高等代数 (II) 第三次习题课

李卓远 数学科学学院

zy.li@stu.pku.edu.cn

1 内容概要

- 用初等对称多项式表示对称多项式;
- 求多项式的判别式 (Newton 公式, 结式);
- 求矩阵的不变因子, 初等因子, 有理标准型, 视为复矩阵的 Jordan 标准型.

2 补充知识

2.1 对称多项式

Definition 2.1.1. 定义

$$W = \{ f \in R[x_1, \dots, x_n] | f(x_1, \dots, x_n) = f(x_{\pi(1)}, \dots, x_{\pi(n)}), \forall \pi \in S(n) \}$$

为对称多项式之集, 其继承 $R[x_1, \cdots, x_n]$ 上的运算构成子环. 又称

$$\sigma_1 = x_1 + \dots + x_n, \ \sigma_2 = x_1 x_2 + \dots + x_{n-1} x_n, \ \sigma_k = \sum_{1 \le j_1 < \dots < j_k \le n} x_{j_1} \cdots x_{j_k}, \ \sigma_n = x_1 \cdots x_n$$

为初等对称多项式 (elementary symmetric polynomials).

Theorem 2.1.2. 任意对称多项式可由初等对称多项式唯一标出,即

$$T: R[x_1, \cdots, x_n] \to W$$

$$x_k \mapsto \sigma_k$$

$$h(x_1, \cdots, x_n) \mapsto h(\sigma_1, \cdots, \sigma_n)$$

为环同构. 这同时也说明了 $W = R[\sigma_1, \cdots, \sigma_n]$.

证明. 直接构造. 令 $ax_1^{l_1}\cdots x_n^{l_n}$ 为 $f\in W$ 的首项, $a\in R, l_1\geq l_2\geq \cdots \geq l_n$. 构造

$$\varphi_1 = a\sigma_1^{l_1 - l_2} \sigma_2^{l_2 - l_3} \cdots \sigma_n^{l_n},$$

其拥有首项

$$ax_1^{l_1-l_2}(x_1x_2)^{l_2-l_3}\cdots(x_1\cdots x_n)^{l_n}=ax_1^{l_1}\cdots x_n^{l_n}.$$

令 $f_1=f-\varphi_1$, 其首项次数严格小于 f 的首项次数 (in canonical order). 依次构造 $f_1,\,f_2,\,\cdots$, 直至 $f_q=0$. 取 $f=f_1+\cdots+f_q\in W$ 即可. 下面再证 T 为单射, 只需说明 $\ker T=0$. 考虑

$$f(\sigma_1,\cdots,\sigma_n)=0$$

中次数最高的项, 得必为 0 单项式.

2.2 重根与判别式

考虑首一多项式

$$f(x) = x^n + a_{n-1}x^{n-1} + \dots + a_1x_1 + a_0 = \prod_{k=1}^n (x - z_k).$$

那么

$$f$$
 有重根 $\iff z_k = z_l$ for some $k \neq l \iff \prod_{k < l} (z_k - z_l)^2 = 0$.

Vieta's formulas 给出

$$\sigma_k(z_1, \dots, z_n) = (-1)^k a_{n-k}, \ k = 1, 2, \dots, n.$$

令

$$g = T^{-1} \left(\prod_{k < l} (x_k - x_l)^2 \right),$$

有

$$\left. \prod_{k < l} (z_k - z_l)^2 = TT^{-1} \left(\prod_{k < l} (x_k - x_l)^2 \right) \right|_{x_j = z_j} = g(\sigma_1, \dots, \sigma_n)|_{x_j = z_j} = g(-a_{n-1}, a_{n-2}, \dots, (-1)^n a_0),$$

记为 f 的判别式 D(f) (discriminant). 一般地, 对于非首一的多项式 f, 约定

$$D(f) = a_n^{2n-2} D(a_n^{-1} f).$$

2.3 Newton 公式

$$D(f) = g(-a_{n-1}, a_{n-2}, \cdots, (-1)^n a_0) = V(z_1, \cdots, z_n)^2$$

$$= \begin{vmatrix} 1 & 1 & \cdots & 1 \\ z_1 & z_2 & \cdots & z_n \\ \vdots & \vdots & & \vdots \\ z_1^{n-1} & z_2^{n-1} & \cdots & z_n^{n-1} \\ \end{vmatrix} \begin{vmatrix} 1 & z_1 & \cdots & z_1^{n-1} \\ 1 & z_2 & \cdots & z_2^{n-1} \\ \vdots & \vdots & & \vdots \\ 1 & z_n & \cdots & z_n^{n-1} \end{vmatrix}$$

$$= \begin{vmatrix} s_0 & s_1 & \cdots & s_{n-1} \\ s_1 & s_2 & \cdots & s_n \\ \vdots & \vdots & & \vdots \\ s_{n-1} & s_n & \cdots & s_{2n-2} \end{vmatrix}$$

Proposition 2.3.1 (Newton's identities). The relation of

$$s_l = \sum_{k=1}^n x_k^l$$
 and $\sigma_l = \sum_{1 \le j_1 < \dots < j_l \le n} x_{j_1} \cdots x_{j_l}$

are given as follows:

$$s_{1} = \sigma_{1},$$

$$s_{2} = \sigma_{1}s_{1} - 2\sigma_{2},$$

$$s_{3} = \sigma_{1}s_{2} - \sigma_{2}s_{1} + 3\sigma_{3},$$

$$s_{k} = \cdots \cdot \cdot \cdot (k \le n)$$

$$s_{n} = \sigma_{1}s_{n-1} - \sigma_{2}s_{n-2} + \cdots + (-1)^{n}\sigma_{n-1}s_{1} + (-1)^{n+1}n\sigma_{n},$$

$$s_{n} = \sigma_{1}s_{n-1} - \sigma_{2}s_{n-2} + \cdots + (-1)^{n}\sigma_{n-1}s_{1} + (-1)^{n+1}\sigma_{n}s_{0},$$

$$s_{n+1} = \sigma_{1}s_{n} - \sigma_{2}s_{n-1} + \cdots + (-1)^{n}\sigma_{n-1}s_{2} + (-1)^{n+1}\sigma_{n}s_{1},$$

$$s_{n+2} = \sigma_{1}s_{n+1} - \sigma_{2}s_{n} + \cdots + (-1)^{n}\sigma_{n-1}s_{3} + (-1)^{n+1}\sigma_{n}s_{2},$$

$$s_{n+k} = \cdots \cdots (at \ most \ n \ terms)$$

2.4 矩阵环

对于数阵, 其矩阵加法和乘法只依赖于矩阵元素所在环的结构. 类比 \mathbb{Z} 和 \mathbb{R} 上的矩阵, 可类似定义一般的环 R 上的元素组成的矩阵, 构成的矩阵全体记为 $M_n(R)$. 不难验证 $M_n(R)$ 自身也为一个环.

- 子环 (subring). 按矩阵结构, 可单独取出对角/上三角/下三角矩阵全体构成 $M_n(R)$ 的子环; 亦可利用 R 本身的结构, 若 R' 为 R 的子环, 则 $M_n(R')$ 也为 $M_n(R)$ 的子环;
- 中心 (center). 与所有元素可交换的元素之集, 也构成 $M_n(R)$ 的一个子环, 记为 $Z(M_n(R))$. 例如 $Z(M_n(\mathbb{R}))$ 为全体数量阵之集;
- 单位 (unit). 所有可逆元之集, 也构成 $M_n(R)$ 的一个子环, 记为 $GL_n(R)$.

Theorem 2.4.1. (Smith Normal Form, SNF) 对任意 $A \in M_n(\mathbb{F}[\lambda])$, 存在可逆 P, Q(初等 λ -阵) 使得

$$PAQ = diag(d_1, d_2, \dots, d_n), d_1 \mid d_2 \mid \dots \mid d_n.$$

Remark 2.4.2. SNFs exist for general PIDs and non-square matrices. See https://en.wikipedia.org/wiki/Smith normal form for a constructive proof.

3 典型例题

Problem 3.1. $\not x T^{-1}(\sum x_1^2 x_2^2)$.

证明. 取 $\sum x_1^2 x_2^2$ 的首项次数 $(2,2,0\cdots,0)$, 得以构造 σ_2^2 .

$$\sum x_1^2 x_2^2 - \sigma_2^2 = -2 \sum x_1^2 x_2 x_3 - 6 \sum x_1 x_2 x_3 x_4$$
$$\sum x_1^2 x_2 x_3 = \sigma_1 \sigma_3 - 4 \sum x_1 x_2 x_3 x_4.$$

于是

$$\sum x_1^2 x_2^2 = \sigma_2^2 - 2 \sum x_1^2 x_2 x_3 - 6\sigma_4 = \sigma_2^2 - 2\sigma_1 \sigma_3 + 2\sigma_4.$$

注意当 n < 4 时, 需要删去 σ_3 或/和 σ_4 .

Remark 3.2. 经过上述计算可总结以下规律:

1. 任一对称多项式可写为若干单项式之和, 而这些单项式可视为若干代表元循环求和, 即若已知 $f \in W$ 的最高次项为 $x_1^2 x_2^2$, 则必可写为

$$f = a \sum x_1^2 x_2^2 + b \sum x_1^2 x_2 x_3 + c \sum x_1 x_2 x_3 x_4.$$

- 2. 上述分解中求和号前系数与任一代表元在 f 中的系数一致.
- 3. 要计算某单项式在 f 中的系数, 只需统计生成方式的可能组合. 例如

$$f = \sum x_1^2 x_2^2 - (x_1 x_2 + \dots + x_{n-1} x_n)^2 = a \sum x_1^2 x_2 x_3 + b \sum x_1 x_2 x_3 x_4.$$

- 对于系数 $a, x_1^2 x_2 x_3 = x_1 x_2 \cdot x_1 x_3$, 对应 a = -2;
- 对于系数 b, $x_1x_2x_3x_4 = x_1x_2 \cdot x_3x_4 = x_1x_3 \cdot x_2x_4 = x_1x_4 \cdot x_2x_3$, 对应 b = -6.

进一步思考, 原计算过程是否可以进行简化? 为叙述方便, 令

$$[2,2] = \sum x_1^2 x_2^2, \, [2,1,1] = \sum x_1^2 x_2 x_3, \, [\alpha] = \sum x_1^{\alpha_1} x_2^{\alpha_2} \cdot x_n^{\alpha_n},$$

那么对于任一 α 满足 $\alpha_1 \geq \alpha_2 \geq \cdots \geq \alpha_n$,

$$[\alpha] = \sigma_1^{\alpha_1 - \alpha_2} \sigma_2^{\alpha_2 - \alpha_3} \cdots \sigma_n^{\alpha_n} + \sum_{\substack{\beta < \alpha \\ |\beta| = |\alpha| \\ \beta_1 > \cdots > \beta_n}} C(\alpha, \beta)[\beta],$$

其中 $C(\alpha, \beta)$ 为相应系数. 归纳易得

$$[\alpha] = \sum_{\substack{\beta < \alpha \\ |\beta| = |\alpha| \\ \beta_1 > \dots > \beta_n}} \tilde{C}(\alpha, \beta) \sigma_1^{\beta_1 - \beta_2} \cdots \sigma_n^{\beta_n}.$$

故仅需

- 1. 列举所有 β 满足 $\beta < \alpha$, $|\beta| = |\alpha|$ 且 $\beta_1 \ge \cdots \ge \beta_n$;
- 2. 求出 $\sigma_1^{\beta_1-\beta_2}\cdots\sigma_n^{\beta_n}$;
- 3. 求出相应系数.

Problem 3.3. $\not x T^{-1}((x_1^2 + 2x_2x_3)(x_2^2 + 2x_3x_1)(x_3^2 + 2x_1x_2)).$

证明. 令 $f = (x_1^2 + 2x_2x_3)(x_2^2 + 2x_3x_1)(x_3^2 + 2x_1x_2)$, 得首项次数为 4, 1, 1, 于是有

$$f = c_1[4, 1, 1] + c_2[3, 3, 0] + c_3[3, 2, 1] + c_4[2, 2, 2] = d_1\sigma_1^3\sigma_3 + d_2\sigma_2^3 + d_3\sigma_1\sigma_2\sigma_3 + d_4\sigma_3^2$$

其中 $d_1 = 4$ 可通过观察 f 的首项系数直接得到. 接下来利用待定系数法:

• $(x_1, x_2, x_3) = (1, 1, 0)$: $\sigma_1 = 2$, $\sigma_2 = 1$, $\sigma_3 = 0$.

$$f(x_1, x_2, x_3) = 2 = d_1 \cdot 2^3 \cdot 0 + d_2 \cdot 1^3 + d_3 \cdot 2 \cdot 1 \cdot 0 + d_4 \cdot 0^2;$$

• $(x_1, x_2, x_3) = (1, 1, 1)$: $\sigma_1 = 3$, $\sigma_2 = 3$, $\sigma_3 = 1$.

$$f(x_1, x_2, x_3) = 3^3 = d_1 \cdot 2^3 \cdot 1 + d_2 \cdot 3^3 + d_3 \cdot 3 \cdot 3 \cdot 1 + d_4 \cdot 1^2;$$

•
$$(x_1, x_2, x_3) = (1, 1, -1)$$
: $\sigma_1 = 1$, $\sigma_2 = -1$, $\sigma_3 = -1$.

$$f(x_1, x_2, x_3) = 3 = d_1 \cdot 1^3 \cdot (-1) + d_2 \cdot (-1)^3 + d_3 \cdot 1 \cdot (-1) \cdot (-1) + d_4 \cdot (-1)^2$$
.

解得 $d_2 = 2$, $d_3 = -18$, $d_4 = 27$.

Problem 3.4. 求方程 $x^3 + a_2x^2 + a_1x + a_0 = 0$ 的三个复根构成等差数列的充要条件.

证明. 令

$$f(x_1, x_2, x_3) = (2x_1 - x_2 - x_3)(2x_2 - x_3 - x_1)(2x_3 - x_1 - x_2).$$

求 $T^{-1}f$ 得

$$f = 2\sigma_1^3 - 9\sigma_1\sigma_2 + 27\sigma_3$$
.

结合 Vieta's formulas:

$$f(z_1, z_2, z_3) = 0 \iff 2(-a_2)^3 - 9a_1(-a_2) + 27(-a_0) = 0.$$

Remark 3.5. 类似地, 可以求得"三个复根某一个平方成为另两个平方和"的充要条件.

Problem 3.6. $\not x D(x^n + a)$.

证明. 利用 Vieta's formulas:

$$\sigma_1 = \sigma_2 = \dots = \sigma_{n-1} = 0, \ \sigma_n = (-1)^n a.$$

结合 Newton's identities:

$$s_{1} = \sigma_{1} = 0,$$

$$s_{2} = \sigma_{1}s_{1} - 2\sigma_{2} = 0,$$

$$\dots$$

$$s_{n-1} = \sigma_{1}s_{n-1} - \sigma_{2}s_{n-2} + \dots + (-1)^{n-1}\sigma_{n-2}s_{1} + (-1)^{n}(n-1)\sigma_{n-1} = 0,$$

$$s_{n} = \sigma_{1}s_{n-1} - \sigma_{2}s_{n-2} + \dots + (-1)^{n}\sigma_{n-1}s_{1} + (-1)^{n+1}n\sigma_{n} = -na,$$

$$s_{n+1} = \sigma_{1}s_{n} - \sigma_{2}s_{n-1} + \dots + (-1)^{n}\sigma_{n-1}s_{2} + (-1)^{n+1}\sigma_{n}s_{1} = 0,$$

$$s_{n+2} = \sigma_{1}s_{n+1} - \sigma_{2}s_{n} + \dots + (-1)^{n}\sigma_{n-1}s_{3} + (-1)^{n+1}\sigma_{n}s_{2} = 0,$$

$$\dots$$

$$s_{2n-2} = \sigma_{1}s_{2n-3} - \sigma_{2}s_{2n-4} + \dots + (-1)^{n}\sigma_{n-1}s_{n-1} + (-1)^{n+1}\sigma_{n}s_{n-2} = 0.$$

于是

$$D(x^{n} + a) = \begin{vmatrix} n & & & 0 \\ & & \ddots & -na \\ & & \ddots & & \\ & & \ddots & & \\ 0 & -na & & \end{vmatrix} = n(-na)^{n-1}(-1)^{\frac{(n-1)(n-2)}{2}} = (-1)^{\frac{n(n-1)}{2}}n^{n}a^{n-1}.$$

Remark 3.7. 判别式亦可通过结式给出, 主要依赖以下结论:

$$R(f, f') = a_0^{n-1} \prod_{k=1}^n f'(x_k)$$

$$= a_0^{n-1} \prod_{k=1}^n a_0 \sum_{j=1}^n \prod_{l \neq j} (x_k - x_l)$$

$$= a_0^{2n-1} \prod_{k=1}^n \prod_{l \neq k} (x_k - x_l) = (-1)^{\frac{n(n-1)}{2}} a_0^{2n-1} \prod_{k < l} (x_k - x_l)^2,$$

即

$$R(f, f') = (-1)^{\frac{n(n-1)}{2}} a_0 D(f).$$

对上题另解:

Remark 3.8. 利用结式还可求稍微复杂的 $D(x^n + a_1x + a_0)$, 留作思考.

Problem 3.9. 考虑

$$f(x) = x^3 + a_2 x^2 + a_1 x + a_0 \in \mathbb{R}[x].$$

讨论 D(f) 的符号与 f 的零点的关系.

证明. 当 D(f)=0 时 f 在 $\mathbb C$ 上有重根, 而由于虚根总是成对出现, 可知重根为实根, 故而 f 的另一零点也为实数. 当 $D(f)\neq 0$ 时, f 的三个零点互不相同. 若全为实数则必有 D(f)>0, 否则不妨设三个零点分别为

$$z_1 \in \mathbb{R}, z_2 = a + bi, z_3 = a - bi, a, b \in \mathbb{R}, b \neq 0.$$

此时

$$D(f) = (a_b i - z_1)^2 (a - b i - z_1)^2 (2bi)^2 = -4b^2 ((a - z_1)^2 + b^2) < 0.$$

综上所述, 当 $D(f) \ge 0$ 对应三个实根, 且 D(f) = 0 时有重根; D(f) < 0 对应一个实根和两个共轭复根.

Remark 3.10. 特别地, 当 $a_2=0$ 时可通过 $D(f)=-4a_1^3-27a_0^2$ 讨论 f=0 的根的情况. 称代数曲线

$$u^2 = x^3 + a_1 x + a_0$$

为椭圆曲线, 若 $4a_1^3 + 27a_0^2 > 0$, 即代数曲线与 x 轴在 \mathbb{R}^2 中只有一个交点. 椭圆曲线加密算法 (elliptic curve cryptography, ECC) 的实现以椭圆曲线上的群结构为基础, 已在现代密码学领域得到了广泛的应用.

Problem 3.11. 计算下列 λ-矩阵的标准形.

$$A(\lambda) = \begin{pmatrix} 3 & -6 & \lambda - 8 \\ \lambda + 2 & 0 & -3 \\ -2 & \lambda + 2 & 4 \end{pmatrix} \qquad B(\lambda) = \begin{pmatrix} \lambda & & & \\ & \lambda^2 - 1 & \lambda + 1 \\ & & 1 - \lambda^3 \end{pmatrix}$$

证明. 直接进行行列变换

$$A(\lambda) \to \begin{pmatrix} 1 & \lambda - 4 & \lambda - 4 \\ \lambda + 2 & 0 & -3 \\ -2 & \lambda + 2 & 4 \end{pmatrix} \to \begin{pmatrix} 1 & 0 & 0 \\ \lambda + 2 & -\lambda^2 + 2\lambda + 8 & -\lambda^2 + 2\lambda + 5 \\ -2 & \lambda + 2 & 4 \end{pmatrix}$$

$$\to \begin{pmatrix} 1 & & & \\ -\lambda^2 + 2\lambda + 8 & -\lambda^2 + 2\lambda + 5 \\ 3\lambda - 6 & 2\lambda - 4 \end{pmatrix} \to \begin{pmatrix} 1 & & \\ -\lambda^2 + 2\lambda + 8 & -3 \\ 3\lambda - 6 & -\lambda + 2 \end{pmatrix}$$

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

$$\to \begin{pmatrix} 1 & & \\ & (\lambda - 1)^2 & -3 \\ & & \frac{1}{3}(\lambda - 1)^2(\lambda - 2) \end{pmatrix} \to \begin{pmatrix} 1 & & \\ & 1 & \\ & (\lambda - 1)^2(\lambda - 2) \end{pmatrix}$$

 $B(\lambda)$ 对应的行列式因子

$$D_1 = 1, D_2 = \left(\begin{vmatrix} \lambda & & \\ & \lambda^2 - 1 \end{vmatrix}, \begin{vmatrix} \lambda & & \\ & \lambda + 1 \end{vmatrix}, \begin{vmatrix} \lambda & & \\ & 1 - \lambda^3 \end{vmatrix}, \begin{vmatrix} \lambda^2 - 1 & \lambda + 1 \\ & 1 - \lambda^3 \end{vmatrix} \right) = 1$$

得 $B(\lambda)$ 相抵于 diag $(1, 1, \det B(\lambda))$.

Problem 3.12. 给出

$$A = \begin{pmatrix} 8 & -3 & 3 & -2 \\ 0 & 3 & 0 & 2 \\ -10 & 6 & -3 & 4 \\ 2 & 0 & 1 & 2 \end{pmatrix}$$

的相似标准形.

证明. 计算其对应 λ -矩阵的相抵标准形:

注意与数阵不同, 由最后的相抵标准形一般无法得到其相抵于 $diag(1, \lambda-2, \lambda-3, (\lambda-2)(\lambda-3))$, 但往往可借此探究行列式因子的取值. 对于右下角的三阶子阵有

$$D_1 = 1, D_2 \mid \left(\begin{pmatrix} \lambda - 2 & \\ & \lambda - 3 \end{pmatrix}, \begin{pmatrix} 3 & -3\lambda + 8 \\ \lambda - 3 & -2 \end{pmatrix} \right) = 1,$$

于是 $\lambda I - A$ 与 diag $(1,1,1,(\lambda-2)^2(\lambda-3)^2)$ 相抵, 原矩阵 A 的相似标准形为

$$A \sim \begin{pmatrix} 2 & & & \\ 1 & 2 & & \\ & & 3 & \\ & & 1 & 3 \end{pmatrix}.$$