§1.1 域

线性代数的基本研究对象是线性空间和其间的线性映射. 线性空间的概念需要以一个域为基础来建立. 本节讨论域的基本概念. 粗略地说, 域是一个这样的集合, 其中的元素可以做加法、减法、乘法和除法运算, 并且这些运算满足常见的性质. 严格定义如下.

定义 1.1. 设非空集合F上给定了两个运算, 称为**加法**和**乘法**, 分别对F中任意两个元素x, y给出F中的元素$x+y$和xy, 并且满足下列性质:

(1) 对任意$x, y, z \in F$有

$$x+y=y+x, \quad xy=yx, \quad (x+y)+z=x+(y+z), \quad x(yz)=(xy)z, \quad x(y+z)=xy+xz.$$

(2) 存在两个不同的元素$0_F, 1_F \in F$满足对任意$x \in F$有$0_F+x=x, 1_Fx=x$.

(3) 对任意$x \in F$, 存在$y \in F$满足$x+y=0_F$, 并且当$x \neq 0_F$时, 存在$z \in F$满足$xz=1_F$.

则称集合F(连同它上面的加法和乘法运算)为一个**域**(field).

注 1.1.
- 加法和乘法运算指两个映射$\alpha, \mu : F \times F \to F$, $\alpha(x,y)=x+y$, $\mu(x,y)=xy$. 这里$F \times F$指F中元素的有序对构成的集合, 即

$$F \times F = \{(x,y) \mid x, y \in F\}.$$

- 性质(1)中的前两个式子分别称为加法和乘法的交换率, 接下来的两个式子分别称为加法和乘法的结合率, 最后一个式子称为乘法对加法的分配律. 由交换率和结合律, 对任意有限个$x_1, \ldots, x_n \in F$, 表达式$\sum_{i=1}^n x_i$和$\prod_{i=1}^n x_i$有意义. 由分配律, $(\sum_{i=1}^n x_i)(\sum_{j=1}^m y_j) = \sum_{i,j} x_i y_j$.

引理 1.1. 满足性质(2)的元素0_F和1_F是唯一的.

证明. 设$0_F, 0'_F \in F$满足对任意$x \in F$有$0_F+x=x$, $0'_F+x=x$. 在第一个式子中取$x=0'_F$, 得$0_F+0'_F=0'_F$. 在第二个式子中取$x=0_F$, 得$0'_F+0_F=0_F$. 但$0_F+0'_F=0'_F+0_F$. 所以$0'_F=0_F$.

类似地, 设$1_F, 1'_F \in F$满足对任意$x \in F$有$1_Fx=x$, $1'_Fx=x$. 在第一个式子中取$x=1'_F$, 得$1_F1'_F=1'_F$. 在第二个式子中取$x=1_F$, 得$1'_F1_F=1_F$. 但$1_F1'_F=1'_F1_F$. 所以$1'_F=1_F$. □

满足性质(2)的唯一元素0_F和1_F分别称为F中的**零元素**和**壹元素**. 当无歧义时, 分别记为0和1. 此时需要注意, 它们一般并不是真正的数, 只是F中的两个特殊元素, 分别用“0”和“1”这两个记号来表示.

引理 1.2. (1) 对任意$x \in F$, 满足性质(3)的元素y是唯一的.

(2) 对任意$x \in F \setminus \{0\}$, 满足性质(3)的元素z是唯一的.

证明. (1) 设$y, y' \in F$满足$x+y=x+y'=0$. 则

$$y' = 0+y' = (y+x)+y' = y+(x+y') = y+0 = y.$$

(2) 设$z, z' \in F$满足$xz=xz'=1$. 则

$$z' = 1z' = (zx)z' = z(xz') = z1 = z.$$

□

对$x \in F$, 满足$x+y=0$的唯一元素$y \in F$称为x在F中的**负元素**, 记为$-x$. 当$x \neq 0$时, 满足$xz=1$的唯一元素$z \in F$称为x在F中的**逆元素**, 记为x^{-1}. 我们可以把这里的负号视为取负元

素的操作, 把取逆符号视为取逆元素的操作, 即考虑映射

$$F \to F, x \mapsto -x, \qquad F \setminus \{0\} \to F, x \mapsto x^{-1}.$$

于是, 形如$-(-x)$ 和$-x^{-1}$的表达式有意义. 我们定义减法和除法运算为

$$x - y = x + (-y), \qquad x/y = xy^{-1}\ (y \neq 0).$$

命题 1.3. 给定域F.

(1) (加法消去律)设$x, y, z \in F$. 如果$x + z = y + z$, 则$x = y$.

(2) 对任意$x \in F$有$0x = 0$.

(3) 设$x, y \in F$满足$xy = 0$. 则$x = 0$或$y = 0$.

证明. (1) 两边同时加上$-z$, 得

$$(x + z) + (-z) = (y + z) + (-z).$$

而

$$(x + z) + (-z) = x + (z + (-z)) = x + 0 = x.$$

类似地,

$$(y + z) + (-z) = y.$$

因此$x = y$.

(2) 注意到$0 + 0 = 0$. 因此$(0 + 0)x = 0x$. 这推出$0x + 0x = 0x + 0$. 由(1)得$0x = 0$.

(3) 若$x \neq 0$, 则$x^{-1}(xy) = (x^{-1}x)y = 1y = y$. 另一方面, 由(2)有$x^{-1}(xy) = x^{-1}0 = 0$. 因此$y = 0$. □

注 1.2. 在域的定义中, 我们要求$1 \neq 0$. 从而域中至少有两个元素. 如果不做此要求, 当$1 = 0$时, 由上面命题中的(3)(其证明没有用到$1 \neq 0$), 对任意$x \in F$有$x = 1x = 0x = 0$. 因此$F = \{0\}$. 定义中的要求即是为了排除掉这种情况.

例 1.1.

- 有理数集$\mathbb{Q}$、实数集$\mathbb{R}$和复数集$\mathbb{C}$在数的加法和乘法下是域, 分别称为有理数域、实数域和复数域.
- $\mathbb{Q}(\sqrt{2}) := \{x + \sqrt{2}y \mid x, y \in \mathbb{Q}\}$在数的加法和乘法下是域.
- 整数集$\mathbb{Z}$和正整数集$\mathbb{N}$在数的加法和乘法下不是域.
- 设p是素数. 对$x \in \mathbb{Z}$, 记

$$\bar{x} = \{n \in \mathbb{Z} \mid n \equiv x \mod p\},$$

$$\mathbb{F}_p := \{\bar{x} \mid x \in \mathbb{Z}\}.$$

容易看出, $\mathbb{F}_p = \{\bar{0}, \bar{1}, \ldots, \overline{p-1}\}$. 因此$|\mathbb{F}_p| = p$. 定义

$$\bar{x} + \bar{y} = \overline{x + y}, \qquad \bar{x}\bar{y} = \overline{xy}.$$

容易证明该定义良定, 并且$\mathbb{F}_p$是域(主要是$\bar{x}^{-1}$存在). □

定义 1.2. 域F的子集F'称为F的**子域**(subfield), 如果

(1) $0_F, 1_F \in F'$,

(2) $x, y \in F' \Longrightarrow x + y, -x, xy, x^{-1}(当x \neq 0时) \in F'$.

注意到子域是域.

例 1.2. $\mathbb{Q}, \mathbb{R}, \mathbb{Q}(\sqrt{2})$都是$\mathbb{C}$的子域.

对$n \in \mathbb{N}$, 记$n_F = \overbrace{1_F + \cdots + 1_F}^{n\text{个}}$. 容易看出, 对任意$m, n \in \mathbb{N} \cup \{0\}$有

$$(m+n)_F = m_F + n_F, \qquad (mn)_F = m_F n_F.$$

定义 1.3. 对于域F, 其**特征**(characteristic)char$F$定义如下: 如果对任意$n \in \mathbb{N}$有$n_F \neq 0_F$, 则定义char$F = 0$; 否则, 定义

$$\mathrm{char}F = \min\{n \in \mathbb{N} \mid n_F = 0_F\}.$$

命题 1.4. 假设$p = \mathrm{char}F \neq 0$. 则

(1) p为素数.

(2) 对$n \in \mathbb{N}$有$n_F = 0_F \iff p \mid n$.

证明. (1) 假设p不是素数. 注意到$p \geq 2$. 于是存在$r, s \in \{2, \ldots, n-1\}$满足$n = rs$. 这推出$r_F s_F = n_F = 0_F$. 另一方面, 由特征定义知$r_F, s_F \neq 0_F$, 从而$r_F s_F \neq 0_F$. 矛盾.

(2) "$\Longleftarrow$". 设$n = pq$. 则$n_F = p_F q_F = 0_F q_F = 0_F$.

"$\Longrightarrow$". 设$n = dp + r$, $0 \leq r < p$. 则$0_F = n_F = d_F p_F + r_F = d_F 0_F + r_F = r_F$. 这推出$r = 0$, 即$p \mid n$. □

命题 1.5. 设F'是域F的子域. 则charF' = char$F$.

证明. 注意到对任意$n \in \mathbb{N} \cup \{0\}$有$n_F = n_{F'}$. 因此$n_F = 0_F \iff n_{F'} = 0_{F'}$. □

例 1.3. $\mathbb{R}, \mathbb{C}, \mathbb{Q}$的特征为0, $\mathbb{F}_p$的特征为p.

对$n \in \mathbb{N}$和$x \in F$, 定义

$$nx := \overbrace{x + \cdots + x}^{n\text{个}} = \overbrace{1_F x + \cdots + 1_F x}^{n\text{个}} = (\overbrace{1_F + \cdots + 1_F}^{n\text{个}})x = n_F x.$$

注意$n1_F = n_F$. 由于下面的命题, 有时我们需要把特征为0的域与特征非零的域区别对待.

命题 1.6. (1) 如果char$F = 0$, 则$nx = 0_F \Longrightarrow x = 0_F$.

(2) 如果char$F = p > 0$并且$p \nmid n$, 则$nx = 0_F \Longrightarrow x = 0_F$.

(3) 如果char$F = p > 0$并且$p \mid n$, 则对任意$x \in F$有$nx = 0_F$.

证明. (1)+(2) $nx = 0_F \iff n_F x = 0_F$. 在(1)和(2)的条件下有$n_F \neq 0_F$. 因此$x = 0_F$.

(3) 此时总有$n_F = 0$. 因此$nx = n_F x = 0$. □

对于$n \in \mathbb{N}$, $a \in F$, 如果$n_F \neq 0_F$, 定义$\frac{1}{n}a := n_F^{-1}a$. 当char$F = 0$时, $\frac{1}{n}a$总是有定义的. 当char$F = p > 0$时, $\frac{1}{n}a$有定义的充要条件是$p \nmid n$. 容易看出, 对这两种情况, 关于$x \in F$的方程$nx = a$有唯一解$x = \frac{1}{n}a$.

§2.1 线性空间

取定域F和正整数n. 如果$x_1,\ldots,x_n\in F$, 则称n元有序组$(x_1,\ldots,x_n)$为一个域F上的n**维向量**或n**维**F**-向量**, 并称每个x_i为该向量的一个**分量**. 考虑集合

$$F^n:=\{(x_1,\ldots,x_n)\mid x_i\in F\},$$

并在它上面定义

- 向量加法: 对$\alpha=(x_1,\ldots,x_n),\beta=(y_1,\ldots,y_n)\in F^n$, 定义
$$\alpha+\beta=(x_1+y_1,\ldots,x_n+y_n).$$
- 纯量乘法: 对$c\in F$, $\alpha=(x_1,\ldots,x_n)\in F^n$, 定义
$$c\alpha=(cx_1,\ldots,cx_n).$$

例 2.1. 集合$\mathbb{R}^2$或$\mathbb{R}^3$中的向量可以理解为平面或空间中的点, 也可以理解为从原点出发指向该点的“箭头”. $\alpha+\beta$按照平行四边形法则. $c\alpha$与α方向相同或相反, 长度为α的长度的$|c|$倍. □

我们把F^n上的向量加法和纯量乘法的性质提炼出来, 引入下面的定义.

定义 2.1. 取定域F. 设集合V上给定了两个运算:

- 向量加法$V\times V\to V$, $(\alpha,\beta)\mapsto\alpha+\beta$;
- 纯量乘法$F\times V\to V$, $(c,\alpha)\mapsto c\alpha$,

并且满足下面的性质:

(1) 对任意$\alpha,\beta\in V$有$\alpha+\beta=\beta+\alpha$.

(2) 对任意$\alpha,\beta,\gamma\in V$有$(\alpha+\beta)+\gamma=\alpha+(\beta+\gamma)$.

(3) 存在$0_V\in V$满足对任意$\alpha\in V$有$\alpha+0_V=\alpha$.

(4) 对任意$\alpha\in V$, 存在$\beta\in V$满足$\alpha+\beta=0_V$.

(5) 对任意$\alpha\in V$有$1_F\alpha=\alpha$.

(6) 对任意$c_1,c_2\in F$和$\alpha\in V$有$(c_1c_2)\alpha=c_1(c_2\alpha)$.

(7) 对任意$c\in F$和$\alpha,\beta\in V$有$c(\alpha+\beta)=c\alpha+c\beta$.

(8) 对任意$c_1,c_2\in F$和$\alpha\in V$有$(c_1+c_2)\alpha=c_1\alpha+c_2\alpha$.

则称V为域F上的**线性空间**或**向量空间**, 简称为F**-线性空间**或F**-向量空间**. V中的元素称为**向量**. 性质(3)中的0_V称为**零向量**或**原点**(当无歧义时简记为0). 性质(4)中的β称为α的**负向量**, 记为$-\alpha$.

注意这里与教材的区别: 我们在线性空间的定义中不要求零向量和负向量的唯一性, 而是将由上面的定义推出它们的唯一性. 这样做的优点是简化了例子的验证.

引理 2.1. 在线性空间中, 零向量是唯一的, 任意向量的负向量也是唯一的.

证明. 与引理1.1和1.2的证明类似. 设$0_V,0'_V\in V$满足对任意$\alpha\in V$有$0_V+\alpha=\alpha$, $0'_V+\alpha=\alpha$. 在第一个式子中取$\alpha=0'_V$, 得$0_V+0'_V=0'_V$. 在第二个式子中取$\alpha=0_V$, 得$0'_V+0_V=0_V$. 但$0_V+0'_V=0'_V+0_V$. 所以$0'_V=0_V$. 为证明负向量的唯一性, 设$\alpha\in V$, 并设$\beta,\beta'\in V$满足$\alpha+\beta=\alpha+\beta'=0_V$. 则

$$\beta'=0_V+\beta'=(\beta+\alpha)+\beta'=\beta+(\alpha+\beta')=\beta+0_V=\beta.$$

□

注 2.1. 类似的证明出现了三次(其他两次见引理1.1和1.2的证明), 原因是我们没有进行进一步的"公理化". 事实上, 引理1.1、1.2和2.1中的唯一性都可以归结为群中单位元和逆元的唯一性. 也许在以后的讲义中会先介绍群的概念.

由性质(1)和(2), 对任意有限个$\alpha_1, \ldots, \alpha_n \in V$, 表达式$\sum_{i=1}^n \alpha_i$有意义(与相加次序无关). 容易验证:

$$\sum c_i\alpha_i + \sum d_i\alpha_i = \sum(c_i + d_i)\alpha_i,$$
$$c\sum c_i\alpha_i = \sum(cc_i)\alpha_i.$$

例 2.2. F^n是线性空间. 其中

$$0_{F^n} = (0_F, \ldots, 0_F),$$

并且对$\alpha = (x_1, \ldots, x_n)$有

$$-\alpha = (-x_1, \ldots, -x_n).$$

□

线性空间中的元素不一定是真的"向量". 例如下面的例子.

例 2.3. 给定集合S. 考虑映射(函数)的集合

$$F^S := \{f : S \to F\}.$$

定义

$$(f+g)(s) = f(s) + g(s), \qquad \forall f, g \in F^S, s \in S,$$
$$(cf)(s) = cf(s), \qquad \forall s \in S, c \in F, f \in F^S.$$

验证性质:

(1).

$$(f+g)(s) = f(s) + g(s) = g(s) + f(s) = (g+f)(s), \forall s \in S \quad \Longrightarrow \quad f + g = g + f.$$

(2).

$$((f+g)+h)(s) = (f+g)(s) + h(s) = (f(s) + g(s)) + h(s),$$
$$(f+(g+h))(s) = f(s) + (g+h)(s) = f(s) + (g(s) + h(s)).$$

两式右边相等. 所以$(f+g)+h = f+(g+h)$.

(3). 取0_{F^S}为在S上恒等于0_F的函数.

(4). 对$f \in F^S$, 令$(-f)(s) = -f(s)$.

(5)–(8)显然.

注意到当$S = \{1, \ldots, n\}$时, $f : \{1, \ldots, n\} \to F$可以等同于$(f(1), \ldots, f(n)) \in F^n$. 反过来, $(x_1, \ldots, x_n) \in F^n$可等同于由$f(i) = x_i$定义的$f : \{1, \ldots, n\} \to F$. 这样$F^S \cong F^n$. □

例 2.4. 设

$$V = \{f : F \to F \mid f\text{是多项式函数}\}.$$

这里多项式函数指形如$f(x) = c_0 + c_1x + \cdots + c_nx^n (c_i \in F)$的函数. 在类似定义的运算下, V是线性空间. □

例 2.5. $\mathbb{C}$是$\mathbb{R}$上的线性空间. $\mathbb{C}^n$也是$\mathbb{R}$上的线性空间. 注意$\mathbb{C}^n$作为实线性空间和复线性空间是不同的线性空间. 一般地, 若V是复线性空间, 则通过执行"忘掉复结构"的操作(即在做纯量乘法$(c, \alpha) \mapsto c\alpha$时只允许$c \in \mathbb{R}$), 可以把$V$视为实线性空间. 更一般地, 若$F'$是$F$的子域, 则$F$是$F'$-

线性空间, F^n也是F'-线性空间, 任意F-线性空间被"忘掉F-结构"后也可视为F'-线性空间. $\square$

命题 2.2. 设V是域F上的线性空间. 则

(1) (加法消去率) 设$\alpha, \beta, \gamma \in V$. 如果$\alpha + \gamma = \beta + \gamma$, 则$\alpha = \beta$.

(2) $c0_V = 0_V, \forall c \in F$.

(3) $0_F\alpha = 0_V, \forall \alpha \in V$.

(4) $c\alpha = 0_V \Longrightarrow c = 0_F$或$\alpha = 0_V$.

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

证明. (1) 两边同时加上$-\gamma$, 得

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

而

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

类似地,

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

因此$\alpha = \beta$. (注意这里的证明与命题1.3(1)的证明类似, 两者都可以归结为群运算中的消去律.)

(2). 注意到$0_V + 0_V = 0_V$. 因此$c(0_V + 0_V) = c0_V$. 这推出$c0_V + c0_V = c0_V + 0_V$. 由(1)得$c0_V = 0_V$.

(3). 注意到$0_F + 0_F = 0_F$. 因此$(0_F + 0_F)\alpha = 0_F\alpha$. 这推出$0_F\alpha + 0_F\alpha = 0_F\alpha + 0_V$. 由(1)得$0_F\alpha = 0_V$.

(4) 若$c \neq 0_F$, 则$c^{-1}(c\alpha) = (c^{-1}c)\alpha = 1_F\alpha = \alpha$. 另一方面, 由(2)有$c^{-1}(c\alpha) = c^{-1}0_V = 0_V$. 因此$\alpha = 0_V$.

(5) 由负向量的唯一性, 只需证明$\alpha + (-1_F)\alpha = 0_V$. 这可以利用(3)验证如下:

$$\alpha + (-1_F)\alpha = 1_F\alpha + (-1_F)\alpha = (1_F + (-1_F))\alpha = 0_F\alpha = 0_V.$$

$\square$

§2.2 子空间

定义 2.1. 设V是域F上的线性空间, W是V的子集. 如果

(1) $0 \in W$;

(2) 对任意$\alpha, \beta \in W$总有$\alpha + \beta \in W$;

(3) 对任意$\alpha \in W$和$c \in F$总有$c\alpha \in W$.

则称W是V的**(线性)子空间**.

容易看出, 如果W是V的子空间, 则V上的向量加法和纯量乘法可以限制在W上, 从而使W成为线性空间.

例 2.1.

- $W = V$或$\{0\}$称为V的**平凡子空间**. 注意空集不是V的子空间.
- $F = \mathbb{R}$, $V = \mathbb{R}^2$, $W =$过原点的直线.
- $F = \mathbb{R}$, $V = \mathbb{R}^3$, $W =$过原点的直线或平面.
- $V = F^n$.

$$W = \{(x_1, \ldots, x_n) \in V \mid x_1 = 0\}$$

和

$$W = \{(x_1, \ldots, x_n) \in V \mid x_1 = 2x_2\}$$

是子空间. 而

$$W = \{(x_1, \ldots, x_n) \in V \mid x_1 = 1 + x_2\}$$

不是子空间.

- $V = F^F = \{$任意函数$f : F \to F\}$, $W = \{$多项式函数$f : F \to F\}$.

定义 2.2. $\beta \in V$称为$\alpha_1, \ldots, \alpha_n \in V$的**线性组合**, 如果存在$c_1, \ldots, c_n \in F$满足$\beta = \sum_{i=1}^n c_i\alpha_i$.

命题 2.1. 对于V的子集W, TFAE:

(1) W是子空间.

(2) W非空, 并且W中任意有限个向量的线性组合仍在W中.

(3) $0 \in W$, 并且对任意$\alpha, \beta \in W$和$c \in F$总有$c\alpha + \beta \in W$.

证明. "(1)$\Longrightarrow$(2)". 显然W非空. 若$\alpha_1, \ldots, \alpha_n \in W$, $c_1, \ldots, c_n \in F$, 则$c_i\alpha_i \in W$, 从而$\sum_{i=1}^n c_i\alpha_i \in W$.

"(2)$\Longrightarrow$(3)". 由于W非空, 所以可以取$\alpha \in W$. 而$0 = 0\alpha$是α的线性组合. 所以$0 \in W$. 另一结论显然.

"(3)$\Longrightarrow$(1)". 设$\alpha, \beta \in W$. 则$\alpha + \beta = 1\alpha + \beta \in W$. 另一方面, 设$\alpha \in W$, $c \in F$, 则$c\alpha = c\alpha + 0 \in W$. □

在证明V的子集是子空间时, 利用命题2.1(3)会带来方便.

下面讨论子空间的交与和. 对于子集$S_1, \ldots, S_k \subset V$, 记

$$\sum_{i=1}^k S_i := \left\{ \sum_{i=1}^k \alpha_i \mid \alpha_i \in S_i \right\}.$$

命题 2.2. (1) 设$\{W_a \mid a \in A\}$是V的一族子空间. 则$\bigcap_{a \in A} W_a$也是V的子空间.

(2) 设$W_1, \ldots, W_k$是V的子空间. 则$\sum_{i=1}^k W_i$也是V的子空间.

证明. (1). 首先, 对任意$a \in A$有$0 \in W_a$, 所以$0 \in \bigcap_{a\in A} W_a$. 另一方面, 设$\alpha, \beta \in \bigcap_{a\in A} W_a, c \in F$. 则对任意$a \in A$有$\alpha, \beta \in W_a$, 从而$c\alpha + \beta \in W_a$. 因此$c\alpha + \beta \in \bigcap_{a\in A} W_a$.

(2). 显然$0 \in \sum_{i=1}^k W_i$. 设$\alpha, \beta \in \sum_{i=1}^k W_i, c \in F$. 则$\alpha = \sum_{i=1}^k \alpha_i$, $\beta = \sum_{i=1}^k \beta_i$, 其中$\alpha_i, \beta_i \in W_i$. 注意到$c\alpha_i + \beta_i \in W_i$. 于是

$$c\alpha + \beta = \sum_{i=1}^k (c\alpha_i + \beta_i) \in \sum_{i=1}^k W_i.$$

□

例 2.2. 设$V = F^4$,

$$W_1 = \{(x, y, z, 0) \mid x, y, z \in F\},$$
$$W_2 = \{(x, 0, 0, y) \mid x, y \in F\}.$$

则

$$W_1 + W_2 = V, \qquad W_1 \cap W_2 = \{(x, 0, 0, 0) \mid x \in F\}.$$

□

命题 2.3. 设$S \subset V$非空. 记

$$\mathrm{span} S := \{S\text{中任意有限个向量的线性组合}\}.$$

则$\mathrm{span} S$是子空间.

证明. 显然$0 \in \mathrm{span} S$. 设$\alpha, \beta \in \mathrm{span} S, c \in F$. 则$\alpha = \sum_{i=1}^m c_i\alpha_i$, $\beta = \sum_{j=1}^n d_j\beta_j$, 其中$\alpha_i, \beta_j \in S$. 从而

$$c\alpha + \beta = \sum_{i=1}^m (cc_i)\alpha_i + \sum_{j=1}^n d_j\beta_j \in \mathrm{span} S.$$

□

定义 2.3. $\mathrm{span} S$称为**由S生成的子空间**. 如果$S = \{\alpha_1, \ldots, \alpha_n\}$有限, 也称$\mathrm{span}\{\alpha_1, \ldots, \alpha_n\}$为由$\alpha_1, \ldots, \alpha_n$生成的子空间. 我们约定$\mathrm{span}\varnothing = \{0\}$.

命题 2.4. 设$S \subset V$. 则

(1) 如果W是V的子空间并且包含S, 则W包含$\mathrm{span} S$. 从而$\mathrm{span} S$是包含S的最小子空间.

(2)

$$\mathrm{span} S = \bigcap_{\substack{W\text{是}V\text{的包}\\ \text{含}S\text{的子空间}}} W.$$

证明. (1) 若$S = \varnothing$则显然. 设$S \neq \varnothing$. 则对任意$\alpha \in \mathrm{span} S$有$\alpha = \sum_{i=1}^n c_i\alpha_i$, 其中$\alpha_i \in S \subset W$. 因此$\alpha \in W$. 这说明$\mathrm{span} S \subset W$.

(2) 记命题中的交集为W'. 它是V的子空间并且包含S. 由(1)即得$\mathrm{span} S \subset W'$. 另一方面, 由$W'$的定义, 它包含在每个包含$S$的子空间之中. 而$\mathrm{span} S$是包含$S$的子空间. 因此$W' \subset \mathrm{span} S$. 这就说明了$\mathrm{span} S = W'$. □

命题 2.5. 设$W_1, \ldots, W_k$是V的子空间. 则$\sum_{i=1}^k W_i = \mathrm{span} \bigcup_{i=1}^k W_i$.

证明. 注意到$\bigcup_{i=1}^k W_i \subset \sum_{i=1}^k W_i$. 由命题2.4(1)即得$\mathrm{span} \bigcup_{i=1}^k W_i \subset \sum_{i=1}^k W_i$. 另一方面, 对任意$\alpha \in \sum_{i=1}^k W_i$有$\alpha = \sum_{i=1}^k \alpha_i$, 其中$\alpha_i \in W_i \subset \bigcup_{i=1}^k W_i$. 所以$\alpha \in \mathrm{span} \bigcup_{i=1}^k W_i$. 这说明$\sum_{i=1}^k W_i \subset \mathrm{span} \bigcup_{i=1}^k W_i$. 因此有$\sum_{i=1}^k W_i = \mathrm{span} \bigcup_{i=1}^k W_i$. □

例 2.3. 设$F=\mathbb{R}$, $V=\mathbb{R}^5$, $\alpha_1=(1,2,0,3,0)$, $\alpha_2=(0,0,1,4,0)$, $\alpha_3=(0,0,0,0,1)$. 则

$$\begin{aligned}\operatorname{span}\{\alpha_1,\alpha_2,\alpha_3\}&=\{c_1\alpha_1+c_2\alpha_2+c_3\alpha_3\mid c_i\in\mathbb{R}\}\\&=\{(c_1,2c_1,c_2,3c_1+4c_2,c_3)\mid c_i\in\mathbb{R}\}\\&=\{(x_1,\ldots,x_5)\mid x_2=2x_1,x_4=3x_1+4x_3\}.\end{aligned}$$

□

例 2.4. 设$V=\{f:F\to F$是多项式函数$\}$. 定义$f_n\in V$为$f_n(x)=x^n$, $n=0,1,\ldots$. 则

$$\operatorname{span}\{f_0,f_1,\ldots\}=V.$$

□

§2.3 基和维数

给定域F和F-线性空间V.

定义 2.1. 设S是V的非空子集. 称S是**线性相关**的, 如果存在有限个互不相同的$\alpha_1, \ldots, \alpha_n \in S$和不全为0的$c_1, \ldots, c_n \in F$满足$\sum_{i=1}^n c_i\alpha_i = 0$. 否则, 称$S$是**线性无关**的. 约定空集是线性无关的.

注 2.1.

- S线性无关$\Longleftrightarrow$对任意有限个互不相同的$\alpha_1, \ldots, \alpha_n \in S$, 如果$c_1, \ldots, c_n \in F$满足$\sum_{i=1}^n c_i\alpha_i = 0$, 则$c_1 = \cdots = c_n = 0$.
- 若$S = \{\alpha_1, \ldots, \alpha_n\}$有限, 我们也称$\alpha_1, \ldots, \alpha_n$线性相关或线性无关. 此时,
  S线性相关$\Longleftrightarrow$存在不全为0的$c_1, \ldots, c_n \in F$满足$\sum_{i=1}^n c_i\alpha_i = 0$,
  S线性无关$\Longleftrightarrow$如果$c_1, \ldots, c_n \in F$满足$\sum_{i=1}^n c_i\alpha_i = 0$, 则$c_1 = \cdots = c_n = 0$.
- S线性无关$\Longleftrightarrow$ S的任意有限子集线性无关.
- 若$S \subset T \subset V$, 则S线性相关$\Longrightarrow$ T线性相关, T线性无关$\Longrightarrow$ S线性无关.
- 若$0 \in S$则S线性相关: $1_F 0_V = 0_V$.

例 2.1. 设$F = \mathbb{C}$, $V = \mathbb{C}^3$, $\alpha_1, \ldots, \alpha_4$见P41. 则它们线性相关: $2\alpha_1 + 2\alpha_2 - \alpha_3 + 0\alpha_4 = 0$. □

定义 2.2. 设$S \subset V$. 如果S线性无关并且生成V, 则称S是V的**基**. 如果V存在有限基, 则称V是**有限维**的; 否则, 称V是**无穷维**的.

例 2.2. 空集是$\{0\}$的基. □

例 2.3. 设$V = F^n$. 取$\epsilon_1 = (1, 0, \ldots, 0), \ldots, \epsilon_n = (0, \ldots, 0, 1)$. 则$\{\epsilon_1, \ldots, \epsilon_n\}$是$F^n$的基, 称为$F^n$的**标准基**. 验证如下:
线性无关: 设$c_1, \ldots, c_n \in F$满足$\sum_{i=1}^n c_i\epsilon_i = 0$, 则$(c_1, \ldots, c_n) = 0$, 即$c_1 = \cdots = c_n = 0$.
$\mathrm{span}\{\epsilon_1, \ldots, \epsilon_n\} = V$: 对任意$\alpha = (x_1, \ldots, x_n) \in F^n$, 有$\alpha = \sum_{i=1}^n x_i\epsilon_i \in \mathrm{span}\{\epsilon_1, \ldots, \epsilon_n\}$.
这说明F^n是有限维的. □

例 2.4. 设F是$\mathbb{C}$的子域, $V = \{$多项式函数$f : F \to F\}$. 定义$f_k \in V$为$f_k(x) = x^k$, $k \in \mathbb{N} \cup \{0\}$. 则$S := \{f_k \mid k \in \mathbb{N} \cup \{0\}\}$是$V$的基:
$\mathrm{span}S = V$: 见上一节.
S线性无关: 对任意互不相同的$f_{k_1}, \ldots, f_{k_n}$, 如果$c_1, \ldots, c_n \in F$满足$\sum_{i=1}^n c_i f_{k_i} = 0$, 则

$$\sum_{i=1}^n c_i x^{k_i} = \sum_{i=1}^n c_i f_{k_i}(x) = \left(\sum_{i=1}^n c_i f_{k_i}\right)(x) = 0, \qquad \forall x \in F.$$

由于$F \subset \mathbb{C}$, 所以含有无穷多个元素. 而非零复多项式只有有限多个根, 所以$c_i = 0$. □

例2.4中的空间V是无穷维的: 对任意有限集$T = \{g_1, \ldots, g_r\} \subset V$, 存在正整数$N$满足$\deg g_i < N$, 于是$f_N \notin \mathrm{span}T$, 从而$\mathrm{span}T \neq V$. 类似的断言总成立:

定理 2.1. 设$S, T \subset V$, $\mathrm{span}S = V$, T线性无关. 若S有限, 则T有限, 并且$|T| \leq |S|$.

证明. 设$|S| = n$. 若结论不成立, 则可以取T的$n+1$元子集$T' = \{\alpha_1, \ldots, \alpha_{n+1}\}$. 先归纳证明下面的断言: 对任意$k \in \{0, \ldots, n\}$, 存在$S_k \subset S$, $|S_k| = n-k$使得$\mathrm{span}(S_k \cup \{\alpha_1, \ldots, \alpha_k\}) = V$. $k = 0$时显然. 假设$1 \leq k \leq n$, 并且当把k替换为$k-1$时断言成立, 即存在$S_{k-1} \subset S$, $|S_{k-1}| = n-k+1$使得$\mathrm{span}(S_{k-1} \cup \{\alpha_1, \ldots, \alpha_{k-1}\}) = V$. 设$S_{k-1} = \{\beta_1, \ldots, \beta_{n-k+1}\}$. 则$\alpha_k$可以表示为

$$\alpha_k = \sum_{i=1}^{k-1} c_i\alpha_i + \sum_{j=1}^{m-k+1} d_j\beta_j.$$

由于T'线性无关, 必有某个$d_{j_0} \neq 0$. 取$S_k = S_{k-1} \setminus \{\beta_{j_0}\}$. 我们证明

$$\mathrm{span}(S_k \cup \{\alpha_1, \ldots, \alpha_k\}) = V.$$

事实上, 由于集合$S_{k-1} \cup \{\alpha_1, \ldots, \alpha_{k-1}\}$生成$V$, 只需证明该集合包含在$\mathrm{span}(S_k \cup \{\alpha_1, \ldots, \alpha_k\})$中, 为此只需验证$\beta_{j_0} \in \mathrm{span}(S_k \cup \{\alpha_1, \ldots, \alpha_k\})$, 这由

$$\beta_{j_0} = d_{j_0}^{-1} \left(\alpha_k - \sum_{i=1}^{k-1} c_i \alpha_i - \sum_{\substack{1 \leq j \leq m-k+1 \\ j \neq j_0}} d_j \beta_j \right)$$

所保证. 这就证明了断言.

在断言中取$k = n$, 得$\mathrm{span}\{\alpha_1, \ldots, \alpha_n\} = V$. 特别地, $\alpha_{n+1} \in \mathrm{span}\{\alpha_1, \ldots, \alpha_n\}$, 与$T'$线性无关矛盾. □

推论 2.2. 设V是有限维线性空间, 则V的任意基有限, 并且任意两组基含有相同的向量个数.

证明. 由于V存在有限基, 由上面定理, V的任意基有限. 进一步, 上面定理推出对V的任意两组基S_1, S_2有$|S_1| \leq |S_2|$并且$|S_2| \leq |S_1|$. 因此$|S_1| = |S_2|$. □

定义 2.3. 设V是有限维线性空间, V的基含有的共同向量个数称为V的**维数**, 记为$\dim V$.

上面的例子已经说明, $\dim F^n = n$, $\dim\{0\} = 0$.

推论 2.3. 设V是有限维线性空间, $S \subset V$.

(1) 如果S线性无关, 则$|S| \leq \dim V$. 因此$\dim V$是V的线性无关子集中向量个数的最大值.

(2) 如果S生成V, 则$|S| \geq \dim V$. 因此$\dim V$是V的生成全空间的子集中向量个数的最小值.

证明. 由定理2.1显然. □

定理 2.4. 设V是有限维线性空间, $S_1 \subset S_2 \subset V$. 假设$S_1$线性无关, S_2生成V. 则存在V的基S满足$S_1 \subset S \subset S_2$. 特别地, V的任意线性无关子集可以扩充为V的基.

先证明:

引理 2.5. 设V是任意线性空间, $S \subset V$线性无关, $\beta \in V$. 如果$\beta \notin \mathrm{span}S$, 则$S \cup \{\beta\}$线性无关.

证明. 只需证明$S \cup \{\beta\}$的任意有限子集T线性无关. 若$\beta \notin T$, 则$T \subset S$, 从而显然线性无关. 假设$\beta \in T$. 设$T = \{\alpha_1, \ldots, \alpha_n, \beta\}$, 其中$\alpha_i \in S$, 并且$\sum_{i=1}^n a_i \alpha_i + b\beta = 0$. 如果$b \neq 0$, 则$\beta = \sum_{i=1}^n (-b^{-1} a_i) \alpha_i \in \mathrm{span}S$, 矛盾. 因此$b = 0$. 这推出$\sum_{i=1}^n a_i \alpha_i = 0$. 由于$S$线性无关, 这推出$a_i = 0$. □

定理2.4的证明. 考虑V的子集族

$$\mathcal{F} = \{S' \subset V \mid S'\text{线性无关}, S_1 \subset S' \subset S_2\}.$$

由推论2.3(1), V的线性无关子集至多含有$\dim V$个向量. 于是, $\mathcal{F}$中的集合含有的向量个数存在最大值, 设为n. 取$S \in \mathcal{F}$满足$|S| = n$. 我们断言$S_2 \subset \mathrm{span}S$. 若不然, 则可以取$\beta \in S_2 \setminus \mathrm{span}S$. 由引理2.5, $S \cup \{\beta\}$线性无关, 并且$S_1 \subset S \cup \{\beta\} \subset S_2$. 因此$S \cup \{\beta\} \in \mathcal{F}$. 但是$|S \cup \{\beta\}| = n+1$, 与$n$为最大值矛盾. 这就证明了$S_2 \subset \mathrm{span}S$. 从而$V = \mathrm{span}S_2 \subset \mathrm{span}S$, 即$\mathrm{span}S = V$. 因此$S$是$V$的基. □

推论 2.6. 设V是有限维线性空间, $S \subset V$满足$|S| = \dim V$. 假设S线性无关或者S生成V. 则S是V的基.

证明. 如果S线性无关, 则S可以扩充为V的基S_1. 但是$|S| = \dim V = |S_1|$. 因此$S = S_1$是基. 类似地, 如果S生成V, 则存在V的基$S_2 \subset S$. 但是$|S| = \dim V = |S_2|$. 因此$S = S_2$是基. □

接下来考虑子空间的维数性质.

命题 2.7. 设V是有限维线性空间, $W \subset V$是子空间. 则

(1) W是有限维线性空间, 并且$\dim W \leq \dim V$;

(2) 若W是真子空间, 则$\dim W < \dim V$.

证明. 注意到W的线性无关子集也是V的线性无关子集, 从而至多含有$\dim V$个向量. 于是, W的线性无关子集中的向量个数存在最大值n, 并且$n \leq \dim V$. 取W的含有n个向量的线性无关子集S. 我们断言$\mathrm{span} S = W$. 事实上, 如果$\mathrm{span} S$是W的真子空间, 则可以取$\beta \in W \setminus \mathrm{span} S$, 由引理2.5, $S \cup \{\beta\}$是W的线性无关子集, 并且$|S \cup \{\beta\}| = n + 1$, 矛盾. 断言推出$S$是$W$的基, 并且$\dim W = |S| = n \leq \dim V$. 这就证明了(1). 如果还有$\dim W = \dim V$, 则$|S| = \dim V$. 由于$S$线性无关, 由推论2.6(1), S是V的基, 从而$W = \mathrm{span} S = V$. 因此(2)成立. □

定理 2.8. 设W_1, W_2是线性空间V的有限维子空间, 则$W_1 + W_2$是有限维的, 并且

$$\dim W_1 + \dim W_2 = \dim(W_1 + W_2) + \dim(W_1 \cap W_2).$$

证明. 因为W_1, W_2有限维, 所以$W_1 \cap W_2$有限维. 取$W_1 \cap W_2$的基$\{\alpha_1, \ldots, \alpha_d\}$, 并分别扩充为$W_1$的基$\{\alpha_1, \ldots, \alpha_d, \beta_1, \ldots, \beta_m\}$和$W_2$的基$\{\alpha_1, \ldots, \alpha_d, \gamma_1, \ldots, \gamma_n\}$. 我们断言:

$$\{\alpha_1, \ldots, \alpha_d, \beta_1, \ldots, \beta_m, \gamma_1, \ldots, \gamma_n\}$$

是$W_1 + W_2$的基. 验证如下:

线性无关: 设$\sum x_i\alpha_i + \sum y_j\beta_j + \sum z_k\gamma_k = 0$. 注意到$\sum x_i\alpha_i + \sum y_j\beta_j \in W_1$, 所以$\sum z_k\gamma_k \in W_1$. 另一方面, 还有$\sum z_k\gamma_k \in W_2$. 所以$\sum z_k\gamma_k \in W_1 \cap W_2$. 于是存在$c_i$使$\sum z_k\gamma_k = \sum c_i\alpha_i$. 但是$\{\alpha_i, \gamma_k\}$线性无关. 所以$z_k = c_i = 0$. 这进一步推出$\sum x_i\alpha_i + \sum y_j\beta_j = 0$. 由于$\{\alpha_i, \beta_j\}$线性无关, 所以$x_i = y_j = 0$. 因此$x_i = y_j = z_k = 0$.

$\mathrm{span}\{\alpha_i, \beta_j, \gamma_k\} = W_1 + W_2$: "$\subset$": $\mathrm{span}\{\alpha_i, \beta_j, \gamma_k\}$中的任意向量$\alpha$形如

$$\alpha = \sum x_i\alpha_i + \sum y_j\beta_j + \sum z_k\gamma_k.$$

由于$\sum x_i\alpha_i + \sum y_j\beta_j \in W_1$, $\sum z_k\gamma_k \in W_2$, 所以$\alpha \in W_1 + W_2$.

"$\supset$": $W_1 + W_2$中的任意向量α形如$\alpha = \beta + \gamma$, 其中$\beta \in W_1$, $\gamma \in W_2$. 而

$$W_1 = \mathrm{span}\{\alpha_i, \beta_j\} \subset \mathrm{span}\{\alpha_i, \beta_j, \gamma_k\}, \qquad W_2 = \mathrm{span}\{\alpha_i, \gamma_k\} \subset \mathrm{span}\{\alpha_i, \beta_j, \gamma_k\}.$$

所以$\beta, \gamma \in \mathrm{span}\{\alpha_i, \beta_j, \gamma_k\}$. 因此$\alpha = \beta + \gamma \in \mathrm{span}\{\alpha_i, \beta_j, \gamma_k\}$.

由断言即得$W_1 + W_2$是有限维的, 并且

$$\dim(W_1 + W_2) = d + m + n = (d + m) + (d + n) - d = \dim W_1 + \dim W_2 - \dim(W_1 \cap W_2).$$

□

§2.4 坐标

定义 2.1. 设V是有限维F-线性空间. 由V中向量构成的n元有序组$(\alpha_1, \ldots, \alpha_n)$称为$V$的**有序基**, 如果$\alpha_1, \ldots, \alpha_n$互不相同, 并且集合$\{\alpha_1, \ldots, \alpha_n\}$是$V$的基.

为了与书上记号一致, 当无歧义时, 我们对有序基也采用集合的记号, 即称$\mathcal{B} = \{\alpha_1, \ldots, \alpha_n\}$为$V$的有序基.

命题 2.1. 设V是有限维线性空间, $\mathcal{B} = \{\alpha_1, \ldots, \alpha_n\}$是$V$的有序基. 则任意$\alpha \in V$表为线性组合$\alpha = \sum_{i=1}^n x_i\alpha_i$的方式唯一, 即

$$\alpha = \sum_{i=1}^n x_i\alpha_i = \sum_{i=1}^n x_i'\alpha_i \qquad \Longrightarrow \qquad x_i = x_i'.$$

证明. 若$\sum_{i=1}^n x_i\alpha_i = \sum_{i=1}^n x_i'\alpha_i$, 则$\sum_{i=1}^n (x_i - x_i')\alpha_i = 0$. 由于$\{\alpha_1, \ldots, \alpha_n\}$线性无关, 所以$x_i - x_i' = 0$. □

定义 2.2. 在命题的条件下, 系数x_i称为向量α关于有序基$\mathcal{B}$的**第i个坐标**, n-维向量$(x_1, \ldots, x_n) \in F^n$称为$\alpha$关于有序基$\mathcal{B}$的**坐标**.

注 2.1. 由于某种原因, 以后我们会把列向量$\begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$称为$\alpha$的坐标.

坐标的概念给出了一个映射

$$\Gamma_{\mathcal{B}} : V \to F^n, \qquad \Gamma_{\mathcal{B}}(\alpha) = (x_1, \ldots, x_n),$$

其中$\alpha = \sum_{i=1}^n x_i\alpha_i$. 它有下面的性质:
单: 设$\Gamma_{\mathcal{B}}(\alpha) = \Gamma_{\mathcal{B}}(\beta)$. 如果$\alpha = \sum_{i=1}^n x_i\alpha_i$, $\beta = \sum_{i=1}^n y_i\alpha_i$, 则$(x_1, \ldots, x_n) = \Gamma_{\mathcal{B}}(\alpha) = \Gamma_{\mathcal{B}}(\beta) = (y_1, \ldots, y_n)$. 这说明$x_i = y_i$. 因此$\alpha = \beta$.
满: 对任意$(x_1, \ldots, x_n) \in F^n$, 取$\alpha = \sum_{i=1}^n x_i\alpha_i$, 则$\Gamma_{\mathcal{B}}(\alpha) = (x_1, \ldots, x_n)$.
保持线性结构: $\Gamma_{\mathcal{B}}(\alpha) + \Gamma_{\mathcal{B}}(\beta) = \Gamma_{\mathcal{B}}(\alpha + \beta)$, $\Gamma_{\mathcal{B}}(c\alpha) = c\Gamma_{\mathcal{B}}(\alpha)$.

这样的映射$\Gamma_{\mathcal{B}}$称为线性同构. 此时我们称V与F^n同构, 记为$V \cong F^n$. 同构的线性空间有很多相同的性质.

例 2.1. $V = F^n$, $\mathcal{B} = \{\epsilon_1, \ldots, \epsilon_n\}$. 则对$\alpha = (x_1, \ldots, x_n)$有$\alpha = \sum_{i=1}^n x_i\epsilon_i$, 即$\alpha$的第$i$个坐标为$x_i$. 这说明$\Gamma_{\mathcal{B}}(\alpha) = \alpha$, 即$\Gamma_{\mathcal{B}} = \mathrm{id}$. □

第五章　行列式

域F上方阵的行列式是F中的元素, 满足一些很好的性质. 例如, 方阵可逆的充分必要条件是它的行列式非零, 方阵乘积的行列式等于行列式的乘积等. 在对矩阵的研究中, 行列式起到了非常重要的作用.

这里我们采用线性映射的观点. 通过与列向量或行向量做乘法, 可以把矩阵视为线性映射. 我们首先利用多重交错线性函数的性质, 对任意有限维线性空间到自身的线性映射定义它的行列式, 然后把方阵的行列式定义为相应的线性映射的行列式. 方阵行列式的一些基本性质也将由线性映射行列式的性质导出.

§5.1　对称群

集合$\{1,\dots,n\}$到自身的可逆映射称为n**次置换**(permutation of degree n). 所有n次置换在映射复合下构成一个群, 称为n**次对称群**(symmetric group of degree n), 记为S_n. 容易看出, $|S_n| = n!$. 如果置换$\sigma \in S_n$互换$\{1,\dots,n\}$中的某两个数字, 而保持其他数字不动, 则称σ为**对换**(transposition). 对于置换$\sigma \in S_n$, 我们记

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

如果σ是互换数字s和t的对换, 则记$\sigma = (s,t)$.

例 5.1. $S_1 = \{\mathrm{id}\}$, $S_2 = \{\mathrm{id}, (1,2)\}$,

$$S_3 = \left\{ \mathrm{id}, (1,2), (2,3), (3,1), \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix} \right\},$$

这里id是相应集合上的恒同置换.

我们称S_n中的$n-1$个对换

$$\{(i,i+1) \mid 1 \le i \le n-1\}$$

为**相邻数字的对换**(transposition of adjacent digits).

命题 5.1. 置换群S_n中相邻数字的对换生成S_n. 也就是说, 对任意$\sigma \in S_n$, 存在相邻数字的对换$\sigma_1,\dots,\sigma_k$使得$\sigma = \sigma_1 \cdots \sigma_k$.

证明. 用归纳法. 命题在$n=1$时无需证明. 假设$n \ge 2$, 并且命题对$n-1$成立. 设$\sigma \in S_n$. 我们分两种情形证明σ可以表示为相邻数字对换的乘积.

情形1. 设$\sigma(n) = n$. 则σ限制在$\{1,\dots,n-1\}$上是S_{n-1}中的元素, 从而归纳假设保证了σ可以表示为相邻数字对换的乘积.

情形2. 设$\sigma(n) < n$. 记$\tau_i = (i,i+1)$. 则$\tau_{n-1}\cdots\tau_{\sigma(n)}\sigma(n) = n$. 由情形1, 存在相邻数字的对换$\sigma_1,\dots,\sigma_k$使得$\tau_{n-1}\cdots\tau_{\sigma(n)}\sigma = \sigma_1\cdots\sigma_k$, 即$\sigma = \tau_{\sigma(n)}\cdots\tau_{n-1}\sigma_1\cdots\sigma_k$. □

对于$\sigma \in S_n$, 我们称

$$\ell(\sigma) = \#\{(i,j) \mid i,j \in \{1,\dots,n\}, i<j, \sigma(i) > \sigma(j)\}$$

为σ的**逆序数**(number of inversions), 并称$\mathrm{sgn}(\sigma) = (-1)^{\ell(\sigma)}$为$\sigma$的**符号**(sign). 如果$\mathrm{sgn}(\sigma) = 1$, 则

称σ为**偶置换**(even permutation); 如果$\mathrm{sgn}(\sigma)=-1$, 则称σ为**奇置换**(odd permutation).

例 5.2. 设$n\geq 2$, I是$\{1,\dots,n\}$的子集, $1\leq |I|<n$. 我们定义置换$\sigma_I\in S_n$如下: 记$k=|I|$, $\Sigma_I=\sum_{i\in I} i$, $I^c=\{1,\dots,n\}\setminus I$. 如果

$$I=\{i_1,\dots,i_k\},\quad i_1<\cdots<i_k,$$
$$I^c=\{i'_1,\dots,i'_{n-k}\},\quad i'_1<\cdots<i'_{n-k},$$

则定义

$$\sigma_I=\begin{pmatrix}1 & \cdots & k & k+1 & \cdots & n\\ i_1 & \cdots & i_k & i'_1 & \cdots & i'_{n-k}\end{pmatrix}. \tag{5.1}$$

置换σ_I的逆序数为

$$\begin{aligned}\ell(\sigma_I)&=\#\{(r,s)\mid i_r>i'_s\}=\sum_{r=1}^{k}\#\{s\mid i_r>i'_s\}\\&=\sum_{r=1}^{k}\#(\{1,\dots,i_r\}\cap I^c)=\sum_{r=1}^{k}(i_r-r)=\Sigma_I-\frac{k(k+1)}{2}.\end{aligned}$$

因此

$$\mathrm{sgn}(\sigma_I)=(-1)^{\Sigma_I-\frac{k(k+1)}{2}}. \tag{5.2}$$

置换的符号可以用下面的表达式来表示.

命题 5.2. 对任意$\sigma\in S_n$, 有

$$\mathrm{sgn}(\sigma)=\prod_{1\leq i<j\leq n}\frac{\sigma(j)-\sigma(i)}{j-i}. \tag{5.3}$$

更一般地, 如果$(d_1,\dots,d_n)$是数字$\{1,\dots,n\}$的排列, 则

$$\mathrm{sgn}(\sigma)=\prod_{1\leq i<j\leq n}\frac{\sigma(d_j)-\sigma(d_i)}{d_j-d_i}. \tag{5.4}$$

证明. 容易看出, (5.3)式右边的乘积中有$\ell(\sigma)$项取负号. 因此该乘积与$\mathrm{sgn}(\sigma)$ 同号. 而该乘积的绝对值等于1. 因此(5.3)成立. 另一方面, 在不计次序的意义下, (5.4)式与(5.3)式右边乘积中的$\frac{n(n-1)}{2}$项相同. 因此(5.4) 也成立. □

我们会多次用到下面的性质.

命题 5.3. (1) 映射$\mathrm{sgn}: S_n\to\{1,-1\}$是群同态, 即对任意$\sigma,\tau\in S_n$有

$$\mathrm{sgn}(\sigma\tau)=\mathrm{sgn}(\sigma)\mathrm{sgn}(\tau).$$

(2) 如果$\sigma\in S_n$是k个对换的乘积, 则$\mathrm{sgn}(\sigma)=(-1)^k$. 特别地, 对换是奇置换.

证明. (1) 由于$(\tau(1),\dots,\tau(n))$是数字$\{1,\dots,n\}$的排列, 由命题5.2,

$$\mathrm{sgn}(\sigma)=\prod_{1\leq i<j\leq n}\frac{\sigma(\tau(j))-\sigma(\tau(i))}{\tau(j)-\tau(i)}.$$

因此

$$\mathrm{sgn}(\sigma\tau)=\prod_{1\leq i<j\leq n}\frac{\sigma(\tau(j))-\sigma(\tau(i))}{\tau(j)-\tau(i)}\prod_{1\leq i<j\leq n}\frac{\tau(j)-\tau(i)}{j-i}=\mathrm{sgn}(\sigma)\mathrm{sgn}(\tau).$$

(2) 首先证明如果$\sigma\in S_n$是对换, 则$\mathrm{sgn}(\sigma)=-1$. 设$\sigma=(s,t)$, $s<t$. 则

$$\begin{aligned}&\{(i,j)\mid i,j\in\{1,\dots,n\},i<j,\sigma(i)>\sigma(j)\}\\&=\{(s,t),(s,s+1),\dots,(s,t-1),(s+1,t),\dots,(t-1,t)\}.\end{aligned}$$

从而$\ell(\sigma)=2(t-s-1)+1$是奇数, 因此$\mathrm{sgn}(\sigma)=-1$. 现在设$\sigma=\sigma_1\cdots\sigma_k$, 其中$\sigma_i$是对换. 由(1), 有$\mathrm{sgn}(\sigma)=\mathrm{sgn}(\sigma_1)\cdots\mathrm{sgn}(\sigma_k)=(-1)^k$. □

推论 5.4. 所有n次偶置换构成的集合A_n是S_n的子群, 并且当$n\geq 2$时有$|A_n|=\frac{1}{2}n!$.

证明. 为了证明A_n是S_n的子群, 我们需要说明: (1)S_n中的恒同置换id是偶置换; (2) 如果$\mathrm{sgn}(\sigma)=\mathrm{sgn}(\tau)=1$, 则$\mathrm{sgn}(\sigma\tau)=1$; (3) 如果$\mathrm{sgn}(\sigma)=1$, 则$\mathrm{sgn}(\sigma^{-1})=1$. (1)是显然的, (2)和(3)是命题5.3(1)的直接推论.

假设$n\geq 2$. 由命题5.3(2), $A_n\neq S_n$. 取定$\tau\in S_n\setminus A_n$. 由命题5.3(1), 可以定义映射

$$\rho_1: A_n\to S_n\setminus A_n,\quad \rho_1(\sigma)=\sigma\tau,$$
$$\rho_2: S_n\setminus A_n\to A_n,\quad \rho_2(\sigma)=\sigma\tau^{-1}.$$

注意到$\rho_1\circ\rho_2$和$\rho_2\circ\rho_1$是恒同映射. 因此ρ_1与ρ_2可逆. 由于集合A_n与$S_n\setminus A_n$之间存在可逆映射, 所以$|A_n|=|S_n\setminus A_n|=\frac{1}{2}|S_n|=\frac{1}{2}n!$. □

群A_n称为n**次交错群**(alternating group of degree n).

例 5.3. $A_1=\{\mathrm{id}\}$, $A_2=\{\mathrm{id}\}$,

$$A_3=\left\{\mathrm{id},\begin{pmatrix}1&2&3\\2&3&1\end{pmatrix},\begin{pmatrix}1&2&3\\3&1&2\end{pmatrix}\right\}.$$

习题 5.1.

1. 列出A_4中的所有元素.
2. 求所有正整数n, 使得置换$\begin{pmatrix}1&\cdots&n\\n&\cdots&1\end{pmatrix}\in A_n$.
3. 设$\{\epsilon_1,\ldots,\epsilon_n\}$是$F^{n\times 1}$的标准基. 对$\sigma\in S_n$, 记$R(\sigma)\in F^{n\times n}$为第$j$列等于$\epsilon_{\sigma(j)}$ 的(可逆)矩阵. 证明映射$R: S_n\to \mathrm{GL}_n(F)$是群同态, 即对任意$\sigma,\tau\in S_n$有$R(\sigma\tau)=R(\sigma)R(\tau)$.
4. 设$\{\delta_1,\ldots,\delta_n\}$是$F^n$的标准基. 对$\sigma\in S_n$, 记$R'(\sigma)\in F^{n\times n}$为第$i$行等于$\delta_{\sigma(i)}$的矩阵. 等式$R'(\sigma\tau)=R'(\sigma)R'(\tau)$是否成立?

§5.2 多重线性函数

设V是域F上的线性空间, r是正整数.

定义 5.1. 映射$L: V^r\to F$称为V上的r**重线性函数**(r-linear function)或r**重线性形式**(r-linear form), 如果对任意指标$1\leq i\leq r$和任意给定的向量$\alpha_1,\ldots,\alpha_{i-1},\alpha_{i+1},\ldots,\alpha_r\in V$, 映射

$$V\to F,\quad \alpha_i\mapsto L(\alpha_1,\ldots,\alpha_r)$$

是线性函数.

我们把V上的所有r重线性函数的集合记为$(V^*)^{\otimes r}$或$\bigotimes^r(V^*)$(用$\bigotimes^r(V^*)$, 书上记为$\mathrm{M}^r(V)$), 并在它上面定义加法和纯量乘法如下:

$$(L_1+L_2)(\alpha_1,\ldots,\alpha_r)=L_1(\alpha_1,\ldots,\alpha_r)+L_2(\alpha_1,\ldots,\alpha_r),\quad \forall L_1,L_2\in(V^*)^{\otimes r},$$
$$(cL)(\alpha_1,\ldots,\alpha_r)=cL(\alpha_1,\ldots,\alpha_r),\quad \forall c\in F, L\in(V^*)^{\otimes r}.$$

容易看出, $(V^*)^{\otimes r}$是线性空间. 注意$(V^*)^{\otimes 1}=V^*$.

例 5.4. 设$f_1,\ldots,f_r\in V^*$, 并定义$f_1\otimes\cdots\otimes f_r:V^r\to F$为

$$f_1\otimes\cdots\otimes f_r(\alpha_1,\ldots,\alpha_r)=f_1(\alpha_1)\cdots f_r(\alpha_r).$$

则$f_1\otimes\cdots\otimes f_r\in(V^*)^{\otimes r}$.

例 5.5. 线性空间上的2重线性函数称为**双线性函数**(bilinear function)或**双线性形式**(bilinear form). 设$A\in F^{n\times n}$. 则映射

$$L:(F^{n\times 1})^2\to F,\quad L(X,Y)=Y^tAX$$

是$F^{n\times 1}$上的双线性函数.

为了定义线性映射的行列式, 我们需要下面的概念.

定义 5.2. 多重线性函数$L\in(V^*)^{\otimes r}$称为**交错**(alternating)的, 如果某两个变量相同时L取值是0, 即对任意不同的$s,t\in\{1,\ldots,r\}$, 有

$$\alpha_s=\alpha_t\quad\Longrightarrow\quad L(\alpha_1,\ldots,\alpha_r)=0.$$

我们把V上的r重交错线性函数的集合记为$\Lambda^r(V^*)$(书上记为$\Lambda^r(V)$). 容易看出, 它是$(V^*)^{\otimes r}$的子空间. 我们约定$\Lambda^1(V^*)=V^*$.

例 5.6. 设$f_1,f_2\in V^*$. 定义

$$f_1\wedge f_2=f_1\otimes f_2-f_2\otimes f_1,$$

即

$$f_1\wedge f_2(\alpha_1,\alpha_2)=f_1(\alpha_1)f_2(\alpha_2)-f_1(\alpha_2)f_2(\alpha_1).$$

则$f_1\wedge f_2\in\Lambda^2(V^*)$. 注意到

$$f_1\wedge f_2=-f_2\wedge f_1.$$

例 5.7. 设$\{f_1,\ldots,f_{2n}\}$是F^{2n}的标准基的对偶基. F^{2n}上的交错双线性函数

$$\omega_0=\sum_{i=1}^n f_i\wedge f_{n+i}$$

称为**标准辛形式**(standard symplectic form). 如果$\alpha=(x_1,\ldots,x_{2n})$, $\beta=(y_1,\ldots,y_{2n})$, 则

$$\omega_0(\alpha,\beta)=\sum_{i=1}^n(x_iy_{n+i}-x_{n+i}y_i).$$

命题 5.5. 设$r\geq 2$.

(1) 设$L\in(V^*)^{\otimes r}$. 假设某两个相邻变量相同时L取值是0, 即对任意$1\leq i\leq r-1$有

$$\alpha_i=\alpha_{i+1}\quad\Longrightarrow\quad L(\alpha_1,\ldots,\alpha_r)=0.$$

则$L\in\Lambda^r(V^*)$.

(2) 设$L\in\Lambda^r(V^*)$. 则对任意$\sigma\in S_r$, 有

$$L(\alpha_{\sigma(1)},\ldots,\alpha_{\sigma(r)})=\operatorname{sgn}(\sigma)L(\alpha_1,\ldots,\alpha_r),\quad\forall\alpha_1,\ldots,\alpha_r\in V. \tag{5.5}$$

证明. 我们把证明分为以下几步.

第一步. 假设$L\in(V^*)^{\otimes r}$满足(1)的条件, 即当某两个相邻变量相同时L取值是0. 我们证明(5.5)当σ是相邻数字的对换时成立. 设$\sigma=(i,i+1)$. 对于取定的$r-2$个向量α_j, $j\in\{1,\ldots,r\}\setminus\{i,i+1\}$, 记

$$L'(\alpha_i,\alpha_{i+1})=L(\alpha_1,\ldots,\alpha_r).$$

则(1)的条件推出

$$\begin{aligned}0 &= L'(\alpha_i+\alpha_{i+1},\alpha_i+\alpha_{i+1})\\ &= L'(\alpha_i,\alpha_i)+L'(\alpha_{i+1},\alpha_{i+1})+L'(\alpha_i,\alpha_{i+1})+L'(\alpha_{i+1},\alpha_i)\\ &= L'(\alpha_i,\alpha_{i+1})+L'(\alpha_{i+1},\alpha_i),\end{aligned}$$

即$L'(\alpha_{i+1},\alpha_i)=-L'(\alpha_i,\alpha_{i+1})$. 由命题5.3(2), $\mathrm{sgn}(\sigma)=-1$. 因此(5.5)对$\sigma=(i,i+1)$成立.

第二步. 假设$L\in(V^*)^{\otimes r}$满足(1)的条件. 我们证明(5.5)对于一般的$\sigma\in S_n$也成立. 由命题5.1, σ可以表示为相邻数字的对换的乘积$\sigma=\sigma_1\cdots\sigma_k$. 对这些对换应用第一步的结果, 得到

$$\begin{aligned}L(\alpha_{\sigma(1)},\ldots,\alpha_{\sigma(r)}) &= L(\alpha_{\sigma_1\cdots\sigma_k(1)},\ldots,\alpha_{\sigma_1\cdots\sigma_k(r)})\\ &= -L(\alpha_{\sigma_2\cdots\sigma_k(1)},\ldots,\alpha_{\sigma_2\cdots\sigma_k(r)})\\ &= \cdots\cdots\\ &= (-1)^k L(\alpha_1,\ldots,\alpha_r).\end{aligned}$$

由命题5.3(2), $\mathrm{sgn}(\sigma)=(-1)^k$. 因此(5.5)对一般的$\sigma\in S_n$也成立. 注意到如果$L\in\Lambda^r(V^*)$, 则(1)的条件成立. 因此我们已经完成了(2)的证明. □

第三步. 最后我们证明(1). 假设$L\in(V^*)^{\otimes r}$满足(1)的条件, $\alpha_1,\ldots,\alpha_r\in V$. 我们需要证明: 对任意不同的$s,t\in\{1,\ldots,r\}$, 如果$\alpha_s=\alpha_t$, 则$L(\alpha_1,\ldots,\alpha_r)=0$. 不妨设$s<t$. 如果$s+1=t$, 则(1)的条件已经保证$L(\alpha_1,\ldots,\alpha_r)=0$. 假设$s+1<t$. 对于对换$\sigma=(s+1,t)$应用第二步的结果, 得到$L(\alpha_1,\ldots,\alpha_r)=-L(\alpha_{\sigma(1)},\ldots,\alpha_{\sigma(r)})$. 另一方面, 在排列$(\sigma(1),\ldots,\sigma(n))$中, 数字$s$与$t$相邻. 所以(1)的条件推出$L(\alpha_{\sigma(1)},\ldots,\alpha_{\sigma(r)})=0$. 因此$L(\alpha_1,\ldots,\alpha_r)=0$. 这就完成了证明. □

现在设V是有限维的, $\dim V=n\geq 1$. 我们考察V上的n重交错线性函数的可能形式. 设$\{\alpha_1,\ldots,\alpha_n\}$是$V$的基, $\{f_1,\ldots,f_n\}$是它在V^*中的对偶基. 则对任意的$L\in(V^*)^{\otimes n}$和$\beta_1,\ldots,\beta_n\in V$, 由于$\beta_i=\sum_{j=1}^n f_j(\beta_i)\alpha_j$, 所以

$$\begin{aligned}L(\beta_1,\ldots,\beta_n) &= L\left(\sum_{j_1=1}^n f_{j_1}(\beta_1)\alpha_{j_1},\ldots,\sum_{j_n=1}^n f_{j_n}(\beta_n)\alpha_{j_n}\right)\\ &= \sum_{1\leq j_1,\ldots,j_n\leq n} f_{j_1}(\beta_1)\cdots f_{j_n}(\beta_n)L(\alpha_{j_1},\ldots,\alpha_{j_n})\\ &= \sum_{1\leq j_1,\ldots,j_n\leq n} L(\alpha_{j_1},\ldots,\alpha_{j_n})f_{j_1}\otimes\cdots\otimes f_{j_n}(\beta_1,\ldots,\beta_n),\end{aligned}$$

即

$$L=\sum_{1\leq j_1,\ldots,j_n\leq n} L(\alpha_{j_1},\ldots,\alpha_{j_n})f_{j_1}\otimes\cdots\otimes f_{j_n}.$$

如果L是交错的, 则只有当$j_1,\ldots,j_n$互不相同时, $L(\alpha_{j_1},\ldots,\alpha_{j_n})$才可能非零. 此时, 存在唯一的$\sigma\in S_n$使得$j_i=\sigma(i)$. 因此, 由命题5.5(2),

$$\begin{aligned}L &= \sum_{\sigma\in S_n} L(\alpha_{\sigma(1)},\ldots,\alpha_{\sigma(n)})f_{\sigma(1)}\otimes\cdots\otimes f_{\sigma(n)}\\ &= L(\alpha_1,\ldots,\alpha_n)\sum_{\sigma\in S_n}\mathrm{sgn}(\sigma)f_{\sigma(1)}\otimes\cdots\otimes f_{\sigma(n)}.\end{aligned}$$

如果我们定义n重线性函数

$$f_1\wedge\cdots\wedge f_n=\sum_{\sigma\in S_n}\mathrm{sgn}(\sigma)f_{\sigma(1)}\otimes\cdots\otimes f_{\sigma(n)}, \tag{5.6}$$

即

$$f_1\wedge\cdots\wedge f_n(\beta_1,\dots,\beta_n)=\sum_{\sigma\in S_n}\operatorname{sgn}(\sigma)f_{\sigma(1)}(\beta_1)\cdots f_{\sigma(n)}(\beta_n), \tag{5.7}$$

上面的讨论说明:

引理 5.6. 设$L\in\Lambda^n(V^*)$, $\{\alpha_1,\dots,\alpha_n\}$是$V$的基, $\{f_1,\dots,f_n\}$是它的对偶基. 则

$$L=L(\alpha_1,\dots,\alpha_n)f_1\wedge\cdots\wedge f_n.$$

接下来我们证明:

引理 5.7. 设$\{\alpha_1,\dots,\alpha_n\}$是$V$的基, $\{f_1,\dots,f_n\}$是其对偶基. 则

(1) $f_1\wedge\cdots\wedge f_n$是交错的.

(2) $f_1\wedge\cdots\wedge f_n(\alpha_1,\dots,\alpha_n)=1$. 特别地, $f_1\wedge\cdots\wedge f_n\neq 0$.

证明. (1) 当$n=1$时无需证明. 假设$n\geq 2$. 设$\beta_1,\dots,\beta_n\in V$, $\beta_s=\beta_t$, 这里$1\leq s<t\leq n$. 我们希望证明$f_1\wedge\cdots\wedge f_n(\beta_1,\dots,\beta_n)=0$. 考虑对换$\tau=(s,t)$. 由推论5.4的证明可知

$$S_n\setminus A_n=\{\sigma\tau\mid\sigma\in A_n\}.$$

从而

$$\begin{aligned}
&f_1\wedge\cdots\wedge f_n(\beta_1,\dots,\beta_n)\\
=&\sum_{\sigma\in A_n}f_{\sigma(1)}\otimes\cdots\otimes f_{\sigma(n)}(\beta_1,\dots,\beta_n)-\sum_{\sigma\in A_n}f_{\sigma\tau(1)}\otimes\cdots\otimes f_{\sigma\tau(n)}(\beta_1,\dots,\beta_n)\\
=&\sum_{\sigma\in A_n}\left(f_{\sigma(1)}(\beta_1)\cdots f_{\sigma(n)}(\beta_n)-f_{\sigma\tau(1)}(\beta_1)\cdots f_{\sigma\tau(n)}(\beta_n)\right)\\
=&\sum_{\sigma\in A_n}\left(\prod_{i\in\{1,\dots,n\}\setminus\{s,t\}}f_{\sigma(i)}(\beta_i)\right)\left(f_{\sigma(s)}(\beta_s)f_{\sigma(t)}(\beta_t)-f_{\sigma(t)}(\beta_s)f_{\sigma(s)}(\beta_t)\right)=0.
\end{aligned}$$

因此$f_1\wedge\cdots\wedge f_n$是交错的.

(2) 由(5.7)即得$f_1\wedge\cdots\wedge f_n(\alpha_1,\dots,\alpha_n)=f_1(\alpha_1)\cdots f_n(\alpha_n)=1$. □

由上面两个引理, 我们很容易得到:

定理 5.8. 设V是域F上的有限维线性空间, $\dim V=n\geq 1$. 则$\dim\Lambda^n(V^*)=1$.

证明. 取定V的一组基和它在V^*中的对偶基$\{f_1,\dots,f_n\}$. 由引理5.7, $f_1\wedge\cdots\wedge f_n$是$\Lambda^n(V^*)$中的非零元素. 再由引理5.6, 可知$\{f_1\wedge\cdots\wedge f_n\}$是$\Lambda^n(V^*)$的基. 因此$\dim\Lambda^n(V^*)=1$. □

一个n维线性空间V上的非零n重交错线性函数可以用来判断V中的n个向量是否构成一组基.

推论 5.9. 设$\dim V=n\geq 1$, $L\in\Lambda^n(V^*)\setminus\{0\}$, $\alpha_1,\dots,\alpha_n\in V$. 则$\{\alpha_1,\dots,\alpha_n\}$是$V$的基的充分必要条件是$L(\alpha_1,\dots,\alpha_n)\neq 0$.

证明. 设$\{\alpha_1,\dots,\alpha_n\}$是基, $\{f_1,\dots,f_n\}$是其对偶基. 由引理5.7(2), $f_1\wedge\cdots\wedge f_n(\alpha_1,\dots,\alpha_n)=1$. 再由定理5.8可知, 存在$c\in F\setminus\{0\}$使得$L=cf_1\wedge\cdots\wedge f_n$. 因此

$$L(\alpha_1,\dots,\alpha_n)=cf_1\wedge\cdots\wedge f_n(\alpha_1,\dots,\alpha_n)=c\neq 0.$$

反过来, 假设$\{\alpha_1,\dots,\alpha_n\}$不是基. 取$\operatorname{span}\{\alpha_1,\dots,\alpha_n\}$的基$\{\beta_1,\dots,\beta_r\}$, $r<n$, 并扩充为V的基$\{\beta_1,\dots,\beta_n\}$. 设$\{g_1,\dots,g_n\}$是$\{\beta_1,\dots,\beta_n\}$在$V^*$中的对偶基. 由于$g_n(\beta_1)=\cdots=g_n(\beta_r)=0$,

所以$g_n(\alpha_1)=\cdots=g_n(\alpha_n)=0$. 由引理5.7(2)和定理5.8, 存在$c'\in F$使得$L=c'g_1\wedge\cdots\wedge g_n$. 因此

$$L(\alpha_1,\ldots,\alpha_n)=c'g_1\wedge\cdots\wedge g_n(\alpha_1,\ldots,\alpha_n)=c'\sum_{\sigma\in S_n}\operatorname{sgn}(\sigma)g_{\sigma(1)}(\alpha_1)\cdots g_{\sigma(n)}(\alpha_n)=0.$$

□

注 5.1. 当V是n维实线性空间时, V上的非零n重交错线性函数可以视为广义平行多面体的有向体积函数. 对于向量$\alpha_1,\ldots,\alpha_n\in V$, 考虑$V$中的**广义平行多面体**(parallelotope)

$$P=P(\alpha_1,\ldots,\alpha_n)=\left\{\sum_{i=1}^{n}c_i\alpha_i \mid 0\le c_i\le 1\right\}.$$

给定$L\in\Lambda^n(V^*)\setminus\{0\}$, 我们把$|L(\alpha_1,\ldots,\alpha_n)|$视为$P$的体积. 注意到$L$的交错性意味着, 如果$P$的某两个顶点$\alpha_s$和$\alpha_t$重合, 则$P$的体积是0. 这是对体积的定义的合理要求. 更一般地, 推论5.9表明, P的体积是0等价于$\{\alpha_1,\ldots,\alpha_n\}$不是基, 即$P$落在$V$的某个真子空间中. 我们称这种情况是退化的. 当$P$非退化时, 在它上面有两个定向. 定向有几种等价的定义方式. 我们采用如下定义: P的顶点$\alpha_1,\ldots,\alpha_n$的两个排列$(\alpha_{i_1},\ldots,\alpha_{i_n})$和$(\alpha_{j_1},\ldots,\alpha_{j_n})$称为是等价的, 如果满足$\sigma(i_k)=j_k$的置换$\sigma\in S_n$是偶置换. 于是, 这$n$个顶点的所有排列被划分为两个等价类. 每个等价类称为P 的一个**定向**(orientation), 每个排列所在的等价类称为由这个排列决定的定向. 这时, 我们把$L(\alpha_1,\ldots,\alpha_n)$视为在$P$上选取由排列$(\alpha_1,\ldots,\alpha_n)$决定的定向后它的有向体积. 有向体积可以取负值, 改变P的定向导致它的有向体积反号. 需要指出的是, 有向体积依赖于L的选取. 如果在V上没有附加其他结构(例如内积和V自身的定向等), 则L没有占特殊地位的选取方式. 而定理5.8说明, 有向体积的赋予方式在差一个非零因子的意义下是唯一的.

习题 5.2.

1. 集合$\{f_1\otimes f_2 \mid f_1,f_2\in V^*\}$是否为$(V^*)^{\otimes 2}$的子空间? 试对$V=F^2$刻画这个集合.
2. 设$\dim V=n$. 证明$\dim(V^*)^{\otimes 2}=n^2$.
3. 证明$F^{n\times 1}$上的双线性函数总是例5.5的形式.
4. 线性空间V上的双线性函数L称为**非退化的**(nondegenerate), 如果对任意$\alpha\in V\setminus\{0\}$, 存在$\beta\in V$使得$L(\alpha,\beta)\ne 0$. 证明例5.5中的双线性函数非退化的充分必要条件是矩阵A可逆.
5. 证明F^{2n}上的标准辛形式是非退化的.
6. 设V是有限维的, L是V上的非退化双线性函数. 对于子空间$W\subset V$, 定义
$$W^{\perp}=\{\alpha\in V \mid L(\alpha,\beta)=0,\forall\beta\in W\}.$$
证明$W^{\perp}$是V的子空间, 并且$\dim W+\dim W^{\perp}=\dim V$.
7. 多重线性函数$L\in(V^*)^{\otimes r}$称为是**反对称**(anti-symmetric)或**斜对称**(skew-symmetric)的, 如果交换两个变量使L的取值反号, 即(5.5)式对任意S_n中的对换σ成立. 证明: 如果$\operatorname{char}F\ne 2$, 则$L$反对称的充分必要条件是它是交错的.
8. 设$\operatorname{char}F=2$. 说明在F^n上存在反对称但不交错的双线性函数.
9. 设$\dim V=n$.
   (1) 证明$\dim\Lambda^2(V^*)=\frac{1}{2}n(n-1)$.
   (2) 设$r>n$. 证明$\Lambda^r(V^*)=\{0\}$.

§5.3 线性映射的行列式

设V和W是域F上的线性空间, $T \in \mathrm{L}(V, W)$, r是正整数. 与转置映射$T^t : W^* \to V^*$类似, 我们定义映射$\Lambda^r(T^t) : \Lambda^r(W^*) \to \Lambda^r(V^*)$为

$$\Lambda^r(T^t)(L)(\alpha_1, \ldots, \alpha_r) = L(T\alpha_1, \ldots, T\alpha_r), \quad \forall L \in \Lambda^r(W^*), \alpha_1, \ldots, \alpha_r \in V.$$

容易看出, $\Lambda^r(T^t)(L)$确实属于$\Lambda^r(V^*)$, 并且$\Lambda^r(T^t)$是线性映射. 注意$\Lambda^1(T^t) = T^t$.

命题 5.10. 设$T \in \mathrm{L}(V, W)$, $U \in \mathrm{L}(W, Z)$. 则

$$\Lambda^r(T^t) \circ \Lambda^r(U^t) = \Lambda^r((UT)^t).$$

证明. 对任意$L \in \Lambda^r(Z^*)$和$\alpha_1, \ldots, \alpha_r \in V$, 有

$$\begin{aligned} \Lambda^r(T^t)(\Lambda^r(U^t)(L))(\alpha_1, \ldots, \alpha_r) &= \Lambda^r(U^t)(L)(T\alpha_1, \ldots, T\alpha_r) \\ &= L(UT\alpha_1, \ldots, UT\alpha_r) = \Lambda^r((UT)^t)(L)(\alpha_1, \ldots, \alpha_r), \end{aligned}$$

即

$$\Lambda^r(T^t)(\Lambda^r(U^t)(L)) = \Lambda^r((UT)^t)(L), \quad \forall L \in \Lambda^r(Z^*).$$

因此命题成立. □

下面假设V是有限维的, $\dim V = n \geq 1$. 对于$T \in \mathrm{L}(V, V)$, 由定理5.8, $\Lambda^n(T^t)$是一维线性空间$\Lambda^n(V^*)$到自身的线性映射, 从而是恒同映射的常数倍(这里的"常数"指域F中的元素). 我们把这个常数定义为T的行列式.

定义 5.3. 设n是正整数, V是域F上的n维线性空间, $T \in \mathrm{L}(V, V)$. 满足

$$\Lambda^n(T^t) = \det(T)\, \mathrm{id}_{\Lambda^n(V^*)} \tag{5.8}$$

的域F中的元素$\det(T)$称为T的**行列式**(determinant).

把行列式的定义式(5.8)写得具体一些, 即

$$\Lambda^n(T^t)(L) = \det(T)L, \quad \forall L \in \Lambda^n(V^*),$$

或者

$$L(T\alpha_1, \ldots, T\alpha_n) = \det(T)L(\alpha_1, \ldots, \alpha_n), \quad \forall L \in \Lambda^n(V^*), \alpha_1, \ldots, \alpha_n \in V. \tag{5.9}$$

线性映射的行列式满足下面的基本性质.

命题 5.11. 设V是域F上的n维线性空间.

(1) $\det(\mathrm{id}_V) = 1$.

(2) 对任意$T, U \in \mathrm{L}(V, V)$, 有$\det(TU) = \det(T)\det(U)$.

(3) $T \in \mathrm{L}(V, V)$可逆的充分必要条件是$\det(T) \neq 0$. 此时, $\det(T^{-1}) = \det(T)^{-1}$.

(4) 设W是另一个n维线性空间, $\Phi : V \to W$是线性同构. 如果$T \in \mathrm{L}(V, V)$和$U \in \mathrm{L}(W, W)$满足$\Phi \circ T = U \circ \Phi$, 则$\det(T) = \det(U)$.

(5) 设$\{\alpha_1, \ldots, \alpha_n\}$是$V$的基, $\{f_1, \ldots, f_n\}$是它在V^* 中的对偶基, $T \in \mathrm{L}(V, V)$. 则

$$\det(T) = f_1 \wedge \cdots \wedge f_n(T\alpha_1, \ldots, T\alpha_n).$$

证明. (1) 容易看出$\Lambda^n(\mathrm{id}_V^t) = \mathrm{id}_{\Lambda^n(V^*)}$. 因此$\det(\mathrm{id}_V) = 1$.

(2) 由命题5.10,

$$\det(TU)\mathrm{id}_{\Lambda^n(V^*)} = \Lambda^n((TU)^t) = \Lambda^n(U^t)\circ\Lambda^n(T^t)$$
$$= (\det(U)\mathrm{id}_{\Lambda^n(V^*)})\circ(\det(T)\mathrm{id}_{\Lambda^n(V^*)}) = \det(T)\det(U)\mathrm{id}_{\Lambda^n(V^*)}.$$

因此$\det(TU) = \det(T)\det(U)$.

(3) 假设T可逆. 由(1)和(2),

$$\det(T)\det(T^{-1}) = \det(\mathrm{id}_V) = 1.$$

因此$\det(T)\neq 0$, 并且$\det(T^{-1}) = \det(T)^{-1}$. 反过来, 假设$\det(T)\neq 0$. 取$L\in\Lambda^n(V^*)\setminus\{0\}$ 和V的基$\{\alpha_1,\dots,\alpha_n\}$. 由推论5.9, $L(\alpha_1,\dots,\alpha_n)\neq 0$. 因此(5.9)推出

$$L(T\alpha_1,\dots,T\alpha_n) = \det(T)L(\alpha_1,\dots,\alpha_n)\neq 0.$$

再次利用推论5.9, 即知$\{T\alpha_1,\dots,T\alpha_n\}$是$V$的基. 这说明$T$把$V$的基映为基. 因此$T$可逆.

(4) 由命题5.10,

$$\Lambda^n((\Phi\circ T)^t) = \Lambda^n(T^t)\circ\Lambda^n(\Phi^t) = \det(T)\Lambda^n(\Phi^t),$$
$$\Lambda^n((U\circ\Phi)^t) = \Lambda^n(\Phi^t)\circ\Lambda^n(U^t) = \det(U)\Lambda^n(\Phi^t).$$

上面两个等式最左边的表达式相等. 所以

$$\det(T)\Lambda^n(\Phi^t) = \det(U)\Lambda^n(\Phi^t).$$

另一方面,

$$\Lambda^n(\Phi^t)\circ\Lambda^n((\Phi^{-1})^t) = \Lambda^n((\Phi^{-1}\circ\Phi)^t) = \Lambda^n(\mathrm{id}_V^t) = \mathrm{id}_{\Lambda^n(V^*)}.$$

特别地, $\Lambda^n(\Phi^t)\neq 0$. 因此$\det(T) = \det(U)$.

(5) 由(5.9), 有

$$f_1\wedge\cdots\wedge f_n(T\alpha_1,\dots,T\alpha_n) = \det(T)f_1\wedge\cdots\wedge f_n(\alpha_1,\dots,\alpha_n).$$

另一方面, 由引理5.7(2), $f_1\wedge\cdots\wedge f_n(\alpha_1,\dots,\alpha_n) = 1$. 所以结论成立. □

注 5.2. 当V是n维实线性空间时, 线性映射$T\in\mathrm{L}(V,V)$的行列式可以视为T把广义平行多面体的有向体积扩大的倍数. 设$P = P(\alpha_1,\dots,\alpha_n)$是$V$中非退化的广义平行多面体(见注5.1). 容易看出, P在映射T下的像$T(P)$即为广义平行多面体$P(T\alpha_1,\dots,T\alpha_n)$, 并且$T(P)$非退化意味着$T$可逆. 另一方面, 通过选取$L\in\Lambda^n(V^*)\setminus\{0\}$并应用(5.9), 可以看出$T(P)$非退化等价于$\det(T)\neq 0$. 此时, 在$P$和$T(P)$上面分别选取由顶点的排列$(\alpha_1,\dots,\alpha_n)$和$(T\alpha_1,\dots,T\alpha_n)$决定的定向, 则它们的有向体积分别是$L(\alpha_1,\dots,\alpha_n)$和$L(T\alpha_1,\dots,T\alpha_n)$. 因此, $\det(T)$即为$T(P)$ 与P的有向体积的比值. 值得注意的是, 这一比值不仅与P的定向的选择无关, 也与P自身无关, 即T把所有广义平行多面体的有向体积扩大了相同的倍数. 另外, 虽然有向体积函数L不唯一, 但T把有向体积扩大的倍数不依赖于L的选取.

行列式是1的线性映射经常有特殊的重要性. 对于有限维线性空间V, 我们记

$$\mathrm{SL}(V) = \{T\in\mathrm{L}(V,V)\mid\det(T) = 1\}.$$

由命题5.11(1)–(3)容易看出, $\mathrm{SL}(V)$是一般线性群$\mathrm{GL}(V)$的子群, 称为V **的特殊线性群**(special linear group of V).

习题 5.3.

1. 设$B\in F^{n\times n}$. 考虑线性映射$T_B: F^{n\times n}\to F^{n\times n}$, $T_B(A) = AB - BA$. 证明$\det(T_B) = 0$.

2. 设V是n维实线性空间, $n \geq 1$. 集合$\Lambda^n(V^*) \setminus \{0\}$的非空子集$\mathcal{O}$称为一个**连通分支**(connected component), 如果对任意$L \in \mathcal{O}$有$\mathcal{O} = \{cL \mid c > 0\}$.
   (1) 证明$\Lambda^n(V^*) \setminus \{0\}$恰好有两个连通分支. 每个连通分支称为$V$的一个**定向**(orientation).
   (2) 假设取定了V的一个定向$\mathcal{O}^+$, 并记V的另一个定向为$\mathcal{O}^-$. 设$T \in \mathrm{L}(V, V)$可逆. 如果$\Lambda^n(T^t)(\mathcal{O}^+) = \mathcal{O}^+$, 则称$T$是**保定向的**(orientation-preserving); 如果$\Lambda^n(T^t)(\mathcal{O}^+) = \mathcal{O}^-$, 则称$T$是**反定向的**(orientation-reversing). 证明T保定向的充分必要条件是$\det(T) > 0$, T反定向的充分必要条件是$\det(T) < 0$.

§5.4 方阵的行列式

现在我们利用线性映射的行列式来定义方阵的行列式. 设$A \in F^{n\times n}$. 它诱导了两个线性映射

$$L_A : F^{n\times 1} \to F^{n\times 1}, \quad L_A(X) = AX,$$
$$R_A : F^n \to F^n, \quad R_A(\alpha) = \alpha A.$$

引理 5.12. $\det(L_A) = \det(R_A)$.

证明. 设$\{\epsilon_1, \ldots, \epsilon_n\}$是$F^{n\times 1}$的标准基, $\{f_1, \ldots, f_n\}$是它在$(F^{n\times 1})^*$中的对偶基, A的第j个列向量为A_j. 由命题5.11(5),

$$\begin{aligned}
\det(L_A) &= f_1 \wedge \cdots \wedge f_n(A\epsilon_1, \ldots, A\epsilon_n)\\
&= f_1 \wedge \cdots \wedge f_n(A_1, \ldots, A_n)\\
&= \sum_{\sigma \in S_n} \mathrm{sgn}(\sigma) f_{\sigma(1)}(A_1) \cdots f_{\sigma(n)}(A_n)\\
&= \sum_{\sigma \in S_n} \mathrm{sgn}(\sigma) A_{\sigma(1)1} \cdots A_{\sigma(n)n}.
\end{aligned}$$

另一方面, 设$\{\delta_1, \ldots, \delta_n\}$是$F^n$的标准基, $\{g_1, \ldots, g_n\}$是其对偶基, A的第i个行向量为α_i. 类似地有

$$\begin{aligned}
\det(R_A) &= g_1 \wedge \cdots \wedge g_n(\delta_1 A, \ldots, \delta_n A)\\
&= g_1 \wedge \cdots \wedge g_n(\alpha_1, \ldots, \alpha_n)\\
&= \sum_{\sigma \in S_n} \mathrm{sgn}(\sigma) g_{\sigma(1)}(\alpha_1) \cdots g_{\sigma(n)}(\alpha_n)\\
&= \sum_{\sigma \in S_n} \mathrm{sgn}(\sigma) A_{1\sigma(1)} \cdots A_{n\sigma(n)}.
\end{aligned}$$

对于$\sigma \in S_n$, 数对的集合$\{(\sigma(1), 1), \ldots, (\sigma(n), n)\}$与$\{(1, \sigma^{-1}(1)), \ldots, (n, \sigma^{-1}(n))\}$相等. 所以

$$A_{\sigma(1)1} \cdots A_{\sigma(n)n} = A_{1\sigma^{-1}(1)} \cdots A_{n\sigma^{-1}(n)}.$$

另外, $\mathrm{sgn}(\sigma) = \mathrm{sgn}(\sigma^{-1})$. 因此

$$\begin{aligned}
\sum_{\sigma \in S_n} \mathrm{sgn}(\sigma) A_{\sigma(1)1} \cdots A_{\sigma(n)n} &= \sum_{\sigma \in S_n} \mathrm{sgn}(\sigma^{-1}) A_{1\sigma^{-1}(1)} \cdots A_{n\sigma^{-1}(n)}\\
&= \sum_{\sigma \in S_n} \mathrm{sgn}(\sigma) A_{1\sigma(1)} \cdots A_{n\sigma(n)}.
\end{aligned}$$

这就证明了$\det(L_A) = \det(R_A)$. □

定义 5.4. 设$A \in F^{n\times n}$. 我们定义A的**行列式**为

$$\det(A) = \det(L_A) = \det(R_A).$$

在引理5.12的证明中, 我们已经得到了行列式的表达式.

推论 5.13. 设$A \in F^{n\times n}$. 则

$$\det(A) = \sum_{\sigma\in S_n} \operatorname{sgn}(\sigma) A_{1\sigma(1)} \cdots A_{n\sigma(n)} = \sum_{\sigma\in S_n} \operatorname{sgn}(\sigma) A_{\sigma(1)1} \cdots A_{\sigma(n)n}. \tag{5.10}$$

如果$A = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \vdots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{bmatrix}$, 我们也记$\det(A) = \begin{vmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \vdots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{vmatrix}$.

例 5.8. 由(5.10), 我们有

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

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

下面给出方阵行列式的几个基本性质.

命题 5.14. (1) $\det(I_n) = 1$.

(2) 设$A, B \in F^{n\times n}$. 则$\det(AB) = \det(A)\det(B)$.

(3) 方阵A可逆的充分必要条件是$\det(A) \neq 0$. 此时, $\det(A^{-1}) = \det(A)^{-1}$.

(4) 相似的方阵行列式相等.

(5) $\det(A^t) = \det(A)$.

(6) 映射$(F^{n\times 1})^n \to F$, $(A_1, \ldots, A_n) \mapsto \det[A_1, \ldots, A_n]$是$n$ 重交错线性函数.

(7) 映射$(F^n)^n \to F$, $(\alpha_1, \ldots, \alpha_n) \mapsto \det \begin{bmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{bmatrix}$是$n$重交错线性函数.

证明. (1) 由命题5.11(1), 有

$$\det(I_n) = \det(L_{I_n}) = \det(\mathrm{id}_{F^{n\times 1}}) = 1.$$

(2) 由命题5.11(2), 有

$$\det(AB) = \det(L_{AB}) = \det(L_A L_B) = \det(L_A)\det(L_B) = \det(A)\det(B).$$

(3) 由命题5.11(3),

$$A\text{可逆} \iff L_A\text{可逆} \iff \det(A) = \det(L_A) \neq 0.$$

此时,

$$\det(A^{-1}) = \det(L_{A^{-1}}) = \det(L_A^{-1}) = \det(L_A)^{-1} = \det(A)^{-1}.$$

(4) 设$A, B \in F^{n\times n}$相似, 即存在可逆矩阵$P \in F^{n\times n}$使得$A = PBP^{-1}$. 则由(3)和(4),

$$\det(A) = \det(PBP^{-1}) = \det(P)\det(B)\det(P)^{-1} = \det(B).$$

(5) 考虑线性同构$\Phi : F^{n\times 1} \to F^n$, $\Phi(X) = X^t$. 则对$X \in F^{n\times 1}$有

$$\Phi(L_A(X)) = \Phi(AX) = (AX)^t = X^tA^t = R_{A^t}(X^t) = R_{A^t}(\Phi(X)).$$

即$\Phi \circ L_A = R_{A^t} \circ \Phi$. 由命题5.11(4), $\det(L_A) = \det(R_{A^t})$. 因此$\det(A) = \det(A^t)$.

(6)+(7) 在引理5.12的证明中, 我们已经得到

$$\det[A_1, \ldots, A_n] = f_1 \wedge \cdots \wedge f_n(A_1, \ldots, A_n),$$

$$\det \begin{bmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{bmatrix} = g_1 \wedge \cdots \wedge g_n(\alpha_1, \ldots, \alpha_n).$$

因此(6)和(7)中的映射是n重交错线性函数. □

与线性映射的情况类似, 行列式是1的方阵也是很重要的. 命题5.14(1)–(3)表明, 集合

$$\mathrm{SL}_n(F) = \{A \in F^{n\times n} \mid \det(A) = 1\}$$

是$\mathrm{GL}_n(F)$的子群, 称为**F上的n次特殊线性群**(special linear group of degree n over F).

命题5.14的(6)和(7)提供了计算行列式的有效工具:

推论 5.15. (1) 若方阵有零行或零列, 则行列式为0.

(2) 若方阵有两行(或两列)相同或成比例, 则行列式为0.

(3) 互换两行或两列, 方阵的行列式反号.

(4) 把某行(或列)的倍数加到另一行(或列)上, 方阵的行列式不变.

(5) 某行(或列)乘以某个常数, 导致方阵的行列式也乘以相同的常数.

证明. 这些都是命题5.14(6)和(7)的明显推论. 例如, 对于(4)中列的情况, 可以验证如下:

$$\begin{aligned} & \det[A_1, \ldots, A_s, \ldots, cA_s + A_t, \ldots, A_n] \\ = \ & c\det[A_1, \ldots, A_s, \ldots, A_s, \ldots, A_n] + \det[A_1, \ldots, A_s, \ldots, A_t, \ldots, A_n] \\ = \ & \det[A_1, \ldots, A_n]. \end{aligned}$$

□

推论5.15(3)–(5)给出了方阵在初等行(列)变换下, 行列式的改变方式. 利用初等行(列)变换, 我们总是可以把方阵化为上三角或下三角的形式. 这里矩阵$A \in F^{n\times n}$称为**上三角的**(upper triangular), 如果当$i > j$时有$A_{ij} = 0$; 矩阵A称为**下三角的**(lower triangular), 如果当$i < j$时有$A_{ij} = 0$. 对于这两种方阵, 我们有:

命题 5.16. 上三角或下三角矩阵的行列式等于所有对角元的乘积. 特别地, 对角矩阵的行列式等于所有对角元的乘积.

证明. 设A是上三角的, 即当$i > j$时有$A_{ij} = 0$. 此时, 在表达式

$$\det(A) = \sum_{\sigma \in S_n} \mathrm{sgn}(\sigma) A_{1\sigma(1)} \cdots A_{n\sigma(n)}$$

中, 如果$A_{1\sigma(1)} \cdots A_{n\sigma(n)} \neq 0$, 则对所有$1 \le i \le n$有$i \le \sigma(i)$. 这推出$\sigma = \mathrm{id}$, 即$\sigma(i) = i$. 所以$\det(A) = A_{11} \cdots A_{nn}$. 下三角的情况可以类似证明. □

例 5.9. 丘老师书上册32页例1; 教材158页例6.

作为命题5.16的推广, 我们证明下面的结果.

命题 5.17. 设方阵A是分块上三角矩阵, 即

$$A = \begin{bmatrix} A_{11} & A_{12} & \cdots & A_{1,k-1} & A_{1k} \\ 0 & A_{22} & \cdots & A_{2,k-1} & A_{2k} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & A_{k-1,k-1} & A_{k-1,k} \\ 0 & 0 & \cdots & 0 & A_{kk} \end{bmatrix},$$

其中A_{ii}为方阵, $A_{ij}\ (i<j)$和0为合适尺寸的矩阵. 则

$$\det(A) = \det(A_{11}) \cdots \det(A_{kk}).$$

特别地, 如果$B \in F^{r\times r}$, $C \in F^{r\times s}$, $D \in F^{s\times s}$, 则

$$\det \begin{bmatrix} B & C \\ 0 & D \end{bmatrix} = \det(B)\det(D). \tag{5.11}$$

对于分块下三角矩阵和分块对角矩阵, 类似的结论也成立.

证明. 我们先证明(5.11)式. 设矩阵$\begin{bmatrix} B & C \\ 0 & D \end{bmatrix}$的$(i,j)$元为$a_{ij}$, $i,j \in \{1,\ldots,r+s\}$. 则

$$\det \begin{bmatrix} B & C \\ 0 & D \end{bmatrix} = \sum_{\sigma \in S_{r+s}} \operatorname{sgn}(\sigma) a_{1\sigma(1)} \cdots a_{r+s,\sigma(r+s)}.$$

注意到当$i \geq r+1$并且$\sigma(i) \leq r$时, 有$a_{i\sigma(i)} = 0$. 所以, 如果$\sigma \in S_{r+s}$使得$a_{1\sigma(1)} \cdots a_{r+s,\sigma(r+s)} \neq 0$, 则当$i \geq r+1$时有$\sigma(i) \geq r+1$. 这推出当$i \leq r$时有$\sigma(i) \leq r$. 因此, 存在$\sigma_1 \in S_r$和$\sigma_2 \in S_s$使得

$$\sigma(i) = \begin{cases} \sigma_1(i), & 1 \leq i \leq r, \\ r + \sigma_2(i-r), & r+1 \leq i \leq r+s. \end{cases}$$

此时有$\operatorname{sgn}(\sigma) = \operatorname{sgn}(\sigma_1)\operatorname{sgn}(\sigma_2)$. 从而

$$\begin{aligned} \det \begin{bmatrix} B & C \\ 0 & D \end{bmatrix} &= \sum_{\sigma_1 \in S_r, \sigma_2 \in S_s} \operatorname{sgn}(\sigma_1)\operatorname{sgn}(\sigma_2) a_{1\sigma_1(1)} \cdots a_{r\sigma_1(r)} a_{r+1,r+\sigma_2(1)} \cdots a_{r+s,r+\sigma_2(s)} \\ &= \left(\sum_{\sigma_1 \in S_r} \operatorname{sgn}(\sigma_1) a_{1\sigma_1(1)} \cdots a_{r\sigma_1(r)} \right) \left(\sum_{\sigma_2 \in S_s} \operatorname{sgn}(\sigma_2) a_{r+1,r+\sigma_2(1)} \cdots a_{r+s,r+\sigma_2(s)} \right) \\ &= \det(B)\det(D). \end{aligned}$$

这就证明了(5.11).

接下来用归纳法证明分块上三角矩阵的结论. 当$k=2$时结论已经证明. 假设$k \geq 3$, 并且结论对$k-1$成立. 则

$$\det(A) = \det \begin{bmatrix} A_{11} & A_{12} & \cdots & A_{1,k-1} \\ 0 & A_{22} & \cdots & A_{2,k-1} \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & A_{k-1,k-1} \end{bmatrix} \det(A_{kk}) = \det(A_{11}) \cdots \det(A_{kk}).$$

这就完成了证明. 分块下三角矩阵情况的证明类似. □

我们讨论了线性映射和矩阵的行列式. 线性映射自身的行列式与它关于有序基的矩阵的行列式有什么关系? 下面命题的前一部分回答了这一问题. 它也可以用来把线性映射行列式的计算转化为方阵行列式的计算.

命题 5.18. 设V是$n \geq 1$维线性空间, $T \in \mathrm{L}(V,V)$.

(1) 设$\mathcal{B}$是V的有序基, $[T]_{\mathcal{B}}$是T关于该有序基的矩阵. 则$\det([T]_{\mathcal{B}}) = \det(T)$.

(2) 设$T^t \in \mathrm{L}(V^*, V^*)$是$T$的转置映射. 则$\det(T^t) = \det(T)$.

证明. (1) 回忆坐标映射$\Gamma : V \to F^{n\times 1}$, $\alpha \mapsto [\alpha]_{\mathcal{B}}$是线性同构, 并且$\Gamma \circ T = L_{[T]_{\mathcal{B}}} \circ \Gamma$. 由命题5.11(4), 即得

$$\det(T) = \det(L_{[T]_{\mathcal{B}}}) = \det([T]_{\mathcal{B}}).$$

(2) 设$\mathcal{B}$是V的有序基, $\mathcal{B}^*$是$\mathcal{B}$在V^*中的对偶基, $[T]_{\mathcal{B}}$和$[T^t]_{\mathcal{B}^*}$分别是T和T^t关于相应有序基的矩阵. 我们知道, $[T^t]_{\mathcal{B}^*} = [T]_{\mathcal{B}}^t$. 由(1)和命题5.14(5), 即得

$$\det(T^t) = \det([T^t]_{\mathcal{B}^*}) = \det([T]_{\mathcal{B}}) = \det(T).$$

这就完成了证明. □

习题 5.4.

1. 教材162–163页习题3,4,5.
2. 丘老师书上册26页1(2); 35页1(3), 2(1), 3(1), 4(2).
3. 利用方阵行列式的表达式(5.10)验证命题5.14中的(1),(5),(6),(7).
4. 证明$\mathrm{sgn}(\sigma) = \det R(\sigma)$, 这里$R(\sigma)$是习题5.1.3中定义的矩阵.
5. 利用命题5.11(5)证明命题5.18(2).
6. 设K是域F的子域, 满足F作为K上的线性空间是有限维的. 对于$x \in F$, K-线性映射
$$T_x : F \to F, \quad y \mapsto xy$$
的行列式称为x的(从F到K的)**范数**(norm), 记为$\mathrm{N}_{F/K}(x) = \det(T_x)$. 对于下面的情况, 给出范数的表达式.
   (1) $K = \mathbb{R}$, $F = \mathbb{C}$.
   (2) $K = \mathbb{Q}$, $F = \mathbb{Q}(\sqrt{2}) := \{a + b\sqrt{2} \mid a, b \in \mathbb{Q}\}$.
   (3) $K = \mathbb{Q}$, $F = \mathbb{Q}(\sqrt[3]{2}) := \{a + b\sqrt[3]{2} + c\sqrt[3]{4} \mid a, b, c \in \mathbb{Q}\}$.
7. 设V是有限维复线性空间, $T \in \mathrm{L}(V,V)$. 我们把对V执行“忘掉复结构”的操作后得到的实线性空间记为$V_{\mathbb{R}}$, 从而T可以视为实线性映射$T_{\mathbb{R}} \in \mathrm{L}(V_{\mathbb{R}}, V_{\mathbb{R}})$. 证明$\det(T_{\mathbb{R}}) = |\det(T)|^2$. (提示: 若$T$关于$V$的某个基的矩阵为$A$, 则$T_{\mathbb{R}}$关于$V_{\mathbb{R}}$的某个基的矩阵为$\begin{bmatrix} \mathrm{Re}(A) & -\mathrm{Im}(A) \\ \mathrm{Im}(A) & \mathrm{Re}(A) \end{bmatrix}$. 设$P = \begin{bmatrix} I & iI \\ iI & I \end{bmatrix}$. 则$P \begin{bmatrix} \mathrm{Re}(A) & -\mathrm{Im}(A) \\ \mathrm{Im}(A) & \mathrm{Re}(A) \end{bmatrix} P^{-1} = \begin{bmatrix} A & \\ & \overline{A} \end{bmatrix}$.)
8. 设K和F如习题6, V是F上的有限维线性空间, $T \in \mathrm{L}(V,V)$. 把V视为K上的线性空间, 记为V_K. 从而T可以视为K-线性映射$T_K \in \mathrm{L}(V_K, V_K)$. 证明$\det(T_K) = \mathrm{N}_{F/K}(\det(T))$. 特别地, 如果$F$是$L$的子域, 则$\mathrm{N}_{L/K} = \mathrm{N}_{F/K} \circ \mathrm{N}_{L/F}$. (注: 此题暂时知道结论即可. 以后将利用有理标准形来证明. 更一般的结果见Bourbaki, Algebra I, p. 546, Prop. 6.)
9. 设$m < n$, $A \in F^{n\times m}$, $B \in F^{m\times n}$. 证明$\det(AB) = 0$.

§5.5 行列式按行或列展开

这一节考察的性质可以用来更加有效地计算行列式, 并且具有重要的理论意义.

设$A = [a_{ij}] \in F^{n\times n}$, $n \geq 2$. 对于正整数$k < n$和指标集$I, J \subset \{1, \ldots, n\}$, $|I| = |J| = k$, 记$A_{I,J} \in F^{k\times k}$为由$A$的矩阵元$\{a_{ij} \mid i \in I, j \in J\}$ 构成的矩阵. 我们同时考虑矩阵A_{I^c,J^c}, 这里$I^c = \{1, \ldots, n\} \setminus I$. 由矩阵$A$得到$A_{I,J}$和$A_{I^c,J^c}$时, 行和列的排列顺序不变. 也就是说, 如果$\sigma_I \in S_n$是由(5.1)定义的置换, 则$A_{I,J}$和$A_{I^c,J^c}$的矩阵元分别为

$$(A_{I,J})_{rs} = a_{\sigma_I(r),\sigma_J(s)}, \qquad r, s \in \{1, \ldots, k\},$$
$$(A_{I^c,J^c})_{rs} = a_{\sigma_I(k+r),\sigma_J(k+s)}, \qquad r, s \in \{1, \ldots, n-k\}.$$

关于A的行列式, 我们有下面的Laplace展开定理.

定理 5.19(Laplace). 设n, k是正整数, $k < n$, $A \in F^{n\times n}$. 对于$I \subset \{1, \ldots, n\}$, 记$\Sigma_I = \sum_{i\in I} i$.

(1) (行列式按k行展开) 取定$I_0 \subset \{1, \ldots, n\}$, $|I_0| = k$. 则

$$\det(A) = \sum_{\substack{J\subset\{1,\ldots,n\}\\ |J|=k}} (-1)^{\Sigma_{I_0}+\Sigma_J} \det(A_{I_0,J}) \det(A_{I_0^c,J^c}).$$

(2) (行列式按k列展开) 取定$J_0 \subset \{1, \ldots, n\}$, $|J_0| = k$. 则

$$\det(A) = \sum_{\substack{I\subset\{1,\ldots,n\}\\ |I|=k}} (-1)^{\Sigma_I+\Sigma_{J_0}} \det(A_{I,J_0}) \det(A_{I^c,J_0^c}).$$

行列式$\det(A_{I,J})$称为A的行列式的**子式**(minor), 行列式$\det(A_{I^c,J^c})$称为子式$\det(A_{I,J})$的**余子式**(complementary minor), 定理中的因子

$$(-1)^{\Sigma_I+\Sigma_J} \det(A_{I^c,J^c})$$

称为子式$\det(A_{I,J})$的**余因子**或**代数余子式**(cofactor). 在这样的词汇下, 定理5.19可以叙述为: A的行列式等于固定行(或列)指标集的所有子式与相应的余因子的乘积之和.

在证明定理前, 我们先来考察它的特殊情况.

例 5.10. 设$B \in F^{k\times k}$, $C \in F^{k\times(n-k)}$, $D \in F^{(n-k)\times(n-k)}$, $A = \begin{bmatrix} B & C \\ 0 & D \end{bmatrix}$. 我们把$A$的行列式按前$k$列展开, 即对$J_0 = \{1, \ldots, k\}$应用定理5.19(2). 对于$I \subset \{1, \ldots, n\}$, $|I| = k$, 如果$I \neq J_0$, 则矩阵A_{I,J_0}总有零行, 从而行列式为0. 另一方面, 我们有$A_{J_0,J_0} = B$, $A_{J_0^c,J_0^c} = D$. 因此, 定理5.19(2)推出

$$\det(A) = (-1)^{2\Sigma_{J_0}} \det(A_{J_0,J_0}) \det(A_{J_0^c,J_0^c}) = \det(B)\det(D).$$

这样我们就重新得到了(5.11)式.

接下来我们考虑$k = 1$时的情况. 此时, 对于$i, j \in \{1, \ldots, n\}$, 一阶矩阵$A_{\{i\},\{j\}}$ 的行列式即为a_{ij}. 为了简化记号, 记$M_{ij} = A_{\{i\}^c,\{j\}^c}$. 它是在$A$中去掉第$i$ 行和第j列后得到的$n-1$阶矩阵. 也就是说, 如果定义置换

$$\sigma_i = \begin{pmatrix} 1 & \cdots & i-1 & i & \cdots & n-1 & n \\ 1 & \cdots & i-1 & i+1 & \cdots & n & i \end{pmatrix}, \tag{5.12}$$

则M_{ij}的矩阵元为

$$(M_{ij})_{rs} = a_{\sigma_i(r),\sigma_j(s)}, \qquad r, s \in \{1, \ldots, n-1\}.$$

注意这里的σ_i实际上是(5.1)中的$\sigma_{\{i\}^c}$. 这样, 定理5.19在$k = 1$时的特殊情况可以叙述如下.

定理 5.20. 设$A = [a_{ij}] \in F^{n\times n}$, $n \geq 2$, 并沿用上面的记号.

(1) (行列式按一行展开) 对任意取定的$1 \leq i_0 \leq n$, 有

$$\det(A) = \sum_{j=1}^{n} (-1)^{i_0+j} a_{i_0 j} \det(M_{i_0 j}).$$

(2) (行列式按一列展开) 对任意取定的$1 \leq j_0 \leq n$, 有

$$\det(A) = \sum_{i=1}^{n} (-1)^{i+j_0} a_{i j_0} \det(M_{i j_0}).$$

我们把对应(i, j)-元的余因子记为

$$C_{ij} = (-1)^{i+j} \det(M_{ij}).$$

推论 5.21. 对任意$i, j \in \{1, \ldots, n\}$, 有

$$\sum_{k=1}^{n} a_{ik} C_{jk} = \sum_{k=1}^{n} a_{ki} C_{kj} = \delta_{ij} \det(A).$$

证明. 当$i = j$时, 等式即为定理5.20. 设$i \neq j$. 记B为把A的第j行替换为第i 行得到的矩阵. 由于B有两行相同, 所以$\det(B) = 0$. 另一方面, 把B的行列式按第j行展开, 并注意到B的第j行的每个矩阵元与A的同位置矩阵元具有相同的余因子, 即得

$$\sum_{k=1}^{n} a_{ik} C_{jk} = \det(B) = 0.$$

类似地,

$$\sum_{k=1}^{n} a_{ki} C_{kj} = 0.$$

□

以余因子C_{ij}为(i, j)-元的n阶矩阵的转置称为A的**伴随矩阵**(adjugate, adjunct 或classical adjoint), 记为adj(A), 即

$$\operatorname{adj}(A)_{ij} = C_{ji} = (-1)^{i+j} \det(M_{ji}).$$

例如,

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}, \quad \operatorname{adj}(A) = \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}.$$

推论5.21可以重新叙述为:

推论 5.22. 设$A \in F^{n\times n}$, $n \geq 2$. 则

$$A \operatorname{adj}(A) = \operatorname{adj}(A) A = \det(A) I_n.$$

特别地, 如果A可逆, 则

$$A^{-1} = \det(A)^{-1} \operatorname{adj}(A).$$

□

推论 5.23(Cramer法则). 设$n \geq 2$, $A \in F^{n\times n}$可逆, $Y \in F^{n\times 1}$. 记B_j为把A的第j列替换为Y后得到的矩阵. 则线性方程组$AX = Y$的唯一解为

$$X = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}, \qquad x_j = \frac{\det(B_j)}{\det(A)}.$$

证明. 我们给出两个证明.

证明一. 方程组的唯一解为

$$X = A^{-1}Y = \det(A)^{-1}\mathrm{adj}(A)Y.$$

两边取第j行, 并对B_j应用定理5.20(2), 即得

$$x_j = \det(A)^{-1}\sum_{i=1}^{n} y_i C_{ij} = \det(A)^{-1}\det(B_j).$$

证明二. 记A的第i列为A_i. 则

$$Y = [A_1, \ldots, A_n]\begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = \sum_{i=1}^{n} x_i A_i.$$

(注意这说明x_j是Y在$F^{n\times 1}$的基$\{A_1, \ldots, A_n\}$下的第j个坐标.) 因此

$$\begin{aligned}\det(B_j) &= \det[A_1, \ldots, A_{j-1}, \sum_{i=1}^{n} x_i A_i, A_{j+1}, \ldots, A_n] \\ &= \sum_{i=1}^{n} x_i \det[A_1, \ldots, A_{j-1}, A_i, A_{j+1}, \ldots, A_n] \\ &= x_j \det(A).\end{aligned}$$

这就完成了证明. □

例 5.11. 考虑方程组$\begin{bmatrix} a & b \\ c & d \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}$. 如果$\begin{vmatrix} a & b \\ c & d \end{vmatrix} = ad - bc \neq 0$, 则唯一解为

$$x_1 = \frac{\begin{vmatrix} y_1 & b \\ y_2 & d \end{vmatrix}}{\begin{vmatrix} a & b \\ c & d \end{vmatrix}} = \frac{dy_1 - by_2}{ad - bc}, \qquad x_2 = \frac{\begin{vmatrix} a & y_1 \\ c & y_2 \end{vmatrix}}{\begin{vmatrix} a & b \\ c & d \end{vmatrix}} = \frac{ay_2 - cy_1}{ad - bc}.$$

考虑方程组$\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_2 \end{bmatrix} = \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix}$. 如果$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} \neq 0$, 则唯一解为

$$x_1 = \frac{\begin{vmatrix} y_1 & a_{12} & a_{13} \\ y_2 & a_{22} & a_{23} \\ y_3 & a_{32} & a_{33} \end{vmatrix}}{\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}}, \qquad x_2 = \frac{\begin{vmatrix} a_{11} & y_1 & a_{13} \\ a_{21} & y_2 & a_{23} \\ a_{31} & y_3 & a_{33} \end{vmatrix}}{\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}}, \qquad x_3 = \frac{\begin{vmatrix} a_{11} & a_{12} & y_1 \\ a_{21} & a_{22} & y_2 \\ a_{31} & a_{32} & y_3 \end{vmatrix}}{\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}}.$$

当方程组的阶数较大时, Cramer法则的重要性更多在于其理论价值. 在具体计算时, 初等行变换的方法具有更高的效率.

现在我们回到定理的证明. 我们先证明定理5.20.

定理5.20的证明. (1) 考虑S_n的n个子集

$$P_j = \{\sigma \in S_n \mid \sigma(i_0) = j\}, \quad 1 \leq j \leq n.$$

对于$\tau \in S_{n-1}$, 定义$\tilde{\tau} \in S_n$为

$$\tilde{\tau} = \begin{pmatrix} 1 & \cdots & n-1 & n \\ \tau(1) & \cdots & \tau(n-1) & n \end{pmatrix}. \tag{5.13}$$

容易看出,

$$P_j = \{\sigma_j \tilde{\tau} \sigma_{i_0}^{-1} \mid \tau \in S_{n-1}\},$$

这里σ_j, σ_{i_0}的定义如(5.12). 注意到$\operatorname{sgn}(\tilde{\tau}) = \operatorname{sgn}(\tau)$, $\operatorname{sgn}(\sigma_j) = (-1)^{n-j}$. 所以

$$\operatorname{sgn}(\sigma_j \tilde{\tau} \sigma_{i_0}^{-1}) = (-1)^{i_0+j} \operatorname{sgn}(\tau).$$

因此

$$\begin{aligned} \det(A) &= \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) \prod_{i=1}^{n} a_{i\sigma(i)} \\ &= \sum_{j=1}^{n} a_{i_0 j} \sum_{\sigma \in P_j} \operatorname{sgn}(\sigma) \prod_{r=1}^{n-1} a_{\sigma_{i_0}(r), \sigma\sigma_{i_0}(r)} \\ &= \sum_{j=1}^{n} a_{i_0 j} \sum_{\tau \in S_{n-1}} \operatorname{sgn}(\sigma_j \tilde{\tau} \sigma_{i_0}^{-1}) \prod_{r=1}^{n-1} a_{\sigma_{i_0}(r), \sigma_j \tau(r)} \\ &= \sum_{j=1}^{n} (-1)^{i_0+j} a_{i_0 j} \sum_{\tau \in S_{n-1}} \operatorname{sgn}(\tau) \prod_{r=1}^{n-1} (M_{i_0 j})_{r, \tau(r)} \\ &= \sum_{j=1}^{n} (-1)^{i_0+j} a_{i_0 j} \det(M_{i_0 j}). \end{aligned}$$

(2) 可以类似证明, 也可以通过对A^t应用(1)来证明. □

下面我们证明定理5.19. 读者可以与定理5.20的证明做比较.

定理5.19的证明. 与定理5.20的情况类似, 我们只需对(1)给出证明. 对于$\{1, \ldots, n\}$的k元子集J, 考虑S_n的子集

$$P_J = \{\sigma \in S_n \mid \sigma(I_0) = J\}.$$

则

$$\det(A) = \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) \prod_{i=1}^{n} a_{i\sigma(i)} = \sum_{\substack{J \subset \{1, \ldots, n\} \\ |J| = k}} \sum_{\sigma \in P_J} \operatorname{sgn}(\sigma) \prod_{i=1}^{n} a_{i\sigma(i)}. \tag{5.14}$$

对于$\tau \in S_k$和$\upsilon \in S_{n-k}$, 定义$(\tau, \upsilon) \in S_n$为

$$(\tau, \upsilon) = \begin{pmatrix} 1 & \cdots & k & k+1 & \cdots & n \\ \tau(1) & \cdots & \tau(k) & k+\upsilon(1), & \cdots & k+\upsilon(n-k) \end{pmatrix}.$$

容易看出,

$$P_J = \{\sigma_J (\tau, \upsilon) \sigma_{I_0}^{-1} \mid \tau \in S_k, \upsilon \in S_{n-k}\},$$

这里σ_J, σ_{I_0}的定义如(5.1). 容易看出, $\operatorname{sgn}(\tau, \upsilon) = \operatorname{sgn}(\tau)\operatorname{sgn}(\upsilon)$. 结合(5.2), 即得

$$\operatorname{sgn}(\sigma_J (\tau, \upsilon) \sigma_{I_0}^{-1}) = (-1)^{\Sigma_{I_0} + \Sigma_J} \operatorname{sgn}(\tau) \operatorname{sgn}(\upsilon).$$

因此

$$\begin{aligned}
\sum_{\sigma\in P_J}\operatorname{sgn}(\sigma)\prod_{i=1}^{n}a_{i\sigma(i)} &= \sum_{\sigma\in P_J}\operatorname{sgn}(\sigma)\prod_{r=1}^{k}a_{\sigma_{I_0}(r),\sigma\sigma_{I_0}(r)}\prod_{s=1}^{n-k}a_{\sigma_{I_0}(k+s),\sigma\sigma_{I_0}(k+s)}\\
&= \sum_{\tau\in S_k,\upsilon\in S_{n-k}}\operatorname{sgn}(\sigma_J(\tau,\upsilon)\sigma_{I_0}^{-1})\prod_{r=1}^{k}a_{\sigma_{I_0}(r),\sigma_J(\tau,\upsilon)(r)}\prod_{s=1}^{n-k}a_{\sigma_{I_0}(k+s),\sigma_J(\tau,\upsilon)(k+s)}\\
&= (-1)^{\Sigma_{I_0}+\Sigma_J}\sum_{\tau\in S_k,\upsilon\in S_{n-k}}\operatorname{sgn}(\tau)\operatorname{sgn}(\upsilon)\prod_{r=1}^{k}a_{\sigma_{I_0}(r),\sigma_J(\tau(r))}\prod_{s=1}^{n-k}a_{\sigma_{I_0}(k+s),\sigma_J(k+\upsilon(s))}\\
&= (-1)^{\Sigma_{I_0}+\Sigma_J}\sum_{\tau\in S_k,\upsilon\in S_{n-k}}\operatorname{sgn}(\tau)\operatorname{sgn}(\upsilon)\prod_{r=1}^{k}(A_{I_0,J})_{r\tau(r)}\prod_{s=1}^{n-k}(A_{I_0^c,J^c})_{s\upsilon(s)}\\
&= (-1)^{\Sigma_{I_0}+\Sigma_J}\left(\sum_{\tau\in S_k}\operatorname{sgn}(\tau)\prod_{r=1}^{k}(A_{I_0,J})_{r\tau(r)}\right)\left(\sum_{\upsilon\in S_{n-k}}\operatorname{sgn}(\upsilon)\prod_{s=1}^{n-k}(A_{I_0^c,J^c})_{s\upsilon(s)}\right)\\
&= (-1)^{\Sigma_{I_0}+\Sigma_J}\det(A_{I_0,J})\det(A_{I_0^c,J^c}).
\end{aligned}$$

将此式代入(5.14), 就完成了证明. □

在下一节中, 我们将给出定理5.19的另一个证明.

例 5.12. Vandermonde行列式(丘老师书上册41页).

习题 5.5.

1. 教材162–163页习题1,2,9,12.
2. 丘老师书上册43–44页1(2,4),2,4,6; 51页4; 56页3.

§5.6 对偶空间上的外代数

在这一节中, 我们重点考察多重交错线性函数的一种交错乘积, 称为外积. 这一构造在微分几何等领域中是非常重要的. 作为准备, 我们先考虑一般多重线性函数的张量积.

定义 5.5. 设V是域F上的线性空间, r,s是正整数, $L\in(V^*)^{\otimes r}$, $M\in(V^*)^{\otimes s}$. 我们定义L与M的**张量积**(tensor product)$L\otimes M\in(V^*)^{\otimes(r+s)}$ 为

$$L\otimes M(\alpha_1,\dots,\alpha_{r+s})=L(\alpha_1,\dots,\alpha_r)M(\alpha_{r+1},\dots,\alpha_{r+s}),\quad \alpha_1,\dots,\alpha_{r+s}\in V.$$

命题 5.24. (1) 映射$(V^*)^{\otimes r}\times(V^*)^{\otimes s}\to(V^*)^{\otimes(r+s)}$, $(L,M)\mapsto L\otimes M$是双线性的.

(2) 设$L\in(V^*)^{\otimes r}$, $M\in(V^*)^{\otimes s}$, $N\in(V^*)^{\otimes t}$. 则

$$(L\otimes M)\otimes N=L\otimes(M\otimes N).$$

证明. (1) 对任意$c\in F$, $L_1,L_2\in (V^*)^{\otimes r}$, $M\in (V^*)^{\otimes s}$, $\alpha_1,\dots,\alpha_{r+s}\in V$有

$$\begin{aligned}&(cL_1+L_2)\otimes M(\alpha_1,\dots,\alpha_{r+s})\\ =&(cL_1+L_2)(\alpha_1,\dots,\alpha_r)M(\alpha_{r+1},\dots,\alpha_{r+s})\\ =&cL_1(\alpha_1,\dots,\alpha_r)M(\alpha_{r+1},\dots,\alpha_{r+s})+L_2(\alpha_1,\dots,\alpha_r)M(\alpha_{r+1},\dots,\alpha_{r+s})\\ =&cL_1\otimes M(\alpha_1,\dots,\alpha_{r+s})+L_2\otimes M(\alpha_1,\dots,\alpha_{r+s})\\ =&(cL_1\otimes M+L_2\otimes M)(\alpha_1,\dots,\alpha_{r+s}).\end{aligned}$$

因此

$$(cL_1+L_2)\otimes M=cL_1\otimes M+L_2\otimes M.$$

类似地, 对任意$c\in F$, $L\in (V^*)^{\otimes r}$, $M_1,M_2\in (V^*)^{\otimes s}$有

$$L\otimes (cM_1+M_2)=cL\otimes M_1+L\otimes M_2.$$

(2) 对任意$\alpha_1,\dots,\alpha_{r+s+t}\in V$, 我们有

$$\begin{aligned}&(L\otimes M)\otimes N(\alpha_1,\dots,\alpha_{r+s+t})\\ =&L\otimes M(\alpha_1,\dots,\alpha_{r+s})N(\alpha_{r+s+1},\dots,\alpha_{r+s+t})\\ =&L(\alpha_1,\dots,\alpha_r)M(\alpha_{r+1},\dots,\alpha_{r+s})N(\alpha_{r+s+1},\dots,\alpha_{r+s+t})\\ =&L(\alpha_1,\dots,\alpha_r)M\otimes N(\alpha_{r+1},\dots,\alpha_{r+s+t})\\ =&L\otimes (M\otimes N)(\alpha_1,\dots,\alpha_{r+s+t}).\end{aligned}$$

因此$(L\otimes M)\otimes N=L\otimes (M\otimes N)$. □

由命题5.24(2), 对于$L_1\in (V^*)^{\otimes r_1},\dots,L_n\in (V^*)^{\otimes r_n}$, 可以定义它们的张量积$L_1\otimes\cdots\otimes L_n$, 所得结果与做乘法的次序无关. 例5.4中的$r$重线性函数即为$f_1,\dots,f_r$的张量积, 那里的记号$f_1\otimes\cdots\otimes f_r$与这里也是一致的. 注意张量积不满足乘法交换率.

下面我们在空间有限维的情况给出$(V^*)^{\otimes r}$的基和维数.

定理 5.25. 设V是有限维线性空间, $\dim V=n$, $\{f_1,\dots,f_n\}$是V^*的基.

(1) 集合

$$\{f_{i_1}\otimes\cdots\otimes f_{i_r}\mid 1\le i_1,\dots,i_r\le n\} \tag{5.15}$$

是$(V^*)^{\otimes r}$的基. 特别地, $\dim (V^*)^{\otimes r}=n^r$.

(2) 如果$\{\alpha_1,\dots,\alpha_n\}$是$\{f_1,\dots,f_n\}$在$V$中的对偶基, 则对任意$L\in (V^*)^{\otimes r}$, 有

$$L=\sum_{1\le i_1,\dots,i_r\le n}L(\alpha_{i_1},\dots,\alpha_{i_r})f_{i_1}\otimes\cdots\otimes f_{i_r}.$$

证明. 注意到对$\beta\in V$有$\beta=\sum_{i=1}^n f_i(\beta)\alpha_i$. 于是, 对任意$L\in (V^*)^{\otimes r}$和$\beta_1,\dots,\beta_r\in V$有

$$\begin{aligned}L(\beta_1,\dots,\beta_r)&=L\left(\sum_{i_1=1}^n f_{i_1}(\beta_1)\alpha_{i_1},\dots,\sum_{i_r=1}^n f_{i_r}(\beta_r)\alpha_{i_r}\right)\\ &=\sum_{1\le i_1,\dots,i_r\le n}f_{i_1}(\beta_1)\cdots f_{i_r}(\beta_r)L(\alpha_{i_1},\dots,\alpha_{i_r})\\ &=\sum_{1\le i_1,\dots,i_r\le n}L(\alpha_{i_1},\dots,\alpha_{i_r})f_{i_1}\otimes\cdots\otimes f_{i_r}(\beta_1,\dots,\beta_r).\end{aligned}$$

这就证明了(2), 并且集合(5.15)生成$(V^*)^{\otimes r}$. 另一方面, 如果$c_{i_1,\dots,i_r}\in F$满足

$$\sum_{1\le i_1,\dots,i_r\le n} c_{i_1,\dots,i_r} f_{i_1}\otimes\cdots\otimes f_{i_r}=0,$$

则对任意$j_1,\dots,j_r\in\{1,\dots,n\}$, 考虑此式左边在$(\alpha_{j_1},\dots,\alpha_{j_r})$的取值, 即得$c_{j_1,\dots,j_r}=0$. 所以集合(5.15)线性无关, 因此是$(V^*)^{\otimes r}$的基. □

我们约定$(V^*)^{\otimes 0}=F$, 并考虑直和

$$\bigoplus_{r=0}^{\infty}(V^*)^{\otimes r}=\{(L_0,L_1,\dots)\mid L_r\in (V^*)^{\otimes r}, \text{只有有限个}r\text{使}L_r\neq 0\}.$$

通过把$L_r\in (V^*)^{\otimes r}$等同为$(0,\dots,0,L_r,0,\dots)$, 可以把所有$(V^*)^{\otimes r}$视为$\bigoplus_{r=0}^{\infty}(V^*)^{\otimes r}$的子空间. 于是

$$(L_0,L_1,\dots)=\sum_{r=0}^{\infty}(0,\dots,0,L_r,0,\dots)=\sum_{r=0}^{\infty}L_r.$$

注意到只有有限个L_r非零, 所以这里的求和是有限和. 我们在$\bigoplus_{r=0}^{\infty}(V^*)^{\otimes r}$上定义张量积为

$$\left(\sum_{r=0}^{\infty}L_r\right)\otimes\left(\sum_{s=0}^{\infty}M_s\right)=\sum_{r,s=0}^{\infty}L_r\otimes M_s.$$

则$\bigoplus_{r=0}^{\infty}(V^*)^{\otimes r}$成为一个域$F$上的代数, 称为$V^*$的**张量代数**(tensor algebra), 记为$\bigotimes(V^*)$或$T(V^*)$.

接下来我们考虑多重交错线性函数的乘法. 对于$L\in\Lambda^r(V^*)$, $M\in\Lambda^s(V^*)$, 张量积$L\otimes M$一般不是交错的. 为了定义交错的乘法, 我们首先考虑一种把一般多重线性函数化为交错函数的方法. 对于$L\in (V^*)^{\otimes r}$, 定义$\operatorname{Alt}(L)\in (V^*)^{\otimes r}$为

$$\operatorname{Alt}(L)(\alpha_1,\dots,\alpha_r)=\sum_{\sigma\in S_r}\operatorname{sgn}(\sigma)L(\alpha_{\sigma(1)},\dots,\alpha_{\sigma(r)}),\quad \alpha_1,\dots,\alpha_r\in V. \tag{5.16}$$

与引理5.7(1)的证明类似, 不难证明$\operatorname{Alt}(L)$是交错的, 称为L的**交错化**(alternation). 事实上, (5.6)式定义的$f_1\wedge\cdots\wedge f_n$即为$f_1\otimes\cdots\otimes f_n$的交错化. 也就是说, 我们有

引理 5.26. 设$f_1,\dots,f_n\in V^*$. 则

$$\operatorname{Alt}(f_1\otimes\cdots\otimes f_n)=\sum_{\sigma\in S_n}\operatorname{sgn}(\sigma)f_{\sigma(1)}\otimes\cdots\otimes f_{\sigma(n)}.$$

证明. 直接验证即得

$$\begin{aligned}\operatorname{Alt}(f_1\otimes\cdots\otimes f_n)(\alpha_1,\dots,\alpha_n)&=\sum_{\sigma\in S_n}\operatorname{sgn}(\sigma)f_1\otimes\cdots\otimes f_n(\alpha_{\sigma(1)},\dots,\alpha_{\sigma(n)})\\&=\sum_{\sigma\in S_n}\operatorname{sgn}(\sigma)f_1(\alpha_{\sigma(1)})\cdots f_n(\alpha_{\sigma(n)})\\&=\sum_{\sigma\in S_n}\operatorname{sgn}(\sigma)f_{\sigma(1)}(\alpha_1)\cdots f_{\sigma(n)}(\alpha_n)\\&=\sum_{\sigma\in S_n}\operatorname{sgn}(\sigma)f_{\sigma(1)}\otimes\cdots\otimes f_{\sigma(n)}(\alpha_1,\dots,\alpha_n).\end{aligned}$$

□

引理5.26启发我们尝试定义$L\in\Lambda^r(V^*)$与$M\in\Lambda^s(V^*)$的“交错乘积”为$\operatorname{Alt}(L\otimes M)$. 但是, 进一步分析以后, 我们发现$\operatorname{Alt}(L\otimes M)$的展开式中每项会重复$r!s!$次. 这是由$L$和$M$的交错性导致的. 另外, 这样定义的“交错乘积”只在差一个因子的意义下满足乘法结合律, 即对于$L\in\Lambda^r(V^*)$, $M\in\Lambda^s(V^*)$, $N\in\Lambda^t(V^*)$, 有

$$(s+t)!\operatorname{Alt}(\operatorname{Alt}(L\otimes M)\otimes N)=(r+s)!\operatorname{Alt}(L\otimes\operatorname{Alt}(M\otimes N)). \tag{5.17}$$

如果域F的特征为0, 有重复项是无关紧要的. 同时, 我们可以把L与M的"交错乘积"定义为

$$\frac{1}{(r+s)!}\mathrm{Alt}(L\otimes M) \tag{5.18}$$

或

$$\frac{1}{r!s!}\mathrm{Alt}(L\otimes M), \tag{5.19}$$

从而使它满足乘法结合律. 这两种定义都是被广泛采用的, 并且各有利弊. 注意(5.19) 中除掉的因子$r!s!$恰好起到了"消掉重复项"的作用, 并且对讨论行列式来说更加适合. 但是, 如果域F的特征非零, 因子$(r+s)!$或$r!s!$在F中一般不可逆, 从而(5.18)和(5.19)没有意义. 为了克服这一困难, 我们跳过交错化的过程, 直接采用"对每个重复项只加一次"的办法来定义交错乘法.

为此, 考虑S_{r+s}的子集

$$\mathrm{Sh}(r,s)=\{\sigma\in S_{r+s}\mid \sigma(1)<\cdots<\sigma(r), \sigma(r+1)<\cdots<\sigma(r+s)\}.$$

$\mathrm{Sh}(r,s)$中的置换称为**(r,s)-洗牌**((r,s)-shuffle)(谁有更好的翻译请告诉我). 容易看出,

$$|\mathrm{Sh}(r,s)|=\binom{r+s}{r}=\frac{(r+s)!}{r!s!}.$$

我们定义$L\wedge M\in (V^*)^{\otimes(r+s)}$为

$$L\wedge M(\alpha_1,\ldots,\alpha_{r+s})=\sum_{\sigma\in\mathrm{Sh}(r,s)}\mathrm{sgn}(\sigma)L(\alpha_{\sigma(1)},\ldots,\alpha_{\sigma(r)})M(\alpha_{\sigma(r+1)},\ldots,\alpha_{\sigma(r+s)}), \tag{5.20}$$

这里$\alpha_1,\ldots,\alpha_{r+s}\in V$.

例 5.13. 设$f_1,f_2\in V^*$. 则$f_1\wedge f_2$与例5.6中的记号和表达式一致. 注意到$f\wedge f=0$. (此性质在以后版本中应该加在例5.6中.)

例 5.14. 设$n>k\geq 1$. 容易看出,

$$\mathrm{Sh}(k,n-k)=\{\sigma_I\mid I\subset\{1,\ldots,n\}, |I|=k\}, \tag{5.21}$$

这里$\sigma_I\in S_n$是由(5.1)定义的置换. 特别地,

$$\mathrm{Sh}(n-1,1)=\{\sigma_i\mid 1\leq i\leq n\},$$

其中σ_i的定义如(5.12). 因此, 如果$L\in\Lambda^{n-1}(V^*)$, $f\in V^*$, 则

$$\begin{aligned}L\wedge f(\alpha_1,\ldots,\alpha_n)&=\sum_{i=1}^{n}\mathrm{sgn}(\sigma_i)L(\alpha_{\sigma_i(1)},\ldots,\alpha_{\sigma_i(n-1)})f(\alpha_{\sigma_i(n)}) \\ &=\sum_{i=1}^{n}(-1)^{n-i}L(\alpha_1,\ldots,\alpha_{i-1},\alpha_{i+1},\ldots,\alpha_n)f(\alpha_i).\end{aligned} \tag{5.22}$$

容易证明, 当$\mathrm{char}\,F=0$时, 由(5.20)定义的$L\wedge M$就是(5.19)式, 从而是交错的. 但是, 在一般情况下, 我们需要重新证明这一事实.

引理 5.27. 对于$L\in\Lambda^r(V^*)$, $M\in\Lambda^s(V^*)$, 由(5.20)给出的多重线性函数$L\wedge M$是交错的.

证明. 由命题5.5(1), 只需证明某两个相邻变量相同时, $L\wedge M$的取值是0. 设$\alpha_1,\ldots,\alpha_{r+s}\in V$, $\alpha_i=\alpha_{i+1}$, 这里$1\leq i\leq r+s-1$. 我们需要证明$L\wedge M(\alpha_1,\ldots,\alpha_{r+s})=0$. 考虑$\mathrm{Sh}(r,s)$的子集

$$\begin{aligned}P&=\{\sigma\in\mathrm{Sh}(r,s)\mid i\in\{\sigma(1),\ldots,\sigma(r)\}, i+1\in\{\sigma(r+1),\ldots,\sigma(r+s)\}\},\\ P'&=\{\sigma\in\mathrm{Sh}(r,s)\mid i\in\{\sigma(r+1),\ldots,\sigma(r+s)\}, i+1\in\{\sigma(1),\ldots,\sigma(r)\}\}.\end{aligned}$$

如果$\sigma\in\mathrm{Sh}(r,s)\smallsetminus(P\cup P')$, 则$i$和$i+1$同时落在$\{\sigma(1),\ldots,\sigma(r)\}$或$\{\sigma(r+1),\ldots,\sigma(r+s)\}$中. 此时, 由$L$和$M$的交错性有

$$L(\alpha_{\sigma(1)},\ldots,\alpha_{\sigma(r)})M(\alpha_{\sigma(r+1)},\ldots,\alpha_{\sigma(r+s)})=0.$$

另一方面, 考虑对换$\tau = (i, i+1)$. 容易看出, 如果$\sigma \in P$, 则$\tau\sigma \in \mathrm{Sh}(r,s)$, 而且进一步有

$$P' = \{\tau\sigma \mid \sigma \in P\}.$$

并且由于$\alpha_i = \alpha_{i+1}$, 当$\sigma \in P$时有

$$L(\alpha_{\sigma(1)}, \ldots, \alpha_{\sigma(r)}) = L(\alpha_{\tau\sigma(1)}, \ldots, \alpha_{\tau\sigma(r)}),$$
$$M(\alpha_{\sigma(r+1)}, \ldots, \alpha_{\sigma(r+s)}) = M(\alpha_{\tau\sigma(r+1)}, \ldots, \alpha_{\tau\sigma(r+s)}).$$

利用这些事实, 我们得到

$$\begin{aligned}
&L \wedge M(\alpha_1, \ldots, \alpha_{r+s}) \\
=& \sum_{\sigma \in P \cup P'} \mathrm{sgn}(\sigma) L(\alpha_{\sigma(1)}, \ldots, \alpha_{\sigma(r)}) M(\alpha_{\sigma(r+1)}, \ldots, \alpha_{\sigma(r+s)}) \\
=& \sum_{\sigma \in P} \mathrm{sgn}(\sigma) (L(\alpha_{\sigma(1)}, \ldots, \alpha_{\sigma(r)}) M(\alpha_{\sigma(r+1)}, \ldots, \alpha_{\sigma(r+s)}) - L(\alpha_{\tau\sigma(1)}, \ldots, \alpha_{\tau\sigma(r)}) M(\alpha_{\tau\sigma(r+1)}, \ldots, \alpha_{\tau\sigma(r+s)})) \\
=& 0.
\end{aligned}$$

这就完成了证明. □

由于引理5.27, 我们可以做出下面的定义.

定义 5.6. 设$L \in \Lambda^r(V^*)$, $M \in \Lambda^s(V^*)$. 由(5.20)给出的$L \wedge M \in \Lambda^{r+s}(V^*)$称为$L$与$M$的**外积**(exterior product或wedge product).

我们可以利用引理5.27给出行列式的Laplace展开定理的另一个证明.

定理5.19的另一证明. 这里只证明按k行展开的部分. 按k列展开的部分可以类似证明. 对于矩阵$B = [b_{ij}] \in F^{n \times l}$ $(1 \le l < n)$和行指标集$I \subset \{1, \ldots, n\}$, $|I| = l$, 记$B_I \in F^{l \times l}$为由B的矩阵元$\{b_{ij} \mid i \in I, 1 \le j \le l\}$构成的矩阵, 即$B_I$的矩阵元为

$$(B_I)_{rs} = b_{\sigma_I(r),s}, \quad r, s \in \{1, \ldots, l\}.$$

我们知道, 映射$(A_1, \ldots, A_n) \mapsto \det[A_1, \ldots, A_n]$是$F^{n \times 1}$上的$n$重交错线性函数. 因此, 对于取定的行指标集$I_0 \subset \{1, \ldots, n\}$, $|I_0| = k$, 映射

$$L : (F^{n \times 1})^k \to F, \quad L(B_1, \ldots, B_k) = \det[B_1, \ldots, B_k]_{I_0}$$

是$F^{n \times 1}$上的k重交错线性函数, 映射

$$M : (F^{n \times 1})^{n-k} \to F, \quad M(B_1, \ldots, B_{n-k}) \mapsto \det[B_1, \ldots, B_{n-k}]_{I_0^c}$$

是$F^{n \times 1}$上的$n-k$重交错线性函数. 由引理5.27, 外积$L \wedge M$是$F^{n \times 1}$上的n重交错线性函数. 而这样的n重交错线性函数空间的维数是1. 因此存在常数$c \in F$使得

$$L \wedge M(A_1, \ldots, A_n) = c \det[A_1, \ldots, A_n], \quad \forall A_1, \ldots, A_n \in F^{n \times 1}. \tag{5.23}$$

下面我们计算$L \wedge M$的表达式, 并且给出常数c. 由(5.21), 我们有

$$\begin{aligned}
L \wedge M(A_1, \ldots, A_n) &= \sum_{\substack{J \subset \{1, \ldots, n\} \\ |J| = k}} \mathrm{sgn}(\sigma_J) L(A_{\sigma_J(1)}, \ldots, A_{\sigma_J(k)}) M(A_{\sigma_J(k+1)}, \ldots, A_{\sigma_J(n)}) \\
&= \sum_{\substack{J \subset \{1, \ldots, n\} \\ |J| = k}} \mathrm{sgn}(\sigma_J) \det[A_{\sigma_J(1)}, \ldots, A_{\sigma_J(k)}]_{I_0} \det[A_{\sigma_J(k+1)}, \ldots, A_{\sigma_J(n)}]_{I_0^c} \\
&= \sum_{\substack{J \subset \{1, \ldots, n\} \\ |J| = k}} \mathrm{sgn}(\sigma_J) \det[A_1, \ldots, A_n]_{I_0, J} \det[A_1, \ldots, A_n]_{I_0^c, J^c}.
\end{aligned}$$

将此式代入(5.23), 得到

$$c\det(A)=\sum_{\substack{J\subset\{1,\dots,n\}\\|J|=k}}\operatorname{sgn}(\sigma_J)\det(A_{I_0,J})\det(A_{I_0^c,J^c}),\quad \forall A\in F^{n\times n}.$$

在上式中取$A=I_n$, 即得$c=\operatorname{sgn}(\sigma_{I_0})$. 再利用(5.2), 就完成了证明. □

我们继续讨论交错多重线性函数的外积.

命题 5.28. (1) 映射$\Lambda^r(V^*)\times\Lambda^s(V^*)\to\Lambda^{r+s}(V^*)$, $(L,M)\mapsto L\wedge M$是双线性的.

(2) 设$L\in\Lambda^r(V^*),M\in\Lambda^s(V^*),N\in\Lambda^t(V^*)$. 则

$$(L\wedge M)\wedge N=L\wedge(M\wedge N).$$

(3) 设$L\in\Lambda^r(V^*),M\in\Lambda^s(V^*)$. 则

$$M\wedge L=(-1)^{rs}L\wedge M.$$

证明. (1) 类似于命题5.24(1)的证明. 这里从略.

(2) 我们验证$(L\wedge M)\wedge N$和$L\wedge(M\wedge N)$在$(\alpha_1,\dots,\alpha_{r+s+t})\in V^{r+s+t}$的取值都等于

$$\sum_{\sigma\in\operatorname{Sh}(r,s,t)}\operatorname{sgn}(\sigma)L(\alpha_{\sigma(1)},\dots,\alpha_{\sigma(r)})M(\alpha_{\sigma(r+1)},\dots,\alpha_{\sigma(r+s)})N(\alpha_{\sigma(r+s+1)},\dots,\alpha_{\sigma(r+s+t)}),\quad(5.24)$$

这里

$$\operatorname{Sh}(r,s,t)=\{\sigma\in S_{r+s+t}\mid\sigma(1)<\cdots<\sigma(r),\sigma(r+1)<\cdots<\sigma(r+s),\sigma(r+s+1)<\cdots<\sigma(r+s+t)\}.$$

首先, 对于$\tau\in\operatorname{Sh}(r,s)$, 定义$\tilde\tau\in\operatorname{Sh}(r,s,t)$为

$$\tilde\tau(i)=\begin{cases}\tau(i), & 1\le i\le r+s,\\ i, & r+s+1\le i\le r+s+t.\end{cases}$$

则$\operatorname{sgn}(\tilde\tau)=\operatorname{sgn}(\tau)$. 容易看出, 如果$\sigma\in\operatorname{Sh}(r+s,t)$, $\tau\in\operatorname{Sh}(r,s)$, 则$\sigma\tilde\tau\in\operatorname{Sh}(r,s,t)$, 并且映射$\operatorname{Sh}(r+s,t)\times\operatorname{Sh}(r,s)\to\operatorname{Sh}(r,s,t)$, $(\sigma,\tau)\mapsto\sigma\tilde\tau$可逆. 从而

$$\begin{aligned}&(L\wedge M)\wedge N(\alpha_1,\dots,\alpha_{r+s+t})\\
=&\sum_{\sigma\in\operatorname{Sh}(r+s,t)}\operatorname{sgn}(\sigma)L\wedge M(\alpha_{\sigma(1)},\dots,\alpha_{\sigma(r+s)})N(\alpha_{\sigma(r+s+1)},\dots,\alpha_{\sigma(r+s+t)})\\
=&\sum_{\substack{\sigma\in\operatorname{Sh}(r+s,t)\\\tau\in\operatorname{Sh}(r,s)}}\operatorname{sgn}(\sigma)\operatorname{sgn}(\tau)L(\alpha_{\sigma(\tau(1))},\dots,\alpha_{\sigma(\tau(r))})M(\alpha_{\sigma(\tau(r+1))},\dots,\alpha_{\sigma(\tau(r+s))})N(\alpha_{\sigma(r+s+1)},\dots,\alpha_{\sigma(r+s+t)})\\
=&\sum_{\substack{\sigma\in\operatorname{Sh}(r+s,t)\\\tau\in\operatorname{Sh}(r,s)}}\operatorname{sgn}(\sigma\tilde\tau)L(\alpha_{\sigma\tilde\tau(1)},\dots,\alpha_{\sigma\tilde\tau(r)})M(\alpha_{\sigma\tilde\tau(r+1)},\dots,\alpha_{\sigma\tilde\tau(r+s)})N(\alpha_{\sigma\tilde\tau(r+s+1)},\dots,\alpha_{\sigma\tilde\tau(r+s+t)})\\
=&(5.24)\text{式}.\end{aligned}$$

类似地, 对$\tau\in\operatorname{Sh}(s,t)$, 定义$\hat\tau\in\operatorname{Sh}(r,s,t)$为

$$\hat\tau(i)=\begin{cases}i, & 1\le i\le r,\\ r+\tau(i-r), & r+1\le i\le r+s+t.\end{cases}$$

则$\mathrm{sgn}(\hat{\tau})=\mathrm{sgn}(\tau)$, 并且$(\sigma,\tau)\mapsto\sigma\hat{\tau}$是从$\mathrm{Sh}(r,s+t)\times\mathrm{Sh}(s,t)$到$\mathrm{Sh}(r,s,t)$的可逆映射. 因此

$$\begin{aligned}
&L\wedge(M\wedge N)(\alpha_1,\ldots,\alpha_{r+s+t})\\
=&\sum_{\sigma\in\mathrm{Sh}(r,s+t)}\mathrm{sgn}(\sigma)L(\alpha_{\sigma(1)},\ldots,\alpha_{\sigma(r)})M\wedge N(\alpha_{\sigma(r+1)},\ldots,\alpha_{\sigma(r+s+t)})\\
=&\sum_{\substack{\sigma\in\mathrm{Sh}(r,s+t)\\ \tau\in\mathrm{Sh}(s,t)}}\mathrm{sgn}(\sigma)\mathrm{sgn}(\tau)L(\alpha_{\sigma(1)},\ldots,\alpha_{\sigma(r)})M(\alpha_{\sigma(r+\tau(1))},\ldots,\alpha_{\sigma(r+\tau(s))})N(\alpha_{\sigma(r+\tau(s+1))},\ldots,\alpha_{\sigma(r+\tau(s+t))})\\
=&\sum_{\substack{\sigma\in\mathrm{Sh}(r,s+t)\\ \tau\in\mathrm{Sh}(s,t)}}\mathrm{sgn}(\sigma\hat{\tau})L(\alpha_{\sigma\hat{\tau}(1)},\ldots,\alpha_{\sigma\hat{\tau}(r)})M(\alpha_{\sigma\hat{\tau}(r+1)},\ldots,\alpha_{\sigma\hat{\tau}(r+s)})N(\alpha_{\sigma\hat{\tau}(r+s+1)},\ldots,\alpha_{\sigma\hat{\tau}(r+s+t)})\\
=&(5.24)\text{式}.
\end{aligned}$$

这就完成了证明.

(3) 考虑置换

$$\tau=\begin{pmatrix}1&\cdots&r&r+1&\cdots&r+s\\ s+1&\cdots&r+s&1&\cdots&s\end{pmatrix}\in S_{r+s}.$$

容易看出, $\mathrm{sgn}(\tau)=(-1)^{rs}$, 并且对于$\sigma\in S_{r+s}$, $\sigma\in\mathrm{Sh}(s,r)$的充分必要条件是$\sigma\tau\in\mathrm{Sh}(r,s)$. 从而对$\alpha_1,\ldots,\alpha_{r+s}\in V$, 我们有

$$\begin{aligned}
M\wedge L(\alpha_1,\ldots,\alpha_{r+s})&=\sum_{\sigma\in\mathrm{Sh}(s,r)}\mathrm{sgn}(\sigma)L(\alpha_{\sigma(s+1)},\ldots,\alpha_{\sigma(r+s)})M(\alpha_{\sigma(1)},\ldots,\alpha_{\sigma(s)})\\
&=\mathrm{sgn}(\tau)\sum_{\sigma\in\mathrm{Sh}(s,r)}\mathrm{sgn}(\sigma\tau)L(\alpha_{\sigma\tau(1)},\ldots,\alpha_{\sigma\tau(r)})M(\alpha_{\sigma\tau(r+1)},\ldots,\alpha_{\sigma\tau(r+s)})\\
&=(-1)^{rs}\sum_{\sigma\in\mathrm{Sh}(r,s)}\mathrm{sgn}(\sigma)L(\alpha_{\sigma(1)},\ldots,\alpha_{\sigma(r)})M(\alpha_{\sigma(r+1)},\ldots,\alpha_{\sigma(r+s)})\\
&=(-1)^{rs}L\wedge M(\alpha_1,\ldots,\alpha_{r+s}).
\end{aligned}$$

因此$M\wedge L=(-1)^{rs}L\wedge M$. □

由外积的乘法结合律, 即命题5.28(2), 对于$L_1\in\Lambda^{r_1}(V^*),\ldots,L_n\in\Lambda^{r_n}(V^*)$, 可以定义它们的外积$L_1\wedge\cdots\wedge L_n$, 所得结果与做外积的次序无关. 如果$L_i$都是线性函数, 它们的外积有下面的表达式.

命题 5.29. 设$f_1,\ldots,f_n\in V^*$. 则

$$f_1\wedge\cdots\wedge f_n=\mathrm{Alt}(f_1\otimes\cdots\otimes f_n)=\sum_{\sigma\in S_n}\mathrm{sgn}(\sigma)f_{\sigma(1)}\otimes\cdots\otimes f_{\sigma(n)}.$$

证明. 第二个等式就是引理5.26. 我们用归纳法证明第一个等式. 当$n=1$时无需证明. 假设$n\geq 2$, 并且第一个等式对$n-1$成立. 对于$1\leq i\leq n$, 记

$$P_i=\{\sigma\in S_n\mid\sigma(n)=i\}.$$

则

$$P_i=\{\sigma_i\tilde{\tau}\mid\tau\in S_{n-1}\},$$

这里σ_i由(5.12)定义, $\tilde{\tau}$由(5.13)定义. 由(5.22)式和归纳假设, 对$\alpha_1,\ldots,\alpha_n\in V$有

$$\begin{aligned}
&f_1\wedge\cdots\wedge f_n(\alpha_1,\ldots,\alpha_n)\\
=&\sum_{i=1}^n \operatorname{sgn}(\sigma_i)\operatorname{Alt}(f_1\otimes\cdots\otimes f_{n-1})(\alpha_{\sigma_i(1)},\ldots,\alpha_{\sigma_i(n-1)})f_n(\alpha_{\sigma_i(n)})\\
=&\sum_{i=1}^n \operatorname{sgn}(\sigma_i)\sum_{\tau\in S_{n-1}}\operatorname{sgn}(\tau)f_1\otimes\cdots\otimes f_{n-1}(\alpha_{\sigma_i(\tau(1))},\ldots,\alpha_{\sigma_i(\tau(n-1))})f_n(\alpha_{\sigma_i(n)})\\
=&\sum_{i=1}^n\sum_{\tau\in S_{n-1}}\operatorname{sgn}(\sigma_i\tilde{\tau})f_1\otimes\cdots\otimes f_n(\alpha_{\sigma_i\tilde{\tau}(1)},\ldots,\alpha_{\sigma_i\tilde{\tau}(n)})\\
=&\sum_{i=1}^n\sum_{\sigma\in P_i}\operatorname{sgn}(\sigma)f_1\otimes\cdots\otimes f_n(\alpha_{\sigma(1)},\ldots,\alpha_{\sigma(n)})\\
=&\operatorname{Alt}(f_1\otimes\cdots\otimes f_n)(\alpha_1,\ldots,\alpha_n).
\end{aligned}$$

这就完成了证明. □

命题5.29说明, 我们在(5.6)中引入的n重交错线性函数即为$f_1,\ldots,f_n$的外积. 那里的记号$f_1\wedge\cdots\wedge f_n$与这里是一致的.

推论 5.30. 设$f_1,\ldots,f_n\in V^*$. 则

(1) 对任意$\sigma\in S_n$有

$$f_{\sigma(1)}\wedge\cdots\wedge f_{\sigma(n)}=\operatorname{sgn}(\sigma)f_1\wedge\cdots\wedge f_n.$$

(2) 对任意$\alpha_1,\ldots,\alpha_n\in V$有

$$f_1\wedge\cdots\wedge f_n(\alpha_1,\ldots,\alpha_n)=\det[f_i(\alpha_j)],$$

这里$[f_i(\alpha_j)]\in F^{n\times n}$是$(i,j)$-元为$f_i(\alpha_j)$的方阵.

证明. (1) 由命题5.29, 有

$$\begin{aligned}
f_{\sigma(1)}\wedge\cdots\wedge f_{\sigma(n)}&=\sum_{\tau\in S_n}\operatorname{sgn}(\tau)f_{\sigma\tau(1)}\otimes\cdots\otimes f_{\sigma\tau(n)}\\
&=\sum_{\tau\in S_n}\operatorname{sgn}(\sigma^{-1}\tau)f_{\tau(1)}\otimes\cdots\otimes f_{\tau(n)}\\
&=\operatorname{sgn}(\sigma)f_1\wedge\cdots\wedge f_n.
\end{aligned}$$

我们也可以对多重交错线性映射$(V^*)^n\mapsto\Lambda^n(V^*)$, $(f_1,\ldots,f_n)\mapsto f_1\wedge\cdots\wedge f_n$应用命题5.5(2)的推广来证明(1).

(2) 利用方阵行列式的表达式(5.10), 即得

$$\begin{aligned}
f_1\wedge\cdots\wedge f_n(\alpha_1,\ldots,\alpha_n)&=\sum_{\sigma\in S_n}\operatorname{sgn}(\sigma)f_{\sigma(1)}\otimes\cdots\otimes f_{\sigma(n)}(\alpha_1,\ldots,\alpha_n)\\
&=\sum_{\sigma\in S_n}\operatorname{sgn}(\sigma)f_{\sigma(1)}(\alpha_1)\cdots f_{\sigma(n)}(\alpha_n)=\det[f_i(\alpha_j)].
\end{aligned}$$

□

上面的内容可以帮助我们进一步理解Laplace展开定理的第二个证明. 设$\{f_1,\ldots,f_n\}$是$F^{n\times 1}$的标准基在$(F^{n\times 1})^*$中的对偶基. 取定指标集$I_0\subset\{1,\ldots,n\}$, $|I_0|=k$. 设

$$I_0=\{i_1,\ldots,i_k\},\quad i_1<\cdots<i_k,$$

$$I_0^c = \{i'_1, \ldots, i'_{n-k}\}, \quad i'_1 < \cdots < i'_{n-k}.$$

由推论5.30(1), 我们有

$$f_1 \wedge \cdots \wedge f_n = \operatorname{sgn}(\sigma_{I_0})(f_{i_1} \wedge \cdots \wedge f_{i_k}) \wedge (f_{i'_1} \wedge \cdots \wedge f_{i'_{n-k}}).$$

容易看出, $f_{i_1} \wedge \cdots \wedge f_{i_k}$和$f_{i'_1} \wedge \cdots \wedge f_{i'_{n-k}}$ 分别是证明中的函数L和M, 而上式就是(5.23)式. 也就是说, 我们把$f_{i_1} \wedge \cdots \wedge f_{i_k}$与$f_{i'_1} \wedge \cdots \wedge f_{i'_{n-k}}$的外积按定义做展开, 就得到了Laplace展开定理.

接下来我们给出$\Lambda^r(V^*)$的基和维数.

定理 5.31. 设V是有限维线性空间, $\dim V = n$, $\{f_1, \ldots, f_n\}$是V^*的基.

(1) 如果$1 \le r \le n$, 则

$$\{f_{i_1} \wedge \cdots \wedge f_{i_r} \mid 1 \le i_1 < \cdots < i_r \le n\} \tag{5.25}$$

是$\Lambda^r(V^*)$的基. 特别地, $\dim \Lambda^r(V^*) = \binom{n}{r}$.

(2) 如果$r > n$, 则$\Lambda^r(V^*) = \{0\}$.

证明. 设$L \in \Lambda^r(V^*)$, $\{\alpha_1, \ldots, \alpha_n\}$是$\{f_1, \ldots, f_n\}$在$V$中的对偶基. 由定理5.25(2),

$$L = \sum_{1 \le i_1, \ldots, i_r \le n} L(\alpha_{i_1}, \ldots, \alpha_{i_r}) f_{i_1} \otimes \cdots \otimes f_{i_r}.$$

如果$r > n$, 则集合$\{1, \ldots, n\}$中的数字$i_1, \ldots, i_r$总会有两个相同, 从而$L(\alpha_{i_1}, \ldots, \alpha_{i_r}) = 0$. 因此$L = 0$. 这就证明了(2). 假设$1 \le r \le n$. 则

$$\begin{aligned}
L &= \sum_{\substack{i_1, \ldots, i_r \in \{1, \ldots, n\} \\ \text{并且互不相同}}} L(\alpha_{i_1}, \ldots, \alpha_{i_r}) f_{i_1} \otimes \cdots \otimes f_{i_r} \\
&= \sum_{1 \le i_1 < \cdots < i_r \le n} \sum_{\sigma \in S_r} L(\alpha_{i_{\sigma(1)}}, \ldots, \alpha_{i_{\sigma(r)}}) f_{i_{\sigma(1)}} \otimes \cdots \otimes f_{i_{\sigma(r)}} \\
&= \sum_{1 \le i_1 < \cdots < i_r \le n} L(\alpha_{i_1}, \ldots, \alpha_{i_r}) \sum_{\sigma \in S_r} \operatorname{sgn}(\sigma) f_{i_{\sigma(1)}} \otimes \cdots \otimes f_{i_{\sigma(r)}} \\
&= \sum_{1 \le i_1 < \cdots < i_r \le n} L(\alpha_{i_1}, \ldots, \alpha_{i_r}) f_{i_1} \wedge \cdots \wedge f_{i_r}.
\end{aligned}$$

因此集合(5.25)生成$\Lambda^r(V^*)$. 另一方面, 集合(5.15)线性无关推出集合(5.25)也线性无关. 这就证明了(5.25)是$\Lambda^r(V^*)$的基. □

我们约定$\Lambda^0(V^*) = F$, 并考虑直和

$$\bigoplus_{r=0}^{\infty} \Lambda^r(V^*) = \{(L_0, L_1, \ldots) \mid L_r \in \Lambda^r(V^*), \text{只有有限个}r\text{使}L_r \neq 0\}.$$

与张量代数的情况类似, 我们把所有$\Lambda^r(V^*)$视为$\bigoplus_{r=0}^{\infty} \Lambda^r(V^*)$的子空间, 并且在$\bigoplus_{r=0}^{\infty} \Lambda^r(V^*)$上定义外积为

$$\left(\sum_{r=0}^{\infty} L_r\right) \wedge \left(\sum_{s=0}^{\infty} M_s\right) = \sum_{r,s=0}^{\infty} L_r \wedge M_s.$$

则$\bigoplus_{r=0}^{\infty} \Lambda^r(V^*)$成为一个域$F$上的代数, 称为$V^*$的**外代数**(exterior algebra)或**Grassmann代数**, 记为$\Lambda(V^*)$. 如果V是有限维的, $\dim V = n$, 则由定理5.31, 我们有$\Lambda(V^*) = \bigoplus_{r=0}^{n} \Lambda^r(V^*)$, 并且

$$\dim \Lambda(V^*) = \sum_{r=0}^{n} \dim \Lambda^r(V^*) = \sum_{r=0}^{n} \binom{n}{r} = 2^n.$$

最后, 我们证明一个线性函数外积的重要性质, 即外积是否为零可以用来判断这些线性函数是否线性相关.

定理 5.32. 设$f_1,\dots,f_n\in V^*$. 则$f_1\wedge\cdots\wedge f_n\neq 0$的充分必要条件是$\{f_1,\dots,f_n\}$线性无关.

证明. "$\Longrightarrow$". 假设$f_1\wedge\cdots\wedge f_n\neq 0$. 则存在$\alpha_1,\dots,\alpha_n\in V$使$f_1\wedge\cdots\wedge f_n(\alpha_1,\dots,\alpha_n)\neq 0$. 由推论5.30(2), 这推出矩阵$[f_i(\alpha_j)]$ 可逆. 如果$c_1,\dots,c_n\in F$满足$\sum_{i=1}^n c_if_i=0$, 则

$$(c_1,\dots,c_n)[f_i(\alpha_j)]=\Big(\sum_{i=1}^n c_if_i(\alpha_1),\dots,\sum_{i=1}^n c_if_i(\alpha_n)\Big)=0.$$

所以$(c_1,\dots,c_n)=0$. 因此$\{f_1,\dots,f_n\}$线性无关.

"$\Longleftarrow$". 假设$\{f_1,\dots,f_n\}$线性无关. 我们先证明线性映射

$$T:V\to F^n,\quad T(\alpha)=(f_1(\alpha),\dots,f_n(\alpha))$$

是满射. 为此, 只需证明$\mathrm{Im}(T)$在$(F^n)^*$中的零化子空间$\mathrm{Im}(T)^0$为零子空间. 设$g\in\mathrm{Im}(T)^0$. 由于g是F^n上的线性函数, 所以存在$c_1,\dots,c_n\in F$使得

$$g(x_1,\dots,x_n)=\sum_{i=1}^n c_ix_i,\quad \forall(x_1,\dots,x_n)\in F^n.$$

而$g\in\mathrm{Im}(T)^0$说明$g\circ T=0$, 即对任意$\alpha\in V$有

$$g(T(\alpha))=\sum_{i=1}^n c_if_i(\alpha)=0,$$

也就是说$\sum_{i=1}^n c_if_i=0$. 由于$\{f_1,\dots,f_n\}$线性无关, 这推出$c_1=\cdots=c_n=0$, 即$g=0$. 这就证明了$\mathrm{Im}(T)^0=0$. 因此T是满射. 现在, 考虑F^n的标准基$\{\epsilon_1,\dots,\epsilon_n\}$. 由于$T$是满射, 存在$\alpha_1,\dots,\alpha_n\in V$ 使得$T(\alpha_j)=\epsilon_j$, 即$f_i(\alpha_j)=\delta_{ij}$. 从而矩阵$[f_i(\alpha_j)]$为单位矩阵. 由推论5.30(2), 我们得到$f_1\wedge\cdots\wedge f_n(\alpha_1,\dots,\alpha_n)=1$. 因此$f_1\wedge\cdots\wedge f_n\neq 0$. □

下面的推论是显然的.

推论 5.33. 设$\dim V=n$, $f_1,\dots,f_n\in V^*$. 则$f_1\wedge\cdots\wedge f_n\neq 0$的充分必要条件是$\{f_1,\dots,f_n\}$是$V^*$的基.

我们利用推论5.9给出推论5.33的另一个证明.

证明. 注意到$\dim\Lambda^n(V^*)=1$. 取定同构映射$\phi:\Lambda^n(V^*)\to F$. 则映射

$$L:(V^*)^n\to F,\quad L(f_1,\dots,f_n)=\phi(f_1\wedge\cdots\wedge f_n)$$

是V^*上的n重线性函数. 由于对任意$f\in V^*$有$f\wedge f=0$, 所以L在相邻变量相同时取值是零. 由命题5.5(1), L是交错的. 另一方面, 由引理5.7(2), L不恒等于零. 总之, 我们有$L\in\Lambda^n(V^{**})\setminus\{0\}$. 由推论5.9, 即得

$$\{f_1,\dots,f_n\}\text{是}V^*\text{的基}\quad\Longleftrightarrow\quad L(f_1,\dots,f_n)\neq 0\quad\Longleftrightarrow\quad f_1\wedge\cdots\wedge f_n\neq 0.$$

这就完成了证明. □

习题 5.6.

1. 验证(5.16)中定义的$\mathrm{Alt}(L)$是交错的.
2. 不用上题结果, 直接验证(5.16)中定义的$\mathrm{Alt}(L)$是反对称的.
3. 不用命题5.28(2)直接验证(5.17)式.
4. 假设$\mathrm{char}\,F=0$. 利用(5.17)验证(5.18)和(5.19)中定义的"交错乘积"都满足乘法结合率.
5. 假设$\mathrm{char}\,F=0$. 验证(5.20)中定义的$L\wedge M$与(5.19)式一致.

6. 设r是奇数, $L \in \Lambda^r(V^*)$. 证明$L \wedge L = 0$.
7. 设$f_1, \ldots, f_r \in V^*$线性无关. 证明对$f \in V^*$, $f \in \operatorname{span}\{f_1, \ldots, f_r\}$的充分必要条件是$f \wedge f_1 \wedge \cdots \wedge f_r = 0$.
8. 设$g_1, \ldots, g_r \in W^*$, $T \in \mathrm{L}(V, W)$. 证明
$$\Lambda^r(T^t)(g_1 \wedge \cdots \wedge g_r) = T^t g_1 \wedge \cdots \wedge T^t g_r.$$
9. 设$\dim V = n$, $f_1, \ldots, f_n \in V^*$, $T \in \mathrm{L}(V, V)$. 证明
$$T^t f_1 \wedge \cdots \wedge T^t f_n = \det(T) f_1 \wedge \cdots \wedge f_n.$$

第六章　初等标准形

§6.1　对角化

除非做出特别说明, 我们总是假定F是任意一个域, V是F上的有限维线性空间. 我们考虑下面的问题:

- 给定线性变换$T \in L(V)$, 寻找V的有序基$\mathcal{B}$, 使得矩阵$[T]_{\mathcal{B}}$尽量简单.
- 给定矩阵$A \in F^{n\times n}$, 寻找可逆矩阵$P \in \mathrm{GL}_n(F)$, 使得矩阵$P^{-1}AP$尽量简单.

矩阵$[T]_{\mathcal{B}}$和$P^{-1}AP$越简单, 关于T和A的很多问题(例如求秩、行列式、核、像集等)就越容易回答.

这两个问题在本质上是同一个问题. 事实上, T在V的所有有序基下的矩阵构成一个相似等价类. 我们选择对线性变换讨论问题. 对于矩阵A的情况, 我们把对线性变换的结果应用到L_A来回答.

我们先把“简单”的矩阵取为对角矩阵.

定义 6.1. (1) 如果$T \in L(V)$在V的某个有序基下的矩阵是对角矩阵, 则称T是**可对角化**的.

(2) 如果$A \in F^{n\times n}$相似于对角矩阵, 则称A是**可对角化**的.

容易看出, A可对角化$\Longleftrightarrow L_A$可对角化.

是否每个T总可对角化? 如果不是, 如何判断T是否可对角化? 我们引入下面的概念.

定义 6.2. 设$T \in L(V)$. 对于$c \in F$, 我们记

$$V_c := \mathrm{Ker}(c\,\mathrm{id}_V - T) = \{\alpha \in V \mid T\alpha = c\alpha\}.$$

如果$V_c \neq \{0\}$, 则称c是T的**特征值**(eigenvalue), V_c是T关于特征值c的**特征子空间**(eigenspace), V_c中的非零向量是T关于特征值c的**特征向量**(eigenvector). T的特征值的集合称为T的**谱**(spectrum), 记为$\sigma(T)$.

命题 6.1. 设$T \in L(V)$, $\mathcal{B}$是V的有序基. 则$[T]_{\mathcal{B}}$是对角矩阵$\Longleftrightarrow$ $\mathcal{B}$中的向量都是T的特征向量. 因此, T可对角化$\Longleftrightarrow$存在由T的特征向量构成的基.

证明. 设$\mathcal{B} = \{\alpha_1, \dots, \alpha_n\}$, $A = [T]_{\mathcal{B}}$. 则

$$T\alpha_j = \sum_{i=1}^{n} A_{ij}\alpha_i.$$

因此, $[T]_{\mathcal{B}}$是对角矩阵$\Longleftrightarrow$存在$c_j \in F$使得$T\alpha_j = c_j\alpha_j \Longleftrightarrow$每个$\alpha_j$是$T$的特征向量. □

对于矩阵, 我们引入类似的概念.

定义 6.3. 设$A \in F^{n\times n}$. 对于$c \in F$, 考虑$F^{n\times 1}$的子空间

$$\mathrm{Ker}(cI - A) = \{X \in F^{n\times 1} \mid AX = cX\}.$$

如果$\mathrm{Ker}(cI - A) \neq \{0\}$, 则称$c$是$A$的**特征值**, $\mathrm{Ker}(cI - A)$是A关于特征值c的**特征子空间**, 其中的非零向量是A关于特征值c的**特征向量**. A的特征值的集合称为A的**谱**, 记为$\sigma(A)$.

下面讨论寻找特征值的方法.

引理 6.2. (1) 设$T \in L(V)$. 则

$$\sigma(T) = \{c \in F \mid \det(c\,\mathrm{id}_V - T) = 0\}.$$

(2) 设$A \in F^{n\times n}$. 则

$$\sigma(A) = \{c \in F \mid \det(cI - A) = 0\}.$$

证明. (1) $c \in \sigma(T) \iff \mathrm{Ker}(c\,\mathrm{id}_V - T) \neq 0 \iff c\,\mathrm{id}_V - T$不可逆$\iff \det(c\,\mathrm{id}_V - T) = 0$.

(2) 在(1)中取$T = L_A$. □

容易看出, 映射$F \to F$, $c \mapsto \det(c\,\mathrm{id}_V - T)$是多项式函数, 特征值即为它的"根". 这启发我们考虑"多项式"$f = \det(x\,\mathrm{id}_V - T)$. 但是, 根据我们对多项式的定义, x是未定元, 并不是域F中的元素. 这时, 表达式$x\,\mathrm{id}_V - T$是没有意义的. 事实上, 如果F是有限域, 满足对任意$c \in F$有$f(c) = \det(c\,\mathrm{id}_V - T)$的多项式$f$并不唯一. 我们利用矩阵来解决这一问题[1].

定义 6.4. 设$A \in F^{n\times n}$. 考虑$xI - A \in F(x)^{n\times n}$(这里$F(x)$见附录). 我们称

$$f_A := \det(xI - A) \in F[x]$$

为A的**特征多项式**(characteristic polynomial).

容易看出, 矩阵的特征多项式具有下面的性质:

- f_A的首项系数是1(这是我们写$xI-A$而不是$A-xI$的原因), $\deg f_A = n$, 常数项为$(-1)^n\det(A)$, x^{n-1}的系数为$-\mathrm{tr}(A)$.
- A的特征值即为f_A的根. 因此$|\sigma(A)| \leq n$.
- 相似的矩阵有相同的特征多项式. 事实上, 若$B = P^{-1}AP$, 则$xI - B = P^{-1}(xI - A)P$, 因此$xI - B$与$xI - A$的行列式相等.
- 可以证明AB与BA有相同的特征多项式:

$$\begin{bmatrix} xI & A \\ B & I \end{bmatrix}\begin{bmatrix} I & 0 \\ -B & xI \end{bmatrix} = \begin{bmatrix} xI - AB & xA \\ 0 & xI \end{bmatrix},$$
$$\begin{bmatrix} I & 0 \\ -B & xI \end{bmatrix}\begin{bmatrix} xI & A \\ B & I \end{bmatrix} = \begin{bmatrix} xI & A \\ 0 & xI - BA \end{bmatrix},$$

因此$x^n\det(xI - AB) = x^n\det(xI - BA)$. 还可以这样证明: 设$A = P\begin{bmatrix} I_r & 0 \\ 0 & 0 \end{bmatrix}Q$, P, Q可逆. 设$B = Q^{-1}\begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix}P^{-1}$. 则

$$AB = P\begin{bmatrix} B_{11} & B_{12} \\ 0 & 0 \end{bmatrix}P^{-1}, \qquad BA = Q^{-1}\begin{bmatrix} B_{11} & 0 \\ B_{21} & 0 \end{bmatrix}Q.$$

于是$f_{AB} = f_{BA} = x^{n-r}f_{B_{11}}$.

定义 6.5. 设$T \in L(V)$, $\mathcal{B}$是V的有序基. 定义T的**特征多项式**为$f_T = f_{[T]_{\mathcal{B}}}$.

对于不同的$\mathcal{B}$, $[T]_{\mathcal{B}}$是相似的, 从而有相同的$f_{[T]_{\mathcal{B}}}$. 因此f_T的定义不依赖于$\mathcal{B}$的选取. 我们有:

- f_T的首项系数是1, $\deg f_T = \dim V$.

[1]另一解决途径如下: 考虑张量积$V \otimes_F F(x)$, 它是域$F(x)$上的线性空间, $x\,\mathrm{id} - T$可以视为$V \otimes_F F(x)$上的线性变换, 于是可以定义T的特征多项式为$f_T = \det(x\,\mathrm{id} - T)$.

- 由于$\sigma(T)=\sigma([T]_{\mathcal{B}})$, 所以$T$的特征值就是$f_T$的根, 因此$|\sigma(T)|\le\dim V$.
- 如果T可对角化, $[T]_{\mathcal{B}}=\operatorname{diag}(c_1,\ldots,c_n)$, 则$f_T=(x-c_1)\cdots(x-c_n)$. 因此对角矩阵$[T]_{\mathcal{B}}$的对角元在差一个排列的意义下被$T$决定. 另一方面, 通过调整有序基$\mathcal{B}$中向量的排列, 可以得到对角矩阵$[T]_{\mathcal{B}}$的对角元的任意排列.

为了讨论特征子空间的性质, 我们插入关于直和的部分内容(见教材§6.6). 设$V_1,\ldots,V_k$是域F上的线性空间(可以是无穷维的). 考虑集合

$$V_1\times\cdots\times V_k:=\{(\alpha_1,\ldots,\alpha_k)\mid\alpha_i\in V_i\}.$$

这里$(\alpha_1,\ldots,\alpha_k)$为向量的$n$元有序组. 严格地说, 它指映射$\{1,\ldots,k\}\to\bigcup_{i=1}^k V_i$, $i\mapsto\alpha_i$. 在$V_1\times\cdots\times V_k$上定义加法和纯量乘法为

$$(\alpha_1,\ldots,\alpha_k)+(\beta_1,\ldots,\beta_k)=(\alpha_1+\beta_1,\ldots,\alpha_k+\beta_k),$$
$$c(\alpha_1,\ldots,\alpha_k)=(c\alpha_1,\ldots,c\alpha_k).$$

则$V_1\times\cdots\times V_k$成为域$F$上的线性空间, 称为$V_1,\ldots,V_k$的**外直和**(external direct sum), 记为$V_1\oplus\cdots\oplus V_k$或$\bigoplus_{i=1}^k V_i$. 对任意$i$, 考虑单线性映射

$$\tau_i: V_i\to\bigoplus_{i=1}^k V_i,\quad \tau_i(\alpha_i)=(0,\ldots,\alpha_i,\ldots,0). \tag{6.1}$$

引理 6.3. 如果$\mathcal{B}_i$是V_i的基, 则$\bigcup_{i=1}^k\tau_i(\mathcal{B}_i)$是$\bigoplus_{i=1}^k V_i$的基. 从而如果$\dim V_i<\infty$, 则$\dim\bigoplus_{i=1}^k V_i<\infty$, 并且

$$\dim\bigoplus_{i=1}^k V_i=\sum_{i=1}^k\dim V_i.$$

证明. 首先, 对任意$(\alpha_1,\ldots,\alpha_k)\in\bigoplus_{i=1}^k V_i$, 由于$\alpha_i\in\operatorname{span}\mathcal{B}_i$, 所以$\tau_i(\alpha_i)\in\operatorname{span}\tau_i(\mathcal{B}_i)$, 从而

$$(\alpha_1,\ldots,\alpha_k)=\sum_{i=1}^k\tau_i(\alpha_i)\in\operatorname{span}\bigcup_{i=1}^k\tau_i(\mathcal{B}_i).$$

因此$\bigcup_{i=1}^k\tau_i(\mathcal{B}_i)$生成全空间. 其次, 如果$\alpha_{ij}\in\mathcal{B}_i$, 并且有限线性组合$\sum_{i=1}^k\sum_j c_{ij}\tau_i(\alpha_{ij})=0$, 由于

$$\sum_{i=1}^k\sum_j c_{ij}\tau_i(\alpha_{ij})=\sum_{i=1}^k\tau_i\left(\sum_j c_{ij}\alpha_{ij}\right)=\left(\sum_j c_{1j}\alpha_{1j},\ldots,\sum_j c_{kj}\alpha_{kj}\right),$$

所以$\sum_j c_{ij}\alpha_{ij}=0$, 从而$c_{ij}=0$. 因此$\bigcup_{i=1}^k\tau_i(\mathcal{B}_i)$线性无关. □

现在设$V_1,\ldots,V_k$是域F上的线性空间V的子空间. 考虑从外直和$\bigoplus_{i=1}^k V_i$到V的线性映射

$$\Phi:\bigoplus_{i=1}^k V_i\to V,\quad(\alpha_1,\ldots,\alpha_k)\mapsto\sum_{i=1}^k\alpha_i.$$

容易看出, $\operatorname{Im}(\Phi)=\sum_{i=1}^k V_i$. 如果$\Phi$是单映射, 则称子空间$V_1,\ldots,V_k$是**无关**的, 并称它们的和$\sum_{i=1}^k V_i$为$V_1,\ldots,V_k$的**内直和**(internal direct sum). 此时, Φ诱导了外直和与内直和之间的线性同构

$$\bigoplus_{i=1}^k V_i\cong\sum_{i=1}^k V_i. \tag{6.2}$$

引理 6.4. 设$V_1,\ldots,V_k\subset V$是子空间. 以下断言等价:

(1) $V_1,\ldots,V_k$无关.

(2) $\alpha_i\in V_i,\sum_{i=1}^k\alpha_i=0\Longrightarrow\alpha_i=0$.

(3) 对任意$2 \le i \le k$有$V_i \cap (V_1 + \cdots + V_{i-1}) = \{0\}$.

(4) 任给V_i的基$\mathcal{B}_i$, 总有$\mathcal{B}_1, \ldots, \mathcal{B}_k$互不相交, 并且$\bigcup_{i=1}^k \mathcal{B}_i$是$\sum_{i=1}^k V_i$的基.

如果$\dim V < \infty$, 则它们还等价于

(5) $\dim \sum_{i=1}^k V_i = \sum_{i=1}^k \dim V_i$.

证明. "(1)$\Longleftrightarrow$(2)". 显然.

"(1)$\Longrightarrow$(3)". 设$\alpha \in V_i \cap (V_1 + \cdots + V_{i-1})$. 则存在$\alpha_1 \in V_1, \ldots, \alpha_{i-1} \in V_{i-1}$使$\alpha = \alpha_1 + \cdots + \alpha_{i-1}$, 即$\Phi(\alpha_1, \ldots, \alpha_{i-1}, -\alpha, 0, \ldots, 0) = 0$. 由于$\Phi$是单射, 所以$\alpha = 0$.

"(3)$\Longrightarrow$(2)". 设$\alpha_i \in V_i$, $\sum_{i=1}^k \alpha_i = 0$. 如果某个$\alpha_i \ne 0$, 取最大指标$i_0$使$\alpha_{i_0} \ne 0$. 则$i_0 \ge 2$, 并且

$$\alpha_{i_0} = -\sum_{i=1}^{i_0-1} \alpha_i \in V_{i_0} \cap (V_1 + \cdots + V_{i_0-1}) = \{0\},$$

矛盾. 因此所有$\alpha_i = 0$.

"(1)$\Longleftrightarrow$(4)". 注意Φ把$\bigcup_{i=1}^k \tau_i(\mathcal{B}_i)$映为$\bigcup_{i=1}^k \mathcal{B}_i$. 所以$\Phi$单$\Longleftrightarrow\mathcal{B}_1, \ldots, \mathcal{B}_k$互不相交并且$\bigcup_{i=1}^k \mathcal{B}_i$线性无关.

"(1)$\Longleftrightarrow$(5)". 我们总有

$$\dim \sum_{i=1}^k V_i = \dim \mathrm{Im}(\Phi) = \dim \bigoplus_{i=1}^k V_i - \dim \mathrm{Ker}(\Phi) = \sum_{i=1}^k \dim V_i - \dim \mathrm{Ker}(\Phi).$$

因此Φ单$\Longleftrightarrow$(5)成立. □

注意当$k = 2$时, V_1与V_2无关$\Longleftrightarrow V_1 \cap V_2 = \{0\}$; 当$k > 2$时, $V_1, \ldots, V_k$无关$\Longrightarrow \bigcap_{i=1}^k V_i = \{0\}$, 但反过来不对.

内直和与外直和有非常紧密的联系. 首先, 给定线性空间$V_1, \ldots, V_k$, 我们可以利用(6.1)中定义的单映射把$\alpha_i \in V_i$与$(0, \ldots, \alpha_i, \ldots, 0) \in \bigoplus_{i=1}^k V_i$等同起来, 从而把$V_i$视为$\bigoplus_{i=1}^k V_i$的子空间. 这样, 我们有$(\alpha_1, \ldots, \alpha_k) = \sum_{i=1}^k \alpha_i$. 注意到$\sum_{i=1}^k \alpha_i = 0 \Longrightarrow \alpha_i = 0$, 即子空间$V_1, \ldots, V_k$是无关的, 并且$\bigoplus_{i=1}^k V_i = \sum_{i=1}^k V_i$. 因此, 在这一观点下, 外直和$\bigoplus_{i=1}^k V_i$成为它的子空间$V_1, \ldots, V_k$的内直和. 另一方面, 给定线性空间$V$的无关的子空间$V_1, \ldots, V_k$, 同构(6.2)说明内直和$\sum_{i=1}^k V_i$中的任意向量$\alpha$可以唯一分解成$\alpha = \sum_{i=1}^k \alpha_i$, $\alpha_i \in V_i$. 我们利用同构(6.2)把α与$(\alpha_1, \ldots, \alpha_k)$等同起来, 从而把内直和$\sum_{i=1}^k V_i$等同为外直和$\bigoplus_{i=1}^k V_i$. 基于上述原因, 我们把内直和$\sum_{i=1}^k V_i$也记为$\bigoplus_{i=1}^k V_i$, 并且在没有歧义的情况下, 把内直和与外直和都简称为**直和**.

例 6.1. 设$\mathrm{char}\, F \ne 2$, $V = F^{n \times n}$,

$$V_1 = \{A \in F^{n\times n} \mid A^t = A\}, \qquad \text{(对称矩阵)}$$
$$V_2 = \{A \in F^{n\times n} \mid A^t + A = 0\}. \qquad \text{(反对称矩阵)}$$

则$V_1 \cap V_2 = \{0\}$, 从而$V_1 + V_2$是直和. 我们断言$V = V_1 \oplus V_2$. 事实上, 对任意$A \in V$有

$$A_1 = \frac{1}{2}(A + A^t) \in V_1, \quad A_2 = \frac{1}{2}(A - A^t) \in V_2,$$

并且$A = A_1 + A_2$. □

例 6.2. 设$\{\alpha_1, \ldots, \alpha_n\}$是$V$的基. 记$V_i = \mathrm{span}\{\alpha_i\}$. 则$V = \bigoplus_{i=1}^n V_i$. □

下面我们继续假设$\dim V < \infty$. 设$T \in L(V)$. V的子空间W称为是T的**不变子空间**, 如果$T(W) \subset W$. 例如, $\mathrm{Ker}(T)$和$\mathrm{Im}(T)$是不变子空间.

命题 6.5. 设$T \in L(V)$. 则T可对角化$\iff V$可以分解成一维不变子空间的直和.

证明. "$\Longrightarrow$". 设T可对角化. 则存在由T的特征向量构成的基$\{\alpha_1, \ldots, \alpha_n\}$. 于是$V_i = \operatorname{span}\{\alpha_i\}$是一维不变子空间, 并且$V = \bigoplus_{i=1}^n V_i$.

"$\Longleftarrow$". 设V可以分解成一维不变子空间的直和$V = \bigoplus_{i=1}^n V_i$. 取$\alpha_i \in V_i \setminus \{0\}$. 则$\alpha_i$是特征向量, 并且$\{\alpha_1, \ldots, \alpha_n\}$是$V$的基. 因此$T$可对角化. □

特征子空间的一个基本性质如下.

命题 6.6. 设$T \in L(V)$. 则特征子空间$\{V_c \mid c \in \sigma(T)\}$是无关的. 于是, 可以考虑直和$\bigoplus_{c\in\sigma(T)} V_c$.

证明. 设$\sigma(T) = \{c_1, \ldots, c_k\}$, $\alpha_i \in V_{c_i}$, $\sum_{i=1}^k \alpha_i = 0$. 此式两边取$T^j$ $(0 \le j \le k-1)$, 得

$$\sum_{i=1}^k c_i^j \alpha_i = 0,$$

即

$$(\alpha_1, \ldots, \alpha_k) \begin{bmatrix} 1 & c_1 & \cdots & c_1^{k-1} \\ \cdots & \cdots & \cdots & \cdots \\ 1 & c_k & \cdots & c_k^{k-1} \end{bmatrix} = (0, \ldots, 0).$$

由于$c_1, \ldots, c_k$互不相同, 所以上式中的Vandermonde矩阵可逆. 两边右乘它的逆, 即得$(\alpha_1, \ldots, \alpha_k) = (0, \ldots, 0)$, 即$\alpha_i = 0$. □

对于$f \in F[x]$, 我们把f的根c的重数记为$m(c, f)$. 下面是特征子空间的另一个基本性质.

命题 6.7. 设$T \in L(V)$, $c \in \sigma(T)$. 则

$$1 \le \dim V_c \le m(c, f_T).$$

证明. 显然$\dim V_c \ge 1$. 取V_c的有序基$\{\alpha_1, \ldots, \alpha_d\}$, 并扩充为$V$的有序基$\mathcal{B} = \{\alpha_1, \ldots, \alpha_n\}$. 由于当$1 \le j \le d$时有$T\alpha_j = c\alpha_j$, 所以$[T]_{\mathcal{B}}$形如$\begin{bmatrix} cI_d & B \\ 0 & D \end{bmatrix}$, 这里$B \in F^{d\times(n-d)}$, $D \in F^{(n-d)\times(n-d)}$. 因此

$$\begin{aligned} f_T = f_{[T]_{\mathcal{B}}} &= \det\left(xI_n - \begin{bmatrix} cI_d & B \\ 0 & D \end{bmatrix}\right) = \det \begin{bmatrix} (x-c)I_d & -B \\ 0 & xI_{n-d} - D \end{bmatrix} \\ &= \det((x-c)I_d)\det(xI_{n-d} - D) = (x-c)^d f_D. \end{aligned}$$

因此$m(c, f_T) \ge d$. □

$m(c, f_T)$和$\dim V_c$分别称为T的特征值c的**代数重数**和**几何重数**.

下面我们讨论线性映射可对角化的条件.

定理 6.8. 设$T \in L(V)$. 以下论断等价:

(1) T可对角化.

(2) $V = \bigoplus_{c\in\sigma(T)} V_c$.

(3) $\dim V = \sum_{c\in\sigma(T)} \dim V_c$.

(4) $f_T = \prod_{c\in\sigma(T)} (x-c)^{\dim V_c}$.

证明. "(1)$\iff$(2)".

$$T\text{可对角化} \iff \text{存在由}T\text{的特征向量构成的基} \iff \bigoplus_{c\in\sigma(T)} V_c = V.$$

“(2)⟹(4)”. 设$\sigma(T)=\{c_1,\dots,c_k\}$, 并记$d_i=\dim V_{c_i}$. 设$\mathcal{B}_i=\{\alpha_{i1},\dots,\alpha_{id_i}\}$是$V_{c_i}$的有序基. 则$\mathcal{B}=\{\alpha_{11},\dots,\alpha_{1d_1},\dots,\alpha_{k1},\dots,\alpha_{kd_k}\}$是$V$的有序基. 由于$T\alpha_{ij}=c_i\alpha_{ij}$, 所以$[T]_{\mathcal{B}}=\operatorname{diag}(c_1I_{d_1},\dots,c_kI_{d_k})$. 因此$f_T=f_{[T]_{\mathcal{B}}}=(x-c_1)^{d_1}\cdots(x-c_k)^{d_k}$.

“(4)⟹(3)”. $\dim V=\deg f_T=\sum_{c\in\sigma(T)}\dim V_c$.

“(3)⟹(2)”. $\dim\bigoplus_{c\in\sigma(T)}V_c=\sum_{c\in\sigma(T)}\dim V_c=\dim V$. 因此$V=\bigoplus_{c\in\sigma(T)}V_c$. □

推论 6.9. 设$T\in L(V)$. 以下论断等价:

(1) T可对角化.

(2) f_T可以分解为一次式的乘积, 并且任意特征值的代数重数等于几何重数.

(3) f_T可以分解为一次式的乘积, 并且任意代数重数大于1的特征值的代数重数等于几何重数.

特别地, 如果$|\sigma(T)|=\dim V$, 则T可对角化.

证明. “(1)⟺(2)”和“(2)⟹(3)”. 显然.

“(3)⟹(2)”. 只需注意到当$m(c,f_T)=1$时, 由命题6.7总有$\dim V_c=1$.

如果$|\sigma(T)|=\dim V$, 则f_T可以分解为一次式的乘积并且无重根, 从而T可对角化. □

推论 6.10. 设$A\in F^{n\times n}$.

(1) A可对角化⟺ f_A可以分解为一次式的乘积, 并且任意代数重数大于1的特征值c的代数重数等于其几何重数$\dim\operatorname{Ker}(cI-A)$. 特别地, 如果$|\sigma(A)|=n$, 则$A$可对角化.

(2) 当A可对角化时, 如果$f_A=(x-c_1)^{d_1}\cdots(x-c_k)^{d_k}$, $\{\alpha_{i1},\dots,\alpha_{id_i}\}$是$\operatorname{Ker}(c_iI-A)$的基, 则矩阵$P=[\alpha_{11},\dots,\alpha_{1d_1},\dots,\alpha_{k1},\dots,\alpha_{kd_k}]\in F^{n\times n}$可逆, 并且
$$P^{-1}AP=\operatorname{diag}(c_1I_{d_1},\dots,c_kI_{d_k}).$$

证明. 在推论6.9中取$T=L_A$即得(1). 当(2)的条件成立时, 由于$F^{n\times 1}=\bigoplus_{i=1}^k\operatorname{Ker}(c_iI-A)$, 所以$P$的列向量构成$F^{n\times 1}$的基, 因此$P$可逆. 由于$A\alpha_{ij}=c_i\alpha_{ij}$, 所以
$$\begin{aligned}AP&=[A\alpha_{11},\dots,A\alpha_{1d_1},\dots,A\alpha_{k1},\dots,A\alpha_{kd_k}]\\&=[c_1\alpha_{11},\dots,c_1\alpha_{1d_1},\dots,c_k\alpha_{k1},\dots,c_k\alpha_{kd_k}]\\&=P\operatorname{diag}(c_1I_{d_1},\dots,c_kI_{d_k}),\end{aligned}$$
即$P^{-1}AP=\operatorname{diag}(c_1I_{d_1},\dots,c_kI_{d_k})$. □

注意A不可对角化可能由两个原因导致:

- 如果f_A不能分解为一次式的乘积, 则A不可对角化. (可以通过扩大F来解决)
- 如果某特征值的几何重数小于代数重数, 即特征向量不足以构成空间的基, 则A不可对角化. (不能通过扩大F来解决,更本质)

如果F是代数闭域, 则第一个原因不会出现.

例 6.3. 设$A=\begin{bmatrix}\cos\theta&-\sin\theta\\\sin\theta&\cos\theta\end{bmatrix}\in\mathbb{R}^{2\times 2}$, 其中$\theta\in\mathbb{R}\setminus\pi\mathbb{Z}$. 则
$$f_A=\det(xI-A)=\det\begin{bmatrix}x-\cos\theta&\sin\theta\\-\sin\theta&x-\cos\theta\end{bmatrix}=x^2-(2\cos\theta)x+1.$$
容易看出, f_A无实根. 所以A(在$\mathbb{R}$上)不可对角化.

如果视$A\in\mathbb{C}^{2\times 2}$, 则$\sigma(A)=\{e^{i\theta},e^{-i\theta}\}$, $|\sigma(A)|=2$, 所以A在$\mathbb{C}$上可对角化. 关于特征

值$e^{i\theta}$和$e^{-i\theta}$的特征子空间

$$\operatorname{Ker}(e^{i\theta}I-A)=\operatorname{Ker}\begin{bmatrix} i\sin\theta & \sin\theta \\ -\sin\theta & i\sin\theta \end{bmatrix}$$

和

$$\operatorname{Ker}(e^{-i\theta}I-A)=\operatorname{Ker}\begin{bmatrix} -i\sin\theta & \sin\theta \\ -\sin\theta & -i\sin\theta \end{bmatrix}$$

分别有基

$$\left\{\alpha_1=\begin{bmatrix} 1 \\ -i \end{bmatrix}\right\},\quad \left\{\alpha_2=\begin{bmatrix} 1 \\ i \end{bmatrix}\right\}.$$

设

$$P=[\alpha_1,\alpha_2]=\begin{bmatrix} 1 & 1 \\ -i & i \end{bmatrix}.$$

则

$$P^{-1}AP=\operatorname{diag}(e^{i\theta},e^{-i\theta}).$$

注意P与θ无关. □

例 6.4. 设$A=\begin{bmatrix} \lambda & a & b \\ 0 & \lambda & c \\ 0 & 0 & \lambda \end{bmatrix}\in\mathbb{R}^{3\times 3}$, 这里$\lambda,a,b,c\in\mathbb{R}$并且$a,b,c$不全为0. 则

$$f_A=\det(xI-A)=\det\begin{bmatrix} x-\lambda & -a & -b \\ 0 & x-\lambda & -c \\ 0 & 0 & x-\lambda \end{bmatrix}=(x-\lambda)^3.$$

因此$\sigma(A)=\{\lambda\}$. 特征值λ的代数重数是3, 几何重数是

$$\dim\operatorname{Ker}(\lambda I-A)=\dim\operatorname{Ker}\begin{bmatrix} 0 & -a & -b \\ 0 & 0 & -c \\ 0 & 0 & 0 \end{bmatrix}=3-\operatorname{rank}\begin{bmatrix} 0 & -a & -b \\ 0 & 0 & -c \\ 0 & 0 & 0 \end{bmatrix}<3.$$

所以A不可对角化. 并且即使视$A\in\mathbb{C}^{3\times 3}$, 它也不可对角化. □

例 6.5. 设$A=\begin{bmatrix} 5 & -6 & -6 \\ -1 & 4 & 2 \\ 3 & -6 & -4 \end{bmatrix}\in\mathbb{R}^{3\times 3}$. 则

$$f_A=\det(xI-A)=\det\begin{bmatrix} x-5 & 6 & 6 \\ 1 & x-4 & -2 \\ -3 & 6 & x+4 \end{bmatrix}=(x-1)(x-2)^2.$$

因此$\sigma(A)=\{1,2\}$. 特征值2的代数重数是2, 几何重数是

$$\dim\operatorname{Ker}(2I-A)=\dim\operatorname{Ker}\begin{bmatrix} -3 & 6 & 6 \\ 1 & -2 & -2 \\ -3 & 6 & 6 \end{bmatrix}=3-\operatorname{rank}\begin{bmatrix} -3 & 6 & 6 \\ 1 & -2 & -2 \\ -3 & 6 & 6 \end{bmatrix}=2.$$

所以A可对角化. 关于特征值1和2的特征子空间

$$\operatorname{Ker}(I-A)=\operatorname{Ker}\begin{bmatrix}-4 & 6 & 6\\ 1 & -3 & -2\\ -3 & 6 & 5\end{bmatrix}$$

和

$$\operatorname{Ker}(2I-A)=\operatorname{Ker}\begin{bmatrix}-3 & 6 & 6\\ 1 & -2 & -2\\ -3 & 6 & 6\end{bmatrix}$$

分别有基

$$\left\{\alpha_1=\begin{bmatrix}3\\-1\\3\end{bmatrix}\right\},\quad \left\{\alpha_2=\begin{bmatrix}2\\1\\0\end{bmatrix},\alpha_3=\begin{bmatrix}2\\0\\1\end{bmatrix}\right\}.$$

设

$$P=[\alpha_1,\alpha_2,\alpha_3]=\begin{bmatrix}3 & 2 & 2\\ -1 & 1 & 0\\ 3 & 0 & 1\end{bmatrix}.$$

则

$$P^{-1}AP=\operatorname{diag}(1,2,2).$$

□

附录: 有理函数域$F(x)$.

考虑集合

$$F[x]\times(F[x]\setminus\{0\})=\{(f,g)\mid f,g\in F[x],g\neq 0\}$$

上的关系: $(f_1,g_1)\sim(f_2,g_2)\iff f_1g_2=f_2g_1$. 容易验证它是等价关系:

- $(f,g)\sim(f,g)$,
- $(f_1,g_1)\sim(f_2,g_2)\Longrightarrow(f_2,g_2)\sim(f_1,g_1)$,
- $(f_1,g_1)\sim(f_2,g_2),(f_2,g_2)\sim(f_3,g_3)\Longrightarrow(f_1,g_1)\sim(f_3,g_3)$.

记$F(x)$为等价类的集合, 并把(f,g)所在的等价类记为$\frac{f}{g}$. 注意到$\frac{f_1}{g_1}=\frac{f_2}{g_2}\iff f_1g_2=f_2g_1$. 在$F(x)$中定义加法和乘法运算为

$$\frac{f_1}{g_1}+\frac{f_2}{g_2}=\frac{f_1g_2+f_2g_1}{g_1g_2},\qquad \frac{f_1}{g_1}\cdot\frac{f_2}{g_2}=\frac{f_1f_2}{g_1g_2}.$$

则$F(x)$构成一个域, 称为F系数的**有理函数域**. 其中的零元素为$\frac{0}{1}$, 壹元素为$\frac{1}{1}$, $\frac{f}{g}$的负元素为$\frac{-f}{g}$, $\frac{f}{g}(f\neq 0)$的逆元素为$\frac{g}{f}$. 虽然$\frac{f}{g}$名为"有理函数", 我们并不把它看成函数. 通过把$f\in F[x]$等同于$\frac{f}{1}\in F(x)$, 我们把$F[x]$看成$F(x)$的子集. 这一等同是保持加法和乘法运算的, 因此$F[x]$是$F(x)$的子环, F是$F(x)$的子域.

§6.2 Cayley-Hamilton定理

定理 6.11(Cayley-Hamilton). 设V是域F上的有限维线性空间, $T\in L(V)$. 则$f_T(T)=0$.

推论 6.12. 设$A\in F^{n\times n}$. 则$f_A(A)=0$

定理6.11$\Longrightarrow$推论6.12的证明. $L_{f_A(A)} = f_A(L_A) = f_{L_A}(L_A) = 0 \Longrightarrow f_A(A) = 0.$ □

注 6.1.
- 容易看出, 推论6.12$\Longrightarrow$定理6.11.
- 推论的错误证明: $f_A(A) = \det(AI - A) = \det(0) = 0.$

为了证明定理6.11, 我们利用环上的模做工具.

定义 6.6. 设集合R上给定了两个运算, 称为加法和乘法, 分别对R中任意两个元素x, y给出R中的元素$x + y$和xy, 并且满足如下性质:

(1) R对加法运算构成交换群, 即:
   - 对任意$x, y \in R$有$x + y = y + x$.
   - 对任意$x, y, z \in R$有$x + (y + z) = (x + y) + z$.
   - 在R中存在唯一的元素, 称为**零元素**并记为0, 满足$0 + x = x$对任意$x \in R$成立.
   - 对任意$x \in R$, 存在唯一的$(-x) \in R$, 称为x的**负元素**, 满足$x + (-x) = 0$.

(2) 对任意$x, y, z \in R$有$x(yz) = (xy)z$, $x(y + z) = xy + xz$, $(y + z)x = yx + zx$.

则称集合R(连同它上面的加法和乘法运算)为**环**. 如果环R中的乘法运算是交换的, 即对任意$x, y \in R$有$xy = yx$, 则称R为**交换环**. 如果在环R中存在唯一的元素, 称为**壹元素**并记为1, 满足$1x = x1 = x$对任意$x \in R$成立, 则称R为**幺环**.

如果$\mathcal{A}$是域F上的代数(例如$F^{n\times n}$, $L(V)$, $F[x]$), 则$\mathcal{A}$是环.

我们主要考虑交换幺环. 注意我们允许$1 = 0$, 即$R = \{0\}$的情况(这主要是因为我们希望交换幺环模掉所有理想(包括R自身)还是交换幺环, 见**??**). 除此之外, 交换幺环与域的唯一区别在于非零元素在乘法运算下的逆元素不一定存在. 也就是说:

- 域是交换幺环.
- 在交换幺环R中, 如果$1 \neq 0$, 并且对任意$x \in R \setminus \{0\}$, 存在唯一的$x^{-1} \in R$满足$xx^{-1} = 1$, 则R是域.

例如, 整数环$\mathbb{Z}$是交换幺环但不是域. 域F上的多项式环$F[x]$是交换幺环但不是域.

设R是交换幺环. 记$R^{m\times n}$为矩阵元在R中的$m \times n$矩阵的集合. 对$A, B \in R^{m\times n}$, 定义$A + B \in R^{m\times n}$为

$$(A + B)_{ij} = A_{ij} + B_{ij}.$$

对$A \in R^{m\times n}, B \in R^{n\times p}$, 定义$AB \in R^{m\times p}$为

$$(AB)_{ij} = \sum_{k=1}^{n} A_{ik}B_{kj}.$$

则

$$A(B + C) = AB + AC, \quad A(BC) = (AB)C, \quad \ldots$$

定义 6.7. 给定交换幺环R. 设对于集合V, 给定了如下两个运算:

- 加法运算$V \times V \to V$, $(\alpha_1, \alpha_2) \mapsto \alpha_1 + \alpha_2$, 使$V$成为交换群.
- 乘法运算$R \times V \to V$, $(c, \alpha) \mapsto c\alpha$, 满足

$$(c_1 + c_2)\alpha = c_1\alpha + c_2\alpha, \quad c(\alpha_1 + \alpha_2) = c\alpha_1 + c\alpha_2, \quad (c_1c_2)\alpha = c_1(c_2\alpha), \quad 1\alpha = \alpha$$

对任意$c, c_1, c_2 \in R$和$\alpha, \alpha_1, \alpha_2 \in V$成立.

则称集合V(连同这两个运算)为**R上的模**(或**R-模**).

例如R^n, $R^{n\times 1}$, $R^{m\times n}$在自然定义的加法和乘法运算下是R-模. 注意如果F是域, 则F-模$\Longleftrightarrow$ F-线性空间.

例 6.6. 设G是交换群, 群运算为加法. 则G自然成为$\mathbb{Z}$-模. 这里$nx = x + \cdots + x$(n个). □

下面的例子对我们是最重要的.

例 6.7. 设V是域F上的线性空间, $T \in L(V)$. 定义乘法$F[x] \times V \to V$为$(f, \alpha) \mapsto f\alpha = f(T)\alpha$. 则$V$成为$F[x]$-模. 以后我们在考虑线性空间到自身的线性映射时, 同时考虑这个模结构是有好处的. □

注 6.2. 事实上V是$L(V)$-模. V上的$F[x]$-模结构由环同态$F[x] \to L(V)$, $f \mapsto f(T)$诱导.

设V为交换幺环R上的模. 我们定义形式矩阵乘法$V^n \times R^{n\times p} \to V^p$为

$$(\alpha_1, \ldots, \alpha_n)A = (\beta_1, \ldots, \beta_p),$$

其中

$$\beta_j = \sum_{i=1}^{n} A_{ij}\alpha_i.$$

容易验证下面的乘法结合律:

$$[(\alpha_1, \ldots, \alpha_n)A]B = (\beta_1, \ldots, \beta_p)(AB), \qquad \forall(\alpha_1, \ldots, \alpha_n) \in V^n, A \in R^{n\times p}, B \in R^{p\times q}.$$

定理6.11的证明. 取V的有序基$\mathcal{B} = \{\alpha_1, \ldots, \alpha_n\}$, 并设$A = [T]_{\mathcal{B}}$. 视$V$为$F[x]$-模, 即对$f \in F[x]$, $\alpha \in V$定义$f\alpha = f(T)\alpha$. 则

$$(\alpha_1, \ldots, \alpha_n)A = (T\alpha_1, \ldots, T\alpha_n) = (x\alpha_1, \ldots, x\alpha_n) = (\alpha_1, \ldots, \alpha_n)xI_n,$$

即

$$(\alpha_1, \ldots, \alpha_n)(xI_n - A) = (0, \ldots, 0).$$

另一方面, 存在矩阵$B \in F[x]^{n\times n}$使得

$$(xI_n - A)B = \det(xI_n - A)I_n = f_T I_n.$$

事实上, 取$B = \mathrm{adj}(xI_n - A)$(伴随矩阵, 即余因子矩阵的转置)即可. 于是,

$$\begin{aligned} &[(\alpha_1, \ldots, \alpha_n)(xI_n - A)]B = (0, \ldots, 0) \\ &\qquad\qquad\qquad \| \\ &(\alpha_1, \ldots, \alpha_n)[(xI_n - A)B] = (\alpha_1, \ldots, \alpha_n)f_T I_n = (f_T\alpha_1, \ldots, f_T\alpha_n) = (f_T(T)\alpha_1, \ldots, f_T(T)\alpha_n). \end{aligned}$$

因此$f_T(T)\alpha_1 = \cdots = f_T(T)\alpha_n = 0$. 这推出$f_T(T) = 0$. □

§6.3 最小多项式

设V是有限维F-线性空间, $T \in L(V)$. 如果$f \in F[x]$满足$f(T) = 0$, 则称f是T的**零化多项式**(annihilating polynomial). 由Cayley-Hamilton定理, f_T是T的零化多项式. 我们把T的所有零化多项式的集合记为M_T.

引理 6.13. M_T是$F[x]$理想.

证明. 显然M_T非空.
M_T是子空间: $f, g \in M_T, c \in F \Longrightarrow f + cg \in M_T$.
M_T是理想: $f \in M_T, g \in F[x] \Longrightarrow fg \in M_T$. □

注 6.3. 由于$f_T \in M_T$, 所以M_T是非零理想. 该事实也可以不用Cayley-Hamilton定理证明: 设$\dim V = n$, 则$\dim L(V) = n^2$, 从而$I, T, \ldots, T^{n^2}$线性相关, 即存在不全为0的$c_0, \ldots, c_{n^2}$使得

$$c_0 I + c_1 T + \cdots + c_{n^2} T^2 = 0.$$

令$f = \sum_{i=0}^{n^2} c_i x^i \neq 0$. 则$f \in M$. □

回忆$F[x]$的非零理想都是主理想, 并且存在唯一的首项系数是1的生成元p_T, 即

$$f(T) = 0 \iff p_T \mid f.$$

p_T称为T的**最小多项式**. 当$V = \{0\}$时, 我们约定$f_T = p_T = 1$. 注意当$V \neq \{0\}$时, 总有$\deg p_T \geq 1$并且$p_T | f_T$. 由理想生成元的性质, $p \in F[x]$是T的最小多项式$\iff$以下两条成立:

- p是T的零化多项式并且首项系数是1;
- 任何T的非零零化多项式的次数$\geq \deg p$.

类似地, 如果$A \in F^{n\times n}$, 则

$$M_A := \{f \in F[x] \mid f(A) = 0\}$$

是$F[x]$的非零理想. 它的首项系数是1的生成元p_A称为A的**最小多项式**. 性质:

- $p_A | f_A$.
- $p_A = p_{L_A}$, $p_T = p_{[T]_{\mathcal{B}}}$.
- A与B相似$\Longrightarrow p_A = p_B$: $f(P^{-1}AP) = P^{-1}f(A)P$.
- 设有两个域$F_1 \subset F_2$, $A \in F_1^{n\times n}$, 则A作为$F_1^{n\times n}$和$A \in F_2^{n\times n}$中的矩阵有相同的最小多项式: 设两种看法的最小多项式分别为$p_1 \in F_1[x]$和$p_2 \in F_2[x]$. 显然有$\deg p_2 \leq \deg p_1$. 设$\deg p_2 = k$, $p_2 = x^k + a_{k-1}x^{k-1} + \cdots + a_1 x + a_0$. 则
$$A^k + a_{k-1}A^{k-1} + \cdots + a_1 A + a_0 I = 0.$$
把这一等式视为关于未知量$a_{k-1}, \ldots, a_0$的线性方程组, 它有n^2个方程, 系数和常数项为F_1中的元素, 方程组在F_2中有解, 因此在F_1中也有解, 即存在$b_{k-1}, \ldots, b_0 \in F_1$满足$A^k + b_{k-1}A^{k-1} + \cdots + b_1 A + b_0 I = 0$. 于是$q := x^k + b_{k-1}x^{k-1} + \cdots + b_1 x + b_0 \in F_1[x]$满足$q(A) = 0$. 因此$\deg p_2 = \deg q \geq \deg p_1$. 这推出$\deg p_1 = \deg p_2$. 若$p_1 \neq p_2$, 则$p_1 - p_2 \in F_2[x] \setminus \{0\}$满足$(p_1 - p_2)(A) = 0$但$\deg(p_1 - p_2) < \deg p_2$, 矛盾.

命题 6.14. p_T与f_T有相同的根.

证明. 由于$p_T | f_T$, 所以p的根总是f的根. 反过来, 设$f(c) = 0$. 则c是T的特征值, 既存在$\alpha \in V \setminus \{0\}$满足$T\alpha = c\alpha$. 设$p = x^k + a_{k-1}x^{k-1} + \cdots + a_1 x + a_0$. 则

$$\begin{aligned}
p(T)\alpha &= (T^k + a_{k-1}T^{k-1} + \cdots + a_1 T + a_0 I)\alpha \\
&= T^k\alpha + a_{k-1}T^{k-1}\alpha + \cdots + a_1 T\alpha + a_0\alpha \\
&= c^k\alpha + a_{k-1}c^{k-1}\alpha + \cdots + a_1 c\alpha + a_0\alpha \\
&= (c^k + a_{k-1}c^{k-1} + \cdots + a_1 c + a_0)\alpha \\
&= p(c)\alpha.
\end{aligned}$$

由于$p(T) = 0$, 所以$p(c)\alpha = 0$. 但$\alpha \neq 0$. 所以$p(c) = 0$. □

在命题中取$T = L_A$, 即得p_A与f_A有相同的根.

由Cayley-Hamilton定理和命题6.14, 如果$\sigma(T) = \{c_1, \ldots, c_k\}$, $f_T = (x - c_1)^{d_1} \cdots (x - c_k)^{d_k}$,

则$p_T=(x-c_1)^{r_1}\cdots(x-c_k)^{r_k}$, 其中$1\le r_i\le d_i$. 实际上, 任何在这个范围中的$r_i$都可以出现.

定理 6.15. 对于$T\in L(V)$, TFAE:

(1) T可对角化.

(2) 设$\sigma(T)=\{c_1,\dots,c_k\}$, 则$p_T=\prod_{i=1}^k(x-c_i)$.

(3) p_T可以分解为互不相同(应该是相伴!)的一次式的乘积.

为了证明定理6.15, 我们先证明:

引理 6.16. 设$\dim V<\infty$, $T_1,\dots,T_k\in L(V)$. 则

$$\dim\operatorname{Ker}(T_1\cdots T_k)\le\sum_{i=1}^k\dim\operatorname{Ker}(T_i).$$

证明. 当$k=1$时无需证明. 设$k=2$. 考虑$T_2|_{\operatorname{Ker}(T_1T_2)}:\operatorname{Ker}(T_1T_2)\to V$. 注意到

$$\operatorname{Ker}(T_2|_{\operatorname{Ker}(T_1T_2)})=\operatorname{Ker}(T_1T_2)\cap\operatorname{Ker}(T_2)=\operatorname{Ker}(T_2)$$

并且$\operatorname{Im}(T_2|_{\operatorname{Ker}(T_1T_2)})\subset\operatorname{Ker}(T_1)$. 因此,

$$\begin{aligned}\dim\operatorname{Ker}(T_1T_2)&=\dim\operatorname{Ker}(T_2|_{\operatorname{Ker}(T_1T_2)})+\dim\operatorname{Im}(T_2|_{\operatorname{Ker}(T_1T_2)})\\&\le\dim\operatorname{Ker}(T_2)+\dim\operatorname{Ker}(T_1).\end{aligned}$$

假设$k>2$, 并且引理当$<k$时成立. 则

$$\begin{aligned}\dim\operatorname{Ker}(T_1\cdots T_k)&\le\dim\operatorname{Ker}(T_1\cdots T_{k-1})+\dim\operatorname{Ker}(T_k)\\&\le\sum_{i=1}^{k-1}\dim\operatorname{Ker}(T_i)+\dim\operatorname{Ker}(T_k)=\sum_{i=1}^{k}\dim\operatorname{Ker}(T_i).\end{aligned}$$

□

定理6.15的证明. "(1)$\Longrightarrow$(2)". 记$p=\prod_{i=1}^k(x-c_i)$. 由命题6.14有$p|p_T$. 取$\mathcal{B}$使$[T]_{\mathcal{B}}=\operatorname{diag}(c_1I_{d_1},\dots,c_kI_{d_k})$. 容易验证$p([T]_{\mathcal{B}})=0$. 从而$p(T)=0$. 因此$p_T|p$. 这推出$p_T=p$.

"(2)$\Longleftarrow$(1)". 设$p_T=\prod_{i=1}^k(x-c_i)$. 则$\prod_{i=1}^k(T-c_iI)=0$. 由引理, 有

$$\dim V=\dim\operatorname{Ker}\left(\prod_{i=1}^k(T-c_iI)\right)\le\sum_{i=1}^k\dim\operatorname{Ker}(T-c_iI)=\sum_{i=1}^k\dim V_{c_i}=\dim\bigoplus_{i=1}^kV_{c_i}.$$

所以只能有$V=\bigoplus_{i=1}^kV_{c_i}$. 因此$T$可对角化.

"(2)$\Longleftarrow$(3)". 显然.

"(3)$\Longleftarrow$(2)". 由命题6.14容易看出. □

在定理中取$T=L_A$, 即得

推论 6.17. 对于$A\in F^{n\times n}$, TFAE:

(1) A可对角化.

(2) 设$\sigma(A)=\{c_1,\dots,c_k\}$, 则$p_A=\prod_{i=1}^k(x-c_i)$.

(3) p_A可以分解为互不相同的一次式的乘积. □

推论6.17可以用来判断A是否可对角化. 假设f_A可以分解为一次式的乘积

$$f_A=\prod_{i=1}^k(x-c_i)^{d_i}$$

(否则不能对角化). 我们知道,

$$A\text{可对角化} \iff \text{当}d_i > 1\text{时总有}\dim\operatorname{Ker}(A - c_iI) = d_i.$$

推论6.17给出了另一种判断方法:

$$A\text{可对角化} \iff \prod_{i=1}^{k}(A - c_iI) = 0.$$

反过来, 推论6.17也对求最小多项式有帮助.

例 6.8. 设$A = \begin{bmatrix}\cos\theta & -\sin\theta \\ \sin\theta & \cos\theta\end{bmatrix} \in \mathbb{R}^{2\times 2}$, 其中$\theta \in \mathbb{R} \setminus \pi\mathbb{Z}$. 我们知道,

$$f_A = x^2 - (2\cos\theta)x + 1.$$

由于f_A无实根, 所以在$\mathbb{R}[x]$中不可约. 由Cayley-Hamilton定理, $p_A | f_A$. 由于$\deg p_A \geq 1$, 所以$p_A = f_A$.

也可以这样求p_A(不用Cayley-Hamilton定理): 视$A \in \mathbb{C}^{2\times 2}$, 则$\sigma(A) = \{e^{i\theta}, e^{-i\theta}\}$, $|\sigma(A)| = 2$, 所以A在$\mathbb{C}$上可对角化. 由推论6.17,

$$p_A = (x - e^{i\theta})(x - e^{-i\theta}) = f_A.$$

这里用到了最小多项式与域无关这一事实. □

例 6.9. 设$A = \begin{bmatrix}\lambda & 0 & 1 \\ 0 & \lambda & 0 \\ 0 & 0 & \lambda\end{bmatrix} \in \mathbb{R}^{3\times 3}$, 这里$\lambda \in \mathbb{R}$. 我们知道, $f_A = (x - \lambda)^3$, 由命题6.14, $p_A = (x - \lambda)^d$, $d \geq 1$. 容易验证, $A - \lambda I \neq 0$, $(A - \lambda I)^2 = 0$. 所以$p_A = (x - \lambda)^2$. 这也推出A不可对角化. 注意这里没有用到Cayley-Hamilton定理. □

例 6.10. 设$A = \begin{bmatrix}3 & 1 & -1 \\ 2 & 2 & -1 \\ 2 & 2 & 0\end{bmatrix} \in \mathbb{R}^{3\times 3}$. 则

$$f_A = \det(xI - A) = \det\begin{bmatrix}x-3 & -1 & 1 \\ -2 & x-2 & 1 \\ -2 & -2 & x\end{bmatrix} = (x-1)(x-2)^2.$$

由Cayley-Hamilton定理和命题6.14, $p_A = (x-1)(x-2)^d$, $d = 1$或2. 容易验证,

$$(A - I)(A - 2I) = \begin{bmatrix}2 & 1 & -1 \\ 2 & 1 & -1 \\ 2 & 2 & -1\end{bmatrix}\begin{bmatrix}1 & 1 & -1 \\ 2 & 0 & -1 \\ 2 & 2 & -2\end{bmatrix} = \begin{bmatrix}2 & * & * \\ * & * & * \\ * & * & *\end{bmatrix} \neq 0.$$

所以$d = 2$, $p_A = f_A$. 这也推出A不可对角化. □

例 6.11. 设$A=\begin{bmatrix}0&1&0&1\\1&0&1&0\\0&1&0&1\\1&0&1&0\end{bmatrix}\in\mathbb{Q}^{4\times4}$. 容易得到:

$$A^2=\begin{bmatrix}2&0&2&0\\0&2&0&2\\2&0&2&0\\0&2&0&2\end{bmatrix},\quad A^3=\begin{bmatrix}0&4&0&4\\4&0&4&0\\0&4&0&4\\4&0&4&0\end{bmatrix}.$$

从而$A^3=4A$. 令

$$p=x^3-4x=x(x+2)(x-2).$$

则$p(A)=0$. 这说明$p_T|p$. 容易看出, $\deg p_T\neq 1$. 我们断言, $\deg p_T\neq 2$. 事实上, 如果$\deg p_T=2$, 则$p_T=x(x+2),x(x-2)$或$(x+2)(x-2)$. 但是$A^2\neq\pm 2A,4I$. 因此$\deg p_T\geq 3$. 只能有$p_T=p$.

我们进一步求f_A. 注意到$\deg f_A=4$. 由命题6.14, 只能有

$$f_A=x^2(x+2)(x-2),\quad x(x+2)^2(x-2)\quad \text{或}\quad x(x+2)(x-2)^2.$$

注意到特征值0的几何重数是

$$\dim\operatorname{Ker}(A)=4-\operatorname{rank}(A)=2.$$

因此特征值0的代数重数≥ 2. 从而只能有$f_A=x^2(x+2)(x-2)$. 另外, 由于特征值0的代数重数等于几何重数, 所以A可对角化, 并且相似于$\operatorname{diag}(0,0,2,-2)$. □

§6.4 不变子空间

回忆: 给定$T\in L(V)$, V的子空间W称为是T的不变子空间, 如果$T(W)\subset W$. 例如, $\{0\}$, V, $\operatorname{Ker}(T)$和$\operatorname{Im}(T)$是不变子空间. 性质:

- 设$W\subset V$是一维子空间. 则W不变$\Longleftrightarrow$ W中的非零向量都是特征向量. 例如, 设
$$A=\begin{bmatrix}\cos\theta&-\sin\theta\\\sin\theta&\cos\theta\end{bmatrix}\in\mathbb{R}^{2\times2},$$
其中$\theta\in\mathbb{R}\setminus\pi\mathbb{Z}$. 则$L_A$只有平凡不变子空间. 事实上, 如果$W\subset V$是非平凡不变子空间, 则$\dim W=1$, 从而$W$中的非零向量都是$A$的特征向量. 但$A$在$\mathbb{R}$中无特征值, 从而无特征向量.
- 若$W\subset V$是T的不变子空间, 则对任意$f\in F[x]$, W是$f(T)$的不变子空间: 设$f=\sum_{i=0}^k a_ix^i$. 则对$\alpha\in W$有
$$f(T)\alpha=\sum_{i=0}^k a_iT^i\alpha\in W.$$
- 设$T,U\in L(V)$, $TU=UT$. 则$\operatorname{Ker}(U)$和$\operatorname{Im}(U)$是T的不变子空间:
$$\alpha\in\operatorname{Ker}(U)\quad\Longrightarrow\quad UT\alpha=TU\alpha=0\quad\Longrightarrow\quad T\alpha\in\operatorname{Ker}(U),$$
$$\alpha\in\operatorname{Im}(U)\quad\Longrightarrow\quad\exists\beta\in V,\alpha=U\beta\quad\Longrightarrow\quad T\alpha=TU\beta=UT\beta\in\operatorname{Im}(U).$$
注意到U与T可交换$\Longrightarrow$ U与$f(T)$可交换$\Longrightarrow$ $\operatorname{Ker}(f(T))$是U的不变子空间. 特别地, T的特征子空间$V_c=\operatorname{Ker}(cI-T)$是$U$(特别地, T)的不变子空间.

设$W \subset V$是T的不变子空间, 则T诱导了$T_W \in L(W,W)$和$T_{V/W} \in L(V/W, V/W)$如下:

- 对$\alpha \in W$, 定义$T_W\alpha = T\alpha$.
- 对$\alpha + W \in V/W$, 定义$T_{V/W}(\alpha + W) = T\alpha + W$. 良定.

设$\dim V < \infty$, $\mathcal{B}' = \{\alpha_1, \ldots, \alpha_r\}$是$W$的有序基, 并扩充为$V$的有序基$\mathcal{B} = \{\alpha_1, \ldots, \alpha_n\}$. 则$\mathcal{B}'' = \{\alpha_{r+1} + W, \ldots, \alpha_n + W\}$是$V/W$的有序基. 设$A = [T]_{\mathcal{B}}$. 则

$$T\alpha_j = \sum_{i=1}^{n} A_{ij}\alpha_i.$$

当$j \le r$时有$T\alpha_j = \sum_{i=1}^n A_{ij}\alpha_i \in W$. 所以当$i > r$时有$A_{ij} = 0$. 因此$A$形如

$$A = \begin{bmatrix} B & C \\ 0 & D \end{bmatrix},$$

这里$B \in F^{r\times r}$, $C \in F^{r\times(n-r)}$, $D \in F^{(n-r)\times(n-r)}$. 我们有:

- $[T_W]_{\mathcal{B}'} = B$. 这是因为
$$T_W\alpha_j = T\alpha_j = \sum_{i=1}^{r} A_{ij}\alpha_i = \sum_{i=1}^{r} B_{ij}\alpha_i, \qquad j \in \{1, \ldots, r\}.$$
- $[T_{V/W}]_{\mathcal{B}''} = D$. 这是因为
$$\begin{aligned} T_{V/W}(\alpha_{r+j} + W) = T\alpha_{r+j} + W &= \sum_{i=1}^{r} A_{i,r+j}\alpha_i + \sum_{i=r+1}^{n} A_{i,r+j}\alpha_i + W = \sum_{i=r+1}^{n} A_{i,r+j}\alpha_i + W \\ &= \sum_{i=r+1}^{n} A_{i,r+j}(\alpha_i + W) = \sum_{i=1}^{n-r} D_{ij}(\alpha_{r+i} + W), \qquad j \in \{1, \ldots, n-r\}. \end{aligned}$$

这推出:

命题 6.18. $\det(T) = \det(T_W)\det(T_{V/W})$.

证明. $\det(T) = \det(A) = \det(B)\det(D) = \det(T_W)\det(T_{V/W})$. □

类似地,

命题 6.19. $f_T = f_{T_W} f_{T_{V/W}}$.

证明.

$$f_T = f_A = \det(xI - A) = \det\begin{bmatrix} xI - B & -C \\ 0 & xI - D \end{bmatrix} = \det(xI - B)\det(xI - D) = f_B f_D = f_{T_W} f_{T_{V/W}}.$$

□

对于$f_1, \ldots, f_k \in F[x]$, 定义它们的**最小公倍式**(least common multiple)为理想$\bigcap_{i=1}^k (f_i)$的唯一的首项系数是1的生成元, 记为$\mathrm{lcm}(f_1, \ldots, f_k)$.

命题 6.20. $\mathrm{lcm}(p_{T_W}, p_{T_{V/W}}) | p_T | p_{T_W} p_{T_{V/W}}$. 特别地, 如果$p_{T_W}$与$p_{T_{V/W}}$互素, 则$p_T = p_{T_W} p_{T_{V/W}}$.

证明. 容易看出, $p_T(T_W) = 0$, $p_T(T_{V/W}) = 0$. 因此$p_{T_W} | p_T$, $p_{T_{V/W}} | p_T$. 这说明$\mathrm{lcm}(p_{T_W}, p_{T_{V/W}}) | p_T$. 另一方面, 对任意$\alpha \in V$有$p_{T_{V/W}}\alpha \in W$, 从而$p_{T_W} p_{T_{V/W}}\alpha = 0$. 这说明$p_T | p_{T_W} p_{T_{V/W}}$. □

注 6.4. 一般$p_T = p_{T_W} p_{T_{V/W}}$不成立.

推论 6.21. 设T可对角化, 则T_W和$T_{V/W}$可对角化.

证明. 由于T可对角化, 由定理6.15, p_T为互不相同的一次式的乘积. 上一命题推出p_{T_W}和$p_{T_{V/W}}$都是互不相同的一次式的乘积. 再由定理6.15, 即得T_W和$T_{V/W}$可对角化. □

我们利用不变子空间讨论下面的性质.

定义 6.8. • 设V是有限维F-线性空间, $\mathcal{F}$是$L(V)$的子集. 如果存在V的有序基$\mathcal{B}$, 使得所有$T \in \mathcal{F}$的矩阵$[T]_{\mathcal{B}}$都是对角矩阵, 则称$\mathcal{F}$中的线性变换**可同时对角化**.

- 设$\mathcal{M}$是$F^{n\times n}$的子集. 如果存在可逆矩阵$P \in \mathrm{GL}_n(F)$, 使得对所有$A \in \mathcal{M}$, 总有$P^{-1}AP$是对角矩阵, 则称$\mathcal{M}$中的矩阵**可同时对角化**.

定理 6.22. 对于$\mathcal{F} \subset L(V)$, TFAE:

(1) $\mathcal{F}$中的线性变换可同时对角化.

(2) $\mathcal{F}$中的线性变换都是可对角化的并且两两可交换.

证明. "(1)$\Longrightarrow$(2)". 只需证明$\mathcal{F}$中的线性变换两两可交换. 取V的有序基$\mathcal{B}$使所有$T \in \mathcal{F}$的矩阵$[T]_{\mathcal{B}}$是对角矩阵. 则对任意$T, U \in \mathcal{F}$有$[TU]_{\mathcal{B}} = [T]_{\mathcal{B}}[U]_{\mathcal{B}} = [U]_{\mathcal{B}}[T]_{\mathcal{B}} = [UT]_{\mathcal{B}}$. 因此$TU = UT$.

"(2)$\Longrightarrow$(1)". 对$\dim V$做归纳. 当$\dim V = 1$时无需证明. 设$\dim V = n \geq 2$, 并且"(2)$\Longrightarrow$(1)"当$\dim V < n$时成立. 如果$\mathcal{F}$中的线性变换都是恒同映射的常数倍, 则(1)自动成立. 因此假设存在$T \in \mathcal{F}$不是恒同映射的常数倍. 则对任意$c \in \sigma(T)$有$\dim V_c < \dim V$. 由于$\mathcal{F}$中的线性变换都与T可交换, 所以V_c是$\mathcal{F}$中所有线性变换的不变子空间. 考虑$L(V_c, V_c)$的子集

$$\mathcal{F}_c = \{U_{V_c} \mid U \in \mathcal{F}\}.$$

其中的线性变换都是可对角化的并且两两可交换. 由归纳假设, $\mathcal{F}_c$中的线性变换可同时对角化, 即存在V_c的基$\mathcal{B}_c$, 其中的每个向量是$\mathcal{F}_c$中所有线性变换的特征向量, 从而是$\mathcal{F}$中所有线性变换的特征向量. 由于T可对角化, 所以$V = \bigoplus_{c\in\sigma(T)} V_c$. 因此$\mathcal{B} = \bigcup_{c\in\sigma(T)} \mathcal{B}_c$是$V$的基. 由于$\mathcal{B}$中每个向量是$\mathcal{F}$中所有线性变换的特征向量, 所以$\mathcal{F}$中的线性变换可同时对角化. □

推论 6.23. 对于$\mathcal{M} \subset F^{n\times n}$, TFAE:

(1) $\mathcal{M}$中的矩阵可同时对角化.

(2) $\mathcal{M}$中的矩阵都是可对角化的并且两两可交换.

我们还考虑:

定义 6.9. (1) 如果$T \in L(V)$在V的某个有序基下的矩阵是上三角矩阵, 则称T是**可三角化**的.

(2) 如果$A \in F^{n\times n}$相似于上三角矩阵, 则称A是**可三角化**的.

注 6.5. 记

$$J_n = \begin{bmatrix} 0 & 0 & \cdots & 0 & 1 \\ 0 & 0 & \cdots & 1 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 1 & \cdots & 0 & 0 \\ 1 & 0 & \cdots & 0 & 0 \end{bmatrix} \in F^{n\times n}.$$

容易验证$J_n^{-1}=J_n$, 并且对任意$A\in F^{n\times n}$, J_nAJ_n是A的“中心对称”矩阵. 特别地,

$$J_n\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1,n-1} & a_{1n} \\ 0 & a_{22} & \cdots & a_{2,n-1} & a_{2n} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & a_{n-1,n-1} & a_{n-1,n} \\ 0 & 0 & \cdots & 0 & a_{nn} \end{bmatrix}J_n=\begin{bmatrix} a_{nn} & 0 & \cdots & 0 & 0 \\ a_{n-1,n} & a_{n-1,n-1} & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{2n} & a_{2,n-1} & \cdots & a_{22} & 0 \\ a_{1n} & a_{1,n-1} & \cdots & a_{12} & a_{11} \end{bmatrix}.$$

因此在可三角化的定义中, 利用上三角矩阵和下三角矩阵是等价的. □

定理 6.24. 设$\dim V=n$, $T\in L(V)$. TFAE:

(1) T可三角化.

(2) f_T可以分解为一次式的乘积.

(3) p_T可以分解为一次式的乘积.

(4) 存在T的不变子空间序列$\{0\}=W_0\subsetneq\cdots\subsetneq W_n=V$满足$\dim W_i=i$.

推论 6.25. 如果F是代数闭域, 则任意$A\in F^{n\times n}$可三角化.

定理6.24的证明. “(1)$\Longrightarrow$(2)”. 设$[T]_{\mathcal{B}}=A$是上三角矩阵. 则$f_T=\prod_{i=1}^n(x-A_{ii})$.

“(2)$\Longrightarrow$(3)”. 只需注意到$p_T|f_T$.

“(3)$\Longrightarrow$(4)”. 用归纳法构造W_k. 当$k=0$时, 只需取$W_0=\{0\}$. 假设$1\le k\le n$, 并且已构造出T的不变子空间序列$\{0\}=W_0\subsetneq\cdots\subsetneq W_{k-1}\subset V$ 满足$\dim W_i=i$. 考虑$T_{V/W_{k-1}}$. 由于$p_{T_{V/W_{k-1}}}|p_T$, $p_{T_{V/W_{k-1}}}$可以分解为一次式的乘积. 特别地, $\sigma(T_{V/W_{k-1}})\ne\varnothing$. 取$c\in\sigma(T_{V/W_{k-1}})$和关于特征值$c$的特征向量$\alpha+W_{k-1}\in V/W_{k-1}$, 这里$\alpha\in V\setminus W_{k-1}$. 则

$$T\alpha+W_{k-1}=T_{V/W_{k-1}}(\alpha+W_{k-1})=c(\alpha+W_{k-1})=c\alpha+W_{k-1},$$

即$T\alpha-c\alpha\in W_{k-1}$. 设$W_k=W_{k-1}\oplus F\alpha$. 则$T\alpha\in W_k$, 从而$W_k$是不变子空间. 这就完成了构造.

“(4)$\Longrightarrow$(1)”. 取V的有序基$\mathcal{B}=\{\alpha_1,\ldots,\alpha_n\}$使得$\{\alpha_1,\ldots,\alpha_i\}$是$W_i$的基, $1\le i\le n$. 设$A=[T]_{\mathcal{B}}$. 则

$$T\alpha_j=\sum_{i=1}^n A_{ij}\alpha_i\in W_j=\operatorname{span}\{\alpha_1,\ldots,\alpha_j\}.$$

因此当$i>j$时有$A_{ij}=0$, 即A上三角. □

注 6.6.
- V的一个子空间序列$\{0\}=W_0\subsetneq\cdots\subsetneq W_k=V$称为$V$的一个**旗**(flag). 如果进一步还有$\dim W_i=i$(即$k=\dim V$), 则称该子空间序列为$V$的一个**全旗**(full flag)或**完备旗**(complete flag). 因此(4)就是说, 存在由T的不变子空间构成的全旗(即T-不变的全旗).
- 如果不用分裂域或代数闭包等工具, “(3)$\Longrightarrow$(2)”并不显然.
- 推论6.25是较弱的结论. 一方面, $P^{-1}AP$可以更好(Jordan标准型). 另一方面, 如果只希望$P^{-1}AP$上三角, 则P可以取得更特殊(例如复矩阵的Schur三角化定理).

下面的结论是可解李代数的Lie定理的特殊情况. 为了简单起见, 我们只考虑代数闭域的情况.

定理 6.26. 假设F是代数闭域, 并且集合$\mathcal{F}\subset L(V)$中的线性变换两两可交换. 则$\mathcal{F}$中的线性变换可同时三角化, 即存在V的有序基$\mathcal{B}$, 使得所有$T\in\mathcal{F}$的矩阵$[T]_{\mathcal{B}}$都是上三角矩阵.

推论 6.27. 假设F是代数闭域, 并且$\mathcal{M}\subset F^{n\times n}$中的矩阵两两可交换. 则$\mathcal{M}$中的矩阵可同时三角

化, 即存在可逆矩阵$P \in \mathrm{GL}_n(F)$, 使得对所有$A \in \mathcal{M}$, 总有$P^{-1}AP$是上三角矩阵.

先证明:

引理 6.28. 假设F是代数闭域, 并且$\mathcal{F} \subset L(V)$中的线性变换两两可交换. 则$\mathcal{F}$中的线性变换存在公共特征向量.

证明. 首先, 不妨假设$\mathcal{F}$是有限集. 事实上, 对于一般的$\mathcal{F}$, 取$L(V)$的子空间$\mathrm{span}\mathcal{F}$的基$\mathcal{F}_0$, 则$\mathcal{F}_0$是有限集, 并且$\mathcal{F}$中的线性变换两两可交换$\Longleftrightarrow$ $\mathcal{F}_0$中的线性变换两两可交换. 假设引理对$\mathcal{F}_0$成立, 即存在$\alpha \in V \setminus \{0\}$使得对任意$T \in \mathcal{F}_0$有$T\alpha \in F\alpha$. 则对任意$U \in \mathcal{F}$, 存在$T_1, \ldots, T_k \in \mathcal{F}_0$使得$U = \sum_{i=1}^k c_iT_i$. 于是$U\alpha = \sum_{i=1}^k c_iT_i\alpha \in F\alpha$. 因此$\alpha$是$\mathcal{F}$中的线性变换的公共特征向量.

假设$\mathcal{F}$是有限集. 对$|\mathcal{F}|$做归纳. 当$|\mathcal{F}| = 0$时无需证明. 假设$|\mathcal{F}| = k \geq 1$, 并且引理当$|\mathcal{F}| = k-1$时成立. 设$\mathcal{F} = \{T_1, \ldots, T_k\}$. 由归纳假设, $T_1, \ldots, T_{k-1}$存在公共特征向量α, 即存在$c_1, \ldots, c_{k-1}$使得$(T_i - c_iI)\alpha = 0$, $1 \leq i \leq k-1$. 于是$W := \bigcap_{i=1}^{k-1} \mathrm{Ker}(T_i - c_iI) \neq \{0\}$. 由于$T_k$与$T_1, \ldots, T_{k-1}$可交换, 所以$W$是$T_k$的不变子空间. 由于$F$是代数闭域, $(T_k)_W$有特征值, 从而有特征向量$\beta \in W \setminus \{0\}$. β即为$T_1, \ldots, T_k$的公共特征向量. □

定理6.26的证明. 先证明存在子空间序列$\{0\} = W_0 \subsetneq \cdots \subsetneq W_n = V$满足$\dim W_i = i$, 并且每个$W_i$是所有$T \in \mathcal{F}$的不变子空间. 用归纳法构造$W_k$. 当$k = 0$时, 只需取$W_0 = \{0\}$. 假设$1 \leq k \leq n$, 并且已构造出子空间序列$\{0\} = W_0 \subsetneq \cdots \subsetneq W_{k-1} \subset V$, 满足当$0 \leq i \leq k-1$时, $\dim W_i = i$, 并且W_i是所有$T \in \mathcal{F}$的不变子空间. 考虑$L(V/W_{k-1}, V/W_{k-1})$的子集

$$\mathcal{F}_{k-1} := \{T_{V/W_{k-1}} \mid T \in \mathcal{F}\}.$$

它是可交换的. 对$\mathcal{F}_{k-1}$用引理, 可知存在$\alpha \in V \setminus W_{k-1}$使得$\alpha + W_{k-1} \in V/W_{k-1}$是所有$T_{V/W_{k-1}}(T \in \mathcal{F})$的公共特征向量. 所以对任意$T \in \mathcal{F}$, 存在$c \in F$ 使得

$$T\alpha + W_{k-1} = T_{V/W_{k-1}}(\alpha + W_{k-1}) = c(\alpha + W_{k-1}) = c\alpha + W_{k-1}.$$

设$W_k = W_{k-1} \oplus F\alpha$. 则$T\alpha \in W_k$, 从而$W_k$是$T$的不变子空间. 这就完成了满足要求的子空间序列的构造.

现在, 取V的有序基$\mathcal{B} = \{\alpha_1, \ldots, \alpha_n\}$使得$\{\alpha_1, \ldots, \alpha_i\}$是$W_i$的基, $1 \leq i \leq n$. 则对任意$T \in \mathcal{F}$, $[T]_{\mathcal{B}}$是上三角矩阵. □

注 6.7. 从证明过程容易看出, 如果F不是代数闭域, 但是$\mathcal{F}$中的线性变换的特征多项式都可以分解为一次式的乘积(这等价于$\mathcal{F}$中的线性变换都是可三角化的), 则定理6.26的结论也成立.

§6.5 投影映射

给定$T \in L(V)$. 假设V可以分解为T的不变子空间的直和$V = \bigoplus_{i=1}^k W_i$. 记$T_i = T_{W_i}$. 则对$\alpha = \sum_{i=1}^k \alpha_i$, 有$T\alpha = \sum_{i=1}^k T_i\alpha_i$. 我们称$T$为$T_1, \ldots, T_k$的**直和**. T的性质被这些T_i的性质完全决定. 因此, 如果我们理解了每个T_i, 我们就理解了T. 例如:

- 设$\mathcal{B}_i = \{\alpha_{i1}, \ldots, \alpha_{id_i}\}$是$W_i$的有序基, 则$\mathcal{B} = \{\alpha_{11}, \ldots, \alpha_{1d_1}, \ldots, \alpha_{k1}, \ldots, \alpha_{kd_k}\}$是$V$的有序基, 并且

$$[T]_{\mathcal{B}} = \mathrm{diag}([T_1]_{\mathcal{B}_1}, \ldots, [T_k]_{\mathcal{B}_k}).$$

  特别地, 如果我们可以取$\mathcal{B}_i$使$[T_i]_{\mathcal{B}_i}$简单, 我们就认为$[T]_{\mathcal{B}}$简单.

- $\det(T) = \prod_{i=1}^k \det(T_i)$. 更一般地, $f_T = \prod_{i=1}^k f_{T_i}$.

- $p_T = \mathrm{lcm}(p_{T_1}, \ldots, p_{T_k})$.
- $\mathrm{Ker}(T) = \bigoplus_{i=1}^k \mathrm{Ker}(T_i)$, $\mathrm{Im}(T) = \bigoplus_{i=1}^k \mathrm{Im}(T_i)$.

我们后面的主要目的之一就是选择适当的直和分解, 使每个T_i具有比较基本的形式.

注 6.8.
- 反过来, 如果给定$T_i = L(W_i, W_i)$, 可以定义它们的直和$\bigoplus_{i=1}^k T_i \in L(V)$为

$$\bigoplus_{i=1}^k T_i \left(\sum_{j=1}^k \alpha_j \right) = \sum_{i=1}^k T_i \alpha_i.$$

  这也可以对外直和来定义.
- 不变直和分解就是$F[x]$-子模的直和分解.

§6.6 准素分解

我们准备对T做分解: 准素分解和循环分解. 加在一起, 得到准素循环分解. 其中循环分解推出有理标准形, 准素循环分解推出Jordan标准形.

设V是有限维F-线性空间. 称$T \in L(V)$是**准素的**(primary), 如果p_T是素多项式的幂. 此时, 我们也称V在T的作用下是准素的, 或称V作为$F[x]$-模是准素的.

注 6.9.
- 以后将看到, p_T是素多项式的幂等价于f_T是素多项式的幂.
- 若p素, 则$F[x]$的理想(p^r)称为**准素理想**(也可以称p^r为准素多项式, 但这么叫的不多).

定理 6.29(准素分解). 给定$T \in L(V)$. 将p_T分解为互不相同的首项系数是1的素多项式的幂的乘积$p_T = \prod_{i=1}^k p_i^{r_i}$, 其中$r_i \geq 1$. 则

$$V = \bigoplus_{i=1}^k \mathrm{Ker}(p_i^{r_i}(T)),$$

其中$W_i := \mathrm{Ker}(p_i^{r_i}(T)) \neq \{0\}$, 并且$p_{T_{W_i}} = p_i^{r_i}$.

证明. 记$f_i = \prod_{j \neq i} p_j^{r_j}$. 注意到$p_i^{r_i}(T) f_i(T) = p_T(T) = 0$但$f_i(T) \neq 0$. 所以$p_i^{r_i}(T)$不可逆. 因此$W_i \neq \{0\}$.

W_i无关: 设$\alpha_i \in W_i$, $\sum_{j=1}^k \alpha_j = 0$. 两边取$f_i$得$f_i \alpha_i = 0$. 注意到还有$p_i^{r_i} \alpha_i = 0$并且$f_i$与$p_i^{r_i}$互素. 所以存在$a, b \in F[x]$使$1 = a f_i + b p_i^{r_i}$. 因此$\alpha_i = a f_i \alpha_i + b p_i^{r_i} \alpha_i = 0$.

$V = \sum_{i=1}^k W_i$: 注意到$f_1, \ldots, f_k$互素. 于是存在$g_1, \ldots, g_k$使得$\sum_{i=1}^k g_i f_i = 1$. 设$\alpha \in V$. 则$f_i \alpha \in W_i$. 因此$\alpha = \sum_{i=1}^k g_i f_i \alpha \in \sum_{i=1}^k W_i$.

$p_{T_{W_i}} = p_i^{r_i}$: 首先, 对任意$\alpha_i \in W_i$有$p_i^{r_i} \alpha_i = 0$. 因此$p_{T_{W_i}} | p_i^{r_i}$. 另一方面, 对$\alpha \in V$, 如果$\alpha = \sum_{j=1}^k \alpha_j$, $\alpha_j \in W_j$, 则

$$\prod_{i=1}^k p_{T_{W_i}} \alpha = \sum_{j=1}^k \prod_{i=1}^k p_{T_{W_i}} \alpha_j = 0.$$

所以$p_T | \prod_{i=1}^k p_{T_{W_i}}$. 这推出$p_{T_{W_i}} = p_i^{r_i}$. □

注意到每个W_i是T-不变子空间, 称为V的p_i-**准素分量**(primary component).

第七章　有理标准形和Jordan标准形

§7.1　循环子空间和零化子

给定域F, 记$R = F[x]$. 设V是有限维F-线性空间. 则$T \in L(V)$诱导了V上的R-模结构: $f\alpha = f(T)\alpha$. 对于$\alpha \in V$, V的子空间

$$R\alpha := \{f\alpha \mid f \in R\} = \operatorname{span}\{\alpha, T\alpha, T^2\alpha, \ldots\}$$

称为由α生成的**循环子空间**. 注意它是V的子模, 也称为由α生成的**循环子模**. 容易看出:

- $R\alpha$是包含α的最小的T-不变子空间.
- $\dim R\alpha = 1 \iff \alpha$是特征向量.

如果$V = R\alpha$, 则称α是**循环向量**. 如果V中存在循环向量, 则称T是**循环**的, 并称V是**循环模**.

例 7.1. 设$A = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}$, $T = L_A$. 设$\epsilon_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$, $\epsilon_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$. 注意到$T\epsilon_1 = \epsilon_2$, $T\epsilon_2 = 0$. 因此

$$R\epsilon_1 = V, \quad R\epsilon_2 = F\epsilon_2.$$

因此ϵ_1是循环向量, ϵ_2不是循环向量.

由这个例子可以看出, 当T循环时, 不一定所有非零向量都是循环向量. 为了鉴别一个向量是否是循环向量, 考虑函数

$$\Delta = \Delta_T : V \to \mathbb{Z}, \qquad \Delta(\alpha) = \dim R\alpha.$$

容易看出:

- T循环$\iff \Delta$的最大值为$\dim V$. 此时, $\alpha \in V$是循环向量$\iff \Delta$在α处取得最大值.

对于$\alpha \in V$, 为了考察$\Delta(\alpha)$, 记

$$M(\alpha) := \{f \in R \mid f\alpha = 0\}.$$

引理 7.1. $M(\alpha)$是理想, 并且$p_T \in M(\alpha)$.

证明. $p_T \in M(\alpha)$; $f, g \in M(\alpha), c \in F \Longrightarrow f + cg \in M(\alpha)$; $f \in M(\alpha), g \in R \Longrightarrow fg \in M(\alpha)$. □

我们称$M(\alpha)$为α的**零化理想**, 并称$M(\alpha)$的唯一的monic生成元p_α为α的**零化子**(annihilator). 注意$p_\alpha | p_T$, 并且$\alpha = 0 \iff M(\alpha) = R \iff p_\alpha = 1$. p_α与$\Delta(\alpha)$的联系如下:

引理 7.2. 记$d = \deg p_\alpha$, 则$\{\alpha, T\alpha, \ldots, T^{d-1}\alpha\}$是$R\alpha$的基. 特别地, $\Delta(\alpha) = \deg p_\alpha$.

证明. 线性无关: 设$\sum_{i=0}^{d-1} c_i T^i \alpha = 0$. 记$g = \sum_{i=0}^{d-1} c_i x^i$. 则$g\alpha = 0$. 因此$p_\alpha | g$. 由于$\deg g < \deg p_\alpha$, 只能有$g = 0$, 即$c_i = 0$.

$\operatorname{span} = R\alpha$: "$\subset$"显然. "$\supset$": 对任意$\beta \in R\alpha$, 存在$f \in R$ 使得$\beta = f\alpha$. 设$f = qp_\alpha + r$, 这里$q, r \in R$, $\deg r < d$. 则$\beta = f\alpha = r\alpha \in \operatorname{span}$. □

由这个引理, 考察$\Delta(\alpha)$的最大值转化为考察$\deg p_\alpha$的最大值. 由于$p_\alpha | p_T$, 总有$\Delta(\alpha) \le \deg p_T$. 下面说明等号是可以成立的.

命题 7.3. 存在$\alpha \in V$满足$p_\alpha = p_T$.

证明. 首先考虑T准素的情况. 设$p_T = p^r$, 其中p素, $r \ge 1$. 此时, 对任意$\alpha \in V$, 有$p_\alpha = p^{r_\alpha}$, $0 \le r_\alpha \le r$. 于是一定有α满足$r_\alpha = r$. 否则, 如果对所有α有$r_\alpha \le r - 1$, 则$p^{r-1}\alpha = 0$, 从

而$p^{r-1}(T)=0$, 矛盾.

对一般情况, 设$p_T=\prod_{i=1}^k p_i^{r_i}$, 其中$p_i$为互不相同的monic素多项式, $r_i\geq 1$. 记$W_i:=\mathrm{Ker}(p_i^{r_i}(T))$. 根据准素分解, $V=\bigoplus_{i=1}^k W_i$, 并且$p_{T_{W_i}}=p_i^{r_i}$. 取$\alpha_i\in W_i$使$p_{\alpha_i}=p_i^{r_i}$. 设$\alpha=\sum_{i=1}^k \alpha_i$. 由$p_\alpha\alpha=0$推出$p_\alpha\alpha_i=0$. 从而$p_i^{r_i}|p_\alpha$. 这推出$p_T|p_\alpha$. 只能有$p_\alpha=p_T$. □

推论 7.4. 函数Δ的最大值为$\deg p_T$, 并且$\Delta(\alpha)=\deg p_T\iff p_\alpha=p_T$.

证明. 由于总有$p_\alpha|p_T$, 所以$\Delta(\alpha)=\deg p_\alpha\leq\deg p_T$. 上面命题说明存在$\alpha$使等号成立. 因此$\Delta$的最大值为$\deg p_T$. 进一步地, $\Delta(\alpha)=\deg p_T\iff\deg p_\alpha=\deg p_T\iff p_\alpha=p_T$. □

由此, 容易得到:

命题 7.5. T循环$\iff\deg p_T=\dim V\iff p_T=f_T$. 此时, $\alpha\in V$是循环向量$\iff p_\alpha=p_T$.

证明.

$$T\text{循环}\iff\Delta\text{的最大值等于}\dim V\iff\deg p_T=\dim V\iff p_T=f_T.$$

此时, $\alpha\in V$是循环向量$\iff\Delta(\alpha)=\deg p_T\iff p_\alpha=p_T$. □

现在我们讨论当T循环时它的矩阵的形式. 对于首项系数是1的n次多项式

$$f=x^n+a_{n-1}x^{n-1}+\cdots+a_1x+a_0,$$

定义它的**相伴矩阵**或**友阵**(companion matrix)$C_f\in F^{n\times n}$为

$$C_f=\begin{bmatrix}0&0&\cdots&0&-a_0\\1&0&\cdots&0&-a_1\\0&1&\cdots&0&-a_2\\\vdots&\vdots&\vdots&\vdots&\vdots\\0&0&\cdots&1&-a_{n-1}\end{bmatrix}.$$

下面说明T循环$\iff$它在某个基下的矩阵是某多项式的相伴矩阵.

命题 7.6. 设$\dim V=n$, $T\in L(V)$.

(1) 如果T循环, α是循环向量, 则T在有序基$\{\alpha,T\alpha,\ldots,T^{n-1}\alpha\}$下的矩阵为$C_{p_\alpha}=C_{p_T}$.

(2) 如果T在某有序基下的矩阵为C_f, 其中f首项系数是1并且次数是n, 则T循环并且$p_T=f$.

证明. 注意到对于$f=x^n+a_{n-1}x^{n-1}+\cdots+a_1x+a_0$和$V$的有序基$\mathcal{B}=\{\alpha_1,\ldots,\alpha_n\}$, $[T]_{\mathcal{B}}=C_f$意味着

$$T\alpha_i=\alpha_{i+1},\qquad 1\leq i\leq n-1,\tag{7.1}$$

$$T\alpha_n=-a_0\alpha_1-a_1\alpha_2-\cdots-a_{n-1}\alpha_n.\tag{7.2}$$

(1) 记$\alpha_i=T^{i-1}\alpha$, $1\leq i\leq n$. 则(7.1)和(7.2)对$f=p_\alpha$成立. 这说明T的矩阵为C_{p_α}.

(2) 设对于上面的f和$\mathcal{B}$有$[T]_{\mathcal{B}}=C_f$, 则(7.1)和(7.2)成立. (7.1)说明α_1是循环向量, 从而T循环并且$p_{\alpha_1}=p_T$. (7.2)说明$f\alpha_1=0$, 即$p_{\alpha_1}|f$. 两者结合, 得$p_T|f$. 但$\deg p_T=\deg f=n$. 所以只能有$p_T=f$. □

推论 7.7. 设f是首项系数是1的多项式. 则$f_{C_f}=p_{C_f}=f$.

证明. 对T_{C_f}和$F^{n\times 1}$的标准基应用(2). □

注 7.1. $f_{C_f}=f$也可以直接计算得到. 这一计算实际上给出了Cayley-Hamilton定理的另一证明: 为证明$f_T(T)=0$, 只需证对任意$\alpha\in V$有$f_T\alpha=0$. 考虑循环子空间$R\alpha$. 这一计算给

出$f_{T_{R\alpha}}=f_{C_{p_\alpha}}=p_\alpha$, 从而$p_\alpha|f_T$, 即$f_T\alpha=0$.

§7.2　循环分解和有理标准形

定理 7.8(循环分解). 设V是有限维F-线性空间, 设$T\in L(V)$. 则存在$\alpha_1,\dots,\alpha_r\in V\setminus\{0\}$满足:

- $V=\bigoplus_{i=1}^r R\alpha_i$;
- 设$p_i=p_{\alpha_i}$. 则$p_r|p_{r-1}|\cdots|p_1$.

整数r和序列$p_1,\dots,p_r$被T唯一决定, 并且

$$p_T=p_1,\qquad f_T=\prod_{i=1}^r p_i.$$

定理中的$p_1,\dots,p_r$称为T的**不变因子**(invariant factors). 对$A\in F^{n\times n}$, L_A的不变因子也称为A的不变因子.

注 7.2.

- p_i也等于$T_{R\alpha_i}$的特征多项式和最小多项式.
- 有的文献中用$p_1|p_2|\cdots|p_r$.

设$\alpha\in V\setminus\{0\}$, $R\alpha\neq V$. 我们希望找$\beta\in V\setminus R\alpha$满足$R\alpha\cap R\beta=\{0\}$. 具体来说, 考虑商空间$V/R\alpha$和$T$诱导的映射$T_{V/R\alpha}$, 对于$L\in(V/R\alpha)\setminus\{0\}$, 我们希望找$\beta\in L$满足$R\alpha\cap R\beta=\{0\}$. 这一般是做不到的, 与$\alpha$的选取有关.

引理 7.9. 设$\alpha\in V$.

(1) 对$L\in V/R\alpha$和$\beta\in L$, 有$R\alpha\cap R\beta=\{0\}\iff p_\beta=p_L$.

(2) 如果$p_\alpha=p_T$, 则对任意$L\in V/R\alpha$, 存在$\beta\in L$满足$p_\beta=p_L$.

证明. (1) "$\Longrightarrow$". 显然总有$p_L|p_\beta$. 另一方面$p_LL=0$推出$p_L\beta\in R\alpha$. 所以$p_L\beta\in R\alpha\cap R\beta=\{0\}$. 因此$p_\beta|p_L$.

"$\Longleftarrow$". 设$\delta\in R\alpha\cap R\beta$, 则$\delta=q\beta\in R\alpha$. 投到$V/R\alpha$上得$qL=0$. 所以$p_\beta=p_L|q$. 因此$\delta=q\beta=0$.

(2) 首先任取$\beta_0\in L$. 则$p_L\beta_0\in R\alpha$. 设$p_L\beta_0=f\alpha$, $p_T=gp_L$. 则$gf\alpha=gp_L\beta_0=0$. 于是$gp_L=p_T=p_\alpha|gf$, 因此$p_L|f$. 设$f=p_Lh$, $\beta=\beta_0-h\alpha\in L$. 则$p_L\beta=0$, 即$p_\beta|p_L$. 另一方面, 总有$p_L|p_\beta$, 所以$p_\beta=p_L$. □

定理7.8存在性部分的证明. 对$\dim V$归纳. $\dim V=1$时显然. 一般情况, 取$\alpha_1\in V$使$p_{\alpha_1}=p_T$. 由于$\dim V/R\alpha_1<\dim V$, 由归纳假设, 存在$L_2,\dots,L_r\in(V/R\alpha_1)\setminus\{0\}$使得$V/R\alpha_1=\bigoplus_{i=2}^r RL_i$, 并且$p_{L_r}|\cdots|p_{L_2}$. 由引理7.9, 对$2\le i\le r$, 存在$\alpha_i\in L_i$满足$p_{\alpha_i}=p_{L_i}$并且$R\alpha_1\cap R\alpha_i=\{0\}$. 显然$p_{\alpha_r}|\cdots|p_{\alpha_1}$. 我们验证$V=\bigoplus_{i=1}^r R\alpha_i$. 设$\sum_{i=1}^r g_i\alpha_i=0$. 投到$V/R\alpha_1$上, 有$\sum_{i=2}^r g_iL_i=0$. 因此$g_iL_i=0$, 即$g_i\alpha_i\in R\alpha_1$. 由于$R\alpha_1\cap R\alpha_i=\{0\}$, 这推出$g_i\alpha_i=0$. 进而$g_1\alpha_1=0$. 因此$R\alpha_1,\dots,R\alpha_r$无关. 另一方面, 对任意$\gamma\in V$, 把$\gamma+R\alpha_1\in V/R\alpha_1$分解为$\gamma+R\alpha_1=\sum_{i=2}^r g_iL_i$, 则$\gamma-\sum_{i=2}^r g_i\alpha_i\in R\alpha_1$. 因此$V=\bigoplus_{i=1}^r R\alpha_i$. □

下面考虑唯一性和关于p_T, f_T的部分.

定理7.8其余部分的证明. 设$\alpha_1,\dots,\alpha_r\in V\setminus\{0\}$满足$V=\bigoplus_{i=1}^r R\alpha_i$, 并且$p_i:=p_{\alpha_i}$满足$p_r|p_{r-1}|\cdots|p_1$. 注意到:

- $p_1|p_T$, 并且

$$p_i = p_{T_{R\alpha_i}} \Longrightarrow p_i(T_{R\alpha_i}) = 0 \Longrightarrow p_1(T_{R\alpha_i}) = 0 \Longrightarrow p_1(T) = 0 \Longrightarrow p_T|p_1.$$

  因此$p_T = p_1$.

- $f_T = \prod_{i=1}^r f_{T_{R\alpha_i}} = \prod_{i=1}^r p_i$.

如果还有$\beta_1, \dots, \beta_s \in V \setminus \{0\}$满足$V = \bigoplus_{i=1}^s R\beta_i$, 并且$q_i := p_{\beta_i}$满足$q_s|q_{s-1}|\cdots|q_1$, 则类似地有$q_T = q_1$, $f_T = \prod_{i=1}^s q_i$. 因此$p_1 = q_1$, $\prod_{i=1}^r p_i = \prod_{i=1}^s q_i$. 我们需要证明$r = s$并且$p_i = q_i$. 如果这不成立, 设存在$2 \le t \le \min\{r, s\}$满足$p_t \ne q_t$, 并且当$i < t$时有$p_i = q_i$. 不妨设$q_t \nmid p_t$. 在$\bigoplus_{i=1}^r R\alpha_i = V = \bigoplus_{i=1}^s R\beta_i$两边乘$p_t$, 得

$$\bigoplus_{i=1}^{t-1} Rp_t\alpha_i = p_tV = \bigoplus_{i=1}^{t-1} Rp_t\beta_i \oplus \bigoplus_{i=t}^{s} Rp_t\beta_i.$$

容易看出:

- f, g monic, $p_\alpha = fg \Longrightarrow p_{f\alpha} = g$: 事实上, $p_{f\alpha}|h \Longleftrightarrow h(f\alpha) = 0 \Longleftrightarrow p_\alpha = fg|fh \Longleftrightarrow g|h$.

因此, 当$i < t$时有

$$\dim Rp_t\alpha_i = \deg p_{p_t\alpha_i} = \deg(p_i/p_t) = \deg(q_i/p_t) = \deg p_{p_t\beta_i} = \dim Rp_t\beta_i.$$

这推出$\bigoplus_{i=t}^s Rp_t\beta_i = \{0\}$. 特别地, $p_t\beta_t = 0$. 因此$q_t|p_t$. 矛盾. □

推论 7.10. p_T与f_T有相同的素因子. □

在定理7.8的结论下, $\mathcal{B}_i := \{\alpha_i, T\alpha_i, \dots, T^{d_i-1}\alpha_i\}$是$R\alpha_i$的有序基, 这里$d_i = \deg p_i = \dim R\alpha_i$. 于是, $[T_{R\alpha_i}]_{\mathcal{B}_i} = C_{p_i}$. 因此, 在$V$的有序基$\mathcal{B} = (\mathcal{B}_1, \dots, \mathcal{B}_r)$下,

$$[T]_{\mathcal{B}} = \operatorname{diag}(C_{p_1}, \dots, C_{p_r}).$$

定义 7.1. 矩阵$A \in F^{n\times n}$称为**有理形矩阵**, 如果存在非常数的首一多项式$p_1, \dots, p_r \in F[x]$, 满足$\sum_{i=1}^r \deg p_i = n$并且$p_r|\cdots|p_1$, 使得

$$A = \operatorname{diag}(C_{p_1}, \dots, C_{p_r}).$$

定理 7.11. 存在V的有序基$\mathcal{B}$使$[T]_{\mathcal{B}}$是有理形矩阵, 并且这样的有理形矩阵是唯一的.

证明. 只需证明唯一性. 设T在某个有序基$\mathcal{B}'$下的矩阵是有理形矩阵$\operatorname{diag}(C_{q_1}, \dots, C_{q_s})$, 其中$q_s|\cdots|q_1$. 设$\mathcal{B}' = (\mathcal{B}'_1, \dots, \mathcal{B}'_s)$, $|\mathcal{B}'_i| = \deg q_i$. 则$\mathcal{B}_i$中的第一个向量$\beta_i$满足$R\beta_i = \operatorname{span}\mathcal{B}_i$并且$p_{\beta_i} = q_i$. 由定理7.8的唯一性部分, $q_1, \dots, q_s$只能是T的不变因子序列, 因此C_{q_i}被T决定. □

推论 7.12. 任意矩阵$A \in F^{n\times n}$相似于唯一的有理形矩阵, 称为A的**有理标准形**(rational canonical form或rational normal form).

证明. 取$T = L_A$. □

叫有理标准形的原因: 若$A \in \mathbb{Q}^{n\times n}$则有理标准形也$\in \mathbb{Q}^{n\times n}$.

推论 7.13. 两个矩阵相似$\Longleftrightarrow$有相同的有理标准形$\Longleftrightarrow$有相同的不变因子序列. □

下面说明矩阵的有理标准形和不变因子不依赖于域的选取.

推论 7.14. 设$A \in F^{n\times n}$的有理标准形为A', 不变因子序列为$p_1, \dots, p_r$. 如果对某个子域$K \subset F$有$A \in K^{n\times n}$, 则$A' \in K^{n\times n}$并且$A \overset{K}{\sim} A'$, $p_1, \dots, p_r \in K[x]$. 特别地, $p_A \in K[x]$.

证明. 设A在$K^{n\times n}$中的有理标准形为$A'' \in K^{n\times n}$. A'和A''都是$F^{n\times n}$中的有理形矩阵并且都与A相似, 由推论7.12的唯一性, $A' = A''$. 因此$A' \in K^{n\times n}$并且$A \overset{K}{\sim} A'$. 由$A' \in K^{n\times n}$即得

到$p_1,\dots,p_r\in K[x]$. $\square$

下面看有理标准形和不变因子的两个应用.

命题 7.15. 设$A,B\in F^{n\times n}$相似. 如果对子域$K\subset F$有$A,B\in K^{n\times n}$, 则A,B在K上也相似.

证明. A,B相似推出它们在$F^{n\times n}$中的有理标准形相同, 设为C. 由$A,B\in K^{n\times n}$推出$C\in K^{n\times n}$并且$A\overset{K}{\sim}C$, $B\overset{K}{\sim}C$. 因此$A\overset{K}{\sim}B$. $\square$

命题 7.16. 设$A\in F^{n\times n}$. 则A与A^t相似.

证明. 先假设A是某个首项系数是1的多项式f的相伴矩阵. 此时有$f_A=p_A=f$. 因此A只有一个不变因子f. 而

$$f_{A^t}=\det(xI-A^t)=\det((xI-A)^t)=\det(xI-A)=f_A=f,$$
$$g(A^t)=g(A)^t\Longrightarrow \text{“}g(A)=0\Longleftrightarrow g(A^t)=0\text{”}\Longrightarrow p_{A^t}=p_A=f.$$

因此A^t也只有一个不变因子f. 因此A^t与A相似.

一般情况, 设A的有理标准形为$\mathrm{diag}(C_{p_1},\dots,C_{p_r})$. 则$A^t\sim\mathrm{diag}(C_{p_1}^t,\dots,C_{p_r}^t)$. 我们已经证明$C_{p_i}^t\sim C_{p_i}$. 所以$\mathrm{diag}(C_{p_1}^t,\dots,C_{p_r}^t)\sim\mathrm{diag}(C_{p_1},\dots,C_{p_r})$. 因此$A^t\sim A$. $\square$

例 7.2. 2×2的有理形矩阵只能是cI_2或$\begin{bmatrix}0&*\\1&*\end{bmatrix}$.

给定$A\in F^{n\times n}$, 如果知道了A的不变因子为$p_1,\dots,p_r$, 则A的有理标准形为$\mathrm{diag}(C_{p_1},\dots,C_{p_r})$. 如何求可逆矩阵$P$使得$P^{-1}AP=\mathrm{diag}(C_{p_1},\dots,C_{p_r})$?

命题 7.17. 设$A\in F^{n\times n}$, $\mathcal{B}$是$F^{n\times 1}$的有序基, P是由$\mathcal{B}$的列向量按次序排列得到的$n\times n$可逆矩阵. 则$P^{-1}AP=[L_A]_{\mathcal{B}}$.

证明. 设$\mathcal{B}=\{\alpha_1,\dots,\alpha_n\}$. 则$P=[\alpha_1,\dots,\alpha_n]$. 记$B=[L_A]_{\mathcal{B}}$. 则

$$A\alpha_j=\sum_{i=1}^n B_{ij}\alpha_i.$$

这意味着

$$AP=[A\alpha_1,\dots,A\alpha_n]=\left[\sum_{i=1}^n B_{i1}\alpha_i,\dots,\sum_{i=1}^n B_{in}\alpha_i\right]=PB,$$

即$P^{-1}AP=B$. $\square$

因此, 如果$\alpha_1,\dots,\alpha_r\in F^{n\times 1}\setminus\{0\}$满足$F^{n\times 1}=\bigoplus_{i=1}^r R\alpha_i$并且$p_{\alpha_i}=p_i$, 则可逆矩阵

$$P=[\alpha_1,A\alpha_1,\dots,A^{d_1-1}\alpha_1,\dots,\alpha_r,A\alpha_r,\dots,A^{d_r-1}\alpha_r],\qquad d_i=\deg p_i=\dim R\alpha_i,$$

满足

$$P^{-1}AP=\mathrm{diag}(C_{p_1},\dots,C_{p_r}).$$

例 7.3. 设$A=\begin{bmatrix}5&-6&-6\\-1&4&2\\3&-6&-4\end{bmatrix}\in\mathbb{R}^{3\times 3}$. 习题课中求过(见P188例3):

$$f_A=(x-1)(x-2)^2.$$

容易验证$(A-I)(A-2I)=0$. 因此

$$p_A=(x-1)(x-2).$$

所以不变因子为

$$p_1 = p_A = (x-1)(x-2) = x^2 - 3x + 2,$$
$$p_2 = f_A/p_A = x - 2.$$

因此A的有理标准形为

$$\operatorname{diag}(C_{p_1}, C_{p_2}) = \operatorname{diag}\left(\begin{bmatrix} 0 & -2 \\ 1 & 3 \end{bmatrix}, 2\right) = \begin{bmatrix} 0 & -2 & 0 \\ 1 & 3 & 0 \\ 0 & 0 & 2 \end{bmatrix}.$$

为了寻找P使$P^{-1}AP = B$, 由上面的分析, 只需寻找$\alpha_1, \alpha_2 \in F^{3\times 1} \setminus \{0\}$满足$F^{3\times 1} = R\alpha_1 \oplus R\alpha_2$并且$p_{\alpha_i} = p_i$. 此时可逆矩阵$P = [\alpha_1, A\alpha_1, \alpha_2]$就满足$P^{-1}AP = \operatorname{diag}(C_{p_1}, C_{p_2})$. 注意到:

- $p_{\alpha_1} = p_1 \iff \alpha_1$不是特征向量.
- $p_{\alpha_2} = p_2 \iff \alpha_2$是特征值2的特征向量.

此时,

- $F^{3\times 1} = R\alpha_1 \oplus R\alpha_2 \iff \alpha_2 \notin R\alpha_1$.

尝试$\alpha_1 = \epsilon_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$. 由于$A\alpha_1 = \begin{bmatrix} 5 \\ -1 \\ 3 \end{bmatrix}$, 所以$\alpha_1$不是特征向量. 因此可选这个$\alpha_1$. 习题课求过(见P188):

$$V_2 = \{(x_1, x_2, x_3)^t \mid x_1 = 2x_2 + 2x_3\}.$$

我们需要找$\alpha_2 \in V_2$使得$\alpha_2 \notin R\alpha_1$. 注意到$R\alpha_1 = \operatorname{span}\{\alpha_1, A\alpha_1\}$, 其中的向量总有$x_2, x_3$同时为0或同时不为0. 取$\alpha_2 = \begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix}$. 则$\alpha_2 \in V_2 \setminus R\alpha_1$. 此时, $P = \begin{bmatrix} 1 & 5 & 2 \\ 0 & -1 & 1 \\ 0 & 3 & 0 \end{bmatrix}$.

事实上, 因为A可对角化, 有$F^{3\times 1} = V_1 \oplus V_2$. 取$V_1$的基$\{\beta_1\}$和$V_2$ 的基$\{\gamma_1, \gamma_2\}$. 则取$\alpha_1 = \beta_1 + \gamma_1$, $\alpha_2 = \gamma_2$即可. □

命题 7.18. 设T可对角化, $\sigma(T) = \{c_1, \ldots, c_k\}$.

(1) 设$\alpha = \sum_{i=1}^k \beta_i$, 其中$\beta_i \in V_{c_i}$, 则$R\alpha = \operatorname{span}\{\beta_1, \ldots, \beta_k\}$, $p_\alpha = \prod_{\beta_i \neq 0}(x - c_i)$.

(2) 设$d_i = \dim V_{c_i}$, 则不变因子为$p_j := \prod_{d_i \geq j}(x - c_i)$, 其个数为$r := \max d_i$.

证明. (1) 注意到$V = \bigoplus_{i=1}^k V_{c_i}$. 由于$R\alpha = \{f\alpha \mid f \in F[x]\}$, 而

$$f\alpha = \sum_{i=1}^k f\beta_i = \sum_{i=1}^k f(c_i)\beta_i,$$

所以$R\alpha \subset \operatorname{span}\{\beta_1, \ldots, \beta_k\}$. 另一方面, 对任意$\sum_{i=1}^k t_i\beta_i \in \operatorname{span}\{\beta_1, \ldots, \beta_k\}$, 存在$f$满足$f(c_i) = t_i$, 此时$\sum_{i=1}^k t_i\beta_i = f\alpha \in R\alpha$. 因此$R\alpha \supset \operatorname{span}\{\beta_1, \ldots, \beta_k\}$. 我们还有

$$f\alpha = 0 \iff f(c_i)\beta_i = 0 \iff \beta_i \neq 0\text{时有}f(c_i) = 0 \iff \prod_{\beta_i \neq 0}(x - c_i) | f.$$

因此$p_\alpha = \prod_{\beta_i \neq 0}(x - c_i)$.

(2) 取V_{c_i}的有序基$\mathcal{B}_i = \{\beta_{i1}, \ldots, \beta_{id_i}\}$. 对于$1 \leq j \leq r$, 考虑$\alpha_j = \sum_{d_i \geq j} \beta_{ij}$. 由(1),

$$R\alpha_j = \operatorname{span}\{\beta_{ij} \mid d_i \geq j\}, \qquad p_{\alpha_j} = p_j.$$

只需注意到$V = \bigoplus_{j=1}^r R\alpha_j$并且$p_r | p_{r-1} | \cdots | p_1$. □

§7.3 准素循环分解和Jordan标准形

称T为**不可分解的**(indecomposable), 如果V不能分解成两个非零不变子空间的直和.

命题 7.19. 任何T总是有限个不可分解变换的直和.

证明. 对$\dim V$归纳. $\dim V = 1$时显然. 一般情况, 如果T不可分解, 则无需证明. 否则, V是两个非零不变子空间的直和, 于是可以对每个直和项用归纳假设. □

定理 7.20. 设V是有限维F-线性空间, $T \in L(V,V)$. 则T不可分解$\Longleftrightarrow$ T是准素循环的.

证明. "$\Longrightarrow$". 设T不可分解. 则T的准素分解和循环分解中都只有一项. 因此T是准素循环的.

"$\Longleftarrow$". 设T是准素循环的. 则$f_T = p_T = p^r$, 其中p素. 假设有不变直和分解$V = V_1 \oplus V_2$. 则$f_{T_{V_1}} f_{T_{V_2}} = f_T = p^r$, 从而$f_i = p^{r_i}$, $r_i \geq 0$, $r_1 + r_2 = r$. 另一方面, 设$p_i = p^{s_i}$, $s_i \leq r_i$. 则$p^r = p_T = \operatorname{lcm}(p_1, p_2) = p^{\max\{s_1, s_2\}}$. 从而$r_1 + r_2 = r = \max\{s_1, s_2\} \leq \max\{r_1, r_2\}$. 只能有$r_1 = 0$或$r_2 = 0$. 因此$V_1 = \{0\}$ 或$V_2 = \{0\}$. □

因此, 我们得到:

定理 7.21(准素循环分解). 设V是有限维F-线性空间, 设$T \in L(V)$. 则V可以分解为一些准素循环的非零不变子空间的直和$V = \bigoplus_{i=1}^{s} V_i$. 设$q_i = p_{T_i}$. 准素多项式序列$q_1, \ldots, q_s$在不计排列次序的意义下被$T$唯一决定, 称为$T$的**基本因子**或**初等因子**(elementary divisors).

证明. 存在性已证明. 只需证唯一性. 设$q_1, \ldots, q_s$的所有素因子为$p_1, \ldots, p_k$. 也就是说, $p_1, \ldots, p_k$是互不相同的素多项式, 每个p_i是某个q_j的素因子, 并且每个q_j都是某个p_i的幂. 将序列$(V_1, q_1), \ldots, (V_s, q_s)$重新排列为下表:

$$\begin{array}{ccc}
(V_{11}, p_1^{r_{11}}), & \ldots, & (V_{1,d_1}, p_1^{r_{1,d_1}}), \\
\vdots & \vdots & \vdots \\
(V_{k1}, p_k^{r_{k1}}), & \ldots, & (V_{k,d_k}, p_k^{r_{k,d_k}}),
\end{array}$$

满足对任意$1 \leq i \leq k$有$r_{i1} \geq \cdots \geq r_{i,d_i}$. 设$d = \max\{d_1, \ldots, d_k\}$. 对$1 \leq j \leq d$, 设$W_j = \bigoplus_{i: d_i \geq j} V_{ij}$. 则:

- $V = \bigoplus_{j=1}^{d} W_j$.
- W_j是循环的:

$$f_{T_{W_j}} = \prod_{i: d_i \geq j} f_{T_{V_{ij}}} = \prod_{i: d_i \geq j} p_{T_{V_{ij}}} = \prod_{i: d_i \geq j} p_i^{r_{ij}},$$

$$p_{T_{W_j}} = \operatorname{lcm}\{p_{T_{V_{ij}}} \mid i : d_i \geq j\} = \operatorname{lcm}\{p_i^{r_{ij}} \mid i : d_i \geq j\} = \prod_{i: d_i \geq j} p_i^{r_{ij}}.$$

  因此$f_{T_{W_j}} = p_{T_{W_j}}$.
- $p_{T_{W_j}}$的公式也说明$p_{T_{W_d}} | \cdots | p_{T_{W_1}}$.

因此$p_{T_{W_1}}, \ldots, p_{T_{W_d}}$就是$T$的不变因子序列. 它是被$T$唯一决定的. 因此上表中的第$j$列是$p_{T_{W_j}}$的所有准素因子, 也是被$T$决定的. □

我们可以从四个角度理解准素循环分解:

- 直接把V分解为不可分解的不变子空间的直和.
- 先对V做准素分解, 再对每个直和项做循环分解, 此时每个小块都是准素循环的. 注意准素的不变子空间还是准素的.

- 先对V做循环分解, 再对每个直和项做准素分解, 此时每个小块都是准素循环的. 注意循环子空间的不变子空间还是循环子空间. 证明如下: 设$W \subset R\alpha$是不变子空间. 考虑$\beta := p_{\alpha+W}\alpha \in W$. 我们验证$W = R\beta$. 显然$R\beta \subset W$. 另一方面, 对任意$\gamma \in W$, 存在$f \in R$满足$\gamma = f\alpha$. $\gamma \in W$推出$p_{\alpha+W}|f$. 设$f = gp_{\alpha+W}$. 则$\gamma = gp_{\alpha+W}\alpha = g\beta \in R\beta$. 因此$W \subset R\beta$.
- 同时对V做准素分解$V = \bigoplus_i V_i$和循环分解$V = \bigoplus_j W_i$. 则$V = \bigoplus_{i,j} V_i \cap W_j$是准素循环分解. (容易验证$V$确实等于$\sum_{i,j} V_i \cap W_j$.)

设$V = \bigoplus_{i=1}^s V_i$是准素循环分解, 相应的基本因子序列为$q_1, \ldots, q_s$. 则在适当有序基下, T_{V_i}的矩阵为C_{q_i}, 从而T的矩阵为$\mathrm{diag}(C_{q_1}, \ldots, C_{q_s})$. 这样得到的"准素有理标准形"与域$F$有关, 因此比有理标准形应用少. 此时更常用的是取另一种有序基得到的Jordan标准形. 为此, 我们引入下面的概念.

定义 7.2. 设$T \in L(V)$. 如果存在正整数r满足$T^r = 0$, 则称T是**幂零**的.

命题 7.22. 设$\dim V = n$, $T \in L(V)$. TFAE:

(1) T幂零.

(2) p_T是x的幂.

(3) $f_T = x^n$.

(4) $T^n = 0$.

证明. "(1)$\Longrightarrow$(2)". 设$N^r = 0$. 则$p_T|x^r$. 因此p_T是x的幂.

"(2)$\Longrightarrow$(3)". 由推论7.10, f_T的素因子只有x. 但$\dim f_T = n$. 所以只能有$f_T = x^n$.

"(3)$\Longrightarrow$(4)". $T^n = f_T(T) = 0$.

"(4)$\Longrightarrow$(1)". 显然. □

现在假设T准素循环, 并且$f_T = p_T = (x-c)^n$. 我们的做法是考虑幂零变换$N = T - cI$的有理标准形. 我们称形如

$$J_n(c) := cI_n + C_{x^n} = \begin{bmatrix} c & 0 & 0 & \cdots & 0 & 0 \\ 1 & c & 0 & \cdots & 0 & 0 \\ 0 & 1 & c & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & c & 0 \\ 0 & 0 & 0 & \cdots & 1 & c \end{bmatrix}$$

的矩阵为**基本Jordan矩阵**或特征值为c的n阶**Jordan块**(Jordan block).

命题 7.23. 设$\dim V = n$, $T \in L(V)$, $c \in F$.

(1) 如果T准素循环, 并且$f_T = p_T = (x-c)^n$, 则T在某有序基下的矩阵为$J_n(c)$.

(2) 如果T在某有序基下的矩阵为$J_n(c)$, 则T准素循环, 并且$f_T = p_T = (x-c)^n$.

特别地, 如果F是代数闭域, 则T准素循环$\Longleftrightarrow$它在某个基下的矩阵是Jordan 块.

证明. 设$N = T - cI$. 则对任意有序基$\mathcal{B}$有

$$[T]_{\mathcal{B}} = [cI + N]_{\mathcal{B}} = c[I]_{\mathcal{B}} + [N]_{\mathcal{B}} = I_n + [N]_{\mathcal{B}}.$$

因此,

$$[T]_{\mathcal{B}} = J_n(c) \Longleftrightarrow [N]_{\mathcal{B}} = J_n(0) = C_{x^n}.$$

此外,

$$f_T = p_T = (x-c)^n \Longleftrightarrow f_N = p_N = x^n.$$

(1) 若T准素循环并且$f_T = p_T = (x-c)^n$, 则$f_N = p_N = x^n$. 因此N循环并且存在$\mathcal{B}$使得$[N]_{\mathcal{B}} = C_{x^n}$. 从而$[T]_{\mathcal{B}} = J_n(c)$.

(2) 如果存在$\mathcal{B}$使$[T]_{\mathcal{B}} = J_n(c)$, 则$[N]_{\mathcal{B}} = C_{x^n}$, 从而$f_N = p_N = x^n$, 因此$f_T = p_T = (x-c)^n$. 这推出$T$准素循环. □

对于一般情况, 假设f_T可以分解为一次式的乘积. 取准素循环分解$V = \bigoplus_{i=1}^s V_i$. 则基本因子形如$q_i = p_{T_{V_i}} = (x-c_i)^{n_i}$. 于是$T_{V_i}$在$V_i$的某有序基下的矩阵为$J_{n_i}(c_i)$. 因此, T在V的某有序基下的矩阵为

$$\operatorname{diag}(J_{n_1}(c_1), \ldots, J_{n_m}(c_m)).$$

这样的矩阵称为**Jordan形矩阵**.

定理 7.24. 假设$T \in L(V)$的特征多项式可以分解为一次式的乘积. 则存在V的有序基$\mathcal{B}$使得$[T]_{\mathcal{B}}$是Jordan形矩阵, 并且Jordan块在相差次序的意义下是唯一的, 称为T的**Jordan标准形**(Jordan canonical form或Jordan normal form). 特别地, 如果F是代数闭域, 则任意T存在Jordan标准形.

证明. 只需证明唯一性. 设T在某个有序基$\mathcal{B}$下的矩阵是Jordan形矩阵$\operatorname{diag}(J_{n_1}(c_1), \ldots, J_{n_m}(c_m))$. 设$\mathcal{B} = (\mathcal{B}_1, \ldots, \mathcal{B}_s)$, $|\mathcal{B}_i| = n_i$. 则$V_i := \operatorname{span}\mathcal{B}_i$是非零不变子空间, 并且$[T_{V_i}]_{\mathcal{B}_i} = J_{n_i}(c_i)$. 因此$T_{V_i}$ 准素循环, 并且$p_{T_{V_i}} = (x-c_i)^{n_i}$. 由定理7.21的唯一性部分, $(x-c_1)^{n_1}, \ldots, (x-c_m)^{n_m}$ 只能是T的基本因子序列, 它们在相差次序的意义下被T 决定. 因此$J_{n_i}(c_i)$相差次序的意义下被T决定. □

推论 7.25. 假设$A \in F^{n\times n}$的特征多项式可以分解为一次式的乘积. 则A相似于某个Jordan形矩阵, 并且Jordan块在相差次序的意义下是唯一的, 称为A的**Jordan标准形**. 特别地, 如果F是代数闭域, 则任意方阵存在Jordan标准形.

注 7.3.

- 这里与书上稍有区别: 书上要求相同特征值在一起, 并且从大到小排列.
- 更多文献中把“1”写在主对角线上方. 我们为了方便, 与教材保持一致. 注意到两者是相似的. 事实上, 如果T在有序基$\{\alpha_1, \ldots, \alpha_n\}$下的矩阵是Jordan形矩阵$J$, 则$T$在有序基$\{\alpha_n, \ldots, \alpha_1\}$下的矩阵是$J^t$.

从T的Jordan标准形$[T]_{\mathcal{B}}$可以直接看出很多T的性质:

- $[T]_{\mathcal{B}}$的对角元集合为$\sigma(T)$. 设为$\{c_1, \ldots, c_k\}$. 如果c_i在$[T]_{\mathcal{B}}$的对角元中出现d_i次, 则$d_i = \dim \operatorname{Ker}((T-c_iI)^n)$, 其中$n = \dim V$, 并且$f_T = \prod_{i=1}^k (x-c_i)^{d_i}$.
- 设特征值为c_i的Jordan块的最大阶数为r_i. 则$p_T = \prod_{i=1}^k (x-c_i)^{r_i}$. 特别地, T可对角化$\Longleftrightarrow$所有Jordan块都是1阶的.
- 特征值为c_i的Jordan块的个数为$\dim V_{c_i}$.

例 7.4. 2×2的Jordan形矩阵只能是$\begin{bmatrix} c_1 & 0 \\ 0 & c_2 \end{bmatrix}$或$\begin{bmatrix} c & 0 \\ 1 & c \end{bmatrix}$.

例 7.5. $\begin{bmatrix} 2 & 0 & 0 & 0 \\ 1 & 2 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 1 & 2 \end{bmatrix}$和$\begin{bmatrix} 2 & 0 & 0 & 0 \\ 1 & 2 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 2 \end{bmatrix}$都是Jordan形矩阵, 特征多项式都是$(x-2)^4$, 最小多项式

都是$(x-2)^2$, 但它们不相似. (考察$\dim\operatorname{Ker}(A-2I)$.)

例 7.6. 设$A=\begin{bmatrix}2&0&0\\a&2&0\\b&c&-1\end{bmatrix}\in\mathbb{C}^{3\times3}$. 则$f_A=(x-2)^2(x+1)$. 如果$p_A=f_A$, 则$A$的Jordan标准形为$\begin{bmatrix}2&0&0\\1&2&0\\0&0&-1\end{bmatrix}$. 如果$p_A=(x-2)(x+1)$, 则$A$的Jordan标准形为$\begin{bmatrix}2&0&0\\0&2&0\\0&0&-1\end{bmatrix}$. 后者发生

$$\Longleftrightarrow (A-2I)(A+I)=\begin{bmatrix}0&0&0\\a&0&0\\b&c&-3\end{bmatrix}\begin{bmatrix}3&0&0\\a&3&0\\b&c&0\end{bmatrix}=\begin{bmatrix}0&0&0\\3a&0&0\\ac&0&0\end{bmatrix}=0\Longleftrightarrow a=0.$$

现在考察与Jordan标准形密切相关的一个问题. 为了简单起见, 假设F是代数闭域. 设$A\in F^{n\times n}$的Jordan标准形为$J=P^{-1}AP$, 其中$P\in GL_n(F)$. 记J_d是把J中不在主对角线上的矩阵元"1"都替换为"0"后得到的矩阵, J_n是把J的对角元都替换为"0"后得到的矩阵. 则J_d是对角矩阵, J_n幂零, J_d与J_n可交换, 并且$J=J_d+J_n$. 记$A_d=PJ_dP^{-1}$, $A_n=PJ_nP^{-1}$. 则A_d可对角化, A_n幂零, A_d与A_n可交换, 并且$A=A_d+A_n$. 下面说明这样得到的A_d和A_n与J和P的选取无关.

定理 7.26. 设V是代数闭域F上的有限维线性空间, $T\in L(V)$.

(1) 存在唯一的$D,N\in L(V)$满足$T=D+N$, D可对角化, N幂零, 并且D与N可交换.

(2) D和N都是T的多项式.

证明. 设$p_T=\prod_{i=1}^k(x-c_i)^{r_i}$, 其中$c_i$互不相同并且$r_i\geq 1$. 记$q_i=(x-c_i)^{r_i}$. 我们断言存在$f\in F[x]$满足

$$f\equiv c_i \pmod{q_i},\qquad i=1,\dots,k.$$

这由多项式的中国剩余定理是显然的, 还可以直接证明如下: 由于$h_i:=\prod_{j\neq i}q_j$与q_i 互素, 存在g_i满足$g_ih_i\equiv 1\pmod{q_i}$. 取$f=\sum_{i=1}^k c_ig_ih_i$即可. 设$D=f(T)$, $N=T-D$. 则D,N都是T的多项式, 从而可交换. 我们验证D可对角化, N幂零. 记$V_i=\operatorname{Ker}(q_i(T))$. 由准素分解, $V=\bigoplus_{i=1}^k V_i$并且$p_{T_{V_i}}=q_i$. 由f的性质有$D_{V_i}=f(T_{V_i})=c_i\operatorname{id}_{V_i}$. 这说明$c_i$是$D$的特征值, 并且相应的特征子空间是$V_i$. 因此$D$可对角化. 另一方面, $p_{T_{V_i}}=q_i$推出$N_{V_i}=T_{V_i}-c_i\operatorname{id}_{V_i}$的最小多项式为$x^{r_i}$. 所以$N_{V_i}$幂零. 因此$N$幂零.

设D',N'也满足(1). 由于上面的D,N都是T的多项式, 所以D,N,D',N'两两可交换. 注意到

$$D-D'=N'-N.$$

由于D,D'可对角化并且可交换, 它们可同时对角化, 从而$D-D'$可对角化. 另一方面, 设$N^r=(N')^r=0$, 则

$$(N'-N)^{2r}=\sum_{i=0}^{2r}C_{2r}^i(N')^i(-N)^{2r-i}=0.$$

因此$N'-N$幂零. 而幂零的可对角化变换只能是0. 因此$D-D'=N'-N=0$. □

这里的D和N分别称为T的**可对角化部分**(或**半单部分**)和**幂零部分**. 分解$T=D+N$称为T的**Jordan分解**.

注 7.4. $\sigma(D)=\sigma(T)$, $f_D=f_T$. 如果$p_T=\prod_{i=1}^k p_i^{r_i}$, 其中$p_i$是互不相同的首项系数是1的素多项式并且$r_i\geq 1$, 则$p_D=\prod_{i=1}^k p_i$.

§0.1　半单变换

对应教材7.5节.

定义 0.1. 设V是域F上的非零有限维线性空间, $T \in L(V)$.

(1) 如果V没有非平凡的T-不变子空间, 则称T为**单**(simple)的或**不可约**(irreducible)的.

(2) 如果对任意T-不变子空间$W \subset V$, 存在T-不变子空间$Z \subset V$满足$V = W \oplus Z$, 则称T为**半单**(semisimple)的或**完全可约**(completely reducible)的.

容易看出, T单$\Longleftrightarrow V$中所有非零向量是循环向量, T单$\Longrightarrow T$不可分解.

引理 0.1. 设$T \in L(V)$.

(1) T单$\Longrightarrow T$半单.

(2) 设T半单, $V' \subset V$是T-不变子空间. 则$T_{V'}$半单.

(3) 设$V = \bigoplus_{i=1}^{k} V_i$, 其中每个$V_i$是$T$-不变子空间. 假设每个$T_{V_i}$半单. 则$T$半单.

证明. (1) 显然.

(2) 设$W \subset V'$是不变子空间. 取V的不变子空间Z满足$V = W \oplus Z$. 我们断言$V' = W \oplus (V' \cap Z)$. 显然$W \cap (V' \cap Z) = \{0\}$. 另一方面, 对$v \in V'$, 设$v = w + z$, 其中$w \in W$, $z \in Z$. 则$z = v - w \in V'$. 所以$z \in V' \cap Z$. 因此$v = w + z \in W + (V' \cap Z)$.

(3) 只需证明$k = 2$的情况. 设$W \subset V$是不变子空间. 由于T_{V_i}半单, 可以取不变子空间$Z_1 \subset V_1$和$Z_2 \subset V_2$满足$V_1 = (W \cap V_1) \oplus Z_1$, $V_2 = ((W + V_1) \cap V_2) \oplus Z_2$. 我们验证$V = W \oplus (Z_1 + Z_2)$. 首先, 如果$w = z_1 + z_2 \in W \cap (Z_1 + Z_2)$, 其中$z_i \in Z_i$, 则$z_2 = w - z_1 \in (W + V_1) \cap Z_2 = \{0\}$, 从而$z_2 = 0$. 这说明$w = z_1 \in W \cap Z_1 = \{0\}$, 从而$w = 0$. 因此$W \cap (Z_1 + Z_2) = \{0\}$. 另一方面,

$$\begin{aligned} V &= V_1 + V_2 = V_1 + ((W + V_1) \cap V_2) + Z_2 \subset V_1 + (W + V_1) + Z_2 = V_1 + W + Z_2 \\ &= ((W \cap V_1) + Z_1) + W + Z_2 \subset (W + Z_1) + W + Z_2 = W + (Z_1 + Z_2). \end{aligned}$$

因此$V = W \oplus (Z_1 + Z_2)$. □

命题 0.2. T半单$\Longleftrightarrow$存在T-不变直和分解$V = \bigoplus_{i=1}^{k} V_i$使得每个$T_{V_i}$单.

证明. "$\Longrightarrow$": 由(2). "$\Longleftarrow$": 由(1)和(3). □

命题 0.3. (1) T单$\Longleftrightarrow f_T$素.

(2) T半单$\Longleftrightarrow p_T$无平方因子.

证明. (1) 设T单. 则T循环, 因此$p_T = f_T$. 如果f_T不素, 设$f_T = gh$, $\deg g, \deg h \geq 1$. 则$g(T)h(T) = 0$. 因此$g(T)$或$h(T)$不可逆. 不妨设$g(T)$不可逆. 则$\mathrm{Ker}(g(T))$是非零不变子空间, 因此为全空间. 这说明$g(T) = 0$. 因此$f_T = p_T | g$, 矛盾. 反过来, 设T不单, 即存在非平凡不变子空间$W \subset V$. 则$f_T = f_{T_W} f_{T_{V/W}}$. 因此$f_T$不素.

(2) 设T半单. 则$T = \bigoplus_{i=1}^{k} T_i$, 其中$T_i$单. 于是$p_{T_i}$素. 因此$p_T = \mathrm{lcm}(p_{T_1}, \ldots, p_{T_k})$无平方因子. 反过来, 设$p_T = \prod_{i=1}^{k} p_i$, p_i素并且互不相同. 则在准素循环分解中, 对每个直和项有$p = f$素. 因此每个直和项单. 因此T半单. □

推论 0.4. 设f_T可以分解为$F[x]$中一次式的乘积(例如F是代数闭域).

(1) T单$\Longleftrightarrow \dim V = 1$.

(2) T半单$\Longleftrightarrow T$可对角化.

证明. 对(2), 只需注意到T可对角化$\iff p_T$可以分解为$F[x]$中互不相伴的一次式的乘积. $\square$

下面承认F的代数闭包$\overline{F}$的存在性. 很多重要的问题涉及到$f \in F[x] \setminus F$在$\overline{F}$中是否有重根. f在$\overline{F}$内有重根可以有两个原因造成: 如果f在$F[x]$内有相同的素因子(即有平方因子), 则显然f在$\overline{F}$内有重根. 另一个原因是由域F本身的性质造成的.

定义 0.2. 如果域F上的任意无平方因子的多项式$f \in F[x] \setminus F$在$\overline{F}$内无重根, 则称F是**完全域**.

可以证明:

- F是完全域$\iff F$上的任意素多项式在$\overline{F}$内无重根.
- 代数闭域、特征是0的域和有限域都是完全域.

但$F = \mathbb{F}_p(t)$不是完全域: $f = x^p - t$是$F[x]$中的素多项式, 但如果$a \in \overline{F}$是f的根则$f = (x-a)^p$. 完全域的一个重要性质为: 设F是完全域, $a \in \overline{F}$. 则$a \in F \iff$对所有$\sigma \in \mathrm{Gal}(\overline{F}/F)$有$\sigma(a) = a$.

推论 0.5. 设$A \in F^{n\times n}$.

(1) A在$\overline{F}$上可对角化$\Longrightarrow L_A$半单.

(2) 如果F是完全域, 则A在$\overline{F}$上可对角化$\iff L_A$半单.

证明. 只需注意到: A在$\overline{F}$上可对角化$\iff p_A$在$\overline{F}$中无重根; L_A半单$\iff p_A$无平方因子. $\square$

第八章　内积空间

§8.1　内积

本章我们总是假设$F=\mathbb{R}$或$\mathbb{C}$.

定义 8.1. 设V是F-线性空间. V上的一个**内积**指一个函数$\langle\cdot,\cdot\rangle: V\times V\to F$, $(\alpha,\beta)\mapsto\langle\alpha,\beta\rangle$满足:

(a) $\langle\alpha+\beta,\gamma\rangle=\langle\alpha,\gamma\rangle+\langle\beta,\gamma\rangle$,

(b) $\langle c\alpha,\beta\rangle=c\langle\alpha,\beta\rangle$,

(c) $\langle\beta,\alpha\rangle=\overline{\langle\alpha,\beta\rangle}$, ($\Longrightarrow\langle\alpha,\beta+\gamma\rangle=\langle\alpha,\beta\rangle+\langle\beta,\gamma\rangle$, $\langle\alpha,c\beta\rangle=\bar{c}\langle\alpha,\beta\rangle$)

(d) $\alpha\neq 0\Longrightarrow\langle\alpha,\alpha\rangle>0$.

取定了内积的F-线性空间称为**内积空间**. 有限维实内积空间也叫**Euclid空间**(Euclidean space). 复内积空间也叫**酉空间**.

注 8.1. 当$F=\mathbb{C}$时, 只能要求$\langle\alpha,c\beta\rangle=\bar{c}\langle\alpha,\beta\rangle$($1\frac{1}{2}$-线性). 如果要求双线性, 则与(d)不相容:

$$\langle\alpha,\alpha\rangle>0\Longrightarrow\langle i\alpha,i\alpha\rangle=-\langle\alpha,\alpha\rangle<0.$$

例 8.1. 设$V=F^{n\times 1}$. 对$\alpha=\begin{bmatrix}x_1\\ \vdots\\ x_n\end{bmatrix}$, $\beta=\begin{bmatrix}y_1\\ \vdots\\ y_n\end{bmatrix}$,

$$\langle\alpha,\beta\rangle=\sum_{j=1}^{n}x_j\bar{y}_j=\alpha^t\bar{\beta}=\beta^*\alpha$$

定义了$F^{n\times 1}$上的内积, 称为**标准内积**. 其中对$A\in F^{m\times n}$, $A^*:=\overline{A^t}\in F^{n\times m}$为$A$的转置共轭. 类似地, 可以定义$F^n$上的标准内积.

例 8.2. 设$V=F^{m\times n}$.

$$\langle A,B\rangle=\operatorname{tr}(AB^*)=\operatorname{tr}(B^*A)=\sum_{j,k}A_{jk}\overline{B_{jk}}$$

定义了$F^{m\times n}$上的内积.

定义 8.2. 矩阵$A\in F^{n\times n}$称为**Hermite矩阵**(Hermitian matrix)($F=\mathbb{R}$时称为**对称矩阵**), 如果$A^*=A$. 如果进一步还有

$$X^*AX>0,\qquad\forall X\in F^{n\times 1}\smallsetminus\{0\},$$

则称A是**正定**的.

例 8.3. 设$Q\in\mathrm{GL}_n(F)$. 则$A=Q^*Q$正定Hermite.

引理 8.1. 设$\dim V=n$, $\mathcal{B}=\{\alpha_1,\dots,\alpha_n\}$是$V$的有序基. 对于$\alpha,\beta\in V$, 记$X=[\alpha]_{\mathcal{B}}$, $Y=[\beta]_{\mathcal{B}}$.

(1) 若$A\in F^{n\times n}$正定Hermite, 则

$$\langle\alpha,\beta\rangle=Y^*AX=\sum_{j,k}A_{kj}x_j\bar{y}_k\tag{8.1}$$

定义了V上的内积.

(2) 对任意V上的内积$\langle\cdot,\cdot\rangle$, 存在唯一的正定Hermite矩阵$A\in F^{n\times n}$使得(8.1)成立, 称为该**内积**

在有序基$\mathcal{B}$下的矩阵.

证明. (1) 容易验证(a)-(d)成立.

(2) 唯一性: 如果矩阵$A \in F^{n\times n}$满足(8.1), 取$\alpha = \alpha_j$, $\beta = \alpha_k$得

$$A_{kj} = \langle \alpha_j, \alpha_k \rangle.$$

存在性: 由此式定义矩阵A. 则

$$\langle \alpha, \beta \rangle = \left\langle \sum_j x_j \alpha_j, \sum_k y_k \alpha_k \right\rangle = \sum_{j,k} x_j \bar{y}_k \langle \alpha_j, \alpha_k \rangle = \sum_{j,k} x_j \bar{y}_k A_{kj} = Y^* A X,$$

即(8.1)成立. 由内积定义即得A正定Hermite. □

例 8.4. 设$A = \begin{bmatrix} 1 & -1 \\ -1 & 4 \end{bmatrix} \in \mathbb{R}^{2\times 2}$. 显然$A$对称. 由于对$X = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \neq 0$有

$$X^t A X = (x_1, x_2) \begin{bmatrix} 1 & -1 \\ -1 & 4 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = x_1^2 - 2x_1x_2 + 4x_2^2 = (x_1 - x_2)^2 + 3x_2^2 > 0,$$

所以A正定. 因此,

$$\langle \alpha, \beta \rangle = \beta^t A \alpha = \alpha^t A \beta = (x_1, x_2) \begin{bmatrix} 1 & -1 \\ -1 & 4 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = x_1y_1 - x_1y_2 - x_2y_1 + 4x_2y_2$$

定义了$\mathbb{R}^{2\times 1}$上的内积, 它在标准基下的矩阵是A.

例 8.5. 设$T : V \to W$是单线性映射, $\langle \cdot, \cdot \rangle_0$是$W$上的内积. 则

$$\langle \alpha, \beta \rangle := \langle T\alpha, T\beta \rangle_0$$

定义了V上的内积. 例如:

- 设$\langle \cdot, \cdot \rangle_0$是$F^{n\times 1}$上的标准内积, $Q \in \mathrm{GL}_n(F)$, $T = L_Q$. 则

  $$\langle \alpha, \beta \rangle = \langle Q\alpha, Q\beta \rangle_0 = (Q\beta)^* Q\alpha = \beta^*(Q^*Q)\alpha$$

  也是$F^{n\times 1}$上的内积. 它在标准基下的矩阵为Q^*Q.
- 设$\dim V = n$, $\mathcal{B} = \{\alpha_1, \ldots, \alpha_n\}$是$V$的有序基, $T = \Gamma_{\mathcal{B}} : V \to F^{n\times 1}$是坐标映射, $\langle \cdot, \cdot \rangle_0$是$F^{n\times 1}$上的标准内积. 则$\langle \alpha, \beta \rangle = \langle T\alpha, T\beta \rangle_0$, 即

  $$\left\langle \sum_j x_j \alpha_j, \sum_k y_k \alpha_k \right\rangle = \sum_j x_j \bar{y}_j$$

  是V上的内积. 它在有序基$\mathcal{B}$下的矩阵为I_n. 特别地, $F^{n\times 1}$上的标准内积在标准基下的矩阵是I_n.

设V是内积空间. 我们定义$\alpha \in V$的**长度**为$\|\alpha\| := \langle \alpha, \alpha \rangle^{\frac{1}{2}}$. 实际上, 内积被长度决定:

引理 8.2(极化恒等式). (1) 若$F = \mathbb{R}$, 则

$$\langle \alpha, \beta \rangle = \frac{1}{4}(\|\alpha + \beta\|^2 - \|\alpha - \beta\|^2).$$

(2) 若$F = \mathbb{C}$, 则

$$\langle \alpha, \beta \rangle = \frac{1}{4} \sum_{k=1}^{4} i^k \|\alpha + i^k \beta\|^2.$$

证明. 略. □

§8.2 内积空间

长度有下面的性质:

引理 8.3. (1) $\|c\alpha\| = |c|\|\alpha\|$.

(2) $\alpha \neq 0 \Longrightarrow \|\alpha\| > 0$.

(3) (Cauchy-Schwarz不等式) $|\langle \alpha, \beta \rangle| \leq \|\alpha\|\|\beta\|$, 并且"="成立$\Longleftrightarrow$ α与β线性相关.

(4) (三角不等式) $\|\alpha + \beta\| \leq \|\alpha\| + \|\beta\|$.

证明. (1), (2)显然.

(3) 不妨设$\alpha \neq 0$. 设$\gamma = \beta - \frac{\langle \beta, \alpha \rangle}{\|\alpha\|^2}\alpha$(是$\beta$在$\alpha^\perp$上的正交投影). 则$\langle \gamma, \alpha \rangle = 0$. 从而

$$0 \leq \|\gamma\|^2 = \left\langle \gamma, \beta - \frac{\langle \beta, \alpha \rangle}{\|\alpha\|^2}\alpha \right\rangle = \langle \gamma, \beta \rangle = \left\langle \beta - \frac{\langle \beta, \alpha \rangle}{\|\alpha\|^2}\alpha, \beta \right\rangle = \langle \beta, \beta \rangle - \frac{\langle \beta, \alpha \rangle \langle \alpha, \beta \rangle}{\|\alpha\|^2} = \|\beta\|^2 - \frac{|\langle \alpha, \beta \rangle|^2}{\|\alpha\|^2}.$$

这就推出想要的不等式. "="成立$\Longleftrightarrow \gamma = 0 \Longleftrightarrow \alpha$与$\beta$线性相关.

(4) 由(3)有

$$\begin{aligned}\|\alpha + \beta\|^2 &= \|\alpha\|^2 + \|\beta\|^2 + \langle \alpha, \beta \rangle + \langle \beta, \alpha \rangle = \|\alpha\|^2 + \|\beta\|^2 + 2\mathrm{Re}\langle \alpha, \beta \rangle \\ &\leq \|\alpha\|^2 + \|\beta\|^2 + 2|\langle \alpha, \beta \rangle| \leq \|\alpha\|^2 + \|\beta\|^2 + 2\|\alpha\|\|\beta\| = (\|\alpha\| + \|\beta\|)^2.\end{aligned}$$

□

例 8.6. 对$\mathbb{R}^n$上的标准内积应用Cauchy-Schwarz不等式, 得

$$\left(\sum_{j=1}^n x_j y_j\right)^2 \leq \left(\sum_{j=1}^n x_j^2\right)\left(\sum_{j=1}^n y_j^2\right).$$

定义 8.3. 设V是内积空间.

- 对于$\alpha, \beta \in V$, 如果$\langle \alpha, \beta \rangle = 0$, 则称$\alpha$与$\beta$**垂直**或**正交**, 记为$\alpha \perp \beta$.
- 如果V的子集S中的向量两两正交, 则称S为**正交集**(orthogonal set). 如果进一步还有$\|\alpha\| = 1, \forall \alpha \in S$, 则称$S$为**标准正交集**(orthonormal set).
- 是正交集的基称为**正交基**, 是标准正交集的基称为**标准正交基**.

例 8.7. 注意0与任何向量垂直. $F^{n\times 1}$的标准基在标准内积下是标准正交基.

当$F = \mathbb{R}$时, 对于$\alpha, \beta \in V \setminus \{0\}$, 由Cauchy-Schwarz不等式, 可以定义α与β之间的**角度**为

$$\angle(\alpha, \beta) := \arccos \frac{\langle \alpha, \beta \rangle}{\|\alpha\|\|\beta\|}.$$

则$\alpha \perp \beta \Longleftrightarrow \angle(\alpha, \beta) = \pi$. 注意有

$$\langle \alpha, \beta \rangle = \|\alpha\|\|\beta\| \cos \angle(\alpha, \beta).$$

引理 8.4. 如果$\alpha \perp \beta$, 则

$$\|\alpha + \beta\|^2 = \|\alpha\|^2 + \|\beta\|^2.$$

更一般地, 如果$\{\alpha_1, \ldots, \alpha_m\}$是正交集, 则

$$\left\|\sum_{j=1}^m \alpha_j\right\|^2 = \sum_{j=1}^m \|\alpha_j\|^2.$$

证明. 由引理8.3(4)的证明显然. □

注意如果$\{\alpha_1,\dots,\alpha_n\}$是标准正交基, 则

$$\left\langle \sum_{j=1}^{n} x_j\alpha_j, \sum_{k=1}^{n} y_k\alpha_k \right\rangle = \sum_{j=1}^{n} x_j\bar{y}_j.$$

特别地,

$$\left\| \sum_{j=1}^{n} x_j\alpha_j \right\|^2 = \sum_{j=1}^{n} |x_j|^2.$$

命题 8.5. 不含零向量的正交集总是线性无关的.

证明. 设S是不含零向量的正交集, $\alpha_1,\dots,\alpha_m\in S$. 假设$\sum_{j=1}^{m} c_j\alpha_j=0$. 则

$$0=\left\langle \sum_{j=1}^{m} c_j\alpha_j, \alpha_k \right\rangle = c_k\|\alpha_k\|^2 \Longrightarrow c_k=0.$$

□

因此, 为了说明一个不含零向量的正交集是正交基, 只需说明它生成全空间.

命题 8.6. 设V是有限维内积空间, $\{\beta_1,\dots,\beta_n\}$是$V$的有序基.

(1) 存在唯一的V的正交基$\{\alpha_1,\dots,\alpha_n\}$和对角元为1的上三角矩阵$N$满足

$$[\beta_1,\dots,\beta_n]=[\alpha_1,\dots,\alpha_n]N. \tag{8.2}$$

(2) 存在唯一的V的标准正交基$\{\alpha_1',\dots,\alpha_n'\}$, 对角元为1的上三角矩阵$N$和对角元为正数的对角矩阵$A$满足

$$[\beta_1,\dots,\beta_n]=[\alpha_1',\dots,\alpha_n']AN. \tag{8.3}$$

此时有

$$\operatorname{span}\{\alpha_1,\dots,\alpha_k\}=\operatorname{span}\{\alpha_1',\dots,\alpha_k'\}=\operatorname{span}\{\beta_1,\dots,\beta_k\}, \qquad k=1,\dots,n.$$

证明. (1) 将(8.2)式重写为

$$\begin{aligned}
\beta_1 &= \alpha_1,\\
\beta_2 &= N_{12}\alpha_1+\alpha_2,\\
&\ \vdots\\
\beta_n &= N_{1n}\alpha_1+\cdots+N_{n-1,n}\alpha_{n-1}+\alpha_n.
\end{aligned}$$

注意前m个等式推出

$$\operatorname{span}\{\alpha_1,\dots,\alpha_m\}=\operatorname{span}\{\beta_1,\dots,\beta_m\}.$$

我们归纳证明: 对任意$1\le m\le n$, 存在唯一的$\alpha_1,\dots,\alpha_m\in V\setminus\{0\}$和$N_{jk}\in F(1\le j<k\le m)$使前$m$个等式成立, 并且$\{\alpha_1,\dots,\alpha_m\}$是正交集. $m=1$时无需证明. 假设$1<m\le n$, 并且结论

对$m-1$成立. 则

α_m和$N_{1m},\ldots,N_{m-1,m}$使得前m个等式成立并且$\{\alpha_1,\ldots,\alpha_m\}$是正交集

$\iff \alpha_m$和$N_{1m},\ldots,N_{m-1,m}$使得第m个等式成立(即$\beta_m=\alpha_m+\sum\limits_{j=1}^{m-1}N_{jm}\alpha_j$)并且$\alpha_m\perp\alpha_k(1\le k\le m-1)$

$\iff \alpha_m=\beta_m-\sum\limits_{j=1}^{m-1}N_{jm}\alpha_j$并且$0=\langle\alpha_m,\alpha_k\rangle=\langle\beta_m-\sum\limits_{j=1}^{m-1}N_{jm}\alpha_j,\alpha_k\rangle=\langle\beta_m,\alpha_k\rangle-N_{km}\|\alpha_k\|^2(1\le k\le m-1)$

$\iff N_{km}=\dfrac{\langle\beta_m,\alpha_k\rangle}{\|\alpha_k\|^2}(1\le k\le m-1)$并且$\alpha_m=\beta_m-\sum\limits_{j=1}^{m-1}N_{jm}\alpha_j$.

因此α_m和$N_{1m},\ldots,N_{m-1,m}$被唯一决定, 即结论的唯一性部分对m成立. 另外, 注意到这样决定出的$\alpha_m\neq 0$. (否则, 将有

$$\beta_m=\sum_{j=1}^{m-1}N_{jm}\alpha_j\in\operatorname{span}\{\alpha_1,\ldots,\alpha_{m-1}\}=\operatorname{span}\{\beta_1,\ldots,\beta_{m-1}\},$$

与$\{\beta_1,\ldots,\beta_n\}$是基矛盾. 这里最后一个等号由归纳假设推出.) 因此, 结论的存在性部分对m也成立. 在结论中取$m=n$即得(1).

(2) 存在性: 由(1)的存在性部分, 可取正交基$\{\alpha_1,\ldots,\alpha_n\}$和对角元为1的上三角矩阵$N$满足(8.2). 取$\alpha_j'=\frac{\alpha_j}{\|\alpha_j\|}$, $A=\operatorname{diag}(\|\alpha_1\|,\ldots,\|\alpha_n\|)$. 则(8.3)成立. 注意到

$$\operatorname{span}\{\alpha_1,\ldots,\alpha_k\}=\operatorname{span}\{\alpha_1',\ldots,\alpha_k'\}.$$

唯一性: 设(8.3)成立. 设$A=\operatorname{diag}(a_1,\ldots,a_n)$, $\alpha_j=a_j\alpha_j'$. 则(8.2)成立. 由(1)的唯一性部分, α_j和N被β_j唯一决定. 另一方面, α_j'和a_j被α_j唯一决定, 从而也被β_j唯一决定. □

在(1)的证明中, 利用

$$N_{km}=\frac{\langle\beta_m,\alpha_k\rangle}{\|\alpha_k\|^2},\qquad \alpha_m=\beta_m-\sum_{j=1}^{m-1}N_{jm}\alpha_j$$

可以由β_j归纳构造α_j. 这一过程称为**Gram-Schmidt正交化**. 由任意m个向量$\beta_1,\ldots,\beta_m$出发, 应用Gram-Schmidt正交化过程, 如果在某步有$\alpha_j=0$, 则$\beta_1,\ldots,\beta_m$线性相关. 如果每个$\alpha_j\neq 0$, 则$\beta_1,\ldots,\beta_m$线性无关, 并且对任意$k\in\{1,\ldots,m\}$, $\{\alpha_1,\ldots,\alpha_k\}$是$\operatorname{span}\{\beta_1,\ldots,\beta_k\}$的正交基.

推论 8.7. 设V是有限维内积空间. 则V中任意不含零向量的正交集可以扩充为V的正交基, V中的任意标准正交集可以扩充为V的标准正交基. 特别地, V存在标准正交基. □

命题 8.8. 设$S=\{\alpha_1,\ldots,\alpha_m\}$是不含零向量的正交集, $\beta\in\operatorname{span}S$. 则

$$\beta=\sum_{k=1}^{m}\frac{\langle\beta,\alpha_k\rangle}{\|\alpha_k\|^2}\alpha_k.$$

特别地, 如果$\{\alpha_1,\ldots,\alpha_n\}$是$V$的正交基, 则

$$\beta=\sum_{k=1}^{n}\frac{\langle\beta,\alpha_k\rangle}{\|\alpha_k\|^2}\alpha_k.$$

如果$\{\alpha_1,\ldots,\alpha_n\}$还是$V$的标准正交基, 则

$$\beta=\sum_{k=1}^{n}\langle\beta,\alpha_k\rangle\alpha_k.$$

证明. 设$\beta = \sum_{j=1}^{m} c_j \alpha_j$. 则
$$\langle \beta, \alpha_k \rangle = \left\langle \sum_{j=1}^{m} c_j \alpha_j, \alpha_k \right\rangle = c_k \|\alpha_k\|^2 \Longrightarrow c_k = \frac{\langle \beta, \alpha_k \rangle}{\|\alpha_k\|^2}.$$
□

命题 8.9(Bessel不等式). 设$S = \{\alpha_1, \dots, \alpha_m\}$是不含零向量的正交集, $\beta \in V$. 则
$$\sum_{k=1}^{m} \frac{|\langle \beta, \alpha_k \rangle|^2}{\|\alpha_k\|^2} \le \|\beta\|^2.$$
并且"="成立$\Longleftrightarrow \beta \in \mathrm{span} S$.

证明. 将S扩充为V的正交基$\{\alpha_1, \dots, \alpha_n\}$. 则
$$\beta = \sum_{k=1}^{n} \frac{\langle \beta, \alpha_k \rangle}{\|\alpha_k\|^2} \alpha_k.$$
因此
$$\|\beta\|^2 = \sum_{k=1}^{n} \frac{|\langle \beta, \alpha_k \rangle|^2}{\|\alpha_k\|^2} \ge \sum_{k=1}^{m} \frac{|\langle \beta, \alpha_k \rangle|^2}{\|\alpha_k\|^2}.$$
等号成立$\Longleftrightarrow \langle \beta, \alpha_{m+1} \rangle = \cdots = \langle \beta, \alpha_n \rangle = 0 \Longleftrightarrow \beta = \sum_{k=1}^{m} \frac{\langle \beta, \alpha_k \rangle}{\|\alpha_k\|^2} \alpha_k \Longleftrightarrow \beta \in \mathrm{span} S$. □

这推出: 如果$S = \{\alpha_1, \dots, \alpha_m\}$是标准正交集, $\beta \in V$, 则
$$\sum_{k=1}^{m} |\langle \beta, \alpha_k \rangle|^2 \le \|\beta\|^2,$$
并且"="成立$\Longleftrightarrow \beta \in \mathrm{span} S$.

例 8.8. P282-283例12,13. 略.

定义 8.4. 设S是内积空间V的子集. 定义S在V中的正交补为
$$S^{\perp} := \{\alpha \in V \mid \alpha \perp \beta, \forall \beta \in S\}.$$

容易验证$S^{\perp}$总是子空间, 并且$S^{\perp} = (\mathrm{span} S)^{\perp}$.

命题 8.10. 设V是有限维内积空间, $W \subset V$是子空间. 则$\dim W + \dim W^{\perp} = \dim V$.

证明. 取W的标准正交基$\{\alpha_1, \dots, \alpha_m\}$, 并扩充为$V$的标准正交基$\{\alpha_1, \dots, \alpha_n\}$. 对于$\beta \in V$, 有
$$\beta = \sum_{k=1}^{n} \langle \beta, \alpha_k \rangle \alpha_k.$$
因此
$$\begin{aligned} \beta \in W^{\perp} &\Longleftrightarrow \langle \beta, \alpha_1 \rangle = \cdots = \langle \beta, \alpha_m \rangle = 0 \\ &\Longleftrightarrow \beta = \sum_{k=m+1}^{n} \langle \beta, \alpha_k \rangle \alpha_k \\ &\Longleftrightarrow \beta \in \mathrm{span}\{\alpha_{m+1}, \dots, \alpha_n\}. \end{aligned}$$
从而$W^{\perp} = \mathrm{span}\{\alpha_{m+1}, \dots, \alpha_n\}$, 因此$\{\alpha_{m+1}, \dots, \alpha_n\}$是$W^{\perp}$的基. □

注 8.2. 我们以前曾证明: 设W是有限维线性空间V的子空间, 并定义$W^0 = \{f \in V^* \mid f|_W = 0\}$, 则$\dim W + \dim W^0 = \dim V$. 上面的命题可以视为这一结论的特殊情况. 实际上, 当$F = \mathbb{R}$或$\mathbb{C}$时, V上的一个内积$\langle \cdot, \cdot \rangle$决定了一个映射$\Phi : V \to V^*$, $\Phi(\beta)(\alpha) = \langle \alpha, \beta \rangle$. (书上记$\Phi(\beta) = f_\beta$.) 容易看出, Φ即单又满, 并且是实线性的. (注意当$F = \mathbb{C}$时有$\Phi(c\beta) = \bar{c}\Phi(\beta)$, 因此$\Phi$不是复线性的.) 此外, 如果$W \subset V$是子空间, 则$\Phi(W^{\perp}) = W^0$. 因此$\dim W^{\perp} = \dim W^0$.

推论 8.11. 设V是有限维内积空间, $W \subset V$是子空间. 则$(W^\perp)^\perp = W$.

证明. 显然$W \subset (W^\perp)^\perp$. 而$\dim W = \dim (W^\perp)^\perp$. 所以$(W^\perp)^\perp = W$. □

推论 8.12. 设V是有限维内积空间, $W \subset V$是子空间. 则$V = W \oplus W^\perp$.

证明. 显然$W \cap W^\perp = \{0\}$. 而

$$\dim(W \oplus W^\perp) = \dim W + \dim W^\perp = \dim V.$$

所以$W \oplus W^\perp = V$. □

我们把沿$W^\perp$到W上的投影称为**到W上的正交投影**, 记为P_W. 对于$\beta \in V$, 我们也称$P_W\beta$为β在W上的正交投影. 容易看出, 如果$\{\alpha_1, \ldots, \alpha_m\}$是$W$的正交基, 则

$$P_W\beta = \sum_{j=1}^{m} \frac{\langle \beta, \alpha_j \rangle}{\|\alpha_j\|^2} \alpha_j.$$

如果$\{\alpha_1, \ldots, \alpha_m\}$还是$W$的标准正交基, 则

$$P_W\beta = \sum_{j=1}^{m} \langle \beta, \alpha_j \rangle \alpha_j.$$

注 8.3.
- 我们知道, 如果$V = R \oplus N$, E是沿N到R上的投影, 则$I - E$是沿R到N上的投影. 因此, $P_{W^\perp} = I - P_W$.
- 在Gram-Schmidt正交化过程中, 记$W_k = \operatorname{span}\{\beta_1, \ldots, \beta_k\}$, 则$\alpha_k = P_{W_{k-1}^\perp}\beta_k$.

命题 8.13. $P_W\beta$是β在W中的**最佳逼近**, 即函数$W \to \mathbb{R}$, $\alpha \mapsto \|\beta - \alpha\|$在且只在$P_W\beta$处达到最小值.

证明. 对$\alpha \in W$有$\beta - \alpha = (\beta - P_W\beta) + (P_W\beta - \alpha)$. 由于$\beta - P_W\beta \in W^\perp$, $P_W\beta - \alpha \in W$, 所以$(\beta - P_W\beta) \perp (P_W\beta - \alpha)$. 因此

$$\|\beta - \alpha\|^2 = \|\beta - P_W\beta\|^2 + \|P_W\beta - \alpha\|^2.$$

因此$\|\beta - \alpha\|$达到最小值$\Longleftrightarrow \|P_W\beta - \alpha\|$达到最小值$\Longleftrightarrow \alpha = P_W\beta$. □

§8.3 线性函数和伴随变换

设V是有限维内积空间. 考虑上一节定义的映射$\Phi : V \to V^*$, $\Phi(\beta)(\alpha) = \langle \alpha, \beta \rangle$.

引理 8.14. Φ即单又满. 并且共轭线性, 即$\Phi(c\beta_1 + \beta_2) = \bar{c}\Phi(\beta)_1 + \Phi(\beta_2)$.

证明. 为证即单又满, 只需证明对任意$f \in V^*$, 存在唯一的β使$\Phi(\beta) = f$. 取V的标准正交基$\{\alpha_1, \ldots, \alpha_n\}$. 对于$\beta = \sum_{j=1}^{n} c_j\alpha_j$,

$$\Phi(\beta) = f \Longleftrightarrow \Phi(\beta)(\alpha_k) = f(\alpha_k), \forall k.$$

而

$$\Phi(\beta)(\alpha_k) = \langle \alpha_k, \beta \rangle = \langle \alpha_k, \sum_{j=1}^{n} c_j\alpha_j \rangle = \sum_{j=1}^{n} \bar{c}_j \langle \alpha_k, \alpha_j \rangle = \bar{c}_k.$$

所以

$$\Phi(\beta) = f \Longleftrightarrow c_k = \overline{f(\alpha_k)}, \forall k \Longleftrightarrow \beta = \sum_{j=1}^{n} \overline{f(\alpha_j)}\alpha_j.$$

因此满足$\Phi(\beta) = f$的β存在唯一.

共轭线性: 对任意α有

$\Phi(c\beta_1+\beta_2)(\alpha)=\langle\alpha,c\beta_1+\beta_2\rangle=\bar{c}\langle\alpha,\beta_1\rangle+\langle\alpha,\beta_2\rangle=\bar{c}\Phi(\beta_1)(\alpha)+\Phi(\beta_2)(\alpha)=(\bar{c}\Phi(\beta)_1+\Phi(\beta_2))(\alpha)$.
因此$\Phi(c\beta_1+\beta_2)=\bar{c}\Phi(\beta)_1+\Phi(\beta_2)$. □

注 8.4. • 在共轭线性的基础上, 为证即单又满, 只需证明$\mathrm{Ker}(\Phi)=0$(注意到$\dim_{\mathbb{R}} V=\dim_{\mathbb{R}} V^*$). 而这是很容易的.

• $\{\Phi(\beta)\}^0=\mathrm{Ker}(\Phi(\beta))=\beta^{\perp}$:

$$\mathrm{Ker}(\Phi(\beta))=\{\alpha\in V\mid\Phi(\beta)(\alpha)=0\}=\{\alpha\in V\mid\langle\alpha,\beta\rangle=0\}=\beta^{\perp}.$$

这推出对任意$S\subset V$有

$$\Phi(S)^0=\bigcap_{\beta\in S}\mathrm{Ker}(\Phi(\beta))=\bigcap_{\beta\in S}\beta^{\perp}=S^{\perp}.$$

还有

$$\Phi(S^{\perp})=S^0.$$

实际上就是同一个空间有两个同构的对偶.

引理 8.15. 设V是有限维内积空间. 则对任意$T\in L(V)$, 存在唯一的$T^*\in L(V)$满足

$$\langle T\alpha,\beta\rangle=\langle\alpha,T^*\beta\rangle,\qquad\forall\alpha,\beta\in V.$$

如果$\mathcal{B}=\{\alpha_1,\ldots,\alpha_n\}$是有序标准正交基, 则$[T^*]_{\mathcal{B}}=[T]_{\mathcal{B}}^*$.

证明.

$$\langle T\alpha,\beta\rangle=\langle\alpha,T^*\beta\rangle\iff\Phi(\beta)(T\alpha)=\Phi(T^*\beta)(\alpha)\iff\Phi(\beta)\circ T=\Phi(T^*\beta)\iff T^t(\Phi(\beta))=\Phi(T^*\beta)$$
$$\iff T^t\circ\Phi=\Phi\circ T^*\iff T^*=\Phi^{-1}\circ T^t\circ\Phi.$$

容易验证$\Phi^{-1}\circ T^t\circ\Phi$线性. 注意$T^*$是使得图表

$$\begin{tikzcd} V^* \arrow[r, "T^t"] & V^* \\ V \arrow[u, "\Phi", "\cong"'] \arrow[r, "T^*"'] & V \arrow[u, "\cong", "\Phi"'] \end{tikzcd}$$

可交换的唯一映射.

如果$\mathcal{B}=\{\alpha_1,\ldots,\alpha_n\}$是有序标准正交基, 则$\Phi(\mathcal{B}):=\{\Phi\alpha_1,\ldots,\Phi\alpha_n\}$是对偶基. 所以$[T^t]_{\Phi(\mathcal{B})}=[T]_{\mathcal{B}}^t$. 因此, 只需证明$[T^*]_{\mathcal{B}}=\overline{[T^t]_{\Phi(\mathcal{B})}}$. 当$F=\mathbb{R}$时, 由于$\Phi$是线性的并且图表可交换, 这是显然的. 当$F=\mathbb{C}$时, 证明如下. 设$[T^*]_{\mathcal{B}}=A$. 则

$$T^*\alpha_k=\sum_{j=1}^{n}A_{jk}\alpha_j.$$

两边取Φ, 得

$$T^t\Phi(\alpha_k)=\Phi(T^*\alpha_k)=\sum_{j=1}^{n}\overline{A_{jk}}\Phi(\alpha_j),$$

即$[T^t]_{\Phi(\mathcal{B})}=\overline{A}$. □

T^*称为T的(关于内积的)**伴随变换**.

注: 可以对一般的$T\in L(V,W)$定义T^*.

例 8.9. 在$F^{n\times 1}$上的标准内积下, $L_A^*=L_{A^*}$. 只需验证$\langle L_AX,Y\rangle=\langle X,L_{A^*}Y\rangle$, 即$\langle AX,Y\rangle=$

$\langle X, A^*Y\rangle$. 而$\langle X, Y\rangle = Y^*X$. 所以后者等价于$Y^*(AX) = (A^*Y)^*X$. 这是显然的.

伴随变换有下面的性质:

引理 8.16. (1) $(T+U)^* = T^* + U^*$.

(2) $(cT)^* = \bar{c}T^*$.

(3) $(TU)^* = U^*T^*$.

(4) $(T^*)^* = T$.

证明. 可以利用$T^* = \Phi^{-1}\circ T^t\circ\Phi$. 也可以直接验证. 例如, 为了直接验证(4), 只需验证$\langle T^*\alpha, \beta\rangle = \langle \alpha, T\beta\rangle$. 这是显然的. □

这些性质说明: $F = \mathbb{C}$时, 映射$T \mapsto T^*$是$\mathbb{C}$-代数的$L(V)$的共轭线性反自同构.

如果$T^* = T$, 即

$$\langle T\alpha, \beta\rangle = \langle \alpha, T\beta\rangle, \qquad \forall \alpha, \beta \in V,$$

则称T是**自伴**的.

引理 8.17. 设$\mathcal{B} = \{\alpha_1, \ldots, \alpha_n\}$是内积空间$V$的标准正交基. 则$T \in L(V)$自伴$\Longleftrightarrow [T]_{\mathcal{B}}$ Hermite.

证明. 由于$[T^*]_{\mathcal{B}} = [T]_{\mathcal{B}}^*$, 所以$T^* = T \Longleftrightarrow [T^*]_{\mathcal{B}} = [T]_{\mathcal{B}} \Longleftrightarrow [T]_{\mathcal{B}}^* = [T]_{\mathcal{B}} \Longleftrightarrow [T]_{\mathcal{B}}$ Hermite. □

例 8.10. 设W是有限维内积空间V的子空间. 则P_W自伴. 事实上, 取W的有序标准正交集$\{\alpha_1, \ldots, \alpha_m\}$, 并扩充为$V$的有序标准正交基$\{\alpha_1, \ldots, \alpha_n\}$. 则$[P]_{\mathcal{B}} = \mathrm{diag}(I_m, 0)$是Hermite矩阵.

注意对任意$T \in L(V)$, 存在唯一的自伴变换$T_1, T_2 \in L(V)$满足$T = T_1 + iT_2$. 事实上, $T_1 = \frac{1}{2}(T+T^*)$, $T_2 = \frac{1}{2i}(T - T^*)$.

§8.4 正交变换和酉变换

定义 8.5. 设V, W是内积空间. 如果$T \in L(V, W)$满足$\langle T\alpha, T\beta\rangle = \langle \alpha, \beta\rangle$, 则称$T$**保内积**. 如果$T$保内积并且是线性空间的同构, 则称$T$是**内积空间的同构**. 此时$T^{-1}$也是内积空间的同构. 如果在$V$与$W$间存在内积空间的同构, 则称$V$与$W$作为内积空间**同构**.

命题 8.18. T保内积$\Longleftrightarrow T$保长度, 即$\|T\alpha\| = \|\alpha\|$, $\forall \alpha \in V$. 特别地, 保内积$\Longrightarrow$单.

证明. "$\Longrightarrow$" 显然.

"$\Longleftarrow$" 只需应用极化恒等式.

如果T保长度, 则$\mathrm{Ker}(T) = \{0\}$, 从而T单. □

因此, 保内积变换也叫**等距变换**(isometry).

命题 8.19. 设V, W是有限维内积空间, $\dim V = \dim W$, $T \in L(V, W)$. TFAE:

(1) T保内积.

(2) T是内积空间的同构.

(3) T把每个V的标准正交基映为W的标准正交基.

(4) T把某个V的标准正交基映为W的标准正交基.

证明. "(1)$\Longrightarrow$(2)" T保内积$\Longrightarrow T$单. 而$\dim V = \dim W$, 所以T满.

"(2)$\Longrightarrow$(3)" 设$\{\alpha_1, \ldots, \alpha_n\}$是$V$的标准正交基. 则$\{T\alpha_1, \ldots, T\alpha_n\}$是$W$的基, 并且$\langle T\alpha_j, T\alpha_k\rangle =$

$\langle\alpha_j,\alpha_k\rangle=\delta_{jk}$. 因此$\{T\alpha_1,\ldots,T\alpha_n\}$是标准正交基.

"(3)$\Longrightarrow$(4)" 显然.

"(4)$\Longrightarrow$(1)" 设T把V的标准正交基$\{\alpha_1,\ldots,\alpha_n\}$映为$W$的标准正交基$\{T\alpha_1,\ldots,T\alpha_n\}$. 则对$\alpha,\beta\in V$, 设$\alpha=\sum_{j=1}^n x_j\alpha_j$, $\beta=\sum_{k=1}^n y_k\alpha_k$. 则

$$\langle T\alpha,T\beta\rangle=\langle\sum_{j=1}^n x_jT\alpha_j,\sum_{k=1}^n y_kT\alpha_k\rangle=\sum_{j=1}^n x_j\bar{y}_j=\langle\sum_{j=1}^n x_j\alpha_j,\sum_{k=1}^n y_k\alpha_k\rangle=\langle\alpha,\beta\rangle.$$

□

推论 8.20. 设V,W是有限维内积空间. 则V与W作为内积空间同构$\Longleftrightarrow\dim V=\dim W$.

证明. "$\Longrightarrow$" 显然.

"$\Longleftarrow$" 取V的标准正交基$\{\alpha_1,\ldots,\alpha_n\}$和$W$的标准正交基$\{\beta_1,\ldots,\beta_n\}$. 取$T\in L(V,W)$满足$T\alpha_j=\beta_j$. 则$T$是内积空间的同构. □

例 8.11. 设V是有限维内积空间, $\mathcal{B}=\{\alpha_1,\ldots,\alpha_n\}$是有序标准正交基. 则坐标映射$\Gamma_{\mathcal{B}}:V\to F^{n\times 1}$, $\alpha\mapsto[\alpha]_{\mathcal{B}}$是内积空间的同构(关于$F^{n\times 1}$上的标准内积).

定义 8.6.
- 实内积空间V作为内积空间的自同构称为**正交变换**. 记V上所有正交变换在映射复合下构成的群为$O(V)$, 称为V上的**正交群**.
- 复内积空间V作为内积空间的自同构称为**酉变换**. 记V上所有酉变换在映射复合下构成的群为$U(V)$, 称为V上的**酉群**.

命题 8.21. $\dim V<\infty$. 则对于$T\in L(V)$, TFAE:

(1) T是内积空间的自同构.

(2) T保内积.

(3) $T^*T=I$.

证明. "(1)$\Longleftrightarrow$(2)" 由上一命题显然.

"(2)$\Longleftrightarrow$(3)" 注意到总有$\langle T\alpha,T\beta\rangle=\langle\alpha,T^*T\beta\rangle$. 因此

$$(2)\Longleftrightarrow\langle\alpha,T^*T\beta\rangle=\langle\alpha,\beta\rangle\Longleftrightarrow T^*T=I.$$

□

定义 8.7. 设$A\in\mathbb{C}^{n\times n}$.

- 如果$A^tA=I$, 则称A称为正交矩阵. 记所有n阶实正交矩阵在矩阵乘法下构成的群为$O(n)$, 称为n**阶正交群**. 记所有n阶复正交矩阵在矩阵乘法下构成的群为$O(n,\mathbb{C})$, 称为n**阶复正交群**.
- 如果$A^*A=I$, 则称A称为酉矩阵. 记所有n阶酉矩阵在矩阵乘法下构成的群为$U(n)$, 称为n**阶酉群**.

显然, 如果A实, 则A正交$\Longleftrightarrow A$酉.

命题 8.22. 设V是域F上的有限维内积空间, $\mathcal{B}=\{\alpha_1,\ldots,\alpha_n\}$是有序标准正交基, $T\in L(V)$.

(1) 如果$F=\mathbb{R}$, 则T正交$\Longleftrightarrow[T]_{\mathcal{B}}$正交.

(2) 如果$F=\mathbb{C}$, 则T酉$\Longleftrightarrow[T]_{\mathcal{B}}$酉.

证明. 只需证明$T^*T=I\Longleftrightarrow[T]_{\mathcal{B}}^*[T]_{\mathcal{B}}=I$. 而总有$[T^*T]_{\mathcal{B}}=[T]_{\mathcal{B}}^*[T]_{\mathcal{B}}$. 因此命题成立. □

命题 8.23. 设$A \in \mathbb{R}^{n\times n}$. TFAE:

(1) A正交.

(2) A^t正交.

(3) A的行向量构成$\mathbb{R}^n$的标准正交基(关于标准内积).

(4) A的列向量构成$\mathbb{R}^{n\times 1}$的标准正交基(关于标准内积).

证明. "(1)$\Longleftrightarrow$(4)". 由于$\mathbb{R}^{n\times 1}$的标准基关于标准内积是标准正交基, 并且L_A在标准基下的矩阵为A, 所以A正交$\Longleftrightarrow$ L_A正交. 而L_A把标准基映为A的列向量集合. 所以L_A正交$\Longleftrightarrow$(4).

"(2)$\Longleftrightarrow$(3)". 对A^t应用"(1)$\Longleftrightarrow$(4)".

"(1)$\Longleftrightarrow$(2)". 由定义, A正交$\Longleftrightarrow$ $A^tA = I$, A^t正交$\Longleftrightarrow$ $AA^t = I$. 因此两者都等价于: A可逆并且$A^{-1} = A^t$.

注 8.5. 直接证明"(3)$\Longleftrightarrow$(4)"并不容易!

类似地, 可以证明:

命题 8.24. 设$A \in \mathbb{C}^{n\times n}$. TFAE:

(1) A酉.

(2) A^*酉.

(3) A的行向量构成$\mathbb{C}^n$的标准正交基(关于标准内积).

(4) A的列向量构成$\mathbb{C}^{n\times 1}$的标准正交基(关于标准内积).

注 8.6. 还有A酉$\Longleftrightarrow$ A^t酉$\Longleftrightarrow$ $\overline{A}$酉: $A^*A = I \Longleftrightarrow \overline{A^*A} = \overline{A^*}\ \overline{A} = (\overline{A})^*\overline{A} = I \Longleftrightarrow \overline{A}$酉$\Longleftrightarrow$ $(\overline{A})^* = A^t$酉.

容易看出, A正交$\Longrightarrow \det(A) = \pm 1$, A酉$\Longrightarrow |\det(A)| = 1$. 记

$$SO(n) = \{A \in O(n) \mid \det(A) = 1\},$$

$$SU(n) = \{A \in U(n) \mid \det(A) = 1\},$$

分别称为n阶**特殊正交群**和n阶**特殊酉群**. 对任意$A_0 \in O(n) \setminus SO(n)$, 有$O(n) = SO(n) \sqcup A_0SO(n)$. 例如, 取$A_0 = \mathrm{diag}(1, \ldots, 1, -1)$. 还有

$$U(n) = \{zA \mid A \in SU(n), z \in \mathbb{C}, |z| = 1\}.$$

例 8.12.

$$O(1) = \{1, -1\}.$$

$$U(1) = \{z \in \mathbb{C} \mid |z| = 1\} = \{e^{i\theta} \mid \theta \in \mathbb{R}\}.$$

$$SO(1) = SO(1) = \{1\}.$$

$$SO(2) = \left\{ \begin{bmatrix} x & -y \\ y & x \end{bmatrix} \mid x, y \in \mathbb{R}, x^2 + y^2 = 1 \right\} = \left\{ \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \mid \theta \in \mathbb{R} \right\}.$$

$$SU(2) = \left\{ \begin{bmatrix} z & -\bar{w} \\ w & \bar{z} \end{bmatrix} \mid z, w \in \mathbb{C}, |z|^2 + |w|^2 = 1 \right\}.$$

定理 8.25(QR分解或Iwasawa分解). 设$F = \mathbb{R}$或$\mathbb{C}$. 则对任意$A \in GL_n(F)$, 存在唯一的$A_k \in O(n)$(若$F = \mathbb{R}$)或$U(n)$(若$F = \mathbb{C}$), 对角元为正数的对角矩阵A_a, 对角元为1的上三角矩阵A_n满足$A = A_kA_aA_n$.

证明. 设$A=[\beta_1,\dots,\beta_n]$. 由命题8.6(2), 存在唯一的$F^{n\times 1}$的有序标准正交基$\{\alpha_1,\dots,\alpha_n\}$, 对角元为正数的对角矩阵$A_a$, 对角元为1的上三角矩阵$A_n$满足$A=[\alpha_1,\dots,\alpha_n]A_aA_n$. 只需再注意到$\{\alpha_1,\dots,\alpha_n\}$是标准正交基$\Longleftrightarrow A_k:=[\alpha_1,\dots,\alpha_n]\in O(n)$或$U(n)$. □

现在考虑内积空间上的坐标变换. 设V是有限维内积空间, $\mathcal{B}=\{\alpha_1,\dots,\alpha_n\}$和$\mathcal{B}'=\{\alpha'_1,\dots,\alpha'_n\}$是有序标准正交基. 则坐标映射$\Gamma_{\mathcal{B}}:V\to F^{n\times 1}$, $\alpha\mapsto[\alpha]_{\mathcal{B}}$和$\Gamma_{\mathcal{B}'}:V\to F^{n\times 1}$, $\alpha\mapsto[\alpha]_{\mathcal{B}'}$是内积空间的同构(关于$F^{n\times 1}$上的标准内积). 我们知道, 存在唯一的$P\in GL_n(F)$满足$\Gamma_{\mathcal{B}}=L_P\circ\Gamma_{\mathcal{B}'}$, 即$[\alpha]_{\mathcal{B}}=P[\alpha]_{\mathcal{B}'}$. 则$L_P$是内积空间$F^{n\times 1}$的同构, 即$P$(作为$L_P$在标准基下的矩阵)是正交矩阵(当$F=\mathbb{R}$时)或酉矩阵(当$F=\mathbb{C}$时). 我们还知道, 对任意$T\in L(V)$, 有

$$[T]_{\mathcal{B}'}=P^{-1}[T]_{\mathcal{B}}P=P^*[T]_{\mathcal{B}}P.$$

定义 8.8.
- 称$A,B\in\mathbb{R}^{n\times n}$**正交相似**, 如果存在$P\in O(n)$满足$B=P^{-1}AP$.
- 称$A,B\in\mathbb{C}^{n\times n}$**酉相似**, 如果存在$P\in U(n)$满足$B=P^{-1}AP$.

因此, 在$F=\mathbb{C}$的情况, 对于$T\in L(V)$, 取有序标准正交基$\mathcal{B}$得到的矩阵$[T]_{\mathcal{B}}$是酉相似的, 而取所有有序标准正交基$\mathcal{B}$得到的矩阵$[T]_{\mathcal{B}}$构成一个酉相似等价类. 对于$F=\mathbb{R}$时情况类似. 因此, 我们可以提下面两个等价的问题:

- 对有限维内积空间V和$T\in L(V)$, 寻找V的有序标准正交基$\mathcal{B}$, 使得矩阵$[T]_{\mathcal{B}}$尽量简单.
- 给定矩阵$A\in\mathbb{C}^{n\times n}$(或$\mathbb{R}^{n\times n}$), 寻找$P\in U(n)$(或$O(n)$), 使得矩阵$P^{-1}AP$尽量简单.

最后, 为了做习题, 我们补充下面的概念:

定义 8.9. 设V是有限维内积空间. 称$T\in L(V)$为**正定的**(positive definite)或**正的**(positive), 如果T自伴, 并且$\langle T\alpha,\alpha\rangle>0$, $\forall\alpha\in V\smallsetminus\{0\}$.

容易看出:

- T正$\Longleftrightarrow$函数$(\alpha,\beta)\mapsto\langle T\alpha,\beta\rangle$是内积.
- 如果$\mathcal{B}$是V的有序标准正交基, 则T正$\Longleftrightarrow[T]_{\mathcal{B}}$正定Hermite.

§8.5 正规变换

本节回答上一节最后提出的问题. 我们先考虑在正交相似或酉相似意义下的可对角化问题.

定义 8.10.
- 有限维内积空间V上的线性变换T如果与T^*可交换, 则称T为**正规变换**(normal).
- 矩阵$A\in\mathbb{C}^{n\times n}$如果与$A^*$可交换, 则称$A$是**正规**的.

注 8.7. 若T正规, 设$T_1=\frac{T+T^*}{2}$, $T_2=\frac{T-T^*}{2i}$. 则T_1, T_2自伴可交换, 并且$T=T_1+iT_2$. 因此T_1,T_2可酉对角化$\Longrightarrow T$ 可酉对角化.

自伴变换、“反自伴”变换(即$T^*+T=0$)和酉变换(正交变换)都是正规的. 容易看出, 如果$\mathcal{B}$是V的有序标准正交基, 则T正规$\Longleftrightarrow[T]_{\mathcal{B}}$正规. 我们证明:

定理 8.26. 设V是F上的有限维内积空间, $T\in L(V)$.

(1) 若$F=\mathbb{R}$, 则存在V的有序标准正交基$\mathcal{B}$使$[T]_{\mathcal{B}}$对角$\Longleftrightarrow T$自伴.

(2) 若$F=\mathbb{C}$, 则存在V的有序标准正交基$\mathcal{B}$使$[T]_{\mathcal{B}}$对角$\Longleftrightarrow T$正规.

下面的推论是显然的.

推论 8.27. (1) $A\in\mathbb{R}^{n\times n}$正交相似于对角矩阵$\Longleftrightarrow A$对称.

(2) $A\in\mathbb{C}^{n\times n}$酉相似于对角矩阵$\Longleftrightarrow$ A正规.

注 8.8. 实正规矩阵(例如正交矩阵)一般不相似于实对角矩阵. 可以证明: 任意实正规矩阵正交相似于分块对角矩阵, 其中每块的阶数至多为2.

定理8.26的"$\Longrightarrow$"部分的证明. 我们知道, $[T]_{\mathcal{B}}$对角$\Longleftrightarrow$ $\mathcal{B}$中的向量都是T的特征向量. 因此, 条件推出存在有序标准正交基$\mathcal{B}=\{\alpha_1,\dots,\alpha_n\}$使得每个$\alpha_j$都是$T$的特征向量. 设$T\alpha_j=c_j\alpha_j$. 则$[T]_{\mathcal{B}}=\operatorname{diag}(c_1,\dots,c_n)$, 从而$[T^*]_{\mathcal{B}}=[T]_{\mathcal{B}}^*=\operatorname{diag}(\bar{c}_1,\dots,\bar{c}_n)$. 如果$F=\mathbb{R}$, 则每个$c_j\in\mathbb{R}$, 于是$[T^*]_{\mathcal{B}}=[T]_{\mathcal{B}}$, 从而$T^*=T$, 即$T$自伴. 如果$F=\mathbb{C}$, 则$[T]_{\mathcal{B}}$与$[T^*]_{\mathcal{B}}$可交换. 这推出$T$与$T^*$可交换, 即$T$正规. □

为了证明定理8.26的"$\Longleftarrow$"部分, 先证明几个引理.

引理 8.28. 设V是有限维内积空间, $T\in L(V)$自伴. 则$f_T\in\mathbb{R}[x]$, 并且在$\mathbb{R}[x]$中完全分解为一次式的乘积. 特别地, $\sigma(T)\subset\mathbb{R}$.

证明. 设$f_T=\prod_{j=1}^k(x-c_j)^{r_j}$, $c_j\in\mathbb{C}$. 只需证明$c_j\in\mathbb{R}$. 取V的有序标准正交基$\mathcal{B}$. 则$A:=[T]_{\mathcal{B}}$ Hermite. 取$X\in\mathbb{C}^{n\times 1}\setminus\{0\}$满足$AX=c_jX$. 则

$$c_jX^*X=X^*(AX)=(X^*A)X=(AX)^*X=(c_jX)^*X=\bar{c}_jX^*X.$$

由于$X\neq 0$, 所以$X^*X>0$. 因此$c_j=\bar{c}_j$, 即$c_j\in\mathbb{R}$. □

引理 8.29. 设V是有限维内积空间, $T\in L(V)$. 如果$W\subset V$是T-不变子空间, 则$W^\perp$是T^*-不变子空间.

证明. 设$\alpha\in W^\perp$. 则对任意$\beta\in W$, 有$\langle T^*\alpha,\beta\rangle=\langle\alpha,T\beta\rangle=0$. 因此$T^*\alpha\in W^\perp$. 这就证明了$W^\perp$是$T^*$-不变子空间. □

引理 8.30. 设V是有限维内积空间, $T\in L(V)$正规.

(1) 如果$W\subset V$是T-不变子空间, 则$W^\perp$是T-不变子空间, W是T^*-不变子空间.

(2) T的特征子空间两两正交.

证明. (1) 分别取W的标准正交基$\mathcal{B}_1$和$W^\perp$的标准正交基$\mathcal{B}_2$, 则$\mathcal{B}=(\mathcal{B}_1,\mathcal{B}_2)$是$V$的标准正交基, 从而$A:=[T]_{\mathcal{B}}$正规. 由于$W$是$T$-不变的, 所以$A$形如

$$A=\begin{bmatrix}B & C\\ 0 & D\end{bmatrix}.$$

计算得

$$AA^*=\begin{bmatrix}BB^*+CC^* & *\\ * & DD^*\end{bmatrix},\qquad A^*A=\begin{bmatrix}B^*B & *\\ * & D^*D+C^*C\end{bmatrix}.$$

A正规推出

$$BB^*+CC^*=B^*B,\qquad DD^*=D^*D+C^*C.$$

对第一个式子两边取矩阵的迹, 得$\operatorname{tr}(CC^*)=0$, 从而$C=0$. 因此$W^\perp$是$T$-不变子空间. 再由引理8.29, 即知$W=(W^\perp)^\perp$是$T^*$-不变子空间.

(2) 设$c_1,c_2\in\sigma(T)$, $c_1\neq c_2$. 需要证明$V_{c_1}\perp V_{c_2}$. 设$\alpha\in V_{c_1}$, $\beta\in V_{c_2}$. 由(1), $F\beta$是T^*-不变子空间, 即$T^*\beta\in F\beta$, 并且

$$\langle T^*\beta,\beta\rangle=\langle\beta,T\beta\rangle=\bar{c}_2\langle\beta,\beta\rangle.$$

因此$T^*\beta=\bar{c}_2\beta$. 这推出

$$c_1\langle\alpha,\beta\rangle=\langle c_1\alpha,\beta\rangle=\langle T\alpha,\beta\rangle=\langle\alpha,T^*\beta\rangle=\langle\alpha,\bar{c}_2\beta\rangle=c_2\langle\alpha,\beta\rangle.$$

由于$c_1\neq c_2$, 所以$\langle\alpha,\beta\rangle=0$. □

注 8.9. 对于自伴变换, 引理8.30(1)是引理8.29的直接推论, 引理8.30(2)的证明也更加简单(结合引理8.28).

定理8.26的"⟸"部分的证明. 由引理8.30(1), T半单. 再由推论7.43和引理8.28, 可知T可对角化. 设$\sigma(T)=\{c_1,\ldots,c_k\}$. 则有直和分解$V=\bigoplus_{j=1}^k V_{c_j}$. 由引理8.30(2), 该直和分解中的直和项两两正交. 取V_{c_j}的有序标准正交基$\mathcal{B}_j$. 则$\mathcal{B}=(\mathcal{B}_1,\ldots,\mathcal{B}_k)$是$V$的有序标准正交基. 由于$\mathcal{B}$中的向量都是$T$的特征向量, 所以$[T]_{\mathcal{B}}$对角. □

注 8.10. 为了得到直和分解$V=\bigoplus_{j=1}^k V_{c_j}$, 也可以不用"半单⟹可对角化": 假设$W:=\bigoplus_{j=1}^k V_{c_j}\neq V$. 为了得到矛盾, 只需说明$W^\perp$中存在$T$的特征向量. 首先, 由引理8.30(1), $W^\perp$是非零T-不变子空间. 当$F=\mathbb{C}$时, $W^\perp$中显然有T的特征向量. 当$F=\mathbb{R}$时并且T自伴时, 容易看出$T_{W^\perp}$也自伴, 于是由引理8.28, $W^\perp$中也存在T的特征向量. □

推论 8.31. 设V是F上的有限维内积空间, $T\in L(V)$. 假设或者$F=\mathbb{R}$并且T自伴, 或者$F=\mathbb{C}$并且T正规. 则有正交直和分解$V=\bigoplus_{c\in\sigma(T)}V_c$. □

可以利用讲义中命题7.17来求正交矩阵或酉矩阵P使得$P^{-1}AP$对角. 此时需要找每个特征子空间V_c的标准正交基.

在复内积空间的情况, 自伴变换、反自伴变换和酉变换是三种重要的正规变换. 它们可以利用谱集来刻画.

推论 8.32. 设V是有限维复内积空间, $T\in L(V)$正规. 则:

(1) T自伴$\Longleftrightarrow\sigma(T)\subset\mathbb{R}$.

(2) T反自伴$\Longleftrightarrow\sigma(T)\subset i\mathbb{R}$.

(3) T酉$\Longleftrightarrow\sigma(T)\subset\{z\in\mathbb{C}\mid|z|=1\}$.

证明. 取V的有序标准正交基$\mathcal{B}$使得$A:=[T]_{\mathcal{B}}$对角. 则$\sigma(T)$是A的对角元的集合.

(1) T自伴$\Longleftrightarrow A$ Hermite $\Longleftrightarrow A$实对角$\Longleftrightarrow\sigma(T)\subset\mathbb{R}$.

(2) 可以类似(1)来证明, 也可以注意T自伴$\Longleftrightarrow iT$反自伴并利用(1).

(3) T酉$\Longleftrightarrow A$酉$\Longleftrightarrow A$的对角元的模都是1 $\Longleftrightarrow\sigma(T)\subset\{z\in\mathbb{C}\mid|z|=1\}$. □

推论 8.33. 设$A\in\mathbb{C}^{n\times n}$.

(1) A Hermite $\Longleftrightarrow A$酉相似于实对角矩阵.

(2) A反Hermite $\Longleftrightarrow A$酉相似于纯虚对角矩阵.

(3) A酉$\Longleftrightarrow A$酉相似于对角元的模是1的对角矩阵.

接下来讨论一般矩阵在正交相似或酉相似下的标准形.

定理 8.34(Schur三角化定理). 设V是有限维复内积空间, $T\in L(V)$. 则存在V的有序标准正交基$\mathcal{B}$使得$[T]_{\mathcal{B}}$上三角.

证明. 由讲义的定理6.24, 存在T-不变的全旗$\{0\}=W_0\subsetneq\cdots\subsetneq W_n=V$. 取$V$的有序标准正交基$\mathcal{B}=\{\alpha_1,\ldots,\alpha_n\}$使得$W_j=\operatorname{span}\{\alpha_1,\ldots,\alpha_j\}$. 则$[T]_{\mathcal{B}}$上三角(见定理6.24的证明中

的“(4)$\Longrightarrow$(1)”). $\square$

推论 8.35. 任意$A\in\mathbb{C}^{n\times n}$酉相似于上三角矩阵.

注 8.11. 对于实矩阵, 与Schur三角化定理直接类似的结果不成立. 可以证明: 任意$A\in\mathbb{R}^{n\times n}$正交相似于分块上三角矩阵, 其中每块的行数和列数至多为2.

定理8.26(2)的“$\Longleftarrow$”部分还可以从Schur三角化定理的角度重新理解(证明).

引理 8.36. 设$A\in\mathbb{C}^{n\times n}$是上三角矩阵. 则$A$正规$\Longleftrightarrow A$对角.

证明. “$\Longleftarrow$”: 若A对角, 则$A^*=\bar{A}$对角, 从而与A可交换. 因此A正规.

“$\Longrightarrow$”: A正规$\Longrightarrow L_A$正规(关于$\mathbb{C}^{n\times 1}$上的标准内积). 设$\{\epsilon_1,\ldots,\epsilon_n\}$是标准基. 由引理8.30(1),

$$\begin{aligned} A\text{上三角} &\Longrightarrow \mathrm{span}\{\epsilon_1,\ldots,\epsilon_k\}\text{是}L_A\text{-不变的}, 1\le k\le n-1\\ &\Longrightarrow \mathrm{span}\{\epsilon_{k+1},\ldots,\epsilon_n\}=\mathrm{span}\{\epsilon_1,\ldots,\epsilon_k\}^{\perp}\text{是}L_A\text{-不变的}, 1\le k\le n-1\\ &\Longrightarrow A\text{下三角}. \end{aligned}$$

因此A对角. $\square$

定理8.26(2)的“$\Longleftarrow$”部分的另一证明. 假设T正规. 取V的有序标准正交基$\mathcal{B}$使得$[T]_{\mathcal{B}}$上三角. 则$[T]_{\mathcal{B}}$正规, 从而对角. $\square$

第九章　内积空间上的线性变换

§9.1　线性空间上的形式

设$F=\mathbb{R}$或$\mathbb{C}$, V是F-线性空间. 我们考虑:

- 当$F=\mathbb{R}$时V上的双线性形式;
- 当$F=\mathbb{C}$时V上的$1\frac{1}{2}$-线性形式.

后者的定义如下:

定义 9.1. 设V是复线性空间. V上的函数$f: V\times V\to\mathbb{C}$如果满足:

(a) $f(c\alpha+\beta,\gamma)=cf(\alpha,\gamma)+f(\beta,\gamma)$,

(b) $f(\alpha,c\beta+\gamma)=\bar{c}f(\alpha,\beta)+f(\beta,\gamma)$,

则称f是V上的$1\frac{1}{2}$**-线性函数**(sesqui-linear function)或$1\frac{1}{2}$**-线性形式**(sesqui-linear form).

我们把这两者都简称为V上的形式, 并把V上的所有形式的集合记为$\mathrm{Form}(V)$, 它在自然定义的运算下构成线性空间.

设$\dim V<\infty$, $f\in\mathrm{Form}(V)$, $\mathcal{B}=\{\alpha_1,\dots,\alpha_n\}$是$V$的有序基. 称$(j,k)$-元为$f(\alpha_k,\alpha_j)$的$n$阶方阵为$f$**在**$\mathcal{B}$**下的矩阵**, 记为$[f]_{\mathcal{B}}$. 对于$\alpha=\sum_k x_k\alpha_k, \beta=\sum_j y_j\alpha_j\in V$, 有

$$f(\alpha,\beta)=f(\sum_k x_k\alpha_k,\sum_j y_j\alpha_j)=\sum_{j,k}x_k\bar{y}_j f(\alpha_k,\alpha_j)=[\beta]_{\mathcal{B}}^*[f]_{\mathcal{B}}[\alpha]_{\mathcal{B}}.$$

特别地, $[f]_{\mathcal{B}}$的定义与内积的矩阵的定义是一致的.

例 9.1. 设$A\in F^{n\times n}$. 定义$f\in\mathrm{Form}(F^{n\times 1})$为

$$f(X,Y)=Y^*AX.$$

设$\mathcal{B}$是$F^{n\times 1}$的标准有序基. 则$[f]_{\mathcal{B}}=A$.

命题 9.1. 如果$\mathcal{B}'=\{\alpha_1',\dots,\alpha_n'\}$是另一组有序基, $[\alpha_1',\dots,\alpha_n']=[\alpha_1,\dots,\alpha_n]P$, 则

$$[f]_{\mathcal{B}'}=P^*[f]_{\mathcal{B}}P.$$

证明. 此时有$[\alpha]_{\mathcal{B}}=P[\alpha]_{\mathcal{B}'}$. 所以

$$[\beta]_{\mathcal{B}'}^*[f]_{\mathcal{B}'}[\alpha]_{\mathcal{B}'}=f(\alpha,\beta)=[\beta]_{\mathcal{B}}^*[f]_{\mathcal{B}}[\alpha]_{\mathcal{B}}=[\beta]_{\mathcal{B}'}^*P^*[f]_{\mathcal{B}}P[\alpha]_{\mathcal{B}'},\qquad\forall\alpha,\beta\in V.$$

因此$[f]_{\mathcal{B}'}=P^*[f]_{\mathcal{B}}P$. □

定义 9.2. 设$f\in\mathrm{Form}(V)$.

(1) 如果$F=\mathbb{R}$并且$f(\alpha,\beta)=f(\beta,\alpha)$, 则称$f$是**对称的**.

(2) 如果$F=\mathbb{C}$并且$f(\alpha,\beta)=\overline{f(\beta,\alpha)}$, 则称$f$是**Hermite的**.

为了与书上统一, 我们把两种情况都称为Hermite的.

引理 9.2. 如果$F=\mathbb{C}$, 则f Hermite $\iff f(\alpha,\alpha)\in\mathbb{R},\forall\alpha\in V$.

证明. "$\Longrightarrow$": 显然. "$\Longleftarrow$": 把$f(\alpha+\beta,\alpha+\beta)$展开, 条件推出

$$f(\alpha,\beta)+f(\beta,\alpha)\in\mathbb{R},\qquad\forall\alpha,\beta\in V,$$

从而

$$f(\alpha,\beta)+f(\beta,\alpha)=\overline{f(\alpha,\beta)}+\overline{f(\beta,\alpha)},\qquad\forall\alpha,\beta\in V.$$

将此式中的β替换为$i\beta$, 得

$$-if(\alpha,\beta)+if(\beta,\alpha)=i\overline{f(\alpha,\beta)}-i\overline{f(\beta,\alpha)},\qquad\forall\alpha,\beta\in V.$$

由此得

$$f(\alpha,\beta)=\overline{f(\beta,\alpha)},\qquad\forall\alpha,\beta\in V.$$

□

定义 9.3. 设$f\in\mathrm{Form}(V)$是Hermite的.

- 如果$f(\alpha,\alpha)>0,\forall\alpha\in V\setminus\{0\}$, 则称$f$是**正定的**(positive definite)或**正的**(positive).
- 如果$f(\alpha,\alpha)\geq 0,\forall\alpha\in V$, 则称$f$是**半正定的**(positive semi-definite)或**非负的**(non-negative).

注意正定Hermite形式就是内积.

定义 9.4. 设$A\in F^{n\times n}$是Hermite矩阵.

- 如果$X^*AX>0,\forall X\in F^{n\times 1}\setminus\{0\}$, 则称$A$是**正定的**或**正的**.
- 如果$X^*AX\geq 0,\forall X\in F^{n\times 1}$, 则称$A$是**半正定的**或**非负的**.

命题 9.3. 设$\dim V<\infty$, $\mathcal{B}$是有序基, 则f Hermite $\Longleftrightarrow$ $[f]_{\mathcal{B}}$ Hermite. 此时, f正定$\Longleftrightarrow$ $[f]_{\mathcal{B}}$正定, f半正定$\Longleftrightarrow$ $[f]_{\mathcal{B}}$半正定.

证明. "$\Longrightarrow$": 显然. "$\Longleftarrow$": 只需注意到

$$\begin{aligned}f(\alpha,\beta)&=[\beta]_{\mathcal{B}}^*[f]_{\mathcal{B}}[\alpha]_{\mathcal{B}},\\ \overline{f(\beta,\alpha)}&=([\alpha]_{\mathcal{B}}^*[f]_{\mathcal{B}}[\beta]_{\mathcal{B}})^*=[\beta]_{\mathcal{B}}^*[f]_{\mathcal{B}}^*[\alpha]_{\mathcal{B}}.\end{aligned}$$

其余结论由

$$f(\alpha,\alpha)=[\alpha]_{\mathcal{B}}^*[f]_{\mathcal{B}}[\alpha]_{\mathcal{B}}$$

是显然的. □

下面讨论正定矩阵. 容易看出:

- A正定$\Longrightarrow$ A可逆.
- A正定, $P\in GL_n(F)\Longrightarrow P^*AP$正定. 特别地, $P\in GL_n(F)\Longrightarrow P^*P$正定.

命题 9.4(Cholesky分解). 设$A\in F^{n\times n}$正定. 则存在唯一的对角元为正数的上三角矩阵R满足$A=R^*R$.

证明. 考虑$F^{n\times 1}$上的内积$f(X,Y)=Y^*AX$和标准内积$f_0(X,Y)=Y^*X$. 则对于$R\in GL_n(F)$,

$$\begin{aligned}&\text{映射}L_R:(F^{n\times 1},f)\to(F^{n\times 1},f_0)\text{是内积空间的同构}\\ \Longleftrightarrow\ &f_0(RX,RY)=f(X,Y),\qquad\forall X,Y\in F^{n\times 1}\\ \Longleftrightarrow\ &Y^*R^*RX=Y^*AX,\qquad\forall X,Y\in F^{n\times 1}\\ \Longleftrightarrow\ &A=R^*R.\end{aligned}$$

取定$P\in GL_n(F)$使得$L_P:(F^{n\times 1},f)\to(F^{n\times 1},f_0)$是内积空间的同构. 则对于$R\in GL_n(F)$,

$$\begin{aligned}&L_R:(F^{n\times 1},f)\to(F^{n\times 1},f_0)\text{是内积空间的同构}\\ \Longleftrightarrow\ &L_R\circ L_P^{-1}:(F^{n\times 1},f_0)\to(F^{n\times 1},f_0)\text{是内积空间的同构}\\ \Longleftrightarrow\ &RP^{-1}\in O(n)(\text{或}U(n))\\ \Longleftrightarrow\ &PR^{-1}\in O(n)(\text{或}U(n)).\end{aligned}$$

由QR分解, P可唯一分解为$O(n)$(或$U(n)$)中矩阵与对角元为正数的上三角矩阵的乘积, 即存在唯一的对角元为正数的上三角矩阵R使得$PR^{-1} \in O(n)$(或$U(n)$). □

推论 9.5. A正定$\Longrightarrow \det(A) > 0$.

证明. 设$A = P^*P$, P可逆. 则$\det(A) = |\det(P)|^2 > 0$. □

下面给出一个判断Hermite矩阵何时正定的方法.

定义 9.5. 设$A \in F^{n\times n}$, $1 \le k \le n$. 行列式

$$\Delta_k(A) := \det \begin{bmatrix} A_{11} & \cdots & A_{1k} \\ \vdots & \vdots & \vdots \\ A_{k1} & \cdots & A_{kk} \end{bmatrix}$$

称为A的第k个**顺序主子式**(leading principal minor).

定理 9.6. 设$A \in F^{n\times n}$ Hermite. 则A正定$\Longleftrightarrow \Delta_k(A) > 0, k = 1, \ldots, n$.

引理 9.7(LU分解). 设F是任意域. 则对$A \in GL_n(F)$, TFAE:

(1) 存在$L, U \in GL_n(F)$, L下三角, U上三角, 满足$A = LU$.

(2) $\Delta_k(A) \neq 0$, $k = 1, \ldots, n-1$.

此时, 如果还要求L(或U)的对角元都是1, 则满足(1)的L, U还是唯一的.

证明. "(1)$\Longrightarrow$(2)": 把L, U写成分块矩阵$L = \begin{bmatrix} L_k & 0 \\ * & * \end{bmatrix}$, $U = \begin{bmatrix} U_k & * \\ 0 & * \end{bmatrix}$, 其中$L_k, U_k \in F^{k\times k}$. 则$A = LU = \begin{bmatrix} L_kU_k & * \\ * & * \end{bmatrix}$. 因此

$$\Delta_k(A) = \det(L_kU_k) = \det(L_k)\det(U_k) \neq 0.$$

"(2)$\Longrightarrow$(1)和分解的唯一性": 只需证明存在唯一的严格上三角矩阵(即对角元都为0)$N \in F^{n\times n}$使得$A(N+I)$下三角, 即$A(N+I)$的第$k+1$列的前k行为0, $1 \le k \le n-1$. 设$A = \begin{bmatrix} A_k & * \\ * & * \end{bmatrix}$, 其中$A_k \in F^{k\times k}$. 设$A, N$的第$k+1$列分别为$\alpha_{k+1}, N_{k+1}$, 它们的前$k$行分别为$\alpha'_{k+1}, N'_{k+1} \in F^{k\times 1}$. 则$N_{k+1} = \begin{bmatrix} N'_{k+1} \\ 0 \end{bmatrix}$. 于是$A(N+I)$的第$k+1$列为

$$AN_{k+1} + \alpha_{k+1} = \begin{bmatrix} A_kN'_{k+1} + \alpha'_{k+1} \\ * \end{bmatrix},$$

从而其前k行为$A_kN'_{k+1} + \alpha'_{k+1}$. 因此只需证明: 对任意$1 \le k \le n-1$, 存在唯一的$N'_{k+1} \in F^{k\times 1}$满足

$$A_kN'_{k+1} + \alpha'_{k+1} = 0.$$

这在(2)的假设下是显然的. (实际上可求得$N'_{k+1} = -A_k^{-1}\alpha'_{k+1}$.) □

注 9.1. 集合

$$\{LU \mid L, U \in GL_n(F), L\text{下三角}, U\text{上三角}\}$$

称为$GL_n(F)$中的一个Bruhat big cell. 它与群$GL_n(F)$的Bruhat分解有关.

定理9.6的证明. "$\Longrightarrow$": 设A正定. 我们已经知道$\Delta_n(A) = \det(A) > 0$. 设$1 \le k \le n-1$,

$A=\begin{bmatrix}A_k & * \\ * & *\end{bmatrix}$. 则$A_k$ Hermite, 并且对任意$X\in F^{k\times 1}\setminus\{0\}$有

$$X^*A_kX=\begin{bmatrix}X\\0\end{bmatrix}^* A\begin{bmatrix}X\\0\end{bmatrix}>0.$$

所以A_k正定, 因此$\Delta_k(A)=\det(A_k)>0$.

"$\Longrightarrow$": 假设$A\in F^{n\times n}$ Hermite 并且$\Delta_k(A)>0, k=1,\dots,n$. 我们证明$A$正定. 由$LU$分解, 存在$L,U\in GL_n(F)$, L下三角, U上三角且对角元都为1, 满足$A=LU$. 记$D=(U^*)^{-1}L$. 则$A=U^*DU$, 从而D Hermite. 注意到D还是下三角的, 从而D对角. 由

$$\begin{bmatrix}A_k & * \\ * & *\end{bmatrix}=\begin{bmatrix}U_k^* & 0 \\ * & *\end{bmatrix}\begin{bmatrix}D_k & * \\ * & *\end{bmatrix}\begin{bmatrix}U_k & * \\ 0 & *\end{bmatrix}$$

得$A_k=U_k^*D_kU_k$. 因此

$$\Delta_k(A)=\det(A_k)=\det(U_k^*D_kU_k)=\det(D_k)=\Delta_k(D)>0.$$

这推出D的对角元都>0, 从而正定. 因此$A=U^*DU$正定. □

§9.2 内积空间上的形式

本节假设V是$F=\mathbb{R}$或$\mathbb{C}$上的有限维内积空间. 此时, 取定的内积给出了$L(V)$与$\mathrm{Form}(V)$的一一对应. 首先, 对任意$T\in L(V)$, $f(\alpha,\beta)=\langle T\alpha,\beta\rangle$是形式. 这给出了线性映射

$$\mathcal{F}: L(V)\to \mathrm{Form}(V),\qquad \mathcal{F}(T)(\alpha,\beta)=\langle T\alpha,\beta\rangle.$$

另一方面, 对任意$f\in\mathrm{Form}(V)$, 考虑共轭线性映射

$$\Phi_f: V\to V^*,\qquad \Phi_f(\alpha)(\beta)=\overline{f(\alpha,\beta)}.$$

回忆我们还有共轭线性同构

$$\Phi: V\to V^*,\qquad \Phi(\alpha)(\beta)=\langle\beta,\alpha\rangle=\overline{\langle\alpha,\beta\rangle}.$$

则

$$T_f:=\Phi^{-1}\circ\Phi_f$$

线性. 这给出了线性映射

$$\mathcal{T}: \mathrm{Form}(V)\to L(V),\qquad \mathcal{T}(f)=T_f=\Phi^{-1}\circ\Phi_f.$$

注意T_f满足性质

$$\langle T_f\alpha,\beta\rangle=\overline{\Phi(T_f\alpha)(\beta)}=\overline{\Phi_f(\alpha)(\beta)}=f(\alpha,\beta).$$

引理 9.8. $\mathcal{F}$与$\mathcal{T}$互为逆映射.

证明. 首先,

$$(\mathcal{F}\circ\mathcal{T}(f))(\alpha,\beta)=\mathcal{F}(T_f)(\alpha,\beta)=\langle T_f\alpha,\beta\rangle=f(\alpha,\beta),$$

即$\mathcal{F}\circ\mathcal{T}(f)=f$. 另一方面,

$$\langle(\mathcal{T}\circ\mathcal{F}(T))\alpha,\beta\rangle=\langle T_{\mathcal{F}(T)}\alpha,\beta\rangle=\mathcal{F}(T)(\alpha,\beta)=\langle T\alpha,\beta\rangle.$$

因此$\mathcal{T}\circ\mathcal{F}(T)=T$. □

命题 9.9. 设$f\in\mathrm{Form}(V)$. 则f Hermite $\Longleftrightarrow T_f$自伴. 此时, f正定$\Longleftrightarrow T_f$正定, f半正定$\Longleftrightarrow T_f$半正定.

这里T_f正定或半正定指:

定义 9.6. 设$T \in L(V)$.

- 如果T自伴并且$\langle T\alpha, \alpha\rangle > 0, \forall \alpha \in V \setminus \{0\}$, 则称$T$是**正定的**或**正的**.
- 如果T自伴并且$\langle T\alpha, \alpha\rangle \geq 0, \forall \alpha \in V$, 则称$T$是**半正定的**或**非负的**.

先证明:

引理 9.10. 设$\mathcal{B} = \{\alpha_1, \dots, \alpha_n\}$是$V$的有序标准正交基.

(1) 设$T \in L(V)$, $A = [T]_{\mathcal{B}}$. 则$A_{kj} = \langle T\alpha_j, \alpha_k\rangle$.

(2) 设$f \in \mathrm{Form}(V)$. 则$[T_f]_{\mathcal{B}} = [f]_{\mathcal{B}}$.

证明. (1) 对任意$\alpha \in V$有$\alpha = \sum_k \langle \alpha, \alpha_k\rangle \alpha_k$. 因此
$$T\alpha_j = \sum_k \langle T\alpha_j, \alpha_k\rangle \alpha_k.$$
而A由
$$T\alpha_j = \sum_k A_{kj}\alpha_k$$
决定. 比较系数即得.

(2)
$$[f]_{\mathcal{B}}\text{的}(j,k)\text{-元} = f(\alpha_k, \alpha_j) = \langle T_f\alpha_k, \alpha_j\rangle = [T_f]_{\mathcal{B}}\text{的}(j,k)\text{-元}.$$

□

注 9.2. $\langle f, g\rangle = \mathrm{tr}(T_f T_g^*)$定义了$\mathrm{Form}(V)$上的内积, 并且对任意$V$的标准正交基$\mathcal{B} = \{\alpha_1, \dots, \alpha_n\}$, 有
$$\langle f, g\rangle = \sum_{j,k} f(\alpha_j, \alpha_k)\overline{g(\alpha_j, \alpha_k)}.$$
事实上, 只需注意到$(T, U) \mapsto \mathrm{tr}(TU^*)$是$L(V)$上的内积, 并且, 记$A = [f]_{\mathcal{B}}$, $B = [g]_{\mathcal{B}}$, 则
$$\mathrm{tr}(T_f T_g^*) = \mathrm{tr}([T_f T_g^*]_{\mathcal{B}}) = \mathrm{tr}(AB^*) = \sum_{j,k} A_{jk}\overline{B_{jk}} = \sum_{j,k} f(\alpha_j, \alpha_k)\overline{g(\alpha_j, \alpha_k)}.$$

命题9.9的证明. 取V的有序标准正交基$\mathcal{B}$. 由引理,
$$f \text{ Hermite} \iff [f]_{\mathcal{B}} \text{ Hermite} \iff [T_f]_{\mathcal{B}} \text{ Hermite} \iff T_f\text{自伴}.$$
关于正定性和半正定性的结论由定义是显然的. □

推论 9.11. 设$F = \mathbb{C}$. 则$T \in L(V)$自伴$\iff \langle T\alpha, \alpha\rangle \in \mathbb{R}, \forall \alpha \in V$.

证明. 取$f \in \mathrm{Form}(V)$使$T_f = T$. 则T_f自伴$\iff f$ Hermite $\iff \langle T_f\alpha, \alpha\rangle = f(\alpha, \alpha) \in \mathbb{R}$. □

推论 9.12. 设$f \in \mathrm{Form}(V)$.

(1) (主轴定理) 如果f Hermite, 则存在有序标准正交基$\mathcal{B}$使得$[f]_{\mathcal{B}}$实对角.

(2) 如果$F = \mathbb{C}$, 则存在有序标准正交基$\mathcal{B}$使得$[f]_{\mathcal{B}}$上三角.

证明. (1) f Hermite $\implies T_f$自伴. 因此可取$\mathcal{B}$使得$[f]_{\mathcal{B}} = [T_f]_{\mathcal{B}}$实对角.

(2) 由Schur三角化定理, 可取$\mathcal{B}$使得$[f]_{\mathcal{B}} = [T_f]_{\mathcal{B}}$上三角. □

注 9.3. 在主轴定理中, 如果$\mathcal{B} = \{\alpha_1, \dots, \alpha_n\}$并且$[f]_{\mathcal{B}} = \mathrm{diag}(c_1, \dots, c_n)$, 则
$$f\Big(\sum_k x_k\alpha_k, \sum_j y_j\alpha_j\Big) = \sum_j c_j x_j \bar{y}_j.$$

§9.3 谱分解

设$F=\mathbb{R}$或$\mathbb{C}$, V是F上的有限维内积空间. 回忆:

- 对于子空间$W\subset V$有$V=W\oplus W^{\perp}$, 从而任意$\alpha\in V$可唯一分解为$\alpha=\beta+\gamma$, 其中$\beta\in W$, $\gamma\in W^{\perp}$. 映射$P_W\in L(V)$, $P_W(\alpha)=\beta$称为**到W上的正交投影**.
- 如果直和分解$V=\bigoplus_{i=1}^{k}W_i$中的直和项W_i两两正交, 则称它为**正交直和分解**. 如果$F=\mathbb{R}$并且T自伴, 或者$F=\mathbb{C}$并且T正规, 则有正交直和分解$V=\bigoplus_{c\in\sigma(T)}V_c$.

定理 9.13(正规变换的谱分解). 设$T\in L(V)$. 假设当$F=\mathbb{R}$时假设T自伴, 当$F=\mathbb{C}$时假设T正规. 设$\sigma(T)=\{c_1,\dots,c_k\}$, P_i为到V_{c_i}上的正交投影. 则

$$T=\sum_{i=1}^{k}c_iP_i,$$

并且对任意$f\in F[x]$有

$$f(T)=\sum_{i=1}^{k}f(c_i)P_i.$$

证明. 取V_{c_i}的有序标准正交基$\mathcal{B}_i$. 则$\mathcal{B}=(\mathcal{B}_1,\dots,\mathcal{B}_k)$是$V$的有序标准正交基, 并且$[T]_{\mathcal{B}}=\operatorname{diag}(c_1I_{d_1},\dots,c_kI_{d_k})$, 其中$d_i=\dim V_{c_i}$. 另一方面, 任意$\alpha\in V$可唯一分解为$\alpha=\sum_{i=1}^{k}\alpha_i$, 其中$\alpha_i\in V_{c_i}$. 注意到$V_{c_i}^{\perp}=\bigoplus_{j\neq i}V_{c_j}$, 从而$P_i\alpha=\alpha_i$. 这说明$[P_i]_{\mathcal{B}}=\operatorname{diag}(\delta_{i1}I_{d_1},\dots,\delta_{ik}I_{d_k})$. 因此

$$[f(T)]_{\mathcal{B}}=\operatorname{diag}(f(c_1)I_{d_1},\dots,f(c_k)I_{d_k})=\sum_{i=1}^{k}f(c_i)[P_i]_{\mathcal{B}}=\left[\sum_{i=1}^{k}f(c_i)P_i\right]_{\mathcal{B}},$$

即$f(T)=\sum_{i=1}^{k}f(c_i)P_i$. 取$f=x$即得$T=\sum_{i=1}^{k}c_iP_i$. □

推论 9.14. 每个P_i为T的多项式.

证明. 取$f_i\in F[x]$使$f_i(c_j)=\delta_{ij}$. 则$P_i=f_i(T)$. □

谱分解定理还可以用来定义T的其他函数: 如果F的子集S包含$\sigma(T)$, $\phi:S\to F$是任意函数, 则可以定义

$$\phi(T)=\sum_{i=1}^{k}\phi(c_i)P_i.$$

例如, 如果T半正定, 则$\sigma(T)\subset[0,+\infty)$, 从而可以定义

$$\sqrt{T}=\sum_{i=1}^{k}\sqrt{c_i}P_i.$$

命题 9.15. (1) $\phi(T)$可对角化, 并且$\sigma(\phi(T))=\phi(\sigma(T))$.

(2) 当$F=\mathbb{R}$时, T自伴$\Longrightarrow\phi(T)$自伴. 当$F=\mathbb{C}$时, T正规$\Longrightarrow\phi(T)$正规.

证明. 利用上面定理证明中的记号, 显然有$[\phi(T)]_{\mathcal{B}}=\operatorname{diag}(\phi(c_1)I_{d_1},\dots,\phi(c_k)I_{d_k})$. □

对于矩阵, 假设$A\in F^{n\times n}$可对角化, $A=P\operatorname{diag}(a_1,\dots,a_n)P^{-1}$, 则可以定义

$$\phi(A)=P\operatorname{diag}(\phi(a_1),\dots,\phi(a_n))P^{-1}.$$

注意定义与$a_1,\dots,a_n$的排序和P的选取无关: 取多项式f满足$f(a_i)=\phi(a_i)$, 则$\phi(A)=f(A)$.

作为应用, 我们证明:

定理 9.16. 设$T\in L(V)$半正定. 则存在唯一的半正定$\sqrt{T}\in L(V)$满足$(\sqrt{T})^2=T$. 此时, T正定$\Longleftrightarrow\sqrt{T}$正定.

证明. 存在性: 设$T=\sum_{i=1}^k c_iP_i$是谱分解. 则$c_i\geq 0$. 定义
$$\sqrt{T}=\sum_{i=1}^k\sqrt{c_i}P_i.$$
则$\sqrt{T}$半正定, 并且$(\sqrt{T})^2=\sum_{i=1}^k(\sqrt{c_i})^2P_i=T$ (注意$P_i^2=P_i$并且当$i\neq j$时有$P_iP_j=0$). 容易看出T正定$\Longleftrightarrow\sqrt{T}$正定.

唯一性: 设N,N'半正定并且$N^2=N'^2=T$. 考虑谱分解$N=\sum_{i=1}^k c_iP_i$, $N'=\sum_{i=1}^{k'}c_i'P_i'$, 其中$c_i,c_i'\geq 0$. 则$\sum_{i=1}^k c_i^2P_i=T=\sum_{i=1}^{k'}(c_i')^2P_i'$. 从而$\{c_1^2,\dots,c_k^2\}=\sigma(T)=\{(c_1')^2,\dots,(c_{k'}')^2\}$. 这推出$k'=k$并且重新排序后有$c_i=c_i'$. 现在有$\sum_{i=1}^k c_i^2P_i=T=\sum_{i=1}^k(c_i)^2P_i'$. 这推出$\mathrm{Im}(P_i)=\mathrm{Ker}(T-c_i^2\,\mathrm{id})=\mathrm{Im}(P_i')$. 从而$P_i=P_i'$. 因此$N=N'$. □

注意到对任意$T\in L(V)$, T^*T半正定. 因此可以考虑$N:=\sqrt{T^*T}$. N的特征值称为T的**奇异值**(singular value), 在很多方面有重要的应用.

引理 9.17. 设$T\in L(V)$, $N=\sqrt{T^*T}$. 则
$$\|T\alpha\|=\|N\alpha\|,\qquad\forall\alpha\in V.$$
特别地, $\mathrm{Ker}(T)=\mathrm{Ker}(N)$.

证明.
$$\|N\alpha\|^2=\langle N\alpha,N\alpha\rangle=\langle N^2\alpha,\alpha\rangle=\langle T^*T\alpha,\alpha\rangle=\langle T\alpha,T\alpha\rangle=\|T\alpha\|^2.$$
□

定理 9.18(极分解或Cartan分解). 设$T\in L(V)$. 则存在酉变换U(当$F=\mathbb{R}$时为正交变换)和唯一的半正定变换N满足$T=UN$. T可逆$\Longleftrightarrow N$正定, 此时U也唯一.

证明. 唯一性: 设$T=UN$, U酉, N半正定. 则$T^*=N^*U^*=NU^*$. 于是$T^*T=(NU^*)(UN)=N^2$. 注意到T^*T半正定. 所以$N=\sqrt{T^*T}$唯一. 显然有T可逆$\Longleftrightarrow N$可逆$\Longleftrightarrow N$正定. 此时$U=TN^{-1}$也唯一.

存在性: 设$N=\sqrt{T^*T}$. 由于$\mathrm{Ker}(T)=\mathrm{Ker}(N)$, 所以存在$U_1\in L(\mathrm{Im}(N),\mathrm{Im}(T))$满足$T=U_1N$. 由于$\|U_1N\alpha\|=\|T\alpha\|=\|N\alpha\|$, 所以对任意$\beta\in\mathrm{Im}(N)$有$\|U_1\beta\|=\|\beta\|$, 即$U_1$保长度, 从而保内积. 再注意到$\dim\mathrm{Im}(N)=\dim\mathrm{Im}(T)$, 所以$U_1$是内积空间的同构. 再任取内积空间的同构$U_2:\mathrm{Im}(N)^\perp\to\mathrm{Im}(T)^\perp$. 则$U:=U_1\oplus U_2:V\to V$也是内积空间的同构(可以取$\mathrm{Im}(N)$和$\mathrm{Im}(N)^\perp$的标准正交基来看), 从而酉. □

推论 9.19(奇异值分解). 设$A\in F^{n\times n}$. 则存在非负对角矩阵D和酉矩阵(当$F=\mathbb{R}$时为正交矩阵)U_1,U_2满足$A=U_1DU_2$.

证明. 由极分解, 存在酉矩阵U和半正定矩阵N满足$A=UN$. 设$N=VDV^{-1}$, 其中V酉, D非负对角. 则$A=UVDV^{-1}$. 取$U_1=UV$, $U_2=V^{-1}$即可. □

推论 9.20. 设$T\in L(V)$可逆. 则T把V的某个正交基映为正交基, 即存在正交基$\{\alpha_1,\dots,\alpha_n\}$使得$\{T\alpha_1,\dots,T\alpha_n\}$也是正交基.

证明. 设$T=UN$, U酉, N正定. 取有序标准正交基$\mathcal{B}=\{\alpha_1,\dots,\alpha_n\}$使得$[N]_\mathcal{B}$正对角, 即$N\alpha_i=c_i\alpha_i$, $c_i>0$. 则$T\alpha_i=UN\alpha_i=c_iU\alpha_i$. 显然$\{U\alpha_1,\dots,U\alpha_n\}$是标准正交基. 所以$\{c_1U\alpha_1,\dots,c_nU\alpha_n\}$是正交基. □

§9.4 正规变换的进一步性质

本节的主要目的之一是讨论实正规变换的性质. 我们将证明:

定理 9.21. 设V是有限维实内积空间, $T \in L(V)$正规. 则存在V的有序标准正交基$\mathcal{B}$使得

$$[T]_{\mathcal{B}} = \operatorname{diag}\left(a_1, \dots, a_l, r_1 \begin{bmatrix} \cos\theta_1 & -\sin\theta_1 \\ \sin\theta_1 & \cos\theta_1 \end{bmatrix}, \dots, r_m \begin{bmatrix} \cos\theta_m & -\sin\theta_m \\ \sin\theta_m & \cos\theta_m \end{bmatrix}\right),$$

其中$a_i \in \mathbb{R}$, $r_j > 0$, $\theta_j \in (0, \pi)$.

容易看出, T正交$\iff a_i = \pm 1$并且$r_j = 1$. 因此得到:

推论 9.22. 设V是有限维实内积空间, $T \in L(V)$正交. 则存在V的有序标准正交基$\mathcal{B}$使得

$$[T]_{\mathcal{B}} = \operatorname{diag}\left(I_{l_1}, -I_{l_2}, \begin{bmatrix} \cos\theta_1 & -\sin\theta_1 \\ \sin\theta_1 & \cos\theta_1 \end{bmatrix}, \dots, \begin{bmatrix} \cos\theta_m & -\sin\theta_m \\ \sin\theta_m & \cos\theta_m \end{bmatrix}\right),$$

其中$\theta_j \in (0, \pi)$.

下面的推论也是显然的.

推论 9.23. 设$A \in \mathbb{R}^{n \times n}$.

(1) 如果A正规, 则A正交相似于定理9.21中的矩阵.

(2) 如果A正交, 则A正交相似于推论9.22中的矩阵. □

我们给出定理9.21的三个证明, 分别利用正规变换的半单性、准素循环分解和比较特征多项式. 前两个证明本质上是等价的. 但是, 为了介绍证明过程中一些有意思的中间结果, 我们给出每个证明的细节. 实际上, 我们不只满足于中间结果对证明定理有用, 而是给出更一般的形式. 在本节中, 如果未做其他说明, 我们总是假设$F = \mathbb{R}$或$\mathbb{C}$, V是F上的有限维内积空间.

第一个证明利用正规变换的半单性. 我们先证明:

引理 9.24. 设$T \in L(V)$正规, $W \subset V$是不变子空间. 则T_W正规.

证明. 我们知道, W也是T^*-不变子空间. 对任意$\alpha, \beta \in W$, 有

$$\langle (T_W)^*\alpha, \beta \rangle = \langle \alpha, T_W\beta \rangle = \langle \alpha, T\beta \rangle = \langle T^*\alpha, \beta \rangle.$$

由于$T^*\alpha \in W$, 所以$(T_W)^*\alpha = T^*\alpha$对任意$\alpha \in W$成立, 即$(T_W)^* = (T^*)_W$. 这推出

$$T_W(T_W)^* = T_W(T^*)_W = (TT^*)_W = (T^*T)_W = (T^*)_W T_W = (T_W)^* T_W,$$

所以T_W正规. □

引理 9.25. 设$T \in L(V)$正规. 则V可分解为不变子空间的正交直和$V = \bigoplus_{i=1}^k V_i$使得每个$T_{V_i}$单正规.

证明. 不妨假设T不单, 即存在非平凡不变子空间$W \subset V$. 由于T正规, $W^\perp$也是不变子空间. 由引理9.24, T_W和$T_{W^\perp}$正规. 由于$\dim W, \dim W^\perp < \dim V$, 对$\dim V$用归纳法, 可以假设引理对$T_W$和$T_{W^\perp}$成立. 这推出引理对$T$也成立. □

由引理9.25, 我们只需考察单正规变换.

引理 9.26. 设V是有限维实内积空间, $T \in L(V)$单正规. 假设$\dim V > 1$. 则$\dim V = 2$, 并且存在$r > 0$, $\theta \in (0, \pi)$和V的有序标准正交基$\mathcal{B}$使得

$$[T]_{\mathcal{B}} = r \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}.$$

其中r和θ被T唯一决定.

证明. 由T单可知f_T是$\mathbb{R}[x]$中的素多项式. 而$\dim V > 1$推出$\deg f_T > 1$. 因此$f_T = (x-c)(x-\bar{c})$, 其中$c \in \mathbb{C} \setminus \mathbb{R}$. 特别地, $\dim V = \deg f_T = 2$. 取V的有序标准正交基$\mathcal{B} = \{\alpha_1, \alpha_2\}$. 则$[T]_{\mathcal{B}}$正规, 从而酉相似于$\mathrm{diag}(c, \bar{c})$. 记$r = |c| > 0$. 则$A := r^{-1}[T]_{\mathcal{B}}$酉相似于酉矩阵$\mathrm{diag}(r^{-1}c, r^{-1}\bar{c})$, 从而$A$也是酉矩阵. 而$A$是实矩阵, 所以$A$正交. 此外还有$\det(A) = (r^{-1}c)(r^{-1}\bar{c}) = 1$. 因此只能有$A = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$, 其中$\theta \in [0, 2\pi)$. 注意到$T$单$\Longrightarrow T$无特征向量$\Longrightarrow \langle T\alpha_1, \alpha_2 \rangle \neq 0$. 必要时用$-\alpha_2$替换$\alpha_2$, 不妨设$\langle T\alpha_1, \alpha_2 \rangle > 0$. 而$r \sin\theta = [T]_{\mathcal{B}}$的$(2,1)$-元$= \langle T\alpha_1, \alpha_2 \rangle$. 所以$\sin\theta > 0$, 即$\theta \in (0, \pi)$. 另外, $r > 0$和$\theta \in (0, \pi)$在各自的取值范围内被矩阵rA的特征多项式唯一决定, 因此被T唯一决定. □

定理9.21的第一个证明. 由引理9.25, 存在正交直和分解$V = \bigoplus_{i=1}^k V_i$使得每个$T_{V_i}$单正规. 由引理9.26, 不妨设当$1 \le i \le l$时$\dim V_i = 1$, 当$1 \le j \le m$时$\dim V_{l+j} = 2$, 其中$l + m = k$. 对$1 \le i \le l$, 任取$V_i$的有序标准正交基$\mathcal{B}_i$. 对$1 \le j \le m$, 取$V_{l+j}$的有序标准正交基$\mathcal{B}_{l+j}$使得

$$[T_{V_{l+j}}]_{\mathcal{B}_{l+j}} = r_j \begin{bmatrix} \cos\theta_j & -\sin\theta_j \\ \sin\theta_j & \cos\theta_j \end{bmatrix},$$

其中$r_j > 0$, $\theta_j \in (0, \pi)$. 则V的有序标准正交基$\mathcal{B} = (\mathcal{B}_1, \ldots, \mathcal{B}_k)$满足定理的要求. □

第二个证明利用准素循环分解. 我们先证明一般线性变换的几个性质.

命题 9.27. 设$T \in L(V)$. 则$\mathrm{Ker}(T)^\perp = \mathrm{Im}(T^*)$, $\mathrm{Im}(T)^\perp = \mathrm{Ker}(T^*)$.

证明.

$$\begin{aligned} \alpha \in \mathrm{Ker}(T) &\Longleftrightarrow \langle T\alpha, \beta \rangle = 0, \forall \beta \in V \\ &\Longleftrightarrow \langle \alpha, T^*\beta \rangle = 0, \forall \beta \in V \\ &\Longleftrightarrow \alpha \in \mathrm{Im}(T^*)^\perp. \end{aligned}$$

所以$\mathrm{Ker}(T) = \mathrm{Im}(T^*)^\perp$. 两边取正交补即得第一个结论, 用$T^*$替换$T$即得第二个结论. □

注 9.4. 上一命题可以与一般域上有限维线性空间的结论$\mathrm{Ker}(T)^0 = \mathrm{Im}(T^t)$, $\mathrm{Im}(T)^0 = \mathrm{Ker}(T^t)$作比较.

命题 9.28. 设$T \in L(V)$. 则$\sigma(T^*) = \overline{\sigma(T)}$, 并且对$c \in \sigma(T)$有$\dim \mathrm{Ker}(T^* - \bar{c}I) = \dim \mathrm{Ker}(T - cI)$.

证明. 对$c \in F$有$(T - cI)^* = T^* - \bar{c}I$. 所以$\mathrm{Ker}(T - cI)^\perp = \mathrm{Im}(T^* - \bar{c}I)$. 因此

$$\dim \mathrm{Ker}(T - cI) = n - \dim \mathrm{Ker}(T - cI)^\perp = n - \dim \mathrm{Im}(T^* - \bar{c}I) = \dim \mathrm{Ker}(T^* - \bar{c}I).$$

这推出

$$c \in \sigma(T) \Longleftrightarrow \dim \mathrm{Ker}(T - cI) \neq 0 \Longleftrightarrow \dim \mathrm{Ker}(T^* - \bar{c}I) \neq 0 \Longleftrightarrow \bar{c} \in \sigma(T^*).$$

因此$\sigma(T^*) = \overline{\sigma(T)}$. □

下面讨论正规变换.

命题 9.29. 设$T \in L(V)$正规, 则对任意$c \in F$有$\mathrm{Ker}(T - cI) = \mathrm{Ker}(T^* - \bar{c}I)$. 特别地, $\mathrm{Ker}(T^*) = \mathrm{Ker}(T)$, 从而$\mathrm{Ker}(T) = \mathrm{Im}(T)^\perp$.

证明. 记$W = \mathrm{Ker}(T - cI)$. 则$(T^*)_W = (T_W)^* = (c\,\mathrm{id}_W)^* = \bar{c}\,\mathrm{id}_W$. 这说明$\mathrm{Ker}(T - cI) \subset \mathrm{Ker}(T^* - \bar{c}I)$. 由维数相等即得等式. (也可以把$T$和$c$分别替换为$T^*$和$\bar{c}$看出反过来的包含关系.)
□

注 9.5. 更直接的证法是利用$\|T\alpha\| = \|T^*\alpha\|$:
$$\|T\alpha\|^2 = \langle T\alpha, T\alpha\rangle = \langle \alpha, T^*T\alpha\rangle = \langle \alpha, TT^*\alpha\rangle = \langle T^*\alpha, T^*\alpha\rangle = \|T^*\alpha\|^2.$$

注 9.6. 书上有一引理说: 设$T \in L(V)$正规, 则$T^2\alpha = 0 \Longrightarrow T\alpha = 0$. 证明如下: 由$\mathrm{Ker}(T)\perp\mathrm{Im}(T)$知$T\alpha \in \mathrm{Ker}(T)\cap\mathrm{Im}(T) = \{0\}$. 事实上这对任何半单的$T$都成立: $T^2\alpha = 0 \Longrightarrow p_\alpha | x^2$. 但$T$半单推出$p_T$无平方因子, 从而$p_\alpha$无平方因子. 因此$p_\alpha | x$, 即$T\alpha = 0$.

我们知道正规变换的特征子空间互相正交. 这可以推广如下:

命题 9.30. 设$T \in L(V)$正规, $f, g \in F[x]$互素. 则$\mathrm{Ker}(f(T))$与$\mathrm{Ker}(g(T))$正交.

证明. 记$W = \mathrm{Ker}(f(T))$, 它是$g(T)$的不变子空间. 由于存在$a, b \in F[x]$满足$af + bg = 1$, 所以$\mathrm{id}_W = a(T)_W f(T)_W + b(T)_W g(T)_W = b(T)_W g(T)_W$, 于是$g(T)_W$可逆. 因此$W \subset \mathrm{Im}(g(T))$. 容易看出$g(T)$正规(利用$g(T)^* = \bar{g}(T^*)$). 所以$\mathrm{Im}(g(T))\perp\mathrm{Ker}(g(T))$. 综合起来即得$W\perp\mathrm{Ker}(g(T))$.
□

注 9.7. 类似可以证明, 对任意F, 任意$T \in L(V)$, 如果$f, g \in F[x]$互素, 则$\mathrm{Ker}(f(T))\cap\mathrm{Ker}(g(T)) = \{0\}$.

命题 9.31. 设$T \in L(V)$正规.

(1) T的准素分解是正交直和分解.

(2) T的循环分解可以取为正交直和分解.

(3) T的准素循环分解可以取为正交直和分解.

证明. (1) 由上一命题显然.

(2) 对$\dim V$作归纳. 取$\alpha_1 \in V$使$p_{\alpha_1} = p_T$. 则$(R\alpha_1)^\perp$是不变子空间. 对$T_{(R\alpha_1)^\perp}$用归纳假设即可.

(3) 由(1)和(2), 显然. (先做准素分解, 再对每个准素分量做正交循环分解.) □

引理 9.32. 设V是有限维实内积空间, $T \in L(V)$正规准素循环. 假设$\dim V > 1$. 则$\dim V = 2$, 并且存在$r > 0$, $\theta \in (0, \pi)$和V的有序标准正交基$\mathcal{B}$使得
$$[T]_{\mathcal{B}} = r\begin{bmatrix}\cos\theta & -\sin\theta \\ \sin\theta & \cos\theta\end{bmatrix}.$$

证明. 由引理9.26, 只需证明T单. 由T准素循环可知$f_T = p_T$为$\mathbb{R}[x]$中素多项式的幂. 而T正规$\Longrightarrow T$半单, 所以p_T在$\mathbb{R}[x]$中无平方因子. 因此f_T是$\mathbb{R}[x]$中的素多项式, 从而T单. □

定理9.21的第二个证明. 做正交的准素循环分解$V = \bigoplus_{i=1}^k V_i$. 则每个$T_{V_i}$正规. 其余类似于第一个证明. □

下面为定理9.21的第三个证明做准备. 回忆两个实方阵如果在$\mathbb{C}$上相似, 则它们在$\mathbb{R}$上相似. 酉相似与正交相似也有类似的关系.

定理 9.33. 如果$A, B \in \mathbb{R}^{n\times n}$酉相似, 则它们正交相似.

先证明:

引理 9.34(酉矩阵的QS分解). 对任意$U \in U(n)$, 存在$Q \in O(n)$和$S \in U(n)$满足:

- $U = QS$,
- 存在$f \in \mathbb{C}[x]$满足$S = f(U^tU)$.

注意S是对称酉矩阵. 引理的结果和证明可以与极分解做比较.

证明. 设$\sigma(U^tU)=\{c_1,\dots,c_k\}$. 取$f\in\mathbb{C}[x]$满足$f(c_i)^2=c_i$. 设$S=f(U^tU)$. 我们断言$S\in U(n)$, 并且$S^2=U^tU$. 事实上, 由$|c_i|=1$推出$|f(c_i)|=1$. 如果

$$U^tU=P\mathrm{diag}(c_1I_{d_1},\dots,c_kI_{d_k})P^{-1},$$

其中$P\in U(n)$, 则

$$S=P\mathrm{diag}(f(c_k)I_{d_1},\dots,f(c_k)I_{d_k})P^{-1}\in U(n),$$

并且

$$S^2=P\mathrm{diag}(f(c_k)^2I_{d_1},\dots,f(c_k)^2I_{d_k})P^{-1}=U^tU.$$

设$Q=US^{-1}\in U(n)$. 注意到$S=f(U^tU)$推出$S^t=S$. 所以

$$\overline{Q}Q^{-1}=\overline{U}S^tSU^{-1}=(U^t)^{-1}S^2U^{-1}=(U^t)^{-1}(U^tU)U^{-1}=I,$$

即$\overline{Q}=Q$, 从而$Q\in O(n)$. □

定理9.33的证明. 设$B=UAU^{-1}$, $U\in U(n)$. 两边取复共轭, 得$B=\overline{U}AU^t$. 由$UAU^{-1}=\overline{U}AU^t$得$A$与$U^tU$可交换. 由引理, 可设$U=QS$, 其中$Q\in O(n)$, $S=f(U^tU)\in U(n)$, $f\in\mathbb{C}[x]$. 则A与S可交换. 从而

$$B=UAU^{-1}=QSAS^{-1}Q^{-1}=QAQ^{-1}.$$

因此A与B正交相似. □

注意书上P356的推论(即正规矩阵酉相似于其有理标准形)是错的(这将推出Hermite矩阵的有理标准形是Hermite的, 这明显不对). 但它下面的两个结论是对的(即有理标准形相同的复(或实)正规矩阵酉(或正交)相似), 还可加强如下:

命题 9.35. 设$F=\mathbb{R}$或$\mathbb{C}$, $A,B\in F^{n\times n}$正规. TFAE:

(1) A与B酉相似($F=\mathbb{C}$时)或正交相似($F=\mathbb{R}$时).

(2) A与B在$\mathbb{F}$上相似.

(3) $f_A=f_B$.

证明. "(1)$\Longrightarrow$(2)"和"(2)$\Longrightarrow$(3)"显然. 只需证明"(3)$\Longrightarrow$(1)". 先假设$F=\mathbb{C}$. 则A,B分别酉相似于对角矩阵D_1,D_2. $f_A=f_B$推出$f_{D_1}=f_{D_2}$. 因此D_1与D_2在重排对角元后相等, 从而酉相似. 这推出A与B酉相似. 再假设$F=\mathbb{R}$. 由已经证明的$F=\mathbb{C}$的情况, A与B酉相似. 再由定理9.33, A与B正交相似. □

推论 9.36. 设$F=\mathbb{R}$或$\mathbb{C}$, $A\in F^{n\times n}$正规. 则A与A^T酉相似($F=\mathbb{C}$时)或正交相似($F=\mathbb{R}$时).

证明. 只需注意到A^T正规并且$f_A=f_{A^T}$. □

推论 9.37. 设V是F上的有限维内积空间, $T\in L(V)$正规. 设$A\in F^{n\times n}$正规. 如果$f_A=f_T$, 则存在V的有序标准正交基$\mathcal{B}$使得$[T]_{\mathcal{B}}=A$.

证明. 只证明$F=\mathbb{R}$的情况. $F=\mathbb{C}$的情况类似. 任取V的有序标准正交基$\mathcal{B}_0$. 设$A_0=[T]_{\mathcal{B}_0}$. 则A_0正规, 并且$f_{A_0}=f_A$. 因此A_0与A正交相似. 而

$$\{[T]_{\mathcal{B}}\mid \mathcal{B}\text{是}V\text{的有序标准正交基}\}=\{P^{-1}A_0P\mid P\text{正交}\}.$$

A属于后一集合, 所以也属于前一集合. □

定理9.21的另一证明. f_T在$\mathbb{R}[x]$中可以分解为

$$f_T = \prod_{i=1}^{k}(x-a_i)\prod_{j=1}^{m}(x^2-(2r_j\cos\theta_j)x+r_j^2),$$

其中$a_i \in \mathbb{R}$, $r_j > 0$, $\theta_j \in (0,\pi)$. 取

$$A = \operatorname{diag}\left(a_1,\ldots,a_k, r_1\begin{bmatrix}\cos\theta_1 & -\sin\theta_1\\ \sin\theta_1 & \cos\theta_1\end{bmatrix},\ldots, r_m\begin{bmatrix}\cos\theta_m & -\sin\theta_m\\ \sin\theta_m & \cos\theta_m\end{bmatrix}\right).$$

则$f_A = f_T$. 由上一推论, 存在V的有序标准正交基$\mathcal{B}$使得$[T]_{\mathcal{B}} = A$. □

最后再证明一个正规变换的性质.

命题 9.38. 设$T \in L(V)$正规, $S \in L(V)$与T可交换. 则S与T^*可交换.

证明. 先假设$F = \mathbb{C}$. 设$\sigma(T) = \{c_1,\ldots,c_k\}$. 取$f \in \mathbb{C}[x]$满足$f(c_i) = \bar{c}_i$. 取$V$的有序标准正交基$\mathcal{B}$使得$[T]_{\mathcal{B}} = D$对角. 则

$$[f(T)]_{\mathcal{B}} = f([T]_{\mathcal{B}}) = [T]_{\mathcal{B}}^* = [T^*]_{\mathcal{B}}.$$

这推出$f(T) = T^*$. 由此即得结论.

再假设$F = \mathbb{R}$. 任取V的有序标准正交基$\mathcal{B}$. 则$A := [T]_{\mathcal{B}}$正规, 并且$B := [S]_{\mathcal{B}}$与A可交换. 只需证明B与$A^* = [T^*]_{\mathcal{B}}$可交换. 视这些矩阵为复矩阵. 在$\mathbb{C}^n$上赋予标准内积. 则$L_A \in L(\mathbb{C}^n)$正规, 并且与$L_B \in L(\mathbb{C}^n)$可交换. 由已经证明的$F = \mathbb{C}$的情况, L_B与$(L_A)^* = L_{A^*}$可交换, 即B与A^*可交换. □

下面的结果在书上出现了三次.

命题 9.39. 设V是任意域上的有限维线性空间, $T \in L(V)$. 设$V = \bigoplus_{i=1}^k V_i$是准素分解, $W \subset V$是不变子空间. 则

$$W = \bigoplus_{i=1}^{k}(W \cap V_i).$$

特别地, 如果$W \cap V_1 = \{0\}$, 则$W \subset \bigoplus_{i=2}^k V_i$.

证明. 考虑T_W的准素分解并注意$p_{T_W} | p_T$即可. □

第十章　双线性形式

§10.1　双线性形式

参考讲义5.2和5.6节.

设F是任意域, V是有限维F-线性空间, $f \in (V^*)^{\otimes 2}$(书上记为$L(V,V,F)$或$M^2(V)$), $\mathcal{B} = \{\alpha_1,\dots,\alpha_n\}$是$V$的有序基. 称$(i,j)$-元为$f(\alpha_i,\alpha_j)$的$n$阶方阵为**$f$在$\mathcal{B}$下的矩阵**, 记为$[f]_{\mathcal{B}}$ (注意与前面不一致, 这里的$[f]_{\mathcal{B}}$是前面的$[f]_{\mathcal{B}}^t$). 对于

$$\alpha = \sum_i x_i\alpha_i, \beta = \sum_j y_j\alpha_j \in V,$$

有

$$f(\alpha,\beta) = f(\sum_i x_i\alpha_i, \sum_j y_j\alpha_j) = \sum_{i,j} x_iy_jf(\alpha_i,\alpha_j) = [\alpha]_{\mathcal{B}}^t[f]_{\mathcal{B}}[\beta]_{\mathcal{B}}.$$

(这与前面的$f(\alpha,\beta) = [\beta]_{\mathcal{B}}^*[f]_{\mathcal{B}}[\alpha]_{\mathcal{B}}$也不一致.)

引理 10.1. 映射$(V^*)^{\otimes 2} \to F^{n\times n}$, $f \mapsto [f]_{\mathcal{B}}$是线性同构.

证明. 线性显然.
单: 如果$[f]_{\mathcal{B}} = 0$, 则由$f(\alpha,\beta) = [\alpha]_{\mathcal{B}}^t[f]_{\mathcal{B}}[\beta]_{\mathcal{B}}$知$f = 0$.
满: 对任意$A \in F^{n\times n}$, $f(\alpha,\beta) = [\alpha]_{\mathcal{B}}^tA[\beta]_{\mathcal{B}}$满足$[f]_{\mathcal{B}} = A$. □

命题 10.2. 如果$\mathcal{B}' = \{\alpha_1',\dots,\alpha_n'\}$是另一组有序基, $[\alpha_1',\dots,\alpha_n'] = [\alpha_1,\dots,\alpha_n]P$, 则

$$[f]_{\mathcal{B}'} = P^t[f]_{\mathcal{B}}P.$$

证明. 此时有$[\alpha]_{\mathcal{B}} = P[\alpha]_{\mathcal{B}'}$. 所以

$$[\alpha]_{\mathcal{B}'}^t[f]_{\mathcal{B}'}[\beta]_{\mathcal{B}'} = f(\alpha,\beta) = [\alpha]_{\mathcal{B}}^t[f]_{\mathcal{B}}[\beta]_{\mathcal{B}} = [\alpha]_{\mathcal{B}'}^tP^t[f]_{\mathcal{B}}P[\beta]_{\mathcal{B}'}, \qquad \forall\alpha,\beta \in V.$$

因此$[f]_{\mathcal{B}'} = P^t[f]_{\mathcal{B}}P$. □

定义 10.1. $A, B \in F^{n\times n}$称为是**合同的**(congruent), 如果存在$P \in GL_n(F)$满足$B = P^tAP$.

容易看出, 对于$f \in (V^*)^{\otimes 2}$, 取V的所有有序基得到矩阵的一个合同等价类, 即对$A_0 = [f]_{\mathcal{B}_0}$,

$$\{[f]_{\mathcal{B}} \mid \mathcal{B}是V的有序基\} = \{P^tA_0P \mid P \in GL_n(F)\}.$$

命题 10.3. 设$f \in (V^*)^{\otimes 2}$, $\mathcal{B} = \{\alpha_1,\dots,\alpha_n\}$是$V$的有序基. 定义线性映射

$$L_f, R_f : V \to V^*, \qquad L_f(\alpha)(\beta) = R_f(\beta)(\alpha) = f(\alpha,\beta).$$

则$\operatorname{rank}(L_f) = \operatorname{rank}(R_f) = \operatorname{rank}([f]_{\mathcal{B}})$.

证明. 记$A = [f]_{\mathcal{B}}$. 则$f(\alpha,\beta) = [\alpha]_{\mathcal{B}}^tA[\beta]_{\mathcal{B}}$. 于是

$$\alpha \in \operatorname{Ker}(L_f) \iff [\alpha]_{\mathcal{B}}^tA[\beta]_{\mathcal{B}} = 0\forall\beta \iff [\alpha]_{\mathcal{B}}^tA = 0 \iff [\alpha]_{\mathcal{B}} \in \operatorname{Ker}(A^t).$$

所以$\dim\operatorname{Ker}(L_f) = \dim\operatorname{Ker}(A^t)$. 这推出$\operatorname{rank}(L_f) = \operatorname{rank}(A)$. 类似地,

$$\beta \in \operatorname{Ker}(R_f) \iff [\alpha]_{\mathcal{B}}^tA[\beta]_{\mathcal{B}} = 0\forall\alpha \iff A[\beta]_{\mathcal{B}} = 0 \iff [\beta]_{\mathcal{B}} \in \operatorname{Ker}(A).$$

所以$\dim\operatorname{Ker}(R_f) = \dim\operatorname{Ker}(A)$. 这推出$\operatorname{rank}(R_f) = \operatorname{rank}(A)$. □

$\operatorname{rank}(L_f) = \operatorname{rank}(R_f) = \operatorname{rank}([f]_{\mathcal{B}})$称为$f$的**秩**, 记为$\operatorname{rank}(f)$.

推论 10.4. 对于$f \in (V^*)^{\otimes 2}$, TFAE:

(1) $\operatorname{rank}(f) = \dim V$.

(2) 对任意$\alpha \in V \setminus \{0\}$, 存在$\beta \in V$使得$f(\alpha, \beta) \neq 0$.

(3) 对任意$\beta \in V \setminus \{0\}$, 存在$\alpha \in V$使得$f(\alpha, \beta) \neq 0$.

(4) 对某个V的有序基$\mathcal{B}$, $[f]_{\mathcal{B}}$可逆.

(5) 对任意V的有序基$\mathcal{B}$, $[f]_{\mathcal{B}}$可逆.

证明. (2)$\Longleftrightarrow \operatorname{Ker}(L_f) = \{0\} \Longleftrightarrow \operatorname{rank}(L_f) = \dim V \Longleftrightarrow$(1).
(3)$\Longleftrightarrow \operatorname{Ker}(R_f) = \{0\} \Longleftrightarrow \operatorname{rank}(R_f) = \dim V \Longleftrightarrow$(1).
(4)(或(5))$\Longleftrightarrow \operatorname{rank}([f]_{\mathcal{B}}) = \dim V \Longleftrightarrow$(1). □

满足这几个等价条件的f称为**非退化的**(non-degenerate).

§10.2 对称、反对称双线性形式和二次型

定义 10.2. 设$f \in (V^*)^{\otimes 2}$.

(1) 如果对任意$\alpha, \beta \in V$有$f(\alpha, \beta) = f(\beta, \alpha)$, 则称$f$是**对称**的(symmetric).

(2) 如果对任意$\alpha, \beta \in V$有$f(\alpha, \beta) + f(\beta, \alpha) = 0$, 则称$f$是**反对称**的(anti-symmetric或shew-symmetric).

(3) 如果对任意$\alpha \in V$有$f(\alpha, \alpha) = 0$, 则称f是**交错**的(alternating).

我们把V上的所有对称、反对称和交错双线性形式的集合分别记为$S^2(V^*)$, $A^2(V^*)$和$\Lambda^2(V^*)$. 显然它们都是$(V^*)^{\otimes 2}$的子空间.

注 10.1. 如果$\operatorname{char} F = 2$, 则$S^2(V^*) = A^2(V^*)$. 这种情况在这里不是讨论的重点. 所以必要时假设$\operatorname{char} F \neq 2$.

命题 10.5. (1) $\Lambda^2(V^*) \subset A^2(V^*)$.

(2) 如果$\operatorname{char} F \neq 2$, 则$\Lambda^2(V^*) = A^2(V^*)$, 并且$(V^*)^{\otimes 2} = S^2(V^*) \oplus A^2(V^*)$.

证明. (1) (以前已证) 若f交错, 则对任意$\alpha, \beta \in V$有

$$0 = f(\alpha + \beta, \alpha + \beta) = f(\alpha, \alpha) + f(\beta, \beta) + f(\alpha, \beta) + f(\beta, \alpha) = f(\alpha, \beta) + f(\beta, \alpha).$$

因此f反对称.

(2) 若f反对称, 则对任意$\alpha \in V$有

$$2f(\alpha, \alpha) = f(\alpha, \alpha) + f(\alpha, \alpha) = 0.$$

由于$\operatorname{char} F \neq 2$, 这推出$f(\alpha, \alpha) = 0$, 即$f$交错. 此外, 若$f \in S^2(V^*) \cap A^2(V^*)$, 则$f(\alpha, \beta) = f(\beta, \alpha) = -f(\alpha, \beta)$, 从而$f = 0$. 另一方面, 对任意$f \in (V^*)^{\otimes 2}$, 取

$$f_1(\alpha, \beta) = \frac{1}{2}(f(\alpha, \beta) + f(\beta, \alpha)), \qquad f_2(\alpha, \beta) = \frac{1}{2}(f(\alpha, \beta) - f(\beta, \alpha)).$$

则$f_1 \in S^2(V^*)$, $f_2 \in A^2(V^*)$, 并且$f = f_1 + f_2$. 因此$(V^*)^{\otimes 2} = S^2(V^*) \oplus A^2(V^*)$. □

命题 10.6. 设$\mathcal{B}$是有序基.

(1) f对称$\Longleftrightarrow [f]_{\mathcal{B}}$对称.

(2) f反对称$\Longleftrightarrow [f]_{\mathcal{B}}$反对称.

(3) f交错$\Longleftrightarrow [f]_{\mathcal{B}}$反对称并且对角元为0(这种矩阵称为是**交错的**).

证明. "$\Longrightarrow$"对于三种情况都是显然的. 下面证明"$\Longleftarrow$". 记$A = [f]_{\mathcal{B}}$.

(1) 假设A对称. 则
$$f(\alpha,\beta)=[\alpha]_{\mathcal{B}}^tA[\beta]_{\mathcal{B}}=([\alpha]_{\mathcal{B}}^tA[\beta]_{\mathcal{B}})^t=[\beta]_{\mathcal{B}}^tA[\alpha]_{\mathcal{B}}=f(\beta,\alpha).$$
(2) 假设A反对称. 则
$$f(\alpha,\beta)=[\alpha]_{\mathcal{B}}^tA[\beta]_{\mathcal{B}}=([\alpha]_{\mathcal{B}}^tA[\beta]_{\mathcal{B}})^t=-[\beta]_{\mathcal{B}}^tA[\alpha]_{\mathcal{B}}=-f(\beta,\alpha).$$
(3) 假设A反对称并且对角元为0. 记$[\alpha]_{\mathcal{B}}=X=(x_1,\ldots,x_n)^t$. 则
$$f(\alpha,\alpha)=X^tAX=\sum_{i=1}^n A_{ii}x_i^2+\sum_{1\le i<j\le n}(A_{ij}+A_{ji})x_ix_j=0.$$
□

定义 10.3. 设$q:V\to F$. 如果存在$f\in(V^*)^{\otimes 2}$满足$q(\alpha)=f(\alpha,\alpha)$, 则称$q$是$V$上的**二次型**.

注意$F^{n\times 1}$上的二次型就是二次齐次多项式函数.

记V上所有二次型的集合为$Q(V)$. 它在明显的运算下构成线性空间. 我们有满线性映射$\Phi:(V^*)^{\otimes 2}\to Q(V)$, $\Phi(f)(\alpha)=f(\alpha,\alpha)$.

命题 10.7. 设$\operatorname{char}F\ne 2$.

(1) $\Phi|_{S^2(V^*)}:S^2(V^*)\to Q(V)$是线性同构.

(2) (**极化恒等式**) 设$q\in Q(V)$. 如果$f\in S^2(V^*)$满足$q(\alpha)=f(\alpha,\alpha)$, 则
$$f(\alpha,\beta)=\frac{1}{4}(q(\alpha+\beta)-q(\alpha-\beta)). \tag{10.1}$$

证明. (1) 显然有$\operatorname{Ker}(\Phi)=\Lambda^2(V^*)$. 只需注意到$\operatorname{char}F\ne 2$时有$(V^*)^{\otimes 2}=S^2(V^*)\oplus\Lambda^2(V^*)$.

(2) 我们有
$$q(\alpha+\beta)=f(\alpha+\beta,\alpha+\beta)=q(\alpha)+q(\beta)+2f(\alpha,\beta),$$
$$q(\alpha-\beta)=f(\alpha-\beta,\alpha-\beta)=q(\alpha)+q(\beta)-2f(\alpha,\beta).$$
由此即得(10.1). □

这说明, 对任意$q\in Q(V)$, 则存在唯一的$f\in S^2(V^*)$满足$\Phi(f)=q$, f的表达式由(10.1)给出. 对于V的有序基$\mathcal{B}$, f在$\mathcal{B}$下的矩阵称为**q在$\mathcal{B}$下的矩阵**, 记为$[q]_{\mathcal{B}}$. 注意$[q]_{\mathcal{B}}$对称, 并且
$$q(\alpha)=[\alpha]_{\mathcal{B}}^t[q]_{\mathcal{B}}[\alpha]_{\mathcal{B}}.$$

设$f\in(V^*)^{\otimes 2}$对称或反对称. 我们希望寻找V的有序基$\mathcal{B}$使得$[f]_{\mathcal{B}}$充分简单.

定理 10.8. (1) 假设$\operatorname{char}F\ne 2$. 如果$f\in S^2(V^*)$, 则存在有序基$\mathcal{B}$使得$[f]_{\mathcal{B}}$对角.

(2) 如果$f\in\Lambda^2(V^*)$, 则存在有序基$\mathcal{B}$使得
$$[f]_{\mathcal{B}}=\operatorname{diag}\left(\begin{bmatrix}0&1\\-1&0\end{bmatrix},\ldots,\begin{bmatrix}0&1\\-1&0\end{bmatrix},0,\ldots,0\right).$$

注 10.2. 如果存在有序基$\mathcal{B}$使得$[f]_{\mathcal{B}}$对角, 则$[f]_{\mathcal{B}}$对称, 从而f对称.

先证明:

引理 10.9. 设$f\in(V^*)^{\otimes 2}$对称或反对称, $W\subset V$是子空间,
$$W^{\perp}:=\{\beta\in V\mid f(\alpha,\beta)=0,\forall\alpha\in W\}.$$
假设$f|_W$非退化. 则$V=W\oplus W^{\perp}$. 并且如果$\mathcal{B}_1$, $\mathcal{B}_2$分别是W和$W^{\perp}$的有序基, $\mathcal{B}=(\mathcal{B}_1,\mathcal{B}_2)$, 则
$$[f]_{\mathcal{B}}=\operatorname{diag}([f|_W]_{\mathcal{B}_1},[f|_{W^{\perp}}]_{\mathcal{B}_2}).$$

证明. 我们有

$$W^{\perp}=\{\beta\in V\mid L_f(\alpha)(\beta)=0,\forall\alpha\in W\}=\bigcap_{\alpha\in W}\mathrm{Ker}(L_f(\alpha))=L_f(W)^0.$$

因此

$$\dim W^{\perp}=\dim L_f(W)^0=\dim V-\dim L_f(W)\geq\dim V-\dim W.$$

另一方面, $f|_W$非退化$\Longrightarrow W\cap W^{\perp}=\{0\}$. 两者结合起来, 得$V=W\oplus W^{\perp}$. 为了得到$[f]_{\mathcal{B}}$的表达式, 只需说明对任意$\alpha\in\mathcal{B}_1$和$\beta\in\mathcal{B}_2$有$f(\alpha,\beta)=f(\beta,\alpha)=0$. $f(\alpha,\beta)=0$由$W^{\perp}$的定义是显然的. 由此和f的对称性或反对称性, 即得$f(\beta,\alpha)=0$. □

定理10.8的证明. 对$\dim V$用归纳法. $\dim V=1$时显然. 假设$\dim V=n>1$, 并且定理当$\dim V<n$成立. 不妨设$f\neq 0$. 下面对两种情况分别证明.

(1) 由于$\mathrm{char}\,F\neq 2$, 由命题10.7(1), $\Phi(f)$非零, 即存在$\alpha\in V$满足$f(\alpha,\alpha)\neq 0$. 设$W=F\alpha$. 则$f|_W$非退化. 由引理10.9, $\dim W^{\perp}=n-1$. 从而由归纳假设, 存在$W^{\perp}$的有序基$\mathcal{B}_2$使$[f|_{W^{\perp}}]_{\mathcal{B}_2}$对角. 取$W$的有序基$\mathcal{B}_1=\{\alpha\}$和$V$的有序基$\mathcal{B}=(\mathcal{B}_1,\mathcal{B}_2)$, 则$[f]_{\mathcal{B}}=\mathrm{diag}([f|_W]_{\mathcal{B}_1},[f|_{W^{\perp}}]_{\mathcal{B}_2})$对角.

(2) 由$f\neq 0$推出存在线性无关的$\alpha,\beta\in V$满足$f(\alpha,\beta)=1$. 取$W=\mathrm{span}\{\alpha,\beta\}$, $\mathcal{B}_1=\{\alpha,\beta\}$. 则$[f|_W]_{\mathcal{B}_1}=\begin{bmatrix}0&1\\-1&0\end{bmatrix}$. 从而$f|_W$非退化. 由引理10.9, $\dim W^{\perp}=n-2$. 从而由归纳假设, 存在$W^{\perp}$的有序基$\mathcal{B}_2$使$[f|_{W^{\perp}}]_{\mathcal{B}_2}$为定理中的形式. 设$\mathcal{B}=(\mathcal{B}_1,\mathcal{B}_2)$. 则$[f]_{\mathcal{B}}=\mathrm{diag}([f|_W]_{\mathcal{B}_1},[f|_{W^{\perp}}]_{\mathcal{B}_2})$也为定理中的形式. □

非退化的交错双线性形式称为**辛形式**.

推论 10.10. 设V上存在辛形式. 则$\dim V$为偶数, 并且存在有序基$\mathcal{B}$使得

$$[f]_{\mathcal{B}}=\begin{bmatrix}0&I_m\\-I_m&0\end{bmatrix}.$$

证明. 对定理10.8(2)中的有序基重新排序即可. □

推论 10.11. 设$\mathrm{char}\,F\neq 2$, $A\in F^{n\times n}$.

(1) 如果A对称, 则A合同于对角矩阵.

(2) 如果A反对称, 则A合同于形如

$$\mathrm{diag}\left(\begin{bmatrix}0&1\\-1&0\end{bmatrix},\ldots,\begin{bmatrix}0&1\\-1&0\end{bmatrix},0,\ldots,0\right)$$

的矩阵.

注 10.3. 当$\mathrm{char}\,F=2$时, (1)不成立. 例如此时$\begin{bmatrix}0&1\\1&0\end{bmatrix}$不合同于对角矩阵. 而此时, (2)对反对称并且对角元为0的矩阵成立.

注 10.4. 一般矩阵(既不对称也不反对称)的合同标准形问题比较复杂. 见Horn & Johnson.

接下来我们对特殊的域F进一步讨论对称双线性形式的矩阵.

定理 10.12. 设$f\in S^2(V^*)$.

(1) 如果F是代数闭域并且$\mathrm{char}\,F\neq 2$(例如$F=\mathbb{C}$), 则存在V的有序基$\mathcal{B}$使得$[f]_{\mathcal{B}}=\mathrm{diag}(I_r,0)$, 其中$r=\mathrm{rank}(f)$.

(2) 如果$F=\mathbb{R}$, 则存在V的有序基$\mathcal{B}$使得$[f]_{\mathcal{B}}=\mathrm{diag}(I_{r_1},-I_{r_2},0)$, 并且$r_1,r_2$被$f$决定.

证明. (1) 由于$\operatorname{char} F \neq 2$, 由定理10.8(1), 存在有序基$\mathcal{B}_0 = \{\alpha_1, \ldots, \alpha_n\}$使得

$$[f]_{\mathcal{B}_0} = \operatorname{diag}(f(\alpha_1, \alpha_1), \ldots, f(\alpha_n, \alpha_n)).$$

对$\alpha_1, \ldots, \alpha_n$重新排序, 不妨设$f(\alpha_i, \alpha_i) \neq 0 \iff 1 \leq i \leq r$. 由于$F$是代数闭域, 当$1 \leq i \leq r$时, 可以取$c_i \in F$满足$c_i^2 = f(\alpha_i, \alpha_i)$. 设$\mathcal{B} = \{c_1^{-1}\alpha_1, \ldots, c_r^{-1}\alpha_r, \alpha_{r+1}, \ldots, \alpha_n\}$. 则$[f]_{\mathcal{B}} = \operatorname{diag}(I_r, 0)$.

(2) 与(1)的证明类似, 存在有序基$\mathcal{B}_0 = \{\alpha_1, \ldots, \alpha_n\}$使得

$$[f]_{\mathcal{B}_0} = \operatorname{diag}(f(\alpha_1, \alpha_1), \ldots, f(\alpha_r, \alpha_r), 0, \ldots, 0),$$

并且对某个$0 \leq r_1 \leq r$有

$$1 \leq i \leq r_1 \implies f(\alpha_i, \alpha_i) > 0,$$
$$r_1 + 1 \leq i \leq r \implies f(\alpha_i, \alpha_i) < 0.$$

当$1 \leq i \leq r$时, 定义$c_i = \sqrt{|f(\alpha_i, \alpha_i)|}$. 设$\mathcal{B} = \{c_1^{-1}\alpha_1, \ldots, c_r^{-1}\alpha_r, \alpha_{r+1}, \ldots, \alpha_n\}$. 则$[f]_{\mathcal{B}} = \operatorname{diag}(I_{r_1}, -I_{r_2}, 0)$, 其中$r_2 = r - r_1$.

现在设V的有序基$\mathcal{B} = \{\alpha_1, \ldots, \alpha_n\}$满足$[f]_{\mathcal{B}} = \operatorname{diag}(I_{r_1}, -I_{r_2}, 0)$. 我们证明$r_1, r_2$被$f$决定. 取使得$f|_W$正定的维数最大的子空间$W \subset V$. 记$V_1 = \operatorname{span}\{\alpha_1, \ldots, \alpha_{r_1}\}$, $V_2 = \operatorname{span}\{\alpha_{r_1+1}, \ldots, \alpha_n\}$. 注意到:

$$f|_{V_1}\text{关于}V_1\text{的基}\{\alpha_1, \ldots, \alpha_{r_1}\}\text{的矩阵为}I_{r_1} \implies f|_{V_1}\text{正定} \implies r_1 = \dim V_1 \leq \dim W,$$

$$\begin{aligned} f|_{V_2}\text{关于}V_2\text{的基}\{\alpha_{r_1+1}, \ldots, \alpha_n\}\text{的矩阵为}\operatorname{diag}(-I_{r_2}, 0) &\implies f|_{V_2}\text{半负定} \\ &\implies V_2 \cap W = \{0\} \\ &\implies r_1 = n - \dim V_2 \geq \dim W. \end{aligned}$$

因此$r_1 = \dim W$被f决定. 类似可证明r_2被f决定. □

定理10.12(2)称为Sylvester惯性定理(Sylvester's law of inertia). 这里"惯性"一词指r_1, r_2被f决定, 与物理中的"惯性"没有直接关系. r_1 与r_2分别称为f的**正惯性指数**(positive index of inertia)和**负惯性指数**(negative index of inertia). 数对(r_1, r_2)称为f的**符号**(signature)[1]. 可以证明, 对任何V的有序基$\mathcal{B}$, r_1(或r_2)等于对称矩阵$[f]_{\mathcal{B}}$的正(或负)特征值的重数之和. 注意这些定义也对f诱导的二次型适用.

注 10.5. 由主轴定理, 定理10.12(2)中的$\mathcal{B}$可以取为(关于任何内积的)正交基.

推论 10.13. 设$A \in F^{n \times n}$对称.

(1) 如果$F = \mathbb{C}$, 则A合同于$\operatorname{diag}(I_r, 0)$. 特别地, 任何$GL_n(\mathbb{C})$中的对称矩阵形如$P^tP$.

(2) 如果$F = \mathbb{R}$, 则A合同于$\operatorname{diag}(I_{r_1}, -I_{r_2}, 0)$, 其中$r_1, r_2$被$A$决定.

注 10.6. • 对于推论10.13(2), 由正交相似的结果, 存在P为正交矩阵乘正对角矩阵, 使得$P^tAP = \operatorname{diag}(I_{r_1}, -I_{r_2}, 0)$.

§10.3 双线性形式的自同构群

设V是n维F-线性空间, $f \in (V^*)^{\otimes 2}$. 记

$$\operatorname{Aut}(V, f) = \{T \in \operatorname{GL}(V) \mid f(T\alpha, T\beta) = f(\alpha, \beta), \forall \alpha, \beta \in V\}.$$

显然$\operatorname{Aut}(V, f)$是$\operatorname{GL}(V)$的子群, 称为f或(V, f)的**自同构群**. 对于矩阵的情况, 设$A \in F^{n \times n}$. 容易

[1]在有的文献中, f的符号定义为$r_1 - r_2$.

验证,

$$G_A := \{M \in \mathrm{GL}_n(F) \mid M^t A M = A\}$$

是$\mathrm{GL}_n(F)$的子群.

引理 10.14. 设$f \in (V^*)^{\otimes 2}$, $\mathcal{B}$是V的有序基, $A = [f]_{\mathcal{B}}$. 则对于$T \in \mathrm{GL}(V)$, 有$T \in \mathrm{Aut}(V, f) \iff [T]_{\mathcal{B}} \in G_A$. 特别地, 有群同构$\mathrm{Aut}(V, f) \cong G_A$.

证明. 记$M = [T]_{\mathcal{B}}$. 由于$f(\alpha, \beta) = [\alpha]_{\mathcal{B}}^t A[\beta]_{\mathcal{B}}$, $[T\alpha]_{\mathcal{B}} = M[\alpha]_{\mathcal{B}}$, 所以

$$\begin{aligned}
T \in \mathrm{Aut}(V, f) &\iff f(T\alpha, T\beta) = f(\alpha, \beta), \forall \alpha, \beta \in V \\
&\iff [T\alpha]_{\mathcal{B}}^t A[T\beta]_{\mathcal{B}} = [\alpha]_{\mathcal{B}}^t A[\beta]_{\mathcal{B}}, \forall \alpha, \beta \in V \\
&\iff [\alpha]_{\mathcal{B}}^t M^t A M[\beta]_{\mathcal{B}} = [\alpha]_{\mathcal{B}}^t A[\beta]_{\mathcal{B}}, \forall \alpha, \beta \in V \\
&\iff M^t A M = A \\
&\iff M \in G_A.
\end{aligned}$$

□

例 10.1. 假设f对称非退化. 记$\mathrm{O}(V, f) := \mathrm{Aut}(V, f)$, 称为$f$或$(V, f)$的**正交群**. 如果$\mathrm{char}\, F \neq 2$, 则$T$是$f$的自同构$\iff T$是$f$诱导的二次型的自同构. 记

$$\mathrm{O}_n(F) := G_I = \{M \in \mathrm{GL}_n(F) \mid M^t M = I\},$$

称为域F上的n**阶正交群**. 如果F是代数闭域并且$\mathrm{char}\, F \neq 2$(例如$F = \mathbb{C}$), 则存在$\mathcal{B}$满足$[f]_{\mathcal{B}} = I$. 此时$\mathrm{Aut}(V, f) \cong \mathrm{O}_n(F)$. 如果$F = \mathbb{R}$, 则存在$\mathcal{B}$满足$[f]_{\mathcal{B}} = I_{p,q} := \mathrm{diag}(I_p, -I_q)$, 并且符号$(p, q)$被$f$决定. 此时$\mathrm{Aut}(V, f)$同构于

$$\mathrm{O}(p, q) := G_{I_{p,q}} = \{M \in \mathrm{GL}_n(\mathbb{R}) \mid M^t I_{p,q} M = I_{p,q}\}.$$

容易看出, $\mathrm{O}(p, q) \cong \mathrm{O}(q, p)$. 当$p, q$都非零时, $\mathrm{O}(p, q)$称为**不定正交群**或**伪正交群**. $\mathrm{O}(3, 1)$称为**Lorentz群**. □

例 10.2. 假设f交错非退化. 记$\mathrm{Sp}(V, f) := \mathrm{Aut}(V, f)$, 称为$f$或$(V, f)$的**辛群**. 由于总存在$\mathcal{B}$满足$[f]_{\mathcal{B}} = J_{2m} := \begin{bmatrix} 0 & I_m \\ -I_m & 0 \end{bmatrix}$(其中$2m = n$), 所以$\mathrm{Aut}(V, f)$同构于域$F$上的$2m$**阶辛群**

$$\mathrm{Sp}_{2m}(F) := G_{J_{2m}} = \{M \in \mathrm{GL}_{2m}(F) \mid M^t J_{2m} M = J_{2m}\}.$$

□