代数学

琉璃色岛屿 Dearee

2024 年 6 月 2 日

目录

第一部分

基础代数

第一章　多项式预先介绍

1.1　数域

定义 1.1.1 P 是**数域**，如果

1. $P \subset \mathbb{C}$,
2. $0, 1 \in P$.
3. $\forall a, b \in P$, 有 $a \pm b \in P, a \cdots b \in P, \frac{a}{b} \in P(b \neq 0)$.(即 a, b 在 P 上的四则运算**封闭**).

1.2　一元多项式

定义 1.2.1 (一元多项式) 设 $a_0, a_1, \cdots a_n$ 属于数域 P. 则形如

$$a_nX^n + a_{n-1}X^{n-1} + \cdots + a_0, (n \in \mathbb{N})$$

的式子被称为数域 P 上关于**符号** X 的**多项式**.

注 1.2.1 形式不唯一，可以是 $a_0 + a_1X + a_2X^2 + \cdots a_nX^n$ 或者 $a_0X^n + a_1X^{n-1} + \cdots a_n$.

我们可以记 $f(X) = a_nX^n + a_{n-1}X^{n-1} + \cdots + a_0, f(x) \in P[x]$ 来表示系数在数域 P 中关于 X 的多项式. 其中 $P[X]$ 是由全体系数属于数域 P 的多项式的集合连同它的运算所形成的**多项式环**.

定义 1.2.2 (多项式的次数) 设 $f(X) = a_nX^n + a_{n-1}X^{n-1} + \cdots + a_0$. 则称 a_nX^n(最高此项) 为该多项式的**首项**，a_n 为**首项系数**，并定义 X 的最高次数为多项式 $f(X) \in P[X]$ 的次数，并记为 $\deg(f(X))$ 或简记 $\deg(f)$.

定义 1.2.3 (多项式的相等) 给定数域 P 上的多项式 $f(X), g(X)$. 若 $f(X), g(X)$ 除去系数为 0 的项以外所有次数相同的项的系数都相等，则称这两个多项式**相等**.

若一个多项式的所有系数都为 0 则称该多项式为**零多项式**. 我们可以认为零多项式的次数为 $-\infty$ 或者干脆认为它**没有意义**. 本书定义零多项式的次数为 $-\infty$.

1.2.1　多项式的和与乘

设属于 P 上的多项式 $f(X)=a_nX^n+a_{n-1}X^{n-1}+\cdots+a_0, g(x)=b_mX^m+b_{m-1}X^{m-1}+\cdots+b_0$，且 $n\geqslant m$. 则

$$f(X)+g(X)=\sum_{k=0}^{n}(a_k+b_k)X^k$$
$$f(X)g(X)=\sum_{k=0}^{n+m}\sum_{i+j=k}(a_ib_j)X^k=\sum_{k=0}^{n+m}\sum_{i=0}^{k}(a_ib_{k-i})X^k$$

定理 1.2.1 设 $f(X),g(X)\in P[X]$. 则有

$$\deg(f(X)+g(X))\leqslant\max\{\deg(f(X)),\deg(g(X))\},\deg(f(X)g(X))\leqslant\deg(f(X))+deg(f(X)).$$

特别地，若两个多项式非零，则第二个式子为 $\deg(f(X)g(X))=\deg(f(X))+\deg(f(X))$.

证明 直接取自定义. □

注 1.2.2 显然对两个多项式做以上运算后的新多项式仍属于 $P[X]$.

介于 $P[X]$ 是一个多项式环，那么属于它的多项式 $f(X),g(X),h(X)$ 还满足以下运算法则.

定理 1.2.2

1. 加法结合律:
$$f(X)+(g(X)+h(X))=(f(X)+g(X))+h(X).$$
2. 加法交换律:
$$f(X)+g(X)=g(X)+f(X).$$
3. 乘法结合律:
$$f(X)(g(X)h(X))=(f(X)g(X))h(X)$$
4. 乘法交换律:
$$f(X)g(X)=g(X)f(X).$$
5. 乘法分配律:
$$f(X)(g(X)+h(X))=f(X)g(X)+f(X)h(X).$$

1.3　多项式的整除

定理 1.3.1 (带余除法) $P[X]$ 中的多项式 $f(x),g(x),(g(x)\neq 0)$ 则一定有 $P[X]$ 中的多项式 $q(X),r(X)$ 满足

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

并且 $q(X),r(X)$ 是唯一确定的.

证明 如果 $deg(f)<deg(g)$. 则一定有 $q(X)=0,r(X)=f(X)$.

如果 $deg(f)\geqslant\deg(g)$. 利用“综合除法”总能找到 $q(X),r(X)$. □

注 1.3.1 称 $q(X)$ 为 $f(X)$ 除以 $g(X)$ 的**商式**或**商**.$r(X)$ 为 $f(X)$ 除以 $g(X)$ 的**余式**或**余**.

定义 1.3.1 (多项式的整除) 给定属于 $P[X]$ 的多项式 $f(X), g(X)$. 如果可以找到属于 $P[X]$ 的多项式 $q(X)$ 满足

$$f(X) = q(X)g(X).$$

则称 $g(X)$ **整除** $f(X)$.，记为 $g(X)|f(X)$. 并称 $f(X)$ 是 $g(X)$ 的**因式**，$g(X)$ 是 $f(X)$ 的**倍式**. 如果不整除可以记为 $g(X) \nmid f(X)$.

注 1.3.2 任意多项式整除 0，即 $f(X)|0$，因为 $0 = f(X)0$. 如果 $g(X)$ 不等式 0，整除可记为 $\frac{f(X)}{g(X)}$. 如果 $q(x)|f(x)$ 则有 $mq(x)|nf(x)$, 其中 m, n 是非零常数.

定理 1.3.2 若 $f(X), g(X) \in P[X]$，且 $f(X)|g(X), g(X)|f(X)$ 同时成立，那么 $f(X) = cg(X)$(其中 c 是非零常数).

证明 由 $f(X)|g(X), g(X)|f(X)$，则存在 $h_1(X), h_2(X) \in P[X]$ 满足

$$g(X) = h_1(X)f(X), f(X) = h_2(X)g(X).$$

于是

$$f(X) = h_1(X)h_2(X)f(X).$$

如果 $f(X) = 0$ 则显然成立，否则 $1 = h_1(X)h_2(X)$. 于是有 $0 = \deg(1) = \deg h_1 + \deg h_2$. 于是 $h_1(X)h_2(X)$ 为非零常数.□

定理 1.3.3 (传递性) 已知 $f(X), g(X), h(X) \in P[X]$ 以及 $f(X)|g(X), g(X)|h(X)$. 那么就有

$$f(X)|h(X).$$

证明 存在 $q_1(X), q_2(X) \in P[X]$. 满足

$$g(X) = q_1(X)f(X), h(X) = q_2(X)g(X).$$

有

$$h(X) = q_1(X)q_2(X)f(X).$$

于是 $f(X)|h(X)$. □

定义 1.3.2 (组合) 多项式 $f_i(X) \in P[X], i = 1, 2 \cdots n$. 那么则称 $\sum_{i=1}^{n} u_i(X)f_i(X)$ 是 $f_i(X)$ 的一个**组合**.

定理 1.3.4 若 $h(X) \in P[X]$ 满足 $h(X)|f_i(X), i = 1, 2, \cdots n$ 那么它一定整除 $f_i(X)$ 的任意组合 $\sum_{i=1}^{n} u_i(X)f_i(X), u_i \in P[X], i = 1, 2, \cdots, n$.

证明 由整除得 $f_i(X) = q_i(X)h(X)$. 则

$$\sum_{i=1}^{n} u_i(X)f_i(X) = h(X)\sum_{i=1}^{n} u_i(X)q_i(X). \tag{1.3.1}$$

注 1.3.3 两个多项式得整除不会因为系数域得扩大而改变.

1.4　最大公因式

定义 1.4.1 (公因式) $\forall f(X), g(X) \in P[X]$. 若存在 $h(X) \in P[X]$ 满足

$$h(X)|f(X) \wedge h(X)|g(X). \tag{1.4.1}$$

则称 $h(X)$ 是 $f(X), g(X)$ 的**公因式**.

定义 1.4.2 (最大公因式) $f(X), g(X), h(X) \in P[X]. h(X)$ 是 $f(X), g(X)$ 的**最大公因式**如果

1. $h(X)|f(X) \wedge h(X)|g(X)$.
2. $c(X) \in P[X](c(X)|f(X), c(X)|g(X) \Rightarrow c(X)|h(X))$.

并且记 $(f(X), g(X))$ 为 $f(X), g(X)$ 首项系数为 1 的最大公因式.

注 1.4.1 整环 R 中的元素 a, b 的最大公因记为 $\gcd(a, b)$. 且有以下性质

1. $\gcd(a, 0) = a$.
2. $t\gcd(a, b) = \gcd(ta, tb)$.
3. $\gcd(a, b) = a \Leftrightarrow a|b$.
4. $\gcd(a, \gcd(b, c)) = \gcd(\gcd(a, b), c)$.

引理 1.4.1 $\forall f(X) \in P[X], (f(X), 0) = f(X)$.

引理 1.4.2 $f(X), g(X), q(X), r(X) \in P[X]$ 满足等式

$$f(X) = q(X)g(X) + r(X).$$

则 $\gcd(f(X), g(X)) = \gcd(g(X), r(X))$.

证明 $\gcd(g(X), r(X))$ 则说明 $\gcd(g(X), r(X))|q(X)g(X) + r(X)$, 说明 $\gcd(g(X), r(X))|f(X)$. 同理 $\gcd(f(X), g(X)) \Rightarrow \gcd(f(X), g(X))|f(X) - q(X)g(X) = \gcd(f(X), g(X))|r(X)$. □

注 1.4.2 将 $\gcd(f(X), g(X)) = \gcd(g(X), r(X))$ 换为 $(f(X), g(X)) = (g(X), r(X))$ 也成立.

定理 1.4.1 $\forall f(X), g(X) \in P[X]$ 一定有 $\gcd(f(X), g(X))$. 并且 $\exists u(X), v(X) \in P[X], \gcd(f(X), g(X)) = u(X)f(X) + v(X)g(X)$.

证明 利用**辗转相除法**使得多项式次数不断下降，得到最大公因式，再回代得到组合. □

定义 1.4.3 (互素) 两个环 $P[X]$ 中的多项式 $f(X), g(X)$ **互素**，如果

$$(f(X), g(X)) = 1. \tag{1.4.2}$$

定理 1.4.2 $\forall f(X), g(X) \in P[X], (f(X), g(X)) = 1 \Leftrightarrow \exists u(X), v(X) \in P[X](u(X)f(X) + v(X)g(X) = 1)$.

证明 必要性：$u(X)f(X)+v(X)g(X)=(f(X),g(X))=1$. 充分性：如果已知 $\exists u(X),v(X)\in P[X](u(X)f(X)+v(X)g(X)=1)$. 并由 $(f(X),g(X))|f(X)\wedge(f(X),g(X))|g(X)$ 可得 $(f(X),g(X))|u(X)f(X)+v(X)g(X)$. 于是 $(f(X),g(X))|1$ 故 $(f(X),g(X))=1$. □

定理 1.4.3 $\forall f(X),g(X),h(X)\in P[X]$, 如果有 $(f(X),g(X))=1, f(X)|g(X)h(X)$. 那么 $f(X)|h(X)$.

证明 $\exists u(X),v(X)\in P[X](u(X)f(X)+v(X)g(X)=1.)$ 两边同时乘以 $h(X)$ 可得

$$[u(X)h(X)]f(X)+v(X)g(X)h(X)=h(X). \tag{1.4.3}$$

由于 $f(X)|g(X)h(X)$ 故 $f(X)|h(X)$. □

定理 1.4.4 对于 $P[X]$ 中的 $f_1(X),f_2(X),h(X)$. 如果有 $(f_1(X),f_2(X))=1\wedge f_1(X)|h(X),f_2(X)|h(X)$,那么 $f_1(X)f_2(X)|h(X)$.

证明 $\exists q_1(X)$ 满足 $h(X)=q_1(X)f_1(X)$. 由于 $f_2(X)|f_1(X)q_1(X),(f_1(X),f_2(X))=1$ 故 $f_2(X)|q_1(X)$. 于是

$$h(X)=q_2(X)f_1(X)f_2(X). \tag{1.4.4}$$

就得到了我们要的结论. □

1.5 因式分解定理

一个多项式能否因式分解以及因式分解什么时候终止取决于它所在的环.

定义 1.5.1 (既约多项式) 环 $P[X]$ 上次数大于等于 1 的多项式如果不能分解为两个次数小于它的多项式（这些多项式也在同一个环中）的乘积，那么该多项式就是环 $P[X]$ 中的**既约多项式**或**不可约多项式**.

注 1.5.1 显然次数为 1 的多项式不论在什么环中都为不可约多项式.

引理 1.5.1 对于环 $P[X]$ 上的既约多项式，如果存在 $f(X)\in P[X](f(X)|p(X))$ 则 $f(X)$ 为非零常数或者 $f(X)=cp(X)$，其中 c 为非零常数.

证明 $f(X)|p(X)\Rightarrow\exists q(X)\in P[X](p(X)=q(X)f(X).)$ 于是 $\deg(p)=\deg(qf)=\deg(q)+\deg(f)$. 如果 $\deg(q)\geqslant 1,\deg(f)\geqslant 1$ 则 $P(X)$ 不是既约多项式. 故必有 $\deg(q)=0\vee\deg(f)=0$. □

引理 1.5.2 $\forall f(X)\in P[X]$. 那么 $P[X]$ 中的既约多项式 $p(X)$ 总满足 $p(X)|f(X)$ 或者 $(p(X),f(X))=1$.

证明 $(p(X),f(X))\Rightarrow(p(X),f(X))|p(X)\wedge(p(X),f(X))|f(X)$. 由于 $\gcd(p(X),f(X))$ 只能为非零常数或者为 $p(X)$ 的倍数，故 $(p(X),f(X))=1\vee(p(X),f(X))=cp(X)$. □

定理 1.5.1 如果 $p(X)\in P[X]$ 是不可约多项式，且对于 $f(X),g(X)\in P[X]$ 有 $p(X)|f(X)g(X)$. 那么有 $p(X)|f(X)\vee p(X)|g(X)$.

证明 如果 $p(X)|f(X)$，那么直接得出结论. 如果 $(p(X),f(X))=1$ 则 $p(X)|g(X)$. □

注 1.5.2 定理可以推广为：$p(X)\in P[X],f_i(X),i=1,2,\cdots n\in P[X]$ 有 $p(X)|\prod\limits_{i=1}^{n}f_i(X)$. 则存在 $j\in\mathbb{N}(p(X)|f_j(X))$.

定理 1.5.2 (唯一分解定理) $\forall f(X)\in P[X]$, 都可以分解为不可约多项式的乘积，

$$f(X)=p_1(X)p_2(X)\cdots p_s(X)=q_1(X)q_2(X)\cdots q_t(X),p_i(X),q_j(X)\in P[X].$$

并且，该分解是**唯一的**，也就是 $s=t$ 并且在式子适当排序后总有

$$p_k(X)=cq_k(X).$$

证明 利用数学归纳法. 如果 $f(X)$ 已经是不可约多项式那么已经无法继续分解，它本身就是一个分解. 如果可以继续分解，那么有

$$f(X)=g_1(X)g_2(X),g_1(X),g_2(X)\in P[X] \tag{1.5.1}$$

如果 $g_1(X),g_2(X)$ 以及既约，那么分解结束，否则它可以继续分解. 不断重复，直到分解出的所有多项式既约.

然后我们证明整除唯一. 首先有 $q_1(X)|q_t(X),p_i(X),q_j(X)$ 于是 $\exists p_w(X),1\leqslant w\leqslant j,q_1(X)|p_w(X)$. 于是等号两边消去 $p_1(X)$. 再找 $p_2(X)$ 不断重复. □

1.5.1　标准因式分解

为了方便，我们将上述因式分解 $f(X)=p_1(X)p_2(X)\cdots p_s(X)$ 的所有不可约因式的首项系数提出使之变为首项系数为 1 的因式，然后将相同的项合并得到

$$f(X)=ch^{r_1}(X)h^{r_2}(X)\cdots h^{r_l}(X). \tag{1.5.2}$$

显然，两个标准因式分解的多项式的最大公因式就是将两个多项式中都出现过的因式提出来，然后取两个多项式中那个指数更小的作为该因式的指数，然后再相乘, 得到新的多项式就是最大公因式.

1.6　重因式

定义 1.6.1 $p(X)\in P[X]$ 是不可约多项式，如果对于 $f(X)\in P[X]$ 有 $p^k(X)|f(X)\wedge p^{k+1}(X)\nmid f(X)$ 则称 $p(X)$ 是 $f(X)$ 的 k **重因式**

注 1.6.1 如果 $k=1$，$p(X)$ 是 $f(X)$ 的**单因式**，如果 $k>1$ 则 $p(X)$ 是 $f(X)$ 的**重因式**.

定理 1.6.1 对于 $P[X]$ 中的 $p(X),f(X)$，如果 $p(X)$ 是 $f(X)$ 的重因式，那么 $p(X)$ 是 $f'(X)$ 的 $k-1$ 重因式.

证明 $f(X)=p^k(X)g(X)$ 其中 $p(X)\nmid g(X)$ 那么 $f'(X)=kp^{k-1}(X)p'(X)g(X)+p^k(X)g'(X)=p^{k-1}(X)(kp'(X)g(X)+p(X)g'(X))$ 于是 $p^{k-1}|f'(X)$. 但 $p^k(X)|f'(X)$. □

推论 1.6.1 如果 $p(X)$ 是 $f(X)$ 的 k 重因式，那么 $p(X)$ 是 $f(X),f'(X),\cdots,f^{(k-1)}(X)$ 的因式，但不是 $f^{(k)}(X)$ 的因式.

定理 1.6.2 $P[X]$ 中 $p(X)$ 是 $f(X)$ 的重因式等价于 $p(X)$ 是 $f(X),f'(X)$ 的公因式. 反之如果 $f(X)$ 没有重因式则 $(f(X),f'(X))=1$.

证明 如果 $p(X)$ 是 $f(X)$ 的重因式那么至少有 $p^2(X)|f(X)$，于是有 $p(X)|f'(X)$. 反之，如果 $p^k(X)$ 是它们的公因式，那么有 $p^k(X)|f'(X)$ 则至少有 $p^{k+1}|f(X)$. □

如果要让一个多项式 $f(X)\in P[X]$ 没有重因式，只需

$$\frac{f(X)}{(f(X),f'(X))}. \tag{1.6.1}$$

1.7　多项式函数

定义 1.7.1 我们设 $a_i, x \in \mathbb{R}, i = 1, 2, \cdots, n$ 则称

$$f : \mathbb{R} \to \mathbb{R}, x \mapsto f(x) = a^n x^n + a^{n-1} x^{n-1} + \cdots + x_1 + a_0.$$

为实数域上的一个**多项式函数**

我们将 $x = \alpha \in \mathbb{R}$ 替换 x 则得到了一个实数值.

定义 1.7.2 如果 $x = \alpha \in \mathbb{R}$ 代入得到 $f(\alpha) = 0$. 则称 α 为 $f(X)$ 的根.

多项式 $f(x)$ 除以 $x - \alpha$ 总能得到

$$q(x) \in R[X], f(x) = q(x)(x - \alpha) + r(x).$$

由于余式的次数一定小于除式的次数，故以上式子一定为

$$f(x) = q(x)(x - \alpha) + c, c \in \mathbb{R}. \tag{1.7.1}$$

推论 1.7.1 $\alpha \in \mathbb{R}$ 是 $f(x) \in \mathbb{R}[X] \Leftrightarrow (x - \alpha) | f(x)$.

定义 1.7.3 1 如果 $x - \alpha$ 是 $f(x)$ 的 k 重因式，那么 α 就是 $f(x)$ 的 k **重根**. 如果 $k = 1$ 则 α 为**单根**，如果 $k > 1$ 则 α 为**重根**.

定理 1.7.1 一个 n 次多项式的根的个数不超过 n，其中重根按重数算.

证明 如果重根 α 超过 n 个，则 $(x - \alpha)^{n+1} \nmid f(x)$ 得到矛盾. □

定理 1.7.2 对于两个多项式函数 $f(x), g(x) \in \mathbb{C}[x]$ 的次数不超过 n，而对于 $n+1$ 个数有 $f(\alpha_i) = g(\alpha_i), i = 1, 2, \cdots, n, n+1$ 则有

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

证明 设 $F(x) = f(x) - g(x) \neq 0$ 而对于由前面得它有 $n + 1$ 个根，而这时不可能的. □

1.8　复系数与实系数的多项式分解

定理 1.8.1 (代数基本定理) 任何复数域上的 $n, n \geqslant 1$ 次多项式都恰有 n 个复数根. 这个说法等价于任何复数域上的次数大于等于 1 的多项式都必有一个复数根.

这表明任何复数域上的多项式 $f(x)$ 都可以唯一的分解成

$$f(x) = c(x - \alpha_1)^{r_1}(x - \alpha_2)^{r_2} \cdots (x - \alpha_n)^{r_n}. \tag{1.8.1}$$

接下来看实数域上的多项式.

定理 1.8.2 对于任意多项式 $f(x) \in \mathbb{R}[x]$，且复数 α 是它的根则它的共轭复数 $\bar{\alpha}$ 也是它的根.

上面的证明是显然的，由此可以证明以下定理.

定理 1.8.3 对于任意 $f(x)\in\mathbb{R}[x]$. 它总能分解成一些一次多项式和一些二次多项式的乘积.

注 1.8.1 可能只能分解成一次多项式的乘积或者二次多项式的乘积.

证明 对于一次多项式，这时显然的. 对于次数大于等于三的多项式，如果它有一实数根，则

$$f(x)=(x-\alpha)g(x).$$

说明它可以被分解，当然 $g(x)$ 可以再分解下去.

如果它有一复根，由上一个定理可知

$$f(x)=(x-\alpha)(x-\bar{\alpha})h(x).$$

其中 $(x-\alpha)(x-\bar{\alpha})$ 是一个二次多项式. 同理 $h(x)$ 也可以再分解. 最后由数学归纳法可证上面的定理. □

于是我们可以得到实数域上多项式函数的标准分解式

$$\mathbb{R}[x]\ni f(x)=a(x-c_1)^{l_1}(x-c_2)^{l_2}\cdots(x-c_s)^{l_s}(x^2+p_1x+q_1)^{k_1}(x^2+p_2x+q_2)^{k_2}\cdots(x^2+p_rx+q_r)^{k_r}. \tag{1.8.2}$$

1.9 有理数域上的多项式

定义 1.9.1 (本原多项式) 非零多项式 $f(X)\in\mathbb{Z}[X]$. 若其系数满足 $(a_0,a_1,\cdots,a_n)=1$ 即互素的. 则称该多项式为**本原多项式**. 任何非零多项式 $g(X)\in\mathbb{Q}[X]$ 都可以表示为 $g(X)=rf(X),r\in\mathbb{Q},f(X)\in\mathbb{Q}[X]$.

引理 1.9.1 (Gauss 引理) 两个本原多项式的积还是本原多项式.

证明 直接由定义可证. □

定理 1.9.1 如果一个非零的整系数多项式可以分解为两个次数较低的有理系数多项式，则它一定可以分解为两个次数较低的整系数多项式.

推论 1.9.1 若 $f(X)\in\mathbb{Z}[X],h(X)\in\mathbb{Q}[X]$. 满足等式

$$f(X)=g(X)h(X).$$

此时若 $g(X)$ 是本原多项式，则 $h(X)$ 一定为整系数多项式.

定理 1.9.2 (Eisenstein 判别法) 多项式 $f(X)=a_nX^n+\cdots+a_1X+a_0\in\mathbb{Z}[X]$.

若存在一个素数 p. 满足

$$p\nmid a_n,p|a_{n-1},\cdots,a_0,p^2\nmid a_0$$

则 $f(X)$ 为有理数域上的不可约多项式.

第二章　代数预备内容

2.1　线性方程组

我们在中学的时候已经学会解带有实系数 a,b,c,e,f,g 以及未知数 x,y 的方程组

$$\begin{cases} ax+by=c \\ ex+fy=g \end{cases}.$$

现在我们将学习解形如

$$\begin{gathered} a_{11}x_1+a_{12}x_2+\cdots+a_{1n}x_n=b_1, \\ a_{21}x_2+a_{22}x_2+\cdots+a_{2n}x_n=b_2, \\ \cdots\cdots \\ a_{m1}x_1+a_{m2}x_2+\cdots+a_{mn}x_n=b_m. \end{gathered} \tag{2.1.1}$$

系数 a_{ij} 读作 $a-i-j$，表示第 i 个方程中第 j 个未知数的系数. 数字 b_i 称为第 i 个方程的常数项. 我们称 $b_1=b_2=\cdots=b_m=0$ 的方程组.

$$\begin{gathered} a_{11}x_1+a_{12}x_2+\cdots+a_{1_n}x_n=0, \\ a_{21}x_2+a_{22}x_2+\cdots+a_{2_n}x_n=0, \\ \cdots\cdots \\ a_{m1}x_1+a_{m2}x_2+\cdots+a_{mn}x_n=0. \end{gathered} \tag{2.1.2}$$

为**齐次方程组**，也是**方程组**(2.1.1)的**诱导组.** 然后我们将各未知数的系数拿出来组成一个表

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} \tag{2.1.3}$$

叫做一个 $m\times n$ **矩阵** (如果 $m=n$ 则称为 n 阶**方阵**). 并用 (a_{ij}) 或用字母 A 来简记，并且我们将所有 $a_{ij}\in K$ 的 $m\times n$ 阶矩阵组成的集合记为 $M_{m,n}(K)$. 自然，$(a_{i1},a_{i_2},\cdots,a_{in})$ 叫做改矩阵的第 i 行，而第 j 列

$$\begin{pmatrix} a_1j \\ a_2j \\ \cdots \\ a_{mj} \end{pmatrix}$$

则用 $[a_1j, a_2j \cdots a_{mj}]$ 表示. 当矩阵是方阵时，我们称 $a_{11}, a_{22}, \cdots, a_{nn}$ 为该矩阵的**主对角线**. 如果一个方阵除了主对角线以外的值都是 0，则称该方阵 (a_{ij}) 为**对角矩阵**，有时记作

$$\mathrm{diag}(a_{11}, a_{22}, \cdots, a_{nn}),$$

并且如果 $a_{11} = a_{22} = \cdots = a_{nn} = c$ 时，则称该矩阵为**纯量矩阵**，并记为 $\mathrm{diag}(a)$. 特别地，$\mathrm{diag}(1)$ 被称为**单位矩阵**，通常记为 E_n，当矩阵阶数确定时可以记为 E.

考虑矩阵 (2.1.3) 的同时也考虑它的**增广矩阵**，即在原矩阵的最右侧加上一列 $[b_1, b_2, \cdots, b_m]$ 使之变成

$$\left(\begin{array}{cccc|c} a_{11} & a_{12} & \cdots & a_{1n} & b_1 \\ a_{21} & a_{22} & \cdots & a_{2n} & b_2 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} & b_m \end{array}\right) \tag{2.1.4}$$

或简记为 $(a_{ij}|b_i)$. 写矩阵时，该竖线可要可不要，为了清楚可以画上该竖线.

如果用实数 x_i^0 代替未知数 x_i 时，方程组 (2.1.1) 中每一个方程都变成了恒等式，则称 n 个数 $x_1^0, x_2^0, \cdots, x_n^0$ 为方程组(2.1.1)的一个**解**, 而称 x_i^0 称为解的**第 i 个分量**. 没有解的方程组方程组叫作**不相容的**. 有解的方程组叫**相容的**. 如果该方程组只有唯一解，则叫它**确定的**. 解的个数多于一个的方程组叫作**不定的**.

注 2.1.1 一个方程组是否相容，如果相容，它的解是什么样的，这就是我们当前应该回答的问题.

2.1.1　线性方程组的等价

假设我们还有一个与方程组(2.1.1)未知数个数相同的线性方程组

$$\begin{aligned} a'_{11}x_1 + a'_{12}x_2 + \cdots + a'_{1n}x_n &= b'_1, \\ a'_{21}x_2 + a'_{22}x_2 + \cdots + a'_{2n}x_n &= b'_2, \\ &\cdots\cdots \\ a'_{m1}x_1 + a'_{m2}x_2 + \cdots + a'_{mn}x_n &= b'_m. \end{aligned} \tag{2.1.5}$$

如果方程组中除了第 i, k 个之外的方程所有方程组不动，然后把第 i, k 个方程交换位置，则称方程组(**??**)由方程组(2.1.1)通过 **(I) 型初等变换**得到的. 如果把方程组的第 k 个方程加上第 i 个方程的 λ 倍，$\lambda \in K$ 则称为 **(II) 型初等变换**. 如果单把第 k 个方程左右乘上 $\lambda \in K$ 则称为 **(III) 型初等变换**. 上述变换在后面会帮助我们解方程.

定义 2.1.1 如果两个线性方程组同时相容或不相容，且相容时有相同的解，则称这两个线性方程组**等价.**

定理 2.1.1 如果一个方程组由另一个线性方程组通过有限多次初等变换得到，则这两个线性方程组等价.

证明 如果线性方程组(2.1.1)通过 (I) 型初等变换得到方程组(2.1.5)，即第 i, j 个方程互换位置. 只需将新线性方程组的第 i, j 个方程再次换位置即可得到原线性方程组.

如果线性方程组(2.1.1)通过 (II) 型初等变换得到方程组(2.1.5)，即第 k 个方程加上第 i 个方程的 λ 倍. 只需将新线性方程组第 k 个方程加上第 i 个方程的 $-\lambda$ 倍即得到原来的方程. 最后 (III) 型初等变换只需两边同时乘以 $\frac{1}{\lambda}$ 即可变回原方程.

其实上述证明就是在说明线性方程组的行初等变换是**可逆的.**

接下来，证明方程组(2.1.1)的解 $(x_1^0, x_2^0, \cdots, x_n^0)$ 也是方程组 (2.1.5) 的解. 显然，如果做的 (I) 型初等变换，仅仅是换了两个方程的位置，两个线性方程组的解显然相同的. 如果做的是 (II) 型初等变换，也仅仅是第 k 的方程变成了

$$(a_{k1}+\lambda a_{i1})x_1+(a_{k2}+\lambda a_{i2})x_2+\cdots+(a_{kn}+\lambda a_{in})x_n=b_k+\lambda b_i.$$

该方程等价于

$$(a_{k1}x_1+a_{k2}x_2+\cdots a_{kn}x_n)+(\lambda a_{i1}x_1+\lambda a_{i2}x_2+\cdots \lambda a_{in}x_n)=b_k+\lambda b_i.$$

由于原方程组的解已经满足除了第 k 个方程外的所有方程. 又因

$$\begin{gathered} a_{k1}x_1^0+a_{k2}x_2^0+\cdots a_{kn}x_n^0=b_k \\ a_{i1}x_1^0+a_{i2}x_2^0+\cdots a_{in}x_n^0=b_i. \end{gathered}$$

只需将上述第二个方程等式左右乘一个 λ，再加到第一个方程就得到了我们的第 k 个方程. 故解也相同. 由于所有初等变换也是可逆的，用同样的办法我们也能证明新线性方程组的解与原线性方程组的解相同. □

然后又注意到，当我们在做初等变换时，未知数其实是没有在变化的，仅仅是系数在不断变化. 也就是说，我们可以把方程组看为

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

我们发现，把每个系数前面的矩阵和常数矩阵拿出来，拼成一个大矩阵，就是我们前面系数的增广矩阵. 于是我们只需对系数矩阵做初等变换，再还原回线性方程组的样子，就可以得到方程组的解. 特别注意，我们在对线性方程组做初等变换时，都是以方程为单位做变换. 这说明，系数矩阵在做类似的初等变换时，只能对行做类似的初等变换.

定义 2.1.2 (矩阵的行初等变换) 设矩阵 $A\in M_{m,n}(K)$. 如果将第 i,j 两行互换位置，则称为**矩阵的 (I) 型行初等变换**. 如果为第 i 行加上第 j 行的 λ 倍，则称为**矩阵的 (II) 型行初等变换**，如果是将第 i 行乘以一个 λ，则称为**矩阵的 (III) 型行初等变换**. 三种初等变换可以统称为矩阵的行初等变换.

接下来我们先解一个线性方程组.

例 2.1.1

$$\begin{cases} \quad 2x_2 - x_3 = 1, \\ x_1 - x_2 + x_3 = 0, \\ 2x_1 + x_2 - x_3 = -2. \end{cases}$$

调换 1,2 两方程的位置得

$$\begin{cases} x_1 - x_2 + x_3 = 0, \\ \quad 2x_2 - x_3 = 1, \\ 2x_1 + x_2 - x_3 = -2. \end{cases}$$

再把第 3 个方程加上第 1 个方程的 -2 倍，得

$$\begin{cases} x_1 - x_2 + x_3 = 0, \\ \quad 2x_2 - x_3 = 1, \\ \quad 3x_2 - 3x_3 = -2. \end{cases}$$

上面两次初等变换得目的是只有第 1 个方程有 x_1，其它两个方程不出现 x_1.

再将第 3 个方程加上第 2 个方程得 -1 倍得到

$$\begin{cases} x_1 - x_2 + x_3 = 0, \\ 2x_2 - x_3 = 1, \\ x_2 - 2x_3 = -3. \end{cases}$$

再将 2,3 两个方程互换位置得到

$$\begin{cases} x_1 - x_2 + x_3 = 0, \\ x_2 - 2x_3 = -3, \\ 2x_2 - x_3 = 1. \end{cases}$$

最后将第 3 个方程加上第 2 个方程的 -2 倍得

$$\begin{cases} x_1 - x_2 + x_3 = 0, \\ x_2 - 2x_3 = -3, \\ 3x_3 = 7. \end{cases}$$

观察最后所得的方程组. 自上而下看，未知量的个数依次减少，形成一个阶梯型. 只要从它的第 3 个方程解出 x_3，带入第 2 个方程解出 x_2，再代入第 1 个方程解出 x_1. 得到方程组的解

$$x_1 = -\frac{2}{3}, x_2 = \frac{5}{3}, x_3 = \frac{7}{3}.$$

前面我们说到，线性方程组在做初等变换时，未知数其实并没有改变，于是它等价于对矩阵做初等行变换. 我们将系数矩阵的增广矩阵拿出来得到

$$\begin{pmatrix} 0 & 2 & -1 & 1 \\ 1 & -1 & 1 & 0 \\ 2 & 1 & -1 & -2 \end{pmatrix}.$$

交换 1，2 两行得到

$$\begin{pmatrix} 1 & -1 & 1 & 0 \\ 0 & 2 & -1 & 1 \\ 2 & 1 & -1 & -2 \end{pmatrix}.$$

再将第 3 行加上第 1 行的 -2 倍得到

$$\begin{pmatrix} 1 & -1 & 1 & 0 \\ 0 & 2 & -1 & 1 \\ 0 & 3 & -3 & -2 \end{pmatrix}.$$

再将上面矩阵的第 3 行加上第 2 行的 -1 倍得

$$\begin{pmatrix} 1 & -1 & 1 & 0 \\ 0 & 2 & -2 & -3 \\ 0 & 2 & -1 & 1 \end{pmatrix}.$$

最后把第三行加上第二行的 -2 倍得

$$\begin{pmatrix}1 & -1 & 1 & 0\\ 0 & 1 & -2 & -3\\ 0 & 0 & 3 & 7\end{pmatrix}.$$

再将矩阵还原为方程组

$$\begin{cases}x_1 - x_2 + x_3 = 0,\\ \quad x_2 - 2x_3 = -3,\\ \qquad 3x_3 = 7.\end{cases}$$

一样可得答案.

上述第二种方法也称为**矩阵消元法**. 在熟悉了矩阵消元法的思路和计算格式后，在实际计算中可以把几次初等行变换一次完成.

如果遇到方程类似于

$$\begin{cases}x_1 - \frac{8}{15}x_2 - \frac{2}{3}x_3 = 0\\ \quad -\frac{8}{15}x_2 - \frac{5}{6}x_3 = 0\end{cases} \tag{2.1.6}$$

. 这时解得

$$x_2 = \frac{25}{16}x_3, x_1 = \frac{3}{2}x_3.$$

我们发现有一个未知量不受限制. 则称该未知量为**自由未知量**. 上面那个解中，x_3 就是自由未知量. 当然，自由未知量的取值不唯一.

2.2 齐次线性方程组

在最开始我们已经定义了齐次线性方程组. 不难看出，齐次线性方程组

$$\begin{gathered}a_{11}x_1 + a_{12}x_2 + \cdots + a_{1_n}x_n = 0,\\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2_n}x_n = 0,\\ \cdots\cdots\\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = 0.\end{gathered}$$

显然有一组解

$$x_1 = 0, x_2 = 0, \cdots, x_n = 0.$$

这组解称为**零解**或**平凡解**. 除此之外如果还存在其它解的话，则称为**非零解**或**非平凡解**.

齐次线性方程组需要讨论的问题是它再什么情况下有非零解，什么情况下只有零解.

命题 2.2.1 数域 K 上的上述齐次线性方程组中，如果方程的个数 m 小于未知量个数 n，则它必有非零解.

证明 对方程个数 m 做数学归纳法. 当 $m = 1$ 时，若 $a_{11} = 0$，则令 $x_1 = 1, x_2 = x_3 = \cdots = x_n = 0$ 即为方程的一组非零解.

现设有 $m-1$ 个方程的齐次线性方程组，当 $m-1$ 小于未知量个数时必非零解. 然后对于有 m 个方程时做讨论.

若上述齐次线性方程组中 x_1 的系数全为 0，则取 $x_1=1, x_2=\cdots=x_n=0$ 作为方程的一组非零解. 否则，对方程做 (I) 型初等变换，总可使第一个方程 x_1 的系数不为 0. 这时再将第一个方程适当倍数加到其它方程上，可把上述方程化为如下形式:

$$\begin{cases} \overline{a_{11}}x_1+\overline{a_{12}}x_2+\cdots+\overline{a_{1_n}}x_n=0, \\ \qquad\quad \overline{a_{22}}x_2+\cdots+\overline{a_{2_n}}x_n=0, \\ \qquad\quad \cdots\cdots\cdots \\ \qquad\quad \overline{a_{m2}}x_2+\cdots+\overline{a_{mn}}x_n=0. \end{cases}$$

上述方程后面 $m-1$ 个方程是有 $n-1$ 个未知量 $x_2,\cdots,x_n$ 和 $m-1$ 个方程的齐次线性方程组. 因为 $m<n$ 故 $m-1<n-1$，按归纳假设，必有一组非零解

$$x_2=k_2,\cdots,x_n=k_n.$$

再把它带回原来未经变换的方程，因 $\overline{a_{11}}\neq 0$ 可解出唯一解 $x_1=k_1$. 于是就得到了原方程组的一组非零解. □

在本章最后，提醒读者注意，在使用矩阵消元法解线性方程组时，只进行加减乘除四种运算，该运算对数域 K 是封闭的，故在初等变换后得到的方程组仍然为数域 K 上的方程组. 所求出的解也是数域 K 上的数. 如果有自由未知量，则自由未知量限制在数域 K 上讨论.

第三章　矩阵与向量空间

3.1　n 维向量空间

定义 3.1.1 设 $\mathbb{K}$ 是一个数域.$\mathbb{K}$ 中 n 个数 $a_1, \cdots, a_n$ 所组成的一个 n 元有序数组

$$\alpha = (a_1, a_2, \cdots, a_n), a_i \in \mathbb{K}, i = 1, 2, \cdots, n$$

称为一个 m **维行向量**，a_i 称为它的第 i 个**分量**或**坐标.**

由所有 n 维行向量组成的集合称为 n **维行向量空间**或 n **维行线性空间**，记为 $\mathbb{K}^n$.

其实，上述的行向量可以依次从上往下写为一竖，即

$$\begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix}.$$

即一个 n 维**列向量**，那么其组成的向量空间为一个 n **维列向量空间**或 n **维列线性空间.** 同样还是记为 $\mathbb{K}^n$. 一般行列可以互相转换，所以记号一般不作区分. 如果要区分，会在前面加以文字说明. 有些书可能会将列线性空间记为 $\mathbb{K}^{n\times 1}$. 此外列向量还可记为 $[a_1, a_2, \cdots, a_n]$. 行向量空间和列向量空间统称为**向量空间.** 有时不做区分.

事实上，向量 $X = (x_1, x_2, \cdots, x_n)$ 可以看成是一个 $1 \times n$ 矩阵. 设 $Y = (y_1, y_2, \cdots, y_n)$ 也是一个向量，$\lambda \in K$ 是一个纯量，于是我们定义向量的加法和数乘.

$$\begin{aligned} X + Y &:= (x_1 + y_1, x_2 + y_2, \cdots, x_n + y_n), \\ \lambda X &:= (\lambda x_1, \lambda x_2, \cdots, \lambda x_n). \end{aligned}$$

此外我们将 $X + (-Y)$ 定义为 $X - Y$. 零向量 $(0, 0, \cdots, 0)$ 今后用 0 表示.

由定义，我们可以证明以下运算法则.

1. $X + Y = Y + X$ 对任意向量 $X, Y \in \mathbb{K}^n$ 成立.
2. $(X + Y) + Z = X + (Y + Z)$ 对任意三个向量 $X, Y, Z \in \mathbb{K}^n$ 成立.
3. 存在一个特别的零向量 0，使得 $X + 0 = 0 + X = X$ 对所有的 $X \in \mathbb{K}^n$ 成立.
4. 每一个向量 $X \in K^n$ 都存在一个负元 $(-X) \in \mathbb{K}^n$ 使得 $X + (-X) = (-X) + X = 0$ 成立.
5. $1X = X1 = X$ 对所有的 $X \in \mathbb{K}^n$ 成立.

6. $(\alpha\beta)X = \alpha(\beta X)$ 对所有的 $\alpha, \beta \in \mathbb{K}, X \in \mathbb{K}^n$ 成立.

7. $(\alpha+\beta)X = \alpha X + \beta X$.

8. $\alpha(X+Y) = \alpha X + \alpha Y$.

我们之所以称全体 n 维向量组成的集合为向量空间，是因为向量之间存在着运算关系，并且集合对上述运算是封闭的.

3.2 线性组合与线性子空间

3.2.1 线性组合与线性相关

定义 3.2.1 设 $X_1, X_2, \cdots, X_k \in \mathbb{K}^n$，纯量 $\alpha_1, \alpha_2.\alpha_k \in \mathbb{K}$. 向量 $Y = \alpha_1 X_1 + \alpha_2 X_2 + \cdots + \alpha_k X_k$ 称为向量 $X_1, X_2, \cdots, X_k$ 带有系数 $\alpha_1, \cdots, \alpha_k$ 的**线性组合**.

于是，我们就可以把前面的线性方程组转换为关于列向量 α_i, β 的向量方程

$$\alpha_1 x_1 + \alpha_2 x_2 + \cdots + \alpha_n x_n = \beta.$$

这里可称 β 是向量 $X_1, X_2, \cdots, X_n$ 的**线性表示**. 如果向量组 $\{X_1, X_2, \cdots, X_n\}$ 中的每一个向量都可以被 $\{Y_1, Y_2, \cdots, Y_k\}$ 线性表示，则称 $\{X_1, X_2, \cdots, X_n\}$ 可被 $\{Y_1, Y_2, \cdots, Y_k\}$ 线性表示.

定义 3.2.2 (线性相关) 关于空间 $\mathbb{K}^n$ 中的向量组 $X_1, X_2, \cdots, X_n$，如果可以找到 n 个 $\mathbb{K}$ 中不全为 0 的数 $a_1, a_2, \cdots, a_n$,，使得

$$a_1 X_1 + a_2 X_2 + \cdots a_n X_n = 0.$$

则称 $X_1, X_2, \cdots, X_n$ **线性相关**. 并称上式是**非平凡的**. 如果

$$a_1 X_1 + a_2 X_2 + \cdots a_n X_n = 0 \Rightarrow a_1 = a_2 = \cdots = a_n = 0$$

则称 $X_1, X_2, \cdots, X_n$ **线性无关.**

定义 3.2.3 定义空间 $\mathbb{K}^n$ 中的**单位行向量**

$$E_1 = (1, 0, \cdots, 0), E_2 = (0, 1, 0, \cdots, 0), \cdots, E_n = (0, 0, 0, \cdots, 1).$$

所以任意向量 $X = (x_1, x_2, \cdots, x_n)$ 可唯一表示成

$$X = x_1 E_1 + x_2 E_2 + \cdots + x_n E_n.$$

所以

$$\mathbb{K}^n = \operatorname{span}\{E_1, E_2, \cdots, E_n\}.$$

将**单位列向量**记为

$$E^1 = [1, 0, \cdots, 0], E^2 = [0, 1, 0, \cdots, 0], \cdots, E^n = [0, 0, \cdots, 1].$$

不难看出，$E_1, E_2, \cdots, E_n$ 是线性无关的. 一个向量 $X \neq 0$ 总是线性无关的.$X_1, X_2, \cdots, X_n$ 线性无关的性质与向量排列的顺序无关，因为定义中的式子中的求和顺序是任意的. 上述定义和前面关于线性方程组的定义比起来更加的抽象，因为向量不涉及到具体的分量，也正因如此，它具有更大的普遍性.

我们发现，向量间的线性组合可以产生新的向量. 于是一个自然的想法就是能不能用一部分向量把其它线性空间内所有的向量表示出来？答案是肯定的.

我们发现，线性空间中的向量对运算封闭，于是给出以下定义.

3.2.2 线性子空间的定义

定义 3.2.4 (线性子空间) 由向量 $X_1, X_2, \cdots, X_k \in \mathbb{K}^n$ 的所有线性组合构成的集合 V 称为由 $X_1, X_2, \cdots, X_k$ **张成**或**生成**的**线性子空间**. 记为，$\mathrm{span}\{X_1, X_2, \cdots, X_k\}$.

之所以称它为线性子空间，是因为生成它的向量都是 $\mathbb{K}^n$ 中的向量. 可以定义任意子集 $S \subset \mathbb{K}^n$ 的线性子空间 $\mathrm{span}\ S$ 是 S 中任意有限个向量线性组合构成的集合.

显然，如果 V 是 $\mathbb{K}^n$ 的一个线性子空间，那么由 V 中任意有限个向量的线性组合生成的线性子空间 $V' = V$. 特别地，如果 $S \subset V$，那么 $\mathrm{span}\ S \subset V$. 于是 $\mathrm{span}\ S$ 可以定义为 $\mathbb{K}^n$ 中含有 S 任意向量的集合的线性子空间的交

$$\mathrm{span}\ S = \bigcap_{S \subset V} V.$$

需要验证右边线性子空间的交集还是一个线性子空间. 因为 $X, Y \in \cap V$ 于是对于所有 $\alpha, \beta \in \mathbb{K}$, 都有 $\alpha X + \beta Y \in V$. 这说明 $\alpha X + \beta Y \in \cap V$. 但是，两个线性子空间 U, V 的交 $U \cup V$ 一般来说不是一个线性子空间. 比如在 $\mathbb{R}^2$ 中，$U = \{(x, 0)|, x \in \mathbb{R}\}, V = \{(0, x)|x \in \mathbb{R}\}$.

定理 3.2.1 1. 如果向量组 $\{X_1, X_2, \cdots, X_n\}$ 的某一个部分组是线性相关的，则向量组本身也线性相关.

2. 线性无关的向量组 $\{X_1, \cdots, X_n\}$ 的任意部分组也线性无关.

3. 在线性相关的向量 $X_1, \cdots, X_k$ 中间，至少有一个向量是其余向量的线性组合.

4. 如果 $X_1, \cdots, X_n$ 中间有一个向量是其余向量的线性表示，则向量 $X_1, \cdots, X_n$ 是线性相关的.

5. 如果向量 $X_1, \cdots, X_n$ 线性无关，而向量 $X_1, \cdots, X_n, X$ 线性相关，则 X 可被 $X_1, \cdots, X_n$ 线性表示.

6. 如果向量 $X_1, \cdots, X_n$ 线性无关，而 X 不可被 $X_1, \cdots, X_n$ 线性表示，则向量组 $\{X_1, \cdots, X_n, X\}$ 是线性无关的.

证明 1. 不失一般性，设前 s 个向量 $X_1, \cdots, X_s, s < k$ 是线性相关的，即存在不全为 0 的 a_i 使得

$$a_1X_1 + \cdots + a_sX_s = 0.$$

再令 $a_{s+1} = a_{s+2} = \cdots = a_n = 0$，于是我们就得到了一个非平凡的等式

$$a_1X_1 + \cdots + a_sX_s + a_{s+1} + \cdots + a_k = 0.$$

2. 利用反证法，如果存在一个部分组是线性相关的，则由第 1 个定理可得，整个向量组本身也线性相关，得到矛盾.

3. 设 $a_1X_1 + a_2X_2 + \cdots a_nX_n = 0$ 中的 $a_n \neq 0$. 于是有

$$X_n = -\frac{a_1}{a_n}X_1 - \frac{a_2}{a_n}X_2 - \cdots - \frac{a_{n-1}}{a_n}X_n.$$

4. 设 $X_n = b_1X_1 + \cdots + b_{n-1}X_{n-1}$. 再令 $a_1 = b_1, \cdots, a_{n-1} = b_{n-1}, a_n = -1$，就得到

$$a_1X_1 + \cdots + a_nX_n = 0, a_n \neq 0.$$

5. 如果有非平凡的关系式，

$$b_1X_1 + \cdots + b_nX_n + bX = 0, b \neq 0.$$

则 X 可被 $X_1, \cdots, X_n$ 线性表示. 如果 $b = 0$，则 $X_1, \cdots, X_n$ 线性无关，得到矛盾.

6. 如果 X 不能被 $X_1, \cdots, X_n$ 线性表示，而 $X_1, \cdots, X_n, X$ 线性相关，则由上一条，X 可被 $X_1, \cdots, X_n$ 线性表示，得矛盾. □

定理 3.2.2 对于两组 $\mathbb{K}^n$ 中的向量

$$\alpha_1, \alpha_2, \cdots, \alpha_r,$$
$$\beta_1, \beta_2, \cdots, \beta_s.$$

如果 $\{\alpha_i\}_{i=1}^r$ 中的每一个向量都可以被 $\{\beta_i\}_{i=1}^s$ 线性表示，且 $r > s$，则 $\{\alpha_1, \cdots, \alpha_s\}$ 是线性相关的.

证明 设

$$\begin{aligned}
\alpha_1 &= a_{11}\beta_1 + a_{12}\beta_2 + \cdots + a_{1s}\beta_s,\\
\alpha_2 &= a_{21}\beta_1 + a_{22}\beta_2 + \cdots + a_{2s}\beta_s,\\
&\cdots\cdots\cdots\cdots\cdots\cdots\\
\alpha_r &= a_{r1}\beta_1 + a_{r2}\beta_2 + \cdots + a_{rs}\beta_s.
\end{aligned}$$

分别用 $x_1, x_2, \cdots, x_r$ 乘上面第 $1, 2, \cdots, r$ 个等式，然后相加，得

$$\sum_{m=1}^r x_m\alpha_m = \left(\sum_{m=1}^r a_{m1}x_m\right)\beta_1 + \left(\sum_{m=1}^r a_{m2}x_m\right)\beta_2 + \cdots + \left(\sum_{m=1}^r a_{m1}x_m\right)\beta_s$$

再考察线性方程组

$$\begin{cases}
\sum\limits_{m=1}^r a_{m1}x_m = a_{11}x_1 + a_{21}x_2 + \cdots + a_{r1}x_r = 0,\\
\sum\limits_{m=1}^r a_{m2}x_m = a_{12}x_1 + a_{22}x_2 + \cdots + a_{r2}x_r = 0,\\
\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\\
\sum\limits_{m=1}^r a_{ms}x_m = a_{1s}x_1 + a_{2s}x_2 + \cdots + a_{rs}x_r = 0,
\end{cases}$$

由于 $r > s$，则上述齐次线性方程组一定有非零解.

于是关系式

$$\sum_{m=1}^r x_m\alpha_m = 0$$

是非平凡的，故 $\alpha_1, \cdots, \alpha_r$ 是线性相关的. □

对上述定理使用逆否命题就可以得到下述推论.

推论 3.2.1 对于两组 $\mathbb{K}^n$ 中的向量

$$\alpha_1, \alpha_2, \cdots, \alpha_r,$$
$$\beta_1, \beta_2, \cdots, \beta_s.$$

如果 $\{\alpha_1, \cdots, \alpha_r\}$ 线性无关，且 $\{\alpha_i\}_{i=1}^r$ 中的每一个向量都可以被 $\{\beta_i\}_{i=1}^s$ 线性表示，则 $r \leqslant s$.

于是我们又可以证明以下推论.

推论 3.2.2 设 $V \subset \mathbb{K}^n$ 是一个 $X_1, X_2, \cdots, X_s$ 为基的线性子空间. 则对于 V 中任意线性无关的向量 $Y_1, \cdots, Y_r$ 都有 $r \leqslant s$.

证明 由于 $\{Y_1, Y_2, \cdots, Y_r\}$ 可被 $\{X_1, X_2, \cdots, X_s\}$ 线性表示，如果此时 $r > s$，则由**定理**(3.2.2)可得，$\{Y_1, \cdots, Y_s\}$ 线性相关，得到了矛盾，故 $r \leqslant s$. □

定理 3.2.3 $\mathbb{K}^n$ 中的每一个非零线性子空间 $V \subset \mathbb{K}^n$ 都有一组有限基. 线性子空间 V 中所有的基都含有相同个数的向量. 这个个数 $r \leqslant n$ 称为线性子空间 V 的**维数**，记为 $\dim_{\mathbb{K}} V$ 或简记为 $\dim V$.

证明 利用筛法，由条件可得线性子空间 V 中至少有一个非零向量 X_1（行或列）. 然后我们再从 V 找出一个与 X_1 线性无关的向量 X_2，把它们放在一起得到 X_1, X_2. 现在假设已经找到了一个线性无关的向量组 $\{X_1, X_2, \cdots, X_k\}$. 如果 $\mathrm{span}\{X_1, X_2, \cdots, X_k\} \neq V$，那么可以在 V 中选出一个向量 $X_{k+1} \notin \mathrm{span}\{X_1, \cdots, X_k\}$. 于是 X_{k+1} 不可以被 $X_1, \cdots, X_k$ 线性表示，即 $\{X_1, \cdots, X_k, X_{k+1}\}$ 是一个线性无关组.

这一过程可以无限制进行下去，但是所有的向量都在 $\mathbb{K}^n = \mathrm{span}\{E_1, E_2, \cdots, E_n\}$ 中，而上述引理可知，$\mathbb{K}^n$ 中任意线性无关向量组最多包含 n 个向量. 于是对于某个自然数 $r \leqslant n$，线性无关组 $X_1, \cdots, X_k, \cdots, X_r \in V$ 成为**极大的**线性无关部分组，即任取 $0 \neq X \in V$，向量组 $\{X_1, \cdots, X_r, X\}$ 都是线性相关的. 于是 X 可被 $X_1, X_2, \cdots, X_r$ 线性表示，即 $X \in \mathrm{span}\{X_1, \cdots, X_r\}$. 所以 $V = \mathrm{span}\{X_1, \cdots, X_r\}$，也就是说，$X_1, \cdots, X_r$ 构成 V 的一组基.

现在设 $Y_1, \cdots, Y_s$ 也是 V 的一组基，再根据上述定理，于是有 $s \leqslant r$. 交换 $X_1, \cdots, X_r$ 和 $Y_1, \cdots, Y_s$ 的位置，又有 $r \leqslant s$. 于是 $s = r$. □

注 3.2.1 上述所有讨论既适合行空间，也适合列空间.

从上述讨论可知，$\mathbb{K}^n$ 中的每一个非零线性子空间 V 都有一个正整数 $r \leqslant n$，被称为 V 的维数，既 $r = \dim V$. 特别地，$\dim \mathbb{K}^n = n$.

如果我们不考虑向量空间或者线性空间，而是只考虑有限个向量组成的向量组 $A = \{X_1, X_2, \cdots, X_n\}$，则该向量组中的极大线性无关组的个数 $r \in \mathbb{Z}^+$ 被称为该向量组的**秩 (rank).** 可记为 $r = \mathrm{rank}\{X_1, X_2, \cdots, X_n\} = \mathrm{rank}\ A$ 或者 $r = r(A)$. 于是显然有

$$\dim \mathrm{span}\ A = \mathrm{rank}\ A.$$

当 $V = \{0\}$ 时，认为 $\dim V = 0$.

定理 3.2.4 给定两个包含于 $\mathbb{K}^n$ 的向量组

$$\{\alpha_1, \alpha_2, \cdots, \alpha_r\} \tag{1}$$
$$\{\beta_1, \beta_2, \cdots, \beta_s\}. \tag{2}$$

如果 (1) 可以被 (2) 线性表示，则向量组 (1) 的秩小于等于向量组 (2) 的秩，即有

$$\dim \mathrm{span}\{\alpha_1, \alpha_2, \cdots, \alpha_r\} \leqslant \dim \mathrm{span}\{\beta_1, \beta_2, \cdots, \beta_s\}$$

证明 从 (1) 中取 $\alpha_{i_1}, \cdots, \alpha_{i_k}$ 作为一组基，从 (2) 中取 $\beta_{i_1}, \cdots, \beta_{i_m}$ 作为一组基. 由条件可知 $\alpha_{i_1}, \cdots, \alpha_{i_k}$ 可被 $\beta_{i_1}, \cdots, \beta_{i_m}$ 线性表示. 如果此时 $k > m$，则 $\alpha_{i_1}, \cdots, \alpha_{i_k}$ 线性相关，得到矛盾. 于是一定有 $k \leqslant m$. □

3.2.3　线性子空间的和

定义 3.2.5 设 U,V 是 $\mathbb{K}^n$ 中的两个线性子空间，线性子空间 span $(U\cup V)$ 叫做 U 与 V 的**和**:

$$U+V=\text{span }(U\cup V)=\{u+v|u\in U,v\in V\}.$$

如果 $(U\cap V)=\{0\}$，则称 $U+V$ 为**直和**，记为 $U\oplus V$.

我们接下来证明，$U+V=\text{span }(U\cup V)$.

证明 **证法一：**设 $u_1,u_2,\cdots,u_s$ 是 U 的一组基，$v_1,v_2,\cdots,v_t$ 是 V 的一组基. 于是

$$U=\text{span}\{u_1,u_2,\cdots,u_s\},V=\text{span}\{v_1,v_2,\cdots,v_t\}.$$

所以

$$U+V=\{u+v|u\in U,v\in V\}=\left\{\sum_{i=1}^{s}k_iu_i+\sum_{j=1}^{t}c_jv_j\middle|u_i\in U,v_i\in V\right\}=\text{span}\{u_1,u_2,\cdots,u_s,v_1,v_2,\cdots,v_t\}=\text{span}(U\cup V).$$

证法二：显然 $u+v,u\in U,v\in V$ 属于线性子空间 span $U\cup V$. 故 $\{u+v|u\in U,v\in V\}\subset\text{span }(U\cup V)$. 又因为 $u\in U$ 时，只需令 $v=0$，于是 $u\in\{u+v|u\in U,v\in V\}$. 同理可得 $v\in\{u+v|u\in U,v\in V\}$. 对于 $U\cap V$ 中元素，也属于上述集合，故 span $(U\cup V)\subset\{u+v|u\in U,v\in V\}$. 于是就证明了我们上述的等式. □

命题 3.2.1 设 $V=V_1\oplus V_2$，且 $X=X_1+X_2=X_1'+X_2'$ 是向量 $X\in V$ 的两种线性组合，此处 $X_1,X_1'\in V_1,X_2,X_2'\in V_2$，则有 $X_1-X_1'=X_2'-X_2\in V_1\cap V_2$，又因为 $V_1\cap V_2=0$，所以 $X_1=X_1',X_2=X_2'$.

对于是逆命题也成立. 如果对于每一个向量 $X\in V$，写法 X_1+X_2 都是唯一的，此处 $X_i\in V_i$，则 $V=V_1+V_2$ 是直和. 更一般地，如果 $V_1,\cdots,V_k\in\mathbb{K}^n$，并且对于每一个向量 $X_i\in V_i$，线性表示 $X=X_1+\cdots+X_k$ 是唯一地，则 $V=V_1+V_2+\cdots V_k$ 为直和.

证明 只需证明逆命题. □

3.3　矩阵的秩

前面我们已经将数域 $\mathbb{K}$ 上的线性方程组和 $\mathbb{K}$ 上 m 维向量空间建立了联系. 由于一个线性方程组可以用矩阵来表示，接下来我们将系统研究矩阵，再将研究矩阵所得到的结果用于线性方程组.

定义 3.3.1 给定数域 $\mathbb{K}$ 上的一个 $m\times n$ 矩阵

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

它的每一列都可以看成一个 m 维向量 $A^{(1)},A^{(2)},\cdots,A^{(n)}$，它有 n 列，组成一个 m 维向量组，称为矩阵 A 的**列向量组**. 同理，它的每一行可以看为一个 n 维行向量，它有 m 行，组成一个 n 维向量组 $A_{(1)},A_{(2)},\cdots,A_{(m)}$，称为矩阵 A 的**行向量组**.

于是线性子空间 $V_c=\text{span}\{A^{(1)},A^{(2)},\cdots,A^{(n)}\}$ 被称为矩阵 A 的**列空间**，$V_r=\text{span}\{A_{(1)},A_{(2)},\cdots,A_{(m)}\}$ 被称为矩阵 A 的**行空间**

我们前面已经定义了向量组的秩，于是我们马上可以给出如下定义.

定义 3.3.2 一个矩阵 A 的行向量组的秩称为 A **行秩**，它的列向量组的秩称为 A **列秩**. 其中行秩就是矩阵 A 行空间的维数，列秩就是矩阵 A 列空间的维数.

如果记矩阵 A 的列向量组为 $\alpha_1, \alpha_2, \cdots, \alpha_n$（这里每个向量都竖着写），则可以把 A 简单的记为

$$A = (\alpha_1, \alpha_2, \cdots, \alpha_n).$$

如果把矩阵 A 的行向量组记为 $\beta_1, \beta_2, \cdots, \beta_m$（这里每个向量都横着写），则可以把 A 记为

$$A = \begin{pmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_m \end{pmatrix}.$$

矩阵 A 行向量组和列向量组属于两个不同的向量空间，但是组成它们的分量确实一样的. 这时我们需要对 A 进行化简，从而找到行向量组的秩和列向量组的秩之间的关系.

类似于定义线性方程组的初等变换，我们定义矩阵的初等变换.

定义 3.3.3 1. 互换两行（列）的位置.

2. 把第 j 行（列）加上第 i 行（列）的 $k \in \mathbb{K}$ 倍 $(i \neq j)$.

3. 把某一行（列）乘以常数 $c \in \mathbb{K}, c \neq 0$.

上述的每一种变换都被称为矩阵 A 的**初等行（列）变换.**

注 3.3.1 如果把 A 看为某个线性方程组的增广矩阵，那么它的行变换与线性方程组的初等变换一致. 不过特别注意，解方程组时不允许做初等列变换，而矩阵本身可以做初等列变换.

矩阵的初等行（列）变换都是可逆的.

1. 如果矩阵 A 经过互换 i, j 两行（列）变成矩阵 B，则 B 再互换 i, j 两（行）就又会变回矩阵 A.
2. 如果矩阵 A 的第 j 行（列）加上第 i 行（列）的 k 倍后变成矩阵 B，则矩阵 B 的第 j 行（列）加上第 i 行（列）的 $-k$ 倍后就变回矩阵 A.
3. 如果 A 的第 i 行（行）乘以非零常数 c 后变成矩阵 B，则 B 的第 i 行（列）乘以 $\frac{1}{c}$ 即变回矩阵 A.

前面我们已经证明，线性方程组做初等变换时，方程组的解不会发生改变，类似的，我们也要证明矩阵在做初等变换时的秩也不会被改变.

命题 3.3.1 矩阵 A 的行秩在初等行变换下不变，矩阵 A 的列秩在初等列变换下也不会改变.

为了方便叙述，在证明之前我们先定义以下内容.

定义 3.3.4 给定 $\mathbb{K}^n$ 中两个向量组，如果这两个向量组可以互相线性表示，则称这两个向量组**线性等价**或**等价.**

显然，如果两个向量组等价，则这两个向量组的秩相等，张成的子空间维数相同.

证明 先证明行秩在初等行变换下不变.

1. 设 A 的行向量组 $\{\alpha_1,\alpha_2,\cdots,\alpha_m\}$. 互换 i,j 两行相当于把行向量组的 α_i,α_j 换一下位置，显然这不会改变行秩.

2. 把 A 的第 i 行乘以 $c\neq 0$，得到新矩阵的行向量组为 $\{c\alpha_1,c\alpha_2,\cdot,c\alpha_m\}$. 由于

$$\alpha_i=\frac{1}{c}(c\alpha_i),(c\alpha_i)=c\alpha_i$$

所以新向量组可以和原向量组互相线性表示，它们等价，于是秩相等.

3. 把矩阵 A 的第 j 行加上第 i 行的 k 倍后，所得新矩阵的行向量组为

$$\{\alpha_1,\cdots,\alpha_i,\cdots,\alpha_j+k\alpha_i,\cdots,\alpha_m\}.$$

因为

$$\alpha_j=\alpha_j+k\alpha_i+(-k)\alpha_i,(\alpha_j+k\alpha_i)=\alpha_j+k\alpha_i.$$

所以新向量组也可以和原向量组互相线性表示，它们等价，于是秩还是相等.

对于矩阵的初等列变换，效果也是一样的. 不做多余的证明. □

接下来我们再引进一个概念. 把矩阵 A 行列互换得到的矩阵 A^t，被称为矩阵 A 的**转置矩阵**. 在本书中，固定用 A^t 表示 A 的转置矩阵.

例 3.3.1 设

$$A=\begin{pmatrix}-1&0&1\\2&4&1\\0&1&-1\end{pmatrix}$$

则

$$A^t=\begin{pmatrix}-1&2&0\\0&4&1\\1&1&-1\end{pmatrix}.$$

不难看出，转置矩阵的列向量其实就是原矩阵的行向量. 于是原矩阵的行秩为转置矩阵的列秩.

命题 3.3.2 矩阵 A 的行秩在初等列变换下保持不变；矩阵 A 的列秩在初等行变换下保持不变.

证明 我们先证明，矩阵的列秩在初等行变化下不变.

设 A 的列向量组 $\alpha_1,\alpha_2,\cdots,\alpha_n$，其列秩为 r. 再设 $\alpha_{i_1},\cdots,\alpha_{i_r}$ 为列向量组的一个极大线性无关部分组. 假设 A 经过初等行变换后所得的新矩阵的列向量组为 $\alpha_1',\alpha_2',\cdots,\alpha_n'$. 我们只需证明，$\alpha_{i_1}',\cdots,\alpha_{i_r}'$ 是它的一个极大线性无关部分组.

于是我们先证明 $\alpha_{i_1}',\cdots,\alpha_{i_r}'$ 线性无关. 以 $\alpha_{i_1},\cdots,\alpha_{i_r}$ 为列向量排成一个矩阵 B. 由于 $\alpha_{i_1},\cdots,\alpha_{i_r}$ 线性无关，于是以 B 为系数矩阵的齐次线性方程组只有零解. 再将 $\alpha_{i_1}',\cdots,\alpha_{i_r}'$ 为列向量排成矩阵 B_1. 因为 B 是由 A 的 $i_1,\cdots,i_r$ 列组成的，所以对 A 做行变换其实也在对 B 做同样的行变换. 即 B_1 是对 B 做初等行变化得到. 由于齐次线性方程组在做初等变换前后的解相同，故以 B_1 为系数矩阵的齐次线性方程组也只有零解. 于是 $\alpha_{i_1}',\cdots,\alpha_{i_r}'$ 也线性无关.

再证任一 a_i' 可被 $\alpha_{i_1}',\cdots,\alpha_{i_r}'$ 线性表示. 考虑以 $\alpha_{i_1},\cdots,\alpha_{i_r},\alpha_i$ 为列向量的矩阵 $\bar{B}$. 因为 α_i 可以被 $\alpha_{i_1},\cdots,\alpha_{i_r}$ 表示，于是以 $\bar{B}$ 为增广矩阵的线性方程组有解. 另外，$\alpha_{i_1}',\cdots,\alpha_{i_r}',\alpha_i'$ 为列向量的矩阵 $\bar{B}_1$ 是由 $\bar{B}$ 做初等行变化得到，于是以 $\bar{B}_1$ 为增广矩阵的线性方程组和以 $\bar{B}$ 为增广矩阵的线性方程组同解. 这就说明 α_i' 可以被 $\alpha_{i_1}',\cdots,\alpha_{i_r}'$ 线性表示.

接下来证明，A 的行秩再初等列变换下不变. 为此考察 A^t.A 的列向量组是 A^t 的行向量组，于是对 A 做初等列变换等价于对 A^t 做初等行变换.A^t 的列秩在初等行变换下不变. 再设 A 经过初等列变换变成 B. 于是 A 的行秩 $=A^t$ 的列秩 $=B^t$ 的列秩 $=B$ 的行秩. 这说明 A 的行秩在初等列变换下也保持不变. □

推论 3.3.1 设 A 是数域 $\mathbb{K}$ 上的 $m\times n$ 矩阵，A 经过若干次初等行变换化为矩阵 B. 设 A 的列向量组是 $\alpha_1,\alpha_2,\cdots,\alpha_n$，$B$ 的列向量组为 $\alpha_1',\alpha_2',\cdots,\alpha_n'$. 则有如下结论：

1. 如果 $\alpha_{i_1},\alpha_{i_2},\cdots,\alpha_{i_r}$ 是 A 的列向量组的一个极大线性无关部分组，则 $\alpha_{i_1}',\alpha_{i_2}',\cdots,\alpha_{i_r}'$ 是 B 的列向量组的一个极大线性无关部分组. 而且，当
$$\alpha_i=k_1\alpha_{i_1}+k_2\alpha_{i_2}+\cdots+k_r\alpha_{i_r}$$
时，有
$$\alpha_i'=k_1\alpha_{i_1}'+k_2\alpha_{i_2}'+\cdots+r\alpha_{i_r}'.$$
2. 如果 $\alpha_{i_1}',\alpha_{i_2}',\cdots,\alpha_{i_r}'$ 是 B 的列向量的一个极大线性无关部分组，当
$$\alpha_i'=k_1\alpha_{i_1}'+k_2\alpha_{i_2}'+\cdots+k_r\alpha_{i_r}'$$
时，有
$$\alpha_i=k_1\alpha_{i_1}+k_2\alpha_{i_2}+\cdots+k_r\alpha_{i_r}.$$

证明 前一个命题直接得出这里的第一个结论，由于初等变换是可逆的，于是 B 也可变回 A，故第二个结论也成立. □

如果一个 $m\times n$ 矩阵的所有元素都为 0，则称为**零矩阵**，直接记为 0. 显然零矩阵的行秩和列秩都为 0.

第四章　群论

4.1　群的概念

现在将视角转到一般的带运算的集合上.

定义 4.1.1 (半群) 设 S 是一个非空集，S 上存在一个二元运算 $S\times S\to S$，并记为 $\cdot$. 如果对任意的 $x,y\in S$，总存在唯一的元素 $x\cdot y\in S$. 并且该运算满足结合律，则称 S 连同运算 $\cdot$ 为一个**半群 (semigroup)**.

称运算 $\cdot$ 为乘法，并且有时记 $a\cdot b$ 为 ab. 如果一个半群的乘法还满足交换律，则称该半群为**交换半群 (commutative semigroup)**.

定义 4.1.2 (幺半群) 如果一个半群 S 中存在一个元素 e，满足对于任意 $a\in S$ 有

$$ea=ae=a$$

则称该半群为**幺半群 (monoid)**. 其中 e 称为 S 的**幺元**或**恒等元 (identify element)**.

定义 4.1.3 (群) 一个幺半群 G，其幺元为 e，且 G 还满足，对于任意 $a\in G$，存在 a^{-1} 使得

$$aa^{-1}=a^{-1}a=e.$$

则称 G 连同运算构成一个**群 (group)**，记为 $(G,\cdot)$. 其 a^{-1} 称为 a 的**逆元 (inverse element)**.

如果 G 上的运算还满足交换律，则称 G 为一个 **Abel 群**或**交换群**.

定义 4.1.4 (群的阶) 一个群中如果只有有限多个元素则称该群为**有限群**，否则就是**无限群**. 通常用 $|G|$ 或者 $\mathrm{ord}(G)$ 表示 G 中元素的个数. 如果 $|G|=n$，则称 G 为 n **阶群 (n order of group)**.

注 4.1.1 先提醒一下，这里的关于群的阶的定义不要和后面 n 阶对称群的阶搞混，请注意辨别.

群是一类普遍的代数体系，我面已经学过的许多代数结构实际上都是一种群结构. 接下来看一些例子.

例 4.1.1 整数集 $\mathbb{Z}=\{0,\pm1,\pm2,\cdots,\pm n,\cdots\}$ 在通常数的加法（小学就学的那个加法）下成为一个群 $(\mathbb{Z},+)$. 这个群是加法群，幺元为 0. 对于任意的 $n\in\mathbb{Z}$，其逆元为 $-n$. 不过，整数集连同乘法运算不构成群，因为幺元为 1 而 0 没有逆元素. 但是 $\mathbb{Z}$ 连同乘法是构成半群的，且为交换幺半群.

例 4.1.2 有理数集 $\mathbb{Q}$ 在通常数的加法下也是一个加法群 $(\mathbb{Q},+)$. 其幺元为 0. 同理，实数集和复数集在数的加法下都构成群.$\mathbb{Q}$ 在数的乘法下构成交换幺半群.

$\mathbb{Q}^*:=\mathbb{Q}\backslash\{0\}$ 在数的乘法下构成一个交换群，幺元为 1，$\frac{p}{q}$ 的逆元为 $\frac{q}{p}$. 同理 $\mathbb{R}^*,\mathbb{C}^*$ 在数的乘法下都是群.

例 4.1.3 $G = \{1, -1\}$ 在数的乘法下构成一个群，其幺元为 1，所有元素的逆元都是它自身.

例 4.1.4 设 S 是一个集合，$\varphi : S \to S$ 是 S 上的一个一一对应. 将 S 上全体一一对应构成的集合记为 $A(S)$，并定义 $A(S)$ 上的运算为合成，由于映射的合成满足结合律，且恒等映射 id_S 为幺元，故 $A(S)$ 是一个幺半群.

对 $A(S)$ 中的任意双射 φ，有逆映射 φ^{-1} 满足

$$\varphi \circ \varphi^{-1} = \varphi^{-1} \circ \varphi = id_S$$

故 $A(S)$ 是一个群. 这个群称为集合 S 上的**变换群.**

例 4.1.5 在上一个例子中，如果 S 是一个有限集且有 n 个元素，我们把这样的群称为 **n 阶对称群 (n symbols symmetric group)**，通常用 S_n 表示. 由于 S 全体元素的任意一个排列决定了 S 的一个变换，而且 S 的任意一个变换决定了一个全排列，因此 $|S_n| = n!$(注意 n 阶对称群的阶不是 n).

例 4.1.6 V 是 $\mathbb{R}$ 上的 n 维向量空间，V 上的向量全体在加法下也构成一个群，其幺元为零向量.

例 4.1.7

第五章　复数和多项式

第六章　多项式的根

第二部分

线性代数

第七章　线性空间与线性变换

我们前面已经简单介绍了线性子空间，那么在本章我们正式研究线性空间与线性变换. 在本章中，我们只抓住向量空间的本质，即八条运算法则，其它非本质的表现形式则舍弃. 先看一个例子.

例 7.0.1 考察闭区间 $[a,b]$ 上全体实连续函数 $f(x)$ 组成的集合，记为 $C[a,b]$. 这个集合内任意两个元素 $f(x),g(x)$，它们可以做加法运算 $f(x)+g(x)\in C[a,b]$. 以及数乘 $kf(x)\in C[a,b]$. 从这两个性质出发，不难验证 $C[a,b]$ 中的元素满足 8 条运算法则，而 $C[a,b]$ 中的元素和 m 维向量空间中的元素完全不同. 但是又有一些共同点，两种运算都满足 8 条运算法则.

像上述的例子还可以举出很多，这也提示我们应该将向量空间进一步抽象化.

7.1　线性空间的基本概念

7.1.1　线性空间的定义

定义 7.1.1 (线性空间) 设 V 是一个非空集合,$\mathbb{F}$ 是任意域. 在 V 中定义一种二元运算(加法)$+:V\times V\to V,(\alpha,\beta)\mapsto\alpha+\beta$. 再定义 V 具有 Abel 群结构 (**空间 V 的加法群**)，以及在集合 V 上满足以下 8 条运算法则：

1. $\alpha+\beta=\beta+\alpha$ 对任意向量 $\alpha,\beta\in V$ 成立.（交换律）
2. $(\alpha+\beta)+\gamma=\alpha+(\beta+\gamma)$ 对任意三个向量 $\alpha,\beta,\gamma\in V$ 成立.（结合律）
3. 存在一个特别的**负元素** 0，使得 $\alpha+0=0+\alpha=\alpha$ 对所有的 $\alpha\in V$ 成立.
4. 每一个向量 $\alpha\in V$ 都存在一个**逆元素** $(-\alpha)\in V$ 使得 $\alpha+(-\alpha)=(-\alpha)+\alpha=0$ 成立.
5. $1\alpha=\alpha1=\alpha$ 对所有的 $\alpha\in V$ 成立.（酋性）
6. $(kl)\alpha=k(l\alpha)$ 对所有的 $k,l\in\mathbb{F},\alpha\in V$ 成立.
7. $(k+l)\alpha=k\alpha+l\alpha,k,l\in\mathbb{F},\alpha\in V.$
8. $k(\alpha+\beta)=k\alpha+k\beta,k\in\mathbb{F},\alpha,\beta\in V.$

我们就称这样的集合 V 为域 $\mathbb{F}$ 上的**线性空间 (linear space)**.

注 7.1.1 值得注意的是，像第 7 条，等号左边的加法是对域 $\mathbb{F}$ 的元素而言的，而等号右侧的符号是对于 V 中的元素而言的. 严格上讲，我们应当使用两种不同的符号作为区分，但在不引起歧义的情况下我们用同一个符号.

由于线性空间是从向量空间中抽象出来的，我们仍然称线性空间中的元素为**向量**，于是零元素称为**零向量**，负元素称为**负向量.** 值得注意的是，此时的向量已经不再是有序数组，即不再是 n 维向量，也可能没有具体的形式.

如果在第一部分讲过的一样，我们称 $\mathrm{span}M$ 是由集合 M 中的元素所**张成**的子空间，这显然是包含集合 M 最小的子空间.

由于 $\mathbb{F}$ 是一个一般的域，所以我们称 $k,l\in\mathbb{F}$ 为**纯量 (scalar)**.

下面可以给出一些基本的命题.

命题 7.1.1 设 V 是 $\mathbb{K}$ 上的线性空间，则有:

1. V 中零向量是唯一的.
2. V 中任意向量的负向量也是唯一的.
3. 加法有消去律，即 $\alpha+\beta=\alpha+\gamma\Rightarrow\beta=\gamma$.
4. 加法可以移项，即 $\alpha+\beta=\gamma\Rightarrow\alpha=\gamma-\beta$.
5. $0\alpha=0;(-1)\alpha=-\alpha,k\cdots0=0$. 其中 $k\in\mathbb{K}$.
6. 若 $k\alpha=\beta,k\neq0$ 则 $\alpha=\dfrac{1}{k}\beta$.

证明 1. 设 $0'$ 也是 V 中另一个零向量，于是

$$0'=0'+0=0+0'=0.$$

2. 设 $(-\alpha)_1,(-\alpha)$ 都为 α 的负元素，于是

$$(-\alpha)=(-\alpha)+0=(-\alpha)+(\alpha+(-\alpha)_1)=((-\alpha)+\alpha)+(-\alpha)_1=0+(-\alpha)_1=(-\alpha)_1.$$

以后我们记 $\beta+(-\alpha)$ 为 $\beta-\alpha$. 并称为**减法.**

3. 将 $\alpha+\beta=\alpha+\gamma$ 的等号两边同时加上 $-\alpha$ 得到

$$((-\alpha)+\alpha)+\beta=((-\alpha)+\alpha)+\gamma$$

得到

$$\beta=0+\beta=0+\gamma=\gamma.$$

4. 只需

$$\alpha=\alpha+0=\alpha+(\beta+(-\beta))=(\alpha+\beta)+(-\beta)=\gamma+(-\beta)=\gamma-\beta.$$

5. 关于第一个等式. $0\alpha=(0+0)\alpha+0\alpha+0\alpha$，再移项得

$$0\alpha=0\alpha-0\alpha=0.$$

关于第二个等式.$\alpha+(-1)\alpha=1\alpha+(-1)\alpha=(1+(-1))\alpha=0\alpha=0$ 由定义可知，$(-1)\alpha$ 为 α 的负元素，又由负元素的唯一性可知 $(-1)\alpha=-\alpha$.

关于第三个等式.$k0+k0=k(0+0)=k0=k0+0$，由消去律可得 $k0=0$.

6. 由于 $\alpha=1\alpha=\left(\dfrac{1}{k}k\right)\alpha=\dfrac{1}{k}(k\alpha)=\dfrac{1}{k}\beta$. 由此也一定可以推出，当 $k\neq0$ 且 $k\neq0$ 时一定有 $\alpha=\dfrac{1}{k}0=0$. □

对于前面的一些不依赖线性方程组的解的命题与定理都是成立的，我们在下面使用时会阐明. 对于一些基本概念我们再用线性空间的形式叙述一遍.

定义 7.1.2 1. 给定 V 内的一个向量组 $\alpha_1,\alpha_2,\cdots,\alpha_s$，如果对于 V 内的一个向量 β，存在 $\mathbb{K}$ 内的 s 个数 $k_1,\cdots,k_s$ 使

$$\beta = k_1\alpha_1 + k_2\alpha_2 + \cdots + k_s\alpha_s$$

则称 β 可被 $\alpha_1,\cdots,\alpha_s$ **线性表示.**

2. 给定 V 内的一个向量组 $\alpha_1,\cdots,\alpha_s$，如果存在 $\mathbb{K}$ 内不全为 0 的 s 个数 $k_1,\cdots,k_s$，使

$$k_1\alpha_1 + k_2\alpha_2 + \cdots + k_s\alpha_s = 0,$$

则称向量组 $\alpha_1,\cdots,\alpha_s$ **线性相关.** 如果

$$k_1\alpha_1 + k_2\alpha_2 + \cdots + k_s\alpha_s = 0 \Rightarrow k_1 = k_2 = \cdots = k_s = 0,$$

则称向量组 $\alpha_1,\cdots,\alpha_s$ **线性无关.**

3. 给定 V 内两个向量组

$$\alpha_1,\alpha_2,\cdots,\alpha_r, \tag{1}$$

$$\beta_1,\beta_2,\cdots,\beta_s, \tag{2}$$

如果 (1) 中的任意向量都能被 (2) 线性表示，且 (2) 中的向量都能被 (1) 线性表示，则称两个向量组**等价.**

4. 给定 V 中的一个向量组 $\alpha_1,\cdots,\alpha_s$，如果它有一个部分组 $\alpha_{i_1},\cdots,\alpha_{i_r}$，满足 $\alpha_{i_1},\cdots,\alpha_{i_r}$ 线性无关且原向量组中任意向量都可以被这个部分组表示. 则称此部分组为原向量的一个**极大线性无关部分组.** 一个向量组的任意极大线性无关部分组都有相同个数的向量，该向量数目称为该向量组的**秩 (rank).**

在之前我们可以用线性方程组有无非零解来判断 $\mathbb{K}^m$ 上的向量组是否线性相关，但是对于一般的线性空间就不能再使用该方法了，而是要具体情况具体分析.

例 7.1.1 在 $\mathbb{R}$ 上线性空间 $C[a,b]$ 内给定向量组

$$e^{\lambda_1 x}, e^{\lambda_2 x}, \lambda_1 \neq \lambda_2$$

证明该向量组线性无关.

证明 按照定义写出

$$k_1 e^{\lambda_1 x} + k_2 e^{\lambda_2 x} = 0.$$

我们先分析等式的含义，两个元素都为 $[a,b]$ 上的连续函数，0 为恒等于 0 的函数. 下面采用反证法，如果有 $k_1 \neq 0$ 于是可得

$$e^{(\lambda_1-\lambda_2)x = -\frac{k_2}{k_1}}.$$

由于 $\lambda_1 \neq \lambda_2$，左端为指数函数，而右边为常数函数，这与指数函数为单调函数矛盾.

下面再看一个证明方法，对式子等号两边同时求导可得

$$k_1\lambda_1 \mathrm{e}^{\lambda_1 x} + k_2\lambda_2 \mathrm{e}^{\lambda_2 x} = 0.$$

上述两个式子联立，看为关于 k_1,k_2 的齐次线性方程组. 其系数矩阵的行列式为

$$\begin{vmatrix} e^{\lambda_1 x} & e^{\lambda_2 x} \\ \lambda_1 e^{\lambda_1 x} & \lambda_2 e^{\lambda_2 x} \end{vmatrix} = e^{(\lambda_1+\lambda_2)x}(\lambda_2-\lambda_1).$$

而对于 (a,b) 内的任意 x 均不为 0，k_1,k_2 对于 $\forall x\in(a,b)$ 都成立，于是只有 $k_1=k_2=0$.

例 7.1.2 (将域本身作为一维空间) 定义 $V=\mathbb{F}.$，V 的基本运算与 $\mathbb{F}$ 的运算一致. 如果 1 是域的单位元，可以认为 $\mathbb{F}=\mathrm{span}\{(1)\}$，是 (1) 张成的线性空间. 其实 $\mathbb{C}$ 可以看作是 $\mathbb{R}$ 上的线性空间. 更一般地，如果域 $\mathbb{F}$ 是自己子域 $\mathbb{B}$ 的一个扩域，则 $\mathbb{F}$ 是域 $\mathbb{B}$ 上的线性空间.

例 7.1.3 (函数空间) 定义函数环 $R^X=\{f:X\to R\}$，在把环 R 用域 $\mathbb{F}$ 代替后，再在定义域上逐点定义加法和纯量乘

$$\begin{aligned}(f+g)(x)&\coloneqq f(x)+g(x),\quad x\in X,\\ (kf)(x)&=k(f(x)),\quad x\in X.\end{aligned}$$

这样的集合 $\mathbb{F}^X$ 其实是 $\mathbb{F}$ 上的线性空间.

我们再定义 $\{x\}$ 上的 Δ 函数 δ_x，即当 $x_1=x$ 时有

$$\delta_x(x_1)=1,$$

当 $x_2\neq x$ 时有

$$\delta_x(x_2)=0.$$

如果 $X=\{1,2,\cdots,n\}$，此时则有 $F^X=F^n$，函数 f 就可以表达成它所有值的向量 $(f(1),f(2),\cdots,f(n))$，而且可以把函数 f 的单个值表达成 Δ 函数的线性组合

$$f=f(1)\delta_1+f(2)\delta_2+\cdots+f(n)\delta_n.$$

对于无穷集 X，我们一般无法这样表达，因为无穷多个向量的和一般没有意义 (除非特别涉及拓扑学).

类似地，我们可以研究定义在整个直线或者区间 (a,b) 上的函数空间 $\mathbb{R}^{(a,b)}$. 其连续函数张成的空间 $C(a,b)$ 是 $\mathbb{R}^{(a,b)}$ 的一个子空间，其所有 (a,b) 上所有可微函数张成的空间也是它的一个子空间.

例 7.1.4 (Amer.Math.Monthly,1990,V.94,p60-62) 设方阵 $A\in M_n(\mathbb{Q})$，如果 A 的每一行的有理数之和与每一列的有理数之和都相同，且都等于一个有理数 $\sigma(A)$，则称 A 是一个**半幻方 (semi-magic square)** 或**半幻的**. 如果还有

$$\sum_{i=1}^{n} a_{i,n+1-i}=\sigma(A)$$

则称 A 是**魔幻方 (magic square)** 或**魔幻的**. 我们把 $\mathbb{Q}$ 上的所有 n 阶半幻方组成的集合记为 $\mathrm{SMag}_n(Q)$，所有 n 阶魔幻方组成的集合记为 $\mathrm{Mag}_n(\mathbb{Q})$. 不难看出，当两个半幻方相加后，其每一行每一列的数相加还都是相等的，所以得到的新方阵还是半幻的. 对于魔幻方也是如此. 所以不难验证 $\mathrm{SMag}_n(Q),\mathrm{Mag}_n(\mathbb{Q})$ 都是 $\mathbb{Q}$ 上的线性空间，并且还有

$$\mathrm{Mag}_n(\mathbb{Q})\subset\mathrm{SMag}_n(\mathbb{Q})\subset M_n(\mathbb{Q}).$$

关于线性空间的几何解释说明

如果 $\mathbb{F}=\mathbb{R}$，我们就称线性空间 V 是**实的**，如果 $\mathbb{F}=\mathbb{C}$，则称 V 为**复的**. 在平时的使用中，这两种情形最让人感兴趣，尽管很多理论并不是域 $\mathbb{F}$ 的特性.

在我们生活的三维空间中，从一个固定的点引出的所有线段的集合可以作为一个线性空间的模型. 我们用数 $k\in\mathbb{R}$ 来乘一条线段，可以让这条线段边长 $k>1$ 倍，也可以压缩 $k\leqslant 1$ 倍，当 $k<0$ 时取反方向. 有向线段的加法按照平行四边形法则建立. 在约定两个有向线段平移后能重合即为相等后，这个实线性空间与自由几何向量的集合相同.

物理的三维空间 $\mathbb{R}^3_\Phi$ 中的对象可以利用图像来表达. 如果将维数升高，我们很难再通过直觉对对象进行描述. 这时我们有必要系统地建立起空间与几何之间的关系，使它们形成稳定的结合，使得理论更具有活力. 例如 $\mathbb{F}=\mathbb{C}$ 时，$\mathbb{C}$ 上的直线就是 1 维坐标空间 $\mathbb{C}^1$. 然后我们可以把复平面 $\mathbb{R}^2$ 作为它的几何表示. 数 $z=x+iy\in\mathbb{C}^1$ 对应于点 $(x,y)\in\mathbb{R}^2$. 如果用 $\mathbb{C}\ni a\neq 0$ 乘它，则可以把复平面上点到原点连线延长 $|a|$ 倍再旋转角度 $\arg a$. 所以，在后面我们会类似的知道 n 维复空间 $\mathbb{C}^n$ 可以表示成 $2n$ 维的实空间 $\mathbb{R}^{2n}$.

值得一说的是，3 维物理空间 $\mathbb{R}^3_\Phi$ 比同样维数的坐标空间内容要丰富得多，因为在 3 维物理空间中定义了向量的长度，向量间的角度，图形的面积和体积. 这些定义的内容在不知不觉中就会用于定义抽象空间的图形上，这个抽象空间的公理系统暂时还没有给出，等到度量概念丰富起来后再研究这样的抽象空间.

我们再来说说线性空间 V 受纯量域性质的影响. 如果 $\mathbb{F}$ 是一个有限域，那么由 $\mathbb{F}^3$ 带来的几何结构是“有窟窿的”，这来自 $\mathbb{F}$ 离散性的推论. 但 $\mathbb{F}$ 的离散性有时可以结合 $\mathbb{F}$ 上的线性几何学，从而可以使用一种离散图. 比方说，在二元域 $\mathbb{F}_2=\{0,1\}$ 上的 n 维坐标空间 $\mathbb{F}_2^n$ 允许与 $\mathbb{R}^n$ 中 n 维立方体顶点的集合 $\{(\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n)|\varepsilon_i=1\vee\varepsilon_i=0\}$. 自然的等同起来.（尝试画图）

7.1.2　基和维数

接下来介绍线性空间的一个核心概念：基和维数. 先给出一个基本命题.

命题 7.1.2 设 V 是 $\mathbb{F}$ 上的一个线性空间. 如果在 V 中存在 n 个线性无关的向量

$$\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$$

使得 V 中任一向量均能被上述向量线性表示，那么有

1. 任给 V 中 n 个线性无关向量

   $$\eta_1,\eta_2,\cdots,\eta_n$$

   则 V 中任意向量也可以被上述向量线性表示.

2. 如果 $\alpha_1,\cdots,\alpha_s$ 是 V 中的一个线性无关向量组，且 V 中任一向量可以被它们线性表示，则 $s=n$.

证明 任给 $\beta\in V$，考察向量组

$$\eta_1,\eta_2,\cdots,\eta_n,\beta$$

它们可以被 $\varepsilon_1,\cdots,\varepsilon_n$ 线性表示，并且个数大于 n，于是它们必定线性相关. 并且 β 可以被 $\eta_1,\cdots,\eta_n$ 线性表示.

如果 $\alpha_1,\cdots,\alpha_s$ 和 $\varepsilon_1,\cdots,\varepsilon_n$ 线性等价，则它们可以互相线性表示，既秩相等. □

定义 7.1.3 设 V 是 $\mathbb{F}$ 上的一个线性空间. 如果 V 中存在 n 个线性无关的向量

$$\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n.$$

并且 V 中的任意向量都可以被此向量组线性表示，则称 V 为 **n 维线性空间**，记作 $\dim V = n$. 上述向量组称为 V 的一组**基.** 特别地，单由零向量组构成的线性空间称为**零空间.** 定义零空间的维数为 0.

由前一个命题立即可以得到，在 n 维线性空间中，任意 n 个线性无关的向量都是它的一组基. 反过来说，它的任意一组基都恰好有 n 个向量. 显然在一个 n 维线性空间中，任意 $n+1$ 个向量都是线性相关的，所以在 n 维线性空间中最多只有 n 个线性无关的向量.

同时，如果一个线性空间中有 n 个线性无关的向量，而任意 $n+1$ 个向量都是线性相关的，则它为 n 维线性空间.

如果一个线性空间中存在任意多个线性无关的向量，则称为**无限维线性空间**. 而 n 维线性空间称为**有限维线性空间.** 我们目前主要研究后者.

下面的定理指出用如何由给定的基实际地建立新的基.

定理 7.1.1 设 V 是域 $\mathbb{F}$ 上的线性空间，有基 $\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n$，则有

1. 每个向量 $v \in V$ 都可以唯一的表示成向量 $\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n$ 的线性组合.
2. V 的任意线性无关向量组 $f_1, f_2, \cdots, f_s, s \leqslant n$ 都可以扩充成 V 的一组基. 特别地，任意一个非零向量 v 都可以包含于某一组基之中.

证明 1. 把给定的向量 $v \in V$ 用给定的基表示，得到

$$\alpha v + \alpha_1 \varepsilon_1 + \cdots + \alpha_n + \varepsilon_n = 0,$$

且 $\alpha \neq 0$，于是有

$$v = (-\alpha^{-1}\alpha_1)\varepsilon_1 + \cdots + (-\alpha^{-1}\alpha)\varepsilon_n.$$

于是我们证明了，任意 $v \in V$ 可表示为基的线性组合.

再证明唯一性，如果 v 有两个分解式

$$\beta_1\varepsilon_1 + \cdots + \beta_n\varepsilon_n = v = \gamma_1\varepsilon_1 + \cdots + \gamma_n\varepsilon_n,$$

则将两式相减得到

$$(\beta_1 - \gamma_1)\varepsilon_1 + \cdots + (\beta_n - \gamma_n)\varepsilon_n = 0.$$

由于基向量的线性无关性，我们有

$$\beta_1 - \gamma_1 = \cdots = \beta_n - \gamma_n = 0.$$

于是我们就证明了唯一性.

2. 考察向量组

$$f_1, \cdots, f_s, \varepsilon_1, \cdots, \varepsilon_n,$$

从中选出一些向量，使得每个向量都不能由它前面的那些向量表示出来. 根据条件，$f_1, \cdots, f_s$ 线性无关，及其任意部分组线性无关，所以它的每一个向量都不可以用它前面的向量表示出来. 从而可以假设选出的向量构成的向量组如下:

$$f_1, \cdots, f_s, \varepsilon_{i_1}, \cdots, \varepsilon_{i_s}.$$

对于任意非平凡关系式

$$\alpha_1 f_1 + \cdots + \alpha_s f_s + \beta_1\varepsilon_{i_1} + \beta_t\varepsilon_{i_t} = 0,$$

如果 $\beta_k \neq 0, k < t$，且 k 是脚标中的最大者，ε_{i_k} 可被它前面的向量线性表示，但是按照前面的选取方式，这是不可能的. 另外，从 1. 可以得到 V 中的每一个向量均可以由基 $\varepsilon_1, \cdots, \varepsilon_n$ 线性表示，从而它就能由 $f_1, f_2, \cdots, f_s, \varepsilon_1, \cdots, \varepsilon_n$ 线性表示，进而它可由 $f_1, \cdots, f_s, \varepsilon_{i_1}, \cdots, \varepsilon_{i_t}$ 线性表示. 这样我们就找到了 f_i 的补足者 $e_{i_1}, \cdots, e_{i_t}$. □

注 7.1.2 在证明 2. 的时候，我们用到的论断实际上就是 Steinitz 替换定理. 即在 $\mathbb{F}$ 上的线性空间 V 中，向量 $u_1,\cdots,u_m$ 线性无关且可由向量 $v_1,\cdots,v_n$ 线性表示，则 $m\leqslant n$，并且用 $u_1,\cdots,u_m$ 替换 $v_1,\cdots,v_n$ 中某 m 个向量后所得的向量组于原向量组等价.

例 7.1.5 证明 $C[a,b]$ 是 $\mathbb{R}$ 上的一个无限维线性空间.

证明 对于任意 $n\in\mathbb{N}^+$，取 n 个两两不等的实数 $\lambda_1,\lambda_2,\cdots,\lambda_n$，我们证明 $C[a,b]$ 内的向量组

$$\mathrm{e}^{\lambda_1 x},\mathrm{e}^{\lambda_2 x},\cdots,\mathrm{e}^{\lambda_n x}$$

线性无关.

设

$$k_1\mathrm{e}^{\lambda_1 x}+k_2\mathrm{e}^{\lambda_2 x}+\cdots+k_n\mathrm{e}^{\lambda_n x}=0,$$

在 (a,b) 内求 $n-1$ 次导数得到

$$\begin{aligned}
&k_1\lambda_1\mathrm{e}^{\lambda_1 x}+k_2\lambda_2\mathrm{e}^{\lambda_2 x}+\cdots+k_n\lambda_n\mathrm{e}^{\lambda_n x}=0,\\
&k_1\lambda_1^2\mathrm{e}^{\lambda_1 x}+k_2\lambda_2^2\mathrm{e}^{\lambda_2 x}+\cdots+k_n\lambda_n^2\mathrm{e}^{\lambda_n x}=0,\\
&\cdots\cdots\cdots\cdots\\
&k_1\lambda_1^{n-1}\mathrm{e}^{\lambda_1 x}+k_2\lambda_2^{n-1}\mathrm{e}^{\lambda_2 x}+\cdots+k_n\lambda_n^{n-1}\mathrm{e}^{\lambda_n x}=0.
\end{aligned}$$

把 $k_1,k_2,\cdots,k_n$ 看作未知数，于是上述方程是 n 个方程的齐次线性方程组，其系数矩阵的行列式为.

$$\begin{aligned}
\begin{vmatrix}
\mathbf{e}^{\lambda_1 x} & \mathbf{e}^{\lambda_2 x} & \cdots & \mathbf{e}^{\lambda_n x}\\
\lambda_1\mathbf{e}^{\lambda_1 x} & \lambda_2\mathbf{e}^{\lambda_2 x} & \cdots & \lambda_n\mathbf{e}^{\lambda_n x}\\
\lambda_1^2\mathbf{e}^{\lambda_1 x} & \lambda_2^2\mathbf{e}^{\lambda_2 x} & \cdots & \lambda_n^2\mathbf{e}^{\lambda_n x}\\
\vdots & \vdots & & \vdots\\
\lambda_1^{n-1}\mathbf{e}^{\lambda_1 x} & \lambda_2^{n-1}\mathbf{e}^{\lambda_2 x} & \cdots & \lambda_n^{n-1}\mathbf{e}^{\lambda_n x}
\end{vmatrix}
&=\mathrm{e}^{\lambda x}\begin{vmatrix}
1 & 1 & \cdots & 1\\
\lambda_1 & \lambda_2 & \cdots & \lambda_n\\
\lambda_1^2 & \lambda_2^2 & \cdots & \lambda_n^2\\
\vdots & \vdots & & \vdots\\
\lambda_1^{n-1} & \lambda_2^{n-1} & \cdots & \lambda_n^{n-1}
\end{vmatrix}\\
&=\mathrm{e}^{(\lambda_1+\cdots+\lambda_n)x}\prod_{1\leqslant j<i\leqslant n}(\lambda_i-\lambda_j)\neq 0.
\end{aligned}$$

由 Cramer 法则，得到 $k_1=k_2=\cdots=k_n=0$. □

我们再给出一个显然的推论.

推论 7.1.1 设 V 是 $\mathbb{F}$ 上的一个 n 维线性空间.$\alpha_1,\alpha_2,\cdots,\alpha_n$ 是 V 中的一个向量组，且 V 中的任意向量都可以被它线性表示，则该向量组为 V 中的一组基.

证明 此时 $\alpha_1,\cdots,\alpha_n$ 与 V 中的任意一组基等价，秩都为 n，故线性无关. □

7.1.3 向量的坐标

设 V 是 $\mathbb{F}$ 上的 n 维线性空间，且

$$\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$$

是它的一组基. 任给 $\alpha\in V$，有线性表示

$$\alpha=a_1\varepsilon_1+a_2\varepsilon_2+\cdots+a_n\varepsilon_n.$$

由于 $\varepsilon_1,\cdots,\varepsilon_n$ 线性无关，于是这个表达式是唯一确定的. 我们称 $a_1,a_2,\cdots,a_n$ 为 α 在基 $\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$ 下的**坐标.**

为了书写方便，我们借助矩阵的乘法，把 α 的坐标写为

$$\alpha=(\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n)\begin{pmatrix}a_1\\a_2\\\vdots\\a_n\end{pmatrix}.$$

当然这只是一个约定，是形式上的，并不是真正的矩阵的乘法. 但是不难看出，矩阵乘法的某些规律比如结合律对这种形式的写法也是适用的.

例 7.1.6 线性空间 $\mathbb{K}[x]_n$(这表示所有多项式次数小于 n 的多项式组成的集合，显然它是一个线性空间)，在它里面挑出 n 个向量 (n 个一元多项式)

$$1,x,x^2,\cdots,x^{n-1}$$

证明它们线性无关.

证明 设

$$k_0\cdot 1+k_1x+k_2x^2+\cdots+k_{n-1}x^{n-1}=0.$$

等号坐标是一个多项式函数，右边是一个零函数，于是对于任意的 $x\in\mathbb{K}$ 代入等式都成立. 如果 $k_0,\cdots,k_{n-1}$ 不全为零，那么等式左端是一个 $\deg\leqslant n-1$ 的多项式，它最多有 $n-1$ 个根. 现在数域 $\mathbb{K}$ 内的数都是它的根，有无穷多个，于是产生了矛盾，这说明 $k_1=k_2=\cdots=k_{n-1}=0$. □

例 7.1.7 考虑 n 维线性空间 $\mathbb{K}[x]_n$ 中的向量 (多项式函数)

$$f(x)=a_0+a_1x+\cdots+a_{n-1}x^{n-1}\quad(a_i\in\mathbb{K}).$$

于是它在基 $1,x,x^2,\cdots,x^{n-1}$ 下的坐标 $a_0,a_1,\cdots,a_{n-1}$. 用形式写法可写为

$$f(x)=(1,x,\cdots,x^{n-1})\begin{pmatrix}a_0\\a_1\\\vdots\\a_{n-1}\end{pmatrix}.$$

并且任给 $a\in\mathbb{K}$，利用 Newton 二项展开公式，有

$$x^k=(x-a+a)^k=\sum_{i=0}^{k}\binom{k}{i}a^i(x-a)^{k-i}.$$

这说明 $\mathbb{K}[x]_n$ 内以下向量组

$$1,x-a,(x-a)^2,\cdots,(x-a)^{n-1}$$

和基 $1,x,x^2,\cdots,x^{n-1}$ 等价所以它也是一组基.

为了进一步讨论多项式，我们引入多项式的**形式微商.** 因为在一个一般数域中没有极限的概念，数域对求极限运算一般也不封闭. 比如一个有理数列当极限存在时，其极限值一般不再是有理数. 但是我们可以“形式”地将数学分析中实系数多项式的求导公式搬过来.

定义 7.1.4 对于 $\mathbb{K}$ 上任意多项式

$$f(x)=a_0x^m+a_1x^{m-1}+\cdots+a_{m-1}x+a_m,$$

我们定义

$$f'(x)=a_0mx^{m-1}+a_1(m-1)x^{m-2}+a_2(m-2)x^{m-3}+\cdots+a_{m-1},$$

为多项式 $f(x)$ 的**形式微商.**

这份形式微商显然也有数学分析中微商的基本性质：对于任意 $g(x),f(x)\in\mathbb{K}[x]$，以及任意 $k\in\mathbb{K}$，有

$$\begin{aligned}&(f(x)+g(x))'=f'(x)+g'(x);\\&(kf(x))'=kf'(x);\\&(f(x)g(x))'=f'(x)g(x)+f(x)g'(x).\end{aligned}$$

类似数学分析中的做法，对 $f(x)$ 求高阶微商，再将 $x=a$ 代入，立即有

$$\begin{aligned}&a_0=f(a),\\&a_1=f'(a),\\&a_2=\frac{1}{2!}f''(a),\\&\cdots\\&a_{n-1}=\frac{1}{(n-1)!}f^{(n-1)}(a).\end{aligned}$$

于是我们就得到了 $\mathbb{K}$ 上 $f(x)$ 在 $a\in\mathbb{K}$ 处的 Taylor 展开式

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

由此可知，$f(x)\in\mathbb{K}[x]$ 在基 $1,(x-a),(x-a)^2,\cdots,(x-a)^{n-1}$ 下的坐标表达式为

$$f(x)=(1,(x-a),(x-a)^2,\cdots,(x-a)^{n-1})\begin{pmatrix}f(x)\\f'(a)\\\frac{f''(a)}{2!}\\\vdots\\\frac{f^{(n-1)}(a)}{(n-1)!}\end{pmatrix}.$$

例 7.1.8 在 $\mathbb{K}^n$ 中给定一组基及一向量 β:

$$\alpha_1=\begin{pmatrix}a_{11}\\a_{21}\\\vdots\\a_{n1}\end{pmatrix},\alpha_2=\begin{pmatrix}a_{12}\\a_{22}\\\vdots\\a_{n2}\end{pmatrix},\cdots,\alpha_n=\begin{pmatrix}a_{1n}\\a_{2n}\\\vdots\\a_{nn}\end{pmatrix},\beta=\begin{pmatrix}b_1\\b_2\\\vdots\\b_n\end{pmatrix}$$

求 β 在基 $\alpha_1,\alpha_2,\cdots,\alpha_n$ 下的坐标.

解 按照定义，其实我们需要求解以下向量方程

$$x_1\alpha_1 + x_2\alpha_2 + \cdots + x_n\alpha_n = \beta.$$

考虑系数矩阵的增广矩阵，并对其做初等行变换得到一个新的增广矩阵.

$$\bar{A} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} & b_1 \\ a_{21} & a_{22} & \cdots & a_{2n} & b_2 \\ \vdots & \vdots & & \vdots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} & b_n \end{pmatrix} \longrightarrow \begin{pmatrix} & c_1 \\ & \\ E & c_2 \\ & \vdots \\ & c_n \end{pmatrix}.$$

于是解 $c_1, c_2, \cdots, c_n$ 就是 β 在基 $\alpha_1, \alpha_2, \cdots, \alpha_n$ 下的坐标.

接下来我们再看一个具体一点的例子.

例 7.1.9 在 $\mathbb{K}^4$ 中给定向量组

$$\begin{aligned} \alpha_1 &= (2, 1, 0, 1) \\ \alpha_2 &= (1, -3, 2, 4) \\ \alpha_3 &= (-5, 0, -1, -7) \\ \alpha_4 &= (1, -6, 2, 6). \end{aligned}$$

证明上述四个向量构成 $\mathbb{K}^4$ 的一组基，并且求向量

$$\beta = (8, 9, -5, 0)$$

在此组基下的坐标.

证明 将 $\alpha_1, \alpha_2, \alpha_3, \alpha_4$ 按照列的方式排入矩阵 A 中，由于

$$\det A = \begin{vmatrix} 2 & 1 & -5 & 1 \\ 1 & -3 & 0 & -6 \\ 0 & 2 & -1 & 2 \\ 1 & 4 & -7 & 6 \end{vmatrix} = 27 \neq 0$$

所以上述向量组线性无关，且刚好有 4 个，于是构成 $\mathbb{K}^4$ 的一组基. □

再考虑 A 的增广矩阵

$$\bar{A} = \begin{pmatrix} 2 & 1 & -5 & 1 & 8 \\ 1 & -3 & 0 & -6 & 9 \\ 0 & 2 & -1 & 2 & -5 \\ 1 & 4 & -7 & 6 & 0 \end{pmatrix},$$

对其做初等行变换可得

$$\begin{pmatrix} 1 & 0 & 0 & 0 & 3 \\ 0 & 1 & 0 & 0 & -4 \\ 0 & 0 & 1 & 0 & -1 \\ 0 & 0 & 0 & 1 & 1 \end{pmatrix}.$$

于是 β 相对于 $\alpha_1,\alpha_2,\alpha_3,\alpha_4$ 的坐标为 $3,-4,-1,1$.

故

$$\beta=3\alpha_1-4\alpha_2-\alpha_3+\alpha_4=(\alpha_1,\alpha_2,\alpha_3,\alpha_4)\begin{pmatrix}3\\-4\\-1\\1\end{pmatrix}.$$

现在取 $\mathbb{K}$ 上 n 维线性空间 V 的一组基 $\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$，设 α,β 在此组基下的坐标为

$$\alpha=(\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n)\begin{pmatrix}a_1\\a_2\\\vdots\\a_n\end{pmatrix},$$

$$\beta=(\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n)\begin{pmatrix}b_1\\b_2\\\vdots\\b_n\end{pmatrix}.$$

那么

$$\alpha+\beta=(a_1\varepsilon_1+a_2\varepsilon_2+\cdots+a_n\varepsilon_n)+(b_1\varepsilon_1+b_2\varepsilon_2+\cdots+v_n\varepsilon_n)=(a_1+b_1)\varepsilon+(a_2+b_2)\varepsilon_2+\cdots+(a_n+b_n)\varepsilon_n.$$

利用形式的矩阵乘法得到坐标表示

$$(\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n)\begin{pmatrix}a_1+b_1\\a_2+b_2\\\vdots\\a_n+b_n\end{pmatrix}.$$

同样对于数乘，$\forall k\in\mathbb{K}$，有

$$\begin{aligned}k\alpha&=ka_1\varepsilon_1+ka_2\varepsilon_2+\cdots+ka_n\varepsilon_n\\&=(\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n)\begin{pmatrix}ka_1\\ka_2\\\vdots\\ka_n\end{pmatrix}.\end{aligned}$$

所以，V 内一个向量 α 在基 $\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$ 下的坐标可以看作是 $\mathbb{K}^n$ 中的一个向量，V 中两个向量相加对应于它们在 $\mathbb{K}^n$ 中的坐标向量相加，数乘相当于它们在 $\mathbb{K}^n$ 中的坐标向量数乘. 这样我们可以把 V 中向量的抽象加法以及数乘具体化为它们的坐标作为 $\mathbb{K}^n$ 中的向量，按 $\mathbb{K}^n$ 的向量加法和数乘做具体的运算. 这样又可以把抽象的东西化为具体的东西. 其实这种化具体为抽象再将抽象重新化为具体，不断重复，是代数学乃至数学的基本研究方法.

接下来我们将看到线性空间 V 中抽象的线性无关可以化为其“坐标向量”的线性无关.

定理 7.1.2 设 V 是数域 $\mathbb{F}$ 上的 n 维线性空间. 在 V 内取定一组基 $\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$. 设 V 内向量组 $\alpha_1,\alpha_2,\cdots,\alpha_s$ 中各向量在此组基下的坐标表达式为

$$\alpha_i=(\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n)A_i,\quad i=1,2,\cdots,s,\quad A_i\in\mathbb{K}^n.$$

那么 $\alpha_1,\alpha_2,\cdots,\alpha_s$ 线性相关（无关）的充要条件是 $A_1,A_2,\cdots,A_s$ 在 $\mathbb{K}^n$ 内线性相关（无关）.

证明 令 $A_i=(a_{1i},a_{2i},\cdots,a_{ni})\in\mathbb{F}^n$. 于是对于任意的 $k_i\in\mathbb{F}$，我们有

$$k_1\alpha_1+k_2\alpha_2+\cdots+k_s\alpha_s=(\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n)[k_1A_1+k_2A_2+\cdots+k_sA_s].$$

令上式等于 0，其条件等价于

$$k_1A_1+k_2A_2+\cdots+k_sA_s=0.$$

那么 $\alpha_i,i=1,2,\cdots,s$ 线性相关的等价条件为存在一组不全为 0 的 $k_i,i=1,\cdots,s$ 使得上述式子等于 0. □

7.1.4 基变换与坐标变换

我们知道，线性空间 V 的基的取法可以有很多种，于是自然而然地我们想要找到不同组基之间的联系，以及找到一个向量在一个基下的坐标表示变换到另一个基下的坐标表示的方法.

首先我们来看基变换，即一个基在另一个基下的坐标表示.

设 n 维线性空间 V 内有两组基

$$\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n,$$
$$\eta_1,\eta_2,\cdots,\eta_n.$$

每个 η_i 都可以被 $\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$ 线性表示，即表示出它们的坐标. 设

$$\begin{aligned}\eta_1&=t_{11}\varepsilon_1+t_{21}\varepsilon_2+\cdots+t_{n1}\varepsilon_n,\\ \eta_2&=t_{12}\varepsilon_1+t_{22}\varepsilon_2+\cdots+t_{n2}\varepsilon_n,\\ &\cdots\cdots\cdots\cdots\cdots\cdots\\ \eta_n&=t_{1n}\varepsilon_1+t_{2n}\varepsilon_2+\cdots+t_{nn}\varepsilon_n.\end{aligned}$$

再利用矩阵的乘法，上述方程组可以形式的写成

$$(\eta_1,\eta_2,\cdots,\eta_n)=(\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n)\begin{pmatrix}t_{11}&t_{12}&\cdots&t_{1n}\\ t_{21}&t_{22}&\cdots&t_{2n}\\ \vdots&\vdots&&\vdots\\ t_{n1}&t_{n2}&\cdots&t_{nn}\end{pmatrix}.$$

于是我们可以定义

$$T=\begin{pmatrix}t_{11}&t_{12}&\cdots&t_{1n}\\ t_{21}&t_{22}&\cdots&t_{2n}\\ \vdots&\vdots&&\vdots\\ t_{n1}&t_{n2}&\cdots&t_{nn}\end{pmatrix}$$

为从基 $\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$ 到基 $\eta_1,\eta_2,\cdots,\eta_n$ 的**过渡矩阵**.

定理 7.1.3 在 $\mathbb{F}$ 上的 n 维线性空间 V 内给定一组基 $\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$，设 T 是 $\mathbb{F}$ 上的一个方阵，并且有等式

$$(\eta_1,\eta_2,\cdots,\eta_n)=(\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n)T.$$

则 $\eta_1,\eta_2,\cdots,\eta_n$ 是 V 的一组基的充分必要条件是 T 可逆.

证明 设 T 的列向量组为

$$T_1=\begin{pmatrix}t_{11}\\t_{21}\\\vdots\\t_{n1}\end{pmatrix},\quad T_2=\begin{pmatrix}t_{12}\\t_{22}\\\vdots\\t_{n2}\end{pmatrix},\cdots,T_n=\begin{pmatrix}t_{1n}\\t_{2n}\\\vdots\\t_{nn}\end{pmatrix},$$

再把 η_i 用这些坐标表示出来，于是有

$$\begin{gathered}\eta_1=(\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n)T_1,\\ \eta_2=(\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n)T_2,\\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\\ \eta_n=(\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n)T_n.\end{gathered}$$

由于 $\eta_1,\cdots,\eta_n$ 是 V 的一组基的充分必要条件是 $\eta_1,\cdots,\eta_n$ 线性无关，这等价于 $T_1,T_2,\cdots,T_n$ 在 $\mathbb{F}^n$ 内线性无关，而后者等价于 T 满秩或可逆. □

由上面的定理可知，两组基之间的过渡矩阵是可逆矩阵. 反过来说，从给定的一组基出发，通过某一个可逆矩阵 T，就可以获得一组新的基.

在讲完基变换后我们看一下坐标变换.

设线性空间 V 中一个向量 α 在第一组基 $\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$ 下的坐标为 $x_1,x_2,\cdots,x_n$ 即

$$\alpha=x_1\boldsymbol{\varepsilon}_1+x_2\boldsymbol{\varepsilon}_2+\cdots+x_n\boldsymbol{\varepsilon}_n=(\boldsymbol{\varepsilon}_1,\boldsymbol{\varepsilon}_2,\cdots,\boldsymbol{\varepsilon}_n)\begin{pmatrix}x_1\\x_2\\\vdots\\x_n\end{pmatrix}.$$

再设 α 在第二组基 $\eta_1,\eta_2,\cdots,\eta_n$ 下的坐标为 $y_1,y_2,\cdots,y_n$，即

$$\alpha=y_1\eta_1+y_2\eta_2+\cdots+y_n\eta_n=(\eta_1,\eta_2,\cdots,\eta_n)\begin{pmatrix}y_1\\y_2\\\vdots\\y_n\end{pmatrix}.$$

然后再设两组基之间的过渡矩阵 T，即

$$(\eta_1,\eta_2,\cdots,\eta_n)=(\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n)T.$$

为了方便，我们令

$$X=\begin{pmatrix}x_1\\x_2\\\vdots\\x_n\end{pmatrix},\quad Y=\begin{pmatrix}y_1\\y_2\\\vdots\\y_n\end{pmatrix}.$$

于是

$$\alpha=(\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n)X=(\eta_1,\eta_2,\cdots,\eta_n)Y.$$

然后将 $(\eta_1,\eta_2,\cdots,\eta_n)=(\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n)T$ 代入并利用形式矩阵乘法的结合律（利用线性空间八条运算法则不难验证）可得

$$\begin{aligned}(\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n)X=&[(\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n)T]Y\\=&(\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n)(TY).\end{aligned}$$

由于 $\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$ 是一组基，所以它们线性无关，它们两个线性组合相等时对应系数相等，故有

$$X=TY.$$

这就是**坐标变换公式.**

例 7.1.10 在 $\mathbb{K}^3$ 中给定两组基

$$\begin{aligned}&\varepsilon_1=(1,0,-1),\varepsilon_2=(2,1,1),\quad\varepsilon_3=(1,1,1);\\&\eta_1=(0,1,1),\quad\eta_2=(-1,1,0),\eta_3=(1,2,1).\end{aligned}$$

求从 ε_i 到 η_i 的过渡矩阵.

解 设过渡矩阵为 T，然后解形式矩阵方程

$$(\eta_1,\eta_2,\eta_3)=(\varepsilon_1,\varepsilon_2,\varepsilon_3)T.$$

将向量按列排入矩阵中得到

$$\begin{pmatrix}0&-1&1\\1&1&2\\1&0&1\end{pmatrix}=\begin{pmatrix}1&2&1\\0&1&1\\-1&1&1\end{pmatrix}T$$

将两个矩阵并排放在一起然后不断做初等行变换可得

$$\begin{pmatrix}1&2&1&0&-1&1\\0&1&1&1&1&2\\-1&1&1&1&0&1\end{pmatrix}\longrightarrow\begin{pmatrix}1&0&0&0&1&1\\0&1&0&-1&-3&-2\\0&0&1&2&4&4\end{pmatrix}.$$

于是过渡矩阵

$$T=\begin{pmatrix}0&1&1\\-1&-3&-2\\2&4&4\end{pmatrix}.$$

7.2　子空间

现在给出一个和我们之前定义等价的新定义.

定义 7.2.1 (子空间) 设 V 是 $\mathbb{F}$ 上的一个线性空间，M 是 V 的一个非空子集. 如果 M 关于 V 内的加法与数乘运算也组成 $\mathbb{K}$ 上的一个线性空间，则 M 称为 V 的一个**线性子空间 (subspace).**

命题 7.2.1 线性空间 V 的一个非空子集 M 是一个子空间的充分必要条件是，它满足以下两个条件：

1. 它对加法封闭:$\forall\alpha,\beta\in M\Rightarrow\alpha+\beta\in M$.

2. 对数乘封闭：$\forall\alpha\in M,k\in\mathbb{K}\Rightarrow k\alpha\in M$.

证明 必要性显然. 证明充分性.$0\alpha \in M \Rightarrow 0 \in M$. 另一方面，$\alpha \in M \Rightarrow (-1)\alpha = -\alpha \in M$. □

正如前面所说，我们把 $\mathrm{span}\{\alpha_1, \alpha_2, \cdots, \alpha_n\}$ 称为**由** $\alpha_1, \alpha_2, \cdots, \alpha_n$ **生成的子空间.** 容易看出，向量组 $\alpha_1, \alpha_2, \alpha_n$ 的任一极大线性无关组都是该子空间的一组基，而其向量组的秩为该子空间的维数.

下面看几个例子.

例 7.2.1 给定 $\mathbb{K}$ 上的一个齐次线性方程组

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = 0, \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = 0, \\ \cdots\cdots\cdots\cdots\cdots\cdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = 0, \end{cases} \quad (a_{ij} \in K)$$

它的全体解向量是 $\mathbb{K}^n$ 中的一个非空子集 M. 该 M 满足**命题7.2.1**中的两个条件，所以它是一个子空间，被称为**解空间.** 该齐次线性方程组的任意一组基础解系就是线性空间 M 的一组基. 如果方程组系输出矩阵的秩为 r，那么

$$\dim M = n - r.$$

例 7.2.2 考察 $\mathbb{K}$ 上全体 n 阶方阵所构成的线性空间 $M_n(\mathbb{K})$. 设 M 是 $M_n(\mathbb{K})$ 内主对角线元素之和为 0 的方阵所成的子集，即

$$M = \left\{ (a_{ij}) \in M_n(\mathbb{K}) | \sum_{i=1}^{n} (a_{ii}) = 0 \right\}.$$

显然 M 也满足**命题7.2.1**中的两个条件，故 M 也是 $M_n(\mathbb{K})$ 的一个子空间.

前面我们用过的 $\{0\}$ 被称为**零子空间**或**平凡子空间**. 除此之外的子空间被称为**非平凡子空间.**

7.2.1 子空间的交与和

类似前面向量空间的交与和，我们给出线性空间的交与和.

定义 7.2.2 设 M_1, M_2 为线性空间 V 中的两个子空间，称 $M_1 \cap M_2$ 为它们的**交**. 又定义

$$M = \{\alpha_1 + \alpha_2 | \alpha_1 \in M_1, \alpha_2 \in M_2\}$$

为 M_1 与 M_2 的**和**，记为 $M_1 + M_2$.

显然有 $M \subset M_1 + M_2; M_2 \subset M_1 + M_2$. 但还是要注意，$M_1 + M_2$ 和 $M_1 \cup M_2$ 不同，后者只是简单地把两个子空间聚拢在一起，它们之间的向量并不相加，在一般情况下，$M_1 \cup M_2 \neq M_1 + M_2$ 并且 $M_1 \cup M_2$ 可能已经不是线性空间了. 不过 $\mathrm{span}(M_1 \cup M_2) = M_1 + M_2$.

命题 7.2.2 设 M_1, M_2 是线性空间 V 的两个子空间，则它们的交 $M_1 \cap M_2$ 与和 $M_1 + M_2$ 仍是 V 的子空间.

证明 县先证明 $M_1 \cap M_2$ 是子空间. 首先 $0 \in M_1 \cap M_2$ 所以 $M_1 \cap M_2 \neq \varnothing$. 设 $\alpha, \beta \in M_1 \cap M_2$，于是 $\alpha + \beta \in M_1, \alpha + \beta \in M_2$ 同时成立，所以 $\alpha + \beta \in M_1 \cap M_2$. 又因为任一 $k \in \mathbb{K}$，以及 $k\alpha \in M_1$ 且 $k\alpha \in M_2$ 所以 $k\alpha \in M_1 \cap M_2$. 根据**定理 7.2.1**，$M_1 \cap M_2$ 是 V 的子空间.

接下来证 M_1+M_2 是子空间. 首先，$0\in M_1+M_2\Rightarrow M_1+M_2\neq\varnothing$. 若 $\alpha,\beta\in M_1+M_2$，则有

$$\alpha=\alpha_1+\alpha_2,\quad(\alpha_1\in M_1,\alpha_2\in M_2);$$
$$\beta=\beta_1+\beta_2,\quad(\beta_1\in M,\beta_2\in M).$$

于是

$$\alpha+\beta=(\alpha_1+\alpha_2)+(\beta_1+\beta_2)=(\alpha_1+\beta_1)+(\alpha_2+\beta_2),$$

其中 $\alpha_1+\beta_1\in M_1,\alpha_2+\beta_2\in M_2$，所以 $\alpha+\beta\in M_1+M_2$. 再对于任意 $\alpha\in M_1+M_2$，因为

$$\alpha=\alpha_1+\alpha_2,\quad(\alpha_1\in M_1,\alpha_2\in M_2),$$

于是有

$$k\alpha=k\alpha_1+k\alpha_2.$$

又因为 $k\alpha_1\in M_1,k\alpha_2\in M_2$，所以 $k\alpha\in M_1+M_2$. 再由**定理7.2.1**可知 M_1+M_2 也是 V 的子空间. □

有了两个子空间的交与和后，可以类比定义出多个子空间的交与和. 设 $M_1,M_2,\cdots,M_k$ 为 V 的 k 个子空间，它们的交定义为

$$M_1\cap M_2\cap\cdots\cap M_k,$$

即这 k 个子空间的公共向量所组成的集合. 显然，它是 V 的一个子空间.

再定义

$$M=\left\{\sum_{i=1}^{n}\alpha_i\,\middle|\,\alpha_i\in M,i=1,2,\cdots,n\right\}$$

称为 n 个子空间的和，记为 $M_1+M_2+\cdots+M_n$，显然归纳的可以看出它也是一个子空间.

例 7.2.3 考察三维几何空间中过坐标原点 O 两张不重合的的平面，两个平面上的向量张成子空间 M_1,M_2. 不难看出，M_1+M_2 是整个三维几何空间. 而它们的交 $M_1\cap M_2$ 是两平面交线 l 上全体向量组成的一维子空间.

例 7.2.4 考察三维几何空间中过坐标原点 O 的两条不重合直线 l_1,l_2. 它们上面的向量分别张成两个一维子空间 L_1,L_2. 于是 L_1+L_2 是过 l_1,l_2 的一张平面上的全体向量张成的二维子空间. 而 $L_1\cup L_2$ 只有零向量，为零子空间 $\{0\}$.

例 7.2.5 给定 $\mathbb{K}$ 上的两个齐次线性方程组

$$\begin{cases}a_{11}x_1+a_{12}x_2+\cdots+a_{1n}x_n=0,\\ a_{21}x_1+a_{22}x_2+\cdots+a_{2n}x_n=0,\\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\\ a_{m1}x_1+a_{m2}x_2+\cdots+a_{mn}x_n=0\ ;\end{cases}\tag{7.2.1}$$

$$\begin{cases}b_{11}x_1+b_{12}x_2+\cdots+b_{1n}x_n=0,\\ b_{21}x_1+b_{22}x_2+\cdots+b_{2n}x_n=0,\\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\\ b_{l_1}x_1+b_{l_2}x_2+\cdots+b_{l_n}x_n=0;\end{cases}\tag{7.2.2}$$

其中方程组(7.2.1)的解空间设为 M_1，方程组(7.2.2)的解空间设为 M_2，它们都是 $\mathbb{K}^n$ 的子空间. 所以 $M_1\cap M-2$ 中每个向量了既是(7.2.1)的解向量，同时也是方程组(7.2.2)的解向量. 所以 $M_1\cap M_2$ 就是把方程组(7.2.1)和(7.2.2)并在一起得到 n 个未知量，$m+l$ 个方程的新方程组的解空间.

定理 7.2.1 (维数公式) 设 M_1, M_2 是线性空间 V 的两个有限维子空间，则有

$$\dim(M_1+M_2)=\dim M_1+\dim M_2-\dim(M_1\cap M_2).$$

证明 在 $M_1\cap M_2$ 内取一组基

$$\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_r.$$

于是它可以扩充为 M_1 的一组基

$$\varepsilon_1,\cdots,\varepsilon_r,\alpha_1,\cdots,\alpha_s.$$

也可以扩充为 M_2 的一组基

$$\varepsilon_1,\cdots,\varepsilon_r,\beta_1,\cdots,\beta_t.$$

这时如果 $M_1\cap M_2=\{0\}$，则直接取 $r=0$. 然后我们证明向量组

$$\varepsilon_1,\cdots,\varepsilon_r,\alpha_1,\cdots,\alpha_s,\beta_1,\cdots,\beta_t$$

是 M_1+M_2 的一组基.

首先对于任意 $\alpha\in M_1+M_2$，有

$$\alpha=\gamma_1+\gamma_2,\quad(\gamma_1\in M_1,\gamma_2\in M_2).$$

而

$$\begin{aligned}\gamma_1&=k_1\varepsilon_1+\cdots+k_r\varepsilon_r+a_1\alpha_1+\cdots+a_s\alpha_s\\\gamma_2&=l_1\varepsilon_1+\cdots+l_r\varepsilon_r+b_1\beta_1+\cdots+b_t\beta_t.\end{aligned}$$

于是有

$$\alpha=\gamma_1+\gamma_2=(k_1+l_1)\varepsilon_1+\cdots+(k_r+l_r)\varepsilon_r+a_1\alpha_1+\cdots+a_s\alpha_s+b_1\beta_1+\cdots+b_t\beta_t.$$

故 α 可被向量组 $\varepsilon_1,\cdots,\varepsilon_r,\alpha_1,\cdots,\alpha_s,\beta_1,\cdots,\beta_t$ 线性表示.

然后再证明 $\varepsilon_1,\cdots,\varepsilon_r,\alpha_1,\cdots,\alpha_s,\beta_1,\cdots,\beta_t$ 线性无关. 令

$$k_1\varepsilon_1+\cdots+k_r\varepsilon_r+a_1\alpha_1+\cdots+a_s\alpha_s+b_1\beta_1+\cdots+b_t\beta_t=0,$$

移项可得

$$k_1\varepsilon_1+\cdots+k_r\varepsilon_r+a_1\alpha_1+\cdots+a_s\alpha_s=-b_1\beta_1-\cdots-b_t\beta_t.$$

由于

$$\begin{aligned}k_1\varepsilon_1+\cdots+k_r\varepsilon_r+a_1\alpha_1+\cdots+a_s\alpha_s&\in M_1,\\-b_1\beta_1-\cdots-b_t\beta_t&\in M_2\end{aligned}$$

而两者相等，所以

$$\gamma=k_1\varepsilon_1+\cdots+k_r\varepsilon_r+a_1\alpha_1+\cdots+a_s\alpha_s\in M_1\cap M_2.$$

于是 γ 可以被 $\varepsilon_1,\cdots,\varepsilon_r$ 线性表示. 然后有

$$\begin{aligned}\gamma&=f_1\varepsilon_1+\cdots+f_r\varepsilon_r,\\\gamma&=-b_1\beta_1-\cdots-b_t\beta_t,\end{aligned}$$

两式相减可得

$$0 = f_1\varepsilon_1 + \cdots + f_r\varepsilon_r + b_1\beta_1 + \cdots + b_t\beta_t.$$

由于 $\varepsilon_1, \cdot, \varepsilon_r, \beta_1, \cdots, \beta_t$ 线性无关，所以 $b_1 = \cdots = b_t = 0$. 代入 $k_1\varepsilon_1 + \cdots + k_r\varepsilon_r + a_1\alpha_1 + \cdots + a_s\alpha_s + b_1\beta_1 + \cdots + b_t\beta_t = 0$ 可得

$$k_1\varepsilon_1 + \cdots + k_r\varepsilon_r + a_1\alpha_1 + \cdots + a_s\alpha_s = 0.$$

再由 $\varepsilon_1, \cdots, \varepsilon_r, \alpha_1, \cdots, \alpha_s$ 线性无关，所以 $k_1 = \cdots = k_r = a_1 = \cdots = a_s = 0$. 故向量组线性无关.

综上，我们设的向量组 $\varepsilon_1, \cdots, \varepsilon_r, \alpha_1, \cdots, \alpha_s, \beta_1, \cdots, \beta_t$ 的确是 $M_1 + M_2$ 的一组基，所以 $\dim(M_1 + M_2) = r + s + t$，而 $\dim(M_1 \cap M_2) = r, \dim M_1 = r + s, \dim M_2 = r + t$. 所以

$$\dim(M_1 + M_2) = \dim M_1 + \dim M_2 - \dim(M_1 \cap M_2).$$

□

注 7.2.1 类似容斥原理，但在 4 个子空间及以上该公式失效，因为 4 个子空间之和就可以构造出无穷多个子空间. 比如 $\mathrm{span}\{e_1\}, \mathrm{span}\{e_2\}, \mathrm{span}\{e_3\}, \mathrm{span}\{e_1 + e_2 + e_3\}$.

从几何的角度上看，对于一个 3 维的几何空间 V，两个相交的 2 维平面 U, W，总有

$$\dim U + \dim W > \dim W,$$

因为它们相交处有一条 1 维的直线，受此启发，我们可以证明以下推论.

推论 7.2.1 设 $M_1, M_2, \cdots, M_n$ 是线性空间 V 中 n 个有限维子空间，则有

$$\dim(M_1 + M_2 \cdots + M_n) \leqslant \dim M_1 + \dim M_2 + \cdots \dim M_n.$$

证明 由前一个定理可知 $\dim(M_1 + M_2) \leqslant \dim M_1 + \dim M_2$，然后使用数学归纳法即可. □

借此，我们再给出一些在未来可能会用上的几何定义. 在 n 维线性空间 V 中存在更小维数的子空间，容易看出，可以在 V 中插入子空间链

$$0 \subset V_1 \subset V_2 \subset \cdots \subset V_{n-1} \subset V_n = V = \mathrm{span}\{e_1, e_2, \cdots, e_n\},$$

其中 $V_i = \mathrm{span}\{e_1, \cdots, e_i\}$. 1 维的线性空间被称为**直线 (line)**，2 维的线性空间则被称为**平面 (plane)**. 当维数大于等于 3 时，就称为是 **k 维平面**. 设 U 是线性空间 V 的子空间. 差

$$\mathrm{codim}\ U = \dim V - \dim U$$

被称为子空间 U 的**余维数 (codimension)**. 我们再定义，余维数为 1 的字空间被称为**超平面 (hyperplane)**. 超平面是一个相对的概念，比如直线是 2 维线性空间的超平面，而 W 可以看成是比它维数大 1 且包含它的一个线性空间的一个超平面.

7.2.2 子空间的直和

我们先考察一个三维几何空间 V. 设过坐标原点 O 的两张不重合的平面分别代表 V 的两个子空间 M_1, M_2. 显然 $M_1+M_2=V$. 所以任意 V 内向量 α 可以表示为 $\alpha=\alpha_1+\alpha_2,(\alpha_1\in M_1,\alpha_2\in M_2)$. 不难看出这个表示方法不唯一.

我们再考察 V 中的零向量. 设 $\forall\beta\in M_1\cap M_2$，由于 $M_1\cap M_2\neq\varnothing$ 所以 V 中的零向量 $0=\beta+(-\beta)$. 显然，这样的 β 有无穷多种取法.

只要零向量的表示法不唯一，那么任意向量的表示法也不唯一. 只要 α 有一个表示 $\alpha_1+\alpha_2$ 那么就有

$$\alpha=\alpha+0=\alpha_1+\alpha_2+(\beta+(-\beta))=(\alpha_1+\beta)+(\alpha_2-\beta),$$

其中 $\alpha_1+\beta\in M_1,\alpha_2-\beta\in M_2$，且 $\alpha_1+\beta\neq\alpha_1,\alpha_2-\beta\neq\alpha_2$.

这种向量表示法不唯一的情况无疑会给我们研究线性空间造成一定的困难. 所以我们需要定义一种子空间的加法，使得向量的表示法也是唯一的. 也就是我们前面提到的直和，我们再在一般的线性空间上给出定义.

定义 7.2.3 (直和) 设 M_1, M_2 是线性空间 V 的两个子空间. 令 $M=M_1+M_2$，如果 M 内的任意向量 α 其表达式

$$\alpha=\alpha_1+\alpha_2,\quad(\alpha_1\in M_1,\alpha_2\in M_2)$$

是唯一的, 则称 M 是 M_1 与 M_2 的**直和 (direct sum)**，记为 $M_1\oplus M_2$

我们接下来证明这个定义和我们前面的定义是等价的.

定理 7.2.2 设 M_1, M_2 是 $\mathbb{F}$ 上线性空间 V 的两个有限维子空间，则以下命题等价：

1. $M_1\oplus M_2$.
2. 向量表示法唯一，即 $0=\alpha_1+\alpha_2,(\alpha_1\in M_1,\alpha_2\in M_2)$ 必有 $\alpha_1=\alpha_2=0$.
3. $M_1\cap M_2=\{0\}$.
4. $\dim M_1+\dim M_2=\dim(M_1+M_2)$.

证明 1. $\Rightarrow$ 2. 按照定义即可说明零向量表示法唯一.

2. $\Rightarrow$ 3 如果 $M_1\cap M_2\neq\{0\}$，则存在 $\beta\in M_1\cap M_2$ 且 $\beta\neq0$，于是存在 $\beta\in M_1.-\beta\in M_2$ 使得

$$0=\beta+(-\beta).$$

于是零向量表示法不唯一，得矛盾.

3. $\Rightarrow$ 4. 利用维数公式即得.

4. $\Rightarrow$ 1. 利用维数公式，$\dim(M_1\cap M_2)=0$ 说明 $M_1\cap M_2=\{0\}$. 于是对任意 $\alpha\in M_1+M_2$，如果有

$$\alpha=\alpha_1+\alpha_2,\quad(\alpha_1\in M_1,\alpha_2\in M_2)$$

以及

$$\alpha=\alpha_1'+\alpha_2',\quad(\alpha_1'\in M_1,\alpha_2'\in M_2),$$

做差可得

$$\alpha_1-\alpha'=\alpha_2-\alpha_2'.$$

这说明 $\alpha_1-\alpha_1'=\alpha_2-\alpha_2'\in M_1\cap M_2=\{0\}$ 所以 $\alpha_1=\alpha_1',\alpha_2=\alpha_2'$ 即 α 的表示法唯一. □

举两个关于直和的例子.

例 7.2.6 考虑线性空间 $M_n(\mathbb{K})$. 令 M 为 $M_n(\mathbb{K})$ 内主对角线上元素之和为 0 的所有方阵张成的子空间. 又命 N 为 n 阶全体**数量矩阵 (scalar matrix)**$dE, d \in \mathbb{K}$ 张成的子空间. 如果有 $A \in M \cap N$，那么同时有 $A = dE$ 和 A 的主对角线上之和 $nd = 0$，所以 $d = 0$，于是 $A \equiv 0$. 所以 $M \cap N = 0$.

例 7.2.7 如果上一个例子中，保持 M 不变，把 N 改为 $M_n(\mathbb{K})$ 内全体对角矩阵组成的子空间，那么 $M + N$ 就不是直和，因为

$$\begin{pmatrix} 1 & & & & \\ & -1 & & & \\ & & 0 & & \\ & & & \ddots & \\ & & & & 0 \end{pmatrix} \in M \cap N$$

所以 $M \cap N \neq \{0\}$.

接下来考虑多个子空间的直和.

定义 7.2.4 设 $M_1, M_2, \cdots, M_n$ 为线性空间 V 的子空间. 令 $M = \sum_{i=1}^{n} M_i$. 如果对于 M 中任意向量 α 有唯一的表达式

$$\alpha = \sum_{i=1}^{n} \alpha_n, \quad (\alpha_i \in M_i, i = 1, 2, \cdots, n).$$

则称 M 是 $M_1, \cdots, M_n$ 的直和. 记为

$$M_1 \oplus M_2 \oplus \cdots \oplus M_n = \bigoplus_{i=1}^{n} M_n.$$

定理 7.2.3 设 $M_1, M_2, \cdots, M_n$ 是 $\mathbb{K}$ 上线性空间 V 的有限维子空间，则以下命题互相等价：

1. $M = \sum_{i=1}^{n} M_i$ 是直和.
2. 零向量的表示法唯一，即如果
$$0 = \sum_{i=1}^{n} \alpha_i, \quad (\alpha_i \in M_i, i = 1, 2, \cdots, n)$$
则 $\alpha_1 = \alpha_2 = \cdots = \alpha_n$.
3. 任意一个子空间与除了它以外子空间的和的交集是 $\{0\}$，即
$$M_i \cap \left(\sum_{j=1, i \neq j}^{M_j} \right) = \{0\}, \quad (i = 1, 2, \cdots, n).$$
4. 子空间和的维数等于子空间的维数和，即
$$\dim \sum_{i=1}^{n} M_n = \sum_{i=1}^{n} \dim M_i.$$

证明 1. ⇒ 2. 由定义直接可得.

2. ⇒ 3. 若存在一个 $0 < i \leqslant n$，使得

$$M_i \cap \left(\sum_{j=1,i\neq j}^{n} M_i \neq \{0\}\right),$$

则存在 $\beta \in M_i \cap \left(\sum_{j=1,i\neq j}^{n} M_i \neq \{0\}\right)$，则 $-\beta \in M_i, \beta \in \sum_{j=1,j\neq i}^{n} M_j$. 于是 β 可被表示为

$$\beta = \alpha_1 + \cdots + \alpha_{i-1} + \alpha_{i+1} + \cdots + \alpha_n, \quad (\alpha_j \in M_j, j \neq i.).$$

于是就有

$$0 = \beta + (-\beta) = \alpha_1 + \cdots + \alpha_{i-1} + (-\beta) + \alpha_{i+1} + \cdots + \alpha_n$$

其中至少有 $-\beta \neq 0$ 这与零向量表示法唯一矛盾.

3. ⇒ 4. 使用归纳法. 当 $n = 1$ 时显然成立. 设对 $n - 1$ 个子空间的情况成立.

因为 $M_1 \cap \left(\sum_{j=2}^{n} M_j\right) = \{0\}$，使用维数公式就有

$$\dim \sum_{i=1}^{k} M_i = \dim M_1 + \dim \sum_{j=2}^{n} M_j.$$

对于任意 $M_j (j \geqslant 2)$ 由条件也有

$$M_j \cap \left(\sum_{s=2,s\neq j}^{n} M_s\right) \subseteq M_j \cap \left(\sum_{s=1,s\neq j}^{n} M_s\right) = \{0\}.$$

于是由归纳假设，有

$$\dim \sum_{j=2}^{n} M_j = \sum_{j=2}^{n} \dim M_j.$$

所以

$$\dim \sum_{i=1}^{n} M_i = M_1 + \sum_{j=2}^{n} \dim M_j = \sum_{i=1}^{n} \dim M_i.$$

4. ⇒ 1. 先证明，

$$M_i \cap \left(\sum_{j=1,j\neq i}^{n} M_j\right) = \{0\} \quad (i = 1, 2, \cdots, n).$$

对任意 $1 \leqslant i \leqslant n$，由维数公式及其推论可得

$$\begin{aligned}
0 \leqslant \dim \left\{M_i \cap \left(\sum_{j=1,j\neq i}^{n} M_j\right)\right\} &= \dim M_i + \dim \sum_{j=1,j\neq i}^{\mathrm{n}} M_j - \dim \sum_{j=1}^{\mathrm{n}} M_j \\
&\leqslant \dim M_i + \sum_{j=1,j\neq i}^{n} M_j - \dim \sum_{j=1}^{k} M_j. \\
&= \sum_{j=1}^{n} M_j - \dim \sum_{j=1}^{n} M_j.
\end{aligned}$$

由假设，$\sum\limits_{j=1}^{n} M_j = \dim \sum\limits_{j=1}^{n} M_j$，所以

$$\dim \left\{ M_i \cap \left(\sum_{j=1,j\neq i}^{n} M_j \right) \right\} = 0$$

即

$$M_i \cap \left(\sum_{j=1,j\neq i}^{n} M_j \right) = \{0\}.$$

再设 $\alpha \in \sum\limits_{i=1}^{n} M_i$，证明 α 的表示法唯一. 令

$$\alpha = \alpha_1 + \alpha_2 + \cdots + \alpha_n = \alpha_1' + \alpha_2' + \cdots + \alpha_n',$$

其中 $\alpha_i, \alpha_i' \in M_i (i = 1, 2 \cdots, n)$. 对任意 i，有

$$\begin{aligned} \alpha_i - \alpha_i' &= (\alpha_1' - \alpha_1) + \cdots + (\alpha_{i-1}' - \alpha_{i-1}) \\ &\quad + (\alpha_{i+1}' - \alpha_{i+1}) + \cdots + (\alpha_k' - \alpha_k) \in \sum_{j=1,j\neq i}^{k} M_j. \end{aligned}$$

于是 $\alpha_i - \alpha_i' \in M_i \cap \left(\sum\limits_{j=1,j\neq i}^{n} M_j = \{0\}. \right)$ 即 $\alpha_i - \alpha_i' = 0 (i = 1, 2, \cdots, n)$. 这又说明了 α 的表示法唯一，即 $\sum\limits_{i=1}^{n} M_i$ 是直和. □

注 7.2.2 上述两个定理在两个子空间和多个子空间是否为直和的判断上较为有用.

推论 7.2.2 设 V 是 $\mathbb{F}$ 上的 n 维线性空间，$M_1, M_2, \cdots, M_k$ 是 V 的子空间，且

$$V = \bigoplus_{i=1}^{k} M_i.$$

则在每个子空间 M_i 内取一组基 $\{\varepsilon_{i1}, \varepsilon_{i2}, \cdots, \varepsilon_{in_i}\}$，则这些基合并后即成 V 的一组基.

证明 每个 M_i 内取定一组基 $\{\varepsilon_{i1}, \varepsilon_{i2}, \cdots, \varepsilon_{in_i}\}$，则 $n_i = \dim M_i$. 由上述定理可知 $\sum_{i=1}^{n} n_i = \dim V = n$. 所以这些向量组合并后有 n 个向量. 对于任意 $\alpha \in V$，有

$$\alpha = \sum_{i=1}^{k} \alpha_i, \quad (\alpha_i \in M_i).$$

而每个 α_i 可被 M_i 上的基线性表示，所以 α 可被合并后的基线性表示，所以合并后的基就是 V 中的一组基. □

定理 7.2.4 设 M 是 $\mathbb{K}$ 上有限维线性空间 V 的一个子空间，则必存在 V 的子空间 N，使得

$$V = M \oplus N.$$

我们称 N 为 M 的**补空间 (complementary space)**.

证明 若 $M = \{0\}$，直接取 $N = V$. 现设 $M \neq \{0\}$. 在 M 内取一组基 $\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_r$，则 $M = \mathrm{span}\{\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_r\}$. 由于 M 的基可以扩充为 V 的一组基

$$\varepsilon_1, \cdots, \varepsilon_r, \varepsilon_{r+1}, \cdots, \varepsilon_n.$$

所以取 $N = \mathrm{span}\{\varepsilon_{r+1}, \cdots, \varepsilon_n\}$，显然有 $V = M + N$. 又因为 $\dim V = n = r + (n-r) = \dim M + \dim N$ 所以 $V = M \oplus N$.□

注 7.2.3 一个子空间的补空间不是唯一的，因为 M 的基实际上有无穷多种方式扩充为 V 的基.

例 7.2.8 考虑平面上以坐标原点 O 为起点的全体向量所组成的 $\mathbb{R}$ 上二维线性空间 V. 过 O 点的一条直线上的全体向量组成的一个一维子空间 M. 过 O 点再任意做一条不与 M 重合的直线，它上面的全体向量组成一个一维子空间 N, 显然 $M\cap N=\{0\}$，所以 $V=M\oplus N$. 于是 N 是 M 的一个补空间. 任意一个 α 上的向量在 N 取定后都可以被唯一的分解成

$$\alpha=\alpha_1+\alpha_2,\quad (\alpha_1\in N,\alpha_2\in M).$$

由于 N 的取法实际上有无穷多种，每改变一次 N，上述分解式中的 α_1,α_2 都要跟着变.

如果两个 $\mathbb{F}$ 上的空间 U,W 不是其 V 的子空间，我们也定义一种和. 考虑有序对 $(u,w)\in W\times U$ 组成的集合，再定义上面的加法

$$(u',w')+(u,w):=(u'+u,w'+w).$$

以及数量乘法，设 $\in\mathbb{F}$

$$k(u,w):=(ku,kw).$$

不难验证这样的空间是一个线性空间，令 $V=U\times W$，这 V 就称为 U,W 的**外直和**. 从直观上看，就好像两个坐标轴构成了一个平面，其元素相当于上面的向量.

所有共线向量 $(u,0)$ 构成 V 的一个子空间 $\widetilde{U}$ 然后 $(0,w)$ 组成的集合也构成 V 的一个子空间 $\widetilde{W}$. 显然有同构映射 $u\mapsto(u,0)$ 和 $w\mapsto(w,0)$. 于是 $U,\widetilde{U}$ 同构，$W,\widetilde{W}$ 也同构. 这样，在不考虑单个空间中具体集合的情况下，我们可以将两种直和等同起来，所以我们也记 U,W 的外直和为

$$V=U\oplus W.$$

由于关于内直和的讨论更多，所以在不做特别说明的情况下，我们说两个空间的直和都是指内直和.

7.3 商空间

定义 7.3.1 设 V 是 $\mathbb{K}$ 上的一个线性空间，M 是 V 的一个子空间，我们定义关系 $\mathcal{R}$

$$\alpha\mathcal{R}\beta:\Longleftrightarrow \alpha-\beta\in M.$$

由于

1. $\alpha\mathcal{R}\alpha$，因为 $0\in M$.

2. 若 $\alpha\mathcal{R}\beta$，则 $\beta\mathcal{R}\alpha$，因为 $\alpha-\beta\in M\Rightarrow\beta-\alpha\in M$.

3. 若 $\alpha\mathcal{R}\beta$ 且 $\beta\mathcal{R}\gamma$，则 $\alpha\mathcal{R}\gamma$，因为 $\alpha-\beta\in M,\beta-\gamma\in M\Rightarrow(\alpha-\beta)+(\beta-\gamma)=\alpha-\gamma\in M$.

所以，关系 $\mathcal{R}$ 是一个等价关系. 我们再定义 V 的子集 $\alpha+M:=\{\alpha+m|\alpha\in V,m\in M\}$. 不难证明 $\alpha+M$ 是等价类，因为我们可以得到以下简单性质.

命题 7.3.1 1. $\alpha'\mathcal{R}\alpha\Leftrightarrow\alpha'-\alpha\in M\Leftrightarrow\alpha'\in\alpha+M$.

2. $\alpha'\in\alpha+M\Leftrightarrow\alpha'+M=\alpha+M$.

3. $\alpha + M = 0 + M \Leftrightarrow \alpha \in M$.

4. $\alpha' + M \neq \alpha + M \Rightarrow (\alpha' + M) \cap (\alpha + M) = \varnothing$.

证明 1. 第一个等价符号直接可由定义得到. 若 $\alpha' - \alpha \in M$ 于是存在 $m \in M$ 使得 $\alpha' = \alpha + m$ 所以 $\alpha' \in \alpha + M$. 若 $\alpha' \in \alpha + M$ 则存在 m 使得 $\alpha' = \alpha + m$，于是 $\alpha' - \alpha = m \in M$.

2. 如果 $\alpha' \in \alpha + M$，则存在 $m \in M$ 使得 $\alpha' = \alpha + m$. 于是对于任意 $m' \in M$ 都有 $\alpha' + m' = \alpha + (m + m') \in \alpha + M$ 所以

$$\alpha' + M \subset \alpha + M.$$

同时，因为 $\alpha = \alpha' - m$，所以 $\alpha + m' = \alpha' + (m - m') \in \alpha' + M$ 所以

$$\alpha + M \subset \alpha' + M.$$

所以 $\alpha' + M = \alpha + M$.

3. 由于性质 2. 可得成立.

4. 如果 $(\alpha' + M) \cap (\alpha + M) \neq \varnothing$，则存在 $\beta \in (\alpha' + M) \cap (\alpha + M)$ 再由 2. 可知

$$\alpha' + M = \beta + M = \alpha + M.$$

同时注意到，对于任意 V 中的向量 α，都会属于一个集合 $\alpha + M$（即以它自己为代表的那个 $\alpha + M$），所以所有集合 $\alpha + M$ 的并集就是 V，故 $\alpha + M$ 的确是一个等价类.

定义 7.3.2 (商空间) 设 V 是 $\mathbb{F}$ 上的线性空间，再设 M 是它的子空间. 定义等价关系

$$\alpha \mathcal{R} \beta :\Longleftrightarrow \alpha - \beta \in M.$$

再定义被称为 V 关于 M 的**陪集 (coset)**$\alpha + M \coloneqq \{\alpha + m | \alpha \in V, m \in M\}$ 我们将商集 $V/\mathcal{R}$ 记为 V/M，并称为 V 关于子空间 M 的**商空间 (quotient space)**，有时也称 V 模去子空间 M，同时称 α 与 β **模** M **同余.**

商空间作为我们研究的一个代数对象，自然要想到定义商空间上的加法运算和数乘运算.

定义 7.3.3 (商空间上的运算) 设 V 是 $\mathbb{K}$ 上的线性空间，V/M 是它一个商空间. 再设 $\alpha + M, \beta + M \in V/M, k \in \mathbb{K}$. 我们定义

$$\begin{aligned} &(\alpha + M) + (\beta + M) \coloneqq (\alpha + \beta) + M, \\ &k(\alpha + M) \coloneqq k\alpha + M. \end{aligned}$$

我们需要证明，在选取不同的代表元时，上述定义的运算结果保持不变，即上述定义是**良定义 (well-defined)** 的.

证明 设

$$\alpha + M = \alpha' + M, \quad \beta + M = \beta' + M, \quad (\alpha \neq \alpha', \beta \neq \beta').$$

则 $\alpha' \mathcal{R} \alpha, \beta' \mathcal{R} \beta$. 所以

$$\alpha' = \alpha + m_1, \quad \beta' = \beta + m_2, \quad (m_1, m_2 \in M).$$

故

$$\alpha' + \beta' = \alpha + \beta + (m_1 + m_2) \in \alpha + \beta + M \Leftrightarrow (\alpha' + \beta') + M = (\alpha + \beta) + M.$$

这说明

$$(\alpha+M)+(\beta+M)=(\alpha'+M)+(\beta'+M).$$

所以加法运算不会因为代表元选取的不同而发生改变.

设 $\alpha+M=\alpha'+M$，则 $\alpha'\mathcal{R}\alpha$. 所以

$$\alpha'=\alpha+m,\quad (m\in M).$$

等式两边同乘 $k\in\mathbb{K}$ 得到

$$k\alpha'=k\alpha+km\in k\alpha+M.$$

所以 $k\alpha'+M=k\alpha+M$ 即

$$k(\alpha+M)=k(\alpha'+M).$$

综上，我们定义的加法与数乘都是良定义的. □

定理 7.3.1 V 是 $\mathbb{F}$ 上的线性空间，M 是 V 的子空间，则商空间 V/M 是线性空间.

证明 对于 $\alpha,\beta,\gamma\in V$ 有

1.

$$\begin{aligned}&(\alpha+M)+[(\beta+M)+(\gamma+M)]\\&=(\alpha+M)+[(\beta+Y)+M]\\&=(\alpha+\beta+\gamma)+M=[(\alpha+\beta)+M]+(\gamma+M)\\&=[(\alpha+M)+(\beta+M)]+(\gamma+M)\,;\end{aligned}$$

2.

$$\begin{aligned}(\alpha+M)+(\beta+M)&=(\alpha+\beta)+M=(\beta+\alpha)+M\\&=(\beta+M)+(\alpha+M);\end{aligned}$$

3. 以零向量为代表元的等价类 $0+M=M$ 满足

$$(\alpha+M)+(0+M)=\alpha+M.$$

故 $0+M=M$ 是 V/M 中的零元素.

4.

$$(\alpha+M)+[(-\alpha)+M]=[\alpha+(-\alpha)]+M=0+M,$$

即 $(-\alpha)+M$ 是 $\alpha+M$ 的负元素.

$5.1\cdot(\alpha+M)=1\alpha+M=\alpha+M.$

6. 对于 $k,l\in\mathbb{F}$ 有

$$\begin{aligned}(kl)(a+M)&=(kl)a+M=k(l\alpha)+M\\&=k(l\alpha+M)=k[l(\alpha+M)]\end{aligned}$$

7. 对于 $k,l\in\mathbb{F}$ 有

$$\begin{aligned}(k+l)(\alpha+M)&=(k+l)\alpha+M=(k\alpha+l\alpha)+M\\&=(k\alpha+M)+(l\alpha+M)\\&=k(\alpha+M)+l(\alpha+M).\end{aligned}$$

8. 对于 $k \in \mathbb{F}$ 有

$$\begin{aligned} k[(\alpha+M)+(\beta+M)] &= k[(\alpha+\beta)+M] \\ &= k(\alpha+\beta)+M = (k\alpha+k\beta)+M \\ &= (k\alpha+M)+(k\beta+M) \\ &= k(\alpha+M)+k(\beta+M). \end{aligned}$$

所以 V/M 满足八条线性空间运算法则，故它是一个线性空间. □

为了把商空间的符号写的简单一点我们定义符号 $\overline{\alpha} := \alpha + M$. 于是由商空间的加法定义可以得到 $\overline{\alpha}+\overline{\beta}=\overline{\alpha+\beta}$ 以及数乘的定义可得 $k\overline{\alpha}=\overline{k\alpha}$. 于是更一般地可以得到

$$k_1\overline{\alpha}_1+k_2\overline{\alpha}_2+\cdots+k_s\overline{\alpha}_s=\overline{k_1\alpha_1+k_2\alpha_2+\cdots+k_s\alpha_s}.$$

例 7.3.1 设 A 是 $\mathbb{K}$ 上的 $m\times n$ 矩阵，令 X 为全体 $n\times 1$ 矩阵 $[x_1,x_2,\cdots,x_n]$. 齐次线性方程组 $AX=0$ 的全体解向量组成 $\mathbb{K}^n$ 的一个子空间 M. 对于 $\alpha\in\mathbb{K}^n$，满足 $B=A\alpha$，那么 α 为线性方程组 $AX=B$ 的一个特解，则等价类 $\alpha+M$ 恰好为 $AX=B$ 的全体解向量组成的集合. 这样，商空间 $\mathbb{K}^n/M$ 的每一个元素就代表了 $\mathbb{K}$ 上一个线性方程组 $AX=B$ 的全体解向量.

例 7.3.2 设 $V=\mathbb{R}^2$ 是坐标平面，设子空间 L 为与 x 轴. 如果把平面中的点作为元素的话，那么平行于 x 轴的直线组成的集合就是 V/L.

接下来讨论商空间的基与维数

定理 7.3.2 设 V 是 $\mathbb{F}$ 上的 n 维线性空间，M 是 V 的一个 m 维子空间，则 $\dim V/M=\dim V-\dim M=n-m$.

证明 在 M 内取一组基 $\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_m$，扩充为 V 的一组基 $\varepsilon_1,\cdots,\varepsilon_m,\varepsilon_{m+1},\cdots,\varepsilon_n$. 我们只需证明，$\overline{\varepsilon}_{m+1},\cdots,\overline{\varepsilon}_n$ 是 V/M 的一组基.

先证明 $\overline{\varepsilon}_{m+1},\cdots,\overline{\varepsilon}_n$ 线性无关. 设

$$k_{m+1}\overline{\varepsilon}_{m+1}+k_{m+2}\overline{\varepsilon}_{m+1}+\cdots+k_n\overline{\varepsilon}_n=\overline{0}.$$

于是

$$\overline{k_{m+1}\varepsilon_{m+1}+k_{m+2}\varepsilon_{m+2}+\cdots+k_n\varepsilon_n}=\overline{0}=0+M.$$

所以

$$k_{m+1}\varepsilon_{m+1}+k_{m+2}\varepsilon_{m+2}+\cdots+k_n\varepsilon_n\in M.$$

由于一个线性无关的向量组的任意部分组都是线性无关的，所以

$$k_{m+1}=k_{m+2}=\cdots=k_n=0.$$

这说明 $\overline{\varepsilon}_{m+1},\cdots,\overline{\varepsilon}_n$ 在 V/M 内线性无关.

再证明 $\forall\overline{\alpha}\in V/M$，其 $\overline{\alpha}$ 可被 $\overline{\varepsilon}_{m+1},\cdots,\overline{\varepsilon}_n$ 线性表示. 注意到

$$V\ni\alpha=k_1\varepsilon_1+\cdots+k_m\varepsilon_m+k_{m+1}\varepsilon_{m+1}+\cdots+k_n\varepsilon_n.$$

于是

$$\overline{\alpha}=\overline{k_1\varepsilon_1+\cdots+k_m\varepsilon_m+k_{m+1}\varepsilon_{m+1}+\cdots+k_n\varepsilon_n}=k_1\overline{\varepsilon}_1+\cdots+k_m\overline{\varepsilon}_m+k_{m+1}\overline{\varepsilon}_{m+1}+\cdots+k_n\overline{\varepsilon}_n.$$

由于 $\varepsilon_i \in M \Leftrightarrow \overline{\varepsilon}_i = \overline{0}, (i = 1, 2, \cdots, m)$. 所以上式

$$\overline{\alpha} = k_{m+1}\overline{\varepsilon}_{m+1} + \cdots + k_n\overline{\varepsilon}_n.$$

综上 $\overline{\varepsilon}_{m+1}, \cdots, \overline{\varepsilon}_n$ 是 V/M 的一组基. □

注 7.3.1 不难发现，$\dim V/M = \mathrm{codim}_V M$.

7.4 线性映射与线性变换

7.4.1 线性映射

定义 7.4.1 设 U, V 为 $\mathbb{K}$ 上两个线性空间，定义映射 $f : U \to V$，满足以下条件

1. 对于任意 $\alpha, \beta \in V$，有
$$f(\alpha + \beta) = f(\alpha) + f(\beta).$$
2. 对于任意 $\alpha \in U, k \in \mathbb{K}$，有
$$f(k\alpha) = kf(\alpha).$$

则称 f 为 U 到 V 的一个**线性映射 (linear map)**.

注 7.4.1 线性映射也称为**线性算子 (linear operator)**，也称为线性空间的**同态映射 (homomorphic map)**，简称**同态 (homomorphism)**. 所以我们把 U 到 V 的所有线性映射组成的集合记为 $\hom(U, V)$.

从线性映射的定义立即可以推出一个等价的等式：对于任意 $\alpha, \beta \in U; k, l \in \mathbb{K}$ 有

$$f(k\alpha + l\beta) = kf(\alpha) + l(\beta).$$

一般地，若 $f \in \hom(U, V)$，对任意 $\alpha_1, \cdots, \alpha_n \in U; k_1, \cdots, k_n \in \mathbb{K}$ 有

$$f(k_1\alpha_1 + k_2\alpha_2 + \cdots + k_n\alpha_n) = k_1f(\alpha_1) + k_2f(\alpha_2) + \cdots + k_nf(\alpha_n).$$

上式为线性映射最基本的属性，线性映射的所有理论都以此为立足点.

从线性映射定义中的第二个条件还可以得到 $f(0) = 0, f(-\alpha) = -f(\alpha)$.

下面看一些线性映射的实例.

例 7.4.1 $\mathbb{K}$ 上的线性空间 $M_{m,n}(\mathbb{K}), M_{m,s}(\mathbb{K})$. 取定 $\mathbb{K}$ 上的一个 $n \times s$ 矩阵 A. 定义映射

$$f : M_{m,n}(\mathbb{K}) \to M_{m,s}(\mathbb{K}), X \mapsto XA.$$

于是对于任意 $X, Y \in M_{m,n}(\mathbb{K}); k, l \in \mathbb{K}$. 有

$$f(kX + lY) = (kX + lY)A = kXA + lYA = kf(X) + lf(Y).$$

故 f 为线性映射.

例 7.4.2 考察一个由开区间 (a,b) 上全体连续函数张成的 $\mathbb{R}$ 上线性空间 $C(a,b)$ 的两个子空间

$$M = \operatorname{span}\{1, \sin x, \sin 2x, \cdots, \sin nx\}$$
$$N = \operatorname{span}\{1, \cos x, \cos 2x, \cdots, \cos nx\}.$$

定义 $D: M \to N, M \ni \alpha \mapsto \dfrac{\mathrm{d}\alpha}{\mathrm{d}x}$. 于是对于 $\forall \alpha, \beta \in M; k, l \in \mathbb{K}$ 有

$$D(k\alpha + l\beta) = k\alpha' + l\beta' = kD(\alpha) + lD(\beta).$$

所以我们定义的映射为线性映射.

定义 7.4.2 (商映射) 设 V 是 $\mathbb{K}$ 上的线性空间,M 是它的一个子空间，定义映射

$$\varphi: V \to V/M, \alpha \mapsto \alpha + M.$$

这个映射称为 V 到商空间 V/M 的**商映射 (quotient map)** 或**商映射 (natural map)**.

注意到，对于任意 $\alpha, \beta \in M; k, l \in \mathbb{K}$

$$\varphi(k\alpha + l\beta) = \overline{k\alpha + l\beta} = k\overline{\alpha} + l\overline{\beta} = k\varphi(\alpha) + l\varphi(\beta).$$

所以商映射也是一个线性映射. 显然商映射是一个满射，但在 $M \neq \{0\}$ 时不是单射. 此外，

$$\varphi(\alpha) = \overline{0} \Leftrightarrow \alpha \in M.$$

命题 7.4.1 设 U, V 是 $\mathbb{K}$ 上的线性空间，单射 $f \in \hom(U, V)$. U 内向量组 $\alpha_1, \alpha_2 \cdots, \alpha_n$ 线性无关的充要条件是 $f(\alpha_1), f(\alpha_2), \cdots, f(\alpha_n)$ 在 V 内线性无关.

证明 先证明必要性. 设

$$k_1\alpha_1 + k_2\alpha_2 + \cdots + k_n\alpha_n = 0.$$

于是

$$f(k_1\alpha_1 + k_2\alpha_2 + \cdots + k_n\alpha_n) = k_1 f(\alpha_1) + k_2 f(\alpha_2) + \cdots + k_n f(\alpha_n) = f(0) = 0 \in V.$$

而此时已知 $\alpha_1, \cdots, \alpha_n$ 线性无关，所以有 $k_1 = k_2 = \cdots = k_n = 0$. 所以 $f(\alpha_1), f(\alpha_2), \cdots, f(\alpha_n)$ 线性无关.

再证明充分性. 设

$$k_1 f(\alpha_1) + k_2 f(\alpha_2) + \cdots + k_n f(\alpha_n) = 0 \in U.$$

于是

$$k_1 f(\alpha_1) + k_2 f(\alpha_2) + \cdots + k_n f(\alpha_n) = f(k_1\alpha_1 + k_2\alpha_2 + \cdots + k_n\alpha_n) = f(0) = 0 \in V.$$

已知 $f(\alpha_1), \cdots, f(\alpha_n)$ 线性无关，所以 $k_1 = k_2 = \cdots = k_n = 0$，所以

$$k_1\alpha_1 + k_2\alpha_2 + \cdots + k_n\alpha_n = 0$$

中 $\alpha_1, \cdots, \alpha_n$ 线性无关. □

推论 7.4.1 设 U, V 是 $\mathbb{K}$ 上的线性空间,如果 $\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n$ 是 U 上的一组基,且有双射 $f \in \hom(U, V)$,则 $f(\varepsilon_1), f(\varepsilon_2), \cdots, f(\varepsilon_n)$ 是 V 上的一组基, 从而 $\dim V = \dim U$.

证明 在**命题7.4.1**中已知 $f(\varepsilon_1),\cdots,f(\varepsilon_n)$ 线性无关. 任给 $\beta\in V$，由于 f 为满射，故存在 $\alpha\in V$ 使得 $f(\alpha)=\beta$. 设

$$\alpha=k_1\varepsilon_1+k_2\varepsilon_2+\cdots+k_n\varepsilon_n,$$

则

$$\beta=f(\alpha)=k_1f(\varepsilon_1)+k_2f(\varepsilon_2)+\cdots+k_nf(\varepsilon_n).$$

所以 $f(\varepsilon_1),f(\varepsilon_2),\cdots,f(\varepsilon_n)$ 为 V 的一组基. □

7.4.2 线性空间的同构

从**命题7.4.1**中可知，如果两个线性空间内存在一个线性映射，且为双射，那么这两个线性空间的维数相等，且原来线性空间的基在被映到新线性空间后仍然是基，于是我们给出以下定义.

定义 7.4.3 设 U,V 是 $\mathbb{K}$ 上的线性空间，如果存在双射 $f\in\hom(U,V)$，则称 U 与 V **同构 (isomorphism)**，称 f 为 U 到 V 的同构映射.

例 7.4.3 把复数域 $\mathbb{C}$ 看为实数域 $\mathbb{R}$ 上的线性空间. 定义 $M_2(\mathbb{R})$ 的一个子集

$$A\coloneqq\left\{\begin{pmatrix}a&b\\-b&a\end{pmatrix}\middle|\,a,b\in\mathbb{R}\right\}.$$

显然 $M_2(\mathbb{R})$ 是 $\mathbb{R}$ 上的线性空间，A 是它的一个子空间. 然后我们定义一个映射

$$f:\mathbb{C}\to A,a+bi\mapsto\begin{pmatrix}a&b\\-b&a\end{pmatrix}.$$

由于

$$f[(a+bi)+(c+di)]=f[(a+b)+(c+d)i]=\begin{pmatrix}a+c&b+d\\-(b+d)&a+c\end{pmatrix}=\begin{pmatrix}a&b\\-b&a\end{pmatrix}+\begin{pmatrix}c&d\\-d&c\end{pmatrix}=f(a+bi)+f(c+di).$$

以及对于任意 $r\in\mathbb{R}$，有

$$f[(a+bi)]=f(ka+kbi)=\begin{pmatrix}ka&kb\\-kb&ka\end{pmatrix}=k\begin{pmatrix}a&b\\-b&a\end{pmatrix}=kf(a+bi).$$

所以 f 是一个线性映射. 又因为对于任意 $\mathbb{C}$ 中的元素，都有唯一一个矩阵与之对应，所以显然 f 是一个双射. 也就是说，$\mathbb{C}$ 与 $M_2(\mathbb{R})$ 是同构的，f 为从 $\mathbb{C}$ 到 $M_2(\mathbb{R})$ 的一个同构映射.

定理 7.4.1 设 U,V 是 $\mathbb{K}$ 上的线性空间，f 是 U 到 V 的同构映射，则其逆映射 f^{-1} 是 V 到 U 的同构映射.

证明 由于 f 是双射，所以 f 可逆，即 f^{-1} 存在，且仍然为双射. 对于任意 $\alpha,\beta\in V;k,l\in\mathbb{K}$ 有

$$f[kf^{-1}(\alpha)+lf^{-1}(\beta)]=k\alpha+l\beta.$$

所以

$$f^{-1}(k\alpha+l\beta)=f^{-1}\circ f[kf^{-1}(\alpha)+lf^{-1}(\beta)]=kf^{-1}(\alpha)+lf^{-1}(\beta).$$

于是 $f^{-1}:U\to V$ 是线性映射，且为双射. □

不难得到以下推论.

推论 7.4.2 U,V,W 是 $\mathbb{F}$ 上的线性空间. 如果 $f:U\to V$ 和 $g:V\to W$ 都是同构映射，则 $g\circ f:U\to W$ 也是同构映射.

上述事实说明，同构关系是一个等价关系. 所以当我们只研究线性空间的加法、数乘运算而不涉及具体内容时，同构的线性空间可以看作是“同一个”线性空间. 如果 U 是一个 n 维线性空间，则与之同构的所有线性空间都是 n 维线性空间. 如果 U,V 都与 $\mathbb{F}^n$ 同构，则 U 与 V 同构.

例 7.4.4 设 V 是 $\mathbb{K}$ 上的 n 维线性空间. 证明 V 内两组基 $\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$ 和 $\eta_1,\eta_2,\cdots,\eta_n$ 之间的过渡矩阵 T 是可逆的.

证明 我们可以找到 V 到 $\mathbb{K}^n$ 的同构映射 $f:V\to\mathbb{K}^n$，对于 V 中的每一个元素 α 都映到它关于 $\varepsilon_1,\cdots,\varepsilon_n$ 的坐标，即 $f(\alpha)=(a_1,a_2,\cdots,a_n)$. 通过 $f(\eta_1),f(\eta_2),\cdots,f(\eta_n)$ 我们分别可以得到 T 每一列列向量，由于 $\eta_1,\cdots,\eta_n$ 也是基，所以 T 的列向量线性无关，故可逆. □

注 7.4.2 在两个线性空间 V,W 中，如果只有一个唯一确定的同构，那么只有两种情况：1.$V=W=\{0\}$;2.$\dim V=\dim W=1$ 且域 $\mathbb{F}$ 是一个二元域.

7.4.3　线性映射的核与余核

定义 7.4.4 (核与余核) 设 U,V 是 $\mathbb{F}$ 上的线性空间，$f\in\hom(U,V)$.

$$\ker f:=\{\alpha\in U|f(\alpha)=0\in V\}.$$

称 $\ker f$ 为线性映射 f 的**核 (kernel)**. 再定义商空间

$$\mathrm{coker}f:=V/\mathrm{Im}f$$

称为 f 的**余核 (cokernel)**.

定理 7.4.2 设 U,V 是 $\mathbb{F}$ 上的线性空间，$f\in\hom(U,V)$ 则

1. $\ker f$ 是 U 的子空间，f 是单射的充分必要条件是 $\ker f=\{0\}$;
2. $\mathrm{Im}f$ 是 V 的子空间，f 是满射的充分必要条件是 $\mathrm{coker}f=\{\bar{0}\}$.

证明 1. 先证明 $\ker f$ 是 U 的子空间. 设 $\alpha,\beta\in\ker f$. 按定义 $f(\alpha)=0,f(\beta)=0$. 由于 $f\in\hom(U,V)$ 所以 $f(\alpha)+f(\beta)=f(\alpha+\beta)=0$ 故 $\alpha+\beta\in\ker f$. 又对于任意的 $k\in\mathbb{F}$，$kf(\alpha)=f(k\alpha)=0$. 所以 $\ker f$ 为 U 的子空间. 接下来证明必要性，如果 f 是单射，若存在 $\alpha,\beta\in U$ 使得 $f(\alpha)=f(\beta)=0$，则 $\alpha=\beta$，而 $f(0)=0$ 所以 $\alpha=\beta=0$. 再证明充分性，若 $\ker f=\{0\}$. 再设 $\alpha,\beta\in\ker f$，则 $f(\alpha)-f(\beta)=f(\alpha-\beta)=0$，而 $\alpha-\beta\in\{0\}$ 所以 $\alpha=\beta=0$.

2. 先证明 $\mathrm{Im}f$ 是子空间. 对于任意 $\alpha,\beta\in\mathrm{Im}f;k,l\in\mathbb{F}$,存在 $\alpha',\beta'\in U$ 满足 $f(\alpha')=\alpha,f(\beta')=\beta$. 由于 $f\in\hom(U,V)$ 所以 $kf(\alpha')+lf(\beta')=f(k\alpha'+l\beta')=k\alpha+l\beta\in\mathrm{Im}f$. 于是 $\mathrm{Im}f$ 对加法、数乘封闭，故为 V 的子空间.f 为满射等价于 $V=\mathrm{Im}f$，这等价于 $\mathrm{coker}f=V/\mathrm{Im}f=0+\mathrm{Im}f=\{\bar{0}\}$. □

例 7.4.5 给定数域 $\mathbb{K}$ 上的 $m\times n$ 矩阵 A，并给出线性映射 $f_A:\mathbb{K}^n\to\mathbb{K}^m,f_A(X)=AX$. 于是对应的余核就是

$$\ker f=\{X\in\mathbb{K}^n|AX=0\}.$$

即齐次线性方程组 $AX=0$ 的解空间. 再设单位列向量组 $E_1,E_2,\cdots,E_n$，以及 A 的列向量组 $\alpha_1,\alpha_2,\cdots,\alpha_n$ 于是映射 $f_A(E_i)=\alpha_i,(i=1,2,\cdots,n)$. 由于 $E_1,E_2,\cdots,E_n$ 张成 $\mathbb{K}^n$ 所以

$$\mathrm{Im}f_A=\mathrm{span}\{\alpha_1,\alpha_2,\cdots,\alpha_n\}.$$

由于线性方程组 $AX=B$ 有解等价于 $B\in\mathrm{Im}f_A$. 而

$$\mathrm{coker}f_A=\mathbb{K}^m/\mathrm{span}\{\alpha_1,\alpha_2,\cdots,\alpha_n\}$$

所以当 $\mathrm{coker}f_A=\{0\}$ 时，任取 $\mathbb{K}^m$ 中的一个元素，都有对应的解 X（可能不止一个）. 而 $\mathrm{coker}f_A$ 中的元素越多，则以系数矩阵为 A 的方程有解的可能性就越小. 所以我们发现，可以用 coker 的大小来衡量以 A 为矩阵的线性方程组有解的可能性.

我们再指出，关于 V 到 V/M 的商映射 φ，$\ker\varphi=M,\mathrm{Im}\varphi=V/M$.

定理 7.4.3 设 U,V,W 是 $\mathbb{F}$ 上的线性空间，且 $V=U\oplus W$，则映射 $f:u\mapsto u+W,u\in U$ 是 U 和 U/L 之间的同构映射.

证明 由于

$$f(\alpha x+\beta y)=(\alpha x+\beta y)+W=(\alpha x+W)+(\beta y+W)=\alpha f(x)+\beta f(y),$$

所以 f 是一个线性映射.

任取 $g\in U/W$，于是存在 $x\in U,y\in W$ 使得 $g=x+y$. 于是

$$g\in(x+W)+(y+W)=(x+W)+W=x+W=f(x),$$

这说明 f 是满射.

最后，对于任意 $u\in\ker f$. 于是 $u+W=W$，再由于 $u\in U$ 以及 $U\cap W=\{0\}$，所以 $u=0$. 故 $\ker f=\{0\}$，于是 f 是单射. □

定理 7.4.4 设 U,V 是 $\mathbb{F}$ 上的线性空间，$f\in\hom(U,V)$，则 $U/\ker f$ 与 $\mathrm{Im}f$ 同构.

证明 定义 $\tau:U\ker f\to\mathrm{Im}f,\tau(\alpha+\ker f)=f(\alpha)$. 设 $\alpha,\beta\in U$ 满足

$$\alpha+\ker f=\beta+\ker f.$$

再设 $m\in\ker f$，满足 $\beta=\alpha+m$ 于是有

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

这说明

$$\tau(\alpha+\ker f)=\tau(\beta+\ker f).$$

所以映射 τ 是良定义的.

接下来证明 τ 是线性映射. 对于任意 $\overline{\alpha},\overline{\beta}\in U/kerf$. 以及任意 $k,l\in\mathbb{F}$

$$\tau(k\overline{\alpha}+l\overline{\beta})=f(k\alpha+l\beta)=kf(\alpha)+lf(\beta)=k\tau(\overline{\alpha})+l\tau(\overline{\beta}).$$

再证明 τ 是单射. 设 $\overline{\alpha},\overline{\beta}\in U/kerf$.

$$\tau(\overline{\alpha})=\tau(\overline{\beta})\Rightarrow f(\alpha)=f(\beta).$$

再加上 $f(\alpha)-f(\beta)=f(\alpha-\beta)=0$，所以 $\alpha-\beta\in\ker f$，于是

$$\alpha+\ker f=\beta+\ker f.$$

τ 也为满射，因为对于任意 $\alpha'\in\mathrm{Im}f$ 都存在 $\alpha\in U$，满足 $f(\alpha)=\tau(\alpha+\ker f)=\alpha'$.

综上，映射 τ 为 $U/\ker f$ 到 $\mathrm{Im}f$ 的同构映射. □

推论 7.4.3 设 U,V 是 $\mathbb{F}$ 上的两个线性空间，$\dim U=n, f\in\hom(U,V)$，则

$$\dim\ker f+\dim\mathrm{Im}f=\dim U.$$

推论 7.4.4 设 $A\in M_{m,n}(\mathbb{K}), \mathrm{rank}A=r$，则齐次方程组 $AX=0$ 基础解系中向量的个数为 $n-r$.

证明 由于 $\mathrm{Im}f_A=\mathrm{span}\{\alpha_1,\alpha_2,\cdots,\alpha_n\}$ 其中 $\alpha_1,\cdots,\alpha_n$ 是 A 的列向量. 所以 $\dim\mathrm{Im}f=r$. 由于

$$\dim\mathbb{K}^n/\ker f_A=\dim\mathrm{Im}f\Leftrightarrow n-r=\dim\ker f_A,$$

并且 $\ker f_A$ 刚好是 $AX=0$ 的解空间，故推论得证. □

注 7.4.3 我们在实际运用中可能会用到上述公式的变形：$\mathrm{rank}A+\dim\ker f_A=n$.

在介绍完线性映射核的维数公式后，我们又可以给出以下线性映射的定理.

引理 7.4.1 设 U,V 是 $\mathbb{F}$ 上的两个线性映射，再设 $f\in\hom(U,V)$ 以及 $\varepsilon_1,\cdots,\varepsilon_n$ 是 U 上的一组基，则

$$\mathrm{Im}f=\mathrm{span}\{f(\varepsilon_1),\cdots,f(\varepsilon_n)\}.$$

证明 任给 $\alpha\in U$，设

$$\alpha=k_1\varepsilon_1+\cdots+k_n\varepsilon_n.$$

两边同时作用 f，则

$$f(\alpha)=k_1f(\varepsilon_1)+\cdots+k_nf(\varepsilon_n)\in\mathrm{span}\{f(\varepsilon_1),\cdots,f(\varepsilon_n)\}.$$

所以 $\mathrm{Im}f\subset\mathrm{span}\{f(\varepsilon_1),\cdots,f(\varepsilon_n)\}$. 又因为 $f(\varepsilon_i)$ 的任意线性组合

$$l_1f(\varepsilon_1)+\cdots+l_nf(\varepsilon_n)=f(l_1\varepsilon_1+\cdots+l_n\varepsilon_n)\in\mathrm{Im}f.$$

故 $\mathrm{span}\{f(\varepsilon_1),\cdots,f(\varepsilon_n)\}\subset\mathrm{Im}f$. 于是我们就证明了等号成立. □

这里再提一下求 $\mathrm{Im}f$ 基的一个方法. 在 $\ker f$ 中取一组基 $\alpha_1,\cdots,\alpha_m$，将其扩充为 U 的一组基 $\alpha_1,\cdots,\alpha_m,\alpha_{m+1},\alpha_n$，那么 $\mathrm{Im}f=\mathrm{span}\{f(\alpha_1),\cdots,f(\alpha_m),f(\alpha_{m+1},\cdots,f(\alpha_n))\}=\mathrm{span}\{f(\alpha_{m+1}),\cdots,f(\alpha_n)\}$. 由于 $\dim(\mathrm{Im}f)=\dim V=\dim(\ker f)=n-m$，所以 $f(\alpha_{m+1}),\cdots,f(\alpha_n)$ 就是 $\mathrm{Im}f$ 的一组基.

定理 7.4.5 如果 U,V 都是 $\mathbb{F}$ 上的 n 维线性空间，$f\in\hom(U,V)$，则 f 为单射当且仅当 f 为满射.

证明 f 为单射等价于 $\ker f=\{0\}$，这等价于

$$\dim(\mathrm{Im}f)=\dim U=\dim V.$$

这又等价于 $\mathrm{Im}f=V$，等价于 f 是满射. □

7.4.4 线性映射的运算

为了更好的研究线性映射，我们定义线性映射的运算.

定义 7.4.5 设 U, V 是 $\mathbb{F}$ 上的线性空间.

1. 给定 $f, g \in \hom(U, V)$，定义它们之间**加法**:

$$(f+g)(\alpha) \coloneqq f(\alpha) + g(\alpha), \quad \forall \alpha \in V.$$

在这个定义下，$f + g \in \hom(U, V)$, 因为

$$(f+g)(k\alpha + l\beta) = k(f+g)(\alpha) + l(f+g)(\beta).$$

2. 给定 $f \in \hom(U, V)$，我们定义 k 与 f 的**数乘**:

$$(kf)(\alpha) \coloneqq kf(\alpha), \quad \forall \alpha \in U.$$

3. 设 U, V, W 是 K 上线性空间. 设 $f \in \hom(U, V), g \in \hom(V, W)$，我定义它们的 g 与 f 的复合为**乘法**:

$$(gf)(\alpha) = g(f(\alpha)), \quad \alpha \in U.$$

显然 $gf \in \hom(U, W)$, 因为

$$(gf)(k\alpha + l\beta) = k(gf)(\alpha) + l(gf)(\beta).$$

注意，只有当 f, g 满足复合的条件时才可以相乘. 在乘法有意义时，乘法运算还满足结合律和分配律，

$$(fg)h = f(gh), \quad f(g+h) = fg + fh, \quad (f+g)h = fh + gh, \quad k \in \mathbb{K}, k(fg) = (kf)g = f(kg).$$

注意，线性映射的乘法一般没有交换律和消去律.

在定义 $U \to V$ 的**零元** $\hat{0}(\alpha) = 0 \in V$ 后不难验证，$\hom(U, V)$ 上关于线性映射的加法和数乘满足 8 条运算法则，故 $\hom(U, V)$ 是一个关于我们前面所定义的加法和数乘的线性空间，并且显然还有 $f + \hat{0} = f, (-f) + f = \hat{0}$.

7.4.5 线性映射的矩阵

在把矩阵抽象为线性映射后，我们应该还要从抽象的线性映射回到具体的矩阵.

定理 7.4.6 设 U, V 分别是 $\mathbb{F}$ 上的 n 维和 m 维线性空间，取 U 上的一组基 $\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n$.

1. 设 $f, g \in \hom(U, V)$. 如果有 $f(\varepsilon_i) = g(\varepsilon_i), i = 1, 2, \cdots, n$ 则有 $f \equiv g$. 即如果两个线性映射 f, g 作用在某一组基上的效果相同，则这两个线性映射其实是同一个线性映射.

2. 对于 V 中的任意 n 个向量 $\alpha_1, \alpha_2, \cdots, \alpha_n$，则一定可以找到唯一的 $f \in \hom(U, V)$，满足 $f(\varepsilon_i) = \alpha_i$, $(i = 1, 2 \cdots, n)$.

证明 1. 设任意的 $\alpha = a_1\varepsilon_1 + \cdots + a_n\varepsilon_n$. 则

$$f(\alpha) = f\left(\sum_{i=1}^{n} a_i\varepsilon_i\right) = \sum_{i=1}^{n} a_i f(\varepsilon_i) = \sum_{i=1}^{n} a_i g(\varepsilon_i) = g(\alpha).$$

2. 设 $\alpha = x_1\varepsilon_1 + \cdots + x_n\varepsilon_n$，再定义映射 $f(\alpha) = \sum\limits_{i=1}^{n} x_1\alpha_i$. 再设 $\beta = \sum_{i=1}^{n} y_i\beta_i$ 由于

$$\begin{aligned} f(k\alpha + l\beta) &= f\left(\sum_{i=1}^{n}(kx_i + ly_i)\right) = \sum_{i=1}^{n}(kx_i + ly_i)\alpha_i = k\left(\sum_{i=1}^{n} x_i\alpha_i\right) + l\left(\sum_{i=1}^{n} y_i\alpha_i\right) \\ &= kf(\alpha) + lf(\beta). \end{aligned}$$

故我们定义的映射是一个线性映射. 现在显然有 $f(\varepsilon_i) = \alpha_i$，其唯一性可由 (1.) 直接得证. □

从**定理** 7.4.6可得，只要确定映射在 U 中一组基上的作用效果，那么这个映射就唯一确定.$f(\varepsilon_i)$ 在 V 中可以表示为一个基底 η_i 的线性组合，即得到线性方程组

$$f(\varepsilon_i) = \sum_{j=1}^{m} a_{ji}\eta_j, i = 1, 2, \cdots, n. \tag{7.4.1}$$

再令矩阵 $A = (a_{ij})$. 于是上面的 n 个等式可以借助矩阵乘法形式的表示为

$$(f(\varepsilon_1), \cdots, f(\varepsilon_n)) = (\eta_1, \cdots, \eta_m)A.$$

我们称 $m \times n$ 矩阵为**线性映射 f 在基 η_i 下的矩阵**. 类似基变换一样，为了满足矩阵的形式乘法，我们把方程第 i 行系数排入矩阵的第 i 列.

显然在 U, V 中的基都取定之后，任取一个矩阵 A，就可以确定一个方程组(7.4.1)，于是 f 在基 ε_i 上的作用效果确定，f 就被唯一确定. 反过来，如果 f 在基 ε_i 下的作用效果确定，反之，若线性映射 f 已经确定，则 f 作用在基上 ε_i 上的效果 $f(\varepsilon_i)$ 唯一确定，由方程组(7.4.1)可知，矩阵 A 的每一列都唯一确定，故 f 在取定基后也唯一决定了一个矩阵 A.

于是，就可以构造一个双射 $\sigma : \hom(U, V) \to M_{m,n}(\mathbb{F})$. 即在 U, V 中基取定之后，线性映射和线性映射在给的基下的矩阵是一一对应的.

为了方便，我们再引进一个记号：

$$f(\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n) \coloneqq (f(\varepsilon_1), f(\varepsilon_2), \cdots, f(\varepsilon_n)).$$

这个记号将使我们以后的推理简化，因为它具有结合律: 令 $\alpha \in U$

$$\begin{aligned} \alpha &= x_1\varepsilon_1 + \cdots + x_n\varepsilon_n = (\varepsilon_1, \cdot, \varepsilon_n)X, \\ f(\alpha) &= \sum_{i=1}^{n} x_i\varepsilon_i = (f(\varepsilon_1), \cdots, f(\varepsilon_n))X = [f(\varepsilon_1, \cdots, \varepsilon_n)]X. \end{aligned}$$

上面两个式子写在一起就是

$$[f(\varepsilon_1, \cdots, \varepsilon_n)]X = f[(\varepsilon_1, \cdots, \varepsilon_n)X].$$

接下来这个命题将把两个线性空间的维数与线性映射构成的集合的维数联系起来.

命题 7.4.2 1. 对于上面所说的双射 $\sigma : \hom(U, V) \to M_{m,n}(\mathbb{F})$. 任意 $f, g \in \hom(U, V)$ 和 $k, l \in \mathbb{F}$ 有

$$\sigma(kf + lg) = k\sigma(f) + l\sigma(g).$$

即 σ 是一个 $\hom(U, V)$ 到 $M_{m,n}(\mathbb{F})$ 的一个线性同构. 于是有

$$\dim \hom(U, V) = \dim M_{m,n}(\mathbb{F}) = mn.$$

2. 若 $f\in\hom(U,V),g\in\hom(V,W)$ 则在用 σ 同时表示映射 $\hom(U,V)\to M_{m,n}(F)$ 和 $\hom(V,W)\to M_{s,n}(\mathbb{F})$ 时有

$$\sigma(gf)=\sigma(g)\sigma(f).$$

证明 1. 设 $\sigma(f)=A,\sigma(g)=B$. 然后

$$\begin{aligned}[(kf+lg)\varepsilon_1,\cdots,(kf+lg)\varepsilon_n]&=[kf(\varepsilon_1)+lg(\varepsilon_1),\cdots,kf(\varepsilon_n)+lg(\varepsilon_n)]\\&=k[f(\varepsilon_1,),\cdots,f(\varepsilon_n)]+l[f(\varepsilon_1),\cdots,f(\varepsilon_n)]\\&=k(\eta_1,\eta_2,\cdots,\eta_m)A+l(\eta_1,\eta_2,\cdots,\eta_m)B\\&=(\eta_1,\eta_2,\cdots,\eta_m)(kA)+(\eta_1,\eta_2,\cdots,\eta_m)(lB)\\&=(\eta_1,\eta_2,\cdots,\eta_m)(kA+lB),\end{aligned}$$

这说明，$kf+lg$ 的矩阵是 $kA+lB=k\sigma(f)+l\sigma(g)$. 就得到了我们要的结论.

2. 在 U,V,W 内分别取一组基

$$\varepsilon_1,\cdots,\varepsilon_n;\quad \eta_1,\cdots,\eta_m;\quad \delta_1,\cdots,\delta_s.$$

再设

$$\begin{aligned}(f(\varepsilon_1),f(\varepsilon_2),\cdots,f(\varepsilon_n))&=(\eta_1,\eta_2,\cdots,\eta_m)A,\\(g(\eta_1),g(\eta_2),\cdots,g(\eta_m))&=(\delta_1,\delta_2,\cdots,\delta_s)B.\end{aligned}$$

于是

$$\begin{aligned}(gf(\varepsilon_1),gf(\varepsilon_2),\cdots,gf(\varepsilon_n))&=g(f(\varepsilon_1),f(\varepsilon_2),\cdots,f(\varepsilon_n))\\&=g[(\eta_1,\eta_2,\cdots,\eta_m)A]=[g(\eta_1,\eta_2,\cdots,\eta_m)]A\\&=[(g(\eta_1),g(\eta_2),\cdots,g(\eta_m))]A\\&=[(\delta_1,\delta_2,\cdots,\delta_s)B]A=(\delta_1,\delta_2,\cdots,\delta_s)(BA).\end{aligned}$$

这说明 $gf\in\hom(U,W)$，并且在给定基下的矩阵 $BA=\sigma(g)\sigma(f)$. 故得结论. □

定理 7.4.7 设 U,V 是 $\mathbb{F}$ 上两个有限维线性空间，则

$$\dim\hom(U,V)=\dim U\cdot\dim V.$$

证明 设 $\dim U=n,\dim V=m$，于是对于任意线性映射 $f\in\hom(U,V)$. 在给定基下得矩阵总为 $m\times n$ 矩阵，而从**命题7.4.2**可知，

$$\dim\hom(U,V)=mn=\dim U\cdot\dim V.$$

□

7.4.6　线性变换的概念

接下来考虑在同一个空间中的线性映射，即令 $U=V$. 这是非常重要的一种线性映射.

定义 7.4.6 (线性变换) 设 V 是 $\mathbb{F}$ 上的一个线性空间，则称线性映射 $\hat{A}:V\to V$ 为 V 上的**线性变换** **(linear transformation)**.V 上所有线性变换构成的集合记为 $\mathrm{end}(V)$.

注 7.4.4 在量子力学中，每一个可观测量都是一个线性变换，故我们采用相同的记号来记一个线性变换.

为了方便，有时我们记 $\hat{A}(\alpha)$ 为 $\hat{A}\alpha$.

接下来看几个例子.

例 7.4.6 设 $D^{\infty}(a,b)$ 是区间 (a,b) 全体任意次可微的实变函数 $f(x)$ 所组成的集合，它关于普通函数的加法和实数的乘法组成实数域 $\mathbb{R}$ 上的线性空间. 在 $D^{\infty}(a,b)$ 内定义一个变换

$$\hat{D} = \frac{\mathrm{d}}{\mathrm{d}x} : f(x) \mapsto f'(x).$$

即让 $D^{\infty}(a,b)$ 中的每一个向量 $f(x)$ 在变换 $\hat{D}$ 的作用下变成该函数的导函数. 根据数学分析中微商的性质，有

$$\begin{aligned}\hat{D}[f(x)+g(x)] &= \hat{D}[f(x)] + \hat{D}[g(x)],\\ \hat{D}[kf(x)] &= k\hat{D}[f(x)].\end{aligned}$$

这说明微商运算是线性空间 $D^{\infty}(a,b)$ 上的一个线性变换.

例 7.4.7 设 $C[a,b]$ 为闭区间 $[a,b]$ 内全体连续函数所组成的一个实数域上的线性空间. 在 $C[a,b]$ 上的定义的变换如下

$$\hat{S} : f(x) \mapsto \int_a^x f(t)\,\mathrm{d}t = F(x).$$

即让 $C[a,b]$ 内每一个向量 $f(x)$ 都对应到它的变上限积分. 根据定积分的性质，有

$$\begin{aligned}\hat{S}(f(x)+g(x)) &= \hat{S}f(x) + Sg(x);\\ \hat{S}(kf(x)) &= kSf(x).\end{aligned}$$

这说明变上限积分是线性空间 $C[a,b]$ 内的一个线性变换.

线性变换的运算与前面线性映射的运算是一致的，接下来介绍几种比较特殊的运算.

定义 7.4.7 设 V 是一个线性空间

1. 零变换 $\hat{0}$. 对于任意 $\alpha \in V$, 有 $\hat{0}\alpha = 0$. 这是 V 内的一个线性变换，称为**零变换 (zero transformation.)**

2. 单位变换 $\hat{E}$. 对任意 $\alpha \in V, \hat{E}\alpha = \alpha$. 这显示也是 V 内的一个线性变换，称为**单位变换**或**恒等变换 (identify transformation)**.

3. 数乘变换 $\hat{k}$. 定义 k 是 $\mathbb{F}$ 中一个固定的数. 对于任意 $\alpha \in V$，定义 $\hat{k}\alpha = k\alpha$. 显然这还是一个线性变换，称为**数乘变换**.

4. **投影变换 (projection transformation)**$\hat{P}$. 设 M 是 V 的一个子空间. 于是存在 M 的一个补空间 N，故对于任意 α，存在唯一的分解式

$$\alpha = \alpha_1 + \alpha_2, \quad (\alpha+1 \in M, \alpha_2 \in N).$$

定义

$$\hat{P}\alpha = \alpha_1.$$

不难验证这也是一个线性变换. 值得注意的是，投影变换依赖于直和分解式的具体形式.

由于线性变换是一类特殊且重要的线性映射，前面关于线性映射的知识都适用于线性变换，现在把一些重点再罗列一下.

1. 如果 $\hat{A} \in \mathrm{end}(V) = \mathrm{hom}(V, V)$，那么 $\ker \hat{A}$ 和 $\mathrm{Im}\hat{A}$ 都是 V 的子空间，且有

$$\dim \ker \hat{A} + \dim \mathrm{Im}\hat{A} = \dim V.$$

2.$\mathrm{end}(V)$ 内有加法、数乘运算，它关于此两种运算称为域 $\mathbb{F}$ 上的线性空间.$\mathrm{end}(V)$ 内任意两元素 $\hat{A}, \hat{B}$ 可做乘法 $\hat{A}\hat{B}$. 对于 $\hat{A} \in \mathrm{end}(V)$ 及 $k \in \mathbb{N}^+$ 定义 $\hat{A}^k = \underbrace{\hat{A}\hat{A}\cdots\hat{A}}_{k\text{个}\hat{A}}$. 并定义当 $\hat{A} \neq \hat{0}$ 时 $\hat{A}^0 = \hat{E}$. 对 $\mathbb{F}$ 上多项式 $f(x) = \sum\limits_{i=0}^{n} a_i x^i$. 定义

$$f(\hat{A}) = a_n \hat{A}^n + a_{n-1}\hat{A}^{n-1} + \cdots + a_0 \hat{E}.$$

并称之为**线性变换的多项式**.

3. 在线性空间 $U = V$ 下取一组基 $\varepsilon_1, \cdots, \varepsilon_n$，那么任意 $\hat{A} \in \mathrm{end}(V)$ 在此组基下的矩阵由下列式子定义

$$(\hat{A}\varepsilon_1, \cdots, \hat{A}\varepsilon_n) = (\varepsilon_1, \cdots, \varepsilon_n)A.$$

其中 A 的第 i 个列向量是 $\hat{A}\varepsilon_i$ 在基 $\varepsilon_1, \cdots, \varepsilon_n$ 下的坐标. 现在矩阵 A 和线性变换 $\hat{A}$ 之间是一一对应的.

于是我们也可以构造双射 $\sigma : \mathrm{end}(V) \to M_n(\mathbb{F})$. 满足以下性质

$$\sigma(k\hat{A} + l\hat{B}) = k\sigma(\hat{A}) + l\sigma(\hat{B}), \quad \sigma(\hat{A}\hat{B}) = \sigma(\hat{A})\sigma(\hat{B})$$

其中 $\hat{A}, \hat{B}$ 都是 V 上的线性变换.

因为单位变换 $\hat{E}$ 在基 ε_i 下的矩阵为单位矩阵 E，我们有

$$\hat{E} = \hat{A}\hat{B} \Leftrightarrow \sigma(\hat{A})\sigma(\hat{B}) = \sigma(\hat{B})\sigma(\hat{A}).$$

所以 $\hat{A}$ 为可逆线性变换的充分必要条件是 $\sigma(\hat{A})$ 是 $\mathbb{F}$ 上可逆 n 阶方阵，且 $\sigma(\hat{A}^{-1}) = [\sigma(A)]^{-1}$.

接下来给出两个有用的定理.

定理 7.4.8 设 $\hat{A}$ 是 $\mathbb{F}$ 上 n 维线性空间 V 上的线性变换，它在 V 的一组基 $\varepsilon_1, \cdots, \varepsilon_n$ 下的矩阵为 A，则

$$\dim \mathrm{Im}\hat{A} = \mathrm{rank}A.$$

证明 由于 $\hat{A}\alpha$ 可由 $\hat{A}\varepsilon_1, \cdots, \hat{A}\varepsilon_n$ 线性表示，因此

$$\mathrm{Im}A = \mathrm{span}\{\hat{A}\varepsilon_1, \cdots, \hat{A}\varepsilon_n\}.$$

由**定理7.1.2**可得

$$\dim \mathrm{span}\{\hat{A}\varepsilon_1, \cdots, \hat{A}\varepsilon_n\} = \mathrm{rank}A.$$

所以

$$\dim \mathrm{Im}\hat{A} = rankA.$$

注 7.4.5 有时为了方便，直接称 $\hat{A}$ 像的维数为线性变换 $\hat{A}$ 的**秩**，于是有 $\mathrm{rank}\hat{A} = \mathrm{rank}A$.

定理 7.4.9 设线性变换 $\hat{A}$ 在一组基下的矩阵为 A. 再设 α 在这组基下的坐标为 X，则 $A\alpha$ 在这组基下的坐标为 AX.

证明 设基为 $\varepsilon_1, \cdots, \varepsilon_n$，则

$$\begin{aligned} \hat{A}\alpha &= \hat{A}[(\varepsilon_1, \cdots, \varepsilon_n)X] = [\hat{A}(\varepsilon_1, \cdots, \varepsilon_n)]X \\ &= [(\varepsilon_1, \cdots, \varepsilon_n)A]X = (\varepsilon_1, \cdots, \varepsilon_n)(AX). \end{aligned}$$

就得到了我们要的结论. □

7.4.7 线性变换在不同基下的矩阵

同一个线性变换在不同基下的矩阵是不一样的，我们来探究这些矩阵之间的关系.

定理 7.4.10 设

$$\varepsilon_1, \cdots, \varepsilon_n,$$
$$\eta_1, \cdots, \eta_n.$$

是 $\mathbb{F}$ 上线性空间 V 的两组基. 由 ε_i 到 η_i 的过渡矩阵为 T，即 $(\eta_1, \cdots, \eta_n) = (\varepsilon_1, \cdots, \varepsilon_n)T$. 再设 A, B 分别是线性变换 $\hat{A} \in \mathrm{end}(V)$ 在基 ε_i, η_i 下的矩阵. 则有

$$B = T^{-1}AT.$$

证明 由线性变换在基下矩阵的定义，我们有

$$\hat{A}(\varepsilon_1, \cdots, \varepsilon_n) = (\varepsilon_1, \cdots, \varepsilon_n)A,$$
$$\hat{A}(\eta_1, \cdots, \eta_n) = (\eta_1, \cdots, \eta_n)B.$$

替换可得

$$\hat{A}[(\varepsilon_1, \cdots, \varepsilon_n)T] = [(\varepsilon_1, \cdots, \varepsilon_n)T]B$$

运用形式结合律，可得

$$\hat{A}[(\varepsilon_1, \cdots, \varepsilon_n)T] = [\hat{A}(\varepsilon_1, \cdots, \varepsilon_n)]T = [(\varepsilon_1, \cdots, \varepsilon_n)A]T = (\varepsilon_1, \cdots, \varepsilon_n)(AT) = (\varepsilon_1, \cdots, \varepsilon_n)(TB).$$

于是可得

$$AT = TB,$$

即 $B = T^{-1}AT$. □

有了这个定理，我们就可以给出以下定义.

定理 7.4.11 对 $\mathbb{F}$ 上的两个的两个 n 阶方阵 A, B，如果存在 $\mathbb{F}$ 上的可逆矩阵 T，使得 $B = T^{-1}AT$. 则称 A, B 矩阵**相似 (similar)**. 并记为 $A \backsim B$.

不难验证，矩阵相似具有以下性质.

1. 自反性：$A \backsim A$，因为 $A = E^{-1}BE$.
2. 对称性：$A \backsim B$ 则 $B \backsim A$. 因为 $A = T^{-1}BT$ 则有 $B = (T^{-1})^{-1}AT^{-1}$.
3. 传递性：$A \backsim B, B \backsim C$ 则 $A \backsim C$. 因为 $A = T_1^{-1}BT_1, C = T_2^{-1}BT_2$ 时，有

$$A = T_1^{-1}(T_2^{-1}CT_2)T_1 = (T_2T_1)^{-1}C(T_2T_1).$$

这说明矩阵之间的相似关系是一个等价关系，于是可以在 $M_n(\mathbb{F})$ 上构造矩阵的等价类，称为**相似类 (similar class)**. 下面这个定理阐明了相似类的实际意义.

定理 7.4.12 $\mathbb{F}$ 上两个 n 阶方阵相似的充分必要条件是，它们是 V 内某一线性变换 $\hat{A}$ 在两个基底下的矩阵.

证明 充分性已证明，现在证明必要性. 现在我们可以找到 V 内的一个线性变换 $\hat{A}$，使它在 V 的某一组基 $\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n$ 下的矩阵为 A. 由于 $B \backsim A$，故存在可逆矩阵 T，使得 $B = T^{-1}AT$. 令

$$(\eta_1, \eta_2, \cdots, \eta_n) = (\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n)T.$$

则 $\eta_1, \cdots, \eta_n$ 是 V 的一个基底，则 $\hat{A}$ 在这组新基下的矩阵为 $T^{-1}AT = B$. □

相似矩阵还有如下简单性质，可由定义直接证得.

命题 7.4.3 1. 如果 $A \backsim B$ 则 $A^m \backsim B^m, m \in \mathbb{N}^+$.

2. 如果 A, B 可逆，且 $A \backsim B$，则 $A^{-1} \backsim B^{-1}$.

3. 如果 $A \backsim B$，则它们的伴随矩阵也相似.

7.5 线性变换的特征值与特征向量

7.5.1 特征值与特征向量的概念

前面我们知道，矩阵之间的相似关系是一个等价关系. 一个自然的想法就是从等价类中找到一个最简单的矩阵作为代表元，显然，除了单位矩阵以外，最简单的矩阵就是对角矩阵. 在解决这个问题之前，我们需要一些前置知识.

定义 7.5.1 (特征值与特征向量) 设 V 是域 $\mathbb{F}$ 上的一个线性空间，且 $\hat{A}$ 是 V 上的一个线性变换，若对于 V 中的一个非零向量 α，存在 $\lambda \in \mathbb{F}$ 满足

$$\hat{A}\alpha = \lambda\alpha$$

则称 λ 是 $\hat{A}$ 的一个**特征值** (**eigenvalue**). 而 α 称为属于 λ 的一个**特征向量** (**eigenvector**).

注 7.5.1 1. 有时物理上可能会称特征值为**本征值**，称特征向量为**本征向量**.

2. 注意特征向量 α 必须是非零向量.

3. λ 必须属于域 $\mathbb{F}$，否则没有意义.

从几何上看，如果 V 上的线性变换 $\hat{A}$ 作用在 α 上的效果为 $\lambda\alpha$，那么该线性变换其实就是将向量 α 拉长了 λ 倍（也有可能是压缩）.

注意到，对于上述的一个特征值 λ，可能会有好几个属于该特征值的特征向量，这些向量可能组成一个集合，我们记该集合为

$$V_\lambda = \{\alpha \in V | \hat{A}\alpha = \lambda\alpha\}.$$

显然，$0 \in V_\lambda$，且 V_λ 关于向量的数乘，加法封闭，故 V_λ 是 V 的一个子空间，于是我们给出如下定义

定义 7.5.2 (特征子空间) 设 V 是 $\mathbb{F}$ 上的一个线性空间，$\hat{A}$ 是 V 上的一个线性变换，$\lambda \in \mathbb{F}$ 是 V 的一个特征值，则称

$$V_\lambda = \{\alpha \in V | \hat{A}\alpha = \lambda\alpha\}$$

为 V 的一个**特征子空间** (**eigensubspace**)

立即可以得到两个明显的事实：

1. $\lambda \in \mathbb{F}$ 是 $\hat{A} \in \hom(V,V)$ 的一个特征值的充分必要条件是 $V_\lambda \neq \{0\}$. 如果 $\alpha \in V_\lambda$，则 α 是属于 λ 的特征向量.

2. 属于 λ 的所有特征向量都可以被 V_λ 上的一组基表示.

接下来需要解决两个问题，一是找出所有满足 $V_\lambda \neq \{0\}$ 的 $\lambda \in \mathbb{F}$，即找出 V 上线性变换 $\hat{A}$ 的所有特征值. 二是找出特征子空间中的一组基.

由**定理7.4.9**可知以下命题.

命题 7.5.1 设 V 是 $\mathbb{F}$ 上 n 维线性空间，$\hat{A}$ 是 V 上的一个线性变换，且 $\varepsilon_1,\cdots,\varepsilon_n$ 是 V 的一个基底. 又有 $\alpha \in V$ 在基底 ε_i 下的坐标为 X，则

$$\hat{A}\alpha = \lambda X \Longleftrightarrow AX = \lambda X,$$

其中 A 是 $\hat{A}$ 在基 $\varepsilon_1,\cdots,\varepsilon_n$ 下的矩阵.

证明 由

$$\alpha = (\varepsilon_1,\cdots,\varepsilon_n)X.$$

可得

$$\hat{A}\alpha = \hat{A}[(\varepsilon_1,\cdots,\varepsilon_n)X] = [\hat{A}(\varepsilon_1,\cdots,\varepsilon_n)]X = [(\varepsilon_1,\cdots,\varepsilon_n)A]X,$$

其中 A 是 $\hat{A}$ 在基 ε_i 下的矩阵. 再由于

$$\hat{A}[(\varepsilon_1,\cdots,\varepsilon_n)X] = \lambda[(\varepsilon_1,\cdots,\varepsilon_n)X] = (\varepsilon_1,\cdots,\varepsilon_n)(\lambda X),$$

代回可得

$$(\varepsilon_1,\cdots,\varepsilon_n)(AX) = (\varepsilon_1,\cdots,\varepsilon_n)(\lambda X).$$

所以

$$AX = \lambda X,$$

其中 $\alpha \neq 0 \Leftrightarrow X \neq 0$. 而矩阵 A 和 $\hat{A}$ 在基 ε_i 下是一一对应的，于是就得到了我们要的等价关系. □

由这个命题可以得到以下等价关系

$$V_\lambda \neq \{0\} \Leftrightarrow \exists \alpha \in V, \hat{A}\alpha = \lambda\alpha \Leftrightarrow \exists X \in \mathbb{F}^n, X \neq 0, AX = \lambda X.$$

考察等式 $AX = \lambda X$，移项可得 $(\lambda E - A)X = 0$. 这显然是一个系数矩阵 A 为 $n \times n$ 的齐次线性方程组，其有非零解的充要条件是 $|\lambda E - A| = 0$. 将 λ 看作未知量，由上面的等价关系，我们知道以下结论.

命题 7.5.2 1. $\lambda \in \mathbb{F}$ 为 $\hat{A}$ 的特征值的充分必要条件是 $|\lambda E - A| = 0$.

2. $\alpha \in V$，α 满足 $\hat{A}\alpha = \lambda\alpha$ 的充分必要条件是 α 在基 ε_i 下的坐标 X 满足 $(\lambda E - A)X = 0$.

前面已经指出，在一组给定的基下 α 与坐标 X 是一一对应的，于是我们可以构造这样一个同构映射 $f: V \to \mathbb{F}^n$. 如果 $\alpha \in V_\lambda$，那么其坐标 $X = f(\alpha)$ 满足 $(\lambda E - A)X = 0$. 于是 $f(V_\lambda)$ 刚好就是 $(\lambda E - A)X = 0$ 的解空间. 然后我们找到该解空间的一组基，其原像 $f^{-1}(X_i) = \varepsilon_i$ 就是子空间的的一组基. 这样，寻找特征向量的问题就转换为寻找解空间基的问题，而后者是很好解决的.

再讲具体的计算之前，我们注意到 $|\lambda E - A|$ 这个行列式起到了关键性的作用，于是我们把它抽象出来，就有了如下定义.

定义 7.5.3 (矩阵的特征值与特征向量) 设 $A=(a_{ij})$ 是 $\mathbb{F}$ 上的 n 阶方阵，令

$$f(\lambda)=|\lambda E-A|,$$

从行列式的展开式可知 $f(\lambda)$ 为 λ 的一个多项式，我们称之为**特征多项式 (characteristic polynomial)**，并称多项式在 $\mathbb{F}$ 上的所有根都为 A 的**特征值**. 再称满足 $(\lambda E-A)X=0$ 的不为 0 的 $X\in\mathbb{F}^n$ 为 A 的**特征向量**.

注 7.5.2 这里定义的是矩阵的特征值和特征向量，请与线性变换的特征值与特征向量区分开.

7.5.2 特征值与特征向量的计算

首先给出线性变换关于特征值的一些运算法则.

命题 7.5.3 设 V 是域 $\mathbb{F}$ 上的线性空间（可以是无线维的），$\hat{A}$ 是 V 上的一个可逆线性变换，则

1. 0 不是 $\hat{A}$ 的特征值.
2. 如果 λ_0 是 $\hat{A}$ 的一个特征值，那么 λ_0^{-1} 是 $\hat{A}^{-1}$ 的一个特征值.

证明 1. 如果 0 是 $\hat{A}$ 的一个特征值，则存在 $\xi\in V$，且 $\xi\neq 0$ 使得 $\hat{A}\xi=0\xi=0$. 又因为 $\hat{A}(0)=0$，于是 $\hat{A}$ 不是单射，这与 $\hat{A}$ 可逆矛盾. 故 0 不是 $\hat{A}$ 的特征值.

2. 设 λ_0 是 $\hat{A}$ 的一个特征值，则有 $\xi\in V$，且 $\xi\neq 0$ 使得 $\hat{A}\xi=\lambda_0\xi$. 在等号两边同时作用 $\hat{A}^{-1}$，得到

$$\hat{A}^{-1}(\hat{A}\xi)=\hat{A}^{-1}(\lambda_0\xi),$$

从而 $\xi=\lambda_0\hat{A}^{-1}\xi_0$. 由于 $\lambda_0\neq 0$，因此 $\hat{A}^{-1}\xi=\lambda_0^{-1}\xi$，即 λ_0^{-1} 是 $\hat{A}^{-1}$ 的一个特征值. □

命题 7.5.4 设 V 是 $\mathbb{F}$ 上的一个线性空间，$\hat{A}$ 是 V 上的一个线性变换，$\lambda_0\in\mathbb{F}$ 是 $\hat{A}$ 的一个特征值，则 λ_0^m 是 $\hat{A}^m, m\in\mathbb{N}$ 的特征值.

证明 设 λ_0 是 $\hat{A}$ 的一个特征值，则存在 $\xi\in V$，且 $\xi\neq 0$ 使得 $\hat{A}\xi=\lambda_0\xi$. 那么

$$\hat{A}^m\xi=\hat{A}(\hat{A}(\hat{A}\cdots\hat{A}(\xi))\cdots)=\hat{A}(\hat{A}(\hat{A}\cdots(\lambda_0\xi))\cdots)=\cdots=\lambda_0^m\xi.$$

故 λ_0^m 是 $\hat{A}^m$ 的一个特征值. □

接下来再看一看矩阵关于特征值的一些运算.

命题 7.5.5 设 A 是 $\mathbb{F}$ 上的一个 n 阶可逆矩阵,，则

1. 如果 A 有特征值, 则 0 不是 A 的特征值.
2. 如果 λ_0 是 A 的一个 l 重特征值，则 λ_0^{-1} 是 A^{-1} 的一个特征值.

证明 1. 由于 A 是 n 阶可逆矩阵，所以

$$|0E-A|=|-A|=(-1)^n|A|\neq 0.$$

故 0 不可能是 A 的特征值.

2. 由 λ_0 是 A 的一个 l 重特征值，所以它是多项式 $|\lambda E-A|=0$ 的一个 l 重根. 于是有

$$|\lambda E-A|=(\lambda-\lambda_0)g(\lambda),$$

其中 $g(\lambda)$ 是一个关于 λ 的 $n-l$ 次多项式，并且不含因式 $(\lambda-\lambda_0)$.

把 $g(\lambda)$ 在复数域上做因式分解，由代数基本定理，可得

$$|\lambda E-A|=(\lambda-\lambda_0)^l(\lambda-\lambda_1)^{l_1}\cdots(\lambda-\lambda_m)^{l_m},$$

其中 $\lambda_1,\cdots,\lambda_m$ 是两两不相等的复数，且它们都不等于 λ_0，以及 $l_1+\cdots+l_m=n-l$.

在将 λ 用 $\frac{1}{\lambda}$ 代入，得到

$$\left|\frac{1}{\lambda}E-A\right|=\left(\frac{1}{\lambda}-\lambda_0\right)^l\left(\frac{1}{\lambda}-\lambda_1\right)^{l_1}\cdots\left(\frac{1}{\lambda}-\lambda_m\right)^{l_m}.$$

从而 A^{-1} 的特征多项式为

$$\begin{aligned}|\lambda E-A^{-1}|&=\left|A^{-1}(-\lambda)\left(\frac{1}{\lambda}E-A^{-1}\right)\right|=(-1)^n\lambda^n|A^{-1}|\left|\frac{1}{\lambda}E-A\right|\\&=(-1)^n\lambda^n\left(\frac{1}{\lambda}-\lambda_0\right)^l\left(\frac{1}{\lambda}-\lambda_1\right)^{l_1}\cdots\left(\frac{1}{\lambda}-\lambda_m\right)^{l_m}\\&=|A^{-1}|(-1+\lambda_0\lambda)^l(-1+\lambda_0\lambda)^{l_1}\cdots(-1+\lambda_0\lambda)^{l_m}\\&=|A^{-1}|\lambda_0^l\lambda_1^{l_1}\cdots\lambda_m^{l_m}\left(\lambda-\frac{1}{\lambda_0}\right)^l\left(\lambda-\frac{1}{\lambda_1}\right)^{l_1}\cdots\left(\lambda-\frac{1}{\lambda_m}\right)^{l_m}.\end{aligned}$$

这说明 $\frac{1}{\lambda_0}$ 是 A^{-1} 的特征多项式的一个 l 重根，表明它是 A^{-1} 的 l 重特征值. □

注 7.5.3 关于 λ 用 $\frac{1}{\lambda}$ 代入的合理性见前面的多项式环.

命题 7.5.6 A 是 $\mathbb{F}$ 上的一个 n 阶方阵，若 λ_0 是 A 的一个特征值，则 λ_0^m 是 A^m 的一个特征值.

证明 由 λ_0 是 A 的一个特征多项式可得，存在 $X\in\mathbb{F}^n$ 使得 $AX=\lambda_0X$. 于是

$$A^mX=A(A(A\cdots(AX))\cdots)=A(A(A\cdots(\lambda_0X))\cdots)=\cdots==\lambda_0(\lambda_0(\lambda_0\cdots(\lambda_0X))\cdots)=\lambda_0^mX.$$

故 λ_0^m 是方阵 A^m 的一个特征值. □

注 7.5.4 如果 λ_0 是 A 的一个 l 重特征值，且 $m=2$，则 λ_0^2 也是 A^2 的一个 l 重特征值.

以上命题在特征值的计算中经常用到.

现在将求线性变换 $\hat{A}$ 的特征值和特征向量的的方法归纳如下：

1. 在 V 中给定一组基 $\varepsilon_1,\cdots,\varepsilon_n$，求 $\hat{A}$ 在这组基下的矩阵 A.
2. 计算特征多项式 $|\lambda E-A|=0$，得到根 $\lambda_1,\cdots,\lambda_n$，就是 $\hat{A}$ 的所有特征值 (注意可能会有重特征值).
3. 将求得的每一个特征值代入 $|\lambda_iE-A|X=0,i=1,2,\cdots,n$，对每一个特征值，都会得到一个齐次线性方程组.
4. 求得每一个特征值所对应的线性方程组的基础解系.
5. 将每个基础解系作为基 ε_i 的坐标写出一组向量，这组向量就是 V_{λ_i} 的一组基.

例 7.5.1 设三维空间 V 上的一个线性变换 $\hat{A}$ 在基 $\varepsilon_1,\varepsilon_2,\varepsilon_3$ 下的矩阵为

$$A=\begin{pmatrix}1&2&2\\2&1&2\\2&2&1\end{pmatrix}.$$

求 $\hat{A}$ 的特征值与对应的特征向量.

解 先求特征多项式的特征根.

$$|\lambda E-A|=\begin{vmatrix}\lambda-1&-2&-2\\-2&\lambda-1&-2\\-2&-2&\lambda-1\end{vmatrix}=(\lambda-5)(\lambda+2)^2=0.$$

所以特征多项式的根为 $\lambda_1=5,\lambda_2=\lambda_3=-1$(二重根). 所以线性变换的特征值分别为 $5,-1,-1$.

再求每个特征值对应的特征向量. 当 $\lambda_1=1$ 时，考虑齐次线性方程组 $(5E-A)X=0$. 得到线性方程组

$$\begin{cases}x_1+x_2=x_3,\\ \qquad x_2=x_3.\end{cases}$$

令 $x_3=1$，得到一个基础解系 $\eta=(1,1,1)$，它对应于 A 的特征向量为 $\varepsilon_1+\varepsilon_2+\varepsilon_3$，它就是 V_{λ_1} 的一组基.

当 $\lambda_2=\lambda_3=-1$ 时，可得线性方程租

$$x_1=-x_2-x_3.$$

同理可得基础解系 $\eta_1=(-1,1,0),\eta=(-1,0,1)$. 该基础解系对应的两个基向量为 $-\varepsilon_1+\varepsilon_2,-\varepsilon_1+\varepsilon_3$. 这就是 $V_{\lambda_2}=V_{\lambda_3}$ 的一组基. □

7.5.3 特征值的基本性质

定理 7.5.1 相似的矩阵有相同的特征多项式.

证明 设 $B=T^{-1}AT$ 则

$$\begin{aligned}|\lambda E-B|&=|\lambda E-T^{-1}AT|=|T^{-1}(\lambda E-A)T|\\&=|T^{-1}||\lambda E-A||T|=|\lambda E-A||T^{-1}T|=|\lambda E-A||E|=|\lambda E-A|.\end{aligned}$$

此命题的逆命题一般不成立，即两个方阵有相同的特征多项式，但它们未必相似. 例如

$$A=\begin{pmatrix}1&1\\0&1\end{pmatrix},\quad B=\begin{pmatrix}1&0\\0&1\end{pmatrix}.$$

这两个矩阵有相同的特征多项式，但是矩阵 B 的相似矩阵只有单位矩阵.

由于线性变换 $\hat{A}$ 在不同基下的矩阵是相似的，根据上述命题，它们的特征多项式相同. 因而，我们把 $\hat{A}$ 在任意一组基下矩阵的特征多项式称为 $\hat{A}$ 的特征多项式. 这个命题也从理论上指明：在用前面讲的办法计算线性变换的特征值和特征向量时，不会因为所选择的基不相同而得到不同的结果. 即它是线性变换的一个不变量.

定理 7.5.2 设 A 是域 $\mathbb{F}$ 上的 n 阶方阵，则

$$f(\lambda)=|\lambda E-A|=\lambda^n-\operatorname{tr}(A)\lambda^{n-1}+\cdots+(-1)^n|A|.$$

其中 $\operatorname{tr}(A)$ 为 A 的主对角线元素之和，即为 A 的迹.

证明 已知 $f(\lambda)$ 为 λ 的多项式，设

$$f(\lambda)=a_0\lambda^m+a_1\lambda^{m-1}+\cdots+a_m.$$

显然 $a_m=f(0)=|-A|=(-1)^n|A|$. 故只需证

$$m=n,a_0=1,a_1=-\operatorname{tr}(A).$$

对 n 做数学归纳法. 当 $n=1$ 时，$f(\lambda)=|\lambda-a_{11}|=\lambda-a_{11}$，显然成立. 下面设对 $\mathbb{F}$ 上 $n-1$ 阶方阵命题成立. 当 $A=(a_{ij})\in M_n(\mathbb{F})$ 时，我们对 $f(\lambda)$ 求形式微商. 一方面

$$f'(\lambda)=a_0m\lambda^{m-1}+a_1(m_-1)\lambda^{m-2}+\cdots+a_{m-1}.$$

另一方面，对行列式作形式微商，可得

$$\frac{\mathrm{d}}{\mathrm{d}\lambda}f(\lambda)=\frac{\mathrm{d}}{\mathrm{d}\lambda}|\lambda E-A|=\sum_{i=1}^n|(\lambda E-A)_i|,$$

其中 $(\lambda E-A)_i$ 表示把 $\lambda E-A$ 的第 i 行对 λ 求形式微商. 我们把 $|(\lambda E-A)_i|$ 对第 i 行展开，再利用归纳假设，有

$$\begin{aligned}
|(\lambda E-A)_i|&=\begin{array}{c}\begin{array}{cccccc} & & & i & & \end{array}\\ \begin{vmatrix}\lambda-a_{11} & -a_{12} & \cdots & \vdots & \cdots & -a_{1n}\\ 0 & 0 & \cdots & 1 & \cdots & 0\\ -a_{n1} & -a_{n2} & \cdots & \vdots & \cdots & \lambda-a_{nn}\end{vmatrix}\ i\end{array}\\
&=|\lambda E_{n-1}=A\tbinom{i}{i}|\\
&=\lambda^{n-1}-\operatorname{tr}(A\tbinom{i}{i})\lambda^{n-2}+\cdots+\lambda^{n-1}-(\operatorname{tr}(A)-a_{ii})\lambda^{n-2}+\cdots.
\end{aligned}$$

带回原式可得

$$\begin{aligned}
\frac{\mathrm{d}f(\lambda)}{\mathrm{d}\lambda}&=\sum_{i=1}^n[\lambda^{n-1}-(\operatorname{tr}(A)-a_ii)\lambda^{n-2}+\cdots]\\
&=n\lambda^{n-1}-(n-1)\operatorname{tr}(A)\lambda^{n-2}+\cdots\\
&=f'(\lambda)=a_0m\lambda^{m-1}+a_1(m-1)\lambda^{m-2}+\cdots.
\end{aligned}$$

这里由蓝以中第一册第一章命题 2.2 的推论 2，我们有 $m=n$ 且

$$a_0m=n,-(n-1)\operatorname{tr}(A)=a_1(m-1).$$

于是 $a_0=1,a_1=-\operatorname{tr}(A)$. □

　　我们知道，两个 n 阶方阵 A,B 一般不可交换，但我们可以证明它们的特征多项式却是一样的.

命题 7.5.7 设 A 是 $\mathbb{F}$ 上 $n\times m$ 矩阵，B 是 $\mathbb{F}$ 上 $m\times n$ 矩阵，则

$$\lambda^m|\lambda E_n-AB|=\lambda^n|\lambda E_m-BA|.$$

证明 利用分块矩阵，我们有

$$\begin{pmatrix}E_m & 0\\ -A & E_n\end{pmatrix}\begin{pmatrix}\lambda E_m & B\\ \lambda A & \lambda E_n\end{pmatrix}=\begin{pmatrix}\lambda E_m & B\\ 0 & \lambda E_n-AB\end{pmatrix}.$$

两边同时取行列式，则有

$$\begin{vmatrix}\lambda E_m & B\\ \lambda A & \lambda E_n\end{vmatrix}=\begin{vmatrix}\lambda E_m & B\\ 0 & \lambda E_n-AB\end{vmatrix}=\lambda^m|\lambda E_n-AB|.$$

另外，我们又有

$$\begin{pmatrix}\lambda E_m & B\\ \lambda A & \lambda E\end{pmatrix}\begin{pmatrix}E_m & 0\\ -A & E_n\end{pmatrix}=\begin{pmatrix}\lambda E_m-BA & B\\ 0 & \lambda E_n\end{pmatrix}.$$

两边再同时取行列式，有

$$\begin{vmatrix}\lambda E_m & B\\ \lambda A & \lambda E_n\end{vmatrix}=\begin{vmatrix}\lambda E_m-BA & B\\ 0 & \lambda E_n\end{vmatrix}=\lambda^n|\lambda E_m-BA|.$$

与前面的式子比较，结论得证. □

由代数学基本定理，一个 n 次多项式在复数域内恰有 n 个根 (重根按重数算). 设 $f(\lambda)$ 在复数域内的 n 个根是 $\lambda_1,\cdots,\lambda_n$，则有

$$\lambda_1+\lambda_2+\cdots+\lambda_n=\mathrm{tr}(A),$$
$$\lambda_1\lambda_2\cdots\lambda_n=|A|.$$

上述性质在讨论一些问题有用，同时借此我们可以得到可逆方阵 A 的伴随矩阵的特征值为

$$\frac{\lambda_1\lambda_2\cdots\lambda_n}{\lambda_i},i=1,2,\cdots,n.$$

7.5.4 具有对角矩阵的线性变换

定义 7.5.4 设 $\hat{A}$ 是 n 维线性空间 V 内的一个线性变换. 如果 V 内存在一组基 $\eta_1,\eta_2,\cdots,\eta_n$ 使得 $\hat{A}$ 在这组基下的矩阵成对角形，则称 $\hat{A}$ 的矩阵**可对角化 (diagonalize)**.

定理 7.5.3 $\mathbb{F}$ 上 n 维线性空间 V 内的一个线性变换 $\hat{A}$ 的矩阵可对角化的充分必要条件是 $\hat{A}$ 有 n 个线性无关的特征向量.

证明 必要性：若 $\hat{A}$ 在基 η_i 下的矩阵成对角形，即

$$(\hat{A}\eta_1,\cdots,\hat{A}\eta_n)=(\eta_1,\cdots,\eta_n)\begin{pmatrix}\lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n\end{pmatrix},$$

则有

$$\hat{A}\eta_1=\lambda_1\eta_1,\cdots,\hat{A}\eta_n=\lambda_n\eta_n.$$

充分性：如果 $\hat{A}$ 有 n 个线性无关的特征向量 $\eta_1,\cdots,\eta_n$，把他们取作 V 的一组基，$\hat{A}$ 在这组基下的矩阵成对角形.□

什么样的线性变换有 n 个线性无关的特征向量呢？

命题 7.5.8 线性变换 $\hat{A}$ 的属于不同特征值的特征向量线性无关.

证明 取 $\hat{A}$ 的 k 个不同特征值 $\lambda_1, \lambda_2, \cdots, \lambda_k$, 它们分别对应于特征向量 $\xi_1, \xi_2, \cdots, \xi_k$. 对 k 作数学归纳法.

当 $k=1$ 时，$\xi_1 \neq 0$，显然线性无关. 设命题在 $k-1$ 个不同特征值的情况下已经成立，证明 个不同特征值的情况下也成立.

考虑等式

$$l_1\xi_1 + \cdots + l_k\xi_k = 0.$$

两边作用 $\hat{A}$ 容易得到

$$l_1\hat{A}\xi_1 + \cdots + l_k + \hat{A}\xi_k = 0$$

由于 $\hat{A}\xi_i = \lambda_i\xi_i$，所以

$$l_1\lambda_1\xi_1 + \cdots + l_k\lambda_k\xi_k = 0.$$

于是有

$$\lambda_1[l_1\xi_1 + \cdots + l_k\xi_k] - l_1\lambda_1\xi_1 + \cdots + l_k\lambda_k\xi_k = l_2(\lambda_1 - \lambda_2)\xi_2 + \cdots + l_k(\lambda_1 - \lambda_k)\xi_k = 0.$$

按照归纳假设，$\xi_2, \cdots, \xi_k$ 线性无关，所以

$$l_2(\lambda_1 - \lambda_2) = \cdots = l_k(\lambda_1 - \lambda_k) = 0.$$

因为 k 个特征值互补相同，故

$$l_2 = \cdots = l_k = 0.$$

代回第一个式子，由于 $\xi \neq 0$ 故 $l_1 = 0$. 这就证明 $\xi_1, \cdots, \xi_k$ 是线性无关的. □

将上述命题换个说法，立马有以下推论.

推论 7.5.1 如果 n 维线性空间 V 内的线性变换 $\hat{A}$ 有 n 个不同的特征值，则它的矩阵可对角化.

不过，如果一个线性把变换没有 n 个不同的特征值时，它的矩阵也有可能可以对角化.

命题 7.5.9 设 $\hat{A}$ 是 $\mathbb{F}$ 上的 n 维线性空间 V 上的线性变换,$\lambda_1, \cdots, \lambda_k, k \leqslant n$ 是 $\mathbb{F}$ 内 k 个不同的纯量,令 $V_{\lambda_i} = \{\hat{A}\alpha = \lambda_i\alpha\}$, 则子空间的和 $\sum\limits_{i=1}^{k} V_{\lambda_i}$ 是直和.

证明 试图证明 0 向量表示法唯一. 设

$$0 = \alpha_1 + \cdots + \alpha_k, (\alpha_i \in V_{\lambda_i}).$$

由于 $\alpha_i \in V_{\lambda_i}$，当 $\alpha_i \neq 0$ 时，它是属于特征值 λ_i 的特征向量. 上式表示 $\alpha_1, \cdots, \alpha_k$ 中不为 0 的向量线性相关，而属于不同特征子空间的向量线性无关，故矛盾. 所以 $\alpha_1 = \cdots = \alpha_n = 0$. 即表示法唯一. □

接下来 $\hat{A}$ 矩阵可对角化的充分必要条件.

定理 7.5.4 设 $\hat{A}$ 是 $\mathbb{F}$ 上 n 维线性空间 V 的线性变换,$\lambda_1, \lambda_2, \cdots, \lambda_k$ 是 $\hat{A}$ 全部互不相同的特征值. 则 $\hat{A}$ 的矩阵可对角化的充分必要条件是

$$V = \bigoplus_{i=1}^{k} V_{\lambda_i}.$$

证明 必要性：设

$$M = V_{\lambda_1} \oplus \cdots \oplus V_{\lambda_k} \subset V.$$

如果 $\hat{A}$ 在基 $\varepsilon_1, \cdots, \varepsilon_n$ 下的矩阵成对角形，则每个 ε_i 均为 $\hat{A}$ 的特征向量，必属于某个特征子空间 $V\lambda_j$，从而属于 M. 故对任意 $\alpha \in V$，有

$$\alpha = a_1\varepsilon_1 + \cdots + a_n\varepsilon_n \in M.$$

即 $V \subset M$，所以 $V = M = V_{\lambda_1} \oplus \cdots \oplus V_{\lambda_k}$.

充分性：设

$$V = V_{\lambda_1} \oplus \cdots \oplus V_{\lambda_k},$$

在每个 V_{λ_i} 中取一组基，合并得到一个 V 的一个向量租. 该向量组是 V 的一组基，这直接就说明了 $\hat{A}$ 在该组基下的矩阵为对角矩阵. □

接下来给出线性变换矩阵对角化的具体方法.

在 V 中取一组基 $\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n$，设

$$(\hat{A}\varepsilon_1, \hat{A}\varepsilon_2, \cdots, \hat{A}\varepsilon_n) = (\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n)A.$$

对 $\hat{A}$ 的每个特征值 λ_i 求解齐次线性方程组

$$(\lambda_i E - A)X = 0$$

得到一个基础解系 $X_{i1}, X_{i2}, \cdots, X_{it_i}$，这里 $t_i = n - \mathrm{rank}(\lambda_i E - A)$. 可以得到在基 ε_i 下以 X_{jt_j} 为坐标的一组向量

$$\begin{cases} \eta_{i1} = (\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n)X_{i1}, \\ \eta_{i2} = (\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n)X_{i2}, \\ \cdots\cdots\cdots\cdots\cdots\cdots \\ \eta_{it_i} = (\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n)X_{it_i}. \end{cases}.$$

这里 $\eta_{i1}, \cdots, \eta_{it_i}$ 即为 $V\lambda_i$ 的一组基，把它们合并得 V 的一组基，$\hat{A}$ 在该组基下的矩阵为对角矩阵.

$$D = \begin{pmatrix} \lambda_1 &&&&&&&& \\ & \ddots &&&&&&& \\ && \lambda_1 &&&&&& \\ &&& \lambda_2 &&&&& \\ &&&& \ddots &&&& \\ &&&&& \lambda_2 &&& \\ &&&&&& \ddots && \\ &&&&&&& \lambda_k & \\ &&&&&&&& \ddots & \\ &&&&&&&&& \lambda_k \end{pmatrix},$$

其中分别有 $t_1, t_2, \cdots, t_k$ 个 $\lambda_1, \lambda_2, \cdots, \lambda_k$. 从 ε 到 η 的过渡矩阵的列向量为向量 η_{ij} 在 ε_i 下的坐标，即

$$T = (X_{11}X_{12}\cdots X_{1t_1}X_{21}X_{22}\cdots X_{2t_2}\cdots X_{k1}X_{k2}\cdots X_{kt_k}).$$

此时有 $T^{-1}AT=D$. 现在 A 的特征多项式

$$f(\lambda)=|\lambda E-D|=(\lambda-\lambda_1)^{t_1}(\lambda-\lambda_2)^{t_2}\cdots(\lambda-\lambda_k)^{t_k}.$$

我们把 $f(\lambda)$ 每个根的重数 t_i 称为**代数重数 (algebraic multiplicity)**，这时它等于 $p_i=\dim V_{\lambda_i}$. 我们称 p_i 为**几何重数 (geometry multiplicity)** 这样，我们又可以得到一个充要条件即线性变换 $\hat{A}$ 可对角化的充要条件是它的每个特征值的代数重数和几何重数都相等.

同时，还注意到，上述求线性变换的过程抽象出来单对矩阵使用. 这时我们可以定义矩阵的对角化.

定义 7.5.5 设 n 阶方阵 $A\in M_n(\mathbb{F})$，如果存在可逆矩阵 T 和对角矩阵 D 满足

$$D=T^{-1}AT$$

则称矩阵 A **可对角化.**

这样，结合前面线性变换的矩阵对角化过程，我们就得到了矩阵求对角化的过程.

7.6　对偶空间

7.6.1　线性函数

注意到线性映射可以定义其加法和数乘，于是我们可以做进一步的研究.

定义 7.6.1 (线性函数) V 是 $\mathbb{F}$ 上的线性空间，若线性映射 $f:V\to F$，则称 f 为 V 上的**线性函数 (linear function)**(有时称为线性型或者线性泛函，线性泛函多用于无穷维函数).

设线性函数 $f:V\to\mathbb{F}$，在 V 中取一组基底 $\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$，把线性函数 f 作用到 $x=a_1\varepsilon_1+a_2\varepsilon_2+\cdots+a_n\varepsilon_n$ 上得到

$$f(x)=a_1\beta_1+a_2\beta_2+\cdots+a_n\beta_n,$$

其中 $f(\varepsilon_i)=\beta_i$ 是纯量. 在确定 f 后，其 β_i 只与选取的基有关. 反过来，如果 $\beta_i\in\mathbb{F},i=1,\cdot,n$ 和基确定后，只有唯一的一个映射 f 满足 $f(\varepsilon_i)=\beta_i$.

注意到线性函数的定义是基的选择无关，所以我们应该知道线性函数在不同基下作用效果的一个转换. 设

$$\mathrm{span}\{\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n\}=V=\mathrm{span}\{\varepsilon_1',\varepsilon_2',\cdots,\varepsilon_n'\}.$$

于是有转换的坐标公式

$$\varepsilon_j'=x_{j1}\varepsilon_1+x_{j2}\varepsilon_2+\cdots+x_{jn}\varepsilon_n,\quad j=1,2,\cdots,n.$$

将 f 作用在不同基表法下的同一个向量 v，就可以得到

$$a_1\beta_1+\cdots+a_n\beta_n=f(v)=a_1'\beta_1'+\cdots+a_n'\beta_n',$$

其中

$$\beta_j'=f(e_j')=f(x_{j1}\varepsilon_1+x_{j2}\varepsilon_2+\cdots+x_{jn}\varepsilon_n)=x_{j1}\beta_1+\cdots+x_{jn}\beta_n.$$

不难看出，在基变换时，基向量的系数和线性函数的系数按照相同的坐标系数变换，也就是一致的，或者说同时的.

7.6.2 对偶空间与对偶基底

既然谈到空间，我们需要知道，空间里的元素是什么，空间中元素的加法和数乘运算是如何定义的？于是我们定义线性函数之间的加法和数乘.

定义 7.6.2 设 f, g 是线性空间 V 上的两个线性函数，x 是 V 中的元素，$\alpha, \beta \in \mathbb{F}$，定义

$$(\alpha f + \beta g)(x) = \alpha f(x) + \beta g(x).$$

于是我们可以定义对偶空间.

定义 7.6.3 (对偶空间) 对于刚才的运算,所有线性空间 V 上的线性函数组成一个**对偶**于 V 的**对偶空间** **(dual space)**$V^* = \mathcal{L}(V, \mathbb{F})$.

在同时处理 V 和 V^* 的时候，为了区分，我们称 V 上的向量为**反变向量 (contravariant vector)**，称 V^* 上的向量为**余向量 (cevector)** 或者**共变向量 (covariant)**. 在未来学习更一般的张量理论的时候，我们会知道余向量与 $(1,0)$ 型张量 (1 阶反变张量) 有着联系，而共边向量与 $(0,1)$ 型张量 (1 阶共变张量) 有着联系. “共”，“反” 二字经常在数学中出现，关于它们的意义都大致相同，其中最一般的是共变函子 (covariant functor) 和反变函子 (contravariant functor).

在空间 V 给定的基底 $(\varepsilon_1, \cdots, \varepsilon_n)$ 下 (这里采用圆括号说明这组基向量是有顺序的)，可以构造双射

$$\Phi : \mathcal{L}(V, \mathbb{F}) \to \mathbb{F}, \quad f \mapsto (f(\varepsilon_1), f_(\varepsilon_2), \cdots, f(\varepsilon_n)).$$

并且定义

$$f + g \mapsto ((f+g)(\varepsilon_1), (f+g)(\varepsilon_2), \cdots, (f+g)(\varepsilon_n)), \quad \lambda f \mapsto (\lambda f(\varepsilon_1), \lambda f_(\varepsilon_2), \cdots, \lambda f(\varepsilon_n)).$$

这样，Φ 就是 V^* 和 $\mathbb{F}$ 之间的同构，特别地，$\dim V^* = \dim V \cdot \dim \mathbb{F} = \dim \mathbb{F}^n = n$.

接下来我们利用 Kronecker 符号构造一个比较特殊的函数. 定义空间 V 上的线性函数

$$\varepsilon^i \coloneqq \varepsilon^i(\varepsilon_j) = \delta_{ij}, \quad j = 1, 2 \cdot, n.$$

于是我们就得到了一个 V^* 上的元素. 同时注意到对于 $\lambda_i \in \mathbb{F}$，由

$$(\lambda_1 \varepsilon^1 + \cdots + \lambda_n \varepsilon^n)(\varepsilon_j) = \lambda_j = 0$$

可知 $\varepsilon^1, \cdots, \varepsilon^n$ 是 $V*$ 上的线性无关元素. 同时不难看出,$\varepsilon^1, \cdots, \varepsilon^n$ 其实分别对应于 $\mathbb{F}^n$ 上的单位向量组 $(0, \cdots, 0, 1, 0, \cdots, 0)$. 这就证明了以下定理.

定理 7.6.1 设 V 是 $\mathbb{F}$ 上的 n 维线性空间. 则对偶空间 $V*$ 也是 n 维线性空间. 如果 $(\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n)$ 是一个基底，并且有线性函数 $\varepsilon^1, \cdots, \varepsilon^n$ 使得

$$\varepsilon^i(\varepsilon_j) = \delta_{ij} = \begin{cases} 1, & i = j \\ 0, & i \neq j \end{cases}.$$

则 $(\varepsilon^1, \cdots, \varepsilon^n)$ 是 V^* 上的一个基底.

基于上面这个定理，我们给出如下定义

定义 7.6.4 (对偶基底) 按照上述定理给出的 $(\varepsilon^1, \varepsilon^2, \cdots, \varepsilon^n)$ 是唯一相对于 $(\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n)$ 的**对偶基底 (dual basis)**(或者称**相互的对偶基底**).

我们使用“对偶”一词，其实已经暗含空间 V 和 V^* 的双边对称关系. 于是可以暂时约定用 $(f,x), f\in V^*, x\in V$ 来代替 $f(x)$，于是我们又得到了一个映射 $V^*\times V\to\mathbb{F}$，并且有

$$\alpha,\beta\in\mathbb{F}, g\in V^*, y\in V,\quad (\alpha f+\beta g, x)=\alpha(f,x)+\beta(g,x),\quad (f,\alpha x+\beta y)=\alpha(f,x)+\beta(f,y).$$

称这种对 f,x 都有线性关系的函数为**双线性型 (bilinear form)**. 在下一章我们会再作讨论.

选取线性空间 V 上的一组基，$(\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n)$，并取其 V^* 上的对偶基 $(\varepsilon^1,\varepsilon^2,\cdots,\varepsilon^n)$，可以用 V 中的基将 V^* 上的元素 f 唯一的表示出来

$$f=f(\varepsilon_1)\varepsilon^1+f(\varepsilon_2)\varepsilon^2+\cdots+f(\varepsilon_n)\varepsilon^n.$$

例 7.6.1 设 $V=P_n=\mathrm{span}\{1,t,\cdots,t^{n-1}\}$ 是次数 $\leqslant n<1$ 的 n 维多项式空间. 映射 $f_\lambda:\varphi\to\varphi(\lambda)$. 在每个点 $\lambda\in\mathbb{R}$, 多项式 $\varphi(t)=\varphi_0+\varphi_1 t+\cdots+\varphi_{n-1}t^{n-1}$ 都对应它在该点的值，f_λ 线性的，通过变动 λ 就可以得到 V^* 的基底，只需按照定义，直接设

$$\sum_{i=0}^{n-1}x_i f_{\lambda_i}=0.$$

展开后重新整理得

$$\sum_{k=0}^{n-1}\varphi_k\sum_{i=0}^{n-1}x_i\lambda_i^k=0.$$

于是我们只需找到一种能让 $x_0=\cdots=x_{n-1}=0$ 的 $\lambda_0,\cdots,\lambda_{n-1}$ 的取法即可. 考察线性方程组

$$\begin{pmatrix}1&1&\cdots&1\\ \lambda_0&\lambda_1&\cdots&\lambda_{n-1}\\ \vdots&\vdots&&\vdots\\ \lambda_0^{n-1}&\lambda_0^{n-1}&\cdots&\lambda_{n-1}^{n-1}\end{pmatrix}\begin{pmatrix}x_0\\ x_1\\ \vdots\\ x_{n-1}\end{pmatrix}=0.$$

只需让该方程组只有 0 解. 即系数矩阵的行列式不等于 0. 同时注意到该系数矩阵的行列式为 Vandermonde 行列式，于是只需行列式的值为

$$\prod_{0\leqslant i\leqslant j\leqslant n-1}(\lambda_j-\lambda_i)\neq 0.$$

于是，只需保证 $\lambda_i, i=0,1,\cdots,n-1$ 中任意两个 λ_i 的取值都不相等即可.

再或者，直接构造函数

$$\varepsilon^k:\varphi\to\frac{\varphi^{(k)}(0)}{k!}=\varphi_k,\quad k=0,1,\cdots,n-1$$

它是对偶于 $1,t,\cdots,t^{n-1}$ 的 V^* 中的基底.

例 7.6.2 设 V 是域 $\mathbb{F}$ 上 n 维线性空间，其 V^* 是 V 的对偶空间，再设 $f,g\in V*$，且 $\ker f=\ker g$. 证明：存在 $\lambda\in\mathbb{F}$ 使得 $g=\lambda f$.

证明 证法一：直接取 $\ker f=\ker g$ 的一组基 $\{\varepsilon_1,\cdots,\varepsilon_{n-1}\}$，将其扩充为 V 的基 $\{\varepsilon_1,\cdots,\varepsilon_{n-1},x_0\}$. 于是对于任意的 $x\in V$ 都有

$$f(x)=k_1f(\varepsilon_1)+\cdots k_{n-1}f(\varepsilon_{n-1})+k_nf(x_0),$$
$$g(x)=k_1g(\varepsilon_1)+\cdots k_{n-1}g(\varepsilon_{n-1})+k_ng(x_0).$$

显然 $g = \lambda f$.

证法二：如果 f, g 都为零函数，那么显然 $\lambda = 1$. 如果 f, g 都不为零函数，由于 $\ker f = \ker g$，于是存在 $x_0 \in V$ 使得 $f(x_0), g(x_0)$ 都不为 0. 不妨设 $c = \dfrac{f(x_0)}{g(x_0)}$. 再设 $\{\varepsilon_1, \cdots, \varepsilon_{n-1}\}$ 为 $\ker f = \ker g$ 上的一组基，将其扩充为 V 上的一组基 $\{\varepsilon_1, \cdots, \varepsilon_{n-1}, x_0\}$.

于是对于任意的 $x \notin \ker f = \ker g$，用反证法可以证明，x 关于基的线性表示中，必定含有 x_0 的项. 最后

$$(f - cg)(x) = kf(x_0) - kcf(x_0) = 0$$

故 $x \in \ker(f - cg)$，所以 $\ker(f - cg) = V$，于是 c 就是要求的 λ. □

7.6.3 自反性

接下来考虑 V^* 的对偶空间即 V 的**双重对偶空间 (double dual space)**$V^{**} = (V^*)^*$. 在 V 是有限维空间的情况下，我们知道 $V \cong V^*$. 同理，可以猜测 V^{**} 与 V^* 也有这样的同构. 也就是说，V^{**} 是 V^* 的对偶空间，即 V^* 上的线性函数组成的空间. 不过，似乎用 V 中的元素直接描述 V^{**} 中的元素是困难的. 其实 V 于 V^{**} 有着**自然同构 (natural isomorphism)**，即该同构不依赖于 V 上基的选取.

定理 7.6.2 存在自然同构 $\mathcal{N} : V \to V^{**}$，它由以下公式所决定

$$\mathcal{N}(x) = \mathcal{N}_x, \quad \mathcal{N}_x(f) = f(x),$$

其中 $x \in V, f \in V^*, \mathcal{N}_x \in V^{**}$.

证明 对于任意线性函数 $f : V \to \mathbb{F}$ 有

$$\mathcal{N}_{\alpha x + \beta y}(f) = f(\alpha x + \beta y) = \alpha f(x) + \beta f(y) = \alpha \mathcal{N}_x(f) + \beta \mathcal{N}_y(f) = (\alpha \mathcal{N}_x + \beta \mathcal{N}_y)(f),$$

所以 $\mathcal{N}(\alpha x + \beta y) = \alpha \mathcal{N}(x) + \beta \mathcal{N}(y)$. 所以 $\mathcal{N}$ 是一个线性映射.

接下来证明 ε 是双射. 在 V 中选取互相对偶的基底 $(\varepsilon_1, \cdots, \varepsilon_n)$ 和 $(\varepsilon^1, \varepsilon^2, \cdots, \varepsilon^n)$. 于是

$$\mathcal{N}_{\varepsilon_j}(\varepsilon^i) = \varepsilon^i(\varepsilon_j) = \delta_{ij}.$$

这说明

$$V^{**} = \operatorname{span}\{\mathcal{N}_{\varepsilon_1}, \mathcal{N}_{\varepsilon_2}, \cdots, \mathcal{N}_{\varepsilon_n}\}.$$

于是我们在 V^{**} 上找到了对偶于 $(\varepsilon^1, \varepsilon^2, \cdots, \varepsilon^n)$ 的基底. 于是任取 V^{**} 中的元素 $\mathcal{N}_{x_0}$，则直接取 V 中的元素 x_0 使得 $\mathcal{N}(x_0) = \mathcal{N}_{x_0}$.，故 $\mathcal{N}$ 为满射. 再对于任意 $x \in \ker\mathcal{N}$，则对于任意 $f \in V^*$ 都有 $f(x) = 0$ 这等价于 $x = 0 \in V$，所以 $\ker\mathcal{N} = \{0\}$. 故 $\mathcal{N}$ 为单射. □

定义 7.6.5 像上述一样，两个线性空间因存在着自然同构而显现出来的性质被称为**自反性 (reflexive)**.

类似于数学分析中，由单射 $\mathbb{N} \to \mathbb{Z}, n \mapsto (n, 0)$ 给出的自然数在整数集中的像等同于自然数然后有 $\mathbb{N} \subset \mathbb{Z}$ 一样，我们将 V^{**} 中的元素也于 V 中的元素等同起来，这样 V 和 V^{**} 的地位就完全平等，然后就可以把 V 看作是 V^* 上线性函数的空间. 于是可以构造函数 $x(f) = (f, x) = \mathcal{N}_x(f) = f(x)$. 这也说明了，对于任意 V^* 中的一组基底，必定有唯一确定的与之对偶的 V 中的基底.

7.6.4　利用对偶空间判断线性无关性

利用对偶空间 V^* 的概念可以来判断 V 中的元素是否线性相关. 其本质是要考察由线性函数 $f_i : V \to \mathbb{F}$ 作用在 V 中不同向量 x_i 而得到的矩阵的行列式.

为了证明最后的那个定理，我们先证明几个引理.

引理 7.6.1 如果 $x_1, x_2, \cdots, x_m$ 是 V 中线性相关的量，而 $f_1, f_2, \cdots, f_m$ 是 V 上任意 m 个线性函数，则

$$\begin{vmatrix} f_1(x_1) & f_1(x_2) & \cdots & f_1(x_m) \\ f_2(x_1) & f_2(x_2) & \cdots & f_2(x_m) \\ \vdots & \vdots & & \vdots \\ f_m(x_1) & f_m(x_2) & \cdots & f_m(x_m) \end{vmatrix} = 0.$$

证明 由于 $x_1, \cdots, x_m$ 线性相关，则其中至少有一个，不妨设是 x_m 是其余向量的线性组合. 设

$$x_m = a_1 x_1 + \cdots + a_{m-1} x_{m-1}.$$

再加上 f 的线性性质，就有

$$\begin{vmatrix} f_1(x_1) & f_1(x_2) & \cdots & f_1(x_m) \\ f_2(x_1) & f_2(x_2) & \cdots & f_2(x_m) \\ \vdots & \vdots & & \vdots \\ f_m(x_1) & f_m(x_2) & \cdots & f_m(x_m) \end{vmatrix} = \begin{vmatrix} f_1(x_1) & f_1(x_2) & \cdots & \sum\limits_{i=1}^{m-1} a_i f_1(x_i) \\ f_2(x_1) & f_2(x_2) & \cdots & \sum\limits_{i=1}^{m-1} a_i f_2(x_i) \\ \vdots & \vdots & & \vdots \\ f_m(x_1) & f_m(x_2) & \cdots & \sum\limits_{i=1}^{m-1} a_i f_m(x_i) \end{vmatrix}$$

用最后一列减去第一列乘以 a_1 得到

$$\begin{vmatrix} f_1(x_1) & f_1(x_2) & \cdots & \sum\limits_{i=2}^{m-1} a_i f_1(x_i) \\ f_2(x_1) & f_2(x_2) & \cdots & \sum\limits_{i=2}^{m-1} a_i f_2(x_i) \\ \vdots & \vdots & & \vdots \\ f_m(x_1) & f_m(x_2) & \cdots & \sum\limits_{i=2}^{m-1} a_i f_m(x_i) \end{vmatrix}.$$

再用最后一列减去第二列乘以 a_2，减去第三列乘以 a_3，不断反复，直到最后一列减去 $m-1$ 列乘以 a_{m-1}. 行列式就变成了

$$\begin{vmatrix} f_1(x_1) & f_1(x_2) & \cdots & 0 \\ f_2(x_1) & f_2(x_2) & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ f_m(x_1) & f_m(x_2) & \cdots & 0 \end{vmatrix} = 0.$$

□

引理 7.6.2 如果 $(f_1,f_2,\cdots,f_n)$ 是线性空间 V 的对偶空间 V^* 的一个基底，那么，向量 $x_1,x_2,\cdots,x_n\in V$ 线性无关当且仅当

$$\begin{vmatrix} f_1(x_1) & f_1(x_2) & \cdots & f_1(x_n) \\ f_2(x_1) & f_2(x_2) & \cdots & f_2(x_n) \\ \vdots & \vdots & & \vdots \\ f_m(x_1) & f_m(x_2) & \cdots & f_n(x_n) \end{vmatrix} \neq 0.$$

证明 先证充分性，如果上述行列式不等于 0 而 $x_1,\cdots,x_n$ 线性相关，则由**引理7.6.1**可知，上述行列式为 0 得矛盾. 再证必要性，如果 $x_1,\cdots,x_n$ 线性无关，则 $V=\mathrm{span}\{x_1,x_2,\cdots,x_n\}$. 再用 $(\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n)$ 表示 V 中对偶于 $(f_1,f_2,\cdots,f_n)$ 的基底. 于是

$$x_j=a_{1j}\varepsilon_1+a_{2j}\varepsilon_2+\cdots+a_{nj}\varepsilon_j,\quad j=1,2,\cdots,n.$$

立即得到 ε_i 到 x_i 的过渡矩阵

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & & a_{nn} \end{pmatrix}.$$

由于上述矩阵可逆，所以 $\det(a_{ij})\neq 0$. 又因为

$$f_i(x_j)=f_i(a_{1j}\varepsilon_1+a_{2j}\varepsilon_2+\cdots+a_{nj}\varepsilon_j)=a_{ij}f_i(\varepsilon_i)=a_{ij},$$

所以 $\det(f_i(x_j))\neq 0$. □

接下来终于要给出我们的重要定理.

定理 7.6.3 设 $(f_1,f_2,\cdots,f_n)$ 是对偶于 V 的空间 V^* 的一个基底. 那么向量组 $x_1,\cdots,x_k\in V$ 的秩等于矩阵

$$A=\begin{pmatrix} f_1(x_1) & f_1(x_2) & \cdots & f_1(x_k) \\ f_2(x_1) & f_2(x_2) & \cdots & f_2(x_k) \\ \vdots & \vdots & & \vdots \\ f_n(x_1) & f_n(x_2) & \cdots & f_n(x_k) \end{pmatrix}$$

的秩.

证明 我们只需要证明，$\det A$ 的最高阶非零子式的阶数刚好就是 $x_1,\cdots,x_k$ 的秩. 设 r 为 $x_1,\cdots,x_k$ 的秩. 对于任意 $m>r$,$x_{j1},\cdots,x_{jm}$ 必定线性相关，根据**引理7.6.1**可知大于 r 阶的子式 $\det(f_i(x_j))=0$.

剩下的只需证明，存在一个阶为 r 的子式非 0. 我们将 $f_1,\cdots,f_n$ 限制在子空间 $U=\mathrm{span}\{x_1,\cdots,x_k\}$ 上得到 $\bar f_1,\bar f_2,\cdots,\bar f_n$. 首先要证明的就是

$$\mathrm{span}\{\bar f_1,\cdots,\bar f_n\}=U^*, \tag{7.6.1}$$

其中 U^* 是对偶于 U 的子空间.

由于 $\bar f_1,\cdots,\bar f_n$ 都是 U 上的线性函数，所以 $\mathrm{span}\{\bar f_1,\cdots,\bar f_n\}\subset U^*$. 其次，再设 $\widetilde f$ 是 U^* 中的任意一个向量，$(\varepsilon_1,\cdots,\varepsilon_r)$ 是 U 中一个基底，而它的扩充 $(\varepsilon_1,\cdots,\varepsilon_r,\varepsilon_{r+1},\cdots,\varepsilon_n)$ 是 V 中的一个基底. 再构造线性函数 $f\in V^*$ 满足

$$f(\varepsilon_i)=\begin{cases}\widetilde f(\varepsilon_i), & i=1,2\cdots,r;\\ 0, & i=r+1,r+2,\cdots,n.\end{cases}$$

因为 $V^*=\mathrm{span}\{f_1,f_2,\cdots,f_n\}$，所以

$$f=f(\varepsilon_1)f_1+\cdots+f(\varepsilon_n)f_n. \tag{7.6.2}$$

又因为任取一个基向量 $\varepsilon_i\in U$，有

$$f(\varepsilon_i)=\widetilde{f}(\varepsilon_i),\quad i=1,2,\cdots,r$$

以及 $f|_U(\varepsilon_i)=\bar{f}(\varepsilon_i)$ 所以 $\bar{f}=\widetilde{f}$. 再把等式(7.6.2)中的所有函数都限制在 U 上，就有

$$\widetilde{f}=\bar{f}=\widetilde{f}(\varepsilon_1)\bar{f}_1+\cdots+\widetilde{f}(\varepsilon_r)\bar{f}_r.$$

故 $\widetilde{f}\in\mathrm{span}\{\bar{f}_1,\cdots,\bar{f}_n\}$. 从而证明了(7.6.1).

最后，在 $x_1,\cdots,x_k$ 中选择 r 个线性无关的向量，设为 $x_{j1},\cdots,x_{jr}$; 在 $\bar{f}_1,\cdots,\bar{f}_n$ 中也选择 r 个线性无关的向量，设为 $\bar{f}_{i1},\cdots,\bar{f}_{ir}$. 它们分别组成 U 和 U^* 的基底，根据**引理7.6.2**可知

$$\begin{vmatrix}\bar{f}_{i1}(x_{j1}) & \bar{f}_{i1}(x_{j2}) & \cdots & \bar{f}_{i1}(x_{jr})\\ \bar{f}_{i2}(x_{j1}) & \bar{f}_{i1}(x_{j2}) & \cdots & \bar{f}_{i2}(x_{jr})\\ \vdots & \vdots & & \vdots\\ \bar{f}_{ir}(x_{j1}) & \bar{f}_{ir}(x_{jr}) & \cdots & \bar{f}_{ir}(x_{jr})\end{vmatrix}\neq 0.$$

又因为 $\bar{f}=f|_U,U\subset V$，所以 $\bar{f}_i(x_j)=f_i(x_j)$. 于是我们就找到了 r 阶的非零子式. □

7.6.5　齐次线性方程组的几何解释

回忆一下，n 个未知量的齐次线性方程组有解时，可以将它的解用 $\mathbb{F}^n$ 上的向量表示出来. 所以再在 $\mathbb{F}^n$ 中可以由方程组的解生成解子空间. 接下来将采用更抽象的观点. 对于一个有 m 个方程，n 个未知数的线性方程组. 我们令 $x\in V$，再将每一个方程都看作是一个 V 上的线性函数. 所以一个线性方程组就可以写为

$$f_1(x)=0,f_2(x)=0,\cdots,f_n(x)=0. \tag{7.6.3}$$

注意到 $\ker f_i$ 是 V 的超平面，所以上述齐次线性方程其实也可以等价描述为 m 个 V 的超平面的交. 因为 $f_1,\cdots,f_m\in V^*$，所以我们需要在 V 中取一个基底.

定理 7.6.4 如果向量组 $f_1,\cdots,f_m$ 的秩为 r，则齐次线性方程组(7.6.3)的解的子空间 $U\subset V$ 的维数为 $\dim V-r$. 并且，任意子空间 $U\subset V$ 必定某个线性方程组(7.6.3)的解子空间.

证明 第一个结论之前已经证明过，但现在的推断看起来会更自然. 不失一般性，假设 $f_1,\cdots,f_r$ 是线性无关的. 所以若 $x\in V$ 满足 $f_1(x)=\cdots=f_r(x)=0$，那么对于任意 $f_i(x),i=r+1,\cdots,m$ 有

$$f_i(x)=k_1f_1(x)+\cdots+k_rf_r(x)=0,\quad ,k_i\in\mathbb{F}.$$

于是方程组(7.6.3)等价于

$$f_1(x)=0,\cdots,f_r(x)=0. \tag{7.6.4}$$

再将 $f_1,\cdots,f_r$ 扩充为空间 V^* 的基底 $\varepsilon^1,\cdots,\varepsilon^n$, 且该基底满足在 $i=1,\cdots,r$ 时有 $e^i=f_i$. 再设 $(\varepsilon_1,\cdots,\varepsilon_n)$ 是 V 中对偶于 $(\varepsilon^1,\cdots,\varepsilon^n)$ 的基底. 由对偶基底的定义，对于 $x=a_1\varepsilon_1+\cdots+a_n\varepsilon_n$ 必然有 $a_1=\cdots=a_r=0$. 于是方程组(7.6.4)的

解空间 U 中的向量 x 由向量 $\varepsilon_{r+1},\cdots,\varepsilon_n$ 的线性组合构成. 即 $U=\operatorname{span}\{\varepsilon_{r+1},\cdots,\varepsilon_n\}$. 再加上 $\varepsilon_{r+1},\cdots,\varepsilon_n$ 线性无关，所以 $\dim U=n-r=\dim V-r$.

对于第二个结论. 对于 n 维线性空间 V，设 $(\varepsilon_1,\cdots,\varepsilon_s)$ 是子空间 $U\subset V$ 的一个基底，同时是 V 中基底 $(\varepsilon_1,\cdots,\varepsilon_n)$ 的一部分. 则向量 $x=a_1\varepsilon_1+\cdots a_n\varepsilon_n$ 属于 U 当且仅当 $a_{s+1}=\cdots=a_n=0$. 选取 V^* 中与 $(\varepsilon_1,\cdots,\varepsilon_n)$ 对偶的一个基底 $(f_1,\cdots,f_n)$，那么 $a_i=f_i(x)$. 于是 $x\in U$ 等价于 $f_{s+1}(x)=0,\cdots,f_n(x)=0$. 于是我们就构造出了可以让 U 作为解子空间的方程组. $\square$

第八章　双线性函数与二次型

8.1　多重线性映射

前面我们已经探索了 V 上的与向量，现在介绍更一般的情况.

定义 8.1.1 (多重线性映射) 设 $V_1, V_2, \cdots, V_p; U$ 是 $\mathbb{F}$ 上的线性空间. 映射 $f: V_1 \times V_2 \times V_p \to U$ 是**多重线性映射 (multilinear map)**，如果对于任意指数 $i = 1, 2 \cdots, p$，以及任意固定的向量 $a_j \in V_j, 1 \leqslant j \leqslant p, i \neq j$，

$$f_i : v \mapsto f(a_1, \cdots, a_{i-1}, v, a_{i+1}, \cdots, a_p)$$

都是线性映射.

注 8.1.1 多重线性映射如果在 p 确定时，也可称为 p-线性映射

其实，就是 $f(v_1, \cdots, v_n)$ 关于括号中的每一个分量，都是线性的，即

$$f_i(kx + ly) = kf_i(x) + lf_i(y), \quad x, y \in V_i; k, l \in \mathbb{F}.$$

这种映射本质上就是将多个线性空间中的向量映到另一个线性空间. 所以说它比线性映射更一般.

不难验证，类似线性映射，两个 p-线性映射的线性组合 $\alpha f + \beta g$ 仍是 p-线性映射. 于是我们可以把所有 $V_1 \times \cdots \times V_p \to U$ 的 p-线性映射组成的集合 $\hom(V_1, \cdots, V_p; U)$ 作为研究的对象.

接下来看一个简单的例子.

例 8.1.1 令 $V_1 \cdots = V_p = U = \mathbb{F}$. 定义

$$f(v_1, \cdots, v_p) = v_1 \cdots v_p.$$

这显然是一个多重线性型.

类似于线性函数，我们也可以给出一个更一般的定义.

定义 8.1.2 (多重线性型) 任意线性空间 $V_1, \cdots, V_p$，称 $V_1 \times \cdots \times V_p$ 到域 $\mathbb{F}$ 的多重线性映射为**多重线性型 (multilinear form)** 或者**多重线性函数 (multilinear function)**.

在这里顺便引出简单的**张量 (tensor)** 概念. 设 $l^i : v_i \mapsto l^i(v_i), i = 1, 2, \cdots, p$ 是 V_i 上的线性函数，那么用关系式

$$f(v_1, \cdots, v_p) = l^1(v_1) \cdots l^p(v_p)$$

定义的函数 f 就是 $V_1 \times \cdots \times V_p$ 上的多重线性型. 称它为线性型 $l^1, \cdots, l^p$ 的**张量积 (tensor product)**，且可以表示为 $f = l^1 \otimes l^2 \otimes \cdots \otimes l^p$ 或简记为 $l^1 l^2 \cdots l^p$. 还需注意的是，正如笛卡尔积是有序的，这里张量积也是有序的，不能随意交换序号.

可以证明，任意多重线性型都可以表示为线性函数的张量积，不过目前我们不需要这个事实. 为了方便，在 $V_1 = \cdots = V_p = V$ 时，记 $V^p = \underbrace{V \times \cdots \times V}_{p\text{个}V}$. 此时可以记

$$\hom_p(V, \mathbb{F}) \coloneqq \hom(V, \cdots, V; \mathbb{F}).$$

在 $V^p \times V^{*q}$ 上的多重线性型随后被说成是 $p+q$ 阶的 (p,q) 型张量. 正如上一节所说，V^{**} 中的向量可以和 V 中的向量等同起来，于是 V 中的向量可以看作是 V^* 上的线性函数，即 $(0,1)$ 型张量可以看作是 V 中的向量.

V^p 上的多重线性型 $f((p,0)$ 型张量) 称为**对称的**. 如果对任意 $v_1, \cdots, v_p \in V$，以及任意置换 $\pi \in S_p$，都有

$$f(v_{\pi(1)}, \cdots, v_{\pi(p)}) = f(v_1, \cdots, v_p).$$

设 ε_π 是置换 π 的奇偶性，如果有

$$f(v_{\pi(1)}, \cdots, v_{\pi(p)}) = \varepsilon_{pi} f(v_1, \cdots, v_p),$$

则称 f 是**斜对称的 (skew-symmetric)** 或者**反对称的**再或者**交错的**. 再之前已经给出了一个很好的例子，就是在 $\dim V = p$ 时，将矩阵的行列式看作是行或列的函数. 在上面的叙述中，先为大家介绍了更一般的定义，具体的详解介绍还是要放到后面.

8.2 双线性型

现在，将上面的定义再做一个限制，考虑 $V_1 = V_2 = V$ 的情形.

定义 8.2.1 (双线性型) 对于任意 $u, v, w \in V; \alpha, \beta \in \mathbb{F}$ 有

$$\begin{aligned} f(\alpha u + \beta v, w) &= \alpha f(u, w) + \beta f(v, w), \\ f(w, \alpha u + \beta v) &= \alpha f(w, u) + \beta f(w, v). \end{aligned}$$

则称 f 为 V 上的一个**双线性型 (bilinear form)**.

在 V 中选择一个基底 $\varepsilon_1, \cdots, \varepsilon_n$，对于 $u, v \in V$，有坐标表示

$$u = x_1\varepsilon_1 + \cdots + x_n\varepsilon_n, \quad v = y_1\varepsilon_1 + \cdots + y_n\varepsilon_n.$$

根据定义，将上述式子代入得

$$f(u,v) = f\left(\sum_{i=1}^n x_i\varepsilon_i, \sum_{j=1}^n y_i\varepsilon_i.\right) = \sum_{i=1}^n x_i f\left(\varepsilon_i, \sum_{j=1}^n y_j\varepsilon_j\right) = \sum_{i=1}^n x_i \sum_{j=1}^n y_j f(\varepsilon_i, \varepsilon_j) = \sum_{i,j=1}^n f(\varepsilon_i, \varepsilon_j) x_i y_j.$$

于是我们就用 n^2 个纯量 $f(\varepsilon_i, \varepsilon_j)$ 表示出了一个双线性型. 我们以后可能会记 $f_{ij} = f(\varepsilon_i, \varepsilon_j)$.

令矩阵 $F = (f_{ij})$，该矩阵称为 V 上双线性型 f 在基 $(\varepsilon_1, \cdots, \varepsilon_n)$ 下的矩阵. 利用矩阵的形式乘法，令 u 在基 ε_i 下的坐标为 $X = [x_1, \cdots, x_n]$，v 在基 ε_i 下的坐标为 $Y = [y_1, \cdots, y_i]$. 于是 $f(u,v)$ 可以表示为

$$f(u,v) = X^t \cdot F \cdot Y.$$

反之，如果有方阵 $F = (f_{ij})$，借助前面那个等式，令 $f(\varepsilon_i, \varepsilon_j) = f_{ij}$，即可定义 V 上的一个双线性型 f. 因此，在给定 $\mathbb{F}$ 上向量空间的基底 $\varepsilon_1, \cdots, \varepsilon_n$ 后，有矩阵空间 $M_n(\mathbb{F})$ 到 V 上双线性型构成的集合之间的一个一一对应. 其实这就是一个同构映射. 如果

$$f(u,v) = X^t F Y, \quad g(u,v) = X^t G Y.$$

则

$$\alpha f(u,v)+\beta g(u,v)=X^t(\alpha F+\beta G)Y.$$

8.3 双线性型在不同基下矩阵的转换

双线性型 f 的公理性定义与 V 中基的选取是无关的. 为了让前面 f 的矩阵表示有价值，我们补充 $f\mapsto F$ 在新基底下转换时矩阵的转换规则. 设 $\varepsilon_1,\cdots,\varepsilon_n$ 和 $\varepsilon_1',\cdots,\varepsilon_n'$ 是 V 中的两个基底，并有过渡矩阵 $A=(a_{ij})$，则

$$\varepsilon_j=\sum_{i=1}^n a_{ij}\varepsilon_j, j=1,2,\cdots,n.$$

如果 $x_1\varepsilon_1+\cdots+x_n\varepsilon_n=v=x_1'\varepsilon_1'+\cdots+x_n'\varepsilon_n'$. 那么，坐标列 X,X' 可用关系 $X=AX'$ 联系起来. 现在，设 $F=(f_{ij})$ 是双线型 f 在基底 ε_i 下的矩阵，$F'=(f_{ij}')$ 是同一双线性型在基底 ε_i' 下的矩阵，换句话说 $f_{ij}=f(\varepsilon_i,\varepsilon_j)$，$f_{ij}'=f(\varepsilon_i',\varepsilon_j')$. 因为 $(AX')^t=X'^tA^t$，所以

$$X'^tF'Y'=f(u,v)=X^tFY=(AX')^tF(AY)=X'^t(A^tFA)Y'.$$

比较这个等式的最左端和最右端，我们就得到了以下结论.

定理 8.3.1 V 上双线性型 f 在基底 $\varepsilon_1,\cdots,\varepsilon_n$ 和 $\varepsilon_1',\cdots,\varepsilon_n'$ 下的矩阵 F 和 F' 可由关系式

$$F'=A^tFA$$

联系起来，其中 A 是从 ε_i 到 ε_i' 的过渡矩阵.

将上述定理抽象出来，有以下定义.

定义 8.3.1 (合同矩阵) 两个矩阵 F,F'，若存在可逆矩阵 A 使得 $F'=A^tFA$，则称矩阵 F,F' 是**合同的** (**congruence**).

由于可逆矩阵实际上是初等变换而来，而初等变换不会改变矩阵的秩，因此，我们给出 f 的秩的定义.

定义 8.3.2 (双线性型的秩) 双线性型 f 在任意一个基底下对应矩阵的秩称为 f 的**秩** (**rank**).

不难看出，f 的秩是一个不变量，与基底的选择无关.

关于双线性型的不变量还有一个结论.

定义 8.3.3 (左核) 设 V 是 $\mathbb{F}$ 上的线性空间，f 是 V 上的双线性型，定义

$$L_f=\{u\in V|\forall v\in V\ f(u,v)=0\}.$$

显然 L_f 是 V 的一个子空间，称它是 f 的**左根** (**left radical**) 或者**左核** (**left kernel**). 同理可以定义右根与右核.

设 $\varepsilon_1,\cdots,\varepsilon_n$ 是 V 中的一个基底. 条件 $v\in L_f$ 等价于

$$f(v,\varepsilon_1)=0,\cdots,f(v,\varepsilon_n)=0.$$

这是由线性函数 $v \mapsto f_j(v) = f(v, \varepsilon_j) = 0, j = 1, 2, \cdots, n$ 决定的一个方程组. 现在将 $f_j \in V^*$ 按照 ε_i 的对偶基展开，那么它关于对偶基的坐标就是 $f_j(\varepsilon_i), i = 1, 2, \cdots, n$. 即矩阵 F 的第 j 行. 那么，函数组 $f_1, \cdots, f_n$ 的秩和 F 的秩一样，设它们都为 r. 由方程组

$$\begin{cases} f_1(u) = 0, \\ \vdots \\ f_n(u) = 0. \end{cases}$$

设 u 在基 $\varepsilon_1, \cdots, \varepsilon_n$ 下的坐标为 $(x_1, \cdots, x_n)$，就得到线性方程组

$$\begin{cases} f_1(e_1)x_1 + \cdots + f_1(e_n)x_n = 0, \\ \cdots\cdots\cdots \\ f_n(e_1)x_n + \cdots + f_n(e_n)x_n = 0. \end{cases}$$

该齐次线性方程组解空间的维数为 $n - r$，而坐标和向量之间一一对应，所以

$$\dim L_f = n - r = n - \mathrm{rank} F = n - \mathrm{rank} F^t = \dim R_f = n - \mathrm{rank} f.$$

双线性型 f 称为**非退化 (nondegenerate)** 如果 $L_f = \{0\}$，等价地 $R_f = \{0\}$，还等价于 $\mathrm{rank} f = \dim V$ (即 f 是满秩的).

8.4 二次型

考虑具体的对称双线性函数. 设 $f(\alpha, \beta)$ 是 $\mathbb{K}$ 上 n 维线性空间内的一个对称双线性函数,在 V 内取一组基 $\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n$，则 $f(\alpha, \beta)$ 在此组基上的 n 阶对称矩阵为 A. 如果此时 $\alpha, \beta \in V$ 在此组基上的坐标分别为 X, Y，令 $f(\varepsilon_i, \varepsilon_j) = a_{ij}$ 则有

$$f(\alpha, \beta) = \sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij} x_i y_j = X^t A Y.$$

与之对应的二次型函数

$$Q_f(\alpha) = \sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij} x_i x_j = X^t A X.$$

这样，V 上抽象的函数被转换为依赖基选取的一个解析表达式了.

定义 8.4.1 (二次型) 将 $\mathbb{K}$ 内的元素做系数的 n 个变量 $x_1, x_2, \cdots, x_n$ 的二次齐次函数

$$f(x_1, x_2, \cdots, x_n) = \sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij} x_i x_j, \quad (a_{ij} = a_{ji})$$

称为**二次型 (quadratic form)**. 其系数组成的矩阵

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix}$$

是 $\mathbb{K}$ 上的 n 阶对称矩阵，称为此**二次型的矩阵**. 我们将 $f(x_1, x_2, \cdots, x_n)$ 的秩定义为 $\mathrm{rank} A$.

令

$$X = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix},$$

利用矩阵的形式乘法我们可以将二次型表示为

$$f(x_1, x_2, \cdots, x_n) = X^t AX.$$

显然，$\mathbb{K}$ 上的二次型 f 就是 V 内对称线性函数 $Q_f(\alpha)$ 在基 $\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n$ 下的解析表达式. 这样，前面关于对称双线性函数的结果可以全部直接用到二次型 f 上面来.

在实际应用中碰到的二次型，由数域中乘法的交换律知可以把 $x_i x_j$ 和 $x_j x_i$ 两项合并到一起，为了使用前面得到的结论，我们需要把合并后的系数一分为二，得出两项 $a_{ij}x_ix_j$ 和 $a_{ji}x_jx_i$，并且 $a_{ij} = a_{ji}$ 然后再写出二次型的矩阵 $A = (a_{ij})$.

例 8.4.1 给定二次型

$$f(x_1, x_2, x_3) = x_1^2 - 2x_1x_2 + 3x_1x_3 - 2x_2x_3 + 4x_3^2.$$

它的矩阵为

$$A = \begin{pmatrix} 1 & -1 & \frac{3}{2} \\ -1 & 0 & -\frac{1}{2} \\ \frac{3}{2} & -\frac{1}{2} & 4 \end{pmatrix}$$

这是一个对称矩阵. 再利用矩阵的乘法，它又可以表示为

$$f(x_1, x_2, x_3) = (x_1, x_2, x_3) \begin{pmatrix} 1 & -1 & \frac{3}{2} \\ -1 & 0 & -\frac{1}{2} \\ \frac{3}{2} & -\frac{1}{2} & 4 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}.$$

对于任意一个二次型，我们应当能熟练的写出它的矩阵.

对于任意一个 $\mathbb{K}$ 上的二次型，我们希望只有一个矩阵 $A = (a_{ij})$ 与之一一对应. 由于对于任意 $\mathbb{K}$ 上的 n 阶矩阵 A，我们以 a_{ij} 为系数，马上就可以得到一个二次型. 于是我们只需证明，对于任意两个不同的二次型，其定义中所给出的两个矩阵也是不同的.

定理 8.4.1 给定 $\mathbb{K}$ 上的两个二次型

$$f(x_1, x_2, \cdots, x_n) = \sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}x_ix_j = X^t AX, \quad (a_{ij} = a_{ji})$$
$$g(x_1, x_2, \cdots, x_n) = \sum_{i=1}^{n}\sum_{j=1}^{n} b_{ij}x_ix_j = X^t BX, \quad (b_{ij} = b_{ji}).$$

如果有 $f \equiv g$，则 $A = B$.

证明 设 V 内对称双线性函数 $f(\alpha, \beta), g(\alpha, \beta)$ 在基 $\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n$ 下的矩阵分别为 A, B，则二次型函数在该基下的表达式为 $Q_f(\alpha), Q_g(\alpha)$，由于 $f \equiv g$ 所以 $Q_f(\alpha) = Q_g(\alpha)$，即 $f(\alpha, \beta) = g(\alpha, \beta)$. 所以 $A = B$. □

从上面的分析可知，$\mathbb{K}$ 上全体 n 元二次型构成的集合与 $\mathbb{K}$ 上全体 n 阶对称矩阵所构成的集合存在一个一一对应的关系. 在前面也指出，$\mathbb{K}$ 上 n 维线性空间 V 取定一组基 $\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$ 后，$\mathbb{K}$ 上全体 n 阶对称矩阵所成的集合与 V 内全体对称双线性函数组成的集合之间也存在着一一对应关系. 所以，在取定一组基后，对称双线性函数与二次型函数也是一一对应的.

所以我们现在所讨论的东西可以用以下四种等价的语言来描述.

$$\text{二次型的语言} \Leftrightarrow \text{对称矩阵的语言} \Leftrightarrow \text{对称双线性函数的语言} \Leftrightarrow \text{二次型函数的语言}.$$

从函数的观点来看，一个函数的自变量用什么字母表示是无关紧要的. 比如 $\sin x$ 和 $\sin y$ 表示的是同一个函数. 所以二次型中只把 x 换为 y，其它保持不变，则二次型保持不变.

8.4.1 二次型的标准型

我们可以把 $\mathbb{K}$ 上的二次型 $f=X^tAX(A^t=A)$ 看作是 V 内二次型函数 $Q_f(\alpha)$ 在基 $\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$ 下的解析表达式. 现在，在 V 内通过过渡矩阵 T 做变换

$$(\eta_1,\eta_2,\cdots,\eta_n)=(\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_2)T,$$

若 α 在 η_i 下的坐标为 Y，那么 $Q_f(\alpha)$ 在 $\eta_1,\eta_2,\cdots,\eta_n$ 下的解析表达式为 $Y^tBY(B^t=B)$. 现在 A,B 分别是 V 内双线性函数 $f(\alpha,\beta)$ 在基 $\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$ 和在基 $\eta_1,\eta_2,\cdots,\eta_n$ 下的矩阵，于是 $B=C^tAC$. 而且由坐标变换可得 $X=CY$. 即

$$Q_f(\alpha)=X^tAX \xlongequal{X=CY} (CY)^tACY=Y^t(C^tAC)Y=Y^tBY.$$

于是我们可以引进如下概念

定义 8.4.2 (可逆线性替换) 考察系数属于 $\mathbb{K}$ 的如下 n 个变量的线性变数替换

$$\begin{cases} x_1=c_{11}y_1+c_{12}y_2+\cdots+c_{1n}y_n \\ x_2=c_{21}y_1+c_{22}y_2+\cdots+c_{2n}y_n \\ \cdots\cdots\cdots\cdots\cdots\cdots \\ x_n=c_{n1}y_1+c_{n2}y_2+\cdots+c_{nn}y_n \end{cases}.$$

如果其系数矩阵

$$C=\begin{pmatrix} c_{11} & c_{12} & \cdots & c_{1n} \\ c_{21} & c_{22} & \cdots & c_{2n} \\ \vdots & \vdots & & \vdots \\ c_{n1} & c_{n2} & \cdots & c_{nn} \end{pmatrix}$$

可逆，则称该线性变数替换为 $\mathbb{K}$ 上**可逆线性替换 (nondegenerate linear substitution)**.

注 8.4.1 注意，$\mathbb{K}$ 上的二次型 f 只能用可逆方阵 C 来做线性替换.

令

$$X=\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}, Y=\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}$$

则上述线性替换可写为

$$X = CY.$$

我们还可以令 $S = T^{-1}$，这样就有**逆变换 (inverse substitution)**$Y = SX$.

由于 $(C^tAC)^t = C^tA^tC = C^tAC$，所以 C^tAC 仍为 $\mathbb{K}$ 上的对称矩阵. 于是当 $\mathbb{K}$ 上的二次型 $f = X^tAX$ 在做可逆线性替换后得到的 $g = Y^t(C^tBC)Y$ 仍然为二次型. 我们有如下事实.

定理 8.4.2 给定 $\mathbb{K}$ 内两个二次型

$$,f(x_1, x_2, \cdots, x_n) = \sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}x_ix_j \quad (a_{ij} = a_{ji}),$$

$$g(y_1, y_2, \cdots, y_n) = \sum_{i=1}^{n}\sum_{j=1}^{n} b_{ij}y_iy_j \quad (b_{ij} = b_{ji}).$$

它们的矩阵分别为 A, B. 且存在 $\mathbb{K}$ 上的可逆线性替换 $X = CY$，使得 f 变成 g 的充要条件为 A 与 B 在 $\mathbb{K}$ 内合同，即

$$B = C^tAC.$$

证明 如果已经存在 C，使得 f 经变换 $X = CY$ 变成 g，则

$$f = X^tAX \xlongequal{X=CY} Y^t(X^tAX)Y = Y^tBY = g.$$

由**定理 8.4.1**可知

$$B = C^tAC.$$

如果已知存在可逆矩阵 C 使得 $B = C^tAC$，则作线性替换 $X = CY$ 有

$$f = X^tAX \xlongequal{X=CY} Y^t(X^tAX)Y = g = Y^tBY.$$

推论 8.4.1 如果 $\mathbb{K}$ 上的二次型 $f = X^tAX$ 在经过可逆线性替换变成 $g = Y^tBY$，则 f, g 为 V 内同一个二次型函数 $Q_f(\alpha)$ 在两组基下的解析表达式.

证明 如果 A 与 B 合同，等价于它们是对称双线性函数 $f(\alpha, \beta)$ 在两组基下的矩阵，而 $f(\alpha, \beta)$ 又一一对应一个二次型函数 $Q_f(\alpha)$，故 A 和 B 是 $Q_f(\alpha)$ 在两组基下的矩阵，它们的解析表达式分别为 X^tAX, Y^tBY. □

我们在给定 $\mathbb{K}$ 上的两个二次型 f, g. 如果 f 可经 $\mathbb{K}$ 上可逆线性变换变成 g，则称 f, g 满足关系 $\backsim$. 显然，关系 $\backsim$ 是一个等价关系，于是我们直接说满足该关系的 f, g **等价.** 由**定理 8.4.2**可知 f, g 等价的充要条件是 A 与 B 合同. 而合同关系又是 $\mathbb{K}$ 上全体 n 阶对称矩阵集合中的一个等价关系. 因此二次型的合同关系也是全体二次型构成的集合中的一个等价关系. 于是我们在二次型构成的集合中划分出不同的等价类. 于是自然想到，要在每个等价类中选出最简单的一个二次型作为代表元. 而接下来介绍的标准形就是最简单的二次型.

定义 8.4.3 (标准形) 形如

$$\sum_{i=1}^{n} d_i^2 z_i^2 = Z^tDZ \quad (d \in \mathbb{K})$$

的二次型被称为**标准形 (canonical forms)**，其矩阵为对角阵

$$D = \begin{pmatrix} d_1 & & & \\ & d_2 & & \\ & & \ddots & \\ & & & d_n \end{pmatrix}.$$

则前面的定理用二次型的语言可以叙述如下

定理 8.4.3 对于任意一个 $\mathbb{K}$ 上的二次型

$$f(x_1, x_2, \cdots, x_n) \sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij} x_i x_j \quad (a_{ij} = a_{ji}),$$

都存在一个 $\mathbb{K}$ 上的可逆方阵 C，使得 f 在线性替换 $X = CY$ 下变成标准形.

结合前面的内容，我们得到了上述定理三种等价的描述：

1. 用双线性函数语言：$\mathbb{K}$ 上 n 维线性空间 V 中的任意一对称双线性函数的矩阵都可以对角化. 即 V 内存在一组基，使得该对称双线性函数的矩阵是对角阵.

2. 用矩阵论的语言：$\mathbb{K}$ 上任一对称的 n 阶方阵都合同于一个对角阵.

3. 用二次型的语言：$\mathbb{K}$ 上任一二次型都可以经过可逆线性替换变为标准型.

所以我们在做相关的证明时，只需证明上述命题其中一个等价，其它两个就自动成立.

8.4.2 二次型化标准形的计算方法

在拿到一个二次型时，从**定理 8.4.3**可知，它总能化为标准形. 接下来介绍通过可逆线性替换化标准形的方法.

首先介绍的是配方法. 我们直接从例子入手.

例 8.4.2 化二次型

$$f(x_1, x_2, x_3) = 2x_1x_2 - 6x_1x_3 + 2x_2x_3 + x_3^2.$$

为标准形.

解 在化时需注意的是，每一次线性变数替换都要是可逆的. 所以我们可以按照 x_1, x_2, x_3 的顺序进行配方（或者反过来也行）.

首先对 x_3 进行配方.

$$f = -9x_1^2 + 8x_1x_2 - x_2^2 + (-3x_1 + x_2 + x_3)^2.$$

在对 x^3 配方时，有个目标就是要除了在平方式中出现 x_3，平方式以外就没有 x^3 了. 这时我们做可逆线性替换

$$\begin{cases} y_1 = x_1, \\ y_2 = x_2, \\ y_3 = -3x_1 + x_2 + x_3. \end{cases}$$

于是原式子可化为

$$-9y_1^2 + 8y_1y_2 - y_2^2 + 3y_3^2.$$

由于 y_1y_2 系数不为 0，所以我们再对 y_1 进行配方，得

$$-9\left(y_1 - \frac{4}{9}\right)^2 + \frac{7}{9}y_2^2 + y_3^3.$$

最后再做可逆线性替换即可得

$$-9z_1^2 + \frac{7}{9}z_2^2 + z_3^2.$$

上面的题目我们还可以用矩阵来表示，即由 $X = CY$，于是我们只需找出 C 即可. 于是写出二次型的矩阵 A. 把它和 3 阶单位矩阵 E 放在一起得到

$$(A|E).$$

然后做对称初等变换，让 A 变为对角矩阵，然后 E 就会变成 C^t（如果是竖着放则变成 C）. 同样可以得到答案. 但是值得注意的是，化为标准形后，可能标准形的样子会因为线性替换方式的不同而改变. 所以我们在下一节会深入研究.

第九章　Euclid 空间