抽象代数观点下的标准型理论

Eliauk

2024 年 7 月 14 日

目录

1 向量空间作为 $\mathbb{F}[x]$-模

令 V 是域 $\mathbb{F}$ 上的 n 维向量空间, $\varphi : V \to V$ 是线性变换. 对于 $v \in V$, 通过定义 $x \cdot v = \varphi(v)$, 使得 V 成为一个 $\mathbb{F}[x]$-模. 也就是说, 对于 $f(x) = a_m x^m + \cdots + a_1 x + a_0 \in \mathbb{F}[x]$, 我们定义

$$f(x) \cdot v = f(\varphi)(v) = a_m \varphi^m(v) + \cdots + a_1 \varphi(v) + a_0 v.$$

不难验证这确实满足模公理. 注意, V 的模结构是和线性变换 φ 有关的, 所以线性变换 φ 的一些相关概念可以在 V 的 $\mathbb{F}[x]$-模结构中有所体现. 从此开始, 后面提到的 V 都指代其相关于 φ 的模结构.

我们考虑 V 的零化子 $\mathrm{Ann}(V)$. 我们知道 V 上线性变换全体构成 n^2 维向量空间 $\mathrm{End}(V)$, 所以 $\mathbb{1}_V, \varphi, \ldots, \varphi^{n^2}$ 一定 $\mathbb{F}$-线性相关. 这就表明存在不全为零的 $a_1, \ldots, a_{n^2} \in \mathbb{F}$ 使得

$$a_{n^2} \varphi^{n^2} + \cdots + a_1 \varphi + a_0 \mathbb{1}_V = 0,$$

所以 $\mathrm{Ann}(V) \neq 0$. 又因为 $\mathbb{F}[x]$ 是 Euclid 整环, 进而是主理想整环, 故存在首一多项式 $m(x) \in \mathbb{F}[x]$ 使得 $\mathrm{Ann}(V) = (m(x))$. 这里的 $m(x)$ 被称为 φ 的**最小多项式**. 任意 $f(x) \in \mathrm{Ann}(V)$ 被称为 φ 的**零化多项式**. 由此可得: 对于任意零化多项式 $f(x)$, 一定有 $m(x) \mid f(x)$. 下面我们用模的语言重新证明 Cayley-Hamilton 定理.

定理 1.1 (Cayley–Hamilton 定理). 设 $\varphi : V \to V$ 是线性变换，那么特征多项式 $c(x) = \det(x\mathbb{1}_V - \varphi)$ 是 φ 的零化多项式，即 $c(x) \in \mathrm{Ann}(V)$，从而 $m(x) \mid c(x)$.

Proof. 令 $e_1, \dots, e_n$ 是 V 的一组基, 设 φ 在这组基下的表示矩阵为 $A = (a_{ij})$. 也就是说, $x \cdot e_j = \varphi(e_j) = \sum_i a_{ij} e_i$, 即 $\sum_i (x\delta_{ij} - a_{ij}) e_i = 0$, 其中 δ_{ij} 表示 Kronecker 符号. 考虑 $\mathbb{F}[x]$ 上的矩阵 $xI_n - A$, 设 $xI_n - A$ 的伴随矩阵为 B, 那么 $C = B(xI_n - A) = \det(xI_n - A) I_n$. C 的 (i, j)-元为

$$c_{ij} = \sum_{k=1}^{n} b_{ik}(x\delta_{jk} - a_{jk}) = \det(xI_n - A)\delta_{ij},$$

所以

$$\sum_{j=1}^{n} c_{ij} e_j = \sum_{j=1}^{m} \sum_{k=1}^{n} b_{ik}(x\delta_{jk} - a_{jk}) e_j = 0 = \det(xI_n - A) e_i,$$

所以对于任意的 $v \in V$, 都有 $\det(xI_n - A)v = 0$, 故 $c(x) = \det(xI_n - A) \in \mathrm{Ann}(V)$. □

2 Euclid 整环上的有限生成模

在本节我们暂时离开向量空间, 讨论 Euclid 整环上的有限生成模是如何发挥作用的.

引理 2.1. M 是有限生成 R-模当且仅当对于某个整数 n，存在满的 R-模同态 $f : R^n \to M$.

Proof. 若 M 是有限生成的, 设 $x_1, \dots, x_n$ 是一组生成元, 考虑映射 $f : R^n \to M$ 为

$$(a_1, \dots, a_n) \mapsto a_1 x_1 + \cdots + a_n x_n,$$

容易验证这是一个模同态, 并且由于 $x_1, \dots, x_n$ 是生成元, 所以 f 是满同态.

反过来, 记 $e_i = (0, \dots, 0, 1, 0, \dots, 0) \in R^n$, 其中第 i 个分量为 1. 那么任取 $m \in M$, 存在 $(a_1, \dots, a_n) \in R^n$ 使得 $f(a_1, \dots, a_n) = m$, 即

$$m = f(a_1, \dots, a_n) = f(a_1 e_1 + \cdots + a_n e_n) = a_1 f(e_1) + \cdots + a_n f(e_n),$$

故 M 由 $f(e_1), \dots, f(e_n)$ 生成. □

本节的主要目标是证明 Euclid 整环上的有限生成模的结构定理:

定理 2.2 (基本定理：不变因子型). R 是 Euclid 整环，M 是有限生成 R-模，那么 M 同构于有限多个循环模的直和，即

$$M \simeq R^r \oplus R/(a_1) \oplus R/(a_2) \oplus \cdots \oplus R/(a_m),$$

其中 $r \geq 0$，$a_1, a_2, \ldots, a_m$ 是 R 中非零非单位的元素，并且满足关系

$$a_1 \mid a_2 \mid \cdots \mid a_m.$$

此时 r 称为 M 的**自由秩**，$a_1, a_2, \ldots, a_m$ 称为 M 的**不变因子**.

实际上上述定理对于主理想整环上的有限生成模也成立, 但是主理想整环没有一个好的算法来寻找 $a_1, a_2, \ldots, a_m$, 我们将看到 Euclid 整环上的 Euclid 算法会发挥重要作用, 并且这也是线性代数中有理标准型和 Jordan 标准型理论的来源. 从现在开始, 本节所指的 R 均代表 Euclid 整环.

设 M 是有限生成 R-模, 根据引理 2.1, 我们知道存在满同态 $f : R^n \to M$, 即 $M \simeq R^n / \ker f$. 由于 R 是 Noether 环, 所以 R^n 是 Noether 模, 那么 R^n 和 $\ker f$ 都是有限生成 R-模, 设 $x_1, \ldots, x_n$ 是 R^n 的基（与向量空间中基的定义相同, 此时代表 $R^n = Rx_1 \oplus \cdots \oplus Rx_n$）, $y_1, \ldots, y_m$ 是 $\ker f$ 的生成元, 注意此时我们考虑生成元的顺序. 那么对于每个 $1 \leq i \leq m$, 都存在 $a_{i1}, \ldots, a_{in} \in R$, 使得

$$y_i = a_{i1}x_1 + \cdots + a_{in}x_n,$$

记矩阵（与线性代数中基的过渡矩阵的写法是一致的）

$$A = (a_{ji}) = \begin{pmatrix} a_{11} & a_{21} & \cdots & a_{m1} \\ a_{12} & a_{22} & \cdots & a_{m2} \\ \vdots & \vdots & \ddots & \vdots \\ a_{1n} & a_{2n} & \cdots & a_{mn} \end{pmatrix} \in M_{n,m}(R).$$

那么矩阵 A 表述了 R^n 的基和 $\ker f$ 的生成元的关系, 我们称其为 R^n 到 $\ker f$ 的**关系矩阵**.

下面我们观察对关系矩阵 A 做初等变换会有什么样的后果. 下面是初等列变换.

(1) 如果用 $u \in R^\times$ 乘以 A 的第 i 列, 那么相当于将 $\ker f$ 的生成元 y_i 变为 uy_i, 即给出了 $\ker f$ 的一组新的生成元 $y_1, \ldots, uy_i, \ldots, y_m$.

(2) 如果交换 A 的第 i 列和第 j 列, 那么相当于交换 $\ker f$ 的生成元 y_i 和 y_j, 即给出了 $\ker f$ 的一组新的生成元 $y_1, \ldots, y_j, \ldots, y_i, \ldots, y_m$.

(3) 如果将 A 的第 i 列乘以 $a \in R$ 加到第 j 列, 那么相当于将 $\ker f$ 的生成元 y_j 替换为 $y_j + ay_i$, 即给出了 $\ker f$ 的一组新的生成元 $y_1, \ldots, y_i, \ldots, y_j + ay_i, \ldots, y_m$.

然后是初等行变换.

(1) 如果用 $u \in R^\times$ 乘以 A 的第 i 行, 那么相当于将 R^n 的基 x_i 变为 $u^{-1}x_i$, 即给出了 R^n 的一组新的基 $x_1, \ldots, u^{-1}x_i, \ldots, x_n$.

(2) 如果交换 A 的第 i 行和第 j 行, 那么相当于交换 R^n 的基 x_i 和 x_j, 即给出了 R^n 的一组新的基 $x_1,\ldots,x_j,\ldots,x_i,\ldots,x_n$.
(3) 如果将 A 的第 i 行乘以 $a\in R$ 加到第 j 行, 那么相当于将 R^n 的基 x_i 替换为 x_i-ax_j, 即给出了 R^n 的一组新的基 $x_1,\ldots,x_i-ax_j,\ldots,x_j,\ldots,x_n$.

下面的定理告诉我们, 通过有限次初等变换, 总能将矩阵 A 变为对角阵, 这意味着 $x_1,\ldots,x_n$ 的某些倍数将成为 $\ker f$ 的生成元, 而基的倍数还是基 (这里使用了 R 是整环的性质), 这就表明 $\ker f$ 也可以写为某些循环子模的直和, 从而证明 定理 2.2.

定理 2.3 (Smith 标准型). 如果关系矩阵为零矩阵, 那么 $\ker f=0$, 从而 $M\simeq R^n$. 所以我们假设 $\ker f\neq 0$, 令 a_1 是初始关系矩阵 A 中所有元素的最大公因子. 那么通过上述六种初等变换, 可以将 A 变换为

$$\begin{pmatrix} D_k & \\ & O_{n-k,m-k}\end{pmatrix},$$

其中 D_k 是对角阵, 对角线为 $a_1,a_2,\ldots,a_k$ $(k\leq m)$, 并且满足 $a_1\mid a_2\mid\cdots\mid a_k$. 上述对角阵被称为 A 的 Smith 标准型.

Proof. I. 首先注意到初等行或者列变换并不会改变所有元素的最大公因子, 因为新矩阵的元素显然被原矩阵元素的最大公因子整除, 而初等变换的操作是可逆的, 所以原矩阵的元素又被新矩阵元素的最大公因子整除, 所以新矩阵和原矩阵元素的最大公因子相同.

我们先考虑对第一行的操作: 假设第一行的最大公因子是 $d=\gcd(a_{11},\ldots,a_{m1})$, 那么我们可以按照下面的方法将 d 放到第一行第一列:

- 首先仅涉及前两列, 使用辗转相除法, 将 $\gcd(a_{11},a_{21})$ 放到第一行第一列;
- 然后仅涉及第一列和第三列, 使用辗转相除法, 将 $\gcd(\gcd(a_{11},a_{21}),a_{31})=\gcd(a_{11},a_{21},a_{31})$ 放到第一行第一列;
- 重复上述步骤, 直到把 d 放到第一行第一列.

上述操作可以导出下面的算法:

(1) 使用上述操作将 $d=\gcd(a_{11},\ldots,a_{m1})$ 放到第一行第一列;
(2) 此时 d 整除 $a_{21},\ldots,a_{m1}$, 那么我们可以将第一行除开 d 之外全变为零;
(3) 如果此时矩阵的所有元素都被 d 整除, 那么操作停止. 否则, 存在元素 a_{ji} $(i>1)$, 使得 $d\nmid a_{ji}$;
(4) 如果 $d\nmid a_{1i}$, 那么我们跳过这一步. 否则, 我们将第 j 列加到第 1 列使得 $a\nmid a_{1i}$, 注意此时并没有改变第一行;
(5) 计算新的最大公因子 $d'=\gcd(d,a_{1i})$, 通过仅涉及第 1 行和第 i 行的操作, 可以将 d' 放到第一行第一列, 注意到通过第四步, 有 $d\nmid a_{1i}$, 所以 d' 的次数严格小于 d 和 a_{1i} 的次数;
(6) 回到操作 1.

这个算法必然在有限步结束, 因为每进行一次第 (5) 步, 第一行第一列的元素的次数就会严格变小, 最多也只能变到零次, 从而在第 (3) 步退出操作. 最后我们就得到了 a_{11} 整除所有其他元素的新矩阵, 而根据前面的观察, 这个新矩阵元素的最大公因子等于初始矩阵元素的最大公因子, 而新矩阵元素的最大公因子显然就是 a_{11}, 这就完成了第 I. 步.

II. 在 I. 的基础上, 使用初等变换将第一行第一列的非 $(1,1)$-元全变成零即可.

III. 在 II. 的基础上, 右下角的分块矩阵又回到了 I. 的形式, 注意到对该分块矩阵的操作不会影响大矩阵的第一行第一列, 所以重复 I. 的操作, 又可以将这个分块矩阵化为 II. 的形式, 即使得 $a_1 \mid a_2$ 并且 a_2 整除其他所有元素.

IV. 重复前三步的操作, 便可以得到形如 $\begin{pmatrix} D_k & 0 \\ 0 & O_{n-k,m-k} \end{pmatrix}$ 的关系矩阵. □

Proof of 定理 2.2. 根据前面的叙述, $M \simeq R^n/\ker f$, 通过初等变换将关系矩阵变为 Smith 标准型后, 此时存在 R^n 的一组基 $x_1,\dots,x_n$ 和满足 $a_1 \mid \cdots \mid a_k$ 的元素 $a_1,\dots,a_k \in R$, 使得 $a_1x_1,\dots,a_kx_k$ 是 $\ker f$ 的生成元 (从而是 $\ker f$ 的基), 那么就有

$$\begin{aligned} M &\simeq R^n/\ker f = (Rx_1 \oplus \cdots \oplus Rx_n)/(Ra_1x_1 \oplus \cdots \oplus Ra_kx_k \oplus 0 \oplus \cdots \oplus 0) \\ &\simeq R/(a_1) \oplus R/(a_2) \oplus \cdots \oplus R/(a_k) \oplus R^{n-k}. \end{aligned}$$ □

定理 2.2 的结果还可以进一步改进, 设 a 是 R 中的非零非单位的元素, 那么 a 可以唯一分解为 $a = p_1^{\alpha_1}\cdots p_s^{\alpha_s}$, 其中 p_i 为互不相同的素元, 那么根据中国剩余定理, 有环同构 (同时也是 R-模同构)

$$R/(a) \simeq R/(p_1^{\alpha_1}) \oplus R/(p_2^{\alpha_2}) \oplus \cdots \oplus R/(p_s^{\alpha_s}).$$

定理 2.2 的右端每一项都可以做类似的分解, 于是我们得到了下面定理.

定理 2.4 (基本定理: 初等因子型). R 是 Euclid 整环, M 是有限生成 R-模, 那么 M 同构于有限多个循环模的直和, 即

$$M \simeq R^r \oplus R/(p_1^{\alpha_1}) \oplus R/(p_2^{\alpha_2}) \oplus \cdots \oplus R/(p_t^{\alpha_t}),$$

其中 $r \geq 0$, $p_1^{\alpha_1}, p_2^{\alpha_2}, \dots, p_t^{\alpha_t}$ 是 R 中素元的幂次 (不需要互不相同). $p_1^{\alpha_1}, p_2^{\alpha_2}, \dots, p_t^{\alpha_t}$ 被称为 M 的**初等因子**.

虽然我们已经得到了 Euclid 整环上有限生成模的两种分解, 但是还有一个问题, 这两种分解是否唯一? 答案是肯定的, 但是证明略为繁琐, 见参考文献 [1].

3 有理标准型和 Jordan 标准型

从名字就可以看出, Euclid 整环上有限生成模的不变因子分解和初等因子分解与线性代数中的标准型理论密切相关, 事实上也确实如此. 下面我们回到对向量空间的研究.

给定 n 维 $\mathbb{F}$-向量空间 V, 设有一组基 $e_1, \ldots, e_n$, φ 是 V 上的线性变换, 那么 V 可以成为一个 $\mathbb{F}[x]$-模, 通过定义 $x \cdot v = \varphi(v)$. V 当然是一个有限生成 $\mathbb{F}[x]$-模, 而 $\mathbb{F}[x]$ 为 Euclid 整环, 这允许我们使用前一节的所有结果.

根据 定理 2.2, 我们知道存在多项式 $a_1(x), \ldots, a_m(x) \in \mathbb{F}[x]$, 满足

$$a_1(x) \mid \cdots \mid a_m(x),$$

使得我们有 $\mathbb{F}[x]$-模同构（自然也是 $\mathbb{F}$-向量空间同构）

$$\sigma : V \simeq \mathbb{F}[x]/(a_1(x)) \oplus \cdots \oplus \mathbb{F}[x]/(a_m(x)). \tag{1}$$

注意, 由于 V 上的任意线性变换都存在非零的最小多项式, 所以 V 作为 $\mathbb{F}[x]$-模是扭模, 故自由秩部分为零. 观察 (1) 式的右端, 容易发现

$$\mathrm{Ann}(V) = \mathrm{Ann}\big(\mathbb{F}[x]/(a_1(x)) \oplus \cdots \oplus \mathbb{F}[x]/(a_m(x))\big) = (a_m(x)).$$

此时我们进一步规定不变因子 $a_1, \ldots, a_m$ 都是首一多项式, 那么根据不变因子分解的唯一性, 我们便得到了下面的命题.

命题 3.1. 线性变换 φ 的最小多项式 $m(x)$ 就是 V 作为 $\mathbb{F}[x]$-模的最大的不变因子，并且 V 的所有不变因子都整除 $m(x)$.

我们继续发掘 (1) 式带给我们的信息. 注意到 (1) 式也是向量空间同构, 所以这意味着我们通过 V 的 $\mathbb{F}[x]$-模结构给出了一种新的向量空间的直和分解, 那么我们自然会寻找 $\mathbb{F}[x]/(a_i(x))$ 的一组基, 然后将其拼成 V 的一组基, 观察 φ 在这组基下的表示矩阵是什么形态, 实际上, 这正是 φ 的有理标准型.

对于线性变换 $\varphi : V \to V$, $\sigma \circ \varphi \circ \sigma^{-1}$ 是向量空间 $\mathbb{F}[x]/(a_1(x)) \oplus \cdots \oplus \mathbb{F}[x]/(a_m(x))$ 上的线性变换. 对于任意 $f(\bar{x}) \in \mathbb{F}[x]/(a_i(x))$, 那么

$$\sigma \circ \varphi \circ \sigma^{-1}\big(f(\bar{x})\big) = \sigma\big(x\sigma^{-1}\big(f(\bar{x})\big)\big) = x\sigma\big(\sigma^{-1}\big(f(\bar{x})\big)\big) = x \cdot f(\bar{x}) = \bar{x} f(\bar{x}).$$

这表明 φ 实际上对应着 $\mathbb{F}[x]/(a_i(x))$ 上的线性变换 $f(\bar{x}) \mapsto \bar{x} f(\bar{x})$, 我们只需要寻找这个线性变换在给定基下的表示矩阵即可.

对于首一多项式 $a(x) \in \mathbb{F}[x]$, 我们考察 $\mathbb{F}[x]$-模 $\mathbb{F}[x]/(a(x))$, 这自然是 $\mathbb{F}$-模, 即 $\mathbb{F}$-向量空间. 如果 $a(x) = x^k + a_{k-1}x^{k-1} + \cdots + a_0$, 读者可以自行验证 $\mathbb{F}[x]/(a(x))$ 有一组基 $\bar{1}, \bar{x}, \ldots, \bar{x}^{k-1}$（利用带余除法). 那么对于线性变换 $\mathbb{F}[x]/(a(x))$ 上的线性变换 $\psi : f(\bar{x}) \mapsto \bar{x} f(\bar{x})$, 我们有

$$\psi(\bar{1}) = \bar{x}, \psi(\bar{x}) = \bar{x}^2, \ldots, \psi(\bar{x}^{k-1}) = \bar{x}^k = -a_0 - a_1\bar{x} - \cdots - a_{k-1}\bar{x}^{k-1},$$

故 ψ 在基 $\bar{1}, \bar{x}, \ldots, \bar{x}^{k-1}$ 下的表示矩阵为

$$\mathcal{C}_{a(x)} = \begin{pmatrix} 0 & 0 & \cdots & 0 & -a_0 \\ 1 & 0 & \cdots & 0 & -a_1 \\ 0 & 1 & \cdots & 0 & -a_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & -a_{k-1} \end{pmatrix}.$$

上述矩阵 $\mathcal{C}_{a(x)}$ 被称为首一多项式 $a(x)$ 的**友阵**, 不难证明 $\mathcal{C}_{a(x)}$ 的特征多项式恰好就是 $a(x)$.

综合上面的叙述, 我们得到如下重要的定理:

定理 3.2 (线性变换的有理标准型). V 是域 $\mathbb{F}$ 上的有限维向量空间，$\varphi : V \to V$ 是线性变换，那么存在 V 的一组基和不变因子 $a_1(x), \ldots, a_m(x) \in \mathbb{F}[x]$，使得 φ 在这组基下的表示矩阵是有 m 块的分块对角阵，且每个分块是首一多项式 $a_i(x)$ 的友阵. 这个分块对角阵我们称之为 φ 的**有理标准型**. 此外，根据不变因子的唯一性，线性变换的有理标准型也是唯一的.

“有理”的含义为对于任意域 $\mathbb{F}$, 这样的标准型总是可以计算的, 我们将马上看到 Jordan 标准型丢失了这种性质. 有理标准型的重要性在于其完全解决了线性变换（矩阵）在相似意义下的分类问题, 即下面的定理.

定理 3.3. 设 φ, ψ 是 V 上的线性变换，那么:

(1) φ 和 ψ 相似;

(2) 通过 φ 得到的 $\mathbb{F}[x]$–模 V_1 和通过 ψ 得到的 $\mathbb{F}[x]$–模 V_2 是同构的;

(3) φ 和 ψ 的有理标准型相同.

Proof. (1) $\Rightarrow$ (2) 若 φ 和 ψ 相似, 即存在同构 σ 使得 $\varphi = \sigma \circ \psi \circ \sigma^{-1}$. 对于 $\sigma : V_2 \to V_1$, 我们说明 σ 是 $\mathbb{F}[x]$-模同构即可. 只需验证 σ 是模同态, 任取 $v \in V$, 我们有

$$\sigma(xv) = \sigma(\psi(v)) = \varphi(\sigma(v)) = x\sigma(v),$$

所以 σ 是 $\mathbb{F}[x]$-模同态.

(2) $\Rightarrow$ (3) 根据不变因子的唯一性即得.

(3) $\Rightarrow$ (1) 显然. □

由于 $a(x)$ 的友阵的特征多项式就是 $a(x)$, φ 的有理标准型又是友阵构成的分块对角阵, 所以 φ 的特征多项式就是这些友阵的特征多项式之积, 进而是 φ 的不变因子之积, 而 φ 的最小多项式是最大的不变因子, 所以从这一点我们可以再次推出 Cayley-Hamilton 定理. 此外, 由于不变因子的整除关系, 所以 φ 的最小多项式和特征多项式的根一定相同（不考虑重数）. 总结起来, 我们便得到了下面的命题.

命题 3.4. φ 是向量空间 V 上的线性变换，那么:

(1) φ 的特征多项式是 φ 的不变因子之积.
(2) φ 的特征多项式整除极小多项式的某个幂次. 特别地, A 的极小多项式和特征多项式有相同的根 (不考虑重数).

由前面的讨论可知, 如果 φ 的特征多项式在中 $\mathbb{F}[x]$ 能完全分裂为一次多项式, 即域 $\mathbb{F}$ 包含 φ 的所有特征值, 那么其不变因子便能全部分解为一次多项式的乘积, 这就方便我们使用 定理 2.4. 为了方便起见, 我们直接假设 $\mathbb{F}$ 是代数闭域, 即 $\mathbb{F}[x]$ 中的不可约多项式只有一次多项式. 我们应用 定理 2.4, 此时 V 的初等因子只可能形如 $(x-\lambda)^k$, 于是 定理 2.4 告诉我们: 存在 $k_1, \ldots, k_t$ 和 $\lambda_1, \ldots, \lambda_t \in \mathbb{F}$ (其中 $\lambda_1, \ldots, \lambda_t$ 包含了 φ 的所有特征值, 但允许重复出现), 使得

$$\tau : V \simeq \mathbb{F}[x]/(x-\lambda_1)^{k_1} \oplus \cdots \oplus \mathbb{F}[x]/(x-\lambda_t)^{k_t}. \tag{2}$$

类似有理标准型的做法, 我们需要寻找 $\mathbb{F}[x]/(x-\lambda)^k$ 上线性变换 $\psi : f(\bar{x}) \mapsto \bar{x} f(\bar{x})$ 在给定基下的表示矩阵. 此时我们取基为 $(\bar{x}-\lambda)^{k-1}, \ldots, \bar{x}-\lambda, 1$, 那么

$$\begin{aligned}
\psi(\bar{x}-\lambda)^{k-1} &= (\bar{x}-\lambda+\lambda)(\bar{x}-\lambda)^{k-1} = \lambda(\bar{x}-\lambda)^{k-1}, \\
\psi(\bar{x}-\lambda)^{k-2} &= (\bar{x}-\lambda+\lambda)(\bar{x}-\lambda)^{k-2} = \lambda(\bar{x}-\lambda)^{k-2} + (\bar{x}-\lambda)^{k-1}, \\
&\cdots \\
\psi(1) &= \bar{x} = \lambda \cdot 1 + (\bar{x}-\lambda),
\end{aligned}$$

也就是说, ψ 在基 $(\bar{x}-\lambda)^{k-1}, \ldots, \bar{x}-\lambda, 1$ 下的表示矩阵为

$$J_k = \begin{pmatrix} \lambda & 1 & & & \\ & \lambda & \ddots & & \\ & & \ddots & 1 & \\ & & & \lambda & 1 \\ & & & & \lambda \end{pmatrix}.$$

上述矩阵被称为 k 阶 **Jordan 块**. 于是我们得到了下面的定理.

定理 3.5 (线性变换的 Jordan 标准型). V 是域 $\mathbb{F}$ 上的有限维向量空间, $\varphi : V \to V$ 是线性变换, 如果 $\mathbb{F}$ 包含 φ 的所有特征值, 那么存在 V 的一组基和初等因子 $(x-\lambda_1)^{k_1}, \ldots, (x-\lambda_t)^{k_t} \in \mathbb{F}[x]$, 使得 φ 在这组基下的表示矩阵是有 t 块的分块对角阵, 且每个分块是 k_t 阶 Jordan 块. 这个分块对角阵我们称之为 φ 的 **Jordan 标准型**. 此外, 根据初等因子的唯一性, 线性变换的 Jordan 标准型也是唯一的.

借助 定理 3.3, 我们知道线性变换 φ, ψ 相似当且仅当它们有相同的有理标准型, 而 Jordan 标准型由有理标准型唯一确定, 所以 φ, ψ 相似当且仅当它们有相同的 Jordan 标准型.

4 标准型的计算

在线性代数中，我们知道计算线性变换 $\varphi : V \to V$ 的有理标准型或者 Jordan 标准型的一般步骤是对特征矩阵 $xI_n - A$ 做初等变换，变换为 Smith 标准型后对角线便是所有的不变因子，现在我们来揭示线性变换的特征矩阵和不变因子之间的关系. 取 $\mathbb{F}[x]^n$ 的一组基为 $\xi_1, \dots, \xi_n$，V 的一组基 $e_1, \dots, e_n$. 定义 $f : \mathbb{F}[x]^n \to V$ 满足 $f(\xi_i) = e_i$，那么 f 是满同态. 回顾 定理 2.3，我们说明 φ 的特征矩阵实际上就是一个关系矩阵！

命题 4.1. 如果 φ 在上述基 $e_1, \dots, e_n$ 下的表示矩阵为 A，特征矩阵

$$xI_n - A = \begin{pmatrix} x - a_{11} & -a_{12} & \cdots & -a_{1n} \\ -a_{21} & x - a_{22} & \cdots & -a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ -a_{n1} & -a_{n2} & \cdots & x - a_{nn} \end{pmatrix},$$

那么 $xI_n - A$ 是 $\mathbb{F}[x]^n$ 到 $\ker f$ 的关系矩阵.

Proof. 也就是说我们要证明 $1 \le i \le n$ 时，

$$\mu_i = -a_{1i}\xi_1 - \cdots - a_{i-1,i}\xi_{i-1} + (x - a_{ii})\xi_i - a_{i+1,i}\xi_{i+1} - \cdots - a_{ni}\xi_n$$

组成了 $\ker f$ 的一组生成元.

直接验证可知

$$f(\mu_i) = -(a_{1i}e_1 + \cdots + a_{ni}e_n) + \varphi(e_i) = 0,$$

所以 $\mu_i \in \ker f$.

由于

$$x\xi_i = \mu_i + \sum_{j=1}^{n} a_{ji}\xi_j \in (\mathbb{F}[x]\mu_1 + \cdots \mathbb{F}[x]\mu_n) + (\mathbb{F}\xi_1 + \cdots + \mathbb{F}\xi_n),$$

所以

$$x^2\xi_i = x(x\xi_i) = x\mu_i + \sum_{j=1}^{n} a_{ji}(x\xi_j) \in (\mathbb{F}[x]\mu_1 + \cdots \mathbb{F}[x]\mu_n) + (\mathbb{F}\xi_1 + \cdots + \mathbb{F}\xi_n),$$

以此类推，可知任意的 $k \ge 0$ 有

$$x^k\xi_i \in (\mathbb{F}[x]\mu_1 + \cdots \mathbb{F}[x]\mu_n) + (\mathbb{F}\xi_1 + \cdots + \mathbb{F}\xi_n),$$

所以

$$\mathbb{F}[x]^n = \mathbb{F}[x]\xi_1 + \cdots + \mathbb{F}[x]\xi_n = (\mathbb{F}[x]\mu_1 + \cdots \mathbb{F}[x]\mu_n) + (\mathbb{F}\xi_1 + \cdots + \mathbb{F}\xi_n).$$

那么任取 $\sum_{i=1}^n g_i(x)\mu_i + \sum_{i=1}^n a_i\xi_i \in \ker f$, 就有

$$\sum_{i=1}^n a_i e_i = \varphi\left(\sum_{i=1}^n g_i(x)\mu_i + \sum_{i=1}^n a_i\xi_i\right) = 0,$$

故 $a_i = 0$, 这就证明了 $\mu_1, \ldots, \mu_n$ 是 $\ker f$ 的生成元. $\square$

至此, 结合定理 2.3, 便介绍完了模论在两大标准型中的应用以及标准型的计算方法. 下面我们用两个例子结束本文, 这两个例子足以体现一般情况下两种标准型包括变换矩阵的计算方法. 需要注意的是, 在计算 Jordan 标准型的时候, 需要读者熟悉中国剩余定理中的同构关系如何给出.

例 4.2. 给定 $\mathbb{R}$ 上的矩阵

$$A = \begin{pmatrix} 2 & -2 & 14 \\ 0 & 3 & -7 \\ 0 & 0 & 2 \end{pmatrix},$$

计算 A 的有理标准型并给出变换矩阵.

Solution. A 的特征矩阵为

$$xI_3 - A = \begin{pmatrix} x-2 & 2 & -14 \\ 0 & x-3 & 7 \\ 0 & 0 & x-2 \end{pmatrix},$$

任取 $\mathbb{R}[x]^3$ 的基 $[\xi_1, \xi_2, \xi_3]$, 下面我们对 $xI_3 - A$ 进行初等变换, 并记录基的变化. 我们用 R 指代行, C 指代列. 可以验证依次通过下面的初等变换:

$$C_1 \leftrightarrow C_2, R_2 + \frac{-(x-3)}{2}R_1 \mapsto R_2, C_2 + \frac{-(x-2)}{2}C_1 \mapsto C_2, C_3 + 7C_1 \mapsto C_3,$$
$$C_2 \leftrightarrow C_3, R_3 - \frac{1}{7}R_2 \mapsto R_3, C_3 + \frac{x-3}{14}C_2 \mapsto C_3, \frac{1}{2}C_1, \frac{1}{7}C_2, 14C_3,$$

便可以将 $xI_3 - A$ 变为 Smith 标准型

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & x-2 & 0 \\ 0 & 0 & x^2-5x+6 \end{pmatrix}.$$

我们仅关注初等行变换, 其相当于对基进行了变换:

$$[\xi_1, \xi_2, \xi_3] \to \left[\xi_1 + \frac{(x-3)}{2}\xi_2, \xi_2, \xi_3\right] \to \left[\xi_1 + \frac{(x-3)}{2}\xi_2, \xi_2 + \frac{1}{7}\xi_3, \xi_3\right],$$

令 $[\mu_1, \mu_2, \mu_3]$ 为上述变换后的基, 此时有同构

$$\begin{aligned} &\mathbb{R}[x]^3/\ker f \\ =\ &(\mathbb{R}[x]\mu_1 \oplus \mathbb{R}[x]\mu_2 \oplus \mathbb{R}[x]\mu_3)/(\mathbb{R}[x]\mu_1 \oplus \mathbb{R}[x](x-2)\mu_2 \oplus \mathbb{R}[x](x^2-5x+6)\mu_3) \\ \simeq\ &\mathbb{R}[x]/(x-2) \oplus \mathbb{R}[x]/(x^2-5x+6), \end{aligned}$$

同构映射为

$$h_1\mu_1 + h_2\mu_2 + h_3\mu_3 + \ker f \mapsto (h_2 + (x-2), h_3 + (x^2-5x+6)),$$

另一方面, 又有 $\mathbb{R}[x]^3/\ker f \simeq \mathbb{R}^3$, 取 $\mathbb{R}^3$ 的基 e_1, e_2, e_3, 此时同构映射为

$$h_1\xi_1 + h_2\xi_2 + h_3\xi_3 + \ker f \mapsto h_1e_1 + h_2e_2 + h_3e_3.$$

所以 $\mathbb{R}[x]/(x-2) \oplus \mathbb{R}[x]/(x^2-5x+6) \simeq \mathbb{R}^3$ 的同构映射满足

$$\begin{aligned}
&(1,0) \mapsto \mu_2 \mapsto e_2 + \frac{1}{7}e_3,\\
&(0,1) \mapsto \mu_3 \mapsto e_3,\\
&(0,\bar{x}) \mapsto x\mu_3 \mapsto xe_3 = 14e_1 - 7e_2 + 2e_3.
\end{aligned}$$

所以 A 在 $[e_2 + \frac{1}{7}e_3, e_3, 14e_1 - 7e_2 + 2e_3]$ 这组基下变换为有理标准型, 即

$$P = \begin{pmatrix} 0 & 0 & 14 \\ 1 & 0 & -7 \\ \frac{1}{7} & 1 & 2 \end{pmatrix}, \quad P^{-1}AP = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 0 & -6 \\ 0 & 1 & 5 \end{pmatrix}.$$ ■

例 4.3. 将上例中的矩阵 A 变换为 Jordan 标准型.

Solution. 初等因子为 $x-2, x-2, x-3$. 故我们有同构

$$\mathbb{R}[x]/(x-2) \oplus \mathbb{R}[x]/(x-2) \oplus \mathbb{R}[x]/(x-3) \simeq \mathbb{R}[x]/(x-2) \oplus \mathbb{R}[x]/(x^2-5x+6) \simeq \mathbb{R}^3,$$

由于 $1\cdot(x-2) + (-1)\cdot(x-3) = 1$, 所以同构映射满足

$$\begin{aligned}
&(1,0,0) \mapsto (1,0) \mapsto \mu_2 \mapsto e_2 + \frac{1}{7}e_3,\\
&(0,1,0) \mapsto (0,-\bar{x}+3) \mapsto (-x+3)\mu_3 \mapsto (-x+3)e_3 = -14e_1 + 7e_2 + e_3,\\
&(0,0,1) \mapsto (0,\bar{x}-2) \mapsto (x-2)\mu_3 \mapsto (x-2)e_3 = 14e_1 - 7e_2,
\end{aligned}$$

所以 A 在 $[e_2 + \frac{1}{7}e_3, -14e_1 + 7e_2 + e_3, 14e_1 - 7e_2]$ 这组基下变换为 Jordan 标准型, 即

$$P = \begin{pmatrix} 0 & -14 & 14 \\ 1 & 7 & -7 \\ \frac{1}{7} & 1 & 0 \end{pmatrix}, \quad P^{-1}AP = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{pmatrix}.$$ ■

例 4.4. 给定 $\mathbb{R}$ 上的矩阵

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

计算 A 的有理标准型和 Jordan 标准型, 并给出变换矩阵.

Solution. A 的特征矩阵为

$$xI_4 - A = \begin{pmatrix} x-1 & 0 & 0 & 0 \\ 0 & x-1 & 0 & 0 \\ 2 & 2 & x & -1 \\ 2 & 0 & 1 & x+2 \end{pmatrix},$$

任取 $\mathbb{R}[x]^4$ 的一组基 $[\xi_1, \xi_2, \xi_3, \xi_4]$. 依次通过初等变换

$$R_3 \leftrightarrow R_1, C_2 - C_1 \mapsto C_2, C_3 - \frac{x}{2}C_1 \mapsto C_3, C_4 + \frac{1}{2}C_1 \mapsto C_4,$$
$$R_3 - \frac{x-1}{2}R_1 \mapsto R_3, R_4 - R_1 \mapsto R_4, R_4 \leftrightarrow R_2, R_3 + \frac{-x+1}{2}R_2 \mapsto R_3,$$
$$R_4 + \frac{x-1}{2}R_2 \mapsto R_4, C_3 + \frac{-x+1}{2}C_2 \mapsto C_3, C_4 + \frac{x+3}{2}C_2 \mapsto C_4,$$
$$R_4 - (x-1)R_3 \mapsto R_4, C_4 - (x+2)C_3 \mapsto C_4, \frac{1}{2}C_1, -\frac{1}{2}C_2, -2C_3, 2C_4,$$

将 $xI_4 - A$ 变为 Smith 标准型

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & x-1 & 0 \\ 0 & 0 & 0 & (x-1)(x+1)^2 \end{pmatrix}.$$

对应的基变换为

$$\begin{aligned}
[\xi_1, \xi_2, \xi_3, \xi_4] &\to [\xi_3, \xi_2, \xi_1, \xi_4] \to \left[\xi_3 + \frac{x-1}{2}\xi_1, \xi_2, \xi_1, \xi_4\right] \\
&\to \left[\xi_3 + \frac{x-1}{2}\xi_1 + \xi_4, \xi_2, \xi_1, \xi_4\right] \to \left[\xi_3 + \frac{x-1}{2}\xi_1 + \xi_4, \xi_4, \xi_1, \xi_2\right] \\
&\to \left[\xi_3 + \frac{x-1}{2}\xi_1 + \xi_4, \xi_4 + \frac{x-1}{2}\xi_1, \xi_1, \xi_2\right] \\
&\to \left[\xi_3 + \frac{x-1}{2}\xi_1 + \xi_4, \xi_4 + \frac{x-1}{2}\xi_1 - \frac{x-1}{2}\xi_2, \xi_1, \xi_2\right] \\
&\to \left[\xi_3 + \frac{x-1}{2}\xi_1 + \xi_4, \xi_4 + \frac{x-1}{2}\xi_1 - \frac{x-1}{2}\xi_2, \xi_1 + (x-1)\xi_2, \xi_2\right].
\end{aligned}$$

这表明 $\mathbb{R}[x]/(x-1) \oplus \mathbb{R}[x]/(x^3 + x^2 - x - 1) \simeq \mathbb{R}^3$ 的同构映射满足:

$$\begin{aligned}
(1, 0) &\mapsto e_1 + (x-1)e_2 = e_1 - 2e_3, \\
(0, 1) &\mapsto e_2, \\
(0, \bar{x}) &\mapsto xe_2 = e_2 - 2e_3, \\
(0, \bar{x}^2) &\mapsto x^2 e_2 = e_2 - 2e_3 + 2e_4,
\end{aligned}$$

所以 A 在 $[e_1-2e_3, e_2, e_2-2e_3, e_2-2e_3+2e_4]$ 这组基下变换为有理标准型, 即

$$P=\begin{pmatrix}1&0&0&0\\0&1&1&1\\-2&0&-2&-2\\0&0&0&2\end{pmatrix},\quad P^{-1}AP=\begin{pmatrix}1&0&0&0\\0&0&0&1\\0&1&0&1\\0&0&1&-1\end{pmatrix}.$$

由于 $\frac{-x-3}{4}(x-1)+\frac{1}{4}(x+1)^2=1$, 所以

$$\mathbb{R}[x]/(x-1)\oplus\mathbb{R}[x]/(x-1)\oplus\mathbb{R}[x]/(x+1)^2\simeq\mathbb{R}[x]/(x-1)\oplus\mathbb{R}[x]/(x^3+x^2-x-1)\simeq\mathbb{R}^3$$

的同构映射满足:

$$\begin{aligned}
(1,0,0)&\mapsto(1,0)\mapsto e_1-2e_3,\\
(0,1,0)&\mapsto\left(0,\frac{1}{4}(\bar{x}+1)^2\right)\mapsto\frac{(x+1)^2}{4}e_2=e_2-\frac{3}{2}e_3+\frac{1}{2}e_4,\\
(0,0,\bar{x}+1)&\mapsto\left(0,-\frac{1}{4}(\bar{x}^3+3\bar{x}^2-\bar{x}-3)\right)\mapsto-\frac{x^3+3x^2-x-3}{4}e_2=e_3-e_4,\\
(0,0,1)&\mapsto\left(0,-\frac{1}{4}(\bar{x}^2+2\bar{x}-3)\right)\mapsto-\frac{x^2+2x-3}{4}e_2=\frac{3}{2}e_3-\frac{1}{2}e_4,
\end{aligned}$$

所以 A 在 $[e_1-2e_3, e_2-\frac{3}{2}e_3+\frac{1}{2}e_4, e_3-e_4, \frac{3}{2}e_3-\frac{1}{2}e_4]$ 这组基下变换为 Jordan 标准型, 即

$$P=\begin{pmatrix}1&0&0&0\\0&1&0&0\\-2&-\frac{3}{2}&1&\frac{3}{2}\\0&\frac{1}{2}&-1&-\frac{1}{2}\end{pmatrix},\quad P^{-1}AP=\begin{pmatrix}1&0&0&0\\0&1&0&0\\0&0&-1&1\\0&0&0&-1\end{pmatrix}.$$ ■

参考文献

[1] Dummit DS, Foote RM. Abstract Algebra. Hoboken: Wiley; 2004.