x)$ 与任一多项式 $f(x)$ 之间的关系只有两种可能:或者 $(p(x),f(x))=1$,或者 $p(x)\mid f(x)$.这是因为 $(p(x),f(x))$ 是 $p(x)$ 的因式,所以 $(p(x),f(x))$ 或者等于 1,或者等于 $p(x)$ 的相伴元.在后一情形有 $p(x)\mid(p(x),f(x))$,从而 $p(x)\mid f(x)$.

由上一段的结论可以推出不可约多项式的一条重要性质:

**定理** 7.4.1 如果 $p(x)$ 是数域 $K$ 上的一个不可约多项式,那么对于 $K[x]$ 中任意两个多项式 $f(x),g(x)$,从 $p(x)\mid f(x)g(x)$ 可以推出 $p(x)\mid f(x)$,或者 $p(x)\mid g(x)$.

**证明** 如果 $p(x)\mid f(x)$,则结论已经成立.

如果 $p(x)$ 不能整除 $f(x)$，则据上一段的结论知，$(p(x),f(x))=1$. 又由于 $p(x)\mid f(x)g(x)$，所以 $p(x)\mid g(x)$. ▌

利用数学归纳法，定理 7.4.1 可以推广为：在 $K[x]$ 中，如果不可约多项式 $p(x)$ 整除一些多项式 $f_1(x),f_2(x),\cdots,f_s(x)$ 的乘积 $f_1(x)f_2(x)\cdots f_s(x)$，则 $p(x)$ 一定整除这些多项式之中的一个.

据定义 1，数域 $K$ 上一个不可约多项式 $p(x)$ 的因式的次数只可能是零或者等于 $\deg p(x)$，因此域 $K$ 上不可约多项式 $p(x)$ 不能表成 $K[x]$ 中两个次数都比 $p(x)$ 低的多项式的乘积. 反之，如果数域 $K$ 上一个次数大于零的多项式 $f(x)$ 不能表成 $K[x]$ 中两个次数都比 $f(x)$ 低的多项式的乘积，那么 $f(x)$ 一定是域 $K$ 上的不可约多项式. 证明如下：

任取 $f(x)$ 的一个因式 $g(x)$，则有 $h(x)\in K[x]$，使得

$$f(x)=g(x)h(x)$$

据假设，$\deg g(x)$ 与 $\deg h(x)$ 中至少有一个与 $\deg f(x)$ 相等. 如果 $\deg g(x)=\deg f(x)$，则 $\deg h(x)=0$，从而 $g(x)$ 与 $f(x)$ 相伴；如果 $\deg h(x)=\deg f(x)$，则 $\deg g(x)=0$，即 $g(x)$ 是 $K$ 中非零数. 这证明了 $f(x)$ 是不可约多项式.

从上一段知道，数域 $K$ 上的不可约多项式就是在 $K[x]$ 中不能分解的多项式(指不能分解成两个次数都比它低的多项式的乘积). 于是从开头的例子看出，一个多项式的不可约性会随着系数域的扩大而改变.

从多项式不可约的充分必要条件容易看出，对于任意一个数域 $K$，$K[x]$ 中的一次多项式一定是不可约的.

现在我们来证明一个非常重要的结果：

**定理 7.4.2(唯一因式分解定理)** $K[x]$ 中每一个次数大于零的多项式 $f(x)$ 都能唯一地分解成数域 $K$ 上有限个不可约多项式的乘积，所谓唯一性是说，如果 $f(x)$ 有两个这样的分解式

$$f(x)=p_1(x)p_2(x)\cdots p_s(x)=q_1(x)q_2(x)\cdots q_t(x)$$

则一定有 $s=t$，并且适当排列因式的次序后有

$$p_i(x) \sim q_i(x), \quad i = 1, 2, \cdots, s$$

**证明** 先证分解式的存在性. 对多项式的次数 $n$ 作数学归纳法.

因为一次多项式都是不可约的，所以 $n = 1$ 时结论成立.

假设对于次数小于 $n$ 的多项式结论成立，现在来看 $n$ 次多项式 $f(x)$.

如果 $f(x)$ 是不可约多项式，则结论显然成立. 如果 $f(x)$ 是可约多项式，则有

$$f(x) = f_1(x) f_2(x)$$

其中 $f_i(x) \in K[x]$，并且 $\deg f_i(x) < \deg f(x)$，$i = 1, 2$. 由归纳假设，$f_1(x)$ 与 $f_2(x)$ 都可以分解成数域 $K$ 上有限个不可约多项式的乘积，把 $f_1(x)$ 与 $f_2(x)$ 的分解式合起来就得到 $f(x)$ 的一个分解式.

由归纳法原理，结论普遍成立.

现在证唯一性. 假设 $f(x)$ 有两个这样的分解式

$$f(x) = p_1(x) p_2(x) \cdots p_s(x) = q_1(x) q_2(x) \cdots q_t(x) \tag{1}$$

其中 $p_i(x), q_j(x), i = 1, \cdots, s, j = 1, \cdots, t$ 都是数域 $K$ 上的不可约多项式.

我们对分解式中不可约因式个数 $s$ 作归纳法.

当 $s = 1$ 时，则 $f(x)$ 是不可约多项式. 由于不可约多项式的因式只有它的相伴元及 $K$ 中非零数，所以 $t = 1$，并且

$$q_1(x) \sim f(x) = p_1(x)$$

假设分解式中不可约因式的个数为 $s - 1$ 时唯一性成立. 现在来看不可约因式的个数为 $s$ 的情形.

由(1)式，$p_1(x) \mid q_1(x) q_2(x) \cdots q_t(x)$，因此 $p_1(x)$ 必能整除其中的一个，不妨设

$$p_1(x) \mid q_1(x)$$

因为 $q_1(x)$ 也是不可约多项式，所以 $p_1(x) \sim q_1(x)$，即

$$p_1(x) = c_1 q_1(x), \quad c_1 \in K, \quad c_1 \neq 0 \tag{2}$$

把(2)代入(1)式，并且两边消去 $q_1(x)$，得

$$p_2(x)\cdots p_s(x)=(c_1^{-1}q_2(x))\cdots q_t(x) \tag{3}$$

由归纳假设，有

$$s-1=t-1,\quad 即\ s=t \tag{4}$$

并且适当排列因式的次序之后有

$$p_2(x)\sim c_1^{-1}q_2(x),p_3(x)\sim q_3(x),\cdots,p_s(x)\sim q_s(x) \tag{5}$$

从而有

$$p_i(x)\sim q_i(x),\quad i=1,2,\cdots,s \tag{6}$$

根据数学归纳法原理，唯一性得证. ∎

从唯一性的证明中可看出，$f(x)$ 的任一不可约因式一定与 $f(x)$ 的分解式中的某一个不可约因式相伴. 因此 $f(x)$ 的分解式给出了它的全部不可约因式(在相伴意义下).

$K[x]$ 中的唯一因式分解定理在理论上非常重要，但是至今仍没有一个统一的方法来做因式分解，也就是没有统一的方法求出一个次数大于零的多项式的所有不可约因式.

在多项式 $f(x)$ 的分解式中，可以把每一个不可约因式的首项系数提出来，使它们成为首项系数为 1 的多项式，再把相同的不可约因式的乘积写成乘幂的形式，于是 $f(x)$ 的分解式成为

$$f(x)=cp_1^{r_1}(x)p_2^{r_2}(x)\cdots p_m^{r_m}(x) \tag{7}$$

其中 $c$ 是 $f(x)$ 的首项系数，$p_1(x),p_2(x),\cdots,p_m(x)$ 是不同的首项系数为 1 的不可约多项式，$r_1,r_2,\cdots,r_m$ 是正整数. 这种分解式(7)称为 $f(x)$ 的**标准分解式**.

从理论研究的角度，如果已知两个多项式 $f(x)$ 与 $g(x)$ 的标准分解式，那么 $(f(x),g(x))$ 等于那些同时在 $f(x)$ 与 $g(x)$ 的标准分解式中出现的不可约多项式的方幂的乘积，所带的方幂的指数等于它在 $f(x)$ 与 $g(x)$ 中所带的方幂中的较小的一个. 即设

$$f(x)=ap_1^{k_1}(x)\cdots p_l^{k_l}(x)p_{l+1}^{k_{l+1}}(x)\cdots p_m^{k_m}(x) \tag{8}$$

$$g(x)=bp_1^{l_1}(x)\cdots p_l^{l_l}(x)q_{l+1}^{l_{l+1}}(x)\cdots q_s^{l_s}(x) \tag{9}$$

则 $$(f(x),g(x)) = p_1^{\min\{k_1,l_1\}}(x)\cdots p_l^{\min\{k_l,l_l\}}(x) \tag{10}$$

理由如下：显然(10)式右端的多项式的乘积 $d(x)$ 是 $f(x)$ 与 $g(x)$ 的一个公因式. 任取 $f(x)$ 与 $g(x)$ 的一个次数大于零的公因式 $c(x)$，在 $c(x)$ 的标准分解式中任取一个不可约因式 $u(x)$，由于 $u(x)|f(x)$ 且 $u(x)|g(x)$，所以 $u(x)$ 等于某个 $p_i(x)$，其中 $1 \leqslant i \leqslant l$；并且 $u(x)$ 在 $c(x)$ 的标准分解式中的幂指数不超过 $\min\{k_i,t_i\}$. 于是得出，$c(x)|d(x)$. 这证明了 $d(x)$ 是 $f(x)$ 与 $g(x)$ 的最大公因式.

由于把多项式分解成不可约因式的乘积没有统一的方法，因此上述求最大公因式的方法不能代替辗转相除法.

我们举一个简单例子来看一下唯一因式分解定理的应用.

**例1** 证明 $x^2-2$ 在有理数域上不可约.

**证明** 若 $x^2-2$ 在有理数域 $Q$ 上可约，则它的标准分解式为

$$x^2-2=(x+a)(x+b),\quad a,b\in Q \tag{11}$$

(11)式也可以看成 $x^2-2$ 在实数域 $R$ 上的一个不可约因式分解. 另一方面我们知道 $x^2-2$ 在 $R$ 上有如下的不可约因式分解

$$x^2-2=(x+\sqrt{2})(x-\sqrt{2}) \tag{12}$$

由 $R[x]$ 中唯一因式分解定理得

$$x+a=x+\sqrt{2} \quad 或 \quad x+a=x-\sqrt{2}$$

由此推出 $a=\sqrt{2}$ 或 $a=-\sqrt{2}$，这与 $a\in Q$ 矛盾. 因此 $x^2+2$ 在 $Q$ 上不可约.

请读者给出 $x^2+2$ 在有理数域上不可约的另一种证法.（提示：仍用反证法，从(11)式根据多项式相等的定义去推出矛盾.）

## 阅读材料六

整数环 $Z$ 与数域 $K$ 上的一元多项式环 $K[x]$ 一样有唯一因子分解定理.

在 $Z$ 中与 $K[x]$ 中不可约多项式类似的概念是素数.

**定义 1** 一个大于 1 的整数，如果它的正因数只有 1 和它本身，就叫做**素数**，否则就叫做**合数**.

从定义 1 立即得出，一个素数 $p$ 与任一整数 $a$ 之间的关系只有两种可能：或者 $(p,a)=1$，或者 $p|a$.

由上述关系，立即得出

**定理 1** 如果 $p$ 是素数，那么对于任意整数 $a,b$，从 $p|ab$ 可以推出 $p|a$ 或者 $p|b$.

定理 1 可以推广为：如果素数 $p$ 能整除一些整数 $a_1,a_2,\cdots,a_s$ 的乘积 $a_1a_2\cdots a_s$，则 $p$ 一定整除其中的一个.

**定理 2（整数的唯一因子分解定理）** 任一大于 1 的整数 $a$ 能唯一地分解成有限个素数的乘积. 所谓唯一性是说，如果 $a$ 有两个这样的分解式

$$a=p_1p_2\cdots p_s=q_1q_2\cdots q_t$$

则一定有 $s=t$，并且适当排列因子的次序后有

$$p_i=q_i,\quad i=1,\cdots,s$$

**证明** 先证分解式的存在性，用数学归纳法.

当 $a=2$ 时，2 是素数，因此结论显然成立.

假设对小于 $a$ 且大于 1 的一切整数，结论成立，考虑 $a(a>1)$. 如果 $a$ 是素数，则结论成立. 如果 $a$ 是合数，则有两个大于 1 且小于 $a$ 的整数 $b,c$，使得 $a=bc$.

由归纳假设，$b$ 与 $c$ 分别能分解成有限个素数乘积，从而 $a$ 能分解成有限个素数的乘积.

由归纳法原理，存在性得证.

唯一性证明与定理 7.4.2 的唯一性证明类似，请读者完成. ▌

定理 2 称为**算术基本定理**，因为它在整数的整除性理论中起着基本重要的作用. 算术基本定理告诉我们，任一大于 1 的整数 $a$ 能唯一地写成（除了因子的排列次序外）

$$a=p_1^{r_1}p_2^{r_2}\cdots p_m^{r_m}\tag{1}$$

其中 $p_1,p_2,\cdots,p_m$ 是不同的素数，$r_1,\cdots,r_m$ 是正整数，(1) 式称为 $a$ 的标准分解式.

如果已知两个大于 1 的整数 $a,b$ 的标准分解式

$$a=p_1^{k_1}\cdots p_s^{k_s}p_{s+1}^{k_{s+1}}\cdots p_m^{k_m}\tag{2}$$

$$b = p_1^{l_1} \cdots p_s^{l_s} q_{s+1}^{l_{s+1}} \cdots q_n^{l_n} \tag{3}$$

则

$$(a,b) = p_1^{r_1} \cdots p_s^{r_s}$$

$$r_i = \min\{k_i, l_i\}, \quad i = 1, \cdots, s \tag{4}$$

$$[a,b] = p_1^{\delta_1} \cdots p_s^{\delta_s} p_{s+1}^{k_{s+1}} \cdots p_m^{k_m} q_{s+1}^{l_{s+1}} \cdots q_n^{l_n}$$

$$\delta_i = \max\{k_i, l_i\}, \quad i = 1, \cdots, s \tag{5}$$

(4)式与(5)式的证明留给读者作为练习.

## 习 题 7.4

1. 证明多项式 $x^2 + 1$ 在有理数域上不可约.

2. 分别在复数域,实数域和有理数域上分解下列多项式为不可约因式的乘积.

(1) $x^4 + 1$;　　(2) $x^4 + 4$.

3. 证明:在 $K[x]$ 中,$g^2(x) | f^2(x)$ 当且仅当 $g(x) | f(x)$.

4. 证明定理7.4.1的逆命题为真,即设 $p(x)$ 是 $K[x]$ 中一个次数大于零的多项式,如果对于任意 $f(x), g(x) \in K[x]$,从 $p(x) | f(x)g(x)$ 可以推出 $p(x) | f(x)$ 或者 $p(x) | g(x)$,那么 $p(x)$ 是不可约多项式.

*5. 证明:数域 $K$ 上一个次数大于零的多项式 $f(x)$ 是 $K[x]$ 中某一不可约多项式的方幂的充分必要条件是对于任意 $g(x) \in K[x]$,必有 $(f(x), g(x)) = 1$,或者存在一个正整数 $m$,使得 $f(x) | g^m(x)$.

*6. 证明:数域 $K$ 上一个次数大于零的多项式 $f(x)$ 是 $K[x]$ 中某一不可约多项式的方幂的充分必要条件是对于任意 $g(x), h(x) \in K[x]$,从 $f(x) | g(x)h(x)$ 可以推出 $f(x) | g(x)$,或者存在一个正整数 $m$,使得 $f(x) | h^m(x)$.

7. 在 $K[x]$ 中,设 $(f, g_i) = 1, i = 1, 2$,证明:

$$(fg_1, g_2) = (g_1, g_2)$$

*8. 证明本节阅读材料六中定理1的逆命题为真,即设 $p$ 是大于1的整数,如果对于任意整数 $a, b$,,从 $p | ab$ 可以推出 $p | a$ 或者 $p | b$,那么 $p$ 是素数.

## §5 重 因 式

上一节我们已证明 $K[x]$ 中每一个次数大于零的多项式 $f(x)$ 能唯一地分解成数域 $K$ 上有限个不可约多项式的乘积. 如果 $f(x)$ 的分解式中每一个不可约因式只出现 1 次, 这种情形是特别重要的情形. 这一节我们要给出识别这种情形的一个统一的方法.

**定义 1** 在 $K[x]$ 中, 不可约多项式 $p(x)$ 称为多项式 $f(x)$ 的 $k$ **重因式**, 如果 $p^k(x)\mid f(x)$, 而 $p^{k+1}(x)\nmid f(x)$(这个符号表示 $p^{k+1}(x)$ 不能整除 $f(x)$).

如果 $k=0$, 则 $p(x)\nmid f(x)$, 因此 $p(x)$ 不是 $f(x)$ 的因式; 如果 $k=1$, 则 $p(x)$ 称为 $f(x)$ 的**单因式**; 如果 $k>1$, 则 $p(x)$ 称为 $f(x)$ 的**重因式**.

显然, 如果 $f(x)$ 的标准分解式为

$$f(x)=cp_1^{r_1}(x)p_2^{r_2}(x)\cdots p_m^{r_m}(x)$$

则 $p_1(x),p_2(x),\cdots,p_m(x)$ 分别是 $f(x)$ 的 $r_1$ 重, $r_2$ 重, $\cdots$, $r_m$ 重因式. 指数 $r_i=1$ 的那些不可约因式是单因式; 指数 $r_i>1$ 的那些不可约因式是重因式. 所以 $f(x)$ 的分解式中每一个不可约因式只出现 1 次的情形也就是 $f(x)$ 没有重因式的情形. 如何判别一个多项式有没有重因式呢? 由于没有一般的方法来求一个多项式的标准分解式, 因此我们必须寻找别的方法来判断一个多项式有没有重因式.

我们先看一个简单例子, 以便从中受到启发.

设 $f(x)=(x+1)^3\in R[x]$, 这时 $f(x)$ 有重因式. 如果我们把 $f(x)$ 看成数学分析中讨论的多项式函数, 那么对 $f(x)$ 可以求导数, 得 $f'(x)=3(x+1)^2$. 于是 $(f(x),f'(x))=(x+1)^2$. 从这个例子受到启发, 有可能运用导数概念以及最大公因式的求法来讨论一个多项式有没有重因式的问题. 由于我们现在讲的多项式

是任意数域 $K$ 上一个不定元的多项式，而数学分析中的多项式函数是实变量 $x$ 的函数，其导数概念涉及极限概念，因此我们不能直接引用数学分析中多项式函数的导数概念，我们必须给任意数域 $K$ 上一元多项式的导数下个定义，当然这个定义是从分析中多项式函数的导数公式受到启发的.

**定义** 2　对于 $K[x]$ 中的多项式

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

我们把 $K[x]$ 中的多项式

$$na_n x^{n-1} + (n-1)a_{n-1}x^{n-2} + \cdots + a_1$$

叫做 $f(x)$ 的**导数**(**或一阶导数**)，记作 $f'(x)$.

$f'(x)$ 的导数叫做 $f(x)$ 的**二阶导数**，记作 $f''(x)$；$f''(x)$ 的导数叫做 $f(x)$ 的**三阶导数**，记作 $f'''(x)$；等等. $f(x)$ 的 $k$ 阶导数也记作 $f^{(k)}(x)$.

从定义 2 立即得出，一个 $n$ 次多项式的导数是一个 $n-1$ 次多项式；它的 $n$ 阶导数是 $K$ 中一个非零数；它的 $n+1$ 阶导数等于零. 零多项式的导数是零多项式.

根据定义 2，通过直接验证，可以得出 $K[x]$ 中多项式的导数的基本公式：

$$[f(x)+g(x)]' = f'(x) + g'(x)$$

$$[cf(x)]' = cf'(x), \quad c \in K$$

$$[f(x)g(x)]' = f'(x)g(x) + f(x)g'(x)$$

$$[f^m(x)]' = mf^{m-1}(x)f'(x)$$

仍看从前的简单例子. 不可约多项式 $x+1$ 是 $f(x) = (x+1)^3$ 的 3 重因式. 由于按定义 2 和上述公式可得出，$f'(x) = 3(x+1)^2$，因此 $x+1$ 是 $f'(x)$ 的 2 重因式. 这个例子的结论有一般性，即

**定理** 7.5.1　在 $K[x]$ 中，如果不可约多项式 $p(x)$ 是 $f(x)$ 的一个 $k(k \geqslant 1)$ 重因式，则 $p(x)$ 是 $f(x)$ 的导数 $f'(x)$ 的一个 $k-1$

重因式. 特别地,多项式 $f(x)$ 的单因式不是 $f(x)$ 的导数 $f'(x)$ 的因式.

**证明** 因为 $p(x)$ 是 $f(x)$ 的 $k$ 重因式,所以存在 $g(x)\in K[x]$,使

$$f(x)=p^k(x)g(x),\quad p(x)\nmid g(x)$$

求 $f(x)$ 的导数,得

$$f'(x)=p^{k-1}(x)[kp'(x)g(x)+p(x)g'(x)]$$

因为 $p(x)\nmid kp'(x)$, $p(x)\nmid g(x)$,并且 $p(x)$ 是不可约多项式,所以 $p(x)\nmid kp'(x)g(x)$. 但是 $p(x)\mid p(x)g'(x)$,所以 $p(x)$ 不能整除上述等式右端括号里的和. 因此 $p(x)$ 是 $f'(x)$ 的 $k-1$ 重因式. ▌

**推论** 7.5.1 在 $K[x]$ 中,不可约多项式 $p(x)$ 是 $f(x)$ 的重因式的充分必要条件为 $p(x)$ 是 $f(x)$ 与 $f'(x)$ 的公因式.

**证明** 必要性. 设不可约多项式 $p(x)$ 是 $f(x)$ 的 $k$ 重因式,其中 $k>1$,则据定理 7.5.1 得, $p(x)$ 是 $f'(x)$ 的 $k-1$ 重因式,从而 $p(x)$ 是 $f(x)$ 与 $f'(x)$ 的公因式.

充分性. 设不可约多项式 $p(x)$ 是 $f(x)$ 与 $f'(x)$ 的公因式. 据定理 7.5.1 知, $p(x)$ 不是 $f(x)$ 的单因式,所以 $p(x)$ 是 $f(x)$ 的重因式. ▌

从推论 7.5.1 立即得到

**定理** 7.5.2 $K[x]$ 中次数大于零的多项式 $f(x)$ 没有重因式的充分必要条件是 $f(x)$ 与它的导数 $f'(x)$ 互素.

定理 7.5.2 表明,判断一个多项式 $f(x)$ 有没有重因式,只要去计算 $(f(x),f'(x))$,而求最大公因式有统一的方法:辗转相除法. 所以我们有统一的方法 —— 辗转相除法判断一个多项式有没有重因式. 不仅如此,由于在数域扩大时,两个多项式的最大公因式不改变,一个多项式的导数也不改变,因此我们还有下述结论:

设数域 $F$ 包含数域 $K$,对于 $f(x)\in K[x]$, $f(x)$ 在 $K[x]$ 中没有重因式的充分必要条件是把 $f(x)$ 看成 $F[x]$ 中的多项式时, $f(x)$ 在 $F[x]$ 中没有重因式.

在一些问题中，如果多项式 $f(x)$ 有重因式，我们希望求出一个多项式 $g(x)$，它没有重因式，并且不计重数时，它与 $f(x)$ 含有完全相同的不可约因式. 从下面的讨论中可得到 $g(x)$ 的求法.

设 $K[x]$ 中的多项式 $f(x)$ 的标准分解式是

$$f(x)=cp_1^{r_1}(x)p_2^{r_2}(x)\cdots p_m^{r_m}(x)$$

据定理 7.5.1 得

$$f'(x)=p_1^{r_1-1}(x)p_2^{r_2-1}(x)\cdots p_m^{r_m-1}(x)h(x)$$

其中 $h(x)$ 不能被任何 $p_i(x)$ 整除，$i=1,\cdots,m$. 于是

$$(f(x),f'(x))=p_1^{r_1-1}(x)p_2^{r_2-1}(x)\cdots p_m^{r_m-1}(x)$$

因此用 $(f(x),f'(x))$ 除 $f(x)$ 所得商式是

$$cp_1(x)p_2(x)\cdots p_m(x)$$

把它记作 $g(x)$，我们便得到一个没有重因式的多项式 $g(x)$，它与 $f(x)$ 含有完全相同的不可约因式(不计重数).

上面一段给出了去掉 $f(x)$ 的不可约因式重数的方法：先用辗转相除法求出 $(f(x),f'(x))$，然后对 $f(x)$ 与 $(f(x),f'(x))$ 做带余除法，所得商式 $g(x)$ 即为所求的没有重因式的多项式.

去掉 $f(x)$ 的不可约因式的重数有不少好处. 例如，为了求 $f(x)$ 的所有不可约因式，我们可以先用上述方法得到一个没有重因式的多项式 $g(x)$，它与 $f(x)$ 含有完全相同的不可约因式(不计重数)，但由于 $g(x)$ 的次数小于 $f(x)$ 的次数，所以 $g(x)$ 的不可约因式可能比较容易求得. 如果我们求出了 $g(x)$ 的一个不可约因式 $p_i(x)$，那么用带余除法可求出 $p_i(x)$ 在 $f(x)$ 中的重数. 又如，在实际问题中常常要求出一个多项式 $f(x)$ 的根(关于多项式的根，我们在下节将详细讨论. 实际上，读者在中学代数里已知道多项式的根的概念)，由于有些求多项式的根的算法只对没有重因式的多项式适用，因此我们可以先去掉 $f(x)$ 的不可约因式的重数，得到一个没有重因式的多项式 $g(x)$，而 $g(x)$ 与 $f(x)$ 有完全相同的根(不计重数).

**例 1** 证明：$Q[x]$ 中的多项式

$$f(x)=1+x+\frac{x^2}{2!}+\cdots+\frac{x^n}{n!}$$

没有重因式.

**证明** 求 $f(x)$ 的导数得

$$f'(x)=1+x+\cdots+\frac{x^{n-1}}{(n-1)!}$$

于是 $f(x)=f'(x)+\frac{x^n}{n!}$. 据习题 7.3 的第 7 题的结果得

$$(f(x),f'(x))=\left(f'(x)+\frac{1}{n!}x^n,f'(x)\right)=\left(\frac{x^n}{n!},f'(x)\right)$$

由于 $\frac{x^n}{n!}$ 的不可约因式只有 $x$(不计重数)，而 $x\nmid f'(x)$，所以

$$\left(\frac{x^n}{n!},f'(x)\right)=1$$

从而 $(f(x),f'(x))=1$. 因此 $f(x)$ 没有重因式. ∎

## 习 题 7.5

1. 判别下列有理系数多项式有无重因式：

(1) $f(x)=x^5-x^4-3x^3-5x^2+8x+12$；

(2) $f(x)=x^4+2x^3-5x^2-22x-24$.

2. $a,b$ 应该满足什么条件，下述实系数多项式才能有重因式：

$$f(x)=x^3+2ax+b$$

3. 举例说明：$K[x]$ 中，一个不可约多项式 $p(x)$ 是 $f(x)$ 的导数 $f'(x)$ 的 $k-1$ 重因式($k\geqslant 2$)，但是 $p(x)$ 不是 $f(x)$ 的 $k$ 重因式.

4. 证明：$K[x]$ 中，若不可约多项式 $p(x)$ 是 $f(x)$ 的导数 $f'(x)$ 的 $k-1$ 重因式($k\geqslant 1$)，并且 $p(x)$ 是 $f(x)$ 的因式，则 $p(x)$ 是 $f(x)$ 的 $k$ 重因式.

5. 证明：$K[x]$ 中，不可约多项式 $p(x)$ 是 $f(x)$ 的 $k$ 重因式($k\geqslant 1$) 的充分必要条件为：$p(x)$ 是 $f(x),f'(x),\cdots,f^{(k-1)}(x)$ 的因式，但不是 $f^{(k)}(x)$ 的因式.

6. 在 $Q[x]$ 中求一个没有重因式的多项式 $g(x)$，使它与 $f(x)$ 含有完全相同的不可约因式(不计重数)；然后求 $f(x)$ 的标准分解式：

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

* 7. 证明：$K[x]$ 中一个 $n$ 次 $(n\geqslant 1)$ 多项式 $f(x)$ 能被它的导数整除的充分必要条件是它与一个一次因式的 $n$ 次幂相伴.

## §6 多项式的根·多项式函数·代数基本定理

从唯一因式分解定理知道，$K[x]$ 中每一个次数大于 0 的多项式都能唯一地分解成数域 $K$ 上有限个不可约多项式的乘积. 由此看出，不可约多项式起着基本建筑块的作用. 这促使我们去搞清楚 $K[x]$ 中的不可约多项式有哪些？为此我们需要多项式的根的概念，并研究多项式的根与它的因式之间的联系. 为了能运用函数论的知识来研究实(复)系数多项式的根，我们要弄清楚多项式函数与多项式之间的关系.

多项式的根的理论曾经是代数学研究的中心问题，虽然现在它不再统治代数学，但是它的重要性仍然是无可置疑的. 许多问题都归结为确定某一特定多项式的根，或者归结为了解根集合的某些信息. 例如在第五章中，讨论数域 $K$ 上一个 $n$ 级矩阵 $A$ 可不可以对角化时，就需要首先求出 $A$ 的特征多项式 $|\lambda I-A|$ 在 $K$ 中的根.

首先我们给出多项式的根的定义：

**定义 1** 设 $K$ 是一个数域，设 $R$ 是(或者可以看成是)$K$ 的一个交换扩环，$f(x)$ 是 $K[x]$ 中一个多项式，如果 $R$ 中一个元素 $c$ 使得 $f(c)=0$，则称 $c$ 是 $f(x)$ 在 $R$ 中的一个**根**(或**零点**).

例如，$Q[x]$ 中的多项式 $f(x)=x^2+1$ 在有理数域中没有根，在实数域中也没有根，但它在复数域中有两个根：i 与 $-$i.

我们常常把一个多项式 $f(x)$ 在 $Q$ 中的根(如果有的话)叫做

有理根，在实数域$R$中的根(如果有的话)叫做实根，在$C$中的根叫做复根.

又如，设实数域$R$上的一个2级矩阵

$$A=\begin{pmatrix}1 & -1\\ 1 & 1\end{pmatrix} \tag{1}$$

我们在本章§1的例10已指出，$R[A]$可以看成是$R$的一个交换扩环. 设$f(x)=x^2-2x+2\in R[x]$，直接计算得

$$f(A)=A^2-2A+2I=0$$

于是(1)式所给的2级矩阵$A$是$f(x)$在$R[A]$中的一个根.

特别重要的情形是讨论$K[x]$中的多项式在$K$中的根. 在这种情形，多项式$f(x)$的根与$f(x)$的一次因式有密切关系. 为了揭示这一关系，我们先用带余除法证明下述余数定理：

**定理** 7.6.1(**余数定理**) 在$K[x]$中，用一次多项式$x-a$去除多项式$f(x)$，所得的余式是$K$中一个数，这个数等于$f(a)$.

**证明** 据带余除法，用$x-a$去除$f(x)$，所得的余式的次数小于$x-a$的次数，因此余式为0或$K$中一个非零数. 把零多项式与$K$中数0等同，则余式可统一说成是$K$中一个数$r$. 设商为$h(x)$，则

$$f(x)=(x-a)h(x)+r$$

$x$用$a$代入，得

$$f(a)=r$$ ▌

利用余数定理我们可得到$K[x]$中的多项式$f(x)$在$K$中的根与$f(x)$在$K[x]$中的一次因式的关系：

**定理**7.6.2(Bezout **定理**) 设$f(x)\in K[x]$，$a\in K$是$f(x)$在$K$中的根的充分必要条件为$x-a$是$f(x)$在$K[x]$中的一次因式.

**证明** 设$f(x)\in K[x]$，$a\in K$. 据余数定理得

$$f(x)=(x-a)h(x)+f(a)$$

于是

$$f(a)=0\Leftrightarrow(x-a)\mid f(x)$$ ▌

利用根与一次因式的关系，对于 $K[x]$ 中的多项式在 $K$ 中的根，我们可以定义重根的概念：$a \in K$ 称为 $f(x) \in K[x]$ 的一个 $k$ **重根**，如果 $(x-a)$ 是 $f(x)$ 的 $k$ 重因式. 当 $k=1$ 时，$a$ 称为**单根**；当 $k>1$ 时，$a$ 称为**重根**.

利用根与一次因式的关系，我们可以得到 $K[x]$ 中的多项式 $f(x)$ 在 $K$ 中的根的数目的一个上界：

**定理 7.6.3** $K[x]$ 中的 $n$ 次 $(n \geqslant 0)$ 多项式在 $K$ 中至多有 $n$ 个根（重根按重数计算）.

**证明** 零次多项式没有根，因此结论成立.

设 $f(x)$ 是一个 $n$ 次多项式，$n>0$. 把 $f(x)$ 分解成不可约多项式的乘积. 据定理 7.6.2 以及根的重数的定义容易看出，$f(x)$ 在 $K$ 中的根的数目（重根按重数计算）等于分解式中一次因式的数目（重因式按重数计算），这个数目当然不超过 $f(x)$ 的次数 $n$. ▌

从定理 7.6.3 可以得到一个重要推论：

**推论 7.6.1** 设 $K[x]$ 中两个多项式 $f(x)$ 与 $g(x)$ 的次数都不超过 $n$. 如果 $x$ 分别用 $K$ 中 $n+1$ 个不同元素 $a_1, a_2, \cdots, a_{n+1}$ 代入，有 $f(a_i)=g(a_i), i=1,2,\cdots,n+1$，则 $f(x)=g(x)$.

**证明** 设 $h(x)=f(x)-g(x)$，则

$$\deg h \leqslant \max\{\deg f, \deg g\} \leqslant n$$

因为

$$h(a_i)=f(a_i)-g(a_i)=0, \quad i=1,2,\cdots,n+1$$

所以 $h(x)$ 在 $K$ 中至少有 $n+1$ 个不同的根. 据定理 7.6.3 得，$h(x)=0$. 于是 $f(x)=g(x)$. ▌

我们将在以后几处指出推论 7.6.1 的用处.

为了研究多项式的实根和复根以及研究其他一些问题，我们需要多项式函数的概念，并且研究多项式函数与多项式之间的关系.

设 $f(x) \in K[x]$，对于 $K$ 中每一个元素 $a$，$x$ 用 $a$ 代入得，$f(a) \in K$. 于是 $K[x]$ 中一个多项式 $f(x)$ 确定了 $K$ 到 $K$ 的一个

映射(即 $K$ 上的一个函数),仍用 $f$ 表示,即

$$f: K \longrightarrow K$$
$$a \longmapsto f(a)$$

这种由 $K[x]$ 中的多项式确定的 $K$ 上的函数称为 $K$ 上的**多项式函数**.

当 $K$ 是实数域 $R$ 时,上述定义的多项式函数就是数学分析中所讨论的多项式函数.

我们已经知道,$K[x]$ 中的每一个多项式都确定一个 $K$ 上的多项式函数. 现在要问:$K[x]$ 中两个不相等的多项式 $f(x)$ 与 $g(x)$,它们确定的 $K$ 上的多项式函数 $f$ 与 $g$ 是否不相等?回答是肯定的,即我们有

**定理** 7.6.4　数域 $K$ 上的两个多项式 $f(x)$ 与 $g(x)$ 如果不相等,则它们确定的 $K$ 上的多项式函数 $f$ 与 $g$ 也不相等.

**证明**　证逆否命题. 设 $K$ 上的多项式函数 $f$ 与 $g$ 相等,则对一切 $a \in K$,有 $f(a) = g(a)$. 由于 $K$ 是数域,它有无穷多个元素,于是据推论 7.6.1 得,$f(x) = g(x)$. 即多项式 $f(x)$ 与 $g(x)$ 相等. ▌

注意,定理 7.6.4 证明中的关键是 $K$ 有无穷多个元素.

我们把数域 $K$ 上的多项式函数组成的集合记作 $K_{pol}$. 让 $K[x]$ 中的多项式 $f(x)$ 对应到它确定的 $K$ 上的函数 $f$,这是 $K[x]$ 到 $K_{pol}$ 的一个映射,记作 $\sigma$. 显然它是满射. 定理 7.6.4 表明这个映射是单射,因此它是双射.

在集合 $K_{pol}$ 中规定加法与乘法如下:设 $f,g \in K_{pol}$,规定

$$(f+g)(a) := f(a) + g(a), \quad \forall\, a \in K$$
$$(fg)(a) := f(a)g(a), \quad \forall\, a \in K$$

容易验证,$f+g, fg \in K_{pol}$(设 $f,g$ 分别是多项式 $f(x),g(x)$ 确定的函数,记

$$h(x) = f(x) + g(x)$$

则对任意 $a \in K$,有

$$h(a) = f(a) + g(a) = (f+g)(a)$$

因此 $h=f+g$ .这表明 $f+g$ 是多项式 $f(x)+g(x)$ 确定的函数.类似可证 $fg$ 是多项式 $f(x)g(x)$ 确定的函数),并且 $K_{pol}$ 是一个环,称它为 $K$ **上的多项式函数环**.它的零元素是零多项式确定的函数 0,称为零函数;它的单位元素是多项式 1 确定的函数 1,它是一个常值函数,它把 $K$ 中每个元素 $a$ 映成 1.容易看出 $K_{pol}$ 是一个交换环.

环 $K[x]$ 与环 $K_{pol}$ 有密切关系:

**定理** 7.6.5 设 $K$ 是数域,让 $K[x]$ 中的多项式 $f(x)$ 对应到它确定的 $K$ 上的函数 $f$,则这个映射 $\sigma$ 是环 $K[x]$ 到环 $K_{pol}$ 的双射,并且 $\sigma$ 保持加法和乘法运算.

**证明** 前面已证 $\sigma$ 是双射.现在证 $\sigma$ 保持运算.

设 $f(x),g(x)\in K[x]$,前已证 $f(x)+g(x)$ 确定的函数是 $f+g$,$f(x)g(x)$ 确定的函数是 $fg$.因此

$$\sigma(f(x)+g(x))=f+g=\sigma(f(x))+\sigma(g(x))$$
$$\sigma(f(x)g(x))=fg=\sigma(f(x))\sigma(g(x))$$

这证明了 $\sigma$ 保持加法,并且保持乘法. ▌

*定理 7.6.5 表明,数域 $K$ 上的一元多项式环 $K[x]$ 与数域 $K$ 上的多项式函数环 $K_{pol}$ 是同构的.同构的环,它们的元素之间有一个一一对应,并且这个对应保持加法与乘法运算.因此在研究有关加法与乘法运算的问题时,同构的环具有相同的性质,从而可以把同构的环等同看待.*

* 我们举一个例子:我们已经知道,数域 $K$ 上的一元多项式环 $K[x]$ 是无零因子环.由于零因子的概念只用到乘法运算以及零元素的概念,而零元素是用加法来定义的,因此一个环有没有零因子是关于加法与乘法运算的性质.由于 $K[x]$ 与 $K_{pol}$ 是同构的环,因此 $K_{pol}$ 也是无零因子环.当然我们也可以直接证明 $K_{pol}$ 是无零因子环:设 $f,g\in K_{pol}$,并且 $f\neq 0,g\neq 0$.于是多项式 $f(x)$ 与 $g(x)$ 不是零多项式.由于 $K[x]$ 是无零因子环,因此 $f(x)g(x)\neq 0$.再据推论 7.6.1 得,$f(x)g(x)$ 确定的函数不等于

零函数，据定理 7.6.5，多项式 $f(x)g(x)$ 确定的函数是 $fg$，因此 $fg \neq 0$. 这证明了 $K_{\mathrm{pol}}$ 是无零因子环.

现在我们可以用多项式函数的概念来刻画 $K[x]$ 中的多项式在 $K$ 中的根：$a \in K$ 是 $K[x]$ 中多项式 $f(x)$ 在 $K$ 中的根当且仅当 $K$ 上的多项式函数 $f$ 在 $a$ 处的函数值 $f(a)=0$. 于是我们可以运用函数论的知识来研究 $K[x]$ 中的多项式在 $K$ 中的根，特别是当 $K$ 为实数域 $R$ 或复数域 $C$ 的情形. 譬如，实系数多项式 $f(x)$ 的实根就是多项式函数 $y=f(x)$ 的图象与 $x$ 轴的交点的横坐标. 据数学分析的知识，多项式函数是连续函数，而连续函数具有这样的性质：在区间 $[a,b]$ 上，它取 $f(a)$ 与 $f(b)$ 之间的每一个值. 因此如果 $f(a)<0$，而 $f(b)>0$，则必有 $c \in (a,b)$，使得 $f(c)=0$，也就是说，在区间 $(a,b)$ 上一定有多项式 $f(x)$ 的一个根.

运用复变函数论的知识可以证明下面重要的定理.

**定理** 7.6.6（**代数基本定理**） 每个次数大于零的复系数多项式在复数域中至少有一个根.

证明代数基本定理所用到的复变函数论的知识主要是下面的定理.

* **定理** 7.6.7（Liouville **定理**） 若函数 $f(z)$ 在复平面 $C$ 上解析，并且有界，则 $f(z)$ 必为一常值函数.

* Liouville 定理的证明可以在复变函数的教材中找到. 这里我们只指出，函数 $f(z)$ 在复平面 $C$ 上解析的意思是：$f(z)$ 在复平面 $C$ 上的每一点的导数都存在. 复变量函数的导数定义以及求导公式都与数学分析中实变量函数一样. 例如，若函数 $f(z)$，$g(z)$ 在复平面 $C$ 上解析，并且对于 $C$ 上每一点 $z$，有 $g(z) \neq 0$，则 $f(z)g(z)$ 在复平面 $C$ 上解析，并且

$$\left(\frac{f(z)}{g(z)}\right)' = \frac{f'(z)g(z)-f(z)g'(z)}{g^2(z)}$$

* **代数基本定理的证明** 假如一个 $n$ 次 $(n>0)$ 复系数多项式 $f(x)$ 在复数域中没有根，则对于复平面 $C$ 上每一点 $z$，有

$f(z) \neq 0$. 由于多项式函数在复平面 $C$ 上是解析的，于是函数

$$\varphi(z) = \frac{1}{f(z)}$$

在复平面上解析. 由于 $z \to \infty$ 时，$\varphi(z) \to 0$，所以存在正实数 $R$ 和 $M_1$，使得当 $|z| > R$ 时，$|\varphi(z)| \leqslant M_1$. 而当 $|z| \leqslant R$ 时，连续函数 $\varphi(z)$ 在 $|z| \leqslant R$ 上有界 $M_2$. 因此对一切 $z \in C$，有

$$|\varphi(z)| \leqslant \max\{M_1, M_2\}$$

据 Liouville 定理得，$\varphi(z)$ 必为一常值函数，从而 $f(z)$ 为一个常值函数，这与多项式 $f(x)$ 的次数大于零矛盾. ▌

定理 7.6.6 被称为"代数基本定理"是因为在 19 世纪以前解代数方程是代数学的最重要课题. 在现代，定理 7.6.6 被看成代数学的基本定理之一(但不是最基本的). 这个定理的第一个严格证明是高斯(Gauss)在 1799 年给出的，后来他又给出了四个证明. Jordan，Weyl 等人也给过证明.

利用根与一次因式的关系(即 Bezout 定理)，代数基本定理可以等价地叙述为：

**定理** 7.6.8　每一个次数大于零的复系数多项式，在复数域上一定有一个一次因式.

由此可知，复数域上所有次数大于 1 的多项式全是可约的. 换句话说，复数域上的不可约多项式只有一次多项式. 于是唯一因式分解定理在复数域上可以叙述成：

**定理** 7.6.9(**复系数多项式唯一因式分解定理**)　每个次数大于零的复系数多项式在复数域上都可以唯一地分解成一次因式的乘积. ▌

因此，复系数多项式 $f(x)$ 具有标准分解式

$$f(x) = a(x - c_1)^{r_1}(x - c_2)^{r_2}\cdots(x - c_m)^{r_m} \tag{2}$$

其中 $a$ 是 $f(x)$ 的首项系数，$c_1, c_2, \cdots, c_m$ 是不同的复数，$r_1, r_2, \cdots, r_m$ 是正整数，从标准分解式(2) 得出

**定理** 7.6.10　每个 $n$ 次($n \geqslant 0$) 复系数多项式恰有 $n$ 个复根

(重根按重数计算). ∎

* 我们先举一个例子说明复系数多项式唯一因式分解定理的应用：

* **例** 1　设 $f(x)\in C[x]$，对于 $a\in C$，$f^{-1}(a)$ 表示 $a$ 在映射 $f$ 下的原象(见第四章 §2 的定义 6)．设 $f(x),g(x)\in C[x]$，证明：如果 $f^{-1}(0)=g^{-1}(0)$，并且 $f^{-1}(1)=g^{-1}(1)$，则 $f(x)=g(x)$.

**证明**　设 $\max\{\deg f(x),\deg g(x)\}=n$，不妨设

$$\deg f(x)=n$$

显然 $f^{-1}(0)\cap f^{-1}(1)=\varnothing$．如果能证明

$$|f^{-1}(0)\cup f^{-1}(1)|\geqslant n+1$$

则据推论 7.6.1 得，$f(x)=g(x)$.

设多项式 $f(x)$，$f(x)-1$ 的标准分解式分别是

$$f(x)=a\prod_{i=1}^{m}(x-c_i)^{r_i},\quad c_i\in C$$

$$f(x)-1=a\prod_{j=1}^{s}(x-d_j)^{t_j},\quad d_j\in C$$

于是　$\sum\limits_{i=1}^{m}r_i=n=\sum\limits_{j=1}^{s}t_j$，　$m+s=|f^{-1}(0)\cup f^{-1}(1)|$

根据定理 7.5.1，我们有

$$\begin{aligned}f'(x)&=(f(x)-1)'\\&=\prod_{i=1}^{m}(x-c_i)^{r_i-1}\cdot\prod_{j=1}^{s}(x-d_j)^{t_j-1}\cdot h(x)\end{aligned}$$

其中 $h(x)$ 不能被所有 $x-c_i$ 整除，也不能被所有 $x-d_j$ 整除.

于是　$\sum\limits_{i=1}^{m}(r_i-1)+\sum\limits_{j=1}^{s}(t_j-1)\leqslant\deg f'(x)=n-1$

另一方面我们有

$$\begin{aligned}&\sum_{i=1}^{m}(r_i-1)+\sum_{j=1}^{s}(t_j-1)\\&=\sum_{i=1}^{m}r_i-m+\sum_{j=1}^{s}t_j-s=2n-(m+s)\end{aligned}$$

因此 $2n-(m+s)\leqslant n-1$. 由此得出 $m+s\geqslant n+1$. ▌

现在我们用复系数多项式唯一因式分解定理来讨论数域 $K$ 上的多项式 $f(x)$ 的复根与 $f(x)$ 的系数之间的关系.

设 $f(x)$ 是 $K[x]$ 中首项系数为 1 的 $n$ 次 $(n\geqslant 1)$ 多项式，把 $f(x)$ 看成复系数多项式，则 $f(x)$ 有 $n$ 个复根 $c_1,c_2,\cdots,c_n$(这里可能有相同的)，于是在复数域上 $f(x)$ 有因式分解

$$f(x)=(x-c_1)(x-c_2)\cdots(x-c_n) \tag{3}$$

设
$$f(x)=x^n+a_{n-1}x^{n-1}+\cdots+a_1x+a_0$$
$$a_i\in K,\quad i=0,1,\cdots,n-1$$

把(3)式的右端展开，并且比较(3)式两端的多项式的系数便得到

$$\begin{aligned}
a_{n-1}&=-(c_1+c_2+\cdots+c_n)\\
&\cdots\quad\cdots\quad\cdots\quad\cdots\quad\cdots\quad\cdots\\
a_{n-k}&=(-1)^k\sum_{1\leqslant i_1<i_2<\cdots<i_k\leqslant n}c_{i_1}c_{i_2}\cdots c_{i_k}\\
&\cdots\quad\cdots\quad\cdots\quad\cdots\quad\cdots\quad\cdots\\
a_0&=(-1)^n c_1c_2\cdots c_n
\end{aligned} \tag{4}$$

这些公式是用多项式 $f(x)$ 的复根表示它的系数的表达式，称为 Vieta 公式. 注意公式中的 $c_1,c_2,\cdots,c_n$ 是复数，但是 $a_{n-1},a_{n-2},\cdots,a_0$ 是数域 $K$ 中的元素.

如果 $f(x)$ 首项系数 $a_n\neq 1$，则公式(4)给出的是比值 $\frac{a_i}{a_n}$ 的表达式.

## 习　题　7.6

1. 设 $f(x)=x^5+7x^4+19x^3+26x^2+20x+8\in Q[x]$，判断 $-2$ 是不是多项式 $f(x)$ 的根，如果是的话，它是几重根？

(提示：用综合除法看 $x+2$ 能否整除 $f(x)$. 关于综合除法见 §2).

2. 设 $f(x)=2x^3-7x^2+4x+a\in Q[x]$，求 $a$ 值，使 $f(x)$ 有重根，并

且求出相应的重根及其重数.

(提示:$c \in Q$是$f(x)$的重根$\Leftrightarrow x-c$是$f(x)$的重因式$\Leftrightarrow x-c$是$f(x)$与$f'(x)$的公因式$\Leftrightarrow x-c$是$(f(x),f'(x))$的因式.)

3. 求下列复系数多项式在复数域中的公共根:

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

(提示:$a \in C$是$f(x)$与$g(x)$公共根$\Leftrightarrow x-a$是$f(x)$与$g(x)$的公因式$\Leftrightarrow x-a$是$(f(x),g(x))$的因式.)

4. 设$f(x)=x^4-5x^3+ax^2+bx+9 \in Q[x]$,如果3是$f(x)$的二重根,求$a,b$.

5. 证明:在$K[x]$中,如果$(x+1)\mid f(x^{2n+1})$,则$(x^{2n+1}+1)\mid f(x^{2n+1})$.

6. 设$f_1(x),f_2(x) \in Q[x]$,证明:如果

$$(x^2+x+1)\mid f_1(x^3)+xf_2(x^3)$$

则1是$f_i(x)$的根,$i=1,2$.

7. 设$C[x]$中的多项式$f(x)\neq 0$,并且$f(x)\mid f(x^m)$,其中$m$是一个大于1的整数.证明$f(x)$在$C$中的根只能是零或单位根.

8. 求复系数多项式$x^n-1$的标准分解式.

*9. 设$K$是一个数域,设$R$是(或者可看成是)$K$的一个交换扩环.设$a \in R$,我们用$J_a$表示$K[x]$中那些多项式组成的集合,它们中每一个在$R$中有根$a$,即

$$J_a=\{f(x)\in K[x]\mid f(a)=0\}$$

证明:(1) $J_a$中存在唯一的首项系数为1的多项式$m(x)$,使得$J_a$的每个多项式都是$m(x)$的倍式;

(提示:取$m(x)$是$J_a$中次数最低的首项系数为1的多项式.)

(2) 如果$R$是无零因子环,则$m(x)$在$K[x]$中不可约.

*10. 取$K$为复数域$C$,取$R=C[A]$,其中

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

求第9题的$J_A$中的多项式$m(x)$.

(提示:注意本节开头的例子.)

11. 证明:数域$K$上任意一个不可约多项式在复数域内没有重根.

12. 设$K[x]$中$n$次多项式

$$f(x)=a_nx^n+a_{n-1}x^{n-1}+\cdots+a_1x+a_0$$

的 $n$ 个复根是 $c_1, c_2, \cdots, c_n$，求以 $bc_1, bc_2, \cdots, bc_n$ 为复根的 $K[x]$ 中的多项式，这里 $b \in K$.

（提示：用 Vieta 公式.）

* 13. 设 $A \in M_n(K)$，$A$ 的不等于零的主子式的最高级数称为 $A$ 的主秩，记作 $\mathrm{pr}(A)$. 证明：$A$ 的非零特征值的数目（重根按重数计算）不超过 $\mathrm{pr}(A)$，也不超过 $\mathrm{rank}(A)$.

（提示：利用上册第五章命题 5.6.1.）

# 阅读材料七

我们来介绍复数域 $C$ 上的一元多项式环 $C[\lambda]$ 的唯一因式分解定理在 $C[\lambda]$ 上的矩阵（$\lambda$-矩阵）中的应用.

在阅读材料三和五中我们讲了每一个 $n$ 级非零的 $\lambda$-矩阵 $A(\lambda)$ 有唯一的相抵标准形，其主对角线上非零元称为 $A(\lambda)$ 的不变因子. $A(\lambda)$ 的不变因子的数目等于 $A(\lambda)$ 的秩 $r$. $A(\lambda)$ 的不变因子 $d_1(\lambda), d_2(\lambda), \cdots, d_r(\lambda)$ 都是首项系数为 1 的多项式，并且满足

$$d_i(\lambda) \mid d_{i+1}(\lambda), \qquad i = 1, \cdots, r-1$$

现在考虑 $C[\lambda]$ 上的 $n$ 级非零矩阵 $A(\lambda)$，由于每个次数大于零的复系数多项式在复数域上都可以唯一地分解成一次因式的乘积，因此 $A(\lambda)$ 的每个次数大于零的不变因子都可以唯一地分解成互不相同的一次因式方幂的乘积. 从而可以引进下述概念.

**定义 1** 设 $A(\lambda)$ 是 $C[\lambda]$ 上的 $n$ 级非零矩阵. 把 $A(\lambda)$ 的每个次数大于零的不变因子分解成互不相同的一次因式方幂的乘积，所有这些一次因式方幂（相同的必须按出现的次数计算）称为 $A(\lambda)$ 的**初等因子**.

从定义 1 看出，如果知道了 $A(\lambda)$ 的不变因子，那么可以唯一确定出 $A(\lambda)$ 的初等因子（据 $C[\lambda]$ 中因式分解的唯一性）. 反过来，如果知道了 $A(\lambda)$ 的初等因子，能否唯一确定出 $A(\lambda)$ 的不变因子？考虑到下面的应用，我们假定 $A(\lambda)$ 的秩为 $n$，于是 $A(\lambda)$ 的不变因子有 $n$ 个. 我们把 $A(\lambda)$ 的初等因子中同一个一次因式的方幂按降幂排列，并且当同一个一次因式的方幂不到 $n$ 个时，就在后面补上适当个数的零次幂（即 1），以便凑齐 $n$ 个方幂. 于是有

$$
\begin{matrix}
(\lambda-\lambda_1)^{k_{11}},(\lambda-\lambda_1)^{k_{12}},\cdots,(\lambda-\lambda_1)^{k_{1n}} \\
(\lambda-\lambda_2)^{k_{21}},(\lambda-\lambda_2)^{k_{22}},\cdots,(\lambda-\lambda_2)^{k_{2n}} \\
\cdots\quad\cdots\quad\cdots\quad\cdots\quad\cdots\quad\cdots \\
(\lambda-\lambda_s)^{k_{s1}},(\lambda-\lambda_s)^{k_{s2}},\cdots,(\lambda-\lambda_s)^{k_{sn}}
\end{matrix} \tag{1}
$$

由于(1)中出现的一次因式方幂(除去零次幂以外)就是 $A(\lambda)$ 的全部次数大于零的不变因子的标准分解式中的一次因式方幂,所以(1) 中的一次因式方幂应分别属于 $A(\lambda)$ 的各个不变因子. 注意到 $A(\lambda)$ 的不变因子具有性质:

$$d_i(\lambda)\,|\,d_{i+1}(\lambda), \qquad i=1,2,\cdots,n-1$$

因此 $d_n(\lambda)$ 应当包含(1) 中的每个一次因式的最高方幂,$d_{n-1}(\lambda)$ 应当包含(1)中每个一次因式的次高方幂,如此下去. 换句话说,(1) 中第 1 列的一次因式方幂的乘积就是 $d_n(\lambda)$,第 2 列的一次因式方幂的乘积就是 $d_{n-1}(\lambda),\cdots$,第 $n$ 列的一次因式方幂(可能是零次幂) 的乘积就是 $d_1(\lambda)$. 这样,从 $A(\lambda)$ 的初等因子便唯一确定出了 $A(\lambda)$ 的不变因子.

既然秩为 $n$ 的 $\lambda$- 矩阵 $A(\lambda)$ 的不变因子与初等因子可以互相唯一确定,于是从阅读材料五的定理 3 立即得到

**定理** 1　$C[\lambda]$ 上两个满秩的 $n$ 级矩阵相抵的充分必要条件是它们有相同的初等因子. ▌

设 $A(\lambda)$ 是 $C[\lambda]$ 上的 $n$ 级满秩矩阵,$A(\lambda)$ 的初等因子比起不变因子较容易求出. 现在我们来介绍初等因子的求法:

**定理** 2　设 $A(\lambda)$ 是 $C[\lambda]$ 上的 $n$ 级满秩矩阵,通过初等变换把 $A(\lambda)$ 化成对角形,然后把主对角线上每个次数大于零的多项式分解成互不相同的一次因式方幂的乘积,则所有这些一次因式方幂(相同的按出现的次数计算) 就是 $A(\lambda)$ 的初等因子.

**证明**　设 $A(\lambda)$ 相抵于

$$G(\lambda)=\operatorname{diag}\{h_1(\lambda),h_2(\lambda),\cdots,h_n(\lambda)\}$$

其中每个 $h_i(\lambda)$ 的首项系数为 1. 设 $h_i(\lambda)$ 的标准分解式为

$$h_i(\lambda)=(\lambda-\lambda_1)^{k_{1i}}(\lambda-\lambda_2)^{k_{2i}}\cdots(\lambda-\lambda_m)^{k_{mi}}, \quad i=1,2,\cdots,n$$

若 $h_i(\lambda)=1$,则各个一次因式方幂的指数为 0. 现在把 $G(\lambda)$ 的主对角线上每个相同的一次因式方幂按递升幂次排列,使幂指数最高的方幂位于 $(n,n)$ 元,次高的位于 $(n-1,n-1)$ 元. 这样得到的 $\lambda$- 矩阵记作 $L(\lambda)$. 这时 $L(\lambda)$ 的

主对角线上元素具有$(i,i)$元整除$(i+1,i+1)$元的特点，$i=1,\cdots,n-1$. 如果能证明 $G(\lambda)$ 与 $L(\lambda)$ 相抵，那么 $L(\lambda)$ 就是 $A(\lambda)$ 的相抵标准形，从而 $L(\lambda)$ 的主对角线上全部一次因式方幂就是 $A(\lambda)$ 的初等因子. 而 $L(\lambda)$ 的这些一次因式方幂正是 $G(\lambda)$ 的所有一次因式方幂.

首先对 $\lambda-\lambda_1$ 的方幂进行讨论. 考虑$G(\lambda)$的相邻的两个主对角元$h_i(\lambda)$，$h_{i+1}(\lambda)$，其中 $k_{1i}>k_{1,i+1}$. 设

$$h_i(\lambda)=(\lambda-\lambda_1)^{k_{1i}}g_1(\lambda),\qquad h_{i+1}(\lambda)=(\lambda-\lambda_1)^{k_{1,i+1}}g_2(\lambda)$$

其中 $g_j(\lambda)$ 与 $\lambda-\lambda_1$ 互素，$j=1,2$. 考虑两个 2 级对角 $\lambda$- 矩阵

$$Q_1(\lambda)=\begin{pmatrix}(\lambda-\lambda_1)^{k_{1i}}g_1(\lambda) & 0\\ 0 & (\lambda-\lambda_1)^{k_{1,i+1}}g_2(\lambda)\end{pmatrix}$$

$$Q_2(\lambda)=\begin{pmatrix}(\lambda-\lambda_1)^{k_{1,i+1}}g_1(\lambda) & 0\\ 0 & (\lambda-\lambda_1)^{k_{1i}}g_2(\lambda)\end{pmatrix}$$

显然 $Q_1(\lambda)$ 与 $Q_2(\lambda)$ 的 2 级行列式因子相同. $Q_1(\lambda)$ 的 1 级行列式因子为

$$\begin{aligned}&\left((\lambda-\lambda_1)^{k_{1i}}g_1(\lambda),(\lambda-\lambda_1)^{k_{1,i+1}}g_2(\lambda)\right)\\ &=(\lambda-\lambda_1)^{k_{1,i+1}}\left((\lambda-\lambda_1)^{k_{1i}-k_{1,i+1}}g_1(\lambda),g_2(\lambda)\right)\\ &=(\lambda-\lambda_1)^{k_{1,i+1}}(g_1(\lambda),g_2(\lambda))\end{aligned}$$

上述最后一个等号是根据习题 7.4 第 7 题结果. 同理可求出 $Q_2(\lambda)$ 的 1 级行列式因子为

$$(\lambda-\lambda_1)^{k_{1,i+1}}(g_1(\lambda),g_2(\lambda))$$

这表明 $Q_1(\lambda)$ 与 $Q_2(\lambda)$ 有相同的各级行列式因子，因此它们相抵. 通过每次调整主对角线上相邻两个多项式的$(\lambda-\lambda_1)$ 的方幂的位置，可以得到一个对角 $\lambda$- 矩阵 $G_1(\lambda)$，它的主对角线上元素是按照 $\lambda-\lambda_1$ 的升幂排列的. 据上述知，$G(\lambda)$ 与 $G_1(\lambda)$ 相抵.

现在对于 $G_1(\lambda)$，考虑 $\lambda-\lambda_2$ 的方幂. 依次进行下去，最后便得到 $L(\lambda)$. 从而 $G(\lambda)$ 与 $L(\lambda)$ 相抵. ▌

设 $A$ 是数域 $K$ 上的 $n$ 级矩阵，它的特征矩阵 $\lambda I-A$ 是满秩的(即秩为 $n$)，现在取 $K$ 为复数域，则由定理 1 和阅读材料五的定理 3 立即得到

**定理 3** 复数域上两个 $n$ 级矩阵的特征矩阵相抵的充分必要条件为，它们有相同的不变因子，或者它们有相同的初等因子. ▌

在阅读材料五中已指出，$\lambda I-A$ 的 $n$ 个不变因子的乘积等于 $A$ 的特征多项式. 从定义 1 即得，$\lambda I-A$ 的所有初等因子的乘积等于 $A$ 的特征多项式.

今后我们把数域 $K$ 上矩阵 $A$ 的特征矩阵 $\lambda I-A$ 的不变因子就叫做 $A$ **的不变因子**；把复矩阵 $A$ 的特征矩阵 $\lambda I-A$ 和初等因子就叫做 $A$ **的初等因子**.

我们写阅读材料三、五、七的目的，一方面是分别介绍带余除法、最大公因式、唯一因式分解定理的应用，另一方面是要寻求数域 $K$ 上 $n$ 级矩阵的集合 $M_n(K)$ 在相似关系下的一组完全不变量，以及每个 $n$ 级复矩阵的相似标准形. 现在来做这第二件事情.

$K[\lambda]$ 上的 $n$ 级矩阵可以按 $\lambda$ 的方幂"展开"，譬如，

$$A(\lambda)=\begin{pmatrix}5\lambda^3+2\lambda-1 & 3\lambda^2+1\\ \lambda^3+4\lambda & 7\lambda^3-2\lambda^2\end{pmatrix}$$

$$=\lambda^3\begin{pmatrix}5 & 0\\ 1 & 7\end{pmatrix}+\lambda^2\begin{pmatrix}0 & 3\\ 0 & -2\end{pmatrix}+\lambda\begin{pmatrix}2 & 0\\ 4 & 0\end{pmatrix}+\begin{pmatrix}-1 & 1\\ 0 & 0\end{pmatrix}$$

这从 $K[\lambda]$ 上矩阵的加法和纯量乘法(用环 $K[\lambda]$ 中一个元素去乘 $K[\lambda]$ 上的一个矩阵)可立即得出. 上述 $A(\lambda)$ 的展开式中的 $\lambda$ 的方幂的"系数"是数域 $K$ 上的矩阵.

当我们把 $A(\lambda)$ 与 $B(\lambda)$ 分别按 $\lambda$ 的方幂展开时，根据两个 $\lambda$-矩阵相等的定义以及两个一元多项式相等的定义，可以推出，$A(\lambda)$ 与 $B(\lambda)$ 相等当且仅当它们的展开式中 $\lambda^i$ 的"系数矩阵"对应相等，$i=0,1,2,\cdots$.

**引理 1** 对于 $K[\lambda]$ 上的 $n$ 级非零矩阵 $U(\lambda)$ 以及 $K$ 上 $n$ 级矩阵 $A$，一定存在 $K[\lambda]$ 上的两个 $n$ 级矩阵 $H(\lambda)$，$G(\lambda)$ 以及 $K$ 上两个 $n$ 级矩阵 $R_1$，$R_2$，使得

$$U(\lambda)=(\lambda I-A)H(\lambda)+R_1 \tag{2}$$

$$U(\lambda)=G(\lambda)(\lambda I-A)+R_2 \tag{3}$$

**证明** 把 $U(\lambda)$ 按 $\lambda$ 的方幂展开：

$$U(\lambda)=D_0\lambda^m+D_1\lambda^{m-1}+\cdots+D_{m-1}\lambda+D_m,\qquad D_0\neq 0 \tag{4}$$

若 $m=0$，则 $U(\lambda)=D_0$，此时有

$$D_0=(\lambda I-A)\cdot 0+D_0$$

下面设 $m>0$，假设

$$H(\lambda)=Q_0\lambda^{m-1}+Q_1\lambda^{m-2}+\cdots+Q_{m-2}\lambda+Q_{m-1} \tag{5}$$

其中 $Q_0,Q_1,\cdots,Q_{m-1}$ 是待定的 $K$ 上的矩阵. 于是

$$\begin{aligned}(\lambda I-A)H(\lambda)=&Q_0\lambda^m+(Q_1-AQ_0)\lambda^{m-1}+\cdots+(Q_k-AQ_{k-1})\lambda^{m-k}\\&+\cdots+(Q_{m-1}-AQ_{m-2})\lambda-AQ_{m-1}\end{aligned} \tag{6}$$

将(4)和(6)代入(2)中，比较两边 $\lambda^i$ 的系数，得

$$D_0 = Q_0,\quad D_1 = Q_1 - AQ_0,\quad \cdots,\quad D_k = Q_k - AQ_{k-1}$$

$$\cdots,\quad D_{m-1} = Q_{m-1} - AQ_{m-2},\quad D_m = -AQ_{m-1} + R_1$$

从而

$$Q_0 = D_0,\quad Q_1 = D_1 + AQ_0,\quad \cdots,\quad Q_k = D_k + AQ_{k-1}$$

$$\cdots,\quad Q_{m-1} = D_{m-1} + AQ_{m-2},\quad R_1 = D_m + AQ_{m-1}$$

这证明了存在 $H(\lambda)$ 和 $R_1$ 满足(2)式.

同理可证，存在 $G(\lambda)$ 和 $R_2$ 满足(3)式. ▎

现在我们来证明主要定理：

**定理 4** 数域 $K$ 上两个 $n$ 级矩阵 $A$ 与 $B$ 相似的充分必要条件是它们的特征矩阵 $\lambda I - A$ 与 $\lambda I - B$ 相抵.

**证明** 必要性. 若 $A \sim B$，则存在 $K$ 上可逆矩阵 $S$，使得 $B = S^{-1}AS$.

从而

$$S^{-1}(\lambda I - A)S = S^{-1}(\lambda I)S - S^{-1}AS = \lambda I - B$$

由于 $S$ 也可看成是 $K[\lambda]$ 上的可逆矩阵. 因此上式表明：$\lambda I - A$ 与 $\lambda I - B$ 相抵(据阅读材料五的推论 1).

充分性. 若 $\lambda I - A$ 与 $\lambda I - B$ 相抵，则存在 $K[\lambda]$ 上的可逆矩阵 $P(\lambda)$，$Q(\lambda)$，使得

$$P(\lambda)(\lambda I - A)Q(\lambda) = \lambda I - B \tag{7}$$

从而有

$$(\lambda I - A)Q(\lambda) = P(\lambda)^{-1}(\lambda I - B) \tag{8}$$

$$P(\lambda)(\lambda I - A) = (\lambda I - B)Q(\lambda)^{-1} \tag{9}$$

据引理 1，存在 $K[\lambda]$ 上矩阵 $H(\lambda)$，$G(\lambda)$ 以及 $K$ 上矩阵 $R_1$，$R_2$，使得

$$P(\lambda) = (\lambda I - B)H(\lambda) + R_1 \tag{10}$$

$$Q(\lambda) = G(\lambda)(\lambda I - B) + R_2 \tag{11}$$

于是我们有

$$\begin{aligned}
&R_1(\lambda I - A)R_2\\
&= [P(\lambda) - (\lambda I - B)H(\lambda)](\lambda I - A)[Q(\lambda) - G(\lambda)(\lambda I - B)]\\
&= [P(\lambda)(\lambda I - A) - (\lambda I - B)H(\lambda)(\lambda I - A)][Q(\lambda) - G(\lambda)(\lambda I - B)]\\
&= [(\lambda I - B)Q(\lambda)^{-1} - (\lambda I - B)H(\lambda)(\lambda I - A)][Q(\lambda) - G(\lambda)(\lambda I - B)]\\
&= (\lambda I - B)[I - H(\lambda)(\lambda I - A)Q(\lambda) - Q(\lambda)^{-1}G(\lambda)(\lambda I - B)\\
&\quad + H(\lambda)(\lambda I - A)G(\lambda)(\lambda I - B)]\\
&= (\lambda I - B)\{I - [H(\lambda)P(\lambda)^{-1} + Q(\lambda)^{-1}G(\lambda)\\
&\quad - H(\lambda)(\lambda I - A)G(\lambda)](\lambda I - B)\}\\
&= (\lambda I - B)[I - L(\lambda)(\lambda I - B)]
\end{aligned} \tag{12}$$

其中　$L(\lambda)=H(\lambda)P(\lambda)^{-1}+Q(\lambda)^{-1}G(\lambda)-H(\lambda)(\lambda I-A)G(\lambda)$

假如 $L(\lambda)\neq 0$，则 $L(\lambda)$ 按 $\lambda$ 的方幂的展开式中 $\lambda$ 的最高次幂的指数 $r\geqslant 0$. 于是(12)式右端的 $\lambda$ 的最高次幂的指数为 $r+2\geqslant 2$. 然而(12)式左端的 $\lambda$ 的最高次幂的指数为 1，矛盾. 因此 $L(\lambda)=0$. 从而(12)式成为

$$R_1(\lambda I-A)R_2=\lambda I-B \tag{13}$$

由此推出

$$R_1R_2=I \tag{14}$$

$$R_1AR_2=B \tag{15}$$

(14)式表明 $R_2$ 可逆且 $R_2^{-1}=R_1$. 于是(15)式成为

$$R_2^{-1}AR_2=B$$

这证明了 $A$ 与 $B$ 相似. ▌

从定理 4 和定理 3 以及阅读材料五的定理 3 立即得到

**定理 5**　数域 $K$ 上两个 $n$ 级矩阵相似的充分必要条件为它们有相同的不变因子；两个 $n$ 级复矩阵相似的充分必要条件为它们有相同的不变因子，或者有相同的初等因子. ▌

这样我们得到了：不变因子是 $M_n(K)$ 在相似关系下的一组完全不变量，其中 $K$ 是任意一个数域.

我们还得到了：初等因子是 $M_n(C)$ 在相似关系下的一组完全不变量.

利用初等因子是 $M_n(C)$ 在相似关系下的一组完全不变量，我们可以解决任一 $n$ 级复矩阵的相似标准形的问题.

首先我们问：什么样的复矩阵，其初等因子只有一个 $(\lambda-\lambda_0)^r$？考察下述形式的 $r$ 级复矩阵：

$$J_0=\begin{pmatrix}\lambda_0 & 1 & 0 & \cdots & 0 & 0\\ 0 & \lambda_0 & 1 & \cdots & 0 & 0\\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots\\ 0 & 0 & 0 & \cdots & \lambda_0 & 1\\ 0 & 0 & 0 & \cdots & 0 & \lambda_0\end{pmatrix} \tag{16}$$

$J_0$ 的特征矩阵为

$$\lambda I-J_0=\begin{pmatrix}\lambda-\lambda_0 & -1 & 0 & \cdots & 0 & 0\\ 0 & \lambda-\lambda_0 & -1 & \cdots & 0 & 0\\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots\\ 0 & 0 & 0 & \cdots & \lambda-\lambda_0 & -1\\ 0 & 0 & 0 & \cdots & 0 & \lambda-\lambda_0\end{pmatrix}$$

由于 $|\lambda I - J_0| = (\lambda - \lambda_0)^r$，所以 $\lambda I - J_0$ 的 $r$ 级行列式因子为 $(\lambda - \lambda_0)^r$. 由于 $\lambda I - J_0$ 的右上角的 $r-1$ 级子式为 $(-1)^{r-1}$，因此 $\lambda I - J_0$ 的 $r-1$ 级行列式因子为 1. 从而 $\lambda I - J_0$ 的 1 级，2 级，…，$r-1$ 级行列式因子全为 1 . 由此得出 $\lambda I - J_0$ 的不变因子为

$$d_1(\lambda) = 1, \quad d_2(\lambda) = 1, \cdots, \quad d_{n-1}(\lambda) = 1, \quad d_n(\lambda) = (\lambda - \lambda_0)^r$$

从而 $\lambda I - J_0$ 的初等因子只有一个 $(\lambda - \lambda_0)^r$. 这样我们找到了一个复矩阵 $J_0$，它的初等因子只有一个 $(\lambda - \lambda_0)^r$. $J_0$ 的主对角线上元素都为 $\lambda_0$，与主对角线平行的第一条斜线上的元素都为 1，其余元素全为 0. $J_0$ 的级数是初等因子的幂指数.

**定义 2** 形如(16)的复矩阵称为一个 $r$ 级 Jordan **块**，记作 $J_r(\lambda_0)$. 由若干个 Jordan 块组成的分块对角矩阵称为 Jordan **形矩阵**.

定义 2 前面一段说明：一个 Jordan 块被它的初等因子唯一决定.

现在我们来求 Jordan 形矩阵

$$J = \mathrm{diag}\{J_{r_1}(\lambda_1), J_{r_2}(\lambda_2), \cdots, J_{r_s}(\lambda_s)\}$$

的初等因子，其中 $r_1 + r_2 + \cdots + r_s = n$. 由于 $J_{r_i}(\lambda_i)$ 的初等因子是 $(\lambda - \lambda_i)^{r_i}$，而对角 $\lambda$- 矩阵

$$G_i(\lambda) = \mathrm{diag}\{1, \cdots, 1, (\lambda - \lambda_i)^{r_i}\}$$

的初等因子也是 $(\lambda - \lambda_i)^{r_i}$，所以 $J_{r_i}(\lambda_i)$ 的特征矩阵 $\lambda I_{r_i} - J_{r_i}(\lambda_i)$ 与 $G_i(\lambda)$ 相抵(据定理 1). 于是 $J$ 的特征矩阵

$$\lambda I - J = \mathrm{diag}\{\lambda I_{r_1} - J_{r_1}(\lambda_1), \lambda I_{r_2} - J_{r_2}(\lambda_2), \cdots, \lambda I_{r_s} - J_{r_s}(\lambda_s)\}$$

与对角 $\lambda$- 矩阵

$$G(\lambda) = \mathrm{diag}\{G_1(\lambda), G_2(\lambda), \cdots, G_s(\lambda)\}$$

相抵. 据定理 2 得，$\lambda I - J$ 的初等因子为

$$(\lambda - \lambda_1)^{r_1}, \quad (\lambda - \lambda_2)^{r_2}, \quad \cdots, \quad (\lambda - \lambda_s)^{r_s} \tag{17}$$

这表明：Jordan 形矩阵的初等因子由它的全部 Jordan 块的初等因子组成. 由于一个 Jordan 块被它的初等因子唯一决定，因此一个 Jordan 形矩阵除去其中 Jordan 块的排列次序外被它的初等因子唯一决定.

现在我们来证明第二个主要定理：

**定理 6** 每个 $n$ 级复矩阵 $A$ 都与一个 Jordan 形矩阵相似，这个 Jordan 形矩阵除去其中 Jordan 块的排列次序外被矩阵 $A$ 唯一决定，称它为 $A$ 的 Jordan 标准形.

**证明** 设 $n$ 级复矩阵 $A$ 的初等因子为

$$(\lambda-\lambda_1)^{r_1},\quad (\lambda-\lambda_2)^{r_2},\quad \cdots,\quad (\lambda-\lambda_s)^{r_s} \tag{18}$$

由于 $A$ 的所有初等因子的乘积等于 $A$ 的特征多项式，所以

$$r_1+r_2+\cdots+r_s=n$$

每一个初等因子 $(\lambda-\lambda_i)^{r_i}$ 决定一个 Jordan 块 $J_{r_i}(\lambda_i)$，这些 Jordan 块组成一个 Jordan 形矩阵

$$J=\operatorname{diag}\{J_{r_1}(\lambda_1),J_{r_2}(\lambda_2),\cdots,J_{r_s}(\lambda_s)\}$$

由前面一段知道，$J$ 的初等因子也是(18). 据定理 5 得，$A$ 与 $J$ 相似.

如果另一个 Jordan 形矩阵 $J_1$ 也与 $A$ 相似，那么 $J_1$ 与 $A$ 有相同的初等因子，从而 $J_1$ 与 $J$ 有相同的初等因子. 于是 $J_1$ 与 $J$ 除了其中 Jordan 块的排列次序外是相同的. 这证明了唯一性. ▌

$n$ 级复矩阵能相似于 Jordan 形矩阵的关键是：$C[\lambda]$ 中的每个次数大于零的多项式能唯一地分解成一次因式的乘积. 这就是我们把这部分知识放在复系数多项式唯一因式分解定理之后作为阅读材料的理由. 在第十章我们将用几何方法给出定理 6 的另一个证明.

**例 1** 求下述复矩阵的 Jordan 标准形

$$A=\begin{pmatrix}2 & 3 & 2\\ 1 & 8 & 2\\ -2 & -14 & -3\end{pmatrix}$$

**解** 先把 $A$ 的特征矩阵 $\lambda I-A$ 通过初等变换化成对角形：

$$\lambda I-A$$

$$=\begin{pmatrix}\lambda-2 & -3 & -2\\ -1 & \lambda-8 & -2\\ 2 & 14 & \lambda+3\end{pmatrix}\longrightarrow\begin{pmatrix}-1 & \lambda-8 & -2\\ 0 & \lambda^2-10\lambda+13 & -2\lambda+2\\ 0 & 2\lambda-2 & \lambda-1\end{pmatrix}$$

$$\longrightarrow\begin{pmatrix}1 & 0 & 0\\ 0 & \lambda^2-10\lambda+13 & -2(\lambda-1)\\ 0 & 2(\lambda-1) & \lambda-1\end{pmatrix}\xrightarrow{②+③2}\begin{pmatrix}1 & 0 & 0\\ 0 & \lambda^2-6\lambda+9 & 0\\ 0 & 2(\lambda-1) & \lambda-1\end{pmatrix}$$

$$\longrightarrow\begin{pmatrix}1 & 0 & 0\\ 0 & (\lambda-3)^2 & 0\\ 0 & 0 & \lambda-1\end{pmatrix}$$

据定理 2，$\lambda I-A$ 的初等因子为

$$\lambda-1,\qquad (\lambda-3)^2$$

因此 $A$ 的 Jordan 标准形是

$$\begin{pmatrix}1&0&0\\0&3&1\\0&0&3\end{pmatrix}$$

根据前面讲的一个 Jordan 块被它的初等因子唯一决定可知，初等因子为 $\lambda-\lambda_1$ 的 Jordan 块就是 1 级矩阵 $(\lambda_1)$. 因此从定理 6 立即得到

**定理 7** $n$ 级复矩阵 $A$ 可对角化的充分必要条件是 $A$ 的初等因子全为一次的.

**证明** 充分性. 由上一段议论即得.

必要性. 设 $A\sim\mathrm{diag}\{\lambda_1,\lambda_2,\cdots,\lambda_n\}$，则 $\lambda I-A$ 相抵于

$$\mathrm{diag}\{\lambda-\lambda_1,\lambda-\lambda_2,\cdots,\lambda-\lambda_n\}$$

从而据定理 2 得，$\lambda I-A$ 的初等因子为

$$\lambda-\lambda_1,\quad \lambda-\lambda_2,\quad \cdots,\quad \lambda-\lambda_n$$ ▌

练习：求下列复矩阵的 Jordan 标准形

(1) $\begin{pmatrix}0&1&0\\-4&4&0\\-2&1&2\end{pmatrix}$；　　(2) $\begin{pmatrix}4&-5&2\\5&-7&3\\6&-9&4\end{pmatrix}$；

(3) $\begin{pmatrix}1&-3&3\\-2&-6&13\\-1&-4&8\end{pmatrix}$；　　(4) $\begin{pmatrix}1&-3&4\\4&-7&8\\6&-7&7\end{pmatrix}$.

# §7 实系数多项式

这一节我们要找出实数域上的所有不可约多项式.

首先我们指出实系数多项式的复根的性质：

**定理 7.7.1** 如果 $c$ 是实系数多项式 $f(x)$ 的一个复根，则 $c$ 的共轭复数 $\bar{c}$ 也是一个复根.

**证明** 设

$$f(x)=a_nx^n+a_{n-1}x^{n-1}+\cdots+a_1x+a_0,$$
$$a_i\in R,\quad i=0,1,\cdots,n$$

因为 $c$ 是 $f(x)$ 的复根，所以

$$f(c)=a_nc^n+a_{n-1}c^{n-1}+\cdots+a_1c+a_0=0 \tag{1}$$

在(1)式两边取共轭复数得

$$a_n\bar{c}^n+a_{n-1}\bar{c}^{n-1}+\cdots+a_1\bar{c}+a_0=0 \tag{2}$$

即 $f(\bar{c})=0$. 因此 $\bar{c}$ 是 $f(x)$ 的一个复根. ▌

利用定理 7.7.1 我们可以得出

**定理** 7.7.2 实数域上的不可约多项式都是一次或某些二次的；实系数二次多项式

$$f(x)=ax^2+bx+c$$

不可约当且仅当它的判别式 $b^2-4ac<0$.

**证明** 设 $f(x)\in R[x]$ 是不可约的，把 $f(x)$ 看成复系数多项式，根据代数基本定理，$f(x)$ 有一个复根 $c$. 如果 $c$ 是实数，则 $f(x)$ 在 $R[x]$ 中有一次因式 $x-c$. 因为 $f(x)$ 不可约，所以

$$(x-c)\sim f(x)$$

即 $f(x)=a(x-c)$，$a$ 是非零实数. 因此 $f(x)$ 是一次多项式.

如果 $c$ 是虚数，据定理 7.7.1，$\bar{c}$ 也是 $f(x)$ 的一个复根. 由于 $\bar{c}\neq c$，所以 $(x-c,x-\bar{c})=1$. 在 $C[x]$ 中

$$(x-c)\mid f(x),\qquad (x-\bar{c})\mid f(x)$$

从而

$$(x-c)(x-\bar{c})\mid f(x)$$

我们有

$$(x-c)(x-\bar{c})=x^2-(c+\bar{c})x+c\bar{c}$$

而 $c+\bar{c}$ 与 $c\bar{c}$ 都是实数. 既然在 $C[x]$ 中

$$(x^2-(c+\bar{c})x+c\bar{c})\mid f(x)$$

于是 在 $R[x]$ 中

$$(x^2-(c+\bar{c})x+c\bar{c})\,|\,f(x)$$

由于 $f(x)$ 在 $R[x]$ 中不可约，因此

$$f(x)=a(x^2-(c+\bar{c})x+c\bar{c})$$

$a$ 是非零实数. 此时 $f(x)$ 是二次多项式.

实系数二次多项式 $f(x)=ax^2+bx+c$ 不可约

$\Leftrightarrow f(x)$ 在 $R[x]$ 中没有一次因式

$\Leftrightarrow f(x)$ 没有实根

$\Leftrightarrow b^2-4ac<0$ ▍

从定理 7.7.2 和唯一因式分解定理立即得出

**定理 7.7.3(实系数多项式唯一因式分解定理)** 每个次数大于零的实系数多项式在实数域上都可以唯一地分解成一次因式与判别式小于零的二次因式的乘积. ▍

因此，实系数多项式具有标准分解式

$$f(x)=a(x-c_1)^{r_1}\cdots(x-c_s)^{r_s}\cdot(x^2+p_1x+q_1)^{k_1}\cdots(x^2+p_tx+q_t)^{k_t} \qquad (3)$$

其中 $a$ 是 $f(x)$ 的首项系数；$c_1,\cdots,c_s$ 是不同的实数；$p_1,q_1;p_2,q_2;\cdots;p_t,q_t$ 是不同的实数对，并且满足 $p_i^2-4q_i<0, i=1,\cdots,t$；$r_1,\cdots,r_s,k_1,\cdots,k_t$ 都是非负整数.

从实系数多项式 $f(x)$ 的标准分解式(3) 看出，如果虚数 $b$ 是 $f(x)$ 的一个复根，则 $b$ 的共轭复数 $\bar{b}$ 也是 $f(x)$ 的一个复根，并且它们的重数相同(此时把 $f(x)$ 看成复系数多项式). 因此通常我们说，实系数多项式的虚根共轭成对出现.

# 阅读材料八

现在我们来讨论实系数多项式的实根的计算问题. 按实数的定义，所谓计算实根就是求与之逼近的有理数.

上一节我们已指出，可以运用函数论的知识来研究实系数多项式的实根. 一个实系数多项式 $f(x)$ 的实根就是多项式函数 $y = f(x)$ 的图象与 $x$ 轴的交点的横坐标. 因此通过画出多项式函数 $y = f(x)$ 的图象可以提供实根的大致位置的信息. 下面的 Weierstrass 关于连续函数的零点定理是所有关于实系数多项式的实根的定理的基础.

**定理 1（Weierstrass 关于连续函数的零点定理）** 设 $f(x)$ 是区间 $[a,b]$ 上的连续（实值）函数，如果 $f(a)f(b) < 0$，则 $f(x)$ 在 $a$ 与 $b$ 之间有一个零点（即根）.

众所周知，多项式函数 $y = f(x)$ 是到处连续的函数. 运用 Weierstrass 零点定理，在 $f(x)$ 的相邻的两个极值点之间至多有一个实根（不计重数）. 而 $f(x)$ 的极值点又是 $f'(x)$ 的实根，$f'(x)$ 的次数比 $f(x)$ 少 1. 这样从较低次的多项式的实根的位置可以推出 $f(x)$ 的实根的大致位置.

例如

$$f(x) = x^3 - x + 1$$

因为

$$f'(x) = 3x^2 - 1$$

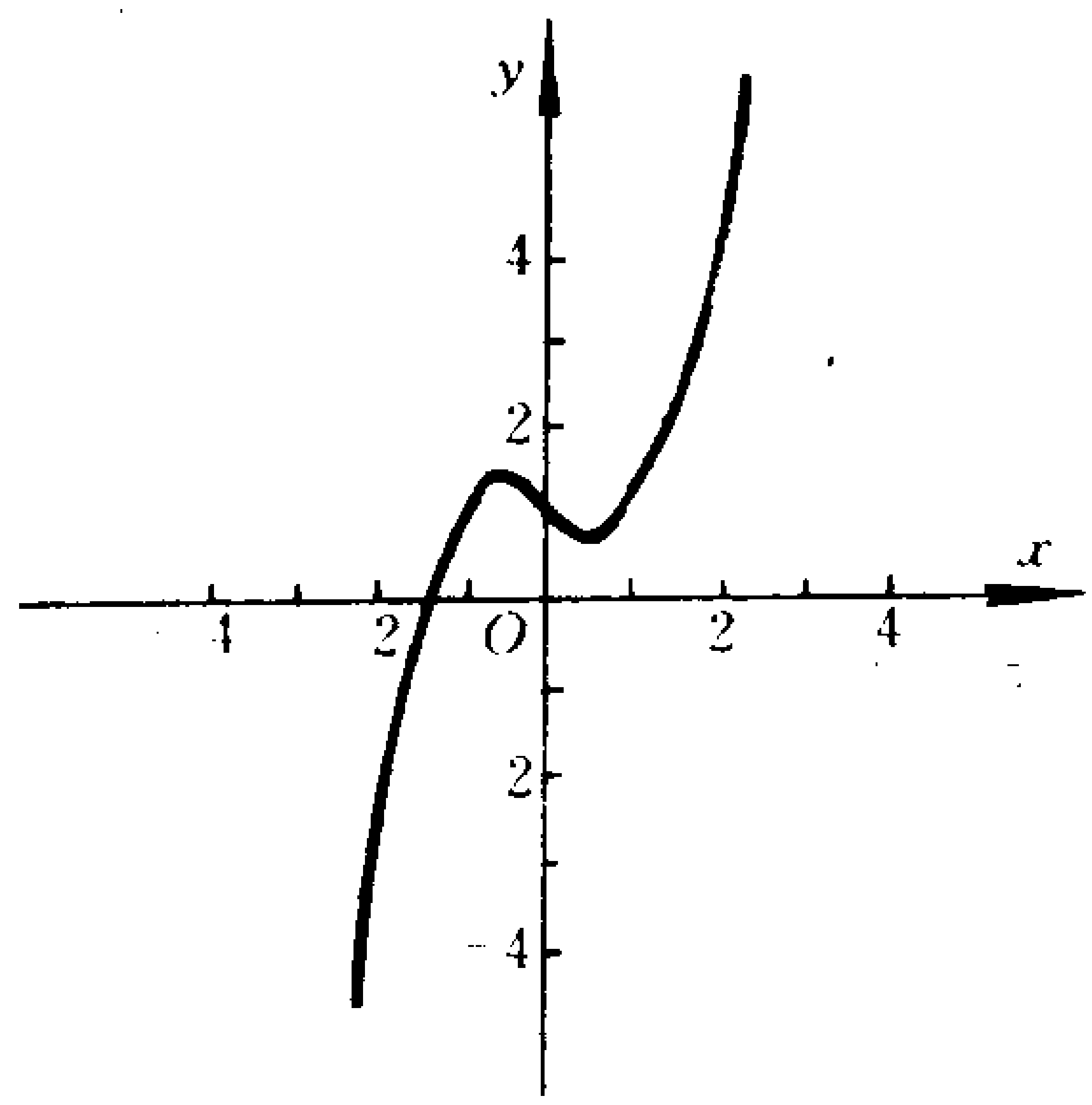
有两个实根：$-\dfrac{\sqrt{3}}{3}$，$\dfrac{\sqrt{3}}{3}$，并且 $f''(x) = 6x$，所以 $-\dfrac{\sqrt{3}}{3}$ 是 $f(x)$ 的极大值点，$\dfrac{\sqrt{3}}{3}$ 是 $f(x)$ 的极小值点. 从而 $f(x)$ 在 $-\dfrac{\sqrt{3}}{3}$ 与 $\dfrac{\sqrt{3}}{3}$ 之间至多有一个根. 由于

$$f\left(-\frac{\sqrt{3}}{3}\right) > 0, \qquad f\left(\frac{\sqrt{3}}{3}\right) > 0$$

所以 $f(x)$ 在 $-\dfrac{\sqrt{3}}{3}$ 与 $\dfrac{\sqrt{3}}{3}$ 之间没有根. 由于 $f(-2) < 0$，所以 $f(x)$ 在 $-2$ 与 $-\dfrac{\sqrt{3}}{3}$ 之间有一个根，并且这个根是单根（因为 $f(x)$ 的重根必是 $f'(x)$ 的根）. 而在 $(-\infty, -2)$ 内没有根，这是因为 $f(x)$ 在 $\left(-\infty, -\dfrac{\sqrt{3}}{3}\right)$ 是增函数. 由于 $\dfrac{\sqrt{3}}{3}$ 是 $f(x)$ 的极小值点，并且 $f(x)$ 只有 $\pm\dfrac{\sqrt{3}}{3}$ 是极值点，所以 $f(x)$ 在 $\left(\dfrac{\sqrt{3}}{3}, +\infty\right)$ 内没有根.

总之 $f(x) = x^3 - x + 1$ 只有一个实根，它是单根，且位于 $-2$ 与

$-\frac{\sqrt{3}}{3}$ 之间. $f(x)$ 的图象如下：




一般地，我们要问：一个实系数多项式 $f(x)$ 有多少个不同的实根？$f(x)$ 的所有实根在哪个区间内？即实根的界的问题. 对每一个实根能否找一个区间包含这个根而不包含其它根，即把实根分离开. 然后我们再去求每个实根的近似值.

先看实系数多项式 $f(x)$ 的实根的界的问题. 我们可以更一般地讨论复系数多项式的复根的范围，再由此得出实系数多项式的实根的界.

**定理 2** 设

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

是一个复系数多项式，其次数 $n \geqslant 1$. 令

$$M = \max\{|a_{n-1}|, |a_{n-2}|, \cdots, |a_1|, |a_0|\}$$

则当 $z \in C$ 且 $|z| \geqslant 1 + \frac{M}{|a_n|}$ 时，我们有

$$|a_n z^n| > |a_{n-1} z^{n-1} + \cdots + a_1 z + a_0|$$

**证明** 当 $M = 0$ 时，结论显然成立.

下设 $M \neq 0$. 当 $z \in C$ 且 $|z| \geqslant 1 + \frac{M}{|a_n|}$ 时，有

$$|z| > 1, \quad \text{且 } |a_n| \geqslant \frac{M}{|z| - 1}$$

从而有

$$
\begin{aligned}
|a_n z^n| &= |a_n||z|^n \geqslant \frac{M|z|^n}{|z|-1} > \frac{M(|z|^n-1)}{|z|-1} \\
&= M(|z|^{n-1} + \cdots + |z| + 1) \\
&\geqslant |a_{n-1}||z|^{n-1} + \cdots + |a_1||z| + |a_0| \\
&= |a_{n-1}z^{n-1}| + \cdots + |a_1 z| + |a_0| \\
&\geqslant |a_{n-1}z^{n-1} + \cdots + a_1 z + a_0|
\end{aligned}
$$ ▌

由定理 2 得出，当 $|z| \geqslant 1 + \dfrac{M}{|a_n|}$ 时，有

$$
\begin{aligned}
|f(z)| &= |a_n z^n + a_{n-1}z^{n-1} + \cdots + a_1 z + a_0| \\
&\geqslant |a_n z^n| - |a_{n-1}z^{n-1} + \cdots + a_1 z + a_0| > 0
\end{aligned}
$$

因此 $f(x)$ 的复根全都在以原点为圆心，以 $1 + \dfrac{M}{|a_n|}$ 为半径的圆内. 把这一结论用到实系数多项式上便得到

**推论 1** 设

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

是一个实系数多项式，其次数 $n \geqslant 1$. 令

$$M = \max\{|a_{n-1}|, |a_{n-2}|, \cdots, |a_1|, |a_0|\}$$

则 $f(x)$ 的实根全都在区间 $\left(-1 - \dfrac{M}{|a_n|}, 1 + \dfrac{M}{|a_n|}\right)$ 内. ▌

从定理 2 还可以得到

**推论 2** 设

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

是一个次数大于零的实系数多项式，则对一切充分大的正数 $r$，$f(r)$ 的符号与 $a_n r^n$ 的符号一样. ▌

例如，设 $f(x) = x^3 - x + 1$. 我们有 $M = 1, 1 + \dfrac{M}{|a_n|} = 2$. 因此 $f(x)$ 的实根全都在区间 $(-2, 2)$ 内.

注意：求出了一个实系数多项式 $f(x)$ 的实根的界只是表明：如果 $f(x)$ 有实根，则所有实根都在这个区间内. 但是不能肯定 $f(x)$ 一定有实根. 如何知道 $f(x)$ 有没有实根？如果有的话，实根的数目（不计重数）是多少？如何把实根分离开？对这些问题的第一个令人满意的回答是在 1829 年由 Sturm 给出的. 下面我们介绍 Sturm 的方法. 先给出一个概念：

**定义 1** 设 $c_1, c_2, \cdots, c_m$ 是一个非零实数的有限序列，如果 $c_i c_{i+1} < 0$，则

我们说，在第 $i+1$ 项有一个**变号**. 这个序列中变号的总数称为它的**变号数**. 一个有限的实数序列的变号数定义为去掉这个序列中的 0 以后得到的序列的变号数.

例如，序列 $-2,0,1,0,0,3,-4,5$ 的变号数是 3.

**定理** 3(Sturm **定理**) 设 $f(x)$ 是一个次数大于零的实系数多项式. 对 $f(x)$ 与 $f'(x)$ 做下述略微修改的辗转相除法

$$\begin{aligned} f(x) &= q_1(x)f'(x) - f_2(x), \quad \deg f_2 < \deg f' \\ f'(x) &= q_2(x)f_2(x) - f_3(x), \quad \deg f_3 < \deg f_2 \\ &\cdots\ \cdots\ \cdots\ \cdots\ \cdots\ \cdots\ \cdots \\ f_{s-1}(x) &= q_s(x)f_s(x) \end{aligned} \tag{1}$$

得到一个多项式序列

$$f_0 = f, \quad f_1 = f', \quad f_2, \quad f_3, \cdots, \quad f_s \tag{2}$$

称序列(2)是 $f(x)$ 的**标准序列**. 假设区间$[a,b]$使得 $f(a)\neq 0, f(b)\neq 0$，则 $f(x)$ 在区间$(a,b)$内的不同的实根的数目是 $V_a - V_b$，其中 $V_c$ 表示序列 $f_0(c), f_1(c), f_2(c), \cdots, f_s(c)$ 的变号数.

我们先看一个例子. 设 $f(x)=x^3-x+1$，则 $f'(x)=3x^2-1$. 对 $f(x)$ 与 $f'(x)$ 做略微修改的辗转相除法，即把每次得到的余式反号以后去除除式：

| | | | |
|---|---|---|---|
| $\frac{9}{2}x+\frac{27}{4}$ | $f'(x)= 3x^2 - 1$ | $f(x)= x^3 - x + 1$ | $\frac{1}{3}x$ |
| | $3x^2 - \frac{9}{2}x$ | $x^3 - \frac{1}{3}x$ | |
| | $\frac{9}{2}x - 1$ | $-\frac{2}{3}x + 1$ | |
| | $\frac{9}{2}x - \frac{27}{4}$ | $f_2(x)= \frac{2}{3}x - 1$ | |
| | $\frac{23}{4}$ | | |
| | $f_3(x)= -\frac{23}{4}$ | | |

因此

$$f_0 = f, \quad f_1 = f', \quad f_2 = \frac{2}{3}x - 1, \quad f_3 = -\frac{23}{4}$$

从 $f_3 = -\frac{23}{4}$ 知道，$(f(x), f'(x)) = 1$，因此 $f(x)$ 没有重根. 在前面我们已计

算过 $f(x)$ 的实根全在区间 $(-2,2)$ 内，现在计算 $f(x)$ 的标准序列在 $-2$ 与 2 处的变号数：

| | $f_0$ | $f_1$ | $f_2$ | $f_3$ | $V$ |
|---|---|---|---|---|---|
| $-2$ | $-$ | $+$ | $-$ | $-$ | 2 |
| 2 | $+$ | $+$ | $+$ | $-$ | 1 |

于是 $f(x)$ 在 $(-2,2)$ 内的不同的实根的数目是 $V_{-2}-V_2=1$. 由于 $f(x)$ 没有重根，所以 $f(x)$ 的实根的总数是 1. 为了进一步确定这个实根的位置，可以再去计算 $V_0, V_{-1}, V_{-\frac{3}{2}}$：

| | $f_0$ | $f_1$ | $f_2$ | $f_3$ | $V$ |
|---|---|---|---|---|---|
| 0 | $+$ | $-$ | $-$ | $-$ | 1 |
| $-1$ | $+$ | $+$ | $-$ | $-$ | 1 |
| $-\frac{3}{2}$ | $-$ | $+$ | $-$ | $-$ | 2 |

因为 $V_{-\frac{3}{2}}-V_{-1}=1$，所以 $f(x)$ 的唯一实根在 $\left(-\dfrac{3}{2},-1\right)$ 内.

Sturm **定理的证明**　$f(x)$ 的标准序列的最后一个多项式 $f_s$ 是 $f(x)$ 与 $f'(x)$ 的一个最大公因式，从而它是每个 $f_i$ 的因式 $(i=0,1,\cdots,s)$. 用 $f_s$ 去除 $f_i$ 得到的商组成另一个多项式序列：$g_0, g_1, g_2, \cdots, g_s$，其中 $g_s=1$. 在 §5 中我们已指出，$g_0(x)$ 没有重因式，并且 $g_0(x)$ 与 $f(x)$ 有完全相同的不可约因式（不计重数）. 因此 $g_0(x)$ 没有重根，并且 $g_0(x)$ 与 $f(x)$ 含有完全相同的根（不计重数）. 因为 $f_i=g_if_s$，所以

$$f_i(c)=g_i(c)f_s(c), \qquad \forall\, c\in R$$

因为 $f(a)\neq 0$，所以 $f_s(a)\neq 0$. 同理 $f_s(b)\neq 0$. 因此数列 $f_0(a), f_1(a), \cdots, f_s(a)$ 与数列 $g_0(a), g_1(a), \cdots, g_s(a)$ 的变号数相同. 我们用 $V_a(g)$ 表示第二个数列的变号数，则 $V_a=V_a(g)$. 同理 $V_b=V_b(g)$. 因此

$$V_a-V_b=V_a(g)-V_b(g)$$

下面我们来证 $g_0(x)$ 在区间 $(a,b)$ 内的实根的数目等于 $V_a(g)-V_b(g)$，从而 $f(x)$ 在 $(a,b)$ 内的不同的实根的数目等于 $V_a-V_b$.

多项式序列 $g_0,g_1,g_2,\cdots,g_s$ 称为 $g_0(x)$ 对于区间$[a,b]$的一个 Sturm 序列. 从(1)式可得出

$$\begin{aligned}
g_0 &= q_1g_1 - g_2\\
g_1 &= q_2g_2 - g_3\\
&\cdots\quad\cdots\quad\cdots\quad\cdots\\
g_{s-1} &= q_sg_s
\end{aligned}\tag{3}$$

其中 $g_s = 1$. 显然 $g_0(a) \neq 0, g_0(b) \neq 0$. 序列 $g_0,g_1,\cdots,g_s$ 中相邻两个多项式不可能有公共根,否则将与 $g_s$ 是常数矛盾. 我们把序列 $g_0,g_1,\cdots,g_s$ 中各个多项式在$(a,b)$内的全部实根按由小到大的顺序排列成:$c_1 < c_2 < \cdots < c_m$,并且令 $c_0 = a, c_{m+1} = b$. 于是区间$(a,b)$包含 $m+1$ 个小区间:

$$(c_0,c_1),(c_1,c_2),\cdots,(c_m,c_{m+1})$$

证明的基本思路是把变号数 $V_u(g)$ 看成 $u$ 的函数. 如果我们能证明:当 $u$ 从 $a$ 变到 $b$ 时,每经过 $g_0(x)$ 的一个实根,$V_u(g)$ 的值就减少 1,而在其他情况下,$V_u(g)$ 的值保持不变,那么 $g_0(x)$ 在区间$(a,b)$内的实根的数目就等于 $V_a(g) - V_b(g)$.

由于每一个多项式 $g_i(i=0,1,\cdots,s)$ 在区间$(c_j,c_{j+1})$内$(j=0,1,\cdots,m)$都没有实根,因此多项式函数 $g_i$ 在$(c_j,c_{j+1})$内的值保持相同的符号. 于是对于$(c_j,c_{j+1})$内任意两点 $u_1,u_2$,有 $V_{u_1}(g) = V_{u_2}(g)$. 即当 $u$ 跑遍$(c_j,c_{j+1})$内每一点时,$V_u(g)$ 的值保持不变,$j = 0,1,\cdots,m$. 也容易证明当 $u$ 跑遍$[a,c_1)$的每一点时,$V_u(g)$ 的值保持不变:设 $u \in (a,c_1)$,我们来证 $V_a(g) = V_u(g)$. 如果 $a$ 不是序列 $g_0,g_1,\cdots,g_s$ 中任何一个多项式的根,则每个 $g_i$ 在区间$[a,u]$上没有根. 同上理得,$V_a(g) = V_u(g)$. 现在假设 $a$ 是某个多项式 $g_k$ 的根,显然 $0 < k < s$. 因为 $g_{k-1} = q_kg_k - g_{k+1}$,并且相邻的两个多项式没有公共根,所以 $g_{k-1}(a)g_{k+1}(a) < 0$. 因为在区间$[a,c_1)$上 $g_{k-1}$ 与 $g_{k+1}$ 都没有根,所以 $g_{k-1}(a)$ 与 $g_{k-1}(u)$ 同号,$g_{k+1}(a)$ 也与 $g_{k+1}(u)$ 同号,从而 $g_{k-1}(u)g_{k+1}(u) < 0$. 于是数列 $g_{k-1}(a),0,g_{k+1}(a)$ 的变号数是 1;数列 $g_{k-1}(u),g_k(u),g_{k+1}(u)$ 的变号数也是 1. 由此推出 $V_a(g) = V_u(g)$. 同理,$u$ 跑遍$(c_m,b]$的每一点时,$V_u(g)$ 的值保持不变. 因此我们只要考察 $u$ 经过 $c_j(j=1,2,\cdots,m)$ 时 $V_u(g)$ 的值如何改变.

情形 1　$c_j$ 不是 $g_0(x)$ 的根,则 $c_j$ 是某个 $g_k(x)$ 的根,其中 $0 < k < s$. 在$(c_{j-1},c_j)$内任取一点 $u_1$. 同上理得出,$V_{u_1}(g) = V_{c_j}(g)$. 在$(c_j,c_{j+1})$内任取一点 $u_2$,同理 $V_{c_j}(g) = V_{u_2}(g)$. 因此当 $u$ 经过 $c_j$ 时,$V_u(g)$ 的值保持不变.

情形 2　$c_j$ 是 $g_0(x)$ 的根. 设 $c_j$ 是 $f(x)$ 的 $l$ 重根,则

$$f(x) = (x - c_j)^l h(x), \qquad h(c_j) \neq 0$$

于是
$$f'(x) = l(x - c_j)^{l-1} h(x) + (x - c_j)^l h'(x)$$

由于 $(x - c_j)^{l-1}$ 是 $f(x)$ 与 $f'(x)$ 的公因式,所以它是 $f_s(x)$ 的一个因式,从而

$$f_s(x) = (x - c_j)^{l-1} k(x), \qquad k(c_j) \neq 0$$

因为 $f_s(x) | f(x)$, 所以 $k(x) | h(x)$. 设 $h(x) = k(x)p(x)$. 又因为 $f_s(x) | f'(x)$,所以 $k(x) | h'(x)$. 设 $h'(x) = k(x)m(x)$. 于是

$$g_0(x) = (x - c_j)p(x), \qquad p(c_j) \neq 0$$

$$g_1(x) = (x - c_j)m(x) + lp(x)$$

从而 $g_1(c_j) = lp(c_j) \neq 0$. 因此 $g_1$ 在 $(c_{j-1}, c_{j+1})$ 内没有根,从而 $g_1$ 在 $(c_{j-1}, c_{j+1})$ 内的值保持相同的符号,即与 $p(c_j)$ 同号. 在 $(c_{j-1}, c_j)$ 内取一点 $u_1$,在 $(c_j, c_{j+1})$ 内取一点 $u_2$,使得 $p(x)$ 在区间 $[u_1, u_2]$ 上没有实根. 于是 $p(x)$ 在区间 $[u_1, u_2]$ 上的值保持相同的符号,即与 $p(c_j)$ 同号. 从而 $g_0(u_1)$ 与 $p(c_j)$ 反号,$g_0(u_2)$ 与 $p(c_j)$ 同号. 所以

$$g_0(u_1)g_1(u_1) < 0, \qquad g_0(u_2)g_1(u_2) > 0$$

于是数列 $g_0(u_1), g_1(u_1)$ 的变号数是1;而数列 $g_0(u_2), g_1(u_2)$ 的变号数为0. $c_j$ 当然也有可能是序列 $g_0, g_1, \cdots, g_s$ 的中间某些多项式 $g_k(0 < k < s)$ 的根,但情形 1 中已证明数列 $g_{k-1}(u_1), g_k(u_1), g_{k+1}(u_1)$ 的变号数与数列 $g_{k-1}(u_2), g_k(u_2), g_{k+1}(u_2)$ 的变号数相同. 因此 $V_{u_1}(g) - V_{u_2}(g) = 1$,即当 $u$ 经过 $g_0(x)$ 的一个根 $c_j$ 时,$V_u(g)$ 的值减少 1. 至此我们完成了 Sturm 定理的证明. ▌

Sturm 定理既能求出一个实系数多项式 $f(x)$ 的不同的实根的数目,又能把实根分离开(象前面的例子中那样的做法). 当我们把实根分离开后,如果想进一步求出实根的近似值,那么可以用计算机来做. 有关实根近似值的算法在计算数学的书中可以找到,这里就不介绍了.

## 习　题　7.7

1. 证明:实系数的奇次多项式至少有一个实根.
2. 求多项式 $x^n - 1$ 在实数域上的标准分解式.

* 3. 证明 $x^3 - 7x - 7$ 在 $(-2, -1)$ 内有两个不同的实根.

* 4. 求 $x^4 + 12x^2 + 5x - 9$ 的不同的实根的数目.

* 5. 求 $x^3 - 5x^2 + 8x - 8$ 的不同的实根的数目,这些根在哪些相邻的整数之间?

6. (1) 证明:实数域上斜对称矩阵的特征多项式的复根都是纯虚数或零.

(提示:参看定理 5.9.2 的证明方法);

(2) 证明:可逆的实斜对称矩阵的特征多项式的不可约因式都是二次的.

* 7. 证明:如果 $n$ 级实矩阵 $A$ 与 $B$ 不相似,则把它们看成复矩阵后仍然不相似.

(提示:假如 $A$ 和 $B$ 作为复矩阵是相似的,则存在复数域上的 $n$ 级可逆矩阵 $U$,使 $U^{-1}AU = B$. 设 $U = P + \mathrm{i}Q$,其中 $P,Q$ 是实矩阵. 设法找一个实可逆矩阵 $S$,使 $S^{-1}AS = B$.)

## §8 有理系数多项式

这一节讨论有理数域上的不可约多项式有哪些?如何判别一个有理系数多项式是否不可约?这个问题的回答比复系数多项式与实系数多项式困难得多. 有理数域上的不可约多项式在代数数论等数学分支中起着重要的作用.

设 $f(x) \in Q[x]$,由于 $f(x)$ 的任一相伴元与 $f(x)$ 只相差一个非零有理数因子,因此 $f(x)$ 与它的任一相伴元在有理数域上有相同的因式. 从而 $f(x)$ 在 $Q[x]$ 中不可约当且仅当它的相伴元在 $Q[x]$ 中不可约. 这样我们可以从 $f(x)$ 的相伴元中选择一个最简单的多项式作为代表来研究它的不可约性. 这个代表可以按下述方法来选取:设 $f(x)$ 的各项系数的分母的最小公倍数为 $m$,则 $mf(x)$ 是一个整系数多项式. 设 $mf(x)$ 的各项系数的最大公因数为 $d$,则 $\frac{m}{d}f(x)$ 是一个各项系数的最大公因数为 $\pm 1$ 的整系数多

项式. $\frac{m}{d}f(x)$ 是 $f(x)$ 的相伴元,并且又最简单,我们就取$\frac{m}{d}f(x)$ 作为代表. 我们给这种类型的多项式取一个名字:

**定义 1** 一个非零的整系数多项式

$$g(x) = b_nx^n + \cdots + b_1x + b_0$$

如果它的各项系数 $b_n, b_{n-1}, \cdots, b_1, b_0$ 的最大公因数为 $\pm 1$,则称 $g(x)$ 是一个**本原多项式**.

从上面一段知道,任何一个非零的有理系数多项式 $f(x)$ 与一个本原多项式相伴(因为$\frac{m}{d}f(x)$ 就是一个本原多项式). 现在我们要进一步指出,与 $f(x)$ 相伴的本原多项式在相差一个正负号下是唯一的. 即如果 $g(x)$ 和 $g_1(x)$ 都是与 $f(x)$ 相伴的本原多项式,则 $g(x) = \pm g_1(x)$. 证明如下:

由已知条件可设

$$f(x) = rg(x) = r_1g_1(x)$$

其中 $r, r_1 \in Q$,且 $r \neq 0, r_1 \neq 0$. 于是

$$\frac{r}{r_1}g(x) = g_1(x)$$

假如 $\frac{r}{r_1} \neq \pm 1$,则可设$\frac{r}{r_1} = \frac{q}{p}$,其中$(p,q) = 1$,并且 $p,q$ 中至少有一个 不等于 $\pm 1$. 由于 $g(x)$ 与 $g_1(x)$ 的地位对称,所以不妨设 $p \neq \pm 1$. 因为 $g_1(x)$ 是整系数多项式,所以$\frac{q}{p}b_i$ 为整数,其中 $b_i (i = 0, 1, \cdots, n)$ 是 $g(x)$ 的各项系数. 于是 $p | qb_i$. 由于$(p,q) = 1$,所以$p | b_i, i = 0, 1, \cdots, n$. 这与$g(x)$ 的各项系数的最大公因数为 $\pm 1$ 矛盾. 因此$\frac{r}{r_1} = \pm 1$,即 $r = \pm r_1, g(x) = \pm g_1(x)$.

上述议论说明,研究有理系数多项式是否不可约的问题可以归结为研究本原多项式在有理数域上是否不可约的问题. 首先我们指出本原多项式的一个重要性质:

**高斯(Gauss)引理** 两个本原多项式的乘积还是本原多

项式.

**证明** 设

$$f(x)=a_nx^n+\cdots+a_1x+a_0$$
$$g(x)=b_mx^m+\cdots+b_1x+b_0$$

是两个本原多项式.设

$$h(x)=f(x)g(x)=c_{n+m}x^{n+m}+\cdots+c_1x+c_0$$

假如 $h(x)$ 不是本原多项式,即 $(c_{n+m},\cdots,c_1,c_0)\neq 1$,则存在一个素数 $p$,使得 $p\mid c_i,i=0,1,\cdots,n+m$. 因为 $f(x)$ 是本原的,所以 $p$ 不能同时整除 $f(x)$ 的每一个系数.设 $a_k$ 是 $a_n,\cdots,a_1,a_0$ 中第一个不能被 $p$ 整除的系数,即

$$p\mid a_n,\quad \cdots,\quad p\mid a_{k+1},\quad p\nmid a_k$$

同样地,因为 $g(x)$ 是本原的,所以可设 $b_l$ 是 $b_m,\cdots,b_1,b_0$ 中第一个不能被 $p$ 整除的系数,即

$$p\mid b_m,\quad \cdots,\quad p\mid b_{l+1},\quad p\nmid b_l$$

由多项式的乘法的定义知,$h(x)$ 的系数 $c_{k+l}$ 为

$$\begin{aligned}c_{k+l}=&a_{k+l}b_0+a_{k+l-1}b_1+\cdots+a_{k+1}b_{l-1}\\&+a_kb_l+a_{k-1}b_{l+1}+\cdots+a_0b_{k+l}\end{aligned}\tag{1}$$

其中若 $i>n$,则令 $a_i=0$;若 $j>m$,则令 $b_j=0$. 由上面的假设,$p$ 整除(1)式右端除去 $a_kb_l$ 以外的每一项,$p$ 又能整除(1)式左端,因此 $p\mid a_kb_l$,由此推出 $p\mid a_k$ 或者 $p\mid b_l$. 这与 $p\nmid a_k$ 且 $p\nmid b_l$ 矛盾.这证明了 $h(x)$ 也是本原多项式. ▌

由数学归纳法立即得出,有限多个本原多项式的乘积仍是本原多项式.

本原多项式是整系数多项式,自然会问:能否把本原多项式 $f(x)$ 在有理数域 $Q$ 上是否不可约的问题转化为研究 $f(x)$ 在整数环 $Z$ 上是否不可约的问题,为此首先需要搞清楚几个概念.

用 $Z[x]$ 表示所有整系数多项式组成的集合,容易看出 $Z[x]$ 对于多项式的减法和乘法都封闭,从而 $Z[x]$ 是 $Q[x]$ 的子环.显

然 $Q[x]$ 的单位元 1 也是 $Z[x]$ 的单位元. 于是由 $Q[x]$ 是整环可以推出 $Z[x]$ 也是整环. 根据可逆元的定义(见习题 7.1 的第 2 题), 容易证明 $Z[x]$ 中的可逆元只有 $\pm 1$.

**定义 2** 设 $f(x), g(x) \in Z[x]$, 如果存在 $Z[x]$ 中一个多项式 $h(x)$, 使得 $f(x) = g(x)h(x)$, 则称 $g(x)$ 整除 $f(x)$, 记作 $g(x) \mid f(x)$. 此时 $g(x)$ 称为 $f(x)$ 在 $Z[x]$ 中的一个**因式**.

**定义 3** 设 $f(x), g(x) \in Z[x]$, 如果 $f(x) \mid g(x)$, 并且有 $g(x) \mid f(x)$, 则称 $f(x)$ 与 $g(x)$ 在 $Z[x]$ **中相伴**.

**命题 7.8.1** 设 $f(x), g(x) \in Z[x]$, 则 $f(x)$ 与 $g(x)$ 在 $Z[x]$ 中相伴的充分必要条件是 $f(x) = \pm g(x)$.

**证明** 充分性是显然的. 下面证必要性. 由已知条件得, 存在 $h_i(x) \in Z[x], i = 1, 2$, 使得

$$f(x) = h_1(x)g(x), \qquad g(x) = h_2(x)f(x)$$

从而有

$$f(x) = h_1(x)h_2(x)f(x) \tag{2}$$

若 $f(x) = 0$, 则 $g(x) = 0$, 结论成立. 若 $f(x) \neq 0$, 由 (2) 得, $1 = h_1(x)h_2(x)$. 由此推出 $h_1(x)$ 是 $Z[x]$ 中的可逆元, 所以 $h_1(x) = \pm 1$. 从而 $f(x) = \pm g(x)$. ▌

**定义 4** 一个整系数多项式 $q(x)$, 如果在 $Z[x]$ 中的因式只有 $\pm 1$(即 $Z[x]$ 中的可逆元) 和 $\pm q(x)$(即 $q(x)$ 的相伴元), 则称 $q(x)$ 是 $Z$ **上的不可约多项式**, 否则称它**在 $Z$ 上可约**.

我们有下面的结果:

**命题 7.8.2** 一个次数大于零的整系数多项式 $q(x)$, 如果在 $Z$ 上不可约, 则它在 $Q$ 上也不可约.

**证明** 假如 $q(x)$ 在 $Q$ 上可约, 则它可以分解成两个次数较低的有理系数多项式的乘积, 即

$$q(x) = h_1(x)h_2(x), \qquad h_i(x) \in Q[x]$$
$$\deg h_i < \deg q, \qquad i = 1, 2$$

设 $g_i(x)$ 是与 $h_i(x)$ 相伴的本原多项式, $q_1(x)$ 是与 $q(x)$ 相伴的本

原多项式，则

$$h_i(x) = r_i g_i(x), \quad r_i \in Q, \quad i = 1,2$$

$$q(x) = m q_1(x), \quad m \in Z$$

于是有 $$m q_1(x) = r_1 r_2 g_1(x) g_2(x)$$

因为 $q_1(x)$ 本原，$g_1(x)g_2(x)$ 也本原，并且它们与同一个多项式相伴，所以

$$q_1(x) = \pm\, g_1(x) g_2(x)$$

不妨设 $q_1(x) = g_1(x)g_2(x)$，于是有

$$q(x) = m g_1(x) g_2(x) \tag{3}$$

这表明 $g_i(x)$ 是 $q(x)$ 在 $Z[x]$ 中的因式. 由于

$$\deg g_i = \deg h_i < \deg q, \quad i = 1,2$$

所以 $$\deg g_i > 0, \quad i = 1,2$$

因此 $$g_i(x) \neq \pm 1, \quad i = 1,2$$

从而 $$g_i(x) \neq \pm\, q(x), \quad i = 1,2$$

这表明 $q(x)$ 在 $Z$ 上可约. 矛盾. ▌

从命题 7.8.2 的证明中我们还可得到下述两个结论：

**命题** 7.8.3 设 $q(x)$ 是一个次数大于零的整系数多项式，如果 $h(x)$ 是 $q(x)$ 在 $Q[x]$ 中的一个因式，则与 $h(x)$ 相伴的本原多项式 $g(x)$ 是 $q(x)$ 在 $Z[x]$ 中的一个因式. ▌

**命题** 7.8.4 如果一个次数大于零的整系数多项式在 $Q$ 上可约，则它一定能分解成两个次数较低的整系数多项式的乘积.

**证明** 从命题 7.8.2 证明中的公式(3)可以看出. ▌

* 命题 7.8.2 给出了整系数多项式在 $Q$ 上不可约的一个充分条件. 对于本原多项式来说，这个条件也是必要条件，即我们有

* **命题** 7.8.5 一个次数大于零的本原多项式在 $Q$ 上不可约的充分必要条件是它在 $Z$ 上不可约.

必要性的证明作为习题.

现在我们给出整系数多项式在有理数域上不可约的一个充分条件：

**定理** 7.8.1(Eisenstein **判别法**) 设

$$f(x)=a_nx^n+\cdots+a_1x+a_0$$

是一个次数 $n$ 大于零的整系数多项式,如果存在一个素数 $p$,使得

1) $p\nmid a_n$;

2) $p\mid a_i, i=0,1,\cdots,n-1$;

3) $p^2\nmid a_0$;

则 $f(x)$ 在有理数域上是不可约的.

**证明** 假如 $f(x)$ 在 $Q$ 上可约,则据命题 7.8.4 得

$$f(x)=(b_mx^m+\cdots+b_1x+b_0)(c_lx^l+\cdots+c_1x+c_0) \quad (4)$$

其中 $b_i, c_j$ 都是整数($i=0,1,\cdots,m; j=0,1,\cdots,l$),$b_m\neq 0, c_l\neq 0$,$m<n, l<n$,并且 $m+l=n$. 因此

$$a_n=b_mc_l,\quad a_0=b_0c_0$$

已知 $p\mid a_0$,所以 $p\mid b_0$ 或者 $p\mid c_0$. 又因为 $p^2\nmid a_0$,所以 $p$ 不能同时整除 $b_0$ 和 $c_0$. 不妨设 $p\mid b_0$ 但 $p\nmid c_0$. 因为 $p\nmid a_n$,所以 $p\nmid b_m$. 假设 $b_0, b_1, \cdots, b_m$ 中第一个不能被 $p$ 整除的是 $b_k$,即

$$p\mid b_0,\quad p\mid b_1,\quad \cdots,\quad p\mid b_{k-1},\quad p\nmid b_k,\quad 0<k\leqslant m \quad (5)$$

比较(4)式两边 $x_k$ 的系数,得

$$a_k=b_kc_0+b_{k-1}c_1+\cdots+b_0c_k \quad (6)$$

因为 $k\leqslant m<n$,所以 $p\mid a_k$. 于是从(5)和(6)式得,$p\mid b_kc_0$. 由于 $p\nmid b_k$,从而 $p\mid c_0$,矛盾. 所以 $f(x)$ 在 $Q$ 上不可约. ▌

利用定理 7.8.1 我们可以得到

**定理** 7.8.2 在 $Q[x]$ 中存在任意次数的不可约多项式.

**证明** 对任意的正整数 $n$,设 $f(x)=x^n+2$. 素数 2 符合定理 7.8.1 的条件,因此 $x^n+2$ 在 $Q$ 上不可约. ▌

有时直接用 Eisenstein 判别法无法判断 $f(x)$ 是否在 $Q$ 上不可约,我们可以通过不定元 $x$ 用 $Q[x]$ 中的元素 $x+b$ 代入,得到另一个多项式

$$g(x):=f(x+b)$$

对 $g(x)$ 可能用 Eisenstein 判别法能判断它不可约，这时原来的 $f(x)$ 是否不可约？下面的命题回答了这个问题：

**命题** 7.8.6　设 $f(x)=a_nx^n+\cdots+a_1x+a_0$ 是次数 $n$ 大于零的整系数多项式，设 $b$ 是任意给定的一个整数. 令

$$\begin{aligned} g(x):&=f(x+b) \\ &=a_n(x+b)^n+\cdots+a_1(x+b)+a_0 \end{aligned}$$

则 $f(x)$ 在 $Q$ 上不可约的充分必要条件是 $g(x)$ 在 $Q$ 上不可约.

**证明**　先证充分性. 因为 $g(x)$ 的首项为 $a_nx^n$，所以

$$\deg g(x)=n=\deg f(x)$$

假如 $f(x)$ 在 $Q$ 上可约，则

$$f(x)=f_1(x)f_2(x),\qquad \deg f_i<\deg f,\quad i=1,2 \tag{7}$$

$x$ 用 $x+b$ 代入，从(7)式得

$$f(x+b)=f_1(x+b)f_2(x+b) \tag{8}$$

(8)式表明 $g(x)$ 在 $Q$ 上可约.

再证必要性. $x$ 用 $x-b$ 代入，则从

$$g(x)=f(x+b)$$

得

$$g(x-b)=f(x)$$

从上面证得的结论知，如果 $f(x)$ 在 $Q$ 上不可约，则 $g(x)$ 在 $Q$ 上也不可约. ▌

**例** 1　设 $p$ 是一个素数，多项式

$$f(x)=x^{p-1}+x^{p-2}+\cdots+x+1$$

称为一个分圆多项式. 证明 $f(x)$ 在 $Q$ 上不可约.

**证明**　我们有

$$\begin{aligned} (x-1)f(x)&=(x-1)(x^{p-1}+x^{p-2}+\cdots+x+1) \\ &=x^p-1 \end{aligned} \tag{9}$$

$x$ 用 $x+1$ 代入，从(9)式得

$$xf(x+1)=(x+1)^p-1 \tag{10}$$

从(10)式得

$$f(x+1) \doteq x^{p-1} + px^{p-2} + \cdots + C_p^k x^{p-k-1} + \cdots + p$$

令 $$g(x) := f(x+1)$$
$$= x^{p-1} + px^{p-2} + \cdots + C_p^k x^{p-k-1} + \cdots + p$$

因为 $$C_p^k = \frac{p(p-1)\cdots(p-k+1)}{k!}, \quad 1 \leqslant k < p$$

并且 $(p, k!) = 1$，所以

$$k! \mid (p-1)\cdots(p-k+1)$$

从而 $$p \mid C_p^k, \quad 1 \leqslant k < p$$

于是对于 $g(x)$，素数 $p$ 满足 Eisenstein 判别法的条件，所以 $g(x)$ 在 $Q$ 上不可约. 从而 $f(x)$ 在 $Q$ 上不可约. ▌

对于有些整系数多项式，即使用命题 7.8.6，仍然不能用 Eisenstein 判别法判别其是否不可约. 对于二次或三次整系数多项式 $f(x)$，如果它可约，则它一定有一次因式，从而它必有有理根. 如果我们会求整系数多项式的有理根，那么就可以解决二次或三次整系数多项式是否不可约的问题.

利用命题 7.8.3 我们可以得到求整系数多项式的全部有理根的方法：

**定理** 7.8.3 设

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

是一个次数 $n$ 大于零的整系数多项式. 如果 $\frac{q}{p}$ 是 $f(x)$ 的一个有理根，其中 $p, q$ 是互素的整数，那么 $p \mid a_n, q \mid a_0$.

**证明** 由所设，$\frac{q}{p}$ 是 $f(x)$ 的一个有理根，因此 $x - \frac{q}{p}$ 是 $f(x)$ 在 $Q[x]$ 中的一个因式. 因为 $(p, q) = 1$，所是 $px - q$ 是与 $x - \frac{q}{p}$ 相伴的一个本原多项式. 据命题 7.8.3，$px - q$ 是 $f(x)$ 在 $Z[x]$ 中的一个因式. 由定义 2，存在一个整系数多项式

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

使得

$$f(x) = (px - q)g(x) \tag{11}$$

比较(11)式两边多项式的首项系数与常数项得

$$a_n = pb_{n-1}, \quad a_0 = -qb_0$$

因此 $$p \mid a_n, \quad q \mid a_0 \qquad \blacksquare$$

定理 7.8.3 的证明中表明，如果 $\frac{q}{p}$ 是一个次数大于零的整系数多项式 $f(x)$ 的有理根，并且 $(p,q)=1$，那么存在一个整系数多项式 $g(x)$，使得

$$f(x) = (px - q)g(x)$$

当有理根 $\frac{q}{p} \neq \pm 1$ 时，$x$ 分别用 $\pm 1$ 代入，由上式得

$$f(1) = (p - q)g(1), \qquad f(-1) = -(p + q)g(-1)$$

由于 $g(x)$ 是整系数多项式，因此 $\frac{f(1)}{p-q}$，$\frac{f(-1)}{p+q}$ 都是整数，这一结论在求整系数多项式的有理根时有用，可从下面的例子中看到.

**例** 2 求 $$f(x) = 3x^4 + 8x^3 + 6x^2 + 3x - 2$$
的全部有理根.

**解** $a_n = 3$ 的因子只有 $\pm 1, \pm 3$；$a_0 = -2$ 的因子只有 $\pm 1$，$\pm 2$. 据定理 7.8.3，$f(x)$ 的有理根只可能是：

$$\pm 1, \quad \pm 2, \quad \pm \frac{1}{3}, \quad \pm \frac{2}{3}$$

因为 $f(1) = 18 \neq 0$，$f(-1) = -4 \neq 0$
所以 $\pm 1$ 不是 $f(x)$ 的根.

考虑 2，因为

$$\frac{f(-1)}{p+q} = -\frac{4}{3}$$

不是整数，所以 2 不是 $f(x)$ 的根.

考虑 $-2$，因为

$$\frac{f(1)}{p-q} = \frac{18}{3} = 6, \qquad \frac{f(-1)}{p+q} = \frac{-4}{-1} = 4$$

所以需要进一步用综合除法来判断 $-2$ 是不是 $f(x)$ 的根：

$$\begin{array}{ccccc|c}
3 & 8 & 6 & 3 & -2 & -2 \\
 & -6 & -4 & -4 & 2 & \\
\hline
3 & 2 & 2 & -1 & 0 & \\
 & -6 & 8 & -20 & & \\
\hline
3 & -4 & 10 & -21 & &
\end{array}$$

这表明 $-2$ 是 $f(x)$ 的单根. 于是

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

考虑 $\frac{1}{3}$,因为

$$\frac{f(1)}{p-q}=9,\qquad \frac{f(-1)}{p+q}=-1$$

所以需作综合除法:

$$\begin{array}{cccc|c}
3 & 2 & 2 & -1 & \frac{1}{3} \\
 & 1 & 1 & 1 & \\
\hline
3 & 3 & 3 & 0 & \\
 & 1 & \frac{4}{3} & & \\
\hline
3 & 4 & \frac{13}{3} & &
\end{array}$$

这表明 $\frac{1}{3}$ 是 $f(x)$ 的单根. 于是

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

显然 $x^2+x+1$ 没有有理根. 因此 $f(x)$ 的全部有理根是 $-2$, $\frac{1}{3}$,它们都是单根.

**例 3** 证明 $f(x)=x^3+2x+1$ 在有理数域上不可约.

**证明** 如果 $f(x)$ 在 $Q$ 上可约,则 $f(x)$ 可以分解成两个次数较低的有理系数多项式的乘积. 由于 $f(x)$ 的次数是 3,所以 $f(x)$ 在 $Q[x]$ 中至少有一个一次因式,从而 $f(x)$ 至少有一个有理根. 因

为 $a_n=1, a_0=1$，所以 $f(x)$ 的有理根只可能是 $\pm 1$. 但

$$f(1)=4\neq 0,\quad f(-1)=-2\neq 0$$

因此 $\pm 1$ 不是 $f(x)$ 的根. 所以 $f(x)$ 在 $Q$ 上不可约. ▌

注意：如果整系数多项式 $f(x)$ 的次数大于 3，那么不能从 $f(x)$ 没有有理根便得出 $f(x)$ 不可约的结论. 这是因为 $f(x)$ 没有有理根，只是说明 $f(x)$ 没有一次因式，但是 $f(x)$ 可能有次数大于 1 的因式，从而 $f(x)$ 可能可约.

## 习 题 7.8

1. 求下列多项式的全部有理根：

(1) $2x^3+x^2-3x+1$；

(2) $2x^4-x^3-19x^2+9x+9$.

2. 下列整系数多项式在有理数域上是否可约？

(1) $x^4-6x^3+2x^2+10$；　　(2) $7x^5+18x^4+6x-6$；

(3) $x^5+5x^3+1$；　　(4) $x^4-2x^3+2x-3$；

(5) $x^p+px^2+1$，$p$ 为奇素数；

(6) $x^3+x^2-3x+2$；

(7) $2x^3-x^2+x+1$；　　(8) $x^4-5x+1$.

3. 设 $n>1$，证明：$n$ 个互不相同的素数的几何平均数一定是无理数.

4. 设 $f(x)=a_nx^n+\cdots+a_1x+a_0$ 是一个次数大于零的整系数多项式，证明：如果 $a_n+a_{n-1}+\cdots+a_1+a_0$ 是一个奇数，则 1 和 $-1$ 都不是 $f(x)$ 的根.

5. 设 $f(x)$ 是一个次数大于零且首项系数为 1 的整系数多项式，证明：如果 $f(0)$ 与 $f(1)$ 都是奇数，则 $f(x)$ 没有有理根.

6. 设 $f(x)=x^3+ax^2+bx+c$ 是整系数多项式，证明：如果 $(a+b)c$ 是奇数，则 $f(x)$ 在有理数域上不可约.

7. 设 $a_1,a_2,\cdots,a_n$ 是 $n$ 个不同的整数，设

$$f(x)=(x-a_1)(x-a_2)\cdots(x-a_n)+1$$

证明:如果 $n$ 是奇数,则 $f(x)$ 在有理数域上不可约.如果 $n$ 是偶数,$f(x)$ 是否在有理数域上不可约?

(提示:用命题 7.8.4 以及推论 7.6.1.)

*8. 设 $a_1,a_2,\cdots,a_n$ 是 $n$ 个不同的整数,设

$$f(x)=(x-a_1)(x-a_2)\cdots(x-a_n)-1$$

证明 $f(x)$ 在有理数域上不可约.

(提示:分情形讨论,除一种情形外,都类似于第 6 题的方法.)

*9. 设 $a_1,a_2,\cdots,a_n$ 是 $n$ 个不同整数,设

$$f(x)=\prod_{i=1}^{n}(x-a_i)^2+1$$

证明 $f(x)$ 在 $Q$ 上不可约.

10. 证明:$Z[x]$ 中的可逆元只有 $\pm 1$.

*11. 证明:一个本原多项式如果在 $Q$ 上不可约,则它在 $Z$ 上也不可约.

*12. 证明:一个次数大于零的整系数多项式如果在 $Z$ 上不可约,则它一定是本原的.

## §9 插值法

在 §6 我们运用函数论知识证明了代数基本定理,进而决定了复数域上的所有不可约多项式.现在反过来我们要运用多项式的理论来解决函数论中的一些问题,其中要用到推论 7.6.1:$K[x]$ 中两个次数不超过 $n$ 的多项式 $f(x)$ 与 $g(x)$,如果它们在 $K$ 的 $n+1$ 个不同元素 $a_1,\cdots,a_{n+1}$ 上有

$$f(a_i)=g(a_i),\quad i=1,\cdots,n+1$$

则这两个多项式相等.这说明:数域 $K$ 上一个次数不超过 $n$ 的多项式,被它在 $K$ 的 $n+1$ 个不同元素上的值所唯一确定.

在实际问题中,我们常常遇到要研究两个变量 $y$ 与 $x$ 之间的依赖关系.通过实验或观测可得到这两个变量对应数值:$x$ 取值 $c_i$ 时,$y$ 取值 $d_i$,$i=0,1,2,\cdots,n$.我们希望找一个函数 $y=f(x)$,使

得 $f(c_i)=d_i, i=0,1,2,\cdots,n$，并且它能尽量准确地反映 $y$ 与 $x$ 之间的依赖关系，而计算又比较简单. 变量 $y$ 与 $x$ 之间的依赖关系是客观存在的，设为 $y=\varphi(x)$. 所找的函数 $f(x)$ 称为 $\varphi(x)$ 的一个**插值函数**. 求插值函数的问题称为**插值问题**，求插值函数的方法称为**插值法**.

实际问题中，如果我们用某种办法已经知道变量 $y$ 与 $x$ 之的依赖关系是一个很"光滑"的函数，那么我们常常取多项式 $f(x)$ 确定的函数作为这个函数的插值函数，此时 $f(x)$ 称为**插值多项式**. 现在的插值问题是：设 $c_0,c_1,\cdots,c_n$ 是 $K$ 的 $n+1$ 个不同元素，$d_0,d_1,\cdots,d_n$ 是 $K$ 的 $n+1$ 个元素. 我们要在 $K[x]$ 中找一个次数 $\leqslant n$ 的多项式 $f(x)$，使得 $f(c_i)=d_i, i=0,1,\cdots,n$. 据上一段提到的推论 7.6.1，如果这个问题的解存在，则它是唯一的. 现在我们来证明插值问题的解一定存在.

**定理** 7.9.1　设 $c_0,c_1,\cdots,c_n$ 是数域 $K$ 的 $n+1$ 个不同的元素，$d_0,d_1,\cdots,d_n$ 是 $K$ 的 $n+1$ 个元素，则 $K[x]$ 中存在唯一的一个次数不超过 $n$ 的多项式 $f(x)$，使得

$$f(c_i)=d_i,\qquad i=0,1,\cdots,n$$

**证明**　上面已指出，据推论 7.6.1，如果这个插值问题的解存在，则它是唯一的. 现在来证存在性.

先看一个特殊情形：

$$d_0=\cdots=d_{i-1}=d_{i+1}=\cdots=d_n=0$$

如果存在一个次数 $\leqslant n$ 的多项式 $f_i(x)$，使得

$$f_i(c_j)=d_j,\quad j=0,1,\cdots,n$$

则

$$f_i(c_0)=0,\quad \cdots,\quad f_i(c_{i-1})=0$$

$$f_i(c_{i+1})=0,\quad \cdots,\quad f_i(c_n)=0$$

这表明 $c_0,\cdots,c_{i-1},c_{i+1},\cdots,c_n$ 是 $f_i(x)$ 的 $n$ 个不同的根. 由于 $\deg f_i\leqslant n$，所以

$$f_i(x)=a_i(x-c_0)\cdots(x-c_{i-1})(x-c_{i+1})\cdots(x-c_n)\quad (1)$$

因为 $f_i(c_i)=d_i$，所以由(1) 式得

$$a_i = \frac{d_i}{(c_i - c_0)\cdots(c_i - c_{i-1})(c_i - c_{i+1})\cdots(c_i - c_n)} \tag{2}$$

从而得

$$f_i(x) = d_i \frac{(x - c_0)\cdots(x - c_{i-1})(x - c_{i+1})\cdots(x - c_n)}{(c_i - c_0)\cdots(c_i - c_{i-1})(c_i - c_{i+1})\cdots(c_i - c_n)} \tag{3}$$

现在看一般情形. 令

$$\begin{aligned} f(x) &= \sum_{i=0}^{n} f_i(x) \\ &= \sum_{i=0}^{n} d_i \frac{(x - c_0)\cdots(x - c_{i-1})(x - c_{i+1})\cdots(x - c_n)}{(c_i - c_0)\cdots(c_i - c_{i-1})(c_i - c_{i+1})\cdots(c_i - c_n)} \end{aligned} \tag{4}$$

则 $\deg f(x) \leqslant n$，并且对任意 $j(j = 0,1,\cdots,n)$ 有

$$f(c_j) = \sum_{i=0}^{n} d_i \frac{(c_j - c_0)\cdots(c_j - c_{i-1})(c_j - c_{i+1})\cdots(c_j - c_n)}{(c_i - c_0)\cdots(c_i - c_{i-1})(c_i - c_{i+1})\cdots(c_i - c_n)} = d_j \tag{5}$$

所以由(4)式给出的 $f(x)$ 是插值问题的解. ▌

公式(4)称为**拉格朗日**(Lagrange)**插值公式**.

定理 7.9.1 也可以采用**待定系数法**来证，设

$$f(x) = a_n x^n + \cdots + a_1 x + a_0$$

是插值问题的解，由 $f(c_i) = d_i, i = 0,1,\cdots,n$，可以得到一个含 $n+1$ 个未知量 $a_0, a_1, \cdots, a_n$ 的由 $n+1$ 个方程组成的线性方程组，它的系数行列式是范德蒙行列式，由于 $c_0, c_1, \cdots, c_n$ 是 $n+1$ 个不同的数，所以这个范德蒙行列式不等于零. 据 Cramer 法则，该方程组有唯一解，并且解有一个公式表示. (这个解也可以用消元法求出). 但是拉格朗日插值公式更为方便，因为它既简单又容易记.

上述插值问题的解还可以用**牛顿**(Newton)**插值公式**给出：

$$\begin{aligned} f(x) = & u_0 + u_1(x - c_0) + u_2(x - c_0)(x - c_1) + \cdots \\ & + u_n(x - c_0)(x - c_1)\cdots(x - c_{n-1}) \end{aligned} \tag{6}$$

其中系数 $u_0, u_1, \cdots, u_n$ 可以通过把 $x$ 逐次用 $c_0, c_1, \cdots, c_n$ 代入而从(6)式求出.

**例 1** 求一个次数不超过 3 的多项式，使

$$f(0)=5,\quad f(1)=7,\quad f(-1)=9,\quad f(-2)=13$$

**解法一** 用拉格朗日插值公式得

$$\begin{aligned}
f(x) =& 5\frac{(x-1)(x+1)(x+2)}{(0-1)(0+1)(0+2)} \\
&+7\frac{(x-0)(x+1)(x+2)}{(1-0)(1+1)(1+2)} \\
&+9\frac{(x-0)(x-1)(x+2)}{(-1-0)(-1-1)(-1+2)} \\
&+13\frac{(x-0)(x-1)(x+1)}{(-2-0)(-2-1)(-2+1)} \\
=& x^3+3x^2-2x+5
\end{aligned}$$

**解法二** 用牛顿插值公式．设

$$f(x)=u_0+u_1x+u_2x(x-1)+u_3x(x-1)(x+1)$$

因为 $f(0)=5$，所以 $u_0=5$. 因为 $f(1)=7$，所以 $7=5+u_1$. 于是 $u_1=2$. 因为 $f(-1)=9$，所以 $9=5-2+2u_2$. 于是得 $u_2=3$. 因为 $f(-2)=13$，所以 $13=5-4+18-6u_3$. 于是得 $u_3=1$. 所以

$$\begin{aligned}
f(x) &= 5+2x+3x(x-1)+x(x-1)(x+1) \\
&= x^3+3x^2-2x+5
\end{aligned}$$

**解法三** 设

$$f(x)=a_3x^3+a_2x^2+a_1x+a_0$$

用待定系数法．细节留给读者．

从例 1 的三种解法看出，用牛顿插值公式最简便．

## 习 题 7.9

1. 求一个次数不超过 3 的多项式 $f(x)$，使

$$f(1)=2,\quad f(2)=-3,\quad f(3)=1,\quad f(4)=3$$

然后求 $f\left(\frac{3}{2}\right), f\left(\frac{7}{2}\right)$.

2. 求一个次数尽可能低的多项式 $f(x)$，使

$$f(0)=3,\quad f(1)=4,\quad f(2)=9,\quad f(3)=18$$

然后求 $f\left(\frac{1}{2}\right), f\left(\frac{5}{2}\right)$.

3. 设 $c_0, c_1, c_2, \cdots, c_n$ 是数域 $K$ 的 $n+1$ 个不同的元素，令

$$F(x)=(x-c_0)(x-c_1)\cdots(x-c_n)$$

把拉格朗日插值公式用 $F(x)$ 以及它的导数 $F'(x)$ 在 $c_i$ 处的值 $F'(c_i)(i=0, 1, \cdots, n)$ 来表示.

4. 求所有 4 次多项式，使它在 $0,1,2,\cdots,n,\cdots$ 的值都是整数.

(提示：用牛顿插值公式.)

## §10 多元多项式环

迄今为止我们讨论了数域 $K$ 上一元多项式环的性质. 这节我们将讨论数域 $K$ 上多元多项式环的基本性质. 多元多项式是代数几何的基本研究对象.

**定义 1** 设 $K$ 是一个数域，用 $n$ 个不属于 $K$ 的符号 $x_1, x_2, \cdots, x_n$ 作表达式

$$\sum_{i_1, i_2, \cdots, i_n} a_{i_1 i_2 \cdots i_n} x_1^{i_1} x_2^{i_2} \cdots x_n^{i_n} \tag{1}$$

其中 $a_{i_1 i_2 \cdots i_n} \in K, i_1, i_2, \cdots, i_n$ 是非负整数；(1) 式中的每一项

$$a_{i_1 i_2 \cdots i_n} x_1^{i_1} x_2^{i_2} \cdots x_n^{i_n}$$

称为一个**单项式**，$a_{i_1 i_2 \cdots i_n}$ 称为这个单项式的**系数**；系数为零的单项式允许任意删去和添进来，(1) 式中只有有限多个单项式的系数不等于零；如果两个单项式中 $x_j (j=1,2,\cdots,n)$ 的幂指数都对应相同，则称它们为**同类项**；我们约定表达式(1) 中的单项式都是不同类的；两个形如(1) 式的表达式相等规定为它们的同类项的系数都相等(即它们含有完全相同的单项式)，则形如(1) 的表达式

称为**系数在数域 $K$ 中的 $n$ 元多项式**，或者简称为**数域 $K$ 上的 $n$ 元多项式**，符号 $x_1,x_2,\cdots,x_n$ 称为**数域 $K$ 上的 $n$ 个无关不定元**.

数域 $K$ 上的一个 $n$ 元多项式如果它的所有系数全为零，则称它为**零多项式**，记为 0.

我们常用 $f(x_1,x_2,\cdots,x_n),g(x_1,x_2,\cdots,x_n),\cdots$ 或 $f,g,\cdots$ 等来代表 $n$ 元多项式.

非负整数 $i_1+i_2+\cdots+i_n$ 称为单项式 $a_{i_1i_2\cdots i_n}x_1^{i_1}x_2^{i_2}\cdots x_n^{i_n}$ 的**全次数**(或者简称为**次数**). 一个 $n$ 元多项式 $f(x_1,x_2,\cdots,x_n)$ 的系数不为零的单项式的全次数的最大值称为这个**多项式 $f$ 的全次数**(或者简称为**次数**)，用 $\deg f$ 来表示. 零多项式的全次数(简称为次数)规定为 $-\infty$. 例如多项式

$$5x_1^4+3x_1^3x_2+2x_1x_2x_3^2+x_2^3+x_2x_3$$

的次数是 4. 从这个例子看出，当 $n>1$ 时，$n$ 元多项式 $f(x_1,x_2,\cdots,x_n)$ 可能有许多单项式具有相同的最大次数，因此我们不能把次数最大的单项式称为 $f(x_1,x_2,\cdots,x_n)$ 的首项. 首项的概念在后面用另外的办法给出.

数域 $K$ 上所有 $n$ 元多项式组成的集合记作 $K[x_1,x_2,\cdots,x_n]$. 在 $K[x_1,x_2,\cdots,x_n]$ 中定义加法与乘法如下：

$$\sum_{i_1,i_2,\cdots,i_n}a_{i_1i_2\cdots i_n}x_1^{i_1}x_2^{i_2}\cdots x_n^{i_n}+\sum_{i_1,i_2,\cdots,i_n}b_{i_1i_2\cdots i_n}x_1^{i_1}x_2^{i_2}\cdots x_n^{i_n}$$

$$:=\sum_{i_1,i_2,\cdots,i_n}(a_{i_1i_2\cdots i_n}+b_{i_1i_2\cdots i_n})x_1^{i_1}x_2^{i_2}\cdots x_n^{i_n}\tag{2}$$

$$\Big(\sum_{i_1,i_2,\cdots,i_n}a_{i_1i_2\cdots i_n}x_1^{i_1}x_2^{i_2}\cdots x_n^{i_n}\Big)\Big(\sum_{j_1,j_2,\cdots,j_n}b_{j_1j_2\cdots j_n}x_1^{j_1}x_2^{j_2}\cdots x_n^{j_n}\Big)$$

$$:=\sum_{s_1,s_2,\cdots,s_n}c_{s_1s_2\cdots s_n}x_1^{s_1}x_2^{s_2}\cdots x_n^{s_n}\tag{3}$$

其中
$$c_{s_1s_2\cdots s_n}=\sum_{i_1+j_1=s_1}\ \sum_{i_2+j_2=s_2}\cdots\sum_{i_n+j_n=s_n}a_{i_1i_2\cdots i_n}b_{j_1j_2\cdots j_n}\tag{4}$$

不难验证 $K[x_1,x_2,\cdots,x_n]$ 对于如上定义的加法与乘法成为一个环. 它的零元素是零多项式. 它有单位元素，即零次多项式 1. 它是

交换环,这个环称为**数域 $K$ 上的 $n$ 元多项式环**.

显然有

$$\deg(f+g) \leqslant \max\{\deg f, \deg g\}$$

稍后我们将讨论 $\deg fg$ 与 $\deg f$,$\deg g$ 的关系. 先来讨论 $K[x_1, x_2, \cdots, x_n]$ 是不是无零因子环,为此我们需要对 $n$ 元多项式 $f(x_1, x_2, \cdots, x_n)$ 的各项规定一个排列顺序,从而给出首项的概念.

每一类单项式 $ax_1^{i_1}x_2^{i_2}\cdots x_n^{i_n}$ 对应一个 $n$ 元有序非负整数组 $(i_1, i_2, \cdots, i_n)$,这个对应是一一对应. 为了给出各类单项式之间的一个排列顺序的方法,就只要对于 $n$ 元有序非负整数组定义一个先后顺序就行了.

对于两个 $n$ 元有序非负整数组 $(i_1, i_2, \cdots, i_n)$ 与 $(j_1, j_2, \cdots, j_n)$,如果 $i_1 = j_1, i_2 = j_2, \cdots, i_{s-1} = j_{s-1}, i_s > j_s (1 \leqslant s \leqslant n)$,则称 $(i_1, i_2, \cdots, i_n)$ **先于** $(j_1, j_2, \cdots, j_n)$,记作

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

由上述定义立即看出,对于任意两个 $n$ 元有序非负整数组 $(i_1, i_2, \cdots, i_n)$,$(j_1, j_2, \cdots, j_n)$,关系

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

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

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

中有一个且仅有一个成立.

关系"$>$"具有传递性,即如果

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

并且 $$(j_1, j_2, \cdots, j_n) > (k_1, k_2, \cdots, k_n)$$

则 $$(i_1, i_2, \cdots, i_n) > (k_1, k_2, \cdots, k_n)$$

这是因为 $i_l - k_l = (i_l - j_l) + (j_l - k_l)$,由此可得结论. 例如,

$$(4,2,3,3) > (4,2,2,4), \qquad (4,2,2,4) > (4,1,4,3)$$

由传递性得 $(4,2,3,3) > (4,1,4,3)$. 直接用定义也可看出 $(4,2,3,3) > (4,1,4,3)$.

这样我们的确给出了 $n$ 元有序非负整数组之间的一个顺序. 相应地, $n$ 元各类单项式之间也有了一个先后顺序. 这种排列顺序的方法是模仿字典中单词的排列原则得出的, 因而称之为**字典排列法**.

例如, 多项式

$$2x_1^4x_2x_3 + x_1x_2^5x_3 + 6x_1^3$$

按字典排列法写出来就是

$$2x_1^4x_2x_3 + 6x_1^3 + x_1x_2^5x_3$$

按字典排列法写出来的第一个系数不为零的单项式称为 $n$ 元多项式的**首项**. 例如, 上面的多项式的首项是 $2x_1^4x_2x_3$. 要注意, 首项不一定具有最大的次数. 例如, 上面多项式的次数是 7, 而首项的次数是 6.

**定理** 7.10.1 在 $K[x_1,x_2,\cdots,x_n]$ 中两个非零多项式的乘积的首项等于它们的首项的乘积. 从而两个非零多项式的乘积仍是非零多项式, 即 $K[x_1,x_2,\cdots,x_n]$ 是无零因子环.

**证明** 设 $f(x_1,x_2,\cdots,x_n)$, $g(x_1,x_2,\cdots,x_n)$ 是 $K[x_1,x_2,\cdots,x_n]$ 中两个非零多项式. 设 $f(x_1,x_2,\cdots,x_n)$ 的首项为 $ax_1^{p_1}x_2^{p_2}\cdots x_n^{p_n}$, $a\neq 0$. 设 $g(x_1,x_2,\cdots,x_n)$ 的首项是 $bx_1^{q_1}x_2^{q_2}\cdots x_n^{q_n}$, $b\neq 0$. 为了证明 $fg$ 的首项是 $abx_1^{p_1+q_1}x_2^{p_2+q_2}\cdots x_n^{p_n+q_n}$, 只要证明

$$(p_1+q_1,p_2+q_2,\cdots,p_n+q_n)$$

先于 $fg$ 中其他单项式的幂指数组就行了. $fg$ 的其他单项式的幂指数组只有三种可能情形.

$$(p_1+j_1,p_2+j_2,\cdots,p_n+j_n)$$

或者 $\quad (i_1+q_1,i_2+q_2,\cdots,i_n+q_n)$

或者 $\quad (i_1+j_1,i_2+j_2,\cdots,i_n+j_n)$

其中 $\quad (p_1,p_2,\cdots,p_n) > (i_1,i_2,\cdots,i_n)$

$$(q_1,q_2,\cdots,q_n) > (j_1,j_2,\cdots,j_n)$$

显然有

$$(p_1+q_1,p_2+q_2,\cdots,p_n+q_n)>(p_1+j_1,p_2+j_2,\cdots,p_n+j_n)$$
$$(p_1+q_1,p_2+q_2,\cdots,p_n+q_n)>(i_1+q_1,i_2+q_2,\cdots,i_n+q_n)$$
$$(i_1+q_1,i_2+q_2,\cdots,i_n+q_n)>(i_1+j_1,i_2+j_2,\cdots,i_n+j_n)$$

由传递性得

$$(p_1+q_1,p_2+q_2,\cdots,p_n+q_n)>(i_1+j_1,i_2+j_2,\cdots,i_n+j_n)$$

这就证明了 $abx_1^{p_1+q_1}x_2^{p_2+q_2}\cdots x_n^{p_n+q_n}$ 不可能与 $fg$ 中其他的单项式相消，而且它先于 $fg$ 其他的单项式，因此它是 $fg$ 的首项. ▌

因为 $K[x_1,x_2,\cdots,x_n]$ 是无零因子环，所以消去律成立.

用数学归纳法可得出

**推论** 7.10.1　在 $K[x_1,x_2,\cdots,x_n]$ 中，如果 $f_i\neq 0,i=1,2,\cdots,m$，则 $f_1f_2\cdots f_m$ 的首项等于每个 $f_i$ 的首项的乘积. ▌

我们要引进齐次多项式这一重要概念：

**定义** 2　数域 $K$ 上的 $n$ 元多项式 $g(x_1,x_2,\cdots,x_n)$ 称为 $m$ **次齐次多项式**，如果它的每个单项式都是 $m$ 次的.

例如，$g(x_1,x_2,x_3)=2x_1^4+3x_1^2x_2x_3+x_1x_2x_3^2$ 是一个 4 次齐次多项式.

显然，$K[x_1,x_2,\cdots,x_n]$ 中两个齐次多项式的乘积仍是齐次多项式，它的次数等于这两个多项式的次数的和.

对于任何一个 $n$ 元多项式 $f(x_1,x_2,\cdots,x_n)$，如果把 $f$ 中所有次数相同的单项式并在一起，则 $f$ 可以唯一地表示成

$$f(x_1,x_2,\cdots,x_n)=\sum_{i=0}^{m}f_i(x_1,x_2,\cdots,x_n),\quad m=\deg f \tag{5}$$

其中 $f_i(x_1,x_2,\cdots,x_n)$ 是零多项式或 $i$ 次齐次多项式，它称为 $f(x_1,x_2,\cdots,x_n)$ 的 $i$ **次齐次成分**.

利用(5)式我们可以证明下述定理.

**定理** 7.10.2　设

$$f(x_1,x_2,\cdots,x_n),\quad g(x_1,x_2,\cdots,x_n)\in K[x_1,x_2,\cdots,x_n]$$

则
$$\deg fg=\deg f+\deg g \tag{6}$$

**证明**　若 $f,g$ 中有一个是零多项式，则(6) 式成立.

现在设

$$f \neq 0, \quad g \neq 0, \quad \deg f = m, \quad \deg g = s$$

则 $\quad f = f_0 + f_1 + \cdots + f_m, \quad g = g_0 + g_1 + \cdots + g_s$

其中 $f_i$ 是 $f$ 的 $i$ 次齐次成分，$g_j$ 是 $g$ 的 $j$ 次齐次成分，$f_m \neq 0$，$g_s \neq 0$，我们有

$$\begin{aligned} fg = & f_0g_0 + \cdots + f_0g_s + f_1g_0 + \cdots + f_1g_s \\ & + \cdots + f_mg_0 + \cdots + f_mg_s \end{aligned} \tag{7}$$

其中 $f_ig_j$ 是零多项式或 $i+j$ 次齐次多项式（$i=0,1,\cdots,m$；$j=0,1,\cdots,s$）. 因为 $f_m \neq 0, g_s \neq 0$，所以 $f_mg_s \neq 0$. 于是 $f_mg_s$ 是 $m+s$ 次齐次多项式. 所以

$$\deg fg = m + s = \deg f + \deg g$$ ▌

* 当 $n > 1$ 时，$K[x_1, x_2, \cdots, x_n]$ 中没有带余除法，但是唯一因式分解定理仍然成立. 我们将在抽象代数课程中讨论这方面的问题.

在本章 §1.3 我们指出，数域 $K$ 上的一元多项式环 $K[x]$ 有通用性质. 现在我们要指出，数域 $K$ 上的 $n$ 元多项式环 $K[x_1, x_2, \cdots, x_n]$ 也具有通用性质，即

**定理** 7.10.3 设 $K$ 是一个数域，设 $R$ 是一个交换环，并且 $R$ 可以看成是 $K$ 的一个扩环（即 $R$ 有一个子环与 $K$ 同构）. 设 $t_1, \cdots, t_n$ 是 $R$ 的元素，令

$$\begin{aligned} \sigma_{t_1,\cdots,t_n}: K[x_1, x_2, \cdots, x_n] & \longrightarrow R \\ f(x_1, x_2, \cdots, x_n) & \longmapsto f(t_1, t_2, \cdots, t_n) \end{aligned} \tag{8}$$

其中设 $f(x_1, x_2, \cdots, x_n) = \sum\limits_{i_1, i_2, \cdots, i_n} a_{i_1 i_2 \cdots i_n} x_1^{i_1} x_2^{i_2} \cdots x_n^{i_n}$，而

$$f(t_1, t_2, \cdots, t_n) := \sum_{i_1, i_2, \cdots, i_n} a_{i_1 i_2 \cdots i_n} t_1^{i_1} t_2^{i_2} \cdots t_n^{i_n}$$

则 $\sigma_{t_1,\cdots,t_n}$ 是 $K[x_1, x_2, \cdots, x_n]$ 到 $R$ 的一个映射，它使得

$$\sigma_{t_1,\cdots,t_n}(x_i) = t_i, \quad i = 1, \cdots, n; \quad \sigma_{t_1,\cdots,t_n}(a) = a, \quad \forall\, a \in K$$

并且它保持加法与乘法运算，即如果

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

$$f(x_1,\cdots,x_n)g(x_1,\cdots,x_n)=p(x_1,\cdots,x_n)$$

那么
$$f(t_1,\cdots,t_n)+g(t_1,\cdots,t_n)=h(t_1,\cdots,t_n)$$
$$f(t_1,\cdots,t_n)g(t_1,\cdots,t_n)=p(t_1,\cdots,t_n)$$

由(8)式定义的映射 $\sigma_{t_1,\cdots,t_n}$ 称为 $x_1,\cdots,x_n$ **用** $t_1,\cdots,t_n$ **代入**.

**证明** 与本章§1 定理 7.1.1 的证明类似. ▎

$K[x_1,x_2,\cdots,x_n]$ 中所有零次多项式添上零多项式组成的子集是 $K[x_1,x_2,\cdots,x_n]$ 的一个子环，它与 $K$ 是环同构的，因此 $K[x_1,x_2,\cdots,x_n]$ 可以看成是 $K$ 的一个扩环. 从而 $x_1,x_2,\cdots,x_n$ 可以用 $K[x_1,x_2,\cdots,x_n]$ 中任意 $n$ 个元素代入，这种代入是保持加法与乘法运算的.

特别重要的一种情形是：$x_1,x_2,\cdots,x_n$ 用 $K$ 中任意 $n$ 个元素 $c_1,c_2,\cdots,c_n$ 代入，由此我们可引进多元多项式函数的概念：

设 $f(x_1,x_2,\cdots,x_n)$ 是数域 $K$ 上的一个 $n$ 元多项式，对于 $K$ 中任意 $n$ 个元素 $c_1,c_2,\cdots,c_n$，将 $x_1,x_2,\cdots,x_n$ 用 $c_1,c_2,\cdots,c_n$ 代入，得 $f(c_1,c_2,\cdots,c_n)\in K$. 于是 $n$ 元多项式 $f(x_1,x_2,\cdots,x_n)$ 确定了集合 $K^n$ 到数域 $K$ 的一个映射(即 $K$ 上的 $n$ 元函数)，仍用 $f$ 表示，即

$$\begin{aligned} f:\ & K^n \longrightarrow K \\ & (c_1,c_2,\cdots,c_n) \longmapsto f(c_1,c_2,\cdots,c_n) \end{aligned} \tag{9}$$

这种由数域 $K$ 上的 $n$ 元多项式确定的 $K$ 上的 $n$ 元函数称为**数域 $K$ 上的 $n$ 元多项式函数**.

设 $f(x_1,x_2,\cdots,x_n)$，$g(x_1,x_2,\cdots,x_n)$ 是数域 $K$ 上的两个 $n$ 元多项式. 显然，如果 $f(x_1,x_2,\cdots,x_n)$ 与 $g(x_1,x_2,\cdots,x_n)$ 相等，则它们确定的 $n$ 元多项式函数 $f$ 与 $g$ 也相等，反之如何？为此我们先证一个命题：

**命题** 7.10.1 设 $h(x_1,\cdots,x_n)$ 是数域 $K$ 上的一个 $n$ 元多项式，如果 $h(x_1,\cdots,x_n)\neq 0$，则 $K$ 中存在 $n$ 个元素 $c_1,c_2,\cdots,c_n$，使得 $h(c_1,c_2,\cdots,c_n)\neq 0$.

**证明** 对不定元的数目 $n$ 作归纳法. $n=1$ 时，由于数域 $K$ 上

非零的一元多项式$h(x)$在$K$中的根的数目不超过多项式的次数，而数域$K$有无穷多个元素，因此如果$h(x)\neq 0$，则存在$c\in K$，使$h(c)\neq 0$.

假设命题对$K[x_1,\cdots,x_{n-1}]$中的多项式成立，现在看$K[x_1,\cdots,x_{n-1},x_n]$中的多项式$h(x_1,\cdots,x_n)$. 把$h(x_1,\cdots,x_n)$写成

$$\begin{aligned}h(x_1,\cdots,x_n)=&u_0(x_1,\cdots,x_{n-1})+u_1(x_1,\cdots,x_{n-1})x_n\\&+\cdots+u_s(x_1,\cdots,x_{n-1})x_n^s\end{aligned}$$

其中$u_i(x_1,\cdots,x_{n-1})\in K[x_1,\cdots,x_{n-1}]$，$i=0,1,\cdots,s$，且$u_s(x_1,\cdots,x_{n-1})\neq 0$. 据归纳假设，存在$c_1,\cdots,c_{n-1}\in K$，使得$u_s(c_1,\cdots,c_{n-1})\neq 0$. 于是$K[x_n]$中的多项式

$$\begin{aligned}h(c_1,\cdots,c_{n-1},x_n)=&u_0(c_1,\cdots,c_{n-1})+u_1(c_1,\cdots,c_{n-1})x_n\\&+\cdots+u_s(c_1,\cdots,c_{n-1})x_n^s\end{aligned}$$

是非零多项式，因此存在$c_n\in K$，使得

$$\begin{aligned}h(c_1,\cdots,c_{n-1},c_n)=&u_0(c_1,\cdots,c_{n-1})+u_1(c_1,\cdots,c_{n-1})c_n\\&+\cdots+u_s(c_1,\cdots,c_{n-1})c_n^s\neq 0\end{aligned}$$ ▌

注意：命题 7.10.1 证明中的关键是$K$有无穷多个元素.

从命题 7.10.1 立即得到

**定理 7.10.4** 设$f(x_1,\cdots,x_n),g(x_1,\cdots,x_n)\in K[x_1,\cdots,x_n]$，如果多项式$f(x_1,\cdots,x_n)$与$g(x_1,\cdots,x_n)$不相等，则由它们确定的$n$元多项式函数$f$与$g$也不相等.

**证明** 考虑多项式

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

如果多项式$f$与$g$不相等，则$h(x_1,\cdots,x_n)\neq 0$. 据命题 7.10.1，存在$c_1,\cdots,c_n\in K$，使得$h(c_1,\cdots,c_n)\neq 0$. $x_1,\cdots,x_n$用$c_1,\cdots,c_n$代入，由上述式子可推出

$$f(c_1,\cdots,c_n)\neq g(c_1,\cdots,c_n)$$

所以映射$f$与$g$不相等. ▌

我们把数域$K$上所有$n$元多项式函数组成的集合记作$K_{\text{npol}}$，

在这个集合中规定加法与乘法如下：$\forall\ (c_1,\cdots,c_n)\in K^n$，

$$(f+g)(c_1,\cdots,c_n):=f(c_1,\cdots,c_n)+g(c_1,\cdots,c_n) \tag{10}$$

$$(fg)(c_1,\cdots,c_n):=f(c_1,\cdots,c_n)g(c_1,\cdots,c_n) \tag{11}$$

容易验证 $K_{\mathrm{npol}}$ 是一个环，称它为**数域 $K$ 上的 $n$ 元多项式函数环**. 类似于定理 7.6.5 的证法可以证明：数域 $K$ 上的 $n$ 元多项式环 $K[x_1,\cdots,x_n]$ 与 $K$ 上的 $n$ 元多项式函数环 $K_{\mathrm{npol}}$ 是同构的. 因此我们可以把数域 $K$ 上的 $n$ 元多项式与 $K$ 上的 $n$ 元多项式函数等同看待.

设 $f(x_1,\cdots,x_n)\in K[x_1,\cdots,x_n]$，对于 $c_1,\cdots,c_n\in K$，如果 $f(c_1,\cdots,c_n)=0$，则称 $(c_1,\cdots,c_n)$ 是 $f(x_1,\cdots,x_n)$ 的一个**零点**. 当 $K$ 取实数域，若 $n=2$，则 $f(x,y)$ 的零点组成的集合就是平面上的一条**代数曲线**；若 $n=3$，则 $f(x,y,z)$ 的零点组成的集合是空间中的一个**代数曲面**. 研究数域 $K$ 上 $n$ 元多项式的零点组成的集合是代数几何的一个基本内容.

利用不定元 $x_1,\cdots,x_n$ 用 $K[x_1,\cdots,x_n]$ 的 $n$ 个元素代入是保持运算的，我们可以得出齐次多项式的一个特征性质：

**定理** 7.10.5　设 $f(x_1,\cdots,x_n)\in K[x_1,\cdots,x_n]$ 且 $f\neq 0$，$m$ 是一个非负整数. 则 $f(x_1,\cdots,x_n)$ 是 $m$ 次齐次多项式的充分必要条件为对一切 $t\in K$，有

$$f(tx_1,tx_2\cdots,tx_n)=t^m f(x_1,x_2,\cdots,x_n) \tag{12}$$

* **证明**　必要性. 设 $f(x_1,\cdots,x_n)$ 是 $m$ 次齐次多项式，即

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

$$i_1+i_2+\cdots+i_n=m$$

任取 $t\in K$，不定元 $x_1,x_2,\cdots,x_n$ 用 $tx_1,tx_2\cdots,tx_n$ 代入，从上式得

$$f(tx_1,\cdots,tx_n)=\sum_{i_1,i_2,\cdots,i_n} a_{i_1i_2\cdots i_n}(tx_1)^{i_1}\cdots(tx_n)^{i_n}$$

$$=t^m f(x_1,\cdots,x_n)$$

充分性. 设对一切 $t\in K$，有(12)式成立. 将 $f(x_1,\cdots,x_n)$

写成

$$f = f_0 + f_1 + \cdots + f_s \tag{13}$$

其中 $f_i$ 是 $f$ 的 $i$ 次齐次成分，$i = 0,1,\cdots,s$. 任取 $t \in K$，$x_1,x_2,\cdots,x_n$ 用 $tx_1,tx_2\cdots,tx_n$ 代入，从(13) 式得出

$$\begin{aligned} f(tx_1,\cdots,tx_n) = & f_0(tx_1,\cdots,tx_n) + f_1(tx_1,\cdots,tx_n) \\ & + \cdots + f_s(tx_1,\cdots,tx_n) \end{aligned}$$

据刚证的必要性以及充分性的假设得

$$\begin{aligned} t^m f(x_1,\cdots,x_n) = & f_0(x_1,\cdots,x_n) + tf_1(x_1,\cdots,x_n) \\ & + \cdots + t^s f_s(x_1,\cdots,x_n) \end{aligned} \tag{14}$$

将(13)式代入(14)式左端，并且根据 $K[x_1,\cdots,x_n]$ 中两个多项式相等的规定，我们得到

$$t^m f_i(x_1,\cdots,x_n) = t^i f_i(x_1,\cdots,x_n), \quad i = 0,1,\cdots,s \tag{15}$$

任取 $i \in \{0,1,\cdots,s\}$ 且 $i \neq m$. 如果 $f_i \neq 0$，则从(15) 式两边消去 $f_i$ 得 $t^m = t^i$，$\forall\, t \in K$. 据定理 7.6.4 得，$x^m = x^i$，这与 $i \neq m$ 矛盾. 因此 $f_i = 0$，当 $i \neq m$. 所以 $f = f_m$. 已知 $f \neq 0$，因此 $f$ 是 $m$ 次齐次多项式. ▌

我们指出，设 $R$ 是(或可看成是)$K$ 的一个交换扩环，任取 $R$ 中 $n$ 个元素 $t_1,t_2,\cdots,t_n$，对于 $f(x_1,x_2,\cdots,x_n) \in K[x_1,x_2,\cdots,x_n]$，把 $x_1,x_2,\cdots,x_n$ 用 $t_1,t_2,\cdots,t_n$ 代入，得到的 $f(t_1,t_2,\cdots,t_n)$ 称为 $t_1,t_2,\cdots,t_n$ **在 $K$ 上的一个多项式**. 尽管从形式上看，$t_1,t_2,\cdots,t_n$ 的多项式 $f(t_1,t_2,\cdots,t_n)$ 与 $n$ 元多项式 $f(x_1,x_2,\cdots,x_n)$ 类似，但它们有本质不同：$t_1,t_2,\cdots,t_n$ 的一个多项式，其表法可能不唯一，而一个 $n$ 元多项式，其表法唯一.

## 习　题　7.10

1. 将下列 4 元多项式按字典排列法排列各单项式的顺序：

(1) $f(x_1,x_2,x_3,x_4)=x_3^4x_4-x_1^3x_2+5x_2x_3x_4+2x_2^4x_3x_4$;

(2) $f(x_1,x_2,x_3,x_4)=x_1^3+x_3^2+3x_1x_2^2x_4-5x_1^2x_3x_4^2-2x_2^3x_3$.

2. 设 $f(x_1,\cdots,x_n)$ 是数域 $K$ 上一个 $n$ 元齐次多项式，证明：如果 $g(x_1,\cdots,x_n),h(x_1,\cdots,x_n)\in K[x_1,\cdots,x_n]$，使得

$$f(x_1,\cdots,x_n)=g(x_1,\cdots,x_n)h(x_1,\cdots,x_n)$$

则 $g(x_1,\cdots,x_n),h(x_1,\cdots,x_n)$ 也是齐次多项式.

3. 把 3 元多项式 $f(x_1,x_2,x_3)=x_1^3+x_2^3+x_3^3-3x_1x_2x_3$ 写成两个 3 元多项式的乘积.

4. 设 $f,g\in K[x_1,\cdots,x_n]$，且 $g\neq 0$. 证明：如果对于使 $g(c_1,\cdots,c_n)\neq 0$ 的任意一组元素 $c_1,\cdots,c_n\in K$，都有 $f(c_1,\cdots,c_n)=0$，则 $f(x_1,\cdots,x_n)=0$.

*5. 设 $f,g\in K[x_1,\cdots,x_n]$，如果存在 $h\in K[x_1,\cdots,x_n]$，使得 $f=gh$，则称 $g$ 整除 $f$，记作 $g|f$. 此时称 $g$ 是 $f$ 的一个因式. 如果 $f|g$，并且 $g|f$，则称 $f$ 与 $g$ 是相伴的，记作 $f\sim g$.

(1) 证明：$f\sim g\Leftrightarrow f$ 与 $g$ 相差一个非零数因子；

(2) $f\in K[x_1,\cdots,x_n]$ 称为不可约的，如果 $f$ 的因式只有 $K$ 中非零数以及 $f$ 的相伴元. 否则称为可约的. 证明：在 $K[x,y]$ 里，多项式 $x^2-y$ 是不可约的.

## §11 对称多项式

这一节我们讨论一类重要的多元多项式，即对称多项式. 在给出对称多项式的概念之前，先介绍置换的概念.

**定义 1** 一个有限集合 $\Omega$ 到自身的一个双射叫做 $\Omega$ 的一个**置换**.

设 $\Omega$ 是具有 $n$ 个元素的有限集. 因为我们讨论的问题与这些元素本身无关，所以我们不妨假定 $\Omega=\{1,2,\cdots,n\}$. 设 $\sigma$ 是 $\Omega$ 的一个置换(也称 $\sigma$ 为 $n$ 元置换)，设 $i$ 在映射 $\sigma$ 下的象是 $a_i$，即

$$\sigma(i)=a_i,\qquad i=1,2,\cdots,n$$

我们可以把置换 $\sigma$ 表示如下：

$$\sigma = \begin{pmatrix} 1 & 2 & \cdots & n \\ a_1 & a_2 & \cdots & a_n \end{pmatrix} \tag{1}$$

由于 $\sigma$ 是 $\Omega$ 到自身的一个双射，所以 $a_1a_2\cdots a_n$ 是自然数 $1,2,\cdots,n$ 的一个全排列，即 $n$ 元排列. 反之，如果给了一个 $n$ 元排列 $a_1a_2\cdots a_n$，则 $\sigma: i \longmapsto a_i$ 是 $\Omega$ 的一个置换. 因此 $n$ 元置换组成的集合与 $n$ 元排列组成的集合之间有一个一一对应. 我们用 $S_n$ 表示所有 $n$ 元置换组成的集合，由刚才的一一对应知道，$|S_n| = n!$.

例如，

$$\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 3 & 1 & 4 & 2 \end{pmatrix} \tag{2}$$

是一个 4 元置换，$\sigma(1) = 3, \sigma(2) = 1, \sigma(3) = 4, \sigma(4) = 2$.

设 $\sigma$ 是 $\Omega$ 的一个置换，因为 $\sigma$ 是 $\Omega$ 到自身的一个双射，所以 $\sigma$ 是可逆映射. $\sigma$ 的逆映射 $\sigma^{-1}$ 称为置换 $\sigma$ 的**逆置换**. 例如，(2) 式给出的置换 $\sigma$ 的逆置换 $\sigma^{-1}$ 如下：

$$\sigma^{-1}(3) = 1, \quad \sigma^{-1}(1) = 2, \quad \sigma^{-1}(4) = 3, \quad \sigma^{-1}(2) = 4$$

于是 $\sigma^{-1}$ 可以表示成

$$\sigma^{-1} = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 4 & 1 & 3 \end{pmatrix} \tag{3}$$

请读者直接从(2)式和(3)式去验证 $\sigma^{-1}\sigma = 1_\Omega, \sigma\sigma^{-1} = 1_\Omega$，其中 $1_\Omega$ 表示 $\Omega$ 到自身的恒等映射.

如果一个 $n$ 元置换把 $i$ 映成 $j$，把 $j$ 映成 $i$，而其余元素不变，则称这个置换是一个**对换**，记作$(ij)$. 显然$(ij)^{-1} = (ij)$.

任意给定一个 $\sigma \in S_n$，由于 $K[x_1,\cdots,x_n]$ 可以看成是 $K$ 的一个扩环，据 $n$ 元多项式环的通用性质，$x_1,x_2,\cdots,x_n$ 用 $x_{\sigma^{-1}(1)}$，$x_{\sigma^{-1}(2)},\cdots,x_{\sigma^{-1}(n)}$ 代入是 $K[x_1,x_2,\cdots,x_n]$ 到 $K[x_1,x_2,\cdots,x_n]$ 的一个映射，并且它保持加法与乘法运算，这个映射记作 $\widetilde{\sigma}$. 即

$$\begin{aligned} \widetilde{\sigma}: K[x_1,x_2,\cdots,x_n] &\longrightarrow K[x_1,x_2,\cdots,x_n] \\ f(x_1,x_2,\cdots,x_n) &\longmapsto f(x_{\sigma^{-1}(1)},x_{\sigma^{-1}(2)},\cdots,x_{\sigma^{-1}(n)}) \end{aligned}$$

也就是说

$$(\tilde{\sigma} f)(x_1, x_2, \cdots, x_n) = f(x_{\sigma^{-1}(1)}, x_{\sigma^{-1}(2)}, \cdots, x_{\sigma^{-1}(n)}) \qquad (4)$$

我们把上述 $\tilde{\sigma}$ 称为**置换 $\sigma$ 诱导的代入**.

例如，设 $\sigma$ 是(2) 式给出的 4 元置换，设

$$f(x_1, x_2, x_3, x_4) = x_1^4 + 3x_1^2 x_2^3 x_4 - 5x_1 x_2^3 x_3^4 x_4^2 + x_3^5 x_4$$

则 $(\tilde{\sigma} f)(x_1, x_2, x_3, x_4) = x_2^4 + 3x_2^2 x_4^3 x_3 - 5x_2 x_4^3 x_1^4 x_3^2 + x_1^5 x_3$

现在我们来讨论对称多项式.

**定义 2** 设 $f(x_1, x_2, \cdots, x_n) \in K[x_1, x_2, \cdots, x_n]$，如果对一切 $\sigma \in S_n$ 都有 $\tilde{\sigma} f = f$，则称 $f(x_1, x_2, \cdots, x_n)$ 是数域 $K$ 上的一个 **$n$ 元对称多项式.**

从定义 2 得出，如果一个 $n$ 元对称多项式 $f(x_1, x_2, \cdots, x_n)$ 含有一项 $a x_1^{i_1} x_2^{i_2} \cdots x_n^{i_n}$，那么它也含有项

$$a x_{\sigma^{-1}(1)}^{i_1} x_{\sigma^{-1}(2)}^{i_2} \cdots x_{\sigma^{-1}(n)}^{i_n}, \qquad \forall\, \sigma \in S_n$$

由于当 $\sigma$ 跑遍一切 $n$ 元置换时，$\sigma^{-1}(1), \sigma^{-1}(2), \cdots, \sigma^{-1}(n)$ 便取遍 $1, 2, \cdots, n$ 的所有 $n$ 元排列，因此 $f(x_1, x_2, \cdots, x_n)$ 含有一切形如

$$a x_{j_1}^{i_1} x_{j_2}^{i_2} \cdots x_{j_n}^{i_n}$$

的项，其中 $j_1 j_2 \cdots j_n$ 是 $1, 2, \cdots, n$ 的任一 $n$ 元排列.

我们在抽象代数课程中将证明：任意一个 $n$ 元置换 $\sigma$ 都可以用 $n-1$ 个对换 $(12), (13), \cdots, (1n)$ 中某一些对换的乘积表示出来. 因此为了判断一个 $n$ 元多项式 $f(x_1, x_2, \cdots, x_n)$ 是否为对称多项式，就**只要看上述** $n-1$ **个对换诱导的代入是否把** $f$ **映成** $f$ **自身**.

例如，$S_3$ 有 6 个元素，它们的每一个都可用对换(12)，(13) 的乘积表示出：

$$\begin{pmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \end{pmatrix} = (12)(12); \qquad \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix} = (13)(12)(13)$$

$$\begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix} = (12); \qquad \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix} = (13)(12)$$

$$\begin{pmatrix}1 & 2 & 3\\3 & 1 & 2\end{pmatrix} = (12)(13); \qquad \begin{pmatrix}1 & 2 & 3\\3 & 2 & 1\end{pmatrix} = (13)$$

上述 6 个等式请读者验证，注意置换的乘法就是映射的乘法，而两个映射的乘积是先作右边的映射，接着作左边的映射（请看第四章 §2）. 记 $\sigma = (12), \tau = (13)$. 于是为了判断数域 $K$ 上的一个 3 元多项式 $f(x_1, x_2, x_3)$ 是否为对称多项式，就只要看 $\tilde{\sigma}f$ 是否等于 $f$，以及 $\tilde{\tau}f$ 是否等于 $f$. 例如，设 3 元多项式

$$\begin{aligned}&f(x_1, x_2, x_3)\\&= x_1^2x_2 + x_1^2x_3 + x_1x_2^2 + x_1x_3^2 + x_2^2x_3 + x_2x_3^2\end{aligned}$$

因为

$$\begin{aligned}&(\tilde{\sigma}f)(x_1, x_2, x_3)\\&= x_2^2x_1 + x_2^2x_3 + x_2x_1^2 + x_2x_3^2 + x_1^2x_3 + x_1x_3^2\\&= f(x_1, x_2, x_3)\\&(\tilde{\tau}f)(x_1, x_2, x_3)\\&= x_3^2x_2 + x_3^2x_1 + x_3x_2^2 + x_3x_1^2 + x_2^2x_1 + x_2x_1^2\\&= f(x_1, x_2, x_3)\end{aligned}$$

所以 $f(x_1, x_2, x_3)$ 是对称多项式.

**命题** 7.11.1　$K[x_1, x_2, \cdots, x_n]$ 中的对称多项式组成的集合 $W$ 是 $K[x_1, x_2, \cdots, x_n]$ 的一个子环.

**证明**　显然 $W$ 是非空的. 为了证 $W$ 是 $K[x_1, x_2, \cdots, x_n]$ 的子环，只要证 $W$ 对减法、乘法封闭. 任取 $f, g \in W$，则对任意 $\sigma \in S_n$，有 $\tilde{\sigma}f = f, \tilde{\sigma}g = g$. 因为 $\tilde{\sigma}$ 是保持加法与乘法运算的，并且容易看出 $-g$ 仍是对称多项式，所以有

$$\begin{aligned}\tilde{\sigma}(f - g) &= \tilde{\sigma}(f + (-g)) = \tilde{\sigma}f + \tilde{\sigma}(-g)\\&= f + (-g) = f - g\\\tilde{\sigma}(fg) &= (\tilde{\sigma}f)(\tilde{\sigma}g) = fg\end{aligned}$$

这证明了 $f - g \in W, fg \in W$. 因此 $W$ 是一个子环. ▌

命题 7.11.1 说明：$K[x_1, x_2, \cdots, x_n]$ 中的对称多项式的和、

差、积仍是对称多项式.

我们下一个任务是要研究这个子环 $W$ 的结构.

**命题** 7.11.2 设 $f_1,f_2,\cdots,f_n$ 都是 $K[x_1,x_2,\cdots,x_n]$ 中的对称多项式，则对于任意一个 $g(x_1,x_2,\cdots,x_n)\in K[x_1,x_2,\cdots,x_n]$，$x_1,x_2,\cdots,x_n$ 用 $f_1,f_2,\cdots,f_n$ 代入，得到的 $K[x_1,x_2,\cdots,x_n]$ 中多项式

$$g(f_1,f_2,\cdots,f_n):=h(x_1,x_2,\cdots,x_n) \tag{5}$$

仍然是对称多项式.

**证明** 设

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

$x_1,x_2,\cdots,x_n$ 用 $f_1,f_2,\cdots,f_n$ 代入，把 $g(x_1,x_2,\cdots,x_n)$ 映成

$$g(f_1,f_2,\cdots,f_n)=\sum_{i_1,\cdots,i_n}a_{i_1\cdots i_n}f_1^{i_1}f_2^{i_2}\cdots f_n^{i_n} \tag{6}$$

展开后这仍是 $K[x_1,x_2,\cdots,x_n]$ 中的多项式，记作 $h(x_1,x_2,\cdots,x_n)$. 由于 $K[x_1,x_2,\cdots,x_n]$ 中的对称多项式的和、差、积仍是对称多项式，所以从(6)式的右端看出 $h(x_1,x_2,\cdots,x_n)$ 是对称多项式. ∎

命题 7.11.2 表明：若知道了 $K[x_1,x_2,\cdots,x_n]$ 中的 $n$ 个对称多项式 $f_1,f_2,\cdots,f_n$，那么我们可以得到许许多多的对称多项式. 这启发我们在研究子环 $W$ 的结构时，第一步要找出 $n$ 个最基本的对称多项式，第二步要研究 $K[x_1,x_2,\cdots,x_n]$ 中任何一个对称多项式能不能用这 $n$ 个最基本的对称多项式运用命题 7.11.2 所指出的方法得到.

$K[x_1,x_2,\cdots,x_n]$ 中 $n$ 个最基本的对称多项式是哪些？本章§6 最后一段在研究多项式根与系数的关系时，曾得到 $f(x)$ 的 $n$ 个复根 $c_1,c_2,\cdots,c_n$ 的 $n$ 个表达式，其中每一个表达式关于 $c_1,c_2,\cdots,c_n$ 是对称的. 从这里受到启发，我们考虑下列 $n$ 个 $n$ 元多项式：

$$\sigma_1(x_1,x_2,\cdots,x_n)=x_1+x_2+\cdots+x_n$$

$$\sigma_2(x_1,x_2,\cdots,x_n)=x_1x_2+x_1x_3+\cdots+x_1x_n+x_2x_3 \\ +\cdots+x_2x_n+\cdots+x_{n-1}x_n$$

$$\cdots\quad\cdots\quad\cdots\quad\cdots\quad\cdots\quad\cdots\quad\cdots$$

$$\sigma_k(x_1,x_2,\cdots,x_n)=\sum_{1\leqslant i_1<\cdots<i_k\leqslant n}x_{i_1}x_{i_2}\cdots x_{i_k} \tag{7}$$

$$\cdots\quad\cdots\quad\cdots\quad\cdots\quad\cdots\quad\cdots\quad\cdots$$

$$\sigma_n(x_1,x_2,\cdots,x_n)=x_1x_2\cdots x_n$$

用 $\tau_j$ 表示对换 $(1j)\in S_n, j=2,\cdots,n$. 我们有

$$\tilde{\tau}_j\sigma_k(x_1,x_2,\cdots,x_n)=\sum_{1\leqslant i_1<\cdots<i_k\leqslant n}x_{\tau_j^{-1}(i_1)}x_{\tau_j^{-1}(i_2)}\cdots x_{\tau_j^{-1}(i_k)} \tag{8}$$

显然 $\tau_j^{-1}=(1j)$. 如果 $i_1i_2\cdots i_k$ 中不含 1 和 $j$,或者既含 1 又含 $j$,则 $x_{\tau_j^{-1}(i_1)}x_{\tau_j^{-1}(i_2)}\cdots x_{\tau_j^{-1}(i_k)}=x_{i_1}x_{i_2}\cdots x_{i_k}$. 如果 $i_1i_2\cdots i_k$ 中含 1 但不含 $j$,则 $x_{\tau_j^{-1}(i_1)}x_{\tau_j^{-1}(i_2)}\cdots x_{\tau_j^{-1}(i_k)}$ 中含 $x_j$ 但不含 $x_1$;如果 $i_1i_2\cdots i_k$ 中含 $j$ 但不含 1,则 $x_{\tau_j^{-1}(i_1)}x_{\tau_j^{-1}(i_2)}\cdots x_{\tau_j^{-1}(i_k)}$ 中含 $x_1$ 但不含 $x_j$. 由此可知 $\tilde{\tau}_j\sigma_k=\sigma_k$,由于 $j=2,\cdots,n$,所以 $\sigma_k$ 是对称多项式,$k=1,2,\cdots,n$. 这 $n$ 个 $n$ 元对称多项式称为**初等对称多项式**.

据命题 7.11.2,对于 $K[x_1,x_2,\cdots,x_n]$ 中任一多项式 $g(x_1,x_2,\cdots,x_n)$,$x_1,x_2,\cdots,x_n$ 用 $\sigma_1,\sigma_2,\cdots,\sigma_n$ 代入,得到的 $K[x_1,x_2,\cdots,x_n]$ 中的多项式 $g(\sigma_1,\sigma_2,\cdots,\sigma_n)$(展开后)是对称多项式. 这个结论可简洁地说成:初等对称多项式的多项式是对称多项式. 下面我们来证明,$K[x_1,x_2,\cdots,x_n]$ 中任何一个对称多项式都能用这种方式得到.

**定理** 7.11.1(**对称多项式基本定理**) 对于 $K[x_1,x_2,\cdots,x_n]$ 中任意一个对称多项式 $f(x_1,x_2,\cdots,x_n)$,都存在 $K[x_1,x_2,\cdots,x_n]$ 中唯一的一个多项式 $g(x_1,x_2,\cdots,x_n)$,使得

$$f(x_1,x_2,\cdots,x_n)=g(\sigma_1,\sigma_2,\cdots,\sigma_n)$$

**证明** 存在性. 采取消去首项法. 设 $n$ 元对称多项式 $f(x_1,x_2,\cdots,x_n)$ 按字典排列法的首项是

$$ax_1^{l_1}x_2^{l_2}\cdots x_n^{l_n},\qquad a\neq 0 \tag{9}$$

我们断言 $l_1\geqslant l_2\geqslant\cdots\geqslant l_n$. 理由如下：假如 $l_i<l_{i+1}$，由于 $f(x_1,x_2,\cdots,x_n)$ 是对称多项式，因此对换 $\tau=(i,i+1)$ 诱导的代入 $\tilde{\tau}$ 使得 $\tilde{\tau}f=f$. 于是项(9) 在 $\tilde{\tau}$ 下的象

$$ax_1^{l_1}x_2^{l_2}\cdots x_{i+1}^{l_i}x_i^{l_{i+1}}\cdots x_n^{l_n} \tag{10}$$

也是 $f$ 的一项. 而项(10) 的幂指数组 $(l_1,\cdots,l_{i-1},l_{i+1},l_i,\cdots,l_n)$ 先于项(9) 的幂指数组 $(l_1,\cdots,l_{i-1},l_i,l_{i+1},l_i,\cdots,l_n)$，这与项(9) 是 $f(x_1,x_2,\cdots,x_n)$ 的首项矛盾. 所以 $l_1\geqslant l_2\geqslant\cdots\geqslant l_n$.

为了消去 $f(x_1,x_2,\cdots,x_n)$ 的首项，同时又要出现 $\sigma_1,\sigma_2,\cdots,\sigma_n$，我们作多项式

$$\varphi_1(x_1,x_2,\cdots,x_n)=a\sigma_1^{l_1-l_2}\sigma_2^{l_2-l_3}\cdots\sigma_{n-1}^{l_{n-1}-l_n}\sigma_n^{l_n} \tag{11}$$

因为 $K[x_1,x_2,\cdots,x_n]$ 中对称多项式的乘积还是对称多项式，所以 $\varphi_1(x_1,x_2,\cdots,x_n)$ 是对称多项式. 又由于多项式的乘积的首项等于它们的首项的乘积，因此 $\varphi_1(x_1,x_2,\cdots,x_n)$ 的首项是

$$\begin{aligned}&ax_1^{l_1-l_2}(x_1x_2)^{l_2-l_3}\cdots(x_1x_2\cdots x_{n-1})^{l_{n-1}-l_n}(x_1x_2\cdots x_{n-1}x_n)^{l_n}\\&=ax_1^{l_1}x_2^{l_2}\cdots x_{n-1}^{l_{n-1}}x_n^{l_n}\end{aligned}$$

它等于 $f(x_1,x_2,\cdots,x_n)$ 的首项. 令

$$\begin{aligned}&f_1(x_1,x_2,\cdots,x_n)\\&=f(x_1,x_2,\cdots,x_n)-\varphi_1(x_1,x_2,\cdots,x_n)\end{aligned} \tag{12}$$

则 $f_1$ 的首项“小于”$f$ 的首项(即 $f$ 的首项的幂指数组先于 $f_1$ 的首项的幂指数组)，并且由于对称多项式的差仍是对称多项式，所以 $f_1(x_1,x_2,\cdots,x_n)$ 是 $K[x_1,x_2,\cdots,x_n]$ 中的对称多项式.

对 $f_1(x_1,x_2,\cdots,x_n)$ 重复上述作法，我们又得到 $K[x_1,x_2,\cdots,x_n]$ 中的一个对称多项式

$$f_2(x_1,x_2,\cdots,x_n)=f_1(x_1,x_2,\cdots,x_n)-\varphi_2(x_1,x_2,\cdots,x_n)$$

其中 $\varphi_2(x_1,x_2,\cdots,x_n)$ 是 $\sigma_1,\sigma_2,\cdots,\sigma_n$ 的适当方幂的乘积，并且其系数等于 $f_1$ 的首项系数(类似于(11))，而 $f_2$ 的首项“小于”$f_1$ 的首项.

如此继续下去，我们得到$K[x_1,x_2,\cdots,x_n]$中一系列的对称多项式

$$f, \qquad f_1=f-\varphi_1, \qquad f_2=f_1-\varphi_2, \qquad \cdots \tag{13}$$

它们的首项一个比一个"小"，其中$\varphi_i$是$\sigma_1,\sigma_2,\cdots,\sigma_n$适当方幂的乘积，并且其系数等于$f_{i-1}$的首项系数. 设$f_i$的首项的幂指数组为$(p_1,p_2,\cdots,p_n)$，则

$$(l_1,l_2,\cdots,l_n)>(p_1,p_2,\cdots,p_n)$$

于是$l_1\geqslant p_1$. 又因为$f_i$是对称多项式，所以$p_1\geqslant p_2\geqslant\cdots\geqslant p_n$. 因此$l_1\geqslant p_1\geqslant p_2\geqslant\cdots\geqslant p_n$. 满足这个条件的非负整数组$(p_1,p_2,\cdots,p_n)$只有有限多个，因此序列(13)中只能有有限多个$f_i$不为零. 即有正整数$s$，使得$f_s=0$. 于是有

$$f_1=f-\varphi_1, \qquad f_2=f_1-\varphi_2, \qquad \cdots$$
$$f_{s-1}=f_{s-2}-\varphi_{s-1}, \qquad f_s=f_{s-1}-\varphi_s=0$$

从而得到

$$f=\varphi_1+\varphi_2+\cdots+\varphi_s \tag{14}$$

设$\varphi_i(x_1,x_2,\cdots,x_n)=a_i\sigma_1^{t_{i1}}\sigma_2^{t_{i2}}\cdots\sigma_n^{t_{in}}$，令

$$g(x_1,x_2,\cdots,x_n)=\sum_{i=1}^{s}a_ix_1^{t_{i1}}x_2^{t_{i2}}\cdots x_n^{t_{in}}\in K[x_1,x_2,\cdots,x_n] \tag{15}$$

则 $$g(\sigma_1,\sigma_2,\cdots,\sigma_n)=\sum_{i=1}^{s}a_i\sigma_1^{t_{i1}}\sigma_2^{t_{i2}}\cdots\sigma_n^{t_{in}}=\sum_{i=1}^{s}\varphi_i(x_1,x_2,\cdots,x_n)$$
$$=f(x_1,x_2,\cdots,x_n)$$

存在性证毕.

*唯一性. 如果$K[x_1,x_2,\cdots,x_n]$中有两个不同的多项式$g_1(x_1,x_2,\cdots,x_n)$，$g_2(x_1,x_2,\cdots,x_n)$，使得

$$f(x_1,x_2,\cdots,x_n)=g_1(\sigma_1,\sigma_2,\cdots,\sigma_n)$$
$$f(x_1,x_2,\cdots,x_n)=g_2(\sigma_1,\sigma_2,\cdots,\sigma_n)$$

则 $$g_1(\sigma_1,\sigma_2,\cdots,\sigma_n)-g_2(\sigma_1,\sigma_2,\cdots,\sigma_n)=0$$

令 $$g(x_1,x_2,\cdots,x_n)$$
$$=g_1(x_1,x_2,\cdots,x_n)-g_2(x_1,x_2,\cdots,x_n) \tag{16}$$

则 $$g(\sigma_1,\sigma_2,\cdots,\sigma_n)$$
$$=g_1(\sigma_1,\sigma_2,\cdots,\sigma_n)-g_2(\sigma_1,\sigma_2,\cdots,\sigma_n)=0 \tag{17}$$
记 $$h(x_1,x_2,\cdots,x_n)$$
$$=g(\sigma_1(x_1,x_2,\cdots,x_n),\cdots,\sigma_n(x_1,x_2,\cdots,x_n)) \tag{18}$$
则由(17)式得
$$h(x_1,x_2,\cdots,x_n)=0 \tag{19}$$
由假设以及(16)式得
$$g(x_1,x_2,\cdots,x_n)\neq 0$$
于是据命题 7.10.1 得，存在 $b_1,b_2,\cdots,b_n\in K$，使得
$$g(b_1,b_2,\cdots,b_n)\neq 0 \tag{20}$$
令 $\varphi(x)=x^n-b_1x^{n-1}+\cdots+(-1)^kb_kx^{n-k}+\cdots+(-1)^nb_n$

设 $\varphi(x)$ 的 $n$ 个复根是 $\alpha_1,\alpha_2,\cdots,\alpha_n$，则从 Vieta 公式推出
$$b_1=\sigma_1(\alpha_1,\alpha_2,\cdots,\alpha_n),\cdots$$
$$b_k=\sigma_k(\alpha_1,\alpha_2,\cdots,\alpha_n),\cdots,b_n=\sigma_n(\alpha_1,\alpha_2,\cdots,\alpha_n) \tag{21}$$
$x_1,x_2,\cdots,x_n$ 用 $\alpha_1,\alpha_2,\cdots,\alpha_n$ 代入，从(19) 式得
$$h(\alpha_1,\alpha_2,\cdots,\alpha_n)=0$$
而从(18)、(21)和(20)式得
$$h(\alpha_1,\alpha_2,\cdots,\alpha_n)$$
$$=g(\sigma_1(\alpha_1,\alpha_2,\cdots,\alpha_n),\cdots,\sigma_n(\alpha_1,\alpha_2,\cdots,\alpha_n))$$
$$=g(b_1,b_2,\cdots,b_n)\neq 0$$
矛盾. 唯一性得证. ▮

对称多项式基本定理完全解决了 $K[x_1,x_2,\cdots,x_n]$ 中所有对称 多项式组成的子环的结构问题. 定理中存在性的证明是构造性的，可以实际地利用它去求多项式 $g(x_1,x_2,\cdots,x_n)$，使得
$$f(x_1,x_2,\cdots,x_n)=g(\sigma_1,\sigma_2,\cdots,\sigma_n)$$

**例 1** 在 $K[x_1,x_2,x_3]$ 中，用初等对称多项式表出对称多项式
$$f(x_1,x_2,x_3)=x_1^2x_2^2+x_1^2x_3^2+x_2^2x_3^2$$

**解** $f(x_1,x_2,x_3)$ 的首项是 $x_1^2x_2^2$，它的幂指数组为 $(2,2,0)$. 作多项式

$$\varphi_1(x_1,x_2,x_3)=\sigma_1^{2-2}\sigma_2^{2-0}\sigma_3^0=\sigma_2^2$$

令
$$\begin{aligned}f_1(x_1,x_2,x_3)&=f(x_1,x_2,x_3)-\varphi_1(x_1,x_2,x_3)\\&=x_1^2x_2^2+x_1^2x_3^2+x_2^2x_3^2-(x_1x_2+x_1x_3+x_2x_3)^2\\&=-2(x_1^2x_2x_3+x_1x_2^2x_3+x_1x_2x_3^2)\\&=-2\sigma_1\sigma_3\end{aligned}$$

$f_1(x_1,x_2,x_3)$ 的首项是 $-2x_1^2x_2x_3$，它的幂指数组为 $(2,1,1)$. 令

$$\varphi_2(x_1,x_2,x_3)=-2\sigma_1^{2-1}\sigma_2^{1-1}\sigma_3^1=-2\sigma_1\sigma_3$$
$$f_2(x_1,x_2,x_3)=f_1(x_1,x_2,x_3)-\varphi_2(x_1,x_2,x_3)=0$$

于是
$$f(x_1,x_2,x_3)=\sigma_2^2-2\sigma_1\sigma_3$$

对于较复杂的 $n$ 元对称多项式 $f(x_1,x_2,\cdots,x_n)$，求一个多项式 $g(x_1,x_2,\cdots,x_n)$，使得 $f(x_1,x_2,\cdots,x_n)=g(\sigma_1,\sigma_2,\cdots,\sigma_n)$，采用待定系数法更为简便. 我们举一个例子来说明这种方法.

**例** 2 在 $K[x_1,x_2,\cdots,x_n]$ 中，用初等对称多项式表出对称多项式

$$f(x_1,x_2,\cdots,x_n)=\sum x_1^2x_2^2$$

这里 $\sum x_1^2x_2^2$ 表示 $x_1^2x_2^2$ 经过 $S_n$ 中任一置换诱导的代入所得到的项的和.

注：由于 $S_n$ 中任一置换可以表示成 $(12),(13),\cdots,(1n)$ 中某些对换的乘积，因此在求 $\sum x_1^2x_2^2$ 的各项时，可以先求 $x_1^2x_2^2$ 在 $(1j)$ 诱导的代入 $\tilde{\tau}_j$ 下得到的项 $(j=2,\cdots,n)$，接着求所得到的项在 $\tilde{\tau}_j$ 下的象，如此反复进行，直到得不到新的项为止. 例如，$n=5$ 时，

$$\begin{aligned}f(x_1,x_2,x_3,x_4,x_5)=&x_1^2x_2^2+x_1^2x_3^2+x_1^2x_4^2+x_1^2x_5^2\\&+x_2^2x_3^2+x_2^2x_4^2+x_2^2x_5^2\\&+x_3^2x_4^2+x_3^2x_5^2\\&+x_4^2x_5^2\end{aligned}$$

**解** 基本定理证明中已指出，序列(13)中多项式 $f_i$ 的首项的幂指数组 $(p_1,p_2,\cdots,p_n)$ 应当满足

$$l_1 \geqslant p_1 \geqslant p_2 \geqslant \cdots \geqslant p_n$$

其中 $(l_1,l_2,\cdots,l_n)$ 是 $f$ 的首项的幂指数组. 本题 $f$ 的首项 $x_1^2x_2^2$ 的幂指数组为 $(2,2,0,\cdots,0)$. 所以 $(p_1,p_2,\cdots,p_n)$ 应满足

$$2 \geqslant p_1 \geqslant p_2 \geqslant \cdots \geqslant p_n$$

因为初等对称多项式 $\sigma_1,\sigma_2,\cdots,\sigma_n$ 都是齐次多项式，而齐次多项式的乘积仍是齐次多项式，所以基本定理证明中的 $\varphi_i$ 都是齐次多项式. 由于 $\varphi_1$ 的首项等于 $f$ 的首项 $x_1^2x_2^2$，所以 $\varphi_1$ 是 4 次齐次多项式. 本题的 $f$ 显然是 4 次齐次多项式. 于是 $f_1 = f - \varphi_1$ 也是 4 次齐次多项式. 同理 $f_2 = f_1 - \varphi_2$ 也是 4 次齐次多项式. 依次下去，对于本题，序列(13)中非零的 $f_i$ 都是 4 次齐次多项式. 因此应当有

$$p_1 + p_2 + \cdots + p_n = 4$$

从上两段可知，可能的幂指数组只有三个：

$$(2,2,0,\cdots,0), \qquad (2,1,1,0,\cdots,0), \qquad (1,1,1,1,0,\cdots,0)$$

它们分别是 $f,f_1,f_2$ 的首项的幂指数组. 由 $\varphi_i$ 的构造方法得

$$\varphi_1 = \sigma_1^{2-2}\sigma_2^{2-0}\sigma_3^{0-0}\cdots\sigma_{n-1}^{0-0}\sigma_n^0 = \sigma_2^2$$

$$\varphi_2 = a\sigma_1^{2-1}\sigma_2^{1-1}\sigma_3^{1-0}\sigma_4^{0-0}\cdots\sigma_{n-1}^{0-0}\sigma_n^0 = a\sigma_1\sigma_3$$

$$\varphi_3 = b\sigma_1^{1-1}\sigma_2^{1-1}\sigma_3^{1-1}\sigma_4^{1-0}\sigma_5^{0-0}\cdots\sigma_{n-1}^{0-0}\sigma_n^0 = b\sigma_4$$

其中 $a,b$ 是待定的系数. 据基本定理的证明，得

$$f(x_1,x_2,\cdots,x_n) = \sigma_2^2 + a\sigma_1\sigma_3 + b\sigma_4 \tag{22}$$

为了决定 $a,b$，先让 $x_1,x_2,x_3,x_4,\cdots,x_n$ 用 $1,1,1,0,\cdots,0$ 代入，$f$ 的值为 3，$\sigma_1,\sigma_2,\sigma_3,\sigma_4$ 的值依次是 $3,3,1,0$，于是从(22)式得到

$$3 = 3^2 + a\cdot 3\cdot 1 + b\cdot 0$$

解得 $a = -2$. 于是(22)式变成

$$f(x_1,x_2,\cdots,x_n) = \sigma_2^2 - 2\sigma_1\sigma_3 + b\sigma_4 \tag{23}$$

$x_1,x_2,x_3,x_4,\cdots,x_n$ 用 $1,1,1,1,0,\cdots,0$ 代入，从(23)式得

$$6 = 6^2 - 2\cdot 4\cdot 4 + b\cdot 1$$

所以 $b=2$. 于是

$$f(x_1,x_2,\cdots,x_n)=\sigma_2^2-2\sigma_1\sigma_3+2\sigma_4$$

例 2 中的对称多项式 $f(x_1,x_2,\cdots,x_n)$ 是齐次多项式. 如果所给的对称多项式 $f$ 不是齐次的，那么把 $f$ 表示成它的齐次成分之和，对每一个齐次成分(它仍然是对称多项式)采用例 2 的做法，最后把所得结果相加即可.

设 $f(x_1,x_2,\cdots,x_n)$ 是数域 $K$ 上的 $n$ 元对称多项式，数域 $F$ 包含 $K$. 任取 $F$ 中 $n$ 个元素 $c_1,c_2,\cdots,c_n$，不定元 $x_1,x_2,\cdots,x_n$ 用 $c_1,c_2,\cdots,c_n$ 代入，$f(x_1,x_2,\cdots,x_n)$ 在此代入下的象 $f(c_1,c_2,\cdots,c_n)$ 称为 $c_1,c_2,\cdots,c_n$ **在 $K$ 上的一个对称多项式**.

对称多项式基本定理的一个重要应用是研究一元多项式的根. 设

$$f(x)=x^n+a_{n-1}x^{n-1}+\cdots+a_1x+a_0\in K[x]$$

它的 $n$ 个复根为 $c_1,c_2,\cdots,c_n$. 据 Vieta 公式得

$$\begin{aligned}
-a_{n-1}&=\sigma_1(c_1,c_2,\cdots,c_n)\\
a_{n-2}&=\sigma_2(c_1,c_2,\cdots,c_n)\\
&\cdots\quad\cdots\quad\cdots\quad\cdots\quad\cdots\\
(-1)^k a_{n-k}&=\sigma_k(c_1,c_2,\cdots,c_n)\\
&\cdots\quad\cdots\quad\cdots\quad\cdots\quad\cdots\\
(-1)^n a_0&=\sigma_n(c_1,c_2,\cdots,c_n)
\end{aligned}\tag{24}$$

(24)式表明，$f(x)$ 的 $n$ 个复根的初等对称多项式分别是 $f(x)$ 的各项系数乘以$(-1)$的适当次幂.

现在任取 $K$ 上的一个 $n$ 元对称多项式 $h(x_1,x_2,\cdots,x_n)$，据对称多项式基本定理，存在 $K$ 上唯一的一个 $n$ 元多项式 $g(x_1,x_2,\cdots,x_n)$，使得

$$h(x_1,x_2,\cdots,x_n)=g(\sigma_1,\sigma_2,\cdots,\sigma_n)\tag{25}$$

$x_1,x_2,\cdots,x_n$ 用 $f(x)$ 的 $n$ 个复根 $c_1,c_2,\cdots,c_n$ 代入，从(25)和(24)式得

$$h(c_1,c_2,\cdots,c_n)$$
$$=g(-a_{n-1},\cdots,(-1)^k a_{n-k},\cdots,(-1)^n a_0) \qquad (26)$$

(26)式表明，$f(x)$ 的 $n$ 个复根的任一对称多项式等于 $f(x)$ 的各项系数的某一个多项式.

特别地，考虑 $f(x)$ 的 $n$ 个复根 $c_1,c_2,\cdots,c_n$ 的下述对称多项式

$$D(c_1,c_2,\cdots,c_n)=\prod_{1\leqslant j<i\leqslant n}(c_i-c_j)^2 \qquad (27)$$

据上一段知道，存在 $K$ 上唯一的一个 $n$ 元多项式 $Dis(x_1,x_2,\cdots,x_n)$，使得

$$D(c_1,c_2,\cdots,c_n)$$
$$=Dis(-a_{n-1},\cdots,(-1)^k a_{n-k},\cdots,(-1)^n a_0) \qquad (28)$$

注意到(27)式给出的 $n$ 个复根的对称多项式 $D(c_1,c_2,\cdots,c_n)$ 能用于判断 $f(x)$ 在复数域中有无重根：$f(x)$ 在复数域中有重根的充分必要条件为 $D(c_1,c_2,\cdots,c_n)=0$. 再据(28)式便得出

**命题** 7.11.3 $f(x)\in K[x]$ 在复数域中有重根的充分必要条件为

$$Dis(-a_{n-1},\cdots,(-1)^k a_{n-k},\cdots,(-1)^n a_0)=0 \qquad (29)$$

▌

上述命题表明，可以用 $f(x)$ 的系数 $a_{n-1},\cdots,a_0$ 的一个多项式

$$Dis(-a_{n-1},\cdots,(-1)^k a_{n-k},\cdots,(-1)^n a_0)$$

来判别 $f(x)$ 有无重根. 因此我们把

$$Dis(-a_{n-1},\cdots,(-1)^k a_{n-k},\cdots,(-1)^n a_0)$$

称为 $f(x)$ 的**判别式**，简记为 $D(f)$.

如何求出 $f(x)$ 的判别式 $D(f)$ 呢？从(28)和(27)式得

$$D(f)=Dis(-a_{n-1},\cdots,(-1)^k a_{n-k},\cdots,(-1)^n a_0)$$
$$=\prod_{1\leqslant j<i\leqslant n}(c_i-c_j)^2=|B||B'|=|BB'| \qquad (30)$$

其中

$$B=\begin{pmatrix}1 & 1 & \cdots & 1\\ c_1 & c_2 & \cdots & c_n\\ \cdots & \cdots & \cdots & \cdots\\ c_1^{n-1} & c_2^{n-1} & \cdots & c_n^{n-1}\end{pmatrix}$$

于是

$$D(f)=\left|\begin{pmatrix}1 & 1 & \cdots & 1\\ c_1 & c_2 & \cdots & c_n\\ \cdots & \cdots & \cdots & \cdots\\ c_1^{n-1} & c_2^{n-1} & \cdots & c_n^{n-1}\end{pmatrix}\begin{pmatrix}1 & c_1 & \vdots & c_1^{n-1}\\ 1 & c_2 & \vdots & c_2^{n-1}\\ \vdots & \vdots & \vdots & \vdots\\ 1 & c_n & \vdots & c_n^{n-1}\end{pmatrix}\right|$$

$$=\begin{vmatrix}n & \sum\limits_{i=1}^{n}c_i & \cdots & \sum\limits_{i=1}^{n}c_i^{n-1}\\ \sum\limits_{i=1}^{n}c_i & \sum\limits_{i=1}^{n}c_i^2 & \cdots & \sum\limits_{i=1}^{n}c_i^{n}\\ \cdots & \cdots & \cdots & \cdots\\ \sum\limits_{i=1}^{n}c_i^{n-1} & \sum\limits_{i=1}^{n}c_i^{n} & \cdots & \sum\limits_{i=1}^{n}c_i^{2n-2}\end{vmatrix} \tag{31}$$

(31)式表明，为了求出 $D(f)$，就需要计算

$$\sum_{i=1}^{n}c_i^k,\qquad k=0,1,2,\cdots,2n-2 \tag{32}$$

由于 $f(x)$ 的 $n$ 个复根 $c_1,c_2,\cdots,c_n$ 是未知的，因此必须想办法通过 $f(x)$ 的系数来计算 $\sum c_i^k$，注意到 $\sum c_i^k$ 是 $c_1,c_2,\cdots,c_n$ 的对称多项式，据前面所述，$\sum c_i^k$ 等于 $f(x)$ 的系数的某一多项式. 问题是如何求出这个多项式？仍然是要运用对称多项式基本定理. 为此我们考虑下列 $n$ 元对称多项式

$$\begin{aligned}&s_k(x_1,x_2,\cdots,x_n)\\ &=x_1^k+x_2^k+\cdots+x_n^k,\quad k=0,1,2,\cdots\end{aligned} \tag{33}$$

这些 $n$ 元对称多项式称为**幂和**.

据对称多项式基本定理，幂和 $s_k$ 能表示成初等对称多项式的

多项式. 具体的表示方法可以用下述递推公式求出.

**牛顿**(Newton)**公式**　当 $1 \leqslant k \leqslant n$ 时,

$$s_k - \sigma_1 s_{k-1} + \sigma_2 s_{k-2} + \cdots + (-1)^{k-1}\sigma_{k-1}s_1 + (-1)^k k\sigma_k = 0 \tag{34}$$

当 $k > n$ 时,

$$s_k - \sigma_1 s_{k-1} + \sigma_2 s_{k-2} + \cdots + (-1)^{n-1}\sigma_{n-1}s_{k-n+1} + (-1)^n \sigma_n s_{k-n} = 0 \tag{35}$$

* **证明**　考虑 $K[x_1,x_2,\cdots,x_n,x]$ 的一个多项式

$$f(x_1,x_2,\cdots,x_n,x) = (x-x_1)(x-x_2)\cdots(x-x_n) \tag{36}$$

直接计算得

$$f(x_1,x_2,\cdots,x_n,x) = x^n - \sigma_1 x^{n-1} + \sigma_2 x^{n-2} + \cdots + (-1)^n \sigma_n \tag{37}$$

$x_1,x_2,\cdots,x_n,x$ 用 $x_1,x_2,\cdots,x_n,x_i$ 代入,从(36)与(37)式得

$$x_i^n - \sigma_1 x_i^{n-1} + \sigma_2 x_i^{n-2} + \cdots + (-1)^n \sigma_n = 0 \tag{38}$$

让 $i = 1,2,\cdots,n$,则(38)给出了 $n$ 个等式. 当 $k \geqslant n$ 时,用 $x_i^{k-n}$ 乘(38)式两边,然后让 $i$ 从 1 变到 $n$ 求和,得

$$\sum_{i=1}^{n}[x_i^k - \sigma_1 x_i^{k-1} + \sigma_2 x_i^{k-2} + \cdots + (-1)^n \sigma_n x^{k-n}] = 0 .$$

即

$$s_k - \sigma_1 s_{k-1} + \sigma_2 s_{k-2} + \cdots + (-1)^n \sigma_n s_{k-n} = 0 \tag{39}$$

现在讨论 $1 \leqslant k \leqslant n$ 的情形. 考虑 $K[x_1,x_2,\cdots,x_n]$ 的一个多项式

$$f_{k,n}(x_1,x_2,\cdots,x_n) = s_k - \sigma_1 s_{k-1} + \sigma_2 s_{k-2} + \cdots + (-1)^{k-1}\sigma_{k-1}s_1 + (-1)^k k\sigma_k \tag{40}$$

由于对称多项式的和、差、积仍是对称多项式,所以 $f_{k,n}$ 是对称多项式. 利用次数公式可得出,$f_{k,n}$ 或者是 $k$ 次齐次多项式,或者是零多项式. 我们来证 $f_{k,n}$ 一定是零多项式. 对于 $m = n - k$ 作归纳法. $m = 0$,即 $k = n$ 时,由已证的(39)式知,$f_{k,n} = 0$. 假设当不定元的数目 $n$ 与 $k$ 的差 $n - k$ 小于 $m(>0)$ 时命题为真,现在看等于 $m$ 的

情形.

$x_1,\cdots,x_{n-1},x_n$ 用 $x_1,\cdots,x_{n-1},0$ 代入,从(40) 式得

$$f_{k,n}(x_1,\cdots,x_{n-1},0)=(s_k)_0-(\sigma_1)_0(s_{k-1})_0+\cdots+(-1)^{k-1}(\sigma_{k-1})_0(s_1)_0+(-1)^k k(\sigma_k)_0 \tag{41}$$

其中 $(s_j)_0,(\sigma_i)_0$ 恰好是 $K[x_1,\cdots,x_{n-1}]$ 中的幂和以及初等对称多项式,$j=1,\cdots,k;i=1,\cdots,k$. 因此(41) 式右端为 $f_{k,n-1}(x_1,\cdots,x_{n-1})$. 因为

$$(n-1)-k=m-1<m$$

所以据归纳法假设得

$$f_{k,n-1}(x_1,\cdots,x_{n-1})=0$$

从而 $$f_{k,n-1}(x_1,\cdots,x_{n-1},0)=0$$

我们把 $f_{k,n}(x_1,\cdots,x_{n-1},x_n)$ 写成下述形式:

$$f_{k,n}(x_1,\cdots,x_{n-1},x_n)=u_0(x_1,\cdots,x_{n-1})+u_1(x_1,\cdots,x_{n-1})x_n+\cdots+u_r(x_1,\cdots,x_{n-1})x_n^r \tag{42}$$

让 $x_1,\cdots,x_{n-1},x_n$ 用 $x_1,\cdots,x_{n-1},0$ 代入,从(42) 式得

$$u_0(x_1,\cdots,x_{n-1})=f_{k,n}(x_1,\cdots,x_{n-1},0)=0$$

因此 $x_n|f_{k,n}$(见习题 7.10 第 5 题关于整除的定义),即 $f_{k,n}=x_nh$. 设 $\tau_j=(1j),2\leqslant j\leqslant n$. 因为 $f_{k,n}$ 是对称多项式,所以

$$f_{k,n}=\tilde{\tau}_nf_{k,n}=x_1(\tilde{\tau}_nh)$$

$$f_{k,n}=\tilde{\tau}_jf_{k,n}=x_j(\tilde{\tau}_j(\tilde{\tau}_nh)),\quad j=2,\cdots,n-1$$

这表明 $x_j|f_{k,n},j=1,2,\cdots,n$. 即 $x_1,x_2,\cdots,x_n$ 都是 $f_{k,n}$ 的因式,从而

$$\begin{aligned}f_{k,n}(x_1,\cdots,x_n)&=x_1x_2\cdots x_ng(x_1,\cdots,x_n)\\&=\sigma_ng(x_1,\cdots,x_n)\end{aligned} \tag{43}$$

假如 $f_{k,n}\neq0$,则 $\deg f_{k,n}=k$,并且 $g\neq0$. 从(43) 式得

$$\deg f_{k,n}=\deg\sigma_n+\deg g$$

即 $k=n+\deg g$. 从而 $\deg g=k-n=-m<0$,矛盾. 所以 $f_{k,n}=0$.

据归纳法原理，对一切 $m \geqslant 0$，有 $f_{k,n}=0$. 从而当 $1 \leqslant k \leqslant n$ 时，(34) 式成立. ▌

利用牛顿公式，可以从 $s_{k-1}(c_1,\cdots,c_n),\cdots,s_1(c_1,\cdots,c_n)$ 以及 $\sigma_1(c_1,\cdots,c_n),\cdots,\sigma_n(c_1,\cdots,c_n)$ 计算出 $s_k(c_1,\cdots,c_n)$. 再从 (31) 式就可以求出 $f(x)$ 的判别式 $D(f)$.

$f(x)$ 的判别式 $D(f)$ 也称为方程 $f(x)=0$ 的判别式.

注意：在上述讨论中，$f(x)$ 的首项系数为 1. 如果 $f(x)$ 的首项系数为 $a_n$，则可以先对 $a_n^{-1}f(x)$ 运用上述方法求出它的判别式 $D(a_n^{-1}f)$，然后规定 $f(x)$ 的判别式为

$$D(f):=a_n^{2n-2}D(a_n^{-1}f) \tag{44}$$

**例** 3　求数域 $K$ 上二次方程 $f(x)=x^2+bx+c=0$ 的判别式.

**解**　设 $f(x)$ 的复根是 $c_1,c_2$，则

$$\sigma_1(c_1,c_2)=-b,\qquad \sigma_2(c_1,c_2)=c$$

据牛顿公式得

$$s_1=\sigma_1,\qquad s_2=\sigma_1 s_1-2\sigma_2=\sigma_1^2-2\sigma_2$$

于是　$$s_1(c_1,c_2)=-b,\qquad s_2(c_1,c_2)=b^2-2c$$

因此　$$D(f)=\begin{vmatrix} 2 & -b \\ -b & b^2-2c \end{vmatrix}=b^2-4c \tag{45}$$

读者已从中学代数里知道 $b^2-4c$ 是 $x^2+bx+c=0$ 的判别式，现在知道了"判别式"一词来源在这里.

**例** 4　求数域 $K$ 上不完全三次方程

$$f(x)=x^3+ax+b=0 \tag{46}$$

的判别式.

**解**　设 $f(x)$ 的复根是 $c_1,c_2,c_3$，则

$$\sigma_1(c_1,c_2,c_3)=0,\qquad \sigma_2(c_1,c_2,c_3)=a,$$
$$\sigma_3(c_1,c_2,c_3)=-b$$

据牛顿公式得

$$s_1(c_1,c_2,c_3) = \sigma_1(c_1,c_2,c_3) = 0, \qquad s_2(c_1,c_2,c_3) = -2a$$

$$s_3(c_1,c_2,c_3) = -3b, \qquad s_4(c_1,c_2,c_3) = 2a^2$$

所以

$$D(f) = \begin{vmatrix} 3 & 0 & -2a \\ 0 & -2a & -3b \\ -2a & -3b & 2a^2 \end{vmatrix} = -4a^3 - 27b^2 \qquad (47)$$

*如果考虑完全三次方程

$$f(x) = x^3 + a_1x^2 + a_2x + a_3 = 0 \qquad (48)$$

则 $D(f)$ 具有比(47)更复杂的形式. 但是我们可以借助下述办法把完全三次方程化为不完全三次方程：$x$ 用 $x - \frac{a_1}{3}$ 代入，令

$$g(x) := f(x - \frac{a_1}{3})$$

直接计算 $f(x - \frac{a_1}{3})$ 可知，$g(x)$ 是不完全三次多项式($x^2$ 的系数为 0). 如果我们求出了 $g(x)$ 的一个根 $c_0$，则

$$f(c_0 - \frac{a_1}{3}) = g(c_0) = 0$$

从而 $c_0 - \frac{a_1}{3}$ 是 $f(x)$ 的一个根. 因此把 $g(x)$ 的根弄清楚了，随之 $f(x)$ 的根也就清楚了. 所以不失一般性，我们总可以假定 $x^2$ 的系数 $a_1 = 0$.

## 习 题 7.11

1. 设 $f(x_1,x_2,x_3)$ 是数域 $K$ 上的一个三元多项式：

$$f(x_1,x_2,x_3) = x_1^3x_2^2 + x_1^3x_3^2 + x_1^2x_2^3 + x_1^2x_3^3 + x_2^3x_3^2 + x_2^2x_3^3$$

证明 $f(x_1,x_2,x_3)$ 是对称多项式.

2. 在 $K[x_1,x_2,x_3]$ 的含有项 $x_1^3x_2$ 的对称多项式中，写出项数最少的那个对称多项式.

3. 在 $K[x_1,x_2,x_3]$ 中，用初等对称多项式表出下列对称多项式：

(1) $x_1^3x_2+x_1x_2^3+x_1^3x_3+x_1x_3^3+x_2^3x_3+x_2x_3^3$；

(2) $x_1^4+x_2^4+x_3^4$；

(3) $(x_1x_2+x_3^2)(x_2x_3+x_1^2)(x_3x_1+x_2^2)$.

4. 在 $K[x_1,\cdots,x_n]$ 中，用初等对称多项式表出下列对称多项式：

(1) $\sum x_1^3$；　　(2) $\sum x_1^2x_2^2x_3$.

5. 在 $K[x_1,x_2,x_3,x_4]$ 中，利用牛顿公式把幂和 $s_2,s_3,s_4,s_5,s_6$ 用初等对称多项式表示.

6. 设 $f(x)$ 是实系数三次多项式，证明：当 $D(f)=0$ 时，$f(x)$ 在 $C$ 中有重根，并且 $f(x)$ 在 $C$ 中的根都是实数；当 $D(f)>0$ 时，$f(x)$ 有三个互不相同的实根；当 $D(f)<0$ 时，$f(x)$ 有一个实根，一对共轭虚根.

7. 设 $f(x)\in K[x]$，并且 $f(x)$ 的首项系数为 1，$a\in K$，$g(x)=(x-a)f(x)$，证明：

$$D(g)=D(f)f(a)^2$$

8. 求 $f(x)=x^n+a\in K[x]$ 的判别式.

9. 求数域 $K$ 上完全三次方程

$$f(x)=x^3+a_1x^2+a_2x+a_3=0$$

的判别式.

## §12 结式·二元高次方程组

上一节的命题 7.11.3 指出，$f(x)\in K[x]$ 的判别式 $D(f)=0$ 当且仅当 $f(x)$ 在复数域中有重根，从而 $f(x)$ 在 $K[x]$ 中有重因式. §5 的定理 7.5.2 指出，$K[x]$ 中的多项式 $f(x)$ 没有重因式的充分必要条件是 $f(x)$ 与它的导数 $f'(x)$ 互素. 把两者结合起来便得出，$f(x)\in K[x]$ 的判别式 $D(f)=0$ 当且仅当 $f(x)$ 与 $f'(x)$ 不互素(即 $f(x)$ 与 $f'(x)$ 在复数域中有公共根). 也就是说，利用多项式 $f(x)$ 的判别式 $D(f)$ 可以给出 $f(x)$ 与它的导数 $f'(x)$ 是否不互素(即在复数域中有没有公共根）的第二种判别法(第一种

判别法是利用辗转相除法去求 $f(x)$ 与 $f'(x)$ 的最大公因式). 这第二种判别法是直接从 $f(x)$ 的系数去判断(因为 $D(f)$ 是 $f(x)$ 的系数 $a_0, a_1, \cdots, a_{n-1}$ 的一个多项式,这里 $f(x)$ 的首项系数为1). 从这受到启发,一般地,为了判断两个多项式 $f(x), g(x) \in K[x]$ 是否不互素(即在复数域中有无公共根),除了用辗转相除法求出它们的最大公因式这种方法外,能不能直接从 $f(x)$ 与 $g(x)$ 的系数去判断呢?本节就来讨论这个问题.

设

$$f(x) = a_0x^n + a_1x^{n-1} + \cdots + a_{n-1}x + a_n$$

$$g(x) = b_0x^m + b_1x_{m-1} + \cdots + b_{m-1}x + b_m$$

是 $K[x]$ 中两个非零多项式,其中 $n > 0, m > 0$,并且允许 $a_0 = 0$ 或 $b_0 = 0$(包括 $a_0 = 0 = b_0$) 这些可能性.

我们首先来求 $f(x)$ 与 $g(x)$ 不互素(即它们在 $K[x]$ 中有次数大于零的公因式) 的必要条件. 假设 $f(x)$ 与 $g(x)$ 在 $K[x]$ 中有次数大于零的公因式 $d(x)$,则存在 $f_1(x), g_1(x) \in K[x]$,使得

$$f(x) = f_1(x)d(x), \qquad g(x) = g_1(x)d(x) \tag{1}$$

由于 $\deg d(x) > 0$,所以

$$\deg f_1(x) < \deg f(x) \leqslant n, \qquad \deg g_1(x) < \deg g(x) \leqslant m$$

从(1)式得

$$g_1(x)f(x) = f_1(x)g(x) \tag{2}$$

设

$$f_1(x) = u_0x^{n-1} + u_1x^{n-2} + \cdots + u_{n-1} \tag{3}$$

$$g_1(x) = v_0x^{m-1} + v_1x^{m-2} + \cdots + v_{m-1} \tag{4}$$

比较(2)式两边多项式的各次项的系数,得

$$\begin{cases} a_0v_0 = b_0u_0 \\ a_1v_0 + a_0v_1 = b_1u_0 + b_0u_1 \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ a_nv_{m-2} + a_{n-1}v_{m-1} = b_mu_{n-2} + b_{m-1}u_{n-1} \\ a_nv_{m-1} = b_mu_{n-1} \end{cases} \tag{5}$$

由于 $f(x) \neq 0$ 且 $g(x) \neq 0$，所以 $f_1(x) \neq 0$ 且 $g_1(x) \neq 0$. 因此

$$(u_0, u_1, \cdots, u_{n-1}) \neq 0, \qquad (v_0, v_1, \cdots, v_{m-1}) \neq 0$$

从而下述齐次线性方程组

$$\begin{cases} a_0x_0 & & + b_0y_0 & & = 0 \\ a_1x_0 + a_0x_1 & & + b_1y_0 + b_0y_1 & & = 0 \\ \cdots \cdots & \cdots \cdots \quad \cdots \cdots & \cdots \cdots \cdots & \cdots \cdots \cdots & \cdots \\ & a_nx_{m-2} + a_{n-1}x_{m-1} & & + b_my_{n-2} + b_{m-1}y_{n-1} & = 0 \\ & a_nx_{m-1} & & + \quad b_my_{n-1} & = 0 \end{cases} \tag{6}$$

有非零解：

$$(v_0, v_1, \cdots, v_{m-1}, -u_0, -u_1, \cdots, -u_{n-1})$$

所以它的系数矩阵 $A$ 的行列式等于零，从而 $|A'| = 0$，即

$$\begin{array}{r} m\text{ 行}\left\{\begin{array}{l} \\ \\ \\ \\ \end{array}\right. \\ n\text{ 行}\left\{\begin{array}{l} \\ \\ \\ \\ \end{array}\right. \end{array}
\begin{vmatrix} a_0 & a_1 & \cdots & \cdots & \cdots & \cdots & a_n & & & & & \\ & a_0 & a_1 & \cdots & \cdots & \cdots & \cdots & a_n & & & & \\ & & & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ & & & & a_0 & a_1 & \cdots & \cdots & \cdots & \cdots & a_n \\ b_0 & b_1 & \cdots & \cdots & \cdots & b_m & & & & & \\ & b_0 & b_1 & \cdots & \cdots & \cdots & b_m & & & & \\ & & & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ & & & & & b_0 & b_1 & \cdots & \cdots & \cdots & b_m \end{vmatrix} = 0 \tag{7}$$

**定义 1** 设

$$f(x) = a_0x^n + a_1x^{n-1} + \cdots + a_n$$

$$g(x) = b_0x^m + b_1x^{m-1} + \cdots + b_m$$

是 $K[x]$ 中两个多项式，其中 $n > 0$ 且 $m > 0$. (7) 式左端的行列式称为 $f(x)$ 与 $g(x)$ 的**结式**，记作 $\mathrm{Res}(f, g)$.

上面我们求出了 $K[x]$ 中两个非零多项式 $f(x)$ 与 $g(x)$ 不互素的必要条件是 $\mathrm{Res}(f, g) = 0$. 现在来看这是否为充分条件.

设 $\mathrm{Res}(f, g) = 0$，则齐次线性方程组(6) 有非零解：

$$(v_0, v_1, \cdots, v_{m-1}, -u_0, -u_1, \cdots, -u_{n-1})$$

从而(5)式成立. 令 $f_1(x), g_1(x)$ 分别如(3)、(4) 式，则(2) 式成立，并且有 $\deg f_1 < n, \deg g_1 < m$. 现在我们增加一个条件：$a_0$ 与 $b_0$ 不全为零. 不妨设 $a_0 \neq 0$，则 $\deg f = n$. 从(2) 式得 $f \mid f_1g$. 假如 $(f, g) = 1$，则 $f \mid f_1$. 从而

$$\deg f \leqslant \deg f_1 < n$$

矛盾. 因此在 $\mathrm{Res}(f,g)=0$ 且 $a_0$ 与 $b_0$ 不全为 0 的条件下, $f(x)$ 与 $g(x)$ 不互素.

综合上述讨论,我们可以得到下面的结论:

**定理** 7.12.1 设

$$f(x) = a_0x^n + a_1x^{n-1} + \cdots + a_n$$

$$g(x) = b_0x^m + b_1x^{m-1} + \cdots + b_m$$

是 $K[x]$ 中两个多项式,其中 $n>0$ 且 $m>0$,则 $\mathrm{Res}(f,g)=0$ 的充分必要条件是 $a_0=0=b_0$,或者 $f(x)$ 与 $g(x)$ 不互素(即它们在 $K[x]$ 中有次数大于零的公因式).

**证明** 如果 $f(x)$ 与 $g(x)$ 都是零多项式,则命题显然成立.

如果 $f(x)$ 与 $g(x)$ 有且只有一个是零多项式,不妨设 $g(x)=0, f(x)\neq 0$. 此时 $\mathrm{Res}(f,g)=0$. 若 $a_0\neq 0$,则 $\deg f=n>0$. 从而 $f(x)$ 就是 $f(x)$ 与 $g(x)$ 的一个次数大于零的公因式. 因此命题成立.

下面设 $f(x)$ 与 $g(x)$ 都是非零多项式. 必要性从定理 7.12.1 的上面一段立即得到. 充分性有一半已证(即本节的第三段). 现在证另一半:若 $a_0=0=b_0$,则 $\mathrm{Res}(f,g)$ 的第 1 列全为零,从而 $\mathrm{Res}(f,g)=0$. ▌

**例** 1 判断

$$f(x) = x^2 + 2x + 2$$

与 $$g(x) = x^2 + x + 1$$

在复数域中有没有公共根.

**解**

$$\mathrm{Res}(f,g) = \begin{vmatrix} 1 & 2 & 2 & 0 \\ 0 & 1 & 2 & 2 \\ 1 & 1 & 1 & 0 \\ 0 & 1 & 1 & 1 \end{vmatrix} = 1$$

所以 $f(x)$ 与 $g(x)$ 互素，从而它们在复数域中没有公共根.

定理 7.12.1 不仅可以用来判断 $K[x]$ 中两个多项式是否不互素(即在复数域中有无公共根)，而且还提供了解二元高次方程组的一个一般方法. 设 $f(x,y),g(x,y)\in K[x,y]$，我们来求方程组

$$\begin{cases} f(x,y)=0 \\ g(x,y)=0 \end{cases} \tag{8}$$

在复数域中的全部解.

把 $f(x,y),g(x,y)$ 都按 $x$ 的降幂写出：

$$f(x,y)=a_0(y)x^n+a_1(y)x^{n-1}+\cdots+a_n(y) \tag{9}$$

$$g(x,y)=b_0(y)x^m+b_1(y)x^{m-1}+\cdots+b_m(y) \tag{10}$$

其中 $a_i(y),b_j(y),i=0,1,\cdots,n,j=0,1,\cdots,m$ 都是 $y$ 的多项式.

如果 $(x_0,y_0)$ 是方程组(8)在复数域中的一个解，则 $x_0$ 是 $x$ 的复系数多项式 $f(x,y_0)$ 与 $g(x,y_0)$ 的一个公共根. 据定理 7.12.1，$\mathrm{Res}(f(x,y_0),g(x,y_0))=0$. 令

$$R_x(f,g)=$$

$$\left|\begin{array}{ccccccccccc}
a_0(y) & a_1(y) & \cdots & \cdots & \cdots & \cdots & a_n(y) & & & & \\
 & a_0(y) & a_1(y) & \cdots & \cdots & \cdots & \cdots & a_n(y) & & & \\
 & & & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\
 & & & & a_0(y) & a_1(y) & \cdots & \cdots & \cdots & \cdots & a_n(y) \\
b_0(y) & b_1(y) & \cdots & \cdots & \cdots & \cdots & b_m(y) & & & & \\
 & b_0(y) & b_1(y) & \cdots & \cdots & \cdots & \cdots & b_m(y) & & & \\
 & & & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\
 & & & & b_0(y) & b_1(y) & \cdots & \cdots & \cdots & \cdots & b_m(y)
\end{array}\right|
\begin{array}{l}
\left.\vphantom{\begin{array}{c}1\\1\\1\\1\end{array}}\right\} m\text{ 行} \\
\left.\vphantom{\begin{array}{c}1\\1\\1\\1\end{array}}\right\} n\text{ 行}
\end{array} \tag{11}$$

则 $R_x(f,g)$ 是 $y$ 的一个多项式. 不定元 $y$ 用 $y_0$ 代入，$R_x(f,g)$ 的象就是 $\mathrm{Res}(f(x,y_0),g(x,y_0))$. 因此，如果 $(x_0,y_0)$ 是方程组(8)在复数域中的一个解，则 $y_0$ 是多项式 $R_x(f,g)$ 的一个复根.

反之，如果 $y_0$ 是多项式 $R_x(f,g)$ 的一个复根，则

$$\mathrm{Res}(f(x,y_0),g(x,y_0))=0$$

据定理 7.12.1，$a_0(y_0)=b_0(y_0)=0$，或者 $f(x,y_0)$ 与 $g(x,y_0)$ 有

公共根，在后一情形，设 $x_0$ 是 $f(x,y_0)$ 与 $g(x,y_0)$ 的一个公共根，那么 $(x_0,y_0)$ 就是方程组(8)的一个解.

上述讨论给出了解二元高次方程组的一个一般方法：

先求 $R_x(f,g)$ 的所有复根，然后把 $R_x(f,g)$ 的每一个复根 $y_0$ 代入(8)，求出 $x$ 的值. 这样就可得到方程组(8)在复数域中的全部解.

由于在(8)中 $x$ 与 $y$ 的地位是对称的，所以也可以类似于(11)去定义 $R_y(f,g)$. 先求 $R_y(f,g)$ 的所有复根，然后把 $R_y(f,g)$ 的每一个复根 $x_0$ 代入(8)，求出 $y$ 的值. 从而得到方程组(8)的全部解.

**例 2** 解方程组

$$\begin{cases} x^2 - 7xy + 4y^2 + 6y - 4 = 0 \\ x^2 - 14xy + 9y^2 - 2x + 14y - 8 = 0 \end{cases} \tag{12}$$

**解** 把(12)式左端的两个多项式 $f(x,y)$, $g(x,y)$ 按 $x$ 的降幂写出：

$$\begin{aligned} f(x,y) &= x^2 - 7yx + (4y^2 + 6y - 4) \\ g(x,y) &= x^2 - (14y + 2)x + (9y^2 + 14y - 8) \end{aligned} \tag{13}$$

$$\begin{aligned} &R_x(f,g) \\ &= \begin{vmatrix} 1 & -7y & 4y^2 + 6y - 4 & 0 \\ 0 & 1 & -7y & 4y^2 + 6y - 4 \\ 1 & -14y - 2 & 9y^2 + 14y - 8 & 0 \\ 0 & 1 & -14y - 2 & 9y^2 + 14y - 8 \end{vmatrix} \\ &= (5y^2 + 8y - 4)^2 - (7y + 2)(7y^3 + 6y^2 - 12y + 8) \\ &= -24y(y - 1)(y - 2)(y + 2) \end{aligned} \tag{14}$$

于是 $R_x(f,g)$ 的 4 个根是：0, 1, 2, $-2$. 从(13)式看出，$a_0(y) = 1$，$b_0(y) = 1$. 因此对 $R_x(f,g)$ 的每一个根都可求得方程组(12)

的解.

用 $y=0$ 代入方程组(12),得

$$\begin{cases} x^2-4=0 \\ x^2-2x-8=0 \end{cases} \tag{15}$$

解方程组(15)得,$x=-2$. 因此$(-2,0)$是方程组(12)的一个解.

用 $y=1$ 代入(12),解得 $x=1$;用 $y=2$ 代入(12),解得 $x=2$;用 $y=-2$ 代入(12),解得$x=0$. 因此$(1,1)$,$(2,2)$,$(0,-2)$也是方程组(12)的解.上述四个解是方程组(12)的全部解.

从定义1看出,为了计算 $\mathrm{Res}(f,g)$,就需要算一个 $n+m$ 阶行列式.当$n$或$m$比较大时,计算$\mathrm{Res}(f,g)$是不容易的.下面我们给出结式的一条性质,它表明:如果 $f(x)$(或 $g(x)$)的复根容易求出,那么就可以很快求出 $\mathrm{Res}(f,g)$.

**性质1** 设 $f(x)$,$g(x)$ 同定义1中所设,并且$a_0\neq 0$,$b_0\neq 0$.如果 $f(x)$,$g(x)$ 在 $C[x]$ 中的分解式为

$$\begin{aligned} f(x)&=a_0(x-\alpha_1)\cdots(x-\alpha_n) \\ g(x)&=b_0(x-\beta_1)\cdots(x-\beta_m) \end{aligned} \tag{16}$$

那么

$$\begin{aligned} \mathrm{Res}(f,g)&=a_0^m\prod_{i=1}^{n}g(\alpha_i)=(-1)^{mn}b_0^n\prod_{j=1}^{m}f(\beta_j) \\ &=a_0^m b_0^n\prod_{i=1}^{n}\prod_{j=1}^{m}(\alpha_i-\beta_j) \end{aligned} \tag{17}$$

***证明** 在数域$K$上的$n+m+1$元多项式环$K[x,y_1,\cdots,y_n,z_1,\cdots,z_m]$中,令

$$\tilde{f}(x,y_1,\cdots,y_n)=a_0(x-y_1)\cdots(x-y_n) \tag{18}$$

$$\tilde{g}(x,z_1,\cdots,z_m)=b_0(x-z_1)\cdots(x-z_m) \tag{19}$$

把$\tilde{f}$按$x$的降幂排列,则$x^n$的系数为$a_0$,$x^{n-k}$的系数为

$$a_k(y_1,\cdots,y_n)=(-1)^k a_0\sigma_k(y_1,\cdots,y_n),\quad k=1,2,\cdots,n$$

同理,把$\tilde{g}$按$x$的降幂排列,则$x^m$的系数为$b_0$,$x^{m-k}$的系数为

$$b_k(z_1,\cdots,z_m)=(-1)^k b_0\sigma_k(z_1,\cdots,z_m),\quad k=1,2,\cdots,m$$

仿照定义 1，规定 $\widetilde{f}$ 与 $\widetilde{g}$ 对 $x$ 的结式为

$$
\mathrm{Res}_x(\widetilde{f},\widetilde{g})=
\begin{vmatrix}
a_0 & a_1(y_1,\cdots,y_n)\cdots & \cdots & \cdots a_n(y_1,\cdots,y_n) & & & \\
 & a_0 & a_1(y_1,\cdots,y_n)\cdots & \cdots & \cdots a_n(y_1,\cdots,y_n) & & \\
\cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\
 & & & a_0 & a_1(y_1,\cdots,y_n)\cdots & \cdots & \cdots a_n(y_1,\cdots,y_n) \\
b_0 & b_1(z_1,\cdots,z_m)\cdots & \cdots & \cdots b_m(z_1,\cdots,z_m) & & & \\
 & b_0 & b_1(z_1,\cdots,z_m)\cdots & \cdots & \cdots b_m(z_1,\cdots,z_m) & & \\
\cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\
 & & & b_0 & b_1(z_1,\cdots,z_m) & \cdots & \cdots\ \cdots b_m(z_1,\cdots,z_m)
\end{vmatrix}
\tag{20}
$$

显然，$\mathrm{Res}_x(\widetilde{f},\widetilde{g})\in K[y_1,\cdots,y_n,z_1,\cdots,z_m]$. 如果我们能够证明

$$\mathrm{Res}_x(\widetilde{f},\widetilde{g})=a_0^m\prod_{i=1}^{n}\widetilde{g}(y_i,z_1,\cdots,z_m)\tag{21}$$

那么不定元 $y_1,\cdots,y_n,z_1,\cdots,z_m$ 用 $\alpha_1,\cdots,\alpha_n,\beta_1,\cdots,\beta_m$ 代入，从(21)式就得到

$$\mathrm{Res}(f,g)=a_0^m\prod_{i=1}^{n}g(\alpha_i)\tag{22}$$

下面来证明(21)式成立. 首先证明 $\mathrm{Res}_x(\widetilde{f},\widetilde{g})$ 是 $K[y_1,\cdots,y_n,z_1,\cdots,z_m]$ 中的 $nm$ 次齐次多项式.

把(20)式右端的行列式 $D$ 的第 1 行乘以 $y_1$，第 2 行乘以 $y_1^2$，$\cdots$，第 $m$ 行乘以 $y_1^m$，第 $m+1$ 行乘以 $y_1$，第 $m+2$ 行乘以 $y_1^2$，$\cdots$，第 $m+n$ 行乘以 $y_1^n$，得到一个行列式 $\widetilde{D}$. 容易看出，$\widetilde{D}$ 的第 1 列元素都是 1 次齐次多项式或 0；第 2 列元素都是 2 次齐次多项式或0；$\cdots$；第 $n+m$ 列元素都是 $n+m$ 次齐次多项式或 0. 由于 $\widetilde{D}$ 的每一项是从 $\widetilde{D}$ 的第 $1,2,\cdots,n+m$ 列中各取一个元素（它们位于不同行）作成的乘积，因此 $\widetilde{D}$ 的每一个非零项的次数为

$$1+2+\cdots+(n+m)=\frac{1}{2}(n+m+1)(n+m)$$

据行列式的性质得到

$$\widetilde{D}=y_1y_1^2\cdots y_1^my_1y_1^2\cdots y_1^nD$$

从而 $D$ 的每一个非零项的次数是

$$\frac{1}{2}(n+m+1)(n+m)-(1+2+\cdots+m)$$
$$-(1+2+\cdots+n)=mn$$

这表明 $\mathrm{Res}_x(\widetilde{f},\widetilde{g})$ 是 $y_1,\cdots,y_n,z_1,\cdots,z_m$ 的 $mn$ 次齐次多项式.

其次证明：对每个 $i\in\{1,2,\cdots,n\}$，有

$$\widetilde{g}(y_i,z_1,\cdots,z_m)\mid \mathrm{Res}_x(\widetilde{f},\widetilde{g})$$

注意

$$\begin{aligned}\widetilde{g}(y_i,z_1,\cdots,z_m)&=b_0(y_i-z_1)\cdots(y_i-z_m)\\&=b_0y_i^m+\cdots+b_k(z_1,\cdots,z_m)y_i^{m-k}+\cdots+b_m(z_1,\cdots,z_m)\end{aligned}\tag{23}$$

$$\begin{aligned}\widetilde{f}(y_i,y_1,\cdots,y_n)&=a_0(y_i-y_1)\cdots(y_i-y_n)\\&=a_0y_i^n+\cdots+a_k(y_1,\cdots,y_n)y_i^{n-k}+\cdots+a_n(y_1,\cdots,y_n)\\&=0\end{aligned}\tag{24}$$

由此受到启发，我们把 $D$ 的第 1 列乘以 $y_i^{n+m-1}$，第 2 列乘以 $y_i^{n+m-2}$，$\cdots$，第 $n+m-1$ 列乘以 $y_i$，并且把它们都加到第 $n+m$ 列上，得到一个行列式 $D^*$. 于是 $D=D^*$. 利用(23)和(24)式可得出 $D^*$ 的第 $n+m$ 列为

$$\begin{pmatrix} y_i^{m-1}\cdot 0\\ y_i^{m-2}\cdot 0\\ \vdots\\ 0\\ y_i^{n-1}\widetilde{g}(y_i,z_1,\cdots,z_m)\\ y_i^{n-2}\widetilde{g}(y_i,z_1,\cdots,z_m)\\ \vdots\\ \widetilde{g}(y_i,z_1,\cdots,z_m)\end{pmatrix}$$

从 $D^*$ 的第 $n+m$ 列提出公因子 $\widetilde{g}(y_i,z_1,\cdots,z_m)$，由此得出

$$\widetilde{g}(y_i,z_1,\cdots,z_m)\mid \mathrm{Res}_x(\widetilde{f},\widetilde{g})\tag{25}$$

由于

$$\tilde{g}(y_i,z_1,\cdots,z_m) = b_0(y_i - z_1)\cdots(y_i - z_m)$$
$$\tilde{g}(y_j,z_1,\cdots,z_m) = b_0(y_j - z_1)\cdots(y_j - z_m)$$

因此当 $i \neq j$ 时,$\tilde{g}(y_i,z_1,\cdots,z_m)$ 与 $\tilde{g}(y_j,z_1,\cdots,z_m)$ 没有次数大于零的公因式. 从而

$$\prod_{i=1}^{n}\tilde{g}(y_i,z_1,\cdots,z_m) \mid \mathrm{Res}_x(\tilde{f},\tilde{g}) \tag{26}$$

由于 $\prod_{i=1}^{n}\tilde{g}(y_i,z_1,\cdots,z_m)$ 与 $\mathrm{Res}_x(\tilde{f},\tilde{g})$ 都是 $mn$ 次多项式,所以可设

$$\mathrm{Res}_x(\tilde{f},\tilde{g}) = c\prod_{i=1}^{n}\tilde{g}(y_i,z_1,\cdots,z_m) \tag{27}$$

$y_1,\cdots,y_n,z_1,\cdots,z_m$ 用 $0,\cdots,0,1,\cdots,1$ 代入,从(27)和(20)式得

$$a_0^m(-1)^{mn}b_0^n = c\prod_{i=1}^{n}b_0(-1)^m \tag{28}$$

从(28)式得,$c = a_0^m$. 代入(27)式得

$$\mathrm{Res}_x(\tilde{f},\tilde{g}) = a_0^m\prod_{i=1}^{n}\tilde{g}(y_i,z_1,\cdots,z_m) \tag{21}$$

这证明了(21)式成立,从而(22)式成立.

由定义 1 和行列式的性质容易推出

$$\mathrm{Res}(f,g) = (-1)^{mn}\mathrm{Res}(g,f)$$

因此据刚才证得的结论可以得出

$$\mathrm{Res}(f,g) = (-1)^{mn}b_0^n\prod_{j=1}^{m}f(\beta_j)$$ ▌

利用性质 1,我们来给出 $K[x]$ 中任一多项式 $f(x)$ 的判别式的求法. 设

$$f(x) = a_0x^n + a_1x^{n-1} + \cdots + a_n \in K[x], \qquad a_0 \neq 0$$

我们规定 $f(x)$ 的判别式 $D(f)$ 为:

$$\begin{aligned} D(f) :&= a_0^{2n-2}\prod_{1\leqslant j<i\leqslant n}(\alpha_i - \alpha_j)^2 \\ &= [a_0^{n-1}\prod_{1\leqslant j<i\leqslant n}(\alpha_i - \alpha_j)]^2 \end{aligned} \tag{29}$$

其中 $\alpha_1,\cdots,\alpha_n$ 是 $f(x)$ 的 $n$ 个复根.

**性质** 2　下述公式成立:

$$D(f) = (-1)^{\frac{n(n-1)}{2}} a_0^{-1} \mathrm{Res}(f,f') \tag{30}$$

**证明**　据性质 1,我们有

$$\mathrm{Res}(f,f') = a_0^{n-1} \prod_{i=1}^{n} f'(\alpha_i) \tag{31}$$

因为

$$f(x) = a_0(x-\alpha_1)(x-\alpha_2)\cdots(x-\alpha_n)$$

所以

$$f'(x) = \sum_{j=1}^{n} a_0(x-\alpha_1)\cdots(x-\alpha_{j-1})(x-\alpha_{j+1})\cdots(x-\alpha_n)$$

从而

$$\begin{aligned} f'(\alpha_i) &= a_0(\alpha_i-\alpha_1)\cdots(\alpha_i-\alpha_{i-1})(\alpha_i-\alpha_{i+1})\cdots(\alpha_i-\alpha_n) \\ &= a_0 \prod_{j\neq i}(\alpha_i-\alpha_j) \end{aligned} \tag{32}$$

于是

$$\begin{aligned} \mathrm{Res}(f,f') &= a_0^{n-1} \prod_{i=1}^{n} \left[a_0 \prod_{j\neq i}(\alpha_i-\alpha_j)\right] \\ &= a_0^{2n-1} \prod_{i=1}^{n} \prod_{j\neq i}(\alpha_i-\alpha_j) \end{aligned} \tag{33}$$

在这个乘积里,对于任意 $i$ 和 $j(i>j)$ 都出现两个因式:$\alpha_i-\alpha_j$ 和 $\alpha_j-\alpha_i$,它们的乘积为 $-(\alpha_i-\alpha_j)^2$. 由于满足条件 $1\leqslant j<i\leqslant n$ 的指标 $i$ 和 $j$ 一共有 $(n-1)+(n-2)+\cdots+1=\dfrac{n(n-1)}{2}$ 对,所以

$$\begin{aligned} \mathrm{Res}(f,f') &= a_0^{2n-1}(-1)^{\frac{n(n-1)}{2}} \prod_{1\leqslant j<i\leqslant n}(\alpha_i-\alpha_j)^2 \\ &= a_0(-1)^{\frac{n(n-1)}{2}} D(f) \end{aligned}$$

▌

**例** 3　求二次多项式 $f(x)=ax^2+bx+c$ 的判别式.

**解**　$f'(x)=2ax+b$. 于是

$$\mathrm{Res}(f,f') = \begin{vmatrix} a & b & c \\ 2a & b & 0 \\ 0 & 2a & b \end{vmatrix} = -a(b^2-4ac)$$

$$D(f) = (-1)^{\frac{2(2-1)}{2}} a^{-1}[-a(b^2-4ac)]$$

$$= b^2 - 4ac$$

利用 $f$ 和 $f'$ 的结式来求 $f$ 的判别式，比上一节讲的求首项系数为 1 的多项式的判别式的方法较简便一些.

## 习 题 7.12

1. $f(x)$ 与 $g(x)$ 同定义 1，问：$\mathrm{Res}(f,g)$ 与 $\mathrm{Res}(g,f)$ 的关系如何？

2. 判断 $f(x) = 2x^3 + 3x^2 - 8x + 3$ 与 $g(x) = 4x^2 + 7x - 15$ 在复域中有无公共根？

3. 解下列方程组：

(1) $\begin{cases} 2x^2 - xy + y^2 - 2x + y - 4 = 0 \\ 5x^2 - 6xy + 5y^2 - 6x + 10y - 11 = 0 \end{cases}$

(2) $\begin{cases} x^2 + y^2 + 4x + 2 = 0 \\ x^2 + 4xy - y^2 + 4x + 8y = 0 \end{cases}$

4. 求下列曲线的直角坐标方程：

(1) $x = t^2 - t, \qquad y = 2t^2 + t - 2$

(2) $x = \dfrac{-t^2 + 2t}{t^2 + 1}, \qquad y = \dfrac{2t^2 + 2t}{t^2 + 1}$

(提示：(2) 任取曲线上一点 $P(x,y)$，则存在实数 $t_0$，使得

$$x = \frac{-t_0^2 + 2t_0}{t_0^2 + 1}, \qquad y = \frac{2t_0^2 + 2t_0}{t_0^2 + 1}$$

即

$$(t_0^2 + 1)x + t_0^2 - 2t_0 = 0$$

$$(t_0^2 + 1)y - 2t_0^2 - 2t_0 = 0$$

这表明 $t$ 的多项式

$$f(t) = (t^2 + 1)x + t^2 - 2t, \qquad g(t) = (t^2 + 1)y - 2t^2 - 2t$$

有公共根 $t_0$，因此 $\mathrm{Res}(f,g)=0$. 这就是曲线上任一点 $P$ 的坐标 $x,y$ 适合的方程. 反之，考虑坐标适合方程 $\mathrm{Res}(f,g)=0$ 的点 $Q(x,y)$，因为 $\mathrm{Res}(f,g)=0$，所以 $x+1=0=y-2$，或者 $f(t)$ 与 $g(t)$ 有次数大于 0 的公因式. 由于 $f(t)$ 与 $g(t)$ 的次数至多为 2 次，且 $f(t)$ 与 $g(t)$ 不相伴，所以后一情形时，$f(t)$ 与 $g(t)$ 有一次公因式，从而有公共实根，于是 $Q$ 点在曲线上. 前一情形，即 $x=-1,y=2$ 时，直接验证可知 $(-1,2)$ 不是曲线上的点. 所以所求曲线的直角坐标方程为 $\mathrm{Res}(f,g)=0$，并且 $(x,y)\neq(-1,2)$.)

5. 求结式

(1) $x^4+x^3+x^2+1$ 与 $x^6+x^5+x^4+x^3+x^2+x+1$

(2) $x^n+2x+1$ 与 $x^2-x-6$

(3) $x^n+2$ 与 $(x-1)^n$

(提示:用结式的性质 1.)

6. 设 $f(x),g_1(x),g_2(x)\in C[x]$，证明

$$\mathrm{Res}(f,g_1g_2)=\mathrm{Res}(f,g_1)\mathrm{Res}(f,g_2)$$

7. 设 $f(x),x-a\in C[x]$，且 $\deg f=n$，求 $\mathrm{Res}(f,x-a)$.

8. 讨论 $f(x)=x^2+1$ 与 $g(x)=x^{2m}+1$ 是否互素.

*9. 设 $f(x)=a_0x^3+a_1x^2+a_2x+a_3\in K[x]$，$a_0\neq 0$，求 $D(f)$.

*10. 设 $f(x),g(x)\in K[x]$，证明

$$D(fg)=D(f)D(g)[\mathrm{Res}(f,g)]^2$$

*11. 设 $f(x)=x^{n-1}+x^{n-2}+\cdots+x+1$，证明

$$D(f)=(-1)^{\frac{(n-1)(n-2)}{2}}n^{n-2}$$

(提示：利用 $(x-1)f(x)=x^n-1$，第 10 题的结论，以及习题 7.11 的第 8 题的结果.)

# * §13 有理函数域

这一章从一开始到上一节，我们详细讨论了数域 $K$ 上的一元多项式环 $K[x]$ 以及多元多项式环 $K[x_1,x_2,\cdots,x_n]$ 的性质. 它们都有加法、减法与乘法运算，但是没有除法运算. $K$ 上的一元多项

式环 $K[x]$ 与整数环 $Z$ 有许多共同的性质，譬如，都有带余除法，任意两个元素都有最大公因子，都有唯一因子分解定理等. $Z$ 中也没有除法，但是把整数集 $Z$ 扩充成有理数集 $Q$ 后，在 $Q$ 中就有了加、减、乘、除四种运算（除数不为零）. 自然要问：$K[x]$ 是否也可以扩充成更大的集合，使它有除法运算呢？这一节就来讨论这个问题.

先分析一下从 $Z$ 扩充成 $Q$ 的情况，任一有理数可以写成分数 $\frac{a}{b}$ 的形式. $\frac{a}{b}$ 可以看成是一个有序整数对 $(a,b)$. 一个有理数写成分数时有多种表示方式，例如

$$\frac{2}{3}=\frac{4}{6}=\frac{6}{9}=\cdots$$

我们知道 $\frac{a}{b}=\frac{c}{d}$ 当且仅当 $ad=bc$，且 $d\neq 0,b\neq 0$. 从这里受到启发，我们来把 $K[x]$ 加以扩充.

我们用 $K[x]^*$ 表示 $K[x]$ 中所有非零多项式组成的集合. 令

$$T=K[x]\times K[x]^*=\{(f,g)\mid f\in K[x],g\in K[x]^*\}\tag{1}$$

在集合 $T$ 中规定一个二元关系 $\sim$ 如下：

$$(f_1,g_1)\sim(f_2,g_2)\text{ 当且仅当 } f_1g_2=g_1f_2\tag{2}$$

容易验证 $\sim$ 是一个等价关系. 证明如下：

任意 $(f,g)\in T$，因为 $fg=gf$，所以 $(f,g)\sim(f,g)$. 即 $\sim$ 具有反身性.

如果 $(f_1,g_1)\sim(f_2,g_2)$，则 $f_1g_2=g_1f_2$，从而 $f_2g_1=g_2f_1$，于是 $(f_2,g_2)\sim(f_1,g_1)$. 即 $\sim$ 具有对称性.

如果 $(f_1,g_1)\sim(f_2,g_2)$，且 $(f_2,g_2)\sim(f_3,g_3)$，则

$$f_1g_2=g_1f_2,\qquad f_2g_3=g_2f_3$$

从而得 $f_1g_2g_3=g_1f_2g_3=g_1g_2f_3$. 因为 $g_2\neq 0$，所以消去 $g_2$ 以后得 $f_1g_3=g_1f_3$. 由此得 $(f_1,g_1)\sim(f_3,g_3)$. 这证明了 $\sim$ 具有传递性.

我们把 $(f,g)$ 确定的等价类记作$\dfrac{f}{g}$,于是

$$\frac{f_1}{g_1}=\frac{f_2}{g_2}\Leftrightarrow(f_1,g_1)\sim(f_2,g_2)\Leftrightarrow f_1g_2=g_1f_2 \tag{3}$$

把所有等价类组成的集合记作 $K(x)$(注意这里是圆括号),即

$$K(x):=\left\{\frac{f}{g}\,\middle|\,(f,g)\in T\right\} \tag{4}$$

集合 $K(x)$ 是集合 $T$ 的一个划分,它称为 $T$ 对于关系 $\sim$ 的商集,可记作 $T/\sim$ (参看第五章 §2).

我们在商集 $K(x)$ 中定义加法与乘法运算如下:

$$\frac{f_1}{g_1}+\frac{f_2}{g_2}=\frac{f_1g_2+g_1f_2}{g_1g_2} \tag{5}$$

$$\frac{f_1}{g_1}\cdot\frac{f_2}{g_2}=\frac{f_1f_2}{g_1g_2} \tag{6}$$

我们需要说明(5)式和(6)式是有意义的,即必须说明 $K(X)$ 中两个元素(即 $\sim$ 等价类)相加(或相乘)的结果不依赖于等价类中代表元的选取. 证明如下:

设 $\dfrac{f'_1}{g'_1}=\dfrac{f_1}{g_1},\dfrac{f'_2}{g'_2}=\dfrac{f_2}{g_2}$,则

$$f'_1g_1=g'_1f_1,\qquad f'_2g_2=g'_2f_2$$

于是有

$$f'_1g_1(g'_2g_2)=g'_1f_1(g'_2g_2)$$

$$f'_2g_2(g'_1g_1)=g'_2f_2(g'_1g_1)$$

这两个式子相加得

$$g_1g_2(f'_1g'_2+f'_2g'_1)=g'_1g'_2(f_1g_2+f_2g_1)$$

所以

$$\frac{f'_1g'_2+f'_2g'_1}{g'_1g'_2}=\frac{f_1g_2+f_2g_1}{g_1g_2}$$

即

$$\frac{f'_1}{g'_1}+\frac{f'_2}{g'_2}=\frac{f_1}{g_1}+\frac{f_2}{g_2} \tag{7}$$

(7)式说明用(5)式定义 $K(x)$ 中元素的加法是合理的.

关于(6)式的合理性. 把 $f'_1g_1=g'_1f_1$ 与 $f'_2g_2=g'_2f_2$ 相乘得

$$f'_1g_1f'_2g_2=g'_1f_1g'_2f_2$$

所以 $$\frac{f'_1f'_2}{g'_1g'_2}=\frac{f_1f_2}{g_1g_2} \tag{8}$$

容易验证，上述定义的加法、乘法都满足交换律、结合律，还满足分配律. $\frac{0}{1}$ 是 $K(x)$ 中的零元素，这是因为

$$\frac{0}{1}+\frac{f_1}{g_1}=\frac{f_1}{g_1}$$

把 $\frac{0}{1}$ 记作 0. $\frac{f}{g}$ 的负元素为$\frac{-f}{g}$，这是因为

$$\frac{f}{g}+\frac{-f}{g}=\frac{fg+g(-f)}{gg}=\frac{0}{g^2}=\frac{0}{1}=0$$

把 $\frac{f}{g}$ 的负元素记作 $-\frac{f}{g}$. $\frac{1}{1}$ 是 $K(x)$ 的单位元素，这是因为

$$\frac{1}{1}\cdot\frac{f_1}{g_1}=\frac{f_1}{g_1}$$

把单位元素记成 1. 综上述 $K(x)$ 成为一个有单位元素的交换环，并且 $1\neq 0$（假如 $K(x)$ 中 $1=0$，则$\frac{1}{1}=\frac{0}{1}$. 由此推出 $K[x]$ 中的 $1=0$，矛盾）. 不仅如此，而且 $K(x)$ 中每一个非零元$\frac{f}{g}$ 都具有性质：存在 $K(x)$ 中一个元素$\frac{g}{f}$，使

$$\frac{f}{g}\cdot\frac{g}{f}=\frac{fg}{gf}=\frac{1}{1}=1,\qquad \frac{g}{f}\cdot\frac{f}{g}=\frac{gf}{fg}=\frac{1}{1}=1 \tag{9}$$

若元素 $\frac{f}{g}$ 具有性质(9)，则称$\frac{f}{g}$ 是**可逆的**，把$\frac{g}{f}$ 称为$\frac{f}{g}$ 的**逆**，记作 $\left(\frac{f}{g}\right)^{-1}$. 由于 $K(x)$ 的每个非零元都可逆，因此可以在 $K(x)$ 中定义除法如下：设$\frac{f_2}{g_2}\neq 0$，对于任一$\frac{f_1}{g_1}\in K(x)$，

$$\frac{f_1}{g_1}\div\frac{f_2}{g_2}:=\frac{f_1}{g_1}\cdot\left(\frac{f_2}{g_2}\right)^{-1} \tag{10}$$

至于 $K(x)$ 中的减法跟通常一样的方式定义

$$\frac{f_1}{g_1}-\frac{f_2}{g_2}:=\frac{f_1}{g_1}+\left(-\frac{f_2}{g_2}\right) \tag{11}$$

这样在集合 $K(x)$ 中有加、减、乘、除四种运算(除式不为 0).

现在我们说明 $K(x)$ 可看成是 $K[x]$ 的扩环,也就是说,$K[x]$ 可看成是 $K(x)$ 的子环. 理由如下:考虑 $K[x]$ 到 $K(x)$ 的一个映射

$$\sigma: f \longmapsto \frac{f}{1}$$

容易看出 $\sigma$ 是单射(假如 $\frac{f_1}{1}=\frac{f_2}{1}$,则 $f_1=f_2$). $\sigma$ 保持加法与乘法,这是因为对于 $f,g \in K[x]$,有

$$\sigma(f+g)=\frac{f+g}{1}=\frac{f}{1}+\frac{g}{1}=\sigma(f)+\sigma(g)$$

$$\sigma(fg)=\frac{fg}{1}=\frac{f}{1}\cdot\frac{g}{1}=\sigma(f)\sigma(g)$$

$\sigma$ 把 $K[x]$ 的单位元 1 映成 $K(x)$ 的单位元 1(因为 $\sigma(1)=\frac{1}{1}=1$). 这证明了 $K[x]$ 与 $K(x)$ 的一个子环 $\sigma(K[x])$ 同构,所以 $K[x]$ 可以看成是 $K(x)$ 的一个子环. 今后我们把 $\frac{f}{1}$ 与 $f$ 等同.

这样我们把数域 $K$ 上的一元多项式环 $K[x]$ 扩充成一个集合 $K(x)$,$K(x)$ 不仅是有单位元的交换环,而且 $K(x)$ 中每个非零元都可逆,从而 $K(x)$ 中可以做除法(除式不为零). $K(x)$ 中有加、减、乘、除四种运算(除式不为零),这与数域类似,我们把 $K(x)$ 称为以 $K$ 为系数域的一元**有理函数域**或数域 $K$ 上的一元**有理函数域**. 把 $K(x)$ 中的元素 $\frac{f}{g}$ 称为 $K$ 上的一元**有理函数**(或者**分式**),其中 $f$ 称为有理函数 $\frac{f}{g}$ 的**分子**,$g$ 称为**分母**. 容易验证,分子与分母乘同一个非零多项式或消去同一个非零公因式,有理函数不改变. 特别地,对于一个非零的有理函数 $\frac{f}{g}$,分子的次数减去分母的次数所得的差 $\deg f-\deg g$ 不依赖于以何种特殊的方式写形式 $\frac{f}{g}$,它由有理函数本身决定. 这个整数叫做有理函数 $\frac{f}{g}$ 的**次数**. 有理函

数 0 与零多项式是等同的，因此有理函数 0 的次数是 $-\infty$. 一个有理数，如果它的分子分母是互素的，则称它为**既约分式**. 任给一个有理函数$\dfrac{f}{g}$，我们可以在分子与分母中消去 $f$ 与 $g$ 的最大公因式，从而得到一个既约分式. 如果同一个有理函数有两个既约分式 $\dfrac{f}{g}$ 与$\dfrac{f_1}{g_1}$，则$\dfrac{f}{g}=\dfrac{f_1}{g_1}$. 从而得到 $fg_1=gf_1$. 因为$(f,g)=1$，所以 $f\mid f_1$. 因为$(f_1,g_1)=1$，所以 $f_1\mid f$. 从而 $f$ 与 $f_1$ 相伴. 同理可证 $g$ 与 $g_1$ 相伴.

如果 $\deg\left(\dfrac{f}{g}\right)=\deg f-\deg g<0$，并且$\dfrac{f}{g}$ 是既约分式，则称$\dfrac{f}{g}$ 是一个**真分式**. 有理函数 0 也被认为是真分式，因为 $\deg 0=-\infty$，并且$(0,1)=1$，即$\dfrac{0}{1}$ 是既约分式.

**定理** 7.13.1　$K(x)$中每一个有理函数都可以唯一地写成一个多项式与一个真分式的和.

**证明**　设$\dfrac{f}{g}$ 是一个既约分式，对$\dfrac{f}{g}$ 的分子和分母作带余除法，得到

$$f=gh+r,\qquad \deg r<\deg g \tag{12}$$

于是

$$\frac{f}{g}=\frac{gh+r}{g}=\frac{gh}{g}+\frac{r}{g}=\frac{h}{1}+\frac{r}{g}=h+\frac{r}{g} \tag{13}$$

因为$(f,g)=1$，所以从(12)式得$(g,r)=(f,g)=1$. 于是$\dfrac{r}{g}$ 是一个真分式. 可表性证毕.

唯一性. 假设还有$\dfrac{f}{g}=h_1+\dfrac{r_1}{g_1}$，其中 $h_1,r_1,g_1\in K[x]$，$\deg r_1<\deg g_1$，且$(r_1,g_1)=1$. 则得

$$h-h_1=\frac{r_1}{g_1}-\frac{r}{g}=\frac{r_1g-g_1r}{g_1g} \tag{14}$$

假如 $h\neq h_1$，则

$$\deg(h-h_1)\geqslant 0$$

$$\deg\left(\frac{r_1g - g_1r}{g_1g}\right) = \deg(r_1g - g_1r) - \deg(g_1g) < 0$$

矛盾. 所以 $h = h_1$. 从而 $r_1g - g_1r = 0$，由此得出 $\frac{r}{g} = \frac{r_1}{g_1}$. ▌

类似于一元有理函数域的构造方法，我们还可以构造出以 $K$ 为系数域的 $n$ 元有理函数域，记作 $K(x_1, x_2, \cdots, x_n)$.

本章从一开始到上一节讨论的数域 $K$ 上一元多项式环 $K[x]$ 以及多元多项式环 $K[x_1, x_2, \cdots, x_n]$ 的性质，除了特别声明是对复数域或实数域或有理数域而言的以外，其他性质，把数域 $K$ 换成一元或 $n$ 元有理函数域以后，仍然成立.

## 习 题 7.13

1. 验证 $K(x)$ 中加法、乘法都满足交换律、结合律，还满足分配律.

2. 设数域 $K$ 上的有理函数域 $K(x)$ 中的非零既约分式 $\frac{f}{g}$ 满足方程

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

其中 $a_i(x) \in K[x], i = 0, 1, \cdots, n$, 且 $a_0(x) \neq 0$. 证明

$$f(x) \mid a_n(x), \qquad g(x) \mid a_0(x)$$

3. 设 $K(x)$ 中非零既约分式 $\frac{f}{g}$ 满足方程

$$y^n + a_1(x)y^{n-1} + \cdots + a_{n-1}(x)y + a_n(x) = 0$$

其中 $a_i(x) \in K[x], i = 1, \cdots, n; n \geqslant 1$. 证明 $\frac{f}{g} \in K[x]$.

## §14 域的概念·有限域·域的特征

上一节讲的有理函数域 $K(x)$ 与任一数域都有加、减、乘、除四种运算(除式不为零)，并且加法、乘法满足交换律、结合律、分配律等. 其实减法由加法确定，除法由乘法以及每个非零元可逆确

定.因此 $K(x)$ 与任一数域都有加法和乘法两种运算,并且满足上述运算规律,而且每个非零元可逆.具有加法和乘法两种运算以及这些性质的集合比有单位元的交换环多一条性质:即每个非零元可逆.我们把这样的集合称为域.确切地说,

**定义 1** 一个有单位元素 $1(\neq 0)$ 的交换环 $F$,如果它的每一个非零元素都可逆,即任给 $a \in F$ 且 $a \neq 0$,存在 $b \in F$,使得 $ab = ba = 1$,则称 $F$ 是一个**域**.

容易证明,对于 $a \neq 0$,使得 $ab = ba = 1$ 的元素 $b \in F$ 是唯一的,称它是 $a$ 的**逆元**,记作 $a^{-1}$.

由于域 $F$ 的每个非零元都可逆,因此域 $F$ 没有非平凡的零因子,从而域 $F$ 一定是整环.

例如,任一数域是域.数域 $K$ 上的一元有理函数域 $K(x)$ 是一个域.任一数域有无限多个元素,一元有理函数域 $K(x)$ 也有无限多个元素(因为 $K$ 可看成是 $K[x]$ 的子环,而 $K[x]$ 又可看成是 $K(x)$ 的子环).它们称为**无限域**.$n$ 元有理函数域 $K(x_1,x_2,\cdots,x_n)$ 也是无限域.

有没有只含有限多个元素的域?回答是肯定的.下面我们要来举一个例子.

给定一个非零整数 $m$,在 $Z$ 上定义一个二元关系如下:$a \sim b$ 当且仅当 $a$ 与 $b$ 被 $m$ 除所得余数相同,即 $m|a-b$.

容易验证,$\sim$ 是一个等价关系.反身性和对称性是显然的.现在看传递性:设 $a \sim b$,且 $b \sim c$,则 $m|a-b$,且 $m|b-c$.于是 $m|(a-b)+(b-c)$,即 $m|a-c$.所以 $a \sim c$.

我们把 $a \sim b$ 记成 $a \equiv b(\mathrm{mod}\, m)$,读作 $a$ 与 $b$ 模 $m$ 同余.$a$ 确定的等价类 $\bar{a}$ 称为 $a$ 的模 $m$ **剩余类**.显然

$$\bar{a} = \{a + mk \mid k \in Z\} \tag{1}$$

因为任一整数被 $m$ 除所得余数只有 $m$ 种不同的可能性:$0,1,\cdots,m-1$,所以 $Z$ 被划分成 $m$ 个剩余类:

$$Z = \bar{0} \cup \bar{1} \cup \cdots \cup \overline{m-1} \tag{2}$$

这 $m$ 个剩余类组成的集合是 $Z$ 对于关系 $\sim$ 的商集，记作 $Z/\sim$. 为了突出模 $m$ 的剩余类组成的商集，记作 $Z/mZ$ 或 $Z/(m)$.

容易验证：如果 $a \equiv b(\mathrm{mod}m)$，$c \equiv d(\mathrm{mod}m)$，则

$$a \pm c \equiv b \pm d(\mathrm{mod}m), \qquad ac \equiv bd(\mathrm{mod}m) \tag{3}$$

特别地，对任意整数 $k$，若 $a \equiv b(\mathrm{mod}m)$，则 $ka \equiv kb(\mathrm{mod}m)$.

在集合 $Z/(m)$ 中定义加法与乘法如下：

$$\bar{a} + \bar{b} := \overline{a + b}, \qquad \bar{a} \cdot \bar{b} := \overline{ab} \tag{4}$$

我们必须说明这些定义是合理的，即与剩余类的代表元的选取无关. 设 $\bar{a} = \bar{a}_1$，$\bar{b} = \bar{b}_1$，则

$$a \equiv a_1(\mathrm{mod}m), \qquad b \equiv b_1(\mathrm{mod}m)$$

据(3)式得，$a + b \equiv a_1 + b_1(\mathrm{mod}m)$，所以 $\overline{a + b} = \overline{a_1 + b_1}$. 同理可证 $\overline{ab} = \overline{a_1 b_1}$. 所以(4)式是合理的.

容易验证，$Z/(m)$ 对于上述加法与乘法运算成为一个有单位元的交换环，称它为**模 $m$ 剩余类环**. 它的零元素是 $\bar{0}$，单位元素是 $\bar{1}$. 它一共有 $m$ 个元素：$\bar{0}, \bar{1}, \cdots, \overline{m-1}$. 我们在每一个剩余类中选取最小非负整数作为代表元，并且为记号的简洁起见，用这个代表元表示它所在的剩余类，则 $Z/(m)$ 的 $m$ 个元素可简记作 $0, 1, \cdots, m-1$.

例如，$Z/(2)$ 有两个元素，简记成：0,1. 在 $Z/(2)$ 中我们有

$$0 + 0 = 0, \qquad 0 + 1 = 1, \qquad 1 + 1 = 0$$
$$0 \cdot 0 = 0, \qquad 0 \cdot 1 = 0, \qquad 1 \cdot 1 = 1$$

又如，$Z/(3) = \{0,1,2\}$，在 $Z/(3)$ 中我们有

$$0 + i = i(i = 0,1,2), \quad 1 + 1 = 2, \quad 1 + 2 = 0, \quad 2 + 2 = 1$$
$$0 \cdot i = 0(i = 0,1,2), \quad 1 \cdot i = i(i = 1,2), \quad 2 \cdot 2 = 1$$

在上面出现的加法与乘法都是剩余类的加法与乘法，不是普通整数的加法与乘法，这是要特别注意的.

从上面的例子还可看到，$Z/(2)$ 的唯一非零元 1 是可逆的，因为 $1 \cdot 1 = 1$. $Z/(3)$ 有两个非零元：1,2. 因为 $1 \cdot 1 = 1$，$2 \cdot 2 = 1$，所以 1,2 都是可逆的. 这样，$Z/(2)$，$Z/(3)$ 就不仅是有单位元的交

换环，而且是域. 它们分别含 2 个元素，3 个元素.

只含有限多个元素的域称为**有限域**. 有一类特别简单的有限域，即

**定理** 7.14.1　设 $p$ 是任一素数，则 $Z/(p)$ 是一个域，它是含 $p$ 个元素的有限域.

**证明**　上面已证 $Z/(p)$ 是有单位元 $1(\neq 0)$ 的交换环. 剩下只要证 $Z/(p)$ 的每一个非零元 $a$ 都可逆，这里 $0 < a < p$. 因为 $p$ 是素数，而且 $0 < a < p$，所以 $(p,a) = 1$. 于是存在整数 $u,v$，使得

$$up + va = 1$$

从而 $p \mid va - 1$，即 $va \equiv 1(\mathrm{mod}\, p)$. 所以 $\overline{va} = \overline{1}$，于是 $\overline{v}\overline{a} = \overline{va} = \overline{1}$. 这证明了 $\overline{a}$ 可逆. 因此 $Z/(p)$ 是一个域. ▌

今后我们把域 $Z/(p)$ 记成 $Z_p$，称它为**模 $p$ 剩余类域**.

如果 $m$ 不是素数，则 $Z/(m)$ 不是域. 理由如下：因为 $m$ 不是素数，所以 $m = 1$ 或 $m$ 是合数. 若 $m = 1$，则 $Z/(1) = Z$，显然不是域. 若 $m$ 是合数，则 $m = m_1 m_2$，其中 $|m_i| < |m|, i = 1,2$. 于是

$$\overline{m_1}\,\overline{m_2} = \overline{m_1 m_2} = \overline{m} = \overline{0} \tag{5}$$

因为 $0 < |m_i| < |m|$，所以 $\overline{m_i} \neq \overline{0}, i = 1,2$. 于是(5)式说明 $Z/(m)$ 有非平凡的零因子. $Z/(m)$ 不是域.

有限域除了 $Z_p$ 这种类型外，还有其他的有限域，在本节习题第 4 题举了一个含 9 个元素的域的例子.

域 $F$ 的一个子环 $F_1$ 自身也是域时，就称它为 $F$ 的一个子域，也说 $F$ 是 $F_1$ 的一个扩域.

两个域称为同构是指它们作为环是同构的.

模 $p$ 剩余类域 $Z_p$ 有如下性质：

$$p\overline{1} = \underbrace{\overline{1} + \overline{1} + \cdots + \overline{1}}_{p} = \overline{p} = \overline{0} \tag{6}$$

而当 $0 < l < p$ 时，$l\overline{1} = \overline{l} \neq \overline{0}$.

任一数域 $K$，则对一切正整数 $n$，都有

$$n1 = \underbrace{1 + 1 + \cdots + 1}_{n} = n \neq 0$$

数域 $K$ 上的一元有理函数域 $K(x)$，也对一切正整数 $n$ 有 $n1 \neq 0$，这里的 1 是 $K(x)$ 的单位元素.

任意一个域情况如何？

**定理** 7.14.2 设 $F$ 是一个域，其单位元记为 $e$，则或者对一切正整数 $n$ 都有 $ne \neq 0$，或者存在一个素数 $p$，使得 $pe = 0$，而对一切 $0 < l < p$，有 $le \neq 0$. 两者只居其一.

**证明** 如果存在正整数 $n$，使 $ne = 0$，设 $n$ 是使 $ne = 0$ 成立的最小正整数. 我们来证 $n$ 一定是素数. 假如 $n$ 不是素数，由于 $n \neq 1$，于是 $n$ 必为合数. 设 $n = n_1 n_2$，其中 $0 < n_i < n$，$i = 1,2$. 于是

$$(n_1 e)(n_2 e) = (n_1 n_2)e = ne = 0 \tag{7}$$

因为 $F$ 是域，它没有非平凡的零因子，所以从(7)式得 $n_1 e = 0$ 或 $n_2 e = 0$. 这与 $n$ 的选取矛盾，所以 $n$ 一定是素数. ▎

**定义** 2 设 $F$ 是一个域，其单位元为 $e$，如果对一切正整数 $n$ 都有 $ne \neq 0$，则称域 $F$ 的**特征为零**，记作 $\mathrm{char}F = 0$. 如果存在一个素数 $p$，使得 $pe = 0$，而对于 $0 < l < p$，有 $le \neq 0$，则称域 $F$ 的**特征为** $p$，记作 $\mathrm{char}F = p$.

于是任一数域的特征为零，数域 $K$ 上的一元(或 $n$ 元)有理函数域的特征为零，而模 $p$ 剩余类域 $Z_p$ 的特征为 $p$.

若域 $F$ 的特征为零，则 $F$ 的任一非零元 $a$，对一切正整数 $n$ 都有 $na \neq 0$. 这是因为，假如有正整数 $n$，使得 $na = 0$，则

$$0 = na = n(e \cdot a) = (ne) \cdot a$$

由于 $a \neq 0$，于是推出 $ne = 0$，矛盾.

若域 $F$ 的特征为 $p$，则对于 $F$ 的任一非零元 $a$，都有 $pa = 0$. 这是因为

$$pa = p(e \cdot a) = (pe) \cdot a = 0 \cdot a = 0$$

而对于 $0 < l < p$，有 $la \neq 0$(容易看出这点). 因此在特征为 $p$ 的域 $F$ 中进行运算时要特别小心，注意虽然 $a \neq 0$，但是 $pa = 0$，即 $a$ 的 $p$ 倍是零元素.

今后我们仍然把任一域 $F$ 的单位元素记成 1.

本章关于数域 $K$ 上的一元多项式环与多元多项式环的绝大部分内容在把数域 $K$ 换成任意域时仍然正确(参看补充题七). 只是要注意：如果命题的证明中要用到这个域包含无穷多个元素，那么该命题对于有限域上的多项式就不成立了. 例如，当 $F$ 为有限域时，$F[x]$ 中两个不相等的多项式确定的多项式函数却有可能相等. 譬如，$Z_3$ 上两个多项式 $f(x)=x^3-x$ 与 $g(x)=0$ 不相等，但是它们确定的多项式函数却相等：都是零函数.

## 习 题 7.14

1. 证明：域 $F$ 中没有非平凡的零因子，从而域一定是整环.

2. 写出模 5 剩余类域 $Z_5$ 中每个非零元的逆元.

* 3. 令

$$F=\left\{\begin{pmatrix} a & b \\ -b & a \end{pmatrix}\middle| a,b\in R\right\}$$

证明：$F$ 对于矩阵的加法与乘法成为一个域，并且域 $F$ 与复数域同构.

4. 令

$$F=\left\{\begin{pmatrix} a & b \\ -b & a \end{pmatrix}\middle| a,b\in Z_3\right\}$$

证明：$F$ 是一个有 9 个元素的域，并且 $\mathrm{char}F=3$.

5. 证明：在特征为 $p$ 的域 $F$ 中，下式成立：

$$(a+b)^p=a^p+b^p$$

6. 写出 $Z_2[x]$ 中所有一次多项式和二次不可约多项式.

7. 设 $f(x)=x^5-x^2+1\in Z[x]$，判断 $f(x)$ 在 $Q$ 上是否不可约.

(提示：用 $\bar{f}(x)$ 表示 $f(x)$ 的各项系数模 2 以后得到的 $Z_2[x]$ 中的多项式. 若 $\bar{f}$ 在 $Z_2[x]$ 中不可约且 $f$ 的首项系数不能被 2 整除，则 $f(x)$ 在 $Q$ 上不可约.)

8. 设 $f(x)=x^4+3x^3+3x^2-5\in Z[x]$，判断 $f(x)$ 在 $Q$ 上是否不可约.

(提示：记号 $\bar{f}(x)$ 同第 7 题，利用 $Z_2[x]$ 中唯一因式分解定理可以证明：

若 $f(x)$ 在 $Q$ 上可约，则 $f(x)$ 在 $Q[x]$ 中必有一次因式.）

# 补 充 题 七

1. 设 $F$ 是一个域. 仿照本章 §1 的定义 1 可以给出系数在域 $F$ 中的一元多项式的概念. 用 $F[x]$ 表示域 $F$ 上所有一元多项式组成的集合，类似于 §1 定义 2，可以在 $F[x]$ 中定义加法和乘法两种运算，并且同样可以验证 $F[x]$ 是一个有单位元的交换环. 命题 7.1.1，命题 7.1.2，以及乘法消去律对于 $F[x]$ 也成立. $F[x]$ 也具有通用性质，即定理 7.1.1 对于 $F[x]$ 也成立. 参照数域 $K$ 上一元多项式环 $K[x]$ 的有关定理的证明方法，证明在 $F[x]$ 中，下述定理（或命题、性质）仍然成立：

(1) 带余除法；

(2) 整除的基本性质；

(3) $f(x)$ 与 $g(x)$ 相伴的充分必要条件是 $f(x)=cg(x)$，其中 $c$ 是 $F$ 中某个非零元；

(4) 对于 $F[x]$ 中任意两个多项式 $f(x)$ 与 $g(x)$，存在它们的一个最大公因式 $d(x)$，并且 $d(x)$ 可以表成 $f(x)$ 与 $g(x)$ 的一个组合，即有 $F[x]$ 中多项式 $u(x)$ 与 $v(x)$，使

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

(5) $F[x]$ 中两个多项式 $f(x)$ 与 $g(x)$ 互素的充分必要条件是有 $F[x]$ 中的多项式 $u(x)$，$v(x)$，使

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

(6) 在 $F[x]$ 中，如果 $f(x)\mid g(x)h(x)$，并且 $(f(x),g(x))=1$，则 $f(x)\mid h(x)$.

2. 设 $F$ 是一个域，$F[x]$ 中一个次数大于零的多项式 $p(x)$ 如果在 $F[x]$ 中的因式只有 $F$ 中的非零元以及 $p(x)$ 的相伴元，则称 $p(x)$ 是域 $F$ 上的一个不可约多项式；否则叫做可约多项式.

(1) 设 $p(x)$ 是域 $F$ 上的一个不可约多项式. 证明：如果在 $F[x]$ 中，$p(x)\mid f(x)g(x)$，那么 $p(x)\mid f(x)$ 或者 $p(x)\mid g(x)$.

(2) 证明 $F[x]$ 中有唯一因式分解定理.

(3) $F[x]$ 中的 $n$ 次多项式 $f(x)$ 的导数 $f'(x)$ 是否一定是 $n-1$ 次多项式？

(提示：区分 $\mathrm{char}F\mid n$ 和 $\mathrm{char}F\nmid n$ 两种情况.)

(4) 在 $F[x]$ 中，设不可约多项式 $p(x)$ 是 $f(x)$ 的一个 $k(k\geqslant 1)$ 重因式，证明：

1) 如果 $\mathrm{char}F\nmid k$，则 $p(x)$ 是 $f'(x)$ 的 $k-1$ 重因式；

2) 如果 $\mathrm{char}F\mid k$，则 $p(x)$ 至少是 $f'(x)$ 的 $k$ 重因式.

(5) 证明：$F[x]$ 中次数大于零的多项式 $f(x)$ 没有重因式的充分必要条件是

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

3. 设 $F$ 是一个域.

(1) 证明：在 $F[x]$ 中，用一次多项式 $x-a$ 去除多项式 $f(x)$，所得的余式是 $F$ 中一个元素 $f(a)$；

(2) 设 $f(x)\in F[x]$，$a\in F$. 证明：$a$ 是 $f(x)$ 在 $F$ 中的根的充分必要条件为 $x-a$ 是 $f(x)$ 在 $F[x]$ 中的一次因式；

(3) 证明：$F[x]$ 中的 $n$ 次 $(n\geqslant 0)$ 多项式 $f(x)$ 在 $F$ 中至多有 $n$ 个根（重根按重数计算）.

4. 设 $F$ 是一个域，仿照本章 §10 的定义 1 可以给出系数在域 $F$ 中的 $n$ 元多项式的概念. 用 $F[x_1,x_2,\cdots,x_n]$ 表示域 $F$ 上所有 $n$ 元多项式组成的集合，类似于 $K[x_1,x_2,\cdots,x_n]$，可以在 $F[x_1,x_2,\cdots,x_n]$ 中定义加法和乘法两种运算，并且 $F[x_1,x_2,\cdots,x_n]$ 成为有单位元的交换环.

(1) 证明：$F[x_1,x_2,\cdots,x_n]$ 是无零因子环；

(2) 设 $f(x_1,x_2,\cdots,x_n), g(x_1,x_2,\cdots,x_n)\in F[x_1,x_2,\cdots,x_n]$. 证明

$$\deg fg=\deg f+\deg g$$

(3) 证明在 $F[x_1,x_2,\cdots,x_n]$ 中也有通用性质，即 §10 中定理 7.10.3 对于 $F[x_1,x_2,\cdots,x_n]$ 也成立.

5. 设 $F$ 是一个域. 仿照本章 §13 的方法，从 $F[x]$ 出发，可以构造出以 $F$ 为系数域的一元有理函数域 $F(x)$. $F(x)$ 的元素 $\dfrac{f(x)}{g(x)}$ 称为 $F$ 上的一元有理函数（或分式）. 证明：如果 $\mathrm{char}F=p$（$p$ 是素数），那么 $F(x)$ 是一个特征为 $p$ 的无限域.

6. 设 $F$ 是一个域，设 $f(x_1,x_2,\cdots,x_n)\in F[x_1,x_2,\cdots,x_n]$. 如果对于每一

个 $n$ 元置换 $\sigma$，都有 $\tilde{\sigma}f = f$，则称 $f(x_1, x_2, \cdots, x_n)$ 是域 $F$ 上的一个 $n$ 元对称多项式. 令

$$\sigma_1(x_1, x_2, \cdots, x_n) = x_1 + x_2 + \cdots + x_n$$

$$\cdots \quad \cdots \quad \cdots \quad \cdots \quad \cdots \quad \cdots \quad \cdots \quad \cdots$$

$$\sigma_k(x_1, x_2, \cdots, x_n) = \sum_{1 \leqslant i_1 < \cdots < i_k \leqslant n} x_{i_1} x_{i_2} \cdots x_{i_k}$$

$$\cdots \quad \cdots \quad \cdots \quad \cdots \quad \cdots \quad \cdots \quad \cdots \quad \cdots$$

$$\sigma_n(x_1, x_2, \cdots, x_n) = x_1 x_2 \cdots x_n$$

易知，$\sigma_1, \cdots, \sigma_n$ 都是域 $F$ 上的 $n$ 元对称多项式，称它们为域 $F$ 上的 $n$ 元初等对称多项式. 证明：对于域 $F$ 上的任意一个 $n$ 元对称多项式 $f(x_1, x_2, \cdots, x_n)$，都存在 $F[x_1, x_2, \cdots, x_n]$ 中一个多项式 $g(x_1, x_2, \cdots, x_n)$，使得

$$f(x_1, x_2, \cdots, x_n) = g(\sigma_1, \sigma_2, \cdots, \sigma_n)$$

# 第八章 线性空间

本书的前七章是从代数方程出发展开讨论的. 第一章至第六章从线性方程组的求解问题入手,阐述了线性方程组的理论、矩阵理论、以及 $n$ 元有序数组的向量空间 $K^n$ 的结构. 第七章源于高次方程(组)的求解问题,阐述了一元多项式与多元多项式的理论,并且介绍了环与域的基本概念.

第一章至第六章实际上给出了**线性代数**的初步轮廓. 从这一章开始直到第十四章,我们要系统地阐述线性代数的理论. 什么是线性代数?线性代数是**线性空间**和**线性映射**的理论. 从更一般的观点来说,线性代数是研究表达自然科学中最普遍的观念之一——**线性观念**的数学语言. 自然界中,在几乎每一个小范围内差不多任何自然变化的过程都是线性的. 因此,线性代数的理论有着广泛的实际应用.

从这一章开始的研究方法与前七章不同. 前七章研究的是线性方程组、矩阵、多项式等具体对象. 这一章开始研究的线性空间,它的元素是抽象的,即不考虑这些元素的具体属性,而只考虑这些元素可以进行哪些代数运算,满足什么运算规律. 我们从许许多多的具体事物抽象出线性空间的概念,把它具有的最基本的性质作为公理给出,然后从这些公理出发进行逻辑推理,逐步深入揭示出线性空间的其他性质,研究线性空间的结构. 这种研究方法称为**公理化方法**. 它使所得出的结果有广泛的适用性.

## §1 线性空间的定义与简单性质

我们要从众多的具体对象中抽象出线性空间的概念.

**例** 1　在解析几何中,我们讲了空间中的向量可以按三角形法则相加,还可以跟实数相乘,并且加法适合结合律,交换律,有零向量,每个向量有负向量(即反向量);数乘向量满足:

$$1\vec{\alpha}=\vec{\alpha},\qquad \lambda(\mu\vec{\alpha})=(\lambda\mu)\vec{\alpha}$$

$$(\lambda+\mu)\vec{\alpha}=\lambda\vec{\alpha}+\mu\vec{\alpha},\qquad \lambda(\vec{\alpha}+\vec{\beta})=\lambda\vec{\alpha}+\lambda\vec{\beta}$$

**例** 2　在本书第三章中,我们讲了数域 $K$ 上的 $n$ 元有序数组可以做加法(将对应的分量相加),数量乘法(将数 $k$ 乘以每个分量),并且加法与数量乘法满足类似于例 1 的八条运算规则.

**例** 3　第四章中我们讲了数域 $K$ 上的 $s\times n$ 矩阵可以做加法(将对应的元素相加),数量乘法(将数 $k$ 乘矩阵的每一个元素),并且加法与数量乘法满足类似于例 1 的八条运算规则.

**例** 4　第七章中我们讲了数域 $K$ 上的一元多项式可以做加法(将同次项的系数相加). 由于 $K$ 中的元素可看成是 $K[x]$ 中的元素,于是由一元多项式的乘法可得出 $K$ 中元素与一元多项式的乘法,即数量乘法. 一元多项式的加法满足类似于例 1 的四条运算规则. 由一元多项式的乘法的运算规则可得出数量乘法满足类似于例 1 的四条运算规则.

**例** 5　第七章中我们还讲了数域 $K$ 上的 $n$ 元多项式可以做加法和乘法,从乘法可得出数量乘法. 并且加法和数量乘法也满足类似于例 1 的八条运算规则.

上面这些例子考虑的对象完全不同,但是它们有一个共同点,即它们都有加法和数量乘法这两种运算,并且这两种运算虽然在各类对象中的具体实现不同,但是它们都满足八条运算规则. 为了把这些不同的对象统一起来研究它们的有关运算的性质,我们就

抽象出线性空间的概念.

**定义** 1 设 $V$ 是一个非空集合,$F$ 是一个域. 在集合 $V$ 中定义了一种代数运算,叫做**加法**,即给出了 $V \times V$ 到 $V$ 的一个映射:$(\alpha,\beta) \longmapsto \gamma$,把$(\alpha,\beta)$ 在这个映射下的象 $\gamma$ 称为 $\alpha$ 与 $\beta$ 的**和**,记作 $\gamma = \alpha + \beta$. 在域 $F$ 与集合 $V$ 的元素之间定义了一种运算,叫做**纯量乘法**(当 $F$ 为数域时,也叫**数量乘法**),即给出了 $F \times V$ 到 $V$ 的一个映射:$(k,\alpha) \longmapsto \delta$,把$(k,\alpha)$ 在此映射下的象 $\delta$ 称为 $k$ 与 $\alpha$ 的**纯量乘积**(当 $F$ 为数域时,也叫**数量乘积**),记作 $\delta = k\alpha$. 如果加法与纯量乘法满足下述规则,那么 $V$ 称为域 $F$ 上的一个**线性空间**:

① $\alpha + \beta = \beta + \alpha, \forall\ \alpha,\beta \in V$;

② $(\alpha + \beta) + \gamma = \alpha + (\beta + \gamma), \forall\ \alpha,\beta,\gamma \in V$;

③ 在 $V$ 中有一个元素 0,使得 $\alpha + 0 = \alpha, \forall\ \alpha \in V$.(具有这个性质的元素 0 称为 $V$ 的**零元素**);

④ 对于 $V$ 中每一个元素 $\alpha$,都存在一个元素 $\beta \in V$,使得 $\alpha + \beta = 0$.($\beta$ 称为 $\alpha$ 的**负元素**);

⑤ 对于 $V$ 中任意元素 $\alpha$,有 $1\alpha = \alpha$;

⑥ $k(l\alpha) = (kl)\alpha, \forall\ k,l \in F, \forall\ \alpha \in V$;

⑦ $(k + l)\alpha = k\alpha + l\alpha, \forall\ k,l \in F, \forall\ \alpha \in V$;

⑧ $k(\alpha + \beta) = k\alpha + k\beta, \forall\ k \in F, \forall\ \alpha,\beta \in V$.

我们将以上八条规则作为线性空间的公理.

从例 1 至例 5 知道,几何空间中所有向量组成的集合是实数域上的一个线性空间;$K^n$ 是数域 $K$ 上的一个线性空间;数域 $K$ 上所有 $s \times n$ 矩阵组成的集合 $M_{s\times n}(K)$ 是数域 $K$ 上的一个线性空间;$K[x]$ 和 $K[x_1,\cdots,x_n]$ 都是数域 $K$ 上的线性空间.

下面我们再举一些例子,从中可以看出线性空间确实概括了许许多多的具体对象.

**例** 6 由于 $K[x_1,\cdots,x_n]$ 中两个 $m$ 次齐次多项式的和是 $m$ 次齐次多项式或零多项式,$K$ 中元素与 $m$ 次齐次多项式的乘积是 $m$ 次齐次多项式或零多项式,因此 $K[x_1,\cdots,x_n]$ 中所有 $m$ 次齐次多

项式添上零多项式组成的集合构成数域 $K$ 上一个线性空间(公理①至⑧不必验证,因为现在考虑的集合是线性空间 $K[x_1,\cdots,x_n]$ 的子集,并且该子集对加法和数量乘法封闭,从而这八条公理自动成立).

**例** 7　由于 $K[x_1,\cdots,x_n]$ 中两个对称多项式的和仍是对称多项式,$K$ 中元素与对称多项式的乘积仍是对称多项式,因此 $K[x_1,\cdots,x_n]$ 中所有对称多项式组成的集合构成数域 $K$ 上一个线性空间(公理①至⑧不必验证,理由同例 6).

**例** 8　设 $X$ 为任意的集合,$F$ 是任一域. 从 $X$ 到 $F$ 的每一个映射 $f$ 称为 $X$ 上的一个($F$ 值)函数. 我们把 $X$ 上的所有($F$ 值)函数组成的集合记作 $F^X$. 在 $F^X$ 中定义加法与纯量乘法运算如下:对于 $f,g \in F^X$,$k \in F$,规定

$$(f+g)(x):=f(x)+g(x), \qquad \forall\, x \in X \tag{1}$$

$$(kf)(x):=k(f(x)), \qquad \forall\, x \in X \tag{2}$$

容易验证所有的公理①至⑧都成立. 因此 $F^X$ 是域 $F$ 上的一个线性空间,$F^X$ 的零元素是零函数 0,即 $0(x)=0,\forall\, x \in X$. (1) 式称为函数的加法,(2) 式称为 $F$ 的元素与函数的纯量乘法.

**例** 9　设 $X$ 为实数域 $R$ 的任一子集,据例 8,$X$ 上的所有(实值)函数组成的集合 $R^X$ 按照函数的加法以及实数与函数的数量乘法,构成实数域 $R$ 上的一个线性空间.

**例** 10　设 $[a,b]$ 是实数轴上的一个闭区间,$[a,b]$ 上的连续函数全体记为 $C[a,b]$. 从数学分析课程知道,$[a,b]$ 上的两个连续函数的和仍是连续函数,实数 $k$ 与连续函数 $f$ 的数量乘积 $kf$ 仍是连续函数,因此 $C[a,b]$ 是实数域上的一个线性空间(公理①至⑧不必验证,理由类似于例 6).

**例** 11　类似于例 10,区间 $[a,b]$ 上所有 $n$ 次可微函数(即有 1 阶,2 阶,$\cdots$,$n$ 阶导数的函数)组成的集合是实数域上的一个线性空间,记作 $C^{(n)}[a,b]$.

**例** 12　数域 $K$ 上一元多项式函数组成的集合 $K_{\mathrm{pol}}$ 是数域 $K$

上的一个线性空间.

**例** 13　数域 $K$ 上 $n$ 元多项式函数组成的集合 $K_{\text{npol}}$ 是数域 $K$ 上的一个线性空间.

例 8 至例 13 表明,许多函数类都是线性空间.线性空间的结构虽然不能概括这些函数类的全部内容,却常常能反映出它们的一些重要性质.因此线性空间的概念在现代函数论中起着重要的作用.把数学分析的各种具体问题抽象到线性空间和拓扑空间的形式中进行研究,这正是泛函分析这门数学分支处理分析问题的特色.

**例** 14　数域 $K$ 上所有次数小于 $n$ 的一元多项式组成的集合是 $K$ 上的一个线性空间(请读者思考为什么?),记作 $K[x]_n$.

**例** 15　域 $F$ 可以看成是域 $F$ 上的一个线性空间,其加法就是域 $F$ 中的加法,其纯量乘法是域 $F$ 中的乘法.此时,公理①至⑧可以从域的加法和乘法满足的运算规则得出.

**例** 16　复数域 $C$ 可以看成是实数域 $R$ 上的一个线性空间,其加法是复数的加法,其数量乘法是实数 $a$ 与复数 $z$ 按复数乘法相乘.

**例** 17　考虑模 $p$ 剩余类域 $Z_p$ 上的有序 $n$ 元组的全体

$$Z_p^n = \{(a_1,a_2,\cdots,a_n) \mid a_i \in Z_p, i = 1,2,\cdots,n\}$$

定义加法与纯量乘法如下:

$$(a_1,\cdots,a_n) + (b_1,\cdots,b_n) := (a_1 + b_1,\cdots,a_n + b_n)$$

$$k(a_1,\cdots,a_n) := (ka_1,\cdots,ka_n)$$

易验证公理①至⑧成立,因此 $Z_p^n$ 是域 $Z_p$ 上的一个线性空间.不难看出,这个线性空间只含有限多个元素,它的元素数目是 $p^n$.例如,$Z_2^3$ 含 8 个元素:

$$(0,0,0),\quad (1,0,0),\quad (0,1,0),\quad (0,0,1)$$

$$(1,1,0),\quad (1,0,1),\quad (0,1,1),\quad (1,1,1)$$

上述例子表明,线性空间的概念适用性很广.从现在开始,我们将从线性空间的八条公理出发,作逻辑推理,深入揭示线性空间

的各种性质，这些性质对所有的具体的线性空间都成立．应当特别注意的是，在证明线性空间的各种性质时，只能用定义 1 中的八条公理以及由这些公理已经推导出的结论，绝不能用某些具体的线性空间（譬如 $K^n$）中元素的具体特性以及我们对这些具体的线性空间已经证明了的结论．虽然某些具体的线性空间（譬如，几何空间或者 $n$ 元有序数组的向量空间 $K^n$）中已知的事实可以给我们启发，但是这些不能代替证明．为了强调这一点，我们在讨论线性空间的重要性质之前，从公理 ① 至 ⑧ 推导出几个在几何空间中或 $K^n$ 中熟知的结论．

设 $V$ 是域 $F$ 上的任一线性空间．

1．$V$ 中零元素是唯一的．

**证明** 假设 $0_1, 0_2$ 是 $V$ 中两个零元素．因为 $0_2$ 是 $V$ 中零元素，据公理 ③ 得，$0_1 + 0_2 = 0_1$．又因为 $0_1$ 是 $V$ 中零元素，据公理 ① 和 ③ 得，$0_1 + 0_2 = 0_2 + 0_1 = 0_2$．于是 $0_1 = 0_2$． ▌

2．$V$ 中每个元素 $\alpha$ 的负元素是唯一的．

**证明** 设 $\beta_1$ 和 $\beta_2$ 都是 $\alpha$ 的负元素，据公理 ④，有 $\alpha + \beta_1 = 0$，$\alpha + \beta_2 = 0$．据公理 ①，② 和 ③ 得

$$\begin{aligned}\beta_1 &= \beta_1 + 0 = \beta_1 + (\alpha + \beta_2) = (\beta_1 + \alpha) + \beta_2 \\ &= (\alpha + \beta_1) + \beta_2 = 0 + \beta_2 = \beta_2\end{aligned}$$ ▌

$\alpha$ 的负元素记为 $-\alpha$．

利用负元素，可以定义**减法**如下：

$$\alpha - \beta := \alpha + (-\beta)$$

3．对于 $V$ 中任意两个元素 $\alpha, \beta$，方程 $\alpha + x = \beta$ 有唯一解

$$x = \beta - \alpha$$

**证明** 因为

$$\begin{aligned}&\alpha + (\beta - \alpha) = \alpha + (\beta + (-\alpha)) \\ &= (\alpha + (-\alpha)) + \beta = 0 + \beta = \beta\end{aligned}$$

所以 $\beta - \alpha$ 是方程 $\alpha + x = \beta$ 的一个解．

如果 $\gamma$ 是方程 $\alpha + x = \beta$ 的一个解，则 $\alpha + \gamma = \beta$．从而

$$\gamma = \gamma + 0 = \gamma + (\alpha + (-\alpha)) = (\alpha + \gamma) + (-\alpha)$$
$$= \beta + (-\alpha) = \beta - \alpha \qquad \blacksquare$$

4. $0\alpha = 0, \forall \alpha \in V$.（注意：等号两边的"0"代表不同的对象）.

**证明** $0\alpha = (0+0)\alpha = 0\alpha + 0\alpha$. 据上述性质 3，得

$$0\alpha = 0\alpha - 0\alpha = 0 \qquad \blacksquare$$

5. $k0 = 0, \forall k \in F$.

**证明** $k0 = k(0+0) = k0 + k0$. 据性质 3，得

$$k0 = k0 - k0 = 0 \qquad \blacksquare$$

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

**证明**

$$\alpha + (-1)\alpha = 1\alpha + (-1)\alpha = (1 + (-1))\alpha = 0\alpha = 0$$

从 $\alpha$ 的负元素的唯一性得，$(-1)\alpha = -\alpha$. $\blacksquare$

7. 如果 $k\alpha = 0$，那么 $k = 0$ 或者 $\alpha = 0$.

**证明** 假设 $k \neq 0$，则

$$\alpha = 1\alpha = (k^{-1}k)\alpha = k^{-1}(k\alpha) = k^{-1}0 = 0 \qquad \blacksquare$$

习惯上，借用几何的语言，把线性空间的元素称为向量，把线性空间称为向量空间. 注意不要把现在讲的向量与以前讲的 $n$ 元有序数组（它也叫向量）混淆. 域 $F$ 中的元素称为纯量. 域 $F$ 上的线性空间 $V$ 的加法和纯量乘法统称为线性运算.

本书用 $F$ 表示任一域，用 $K$ 表示数域，不再每次声明.

## 习　题　8.1

1. 检验以下集合对于所指的线性运算是否构成数域 $K$ 上的线性空间：

(1) $K[x]$ 中 $n$ 次多项式的全体，对于 $K[x]$ 中的加法和数量乘法；

(2) 设 $A$ 是 $K$ 上的一个 $n$ 级矩阵，$K[A]$ 对于矩阵的加法和数量乘法；

(3) $K$ 上的全体 $n$ 级对称（斜对称，上三角）矩阵，对于矩阵的加法和数

量乘法；

(4) $K[x]$ 中次数不超过 $n$ 的多项式的全体，对于 $K[x]$ 中的加法和数量乘法.

2. 检验以下集合对于所指的线性运算是否构成实数域 $R$ 上的线性空间：

(1) 全体正实数 $R^+$，加法与数量乘法定义为

$$a \oplus b = ab, \forall\ a,\quad b \in R^+$$

$$k \odot a = a^k,\quad \forall\ a \in R^+, k \in R$$

(2) 设 $V$ 是定义域为 $[a,b]$，伴域为正实数集的所有函数组成的集合，加法与数量乘法定义为

$$f \oplus g = fg,\quad \forall\ f, g \in V$$

$$k \odot f = f^k,\quad \forall\ f \in V, k \in R$$

其中 $f^k(x) := [f(x)]^k, \forall\ x \in [a,b]$.

3. 设 $V$ 是复数域 $C$ 上的一个线性空间，如果加法保持不变，而数量乘法改成

$$k \cdot a = \bar{k}\alpha$$

其中 $\bar{k}$ 是 $k$ 的共轭复数，问：集合 $V$ 对于原来的加法和现在定义的数量乘法是否构成复数域 $C$ 上的一个线性空间.

4. 实数集 $R$ 的下列子集对于通常的加法和数量乘法，是否形成有理数域 $Q$ 上的线性空间：

(1) 全体正实数 $R_r^+$；　(2) 全体负实数；　(3) $\{a + b\pi \mid a, b \in Q\}$.

5. 用 $F^\infty$ 表示域 $F$ 上所有无限序列组成的集合，即

$$F^\infty = \{(a_1, a_2, a_3, \cdots) \mid a_i \in F, i = 1, 2, 3, \cdots\}$$

在 $F^\infty$ 中定义加法与纯量乘法如下：

$$(a_1, a_2, a_3, \cdots) + (b_1, b_2, b_3, \cdots) = (a_1 + b_1, a_2 + b_2, a_3 + b_3, \cdots)$$

$$k(a_1, a_2, a_3, \cdots) = (ka_1, ka_2, ka_3, \cdots)$$

验证 $F^\infty$ 是域 $F$ 上的一个线性空间.

*6. $F^\infty$ 的下列子集对于 $F^\infty$ 的加法与纯量乘法是否构成域 $F$ 上的线性空间：

(1) 只有有限多个分量不为 0 的无限序列组成的子集；

(2) 只有有限多个分量为 0 的无限序列组成的子集；

(3) 没有分量等于 1 的无限序列组成的子集.

7. $R^{\infty}$ 或 $C^{\infty}$ 的下列子集对于第 5 题定义的加法与纯量乘法(取 $F=R$ 或 $C$),是否构成 $R$ 或 $C$ 上的线性空间:

(1) 满足 Cauchy 条件的无限序列组成的子集,Cauchy 条件是:任给 $\varepsilon>0$,$\exists\ N>0$,使得只要 $m,n>N$,就有 $|a_m-a_n|<\varepsilon$.

(2) 满足 Hilbert 条件的无限序列组成的子集,Hilbert 条件是:级数 $\sum\limits_{n=1}^{\infty}|a_n|^2$ 收敛;

(3) 有界的无限序列组成的子集,一个无限序列 $(a_1,a_2,\cdots)$ 称为有界,如果存在一个实数 $c$,使得 $|a_i|<c,\forall\ i$.

* 8. 设 $X$ 是任一集合,$F$ 是一个域. $F^X$ 的下列子集对于函数的加法以及 $F$ 的元素与函数的纯量乘法,是否构成域 $F$ 上的线性空间:

(1) 给定 $a\in X$,集合 $\{f\in F^X\mid f(a)=0\}$;

(2) 给定 $a\in X$,给定 $k\in F$,集合 $\{f\in F^X\mid f(a)=k\}$;

(3) 给定 $X_0\subset X$,集合 $\{f\in F^X\mid f(x)=0,\forall\ x\in X_0\}$;

(4) 给定 $X_0\subset X$,集合 $\{f\in F^X\mid f$ 在 $X_0$ 的至少一个点上的值为 $0\}$.

* 9. $R^R$ 的下列子集对于函数的加法以及实数与函数的数量乘法,是否构成实数域 $R$ 上的线性空间:

(1) $\{f\in R^R\mid \lim\limits_{|x|\to\infty}f(x)=0\}$;

(2) $\{f\in R^R\mid \lim\limits_{|x|\to\infty}f(x)=1\}$;

(3) $\{f\in R^R\mid f$ 只有有限多个间断点$\}$.

* 10. 证明:加法交换律可以由线性空间的其它公理推出.

(提示:分七步. 第一步,证明:若 $\beta$ 是 $\alpha$ 的负元素,则 $\beta+\alpha=0$;第二步,证明 $0+\alpha=\alpha,\forall\ \alpha\in V$;第三步,证明 $V$ 中零元素唯一;第四步,证明 $V$ 中每个元素 $\alpha$ 的负元素唯一;第五步,证明 $0\alpha=0,\forall\ \alpha\in V$;第六步,证明 $(-1)\alpha=-\alpha,\forall\ \alpha\in V$;第七步,证明 $(\alpha+\beta)+[-(\beta+\alpha)]=0$).

## §2 线性相关性与线性无关性

研究线性空间就是要把它的结构搞清楚. 由于域 $F$ 上的线性空间 $V$ 具有加法运算以及 $F$ 与 $V$ 之间的纯量乘法运算,因此我们

首先要研究 $V$ 的元素之间关于这两种运算的性质，即线性相关性与线性无关性. 这些内容在第三章中曾经对 $n$ 元有序数组的向量空间讲过一遍，现在是对一般的线性空间(其元素可以是任意的)来讲.

本节中的 $V$ 都是域 $F$ 上的线性空间，不再每次说明.

**定义 1** 设 $\alpha_1,\alpha_2,\cdots,\alpha_s$ 是 $V$ 中一组向量($s\geqslant 1$)，$k_1,k_2,\cdots,k_s$ 是 $F$ 中元素，那么向量

$$\alpha = k_1\alpha_1 + k_2\alpha_2 + \cdots + k_s\alpha_s$$

称为向量组 $\alpha_1,\alpha_2,\cdots,\alpha_s$ 的以 $k_1,k_2,\cdots,k_s$ 为系数的**线性组合**，也称向量 $\alpha$ 可以用向量组 $\alpha_1,\alpha_2,\cdots,\alpha_s$ **线性表出**.

**定义 2** $V$ 中向量组 $\alpha_1,\cdots,\alpha_s$($s\geqslant 1$) 称为是**线性相关**的，如果在 $F$ 中有不全为零的元素 $k_1,\cdots,k_s$，使得

$$k_1\alpha_1 + \cdots + k_s\alpha_s = 0 \tag{1}$$

否则，向量组 $\alpha_1,\cdots,\alpha_s$ 称为是**线性无关**的. 换句话说，向量组 $\alpha_1,\cdots,\alpha_s$($s\geqslant 1$) 称为线性无关，如果从

$$k_1\alpha_1 + \cdots + k_s\alpha_s = 0$$

可以推出 $k_1 = \cdots = k_s = 0$.

把对应于 $s=0$ 的空向量组定义成线性无关的，这可以为今后的讨论带来方便，而且它与非空向量组的定义 2 一致.

我们把(1)式称为向量组 $\alpha_1,\cdots,\alpha_s$ 的一个**线性关系**. 如果 $k_1,\cdots,k_s$ 全为零，则称此线性关系是**平凡的**；否则称为**非平凡的**.

$V$ 的任一子集称为向量集. 对于有限向量集 $W$，如果给它的元素一种编号(即给它的元素排一个次序)，写成 $\alpha_1,\cdots,\alpha_s$，则称 $\alpha_1,\cdots,\alpha_s$ 是一个向量族. 本书把向量族也称为向量组. 因此今后谈到向量组时，是指它有有限多个向量，并且排了一个次序. 显然，如果向量组 $\alpha_1,\cdots,\alpha_m$ 线性相关(线性无关)，则对于任一 $m$ 元置换 $\sigma$，向量组 $\alpha_{\sigma(1)},\cdots,\alpha_{\sigma(m)}$ 也线性相关(线性无关). 这一点确保了下述定义的正确性：

**定义 3** 有限向量集称为**线性相关(线性无关)**的，如果对其

元素有一种编号(因此每一种编号)使它成为线性相关(线性无关)的.

按上面指出的,空集是线性无关的.

1. 只含一个元素 $\alpha$ 的向量集线性无关 $\Leftrightarrow \alpha \neq 0$.

**证明** 必要性. 假如 $\alpha = 0$,则 $1\alpha = 0$,从而 $\alpha$ 线性相关.

充分性. 设 $\alpha \neq 0$. 如果 $k\alpha = 0$,由于 $\alpha \neq 0$,因此 $k = 0$. 从而 $\alpha$ 线性无关. ▌

2. 如果向量组 $\alpha_1,\cdots,\alpha_s$ 的一个部分组线性相关,则向量组 $\alpha_1,\cdots,\alpha_s$ 线性相关.

**证明** 由线性相关的定义立即得到. ▌

上述性质 2 启发我们可以作出下述定义:

**定义 4** 无限向量集 $W$ 称为**线性相关**的,如果 $W$ 有一个有限子集是线性相关的. 否则,$W$ 称为**线性无关**的,即如果 $W$ 的任何有限子集都是线性无关的,就称 $W$ 是线性无关的.

3. 包含零向量的向量集是线性相关的.

**证明** 取该集合 $W$ 中只含一个零向量的子集,由性质 1 得出这个有限子集是线性相关的,从而 $W$ 是线性相关的. ▌

4. 一个向量集是线性相关的当且仅当它的向量中至少有一个可以由其余向量中的有限个线性表出.

**证明** 必要性. 设向量集 $W$ 是线性相关的,则 $W$ 有一个有限子集 $\{\alpha_1,\alpha_2,\cdots,\alpha_s\}$ 是线性相关的. 于是有 $F$ 中不全为零的元素 $k_1,k_2,\cdots,k_s$,使得

$$k_1\alpha_1 + k_2\alpha_2 + \cdots + k_s\alpha_s = 0$$

不妨设 $k_1 \neq 0$,于是从上式得出,$\alpha_1$ 可以用 $\alpha_2,\cdots,\alpha_s$ 线性表出. 充分性由定义 2 和定义 4 立即得出. ▌

5. 向量组 $\alpha_1,\cdots,\alpha_s$ 是线性无关的当且仅当任何可以用 $\alpha_1,\cdots,\alpha_s$ 线性表出的向量,其表出方式唯一.

**证明** 与第三章 §3 命题 3.3.2 的证明一样. ▌

6. 设向量组 $\alpha_1,\cdots,\alpha_s$ 线性无关,而向量组 $\alpha_1,\cdots,\alpha_s,\beta$ 线性相

关，则 $\beta$ 可以由向量组 $\alpha_1,\cdots,\alpha_s$ 线性表出.

**证明** 与第三章 §3 命题 3.3.3 的证明一样. ▌

下面讨论 $V$ 中两个向量集之间的关系.

**定义 5** 如果向量集 $W_1$ 中每个向量可以由向量集 $W_2$ 中有限多个向量线性表出，则称向量集 $W_1$ 可以由向量集 $W_2$ **线性表出**. 如果这两个向量集可以互相线性表出，则称它们是**等价**的.

7. 如果一个向量 $\alpha$ 可以由向量集 $W_1$ 线性表出，并且向量集 $W_1$ 可以由向量集 $W_2$ 线性表出，则 $\alpha$ 可以由向量集 $W_2$ 线性表出（线性表出的传递性）.

**证明** 与第三章 §4 引理 3.4.2 的证明类似. ▌

由性质 7 立即得出，$V$ 中向量集的等价具有传递性. 显然，向量集的等价具有反身性和对称性. 因此，向量集的等价是 $V$ 中向量集之间的一个等价关系.

8. 设向量组 $\beta_1,\cdots,\beta_r$ 可以由向量组 $\alpha_1,\cdots,\alpha_s$ 线性表出. 如果 $r>s$，那么 $\beta_1,\cdots,\beta_r$ 线性相关.

**证明** 与第三章 §4 引理 3.4.1 的证明一样. 注意：第一章 §2 推论 1.2.2 对于任意域上的齐次线性方程组仍成立. ▌

这里我们要指出，本书第一章至第六章讲的线性方程组的理论、行列式的理论、数域 $K$ 上 $n$ 元有序数组的向量空间 $K^n$ 的理论和矩阵的理论中所有概念和绝大多数结论，在把数域 $K$ 换成特征不等于 2 的任意域 $F$ 后仍然成立，只是要注意识别零元素（在特征为 $p$ 的域 $F$ 中，$pa$ 是零元素，其中 $a\in F$），譬如，在等式两边乘以某个元素的逆时，要先判别这个元素是不是非零元素，因为非零元素才有逆元. 在特征等于 2 的域的情形，第一章至第六章的绝大多数概念仍有，但有不少结论不成立，因为此时，域 $F$ 中任一元素 $a$ 的 2 倍都为零，即 $2a=0$. 当然也有许多结论仍然成立. 这要具体问题具体分析.

由性质 8 立即得到

9. 等价的线性无关的向量组含有相同数目的向量. ▌

利用两个向量集(组)之间的关系的这些性质，我们来讨论向量集(组)的极大线性无关集(组)和向量组的秩这两个概念及其性质.

**定义 6** 向量集(组)的一个子集(部分组)称为一个**极大线性无关集(组)**，如果这个子集(部分组)本身是线性无关的，但是从这个向量集(组)的其余向量(如果还有的话)中任取一个添进去，得到的新的子集(部分组)都线性相关.

不难看出，向量集(组)与它的极大线性无关集(组)等价(请读者作为练习). 然后据线性表出的传递性可得出，向量集(组)的任意两个极大线性无关集(组)等价. 从而向量组的任意两个极大线性无关组含有相同数目的向量(据性质 9).

**定义 7** 向量组 $\alpha_1,\cdots,\alpha_s$ 的一个极大线性无关组包含的向量的数目称为向量组 $\alpha_1,\cdots,\alpha_s$ 的**秩**. 记作 $\operatorname{rank}\{\alpha_1,\cdots,\alpha_s\}$.

据定义 6 和性质 1，全由零向量组成的向量组 $\alpha_1,\cdots,\alpha_s$ 的极大线性无关组是空集，从而它的秩为 0.

由于每个向量组都与它的极大线性无关组等价，于是据向量组等价的对称性和传递性，以及性质 9，得到

10. 等价的向量组有相同的秩. ▌

现在我们把上面对于一般的线性空间介绍的概念和得到的性质运用到由一些函数组成的具体的线性空间上.

我们用 $C^{(n-1)}[a,b]$ 表示区间 $[a,b]$ 上 $n-1$ 次可微函数的全体，据上一节的例 11 知，$C^{(n-1)}[a,b]$ 是实数域上的一个线性空间. 下面我们将给出 $C^{(n-1)}[a,b]$ 中 $n$ 个函数 $f_1(x),f_2(x),\cdots,f_n(x)$ 线性无关的一个充分条件，为此我们先引进一个概念：下述 $n$ 级行列式

$$\begin{vmatrix} f_1(x) & f_2(x) & \cdots & f_n(x) \\ f_1'(x) & f_2'(x) & \cdots & f_n'(x) \\ \cdots & \cdots & \cdots & \cdots \\ f_1^{(n-1)}(x) & f_2^{(n-1)}(x) & \cdots & f_n^{(n-1)}(x) \end{vmatrix} \tag{2}$$

称为函数组 $f_1(x), f_2(x), \cdots, f_n(x)$ 的 Wronsky(**朗斯基**) **行列式**，它是 $x$ 的函数，记作 $W(x)$，或者 $W(f_1, f_2, \cdots, f_n)$.

**命题** 8.2.1 $C^{(n-1)}[a,b]$ 中的函数组 $f_1(x), f_2(x), \cdots, f_n(x)$，如果它的 Wronsky 行列式 $W(x)$ 在某一点 $x_0 \in [a,b]$ 处的函数值 $W(x_0) \neq 0$，则函数组 $f_1(x), f_2(x), \cdots, f_n(x)$ 线性无关.

**证明** 设

$$k_1 f_1(x) + k_2 f_2(x) + \cdots + k_n f_n(x) = 0 \tag{3}$$

上式右端的 0 是零函数，因此上式是对一切 $x \in [a,b]$ 都成立的恒等式. 在上式两边分别求 1 阶，2 阶，…，$n-1$ 阶导数，得到

$$\begin{array}{l} k_1 f_1'(x) + k_2 f_2'(x) + \cdots + k_n f_n'(x) = 0 \\ \cdots\ \cdots\ \cdots\ \cdots\ \cdots\ \cdots\ \cdots\ \cdots\ \cdots \\ k_1 f_1^{(n-1)}(x) + k_2 f_2^{(n-1)}(x) + \cdots + k_n f_n^{(n-1)}(x) = 0 \end{array} \tag{4}$$

在(3)式和(4)式中，考虑各个函数在 $x_0$ 处的函数值，得到

$$\begin{cases} k_1 f_1(x_0) + k_2 f_2(x_0) + \cdots + k_n f_n(x_0) = 0 \\ k_1 f_1'(x_0) + k_2 f_2'(x_0) + \cdots + k_n f_n'(x_0) = 0 \\ \cdots\ \cdots\ \cdots\ \cdots\ \cdots\ \cdots\ \cdots\ \cdots\ \cdots\ \cdots \\ k_1 f_1^{(n-1)}(x_0) + k_2 f_2^{(n-1)}(x_0) + \cdots + k_n f_n^{(n-1)}(x_0) = 0 \end{cases} \tag{5}$$

已知 $W(x_0) \neq 0$，因此以 $W(x_0)$ 为系数矩阵的行列式的齐次线性方程组只有零解. 从而由(5) 得

$$k_1 = k_2 = \cdots = k_n = 0$$

因此，$f_1(x), f_2(x), \cdots, f_n(x)$ 线性无关. ▌

一般说来，命题 8.2.1 所述条件 $W(x_0) \neq 0$ 只是函数组 $f_1(x), f_2(x), \cdots, f_n(x)$ 线性无关的一个充分条件，不是必要条件，这从下面的例 2 可以看出.

给了实数域上的线性空间 $R^{[a,b]}$ 中的函数组 $f_1(x), \cdots, f_n(x)$，如果它们是 $n-1$ 次可微的，那么可以先观察它们的 Wronsky 行列式 $W(x)$ 是否在某点 $x_0 \in [a,b]$ 处 $W(x_0) \neq 0$，若是，则 $f_1(x), \cdots, f_n(x)$ 线性无关. 若 $\forall\, x \in [a,b]$，都有

$W(x)=0$，则需要用定义 2 去判断它们是否线性无关(见例 2)．此外，如果 $W(x)$ 不容易计算，则可以用定义 2 去判断．

**例 1** 判断实数域上的线性空间 $R^R$ 中的函数组 $e^x\cos x,\sin x$ 是否线性无关．

**解** $e^x\cos x,\sin x$ 的 Wronsky 行列式是

$$W(x)=\begin{vmatrix} e^x\cos x & \sin x \\ e^x\cos x - e^x\sin x & \cos x \end{vmatrix}$$

在 $x=0$ 处，

$$W(0)=\begin{vmatrix} 1 & 0 \\ 1 & 1 \end{vmatrix}=1\neq 0$$

因此 $e^x\cos x,\sin x$ 线性无关．

**例 2** 判断实数域上的线性空间 $R^R$ 中的函数组 $x^2,x|x|$ 是否线性无关．

**解** 假设 $k_1x^2+k_2x|x|=0$，等式右端是零函数．

令 $x=1$，得 $k_1+k_2=0$．令 $x=-1$，得 $k_1-k_2=0$．因此，$k_1=k_2=0$．这表明 $x^2,x|x|$ 线性无关．

例 2 中的函数 $f_2(x)=x|x|$ 在 $(-\infty,\infty)$ 上有一阶导数，因此 $x^2,x|x|$ 有 Wronsky 行列式

$$W(x)=\begin{vmatrix} x^2 & x|x| \\ 2x & 2|x| \end{vmatrix}=0,\qquad \forall\, x\in(-\infty,\infty)$$

由此可见，命题 8.2.1 只给出了 $n-1$ 次可微函数组 $f_1(x),\cdots,f_n(x)$ 线性无关的一个充分条件，它不是必要条件．

## 习　题　8.2

1．判断实数域上的线性空间 $R^R$ 中的下列函数组是否线性无关：

(1) $1,\cos^2x,\cos 2x$；

(2) $1,\cos x,\cos 2x,\cos 3x$．

2. 证明:如果 $f_1(x),f_2(x),f_3(x)$ 是数域 $K$ 上线性空间 $K[x]$ 中三个互素的多项式,但其中任意两个都不互素,那么它们线性无关.

3. 证明:域 $F$ 上线性空间 $V$ 中,若向量组 Ⅰ 可由向量组 Ⅱ 线性表出,则 Ⅰ 的秩 $\leqslant$ Ⅱ 的秩.

4. 在域 $F$ 上线性空间 $V$ 中,设向量组 $\alpha_1,\cdots,\alpha_s$ 线性无关,设

$$\beta = b_1\alpha_1 + \cdots + b_s\alpha_s$$

证明:如果对于某个 $i(1 \leqslant i \leqslant s)$,$b_i \neq 0$,则用 $\beta$ 替换 $\alpha_i$ 以后得到的向量组 $\alpha_1,\cdots,\alpha_{i-1},\beta,\alpha_{i+1},\cdots,\alpha_s$ 也线性无关.

5. 证明:在实数域上的线性空间 $R^R$ 中,函数组 $e^{\lambda_1 x},e^{\lambda_2 x},\cdots,e^{\lambda_n x}$ 是线性无关的,其中 $\lambda_1,\lambda_2,\cdots,\lambda_n$ 是互不相同的实数.

6. 证明:在实数域上的线性空间 $R^R$ 中,函数组 $x^{a_1},x^{a_2},\cdots,x^{a_n}$ 是线性无关的,其中 $a_1,a_2,\cdots,a_n$ 是互不相同的实数.

(提示:用定义 2,在线性关系中令 $x = 1,2,2^2,\cdots,2^{n-1}$.)

*7. 证明:在实数域上的线性空间 $R^R$ 中,下列各函数组均线性无关:

(1) $\sin x,\sin 2x,\cdots,\sin nx$;

(2) $1,\cos x,\cos 2x,\cdots,\cos nx$;

(3) $1,\cos x,\sin x,\cos 2x,\sin 2x,\cdots,\cos nx,\sin nx$.

(提示:对 $n$ 用归纳法.在考察 $n$ 的情形时,用定义 2,对线性关系式微分两次,然后将线性关系式的适当倍数与第二次微分后的式子相加,再利用归纳假设).

*8. 证明:在实数域上的线性空间 $R^R$ 中,下列各函数组均线性无关:

(1) $1,\sin x,\sin^2 x,\cdots,\sin^n x$;

(2) $1,\cos x,\cos^2 x,\cdots,\cos^n x$.

(提示:用定义 2;在线性关系式中分别令

$$x = \frac{1}{n+1}\cdot\frac{\pi}{2},\cdots,\frac{n+1}{n+1}\cdot\frac{\pi}{2}$$

得到一个 $n+1$ 元齐次线性方程组,其系数矩阵的行列式为 Vandermonde 行列式).

9. 证明:在实数域上的线性空间 $R^R$ 中,下列各函数组均线性相关:

(1) $1,\sin x,\cos x,\sin^2 x,\cos^2 x,\cdots,\sin^n x,\cos^n x$,其中 $n \geqslant 2$;

(2) $\sin x,\cos x,\sin^2 x,\cos^2 x,\cdots,\sin^n x,\cos^n x$,其中 $n \geqslant 4$.

(提示:只要找一个部分组线性相关即可.注意利用三角函数的恒等式).

10. 判断实数域上的线性空间 $R^R$ 中下列函数组是否线性无关：

(1) $1,\sin x,\cos x$；

(2) $\sin x,\cos x,\sin^2 x,\cos^2 x$.

## §3 基·维数·坐标

在第三章中，读者已看到，在研究 $n$ 元有序数组的向量空间 $K^n$ 的结构时，“基”起了关键作用. 由此受到启发，在研究一般的线性空间时，也应引进基的概念. 如果没有特别声明时，本节中的 $V$ 是任一域 $F$ 上的线性空间.

**定义 1** 如果 $V$ 的一个向量集 $S$ 是线性无关的，并且 $V$ 中每一个向量可以由 $S$ 中有限多个向量线性表出，那么称 $S$ 是 $V$ 的一个**基**.

为了方便起见，把只含有零向量的线性空间的基规定为空集.

**例 1** 在数域 $K$ 上所有 $n\times m$ 矩阵组成的线性空间 $M_{n\times m}(K)$ 中，基本矩阵组 $E_{11},E_{12},\cdots,E_{1m},\cdots,E_{n1},\cdots,E_{nm}$ 是线性无关的（请读者验证），并且每个 $n\times m$ 矩阵 $A=(a_{ij})$ 可表成

$$A=\sum_{i=1}^{n}\sum_{j=1}^{m}a_{ij}E_{ij} \tag{1}$$

因此，$E_{11},E_{12},\cdots,E_{nm}$ 是 $M_{n\times m}(K)$ 的一个基.

**例 2** 在数域 $K$ 上的线性空间 $K[x]$ 中，考虑它的一个子集

$$S=\{x^m\mid m=0,1,2,\cdots\}$$

据一元多项式的定义可得出，$S$ 的任一有限子集是线性无关的，从而 $S$ 是线性无关的. 显然，每个一元多项式可以由 $S$ 中有限多个元素线性表出. 因此，$S$ 是 $K[x]$ 的一个基. 这个基包含无限多个元素. 我们指出，$K[x]$ 不可能有一个基是包含有限多个元素的. 理由如下：假如 $K[x]$ 有一个基包含 $r$ 个元素：$f_1(x),\cdots,f_r(x)$，设这 $r$ 个多项式的次数中最大者为 $n$，则一个 $n+1$ 次多项式 $g(x)$ 就无

法由 $f_1(x),\cdots,f_r(x)$ 线性表出，矛盾.

**定义 2** $V$ 称为**有限维**的，如果它有一个基包含有限多个向量；否则，$V$ 称为**无限维**的.

从例 1、例 2 看到，$M_{n\times m}(K)$ 是有限维的，$K[x]$ 是无限维的.

从定义 2 知道，有限维线性空间一定有一个基，并且这个基包含有限多个向量；无限维线性空间如果有一个基，则这个基一定包含无限多个向量. 但问题是：无限维线性空间是否一定有基呢？我们将在阅读材料九里，运用 Zorn 引理证明无限维线性空间一定有基.

本课程虽然着重讨论有限维线性空间，但是也讨论无限维线性空间. 从今以后，如果不加声明，则所讨论的线性空间既可以是有限维的，也可以是无限维的.

**定理 8.3.1** 在有限维线性空间中，任意两个基所包含的向量的数目相同.

**证明** 设 $V$ 是有限维的，$\alpha_1,\cdots,\alpha_s$ 与 $\beta_1,\cdots,\beta_r$ 是 $V$ 的两个基. 由定义 1，它们等价并且都线性无关，从而它们含有相同数目的向量(据 §2 性质 9). ▌

定理 8.3.1 使我们可以作出下述定义：

**定义 3** 设 $V$ 是有限维的，则 $V$ 的一个基包含的向量的数目称为 $V$ 的**维数**，记作 $\dim_F V$ 或 $\dim V$.

由定义 3 知，只含零向量的线性空间的维数为 0. 若 $V$ 是无限维的，则记 $\dim V=\infty$. 从例 1 知，$\dim M_{n\times m}(K)=nm$.

**例 3** 求数域 $K$ 上所有 $n$ 级斜对称矩阵组成的线性空间 $V$ 的一个基及其维数.

**解** 任一 $n$ 级斜对称矩阵 $A$ 具有形式

$$A=\begin{pmatrix} 0 & a_{12} & \cdots & a_{1n} \\ -a_{12} & 0 & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ -a_{1n} & -a_{2n} & \cdots & 0 \end{pmatrix}$$

因此

$$A = a_{12}(E_{12} - E_{21}) + a_{13}(E_{13} - E_{31}) + \cdots$$
$$+ a_{1n}(E_{1n} - E_{n1}) + \cdots + a_{n-1,n}(E_{n-1,n} - E_{n,n-1}) \quad (2)$$

由于

$$(E_{ij} - E_{ji})' = E'_{ij} - E'_{ji} = E_{ji} - E_{ij} = -(E_{ij} - E_{ji})$$

所以 $E_{12} - E_{21}, E_{13} - E_{31}, \cdots, E_{n-1,n} - E_{n,n-1}$ 都是斜对称矩阵.

假设

$$k_{12}(E_{12} - E_{21}) + k_{13}(E_{13} - E_{31}) + \cdots$$
$$+ k_{n-1,n}(E_{n-1,n} - E_{n,n-1}) = 0 \quad (3)$$

由于 $\{E_{ij} | i = 1, \cdots, n; j = 1, \cdots, n\}$ 是 $M_n(K)$ 的一个基,所以 $E_{12}$, $E_{21}, E_{13}, E_{31}, \cdots, E_{n-1,n}, E_{n,n-1}$ 线性无关,从而由(3)式可推出 $k_{12} = k_{13} = \cdots = k_{n-1,n} = 0$. 这证明了 $(E_{12} - E_{21}), \cdots, (E_{n-1,n} - E_{n,n-1})$ 线性无关. 又由(2)式便可得出,$E_{12} - E_{21}, \cdots, E_{n-1,n} - E_{n,n-1}$ 是 $V$ 的一个基.

$$\dim V = (n-1) + (n-2) + \cdots + 1 = \frac{n(n-1)}{2}$$

**命题** 8.3.1 在 $n$ 维线性空间 $V$ 中,任意 $n$ 个线性无关的向量都是 $V$ 的一个基.

**证明** 设 $\alpha_1, \cdots, \alpha_n$ 是 $V$ 中 $n$ 个线性无关的向量. 任取 $\beta \in V$,只要证 $\beta$ 可由 $\alpha_1, \cdots, \alpha_n$ 线性表出,那么 $\alpha_1, \cdots, \alpha_n$ 便是 $V$ 的一个基. 因为 $\dim V = n$,所以可在 $V$ 中取一个基 $\delta_1, \delta_2, \cdots, \delta_n$. 于是向量组 $\alpha_1, \cdots, \alpha_n, \beta$ 可由 $\delta_1, \cdots, \delta_n$ 线性表出. 因为 $n + 1 > n$,据 §2 性质 8 得,$\alpha_1, \cdots, \alpha_n, \beta$ 线性相关. 由于 $\alpha_1, \cdots, \alpha_n$ 线性无关,据 §2 性质 6 得,$\beta$ 可由 $\alpha_1, \cdots, \alpha_n$ 线性表出. ▌

**定理** 8.3.2 在 $n$ 维线性空间 $V$ 中,任意一个线性无关的向量组 $\alpha_1, \cdots, \alpha_r$ 都可以扩充成 $V$ 的一个基(即存在 $V$ 的一个基包含 $\{\alpha_1, \cdots, \alpha_r\}$).

**证明** 若 $r = n$,则 $\alpha_1, \cdots, \alpha_n$ 是 $V$ 的一个基(据命题8.3.1). 下设 $r < n$. 此时 $V$ 中必有一个向量 $\beta_1$ 不能由 $\alpha_1, \cdots, \alpha_r$ 线性表出(否

则，$\alpha_1,\cdots,\alpha_r$ 将成为 $V$ 的一个基，但 $r\neq n$，矛盾），从而 $\alpha_1,\cdots,\alpha_r,\beta_1$ 线性无关（据 §2 性质 6 的逆否命题）. 如果 $r+1=n$，则 $\alpha_1,\cdots,\alpha_r,\beta_1$ 是 $V$ 的一个基. 如果 $r+1<n$，则 $V$ 中有一个向量 $\beta_2$ 不能由 $\alpha_1,\cdots,\alpha_r,\beta_1$ 线性表出，从而 $\alpha_1,\cdots,\alpha_r,\beta_1,\beta_2$ 线性无关. 依次下去，可得到线性无关的向量组

$$\alpha_1,\quad \cdots,\quad \alpha_r,\quad \beta_1,\quad \cdots,\quad \beta_l$$

其中 $r+l=n$，这就是 $V$ 的一个基. ▌

在有限维向量空间 $V$ 中，把一个基的元素排一个次序，写成 $\alpha_1,\cdots,\alpha_n$，这称为 $V$ 的一个有序基，但是我们仍称它为基，因为它的元素的次序从书写形式中已经表明.

**命题** 8.3.2　$n$ 维线性空间 $V$ 中取一个基 $\alpha_1,\cdots,\alpha_n$，则 $V$ 中每个向量 $\alpha$ 能唯一地表成 $\alpha_1,\cdots,\alpha_n$ 的线性组合.

**证明**　由 §2 性质 5 立即得出. ▌

**定义** 4　在 $n$ 维线性空间 $V$ 中取一个基 $\alpha_1,\cdots,\alpha_n$，$V$ 中每个向量 $\alpha$ 表成 $\alpha_1,\cdots,\alpha_n$ 的线性组合 $\alpha=\sum\limits_{i=1}^{n}a_i\alpha_i$ 的系数组成的有序数组 $(a_1,\cdots,a_n)$ 称为 $\alpha$ 在基 $\alpha_1,\cdots,\alpha_n$ 下的**坐标**，它是由 $\alpha$ 和基 $\alpha_1,\cdots,\alpha_n$ 唯一确定的.

因此，在有限维线性空间 $V$ 中，只要找到了一个基，则每一个向量的形式都清楚了，从而 $V$ 的结构也就清楚了.

**例** 4　把复数域 $C$ 看成实数域 $R$ 上的线性空间，求它的一个基和维数，以及每个复数 $z=a+bi$ 在这个基下的坐标.

**解**　由于每个复数 $z$ 可以写成 $z=a+bi$，其中 $a,b\in R$，所以 $z$ 可由 $1,\mathrm{i}$ 线性表出. 假设 $k_1 1+k_2\mathrm{i}=0$，由复数相等的定义得，$k_1=k_2=0$，这表明 $1,\mathrm{i}$ 是线性无关的. 因此，$1,\mathrm{i}$ 就是一个基. 从而 $\dim_R C=2$.

复数 $z=a+bi$ 在基 $1,\mathrm{i}$ 下的坐标是 $(a,b)$.

注意：若复数域 $C$ 看成自身上的线性空间，则它的一个基是：1，从而 $\dim_C C=1$. 这个例子说明，线性空间的维数是与所考虑的

域有关的.

# 阅读材料九

设 $W$ 是一个集合，设 $S$ 是由 $W$ 的一些子集组成的集合. 在 $S$ 中任取两个元素 $A,B$（它们是 $W$ 的子集），可能它们有包含关系，即 $A\subset B$ 或 $B\subset A$；也可能它们没有包含关系，即 $A\not\subset B$ 且 $B\not\subset A$. 包含关系 $\subset$ 称为 $S$ 的一个**偏序**，它具有以下性质：

1）反身性，即 $A\subset A$ 对所有 $A\in S$；

2）反对称性，即若 $A\subset B$ 且 $B\subset A$，则 $A=B$；

3）传递性，即若 $A\subset B$ 且 $B\subset C$，则 $A\subset C$.

$S$ 称为**偏序集**.

偏序集 $S$ 的一个元素 $A$ 称为 $S$ 的一个**极大元素**，如果不存在 $B\in S$，使得 $A\subset B$ 且 $A\neq B$.

偏序集 $S$ 的一个子集 $T$ 称为 $S$ 的一个**链**，如果 $T$ 中任意两个元素都有包含关系.

设 $U$ 是偏序集 $S$ 的一个子集，$S$ 中的一个元素 $B$ 称为 $U$ 的一个**上界**，如果对所有 $X\in U$ 都有 $X\subset B$.

**Zorn 引理** 若一个偏序集 $S$ 的每个链都有上界，则 $S$ 有一个极大元素.

**定理 1** 域 $F$ 上的线性空间 $V$ 一定有基.

**证明** 设 $S$ 是 $V$ 中所有线性无关的向量集组成的集合，对于包含关系"$\subset$"，$S$ 成为一个偏序集. 任取 $S$ 的一个链 $T=\{B_i\mid i\in I\}$，其中 $I$ 为指标集. 我们令

$$B=\bigcup_{i\in I}B_i$$

假如 $B$ 是线性相关的向量集，则 $B$ 有一个有限子集 $C=\{\alpha_1,\cdots,\alpha_r\}$ 是线性相关的. 设 $\alpha_1\in B_{i_1},\cdots,\alpha_r\in B_{i_r}$. 由于 $T$ 中任意两个元素都有包含关系，因此 $B_{i_1},B_{i_2},\cdots,B_{i_r}$ 中有一个 $B_{i_l}$ 包含了其余 $r-1$ 个. 从而 $C\subset B_{i_l}$. 由于 $B_{i_l}$ 是线性无关的，所以 $C$ 也应当是线性无关的，矛盾. 这证明了 $B$ 是线性无关的向量集，从而 $B\in S$. 显然每个 $B_i\subset B$. 因此 $B$ 是 $T$ 的一个上界. 这表明 $S$ 的每个链都有上界. 据 Zorn 引理，$S$ 有一个极大元素 $A$. 于是 $A$ 是线性无关的向量

集. 在 $V$ 中任取一个向量 $\alpha$，若 $\alpha \in A$，则 $\alpha = 1 \cdot \alpha$. 若 $\alpha \notin A$，则 $A \subset A \cup \{\alpha\}$ 且 $A \neq A \cup \{\alpha\}$. 由于 $A$ 是 $S$ 的一个极大元素，所以 $A \cup \{\alpha\} \notin S$. 从而 $A \cup \{\alpha\}$ 是线性相关的向量集. 于是 $A \cup \{\alpha\}$ 有一个有限子集 $A_1$ 是线性相关的. 必有 $\alpha \in A_1$（否则 $A_1$ 是 $A$ 的子集，矛盾）. 由于 $A_1 \backslash \{\alpha\}$ 是线性无关的（因为 $A_1 \backslash \{\alpha\}$ 是 $A$ 的子集），所以 $\alpha$ 可以由 $A_1 \backslash \{\alpha\}$ 中的向量线性表出. 这证明了 $V$ 中任一向量 $\alpha$ 可以由 $A$ 中有限多个向量线性表出. 因此 $A$ 是 $V$ 的一个基. ▌

## 习　题　8.3

1. 求下列线性空间的一个基和维数：

(1) 数域 $K$ 上所有 $n$ 级对称（上三角）矩阵组成的线性空间；

(2) 习题 8.1 第 2 题的 (1)；

(3) $K[x]_n$；

(4) $K[A]$，其中

$$A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \omega & 0 \\ 0 & 0 & \omega^2 \end{pmatrix}, \qquad \omega = \frac{-1+\sqrt{3}\,\mathrm{i}}{2}$$

(5) $K[x,y]$ 中 $m$ 次齐次多项式添上零多项式组成的线性空间.

2. 设 $V$ 是域 $F$ 上的 $n$ 维线性空间，$\alpha_1, \cdots, \alpha_n$ 是 $V$ 中 $n$ 个向量. 证明：如果 $V$ 中每个向量都可以由 $\alpha_1, \cdots, \alpha_n$ 线性表出，则 $\alpha_1, \cdots, \alpha_n$ 是 $V$ 的一个基.

3. 证明：在 $K[x]_n$ 中，下列多项式组都是一个基，并且分别求多项式 $f$ 在各个基下的坐标：

(1) $1, x-a, (x-a)^2, \cdots, (x-a)^{n-1}$；

(2) 设 $c_1, c_2, \cdots, c_n$ 是 $K$ 中 $n$ 个不同的元素，令

$$g_i(x) = \prod_{j \neq i} (x-c_j)(c_i-c_j)^{-1}, \qquad i = 1,2,\cdots,n$$

考虑多项式组 $g_1(x), g_2(x), \cdots, g_n(x)$.

4. 设 $X = \{a_1, a_2, \cdots, a_n\}$，求域 $F$ 上的线性空间 $F^X$ 的一个基和维数，并且求函数 $f$ 在这个基下的坐标.

5. 在 $K^4$ 中，求向量 $\alpha$ 在基 $\alpha_1, \alpha_2, \alpha_3, \alpha_4$ 下的坐标，设

$\alpha_1 = (1,1,1,1), \quad \alpha_2 = (1,1,1,0), \quad \alpha_3 = (1,1,0,0),$

$\alpha_4 = (1,0,0,0)$, $\quad \alpha = (2, -1,3,4)$.

* 6. 求由 $K[x_1,\cdots,x_n]$ 中 $m$ 次齐次多项式添上零多项式组成的 $K$ 上的线性空间 $W$ 的维数.

(提示：$\dim W$ 等于形如

$$x_1^{i_1}x_2^{i_2}\cdots x_n^{i_n}, \qquad i_1 + i_2 + \cdots + i_n = m$$

的单项式的数目. 把这种形式的一个单项式对应于由 $m$ 个球和 $n$ 根小棍排成的一行：

$$\underbrace{\bigcirc\cdots\bigcirc}_{i_1}\ \underset{x_1}{|}\ \underbrace{\bigcirc\cdots\bigcirc}_{i_2}\ \underset{x_2}{|}\ \cdots\ \underset{x_{n-1}}{|}\ \underbrace{\bigcirc\cdots\bigcirc}_{i_n}\ \underset{x_n}{|}$$

然后计算这些排法的总数. 注意最后一根小棍的位置是固定的.)

* 7. 求由 $K[x_1,\cdots,x_n]$ 中次数 $\leqslant m$ 的多项式组成的线性空间 $V$ 的维数.

(提示：$\dim V$ 等于形如

$$x_1^{i_1}x_2^{i_2}\cdots x_n^{i_n}, \qquad i_1 + i_2 + \cdots + i_n \leqslant m$$

的单项式的数目. 把这样的一个单项式对应于由 $m$ 个球和 $n+1$ 根小棍排成的一行：

$$\underbrace{\bigcirc\cdots\bigcirc}_{i_1}\ \underset{x_1}{|}\ \underbrace{\bigcirc\cdots\bigcirc}_{i_2}\ \underset{x_2}{|}\ \cdots\ |\ \underbrace{\bigcirc\cdots\bigcirc}_{i_n}\ \underset{x_n}{|}\ \underbrace{\bigcirc\cdots\cdots\bigcirc}_{m-(i_1+\cdots+i_n)}\ |$$

然后用第 6 题的结论.)

## §4 基变换与坐标变换

在域 $F$ 上的 $n$ 维线性空间 $V$ 中，取定一个基 $\alpha_1,\cdots,\alpha_n$，则 $V$ 中每个向量 $\alpha$ 在这个基下有坐标. 对于不同的基，同一个向量 $\alpha$ 的坐标一般是不同的. 本节讨论随着基的改变，同一个向量的坐标是如何变化的.

设 $\alpha_1,\cdots,\alpha_n$ 与 $\beta_1,\cdots,\beta_n$ 是 $V$ 的两个基. 由于 $\alpha_1,\cdots,\alpha_n$ 是基，因此 $\beta_1,\cdots,\beta_n$ 中每个向量可唯一地表示成 $\alpha_1,\cdots,\alpha_n$ 的线性组合：

$$\begin{cases} \beta_1 = a_{11}\alpha_1 + a_{21}\alpha_2 + \cdots + a_{n1}\alpha_n, \\ \beta_2 = a_{12}\alpha_1 + a_{22}\alpha_2 + \cdots + a_{n2}\alpha_n, \\ \cdots \quad \cdots \quad \cdots \quad \cdots \quad \cdots \quad \cdots \quad \cdots \\ \beta_n = a_{1n}\alpha_1 + a_{2n}\alpha_2 + \cdots + a_{nn}\alpha_n \end{cases} \tag{1}$$

设向量 $\alpha$ 在基 $\alpha_1,\cdots,\alpha_n$ 与基 $\beta_1,\cdots,\beta_n$ 下的坐标分别是 $(x_1,x_2,\cdots,x_n)$ 与 $(y_1,y_2,\cdots,y_n)$，即

$$\begin{aligned} \alpha &= x_1\alpha_1 + x_2\alpha_2 + \cdots + x_n\alpha_n \\ &= y_1\beta_1 + y_2\beta_2 + \cdots + y_n\beta_n \end{aligned} \tag{2}$$

我们要找出 $(x_1,x_2,\cdots,x_n)$ 与 $(y_1,y_2,\cdots,y_n)$ 之间的关系.

为了使推导过程简洁，我们引进一种形式的写法：把向量组 $\alpha_1,\alpha_2,\cdots,\alpha_n$ 的线性组合写成

$$x_1\alpha_1 + x_2\alpha_2 + \cdots + x_n\alpha_n =: (\alpha_1,\alpha_2,\cdots,\alpha_n)\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \tag{3}$$

进而把(1)式写成

$$(\beta_1,\beta_2,\cdots,\beta_n) = (\alpha_1,\alpha_2,\cdots,\alpha_n)\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix} \tag{4}$$

我们说这种写法是"形式的"，因为在这里是以一般的线性空间 $V$ 中的元素 $\alpha_1,\alpha_2,\cdots.\alpha_n$ 构成有序元素组 $(\alpha_1,\alpha_2,\cdots,\alpha_n)$，不是以域 $F$ 的元素构成有序元素组，但是我们却赋予它与域 $F$ 上的有序元素组一样的运算性质. 对于具体的 $n$ 元数组的向量空间 $K^n$，由于 $\alpha_1,\cdots,\alpha_n$ 都是 $n$ 元数组，因此当 $\alpha_1,\cdots,\alpha_n$ 为列向量时，有序元素组 $(\alpha_1,\alpha_2,\cdots,\alpha_n)$ 表示以 $\alpha_1,\alpha_2,\cdots,\alpha_n$ 为列的矩阵，这时(3)式和(4)式正好是第四章 §3 讲的矩阵乘法的第二种表述方式.

上述形式写法满足以下运算规则：设 $(\alpha_1,\alpha_2,\cdots,\alpha_n)$ 与 $(\beta_1,$

$\beta_2,\cdots,\beta_n$）是 $V$ 中两个向量组，$A=(a_{ij})$，$B=(b_{ij})$ 是域 $F$ 上两个 $n$ 级矩阵，则

$$((\alpha_1,\alpha_2,\cdots,\alpha_n)A)B$$
$$=(\alpha_1,\alpha_2,\cdots,\alpha_n)(AB)$$
$$(\alpha_1,\alpha_2,\cdots,\alpha_n)A+(\alpha_1,\alpha_2,\cdots,\alpha_n)B$$
$$=(\alpha_1,\alpha_2,\cdots,\alpha_n)(A+B)$$
$$(\alpha_1,\alpha_2,\cdots,\alpha_n)A+(\beta_1,\beta_2,\cdots,\beta_n)A$$
$$=(\alpha_1+\beta_1,\alpha_2+\beta_2,\cdots,\alpha_n+\beta_n)A$$

**证明** 由于上述形式写法的定义与矩阵乘法的定义在形式上一样，因此把矩阵乘法的有关运算法则的证明重复一遍，便得出上述三个规则. 譬如，对第一个规则的证明：

$$\text{等号左边的}(1,j)\text{ 元}=\sum_{k=1}^{n}((\alpha_1,\alpha_2,\cdots,\alpha_n)A)(1;k)\cdot b_{kj}$$
$$=\sum_{k=1}^{n}(\sum_{l=1}^{n}\alpha_l a_{lk})b_{kj}=\sum_{k=1}^{n}\sum_{l=1}^{n}a_{lk}b_{kj}\alpha_l;$$
$$\text{等号右边的}(1,j)\text{ 元}=\sum_{l=1}^{n}\alpha_l(AB)(l;j)=\sum_{l=1}^{n}\alpha_l(\sum_{k=1}^{n}a_{lk}b_{kj})$$
$$=\sum_{l=1}^{n}\sum_{k=1}^{n}a_{lk}b_{kj}\alpha_l=\sum_{k=1}^{n}\sum_{l=1}^{n}a_{lk}b_{kj}\alpha_l$$

所以第一个运算规则成立. 类似地可证第二、三个运算规则. ▌

我们把公式(4)中出现的矩阵

$$A=\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix} \tag{5}$$

称为由基 $\alpha_1,\alpha_2,\cdots,\alpha_n$ 到基 $\beta_1,\beta_2,\cdots,\beta_n$ 的**过渡矩阵**，即

$$(\beta_1,\beta_2,\cdots,\beta_n)=(\alpha_1,\alpha_2,\cdots,\alpha_n)A \tag{6}$$

中的矩阵 $A$.

**命题** 8.4.1 基 $\alpha_1,\alpha_2,\cdots,\alpha_n$ 到基 $\beta_1,\beta_2,\cdots,\beta_n$ 的过渡矩阵一定是可逆矩阵.

**证明** 设 $(\beta_1,\beta_2,\cdots,\beta_n)=(\alpha_1,\alpha_2,\cdots,\alpha_n)A$,要证 $A$ 可逆. 用反证法,假如 $A$ 不可逆,则 $|A|=0$. 于是齐次线性方程组 $AX=0$ 有非零解,取一个非零解 $\gamma=\begin{pmatrix}c_1\\ \vdots\\ c_n\end{pmatrix}$. 有

$$c_1\beta_1+c_2\beta_2+\cdots+c_n\beta_n=(\beta_1,\beta_2,\cdots,\beta_n)\begin{pmatrix}c_1\\ \vdots\\ c_n\end{pmatrix}$$

$$=((\alpha_1,\alpha_2,\cdots,\alpha_n)A)\gamma=(\alpha_1,\alpha_2,\cdots,\alpha_n)(A\gamma)=0$$

这表明 $\beta_1,\beta_2,\cdots,\beta_n$ 线性相关,矛盾. 因此 $A$ 可逆. ▌

**命题** 8.4.2 设 $\alpha_1,\alpha_2,\cdots,\alpha_n$ 是 $V$ 的一个基,$\beta_1,\cdots,\beta_n$ 是 $V$ 的一个向量组,设

$$(\beta_1,\beta_2,\cdots,\beta_n)=(\alpha_1,\alpha_2,\cdots,\alpha_n)A$$

如果 $A$ 可逆,则 $\beta_1,\beta_2,\cdots,\beta_n$ 是 $V$ 的一个基.

**证明** 只要证 $\beta_1,\beta_2,\cdots,\beta_n$ 线性无关即可. 假设

$$k_1\beta_1+k_2\beta_2+\cdots+k_n\beta_n=0$$

则

$$0=(\beta_1,\beta_2,\cdots,\beta_n)\begin{pmatrix}k_1\\ \vdots\\ k_n\end{pmatrix}=(\alpha_1,\alpha_2,\cdots,\alpha_n)\left(A\begin{pmatrix}k_1\\ \vdots\\ k_n\end{pmatrix}\right)$$

因为 $\alpha_1,\alpha_2,\cdots,\alpha_n$ 是 $V$ 的一个基,所以从上式得 $A\begin{pmatrix}k_1\\ \vdots\\ k_n\end{pmatrix}=0$. 因为 $A$ 可逆,所以 $\begin{pmatrix}k_1\\ \vdots\\ k_n\end{pmatrix}=0$. 这证明了 $\beta_1,\cdots,\beta_n$ 线性无关. ▌

现在我们来讨论向量 $\alpha$ 在不同基下的坐标之间的关系.

$$\alpha=(\beta_1,\beta_2,\cdots,\beta_n)\begin{pmatrix}y_1\\y_2\\\vdots\\y_n\end{pmatrix}=(\alpha_1,\alpha_2,\cdots,\alpha_n)\left(A\begin{pmatrix}y_1\\y_2\\\vdots\\y_n\end{pmatrix}\right)$$

又 $\alpha=(\alpha_1,\alpha_2,\cdots,\alpha_n)\begin{pmatrix}x_1\\x_2\\\vdots\\x_n\end{pmatrix}$. 由于 $\alpha$ 在基 $\alpha_1,\alpha_2,\cdots,\alpha_n$ 下的坐标是唯一的，因此得到

$$\begin{pmatrix}x_1\\x_2\\\vdots\\x_n\end{pmatrix}=A\begin{pmatrix}y_1\\y_2\\\vdots\\y_n\end{pmatrix}\tag{7}$$

或者

$$\begin{pmatrix}y_1\\y_2\\\vdots\\y_n\end{pmatrix}=A^{-1}\begin{pmatrix}x_1\\x_2\\\vdots\\x_n\end{pmatrix}\tag{8}$$

(7)式和(8)式就是在基变换(6)下，向量的坐标变换公式.

**例 1** 在 $K^3$ 中，设

$$\alpha_1=(1,0,-1),\quad \alpha_2=(2,1,1),\quad \alpha_3=(1,1,1)$$
$$\beta_1=(0,1,1),\quad \beta_2=(-1,1,0),\quad \beta_3=(1,2,1)$$
$$\alpha=(2,5,3)$$

求基 $\alpha_1,\alpha_2,\alpha_3$ 到基 $\beta_1,\beta_2,\beta_3$ 的过渡矩阵 $T$，并且求 $\alpha$ 分别在这两个基下的坐标.

**解** 因为 $\beta=k_1\alpha_1+k_2\alpha_2+k_3\alpha_3$

$$\Leftrightarrow\beta'=k_1\alpha_1'+k_2\alpha_2'+k_3\alpha_3'$$

所以 $$(\beta_1,\beta_2,\beta_3)=(\alpha_1,\alpha_2,\alpha_3)T$$
$$\Leftrightarrow(\beta_1',\beta_2',\beta_3')=(\alpha_1',\alpha_2',\alpha_3')T$$

设 $A=(\alpha_1{}',\alpha_2{}',\alpha_3{}')$，即 $A$ 是以 $\alpha_1{}',\alpha_2{}',\alpha_3{}'$ 为列的矩阵. 同理，设 $B=(\beta_1{}',\beta_2{}',\beta_3{}')$，则从 $(\beta_1{}',\beta_2{}',\beta_3{}')=(\alpha_1{}',\alpha_2{}',\alpha_3{}')T$ 得出，$B=AT$. 为了求 $T$，需要解这个矩阵方程，可按照第四章 §9 讲的方法求解：

$$\begin{pmatrix} 1 & 2 & 1 & 0 & -1 & 1 \\ 0 & 1 & 1 & 1 & 1 & 2 \\ -1 & 1 & 1 & 1 & 0 & 1 \end{pmatrix} \xrightarrow{\text{行变换}} \begin{pmatrix} 1 & 0 & 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & -1 & -3 & -2 \\ 0 & 0 & 1 & 2 & 4 & 4 \end{pmatrix}$$

所以

$$T=\begin{pmatrix} 0 & 1 & 1 \\ -1 & -3 & -2 \\ 2 & 4 & 4 \end{pmatrix}$$

设 $\alpha$ 在基 $\alpha_1,\alpha_2,\alpha_3$ 下的坐标为 $(x_1,x_2,x_3)$，则

$$\alpha=(\alpha_1,\alpha_2,\alpha_3)\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$$

此式成立当且仅当 $\quad \alpha'=(\alpha_1{}',\alpha_2{}',\alpha_3{}')\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$

即 $$A\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}=\begin{pmatrix} 2 \\ 5 \\ 3 \end{pmatrix}$$

解这个线性方程组得 $\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}=\begin{pmatrix} 2 \\ -5 \\ 10 \end{pmatrix}$. 因此 $\alpha$ 在基 $\alpha_1,\alpha_2,\alpha_3$ 下的坐标为 $(2,-5,10)$.

同理，解线性方程组 $B\begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix}=\begin{pmatrix} 2 \\ 5 \\ 3 \end{pmatrix}$ 得，$\begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix}=\begin{pmatrix} 1 \\ 0 \\ 2 \end{pmatrix}$，因此 $\alpha$ 在基 $\beta_1,\beta_2,\beta_3$ 下的坐标为 $(1,0,2)$.

可以用坐标变换公式(7)验证所得结果是正确的.

## 习 题 8.4

1. 在 $K^4$ 中,求由基 $\alpha_1,\alpha_2,\alpha_3,\alpha_4$ 到基 $\beta_1,\beta_2,\beta_3,\beta_4$ 的过渡矩阵,并求向量 $\alpha$ 在所指基下的坐标. 设

$$(1)\quad \begin{cases}\alpha_1=(1,0,0,0),\\ \alpha_2=(0,1,0,0),\\ \alpha_3=(0,0,1,0),\\ \alpha_4=(0,0,0,1),\end{cases}\qquad \begin{cases}\beta_1=(1,1,-1,1),\\ \beta_2=(2,3,1,1),\\ \beta_3=(3,1,-2,0),\\ \beta_4=(4,1,-1,2),\end{cases}$$

$\alpha=(x_1,x_2,x_3,x_4)$ 在基 $\beta_1,\beta_2,\beta_3,\beta_4$ 下的坐标;

$$(2)\quad \begin{cases}\alpha_1=(1,0,0,0),\\ \alpha_2=(4,1,0,0),\\ \alpha_3=(-3,2,1,0),\\ \alpha_4=(2,-3,2,1),\end{cases}\qquad \begin{cases}\beta_1=(1,1,8,3),\\ \beta_2=(0,3,7,2),\\ \beta_3=(1,1,6,2),\\ \beta_4=(-1,4,-1,-1),\end{cases}$$

$\alpha=(1,4,2,3)$ 在基 $\alpha_1,\alpha_2,\alpha_3,\alpha_4$ 下的坐标.

2. 在第 1 题(1)中,求一非零向量 $\alpha$,它在基 $\alpha_1,\alpha_2,\alpha_3,\alpha_4$ 与基 $\beta_1,\beta_2,\beta_3,\beta_4$ 下有相同的坐标.

3. (1) 证明:在 $C[x]_n$ 中,多项式组

$$f_i=(x-a_1)\cdots(x-a_{i-1})(x-a_{i+1})\cdots(x-a_n),\quad i=1,2\cdots,n$$

是一个基,其中 $a_1,a_2,\cdots,a_n$ 是互不相同的数;

(2) 在(1)中,取 $a_1,a_2,\cdots,a_n$ 为全体 $n$ 次单位根,求由基 $1,x,\cdots,x^{n-1}$ 到基 $f_1,f_2,\cdots,f_n$ 的过渡矩阵.

(提示: $x^n-1=\prod_{i=1}^{n}(x-a_i)$).

## §5 线性子空间

在第三章我们看到,数域 $K$ 上 $n$ 元齐次线性方程组 $AX=0$ 的解空间是向量空间 $K^n$ 的一个线性子空间. 这个例子说明,我们

不仅要研究整个向量空间的结构，而且应当研究它的线性子空间. 一方面这些子空间本身有它的应用，另一方面通过研究线性空间的子空间可以更深刻地揭示整个线性空间的结构.

**定义** 1 域 $F$ 上线性空间 $V$ 的一个非空子集 $W$ 称为 $V$ 的一个**线性子空间**(简称为**子空间**)，如果 $W$ 对于 $V$ 的两种运算也构成域 $F$ 上的线性空间.

**定理** 8.5.1 域 $F$ 上线性空间 $V$ 的非空子集 $W$ 是 $V$ 的一个线性子空间当且仅当 $W$ 对于 $V$ 的两种运算封闭，即

1) $\alpha,\beta \in W \Rightarrow \alpha+\beta \in W$;

2) $k \in F, \alpha \in W \Rightarrow k\alpha \in W$.

**证明** 必要性由定义 1 直接得出. 下面看充分性. 由已知条件，$V$ 的两种运算都是 $W$ 的运算. 作为 $W$ 的运算，同样满足加法交换律，加法结合律，以及公理 ⑤ 至 ⑧. 因为 $0 \in F, \alpha \in W$，所以 $0\alpha \in W$. 由于 $V$ 是线性空间，因此 $0\alpha = 0$，从而得 $0 \in W$. 于是 $V$ 的零元素是 $W$ 的零元素. 因为 $-1 \in F, \alpha \in W$，所以 $(-1)\alpha \in W$. 由于 $V$ 是线性空间，因此 $(-1)\alpha = -\alpha$. 从而得 $-\alpha \in W$，由于 $\alpha + (-\alpha) = 0$，于是 $W$ 中每个元素 $\alpha$ 在 $V$ 中的负元素 $-\alpha$ 也是 $\alpha$ 在 $W$ 中的负元素. 这样我们证明了 $W$ 对于 $V$ 的两种运算也成为域 $F$ 上一个线性空间. 因此 $W$ 是 $V$ 的一个线性子空间. ∎

**例** 1 在线性空间 $V$ 中，$\{0\}$ 是 $V$ 的一个线性子空间，称它为零子空间，也记作 0. 显然，$V$ 本身是 $V$ 的一个线性子空间. 0 和 $V$ 称为 $V$ 的平凡子空间，其余的线性子空间称为非平凡子空间.

**例** 2 $C[a,b]$ 是 $R^{[a,b]}$ 的线性子空间. $C^{(n)}[a,b]$ 是 $C[a,b]$ 的线性子空间.

**例** 3 $K[x]_n$ 是 $K[x]$ 的线性子空间.

维数是线性空间最重要的数字特征，既然 $V$ 的子空间也是线性空间，因此它也有维数. 自然要问：$V$ 的子空间的维数与 $V$ 的维数有什么关系？下面的定理回答了这个问题：

**定理** 8.5.2 设 $W$ 是线性空间 $V$ 的一个线性子空间，则

$$\dim W \leqslant \dim V$$

**证明** 情形 1，$\dim V = n$. 此时据 §2 性质 8，$V$ 中任意 $n+1$ 个向量都线性相关. 在 $W$ 中取一个基 $S$，由于 $S$ 是线性无关的，因此 $S$ 包含的向量数目不超过 $n$. 于是 $W$ 是有限维的，且 $\dim W \leqslant n$.

情形 2，$\dim V = \infty$. 若 $W$ 是无限维的，则 $\dim W \leqslant \dim V$. 若 $W$ 是有限维的，则 $\dim W < \dim V$. ▌

**命题** 8.5.1 设 $V$ 是有限维线性空间，$W$ 是 $V$ 的子空间，如果 $\dim W = \dim V$，则 $W = V$.

**证明** 在 $W$ 中取一个基 $\alpha_1,\cdots,\alpha_n$，由于 $\dim W = \dim V$，据命题 8.3.1 得，$\alpha_1,\cdots,\alpha_n$ 就是 $V$ 的一个基. $V$ 中任一向量 $\alpha$，由于它可表成 $\alpha_1,\cdots,\alpha_n$ 的线性组合，因此 $\alpha \in W$. 从而 $V \subset W$. 又 $W \subset V$，所以 $W = V$. ▌

**命题** 8.5.2 设 $V$ 是有限维线性空间，$W$ 是 $V$ 的一个子空间，则 $W$ 的一个基可以扩充成 $V$ 的一个基.

**证明** 从定理 8.5.2 和定理 8.3.2 立即得到. ▌

如何构造域 $F$ 上线性空间 $V$ 的子空间？给出 $V$ 的一个向量组 $\alpha_1,\alpha_2,\cdots,\alpha_s$，我们想构造一个包含 $\alpha_1,\alpha_2,\cdots,\alpha_s$ 的最小的子空间. 由于子空间对加法和纯量乘法封闭，因此包含 $\alpha_1,\alpha_2,\cdots,\alpha_s$ 的子空间一定包含它们的任一线性组合. 为此考虑这个向量组的所有线性组合组成的集合 $W$，即

$$W = \{k_1\alpha_1 + k_2\alpha_2 + \cdots + k_s\alpha_s \mid k_i \in F,\quad i = 1,2,\cdots,s\}$$

容易验证 $W$ 是 $V$ 的一个线性子空间(因为 $W$ 非空，并且对 $V$ 的两种运算封闭). 我们称 $W$ 是由 $\alpha_1,\alpha_2,\cdots,\alpha_s$ **生成(或张成)的线性子空间**，记作 $\langle\alpha_1,\alpha_2,\cdots,\alpha_s\rangle$ 或者 $L(\alpha_1,\alpha_2,\cdots,\alpha_s)$.

从线性子空间的定义知道，如果 $V$ 的一个子空间 $U$ 包含 $\alpha_1,\alpha_2,\cdots,\alpha_s$，那么 $U$ 一定包含 $\alpha_1,\alpha_2,\cdots,\alpha_s$ 的所有的线性组合，也就是说，$U$ 一定包含 $\langle\alpha_1,\alpha_2,\cdots,\alpha_s\rangle$. 因此，由 $\alpha_1,\alpha_2,\cdots,\alpha_s$ 生成的子空间 $\langle\alpha_1,\alpha_2,\cdots,\alpha_s\rangle$ 是 $V$ 中包含向量组 $\alpha_1,\alpha_2,\cdots,\alpha_s$ 的最小的子空间.

在有限维线性空间 $V$ 中，任何一个线性子空间 $W$ 都可以用上

述方法得到，这是因为 $W$ 也是有限维的(据定理 8.5.2)，在 $W$ 中取一个基 $\alpha_1,\alpha_2,\cdots,\alpha_r$，则 $W=\langle\alpha_1,\alpha_2,\cdots,\alpha_r\rangle$.

**定理** 8.5.3 $V$ 中向量组 $\alpha_1,\alpha_2,\cdots,\alpha_s$ 生成的线性子空间 $\langle\alpha_1,\alpha_2,\cdots,\alpha_s\rangle$ 的维数等于向量组 $\alpha_1,\alpha_2,\cdots,\alpha_s$ 的秩；向量组 $\alpha_1,\alpha_2,\cdots,\alpha_s$ 的一个极大线性无关组就是 $\langle\alpha_1,\alpha_2,\cdots,\alpha_s\rangle$ 的一个基.

**证明** 设向量组 $\alpha_1,\alpha_2,\cdots,\alpha_s$ 的一个极大线性无关组是 $\alpha_{j_1},\alpha_{j_2},\cdots,\alpha_{j_r}$. 据线性表出的传递性，$\langle\alpha_1,\alpha_2,\cdots,\alpha_s\rangle$ 中每一个向量可以由 $\alpha_{j_1},\alpha_{j_2},\cdots,\alpha_{j_r}$ 线性表出，从而 $\alpha_{j_1},\alpha_{j_2},\cdots,\alpha_{j_r}$ 是 $\langle\alpha_1,\alpha_2,\cdots,\alpha_s\rangle$ 的一个基. 由此得出，$\langle\alpha_1,\alpha_2,\cdots,\alpha_s\rangle$ 的维数等于向量组 $\alpha_1,\alpha_2,\cdots,\alpha_s$ 的秩. ▌

**命题** 8.5.3 $V$ 中两个向量组 $\alpha_1,\alpha_2,\cdots,\alpha_s$ 与 $\beta_1,\beta_2,\cdots,\beta_m$ 生成的子空间相同的充分必要条件是这两个向量组等价.

**证明** 必要性. $\langle\alpha_1,\alpha_2,\cdots,\alpha_s\rangle=\langle\beta_1,\beta_2,\cdots,\beta_m\rangle$，则 $\alpha_i$ 可以由 $\beta_1,\beta_2,\cdots,\beta_m$ 线性表出，$i=1,\cdots,s$；$\beta_j$ 可以由 $\alpha_1,\alpha_2,\cdots,\alpha_s$ 线性表出，$j=1,\cdots,m$. 因此这两个向量组等价.

充分性. 如果 $\alpha_1,\alpha_2,\cdots,\alpha_s$ 与 $\beta_1,\beta_2,\cdots,\beta_m$ 等价，则由线性表出的传递性，$\langle\alpha_1,\alpha_2,\cdots,\alpha_s\rangle$ 中任一向量可由 $\beta_1,\beta_2,\cdots,\beta_m$ 线性表出，因此 $\langle\alpha_1,\alpha_2,\cdots,\alpha_s\rangle\subset\langle\beta_1,\beta_2,\cdots,\beta_m\rangle$. 同理，有 $\langle\beta_1,\beta_2,\cdots,\beta_m\rangle\subset\langle\alpha_1,\alpha_2,\cdots,\alpha_s\rangle$. 所以 $\langle\alpha_1,\cdots,\alpha_s\rangle=\langle\beta_1,\cdots,\beta_m\rangle$. ▌

## 阅读材料十

在第三章§7 中我们证明了：数域 $K$ 上 $n$ 元齐次线性方程组的解集 $W$ 是 $K^n$ 的一个线性子空间，称它为这个方程组的解空间. 完全一样地可以证明：任一域 $F$ 上 $n$ 元齐次线性方程组的解集是 $F^n$ 的一个线性子空间.(注：$F^n$ 的定义及其运算与 $K^n$ 类似.)

例如，域 $F$ 上的一个 $n$ 元齐次线性方程

$$a_1x_1+a_2x_2+\cdots+a_nx_n=0 \tag{1}$$

的解集 $W$ 是 $F^n$ 的一个线性子空间.

一般地，域 $F$ 上的一个 $n$ 元方程

$$f(x_1,x_2,\cdots,x_n)=0 \tag{2}$$

的解集如果是 $F^n$ 的一个线性子空间，那么称方程(2)是 $F$ 上的一个($n$ 元) **线性条件**.

从上述知，域 $F$ 上任意一个 $n$ 元齐次线性方程(1) 是 $F$ 上的一个线性条件.

请读者思考，域 $F$ 上一个 $n$ 元非齐次线性方程是不是 $F$ 上的一个线性条件?

**例 1** $Z_2$ 上的下述 $n$ 元方程

$$\sum_{i=1}^{n} x_i^2 = 0 \tag{3}$$

是不是 $Z_2$ 上的一个线性条件?

**解** 方程(3)的解集 $W$ 包含 $Z_2^n$ 的零向量，因此 $W$ 是 $Z_2^n$ 的非空子集. 在 $W$ 中任取两个向量

$$\alpha=(a_1,a_2,\cdots,a_n), \qquad \beta=(b_1,b_2,\cdots,b_n)$$

我们来看 $\alpha+\beta=(a_1+b_1,a_2+b_2,\cdots,a_n+b_n)$ 是否属于 $W$?由于

$$\sum_{i=1}^{n}(a_i+b_i)^2=\sum_{i=1}^{n}(a_i^2+2a_ib_i+b_i^2)=\sum_{i=1}^{n}(a_i^2+b_i^2)$$

$$=\sum_{i=1}^{n}a_i^2+\sum_{i=1}^{n}b_i^2=0$$

所以 $\alpha+\beta\in W$. 由于

$$\sum_{i=1}^{n}(ka_i)^2=k^2\sum_{i=1}^{n}a_i^2=0, \quad k\in Z_2$$

所以 $k\alpha\in W$. 于是 $W$ 是 $Z_2^n$ 的一个线性子空间. 因此方程(3) 是 $Z_2$ 上的一个线性条件.

请读者思考，方程(3) 是不是实数域 $R$ 上的一个线性条件?是不是复数域 $C$ 上的一个线性条件?

从线性条件的定义知道，给了域 $F$ 上一个($n$ 元) 线性条件，它就决定了 $F^n$ 的一个线性子空间. 这是构造 $F^n$ 的线性子空间的一种方法.

例如，由 $Z_2$ 上的 3 元线性条件

$$\sum_{i=1}^{3} x_i^2 = 0 \tag{4}$$

确定 $Z_2^3$ 的一个线性子空间是

$$\{(0,0,0),(1,1,0),(1,0,1),(0,1,1)\}$$

# 习　题　8.5

1. 设 $V_1$ 和 $V_2$ 都是域 $F$ 上线性空间 $V$ 的子空间，且 $V_1 \subset V_2$. 证明：

(1)　$\dim V_1 \leqslant \dim V_2$；

(2)　如果 $V_2$ 是有限维的，并且 $\dim V_1 = \dim V_2$，则 $V_1 = V_2$.

2. 判断数域 $K$ 上下列 $n$ 元方程的解集是否为 $K^n$ 的子空间：

(1)　$\sum\limits_{i=1}^{n} a_i x_i = 1$；

(2)　$x_3 = 2x_4$，这里 $n \geqslant 4$.

3. 设 $A \in M_n(F)$，其中 $\operatorname{char} F = 0$.

(1)　证明：域 $F$ 上全体与 $A$ 可交换的矩阵组成 $M_n(F)$ 的一个线性子空间. 把它记作 $C(A)$；

(2)　求 $C(aI)$，其中 $a$ 是 $F$ 中给定的非零元素；

(3)　设

$$A = \begin{pmatrix} a_1 & 0 & 0 & \cdots & 0 \\ 0 & a_2 & 0 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \cdots & a_n \end{pmatrix}$$

其中 $a_1, a_2, \cdots, a_n$ 两两不同，求 $C(A)$ 的维数和一个基.

4. 记号同第 3 题，其中 $n = 3$. 设

$$A = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 4 & -2 & 1 \end{pmatrix}$$

求 $C(A)$ 的维数和一个基.

5. 在域 $F$ 上的线性空间 $V$ 中，如果 $k_1\alpha + k_2\beta + k_3\gamma = 0$，并且 $k_1k_2 \neq 0$，证明：$\langle \alpha, \gamma \rangle = \langle \beta, \gamma \rangle$.

6. 在 $K^4$ 中($K$ 是数域)，求向量组 $\alpha_1, \alpha_2, \alpha_3, \alpha_4$ 生成的子空间的维数和一

个基.设

$$\alpha_1=(1,-3,2,-1),\qquad \alpha_2=(-2,1,5,3)$$
$$\alpha_3=(4,-3,7,1),\qquad \alpha_4=(-1,-11,8,-3)$$

7. 设 $V$ 是域 $F$ 上一个 $n$ 维线性空间，$\alpha_1,\alpha_2,\cdots,\alpha_n$ 是 $V$ 的一个基，$\beta_1,\beta_2,\cdots,\beta_S$ 是 $V$ 的一个向量组，并且

$$(\beta_1,\beta_2,\cdots,\beta_S)=(\alpha_1,\alpha_2,\cdots,\alpha_n)A$$

证明：$\langle\beta_1,\beta_2,\cdots,\beta_S\rangle$ 的维数等于 $n\times s$ 矩阵 $A$ 的秩.

（提示：设 $A$ 的列向量组 $A_1,A_2,\cdots,A_S$ 的一个极大线性无关组是 $A_{j_1},\cdots,A_{j_r}$，去证 $\beta_{j_1},\cdots,\beta_{j_r}$ 是 $\beta_1,\beta_2,\cdots,\beta_S$ 的一个极大线性无关组）.

* 8. $\sum\limits_{i=1}^{n}x_i^p=0$ 是否为 $Z_p$ 上的线性条件，如果是，写出由 $Z_3$ 上的线性条件 $\sum\limits_{i=1}^{2}x_i^3=0$ 确定的 $Z_3^2$ 的线性子空间.

* 9. 分别就 $F=R$ 与 $F=C$ 两种情形，讨论 $\sum\limits_{i=1}^{n}x_i^2=0$ 是否为 $F$ 上的线性条件？如果是，写出由此条件确定的 $F^n$ 的线性子空间；如果不是，请说明理由.

* 10. 设实二次型

$$f(x_1,x_2,\cdots,x_n)=x_1^2+\cdots+x_p^2-x_{p+1}^2-\cdots-x_n^2$$

的符号差 $s>0$，证明：方程 $f(x_1,\cdots,x_n)=0$ 的解集 $W$ 中有一个子集 $W_1$ 是 $R^n$ 的一个线性子空间，且 $\dim W_1=n-p$.

11. 设 $V_1,V_2$ 是域 $F$ 上线性空间 $V$ 的两个真子空间（即 $V_i\neq V$），证明：

$$V_1\cup V_2\neq V$$

* 12. 设 $V_1,V_2,\cdots,V_S$ 是域 $F$ 上线性空间 $V$ 的 $s$ 个真子空间，证明：如果 $\mathrm{char}F=0$，则 $V$ 中至少有一个向量不属于 $V_1,V_2,\cdots,V_S$ 中任何一个. 如果 $\mathrm{char}F=p$，结论是否一定成立？

（提示：$\mathrm{char}F=0$ 时，对 $s$ 用归纳法. 当 $s=2$ 时，第 11 题已证. 在 $s$ 的情形，对 $V_1,\cdots,V_{S-1}$ 用归纳假设，$V$ 中存在向量 $\alpha$ 不属于 $V_1,\cdots,V_{S-1}$ 中任何一个. 假如 $\alpha\notin V_S$，则 $\alpha$ 即为所求. 下设 $\alpha\in V_S$. 设 $\beta\notin V_S$. 考虑集合 $W=\{k\alpha+\beta\mid k\in F\}$，先证 $W$ 中任一向量不属于 $V_S$. 再证 $W$ 中至少有一个向量不属于 $V_1,\cdots,V_{S-1}$ 中任何一个. $\mathrm{char}F=p$ 时，结论不一定成立. 例如，设 $V=Z_2^3$，可举反例.）

## §6 子空间的交与和·子空间的直和

这一节讨论如何利用线性空间 $V$ 的已知的子空间构造出新的子空间. 我们知道,集合有交、并等运算. 线性空间 $V$ 的子空间作为集合,当然可以求交集,并集,问题是:两个子空间的交集、并集是否仍是子空间?

**定理** 8.6.1 设 $V_1,V_2$ 都是域 $F$ 上线性空间 $V$ 的子空间,则 $V_1\cap V_2$ 也是 $V$ 的子空间.

**证明** 因为 $0\in V_1\cap V_2$,所以 $V_1\cap V_2$ 非空集. 设 $\alpha,\beta\in V_1\cap V_2$,则 $\alpha,\beta\in V_i, i=1,2$. 因为 $V_i$ 是子空间,所以 $\alpha+\beta\in V_i$; $k\alpha\in V_i, \forall\, k\in F; i=1,2$. 于是 $\alpha+\beta\in V_1\cap V_2, k\alpha\in V_1\cap V_2$, $\forall\, k\in F$. 因此,$V_1\cap V_2$ 是 $V$ 的子空间. ▌

由集合的交的定义可得出,子空间的交适合下列运算规则:

(i) 交换律:$V_1\cap V_2=V_2\cap V_1$;

(ii) 结合律:$(V_1\cap V_2)\cap V_3=V_1\cap(V_2\cap V_3)$.

由结合律,我们可定义多个子空间的交:

$$V_1\cap V_2\cap\cdots\cap V_s=\bigcap_{i=1}^{s}V_i$$

用归纳法易证,$\bigcap\limits_{i=1}^{s}V_i$ 也是 $V$ 的子空间.

* 类似于定理 8.6.1 的证法可得,设 $I$ 是任一指标集,若对于每个 $i\in I$,$V_i$ 是 $V$ 的子空间;则 $\bigcap\limits_{i\in I}V_i$ 也是 $V$ 的子空间. (注:$\bigcap\limits_{i\in I}V_i:=\{\alpha\mid\alpha\in V_i,\forall\, i\in I\}$).

线性空间 $V$ 的两个子空间 $V_1$ 与 $V_2$ 的并集一般来说不是 $V$ 的子空间. 例如,设 $V$ 是几何空间(即以原点 $O$ 为起点的所有向量组成的 3 维实线性空间),$V_1,V_2$ 是通过原点的两个不同的平面,它们都是 $V$ 的子空间. 由于 $V_1\cup V_2$ 对加法不封闭,因此 $V_1\cup V_2$ 不是

$V$ 的子空间.如果我们想构造一个包含 $V_1 \cup V_2$ 的子空间,那么这个子空间应当包含 $V_1$ 中的任一向量 $\alpha_1$ 与 $V_2$ 中的任一向量 $\alpha_2$ 的和 $\alpha_1 + \alpha_2$. 由此受到启发,我们应当考虑集合

$$\{\alpha_1 + \alpha_2 | \alpha_1 \in V_1, \alpha_2 \in V_2\}$$

**定理** 8.6.2 设 $V_1, V_2$ 是域 $F$ 上线性空间 $V$ 的两个子空间,则集合

$$\{\alpha_1 + \alpha_2 | \alpha_1 \in V_1, \alpha_2 \in V_2\} \tag{1}$$

是 $V$ 的一个子空间,称它是 $V_1$ 与 $V_2$ 的**和**,记作 $V_1 + V_2$.

**证明** 把集合(1)记作 $W$. 显然 $0 \in W$(因为 $0 = 0 + 0$). 在 $W$ 中任取两个向量 $\alpha, \beta$,则

$$\alpha = \alpha_1 + \alpha_2, \qquad \beta = \beta_1 + \beta_2$$

其中 $\alpha_1, \beta_1 \in V_1, \alpha_2, \beta_2 \in V_2$,我们有

$$\alpha + \beta = (\alpha_1 + \beta_1) + (\alpha_2 + \beta_2) \tag{2}$$

由于 $V_1, V_2$ 是 $V$ 的子空间,所以 $\alpha_1 + \beta_1 \in V_1, \alpha_2 + \beta_2 \in V_2$,从而 $\alpha + \beta \in W$.

任取 $k \in F$,我们有

$$k\alpha = k\alpha_1 + k\alpha_2, \quad k\alpha_1 \in V_1, k\alpha_2 \in V_2$$

因此 $k\alpha \in W$. 从而 $W$ 是 $V$ 的一个子空间. ▌

对于 $V_1$ 中任一向量 $\alpha_1$,有 $\alpha_1 = \alpha_1 + 0$. 因此 $V_1 \subset V_1 + V_2$. 同理,$V_2 \subset V_1 + V_2$. 从而 $V_1 \cup V_2 \subset V_1 + V_2$. 所以 $V_1 + V_2$ 是包含 $V_1 \cup V_2$ 的子空间.

设 $U$ 是 $V$ 的子空间且 $U \supset V_1 \cup V_2$,则对于任意 $\alpha_i \in V_i, i = 1,2$,有 $\alpha_i \in U$. 从而 $\alpha_1 + \alpha_2 \in U$. 由此看出 $V_1 + V_2 \subset U$. 这表明 $V_1 + V_2$ 是 $V$ 中包含 $V_1 \cup V_2$ 的最小的子空间.

从定理 8.6.2 知道

$$V_1 + V_2 = \{\alpha_1 + \alpha_2 | \alpha_1 \in V_1, \alpha_2 \in V_2\} \tag{3}$$

从(3)式容易看出,子空间的和适合下列运算规则:

(i) 交换律:$V_1 + V_2 = V_2 + V_1$;

(ii) 结合律:$(V_1 + V_2) + V_3 = V_1 + (V_2 + V_3)$.

由结合律，我们可以定义多个子空间的和：

$$V_1+V_2+\cdots+V_s=\sum_{i=1}^{s}V_i$$

用归纳法易证，$\sum\limits_{i=1}^{s}V_i$ 仍是 $V$ 的子空间，并且

$$\begin{aligned}&V_1+V_2+\cdots+V_s\\&=\{\alpha_1+\alpha_2+\cdots+\alpha_s\mid\alpha_i\in V_i,i=1,\cdots,s\}\end{aligned}\tag{4}$$

**命题** 8.6.1 设 $\alpha_1,\cdots,\alpha_r$ 与 $\beta_1,\cdots,\beta_s$ 是域 $F$ 上线性空间 $V$ 的两个向量组，则

$$\langle\alpha_1,\cdots,\alpha_r\rangle+\langle\beta_1,\cdots,\beta_s\rangle=\langle\alpha_1,\cdots,\alpha_r,\beta_1,\cdots,\beta_s\rangle\tag{5}$$

**证明** 从(3)式得出

$$\begin{aligned}&\langle\alpha_1,\cdots,\alpha_r\rangle+\langle\beta_1,\cdots,\beta_s\rangle\\&=\{(k_1\alpha_1+\cdots+k_r\alpha_r)+(l_1\beta_1+\cdots+l_s\beta_s)\mid k_i,l_j\in F\}\\&=\langle\alpha_1,\cdots,\alpha_r,\beta_1,\cdots,\beta_s\rangle\end{aligned}$$ ▌

设 $V$ 是几何空间，$V_1$ 是过点 $O$ 的一个平面，$V_2$ 是过点 $O$ 的另一个平面，它们相交于一条直线 $L$. 则 $V_1,V_2,L$ 都是 $V$ 的子空间，并且 $V_1\cap V_2=L$. 由于 $V$ 中每个向量 $\alpha$ 可以表示成 $V_1$ 中一个向量与 $V_2$ 中一个向量的和(注意表法不唯一)，所以 $V_1+V_2=V$. 由于 $\dim V_1=\dim V_2=2,\dim L=1,\dim V=3$，因此在本例中，有

$$\dim V_1+\dim V_2=\dim(V_1+V_2)+\dim(V_1\cap V_2)\tag{6}$$

公式(6)对于任一线性空间的任意两个有限维子空间都成立，即，我们有

**定理** 8.6.3(**维数公式**) 如果 $V_1,V_2$ 是域 $F$ 上线性空间 $V$ 的两个有限维子空间，那么 $V_1\cap V_2$ 与 $V_1+V_2$ 也都是有限维的，并且

$$\dim V_1+\dim V_2=\dim(V_1+V_2)+\dim(V_1\cap V_2)\tag{7}$$

**证明** 因为 $V_1$ 是有限维的，而 $V_1\cap V_2$ 是 $V_1$ 的子空间，所以 $V_1\cap V_2$ 也是有限维的. 设 $V_1,V_2$ 的维数分别是 $n_1,n_2,V_1\cap V_2$ 的维数是 $m$. 取 $V_1\cap V_2$ 的一个基 $\alpha_1,\cdots,\alpha_m$. 因为 $V_1\cap V_2\subset V_1$，所以

它可以扩充成 $V_1$ 的一个基 $\alpha_1,\cdots,\alpha_m,\beta_1,\cdots,\beta_{n_1-m}$. 同理，它也可以扩充成 $V_2$ 的一个基 $\alpha_1,\cdots,\alpha_m,\gamma_1,\cdots,\gamma_{n_2-m}$. 据(5)式，我们有

$$
\begin{aligned}
& V_1+V_2 \\
&=\langle\alpha_1,\cdots,\alpha_m,\beta_1,\cdots,\beta_{n_1-m}\rangle+\langle\alpha_1,\cdots,\alpha_m,\gamma_1,\cdots,\gamma_{n_2-m}\rangle \\
&=\langle\alpha_1,\cdots,\alpha_m,\beta_1,\cdots,\beta_{n_1-m},\gamma_1,\cdots,\gamma_{n_2-m}\rangle
\end{aligned}
\tag{8}
$$

于是 $V_1+V_2$ 是有限维的. 如果能证

$$\alpha_1,\cdots,\alpha_m,\beta_1,\cdots,\beta_{n_1-m},\gamma_1,\cdots,\gamma_{n_2-m}$$

线性无关，则它就是 $V_1+V_2$ 的一个基，从而有

$$
\begin{aligned}
\dim(V_1+V_2) &= m+(n_1-m)+(n_2-m) \\
&= n_1+n_2-m \\
&= \dim V_1+\dim V_2-\dim(V_1\cap V_2)
\end{aligned}
$$

假设有等式

$$
\begin{aligned}
& k_1\alpha_1+\cdots+k_m\alpha_m+p_1\beta_1+\cdots+p_{n_1-m}\beta_{n_1-m} \\
& \quad +q_1\gamma_1+\cdots+q_{n_2-m}\gamma_{n_2-m}=0
\end{aligned}
$$

令
$$
\begin{aligned}
\delta &= k_1\alpha_1+\cdots+k_m\alpha_m+p_1\beta_1+\cdots+p_{n_1-m}\beta_{n_1-m} \\
&= -q_1\gamma_1-\cdots-q_{n_2-m}\gamma_{n_2-m}
\end{aligned}
\tag{9}
$$

由(9)的第一个等式知 $\delta\in V_1$，由第二个等式知 $\delta\in V_2$. 于是 $\delta\in V_1\cap V_2$. 因此 $\delta$ 可由 $\alpha_1,\cdots,\alpha_m$ 线性表出，令

$$\delta=l_1\alpha_1+\cdots+l_m\alpha_m \tag{10}$$

由(9)的第二式以及(10)式得

$$l_1\alpha_1+\cdots+l_m\alpha_m+q_1\gamma_1+\cdots+q_{n_2-m}\gamma_{n_2-m}=0 \tag{11}$$

由于 $\alpha_1,\cdots,\alpha_m,\gamma_1,\cdots,\gamma_{n_2-m}$ 线性无关，所以

$$l_1=\cdots=l_m=q_1=\cdots=q_{n_2-m}=0$$

从而 $\delta=0$. 再由(9)的第一式便得到

$$k_1\alpha_1+\cdots+k_m\alpha_m+p_1\beta_1+\cdots+p_{n_1-m}\beta_{n_1-m}=0 \tag{12}$$

由于 $\alpha_1,\cdots,\alpha_m,\beta_1,\cdots,\beta_{n_1-m}$ 线性无关，所以

$$k_1=\cdots=k_m=p_1=\cdots=p_{n_1-m}=0$$

这证明了 $\alpha_1,\cdots,\alpha_m,\beta_1,\cdots,\beta_{n_1-m},\gamma_1,\cdots,\gamma_{n_2-m}$ 线性无关. ▌

从维数公式看出，$V$ 的两个有限维子空间的和的维数与交的维数加起来才等于 $V_1$ 的维数与 $V_2$ 的维数之和. 两个子空间的和的维数等于维数的和的情形是特别重要的，由维数公式立即得到

**推论** 8.6.1 设 $V_1,V_2$ 是域 $F$ 上线性空间 $V$ 的两个有限维子空间，则

$$\dim(V_1+V_2)=\dim V_1+\dim V_2 \Leftrightarrow V_1\cap V_2=0$$

这里 0 表示 $V$ 的零子空间.

现在我们来仔细讨论这种特殊情形. 先看一个例子. 设 $V$ 是几何空间，$V_1$ 是过原点 $O$ 的一个平面，$V_2$ 是过原点 $O$ 的一条直线且 $V_2$ 不在 $V_1$ 上. 于是 $V_1\cap V_2=0$. 容易看出 $V_1+V_2=V$，并且 $V$ 中每个向量 $\alpha$ 能被唯一地表示成

$$\alpha=\alpha_1+\alpha_2,\quad \alpha_1\in V_1,\alpha_2\in V_2$$

由此受到启发，引出下述定义：

**定义** 2 设 $V_1,V_2$ 是域 $F$ 上线性空间 $V$ 的子空间，如果和 $V_1+V_2$ 中每个向量 $\alpha$ 能被唯一地表示成

$$\alpha=\alpha_1+\alpha_2,\quad \alpha_1\in V_1,\alpha_2\in V_2 \tag{13}$$

那么和 $V_1+V_2$ 就称为**直和**，记作 $V_1\oplus V_2$.

**定理** 8.6.4 设 $V_1,V_2$ 是域 $F$ 上线性空间 $V$ 的子空间，则下列命题互相等价：

(i) 和 $V_1+V_2$ 是直和；

(ii) 和 $V_1+V_2$ 中零向量的表法唯一，即如果

$$\alpha_1+\alpha_2=0,\quad \alpha_1\in V_1,\alpha_2\in V_2$$

则 $\alpha_1=0$ 且 $\alpha_2=0$.

(iii) $V_1\cap V_2=0$.

**证明** (i)⇒(ii)是显然的.

(ii)⇒(iii)：任取 $\alpha\in V_1\cap V_2$，于是零向量可表成

$$0=\alpha+(-\alpha),\quad \alpha\in V_1,\ -\alpha\in V_2$$

据(ii)得，$\alpha=0$. 因此 $V_1\cap V_2=0$.

(iii)⇒(i)：任取 $\alpha \in V_1 + V_2$，假设 $\alpha$ 有两种表法：

$$\alpha = \alpha_1 + \alpha_2, \quad \alpha_1 \in V_1, \alpha_2 \in V_2$$
$$\alpha = \beta_1 + \beta_2, \quad \beta_1 \in V_1, \beta_2 \in V_2$$

则 $\alpha_1 - \beta_1 = \beta_2 - \alpha_2 \in V_1 \cap V_2$. 因为 $V_1 \cap V_2 = 0$，所以 $\alpha_1 = \beta_1$，$\alpha_2 = \beta_2$. 因此，和 $V_1 + V_2$ 是直和. ▌

**定理** 8.6.5　设 $V_1, V_2$ 是域 $F$ 上线性空间 $V$ 的两个有限维子空间，则下列命题互相等价：

(i) 和 $V_1 + V_2$ 是直和；

(ii) $\dim(V_1 + V_2) = \dim V_1 + \dim V_2$；

(iii) $V_1$ 的一个基与 $V_2$ 的一个基合起来是 $V_1 + V_2$ 的一个基.

**证明**　据定理 8.6.4 和推论 8.6.1 立即得到(i)⇔(ii).

(iii)⇒(ii)是显然的. 现在证(ii)⇒(iii)：设 $\alpha_1, \cdots, \alpha_s$ 是 $V_1$ 的一个基，$\beta_1, \cdots, \beta_r$ 是 $V_2$ 的一个基，则

$$\begin{aligned} V_1 + V_2 &= \langle \alpha_1, \cdots, \alpha_s \rangle + \langle \beta_1, \cdots, \beta_r \rangle \\ &= \langle \alpha_1, \cdots, \alpha_s, \beta_1, \cdots, \beta_r \rangle \end{aligned}$$

因为　$\dim(V_1 + V_2) = \dim V_1 + \dim V_2 = s + r$

所以向量组 $\alpha_1, \cdots, \alpha_s, \beta_1, \cdots, \beta_r$ 的秩等于 $s + r$，从而它线性无关，因此它是 $V_1 + V_2$ 的一个基. ▌

**定理** 8.6.6　设 $V$ 是域 $F$ 上有限维线性空间，$U$ 是 $V$ 的一个子空间，则存在 $V$ 的一个子空间 $W$，使得

$$V = U \oplus W \tag{14}$$

**证明**　因为 $V$ 是有限维的，所以子空间 $U$ 是有限维的. 取 $U$ 的一个基 $\alpha_1, \cdots, \alpha_s$，把它扩充成 $V$ 的一个基

$$\alpha_1, \cdots, \alpha_s, \alpha_{s+1}, \cdots, \alpha_n$$

令 $W = \langle \alpha_{s+1}, \cdots, \alpha_n \rangle$，则

$$\begin{aligned} U + W &= \langle \alpha_1, \cdots, \alpha_s \rangle + \langle \alpha_{s+1}, \cdots, \alpha_n \rangle \\ &= \langle \alpha_1, \cdots, \alpha_s, \alpha_{s+1}, \cdots, \alpha_n \rangle = V \end{aligned}$$

由于 $U$ 的一个基与 $W$ 的一个基合起来是 $U + W$ 的一个基，因此和 $U + W$ 是直和. 从而 $V = U \oplus W$. ▌

**定义** 2　设 $V$ 是域 $F$ 上的线性空间，$U$ 是 $V$ 的一个子空间，如果存在 $V$ 的一个子空间 $W$，使得 $V = U \oplus W$，则称 $W$ 是 $U$ 在 $V$ 里的补空间. 这时 $U$ 称为 $W$ 在 $V$ 里的补空间.

从定理 8.6.6 知道，如果 $V$ 是有限维的，那么它的每一个子空间都有补空间. 注意一个子空间的补空间不是唯一的. 例如，在几何空间中，设 $\pi$ 是过原点 $O$ 的一个平面，则任意一条经过 $O$ 点但不在 $\pi$ 上的直线都是 $\pi$ 的补空间.

注意：子空间 $U$ 在 $V$ 里的补空间的概念与子集 $U$ 在 $V$ 里的补集的概念是不同的概念，不要混淆.

**例** 1　设 $V = M_n(K)$，其中 $K$ 是数域. 用 $V_1$ 表示 $K$ 上所有 $n$ 级对称矩阵组成的子空间，用 $V_2$ 表示所有 $n$ 级斜对称矩阵组成的子空间，证明：$V = V_1 \oplus V_2$.

**证明**　第一步，证明 $V = V_1 + V_2$. $V \supset V_2 + V_1$ 是显然的. 关键是要证 $V \subset V_1 + V_2$. 任取 $A \in V = M_n(K)$，有

$$A = \frac{A + A'}{2} + \frac{A - A'}{2} \tag{15}$$

易验证 $\frac{A + A'}{2} \in V_1$，$\frac{A - A'}{2} \in V_2$. 因此 $A \in V_1 + V_2$. 这证明了 $V \subset V_1 + V_2$. 因此 $V = V_1 + V_2$.

第二步，证明和 $V_1 + V_2$ 是直和. 为此只要证 $V_1 \cap V_2 = 0$. 任取 $B \in V_1 \cap V_2$，则 $B' = B$，并且 $B' = -B$. 于是 $B = -B$，从而 $2B = 0$. 因为 $\operatorname{char} K = 0$，所以 $B = 0$. 于是 $V_1 \cap V_2 = 0$.

综上述得，$V = V_1 \oplus V_2$.　▌

子空间的直和的概念可以推广到多个子空间的情形.

**定义** 3　设 $V_1, V_2, \cdots, V_s$ 都是域 $F$ 上线性空间 $V$ 的子空间. 如果和 $V_1 + V_2 + \cdots + V_s$ 中每个向量 $\alpha$ 可唯一地表示成

$$\alpha = \alpha_1 + \alpha_2 + \cdots + \alpha_s, \quad \alpha_i \in V_i (i = 1, 2, \cdots, s)$$

则这个和就称为**直和**，记作 $V_1 \oplus V_2 \oplus \cdots \oplus V_s$ 或 $\bigoplus_{i=1}^{s} V_i$.

**定理** 8.6.7　设 $V_1, V_2, \cdots, V_s$ 是域 $F$ 上线性空间 $V$ 的子空间，

则下列命题互相等价：

(i) 和 $V_1+V_2+\cdots+V_s$ 是直和；

(ii) 和 $\sum_{i=1}^{s}V_i$ 中零向量的表法唯一；

(iii) $V_i\cap\sum_{j\neq i}V_j=0$， $i=1,2,\cdots,s$.

**证明** (i)⇒(ii)是显然的.

(ii)⇒(iii)：任取 $\alpha\in V_i\cap\sum_{j\neq i}V_j$，则 $-\alpha\in V_i$ 且 $\alpha\in\sum_{j\neq i}V_j$.
于是 $\alpha=\sum_{j\neq i}\alpha_j$，其中 $\alpha_j\in V_j$. 因此零向量可以表成

$$0=(-\alpha)+\alpha=(-\alpha)+\sum_{j\neq i}\alpha_j$$

据(ii)得，$-\alpha=0$，所以 $\alpha=0$. 于是 $V_i\cap\sum_{j\neq i}V_j=0$.

(iii)⇒(i)：任取 $\alpha\in\sum_{i=1}^{s}V_i$，假设 $\alpha$ 有两种表法：

$$\alpha=\alpha_1+\alpha_2+\cdots+\alpha_s,\quad \alpha_i\in V_i(i=1,2,\cdots,s)$$
$$\alpha=\beta_1+\beta_2+\cdots+\beta_s,\quad \beta_i\in V_i(i=1,2,\cdots,s)$$

任取 $i\in\{1,2,\cdots,s\}$. 从上两式可得出

$$\beta_i-\alpha_i=\sum_{j\neq i}(\alpha_j-\beta_j)\in V_i\cap\sum_{j\neq i}V_j$$

因为 $V_i\cap\sum_{j\neq i}V_j=0$，所以 $\beta_i-\alpha_i=0$，即 $\beta_i=\alpha_i$. ▌

**定理** 8.6.8 设 $V_1,V_2,\cdots,V_s$ 是域 $F$ 上线性空间 $V$ 的有限维子空间，则下列命题互相等价：

(i) 和 $\sum_{i=1}^{s}V_i$ 是直和；

(ii)
$$\begin{aligned}&\dim(V_1+V_2+\cdots+V_s)\\=&\dim V_1+\dim V_2+\cdots+\dim V_s\end{aligned}$$

(iii) $V_i$ 的一个基，当 $i=1,2,\cdots,s$ 时，合起来是 $\sum_{i=1}^{s}V_i$ 的一个基.

**证明** (i)⇒(ii)：因为和 $\sum_{i=1}^{s} V_i$ 是直和，据定理 8.6.7 得，$V_i \cap \sum_{j\neq i} V_j = 0, i = 1,2,\cdots,s$. 于是

$$\begin{aligned}\dim\left(\sum_{i=1}^{s} V_i\right) &= \dim\left(V_1 + \sum_{j\neq 1} V_j\right)\\ &= \dim V_1 + \dim\left(\sum_{j\neq 1} V_j\right) - \dim\left(V_1 \cap \sum_{j\neq 1} V_j\right)\\ &= \dim V_1 + \dim\left(\sum_{j\neq 1} V_j\right)\end{aligned}$$

注意

$$V_2 \cap (V_3 + \cdots + V_s) \subset V_2 \cap (V_1 + V_3 + \cdots + V_s) = 0$$

因此对 $s$ 用归纳法，据归纳假设可得

$$\dim\left(\sum_{j\neq 1} V_j\right) = \sum_{j\neq 1} \dim V_j$$

从而得到

$$\dim\left(\sum_{i=1}^{s} V_i\right) = \sum_{i=1}^{s} \dim V_i$$

(ii)⇒(iii)：类似于定理 8.6.5 证明中的(ii)⇒(iii).

(iii)⇒(i)：易证和 $\sum_{i=1}^{s} V_i$ 中零向量的表法唯一，从而 $\sum_{i=1}^{s} V_i$ 是直和. ▎

下面举一个例子说明在 $K^n$ 中如何具体求两个子空间的和与交的基和维数.

**例 2** 设 $V = K^4, V_1 = \langle \alpha_1, \alpha_2, \alpha_3 \rangle, V_2 = \langle \beta_1, \beta_2 \rangle$，其中

$$\alpha_1 = (1,2,1,0), \quad \alpha_2 = (-1,1,1,1), \quad \alpha_3 = (0,3,2,1),$$
$$\beta_1 = (2,-1,0,1), \quad \beta_2 = (1,-1,3,7)$$

分别求 $V_1$ 与 $V_2$ 的和与交的基和维数.

**解** 因为

$$V_1 + V_2 = \langle \alpha_1, \alpha_2, \alpha_3 \rangle + \langle \beta_1, \beta_2 \rangle = \langle \alpha_1, \alpha_2, \alpha_3, \beta_1, \beta_2 \rangle$$

所以向量组 $\alpha_1, \alpha_2, \alpha_3, \beta_1, \beta_2$ 的一个极大线性无关组就是 $V_1 + V_2$

的一个基，这个向量组的秩就是 $V_1+V_2$ 的维数. 按照第三章 §5 讲的方法，把 $\alpha_1,\alpha_2,\alpha_3,\beta_1,\beta_2$ 写成列向量，组成矩阵 $A$，对 $A$ 作一系列初等行变换，化成简化行阶梯形矩阵：

$$A=\begin{pmatrix}1&-1&0&2&1\\2&1&3&-1&-1\\1&1&2&0&3\\0&1&1&1&7\end{pmatrix}\to\begin{pmatrix}1&0&1&0&-1\\0&1&1&0&4\\0&0&0&1&3\\0&0&0&0&0\end{pmatrix}\qquad(16)$$

由此得出，$\alpha_1,\alpha_2,\beta_1$ 是 $V_1+V_2$ 的一个基，$\dim(V_1+V_2)=3$. 同时也知道，$\beta_2$ 可由 $\alpha_1,\alpha_2,\beta_1$ 线性表出，其系数应当是线性方程组

$$x_1\alpha_1+x_2\alpha_2+x_3\beta_1=\beta_2$$

的解. 而从上述 $A$ 及其简化行阶梯形矩阵的第 1,2,4,5 列可以看出，此方程组的解是 $(-1,4,3)$. 因此

$$\beta_2=-\alpha_1+4\alpha_2+3\beta_1$$

从而 $$\alpha_1-4\alpha_2=3\beta_1-\beta_2\in V_1\cap V_2$$

因为 $$\dim(V_1\cap V_2)=\dim V_1+\dim V_2-\dim(V_1+V_2)$$

而从(16)可以看出，$\dim V_1=2,\dim V_2=2$，所以

$$\dim(V_1\cap V_2)=2+2-3=1$$

于是 $\alpha_1-4\alpha_2=(5,-2,-2,-4)$ 是 $V_1\cap V_2$ 的一个基.

例 2 的求解过程表明，只要把矩阵 $A$ 经过一系列初等行变换化成简化行阶梯形矩阵，那么我们所需要的一切信息都可由此得到.

## 习 题 8.6

1. 设 $V=K^4,V_1=\langle\alpha_1,\alpha_2\rangle$（或 $V_1=\langle\alpha_1,\alpha_2,\alpha_3\rangle$），$V_2=\langle\beta_1,\beta_2\rangle$，分别求 $V_1+V_2$ 与 $V_1\cap V_2$ 的基和维数.

(1) $\alpha_1=(1,-1,0,1)$， $\alpha_2=(-2,3,1,-3)$，

$\beta_1=(1,2,0,-2)$， $\beta_2=(1,3,1,-3)$；

(2) $\alpha_1=(1,1,-1,2)$, $\alpha_2=(2,-1,3,0)$,

$\alpha_3=(0,-3,5,-4)$,

$\beta_1=(1,2,2,1),\beta_2=(4,-3,3,1)$.

2. 设 $V=K^n$,设 $V_1$ 与 $V_2$ 分别是齐次线性方程组

$$x_1+x_2+\cdots+x_n=0$$

与

$$x_1=x_2=\cdots=x_n$$

的解空间,证明 $V=V_1\oplus V_2$.

3. 证明:如果 $V=V_1\oplus V_2,V_1=V_{11}\oplus V_{12}$,那么

$$V=V_{11}\oplus V_{12}\oplus V_2$$

4. 证明:域 $F$ 上每一个 $n$ 维线性空间都可以表示成 $n$ 个一维子空间的直和.

5. 设 $V$ 是域 $F$ 上的线性空间,$V_1,V_2,\cdots,V_s$ 是 $V$ 的子空间,证明:和 $\sum\limits_{i=1}^{s}V_i$ 是直和的充分必要条件是

$$V_i\cap\sum_{j=1}^{i-1}V_j=0,\quad i=2,\cdots,s$$

6. 设 $V_1,V_2,W$ 都是域 $F$ 上线性空间 $V$ 的子空间,并且 $W\subset V_1+V_2$. 问:$W=(W\cap V_1)+(W\cap V_2)$ 是否总是成立?如果 $V_1\subset W$,那么上式是否一定成立?

7. 设 $V_1,W$ 是域 $F$ 上线性空间 $V$ 的子空间,且 $V_1\subset W$. 设 $V_1$ 在 $V$ 中的一个补空间是 $V_2$. 证明:$W=V_1\oplus(V_2\cap W)$.

8. 用 $M_n^0(F)$ 表示 $M_n(F)$ 中迹为零的矩阵组成的集合.

(1) 证明:$M_n^0(F)$ 是线性空间 $M_n(F)$ 的一个子空间;

(2) 证明:如果 $\mathrm{char}F=0$,那么

$$M_n(F)=\langle I\rangle\oplus M_n^0(F)$$

9. 设 $V=K^3$,设 $V_1=\langle\alpha_1,\alpha_2\rangle,\alpha_1=(1,2,3),\alpha_2=(3,2,1)$. 求 $V_1$ 在 $V$ 里的一个补空间.

## §7 线性空间的同构·有限域的元素数目

在前面几节我们列举了许多具体的线性空间,并且指出,给

定域上一个线性空间 $V$，它有许多线性子空间，这些子空间也都是域 $F$ 上的线性空间. 自然会提出一个问题：域 $F$ 上的众多的线性空间中，哪些在本质上是一样的?所谓本质上一样，粗略地说就是：尽管这些线性空间的元素不同，加法与纯量乘法的定义也可能不同，但是它们的元素之间存在一一对应，使得对应的元素关于这两种运算的性质完全一样. 也就是从代数运算的观点来看它们的结构完全相同. 我们用“同构”这一术语表达这些线性空间之间的关系. 确切地说，

**定义 1** 设 $V$ 和 $V'$ 都是域 $F$ 上的线性空间，如果存在 $V$ 到 $V'$ 的一个双射 $\sigma$，并且 $\sigma$ 保持加法与纯量乘法两种运算，即使得

$$\sigma(\alpha+\beta)=\sigma(\alpha)+\sigma(\beta), \quad \forall\ \alpha,\beta\in V \tag{1}$$

$$\sigma(k\alpha)=k\sigma(\alpha), \quad \forall\ k\in F,\alpha\in V \tag{2}$$

则称 $V$ 与 $V'$ 是同构的，记作 $V\cong V'$. 这样的映射 $\sigma$ 称为 $V$ 到 $V'$ 的同构映射.

从定义看出，如果域 $F$ 上的两个线性空间 $V$ 和 $V'$ 是同构的，那么 $V$ 与 $V'$ 的元素之间存在一一对应：$\alpha\mapsto\sigma(\alpha)$，并且这个映射 $\sigma$ 具有下列性质：

1. $\sigma(0)$ 是 $V'$ 的零元素 $0'$.

**证明** 任取 $\alpha'\in V'$，因为 $\sigma$ 是满射，所以存在 $\alpha\in V$，使得 $\alpha'=\sigma(\alpha)$. 于是

$$\alpha'+\sigma(0)=\sigma(\alpha)+\sigma(0)=\sigma(\alpha+0)=\sigma(\alpha)=\alpha'$$

这表明 $\sigma(0)$ 是 $V'$ 的零元素. ▌

2. $\sigma(-\alpha)=-\sigma(\alpha),\forall\ \alpha\in V$.

**证明** 因为

$$\sigma(\alpha)+\sigma(-\alpha)=\sigma(\alpha+(-\alpha))=\sigma(0)=0'$$

所以

$$\sigma(-\alpha)=-\sigma(\alpha) \qquad ▌$$

3. 设 $\alpha_1,\alpha_2,\cdots,\alpha_s$ 是 $V$ 的一个向量组，$k_1,k_2,\cdots,,k_s$ 是 $F$ 中任意 $s$ 个元素，则

$$\begin{aligned}&\sigma(k_1\alpha_1+k_2\alpha_2+\cdots+k_s\sigma_s)\\=&k_1\sigma(\alpha_1)+k_2\sigma(\alpha_2)+\cdots+k_s\sigma(\alpha_s)\end{aligned}\tag{3}$$

**证明** 由(1)和(2)式立即得出. ▌

4. $V$ 中向量组 $\alpha_1,\alpha_2\cdots,\alpha_s$ 线性相关当且仅当它们的象 $\sigma(\alpha_1)$, $\sigma(\alpha_2),\cdots,\sigma(\alpha_s)$ 是 $V'$ 的线性相关的向量组.

**证明** 因为 $\sigma$ 是单射,并且据性质 1,性质 3 得

$$\begin{aligned}&k_1\alpha_1+\cdots+k_s\alpha_s=0\\\Leftrightarrow&\sigma(k_1\alpha_1+\cdots+k_s\alpha_s)=\sigma(0)\\\Leftrightarrow&k_1\sigma(\alpha_1)+\cdots+k_s\sigma(\alpha_s)=0'\end{aligned}$$

所以 $\alpha_1,\cdots,\alpha_s$ 线性相关当且仅当 $\sigma(\alpha_1),\cdots,\sigma(\alpha_s)$ 线性相关. ▌

5. 如果 $\alpha_1,\alpha_2,\cdots,\alpha_n$ 是 $V$ 的一个基,则 $\sigma(\alpha_1),\sigma(\alpha_2),\cdots,\sigma(\alpha_n)$ 是 $V'$ 的一个基.

**证明** 据性质 4, $\sigma(\alpha_1),\cdots,\sigma(\alpha_n)$ 是 $V'$ 的一个线性无关的向量组. 任取 $\beta\in V'$,因为 $\sigma$ 是满射,所以存在 $\alpha\in V$,使得 $\beta=\sigma(\alpha)$. 设 $\alpha=a_1\alpha_1+\cdots+a_n\alpha_n$,则

$$\beta=\sigma(a_1\alpha_1+\cdots+a_n\alpha_n)=a_1\sigma(\alpha_1)+\cdots+a_n\sigma(\alpha_n)$$

因此 $\sigma(\alpha_1),\cdots,\sigma(\alpha_n)$ 是 $V'$ 的一个基. ▌

从以上五条性质我们可以看到,同构的线性空间其对应元素关于线性运算的性质是完全一样的.

从性质 5 我们还可得到

6. 如果 $V$ 是有限维的,那么与 $V$ 同构的线性空间 $V'$ 也是有限维的,并且它们的维数相同.

下面的性质 7 表明同构的线性空间 $V$ 与 $V'$,它们的子空间存在对应关系:

7. 如果 $V_1$ 是 $V$ 的一个线性子空间,则 $V_1$ 在同构映射 $\sigma$ 下的象集

$$\sigma(V_1)=\{\sigma(\alpha)\mid\alpha\in V_1\}$$

是 $V'$ 的子空间;并且如果 $V_1$ 是有限维的,则 $\sigma(V_1)$ 也是有限维的,而且 $V_1$ 与 $\sigma(V_1)$ 的维数相同.

**证明** 易验证 $\sigma(V_1)$ 非空集并且对于 $V'$ 的加法与纯量乘法封闭，因此 $\sigma(V_1)$ 是 $V'$ 的子空间. 容易看出，$\sigma$ 限制到 $V_1$ 上，它是 $V_1$ 到 $\sigma(V_1)$ 的一个双射，并且保持加法与纯量乘法两种运算，从而 $V_1$ 与 $\sigma(V_1)$ 是同构的线性空间. 据性质 6 立即得到性质 7 的后半部分. ▌

性质 6 和性质 7 进一步表明，同构的线性空间的确是具有相同的结构. 那么如何判别两个线性空间是否同构？据定义 1，应当在这两个线性空间之间找一个双射 $\sigma$，并且看 $\sigma$ 是否保持两种运算. 对于两个有限维线性空间，则有更简单的办法判别它们是否同构：

**定理** 8.7.1 域 $F$ 上两个有限维线性空间同构的充分必要条件是它们的维数相同.

**证明** 必要性从性质 6 即得. 下面证充分性. 设

$$\dim V = \dim V' = n$$

取 $V$ 的一个基 $\alpha_1,\alpha_2,\cdots,\alpha_n$；取 $V'$ 的一个基 $\gamma_1,\gamma_2,\cdots,\gamma_n$. 令

$$\sigma: V \longrightarrow V'$$

$$\alpha = a_1\alpha_1 + \cdots + a_n\alpha_n \longmapsto a_1\gamma_1 + \cdots + a_n\gamma_n \tag{4}$$

由于 $V$ 中元素 $\alpha$ 可用基 $\alpha_1,\cdots,\alpha_n$ 唯一地线性表出，因此 $\sigma$ 是 $V$ 到 $V'$ 的一个映射. 由于 $V'$ 中元素用基 $\gamma_1,\cdots,\gamma_n$ 线性表出的表法唯一，所以 $\sigma$ 是单射. 由于 $V'$ 中任一元素可由 $\gamma_1,\cdots,\gamma_n$ 线性表出，因此 $\sigma$ 是满射. 从而 $\sigma$ 是双射.

设 $\beta = b_1\alpha_1 + \cdots + b_n\alpha_n$. 直接计算可以得出，

$$\sigma(\alpha+\beta) = \sigma(\alpha) + \sigma(\beta), \quad \sigma(k\alpha) = k\sigma(\alpha), \quad \forall\, k \in F$$

因此 $\sigma$ 是 $V$ 到 $V'$ 的一个同构映射，从而 $V \cong V'$. ▌

定理 8.7.1 说明，维数相同的两个有限维线性空间一定同构. 因此域 $F$ 上任一 $n$ 维线性空间 $V$ 都与 $F$ 上有序 $n$ 元组的线性空间 $F^n$ 同构. 在 $V$ 中取一个基 $\alpha_1,\alpha_2,\cdots,\alpha_n$；在 $F^n$ 中取标准基

$$\varepsilon_1 = (1,0,\cdots,0), \varepsilon_2 = (0,1,0,\cdots,0), \cdots, \varepsilon_n = (0,\cdots,0,1)$$

则从(4)式得

$$\sigma: V \longrightarrow F^n$$
$$a_1\alpha_1 + \cdots + a_n\alpha_n \longmapsto (a_1, \cdots, a_n) \tag{5}$$
是 $V$ 到 $F^n$ 的一个同构映射.(5) 式表明,$V$ 中取定一个基后,让向量 $\alpha$ 对应到它在这个基下的坐标$(a_1,a_2,\cdots,a_n)$,这是 $V$ 到 $F^n$ 的一个同构映射. 这个同构映射很有用,它使得我们可以通过$V$中向量的坐标来研究 $V$ 的有关线性运算的性质. 由于 $V$ 与 $F^n$ 同构,因此 $F^n$ 中有关线性运算的性质在 $V$ 中仍然成立. 我们看一个例子:

**例 1** 设 $\alpha_1,\cdots,\alpha_n$ 是域 $F$ 上线性空间 $V$ 的一个基,设 $\beta_1,\cdots,\beta_s$ 是 $V$ 的一个向量组,并且
$$(\beta_1,\cdots,\beta_s) = (\alpha_1,\cdots,\alpha_n)A \tag{6}$$
其中 $A$ 是一个 $n \times s$ 矩阵. 证明:$\langle\beta_1,\cdots,\beta_s\rangle$ 的维数等于 $A$ 的秩.

**证明** 用 $\sigma$ 表示 $V$ 到 $F^n$ 的同构映射,它把 $\alpha \in V$ 映到 $\alpha$ 在基 $\alpha_1,\cdots,\alpha_n$ 下的坐标. 从(6) 式知,矩阵 $A$ 的第 $j$ 个列向量 $A_j$ 是 $\beta_j$ 在基 $\alpha_1,\cdots,\alpha_n$ 下的坐标,因此 $\sigma(\beta_j) = A_j$. 据性质 7 得
$$\begin{aligned}&\dim\langle\beta_1,\cdots,\beta_s\rangle = \dim(\sigma\langle\beta_1,\cdots,\beta_s\rangle)\\ =&\dim\langle\sigma(\beta_1),\cdots,\sigma(\beta_s)\rangle = \dim\langle A_1,\cdots,A_s\rangle = \operatorname{rank}(A) \quad \blacksquare\end{aligned}$$

同构是域 $F$ 上线性空间之间的一个关系,显然同构这个关系具有反身性(因为 $V$ 的恒等映射 $1_V$ 是 $V$ 到 $V$ 的一个同构映射),下面的命题 8.7.1 表明 $\cong$ 具有对称性和传递性. 从而 $\cong$ 是一个等价关系,它给出了域 $F$ 上所有线性空间组成的集合 $S$ 的一个划分. 一个等价类就称为一个同构类. 据定理 8.7.1,对于每一个非负整数 $n$,域 $F$ 上所有 $n$ 维线性空间恰好组成一个同构类. 因此,对于域 $F$ 上所有有限维线性空间组成的集合 $S_1$ 来说,维数是在同构关系下的完全不变量(完全不变量的定义见第五章 §3 的定义 2). 也就是说,维数足以区分 $S_1$ 里的不同的同构类.

**命题** 8.7.1 同构映射的逆映射以及两个同构映射的乘积还是同构映射.

**证明** 设 $\sigma$ 是域 $F$ 上线性空间 $V$ 到 $V'$ 的一个同构映射. 显然 $\sigma^{-1}$ 是 $V'$ 到 $V$ 的双射. 我们来证 $\sigma^{-1}$ 保持两种运算. 设 $\alpha',\beta'$ 是 $V'$

中任意两个向量，并且设

$$\sigma^{-1}(\alpha') = \alpha, \quad \sigma^{-1}(\beta') = \beta$$

则

$$\alpha' = \sigma(\alpha), \quad \beta' = \sigma(\beta)$$

于是

$$\begin{aligned}
\sigma^{-1}(\alpha' + \beta') &= \sigma^{-1}(\sigma(\alpha) + \sigma(\beta)) = \sigma^{-1}(\sigma(\alpha + \beta)) \\
&= (\sigma^{-1}\sigma)(\alpha + \beta) = 1_V(\alpha + \beta) = \alpha + \beta \\
&= \sigma^{-1}(\alpha') + \sigma^{-1}(\beta') \\
\sigma^{-1}(k\alpha') &= \sigma^{-1}(k\sigma(\alpha)) = \sigma^{-1}(\sigma(k\alpha)) = (\sigma^{-1}\sigma)(k\alpha) \\
&= k\alpha = k\sigma^{-1}(\alpha'), \forall\, k \in F
\end{aligned}$$

所以 $\sigma^{-1}$ 是 $V'$ 到 $V$ 的同构映射.

设 $\sigma$ 和 $\tau$ 分别是域 $F$ 上线性空间 $V$ 到 $V'$ 与 $V'$ 到 $V''$ 的同构映射. 我们来证 $\tau\sigma$ 是 $V$ 到 $V''$ 的一个同构映射. 显然 $\tau\sigma$ 是 $V$ 到 $V''$ 的双射. 任取 $\alpha,\beta \in V$，任取 $k \in F$，有

$$\begin{aligned}
(\tau\sigma)(\alpha + \beta) &= \tau(\sigma(\alpha + \beta)) = \tau(\sigma(\alpha) + \sigma(\beta)) \\
&= \tau\sigma(\alpha) + \tau\sigma(\beta) \\
(\tau\sigma)(k\alpha) &= \tau(\sigma(k\alpha)) = \tau(k\sigma(\alpha)) \\
&= k\tau(\sigma(\alpha)) = k\tau\sigma(\alpha)
\end{aligned}$$

因此 $\tau\sigma$ 是 $V$ 到 $V''$ 的同构映射. ▌

作为线性空间的同构的一个应用，我们来证明一个非常重要的结论：

**定理** 8.7.2 有限域的元素数目一定是一个素数的方幂.

为了证明定理 8.7.2，先介绍环的一些简单性质以及有限域中的一个结论.

设 $R$ 是任意一个环，则

1. $0a = a0 = 0, \forall\, a \in R$.

2. $(-a)b = a(-b) = -ab, (-a)(-b) = ab, \forall\, a,b \in R$.

这两条性质都可从环的定义出发去证明，参看阅读材料一.

设 $n$ 是正整数，对于 $a \in R$，我们定义

$$na := \underbrace{a + a + \cdots + a}_{(n\text{个})}$$

$$0a := 0 \quad (\text{注意左边是整数 } 0)$$

$$(-n)a := n(-a)$$

3. 对任意整数 $n, m$，任意 $a, b \in R$，有

$$(n + m)a = na + ma$$

$$(nm)a = n(ma)$$

$$n(a + b) = na + nb$$

$$m(ab) = (ma)b = a(mb)$$

性质 3 也参看阅读材料一.

**命题** 8.7.2 任一有限域的特征一定是一个素数；并且如果它的特征为 $p$，则它含有一个子域，其元素数目为 $p$.

**证明** 设 $F$ 是有限域，其单位元记为 $e$. 若 $F$ 的特征为零，则 $e, 2e, \cdots, ne, \cdots$ 是 $F$ 中两两不同的元素，于是 $F$ 为无限域，矛盾. 所以 $F$ 的特征为一个素数 $p$. 令

$$F_p = \{0, e, 2e, 3e, \cdots, (p-1)e\}$$

在 $F_p$ 中任取两个元素 $ie, je$，设

$$i - j = k_1 p + r_1$$

其中 $0 \leqslant r_1 < p$. 则

$$\begin{aligned} ie - je &= (i - j)e = (k_1 p + r_1)e = (k_1 p)e + r_1 e \\ &= k_1(pe) + r_1 e = k_1 0 + r_1 e = r_1 e \in F_p \end{aligned}$$

类似地可证，$(ie)(je) \in F_p$. 因此，$F_p$ 是 $F$ 的一个子环. 显然 $e$ 是 $F_p$ 的单位元，且 $F_p$ 为交换环. 剩下只要证 $F_p$ 的每个非零元 $ie$ 在 $F_p$ 中有逆元，则 $F_p$ 就是一个域，从而命题 8.7.2 得证.

因为 $i$ 与 $p$ 互素，所以存在整数 $u, v$，使得

$$ui + vp = 1$$

从而 $\quad (ui)e + (vp)e = e$

即 $\quad e = u(ie) + v(pe) = u(e \cdot ie) = (ue)(ie)$

设 $\quad u = k_2 p + r_2, \qquad 0 \leqslant r_2 < p$

则 $$ue=(k_2p+r_2)e=r_2e\in F_p$$
因此 $ie$ 在 $F_p$ 中有逆元 $r_2e$. ▌

**定理 8.7.2 的证明** 设 $F$ 是有限域，其特征为 $p$. 则它含有子域 $F_p$. $F$ 可以看成是域 $F_p$ 上的线性空间，其中加法是 $F$ 的加法，纯量乘法是 $F_p$ 中元素与 $F$ 中元素做 $F$ 的乘法，线性空间的八条公理容易验证. 因为 $F$ 只含有限多个元素，所以 $F$ 一定是有限维的，设其维数为 $n$. 则 $F\cong F_p^n$. 于是 $F$ 到 $F_p^n$ 有一个双射. 因为 $|F_p^n|=p^n$，所以 $|F|=p^n$. ▌

## 习 题 8.7

1. 证明：实数域作为它自身上的线性空间与习题 8.1 第 2 题(1) 中的线性空间同构.

2. 证明：域 $F$ 上的线性空间 $M_{s\times n}(F)$ 与 $F^{m}$ 同构，并且写出一个同构映射.

3. 证明：域 $F$ 上次数 $<n$ 的一元多项式组成的线性空间 $F[x]_n$ 与 $F^n$ 同构，并且写出一个同构映射.

4. 设 $c_1,c_2,\cdots,c_k$ 是给定的 $k$ 个不同的实数，$n>k$，在 $R[x]_n$ 中以 $c_1,c_2,\cdots,c_k$ 为根的多项式组成的集合记作 $W$，即
$$W=\{f(x)\in R[x]_n\,|\,f(c_i)=0,i=1,\cdots,k\}$$
证明 $W$ 是 $R[x]_n$ 的一个子空间，并且求 $\dim W$.

5. 设域 $F$ 上线性空间 $V$ 中的向量组 $\alpha_1,\alpha_2,\alpha_3,\alpha_4$ 线性无关，试问：向量组 $\alpha_1+\alpha_2,\alpha_2+\alpha_3,\alpha_3+\alpha_4,\alpha_4+\alpha_1$ 是否线性无关？令
$$W=\langle\alpha_1+\alpha_2,\alpha_2+\alpha_3,\alpha_3+\alpha_4,\alpha_4+\alpha_1\rangle$$
求 $W$ 的一个基和维数.

6. 令
$$L=\left\{\begin{pmatrix} a & b\\ -b & a\end{pmatrix}\middle|\, a,b\in R\right\}$$
(1) 证明 $L$ 是实线性空间 $M_2(R)$ 的子空间，求 $L$ 的一个基和维数；

(2) 证明复数域 $C$ 作为实数域 $R$ 上的线性空间与 $L$ 同构，并且写出一个同构映射.

$*7$. 令

$$H=\left\{\left.\begin{pmatrix}\alpha & \beta\\ -\bar{\beta} & \bar{\alpha}\end{pmatrix}\right|\alpha,\beta\in C\right\}$$

(1) 证明 $H$ 对于矩阵的加法，以及实数与矩阵的数量乘法(注意实数可看成复数) 构成实数域上的一个线性空间；

(2) 求 $H$ 的一个基和维数；

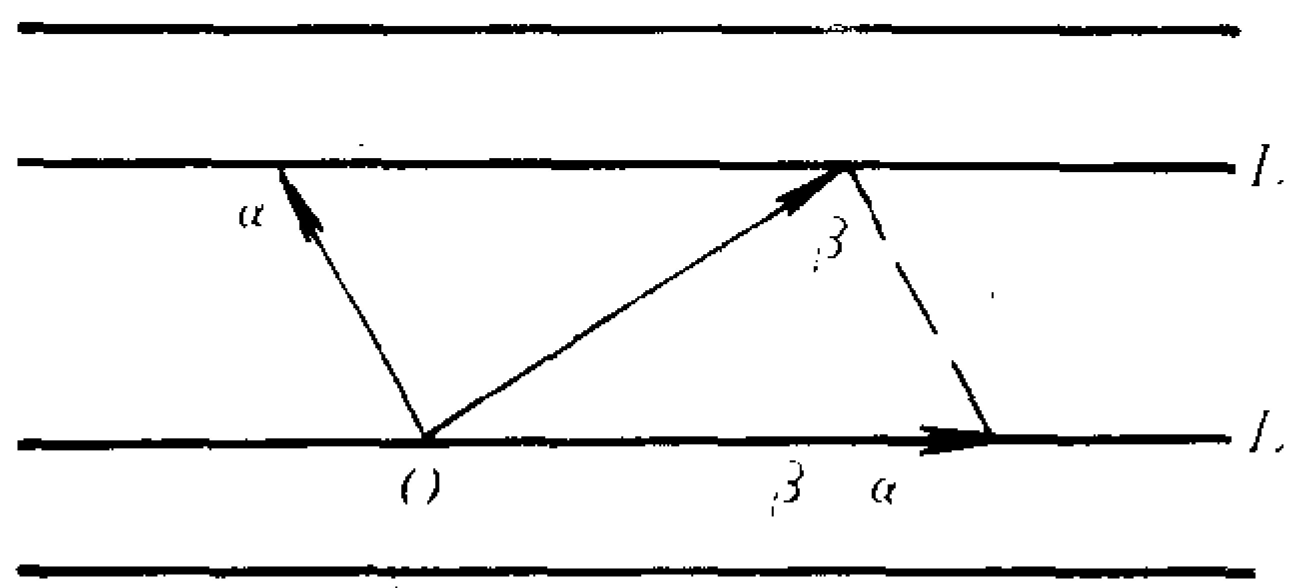
(3) 证明 $H$ 与 $R^4$ 同构，并且写出一个同构映射.

## §8 商空间·余维数

给了域 $F$ 上的一个线性空间 $V$，§5 和 §6 给出了构造 $V$ 的子空间的几种方法. $V$ 的子空间的元素仍是 $V$ 的元素. 本节我们要从 $V$ 和它的一个子空间 $W$ 出发，构造出域 $F$ 上的一个新的线性空间，称它为商空间，它的元素不是 $V$ 的元素，而是 $V$ 的某种类型的子集.

我们先从几何的例子谈起. 平面上以定点 $O$ 为起点的所有向量组成实数域上的一个二维线性空间 $V$. 有的问题中，我们要把 $V$ 中向量组成的某种类型的子集作为研究对象. 譬如，把与过 $O$ 点的一条直线 $L_0$ 平行的每一条直线作为研究对象. 这里我们把一条直线看成是 $V$ 里终点在此直线上的所有向量组成的子集. $L_0$ 是 $V$ 的子空间，与 $L_0$ 平行但不重合的每一条直线都不是 $V$ 的子空间(因为对加法和数量乘法都不封闭). 与 $L_0$ 平行的所有直线(包括 $L_0$) 构成平面 $V$ 的一个划分. 据第五章 §2，可以在 $V$ 中建立一个等价关系，使得与 $L_0$ 平行的每一条直线是一个等价类. 如何建立这个等价关系？设 $L$ 是与 $L_0$ 平行的一条直线，且 $\alpha\in L$(即 $\alpha$ 的终点在 $L$ 上). 从向量加法的平行四边形法则容易得出：向量 $\beta\in L$ 当且仅当 $\beta-\alpha\in L_0$. 由此受到启发，在 $V$ 上定义一个二元关系 $\sim$ 如下：

$$\delta \sim \gamma \text{ 当且仅当 } \delta - \gamma \in L_0$$




容易看出 $\sim$ 是一个等价关系，并且 $L$ 就是 $\alpha$ 确定的等价类 $\bar{\alpha}$(这是因为 $\beta \in L \Leftrightarrow \beta - \alpha \in L_0 \Leftrightarrow \beta \sim \alpha \Leftrightarrow \beta \in \bar{\alpha}$). 于是与 $L_0$ 平行的直线的全体就是 $V$ 的所有 $\sim$ 等价类组成的集合，这个集合称为 $V$ 对于关系 $\sim$ 的**商集**，记作 $V/\sim$（商集的定义见第五章 §2). 注意商集 $V/\sim$ 的元素是等价类，即与 $L_0$ 平行的直线，而不是向量.

从上述几何例子受到启发，一般地，设 $V$ 是域 $F$ 上的线性空间，取 $V$ 的一个子空间 $W$，我们可以在 $V$ 上定义一个二元关系如下：

$$\alpha \sim \beta \text{ 当且仅当 } \alpha - \beta \in W \tag{1}$$

这个关系 $\sim$ 是一个等价关系，理由如下：

任取 $\alpha \in V$，因为 $\alpha - \alpha = 0 \in W$，所以 $\alpha \sim \alpha$. 这表明 $\sim$ 具有反身性.

假设 $\alpha \sim \beta$，则 $\alpha - \beta \in W$. 因为 $W$ 是 $V$ 的子空间，所以

$$\beta - \alpha = -(\alpha - \beta) \in W$$

于是 $\beta \sim \alpha$. 这证明了 $\sim$ 具有对称性.

类似地，利用 $W$ 是 $V$ 的子空间这一事实，可以证明 $\sim$ 具有传递性(请读者练习).

$V$ 的所有 $\sim$ 等价类组成的集合称为 $V$ 对于关系 $\sim$ 的**商集**，记作 $V/\sim$. 由于关系 $\sim$ 是利用 $V$ 的子空间 $W$ 确定的，因此把 $V/\sim$ 也称为 $V$ **对子空间 $W$ 的商集**，记成 $V/W$.

商集 $V/W$ 里的元素是 $\sim$ 等价类. 我们来看 $\alpha \in V$ 确定的等价类 $\bar{\alpha}$ 是由 $V$ 里哪些向量组成的. 因为

$$\begin{aligned}\beta \in \bar{\alpha} &\Leftrightarrow \beta \sim \alpha \Leftrightarrow \beta - \alpha \in W \\ &\Leftrightarrow \beta - \alpha = \gamma, \gamma \in W \\ &\Leftrightarrow \beta = \alpha + \gamma, \gamma \in W\end{aligned} \tag{2}$$

所以
$$\bar{\alpha} = \{\alpha + \gamma \mid \gamma \in W\} =: \alpha + W$$

这里我们用符号 $\alpha + W$ 表示 $V$ 的一个子集 $\{\alpha + \gamma \mid \gamma \in W\}$，称它是 $V$ 的一个 $W$ 型的**线性子簇**. 从(2) 式知道，线性子簇 $\alpha + W$ 等于 $\alpha$ 确定的等价类 $\bar{\alpha}$. 于是

$$V/W = \{\alpha + W \mid \alpha \in V\} \tag{3}$$

即商集 $V/W$ 由 $V$ 的所有 $W$ 型的线性子簇组成. 线性子簇 $\alpha + W$ 中，$\alpha$ 称为代表元. 容易验证，如果 $\beta \in \alpha + W$，那么集合 $\beta + W$ 等于 $\alpha + W$. 因此，线性子簇 $\alpha + W$ 里任何一个元素都可以作为它的代表元. 那么当 $\alpha$ 跑遍 $V$ 时，如何识别哪些线性子族是相同的呢? 由于线性子簇 $\alpha + W$ 等于等价类 $\bar{\alpha}$，根据第五章 §2 命题 5.2.1，两个等价类 $\bar{\alpha}$ 与 $\bar{\beta}$ 相等的充分必要条件是 $\alpha \sim \beta$，因此

$$\alpha + W = \beta + W \Leftrightarrow \alpha - \beta \in W \tag{4}$$

线性空间 $V$ 是具有加法与纯量乘法两种运算的集合，因此我们有理由期望 $V$ 对子空间 $W$ 的商集 $V/W$ 也可以定义加法与纯量乘法两种运算. 但是由于 $V/W$ 不是 $V$ 的子集，因此不能把 $V$ 的两种运算直接搬到 $V/W$ 中来.

**定理** 8.8.1 设 $W$ 是域 $F$ 上线性空间 $V$ 的一个子空间. 在商集 $V/W$ 中规定

$$(\alpha + W) + (\beta + W) = (\alpha + \beta) + W \tag{5}$$

$$k(\alpha + W) = k\alpha + W, k \in F \tag{6}$$

则(5)式和(6)式分别定义了 $V/W$ 中的加法运算与纯量乘法运算，并且 $V/W$ 对于这两种运算成为域 $F$ 上一个线性空间，它的零元素是 $W$(即 $0 + W$)，线性空间 $V/W$ 称为 $V$ 对 $W$ 的**商空间**.

**证明** (5)式和(6)式都是用到线性子簇的代表元，但是一个

线性子簇的代表元可以有很多种选取方式，因此我们必须证明(5)式和(6)式的结果不依赖于代表元的选取. 设

$$\alpha' + W = \alpha + W, \qquad \beta' + W = \beta + W$$

据(4)式得，$\alpha' - \alpha \in W, \beta' - \beta \in W$. 由于 $W$ 是 $V$ 的子空间，所以有

$$(\alpha' + \beta') - (\alpha + \beta) = (\alpha' - \alpha) + (\beta' - \beta) \in W$$
$$k\alpha' - k\alpha = k(\alpha' - \alpha) \in W$$

仍据(4)式得，

$$(\alpha' + \beta') + W = (\alpha + \beta) + W, \quad k\alpha' + W = k\alpha + W$$

这证明了(5)式和(6)式的确定义了 $V/W$ 的加法运算以及 $F$ 与 $V/W$ 的纯量乘法运算.

为了说明 $V/W$ 是域 $F$ 上的一个线性空间，剩下要验证线性空间的八条公理在 $V/W$ 中都成立. 它们都容易从 $V$ 满足这些公理得出，例如公理①，

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

现在验证公理③：任 $\alpha + W \in V/W$，因为

$$(\alpha + W) + (0 + W) = (\alpha + 0) + W = \alpha + W$$

所以 $0 + W = W$ 是 $V/W$ 的零元素. 其余公理都容易验证. ▍

**例** 1　设 $V$ 是平面上以定点 $O$ 为起点的所有向量组成的线性空间，$W$ 是过 $O$ 点的一条直线 $L_0$. 据本节开头一段知道，$V$ 对 $W$ 的商空间 $V/W$ 由与 $L_0$ 平行的所有直线(包括 $L_0$)组成.

**例** 2　第三章 §8 中，我们曾经给出线性流形的概念：设 $W$ 是 $K^n$ 的一个子空间，给定 $\alpha \in K^n$. $K^n$ 的子集

$$\alpha + W := \{\alpha + \eta \mid \eta \in W\}$$

称为一个 $W$ 型的线性流形. 用现在的术语，它就是 $K^n$ 中的一个 $W$ 型的线性子簇. 因此，$K^n$ 的所有 $W$ 型的线性流形组成的集合是 $K^n$ 对 $W$ 的商空间 $K^n/W$.

**例** 3　设 $V$ 是几何空间，$W$ 是过 $O$ 点的一个平面. 容易看出，

任意给定一个向量 $\alpha$，线性子簇 $\alpha+W$ 是由平面 $W$ 沿方向 $\alpha$ 平移得到的平面（据解析几何知道，向量 $\alpha$ 确定的平移是正交变换，而正交变换把平面变成平面）．因此，商空间 $V/W$ 由与 $W$ 平行的所有平面（包括 $W$）组成．

现在我们来讨论商空间的维数．

**定理** 8.8.2　如果 $W$ 是域 $F$ 上有限维线性空间 $V$ 的一个子空间，则

$$\dim(V/W)=\dim V-\dim W \tag{7}$$

**证明**　在 $W$ 中取一个基 $\alpha_1,\cdots,\alpha_s$，把它扩充成 $V$ 的一个基

$$\alpha_1,\cdots,\alpha_s,\alpha_{s+1},\cdots,\alpha_n$$

任取 $\beta+W\in V/W$，设 $\beta=b_1\alpha_1+\cdots+b_n\alpha_n$，则

$$\begin{aligned}\beta+W&=(b_1\alpha_1+\cdots+b_n\alpha_n)+W\\&=(b_1\alpha_1+W)+\cdots+(b_s\alpha_s+W)\\&\quad+(b_{s+1}\alpha_{s+1}+W)+\cdots+(b_n\alpha_n+W)\\&=W+\cdots+W+b_{s+1}(\alpha_{s+1}+W)\\&\quad+\cdots+b_n(\alpha_n+W)\\&=b_{s+1}(\alpha_{s+1}+W)+\cdots+b_n(\alpha_n+W)\end{aligned}\tag{8}$$

这表明 $V/W$ 中任一向量可表成 $\alpha_{s+1}+W,\cdots,\alpha_n+W$ 的线性组合．若能证 $\alpha_{s+1}+W,\cdots,\alpha_n+W$ 线性无关，则它就是 $V/W$ 的一个基，从而

$$\dim(V/W)=n-s=\dim V-\dim W$$

现在假设

$$k_1(\alpha_{s+1}+W)+\cdots+k_{n-s}(\alpha_n+W)=W \tag{9}$$

则　　$(k_1\alpha_{s+1}+\cdots+k_{n-s}\alpha_n)+W=W$

从而　　$k_1\alpha_{s+1}+\cdots+k_{n-s}\alpha_n\in W$

于是　　$k_1\alpha_{s+1}+\cdots+k_{n-s}\alpha_n=l_1\alpha_1+\cdots+l_s\alpha_s$

即　　$$l_1\alpha_1+\cdots+l_s\alpha_s-k_1\alpha_{s+1}-\cdots-k_{n-s}\alpha_n=0 \tag{10}$$

由于 $\alpha_1,\cdots,\alpha_s,\alpha_{s+1},\cdots,\alpha_n$ 线性无关，所以

$$l_1 = \cdots = l_s = k_1 = \cdots = k_{n-s} = 0$$

这证明了 $\alpha_{s+1} + W, \cdots, \alpha_n + W$ 是 $V/W$ 中线性无关的向量组. ▌

数学中许多重要问题会遇到线性空间 $V$ 及其子空间 $W$ 都是无限维,而商空间 $V/W$ 是有限维. 在这种情形,定理 8.8.2 不适用, $\dim(V/W)$ 的计算常常变成非平凡的问题.

**定义** 1　设 $W$ 是域 $F$ 上线性空间 $V$ 的一个子空间,如果 $V$ 对 $W$ 的商空间 $V/W$ 是有限维的,则 $\dim(V/W)$ 称为子空间 $W$ 在 $V$ 中的**余维数**(codimension),记作 $\operatorname{codim}_V W$ 或 $\operatorname{codim} W$.

设 $W$ 是域 $F$ 上线性空间 $V$ 的一个子空间,则 $V$ 到商空间 $V/W$ 有一个很自然的映射 $\pi: \alpha \longmapsto \alpha + W$,称它为标准映射或典范映射(canonical mapping). 显然它是满射,并且 $V/W$ 的一个元素 $\alpha + W$ 上的纤维 $\pi^{-1}(\alpha + W)$ 恰好是线性子簇 $\alpha + W$(作为 $V$ 的一个子集),这是因为

$$\begin{aligned}
&\beta \in \pi^{-1}(\alpha + W) \\
\Leftrightarrow &\pi(\beta) = \alpha + W \Leftrightarrow \beta + W = \alpha + W \\
\Leftrightarrow &\beta - \alpha \in W \Leftrightarrow \beta \sim \alpha \Leftrightarrow \beta \in \bar{\alpha} = \alpha + W
\end{aligned}$$

把这个结果用到例 1 上便得出,与 $L_0$ 平行的每一条直线(它是 $V$ 的一个线性子簇)都是一根纤维.

## 习　题　8.8

1. 设 $U, W$ 都是域 $F$ 上线性空间 $V$ 的子空间,证明: $U + W$ 对 $W$ 的商空间 $U + W/W$ 与 $U$ 对 $U \cap W$ 的商空间 $U/U \cap W$ 同构.

(提示: 规定 $(u + w) + W \longmapsto u + U \cap W$,去证这是 $U + W/W$ 到 $U/U \cap W$ 的一个映射,并且是单射,满射. 再去证此映射保持加法和纯量乘法两种运算).

2. 设 $U, W$ 都是域 $F$ 上线性空间 $V$ 的子空间,并且 $V = U \oplus W$. 证明: $V$ 对 $W$ 的商空间 $V/W$ 与 $U$ 同构.

(提示：去证 $u \to u + W$ 是 $U$ 到 $V/W$ 的同构映射).

3. 设 $V$ 是几何空间，$W$ 是过 $O$ 点的一个平面，$U$ 是过 $O$ 点的一条直线并且它不在平面 $W$ 上.

(1) 求 $V$ 对 $W$ 的商空间 $V/W$ 的维数和一个基；写出 $U$ 到 $V/W$ 的一个同构映射；

(2) $V/U$ 由哪些元素组成？求 $V/U$ 的维数和一个基；写出 $W$ 到 $V/U$ 的一个同构映射.

## 补 充 题 八

1. 设 $V$ 是域 $F$ 上的线性空间，$S$ 是 $V$ 的任一子集. $V$ 中包含 $S$ 的所有子空间的交称为**由 $S$ 生成的子空间**，记作 $\langle S\rangle$，即

$$\langle S\rangle = \bigcap W \tag{1}$$

其中 $W$ 取遍 $V$ 中包含 $S$ 的所有子空间.

(1) 证明：$S \subset \langle S\rangle$；

(2) 用 $T$ 表示由 $S$ 里的任意有限多个向量的所有线性组合组成的集合，证明：$\langle S\rangle = T$.

(提示：第一步，利用第(1)小题结论去证 $T \subset \langle S\rangle$；第二步，去证 $T$ 是 $V$ 的子空间；第三步，去证 $T \supset \langle S\rangle$.)

(3) 当 $S$ 为有限集合时，说明用(1)式定义的 $\langle S\rangle$ 与 §5 关于向量组生成的子空间的定义一致.

2. 设 $V$ 是域 $F$ 上的线性空间，$V_1$ 与 $V_2$ 是 $V$ 的两个子空间，证明：

$$\langle V_1 \cup V_2\rangle = V_1 + V_2$$

3. 在域 $F$ 上的线性空间 $M_{m\times n}(F)$ 中，令

$$M_{m\times n}(F)E_{ii} := \{AE_{ii} \mid A \in M_{m\times n}(F)\}$$

其中 $E_{ii}$ 表示 $(i,i)$ 元为 1，其余元为 0 的矩阵，$1 \leqslant i \leqslant n$. 证明：

(1) $M_{m\times n}(F)E_{ii}$ 是 $M_{m\times n}(F)$ 的子空间，$1 \leqslant i \leqslant n$；

(2) $M_{m\times n}(F) = \bigoplus\limits_{i=1}^{n} M_{m\times n}(F)E_{ii}$.

4. $F[x]_n$ 在 $F[x]$ 中有没有补空间？如果有，请找出来.

5. 设 $A$ 是 $m \times n$ 实矩阵，用 $U$ 表示 $A$ 的列空间，用 $W$ 表示 $AA'$ 的列空

间，证明：$U=W$.

6. 设 $A\in M_n(F)$ 且 $A$ 是可逆的，把 $A$ 和 $A^{-1}$ 分块：

$$A=\begin{array}{cc} & \begin{array}{cc} k & n-k \end{array} \\ & \left(\begin{array}{cc} A_{11} & A_{12} \\ A_{21} & A_{22} \end{array}\right)\begin{array}{l} l \\ n-l \end{array} \end{array} \qquad A^{-1}=\begin{array}{cc} & \begin{array}{cc} l & n-l \end{array} \\ & \left(\begin{array}{cc} B_{11} & B_{12} \\ B_{21} & B_{22} \end{array}\right)\begin{array}{l} k \\ n-k \end{array} \end{array}$$

证明：$A_{12}X=0$ 的解空间 $W$ 同构于 $B_{12}Y=0$ 的解空间 $U$，从而

$$\dim W=\dim U$$

（提示：任取 $\alpha\in W$，有

$$\begin{pmatrix} 0 \\ \alpha \end{pmatrix}=A^{-1}A\begin{pmatrix} 0 \\ \alpha \end{pmatrix}=\begin{pmatrix} B_{12}A_{22}\alpha \\ B_{22}A_{22}\alpha \end{pmatrix}$$

从而得，$B_{12}A_{22}\alpha=0$，即 $A_{22}\alpha\in U$. 于是，令 $\sigma(\alpha)=A_{22}\alpha$，则 $\sigma$ 是 $W$ 到 $U$ 的一个映射. 去证 $\sigma$ 是单射，满射，且保持加法和纯量乘法运算.）

7. 设 $A\in M_n(F)$ 且 $A$ 是可逆的，把 $A$ 分块：

$$A=\begin{array}{cc} & \begin{array}{cc} k & n-k \end{array} \\ & \left(\begin{array}{cc} A_{11} & A_{12} \\ A_{21} & A_{22} \end{array}\right)\begin{array}{l} n-k \\ k \end{array} \end{array}$$

证明：如果 $A_{21}$ 中每个元素在 $A$ 中的余子式等于零，那么 $2k\leqslant n$；并且此时求 $\operatorname{rank}A_{12}$.

（提示：利用 $A^{-1}=|A|^{-1}A^*$，上册习题 2.3 第 5 题的结果以及本补充题八的第 6 题的结果。）

# 第九章　线性映射·线性变换

上一章我们研究了任意域 $F$ 上线性空间的结构.在许多数学问题和实际问题中起着重要作用的是线性空间到线性空间的映射.例如,解析几何中讨论的正交变换和仿射变换;数学分析中讨论的求微商(这是一次可微函数组成的线性空间到实值函数组成的线性空间的一个映射);几何空间沿某一方向到一个平面上的投影;等等.这些映射有一个共同点:保持加法和纯量乘法两种运算,称它们为线性映射.这一章我们来讨论线性映射,它是线性代数的主要研究对象之一.

## §1　线性映射的定义·存在性

**定义 1**　设 $V$ 与 $V'$ 是域 $F$ 上的两个线性空间,$V$ 到 $V'$ 的一个映射 $\underset{\sim}{A}$ 如果保持加法和纯量乘法两种运算,即

$$\underset{\sim}{A}(\alpha+\beta)=\underset{\sim}{A}(\alpha)+\underset{\sim}{A}(\beta),\qquad \forall\ \alpha,\beta\in V$$

$$\underset{\sim}{A}(k\alpha)=k\underset{\sim}{A}(\alpha),\qquad \forall\ \alpha\in V,k\in F$$

那么称 $\underset{\sim}{A}$ 是 $V$ 到 $V'$ 的一个**线性映射**.

与上一章 §7 给出的 $V$ 到 $V'$ 的同构映射的概念比较可看出,线性映射比同构映射少了单射和满射这两条要求.因此线性映射比同构映射广泛得多.$V$ 到 $V'$ 的线性映射也称为**同态映射**,简称为**同态**,如果同态映射是双射,则它就是同构映射.

线性空间 $V$ 到自身的线性映射也称为 $V$ 上的**线性变换**(或 $V$ 上的**线性算子**).

**例** 1　$V$ 到 $V'$ 的**零映射**(即 $\underline{0}(\alpha)=0, \forall\ \alpha\in V$)是线性映射.

**例** 2　$V$ 到自身的**恒等映射**:$\mathrm{id}_v(\alpha)=\alpha, \forall\ \alpha\in V$,是 $V$ 上的一个线性变换. 恒等映射 $\mathrm{id}_v$ 也记作 $I$ 或 $1_v$.

**例** 3　任意给定 $k\in F$,$V$ 到自身的一个映射:

$$\underline{k}(\alpha):=k\alpha,\quad \forall\ \alpha\in V$$

称为 $V$ 上的由 $k$ 决定的**数乘变换**(或**位似变换**). 当 $k=0$ 时,便得到零变换;当 $k=1$ 时,得到恒等变换.

**例** 4　设 $A$ 是域 $F$ 上的一个 $s\times n$ 矩阵,令

$$\underline{A}(\alpha):=A\alpha,\quad \forall\ \alpha\in F^n$$

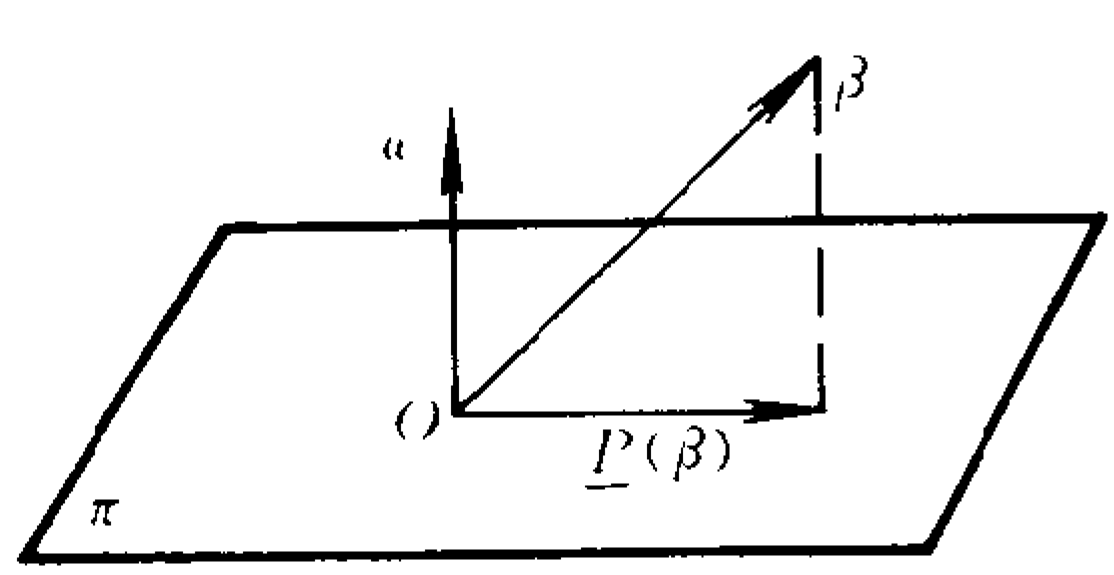
则 $\underline{A}$ 是 $F^n$ 到 $F^s$ 的一个线性映射.

**例** 5　平面上的正交向量变换和仿射向量变换都是平面上的线性变换.

**例** 6　几何空间 $V$ 到经过原点 $O$ 的平面 $\pi$ 上的正投影 $\underline{P}$ 是 $V$ 上的一个线性变换,理由如下:设 $\alpha_0$ 是平面 $\pi$ 的单位法向量,容易看出 $\underline{P}(\beta)$ 是 $\beta$ 沿方向 $\alpha_0$ 的外射影. 根据第四章 §10 的公式(17)可得,

$$\underline{P}(\beta)=\beta-(\beta,\alpha_0)\alpha_0,\quad \forall\ \beta\in V$$

可直接验证 $\underline{P}$ 保持加法和数量乘法.




**例** 7　用 $C^{(1)}(a,b)$ 表示区间 $(a,b)$ 上一次可微函数组成的线性空间,则求微商是 $C^{(1)}(a,b)$ 到 $R^{(a,b)}$ 的一个线性映射,用 $\underline{D}$ 表

示，即

$$\underline{D}(f(x)) = f'(x)$$

**例** 8　线性空间 $C[a,b]$ 到自身上的一个映射

$$\underline{J}(f(x)) = \int_a^x f(t)\mathrm{d}t$$

是 $C[a,b]$ 上的一个线性变换.

* **例** 9　线性空间 $V$ 到商空间 $V/W$ 的标准映射 $\pi: \alpha \mapsto \alpha + W$ 是一个线性映射，理由如下：

$$\begin{aligned}\pi(\alpha+\beta) &= (\alpha+\beta)+W = (\alpha+W)+(\beta+W)\\ &= \pi(\alpha)+\pi(\beta)\end{aligned}$$

$$\pi(k\alpha) = k\alpha + W = k(\alpha+W) = k\pi(\alpha)$$

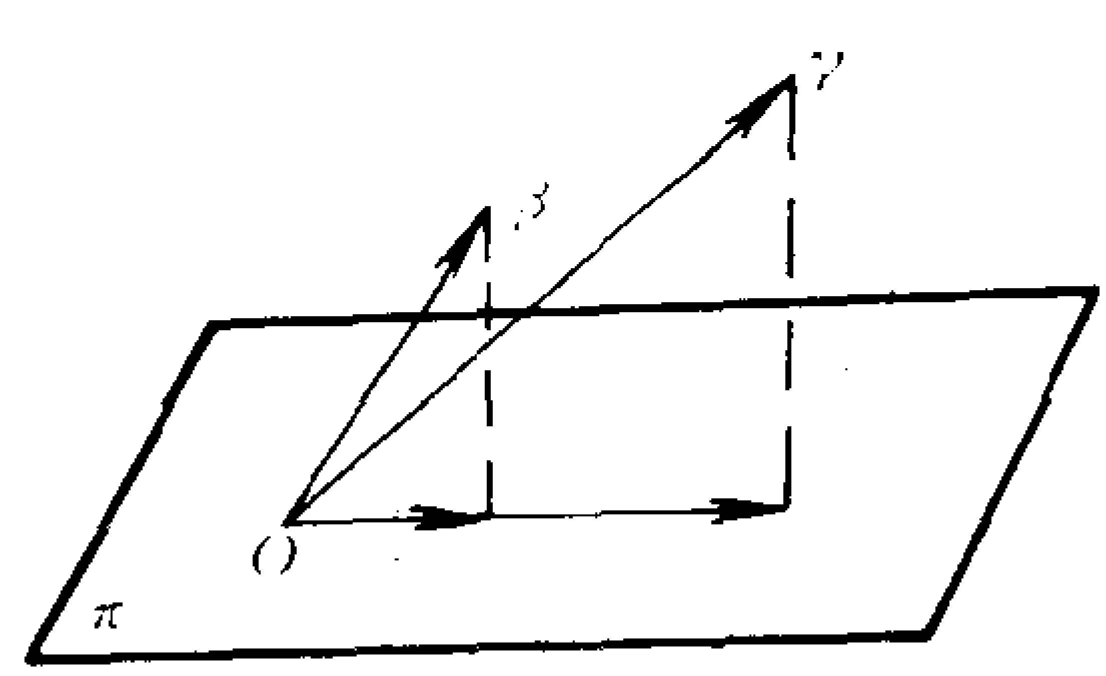
由于线性映射只比同构映射少单射和满射的条件，因此同构映射的性质中，只要证明中没有用到单射和满射的条件，或者可以改变证明方法使得新的证法中不用单射和满射的条件，那么对线性映射也成立. 于是，设 $\underline{A}$ 是线性空间 $V$ 到 $V'$ 的线性映射，则

1. $\underline{A}(0) = 0$；(理由：$\underline{A}(0) = \underline{A}(0\alpha) = 0\underline{A}(\alpha) = 0$.)

2. $\underline{A}(-\alpha) = -\underline{A}(\alpha)$；$\forall\ \alpha \in V$；

3. $\underline{A}(k_1\alpha_1 + \cdots + k_s\alpha_s) = k_1\underline{A}(\alpha_1) + \cdots + k_s\underline{A}(\alpha_s)$；

4. 若 $\alpha_i, \cdots, \alpha_s$ 是 $V$ 的一个线性相关的向量组，则 $\underline{A}(\alpha_1), \cdots, \underline{A}(\alpha_s)$ 是 $V'$ 的一个线性相关的向量组；但是反之不成立，即线性映射可能把线性无关的向量组变成线性相关的向量组.



例如，例 6 的正投影 $\underline{P}$ 把线性无关的向量组 $\beta,\gamma$ 变成线性相关的向量组 $\underline{P}(\beta),\underline{P}(\gamma)$.

从性质 3 看到，如果 $V$ 是有限维的并且在 $V$ 中取定一个基 $\alpha_1,\cdots,\alpha_n$，那么 $V$ 中任一向量

$$\beta = b_1\alpha_1 + \cdots + b_n\alpha_n$$

在线性映射 $\underline{A}$ 下的象

$$\underline{A}\beta = b_1\underline{A}\alpha_1 + b_2\underline{A}\alpha_2 + \cdots + b_n\underline{A}\alpha_n \tag{1}$$

这说明，只要知道了 $V$ 的一个基里的向量在线性映射 $\underline{A}$ 下的象，那么 $V$ 中每一个向量在 $\underline{A}$ 下的象就都确定了. 由此得出

**定理** 9.1.1　设 $V$ 和 $V'$ 都是域 $F$ 上线性空间，且 $V$ 是有限维的，$\underline{A}$ 和 $\underline{B}$ 都是 $V$ 到 $V'$ 的线性映射. 在 $V$ 中取一个基 $\alpha_1,\cdots,\alpha_n$. 如果 $\underline{A}$ 和 $\underline{B}$ 在这个基上的作用相同，即

$$\underline{A}\alpha_i = \underline{B}\alpha_i, \quad i = 1,\cdots,n$$

那么 $\underline{A} = \underline{B}$.

**证明**　任取 $V$ 的一个向量 $\beta = b_1\alpha_1 + \cdots + b_n\alpha_n$，因为

$$\begin{aligned}\underline{A}\beta &= b_1\underline{A}\alpha_1 + \cdots - b_n\underline{A}\alpha_n \\ &= b_1\underline{B}\alpha_1 + \cdots + b_n\underline{B}\alpha_n = \underline{B}\beta\end{aligned}$$

所以 $\underline{A} = \underline{B}$.　▌

定理 9.1.1 表明，$V$ 到 $V'$ 的一个线性映射完全被它在 $V$ 的一个基上的作用所决定. 现在要问：给了域 $F$ 上任意两个线性空间 $V$ 和 $V'$，是否存在 $V$ 到 $V'$ 的线性映射？如果 $V$ 是有限维的，那么回答是肯定的，即

**定理** 9.1.2　设 $V$ 和 $V'$ 都是域 $F$ 上的线性空间，且 $\alpha_1,\cdots,\alpha_n$ 是 $V$ 的一个基. 在 $V'$ 中任意取定 $n$ 个向量 $\gamma_1,\cdots,\gamma_n$（它们中可以有相同的），一定存在 $V$ 到 $V'$ 的唯一的线性映射 $\underline{A}$，使得

$$\underline{A}\alpha_i = \gamma_i, \quad i = 1,\cdots,n$$

**证明**　存在性. 我们来建立 $V$ 到 $V'$ 的一个映射 $\underline{A}$：任取

$$\beta = b_1\alpha_1 + \cdots + b_n\alpha_n \in V$$

规定 $$\underline{A}\beta = b_1\gamma_1 + b_2\gamma_2 + \cdots + b_n\gamma_n \tag{2}$$
由于 $\alpha_1,\cdots,\alpha_n$ 是 $V$ 的一个基,所以(2) 式的确定义了 $V$ 到 $V'$ 的一个映射. 下面来证 $\underline{A}$ 是线性的,在 $V$ 中任取两个向量
$$\beta = \sum_{i=1}^{n} b_i\alpha_i, \qquad \delta = \sum_{i=1}^{n} d_i\alpha_i$$
于是
$$\underline{A}(\beta + \delta) = \underline{A}\left(\sum_{i=1}^{n}(b_i + d_i)\alpha_i\right) = \sum_{i=1}^{n}(b_i + d_i)\gamma_i$$
$$= \sum_{i=1}^{n} b_i\gamma_i + \sum_{i=1}^{n} a_i\gamma_i = \underline{A}\beta + \underline{A}\delta$$
$$\underline{A}(k\beta) = \underline{A}\left(\sum_{i=1}^{n}(kb_i)\alpha_i\right) = \sum_{i=1}^{n}(kb_i)\gamma_i$$
$$= k\sum_{i=1}^{n} b_i\gamma_l = k\underline{A}\beta, \quad \forall\ k \in F$$
因此 $\underline{A}$ 是 $V$ 到 $V'$ 的一个线性映射. 从(2) 式立即得出
$$\begin{aligned}\underline{A}\alpha_i &= \underline{A}(0\alpha_1 + \cdots + 0\alpha_{i-1} + 1\alpha_i + 0\alpha_{i+1} + \cdots + 0\alpha_n) \\ &= 0\gamma_1 + \cdots + 0\gamma_{i-1} + 1\gamma_i + 0\gamma_{i+1} + \cdots + 0\gamma_n \\ &= \gamma_i, \quad i = 1,\cdots,n\end{aligned}$$
唯一性从定理 9.1.1 立即得到. ▌

定理 9.1.2 表明,域 $F$ 上有限维线性空间 $V$ 到域 $F$ 上任意一个线性空间 $V'$ 的线性映射一定存在,并且只要在 $V$ 中取定一个基 $\alpha_1,\cdots,\alpha_n$,则对于 $V'$ 中任意 $n$ 个向量 $\gamma_1,\cdots,\gamma_n$,都有唯一的一个线性映射 $\underline{A}$,使 $\underline{A}(\alpha_i) = \gamma_i, i = 1,\cdots,n$.

如果 $V$ 是域 $F$ 上任一线性空间(有限维,或者无限维),我们来讨论 $V$ 上线性变换的存在性问题.

**定理** 9.1.3 设 $V$ 是域 $F$ 上任一线性空间,如果存在 $V$ 的子空间 $U,W$,使得 $V = U \oplus W$,那么存在 $V$ 上唯一的一个线性变换 $\underline{P}_U$,使得
$$\underline{P}_U(\delta) = \begin{cases}\delta, & \text{当 } \delta \in U \\ 0, & \text{当 } \delta \in W\end{cases} \tag{3}$$
这个线性变换 $\underline{P}_U$ 称为平行于 $W$ 在 $U$ 上的**投影**(**射影**).

**证明** 任取 $\alpha \in V$，设 $\alpha = \alpha_1 + \alpha_2, \alpha_1 \in U, \alpha_2 \in W$. 令

$$\underline{P}_U(\alpha) = \alpha_1$$

则 $\underline{P}_U$ 是 $V$ 到 $V$ 的一个映射(因为 $\alpha$ 写成 $\alpha = \alpha_1 + \alpha_2$ 的表法唯一). 下面来证 $\underline{P}_U$ 是线性的. 设 $\beta = \beta_1 + \beta_2, \beta_1 \in U, \beta_2 \in W$，则

$$\begin{aligned}\underline{P}_U(\alpha + \beta) &= \underline{P}_U((\alpha_1 + \beta_1) + (\alpha_2 + \beta_2)) \\ &= \alpha_1 + \beta_1 = \underline{P}_U(\alpha) + \underline{P}_U(\beta)\end{aligned}$$

$$\begin{aligned}\underline{P}_U(k\alpha) &= \underline{P}_U(k\alpha_1 + k\alpha_2) \\ &= k\alpha_1 = k\underline{P}_U(\alpha), \quad \forall\, k \in F\end{aligned}$$

因此 $\underline{P}_U$ 是 $V$ 上的一个线性变换. 如果 $\delta \in U$，则

$$\underline{P}_U(\delta) = \underline{P}_U(\delta + 0) = \delta$$

如果 $\delta \in W$，则

$$\underline{P}_U(\delta) = \underline{P}_U(0 + \delta) = 0$$

存在性得证.

唯一性. 设 $V$ 上的线性变换 $\underline{A}$ 满足(3)式. 任取 $\alpha \in V$，设

$$\alpha = \alpha_1 + \alpha_2, \alpha_1 \in U, \alpha_2 \in W$$

则 $\quad \underline{A}\alpha = \underline{A}(\alpha_1 + \alpha_2) = \underline{A}\alpha_1 + \underline{A}\alpha_2 = \alpha_1 + 0 = \alpha_1 = \underline{P}_U(\alpha)$

因此，$\underline{A} = \underline{P}_U$. ▌

类似地，定义 $\underline{P}_W(\alpha) = \alpha_2$，则 $\underline{P}_W$ 是线性变换. 称它为平行于 $U$ 在 $W$ 上的投影.

如果 $V$ 为几何空间，$U$ 是过原点 $O$ 的一个平面，$W$ 是过原点 0 且不在平面 $U$ 上的任意一条直线，则 $\underline{P}_U$ 是平行于 $W$ 在 $U$ 上的投影. 特别地，若 $W$ 与平面 $U$ 垂直，则 $\underline{P}_U$ 就是例 6 中的正投影 $\underline{P}$.

投影 $\underline{P}_U$ 是一类非常重要的线性变换.

## 习 题 9.1

1. 判别下面所定义的变换，哪些是线性变换，哪些不是：

(1) 在域 $F$ 上线性空间 $V$ 中，$\underset{\sim}{A}(\beta)=k\beta+\alpha_0$，其中 $\alpha_0$ 是 $V$ 里一个固定的向量，$k$ 是 $F$ 中一个固定元素；

(2) 在 $K^3$ 中，$\underset{\sim}{A}(x_1,x_2,x_3)=(x_1-x_2,x_2+x_3,x_3^2)$；

(3) 在 $K^3$ 中，$\underset{\sim}{A}(x_1,x_2,x_3)=(2x_1-x_2,x_2+x_3,3x_1-x_2+x_3)$；

(4) 在 $F[x]$ 中，$\underset{\sim}{A}f(x)=f(x+a)$，其中 $a$ 是 $F$ 中一个固定元素；

(5) 把复数域分别看作实数域和复数域上的线性空间，$\underset{\sim}{A}z=\bar{z}$；

(6) 在 $M_n(F)$ 中，$\underset{\sim}{A}(X)=BXC$，其中 $B,C\in M_n(F)$ 是两个固定矩阵.

2. 设 $V$ 是习题 8.1 第 2 题(1) 的实线性空间，$V'$ 是实数域 $R$ 看成自身上的线性空间，判断映射

$$\log_a: V \longrightarrow V'$$
$$x \longmapsto \log_a x$$

是不是 $V$ 到 $V'$ 的一个线性映射，其中 $a>0$ 且 $a\neq 1$.

3. 设 $X$ 为任一集合，$x_0\in X$. 证明：域 $F$ 上的线性空间 $F^X$ 到 $F$($F$ 看成自身上的线性空间) 的下述映射

$$\underset{\sim}{A}f=f(x_0),\quad \forall\, f\in F^X$$

是线性映射.

*4. 设 $S\subset X$ 是两个集合，$F$ 是域. 对于 $f\in F^X$，用 $f|S$ 表示函数 $f$ 在 $S$ 上的限制，即 $f|S$ 的定义域是 $S$，并且对任一 $s\in S$，

$$(f|S)(s)=f(s)$$

证明：$F^X$ 到 $F^S$ 的映射 $\sigma: f\longmapsto f|S$ 是线性映射.

*5. 设 $V$ 是 $K[x,y]$ 中所有 $m$ 次齐次多项式添上零多项式组成的线性空间，设 $A=(a_{ij})\in M_2(K)$. 定义 $V$ 到自身的一个映射 $\underset{\sim}{A}$ 如下：

$$(\underset{\sim}{A}f)(x,y):=f(a_{11}x+a_{21}y,a_{12}x+a_{22}y)$$

判断 $\underset{\sim}{A}$ 是不是 $V$ 上的一个线性变换.

## § 2 线性映射的运算

这一节我们来讨论线性映射可以做哪些运算.

设 $V,U,W$ 都是域 $F$ 上的线性空间，我们把 $V$ 到 $U$ 的所有线性映射组成的集合记作 $\mathrm{Hom}(V,U)$. 类似地，$\mathrm{Hom}(U,W)$ 表示 $U$

到 $W$ 的所有线性映射组成的集合.

设 $\underset{\sim}{A}\in \mathrm{Hom}(V,U)$，$\underset{\sim}{B}\in \mathrm{Hom}(U,W)$，由于映射有乘法运算，因此线性映射有乘法运算，乘积映射 $\underset{\sim}{B}\,\underset{\sim}{A}$ 是 $V$ 到 $W$ 的映射. 它是不是线性映射?回答是肯定的，即

**命题** 9.2.1　设 $\underset{\sim}{A}\in \mathrm{Hom}(V,U)$，$\underset{\sim}{B}\in \mathrm{Hom}(U,W)$，则

$$\underset{\sim}{B}\,\underset{\sim}{A}\in \mathrm{Hom}(V,W)$$

**证明**　设 $\alpha,\beta\in V$，$k\in F$，则

$$\begin{aligned}(\underset{\sim}{B}\,\underset{\sim}{A})(\alpha+\beta) &= \underset{\sim}{B}\,(\underset{\sim}{A}\,(\alpha+\beta)) = \underset{\sim}{B}\,(\underset{\sim}{A}\,\alpha+\underset{\sim}{A}\,\beta)\\ &= \underset{\sim}{B}\,(\underset{\sim}{A}\,\alpha)+\underset{\sim}{B}\,(\underset{\sim}{A}\,\beta)\\ &= (\underset{\sim}{B}\,\underset{\sim}{A})\alpha+(\underset{\sim}{B}\,\underset{\sim}{A})\beta\end{aligned}$$

$$\begin{aligned}(\underset{\sim}{B}\,\underset{\sim}{A})(k\alpha) &= \underset{\sim}{B}\,(\underset{\sim}{A}\,(k\alpha)) = \underset{\sim}{B}\,(k(\underset{\sim}{A}\,\alpha))\\ &= k(\underset{\sim}{B}\,(\underset{\sim}{A}\,\alpha)) = k((\underset{\sim}{B}\,\underset{\sim}{A})\alpha)\end{aligned}$$

因此 $\underset{\sim}{B}\,\underset{\sim}{A}$ 是线性映射.　▌

由于映射的乘法适合结合律，因此线性映射的乘法适合结合律. 由于映射的乘法不适合交换律，因此线性映射的乘法不适合交换律. 设 $\underset{\sim}{A}\in \mathrm{Hom}(V,U)$，则

$$\underset{\sim}{A}\,1_V=\underset{\sim}{A},\qquad 1_U\underset{\sim}{A}=\underset{\sim}{A} \tag{1}$$

第四章 § 2 中，我们给出了可逆映射的概念.

**命题** 9.2.2　设 $\underset{\sim}{A}\in \mathrm{Hom}(V,U)$，如果 $\underset{\sim}{A}$ 是可逆的，则

$$\underset{\sim}{A}^{-1}\in \mathrm{Hom}(U,V)$$

**证明**　因为 $\underset{\sim}{A}$ 可逆，所以

$$\underset{\sim}{A}\underset{\sim}{A}^{-1}=1_U,\qquad \underset{\sim}{A}^{-1}\underset{\sim}{A}=1_V$$

于是对任意 $\delta,\gamma\in U$，任意 $k\in F$，有

$$\begin{aligned}\underset{\sim}{A}^{-1}(\delta+\gamma) &= \underset{\sim}{A}^{-1}[(\underset{\sim}{A}\underset{\sim}{A}^{-1})\delta+(\underset{\sim}{A}\underset{\sim}{A}^{-1})\gamma]\\ &= \underset{\sim}{A}^{-1}[\underset{\sim}{A}(\underset{\sim}{A}^{-1}\delta)+\underset{\sim}{A}(\underset{\sim}{A}^{-1}\gamma)]\\ &= \underset{\sim}{A}^{-1}[\underset{\sim}{A}(\underset{\sim}{A}^{-1}\delta+\underset{\sim}{A}^{-1}\gamma)]\\ &= (\underset{\sim}{A}^{-1}\underset{\sim}{A})(\underset{\sim}{A}^{-1}\delta+\underset{\sim}{A}^{-1}\gamma) = \underset{\sim}{A}^{-1}\delta+\underset{\sim}{A}^{-1}\gamma\end{aligned}$$

$$
\begin{aligned}
\underline{A}^{-1}(k\delta) &= \underline{A}^{-1}[k((\underline{A}\underline{A}^{-1})\delta)] = \underline{A}^{-1}[k(\underline{A}(\underline{A}^{-1}\delta))] \\
&= \underline{A}^{-1}[\underline{A}(k(\underline{A}^{-1}\delta))] \\
&= (\underline{A}^{-1}\underline{A})(k\underline{A}^{-1}\delta) = k\underline{A}^{-1}\delta
\end{aligned}
$$

因此 $\underline{A}^{-1} \in \mathrm{Hom}(U,V)$. ▌

因为 $V$ 到 $U$ 的映射 $\underline{A}$ 可逆的充分必要条件为 $\underline{A}$ 是双射，所以 $\underline{A}$ 是 $V$ 到 $U$ 的可逆的线性映射当且仅当 $\underline{A}$ 是 $V$ 到 $U$ 的同构映射.

由于域 $F$ 上两个有限维线性空间 $V$ 和 $U$ 同构的充分必要条件是它们的维数相同，因此 $V$ 到 $U$ 的可逆的线性映射存在的充分必要条件是 $\dim V = \dim U$.

特别地，有限维线性空间 $V$ 一定存在可逆的线性变换. $\underline{A}$ 是 $V$ 上可逆的线性变换当且仅当 $\underline{A}$ 是 $V$ 到自身的一个同构映射. $V$ 到自身的同构映射称为 $V$ 的一个自同构.

由于线性空间有加法运算，因此可以定义线性映射的加法运算.

设 $\underline{A},\underline{B}\in \mathrm{Hom}(V,U)$，定义它们的**和** $\underline{A}+\underline{B}$ 为

$$(\underline{A}+\underline{B})\alpha := \underline{A}\,\alpha + \underline{B}\,\alpha, \qquad \forall\, \alpha \in V \tag{2}$$

容易证明 $\underline{A}+\underline{B}$ 也是 $V$ 到 $U$ 的线性映射. 理由如下：

$$
\begin{aligned}
(\underline{A}+\underline{B})(\alpha+\beta) &= \underline{A}\,(\alpha+\beta) + \underline{B}\,(\alpha+\beta) \\
&= \underline{A}\,\alpha + \underline{A}\,\beta + \underline{B}\,\alpha + \underline{B}\,\beta \\
&= (\underline{A}\,\alpha + \underline{B}\,\alpha) + (\underline{A}\,\beta + \underline{B}\,\beta) \\
&= (\underline{A}+\underline{B})\alpha + (\underline{A}+\underline{B})\beta \\
(\underline{A}+\underline{B})(k\alpha) &= \underline{A}\,(k\alpha) + \underline{B}\,(k\alpha) \\
&= k\underline{A}\,\alpha + k\underline{B}\,\alpha = k(\underline{A}\,\alpha + \underline{B}\,\alpha) \\
&= k(\underline{A}+\underline{B})\alpha
\end{aligned}
$$

这证明了 $\underline{A}+\underline{B}$ 是线性映射.

容易验证，线性映射的加法适合交换律，结合律. 零映射 $\underline{0}$ 具有性质：

$$\underline{A}+\underline{0}=\underline{A},\quad \forall\ \underline{A}\in \mathrm{Hom}(V,U) \tag{3}$$

对于每个 $\underline{A}\in \mathrm{Hom}(V,U)$，可以定义它的负映射 $-\underline{A}$ 为：

$$(-\underline{A})\alpha:=-(\underline{A}\,\alpha),\quad \forall\ \alpha\in V \tag{4}$$

容易验证 $-\underline{A}$ 也是 $V$ 到 $U$ 的线性映射，并且有

$$\underline{A}+(-\underline{A})=\underline{0} \tag{5}$$

线性映射的减法定义为

$$\underline{A}-\underline{B}:=\underline{A}+(-\underline{B})$$

线性映射的乘法对于加法有左右分配律，即设 $\underline{A},\underline{B}\in \mathrm{Hom}(V,U)$，$\underline{C}\in \mathrm{Hom}(U,W)$，$\underline{D}\in \mathrm{Hom}(M,V)$，则

$$\underline{C}\,(\underline{A}+\underline{B})=\underline{C}\,\underline{A}+\underline{C}\,\underline{B} \tag{6}$$

$$(\underline{A}+\underline{B})\underline{D}=\underline{A}\,\underline{D}+\underline{B}\,\underline{D} \tag{7}$$

**左分配律的证明** 对任意 $\alpha\in V$，有

$$\begin{aligned}(\underline{C}\,(\underline{A}+\underline{B}))\alpha &=\underline{C}\,((\underline{A}+\underline{B})\alpha)=\underline{C}\,(\underline{A}\,\alpha+\underline{B}\,\alpha)\\ &=\underline{C}\,(\underline{A}\,\alpha)+\underline{C}\,(\underline{B}\,\alpha)=(\underline{C}\,\underline{A})\alpha+(\underline{C}\,\underline{B})\alpha\\ &=(\underline{C}\,\underline{A}+\underline{C}\,\underline{B})\alpha\end{aligned}$$

所以 $$\underline{C}\,(\underline{A}+\underline{B})=\underline{C}\,\underline{A}+\underline{C}\,\underline{B}$$

右分配律的证明留给读者.

利用线性映射的乘法和数乘变换可以定义线性映射的纯量乘法：

设 $\underline{A}\in \mathrm{Hom}(V,U)$，$k\in F$，定义 $k$ 与 $\underline{A}$ 的**纯量乘积** $k\underline{A}$ 为

$$k\underline{A}:=\underline{k}\,\underline{A} \tag{8}$$

其中 $\underline{k}\in \mathrm{Hom}(U,U)$，据命题 9.2.1 得，$k\underline{A}\in \mathrm{Hom}(V,U)$ 显然有：

$$(k\underline{A})\alpha=k(\underline{A}\,\alpha) \tag{9}$$

直接验证可以知道，对任意 $\alpha\in V$，有

$$(\underline{k}\,\underline{A})\alpha=(\underline{A}\,\underline{k})\alpha$$

其中等号左端的 $\underline{k}\in \mathrm{Hom}(U,U)$，而等号右端的 $\underline{k}\in \mathrm{Hom}(V,V)$，从而有

$$\underline{k}\,\underline{A}=\underline{A}\,\underline{k}$$

注意等号两端的 $\underset{\sim}{k}$ 的不同含义.

容易验证,线性映射的纯量乘法满足以下规则:

$$1\underset{\sim}{A} = \underset{\sim}{A} \tag{10}$$

$$(kl)\underset{\sim}{A} = k(l\underset{\sim}{A}) \tag{11}$$

$$(k+l)\underset{\sim}{A} = k\underset{\sim}{A} + l\underset{\sim}{A} \tag{12}$$

$$k(\underset{\sim}{A} + \underset{\sim}{B}) = k\underset{\sim}{A} + k\underset{\sim}{B} \tag{13}$$

$$k(\underset{\sim}{A}\,\underset{\sim}{C}) = (k\underset{\sim}{A})\underset{\sim}{C} = \underset{\sim}{A}\,(k\underset{\sim}{C}) \tag{14}$$

其中 $k, l \in F$, $\underset{\sim}{A}, \underset{\sim}{B} \in \mathrm{Hom}(V,U)$, $\underset{\sim}{C} \in \mathrm{Hom}(W,V)$.

从以上讨论看出,$\mathrm{Hom}(V,U)$ 有加法和纯量乘法两种运算,并且满足线性空间八条公理,因此 $\mathrm{Hom}(V,U)$ 是域 $F$ 上的一个线性空间.

从以上讨论还可看出,$\mathrm{Hom}(V,V)$ 有加法、纯量乘法、乘法三种运算. 一方面,$\mathrm{Hom}(V,V)$ 是域 $F$ 上的一个线性空间;另一方面,$\mathrm{Hom}(V,V)$ 对于加法、乘法两种运算成为一个有单位元的环;并且乘法与纯量乘法满足:$\forall\ \underset{\sim}{A}, \underset{\sim}{B} \in \mathrm{Hom}(V,V)$, $k \in F$,有

$$k(\underset{\sim}{A}\,\underset{\sim}{B}) = (k\underset{\sim}{A})\underset{\sim}{B} = \underset{\sim}{A}\,(k\underset{\sim}{B}) \tag{15}$$

**定义** 1 一个非空集合 $A$ 如果有加法运算,乘法运算,以及与域 $F$ 的纯量乘法运算,$A$ 对于加法和乘法成为一个环,$A$ 对于加法和纯量乘法成为域 $F$ 上的一个线性空间,并且对一切 $k \in F$, $\alpha$, $\beta \in A$,有

$$k(\alpha\beta) = (k\alpha)\beta = \alpha(k\beta) \tag{16}$$

则 $A$ 称为域 $F$ 上的一个**代数**. 如果 $A$ 作为域 $F$ 上的线性空间是有限维(无限维)的,则称代数 $A$ 是有限维(无限维)的,把 $F$ 上线性空间 $A$ 的维数称为代数 $A$ 的**维数**.

于是 $\mathrm{Hom}(V,V)$ 是域 $F$ 上的一个代数. $\mathrm{Hom}(V,V)$ 也记成 $\mathrm{End}(V)$.

由于线性变换的乘法满足结合律,因此对于正整数 $n$,可以定义线性变换 $\underset{\sim}{A}$ 的 $n$ 次幂:

$$\underset{\sim}{A}^n := \underbrace{\underset{\sim}{A}\underset{\sim}{A}\cdots\underset{\sim}{A}}_{n\text{个}} \tag{17}$$

我们还定义
$$\underset{\sim}{A}^0 := \underset{\sim}{I} \tag{18}$$

容易推出指数法则：
$$\underset{\sim}{A}^{n+m} = \underset{\sim}{A}^n\underset{\sim}{A}^m, \qquad (\underset{\sim}{A}^m)^n = \underset{\sim}{A}^{mn}, \qquad m,n \geqslant 0 \tag{19}$$

当线性变换 $\underset{\sim}{A}$ 可逆时，定义 $\underset{\sim}{A}$ 的负整数幂为
$$\underset{\sim}{A}^{-n} := (\underset{\sim}{A}^{-1})^n \tag{20}$$

这时，指数法则可以推广到负整数幂的情形. 注意，一般说来，$(\underset{\sim}{A}\underset{\sim}{B})^n \neq \underset{\sim}{A}^n\underset{\sim}{B}^n$.

设 $\underset{\sim}{A} \in \mathrm{Hom}(V,V)$，$m$ 是非负整数，形如下述的表达式
$$a_m\underset{\sim}{A}^m + a_{m-1}\underset{\sim}{A}^{m-1} + \cdots + a_0\underset{\sim}{I}, \quad a_i \in F, i = 0,1,\cdots,m \tag{21}$$

称为**线性变换 $\underset{\sim}{A}$ 的多项式**，记成 $f(\underset{\sim}{A})$，它仍是 $V$ 上的一个线性变换. $V$ 上线性变换 $\underset{\sim}{A}$ 的所有多项式组成的集合记作 $F[\underset{\sim}{A}]$，显然，$F[\underset{\sim}{A}]$ 是 $\mathrm{Hom}(V,V)$ 的一个非空子集. 设
$$f(\underset{\sim}{A}) = \sum_{i=0}^{m} a_i\underset{\sim}{A}^i, \quad g(\underset{\sim}{A}) = \sum_{i=0}^{n} b_i\underset{\sim}{A}^i, \quad m \geqslant n$$

从线性变换的运算法则可得出
$$f(\underset{\sim}{A}) \pm g(\underset{\sim}{A}) = \sum_{i=0}^{m}(a_i \pm b_i)\underset{\sim}{A}^i \tag{22}$$

$$f(\underset{\sim}{A})g(\underset{\sim}{A}) = \sum_{s=0}^{m-n}\Big(\sum_{i+j=s} a_ib_j\Big)\underset{\sim}{A}^s \tag{23}$$

因此 $F[\underset{\sim}{A}]$ 对于加法、减法、乘法都封闭，从而 $F[\underset{\sim}{A}]$ 是环 $\mathrm{Hom}(V,V)$ 的一个子环. 从(23)式可得出
$$f(\underset{\sim}{A})g(\underset{\sim}{A}) = g(\underset{\sim}{A})f(\underset{\sim}{A}) \tag{24}$$

因此 $F[\underset{\sim}{A}]$ 是交换环. 称 $F[\underset{\sim}{A}]$ 是**线性变换 $\underset{\sim}{A}$ 的多项式环**. 与第七章 §1 的例 10 类似，$F[\underset{\sim}{A}]$ 可以看成是 $F$ 的一个扩环.

研究线性映射的运算的重要性之一是，线性映射的一些性质，以及线性映射之间的关系可以通过线性映射的运算表示出来. 下面通过几个例子说明这点.

**例** 1　设 $V$ 是域 $F$ 上的线性空间，且 $V=U\oplus W$，则

$$\begin{aligned}&\underset{\sim}{P}_U^2=\underset{\sim}{P}_U,\quad \underset{\sim}{P}_W^2=\underset{\sim}{P}_W,\quad \underset{\sim}{P}_U+\underset{\sim}{P}_W=\underset{\sim}{I}\\&\underset{\sim}{P}_U\underset{\sim}{P}_W=\underset{\sim}{P}_W\underset{\sim}{P}_U=\underset{\sim}{0}\end{aligned}\tag{25}$$

**证明**　任取 $\alpha\in V$，设 $\alpha=\alpha_1+\alpha_2,\alpha_1\in U,\alpha_2\in W$，则

$$\underset{\sim}{P}_U^2(\alpha)=\underset{\sim}{P}_U(\underset{\sim}{P}_U\alpha)=\underset{\sim}{P}_U(\alpha_1)=\alpha_1=\underset{\sim}{P}_U(\alpha)$$

$$\underset{\sim}{P}_W^2(\alpha)=\underset{\sim}{P}_W(\underset{\sim}{P}_W\alpha)=\underset{\sim}{P}_W(\alpha_2)=\alpha_2=\underset{\sim}{P}_W(\alpha)$$

$$(\underset{\sim}{P}_U+\underset{\sim}{P}_W)\alpha=\underset{\sim}{P}_U\alpha+\underset{\sim}{P}_W\alpha=\alpha_1+\alpha_2=\alpha$$

$$(\underset{\sim}{P}_U\underset{\sim}{P}_W)\alpha=\underset{\sim}{P}_U(\underset{\sim}{P}_W\alpha)=\underset{\sim}{P}_U(\alpha_2)=0$$

$$(\underset{\sim}{P}_W\underset{\sim}{P}_U)\alpha=\underset{\sim}{P}_W(\underset{\sim}{P}_U\alpha)=\underset{\sim}{P}_W(\alpha_1)=0$$

所以

$$\begin{aligned}&\underset{\sim}{P}_U^2=\underset{\sim}{P}_U,\quad \underset{\sim}{P}_W^2=\underset{\sim}{P}_W,\quad \underset{\sim}{P}_U+\underset{\sim}{P}_W=\underset{\sim}{I}\\&\underset{\sim}{P}_U\underset{\sim}{P}_W=\underset{\sim}{P}_W\underset{\sim}{P}_U=\underset{\sim}{0}\end{aligned}$$

▌

满足 $\underset{\sim}{A}^2=\underset{\sim}{A}$ 的线性变换 $\underset{\sim}{A}$ 称为**幂等变换**. 如果两个幂等变换 $\underset{\sim}{A},\underset{\sim}{B}$ 满足 $\underset{\sim}{A}\underset{\sim}{B}=\underset{\sim}{B}\underset{\sim}{A}=\underset{\sim}{0}$，则称 $\underset{\sim}{A}$ 与 $\underset{\sim}{B}$ 是**正交的**.

例 1 表明，如果 $V=U\oplus W$，则投影 $\underset{\sim}{P}_U,\underset{\sim}{P}_W$ 是正交的幂等变换，并且它们的和是恒等变换. 这个结果是重要的.

* **例** 2　在解析几何中，我们讲过向量 $\alpha$ 在方向 $\varepsilon$（单位向量）上的内射影 $\alpha_1=(\Pi_\varepsilon\alpha)\varepsilon$，其中 $\Pi_\varepsilon\alpha$ 是 $\alpha$ 在方向 $\varepsilon$ 上的分量. 由分量和内积的关系知道，$(\alpha,\varepsilon)=\Pi_\varepsilon\alpha$. 因此

$$\alpha_1=(\alpha,\varepsilon)\varepsilon\tag{26}$$

我们用 $\underset{\sim}{P}_\delta(\alpha)$ 表示向量 $\alpha$ 在方向 $\delta$ 上的内射影，则从(26)式可得

$$\underset{\sim}{P}_\delta(\alpha)=\frac{(\alpha,\delta)}{(\delta,\delta)}\delta\tag{27}$$

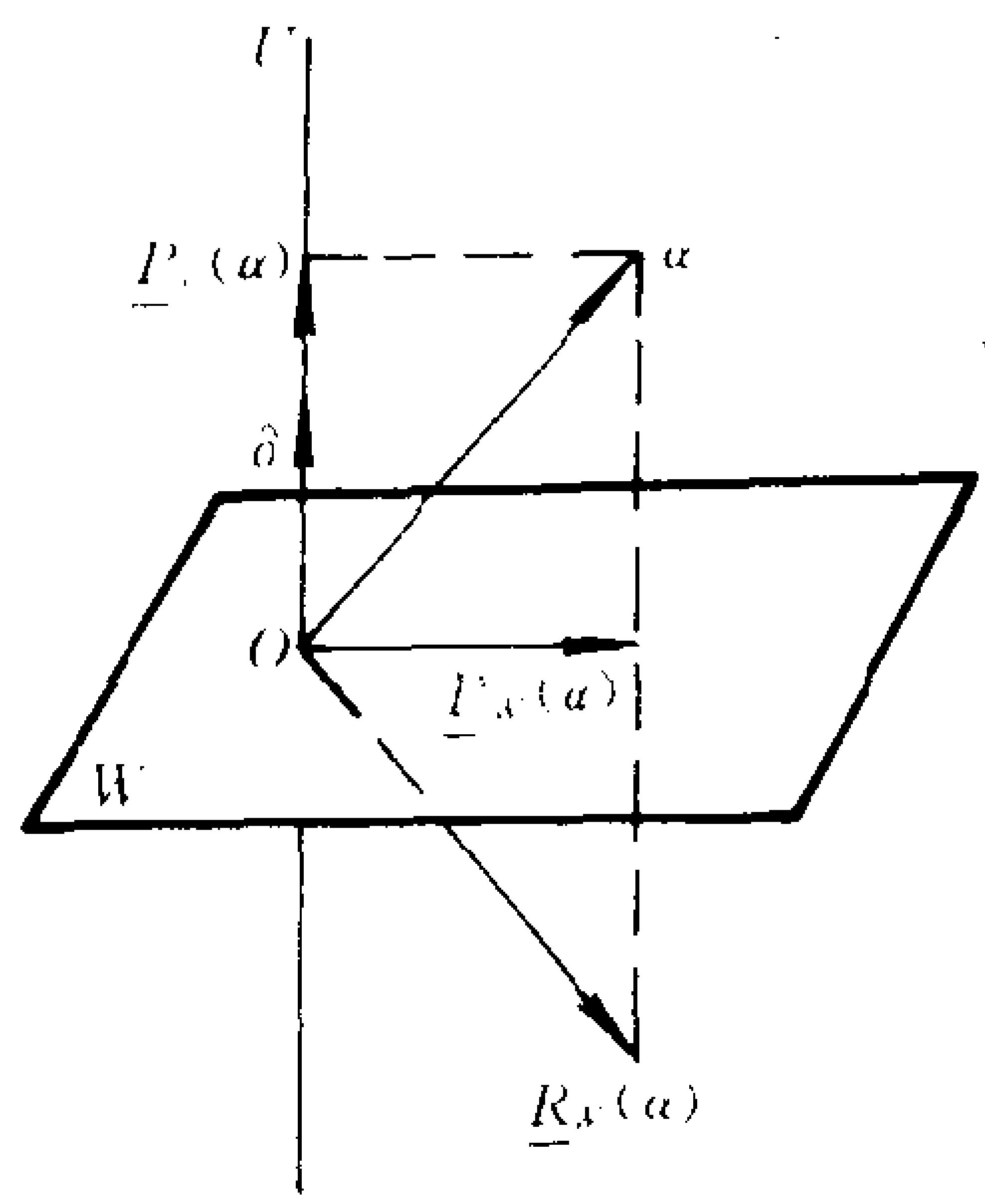
现在设 $V$ 是几何空间，$U$ 是过定点 $O$ 且方向向量为 $\delta$ 的直线，$W$ 是过点 $O$ 且以 $\delta$ 为法向量的平面，则 $V=U\oplus W$，并且 $\underset{\sim}{P}_U=\underset{\sim}{P}_\delta$. 据例 1 得，在平面 $W$ 上的正投影 $\underset{\sim}{P}_W$ 为

$$\underset{\sim}{P}_W=\underset{\sim}{I}-\underset{\sim}{P}_\delta\tag{28}$$

对于平面 $W$ 的反射 $\underset{\sim}{R}_W$ 也是 $V$ 上一个线性变换，容易看出，对于任意 $\alpha\in V$，有

$$\underline{R}_W(\alpha) = \alpha - 2\underline{P}_\delta(\alpha)$$

因此
$$\underline{R}_W = \underline{I} - 2\underline{P}_\delta \tag{29}$$




利用公式(27),(28),(29),我们可以计算任一向量 $\alpha$ 在平面 $W$ 上的射影 $\underline{P}_W(\alpha)$,以及 $\alpha$ 对于平面 $W$ 的反射下的象 $\underline{R}_W(\alpha)$.

**例** 3　在实线性空间 $R[x]_n$ 中,求导数是一个线性变换,用 $\underline{D}$ 表示. 显然有

$$\underline{D}^n = \underline{0} \tag{30}$$

其次,不定元 $x$ 用 $x+a$ 代入($a \in R$):

$$f(x) \longmapsto f(x+a), \quad \forall\ f(x) \in R[x]_n$$

也是一个线性变换,称为**变量的平移**,记作 $\underline{S}_a$. 根据泰勒展开式,对一切 $f(x) \in R[x]_n$,有

$$f(x+a) = f(x) + af'(x) + \frac{a^2}{2!}f''(x) + \cdots + \frac{a^{n-1}}{(n-1)!}f^{(n-1)}(x) \tag{31}$$

从(31)式得到

$$\underline{S}_a = \underline{I} + a\underline{D} + \frac{a^2}{2!}\underline{D}^2 + \cdots + \frac{a^{n-1}}{(n-1)!}\underline{D}^{n-1} \tag{32}$$

这表明，变量的平移 $\underline{S}_a$ 是 $\underline{D}$ 的一个多项式.

## 习 题 9.2

1. 在几何空间中，取右手直角坐标系 $Oxyz$. 用 $\underline{A}$ 表示将空间绕 $x$ 轴按右手螺旋方向旋转 90° 的变换，用 $\underline{B}$ 表示绕 $y$ 轴右旋 90° 的变换，用 $\underline{C}$ 表示绕 $z$ 轴右旋 90° 的变换. 证明：

$$\underline{A}^4 = \underline{B}^4 = \underline{C}^4 = \underline{I},\quad \underline{A}\underline{B} \neq \underline{B}\underline{A},\quad \underline{A}^2\underline{B}^2 = \underline{B}^2\underline{A}^2$$

并检验 $(\underline{A}\underline{B})^2 = \underline{A}^2\underline{B}^2$ 是否成立.

2. 在 $K[x]$ 中，$\underline{D}f(x) = f'(x)$，$\underline{B}f(x) = xf(x)$，证明：

$$\underline{D}\underline{B} - \underline{B}\underline{D} = \underline{I}$$

3. 设 $\underline{A}$，$\underline{B}$ 是 $V$ 上的线性变换，如果 $\underline{A}\underline{B} - \underline{B}\underline{A} = \underline{I}$，证明：

$$\underline{A}^k\underline{B} - \underline{B}\underline{A}^k = k\,\underline{A}^{k-1},\quad k > 1$$

4. 设 $\alpha_1,\cdots,\alpha_n$ 是线性空间 $V$ 的一个基，$\underline{A}$ 是 $V$ 上的一个线性变换，证明：$\underline{A}$ 可逆当且仅当 $\underline{A}\alpha_1,\cdots,\underline{A}\alpha_n$ 线性无关.

5. 设 $\underline{A}$ 是线性空间 $V$ 上的线性变换，如果 $\underline{A}^{m-1}\alpha \neq 0$，但是 $\underline{A}^m\alpha = 0$，证明：$\alpha,\underline{A}\alpha,\cdots,\underline{A}^{m-1}\alpha\,(m > 0)$ 线性无关.

*6. 证明：$M_n(F)$ 是域 $F$ 上的一个代数，并且求这个代数的维数.

*7. 设 $\delta,\gamma$ 是几何空间中的两个向量，证明：$\delta$ 与 $\gamma$ 互相垂直的充分必要条件为 $\underline{P}_\delta\underline{P}_\gamma = \underline{0}$.

8. 设 $V$ 是域 $F$ 上线性空间，$\mathrm{char}F \neq 2$. 设 $\underline{A}$，$\underline{B}$ 是 $V$ 上的幂等变换，证明：

(1) $\underline{A} + \underline{B}$ 是幂等变换的充分必要条件为 $\underline{A}\underline{B} = \underline{B}\underline{A} = \underline{0}$；

(2) 如果 $\underline{A}\underline{B} = \underline{B}\underline{A}$，则 $\underline{A} + \underline{B} - \underline{A}\underline{B}$ 也是幂等变换.

## §3 线性映射的核与象

这一节我们来讨论伴随着线性映射的两类重要的线性子空间：线性映射的核和象。利用它们可以研究线性映射的性质，还可

以研究线性空间的结构.

设 $V$ 和 $V'$ 都是域 $F$ 上的线性空间,不再每次声明.

**定义 1** 设 $\underset{\sim}{A}$ 是 $V$ 到 $V'$ 的一个线性映射. $V'$ 的零向量在 $\underset{\sim}{A}$ 下的原象 $\underset{\sim}{A}^{-1}(0)$ 称为 $\underset{\sim}{A}$ 的**核**,记成 $\mathrm{Ker}\,\underset{\sim}{A}$,即

$$\mathrm{Ker}\,\underset{\sim}{A} = \{\alpha \in V \mid \underset{\sim}{A}\alpha = 0\} \tag{1}$$

**命题 9.3.1** 设 $\underset{\sim}{A}$ 是 $V$ 到 $V'$ 的一个线性映射,则 $\mathrm{Ker}\,\underset{\sim}{A}$ 是 $V$ 的一个线性子空间.

**证明** 因为 $\underset{\sim}{A}(0) = 0$,所以 $\mathrm{Ker}\,\underset{\sim}{A}$ 非空. 设 $\alpha, \beta \in \mathrm{Ker}\,\underset{\sim}{A}$,则

$$\underset{\sim}{A}(\alpha + \beta) = \underset{\sim}{A}\alpha + \underset{\sim}{A}\beta = 0$$

$$\underset{\sim}{A}(k\alpha) = k\,\underset{\sim}{A}\alpha = k0 = 0, \quad \forall\, k \in F$$

因此, $\alpha + \beta, k\alpha \in \mathrm{Ker}\,\underset{\sim}{A}$. 这证明了 $\mathrm{Ker}\,\underset{\sim}{A}$ 是 $V$ 的线性子空间. ▌

**命题 9.3.2** $V$ 到 $V'$ 的线性映射 $\underset{\sim}{A}$ 是单射当且仅当

$$\mathrm{Ker}\,\underset{\sim}{A} = 0$$

**证明** 必要性. 设 $\underset{\sim}{A}$ 是单射. 任取 $\alpha \in \mathrm{Ker}\,\underset{\sim}{A}$,则 $\underset{\sim}{A}\alpha = 0$. 又 $\underset{\sim}{A}(0) = 0$. 因为 $\underset{\sim}{A}$ 是单射,所以 $\alpha = 0$. 于是 $\mathrm{Ker}\,\underset{\sim}{A} = 0$.

充分性. 设 $\alpha_1, \alpha_2 \in V$,如果 $\underset{\sim}{A}\alpha_1 = \underset{\sim}{A}\alpha_2$,则

$$\underset{\sim}{A}(\alpha_1 - \alpha_2) = \underset{\sim}{A}\alpha_1 - \underset{\sim}{A}\alpha_2 = 0$$

从而 $\alpha_1 - \alpha_2 \in \mathrm{Ker}\,\underset{\sim}{A}$. 因为 $\mathrm{Ker}\,\underset{\sim}{A} = 0$,所以 $\alpha_1 - \alpha_2 = 0$,即 $\alpha_1 = \alpha_2$. 这证明了 $\underset{\sim}{A}$ 是单射. ▌

**命题 9.3.3** 设 $\underset{\sim}{A}$ 是 $V$ 到 $V'$ 的一个线性映射,则 $\underset{\sim}{A}$ 的象

$$\mathrm{Im}\,\underset{\sim}{A} = \{\underset{\sim}{A}\alpha \mid \alpha \in V\}$$

是 $V'$ 的一个线性子空间.

**证明** 显然 $\mathrm{Im}\,\underset{\sim}{A}$ 非空集. 任取 $\mathrm{Im}\,\underset{\sim}{A}$ 中两个元素 $\underset{\sim}{A}\alpha, \underset{\sim}{A}\beta$,则

$$\underset{\sim}{A}\alpha + \underset{\sim}{A}\beta = \underset{\sim}{A}(\alpha + \beta) \in \mathrm{Im}\,\underset{\sim}{A}$$

$$k\,\underset{\sim}{A}\alpha = \underset{\sim}{A}(k\alpha) \in \mathrm{Im}\,\underset{\sim}{A}, \quad \forall\, k \in F$$

所以 $\mathrm{Im}\,\underset{\sim}{A}$ 是 $V'$ 的子空间. ▌

$\mathrm{Im}\,\underset{\sim}{A}$ 也记成 $\underset{\sim}{A}V$. $\mathrm{Im}\,\underset{\sim}{A}$ 也称为 $\underset{\sim}{A}$ 的值域.

显然, $V$ 到 $V'$ 的线性映射 $\underset{\sim}{A}$ 是满射当且仅当 $\mathrm{Im}\,\underset{\sim}{A} = V'$.

**定理** 9.3.1　设 $V$ 是有限维线性空间，$\underline{A}$ 是 $V$ 到 $V'$ 的一个线性映射，则 Ker $\underline{A}$ 与 Im $\underline{A}$ 都是有限维的，并且

$$\dim \operatorname{Ker} \underline{A} + \dim \operatorname{Im} \underline{A} = \dim V \tag{2}$$

证明　因为 $V$ 是有限维的，所以它的子空间 Ker $\underline{A}$ 是有限维的. 取 Ker $\underline{A}$ 的一个基 $\alpha_1, \cdots, \alpha_m$，把它扩充成 $V$ 的一个基

$$\alpha_1, \cdots, \alpha_m, \alpha_{m+1}, \cdots, \alpha_n$$

在 Im $\underline{A}$ 中任取一个向量 $\underline{A}\alpha$，其中 $\alpha \in V$. 设 $\alpha = \sum_{i=1}^{n} a_i \alpha_i$，则

$$\underline{A}\alpha = \sum_{i=1}^{n} a_i \underline{A}\alpha_i = a_{m+1}\underline{A}\alpha_{m+1} + \cdots + a_n \underline{A}\alpha_n \tag{3}$$

因此
$$\operatorname{Im}\underline{A} = \langle \underline{A}\alpha_{m+1}, \cdots, \underline{A}\alpha_n \rangle \tag{4}$$

从而 Im $\underline{A}$ 是有限维的. 我们来证 $\underline{A}\alpha_{m+1}, \cdots, \underline{A}\alpha_n$ 线性无关：

假设
$$k_{m+1}\underline{A}\alpha_{m+1} + \cdots + k_n \underline{A}\alpha_n = 0$$

则
$$\underline{A}(k_{m+1}\alpha_{m+1} + \cdots + k_n\alpha_n) = 0$$

于是
$$k_{m+1}\alpha_{m+1} + \cdots + k_n\alpha_n \in \operatorname{Ker} \underline{A}$$

所以
$$k_{m+1}\alpha_{m+1} + \cdots + k_n\alpha_n = l_1\alpha_1 + \cdots + l_m\alpha_m$$

即
$$-l_1\alpha_1 - \cdots - l_m\alpha_m + k_{m+1}\alpha_{m+1} + \cdots + k_n\alpha_n = 0$$

由此即得
$$l_1 = \cdots = l_m = k_{m+1} = \cdots = k_n = 0$$

所以 $\underline{A}\alpha_{m+1}, \cdots, \underline{A}\alpha_n$ 线性无关，从而它是 Im $\underline{A}$ 的一个基. 因此，

$$\dim \operatorname{Im} \underline{A} = n - m = \dim V - \dim \operatorname{Ker} \underline{A}$$ ▎

从(3)式看到，当 $V$ 是有限维时，若 $\alpha_1, \cdots, \alpha_n$ 是 $V$ 的一个基，则

$$\operatorname{Im} \underline{A} = \langle \underline{A}\alpha_1, \cdots, \underline{A}\alpha_n \rangle \tag{5}$$

当 $V$ 是有限维时，$V$ 到 $V'$ 的线性映射 $\underline{A}$ 的核的维数也称为 $\underline{A}$ 的**零度**，$\underline{A}$ 的象 Im $\underline{A}$ 的维数称为 $\underline{A}$ 的**秩**.

**推论** 9.3.1　如果 $V$ 和 $V'$ 都是域 $F$ 上 $n$ 维线性空间，且 $A$ 是 $V$ 到 $V'$ 的一个线性映射，则 $A$ 是单射当且仅当 $A$ 是满射.

**证明**
$$\begin{aligned} A \text{ 是单射 } &\Leftrightarrow \operatorname{Ker} A = 0 \\ &\Leftrightarrow \dim \operatorname{Im} A = \dim V = \dim V' \\ &\Leftrightarrow \operatorname{Im} A = V' \Leftrightarrow A \text{ 是满射} \end{aligned}$$ ▎

**推论** 9.3.2　对于有限维线性空间 $V$ 上的线性变换 $\underset{\sim}{A}$，它是单射当且仅当它是满射. ▌

注意，对于有限维线性空间 $V$ 上的线性变换 $\underset{\sim}{A}$，虽然子空间 $\operatorname{Ker}\underset{\sim}{A}$ 与 $\operatorname{Im}\underset{\sim}{A}$ 的维数之和等于 $\dim V$，但是 $\operatorname{Ker}\underset{\sim}{A}+\operatorname{Im}\underset{\sim}{A}$ 并不一定是整个空间 $V$. 例如，在线性空间 $K[x]_n$ 中，求导数 $\underset{\sim}{D}$ 的象 $\operatorname{Im}\underset{\sim}{D}=K[x]_{n-1}$，$\underset{\sim}{D}$ 的核 $\operatorname{Ker}\underset{\sim}{D}=K$. 显然

$$K+K[x]_{n-1}\neq K[x]_n$$

但是，如果 $\underset{\sim}{A}$ 是 $V$ 上的幂等变换，则有下述结论：

**命题** 9.3.4　如果 $\underset{\sim}{A}$ 是 $V$ 上的幂等变换，则

$$V=\operatorname{Ker}\underset{\sim}{A}\oplus Im\,\underset{\sim}{A} \tag{6}$$

并且 $\underset{\sim}{A}$ 是平行于 $\operatorname{Ker}\underset{\sim}{A}$ 在 $\operatorname{Im}\underset{\sim}{A}$ 上的投影.

**证明**　先证 $V=\operatorname{Ker}\underset{\sim}{A}+\operatorname{Im}\underset{\sim}{A}$. 任取 $\alpha\in V$，若能证

$$\alpha-\underset{\sim}{A}\alpha\in\operatorname{Ker}\underset{\sim}{A}$$

则 $\alpha\in\operatorname{Ker}\underset{\sim}{A}+\operatorname{Im}\underset{\sim}{A}$，从而 $V=\operatorname{Ker}\underset{\sim}{A}+\operatorname{Im}\underset{\sim}{A}$. 因为

$$\begin{aligned}\underset{\sim}{A}(\alpha-\underset{\sim}{A}\alpha)&=\underset{\sim}{A}\alpha-\underset{\sim}{A}(\underset{\sim}{A}\alpha)=\underset{\sim}{A}\alpha-\underset{\sim}{A}^2\alpha\\&=\underset{\sim}{A}\alpha-\underset{\sim}{A}\alpha=0\end{aligned}$$

所以 $\alpha-\underset{\sim}{A}\alpha\in\operatorname{Ker}\underset{\sim}{A}$.

再证和 $\operatorname{Ker}\underset{\sim}{A}+\operatorname{Im}\underset{\sim}{A}$ 是直和. 任取 $\alpha\in\operatorname{Ker}\underset{\sim}{A}\cap\operatorname{Im}\underset{\sim}{A}$. 因为 $\alpha\in\operatorname{Im}\underset{\sim}{A}$，所以存在 $\beta\in V$，使得 $\underset{\sim}{A}\beta=\alpha$. 因为 $\alpha\in\operatorname{Ker}\underset{\sim}{A}$，所以

$$0=\underset{\sim}{A}\alpha=\underset{\sim}{A}(\underset{\sim}{A}\beta)=\underset{\sim}{A}^2\beta=\underset{\sim}{A}\beta=\alpha$$

于是

$$\operatorname{Ker}\underset{\sim}{A}\cap\operatorname{Im}\underset{\sim}{A}=0$$

因此和 $\operatorname{Ker}\underset{\sim}{A}+\operatorname{Im}\underset{\sim}{A}$ 是直和. 这证明了

$$V=\operatorname{Ker}\underset{\sim}{A}\oplus\operatorname{Im}\underset{\sim}{A}$$

如果 $\delta\in\operatorname{Im}\underset{\sim}{A}$，则存在 $\beta\in V$，使得 $\underset{\sim}{A}\beta=\delta$. 从而

$$\underset{\sim}{A}\delta=\underset{\sim}{A}(\underset{\sim}{A}\beta)=\underset{\sim}{A}^2\beta=\underset{\sim}{A}\beta=\delta$$

如果 $\delta\in\operatorname{Ker}\underset{\sim}{A}$，则 $\underset{\sim}{A}\delta=0$. 于是据定理 9.1.3 得，$\underset{\sim}{A}$ 就是平行于 $\operatorname{Ker}\underset{\sim}{A}$ 在 $\operatorname{Im}\underset{\sim}{A}$ 上的投影. ▌

上一节我们曾指出，投影是幂等变换. 现在的命题 9.3.4 则表明，幂等变换是投影.

**推论** 9.3.3 设 $V = U \oplus W$，用 $\underline{P}_W$ 表示平行于 $U$ 在 $W$ 上的投影，则 $\operatorname{Ker} \underline{P}_w = U$，$\operatorname{Im} \underline{P}_W = W$.

**证明** 因为 $\underline{P}_W$ 是幂等变换，据命题 9.3.4 得，$\underline{P}_W$ 是平行于 $\operatorname{Ker} \underline{P}_w$ 在 $\operatorname{Im} \underline{P}_w$ 上的投影，所以 $\operatorname{Ker} \underline{P}_W = U$，$\operatorname{Im} \underline{P}_w = W$. ▌

推论 9.3.3 在几何上看是明显的，请读者画个图想一想.

从推论 9.3.3 容易得出下面的结果：

**命题** 9.3.5 设 $V$ 是域 $F$ 上有限维线性空间，则

(i) $V$ 的任一子空间 $U$ 是 $V$ 上某一线性变换的核；

(ii) $V$ 的任一子空间 $U$ 是 $V$ 上某一线性变换的象.

**证明** (i) 因为 $V$ 是有限维的，所以 $U$ 在 $V$ 里有补空间 $W$，即 $V = U \oplus W$. 据推论 9.3.3，$\operatorname{Ker} \underline{P}_W = U$. 这说明 $U$ 是平行于 $U$ 在 $U$ 的一个补空间 $W$ 上的投影的核.

(ii)的证明留给读者作为习题. ▌

请读者在证明了(ii)之后想一想：有限维线性空间 $V$ 的任一子空间 $U$ 是什么样的投影的象.

命题 9.3.5 表明，有限维线性空间 $V$ 的任一子空间 $U$ 是某一个投影的核，也是另一个投影的象. 由此看到，投影是非常重要的线性变换. 同时也看到，可以通过研究 $V$ 上的线性变换来研究 $V$ 的子空间.

定理 9.3.1 是一个非常有用的结果，它刻画了线性映射 $\underline{A}$ 的核的维数、象的维数与 $\underline{A}$ 的定义域的维数之间的关系. 我们要善于把有关的问题转化为线性映射的问题，从而利用定理 9.3.1. 下面举一个例子：

**例** 1 设 $A \in M_{s\times n}(F)$，令

$$\underline{A}(\alpha) := A\alpha, \quad \forall\ \alpha \in F^n$$

则 $\underline{A}$ 是 $F^n$ 到 $F^s$ 的一个线性映射. 证明：

(1) 齐次线性方程组 $AX = 0$ 的解空间 $W$ 是 $\underline{A}$ 的核；

(2) 矩阵 $A$ 的列空间是 $\underline{A}$ 的象.

**证明** (1) 因为

$$\alpha \in W \Leftrightarrow A\alpha = 0 \Leftrightarrow \underset{\sim}{A}\alpha = 0 \Leftrightarrow \alpha \in \mathrm{Ker}\underset{\sim}{A}$$

所以 $W = \mathrm{Ker}\underset{\sim}{A}$.

(2) 设 $A$ 的列向量组是 $A_1, A_2, \cdots, A_n$,则

$$\begin{aligned}
&\alpha \in \langle A_1, A_2, \cdots, A_n\rangle \\
\Leftrightarrow\ &\alpha = k_1A_1 + k_2A_2 + \cdots + k_nA_n \\
\Leftrightarrow\ &\alpha = (A_1, A_2, \cdots, A_n)\begin{pmatrix} k_1 \\ k_2 \\ \vdots \\ k_n \end{pmatrix} \Leftrightarrow \alpha = A\begin{pmatrix} k_1 \\ k_2 \\ \vdots \\ k_n \end{pmatrix} \\
\Leftrightarrow\ &\alpha = \underset{\sim}{A}\begin{pmatrix} k_1 \\ k_2 \\ \vdots \\ k_n \end{pmatrix} \Leftrightarrow \alpha \in \mathrm{Im}\underset{\sim}{A}
\end{aligned}$$

因此 $\langle A_1, A_2, \cdots, A_n\rangle = \mathrm{Im}\underset{\sim}{A}$. ▌

## 习 题 9.3

*1. 证明:有限维线性空间 $V$ 的任一子空间 $U$ 是 $V$ 上某一线性变换的象.

2. 设 $V$ 和 $V'$ 是域 $F$ 上线性空间,$\underset{\sim}{A}$ 是 $V$ 到 $V'$ 的一个线性映射. 设 $W$ 是 $V$ 的一个子空间,令

$$\underset{\sim}{A}W := \{\underset{\sim}{A}\beta \mid \beta \in W\}$$

证明 $\underset{\sim}{A}W$ 是 $V'$ 的一个线性子空间.

3. 设 $V$ 和 $V'$ 都是域 $F$ 上线性空间,并且 $V$ 是有限维的. 设 $\underset{\sim}{A}$ 是 $V$ 到 $V'$ 的一个线性映射,$W$ 是 $V$ 的一个子空间. 证明

$$\dim \underset{\sim}{A}W + \dim(\mathrm{Ker}\,\underset{\sim}{A} \cap W) = \dim W$$

4. 设 $V, U, W$ 都是域 $F$ 上的线性空间,并且 $V$ 是有限维的. 设 $\underset{\sim}{A} \in \mathrm{Hom}(V, U)$, $\underset{\sim}{B} \in \mathrm{Hom}(U, W)$,证明

$$\dim \mathrm{Ker}(\underset{\sim}{B}\underset{\sim}{A}) \leqslant \dim \mathrm{Ker}\,\underset{\sim}{A} + \dim \mathrm{Ker}\,\underset{\sim}{B}$$

5. 设 $V, U, W$ 都是域 $F$ 上的线性空间,并且 $\dim V = n$, $\dim U = m$. 设

$\underset{\sim}{A} \in \mathrm{Hom}(V,U)$，$\underset{\sim}{B} \in \mathrm{Hom}(U,W)$，证明

$$\mathrm{rank}(\underset{\sim}{B}\underset{\sim}{A}) \geqslant \mathrm{rank}\,\underset{\sim}{A} + \mathrm{rank}\,\underset{\sim}{B} - m$$

*6. 设 $V,U,W,M$ 都是域 $F$ 上的线性空间，并且 $V,U$ 都是有限维的. 设 $\underset{\sim}{A} \in \mathrm{Hom}(V,U)$，$\underset{\sim}{B} \in \mathrm{Hom}(U,W)$，$\underset{\sim}{C} \in \mathrm{Hom}(W,M)$. 证明

$$\mathrm{rank}(\underset{\sim}{C}\underset{\sim}{B}\underset{\sim}{A}) \geqslant \mathrm{rank}(\underset{\sim}{C}\underset{\sim}{B}) + \mathrm{rank}(\underset{\sim}{B}\underset{\sim}{A}) - \mathrm{rank}\,\underset{\sim}{B}$$

*7. 设 $\underset{\sim}{A},\underset{\sim}{B}$ 都是域 $F$ 上线性空间 $V$ 上的幂等变换. 证明：

(1) $\underset{\sim}{A}$ 与 $\underset{\sim}{B}$ 有相同的象的充分必要条件是 $\underset{\sim}{A}\underset{\sim}{B} = \underset{\sim}{B}$，$\underset{\sim}{B}\underset{\sim}{A} = \underset{\sim}{A}$；

(2) $\underset{\sim}{A}$ 与 $\underset{\sim}{B}$ 有相同的的核的充分必要条件是 $\underset{\sim}{A}\underset{\sim}{B} = \underset{\sim}{A}$，$\underset{\sim}{B}\underset{\sim}{A} = \underset{\sim}{B}$.

*8. 设 $V$ 是域 $F$ 上线性空间，$\mathrm{char}F = 0$. 证明：如果 $A_1,\cdots A_s$ 是 $V$ 上 $s$ 个两两不同的线性变换，那么在 $V$ 中至少有一个向量 $\alpha$，使得 $\underset{\sim}{A}_1\alpha,\cdots,\underset{\sim}{A}_s\alpha$ 两两不同.

(提示：利用习题 8.5 的第 12 题).

9. 设 $V$ 和 $V'$ 都是域 $F$ 上有限维线性空间，$\underset{\sim}{A}$ 是 $V$ 到 $V'$ 的一个线性映射. 证明：存在直和分解 $V = U \oplus W$，$V' = M \oplus N$，使得 $\mathrm{Ker}\,\underset{\sim}{A} = U$，并且

$$W \cong M$$

10. 设 $A \in M_{s\times n}(F)$，令 $\underset{\sim}{A}(\alpha) = A\alpha$，$\forall\ \alpha \in F^n$. 证明：线性映射 $\underset{\sim}{A}$ 的秩等于矩阵 $A$ 的秩.

## §4 线性映射(线性变换)与矩阵的关系

研究线性映射的方法有两种：一种是把它作为具有特殊性质(保持线性运算)的映射直接进行研究，例如上一节讨论线性映射的核和象；另一种方法是当线性空间为有限维时，还可以通过矩阵来研究线性映射. 这一节我们来建立线性映射(线性变换)与矩阵的关系.

设 $V$ 和 $V'$ 都是域 $F$ 上有限维线性空间，$\dim V = n$，$\dim V' = s$. 设 $\underset{\sim}{A}$ 是 $V$ 到 $V'$ 的一个线性映射.

从定理 9.1.1 知道，线性映射 $\underset{\sim}{A}$ 完全被它在 $V$ 的一个基上的作用所决定. 因此我们在 $V$ 中取一个基 $\alpha_1,\cdots,\alpha_n$. $V'$ 里的向量组

$\underline{A}\alpha_1,\cdots,\underline{A}\alpha_n$ 就决定了 $\underline{A}$. 由于 $V'$ 也是有限维的，在 $V'$ 中取一个基 $\eta_1,\cdots,\eta_s$，则 $\underline{A}\alpha_1,\cdots,\underline{A}\alpha_n$ 完全被它们在基 $\eta_1,\cdots\eta_s$ 下的坐标所决定，设

$$\begin{cases} \underline{A}\alpha_1 = a_{11}\eta_1 + a_{21}\eta_2 + \cdots + a_{s1}\eta_s \\ \cdots \quad \cdots \quad \cdots \quad \cdots \quad \cdots \quad \cdots \quad \cdots \\ \underline{A}\alpha_n = a_{1n}\eta_1 + a_{2n}\eta_2 + \cdots + a_{sn}\eta_s \end{cases} \tag{1}$$

按照形式写法，(1)式可写成

$$(\underline{A}\alpha_1,\cdots,\underline{A}\alpha_n) = (\eta_1,\cdots,\eta_s)\begin{pmatrix} a_{11} & \cdots & a_{1n} \\ a_{21} & \cdots & a_{2n} \\ \vdots & \vdots & \vdots \\ a_{s1} & \cdots & a_{sn} \end{pmatrix} \tag{2}$$

把(2)式右端的域 $F$ 上的 $s \times n$ 矩阵记作 $A$，它的第 $j$ 列就是 $\underline{A}\alpha_j$ 在 $V'$ 的基 $\eta_1,\cdots,\eta_s$ 下的坐标，$j = 1,\cdots,n$. 矩阵 $A$ 称为**线性映射 $\underline{A}$ 在 $V$ 的基 $\{\alpha_j\}$ 和 $V'$ 的基 $\{\eta_i\}$ 下的矩阵**. 注意 $A$ 的行数 $s = \dim V'$，$A$ 的列数 $n = \dim V$. (2) 式的左端 $(\underline{A}\alpha_1,\cdots,\underline{A}\alpha_n)$ 记成 $\underline{A}(\alpha_1,\cdots,\alpha_n)$. 因此 $A$ 是 $\underline{A}$ 在 $V$ 的基 $\{\alpha_j\}$ 和 $V'$ 的基 $\{\eta_i\}$ 下的矩阵当且仅当下式成立：

$$\underline{A}(\alpha_1,\cdots,\alpha_n) = (\eta_1,\cdots,\eta_s)A \tag{2$'$}$$

上面的议论说明，在 $V$ 中取定一个基 $\{\alpha_j\}$ 和 $V'$ 中取定一个基 $\{\eta_i\}$ 以后，对于 $V$ 到 $V'$ 的每一个线性映射 $\underline{A}$，有唯一确定的 $s \times n$ 矩阵 $A$ 与它对应. 因此，这个对应给出了 $\mathrm{Hom}(V,V')$ 到 $M_{s\times n}(F)$ 的一个映射 $\sigma$. 设 $\underline{B} \in \mathrm{Hom}(V,V')$，$\sigma(\underline{B}) = B$，这里 $B$ 是 $\underline{B}$ 在基 $\{\alpha_j\}$ 和基 $\{\eta_i\}$ 下的矩阵. 如果 $B = A$，则 $\underline{B}\alpha_j = \underline{A}\alpha_j$，$j = 1,\cdots,n$. 据定理 9.1.1 得，$\underline{B} = \underline{A}$. 这表明 $\sigma$ 是单射. 任给 $C \in M_{s\times n}(F)$，$V'$ 中以 $C$ 的第 $j$ 列作为在基 $\{\eta_i\}$ 下的坐标的向量记作 $\boldsymbol{\gamma}_j$，$j = 1,\cdots,n$. 据定理 9.1.2，存在 $V$ 到 $V'$ 的一个线性映射 $\underline{C}$，使得 $\underline{C}\alpha_j = \boldsymbol{\gamma}_j$，$j = 1,\cdots,n$. 从而

$$(\underline{C}\alpha_1,\cdots,\underline{C}\alpha_n) = (\boldsymbol{\gamma}_1,\cdots,\boldsymbol{\gamma}_n) = (\eta_1,\cdots,\eta_s)C$$

于是 $\underset{\sim}{C}$ 是 $C$ 在基 $\{\alpha_j\}$ 和基 $\{\eta_i\}$ 下的矩阵，因此 $\sigma(\underset{\sim}{C}) = C$. 这表明 $\sigma$ 是满射. 因此 $\sigma$ 是 $\mathrm{Hom}(V,V')$ 到 $M_{s\times n}(F)$ 的一个双射. 进一步我们有下述结果：

**定理** 9.4.1　设 $V$ 和 $V'$ 都是域 $F$ 上有限维线性空间，并且 $\dim V = n, \dim V' = s$. 在 $V$ 中取一个基 $\alpha_1,\cdots,\alpha_n$，在 $V'$ 中取一个基 $\eta_1,\cdots,\eta_s$，则 $V$ 到 $V'$ 的每一个线性映射与它在基 $\{\alpha_j\}$ 和 $\{\eta_i\}$ 下的矩阵的对应 $\sigma$ 是线性空间 $\mathrm{Hom}(V,V')$ 到 $M_{s\times n}(F)$ 的同构映射，从而

$$\mathrm{Hom}(V,V') \cong M_{s\times n}(F)$$

**证明**　上一段已证 $\sigma$ 是 $\mathrm{Hom}(V,V')$ 到 $M_{s\times n}(F)$ 的双射. 现在来证 $\sigma$ 保持加法与纯量乘法运算. 任取 $\underset{\sim}{A},\underset{\sim}{B} \in \mathrm{Hom}(V,V')$，设 $\sigma(\underset{\sim}{A}) = A, \sigma(\underset{\sim}{B}) = B$，即

$$\underset{\sim}{A}(\alpha_1,\cdots,\alpha_n) = (\underset{\sim}{A}\alpha_1,\cdots,\underset{\sim}{A}\alpha_n) = (\eta_1,\cdots,\eta_s)A$$

$$\underset{\sim}{B}(\alpha_1,\cdots,\alpha_n) = (\underset{\sim}{B}\alpha_1,\cdots,\underset{\sim}{B}\alpha_n) = (\eta_1,\cdots,\eta_s)B$$

则

$$\begin{aligned}(\underset{\sim}{A}+\underset{\sim}{B})(\alpha_1,\cdots,\alpha_n) &= ((\underset{\sim}{A}+\underset{\sim}{B})\alpha_1,\cdots,(\underset{\sim}{A}+\underset{\sim}{B})\alpha_n)\\ &= (\underset{\sim}{A}\alpha_1+\underset{\sim}{B}\alpha_1,\cdots,\underset{\sim}{A}\alpha_n+\underset{\sim}{B}\alpha_n)\\ &= (\underset{\sim}{A}\alpha_1,\cdots,\underset{\sim}{A}\alpha_n)+(\underset{\sim}{B}\alpha_1,\cdots,\underset{\sim}{B}\alpha_n)\\ &= (\eta_1,\cdots,\eta_s)A+(\eta_1,\cdots,\eta_s)B\\ &= (\eta_1,\cdots\eta_s)(A+B) \qquad (3)\end{aligned}$$

这表明 $\underset{\sim}{A}+B$ 在基 $\{\alpha_j\}$ 和 $\{\eta_i\}$ 下的矩阵是 $A+B$. 因此

$$\sigma(\underset{\sim}{A}+\underset{\sim}{B}) = A+B = \sigma(\underset{\sim}{A})+\sigma(\underset{\sim}{B})$$

类似地，对于任意 $k \in F$，因为

$$(k\underset{\sim}{A})(\alpha_1,\cdots,\alpha_n) = ((k\underset{\sim}{A})\alpha_1,\cdots,(k\underset{\sim}{A})\alpha_n)$$

$$= (k(\underset{\sim}{A}\alpha_1),\cdots,k(\underset{\sim}{A}\alpha_n)) = (k(\sum_{i=1}^{s}a_{i1}\eta_i),\cdots,k(\sum_{i=1}^{s}a_{in}\eta_i))$$

$$= (\sum_{i=1}^{s}(ka_{i1})\eta_i,\cdots,\sum_{i=1}^{s}(ka_{in})\eta_i) = (\eta_1,\cdots\eta_s)(kA) \qquad (4)$$

所以 $k\underset{\sim}{A}$ 在基 $\{\alpha_j\}$ 和 $(\eta_i\}$ 下的矩阵是 $kA$. 因此

$$\sigma(k\underset{\sim}{A}) = kA = k\sigma(\underset{\sim}{A})$$

这证明了 $\sigma$ 是 $\mathrm{Hom}(V,V')$ 到 $M_{s\times n}(F)$ 的同构映射. ∎

如果我们规定

$$k(\gamma_1,\cdots,\gamma_n):=(k\gamma_1,\cdots,k\gamma_n)$$

则从(3)式和(4)式的证明过程中可以顺便看到

$$(\underset{\sim}{A}+\underset{\sim}{B})(\alpha_1,\cdots,\alpha_n)=\underset{\sim}{A}(\alpha_1,\cdots,\alpha_n)+\underset{\sim}{B}(\alpha_1,\cdots,\alpha_n) \tag{5}$$

$$(k\underset{\sim}{A})(\alpha_1,\cdots,\alpha_n)=k(\underset{\sim}{A}(\alpha_1,\cdots\alpha_n)) \tag{6}$$

**推论** 9.4.1　设 $\dim V=n$, $\dim V'=s$,则 $\mathrm{Hom}(V,V')$ 是有限维的,并且

$$\dim\mathrm{Hom}(V,V')=\dim V\cdot\dim V' \tag{7}$$

**证明**　据定理 9.4.1, $\mathrm{Hom}(V,V')\cong M_{s\times n}(F)$. 我们知道 $\dim M_{s\times n}(F)=sn$,据第八章 §7 性质 6, $\mathrm{Hom}(V,V')$ 也是有限维的,并且

$$\dim\mathrm{Hom}(V,V')=\dim M_{s\times n}(F)=sn$$ ∎

在知道了 $V$ 到 $V'$ 的线性映射 $\underset{\sim}{A}$ 在基 $\{\alpha_j\}$ 和 $\{\eta_i\}$ 下的矩阵 $A$ 之后, $V$ 中任一向量 $\alpha$ 在 $\underset{\sim}{A}$ 下的象很容易求出:

**命题** 9.4.1　设 $\alpha_1,\cdots,\alpha_n$ 是 $V$ 的一个基, $\eta_1,\cdots,\eta_s$ 是 $V'$ 的一个基,设 $\underset{\sim}{A}\in\mathrm{Hom}(V,V')$,并且 $\underset{\sim}{A}$ 在基 $\{\alpha_j\}$ 和 $\{\eta_i\}$ 下的矩阵为 $A$. 对于任一 $\alpha\in V$,设 $\alpha$ 在基 $\{\alpha_j\}$ 下的坐标为 $\begin{pmatrix}x_1\\ \vdots\\ x_n\end{pmatrix}$,则 $\underset{\sim}{A}\alpha$ 在基 $\{\eta_i\}$ 下的坐标为 $A\begin{pmatrix}x_1\\ \vdots\\ x_n\end{pmatrix}$.

**证明**　我们有

$$\underset{\sim}{A}\alpha=x_1\underset{\sim}{A}\alpha_1+\cdots+x_n\underset{\sim}{A}\alpha_n=$$

$$(\underset{\sim}{A}\alpha_1,\cdots,\underset{\sim}{A}\alpha_n)\begin{pmatrix}x_1\\ \vdots\\ x_n\end{pmatrix}=((\eta_1,\cdots\eta_s)A)\begin{pmatrix}x_1\\ \vdots\\ x_n\end{pmatrix}=(\eta_1,\cdots,\eta_s)\left(A\begin{pmatrix}x_1\\ \vdots\\ x_n\end{pmatrix}\right)$$

因此，$A\begin{pmatrix}x_1\\ \vdots\\ x_n\end{pmatrix}$是$\underline{A}\alpha$在基$\eta_1,\cdots,\eta_s$下的坐标. ▌

**推论** 9.4.2 设$V$到$V'$的线性映射$\underline{A}$在基$\{\alpha_j\}$和$\{\eta_i\}$下的矩阵为$A$,$V$中任一向量$\alpha$在基$\{\alpha_j\}$下的坐标为$X=\begin{pmatrix}x_1\\ \vdots\\ x_n\end{pmatrix}$,$V'$中向量$\gamma$在基$\{\eta_i\}$下的坐标为$Y=\begin{pmatrix}y_1\\ \vdots\\ y_s\end{pmatrix}$,则

$$\underline{A}\alpha=\gamma\Leftrightarrow AX=Y \tag{8}$$

**证明** 从命题 9.4.1 立即得到. ▌

现在我们来讨论$n$维线性空间$V$上的线性变换与矩阵的关系. 设$\underline{A}$是$V$上的一个线性变换. 我们把上面关于线性映射与矩阵的关系运用到$V$上的线性变换中来. 这时，只需在$V$中取定一个基$\alpha_1,\cdots,\alpha_n$,把基向量$\alpha_j$在$\underline{A}$下的象$\underline{A}\alpha_j$仍然用这个基线性表出，即

$$\begin{cases}\underline{A}\alpha_1=a_{11}\alpha_1+a_{21}\alpha_2+\cdots+a_{n1}\alpha_n\\ \cdots\quad\cdots\quad\cdots\quad\cdots\quad\cdots\quad\cdots\quad\cdots\\ \underline{A}\alpha_n=a_{1n}\alpha_1+a_{2n}\alpha_2+\cdots+a_{nn}\alpha_n\end{cases} \tag{9}$$

按照形式写法，把(9)式写成

$$\begin{aligned}\underline{A}(\alpha_1,\cdots,\alpha_n)&=(\underline{A}\alpha_1,\cdots,\underline{A}\alpha_n)\\ &=(\alpha_1,\cdots,\alpha_n)\begin{pmatrix}a_{11}&\cdots&a_{1n}\\ a_{21}&\cdots&a_{2n}\\ \cdots&\cdots&\cdots\\ a_{n1}&\cdots&a_{nn}\end{pmatrix}\end{aligned} \tag{10}$$

(10)式右端的$n$级矩阵$A=(a_{ij})$称为**线性变换$\underline{A}$在基$\alpha_1,\cdots,\alpha_n$下的矩阵**,它的第$j$列就是$\underline{A}\alpha_j$在基$\alpha_1,\cdots,\alpha_n$下的坐标.

**定理** 9.4.2 设$V$是域$F$上$n$维线性空间，在$V$中取一个基$\alpha_1,\cdots,\alpha_n$,则$V$上的每一个线性变换与它在基$\alpha_1,\cdots,\alpha_n$下的矩阵的

对应 $\sigma$ 是线性空间 $\mathrm{Hom}(V,V)$ 到 $M_n(F)$ 的同构映射，也是环 $\mathrm{Hom}(V,V)$ 到 $M_n(F)$ 的同构映射.

**证明** 结论的前半部分已在定理 9.4.1 中证明. 后半部分中 $\sigma$ 是双射，保持加法也已证明，剩下只要证 $\sigma$ 保持乘法. 设线性变换 $\underline{A},B$ 在基 $\alpha_1,\cdots,\alpha_n$ 下的矩阵分别是 $A,B$. 则

$$\underline{A}(\alpha_1,\cdots\alpha_n)=(\alpha_1,\cdots,\alpha_n)A$$

$$\underline{B}(\alpha_1,\cdots,\alpha_n)=(\alpha_1,\cdots,\alpha_n)B$$

因为

$$(\underline{A}\underline{B})(\alpha_1,\cdots,\alpha_n)=\underline{A}(\underline{B}(\alpha_1,\cdots,\alpha_n))=\underline{A}((\alpha_1,\cdots,\alpha_n)B)$$

$$=\underline{A}\Big(\sum_{i=1}^n b_{i1}\alpha_i,\cdots,\sum_{i=1}^n b_{in}\alpha_i\Big)=\Big(\sum_{i=1}^n b_{i1}\underline{A}\alpha_i,\cdots,\sum_{i=1}^n b_{in}\underline{A}\alpha_i\Big)$$

$$=(\underline{A}\alpha_1,\cdots,\underline{A}\alpha_n)B=((\alpha_1,\cdots,\alpha_n)A)B=(\alpha_1,\cdots,\alpha_n)(AB)$$

所以 $\underline{A}\underline{B}$ 在基 $\alpha_1,\cdots,\alpha_n$ 下的矩阵是 $AB$. 于是

$$\sigma(\underline{A}\underline{B})=AB=\sigma(\underline{A})\sigma(\underline{B})$$

这证明了 $\sigma$ 保持乘法，从而 $\sigma$ 也是环 $\mathrm{Hom}(V,V)$ 到 $M_n(F)$ 的同构映射. ▌

注：从上述证明过程中看出，我们有

$$\underline{A}((\alpha_1,\cdots,\alpha_n)B)=(\underline{A}(\alpha_1,\cdots,\alpha_n))B \tag{11}$$

由于 $\sigma$ 也是环 $\mathrm{Hom}(V,V)$ 到 $M_n(F)$ 的同构映射，所以 $\mathrm{Hom}(V,V)$ 与 $M_n(F)$ 的对应元素之间有关乘法运算的性质是一样的. 从而 $V$ 上的恒等变换 $\underline{I}$ 对应于 $n$ 级单位矩阵 $I$(即恒等变换 $\underline{I}$ 的矩阵是单位矩阵 $I$)；$V$ 上的线性变换 $\underline{A}$ 可逆当且仅当 $A$ 在基 $\alpha_1,\cdots,\alpha_n$ 下的矩阵 $A$ 可逆，并且 $\underline{A}^{-1}$ 对应于 $A^{-1}$，即 $\underline{A}^{-1}$ 在基 $\alpha_1,\cdots,\alpha_n$ 下的矩阵是 $A^{-1}$.

* **定义 1** 设 $M$ 和 $M'$ 都是域 $F$ 上的代数，如果存在 $M$ 到 $M'$ 的一个双射 $\sigma$，使得 $\sigma$ 既是线性空间 $M$ 到 $M'$ 的同构映射，又是环 $M$ 到 $M'$ 的同构映射，那么称代数 $M$ 与 $M'$ 是同构的，并且称 $\sigma$ 是代数 $M$ 到 $M'$ 的一个同构映射.

* $V$ 上每一个线性变换与它在 $V$ 的一个基下的矩阵的对应 $\sigma$ 就是代数 $\mathrm{Hom}(V,V)$ 到 $M_n(F)$ 的一个同构映射. 从而代数 $\mathrm{Hom}(V,V)$ 与 $M_n(F)$ 同构.

**命题** 9.4.2　设 $V$ 是域 $F$ 上 $n$ 维线性空间，$\underline{A}$ 是 $V$ 上的一个线性变换. 如果 $\underline{A}$ 在 $V$ 的基 $\alpha_1,\cdots,\alpha_n$ 下的矩阵为 $A$，$\alpha\in V$ 在基 $\alpha_1,\cdots,\alpha_n$ 下的坐标为 $X$，则 $\underline{A}\alpha$ 在基 $\alpha_1,\cdots,\alpha_n$ 下的坐标为 $AX$.

**证明**　从命题 9.4.1 立即得到. ▌

**例** 1　设 $\underline{A}$ 是域 $F$ 上 $n$ 维线性空间 $V$ 的幂等变换，证明：在 $V$ 中存在一个基，使得 $\underline{A}$ 在这个基下的矩阵 $A$ 形如

$$A=\begin{pmatrix} I_r & 0\\ 0 & 0\end{pmatrix}\tag{12}$$

其中 $r=\operatorname{rank}\underline{A}$.

**证明**　因为 $\underline{A}$ 是幂等变换，据命题 9.3.4 得

$$V=\operatorname{Ker}\underline{A}\oplus\operatorname{Im}\underline{A}$$

并且 $\underline{A}$ 是平行于 $\operatorname{Ker}\underline{A}$ 在 $\operatorname{Im}\underline{A}$ 上的投影.

在 $\operatorname{Im}\underline{A}$ 中取一个基 $\alpha_1,\cdots,\alpha_r$；在 $\operatorname{Ker}\underline{A}$ 中取一个基 $\alpha_{r+1},\cdots,\alpha_n$. 则 $\alpha_1,\cdots,\alpha_r,\alpha_{r-1},\cdots,\alpha_n$ 是 $V$ 的一个基. 因为

$$\underline{A}\alpha_i=\alpha_i,\qquad i=1,\cdots,r$$

$$\underline{A}\alpha_j=0,\qquad j=r+1,\cdots,n$$

所以
$$\underline{A}(\alpha_1,\cdots,\alpha_n)=(\alpha_1,\cdots,\alpha_n)\begin{pmatrix} I_r & 0\\ 0 & 0\end{pmatrix}$$

因此 $\underline{A}$ 在基 $\alpha_1,\cdots,\alpha_n$ 下的矩阵 $A=\begin{pmatrix} I_r & 0\\ 0 & 0\end{pmatrix}$，其中

$$r=\dim\operatorname{Im}\underline{A}=\operatorname{rank}\underline{A}$$ ▌

## 习　题　9.4

1. 求下列线性变换在所指定基下的矩阵：

(1) 在 $K^3$ 中，

$$\underset{\sim}{A}(x_1,x_2,x_3)=(x_1+2x_2,x_3-x_2,x_2-x_3)$$

求 $\underset{\sim}{A}$ 在基 $\varepsilon_1=(1,0,0),\varepsilon_2=(0,1,0),\varepsilon_3=(0,0,1)$ 下的矩阵；

(2) $[0;\varepsilon_1,\varepsilon_2,\varepsilon_3]$ 是空间中一个右手直角坐标系，$\underset{\sim}{A}$ 是对 $xoz$ 面的垂直投影，$\underset{\sim}{B}$ 是对 $z$ 轴的垂直投影，求 $\underset{\sim}{A},\underset{\sim}{B},\underset{\sim}{A}\underset{\sim}{B}$ 在基 $\varepsilon_1,\varepsilon_2,\varepsilon_3$ 下的矩阵；

(3) 在线性空间 $R[x]_n$ 中，用 $\underset{\sim}{D}$ 表示求导数. 求 $\underset{\sim}{D}$ 在基

$$f_0=1,\quad f_i=\frac{(x-a)^i}{i!}\quad(i=1,2,\cdots,n-1)$$

下的矩阵，其中 $a\in R$；

(4) 在 $R^R$ 中，由下述六个函数

$$f_1=\mathrm{e}^{ax}\cos bx,\qquad f_2=\mathrm{e}^{ax}\sin bx$$

$$f_3=x\mathrm{e}^{ax}\cos bx,\qquad f_4=x\mathrm{e}^{ax}\sin bx$$

$$f_5=\frac{1}{2}x^2\mathrm{e}^{ax}\cos bx,\qquad f_6=\frac{1}{2}x^2\mathrm{e}^{ax}\sin bx$$

生成的 6 维线性子空间记作 $V$. 说明求微商 $\underset{\sim}{D}$ 是 $V$ 上的一个线性变换，并且求 $\underset{\sim}{D}$ 在基 $f_1,f_2,\cdots,f_6$ 下的矩阵；

2. 在 $M_2(F)$ 中定义下述变换：

$$\underset{\sim}{A}_1(X)=\begin{pmatrix}a&b\\c&d\end{pmatrix}X,\qquad \underset{\sim}{A}_2(X)=X\begin{pmatrix}a&b\\c&d\end{pmatrix}$$

$$\underset{\sim}{A}_3(X)=\begin{pmatrix}a&b\\c&d\end{pmatrix}X\begin{pmatrix}a&b\\c&d\end{pmatrix}$$

其中 $\begin{pmatrix}a&b\\c&d\end{pmatrix}$ 是给定的一个 2 级矩阵. 说明 $\underset{\sim}{A}_1,\underset{\sim}{A}_2,\underset{\sim}{A}_3$ 都是 $M_2(F)$ 上的线性变换，并且求 $\underset{\sim}{A}_1,\underset{\sim}{A}_2,\underset{\sim}{A}_3$ 在基 $E_{11},E_{12},E_{21},E_{22}$ 下的矩阵.

3. 在 $n$ 维线性空间 $V$ 中，设有线性变换 $\underset{\sim}{A}$ 与向量 $\alpha$，使得 $\underset{\sim}{A}^{n-1}\alpha\neq0$，但是 $\underset{\sim}{A}^n\alpha=0$. 证明：$V$ 中存在一个基，使得 $\underset{\sim}{A}$ 在这个基下的矩阵是

$$\begin{pmatrix}0&1&0&\cdots&0&0\\0&0&1&\cdots&0&0\\\cdots&\cdots&\cdots&\cdots&\cdots&\cdots\\0&0&0&\cdots&0&1\\0&0&0&\cdots&0&0\end{pmatrix}$$

4. 设 $V$ 是域 $F$ 上 $n$ 维线性空间，证明：$V$ 的与全体线性变换可以交换的线性变换是数乘变换.

5. 在 $K[x]_n$ 中($n > 1$),用 $\underline{D}$ 表示求导数,证明:$\underline{D}$ 在任何一个基下的矩阵都不可能是对角矩阵.

6. 设 $V$ 是域 $F$ 上 $n$ 维线性空间,证明

$$\dim \mathrm{Hom}(V,V) = n^2$$

7. 设 $\underline{A}$ 是域 $F$ 上 $n$ 维线性空间 $V$ 上的一个线性变换. 证明:

(1) 在 $F[x]$ 中存在一个次数 $\leqslant n^2$ 的非零多项式 $f(x)$ 使得 $f(\underline{A}) = \underline{0}$;

(2) 如果 $f(\underline{A}) = 0, g(\underline{A}) = 0$,则 $d(\underline{A}) = 0$,这里 $d(x)$ 是 $f(x)$ 与 $g(x)$ 的一个最大公因式;

(3) $\underline{A}$ 可逆的充分必要条件是:有一个常数项不为零的多项式 $f(x)$ 使得 $f(\underline{A}) = 0$.

8. 设 $V$ 是域 $F$ 上的 $n$ 维线性空间,$\alpha_1,\cdots,\alpha_n$ 是 $V$ 的一个基. 设 $V$ 上的一个线性变换 $\underline{A}$ 在基 $\alpha_1,\cdots,\alpha_n$ 下的矩阵是 $A$,证明:

$$\mathrm{rank}\ \underline{A} = \mathrm{rank}\ A$$

* 9. 设 $V$ 和 $V'$ 都是域 $F$ 上有限维线性空间,$\alpha_1,\cdots,\alpha_n$ 是 $V$ 的一个基;$\eta_1,\cdots,\eta_s$ 是 $V'$ 的一个基. $V$ 到 $V'$ 的一个线性映射 $\underline{A}$ 在基 $\{\alpha_j\}$ 和 $\{\eta_i\}$ 下的矩阵为 $A$. 证明

$$\mathrm{rank}\underline{A} = \mathrm{rank}A$$

* 10. 设 $V$ 和 $V'$ 分别是域 $F$ 上 $n$ 维、$s$ 维线性空间,证明:$V$ 到 $V'$ 的每一个秩为 $r$ 的线性映射 $\underline{A}$ 能表示成 $r$ 个秩为 1 的线性映射的和.

* 11. 设 $A,B,C$ 分别是域 $F$ 上 $s \times n, m \times s, r \times m$ 矩阵,证明:

$$\mathrm{rank}(CBA) \geqslant \mathrm{rank}(CB) + \mathrm{rank}(BA) - \mathrm{rank}B$$

* 12. 设 $V$ 和 $V'$ 都是域 $F$ 上有限维线性空间,$\underline{A}$ 是 $V$ 到 $V'$ 的一个线性映射. 证明:存在 $V$ 的一个基和 $V'$ 的一个基,使得 $\underline{A}$ 在这一对基下的矩阵 $A$ 形如

$$\begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}$$

(提示:利用习题 9.3 的第 9 题).

## §5 线性变换在不同基下的矩阵的关系·特征值与特征向量·可对角化的线性变换

这一节我们来讨论 $V$ 上的一个线性变换 $\underline{A}$ 在 $V$ 的不同基下

的矩阵之间的关系.

**定理** 9.5.1 设 $V$ 是域 $F$ 上 $n$ 维线性空间, $V$ 上的一个线性变换 $\underset{=}{A}$ 在 $V$ 的两个基 $\alpha_1,\cdots,\alpha_n$ 与 $\eta_1,\cdots,\eta_n$ 下的矩阵分别为 $A$ 与 $B$, 从基 $\{\alpha_i\}$ 到基 $\{\eta_i\}$ 的过渡矩阵是 $S$, 则

$$B = S^{-1}AS \tag{1}$$

**证明** 由已知条件我们有

$$\underset{=}{A}(\alpha_1,\cdots,\alpha_n) = (\alpha_1,\cdots,\alpha_n)A \tag{2}$$

$$\underset{=}{A}(\eta_1,\cdots,\eta_n) = (\eta_1,\cdots,\eta_n)B \tag{3}$$

$$(\eta_1,\cdots,\eta_n) = (\alpha_1,\cdots,\alpha_n)S \tag{4}$$

于是

$$\begin{aligned}&\underset{=}{A}(\eta_1,\cdots,\eta_n)\\ &=\underset{=}{A}((\alpha_1,\cdots,\alpha_n)S) \overset{*}{=} (\underset{=}{A}(\alpha_1,\cdots,\alpha_n))S\\ &=((\alpha_1,\cdots,\alpha_n)A)S = (\alpha_1,\cdots,\alpha_n)(AS)\\ &=((\eta_1,\cdots,\eta_n)S^{-1})(AS) = (\eta_1,\cdots,\eta_n)(S^{-1}AS)\end{aligned} \tag{5}$$

比较(3)和(5)式得

$$B = S^{-1}AS$$ ▍

(5)式推导过程中加"$*$"的一步的理由请看上一节(11)式.

定理 9.5.1 表明, 同一个线性变换 $\underset{=}{A}$ 在 $V$ 的不同基下的矩阵是相似的, 这就是我们为什么在第五章花相当多的篇幅讨论 $n$ 级矩阵的相似关系的原因.

**定理** 9.5.2 域 $F$ 上 $n$ 维线性空间 $V$ 的同一个线性变换 $\underset{=}{A}$ 在 $V$ 的所有各个基下的矩阵组成的集合恰好是 $M_n(F)$ 的一个相似等价类.

**证明** 设 $\underset{=}{A}$ 在 $V$ 的一个基 $\alpha_1,\cdots,\alpha_n$ 下的矩阵为 $A$, 用 $\bar{A}$ 表示 $M_n(F)$ 中由 $A$ 确定的相似等价类.

任取 $V$ 的一个基 $\eta_1,\cdots,\eta_n$, 设 $\underset{=}{A}$ 在此基下的矩阵是 $B$. 据定理 9.5.1, $B \sim A$, 从而 $B \in \bar{A}$.

反之, 任取 $C \in \bar{A}$, 则有可逆矩阵 $U$, 使得 $C = U^{-1}AU$. 令

$$(\gamma_1,\cdots,\gamma_n) = (\alpha_1,\cdots,\alpha_n)U \tag{6}$$

则据命题 8.4.2 知，$\gamma_1,\cdots,\gamma_n$ 是 $V$ 的一个基. 据定理 9.5.1，$\underset{-}{A}$ 在基 $\gamma_1,\cdots,\gamma_n$ 下的矩阵为 $U^{-1}AU$，即 $C$ 是 $\underset{-}{A}$ 在基 $\{\gamma_i\}$ 下的矩阵. ▌

从定理 9.5.2 得出，同一个线性变换 $\underset{-}{A}$ 在 $V$ 的所有各个基下的矩阵组成的集合是 $M_n(F)$ 的一个相似等价类. 于是 $M_n(F)$ 在相似关系下的不变量就反映了线性变换的内在性质，它们与基的选取无关. 譬如 $n$ 级矩阵的行列式、秩、迹、特征多项式、特征值等都是 $M_n(F)$ 的相似不变量，因此我们可以把线性变换 $\underset{-}{A}$ 在某一个基下的矩阵 $A$ 的行列式、秩、迹、特征多项式、特征值分别称为线性变换 $\underset{-}{A}$ 的行列式、秩、迹、特征多项式、特征值. 这里说明一下：在 §3 我们曾经把 $\underset{-}{A}$ 的象 $\operatorname{Im}\underset{-}{A}$ 的维数称为 $\underset{-}{A}$ 的秩，现在又把 $\underset{-}{A}$ 在某一个基下的矩阵 $A$ 的秩称为 $\underset{-}{A}$ 的秩，这会不会产生矛盾？不会，因为习题 9.4 的第 8 题证明了 $\dim\operatorname{Im}\underset{-}{A}=\operatorname{rank}A$. 用象 $\operatorname{Im}\underset{-}{A}$ 的维数作为 $\underset{-}{A}$ 的秩的定义更具有几何直观，并且它反映了线性变换的内在性质.

刚才我们把线性变换 $\underset{-}{A}$ 在某一个基下的矩阵 $A$ 的特征值称为线性变换 $\underset{-}{A}$ 的特征值. 为了更好理解线性变换的特征值的几何意义，以及对无限维线性空间的线性变换也考虑它的特征值，我们现在给出线性变换的特征值另一个定义，然后我们说明这两个定义在有限维线性空间的情形是一致的.

**定义 1** 设 $\underset{-}{A}$ 是域 $F$ 上线性空间 $V$ 的一个线性变换，如果对于域 $F$ 中一个元素 $\lambda_0$，$V$ 中存在一个非零向量 $\xi$，使得

$$\underset{-}{A}\xi=\lambda_0\xi \tag{7}$$

那么 $\lambda_0$ 称为 $\underset{-}{A}$ 的一个**特征值**，而 $\xi$ 称为 $\underset{-}{A}$ 的属于特征值 $\lambda_0$ 的一个**特征向量**.

从定义 1 看出，线性变换 $\underset{-}{A}$ 的特征向量 $\xi$ 是具有这样性质的非零向量：它在 $\underset{-}{A}$ 下的象 $\underset{-}{A}\xi$ 是它自己的 $\lambda_0$ 倍，而 $\lambda_0$ 就是 $\underset{-}{A}$ 的一个特征值.

现在设 $V$ 是域 $F$ 上的 $n$ 维线性空间，$V$ 中取定一个基 $\alpha_1,\cdots,\alpha_n$. 设线性变换 $\underset{-}{A}$ 在基 $\{\alpha_i\}$ 下的矩阵是 $A$，向量 $\xi$ 在基 $\{\alpha_i\}$ 下的坐标

是 $X,\lambda_0 \in F$. 从 §4 的(8)式得出

$$\underline{A}\xi = \lambda_0\xi \Leftrightarrow AX = \lambda_0 X \tag{8}$$

由此得出:

$$\begin{aligned} &\lambda_0 \text{ 是 } \underline{A} \text{ 的特征值} \\ \Leftrightarrow &\lambda_0 \text{ 是 } A \text{ 的特征值} \end{aligned} \tag{9}$$

$$\begin{aligned} &\xi \text{ 是 } \underline{A} \text{ 的属于 } \lambda_0 \text{ 的特征向量} \\ \Leftrightarrow &X \text{ 是 } A \text{ 的属于 } \lambda_0 \text{ 的特征向量} \end{aligned} \tag{10}$$

因此，只要把矩阵 $A$ 的全部特征值求出来了，它们就是线性变换 $\underline{A}$ 的全部特征值. 只要把矩阵 $A$ 的属于 $\lambda_0$ 的全部特征向量($\in F^n$)求出来了，分别以它们为坐标的向量($\in V$)就是 $\underline{A}$ 的属于 $\lambda_0$ 的全部特征向量. 由于 $V$ 的每个向量与它在基$\{\alpha_i\}$下的坐标的对应是$V$到$F^n$的同构映射，因此只要把矩阵$A$的属于$\lambda_0$的极大线性无关特征向量组求出来了，分别以它们为坐标的向量就是 $\underline{A}$ 的属于 $\lambda_0$ 的极大线性无关特征向量组.

第五章 §6 中我们曾指出，矩阵 $A$ 的属于特征值 $\lambda_0$ 的全部特征向量添上零向量所成的集合是 $F^n$(那时是 $K^n$) 的一个线性子空间，称它是 $A$ 的属于 $\lambda_0$ 的特征子空间. 根据同构映射的性质 7(见第八章 §7) 得到，线性变换 $\underline{A}$ 的属于特征值 $\lambda_0$ 的全部特征向量添上零向量组成的集合是 $V$ 的一个线性子空间，称它是 $\underline{A}$ 的属于 $\lambda_0$ 的**特征子空间**，记作$V_{\lambda_0}$，并且$V_{\lambda_0}$的维数等于矩阵$A$的属于$\lambda_0$的特征子空间的维数，而后者就是齐次线性方程组$(\lambda_0 I - A)X = 0$ 的解空间的维数.

对于域 $F$ 上任意一个线性空间 $V$(有限维或无限维)，可以直接验证线性变换 $\underline{A}$ 的属于特征值 $\lambda_0$ 的全部特征向量添上零向量组成的集合是$V$的一个线性子空间，称它为$\underline{A}$的一个**特征子空间**，记作 $V_{\lambda_0}$. 用集合记号可写成

$$V_{\lambda_0} = \{\alpha \mid \underline{A}\alpha = \lambda_0\alpha, \alpha \in V\} \tag{11}$$

验证如下: 显然$V_{\lambda_0}$非空(因为$0 \in V_{\lambda_0}$). 设$\alpha,\beta \in V_{\lambda_0}$，则$\underline{A}\alpha = \lambda_0\alpha$,

$\underset{\sim}{A}\beta = \lambda_0\beta$. 从而

$$\underset{\sim}{A}(\alpha+\beta) = \underset{\sim}{A}\alpha + \underset{\sim}{A}\beta = \lambda_0\alpha + \lambda_0\beta = \lambda_0(\alpha+\beta)$$

$$\underset{\sim}{A}(k\alpha) = k\underset{\sim}{A}\alpha = k(\lambda_0\alpha) = \lambda_0(k\alpha), \quad k \in F$$

因此，$\alpha+\beta, k\alpha \in V_{\lambda_0}$. 所以 $V_{\lambda_0}$ 是 $V$ 的一个子空间. ▌

**例1** 设 $V$ 是数域 $K$ 上的3维线性空间，$\underset{\sim}{A}$ 是 $V$ 上的一个线性变换，$\underset{\sim}{A}$ 在 $V$ 的一个基 $\alpha_1,\alpha_2,\alpha_3$ 下的矩阵是

$$A = \begin{pmatrix} 2 & -2 & 2 \\ -2 & -1 & 4 \\ 2 & 4 & -1 \end{pmatrix}$$

求 $\underset{\sim}{A}$ 的特征值和特征向量.

**解** $A$ 的特征多项式为

$$|\lambda I - A| = \begin{vmatrix} \lambda-2 & 2 & -2 \\ 2 & \lambda+1 & -4 \\ -2 & -4 & \lambda+1 \end{vmatrix}$$

$$= (\lambda-3)^2(\lambda+6)$$

所以 $\underset{\sim}{A}$ 的特征值是 3(二重) 与 $-6$.

对于特征值3，解齐次线性方程组 $(3I-A)X=0$，得到一个基础解系：

$$\begin{pmatrix} -2 \\ 1 \\ 0 \end{pmatrix}, \qquad \begin{pmatrix} 2 \\ 0 \\ 1 \end{pmatrix}$$

从而 $\underset{\sim}{A}$ 的属于3的极大线性无关特征向量组是

$$\xi_1 = -2\alpha_1 + \alpha_2, \qquad \xi_2 = 2\alpha_1 + \alpha_3$$

于是 $\underset{\sim}{A}$ 的属于3的全部特征向量是 $k_1\xi_1 + k_2\xi_2$，$k_1,k_2$ 取遍数域 $K$ 中不全为零的全部数对.

对于特征值 $-6$，解 $(-6I-A)X=0$ 得到一个基础解系：

$$\begin{pmatrix} 1 \\ 2 \\ -2 \end{pmatrix}$$

从而 $\underline{A}$ 的属于 $-6$ 的极大线性无关特征向量组是

$$\xi_3 = \alpha_1 + 2\alpha_2 - 2\alpha_3$$

于是 $\underline{A}$ 的属于 $-6$ 的全部特征向量是 $k\xi_3$，$k$ 是 $K$ 中任意非零数.

线性变换理论要研究的一个重要问题是：对于 $n$ 维线性空间 $V$ 上的一个线性变换 $\underline{A}$，要在 $V$ 中找一个适当的基，使 $\underline{A}$ 在这个基下的矩阵具有较简单的形式. 由于 $\underline{A}$ 在 $V$ 的所有各个基下的矩阵组成的集合是一个相似等价类，因此上述问题也就是在 $M_n(F)$ 的一个相似等价类里找一个矩阵具有较简单的形式. 这个问题我们在第五章 §5 至 §9 开始进行了讨论，并且对于其中的一种情形，即相似等价类里能找到对角矩阵的情形，作了详细的讨论，给出了这种情形的一些充分必要条件. 现在我们用线性变换的语言来叙述这些结果.

**定义** 2　域 $F$ 上 $n$ 维线性空间 $V$ 上的一个线性变换 $\underline{A}$ 称为**可对角化**的，如果 $V$ 中存在一个基，使得 $\underline{A}$ 在这个基下的矩阵为对角矩阵.

我们在 $V$ 中先任意取定一个基 $\alpha_1, \cdots, \alpha_n$，设线性变换 $\underline{A}$ 在这个基下的矩阵是 $A$，由定义 2 和上一段的议论知道

$$\underline{A} \text{ 可对角化} \Leftrightarrow A \text{ 可对角化} \tag{12}$$

于是，从第五章 §7 给出的 $n$ 级矩阵 $A$ 可对角化的一些充分必要条件容易得出，线性变换 $\underline{A}$ 可对角化的充分必要条件.

**定理** 9.5.3　域 $F$ 上 $n$ 维线性空间 $V$ 上的一个线性变换 $\underline{A}$ 可对角化的充分必要条件是，$\underline{A}$ 有 $n$ 个线性无关的特征向量，也就是，$V$ 中存在由 $\underline{A}$ 的特征向量组成的一个基.

**证法一**　在 $V$ 中取一个基 $\alpha_1, \cdots, \alpha_n$，设 $\underline{A}$ 在这个基下的矩阵是 $A$. 据第五章 §7 定理 5.7.1，$n$ 级矩阵 $A$ 可对角化的充分必要条件是 $A$ 有 $n$ 个线性无关的特征向量. 据(10)式以及线性空间的同构映射的性质 4，立即得到结论.

**证法二**

$\underset{\sim}{A}$ 可对角化

$\Leftrightarrow V$ 中有一个基 $\xi_1,\cdots,\xi_n$,使得

$$\underset{\sim}{A}(\xi_1,\xi_2,\cdots,\xi_n)=(\xi_1,\xi_2,\cdots,\xi_n)\begin{pmatrix}\lambda_1 & 0 & 0 & \cdots & 0\\ 0 & \lambda_2 & 0 & \cdots & 0\\ \cdots & \cdots & \cdots & \cdots & \cdots\\ 0 & 0 & 0 & \cdots & \lambda_n\end{pmatrix}$$

$\Leftrightarrow \underset{\sim}{A}\xi_i=\lambda_i\xi_i, i=1,\cdots,n$,并且 $\xi_1,\cdots,\xi_n$ 线性无关

$\Leftrightarrow \xi_1,\cdots,\xi_n$ 是 $\underset{\sim}{A}$ 的特征向量并且线性无关 ∎

从证法二看出,如果线性变换 $\underset{\sim}{A}$ 在一个基下的矩阵是对角矩阵,那么主对角线上的元素恰好是 $\underset{\sim}{A}$ 的全部特征值(重根按重数计算).因此除了主对角线上元素的排列次序外,这个对角矩阵是由线性变换 $\underset{\sim}{A}$ 唯一决定的.

如何知道一个线性变换 $\underset{\sim}{A}$ 有没有 $n$ 个线性无关的特征向量呢?在第五章 §7 我们对矩阵 $A$ 讨论了这个问题,证明了:$n$ 级矩阵 $A$ 的属于不同特征值的特征向量是线性无关的;$A$ 的分别属于不同特征值 $\lambda_1,\cdots,\lambda_m$ 的线性无关的特征向量组合在一起仍然线性无关.由此得到(据(10)式以及线性空间的同构映射的性质 4)

**定理** 9.5.4　$n$ 维线性空间 $V$ 上线性变换 $\underset{\sim}{A}$ 的属于不同特征值的特征向量是线性无关的;$\underset{\sim}{A}$ 的属于不同特征值 $\lambda_1,\cdots,\lambda_m$ 的线性无关的特征向量组合在一起仍线性无关. ∎

由定理 9.5.3 和定理 9.5.4 立即得到

**推论** 9.5.1　$n$ 维线性空间 $V$ 上的线性变换 $\underset{\sim}{A}$ 如果有 $n$ 个不同的特征值,那么 $\underset{\sim}{A}$ 可对角化.

**推论** 9.5.2　复数域上 $n$ 维线性空间 $V$ 的线性变换 $\underset{\sim}{A}$ 的特征多项式如果没有重根,那么 $\underset{\sim}{A}$ 可对角化.

注意这两个推论是 $\underset{\sim}{A}$ 可对角化的充分条件,但不是必要条件.

从定理 9.5.3 和定理 9.5.4 还可得到

**定理** 9.5.5　设 $n$ 维线性空间 $V$ 上的线性变换 $\underset{\sim}{A}$ 的全部不同的特征值是 $\lambda_1,\cdots,\lambda_m$，则 $\underset{\sim}{A}$ 可对角化的充分必要条件是

$$V = V_{\lambda_1} \oplus V_{\lambda_2} \oplus \cdots \oplus V_{\lambda_m}$$

**证明**　充分性. 设 $V = V_{\lambda_1} \oplus \cdots \oplus V_{\lambda_m}$. 在 $V_{\lambda_1},\cdots,V_{\lambda_m}$ 中各取一个基，由于和 $\sum\limits_{i=1}^{m} V_{\lambda_i}$ 是直和，所以据定理 8.6.8 得，把它们合起来是 $V_{\lambda_1} \oplus \cdots \oplus V_{\lambda_m}$ 的一个基，即 $V$ 的一个基. 这表明 $V$ 中存在由 $\underset{\sim}{A}$ 的特征向量组成的一个基，从而 $\underset{\sim}{A}$ 可对角化.

必要性. 设 $\underset{\sim}{A}$ 可对角化. 在 $V_{\lambda_i}$ 中取一个基 $\alpha_{i1},\cdots,\alpha_{ir_i}$，$i = 1,\cdots,m$，它们合起来是 $\underset{\sim}{A}$ 的最大的线性无关的特征向量组. 因为 $\underset{\sim}{A}$ 可对角化，所以 $r_1 + \cdots + r_m = n$. 从而 $\alpha_{11},\cdots,\alpha_{mr_m}$ 是 $V$ 的一个基. 因此

$$V = V_{\lambda_1} + \cdots + V_{\lambda_m}$$

并且上式右端的和是直和(据定理 8.6.8). ∎

从定理 9.5.3，9.5.4 和定理 9.5.5 又可得到

**定理** 9.5.6　条件同定理 9.5.5，则 $\underset{\sim}{A}$ 可对角化的充分必要条件是

$$\sum_{i=1}^{m} \dim V_{\lambda_i} = \dim V$$

**证明**　必要性. 设 $\underset{\sim}{A}$ 可对角化，则 $V = V_{\lambda_1} \oplus \cdots \oplus V_{\lambda_m}$. 从而

$$\begin{aligned}\dim V &= \dim(V_{\lambda_1} \oplus \cdots \oplus V_{\lambda_m}) \\ &= \dim V_{\lambda_1} + \cdots + \dim V_{\lambda_m}\end{aligned}$$

充分性. 设

$$\sum_{i=1}^{m} \dim V_{\lambda_i} = \dim V = n$$

在 $V_{\lambda_1},\cdots,V_{\lambda_m}$ 中各取一个基，把它们合起来共有 $n$ 个向量. 据定理 9.5.4，这 $n$ 个向量仍线性无关. 从而它们是 $V$ 的一个基. 据定理 9.5.3，$\underset{\sim}{A}$ 可对角化. ∎

**例 2**　对于例 1 中的 3 维线性空间 $V$ 上的线性变换 $\underset{\sim}{A}$，它有 3

个线性无关的特征向量：

$$\xi_1 = -2\alpha_1 + \alpha_2, \quad \xi_2 = 2\alpha_1 + \alpha_3, \quad \xi_3 = \alpha_1 + 2\alpha_2 - 2\alpha_3$$

因此 $\underline{A}$ 可对角化，即 $\underline{A}$ 在基 $\xi_1, \xi_2, \xi_3$ 下的矩阵为

$$B = \mathrm{diag}\{3, 3, -6\}$$

由基 $\alpha_1, \alpha_2, \alpha_3$ 到 $\xi_1, \xi_2, \xi_3$ 的过渡矩阵是

$$S = \begin{pmatrix} -2 & 2 & 1 \\ 1 & 0 & 2 \\ 0 & 1 & -2 \end{pmatrix}$$

于是 $S^{-1}AS = B$.

**例** 3　从§4的例1知道，域 $F$ 上 $n$ 维线性空间 $V$ 的任一幂等变换一定可对角化，并且幂等变换的特征值只可能是1或0. 属于1的特征子空间是这个幂等变换的象；而属于0的特征子空间是它的核.

上面我们给出了线性变换可对角化的三个充分必要条件(即定理9.5.3，定理9.5.5，定理9.5.6). 以后我们还会继续给出线性变换可对角化的充分必要条件.

可对角化的线性变换形成最简单的并且在许多方面是最重要的线性变换类.

在第五章§7的例2中，我们看到，一个矩阵能否对角化与所考虑的域有关. 从而一个线性变换是不是可对角化与域有关，也就是说，线性变换的可对角化性质是与域的性质有关系的.

## 习　题　9.5

1. (1) 已知 $K^3$ 上线性变换 $\underline{A}$ 在基 $\varepsilon_1 = (1,0,0), \varepsilon_2 = (0,1,0), \varepsilon_3 = (0,0,1)$ 下的矩阵是

$$A = \begin{pmatrix} 15 & -11 & 5 \\ 20 & -15 & 8 \\ 8 & -7 & 6 \end{pmatrix}$$

求 $\underset{\sim}{A}$ 在基 $\eta_1=(2,3,1),\eta_2=(3,4,1),\eta_3=(1,2,2)$ 下的矩阵；

(2) 已知$K^3$上线性变换 $\underset{\sim}{A}$ 在基 $\alpha_1=(8,-6,7),\alpha_2=(-16,7,-13)$，$\alpha_3=(9,-3,7)$ 下的矩阵为

$$A=\begin{pmatrix}1 & -18 & 15\\ -1 & -22 & 20\\ 1 & -25 & 22\end{pmatrix}$$

求 $\underset{\sim}{A}$ 在基 $\eta_1=(1,-2,1),\eta_2=(3,-1,2),\eta_3=(2,1,2)$ 下的矩阵.

2. 设域 $F$ 上三维线性空间 $V$ 上的线性变换 $\underset{\sim}{A}$ 在基 $\alpha_1,\alpha_2,\alpha_3$ 下的矩阵为

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

(1) 求 $\underset{\sim}{A}$ 在基 $\alpha_2,\alpha_3,\alpha_1$ 下的矩阵；

(2) 求 $\underset{\sim}{A}$ 在基 $k\alpha_1,\alpha_2,\alpha_3$ 下的矩阵，其中 $k\in F$ 且 $k\neq 0$；

(3) 求 $\underset{\sim}{A}$ 在基 $\alpha_1,\alpha_1+\alpha_2,\alpha_3$ 下的矩阵.

*3. 设 $\underset{\sim}{A}$ 是域 $F$ 上 $n$ 维线性空间 $V$ 上的一个线性变换. 证明：如果 $\underset{\sim}{A}$ 在 $V$ 的所有各个基下的矩阵都相同，则 $\underset{\sim}{A}$ 是数乘变换.

4. 设 $\alpha_1,\alpha_2,\alpha_3,\alpha_4$ 是数域 $K$ 上 4 维线性空间 $V$ 的一个基，已知线性变换 $\underset{\sim}{A}$ 在这个基下的矩阵为

$$A=\begin{pmatrix}1 & 0 & 2 & 1\\ -1 & 2 & 1 & 3\\ 1 & 2 & 5 & 5\\ 2 & -2 & 1 & -2\end{pmatrix}$$

(1) 求 $\underset{\sim}{A}$ 在基 $\eta_1=\alpha_1-2\alpha_2+\alpha_4,\eta_2=3\alpha_2-\alpha_3-\alpha_4,\eta_3=\alpha_3+\alpha_4$，$\eta_4=2\alpha_4$ 下的矩阵；

(2) 求 $\underset{\sim}{A}$ 的核与值域；

(3) 在 $\mathrm{Ker}\,\underset{\sim}{A}$ 中选一个基，把它扩充成 $V$ 的一个基，并求 $\underset{\sim}{A}$ 在这个基下的矩阵；

(4) 在 $\mathrm{Im}\,\underset{\sim}{A}$ 中选一个基，把它扩充成 $V$ 的一个基，并且求 $\underset{\sim}{A}$ 在这个基下的矩阵.

5. 给定 $K^3$ 的两个基 $\alpha_1,\alpha_2,\alpha_3$ 与 $\eta_1,\eta_2,\eta_3$ 同第1(2)题，定义线性变换 $\underset{\sim}{B}$：

$$\underset{\sim}{B}\alpha_i=\eta_i,\qquad i=1,2,3$$

(1) 求出由基 $\alpha_1,\alpha_2,\alpha_3$ 到基 $\eta_1,\eta_2,\eta_3$ 的过渡矩阵；

(2) 求出 $\underline{B}$ 在基 $\alpha_1,\alpha_2,\alpha_3$ 下的矩阵；

(3) 求出 $\underline{B}$ 在基 $\eta_1,\eta_2,\eta_3$ 下的矩阵.

6. 求复数域上线性空间 $V$ 的线性变换 $\underline{A}$ 的全部特征值和特征向量，已知 $\underline{A}$ 在 $V$ 的一个基下的矩阵为

(1) $A=\begin{pmatrix} 2 & 2 & -2 \\ 2 & 5 & -4 \\ -2 & -4 & 5 \end{pmatrix}$；  (2) $A=\begin{pmatrix} 2 & 3 & 2 \\ 1 & 8 & 2 \\ -2 & -14 & -3 \end{pmatrix}$

7. 在第 6 题中，哪些线性变换是可对角化的？对于可对角化的线性变换，求 $V$ 的一个基，使 $\underline{A}$ 在这个基下的矩阵为对角矩阵，并且写出旧基到新基的过渡矩阵 $S$，以及对角矩阵 $S^{-1}AS$.

8. 在 $K[x]_n$ 中 $(n>1)$，写出求导数 $\underline{D}$ 的特征多项式.

9. 设 $\alpha_1,\alpha_2,\alpha_3,\alpha_4$ 是数域 $K$ 上 4 维线性空间 $V$ 的一个基. 线性变换 $\underline{A}$ 在这个基下的矩阵为

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

(1) 求 $\underline{A}$ 的特征值与特征向量；

(2) 求 $V$ 的一个基，使得 $\underline{A}$ 在这个基下的矩阵为对角矩阵，并且写出这个对角矩阵.

10. 设 $V$ 是域 $F$ 上任意一个线性空间（可以是无限维的），$\underline{A}$ 是 $V$ 上的一个线性变换. 证明：$\underline{A}$ 的属于不同特征值的特征向量是线性无关的.

11. 设 $V$ 是域 $F$ 上任意一个线性空间，$\lambda_1,\lambda_2$ 是线性变换 $\underline{A}$ 的两个不同的特征值，$\xi_1,\xi_2$ 分别是属于 $\lambda_1,\lambda_2$ 的特征向量. 证明：$\xi_1+\xi_2$ 不是 $\underline{A}$ 的特征向量.

12. 设 $V$ 是域 $F$ 上任一线性空间，证明：如果 $V$ 上的线性变换 $\underline{A}$ 以 $V$ 中每个非零向量作为它的特征向量，则 $\underline{A}$ 是数乘变换.

13. 设 $V$ 是域 $F$ 上任一线性空间，设 $\underline{A}$ 是 $V$ 上的可逆的线性变换.

(1) 证明 $\underline{A}$ 的特征值一定不为 0；

(2) 证明：如果 $\lambda$ 是 $\underline{A}$ 的特征值，则 $\dfrac{1}{\lambda}$ 是 $\underline{A}^{-1}$ 的特征值.

14. 设 $V$ 是域 $F$ 上的 $n(>1)$ 维线性空间，设线性变换 $\underline{A}$ 在 $V$ 的一个基下的矩阵为

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

证明：$\underset{=}{A}$ 不是可对角化的.

*15. 复数域上 $n$ 维线性空间 $V$ 的线性变换 $\underset{=}{A}$ 的所有特征值组成的 $n$ 元数组 $(\lambda_1,\lambda_2,\cdots,\lambda_n)$ 称为 $\underset{=}{A}$ 的**谱**(spectrum). 如果 $\underset{=}{A}$ 的所有特征值都是 1 重的，则 $\underset{=}{A}$ 的谱称为**单**的.

设 $\underset{=}{A}$ 的谱是单的.

(1) 证明：如果线性变换 $\underset{=}{B}$ 与 $\underset{=}{A}$ 可交换，即 $\underset{=}{B}\underset{=}{A} = \underset{=}{A}\underset{=}{B}$，则 $\underset{=}{B}$ 能表示成 $\underset{=}{A}$ 的一个次数小于 $n$ 的多项式；

（提示：易知 $\underset{=}{A}$ 可对角化，去 $\underset{=}{B}$ 证可对角化，然后用待定系数法.）

(2) 证明：与 $\underset{=}{A}$ 可交换的所有线性变换组成的集合是 $\mathrm{Hom}(V,V)$ 的一个子空间，记作 $C(\underset{=}{A})$，并且 $\dim C(\underset{=}{A}) = \dim V$.

（提示：利用 $\underset{=}{A}$ 可对角化，转化为矩阵的相应问题.）

## §6 线性变换的不变子空间

上一节我们讨论了可对角化的线性变换. 不可以对角化的线性变换，其结构又如何呢？解决这个问题的思路是什么？从定理 9.5.5 看到，$\underset{=}{A}$ 是可对角化的线性变换当且仅当空间 $V$ 能分解成 $\underset{=}{A}$ 的特征子空间的直和. 由此受到启发，研究不可以对角化的线性变换的结构，能不能从研究空间 $V$ 分解成与 $\underset{=}{A}$ 有关的特殊类型的子空间的直和入手?这节就来讨论这个问题. 注意 $\underset{=}{A}$ 的特征子空间 $V_{\lambda_i}$ 具有如下性质：若 $\alpha \in V_{\lambda_i}$，则 $\underset{=}{A}\alpha = \lambda_i\alpha_i \in V_{\lambda_i}$ 这启发我们引入 $\underset{=}{A}$ 的不变子空间的概念.

**定义 1** 设 $\underset{=}{A}$ 是域 $F$ 上线性空间 $V$ 的线性变换，$W$ 是 $V$ 的子空间. 如果 $W$ 中的向量在 $\underset{=}{A}$ 下的象仍在 $W$ 中，即对于任意 $\alpha\in W$，

都有 $\underset{\sim}{A}\alpha\in W$，则称 $W$ 是 $\underset{\sim}{A}$**的不变子空间**，简称 $\underset{\sim}{A}$**—子空间**.

显然，整个空间 $V$ 和零子空间 0，对于 $V$ 上的每个线性变换 $\underset{\sim}{A}$ 来说，都是 $\underset{\sim}{A}$—子空间. 称 $V$ 和 0 是 $\underset{\sim}{A}$ 的平凡的不变子空间.

**命题** 9.6.1　$V$ 上线性变换 $\underset{\sim}{A}$ 的核与象，$\underset{\sim}{A}$ 的特征子空间都是 $\underset{\sim}{A}$—子空间.

**证明**　任取 $\alpha\in\mathrm{Ker}\,\underset{\sim}{A}$，因为 $\underset{\sim}{A}\alpha=0\in\mathrm{Ker}\,\underset{\sim}{A}$，所以 $\mathrm{Ker}\,\underset{\sim}{A}$ 是 $\underset{\sim}{A}$—子空间.

任取 $\alpha\in\mathrm{Im}\,\underset{\sim}{A}$，因为 $\underset{\sim}{A}\alpha\in\mathrm{Im}\,\underset{\sim}{A}$，所以 $\mathrm{Im}\,\underset{\sim}{A}$ 是 $\underset{\sim}{A}$—子空间.

任取 $\alpha\in V_{\lambda_i}$，因为 $\underset{\sim}{A}\alpha=\lambda_i\alpha\in V_{\lambda_i}$，所以 $V_{\lambda_i}$ 是 $\underset{\sim}{A}$—子空间. ▌

**命题** 9.6.2　如果线性变换 $\underset{\sim}{A}$ 与 $\underset{\sim}{B}$ 是可交换的，则 $\underset{\sim}{A}$ 的核与象都是 $\underset{\sim}{B}$— 子空间. $\underset{\sim}{A}$ 的特征子空间也是 $\underset{\sim}{B}$— 子空间.

**证明**　任取 $\alpha\in\mathrm{Ker}\,\underset{\sim}{A}$，则 $\underset{\sim}{A}\alpha=0$. 于是

$$\begin{aligned}\underset{\sim}{A}(\underset{\sim}{B}\alpha)&=(\underset{\sim}{A}\underset{\sim}{B})\alpha=(\underset{\sim}{B}\underset{\sim}{A})\alpha\\&=\underset{\sim}{B}(\underset{\sim}{A}\alpha)=\underset{\sim}{B}(0)=0\end{aligned}$$

所以 $\underset{\sim}{B}\alpha\in\mathrm{Ker}\,\underset{\sim}{A}$. 因此 $\mathrm{Ker}\,\underset{\sim}{A}$ 是 $\underset{\sim}{B}$— 子空间.

任取 $\alpha\in\mathrm{Im}\,\underset{\sim}{A}$，则存在 $\beta\in V$，使得 $\underset{\sim}{A}\beta=\alpha$. 从而

$$\begin{aligned}\underset{\sim}{B}\alpha&=\underset{\sim}{B}(\underset{\sim}{A}\beta)=(\underset{\sim}{B}\underset{\sim}{A})\beta\\&=(\underset{\sim}{A}\underset{\sim}{B})\beta=\underset{\sim}{A}(\underset{\sim}{B}\beta)\in\mathrm{Im}\,\underset{\sim}{A}\end{aligned}$$

因此 $\mathrm{Im}\,\underset{\sim}{A}$ 是 $\underset{\sim}{B}$— 子空间.

在 $\underset{\sim}{A}$ 的特征子空间 $V_{\lambda_i}$ 中任取一个向量 $\alpha$，则 $\underset{\sim}{A}\alpha=\lambda_i\alpha$. 从而

$$\begin{aligned}\underset{\sim}{A}(\underset{\sim}{B}\alpha)&=(\underset{\sim}{A}\underset{\sim}{B})\alpha=(\underset{\sim}{B}\,\underset{\sim}{A})\alpha\\&=\underset{\sim}{B}(\underset{\sim}{A}\alpha)=\underset{\sim}{B}(\lambda_i\alpha)=\lambda_i(\underset{\sim}{B}\alpha)\end{aligned}$$

所以 $\underset{\sim}{B}\alpha\in V_{\lambda_i}$. 因此 $V_{\lambda_i}$ 是 $\underset{\sim}{B}$— 子空间. ▌

由于对任意 $f(x)\in F[x]$，$f(\underset{\sim}{A})$ 与 $\underset{\sim}{A}$ 可交换，从命题 9.6.2 立即得到

**推论** 9.6.1　对于任意 $f(x)\in F[x]$，线性变换 $f(\underset{\sim}{A})$ 的核与象都是 $\underset{\sim}{A}$— 子空间，$f(\underset{\sim}{A})$ 的特征子空间也是 $\underset{\sim}{A}$— 子空间. ▌

**命题** 9.6.3　线性变换 $\underset{\sim}{A}$ 的不变子空间的和与交仍是 $\underset{\sim}{A}$ 的不

变子空间.

**证明** 设 $V_1,\cdots,V_s$ 都是 $\underset{\sim}{A}$— 子空间. 在和 $\sum_{i=1}^{s} V_i$ 中任取一个向量 $\sum_{i=1}^{s}\alpha_i$，其中 $\alpha_i \in V_i, i = 1,\cdots,s$，则

$$\underset{\sim}{A}\left(\sum_{i=1}^{s}\alpha_i\right) = \sum_{i=1}^{s}\underset{\sim}{A}\alpha_i, \qquad \underset{\sim}{A}\alpha_i \in V_i, i = 1,\cdots,s$$

因此 $\underset{\sim}{A}\left(\sum_{i=1}^{s}\alpha_i\right) \in \sum_{i=1}^{s}V_i$. 于是 $\sum_{i=1}^{s}V_i$ 是 $\underset{\sim}{A}$— 子空间.

任取 $\alpha \in \bigcap_{i=1}^{s}V_i$，则 $\underset{\sim}{A}\alpha \in \bigcap_{i=1}^{s}V_i$. 因此 $\bigcap_{i=1}^{s}V_i$ 是 $\underset{\sim}{A}$— 子空间. ▌

**命题** 9.6.4 $V$ 的任何一个子空间都是数乘变换的不变子空间.

**证明** 设 $W$ 是 $V$ 的子空间，设 $k \in F$. 任取 $\alpha \in W$，则 $\underset{\sim}{k}\alpha = k\alpha \in W$. 因此 $W$ 是 $\underset{\sim}{k}$ 的不变子空间. ▌

线性变换 $\underset{\sim}{A}$ 的一维不变子空间与 $\underset{\sim}{A}$ 的特征向量有密切关系：

**命题** 9.6.5 设 $\underset{\sim}{A}$ 是 $V$ 上的线性变换. 如果 $W$ 是 $\underset{\sim}{A}$ 的一维不变子空间，则 $W$ 中任何一个非零向量都是 $\underset{\sim}{A}$ 的特征向量；反之，如果 $\xi$ 是 $\underset{\sim}{A}$ 的一个特征向量，则 $\xi$ 生成的子空间 $\langle\xi\rangle$ 是 $\underset{\sim}{A}$ 的一维不变子空间.

**证明** 设 $W$ 是 $\underset{\sim}{A}$ 的一维不变子空间. 任取 $\alpha \in W$ 且 $\alpha \neq 0$，则 $\alpha$ 是 $W$ 的一个基. 因为 $W$ 是 $\underset{\sim}{A}$— 子空间，所以 $\underset{\sim}{A}\alpha \in W$. 从而 $\underset{\sim}{A}\alpha = k\alpha$，对某个 $k \in F$. 这表明 $\alpha$ 是 $\underset{\sim}{A}$ 的特征向量.

反之，设 $\xi$ 是 $\underset{\sim}{A}$ 的一个特征向量，即 $\underset{\sim}{A}\xi = \lambda_0\xi$. 在 $\langle\xi\rangle$ 中任取一个向量 $b\xi$，我们有

$$\underset{\sim}{A}(b\xi) = b\,\underset{\sim}{A}\xi = b(\lambda_0\xi) = (b\lambda_0)\xi \in \langle\xi\rangle$$

所以 $\langle\xi\rangle$ 是 $\underset{\sim}{A}$ 的不变子空间，显然 $\langle\xi\rangle$ 是一维的. ▌

下面给出 $V$ 的有限维子空间 $W$ 是 $\underset{\sim}{A}$— 子空间的一个判别法则：

**命题** 9.6.6 设 $V$ 的子空间 $W = \langle\alpha_1,\cdots,\alpha_m\rangle$，则 $W$ 是线性变

换 $\underline{A}$ 的不变子空间当且仅当 $\underline{A}\alpha_i \in W, i = 1, \cdots, m.$

**证明** 必要性是显然的. 现在看充分性. 任取 $\alpha \in W$, 则

$$\alpha = \sum_{i=1}^{m} k_i \alpha_i$$

从而 $\underline{A}\alpha = \sum\limits_{i=1}^{m} k_i \underline{A}\alpha_i \in W$. 所以 $W$ 是 $\underline{A}$ 的不变子空间. ▌

如果 $V$ 的子空间 $W$ 是 $V$ 上线性变换 $\underline{A}$ 的不变子空间, 则 $W$ 中向量在 $\underline{A}$ 下的象仍在 $W$ 中. 于是从 $\underline{A}$ 可诱导出一个定义域为 $W$, 伴域也为 $W$ 的映射, 记作 $\underline{A}|W$, 称为 **$\underline{A}$ 在 $W$ 上的限制**, 其对应法则为

$$(\underline{A}|W)\alpha = \underline{A}\alpha, \qquad \forall\, \alpha \in W \tag{1}$$

显然, $\underline{A}|W$ 是 $W$ 上的线性变换.

例如, 设 $\underline{A}$ 是 $V$ 上的线性变换, $V_{\lambda_i}$ 是 $\underline{A}$ 的一个特征子空间, 则

$$\underline{A}|\mathrm{Ker}\underline{A} = \underline{0}; \qquad \underline{A}|V_{\lambda_i} = \lambda_i$$

注意, 如果子空间 $W$ 不是 $\underline{A}$ 的不变子空间, 那么由(1)式定义的 $\underline{A}$ 在 $W$ 上的限制 $\underline{A}|W$ 是 $W$ 到 $V$ 的一个线性映射, 而不是 $W$ 上的线性变换.

* 如果 $W$ 是 $V$ 上线性变换 $\underline{A}$ 的不变子空间, 那么 $\underline{A}$ 还可以诱导出商空间 $V/W$ 上的一个线性变换 $\overline{\underline{A}}$, 即我们规定

$$\begin{aligned} \overline{\underline{A}}: V/W &\longrightarrow V/W \\ \alpha + W &\longmapsto \underline{A}\alpha + W \end{aligned} \tag{2}$$

首先需要说明, (2) 式的确给出了 $V/W$ 到 $V/W$ 的一个映射 $\overline{\underline{A}}$, 即要证明: 线性子簇 $\alpha + W$ 在 $\overline{\underline{A}}$ 下的象不依赖于代表的选择. 假设 $\alpha + W = \beta + W$, 则 $\alpha - \beta \in W$. 由于 $W$ 是 $\underline{A}$— 子空间, 所以

$$\underline{A}\alpha - \underline{A}\beta = \underline{A}(\alpha - \beta) \in W$$

于是

$$\underline{A}\alpha + W = \underline{A}\beta + W$$

这证明了(2)式的定义是合理的. 其次我们来证明 $\overline{\underline{A}}$ 是线性的: 设

$\alpha + W, \gamma + W \in V/W, k \in F$,我们有

$$\bar{\underline{A}}[(\alpha + W) + (\gamma + W)]$$

$$= \bar{\underline{A}}[(\alpha + \gamma) + W]$$

$$= \underline{A}(\alpha + \gamma) + W = (\underline{A}\alpha + \underline{A}\gamma) + W$$

$$= (\underline{A}\alpha + W) + (\underline{A}\gamma + W) = \bar{\underline{A}}(\alpha + W) + \bar{\underline{A}}(\gamma + W)$$

$$\bar{\underline{A}}[k(\alpha + W)] = \bar{\underline{A}}(k\alpha + W)$$

$$= \underline{A}(k\alpha) + W = k\,\underline{A}\alpha + W$$

$$= k(\underline{A}\alpha + W) = k[\bar{\underline{A}}(\alpha + W)]$$

因此 $\bar{\underline{A}}$ 是 $V/W$ 上的线性变换.

* 总而言之,如果 $W$ 是 $V$ 上线性变换 $\underline{A}$ 的不变子空间,那么 $\underline{A}$ 既可以诱导出子空间 $W$ 上的线性变换 $\underline{A}|W$,又可以诱导出商空间 $V/W$ 上的线性变换 $\bar{\underline{A}}$,它们分别由(1)式和(2)式定义.

从现在起,我们假设 $V$ 是有限维的,$\dim V = n$. 我们需利用线性变换 $\underline{A}$ 的不变子空间来研究 $\underline{A}$ 的结构.

**定理** 9.6.1　设 $\underline{A}$ 是域 $F$ 上 $n$ 维线性空间 $V$ 的线性变换. 则 $\underline{A}$ 可对角化的充分必要条件是 $V$ 可以分解成 $\underline{A}$ 的一维不变子空间的直和.

**证明**　必要性. 设 $\underline{A}$ 可对角化,则 $V$ 中存在由 $\underline{A}$ 的特征向量组成的一个基 $\alpha_1, \cdots, \alpha_n$. 因此

$$V = \langle \alpha_1 \rangle \oplus \cdots \oplus \langle \alpha_n \rangle$$

据命题 9.6.5,$\langle \alpha_i \rangle$ 是 $\underline{A}$ 的一维不变子空间,$i = 1, \cdots, n$.

充分性. 设 $V$ 可以分解成 $\underline{A}$ 的一维不变子空间 $W_i (i = 1, \cdots, n)$ 的直和:$V = W_1 \oplus \cdots \oplus W_n$. 在 $W_i$ 中取一个基 $\xi_i$,据命题 9.6.5,$\xi_i$ 是 $\underline{A}$ 的特征向量,$i = 1, \cdots, n$. 由于和 $W_1 \oplus \cdots \oplus W_n$ 是直和,所以 $\xi_1, \cdots, \xi_n$ 是 $W_1 \oplus \cdots \oplus W_n$ 的一个基,即 $V$ 的一个基. 据定理 9.5.3,$\underline{A}$ 可对角化. ▌

**定理** 9.6.2　设 $\underset{=}{A}$ 是域 $F$ 上 $n$ 维线性空间 $V$ 的线性变换. 则 $\underset{=}{A}$ 在 $V$ 的一个基下的矩阵为形如

$$\begin{pmatrix} A_1 & A_3 \\ 0 & A_2 \end{pmatrix} \tag{3}$$

的分块上三角矩阵的充分必要条件是，$\underset{=}{A}$ 有非平凡的不变子空间.

**证明**　必要性. 设 $\underset{=}{A}$ 在 $V$ 的一个基 $\alpha_1,\cdots,\alpha_n$ 下的矩阵是形如 (3) 的分块上三角矩阵 $A=(a_{ij})$，其中 $A_1$ 是 $r$ 级方阵，$0<r<n$. 则对于 $1\leqslant j\leqslant r$，我们有

$$\underset{=}{A}\alpha_j = a_{1j}\alpha_1 + a_{2j}\alpha_2 + \cdots + a_{rj}\alpha_r \tag{4}$$

令 $W \doteq \langle \alpha_1,\alpha_2,\cdots,\alpha_r\rangle$. 从 (4) 式看出，$\underset{=}{A}\alpha_j\in W$，$j=1,\cdots,r$. 据命题 9.6.6，$W$ 是 $\underset{=}{A}$ 的不变子空间. 因为 $0<r<n$，所以 $W$ 是非平凡的.

充分性. 设 $W$ 是 $\underset{=}{A}$ 的非平凡不变子空间，设 $\dim W=r$，$0<r<n$. 在 $W$ 中取一个基 $\alpha_1,\cdots,\alpha_r$，把它扩充成 $V$ 的一个基 $\alpha_1,\cdots,\alpha_r,\alpha_{r+1},\cdots,\alpha_n$. 因为 $\underset{=}{A}\alpha_j\in W$，$j=1,\cdots,r$，所以

$$\begin{aligned}
&\underset{=}{A}(\alpha_1,\cdots,\alpha_n) \\
&= (\alpha_1,\cdots,\alpha_n)\begin{pmatrix}
a_{11} & \cdots & a_{1r} & a_{1,r+1} & \cdots & a_{1n} \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
a_{r1} & \cdots & a_{rr} & a_{r,r+1} & \cdots & a_{rn} \\
0 & \cdots & 0 & a_{r+1,r+1} & \cdots & a_{r+1,n} \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
0 & \cdots & 0 & a_{n,r+1} & \cdots & a_{nn}
\end{pmatrix}
\end{aligned} \tag{5}$$

即 $\underset{=}{A}$ 在基 $\alpha_1,\cdots,\alpha_r,\alpha_{r+1},\cdots,\alpha_n$ 下的矩阵是形如 (3) 的分块上三角矩阵.　▌

从定理 9.6.2 及其证明过程我们可以进一步得到

**推论** 9.6.1　设 $n$ 维线性空间 $V$ 的线性变换 $\underset{=}{A}$ 在 $V$ 的一个基 $\alpha_1,\cdots,\alpha_n$ 下的矩阵 $A$ 是形如 (3) 的分块上三角矩阵，其中 $A_1$ 是 $r$ 级

方阵，设 $W=\langle\alpha_1,\cdots,\alpha_r\rangle$，则 $A$ 的左上角的子矩阵 $A_1$ 是子空间 $W$ 上的线性变换 $\underline{A}|W$ 在基 $\alpha_1,\cdots,\alpha_r$ 下的矩阵；* 而 $A$ 的右下角的子矩阵 $A_2$ 是商空间 $V/W$ 上的线性变换 $\bar{\underline{A}}$ 在基 $\alpha_{r+1}+W,\cdots,\alpha_n+W$ 下的矩阵.

**证明** 据定理 9.6.2，$W$ 是 $\underline{A}$ 的非平凡不变子空间. 据(4)式，我们有

$$(\underline{A}|W)\alpha_j=\underline{A}\alpha_j=a_{1j}\alpha_1+a_{2j}\alpha_2+\cdots+a_{rj}\alpha_r,\quad j=1,\cdots,r$$

同此 $\underline{A}|W$ 在 $W$ 的基 $\alpha_1,\cdots,\alpha_r$ 下的矩阵是

$$\begin{pmatrix} a_{11} & \cdots & a_{1r}\\ a_{21} & \cdots & a_{2r}\\ \cdots & \cdots & \cdots\\ a_{r1} & \cdots & a_{rr}\end{pmatrix}=A_1$$

* 在定理 8.8.2 中，我们已证明了 $\alpha_{r+1}+W,\cdots,\alpha_n+W$ 是商空间 $V/W$ 的一个基. 对于 $j=r+1,\cdots,n$，我们有

$$\begin{aligned}\bar{\underline{A}}(\alpha_j+W)&=\underline{A}\alpha_j+W=(a_{1j}\alpha_1+\cdots+a_{nj}\alpha_n)+W\\&=a_{1j}(\alpha_1+W)+\cdots+a_{rj}(\alpha_r+W)\\&\quad+a_{r+1,j}(\alpha_{r+1}+W)+\cdots+a_{nj}(\alpha_n+W)\\&=a_{r+1,j}(\alpha_{r+1}+W)+\cdots+a_{nj}(\alpha_n+W)\end{aligned}\tag{6}$$

(6)式的最后一步是因为 $\alpha_1+W=W,\cdots,\alpha_r+W=W$，而 $W$ 是商空间 $V/W$ 里的零向量. 从(6)式得出

$$\begin{aligned}&\bar{\underline{A}}(\alpha_{r+1}+W,\cdots,\alpha_n+W)\\&=(\alpha_{r+1}+W,\cdots,\alpha_n+W)\begin{pmatrix}a_{r+1,r+1} & \cdots & a_{r+1,n}\\ \vdots & \vdots & \vdots\\ a_{n,r+1} & \cdots & a_{nn}\end{pmatrix}\end{aligned}\tag{7}$$

(7)式右端的矩阵正好是 $A$ 的右下角的子矩阵 $A_2$. ▌

**定理** 9.6.3 设 $\underline{A}$ 是域 $F$ 上 $n$ 维线性空间 $V$ 的线性变换，则 $\underline{A}$ 在 $V$ 的一个基下的矩阵为分块对角矩阵的充分必要条件是，$V$ 能分解成 $\underline{A}$ 的若干个非平凡不变子空间的直和.

**证明** 必要性. 设 $\underline{A}$ 在 $V$ 的一个基 $\alpha_{11},\cdots,\alpha_{1r_1},\cdots,\alpha_{s1},\cdots,\alpha_{sr_s}$ 下的矩阵 $A=\mathrm{diag}\{A_1,A_2,\cdots,A_s\}$，其中 $A_i$ 是 $r_i$ 级方阵，$i=1,\cdots,s$. 令 $W_1=\langle\alpha_{11},\cdots,\alpha_{1r_1}\rangle,\cdots,W_s=\langle\alpha_{s1},\cdots,\alpha_{sr_s}\rangle$. 由于

$$\underline{A}(\alpha_{j1},\cdots,\alpha_{jr_j})=(\alpha_{j1},\cdots,\alpha_{jr_j})A_j,\quad j=1,\cdots,s \tag{8}$$

因此 $\underline{A}\alpha_{j1},\cdots,\underline{A}\alpha_{jr_j}$ 都属于 $W_j$. 从而 $W_j$ 是 $\underline{A}$— 子空间. 由于 $W_j$ 的一个基，$j=1,\cdots,s$，合起来是 $V$ 的一个基，所以

$$V=W_1\oplus W_2\oplus\cdots\oplus Ws$$

充分性. 设 $V$ 是若干个 $\underline{A}$ 的非平凡不变子空间的直和：

$$V=W_1\oplus W_2\oplus\cdots\oplus Ws \tag{9}$$

在每个 $\underline{A}$— 子空间 $W_j$ 中取一个基 $\alpha_{j1},\cdots,\alpha_{jr_j}$，从(9) 式得出，

$$\alpha_{11},\cdots,\alpha_{1r_1},\cdots,\alpha_{s1},\cdots,\alpha_{sr_s}$$

是 $V$ 的一个基. 由于 $W_j$ 是 $\underline{A}$— 子空间，于是有

$$\underline{A}(\alpha_{j1},\cdots,\alpha_{jr_j})=(\alpha_{j1},\cdots,\alpha_{jr_j})A_j,\quad j=1,\cdots,s \tag{10}$$

所以

$$\underline{A}(\alpha_{11},\cdots,\alpha_{sr_s})=(\alpha_{11},\cdots,\alpha_{sr_s})\begin{pmatrix}A_1&&&\\&A_2&&\\&&\ddots&\\&&&A_s\end{pmatrix} \tag{11}$$

∎

从定理 9.6.3 的证明过程中看出，$A_j$ 就是 $\underline{A}$— 子空间 $W_j$ 上的线性变换 $\underline{A}|W_j$ 在基 $\alpha_{j1},\cdots,\alpha_{jr_j}$ 下的矩阵，$j=1,\cdots,s$.

作为定理 9.6.2 的应用，我们来证明下面一个结论.

* **命题** 9.6.7 设 $\underline{A}$ 是域 $F$ 上 $n$ 维线性空间 $V$ 的线性变换，$\lambda_0$ 是 $\underline{A}$ 的一个 $l$ 重特征值，则

$$\dim V_{\lambda_0}\leqslant l \tag{12}$$

我们把 $\dim V_{\lambda_0}$ 称为特征值 $\lambda_0$ 的**几何重数**，把 $\lambda_0$ 作为 $\underline{A}$ 的特征多项式的根的重数 $l$ 称为 $\lambda_0$ 的**代数重数**. (12) 式也就是，$\underline{A}$ 的特征值 $\lambda_0$ 的几何重数不超过它的代数重数.

**证法一** 设 $\underset{\sim}{A}$ 在 $V$ 的一个基下的矩阵是 $A$. 第五章 §8 的例 8 已证明:$A$ 的属于 $\lambda_0$ 的特征子空间的维数不超过 $\lambda_0$ 作为 $A$ 的特征多项式的根的重数. 由于 $V_{\lambda_0}$ 的维数等于 $A$ 的属于 $\lambda_0$ 的特征子空间的维数,因此(12)式成立.

**证法二** 因为 $\lambda_0$ 是 $\underset{\sim}{A}$ 的特征值,所以 $V_{\lambda_0}$ 是 $\underset{\sim}{A}$ 的不变子空间,且 $V_{\lambda_0}\neq 0$. 如果 $\underset{\sim}{A}$ 是数乘变换,则结论显然成立. 下设 $\underset{\sim}{A}$ 不是数乘变换,于是 $V_{\lambda_0}\neq V$. 设 $\dim V_{\lambda_0}=r$. 据定理 9.6.2,$\underset{\sim}{A}$ 在 $V$ 的一个基下的矩阵为

$$A=\begin{pmatrix} A_1 & A_3 \\ 0 & A_2 \end{pmatrix}$$

其中 $A_1$ 是 $V_{\lambda_0}$ 上的线性变换 $\underset{\sim}{A}|V_{\lambda_0}$ 在 $V_{\lambda_0}$ 的一个基下的矩阵,于是 $A_1$ 是 $r$ 级矩阵. 由于 $\underset{\sim}{A}|V_{\lambda_0}=\underset{\sim}{\lambda_0}$,所以 $A_1=\lambda_0 I_r$. 于是

$$\begin{aligned}|\lambda I-A| &= |\lambda I_r-\lambda_0 I_r|\cdot|\lambda I_{n-r}-A_2| \\ &= (\lambda-\lambda_0)^r|\lambda I_{n-r}-A_2|\end{aligned}$$

所以 $\lambda_0$ 作为 $\underset{\sim}{A}$ 的特征多项式 $|\lambda I-A|$ 的根的重数 $l\geqslant r$. ▌

## 习 题 9.6

1. 在 $K^4$ 中,设线性变换 $\underset{\sim}{A}$ 在基 $\alpha_1,\alpha_2,\alpha_3,\alpha_4$ 下的矩阵是

$$A=\begin{pmatrix} 1 & 0 & 2 & -1 \\ 0 & 1 & 4 & -2 \\ 2 & -1 & 0 & 1 \\ 2 & -1 & -1 & 2 \end{pmatrix}$$

令 $W=\langle\alpha_1+2\alpha_2,\alpha_2+\alpha_3+2\alpha_4\rangle$. 证明:$W$ 是 $\underset{\sim}{A}$—子空间.

2. 设 $W$ 是 $V$ 上可逆线性变换 $\underset{\sim}{A}$ 的有限维不变子空间,证明:

(1) $\underset{\sim}{A}|W$ 是 $W$ 上的可逆线性变换;

(2) $W$ 也是 $\underset{\sim}{A}^{-1}$ 的不变子空间,并且 $(\underset{\sim}{A}|W)^{-1}=\underset{\sim}{A}^{-1}|W$.

3. 设 $V$ 是复数域上的 $n$ 维线性空间,$\underset{\sim}{A},\underset{\sim}{B}$ 是 $V$ 上的线性变换,并且 $\underset{\sim}{A}\underset{\sim}{B}=\underset{\sim}{B}\underset{\sim}{A}$. 证明:$\underset{\sim}{A}$ 与 $\underset{\sim}{B}$ 至少有一个公共的特征向量.

4. 条件同第 3 题，并且设 $\underset{\sim}{A}$ 有 $s$ 个不同的特征值. 证明：$\underset{\sim}{A}$ 与 $\underset{\sim}{B}$ 至少有 $s$ 个公共的特征向量，并且它们线性无关.

5. 设 $V$ 是域 $F$ 上的 $n$ 维线性空间，线性变换 $\underset{\sim}{A}$ 在基 $\alpha_1,\alpha_2,\cdots,\alpha_n$ 下的矩阵为

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

(1) 证明：如果 $\alpha_n$ 属于 $\underset{\sim}{A}$ 的不变子空间 $W$，则 $W=V$；

(2) 证明：$\alpha_1$ 属于 $\underset{\sim}{A}$ 的任意一个非零不变子空间；

(3) 证明：$V$ 不能分解成 $\underset{\sim}{A}$ 的两个非平凡不变子空间的直和；

*(4) 求 $\underset{\sim}{A}$ 的所有不变子空间.

6. 设 $W$ 是 $V$ 上线性变换 $\underset{\sim}{A}$ 的不变子空间，证明：$W$ 在 $\underset{\sim}{A}$ 下的象 $\underset{\sim}{A}W$，以及 $W$ 在 $\underset{\sim}{A}$ 下的原象 $\underset{\sim}{A}^{-1}W$，都是 $\underset{\sim}{A}$ 的不变子空间，其中

$$\underset{\sim}{A}^{-1}W:=\{\alpha\in V\mid \underset{\sim}{A}\alpha\in W\}$$

*7. 在 $R[x]_n$ 中，求出求微商变换 $\underset{\sim}{D}$ 的所有不变子空间.

*8. 设 $V$ 是复数域上 $n$ 维线性空间，设线性变换 $\underset{\sim}{A}$ 有 $n$ 个不同的特征值 $\lambda_1,\lambda_2,\cdots,\lambda_n$，求 $\underset{\sim}{A}$ 的所有不变子空间，并且求出 $\underset{\sim}{A}$ 的不变子空间的数目.

(提示：利用定理 9.5.5；注意对于 $\underset{\sim}{A}$ 的任一非零不变子空间 $W$，$\underset{\sim}{A}|W$ 的特征值一定是 $\underset{\sim}{A}$ 的特征值.)

*9. 设 $\underset{\sim}{A}$ 是复数域上 $n$ 维线性空间 $V$ 的线性变换，证明：$\underset{\sim}{A}$ 可对角化的充分必要条件是，$\underset{\sim}{A}$ 的每一个特征值的几何重数等于它的代数重数.

10. 设 $V$ 是实数域上 $n$ 维线性空间，证明：$V$ 的任一线性变换 $\underset{\sim}{A}$ 必有一个 1 维不变子空间或者 2 维不变子空间.

11. 设 $V$ 是平面上以定点 $O$ 为起点的所有向量组成的实数域上的 2 维线性空间，$\underset{\sim}{A}$ 是绕 $O$ 点转角为 $\theta$ 的旋转，其中 $\theta\neq k\pi,k\in Z$. 证明：$\underset{\sim}{A}$ 没有非平凡的不变子空间.

# 补 充 题 九

1. 设 $A,B$ 分别是域 $F$ 上 $m\times n$ 矩阵和 $l\times n$ 矩阵，用 $U$ 表示 $A$ 的列空

间，用 $W$ 表示 $BX=0$ 的解空间.

(1) 令 $\underset{\sim}{A}(\alpha)=A\alpha, \forall\ \alpha\in W$，证明：$\underset{\sim}{A}$ 是 $W$ 到 $U$ 的一个线性映射；

(2) 令 $S:=\{A\alpha \mid \alpha\in W\}$，证明：$S$ 是 $U$ 的一个子空间；

(3) 证明：$\dim S=\operatorname{rank}\begin{pmatrix} A \\ B \end{pmatrix}-\operatorname{rank}B$.

2. 设 $A,B$ 都是 $n\times m$ 实矩阵，用 $W$ 表示 $B'X=0$ 的解空间，取 $W$ 的一个基：$\beta_1,\cdots,\beta_r$. 把矩阵 $(\beta_1,\cdots,\beta_r)$ 记作 $C$. 用 $U_1,U_2$ 分别表示 $A,B$ 的列空间. 证明：如果 $U_1\cap U_2=0$，则 $A'C$ 的列空间 $V_2$ 与 $A'$ 的列空间 $V_1$ 相等.

（提示：去证 $V_2=\{A'\alpha \mid \alpha\in W\}$，然后利用第 1 题的结论.）

3. 设 $A,B$ 分别是域 $F$ 上 $m\times n$，$n\times m$ 矩阵，证明：

$$\operatorname{rank}(A-ABA)=\operatorname{rank}A+\operatorname{rank}(I_n-BA)-n$$

（提示：用两种方法计算下述分块矩阵的秩：

$$\begin{pmatrix} I_n & I_n-BA \\ A & 0 \end{pmatrix}$$

即得结果.）

4. 设 $A$ 是域 $F$ 上 $m\times n$ 矩阵，证明：$B$ 是 $A$ 的一个广义逆的充分必要条件为

$$\operatorname{rank}A+\operatorname{rank}(I_n-BA)=n$$

5. 设 $V$ 是域 $F$ 上 $n$ 维线性空间，$\underset{\sim}{A}_1,\cdots,\underset{\sim}{A}_s$ 是 $V$ 上的线性变换，令 $\underset{\sim}{A}=\sum_{i=1}^{s}\underset{\sim}{A}_i$，证明：如果 $\underset{\sim}{A}$ 是幂等变换，且 $\operatorname{rank}\underset{\sim}{A}=\sum_{i=1}^{s}\operatorname{rank}\underset{\sim}{A}_i$，则 $\underset{\sim}{A}_1,\cdots,\underset{\sim}{A}_s$ 都是幂等变换，且 $\underset{\sim}{A}_i\underset{\sim}{A}_j=0$，当 $i\neq j$.

（提示：在 $V$ 中取一个基，设 $\underset{\sim}{A}_i$ 在此基下的矩阵为 $A_i, i=1,\cdots,s$，则 $\underset{\sim}{A}$ 在此基下的矩阵为 $A=\sum_{i=1}^{s}A_i$. 令 $D:=\operatorname{diag}\{A_1,\cdots,A_s\}$，$E:=(I_n,\cdots,I_n)$. 去证："$\underset{\sim}{A}_1,\cdots,\underset{\sim}{A}_s$ 都是幂等变换且 $\underset{\sim}{A}_i\underset{\sim}{A}_j=0, i\neq j$" 的充分必要条件为 $E'E$ 是 $D$ 的一个广义逆，然后利用第 4 题的结论.）

6. 设 $V$ 是域 $F$ 上 $n$ 维线性空间，$\underset{\sim}{A}_1,\cdots,\underset{\sim}{A}_s$ 是 $V$ 上的线性变换. 证明：如果 $\sum_{i=1}^{s}\underset{\sim}{A}_i=\underset{\sim}{I}$ 且 $\sum_{i=1}^{s}\operatorname{rank}\underset{\sim}{A}_i=n$，则 $\underset{\sim}{A}_1,\cdots,\underset{\sim}{A}_s$ 都是幂等变换，且 $\underset{\sim}{A}_i\underset{\sim}{A}_j=0$，$i\neq j$.

# 第十章　线性变换的 Jordan 标准形

这一章的主要任务是讨论不可以对角化的线性变换的结构，即在空间 $V$ 中要找一个适当的基，使得线性变换 $\underset{\sim}{A}$ 在这个基下的矩阵具有较简单的形式. 一般认为，对角矩阵是最简单的(除去数量矩阵外). 如果 $\underset{\sim}{A}$ 不是可对角化的，那么自然要考虑分块对角矩阵. 据定理 9.6.3，线性变换 $\underset{\sim}{A}$ 在 $V$ 的一个基下的矩阵为分块对角矩阵 $A = \mathrm{diag}\{A_1, A_2, \cdots, A_s\}$ 的充分必要条件是，空间 $V$ 能分解成 $\underset{\sim}{A}$ 的非平凡不变子空间的直和：

$$V = W_1 \oplus W_2 \oplus \cdots \oplus W_s \tag{1}$$

其中 $A_j$ 就是 $\underset{\sim}{A}|W_j$ 在 $W_j$ 的适当一个基下的矩阵. 沿着这条思路来找 $V$ 的一个基，使 $\underset{\sim}{A}$ 在这个基下的矩阵具有较简单的形式，就需要解决两个问题：第一如何把空间 $V$ 分解成 $\underset{\sim}{A}$ 的不变子空间的直和？(1) 式右端的那些 $W_j$ 是什么？第二，在每个 $W_j$ 里，怎样找一个适当的基，使得 $\underset{\sim}{A}|W_j$ 在该基下的矩阵 $A_j$ 具有较简单的形式？如果每个 $A_j$ 都较简单，那么分块对角矩阵 $\mathrm{diag}\{A_1, A_2, \cdots, A_s\}$ 也就较简单了. 这一章就要来解决这些问题.

## §1　线性变换的多项式的核之间的关系

设 $V$ 是域 $F$ 上的 $n$ 维线性空间，$\underset{\sim}{A}$ 是 $V$ 上的一个线性变换. 如何把空间 $V$ 分解或 $\underset{\sim}{A}$ 的不变子空间的直和呢？我们已经知道，$\mathrm{Ker}\,\underset{\sim}{A}$，$\mathrm{Im}\,\underset{\sim}{A}$，$V_{\lambda_i}$ 都是 $\underset{\sim}{A}$ 的不变子空间. 如果 $\underset{\sim}{A}$ 不是可对角化的，那么 $V$ 不能分解成 $\underset{\sim}{A}$ 的特征子空间的直和. 此外，$\mathrm{Ker}\,\underset{\sim}{A} + \mathrm{Im}\,\underset{\sim}{A}$ 一般来说，不等于 $V$. 所以我们需要找 $\underset{\sim}{A}$ 的其他的不变子空间. 我们又

知道，对于任意 $f(x)\in F[x]$，$\mathrm{Ker}\, f(\underline{A})$，$\mathrm{Im}\, f(\underline{A})$ 都是 $\underline{A}$ 的不变子空间(据推论 9.6.1). 由此受到启发，能不能找到一些多项式 $f_1(x),\cdots,f_s(x)$，使得

$$V=\mathrm{Ker}f_1(\underline{A})\oplus\mathrm{Ker}f_2(\underline{A})\oplus\cdots\oplus\mathrm{Ker}f_s(\underline{A})$$

为此，我们首先来讨论，对于不同的一元多项式 $f(x)$，$g(x)$，$x$ 用 $\underline{A}$ 代入得到的 $f(\underline{A})$，$g(\underline{A})$，它们的核有什么关系?

我们在更一般的情形下来讨论，即设 $V$ 是域 $F$ 上任一线性空间(有限维，或者无限维)，$\underline{A}$ 是 $V$ 上的一个线性变换. 下面出现的一元多项式都是域 $F$ 上的一元多项式.

**命题** 10.1.1　如果 $h(x)\mid f(x)$，则 $\mathrm{Ker}h(\underline{A})\subset\mathrm{Ker}f(\underline{A})$.

**证明**　因为 $h(x)\mid f(x)$，所以存在 $g(x)\in F[x]$，使得

$$f(x)=g(x)h(x)$$

不定元 $x$ 用 $\underline{A}$ 代入，得 $f(\underline{A})=g(\underline{A})h(\underline{A})$.

任取 $\alpha\in\mathrm{Ker}h(\underline{A})$，因为

$$f(\underline{A})\alpha=[g(\underline{A})h(\underline{A})]\alpha=g(\underline{A})[h(\underline{A})\alpha]=g(\underline{A})0=0$$

所以 $\alpha\in\mathrm{Ker}f(\underline{A})$. 因此 $\mathrm{Ker}h(\underline{A})\subset\mathrm{Ker}f(\underline{A})$. ▮

**命题** 10.1.2　设 $d(x)$ 是 $f(x)$ 与 $g(x)$ 的一个最大公因式，则

$$\mathrm{Ker}d(\underline{A})=\mathrm{Ker}f(\underline{A})\cap\mathrm{Ker}g(\underline{A})\tag{1}$$

**证明**　据命题 10.1.1 得到，

$$\mathrm{Ker}d(\underline{A})\subset\mathrm{Ker}f(\underline{A})\cap\mathrm{Ker}g(\underline{A})$$

下面来证明 $\mathrm{Ker}d(\underline{A})\supset\mathrm{Ker}f(\underline{A})\cap\mathrm{Ker}g(\underline{A})$.

据定理 7.3.1，有 $F[x]$ 中的多项式 $u(x)$ 与 $v(x)$，使得

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

$x$ 用 $\underline{A}$ 代入得

$$d(\underline{A})=u(\underline{A})f(\underline{A})+v(\underline{A})g(\underline{A})$$

任取 $\alpha\in\mathrm{Ker}f(\underline{A})\cap\mathrm{Ker}g(\underline{A})$，则

$$\begin{aligned}d(\underline{A})\alpha&=[u(\underline{A})f(\underline{A})+v(\underline{A})g(\underline{A})]\alpha\\&=u(\underline{A})(f(\underline{A})\alpha)+v(\underline{A})(g(\underline{A})\alpha)=0\end{aligned}$$

因此 $\alpha \in \mathrm{Ker}d(\underset{\sim}{A})$. 从而(1)式成立. ▎

**命题** 10.1.3　设 $f(x)=f_1(x)f_2(x)$，如果

$$(f_1(x),f_2(x))=1$$

则
$$\mathrm{Ker}f(\underset{\sim}{A})=\mathrm{Ker}f_1(\underset{\sim}{A})\oplus \mathrm{Ker}f_2(\underset{\sim}{A}) \tag{2}$$

**证明**　先证 $\mathrm{Ker}f(\underset{\sim}{A})=\mathrm{Ker}f_1(\underset{\sim}{A})+\mathrm{Ker}f_2(\underset{\sim}{A})$.

因为
$$f_i(x)\mid f(x)$$

所以
$$\mathrm{Ker}f_i(\underset{\sim}{A})\subset \mathrm{Ker}f(\underset{\sim}{A}),\quad i=1,2$$

从而
$$\mathrm{Ker}f_1(\underset{\sim}{A})+\mathrm{Ker}f_2(\underset{\sim}{A})\subset \mathrm{Ker}f(\underset{\sim}{A})$$

因为
$$(f_1(x),f_2(x))=1$$

所以存在 $u_1(x),u_2(x)\in F[x]$，使得

$$u_1(x)f_1(x)+u_2(x)f_2(x)=1$$

$x$ 用 $\underset{\sim}{A}$ 代入得

$$u_1(\underset{\sim}{A})f_1(\underset{\sim}{A})+u_2(\underset{\sim}{A})f_2(\underset{\sim}{A})=\underset{\sim}{I}$$

任取 $\alpha\in Ker\ f(\underset{\sim}{A})$，则

$$\begin{aligned}\alpha=\underset{\sim}{I}\alpha&=[u_1(\underset{\sim}{A})f_1(\underset{\sim}{A})+u_2(\underset{\sim}{A})f_2(\underset{\sim}{A})]\alpha\\&=u_1(\underset{\sim}{A})f_1(\underset{\sim}{A})\alpha+u_2(\underset{\sim}{A})f_2(\underset{\sim}{A})\alpha\end{aligned}$$

令
$$\alpha_1=u_2(\underset{\sim}{A})f_2(\underset{\sim}{A})\alpha,\alpha_2=u_1(\underset{\sim}{A})f_1(\underset{\sim}{A})\alpha$$

则
$$\begin{aligned}f_1(\underset{\sim}{A})\alpha_1&=f_1(\underset{\sim}{A})[u_2(\underset{\sim}{A})f_2(\underset{\sim}{A})\alpha]\\&=u_2(\underset{\sim}{A})[f_1(\underset{\sim}{A})f_2(\underset{\sim}{A})]\alpha\\&=u_2(\underset{\sim}{A})(f(\underset{\sim}{A})\alpha)=0\end{aligned}$$

所以 $\alpha_1\in \mathrm{Ker}f_1(\underset{\sim}{A})$. 同理可证，$\alpha_2\in\mathrm{Ker}f_2(\underset{\sim}{A})$. 因此

$$\alpha=\alpha_1+\alpha_2\in \mathrm{Ker}f_1(\underset{\sim}{A})+\mathrm{Ker}f_2(\underset{\sim}{A})$$

所以

$$\mathrm{Ker}f(\underset{\sim}{A})=\mathrm{Ker}f_1(\underset{\sim}{A})+\mathrm{Ker}f_2(\underset{\sim}{A})$$

再证和 $\mathrm{Ker}f_1(\underset{\sim}{A})+\mathrm{Ker}f_2(\underset{\sim}{A})$ 是直和. 因为

$$(f_1(x),f_2(x))=1$$

据命题 10.1.2 得

$$\mathrm{Ker}f_1(\underset{\sim}{A})\cap \mathrm{Ker}f_2(\underset{\sim}{A})=\mathrm{Ker}\underset{\sim}{I}=0$$

所以上述和是直和．从而(2)式成立． ▌

用数学归纳法可以把命题 10.1.3 推广成：

**定理** 10.1.1　设 $f(x)=f_1(x)f_2(x)\cdots f_s(x)$，如果 $f_1(x)$，$f_2(x)$，$\cdots$，$f_s(x)$ 两两互素，则

$$\mathrm{Ker}f(\underline{A})=\mathrm{Ker}f_1(\underline{A})\oplus\mathrm{Ker}f_2(\underline{A})\oplus\cdots\oplus\mathrm{Ker}f_s(\underline{A}) \quad (3)$$

**证明**　$s=2$ 时已证．假设命题对于 $s-1$ 时成立，来看

$$f(x)=f_1(x)\cdots f_{s-1}(x)f_s(x)$$

的情形．因为 $f_1(x)$，$\cdots$，$f_s(x)$ 两两互素，所以

$$(f_1(x)\cdots f_{s-1}(x),f_s(x))=1$$

记 $g(x)=f_1(x)\cdots f_{s-1}(x)$，据命题 10.1.3，得

$$\mathrm{Ker}f(\underline{A})=\mathrm{Ker}g(\underline{A})\oplus\mathrm{Ker}f_s(\underline{A}) \quad (4)$$

据归纳假设得

$$\mathrm{Ker}g(\underline{A})=\mathrm{Ker}f_1(\underline{A})\oplus\cdots\oplus\mathrm{Ker}f_{s-1}(\underline{A}) \quad (5)$$

从(4)和(5)式即得(3)式． ▌

由于 $\mathrm{Ker}\,\underline{0}=V$，因此从定理 10.1.1 受到启发，如果能找到一个多项式 $f(x)$，使得 $f(\underline{A})=0$，那么空间 $V$ 就能分解成

$$V=\mathrm{Ker}f_1(\underline{A})\oplus\cdots\oplus\mathrm{Ker}f_s(\underline{A}) \quad (6)$$

其中，$f_1(x)\cdots f_s(x)=f(x)$，并且 $f_1(x)$，$\cdots$，$f_s(x)$ 两两互素．

**定义** 1　设 $\underline{A}$ 是 $V$ 上的线性变换，如果 $F$ 上的一元多项式 $f(x)$，使得 $f(\underline{A})=\underline{0}$，则称 $f(x)$ 是 $\underline{A}$ 的一个**零化多项式**．

换句话说，如果 $f(x)$ 在 $F[\underline{A}]$ 中的一个根是 $\underline{A}$，则称 $f(x)$ 是 $\underline{A}$ 的一个零化多项式．

如果 $V$ 是有限维的，那么 $V$ 上的任意一个线性变换 $\underline{A}$ 都有非零的零化多项式．理由如下：设 $\dim V=n$，则 $\dim\mathrm{Hom}(V,V)=n^2$．从而 $\underline{I}$，$\underline{A}$，$\underline{A}^2$，$\cdots$，$\underline{A}^{n^2}$ 一定线性相关．于是有 $F$ 中不全为零的元素 $k_0$，$k_1$，$\cdots$，$k_{n^2}$，使得

$$k_0\underline{I}+k_1\underline{A}+k_2\underline{A}^2+\cdots+k_{n^2}\underline{A}^{n^2}=\underline{0}$$

令　$$f(x)=k_0+k_1x+k_2x^2+\cdots+k_{n^2}x^{n^2}$$

则由上式得 $f(\underset{\sim}{A})=\underset{\sim}{0}$. 因此 $f(x)$ 是 $\underset{\sim}{A}$ 的一个零化多项式，并且 $f(x)$ 不是零多项式.

上述证明给出了找 $\underset{\sim}{A}$ 的非零的零化多项式的一种方法. 除此之外，还有没有别的方法？我们将在下一节继续讨论.

类似于定义 1，可给出 $n$ 级矩阵 $A$ 的零化多项式的概念. 即如果 $f(x)$ 在 $F[A]$ 中的一个根是 $A$，则称 $f(x)$ 是 $A$ 的一个**零化多项式**.

设线性变换 $\underset{\sim}{A}$ 在 $V$ 的一个基下的矩阵是 $A$，据定理 9.4.2 得

$$f(x) \text{ 是 } \underset{\sim}{A} \text{ 的零化多项式} \Leftrightarrow f(x) \text{ 是 } A \text{ 的零化多项式}$$

## 习 题 10.1

1. 写出零变换的一个非零的零化多项式.

2. 写出数乘变换 $\underset{\sim}{k}$ 的一个非零的零化多项式.

3. 证明：如果线性变换 $\underset{\sim}{A}$ 有一个零化多项式是 1 次的，则 $\underset{\sim}{A}$ 一定是数乘变换.

*4. 求出数域 $K$ 上 2 维线性空间 $V$ 上的下列线性变换 $\underset{\sim}{A}$ 的一个非零的零化多项式，$\underset{\sim}{A}$ 在 $V$ 的一个基下的矩阵分别为

(1) $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$；　(2) $\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$；　(3) $\begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}$.

*5. 求下述 $n$ 级矩阵 $A$ 的一个非零的零化多项式：

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

## § 2 Hamilton-Cayley 定理

上一节我们指出，对于有限维线性空间 $V$，它的每一个线性变换都有非零的零化多项式. 这一节我们来证明，线性变换 $\underline{A}$ 的特征多项式 $f(\lambda)$ 就是 $\underline{A}$ 的一个零化多项式. 这个结论称为 Hamilton-Cayley 定理. 为了证明这个定理，我们先介绍一些预备知识.

我们知道，域 $F$ 上的一个 $n$ 级矩阵是指由 $F$ 中 $n^2$ 个元素排成的 $n$ 行 $n$ 列的一张表. 同样地，我们可以考虑整环 $R$ 上的 $n$ 级矩阵，它是由整环 $R$ 中的元素排成的表. 整环 $R$ 上所有 $n$ 级矩阵组成的集合记作 $M_n(R)$. 我们知道，$M_n(F)$ 有加法、纯量乘法、乘法三种运算，注意到它们都是通过 $F$ 中元素的加法、乘法进行. 因此可以用同样方式定义 $M_n(R)$ 的加法、纯量乘法、乘法运算.

域 $F$ 上 $n$ 级矩阵的行列式的定义只用到 $F$ 中元素的加法和乘法，因此可以用同样方式定义整环 $R$ 上 $n$ 级矩阵的行列式. 第二章 § 4 对于数域 $K$ 上 $n$ 级矩阵证明的行列式的所有性质对于任意域 $F$ 上的矩阵一样成立（只是当 $\mathrm{char}F = 2$ 时，推论 2 要换一个证法），对于整环 $R$ 上的矩阵也成立. 从而行列式按一行（列）展开的定理 2.5.1，定理 2.5.2，定理 2.5.3，定理 2.5.4 也成立. 进而

$$A^*A = AA^* = |A|I$$

仍成立，其中 $A^*$ 是 $A$ 的伴随矩阵.

域 $F$ 上的一元多项式环 $F[\lambda]$ 是一个整环，因此可以考虑环 $F[\lambda]$ 上的 $n$ 级矩阵，称这样的矩阵为 $\lambda$- 矩阵，记成 $A(\lambda)$，$B(\lambda)$，…，等. 例如

$$A(\lambda) = \begin{pmatrix} 2\lambda^3 + \lambda^2 + 1 & \lambda^2 - 3 \\ \lambda^3 - 1 & 2\lambda + 5 \end{pmatrix} \tag{1}$$

是一个 $\lambda$- 矩阵，按照整环上矩阵的加法与纯量乘法，可以把（1）中

的 $\lambda$- 矩阵 $A(\lambda)$ 写成

$$A(\lambda)=\begin{pmatrix}2\lambda^3 & 0\\ \lambda^3 & 0\end{pmatrix}+\begin{pmatrix}\lambda^2 & \lambda^2\\ 0 & 0\end{pmatrix}+\begin{pmatrix}0 & 0\\ 0 & 2\lambda\end{pmatrix}+\begin{pmatrix}1 & -3\\ -1 & 5\end{pmatrix}$$

$$=\lambda^3\begin{pmatrix}2 & 0\\ 1 & 0\end{pmatrix}+\lambda^2\begin{pmatrix}1 & 1\\ 0 & 0\end{pmatrix}+\lambda\begin{pmatrix}0 & 0\\ 0 & 2\end{pmatrix}+\begin{pmatrix}1 & -3\\ -1 & 5\end{pmatrix} \tag{2}$$

其中 $\lambda^i$ 的"系数矩阵"是域 $F$ 上的矩阵.

当我们把两个 $\lambda$- 矩阵 $A(\lambda)$ 和 $B(\lambda)$ 都写成(2) 那样的形式时，根据两个一元多项式相等的定义以及两个 $\lambda$- 矩阵相等的定义，可以推出，$A(\lambda)$ 与 $B(\lambda)$ 相等当且仅当它们的 $\lambda^i$ 的"系数矩阵"对应相等，$i=0,1,2,\cdots$.

有了上述准备知识后，我们来证明 Hamilton-Cayloy 定理.

Hamilton-Cayley **定理** 设 $A$ 是域 $F$ 上的 $n$ 级矩阵，则 $A$ 的特征多项式 $f(\lambda)$ 是 $A$ 的一个零化多项式.

**证明** 设 $B(\lambda)$ 是 $\lambda$- 矩阵 $(\lambda I-A)$ 的伴随矩阵，则

$$B(\lambda)(\lambda I-A)=|\lambda I-A|I=f(\lambda)I \tag{3}$$

因为 $B(\lambda)$ 的元素是矩阵 $\lambda I-A$ 的元素(一次或零次多项式) 的代数余子式，所以 $B(\lambda)$ 的元素都是次数不超过 $n-1$ 的一元多项式. 因此 $B(\lambda)$ 可以写成

$$B(\lambda)=\lambda^{n-1}B_{n-1}+\lambda^{n-2}B_{n-2}+\cdots+\lambda B_1+B_0 \tag{4}$$

其中 $B_{n-1},B_{n-2},\cdots,B_1,B_0$ 都是域 $F$ 上的 $n$ 级矩阵.

直接计算得

$$\begin{aligned}&B(\lambda)(\lambda I-A)\\&=(\lambda^{n-1}B_{n-1}+\lambda^{n-2}B_{n-2}+\cdots+\lambda B_1+B_0)(\lambda I-A)\\&=\lambda^nB_{n-1}+\lambda^{n-1}(B_{n-2}-B_{n-1}A)+\\&\quad\cdots+\lambda(B_0-B_1A)-B_0A\end{aligned} \tag{5}$$

由于 $A$ 的特征多项式 $f(\lambda)$ 的次数是 $n$，于是可以设

$$f(\lambda)=\lambda^n+a_{n-1}\lambda^{n-1}+\cdots+a_1\lambda+a_0 \tag{6}$$

从而
$$f(\lambda)I=\lambda^nI+a_{n-1}\lambda^{n-1}I+\cdots+a_1\lambda I+a_0I \tag{7}$$

据(3)，(5)，(7)式得

$$
\begin{cases}
B_{n-1} = I \\
B_{n-2} - B_{n-1}A = a_{n-1}I \\
\cdots \quad \cdots \quad \cdots \quad \cdots \quad \cdots \\
B_0 - B_1A = a_1I \\
-B_0A = a_0I
\end{cases} \tag{8}
$$

用 $A^n, A^{n-1}, \cdots, A, I$ 依次从右边乘(8)的第 1 式，第 2 式，…，第 $n$ 式，第 $n+1$ 式，得

$$
\begin{cases}
B_{n-1}A^n = A^n \\
B_{n-2}A^{n-1} - B_{n-1}A^n = a_{n-1}A^{n-1} \\
\cdots \quad \cdots \quad \cdots \quad \cdots \quad \cdots \quad \cdots \quad \cdots \\
B_0A - B_1A^2 = a_1A \\
-B_0A = a_0I
\end{cases} \tag{9}
$$

把(9)的 $n+1$ 个式子一起加起来，左边为零，右边即为 $f(A)$. 因此 $f(A)=0$. ▌

用线性变换的语言叙述上述定理，即

Hamilton-Cayley **定理** 设 $\underset{\sim}{A}$ 是域 $F$ 上 $n$ 维线性空间 $V$ 上的一个线性变换，则 $\underset{\sim}{A}$ 的特征多项式 $f(\lambda)$ 是 $\underset{\sim}{A}$ 的一个零化多项式. ▌

## 习 题 10.2

1. 设 $\underset{\sim}{A}$ 是域 $F$ 上 $n$ 维线性空间 $V$ 上的线性变换.

证明：如果 $\underset{\sim}{A}$ 的特征多项式 $f(\lambda)$ 在 $F[\lambda]$ 中可以分解成

$$f(\lambda) = (\lambda-\lambda_1)^{r_1}(\lambda-\lambda_2)^{r_2}\cdots(\lambda-\lambda_s)^{r_s}$$

则 $V = \operatorname{Ker}(\underset{\sim}{A}-\lambda_1\underset{\sim}{I})^{r_1} \oplus \operatorname{Ker}(\underset{\sim}{A}-\lambda_2\underset{\sim}{I})^{r_2} \oplus \cdots \oplus \operatorname{Ker}(\underset{\sim}{A}-\lambda_s\underset{\sim}{I})^{r_s}$

2. 证明：对于域 $F$ 上的 $n$ 级可逆矩阵 $A$，存在 $F$ 中元素 $k_0, k_1, \cdots k_{n-1}$，使得

$$A^{-1} = k_{n-1}A^{n-1} + \cdots + k_1A + k_0I$$

*3. 设 $A,B$ 分别是 $n$ 级，$m$ 级复矩阵，证明：矩阵方程 $AX-XB=0$ 只有零解的充分必要条件是，$A$ 与 $B$ 没有公共的特征值.

（提示：必要性用反证法，假如 $A$ 与 $B$ 有公共特征值，去找一个非零矩阵 $C$，满足 $AX-XB=0$. 充分性：设 $f(\lambda),g(\lambda)$ 分别是 $A,B$ 的特征多项式. 用 Hamilton-Cayley 定理去证 $g(A)$ 可逆. 设 $C$ 是 $AX-XB=0$ 的解，去证 $C=0$.）

*4. 设 $V$ 是域 $F$ 上 $n$ 维线性空间，$F$ 是一个代数封闭域（即 $F[x]$ 中每个多项式在 $F$ 中有根. 譬如，复数域是一个代数封闭域）. 用数学归纳法证明 Hamilton-Cayley 定理.

（提示：设 $\underset{\sim}{A}$ 是 $V$ 上线性变换，$f(\lambda)$ 是 $\underset{\sim}{A}$ 的特征多项式. 对线性空间的维数 $n$ 用数学归纳法. $n=1$ 时，设 $V=\langle\alpha\rangle$，则 $\underset{\sim}{A}\alpha=k\alpha$ 对某个 $k\in F$. 于是 $\underset{\sim}{A}=\underset{\sim}{k}$，从而 $f(\lambda)=\lambda-k$. 所以

$$f(\underset{\sim}{A})=\underset{\sim}{k}-k\underset{\sim}{I}=\underset{\sim}{0}$$

假设命题对于 $n-1$ 维线性空间成立. 设 $V$ 是 $n$ 维. 设 $\lambda_1$ 是 $\underset{\sim}{A}$ 的一个特征值，$\alpha_1$ 是 $\underset{\sim}{A}$ 的属于 $\lambda_1$ 的特征向量. 令 $W=\langle\alpha_1\rangle$. 把 $\alpha_1$ 扩充成 $V$ 的一个基 $\alpha_1,\alpha_2,\cdots,\alpha_n$. 因为 $W$ 是 $\underset{\sim}{A}$ 的不变子空间，所以 $\underset{\sim}{A}$ 在 $V$ 的基 $\alpha_1,\alpha_2,\cdots,\alpha_n$ 下的矩阵 $A$ 是分块上三角矩阵

$$A=\begin{pmatrix}\lambda_1 & A_1\\ 0 & A_2\end{pmatrix}$$

其中 $A_2$ 是商空间 $V/W$ 上的线性变换 $\underset{\sim}{\bar{A}}$ 在基 $\alpha_2+W,\cdots,\alpha_n+W$ 下的矩阵. 于是 $\underset{\sim}{\bar{A}}$ 的特征多项式为 $f_2(\lambda)=|\lambda I_{n-1}-A_2|$. 据归纳假设，$f_2(\underset{\sim}{\bar{A}})=\underset{\sim}{0}$. 去证 $f(\underset{\sim}{A})=\underset{\sim}{0}$. 注意 $f(\lambda)=(\lambda-\lambda_1)f_2(\lambda)$，因此

$$f(\underset{\sim}{A})=(\underset{\sim}{A}-\lambda_1\underset{\sim}{I})f_2(\underset{\sim}{A})$$

任取 $\alpha\in V$，设 $\alpha=\sum_{i=1}^{n}b_i\alpha_i$. 我们有

$$f_2(\underset{\sim}{A})\alpha=\sum_{i=1}^{n}b_if_2(\underset{\sim}{A})\alpha_i$$

而
$$f_2(\underset{\sim}{\bar{A}})(\alpha_i+W)=f_2(\underset{\sim}{A})\alpha_i+W$$

由 $f_2(\underset{\sim}{\bar{A}})=\underset{\sim}{0}$ 得，$f_2(\underset{\sim}{A})\alpha_i\in W$，所以 $f_2(\underset{\sim}{A})\alpha\in W$. 由此得出

$$f(\underset{\sim}{A})\alpha=(\underset{\sim}{A}-\lambda_1\underset{\sim}{I})f_2(\underset{\sim}{A})\alpha=0$$

从而定理得证.）

## §3 线性变换和矩阵的的最小多项式

从 Hamilton-Cayley 定理知道，有限维线性空间 $V$ 上的线性变换 $\underline{A}$ 的特征多项式 $f(\lambda)$ 是 $\underline{A}$ 的一个零化多项式. 这一节我们来讨论 $\underline{A}$ 有没有别的零化多项式?如果还有的话，那么 $\underline{A}$ 的零化多项式之间有什么关系?

**定义** 1　设 $\underline{A}$ 是域 $F$ 上线性空间 $V$ 的一个线性变换，在 $\underline{A}$ 的所有非零的零化多项式中，次数最低的首项系数为 1 的多项式称为 $\underline{A}$ 的**最小多项式**.

**命题** 10.3.1　线性空间 $V$ 上的线性变换 $\underline{A}$ 的最小多项式是唯一的.

**证明**　设 $m_1(\lambda)$ 和 $m_2(\lambda)$ 都是 $\underline{A}$ 的最小多项式，则它们的次数相同并且首项系数都为1，从而 $h(\lambda) = m_1(\lambda) - m_2(\lambda)$ 的次数比 $m_1(\lambda)$ 的低. 由于 $h(\underline{A}) = m_1(\underline{A}) - m_2(\underline{A}) = \underline{0}$，所以 $h(\lambda)$ 也是 $\underline{A}$ 的一个零化多项式. 据最小多项式的定义得，$h(\lambda) = 0$. 即

$$m_1(\lambda) = m_2(\lambda) \qquad \blacksquare$$

**命题** 10.3.2　$\underline{A}$ 的任一零化多项式 $g(\lambda)$ 是 $\underline{A}$ 的最小多项式 $m(\lambda)$ 的倍式.

**证明**　作带余除法，得

$$g(\lambda) = h(\lambda)m(\lambda) + r(\lambda), \quad \deg r(\lambda) < \deg m(\lambda)$$

不定元 $\lambda$ 用 $\underline{A}$ 代入，从上式得

$$g(\underline{A}) = h(\underline{A})m(\underline{A}) + r(\underline{A})$$

由于 $g(\underline{A}) = \underline{0}$，$m(\underline{A}) = \underline{0}$，所以 $r(\underline{A}) = \underline{0}$. 据最小多项式的定义得，$r(\lambda) = 0$. 因此 $g(\lambda)$ 是 $m(\lambda)$ 的倍式.　$\blacksquare$

由命题 10.3.2 和 Hamilton-Cayley 定理得，域 $F$ 上有限维线性空间 $V$ 的线性变换 $\underline{A}$ 的特征多项式 $f(\lambda)$ 是 $\underline{A}$ 的最小多项式 $m(\lambda)$ 的倍式. 从而 $\underline{A}$ 的最小多项式 $m(\lambda)$ 在 $F$ 中的根都是 $\underline{A}$ 的特

征多项式 $f(\lambda)$ 在 $F$ 中的根. 反之，$\underline{A}$ 的特征多项式 $f(\lambda)$ 在 $F$ 中的根都是 $\underline{A}$ 的最小多项式在 $F$ 中的根(重数可以不同). 理由如下：设 $\lambda_0$ 是 $f(\lambda)$ 在 $F$ 中的一个根，则 $\lambda_0$ 是 $\underline{A}$ 的特征值，从而存在 $V$ 中非零向量 $\xi$，使得 $\underline{A}\xi = \lambda_0\xi$. 设

$$m(\lambda) = c_0 + c_1\lambda + \cdots + c_s\lambda^s$$

则

$$\begin{aligned} 0 &= m(\underline{A})\xi = (c_0\underline{I} + c_1\underline{A} + \cdots + c_s\underline{A}^s)\xi \\ &= c_0\xi + c_1\lambda_0\xi + \cdots + c_s\lambda_0^s\xi = m(\lambda_0)\xi \end{aligned}$$

从而 $m(\lambda_0) = 0$，即 $\lambda_0$ 是 $m(\lambda)$ 的一个根. 这样我们证明了：

**命题** 10.3.3 设 $\underline{A}$ 是域 $F$ 上 $n$ 维线性空间 $V$ 的一个线性变换，则 $\underline{A}$ 的特征多项式与 $\underline{A}$ 的最小多项式在 $F$ 中有相同的根(重数可以不同). ▌

类似地，可以定义域 $F$ 上 $n$ 级矩阵 $A$ 的最小多项式.

设 $n$ 维线性空间 $V$ 上的一个线性变换 $\underline{A}$ 在 $V$ 的一个基下的矩阵是 $A$，我们在 §1 的最后一段指出，$g(\lambda)$ 是 $\underline{A}$ 的零化多项式当且仅当 $g(\lambda)$ 是 $A$ 的零化多项式. 由此得出，$\underline{A}$ 的所有零化多项式组成的集合与 $A$ 的所有零化多项式组成的集合相等. 从而 $m(\lambda)$ 是 $\underline{A}$ 的最小多项式当且仅当 $m(\lambda)$ 是 $A$ 的最小多项式.

由于相似的矩阵可以看作是同一个线性变换 $\underline{A}$ 在 $V$ 的不同基下的矩阵，从上一段的结论立即得出，相似的矩阵有相同的最小多项式.

注意，上述结论的逆命题不成立. 即最小多项式相同的矩阵不一定是相似的. 例子可看下面的例 3.

**定义** 2 线性空间 $V$ 上的线性变换 $\underline{A}$，如果存在一个正整数 $l$，使得 $\underline{A}^l = \underline{0}$，则称 $\underline{A}$ 是**幂零变换**. 使得 $\underline{A}^l = \underline{0}$ 成立的最小正整数 $l$ 称为 $\underline{A}$ 的**幂零指数**.

**例** 1 证明：幂零变换 $\underline{A}$ 的最小多项式为 $\lambda^l$，其中 $l$ 是 $\underline{A}$ 的幂零指数.

**证明** 因为 $l$ 是 $\underline{A}$ 的幂零指数，所以 $\underline{A}^l = \underline{0}$. 从而 $\lambda^l$ 是 $\underline{A}$ 的一

个零化多项式. 因此 $\underline{A}$ 的最小多项式 $m(\lambda)$ 是 $\lambda^l$ 的因式. 由于当 $0<r<l$ 时, $\underline{A}^r \neq \underline{0}$, 所以 $m(\lambda)=\lambda^l$. ▌

**定义** 3 域 $F$ 上的一个 $r$ 级矩阵如果形如

$$\begin{pmatrix} a & 1 & 0 & \cdots & 0 & 0 \\ 0 & a & 1 & \cdots & 0 & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \cdots & a & 1 \\ 0 & 0 & 0 & \cdots & 0 & a \end{pmatrix} \tag{1}$$

则称它为一个 $r$ **级 Jordan 块**. 主对角元为 $a$ 的 $r$ 级 Jordan 块, 记作 $J_r(a)$.

**例** 2 证明: 主对角元为 $a$ 的 $r$ 级 Jordan 块 $J_r(a)$ 的最小多项式 $m(\lambda)$ 等于它的特征多项式 $f(\lambda)=(\lambda-a)^r$.

**证明** 显然, $J_r(a)$ 的特征多项式 $f(\lambda)=(\lambda-a)^r$. 于是它的最小多项式 $m(\lambda)$ 为 $(\lambda-a)^r$ 的一个因式 $(\lambda-a)^s$, 其中 $s\leqslant r$. 于是有 $(J_r(a)-aI)^s=0$. 因为

$$J_r(a)-aI=\begin{pmatrix} 0 & 1 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 1 & \cdots & 0 & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \cdots & 0 & 1 \\ 0 & 0 & 0 & \cdots & 0 & 0 \end{pmatrix}=J_r(0)$$

所以 $(J_r(a)-aI)^k\neq 0$, 当 $0\leqslant k\leqslant r-1$ (参见习题 4.3 的第 5 题 (7)). 因此 $s=r$, 即 $m(\lambda)=(\lambda-a)^r$. ▌

**定理** 10.3.1 设 $V$ 是域 $F$ 上的线性空间, $\underline{A}$ 是 $V$ 上的一个线性变换. 如果 $V$ 能分解成 $\underline{A}$ 的若干个非平凡不变子空间的直和

$$V=W_1\oplus W_2\oplus\cdots\oplus W_s \tag{2}$$

则 $\underline{A}$ 的最小多项式

$$m(\lambda)=[m_1(\lambda),m_2(\lambda),\cdots,m_s(\lambda)]$$

其中 $m_j(\lambda)$ 是 $W_j$ 上的线性变换 $\underline{A}|W_j$ 的最小多项式, $j=1,\cdots,s$, $[m_1(\lambda),m_2(\lambda),\cdots,m_s(\lambda)]$ 是 $m_1(\lambda),m_2(\lambda),\cdots,m_s(\lambda)$ 的最小公

倍式.

**证明** 记

$$g(\lambda)=[m_1(\lambda),m_2(\lambda),\cdots,m_s(\lambda)]$$

设 $$g(\lambda)=h_j(\lambda)m_j(\lambda),\quad j=1,\cdots,s$$

任 $\alpha\in V$,设

$$\alpha=\alpha_1+\alpha_2+\cdots+\alpha_s,\alpha_j\in W_j,\quad j=1,\cdots,s$$

因为 $$m_j(\underline{A})\alpha_j=m_j(\underline{A}|W_j)\alpha_j=\underline{0}\alpha_j=0,\quad j=1,\cdots,s$$

所以 $$g(\underline{A})\alpha=g(\underline{A})\alpha_1+g(\underline{A})\alpha_2+\cdots+g(\underline{A})\alpha_s$$

$$=\sum_{j=1}^{s}h_j(\underline{A})m_j(\underline{A})\alpha_j=0$$

这表明 $g(\underline{A})=\underline{0}$,因此,$m(\lambda)\,|\,g(\lambda)$.

因为对任一 $\alpha_j\in W_j$,有

$$m(\underline{A}|W_j)\alpha_j=m(\underline{A})\alpha_j=\underline{0}\alpha_j=0$$

所以 $m(\underline{A}|W_j)=\underline{0}$. 因此 $m_j(\lambda)\,|\,m(\lambda)$,$j=1,\cdots,s$. 从而

$$[m_1(\lambda),\cdots,m_s(\lambda)]\,|\,m(\lambda)$$

即 $g(\lambda)\,|\,m(\lambda)$. 所以 $m(\lambda)=g(\lambda)$. ▍

**推论** 10.3.1 设 $A$ 是域 $F$ 上一个 $n$ 级分块对角矩阵,即 $A=\operatorname{diag}\{A_1,\cdots,A_s\}$,则 $A$ 的最小多项式 $m(\lambda)=[m_1(\lambda),\cdots,m_s(\lambda)]$,其中 $m_j(\lambda)$ 是 $A_j$ 的最小多项式.

**证明** 设 $V$ 是域 $F$ 上 $n$ 维线性空间,于是 $A$ 是 $V$ 上一个线性变换 $\underline{A}$ 在 $V$ 的一个基下的矩阵. 据定理 9.6.3,$V$ 能分解成 $\underline{A}$ 的若干个非平凡不变子空间的直和:

$$V=W_1\oplus\cdots\oplus W_s$$

并且 $A_j$ 就是 $\underline{A}|W_j$ 在 $W_j$ 的适当基下的矩阵,从而 $\underline{A}|W_j$ 的最小多项式就是 $A_j$ 的最小多项式 $m_j(\lambda)$,$j=1,\cdots,s$. 又 $\underline{A}$ 的最小多项式是 $A$ 的最小多项式 $m(\lambda)$. 据定理 10.3.1 得

$$m(\lambda)=[m_1(\lambda),\cdots,m_s(\lambda)]$$ ▍

**例** 3 求数域 $K$ 上下列矩阵的最小多项式:

$$A=\begin{pmatrix}3&1&0&0\\0&3&0&0\\0&0&3&0\\0&0&0&5\end{pmatrix},\qquad B=\begin{pmatrix}3&1&0&0\\0&3&0&0\\0&0&5&0\\0&0&0&5\end{pmatrix}$$

**解**　2 级 Jordan 块 $J_2(3)$ 的最小多项式是 $(\lambda-3)^2$. 据推论 10.3.1 得, $A$ 的最小多项式为

$$[(\lambda-3)^2,\lambda-3,\lambda-5]=(\lambda-3)^2(\lambda-5)$$

$B$ 的最小多项式为

$$[(\lambda-3)^2,\lambda-5,\lambda-5]=(\lambda-3)^2(\lambda-5)$$

例 3 的矩阵 $A$ 与 $B$ 有相同的最小多项式 $(\lambda-3)^2(\lambda-5)$, 但是 $A$ 的特征多项式为 $(\lambda-3)^3(\lambda-5)$, $B$ 的特征多项式为 $(\lambda-3)^2(\lambda-5)^2$, 所以 $A$ 与 $B$ 不相似.

**定义 4**　由若干个 Jordan 块组成的分块对角矩阵称为 **Jordan 形矩阵**.

例 3 中的矩阵 $A$, $B$ 都是 Jordan 形矩阵. 由于一级 Jordan 块就是一级矩阵, 因此对角矩阵可以看成是由一级 Jordan 块组成的 Jordan 形矩阵.

设一个 Jordan 形矩阵

$$A=\operatorname{diag}\{J_{r_1}(a),\cdots,J_{r_s}(a),\cdots,J_{t_1}(b),\cdots,J_{t_m}(b)\} \tag{3}$$

据推论 10.3.1 和例 2 得, $A$ 的最小多项式 $m(\lambda)$ 为

$$\begin{aligned}m(\lambda)&=[(\lambda-a)^{r_1},\cdots,(\lambda-a)^{r_s},\cdots,(\lambda-b)^{t_1},\cdots,(\lambda-b)^{t_m}]\\&=(\lambda-a)^{\max\{r_1,\cdots,r_s\}}\cdots(\lambda-b)^{\max\{t_1,\cdots,t_m\}}\end{aligned} \tag{4}$$

利用(4)式可以立即写出一个 Jordan 形矩阵的最小多项式.

线性变换的最小多项式在研究线性变换的结构中起着十分重要的作用. 现在先利用最小多项式给出线性变换可对角化的一个充分必要条件, 下一节将继续用最小多项式研究不可以对角化的线性变换的结构.

**定理 10.3.2**　设 $V$ 是域 $F$ 上 $n$ 维线性空间, $V$ 上的一个线性

变换 $\underline{A}$ 可对角化的充分必要条件是，$\underline{A}$ 的最小多项式 $m(\lambda)$ 能分解成 $F$ 上不同的一次因式的乘积：

$$m(\lambda)=(\lambda-a_1)(\lambda-a_2)\cdots(\lambda-a_r) \tag{5}$$

其中 $a_1,a_2,\cdots,a_r$ 两两不同.

**证明** 必要性. 设 $\underline{A}$ 可对角化，设 $\underline{A}$ 的全部不同的特征值是 $\lambda_1,\lambda_2,\cdots,\lambda_r$. 据定理 9.5.5，$V$ 可分解成 $\underline{A}$ 的特征子空间 $V_{\lambda_j}$，$j=1,\cdots,r$，的直和：

$$V=V_{\lambda_1}\oplus V_{\lambda_2}\oplus\cdots\oplus V_{\lambda_r}$$

据定理 10.3.1，$\underline{A}$ 的最小多项式 $m(\lambda)=[m_1(\lambda),\cdots,m_r(\lambda)]$，其中 $m_j(\lambda)$ 是 $\underline{A}|V_{\lambda_j}$ 的最小多项式，$j=1,\cdots,r$. 因为 $\underline{A}|V_{\lambda_j}=\underline{\lambda_j}$，而数乘变换 $\underline{\lambda_j}$ 的一个零化多项式是 $\lambda-\lambda_j$，所以 $m_j(\lambda)=\lambda-\lambda_j$，$j=1,\cdots,r$. 因此

$$m(\lambda)=(\lambda-\lambda_1)(\lambda-\lambda_2)\cdots(\lambda-\lambda_r)$$

充分性. 设 $\underline{A}$ 的最小多项式 $m(\lambda)$ 可分解成(5)式，其中 $a_1$，$a_2\cdots,a_r$ 是 $F$ 中的两两不同的元素. 因为 $\lambda-a_1,\cdots,\lambda-a_r$ 两两互素，据定理 10.1.1，得

$$\operatorname{Ker}m(\underline{A})=\operatorname{Ker}(\underline{A}-a_1\underline{I})\oplus\cdots\oplus\operatorname{Ker}(\underline{A}-a_r\underline{I}) \tag{6}$$

显然，$\operatorname{Ker}m(\underline{A})=\operatorname{Ker}\underline{0}=V$. 因为

$$\begin{aligned}\alpha\in\operatorname{Ker}(\underline{A}-a_j\underline{I})&\Leftrightarrow(\underline{A}-a_j\underline{I})\alpha=0\\&\Leftrightarrow\underline{A}\alpha=a_j\alpha\Leftrightarrow\alpha\in V_{a_j}\end{aligned}$$

其中 $V_{a_j}$ 是 $\underline{A}$ 的属于特征值 $a_j$ 的特征子空间，所以

$$\operatorname{Ker}(\underline{A}-a_j\underline{I})=V_{a_j},\qquad j=1,\cdots,r \tag{7}$$

于是从(6)和(7)式得

$$V=V_{a_1}\oplus\cdots\oplus V_{a_r} \tag{8}$$

因此 $\underline{A}$ 可对角化. ▌

**推论 10.3.2** 域 $F$ 上 $n$ 级矩阵 $A$ 可对角化的充分必要条件是，$A$ 的最小多项式能分解成 $F$ 上不同的一次因式的乘积. ▌

**推论 10.3.3** 复数域上 $n$ 级矩阵 $A$ 可对角化的充分必要条

件是 $A$ 的最小多项式没有重根. ▌

利用最小多项式来判断一个线性变换或矩阵是不是可对角化,很简便. 例如,域 $F$ 上 $n$ 维线性空间 $V$ 的幂零变换 $\underset{\sim}{A}$,如果它的幂零指数 $l>1$,那么 $\underset{\sim}{A}$ 一定不可以对角化,这是因为 $\underset{\sim}{A}$ 的最小多项式 $m(\lambda)=\lambda^l$. 同理,级数 $r>1$ 的 Jordan 块 $J_r(a)$ 一定不可以对角化,因为它的最小多项式 $m(\lambda)=(\lambda-a)^r$. 进而可以知道,包含级数大于 1 的 Jordan 块的 Jordan 形矩阵一定不可以对角化.

## 习　题　10.3

1. 求下列复矩阵的最小多项式:

(1) $\begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$;　　(2) $\begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}$;

(3) $\begin{pmatrix} 7 & -1 & -7 & 1 \\ -1 & 7 & 1 & -7 \\ 7 & -1 & -7 & 1 \\ -1 & 7 & 1 & -7 \end{pmatrix}$;　　(4) $\begin{pmatrix} 3 & 1 & -1 \\ 0 & 2 & 0 \\ 1 & 1 & 1 \end{pmatrix}$

2. 第 1 题中各矩阵作为实数域上矩阵是否可对角化?

3. 设 $V$ 是数域 $K$ 上 3 维线性空间,$V$ 上的线性变换 $\underset{\sim}{A}$ 在基 $\alpha_1,\alpha_2,\alpha_3$ 下的矩阵为

$$A=\begin{pmatrix} 1 & 0 & 0 \\ 1 & 2 & 1 \\ -1 & 0 & 1 \end{pmatrix}$$

(1) 求 $\underset{\sim}{A}$ 的最小多项式 $m(\lambda)$;

(2) 对应于 $m(\lambda)$ 的因式分解,写出 $V$ 的直和分解式,并且求出在分解式中出现的每个子空间的一个基.

4. 设 $\underset{\sim}{A}$ 是域 $F$ 上线性空间 $V$ 上的线性变换,设 $\underset{\sim}{A}$ 的最小多项式 $m(\lambda)$ 在 $F[\lambda]$ 中的标准分解式是

$$m(\lambda)=p_1^{\gamma_1}(\lambda)p_2^{\gamma_2}(\lambda)\cdots p_s^{\gamma_s}(\lambda)$$

其中 $p_1(\lambda),p_2(\lambda),\cdots,p_s(\lambda)$ 是域 $F$ 上两两不同的不可约多项式，并且首项系数都为 1.

(1) 证明：$V=\mathrm{Ker}p_1^{\gamma_1}(\underset{\sim}{A})\oplus\cdots\oplus\mathrm{Ker}p_s^{\gamma_s}(\underset{\sim}{A})$；

(2) 记 $W_j=\mathrm{Ker}p_j^{\gamma_j}(\underset{\sim}{A})$，证明：$\underset{\sim}{A}|W_j$ 的最小多项式是 $p_j^{\gamma_j}(\lambda)$；

(3) 令 $\underset{\sim}{B}_j=p_j(\underset{\sim}{A}|W_j)$，证明：$\underset{\sim}{B}_j$ 是 $W_j$ 上的幂零变换，并且它的幂零指数是 $\gamma_j$.

(提示：(2) 注意 $\underset{\sim}{A}|W_j$ 与 $\underset{\sim}{A}$ 在 $W_j$ 上的作用是一样的，因此对于任意 $\alpha_j\in W_j$，有

$$p_j^{\gamma_j}(\underset{\sim}{A}|W_j)\alpha_j=p_j^{\gamma_j}(\underset{\sim}{A})\alpha_j=0$$

从而 $p_j^{\gamma_j}(\underset{\sim}{A}|W_j)=\underset{\sim}{0}$，所以 $p_j^{\gamma_j}(\lambda)$ 是 $\underset{\sim}{A}|W_j$ 的一个零化多项式. 于是 $\underset{\sim}{A}|W_j$ 的最小多项式 $m_j(\lambda)=p_j^{t_j}(\lambda)$，其中 $t_j\leqslant\gamma_j,j=1,\cdots,s$. 由第(1)小题，$V$ 是 $\underset{\sim}{A}$ 的 $s$ 个非平凡不变子空间的直和，据定理 10.3.1 得

$$m(\lambda)=[m_1(\lambda),\cdots,m_s(\lambda)]=p_1^{t_1}(\lambda)\cdots p_s^{t_s}(\lambda)$$

由唯一因式分解定理得，$t_j=\gamma_j,j=1,\cdots,s$. 所以 $m_j(\lambda)=p_j^{\gamma_j}(\lambda)$. (3) 由第(2)小题证明过程中知道，$p_j^{\gamma_j}(\underset{\sim}{A}|W_j)=\underset{\sim}{0}$. 所以 $\underset{\sim}{B}_j^{\gamma_j}=\underset{\sim}{0}$，即 $\underset{\sim}{B}_j$ 是幂零变换. 对于 $l<\gamma_j$，假如 $\underset{\sim}{B}_j^{\,l}=\underset{\sim}{0}$，则 $p_j^{\,l}(\underset{\sim}{A}|W_j)=\underset{\sim}{0}$，从而 $p_j^{\,l}(\lambda)$ 是 $\underset{\sim}{A}|W_j$ 的一个零化多项式，这与 $p_j^{\gamma_j}(\lambda)$ 是 $\underset{\sim}{A}|W_j$ 的最小多项式矛盾. 因此，$\gamma_j$ 是 $\underset{\sim}{B}_j$ 的幂零指数.)

*5. 域 $F$ 上 $n$ 维线性空间 $V$ 称为关于 $V$ 上的线性变换 $\underset{\sim}{A}$ 是**循环的**，如果 $V$ 中存在一个向量 $\xi$，使得 $\xi,\underset{\sim}{A}\xi,\cdots,\underset{\sim}{A}^{n-1}\xi$ 组成 $V$ 的一个基. 此时称 $\xi$ 是 $\underset{\sim}{A}$ 的一个循环向量. 现在设 $V$ 关于 $\underset{\sim}{A}$ 是循环的，$\xi$ 是 $\underset{\sim}{A}$ 的一个循环向量.

(1) 求 $\underset{\sim}{A}$ 在基 $\underset{\sim}{A}^{n-1}\xi,\underset{\sim}{A}^{n-2}\xi,\cdots,\underset{\sim}{A}\xi,\xi$ 下的矩阵，称这个矩阵是一个**循环块**；

(2) 求 $\underset{\sim}{A}$ 的最小多项式 $m(\lambda)$，并且比较 $\deg m(\lambda)$ 与 $\dim V$；

(3) 求 $\underset{\sim}{A}$ 的特征多项式 $f(\lambda)$.

*6. 设域 $F$ 上 $n$ 维线性空间 $V$ 的一个线性变换 $\underset{\sim}{A}$ 在基 $\alpha_1,\alpha_2,\cdots,\alpha_n$ 下的矩阵是一个循环块：

$$A=\begin{pmatrix} a_{n-1} & 1 & 0 & \cdots & 0 \\ a_{n-2} & 0 & 1 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ a_1 & 0 & 0 & \cdots & 1 \\ a_0 & 0 & 0 & \cdots & 0 \end{pmatrix}$$

(1) 证明：$V$ 关于 $\underset{\sim}{A}$ 是循环的；

(2) 求循环块 $A$ 的最小多项式和特征多项式.

7. 设域 $F$ 上 $n$ 维线性空间 $V$ 的一个线性变换 $\underset{\sim}{A}$ 的最小多项式是 $(\lambda - a)^n$.

(1) 证明：$V$ 中存在向量 $\xi$，使得 $(\underset{\sim}{A} - a\underset{\sim}{I})^{n-1}\xi, \cdots, (\underset{\sim}{A} - a\underset{\sim}{I})\xi, \xi$ 组成 $V$ 的一个基；

(2) 求 $\underset{\sim}{A}$ 在这个基下的矩阵.

8. 设 $\underset{\sim}{B}$ 是域 $F$ 上 $n$ 维线性空间 $V$ 上的幂零变换. 证明：如果 $\underset{\sim}{B}$ 的幂零指数是 $n$，则 $V$ 中存在一个向量 $\xi$，使得 $\underset{\sim}{B}^{n-1}\xi, \underset{\sim}{B}^{n-2}\xi, \cdots, \underset{\sim}{B}\xi, \xi$ 组成 $V$ 的一个基；并且求 $\underset{\sim}{B}$ 在这个基下的矩阵.

9. 设 $\underset{\sim}{B}$ 是域 $F$ 上有限维线性空间 $V$ 上的幂零变换，证明：$\underset{\sim}{B}$ 的幂零指数不超过 $\dim V$.

10. 证明：复数域上的周期矩阵（若存在正整数 $m$，使得 $A^m = I$，则 $A$ 称为周期矩阵，使 $A^m = I$ 成立的最小正整数 $m$ 称为 $A$ 的周期）一定可以对角化.

11. 求下列 Jordan 形矩阵的最小多项式：

(1) $\mathrm{diag}\{J_3(\sqrt{2}), J_1(\sqrt{2}), J_2(-1)\}$；

(2) $\mathrm{diag}\{J_3(2), J_2(-1), J_2(-1), J_1(-1), J_1(3)\}$.

*12. 证明：如果域 $F$ 上的 $n$ 级矩阵 $A$ 与 $B$ 都是可对角化的，并且 $A$ 与 $B$ 可交换，则存在域 $F$ 上一个 $n$ 级可逆矩阵 $S$，使得 $S^{-1}AS$ 与 $S^{-1}BS$ 都为对角矩阵.

（提示：设 $A$ 的全部不同的特征值为 $\lambda_1, \cdots, \lambda_s$，其中 $\lambda_i$ 的重数为 $n_i, i = 1, \cdots, s$. 于是有

$$P^{-1}AP = \mathrm{diag}\{\lambda_1 I_{n_1}, \cdots, \lambda_s I_{n_s}\} =: D$$

令 $G = P^{-1}BP$. 然后利用习题 4.6 第 10 题的结果（对域 $F$ 仍成立）以及定理 10.3.2 和推论 10.3.1）.

## §4 Jordan 标准形的存在性，唯一性，计算及应用

这一节我们来解决本章开头提出的问题：设 $\underset{\sim}{A}$ 是域 $F$ 上 $n$ 维

线性空间$V$的一个线性变换,要在空间$V$中找一个适当的基,使得$\underline{A}$在这个基下的矩阵具有较简单的形式.本章开头指出的解决这个问题的思路包含两个步骤:第一步,把空间$V$分解成$\underline{A}$的不变子空间的直和.这一步根据定理10.1.1,只要把$\underline{A}$的一个零化多项式(通常取$\underline{A}$的最小多项式$m(\lambda)$)分解因式,我们假设$m(\lambda)$在$\mathbf{F}[\lambda]$中能分解成一次因式的乘积:

$$m(\lambda) = (\lambda - \lambda_l)^{l_1}(\lambda - \lambda_2)^{l_2}\cdots(\lambda - \lambda_s)^{l_s} \tag{1}$$

于是$V$能分解成$\underline{A}$的不变子空间的直和:

$$\begin{aligned} V &= \mathrm{Ker}\, m(\underline{A}) \\ &= \mathrm{Ker}(\underline{A} - \lambda_l \underline{I})^{l_1} \oplus \cdots \oplus \mathrm{Ker}(\underline{A} - \lambda_s \underline{I})^{l_s} \end{aligned} \tag{2}$$

令 $\quad W_j = \mathrm{Ker}(\underline{A} - \lambda_j \underline{I})^{l_j}, \quad j = 1, \cdots, s$

则 $$V = W_1 \oplus W_2 \oplus \cdots \oplus W_s \tag{3}$$

分别在$W_1, \cdots, W_s$中各取一个基,它们合起来是$V$的一个基,$\underline{A}$在这个基下的矩阵$A$便是分块对角矩阵$A = \mathrm{diag}\{A_1, \cdots, A_s\}$,其中$A_j$是$\underline{A}|W_j$在$W_j$的上述基下的矩阵(见定理9.6.3).为了使$A$较简单,就应当让每个$A_j$较简单,为此需要在每个$W_j$中,选取一个合适的基.这就是解决我们的问题的第二步.本节主要篇幅就是用来进行这第二步工作.下面任意取定一个$W_j = \mathrm{Ker}(\underline{A} - \lambda_j \underline{I})^{l_j}$ $(1 \leqslant j \leqslant s)$.

令$\underline{B}_j = \underline{A}|W_j - \lambda_j \underline{I}$,其中$\underline{I}$是$W_j$上的恒等变换.设$\underline{B}_j$和$\underline{A}|W_j$在$W_j$的一个基下的矩阵分别为$B_j$和$A_j$,据线性变换与矩阵的对应关系得出,$B_j = A_j - \lambda_j I$.即$A_j = B_j + \lambda_j I$.这表明,只要能找到$W_j$的一个基,使$\underline{B}_j$的矩阵$B_j$较简单,那么$\underline{A}|W_j$的矩阵$A_j$也就较简单了,因为$A_j$与$B_j$只相差一个纯量矩阵$\lambda_j I$.所以我们来讨论$\underline{B}_j$.任取$\alpha \in W_j = \mathrm{Ker}(\underline{A} - \lambda_j \underline{I})^{l_j}$,

$$\underline{B}_j^{l_j}\alpha = (\underline{A}|W_j - \lambda_j \underline{I})^{l_j}\alpha = (\underline{A} - \lambda_j \underline{I})^{l_j}\alpha = 0 \tag{4}$$

因此$\underline{B}_j^{l_j}$是$W_j$上的零变换.从而$\underline{B}_j$是$W_j$上的幂零变换.进一步可证$\underline{B}_j$的幂零指数是$l_j$(见习题10.3第4题的提示).这样第二步要

做的工作是，对于 $W_j$ 上的幂零指数为 $l_j$ 的幂零变换 $\underline{B}_j$，要在 $W_j$ 中找一个合适的基，使得 $\underline{B}_j$ 的矩阵较简单. 为了书写简便起见，我们去掉下标 $j$. 即

问题(一)：设 $W$ 是域 $F$ 上的 $r$ 维线性空间，$\underline{B}$ 是 $W$ 上的幂零变换，其幂零指数为 $l$. 要在 $W$ 中找一个合适的基，使得 $\underline{B}$ 的矩阵较简单.

我们指出，$\underline{B}$ 的幂零指数 $l$ 不超过 $W$ 的维数 $r$. 理由如下：因为 $\underline{B}^{l-1} \neq \underline{0}$，所以 $W$ 中存在向量 $\xi$，使得 $\underline{B}^{l-1}\xi \neq 0$. 又有

$$\underline{B}^{l}\xi = \underline{0}\xi = 0$$

据习题 9.2 的第 5 题得，$\xi, \underline{B}\xi, \cdots, \underline{B}^{l-1}\xi$ 线性无关. 从而

$$l \leqslant \dim W$$

如果 $\underline{B}$ 的幂零指数 $l = \dim W = r$，则上一段得到的向量组 $\underline{B}^{r-1}\xi$，$\underline{B}^{r-2}\xi, \cdots, \underline{B}\xi, \xi$ 是 $W$ 的一个基，容易算出，$\underline{B}$ 在这个基下的矩阵 $B$ 是一个主对角元为 0 的 $r$ 级 Jordan 块 $J_r(0)$.

这样，问题的困难之处在于 $\underline{B}$ 的幂零指数 $l$ 小于 $\dim W$ 的情形. 虽然我们能在 $W$ 里找到向量 $\xi$，使得 $\xi, \underline{B}\xi, \cdots, \underline{B}^{l-1}\xi$ 线性无关，但是由于 $l < \dim W$，因此它们不组成 $W$ 的一个基. 注意到它们生成的子空间 $\langle \xi, \underline{B}\xi, \cdots, \underline{B}^{l-1}\xi \rangle$ 是 $\underline{B}$ 的不变子空间，并且 $\underline{B}$ 在这个子空间上的限制在基 $\underline{B}^{l-1}\xi, \cdots, \underline{B}\xi, \xi$ 下的矩阵是一个 Jordan 块 $J_l(0)$. 这促使我们先分析一下这种类型的 $\underline{B}$ 不变子空间.

**定义 1** 对于 $W$ 中一个向量 $\xi$，如果存在正整数 $t$，使得

$$\underline{B}^{t-1}\xi \neq 0, \quad 而\ \underline{B}^{t}\xi = 0$$

则子空间 $\langle \xi, \underline{B}\xi, \cdots, \underline{B}^{t-1}\xi \rangle$ 称为由 $\xi$ 生成的 $\underline{B}$— **循环子空间**，记作 $Z_t(\xi; \underline{B})$. $\underline{B}^{t-1}\xi, \cdots, \underline{B}\xi, \xi$ 是 $Z_t(\xi; \underline{B})$ 的一个基，从而 $Z_t(\xi; \underline{B})$ 的维数是 $t$.

容易看出，$Z_t(\xi; \underline{B})$ 是 $\underline{B}$ 的不变子空间，并且 $\underline{B} | Z_t(\xi; \underline{B})$ 在上述基下的矩阵是一个 $t$ 级 Jordan 块 $J_t(0)$.

因此，只要我们能把 $W$ 分解成若干个 $\underline{B}$— 循环子空间的直

和，那么 $\underset{\sim}{B}$ 的矩阵就是一个 Jordan 形矩阵，而 Jordan 形矩阵是比较简单的. 这样为了解决问题(一) 就只要解决下述的问题(二)：

问题(二)：设 $W$ 是域 $F$ 上的 $r$ 维线性空间，$\underset{\sim}{B}$ 是 $W$ 上的幂零变换，其幂零指数为 $l$. 能否把 $W$ 分解成一些 $\underset{\sim}{B}$— 循环子空间的直和？

* 我们需要引进几个概念：

**定义** 2　设 $V$ 是域 $F$ 上 $n$ 维线性空间，$\underset{\sim}{A}$ 是 $V$ 上的一个线性变换. $\alpha_1,\alpha_2,\cdots,\alpha_s$ 是 $V$ 的一个向量组. 对于 $\alpha\in V$，如果存在 $f_1(\lambda)$，$f_2(\lambda),\cdots,f_s(\lambda)\in F[\lambda]$，使

$$\alpha=f_1(\underset{\sim}{A})\alpha_1+f_2(\underset{\sim}{A})\alpha_2+\cdots+f_s(\underset{\sim}{A})\alpha_s \tag{5}$$

则称 $\alpha$ 可以由 $\alpha_1,\alpha_2,\cdots,\alpha_s F[\lambda]$— **线性表出**.

如果 $V$ 中每个向量 $\alpha$ 都可以由向量组 $\alpha_1,\alpha_2,\cdots,\alpha_s F[\lambda]$— 线性表出，则称 $\alpha_1,\alpha_2,\cdots,\alpha_s$ 是 $V$ 的一个 $\underset{\sim}{A}$— **生成元系**.

如果 $V$ 的一个 $\underset{\sim}{A}$ 生成元系 $\alpha_1,\alpha_2,\cdots,\alpha_s$ 中每个 $\alpha_i$ 都不能由其余的向量 $F[\lambda]$— 线性表出，则称 $\alpha_1,\alpha_2,\cdots,\alpha_s$ 是 $V$ 的一个**最小 $\underset{\sim}{A}$— 生成元系**.

不难看出，$V$ 的任意一个基 $\eta_1,\eta_2,\cdots,\eta_n$ 是 $V$ 的一个 $\underset{\sim}{A}$— 生成元系.

$V$ 的最小 $\underset{\sim}{A}$— 生成元系一定存在. 理由如下：任取 $V$ 的一个 $\underset{\sim}{A}$— 生成元系 $\alpha_1,\alpha_2,\cdots,\alpha_m$，如果它不是最小的，则存在一个 $\alpha_j$ 可以由其余向量 $F[\lambda]$— 线性表出. 去掉这个 $\alpha_j$，剩下的 $m-1$ 个向量仍是 $V$ 的 $A$— 生成元系. 由此看出，$V$ 的最小 $\underset{\sim}{A}$— 生成元系存在.

例如，考虑定义 1 中的 $Z_t(\xi;\underset{\sim}{B})$，它的任一向量 $\alpha$ 可以表示成

$$\begin{aligned}\alpha&=k_0\xi+k_1\underset{\sim}{B}\xi+\cdots+k_{t-1}\underset{\sim}{B}^{t-1}\xi\\&=(k_0\underset{\sim}{I}+k_1\underset{\sim}{B}++k_{t-1}\underset{\sim}{B}^{t-1})\xi\end{aligned}$$

令 $h(\lambda)=k_0+k_1\lambda+\cdots+k_{t-1}\lambda^{t-1}$，则 $\alpha=h(\underset{\sim}{B})\xi$. 这表明 $\xi$ 是 $Z_t(\xi;\underset{\sim}{B})$ 的一个 $\underset{\sim}{B}$— 生成元系. 显然它是最小的.

上述说明 $Z_t(\xi;\underset{\sim}{B})\subset\{h(\underset{\sim}{B})\xi\mid h(\lambda)\in F[\lambda]\}$. 反之，对于任一

$h(\lambda) \in F[\lambda]$，由于 $\xi \in Z_t(\xi;\underline{B})$ 且 $Z_t(\xi;\underline{B})$ 是 $\underline{B}$ 的不变子空间，因此 $h(\underline{B})\xi \in Z_t(\xi;\underline{B})$. 这样我们证明了：

$$Z_t(\xi;\underline{B}) = \{h(\underline{B})\xi \mid h(\lambda) \in F[\lambda]\}$$

现在我们来回答问题(二)：

**定理** 10.4.1　设 $W$ 是域 $F$ 上有限维线性空间，$\underline{B}$ 是 $W$ 上的一个幂零变换，则 $W$ 能分解成一些 $\underline{B}$— 循环子空间的直和.

* **证明**　对 $W$ 的最小 $\underline{B}$— 生成元系所含向量的数目 $s$ 作数学归纳法.

$s = 1$ 时，设 $\xi$ 是 $W$ 的一个最小 $B$— 生成元系，则 $W = \{h(\underline{B})\xi \mid h(\lambda) \in F[\lambda]\}$. 设 $\underline{B}^{t-1}\xi \neq 0$，而 $\underline{B}^t\xi = 0$，则 $Z_t(\xi;\underline{B}) = \{h(\underline{B})\xi \mid h(\lambda) \in F[\lambda]\}$. 从而

$$W = Z_t(\xi;\underline{B})$$

假设对于具有幂零变换 $\underline{B}$ 的有限维线性空间，当它的最小 $\underline{B}$— 生成元系所含向量数目等于 $s-1$ 时，命题成立. 现在来看 $W$，它的一个最小 $\underline{B}$— 生成元系是 $\alpha_1, \alpha_2, \cdots, \alpha_s$.

为了利用归纳假设，我们需要把 $W$ 分解成一个 $\underline{B}$— 循环子空间与一个 $\underline{B}$— 子空间 $U$ 的直和，其中 $U$ 的一个最小 $\underline{B}$— 生成元系含 $s-1$ 个向量. 难点在于如何选取这个 $\underline{B}$— 循环子空间的生成元，使得上述和为直和. 证明分成几个步骤.

第一步. 找出一个 $\underline{B}$— 循环子空间的生成元 $\xi_1$.

对于给定的 $i(1 \leqslant i \leqslant s)$，令

$$\begin{aligned} H_i = \{h_i(\lambda) \in F[\lambda] \mid & h_1(\underline{B})\alpha_1 + \cdots \\ & + h_i(\underline{B})\alpha_i + \cdots + h_s(\underline{B})\alpha_s = 0\} \end{aligned} \qquad (6)$$

显然，$H_i$ 对于减法封闭；并且对于 $h_i(\lambda) \in H_i, f(\lambda) \in F[\lambda]$，有 $f(\lambda)h_i(\lambda) \in H_i$. 在 $H_i$ 的非零多项式中取一个次数最低的且首项系数为 1 的多项式 $m_i(\lambda)$，我们断言 $H_i$ 中任一多项式 $h_i(\lambda)$ 是 $m_i(\lambda)$ 的倍式. 理由如下：作带余除法

$$h_i(\lambda) = p_i(\lambda)m_i(\lambda) + r_i(\lambda), \quad \deg r_i(\lambda) < \deg m_i(\lambda)$$

于是 $$r_i(\lambda)=h_i(\lambda)-p_i(\lambda)m_i(\lambda)\in H_i$$
从 $m_i(\lambda)$ 的取法得出，$r_i(\lambda)=0$. 所以 $h_i(\lambda)=p_i(\lambda)m_i(\lambda)$.

因为 $\underline{B}$ 是 $W$ 的幂零变换，所以可设 $\underline{B}^{t_i-1}\alpha_i\neq 0$，而 $\underline{B}^{t_i}\alpha_i=0$. 于是有

$$0\alpha_1+\cdots+0\alpha_{i-1}+\underline{B}^{t_i}\alpha_i+0\alpha_{i+1}+\cdots+0\alpha_s=0$$

由此得出，$\lambda^{t_i}\in H_i$. 因此 $m_i(\lambda)\mid\lambda^{t_i}$. 从而可设

$$m_i(\lambda)=\lambda^{\mu_i},\quad \mu_i\leqslant t_i$$

于是 $$H_i=\{p(\lambda)\lambda^{\mu_i}\mid p(\lambda)\in F[\lambda]\}\tag{7}$$

适当排列 $\alpha_1,\alpha_2,\cdots,\alpha_s$ 的次序，使得

$$\mu_1\leqslant\mu_2\leqslant\cdots\leqslant\mu_s$$

因为 $\lambda^{\mu_1}\in H_1$，据(6) 式得，存在 $h_2(\lambda),\cdots,h_s(\lambda)\in F[\lambda]$，使得

$$\underline{B}^{\mu_1}\alpha_1+h_2(\underline{B})\alpha_2+\cdots+h_s(\underline{B})\alpha_s=0\tag{8}$$

(8)式表明 $h_j(\lambda)\in H_j,\ j=2,\cdots,s$. 因此从(7) 式得出，$\lambda^{\mu_j}\mid h_j(\lambda)$. 又由于 $\mu_1\leqslant\mu_j$，所以

$$h_j(\lambda)=p_j(\lambda)\lambda^{\mu_1},\quad j=2,\cdots,s$$

由此得出 $$h_j(\underline{B})=p_j(\underline{B})\underline{B}^{\mu_1},\quad j=2,\cdots,s$$

把它们代入(8)式得

$$\underline{B}^{\mu_1}\alpha_1+p_2(\underline{B})\underline{B}^{\mu_1}\alpha_2+\cdots+p_s(\underline{B})\underline{B}^{\mu_1}\alpha_s=0$$

从而有 $$\underline{B}^{\mu_1}[\alpha_1+p_2(\underline{B})\alpha_2+\cdots+p_s(\underline{B})\alpha_s]=0\tag{9}$$

令 $$\xi_1=\alpha_1+p_2(\underline{B})\alpha_2+\cdots+p_s(\underline{B})\alpha_s\tag{10}$$

由于 $\alpha_1,\alpha_2,\cdots,\alpha_s$ 是 $W$ 的最小 $\underline{B}$— 生成元系，所以 $\xi_1\neq 0$. 设 $\underline{B}^{k_1-1}\xi_1\neq 0$，而 $\underline{B}^{k_1}\xi_1=0$，从(9) 式得 $\underline{B}^{\mu_1}\xi_1=0$. 因此 $\mu_1\geqslant k_1$. 又从(10) 式得

$$0=\underline{B}^{k_1}\xi_1=\underline{B}^{k_1}\alpha_1+\underline{B}^{k_1}p_2(\underline{B})\alpha_2++\underline{B}^{k_1}p_s(\underline{B})\alpha_s\tag{11}$$

(11)式表明 $\lambda^{k_1}\in H_1$，从而 $\lambda^{\mu_1}\mid\lambda^{k_1}$. 由此得出，$\mu_1\leqslant k_1$. 这证明了：$\mu_1=k_1$. 从而 $\xi_1$ 生成 $\mu_1$ 维 $\underline{B}$— 循环子空间 $Z_{\mu_1}(\xi_1;\underline{B})$. 这样我们终于找到了一个 $\underline{B}$— 循环子空间的生成元 $\xi_1$.

第二步. 找一个 $\underline{B}$— 子空间 $U$. 令

$$U = \{f_2(\underline{B})\alpha_2 + \cdots + f_s(\underline{B})\alpha_s \mid f_j(\lambda) \in F[\lambda], j = 2, \cdots, s\}$$

易验证 $U$ 是 $W$ 的线性子空间,并且 $U$ 是 $\underline{B}$— 子空间. 从而 $\underline{B}|U$ 是 $U$ 的幂零变换. 容易看出 $\alpha_2, \cdots, \alpha_s$ 是 $U$ 的最小 $\underline{B}|U$— 生成元系. 于是可以对 $U$ 用归纳假设得,$U$ 可以分解成一些 $\underline{B}|U$— 循环子空间的直和:

$$U = Z_{k_2}(\xi_2;\underline{B}|U) \oplus \cdots \oplus Z_{k_m}(\xi_m;\underline{B}|U)$$

显然有 $Z_{k_j}(\xi_j;\underline{B}|U) = Z_{k_j}(\xi_j;\underline{B}), j = 2, \cdots, m$. 因此

$$U = Z_{k_2}(\xi_2;\underline{B}) \oplus \cdots \oplus Z_{k_m}(\xi_m;\underline{B}) \tag{12}$$

第三步. 证明 $W = Z_{\mu_1}(\xi_1;\underline{B}) \oplus U$.

从(10)式看出, $\alpha_1 \in Z_{\mu_1}(\xi_1;\underline{B}) + U$,因此

$$W = Z_{\mu_1}(\xi_1;\underline{B}) + U \tag{13}$$

任取 $\eta \in Z_{\mu_1}(\xi_1;\underline{B}) \cap U$,则存在 $f_i(\lambda) \in F[\lambda], i = 1, \cdots, s$, 使得

$$\eta = f_1(\underline{B})\xi_1, \quad \eta = f_2(\underline{B})\alpha_2 + \cdots + f_s(\underline{B})\alpha_s \tag{14}$$

于是从(10)和(14)式得

$$f_1(\underline{B})\alpha_1 + [f_1(\underline{B})p_2(\underline{B}) - f_2(\underline{B})]\alpha_2 + \cdots$$
$$+ [f_1(\underline{B})p_s(\underline{B}) - f_s(\underline{B})]\alpha_s = 0$$

此式表明 $f_1(\lambda) \in H_1$,从而 $f_1(\lambda) = q(\lambda)\lambda^{\mu_1}$. 因此

$$\eta = f_1(\underline{B})\xi_1 = q(\underline{B})\underline{B}^{\mu_1}\xi_1 = q(\underline{B})0 = 0$$

这证明了 $Z_{\mu_1}(\xi_1;\underline{B}) \cap U = 0$. 所以

$$\begin{aligned} W &= Z_{\mu_1}(\xi_1;\underline{B}) \oplus U \\ &= Z_{\mu_1}(\xi_1;\underline{B}) \oplus Z_{k_2}(\xi_2;\underline{B}) \oplus \cdots \oplus Z_{k_m}(\xi_m;\underline{B}) \end{aligned}$$

据数学归纳法原理,对一切正整数 $s$,命题成立. ▌

有了定理 10.4.1,我们便可以了解幂零变换的结构了,即我们有下述定理.

**定理** 10.4.2　设 $W$ 是域 $F$ 上 $r$ 维线性空间,$\underline{B}$ 是 $W$ 上的一个幂零变换,其幂零指数为 $l$. 则 $W$ 中存在一个基,使得 $\underline{B}$ 在此基下

的矩阵$B$为Jordan形矩阵，其中每个Jordan块的主对角元都是0，并且$t$级Jordan块的数目$N(t)$为

$$N(t) = \text{rank}\underline{B}^{t+1} + \text{rank}\underline{B}^{t-1} - 2\text{rank}\underline{B}^{t} \tag{15}$$

$B$中Jordan块的总数等于$\underline{B}$的特征子空间$V_0$的维数. 这个Jordan形矩阵$B$称为幂零变换$\underline{B}$的Jordan **标准形**. 除去Jordan块的排列次序外，$\underline{B}$的Jordan标准形是唯一的.

**证明** 据定理10.4.1，$W$能分解成

$$W = Z_{k_1}(\xi_1;\underline{B}) \oplus \cdots \oplus Z_{k_s}(\xi_s;\underline{B})$$

在每个$Z_{k_i}(\xi_i;\underline{B})$中取一个基：$\underline{B}^{k_i-1}\xi_i, \cdots, \underline{B}\xi_i, \xi_i$，把它们合起来便成为$W$的一个基. 由于$\underline{B}|Z_{k_i}(\xi_i;\underline{B})$在$Z_{k_i}(\xi_i;\underline{B})$的上述基下的矩阵是一个Jordan块$J_{k_i}(0)$，从而$\underline{B}$在$W$的上述基下的矩阵$B$为Jordan形矩阵：

$$B = \text{diag}\{J_{k_1}(0), \cdots, J_{k_s}(0)\}$$

现在来计算$B$中$t$级Jordan块$J_t(0)$的数目$N(t)$. 由于$\underline{B}$的幂零指数为$l$，所以$k_i \leqslant l, i = 1, \cdots, s$. 这说明当$t > l$时，$N(t) = 0$. 下面设$1 \leqslant t \leqslant l$.

首先，给定一个正整数$m$，我们来计算$\text{rank}B^m$. 当$m \geqslant l$时有

$$J_{k_i}(0)^m = 0, \quad i = 1, \cdots, s$$

从而$B^m = 0$. 此时有$\text{rank}B^m = 0$. 下面设$m < l$. 如果$m \geqslant t$，则$J_t(0)^m = 0$，从而$\text{rank}J_t(0)^m = 0$. 如果$m < t$，则

$$J_t(0)^m = \begin{pmatrix} \overbrace{0 & \cdots & 0}^{m} & 1 & 0 & \cdots & 0 & 0 \\ 0 & \cdots & 0 & 0 & 1 & \cdots & 0 & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & \cdots & 0 & 0 & 0 & \cdots & 0 & 1 \\ 0 & \cdots & 0 & 0 & 0 & \cdots & 0 & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & \cdots & 0 & 0 & 0 & \cdots & 0 & 0 \end{pmatrix}$$

从而 $\operatorname{rank} J_t(0)^m = t - m$. 因此当 $m < l$ 时，有

$$
\begin{aligned}
&\operatorname{rank} B^m \\
&= \sum_{t=1}^{l} [\operatorname{rank} J_t(0)^m] N(t) = \sum_{t=m+1}^{l} (t-m) N(t) \\
&= N(m+1) + 2N(m+2) + \cdots + (l-m)N(l) \qquad (16)
\end{aligned}
$$

当 $1 \leqslant m < l - 1$ 时，从(16)式得

$$
\begin{aligned}
&\operatorname{rank} B^m - \operatorname{rank} B^{m+1} \\
&= N(m+1) + N(m+2) + \cdots + N(l) \qquad (17)
\end{aligned}
$$

从(16)式得 $\operatorname{rank} B^{l-1} = N(l)$，因此(17)式对于 $m = l - 1$ 时也成立.

当 $2 \leqslant m \leqslant l$ 时，从(17)式得

$$
\begin{aligned}
&\operatorname{rank} B^{m-1} - \operatorname{rank} B^m \\
&= N(m) + N(m+1) + \cdots + N(l) \qquad (18)
\end{aligned}
$$

由于 $\operatorname{rank} J_t(0) = t - 1$，所以

$$
\begin{aligned}
\operatorname{rank} B &= r - [N(1) + N(2) + \cdots + N(l)] \\
&= \operatorname{rank} I - [N(1) + N(2) + \cdots + N(l)] \qquad (19)
\end{aligned}
$$

这说明(18)式对 $m = 1$ 时也成立. 因此当 $1 \leqslant m < l$ 时，从(18)式减去(17)式得

$$
N(m) = \operatorname{rank} B^{m+1} + \operatorname{rank} B^{m-1} - 2\operatorname{rank} B^m
$$

显然，上式对于 $m \geqslant l$ 也成立. 从而对一切 $m \geqslant 1$ 有

$$
N(m) = \operatorname{rank} \underline{B}^{m+1} + \operatorname{rank} \underline{B}^{m-1} - 2\operatorname{rank} \underline{B}^m \qquad (20)
$$

现在来求 $B$ 中 Jordan 块的总数：从(19)式得到

$$
\begin{aligned}
&N(1) + N(2) + \cdots + N(l) = \operatorname{rank} I - \operatorname{rank} B \\
&= \dim W - \dim(\operatorname{Im} \underline{B}) = \dim(\operatorname{Ker} \underline{B}) \qquad (21)
\end{aligned}
$$

由于幂零变换 $B$ 的特征值为 0，所以 $\underline{B}$ 的特征子空间 $V_0$ 为

$$
V_0 = \{\alpha \in W \mid \underline{B}\alpha = 0\} = \operatorname{Ker} \underline{B}
$$

于是(21)式表明：$B$ 中 Jordan 块的总数等于 $\underline{B}$ 的特征子空间 $V_0$ 的维数.

从(20)式看出，对于 $m\geqslant 1$，$\underline{B}$ 中 $m$ 级 Jordan 块 $J_m(0)$ 的数目由 $\underline{B}$ 决定，因此 $\underline{B}$ 的 Jordan 标准形除了 Jordan 块的排列次序外，是唯一的. ▌

现在我们把关于幂零变换的定理 10.4.2 应用到一般的线性变换的结构的研究上，可以得到

**定理** 10.4.3 设 $\underline{A}$ 是域 $F$ 上 $n$ 维线性空间 $V$ 的一个线性变换. 如果 $\underline{A}$ 的最小多项式 $m(\lambda)$ 在 $F[\lambda]$ 中能分解成一次因式的乘积

$$m(\lambda)=(\lambda-\lambda_1)^{l_1}(\lambda-\lambda_2)^{l_2}\cdots(\lambda-\lambda_s)^{l_s} \tag{1}$$

则 $V$ 中存在一个基，使得 $\underline{A}$ 在这个基下的矩阵 $A$ 为 Jordan 形矩阵，其中主对角元为 $\lambda_j$ 的 $t$ 级 Jordan 块 $J_t(\lambda_j)$ 的数目 $N(t;\lambda_j)$ 为

$$\begin{aligned}N(t;\lambda_j)=&\operatorname{rank}(\underline{A}-\lambda_j\underline{I})^{t+1}+\operatorname{rank}(\underline{A}-\lambda_j\underline{I})^{t-1}\\&-2\operatorname{rank}(\underline{A}-\lambda_j\underline{I})^{t}\end{aligned} \tag{22}$$

主对角元为 $\lambda_j$ 的 Jordan 块的总数 $N(\lambda_j)$ 为

$$N(\lambda_j)=\dim V-\operatorname{rank}(\underline{A}-\lambda_j\underline{I})\quad j=1,\cdots,s \tag{23}$$

这个 Jordan 形矩阵 $A$ 称为 $\underline{A}$ **的** Jordan **标准形**. 除去 Jordan 块的排列次序外，$\underline{A}$ 的 Jordan 标准形是唯一的.

**证明** 由(1)看出，$\lambda_1,\cdots,\lambda_s$ 是 $\underline{A}$ 的所有不同的特征值. 由(1)得出，$V$ 有直和分解式(2). 令

$$W_j=\operatorname{Ker}(\underline{A}-\lambda_j\underline{I})^{l_j},\quad \underline{B}_j=\underline{A}|W_j-\lambda_j\underline{I},\quad j=1,\cdots,s$$

则 $\underline{B}_j$ 是 $W_j$ 上的幂零变换，并且幂零指数为 $l_j$. 于是据定理 10.4.2，$\underline{B}_j$ 在 $W_j$ 的适当一个基下的矩阵 $B_j$ 为 Jordan 形矩阵，它的每个 Jordan 块的主对角元都为 0. 从而 $\underline{A}|W_j$ 在 $W_j$ 的这个基下的矩阵 $A_j=B_j+\lambda_jI$ 也是 Jordan 形矩阵，$A_j$ 的每个 Jordan 块的主对角元都为 $\lambda_j$. 把 $W_j$，$j=1,\cdots,s$，的基合起来便成为 $V$ 的一个基，$\underline{A}$ 在 $V$ 的这个基下的矩阵

$$A=\operatorname{diag}\{A_1,\cdots,A_s\}$$

因此 $A$ 是 Jordan 形矩阵. $A$ 中主对角元为 $\lambda_j$ 的 $t$ 级 Jordan 块 $J_t(\lambda_j)$

的数目 $N(t;\lambda_j)$ 等于 $B_j$ 中 $t$ 级 Jordan 块的数目，因此

$$
\begin{aligned}
N(t;\lambda_j) &= \operatorname{rank}\underline{B}_j^{t+1} + \operatorname{rank}\underline{B}_j^{t-1} - 2\operatorname{rank}\underline{B}_j^{t} \\
&= \operatorname{rank}(\underline{A}|W_j - \lambda_j\underline{I})^{t+1} + \operatorname{rank}(\underline{A}|W_j - \lambda_j\underline{I})^{t-1} \\
&\quad - 2\operatorname{rank}(\underline{A}|W_j - \lambda_j\underline{I})^{t} \\
&= 2\dim\operatorname{Ker}(\underline{A}|W_j - \lambda_j\underline{I})^{t} - \dim\operatorname{Ker}(\underline{A}|W_j - \lambda_j\underline{I})^{t+1} \\
&\quad - \dim\operatorname{Ker}(\underline{A}|W_j - \lambda_j\underline{I})^{t-1}
\end{aligned}
$$

由于当 $i \leqslant l_j$ 时，我们有

$$
\begin{aligned}
&\alpha \in \operatorname{Ker}(\underline{A} - \lambda_j\underline{I})^{i} \\
&\Leftrightarrow (\underline{A} - \lambda_j\underline{I})^{i}\alpha = 0 \Leftrightarrow \alpha \in W_j \text{ 并且 } (\underline{A} - \lambda_j\underline{I})^{i}\alpha = 0 \\
&\Leftrightarrow (\underline{A}|W_j - \lambda_j\underline{I})^{i}\alpha = 0 \Leftrightarrow \alpha \in \operatorname{Ker}(\underline{A}|W_j - \lambda_j\underline{I})^{i}
\end{aligned}
$$

因此 $$\operatorname{Ker}(\underline{A} - \lambda_j\underline{I})^{i} = \operatorname{Ker}(\underline{A}|W_j - \lambda_j\underline{I})^{i} \tag{24}$$

我们指出，当 $i = l_j + u$（$u$ 是正整数）时，(24) 式仍然成立. 理由如下：从(1) 式得

$$m(\lambda)(\lambda - \lambda_j)^{u} = (\lambda - \lambda_1)^{l_1}\cdots(\lambda - \lambda_j)^{l_j+u}\cdots(\lambda - \lambda_s)^{l_s}$$

于是 $V$ 可以分解成

$$
\begin{aligned}
V = &\operatorname{Ker}(\underline{A} - \lambda_1\underline{I})^{l_1} \oplus \cdots \oplus \operatorname{Ker}(\underline{A} - \lambda_j\underline{I})^{l_j+u} \oplus \cdots \\
&\oplus \operatorname{Ker}(\underline{A} - \lambda_s\underline{I})^{l_s}
\end{aligned}
\tag{25}
$$

任取 $\beta \in \operatorname{Ker}(\underline{A} - \lambda_j\underline{I})^{l_j+u}$，据(2) 式，$\beta$ 可表示成

$$\beta = \beta_1 + \cdots + \beta_j + \cdots + \beta_s$$

其中 $\beta_i \in \operatorname{Ker}(\underline{A} - \lambda_i\underline{I})^{l_i}, i = 1, \cdots, s$. 于是得出

$$0 = \beta_1 + \cdots + (\beta_j - \beta) + \cdots + \beta_s \tag{26}$$

由于 $\operatorname{Ker}(\underline{A} - \lambda^j\underline{I})^{l_j} \subset \operatorname{Ker}(\underline{A} - \lambda_j\underline{I})^{l_j+u}$，因此

$$\beta_j - \beta \in \operatorname{Ker}(\underline{A} - \lambda_j\underline{I})^{l_j+u}$$

从(25)式和(26)式（利用零向量的表法唯一）得

$$\beta_j - \beta = 0$$

从而 $\beta = \beta_j \in \operatorname{Ker}(\underline{A} - \lambda_j\underline{I})^{l_j}$. 由此得出

$$\operatorname{Ker}(\underline{A}-\lambda_j\underline{I})^{l_j+u}=\operatorname{Ker}(\underline{A}-\lambda_j\underline{I})^{l_j} \tag{27}$$

现在任取 $\eta\in\operatorname{Ker}(\underline{A}|W_j-\lambda_j\underline{I})^{l_j+u}$,则 $\eta\in W_j$. 由于

$$W_j=\operatorname{Ker}(\underline{A}-\lambda_j\underline{I})^{l_j}=\operatorname{Ker}(\underline{A}|W_j-\lambda_j\underline{I})^{l_j}$$

因此 $\eta\in\operatorname{Ker}(\underline{A}|W_j-\lambda_j\underline{I})^{l_j}$. 由此得出

$$\operatorname{Ker}(\underline{A}|W_j-\lambda_j\underline{I})^{l_j+u}=\operatorname{Ker}(\underline{A}|W_j-\lambda_j\underline{I})^{l_j}$$

从而结合(27)式便得出

$$\operatorname{Ker}(\underline{A}-\lambda_j\underline{I})^{l_j+u}=\operatorname{Ker}(\underline{A}|W_j-\lambda_j\underline{I})^{l_j+u}$$

所以

$$\begin{aligned}N(t;\lambda_j)&=2\dim\operatorname{Ker}(\underline{A}-\lambda_j\underline{I})^t-\dim\operatorname{Ker}(\underline{A}-\lambda_j\underline{I})^{t+1}\\&\quad-\dim\operatorname{Ker}(\underline{A}-\lambda_j\underline{I})^{t-1}\\&=\operatorname{rank}(\underline{A}-\lambda_j\underline{I})^{t+1}+\operatorname{rank}(\underline{A}-\lambda_j\underline{I})^{t-1}\\&\quad-2\operatorname{rank}(\underline{A}-\lambda_j\underline{I})^t\end{aligned} \tag{22}$$

$\underline{A}$ 中主对角元为 $\lambda_j$ 的 Jordan 块总数 $N(\lambda_j)$ 等于 $B_j$ 中 Jordan 块总数,从而等于 $\underline{B}_j$ 的特征子空间 $V_0$ 的维数. 因为

$$V_0=\operatorname{Ker}\underline{B}_j=\operatorname{Ker}(\underline{A}|W_j-\lambda_j\underline{I})=\operatorname{Ker}(\underline{A}-\lambda_j\underline{I})$$

所以
$$\begin{aligned}N(\lambda_j)&=\dim V_0=\dim\operatorname{Ker}(\underline{A}-\lambda_j\underline{I})\\&=\dim V-\operatorname{rank}(\underline{A}-\lambda_j\underline{I})\end{aligned} \tag{23}$$

$\underline{A}$ 的 Jordan 标准形的主对角线上元素都是 $\underline{A}$ 的特征值,从(22) 式看出,对于 $\underline{A}$ 的每一个特征值 $\lambda_j$,主对角元为 $\lambda_j$ 的 $t$ 级 Jordan 块的数目 $N(t;\lambda_j)$ 由 $\underline{A}$ 决定. 因此除去 Jordan 块的排列次序外,$\underline{A}$ 的 Jordan 标准形是唯一的. ▌

用矩阵的语言来叙述上述结果就是:

**定理** 10.4.4 设 $A$ 是域 $F$ 上的 $n$ 级矩阵,如果 $A$ 的最小多项式 $m(\lambda)$ 在 $F[\lambda]$ 中能分解成一次因式的乘积

$$m(\lambda)=(\lambda-\lambda_1)^{l_1}(\lambda-\lambda_2)^{l_2}\cdots(\lambda-\lambda_s)^{l_s}$$

则 $A$ 与一个 Jordan 形矩阵相似,除了 Jordan 块的排列次序外,这个 Jordan 形矩阵由 $A$ 唯一决定,它称为 $A$ 的 Jordan 标准形,其中

Jordan 块 $J_t(\lambda_j)$ 的数目 $N(t;\lambda_j)$ 为

$$N(t;\lambda_j)=\operatorname{rank}(A-\lambda_j I)^{t+1}+\operatorname{rank}(A-\lambda_j I)^{t-1} - 2\operatorname{rank}(A-\lambda_j I)^{t}$$

主对角元为 $\lambda_j$ 的 Jordan 块的总数 $N(\lambda_j)$ 为

$$N(\lambda_j)=n-\operatorname{rank}(A-\lambda_j I),\quad j=1,\cdots,s$$ ▌

由于复数域上的每个一元多项式都可以分解成一次因式的乘积，所以复数域上有限维线性空间的每一个线性变换都有 Jordan 标准形，或者说，复数域上每一个方阵都有 Jordan 标准形.

* **定义 3** 设 $F$ 是一个域，如果域 $F$ 上每一个一元多项式在 $F$ 中有根，则称 $F$ 是一个**代数封闭域**.

* 从定义 3 立即得出，域 $F$ 是一个代数封闭域当且仅当 $F[x]$ 中每一个多项式都能分解成一次因式的乘积.

* 从而，代数封闭域 $F$ 上的有限维线性空间的每一个线性变换都有 Jordan 标准形. 换句话说，代数封闭域 $F$ 上每一个方阵都有 Jordan 标准形.

根据上面的讨论，对于域 $F$ 上 $n$ 维线性空间 $V$ 的线性变换 $\underline{A}$，判断它有没有 Jordan 标准形，以及如果有 Jordan 标准形，把它求出来的步骤如下：

第一步. 写出 $\underline{A}$ 在 $V$ 的一个基下的矩阵 $A$（这一步常常是已知条件里给出的）；

第二步. 求 $A$ 的特征多项式 $f(\lambda)=|\lambda I-A|$. 如果 $f(\lambda)$ 在 $F[\lambda]$ 中能分解成一次因式的乘积，则 $\underline{A}$ 的最小多项式 $m(\lambda)$ 也能分解成一次因式的乘积，从而 $\underline{A}$ 有 Jordan 标准形；如果 $f(\lambda)$ 在 $F[\lambda]$ 中不能分解成一次因式的乘积，则 $\underline{A}$ 没有 Jordan 标准形（假如 $\underline{A}$ 有 Jordan 标准形 $J$，由于 $J$ 是上三角矩阵，因此 $J$ 的特征多项式能分解成一次因式的乘积. 又由于 $J$ 与 $A$ 相似，它们有相同的特征多项式，从而 $f(\lambda)$ 能分解成一次因式的乘积）.

第三步. 当 $\underline{A}$ 有 Jordan 标准形 $J$ 时，由于 $J$ 与 $A$ 相似，所以 $J$ 的主对角线上元素都是 $A$ 的特征值，并且特征值 $\lambda_j$ 在 $J$ 的主对角

线上出现的次数等于 $\lambda_j$ 作为 $A$ 的特征多项式的根的重数. 对于每个特征值 $\lambda_j$，求出各级 Jordan 块的数目. 首先，求出 $\operatorname{rank}(A-\lambda_j I)$，则主对角元为 $\lambda_j$ 的 Jordan 块的总数

$$N(\lambda_j)=\dim V-\operatorname{rank}(A-\lambda_j I)$$

其次，求 $\operatorname{rank}(A-\lambda_j I)^2$，于是 1 级 Jordan 块 $J_1(\lambda_j)$ 的数目为

$$N(1;\lambda_j)=\operatorname{rank}(A-\lambda_j I)^2+n-2\operatorname{rank}(A-\lambda_j I)$$

比较 $N(1;\lambda_j)$ 与 $N(\lambda_j)$，若 $N(1;\lambda_j)<N(\lambda_j)$，则继续求 $\operatorname{rank}(A-\lambda_j I)^3$，于是 2 级 Jordan 块 $J_2(\lambda_j)$ 的数目为

$$\begin{aligned}&N(2;\lambda_j)\\&=\operatorname{rank}(A-\lambda_j I)^3+\operatorname{rank}(A-\lambda_j I)-2\operatorname{rank}(A-\lambda_j I)^2\end{aligned}$$

比较 $N(1;\lambda_j)+N(2;\lambda_j)$ 是否等于 $N(\lambda_j)$，如果相等，则主对角元为 $\lambda_j$ 的 Jordan 块数目已全部求出. 否则，继续计算 $\operatorname{rank}(\underline{A}-\lambda_j \underline{I})^4$，求出 $N(3;\lambda_j)$. 依次下去，直到已求出的各级 Jordan 块数目之和等于 $N(\lambda_j)$ 为止.

第四步. 根据第三步求出的各级 Jordan 块的数目，就可以写出 $\underline{A}$ 的一个 Jordan 标准形.

下面举一个例子.

**例 1** 设数域 $K$ 上 3 维线性空间 $V$ 的线性变换 $\underline{A}$ 在 $V$ 的一个基 $\alpha_1,\alpha_2,\alpha_3$ 下的矩阵为

$$A=\begin{pmatrix}2&3&2\\1&8&2\\-2&-14&-3\end{pmatrix}$$

求 $\underline{A}$ 的 Jordan 标准形.

**解**

$$\begin{aligned}f(\lambda)=|\lambda I-A|&=\begin{vmatrix}\lambda-2&-3&-2\\-1&\lambda-8&-2\\2&14&\lambda+3\end{vmatrix}\\&=(\lambda-1)(\lambda-3)^2\end{aligned}$$

由此看出，$\underset{\sim}{A}$ 有 Jordan 标准形.

对于特征值 $\lambda_1 = 1$，它是 $f(\lambda)$ 的 1 重根，从而 $\lambda_1$ 在 $\underset{\sim}{A}$ 的 Jordan 标准形 $J$ 的主对角线上出现一次，因此 $J$ 中主对角元为 1 的 Jordan 块只有一个，且它是 1 级的.

对于特征值 $\lambda_2 = 3$，先求 $\mathrm{rank}(A - 3I)$. 因为

$$A - 3I = \begin{pmatrix} -1 & 3 & 2 \\ 1 & 5 & 2 \\ -2 & -14 & -6 \end{pmatrix} \longrightarrow \begin{pmatrix} -1 & 3 & 2 \\ 0 & 8 & 4 \\ 0 & 0 & 0 \end{pmatrix}$$

所以 $\mathrm{rank}(A - 3I) = 2$. 从而 $N(\lambda_2) = 3 - 2 = 1$. 由于 $\lambda_2$ 是二重根，所以 $\lambda_2$ 在 $\underset{\sim}{A}$ 的 Jordan 标准形 $J$ 的主对角线上出现两次，从而 $J$ 中主对角元为 $\lambda_2$ 的唯一的一个 Jordan 块一定是 2 级的.

综上述，$\underset{\sim}{A}$ 的 Jordan 标准形为

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 3 & 1 \\ 0 & 0 & 3 \end{pmatrix}$$

**定义** 4　$V$ 上线性变换 $\underset{\sim}{A}$ 的一个 Jordan **基**是 $V$ 的一个基，它使得 $\underset{\sim}{A}$ 在这个基下的矩阵为 Jordan 形矩阵.

当我们已经求出 $\underset{\sim}{A}$ 的 Jordan 标准形 $J$ 以后，为了求出 $\underset{\sim}{A}$ 的一个 Jordan 基，只要把原来的基到 Jordan 基的过渡矩阵 $S$ 求出即可. 由于 $J = S^{-1}AS$，所以 $S$ 是矩阵方程

$$AX = XJ \tag{28}$$

的解并且应为可逆矩阵. 如果 $\dim V = n$，则(28) 是 $n^2$ 个未知量 $x_{ij}, i, j = 1, \cdots, n$ 的由 $n^2$ 个方程组成的线性方程组，解这个线性方程组，可求出 $X = (x_{ij})$，选取可逆矩阵(因为 $\underset{\sim}{A}$ 的 Jordan 标准形存在，所以满足方程(28) 的可逆矩阵一定存在)，便可作为过渡矩阵 $S$.

**例** 2　求例 1 的线性变换 $\underset{\sim}{A}$ 的一个 Jordan 基.

**解**　设 $X = (X_1, X_2, X_3)$，由 $\underset{\sim}{A}$ 的 Jordan 标准形以及方程(28) 得

$$A(X_1,X_2,X_3)=(X_1,3X_2,X_2+3X_3). \qquad (29)$$

所以 $\quad AX_1=X_1,\quad AX_2=3X_2,\quad AX_3=X_2+3X_3$

由此看出，$X_1$ 是 $A$ 的属于 1 的一个特征向量，解方程组

$$(I-A)Y=0$$

求得 $X_1=(2,0,-1)'$. 同理，$X_2$ 是 $A$ 的属于 3 的一个特征向量，解方程组

$$(3I-A)Y=0$$

求得 $X_2=(1,-1,2)'$. 再去解方程组 $(A-3I)Y=X_2$：

$$\begin{pmatrix}-1 & 3 & 2 & 1\\ 1 & 5 & 2 & -1\\ -2 & -14 & -6 & 2\end{pmatrix}\longrightarrow\begin{pmatrix}1 & 0 & -\frac{1}{2} & -1\\ 0 & 1 & \frac{1}{2} & 0\\ 0 & 0 & 0 & 0\end{pmatrix}$$

它的一个特解是 $Y_0=(-1,0,0)'$. 取 $X_3=(-1,0,0)'$，则

$$X=\begin{pmatrix}2 & 1 & -1\\ 0 & -1 & 0\\ -1 & 2 & 0\end{pmatrix} \qquad (30)$$

容易看出 $X$ 是可逆矩阵. 它就可作为 $V$ 的原来的基 $\alpha_1,\alpha_2,\alpha_3$ 到 $\underline{A}$ 的一个 Jordan 基 $\xi_1,\xi_2,\xi_3$ 的过渡矩阵. 所以

$$(\xi_1,\xi_2,\xi_3)=(\alpha_1,\alpha_2,\alpha_3)\begin{pmatrix}2 & 1 & -1\\ 0 & -1 & 0\\ -1 & 2 & 0\end{pmatrix}$$

即 $\underline{A}$ 的一个 Jordan 基是：

$$\xi_1=2\alpha_1-\alpha_3,\quad \xi_2=\alpha_1-\alpha_2+2\alpha_3,\quad \xi_3=-\alpha_1$$

矩阵的 Jordan 标准形的最重要的应用之一是计算矩阵的多项式，譬如，要计算 $A^m$，其中 $m$ 是很大的正整数. 如果 $A$ 有 Jordan 标准形 $J$，并且用上述方法求出了可逆矩阵 $X$，使得 $X^{-1}AX=J$，那么

$$A^m=(XJX^{-1})^m=XJ^mX^{-1} \qquad (31)$$

每个 Jordan 块 $J_t(\lambda_j)$ 的方幂易于计算：

$$J_t(\lambda_j)^r = (\lambda_j I + J_t(0))^r \tag{32}$$

由于 $\lambda_j I$ 与 $J_t(0)$ 可交换，因此可用二项式定理展开(32) 式右端. 注意：当 $r \geqslant t$ 时，$J_t(0)^r = 0$. 这表明，计算 $J^m$ 比直接计算 $A^m$ 要简便得多. 而且一旦把 $J$ 和 $X$ 求出来后，对一切正整数 $m$，可以用同一个公式(31) 计算 $A^m$.

**例 3** 对于例 1 中的 $A$，计算 $A^{10}$.

**解** 从例 1 知道，$A$ 的 Jordan 标准形是

$$J = \mathrm{diag}\left\{(1), \begin{pmatrix} 3 & 1 \\ 0 & 3 \end{pmatrix}\right\}$$

从例 2 知道，(30) 式给出的矩阵 $X$ 使得 $X^{-1}AX = J$. 从而 $A^m = XJ^mX^{-1}$. 易求出

$$X^{-1} = \begin{pmatrix} 0 & -2 & -1 \\ 0 & -1 & 0 \\ -1 & -5 & -2 \end{pmatrix}$$

当 $r \geqslant 2$，有

$$\begin{pmatrix} 3 & 1 \\ 0 & 3 \end{pmatrix}^r = \left[3I + \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}\right]^r$$

$$= 3^r I + r3^{r-1} I \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 3^r & r3^{r-1} \\ 0 & 3^r \end{pmatrix}$$

所以

$$A^{10} = X\mathrm{diag}\left\{(1)^{10}, \begin{pmatrix} 3 & 1 \\ 0 & 3 \end{pmatrix}^{10}\right\} X^{-1}$$

$$= \begin{pmatrix} 2 & 1 & -1 \\ 0 & -1 & 0 \\ -1 & 2 & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 3^{10} & 10 \cdot 3^9 \\ 0 & 0 & 3^{10} \end{pmatrix} \begin{pmatrix} 0 & -2 & -1 \\ 0 & -1 & 0 \\ -1 & -5 & -2 \end{pmatrix}$$

$$= \begin{pmatrix} -7 \cdot 3^9 & -38 \cdot 3^9 - 4 & -14 \cdot 3^9 - 2 \\ 10 \cdot 3^9 & 53 \cdot 3^9 & 20 \cdot 3^9 \\ -20 \cdot 3^9 & -106 \cdot 3^9 + 2 & -40 \cdot 3^9 + 1 \end{pmatrix}$$

* 计算 $n$ 级矩阵 $A$ 的多项式也可以不用 $A$ 的 Jordan 标准形，而用 Hamilton-Cayley 定理，以及带余除法，请看补充题十第 1 题的提示.

§3 的公式(4)给出了一个 Jordan 形矩阵的最小多项式的求法，因此，如果知道了 $A$ 的 Jordan 标准形 $J$，那么 $A$ 与 $J$ 有相同的最小多项式，从而可以立即写出 $A$ 的最小多项式 $m(\lambda)$ 为

$$m(\lambda)=(\lambda-\lambda_1)^{l_1}(\lambda-\lambda_2)^{l_2}\cdots(\lambda-\lambda_s)^{l_s}$$

其中 $\lambda_1,\lambda_2,\cdots,\lambda_s$ 是 $J$ 的主对角线上的所有不同的元素，$l_j$ 是主对角元为 $\lambda_j$ 的最大 Jordan 块的级数.

**例** 4 设复数域上一个 12 级矩阵 $A$ 的 Jordan 标准形是

$$J=\operatorname{diag}\left\{\begin{pmatrix}3&1\\0&3\end{pmatrix},\begin{pmatrix}3&1\\0&3\end{pmatrix},\begin{pmatrix}3&1\\0&3\end{pmatrix},(-2),(-2),\begin{pmatrix}\mathrm{i}&1\\0&\mathrm{i}\end{pmatrix},\begin{pmatrix}-\mathrm{i}&1\\0&-\mathrm{i}\end{pmatrix}\right\}$$

则 $A$ 的最小多项式 $m(\lambda)$ 为

$$m(\lambda)=(\lambda-3)^2(\lambda+2)(\lambda-\mathrm{i})^2(\lambda+\mathrm{i})^2$$

而 $A$ 的特征多项式 $f(\lambda)$ 为

$$f(\lambda)=(\lambda-3)^6(\lambda+2)^2(\lambda-\mathrm{i})^2(\lambda+\mathrm{i})^2$$

## 习 题 10.4

1. 求下列数域 $K$ 上的矩阵的 Jordan 标准形：

(1) $\begin{bmatrix}13&16&16\\-5&-7&-6\\-6&-8&-7\end{bmatrix}$；　(2) $\begin{bmatrix}3&0&8\\3&-1&6\\-2&0&-5\end{bmatrix}$；

(3) $\begin{bmatrix}0&1&0\\-4&4&0\\-2&1&2\end{bmatrix}$；　(4) $\begin{bmatrix}4&-5&2\\5&-7&3\\6&-9&4\end{bmatrix}$；

(5) $\begin{bmatrix}1&-3&3\\-2&-6&13\\-1&-4&8\end{bmatrix}$；　(6) $\begin{bmatrix}1&-3&4\\4&-7&8\\6&-7&7\end{bmatrix}$；

(7) $\begin{pmatrix} 3 & -4 & 0 & 2 \\ 4 & -5 & -2 & 4 \\ 0 & 0 & 3 & -2 \\ 0 & 0 & 2 & -1 \end{pmatrix}$； (8) $\begin{pmatrix} 3 & -1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 3 & 0 & 5 & -3 \\ 4 & -1 & 3 & -1 \end{pmatrix}$.

2. 设数域 $K$ 上 3 维线性空间 $V$ 的线性变换 $\underline{A}$ 在 $V$ 的一个基 $\alpha_1,\alpha_2,\alpha_3$ 下的矩阵 $A$ 分别是第 1 题的(1),(4),(7),在这三种情形下,分别求 $\underline{A}$ 的一个 Jordan 基.

3. 分别对于第 1 题(1),(4) 中的矩阵 $A$,求 $A^{10}$.

4. 分别对于第 1 题各个小题中的矩阵 $A$,写出 $A$ 的特征多项式和最小多项式.

5. 证明:$J_t(a) \sim J_t(a)'$.

(提示:利用矩阵 $\begin{pmatrix} 0 & 0 & \cdots & 0 & 1 \\ 0 & 0 & \cdots & 1 & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 1 & 0 & \cdots & 0 & 0 \end{pmatrix}$.)

6. 证明:任一 $n$ 级复矩阵 $A$ 与它的转置 $A'$ 相似.

7. 证明:$n$ 级复矩阵 $A$ 的迹等于 $A$ 的 $n$ 个特征值的和.

8. 证明:设 $n$ 级复矩阵 $A$ 的 $n$ 个特征值是 $\lambda_1,\cdots,\lambda_n$,则对于任一正整数 $m$,$A^m$ 的 $n$ 个特征值是 $\lambda_1^m,\cdots,\lambda_n^m$.

9. 证明:设 $n$ 级复矩阵 $A$ 的 $n$ 个特征值是 $\lambda_1,\cdots,\lambda_n$,则对于任一复系数多项式 $g(\lambda)$,矩阵 $A$ 的多项式 $g(A)$ 的 $n$ 个特征值是 $g(\lambda_1),\cdots,g(\lambda_n)$.

(提示:利用 $A$ 的 Jordan 标准形).

10. 证明:对于任一 $n$ 级复矩阵 $A$,存在一个可逆矩阵 $S$,使得 $S^{-1}AS = GH$,其中 $G,H$ 都为对称矩阵,并且 $G$ 可逆.

(提示:利用 $A$ 的 Jordan 标准形,以及利用矩阵 $\begin{pmatrix} 0 & 0 & \cdots & 0 & 1 \\ 0 & 0 & \cdots & 1 & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 1 & 0 & \cdots & 0 & 0 \end{pmatrix}$,注意此矩阵的平方等于 $I$.)

11. 证明:如果域 $F$ 上的 $n$ 级矩阵 $A$ 是幂零矩阵,则对一切正整数 $k$,有 $Tr(A^k) = 0$.

(提示:利用 $A$ 的 Jordan 标准形).

12. 证明:$n$ 级复矩阵 $A$ 是幂零矩阵的充分必要条件是,$A$ 的特征值全为 0.

13. 证明:域 $F$ 上 $n$ 级幂零矩阵的幂零指数不超过 $n$.

* 14. 证明:如果 $n$ 级复矩阵 $A$ 满足 $Tr(A^k)=0$,其中 $k=1,2,\cdots,n$,则 $A$ 是幂零矩阵.

(提示:利用 $A$ 的 Jordan 标准形和牛顿公式,以及复系数多项式的根与系数的关系,去证 $A$ 的特征值全为零.再用第 12 题结果).

* 15. 设 $A,B,C$ 都是 $n$ 级复矩阵,且 $AB-BA=C$,证明:如果 $C$ 与 $A$ 可交换,则 $C$ 是幂零矩阵.

(提示:用第 14 题的结果).

* 16. 证明:如果 $n$ 级复矩阵 $A$ 可对角化,并且 $Tr(A^k)=0,k=1,\cdots,n$,则 $A=0$.

# 补 充 题 十

1. 设

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

求 $A^{100}+2A^{90}+3A^{60}$.

(提示:设

$$g(\lambda)=\lambda^{100}+2\lambda^{90}+3\lambda^{60}$$

则 $$g(A)=A^{100}+2A^{90}+3A^{60}$$

求出 $A$ 的特征多项式 $f(\lambda)=|\lambda I-A|$.然后作带余除法,

$$g(\lambda)=h(\lambda)f(\lambda)+r(\lambda),\quad \deg r(\lambda)<\deg f(\lambda)=3$$

设 $r(\lambda)=c_2\lambda^2+c_1\lambda+c_0$,其中 $c_0,c_1,c_2$ 待定.设法求出 $c_0,c_1,c_2$,则 $g(A)=h(A)f(A)+r(A)=r(A)=c_2A^2+c_1A+c_0I$.)

2. 设 $A=\begin{pmatrix}0&1\\-1&2\end{pmatrix}$,求 $A^{100}+3A^{23}+A^{20}$.

3. 设

$$A=\begin{pmatrix}2&2&1\\1&3&1\\1&2&2\end{pmatrix}$$

求 $A^{1000}$.

4. 设 $A$ 为第 1 题中的 3 级矩阵,证明:当 $k\geqslant 3$ 时,有

$$A^k=A^{k-2}+A^2-I$$

然后利用这个公式计算 $A^{100}$.

5. 设域 $F$ 上矩阵 $J=\mathrm{diag}\{J_t(\lambda_1),J_r(\lambda_2)\}$,且 $\lambda_1\neq\lambda_2$. 证明:如果域 $F$ 上矩阵 $A$ 满足 $AJ=JA$,则 $A=\mathrm{diag}\{A_1,A_2\}$,并且 $A_1$ 与 $J_t(\lambda_1)$ 可交换,$A_2$ 与 $J_r(\lambda_2)$ 可交换.

(提示:利用类似于习题 10.2 第 3 题的结果).

# 第十一章　线性函数·对偶空间·双线性函数

在第九章的 §1 至 §4，我们讲了域 $F$ 上线性空间 $V$ 到 $V'$ 的线性映射的定义、存在性、运算、象与核、及其与矩阵的关系. 从第九章 §5 至第十章讨论了 $V$ 到自身的线性映射，即 $V$ 上的线性变换的结构.

本章来讨论域 $F$ 上线性空间 $V$ 到 $F$（把 $F$ 看成自身上的 1 维线性空间）的线性映射，称它为 $V$ 上的线性函数. 我们将着重讨论 $V$ 上所有线性函数组成的集合 $\mathrm{Hom}(V,F)$ 的结构，及其与 $V$ 的关系.

在讨论 $V$ 上线性函数的基础上，我们将讲述 $V$ 上的双线性函数，这为第十二章至第十四章打下基础.

## §1　线性函数

设 $V$ 是域 $F$ 上的一个线性空间.

**定义 1**　$V$ 到 $F$ 的一个映射 $f$ 称为 $V$ 上的一个**线性函数**，如果

$$f(\alpha+\beta)=f(\alpha)+f(\beta),\quad \forall\, \alpha,\beta\in V$$

$$f(k\alpha)=kf(\alpha),\quad \forall\, \alpha\in V, k\in F$$

线性函数也称为**余向量**(covectors).

由于域 $F$ 可以看成是自身上的线性空间，所以 $V$ 上的线性函数可以看成是线性空间 $V$ 到线性空间 $F$ 的线性映射，因此第九章 §1 至 §4 关于线性映射的结果对于线性函数也成立.

线性函数是十分重要的一类函数，在数学的各个分支和许多

实际问题中都会遇到线性函数. 下面举几个例子.

**例** 1　定积分把每一个连续函数 $f(x)$ 对应到一个实数 $\int_a^b f(x)\mathrm{d}x$，并且有

$$\int_a^b [f(x)+g(x)]\mathrm{d}x = \int_a^b f(x)\mathrm{d}x + \int_a^b g(x)\mathrm{d}x$$

$$\int_a^b [kf(x)]\mathrm{d}x = k\int_a^b f(x)\mathrm{d}x$$

所以定积分是 $C[a,b]$ 上的一个线性函数.

**例** 2　矩阵的迹把域 $F$ 上每一个 $n$ 级矩阵 $A=(a_{ij})$ 对应到 $F$ 中的一个元素 $\sum_{i=1}^{n} a_{ii}$，并且有

$$Tr(A+B) = Tr(A) + Tr(B)$$

$$Tr(kA) = kTr(A)$$

所以矩阵的迹是 $M_n(F)$ 上的一个线性函数.

**例** 3　在域 $F$ 上的一元多项式环 $F[x]$ 中，不定元 $x$ 用 $F$ 中一个元素 $t$ 代入，它把每一个多项式 $f(x)$ 对应到 $F$ 中一个元素 $f(t)$. 由于不定元 $x$ 用 $t$ 代入保持加法与乘法（从而也保持纯量乘法），所以 $x$ 用 $t(t\in F)$ 代入是线性空间 $F[x]$ 上的一个线性函数.

**例** 4　给定 $F$ 中 $n$ 个元素 $a_1,a_2,\cdots,a_n$，考虑 $F^n$ 到 $F$ 的一个映射 $f$：

$$(x_1,x_2,\cdots,x_n) \longmapsto a_1x_1 + a_2x_2 + \cdots + a_nx_n$$

即

$$f(x_1,x_2,\cdots,x_n) = a_1x_1 + a_2x_2 + \cdots + a_nx_n \tag{1}$$

容易直接验证 $f$ 保持加法与纯量乘法两种运算，因此形如(1)的函数 $f$ 是 $F^n$ 上的一个线性函数.

注意：在数学分析里，把形如

$$g(x_1,x_2,\cdots,x_n) = a_1x_1 + \cdots + a_nx_n + b$$

的 $n$ 元函数 $g$ 叫做线性函数. 如果 $b\neq 0$，那么 $g$ 不保持加法运算，也不保持纯量乘法运算，从而 $g$ 不是定义 1 意义上的线性函数. 所以“线性函数”这一术语在分析和代数里有不同的含义. 代数课程

中讲的线性函数是分析课程中的齐次线性函数(即$b=0$时的线性函数).

现在我们来讨论一般的有限维线性空间$V$上的线性函数$f$的表达式.

设$V$是域$F$上的$n$维线性空间,$f$是$V$上的一个线性函数.在$V$中取一个基$\alpha_1,\alpha_2,\cdots,\alpha_n$.由于$f$可以看成是线性空间$V$到线性空间$F$的一个线性映射,因此$f$完全被它在$V$的一个基$\alpha_1,\alpha_2,\cdots,\alpha_n$上的作用所决定.即只要知道$f(\alpha_1),f(\alpha_2),\cdots,f(\alpha_n)$,就可以知道$V$中任一向量$\beta=\sum_{i=1}^{n}x_i\alpha_i$在$f$下的象:

$$f(\beta)=\sum_{i=1}^{n}x_if(\alpha_i)\tag{2}$$

(2)式就是线性函数$f$在基$\alpha_1,\cdots,\alpha_n$下的表达式,它表明:$f$在$\beta$上的函数值$f(\beta)$是$\beta$的坐标$x_1,\cdots,x_n$的一次齐次多项式.

现在我们来讨论如何构造域$F$上$n$维线性空间$V$上的线性函数.

在$V$中取一个基$\alpha_1,\cdots,\alpha_n$.在$F$中任意取定$n$个元素$a_1,a_2,\cdots,a_n$,据第九章定理9.1.2,存在$V$到$F$的唯一的线性映射$f$,使得

$$f(\alpha_i)=a_i,\quad i=1,\cdots,n\tag{3}$$

从定理9.1.2的证明过程知道,$\beta=\sum_{i=1}^{n}x_i\alpha_i$在$f$下的象是

$$f(\beta)=\sum_{i=1}^{n}x_ia_i\tag{4}$$

## 习　题　11.1

1. 设$V=C[a,b]$,$V$上的下列函数哪些是线性函数?

(1) $f(x)\longmapsto\int_a^b f^2(x)\mathrm{d}x$;

(2) $f(x) \mapsto \int_a^b f(x)g(x)\mathrm{d}x$,

其中，$g(x)$ 是固定的一个函数；

(3) $f(x) \mapsto f(0)$.

2. 设 $V$ 是域 $F$ 上的一个 3 维线性空间，$\alpha_1,\alpha_2,\alpha_3$ 是 $V$ 的一个基，$f$ 是 $V$ 上的一个线性函数，已知

$$f(\alpha_1 + 2\alpha_3) = 4,\quad f(\alpha_2 + 3\alpha_3) = 0,\quad f(4\alpha_1 + \alpha_2) = 5$$

求 $f(x_1\alpha_1 + x_2\alpha_2 + x_3\alpha_3)$.

3. $V$ 及 $\alpha_1,\alpha_2,\alpha_3$ 同第 2 题，试找出一个线性函数 $f$，使

$$f(3\alpha_1 + \alpha_2) = 2,\quad f(\alpha_2 - \alpha_3) = 1,\quad f(2\alpha_1 + \alpha_3) = 2$$

4. 设 $V$ 是域 $F$ 上一个线性空间，$\mathrm{char}F = 0$. 设 $f_1,f_2,\cdots,f_s$ 都是 $V$ 上的线性函数，并且它们都不是零函数. 证明：存在 $\alpha \in V$，使得

$$f_i(\alpha) \neq 0,\quad i = 1,2,\cdots,s$$

（提示：利用习题 8.5 的第 12 题结论）.

## §2 对偶空间

设 $V$ 是域 $F$ 上的一个线性空间，由于 $V$ 上的线性函数可以看成是线性空间 $V$ 到线性空间 $F$ 的线性映射，因此运用第九章的记号，可以把 $V$ 上所有线性函数组成的集合记作 $\mathrm{Hom}(V,F)$. 这一节我们来讨论 $\mathrm{Hom}(V,F)$ 的结构，以及它与 $V$ 的关系.

从第九章 § 2 知道，$\mathrm{Hom}(V,F)$ 也是域 $F$ 上的一个线性空间，称它是 $V$ 上的**线性函数空间**，也记作 $T_1(V)$.

从现在起设 $V$ 是有限维的，维数为 $n$. 由于域 $F$ 看成自身上的线性空间是 1 维的，因此据第九章推论 9.4.1 得

$$\dim\mathrm{Hom}(V,F) = \dim V \cdot \dim F = n \tag{1}$$

这表明 $\mathrm{Hom}(V,F)$ 与 $V$ 的维数相同，从而它们同构，即

$$\mathrm{Hom}(V,F) \cong V \tag{2}$$

在 $V$ 中取一个基 $\alpha_1,\alpha_2,\cdots,\alpha_n$，现在我们来找 $\mathrm{Hom}(V,F)$ 的一个基. 由于 $\mathrm{Hom}(V,F)$ 是 $n$ 维的，因此只要找出 $V$ 上的 $n$ 个线性函

数，并且它们线性无关就可以了. 上一节的最后一段指出，在 $F$ 中任意取定 $n$ 个元素 $a_1,a_2,\cdots,a_n$，一定存在 $V$ 到 $F$ 的唯一的线性映射 $f$，使得 $f(\alpha_i)=a_i, i=1,2,\cdots,n$. 于是我们可以按照这个方法找出 $V$ 上的 $n$ 个线性函数：

给定 $F$ 中 $n$ 个元素：$1,0,\cdots,0$，则存在 $V$ 上唯一的线性函数 $f_1$，使得 $f_1(\alpha_1)=1, f_1(\alpha_2)=\cdots=f_1(\alpha_n)=0$；

给定 $F$ 中 $n$ 个元素：$0,1,0,\cdots,0$，则存在 $V$ 上唯一的线性函数 $f_2$，使得 $f_2(\alpha_2)=1, f_2(\alpha_j)=0, j\neq 2$；

… … … … … …

给定 $F$ 中 $n$ 个元素：$0,\cdots,0,1$，则存在 $V$ 上唯一的线性函数 $f_n$，使得 $f_n(\alpha_n)=1, f_n(\alpha_j)=0, j\neq n$.

这样我们找到了 $V$ 上的 $n$ 个线性函数 $f_1,f_2,\cdots,f_n$，其中 $f_i$ $(1\leqslant i\leqslant n)$ 在基向量上的函数值为

$$f_i(\alpha_j)=\begin{cases}1, & j=i\\ 0, & j\neq i\end{cases} \tag{3}$$

采用 Kronecker 记号 $\delta_{ij}$（见第四章 §10），(3) 式可写成

$$f_i(\alpha_j)=\delta_{ij},\quad j=1,\cdots,n \tag{4}$$

现在我们来证明 $f_1,f_2,\cdots,f_n$ 是线性无关的. 设

$$k_1f_1+k_2f_2+\cdots+k_nf_n=0 \tag{5}$$

任给 $j$ $(1\leqslant j\leqslant n)$，计算(5) 式左右两边的函数在 $\alpha_j$ 上的函数值得

$$k_1f_1(\alpha_j)+k_2f_2(\alpha_j)+\cdots+k_nf_n(\alpha_j)=0$$

利用(3)式便得出，$k_j=0$. 由于 $j=1,\cdots,n$，因此 $f_1,f_2,\cdots,f_n$ 线性无关.

综上述，$f_1,f_2,\cdots,f_n$ 是 $\mathrm{Hom}(V,F)$ 的一个基.

我们把上面所得的结果写成一个定理如下：

**定理** 11.2.1 设 $V$ 是域 $F$ 上的 $n$ 维线性空间，则 $V$ 上所有线性函数组成的集合 $\mathrm{Hom}(V,F)$ 也是域 $F$ 上的 $n$ 维线性空间，称它是 $V$ 的**对偶空间**(或**共轭空间**)，简记成 $V^*$，并且 $V^*\cong V$. 如果在

$V$ 中取一个基 $\alpha_1,\alpha_2,\cdots,\alpha_n$，那么由(3)式确定的线性函数 $f_1,f_2,\cdots,f_n$ 是 $V^*$ 的一个基，称它是 $\alpha_1,\alpha_2,\cdots,\alpha_n$ 的**对偶基**. ▌

现在在 $V$ 中取一个基 $\alpha_1,\alpha_2,\cdots,\alpha_n$，在 $V^*$ 中取 $\alpha_1,\alpha_2,\cdots,\alpha_n$ 的对偶基 $f_1,f_2,\cdots,f_n$. 我们来分别讨论 $V$ 中任一向量 $\beta$ 在基 $\alpha_1,\alpha_2,\cdots,\alpha_n$ 下的坐标，以及 $V^*$ 中任一向量 $f$ 在基 $f_1,f_2,\cdots,f_n$ 下的坐标.

$V$ 中任取一个向量 $\beta=\sum_{j=1}^{n}x_j\alpha_j$，从(3)式得

$$f_i(\beta)=\sum_{j=1}^{n}x_jf_i(\alpha_j)=x_i \tag{6}$$

即 $\beta$ 在基 $\alpha_1,\cdots,\alpha_n$ 下的坐标的第 $i$ 个分量等于 $f_i(\beta)$. 因此

$$\beta=\sum_{i=1}^{n}f_i(\beta)\alpha_i \tag{7}$$

$V^*$ 中任取一个向量 $f=\sum_{i=1}^{n}c_if_i$，比较左右两边的函数在 $\alpha_j$ 上的函数值得

$$f(\alpha_j)=\sum_{i=1}^{n}c_if_i(\alpha_j)=c_j \tag{8}$$

这表明 $f$ 在基 $f_1,f_2,\cdots,f_n$ 下的坐标的第 $j$ 个分量等于 $f(\alpha_j)$. 因此

$$f=\sum_{j=1}^{n}f(\alpha_j)f_j \tag{9}$$

**例 1** 设 $V=M_2(F)$，在 $V$ 中取一个基 $E_{11},E_{12},E_{21},E_{22}$，求它的对偶基 $f_{11},f_{12},f_{21},f_{22}$，要求写出 $f_{ij}$ 的表达式；并且求 $V$ 上任一线性函数 $f$ 的表达式.

**解** 从(3)式得

$$f_{11}(E_{11})=1,\quad f_{11}(E_{12})=f_{11}(E_{21})=f_{11}(E_{22})=0$$
$$f_{12}(E_{12})=1,\quad f_{12}(E_{11})=f_{12}(E_{21})=f_{12}(E_{22})=0$$
$$f_{21}(E_{21})=1,\quad f_{21}(E_{11})=f_{21}(E_{12})=f_{21}(E_{22})=0$$
$$f_{22}(E_{22})=1,\quad f_{22}(E_{11})=f_{22}(E_{12})=f_{22}(E_{21})=0$$

任取 $A=(a_{ij})\in M_2(F)$，由于 $A=\sum_{i=1}^{2}\sum_{j=1}^{2}a_{ij}E_{ij}$，所以

$$f_{11}(A)=a_{11},\quad f_{12}(A)=a_{12},\quad f_{21}(A)=a_{21},\quad f_{22}(A)=a_{22}$$

任取 $V$ 上一个线性函数 $f$，设 $f(E_{ij})=c_{ij}, i,j=1,2$，则从(9)式得

$$\begin{aligned}f(A)&=c_{11}f_{11}(A)+c_{12}f_{12}(A)+c_{21}f_{21}(A)+c_{22}f_{22}(A)\\&=c_{11}a_{11}+c_{12}a_{12}+c_{21}a_{21}+c_{22}a_{22}\end{aligned}\tag{10}$$

在现代数学文献中，把 $V$ 的一个基 $\alpha_1,\cdots,\alpha_n$ 的对偶基记成 $\alpha^1,\cdots,\alpha^n$，这里 $\alpha^i$ 中的 $i$ 是上标，不是指数. 考虑到读者可能对上标不熟悉，因此我们在本书中把 $\alpha_1,\cdots,\alpha_n$ 的对偶基记成 $f_1,\cdots,f_n$. 但是建议有兴趣的读者把前面和下面讲的有关对偶基的公式采用记号 $\alpha^1,\alpha^2,\cdots,\alpha^n$ 表示 $\alpha_1,\alpha_2,\cdots,\alpha_n$ 的对偶基.

$V$ 中不同基的对偶基之间有什么关系？

**定理** 11.2.2 设 $V$ 是域 $F$ 上 $n$ 维线性空间，在 $V$ 中取两个基：$\alpha_1,\cdots,\alpha_n$ 与 $\beta_1,\cdots,\beta_n$. 设它们的对偶基分别是 $f_1,\cdots,f_n$ 与 $g_1,\cdots,g_n$. 如果 $V$ 中基 $\alpha_1,\cdots,\alpha_n$ 到基 $\beta_1,\cdots,\beta_n$ 的过渡矩阵是 $A=(a_{ij})$，则 $V^*$ 中基 $f_1,\cdots,f_n$ 到基 $g_1,\cdots,g_n$ 的过渡矩阵为 $(A^{-1})'$.

**证明** 由已知条件，有

$$(\beta_1,\cdots,\beta_n)=(\alpha_1,\cdots,a_n)A\tag{11}$$

于是

$$\beta_i=\sum_{k=1}^{n}a_{ki}\alpha_k\tag{12}$$

设 $f_1,\cdots,f_n$ 到 $g_1,\cdots,g_n$ 的过渡矩阵为 $B=(b_{ij})$，则

$$(g_1,\cdots,g_n)=(f_1,\cdots,f_n)B\tag{13}$$

于是 $g_j=\sum_{k=1}^{n}b_{kj}f_k$. 考虑此式两边的函数在 $\beta_i$ 上的函数值，得

$$\delta_{ij}=g_j(\beta_i)=\sum_{k=1}^{n}b_{kj}f_k(\beta_i)=\sum_{k=1}^{n}b_{kj}a_{ki}\tag{14}$$

在上述推导过程的最后一步用到了：$f_k(\beta_i)$ 等于 $\beta_i$ 在基 $\alpha_1,\cdots,\alpha_n$ 下的坐标的第 $k$ 个分量，而从(12)式知，它是 $a_{ki}$.

由于 $a_{ki}$ 是 $A'$ 的$(i,k)$元，所以(14)式表明

$$A'B = I \tag{15}$$

因此，$B = (A')^{-1} = (A^{-1})'$. ▌

* 现在我们来找 $V$ 到 $V^*$ 的一个同构映射. 因为 $V$ 和 $V^*$ 都是 $n$ 维的，所以它们都与 $F^n$ 同构. 我们知道，在域 $F$ 上一个 $n$ 维线性空间中取定一个基后，让每个向量对应到它在这个基下的坐标就是所给 $n$ 维线性空间到 $F^n$ 的一个同构映射. 我们在 $V$ 中取一个基 $\alpha_1$，$\alpha_2, \cdots, \alpha_n$，在 $V^*$ 中取 $\alpha_1, \alpha_2, \cdots, \alpha_n$ 的对偶基 $f_1, f_2, \cdots, f_n$，则有 $V$ 到 $F^n$ 的一个同构映射 $\sigma_1$：

$$\sigma_1(\sum_{i=1}^{n} a_i\alpha_i) = (a_1, a_2, \cdots, a_n)$$

又有 $F^n$ 到 $V^*$ 的一个同构映射 $\sigma_2$：

$$\sigma_2(a_1, a_2, \cdots, a_n) = \sum_{i=1}^{n} a_i f_i$$

从而有 $V$ 到 $V^*$ 的一个同构映射 $\sigma = \sigma_2\sigma_1$：

$$\sigma(\sum_{i=1}^{n} a_i\alpha_i) = \sum_{i=1}^{n} a_i f_i \tag{16}$$

设 $\alpha = \sum_{i=1}^{n} a_i\alpha_i$，记 $\sigma(\alpha) = f_\alpha$，则从(16)式得

$$f_\alpha = \sum_{i=1}^{n} a_i f_i \tag{17}$$

对于 $V$ 中任一向量 $\beta = \sum_{i=1}^{n} b_i\alpha_i$，从(17)式和(6)式得

$$f_\alpha(\beta) = \sum_{i=1}^{n} a_i f_i(\beta) = \sum_{i=1}^{n} a_i b_i \tag{18}$$

(18)式说明：$\alpha$ 在上述同构映射下的象 $f_\alpha$ 在 $\beta$ 上的函数值 $f_\alpha(\beta)$ 等于 $\alpha$ 与 $\beta$ 的坐标的对应分量乘积之和.

以上的讨论是对域 $F$ 上任一 $n$ 维线性空间进行的. 因此对于域 $F$ 上 $n$ 维线性空间 $V$，我们也可以考虑 $V^*$ 上的所有线性函数组成的线性空间 $\mathrm{Hom}(V^*, F)$（也记成 $T_1(V^*)$），它是 $V^*$ 的对偶空间，记成 $(V^*)^*$，简记成 $V^{**}$. 据定理 11.2.1 得，

$$\dim V^{**} = \dim V^{*} = \dim V$$

因此有
$$V \cong V^{**} \tag{19}$$

$V^{**}$ 称为 $V$ 的**双重对偶空间**.

* 现在我们来求 $V$ 到 $V^{**}$ 的一个同构映射,在 $V$ 中取一个基 $\alpha_1,\cdots,\alpha_n$,设它的对偶基是 $f_1,\cdots,f_n$. 任取 $V$ 中一个向量 $\alpha = \sum_{i=1}^{n} a_i\alpha_i$. 据上两段结论,有 $V$ 到 $V^*$ 的一个同构映射 $\sigma_1$,它把 $\alpha$ 映成 $f_\alpha$. 对 $V^*$,有 $V^*$ 到 $V^{**}$ 的一个同构映射 $\sigma_2$,它把 $f_\alpha$ 映成 $\alpha^{**}$,其中 $\alpha^{**}(f)$ 等于 $f_\alpha$ 与 $f$ 在基 $f_1,\cdots,f_n$ 下的坐标的对应分量乘积之和. 据(17)式和(9)式,有

$$f_\alpha = \sum_{i=1}^{n} a_i f_i, \quad f = \sum_{i=1}^{n} f(\alpha_i) f_i$$

因此
$$\alpha^{**}(f) = \sum_{i=1}^{n} a_i f(\alpha_i) = f(\sum_{i=1}^{n} a_i\alpha_i) = f(\alpha) \tag{20}$$

于是我们找到了 $V$ 到 $V^{**}$ 的一个同构映射 $\sigma = \sigma_2\sigma_1$,它把 $V$ 中向量 $\alpha$ 映成 $V^{**}$ 中元素 $\alpha^{**}$,其中

$$\alpha^{**}(f) = f(\alpha), \quad \forall f \in V^* \tag{21}$$

把上面的结果写成一个定理:

* **定理** 11.2.3 设 $V$ 是域 $F$ 上的 $n$ 维线性空间,$V^{**}$ 是 $V$ 的双重对偶空间,则

$$V \cong V^{**}$$

并且 $V$ 到 $V^{**}$ 的一个同构映射是 $\sigma: \alpha \longmapsto \alpha^{**}$,其中

$$\alpha^{**}(f) = f(\alpha), \quad \forall f \in V^*$$ ▌

* 这里我们要指出,$V$ 到 $V^{**}$ 的上述同构映射不依赖于 $V$ 中基的选择. 理由如下:上面在 $V$ 中取定一个基 $\alpha_1,\cdots,\alpha_n$,我们找到了 $V$ 到 $V^{**}$ 的一个同构映射 $\sigma: \alpha \longmapsto \alpha^{**}$,其中

$$\alpha^{**}(f) = f(\alpha), \quad \forall f \in V^*$$

也就是
$$\sigma(\alpha)f = f(\alpha), \quad \forall f \in V^*$$

现在在 $V$ 中另取一个基 $\beta_1,\cdots,\beta_n$,设它的对偶基是 $g_1,\cdots,g_n$. 与前

面一样的理由，有 $V$ 到 $V^*$ 的一个同构映射 $\tau_1$，它把 $V$ 中向量 $\alpha = \sum_{i=1}^{n} b_i\beta_i$ 映成 $g_\alpha$. 又有 $V^*$ 到 $V^{**}$ 的同构映射 $\tau_2$，它把 $g_\alpha$ 映成 $\tau_2(g_\alpha)$，其中 $\tau_2(g_\alpha)f$ 等于 $g_\alpha$ 与 $f$ 在基 $g_1,\cdots,g_n$ 下的坐标的对应分量乘积之和. 因为 $g_\alpha = \sum_{i=1}^{n} b_i g_i$，并且 $f = \sum_{i=1}^{n} f(\beta_i)g_i$，所以

$$\tau_2(g_\alpha)f = \sum_{i=1}^{n} b_i f(\beta_i) = f(\sum_{i=1}^{n} b_i\beta_i) = f(\alpha), \forall\, f \in V^* \quad (22)$$

于是我们得到 $V$ 到 $V^{**}$ 的又一个同构映射 $\tau = \tau_2\tau_1$，它把 $V$ 中向量 $\alpha$ 映成 $\tau(\alpha)$，其中

$$\tau(\alpha)f = (\tau_2\tau_1(\alpha))f = \tau_2(g_\alpha)f = f(\alpha), \forall\, f \in V^*$$

因此 $\sigma(\alpha)f = \tau(\alpha)f, \forall\, f \in V^*$. 由此得出

$$\sigma(\alpha) = \tau(\sigma), \quad \forall\, \alpha \in V$$

所以 $\sigma = \tau$. 这证明了：$V$ 到 $V^{**}$ 的同构映射：$\alpha \longmapsto \alpha^{**}$，其中 $\alpha^{**}(f) = f(\alpha)$ 不依赖于 $V$ 中基的选择. 这样的同构映射称为**标准同构或自然同构**.

* 由于 $V$ 到 $V^{**}$ 存在自然同构，因此我们可以把 $V^{**}$ 与 $V$ 等同，从而可以把 $V$ 看成 $V^*$ 的对偶空间，这样 $V$ 与 $V^*$ 就互为对偶空间. 这就是为什么把 $V^*$ 称为 $V$ 的对偶空间的原因.

* 由于 $V$ 可以看成是 $V^*$ 的对偶空间 $V^{**}$，而 $V^{**}$ 是 $V^*$ 上所有线性函数组成的空间，因此任一 $n$ 维线性空间可以看成是某个 $n$ 维线性空间上所有线性函数组成的空间.

* 注意：前面建立的 $V$ 到 $V^*$ 的同构映射就会依赖于 $V$ 中基的选择，例子可见本节习题的第 3 题，因此 $V$ 到 $V^*$ 的同构映射不是自然同构.

## 习 题 11.2

1. 设 $V$ 是数域 $K$ 上一个 3 维线性空间，$\alpha_1,\alpha_2,\alpha_3$ 是 $V$ 的一个基，$f_1,f_2,f_3$

是 $\alpha_1,\alpha_2,\alpha_3$ 的对偶基. 设

$$\beta_1=2\alpha_1+\alpha_2+2\alpha_3,\quad \beta_2=\alpha_1+2\alpha_2-2\alpha_3,\quad \beta_3=-2\alpha_1+2\alpha_2+\alpha_3$$

证明：$\beta_1,\beta_2,\beta_3$ 是 $V$ 的一个基，并且求它的对偶基(用 $f_1,f_2,f_3$ 表出).

2. 设 $V=R[x]_3$，对于 $g(x)\in V$，定义

$$f_1(g(x))=\int_0^1 g(x)\mathrm{d}x$$

$$f_2(g(x))=\int_0^2 g(x)\mathrm{d}x$$

$$f_3(g(x))=\int_0^{-1} g(x)\mathrm{d}x$$

证明：$f_1,f_2,f_3$ 是 $V^*$ 的一个基；并且求出 $V$ 的一个基 $g_1(x),g_2(x),g_3(x)$，使得 $f_1,f_2,f_3$ 是它的对偶基.

3. 设 $V$ 是域 $F$ 上 1 维线性空间，$V$ 中取两个基：$\alpha_1$ 和 $\beta_1$，其中 $\beta_1=a\alpha_1$，$a\in F$. 分别求 $\alpha_1$ 和 $\beta_1$ 的对偶基 $f_1$ 和 $g_1$. 证明：如果 $a^2\neq 1$，那么 $V$ 到 $V^*$ 的两个同构映射

$$\sigma:\alpha_1\longmapsto f_1;\qquad \tau:\beta_1\longmapsto g_1$$

是不相同的.

(提示：为了证 $\sigma\neq\tau$，只要证 $\sigma(\alpha_1)\neq\tau(\alpha_1)$.)

4. 设 $V$ 是域 $F$ 上线性空间(不必是有限维的)，$\mathrm{char}F=0$. 在 $V^*$ 中定义乘法运算如下：

$$(fg)(\alpha)=f(\alpha)g(\alpha),\quad \forall\,\alpha\in V$$

证明：如果 $fg=0$，则 $f=0$ 或 $g=0$.

*5. 设 $V$ 是域 $F$ 上线性空间(不必是有限维的)，$f$ 是 $V$ 上非零线性函数.

(1) 证明 $\mathrm{Ker}f$ 是 $V$ 的极大子空间(即如果 $V$ 的子空间 $U$ 满足 $\mathrm{Ker}f\subset U\subset V$，则 $U=\mathrm{Ker}f$ 或 $U=V$)；

(2) 任意给定 $\beta\notin\mathrm{Ker}f$，证明：$V$ 中任一向量 $\alpha$ 可以唯一地表示成

$$\alpha=\eta+k\beta,\quad 其中\ \eta\in\mathrm{Ker}f,k\in F$$

*6. 设 $V$ 是域 $F$ 上线性空间(不必是有限维的)，证明：如果 $V$ 上两个线性函数 $f$ 与 $g$ 有相同的核，则 $f=ag$，其中 $a\in F$ 且 $a\neq 0$.

7. 设 $V=R^n$ 是 $n$ 维欧氏空间(见第四章 §10 的定义 3)，其内积为

$$(\alpha,\beta)=a_1b_1+a_2b_2+\cdots+a_nb_n$$

其中 $\alpha=(a_1,a_2,\cdots,a_n)$，$\beta=(b_1,b_2,\cdots,b_n)$. 对于 $V$ 中确定的向量 $\alpha$，定义 $V$ 上的一个函数 $\alpha^*$ 如下：

$$\alpha^*(\beta) = (\alpha,\beta), \quad \forall\, \beta \in V$$

(1) 证明 $\alpha^*$ 是 $V$ 上的线性函数；

*(2) 证明：$V$ 到 $V^*$ 的映射 $\sigma: \alpha \longmapsto \alpha^*$ 是一个同构映射；

*(3) 在 $V$ 中任取一个标准正交基 $\eta_1,\cdots,\eta_n$. 用定理 11.2.2 后面一段中讲的方法确定出 $V$ 到 $V^*$ 的一个同构映射 $\tau$，证明：$\tau = \sigma$；从而 $V$ 到 $V^*$ 的由上述方法确定的同构映射不依赖于 $V$ 中的标准正交基的选择，因此它是自然同构. 于是可把 $V$ 与 $V^*$ 等同，即可把欧氏空间 $R^n$ 看成自身的对偶空间.

8. 设 $\underline{A}$ 是域 $F$ 上 $n$ 维线性空间 $V$ 上的一个线性变换.

(1) 证明：对于 $f \in V^*$，有 $f\underline{A} \in V^*$；

(2) 定义 $V^*$ 到自身的一个映射 $\underline{A}^*$ 为

$$f \longmapsto f\underline{A}$$

证明 $\underline{A}^*$ 是 $V^*$ 上的线性变换；

(3) 设 $\alpha_1,\cdots,\alpha_n$ 是 $V$ 的一个基，$f_1,\cdots,f_n$ 是它的对偶基，并且 $\underline{A}$ 在 $\alpha_1,\cdots,\alpha_n$ 下的矩阵为 $A$，证明：$\underline{A}^*$ 在 $f_1,\cdots,f_n$ 下的矩阵为 $A'$（因此 $\underline{A}^*$ 称为 $\underline{A}$ 的转置映射）.

9. 设 $V$ 是域 $F$ 上一个线性空间，$f_1,\cdots,f_s$ 是 $V$ 上 $s$ 个线性函数，

(1) 证明下述集合

$$W = \{\alpha \in V \mid f_i(\alpha) = 0, i = 1,\cdots,s\}$$

是 $V$ 的一个子空间. $W$ 称为线性函数 $f_1,\cdots,f_s$ 的零化子空间.

*(2) 证明：若 $V$ 是有限维的，则 $V$ 的任一子空间都是某些线性函数的零化子空间.

*10. 设 $V$ 是域 $F$ 上一个线性空间，$\mathrm{char}F = 0$. 设 $\alpha_1,\alpha_2,\cdots,\alpha_s$ 是 $V$ 中非零向量，证明：存在 $f \in V^*$，使得

$$f(\alpha_i) \neq 0, \quad i = 1,2,\cdots,s$$

（提示：把 $V$ 与 $V^{**}$ 等同，然后用习题 11.1 的第 4 题的结论.）

## §3 双线性函数

在 §1 和 §2 我们讨论了域 $F$ 上线性空间 $V$ 上的线性函数，它是 $V$ 到 $F$ 的映射，并且保持加法和纯量乘法. 在许多问题中我们还

会遇到$V\times V$到$F$的映射.譬如,在解析几何中,两个向量$\vec{a}$与$\vec{b}$的内积是几何空间$V$与自身的笛卡儿积$V\times V$到$R$的一个映射:$(\vec{a},\vec{b})\longmapsto\vec{a}\cdot\vec{b}$.根据内积的性质,我们有

$$\vec{a}\cdot(k_1\vec{b}_1+k_2\vec{b}_2)=k_1\vec{a}\cdot\vec{b}_1+k_2\vec{a}\cdot\vec{b}_2 \tag{1}$$

$$(k_1\vec{a}_1+k_2\vec{a}_2)\cdot\vec{b}=k_1\vec{a}_1\cdot\vec{b}+k_2\vec{a}_2\cdot\vec{b} \tag{2}$$

如果我们把上述映射记作$f$,则(1)和(2)可写成

$$f(\vec{a},k_1\vec{b}_1+k_2\vec{b}_2)=k_1f(\vec{a},\vec{b}_1)+k_2f(\vec{a},\vec{b}_2) \tag{3}$$

$$f(k_1\vec{a}_1+k_2\vec{a}_2,\vec{b})=k_1f(\vec{a}_1,\vec{b})+k_2f(\vec{a}_2,\vec{b}) \tag{4}$$

从这个例子我们抽象出线性空间$V$的双线性函数的概念:

**定义** 1 设$V$是域$F$上的一个线性空间,$V\times V$到$F$的一个映射$f$,如果满足

(i) $f(\alpha,k_1\beta_1+k_2\beta_2)=k_1f(\alpha,\beta_1)+k_2f(\alpha,\beta_2)$

(ii) $f(k_1\alpha_1+k_2\alpha_2,\beta)=k_1f(\alpha_1,\beta)+k_2f(\alpha_2,\beta)$

其中$\alpha,\alpha_1,\alpha_2,\beta,\beta_1,\beta_2$是$V$中任意向量,$k_1,k_2$是$F$中任意元素,那么称$f$是$V$上的一个**双线性函数**.$f$也写成$f(\alpha,\beta)$.

条件(i)表明:当$\alpha$固定时,映射$\beta\longmapsto f(\alpha,\beta)$是$V$上的一个线性函数,把这个线性函数记成$\alpha_L$;

条件(ii)表明:当$\beta$固定时,映射$\alpha\longmapsto f(\alpha,\beta)$是$V$上的一个线性函数,记作$\beta_R$.

这就是"双线性函数"这一词的来由.

**例** 1 欧氏空间$R^n$的内积$(\alpha,\beta)$是$R^n$上的一个双线性函数.(见第四章 §10).

**例** 2 设$V=M_n(F)$,令

$$f(A,B)=Tr(AB),\quad\forall\ A,B\in V$$

则$f$是$V$上的一个双线性函数.

**例** 3 设$V=C[a,b]$,令

$$f(g(x),h(x))=\int_a^b g(x)h(x)\mathrm{d}x,\quad\forall\ g(x),h(x)\in V$$

则$f$是$V$上的一个双线性函数.

现在设$V$是$n$维的，设$\alpha_1,\cdots,\alpha_n$是$V$的一个基，$f$是$V$上的一个双线性函数. 设

$$\alpha = x_1\alpha_1 + \cdots + x_n\alpha_n, \quad \beta = y_1\alpha_1 + \cdots + y_n\alpha_n$$

则
$$f(\alpha,\beta) = f(\sum_{i=1}^{n} x_i\alpha_i, \sum_{j=1}^{n} y_j\alpha_j) = \sum_{i=1}^{n}\sum_{j=1}^{n} x_i y_j f(\alpha_i,\alpha_j) \tag{5}$$

(5)式是双线性函数$f$在基$\alpha_1,\cdots,\alpha_n$下的表达式. 从(5)式得出下面的性质1和性质2.

1. $V$上的一个双线性函数$f$完全被它在$V$的一个基的向量组成的有序对上的函数值$f(\alpha_i,\alpha_j), i,j = 1,2,\cdots,n$，所确定. 即如果$V$上的两个双线性函数$f,g$满足

$$f(\alpha_i,\alpha_j) = g(\alpha_i,\alpha_j), \quad i,j = 1,2,\cdots,n$$

则$f = g$.

2. 对于$V$上的双线性函数$f$，构造一个$n$级矩阵$A$：

$$A = \begin{pmatrix} f(\alpha_1,\alpha_1) & f(\alpha_1,\alpha_2) & \cdots & f(\alpha_1,\alpha_n) \\ f(\alpha_2,\alpha_1) & f(\alpha_2,\alpha_2) & \cdots & f(\alpha_2,\alpha_n) \\ \cdots & \cdots & \cdots & \cdots \\ f(\alpha_n,\alpha_1) & f(\alpha_n,\alpha_2) & \cdots & f(\alpha_n,\alpha_n) \end{pmatrix}$$

称$A$是$f$在基$\alpha_1,\alpha_2,\cdots,\alpha_n$下的**度量矩阵**；也称$A$是基$\alpha_1,\alpha_2,\cdots,\alpha_n$在$f$下的度量矩阵. $f$在给定基下的度量矩阵是唯一的. 据1知，不同的双线性函数在同一个基下的度量矩阵一定不同. 利用度量矩阵$A$，我们可以把双线性函数$f$在基$\alpha_1,\alpha_2,\cdots,\alpha_n$下的表达式(5)写成

$$f(\alpha,\beta) = X'AY \tag{6}$$

其中$X' = (x_1,\cdots,x_n)$，$Y' = (y_1,\cdots,y_n)$分别是$\alpha,\beta$在基$\alpha_1,\alpha_2,\cdots,\alpha_n$下的坐标.

3. 反之，任给域$F$上一个$n$级矩阵$A = (a_{ij})$，在$V$中取定一个基$\alpha_1,\alpha_2,\cdots,\alpha_n$，定义$V \times V$到$F$的一个映射$f$如下：

$$f(\alpha,\beta) = \sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}x_i y_j \tag{7}$$

其中 $(x_1,\cdots,x_n)$ 和 $(y_1,\cdots,y_n)$ 分别是 $\alpha,\beta$ 在基 $\alpha_1,\alpha_2,\cdots,\alpha_n$ 下的坐标，则 $f$ 是 $V$ 上的一个双线性函数，并且 $f$ 在基 $\alpha_1,\alpha_2,\cdots,\alpha_n$ 下的度量矩阵恰好是 $A$.

**证明** 令 $X' = (x_1,\cdots,x_n), Y' = (y_1,\cdots,y_n)$. 从(7)式得

$$f(\alpha,\beta) = X'AY \tag{8}$$

从(8)式容易验证 $f$ 是双线性函数.

从(7)式得

$$f(\alpha_i,\alpha_j) = \sum_{k=1}^{n}\sum_{l=1}^{n} a_{kl}x_k y_l = a_{ij} \tag{9}$$

因此 $f$ 在基 $\alpha_1,\cdots,\alpha_n$ 下的度量矩阵正好是 $A$. ▌

以上说明，如果在 $V$ 中取定一个基，那么 $V$ 上全体双线性函数与 $F$ 上全体 $n$ 级矩阵之间有一个一一对应，即让双线性函数 $f$ 对应到它在给定基下的度量矩阵（上述的性质 2 说明这个对应是映射，性质 1 说明这个映射是单射，第 3 点说明这个映射是满射）.

如果把双线性函数 $f$ 的度量矩阵 $A$ 的 $(i.j)$ 元记成 $a_{ij}$，则从(5)式得

$$f(\alpha,\beta) = \sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}x_i y_j \tag{10}$$

(10)式右端的表达式

$$\sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}x_i y_j \tag{11}$$

称为 $x_1,\cdots,x_n$ 与 $y_1,\cdots,y_n$ 的**双线性型**. (10) 式表明，任一双线性函数能够用坐标 $x_1,\cdots,x_n$ 与 $y_1,\cdots,y_n$ 的双线性型表示出. 注意：双线性型指的是表达式(11)，而双线性函数指的是 $V\times V$ 到 $F$ 的映射.

下面我们来讨论 $V$ 上的同一个双线性函数 $f$ 在 $V$ 的不同基下的度量矩阵之间的关系.

**定理 11.3.1** 设 $\alpha_1,\cdots,\alpha_n$ 和 $\eta_1,\cdots,\eta_n$ 是域 $F$ 上线性空间 $V$ 的两个基，它们的关系是

$$(\eta_1,\cdots,\eta_n) = (\alpha_1,\cdots,\alpha_n)C$$

设双线性函数 $f$ 在基 $\{\alpha_i\}$ 和基 $\{\eta_j\}$ 下的度量矩阵分别为 $A$ 和 $B$，则 $B = C'AC$，即同一个双线性函数在不同基下的度量矩阵是合同的.

**证明** 任取 $\alpha,\beta \in V$，设

$$\alpha = (\alpha_1,\cdots,\alpha_n)X = (\eta_1,\cdots,\eta_n)X_1$$
$$\beta = (\alpha_1,\cdots,\alpha_n)Y = (\eta_1,\cdots,\eta_n)Y_1$$

则 $X = CX_1, Y = CY_1$. 于是

$$f(\alpha,\beta) = X'AY = (CX_1)'A(CY_1) = X_1'(C'AC)Y_1$$

因此由前面的第 3 点知道，$C'AC$ 是 $f$ 在基 $\eta_1,\cdots,\eta_n$ 下的度量矩阵. 又 $B$ 也是 $f$ 在基 $\eta_1,\cdots,\eta_n$ 下的度量矩阵，所以 $B = C'AC$. ▌

**推论** 11.3.1 $V$ 的一个双线性函数 $f$ 在 $V$ 的各个基下的度量矩阵组成的集合恰好是 $M_n(F)$ 的一个合同等价类.

证明留给读者作为练习.

由于合同的矩阵有相同的秩，因此我们把双线性函数 $f$ 在一个基下的度量矩阵的秩称为 $f$ 的**矩阵秩**，记作 $\text{rank}_m f$.

**定义** 2 设 $f$ 是域 $F$ 上线性空间 $V$ 上的一个双线性函数，$V$ 的一个子集合

$$\{\alpha \in V \mid f(\alpha,\beta) = 0, \forall\ \beta \in V\}$$

称为 $f$ 的**左根**，记作 $\text{rad}_L V$. $V$ 的另一子集

$$\{\beta \in V \mid f(\alpha,\beta) = 0, \forall\ \alpha \in V\}$$

称为 $f$ 的**右根**，记作 $\text{rad}_R V$.

容易看出，$V$ 上双线性函数 $f$ 的左根和右根都是 $V$ 的子空间.

**定义** 3 如果 $V$ 上双线性函数 $f$ 的左根和右根都是零子空间，那么称 $f$ 是**非退化的**.

如果 $V$ 是有限维的，那么可以用 $f$ 的度量矩阵来判断 $f$ 是不是非退化的.

**定理** 11.3.2 设 $V$ 是域 $F$ 上 $n$ 维线性空间，$f$ 是 $V$ 上的一个双线性函数，设 $f$ 在基 $\alpha_1,\cdots,\alpha_n$ 下的度量矩阵为 $A$，则 $f$ 是非退化的充分必要条件为 $A$ 是可逆矩阵（即满秩矩阵）.

**证明** 先证：$\mathrm{rad}_L V = 0 \Leftrightarrow A$ 满秩. 设

$$\alpha = (\alpha_1, \cdots, \alpha_n)X, \quad \beta = (\alpha_1, \cdots, \alpha_n)Y$$

则 $f(\alpha, \beta) = X'AY$. 从而

$$\begin{aligned}
&\mathrm{rad}_L V = 0 \\
&\Leftrightarrow \text{从 } f(\alpha,\beta) = 0, \forall\, \beta \in V \text{ 可推出 } \alpha = 0 \\
&\Leftrightarrow \text{从 } X'AY = 0, \forall\, Y \in F^n \text{ 可推出 } X = 0 \\
&\Leftrightarrow \text{从 } X'A = 0 \text{ 可推出 } X = 0 \\
&\Leftrightarrow \text{从 } A'X = 0 \text{ 可推出 } X = 0 \\
&\Leftrightarrow A'X = 0 \text{ 只有零解} \\
&\Leftrightarrow \mathrm{rank} A' = n \\
&\Leftrightarrow \mathrm{rank} A = n
\end{aligned}$$

同理可证：$\mathrm{rad}_R V = 0 \Leftrightarrow A$ 满秩. 因此，$f$ 非退化 $\Leftrightarrow A$ 满秩. ▌

**推论** 11.3.1 设 $f$ 是域 $F$ 上有限维线性空间 $V$ 上的一个双线性函数，则 $f$ 的左根等于零子空间当且仅当 $f$ 的右根等于零子空间. ▌

## 习 题 11.3

1. 设 $A$ 是域 $F$ 上一个 $m$ 级矩阵，设 $V = M_{m\times n}(F)$，定义 $V \times V$ 到 $F$ 的一个映射

$$f(X,Y) = Tr(X'AY), \quad X, Y \in M_{m\times n}(F)$$

(1) 证明 $f$ 是 $M_{m\times n}(F)$ 上的一个双线性函数；

(2) 求 $f$ 在基 $E_{11}, \cdots, E_{1n}, \cdots, E_{m1}, \cdots, E_{mn}$ 下的度量矩阵.

2. 在 $K^4$ 中，设 $\alpha = (x_1, x_2, x_3, x_4)$，$\beta = (y_1, y_2, y_3, y_4)$. 定义 $K^4 \times K^4$ 到 $K$ 的一个映射

$$f(\alpha, \beta) = x_1y_1 - x_2y_2 + x_3y_3 - x_4y_4$$

(1) 证明 $f$ 是 $K^4$ 上的一个双线性函数；

(2) 求 $f$ 在标准基 $\varepsilon_1, \varepsilon_2, \varepsilon_3, \varepsilon_4$ 下的度量矩阵；

(3) 求 $f$ 在基

$$\eta_1=(2,1,-1,1),\qquad \eta_2=(0,3,1,0)$$
$$\eta_3=(5,3,2,1),\qquad \eta_4=(6,6,1,3)$$
下的度量矩阵；

(4) 证明 $f$ 是非退化的；

(5) 求一个向量 $\alpha\neq 0$，使 $f(\alpha,\alpha)=0$.

3. 在 $K^4$ 中，设 $\alpha=(x_1,x_2,x_3,x_4)$，$\beta=(y_1,y_2,y_3,y_4)$，定义 $K^4$ 上的一个双线性函数
$$f(\alpha,\beta)=x_1y_2-2x_2y_1+x_3y_4-3x_4y_2$$
求 $f$ 在基
$$\alpha_1=(1,2,1,1),\qquad \alpha_2=(2,3,1,0)$$
$$\alpha_3=(3,1,1,-2),\qquad \alpha_4=(4,2,-1,-6)$$
下的度量矩阵.

4. 证明：当 $V$ 是有限维线性空间时，$V$ 上的双线性函数 $f$ 的左根与右根的维数相同，都等于 $\dim V-\operatorname{rank}_m f$.

*5. 设 $f$ 是域 $F$ 上 $n$ 维线性空间 $V$ 上的一个双线性函数.

(1) 证明：映射 $L_f:\alpha\longmapsto\alpha_L$ 是 $V$ 到 $V^*$ 的一个线性映射；映射 $R_f:\beta\longmapsto\beta_R$ 是 $V$ 到 $V^*$ 的一个线性映射；

(2) 证明：$\operatorname{Ker}L_f=\operatorname{rad}_L V$，$\operatorname{Ker}R_f=\operatorname{rad}_R V$；

(3) 证明：$\operatorname{rank}L_f=\operatorname{rank}_m f=\operatorname{rank}R_f$；

(4) 证明：$f$ 是非退化的充分必要条件为 $L_f$(或 $R_f$) 是线性空间 $V$ 到 $V^*$ 的同构映射.

6. 证明：$M_n(F)$ 上的双线性函数 $f(A,B)=Tr(AB)$ 是非退化的.

## §4 对称双线性函数·斜对称双线性函数

这一节我们介绍两类重要的常用的双线性函数.

**定义 1** 设 $f$ 是域 $F$ 上线性空间 $V$ 的一个双线性函数，如果
$$f(\alpha,\beta)=f(\beta,\alpha),\quad \forall\,\alpha,\beta\in V \tag{1}$$
则称 $f$ 是对称的；如果
$$f(\alpha,\beta)=-f(\beta,\alpha),\quad \forall\,\alpha,\beta\in V \tag{2}$$

则称 $f$ 是斜对称的(或反对称的).

**例** 1　几何空间中的内积是对称双线性函数,因为

$$\vec{a}\cdot\vec{b}=\vec{b}\cdot\vec{a}$$

**例** 2　欧氏空间 $R^n$ 的内积 $(\alpha,\beta)$ 是对称双线性函数.

**例** 3　在 $M_n(F)$ 中,令

$$f(A,B)=Tr(AB),\quad \forall\ A,B\in M_n(F)$$

则 $f$ 是对称双线性函数.

**例** 4　在 $C[a,b]$ 中,令

$$f(g(x),h(x))=\int_a^b g(x)h(x)\mathrm{d}x$$

则 $f$ 是对称双线性函数.

**例** 5　在 $R^2$ 中,对于任意 $\alpha=(x_1,x_2),\beta=(y_1,y_2)$,令

$$f(\alpha,\beta)=x_1y_2-x_2y_1$$

则 $f$ 是斜对称双线性函数.

对称(或斜对称)双线性函数的度量矩阵有什么特点?

设 $f$ 是域 $F$ 上 $n$ 维线性空间 $V$ 的一个双线性函数. 在 $V$ 中取一个基 $\alpha_1,\cdots,\alpha_n$,设 $f$ 在这个基下的度量矩阵为 $A$.

如果 $f$ 是对称的,则

$$f(\alpha_i,\alpha_j)=f(\alpha_j,\alpha_i),\quad i,j=1,\cdots,n$$

因此 $A$ 是对称矩阵.

反之,如果 $A$ 是对称矩阵,那么对于 $V$ 中任意两个向量

$$\alpha=(\alpha_1,\cdots,\alpha_n)X,\beta=(\alpha_1,\cdots,\alpha_n)Y$$

有 $$f(\alpha,\beta)=X'AY=(X'AY)'=Y'AX=f(\beta,\alpha)$$

因此 $f$ 是对称的.

这证明了:$f$ 是对称的当且仅当它在 $V$ 的任意一个基下的度量矩阵是对称的.

同理可证,$f$ 是斜对称的当且仅当它在 $V$ 的任意一个基下的度量矩阵是斜对称的.

现在我们来讨论对称双线性函数的度量矩阵具有怎样的简单

形式.

**定理** 11.4.1 设 $f$ 是数域 $K$ 上 $n$ 维线性空间 $V$ 上的一个对称双线性函数，则 $V$ 中存在一个基，使得 $f$ 在这个基下的度量矩阵为对角矩阵.

**证明** 任取 $V$ 的一个基 $\alpha_1,\cdots,\alpha_n$，设 $f$ 在此基下的度量矩阵为 $A$，则 $A$ 是对称矩阵. 据第六章 §2 的定理 6.2.2，$A$ 合同于一个对角矩阵 $D$，即存在可逆矩阵 $C$，使得 $D=C'AC$. 令

$$(\eta_1,\cdots,\eta_n)=(\alpha_1,\cdots,\alpha_n)C$$

则 $\eta_1,\cdots,\eta_n$ 也是 $V$ 的一个基. 据定理 11.3.1，$f$ 在基 $\eta_1,\cdots,\eta_n$ 下的度量矩阵为 $C'AC$，这是对角矩阵 $D$. ▌

现在我们把上述结果推广到特征不为 2 的任意域上的有限维线性空间中. 一种方法是先把第六章 §2 的定理 6.2.1 和定理 6.2.2 推广到特征不为 2 的域上. 另一种方法即现在讲的方法.

**定理** 11.4.2 设 $f$ 是特征不为 2 的域 $F$ 上的 $n$ 维线性空间 $V$ 上的对称双线性函数，则 $V$ 中存在一个基，使 $f$ 在此基下的度量矩阵为对角矩阵.

**证明** 对维数 $n$ 作归纳法.

$n=1$ 时，$f$ 在任一个基下的矩阵是 1 级矩阵，这是对角矩阵.

假设 $F$ 上 $n-1$ 维线性空间上的对称双线性函数有此性质，现在看 $n$ 维线性空间 $V$ 上的对称双线性函数 $f$.

假如对一切 $\alpha,\beta\in V$，都有 $f(\alpha,\beta)=0$，则结论显然成立. 下面设 $f(\alpha,\beta)$ 不全为零. 这时一定存在一个向量 $\alpha_1\neq 0$，使得

$$f(\alpha_1,\alpha_1)\neq 0$$

否则，若对一切 $\alpha\in V$，都有 $f(\alpha,\alpha)=0$，则对任意 $\alpha,\beta\in V$，有

$$\begin{aligned}0&=f(\alpha+\beta,\alpha+\beta)=f(\alpha,\alpha+\beta)+f(\beta,\alpha+\beta)\\&=f(\alpha,\alpha)+f(\alpha,\beta)+f(\beta,\alpha)+f(\beta,\beta)\\&=2f(\alpha,\beta)\end{aligned}$$

因为 $\mathrm{char}F\neq 2$，所以从上式得 $f(\alpha,\beta)=0$，矛盾. 记 $f(\alpha_1,\alpha_1)=d_1$，把 $\alpha_1$ 扩充成 $V$ 的一个基 $\alpha_1,\alpha_2,\cdots,\alpha_n$. 用类似于 Schmidt 正交化

方法，令

$$\alpha_i' = \alpha_i - \frac{f(\alpha_1, \alpha_i)}{f(\alpha_1, \alpha_1)}\alpha_1, \quad i = 2, \cdots, n \tag{3}$$

直接计算得

$$f(\alpha_1, \alpha_i') = 0, \quad i = 2, \cdots, n \tag{4}$$

从(3)式可得出 $\alpha_1, \alpha_2', \cdots, \alpha_n'$ 仍是 $V$ 的一个基. 令

$$W = \langle \alpha_2', \cdots, \alpha_n' \rangle$$

则 $V = \langle \alpha_1 \rangle \oplus W$，把 $f$ 看成 $W$ 上的双线性函数，它仍是对称的. 据归纳假设，$W$ 中存在一个基 $\eta_2, \cdots, \eta_n$，使得 $f$ 在这个基下的度量矩阵为对角矩阵 $D_1$，即

$$f(\eta_i, \eta_j) = 0, \quad i \neq j, \quad i, j = 2, \cdots, n \tag{5}$$

由(4)式得出，对任意 $\beta \in W$，有 $f(\alpha_1, \beta) = 0$. 从而 $f(\alpha_1, \eta_i) = 0$，$i = 2, \cdots, n$. 于是 $f$ 在 $V$ 的一个基 $\alpha_1, \eta_2, \cdots, \eta_n$ 下的度量矩阵为对角矩阵 $D = \text{diag}\{(d_1), D_1\}$.

据归纳法原理，对一切自然数 $n$，命题成立. ▍

从定理 11.4.2 得出，如果 $V$ 上的对称双线性函数是非退化的，则存在 $V$ 的一个基 $\eta_1, \cdots, \eta_n$，使得

$$f(\eta_i, \eta_i) \neq 0, \quad i = 1, \cdots, n$$

$$f(\eta_i, \eta_j) = 0, \quad i \neq j, \quad i, j = 1, \cdots, n$$

定理 11.4.2 的意义之一在于可简化计算对称双线性函数在任意一对向量上的函数值. 由于一定存在 $V$ 的一个基 $\eta_1, \eta_2, \cdots, \eta_n$，使得 $f$ 在此基下的度量矩阵为对角矩阵 $\text{diag}\{d_1, \cdots, d_n\}$，从而对任意

$$\alpha = x_1\eta_1 + \cdots + x_n\eta_n, \quad \beta = y_1\eta_1 + \cdots + y_n\eta_n$$

有

$$f(\alpha, \beta) = d_1x_1y_1 + d_2x_2y_2 + \cdots + d_nx_ny_n \tag{6}$$

特别地，当 $F$ 为复数域 $C$ 时，由于任一复对称矩阵合同于一个形式为 $\text{diag}\{1, \cdots, 1, 0, \cdots, 0\}$ 的对角矩阵，因此(6) 式可进一步简化成

$$f(\alpha, \beta) = x_1y_1 + \cdots + x_ry_r \tag{7}$$

其中 $0 \leqslant r \leqslant n$.

当 $F$ 为实数域 $R$ 时,(6) 式可进一步简化成

$$f(\alpha,\beta) = x_1y_1 + \cdots + x_py_p - x_{p+1}y_{p+1} - \cdots - x_ry_r \quad (8)$$

其中 $0 \leqslant p \leqslant r \leqslant n$.

现在我们来讨论斜对称双线性函数. 设 $f$ 是域 $F$ 上 $n$ 维线性空间 $V$ 的斜对称双线性函数.

* 如果 $\mathrm{char}F = 2$,则对任意 $\alpha,\beta \in V$,有 $2f(\alpha,\beta) = 0$. 从而 $f(\alpha,\beta) + f(\alpha,\beta) = 0$. 于是 $f(\alpha,\beta) = - f(\alpha,\beta)$. 由于 $f$ 是斜对称的,所以

$$f(\alpha,\beta) = - f(\beta,\alpha) = f(\beta,\alpha)$$

这表明 $f$ 又是对称的. 因此,在特征为 2 的域上的线性空间中,斜对称双线性函数与对称双线性函数是一致的.

下面设 $\mathrm{char}F \neq 2$. 这时对任意 $\alpha \in V$,有 $f(\alpha,\alpha) = - f(\alpha,\alpha)$. 从而 $2f(\alpha,\alpha) = 0$. 由此得出 $f(\alpha,\alpha) = 0$.

**定理** 11.4.3 设 $f$ 是特征不为 2 的域 $F$ 上的 $n$ 维线性空间 $V$ 上的斜对称双线性函数,则存在 $V$ 的一个基 $\varepsilon_1,\varepsilon_{-1},\cdots,\varepsilon_r,\varepsilon_{-r}$, $\eta_1,\cdots,\eta_s$,使得 $f$ 在这个基下的度量矩阵具有形式

$$\mathrm{diag}\left\{\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix},\cdots,\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix},0,\cdots,0\right\} \quad (9)$$

即

$$\begin{cases} f(\varepsilon_i,\varepsilon_{-i}) = 1, & i = 1,\cdots,r \\ f(\varepsilon_i,\varepsilon_j) = 0, & i + j \neq 0 \\ f(\alpha,\eta_k) = 0, & \alpha \in V, \quad k = 1,\cdots,s \end{cases} \quad (10)$$

**证明** 对维数 $n$ 作归纳法.

$n = 1$ 时,$f$ 的度量矩阵为(0).

假设维数小于 $n$ 时,命题成立. 来看 $n$ 维情形.

如果 $f = 0$,则 $f$ 的度量矩阵为 0.

下面设 $f \neq 0$. 这时 $V$ 中一定存在线性无关的向量 $\varepsilon_1,\alpha_2$,使得 $f(\varepsilon_1,\alpha_2) \neq 0$. 理由是:假如对一切线性无关的向量 $\alpha,\beta$,都有

$$f(\alpha,\beta)=0$$

又对一切线性相关的向量 $\alpha,k\alpha$,有

$$f(\alpha,k\alpha)=kf(\alpha,\alpha)=0$$

则 $f=0$,矛盾.令 $\varepsilon_{-1}=f(\varepsilon_1,\alpha_2)^{-1}\alpha_2$,则 $f(\varepsilon_1,\varepsilon_{-1})=1$.显然 $\varepsilon_1$,$\varepsilon_{-1}$ 仍线性无关.把 $\varepsilon_1,\varepsilon_{-1}$ 扩充成 $V$ 的一个基 $\varepsilon_1,\varepsilon_{-1},\beta_3,\cdots,\beta_n$.令

$$\beta_i'=\beta_i-f(\beta_i,\varepsilon_{-1})\varepsilon_1+f(\beta_i,\varepsilon_1)\varepsilon_{-1},\quad i=3,\cdots,n \tag{11}$$

则

$$\begin{aligned} f(\varepsilon_1,\beta_i') &= f(\varepsilon_1,\beta_i)-f(\beta_i,\varepsilon_{-1})f(\varepsilon_1,\varepsilon_1) \\ &\quad +f(\beta_i,\varepsilon_1)f(\varepsilon_1,\varepsilon_{-1}) \\ &= f(\varepsilon_1,\beta_i)+f(\beta_i,\varepsilon_1)=0 \end{aligned} \tag{12}$$

$$\begin{aligned} f(\varepsilon_{-1},\beta_i') &= f(\varepsilon_{-1},\beta_i)-f(\beta_i,\varepsilon_{-1})f(\varepsilon_{-1},\varepsilon_1) \\ &\quad +f(\beta_i,\varepsilon_1)f(\varepsilon_{-1},\varepsilon_{-1}) \\ &= f(\varepsilon_{-1},\beta_i)+f(\beta_i,\varepsilon_{-1})=0 \end{aligned} \tag{13}$$

$i=3,\cdots,n$.

容易从(11)式看出,$\varepsilon_1,\varepsilon_{-1},\beta_3',\cdots,\beta_n'$ 仍是 $V$ 的一个基,令

$$W=\langle\beta_3',\cdots,\beta_n'\rangle$$

则

$$V=\langle\varepsilon_1,\varepsilon_{-1}\rangle\oplus W$$

把 $f$ 看成 $W$ 上的双线性函数,它仍是斜对称的.据归纳假设,存在 $W$ 的一个基

$$\varepsilon_2,\varepsilon_{-2},\cdots,\varepsilon_r,\varepsilon_{-r},\eta_1,\cdots,\eta_3$$

使得 $f$ 在这个基下的度量矩阵为

$$B_1=\operatorname{diag}\left\{\begin{pmatrix}0&1\\-1&0\end{pmatrix},\cdots,\begin{pmatrix}0&1\\-1&0\end{pmatrix},0,\cdots,0\right\} \tag{14}$$

从(12)和(13)式得出,$\forall\ \beta\in W$,有 $f(\varepsilon_1,\beta)=0$ 且 $f(\varepsilon_{-1},\beta)=0$.从而 $f$ 在 $V$ 的一个基

$$\varepsilon_1,\varepsilon_{-1},\varepsilon_2,\varepsilon_{-2},\cdots,\varepsilon_r,\varepsilon_{-r},\eta_1,\cdots,\eta_s$$

下的度量矩阵为

$$B=\operatorname{diag}\left\{\begin{pmatrix}0&1\\-1&0\end{pmatrix},\begin{pmatrix}0&1\\-1&0\end{pmatrix},\cdots,\begin{pmatrix}0&1\\-1&0\end{pmatrix},0,\cdots,0\right\}$$

由归纳法原理，对一切自然数 $n$，命题成立. ▍

从定理 11.4.3 得出，如果 $V$ 上的斜对称双线性函数 $f$ 是非退化的，则存在 $V$ 的一个基，使得 $f$ 在此基下的度量矩阵为

$$\operatorname{diag}\left\{\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \cdots, \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}\right\} \tag{15}$$

从而得出，$\dim V$ 一定是偶数.

## 习　题　11.4

1. 设 $V$ 是复数域上的 $n$ 维线性空间，$n \geqslant 2$. 设 $f$ 是 $V$ 上的一个对称双线性函数.

(1) 证明 $V$ 中有非零向量 $\xi$，使 $f(\xi,\xi) = 0$；

(2) 如果 $f$ 是非退化的，则必有线性无关的向量 $\xi,\eta$，满足：$f(\xi,\eta) = 1$，$f(\xi,\xi) = f(\eta,\eta) = 0$.

2. 设 $f$ 是特征不为 2 的域 $F$ 上线性空间 $V$ 的双线性函数，证明：$f$ 是斜对称的充分必要条件为：对任意 $\alpha \in V$，都有 $f(\alpha,\alpha) = 0$.

3. 首先给出两个概念：设 $V$ 是域 $F$ 上的线性空间，$f$ 是 $V$ 上的对称的或斜对称的双线性函数. 如果 $V$ 中两个向量 $\alpha,\beta$，满足 $f(\alpha,\beta) = 0$，则称 $\alpha$ 与 $\beta$ **正交**. 设 $W$ 是 $V$ 的一个子空间，把 $f$ 看成 $W$ 上的双线性函数，可定义 $W$ 的**左根** $\mathrm{rad}_L W$ 与 $W$ 的**右根** $\mathrm{rad}_R W$，即

$$\mathrm{rad}_L W := \{\alpha \in W \mid f(\alpha,\beta) = 0, \forall\, \beta \in W\}$$

右根可类似定义. 如果 $f$ 是对称的或斜对称的，则 $\mathrm{rad}_L W$ 与 $\mathrm{rad}_R W$ 相等，记成 $\mathrm{rad} W$，称为 $W$ 的**根**.

* 现在设 $\mathrm{char} F \neq 2$，设 $f$ 是 $V$ 上对称的或斜对称的双线性函数. 设 $W$ 是 $V$ 的一个有限维真子空间(即 $W \neq V$). 设 $\xi \in V$，但是 $\xi \notin W$，并且 $\xi$ 与 $\mathrm{rad} W$ 中的向量正交. 证明：在 $W$ 型的线性子簇 $\xi + W$ 中存在向量 $\eta \neq 0$，使得

$$f(\eta,\beta) = 0, \quad \forall\, \beta \in W$$

(提示：分别用定理 11.4.2 与定理 11.4.3 的结论及其证明中所用的技巧.)

4. 设 $V$ 是域 $F$ 上的线性空间，$f$ 是 $V$ 上的对称的或斜对称的双线性函

数. 设 $W$ 是 $V$ 的一个子空间,令

$$W^{\perp}:=\{\alpha\in V\mid f(\alpha,\beta)=0,\forall\ \beta\in W\}$$

证明:(1) $W^{\perp}$ 是 $V$ 的子空间,称 $W^{\perp}$ 是 $W$ 的**正交补**;

(2) $\mathrm{rad}W\subset W^{\perp}$, $\mathrm{rad}W=W\cap W^{\perp}$.

*5. 设 $V$ 是特征不为 2 的域 $F$ 上的 $n$ 维线性空间,$f$ 是 $V$ 上的对称(或斜对称)双线性函数. 设 $W$ 是 $V$ 的一个非平凡子空间. 证明:$f$ 限制在 $W$ 上是非退化的充分必要条件是 $V=W\oplus W^{\perp}$.

6. 设 $V$ 是域 $F$ 上的 $n$ 维线性空间,$f$ 是 $V$ 上的非退化的对称双线性函数,$W$ 是 $V$ 的一个子空间. 证明

(1) $\dim W+\dim W^{\perp}=\dim V$;

(2) $(W^{\perp})^{\perp}=W$.

(提示: (1) 分析 $\alpha\in W^{\perp}$ 的充分必要条件,然后利用齐次线性方程组的理论. (2) 注意利用第(1) 小题的结论.)

7. 证明:特征不为 2 的域 $F$ 上的斜对称矩阵的行列式是 $F$ 中某个元素的平方.

8. 证明:特征不为 2 的域 $F$ 上两个 $n$ 级斜对称矩阵合同的充分必要条件是它们有相同的秩.

## * §5 双线性函数空间

前两节我们主要是讨论单个的双线性函数. 这一节我们来讨论线性空间 $V$ 上所有双线性函数组成的集合.

设 $V$ 是域 $F$ 上一个线性空间,我们把 $V$ 上所有双线性函数组成的集合记作 $T_2(V)$.

对于 $f,g\in T_2(V)$,定义 $f$ 与 $g$ 的和 $f+g$ 为

$$(f+g)(\alpha,\beta):=f(\alpha,\beta)+g(\alpha,\beta),\quad\forall\ \alpha,\beta\in V$$

对于 $f\in T_2(V)$, $k\in F$,定义 $k$ 与 $f$ 的纯量乘积 $kf$ 为

$$(kf)(\alpha,\beta):=kf(\alpha,\beta),\quad\forall\ \alpha,\beta\in V$$

容易验证,$f+g$, $kf$ 都是 $V$ 上的双线性函数;并且 $T_2(V)$ 对于上

述定义的加法和纯量乘法成为域 $F$ 上的一个线性空间. 称 $T_2(V)$ 是 $V$ 上的**双线性函数空间**.

现在设 $V$ 是 $n$ 维的. 由 §3 知道, $V$ 中取定一个基 $\alpha_1,\cdots,\alpha_n$ 后, $V$ 上的双线性函数 $f$ 到它在基 $\alpha_1,\cdots,\alpha_n$ 下的度量矩阵 $A$ 的对应是 $T_2(V)$ 到 $M_n(F)$ 的一个双射. 由于

$$(f+g)(\alpha_i,\alpha_j)=f(\alpha_i,\alpha_j)+g(\alpha_i,\alpha_j)$$

所以 $f+g$ 在基 $\{\alpha_i\}$ 下的度量矩阵 $C$ 等于 $f$ 与 $g$ 在基 $\{\alpha_i\}$ 下的度量矩阵 $A$ 与 $B$ 之和: $C=A+B$. 这说明上述映射 $\sigma$ 保持加法, 即

$$\sigma(f+g)=\sigma(f)+\sigma(g)$$

又由于

$$(kf)(\alpha_i,\alpha_j)=kf(\alpha_i,\alpha_j)$$

所以 $kf$ 在基 $\{\alpha_i\}$ 下的度量矩阵 $D$ 等于 $k$ 与 $f$ 的度量矩阵 $A$ 的纯量乘积 $kA$, 即 $D=kA$. 这说明映射 $\sigma$ 保持纯量乘法, 即

$$\sigma(kf)=k\sigma(f)$$

因此 $\sigma$ 是 $T_2(V)$ 到 $M_n(F)$ 的一个同构映射. 从而

$$T_2(V)\cong M_n(F),\quad \dim T_2(V)=n^2$$

在本节后面的阅读材料中, 我们将找出双线性函数空间 $T_2(V)$ 的一个基, 并且将求出 $V$ 上任一双线性函数 $f$ 在这个基下的坐标. 在阅读材料中, 我们还给出了双线性函数的秩的概念, 讨论了双线性函数的秩与它的矩阵秩之间的关系. 有兴趣的读者可自己看阅读材料.

我们把 $V$ 上所有对称双线性函数组成的集合记作 $S_2(V)$, 把 $V$ 上所有斜对称双线性函数组成的集合记作 $\Lambda_2(V)$.

容易验证, $S_2(V)$ 与 $\Lambda_2(V)$ 都是 $T_2(V)$ 的子空间.

**定理** 11.5.1 设 $V$ 是特征不为 2 的域 $F$ 上的线性空间, 则

$$T_2(V)=S_2(V)\oplus\Lambda_2(V) \tag{1}$$

**证明** 先证 $T_2(V)=S_2(V)+\Lambda_2(V)$. 任取 $f\in T_2(V)$. 令

$$g(\alpha,\beta)=\frac{f(\alpha,\beta)+f(\beta,\alpha)}{2} \tag{2}$$

$$h(\alpha,\beta)=\frac{f(\alpha,\beta)-f(\beta,\alpha)}{2} \tag{3}$$

(注：这里除以 2 是指乘以$(2e)^{-1}$，其中$e$是域$F$中的单位元素). 容易验证$g \in S_2(V), h \in \Lambda_2(V)$，并且

$$f(\alpha,\beta) = g(\alpha,\beta) + h(\alpha,\beta), \quad \forall\ \alpha,\beta \in V \tag{4}$$

因此$T_2(V) = S_2(V) + \Lambda_2(V)$.

再证$S_2(V) \cap \Lambda_2(V) = 0$. 设$f \in S_2(V) \cap \Lambda_2(V)$，则一方面，$f(\alpha,\beta) = f(\beta,\alpha)$；另一方面又有$f(\alpha,\beta) = -f(\beta,\alpha)$. 由此得出$2f(\alpha,\beta) = 0$. 由于$\mathrm{char}F \neq 2$，所以

$$f(\alpha,\beta) = 0, \quad \forall\ \alpha,\beta \in V$$

因此$f = 0$.

综上述得，$T_2(V) = S_2(V) \oplus \Lambda_2(V)$. ▌

下面我们对$S_2(V)$作进一步讨论，将指出对称双线性函数与二次函数之间的联系.

**定义 1** 设$V$是域$F$上的一个线性空间，$V$上的一个函数$q$称为是**二次函数**，如果存在$V$上的一个双线性函数$f$，使得

$$q(\alpha) = f(\alpha,\alpha), \quad \forall\ \alpha \in V \tag{5}$$

以下设$\mathrm{char}F \neq 2$. 把$f$分解成

$$f = g + h, \quad g \in S_2(V), \quad h \in \Lambda_2(V)$$

从(3)式得，$h(\alpha,\alpha) = 0$. 于是$f(\alpha,\alpha) = g(\alpha,\alpha), \forall\ \alpha \in V$. 因此，不妨设定义 1 中的双线性函数$f$是对称的.

显然，给了$V$上一个对称双线性函数$f$，就有唯一的一个二次函数$q$. 反之，给了$V$上的一个二次函数$q$，我们来证明：存在唯一的对称双线性函数$f$，满足(5)式.

**证明** 由定义 1，存在$V$上的一个对称双线性函数$f$，使得$q(\alpha) = f(\alpha,\alpha), \forall\ \alpha \in V$. 由此得出，对一切$\alpha,\beta \in V$，有

$$\begin{aligned}
&\frac{1}{2}[q(\alpha+\beta) - q(\alpha) - q(\beta)] \\
=&\frac{1}{2}[f(\alpha+\beta,\alpha+\beta) - f(\alpha,\alpha) - f(\beta,\beta)] \\
=&f(\alpha,\beta)
\end{aligned} \tag{6}$$

如果还有一个对称双线性函数 $g$ 也满足(5)式，即

$$q(\alpha)=g(\alpha,\alpha)$$

则同理有 $$\frac{1}{2}[q(\alpha+\beta)-q(\alpha)-q(\beta)]=g(\alpha,\beta) \tag{7}$$

比较(6)和(7)式得

$$f(\alpha,\beta)=g(\alpha,\beta),\quad \forall\ \alpha,\beta\in V$$

因此 $f=g$. ▌

上述证明了：让 $V$ 上的一个对称双线性函数 $f$ 对应到由(5)式确定的 $V$ 上的二次函数 $q$，这个对应是 $S_2(V)$ 到 $V$ 上所有二次函数组成的集合之间的一一对应.

现在设 $V$ 是 $n$ 维的. 在 $V$ 中取定一个基 $\alpha_1,\cdots,\alpha_n$，设 $f$ 是 $V$ 上的一个对称双线性函数，它在基 $\alpha_1,\cdots,\alpha_n$ 下的度量矩阵为 $A=(a_{ij})$，则对于任意 $\alpha=\sum_{i=1}^{n}x_i\alpha_i,\beta=\sum_{i=1}^{n}y_i\alpha_i$，有

$$f(\alpha,\beta)=\sum_{i=1}^{n}\sum_{j=1}^{n}a_{ij}x_iy_j \tag{8}$$

从而得出 $$q(\alpha)=\sum_{i=1}^{n}\sum_{j=1}^{n}a_{ij}x_ix_j \tag{9}$$

$q(\alpha)$ 的表达式

$$\sum_{i=1}^{n}\sum_{j=1}^{n}a_{ij}x_ix_j \tag{10}$$

是 $x_1,\cdots,x_n$ 的二次齐次多项式，它就是我们在第六章讨论过的 $n$ 元二次型. 二次型(10)的矩阵是 $A=(a_{ij})$，它就是对双线性函数 $f$ 在基 $\alpha_1,\cdots,\alpha_n$ 下的度量矩阵. 由此看出，我们可以利用对称双线性函数来研究二次型，也可以用二次型的理论来研究对称双线性函数.

二次函数 $q(\alpha)$ 在基 $\alpha_1,\cdots,\alpha_n$ 下的表达式(10)的矩阵 $A$ 也称为二次函数 $q$ 在基 $\alpha_1,\cdots,\alpha_n$ 下的矩阵.

## 习 题 11.5

1. 设 $V$ 是特征不为 2 的域 $F$ 上的 $n$ 维线性空间，$V$ 上的二次函数也可如下定义：在 $V$ 中取一个基 $\alpha_1,\cdots,\alpha_n$，$V$ 到 $F$ 的一个映射 $q$ 称为 $V$ 上的二次函数，如果对于任意 $\alpha=\sum\limits_{i=1}^{n}x_i\alpha_i\in V$，有

$$q(\alpha)=\sum_{i=1}^{n}\sum_{j=1}^{n}a_{ij}x_ix_j,\quad a_{ij}=a_{ji}$$

证明：这个定义与定义 1 等价.

2. 设 $q$ 是域 $F$ 上线性空间 $V$ 上的一个二次函数，$\mathrm{char}F\neq 2$. 证明：

$$q(a\alpha)=a^2q(\alpha),\quad \forall\, a\in F,\quad \forall\, \alpha\in V$$

# 阅读材料十一

在 §3 开始部分我们指出，给了 $V$ 上一个双线性函数 $f(\alpha,\beta)$，可以得到两类线性函数：

$$\alpha_L:\beta\longmapsto f(\alpha,\beta),\qquad \beta_R:\alpha\longmapsto f(\alpha,\beta)$$

现在把问题反过来，给了 $V$ 上的两个线性函数 $g$ 与 $h$，能不能得到 $V$ 上的一个双线性函数？回答是肯定的. 事实上，令

$$f(\alpha,\beta)=g(\alpha)h(\beta),\quad \forall\, \alpha,\beta\in V \tag{1}$$

容易验证 $f(\alpha,\beta)$ 是 $V$ 上的双线性函数，通常把它记成 $g\otimes h$，即

$$(g\otimes h)(\alpha,\beta):=g(\alpha)h(\beta),\quad \forall\, \alpha,\beta\in V \tag{2}$$

把 $g\otimes h$ 称为**线性函数 $g$ 与 $h$ 的张量积**(tensor product).

利用 $V$ 上两个线性函数的张量积是双线性函数这一事实，我们可以找出 $n$ 维向量空间 $V$ 上的双线性函数空间 $T_2(V)$ 的一个基. 先在 $V$ 中取定一个基 $\alpha_1,\cdots,\alpha_n$，它的对偶基为 $f_1,\cdots,f_n$. 考虑下述 $n^2$ 个双线性函数

$$f_i\otimes f_j,\quad i,j=1,2,\cdots,n$$

我们来证明它们组成 $T_2(V)$ 的一个基. 为此只要证明它们线性无关就够了.

假设

$$\sum_{i=1}^{n}\sum_{j=1}^{n}k_{ij}(f_i\otimes f_j)=0 \tag{3}$$

要证所有 $k_{ij}=0$. 设 $\alpha,\beta$ 在基 $\alpha_1,\cdots,\alpha_n$ 下的坐标分别是 $(x_1,\cdots,x_n)$ 与 $(y_1,\cdots,y_n)$，则

$$(f_i\otimes f_j)(\alpha,\beta)=f_i(\alpha)f_i(\beta)=x_iy_j \tag{4}$$

于是
$$k_{ij}(f_i\otimes f_j)(\alpha,\beta)=k_{ij}x_iy_j \tag{5}$$

考虑(3)式两边的函数在 $(\alpha_r,\alpha_s)$ 上的值，得

$$\sum_{i=1}^{n}\sum_{j=1}^{n}k_{ij}(f_i\otimes f_j)(\alpha_r,\alpha_s)=0$$

利用(5)式，上式变成

$$k_{rs}=0 \tag{6}$$

所以 $f_i\otimes f_j, i,j=1,2,\cdots,n$，线性无关.

现在我们来求 $V$ 上任一双线性函数 $f(\alpha,\beta)=\sum\limits_{i=1}^{n}\sum\limits_{j=1}^{n}a_{ij}x_iy_j$ 在 $T_2(V)$ 的上述基 $\{f_i\otimes f_j\}$ 下的坐标，设

$$f=\sum_{i=1}^{n}\sum_{j=1}^{n}k_{ij}(f_i\otimes f_j) \tag{7}$$

考虑它们在 $(\alpha_r,\alpha_s)$ 上的值，得

$$f(\alpha_r,\alpha_s)=k_{rs} \tag{8}$$

由于 $f(\alpha_r,\alpha_s)=a_{rs}$，所以 $a_{rs}=k_{rs}$. 于是得出

$$f(\alpha,\beta)=\sum_{i=1}^{n}\sum_{j=1}^{n}a_{ij}x_iy_j$$

在基 $\{f_i\otimes f_j\}$ 下的坐标为 $(a_{11},\cdots,a_{1n},\cdots,a_{n1},\cdots,a_{nn})$，即

$$f=\sum_{i=1}^{n}\sum_{j=1}^{n}a_{ij}(f_i\otimes f_j) \tag{9}$$

其中 $a_{ij}$ 是 $f$ 在基 $\alpha_1,\cdots,\alpha_n$ 下的度量矩阵 $A$ 的 $(i,j)$ 元.

我们把上面的结果写成一个定理：

**定理 1** 设 $V$ 是域 $F$ 上 $n$ 维线性空间，$\alpha_1,\cdots,\alpha_n$ 是 $V$ 的一个基，它的对偶基是 $f_1,\cdots,f_n$. $V$ 上双线性函数 $f$ 在基 $\alpha_1,\cdots,\alpha_n$ 下的度量矩阵为 $A=(a_{ij})$，则张量积 $f_i\otimes f_j, i,j=1,2,\cdots,n$，组成 $T_2(V)$ 的一个基；$f$ 在这个基下的坐标的第 $ij$ 个分量为 $A$ 的 $(i,j)$ 元 $a_{ij}$，即

$$f = \sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij}(f_i \otimes f_j)$$ ▌

下面我们来给出双线性函数的秩的概念.

**定义** 1　设 $f$ 是域 $F$ 上线性空间 $V$ 上的一个双线性函数. 任意给定 $\alpha \in V$,线性函数

$$\alpha_L : \beta \longmapsto f(\alpha, \beta)$$

称为 **伴随 $f$ 的线性函数**. 任意给定 $\beta \in V$,

$$\beta_R : \alpha \longmapsto f(\alpha, \beta)$$

也称为 **伴随 $f$ 的线性函数**.

**定义** 2　设 $f$ 是线性空间 $V$ 上的一个双线性函数. 伴随 $f$ 的所有线性函数生成的子空间(它是 $V^*$ 的子空间)称为 $f$ 的**秩空间**.

**定义** 3　设 $f$ 是线性空间 $V$ 上的一个双线性函数,$f$ 的秩空间的维数称为 $f$ 的**秩**,记作 $\operatorname{rank} f$.

如何求出双线性函数 $f$ 的秩?$f$ 的秩与 $f$ 的矩阵秩是什么关系?下面我们来讨论这两个问题.

设 $\xi_1, \xi_2, \cdots, \xi_r$ 是 $V$ 上的一组线性函数. 如果 $V$ 上的双线性函数 $f$ 能表示成

$$f = \sum_{i=1}^{r} \sum_{j=1}^{r} b_{ij} \xi_i \otimes \xi_j \tag{10}$$

那么称 $f$ **能用 $\xi_1, \xi_2, \cdots, \xi_r$ 张量形式表示**.

**命题** 1　设 $V$ 上的双线性函数 $f$ 能用 $V$ 上的线性函数 $\xi_1, \xi_2, \cdots, \xi_r$ 张量形式表出,则 $f$ 的秩空间 $W$ 含于 $\xi_1, \xi_2, \cdots, \xi_r$ 生成的子空间,即

$$W \subset \langle \xi_1, \xi_2, \cdots, \xi_r \rangle$$

**证明**　由已知条件得到(10)式. 任取一个伴随 $f$ 的线性函数

$$\alpha_L : \beta \longmapsto f(\alpha, \beta)$$

我们有

$$\begin{aligned}
\alpha_L(\beta) = f(\alpha, \beta) &= \sum_{i=1}^{r} \sum_{j=1}^{r} b_{ij} \xi_i(\alpha) \xi_j(\beta) \\
&= \sum_{j=1}^{r} \Big( \sum_{i=1}^{r} b_{ij} \xi_i(\alpha) \Big) \xi_j(\beta) = \sum_{j=1}^{r} c_j \xi_j(\beta) \\
&= \Big( \sum_{j=1}^{r} c_j \xi_j \Big) \beta, \quad \forall\, \beta \in V
\end{aligned}$$

其中 $c_j = \sum_{i=1}^{r} b_{ij} \xi_i(\alpha)$,因此 $\alpha_L = \sum_{j=1}^{r} c_j \xi_j$. 这说明 $\alpha_L \in \langle \xi_1, \xi_2, \cdots, \xi_r \rangle$. 同理可证

$\beta_R \in \langle \xi_1, \xi_2, \cdots, \xi_r \rangle, \forall\, \beta \in V$. 因此 $W \subset \langle \xi_1, \xi_2, \cdots, \xi_r \rangle$. ▌

**推论** 1　设 $V$ 上的双线性函数 $f$ 能用 $V$ 上的线性函数 $\xi_1, \xi_2, \cdots, \xi_r$ 张量形式表出，则 $f$ 的秩不超过 $r$.

**证明**　由命题 1 得

$$\operatorname{rank} f = \dim W \leqslant \dim \langle \xi_1, \xi_2, \cdots, \xi_r \rangle \leqslant r$$ ▌

**命题** 2　$n$ 维线性空间 $V$ 上的双线性函数 $f$ 能够用它的秩空间 $W$ 的任意一个基张量形式表出.

**证明**　设 $\xi_1, \cdots, \xi_r$ 是 $W$ 的一个基，把它扩充成 $V^*$ 的一个基 $\xi_1, \cdots, \xi_r, \xi_{r+1}, \cdots, \xi_n$. 设它的对偶基为 $\alpha_1^{**}, \cdots, \alpha_r^{**}, \alpha_{r+1}^{**}, \cdots, \alpha_n^{**}$，这是 $V^{**}$ 的一个基，把 $V^{**}$ 与 $V$ 等同，则 $\alpha_1, \cdots, \alpha_r, \alpha_{r+1}, \cdots, \alpha_n$ 是 $V$ 的一个基. 当 $r < j \leqslant n$ 时，我们有

$$\xi_i(\alpha_j) = \alpha_j^{**}(\xi_i) = 0, \quad 1 \leqslant i \leqslant r$$

从而对于 $W$ 中任一元素 $\xi$，有

$$\xi(\alpha_j) = 0, \quad r < j \leqslant n$$

特别地，由于 $\alpha_L \in W$，所以 $\alpha_L(\alpha_j) = 0, r < j \leqslant n$，即 $f(\alpha, \alpha_j) = 0, r < j \leqslant n$. 同理，由于 $\beta_R \in W$，所以 $\beta_R(\alpha_j) = 0$，从而 $f(\alpha_j, \beta) = 0, r < j \leqslant n$，因此 $f$ 在基 $\alpha_1, \cdots, \alpha_n$ 下的度量矩阵 $A = (a_{ij})$ 形如

$$A = \begin{pmatrix} A_1 & 0 \\ 0 & 0 \end{pmatrix} \tag{11}$$

其中 $A_1 = (f(\alpha_k, \alpha_l))$ 是 $r$ 级矩阵，$1 \leqslant k, l \leqslant r$. 因为

$$\xi_k(\alpha_i) = \alpha_i^{**}(\xi_k) = \delta_{ik}$$

所以 $\xi_1, \cdots, \xi_n$ 是 $\alpha_1, \cdots, \alpha_n$ 的对偶基. 据定理 1 得

$$f = \sum_{i=1}^{r} \sum_{j=1}^{r} a_{ij} \xi_i \otimes \xi_j$$

这表明 $f$ 能用 $W$ 的一个基 $\xi_1, \cdots, \xi_r$ 张量形式表出. ▌

从推论 1 和命题 2 得到

**定理** 2　$n$ 维线性空间 $V$ 上的双线性函数 $f$ 的秩等于能够用张量形式表出 $f$ 的 $V$ 上线性函数的最小数目. ▌

从命题 2 的证明过程中看到

$$\operatorname{rank}_m f = \operatorname{rank} A = \operatorname{rank} A_1 \leqslant r = \dim W = \operatorname{rank} f$$

因此我们得到

**定理** 3　$n$ 维线性空间 $V$ 上的双线性函数 $f$ 的矩阵秩不超过 $f$ 的秩. ▌

的确存在双线性函数 $f$，它的矩阵秩小于它的秩. 例如，在 2 维线性空间 $V$ 中取一个基 $\alpha_1,\alpha_2$，设它的对偶基为 $f_1,f_2$. 设 $V$ 上的一个双线性函数 $f=f_1\otimes f_2$，则由定理 1 可立即写出 $f$ 在基 $\alpha_1,\alpha_2$ 下的度量矩阵为

$$A=\begin{pmatrix}0 & 1\\ 0 & 0\end{pmatrix} \tag{12}$$

因此 $f$ 的矩阵秩等于 1. 由于 $f$ 可用 $f_1,f_2$ 张量形式表出，因此 $\operatorname{rank}f\leqslant 2$. 我们指出，$f$ 不可能用一个线性函数 $\xi_1$ 张量形式表出，否则 $f=a\xi_1\otimes\xi_1$. 这时对任意 $\alpha,\beta\in V$ 有

$$f(\alpha,\beta)=a\xi_1(\alpha)\xi_1(\beta)=f(\beta,\alpha)$$

于是推出 $f(\alpha_1,\alpha_2)=f(\alpha_2,\alpha_1)$，从而 $A$ 应为对称矩阵，这与(12)矛盾. 所以从定理 2 得出，$\operatorname{rank}f=2$.

对于对称双线性函数，则它的矩阵秩等于它的秩，即

**定理** 4　设 $f$ 是 $n$ 维线性空间 $V$ 上的对称双线性函数，则

$$\operatorname{rank}_m f=\operatorname{rank}f$$

**证明**　由于 $f$ 是对称的，所以

$$\beta_R:\alpha\longmapsto f(\alpha,\beta)=f(\beta,\alpha)$$

即 $\beta_R=\beta_L$. 因此伴随 $f$ 的每个线性函数有形式

$$\alpha_L:\beta\longmapsto f(\alpha,\beta)$$

在 $V$ 中取一个基 $\alpha_1,\cdots,\alpha_n$. 设 $\alpha=\sum\limits_{i=1}^{n}x_i\alpha_i$，则

$$\alpha_L(\beta)=f(\alpha,\beta)=\sum_{i=1}^{n}x_if(\alpha_i,\beta)=\sum_{i=1}^{n}x_i\alpha_{i_L}(\beta)$$

从而得 $\alpha_L=\sum\limits_{i=1}^{n}x_i\alpha_{i_L}$. 这表明 $f$ 的秩空间 $W$ 为

$$W=\langle\alpha_{1_L},\alpha_{2_L},\cdots,\alpha_{n_L}\rangle$$

因此 $\dim W$ 等于向量组 $\alpha_{1_L},\cdots,\alpha_{n_L}$ 的秩.

设 $\alpha_1,\cdots,\alpha_n$ 的对偶基为 $f_1,\cdots,f_n$. 因为

$$\alpha_{i_L}(\beta)=f(\alpha_i,\beta)=\sum_{j=1}^{n}f(\alpha_i,\alpha_j)f_j(\beta)=\Big[\sum_{j=1}^{n}f(\alpha_i,\alpha_j)f_j\Big](\beta)$$

所以

$$\alpha_{i_L}=\sum_{j=1}^{n}f(\alpha_i,\alpha_j)f_j$$

用 $A$ 表示 $f$ 在基 $\alpha_1,\cdots,\alpha_n$ 下的度量矩阵，上式表明 $\alpha_{i_L}$ 在基 $f_1,\cdots,f_n$ 下的坐

标为 $A$ 的第 $i$ 行. 因此 $\dim\langle\alpha_{1_L},\cdots,\alpha_{n_L}\rangle$ 等于 $\mathrm{rank}A$. 即

$$\mathrm{rank}f = \mathrm{rank}A = \mathrm{rank}_m f$$ ▌

定理 4 对于斜对称双线性函数也成立，证明方法几乎一样，仅有的改动之处为：$\beta_R = (-\beta)_L$.

# 补充题十一

1. 设 $V$ 是实数域上 $n$ 维线性空间，$q$ 是 $V$ 的二次函数. 如果 $q(\xi) = 0$，则称 $\xi$ 是 $q$ 的**零向量**. 证明：如果 $q(\alpha)$ 的表达式是不定的二次型，则 $V$ 中存在由 $q$ 的零向量组成的一个基.

（提示：由已知条件，$V$ 中存在一个基 $\alpha_1,\cdots,\alpha_n$，使得 $q(\alpha)$ 的表达式为

$$q(\alpha) = x_1^2 + \cdots + x_p^2 - x_{p+1}^2 - \cdots - x_r^2$$

其中 $p \geqslant 1, p < r \leqslant n$. 令

$$\eta_i = \alpha_i + \alpha_{p+1}, \quad i = 1,\cdots,p$$
$$\eta_j = -\alpha_1 + \alpha_j, \quad j = p+1,\cdots,r$$
$$\eta_k = \alpha_k, \quad k = r+1,\cdots,n$$

去证 $\eta_1,\cdots,\eta_n$ 是 $V$ 的一个基，这只要证 $\alpha_1,\cdots,\alpha_n$ 可以由 $\eta_1,\cdots,\eta_n$ 线性表出. 最后证 $q(\eta_i) = 0, i = 1,\cdots,n$.）

2. 设 $V$ 是实数域上 $n$ 维线性空间，$V$ 的二次函数 $q$ 的所有零向量组成的集合 $S$ 称为 $q$ 的**零锥**. 证明：$q$ 的零锥 $S$ 是 $V$ 的子空间的充分必要条件是，$q(\alpha)$ 的表达式为半正定的或者半负定的二次型.

（提示：必要性利用第 1 题的结果.）

3. 设 $q$ 是欧几里得空间 $R^n$ 的一个二次函数. 证明：$q$ 的零锥 $S$ 包含 $R^n$ 的一个标准正交基的充分必要条件是，$q$ 在 $R^n$ 的某一个标准正交基（从而在 $V$ 的任一标准正交基）下的矩阵的迹等于零.

（提示：证充分性时，对 $n$ 用归纳法. 设 $q$ 在 $R^n$ 的一个标准正交基 $\alpha_1,\cdots,\alpha_n$ 下的矩阵 $A$ 的迹等于零. 假如 $\alpha_1,\cdots,\alpha_n$ 不全属于 $S$，则可证存在 $\alpha_i,\alpha_j$，使 $q(\alpha_i) > 0, q(\alpha_j) < 0$. 令 $\eta_1 = \alpha_i + \lambda\alpha_j$ 为单位向量且使 $q(\eta_1) = 0$，其中 $\lambda$ 待定. 然后把 $\eta_1$ 扩充成 $V$ 的标准正交基 $\eta_1,\eta_2,\cdots,\eta_n$. 令 $W = \langle\eta_1,\eta_2,\cdots,\eta_n\rangle$.）

4. 用矩阵的术语叙述第 3 题.

# 第十二章　欧几里得空间

迄今为止，我们对于线性空间的讨论主要是围绕它的线性运算（加法与纯量乘法两种运算）进行. 譬如，线性相关性、基、线性子空间（对线性运算封闭的非空子集）、线性映射（一个线性空间 $V$ 到另一个线性空间 $V'$ 的映射，它保持线性运算）、线性变换（$V$ 到自身的线性映射）、线性函数（$V$ 到 $F$ 的线性映射），等等.

从本章开始一直到第十四章，我们将分别在实数域上、复数域上、任意域上的线性空间中引进度量概念.

从解析几何中知道，在几何空间中，引进了向量的内积的概念后，向量的长度、两个向量的夹角这些度量概念就都可以用内积来表示. 而向量的内积是几何空间（三维实线性空间）上的一个正定对称双线性函数（正定性是指对一切 $\vec{\alpha}$ 都有 $\vec{\alpha}\cdot\vec{\alpha}\geqslant 0$，等号成立当且仅当 $\vec{\alpha}=\vec{0}$）. 由此受到启发，在实数域上的线性空间 $V$ 中只要给定了一个正定对称双线性函数，就有可能在 $V$ 中引进长度、角度等度量概念. 我们将按照这样的思路进行讨论.

## §1　内积·实内积空间

**定义 1**　设 $V$ 是实数域 $R$ 上的一个线性空间，$V$ 上的一个正定对称双线性函数 $f$ 称为 $V$ 上的一个**内积**. 所谓 $f$ 是正定的，是指对一切 $\alpha\in V$ 都有 $f(\alpha,\alpha)\geqslant 0$，当且仅当 $\alpha=0$ 时有 $f(\alpha,\alpha)=0$.

习惯上把内积 $f$ 在有序向量对 $(\alpha,\beta)$ 上的函数值 $f(\alpha,\beta)$ 简记成 $(\alpha,\beta)$，或者记成 $(\alpha|\beta)$. 从而把内积 $f$ 记成（ ， ），或者（ | ）. 也可以把内积记成 $(\alpha,\beta)$，或者 $(\alpha|\beta)$. 一般说来，$V$ 上有

许多内积，为了区别$V$上的不同的内积，可以添写下标，例如，写成$(\ ,\ )_1,(\ ,\ )_2,\cdots$等等.

例如，在$R^3$中，对于

$$\alpha=(x_1,x_2,x_3),\quad \beta=(y_1,y_2,y_3)$$

规定 $$(\alpha,\beta)_1:=x_1y_1+x_2y_2+x_3y_3$$

容易验证$(\ ,\ )_1$是$R^3$上的一个正定对称双线性函数，从而它是$R^3$上的一个内积. 如果规定

$$(\alpha,\beta)_2:=x_1y_1+2x_2y_2+3x_3y_3$$

容易验证$(\ ,\ )_2$也是$R^3$上的一个内积.

**定义** 2 设$V$是实数域$R$上的一个线性空间，如果给定了$V$上的一个内积，则称$V$是一个**实内积空间**. 有限维的实内积空间称为**欧几里得空间**，简称为**欧氏空间**，并且把线性空间$V$的维数称为欧氏空间的**维数**.

例如，在$R^3$中，如果给定了一个内积$(\ ,\ )_1$，则$R^3$成为一个欧几里得空间. 如果在$R^3$中给定另一个内积$(\ ,\ )_2$，则$R^3$又成为另一个欧氏空间.

**例** 1 在$R^n$中，对于任意

$$\alpha=(x_1,x_2,\cdots,x_n),\quad \beta=(y_1,y_2,\cdots,y_n)$$

规定 $$(\alpha,\beta):=x_1y_1+x_2y_2+\cdots+x_ny_n \tag{1}$$

容易验证这是一个正定对称双线性函数. 因此它是$R^n$上的一个内积(我们曾经在第四章 §10 讲过这个内积，称它为$R^n$上的标准内积). $R^n$对于这个内积成为一个$n$维欧几里得空间.

**例** 2 在实线性空间$M_{n\times m}(R)$中规定

$$(A,B):=Tr(AB') \tag{2}$$

容易验证这是$M_{n\times m}(R)$上的一个正定对称双线性函数，从而这是$M_{n\times m}(R)$的一个内积. $M_{n\times m}(R)$对于这个内积成为一个$nm$维欧氏空间.

**例** 3 在$C[a,b]$中，规定

$$(f,g):=\int_a^b f(x)g(x)\mathrm{d}x \tag{3}$$

我们已在第十一章 §3 指出这是 $C[a,b]$ 上的一个双线性函数. 显然它是对称的. 容易验证它是正定的. 因此 $(f,g)$ 是 $C[a,b]$ 上的一个内积. $C[a,b]$ 对于这个内积成为一个实内积空间.

**例** 4 $R^\infty$ 中满足 Hilbert 条件(即级数 $\sum\limits_{i=1}^{\infty}a_n^2$ 收敛) 的无限序列 $(a_1,a_2,\cdots)$ 组成的子集 $H$ 是实数域上的线性空间(见习题 8.1 的第 7 题). 对于 $H$ 中任意两个向量

$$\alpha=(a_1,a_2,\cdots),\quad \beta=(b_1,b_2\cdots)$$

规定

$$(\alpha,\beta):=\sum_{i=1}^{\infty}a_ib_i$$

容易验证这是 $H$ 上的一个正定对称双线性函数(注意:从 $|a_ib_i|\leqslant\frac{1}{2}(a_i^2+b_i^2)$ 可推出级数 $\sum\limits_{i=1}^{\infty}a_ib_i$ 收敛). 从而 $H$ 对于这个内积成为一个实内积空间.

在实内积空间 $V$ 中,由于指定了 $V$ 上的一个内积( , ),因此可以引进向量的长度的概念.

**定义** 3 非负实数 $\sqrt{(\alpha,\alpha)}$ 称为向量 $\alpha$ 的**长度**,记作 $|\alpha|$(或者 $\|\alpha\|$).

显然,零向量的长度为 0,非零向量的长度是正数.

**性质** 1

$$|k\alpha|=|k||\alpha|,k\in R,\alpha\in V \tag{4}$$

**证明** $|k\alpha|=\sqrt{(k\alpha,k\alpha)}=\sqrt{k^2(\alpha,\alpha)}=|k||\alpha|.$ ▌

长度为 1 的向量称为**单位向量**. 如果 $\alpha\neq0$,则根据性质 1 得,$\frac{1}{|\alpha|}\alpha$ 是一个单位向量. 把 $\alpha$ 变成 $\frac{1}{|\alpha|}\alpha$ 称为把 $\alpha$ **单位化**.

为了在实内积空间中引进两个向量的夹角的概念,可以从解析几何中受到启发,但是首先要证明

**定理** 12.1.1(Cauchy-Buniakowski 不等式) 在实内积空间

$V$ 中，对于任意向量 $\alpha,\beta$，有

$$|(\alpha,\beta)| \leqslant |\alpha||\beta| \tag{5}$$

等号成立当且仅当 $\alpha,\beta$ 线性相关.

**证明** 如果 $\alpha,\beta$ 线性相关，则 $\alpha=0$ 或者 $\beta=k\alpha$. 若 $\alpha=0$，则

$$|(0,\beta)| = |0(0,\beta)| = 0 = |0||\beta|$$

若 $\beta=k\alpha$，则

$$\begin{aligned}|(\alpha,\beta)| &= |(\alpha,k\alpha)| = |k(\alpha,\alpha)|\\ &= |k||\alpha|^2 = |\alpha||k\alpha| = |\alpha||\beta|\end{aligned}$$

如果 $\alpha,\beta$ 线性无关，则对一切实数 $t$，都有 $t\alpha+\beta\neq 0$. 据内积的正定性得

$$0 < (t\alpha+\beta,t\alpha+\beta) = t^2|\alpha|^2 + 2t(\alpha,\beta) + |\beta|^2 \tag{6}$$

从而(6)式右端的 $t$ 的二次三项式的判别式

$$4(\alpha,\beta)^2 - 4|\alpha|^2|\beta|^2 < 0$$

即 $|(\alpha,\beta)| < |\alpha||\beta|$. ▌

**推论** 12.1.1 对于任意两组实数 $a_1,\cdots,a_n$ 与 $b_1,\cdots,b_n$，有

$$|a_1b_1+\cdots+a_nb_n| \leqslant \sqrt{a_1^2+\cdots+a_n^2}\sqrt{b_1^2+\cdots+b_n^2} \tag{7}$$

(7)式是著名的 Canchy **不等式**. ▌

**推论** 12.1.2 对于任意 $f,g\in C[a,b]$，有

$$\left|\int_a^b f(x)g(x)\mathrm{d}x\right| \leqslant \left(\int_a^b f^2(x)\mathrm{d}x\right)^{\frac{1}{2}}\left(\int_a^b g^2(x)\mathrm{d}x\right)^{\frac{1}{2}} \tag{8}$$

(8)式是著名的 Schwarz **不等式**. ▌

**定义** 4 实内积空间 $V$ 中，两个非零向量 $\alpha,\beta$ 的**夹角** $\langle\alpha,\beta\rangle$ 规定为

$$\langle\alpha,\beta\rangle = \arccos\frac{(\alpha,\beta)}{|\alpha||\beta|} \tag{9}$$

于是 $0\leqslant\langle\alpha,\beta\rangle\leqslant\pi$.

从(9)式看出，$(\alpha,\beta)=0 \Leftrightarrow \langle\alpha,\beta\rangle=\dfrac{\pi}{2}$，于是我们给出下述定义：

**定义 5** 如果 $(\alpha,\beta)=0$，则称 $\alpha$ 与 $\beta$ **正交**，记为 $\alpha\perp\beta$.

由定义 5 以及内积的正定性得出，只有零向量才与自己正交. 从而若 $\alpha$ 与 $V$ 中一切向量都正交，则 $\alpha$ 一定是零向量. 换句话说，对于内积，有 $\mathrm{rad}V=0$，从而内积是 $V$ 上的非退化的双线性函数.

**推论** 12.1.3 在实内积空间 $V$ 中，三角形不等式成立，即对于任意 $\alpha,\beta\in V$，有

$$|\alpha+\beta|\leqslant|\alpha|+|\beta| \tag{10}$$

**证明**

$$\begin{aligned}|\alpha+\beta|^2&=(\alpha+\beta,\alpha+\beta)=|\alpha|^2+2(\alpha,\beta)+|\beta|^2\\&\leqslant|\alpha|^2+2|\alpha||\beta|+|\beta|^2=(|\alpha|+|\beta|)^2\end{aligned}$$

所以 $$|\alpha+\beta|\leqslant|\alpha|+|\beta|$$ ▍

**推论** 12.1.4 在实内积空间 $V$ 中，勾股定理成立，即如果 $\alpha$ 与 $\beta$ 正交，则

$$|\alpha+\beta|^2=|\alpha|^2+|\beta|^2 \tag{11}$$

**证明** $|\alpha+\beta|^2=(\alpha+\beta,\alpha+\beta)=|\alpha|^2+|\beta|^2$. ▍

利用数学归纳法可把勾股定理推广到多个向量的情形，即如果 $\alpha_1,\alpha_2,\cdots,\alpha_n$ 两两正交，则

$$|\alpha_1+\alpha_2+\cdots+\alpha_n|^2=|\alpha_1|^2+|\alpha_2|^2+\cdots+|\alpha_n|^2 \tag{12}$$

**推论** 12.1.5 在实内积空间 $V$ 中，余弦定理成立，即设 $\alpha,\beta,\gamma$ 是一个三角形的三条边，其中 $\gamma=\beta-\alpha$，则

$$|\gamma|^2=|\alpha|^2+|\beta|^2-2|\alpha||\beta|\cos\langle\alpha,\beta\rangle \tag{13}$$

**证明**

$$\begin{aligned}|\gamma|^2&=(\beta-\alpha,\beta-\alpha)=|\beta|^2-2(\alpha,\beta)+|\alpha|^2\\&=|\alpha|^2+|\beta|^2-2|\alpha||\beta|\cos\langle\alpha,\beta\rangle\end{aligned}$$ ▍

在数学分析课程中，读者已看到，在研究无限性的问题时，极限是最重要的概念. 而刻画极限要用到距离的概念. 我们将首先给出距离的定义，然后指出，由于实内积空间中有了向量的长度的概念，因此就可以定义距离.

**定义 6** 设 $E$ 是一个非空集合，$d$ 是 $E\times E$ 到 $R$ 的一个映射，如果对一切 $x,y,z\in E$，有

(i) $d(x,y)=d(y,x)$ （对称性）；

(ii) $d(x,y)\geqslant 0$，等号成立当且仅当 $x=y$ （正定性）；

(iii) $d(x,z)\leqslant d(x,y)+d(y,z)$ （三角形不等式）；

则称 $d$ 是一个**距离**（或**度量** metric）. 如果集合 $E$ 定义了一个距离 $d$，则称 $E$ 是一个**度量空间**. 把 $d(x,y)$ 称为 $x$ **与** $y$ **之间的距离**.

**命题** 12.1.1 在实内积空间 $V$ 中，对于任意 $\alpha,\beta\in V$，定义

$$d(\alpha,\beta):=|\alpha-\beta| \tag{14}$$

则 $d$ 是一个距离. 从而实内积空间 $V$ 对于这个距离 $d$ 成为一个度量空间.

**证明**

$$d(\alpha,\beta)=|\alpha-\beta|=|\beta-\alpha|=d(\beta,\alpha)$$

$$d(\alpha,\beta)=|\alpha-\beta|\geqslant 0$$

等号成立当且仅当 $\alpha-\beta=0$，即 $\alpha=\beta$.

$$\begin{aligned}d(\alpha,\gamma)&=|\alpha-\gamma|=|\alpha-\beta+\beta-\gamma|\\&\leqslant|\alpha-\beta|+|\beta-\gamma|=d(\alpha,\beta)+d(\beta,\gamma)\end{aligned}$$

所以 $d$ 是一个距离. ▌

综上所述，在实线性空间 $V$ 中，给定了一个内积以后，就有了向量的长度、两个非零向量的夹角、正交性、两个向量之间的距离等度量概念.

在实线性空间 $V$ 中，除了通过给定一个内积来引进距离的概念以外，也可以用其他方法引进距离. 我们从命题 12.1.1 的证明过程可看出，用(14)式定义的 $d$ 之所以是一个距离，关键是利用了向量的长度具有性质 1，正定性，以及三角形不等式. 由此受到启发，只要 $V$ 到 $R$ 的一个映射具有这三条性质，那么利用这个映射也可以定义 $V$ 中的一个距离. 为此我们抽象出范数的概念.

**定义 7** 设 $V$ 是实线性空间，$V$ 到 $R$ 的一个映射称为**范数**，用记号 $\|\cdot\|$ 表示，如果对任意 $\alpha,\beta\in V,k\in R$，有

(i) $\|\alpha\| \geqslant 0$，等号成立当且仅当 $\alpha = 0$ （正定性）；

(ii) $\|k\alpha\| = |k|\,\|\alpha\|$ （绝对齐次性）；

(iii) $\|\alpha+\beta\| \leqslant \|\alpha\| + \|\beta\|$ （三角形不等式）.

在实内积空间 $V$ 中，用内积定义的向量的长度满足定义 7 中的三个条件，因此向量的长度是一个范数. 但是 $V$ 中还可以定义其他的范数，见后面的例 5.

**定义** 8　实线性空间 $V$ 中，如果定义了一个范数，则称 $V$ 是**赋范线性空间**，简称为**赋范空间**.

**命题** 12.1.2　在赋范线性空间 $V$ 中，定义

$$d(\alpha,\beta) := \|\alpha-\beta\|, \quad \forall\, \alpha,\beta \in V \tag{15}$$

则 $d$ 是一个距离.

**证明**　与命题 12.1.1 的证明一样. ▌

**例** 5　在实线性空间 $M_{n\times m}(R)$ 中，定义一个内积为

$$(A,B) = Tr(AB') \tag{16}$$

（参看例 2）. 于是 $A = (a_{ij})$ 的长度是一个范数，这个范数记作 $\|\cdot\|_1$. 从(16) 式可得出

$$\|A\|_1 = \sqrt{Tr(AA')} = \sqrt{\sum_{i=1}^{n}\sum_{j=1}^{m} a_{ij}^2} \tag{17}$$

我们定义 $M_{n\times m}(R)$ 到 $R$ 的一个映射 $\|\cdot\|_2$ 如下：

$$\|A\|_2 := \max\{|a_{ij}| \mid i = 1,\cdots,n;\, j = 1,\cdots,m\} \tag{18}$$

容易验证，$\|\cdot\|_2$ 也是 $M_{n\times m}(R)$ 的一个范数.

我们定义 $M_{n\times m}(R)$ 到 $R$ 的另一个映射 $\|\cdot\|_3$ 为

$$\|A\|_3 := \max\{|A\alpha| \mid \alpha \in R^m, |\alpha| = 1\} \tag{19}$$

其中 $R^m$ 和 $R^n$ 都是定义了标准内积的欧氏空间. 由于 $\alpha \longmapsto A\alpha$ 是 $R^m$ 到 $R^n$ 的一个线性映射，并且容易看出它是连续的，从而它在 $R^m$ 中的有界闭集 $\{\alpha \mid |\alpha| = 1, \alpha \in R^m\}$ 上有最大值. 因此(19) 式右端是有意义的. 不难证明，$\|\cdot\|_3$ 是 $M_{n\times m}(R)$ 的一个范数.

**命题** 12.1.3　在实线性空间 $M_{n\times m}(R)$ 中定义的上述三个范

数具有如下关系：

$$\| A \|_2 \leqslant \| A \|_3 \leqslant \| A \|_1, \quad \forall A \in M_{n\times m}(R) \tag{20}$$

证明留给读者.

(提示：设 $\varepsilon_j' = (0,\cdots,0,1,0,\cdots,0)$，从(19)式得 $|A\varepsilon_j| \leqslant \| A \|_3$，从而 $|a_{ij}| \leqslant \| A \|_3, i = 1,\cdots,n; j = 1,\cdots,m$. 于是 $\| A \|_2 \leqslant \| A \|_3$.

利用 $R^m$ 中的 Cauchy 不等式可证出

$$|A\alpha|^2 \leqslant \| A \|_1^2$$

其中 $\alpha \in R^m$ 且 $|\alpha| = 1$. 从而得出 $\| A \|_3 \leqslant \| A \|_1$.)

现在假定实内积空间 $V$ 是有限维的，即考虑欧几里得空间 $V$. 设 $\dim V = n$.

在 $V$ 中取一个基 $\alpha_1,\cdots,\alpha_n$，设 $V$ 中指定的内积(　,　)在这个基下的度量矩阵为 $A = (a_{ij})$，则 $A$ 是对称矩阵. 设

$$\alpha = (\alpha_1,\cdots,\alpha_n)X, \quad \beta = (\alpha_1,\cdots,\alpha_n)Y$$

则

$$(\alpha,\beta) = X'AY \tag{21}$$

从内积(　,　)可得到二次函数 $q$，其中 $q(\alpha) = (\alpha,\alpha) = X'AX$. 进而得到二次型 $X'AX$. 因为 $(\alpha,\alpha) \geqslant 0$，所以 $X'AX \geqslant 0$. 当 $\alpha \neq 0$ 时有 $(\alpha,\alpha) > 0$，于是当 $X \neq 0$ 时，有 $X'AX > 0$. 这表明二次型 $X'AX$ 是正定的，从而它的矩阵 $A$ 是正定矩阵. 这证明了：内积在 $V$ 的任意一个基下的度量矩阵是正定矩阵.

在欧氏空间 $V$ 中指定的内积在基 $\alpha_1,\cdots,\alpha_n$ 下的度量矩阵 $A$ 也称为基 $\alpha_1,\cdots,\alpha_n$ 的**度量矩阵**.

## 习　题　12.1

1. 在线性空间 $M_n(R)$ 中，规定

$$f(A,B) := T_r(AB), \quad \forall A,B \in M_n(R)$$

试问：$f$ 是不是 $M_n(R)$ 上的一个内积？

2. 在 $R^2$ 中，对于任意 $\alpha=(x_1,x_2),\beta=(y_1,y_2)$，规定

$$(\alpha,\beta)=x_1y_1-x_1y_2-x_2y_1+4x_2y_2$$

证明：( , ) 是 $R^2$ 上的一个内积.

3. 设 $C$ 是一个 $n$ 级实可逆矩阵. 在 $R^n$ 中，对于任意两个列向量 $X,Y$，规定

$$(X,Y)=X'C'CY$$

证明：( , ) 是 $R^n$ 上的一个内积.

4. 设 $V$ 和 $U$ 都是实数域上的线性空间，设 $(\ ,\ )_1$ 是 $U$ 上的一个内积. 设 $\sigma$ 是 $V$ 到 $U$ 的一个线性映射，并且 $\sigma$ 是单射. 对于 $V$ 中任意两个向量 $\alpha,\beta$，规定

$$(\alpha,\beta)=(\sigma\alpha,\sigma\beta)_1$$

证明：( , ) 是 $V$ 上的一个内积.

*5. 设 $V$ 是 $n$ 维实线性空间，$V$ 中取一个基 $\alpha_1,\alpha_2,\cdots,\alpha_n$. 对于 $V$ 中任意两个向量 $\alpha=\sum_{i=1}^{n}x_i\alpha_i,\beta=\sum_{i=1}^{n}y_i\alpha_i$，规定

$$(\alpha,\beta)=x_1y_1+x_2y_2+\cdots+x_ny_n$$

证明：(1) ( , ) 是 $V$ 上的一个内积；

(2) 对于 $V$ 中任意一个基 $\alpha_1,\cdots,\alpha_n$，存在唯一的一个 $V$ 上的内积 ( , )，使得

$$(\alpha_i,\alpha_j)=\delta_{ij},\quad i,j=1,2,\cdots,n$$

（提示：(1)方法一：直接按内积的定义验证，方法二：利用第 4 题的结论.）

6. 设 $V=C[0,1]$，考虑 $V$ 到自身的一个映射 $\sigma: f\longmapsto\sigma f$，其中 $\sigma f$ 的定义为

$$(\sigma f)(t):=tf(t),\quad \forall\, t\in[0,1]$$

(1) 证明 $\sigma$ 是 $V$ 上的线性变换，并且 $\sigma$ 是单射；

(2) 对于任意 $f,g\in V$，规定

$$(f,g)=\int_0^1 f(t)g(t)t^2\mathrm{d}t$$

证明：$(f,g)$ 是 $V$ 上的一个内积.

*7. 设 $V$ 是实线性空间，证明：$V$ 上的两个内积的和仍是 $V$ 上的一个内积；$V$ 上一个内积的正实数倍仍是 $V$ 上的一个内积. 试问：$V$ 上的两个内积的差

是 $V$ 上的内积吗?$V$ 上一个内积的负实数倍是 $V$ 上的内积吗?

*8. 求出 $R^1$ 上的所有内积.

9. 设( , )是 $R^2$ 上的标准内积.

(1) 设 $\alpha=(1,2),\beta=(-1,1)$. 如果 $\gamma\in R^2$ 使得 $(\alpha,\gamma)=-1$,并且 $(\beta,\gamma)=3$,求 $\gamma$;

(2) 证明:对于任意 $\alpha\in R^2$,有

$$\alpha=(\alpha,\varepsilon_1)\varepsilon_1+(\alpha,\varepsilon_2)\varepsilon_2$$

其中 $\varepsilon_1=(1,0),\varepsilon_2=(0,1)$.

10. 设 $A=(a_{ij})$ 是一个 $n$ 级正定矩阵. 在 $R^n$ 中,对于任意两个列向量 $X$, $Y$,规定

$$(X,Y)=X'AY$$

(1) 证明( , )是 $R^n$ 上的一个内积;

(2) 求内积( , )在 $R^n$ 的基 $\varepsilon_1=(1,0,\cdots,0),\cdots,\varepsilon_n=(0,\cdots,0,1)$ 下的度量矩阵.

(3) 具体写出这个欧氏空间中的 Cauchy-Buniakowski 不等式.

11. 在 $R^4$ 中,给定标准内积( , ). 求 $\alpha,\beta$ 之间的夹角 $\langle\alpha,\beta\rangle$,设

$$\alpha=(1,-1,4,0),\quad \beta=(3,1,-2,2)$$

12. 设 $\alpha_1,\alpha_2,\cdots,\alpha_n$ 是欧氏空间 $V$ 的一个基,证明:

(1) 如果 $\eta\in V$ 使 $(\eta,\alpha_i)=0,i=1,2,\cdots,n$,则 $\eta=0$;

(2) 如果 $\eta_1,\eta_2\in V$ 使得对任一 $\alpha\in V$ 有 $(\eta_1,\alpha)=(\eta_2,\alpha)$,那么 $\eta_1=\eta_2$.

*13. 设 $\underset{\sim}{A}$ 是 $R^2$ 上的一个线性变换,使得

$$\underset{\sim}{A}\begin{pmatrix}x_1\\x_2\end{pmatrix}=\begin{pmatrix}x_2\\-x_1\end{pmatrix},\quad \forall\begin{pmatrix}x_1\\x_2\end{pmatrix}\in R^2$$

设( , )是 $R^2$ 上的标准内积,证明:

$$(x,\underset{\sim}{A}x)=0,\quad \forall\, x\in R^2$$

说明线性变换 $\underset{\sim}{A}$ 的几何意义.

14. 设 $V$ 是任一实内积空间,证明:

$$(\alpha,\beta)=\frac{1}{4}|\alpha+\beta|^2-\frac{1}{4}|\alpha-\beta|^2,\quad \forall\,\alpha,\beta\in V$$

这个恒等式称为**极化恒等式**.

15. 设 $V$ 是任一实内积空间,证明:

$$|\alpha+\beta|^2+|\alpha-\beta|^2=2|\alpha|^2+2|\beta|^2,\quad \forall\,\alpha,\beta\in V$$

说明这个恒等式的几何意义.

16. 求出第 2 题中 $R^2$ 的内积在基 $\varepsilon_1=(1,0),\varepsilon_2=(0,1)$ 下的度量矩阵.

17. 设 $V=R[x]$,对于 $f(x)=\sum_{i=0}^{n}a_ix^i,g(x)=\sum_{j=0}^{m}b_jx^j\in R[x]$,规定

$$(f,g)=\sum_{i=0}^{n}\sum_{j=0}^{m}\frac{a_ib_j}{i+j+1}$$

(1) 证明:( , )是 $R[x]$ 上的一个内积;

(2) 设 $W$ 是由 $R[x]$ 中次数不超过 $n$ 的多项式组成的子空间,把 $R[x]$ 中的上述内积( , )限制到 $W$ 上. 求出 $W$ 上的这个内积在 $W$ 的一个基 $1,x,x^2,\cdots,x^n$ 下的度量矩阵.

(提示:(1) 把 $C[0,1]$ 上的内积 $(f,g)=\int_0^1 f(x)g(x)\mathrm{d}x$ 限制到 $R[x]$ 上.)

*18. 证明下述 $n+1$ 级实矩阵 $A$ 是正定矩阵:

$$A=\begin{pmatrix}1 & \frac{1}{2} & \frac{1}{3} & \cdots & \frac{1}{n+1}\\ \frac{1}{2} & \frac{1}{3} & \frac{1}{4} & \cdots & \frac{1}{n+2}\\ \cdots & \cdots & \cdots & \cdots & \cdots\\ \frac{1}{n+1} & \frac{1}{n+2} & \frac{1}{n+3} & \cdots & \frac{1}{2n+1}\end{pmatrix}$$

*19. 设 $V$ 是一个 $n$ 维欧氏空间,$V$ 中取一个基 $\alpha_1,\cdots,\alpha_n$. 设 $c_1,\cdots,c_n$ 是任意给定的一组实数. 证明:$V$ 中存在唯一的一个向量 $\alpha$,使得

$$(\alpha,\alpha_j)=c_j,\quad j=1,2,\cdots,n$$

*20. 设 $\alpha=(x,y),\beta=(-y,x)\in R^2$. 试问:

(1) 若欧氏空间 $R^2$ 中指定的内积是标准内积,$\alpha$ 与 $\beta$ 是否一定正交?

(2) 若欧氏空间 $R^2$ 中指定的内积是第 2 题中的内积,$\alpha$ 与 $\beta$ 是否一定正交?写出 $\alpha$ 与 $\beta$ 正交的充分必要条件.

## §2 正交集·标准正交基

我们已经知道,一个实内积空间 $V$ 就是一个实线性空间 $V$ 连

同$V$上的一个指定的内积. 有了这个内积,$V$中就有了长度、角度、正交、距离等度量概念. 实内积空间$V$中的向量之间的关系,从线性运算的角度看,有线性相关与线性无关之区分;从度量的角度看,有正交与不正交之区分. 本节来讨论这两者之间的联系. 进而给出欧氏空间中标准正交基的定义、存在性、特征、求法.

**定义** 1 实内积空间$V$的一个子集$S$称为**正交集**,如果$S$中任意两个不同的向量都正交. 假如还有$S$中每一个向量都是单位向量,则称$S$是一个**正交规范集**.

正交集中允许有零向量. 当然我们感兴趣的是由非零向量组成的正交集. 第四章 § 10 讲的正交向量组就是由非零向量组成的有限正交集.

**定理** 12.2.1 在实内积空间$V$中,由非零向量组成的正交集是线性无关的.

**证明** 设$S$是由非零向量组成的正交集. 为了证明$S$是线性无关的,只要证$S$的任一有限子集(由不同元素组成)是线性无关的. 设$\alpha_1,\alpha_2,\cdots,\alpha_m$是$S$中的不同的向量. 假设

$$k_1\alpha_1 + k_2\alpha_2 + \cdots + k_m\alpha_m = 0 \tag{1}$$

则$(\sum_{i=1}^{m} k_i\alpha_i,\alpha_j) = 0$. 由于$S$是正交集,因此得出,$k_j(\alpha_j,\alpha_j) = 0$. 由于$(\alpha_j,\alpha_j) \neq 0$,因此$k_j = 0, 1 \leqslant j \leqslant m$. 这证明$S$是线性无关的. ▌

从定理 12.2.1 立即得出,在$n$维欧几里得空间$V$中,彼此正交的非零向量的数目不会超过空间的维数$n$. 进一步我们可以证明:

**定理** 12.2.2 在$n$维欧几里得空间$V$中,存在$n$个非零向量,它们彼此正交.

**证明** 因为内积是对称双线性函数,所以$V$中存在一个基$\alpha_1,\alpha_2,\cdots,\alpha_n$,使得$V$中指定的内积在这个基下的度量矩阵是对角矩阵:$\mathrm{diag}\{d_1,d_2,\cdots,d_n\}$. 于是

$$(\alpha_i,\alpha_j) = 0,\quad 当\ i \neq j,\quad i,j = 1,\cdots,n$$

因此 $\alpha_1,\alpha_2,\cdots,\alpha_n$ 是两两正交的 $n$ 个非零向量. ▎

**定义** 2 在 $n$ 维欧氏空间中，设 $\alpha_1,\alpha_2,\cdots,\alpha_n$ 是两两正交的 $n$ 个非零向量，则 $\alpha_1,\alpha_2,\cdots,\alpha_n$ 称为 $V$ 的一个**正交基**. 如果 $V$ 的一个正交基中每个向量都是单位向量，则称它为 $V$ 的一个**标准正交基**(或**正交规范基**).

由定理 12.2.1 知，$V$ 的每个正交基都是 $V$ 的一个基. 从定理 12.2.2 知道，$n$ 维欧氏空间一定有正交基. 进一步我们可以证明：

**定理** 12.2.3 $n$ 维欧氏空间 $V$ 一定有标准正交基.

**证明** 在定理 12.2.2 的证明中找到了 $V$ 的一个正交基 $\alpha_1,\alpha_2,\cdots,\alpha_n$. 令

$$\eta_i = \frac{1}{|\alpha_i|}\alpha_i, \quad i = 1,\cdots,n \tag{2}$$

显然，$\eta_1,\eta_2,\cdots,\eta_n$ 就是 $V$ 的一个标准正交基. ▎

从定义 2 知道，在 $n$ 维欧氏空间 $V$ 中，向量组 $\eta_1,\eta_2,\cdots,\eta_n$ 是 $V$ 的一个标准正交基当且仅当下列式子成立：

$$(\eta_i,\eta_j) = \delta_{ij}, \quad i,j = 1,\cdots,n \tag{3}$$

换句话说，$V$ 的一个基是标准正交基当且仅当内积在这个基下的度量矩阵是单位矩阵.

从上述结论看出，在 $n$ 维欧氏空间 $V$ 中，标准正交基比一般的基优越的地方在于：内积的表达式比较简单，从而易于计算. 设 $\eta_1,\cdots,\eta_n$ 是欧氏空间 $V$ 的一个标准正交基，由于内积在这个基下的度量矩阵是 $I$，因此对于任意 $\alpha = \sum_{i=1}^{n} x_i\eta_i, \beta = \sum_{i=1}^{n} y_i\eta_i$，有

$$(\alpha,\beta) = X'IY = X'Y = x_1y_1 + x_2y_2 + \cdots + x_ny_n \tag{4}$$

其中 $X' = (x_1,\cdots,x_n), Y' = (y_1,\cdots,y_n)$. 当 $n = 3$ 时，(4) 式正是几何空间中向量的内积在直角坐标系中的表达式.

标准正交基的另一个优越性在于，$V$ 中任一向量在标准正交基下的坐标可以通过内积简单地表示出来. 设 $\eta_1,\cdots,\eta_n$ 是 $V$ 的一个标准正交基，则对于任意 $\alpha \in V$，有

$$\alpha = \sum_{i=1}^{n} (\alpha, \eta_i)\eta_i \tag{5}$$

即 $\alpha$ 在标准正交基 $\eta_1, \cdots, \eta_n$ 下的坐标的第 $i$ 个分量正好是 $\alpha$ 与 $\eta_i$ 的内积，$i = 1, \cdots, n$.

**证明** 设 $\alpha = \sum_{i=1}^{n} x_i \eta_i$，则由(4)式得

$$(\alpha, \eta_j) = x_j, \quad j = 1, 2, \cdots, n$$

于是得到(5)式. ▌

(5)式中的 $\{(\alpha, \eta_i) \mid i = 1, \cdots, n\}$ 也称为 $\alpha$ 关于标准正交基 $\eta_1, \cdots, \eta_n$ 的 Fourier **系数**.

在欧氏空间中，由于标准正交基有这些优越性，所以在讨论欧氏空间中的问题时，总是取它的一个标准正交基，而不取一般的基.

在 $n$ 维欧氏空间 $V$ 中，标准正交基不是唯一的. 例如几何空间中，任意三个两两垂直的单位向量都组成一个标准正交基. 现在来讨论 $V$ 中标准正交基之间的关系.

设 $\varepsilon_1, \cdots, \varepsilon_n$ 与 $\eta_1, \cdots, \eta_n$ 都是 $V$ 的标准正交基，我们来看 $\varepsilon_1, \cdots, \varepsilon_n$ 到 $\eta_1, \cdots, \eta_n$ 的过渡矩阵 $P$ 有什么特点？由于 $V$ 中指定的内积在 $V$ 的任意一个标准正交基下的度量矩阵都是单位矩阵 $I$，并且同一个双线性函数在不同基下的度量矩阵是合同的，因此得到

$$I = P'IP = P'P \tag{6}$$

这表明 $P$ 是正交矩阵. 于是我们证明了

**命题** 12.2.1 欧氏空间 $V$ 中，由标准正交基到标准正交基的过渡矩阵是正交矩阵. ▌

反之，我们有

**命题** 12.2.2 $n$ 维欧氏空间 $V$ 中，设 $\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n$ 是 $V$ 的一个标准正交基，设向量组 $\eta_1, \cdots, \eta_n$ 满足

$$(\eta_1, \cdots, \eta_n) = (\varepsilon_1, \cdots, \varepsilon_n)P$$

其中 $P$ 是正交矩阵，则 $\eta_1, \cdots, \eta_n$ 是 $V$ 的一个标准正交基.

**证明** 因为 $P$ 可逆，所以 $\eta_1,\cdots,\eta_n$ 是 $V$ 的一个基. 因为内积在基 $\eta_1,\cdots,\eta_n$ 下的度量矩阵为 $P'IP=P'P=I$，所以 $\eta_1,\cdots,\eta_n$ 是标准正交基. ▌

在 $n$ 维欧氏空间 $V$ 中，如何找出一个标准正交基?注意第四章§10 的定理 4.10.1 对于任意 $n$ 维欧氏空间也成立(证明完全一样). 因此在 $V$ 中找标准正交基的方法是:

首先，给出 $V$ 的一个基 $\alpha_1,\alpha_2,\cdots,\alpha_n$.

其次，用 Gram-Schmidt **正交化过程**把 $\alpha_1,\alpha_2,\cdots,\alpha_n$ 变成与之等价的正交向量组 $\beta_1,\beta_2,\cdots,\beta_n$. 具体做法是:令

$$\beta_1=\alpha_1,$$

$$\beta_2=\alpha_2-\frac{(\alpha_2,\beta_1)}{(\beta_1,\beta_1)}\beta_1,$$

$$\beta_3=\alpha_3-\frac{(\alpha_3,\beta_1)}{(\beta_1,\beta_1)}\beta_1-\frac{(\alpha_3,\beta_2)}{(\beta_2,\beta_2)}\beta_2,$$

$$\cdots\quad\cdots\quad\cdots\quad\cdots\quad\cdots\quad\cdots\quad\cdots$$

$$\beta_n=\alpha_n-\sum_{i=1}^{n-1}\frac{(\alpha_n,\beta_i)}{(\beta_i,\beta_i)}\beta_i$$

第三步，把 $V$ 的正交基 $\beta_1,\cdots,\beta_n$ 中每个向量**单位化**，即令

$$\eta_i=\frac{1}{|\beta_i|}\beta_i,\quad i=1,\cdots,n$$

则 $\eta_1,\eta_2,\cdots,\eta_n$ 是 $V$ 的一个标准正交基.

## 习　题　12.2

1. 设 $\varepsilon_1,\varepsilon_2,\varepsilon_3$ 是 3 维欧氏空间 $V$ 的一个标准正交基，证明:

$$\alpha_1=\frac{1}{3}(2\varepsilon_1-\varepsilon_2+2\varepsilon_3),\quad \alpha_2=\frac{1}{3}(2\varepsilon_1+2\varepsilon_2-\varepsilon_3)$$

$$\alpha_3=\frac{1}{3}(\varepsilon_1-2\varepsilon_2-2\varepsilon_3)$$

也是 $V$ 的一个标准正交基.

2. 设 $\varepsilon_1,\varepsilon_2,\varepsilon_3,\varepsilon_4,\varepsilon_5$ 是 5 维欧氏空间 $V$ 的一个标准正交基，$V_1=\langle\alpha_1,\alpha_2,\alpha_3\rangle$，其中

$$\alpha_1=\varepsilon_1+2\varepsilon_3-\varepsilon_5,\quad \alpha_2=\varepsilon_2-\varepsilon_3+\varepsilon_4$$

$$\alpha_3=-\varepsilon_2+\varepsilon_3+\varepsilon_5$$

求 $V_1$ 的一个标准正交基.

3. 已知一个 $3\times5$ 实矩阵

$$A=\begin{pmatrix}1&-1&2&0&3\\2&0&-1&1&4\\-1&1&1&0&-2\end{pmatrix}$$

求一个 $5\times2$ 实矩阵 $B$，使 $AB=0$，且 $B$ 的列向量组是正交规范集($R^5$ 中定义标准内积).

4. 在 $R[x]_4$ 中给定一个内积为

$$(f,g)=\int_{-1}^{1}f(x)g(x)\mathrm{d}x$$

求 $R[x]_4$ 的一个标准正交基.

(提示：由基 $1,x,x^2,x^3$ 出发作正交化和单位化.)

5. 设 $V$ 是 3 维欧氏空间，$\alpha_1,\alpha_2,\alpha_3$ 是 $V$ 的一个基. $V$ 中指定的内积在基 $\alpha_1,\alpha_2,\alpha_3$ 下的度量矩阵为

$$A=\begin{pmatrix}1&0&1\\0&10&-2\\1&-4&2\end{pmatrix}$$

求 $V$ 的一个标准正交基.

6. 设 $\alpha_1,\alpha_2,\cdots,\alpha_m$ 是 $n$ 维欧氏空间 $V$ 的一组向量，令

$$A=\begin{pmatrix}(\alpha_1,\alpha_1)&(\alpha_1,\alpha_2)&\cdots&(\alpha_1,\alpha_m)\\(\alpha_2,\alpha_1)&(\alpha_2,\alpha_2)&\cdots&(\alpha_2,\alpha_m)\\\cdots\cdots&\cdots\cdots&\cdots&\cdots\cdots\\(\alpha_m,\alpha_1)&(\alpha_m,\alpha_2)&\cdots&(\alpha_m,\alpha_m)\end{pmatrix}$$

把 $|A|$ 称为向量组 $\alpha_1,\cdots,\alpha_m$ 的 Gram **行列式**，记作 $G(\alpha_1,\cdots,\alpha_m)$. 证明：$G(\alpha_1,\cdots,\alpha_m)\geqslant0$，等号成立当且仅当 $\alpha_1,\cdots,\alpha_m$ 线性相关.

(提示：先证若 $\alpha_1,\cdots,\alpha_m$ 线性相关，则 $G(\alpha_1,\cdots,\alpha_m)=0$. 方法是：用线性相关的定义，然后转化为齐次线性方程组的非零解的判定问题. 再证若 $\alpha_1,\cdots,\alpha_m$ 线性无关，则 $G(\alpha_1,\cdots,\alpha_m)>0$. 方法是令 $V_1=\langle\alpha_1,\cdots,\alpha_m\rangle$，则

$\alpha_1,\cdots,\alpha_m$ 是 $V_1$ 的一个基. 考虑内积在 $V_1$ 上的限制在基 $\alpha_1,\cdots,\alpha_m$ 下的度量矩阵.）

7. 在 $n$ 维欧氏空间 $V$ 中，向量组 $\alpha_1,\alpha_2,\cdots,\alpha_m$ 张成的"$m$ 维平行 $2m$ 面体"的体积 $V(\alpha_1,\cdots,\alpha_m)$ 规定为

$$[V(\alpha_1,\cdots,\alpha_m)]^2 = G(\alpha_1,\cdots,\alpha_m)$$

当 $m=2,3$ 时，分别计算 $V^2$ 的表达式，说明其几何意义.

*8. 在实内积空间 $C[0,2\pi]$ 中，指定的内积为

$$(f,g) = \int_0^{2\pi} f(x)g(x)\mathrm{d}x$$

证明：$C[0,2\pi]$ 的一个子集

$$S = \left\{ \frac{1}{\sqrt{2\pi}}, \frac{1}{\sqrt{\pi}}\cos nx, \frac{1}{\sqrt{\pi}}\sin nx \,\middle|\, n \geqslant 1 \right\}$$

是一个正交规范集.

## §3 正交补·正交投影·最小二乘法

这一节我们要通过实内积空间的子空间来研究整个空间的结构及其应用.

设 $V$ 是一个实内积空间，$V_1$ 是 $V$ 的任一线性子空间. 显然 $V$ 上指定的内积（ , ）可以限制到 $V_1$ 上，从而 $V_1$ 对于这个内积也成为一个实内积空间. 此时称 $V_1$ 是实内积空间 $V$ 的一个子空间.

设 $S$ 是实内积空间 $V$ 的一个子集，$V$ 中向量 $\alpha$ 如果与 $S$ 中每个向量都正交，则称 $\alpha$ **与** $S$ **正交**，记作 $\alpha \perp S$.

**定义 1** 设 $S$ 是实内积空间 $V$ 的一个非空子集，$V$ 中与 $S$ 正交的所有向量组成的集合称为 $S$ 的**正交补**，记作 $S^\perp$，即

$$S^\perp = \{\alpha \in V \mid (\alpha,\beta) = 0, \forall\, \beta \in S\} \tag{1}$$

**命题** 12.3.1 设 $S$ 是实内积空间 $V$ 的任意一个非空子集，则 $S^\perp$ 是 $V$ 的一个子空间.

**证明** 显然，$0 \in S^\perp$. 任 $\alpha,\beta \in S^\perp$，任取 $\gamma \in S$，有

$$(\alpha+\beta,\gamma)=(\alpha,\gamma)+(\beta,\gamma)=0$$

所以 $\alpha+\beta\in S^{\perp}$. 同理可证 $k\alpha\in S^{\perp},\forall\ k\in R,\alpha\in S^{\perp}$. 因此 $S^{\perp}$ 是 $V$ 的一个子空间. ▌

如果把 $S$ 取成 $V$ 的一个有限维子空间，则有一个重要的结论：

**定理** 12.3.1　设 $U$ 是实内积空间 $V$ 的一个有限维子空间，则

$$V=U\oplus U^{\perp} \tag{2}$$

**证明**　先证 $V=U+U^{\perp}$. 在 $U$ 中取一个标准正交基 $\varepsilon_1,\cdots,\varepsilon_m$. 任取 $\alpha\in V$，令

$$\alpha_1=(\alpha,\varepsilon_1)\varepsilon_1+\cdots+(\alpha,\varepsilon_m)\varepsilon_m\in U \tag{3}$$

$$\alpha_2=\alpha-\alpha_1 \tag{4}$$

因为对于 $j=1,\cdots,m$，有

$$\begin{aligned}(\alpha_2,\varepsilon_j)&=(\alpha-\alpha_1,\varepsilon_j)=(\alpha,\varepsilon_j)-(\sum_{i=1}^{m}(\alpha,\varepsilon_i)\varepsilon_i,\varepsilon_j)\\&=(\alpha,\varepsilon_j)-(\alpha,\varepsilon_j)=0\end{aligned}$$

所以 $\alpha_2\perp U$，即 $\alpha_2\in U^{\perp}$. 从 $\alpha=\alpha_1+\alpha_2$ 知道 $V=U+U^{\perp}$.

再证 $U\cap U^{\perp}=0$. 假设 $\gamma\in U\cap U^{\perp}$，则 $(\gamma,\gamma)=0$，从而得 $\gamma=0$.

综上述得，$V=U\oplus U^{\perp}$. ▌

例如，在几何空间 $V$ 中，设 $U$ 是过原点的一个平面，则 $U^{\perp}$ 是过原点并且与平面 $U$ 垂直的直线. 显然，$V=U\oplus U^{\perp}$.

从定理 12.3.1 知道，如果 $U$ 是实内积空间 $V$ 的一个有限维子空间，则 $V=U\oplus U^{\perp}$. 从而 $V$ 中每个向量 $\alpha$ 能唯一地表示成

$$\alpha=\alpha_1+\alpha_2,\quad \alpha_1\in U,\quad \alpha_2\in U^{\perp} \tag{5}$$

令 $\underline{P}_U(\alpha)=\alpha_1$，则 $\underline{P}_U$ 就是平行于 $U^{\perp}$ 在 $U$ 上的投影（见第九章§1），称 $\underline{P}_U$ 是 $V$ **在** $U$ **上的正交投影**，它是 $V$ 上的一个线性变换. 我们把(5)式中的 $\alpha_1$ 称为 $\alpha$ **在** $U$ **上的正交投影**. 换句话说，$\alpha$ 在 $U$ 上的正交投影就是 $\alpha$ 在 $\underline{P}_U$ 下的象. 从定理 12.3.1 的证明中看到，若在 $U$ 中取一个标准正交基 $\varepsilon_1,\cdots,\varepsilon_m$，则 $\alpha$ 在 $U$ 上的正交投影为

$$\alpha_1 = \sum_{i=1}^{m}(\alpha,\varepsilon_i)\varepsilon_i \tag{6}$$

从(5)式看出，$\alpha_1 \in U$是$\alpha$在$U$上的正交投影的充分必要条件是$\alpha - \alpha_1 \in U^{\perp}$.

以几何空间为例，设$U$是过原点$O$的一个平面，$\alpha$在$U$上的正交投影$\alpha_1$具有这样的性质：$\alpha_1$与$\alpha$的距离比$U$上任一其他向量$\gamma$与$\alpha$的距离都短. 这一性质对于$U$是实内积空间$V$的有限维子空间的情形也正确.

**定理** 12.3.2 设$U$是实内积空间$V$的一个有限维子空间，对于$\alpha \in V$，$\alpha_1 \in U$是$\alpha$在$U$上的正交投影的充分必要条件为

$$d(\alpha,\alpha_1) \leqslant d(\alpha,\gamma), \quad \forall\, \gamma \in U \tag{7}$$

**证明** 必要性. 设$\alpha_1$是$\alpha$在$U$上的正交投影，则

$$(\alpha - \alpha_1) \perp (\alpha_1 - \gamma), \quad \forall\, \gamma \in U$$

于是由勾股定理得

$$\begin{aligned}&|\alpha - \alpha_1|^2 + |\alpha_1 - \gamma|^2 \\ = &|(\alpha - \alpha_1) + (\alpha_1 - \gamma)|^2 = |\alpha - \gamma|^2, \quad \forall\, \gamma \in U\end{aligned}$$

由此得出，$|\alpha - \alpha_1| \leqslant |\alpha - \gamma|$，$\forall\, \gamma \in U$.

充分性. 设(7)式成立. 假设$\delta$是$\alpha$在$U$上的正交投影. 据刚才证的必要性，得出

$$d(\alpha,\delta) \leqslant d(\alpha,\alpha_1) \tag{8}$$

由(8)式和(7)式得出，$d(\alpha,\delta) = d(\alpha,\alpha_1)$. 由于$(\alpha - \delta) \in U^{\perp}$，$(\delta - \alpha_1) \in U$，因此

$$\begin{aligned}|\alpha - \alpha_1|^2 &= |(\alpha - \delta) + (\delta - \alpha_1)|^2 \\ &= |\alpha - \delta|^2 + |\delta - \alpha_1|^2\end{aligned}$$

由此得出，$|\delta - \alpha_1|^2 = 0$，所以$\delta = \alpha_1$. ▌

**定义** 2 设$U$是实内积空间$V$的一个子空间，$\alpha$是$V$中向量. 如果$U$中存在一个向量$\delta$，使得$d(\alpha,\delta) \leqslant d(\alpha,\gamma)$，$\forall\, \gamma \in U$，那么$\delta$称为$\alpha$在$U$上的**最佳逼近元**.

定理 12.3.2 表明，如果$U$是$V$的有限维子空间，则$V$中任一

向量$\alpha$在$U$上的最佳逼近元存在并且唯一，它就是$\alpha$在$U$上的正交投影$\alpha_1$.

下面介绍正交投影和最佳逼近元的应用之一：最小二乘法.

在许多实际问题中需要研究一个变量$y$与其他一些变量$x_1$, $x_2,\cdots,x_n$之间的依赖关系. 经过实际观测和分析，假定$y$与$x_1,\cdots$, $x_n$之间呈线性关系：

$$y=k_1x_1+k_2x_2+\cdots+k_nx_n$$

但是系数$k_1,\cdots,k_n$是未知的. 为了定出它们，需要观测数据$m$次，即测得$m$组数：

| $y$ | $x_1$ | $x_2$ | $\cdots$ | $x_n$ |
|---|---|---|---|---|
| $b_1$ | $a_{11}$ | $a_{12}$ | $\cdots$ | $a_{1n}$ |
| $\cdots$ | $\cdots$ | $\cdots$ | $\cdots$ | $\cdots$ |
| $b_m$ | $a_{m1}$ | $a_{m2}$ | $\cdots$ | $a_{mn}$ |

如果观测是绝对精确的话，原则上只要测量$m=n$次，通过线性方程组就可以解出$k_1,\cdots,k_n$. 但是，任何观测都会有误差，这样就需要多观测些次数，即$m>n$. 于是得到的线性方程组

$$\begin{cases} a_{11}k_1+a_{12}k_2+\cdots+a_{1n}k_n=b_1 \\ \cdots\quad\cdots\quad\cdots\quad\cdots\quad\cdots\quad\cdots\quad\cdots \\ a_{m1}k_1+a_{m2}k_2+\cdots+a_{mn}k_n=b_m \end{cases} \tag{8}$$

中，方程个数$m$大于未知量个数$n$. 这时方程组(8)可能无解. 即任何一组实数$k_1,\cdots,k_n$，都可能使

$$\sum_{i=1}^{m}[b_i-(a_{i1}k_1+\cdots+a_{in}k_n)]^2. \tag{9}$$

不等于零. 我们要设法找一组数$k_1^0,\cdots,k_n^0$，使得(9)式最小. 这样的$k_1^0,\cdots,k_n^0$称为方程组(8)的**最小二乘解**. 这种问题叫做**最小二乘法问题**.

现在我们来讨论如何求出方程组(8)的最小二乘解. (9)式是平方和的形式，这使人联想到欧氏空间$R^m$(指定的内积是标准内

积）中向量的长度的表达式. 令

$$A=\begin{pmatrix}a_{11}&a_{12}&\cdots&a_{1n}\\ \cdots&\cdots&\cdots&\cdots\\ a_{m1}&a_{m2}&\cdots&a_{mn}\end{pmatrix},\quad X=\begin{pmatrix}k_1\\ \vdots\\ k_n\end{pmatrix},\quad \beta=\begin{pmatrix}b_1\\ \vdots\\ b_m\end{pmatrix}$$

则(9)式是欧氏空间 $R^m$(指定的内积为标准内积）中向量 $\beta-AX$ 的长度的平方，也就是 $\beta$ 与 $AX$ 的距离的平方. 设 $A$ 的列向量组为 $\alpha_1,\cdots,\alpha_n$，则

$$AX=k_1\alpha_1+\cdots+k_n\alpha_n\in\langle\alpha_1,\cdots,\alpha_n\rangle=:U$$

找 $X^0=\begin{pmatrix}k_1^0\\ \vdots\\ k_n^0\end{pmatrix}$ 使得(9)式最小的问题就是在 $U$ 中找一个向量 $AX^0$，使得它与 $\beta$ 的距离比 $U$ 中任一其他向量与 $\beta$ 的距离都短. 也就是找 $\beta$ 在 $U$ 上的最佳逼近元. 由于 $R^m$ 和 $U$ 都是有限维的，因此 $\beta$ 在 $U$ 上的最佳逼近元就是 $\beta$ 在 $U$ 上的正交投影. 因此 $AX^0$ 应为 $\beta$ 在 $U$ 上的正交投影，换句话说，$\beta-AX^0\in U^\perp$. 而

$$\begin{aligned}&\beta-AX^0\in U^\perp\Leftrightarrow\beta-AX^0\perp\alpha_j,\quad j=1,\cdots,n\\ &\Leftrightarrow(\beta-AX^0,\alpha_j)=0,\quad j=1,\cdots,n\\ &\Leftrightarrow\alpha_j'(\beta-AX^0)=0,\quad j=1,\cdots,n\\ &\Leftrightarrow A'(\beta-AX^0)=0\Leftrightarrow A'AX^0=A'\beta\\ &\Leftrightarrow X^0\text{ 是线性方程组 }A'AX=A'\beta\text{ 的解}\end{aligned}$$

因为

$$\begin{aligned}&\operatorname{rank}((A'A,A'\beta))\\ =&\operatorname{rank}(A'(A,\beta))\leqslant\operatorname{rank}A'=\operatorname{rank}(A'A)\\ &\operatorname{rank}((A'A,A'\beta))\geqslant\operatorname{rank}(A'A)\end{aligned}$$

所以 $$\operatorname{rank}(A'A,A'\beta)=\operatorname{rank}(AA')$$

从而线性方程组 $A'AX=A'\beta$ 有解.

这样把求线性方程组 $AX=\beta$ 的最小二乘解的问题归结为解线性方程组 $A'AX=A'\beta$. 而这个方程组是一定有解的.

# 习 题 12.3

1. 设 $U$ 是欧氏空间 $R^4$(指定标准内积)的一个子空间,$U=\langle\alpha_1,\alpha_2\rangle$,其中

$$\alpha_1=(1,1,2,1,),\quad \alpha_2=(1,0,0,-2)$$

(1) 求 $U^{\perp}$ 的一个标准正交基;

(2) 求向量 $\alpha=(1,-3,2,2)$ 在 $U$ 上的正交投影.

2. 设 $V$ 是一个 $n$ 维欧氏空间,$\alpha\neq 0$ 是 $V$ 中一个固定向量. 令 $V_1=\{\alpha\}^{\perp}$. 求 $\dim V_1$.

3. 设 $U$ 是 $n$ 维欧氏空间 $V$ 的一个子空间,证明:$(U^{\perp})^{\perp}=U$.

4. 设 $V_1,V_2$ 是 $n$ 维欧氏空间 $V$ 的两个子空间,证明:

$$(V_1+V_2)^{\perp}=V_1^{\perp}\cap V_2^{\perp};\qquad (V_1\cap V_2)^{\perp}=V_1^{\perp}+V_2^{\perp}$$

5. 证明:欧氏空间 $R^n$(指定标准内积)的任一子空间 $U$ 是某个实系数齐次线性方程组的解空间.

*6. 设 $V$ 是一个实内积空间,$W$ 是 $V$ 的一个子空间(可能无限维). 设 $\alpha\in V$. 证明

(1) $\beta\in W$ 是 $\alpha$ 在 $W$ 上的最佳逼近元当且仅当 $(\alpha-\beta)\in W^{\perp}$;

(2) 如果 $\alpha$ 在 $W$ 上的最佳逼近元存在,那么它是唯一的.

(提示:(1) 必要性. 设 $\beta$ 是 $\alpha$ 在 $W$ 上的最佳逼近元. 易推出

$$2(\alpha-\beta,\beta-\gamma)+|\beta-\gamma|^2\geqslant 0,\quad \forall\,\gamma\in W$$

从而也有

$$2(\alpha-\beta,k(\beta-\gamma))+|k(\beta-\gamma)|^2\geqslant 0,\quad \forall\,\gamma\in W,\forall\,k\in R\quad (*)$$

任意给定 $\gamma\in W$,且 $\gamma\neq\beta$,选取适当的 $k_0$,使得(*)式左端的两项能合并成一项,从而推出 $(\alpha-\beta,\beta-\gamma)=0$. 由此得出,$\alpha-\beta\in W^{\perp}$.)

*7. 设 $W$ 是实内积空间 $V$ 的一个子空间. 如果 $\alpha\in V$ 在 $W$ 上的最佳逼近元 $\beta$ 存在(此时必唯一),那么把 $\beta$ 称为 **$\alpha$ 在 $W$ 上的正交投影**. 如果 $V$ 中每一个向量都有在 $W$ 上的正交投影,那么让 $V$ 中每一个向量对应到它在 $W$ 上的正交投影的映射称为 **$V$ 在 $W$ 上的正交投影**.(注:当 $W$ 是有限维时,这里的正交投影的定义与本节正文中的一致.)

证明：$V$ 在 $W$ 上的正交投影 $\underset{\sim}{P}$ 如果存在，则它是 $V$ 上的线性变换，并且是幂等的，还有

$$\operatorname{Ker} \underset{\sim}{P} = W^{\perp}, \qquad \operatorname{Im} \underset{\sim}{P} = W$$

8. 设 $U$ 是实内积空间 $V$ 的一个有限维子空间，用 $\underset{\sim}{P}_U$ 表示 $V$ 在 $U$ 上的正交投影，用 $\underset{\sim}{I}$ 表示 $V$ 上的恒等变换，证明：$V$ 在 $U^{\perp}$ 上的正交投影存在，它等于 $\underset{\sim}{I} - \underset{\sim}{P}_U$.

* 9. 设 $W$ 是实内积空间 $V$ 的一个子空间，证明：如果 $V$ 在 $W$ 上的正交投影 $\underset{\sim}{P}$ 存在，则 $V$ 在 $W^{\perp}$ 上的正交投影也存在，它等于 $\underset{\sim}{I} - \underset{\sim}{P}$.

* 10. 设 $W$ 是实内积空间 $V$ 的一个子空间，证明：$V$ 在 $W$ 上的正交投影存在的充分必要条件是

$$V = W \oplus W^{\perp}$$

11. 设 $U$ 是实内积空间 $V$ 的一个有限维子空间，$\alpha_1, \cdots, \alpha_m$ 是 $U$ 的一个正交基，用 $\underset{\sim}{P}_U$ 表示 $V$ 在 $U$ 上的正交投影. 证明：对于 $\alpha \in V$，有

$$\underset{\sim}{P}_U(\alpha) = \sum_{i=1}^{m} \frac{(\alpha, \alpha_i)}{|\alpha_i|^2} \alpha_i$$

* 12. 可以用正交投影的术语几何地描述实内积空间 $V$ 中对于线性无关的向量组 $\alpha_1, \cdots, \alpha_m$ 施行 Gram-Schmidt 正交化的过程：令

$$W_1 = 0, \quad W_i = \langle \alpha_1, \cdots, \alpha_{i-1} \rangle, \quad i = 2, \cdots, m$$

用 $\underset{\sim}{P}_i$ 表示 $V$ 在 $W_i$ 上的正交投影，用 $\widetilde{\underset{\sim}{P}}_i$ 表示 $V$ 在 $W_i^{\perp}$ 上的正交投影. 令

$$\beta_i = \widetilde{\underset{\sim}{P}}_i(\alpha_i), \quad i = 1, \cdots, m$$

验证 $\beta_1, \cdots, \beta_m$ 就是对 $\alpha_1, \cdots, \alpha_m$ 施行 Gram-Schmidt 正交化过程得到的正交向量组.

13. $\{\varepsilon_1, \cdots, \varepsilon_m\}$ 是实内积空间 $V$ 的一个正交规范集. 证明：对于一切 $\alpha \in V$，有

$$\sum_{i=1}^{m} (\alpha, \varepsilon_i)^2 \leqslant |\alpha|^2$$

等号成立当且仅当 $\alpha = \sum_{i=1}^{m} (\alpha, \varepsilon_i) \varepsilon_i$. 这个不等式称为 Bessel **不等式**.

* 14. 实内积空间 $C[0, 2\pi]$，其指定内积为

$$(f, g) = \int_0^{2\pi} f(x) g(x) \mathrm{d}x$$

证明：对于一切 $f \in C[0, 2\pi]$，有

$$\frac{1}{\pi}\sum_{k=1}^{m}\left[\left(\int_0^{2\pi}f(x)\cos kx\mathrm{d}x\right)^2+\left(\int_0^{2\pi}f(x)\sin kx\mathrm{d}x\right)^2\right]$$
$$\leqslant\int_0^{2\pi}f^2(x)\mathrm{d}x-\frac{1}{2\pi}\left(\int_0^{2\pi}f(x)\mathrm{d}x\right)^2$$

(提示:利用习题 12.2 的第 8 题.)

15. 在欧氏空间 $R[x]_4$ 中,其指定的内积为

$$(f,g)=\int_0^1 f(x)g(x)\mathrm{d}x$$

设 $W$ 是由零次多项式和零多项式组成的子空间,求 $W^\perp$ 以及它的一个基.

16. 实内积空间 $M_n(R)$,指定的内积为 $(A,B)=Tr(AB')$. 设 $W$ 是由所有对角矩阵组成的子空间,求 $W^\perp$ 以及 $W^\perp$ 的一个标准正交基.

*17. 设 $V=C[-1,1]$,指定内积为

$$(f,g)=\int_{-1}^1 f(x)g(x)\mathrm{d}x$$

设 $W$ 是 $V$ 中所有奇函数组成的子空间,求 $W^\perp$;试问:$V$ 在 $W$ 上的正交投影存在吗?

18. 设 $W$ 是实内积空间 $V$ 的一个有限维子空间,用 $\underline{P}$ 表示 $V$ 在 $W$ 上的正交投影,证明

$$(\underline{P}\alpha,\beta)=(\alpha,\underline{P}\beta),\quad \forall\,\alpha,\beta\in V$$

19. 设 $S_1,S_2$ 是实内积空间 $V$ 的两个子集,证明:

$$S_1\subset S_2\Rightarrow S_1^\perp\supset S_2^\perp$$

*20. 设 $S$ 是实内积空间 $V$ 的一个子集,用 $\langle S\rangle$ 表示 $V$ 中所有包含 $S$ 的子空间的交,它仍是 $V$ 的子空间(参看第八章 §6),称它是 $S$ 生成的子空间. 证明:

(1) $\langle S\rangle^\perp\subset S^\perp$;

(2) $\langle S\rangle\subset(S^\perp)^\perp$;

(3) 如果 $V$ 是有限维的,则 $\langle S\rangle=(S^\perp)^\perp$.

# §4 实内积空间的同构

在 §1 中我们曾指出同一个实线性空间 $V$,当指定不同的内

积时，$V$ 便成为不同的实内积空间. 这样，从同一个实线性空间 $V$，可以得到许多实内积空间. 至于从不同的实线性空间，当然可以得到许多不同的实内积空间. 对于众多的实内积空间，如何区分哪些在本质上是一样的？作为线性空间，本质上相同就是它们的元素存在 一一对应，使得对应的元素关于加法和数量乘法这两种运算的性质一样，这称为同构的线性空间. 现在实内积空间在线性空间上装备了一个内积，因此本质上相同的实内积空间 $V$ 与 $V'$ 应当比同构的线性空间增加一条要求：若 $\alpha_i$ 对应到 $\alpha_i'$，$i=1,2$，则 $V$ 上的内积在 $(\alpha_1,\alpha_2)$ 上的函数值应当等于 $V'$ 上的内积在 $(\alpha_1',\alpha_2')$ 上的函数值. 这条要求可以简洁地说成：保持内积. 从以上分析我们给出两个实内积空间在本质上相同的确切含义：

**定义** 1　设 $V$ 和 $V'$ 都是实内积空间，如果存在 $V$ 到 $V'$ 的一个双射 $\sigma$，使得 $\forall\ \alpha,\beta\in V, k\in R$，有

$$\sigma(\alpha+\beta)=\sigma(\alpha)+\sigma(\beta),\quad \sigma(k\alpha)=k\sigma(\alpha)$$

$$(\sigma(\alpha),\sigma(\beta))_2=(\alpha,\beta)_1$$

则称 $V$ 与 $V'$ 是**同构的**，记作 $V\cong V'$；把 $\sigma$ 称为 $V$ 到 $V'$ 的一个**同构映射**.

从定义看出：实内积空间 $V$ 到 $V'$ 的同构映射 $\sigma$ 是线性空间 $V$ 到 $V'$ 的同构映射并且保持内积. 因此，一方面 $\sigma$ 具有线性空间的同构映射的一切性质；另一方面 $\sigma$ 还具有与内积有关的性质. 譬如，若 $V$ 是有限维的，则 $V'$ 也是有限维的，而且 $V$ 与 $V'$ 的维数相同；还有，$\sigma$ 把 $V$ 的标准正交基映成 $V'$ 的标准正交基，等等.

**定理** 12.4.1　两个欧几里得空间同构的充分必要条件是它们的维数相同.

**证明**　必要性已在上面指出. 现在证充分性. 在 $n$ 维欧氏空间 $V$ 和 $V'$ 中分别取一个标准正交基：$\varepsilon_1,\cdots,\varepsilon_n$ 和 $\varepsilon_1',\cdots,\varepsilon_n'$. 令

$$\sigma: V\longrightarrow V'$$

$$\alpha=\sum_{i=1}^{n}x_i\varepsilon_i\longmapsto\sigma(\alpha)=\sum_{i=1}^{n}x_i\varepsilon_i'$$

则 $\sigma$ 是线性空间 $V$ 到 $V'$ 的一个同构映射. 设 $\beta = \sum_{i=1}^{n} y_i \varepsilon_i$，则 $\sigma(\beta) = \sum_{i=1}^{n} y_i \varepsilon_i'$. 于是

$$(\sigma(\alpha), \sigma(\beta)) = \sum_{i=1}^{n} x_i y_i = (\alpha, \beta)$$

因此，$\sigma$ 是实内积空间 $V$ 到 $V'$ 的一个同构映射. 从而 $V \cong V'$. ▌

由定理 12.4.1 得出，同一个 $n$ 维实线性空间 $V$，虽然装备上不同的内积，成为不同的欧氏空间，但是这些欧氏空间是同构的. 而且从不同的 $n$ 维实线性空间得到的欧氏空间也是同构的. 特别地，任一 $n$ 维欧氏空间 $V$ 都与装备标准内积的欧氏空间 $R^n$ 同构，它的一个同构映射是

$$\sigma: V \longrightarrow R^n$$

$$\alpha = \sum_{i=1}^{n} x_i \eta_i \longmapsto (x_1, x_2, \cdots, x_n)$$

其中，$\eta_1, \cdots, \eta_n$ 是 $V$ 的一个标准正交基. 即让 $V$ 中每一个向量对应到它在 $V$ 的一个标准正交基下的坐标，这个映射就是 $n$ 维欧氏空间 $V$ 到 $R^n$ 的同构映射. 由于 $V$ 的标准正交基不唯一，因此 $V$ 到 $R^n$ 的同构映射也不唯一.

第八章 § 7 中曾指出，线性空间的同构具有反身性、对称性和传递性. 可以证明，实内积空间的同构也具有反身性、对称性和传递性. 反身性是显然的. 关于对称性，只要证同构映射的逆映射还是同构映射. 关于传递性，只要证同构映射的乘积还是同构映射. 证明留给读者作为习题.

## 习 题 12.4

1. 在 $R^2$ 中指定内积为

$$(\alpha, \beta)_1 = x_1 y_1 + 2 x_2 y_2$$

其中 $\alpha=(x_1,x_2)$，$\beta=(y_1,y_2)$. 把这个欧氏空间记作 $V$. 找出 $V$ 到欧氏空间 $R^2$（指定的内积为标准内积）的一个同构映射.

2. 证明：同构作为实内积空间之间的关系具有反身性、对称性和传递性.

3. 设 $V$ 和 $V'$ 是两个实内积空间，证明：$\alpha$ 是实内积空间 $V$ 到 $V'$ 的同构映射当且仅当 $\alpha$ 是 $V$ 到 $V'$ 上的线性映射，并且 $\alpha$ 保持内积.

4. 设 $V$ 是 $M_3(R)$ 中所有斜对称矩阵组成的线性空间，装备上一个内积：

$$(A,B)=\frac{1}{2}Tr(AB')$$

设 $\sigma$ 是欧氏空间 $R^3$（指定的是标准内积）到 $V$ 的一个映射：

$$\sigma(x_1,x_2,x_3)=\begin{pmatrix}0 & x_1 & x_2\\ -x_1 & 0 & x_3\\ -x_2 & -x_3 & 0\end{pmatrix}$$

证明：$\sigma$ 是欧氏空间 $R^3$ 到 $V$ 的一个同构映射；求 $V$ 的一个标准正交基.

## §5 正交变换

在研究线性空间时我们讨论了保持线性运算的变换，即线性变换. 现在我们研究实内积空间，则需要讨论与内积有关的线性变换. 这一节来讨论保持内积的线性变换，称之为正交变换.

讨论正交变换的背景来自几何空间. 在解析几何中，我们讲了：空间到自身上的变换如果保持点之间的距离不变，则称为正交（点）变换. 类似地我们有

**定义 1** 实内积空间 $V$ 到自身上的一个变换 $\underset{\sim}{A}$，如果保持向量的内积不变，即对于一切 $\alpha,\beta\in V$ 都有

$$(\underset{\sim}{A}\alpha,\underset{\sim}{A}\beta)=(\alpha,\beta)\tag{1}$$

则称 $\underset{\sim}{A}$ 是 $V$ 上的一个**正交变换**.

从定义 1 容易看出，$V$ 上的正交变换保持向量的长度不变，保持两个非零向量的夹角不变，保持正交性不变.

**命题** 12.5.1 实内积空间 $V$ 上的正交变换 $\underline{A}$ 一定是线性变换.

**证明** 先证：$\forall\ \alpha,\beta\in V$，有 $\underline{A}(\alpha+\beta)=\underline{A}\alpha+\underline{A}\beta$. 为此只要证$(\underline{A}(\alpha+\beta)-(\underline{A}\alpha+\underline{A}\beta),\underline{A}(\alpha+\beta)-(\underline{A}\alpha+\underline{A}\beta))=0$.

$$\begin{aligned}
&(\underline{A}(\alpha+\beta)-(\underline{A}\alpha+\underline{A}\beta),\underline{A}(\alpha+\beta)-(\underline{A}\alpha+\underline{A}\beta))\\
&=|\underline{A}(\alpha+\beta)|^2-2(\underline{A}(\alpha+\beta),\underline{A}\alpha+\underline{A}\beta)+|\underline{A}\alpha+\underline{A}\beta|^2\\
&=|\alpha+\beta|^2-2(\underline{A}(\alpha+\beta),\underline{A}\alpha)-2(\underline{A}(\alpha+\beta),\underline{A}\beta)\\
&\quad+|\underline{A}\alpha|^2+|\underline{A}\beta|^2+2(\underline{A}\alpha,\underline{A}\beta)\\
&=|\alpha+\beta|^2-2(\alpha+\beta,\alpha)-2(\alpha+\beta,\beta)\\
&\quad+|\alpha|^2+|\beta|^2+2(\alpha,\beta)\\
&=|\alpha+\beta|^2-2(\alpha+\beta,\alpha+\beta)+|\alpha+\beta|^2=0
\end{aligned}$$

同理可证

$$(\underline{A}(k\alpha)-k\underline{A}\alpha,\underline{A}(k\alpha)-k\underline{A}\alpha)=0$$

从而得出

$$\underline{A}(k\alpha)=k\underline{A}\alpha,\quad \forall\ \alpha\in V,\quad k\in R$$

所以 $\underline{A}$ 是 $V$ 上的一个线性变换. ▌

从命题 12.5.1 和定义 1 容易得出，正交变换保持两个向量之间的距离不变.

**命题** 12.5.2 实内积空间 $V$ 上的正交变换 $\underline{A}$ 一定是单射，从而 $\underline{A}$ 是可逆的.

**证明** 因为 $(\underline{A}\alpha,\underline{A}\alpha)=(\alpha,\alpha)$，所以

$$\begin{aligned}
\alpha\in \operatorname{Ker}\underline{A}&\Leftrightarrow\underline{A}\alpha=0\Leftrightarrow(\underline{A}\alpha,\underline{A}\alpha)=0\\
&\Leftrightarrow(\alpha,\alpha)=0\Leftrightarrow\alpha=0
\end{aligned}$$

从而 $\operatorname{Ker}\underline{A}=0$. 因此 $\underline{A}$ 是单射. 又由定义知，$\underline{A}$ 是满射，所以 $\underline{A}$ 是双射. 从而 $\underline{A}$ 是可逆的. ▌

我们可从另外一个角度来刻画正交变换：

**定理** 12.5.1 实内积空间 $V$ 到自身上的变换 $\underline{A}$ 是正交变换的充分必要条件为，$\underline{A}$ 是保持向量的长度不变的线性变换.

**证明** 必要性从定义 1 和命题 12.5.1 立即得到.

充分性. 设 $\underset{\sim}{A}$ 是映上的线性变换,且保持向量的长度不变. 对于任意 $\alpha,\beta\in V$,有

$$(\underset{\sim}{A}(\alpha+\beta),\underset{\sim}{A}(\alpha+\beta))=(\alpha+\beta,\alpha+\beta) \tag{2}$$

(2)式的左边为

$$\begin{aligned}&(\underset{\sim}{A}\alpha+\underset{\sim}{A}\beta,\underset{\sim}{A}\alpha+\underset{\sim}{A}\beta)\\ &=|\underset{\sim}{A}\alpha|^2+2(\underset{\sim}{A}\alpha,\underset{\sim}{A}\beta)+|\underset{\sim}{A}\beta|^2=|\alpha|^2+2(\underset{\sim}{A}\alpha,\underset{\sim}{A}\beta)+|\beta|^2\end{aligned}$$

(2)式的右边为

$$|\alpha|^2+2(\alpha,\beta)+|\beta|^2$$

由此得出,$(\underset{\sim}{A}\alpha,\underset{\sim}{A}\beta)=(\alpha,\beta)$. 所以 $\underset{\sim}{A}$ 是正交变换. ▌

从以上讨论知道,实内积空间 $V$ 上的正交变换 $\underset{\sim}{A}$ 是可逆的线性变换,并且保持内积不变,因此 $\underset{\sim}{A}$ 是实内积空间 $V$ 到自身的一个同构映射. 反之,$V$ 到自身的每一个同构映射都是 $V$ 上的正交变换. 我们知道,实内积空间的同构关系具有对称性和传递性,即实内积空间的同构映射的逆映射仍是同构映射;同构映射的乘积还是同构映射. 因此,正交变换的逆变换仍是正交变换,正交变换的乘积仍是正交变换. 从实内积空间的同构映射的性质还知道,当 $V$ 是有限维时,正交变换 $\underset{\sim}{A}$ 把 $V$ 的标准正交基映成 $V$ 的标准正交基,这是欧氏空间上的正交变换的一个特征性质. 即我们有下述定理:

**定理** 12.5.2 设 $\underset{\sim}{A}$ 是欧氏空间 $V$ 的一个线性变换,则下列命题等价:

(i) $\underset{\sim}{A}$ 是正交变换;

(ii) 如果 $\varepsilon_1,\cdots,\varepsilon_n$ 是 $V$ 的一个标准正交基,则 $\underset{\sim}{A}\varepsilon_1,\cdots,\underset{\sim}{A}\varepsilon_n$ 也是 $V$ 的标准正交基;

(iii) $\underset{\sim}{A}$ 在 $V$ 的任意一个标准正交基下的矩阵是正交矩阵.

**证明** (i)⇒(ii). 上面已证.

(ii)⇒(iii). 任取 $V$ 的一个标准正交基 $\varepsilon_1,\cdots,\varepsilon_n$. 由假设知,$\underset{\sim}{A}\varepsilon_1,\cdots,\underset{\sim}{A}\varepsilon_n$ 也是 $V$ 的标准正交基. 从而由基 $\varepsilon_1,\cdots,\varepsilon_n$ 到基 $\underset{\sim}{A}\varepsilon_1,\cdots,\underset{\sim}{A}\varepsilon_n$ 的过渡矩阵 $A$ 是正交矩阵,即

$$(\underset{\sim}{A}\varepsilon_1,\cdots,\underset{\sim}{A}\varepsilon_n)=(\varepsilon_1,\cdots,\varepsilon_n)A \tag{3}$$

(3)式说明，$\underset{\sim}{A}$ 在基 $\varepsilon_1,\cdots,\varepsilon_n$ 下的矩阵是 $A$，而 $A$ 是正交矩阵.

(iii)⇒(i). 取 $V$ 的一个标准正交基 $\varepsilon_1,\cdots,\varepsilon_n$. 由假设，$\underset{\sim}{A}$ 在基 $\varepsilon_1,\cdots,\varepsilon_n$ 下的矩阵 $A$ 是正交矩阵. 由于正交矩阵是可逆的，因此线性变换 $\underset{\sim}{A}$ 是可逆的. 从而 $\underset{\sim}{A}$ 是映上的. 剩下只要证 $\underset{\sim}{A}$ 保持向量的内积不变. 任取 $V$ 中两个向量：

$$\alpha=(\varepsilon_1,\cdots,\varepsilon_n)X,\quad \beta=(\varepsilon_1,\cdots,\varepsilon_n)Y$$

则 $$\underset{\sim}{A}\alpha=(\varepsilon_1,\cdots,\varepsilon_n)(AX),\quad \underset{\sim}{A}\beta=(\varepsilon_1,\cdots,\varepsilon_n)(AY)$$

由于 $\varepsilon_1,\cdots,\varepsilon_n$ 是 $V$ 的标准正交基，所以

$$\begin{aligned}(\underset{\sim}{A}\alpha,\underset{\sim}{A}\beta)&=(AX)'(AY)=X'(A'A)Y\\&=X'Y=(\alpha,\beta)\end{aligned}$$

因此 $\underset{\sim}{A}$ 是正交变换. ▌

既然欧氏空间 $V$ 上的正交变换在 $V$ 的任一标准正交基下的矩阵是正交矩阵，而正交矩阵的行列式等于 $+1$ 或者 $-1$，同此，正交变换的行列式等于 $+1$ 或者 $-1$. 行列式等于 $+1$ 的正交变换称为**旋转**，或者称为**第一类的**；行列式等于 $-1$ 的正交变换称为**第二类的**.

$n$ 维线性空间的任意一个 $n-1$ 维子空间称为一个**超平面**.

**例 1**　在欧氏空间 $V$ 中取一个标准正交基 $\varepsilon_1,\cdots,\varepsilon_n$. 定义 $V$ 上的一个线性变换 $\underset{\sim}{A}$ 为

$$\underset{\sim}{A}\varepsilon_1=-\varepsilon_1,\quad \underset{\sim}{A}\varepsilon_i=\varepsilon_i,\quad i=2,\cdots,n$$

则 $\underset{\sim}{A}$ 在基 $\varepsilon_1,\cdots,\varepsilon_n$ 下的矩阵为

$$A=\operatorname{diag}\{-1,1,\cdots,1\}$$

显然 $A$ 是正交矩阵，因此 $\underset{\sim}{A}$ 是正交变换. 由于 $|A|=-1$，因此 $\underset{\sim}{A}$ 是第二类的. 这个正交变换是关于超平面

$$W=\langle\varepsilon_2,\cdots,\varepsilon_n\rangle$$

的一个镜面反射.

关于镜面反射的定义见本节习题第 4 题.

# 习　题　12.5

1. 设 $V$ 是有限维的实内积空间(即欧氏空间),证明:$V$ 的一个变换 $\underset{\sim}{A}$ 如果保持向量内积不变,则 $\underset{\sim}{A}$ 是正交变换.

2. 设 $V$ 是二维欧氏空间,取 $V$ 的一个标准正交基 $\varepsilon_1,\varepsilon_2$. 设 $\underset{\sim}{A}$ 是 $V$ 的一个正交变换.

(1) 证明:$\underset{\sim}{A}$ 在基 $\varepsilon_1,\varepsilon_2$ 下的矩阵 $A$ 是

$$\begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \quad 或者 \quad \begin{pmatrix} \cos\theta & \sin\theta \\ \sin\theta & -\cos\theta \end{pmatrix}$$

其中 $\theta$ 是某个实数,$0 \leqslant \theta < 2\pi$;

(2) 设 $V = R^2$,其内积为标准内积,$\varepsilon_1 = (1,0)$,$\varepsilon_2 = (0,1)$. 分别对于 $\underset{\sim}{A}$ 在基 $\varepsilon_1,\varepsilon_2$ 下矩阵的两种情形说明 $\underset{\sim}{A}$ 的几何意义.

3. 设 $V$ 是实内积空间,证明:$V$ 上的正交变换 $\underset{\sim}{A}$ 如果有特征值,则它的特征值必为 1 或 $-1$.

4. 设 $V$ 是 $n$ 维欧氏空间,$\eta$ 是 $V$ 中一个单位向量,设 $\underset{\sim}{P}$ 是 $V$ 在 $\langle\eta\rangle$ 上的正交投影. 令

$$\underset{\sim}{A} = \underset{\sim}{I} - 2\underset{\sim}{P}$$

则 $\underset{\sim}{A}$ 称为关于超平面 $\langle\eta\rangle^{\perp}$ 的**镜面反射**,简称为镜面反射. 证明:镜面反射 $\underset{\sim}{A}$ 是正交变换,并且是第二类的.

5. 设 $\underset{\sim}{A}$ 是实内积空间 $V$ 上的一个正交变换,$W$ 是 $\underset{\sim}{A}$ 的有限维不变子空间,证明:$W^{\perp}$ 也是 $\underset{\sim}{A}$ 的不变子空间.

6. 设 $\underset{\sim}{A}$ 是 $n$ 维欧氏空间 $V$ 上的一个正交变换,并且 1 是 $\underset{\sim}{A}$ 的一个特征值,$\underset{\sim}{A}$ 的属于 1 的特征子空间 $V_1$ 是 $n-1$ 维的. 证明:$\underset{\sim}{A}$ 是镜面反射.

7. 设 $\alpha,\beta$ 是欧氏空间 $V$ 中两个不同的单位向量,证明:存在一个镜面反射 $\underset{\sim}{A}$,使得 $\underset{\sim}{A}(\alpha) = \beta$.

* 8. 证明:$n$ 维欧氏空间中任一正交变换都可以表示成一系列镜面反射的乘积.

(提示:利用第 7 题的结果.)

9. 证明:正交矩阵的特征多项式在复数域中的根或者是 1,或者是 $-1$,

或者是 $\cos\theta \pm \mathrm{i}\sin\theta$，$\theta$ 为实数，$0<\theta<\pi$.

*10. 设 $\underline{A}$ 是 $n$ 维欧氏空间 $V$ 上的一个正交变换，证明：存在 $V$ 的一个标准正交基，使得 $\underline{A}$ 在这个基下的矩阵 $A$ 具有形式

$$\operatorname{diag}\left\{(b_1),\cdots,(b_s),\begin{pmatrix}\cos\theta_1 & -\sin\theta_1\\ \sin\theta_1 & \cos\theta_1\end{pmatrix},\cdots,\begin{pmatrix}\cos\theta_m & -\sin\theta_m\\ \sin\theta_m & \cos\theta_m\end{pmatrix}\right\}$$

其中 $0<\theta_i<\pi, i=1,\cdots,m; b_j=1$ 或 $-1, i=1,\cdots,s$.

（提示：对维数 $n$ 作数学归纳法，先分别对于 $n=1$ 和 $n=2$ 的情形证明. 然后注意利用习题 9.6 的第 10 题的结论）.

*11. 证明：任一 $n$ 级正交矩阵一定正交相似于一个第 10 题中所写的分块对角矩阵.

*12. 几何空间中，正交点变换称为第一类的，如果它诱导的正交向量变换是第一类的；否则称为第二类的. 证明：保持原点不动的第一类正交点变换一定是绕过原点的某根直线的旋转.

13. 证明：奇数维欧氏空间中的旋转一定以 1 作为它的一个特征值.

14. 证明：欧氏空间中，第二类正交变换一定以 $-1$ 作为它的一个特征值.

*15. 设 $\alpha_1,\cdots,\alpha_m$ 和 $\beta_1,\cdots,\beta_m$ 是 $n$ 维欧氏空间中两个向量组. 证明：存在一个正交变换 $\underline{A}$，使得

$$\underline{A}\alpha_i=\beta_i,\quad i=1,\cdots,m$$

的充分必要条件为

$$(\alpha_i,\alpha_j)=(\beta_i,\beta_j),\quad i,j=1,\cdots,m$$

## §6 对 称 变 换

本节讨论实内积空间中与内积有关的另一类线性变换.

**定义** 1　实内积空间 $V$ 上的线性变换 $\underline{A}$，如果满足

$$(\underline{A}\alpha,\beta)=(\alpha,\underline{A}\beta),\quad \forall\ \alpha,\beta\in V$$

则称 $\underline{A}$ 是**对称变换**.

**例** 1　设 $W$ 是实内积空间 $V$ 的一个有限维子空间，$V$ 在 $W$ 上的正交投影 $\underline{P}$ 就是一个对称变换.

**证明** 任取 $\alpha,\beta\in V$. 设

$$\alpha=\alpha_1+\alpha_2,\quad \alpha_1\in W,\quad \alpha_2\in W^{\perp}$$

$$\beta=\beta_1+\beta_2,\quad \beta_1\in W,\quad \beta_2\in W^{\perp}$$

据正交投影的定义得，$\underline{P}\alpha=\alpha_1,\underline{P}\beta=\beta_1$. 于是

$$(\underline{P}\alpha,\beta)=(\alpha_1,\beta_1+\beta_2)=(\alpha_1,\beta_1)+(\alpha_1,\beta_2)=(\alpha_1,\beta_1)$$

$$(\alpha,\underline{P}\beta)=(\alpha_1+\alpha_2,\beta_1)=(\alpha_1,\beta_1)+(\alpha_2,\beta_1)=(\alpha_1,\beta_1)$$

由此得出，$(\underline{P}\alpha,\beta)=(\alpha,\underline{P}\beta),\forall\ \alpha,\beta\in V$. ▍

**命题** 12.6.1 设 $\underline{A}$ 是实内积空间 $V$ 上的一个对称变换，如果 $W$ 是 $\underline{A}$ 的不变子空间，则 $W^{\perp}$ 也是 $\underline{A}$ 的不变子空间.

**证明** 任取 $\beta\in W^{\perp}$，要证 $A\beta\in W^{\perp}$. 任取 $\alpha\in W$，有 $\underline{A}\alpha\in W$，于是$(\underline{A}\alpha,\beta)=0$. 从而

$$(\alpha,A\beta)=(\underline{A}\alpha,\beta)=0$$

因此 $\underline{A}\beta\in W^{\perp}$，即 $W^{\perp}$ 是 $\underline{A}$ 的不变子空间. ▍

**命题** 12.6.2 $n$ 维欧氏空间 $V$ 上的线性变换 $\underline{A}$ 是对称变换当且仅当 $\underline{A}$ 在 $V$ 的任意一个标准正交基下的矩阵是对称矩阵.

**证明** 必要性. 任取 $V$ 的一个标准正交基 $\varepsilon_1,\cdots,\varepsilon_n$. 设

$$\underline{A}(\varepsilon_1,\cdots,\varepsilon_n)=(\varepsilon_1,\cdots,\varepsilon_n)A,\quad A=(a_{ij})$$

由于向量 $\alpha$ 在标准正交基 $\varepsilon_1,\cdots,\varepsilon_n$ 下的坐标的第 $k$ 个分量是 $(\alpha,\varepsilon_k)$，$k=1,\cdots,n$，所以

$$a_{ij}=\underline{A}\varepsilon_j\text{ 在基 }\varepsilon_1,\cdots,\varepsilon_n\text{ 下坐标的第 }i\text{ 个分量}=(\underline{A}\varepsilon_j,\varepsilon_i)$$

$$a_{ji}=A\varepsilon_i\text{ 在基 }\varepsilon_1,\cdots,\varepsilon_n\text{ 下坐标的第 }j\text{ 个分量}=(\underline{A}\varepsilon_i,\varepsilon_j)$$

由于 $\underline{A}$ 是对称变换，所以

$$a_{ij}=(\underline{A}\varepsilon_j,\varepsilon_i)=(\varepsilon_j,\underline{A}\varepsilon_i)=(A\varepsilon_i,\varepsilon_j)=a_{ji},\quad i,j=1,\cdots,n$$

于是 $A$ 是对称矩阵.

充分性. 设线性变换 $\underline{A}$ 在 $V$ 的一个标准正交基 $\varepsilon_1,\cdots,\varepsilon_n$ 下的矩阵 $A$ 是对称矩阵. 在 $V$ 中任取两个向量：

$$\alpha=(\varepsilon_1,\cdots,\varepsilon_n)X,\quad \beta=(\varepsilon_1,\cdots,\varepsilon_n)Y$$

则 $$\underline{A}\alpha=(\varepsilon_1,\cdots,\varepsilon_n)(AX),\quad \underline{A}\beta=(\varepsilon_1,\cdots,\varepsilon_n)(AY)$$

从而 $(\underline{A}\alpha,\beta)=(AX)'Y=X'A'Y=X'AY=(\alpha,\underline{A}\beta)$

所以 $\underline{A}$ 是对称变换. ▌

第五章 § 9 的定理 5.9.4 指出，实对称矩阵一定正交相似于一个对角矩阵. 由此得出

**定理** 12.6.1 设 $\underline{A}$ 是 $n$ 维欧氏空间 $V$ 上的一个对称变换，则存在 $V$ 的一个标准正交基，使得 $\underline{A}$ 在该基下的矩阵为对角矩阵. ▌

也可以用空间 $V$ 分解成 $\underline{A}$ 的不变子空间的直和的方法给出定理 12.6.1 另一证法，留给读者作为习题. 证明中注意利用对称变换的特征多项式在复数域中的根全是实数(因为实对称矩阵有此性质)，以及利用命题 12.6.1.

## 习 题 12.6

1. 用空间分解的方法证明定理 12.6.1.

2. 证明：实数域上的斜对称矩阵的特征多项式在复数域中的根是零或纯虚数.

## 补充题十二

1. 实内积空间 $V$ 上的线性变换 $\underline{A}$ 称为**斜对称的**，如果

$$(\underline{A}\alpha,\beta)=-(\alpha,\underline{A}\beta),\quad \forall\ \alpha,\beta\in V$$

证明：如果 $W$ 是 $\underline{A}$ 的不变子空间，则 $W^{\perp}$ 也是.

2. 证明：$n$ 维欧氏空间 $V$ 上的线性变换 $\underline{A}$ 是斜对称的当且仅当 $\underline{A}$ 在 $V$ 的任意一个标准正交基下的矩阵是斜对称矩阵.

3. 设 $\underline{A}$ 是 $n$ 维欧氏空间 $V$ 上的一个斜对称变换，证明：存在 $V$ 的一个标准正交基，使得 $\underline{A}$ 在这个基下的矩阵具有如下形式：

$$\mathrm{diag}\left\{\begin{pmatrix}0 & a_1\\ -a_1 & 0\end{pmatrix},\cdots,\begin{pmatrix}0 & a_s\\ -a_s & 0\end{pmatrix},(0),\cdots,(0)\right\}$$

其中，$a_1,\cdots,a_s$ 都是实数.

（提示：对维数 $n$ 用数学归纳法. 先对于 $n=1$ 和 $n=2$ 的情形分别讨论. 然后注意利用习题 9.6 的第 10 题的结论，把空间 $V$ 直和分解.）

# 第十三章　酉　空　间

这一章我们要在复数域上的线性空间中引进度量概念. 从上一章看到,关键是要引进内积的概念. 但是在复线性空间中不能照搬实线性空间中内积的定义,否则会出现矛盾: 设 $\alpha$ 是复线性空间 $V$ 中一个非零向量,据正定性,有 $(\alpha,\alpha) > 0$. 由于 $\mathrm{i}\alpha$ 也是 $V$ 中非零向量,于是也有 $(\mathrm{i}\alpha,\mathrm{i}\alpha) > 0$. 又据双线性,有 $(\mathrm{i}\alpha,\mathrm{i}\alpha) = \mathrm{i}^2(\alpha,\alpha) = -(\alpha,\alpha)$,于是得出 $(\alpha,\alpha) < 0$. 矛盾. 由此可见,为了使复线性空间的内积仍然具有正定性(这条性质对于引进长度等度量概念是必要的),就必须把"双线性函数"这条要求减弱. 如果我们只要求对第一个变量是线性的. 而对第二个变量只要求是半线性的,即

$$(\alpha,\beta_1+\beta_2) = (\alpha,\beta_1) + (\alpha,\beta_2)$$

$$(\alpha,k\beta) = \bar{k}(\alpha,\beta)$$

那么 $(\mathrm{i}\alpha,\mathrm{i}\alpha) = \mathrm{i}\bar{\mathrm{i}}(\alpha,\alpha) = (\alpha,\alpha)$,这样就不会出现上述矛盾了. 此外,为了使 $(\alpha,\alpha)$ 为实数,对称性应当代之以 Hermite 性,即

$$(\alpha,\beta) = \overline{(\beta,\alpha)}$$

我们将在§1具体给出复线性空间上内积的定义. 有了内积后,就可引进长度、角度、正交、距离等度量概念,这些与实内积空间一样.

## §1　内积·复内积空间·正交补

**定义 1**　设 $V$ 是一个复线性空间,$V$ 上的一个函数 $f$(即 $V$ 到 $C$ 的一个映射)称为是**半线性的**,如果

$$f(\alpha+\beta) = f(\alpha) + f(\beta),\quad \forall\, \alpha,\beta \in V \qquad (1)$$

$$f(k\alpha)=\bar{k}f(\alpha),\quad \forall\,\alpha\in V,k\in C \tag{2}$$

$V$ 上的一个二元函数 $g$(即 $V\times V$ 到 $C$ 的一个映射)称为**一个半线性函数**(sesquilinear function),或者**共轭双线性函数**,如果 $g(\alpha,\beta)$ 对于第一个变量是线性的,而对于第二个变量是半线性的,即

$$g(k_1\alpha_1+k_2\alpha_2,\beta)=k_1g(\alpha_1,\beta)+k_2g(\alpha_2,\beta) \tag{3}$$

$$g(\alpha,k_1\beta_1+k_2\beta_2)=\bar{k}_1g(\alpha_1,\beta_1)+\bar{k}_2g(\alpha_2,\beta_2) \tag{4}$$

其中 $\alpha_1,\alpha_2,\beta,\alpha,\beta_1,\beta_2$ 是 $V$ 中任意向量,$k_1,k_2$ 是任意复数.

**定义** 2　复线性空间 $V$ 上的一个二元函数 $f$ 称为是 **Hermite 的**,如果

$$f(\alpha,\beta)=\overline{f(\beta,\alpha)},\quad \forall\,\alpha,\beta\in V \tag{5}$$

**定义** 3　复线性空间 $V$ 上的一个二元函数 $f$ 称为**正定的**,如果对于一切 $\alpha\in V$,有 $f(\alpha,\alpha)$ 是非负实数,并且 $f(\alpha,\alpha)=0$ 当且仅当 $\alpha=0$.

**定义** 4　复线性空间 $V$ 上的一个正定的、Hermite 的、一个半线性函数 $f$ 称为 $V$ 上的一个**内积**.

习惯上,用记号( , )或者( | )表示内积.

**定义** 5　复线性空间 $V$ 上如果指定了一个内积,则称 $V$ 是一个**复内积空间**. 有限维的复内积空间称为**酉空间**.

注:有些文献把无限维的复内积空间也称为酉空间. 本书则只是把有限维的复内积空间称为酉空间,这样在叙述上方便一些.

**例** 1　在 $C^n$ 中,对于任意

$$\alpha=(x_1,\cdots,x_n),\quad \beta=(y_1,\cdots,y_n)$$

定义
$$(\alpha,\beta):=x_1\bar{y}_1+x_2\bar{y}_2+\cdots+x_n\bar{y}_n \tag{6}$$

容易验证,( , )是一个内积,这个内积称为 $C^n$ 上的**标准内积**. $C^n$ 装备上这个标准内积,便成为一个酉空间.

**例** 2　用 $\widetilde{C}[a,b]$ 表示区间$[a,b]$上的所有连续复值函数组成的线性空间. 定义

$$(f,g):=\int_a^b f(x)\overline{g(x)}\mathrm{d}x \tag{7}$$

容易验证,( , )是一个内积.$\tilde{C}[a,b]$ 装备上这个内积,便成为一个复内积空间.

**例** 3 在 $M_n(C)$ 中,定义

$$(A,B):=Tr(AB^*) \tag{8}$$

其中 $B^*$ 表示 $B$ 中所有元素取共轭复数后再转置,容易验证( , )是 $M_n(C)$ 上的一个内积,从而 $M_n(C)$ 连同这个内积一起成为酉空间.

与实内积空间一样,在复内积空间 $V$ 中可以引进向量的长度的概念:

**定义** 6 非负实数 $\sqrt{(\alpha,\alpha)}$ 称为向量 $\alpha$ 的**长度**,记作 $|\alpha|$(或者 $\|\alpha\|$).

显然 $|0|=0$;$\alpha\neq 0$ 时,$|\alpha|>0$.

**性质** 1 $$|k\alpha|=|k||\alpha|,\quad k\in C,\quad \alpha\in V \tag{9}$$

这里 $|k|$ 表示复数 $k$ 的模.

**证明** $|k\alpha|^2=(k\alpha,k\alpha)=k\bar{k}(\alpha,\alpha)=|k|^2|\alpha|^2$

因此 $|k\alpha|=|k||\alpha|$ ▎

**定理** 13.1.1(Cauchy-Buniakowski **不等式**) 在复内积空间 $V$ 中,对于任意向量 $\alpha,\beta$,有

$$|(\alpha,\beta)|\leqslant|\alpha||\beta| \tag{10}$$

等号成立当且仅当 $\alpha,\beta$ 线性相关.

**证明** 当 $\alpha,\beta$ 线性相关时,与实内积空间的情形一样可证出

$$|(\alpha,\beta)|=|\alpha||\beta|$$

如果 $\alpha,\beta$ 线性无关,则对任意复数 $t$,有

$$\alpha+t\beta\neq 0$$

于是 $0<|\alpha+t\beta|^2=|\alpha|^2+\bar{t}(\alpha,\beta)+t(\beta,\alpha)+t\bar{t}|\beta|^2$

特别地,取 $t=-\dfrac{(\alpha,\beta)}{|\beta|^2}$,代入上式得

$$0<|\alpha|^2-\frac{|(\alpha,\beta)|^2}{|\beta|^2}-\frac{|(\alpha,\beta)|^2}{|\beta|^2}+\frac{|(\alpha,\beta)|^2}{|\beta|^2}$$
$$=|\alpha|^2-\frac{|(\alpha,\beta)|^2}{|\beta|^2}$$

于是得出 $|(\alpha,\beta)|<|\alpha||\beta|$. ▌

把 Cauchy-Buniakowski 不等式分别应用到例 1,例 2,例 3 的复内积空间中,便得到

$$|\sum_{i=1}^{n}x_i\bar{y}_i|\leqslant(\sum_{i=1}^{n}|x_i|^2)^{\frac{1}{2}}(\sum_{i=1}^{n}|y_i|^2)^{\frac{1}{2}} \tag{11}$$

$$\left|\int_a^b f(x)\,\overline{g(x)}\mathrm{d}x\right|\leqslant\left(\int_a^b|f(x)|^2\mathrm{d}x\right)^{\frac{1}{2}}\left(\int_a^b|g(x)|^2\mathrm{d}x\right)^{\frac{1}{2}} \tag{12}$$

$$|Tr(AB^*)|\leqslant(Tr(AA^*))^{\frac{1}{2}}(Tr(BB^*))^{\frac{1}{2}} \tag{13}$$

**定义** 7　复内积空间 $V$ 中,两个非零向量 $\alpha,\beta$ 的**夹角** $\langle\alpha,\beta\rangle$ 规定为

$$\langle\alpha,\beta\rangle=\arccos\frac{|(\alpha,\beta)|}{|\alpha||\beta|} \tag{14}$$

于是 $0\leqslant\langle\alpha,\beta\rangle\leqslant\frac{\pi}{2}$.

在自然科学中利用复内积空间的一个非常重要的模型,不仅把 $\frac{|(\alpha,\beta)|}{|\alpha||\beta|}$ 解释成 $\alpha$ 与 $\beta$ 的夹角的余弦,而且把这个量的平方(即 $\cos^2\langle\alpha,\beta\rangle$) 解释成**概率**. 有兴趣的读者可以看量子力学的书.

从(14)式看出, $\langle\alpha,\beta\rangle=\frac{\pi}{2}\Leftrightarrow(\alpha,\beta)=0$.

因此我们引进下述概念:

**定义** 8　在复内积空间 $V$ 中,如果 $(\alpha,\beta)=0$,则称 $\alpha$ 与 $\beta$ **正交**,记为 $\alpha\perp\beta$.

显然, $\alpha$ 与自己正交当且仅当 $\alpha=0$.

我们用 $\mathrm{Re}z$ 表示复数 $z$ 的实部,用 $\mathrm{Im}z$ 表示复数 $z$ 的虚部.

**推论** 13.1.1(**三角形不等式**)　在复内积空间 $v$ 中,对于任意两个向量 $\alpha,\beta$,有

$$|\alpha+\beta|\leqslant|\alpha|+|\beta| \tag{15}$$

**证明** $|\alpha+\beta|^2=|\alpha|^2+(\alpha,\beta)+\overline{(\alpha,\beta)}+|\beta|^2$

$$\begin{aligned}&=|\alpha|^2+2\mathrm{Re}(\alpha,\beta)+|\beta|^2\\&\leqslant|\alpha|^2+2|(\alpha,\beta)|+|\beta|^2\\&\leqslant|\alpha|^2+2|\alpha||\beta|+|\beta|^2=(|\alpha|+|\beta|)^2\end{aligned}$$

由此得出 $|\alpha+\beta|\leqslant|\alpha|+|\beta|$. ▌

**推论** 13.1.2(**勾股定理**) 如果 $\alpha\perp\beta$,则

$$|\alpha+\beta|^2=|\alpha|^2+|\beta|^2 \tag{16}$$

**证明** 从推论 13.1.1 的证明过程可看出. ▌

用归纳法可把勾股定理推广为:如果 $\alpha_1,\alpha_2,\cdots,\alpha_s$ 两两正交,则

$$|\alpha_1+\alpha_2+\cdots+\alpha_s|^2=|\alpha_1|^2+|\alpha_2|^2+\cdots+|\alpha_s|^2 \tag{17}$$

**推论** 13.1.3 在复内积空间 $V$ 中,对于任意两个向量 $\alpha,\beta$,定义

$$d(\alpha,\beta):=|\alpha-\beta| \tag{18}$$

则 $d$ 是一个距离. ▌

与实内积空间一样,在复内积空间中有**正交集**、**正交规范集**的概念. 并且有

**定理** 13.1.2 在复内积空间 $V$ 中,由非零向量组成的正交集是线性无关的.

**证明** 与实内积空间的情况一样证明. ▌

于是,在 $n$ 维酉空间中,两两正交的非零向量的数目不会超过空间的维数 $n$.

酉空间中也有正交基、标准正交基的概念,它们的定义与欧氏空间中一样. 它们是否存在?

**定理** 13.1.3 $n$ 维酉空间 $V$ 中,一定存在标准正交基.

**证明** 先取 $V$ 的一个基 $\alpha_1,\cdots,\alpha_n$. 对它施行 Gram-Schmidt 正交化过程,可得到一个与 $\alpha_1,\cdots,\alpha_n$ 等价的正交向量组 $\beta_1,\cdots,\beta_n$. 于是 $\beta_1,\cdots,\beta_n$ 是 $V$ 的一个正交基. 把每个 $\beta_i$ 单位化,即令

$$\eta_i=\frac{1}{|\beta_i|}\beta_i,\quad i=1,\cdots,n$$

则 $\eta_1,\cdots,\eta_n$ 就是 $V$ 的一个标准正交基. ▌

$n$ 维酉空间 $V$ 中,向量组 $\eta_1,\cdots,\eta_n$ 是 $V$ 的一个标准正交基当且仅当

$$(\eta_i,\eta_j)=\delta_{ij},\quad i,j=1,\cdots,n \tag{19}$$

**命题** 13.1.1 设 $\eta_1,\cdots,\eta_n$ 是酉空间 $V$ 的一个标准正交基,对于 $\alpha=(\eta_1,\cdots,\eta_n)X,\beta=(\eta_1,\cdots,\eta_n)Y$,有

$$(\alpha,\beta)=Y^*X \tag{20}$$

**证明** 设 $X'=(x_1,\cdots,x_n),Y'=(y_1,\cdots,y_n)$,则

$$(\alpha,\beta)=(\sum_{i=1}^n x_i\eta_i,\sum_{j=1}^n y_j\eta_j)$$

$$=\sum_{i=1}^n\sum_{j=1}^n x_i\bar{y}_j(\eta_i,\eta_j)=\sum_{i=1}^n x_i\bar{y}_i=Y^*X \tag{21}$$

▌

**命题** 13.1.2 设 $\varepsilon_1,\cdots,\varepsilon_n$ 是酉空间 $V$ 的一个标准正交基,则对于任意 $\alpha\in V$,有

$$\alpha=\sum_{i=1}^n(\alpha,\varepsilon_i)\varepsilon_i \tag{22}$$

**证明** 设 $\alpha=\sum_{i=1}^n x_i\varepsilon_i$,则由(21)式得

$$(\alpha,\varepsilon_j)=x_j$$ ▌

(22)式中的系数称为 Fourier 系数.

酉空间 $V$ 中标准正交基不唯一. 设 $\varepsilon_1,\cdots,\varepsilon_n$ 与 $\eta_1,\cdots,\eta_n$ 都是 $V$ 的标准正交基,且

$$(\eta_1,\cdots,\eta_n)=(\varepsilon_1,\cdots,\varepsilon_n)B$$

设 $B$ 的列向量组是 $B_1,\cdots,B_n$,则 $B_j$ 是 $\eta_j$ 在基 $\varepsilon_1,\cdots,\varepsilon_n$ 下的坐标. 于是

$$\delta_{ji}=\delta_{ij}=(\eta_i,\eta_j)=B_j^*B_i,\quad i,j=1,\cdots,n \tag{23}$$

由于 $B_j^*$ 是 $B^*$ 的第 $j$ 行,$B_i$ 是 $B$ 的第 $i$ 列,所以

$$B^*B=I \tag{24}$$

**定义** 9　$n$ 级复矩阵 $B$ 如果满足

$$B^* B = I$$

则称 $B$ 是**酉矩阵**.

显然，酉矩阵 $B$ 是可逆的，并且 $B^{-1} = B^*$. 从而 $BB^* = I$. 反之，若 $BB^* = I$，则同理可推出 $B^* B = I$，从而 $B$ 为酉矩阵.

上面我们证明了

**命题** 13.1.3　酉空间中，由标准正交基到标准正交基的过渡矩阵是酉矩阵. ▌

反之，我们有

**命题** 13.1.4　$n$ 维酉空间 $V$ 中，设 $\varepsilon_1,\cdots,\varepsilon_n$ 是 $V$ 的一个标准正交基，设向量组 $\eta_1,\cdots,\eta_n$ 满足

$$(\eta_n,\cdots,\eta_n) = (\varepsilon_1,\cdots,\varepsilon_n)B$$

其中 $B$ 是酉矩阵，则 $\eta_1,\cdots,\eta_n$ 是 $V$ 的一个标准正交基.

**证明**　$\eta_j$ 在基 $\varepsilon_1,\cdots,\varepsilon_n$ 下的坐标是 $B$ 的第 $j$ 个列向量 $B_j$. 由于 $B$ 是酉矩阵，所以

$$(\eta_i,\eta_j) = B_j^* B_i = \delta_{ji} = \delta_{ij},\quad i,j = 1,\cdots,n$$

这表明 $\eta_1,\cdots,\eta_n$ 是 $V$ 的标准正交基. ▌

设 $V$ 是复内积空间，则对于一切 $\alpha,\beta \in V$，有

$$(\alpha,\beta) = \mathrm{Re}(\alpha,\beta) + \mathrm{iIm}(\alpha,\beta) \tag{25}$$

如果 $z$ 是复数，则 $\mathrm{Im}(z) = \mathrm{Re}(-\mathrm{i}z)$. 于是

$$\mathrm{Im}(\alpha,\beta) = \mathrm{Re}[-\mathrm{i}(\alpha,\beta)] = \mathrm{Re}(\alpha,\mathrm{i}\beta) \tag{26}$$

由(25)和(26)得出

$$(\alpha,\beta) = \mathrm{Re}(\alpha,\beta) + \mathrm{iRe}(\alpha,\mathrm{i}\beta) \tag{27}$$

在复内积空间 $V$ 中，由内积(　,　)可诱导出 $V$ 上的一个函数 $q$：

$$q(\alpha) := (\alpha,\alpha) = |\alpha|^2,\quad \forall\,\alpha \in V \tag{28}$$

称 $q$ 是 $V$ 上的 Hermite 二次函数.

从内积的性质得出，对于一切 $\alpha,\beta \in V$，有

$$q(\alpha \pm \beta) = |\alpha \pm \beta|^2 = |\alpha|^2 \pm 2\mathrm{Re}(\alpha,\beta) + |\beta|^2 \tag{29}$$

$$q(\alpha \pm \mathrm{i}\beta) = |\alpha \pm \mathrm{i}\beta|^2 = |\alpha|^2 \pm 2\mathrm{Re}(\alpha,\mathrm{i}\beta) + |\beta|^2 \quad (30)$$

从(27),(29)和(30)式得出

$$\begin{aligned}(\alpha,\beta) &= \frac{1}{4}|\alpha+\beta|^2 - \frac{1}{4}|\alpha-\beta|^2 \\ &\quad + \frac{\mathrm{i}}{4}|\alpha+\mathrm{i}\beta|^2 - \frac{\mathrm{i}}{4}|\alpha-\mathrm{i}\beta|^2 \qquad (31)\\ &= \frac{1}{4}q(\alpha+\beta) - \frac{1}{4}q(\alpha-\beta) \\ &\quad + \frac{\mathrm{i}}{4}q(\alpha+\mathrm{i}\beta) - \frac{\mathrm{i}}{4}q(\alpha-\mathrm{i}\beta) \qquad (32)\end{aligned}$$

(31)式称为**极化恒等式**,它也可以写成

$$(\alpha,\beta) = \frac{1}{4}\sum_{m=1}^{4} i^m |\alpha + \mathrm{i}^m\beta|^2 \quad (33)$$

(32)式表明,内积可以由 Hermite 二次函数 $q$ 确定.

与实内积空间一样,在复内积空间 $V$ 中有子集 $S$ 的**正交补**的概念,并且任一子集 $S$ 的正交补 $S^{\perp}$ 一定是 $V$ 的子空间.

**定理** 13.1.4　设 $U$ 是复内积空间 $V$ 的一个有限维子空间,则

$$V = U \oplus U^{\perp} \quad (34)$$

**证明**　与定理 12.3.1 一样. ▌

如果 $U$ 是复内积空间 $V$ 的有限维子空间,则也有 $V$ 在 $U$ 上的**正交投影**,向量 $\alpha$ 在 $U$ 上的**正交投影**的概念. 也有

**定理** 13.1.5　设 $U$ 是复内积空间 $V$ 的一个有限维子空间,对于 $\alpha \in V$, $\alpha_1 \in U$ 是 $\alpha$ 在 $U$ 上的正交投影的充分必要条件为

$$d(\alpha,\alpha_1) \leqslant d(\alpha,\gamma), \quad \forall\, \gamma \in U \quad (35)$$

**证明**　与定理 12.3.2 一样. ▌

复内积空间中也有向量 $\alpha$ 在子空间 $U$ 上的**最佳逼近元**的概念.

* 利用最佳逼近元可以定义 $V$ 在任一子空间 $U$ 上的**正交投影**的概念. 并且习题 12.3 的第 6,7,8,9,10,11,12 题对于复内积空间仍然成立.

*在酉空间中也有Bessel**不等式**成立(见习题12.3的第13题),注意这时不等式为

$$\sum_{i=1}^{n}|(\alpha,\varepsilon_i)|^2\leqslant|\alpha|^2$$

两个复内积空间**同构**的定义与实内积空间的情形一样.并且有以下定理.

**定理** 13.1.6 两个酉空间同构的充分必要条件是它们的维数相同. ▍

## 习 题 13.1

1. 对于复线性空间的情形,证明习题12.1的第4题.

*2. 求出 $C^1$ 上的所有内积.

*3. 在复内积空间 $\tilde{C}[0,2\pi]$ 中,指定内积为

$$(f,g)=\int_0^{2\pi}f(x)\overline{g(x)}\mathrm{d}x$$

证明:$\tilde{C}[0,2\pi]$ 的一个子集

$$S=\left\{\frac{1}{\sqrt{2\pi}}\mathrm{e}^{\mathrm{i}nx}\,\middle|\,n=0,\pm1,\pm2,\cdots\right\}$$

是一个正交规范集.

*4. 对于复内积空间,证明习题12.3的第6题.

*5. 对于复内积空间,证明习题12.3的第10题.

*6. 对于复内积空间,证明习题12.3的第11题.

7. 对于复内积空间,证明习题12.3的第13题.

8. 在酉空间 $C^3$ 中,指定的是标准内积.设

$$\alpha_1=(\mathrm{i},-1,\mathrm{i}),\quad\alpha_2=(1,0,\mathrm{i}),\quad\alpha_3=(1,1,1)$$

求 $C^3$ 的一个标准正交基.

9. 证明:酉矩阵的特征值的模为1.

10. 证明:上三角的酉矩阵必为对角矩阵,并且主对角元的模等于1.

11. 证明:任一可逆复矩阵 $A$ 一定可以分解成 $A=UT$,其中 $U$ 是酉矩

阵,$T$ 是主对角元都为正实数的上三角矩阵;并且这种分解法是唯一的.

12. 证明:复内积空间 $V$ 中,如果向量 $\beta$,使得 $(\alpha,\beta)=0, \forall\, \alpha\in V$,则 $\beta=0$.

## §2 酉变换・Hermite 变换

从这一节开始我们来讨论复内积空间中,与内积有关的线性变换.

**定义 1** 复内积空间 $V$ 到自身上的一个变换 $\underset{\sim}{A}$,如果满足

$$(\underset{\sim}{A}\,\alpha,\underset{\sim}{A}\,\beta)=(\alpha,\beta),\quad \forall\, \alpha,\beta\in V \tag{1}$$

则称 $\underset{\sim}{A}$ 是 $V$ 上的一个**酉变换**.

**命题** 13.2.1 复内积空间 $V$ 上的酉变换一定是线性变换,并且是单射,从而是可逆的.

**证明** 与实内积空间情形类似. ▌

从定义 1 得出,复内积空间 $V$ 上的酉变换保持向量的长度不变,保持两个非零向量的夹角不变,保持正交性不变. 再由命题 13.2.1 得出,酉变换保持两个向量之间的距离不变.

**定理** 13.2.1 复内积空间 $V$ 到自身上的变换 $\underset{\sim}{A}$ 是酉变换当且仅当 $\underset{\sim}{A}$ 是线性变换,并且保持向量的长度不变.

**证明** 必要性是显然的. 现在证充分性. 对于任意 $\alpha,\beta\in V$,据极化恒等式,有

$$(\alpha,\beta)=\frac{1}{4}\sum_{m=1}^{4}\mathrm{i}^m\,|\alpha+\mathrm{i}^m\beta|^2$$

$$\begin{aligned}(\underset{\sim}{A}\,\alpha,\underset{\sim}{A}\,\beta)&=\frac{1}{4}\sum_{m=1}^{4}\mathrm{i}^m\,|\underset{\sim}{A}\,\alpha+\mathrm{i}^m\underset{\sim}{A}\,\beta|^2\\&=\frac{1}{4}\sum_{m=1}^{4}\mathrm{i}^m\,|\underset{\sim}{A}\,(\alpha+\mathrm{i}^m\beta)|^2\\&=\frac{1}{4}\sum_{m=1}^{4}\mathrm{i}^m\,|\alpha+\mathrm{i}^m\beta|^2=(\alpha,\beta)\end{aligned}$$ ▌

由于复内积空间 $V$ 上的酉变换 $\underset{\sim}{A}$ 是可逆的线性变换，并且保持内积不变，所以 $\underset{\sim}{A}$ 是复内积空间 $V$ 到自身的一个同构映射. 由于同构关系具有对称性和传递性，所以酉变换的逆变换还是酉变换，酉变换的乘积还是酉变换.

**定理** 13.2.2　设 $\underset{\sim}{A}$ 是 $n$ 维酉空间 $V$ 上的一个线性变换，则下列命题等价：

(i) $\underset{\sim}{A}$ 是酉变换；

(ii) $\underset{\sim}{A}$ 把 $V$ 的标准正交基映成标准正交基；

(iii) $\underset{\sim}{A}$ 在 $V$ 的任一标准正交基下的矩阵是酉矩阵.

**证明**　(i) ⇒ (ii). 因为酉空间之间的同构映射把标准正交基变成标准正交基，所以酉变换有此性质.

(ii) ⇒ (iii). 任取 $V$ 的一个标准正交基 $\varepsilon_1\cdots,\varepsilon_n$. 由假设知 $\underset{\sim}{A}\,\varepsilon_1,\cdots,\underset{\sim}{A}\,\varepsilon_n$ 也是标准正交基. 于是 $\varepsilon_1,\cdots,\varepsilon_n$ 到 $\underset{\sim}{A}\,\varepsilon_1,\cdots,\underset{\sim}{A}\,\varepsilon_n$ 的过渡矩阵 $A$ 是酉矩阵，而这就是 $\underset{\sim}{A}$ 在基 $\varepsilon_1,\cdots\varepsilon_n$ 下的矩阵.

(iii) ⇒ (i). 类似于欧氏空间的情形.　▌

**定义** 2　复内积空间 $V$ 上的线性变换 $\underset{\sim}{A}$ 如果满足

$$(\underset{\sim}{A}\,\alpha,\beta)=(\alpha,\underset{\sim}{A}\,\beta),\quad \forall\ \alpha,\beta\in V \tag{2}$$

则称 $\underset{\sim}{A}$ 是 Hermite **变换**，或者**自伴(随)变换**.

从定义看出，复内积空间中的 Hermite 变换与实内积空间中的对称变换类似.

Hermite 变换是重要的，它有许多特殊的性质，并且在实际问题中引出的许多变换是自伴变换.

**定理** 13.2.3　$n$ 维酉空间 $V$ 上的线性变换 $\underset{\sim}{A}$ 是 Hermite 变换当且仅当 $\underset{\sim}{A}$ 在 $V$ 的任意一个标准正交基下的矩阵 $A$ 满足

$$A^*=A \tag{3}$$

满足(3)式的复矩阵称为 Hermite **矩阵**，或者**自伴矩阵**.

**证明**　必要性. 任取 $V$ 的一个标准正交基 $\varepsilon_1,\cdots\varepsilon_n$. 设 $\underset{\sim}{A}$ 在此基下的矩阵为 $A=(a_{ij})$. 于是

$$(\underset{\sim}{A}\,\varepsilon_j,\varepsilon_i)=a_{ij},\quad (\underset{\sim}{A}\,\varepsilon_i,\varepsilon_j)=a_{ji}$$

从而 $a_{ji}=(\underset{\sim}{A}\varepsilon_i,\varepsilon_j)=(\varepsilon_i,\underset{\sim}{A}\varepsilon_j)=\overline{(\underset{\sim}{A}\varepsilon_j,\varepsilon_i)}=\bar{a}_{ij}$

因此 $A^*=A$.

充分性. 设线性变换 $\underset{\sim}{A}$ 在 $V$ 的一个标准正交基 $\varepsilon_1,\cdots\varepsilon_n$ 下的矩阵 $A$ 满足 $A^*=A$. 在 $V$ 中任取

$$\alpha=(\varepsilon_1,\cdots,\varepsilon_n)X,\quad \beta=(\varepsilon_1,\cdots\varepsilon_n)Y$$

则 $(\underset{\sim}{A}\alpha,\beta)=Y^*AX=Y^*A^*X=(AY)^*X=(\alpha,\underset{\sim}{A}\beta)$

所以 $\underset{\sim}{A}$ 是 Hermite 变换. ▌

**定理** 13.2.4 酉空间 $V$ 中的 Hermite 变换 $\underset{\sim}{A}$ 的特征值一定是实数.

**证明** 设 $\lambda$ 是 $\underset{\sim}{A}$ 的任一特征值,设 $\underset{\sim}{A}\alpha=\lambda\alpha,\alpha\neq 0$. 因为

$$(\underset{\sim}{A}\alpha,\alpha)=(\lambda\alpha,\alpha)=\lambda|\alpha|^2$$
$$(\alpha,\underset{\sim}{A}\alpha)=(\alpha,\lambda\alpha)=\bar{\lambda}|\alpha|^2$$

所以 $\lambda|\alpha|^2=\bar{\lambda}|\alpha|^2$. 因为 $\alpha\neq 0$,所以 $|\alpha|\neq 0$. 于是得出 $\lambda=\bar{\lambda}$,即 $\lambda$ 是实数. ▌

## 习 题 13.2

1. 写出 1 级酉矩阵的形式.

2. 证明:酉矩阵的行列式的模为 1.

3. 写出 2 级酉矩阵的形式.

*4. 复内积空间 $V$ 上的线性变换 $\underset{\sim}{A}$ 称为**斜** Hermite **的**,如果

$$(\underset{\sim}{A}\alpha,\beta)=-(\alpha,\underset{\sim}{A}\beta),\quad \forall\ \alpha,\beta\in V$$

证明:$n$ 维酉空间 $V$ 上的线性变换 $\underset{\sim}{A}$ 是斜 Hermite 的充分必要条件是,$\underset{\sim}{A}$ 在 $V$ 的任意一个标准正交基下的矩阵 $A$ 满足

$$A^*=-A$$

这种矩阵称为**斜** Hermite **矩阵**.

*5. 证明:若 $A$ 是 Hermite 矩阵,则 i$A$ 是斜 Hermite 矩阵;若 $B$ 是斜 Hermite 矩阵,则 i$B$ 是 Hermite 矩阵.

6. 设 $n$ 级复矩阵 $A=P+\mathrm{i}Q$,其中 $P,Q$ 都是实矩阵. 证明:$A$ 为酉矩阵

的充分必要条件是，$P'Q$ 为对称矩阵，并且 $P'P + Q'Q = I$.

## §3　线性变换的伴随变换

这一节我们要在复(实)内积空间中，讨论一个线性变换的"伴随变换". 下一节我们将利用线性变换 $\underline{A}$ 的伴随变换来研究 $\underline{A}$ 的性质. 为了证明有限维复(实) 内积空间中每个线性变换都存在伴随变换，我们需要先讨论由内积诱导的线性函数.

设 $V$ 是一个复(实) 内积空间，$\beta$ 是 $V$ 中某个固定的向量，我们定义一个 $V$ 上的函数 $f_\beta$ 如下：

$$f_\beta(\alpha) = (\alpha,\beta), \quad \forall\, \alpha \in V \tag{1}$$

显然 $f_\beta$ 是 $V$ 上的线性函数(这从内积的定义知道). 如果 $V$ 是有限维的，则 $V$ 上的每个线性函数都可用这种方法从某个 $\beta$ 产生，即

**定理** 13.3.1　设 $V$ 是 $n$ 维复(实) 内积空间，$f$ 是 $V$ 上的一个线性函数. 则存在 $V$ 中唯一的向量 $\beta$，使得 $f = f_\beta$，即

$$f(\alpha) = (\alpha,\beta), \quad \forall\, \alpha \in V \tag{2}$$

**证明**　设 $\alpha_1,\cdots,\alpha_n$ 是 $V$ 的一个标准正交基. 令

$$\beta = \sum_{j=1}^{n} \overline{f(\alpha_j)}\alpha_j \tag{3}$$

则

$$f_\beta(\alpha_k) = (\alpha_k, \sum_{j=1}^{n} \overline{f(\alpha_j)}\alpha_j) = f(\alpha_k)$$

因此 $f = f_\beta$.

再证唯一性. 假设还有 $\gamma \in V$ 使得 $f = f_\gamma$，则

$$(\alpha,\beta) = (\alpha,\gamma), \quad \forall\, \alpha \in V$$

于是 $(\alpha,\beta - \gamma) = 0, \forall\, \alpha \in V$. 由此得出 $\beta = \gamma$. ▌

**定理** 13.3.2　对于 $n$ 维复(实) 内积空间 $V$ 上的任一线性变换 $\underline{A}$，存在 $V$ 上唯一的线性变换 $\underline{A}^*$，使得

$$(\underline{A}\,\alpha,\beta) = (\alpha,\underline{A}^*\beta), \quad \forall\, \alpha,\beta \in V \tag{4}$$

**证明** 设 $\beta$ 是 $V$ 中任一向量，则 $f_\beta$ 是 $V$ 上的一个线性函数，从而 $f_\beta \underline{A}$ 也是 $V$ 上的线性函数．我们有

$$(f_\beta \underline{A})\alpha = (\underline{A}\,\alpha,\beta),\quad \forall\, \alpha \in V$$

据定理 13.3.1，存在唯一的向量 $\beta' \in V$，使得 $f_\beta \underline{A} = f_{\beta'}$，即

$$(\underline{A}\,\alpha,\beta) = (\alpha,\beta'),\quad \forall\, \alpha \in V \tag{5}$$

于是我们得到了 $V$ 到自身的一个映射 $A^*:\beta \to \beta'$．由(5)式得

$$(\underline{A}\,\alpha,\beta) = (\alpha,\underline{A}^*\beta),\quad \forall\, \alpha,\beta \in V \tag{6}$$

现在来验证 $\underline{A}^*$ 是 $V$ 上的线性变换．设 $\beta,\gamma \in V, k \in C$ 或 $R$，则对于任意 $\alpha \in V$，有

$$\begin{aligned}
&(\alpha,\underline{A}^*(k\beta+\gamma)) = (\underline{A}\,\alpha,k\beta+\gamma)\\
&=(\underline{A}\,\alpha,k\beta) + (\underline{A}\,\alpha,\gamma) = \bar{k}(\underline{A}\,\alpha,\beta) + (\underline{A}\,\alpha,\gamma)\\
&=\bar{k}(\alpha,\underline{A}^*\beta) + (\alpha,\underline{A}^*\gamma) = (\alpha,k\underline{A}^*\beta) + (\alpha,\underline{A}^*\gamma)\\
&=(\alpha,k\underline{A}^*\beta + \underline{A}^*\gamma)
\end{aligned}$$

因此 $\quad \underline{A}^*(k\beta+\gamma) = k\underline{A}^*\beta + \underline{A}^*\gamma$

从而 $\underline{A}$ 是线性变换．

再证唯一性．假设还有线性变换 $\underline{B}$ 使得

$$(\underline{A}\,\alpha,\beta) = (\alpha,\underline{B}\,\beta),\quad \forall\, \alpha,\beta \in V \tag{7}$$

则从(6)和(7)式得

$$(\alpha,\underline{A}^*\beta) = (\alpha,\underline{B}\,\beta),\quad \forall\, \alpha,\beta \in V$$

由此得出 $\quad (\alpha,(\underline{A}^* - \underline{B})\beta) = 0,\quad \forall\, \alpha,\beta \in V$

从而 $\quad (\underline{A}^* - \underline{B})\beta = 0,\quad \forall\, \beta \in V$

所以 $\underline{A}^* = \underline{B}$． ▌

**定义 1** 设 $\underline{A}$ 是复(实)内积空间 $V$ 上的一个线性变换．如果存在 $V$ 上的一个线性变换 $\underline{A}^*$，使得

$$(\underline{A}\,\alpha,\beta) = (\alpha,\underline{A}^*\beta),\quad \forall\, \alpha,\beta \in V \tag{8}$$

则称 $\underline{A}$ **有一个伴随变换**．

据定理 13.3.2，有限维复(实)内积空间 $V$ 上的每一个线性变换有唯一的一个伴随变换．在无限维复(实)内积空间的情形，线

性变换$\underset{\sim}{A}$不一定存在伴随变换. 但是当$\underset{\sim}{A}$有伴随变换时，是唯一的（见定理 13.3.2 的唯一性的证明）. 这时，我们称它是$\underset{\sim}{A}$ **的伴随变换**.

**定理** 13.3.3　设$V$是$n$维复（实）内积空间，$\varepsilon_1,\cdots\varepsilon_n$是$V$的一个标准正交基. 设$\underset{\sim}{A}$是$V$上的一个线性变换，它在基$\varepsilon_1\cdots\varepsilon_n$下的矩阵是$A=(a_{ij})$. 则$\underset{\sim}{A}$的伴随变换$\underset{\sim}{A}^*$在基$\varepsilon_1,\cdots\varepsilon_n$下的矩阵

$$B=A^*$$

这里$A^*$表示$A$的共轭转置.

**证明**　据定义 1，得

$$(\underset{\sim}{A}\,\varepsilon_i,\varepsilon_j)=(\varepsilon_i,\underset{\sim}{A}^*\varepsilon_j),\quad i,j=1,\cdots,n$$

显然，$(\underset{\sim}{A}\,\varepsilon_i,\varepsilon_j)=a_{ji}$. 设$B=(b_{ij})$，我们有

$$(\underset{\sim}{A}\,\varepsilon_i,\varepsilon_j)=(\varepsilon_i,\underset{\sim}{A}^*\varepsilon_j)=\overline{(\underset{\sim}{A}^*\varepsilon_j,\varepsilon_i)}=\overline{b}_{ij}$$

所以$\overline{b}_{ij}=a_{ji}$. 于是$b_{ij}=\overline{a}_{ji}$，即$B=A^*$　▌

**例** 1　实内积空间中，对称变换$\underset{\sim}{A}$的伴随变换$\underset{\sim}{A}^*=\underset{\sim}{A}$，因为据定义，有

$$(\underset{\sim}{A}\,\alpha,\beta)=(\alpha,\underset{\sim}{A}\,\beta),\quad\forall\,\alpha,\beta\in V$$

**例** 2　复内积空间中，Hermite 变换$\underset{\sim}{A}$的伴随变换$\underset{\sim}{A}^*=\underset{\sim}{A}$. 所以 Hermite 变换也称为自伴变换.

**例** 3　实内积空间中，正交变换$\underset{\sim}{A}$的伴随变换$\underset{\sim}{A}^*=\underset{\sim}{A}^{-1}$，这是因为

$$(\underset{\sim}{A}\,\alpha,\beta)=(\underset{\sim}{A}\,\alpha,\underset{\sim}{A}\,\underset{\sim}{A}^{-1}\beta)=(\alpha,\underset{\sim}{A}^{-1}\beta),\quad\forall\,\alpha,\beta\in V$$

**例** 4　复内积空间中，酉变换$\underset{\sim}{A}$的伴随变换$\underset{\sim}{A}^*=\underset{\sim}{A}^{-1}$.

**定理** 13.3.4　设$V$是有限维复（实）内积空间，如果$\underset{\sim}{A}$和$\underset{\sim}{B}$是$V$上的线性变换，并且$k$是复（实）数，则

(i) $(\underset{\sim}{A}+\underset{\sim}{B})^*=\underset{\sim}{A}^*+\underset{\sim}{B}^*$;

(ii) $(k\underset{\sim}{A})^*=\overline{k}\underset{\sim}{A}^*$;

(iii) $(\underset{\sim}{A}\,\underset{\sim}{B})^*=\underset{\sim}{B}^*\underset{\sim}{A}^*$;

(iv) $(\underset{\sim}{A}^*)^*=\underset{\sim}{A}$.

**证明** (i) 任取 $\alpha,\beta\in V$,有

$$((\underset{\sim}{A}+\underset{\sim}{B})\alpha,\beta)=(\underset{\sim}{A}\alpha+\underset{\sim}{B}\alpha,\beta)$$
$$=(\underset{\sim}{A}\alpha,\beta)+(\underset{\sim}{B}\alpha,\beta)=(\alpha,\underset{\sim}{A}^*\beta)+(\alpha,\underset{\sim}{B}^*\beta)$$
$$=(\alpha,\underset{\sim}{A}^*\beta+\underset{\sim}{B}^*\beta)=(\alpha,(\underset{\sim}{A}^*+\underset{\sim}{B}^*)\beta)$$

从伴随变换的唯一性得出,$(\underset{\sim}{A}+\underset{\sim}{B})^*=\underset{\sim}{A}^*+\underset{\sim}{B}^*$.

(ii) 留给读者.

(iii) $((\underset{\sim}{A}\,\underset{\sim}{B})\alpha,\beta)=(\underset{\sim}{B}\alpha,\underset{\sim}{A}^*\beta)=(\alpha,\underset{\sim}{B}^*(\underset{\sim}{A}^*\beta))$

所以 $(\underset{\sim}{A}\,\underset{\sim}{B})^*=\underset{\sim}{B}^*\underset{\sim}{A}^*$.

(iv) $(\underset{\sim}{A}^*\alpha,\beta)=\overline{(\beta,\underset{\sim}{A}^*\alpha)}=\overline{(\underset{\sim}{A}\beta,\alpha)}=(\alpha,\underset{\sim}{A}\beta)$

因此 $(\underset{\sim}{A}^*)^*=\underset{\sim}{A}$. ▌

* 从定理 13.3.4 看出,从 $\underset{\sim}{A}$ 过渡到 $\underset{\sim}{A}^*$ 有些象复数取共轭. 此外,一个复数 $z$ 可以写成 $z=a+b\mathrm{i}$,其中 $a,b$ 是实数,而实数 $a$ 具有性质 $\bar{a}=a$. 我们现在来看一个线性变换 $\underset{\sim}{A}$ 能否也写成形式

$$\underset{\sim}{A}=\underset{\sim}{A}_1+\mathrm{i}\underset{\sim}{A}_2 \tag{9}$$

其中 $\underset{\sim}{A}_i^*=\underset{\sim}{A}_i, i=1,2$. 这只要令

$$\underset{\sim}{A}_1=\frac{1}{2}(\underset{\sim}{A}+\underset{\sim}{A}^*),\qquad \underset{\sim}{A}_2=\frac{1}{2\mathrm{i}}(\underset{\sim}{A}-\underset{\sim}{A}^*) \tag{10}$$

则容易看出,$\underset{\sim}{A}=\underset{\sim}{A}_1+\mathrm{i}\underset{\sim}{A}_2$,并且 $\underset{\sim}{A}_1^*=\underset{\sim}{A}_1,\underset{\sim}{A}_2^*=\underset{\sim}{A}_2$. 这说明任一线性变换 $\underset{\sim}{A}$ 能写成(9)式,其中 $\underset{\sim}{A}_1,\underset{\sim}{A}_2$ 都是 Hermite 变换. 还可以证明 $\underset{\sim}{A}$ 的这种表法是唯一的. 留给读者作为习题.

## 习　题　13.3

1. 设 $V=M_n(C)$,具有内积

$$(A,B)=Tr(B^*A),\quad \forall\, A,B\in V$$

设 $M$ 是一个固定的 $n$ 级复矩阵. 定义

$$L_M(A)=MA,\quad \forall\, A\in M_n(C)$$

易看出 $L_M$ 是 $M_n(C)$ 上的一个线性变换. 求 $L_M^*$.

* 2. 证明：复内积空间 $V$ 中，任一线性变换 $\underline{A}$ 写成(9)式的形式是唯一的.

## §4 正规变换

这一节要讨论的问题是，对于有限维复(实)内积空间 $V$ 上的线性变换 $\underline{A}$，在什么条件下，存在 $V$ 的一个标准正交基，使得 $\underline{A}$ 在这个基下的矩阵是对角矩阵？

我们先推导 $\underline{A}$ 应当满足的必要条件. 假设存在 $V$ 的一个基 $\varepsilon_1,\cdots\varepsilon_n$，使得 $\underline{A}$ 在这个基下的矩阵 $A$ 为对角矩阵 $\mathrm{diag}\{\lambda_1,\cdots,\lambda_n\}$，则 $\underline{A}$ 的伴随变换 $\underline{A}^*$ 在这个基下的矩阵为

$$A^* = \mathrm{diag}\{\bar{\lambda}_1,\cdots,\bar{\lambda}_n\} \tag{1}$$

从(1)看出，如果 $V$ 是欧氏空间，则有 $A^* = A$，从而 $\underline{A}^* = \underline{A}$，即 $\underline{A}$ 是对称变换，显然对欧氏空间来说，这个必要条件也是充分条件. 如果 $V$ 是酉空间，则有 $AA^* = A^*A$. 于是 $\underline{A}\,\underline{A}^* = \underline{A}^*\underline{A}$.

**定义** 1 设 $V$ 是有限维复(实)内积空间，$\underline{A}$ 是 $V$ 上的线性变换，如果

$$\underline{A}^*\underline{A} = \underline{A}\,\underline{A}^* \tag{2}$$

则称 $\underline{A}$ 是**正规变换**.

容易看出，欧氏空间中的正交变换，对称变换都是正规变换. 酉空间中的酉变换，Hermite 变换也都是正规变换.

显然，如果 $\underline{A}$ 是正规变换，则 $\underline{A}^*$ 也是.

我们已经知道了，在欧氏空间 $V$ 中，存在 $V$ 的一个标准正交基，使得线性变换 $\underline{A}$ 在这个基下的矩阵为对角矩阵的充分必要条件是 $\underline{A}$ 为对称变换.

从前面的分析知道，在酉空间 $V$ 中，存在 $V$ 的一个标准正交基，使得线性变换 $\underline{A}$ 在这个基下的矩阵为对角矩阵的必要条件是 $\underline{A}$ 为正规变换. 本节的中心内容就是要证明这个必要条件也是充

分条件. 为此我们先证明两个引理和两个定理.

**引理** 13.4.1 设 $\underset{\sim}{A}$ 是有限维复(实)内积空间 $V$ 上的正规变换,则对于 $V$ 中任一向量 $\alpha$,有

$$|\underset{\sim}{A}\alpha| = |\underset{\sim}{A}^*\alpha| \tag{3}$$

**证明** 因为 $\underset{\sim}{A}\underset{\sim}{A}^* = \underset{\sim}{A}^*\underset{\sim}{A}$,所以

$$\begin{aligned}|\underset{\sim}{A}\alpha|^2 &= (\underset{\sim}{A}\alpha,\underset{\sim}{A}\alpha) = (\alpha,\underset{\sim}{A}^*(\underset{\sim}{A}\alpha)) = (\alpha,\underset{\sim}{A}(\underset{\sim}{A}^*\alpha)) \\ &= (\underset{\sim}{A}^*\alpha,\underset{\sim}{A}^*\alpha) = |\underset{\sim}{A}^*\alpha|^2\end{aligned}$$

▌

**引理** 13.4.2 设 $\underset{\sim}{A}$ 是有限维复(实)内积空间 $V$ 上的正规变换,$c$ 是任一复(实)数,则 $c\underset{\sim}{I} - \underset{\sim}{A}$ 也是正规变换.

**证明** $(c\underset{\sim}{I} - \underset{\sim}{A})^* = \bar{c}\underset{\sim}{I} - \underset{\sim}{A}^*$. 因为

$$(c\underset{\sim}{I} - \underset{\sim}{A})(\bar{c}\underset{\sim}{I} - \underset{\sim}{A}^*) = c\bar{c}\underset{\sim}{I} - c\underset{\sim}{A}^* - \bar{c}\underset{\sim}{A} + \underset{\sim}{A}\underset{\sim}{A}^*$$

$$(\bar{c}\underset{\sim}{I} - \underset{\sim}{A}^*)(c\underset{\sim}{I} - \underset{\sim}{A}) = \bar{c}c\underset{\sim}{I} - c\underset{\sim}{A}^* - \bar{c}\underset{\sim}{A} + \underset{\sim}{A}^*\underset{\sim}{A}$$

而 $\underset{\sim}{A}\underset{\sim}{A}^* = \underset{\sim}{A}^*\underset{\sim}{A}$,所以 $c\underset{\sim}{I} - \underset{\sim}{A}$ 是正规变换. ▌

**定理** 13.4.1 设 $V$ 是有限维复(实)内积空间,$\underset{\sim}{A}$ 是 $V$ 上的正规变换. 假设 $\alpha$ 是 $V$ 中的一个向量,则 $\alpha$ 是 $\underset{\sim}{A}$ 的属于特征值 $\lambda$ 的一个特征向量当且仅当 $\alpha$ 是 $\underset{\sim}{A}^*$ 的属于特征值 $\bar{\lambda}$ 的特征向量.

**证明** $\alpha$ 是 $\underset{\sim}{A}$ 的属于特征值 $\lambda$ 的一个特征向量当且仅当 $\underset{\sim}{A}\alpha = \lambda\alpha$,而后者等价于 $(\lambda\underset{\sim}{I} - \underset{\sim}{A})\alpha = 0$. 因此为了证明定理 13.4.1,就只要证

$$|(\lambda\underset{\sim}{I} - \underset{\sim}{A})\alpha| = |(\bar{\lambda}\underset{\sim}{I} - \underset{\sim}{A}^*)\alpha| \tag{4}$$

据引理 13.4.2,$\lambda\underset{\sim}{I} - \underset{\sim}{A}$ 是 $V$ 上的正规变换. 由于

$$(\lambda\underset{\sim}{I} - \underset{\sim}{A})^* = \bar{\lambda}\underset{\sim}{I} - \underset{\sim}{A}^*$$

所以据引理 13.4.1 即得出(4)式成立. 从而 $(\lambda\underset{\sim}{I} - \underset{\sim}{A})\alpha = 0$ 当且仅当 $(\bar{\lambda}\underset{\sim}{I} - \underset{\sim}{A}^*)\alpha = 0$. ▌

**定理** 13.4.2 设 $V$ 是有限维复(实)内积空间,$\underset{\sim}{A}$ 是 $V$ 上的任一线性变换. 如果 $W$ 是 $\underset{\sim}{A}$ 的不变子空间,则 $W^{\perp}$ 是 $\underset{\sim}{A}^*$ 的不变子空间.

**证明** 任取 $\beta \in W^{\perp}$,要证 $\underset{\sim}{A}^*\beta \in W^{\perp}$. 任取 $\alpha \in W$,由已知条

件得，$\underline{A}\alpha \in W$. 于是有

$$(\alpha, \underline{A}^* \beta) = (\underline{A}\alpha, \beta) = 0$$

所以 $\underline{A}^* \beta \in W^{\perp}$. ▌

现在我们来证本节的主要定理：

**定理** 13.4.3 设 $V$ 是有限维复内积空间(即酉空间)，$\underline{A}$ 是 $V$ 上的一个正规变换，则存在 $V$ 的一个标准正交基，使得 $\underline{A}$ 在这个基下的矩阵为对角矩阵.

**证明** 对酉空间的维数 $n$ 作数学归纳法.

$n=1$ 时，命题显然成立. 假设对于 $n-1$ 维酉空间，命题成立. 现在看 $n$ 维酉空间 $V$ 上的正规变换 $\underline{A}$.

由于 $V$ 是复数域上的线性空间，因此 $\underline{A}$ 必有特征值. 取 $\underline{A}$ 的一个特征值 $\lambda_1$，设 $\eta_1$ 是 $\underline{A}$ 的属于 $\lambda_1$ 的一个特征向量，并且 $|\eta_1| = 1$. 据定理 13.4.1，$\eta_1$ 也是 $\underline{A}^*$ 的属于特征值 $\overline{\lambda}_1$ 的一个特征向量. 于是 $\langle \eta_1 \rangle$ 既是 $\underline{A}$ 的不变子空间，又是 $\underline{A}^*$ 的不变子空间. 记

$$W = \langle \eta_1 \rangle^{\perp}$$

则

$$V = \langle \eta_1 \rangle \oplus W \tag{5}$$

据定理 13.4.2，$W$ 既是 $\underline{A}^*$ 的不变子空间，又是 $\underline{A}$ 的不变子空间. 从而有 $W$ 上的线性变换 $\underline{A}|W$ 和 $\underline{A}^*|W$. 由线性变换的伴随变换的定义可看出 $\underline{A}^*|W$ 是 $\underline{A}|W$ 的伴随变换. 从

$$\underline{A}\underline{A}^* = \underline{A}^*\underline{A}$$

可得出

$$(\underline{A}|W)(\underline{A}^*|W) = (\underline{A}^*|W)(\underline{A}|W)$$

因此 $\underline{A}|W$ 是 $W$ 上的正规变换. 由于 $\dim W = n-1$，因此据归纳假设，存在 $W$ 的一个标准正交基 $\eta_2, \cdots, \eta_n$，使得 $\underline{A}|W$ 在基 $\eta_2, \cdots, \eta_n$ 下的矩阵为对角矩阵 $\mathrm{diag}\{\lambda_2, \cdots, \lambda_n\}$. 由(5)式知，$\eta_1, \eta_2, \cdots, \eta_n$ 是 $V$ 的一个标准正交基，并且 $\underline{A}$ 在这个基下的矩阵为

$$\mathrm{diag}\{\lambda_1, \lambda_2, \cdots, \lambda_n\}$$

由数学归纳法原理，对一切自然数 $n$，命题成立. ▌

从定理 13.4.3 的证明中看到，这个命题成立是强烈地依赖于

复数域的特性：每个次数大于零的复系数多项式在复数域中有根．因此，对于欧氏空间$V$上的正规变换$\underline{A}$，不一定存在$V$的一个标准正交基，使得$\underline{A}$在这个基下的矩阵为对角矩阵．这是因为$\underline{A}$的特征多项式可能没有实根，从而$\underline{A}$没有特征值，因此$\underline{A}$就没有特征向量．例如，平面上绕原点$O$的旋转是一个正交变换，而正交变换一定是正规变换．显然，转角$\theta \neq k\pi$的旋转没有特征向量．因此当然不存在由特征向量组成的标准正交基．从而平面上不存在一个标准正交基，使得旋转在这个基下的矩阵为对角矩阵．

下面我们想用矩阵的语言叙述定理 13.4.3. 为此先引进正规矩阵的概念．

**定义** 2　一个$n$级复（实）矩阵$A$称为**正规的**，如果

$$AA^* = A^*A$$

正规变换与正规矩阵的关系如下：

**命题** 13.4.1　在有限维复（实）内积空间$V$中，线性变换$\underline{A}$是正规的当且仅当它在$V$的任意一个标准正交基下的矩阵是正规的．

**证明**　任取$V$的一个标准正交基$\varepsilon_1,\cdots,\varepsilon_n$．设$\underline{A}$在这个基下的矩阵为$A$，则$\underline{A}^*$在这个基下的矩阵为$A^*$．由于

$$\underline{A}\underline{A}^* = \underline{A}^*\underline{A} \Leftrightarrow AA^* = A^*A$$

所以，$\underline{A}$正规当且仅当$A$正规．　▌

**推论** 13.4.1　对于复数域上的每一个$n$级正规矩阵$A$，存在一个酉矩阵$U$，使得$U^{-1}AU$为对角矩阵．

**证明**　取一个$n$维酉空间$V$，在$V$中取一个标准正交基$\varepsilon_1,\cdots,\varepsilon_n$．令

$$\underline{A}(\varepsilon_1,\cdots,\varepsilon_n) = (\varepsilon_1,\cdots,\varepsilon_n)A$$

则$\underline{A}$是$V$上的一个线性变换，并且是正规的．据定理 13.4.3，存在$V$的一个标准正交基$\eta_1,\cdots,\eta_n$，使得$\underline{A}$在这个基下的矩阵为对角矩阵$D$．设$\varepsilon_1,\cdots,\varepsilon_n$到$\eta_1,\cdots,\eta_n$的过渡矩阵是$U$，则$U$是酉矩阵，并且$U^{-1}AU = D$．　▌

既然酉空间中，酉变换、Hermite 变换都是正规变换，因此定理 13.4.3 对于这些变换也成立. 从而酉矩阵 $A$ 一定酉相似于一个对角矩阵 $D$,即存在一个酉矩阵 $U$,使得 $U^{-1}AU = D$. 由于

$$
\begin{aligned}
DD^* &= (U^{-1}AU)(U^{-1}AU)^* = U^{-1}AUU^*A^*(U^{-1})^* \\
&= U^{-1}(U^{-1})^* = I
\end{aligned}
$$

所以 $D$ 也是酉矩阵. 设

$$D = \operatorname{diag}\{\lambda_1, \cdots, \lambda_n\}$$

则
$$DD^* = \operatorname{diag}\{\lambda_1\bar{\lambda}_1, \cdots, \lambda_n\bar{\lambda}_n\}$$

由此推出 $|\lambda_j| = 1$,即 $\lambda_j = \cos\theta_j + \mathrm{i}\sin\theta_j$,$\theta_j$ 是实数,$0 \leqslant \theta_j < 2\pi$,$j = 1, \cdots, n$. 利用欧拉(Euler) 公式:

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

可以把 $\cos\theta_j + \mathrm{i}\sin\theta_j$ 写成 $\mathrm{e}^{\mathrm{i}\theta_j}$. 这样我们证明了:

**推论** 13.4.2　每一个 $n$ 级酉矩阵一定酉相似于一个对角矩阵

$$\operatorname{diag}\{\mathrm{e}^{\mathrm{i}\theta_1}, \mathrm{e}^{\mathrm{i}\theta_2}, \cdots, \mathrm{e}^{\mathrm{i}\theta_n}\}$$

其中 $0 \leqslant \theta_j < 2\pi$, $j = 1, \cdots, n$. ▌

同理, Hermite 矩阵 $A$ 一定酉相似于一个对角矩阵 $D$,即存在酉矩阵 $U$,使得

$$U^{-1}AU = D$$

因为
$$U^{-1} = U^*, A^* = A$$

所以
$$
\begin{aligned}
D^* &= (U^{-1}AU)^* = U^*A^*(U^{-1})^* \\
&= U^*A(U^{-1})^* = U^{-1}AU = D
\end{aligned}
$$

于是 $D$ 也是 Hermite 矩阵. 设

$$D = \operatorname{diag}\{\lambda_1, \cdots, \lambda_n\}$$

则
$$D^* = \operatorname{diag}\{\bar{\lambda}_1, \cdots, \bar{\lambda}_n\}$$

从 $D^* = D$ 推出 $\lambda_1, \cdots, \lambda_n$ 都是实数. 于是我们证明了:

**推论** 13.4.3　每一个 $n$ 级 Hermite 矩阵酉相似于一个实对角矩阵,从而 Hermite 矩阵的特征值都是实数. ▌

# 习 题 13.4

1. 证明：正规变换的属于不同特征值的特征向量必正交.

2. 证明：Hermite 矩阵的特征多项式是实系数多项式.

3. 设 $V=\widetilde{C}[0,1]$（即区间$[0,1]$上的所有连续复值函数组成的线性空间），指定内积

$$(f,g)=\int_0^1 f(x)\overline{g(x)}\mathrm{d}x$$

设 $V$ 的一个变换 $\sigma: f\longmapsto \sigma f$，其中

$$(\sigma f)(x):=xf(x),\quad \forall\, x\in[0,1]$$

(1) 证明：$\sigma$ 是 $V$ 上的自伴的线性变换；

(2) 证明：$\sigma$ 没有特征值.

注：这题说明，无限维复内积空间 $V$ 上的自伴变换（即 Hermite 变换）可能没有特征值，从而没有特征向量. 因此可能不存在 $V$ 的一个标准正交基，使这个自伴变换在此基下的矩阵为对角矩阵. 由此例看出，定理13.4.3 中的条件“有限维”是必要的.

4. 设 $V$ 是有限维复（实）内积空间，$\underset{\sim}{A}$ 是 $V$ 上的一个线性变换，$\varepsilon_1,\cdots,\varepsilon_n$ 是 $V$ 的一个标准正交基. 假设 $\underset{\sim}{A}$ 在这个基下的矩阵 $A$ 是上三角矩阵. 证明：$\underset{\sim}{A}$ 是正规的当且仅当 $A$ 是对角矩阵.

5. 证明：复（实）数域上的上三角矩阵是正规的当且仅当它是对角矩阵.

6. 设 $V$ 是有限维复内积空间，证明：对于 $V$ 上的任一线性变换 $\underset{\sim}{A}$，存在 $V$ 的一个标准正交基，使得 $\underset{\sim}{A}$ 在这个基下的矩阵是上三角矩阵.

7. 证明：每一个 $n$ 级复矩阵都酉相似于一个上三角矩阵.

8. 复数域上的对称矩阵是不是自伴矩阵（即 Hermite 矩阵）？是不是正规矩阵？

9. 证明：正规变换与复数的乘积仍是正规变换.

10. 设 $C^2$ 是酉空间，指定的内积是标准内积，设 $C^2$ 上的一个线性变换 $\underset{\sim}{A}$ 在基 $\varepsilon_1=(1,0),\varepsilon_2=(0,1)$ 下的矩阵为

$$A=\begin{pmatrix}1 & \mathrm{i}\\ \mathrm{i} & 1\end{pmatrix}$$

证明：$\underset{\sim}{A}$ 是正规变换，并且求 $C^2$ 的一个标准正交基，使得 $\underset{\sim}{A}$ 在这个基下的矩阵为对角矩阵.

11. 设 $\underset{\sim}{A}$ 是有限维复内积空间 $V$ 上的正规变换，证明：

(1) $\underset{\sim}{A}$ 是自伴的当且仅当 $\underset{\sim}{A}$ 的特征值都是实数；

(2) $\underset{\sim}{A}$ 是酉变换当且仅当 $\underset{\sim}{A}$ 的特征值的模为 1.

12. 证明：线性变换 $\underset{\sim}{A}$ 是正规的充分必要条件是

$$\underset{\sim}{A} = \underset{\sim}{A}_1 + \mathrm{i}\,\underset{\sim}{A}_2$$

其中 $\underset{\sim}{A}_1, \underset{\sim}{A}_2$ 是自伴变换，并且 $\underset{\sim}{A}_1\underset{\sim}{A}_2 = \underset{\sim}{A}_2\underset{\sim}{A}_1$.

13. 证明：实对称矩阵 $A$ 有一个实对称立方根，即存在一个实对称矩阵 $B$，使 $A = B^3$.

14. 证明：正规的幂零变换一定是零变换.

15. 设 $V$ 是有限维复内积空间，$\underset{\sim}{A}, \underset{\sim}{B}$ 是 $V$ 上的正规变换，且 $\underset{\sim}{A}\underset{\sim}{B} = \underset{\sim}{B}\underset{\sim}{A}$. 证明：存在 $V$ 的一个标准正交基，使得 $\underset{\sim}{A}, \underset{\sim}{B}$ 在这个基下的矩阵都是对角矩阵.

（提示：利用习题 9.6 的第 3 题.）

16. 证明：有限维复内积空间 $V$ 上的两个正规变换如果可交换，则它们的乘积也是正规变换.

（提示：利用第 15 题结论.）

17. 证明：有限维复内积空间 $V$ 上的两个线性变换如果可交换，则存在 $V$ 的一个标准正交基，使得它们在这个基下的矩阵都是上三角矩阵.

18. 设 $A, B$ 都是 $n$ 级复矩阵，且 $B$ 是幂零矩阵，并且 $AB = BA$. 证明：

$$|A + B| = |A|$$

* 19. 设 $H$ 是 $n$ 级 Hermite 矩阵. 证明：

(1) $I - \mathrm{i}H$ 与 $I + \mathrm{i}H$ 都可逆；

(2) $A = (I - \mathrm{i}H)(I + \mathrm{i}H)^{-1}$ 是酉矩阵，并且 $-1$ 不是 $A$ 的特征值.

* 20. 设 $U$ 是 $n$ 级酉矩阵，并且 $-1$ 不是 $U$ 的特征值. 证明：$I + U$ 可逆，并且

$$H = \mathrm{i}(I - U)(I + U)^{-1}$$

是 Hermite 矩阵.

注：第 19，20 题说明：在 $n$ 级 Hermite 矩阵的全体与不以 $-1$ 为特征值的 $n$ 级酉矩阵的全体之间存在一一对应. 这个对应称为 Cayley **变换**. 它类似于复平面上的一个变换

$$w = \frac{1 - \mathrm{i}z}{1 + \mathrm{i}z}$$

21. 设 $H$ 是 Hermite 矩阵，证明：

(1) $I + \mathrm{i}H, I - \mathrm{i}H$ 都是正规矩阵；

(2) $I + \mathrm{i}H$ 与 $(I - \mathrm{i}H)^{-1}$ 可交换.

# ＊§5 Hermite 型

欧氏空间 $V$ 中装备的内积（ , ）是一个正定对称双线性函数，设它在 $V$ 的一个基 $\alpha_1, \cdots \alpha_n$ 下的度量矩阵为 $A = (a_{ij})$.

设
$$\alpha = \sum_{i=1}^{n} x_i \alpha_i, \quad \beta = \sum_{i=1}^{n} y_i \alpha_i$$

则
$$(\alpha, \beta) = \sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij} x_i y_j$$

特别地，当 $\beta = \alpha$ 时有
$$(\alpha, \alpha) = \sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij} x_i x_j \tag{1}$$

(1)式右端是 $x_1, \cdots, x_n$ 的一个二次型. 这表明，欧氏空间 $V$ 中取定一个基后，内积在 $(\alpha, \alpha)$ 上的函数值的表达式是 $\alpha$ 在这个基下的坐标 $x_1, \cdots, x_n$ 的一个二次型，这个二次型的矩阵就是内积在给定基下的度量矩阵. 因此欧氏空间 $V$ 中装备的内积与二次型有密切联系.

类似地，我们来看酉空间 $V$ 中的内积. 在 $V$ 中取一个基 $\alpha_1, \cdots \alpha_n$，构造一个矩阵

$$A = \begin{pmatrix} (\alpha_1, \alpha_1) & (\alpha_1, \alpha_2) & \cdots & (\alpha_1, \alpha_n) \\ (\alpha_2, \alpha_1) & (\alpha_2, \alpha_2) & \cdots & (\alpha_2, \alpha_n) \\ \cdots & \cdots & \cdots & \cdots \\ (\alpha_n, \alpha_1) & (\alpha_n, \alpha_2) & \cdots & (\alpha_n, \alpha_n) \end{pmatrix} \tag{2}$$

这个矩阵是由酉空间 $V$ 中装备的内积以及基 $\alpha_1, \cdots, \alpha_n$ 唯一决定

的，称 $A$ 是内积在基 $\alpha_1,\cdots,\alpha_n$ 下的**度量矩阵**. 由于 $A$ 的 $(i,j)$ 元是 $(\alpha_i,\alpha_j)$，而 $A^*$（即 $A$ 的共轭转置）的 $(i,j)$ 元是 $\overline{(\alpha_j,\alpha_i)}=(\alpha_i,\alpha_j)$，所以

$$A^*=A \tag{3}$$

这表明，酉空间 $V$ 中装备的内积在 $V$ 的任意一个基下的度量矩阵 $A$ 是 Hermite 矩阵. 设 $A=(a_{ij})$，$\alpha=\sum x_i\alpha_i$，$\beta=\sum y_i\alpha_i$，则

$$\begin{aligned}(\alpha,\beta)&=\left(\sum_i x_i\alpha_i,\sum_j y_j\alpha_j\right)\\&=\sum_i\sum_j x_i\bar{y}_j(\alpha_i,\alpha_j)=\sum_i\sum_j a_{ij}x_i\bar{y}_j\end{aligned} \tag{4}$$

特别地，当 $\beta=\alpha$ 时，有

$$(\alpha,\alpha)=\sum_i\sum_j a_{ij}x_i\bar{x}_j \tag{5}$$

**定义 1**　$n$ 个复变量 $x_1,\cdots,x_n$ 的表达式

$$f(x_1,\cdots,x_n)=\sum_i\sum_j a_{ij}x_i\bar{x}_j \tag{6}$$

其中 $a_{ji}=\bar{a}_{ij}$，称它为一个 $n$ 元 Hermite **型**. 把矩阵 $A=(a_{ij})$，称为 Hermite 型 $f(x_1,\cdots,x_n)$ 的矩阵，它是一个 Hermite 矩阵.

设 $X'=(x_1,\cdots,x_n)$，则 Hermite 型(6) 可写成

$$\sum_i\sum_j a_{ij}x_i\bar{x}_j=X^*AX \tag{7}$$

从上面的分析知道，酉空间 $V$ 中取定一个基后，$V$ 中装备的内积在 $(\alpha,\alpha)$ 上的函数值的表达式(5) 是 $\alpha$ 在这个基下的坐标 $x_1,\cdots,x_n$ 的一个 Hermite 型，该 Hermite 型的矩阵就是内积在给定基下的度量矩阵. 因此，酉空间中装备的内积与 Hermite 型有密切联系.

由于 Hermite 型(6)的矩阵是 Hermite 矩阵，因此

$$\begin{aligned}\overline{X^*AX}&=X'\overline{AX}=(X'\overline{AX})'=X^*A^*X\\&=X^*AX\end{aligned}$$

这表明 $X^*AX$ 总是实数.

从 §4 知道（见推论 13.4.3），$n$ 级 Hermite 矩阵 $A$ 酉相似于

一个实对角矩阵 $D = \mathrm{diag}\{d_1, \cdots, d_n\}$. 即存在一个酉矩阵 $U$，使得 $U^{-1}AU = D$. 令 $X = UY$，其中 $Y' = (y_1, \cdots, y_n)$，则

$$\begin{aligned} X^*AX &= Y^*U^*AUY = Y^*U^{-1}AUY = Y^*DY \\ &= d_1y_1\overline{y}_1 + d_2y_2\overline{y}_2 + \cdots + d_ny_n\overline{y}_n \end{aligned} \tag{8}$$

这证明了：

**定理** 13.5.1　对于 Hermite 型 $f(x_1, \cdots, x_n) = X^*AX$，存在酉替换 $X = UY$（即，$U$ 是酉矩阵），使得

$$f(x_1, \cdots, x_n) = d_1y_1\overline{y}_1 + \cdots + d_ny_n\overline{y}_n \tag{9}$$

其中 $d_1, \cdots, d_n$ 是 $A$ 的全部特征值，它们都是实数. ▌

**定义** 2　Hermite 型 $X^*AX$，如果对于任意 $\alpha \in C^n$ 且 $\alpha \neq 0$，有

$$\alpha^*A\alpha > 0 \tag{10}$$

则称 $X^*AX$ 是一个**正定** Hermite **型**. 一个正定 Hermite 型 $X^*AX$ 的矩阵 $A$ 称为**正定** Hermite **矩阵**.

正定 Hermite 矩阵与第六章 §5 讲的正定矩阵（即正定对称矩阵）很相象，不过前者是复数域上的矩阵，后者是实数域上的矩阵.

**定理** 13.5.2　设 $A$ 是 $n$ 级 Hermite 矩阵，则下列命题等价：

1) $A$ 是正定 Hermite 矩阵；

2) 对于任意 $n$ 级复矩阵 $P$，有 $P^*AP$ 是正定 Hermite 矩阵；

3) $A$ 的特征值全大于零；

4) 存在 $n$ 级可逆复矩阵 $P$，使 $P^*AP = I$；

5) $A$ 可以分解成 $Q^*Q$，其中 $Q$ 是 $n$ 级可逆复矩阵；

6) $A$ 的所有顺序主子式全大于零；

7) $A$ 的所有主子式全大于零.

**证明**　1) $\Rightarrow$ 2). 任取 $\alpha \in C^n$ 且 $\alpha \neq 0$，则 $P\alpha \neq 0$. 因为 $A$ 是正定 Hermite 矩阵，所以

$$\alpha^*(P^*AP)\alpha = (P\alpha)^*A(P\alpha) > 0$$

因此 $P^*AP$ 是正定 Hermite 矩阵.

2) $\Rightarrow$ 3). 由假设，$A$ 是 Hermite 矩阵. 于是存在酉矩阵 $U$，使

$$U^{-1}AU = \mathrm{diag}\{\lambda_1, \cdots, \lambda_n\} =: D$$

其中 $\lambda_i$ 是实数,$i = 1,\cdots,n$. 由假设,$U^{-1}AU$ 是正定 Hermite 矩阵,由此推出 $\varepsilon_i^* D\varepsilon_i > 0$,即 $\lambda_i > 0, i = 1,\cdots,n$.

3) ⇒ 4). 因为 $A$ 是 Hermite 矩阵,所以存在酉矩阵 $U$,使得

$$U^{-1}AU = \mathrm{diag}\{\lambda_1, \cdots, \lambda_n\}$$

其中 $\lambda_i$ 是实数,$i = 1,\cdots,n$. 由假设,$\lambda_i > 0, i = 1,\cdots,n$. 令

$$Q = \mathrm{diag}\{\sqrt{\lambda_1}, \cdots, \sqrt{\lambda_n}\}$$

则 $$U^{-1}AU = QQ$$

从而 $$Q^{-1}U^{-1}AUQ^{-1} = I$$

令 $$P = UQ^{-1}$$

则 $$P^* = (UQ^{-1})^* = Q^{-1}U^* = Q^{-1}U^{-1}$$

于是 $P^*AP = I$.

4) ⇒ 5). 由假设,$P^*AP = I$. 于是

$$A = (P^*)^{-1}P^{-1}$$

令 $Q = P^{-1}$,则

$$Q^* = (P^{-1})^* = \overline{(P^{-1})'} = (\overline{P}^{-1})' = (\overline{P}')^{-1} = (P^*)^{-1}$$

所以 $A = Q^*Q$.

5) ⇒ 1). 设 $A = Q^*Q$,其中 $Q$ 可逆. 任取 $\alpha \in C^n$ 且 $\alpha \neq 0$,有

$$\alpha^* A\alpha = \alpha^* Q^* Q\alpha$$

设 $(Q\alpha)' = (c_1, \cdots, c_n)$,则

$$\alpha^* A\alpha = c_1\bar{c}_1 + \cdots + c_n\bar{c}_n = |c_1|^2 + \cdots + |c_n|^2 > 0$$

所以 $A$ 是正定的 Hermite 矩阵.

1) ⇔ 6). 类似于第六章 §5 的定理 6.5.3 的证法.

1) ⇔ 7). 类似于习题 6.5 的第 7 题的证法. ▌

## 习 题 13.5

1. 证明:若 $A$ 是正定 Hermite 矩阵,则存在唯一的一个正定 Hermite 矩

阵 $B$,使 $A = B^2$.

2. 证明:可逆的 Hermite 矩阵的逆矩阵仍是 Hermite 矩阵.

3. 证明**极分解定理**:对于任一 $n$ 级可逆复矩阵 $A$,存在一个酉矩阵 $U$ 和两个正定 Hermite 矩阵 $S_1, S_2$,使得

$$A = S_1 U = U S_2$$

并且这种分解是唯一的.

注:当 $n = 1$ 时,一级酉矩阵形如$(e^{i\theta})$,其中 $\theta$ 为实数,一级正定 Hermite 矩阵形如$(r)$,其中 $r$ 是正实数.于是极分解定理说的是:任一非零复数 $z$ 可以分解成 $z = re^{i\theta}$,其中 $r > 0$,$\theta$ 为实数.这是复数 $z$ 在复平面的极坐标系中的表达式.这就是极分解式这个词的由来.

4. 设 $A = (a_{ij})$ 是正定 Hermite 矩阵,证明:

$$|A| \leqslant a_{11}a_{22}\cdots a_{nn}$$

等号成立当且仅当 $A$ 是对角矩阵.

(提示:类似于补充题六的第 6 题的方法).

5. 设 $A$ 是 $n$ 级可逆复矩阵,证明:

$$\| A \|^2 \leqslant \prod_{i=1}^{n} \sum_{j=1}^{n} |a_{ji}|^2$$

这里 $\| A \|$ 表示复数 $|A|$ 的模.

6. 设 $A$ 是 $n$ 级可逆复矩阵,并且 $A$ 的每个元素的模不超过 1,证明:

$$\| A \|^2 \leqslant n^n$$

# 第十四章　正交空间·辛空间·群

## §1　引　　言

在第十二章，我们对于实数域上的线性空间 $V$ 引进了内积的概念：$V$ 上的一个正定对称双线性函数称为 $V$ 上的一个内积. 在第十三章，我们又对复数域上的线性空间 $V$ 引进了内积的概念，这时为了保留正定性，需要把"对称性"换成"Hermite 性"(这样才能使 $(\alpha,\alpha)$ 为实数)，并且需要把"双线性函数"减弱成"一个半线性函数"(这样才能对于 $V$ 中一切非零向量 $\alpha$，都有 $(\alpha,\alpha)$ 为正实数). 即复线性空间 $V$ 上的一个内积是正定的 Hermite 的一个半线性函数. 引进了内积后，便可引进长度、角度、正交、距离等度量概念. 因此我们可以把内积看作线性空间上的一个度量. 依照这样的思路，为了在任意域 $F$ 上的线性空间 $V$ 中引进度量，就只需要引进内积的概念. 但是这个时候，正定性很难保留，因为在一般的域中，没有"正"元素的概念. 即使在具有正元素概念的有序域中，如何使得对于 $V$ 中一切非零向量 $\alpha$，都有 $(\alpha,\alpha)$ 是正元素，这也很困难. 所以在任意域 $F$ 上的线性空间 $V$ 中引进内积的概念，要放弃"正定性"的要求. 然而没有了正定性，就不容易引进向量的长度、两个向量的夹角，以及两个向量的距离概念. 这时我们必须坚持保留两个向量"正交"的概念. 否则，什么度量概念都没有引进到空间 $V$ 中. 两个向量 $\alpha$ 与 $\beta$ 正交的定义自然应当是 $f(\alpha,\beta)=0$，其中 $f$ 是 $V$ 上的双线性函数. 注意正交关系应当是一个对称关系，即若 $f(\alpha,\beta)=0$，应当有 $f(\beta,\alpha)=0, \forall\ \alpha,\beta\in V$. 可以证明，满足这个要求的

双线性函数或者是对称的，或者是斜对称的（见[12]p330, Theorem6.2）. 以上分析说明，为了在任意域 $F$ 上的线性空间 $V$ 中能引进正交等度量概念，就应当在 $V$ 上指定一个对称双线性函数或者斜对称双线性函数 $f$，我们可以把 $f$ 也称为 $V$ 上的一个内积，或者称 $f$ 是 $V$ 上的一个度量. 如果在 $V$ 上指定的内积是一个对称双线性函数，则 $V$ 称为**正交空间**，如果在 $V$ 上指定的内积是一个斜对称双线性函数，则 $V$ 称为**辛空间**(symplectic space). 我们感兴趣的是，$V$ 为有限维并且 $f$ 是非退化的情形，因为这种情形是很有用的.

上面一段讲的是，为了在任意域 $F$ 上的线性空间 $V$ 中引进度量，其内积应当是 $V$ 上的对称双线性函数或者斜对称双线性函数. 另一方面，即便是实数域上的线性空间 $V$，在某些问题里，也会遇到用非退化的对称双线性函数 $f$ 作为度量，而不采用第十二章讲的内积（正定对称双线性函数）. 例如，作为爱因斯坦(Einstein) 相对论的基础的"时 — 空"空间，即 Minkowski **空间** $V$，它是实数域上的四维线性空间并且指定了一个非退化的对称双线性函数作为度量. 它的一个向量 $\alpha$ 在某个基下的坐标$(t,x_1,x_2,x_3)$ 的第一个分量 $t$ 表示时间，后三个分量$(x_1,x_2,x_3)$ 表示通常空间中的位置. 对于 $\alpha,\beta\in V$，$f(\alpha-\beta,\alpha-\beta)$ 称为 $\alpha$ 与 $\beta$ 之间的时 — 空间隔的平方，它可能是正实数，零，或者是负实数，这些分别有物理解释.

这一章的前半部分的内容用于讨论正交空间和辛空间. 现在简单介绍后半部分的内容.

在一个线性空间 $V$ 中定义了一个度量（即内积）后，自然要讨论保持度量不变的线性变换，即使得

$$(\underset{\sim}{A}\alpha,\underset{\sim}{A}\beta)=(\alpha,\beta),\quad \forall\ \alpha,\beta\in V$$

成立的线性变换 $\underset{\sim}{A}$. 譬如，在实内积空间中的正交变换，在复内积空间中的酉变换. 这类线性变换有如下性质：它们的乘积仍是这类变换，逆变换仍是这类变换. 于是这类线性变换组成的集合对于映射的乘法形成一个**群**（群的概念在 §4 中介绍）. 若有限维内积空

间 $V$ 中的内积是对称双线性函数，即 $V$ 是正交空间，则把保持内积的线性变换组成的群称为**正交群**(欧氏空间中的正交变换组成的群也是正交群)；辛空间中保持内积不变的线性变换组成的群称为**辛群**；酉空间中保持内积不变的线性变换组成的群称为**酉群**.

1872 年，F. Klein 提出了著名的"Erlangen program"(Erlangen 纲领)，指出：几何就是研究空间中的图形在某个变换群的作用下不变的性质. 于是研究欧几里得空间中的图形在正交群作用下不变的性质，这种几何称为**欧几里得几何**. 研究酉空间中的图形在酉群的作用下不变的性质，这种几何称为**酉几何**. 研究正交空间中图形在正交群作用下不变的性质，这种几何称为**正交几何**. 研究辛空间中图形在辛群作用下不变的性质，这种几何称为**辛几何**.

## * §2 正 交 空 间

**定义 1** 域 $F$ 上的线性空间 $V$ 如果指定了一个对称双线性函数 $f$，则称 $V$ 是一个**正交空间**，称 $f$ 是 $V$ 上的一个**度量**(或**内积**). 用$(V,f)$表示指定的度量为 $f$ 的正交空间. 如果 $f$ 是非退化的，则 $(V,f)$ 称为**正则的**，否则称为**非正则的**.

**例 1** $R^4$ 中，对于

$$\alpha=(x_1,x_2,x_3,x_4),\quad \beta=(y_1,y_2,y_3,y_4)$$

定义
$$f(\alpha,\beta):=x_1y_1-x_2y_2-x_3y_3-x_4y_4. \tag{1}$$

显然，$f$ 是 $R^4$ 上的一个非退化的对称双线性函数，因此 $f$ 给出了 $R^4$ 上的一个度量. $R^4$ 对于这个度量成为一个正则的正交空间. 在这个空间里，有

$$f(\alpha,\alpha)=x_1^2-x_2^2-x_3^2-x_4^2 \tag{2}$$

于是，若 $\alpha=(2,1,0,0)$，则 $f(\alpha,\alpha)=3$. 若 $\alpha=(1,1,0,0)$，则 $f(\alpha,\alpha)=0$. 若 $\alpha=(0,1,0,0)$，则 $f(\alpha,\alpha)=-1$.

**定义 2** 在正交空间 $(V,f)$ 中，若 $f(\alpha,\beta)=0$，则称 $\alpha$ 与 $\beta$ **正**

**交**,记作 $\alpha \perp \beta$.

由于 $f$ 是对称的双线性函数,于是从 $f(\alpha,\beta)=0$ 可以推出 $f(\beta,\alpha)=0$. 所以若 $\alpha$ 与 $\beta$ 正交,则 $\beta$ 也与 $\alpha$ 正交.

从例 1 中看到,正交空间中的一个非零向量有可能与自身正交.

**定义 3** 在正交空间 $(V,f)$ 中,一个非零向量 $\alpha$ 称为**迷向的**(isotropic),如果 $f(\alpha,\alpha)=0$;否则,$\alpha$ 称为**非迷向的**(anisotropic). 正交空间 $(V,f)$ 称为**迷向的**,如果它包含了一个(非零的)迷向向量;否则称为**非迷向的**. $(V,f)$ 称为**全迷向的**,如果 $V$ 中的所有非零向量都是迷向的.

例如,例 1 中的正交空间 $(R^4,f)$ 是迷向的,但不是全迷向的.

**命题** 14.2.1 如果正交空间 $(V,f)$ 是非迷向的,则它一定是正则的,亦即 $f$ 一定是非退化的.

**证明** 假如 $f$ 是退化的,则 $\mathrm{rad}V\neq 0$. 于是存在 $\alpha\in\mathrm{rad}V$ 且 $\alpha\neq 0$. 从而 $f(\alpha,\alpha)=0$,矛盾. ▌

注意命题 14.2.1 的逆命题不成立. 例如,例 1 中的 $(R^4,f)$ 是正则的,但是它是迷向的.

**命题** 14.2.2 设 $\mathrm{char}F\neq 2$. 若正交空间 $(V,f)$ 是全迷向的,则 $f=0$.

**证明** 任取 $\alpha,\beta\in V$. 由于 $V$ 全迷向,所以

$$\begin{aligned}0&=f(\alpha+\beta,\alpha+\beta)\\&=f(\alpha,\alpha)+2f(\alpha,\beta)+f(\beta,\beta)\\&=2f(\alpha,\beta)\end{aligned}$$

因为 $\mathrm{char}F\neq 2$,所以从上式得出 $f(\alpha,\beta)=0$. 于是 $f=0$. ▌

设 $(V,f)$ 是一个正交空间,$W$ 是 $V$ 的一个线性子空间. 显然,把 $f$ 限制到 $W$ 上就成为 $W$ 上的一个对称双线性函数,从而 $(W,f|W)$ 也是一个正交空间. 称 $(W,f|W)$ 是正交空间 $(V,f)$ 的一个子空间. 值得注意的是,即使 $(V,f)$ 是正则的,但是它的子空间 $(W,f|W)$ 有可能是非正则的. 例如,在例 1 的 $(R^4,f)$ 中,设

$\alpha=(1,1,0,0)$，令 $W=\langle\alpha\rangle$. 则

$$f(\alpha,k\alpha)=kf(\alpha,\alpha)=0,\quad \forall\ k\in R$$

于是 $\alpha\in \mathrm{rad}W$. 这表明 $f|W$ 是退化的. 从而 $(W,f|W)$ 是非正则的. 进一步看到，$\forall\ k\in R$，有

$$f(k\alpha,k\alpha)=k^2f(\alpha,\alpha)=0$$

于是 $(W,f|W)$ 是全迷向的.

**定义** 4　设 $S$ 是正交空间 $(V,f)$ 的一个子集，下述集合

$$\{\alpha\in V\mid f(\alpha,\beta)=0,\forall\ \beta\in S\}$$

称为 $S$ 的**正交补**，记作 $S^{\perp}$.

容易看出，$S^{\perp}$ 是 $V$ 的线性子空间.

**定理** 14.2.1　设 $(V,f)$ 是有限维正则的正交空间，$W$ 是 $V$ 的一个子空间. 则

(i)　$\dim W+\dim W^{\perp}=\dim V$；　(3)

(ii)　$(W^{\perp})^{\perp}=W$.　(4)

**证明**　这实际上就是习题 11.4 的第 6 题，已经给出了证明的思路，这里不再重复写出.　▌

**定义** 5　有限维正交空间 $(V,f)$ 的一个基 $\alpha_1,\cdots,\alpha_n$ 称为**正交基**，如果 $\alpha_1,\cdots,\alpha_n$ 两两正交，即 $f(\alpha_i,\alpha_j)=0$，对一切 $i\neq j$.

注意正交基的定义中，首先要求是基，然后要求基向量两两正交. 这是因为在正交空间中两两正交的向量可能是线性相关的. 例如，在例 1 的 $(R^4,f)$ 中，设 $\alpha=(1,1,0,0)$，则 $\alpha$ 与 $2\alpha$ 是正交的，然而它们是线性相关的.

**定理** 14.2.2　特征不为 2 的域 $F$ 上的有限维的正交空间 $(V,f)$ 一定存在正交基.

**证明**　据定理 11.4.2，存在 $V$ 的一个基 $\alpha_1,\cdots,\alpha_n$，使得 $f$ 在此基下的度量矩阵 $A$ 是对角矩阵 $\mathrm{diag}\{d_1,\cdots,d_n\}$. 这表明 $\alpha_1,\cdots,\alpha_n$ 两两正交，因此它是 $V$ 的一个正交基.　▌

注意当 $(V,f)$ 是正则的时候，上述证明中的 $A$ 是满秩矩阵，从而 $d_i\neq 0,i=1,\cdots,n$. 于是 $f(\alpha_i,\alpha_i)\neq 0$，这表明 $\alpha_i$ 是非迷向的，

$i=1,\cdots,n$. 这也就是说，当$(V,f)$正则时，存在由非迷向向量组成的正交基.

**定义** 6　有限维正交空间$(V,f)$的一个正交基$\varepsilon_1,\cdots,\varepsilon_n$，称为**标准正交基(或正交规范基)**，如果

$$f(\varepsilon_i,\varepsilon_i)=0\text{ 或 }\pm 1,\quad i=1,\cdots,n$$

当$F=R$或$C$时，有限维正交空间$(V,f)$存在标准正交基，并且当$F=C$时，总可以使标准正交基中的每个向量$\varepsilon_i$，满足

$$f(\varepsilon_i,\varepsilon_i)=0\text{ 或 }1$$

对于有限维的正则的正交空间$(V,f)$，它的每个向量$\beta$在$V$的正交基$\alpha_1,\cdots,\alpha_n$下的坐标特别简单好记，即我们有

$$\beta=\sum_{i=1}^{n}\frac{f(\beta,\alpha_i)}{f(\alpha_i,\alpha_i)}\alpha_i \tag{5}$$

请读者证明公式(5).

**定理** 14.2.3　设$\operatorname{char}F\neq 2$，设$(V,f)$是正交空间，如果$W$是$V$的有限维正则子空间，则

$$V=W\oplus W^{\perp} \tag{6}$$

**证明**　任取$\alpha\in W\cap W^{\perp}$. 因为$W$是正则的，所以$\operatorname{rad}W=0$，从而$\alpha=0$. 因此$W\cap W^{\perp}=0$.

再证$V=W+W^{\perp}$. 在$W$中取一个由非迷向向量组成的正交基$\alpha_1,\cdots,\alpha_m$，对于$V$中任一向量$\beta$，令

$$\beta_1=\sum_{i=1}^{m}\frac{f(\beta,\alpha_i)}{f(\alpha_i,\alpha_i)}\alpha_i\in W,\quad \beta_2=\beta-\beta_1$$

因为

$$\begin{aligned}f(\beta_2,\alpha_j)&=f(\beta,\alpha_j)-\sum_{i=1}^{m}\frac{f(\beta,\alpha_i)}{f(\alpha_i,\alpha_i)}f(\alpha_i,\alpha_j)\\&=f(\beta,\alpha_j)-f(\beta,\alpha_j)=0,\quad j=1,\cdots,m\end{aligned}$$

所以$\beta_2\in W^{\perp}$，于是

$$\beta=\beta_1+\beta_2\in W+W^{\perp}$$

所以$V=W\oplus W^{\perp}$.　∎

**定义** 7　设$W_1$和$W_2$是正交空间$(v,f)$的两个子空间，如果

$$f(\beta_1,\beta_2)=0, \quad \forall\ \beta_1\in W_1, \quad \beta_2\in W_2$$

则称子空间 $W_1$ 与 $W_2$ 是**正交的**.

例如,$(V,f)$ 的任一子空间 $W$ 与 $W^{\perp}$ 是正交的.

定理14.2.2的一个等价说法是,特征不为2的域上的有限维正交空间$(V,f)$一定能分解成两两正交的一维子空间的直和(简称为一维子空间的**正交直和**).

**定义8** 设 $(V_1,f_1)$ 和$(V_2,f_2)$ 是域 $F$ 上的两个正交空间,如果存在线性空间 $V_1$ 到 $V_2$ 的一个同构映射 $\sigma$,并且 $\sigma$ 保持内积,即

$$f_2(\sigma\alpha,\sigma\beta)=f_1(\alpha,\beta), \quad \forall\ \alpha,\beta\in V_1$$

则称正交空间 $(V_1,f_1)$ 和$(V_2,f_2)$ **是同构的**,或者称 $V_1$ 与 $V_2$ 是保距同构的(isometric),称 $\sigma$ 是正交空间$(V_1,f_1)$ 到$(V_2,f_2)$ 的一个**同构映射**.

容易看出,恒等映射是同构映射,同构映射的乘积还是同构映射,同构映射的逆映射还是同构映射. 从而域 $F$ 上正交空间之间的同构关系具有反身性、传递性和对称性.

设 $\mathrm{char}F\neq 2$,设$\sigma$是有限维正交空间$(V_1,f_1)$到$(V_2,f_2)$的一个同构映射. 由于 $\sigma$ 也是线性空间 $V_1$ 到 $V_2$ 的同构映射,所以 $\sigma$ 把 $V_1$ 的一个基映成 $V_2$ 的一个基. 又由于 $\sigma$ 保持内积,所以 $\sigma$ 把 $V_1$ 的一个正交基 $\alpha_1,\cdots,\alpha_n$ 映成 $V_2$ 的一个正交基 $\sigma(\alpha_1),\cdots,\sigma(\alpha_n)$;并且

$$f_2(\sigma(\alpha_i),\sigma(\alpha_i))=f_1(\alpha_i,\alpha_i), \quad i=1,\cdots,n$$

域 $F$ 上两个有限维正交空间$(V_1,f_1)$与$(V_2,f_2)$如果同构,则它们的维数相同. 反之,维数相同的正交空间不一定同构. 那么应当再加上什么条件才能同构?这强烈地依赖于域$F$的性质. 下面我们仅讨论实数域 $R$ 和复数域 $C$ 的情形.

对于实数域上的有限维正交空间 $(V,f)$,一定存在标准正交基 $\varepsilon_1,\cdots,\varepsilon_n$. 我们把使得 $f(\varepsilon_i,\varepsilon_i)=0,1,-1$ 的 $\varepsilon_i$ 的数目分别记作 $r_0,r_+,r_-$,则

$$r_0+r_++r_-=n$$

$f$ 在基 $\varepsilon_1,\cdots,\varepsilon_n$ 下的度量矩阵 $A$ 是对角矩阵,主对角线上有

$r_0$ 个 0,$r_+$ 个 1,$r_-$ 个 $-1$. 适当安排基向量的次序,可以设

$$A = \mathrm{diag}\{1,\cdots,1,-1,\cdots,-1,0,\cdots,0\}$$

于是 $$r_+ + r_- = \mathrm{rank}A = \mathrm{rank}f$$

从而 $$r_0 = n - \mathrm{rank}f$$

由于双线性函数 $f$ 在 $V$ 的不同基下的度量矩阵是合同的,而两个 $n$ 级实对称矩阵合同的充分必要条件是它们有相同的秩和相同的正惯性指数(见第六章 §3 的推论 6.3.3),因此 $r_+,r_-$ 不依赖于标准正交基的选择,它们分别是 $f$ 对应的二次型(见第十一章 §5)的正惯性指数,负惯性指数. 而 $r_+ - r_-$ 是二次型的符号差. 我们把 $r_+,r_-,r_+-r_-$ 分别称为 $f$ 的正惯性指数、负惯性指数、符号差. 根据这些分析,我们可以得到

**定理** 14.2.4 实数域上两个有限维正交空间 $(V_1,f_1)$ 与 $(V_2,f_2)$ 同构的充分必要条件是 $V_1$ 与 $V_2$ 有相同的维数,并且 $f_1$ 与 $f_2$ 有相同的秩和相同的符号差.

**证明** 必要性. 在 $V_1$ 中取一个标准正交基 $\varepsilon_1,\cdots,\varepsilon_n$. 设 $f_1(\varepsilon_i,\varepsilon_i)=0,1,-1$ 的 $\varepsilon_i$ 的数目分别为 $r_0,r_+,r_-$. 设 $\sigma$ 是 $(V_1,f_1)$ 到 $(V_2,f_2)$ 的一个同构映射,则 $\sigma\varepsilon_1,\cdots,\sigma\varepsilon_n$ 是 $V_2$ 的一个标准正交基,并且使得 $f_2(\sigma\varepsilon_i,\sigma\varepsilon_i)=0,1,-1$ 的 $\sigma\varepsilon_i$ 的数目也分别等于 $r_0$, $r_+,r_-$. 由于 $r_+ + r_-$ 是 $f_1$ 的秩,也是 $f_2$ 的秩,所以 $f_1$ 与 $f_2$ 有相同的秩. $r_+ - r_-$ 是 $f_1,f_2$ 的符号差.

充分性. 在 $V_1,V_2$ 中分别取一个标准正交基

$$\varepsilon_1,\cdots,\varepsilon_n;\quad \varepsilon_1',\cdots,\varepsilon_n'$$

相应的有 $$r_0,r_+,r_-;\quad r_0',r_+',r_-'$$

由假设得 $$r_+ + r_- = r_+' + r_-',\quad r_+ - r_- = r_+' - r_-'$$

由此得出 $$r_+ = r_+',\quad r_- = r_-',\quad r_0 = r_0'$$

适当安排 $V_1,V_2$ 的基向量的次序,可使前 $r_+$ 个基向量与自己的内积值为 1,中间 $r_-$ 个的内积值为 $-1$,最后 $r_0$ 个的内积值为 0. 令

$$\sigma: V_1 \longrightarrow V_2$$

$$\alpha = \sum_{i=1}^{n} x_i \varepsilon_i \longmapsto \sigma(\alpha) = \sum_{i=1}^{n} x_i \varepsilon_i'$$

则 $\sigma$ 是线性空间 $V_1$ 到 $V_2$ 的一个同构映射. 设 $f_1, f_2$ 在基 $\{\varepsilon_i\}$, 基 $\{\varepsilon_i'\}$ 下的度量矩阵分别为 $A_1, A_2$. 则由上述知

$$A_1 = A_2 = \mathrm{diag}\{1, \cdots, 1, -1, \cdots, -1, 0, \cdots, 0\}$$

于是对于 $V_1$ 中任意向量

$$\alpha = (\varepsilon_i, \cdots, \varepsilon_n)X, \quad \beta = (\varepsilon_1, \cdots, \varepsilon_n)Y$$

有
$$f_1(\alpha, \beta) = X' A_1 Y = X' A_2 Y = f_2(\sigma\alpha, \sigma\beta)$$

因此 $\sigma$ 是正交空间 $(V_1, f_1)$ 到 $(V_2, f_2)$ 的同构映射. ▍

从定理 14.2.4 看出，同一个 $n$ 维实线性空间 $V$, 装备上不同的对称双线性函数 $f_1, f_2$, 得到的两个正交空间 $(V, f_1)$ 与 $(V, f_2)$ 可能是不同构的. 请读者思考：为什么同一个 $n$ 维实线性空间 $V$, 装备上不同的内积(正定对称双线性函数), 得到的欧氏空间却一定是同构的呢? 能不能也从定理 14.2.4 看出来呢?

类似地可以证明：

**定理** 14.2.5　复数域上两个有限维正交空间 $(V_1, f_1)$ 与 $(V_2, f_2)$ 同构的充分必要条件是 $V_1$ 与 $V_2$ 有相同的维数, 并且 $f_1$ 与 $f_2$ 有相同的秩.

**证明**　注意对于复数域上的 $n$ 维正交空间 $(V, f)$, 存在一个标准正交基 $\varepsilon_1, \cdots, \varepsilon_n$, 并且 $f(\varepsilon_i, \varepsilon_i) = 0$ 或 $1, i = 1, \cdots, n$. 另外注意利用 $n$ 级复对称矩阵合同的充分必要条件是它们有相同的秩. 本定理的证明思路与定理 14.2.4 相同. ▍

**定义** 9　设 $(V, f)$ 是域 $F$ 上有限维正则的正交空间, $V$ 上的一个线性变换 $\underline{T}$ 如果保持内积不变, 即

$$f(\underline{T}\alpha, \underline{T}\beta) = f(\alpha, \beta), \quad \forall\, \alpha, \beta \in V \tag{7}$$

则称 $\underline{T}$ 是 $V$ 上的一个**正交变换**.

在第十一章 §5, 我们曾指出, 给了特征不为 2 的域 $F$ 上的线性空间 $V$ 上的一个对称双线性函数 $f$, 就有唯一的一个二次函数 $q$ 与之对应, 其中

$$q(\alpha) := f(\alpha,\alpha), \quad \forall\, \alpha \in V \tag{8}$$

反之亦然. 在 $V$ 中取定一个基 $\alpha_1,\cdots,\alpha_n$ 后，二次函数 $q$ 在 $\alpha$ 上的函数值 $q(\alpha)$ 的表达式是一个二次型 $\sum\limits_i \sum\limits_j a_{ij}x_ix_j$，其中 $A=(a_{ij})$ 是 $f$ 在基 $\alpha_1,\cdots,\alpha_n$ 下的度量矩阵，$(x_1,\cdots,x_n)$ 是 $\alpha$ 在基 $\alpha_1,\cdots,\alpha_n$ 下的坐标. 显然 $V$ 中取定一个基后，二次函数 $q$ 与二次型是一一对应的. 把二次型也记作 $q$.

**命题** 14.2.3 设 $(V,f)$ 是特征不为 2 的域 $F$ 上的有限维正则的正交空间，则 $V$ 上的线性变换 $\underline{T}$ 是正交变换的充分必要条件是，$\underline{T}$ 保持 $f$ 对应的二次型 $q$ 不变，即

$$q(\underline{T}\alpha) = q(\alpha), \quad \forall\, \alpha \in V \tag{9}$$

**证明** 必要性是显然的. 关于充分性. 据第十一章 §5 的公式 (6)，对一切 $\alpha,\beta \in V$，有

$$f(\alpha,\beta) = \frac{1}{2}[q(\alpha+\beta) - q(\alpha) - q(\beta)]$$

由此推出，$f(\underline{T}\alpha,\underline{T}\beta) = f(\alpha,\beta)$. ▍

在 $V$ 中取一个基 $\alpha_1,\cdots,\alpha_n$，设 $f$ 在这个基下的度量矩阵是 $A$，线性变换 $\underline{T}$ 在这个基下的矩阵是 $T$，$\alpha=(\alpha_1,\cdots,\alpha_n)X$，则 $\underline{T}$ 是正交变换的充分必要条件是

$$X'(T'AT)X = X'AX, \quad \forall\, X \in F^n \tag{10}$$

注意：把(9)式具体写出来就是(10)式.

**命题** 14.2.4 设 $(V,f)$ 是域 $F$ 上的有限维正则的正交空间，则 $\underline{T}$ 是 $V$ 上的正交变换的充分必要条件是，$\underline{T}$ 是正交空间 $V$ 到自身的同构映射.

**证明** 充分性是显然的. 关于必要性只要证 $\underline{T}$ 是双射就够了. 设 $\alpha \in \mathrm{Ker}\underline{T}$，则 $\underline{T}\alpha = 0$. 从而对一切 $\beta \in V$，有

$$f(\alpha,\beta) = f(\underline{T}\alpha,\underline{T}\beta) = 0$$

由于 $f$ 是非退化的，因此 $\alpha = 0$. 这证明了 $\underline{T}$ 是单射. 由于 $V$ 是有限维的，所以 $\underline{T}$ 也是满射. ▍

从同构映射的性质得出，恒等变换是正交变换；正交变换的乘积还是正交变换；正交变换的逆变换是正交变换.

**定理** 14.2.6　设 $(V,f)$ 是特征不为 2 的域 $F$ 上的有限维正则的正交空间. 任取 $V$ 的一个基 $\alpha_1,\cdots,\alpha_n$，设 $f$ 在这个基下的度量矩阵是 $A$. 设线性变换 $\underline{T}$ 在这个基下的矩阵是 $T$，则 $\underline{T}$ 是正交变换的充分必要条件是

$$T'AT = A \tag{11}$$

**证明**　据(10)式知，$\underline{T}$ 是正交变换当且仅当

$$X'(T'AT)X = X'AX, \quad \forall\, X \in F^n \tag{10}$$

因为 $A$ 是对称矩阵，所以 $T'AT$ 也是对称矩阵. 于是由(10) 式得出，$T'AT = A$. ▌

**推论** 14.2.1　条件同定理 14.2.6. 正交变换 $\underline{T}$ 在 $V$ 的任意一个基下的矩阵 $T$ 的行列式等于 $+1$ 或者 $-1$.

**证明**　由(11)式得，$|T|^2|A| = |A|$. 因为 $f$ 非退化，所以 $A$ 满秩. 从而得出 $|T|^2 = 1$. 于是 $|T| = 1$ 或 $-1$. 由于线性变换在不同基下的矩阵是相似的，而相似的矩阵有相同的行列式. 所以命题成立. ▌

行列式为 $+1$ 的正交变换称为**第一类的**，或**旋转**；行列式为 $-1$ 的正交变换称为**第二类的**.

## 习　题　14.2

1. $R^2$ 中，对于 $\alpha' = (x_1,x_2)$，$\beta' = (y_1,y_2)$，定义

$$f(\alpha,\beta) = x_1y_1 - x_2y_2$$

(1) 证明：$(R^2,f)$ 是一个正则的正交空间；

(2) 设 $\varepsilon_1' = (1,0)$，$\varepsilon_2' = (0,1)$，证明：$\varepsilon_1,\varepsilon_2$ 是 $(R^2,f)$ 的一个标准正交基；

(3) 求 $f$ 在基 $\varepsilon_1,\varepsilon_2$ 下的度量矩阵；

(4) 设

$$T=\begin{pmatrix}\sqrt{2} & 1\\ 1 & \sqrt{2}\end{pmatrix}$$

令 $\underline{T}$ 是 $R^2$ 上的一个线性变换，它在基 $\varepsilon_1,\varepsilon_2$ 下的矩阵为 $T$. 证明：$\underline{T}$ 是正交空间 $(R^2,f)$ 上的一个正交变换，并且它的特征值是 $\sqrt{2}\pm 1$；

(5) 求 $\underline{T}$ 的全部特征向量，并且说明 $\underline{T}$ 的特征向量都是迷向的.

2. 详细写出定理 14.2.5 的证明.

3. 设 $(V,f)$ 是特征不为 2 的域 $F$ 上的有限维正则的正交空间，任意取定 $V$ 中的一个非迷向向量 $\eta$，定义

$$\tau_\eta(\alpha):=\alpha-\frac{2f(\alpha,\eta)}{f(\eta,\eta)}\eta,\quad \forall\,\alpha\in V$$

证明：$\tau_\eta$ 是 $V$ 上的一个第二类的正交变换. 称 $\tau_\eta$ 是关于超平面 $\langle\eta\rangle^\perp$ 的镜面反射.

4. 设 $(V,f)$ 是域 $F$ 上有限维正则的正交空间. $\underline{T}$ 是 $V$ 上的一个正交变换，证明：如果 $V$ 的子空间 $W$ 是 $\underline{T}$ 的不变子空间，则 $W^\perp$ 也是 $\underline{T}$ 的不变子空间.

5. 设 $\mathrm{char}F\neq 2$，一个 2 维正交空间 $(V,f)$ 如果是正则的而且是迷向的，则称它为一个**双曲平面**(hyperbolic plane). 证明：2 维正交空间 $(V,f)$ 是双曲平面的充分必要条件是，存在 $V$ 的一个基，使得 $f$ 在这个基下的度量矩阵为

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

## *§3 辛 空 间

本节域 $F$ 的特征不为 2. 不再每次声明.

**定义 1** 域 $F$ 上的线性空间 $V$ 如果指定了一个斜对称双线性函数 $f$，则称 $V$ 是一个**辛空间**，用 $(V,f)$ 表示. 称 $f$ 是 $V$ 上的一个**内积**(或**辛内积**). 如果 $f$ 是非退化，则称 $(V,f)$ 是**正则的**；否则称为**非正则的**.

**例1** $R^2$ 中,对于 $\alpha=(x_1,x_2),\beta=(y_1,y_2)$,定义

$$f(\alpha,\beta):=x_1y_2-x_2y_1 \tag{1}$$

显然,$f$ 是 $R^2$ 上的一个非退化斜对称双线性函数,所以$(R^2,f)$ 成为一个辛空间.

与正交空间一样,辛空间中有向量的正交、子空间的正交、子集的正交补、迷向向量、迷向子空间概念.显然,辛空间中每个向量都是迷向的,且任一子集的正交补是子空间.

从第十一章§4的公式(15)知道,有限维的正则的辛空间一定是偶数维的.

从定理11.4.3知道,$n$ 维辛空间$(V,f)$ 中,存在 $V$ 的一个基 $\varepsilon_1,\varepsilon_{-1},\cdots,\varepsilon_r,\varepsilon_{-r},\eta_1,\cdots,\eta_s$,使得 $f$ 在这个基下的度量矩阵 $A$ 为

$$\mathrm{diag}\left\{\begin{pmatrix}0&1\\-1&0\end{pmatrix},\cdots,\begin{pmatrix}0&1\\-1&0\end{pmatrix},0,\cdots,0\right\}$$

即

$$\begin{aligned}&f(\varepsilon_i,\varepsilon_{-i})=1,\quad i=1,\cdots,r\\&f(\varepsilon_i,\varepsilon_j)=0,\quad i+j\neq 0\\&f(\varepsilon_i,\eta_k)=0,\quad i=\pm 1,\cdots,\pm r\\&f(\eta_j,\eta_k)=0,\quad j,k=1,\cdots,s\end{aligned} \tag{2}$$

这个基称为$(V,f)$的**辛基**.

我们把基 $\varepsilon_1,\cdots,\varepsilon_r,\varepsilon_{-1},\cdots,\varepsilon_{-r},\eta_1,\cdots,\eta_s$ 也称为辛基.容易看出,$f$ 在此基下的度量矩阵为

$$\begin{pmatrix}0&I_r&0\\-I_r&0&0\\0&0&0\end{pmatrix}$$

**定理14.3.1** 设$(V,f)$是$n$维辛空间,如果$W$是$V$的正则子空间,则

$$V=W\oplus W^{\perp} \tag{3}$$

**证明** 任取$\alpha\in W\cap W^{\perp}$,因为$W$是正则的,所以 $\mathrm{rad}W=0$,从而 $\alpha=0$.因此 $W\cap W^{\perp}=0$.

在 $W$ 中取一个辛基 $\varepsilon_1,\varepsilon_{-1},\cdots,\varepsilon_m,\varepsilon_{-m}$.把它扩充成 $V$ 的一个

基：$\varepsilon_1,\varepsilon_{-1},\cdots,\varepsilon_m,\varepsilon_{-m},\beta_1,\cdots,\beta_t$，令

$$\beta_j' = \beta_j - \sum_{i=1}^{m} f(\beta_j,\varepsilon_{-i})\varepsilon_i + \sum_{i=1}^{m} f(\beta_j,\varepsilon_i)\varepsilon_{-i}, \quad j = 1,\cdots,t$$

则
$$f(\beta_j',\varepsilon_k) = f(\beta_j,\varepsilon_k) - \sum_{i=1}^{m} f(\beta_j,\varepsilon_{-i})f(\varepsilon_i,\varepsilon_k)$$

$$+ \sum_{i=1}^{m} f(\beta_j,\varepsilon_i)f(\varepsilon_{-i},\varepsilon_k)$$

$$= f(\beta_j,\varepsilon_k) + f(\beta_j,\varepsilon_k)f(\varepsilon_{-k},\varepsilon_k) = 0$$

$$f(\beta_j',\varepsilon_{-k}) = f(\beta_j,\varepsilon_{-k}) - f(\beta_j,\varepsilon_{-k})f(\varepsilon_k,\varepsilon_{-k}) = 0$$

$k = 1,\cdots,m$. 显然 $\varepsilon_1,\varepsilon_{-1},\cdots,\varepsilon_m,\varepsilon_{-m},\beta_1',\cdots,\beta_t'$ 仍是 $V$ 的一个基.

令
$$U = \langle \beta_1',\cdots,\beta_t' \rangle$$

则 $V = W \oplus U$，由上面推导出的结果知道 $\beta_j' \in W^{\perp}$，$j = 1,\cdots,t$. 所以 $U \subset W^{\perp}$. 另一方面，任取 $\alpha \in W^{\perp}$，设

$$\alpha = \alpha_1 + \alpha_2, \quad \alpha_1 \in W, \quad \alpha_2 \in U$$

任取 $\beta \in W$，有

$$0 = f(\alpha,\beta) = f(\alpha_1,\beta) + f(\alpha_2,\beta) = f(\alpha_1,\beta)$$

于是 $\alpha_1 \in W^{\perp}$. 又 $\alpha_1 \in W$，所以 $\alpha_1 = 0$. 从而 $\alpha \in U$. 因此 $U = W^{\perp}$. 于是有 $V = W \oplus W^{\perp}$. ▌

**定理** 14.3.2 有限维辛空间 $(V,f)$ 一定能分解成一维非正则子空间与二维正则子空间的正交直和.

**证明** 由定理 11.4.3 即得. ▌

与正交空间一样，辛空间也有同构的概念. 设 $(V_1,f_1)$ 与 $(V_2,f_2)$ 是两个有限维的辛空间，如果它们同构，则存在一个同构映射 $\sigma$，把 $V_1$ 的辛基映成 $V_2$ 的辛基，从而 $\dim V_1 = \dim V_2$.

**定理** 14.3.3 域 $F$ 上两个有限维辛空间 $(V_1,f_1)$ 与 $(V_2,f_2)$ 同构的充分必要条件是，$V_1$ 与 $V_2$ 有相同的维数，并且 $f_1$ 与 $f_2$ 有相同的矩阵秩.

**证明** 必要性. 在 $V_1$ 中取一个辛基

$$\varepsilon_1,\varepsilon_{-1},\cdots,\varepsilon_r,\varepsilon_{-r},\eta_1,\cdots,\eta_s$$

则它们在同构映射 $\sigma$ 下的象是 $V_2$ 的一个辛基. 由此得出

$$\dim V_1 = \dim V_2$$

并且 $f_1$ 与 $f_2$ 有相同的矩阵秩.

充分性. 由假设, $V_1$ 和 $V_2$ 中分别有辛基:

$$\varepsilon_1,\varepsilon_{-1},\cdots,\varepsilon_r,\varepsilon_{-r},\eta_1,\cdots,\eta_s \quad 和 \quad \varepsilon_1',\varepsilon_{-1}',\cdots,\varepsilon_r',\varepsilon_{-r}',\eta_1',\eta_s'$$

于是 $f_1,f_2$ 分别在这两个基下的度量矩阵 $A_1,A_2$ 都等于

$$\operatorname{diag}\left\{\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix},\cdots,\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix},0,\cdots,0\right\}$$

令 $\sigma$ 是 $V_1$ 到 $V_2$ 的映射. 使得 $\alpha$ 与 $\sigma(\alpha)$ 分别在上述辛基下的坐标是相同的. 于是对于 $V_1$ 中任意向量

$$\alpha = (\varepsilon_1,\varepsilon_{-1},\cdots,\eta_s)X, \quad \beta = (\varepsilon_1,\varepsilon_{-1},\cdots,\eta_s)Y$$

有
$$f_1(\alpha,\beta) = X'A_1Y = X'A_2Y = f_2(\sigma\alpha,\sigma\beta)$$

因此 $\sigma$ 是辛空间 $(V_1,f_1)$ 到 $(V_2,f_2)$ 的同构映射. ▌

**推论** 14.3.1 域 $F$ 上两个有限维的正则辛空间同构的充分必要条件是它们有相同的维数. ▌

辛空间的同构具有反身性、对称性和传递性.

**定义** 2 设 $(V,f)$ 是有限维正则辛空间, $V$ 上的一个线性变换 $\underline{B}$ 如果保持内积不变, 则称为**辛变换**.

**命题** 14.3.1 设 $(V,f)$ 是有限维正则辛空间, 则 $\underline{B}$ 是 $V$ 上辛变换的充分必要条件是, $\underline{B}$ 是辛空间 $V$ 到自身的同构映射.

**证明** 与命题 14.2.4 的证法一样. ▌

从同构映射的性质得出, 恒等变换是辛变换; 辛变换的逆变换是辛变换; 辛变换的乘积是辛变换.

设 $\underline{B}$ 是有限维正则辛空间 $(V,f)$ 上的一个辛变换, 在 $V$ 中取一个辛基 $\varepsilon_1,\cdots,\varepsilon_r,\varepsilon_{-1},\cdots,\varepsilon_{-r}$, 则 $f$ 在这个基下的矩阵 $A$ 为

$$\begin{pmatrix} 0 & I_r \\ -I_r & 0 \end{pmatrix}$$

设 $\underline{B}$ 在这个基下的矩阵为 $B$. 任取 $\alpha,\beta \in V$, 它们在此基下的坐标列向量分别为 $X,Y$. 则

$$f(\alpha,\beta)=X'AY$$

$$f(\underline{B}\,\alpha,\underline{B}\,\beta)=(BX)'A(BY)=X'B'ABY$$

由于 $\underline{B}$ 是辛变换,它保持内积,因此得到

$$X'(B'AB)Y=X'AY,\quad \forall\ X,Y\in F^{2r} \tag{4}$$

从(4)式得出,$B'AB=A$.

反之,如果 $\underline{B}$ 在上述辛基下的矩阵 $B$ 满足

$$B'AB=A \tag{5}$$

则(4)式成立,从而

$$f(\alpha,\beta)=f(\underline{B}\,\alpha,\underline{B}\beta),\quad \forall\ \alpha,\beta\in V$$

因此 $\underline{B}$ 是辛变换.于是我们证明了

**定理** 14.3.4 设 $(V,f)$ 是有限维正则辛空间,则 $V$ 上的线性变换 $\underline{B}$ 是辛变换的充分必要条件是 $\underline{B}$ 在 $V$ 的辛基

$$\varepsilon_1,\cdots,\varepsilon_r,\varepsilon_{-1},\cdots,\varepsilon_{-r}$$

下的矩阵 $B$ 满足

$$B'AB=A \tag{5}$$

其中
$$A=\begin{pmatrix}0 & I_r\\ -I_r & 0\end{pmatrix}$$ ▌

**推论** 14.3.2 辛变换的行列式为 1 或 $-1$.

**证明** 因为 $B'AB=A$,所以 $|B|^2|A|=|A|$.因为 $|A|=1$,所以 $|B|^2=1$.由此得 $|B|=\pm 1$. ▌

进一步可以证明 辛变换的行列式等于 1(见阅读材料十二).

**定义** 3 设 $A=\begin{pmatrix}0 & I_r\\ -I_r & 0\end{pmatrix}$,满足 $B'AB=A$ 的矩阵 $B$ 称为**辛矩阵**.

## 习 题 14.3

1. 设 $(V,f)$ 是域 $F$ 上有限维正则辛空间,证明:

(1) $V$ 上辛变换把 $V$ 的辛基变成辛基；

(2) 如果 $V$ 上线性变换 $\underline{B}$ 把 $V$ 的辛基变成辛基，则 $\underline{B}$ 是辛变换.

2. 设 $A = \begin{pmatrix} 0 & I_r \\ -I_r & 0 \end{pmatrix}$，证明：$B$ 是辛矩阵的充分必要条件是

$$B = -A(B^{-1})'A$$

3. 设 $B$ 是 $2r$ 级矩阵，把 $B$ 分块写成

$$B = \begin{pmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{pmatrix}$$

其中 $B_{ij}$ 是 $r$ 级矩阵，证明：$B$ 是辛矩阵的充分必要条件是

$$\begin{cases} B_{11}'B_{21} = B_{21}'B_{11} \\ B_{12}'B_{22} = B_{22}'B_{12} \\ B_{11}'B_{22} - B_{21}'B_{12} = I_r \end{cases}$$

4. 证明下列矩阵都是辛矩阵：

(1) $\mathrm{diag}\left\{\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \cdots, \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}\right\}$，$r$ 是偶数；

(2) $\begin{pmatrix} 0 & I_r \\ -I_r & 0 \end{pmatrix}$；  (3) $\begin{pmatrix} 0 & -I_r \\ I_r & 0 \end{pmatrix}$；  (4) $\begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & -1 & 0 \\ 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}$.

5. 设 $g(\lambda)$ 是 $2r$ 级辛矩阵 $B$ 的特征多项式. 证明：$g(\lambda) = \lambda^{2r}g(\lambda^{-1})$.

6. 设 $B$ 是实数域上的辛矩阵，$\lambda_1$ 是 $B$ 的特征多项式的一个复根，证明：$\lambda_1^{-1}, \bar{\lambda}_1, \bar{\lambda}_1^{-1}$ 都是 $B$ 的特征多项式的复根.

# 阅读材料十二

**定理 1** 设 $F$ 是特征不为 2 的域，$F_1$ 是 $F$ 的最小子域，$x_{11}, \cdots, x_{1n}, x_{21}, \cdots, x_{2n}, x_{n1}, \cdots, x_{nn}$ 是 $n^2$ 个无关不定元，其中 $n = 2r$，则

(1) 存在唯一的 $F_1$ 上的 $r(2r-1)$ 元多项式

$$pf(x_{12}, \cdots, x_{1n}, x_{23}, \cdots, x_{2n}, \cdots, x_{n-1,n})$$

使得对于域 $F$ 上任一 $2r$ 级斜对称矩阵 $A = (a_{ij})$，有

$$pf^2(A) = detA$$

其中 $pf(A)=pf(a_{12},\cdots,a_{1n},\cdots,a_{n-1,n})$，并且满足

$$pf\begin{pmatrix} 0 & I_r \\ -I_r & 0 \end{pmatrix}=1$$

(2) 多项式 $pf$ 具有下述性质：对于域 $F$ 上任一 $2r$ 级斜对称矩阵 $A$，任一 $2r$ 级矩阵 $B$，有

$$pf(B'AB)=(detB)pf(A)$$

**证明** (1) 用 $E$ 表示以 $F_1$ 为系数域的 $n^2$ 元有理函数域. 由于 $\mathrm{char}F_1\neq 2$，所以 $\mathrm{char}E\neq 2$，令

$$G=\begin{pmatrix} 0 & x_{12} & x_{13} & \cdots & x_{1n} \\ -x_{12} & 0 & x_{23} & \cdots & x_{2n} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ -x_{1n} & -x_{2n} & -x_{3n} & \cdots & 0 \end{pmatrix}$$

$G$ 是域 $E$ 上的一个斜对称矩阵. 据习题 11.4 第 7 题得，$detG$ 是 $E$ 中某个元素的平方. 又从行列式的定义容易看出，$detG$ 是域 $F_1$ 上 $x_{12},\cdots,x_{1n},x_{23},\cdots,x_{2n},\cdots,x_{n-1,n}$ 的一个多项式. 因此

$$detG=f^2(x_{12},\cdots,x_{1n},x_{23},\cdots,x_{2n},\cdots,x_{n-1,n}) \tag{1}$$

其中 $f(x_{12},\cdots,x_{n-1,n})\in E$. 设

$$f(x_{12},\cdots,x_{n-1,n})=\frac{g(x_{12},\cdots,x_{n-1,n})}{h(x_{12},\cdots,x_{n-1,n})}$$

其中，$g(x_{12},\cdots,x_{n-1,n}),h(x_{12},\cdots,x_{n-1,n})$ 是域 $F_1$ 上 $x_{12},\cdots,x_{n-1,n}$ 的多项式，并且 $(g,h)=1$. 从(1)得

$$h^2detG=g^2$$

从而 $h|g^2$. 由于 $(h,g)=1$，所以从 $h|g^2$ 推出 $h|g$. 由此得出 $h\in F_1$. 因此

$$f(x_{12},\cdots,x_{n-1,n})\in F_1[x_{12},\cdots,x_{n-1,n}]$$

从(1)知道，$f$ 或 $-f$ 的平方都等于 $detG$. 我们把 $f$ 与 $-f$ 都叫做 $detG$ 的平方根. 取其中一个平方根记作 $pf(x_{12},\cdots,x_{n-1,n})$，它满足

$$pf\begin{pmatrix} 0 & I_r \\ -I_r & 0 \end{pmatrix}=1$$

于是多项式 $pf(x_{12},\cdots,x_{n-1,n})$ 是唯一确定的.

任取域 $F$ 上一个 $2r$ 级斜对称矩阵 $A=(a_{ij})$，不定元 $x_{12},\cdots,x_{n-1,n}$ 用 $F_1$ 的交换扩环 $F$ 中的元素 $a_{12},\cdots,a_{n-1,n}$ 代入，由于

$$pf^2(x_{12},\cdots,x_{n-1,n})=detG \tag{2}$$

因此 $$pf^2(a_{12},\cdots,a_{n-1,n}) = detA \tag{3}$$

把 $pf(a_{12},\cdots,a_{n-1,n})$ 记作 $pf(A)$，则(3) 式成为

$$pf^2(A) = detA \tag{4}$$

(2) 任取环 $F_1[x_{11},\cdots,x_{1n},\cdots,x_{n1},\cdots,x_{nn}]$ 上的一个 $2r$ 级矩阵 $S=(s_{ij}(x_{11},\cdots,x_{nn}))$，则 $S'GS$ 是此环上的一个斜对称矩阵. 设

$$S'GS = \begin{pmatrix} 0 & h_{12}(x_{11},\cdots,x_{nn}) & \cdots & h_{1n}(x_{11},\cdots,x_{nn}) \\ -h_{12}(x_{11},\cdots,x_{nn}) & 0 & \cdots & h_{2n}(x_{11},\cdots,x_{nn}) \\ \cdots & \cdots & \cdots & \cdots \\ -h_{1n}(x_{11},\cdots,x_{nn}) & -h_{2n}(x_{11},\cdots,x_{nn}) & \cdots & 0 \end{pmatrix}$$

不定元 $x_{12},\cdots,x_{n-1,n}$ 用 $F_1$ 的交换扩环 $F_1[x_{11},\cdots,x_{nn}]$ 中元素

$$h_{12}(x_{11},\cdots,x_{nn}),\cdots,h_{n-1,n}(x_{11},\cdots,x_{nn})$$

代入，从(2)式得到

$$pf^2(S'GS) = det(S'GS) \tag{5}$$

从而 $$pf^2(S'GS) = (detS)^2 detG = (detS)^2 pf^2(G) \tag{6}$$

因此 $$pf(S'GS) = \pm (detS)pf(G) \tag{7}$$

由于 $pf(S'GS)$ 与 $(detS)pf(G)$ 都是 $F_1[x_{11},\cdots,x_{nn}]$ 中的多项式，因此从(7) 式得出

$$pf(S'GS) = (detS)pf(G) \tag{8}$$

或者 $$pf(S'GS) = -(detS)pf(G) \tag{9}$$

两者只居其一.

设

$$A = (a_{ij}) = \begin{pmatrix} 0 & I_r \\ -I_r & 0 \end{pmatrix}, \qquad I_{2r} = (e_{ij}) = \begin{pmatrix} I_r & 0 \\ 0 & I_r \end{pmatrix}$$

取多项式 $s_{ij}(x_{11},\cdots,x_{nn})$，使得

$$s_{ij}(a_{11},\cdots,a_{nn}) = e_{ij}, \quad 1 \leqslant i,j \leqslant n$$

假如

$$pf(S'GS) = -(detS)pf(G) \tag{9}$$

不定元 $x_{11},\cdots,x_{nn}$ 用 $a_{11},\cdots,a_{nn}$ 代入，从(9) 式得出

$$pf(I_{2r}'AI_{2r}) = -(detI_{2r})pf(A) \tag{10}$$

即 $$pf\begin{pmatrix} 0 & I_r \\ -I_r & 0 \end{pmatrix} = -pf\begin{pmatrix} 0 & I_r \\ -I_r & 0 \end{pmatrix}$$

由此得出，$1=-1$，矛盾(因为 $\mathrm{char}F_1\neq 2$). 所以

$$pf(S'GS)=(detS)pf(G) \tag{8}$$

现在设 $A=(a_{ij})$ 是域 $F$ 上任一 $2r$ 级斜对称矩阵，$B=(b_{ij})$ 是域 $F$ 上任一 $2r$ 级矩阵. 多项式 $s_{ij}(x_{11},\cdots,x_{nn})$，使得

$$s_{ij}(a_{11},\cdots,a_{nn})=b_{ij},\quad 1\leqslant i,j\leqslant n$$

不定元 $x_{11},\cdots,x_{nn}$ 用 $a_{11},\cdots,a_{nn}$ 代入，从(8)式得

$$pf(B'AB)=(detB)pf(A) \tag{11}$$

定理的证明已完成. ▌

**推论 1** 设 $F$ 是特征不为 2 的域，则域 $F$ 上辛矩阵 $B$ 的行列式为 1.

**证明** 设

$$A=\begin{pmatrix}0 & I_r\\ -I_r & 0\end{pmatrix}$$

由辛矩阵的定义知，$B'AB=A$. 从定理 1 的(2)得出

$$pf(B'AB)=(detB)pf(A)$$

从而 $pf(A)=(detB)pf(A)$. 由于 $pf(A)=1$，因此，$detB=1$. ▌

从推论 1 得出，辛变换的行列式等于 1.

## §4 群的定义和例·子群·同构

从第十二章到上一节，我们讨论了具有一个度量(即内积)的线性空间中，保持这个度量不变的线性变换. 譬如，欧氏空间中的正交变换，酉空间中的酉变换，正交空间中的正交变换，辛空间中的辛变换. 它们都有如下的性质，以酉空间中的酉变换为例：恒等变换是酉变换，酉变换的逆变换是酉变换，酉变换的乘积是酉变换. 但是酉变换的和不一定是酉变换，酉变换与复数的数量乘积不一定是酉变换(譬如，$\underline{I}$ 与 $-\underline{I}$ 都是酉变换，但是 $\underline{I}+(-\underline{I})=\underline{0}$ 不是酉变换；$2\underline{I}$ 不是酉变换). 这说明，$n$ 级酉空间 $V$ 中的所有酉变换组成的集合 $U(V)$ 中具有一种运算：映射的乘法；它满足结合律；恒等变换属于 $U(V)$；若 $\underline{A}\in U(V)$，则 $\underline{A}^{-1}\in U(V)$. 类似地，$n$ 维

欧氏空间 $V$ 中的所有正交变换组成的集合 $O(V)$，域 $F$ 上的 $n$ 维正则的正交空间 $(V,f)$ 中所有正交变换组成的集合 $O(V,f)$，特征不为 2 的域 $F$ 上 $2r$ 维正则辛空间 $(V,f)$ 中所有辛变换组成的集合 $Sp(V,f)$，也都具有象 $U(V)$ 那样的性质. 从这些例子我们抽象出群的概念：

**定义** 1　设 $G$ 是一个非空集合，如果在 $G$ 上定义了一种代数运算，叫做**乘法**，即给出了 $G\times G$ 到 $G$ 的一个映射：$(a,b)\longmapsto d$，把 $(a,b)$ 在此映射下的象 $d$ 称为 $a$ 与 $b$ 的积，记作 $d=ab$；并且这种代数运算适合下列规则，那么 $G$ 称为一个**群**：

(i) $a(bc)=(ab)c,\quad \forall\, a,b,c\in G$(结合律)；　(1)

(ii) $G$ 中有一个元素 $e$，使得

$$ea=ae=a,\quad \forall\, a\in G \tag{2}$$

(iii) 对于 $G$ 中每一个元素 $a$，都有 $G$ 中一个元素 $b$，使得

$$ab=ba=e \tag{3}$$

**命题** 14.4.1　群 $G$ 中具有性质(2)的元素 $e$ 是唯一的，称 $e$ 是 $G$ 的**单位元素**.

**证明**　设 $e$ 和 $e'$ 都具有性质(2)，则

$$e'=ee'=e$$ ▌

**命题** 14.4.2　对于群 $G$ 中元素 $a$，$G$ 中具有性质(3)的元素 $b$ 是唯一的，称 $b$ 是 $a$ 的**逆元素**，记作 $a^{-1}$.

**证明**　设 $G$ 中元素 $b$ 和 $b'$ 都具有性质(3)，则

$$b'=eb'=(ba)b'=b(ab')=be=b$$ ▌

如果群 $G$ 的运算还满足交换律，即

$$ab=ba,\quad \forall\, a,b\in G$$

则称 $G$ 是**交换群**，或 Abel **群**.

有时，把 Abel 群 $G$ 的运算叫做**加法**，即把 $(a,b)$ 在这个映射下的象记成 $a+b$，称为 $a$ 与 $b$ 的**和**. 此时定义 1 中的条件(i)，(ii)，(iii) 写成

(i)′ $a+(b+c)=(a+b)+c,\quad \forall\, a,b,c\in G$；

(ii)′ $G$ 中有一个元素 0，使得

$$0+a=a+0=a,\quad \forall\, a\in G$$

(iii)′ 对于 $G$ 中每一个元素 $a$，都有 $G$ 中一个元素 $b$，使得

$$a+b=b+a=0$$

这时，把 0 称为 $G$ 的**零元素**，把 $b$ 称为 $a$ 的**负元素**，记作 $-a$.

由于群 $G$ 的运算满足结合律，因此 $G$ 中任意 $m$ 个元素 $a_1,\cdots,a_m$ 的积与括号的添法无关，可以简单地写成 $a_1a_2\cdots a_m$. 从而在群 $G$ 中可以定义元素的**方幂**：对于任意正整数 $m$，

$$a^m:=\underbrace{a\ a\ \cdots\ a}_{m个}$$

$$a^0:=e$$

$$a^{-m}:=(a^{-1})^m$$

容易验证，对于任意 $a,b\in G$，任意整数 $n,m$，有

$$a^na^m=a^{n+m},\qquad (a^n)^m=a^{nm}$$

一般地，$(ab)^m\neq a^mb^m$. 但是对于 abel 群 $G$，有

$$(ab)^m=a^mb^m,\quad \forall\, a,b\in G,\quad \forall\, m\in Z$$

如果 $G$ 是无限集合，则群 $G$ 称为**无限群**；如果 $G$ 是有限集合，则群 $G$ 称为**有限群**，此时，$G$ 中元素的数目称为群 $G$ 的**阶**，记作 $|G|$.

**例** 1　$O(V)$，$U(V)$，$O(V,f)$，$Sp(V,f)$ 都是群.

**例** 2　域 $F$ 上 $n$ 维线性空间 $V$ 上的所有可逆线性变换组成的集合 $GL(V)$，对于映射的乘法成为一个群.

**例** 3　域 $F$ 上 $n$ 级可逆矩阵的全体，对于矩阵的乘法成为一个群，记作 $GL(n,F)$，称它为域 $F$ 上的**一般线性群**.

**例** 4　域 $F$ 上行列式为 1 的 $n$ 级矩阵组成的集合，对于矩阵的乘法成为一个群，记作 $SL(n,F)$，称它为域 $F$ 上的**特殊线性群**.

**例** 5　实数域上 $n$ 级正交矩阵的全体，对于矩阵的乘法成为一个群，记作 $O(n)$，称它为实数域上的**正交群**.

**例** 6　$n$ 级酉矩阵的全体，对于矩阵的乘法成为一个群，记作

$U(n)$，称它为**酉群**.

* **例** 7　设 $\mathrm{char}F \neq 2$，取定 $F$ 上一个 $n$ 级可逆对称矩阵 $A$，令

$$O(A,F) = \{T \in GL(n,F) \mid T'AT = A\}$$

则 $O(A,F)$ 对于矩阵的乘法成为一个群，称它为域 $F$ 上的一个正交群.

* **例** 8　在例 7 中，当 $F$ 为实数域，取 $A$ 为

$$\mathrm{diag}\{1,\cdots,1,-1,\cdots,-1\}$$

其中 1 和 $-1$ 的数目分别为 $p,q$，并且 $q \neq 0$，这时 $O(A,R)$ 也称为伪正交群，记成 $O(p,q)$. 当 $q = 0$ 时，$O(A,R)$ 就是实数域上的正交群 $O(n)$.

* **例** 9　特征不为 2 的域 $F$ 上 $2r$ 级辛矩阵的全体，对于矩阵的乘法成为一个群，称它为辛群，记作 $Sp(2r,F)$.

* 域 $F$ 上的一般线性群 $GL(n,F)$；特征不为 2 的域 $F$ 上的正交群 $O(A,F)$，辛群 $Sp(2r,F)$；以及复数域上的酉群 $U(n)$，都称为**典型群**.

**例** 10　所有 $n$ 元置换组成的集合 $S_n$，对于映射的乘法成为一个群，称 $S_n$ 为 $n$ **元对称群**.

**例** 11　我们知道，正方形比长方形更加对称. 为什么呢？为了回答这个问题，我们设平面上正方形 $E_1$ 与长方形 $E_2$ 的两条对角线的交点都是 $O$ 点. 考虑使正方形 $E_1$ 与其自身重合的那些平面正交点变换，共有八个：恒等变换 $e$、绕 $O$ 点转角为 $90°$ 的旋转 $T$、绕 $O$ 点转角为 $180°$ 的旋转 $T^2$、绕 $O$ 点转角为 $270°$ 的旋转 $T^3$、关于过 $O$ 点与正方形的一边平行的直线 $l_1$ 的反射 $\sigma_1$、关于过 $O$ 点与正方形另一边平行的直线 $l_2$ 的反射 $\sigma_2$、关于正方形的一条对角线的反射 $\sigma_3$、关于正方形的另一条对角线的反射 $\sigma_4$. 可以验证集合

$$\{e,T,T^2,T^3,\sigma_1,\sigma_2,\sigma_3,\sigma_4\}$$

对于变换的乘法成为一个群 $G_1$，称 $G_1$ 为**正方形 $E_1$ 的对称群**. 类似地，考虑使长方形 $E_2$ 与其自身重合的那些正交点变换，共有四个：恒等变换 $e$、绕 $O$ 点转角为 $180°$ 的旋转、关于过 $O$ 点与长方形的一

边平行的直线的反射、关于过 $O$ 点与长方形的另一边平行的直线的反射.这四个变换组成的集合对于变换的乘法成为一个群 $G_2$,称 $G_2$ 为**长方形 $E_2$ 的对称群**.由此看出,正方形比长方形更对称,是因为正方形的对称群比长方形的对称群"大".事实上,群是认识现实世界最深刻的规律性之一——对称性的有力武器.

**定义** 2 如果群 $G$ 的非空子集 $H$ 对于 $G$ 的运算也成一个群,那么 $H$ 称为 $G$ 的**子群**.记作 $H < G$.

显然,$\{e\}$,$G$ 都是 $G$ 的子群,它们称为 $G$ 的平凡子群.

设 $\mathrm{char}F \neq 2$,$O(A,F)$ 是 $GL(n,F)$ 的子群;$Sp(2r,F)$ 是 $GL(2r,F)$ 的子群.$O(p,q)$ 是 $GL(p+q,R)$ 的子群.$U(n)$ 是 $GL(n,C)$ 的子群.对任意域 $F$,$SL(n,F)$ 是 $GL(n,F)$ 的子群.

**定理** 14.4.1 群 $G$ 的非空子集 $H$ 是一个子群的充分必要条件是,由 $a,b \in H$ 可推出 $ab^{-1} \in H$.

**证明** 必要性是显然的.现在证明充分性.因为 $H$ 非空,所以 $H$ 含有一个元素 $a$,于是 $aa^{-1} \in H$.由此得,$e \in H$.显然 $e$ 也是 $H$ 的单位元素.

任取 $b \in H$,由 $e,b \in H$ 得到 $eb^{-1} \in H$,即 $b^{-1} \in H$.

任取 $c,b \in H$,由上述知,$b^{-1} \in H$.由 $c,b^{-1} \in H$ 得出 $c(b^{-1})^{-1} \in H$,即 $cb \in H$.这说明群 $G$ 的乘法运算限制到 $H$ 上,是 $H$ 的运算.结合律是显然的.综上述得,$H$ 是 $G$ 的子群. ▌

为了把群进行分类,看哪些群在本质上是一样的,需要引进群的同构的概念.

**定义** 3 设 $G$ 和 $G'$ 是两个群,如果存在 $G$ 到 $G'$ 的一个双射 $\sigma$,使得

$$\sigma(ab) = \sigma(a)\sigma(b), \quad \forall\, a,b \in G \tag{4}$$

则称 $G$ **同构于** $G'$,记作 $G \cong G'$;映射 $\sigma$ 称为 $G$ 到 $G'$ 的一个**同构映称**(或简称**同构**).

**命题** 14.4.3 如果 $\sigma$ 是群 $G$ 到 $G'$ 的一个同构映射,则

$$\sigma(e) = e'; \quad \sigma(a^{-1}) = \sigma(a)^{-1}, \quad \forall\, a \in G$$

**证明** 任取 $a' \in G'$，因为 $\sigma$ 是双射，所以存在 $a \in G$，使得 $\sigma(a) = a'$. 从而

$$\sigma(e)a' = \sigma(e)\sigma(a) = \sigma(ea) = \sigma(a) = a'$$

同理有 $a'\sigma(e) = a'$. 因此 $\sigma(e) = e'$.

因为 $\qquad \sigma(a)\sigma(a^{-1}) = \sigma(aa^{-1}) = \sigma(e) = e'$

并且 $\qquad \sigma(a^{-1})\sigma(a) = e'$

所以 $\sigma(a^{-1})$ 是 $\sigma(a)$ 的逆元. ▌

不难证明，群的同构作为群之间的一种关系也具有反身性、对称性和传递性.

**例** 12 在域 $F$ 上的 $n$ 维线性空间 $V$ 中取定一个基后，$V$ 上可逆线性变换与它在给定基下的矩阵（可逆矩阵）的对应 $\sigma$ 是 $GL(V)$ 到 $GL(n,F)$ 的双射，且保持乘法运算，因此 $\sigma$ 是 $GL(V)$ 到 $GL(n,F)$ 的一个同构映射. 从而 $GL(V) \cong GL(n,F)$.

**例** 13 $n$ 维欧氏空间 $V$ 中取定一个标准正交基后，$V$ 上正交变换与它在给定的标准正交基下的矩阵（正交矩阵）的对应是 $O(V)$ 到 $O(n)$ 的一个双射，而且保持乘法运算，因此

$$O(V) \cong O(n)$$

**例** 14 $n$ 维酉空间 $V$ 中取定一个标准正交基后，$V$ 上酉变换与它在给定标准正交基下的矩阵（酉矩阵）的对应是 $U(V)$ 到 $U(n)$ 的一个同构映射，因此 $U(V) \cong U(n)$.

* **例** 15 在特征不为 2 的域 $F$ 上的 $n$ 维正则的正交空间 $(V,f)$ 中，取定一个正交基后，设 $f$ 在这个正交基下的度量矩阵为 $A$，则正交变换与它在给定正交基下的矩阵的对应是 $O(V,f)$ 到 $O(A,F)$ 的一个同构映射，因此 $O(V,f) \cong O(A,F)$.

* **例** 16 特征不为 2 的域 $F$ 上的 $2r$ 维正则辛空间 $(V,f)$ 中，取定一个辛基后，辛变换与它在给定辛基下的矩阵（辛矩阵）的对应是 $Sp(V,f)$ 到 $Sp(2r,F)$ 的一个同构映射，因此

$$Sp(V,f) \cong Sp(2r,F)$$

设 $\Omega$ 是任一非空集合，$\Omega$ 到自身的所有双射组成的集合

$S(\Omega)$，对于映射的乘法成为一个群，称它为 $\Omega$ 上的**全变换群**. $S(\Omega)$ 的任一子群称为 $\Omega$ 上的**变换群**. 当 $\Omega$ 为 $n$ 个元素的有限集合时，$\Omega$ 上的全变换群也就是 $n$ 元对称群 $S_n$. $S_n$ 的子群称为**置换群**.

历史上，对群的研究最早是从置换群和变换群开始的. 1771 年，Lagrange 自发地采用置换群以解决用根式解代数方程问题，1799 年 Ruffin，1824 年 Abel 继续这一工作，直到 1830 年，Galois 自觉地应用群的思想（群的术语就是他首先引进的）彻底解决了这个问题，证明了一般的五次和五次以上的方程不能用根式解. 与此独立，19 世纪中叶，出现了多种“几何”，需要弄清楚到底什么叫做几何？如何对各种几何进行分类. 1872 年 Klein 提出了著名的 Erlangen 纲领，用变换群来对几何学分类. 他指出，几何就是研究空间中的图形在某个变换群的作用下不变的性质. 到 19 世纪末叶，人们意识到，在数学的不同领域中独立存在的群论思想，在原则上是统一的. 这种想法引起了研究抽象群的概念. Kelly，Frobenius，Dyck 等最早从事抽象群的研究. Schmidt 于 1916 年出版了《抽象群论》的书. 于是群论成为代数学的一个重要分支.

下面的定理说明了抽象群与变换群的密切关系.

* **定理** 14.4.2（Cayley **定理**） 任何一个群都同构于一个变换群.

**证明** 设 $G$ 是一个群. 对于 $G$ 中每个元素 $a$，定义集合 $G$ 上的一个变换 $\sigma_a$ 如下：

$$\sigma_a(x) := ax, \quad \forall\, x \in G \tag{5}$$

令 $G_L = \{\sigma_a \mid a \in G\}$.

首先证明 $\sigma_a$ 是 $G$ 到自身的双射，为此只要证 $\sigma_a$ 是集合 $G$ 的可逆变换. 因为对一切 $x \in G$，有

$$(\sigma_{a^{-1}}\sigma_a)(x) = \sigma_{a^{-1}}(ax) = a^{-1}(ax) = x$$

$$(\sigma_a\sigma_{a^{-1}})(x) = \sigma_a(a^{-1}x) = a(a^{-1}x) = x$$

所以 $\sigma_{a^{-1}}\sigma_a$ 与 $\sigma_a\sigma_{a^{-1}}$ 都等于 $G$ 的恒等变换. 从而 $\sigma_a$ 是 $G$ 的可逆变换，并且 $\sigma_a^{-1} = \sigma_{a^{-1}}$.

其次证明 $G_L$ 是集合 $G$ 上的变换群，即要证 $G_L$ 是 $S(G)$ 的子群. 上面已证 $G_L$ 是 $S(G)$ 的子集，显然 $G_L$ 非空. 任取 $\sigma_a,\sigma_b \in G_L$，因为

$$(\sigma_a\sigma_b)x = \sigma_a(bx) = (ab)x = \sigma_{ab}(x), \quad \forall\, x \in G$$

所以

$$\sigma_a\sigma_b = \sigma_{ab}, \quad \forall\, a,b \in G \tag{6}$$

于是

$$\sigma_a(\sigma_b)^{-1} = \sigma_a\sigma_{b^{-1}} = \sigma_{ab^{-1}} \in G_L$$

这证明了 $G_L$ 是 $S(G)$ 的子群.

最后证明 $G \cong G_L$. 令

$$\begin{aligned} \psi: G &\longrightarrow G_L \\ a &\longmapsto \sigma_a \end{aligned} \tag{7}$$

显然 $\psi$ 是 $G$ 到 $G_L$ 的满射. 如果 $\sigma_a = \sigma_b$，则

$$a = \sigma_a(e) = \sigma_b(e) = b$$

因此 $\psi$ 是单射. 又从(6)式知，$\psi$ 保持运算. 所以 $\psi$ 是群 $G$ 到 $G_L$ 的一个同构映射. 从而 $G \cong G_L$. ▌

* $\sigma_a$ 就是把 $G$ 中每个元素用 $a$ 左乘，称 $\sigma_a$ 是由元素 $a$ 引起的**左平移**. 抽象群 $G$ 与变换群 $G_L$ 同构，意味着抽象群 $G$ 获得了一个具体实现，因此把变换群 $G_L$ 称为群 $G$ 的**左正则表示**.

* 如果定义

$$\tau_a(x) = xa^{-1}, \quad \forall\, x \in G$$

即把 $G$ 中每个元素用 $a$ 的逆右乘，则同理可证，$G_R = \{\tau_a \mid a \in G\}$ 是集合 $G$ 的一个变换群，并且 $G \cong G_R$. $\tau_a$ 称为由元素 $a$ 引起的**右平移**. 变换群 $G_R$ 称为群 $G$ 的**右正则表示**.

* **推论** 14.4.1　任何一个有限群都同构于一个置换群. ▌

## 习　题　14.4

1. 证明：在群 $G$ 中，对任意元素 $a,b$，方程

$$ax = b$$

有唯一解；方程

$$ya = b$$

也有唯一解.

2. 证明：在群 $G$ 中，消去律成立. 即

$$由\ ax = ay\ 可推出\ x = y$$

$$由\ xa = ya\ 可推出\ x = y$$

3. 在正方形 $E$ 的对称群 $G$ 中，用 $T$ 表示绕正方形的中心且转角为 $90^\circ$ 的旋转，用 $\sigma$ 表示关于过中心且与一边平行的直线 $l_1$ 的反射. 说明 $G$ 中的四个反射是 $\sigma, \sigma T, \sigma T^2, \sigma T^3$.

4. 写出平面上正六边形的对称群的所有元素.

5. 证明：群 $G$ 的任意多个子群的交还是 $G$ 的子群.

*6. 设 $G$ 是一个非空集合，在 $G$ 上面定义了一种运算，称为乘法. 它满足结合律，并且对任意 $a, b \in G$，方程 $ax = b$ 在 $G$ 中有解，方程 $ya = b$ 在 $G$ 中也有解. 证明：$G$ 是一个群.

7. 设 $a$ 是群 $G$ 的任意给定的一个元素，证明：$G$ 中所有与 $a$ 可交换的元素组成的集合是 $G$ 的一个子群，称它为 $a$ **在 $G$ 里的中心化子**，记作 $C_G(a)$.

*8. 设 $\mathrm{char}F \neq 2$，设 $A, B$ 是域 $F$ 上两个合同的可逆对称矩阵，证明：

$$O(A,F) \cong O(B,F)$$

9. 把空间中图形 $E$ 变到自身的空间的旋转变换的全体构成的群，称为**图形 $E$ 的对称群**. 求出空间正四面体的对称群.

10. 在数域 $K$ 上的 $n$ 元多项式环 $K[x_1, \cdots, x_n]$ 中，取一个多项式

$$f(x_1, \cdots, x_n) = \prod_{i<j} (x_i - x_j)$$

对于 $\sigma \in S_n$，用 $\tilde{\sigma}$ 表示置换 $\sigma$ 诱导的代入，即

$$(\tilde{\sigma} f)(x_1, \cdots, x_n) = f(x_{\sigma^{-1}(1)}, \cdots, x_{\sigma^{-1}(n)})$$

显然有 $\qquad \tilde{\sigma} f = f \quad$ 或者 $\quad \tilde{\sigma} f = -f$

如果 $\tilde{\sigma} f = f$，则称 $\sigma$ 是**偶置换**；如果 $\tilde{\sigma} f = -f$，则称 $\sigma$ 是**奇置换**.

(1) 证明：$S_n$ 中全体偶置换组成 $S_n$ 的一个子群，这个群称为 **$n$ 元交错群**，记作 $A_n$；

(2) 证明：$|A_n| = \dfrac{n!}{2}$；

11. 证明：每个 $n$ 元置换都可以表示成一些对换的乘积；设

$$\sigma = \begin{pmatrix} 1 & 2 & \cdots & n \\ a_1 & a_2 & \cdots & a_n \end{pmatrix}$$

则 $\sigma$ 表成对换的乘积时，对换的数目与 $n$ 元排列 $a_1a_2\cdots a_n$ 有相同的奇偶性. 从而 $\sigma$ 表示成对换乘积时，对换的数目的奇偶性由 $\sigma$ 本身唯一决定.

12. 证明：$n$ 元置换 $\sigma$ 为偶置换当且仅当 $\sigma$ 能表示成偶数个对换的乘积.

13. 如果一个 $n$ 元置换 $\sigma$ 将 $1,2,\cdots,n$ 中某 $m$ 个数 $a_1,a_2,\cdots,a_m$ 映成：

$$\sigma(a_1)=a_2,\quad \sigma(a_2)=a_3,\quad \cdots,\quad \sigma(a_{m-1})=a_m,\quad \sigma(a_m)=a_1$$

而保持其余 $n-m$ 个数不变，那么 $\sigma$ 称为一个**轮换**，记作

$$\sigma=(a_1,a_2,\cdots,a_m)$$

把 $m$ 称为轮换 $\sigma$ 的**长度**. 长度为 2 的轮换就是对换.

两个轮换 $(a_1,a_2,\cdots,a_m)$ 与 $(b_1,b_2,\cdots,b_t)$ 称为**不相交**，如果

$$a_i\neq b_j,\quad i=1,2,\cdots,m;\quad j=1,\cdots,t$$

容易看出，不相交的轮换对于乘法是可交换的.

证明：任一 $n$ 元置换 $\sigma$ 都可以分解成一些不相交的轮换的乘积，而且这种分解除了轮换出现的次序外是唯一的.

14. 写出 $A_3,A_4$ 的所有元素(用轮换表法).

15. 证明：正四面体群的对称群与 $A_4$ 同构.

参考文献

1. 北京大学数学系几何与代数教研室代数小组.高等代数(第二版).北京:高等教育出版社.1988

2. A I Kostrikin.代数学引论(上册).张顺燕,蓝以中,丘维声译.北京:高等教育出版社.1988

3. A I Kostrikin & Yu I Manin. Linear Algebra And Geometry. Translated by M E Alferieff. New York. Gordon And Breach Science Publishers. 1989

4. K Hoffman & R Kunze. Linear Algebra. Second Edition. New Jersey. Prentice-Hall Inc. 1971

5. M Postnikov.几何讲义 第二学期 线性代数和微分几何.陈维桓,石生明译.北京:高等教育出版社.1992

6. M Postnikov.几何讲义 第一学期 解析几何.周友成译.北京:高等教育出版社.1992

7. 许以超.代数学引论.上海:上海科技出版社.1965

8. 张禾瑞,郝鈵新.高等代数(第三版).北京:高等教育出版社.1984

9. 李炯生,查建国.线性代数.合肥:中国科学技术大学出版社.1989

10. H B 普罗斯库烈柯夫.线性代数习题集.周晓钟译.北京:人民教育出版社.1981

11. B L Van der Waerden.代数学 I.丁石孙,曾肯成,郝鈵新译.万哲先校.北京:科学出版社.1963

12. N Jacobson. Basic Algebra I. San Francisco. W H Freeman and Company. 1974

13. 聂灵沼,丁石孙.代数学引论.北京:高等教育出版社.1988

14. 莫宗坚,蓝以中,赵春来.代数学上册.北京:北京大学出版社.1986

15. 张恭庆,林源渠.泛函分析讲义(上册).北京:北京大学出版社.1987

16. 丘维声.解析几何.北京:北京大学出版社.1988

17. 王松桂.线性模型的理论及其应用.合肥:安徽教育出版社.1987

［Ｇｅｎｅｒａｌ Ｉｎｆｏｒｍａｔｉｏｎ］
书名＝高等代数（下册）
作者＝丘维声　　编著
页数＝４３１
ＳＳ号＝１０６５２８９１
出版日期＝１９９６年１２月第１版