第三章　特征值与矩阵的Jordan标准形

除特别说明, 本章讨论的矩阵都是复数矩阵.

引　言　如何计算矩阵的高次幂 A^m?

在第二章第六节公式 (2.6.4)(差分方程与线性滤波器) 中, 我们看到计算方阵 A 的任意次幂 A^m 是矩阵理论的基本问题. 我们知道, 如果 A 可以 (相似) 对角化, 即存在可逆矩阵 P 使得 $P^{-1}AP = D$ 为对角矩阵, 则可以方便地计算 $A^m = PD^mP^{-1}$. 但并不是任何方阵都可以对角化, 因此寻找 $P^{-1}AP$ 的最简单形式仍是一个需要解决的问题. 早在1837年, Jacobi 即已证明了任何复矩阵均可以 (相似) 三角化, 一般称此结果为 (矩阵的) **Jacobi 定理** . 本章我们首先证明一个更强的结论, 即任何复矩阵都可以酉三角化, 此即著名的Schur[24] 三角化定理. 以此为基础我们证明Cayley-Hamilton [25] 定理, 进而利用最小多项式给出矩阵可以对角化的一个刻画. 利用最小多项式, 可以得到计算高次幂 A^m 的一个降幂方法. 利用更精细的分块Schur三角化定理即可证明任何矩阵的Jordan [26] 标准形的存在性与唯一性.

第一节　Schur三角化定理: 化简矩阵的基础

定理 3.1.1 (Schur (酉) 三角化定理) 设 $A \in \mathbb{C}^{n\times n}$, 则存在酉矩阵 U 使

$$U^*AU = B, \tag{3.1.1}$$

其中 B 为一个上三角矩阵.

证　对矩阵 A 的阶数 n 施行数学归纳法. 当 $n = 1$ 时, 结论显然成立. 假设结论对阶数小于 n 的矩阵都成立, 下证结论对 n 阶矩阵也成立.

设 $\lambda_1 \in \mathbb{C}$ 为 A 的一个特征值, 相应的单位特征向量为 α_1, 由第二章的讨论可知, α_1 可以扩充成 $\mathbb{C}^n$ 的一组标准正交基 $\alpha_1, \alpha_2, \cdots, \alpha_n$, 因而 $U_1 = (\alpha_1, \alpha_2, \cdots, \alpha_n)$ 是一个酉阵. 因为向量组 $A\alpha_1 = \lambda\alpha_1$, $A\alpha_2, \cdots, A\alpha_n$ 为基 $\alpha_1, \alpha_2, \cdots, \alpha_n$ 的线性组合, 所以

$$AU_1 = A(\alpha_1, \alpha_2, \cdots, \alpha_n) = (\alpha_1, \alpha_2, \cdots, \alpha_n)\begin{pmatrix} \lambda_1 & c_{12} & c_{13} & \cdots & c_{1n} \\ 0 & c_{22} & c_{23} & \cdots & c_{2n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & c_{n2} & c_{n3} & \cdots & c_{nn} \end{pmatrix},$$

其中 c_{ij} 为一些复数. 令

$$C_1 = \begin{pmatrix} c_{22} & c_{23} & \cdots & c_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ c_{n2} & c_{n3} & \cdots & c_{nn} \end{pmatrix}.$$

---

[24]Issai Schur (1875-1941), 著名德国数学家, Frobenius 的学生.

[25]Arthur Cayley(1821-1895), 著名英国数学家, 英国现代数学的奠基人. William Hamilton(1805-1865), 著名爱尔兰物理学家, 天文学家和数学家, 以其名字命名的Hamilton力学是牛顿力学的革新, 数学上以发现四元数 (Quaternions) 而闻名.

[26]Marie Ennemond Camille Jordan (1838-1922), 法国著名数学家, 数学上还有著名的 Jordan Curve Theorem. 第25593号小行星camillejordan即以其名字命名.

则

$$AU_1 = U_1 \begin{pmatrix} \lambda_1 & c_{12}\, c_{13} \cdots c_{1n} \\ 0 & \\ \vdots & C_1 \\ 0 & \end{pmatrix}.$$

由于 C_1 为 $n-1$ 阶矩阵, 由归纳假设, 存在 $n-1$ 阶酉矩阵 U_2 使

$$U_2^* C_1 U_2 = B_1 = \begin{pmatrix} b_{22} & b_{23} & \cdots & b_{2n} \\ & b_{33} & \cdots & b_{3n} \\ & & \ddots & \vdots \\ & & & b_{nn} \end{pmatrix}$$

为上三角矩阵. 令

$$U = U_1 \begin{pmatrix} 1 & \\ & U_2 \end{pmatrix},$$

则

$$\begin{aligned} U^*AU &= \begin{pmatrix} 1 & \\ & U_2^* \end{pmatrix} U_1^* A U_1 \begin{pmatrix} 1 & \\ & U_2 \end{pmatrix} \\ &= \begin{pmatrix} 1 & \\ & U_2^* \end{pmatrix} \begin{pmatrix} \lambda_1 & c_{12}\, c_{13} \cdots c_{1n} \\ 0 & \\ \vdots & C_1 \\ 0 & \end{pmatrix} \begin{pmatrix} 1 & \\ & U_2 \end{pmatrix} \\ &= \begin{pmatrix} \lambda_1 & d_{12} & d_{13} & \cdots & d_{1n} \\ & b_{22} & b_{23} & \cdots & b_{2n} \\ & & b_{33} & \cdots & b_{3n} \\ & & & \ddots & \vdots \\ & & & & b_{nn} \end{pmatrix}. \end{aligned}$$

定理得证. □

公式 (3.1.1) 可变形为

$$A = UBU^* \tag{3.1.2}$$

一般称公式 (3.1.2) 为矩阵 A 的 **Schur 分解**.

Schur 三角化定理是矩阵理论中最重要的结果之一 (实际上在许多领域均是如此), 众多深刻的结论均由其导出, 也是矩阵快速计算的基础, 值得深入研究. 由 Schur 三角化定理, 总可以使 $P^{-1}AP$ 为上三角矩阵. 进一步, 可以使 $P^{-1}AP$ 为分块对角矩阵, 且每个对角块是对角元素均相同的上三角矩阵. 为此只需证明下述引理 (此时 $\mathbb{F}$ 不必是复数域).

引理 3.1.1 设 $A = (a_{ij})$ 是 $\mathbb{F}$ 上的任意上三角矩阵, $1 \le p < q \le n$. 设 $P = I + \alpha E_{pq}$, $\alpha \in \mathbb{F}$, 则 $B = (b_{ij}) = P^{-1}AP$ 是与 A 的主对角线相同 (包括顺序) 的上三角矩阵, 且 $b_{pq} = a_{pq} + \alpha(a_{pp} - a_{qq})$. 特别地, 若 $a_{pp} \ne a_{qq}$, 则可选取适当的 α 使得 $b_{pq} = 0$.

证 注意 P 是第三种初等矩阵, $P^{-1}=I-\alpha E_{pq}$. 故 $P^{-1}A$ 仅将 A 的第 q 行的 $-\alpha$ 倍加到第 p 行, 因此所得矩阵仍是上三角矩阵且不改变 A 的对角线; AP 的意义类似. 因此知 B 是与 A 的主对角线相同 (包括顺序) 的上三角矩阵. 直接计算可得 b_{pq}. □

例 3.1.1 设 $\lambda_1\neq\lambda_2$, $P=I-\frac{c}{(\lambda_1-\lambda_2)}E_{12}$,

$$A=\begin{pmatrix}\lambda_1 & a & b\\ 0 & \lambda_1 & c\\ 0 & 0 & \lambda_2\end{pmatrix},$$

则

$$B=P^{-1}AP=\begin{pmatrix}\lambda_1 & a & x\\ 0 & \lambda_1 & 0\\ 0 & 0 & \lambda_2\end{pmatrix}.$$

若再取 $Q=I-\frac{x}{(\lambda_1-\lambda_2)}E_{13}$, 则 $Q^{-1}BQ$ 是除将 B 中元素 x 变为 0 而其余元素均不变的上三角矩阵. 从而 A 相似于分块对角矩阵

$$\begin{pmatrix}\lambda_1 & a & 0\\ 0 & \lambda_1 & 0\\ 0 & 0 & \lambda_2\end{pmatrix}.$$

例 3.1.1 表明, 重复使用 引理 3.1.1 可得下述

定理 3.1.2 (分块 Schur 三角化定理) 设 n 阶复矩阵 A 的特征多项式 $|\lambda I-A|=\prod\limits_{i=1}^{s}(\lambda-\lambda_i)^{n_i}$, 其中 $\sigma(A)=\{\lambda_i\,|\,1\leq i\leq s\}$, $\sum\limits_{i=1}^{s}n_i=n$. 则存在可逆矩阵 P 使得

$$P^{-1}AP=A_1\oplus A_2\oplus\cdots\oplus A_s,$$

其中 A_i 是特征值均为 λ_i 的 n_i 阶上三角矩阵.

注. 请读者思考, 分块 Schur 三角化定理中的可逆矩阵 P 是否可加强为酉矩阵? 见本节思考题 2.

由分块 Schur 三角化定理即可证明下面的 Cayley-Hamilton 定理, 先考察一个简单的例子.

例 3.1.2 设 A 是 n 阶严格上三角矩阵, 则 $A^n=0$.

定理 3.1.3 (Cayley-Hamilton 定理[27]) 设矩阵 A 的特征多项式为 $f(\lambda)$, 则有 $f(A)=0$.

证 设 A 的特征多项式为

$$f(\lambda)=(\lambda-\lambda_1)^{n_1}(\lambda-\lambda_2)^{n_2}\cdots(\lambda-\lambda_s)^{n_s},$$

[27] 该定理当 $n=2,3$ 时由 Cayley 于 1853 年证明, $n=4$ 时由 Hamilton 于 1850 年左右证明, 但一般形式由 Frobenius 于 1878 年证明.

其中 $\sum_{i=1}^{s} n_i = n$. 首先注意, 若 A 与 B 相似, 则对任意多项式 $g(x)$, 有 $g(A) = 0 \iff g(B) = 0$. 故由分块Schur三角化定理, 可设 $A = A_1 \oplus A_2 \oplus \cdots \oplus A_s$, 其中 A_i 是特征值均为 λ_i 的 n_i 阶上三角矩阵. 则

$$f(A) = (A - \lambda_1 I)^{n_1}(A - \lambda_2 I)^{n_2} \cdots (B - \lambda_s I)^{n_s}.$$

由 例3.1.2 可知, 对每个 i, 均有 $(A_i - \lambda_i I_{n_i})^{n_i} = 0$, 故上式的第 i 个因子 $(A - \lambda_i I)^{n_i}$ 的第 i 个块为 n_i 阶0矩阵, 从而整个乘积等于0矩阵. □

由于 n 阶矩阵 A 的特征多项式是 n 次多项式, Cayley-Hamilton定理表明, A 的 n 次幂可由其较低次幂的线性组合给出, 因此, A 的高于 n 次的幂可由其低于 n 次的幂的线性组合给出, 故对任意自然数 m, 有

$$A^m \in \mathrm{Span}\{I,\ A,\ A^2,\ \cdots,\ A^{n-1}\}.$$

换句话说, n 阶矩阵 A 的任意次幂均属于由 $I, A, A^2, \cdots, A^{n-1}$ 生成的 $M_n(\mathbb{C})$ 的子空间. 这就提供了一种计算高次幂的降幂算法.

例 3.1.3 设

$$A = \begin{pmatrix} 2 & 3 \\ 1 & 2 \end{pmatrix},$$

求 A^2, A^3, A^4.

解 A 的特征多项式为 $f(\lambda) = \lambda^2 - 4\lambda + 1$, 所以 $A^2 - 4A + I = 0$. 故知

$$A^2 = 4A - I,\ A^3 = 4A^2 - A = 15A - 4I,\ A^4 = 15A^2 - 4A = 56A - 15I.$$

命题 3.1.1 (Sylvester降幂公式) 设 A 与 B 分别是 $m \times n$ 与 $n \times m$ 矩阵, $m \geq n$. 则

$$|\lambda I_m - AB| = \lambda^{m-n}|\lambda I_n - BA|.$$

证 注意下述分块矩阵的恒等式:

$$\begin{pmatrix} I & B \\ 0 & I \end{pmatrix}\begin{pmatrix} 0 & 0 \\ A & AB \end{pmatrix} = \begin{pmatrix} BA & BAB \\ A & AB \end{pmatrix} = \begin{pmatrix} BA & 0 \\ A & 0 \end{pmatrix}\begin{pmatrix} I & B \\ 0 & I \end{pmatrix},$$

因此, 矩阵

$$C_1 = \begin{pmatrix} BA & 0 \\ A & 0 \end{pmatrix} \quad \text{与矩阵} \quad C_2 = \begin{pmatrix} 0 & 0 \\ A & AB \end{pmatrix}$$

相似. 而 C_1 的特征值是 BA 的特征值加上 m 个0; C_2 的特征值是 AB 的特征值加上 n 个0. 现因 C_1 与 C_2 的特征值相同, 故 AB 与 BA 的非零特征值相同, 故只相差 $m - n$ 个0. □

上述命题亦称为特征多项式的降阶计算公式.

例 3.1.4 设

$$A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{pmatrix}, \quad B = \begin{pmatrix} 1 & -1 & 0 \\ 0 & 1 & -1 \end{pmatrix}.$$

则

$$AB = \begin{pmatrix} 1 & 1 & -2 \\ 3 & 1 & -4 \\ 5 & 1 & -6 \end{pmatrix}.$$

显然较难直接看出 AB 的特征值 (需要一定的计算), 但是

$$BA = \begin{pmatrix} -2 & -2 \\ -2 & -2 \end{pmatrix},$$

故无须计算即可知 BA 的特征值为 $0, \text{-}4$. 因此, 由 Sylvester 降幂公式知 AB 的特征值为 $0, 0, \text{-}4$. (由于 $r(AB) = 2$, 故还知道 AB 不能对角化, 但显然 BA 可以对角化.)

例 3.1.5 设 u 是 n 维单位向量, 求 n 阶实 Householder 矩阵 $I - 2uu^T$ 的特征值及它的迹和行列式.

解 由 Sylvester 降幂公式,

$$|\lambda I - (I - 2uu^T)| = |(\lambda - 1)I + 2uu^T| = (\lambda - 1)^{n-1}|\lambda - 1 + 2u^T u| = (\lambda - 1)^{n-1}(\lambda + 1).$$

由此知, $\lambda = 1$ 是 $n-1$ 重根, 而 $\lambda = -1$ 是 1 重根. 因此, $\operatorname{tr}(I - 2uu^T) = n - 2$; $|I - 2uu^T| = -1$.

例 3.1.6 设 A 是秩为 1 的 n 阶方阵, 试求其特征多项式.

解 由于 A 的秩为 1, 故它至少有 $n-1$ 个特征值为 0, 因此其特征多项式 $|\lambda I - A| = \lambda^{n-1}(\lambda - \operatorname{tr} A)$.

Sylvester 降幂公式与 Cayley-Hamilton 定理均提供了某种简化矩阵计算的方式. 如 例 3.1.4 所示, 当 m 与 n 均较小或当 $m-n$ 较大时, Sylvester 降幂公式较为有效, 而 Cayley-Hamilton 定理仅仅将 n 次幂简化为 $n-1$ 次幂和较低次幂的线性组合, 显然不是理想的快速计算模式. 然而, Cayley-Hamilton 定理却为我们提供了一种思路, 即如果有次数很低 (设为 d) 的多项式 $g(x)$ 使得 $g(A) = 0$, 则 A 的高次幂就可以通过 A 的 $d-1$ 次幂及更低次幂的线性组合求出.

例 3.1.7 设 3 阶矩阵 A 的特征多项式为 $f(\lambda) = \lambda^3 - 5\lambda^2$. 设 $g(x) = x^2 - 5x$ 满足条件 $g(A) = 0$, 试求 A^4.

解 由特征多项式可知 $A^3 - 5A^2 = 0$, 所以 $A^4 = 25A^2$. 但由 $g(A) = 0$ 可知 $A^2 - 5A = 0$, 所以 $A^4 = 125A$.

定义 3.1.1 设 A 是 n 阶矩阵, $f(x)$ 是多项式. 如果 $f(A) = 0$, 则称 $f(x)$ 是 A 的**零化多项式**. 次数最低的首一零化多项式称为 A 的**最小多项式**, 记为 $m_A(x)$ 或 $m(x)$.

最小多项式存在且唯一, 证明见习题 5.

由于矩阵 A 的特征多项式是它的零化多项式, 因此最小多项式的次数不超过 A 的阶数. 但一般来说, 寻找矩阵的最小多项式是非常困难的工作 (其难度并不亚于计算矩阵的特征多项式), 比如 例 3.1.4 中的矩阵 AB 的最小多项式是什么? 所以需要理解最小多项式的基本性质.

命题 3.1.2 设 $m(x)$ 是 A 的最小多项式, $f(x)$ 是 A 的任意零化多项式, 则 $m(x)|f(x)$. 特别地, $m(x)\,|\,|xI-A|$.

证 设 $f(x)=m(x)q(x)+r(x)$, 其中 $r(x)$ 的次数小于 $m(x)$ 的次数或者 $r(x)$ 为 0 多项式. 则 $0=f(A)=m(A)q(A)+r(A)=r(A)$, 即 $r(x)$ 也是 A 的零化多项式. 由于 $m(x)$ 是 A 的最小多项式, 故只有 $r(x)=0$, 从而 $m(x)\,|\,f(x)$. □

例 3.1.8 求下列 n 阶矩阵的最小多项式:

$$J_n=\begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ 0 & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \cdots & 1 \\ 0 & 0 & 0 & \cdots & 0 \end{pmatrix}.$$

解 矩阵 A 实际上是第一章例 2.2.22 中的线性变换 σ 关于标准基的矩阵. 由于 σ 是幂零指数为 n 的幂零变换, 故由线性变换与矩阵的对应关系可知 $A^n=0$, 但 $A^{n-1}\neq 0$(也可直接计算), 因此 A 的最小多项式为 $m(x)=x^n$, 恰好等于 A 的特征多项式. 矩阵 J_n 称为 n 阶标准幂零矩阵或 n 阶**幂零块**或 n 阶幂零 Jordan 块. 矩阵 J_n 显然是最简单的秩为 $n-1$ 的 n 阶 (严格上三角) 矩阵, 在矩阵的 Jordan 标准形理论和微积分运算中起关键性作用. 为方便起见, 常将 J_n 中元素 1 所在的斜线称为 (第一) 上对角线.

例 3.1.9 试求下列分块矩阵的最小多项式:

$$A=(a)\oplus\begin{pmatrix} a & 1 \\ 0 & a \end{pmatrix}\oplus\begin{pmatrix} a & 1 & 0 \\ 0 & a & 1 \\ 0 & 0 & a \end{pmatrix}.$$

解 A 的特征多项式为 $f(\lambda)=(\lambda-a)^6$. 从而由 命题3.1.2 可知其最小多项式 $m(x)=(x-a)^k$, $k\leq 6$, 所以只要考察 $(A-aI)$ 的乘积即可. 因为

$$A-aI=(0)\oplus\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}\oplus\begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}.$$

即 $A-aI$ 是由三个子块组成的分块对角矩阵, 每一个子块都是幂零块, 因此 $m(x)=(x-a)^3$.

例 3.1.10 设 $a\neq b$, 试求下列分块矩阵的最小多项式:

$$A=\begin{pmatrix} a & 1 \\ 0 & a \end{pmatrix}\oplus\begin{pmatrix} a & 1 & 0 \\ 0 & a & 1 \\ 0 & 0 & a \end{pmatrix}\oplus(b)\oplus\begin{pmatrix} b & 1 \\ 0 & b \end{pmatrix}.$$

解 A 的特征多项式为 $(\lambda-a)^5(\lambda-b)^3$. 所以 A 的最小多项式具有形式 $(\lambda-a)^s(\lambda-b)^t$. 考察矩阵 $A-aI$ 与 $A-bI$ 的乘积

$$A-aI=J_2\oplus J_3\oplus B,\quad A-bI=C\oplus(0)\oplus J_2,$$

其中 B 与 C 是可逆矩阵 (因为 $a \neq b$). 由分块对角矩阵的乘法可知, 只要分别求出使 $A-aI$ 与 $A-bI$ 的幂零块为零的次数即可. 这两个次数分别为3与2. 因此, A 的最小多项式为 $m(x)=(x-a)^3(x-b)^2$.

命题 3.1.3 设 A 是任意方阵, $m(x)$ 是 A 的最小多项式. 设 $\lambda_0 \in \mathbb{C}$. 则 λ_0 是 A 的特征值 $\iff$ $m(\lambda_0)=0$.

证 充分性是显然的, 因为最小多项式的零点也是特征多项式的零点, 故必为特征值.

反之, 设 λ_0 是 A 的特征值, α 是相应的特征向量, 则 $0=m(A)\alpha=m(\lambda_0)\alpha$. 由于 $\alpha \neq 0$, 故只有 $m(\lambda_0)=0$. □

例 3.1.11 矩阵 A 是幂零矩阵 $\iff$ A 的特征值均为0.

这是因为, $A^m=0 \iff$ A 的最小多项式为 x^k, 其中 $k \leq m$.

例 3.1.12 求下列矩阵的最小多项式:

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

解 A 的特征多项式为 $f(\lambda)=(\lambda-3)(\lambda-2)^3$, 故其最小多项式仅有下述三种可能:

$$(x-3)(x-2);\quad (x-3)(x-2)^2;\quad (x-3)(x-2)^3.$$

直接计算得,

$$(A-3I)(A-2I)\neq 0;\quad (A-3I)(A-2I)^2=0,$$

因此最小多项式为 $m(x)=(x-3)(x-2)^2$.

由上面的几个例子可以推测下面的结论 (证明见习题6):

命题 3.1.4 分块对角矩阵的最小多项式等于各个子块的最小多项式的最小公倍式 (对照: 分块对角矩阵的特征多项式等于各个子块的特征多项式的乘积).

命题 3.1.5 相似矩阵具有相同的最小多项式.

证 设矩阵 A 与 B 相似, 即存在可逆矩阵 P 使得 $B=P^{-1}AP$. 设 A 与 B 的最小多项式分别为 $m_A(x)$ 与 $m_B(x)$, 则 $0=m_B(B)=m_B(P^{-1}AP)=P^{-1}m_B(A)P$, 故 $m_B(A)=0$, 即 $m_B(x)$ 是 A 的零化多项式, 于是 $m_A(x)\,|\,m_B(x)$; 同理, $m_B(x)\,|\,m_A(x)$. 由于它们都是首一多项式, 故 $m_A(x)=m_B(x)$. □

定理 3.1.4 矩阵 A 可以对角化 $\iff$ A 的最小多项式没有重根.

证 必要性由 命题3.1.4 和 命题3.1.5 可得, 因为对角矩阵的最小多项式没有重根, 故与其相似的矩阵的最小多项式也没有重根. 下证充分性. 设 A 的最小多项式 $m(x)=$

$(x-\lambda_1)(x-\lambda_2)\cdots(x-\lambda_s)$ 无重根. 需要证明分块 Schur 三角化定理中的每个对角块都是对角矩阵. 由于相似矩阵具有相同的对角化性质, 故由分块 Schur 三角化定理可设矩阵

$$A = A_1 \oplus A_2 \oplus \cdots \oplus A_s,$$

其中 A_i 是特征值均为 λ_i 的 n_i 阶上三角矩阵, $\sum\limits_{i=1}^{s} n_i = n$. 需要证明每个 A_i 均是对角矩阵 (实际上是纯量矩阵 $\lambda_i I$). 对 s 作归纳. 当 $s = 1$ 时, $m(x) = x - \lambda_1$ 是一次的, 因此 $m(A) = 0 \iff A = \lambda_1 I$, 故定理成立. 假设对不同特征值个数 $< s$ 的矩阵定理成立. 由命题 3.1.4 可知, $B = A_1 \oplus A_2 \oplus \cdots \oplus A_{s-1}$ 的最小多项式是 $m(x)$ 的因式故无重根, 因此 B 是对角矩阵. 由 $m(A) = 0$ 得 $m(B) \oplus m(A_s) = 0$, 所以 $m(A_s) = 0$. 但 A_s 仅有一个特征值, 故再由归纳假设知 A_s 也是对角矩阵, 从而 A 是对角矩阵. □

推论 3.1.1 设 A 为方阵, $f(x)$ 是无重因式的多项式. 若 $f(A) = 0$, 则 A 可以对角化.

例 3.1.13 幂等矩阵与对合矩阵均可以对角化.

这是因为幂等矩阵 A 满足 $A^2 = A$, 故其最小多项式 $m(x) \mid (x^2 - x)$, 故无重根; 而对合矩阵 A 满足 $A^2 = I$, 从而其最小多项式 $m(x) \mid (x^2 - 1)$, 也无重根. 因此, 这两类矩阵均可以对角化.

以下我们简要介绍线性变换的特征值与特征向量.

设 V 是 $\mathbb{C}$ 上的线性空间 (可以是无限维), $\sigma \in \mathrm{End}V$, $\lambda \in \mathbb{C}$. 如果存在非零向量 $\alpha \in V$ 使得 $\sigma(\alpha) = \lambda\alpha$, 则称 λ 是 σ 的一个特征值, 而 α 是 σ 的一个 (属于特征值 λ 的) 特征向量.

如果 λ 是 σ 的特征值, 则 V 的子集合 $\{x \in V \mid \sigma(x) = \lambda x\}$ 称为 σ 的 (属于特征值 λ 的) 特征子空间, 记为 V_λ. 这个子集合的确是 V 的子空间, 因为

$$V_\lambda = \mathrm{Ker}(\lambda I - \sigma). \tag{3.1.3}$$

上述定义与矩阵的特征值和特征向量完全一致. 从线性变换的角度看, 特征向量是在线性变换下不改变方向的那些非零向量, 而特征值则是特征向量的伸缩系数. 注意线性变换在其每个特征子空间上的限制是一个位似, 因此特征值与特征向量具有简化线性变换的矩阵表示的作用. 确切地说, 如果有限维线性空间 V 有一组基其基向量全部为 σ 的特征向量, 则 σ 在这组基下的矩阵是对角矩阵. 故有下述定理 (详细证明见习题 7).

定理 3.1.5 设 V 是有限维线性空间, $\sigma \in \mathrm{End}V$, A 是 σ 在某组基下的矩阵. 则

(1) A 与 σ 有完全相同的特征值 (即重数也一样);

(2) 设 σ 的不同特征值为 $\lambda_1, \cdots, \lambda_s$, 则 A 可以对角化 $\iff V = \sum\limits_{i=1}^{s} \oplus V_{\lambda_i}$. 于是 A 可以对角化 $\iff \sigma$ 可以对角化, 即 $\sigma = \sum\limits_{i=1}^{s} \oplus \sigma_i$, 其中 σ_i 是 σ 在 V_{λ_i} 上的限制.

从线性变换的角度理解特征值与特征向量较为直观. 比如考虑旋转矩阵 (设 $\theta \neq 0, \pi$)

$$A = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix}$$

的特征值与特征向量问题. 如果仅从矩阵的角度出发, 那就除计算别无他途. 但是, 矩阵 A 是平面上的旋转, 因此它不改变任何向量的长度, 故其特征值的模只能是1, 特别其实特征值只能是 ± 1. 另一方面, 该旋转改变了任何一个非零向量的方向, 因此它没有实特征向量. 所以将矩阵的特征值与特征向量问题转化为线性变换的相应问题常常是有效的方法.

思考题

1. 实数域上的Schur三角化定理成立吗, 即每个实方阵是否可以正交三角化?
2. 是否每个矩阵都可以分块酉三角化, 即分块Schur三角化定理中的可逆矩阵是否可以加强为酉矩阵?
3. 设 A, B 为同阶方阵, 则由降幂公式知 AB 与 BA 有相同的特征多项式, 它们是否相似?
4. 特征多项式与最小多项式的商多项式有何意义?
5. 如果一个线性变换 σ 的特征值的模均小于1, σ 有何特点?
6. 如果一个线性变换 σ 有一组正交的特征向量, σ 有何特点?

第二节　Jordan标准形: 复数矩阵的一种最简形式

分块Schur三角化定理 (定理3.1.2) 将寻求矩阵标准形的问题化为寻找主对角元素均相同的上三角矩阵的标准形. 这样的矩阵可以写成

$$A = \lambda I + N$$

的形式, 其中 N 是严格上三角矩阵. 因为

$$P^{-1}AP = \lambda I + P^{-1}NP,$$

故只需研究严格上三角矩阵的标准形即可.

例 3.2.1　设 A 是 n 阶严格上三角矩阵且 $r(A) = n-1$, 则 A 与 J_n 相似.

证　将 A 写成分块矩阵的形式有 $A = \begin{pmatrix} 0 & B \\ 0 & \mathbf{0}^T \end{pmatrix}$, 其中 B 是 $n-1$ 阶方阵, 则因 $r(B) = r(A) = n-1$ 知 B 可逆, 因此取 $P = \begin{pmatrix} 1 & \mathbf{0}^T \\ 0 & B^{-1} \end{pmatrix}$, 则有 $P^{-1}AP = J_n$.

例 3.2.2　设 $A = \begin{pmatrix} 0 & J_3 \\ 0 & \mathbf{0}^T \end{pmatrix}$ 是4阶严格上三角矩阵, 注意 A 不与 $B = \begin{pmatrix} J_3 & 0 \\ \mathbf{0}^T & 0 \end{pmatrix}$ 相似, 而与 $C = J_2 \oplus J_2$ 相似.

对 $2 \le r(A) \le n-2$ 的 n 阶严格上三角矩阵 A 来说, 其标准形并不能由其秩完全确定. 比如上例中的矩阵 B 与 C 均为秩2的最简形式. 仔细分析 B 与 C 的结构可知, 它们的一个差别是 $B^2 \neq 0$ 而 $C^2 = 0$. 因此严格上三角矩阵的幂零指数对其标准形有重要意义. 下面列出 J_n 的一些性质 (证明见习题16).

命题 3.2.1　(1) $J_n^T J_n = (0) \oplus I_{n-1}$, $J_n J_n^T = I_{n-1} \oplus (0)$;
(2) $J_n e_i = e_{i-1}$, 其中 e_i 是标准向量, 约定 e_0 是0向量;
(3) 对任意 $x \in \mathbb{C}^n$, $(I_n - J_n^T J_n)x = (x^T e_1)e_1$.

现在可以证明严格上三角矩阵的Jordan标准形定理了.

定理 3.2.1 (幂零矩阵的Jordan标准形定理[28]) 设 A 是 n 阶严格上三角复矩阵. 则存在可逆矩阵 P 和正整数 $n_1 \geq n_2 \geq \cdots \geq n_m$, $n_1 + n_2 + \cdots + n_m = n$, 使得

$$P^{-1}AP = J_{n_1} \oplus J_{n_2} \oplus \cdots \oplus J_{n_m}. \tag{3.2.1}$$

上式右端的矩阵称为**幂零Jordan矩阵**, 它是矩阵 A 的Jordan标准形, 一般记为 J. 如果不计诸 n_j 的次序与大小, 则 A 的Jordan标准形是唯一的.

证 首先我们指出, 任意调整公式 (3.2.1) 右端的每个直和因子的顺序, 得到的矩阵均是相似的, 因此条件

$$n_1 \geq n_2 \geq \cdots \geq n_m, n_1 + n_2 + \cdots + n_m = n$$

总是可以满足的, 故只需证明公式 (3.2.1) 及唯一性.

对 A 的阶数 n 作归纳. 当 $n = 1$ 时, $A = (0)$, 定理显然成立. 假设对阶数 $< n$ 的所有严格上三角矩阵, 定理成立. 现将 A 写成分块矩阵

$$\begin{pmatrix} 0 & \alpha^T \\ 0 & A_1 \end{pmatrix},$$

其中 $\alpha \in \mathbb{C}^{n-1}$ 而 A_1 是 $n-1$ 阶严格上三角矩阵. 由归纳假设, 存在 $n-1$ 阶可逆矩阵 P_1 使得 $P_1^{-1}A_1P_1$ 具有公式 (3.2.1) 的形式, 即

$$A_2 = P_1^{-1}A_1P_1 = J_{k_1} \oplus J_{k_2} \oplus \cdots \oplus J_{k_s},$$

其中 $k_1 \geq k_2 \geq \cdots \geq k_s$, $k_1 + k_2 + \cdots + k_s = n - 1$. 令

$$J = J_{k_2} \oplus \cdots \oplus J_{k_s}$$

并取 $P_2 = (1) \oplus P_1$, 则 $A_2 = J_{k_1} \oplus J$ 且

$$A_3 = P_2^{-1}AP_2 = \begin{pmatrix} 0 & \alpha^T P_1 \\ 0 & A_2 \end{pmatrix} = \begin{pmatrix} 0 & \alpha_1^T & \alpha_2^T \\ 0 & J_{k_1} & 0 \\ 0 & 0 & J \end{pmatrix},$$

其中 $(\alpha_1^T \ \ \alpha_2^T)$ 是矩阵 $\alpha^T P_1$ 的行分块矩阵. 取

$$P_3 = \begin{pmatrix} 1 & \alpha_1^T J_{k_1}^T & 0 \\ 0 & I & 0 \\ 0 & 0 & I \end{pmatrix},$$

则利用 $x \in \mathbb{C}^n$, $(I_n - J_n^T J_n)x = (x^T e_1)e_1$(见 命题3.2.1(3)) 可得

$$A_4 = P_3^{-1}A_3P_3 = \begin{pmatrix} 0 & (\alpha_1^T e_1)e_1^T & \alpha_2^T \\ 0 & J_{k_1} & 0 \\ 0 & 0 & J \end{pmatrix} \tag{3.2.2}$$

[28]该定理及其一般形式即本节 定理3.2.2 由 Jordan 于 1870 年证明.

现若 $\alpha_1^T e_1 = 0$, 则 $A_4 = \begin{pmatrix} 0 & 0 & \alpha_2^T \\ 0 & J_{k_1} & 0 \\ 0 & 0 & J \end{pmatrix}$ 相似于矩阵 $\begin{pmatrix} J_{k_1} & 0 & 0 \\ 0 & 0 & \alpha_2^T \\ 0 & 0 & J \end{pmatrix}$. 再由归纳假设可知矩阵 $\begin{pmatrix} 0 & \alpha_2^T \\ 0 & J \end{pmatrix}$ 相似于一个幂零 Jordan 矩阵 J'. 因此 A 相似于 $J_{k_1} \oplus J'$, 定理成立.

如果数 $a = \alpha_1^T e_1 \neq 0$, 则取可逆矩阵 $P_4 = (a) \oplus I \oplus (aI)$, 直接计算 (见习题 17) 可得

$$A_5 = P_4^{-1} A_4 P_4 = \begin{pmatrix} J'' & e_1 \alpha_2^T \\ 0 & J \end{pmatrix} \tag{3.2.3}$$

其中 $J'' = \begin{pmatrix} 0 & e_1^T \\ 0 & J_{k_1} \end{pmatrix} = J_{k_1+1}$. 由 命题 3.2.1(2) 知, $J'' e_{i+1} = e_i$, 故

$$\begin{pmatrix} I & e_2 \alpha_2^T \\ 0 & I \end{pmatrix} \begin{pmatrix} J'' & e_1 \alpha_2^T \\ 0 & J \end{pmatrix} \begin{pmatrix} I & -e_2 \alpha_2^T \\ 0 & I \end{pmatrix} = \begin{pmatrix} J'' & e_2 \alpha_2^T J \\ 0 & J \end{pmatrix}. \tag{3.2.4}$$

注意上式右端矩阵的右上角与原矩阵相比, 由 $e_1 \alpha_2^T$ 变为 $e_2 \alpha_2^T J$, 因此重复上述步骤 (见习题 18), 可使式 (3.2.4) 右端矩阵的右上角变为 $e_2 \alpha_2^T J^N$, 其中 N 为任意正整数. 而由 k_1 的选取知 $J^{k_1} = 0$, 故 A_5 相似于矩阵 $J'' \oplus J$, 而这是一个幂零 Jordan 矩阵.

再证唯一性. 仍对 n 作归纳. $n = 1$ 时唯一性显然成立. 假设对一切阶数 $< n$ 的严格上三角矩阵唯一性成立, 考虑 n 阶矩阵 A 的两个 Jordan 标准形

$$J = J_{n_1} \oplus J_{n_2} \oplus \cdots \oplus J_{n_s}$$

与

$$\tilde{J} = J_{m_1} \oplus J_{m_2} \oplus \cdots \oplus J_{m_t}$$

其中 $n_1 \geq n_2 \geq \cdots \geq n_s \geq 1$, $m_1 \geq m_2 \geq \cdots \geq m_t \geq 1$. 则 $n_1 = m_1$, 否则, 不妨设 $n_1 > m_1$, 则 $J^{m_1} \neq 0$ 而 $\tilde{J}^{m_1} = 0$, 矛盾! 现由第二章, 推论 2.5.3, 知 $J_{n_2} \oplus \cdots \oplus J_{n_s}$ 与 $J_{m_2} \oplus \cdots \oplus J_{m_t}$ 相似, 由归纳假设, 它们具有相同的 Jordan 标准形, 因此 $s = t$ 且 $n_2 = m_2, \cdots, n_s = m_s$. □

由此定理即可得到判断两个幂零矩阵相似的下述准则 (证明见习题 16):

推论 3.2.1 设 M 与 N 是两个 n 阶幂零矩阵. 则 M 与 N 相似 $\iff$ $r(M^k) = r(N^k), \forall k \geq 1$.

定义 3.2.1 矩阵 $\lambda I_n + J_n$ 称为 n 阶 λ-Jordan 块, 记为 $J_n(\lambda)$.

由分块 Schur 三角化定理以及幂零矩阵的 Jordan 标准形定理可得一般矩阵的 Jordan 标准形定理.

定理 3.2.2 (Jordan 标准形定理) 设 A 是 n 阶复矩阵, 则存在可逆矩阵 P 和正整数 $n_1 \geq n_2 \geq \cdots \geq n_m$, $n_1 + n_2 + \cdots + n_m = n$, 使得

$$P^{-1} A P = J_{n_1}(\lambda_1) \oplus J_{n_2}(\lambda_2) \oplus \cdots \oplus J_{n_m}(\lambda_m). \tag{3.2.5}$$

上式右端的矩阵称为 **Jordan 矩阵**, 它是矩阵 A 的 Jordan 标准形. 如果不计诸 n_j 的次序与大小, 则 A 的 Jordan 标准形是唯一的.

注 1. 公式 (3.2.5) 中的诸特征值 λ_i 可能相同.
注 2. 定理 3.2.2 仅对复数域上的矩阵成立, 见本节思考题 1 与习题 33.

推论 3.2.2 方阵 A 可以对角化 $\iff$ A 的 Jordan 标准形是对角矩阵.

一般而言, 确定矩阵的 Jordan 标准形是极为困难的问题, 因为基本上不可能求出高阶矩阵的全部特征值. 另外, 即使是幂零矩阵, 定理 3.2.1 的证明中给出的算法也是不能付诸实践的, 因为该算法是不稳定的. 参考下面的例子.

例 3.2.3 设 $A_t = \begin{pmatrix} t & 1 \\ 0 & 0 \end{pmatrix}$. 则当 $t \neq 0$ 时, 其 Jordan 标准形为对角矩阵 $D = \begin{pmatrix} t & 0 \\ 0 & 0 \end{pmatrix}$. 但当 $t = 0$ 时, 其 Jordan 标准形为 $A = J_2$. 因此在 $t \to 0$ 的极限过程中, A_t 的 Jordan 标准形并不收敛于其极限的 Jordan 标准形.

上面的例子表明, 矩阵的 Jordan 标准形不具有稳定的数值方法 (这是因为 Jordan 标准形不是矩阵的各个元素的连续函数), 但它在理论上仍是极其重要的, 因为一个矩阵的 Jordan 标准形不仅具有非常简单的形式, 而且包含了该矩阵的几乎所有信息, 比如秩, 特征值, 线性无关的特征向量的个数, 特征子空间的维数等等.

思考题

1. 设 $A \in M_n(\mathbb{Q})$. 是否存在可逆矩阵 $P \in M_n(\mathbb{Q})$ 使得 $P^{-1}AP$ 为 Jordan 标准形? 将 $\mathbb{Q}$ 换成 $\mathbb{R}$ 又如何?
2. 分块矩阵 $\begin{pmatrix} A & B \\ B & A \end{pmatrix}$ 的 Jordan 标准形与 A, B 的 Jordan 标准形有何关系? 特征值有何联系? 特别讨论 $A = 0$ 与 $A = B$ 的情形.
3. 仿照幂零矩阵相似的判别准则 (即 推论 3.2.1) 给出两个同阶矩阵相似的判别准则. 是否能够判断该准则与幂零矩阵相似的判别准则哪个更有意义?

第三节 Jordan 标准形的计算

定理 3.2.1 与 定理 3.2.2 虽然给出了矩阵的 Jordan 标准形的存在性与唯一性, 但是都没有给出 Jordan 标准形的具体数据, 比如一共有多少个对角块? 各阶对角块有几个? 等等. 下面的命题给出了幂零矩阵的 Jordan 标准形的最大块的阶数.

命题 3.3.1 (最大 Jordan 块) 设 A 是严格上三角矩阵, 则其 Jordan 标准形的 Jordan 块的阶数的最大值等于其幂零指数.

证 设 A 的幂零指数为 e, 其 Jordan 标准形为 J, J 的 Jordan 块的阶数的最大值为 f. 则因 $A^e = 0$ 知 $J^e = 0$, 从而 $f \leq e$. 同理, 由 $J^f = 0$ 可知 $A^f = 0$, 故 $e \leq f$. $\square$

例 3.3.1 试证下列矩阵为幂零矩阵, 并求其 Jordan 标准形:

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

解 直接计算便知 $A^2 = 0$. 因此 A 为幂零矩阵, 且幂零指数为 2. 故由 命题 3.3.1 可知, A 的 Jordan 标准形 J 的最大 Jordan 块为 2 阶的. 由于 $r(A) = 2$, 故 A 的零度为 2, 即有 2 个块, 所以 $J = J_2 \oplus J_2$.

上面的例子说明, 对阶数较小的幂零矩阵可以通过其幂零指数和秩来确定其 Jordan 标准形. 但对高阶矩阵, 则需要更为精细的方法, 我们将其列为下面的定理, 证明见习题 20.

定理 3.3.1 设 n 阶严格上三角矩阵 A 的 Jordan 标准形 J 同公式 (3.2.1), 其幂零指数为 e. 则

(1) $e = \max\{n_i \,|\, 1 \le i \le m\}$;

(2) J 中 Jordan 块的个数 m 等于 A 的零度;

(3) 记 J 中 k 阶 Jordan 块的个数为 ℓ_k, A^k 的零度为 η_k, $0 \le k \le m$. 则

$$\ell_k = 2\eta_k - \eta_{k-1} - \eta_{k+1}, \quad 1 \le k \le m. \tag{3.3.1}$$

注 1. 计数公式 (3.3.1) 来源于数列

$$\eta_0 = 0, \eta_1, \eta_2, \cdots, \eta_{k-1}, \eta_k, \eta_{k+1}, \cdots$$

的二阶差分, 即 $\ell_k = (\eta_k - \eta_{k-1}) - (\eta_{k+1} - \eta_k)$.

注 2. 上面定理中的 (2) 等价于说 J 中 Jordan 块的个数等于 A 的线性无关的特征向量的个数. 这是因为每个 Jordan 块的秩恰好比其阶数小 1, 即零度为 1, 因此对应于 1 个线性无关的特征向量.

例 3.3.2 试证下列矩阵为幂零矩阵, 并求其 Jordan 标准形:

$$A = \begin{pmatrix} 2 & 6 & 5 & 0 & 0 & 4 & 0 & 6 \\ 1 & 3 & 2 & 0 & 0 & 2 & 0 & 3 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ -1 & -1 & -2 & 0 & -1 & -4 & 0 & -1 \\ 2 & 5 & 4 & 0 & 0 & 4 & 0 & 5 \\ -1 & -3 & -2 & 0 & 0 & -2 & 0 & -3 \\ 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 \\ -1 & -3 & -2 & 0 & 0 & -2 & 0 & -3 \end{pmatrix}.$$

解 直接计算可知 $A^3 = 0$. 因此 A 为幂零矩阵, 且幂零指数为 3. 再计算可得 $r(A) = 5$. 故由 定理 3.3.1 可知, A 的 Jordan 标准形 J 共有 3 块, 且最大 Jordan 块为 3 阶的. 因此其它两块必为一个 3 阶块和一个 2 阶块, 故知 $J = J_3 \oplus J_3 \oplus J_2$.

例 3.3.3 设 A 是 例 3.3.1 中的矩阵, 求可逆矩阵 P, 使得 $P^{-1}AP = J$.

解 设 $P = (\alpha_1, \alpha_2, \alpha_3, \alpha_4)$, 可得

$$A(\alpha_1, \alpha_2, \alpha_3, \alpha_4) = (\alpha_1, \alpha_2, \alpha_3, \alpha_4)(J_2 \oplus J_2) = (0, \alpha_1, 0, \alpha_3).$$

由此可得 P 的各个列向量应满足的方程组分别为

$$A\alpha_1 = 0, \qquad A\alpha_2 = \alpha_1, \qquad A\alpha_3 = 0, \qquad A\alpha_4 = \alpha_3. \tag{3.3.2}$$

注意到四个方程组的系数矩阵都相同, 可以采用下面的方式统一求解:

$$\left(\begin{array}{cccc|c} 1 & 1 & 0 & 1 & b_1 \\ 0 & -1 & -1 & -1 & b_2 \\ 1 & 1 & 0 & 1 & b_3 \\ -1 & 0 & 1 & 0 & b_4 \end{array}\right) \to \left(\begin{array}{cccc|c} 1 & 0 & -1 & 0 & b_1 + b_2 \\ 0 & 1 & 1 & 1 & -b_2 \\ 0 & 0 & 0 & 0 & -b_1 + b_3 \\ 0 & 0 & 0 & 0 & b_1 + b_2 + b_4 \end{array}\right)$$

由此可知要使方程组 $Ax=\beta$ 有解, 向量 $\beta=(b_1,b_2,b_3,b_4)^T$ 要满足

$$b_1=b_3,\qquad b_1+b_2+b_4=0.$$

解方程组

$$\begin{cases} x_1 \quad -x_3 \qquad = 0, \\ \quad\; x_2 + x_3 + x_4 = 0, \end{cases}$$

得 $\alpha_1=(1,-1,1,0)^T$, $\alpha_3=(0,1,0,-1)^T$. 这两个向量都满足 $Ax=\beta$ 的相容性条件. 解 $Ax=\alpha_1$, 即

$$\begin{cases} x_1 \quad -x_3 \qquad = 0, \\ \quad\; x_2 + x_3 + x_4 = 1, \end{cases}$$

得 $\alpha_2=(1,0,1,0)^T$. 解 $Ax=\alpha_3$, 即

$$\begin{cases} x_1 \quad -x_3 \qquad = 1, \\ \quad\; x_2 + x_3 + x_4 = -1, \end{cases}$$

得 $\alpha_4=(1,0,0,-1)^T$. 因此,

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

方程 (3.3.2) 中的向量 α_1 与 α_3 显然是属于特征值 0 的特征向量, 而向量 α_2 与 α_4 不是特征向量, 它们实际上满足方程 (请验证!)

$$A^2x=0. \tag{3.3.3}$$

一般地, 设 λ 是 n 阶矩阵 A 的特征值, 则将方程 $(A-\lambda I)^kx=0$(k 是任意正整数) 的非零解称为 A 的 (属于特征值 λ 的)**广义特征向量**, 可以证明 (见习题 21-24) 广义特征向量和零向量 (即所有方程 $(A-\lambda I)^kx=0$ 的解空间的并集) 构成 $\mathbb{C}^n$ 的一个子空间, 称为 A 的 (属于特征值 λ 的) **广义特征子空间或根子空间**, 记为 E_λ. 证明矩阵的 Jordan 标准形的存在性的另一个途径即是利用 $\mathbb{C}^n=\sum\limits_{\lambda\in\sigma(A)}\oplus E_\lambda$. 习题 21-24 是这种途径的具体实现, 有兴趣的读者可以尝试完成之. 证明矩阵的 Jordan 标准形的存在性的第三条途径是使用 λ- 矩阵的 Smith [29] 标准形, 我们在习题 56-60 中介绍了这个方法.

结合 Schur 分块三角化定理, 即可由 定理 3.3.1 得到一般矩阵的 Jordan 标准形的结构.

定理 3.3.2 设 n 阶矩阵 A 的 Jordan 标准形 J 同公式 (3.2.5). 设 μ 为 A 的一个特征值, 记 $(A-\mu I)^k$ 的零度为 η_k, J 中对角线元素为 μ 的 k 阶 Jordan 块的个数为 ℓ_k, 则

(1) η_1 等于 J 中对角线元素为 μ 的 Jordan 块的个数;

(2) $\ell_1=2\eta_1-\eta_2$, $\ell_k=2\eta_k-\eta_{k-1}-\eta_{k+1}$, $k\geq 2$.

注意, 与幂零矩阵的相应结果 (定理 3.3.1) 对照, 上面的定理没有太大实用价值 (对照上一节的思考题 3). 因为 n 阶矩阵 A 可以通过 (计算机) 验证 $A^n=0$ 来判断其幂零性后再使用 定理 3.3.1, 但非幂零矩阵在使用 定理 3.3.2 前需要先求出 A 的所有特征值, 而正如先前指出的, 这个工作的难度几乎是不能克服的.

[29]Henry John Stephen Smith(1826-1883), 爱尔兰 - 英国数学家.

例 3.3.4 求下列矩阵的 Jordan 标准形 J, 并求变换矩阵 P, 使 $P^{-1}AP = J$:

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

解 易知 $|\lambda I - A| = (\lambda - 1)^4$, 故 A 的 Jordan 标准形 J 的对角元素均为 1. 再计算 $A - I$ 的幂零度为 3, 故最大子块是 3 阶的, 因此

$$J = J_3(1) \oplus (1).$$

设 $P^{-1}AP = J$. 令 $P = (\alpha_1, \alpha_2, \alpha_3, \alpha_4)$, 则由 $AP = PJ$ 得

$$A(\alpha_1, \alpha_2, \alpha_3, \alpha_4) = (\alpha_1, \alpha_2, \alpha_3, \alpha_4)(J_3(1) \oplus (1)) = (\alpha_1, \alpha_1 + \alpha_2, \alpha_2 + \alpha_3, \alpha_4).$$

于是得到四个方程组:

$$A\alpha_1 = \alpha_1, \quad A\alpha_2 = \alpha_1 + \alpha_2, \quad A\alpha_3 = \alpha_2 + \alpha_3, \quad A\alpha_4 = \alpha_4,$$

即

$$(I - A)\alpha_1 = 0, \quad (I - A)\alpha_2 = -\alpha_1, \quad (I - A)\alpha_3 = -\alpha_2, \quad (I - A)\alpha_4 = 0.$$

解之得 $\alpha_1 = (0, 0, 1, -1)^T$, $\alpha_2 = (1, -2, 0, 0)^T$, $\alpha_3 = (-1, 3, 0, 0)^T$, $\alpha_4 = (-1, 2, 1, 0)^T$.

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

当然, P 不是唯一的 (为什么?).

例 3.3.5 求下列矩阵的 Jordan 标准形:

$$A = \begin{pmatrix} 3 & -1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 5 & -3 & -1 & 0 & 0 & 0 & 0 & 0 \\ -8 & 7 & 3 & 0 & 0 & 0 & 0 & 0 \\ 8 & 0 & 2 & -1 & 1 & 0 & 0 & 0 \\ 5 & -11 & -6 & 0 & 0 & 1 & 0 & 0 \\ 4 & 8 & 5 & 0 & -1 & -2 & 0 & 0 \\ 6 & 13 & 8 & -4 & -1 & -4 & 2 & 1 \\ 14 & -12 & -4 & -3 & -1 & 1 & 0 & 2 \end{pmatrix}.$$

解 容易看出 A 是一个分块下三角矩阵. 因此 A 的特征多项式为

$$|\lambda I - A| = \begin{vmatrix} \lambda - 3 & 1 & 0 \\ -5 & \lambda + 3 & 1 \\ 8 & -7 & \lambda - 3 \end{vmatrix} \cdot \begin{vmatrix} \lambda + 1 & -1 & 0 \\ 0 & \lambda & -1 \\ 0 & 1 & \lambda + 2 \end{vmatrix} \cdot \begin{vmatrix} \lambda - 2 & -1 \\ 0 & \lambda - 2 \end{vmatrix}$$
$$= (\lambda - 1)^3(\lambda + 1)^3(\lambda - 2)^2.$$

注意3重特征值1均来自左上角的主子矩阵 A_1(前三行前三列), 3重特征值-1均来自第 4-6 行与列构成的主子矩阵 A_2, 而两重特征值2则来自后两行两列构成的主子矩阵 A_3. 因此由分块Schur三角化定理知存在严格上三角矩阵 N_i 使得 A 相似于分块对角矩阵

$$B = (I_3 + N_1) \oplus (-I_3 + N_2) \oplus (2I_2 + N_3),$$

其中 N_i 相似于 $A_i - \lambda_i I$. 由于 $A_1 - I$ 的零度为1, 故 N_1 相似于 J_3, 即对角元素为1的Jordan块只有1块. 类似地, $-A-I$ 与 $A-2I$ 的零度都为1. 故 N_2 与 N_3 分别相似于 J_3 与 J_2. 因此 A 的Jordan标准形为

$$J = J_3(1) \oplus J_3(-1) \oplus J_2(2).$$

变换矩阵的求解见习题26.

例 3.3.6 试求秩为1的 n 阶方阵 A 的Jordan标准形.

解 设 $r(A) = 1$. 则 A 的Jordan标准形只有一个位置不为零. 所以 A 至少有 $n-1$ 个特征值都是0, 而另一个特征值必为 $\operatorname{tr} A$. 因此 A 的Jordan标准形为:

(1) $\operatorname{diag}(\operatorname{tr} A, 0, \cdots, 0)$ (若 $\operatorname{tr} A \neq 0$), 或

(2) $\begin{pmatrix} 0 & 1 & & \\ & 0 & & \\ & & \ddots & \\ & & & 0 \end{pmatrix}$ (若 $\operatorname{tr} A = 0$, 此时 A 幂零, 且幂零指数为2).

思考题

1. 两个矩阵的和与积的Jordan标准形是否等于它们的Jordan标准形的和与积?
2. 如果 P 与 Q 均为Jordan标准形中的变换矩阵, 它们之间有何关系?

第四节 盖尔圆定理: 特征值的估计

对高阶矩阵来说, 精确计算矩阵的特征值一般是做不到的(1825年左右, Abel[30]与Galois[31]等人证明了5次及5次以上的代数方程无公式解). 幸运的是, 工程实践中的众多问题常常只需要确定特征值是否满足一定精度, 比如线性系统的稳定性仅需要其所有特征值的实部小于0, 即所有特征值均位于左半平面, 而在第五章中我们将看到矩阵级数 $\sum\limits_{i=0}^{\infty} a_i A^i$ 的敛散性也只需了解 A 的特征值的模是否均小于1. 因此更为合理的办法是对特征值的范围作出估计. 复数域上 n 阶矩阵的特征值可以用复平面上的点来表示, 因此对这些点的位置的估计就是对特征值的估计.

设 $A = (a_{ij})$ 是 n 阶复数矩阵. 在复平面上, 称集合

$$D_i(A) = \{x \in \mathbb{C} \,|\, |x - a_{ii}| \le \sum_{j \neq i} |a_{ij}|\} \quad (i = 1, 2, \cdots, n)$$

为矩阵 A 的第 i 个圆盘. 将 $\sum_{j\neq i} |a_{ij}|$ 记为 $R_i(A)$, 称为 A 的去心绝对行和, 则矩阵 A 的第 i 个圆盘又可写成

$$D_i(A) = \{x \in \mathbb{C} \,|\, |x - a_{ii}| \le R_i(A)\}.$$

[30]Niels Henrik (1802-1829), 著名挪威数学家, 月球上有以其名字命名的火山, 以其名字命名的Abel Prize是世界上奖金 (约一百万美元) 最高的数学奖.

[31]Évariste Galois(1811-1832), 天才的法国数学家, 群论及有限域的创始人.

一般地, 称 A 的所有圆盘的并形成的区域 $\cup_{i=1}^{n} D_i(A)$ 为 A 的 (关于行的) **Gerschgorin**[32] **区域**, 记为 $G(A)$, 而将 $G(A)$ 中的每一个圆盘称为**Gerschgorin 圆盘**或**盖尔圆盘** , 这些圆盘的边界称为Gerschgorin圆或盖尔圆 .

例 3.4.1 求 A 的圆盘, 其中

$$A = \begin{pmatrix} 0 & 1 & -1 \\ 1 & 2 & 0 \\ \mathrm{i} & 1 & 2\mathrm{i} \end{pmatrix}.$$

解 $D_1(A) = \{x \mid |x| \le 2\},\ D_2(A) = \{x \mid |x-2| \le 1\},\ D_3(A) = \{x \mid |x - 2\mathrm{i}| \le 2\}$.

例 3.4.2 由于盖尔圆盘只涉及矩阵元素的模或绝对值, 因此如果矩阵 $A = (a_{ij})$ 的对角元素均为非负实数, 则 A 与其绝对值矩阵 $|A| = (|a_{ij}|)$ 具有完全相同的盖尔圆盘和盖尔区域.

定理 3.4.1 (盖尔圆盘定理[33]) 设 A 是 n 阶矩阵, 则它的特征值 λ 至少满足下列不等式之一:

$$|\lambda - a_{ii}| \le \sum_{j \ne i}^{n} |a_{ij}|,\ i = 1, 2, \cdots, n.$$

等价地, $\sigma(A) \subseteq G(A)$, 即 A 的每个特征值都落在 A 的某个圆盘之内.

证 设 λ 是 A 的一个特征值, $\alpha = (x_1, \cdots, x_n)^T$ 是它的一个特征向量. 则 $A\alpha = \lambda\alpha$, 即

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = \lambda x_1, \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = \lambda x_2, \\ \qquad\qquad\vdots \\ a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n = \lambda x_n. \end{cases} \qquad (*)$$

令 $\max\{|x_1|, \cdots, |x_n|\} = |x_m|$, 则由 $x \ne 0$ 知 $x_m \ne 0$. 将方程组 (*) 中的第 m 个方程改写为

$$(\lambda - a_{mm})x_m = \sum_{j \ne m} a_{mj}x_j.$$

两边取模得

$$|\lambda - a_{mm}||x_m| \le \sum_{j \ne m} |a_{mj}||x_j| \le |x_m| \sum_{j \ne m} |a_{mj}|,$$

所以

$$|\lambda - a_{mm}| \le \sum_{j \ne m} |a_{mj}|.$$

即特征值 λ 落在第 m 个圆盘内. □

注 对 A^T 应用圆盘定理, 可以得到 A 的关于列的相应结果.

[32]Semyon Aranovich Gershgorin(1901-1933), 前苏联 (现白俄罗斯) 数学家.
[33]Gerschgorin 于 1931 年证明了此定理.

例 3.4.3 求 A 的圆盘并估计其特征值, 其中

$$A = \begin{pmatrix} 10 & 1 & 2 & 3 \\ 5 & 30 & 1 & 2 \\ 10 & 3 & -10 & 5 \\ 2 & 3 & 1 & -40 \end{pmatrix}.$$

解 矩阵 A 的圆盘为

$$\begin{aligned} |\lambda - 10| &\le 1+2+3 &= 6, \\ |\lambda - 30| &\le 1+5+2 &= 8, \\ |\lambda + 10| &\le 10+3+5 &= 18, \\ |\lambda + 40| &\le 2+3+1 &= 6. \end{aligned}$$

如下面的示意图:

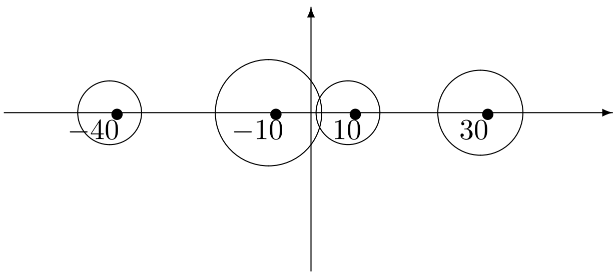

图 3.4.1

因此, 四个特征值均落在四个圆盘之中.

例 3.4.4 设 n 阶矩阵 A 满足对角强优条件 (亦称“严格对角占优”条件)

$$|a_{ii}| > \sum_{j\ne i} |a_{ij}|, \ i = 1, 2, \cdots, n,$$

证明: $|A| \ne 0$.

证 只需证明 A 的特征值全部不为 0. 设 λ 是 A 的一个特征值, 则由圆盘定理, 它必然落在某个圆盘之内, 即存在 k, 使得 $|\lambda - a_{kk}| \le \sum_{j\ne k} |a_{kj}|$. 如果 $\lambda = 0$, 则有 $|a_{kk}| \le \sum_{j\ne k} |a_{kj}|$, 矛盾! 故 A 不可能有零特征值, 从而 A 的行列式不等于 0. □

请注意, 圆盘定理只是说明, 每个特征值必然落在一个圆盘内, 但并未指明落在哪个圆盘内. 因此, 有些圆盘可能不含特征值, 而另一些则可能包含多个特征值.

例 3.4.5 求 A 的圆盘, 其中

$$A = \begin{pmatrix} 10 & -8 \\ 5 & 0 \end{pmatrix}.$$

解 直接计算可知, 矩阵 A 的特征值为

$$\lambda_1 = 5 + \sqrt{15}\mathrm{i},\ \lambda_2 = 5 - \sqrt{15}\mathrm{i}.$$

而 A 的圆盘为

$$|\lambda - 10| \le 8, \quad |\lambda| \le 5.$$

如下图:

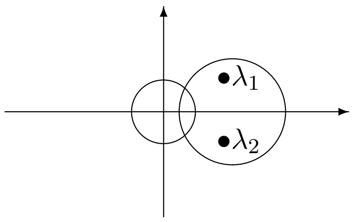

图 3.4.2

因此, 两个特征值全部落在圆盘 $|\lambda - 10| \le 8$ 内, 而在圆盘 $|\lambda| \le 5$ 之外!

实际上有下述一般性的结论:

命题 3.4.1 (精细圆盘定理) 设 C 是盖尔区域的一个由 k 个圆盘组成的连通分支, 则 C 恰好含有 k 个特征值.

证 盖尔区域本身可以是不连通的, 所谓连通分支是指其最大的连通部分. 我们以下给出的证明基于一个基本事实, 即复系数多项式的根是其系数的连续函数. 记 $A = (a_{ij})$ 的对角元素构成的对角矩阵为 D, 令 $B = A - D$. 考察矩阵 $A(\varepsilon) = D + \varepsilon B$, 其中 $0 \le \varepsilon \le 1$. 显然 $A(0) = D$, $A(1) = A$. 注意 $A(\varepsilon)$ 的特征多项式是 ε 的多项式, 因此 $A(\varepsilon)$ 的特征值是 ε 的连续函数. 由圆盘定理, 对任意 ε, 矩阵 $A(\varepsilon)$ 的特征值均位于以 a_{ii} 为圆心, 半径为 $\varepsilon R_i(A)$ 的圆盘之内. 当 ε 从 0 连续地变到 1 时, 特征值也连续地变化.

不妨设 C 由前 k 个圆盘组成, 由于 C 是一个连通分支, 故它与 A 的其它 $n-k$ 个圆盘是分离的. 因此对于每个 $\varepsilon \in [0,1]$, 矩阵 $A(\varepsilon)$ 的盖尔区域也具有相同的性质. 但当 $\varepsilon = 0$ 时, 特征值是 $a_{11}, a_{22}, \cdots, a_{nn}$, 它们之中的前 k 个落在相应于前 k 个圆盘的区域中, 而其余的 $n-k$ 个特征值则落在该区域之外, 因此, 对所有的 $\varepsilon \in [0,1]$, 该结论依然成立. 特别, 当 $\varepsilon = 1$ 时, 结论也成立. $\square$

比如在 例 3.4.3 中, A 的圆盘共有 4 个, 它们共构成 3 个连通部分 (见 例 3.4.3 的图): (1) $|\lambda + 40| \le 6$; (2) $|\lambda + 10| \le 18$; $|\lambda - 10| \le 6$; (3) $|\lambda - 30| \le 8$. 因此可以断言第二个连通部分含有两个特征值, 而在第一、三个由单个圆盘组成的连通部分各含有一个特征值. 进一步, 由于 A 是实数矩阵, 其特征多项式的复数根两两共轭, 而第一、三个圆盘是自共轭的 (即属于该圆盘的复数的共轭仍属于该圆盘), 因此它们所包含的唯一的特征值必然是实特征值. 即有下述

推论 3.4.1 设 n 阶矩阵 A 的主对角线元素均为实数, A 的特征多项式是实系数多项式. 若 A 的每个圆盘均与其余圆盘分离, 则 A 的特征值均为实数.

一般来说, 直接使用圆盘定理得到的特征值估计往往较为粗糙. 比如矩阵

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

的两个盖尔圆盘分别为 $|x-1| \le 4$ 与 $x = 2$, 利用圆盘定理只能知道 A 的两个特征值都落在圆盘 $|x-1| \le 4$ 中, 这个估计当然太差了. 为了得到更为有效的估计, 往往希望每个圆盘只包含 A 的一个特征值, 这就需要将每个圆盘的半径适当缩小. 由于 $\sigma(A) = \sigma(P^{-1}AP)$(相似矩阵的特征值相同), 故可以考虑对矩阵 A 施行相似变换后, 再应用圆盘定理. 为了便于应用, 一般选择对角矩阵作相似变换, 即令 $P = D = \mathrm{diag}\,(d_1, d_2, \cdots, d_n)$, 其中 $d_i > 0$. 于是 $B = D^{-1}AD = (a_{ij}\frac{d_j}{d_i})$. 显然 B 的对角元素与 A 的对角元素完全相同, 因此这样的相似变换 (称为对角相似变换) 只改变盖尔圆盘的半径而不改变它们的中心. (d_j 的选取原则是: 欲使第 j 个圆盘缩小, 可取 $d_j > 1$, 而其余的 d_k 取为 1, 此时 B 的其余圆盘相对放大; 反之, 欲使第 j 个圆盘放大, 可取 $d_j < 1$, 而其余的 d_k 取为 1, 此时 B 的其余圆盘相对缩小.) 利用这个技巧, 再研究前面的矩阵 A, 则得到

$$B = \begin{pmatrix} d_1^{-1} & 0 \\ 0 & d_2^{-1} \end{pmatrix} \begin{pmatrix} 1 & 4 \\ 0 & 2 \end{pmatrix} \begin{pmatrix} d_1 & 0 \\ 0 & d_2 \end{pmatrix} = \begin{pmatrix} 1 & \frac{4d_2}{d_1} \\ 0 & 2 \end{pmatrix},$$

因此 B 的两个盖尔圆盘分别为 $|x-1| \le \frac{4d_2}{d_1}$ 与 $x = 2$, 由于可以将数 $\frac{4d_2}{d_1}$ 取为任意正实数, 因此两个特征值必然为 1 与 2.

例 3.4.6 估计矩阵 A 的特征值, 其中

$$A = \begin{pmatrix} 0.9 & 0.01 & 0.11 \\ 0.01 & 0.8 & 0.12 \\ 0.01 & 0.02 & 0.4 \end{pmatrix}.$$

解 由圆盘定理可知 A 的特征值在下列圆盘之中: $|\lambda - 0.9| \le 0.12$; $|\lambda - 0.8| \le 0.13$; $|\lambda - 0.4| \le 0.03$. 所以第一、二个圆盘构成一个连通部分, 而第三个圆盘单独构成一个连通部分. 这样, 有两个特征值不能分离. 但若作相似变换, $D^{-1}AD = B$, 其中 D 是对角矩阵 $\mathrm{diag}\,(1, 1, \frac{1}{10})$, 则

$$B = \begin{pmatrix} 0.9 & 0.01 & 0.011 \\ 0.01 & 0.8 & 0.012 \\ 0.1 & 0.2 & 0.4 \end{pmatrix}.$$

故 B 的圆盘为

$$|\lambda - 0.9| \le 0.021; \quad |\lambda - 0.8| \le 0.022; \quad |\lambda - 0.4| \le 0.3.$$

从而 B 的三个圆盘都是孤立的, 因而每个圆盘中都有一个特征值. 由于 A 与 B 相似, 从而有相同的特征值, 故 A 的特征值分别在上述三个圆盘中, 且按 推论 3.4.1, 它们都是实数.

例 3.4.7 证明矩阵 A 至少有两个实特征值, 其中

$$A = \begin{pmatrix} 7 & 1 & 2 & -1 \\ 0 & 8 & 1 & 1 \\ -1 & 0 & 5 & 0 \\ 1 & 0 & 0 & 1 \end{pmatrix}.$$

证 A 的四个圆盘为

$$D_1 = |\lambda - 7| \le 4;\quad D_2 = |\lambda - 8| \le 2;\quad D_3 = |\lambda - 5| \le 1;\quad D_4 = |\lambda - 1| \le 1.$$

直接计算可知, 四个圆盘构成两个连通部分, 分别为 $G_1 = D_1 \bigcup D_2 \bigcup D_3$ 与 $G_2 = D_4$. 因此 G_2 包含唯一的特征值, 该特征值只能与自己共轭, 故为实数. 因此, 含在 G_1 中的三个特征值必有一个是实数. 从而 A 至少有两个实特征值.

利用圆盘定理可以得到矩阵谱半径的如下估计.

命题 3.4.2 设 $A = (a_{ij})$ 是 n 阶复数矩阵. 令 $\nu = \max\limits_{1\le k\le n} \sum_{j=1}^n |a_{kj}|$ (此即矩阵行的元素的绝对值之和的最大者), 则 $\rho(A) \le \nu$.

证 设 λ_0 是 A 的一个特征值, 由圆盘定理可知, 一定存在 k 使得 $|\lambda_0 - a_{kk}| \le \sum_{j=1,j\ne k}^n |a_{kj}|$ 成立. 所以 $|\lambda_0| \le |a_{kk}| + \sum_{j=1,j\ne k}^n |a_{kj}| = \sum_{j=1}^n |a_{kj}| \le \nu$. 因此, 任何特征值的模都不超过 ν, 所以 $\rho(A) \le \nu$. □

注意 A 与 A^T 有相同的特征值, 若记矩阵 A 的列的绝对值之和的最大值为 ν', 即 $\nu' = \max\limits_{1\le j\le n} \sum_{k=1}^n |a_{kj}|$, 对 A^T 使用上述命题可得

命题 3.4.3 $\rho(A) \le \min\{\nu, \nu'\}$.

盖尔圆盘定理是最简便的特征值估计方法, 自 1931 年发表以来, 有众多新的拓展方法, 以下介绍其中两种, 证明见习题 50 与 51.

类似于去心绝对行和, 定义矩阵 $A = (a_{ij})$ 的去心绝对列和如下:

$$C_i(A) = \sum_{j\ne i} |a_{ji}|,\ 1 \le\ i \le n.$$

定理 3.4.2 (Ostrowski[34] 圆盘定理) 设 λ 是 n 阶矩阵 A 的一个特征值, $0 \le \alpha \le 1$ 是实数. 则存在 $1 \le i \le n$ 使得

$$|\lambda - a_{ii}| \le R_i(A)^\alpha C_i(A)^{1-\alpha}.$$

可以将上述定理中的圆盘称为第 i 个 Ostrowski 圆盘. 注意对 Ostrowski 圆盘定理, 类似于盖尔圆盘定理的精细圆盘定理也成立, 这依然因为特征值是矩阵各元素的连续函数.

例 3.4.8 设 $A = \begin{pmatrix} 1 & 4 \\ 1 & 6 \end{pmatrix}$. 则 A 的两个盖尔圆盘分别为 $|x-1| \le 4$ 与 $|x-6| \le 1$. 由于这两个圆相切, 故由盖尔圆盘定理不能得出很好的估计. 但取 $\alpha = \frac{1}{2}$, 则可得两个 Ostrowski 圆盘分别为 $|x-1| \le 2$ 与 $|x-6| \le 2$, 这是两个分离的圆盘. 因此可以断言这两个圆盘各包含 A 的一个特征值. 实际上, 将二者结合使用可知, 一个实特征值位于 $|x-1| \le 2$, 而另一个实特征值介于-1 与 7 之间.

[34] Alexander Markowich (1893-1986), 俄裔德国瑞士数学家, 对众多数学分支有重要贡献.

定理 3.4.3 (Brauer[35] 定理) 设 λ 是 n 阶矩阵 A 的一个特征值, 则存在 $1 \le i \ne j \le n$ 使得

$$|\lambda - a_{ii}||\lambda - a_{jj}| \le R_i(A)R_j(A).$$

(上式表达的几何图形称为 Cassini[36] 卵形.)

从纯粹的代数角度看, 定理3.4.3 只不过是两个盖尔圆盘的方程相乘的结果, 但有趣的是, 这个结果不能推广到三个或更多盖尔圆盘的方程相乘的情形, 请读者验证习题 52.

思考题

1. 用盖尔圆盘定理如何估计酉矩阵与正交矩阵的特征值?

第五节 应用: 主元分析法, 商品定价

本章讨论的矩阵的特征值、特征向量与Jordan标准形是矩阵理论的根本，因为矩阵计算的实质是特征值的计算，而矩阵的Jordan标准形则从理论上提供了理解矩阵性质、计算矩阵函数、研究矩阵微积分的一种简便方法 (尽管Jordan标准形不是矩阵计算的实用方法). Schur三角化定理有众多应用, 比如我们将在第四章学习的正规矩阵及其谱分解. 矩阵的Jordan标准形以及盖尔圆盘定理也都具有广泛的应用, 本节我们仅介绍特征值与特征向量的两个应用, 即**主元分析法** 与**商品定价** .

一. 主元分析法

主元分析法也称为主分量分析法, 英文缩写为PCA. 其主要思想如下:

假设有 n 个统计相关的随机变 (向) 量 $\{x_1, \cdots, x_n\}$. 现在希望通过某种方法重新构造 N 个新随机变量 $\{\tilde{x}_1, \cdots, \tilde{x}_N\}$, 使得这些新变量统计独立, 即这些新变量彼此正交 (因而相互之间不再有信息冗余). 这个过程称为**特征提取**. 以矩阵理论的观点看, 特征提取就是将以 $\{x_1, \cdots, x_n\}$ 为基的 n 维复线性空间 (与 $\mathbb{C}^n$ 同构) 转化为以 $\{\tilde{x}_1, \cdots, \tilde{x}_N\}$ 为基的 N 维复线性空间 (与 $\mathbb{C}^N$ 同构). 显然, 为了使新变量继承尽可能多的原始变量的特征, 应首先使用正交变换将原始变量变为一组彼此正交的新变量 (仍有 n 个), 再使用某种适当的线性变换 (称为**降维 (变换)**) 来完成特征提取. 主元分析法是降维的一个特殊情形 (也是应用最广的情形), 其做法是在 n 个新变量中选取其中 N 个**能量** (见下面的定义) 最大者, 并将这 N 个变量 (称为主分量或主成分) 的特征视为 n 个原始变量的特征.

定义 3.5.1 设 $x = (x_1, \cdots, x_n)^T$ 是一个 (复) 随机变量, x 的数学期望是向量 $\mathrm{E}\{x\} = (\mathrm{E}\{x_1\}, \cdots, \mathrm{E}\{x_n\})^T$; x 的能量是随机变量 $|x|^2 = x^*x$ 的数学期望 $\mathrm{E}\{|x|^2\}$, 记为 E_x. x 的**自相关矩阵**是指矩阵 $R_x = \mathrm{E}\{xx^*\} = (\mathrm{E}\{x_i\bar{x}_j\})_{n\times n}$.

注意随机变量 x 的自相关矩阵 R_x 总是Hermite矩阵, 因此其特征值都是实数 (第一章定理 1.5.6), 设为 $\lambda_1 \ge \lambda_2 \ge \cdots \ge \lambda_n$. 它的前 N 个特征值 $\lambda_1, \cdots, \lambda_N$ 称为 x 的主特征值.

主元分析法的步骤如下:

第一步: 降维.

[35]Alfred Brauer(1894-1985), 德裔美籍数学家, 是著名德裔美籍数学家Richard Brauer的哥哥. 北卡罗莱纳大学教堂山分校 (University of North Carolina, Chapel Hill) 的数学图书馆以其名字命名.

[36]Giovanni Domenico Cassini(1625-1712), 意大利法国天文学家, 地心说的代表之一, Cassini卵形即是其描述太阳绕地球旋转的轨迹.

由 n 个原始变量 $\{x_1, \cdots, x_n\}$ 构造 N 个主分量

$$\tilde{x}_j = \sum_{i=1}^{n} \bar{a}_{ij} x_i = \alpha_j^* x, \qquad\qquad j = 1, 2, \cdots, N$$

其中 $x = (x_1, \cdots, x_n)^T$, 而 $\alpha_j = (a_{1j}, a_{2j}, \cdots, a_{nj})^T$ 是待定常数向量.

第二步: 正交化.

为使 N 个主分量成为标准正交组, 即内积 (其中 δ_{ij} 是 Kronecker 符号)

$$(\tilde{x}_i, \tilde{x}_j) = \tilde{x}_j^* \tilde{x}_i = x^* \alpha_j \alpha_i^* x = x^* x \alpha_i^* \alpha_j = |x|^2 \alpha_i^* \alpha_j = \delta_{ij},$$

必须 (因为 $x \neq 0$, 故 $|x|^2 \neq 0$)

$$\alpha_i^* \alpha_j = \frac{\delta_{ij}}{|x|^2}.$$

第三步: 能量最大化.

取 $\alpha_i\,(i = 1, 2, \cdots, N)$ 为 x 的自相关矩阵 R_x 的 N 个主特征值的特征向量 v_i, 则可知诸主分量的能量为

$$\begin{aligned} E_{\tilde{x}_i} &= \mathrm{E}\{|\tilde{x}_i|^2\} = \mathrm{E}\{(\alpha_i^* x)^* \alpha_i^* x\} = \mathrm{E}\{\alpha_i^* x (\alpha_i^* x)^*\} = v_i^* \mathrm{E}\{x x^* v_i\} = v_i^* R_x v_i \\ &= v_i^* (v_1, v_2, \cdots, v_N) \mathrm{diag}\,(\lambda_1, \lambda_2, \cdots, \lambda_N) \begin{pmatrix} v_1^* \\ v_2^* \\ \vdots \\ v_N^* \end{pmatrix} v_i = \lambda_i\,. \end{aligned}$$

由于诸 $\lambda_i\,(i = 1, 2, \cdots, N)$ 是主特征值, 因此

$$E_{\tilde{x}_1} \geq E_{\tilde{x}_2} \geq \cdots \geq E_{\tilde{x}_N}$$

这就保证了新变量 $\tilde{x}_i\,(i = 1, 2, \cdots, N)$ 确为能量最大者.

注意自相关矩阵 R_x 的对角元素恰好是诸原始变量 x_i 的能量 E_{x_i}, 因此若 R_x 仅有 N 个大的特征值 $\lambda_i\,(i = 1, 2, \cdots, N)$, 则

$$\mathrm{tr}(R_x) = E_{x_1} + E_{x_2} + \cdots + E_{x_n} \approx \lambda_1 + \lambda_2 + \cdots + \lambda_N$$

因此

$$E_{\tilde{x}_1} + E_{\tilde{x}_2} + \cdots + E_{\tilde{x}_N} = \lambda_1 + \lambda_2 + \cdots + \lambda_N \approx E_{x_1} + E_{x_2} + \cdots + E_{x_n}$$

这就是说, 将原始 n 个随机变量换成 N 个彼此正交的主分量, 可以近似保持原始随机变量的能量之和.

二. 商品定价

我们先介绍非负矩阵的一些概念.

将单位矩阵作任意次行 (或列) 交换所得的矩阵称为**置换矩阵**. 如果矩阵 A 的所有元素均是正数, 则称 A 是**正矩阵**, 记作 $A > 0$. 设 A 是一个 n 阶非负矩阵 (记作 $A \geq 0$), 如果存在置换矩阵 P 使得 $P^T A P = \begin{pmatrix} B & C \\ 0 & D \end{pmatrix}$, 其中 B, C, D 均为非负矩阵, 则称 A 是可约矩阵, 否则

称 A 为**不可约 (非负) 矩阵**. 比如矩阵 $\begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}$ 是可约的, 而 $\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$ 是不可约的. 对非负不可约矩阵有下述刻画 (证明见习题 62):

定理 3.5.1 设 A 是 n 阶非负矩阵, 则 A 是不可约矩阵 $\iff (I+A)^{n-1} > 0$.

由此定理即可知矩阵 $\begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}$ 是不可约的.

定理 3.5.2 (**Perron**[37]**-Froubenius 定理**[38])) 设 A 是 n 阶不可约非负矩阵, 则
(1) A 的谱半径 $\rho(A)$ 是 A 的一个单特征值;
(2) A 有一个属于特征值 $\rho(A)$ 的正特征向量;
(3) A 的其余特征值 λ 均无正特征向量.

Perron-Froubenius 定理是非负矩阵理论中最精彩的部分, 其证明较难 (可以说是矩阵理论的第一难度), 有兴趣的读者可尝试证明之或阅读参考文献 [17] 或 [14]. Perron-Froubenius 定理中的谱半径 $\rho(A)$ 称为 A 的**Perron 根**, 而属于 Perron 根的正特征向量称为矩阵 A 的**Perron 向量**. 由 Perron-Froubenius 定理, 不可约非负矩阵的 Perron 向量是该矩阵唯一的正特征向量.

理论经济学中的商品定价问题与 Perron 根与 Perron 向量密切相关.

考虑 n 个工厂的封闭经济系统 $\mathcal{E}$(即自产自销, 自给自足的理想经济体系), 每个工厂生产一种商品. 设描述该经济系统的状态矩阵为 $A=(a_{ij})_{n\times n}$, 其中 a_{ij} 表示第 i 个工厂购买第 j 个工厂商品的百分比. 显然有

$$a_{ij} \geq 0, \qquad \sum_{i=1}^{n} a_{ij} = 1,\ j=1,2,\cdots,n. \tag{3.5.1}$$

因此 A 是一个非负矩阵且其每个列和均为 1 (这样的矩阵称为**随机矩阵**). 故 1 是 A 的一个特征值, 且由盖尔圆盘定理可知 A 的其余特征值的模均小于 1. 进一步, 可以假设 A 是不可约的. 现在的问题是: 有无可能制定一个“公平的”价格体系使得每个工厂均能收支平衡? 为解决此问题, 设 P_i 表示第 i 个工厂的总收入, 并记 $P=(P_1,\cdots,P_n)^T$(向量 P 称为**定价向量** 或**定价结构**). 如果定价结构 P 使得每个工厂保持收支平衡, 则称 P 是一个**均衡定价结构** .

定理 3.5.3 (Leontief 定理[39]) 设 A 是封闭经济系统 $\mathcal{E}$ 的状态矩阵, P 是一个 n 元正向量. 则 P 是一个均衡定价结构 $\iff$ P 是 A 的一个 Perron 向量.

证 由于 a_{ij} 是第 i 个工厂购买第 j 个工厂商品的百分比, 故第 i 个工厂的总支出为 $\sum\limits_{j=1}^{n} a_{ij}P_j$. 因此 P 是一个均衡定价结构 $\iff$

$$P_i = \sum_{j=1}^{n} a_{ij}P_j, \quad i=1,2,\cdots,n. \tag{3.5.2}$$

[37] Oskar Perron(1880-1975), 德国数学家, 数学上有著名的**Perron 悖论**, 即若假定有最大的自然数, 则该自然数必然等于 1!(这说明假定存在性是很危险的.)

[38] Perron 于 1907 年对正矩阵, Frobenius 于 1912 年对非负矩阵证明了该定理.

[39] 该定理出自 Leontief 的著名论文"Input-Output Economics", Scientific Amrican,October 1951, pp. 15-21.

即 $AP = P$. □

当然, 一个封闭的经济系统是不可能持续发展的, 也就是说, 必须要出售部分商品给外部方能使系统“盈利”. 设第 i 个工厂的盈利为 $\gamma_i \geq 0$, 即有

$$\gamma_i = P_i - \sum_{j=1}^{n} a_{ij} P_j, \quad i = 1, 2, \cdots, n. \tag{3.5.3}$$

记 $\gamma = (\gamma_1, \cdots, \gamma_n)^T$. 则 γ 是一个非负向量, 且 $\gamma \neq 0$. 将公式 (3.5.3) 写成矩阵形式有

$$(I - A)P = \gamma. \tag{3.5.4}$$

使上述方程对任何非负向量 γ 都有解的一个充分必要条件是 $\rho(A) < 1$(证明见习题 63). 此时矩阵 A 的列和至少有一个小于 1, 即至少有一个工厂向外部出售了商品.

现考虑一个简单情形, 即每个工厂的毛利率都相同 (这看起来比较“公平”), 设为 $r > 0$. 于是 $\gamma_i = rP_i$, 因此方程 (3.5.4) 变为

$$(I - A)P = rP. \tag{3.5.5}$$

即

$$AP = \lambda P, \quad \text{其中 } \lambda = 1 - r,$$

于是定价结构 P 仍是 A 的一个正特征向量, 因此由 Perron-Froubenius 定理可知 P 必然是 A 的一个 Perron 向量! 所以, 无论是一个均衡定价的封闭经济系统还是一个具有正盈利的开放经济系统, 其定价结构必然是其系统矩阵的一个正特征向量即 Perron 向量, 只不过后者的状态矩阵的谱半径即 Perron 根小于 1.

思考题
1. 置换矩阵的行列结构有何特点?
2. 置换矩阵对应的线性变换有何特点?
3. 设 P 是置换矩阵, A 是任意与 P 同阶的矩阵, 问 P^TAP 与 A 的行列结构有何关系?

习 题 三

1. 详细证明分块 Schur 三角化定理.

2. 设 A 为第一章例 1.2.2 中的矩阵.
(1) 利用满秩分解和 Sylvester 降幂公式求 A 的特征多项式与 A^6;
(2) 求与 A 相似的分块对角矩阵, 使得每块恰有唯一的特征值.

3. 设 $\alpha = (a_1, \cdots, a_n)^T$, $\beta = (b_1, \cdots, b_n)^T$, x 为任意常数, $A = xI_n + \alpha\beta^T$.
(1) 直接计算行列式 $|A|$;
(2) 利用 Sylvester 降幂公式计算行列式 $|A|$;
(3) 利用特征值计算行列式 $|A|$.

4. 设 $\alpha \neq \beta$, 试求矩阵

$$\begin{pmatrix} \alpha & a & b & c \\ 0 & \alpha & d & e \\ 0 & 0 & \beta & x \\ 0 & 0 & 0 & \beta \end{pmatrix}$$

的相似分块对角矩阵.

5. (1) 举例说明 Schur 三角化定理在实数域上不成立;

(2) 证明实数域上的 Schur 三角化定理: 设 $A \in M_n(\mathbb{R})$, 则存在正交矩阵 Q 使得

$$Q^TAQ = \begin{pmatrix} A_1 & & * & \\ & A_2 & & \\ & & \ddots & \\ & 0 & & A_k \end{pmatrix}$$

其中每个 A_i 或者是 1 阶实矩阵或者是形如 $\begin{pmatrix} a_i & b_i \\ -b_i & a_i \end{pmatrix}$ 的 2 阶实矩阵 $(b_i \neq 0)$.

(提示: 首先, 如果 λ 是 A 的非实数特征值, $Ax = \lambda x$, 则 $A\bar{x} = \bar{\lambda}\bar{x}$, 由此可知 x 与 $\bar{x}$ 线性无关, 进而 $\mathrm{Re}x$ 与 $\mathrm{Im}x$ 线性无关, 将其正交化后构造正交矩阵, 再利用归纳法.)

6. 设 a 是复常数, $V = \{e^{ax}f(x) \mid f(x) \in \mathbb{C}_n[x]\}$ 是 n 维复线性空间.

(1) 证明求导运算 $\partial : \alpha \mapsto \frac{\mathrm{d}\,\alpha}{\mathrm{d}\,x}$ 是 V 上的线性变换;

(2) 求 ∂ 的 Jordan 标准形.

7. 详细证明 定理 3.1.5.

8. 设 σ 是 $\mathbb{R}^3$ 的线性变换, 设 $x = (x_1, x_2, x_3)^T$, $\sigma(x) = (-2x_2 - 2x_3, -2x_1 + 3x_2 - x_3, -2x_1 - x_2 + 3x_3)^T$. 试求 $\mathbb{R}^3$ 的一个基, 使得 σ 在该基下的矩阵尽可能简单.

9. 已知 $B = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$, 线性空间 $V = \{X = (x_{ij}) \in M_2(\mathbb{R}) \mid \operatorname{tr} X = 0\}$ 的线性变换 σ 为 $\sigma(X) = B^TX - X^TB$, $X \in V$. 试求 V 的一个基, 使得 σ 在该基下的矩阵尽可能简单.

10. 设 A 的特征值为 0, 1, 对应的特征向量为 $(1,2)^T, (2,-1)^T$. 判断 A 是否为对称矩阵并求 A.

11. 求下列矩阵的最小多项式并指出其中可以对角化的矩阵:

(1) $\begin{pmatrix} 3 & 2 \\ 4 & 5 \end{pmatrix}$; (2) $\begin{pmatrix} 2 & 3 & 2 \\ 0 & 5 & 4 \\ 0 & -2 & -1 \end{pmatrix}$; (3) $\begin{pmatrix} 4 & 6 & 0 \\ -3 & -5 & 0 \\ 3 & 6 & 1 \end{pmatrix}$; (4) $\begin{pmatrix} 4 & 6 & 0 \\ -3 & -5 & 0 \\ -3 & 6 & 1 \end{pmatrix}$.

12. 试构造两个同阶矩阵, 使得它们

(1) 具有相同的特征多项式与不同的最小多项式;

(2) 具有相同的最小多项式与不同的特征多项式;

(3) 证明矩阵的最小多项式存在且唯一.

13. 设 n 阶矩阵 A 的特征值均为实数. 证明:

(1) A 的特征多项式的 $n-k$ 次项的系数等于 A 的所有 k 阶主子式之和;

(2) 若 A 的所有一阶主子式之和与所有二阶主子式之和都等于零, 则 A 是幂零矩阵.

14. 设 n 阶矩阵 A 的主对角元全是 1, 且其特征值均为非负数, 证明 $|A| \leq 1$.

15. 设 $AB = BA$, 证明 A 与 B 有公共的特征向量. 该结论的逆命题成立吗?

16. (1) 验证 命题 3.2.1;

(2) 证明幂零矩阵相似的准则即 推论 3.2.1.

17. 计算幂零矩阵的 Jordan 标准形定理中的 (3.2.3).

18. 本题是幂零矩阵的 Jordan 标准形定理的证明中当数 $a = \alpha_1^Te_1 \neq 0$ 时的实际计算. 设

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

计算 A 的 Jordan 标准形.

19. 利用矩阵的 Jordan 标准形定理证明 **Fitting**[40] **引理**(对照第二章公式 (2.5.5)):
设 V 为 n 维线性空间, $\sigma \in \mathrm{End}V$, 则 $V = \mathrm{Im}(\sigma^n) \oplus \mathrm{Ker}(\sigma^n)$.

20. 证明 定理 3.3.1 的 (2) 与 (3).

下面的 21-24 题展示了如何利用广义特征子空间来得到矩阵的 Jordan 标准形, 其中设 $|\lambda I - A| = \prod\limits_{i=1}^{s}(\lambda - \lambda_i)^{n_i}$ 是 n 阶矩阵 A 的特征多项式, $\sigma(A) = \{\lambda_1, \lambda_2, \cdots, \lambda_s\}$, g_i 为 λ_i 的几何重数.

21. 证明广义特征子空间 $E_{\lambda_i} = \{x \in \mathbb{C}^n \mid (A - \lambda I)^{n_i} x = 0\}$.

22. 证明 $\dim_{\mathbb{C}} E_{\lambda_i} = n_i$, 从而 $\mathbb{C}^n = \sum\limits_{\lambda \in \sigma(A)} \oplus E_\lambda$.(此即“谱定理”.)

23. 证明存在 $\alpha_j \in E_{\lambda_i}, 1 \le j \le g_i$, 使得 $\cup_{1\le j\le g_i}\{\alpha_j, (A - \lambda_i I)\alpha_j, \cdots, (A - \lambda_i I)^{m_j - 1}\alpha_j\}$ 构成 E_{λ_i} 的一组基 (称为由诸向量 α_j 生成的循环基), 从而 E_{λ_i} 是 A 的不变子空间.

24. 由每个广义特征子空间的循环基构成的 $\mathbb{C}^n$ 的基称为 Jordan 基. 证明 A 在 $\mathbb{C}^n$ 的 Jordan 基下的矩阵是其 Jordan 标准形 (即将 A 看成是线性变换 $x \mapsto Ax$).

25. 详细证明 定理 3.3.2, 并研究矩阵 A 的特征向量与变换矩阵 P 的特征向量之间的关系.

26. 求 例 3.3.5 的变换矩阵.

27. 求下列矩阵的 Jordan 标准形:

$$(1)\begin{pmatrix} 0 & 1 & 0 \\ -4 & 4 & 0 \\ -2 & 1 & 2 \end{pmatrix};\quad (2)\begin{pmatrix} 2 & 6 & -15 \\ 1 & 1 & -5 \\ 1 & 2 & -6 \end{pmatrix};\quad (3)\begin{pmatrix} 9 & -6 & -2 \\ 18 & -12 & -3 \\ 18 & -9 & -6 \end{pmatrix}$$
$$(4)\begin{pmatrix} 4 & 6 & -15 \\ 1 & 3 & -5 \\ 1 & 2 & -4 \end{pmatrix};\quad (5)\begin{pmatrix} 0 & -4 & 0 \\ 1 & -4 & 0 \\ 1 & -2 & -2 \end{pmatrix};\quad (6)\begin{pmatrix} 4 & -5 & 2 \\ 5 & -7 & 3 \\ 6 & -9 & 4 \end{pmatrix}.$$

28. 求下列矩阵的 Jordan 标准形, 并求变换矩阵 P 使 $P^{-1}AP = J$:

$$(1)\begin{pmatrix} -4 & 2 & 10 \\ -4 & 3 & 7 \\ -3 & 1 & 7 \end{pmatrix};\ (2)\begin{pmatrix} 3 & 2 & 1 \\ 0 & 4 & 0 \\ -1 & 2 & 5 \end{pmatrix};\ (3)\begin{pmatrix} 2 & 2 & 2 & 1 \\ -1 & -1 & -3 & -2 \\ 1 & 2 & 5 & 3 \\ -1 & -2 & -4 & -2 \end{pmatrix}.$$

29. 试判断下面 4 个矩阵, 哪些是相似的:

$$A = \begin{pmatrix} -3 & 3 & -2 \\ -7 & 6 & -3 \\ 1 & -1 & 2 \end{pmatrix},\ B = \begin{pmatrix} 0 & 1 & -1 \\ -4 & 4 & -2 \\ -2 & 1 & 1 \end{pmatrix},\ C = \begin{pmatrix} 0 & -1 & -1 \\ -3 & -1 & -2 \\ 7 & 5 & 6 \end{pmatrix},\ D = \begin{pmatrix} 0 & 1 & 2 \\ 0 & 1 & 1 \\ 0 & 0 & 2 \end{pmatrix}.$$

30. (1) 证明不等于零的幂零矩阵一定不相似于对角矩阵;
(2) 设 A 具有唯一特征值但 A 不是对角矩阵. 证明 A 一定不相似于对角矩阵.

31. 证明任何复矩阵 A 可唯一地分解为 $A = D + N$, 其中 D 为可对角化矩阵, N 是幂零矩阵, 且 $DN = ND$.(此称为矩阵的 Jordan-Chevalley[41] 分解.) 以此解释上题的结论.

32. 设 $A = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & -1 \\ 0 & 1 & 1 \end{pmatrix}$.
(1) 求 A 的特征值及 A^{100};
(2) A 的 Jordan-Chevalley 分解是什么?

---

[40] Hans Fitting(1906-1938), 德国数学家.

[41] Claude Chevalley(1909-1984), 著名法国数学家, 生于南非, 拥有美法两国国籍, 对当代数学的众多分支有重要贡献.

33. 设 $p(\lambda)=(-1)^n[\lambda^n-a_{n-1}\lambda^{n-1}-a_{n-2}\lambda^{n-2}-\cdots-a_1\lambda-a_0]$. 称矩阵

$$C=\begin{pmatrix} a_{n-1} & a_{n-2} & \cdots & a_1 & a_0 \\ 1 & 0 & \cdots & 0 & 0 \\ 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & 0 \end{pmatrix}$$

为多项式 $p(\lambda)$ 的友矩阵. 设 n 阶矩阵 A 的特征多项式为 $(-1)^n p(\lambda)$.

(1) 计算 C 的特征多项式;

(2) 当 $n=2$ 时, 证明：A 与 C 相似当且仅当 A 的最小多项式等于其特征多项式;

(3) 试将 (2) 中的结论推广到一般情形. (提示: 查阅 Frobenius 标准形与有理标准形.)

34. 设 V 是由函数 e^x, xe^x, x^2e^x, e^{2x} 的线性组合生成的线性空间. 定义 V 的一个线性算子如下: $T(f)=f'$. 求 T 的 Jordan 标准形及 Jordan 基.

35. 如果矩阵 A 的特征多项式和最小多项式相同, 问 A 的 Jordan 标准形有何特点?

36. (酉矩阵与离散 Fourier 变换) 设 σ 是 $\mathbb{C}^n$ 的循环位移变换, 即 $\sigma((x_1,x_2,\cdots,x_n)^T)=(x_2,x_3,\cdots,x_n,x_1)^T$. 证明:

(1)σ 的特征值恰好为方程 $\lambda^n=1$ 的所有根 $\lambda_j=e^{\frac{2\pi \mathrm{i}}{n}j}, 1\le j\le n$;

(2)σ 的属于特征值 λ_j 的特征向量为 $\alpha_j=(\lambda_j,\lambda_j^2,\cdots,\lambda_j^n)^T$, 且 $\|\alpha_j\|=\sqrt{n}$;

(3)$\alpha_1,\cdots,\alpha_n$ 是 $\mathbb{C}^n$ 的一组正交基;

(4) 任何向量 $x=(x_1,x_2,\cdots,x_n)^T$ 均是 σ 的特征向量 α_j 的线性组合 $x=\sum\limits_{j=1}^n a_j\alpha_j$, 即 $x_k=\sum\limits_{j=1}^n a_je^{\frac{2\pi\mathrm{i}}{n}j}$;

(5) 上面的系数 $a_j=(x,\alpha_j)/n=\frac{1}{n}\sum\limits_{k=1}^n x_ke^{-\frac{2\pi\mathrm{i}}{n}jk}$;

(6) 研究 σ 与第一章习题 7 中的 Fourier 矩阵的关系, 并由此再求该矩阵的逆.

37. 求矩阵 $A=\begin{pmatrix} 9 & 1 & 1 \\ 1 & i & 1 \\ 1 & 1 & 3 \end{pmatrix}$ 的盖尔圆盘并隔离之.

38. 求矩阵 $A=\begin{pmatrix} 20 & 3 & 1 \\ 2 & 10 & 2 \\ 8 & 1 & 0 \end{pmatrix}$ 的盖尔圆盘并讨论 A 的特征值的范围与性质.

39. 设矩阵

$$A=\begin{pmatrix} 7 & -16 & 8 \\ -16 & 7 & -8 \\ 8 & -8 & -5 \end{pmatrix}.$$

(1) 求 A 的盖尔圆盘并利用对角相似变换改进之;

(2) 同构特征多项式计算 A 的特征值并与 (1) 比较.

40. 证明 Hilbert[42] 矩阵

$$A=\begin{pmatrix} 2 & \frac{1}{2} & \frac{1}{2^2} & \cdots & \frac{1}{2^{n-1}} \\ \frac{2}{3} & 4 & \frac{2}{3^2} & \cdots & \frac{2}{3^{n-1}} \\ \frac{3}{4} & \frac{3}{4^2} & 6 & \cdots & \frac{3}{4^{n-1}} \\ \vdots & \vdots & \vdots & & \vdots \\ \frac{n}{n+1} & \frac{n}{(n+1)^2} & \frac{n}{(n+1)^3} & \cdots & 2n \end{pmatrix}$$

可以对角化, 且 A 的特征值都是实数.

[42] David Hilbert(1862-1943), 著名德国数学家, 对数学与物理的众多分支有杰出贡献, 他于 1900 年提出的数学的 23 个问题几乎确定了此后一个多世纪世界数学的整个发展方向. Hilbert 的名言: We must know, we shall know.

41. 分别利用盖尔圆盘定理和Ostrowski圆盘定理估计下面矩阵的谱半径:

$$A=\begin{pmatrix} 1 & 0.1 & 0.2 & 0.3 \\ 0.5 & 3 & 0.1 & 0.2 \\ 1 & 0.3 & -1 & 0.5 \\ 1.2 & -0.6 & -0.2 & -3.6 \end{pmatrix}.$$

42. 证明 $\sigma(A)\subseteq G(A)\cap G(A^T)$.

43. 设矩阵 A 酉相似于矩阵

$$\begin{pmatrix} 3.05 & -0.06 & 0.02 \\ -0.06 & -6.91 & 0.07 \\ 0.02 & 0.07 & 8.44 \end{pmatrix}.$$

试估计 A 的特征值.

44. 证明 $\sigma(A)=\cap_P G(P^{-1}AP)$, 其中 P 取遍所有可逆矩阵. 如果将 P 限定为可逆对角矩阵如何?

45. 设 $A=(a_{ij})$ 有 s 行严格对角占优, 证明 $r(A)\geq s$.

46. 设 n 阶矩阵 $A=(a_{ij})$ 至少有 $n-1$ 行严格对角占优, 且剩余一行的对角元素非零, 证明 A 可逆.

47. 证明 **Hadamard**[43] **不等式**：对任意 n 阶矩阵 $A=(a_{ij})$ 有 $|A|\leq\prod\limits_{i=1}^{n}(\sum\limits_{i=1}^{n}|a_{ij}|^2)^{1/2}$. 并由此证明:

(1) 若 A 是正定矩阵, 则 $|A|\leq\prod\limits_{i=1}^{n}a_{ii}$;

(2) 设 C 是非负实数, 若 $|a_{ij}|\leq C, 1\leq i,\, j\leq n$, 则 $|A|\leq C^n n^{n/2}$;

(3) 设 $a_{ij}=\pm1, 1\leq i,\, j\leq n$, 则由 (2) 可知 $|A|\leq n^{n/2}$. 如果等号成立, 则称 A 是一个 **Hadamard 矩阵**. 证明 A 是 **Hadamard 矩阵** $\iff$ $A^TA=nI_n$ $\iff$ A 的列两两正交.

48. 设 $A=(a_{ij})$ 是严格对角占优矩阵, $D=\mathrm{diag}\,(a_{11},\cdots,a_{nn})$. 证明 D 可逆且 $\rho(I-D^{-1}A)<1$.

49. 试利用 Ostrowski 圆盘定理和 Brauer 定理各给出一个矩阵可逆的充分条件.

50. 试证明 Ostrowski 圆盘定理.

51. 试证明 Brauer 定理.

52. 研究下面的矩阵, 说明 Brauer 定理不能推广到三个盖尔圆盘的方程相乘的情形.

$$A=\begin{pmatrix} 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

53. 设数列 $a_0=0, a_1=1, a_2,\cdots,a_n,\cdots$ 满足条件 $a_{n+1}=xa_{n-1}+a_n, n\geq1$, 试求 a_n 的通项公式, 其中 x 为实参数. (当 $x=1$ 时, 此数列即为 **Fibonacci**[44] **数列** .)

54. 设 A 是 n 阶矩阵, 称满足条件 $y^TA=\lambda y^T$ 的向量 y 为 A 的属于特征值 λ 的左特征向量. 证明: A 的相应于特征值 λ 的左特征向量与相应于特征值 μ 的特征向量正交 $(\lambda\neq\mu)$.

55. (无限维线性空间的线性变换的特征值与特征向量) 无限次可导的实函数全体构成一个无限维实线性空间, 记为 C^∞. 定义 C^∞ 上的线性变换 $\partial=\frac{\mathrm{d}}{\mathrm{d}x}$:

$$\partial:\quad f(x)\mapsto f'(x).$$

试求 ∂ 的谱 $\sigma(\partial)$ 与特征向量. 比较你的结论与有限维线性空间的相应结论.

---

[43] Jacques Salomon(1865-1963), 著名法国数学家, 对数学的诸多分支有重要贡献, 组合学中有著名的 **Hadamard 猜想**: 对每个正整数 k, 均存在 $4k$ 阶的 Hadamard 矩阵.

[44] Leonardo Pisano 或 Leonardo Bonacci(1170-1250), 意大利数学家, 被称为中世纪最具天赋的西方数学家.

下面的56-60题是证明Jordan标准形存在与唯一性的 λ- 矩阵 (即以 λ 的多项式为元素的矩阵) 方法, 为简单起见, 所有矩阵均假定是 n 阶的. λ- 矩阵的初等变换与通常的线性变换类似 (倍加变换可以使用多项式).

56. 证明任何秩为 r 的 λ- 矩阵 $A(\lambda)$ 均 (在初等变换下) 等价于下面的对角矩阵 (称为 $A(\lambda)$ 的 **Smith 标准形** [45])

$$\mathrm{diag}\,(d_1(\lambda), d_2(\lambda), \cdots, d_r(\lambda), 0, \cdots, 0),$$

其中 $d_i(\lambda)$ 均为首一多项式且 $d_i(\lambda)|d_{i+1}(\lambda), 1 \leq i \leq r-1$. 这些 $d_i(\lambda)$ 称为 $A(\lambda)$ 的**不变因子**.

57. 将 $A(\lambda)$ 的每个正次数的不变因子分解为不同的首一一次多项式的幂的乘积, 这些一次多项式的幂合称为矩阵 $A(\lambda)$ 的**初等因子**. 证明每个Jordan块仅有唯一的不变因子, 从而这个不变因子就是它的唯一初等因子.

58. 证明 $\lambda I - A$ 与 $\lambda I - B$ 等价 $\iff$ 它们有相同的不变因子.

59. 证明矩阵 A 与 B 相似 $\iff$ $\lambda I - A$ 与 $\lambda I - B$ 等价.

60. 证明任何矩阵 A 一定相似于一个Jordan标准形.

61. 判断下面的矩阵是否为不可约矩阵?

$$(1)\begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}, \quad (2)\begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 1 \end{pmatrix}, \quad (3)\begin{pmatrix} 0 & 1 & 0 \\ 1 & 1 & 1 \\ 0 & 1 & 0 \end{pmatrix}.$$

62. 证明 定理3.5.1.

63. 设 A 是非负矩阵, 证明方程 $(I-A)P = \gamma$ 对任何非负向量 γ 总有正向量解 $P \iff \rho(A) < 1$.(提示: 如果 $(I-A)^{-1} \geq 0$, 则 $\rho(A) < 1$.)

[45] Henry John Stephen Smith (1826-1883), 爱尔兰数学家.