高等代数（上）

黄山筱

1 预备知识

1.1 基本记号

- $\mathbb{N}$:自然数；$\mathbb{Z}$:整数；$\mathbb{Q}$:有理数；$\mathbb{R}$:实数；$\mathbb{C}$:复数。
- $\forall$: 对任意；$\exists$存在；$\exists!$存在唯一。
- 大写字母A, B, C表示集合，小写字母a, b, c表示元素（例如：$a \in A$ 或 $a \notin B$ 或 $A \cup B$ 或 $A \cap B$ 或$A \subseteq B$），如果集合A仅包含有限个元素，记$|A|$为A中的元素个数，反之，记$|A| = \infty$表示A包含无限个元素。
- $\{x|P(x)\}$表示所有具有性质P的元素构成的集合。设A是一个集合，$\{x \in A|P(x)\}$表示A中所有具有性质P的元素构成的集合。
- $\mathbb{F} \in \{\mathbb{Q}, \mathbb{R}, \mathbb{C}\}$，记$\mathbb{F}[x]$ 为以$\mathbb{F}$为系数的一元多项式全体构成的集合，即

$$\mathbb{F}[x] := \{a_0 + a_1 x + \cdots + a_n x^n | a_0, \ldots, a_n \in \mathbb{F}\}$$

定义1.1.1. 设A, B为两个集合，A, B的**笛卡尔积**定义为

$$\{(a, b)|a \in A, b \in B\},$$

记作$A \times B$，特别地，$A^2 := A \times A$ 以及$A^n := A^{n-1} \times A$。

例1.1.2. 二维平面$\mathbb{R}^2 := \{(a, b)|a, b \in \mathbb{R}\}$，三维空间$\mathbb{R}^3 := \{(a, b, c)|a, b, c \in \mathbb{R}\}$。

设A, B是两个集合。一个从集合A到集合B的映射$f : A \to B$，直观地来讲，就是一个确定的法则，使得对于A中的每个元素x，都有B中唯一确定的元素y与它对应，用集合的语言严格描述即为：

定义1.1.3. 设A, B是两个集合，若子集$f \subseteq A \times B$ 满足，$\forall a \in A$, $\exists! b \in B$ 使得 $(a, b) \in f$。则称$f : A \to B$ 为从A到B的**映射（函数）**。

例1.1.4. 设A是一个集合，记$\mathrm{id}_A : A \to A$为恒等映射，即$\mathrm{id}_A(a) = a$。则恒等映射id_A上述定义中对应子集

$$\{(a, a)|a \in A\}.$$

注1.1.5.

- 事实上，当$A = B = \mathbb{R}$时，定义中的f即为所谓的函数图像（即描述一个函数等价于描述这个函数的图像）。
- 当我们明确知道函数的对应关系时，可以用以下记号表示

$$\begin{aligned} f : A &\to B \\ x &\mapsto y = f(x) \end{aligned}$$

例如，实数域上的二次函数可记作

$$\begin{aligned} f : \mathbb{R} &\to \mathbb{R} \\ x &\mapsto (y =) x^2 \end{aligned}$$

- 对于映射$f:A\to B$，A称为f的**定义域**，B称为f的**终域**。如果另有一个映射$g:B\to C$，即f的终域是g的定义域，可定义f与g的**复合映射**$g\circ f:A\to C, x\mapsto g(f(x))$。
- A在映射f下在B中的象，即$\{f(a)|a\in A\}\subseteq B$，称为$f$的**值域**，记作$\mathrm{Im}(f)$。
- 若$\mathrm{Im}(f)=B$，则称f是一个**满射**。
- 若$\forall x_1\neq x_2\in A$，都有$f(x_1)\neq f(x_2)$，则称f是一个**单射**。
- 若f即使单射也是满射，则称f是一个**双射（一一对应）**。

定义1.1.6. 设$f:A\to B$是一个映射，

- $\forall y\in B$，记$f^{-1}(y):=\{x\in A|f(x)=y\}$；
- $\forall C\subseteq B$，记$f^{-1}(C):=\{x\in A|f(x)\in C\}$；

称为y（或C）在f之下的**原像**。

注1.1.7. 设$f:A\to B$是一个映射，

- f 是满射$\Leftrightarrow \forall y\in B, |f^{-1}(y)|\geq 1$；
- f 是满射$\Leftrightarrow \forall y\in B, |f^{-1}(y)|\leq 1$；
- f 是双射$\Leftrightarrow \forall y\in B, |f^{-1}(y)| = 1$；

命题1.1.8. 设$f:A\to B$ 是一个映射。f是一个双射当且仅当存在映射$g:B\to A$使得$f\circ g=\mathrm{id}_B$且$g\circ f=\mathrm{id}_A$。此时g称为f的逆映射（我们也将这样的g记为f^{-1}）。

证明. 首先假设存在满足条件的g，我们证明f即使单射又是满射。

1. 单射：使用反证法。假设存在$a_1\neq a_2$使得$f(a_1)=f(a_2)$，则由条件有$a_1=\mathrm{id}_A(a_1)=g\circ f(a_1)=g\circ f(a_2)=\mathrm{id}_A(a_2)=s_2$，矛盾。
2. 满射：$\forall b\in B$，由条件$\mathrm{id}_B(b)=f\circ g(b)=f(g(b))$。因此$g(b)$是$b$的一个原像，从而$f$是满射。

下面假设f是双射，我们构造这样的g。由注1.1.7，$\forall b\in B$，有$f^{-1}(b)$中恰有一个元素，因此我们就令$g(b)$等于这个元素（即$\{g^{-1}(b)\}=f^{-1}(b)$）。根据构造，显然$f\circ g(b)=b$ 且$\forall a\in A, g\circ f(a)=a$。 □

命题1.1.9. 设A是一个集合。映射

$$\begin{aligned} f:(A\times A)\times A&\to A\times(A\times A)\\ ((a,b),c)&\mapsto (a,(b,c))\end{aligned}$$

是一个集合之间的双射。

证明. 首先，$\forall(a,(b,c))\in A\times(A\times A)$，显然有$f((a,b),c)=(a,(b,c))$，从而$f$是满射。

另一方面，若$\exists((a_1,b_1),c_1),((a_2,b_2),c_2)\in(A\times A)\times A$使得$f(((a_1,b_1),c_1))=f(((a_2,b_2),c_2))$，即$(a_1,(b_1,c_1))=(a_2,(b_2,c_2))$。从而$a_1=a_2$且$(b_1,c_1)=(b_2,c_2)\Rightarrow a_1=a_2, b_1=b_2, c_1=c_2$。进而有$((a_1,b_1),c_1)=((a_2,b_2),c_2)$，所以$f$也是一个单射。 □

注1.1.10. 设$f:A\to B$是集合之间的一个双射，有时候我们在f的定义较为"自然"的时候，忽略记号f而简记为$A\cong B$，称为集合A与集合B同构。在上面这个例子中，我们可以记为

$$(A\times A)\times A\cong A\times(A\times A),$$

在同构的意义下，我们不妨认为上面两个集合是相同的，也可以进一步简记为$A\times A\times A$。

一般来说，集合的之间很少只用”同构”一词，这是因为集合本身不具有任何结构，”同构”只能说明两个集合的元素数量”一样多”，不能带来更多的信息。更常见的是用在描述其他带有结构的数学对象上，例如将来要学到的线性空间之间。

例1.1.11. 设A是一个集合，A的所有子集构成的集合，称为A的**幂集**，记作2^A。考虑映射

$$\theta : \{\text{一切映射} f : A \to \{0,1\}\text{构成的集合}\} \to 2^A$$
$$f \mapsto \{a \in A | f(a) = 1\}$$

则θ 是一个双射。

思考：

- θ 的逆映射是什么?
- 如果A是一个有限集，$|2^A|$是多少?

1.2 （等价）关系与商集

对于一个集合A，一个常常值得考虑的问题是，将A中根据的不同性质进行将元素分成若干类，使得同一类的元素具有某些相同的性质。这自然引出下面的定义：

定义1.2.1. 设A是一个集合，I是一个指标集，A的一族子集$\{A_i | i \in I, A_i \subseteq A\}$称为集合$A$的一个**分划**，如果

1. $\forall i, j \in I$，有$A_i \cap A_j = \emptyset$;
2. $\cup_{i\in I} A_i = A$。

记为$A = \sqcup_{i\in I} A_i$。

例1.2.2. 考虑以下分划

1. $\{\text{全体人类}\} = \{\text{女性}\} \sqcup \{\text{男性}\}$;
2. $\{\text{全校本科学生}\} = \{\text{一年级本科生}\} \sqcup \{\text{二年级本科生}\} \sqcup \{\text{三年级本科生}\} \sqcup \{\text{四年级本科生}\}$;
3. $\mathbb{R} = \{0\} \sqcup \mathbb{R}_{>0} \sqcup \mathbb{R}_{<0}$;
4. $\mathbb{Z} = \{\text{奇数}\} \sqcup \{\text{偶数}\}$。

虽然集合的分划是一个直观的概念，但是在实际数学问题中，显式地描述分划往往十分困难，因此我们需要引入集合分划的一个等价描述，也就是下面所说的等价关系。

”关系”指的是同一集合上元素间的关系。严格地来说，所谓**集合A上的一个关系**即为$A \times A$上的一个子集$R \subseteq A \times A$。对于$x, y \in A$，若二元对$(x, y) \in R$，则称x与y有(R)关系，记为$x \sim_R y$（或简记为$x \sim y$）。

定义1.2.3. 设A是一个集合，A上的一个关系$R \subseteq A \times A$称为**等价关系**，如果R满足

1. （自反性）$\forall x \in A, x \sim x$;
2. （对称性）若$x \sim y$，则$y \sim x$;
3. （传递性）若$x \sim y$且$y \sim z$，则$x \sim z$。

若$x \sim y$，我们称x与y（**对于关系R**）**等价**。

例1.2.4. 考虑以下定义的关系是否是等价关系

1. 人类全体集合中的夫妻关系；

2. 人类全体集合中的朋友关系；

3. 人类全体集合中的血缘关系；

4. 全班同学中的同性别关系（即两个具有关系当且仅当两人同性别）；

5. 全校同学中的同性专业关系（即两个具有关系当且仅当两人同专业）；

例1.2.5. 设A是一个集合，考虑以下定义的关系是否是等价关系

1. 若$A=\mathbb{Q}$，定义$x\sim y$当且仅当$x\leq y$；

2. 恒等函数id_A；

3. 对任意$x,y\in A$，都定义$x\sim y$；

4. $A=\mathbb{R}$，定义$x\sim y$当且仅当$x-y\in\mathbb{Z}$。

正如本小节开头所言，等价关系是集合分划的一个等价描述，具体的来说，就是有以下结果：

命题1.2.6. 设A是一个集合，A的所有分划构成的集合和A上所有的等价关系构成的集合有自然的一一对应，即以下映射

$$\begin{aligned}\{\{A_i|i\in I\}\text{是}A\text{的一个分划}\}&\to\{R\subseteq A\times A|R\text{是一个等价关系}\}\\ \mathcal{A}=(A_1,\dots,A_n)&\mapsto R_{\mathcal{A}}\end{aligned}$$

是一个双射。其中，$R_{\mathcal{A}}$定义如下：$\forall a,b\in A$，$(a,b)\in R_{\mathcal{A}}$ 当且仅当$\exists i$使得$a,b\in A_i$。

证明. 首先我们验证$R_{\mathcal{A}}$ 是一个等价关系。

1. 自反性：显然。（由于$\sqcup_i A_i=A$，$\forall a\in A$必然存在i使得$a\in A_i$）；

2. 对称性：显然。（若$a,b\in A$，自然有$b,a\in A$）；

3. 传递性：对于$a,b,c\in A$，若存在i,j使得$a,b\in A_i$且$b,c\in A_j$ ，则$b\in A_i\cap A_j\Rightarrow i=j$。从而$a\sim c$

下面构造命题中映射的逆映射：设R是一个A上的等价关系，$\forall a\in A$，令$[a]:=\{b\in A|a\sim_R b\}$。则$\{[a]|a\in A\}$是一个$A$上的分划（显然有$A=\cup_{a\in A}[a]$，另外若$c\in[a]\cap[b]$，由传递性得$[a]=[b]$）。

请读者自行验证这是逆映射。 □

注1.2.7. 设A是一个集合，R是A上的一个等价关系，记A/R或$A/\sim_R$为此等价关系对应的分划，即

$$A/\sim_R:=\{[a]|a\in A\}.$$

$A/\sim_R$称为A关于等价关系R的**商集**，自然映射$A\to A/\sim_R, a\mapsto[a]$称为**（关于R的）商映射**。

例1.2.8. 设A是一个集合，考虑以下等价关系的商集

1. 恒等函数id_A；

2. 对任意$x,y\in A$，都定义$x\sim y$；

3. $A=\mathbb{R}$，定义$x\sim y$当且仅当$x-y\in\mathbb{Z}$。

4. $A=\mathbb{Z}$，定义$x\sim y$当且仅当$2|(x-y)$。

5. 设B是另一个集合，$f:A\to B$是满射，定义$a\sim b$当且仅当$f(a)=f(b)$。

例1.2.9. （有理数集的构造） 在集合$A:=\mathbb{Z}\times(\mathbb{Z}\backslash\{0\})$上规定关系：$(a,b)\sim(c,d)$当且仅当$ad=bc$。这是一个等价关系：

1. 自反性：$ab = ab$;
2. 对称性：$ad = bc \Rightarrow cb = da$;
3. 传递性：若$(a,b) \sim (c,d), (c,d) \sim (e,f)$，则$ad = bc, cf = de$。从而$adf = bcf = bde$。由于$d \neq 0$，有$af = be$，即$(a,b) \sim (e,f)$。

在集合$A/\sim$上还可以定义两种二元运算，加法+及乘法×，如下：

1. $[(a,b)] + [(c,d)] := [(ad+bc, bd)]$;
2. $[(a,b)] \times [(c,d)] = [(ac, bd)]$。

这两种运算都是良定义的，这里仅验证加法，乘法留给读者作练习。设$(a_1,b_1) \sim (a_2,b_2)$，则

$$\begin{aligned}[(a_1,b_1)] + [(c,d)] &= [(a_1d + b_1c, b_1d)] \\ &= [(b_2a_1d + b_2b_1c, b_2b_1d)] \\ &= [(a_2b_1d + b_2b_1c, b_2b_1d)] \\ &= [(a_2d + b_2c, b_2d)] = [(a_2,b_2)] + [(c,d)]\end{aligned}$$

事实上，上述构造$(A/\sim, +, \times)$即为有理数集以及其上的四则运算。

1.3 习题

习题1.3.1. 设A, B是集合。映射

$$\begin{aligned} f : A \times B &\to B \times A \\ (a,b) &\mapsto (b,a)\end{aligned}$$

是一个集合之间的双射。

习题1.3.2. 描述例1.2.5中的关系作为$A \times A$中的子集。

习题1.3.3. 设p是一个给定素数。在$\mathbb{Z}$上规定关系：$a \sim_p b$当且仅当$p|(a-b)$。证明$\sim_p$是一个等价关系，并计算$|\mathbb{Z}/\sim_p|$。

2 多项式

设$\mathbb{F} \in \{\mathbb{Q}, \mathbb{R}, \mathbb{C}\}$，本章节的主要目标是理解多项式集合$\mathbb{F}[x]$上乘法运算的结构。具体来说，在$\mathbb{F}[x]$上的乘法结构中具有最小单位，称之为不可约多项式，而一般的一个多项式$f(x) \in \mathbb{F}[x]$（在某种意义下）可唯一地进行分解

$$f(x) = p_1(x)^{t_1} \cdots p_k(x)^{t_k},$$

其中q_i是不可约多项式。为了理解和证明上述唯一分解定理和其相关内容，我们需要花费一整个章节的努力，学习多项式理论中的许多知识。

2.1 运算与数域

粗略地来说，人类在对运算的不断探索中，逐渐扩大对“数”的概念的理解。最初人们学会数数的时候，人类“发现”（同时也是“发明”）了自然数，随之创造了加法运算。为了逆运算（即减法）的封闭性，人类将自然数集扩展到了整数集。出于类似的原因，为了乘法的逆运算的封闭性，人们进一步扩大数集得到了有理数集。最后为了极限运算的封闭性人们再次将扩其张得到实数集，并且为了代数的封闭性人们最终得到了复数。

$$\mathbb{N} \xrightarrow{\text{加法逆}} \mathbb{Z} \xrightarrow{\text{乘法逆}} \mathbb{Q} \xrightarrow{\text{极限}} \mathbb{R} \xrightarrow{\text{代数封闭}} \mathbb{C}$$

为了更好地理解上面这段话，我们首先需要引入“二元运算”及“加法群”的概念。设A是一个集合，若$* : A \times A \to A$ 是一个映射，则称元素对$(A, *)$是**（A上的）二元运算**。

例2.1.1. 以下几个例子都是二元运算。

1. $(\mathbb{N}, +)$，$(\mathbb{Z}, +)$；
2. $(\mathbb{N}, *)$，$(\mathbb{Q}, *)$；
3. 设A是一个集合，$(A^A, \circ)$（回忆A^A是A上所有到自身的映射构成的集合）。

定义2.1.2. （加法群）设$(A, *)$是一个二元运算。称$(A, *)$构成一个**加法群**，如果满足

1. （结合律）$\forall a, b, c \in A$，有$(a * b) * c = a * (b * c)$；
2. （交换律）$\forall a, b \in A$，有$a * b = b * a$；
3. （零元）$\exists e \in A$使得$\forall a \in A$，有$e * a = a$；
4. （逆元）$\forall a \in A$，$\exists b \in A$使得$a * b = e$（此时b记为$-a$）。

注2.1.3. 若$(a, *)$是一个加法群，则零元e唯一（设e_1, e_2是A中的零元，有$e_1 = e_1 * e_2 = e_2 * e_1 = e_2$）。对任意$a \in A$，$a$的逆元也唯一（设$b_1, b_2$是$a$的逆元，有$b_1 = b_1 * (a * b_2) = (b_1 * a) * b_2 = (a * b_1) * b_2 = b_2$）。

定义2.1.4. （域）设$(A, +)$是一个是一个加法群，并设$0 \in A$是此加法群上的零元，记$A^* := A \backslash \{0\}$。若$(A^*, \times)$是另一个加法群，满足

1. $\forall a \in A$，有$0 \times a = 0$；
2. (分配律)$\forall a, b, c \in A, a \times (b + c) = a \times b + a \times c$；

则称三元组$(A, +, \times)$是一个**域**。

简单来说，域就是一个集合上有两个相协调的加法群（即满足分配律）。

例2.1.5. 考虑以下域的例子

1. $(\mathbb{Q}, +, \times)$；
2. 定义$\mathbb{Q}(\sqrt{2}) := \{a + b\sqrt{2} | a, b \in \mathbb{Q}\}$；

3. 令$A := \mathbb{Q} \times \mathbb{Q}$，定义

   - $(a_1, a_2) + (b_1, b_2) := (a_1 + a_2, b_1 + b_2)$ ；
   - $(a_1, a_2) \times (b_1, b_2) = (a_1a_2 + 2b_1b_2, a_1b_2 + a_2b_1)$；

4. 令$A := \mathbb{R} \times \mathbb{R}$，定义

   - $(a_1, a_2) + (b_1, b_2) := (a_1 + a_2, b_1 + b_2)$ ；
   - $(a_1, a_2) \times (b_1, b_2) = (a_1a_2 - b_1b_2, a_1b_2 + a_2b_1)$；

   事实上，这里的$(A, +, \times)$就是我们曾经学过的复数域$\mathbb{C}$，其中$i = (0, 1)$。当我们使用复数域的符号$\mathbb{C}$，我们默认上面已经定义了我们所熟知的加法$+$及乘法$\times$（有理数域$\mathbb{Q}$，实数域$\mathbb{R}$同理）。

未来我们将会讨论更多的域的例子，以及从不同的角度理解复数域。

定义2.1.6. 设$A \subseteq \mathbb{C}$是复数域的一个子集，如果$\mathbb{C}$在A诱导了自然的域结构，则称A是一个**数域**（即复数域上的加法$+$和乘法$\times$在A上封闭，且$(A, +, \times)$构成一个域）。

2.2 带余除法与最大公因式

在本小节中，我们固定$\mathbb{F}$是一个数域。我们先回顾多项式集合$\mathbb{F}[x]$中的基本运算和符号。设$f(x) = a_0 + a_1x + \cdots a_nx^n \in \mathbb{F}[x]$且$a_n \neq 0$，

- $\deg(f) = n$（补充定义$\deg(0) = -\infty$）；
- $l(f) = a_n$，若$a_n = 1$称f为**首一多项式**；
- 若另有$g(x) = b_0 + b_1x + \cdots b_mx^m$，定义

$$f(x) + g(x) = (a_0 + b_0) + (a_1 + b_1)x + \cdots (a_n + b_n)x^n$$

  （不妨设$m \leq n$并补充定义$b_{m+1}, \ldots, b_n = 0$），和

$$f(x)g(x) := c_0 + c_1x + \cdots + c_{n+m}x^{n+m},$$

  其中$c_k = a_0b_k + a_1b_{k-1} \cdots a_kb_0$。
- 若另有$g, h \in \mathbb{F}[x]$使得$f = gh$，则记为$g \mid f$（也有$h \mid f$）。

注2.2.1. 我们有

- $\deg(f + g) \leq \max\{\deg(f), \deg(g)\}$。若$\deg(f) \neq \deg(g)$，不等式取等号；
- $\deg(fg) = \deg(f) + \deg(g)$；
- 定义$f \sim g$当且仅当存在$a \in \mathbb{F}^*$使得$f = cg$，称f与g**相伴**。相伴关系是等价关系。存在唯一$g \in [f]$使得$l(g) = 1$；
- 若$f \mid g$且$g \mid f$，则$f \sim g$。

我们将自然数上的带余除法进行类比，得到多项式集上的带余除法。

命题2.2.2. （带余除法）设$f, g \in \mathbb{F}[x]$满足$\deg(g) \leq \deg(f)$，则存在唯一$q, r \in \mathbb{F}[x]$满足

- $f = qg + r$；
- $\deg(r) < \deg(g)$。

证明. 先证明存在性。考虑集合$\{f-hg|h\in\mathbb{F}[x]\}$中次数最低的多项式，设在某个$q\in\mathbb{F}[x]$时取到。若$\deg(f-qg)\geq\deg(g)$，通过调整$q$即可得到矛盾（设$g(x)=a_nx^n+\cdots$及$(f-qg)(x)=a_mx^m+\cdots$，其中$m\geq n$。则$\deg(f-(q+\frac{a_m}{a_n})g)<\deg(f-qg)$）。

下证唯一性。若有两组$(q_1,r_1),(q_2,r_2)$满足命题中的条件，则$(q_1-q_2)g=(r_2-r_1)$。若$q_1\neq q_2$，则$\deg(r_1-r_2)<\deg(g)<\deg((q_1-q_2)g)$矛盾。从而$q_1=q_2$，进而$r_1=r_2$。 □

例2.2.3. 计算x^4+2除x^2+1的带余除法。

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

从而$x^4+2=(x^2-1)(x^2+1)+3$。

推论2.2.4. 设$f(x)\in\mathbb{F}[x]$。对任意$a\in\mathbb{F}$，$(x-a)\mid f(x)$当且仅当$f(a)=0$。

证明. 考虑$f(x)$对$(x-a)$作带余除法，得到

$$f(x)=g(x)(x-a)+r,$$

此处$r\in\mathbb{F}[x]$。因此$(x-a)\mid f(x)$当且仅当$r=0$。注意到$r=f(x)-g(x)(x-a)=f(a)-g(a)(a-a)=f(a)$，从而得到结论。 □

推论2.2.5. 设$f(x)\in\mathbb{F}[x]$。若$\deg(f)=n\neq-\infty$且存在互不相等的$a_1,\ldots,a_n\in\mathbb{F}[x]$使得$f(a_i)=0$，则$f(x)=l(f)(x-a_1)\cdots(x-a_n)$。

证明. 我们对n进行归纳证明。

1. 当$n=1$时，显然。
2. 假设当$n=k-1$时命题成立，考虑$n=k$时的情况。由于$f(a_n)=0$，由上一推论有$(x-a_n)\mid f(x)$。设$f(x)=(x-a_n)g(x)$，则$\deg(g)=k-1$且$g(a_1)=\cdots=g(a_{k-1})=0$。由归纳假设即得结论。

□

推论2.2.6. 设$f(x)\in\mathbb{F}[x]$满足$\deg(f)\leq n$。若存在$a_0,\ldots,a_n$使得$f(a_i)=0$，则$f(x)=0$。

定义2.2.7. 设$f,g\in\mathbb{F}[x]$。若$h\in\mathbb{F}[x]$使得$h\mid f$和$h\mid g$则称h是f,g的**公因式**。r若$h_0\in\mathbb{F}[x]$满足

1. h_0是f,g的公因式;
2. $\forall h$是f,g的公因式，都有$h\mid h_0$;

则称h_0是f,g的**最大公因式**。最大公因式若存在，所有的最大公因式都相伴，其中的首一多项式记为(f,g)。

命题2.2.8. 设$f,g\in\mathbb{F}[x]$。则(f,g)存在且存在$u,v\in\mathbb{F}[x]$使得$uf+vg=(f,g)$。

证明. 取$u,v\in\mathbb{F}[x]$使得$uf+vg$次数最低，首项系数为1且不等于0。显然对任意f,g的公因式h，都有$h\mid(uf+vg)$。

若$uf+vg\nmid f$（注意到$\deg(uf+vg)\leq\deg(f)$，则存在$q,r\in\mathbb{F}$使得$\deg(r)<\deg(uf+vg)$且$f=q(uf+vg)+r$。则$r=(qu-1)f+qvg$，矛盾。因此$uf+vg\mid f$且同理$uf+vg\mid g$。

□

注2.2.9. 从上述的证明中可知，可以通过辗转相除法得到f,g的最大公因式（不妨设$\deg(g) \geq \deg(f)$）。

1. 若$f \mid g$，则$(f,g) = l(f)^{-1}f$。
2. 否则

$$\begin{aligned} g &= q_1 f + r_1 \\ f &= q_2 r_1 + r_2 \\ r_1 &= q_3 r_2 + r_3 \\ &\cdots \\ r_{n-2} &= q_n r_{n-1} + r_n (\text{且} r_n \mid r_{n-1}) \end{aligned}$$

从而r_n可以由f,g生成且由归纳可知$r_n \mid f,g$。从而$(f,g) = l(r)^{-1}r$。

例2.2.10. 求$f(x) = x^3 + x^2 + x + 1$和$g(x) = x^4 + x^3 + x + 1$的最大公因式，并表达成$uf + vg$的形式。

证明.

$$\begin{aligned} x^4 + x^3 + x + 1 &= x(x^3 + x^2 + x + 1) + (-x^2 + 1) \\ x^3 + x^2 + x + 1 &= (-x - 1)(-x^2 + 1) + 2(x + 1) (\text{且} x + 1 \mid x^2 - 1) \end{aligned}$$

从而$x + 1 = (f,g)$且

$$\begin{aligned} x + 1 &= \frac{1}{2}[(x^3 + x^2 + x + 1) - (-x - 1)(-x^2 + 1)] \\ &= \frac{1}{2}[(x^3 + x^2 + x + 1) - (-x - 1)((x^4 + x^3 + x + 1) - x(x^3 + x^2 + x + 1))] \\ &= \frac{1}{2}[(-x^2 - x + 1)(x^3 + x^2 + x + 1) + (x + 1)(x^4 + x^3 + x + 1)] \end{aligned}$$

□

2.3 唯一分解定理

本小节将要证明本章开头所提到的多项式唯一分解定理。为了证明定理，需要一些准备工作。

定义2.3.1. 设$f(x), g(x) \in \mathbb{F}[x]$称为**互素的**当且仅当$(f(x), g(x)) = 1$。

推论2.3.2. $f(x), g(x) \in \mathbb{F}[x]$互素当且仅当存在$u(x), v(x) \in \mathbb{F}[x]$使得$u(x)f(x) + v(x)g(x) = 1$。

证明. 必要性是命题2.2.8的直接推论。下证充分性。显然有$(f,g) \mid 1 = uf + vg$。注意到$\deg((f,g)) \leq \deg(1) = 0$，只能有$(f,g) = 1$。 □

命题2.3.3. 设$f(x), g(x) \in \mathbb{F}[x]$满足$(f(x), g(x)) = 1$。则对任意$m, n \in \mathbb{N}$，都有$(f^m, g^n) = 1$。

证明. 由推论2.3.2知存在$u, v \in \mathbb{F}[x]$使得$uf + vg = 1$，从而

$$\begin{aligned} 1 &= (uf + vg)^{m+n} \\ &= \sum_{i=0}^{m+n} C_{m+n}^i (uf)^i (vg)^{m+n-i} \\ &= \sum_{i=0}^{m} C_{m+n}^i (uf)^i (vg)^{m+n-i} + \sum_{i=m+1}^{m+n} C_{m+n}^i (uf)^i (vg)^{m+n-i} \\ &= \left(\sum_{i=0}^{m} C_{m+n}^i (uf)^i (vg)^{m-i} \right) v^n g^n + \left(\sum_{i=1}^{n} C_{m+n}^{m+i} (uf)^i (vg)^{n-i} \right) u^m f^m \end{aligned}$$

从而$(f^m, g^n) = 1$。 □

命题2.3.4. 设$p_1(x),\dots,p_m(x),f(x)\in\mathbb{F}[x]$满足

1. 对于$i=1,\dots,n$，有$p_i\mid f(x)$;
2. 对于$1\le i\neq j\le m$，有$(p_i,p_j)=1$,

则$p_1\cdots p_m\mid f$。

证明. 对m进行归纳证明。当$m=1$时命题显然成立。假设当$m=k-1$时命题成立，考虑$m=k$的情况。对于$i=1,\dots,k-1$，有$p_i\mid(f/p_k)p_k$。由于$(p_i,p_k)=1$，则存在$u,v\in\mathbb{F}[x]$使得$up_i+vp_k=1$。两边同乘f/p_k得$up_i(f/p_k)+vp_k(f/p_k)=f/p_k$。从而$p_i\mid f/p_k$。由归纳假设得$p_1\cdots p_{k-1}\mid f/p_k$，即$p_1\cdots p_m\mid f$。 □

定义2.3.5. 设$f(x)\in\mathbb{F}[x]_{\deg\ge1}$（此处记号意味着次数$\ge1$的多项式构成的集合），我们称$f(x)$是

1. **（在$\mathbb{F}$上的）可约多项式**，若存在$g(x),h(x)\in\mathbb{F}[x]$满足$\deg(g),\deg(h)<\deg(f)$且$f=gh$;
2. **（在$\mathbb{F}$上的）不可约多项式**如果$f(x)$不是可约多项式。

注2.3.6. 多项式的不可约性依赖$\mathbb{F}$的选取，例如

1. $x^2-2=(x-\sqrt{2})(x+\sqrt{2})$，因此$x^2-2$在$\mathbb{R}$上是可约的，但是在$\mathbb{Q}$上是不可约的;
2. $x^2+1=(x-i)(x+i)$，因此x^2+1在$\mathbb{C}$上是可约的，但是在$\mathbb{R}$上是不可约的。

命题2.3.7. 设$f(x),g(x)\in\mathbb{F}[x]$且$f(x)$不可约，则必有$(f,g)=1$或$f\mid g$。

证明. 由于$f=(f,g)\frac{f}{(f,g)}$，由于f不可约，必有$\deg(f)=\deg(f,g)$或$\deg(f)=\deg(\frac{f}{(f,g)})$。前者意味着$f\sim(f,g)$从而$f\mid g$，后者意味着$(f,g)=1$。 □

命题2.3.8. 设$p(x),f_1(x),\dots,f_m(x)\in\mathbb{F}[x]$且$p(x)$是不可约多项式满足$p(x)\mid f_1(x)\cdots f_m(x)$，则存在$i\in\{1,\dots,m\}$使得$p(x)\mid f_i(x)$。

证明. 假设命题不成立，则对任意$i=1,\dots,m$都有$(p,f_i)=1$。由推论2.3.2知存在$u_i,v_i\in\mathbb{F}[x]$使得$u_ip+v_if_i=1$，从而

$$\begin{aligned}1&=(u_1p+v_1f_1)\cdots(u_mp+v_mf_m)\\&=((u_1p+v_1f_1)\cdots(u_mp+v_mf_m)-(v_1\cdots v_m)\cdot f_1\cdots f_m)+(v_1\cdots v_m)\cdot f_1\cdots f_m.\end{aligned}$$

由于p同时整除$(u_1p+v_1f_1)\cdots(u_mp+v_mf_m)-(v_1\cdots v_m)\cdot f_1\cdots f_m$以及$(v_1\cdots v_m)\cdot f_1\cdots f_m$。这推出$p\mid1$，矛盾！ □

定义2.3.9. 设$p(x),f(x)\in\mathbb{F}[x]$且$p(x)$不可约，且$p(x)\mid f(x)$，则称$p(x)$是$f(x)$的**不可约因式**。称$e\in\mathbb{N}$是$p$在$f$中的**重数**，若满足$p^e\mid f$且$p^{e+1}\nmid f$，记为$e(p|f)$（若$e\nmid f$，也记为$e(p|f)=0$）。

定理2.3.10. （多项式的标准分解）设$f(x)\in\mathbb{F}[x]$且$\deg(f)>0$。则$f(x)$可分解为不可约多项式的乘积，即(在不计次序的意义下)可唯一分解为以下形式

$$f(x)=l(f)p_1(x_1)^{e_1}p_2(x_2)^{e_2}\cdots p_m(x)^{e_m}$$

其中$p_1,\dots,p_m$是互不相同的首一不可约多项式。

证明. 显然$c=l(f)$，用$f(x)/l(f)$替代$f(x)$后我们可不妨假设$f(x)$是首一多项式。

设$p_1(x),\dots,p_m(x)$是所有$f(x)$的不同的不可约因子（f只有有限多个不可约因子，否则存在$p_1,\dots,p_k$都整除f且p_i不可约且$k>\deg(f)$，则由命题2.3.4有$p_1\cdots p_k\mid f$，矛盾！），并且p_i在f中的重数为e_i。由命题2.3.3知$(p_i^{e_i},p_j^{e_j})=1$，从而由命题2.3.4知$p_1^{e_1}\dots p_m^{e_m}\mid f$。若$\frac{f}{p_1^{e_1}\dots p_m^{e_m}}$还有不可约因子$q$，则$q$同时也是$f$的不可约因子。根据定义必有存在$i$使得$q=p_i$。但这意味着$p_i^{e_i+1}\mid f$，与$e_i$是$p_i$在$f$中的重数矛盾！从而$\frac{f}{p_1^{e_1}\dots p_m^{e_m}}$无不可约因子且首项系数为1，得到$f(x)=cp_1(x_1)^{e_1}p_2(x_2)^{e_2}\cdots p_m(x)^{e_m}$。

□

注2.3.11. 上述多项式的标准分解也可以表达为

$$f(x) = l(f) \prod_{p\in\mathbb{F}[x]\text{首一不可约}} p^{e(p|f)}.$$

注2.3.12. 有了多项式的标准分解后，多项式的整除关系和公因式的分析变得清晰明了。具体来说，有以下推论

1. 设$f(x), g(x) \in \mathbb{F}[x]$，则$(f, g) = 1$当且仅当$\forall p(x) \in \mathbb{F}[x]$不可约，都有$e(p|g) = 0$ 或$e(p|f) = 0$。
2. 设$f(x), g(x) \in \mathbb{F}[x]$，则$g \mid f$当且仅当$\forall p(x) \in \mathbb{F}[x]$不可约，都有$e(p|g) \le e(p|f)$。
3. 设$f(x), g(x) \in \mathbb{F}[x]$，则$\forall p(x) \in \mathbb{F}[x]$不可约，都有$e(p|fg) = e(p|f) + e(p|g)$。
4. 设$f(x), g(x) \in \mathbb{F}[x]$，则
$$(f, g) = \prod_{p\in\mathbb{F}[x]\text{首一不可约}} p^{\min\{e(p|f),e(p|g)\}}$$

且之前的推论2.3.4以及命题2.3.7，2.3.8都十分显然。

(对于推论2.3.4，条件1表示$\forall q \in \mathbb{F}[x]$不可约，都有$e(q|p_i) \le e(q|f)$；条件2表示$\forall q \in \mathbb{F}[x]$不可约，至多有一个$i$使得$e(q|p_i) \neq 0$。从而$e(q|p_1 \cdots p_m) = e(q|p_1) + \cdots + e(q|p_m) \le e(q|f)$)。

(对于命题2.3.7，$e(f|g) = 0$意味着$(f, g) = 1$而$e(f|g) > 0$意味着$f \mid g$。)

(对于命题2.3.8，条件意味着$1 \le e(p|f_1 \cdots f_m) = e(p|f_1) + \cdots + e(p|f_m)$。从而必存在某个$i$使得$e(p|f_i) \ge 1$，即$p \mid f_i$。)

2.4 代数基本定理

定义2.4.1. （欧拉公式）设$\theta \in \mathbb{R}$，规定

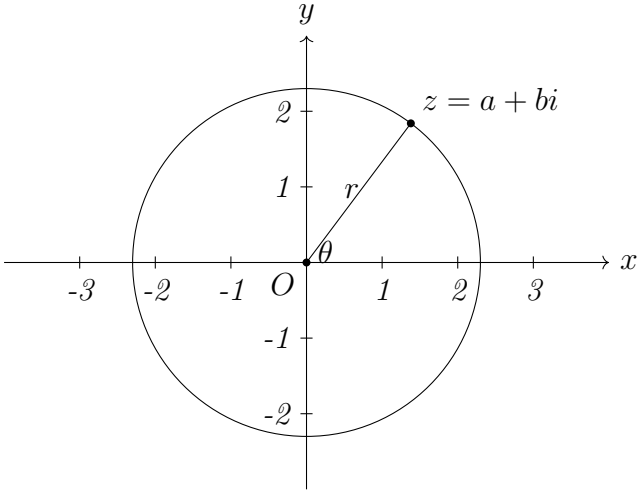
$$e^{i\theta} := \cos\theta + i\sin\theta$$

命题2.4.2. 设$z = a + bi \in \mathbb{C}^*$，则存在$r > 0, \theta \in \mathbb{R}$使得

$$z = re^{i\theta}.$$

实际上，$r = |z| = \sqrt{a^2 + b^2}$ 且θ满足 $\cos\theta = \frac{a}{\sqrt{a^2+b^2}}, \sin\theta = \frac{b}{\sqrt{a^2+b^2}}$。

特别地，对任意$k \in \mathbb{Z}$ 都有$z = re^{i(\theta+2k\pi i)}$ 且$\bar{z} = re^{-i\theta}$。



证明. 直接带入欧拉公式即可。 □

命题2.4.3. 设$z_1 = r_1 e^{i\theta_1}, z_2 = r_2 e^{i\theta_2} \in \mathbb{C}$，则

$$z_1 z_2 = r_1 r_2 e^{i(\theta_1+\theta_2)}$$

证明. 带入欧拉公式，有

$$\begin{aligned} z_1 z_2 &= r_1(\cos\theta_1 + i\sin\theta_1) r_2(\cos\theta_2 + i\sin\theta_2) \\ &= r_1 r_2[(\cos\theta_1\cos\theta_2 - \sin\theta_1\sin\theta_2) + i(\cos\theta_1\sin\theta_2 + \sin\theta_1\cos\theta_2)] \\ &= r_1 r_2[\cos(\theta_1+\theta_2) + i\sin(\theta_1+\theta_2)] \text{ (和角公式)} \\ &= r_1 r_2 e^{i(\theta_1+\theta_2)}. \end{aligned}$$

□

定理2.4.4. （代数基本定理，Gauβ） 设$f(z) \in \mathbb{C}[z]_{\deg\geq 1}$，则存在$z_0 \in \mathbb{C}$使得$f(z_0) = 0$。

证明. 使用反正法假设$f(z)$无零点。设$|f(z)|$在某个ω处取到最小值（最小值点的存在性需要用到数学分析中的结论：紧集上的连续函数存在最大最小值，超出本门课的教学范围），用函数$z \mapsto f(z-\omega)$替代$f(z)$可不妨假设最小值在$z = 0$处取到。

设$f(z) = a_0 + a_k z^k + \cdots + a_n z^n$，其中$a_0, a_k, a_n \neq 0$（若$a_0 = 0$则$f(0) = 0$矛盾）。则

$$f(z) = a_0 + a_k z^k + z^k(a_{k+1}z + \cdots + a_n z^{n-k}).$$

设$-a_0/a_k = re^{i\theta}$，取$z_0 = r_0 e^{i\theta/k}$，其中$1 > r_0 > 0$待定，则

$$\begin{aligned} |f(z_0)| &= |a_0 + a_k z^k + z^k(a_{k+1}z + \cdots + a_n z^{n-k})| \\ &\leq |a_0 + a_k z^k| + |z|^k |a_{k+1}z + \cdots + a_n z^{n-k}| \\ &= |a_0||1 - re^{i\theta} r_0^k e^{i\theta}| + r_0^k |a_{k+1}z + \cdots + a_n z^{n-k}| \\ &\leq |a_0||1 - rr_0^k| + r_0^{k+1} b \quad (\text{其中} b = \max\{|a_{k+1}||i = 1, \ldots, n-k\}) \\ &\leq |a_0||1 - rr_0^k| + r_0^{k+1} b \end{aligned}$$

若进一步要求$\min\{1, 1/r^{1/k}\} > r_0 > 0$，则由上式知$|f(z_0)| \leq |a_0|(1 - rr_0^k) + r_0^{k+1}b$。取$r_0 > 0$充分小时（即$r_0 < |a_0|r/b$）必有$r_0^{k+1}b < |a_0|rr_0^k$。则此时$|f(z_0)| \leq |a_0|(1 - rr_0^k) + |a_0|rr_0^k = |a_0|$，与$|f(z)|$在$z = 0$处取到最小值矛盾！从而必存在$z_0 \in \mathbb{C}$使得$f(z_0) = 0$。 □

推论2.4.5. $f(z) \in \mathbb{C}[z]$是不可约多项式当且仅当$\deg(f) = 1$。

证明. 由代数基本定理，存在$z_0 \in \mathbb{C}$使得$f(z_0) = 0$。由推论2.2.4知$(z - z_0) \mid f(z)$。从而$f(z)$不可约当且仅当$f(z) \sim (z - z_0)$。 □

推论2.4.6. 设$f(z) \in \mathbb{C}[z]$，则存在两两不同的$z_1, \ldots, z_m \in \mathbb{C}(m \leq \deg(f))$使得

$$f(z) = l(f)\prod_{i=1}^{m}(z - z_i)^{e_i}.$$

事实上，$z_1, \ldots, z_m$是函数$f(z)$的所有零点，$e_i = e(z - z_i|f)$称为$f(z)$在零点z_i处的重数。

特别地，$e_1 + \cdots + e_m = \deg(f)$。

证明. 由多项式的标准分解和推论2.4.6即得。 □

推论2.4.7. $f(x) \in \mathbb{R}[x]$的不可约多项式仅有以下两种情形

1. $f(x) \sim x - a$，其中$a \in \mathbb{R}$;
2. $f(x) \sim x^2 + bx + c$，其中$b, c \in \mathbb{R}$满足$b^2 - 4c < 0$。

证明. 设$f(x)$在$\mathbb{R}$上不可约。由代数基本定理知$\exists z_0 \in \mathbb{C}$使得$(x-z_0)|f(x)$。若$z_0 \in \mathbb{R}$则有$f(x) \sim x-z_0$，即为情况1。否则$z_0 \in \mathbb{C}\backslash\mathbb{R}$。注意到$f(\bar{z_0}) = \overline{f(z_0)} = \bar{0} = 0$知$(x-\bar{z_0})|f(x)$。由推论2.3.4有$(x-z_0)(x-\bar{z_0}) \mid f(x)$。由于$(x-z_0)(x-\bar{z_0}) = x^2 - (z_0+\bar{z_0})x + z_0\bar{z_0} = x^2 - 2\mathrm{Re}(z_0)x + |z_0|^2 \in \mathbb{R}[x]$，从而$f(x) \sim x^2 - 2\mathrm{Re}(z_0)x + |z_0|^2$，即为情况2。 □

推论2.4.8. 设$f(x) \in \mathbb{C}[z]_{\deg\geq 1}$，则$f(x)$具有标准分解

$$f(x) = l(f)\prod_{i=1}^{m}(x-a_i)^{e_i}\prod_{j=1}^{n}(x^2+b_jx+c_j)^{l_j},$$

其中$a_i, b_j, c_j \in \mathbb{R}$且$a_i$之间两两不同，数对$(b_j, c_j)$之间两两不同。

特别地，$e_1 + \cdots + e_m + 2l_1 + \cdots + 2l_n = \deg(f)$。

证明. 由多项式的标准分解和推论2.4.7即得。 □

例2.4.9. 分别讨论$f(x) = x^n - 1$在$\mathbb{C}$以及$\mathbb{R}$上的标准分解。

证明. 首先考虑在$\mathbb{C}$上的分解。注意到取$z_k = e^{\frac{2\pi ki}{n}}, k = 0, 1, \ldots, n-1$时，有

$$f(z_k) = (e^{\frac{2\pi ki}{n}})^n - 1 = e^{2\pi i} - 1 = 0,$$

从而

$$f(x) = \prod_{k=0}^{n-1}(x-z_k) = \prod_{k=0}^{n-1}(x-e^{\frac{2\pi ki}{n}}).$$

对于$\mathbb{R}$上的分解。若n是奇数，$z_k \in \mathbb{R}$当且仅当$k=0$，且当$k \neq 0$时，$\bar{z_k} = z_{n-k}$。从而

$$\begin{aligned} f(x) &= (x-z_0)\prod_{k=1}^{\frac{n-1}{2}}(x^2 - (z_k+z_{n-k})x + z_kz_{n-k}) \\ &= (x-1)\prod_{k=1}^{\frac{n-1}{2}}(x^2 - 2\mathrm{Re}(z_k)x + |z_k|^2) \\ &= (x-1)\prod_{k=1}^{\frac{n-1}{2}}(x^2 - 2\cos\frac{2\pi k}{n}x + 1). \end{aligned}$$

若n偶数，$z_k \in \mathbb{R}$当且仅当$k = 0, n/2$，且当$k \neq 0, n/2$时，$\bar{z_k} = z_{n-k}$。从而

$$\begin{aligned} f(x) &= (x-z_0)(x-z_{n/2})\prod_{k=1}^{\frac{n}{2}-1}(x^2 - (z_k+z_{n-k})x + z_kz_{n-k}) \\ &= (x-1)(x+1)\prod_{k=1}^{\frac{n}{2}-1}(x^2 - 2\mathrm{Re}(z_k)x + |z_k|^2) \\ &= (x-1)(x+1)\prod_{k=1}^{\frac{n}{2}-1}(x^2 - 2\cos\frac{2\pi k}{n}x + 1). \end{aligned}$$

□

2.5 重因式

本小节的主要内容是通过引入多项式的导数算子，帮助计算多项式中不可约因子的重数。

定义2.5.1. 设$\mathbb{F}$是数域，定义$\mathbb{F}[x]$上的导数算子$(\ )' : \mathbb{F}[x] \to \mathbb{F}[x]$满足以下条件的唯一映射:

1. $(x)' = 1$;
2. （线性1）$\forall f, g \in \mathbb{F}[x]$，都有$(f+g)' = (f)' + (g)'$;
3. （线性2）$\forall f \in \mathbb{F}[x], c \in \mathbb{F}$，都有$(cf)' = c(f)'$;
4. （Leibniz法则）$\forall f, g \in \mathbb{F}[x]$，都有$(fg)' = f(g)' + (f)'g$。

命题2.5.2. 满足定义*2.5.1*的映射存在，且$\forall f(x) = a_0 + a_1 x + \cdots + a_n x^n$，有

$$(f)' = a_1 + 2a_2 x + \cdots + n a_n x^{n-1}$$

证明. 首先假设满足定义2.5.1的映射存在，由线性性知仅需证明$(x^n)' = nx^{n-1}$。

当$n = 0$时，$(1)' = (1 \times 1)' = 1 \times (1)' + (1)' \times 1 = 2(1)'$，从而$(1)' = 0$。
对于当$n > 1$时，$(x^n)' = x(x^{n-1})' + (x')x^{n-1} = x(x^{n-1})' + x^{n-1}$。通过数学归纳法易知$(x^n)' = nx^{n-1}$。
通过直接计算容易验证该映射满足定义2.5.1中的四个条件。 □

注2.5.3. 我们可归纳定义k阶导数算子$f^{(k)}(x) := (f^{(k-1)}(x))'$。

记号$f^{(k)}(x)$表示对$f(x)$进行k次求导计算，而记号$f^k(x) := \underbrace{f(x) \times \cdots \times f(x)}_{k\text{次连乘}}$表示多项式的$k$次方。

命题2.5.4. 设$f(x) \in \mathbb{F}[x], m \in \mathbb{N}$，则

$$(f^m(x))' = m f'(x) f^{m-1}(x).$$

证明. 对m进行归纳。当$m = 1$时命题显然成立（此处$f^0(x) := x$）。假设$m = k-1$时命题成立，当$m = k$时有

$$\begin{aligned}
(f^k(x))' &= (f(x)f^{k-1}(x))' \\
&= f(x)(f^{k-1}(x))' + f'(x)f^{k-1}(x) \quad \text{(Leibniz)} \\
&= f(x)(k-1)f'(x)f^{k-2}(x) + f'(x)f^{k-1}(x) \quad \text{(归纳假设)} \\
&= k f'(x) f^{k-1}(x).
\end{aligned}$$

因而由归纳假设知命题成立。 □

定理2.5.5. 设$f(x), p(x) \in \mathbb{F}[x]$。假设$p(x)$不可约且$p \mid f$，则

$$e(p|f') = e(p|f) - 1.$$

证明. 令$e = e(p|f)$，并设$f(x) = p^e(x)g(x)$。根据定义有$(p(x), g(x)) = 1$。从而

$$\begin{aligned}
f'(x) &= p^e(x)g'(x) + (p^e(x))'g(x) \\
&= p^e(x)g'(x) + ep^{e-1}(x)p'(x)g(x) \\
&= p^{e-1}(x)(p(x)g'(x) + ep'(x)g(x))
\end{aligned}$$

注意到$\deg(p'(x)) < \deg(p(x))$，则$(p(x), p'(x)) = 1 \Rightarrow (p(x), ep'(x)g(x)) = 1 \Rightarrow (p, pg' + ep'g) = 1$。因此有$e(p|f') = e(p|f) - 1$。

□

推论2.5.6. 设$f(x), p(x) \in \mathbb{F}[x]$。假设$p(x)$不可约且$p \mid f$，则对于$1 \le k \le e := e(p|f)$ 有

$$e(p|f^{(k)}) = e - k$$

特别地，$p(x) \mid f^{(e-1)}(x)$但$(p(x), f^{(e)}(x)) = 1$。

证明. 这是定理2.5.5的直接推论。 □

推论2.5.7. 设$f(x)\in\mathbb{F}[x]_{\deg\geq 1}$，设$f(x)$具有标准分解（$p_i$互不相同）

$$f(x)=l(f)p_1^{e_1}(x)\cdots p_m^{e_m}(x)$$

则

$$(f(x),f'(x))=p_1^{e_1-1}(x)\cdots p_m^{e_m-1}(x)$$

特别地，$f(x)$没有重因式当且仅当$(f(x),f'(x))=1$。

证明. 注意到对于$i=1,\ldots,m$，有$e(p_i|(f,f'))=\min\{e(p_i|f),e(p_i|f)\}=\min\{e_i,e_i-1\}=e_i-1$。 □

注2.5.8. 设$f(x)\in\mathbb{F}[x]_{\deg\geq 1}$，则$\frac{f(x)}{(f(x),f'(x))}$和$f(x)$具有相同的不可约因子，但每个不可约因子$p_i(x)$的重数都是1。有时候这个方法可以在求$f(x)$因式分解的时候降低次数从而减少求解难度。

例2.5.9. 求$f(x)=x^5-15x^3+10x^2+60x-72$在$\mathbb{R}$上的标准分解。

证明. 首先，$f'(x)=5x^4-45x^2+20x+60$。使用辗转相除法

$$\begin{array}{r|l}
5x^4-45x^2+20x+60 & \frac{1}{5}x \\
\hline
x^5-15x^3+10x^2+60x-72 & \\
-x^5+9x^3-4x^2-12x & \\
\hline
-6x^3+6x^2+48x-72 &
\end{array}$$

$$\begin{array}{r|l}
x^3-x^2-8x+12 & 5x+5 \\
\hline
5x^4-45x^2+20x+60 & \\
-5x^4+5x^3+40x^2-60x & \\
\hline
5x^3-5x^2-40x+60 & \\
-5x^3+5x^2+40x-60 & \\
\hline
0 &
\end{array}$$

即

$$x^5-15x^3+10x^2+60x-72=\frac{1}{5}x(5x^4-45x^2+20x+60)+(-6x^3+6x^2+48x-72)$$

且

$$(-6x^3+6x^2+48x-72)\mid(5x^4-45x^2+20x+60)$$

再次利用带余除法得$\frac{f}{(f,f')}=x^2+x-6=(x-2)(x+3)$。从而$f(x)$具有形式$(x-2)^{e_1}(x+3)^{e_2}$满足

$$\begin{cases} e_1+e_2=5 \\ (-2)^{e_1}3^{e_2}=-72 \end{cases}$$

此方程的正整数解仅有$e_1=3,\ e_2=2$。即$f(x)=(x-2)^3(x+3)^2$。 □

2.6 Lagrange插值公式

我们知道两点确定唯一的一条直线（即一次函数），三点确定唯一的一条二次曲线，因此我们有理由推断，平面上一般的$n+1$个点确定唯一的一条n次曲线。这就是Lagrange插值公式。

为此，我们先对特殊情形进行构造。

引理2.6.1. 设$a_1,\ldots,a_n,b\in\mathbb{F}$是数域$\mathbb{F}$上$n+1$个互不相同的数，则存在多项式$f(x)\in\mathbb{F}$使得

1. $\forall i=1,\ldots,n,f(a_i)=0$;
2. $f(b)=1$。

证明. 取$f(x)=\prod_{k=1}^{n}\frac{x-a_i}{b-a_i}$即可。 □

命题2.6.2. （*Lagrange*插值公式）设$a_0,\dots,a_n,b_0\dots,b_n\in\mathbb{F}$且$a_0,\dots,a_n$互不相同，则存在唯一的不超过$n$次多项式$f(x)\in\mathbb{F}[x]_{\deg\le n}$使得对于$i=0,\dots,n$有$f(a_i)=b_i$。

证明. 存在性。对于$i=0,\dots,n$，由引理2.6.1取$f_i(x)\in\mathbb{F}[x]$使得对于$j\in\{0,\dots,n\}\backslash\{i\}$，有$f_i(a_j)=0$且$f_i(a_i)=1$。令$f(x)=b_0f_0(x)+b_1f_1(x)+\cdots+b_nf_n(x)$。则对于$i=0,\dots,n$，有

$$\begin{aligned}f(a_i)&=b_0f(a_i)+\cdots+b_nf_n(a_i)\\&=b_0\times0+\cdots+b_{i-1}\times0+b_i\times1+b_{i+1}\times0+\cdots+b_n\times0\\&=b_i.\end{aligned}$$

唯一性。设$g(x)$是另一个满足条件的多项式。则对于$i=0,\dots,n$，有$f(a_i)-g(a_i)=b_i-b_i=0$，因此$(x-a_i)\mid(f-g)$。由推论2.3.4知$(x-a_0)\cdots(x-a_n)\mid(f-g)$。 □

2.7 有理系数多项式

本小节的主要内容是关于有理系数多项式（即$\mathbb{Q}[x]$）分解的讨论。不同于复（实）系数，有理系数的多项式分解十分困难，因此我们仅能得到相对有限的结果。

设$a_1,\dots,a_n\in\mathbb{Z}$，令$(a_1,\dots,a_n)\in\mathbb{N}$表示$a_1,\dots,a_n$的最大公约数，令$[a_1,\dots,a_n]\in\mathbb{N}$表示$a_1,\dots,a_n$的最小公倍数。

定义2.7.1. 设$f(x)=a_0+a_1x+\cdots+a_nx^n\in\mathbb{Z}[x]$ 满足$(a_0\dots,a_n)=1$，则称$f(x)$是本原多项式。

命题2.7.2. 设$f(x)\in\mathbb{Q}[x]$，则存在（差$\pm$号的意义下）唯一的本原多项式$g(x)$与$f(x)$相伴。

证明. 设$f(x)=\frac{a_0}{b_0}+\frac{a_1}{b_1}x+\cdots+\frac{a_n}{b_n}x^n$，其中$a_i,b_i\in\mathbb{Z}$且$(a_i,b_i)=1$。令$a=(a_0,\dots,a_n)$并令$b=[b_0,\dots,b_n]$。则

$$f(x)\sim g(x):=\frac{b}{a}\left(\frac{a_0}{b_0}+\frac{a_1}{b_1}x+\cdots+\frac{a_n}{b_n}x^n\right)\in\mathbb{Z}[x]$$

（注意到$m=(\frac{b}{b_0}\frac{a_0}{a},\frac{b}{b_1}\frac{a_1}{a},\dots,\frac{b}{b_n}\frac{a_n}{a})=1$。事实上对任意的素数$p$，有$e(p|a)=\min\{e(p|a_i)|i=0,\dots,n\}$且$e(p|b)=\max\{e(p|b_i)|i=0,\dots,n\}$。若$p|m$，考虑$i\in\{0,\dots,n\}$使得$e(p|a_i)=e(p|a)$，从而$p\mid\frac{b}{b_i}$。再考虑$j\in\{0,\dots,n\}$使得$e(p|b_j)=e(p|b)$。此时$p\nmid\frac{b}{b_j}$，因此必有$p\mid b_j$。但由于$(a_j,b_j)=1$，因此$p\nmid a_j$从而$p\nmid\frac{b}{b_j}\frac{a_j}{a}$，矛盾！）

因此$g(x)$是本原多项式。若另有本原多项式$h(x)=c_0+c_1x+\cdots+c_nx^n\in\mathbb{Z}[x]\sim g(x)$。根据定义$\exists\alpha,\beta\in\mathbb{Z}$使得$g(x)=\frac{\alpha}{\beta}h(x)$。从而$\beta\mid(\alpha c_1,\dots,\alpha c_n)=\alpha$。对称地亦有$\alpha\mid\beta$。因此$\beta=\pm\alpha$。 □

命题2.7.3. 设$f(x),g(x)\in\mathbb{Z}[x]$是本原多项式，则$f(x)g(x)$也是本原多项式。

证明. 设$f(x)=a_0+\cdots+a_mx^m,g(x)=b_0+\cdots+b_nx^n$。对任意素数$p$，设$i\in\mathbb{N}$是最小的自然数满足$p\nmid a_i$，设$j\in\mathbb{N}$是最小的自然数满足$p\nmid b_j$。考虑$f(x)g(x)$中$x^{i+j}$项的系数为$c_{i+j}=a_0b_{i+j}+\cdots a_{i+j}b_0$，其中仅有$a_ib_j$不被$p$整除，其余项根据定义都是$p$的倍数。因此$p\nmid c_{i+j}$，即$f(x)g(x)$是本原多项式。 □

推论2.7.4. 设$f(x)\in\mathbb{Z}[x]$，则$f(x)$在$\mathbb{Z}$上不可约当且仅当$\mathbb{Q}$不可约。

证明. 若$f(x)$在$\mathbb{Q}$上不可约，显然$f(x)$在$\mathbb{Z}$上也不可约。

另一方面，假设$f(x)$在$\mathbb{Q}$上可约，即存在$g(x),h(x)\in\mathbb{Q}[x]$使得$f(x)=g(x)h(x)$。设$g_0(x),h_0(x)\in\mathbb{Z}[x]$是本原多项式使得$g(x)\sim g_0(x),h(x)\sim h_0(x)$。由于$f(x)\sim g_0(x)h_0(x)$，则存在$c\in\mathbb{Q}$使得$f=cg_0h_0$。由命题2.7.3知$g_0h_0$是本原多项式，因此$c\in\mathbb{Z}$（否则若$c$有非平凡分母$b$，则$b$整除$g_0h_0$的各项系数，与$g_0h_0$是本原多项式矛盾！）。 □

推论2.7.5. （整系数多项式的标准分解）设$f(x)\in\mathbb{Z}[x]$，则（差$\pm$号的意义下）存在唯一的分解

$$f(x)=cp_1(x)^{e_1}\cdots p_m(x)^{e_m}$$

其中$c\in\mathbb{Z}$使得f/c是本原多项式，$p_1,\dots,p_m$是互不相伴的不可约本原多项式。

证明. 设$f(x)$在$\mathbb{Q}$上有标准分解$f(x)=l(f)q_1^{e_1}\cdots q_m^{e_m}$。设$p_i(x)\in\mathbb{Z}[x]$是与$q_i(x)$相伴的本原不可约多项式，因此

$$f(x)\sim p_1^{e_1}\cdots p_m^{e_m}$$

即存在$c\in\mathbb{Z}$使得$f(x)=cp_1(x)^{e_1}\cdots p_m(x)^{e_m}$。 □

事实上，对于一个一般的整系数多项式$f(x)\in\mathbb{Z}[x]$，要判断$f(x)$是否不可约十分困难。下面给出一个判断$f(x)$不可约的充分条件。

命题2.7.6. （*Eisenstein*判别法）设$f(x)=a_0+a_1x+\cdots+a_nx^n\in\mathbb{Z}[x]$。若存在素数$p$使得

1. $p\nmid a_n$,
2. $p\mid(a_0,\ldots,a_{n-1})$,
3. $p^2\nmid a_0$。

则$f(x)$不可约。

证明. 假设不然，设存在$g(x)=b_0+\cdots+b_sx^s,h(x)=c_0+\cdots+c_tx^t\in\mathbb{Z}[x]$满足$f=gh$且$s,t<n$。由于$p||a_0=b_0c_0$，不妨设$p|b_0$但$p\nmid c_0$。若一切的$b_i$都有$p|b_i$，则$p|a_n=b_sc_t$，矛盾！从而存在最小的$k\in\mathbb{N}$使得$p\nmid b_k$。从而$p\nmid a_k=b_kc_0+b_{k-1}c_1+\cdots+b_0c_k$，矛盾！因此$f(x)$不可约。 □

有时候对于多项式的不可约判别无法直接使用Essentein判别法，需要我们借助平移变量的技巧。

例2.7.7. 设p是素数，证明多项式

$$f(x)=x^{p-1}+x^{p-2}+\cdots+x+1$$

在$\mathbb{Q}$上不可约。

证明. 作变量替换$y=x-1$，则

$$\begin{aligned}f(x)&=\frac{x^p-1}{x-1}\\&=\frac{(y+1)^p-1}{y}\\&=y^{p-1}+\sum_{k=1}^{p-1}C_p^ky^{k-1}.\end{aligned}$$

注意到当$1\le k\le p-1$时，$p|k$且$p^2\nmid C_p^1=p$，由Essentein知$f(x)$不可约。 □

本节的最后给出整系数多项式在有理数域上含有一次不可约因式的必要条件，也就是存在有理根的必要条件。

命题2.7.8. 设$f(x)=a_0+\cdots+a_nx^n$，则有理数$\frac{a}{b}$（其中$(a,b)=1$）是$f(x)$的根的必要条件是$b|a_n$且$a|a_0$。

证明. 假设$\frac{a}{b}$是$f(x)$的根，则

$$\begin{aligned}0&=a_0+a_1(\frac{a}{b})+\cdots+a_n(\frac{a}{b})^n\\&=\frac{1}{b^n}\left(a_0b^n+a_1ab^{n-1}+\cdots+a_na^n\right)\end{aligned}$$

从而

$$b\mid a_0b^n+a_1ab^{n-1}+\cdots+a_na^n\Rightarrow b\mid a_na^n\Rightarrow b\mid a_n$$

且

$$a\mid a_0b^n+a_1ab^{n-1}+\cdots+a_na^n\Rightarrow a\mid a_0b^n\Rightarrow a\mid a_0.$$

□

2.8 习题

习题2.8.1. 证明$\mathbb{Q}$是最小的数域，即$\forall A \subseteq \mathbb{C}$是一个域，有$\mathbb{Q} \subseteq A$。

习题2.8.2. 设$A := \{0, 1\}$，定义

- $\forall a \in A, 0 + a = a$且$1 + 1 = 0$;
- $\forall a \in A, 0 \times a = 0$且$1 \times 1 = 1$;
- $+$和$\times$都满足交换律。

验证$(A, +, \times)$构成一个域（提示：需要验证$+$和$\times$都满足结合律，找到各自的零元，验证逆元存在，验证分配律）。

这是一个最简单的域非数域的例子。

习题2.8.3. 设$f(x) = x^4 + 1, g(x) = x^2 + x + 1$，求$g(x)$除$f(x)$的商和余式。

习题2.8.4. 设$a \neq b \in \mathbb{F}$，证明$f(x)$除$(x - a)(x - b)$的余式

$$r(x) = \frac{f(b) - f(a)}{b - a}x + \frac{bf(a) - af(b)}{b - a}$$

(提示：计算$f(a) - r(a)$以及$f(b) - r(b)$)。

习题2.8.5. 计算下列多项式的最大公因式(f, g)，并表达成$uf + vg$的形式。

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

习题2.8.6. 试在$\mathbb{R}$和$\mathbb{C}$上写出下列多项式的标准分解

1. $f(x) = x^3 - 1$;
2. $f(x) = x^4 - 1$;
3. $f(x) = x^n + 1$。

习题2.8.7. 设$f(x), g(x) \in \mathbb{F}[x]$。证明注2.3.12中的推论。以此证明结论：

1. 对任意$n \in \mathbb{N}$，都有$(f, g)^n = (f^n, g^n)$。
2. * 令$[f, g] = \frac{fg}{(f,g)}$，有$[f, g]$同时整除$f, g$且如果另有$h$同时整除$f, g$必有$[f, g] \mid h$（称$[f, g]$为$f, g$的**最小公倍式**）。（提示：$m + n = \min\{m, n\} + \max\{m, n\}$）

习题2.8.8. 判断下列有理数域上的多项式有无重因式。如果有重因式，试求出一个多项式与它有完全相同的不可约因式（不计重数），且这个多项式没有重因式。

1. $x^4 + 4x^2 - 4x - 3$;
2. $x^5 - 5x^4 + 7x^3 - 2x^2 + 4x - 8$。

习题2.8.9. 设$f(x) \in \mathbb{F}[x]$，若存在$p(x) \backsim q(x) \in \mathbb{F}[x]$不可约使得$p, q$都是$f(x)$的不可约因子，求证：

$$f'(x) \nmid f(x).$$

（提示：试证明$\deg(f') > \deg((f, f'))$。）

习题2.8.10. 判断下列整系数多项式在$\mathbb{Q}$上是否不可约：

1. $f(x) = x^6 - 6x^3 + 9x^2 - 24x + 15$;
2. $g(x) = x^6 + x^3 + 1$;
3. 设p是一个奇素数，$h(x) = x^p - px + 1$。

习题2.8.11. 求下列多项式的有理根

1. $x^3 - 6x^2 + 15x - 14$;
2. $2x^3 + x^2 - 3x + 1$。

2.9 再议数系*

以发明自然数作为起点，随着对运算不断探索的过程中，人类不断地扩展对数的理解和定义。简单来说，有下面这样的扩张过程

$$\mathbb{N} \subseteq \mathbb{Z} \subseteq \mathbb{Q} \subseteq \mathbb{R} \subseteq \mathbb{C}$$

借助目前为止学过的知识，我们可以通过等价关系的语言，以自然数作为基础，层层递进进行严格的构造。

1. $\mathbb{N} \Rightarrow \mathbb{Z}$（加法逆运算的封闭性）

为了在$\mathbb{N}$上加法能进行良好的逆运算（即减法），我们可以用一对自然数(a,b)来表示它们的差$a-b$，从而通过等价关系的方式定义出整数集$\mathbb{Z}$。

定义2.9.1. 规定$\mathbb{Z} := \mathbb{N} \times \mathbb{N}/\sim$，其中$(a,b) \sim (c,d)$当且仅当$a+d=b+c$，并且在$\mathbb{N} \times \mathbb{N}$上规定二元运算$+$以及$\times$:

- $(a,b)+(c,d) := (a+c, b+d)$;
- $(a,b) \times (c,d) := (ac+bd, ad+bc)$。

可通过该等价关系良好地继承到$\mathbb{Z}$上。

注2.9.2. 对于一个的集合A以及其上的等价关系$\sim$和二元运算$*: A \times A \to A$，我们说**此二元运算可通过等价关系$\sim$继承到商集$A/\sim$上**当且仅当

- 对任意$a_1,b_1,a_2,b_2,c,d \in A$满足$(a_1,b_1) \sim (a_2,b_2)$，都有$(a_1,b_1)*(c,d) \sim (a_2,b_2)*(c,d)$;
- 对任意$a,b,c_1,d_1,c_2,d_2 \in A$满足$(c_1,d_1) \sim (c_2,d_2)$，都有$(a,b)*(c_1,d_1) \sim (a,b)*(c_2,d_2)$。

此时在$A/\sim$上继承了的二元运算$*: [(a,b)]*[(c,d)] := [(a,b)*(c,d)]$。

由于表达式的对称性，要验证$\mathbb{Z}$上继承了$+$以及$\times$，只需验证以上第一个条件:

- $+$运算：设$a_1,b_1,a_2,b_2,c,d \in \mathbb{N}$满足$a_1+b_2=b_1+a_2$，有

$$(a_1+c)+(b_2+d) = a_1+b_2+c+d = b_1+a_2+c+d = (b_1+d)+(a_2+c)$$

从而

$$(a_1,b_1)+(c,d) = (a_1+c, b_1+d) \sim (a_2+c, b_2+d) = (a_2,b_2)+(c,d)$$

- $\times$运算：设$a_1,b_1,a_2,b_2,c,d \in \mathbb{N}$满足$a_1+b_2=b_1+a_2$，有

$$(a_1c+b_1d)+(a_2d+b_2c) = (a_1+b_2)c+(a_2+b_1)d \sim (a_2+b_1)c+(a_1+b_2)d = (a_1d+b_1c)+(a_2c+b_2d,)$$

从而

$$(a_1,b_1) \times (c,d) = (a_1c+b_1d, a_1d+b_1c) \sim (a_2c+b_2d, a_2d+b_2c) = (a_2,b_2) \times (c,d).$$

此时$\mathbb{Z}$上关于加法$+$有零元$(0,0)$，且对任意$[(a,b)] \in \mathbb{Z}$有逆元$[(b,a)]$。

2. $\mathbb{Z} \Rightarrow \mathbb{Q}$（乘法逆运算的封闭性）

虽然在$\mathbb{Z}$上有了加法$+$和乘法$\times$，并且加法具有逆运算，但乘法却不能进行良好的逆运算（即除法）。因此我们可以用一对整数(a,b)来表示它们的商$\frac{a}{b}$，从而通过等价关系的方式定义出有理数域$\mathbb{Q}$。

定义2.9.3. 规定$\mathbb{Q} := \mathbb{Z} \times \mathbb{Z}/\sim$，其中$(a,b) \sim (c,d)$当且仅当$ad=bc$，并且在$\mathbb{Z} \times \mathbb{Z}$上规定二元运算$+$以及$\times$:

- $(a,b)+(c,d):=(ad+bc,bd)$;
- $(a,b)\times(c,d):=(ac,bd)$。

可通过该等价关系良好地继承到$\mathbb{Z}$上（留给读者验证，参考例1.2.9）。

3. $\mathbb{Q}\Rightarrow\mathbb{R}$（极限运算的封闭性）

现在有理数域$\mathbb{Q}$上有了良好定义的四则运算（事实上这是包含$\mathbb{N}$的最小域）。另一方面，$\mathbb{Q}$上的绝对值函数$|\cdot|:\mathbb{Q}\to\mathbb{Q}$能够在$\mathbb{Q}$上定义极限运算。但遗憾的是，极限运算在$\mathbb{Q}$上无法封闭。由一定的数学直观上出发，我们知道$\mathbb{Q}$上的Cauchy序列都应该存在极限。首先我们复习一下Cauchy序列（在$\mathbb{Q}$上）的定义：

定义2.9.4. 设$\mathbf{a}=(a_1,\dots,a_n,\dots)\in\mathbb{Q}^{\mathbb{N}}$是有理数序列，称$\mathbf{a}$是**Cauchy**的当且仅当对任意$M\in\mathbb{N}^*$，都存在$N\in\mathbb{N}$使得对任意$n_1,n_2>N$时，有$|a_{n_1}-a_{n_2}|<1/M$。

我们可以通过考虑$\mathbb{Q}$上全体的$Cauchy$序列，来表示一切实数。但不同的$Cauchy$列有可能表示相同的实数，因此需要再次借助等价关系的语言来定义实数。

定义2.9.5. 规定$\mathbb{R}:=\{\mathbf{a}=(a_1,\dots,a_n,\dots)\in\mathbb{Q}^{\mathbb{N}}\mid\mathbf{a}是Cauchy列\}/\sim$，其中$\mathbf{a}\sim\mathbf{b}$当且仅当$\mathbf{a}-\mathbf{b}$的极限存在且为0（$\mathbb{R}$上的加法+和乘法×由序列的加法+和乘法×自然定义）。

4. $\mathbb{R}\Rightarrow\mathbb{C}$（代数扩张的封闭性）

对于一个一般的域$\mathbb{F}$，考虑其上的多项式$f(x)\in\mathbb{F}[x]$，如果无法在$\mathbb{F}$上找到一个元素$a\in\mathbb{F}$使得$f(a)=0$，则人们自然的考虑如何讲域$\mathbb{F}$进行适当扩张使得$f(x)$有零点（一个典型的例子是历史上人们在实数域$\mathbb{R}$添加x^2+1的根从而得到复数域$\mathbb{C}$）。下面我们讲这个过程一般化，从而帮助我们从另一个角度再次理解复数。

定义2.9.6. 设$\mathbb{F}$是一个域，$f(x)\in\mathbb{F}[x]$是一个多项式。在$\mathbb{F}[x]$上定义等价关系$\sim$：对任意$g,h\in\mathbb{F}[x]$，$g\sim h$当且仅当$f\mid(g-h)$，其商集记为$\mathbb{F}[x]/(f(x))$。其上的加法和乘法继承自$\mathbb{F}[x]$上自然的加法和乘法。

命题2.9.7. 设$\mathbb{F}$是一个域，$f(x)\in\mathbb{F}[x]$是一个多项式。则$\mathbb{F}[x]/(f(x))$构成一个域当且仅当$f(x)$在$\mathbb{F}$上不可约。

证明. 设$f(x)$可约，即存在$g,h\in\mathbb{F}[x]$满足$gh=f$且$\deg(g),\deg(h)<\deg(f)$。根据定义$[g][h]=[f]=[0]$ 但$[g],[h]\neq[0]$。从而$[g],[h]$没有乘法逆元，进而$\mathbb{F}[x]/(f(x))$不构成域。

设$f(x)$不可约。要验证$f(x)$是域仅需验证$\forall g(x)\in\mathbb{F}[x]$满足$[g(x)]\neq[0]$，即$f\nmid g$，则$[g(x)]$在$\mathbb{F}[x]/(f(x))$存在乘法逆元（其余条件平凡）。由于$f(x)$不可约，则存在$u,v\in\mathbb{F}[x]$使得$uf+vg=1$。由定义即为$[v][g]=[1-uf]=[1]$，从而$[v]$是$[g]$的乘法逆元。 □

3 线性代数导引

从本章节开始，整份讲义的所有内容将要围绕线性代数展开。因此我们花一些篇幅，从总体的视角向读者介绍线性代数所研究的数学对象以及关心的基础问题。由于所涉及的内容相对抽象，笔者尽可能地结合具体的例子进行讲解。

在本讲义的后续内容中，除非特别说明，固定$\mathbb{F}$为一个（数）域。

3.1 线性空间

线性代数所研究的核心对象称为线性空间。简单来说，线性空间是对三维空间中向量的相关运算进行抽象得到的定义。在实三维空间$\mathbb{R}^3$上，一个向量$\mathbf{x}$可由一个三元实数组(a,b,c)表示。两个向量$\mathbf{x}_1=(a_1,b_1,c_1),\mathbf{x}_2=(a_2,b_2,c_2)$之间可以进行加法运算$\mathbf{x}_1+\mathbf{x}_2:=(a_1+a_2,b_1+b_2,c_1+c_2)$。另外，一个实数$k\in\mathbb{R}$和一个向量$\mathbf{x}=(a,b,c)\in\mathbb{R}^3$之间还可以定义数乘运算$k\cdot\mathbf{x}:=(ka,kb,kc)$。对以上运算进行抽象，就得到下面一般的线性空间的定义。

定义3.1.1. 一个三元组$(V,+,\cdot)$，其中

- V是一个集合；
- $+:V\times V\to V$是V上的加法（加法运算）；
- $\cdot:\mathbb{F}\times V\to V$是一个映射（数乘运算），

称为$\mathbb{F}$**上的线性空间**，如果

1. (单位条件) 对任意$x\in V$，有$1\cdot x=x$；
2. （结合律）对任意$a_1,a_2\in\mathbb{F},x\in\mathbb{V}$，有$(a_1a_2)\cdot x=a_1\cdot(a_2\cdot x)$；
3. （分配律1）对任意$a_1,a_2\in\mathbb{F},x\in\mathbb{V}$，有$(a_1+a_2)\cdot x=a_1\cdot x+a_2\cdot x$；
4. （分配律2）对任意$a\in\mathbb{F},x_1,x_2\in\mathbb{V}$，有$a\cdot(x_1+x_2)=a\cdot x_1+a\cdot x_2$；

在通常情况下，加法$+$和数乘$\cdot$常常默认已知而将此线性空间简记为V。

注3.1.2. 设V是$\mathbb{F}$上的线性空间，记$\mathbf{0}\in V$为加法零元，我们有：

- 对任意$a\in\mathbb{F}$，有$a\cdot\mathbf{0}=\mathbf{0}$。

  证明. 注意到$a\cdot\mathbf{0}=a\cdot(\mathbf{0}+\mathbf{0})=a\cdot\mathbf{0}+a\cdot\mathbf{0}$。两边加上$a\cdot\mathbf{0}$的逆元得$a\cdot\mathbf{0}=\mathbf{0}$。 □

- 对任意$x\in V$，有$0\cdot x=\mathbf{0}$。

  证明. 注意到$0\cdot x=(0+0)\cdot x=0\cdot x+0\cdot x$。两边加上$0\cdot x$的逆元得$0\cdot x=\mathbf{0}$。 □

一个最基本，也是最典型，并且贯穿全书的例子，就是以下的n维标准线性空间$\mathbb{F}^n$。

例3.1.3. （n维标准线性空间）$\mathbb{F}^n$与其上标准的加法和数乘运算，即

- $(a_1,\dots,a_n)+(b_1,\dots,b_n)=(a_1+b_1,\dots,a_n+b_n)$；
- $k\cdot(a_1,\dots,a_n)=(ka_1,\dots,ka_n)$，

构成$\mathbb{F}$上的线性空间，也同样记为$\mathbb{F}^n$。

特别地，当$n=1$时，即域$\mathbb{F}$本身也是$\mathbb{F}$上的线性空间。

例3.1.4. （平凡空间）域$\mathbb{F}$上的平凡线性空间$\mathbb{F}^0$上只有一个元素，即零向量$\mathbf{0}$，上面的加法和数乘定义显然。在不产生歧义的时候，常常也把$\mathbb{F}^0$简记为0。

例3.1.5. 多项式集$\mathbb{F}[x]$关于多项式的加法以及$\mathbb{F}$在$\mathbb{F}[x]$上的数乘运算构成$\mathbb{F}$上的线性空间。

例3.1.6. （直和函子）设V_1, V_2是两个$\mathbb{F}$上的线性空间，则其笛卡尔积$V_1 \times V_2$上有自然的加法和数乘运算：

- 对任意的$x_1, y_1 \in V_1, x_2, y_2 \in V_2$，规定$(x_1, x_2) + (y_1, y_2) := (x_1 + x_2, y_1 + y_2)$；
- 对任意$k \in \mathbb{F}, x_1 \in V_1, x_2 \in V_2$，规定$k \cdot (x_1, x_2) := (k \cdot x_1, k \cdot x_2)$，

从而构成一个线性空间，称为V_1与V_2的直和，记为$V_1 \oplus V_2$。

更一般地，对于$\mathbb{F}$上若干个线性空间$V_1, V_2, \ldots, V_n$，可类似地定义其直和$V_1 \oplus \cdots \oplus V_n$。

定义3.1.7. （子空间）设V是$\mathbb{F}$上的一个线性空间，$W \subseteq V$是V的一个子集，如果W对于V上原有的加法和数乘运算封闭，即

- 对任意$x_1, x_2 \in W$，都有$x_1 + x_2 \in W$；
- 对任意$k \in \mathbb{F}, x \in W$，都有$k \cdot x \in W$。

则W和原有的加法和数乘运算一起，构成一个线性空间，称为V的**子空间**。

例3.1.8. （$\mathbb{R}^2$上的子空间）

- 0维子空间$\mathbf{0}$；
- 任取$(a, b) \neq (0, 0) \in \mathbb{R}^2$，根据子空间的定义，$[a, b] := \{(ka, kb) | k \in \mathbb{R}\} \subseteq \mathbb{R}^2$是过原点和点$(a, b)$的直线构成的子空间；
- 对于$\mathbb{R}^2$上一般的一条直线，如果该直线不过原点，则不包含零向量$\mathbf{0}$，从而不是一个子空间；
- 若$W \subseteq \mathbb{R}^2$是$\mathbb{R}^2$上的子空间，并且包含两个不共线的非零向量$\mathbf{x}_1 = (a_1, b_1), \mathbf{x_2} = (a_2, b_2)$，则由平行四边形法则知平面上的任一向量$y$可表为形式$a\mathbf{x}_1 + b\mathbf{x}_2$，从而根据子空间的定义知$W = \mathbb{R}^2$。

例3.1.9. （$\mathbb{R}^3$上的子空间）

- 零维子空间：$\mathbf{0}$；
- 一维子空间：任取$(a, b, c) \in \mathbb{R}^3 \backslash \mathbf{0}$，过该点的直线$[a, b, c] := \{ka, kb, kc) | k \in \mathbb{R}\}$；
- 二维子空间：任取$\mathbf{x}_1, \mathbf{x}_2 \in \mathbb{R}^3$不共线，由$\mathbf{x}_1, \mathbf{x}_2$张成的2维平面$\mathrm{span}(\mathbf{x}_1, \mathbf{x}_2) := \{k_1\mathbf{x}_1 + k_2\mathbf{x}_2 | k_1, k_2 \in \mathbb{R}\}$；
- 三维子空间：$\mathbb{R}^3$。

定义/命题3.1.10. （商空间）设V是$\mathbb{F}$上的线性空间，$W \subseteq V$是子空间。在V上可规定等价关系$\sim_W$：对任意$x_1, x_2 \in V$，规定$x_1 \sim x_2$当且仅当$x_1 - x_2 \in W$。可以验证商集$V/\sim_W$继承V上的加法和数乘运算，构成一个线性空间，称为（V**关于**W**的**）**商空间**，记作V/W。

证明. 我们只需验证加法和数乘的良定义性，至于加法的交换律，零元逆元以及线性空间的四条性质自动继承自V。

设$a_1, a_2, b \in V$满足$a_1 - a_2 \in W$，则显然有$(a_1 + b) - (a_2 + b) \in W$，从而$[a_1 + b] = [a_2 + b]$，即加法良定义。设$k \in \mathbb{F}, a_1, a_2 \in V$满足$a_1 - a_2 \in W$，则$ka_1 - ka_2 = k(a_1 - a_2) \in W$从而$[ka_1] = [ka_2]$，即数乘良定义。 □

例3.1.11. 设$f(x) \in \mathbb{F}[x]$是一个多项式，则$(f) := \{g(x) \in \mathbb{F}[x] \mid f(x) \mid g(x)\}$构成$\mathbb{F}[x]$的一个子空间，而$\mathbb{F}[x]/(f)$构成$\mathbb{F}[x]$的商空间。

例3.1.12. （限制函子）设$\mathbb{F}_1 \subseteq \mathbb{F}_2$是有两个具有自然包含关系的域。设$V$是$\mathbb{F}_2$上的线性空间，则$V$自然地看作是$\mathbb{F}_1$上的线性空间，即数乘运算限制在$\cdot : \mathbb{F}_1 \times V \to V$上，记此线性空间为$V|_{\mathbb{F}_1}$。

特别地，$\mathbb{F}_2$（即$\mathbb{F}_2^1$）可自然看作$\mathbb{F}_1$上的线性空间。例如

1. $\mathbb{C}$可看作$\mathbb{R}$上的线性空间；
2. $\mathbb{R}$可看作$\mathbb{Q}$上的线性空间；
3. $\mathbb{Q}(\sqrt{2})$可看作$\mathbb{Q}$上的线性空间。

3.2 线性映射

当我们考虑具有特定结构（代数结构、几何结构等等）的一类数学对象时，它们之间"合法"的映射应该在某种一下与自身的结构相容（例如，在分析中考虑实数域$\mathbb{R}$上由极限定义出的距离远近的结构，在考虑连续函数$f : \mathbb{R} \to \mathbb{R}$时应该保持，即$\lim_{x\to a} f(x) = f(\lim_{x\to a} x) = f(x)$）。

定义3.2.1. 设V, W是$\mathbb{F}$上的线性空间，映射$f : V \to W$称为一个**线性映射**如果

1. （加法线性）$f(\mathbf{x}_1 + \mathbf{x}_2) = f(\mathbf{x}_1) + f(\mathbf{x}_2)$
2. （数乘线性）$f(a\mathbf{x}) = af(\mathbf{x})$.

即以下两个图表交换：

$$\begin{array}{ccc} V \times V & \xrightarrow{(-)+(-)} & V \\ {\scriptstyle f\times f}\downarrow & & \downarrow{\scriptstyle f} \\ V \times V & \xrightarrow{(-)+(-)} & V \end{array} \qquad \begin{array}{ccc} \mathbb{F} \times V & \xrightarrow{(-)\cdot(-)} & V \\ {\scriptstyle \mathrm{id}\times f}\downarrow & & \downarrow{\scriptstyle f} \\ \mathbb{F} \times V & \xrightarrow{(-)\cdot(-)} & V \end{array}$$

从V到W的一切线性映射构成的集合记为$\mathcal{L}(V, W)$。

例3.2.2. 设V是一个$\mathbb{F}$上的线性空间，对任意$k \in \mathbb{F}$，映射

$$\begin{aligned} k : V &\to V \\ x &\mapsto k \cdot x \end{aligned}$$

是一个线性映射，称为**数乘映射**。

证明. 定义3.2.1中的加法线性即为定义3.1.1中的分配律2，而定义3.2.1中的数乘线性由定义3.1.1中的结合律以及$\mathbb{F}$上乘法的交换律所保证。 □

例3.2.3. 设$(a, b, c) \in \mathbb{R}^3 \backslash \mathbf{0}$,

- $V = \{k(a, b, c) | k \in \mathbb{R}\}$是过$(a, b, c)$的一维子空间，投影映射

$$\begin{aligned} p_V : \mathbb{R}^3 &\to V \\ (u, v, w) &\mapsto \frac{au + bv + cw}{a^2 + b^2 + c^2}(a, b, c) \end{aligned}$$

  是一个线性映射。

- $V^\perp = \{(x, y, z) \in \mathbb{R}^3 | ax + by + cz = 0\}$是与$(a, b, c)$垂直的二维子空间（平面），投影映射

$$\begin{aligned} p_V : \mathbb{R}^3 &\to V \\ (u, v, w) &\mapsto (u, v, w) - \frac{au + bv + cw}{a^2 + b^2 + c^2}(a, b, c) \end{aligned}$$

  是一个线性映射。

例3.2.4. 以下这些映射都不是线性映射

1.
$$f:\mathbb{R}^3\to\mathbb{R}^3$$
$$(x,y,z)\mapsto(e^x,\cos y,\sin z);$$

2.
$$f:\mathbb{R}^3\to\mathbb{R}^3$$
$$(x,y,z)\mapsto(xy,yz,zx);$$

3.
$$f:\mathbb{R}^3\to\mathbb{R}^3$$
$$(x,y,z)\mapsto(x+1,y+1,z+1);$$

例3.2.5. 对于$V=\mathbb{C}$作为$\mathbb{C}$上的线性空间，考虑映射$f:z\mapsto\bar{z}$。此映射具有加法线性：$\overline{z_1+z_2}=\bar{z}_1+\bar{z}_2$，但不具数乘线性$\overline{z_1\cdot z_2}=\bar{z}_1\cdot\bar{z}_2\neq z_1\cdot\bar{z}_2$（若$z_1\notin\mathbb{R}$）。因此$f$并非$V$上的线性映射，但$f$是$V|_{\mathbb{R}}$上的线性映射。

定义3.2.6. 设V,W是$\mathbb{F}$上的线性空间。若存在$f\in\mathcal{L}(V,W)$是一个双射，则称V与W**同构**，记作$V\cong W$，并称f是一个**线性同构**。一切从V到W的线性同构记作$\mathrm{Isom}(V,W)\subseteq\mathcal{L}(V,W)$。

定义3.2.7. 设V是$\mathbb{F}$上的线性空间。如果存在$n\in\mathbb{N}$以及线性映射$f\in\mathcal{L}(\mathbb{F}^n,V)$是满射，则称$V$是一个**有限维的线性空间**。

从下一个章节开始，如无特殊说明，本讲义将有限维线性空间简称为线性空间。

注3.2.8. 对于$\mathbb{F}$上的线性空间V，若存在$n\in\mathbb{N}$使得$\mathrm{Isom}(\mathbb{F}^n,V)\neq\emptyset$，则我们希望称$V$的维数是$n$。但为了使该定义良好，还需要说明对于$m\neq n$，总有$\mathrm{Isom}(\mathbb{F}^n,\mathbb{F}^m)=\emptyset$。很遗憾这个结论需要等到下个章节学完才能得到，因此我们暂时无法给出维数的定义。

线性映射的复合也是线性映射。

命题3.2.9. 设U,V,W是$\mathbb{F}$上的线性空间，并设$g\in\mathcal{L}(U,V),f\in\mathcal{L}(V,W)$，则$f\circ g\in\mathcal{L}(U,W)$。

证明. 设$x_1,x_2\in U$，则
$$\begin{aligned}f\circ g(x_1+x_2)&=f(g(x_1+x_2))\\&=f(g(x_1)+g(x_2))\\&=f(g(x_1))+f(g(x_2))\\&=f\circ g(x_1)+f\circ g(x_2).\end{aligned}$$
另外，设$x\in U,a\in\mathbb{F}$，则
$$\begin{aligned}f\circ g(a\cdot x)&=f(g(a\cdot x))\\&=f(a\cdot g(x))\\&=a\cdot f(g(x))\\&=a\cdot(f\circ g)(x).\end{aligned}$$
□

定义/命题3.2.10. 设V,W是$\mathbb{F}$上的线性空间，并设$f\in\mathcal{L}(U,V)$，记$\ker(f):=\{x\in V|f(x)=0\}$，称为线性映射映射$f$核。我们有

1. $\ker(f)$是V的线性子空间；

2. $\mathrm{Im}(f)$是W的线性子空间;

3. 自然的映射

$$\begin{aligned} F : V/\ker(f) &\to \mathrm{Im}(f) \\ [x] &\mapsto f(x) \end{aligned}$$

是一个线性同构。

证明. 根据定义，F显然是一个满射。假设$x_1, x_2 \in \mathbb{V}$使得$f(x_1) = f(x_2)$，则$\mathbf{0} = f(x_1) - f(x_2) = f(x_1 - x_2)$。从而$x_1 - x_2 \in \ker(f)$，即有$[x_1] = [x_2]$。因此$F$是单射。 □

有趣的是，两个线性空间V, W之间的线性映射组成的集合本身也具有线性结构，因此也构成一个线性空间。

例3.2.11. 设V, W是$\mathbb{F}$上的线性空间，在$\mathcal{L}(V, W)$上可定义自然的加法以及数乘运算

- 对任意$f, g \in \mathcal{L}(V, W)$，规定

$$\begin{aligned} f + g : V &\to W \\ x &\mapsto f(x) + g(x). \end{aligned}$$

- 对任意$k \in \mathbb{F}, f \in \mathcal{L}(V, W)$，规定

$$\begin{aligned} k \cdot f : V &\to W \\ x &\mapsto k \cdot f(x). \end{aligned}$$

则$\mathcal{L}(V, W)$同此加法和数乘一同构成一个线性空间，证明留作习题。

特别地，当$W = \mathbb{F}$时，空间$\mathcal{L}(V, \mathbb{F})$称为$V$的对偶空间，记作$V^\vee$。空间$V^\vee$中的元素称为$V$上的线性函数。

3.3 矩阵及线性方程组

设$f : \mathbb{R}^n \to \mathbb{R}^m \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$，记$e_j := (\underbrace{0, \ldots, 0, 1, 0, \ldots, 0}_{\text{第}j\text{位为}1}) \in \mathbb{R}^n$。若知道$f$在每个$e_j$上的像$f(e_j)$，则对于$\mathbb{R}^n$上的一个一般元素$(x_1, \ldots, x_n)$，由线性映射的性质知

$$\begin{aligned} f((x_1, \ldots, x_n)) &= f(x_1 e_1 + \cdots + x_n e_n) \\ &= f(x_1 e_1) + \cdots + f(x_n e_n) \\ &= x_1 f(e_1) + \cdots + x_n f(e_n). \end{aligned}$$

因此一个线性映射$f : \mathbb{R}^n \to \mathbb{R}^m$完全由$f$在$e_1, \ldots, e_n$上的像所确定。反正，任意指定$e_1, \ldots, e_n$在$\mathbb{R}^m$中的像，也可以唯一确定一个线性映射$f : \mathbb{R}^n \to \mathbb{R}^m$。对于每一个$i = 1, \ldots, n$，设

$$f(e_i) = (a_{1j}, \ldots, a_{mj})$$

则

$$\begin{aligned} f((x_1, \ldots, x_n)) &= x_1 f(e_1) + \cdots + x_j f(e_j) + \cdots + x_n f(e_n) \\ &= \sum_{j=1}^{n} x_j (a_{1j}, \ldots, a_{mj}) \\ &= (\sum_{j=1}^{n} x_j a_{1j}, \ldots, \sum_{j=1}^{n} x_j a_{ij}, \ldots, \sum_{j=1}^{n} x_j a_{mj}). \end{aligned}$$

根据上述讨论，我们自然引出下面关于矩阵的定义

定义3.3.1. 设$m,n\in\mathbb{N}$。对于$i=1,\dots,m,j=1,\dots,n$，指定$a_{ij}\in\mathbb{F}$并排成一个m行n列的数阵A，称为一个$m\times n$**型矩阵**，记作

$$A:=\begin{pmatrix}a_{11}&\cdots&a_{1n}\\ \vdots&&\vdots\\ a_{m1}&\cdots&a_{mn}\end{pmatrix}$$

或简记为$A=(a_{ij})_{m\times n}$。矩阵A中第i行第j列的元素a_{ij}称为第(i,j)-元。我们将所有以$\mathbb{F}$为系数的$m\times n$型矩阵构成的集合记作$\mathrm{Mat}_{m\times n}(\mathbb{F})$。特别地，对于$n\times n$-型矩阵称为$n$阶方阵。

对于一个列向量$\mathbf{x}:=(x_1,\dots,x_n)^T\in\mathbb{F}$，规定矩阵$A$在$\mathbf{x}$上的作用为

$$A\cdot\mathbf{x}=\begin{pmatrix}a_{11}&\cdots&a_{1n}\\ \vdots&&\vdots\\ a_{m1}&\cdots&a_{mn}\end{pmatrix}\cdot\begin{pmatrix}x_1\\ \vdots\\ x_n\end{pmatrix}:=\begin{pmatrix}x_1a_{11}+\cdots+x_na_{1n}\\ \vdots\\ x_1a_{m1}+\cdots+x_na_{mn}\end{pmatrix}$$

由先前的讨论，我们知矩阵作用和标准线性空间之间的线性变换有如下对应：

命题3.3.2. 设$f\in\mathcal{L}(\mathbb{F}^n,\mathbb{F}^m)$，记$e_j:=(\underbrace{0,\dots,0,1,0,\dots,0}_{\text{第}j\text{位为}1})\in\mathbb{F}^n$。设$f(e_j)=(a_{1j},\dots,a_{mj})$，并记$A_f:=(a_{ij})_{m\times n}$，则对任意$\mathbf{x}\in\mathbb{F}^n$，有

$$f(\mathbf{x})^T=A_f\cdot\mathbf{x}^T$$

注3.3.3. 上述命题意味着下面两个集合存在自然的双射

$$\begin{gathered}\mathcal{L}(\mathbb{F}^n,\mathbb{F}^m)\leftrightarrow\mathrm{Mat}_{m\times n}(\mathbb{F})\\ f\mapsto A_f\\ f_A\hookleftarrow A\end{gathered}$$

特别的，想要证明两个矩阵A,B相等，只需要说明对任意$\mathbf{x}\in\mathbb{F}^n$，都有$f_A(\mathbf{x})=f_B(\mathbf{x})$。

定义/命题3.3.4. 设$A=(a_{ij})_{m\times n},B=(b_{ij})_{m\times n}\in\mathrm{Mat}_{m\times n}(\mathbb{F})$，定义两个矩阵的加法为

$$A+B=(a_{ij}+b_{ij})_{m\times n}$$

从而有

$$f_A+f_B=f_{A+B}$$

证明. 对任意$\mathbf{x}=(x_1,\dots,x_n)\in\mathbb{F}^n$，有

$$\begin{aligned}f_A(\mathbf{x})^T+f_B(\mathbf{x})^T&=A\cdot\mathbf{x}^T+B\cdot\mathbf{x}^T\\ &=\begin{pmatrix}x_1a_{11}+\cdots+x_na_{1n}\\ \vdots\\ x_1a_{m1}+\cdots+x_na_{mn}\end{pmatrix}+\begin{pmatrix}x_1b_{11}+\cdots+x_nb_{1n}\\ \vdots\\ x_1b_{m1}+\cdots+x_nb_{mn}\end{pmatrix}\\ &=\begin{pmatrix}x_1(a_{11}+b_{11})+\cdots+x_n(a_{1n}+b_{1n})\\ \vdots\\ x_1(a_{m1}+b_{m1})+\cdots+x_n(a_{mn}+b_mn)\end{pmatrix}\\ &=f_{A+B}(\mathbf{x})^T\end{aligned}$$

□

定义/命题3.3.5. 设$k\in\mathbb{F}$以及$A=(a_{ij})_{m\times n},\in\mathrm{Mat}_{m\times n}(\mathbb{F})$，定义矩阵的数乘为

$$k\cdot A:=(ka_{ij})_{m\times n}$$

从而有

$$f_{k\cdot A}=k\cdot f_A$$

证明. 对任意$\mathbf{x}=(x_1,\ldots,x_n)\in\mathbb{F}^n$，有

$$\begin{aligned}f_{k\cdot A}(\mathbf{x})^T&=(x_1ka_{i1}+\ldots x_nka_{in})_m^T\\&=k(x_1a_{i1}+\ldots x_na_{in})_m^T\\&=k\cdot f_A(\mathbf{x}).\end{aligned}$$

□

定义/命题3.3.6. 设$A=(a_{ij})_{l\times m}\in\mathrm{Mat}_{l\times m}(\mathbb{F}),B=(b_{st})_{m\times n}\in\mathrm{Mat}_{m\times n}(\mathbb{F})$，定义矩阵的乘法为

$$A\times B:=(\sum_{j=1}^{m}a_{ij}b_{jt})_{l\times n}$$

从而

$$f_A\circ f_B=f_{A\times B}$$

证明. 对任意$\mathbf{x}=(x_1,\ldots,x_n)\in\mathbb{F}^n$，有

$$\begin{aligned}f_A\circ f_B(\mathbf{x})&=f_A((y_s:=\sum_{t=1}^{n}x_tb_{st})_m)\\&=(\sum_{j=1}^{n}y_ja_{ij})_l\\&=(\sum_{j=1}^{n}(a_{ij}\sum_{t=1}^{n}x_tb_{jt}))_l\\&=(\sum_{t=1}^{n}x_t(\sum_{j=1}^{m}a_{ij}b_{jt}))_l\\&=f_{A\times B}(\mathbf{x}).\end{aligned}$$

□

注3.3.7. 对于$A\in\mathrm{Mat}_{l\times m}(\mathbb{F}),B\in\mathrm{Mat}_{m\times n}(\mathbb{F})$，可将$A$看作$l$个$m$维行向量竖直排列，即

$$A=\begin{pmatrix}\mathbf{a}_1\\\vdots\\\mathbf{a}_l\end{pmatrix}$$

并将B看作n个m维列向量水平排列，即$B=(\mathbf{b}_1^T,\ldots,\mathbf{b}_n^T)$，则矩阵乘法可定义为

$$A\times B:=(c_{ij}=\mathbf{a}_i\cdot\mathbf{b}_j^T)_{l\times n},$$

其中$\mathbf{a}\cdot\mathbf{b}^T$表示两个向量之间的内积。

注3.3.8. 矩阵的加法和乘法和域上的加法和乘法有许多相似的性质。具体而言，矩阵的加法是一个加法运算，乘法具有零元，同时加法和乘法也具有分配律。即以下法则（对于适当形状的矩阵）

1. 加法零元是$0:=\begin{pmatrix}0&\cdots&0\\\vdots&\ddots&\vdots\\0&\cdots&0\end{pmatrix}$；
2. 矩阵$A=(a_{ij})$ 的加法逆元是$-A=(-a_{ij})$；
3. 对于$n\times n$的方阵，乘法单位是$E_n:=(a_{ij}=\delta_{ij})n\times n$，其中$\delta_{ij}=\begin{cases}1,&i=j\\0,&i\neq j\end{cases}$，即对角线上为1，其余位置为0的矩阵。

4. $A+B=B+A$;
5. $(A+B)+C=A+(B+C)$;
6. $(A\times B)\times C=A\times(B\times C)$;
7. $(A+B)\times C=(A\times C)+(B\times C)$;
8. $A\times(B+C)=(A\times B)+(A\times C)$。

但不同的是，矩阵不一定具有乘法逆元，同时乘法一般也没有交换律。

注3.3.9. 对于一个有n个变量，m个方程组成的一次线性方程组，一般具有形式：

$$\begin{cases} a_{11}x_1+\cdots+a_{1n}x_n+b_1=0 \\ a_{21}x_1+\cdots+a_{2n}x_n+b_2=0 \\ \vdots \\ a_{m1}x_1+\cdots+a_{mn}x_n+b_m=0 \end{cases}$$

其方程组的解$\mathbf{x}=(x_1,\ldots,x_n)$即对应于线性映射$f_A:\mathbb{F}^n\to\mathbb{F}^m$下向量$\mathbf{b}=-(b_1,\ldots,b_m)$的原像，其中$A=(a_{ij})_{m\times n}$。

本小节的内容十分关键，是理解线性映射和矩阵之间关系的重要桥梁。总结而言，有下面这张表格：

	线性映射$f:\mathbb{F}^n\to\mathbb{F}^m$	矩阵$A_f\in\mathrm{Mat}_{m\times n}(\mathbb{F})$
加法	$f+g$	A_f+A_g
数乘	$k\cdot f$	$k\cdot A_f$
复合	$f\circ g$	$A_f\times A_g$
	常数项向量的原像	线性方程组的解

3.4 矩阵的一些计算

例3.4.1. （对角矩阵）

- 设$k_1,k_2,a,b,c,d\in\mathbb{F}$，则

$$\begin{pmatrix}k_1&0\\0&k_2\end{pmatrix}\begin{pmatrix}a&b\\c&d\end{pmatrix}=\begin{pmatrix}k_1a&k_1b\\k_2c&k_2d\end{pmatrix}\text{且}\begin{pmatrix}a&b\\c&d\end{pmatrix}\begin{pmatrix}k_1&0\\0&k_2\end{pmatrix}=\begin{pmatrix}k_1a&k_2b\\k_1c&k_2d\end{pmatrix}$$

- 设$k_1,k_2,k_3,a_{ij}\in\mathbb{F}$，其中$i,j=1,2,3$，则

$$\begin{pmatrix}k_1&0&0\\0&k_2&0\\0&0&k_3\end{pmatrix}\begin{pmatrix}a_{11}&a_{12}&a_{13}\\a_{21}&a_{22}&a_{23}\\a_{31}&a_{32}&a_{33}\end{pmatrix}=\begin{pmatrix}k_1a_{11}&k_1a_{12}&k_1a_{13}\\k_2a_{21}&k_2a_{22}&k_2a_{23}\\k_3a_{31}&k_3a_{32}&k_3a_{33}\end{pmatrix}$$

且

$$\begin{pmatrix}a_{11}&a_{12}&a_{13}\\a_{21}&a_{22}&a_{23}\\a_{31}&a_{32}&a_{33}\end{pmatrix}\begin{pmatrix}k_1&0&0\\0&k_2&0\\0&0&k_3\end{pmatrix}=\begin{pmatrix}k_1a_{11}&k_2a_{12}&k_3a_{13}\\k_1a_{21}&k_2a_{22}&k_3a_{23}\\k_1a_{31}&k_2a_{32}&k_3a_{33}\end{pmatrix}$$

- 一般地，设$k_1,\ldots,k_n\in\mathbb{F}$，记$\mathrm{diag}(k_1,\ldots,k_n)$为第$(i,i)$-元为$k_i$，其他位置为0的矩阵（这样的矩阵称为对角矩阵）。则对任意$A=(a_{ij})_{n\times m}$，有

$$\begin{pmatrix}k_1&0&\cdots&0\\0&k_2&\cdots&0\\\vdots&\vdots&\ddots&\vdots\\0&0&\cdots&k_n\end{pmatrix}\begin{pmatrix}a_{11}&\cdots&a_{1m}\\a_{21}&\cdots&a_{2m}\\\vdots&&\vdots\\a_{n1}&\cdots&a_{nm}\end{pmatrix}=\begin{pmatrix}k_1a_{11}&\cdots&k_1a_{1m}\\k_2a_{21}&\cdots&k_2a_{2m}\\\vdots&&\vdots\\k_na_{n1}&\cdots&k_na_{nm}\end{pmatrix}$$

即将A的第i行元素整体数乘k_i倍。类似地，对任意$B=(b_{ij})_{m\times n}$，有

$$\begin{pmatrix} b_{11} & \cdots & b_{1n} \\ b_{21} & \cdots & b_{2n} \\ \vdots & & \vdots \\ b_{m1} & \cdots & b_{mn} \end{pmatrix}\begin{pmatrix} k_1 & 0 & \cdots & 0 \\ 0 & k_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & k_n \end{pmatrix}=\begin{pmatrix} k_1b_{11} & \cdots & k_nb_{1n} \\ k_1b_{21} & \cdots & k_nb_{2n} \\ \vdots & & \vdots \\ k_1b_{m1} & \cdots & k_nb_{mn} \end{pmatrix}$$

即将B的第i列元素整体数乘k_i倍。

- 特别地，若$k_1=\cdots=k_n=k$，则左乘和右乘$\mathrm{diag}(k,\ldots,k)$都是将矩阵数乘$k$倍。若$k=1$，则保持矩阵不动，因此$E_n=\mathrm{diag}(1,\ldots,1)$是$\mathrm{Mat_{n\times n}}(\mathbb{F})$中的乘法单位。

注3.4.2. 我们记$E_n(i(k)):=\mathrm{diag}(\underbrace{1,\ldots,1,k,1,\ldots,1}_{\text{第}i\text{位为}k})$。

例3.4.3. $(E(i,j))$

- 设$a,b,c,d\in\mathbb{F}$，则

$$\begin{pmatrix}0&1\\1&0\end{pmatrix}\begin{pmatrix}a&b\\c&d\end{pmatrix}=\begin{pmatrix}c&d\\a&b\end{pmatrix}\text{且}\begin{pmatrix}a&b\\c&d\end{pmatrix}\begin{pmatrix}0&1\\1&0\end{pmatrix}=\begin{pmatrix}b&a\\d&c\end{pmatrix}$$

- 设$(a_{ij})_{3\times3}\in\mathrm{Mat}_{3\times3}(\mathbb{F})$，则

1.
$$\begin{pmatrix}0&1&0\\1&0&0\\0&0&1\end{pmatrix}\begin{pmatrix}a_{11}&a_{12}&a_{13}\\a_{21}&a_{22}&a_{23}\\a_{31}&a_{32}&a_{33}\end{pmatrix}=\begin{pmatrix}a_{21}&a_{22}&a_{23}\\a_{11}&a_{12}&a_{13}\\a_{31}&a_{32}&a_{33}\end{pmatrix}$$
且
$$\begin{pmatrix}a_{11}&a_{12}&a_{13}\\a_{21}&a_{22}&a_{23}\\a_{31}&a_{32}&a_{33}\end{pmatrix}\begin{pmatrix}0&1&0\\1&0&0\\0&0&1\end{pmatrix}=\begin{pmatrix}a_{12}&a_{11}&a_{13}\\a_{22}&a_{21}&a_{23}\\a_{32}&a_{31}&a_{33}\end{pmatrix}$$

2.
$$\begin{pmatrix}1&0&0\\0&0&1\\0&1&0\end{pmatrix}\begin{pmatrix}a_{11}&a_{12}&a_{13}\\a_{21}&a_{22}&a_{23}\\a_{31}&a_{32}&a_{33}\end{pmatrix}=\begin{pmatrix}a_{11}&a_{12}&a_{13}\\a_{31}&a_{32}&a_{33}\\a_{21}&a_{22}&a_{23}\end{pmatrix}$$
且
$$\begin{pmatrix}a_{11}&a_{12}&a_{13}\\a_{21}&a_{22}&a_{23}\\a_{31}&a_{32}&a_{33}\end{pmatrix}\begin{pmatrix}1&0&0\\0&0&1\\0&1&0\end{pmatrix}=\begin{pmatrix}a_{11}&a_{13}&a_{12}\\a_{21}&a_{23}&a_{22}\\a_{31}&a_{33}&a_{32}\end{pmatrix}$$

3.
$$\begin{pmatrix}0&0&1\\0&1&0\\1&0&0\end{pmatrix}\begin{pmatrix}a_{11}&a_{12}&a_{13}\\a_{21}&a_{22}&a_{23}\\a_{31}&a_{32}&a_{33}\end{pmatrix}=\begin{pmatrix}a_{31}&a_{32}&a_{33}\\a_{21}&a_{22}&a_{23}\\a_{11}&a_{12}&a_{13}\end{pmatrix}$$
且
$$\begin{pmatrix}a_{11}&a_{12}&a_{13}\\a_{21}&a_{22}&a_{23}\\a_{31}&a_{32}&a_{33}\end{pmatrix}\begin{pmatrix}0&0&1\\0&1&0\\1&0&0\end{pmatrix}=\begin{pmatrix}a_{13}&a_{12}&a_{11}\\a_{23}&a_{22}&a_{21}\\a_{33}&a_{32}&a_{31}\end{pmatrix}$$

- 一般地，记$E_n(i,j)$为将单位矩阵的第i,j两行互换得到矩阵，则对于$A\in\mathrm{Mat}_{n\times m}(\mathbb{F})$，矩阵$E_n(i,j)\times A$是将矩阵$A$的第$i,j$两行互换；对于$B\in\mathrm{Mat}_{m\times n}(\mathbb{F})$，矩阵$B\times E_n(i,j)$是将矩阵$B$的第$i,j$两列互换。

例3.4.4. $E_n(i,j(k))$

- $$\begin{pmatrix}1 & k\\ 0 & 1\end{pmatrix}\begin{pmatrix}a & b\\ c & d\end{pmatrix}=\begin{pmatrix}a+kc & b+kd\\ c & d\end{pmatrix}\text{且}\begin{pmatrix}a & b\\ c & d\end{pmatrix}\begin{pmatrix}1 & k\\ 0 & 1\end{pmatrix}=\begin{pmatrix}a & b+ka\\ c & d+kb\end{pmatrix}$$
- 一般地，记$E_n(i,j)$为将单位矩阵的(i,j)-元修改为k的矩阵。因此可将$E_n(i,j)$看作将单位矩阵的第j行乘k倍加到第i行，同时也可看作将单位矩阵的第i列乘k倍加到第j列。矩阵$E_n(i,j)\times A$是将矩阵A的第j行乘k倍加到第i行；矩阵$B\times E_n(i,j)$是将矩阵B的第i列乘k倍加到第j列；

例3.4.5. （旋转变换）设$\theta\in\mathbb{R}$，记$R_\theta:=\begin{pmatrix}\cos\theta & -\sin\theta\\ \sin\theta & \cos\theta\end{pmatrix}$。对于平面上的一个向量$\mathbf{x}$，设$|\mathbf{x}|=r$且$\mathbf{x}$与$x$轴的夹角为$\alpha$，则$\mathbf{x}=(r\cos\alpha,r\sin\alpha)$。从而

$$R_\theta\cdot\mathbf{x}^T=\begin{pmatrix}\cos\theta & -\sin\theta\\ \sin\theta & \cos\theta\end{pmatrix}\begin{pmatrix}r\cos\alpha\\ r\sin\alpha\end{pmatrix}=\begin{pmatrix}r\cos(\alpha+\theta)\\ r\sin(\alpha+\theta)\end{pmatrix}$$

即R_θ在$\mathbb{R}$上的作用是将平面逆时针旋转θ角。

3.5 习题

习题3.5.1. 设V是$\mathbb{F}$上的线性空间。证明：对任意$x\in V$，有$(-1)\cdot x=-x$。

习题3.5.2. 设V,W是$\mathbb{F}$上的线性空间，验证$\mathcal{L}(V,W)$及其上自然的加法和数乘运算构成一个线性空间。具体来说，需要验证

- 对任意$f,g\in\mathcal{L}(V,W)$，$(f+g)\in\mathcal{L}(V,W)$；
- 对任意$k\in\mathbb{F},f\in\mathcal{L}(V,W)$，$k\cdot f\in\mathcal{L}(V,W)$；
- $+$在$\mathcal{L}(V,W)$是一个加法运算；
- 加法和数乘运算满足定义3.1.1的四条性质。

习题3.5.3. 设V是$\mathbb{F}$上的线性空间，证明自然映射

$$\begin{aligned}F:V&\to(V^\vee)^\vee\\ x&\mapsto(f\mapsto f(x))\end{aligned}$$

是一个线性映射。

习题3.5.4. 运用矩阵和线性映射之间的对应关系，证明矩阵的加法和乘法具有相应的结合律和分配律，即

1. 设$A\in\mathrm{Mat}_{k\times l},B\in\mathrm{Mat}_{l\times m},C\in\mathrm{Mat}_{m\times n}$，则$(A\times B)\times C=A\times(B\times C)$；
2. 设$A\in\mathrm{Mat}_{l\times m},B\in\mathrm{Mat}_{l\times m},C\in\mathrm{Mat}_{m\times n}$，则$(A+B)\times C=(A\times C)+(B\times C)$。

习题3.5.5. 对于$a,b\in\mathbb{F}$，定义矩阵$A_{(a,b)}:=\begin{pmatrix}a & b\\ -b & a\end{pmatrix}$。设$a_1,b_1,a_2,b_2\in\mathbb{F}$，计算$A_{(a_1,b_1)}\times A_{(a_2,b_2)}$。（这个计算告诉我们，映射$\theta:\mathbb{C}\to\mathrm{Mat}_{2\times2}(\mathbb{R}),a+bi\mapsto A_{(a,b)}$是从复数域到矩阵代数的自然嵌入）

习题3.5.6. 设$A\in\mathrm{Mat}_{n\times n}(\mathbb{F})$是$n$阶方阵，对任意$n\in\mathbb{N}$，定义$A^n:=\underbrace{A\times\cdots\times A}_{n\text{次}}$。因此对任意多项式$f(x)=a_0+a_1+\cdots+a_nx^n\in\mathbb{F}[x]$，可考虑表达式$f(A):=a_0E_n+a_1A+\cdots+a_nA^n$。

1. 设$A=\begin{pmatrix}1 & 2\\ 3 & 4\end{pmatrix},f(x)=x^2-5x-10$，计算$f(A)$。
2. 设$A=\begin{pmatrix}0 & 1 & 0\\ 0 & 0 & 1\\ 1 & 0 & 0\end{pmatrix},f(x)=x^3-1$，计算$f(A)$。
3. 对任意$g,h\in\mathbb{F}[x]$，求证：$g(A)h(A)=h(A)g(A)$。

4 线性方程组求解与矩阵的初等变换

4.1 高斯消元法

考虑一个由n个变元，$\mathbb{F}$系数的m个线性方程组成方程组，将其整理称标准形式

$$\begin{cases} a_{11}x_1+\cdots+a_{1n}x_n=b_1 & ① \\ a_{21}x_1+\cdots+a_{2n}x_n=b_2 & ② \\ \cdots & \\ a_{m1}x_1+\cdots+a_{mn}x_n=b_m & ⓜ \end{cases}$$

之后我们循环进行一下3个步骤进行消元：

1. 首先适当交换方程顺序，使得第1个变量x_1在第1个方程中系数不为0。（否则x_1在所有方程中系数都为0，则用x_2替代x_1）。
2. 将第1个方程$\times\frac{1}{a_{11}}$（即将x_1的系数标准化为1）：

$$\begin{cases} x_1+\cdots+a_{1n}x_n=b_1 & ① \\ a_{21}x_1+\cdots+a_{2n}x_n=b_2 & ② \\ \cdots & \\ a_{m1}x_1+\cdots+a_{mn}x_n=b_m & ⓜ \end{cases}$$

3. 对于一切的$k\neq 1$，将第k个方程替换为$ⓚ-a_{k1}\times ①$消去变量x_1：

$$\begin{cases} x_1+ & a_{12x_2}+\cdots+a_{1n}x_n=b_1 & ① \\ & a_{22}x_2+\cdots+a_{2n}x_n=b_2 & ② \\ & \cdots & \\ & a_{m2}x_2+\cdots+a_{mn}x_n=b_m & ⓜ \end{cases}$$

在之后的第i次循环中，我们将上述步骤中的”1”替换成”i”不断化简方程，直到以下两个情况之一发生则停止变换

1. 出现某个方程具有形式$0=b_k$但$b_k\neq 0$，由矛盾说明方程无解。
2. 否则在第k次循环完成后，剩余的方程全都被消去，方程最终化简为形式

$$\begin{cases} x_1+0+\cdots+0+a_{1k+1}x_{k+1}+\cdots+a_{1n}x_n=b_1 & ① \\ x_2+\cdots+0+a_{2k+1}x_{k+1}+\cdots+a_{2n}x_n=b_2 & ② \\ \cdots & \\ x_k+a_{kk+1}x_{k+1}+\cdots+a_{kn}x_n=b_k & ⓜ \end{cases}$$

此时对任意$x_{k+1},\ldots,x_n$在$\mathbb{F}$上的一组取值$c_{k+1},\ldots,c_n$，都对应原方程的一组解

$$\begin{cases} x_1=b_1-(a_{1k+1}c_{k+1}+\cdots+a_{1n}c_n) \\ \vdots \\ x_k=b_k-(a_{kk+1}c_{k+1}+\cdots+a_{kn}c_n) \\ x_{k+1}=c_{k+1} \\ \vdots \\ x_n=c_n \end{cases}$$

例4.1.1. 求解方程

$$\begin{cases} 2x_1+3x_2+5x_3=1 & ① \\ 3x_1-6x_2+4x_3=3 & ② \\ 3x_1+15x_2+11x_3=2 & ③ \end{cases}$$

证明. 1. 第一次循环：$① \Rightarrow 1/2 \times ①, ② \Rightarrow ② - 3 \times ①, ③ \Rightarrow ③ - 3 \times ①$，得到

$$\begin{cases} x_1 + \frac{3}{2}x_2 + \frac{5}{2}x_3 = \frac{1}{2} & ① \\ -\frac{21}{2}x_2 - \frac{7}{2}x_3 = \frac{3}{2} & ② \\ \frac{21}{2}x_2 + \frac{7}{2}x_3 = \frac{1}{2} & ③ \end{cases}$$

2. 第二次循环：$② \Rightarrow -2/21 \times ②, ③ \Rightarrow ③ - 21/2 \times ②$

$$\begin{cases} x_1 + \frac{3}{2}x_2 + \frac{5}{2}x_3 = \frac{1}{2} & ① \\ x_2 + \frac{1}{3}x_3 = -\frac{1}{7} & ② \\ 0 = 2 & ③ \end{cases}$$

因此原方程无解。 □

例4.1.2. 求解线性方程组

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

证明. 1. 第一次循环：$② \Rightarrow ② - 2 \times ①, ③ \Rightarrow ③ - ①, ④ \Rightarrow ② + 2 \times ④$

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

2. 第二次循环：先交换②③两式

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

再令$① \Rightarrow ① + 2 \times ②, ③ \Rightarrow ③ - 3 \times ②$

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

3. 第三次循环：交换③④两式

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

再令$① \Rightarrow ① + 3③, ② \Rightarrow ② + ③, ④ \Rightarrow ④ - 2 \times ③$

$$\begin{cases} x_1 - 9x_4 = 5 & ① \\ x_2 - 4x_4 = 0 & ② \\ -x_3 - 3x_4 = 1 & ③ \\ 0 = 0 & ④ \end{cases}$$

从而对任意$c \in \mathbb{F}$，生成原方程的一组解

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

□

注4.1.3. 事实上，根据矩阵和线性方程组的关系，上述操作都可以用矩阵的语言重新整理。求解线性方程组

$$\begin{cases} a_{11}x_1 + \cdots + a_{1n}x_n = b_1 & ① \\ a_{21}x_1 + \cdots + a_{2n}x_n = b_2 & ② \\ \cdots & \\ a_{m1}x_1 + \cdots + a_{mn}x_n = b_m & ⓜ \end{cases}$$

对应寻找向量$\mathbf{x} = (x_1, \ldots, x_n) \in \mathbb{F}^n$使得

$$A \cdot \mathbf{x}^T = \mathbf{b}^T \quad (\star)$$

其中$A = (a_{ij})_{m\times n}, \mathbf{b} = (b_1, \ldots, b_m)$。

- 交换i, j两行对即为将$(\star)$两边同左乘$E_m(i, j)$;
- 将$(\star)$中第ⓘ式乘k倍即为将$(\star)$两边同左乘$E_m(i(k))$;
- 将$(\star)$中第ⓙ式乘k倍加到第ⓘ式即为将$(\star)$两边同左乘$E_m(i, j(k))$。

例如，在例4.1.1中，进行的操作为依次左乘$E(1(\frac{1}{2})), E(2, 1(-3)), E(3, 1(-3)), E(2(-\frac{2}{21})), E(3, 2(-\frac{21}{2}))$，从而有

$$\begin{aligned} &E(3, 2(-\frac{21}{2}))E(2(-\frac{2}{21}))E(3, 1(-3))E(2, 1(-3))E(1(\frac{1}{2}))\begin{pmatrix} 2 & 3 & 5 \\ 3 & -6 & 4 \\ 3 & 15 & 11 \end{pmatrix} \\ =&E(3, 2(-\frac{21}{2}))E(2(-\frac{2}{21}))E(3, 1(-3))E(2, 1(-3))\begin{pmatrix} 1 & 3/2 & 5/2 \\ 3 & -6 & 4 \\ 3 & 15 & 11 \end{pmatrix} \\ =&E(3, 2(-\frac{21}{2}))E(2(-\frac{2}{21}))\begin{pmatrix} 1 & 3/2 & 5/2 \\ 0 & -21/2 & -7/2 \\ 0 & 21/2 & 7/2 \end{pmatrix} \\ =&\begin{pmatrix} 1 & 3/2 & 5/2 \\ 0 & 1 & 3 \\ 0 & 0 & 0 \end{pmatrix} \end{aligned}$$

另外，由先前的计算可知$E(3, 2(-\frac{21}{2}))E(2(-\frac{2}{21}))E(3, 1(-3))E(2, 1(-3))E(1(\frac{1}{2}))(1, 3, 2)^T = (1/2, -1/7, 2)^T$。

4.2 简化阶梯矩阵

为了描述和运算矩阵的方便，我们需要引入分块矩阵的概念。

设$m = m_1 + \cdots + m_r, n = n_1 + \cdots + n_s \in \mathbb{N}$。对于一个$m \times n$-型矩阵$A \in \mathrm{Mat}_{m\times n}(\mathbb{F})$，在$A$的若干行以及若干列之间用一些水平和垂直的线将其分割为若干小块，使得第(i, j)-位置的小块是一个$m_i \times n_j$-型矩阵，此时我们可将A记为

$$A = \left(\begin{array}{c:c:c} A_{11} & \cdots & A_{1s} \\ \hdashline \vdots & & \vdots \\ \hdashline A_{rs} & \cdots & A_{rs} \end{array}\right)$$

我们此时称A为$(m_1,\dots,m_r)\times(n_1\text{fi}\cdots\text{fi}n_s)$-型分块矩阵，$A$可简记为$(A_{ij})_{r\times s}$。两个分块矩阵之间的计算遵循以下规则

1. 两个具有同样分块形状的矩阵才能进行分块加法。即若$A=(A_{ij})_{r\times s},B=(B_{ij})_{r\times s}$都是$(m_1,\dots,m_r)\times(n_1,\cdots,n_r)$型分块矩阵，则
$$A+B=(A_{ij}+B_{ij})_{r\times s}$$
2. 分块乘法的要求前一个矩阵的列分块等于后一个矩阵的行分块。即设$A=(A_{ij})_{r\times s}$是$(m_1,\dots,m_r)\times(n_1+\cdots+n_s)$-型且$B=(B_{jk})_{s\times t}$是$(n_1+\cdots+n_s)\times(l_1+\cdots+l_t)$-型分块矩阵，则
$$A\times B=(C_{ik}:=\sum_{j=1}^{s}A_{ij}B_{jk}).$$

对于分块矩阵的进一步解释，见附录小节。

有了上一个小节的计算经验，结合上述分块矩阵的语言，我们可以更加清晰地讨论线性方程组的求解。

定义4.2.1. 设$m,n,k\in\mathbb{N}$且$k<\min\{m,n\}$，并设$0<d_1<d_2<\cdots d_k<\min\{m,n\}$,称以下形状的矩阵是$(d_1,\dots,d_k)$-**型阶梯矩阵**
$$\left(\begin{array}{c:c}A&B\\\hline 0&0\end{array}\right)$$
其中（此处$\mathbf{a}_i:=(\underbrace{0,\dots,0,1}_{\text{第}i\text{位为}1},a_{i,d_i+1},\dots,a_{i,d_k})\in\mathbb{R}^{d_k}$）$A=\begin{pmatrix}\mathbf{a}_1\\\mathbf{a}_2\\\vdots\\\mathbf{a}_k\end{pmatrix}$，$B\in\mathrm{Mat}_{k\times(n-d_k)}$。矩阵的$(i,a_i)$-元为1，称为此阶梯矩阵的**主元**。对于$(d_1,\dots,d_k)$-型阶梯矩阵$M$，称线性方程$M\mathbf{x}^T=\mathbf{b}^T$中未知向量$\mathbf{x}=(x_1,\dots,x_n)$的**主变量**为$x_{d_1},\dots,x_{d_k}$，其余变量为**自由变量**。

进一步地，如果阶梯矩阵的主元所在列的其他元素皆为0，称此矩阵为**简化阶梯矩阵**。

在上一个章节中，我们认识了三种操作矩阵行列变换的矩阵$E(i,j),E(i(k)),E(i,j(k))$，我们将其统称为**初等矩阵**，将初等矩阵与其他矩阵相乘，称为**初等变换**。对于两个的矩阵$A,B\in\mathrm{Mat}_{m\times n}(\mathbb{F})$，如果$A$经过有限次的行初等变换得到$B$，即存在$P_1,P_2,\dots,P_k$是初等矩阵，使得
$$B=P_k\cdots P_2P_1A,$$
则称A与B**行相抵**或**行等价**，记作$A\sim_l B$。

命题4.2.2. 对任意$A\neq 0\in\mathrm{Mat}_{m\times n}(\mathbb{F})$，存在唯一的简化阶梯矩阵$B\in\mathrm{Mat}_{m\times n}(\mathbb{F})$使得$A\sim_l B$。

证明. • 存在性。

对行数m进行归纳。当$m=1$时，设$A=(a_1,a_2,\dots,a_n)$，取d_1是最小的整数使得$a_{d_1}\neq 0$，则$\frac{1}{d_1}A=(0,\dots,0,1,a_{d_1+1},\dots,a_n)$是$(d_1)$-型简化阶梯矩阵。假设当$m=r-1$时命题成立，考虑$m=r$时的情形。设$d_1$是最小的整数使得存在$i\in\{1,\dots,r\}$，满足$a_{d_1,i}\neq 0$。将$A$左乘$E_r(1(1/a_{d_1,i}))E_r(1,i)$，得到的矩阵$A_1=E_r(1(1/a_{d_1,i}))E_r(1,i)A$具有形式
$$A_1=\begin{pmatrix}0&\cdots&0&1&a_{1,d_1+1}&\cdots&a_{1,d_n}\\0&\cdots&0&a_{2,d_1}&a_{2,d_1+1}&\cdots&a_{2,d_n}\\\vdots&&\vdots&\vdots&\vdots&&\vdots\\0&\cdots&0&a_{r,d_1}&a_{r,d_1+1}&\cdots&a_{r,d_n}\end{pmatrix}$$

接着将A_1左乘$\prod_{i=2}^{r-1} E_r(i, 1(\frac{-1}{a_{i,d_1}}))$，得到矩阵$B_1$具有形式

$$B_1 = \left(\begin{array}{cccc:ccc} 0 & \cdots & 0 & 1 & a_{1,d_1+1} & \cdots & a_{1,d_n} \\ 0 & \cdots & 0 & 0 & b_{2,d_1+1} & \cdots & b_{2,d_n} \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ 0 & \cdots & 0 & 0 & b_{r,d_1+1} & \cdots & b_{r,d_n} \end{array}\right)$$

矩阵B_1分块的右下角是一个$(r-1)\times(n-d_1)$-型矩阵C_1。根据归纳假设，存在简化阶梯矩阵$B' \in \mathrm{Mat}_{(r-1)\times(n-d_1)}$使得$C_1 \sim_l B'$，即存在初等矩阵$Q_1, \ldots, Q_k$使得

$$B' = Q_1 \cdots Q_k C.$$

令P_i为如下$(1, r-1)\times(1, r-1)$-型分块矩阵

$$P_i = \left(\begin{array}{c:c} 1 & 0 \\ \hdashline 0 & Q_i \end{array}\right)$$

则$P_1 \ldots P_k B_1$具有形式

$$B_2 = \left(\begin{array}{c:c} e_{d_1} & \mathbf{a} \\ \hdashline 0 & B' \end{array}\right)$$

其中$\mathbf{a} = (a_{1,d_1+1}, \cdots, a_{1,n})$。假设$B'$主元的列指标为$d_2, \ldots, d_s$，则$E_r(1, d_s(\frac{-1}{a_{d_s}})) \cdots E_r(1, d_2(\frac{-1}{a_{d_2}}))B_2$为简化型阶梯矩阵。

- 唯一性。留待以后证明。

□

4.3 线性方程组解的结构

为了进一步研究线性方程组解的结构，我们需要先介绍可逆矩阵的相关知识。

定义4.3.1. 设$A \in \mathrm{Mat}_{m\times n}(\mathbb{F})$，若$f_A \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$是一个双射，则$A$称为**可逆矩阵**。

推论4.3.2. 设$A \in \mathrm{Mat}_{m\times n}(\mathbb{F})$。矩阵$A$可逆当且仅当存在唯一$B \in \mathrm{Mat}_{n\times m}(\mathbb{F})$使得$AB = E_m$且$BA = E_n$。记作$A^{-1} = B$。

证明. 设A_f的逆映射是g，则$B = A_g$。 □

推论4.3.3. 初等矩阵都是可逆矩阵。

证明. 直接计算可知

- $E_n(i, j)^{-1} = E_n(i, j)$;
- $E_n(i(k))^{-1} = E_n(i(1/k))$;
- $E_n(i, j(k)) = E_n(i, j(-k))$。

□

推论4.3.4. 设$A \in \mathrm{Mat}_{l\times m}(\mathbb{F}), B \in \mathrm{Mat}_{m\times n}(\mathbb{F})$是可逆矩阵，则$AB$可逆且$(AB)^{-1} = B^{-1}A^{-1}$。

证明. 直接计算得

$$(AB)(B^{-1}A^{-1}) = A(BB^{-1})A = AE_mA^{-1} = E_l$$

且

$$(B^{-1}A^{-1})(AB) = B^{-1}(A^{-1}A)B = B^{-1}E_mB = E_n.$$

□

在讨论一般的线性方程组之前，让我们先研究一类特殊的线性方程组，即齐次线性方程组。

定义4.3.5. 设$A \in \mathrm{Mat}_{m\times n}(\mathbb{F})$，称方程$A\cdot \mathbf{x}^T = \mathbf{0}^T$为**齐次线性方程组**，其中$\mathbf{x} = (x_1, \ldots, x_n)$。

命题4.3.6. 设$A \in \mathrm{Mat}_{m\times n}(\mathbb{F})$行等价于$(d_1, \ldots, d_k)$-型矩阵$B$，则齐次线性方程组$A\cdot \mathbf{x}^T = \mathbf{0}^T$解集是一个线性空间，且与$B\cdot \mathbf{x}^T = \mathbf{0}^T$ 的相同，且同构于$\mathbb{R}^{n-k}$。

证明. 命题中实际上包含三个结论，我们逐一证明。

1. 注意到$A\cdot \mathbf{x}^T = \mathbf{0}^T$当且仅当$\mathbf{x} \in \ker(f_A)$，从而方程组的解构成一个线性空间。

2. 由于$A \sim_l B$，即存在初等矩阵$P_1, \ldots, P_r$使得

$$B = P_r \cdots P_1 A$$

   记$P = P_r \cdots P_1$，由推论4.3.4知P是可逆矩阵，从而$\ker(f_B) = \ker(f_A)$。

3. 设$B = (b_{ij})_{m\times n}$是$(d_1, \ldots, d_k)$-型可逆矩阵，从而有（其中$D = \{d_1, \ldots, d_k\}$）

$$\begin{cases} x_{d_1} = & \sum_{i>d_1, i\notin D} b_{1,i}x_i \\ x_{d_2} = & \sum_{i>d_1, i\notin D} b_{2,i}x_i \\ & \vdots \\ x_{d_k} = & \sum_{i>d_1, i\notin D} b_{k,i}x_i \end{cases}$$

   从而对任意$\underbrace{(c_1, \ldots, c_{d_1-1}, c_{d_1+1}, \ldots, c_n)}_{\text{下指标}d_1, \ldots, d_k\text{缺失}} \in \mathbb{F}^{n-k}$，都恰好对应原方程组的一组解$x_i = c_i$，其中$i \in \{1, \ldots, n\}\backslash D$，且原方程组的解都有这样的$c_i$唯一确定。

□

推论4.3.7. 设$A \in \mathrm{Mat}_{m\times n}(\mathbb{F})$是可逆矩阵，则$m = n$。特别地，对于$m \neq n$，必有$\mathbb{F}^n \ncong \mathbb{F}^m$。

证明. 反证法。假设$m \neq n$，不妨设$m < n$，否则考虑矩阵$A^{-1} \in \mathrm{Mat}_{n\times m}(\mathbb{F})$。由命题4.3.6知$\ker(f_A)$同构于$\mathbb{R}^{n-k}$。注意到$k \leq m$，从而$n - k \geq n - m > 0$，从而$f_A$并非单射，矛盾。 □

注4.3.8. 记$\mathrm{GL}_n(\mathbb{F}) \subseteq \mathrm{Mat}_{n\times n}(\mathbb{F})$为$n$-阶可逆矩阵构成的集合。对于线性空间$V$，记$\mathrm{GL}_n(V) \subseteq \mathcal{L}(V, V)$为到$V$自身的可逆映射构成的集合。$\mathrm{GL}_n(\mathbb{F}), \mathrm{GL}_n(V)$关于矩阵乘法和映射复合分别构成加法群，但其加法甚至不封闭。

有了上述推论，我们终于能够正式定义线性空间的维数。

定义4.3.9. 设线性空间V同构于$\mathbb{F}^n$，则称V的**维数是**n，记作$\dim(V) = n$。

设W是V上的一个子空间，$\mathbf{b} \in V$。集合

$$\mathbf{b} + W := \{\mathbf{b} + \mathbf{w} | \mathbf{w} \in W\}$$

称为V的一个超平面。若W的维数是k，称为一个k-维超平面。

下面是矩阵的一个关键不变量。

定义4.3.10. 设$A \in \mathrm{Mat}_{m\times n}(\mathbb{F})$，定义矩阵$A$的秩$\mathrm{rk}(A)$为$n - \dim(\ker(A))$。特别地，如果$A$行等价于一个$(d_1, \ldots, d_k)$-型简化阶梯矩阵，则$\mathrm{rk}(A) = k$。

推论4.3.11. 设$A \in \mathrm{Mat}_{m\times n}(\mathbb{F}), \mathbf{b} \in \mathbb{F}^m$，若线性方程组$A\mathbf{x}^T = \mathbf{b}^T$有解，则解集是一个$(n - \mathrm{rk}(A))$-维超平面。

证明. 设$\mathbf{u} \in \mathbb{F}^n$是原方程组一个解，则向量$\mathbf{v} \in \mathbb{F}^n$也是解当且仅当$\mathbf{u} - \mathbf{v}$是齐次方程组$A\mathbf{x}^T = \mathbf{0}^T$的解，从而$A\mathbf{x}^T = \mathbf{0}^T$的解集为$\mathbf{u} + \ker(A)$。 □

命题4.3.12. 设$A \in \mathrm{Mat}_{m\times n}(\mathbb{F})$。有

$$\dim(\mathrm{Im}(A)) = \mathrm{rk}(A).$$

证明. 设$\mathrm{rk}(A) = k$，则存在$(d_1, \ldots, d_k)$-型阶梯矩阵B以及可逆矩阵P使得$B = PA$。注意到

$$B : \mathbb{F}^n \xrightarrow{A} \mathbb{F}^m \xrightarrow[\cong]{P} \mathbb{F}^m.$$

从而$\mathrm{Im}(A) = \mathrm{Im}(B)$。由于$\mathrm{rk}(A) = \mathrm{rk}(B)$，因此只需对简化阶梯矩阵证明原命题。

但显然对于一个$(d_1, \ldots, d_k)$-型简化阶梯矩阵B，设$\mathbf{b} = (b_1, \ldots, b_m)$。方程$B\mathbf{x}^T = \mathbf{b}^T$有解的充分必要条件为$b_{k+1} = \cdots = b_m = 0$。即$\mathrm{Im}(B) \cong \mathbb{F}^k$。 □

推论4.3.13. 设$A \in \mathrm{Mat}_{m\times n}(\mathbb{F})$，且$PA$是简化阶梯矩阵，则方程$A\mathbf{x}^T = \mathbf{b}^T$有解的充分必要条件为$P\mathbf{b}^T$的后$(m - \mathrm{rk}(A))$个分量皆为0。

推论4.3.14. 设$f \in \mathcal{L}(\mathbb{F}^n, \mathbb{F}^m)$，则

$$\dim(\mathrm{Im}(f)) + \dim(\ker(f)) = n.$$

定义4.3.15. 对于矩阵$A \in \mathrm{Mat}_{m\times n}(\mathbb{F})$，我们称$A$是满秩的当且仅当$\mathrm{rk}(A) = \min\{m, n\}$。

注4.3.16. 矩阵A满秩的几何含义可以看作作为线性变换，A的像空间尽可能地大，核空间尽可能地小。特别地：

- 如果$m \le n$，矩阵A满秩意味着A是满射，其核空间$\ker(A) \cong \mathbb{F}^{n-m}$；
- 如果$m \ge n$，矩阵A满秩意味着A是单射，其像空间$\mathrm{Im}(A) \cong \mathbb{F}^n$。

现在补充证明命题4.2.2的唯一性。

证明. (命题4.2.2唯一性) 设存在可逆矩阵P, P'使得$B = PA, B' = P'A$为简化阶梯矩阵。由于$\ker(B) = \ker(A) = \ker(B')$，因此证明唯一性可转变为证明：若两个简化阶梯矩阵$B, B'$的核相等，则$B = B'$。

设B是$(d_1, \ldots, d_k)$-型而B'是$(d'_1, \ldots, d'_k)$-型。首先说明$d_i = d'_i$。否则存在最大的i使得$d_i \ne d'_i$。不妨假设$d_i < d'_i$，则x_{d_i}是方程$B\mathbf{x}^T = \mathbf{0}^T$的主元但是方程$B'\mathbf{x}^T = \mathbf{0}^T$的自由变元。令$B'\mathbf{x}^T = \mathbf{0}^T$中$x_{d_i} = 1$而其余的自由变元为0得到方程的一组解。注意到这组解中$x_{d_i+1} = x_{d_i+2} = x_n = 0$，但这不可能是$B\mathbf{x}^T = \mathbf{0}^T$的解（因为$x_d$应由$x_{d_i+1}, x_{d_i+2}, \ldots, x_n$齐次线性表示），与$\ker(B) = \ker(B')$矛盾！因此$B$与$B'$的主元相同。设$B = (b_{ij})_{m\times n}, B' = (b'_{ij})_{m\times n}$，对任意$i$以及$j > d_i$满足$x_j$是两方程的自由变元，令$x_j = 1$并令其他自由变元为0，分别得到两方程的一组解。此时$B\mathbf{x}^T = \mathbf{0}^T$的解中$x_i = -b_{ij}$而$B'\mathbf{x}^T = \mathbf{0}^T$的解中$x_i = -b'_{ij}$，这推出$b_{ij} = b'_{ij}$，从而$B = B'$。唯一性得证。 □

推论4.3.17. 矩阵$A \in \mathrm{Mat}_{n\times n}(\mathbb{F})$可逆的充分必要条件为$A \sim_l E_n$。特别地，任意可逆矩阵可表为若干个初等矩阵的乘积。

证明. 设B是唯一与A行等价的简化阶梯矩阵。由于A可逆，从而$\ker(B) = \ker(A) = 0$，从而方程$B\mathbf{x}^T = \mathbf{0}^T$无自由变量，由阶梯矩阵的形状即知$B = E_n$。 □

4.4 矩阵的相抵

定义/命题4.4.1. 设$A, B \in \mathrm{Mat}_{m\times n}(\mathbb{F}^n)$，规定关系$A \simeq B$当且仅当存在$P \in \mathrm{GL}_m(\mathbb{F}), Q \in \mathrm{GL}_n(\mathbb{F})$使得$B = PAQ$，此时我们称$A$与$B$相抵。相抵关系是一个等价关系。

证明. 逐条验证等价关系的条件。

- （自反性）取$P = E_m, Q = E_n$，有$A = PAQ$。
- （对称性）设$B = PAQ$，则$A = P^{-1}BQ^{-1}$。

- （传递性）设$B = P_1AQ_1, C = P_2BQ_2$，则$C = P_2P_1AQ_1Q_2$。注意到$\mathrm{GL}_k(\mathbb{F})$构成加法群，从而$P_2P_1, Q_1Q_2$都是可逆矩阵。

□

命题4.4.2. 设$A \in \mathrm{Mat}_{m\times n}(\mathbb{F}^n)$且$\mathrm{rk}(A) = k$，则$A \simeq \left(\begin{array}{c|c} E_k & 0 \\ \hline 0 & 0 \end{array}\right)$

证明. 由命题4.2.2，存在$P \in \mathrm{GL}_n(\mathbb{F})$，使得$B = PA$是$(d_1, \ldots, d_k)$-型简化阶梯矩阵。令

$$Q_1 = E_n(1, d_1)\cdots E_k(k, d_k),$$

则BQ_1具有形式$\left(\begin{array}{c|c} E_k & C \\ \hline 0 & 0 \end{array}\right)$，令$Q_2 = \left(\begin{array}{c|c} E_k & -C \\ \hline 0 & E_{n-k} \end{array}\right)$，则$PAQ_1Q_2 = BQ_1Q_2 = \left(\begin{array}{c|c} E_k & 0 \\ \hline 0 & 0 \end{array}\right)$。 □

命题4.4.3. 设$A \simeq B \in \mathrm{Mat}_{m\times n}(\mathbb{F}^n)$，则$\mathrm{rk}(A) = \mathrm{rk}(B)$。

证明. 设$P \in \mathrm{GL}_m(\mathbb{F}), Q \in \mathrm{GL}_n(\mathbb{F})$，使得$B = PAQ$。由于$\mathrm{Im}(B) = \mathrm{Im}(PAQ) = \mathrm{Im}(AQ) = \mathrm{Im}(A)$，因此由命题4.3.12知$\mathrm{rk}(B) = \mathrm{rk}(A)$。 □

推论4.4.4. 对于$k_1 \neq k_2$，有$\left(\begin{array}{c|c} E_{k_1} & 0 \\ \hline 0 & 0 \end{array}\right)_{m\times n} \not\simeq \left(\begin{array}{c|c} E_{k_2} & 0 \\ \hline 0 & 0 \end{array}\right)_{m\times n}$。

证明. 显然两矩阵都是简化阶梯矩阵，根据定义秩分别为k_1, k_2。由上一个命题立得结论。 □

注4.4.5. 商集$\mathrm{Mat}_{m\times n}(\mathbb{F})/\simeq$中仅有$k+1$个元素:

$$[0_{m\times n}], \left[\left(\begin{array}{c|c} E_1 & 0 \\ \hline 0 & 0 \end{array}\right)_{m\times n}\right], \left[\left(\begin{array}{c|c} E_2 & 0 \\ \hline 0 & 0 \end{array}\right)_{m\times n}\right], \ldots, \left[\left(\begin{array}{c|c} E_k & 0 \\ \hline 0 & 0 \end{array}\right)_{m\times n}\right],$$

其中$k = \min\{m, n\}$。这再次说明$\mathrm{rk}(A)$是矩阵A的重要不变量。

4.5 矩阵的转置

定义4.5.1. 设$A = (a_{ij})_{m\times n} \in \mathrm{Mat}_{m\times n}(\mathbb{F})$，定义$A$的转置为$(b_{ij} := a_{ji})_{n\times m} \in \mathrm{Mat}_{n\times m}(\mathbb{F})$，记作$A^T$

命题4.5.2. 转置作用与矩阵运算之间的关系如下（假设A, B的形状适当）：

1. $(A^T)^T = A$
2. $(A + B)^T = A^T + B^T$
3. $(kA)^T = kA^T$
4. $(AB)^T = B^TA^T$

证明. 设$A = (a_{ij})_{m\times n}$，并记$A^T = (a'_{ij})_{n\times m}$，因此$a_{ij} = a'_{ji}$

1. 根据定义，$(A^T)^T$的(i, j)-元为A^T的(j, i)-元，即为A的(i, j)-元，从而$(A^T)^T = A$
2. 设$B = (b_{ij})_{m\times n}$，并记$B^T = (b'_{ij})_{n\times m}$，因此$b_{ij} = b'_{ji}$。从而$A^T + B^T = (a'_{ij})_{n\times m} + (b'_{ij})_{n\times m} = (a'_{ij} + b'_{ij})_{n\times m} = (A + B)^T$。
3. 显然等式两边的(i, j)-元都为ka_{ji}。
4. 设$B = (b_{ij})_{n\times k}$。则$(AB)^T$的(i, j)-元即为AB的(j, i)-元，因此为$\sum_{t=1}^n a_{jt}b_{ti}$。若记$a'_{ij} = a_{ji}, b'_{ij} = b_{ji}$，则$B^TA^T$的$(i, j)$-元为$\sum_{t=1}^n b'_{it}a'_{tj} = \sum_{t=1}^n a_{jt}b_{ti}$。即等式成立。

□

4.6 习题

习题4.6.1. 求证：在集合$\mathrm{Mat}_{m\times n}$上的关系$\sim_l$是一个等价关系。

习题4.6.2. 设V,W是两个n-维线性空间，设$f\in\mathcal{L}(V,W)$，求证：f是单射当且仅当f是一个满射。

习题4.6.3. 令$a=\sqrt{2}+\sqrt{3}$，求$(c_0,c_1,c_2,c_3,c_4)\in\mathbb{Q}^5\backslash\{\mathbf{0}\}$使得$c_0+c_1a+c_2a^2+c_3a^3+c_4a^4=0$。

习题4.6.4. 设$A=(a_{ij})\in\mathrm{Mat}_{n\times n}(\mathbb{F})$，定义$\mathrm{tr}(A)=a_{11}+\cdots+a_{nn}$（即对角元素的和）。求证：对任意可逆矩阵$B\in\mathrm{Mat}_{n\times n}(\mathbb{F})$，都有$\mathrm{tr}(A)=\mathrm{tr}(BAB^{-1})$。（提示：先对$B$为初等矩阵的情形进行证明，再将一般的$B$分解为初等矩阵的乘积）

习题4.6.5. 设$\mathbf{x}_0,\dots,\mathbf{x}_n\in\mathbb{F}^n$，求证：存在$(a_0,\dots,a_n)\in\mathbb{F}^n\backslash\{0\}$，使得$a_0\mathbf{x}_0+\cdots+a_n\mathbf{x}_0=\mathbf{0}$。并以此证明，对任意$A\in\mathrm{Mat}_{n\times n}(\mathbb{F})$，存在$f(x)\in\mathbb{F}[x]$使得$f(A)=0$。

习题4.6.6. 对于方阵$A\in\mathrm{Mat}_{n\times n}(\mathbb{F})$，若$A$满足

- $A=A^T$，则称A是对称方阵阵，n阶对称方阵构成的集合记为$\mathrm{Mat}_n^{\mathrm{s}}(\mathbb{F})$;
- $A=-A^T$，则称A是反对称方阵阵，n阶反对称方阵构成的集合记为$\mathrm{Mat}_n^{\mathrm{as}}(\mathbb{F})$。

求证：有直和分解$\mathrm{Mat}_{n\times n}(\mathbb{F})\cong\mathrm{Mat}_n^{\mathrm{s}}(\mathbb{F})\oplus\mathrm{Mat}_n^{\mathrm{as}}(\mathbb{F})$

习题4.6.7. 设V是$\mathbb{F}$上的线性空间，$f\in\mathcal{L}(V,V)$满足$f^2=f$。记$V_1=\ker(f),V_2=\mathrm{Im}(f)$。求证：

$$V\cong V_1\oplus V_2.$$

（因此我们也称满足这样条件的映射称为投影映射）

4.7 附录：直和分解与分块矩阵

定义/命题4.7.1. 设V是一个$\mathbb{F}$上的线性空间，$V_1,\dots,V_n\subseteq V$是线性子空间，满足对任意$v\in V$，都存在唯一的$v_1,\dots,v_n\in V$使得

$$v=v_1+\cdots+v_n.$$

称$(V_1,\dots,V_n)$为V的一个直和分解。自然的映射

$$\begin{aligned}\mathrm{pr}_i:V&\to V_i\\ v=v_1+\cdots+v_n&\mapsto v_i\end{aligned}$$

称为V（**在此分解下**）**到第i个分量的投影**。则有以下自然同构

$$\begin{aligned}V&\xrightarrow{\cong}V_1\oplus\cdots\oplus V_n\\ F:v&\mapsto(\mathrm{pr}_1(v),\dots,\mathrm{pr}_n(v))\\ i_1(v_1)+\cdots+i_n(v_n)&\mapsfrom(v_1,\dots,v_n):G\end{aligned}$$

其中$i_1,\dots,i_n$是$V_1,\dots,V_n$到V的自然嵌入。

证明. 根据假设，有

$$\begin{aligned}v&=v_1+\cdots v_n\\ &=i_1\circ\mathrm{pr}_1(v)+\cdots i_n\circ\mathrm{pr}_1(v)\\ &=G\circ F(v)\end{aligned}$$

反之，对任意$v_1\in V_1,\dots,v_n\in V_n$，有

$$\begin{aligned}v_k&=\mathrm{pr}_k(i_1(v_1)+\cdots+i_n(v_n))\\ &=\mathrm{pr}_k\circ G(v_1,\dots,v_n)\end{aligned}$$

从而

$$(v_1,\dots,v_k)=(\mathrm{pr}_1\circ G(v_1,\dots,v_n),\cdots,\mathrm{pr}_n\circ G(v_1,\dots,v_n))=F\circ G(v_1,\dots,v_n).$$

□

最简单的例子是标准线性空间到其各个坐标分量的直和分解。

例4.7.2. 设$V=\mathbb{F}^n$。对于$k=1,\dots,n$，令$V_k=\mathbb{F}\cdot e_k=\{a\cdot e_k|a\in\mathbb{F}\}$（回忆$e_k$是第$k$个分量为1其余分量为0的向量）为$V$的一维子空间。则有标准的直和分解

$$V\cong V_1\oplus\cdots\oplus V_n$$
$$v=(a_1,\dots,a_n)\mapsto(a_1e_1,\dots,a_ne_n)$$

特别需要注意的是，投影映射$\mathrm{pr}_k:V\to V_k$不只依赖V_k，而依赖整个直和分解$(V_1,\cdots,V_k)$。

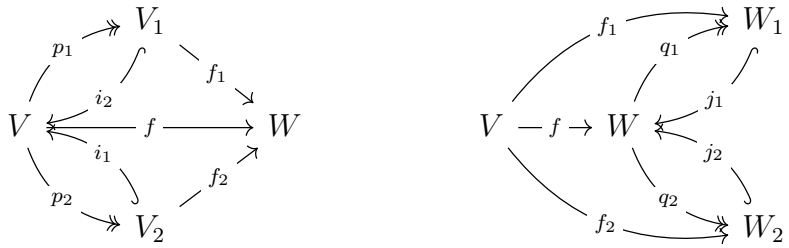
例4.7.3. 设$V=\mathbb{R}^3$，令$V_1=\mathbb{R}\cdot e_1\oplus\mathbb{R}\cdot e_2=\{a_1e_1+a_2e_2|a_1,a_2\in\mathbb{R}\}$及$V_2=\mathbb{R}\cdot e_3$。从而有直和分解$V=V_1\oplus V_2$，则对任意$v=(a_1,a_2,a_3)\in V$，有$\mathrm{pr}_1(v)=(a_1,a_2,0)$。

但若令$V_2'=\mathbb{R}\cdot(1,1,1)$，则对任意的$v=(a_1,a_2,a_3)\in V$，有

$$v=(a_1-a_3,a_2-a_3,0)+(a_3,a_3,a_3)$$

从而在直和分解$V\cong V_1\oplus V_2'$下，有$\mathrm{pr}_1(v)=(a_1-a_3,a_2-a_3,0)$。

直和分解不仅能分解线性空间中的向量，还能进一步诱导线性映射的分解。为了简化思考，我们先考虑$n=2$的情形。事实上，我们可以通过考虑下面两个图表对空间$\mathcal{L}(V,W)$进行直和分解。



命题4.7.4. 假设V,W是$\mathbb{F}$线性空间，其中V有直和分解$V=V_1\oplus V_2$，W有直和分解$W=W_1\oplus W_2$。记$p_1:V\to V_1,p_2:V\to V_2,q_1:W\to W_1,q_2:W\to W_2$是各自在此分解下的投影映射，而$i_1,i_2,j_1,j_2$是$V_1,V_2,W_1,W_2$到$V,W$的自然嵌入。则映射

$$\mathcal{L}(V_1\oplus V_2,W)\to\mathcal{L}(V_1,W)\oplus\mathcal{L}(V_2,W)$$
$$F_1:f\mapsto(f\circ i_1,f\circ i_2)$$
$$f_1\circ p_1+f_2\circ p_2\hookleftarrow(f_1,f_2):G_1$$

及映射

$$\mathcal{L}(V,W_1\oplus W_2)\to\mathcal{L}(V,W_1)\oplus\mathcal{L}(V,W_2)$$
$$F_2:f\mapsto(q_1\circ f,q_2\circ f\circ)$$
$$j_1\circ f_1+j_2\circ f_2\hookleftarrow(f_1,f_2):G_2$$

是线性同构。

证明. 设$f\in\mathcal{L}(V,W)$。由定义/命题4.7.1知，对任意$v\in V$，有$G_1\circ F_1(f(v))=f(v)$，从而$F_1\circ G_1=\mathrm{id}$。同理，对任意$(f_1,f_2)\in\mathcal{L}(V_1,W)\oplus\mathcal{L}(V_2,W)$，以及任意$v_1\in V_1,v_2\in V_2$，有$F_1\circ G_1(f_1(v_1),f_2(v_2))=(f_1(v_1),f_2(v_2))$，从而$F_2\circ G_2=\mathrm{id}$。进而$F_1$是线性同构，并且$G_1=F_1^{-1}$。

另一部分的证明完全对称，留给读者作为练习。 □

注4.7.5.

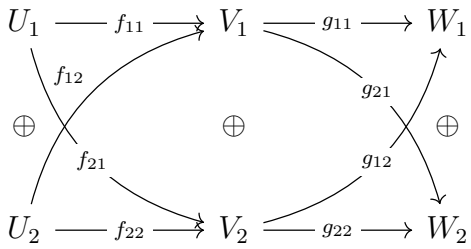
- 将上述两个分解结合起来，我们就得到直和分解

$$\mathcal{L}(V,W)\cong\mathcal{L}(V_1,W)\oplus\mathcal{L}(V_2,W)\cong\begin{matrix}(\mathcal{L}(V_1,W_1)\oplus\mathcal{L}(V_2,W_1))\\ \bigoplus\\ (\mathcal{L}(V_1,W_2)\oplus\mathcal{L}(V_2,W_2))\end{matrix}$$

若$V_1 = \mathbb{F}^{n_1}, V_2 = \mathbb{F}^{n_2}, W_1 = \mathbb{F}^{m_1}, W_2 = \mathbb{F}^{m_2}$，并记$n = n_1 + n_2, m = m_1 + m_2$，则以上分解就对应$m \times n$阶矩阵进行$(n_1, n_2) \times (m_1, m_2)$分块。

$$\begin{aligned} \mathrm{Mat}_{m\times n} \cong & \begin{matrix} (\mathrm{Mat}_{m_1\times n_1} \oplus \mathrm{Mat}_{m_1\times n_2}) \\ \bigoplus \\ (\mathrm{Mat}_{m_2\times n_1} \oplus \mathrm{Mat}_{m_2\times n_2}) \end{matrix} \\ A \mapsto & \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix} \end{aligned}$$

- 假设我们还有$U \cong U_1 \oplus U_2$，我们可以考虑映射复合：



从而有

$$\begin{pmatrix} g_{11} & g_{12} \\ g_{21} & g_{22} \end{pmatrix} \circ \begin{pmatrix} f_{11} & f_{12} \\ f_{21} & f_{22} \end{pmatrix} = \begin{pmatrix} g_{11} \circ f_{11} + g_{12} \circ f_{12} & g_{11} \circ f_{12} + g_{12} \circ f_{22} \\ g_{21} \circ f_{11} + \circ g_{22} \circ f_{21} & g_{21} \circ f_{12} + g_{22} \circ f_{22} \end{pmatrix}$$

这也从另一个视角再次解释了矩阵乘法的法则。

- 显然我们可以将2×2的分块推广到一般的$r \times s$分块的情况。假设$\mathbb{F}$上的线性空间V, W有直和分解$V = V_1 \oplus \cdots \oplus V_s, W = W_1 \oplus \cdots \oplus W_r$，则有自然的线性同构

$$\begin{aligned} \mathcal{L}(V, W) \cong & \bigoplus_{i=1}^{s} \bigoplus_{j=1}^{r} \mathcal{L}(V_i, W_j) \\ f \mapsto & (\mathrm{pr}_j \circ f \circ \iota_i)_{r\times s} \end{aligned}$$

其中$\iota_i : V_i \hookrightarrow V$是自然嵌入而$\mathrm{pr}_j : W_j \twoheadrightarrow W$是投影映射。