多项式矩阵

黄利兵

数学科学学院

2023 年 4 月 6 日

主要内容

导言

- 本章主要解决两个问题:
  (1) 如何快速计算一个矩阵的 Jordan 标准形?
  (2) 如何判断两个矩阵是否相似?
- 解决这两个问题的一种方案是使用多项式矩阵; 它既可看作以多项式作为元素的矩阵, 也可看作以数字矩阵为系数的多项式.
- 初等变换是通往多项式矩阵的诸多不变量的有力武器.

多项式矩阵的定义 (一)

前面我们通常都是在某个数域中讨论矩阵, 但在有些情形, 参与运算的对象不一定构成数域.

例

张丘建《算经》: 鸡翁一, 值钱五; 鸡母一, 值钱三; 鸡雏三, 值钱一. 百钱买百鸡, 则翁、母、雏各几何? 它等价于求线性方程组 $\begin{cases} 15x+9y+z=300 \\ x+y+z=100 \end{cases}$ 的 (非负) 整数解. 容易发现, 最后的结果为 $(x,y,z)=(4t,25-7t,75+3t)$, $t\in\mathbb{Z}$.

例

在 Hamilton-Cayley 定理的证明过程中, 我们使用了这样一种矩阵, 其中的元素都是 λ 的多项式.

例

在计算矩阵 A 的特征值时, 我们通常先对矩阵 $\lambda E_n - A$ 进行初等行变换, 变换的过程中得到的矩阵里每个元素都是 λ 的多项式, 而变换的结果可以用于计算特征向量.

多项式矩阵的定义 (二)

以上这些场景的共同特点都是: 需要考虑的矩阵中的元素都是整数或多项式. 整数和多项式都可以作加法、减法、乘法以及带余除法, 并都有相应的整除理论. 下面将只讨论多项式情形, 整数情形是类似的.

定义

若 $a_{ij}(\lambda) \in P[\lambda]$, $1 \le i \le m$, $1 \le j \le n$, 则称矩阵 $A(\lambda) = (a_{ij}(\lambda))$ 为多项式矩阵或 λ-矩阵. 所有这样的矩阵构成的集合记作 $P[\lambda]^{m\times n}$.

多项式矩阵 $A(\lambda)$ 也可写为 $A_k\lambda^k + A_{k-1}\lambda^{k-1} + \cdots + A_1\lambda + A_0$, 其中 A_k, A_{k-1}, $\cdots$, A_1, $A_0 \in P^{m\times n}$. 若 $A_k \ne 0$, 则称 $A(\lambda)$ 的次数为 k, 记作 $\deg A(\lambda) = k$.

例

多项式矩阵 $\begin{bmatrix} \lambda^2+1 & \lambda-1 \\ -\lambda+3 & \lambda^2+\lambda \\ 2\lambda^2-1 & \lambda-\lambda^2 \end{bmatrix}$ 也可写为

$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 2 & -1 \end{bmatrix}\lambda^2 + \begin{bmatrix} 0 & 1 \\ -1 & 1 \\ 0 & 1 \end{bmatrix}\lambda + \begin{bmatrix} 1 & -1 \\ 3 & 0 \\ -1 & 0 \end{bmatrix}.$$ 它的次数为 2.

多项式矩阵的秩

多项式矩阵的加法、数乘以及乘法等运算都与普通矩阵的运算定义一致，只是将数的运算换成了多项式的运算．多项式矩阵也可以像普通矩阵那样定义行列式，性质也类似．需要特别讨论的概念有两个，即多项式矩阵的秩和可逆矩阵．

定义

设 $A(\lambda) \in P[\lambda]^{m\times n}$．如果 $A(\lambda)$ 有某个 r 级子式不是零多项式，而所有 $r+1$ 级子式全为零，则称 $A(\lambda)$ 的秩为 r．

例

矩阵 $\begin{bmatrix} \lambda^2 & 1 & \lambda-1 \\ \lambda^3 & \lambda & \lambda^2 \\ 0 & 0 & \lambda \end{bmatrix}$ 右上角的子式不为零，而整个矩阵的行列式为零，所以它的秩为 2．

可逆的多项式矩阵

定义

设 $A(\lambda) \in P[\lambda]^{n\times n}$. 如果存在 $B(\lambda)$, 使得 $A(\lambda)B(\lambda) = E_n$, 则称 $A(\lambda)$ 可逆.

命题

n 阶多项式矩阵 $A(\lambda)$ 可逆当且仅当其行列式为非零常数.

证明.

如果 $A(\lambda)$ 可逆, 则有多项式矩阵 $B(\lambda)$ 使得 $A(\lambda)B(\lambda) = E_n$. 取行列式可得 $\det(A(\lambda))\det(B(\lambda)) = 1$, 因此 $\det(A(\lambda))$ 为非零常数.
如果 $\det(A(\lambda)) = d$ 为非零常数, 则由于 $A(\lambda)$ 的伴随矩阵 $C(\lambda)$ 仍为多项式矩阵, 可知 $C(\lambda)/d$ 为 $A(\lambda)$ 的逆矩阵. □

例

多项式矩阵 $A(\lambda) = \begin{bmatrix} \lambda-1 & \lambda+1 & 2\lambda \\ \lambda^2-1 & \lambda^2 & \lambda+1 \\ 1 & 1 & 0 \end{bmatrix}$ 的行列式为 2, 所以它是可逆的.

多项式矩阵的初等变换

对多项式矩阵, 可以作如下三种初等行 (列) 变换:

- 交换两行 (列);
- 将某行 (列) 乘以非零常数 c;
- 将某行 (列) 的 $\varphi(\lambda)$ 倍加到另一行 (列) 上 (其中 $\varphi(\lambda) \in P[\lambda]$).

初等行变换和初等列变换统称为初等变换.

定义

若 $A(\lambda) \in P[\lambda]^{m\times n}$ 能经若干次初等变换变为 $B(\lambda)$, 则称 $A(\lambda)$ 与 $B(\lambda)$ 相抵.

容易看出, 相抵是多项式矩阵之间的等价关系.

相抵标准形

定理 (Smith 标准形)

任意一个多项式矩阵 $A(\lambda) \in P[\lambda]^{m\times n}$ 总可经若干次初等变换变为对角矩阵 $\mathrm{diag}(d_1(\lambda), d_2(\lambda), \cdots, d_r(\lambda), 0, \cdots, 0)$, 其中 $d_1(\lambda)$, $d_2(\lambda)$, $\cdots$, $d_r(\lambda)$ 为首一多项式, 且 $d_1(\lambda)|d_2(\lambda)|\cdots|d_r(\lambda)$.

后面我们将证明这里的对角矩阵是唯一的, 称为 $A(\lambda)$ 的相抵标准形, 或 *Smith* 标准形, 简称标准形; 其中的多项式 d_1, d_2, $\cdots$, d_r 称为 $A(\lambda)$ 的不变因子.

为证明这个定理, 我们需要描述出初等变换的过程. 这个过程将分为两部分, 第一部分实现对角化, 第二部分实现顺次整除的性质. 这两部分各归结为一个引理.

实现对角化的方法

引理

若多项式矩阵 $A(\lambda) \neq 0$, 则可经初等变换使得左上角不为零, 而第一行和第一列的其他位置全为零.

证明.

首先在 $A(\lambda)$ 中找到次数最低的非零多项式, 记作 $g(\lambda)$. 适当交换行和列, 可将 $g(\lambda)$ 置于第一行第一列, 即左上角. 对于第一行第 k 列的多项式 $f(\lambda)$, 可将第一列的适当倍数加到第 k 列, 使得 $f(\lambda)$ 变为 $g(\lambda)$ 除 $f(\lambda)$ 的余式 $r(\lambda)$, $k = 2, \cdots, n$. 类似地, 可将第一列的其他多项式也都替换为被 $g(\lambda)$ 除的余式.

如果所有这些余式都为零, 则引理得证. 否则, 非零余式的次数将小于 $g(\lambda)$ 的次数, 我们可重复上述过程. 由于此过程中左上角多项式的次数严格减小, 所以此过程一定在有限步内结束, 这时第一行和第一列的其他位置都是零. $\square$

反复利用这个引理, 就可得到, 多项式矩阵 $A(\lambda)$ 可经初等变换变为对角矩阵 $\mathrm{diag}(u_1(\lambda), u_2(\lambda), \cdots)$.

实现整除性的方法

引理

对角矩阵 $\begin{bmatrix} f(\lambda) & \\ & g(\lambda) \end{bmatrix}$ 可经初等变换变为 $\begin{bmatrix} (f(\lambda), g(\lambda)) & \\ & [f(\lambda), g(\lambda)] \end{bmatrix}$.

证明.

由 Bezout 定理, 存在多项式 $u(\lambda)$, $v(\lambda)$, 使得 $u(\lambda)f(\lambda) + v(\lambda)g(\lambda) = (f(\lambda), g(\lambda))$. 因此, 有下面的初等变换过程

$$\begin{bmatrix} f & \\ & g \end{bmatrix} \to \begin{bmatrix} f & uf \\ & g \end{bmatrix} \to \begin{bmatrix} f & (f,g) \\ & g \end{bmatrix} \to \begin{bmatrix} 0 & (f,g) \\ -fg/(f,g) & g \end{bmatrix} \to \begin{bmatrix} 0 & (f,g) \\ -fg/(f,g) & \end{bmatrix},$$

再交换两列, 并将第二行乘以适当的非零常数, 就得到了 $\mathrm{diag}((f,g),[f,g])$. □

反复利用这个引理, 就可将对角的多项式矩阵变为定理所要求的逐个整除的形式.

注意作初等变换等价于乘以相应的初等矩阵, 所以上述定理也可叙述为

定理

若 $A(\lambda) \in P[\lambda]^{m\times n}$ 的秩为 r, 则存在可逆的多项式矩阵 $P(\lambda)$ 和 $Q(\lambda)$, 使得 $P(\lambda)A(\lambda)Q(\lambda) = \mathrm{diag}(d_1(\lambda), d_2(\lambda), \cdots, d_r(\lambda), 0, \cdots, 0)$, 其中首一多项式 d_1, d_2, $\cdots$, d_r 满足 $d_1 \mid d_2 \mid \cdots \mid d_r$.

实际操作时, 我们不必完全照搬前面证明的思路, 可以适当灵活处理.

例

求多项式矩阵 $\begin{bmatrix} \lambda-1 & 1 & -\lambda+1 & -\lambda-1 \\ -\lambda^2-2\lambda+2 & \lambda^2-2 & \lambda^2+2\lambda-2 & \lambda+2 \\ -\lambda^2-2\lambda & \lambda^2+\lambda & \lambda^2+2\lambda & 0 \end{bmatrix}$ 的相抵标准形.

解答

依次作如下变换: 第三列乘以 1 加到第一列, 第二列乘以 (-1) 加到第三列, 第三行乘以 (-1) 加到第二行, 第四列乘以 (-1) 加到第三列, 第三列乘以 (-1) 加到第二列, 第一行乘以 2 加到第二行, 第一行乘以 $(-\lambda)$ 加到第三行, 第二行乘以 λ 加到第三行, 第二列乘以 -1 加到第四列, 接下来容易得到相抵标准形为 $\mathrm{diag}(1, \lambda, \lambda)$.

行列式因子

定义

设 $A(\lambda) \in P[\lambda]^{m\times n}$ 的秩为 r. 若 $1 \leq k \leq r$, 则定义 $A(\lambda)$ 的 k 级行列式因子为 $A(\lambda)$ 的所有 k 级子式的最大公因式; 若 $k > r$, 则定义 $A(\lambda)$ 的 k 级行列式因子为 0.

例

多项式矩阵 $A(\lambda) = \begin{bmatrix} \lambda-1 & 1 & \lambda+2 \\ \lambda^2 & \lambda+1 & \lambda^2+2 \\ \lambda & 2 & -\lambda+4 \end{bmatrix}$ 的 1 级行列式因子为 1. 它的伴随矩阵为 $\begin{bmatrix} -3\lambda^2+3\lambda & 3\lambda & -3\lambda \\ 2\lambda^3-4\lambda^2+2\lambda & -2\lambda^2+2\lambda & 2\lambda^2-2\lambda \\ \lambda^2-\lambda & -\lambda & \lambda \end{bmatrix}$, 可见 $A(\lambda)$ 的 2 级行列式因子为 λ. 由于 $\det(A(\lambda)) = 0$, 所以它的 3 级行列式因子为 0.

行列式因子的相抵不变性

引理

初等变换不改变行列式因子. 因此, 相抵的多项式矩阵具有相同的行列式因子.

证明.

分三种情形讨论.

- 交换两行 (列) 时, 所有的 k 级子式与原来相比仅相差正负号, 所以 k 级行列式因子不变;
- 将某行 (列) 乘以非零常数 c 时, 新矩阵的 k 级子式要么与原来相等, 要么成为原来的 c 倍, 它们的最大公因式不变;
- 将某行 (列) 的 $\varphi(\lambda)$ 倍加到另一行 (列) 时, 新矩阵的每个 k 级子式要么与原来相等, 要么等于原来的一个 k 级子式加上另一个 k 级子式的 $\pm\varphi(\lambda)$ 倍. 因此, 原矩阵的 k 级行列式因子整除新矩阵的 k 级行列式因子. 同理可证反过来的整除关系.

□

这个引理的直接推论是: 相抵的多项式矩阵有相同的秩.

相抵标准形的唯一性

利用上述引理, 我们来证明多项式矩阵的行列式因子与不变因子是互相决定的.

命题

若多项式矩阵 $A(\lambda)$ 的行列式因子为 D_1, D_2, $\cdots$, D_r, 不变因子为 d_1, d_2, $\cdots$, d_r, 则有

$$D_1 = d_1, \quad D_2 = d_1 d_2, \quad \cdots, \quad D_r = d_1 d_2 \cdots d_r.$$

证明.

由于 $A(\lambda)$ 可经初等变换变为对角矩阵 $\mathrm{diag}(d_1, d_2, \cdots, d_r, 0, \cdots, 0)$, 所以 $A(\lambda)$ 与这个对角矩阵有相同的行列式因子. 由于 $d_1|d_2|\cdots|d_r$, 易知这个对角矩阵的各级行列式因子分别为 $D_1 = d_1$, $D_2 = d_1 d_2$, $\cdots$, $D_r = d_1 d_2 \cdots d_r$. □

推论

多项式矩阵 $A(\lambda) \in P[\lambda]^{m\times n}$ 的相抵标准形是唯一的.

由于可逆的 n 阶多项式矩阵的 n 级行列式因子为 1, 所以其所有行列式因子都是 1, 不变因子也都是 1, 我们有

推论

$A(\lambda) \in P[\lambda]^{n\times n}$ 是可逆的, 当且仅当 $A(\lambda)$ 能写为一些初等多项式矩阵的乘积.

思考题

- (*****) 给定 $A \in P^{n\times n}$. 设多项式矩阵 $\lambda E_n - A$ 的不变因子为 $d_1(\lambda)$, $d_2(\lambda), \cdots, d_n(\lambda)$. 证明: 所有与 A 可交换的矩阵构成的线性空间的维数是

$$(2n-1)\deg d_1(\lambda) + (2n-3)\deg d_2(\lambda) + \cdots + 3\deg d_{n-1}(\lambda) + \deg d_n(\lambda).$$

- (****) 设 $A \in P^{n\times n}$. 记 W 为所有与 A 可交换的矩阵构成的集合, U 为 A 的所有多项式构成的集合. 证明: $W = U$ 的充要条件是 A 的最小多项式与特征多项式相等.

多项式矩阵的带余除法

前面我们是从矩阵的角度来讨论多项式矩阵. 现在我们从多项式的角度来观察. 一元多项式可以作带余除法, 某些 $n \times n$ 多项式矩阵也可以作带余除法. 设

$$U(\lambda) = U_0 + U_1\lambda + \cdots + U_k\lambda^k, \quad V(\lambda) = V_0 + V_1\lambda + \cdots + V_l\lambda^l \in P[\lambda]^{n\times n},$$

那么, 当 $V(\lambda)$ 的 “首项系数” V_l 可逆时, 存在唯一的多项式矩阵 $Q(\lambda)$ 和 $R(\lambda)$, 使得

$$U(\lambda) = Q(\lambda)V(\lambda) + R(\lambda), \quad \deg R(\lambda) < l;$$

也存在唯一的多项式矩阵 $\tilde{Q}(\lambda)$ 和 $\tilde{R}(\lambda)$, 使得

$$U(\lambda) = V(\lambda)\tilde{Q}(\lambda) + \tilde{R}(\lambda), \quad \deg \tilde{R}(\lambda) < l.$$

分别称 $Q(\lambda)$ 和 $R(\lambda)$($\tilde{Q}(\lambda)$ 和 $\tilde{R}(\lambda)$) 为 $V(\lambda)$ 除 $U(\lambda)$ 的右商和 右余式 (左商和左余式).

在这里我们将不详细证明上述结论, 而只讨论一种简单的情形.

在一元多项式理论中, 有如下常用的余数定理: 多项式 $f(x)$ 除以一次多项式 $x-a$, 所得的余数是 $f(a)$. 现在我们把它推广到 $n\times n$ 多项式矩阵.

定义

对于多项式矩阵 $U(\lambda)=U_0+U_1\lambda+\cdots+U_k\lambda^k\in P[\lambda]^{n\times n}$ 和矩阵 $A\in P^{n\times n}$, 定义 $U(\lambda)$ 在 A 处的右值为

$$U(A)=U_0+U_1A+\cdots+U_kA^k.$$

类似地, 定义 $U(\lambda)$ 在 A 处的左值为

$$\tilde{U}(A)=U_0+AU_1+\cdots+A^kU_k.$$

引理

若 $A\in P^{n\times n}$, $U(\lambda)\in P[\lambda]^{n\times n}$, 则存在唯一的 $Q(\lambda)\in P[\lambda]^{n\times n}$ 和 $R\in P^{n\times n}$, 使得

$$U(\lambda)=Q(\lambda)(\lambda E_n-A)+R.$$

进一步, $R=U(A)$.

证明.

存在性: 在恒等式

$$\lambda^j E_n - A^j = (\lambda^{j-1}E_n + \lambda^{j-2}A + \cdots + \lambda A^{j-2} + A^{j-1})(\lambda E_n - A)$$

两端左乘 U_j, 再对 j 求和, 就得到

$$U(\lambda) - U(A) = Q(\lambda)(\lambda E_n - A).$$

唯一性: 若有 $Q_1(\lambda)$, R_1 和 $Q_2(\lambda)$, R_2 都使第一个等式成立, 则相减得 $(Q_1(\lambda) - Q_2(\lambda))(\lambda E_n - A) = R_2 - R_1$. 比较次数就得到 $Q_1(\lambda) - Q_2(\lambda) = 0$, 进而 $R_2 - R_1 = 0$. □

类似可以证明, 存在唯一的 $\tilde{Q}(\lambda)$, 使得

$$U(\lambda) = (\lambda E_n - A)\tilde{Q}(\lambda) + \tilde{U}(A).$$

利用上述引理, 我们可以把 Hamilton-Cayley 定理的证明简写为: 由 $(\lambda E_n - A)(\lambda E_n - A)^* = f(\lambda)E_n$ 可知, 多项式矩阵 $(\lambda E_n - A)$ 除 $f(\lambda)E_n$ 的左余式为零, 因此 $f(\lambda)E_n$ 在 A 处的左值为零, 即 $f(A) = 0$.

矩阵相似的条件

对于 $A \in P^{n\times n}$, 称 $\lambda E_n - A$ 为它的特征矩阵.

定理

设 $A, B \in P^{n\times n}$. 那么, A 与 B 相似的充分必要条件是 $\lambda E_n - A$ 与 $\lambda E_n - B$ 相抵, 即它们有相同的不变因子.

证明.

必要性: 若 A 与 B 相似, 则有可逆矩阵 T 使得 $A = T^{-1}BT$, 从而 $\lambda E_n - A = T^{-1}(\lambda E_n - B)T$, 可见 $\lambda E_n - A$ 与 $\lambda E_n - B$ 相抵.
充分性: 若 $\mathbb{A} = \lambda E_n - A$ 与 $\mathbb{B} = \lambda E_n - B$ 相抵, 则有可逆的多项式矩阵 U 和 V, 使得 $\mathbb{A} = U\mathbb{B}V$. 由引理, 存在多项式矩阵 Q, R, 使得

$$U = \mathbb{A}Q + U_0, \quad V = R\mathbb{A} + V_0.$$

其中 $U_0 = \tilde{U}(A)$, $V_0 = V(A)$. 我们的第一个目标是证明 $\mathbb{A} = U_0\mathbb{B}V_0$, 即

$$\lambda E_n - A = U_0(\lambda E_n - B)V_0.$$

证明 (续).

事实上

$$\begin{aligned} U_0\mathbb{B}V_0 =&(U-\mathbb{A}Q)\mathbb{B}(V-R\mathbb{A}) \\ =&\mathbb{A}-\mathbb{A}Q\mathbb{B}V-U\mathbb{B}R\mathbb{A}+\mathbb{A}Q\mathbb{B}R\mathbb{A} \\ =&\mathbb{A}-\mathbb{A}QU^{-1}\mathbb{A}-\mathbb{A}V^{-1}R\mathbb{A}+\mathbb{A}Q\mathbb{B}R\mathbb{A} \\ =&\mathbb{A}\{E_n-S(\lambda)\,\mathbb{A}\}, \end{aligned}$$

其中 $S(\lambda)=QU^{-1}+V^{-1}R-Q\mathbb{B}R$ 为多项式矩阵. 因此我们有

$$U_0(\lambda E_n-B)V_0=(\lambda E_n-A)\{E_n-S(\lambda)(\lambda E_n-A)\}.$$

比较两端的次数可知 $S(\lambda)=0$, 因此 $\lambda E_n-A=U_0(\lambda E_n-B)V_0$.
再比较两端一次项与常数项的系数可知, $U_0V_0=E_n$, $U_0BV_0=A$. 因此 A 与 B 相似. □

通常, 将多项式矩阵 λE_n-A 的不变因子 (行列式因子) 也称为矩阵 A 的不变因子 (行列式因子). 这样, 上述定理也可叙述为: A 与 B 相似, 当且仅当它们有相同的不变因子.

由于同一线性变换在不同基下的矩阵是相似的，它们有相同的不变因子 (行列式因子)，因此，我们也把它们称为该线性变换的不变因子 (行列式因子)．特别地，最高阶的行列式因子就是特征多项式.

思考题

- (***) 有同学认为，比较等式 $\lambda E_n - A = U(\lambda)(\lambda E_n - B)V(\lambda)$ 两端的次数，不就能得到 $U(\lambda)$，$V(\lambda)$ 为数字矩阵了吗？这个说法有什么问题？
- (**) 判断矩阵 $A = \begin{bmatrix} 1 & -1 & 1 \\ 1 & 3 & -2 \\ 1 & 2 & -1 \end{bmatrix}$ 与 $B = \begin{bmatrix} 1 & -1 & 1 \\ 0 & 2 & -1 \\ 0 & 1 & 0 \end{bmatrix}$ 是否相似.
- (****) 证明矩阵 $A \in P^{n\times n}$ 的第 n 个不变因子恰好是 A 的最小多项式.

伴侣矩阵

定义

首一多项式 $c(\lambda) = \lambda^m + c_{m-1}\lambda^{m-1} + \cdots + c_1\lambda + c_0 \in P[\lambda]$ 的伴侣矩阵是指如下的 $m \times m$ 矩阵

$$C = \begin{bmatrix} 0 & -c_0 \\ E_{m-1} & -\alpha \end{bmatrix}.$$

其中 $\alpha = (c_1, \cdots, c_{m-1})'$. 特别地, 0 次多项式的伴侣矩阵不存在.

引理

若 $C \in P^{m\times m}$ 是首一多项式 $c(\lambda)$ 的伴侣矩阵, 则多项式矩阵 $\lambda E_m - C$ 的相抵标准形是 $\mathrm{diag}(1, \cdots, 1, c(\lambda))$.

证明.

注意 $\lambda E_m - C$ 的左下角有一个 $(m-1)$ 级子式为 $(-1)^{m-1}$, 所以它的 $(m-1)$ 级行列式因子为 1, 从而 1 到 $(m-1)$ 级行列式因子全为 1, 而 m 级行列式因子为 $(\lambda E_m - C)$ 的行列式, 易知它恰好等于 $c(\lambda)$. □

Frobenius 标准形

利用矩阵相似的条件和上面的引理, 容易得到

定理

设 $A \in P^{n\times n}$, 并设 $\lambda E_n - A$ 与对角矩阵 $\mathrm{diag}(d_1(\lambda), d_2(\lambda), \cdots, d_n(\lambda))$ 相抵, 那么, A 相似于准对角矩阵 $\mathrm{diag}(C_1, C_2, \cdots, C_n)$, 其中 C_j 是首一多项式 $d_j(\lambda)$ 的伴侣矩阵. 特别地, 当 d_1, d_2, $\cdots$, d_n 为不变因子时, 相应的准对角矩阵称为 A 的有理标准形, 也称 *Frobenius* 标准形. 两矩阵相似, 当且仅当它们有相同的有理标准形.

证明.

特征矩阵 $\lambda E_n - \mathrm{diag}(C_1, C_2, \cdots, C_n)$ 仍是准对角矩阵, 我们可分别将各对角块变为标准形 $\mathrm{diag}(1, \cdots, 1, d_i)$. 从而该特征矩阵与 $\mathrm{diag}(d_1, \cdots, d_n)$ 相抵. (请思考: 1 到哪里去了?) □

注意不变因子满足 $d_1 \mid d_2 \mid \cdots \mid d_n$, 所以有理标准形中的对角块 C_1, C_2, $\cdots$, C_n 的阶数是递增的. 自然的问题是, 能否适当选取对角矩阵 $\mathrm{diag}(d_1, \cdots, d_n)$, 使得相应的对角块的阶数不太大呢? 一种可能的办法是考虑所谓初等因子.

初等因子

定义

将 $A \in P^{n\times n}$ 的不变因子分解为不可约多项式的乘积

$$
\begin{aligned}
d_1(\lambda) &= p_1(\lambda)^{e_{11}} \cdots p_s(\lambda)^{e_{s1}} \\
d_2(\lambda) &= p_1(\lambda)^{e_{12}} \cdots p_s(\lambda)^{e_{s2}} \\
\cdots & \quad \cdots \\
d_n(\lambda) &= p_1(\lambda)^{e_{1n}} \cdots p_s(\lambda)^{e_{sn}}
\end{aligned}
$$

其中 $p_i(\lambda)$ 为不同的首一不可约多项式, e_{ij} 为非负整数, $1 \leq i \leq s$, $1 \leq j \leq n$. 称所有形如 $p_i(\lambda)^{e_{ij}}$, $e_{ij} > 0$ 的多项式 (允许相同) 为 A 的初等因子.

例

如果 $A \in \mathbb{R}^{12\times 12}$ 的不变因子是 $1, \cdots, 1$(共 9 个 1), $(\lambda-1)^2$, $(\lambda-1)^2(\lambda+1)$, $(\lambda-1)^2(\lambda+1)(\lambda^2+1)^2$, 则 A 的初等因子为

$$
(\lambda-1)^2, \quad (\lambda-1)^2, \quad (\lambda-1)^2, \quad \lambda+1, \quad \lambda+1, \quad (\lambda^2+1)^2.
$$

由定义可以看出, 初等因子是由不变因子决定的. 那么反过来, 由初等因子能否确定不变因子呢? 我们重新来观察等式

$$\begin{array}{rcllll} d_1 & = & p_1^{e_{11}} & p_2^{e_{21}} & \cdots & p_s^{e_{s1}}, \\ d_2 & = & p_1^{e_{12}} & p_2^{e_{22}} & \cdots & p_s^{e_{s2}}, \\ \vdots & & & \vdots & & \\ d_n & = & p_1^{e_{1n}} & p_2^{e_{2n}} & \cdots & p_s^{e_{sn}}. \end{array}$$

由于 $d_1 \mid d_2 \mid \cdots \mid d_n$, 所以第 i 列的指数 e_{ij} 满足关系

$$e_{i1} \leq e_{i2} \leq \cdots \leq e_{in}.$$

这样, 一旦知道了初等因子, 即知道了 $e_{i1}, e_{i2}, \cdots, e_{in}$ 中大于零的那些数, 那么, 剩下的那些必定全为零. 把这些数按大小关系填入上面的等式, 就能得出左端的不变因子 d_j. 这样, 我们证明了

命题

矩阵的不变因子与初等因子是互相决定的.

例

设 $A \in R^{n\times n}$ 的初等因子为

$$\lambda-1,\ (\lambda-1)^2,\ (\lambda-1)^2,\ \lambda^2+5,\ (\lambda^2+5)^2,\ \lambda^2+\lambda+1,$$

试求 n 及 A 的不变因子.

解答

所有初等因子的乘积等于所有不变因子的乘积, 等于特征多项式, 特征多项式的次数就是 n. 由所给条件易得 $n=13$.

出现的不可约因式有 $(\lambda-1)$, (λ^2+5) 和 $(\lambda^2+\lambda+1)$. 将它们的指数递减排列, 并适当补上 0, 就可逐步得到

$$\begin{aligned}
d_{13} &= (\lambda-1)^2(\lambda^2+5)^2(\lambda^2+\lambda+1),\\
d_{12} &= (\lambda-1)^2(\lambda^2+5),\\
d_{11} &= \lambda-1, \qquad d_{10}=d_9=\cdots=d_1=1.
\end{aligned}$$

由于初等因子与不变因子互相决定, 我们有

定理

设 $A, B \in P^{n\times n}$, 则下面四个条件两两等价

(1) A 与 B 相似;

(2) $\lambda E_n - A$ 与 $\lambda E_n - B$ 相抵;

(3) A 与 B 有相同的不变因子;

(4) A 与 B 有相同的初等因子.

一般地, 先求不变因子再求初等因子比较麻烦, 下面这个结论可以带来一些便利.

命题

设 $A \in P^{n\times n}$. 如果多项式矩阵 $\lambda E_n - A$ 与对角矩阵 $\mathrm{diag}(h_1(\lambda), \cdots, h_n(\lambda))$ 相抵, 那么, 只要将 $h_i(\lambda)$ 分解为不可约首一多项式方幂的乘积, 就可得到 A 的所有初等因子.

证明.

考虑继续将对角矩阵 $\mathrm{diag}(h_1(\lambda), \cdots, h_n(\lambda))$ 化为标准形. 利用前面讲过的引理, 操作过程中每次选择对角线上的两个多项式 $u(\lambda)$, $v(\lambda)$, 可将它们变为 $u(\lambda)$ 与 $v(\lambda)$ 的最大公因式和最小公倍式. 在这个过程中, 每个不可约因式的指数只是发生了交换, 而从未改变. □

推论

把方阵 A_1 与 A_2 的初等因子合起来, 就是准对角矩阵 $\mathrm{diag}(A_1, A_2)$ 的所有初等因子.

证明.

对于准对角矩阵 $\mathrm{diag}(A_1, A_2)$, 可以分别对对角块 $\lambda E - A_1$ 和 $\lambda E - A_2$ 作初等变换, 分别将它们变为标准形. 这两部分合在一起, 一般并不是 $\lambda E - \mathrm{diag}(A_1, A_2)$ 的标准形. 但由上述定理, 只要将两部分多项式分解为不可约多项式的方幂, 就可找到所有初等因子. 因此, 结论成立. □

另一种有理标准形

定理

设矩阵 A 的所有初等因子为 $f_1, f_2, \cdots, f_r$, 各初等因子的伴侣矩阵分别为 $C_1, C_2, \cdots, C_r$, 则 A 相似于 $\mathrm{diag}(C_1, C_2, \cdots, C_r)$. 这个准对角矩阵有时也称为 A 的 (主) 有理标准形.

证明.

注意伴侣矩阵 C_j 的不变因子为 $1, \cdots, 1, f_j$, 从而初等因子仅有一个, 就是 f_j. 利用前面的结论, 准对角矩阵 $\mathrm{diag}(C_1, C_2, \cdots, C_r)$ 的初等因子恰好为 $f_1, f_2, \cdots, f_r$. 因此它与 A 相似. $\square$

思考题

(**) 求矩阵 $\begin{bmatrix} 3 & -4 & 0 & 2 \\ 4 & -5 & -2 & 4 \\ 0 & 0 & 3 & -2 \\ 0 & 0 & 2 & -1 \end{bmatrix}$ 的主有理标准形.

Jordan 标准形

引理

设 $\lambda_0 \in \mathbb{C}$, 则 *Jordan* 块 $J_k(\lambda_0)$ 的初等因子为 $(\lambda-\lambda_0)^k$.

证明.

注意多项式矩阵 $\lambda E_k - J_k(\lambda_0)$ 右上角 $k-1$ 级子式为 $(-1)^{k-1}$, 所以从 1 级到 $k-1$ 级行列式因子全为 1. 它的 k 级行列式因子就是行列式 $(\lambda-\lambda_0)^k$. 因此, 它的不变因子为 $1, \cdots, 1, (\lambda-\lambda_0)^k$, 初等因子仅有一个, 就是 $(\lambda-\lambda_0)^k$. □

利用这个引理, 类似于上一个定理的证明, 我们可得到

定理

若 $A \in \mathbb{C}^{n\times n}$ 的所有初等因子为 $(\lambda-\lambda_1)^{k_1}, \cdots, (\lambda-\lambda_s)^{k_s}$, 则 A 相似于 *Jordan* 矩阵

$$\operatorname{diag}\big(J_{k_1}(\lambda_1), \cdots, J_{k_s}(\lambda_s)\big).$$

推论

n 阶复矩阵 A 可对角化的充分必要条件是所有初等因子都是 1 次的.

例

求矩阵 $B=\begin{bmatrix}2&0&0&0\\-1&1&0&0\\0&-1&0&-1\\1&1&1&2\end{bmatrix}$ 的 Jordan 标准形.

解答

对 B 的特征矩阵 λE_4-B 作初等变换如下

$$\begin{bmatrix}\lambda-2&&&\\1&\lambda-1&&\\&1&\lambda&1\\-1&-1&-1&\lambda-2\end{bmatrix}\to\begin{bmatrix}\lambda-2&&&\\1&\lambda-1&&\\&1&\lambda&1\\-1&1-\lambda&-(\lambda-1)^2&\end{bmatrix}$$

$$\to\begin{bmatrix}\lambda-2&&&\\1&\lambda-1&&\\&&&1\\-1&1-\lambda&-(\lambda-1)^2&\end{bmatrix}\to\begin{bmatrix}\lambda-2&&&\\1&\lambda-1&&\\&&&1\\&&-(\lambda-1)^2&\end{bmatrix}$$

解答 (续)

$$
\rightarrow \begin{bmatrix} & -(\lambda-1)(\lambda-2) & & \\ 1 & \lambda-1 & & \\ & & & 1 \\ & & -(\lambda-1)^2 & \end{bmatrix}
$$

$$
\rightarrow \begin{bmatrix} & -(\lambda-1)(\lambda-2) & & \\ 1 & & & \\ & & & 1 \\ & & -(\lambda-1)^2 & \end{bmatrix}.
$$

可见, $\lambda E_4 - B$ 与对角矩阵 $\mathrm{diag}\big(1, 1, (\lambda-1)(\lambda-2), (\lambda-1)^2\big)$ 相抵, 其初等因子为 $\lambda-2$, $\lambda-1$, $(\lambda-1)^2$. 因此, B 的 Jordan 标准形为 $\mathrm{diag}\big(2, 1, J_2(1)\big)$.

思考题

(**) 用这一节的知识重新证明 $A \in P^{n\times n}$ 的第 n 个不变因子是它的最小多项式.