高等代数（上海市精品课程，制作人：朱胜林，mazhusl@fudan.edu.cn ）

朱胜林

第七章
相似标准型

我的信息

朱胜林

办公室:	光华东楼 1708
办公电话:	55664896
办公时间:	周二下午 1：30 — 3：30
电子邮件:	mazhusl@fudan.edu.cn
个人主页:	homepage.fudan.edu.cn/~zhusl
本课程主页:	jpkc.fudan.edu.cn/s/100/

多项式矩阵

在本节中，设 λ 是一个未定元，多项式在未指明的情况下是指 λ 的多项式。

定义

由数域 K 上 λ 的多项式作为项构成矩阵称为 λ-矩阵，有时也称为 λ 的多项式矩阵。λ-矩阵的加、减、乘、行列式等运算如同 K 中矩阵的相应运算一样定义。

λ-矩阵的许多运算性质在不涉及多项式的逆的情况下和项全部为数域 K 的数的矩阵的性质基本相同。为方便计，称数域 K 的数的矩阵为数字矩阵。
一个 λ-矩阵总是可写成一个矩阵多项式，即以"矩阵作为系数"的多项式。如

$$\begin{pmatrix} 2\lambda^2+3\lambda+1 & -3\lambda+\pi \\ 2\lambda-9 & -\dfrac{3}{2}\lambda \end{pmatrix} = \lambda^2\begin{pmatrix} 2 & 0 \\ 0 & 0 \end{pmatrix} + \lambda\begin{pmatrix} 3 & -3 \\ 2 & -\dfrac{3}{2} \end{pmatrix} + \begin{pmatrix} 1 & \pi \\ -9 & 0 \end{pmatrix}。$$

多项式矩阵

定义 (初等行（列）变换)

对于一个 λ-矩阵，下列变换称为初等变换

1. 将两行（列）对换；
2. 对矩阵的一行（列）乘上非零常数；
3. 将矩阵的一行（列）的多项式倍加到另一行（列）。

一个矩阵 $A(\lambda)$ 在实施了若干次上述的初等变换后得到矩阵 $B(\lambda)$，则称 $A(\lambda)$ 与 $B(\lambda)$ 作为 λ-矩阵相抵。n 阶单位阵 I_n 经过一次初等变换得到的矩阵称为初等矩阵。如同数字矩阵一样，我们仍有"左行右列"的说法：

定理 (左行右列原则)

对一个 λ-矩阵左乘一个初等矩阵，其结果相当于实施了一次相应的初等行变换；对一个 λ-矩阵右乘一个初等矩阵，其结果相当于实施了一次相应的初等列变换。

若一个 λ-方阵的逆矩阵也是 λ-矩阵，则称该方阵为可逆的 λ-矩阵。

定义（可逆 λ-矩阵）

设 $A(\lambda)$ 是一个 n 阶 λ-方阵，如果存在一个 n 阶 λ-方阵 $B(\lambda)$，成立

$$A(\lambda)B(\lambda) = B(\lambda)A(\lambda) = I_n,$$

则称 $A(\lambda)$ 为一个可逆 λ-矩阵，矩阵 $B(\lambda)$ 称为 $A(\lambda)$ 的逆矩阵，记为 $A^{-1}(\lambda)$。

通过矩阵与其伴随矩阵的乘积的等式，以及矩阵乘积的行列式与行列式之积相等，我们得到：

定理

一个 n 阶 λ-方阵 $A(\lambda)$ 可逆当且仅当 $A(\lambda)$ 的行列式 $|A(\lambda)| \in K$ 是一个非零数。

初等矩阵都是可逆矩阵。在下一节我们会看到：一个 λ-矩阵可逆当且仅当它是初等矩阵之积，且两个 λ-矩阵 $A(\lambda)$、$B(\lambda)$ 相抵当且仅当存在可逆的 λ-矩阵 $P(\lambda)$ 和 $Q(\lambda)$，使得

$$P(\lambda)A(\lambda)Q(\lambda) = B(\lambda)。$$

两个矩阵多项式有时也可以进行类似于数字矩阵的带余除法，但条件是作为“除数”的矩阵多项式的“首项”矩阵是可逆矩阵。

引理

设 $A(\lambda)$ 和 $B(\lambda)$ 都是 n 阶 λ-方阵，
且 $B(\lambda)=\lambda^m B_m+\lambda^{m-1}B_{m-1}+\cdots+\lambda B_1+B_0$ 的表达式中的 B_m 是可逆阵，则必存在唯一的 λ-矩阵 $Q(\lambda)$ 和 $R(\lambda)$，使得

$$A(\lambda)=B(\lambda)Q(\lambda)+R(\lambda),$$

其中矩阵 $R(\lambda)$ 的每一项的次数均小于 m；
同时也存在唯一的 λ-矩阵 $\overline{Q(\lambda)}$ 和 $\overline{R(\lambda)}$，使得

$$A(\lambda)=\overline{Q(\lambda)}B(\lambda)+\overline{R(\lambda)},$$

其中矩阵 $\overline{R(\lambda)}$ 的每一项的次数均小于 m。

引理

设 A、B 为 K 上 n 阶方阵，若存在 n 阶方阵 P、Q，满足

$$\lambda I_n - A = P(\lambda I_n - B)Q,$$

则 A 和 B 相似。

有了以上的准备工作，我们就得到了这一节最重要定理。这个定理也是以 λ-矩阵方法推导出 Jordan 标准型理论的重要依据。

定理 (数域 K 上矩阵相似与其 λ-矩阵相抵的关系)

设 A、B 是数域 K 上的 n 阶方阵，则 A、B 相似的充分必要条件是相应的 λ-矩阵 $\lambda I_n - A$ 与 $\lambda I_n - B$ 相抵（作为 λ-矩阵）。

下面一个结论的证明留在以后，这里先叙述结论。

推论

设数域 $K \subseteq L$，并且 A、B 均为 K 上的同阶方阵，若 A、B 在数域 L 中相似，则必在数域 K 中相似。

对于 K 上的矩阵，我们知道相抵的矩阵中有一个标准型 $\begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}$，对于 λ-矩阵来说，因为并不是所有的非零多项式在 $K[\lambda]$ 中可逆，所以其标准型略微有些区别。这一节，我们导出 λ-矩阵的相抵标准型，并在下一节中证明该标准型是唯一的。

引理

设 $A(\lambda) = \left(a_{ij}(\lambda)\right)_{m\times n}$ 是一个非零 λ-矩阵中，则 $A(\lambda)$ 必相抵于一个 λ-矩阵 $B(\lambda) = \left(b_{ij}(\lambda)\right)_{m\times n}$，其中 $b_{11}(\lambda)$ 整除 $B(\lambda)$ 中的每一项。

证明:

对 $k = \min\left\{\deg a_{ij}(\lambda) \mid a_{ij}(\lambda) \neq 0, i = 1, \ldots, m; j = 1, \ldots, n\right\}$ 归纳。

$\square$

这个引理可导出 λ-矩阵的相抵标准型：

定理 (λ-矩阵的相抵标准型定理)

任一 $\lambda-$ 矩阵 $A(\lambda)_{m\times n}$ 都相抵于如下形式的矩阵

$$\begin{pmatrix} d_1(\lambda) & & & & & & \\ & d_2(\lambda) & & & & & \\ & & \ddots & & & & \\ & & & d_r(\lambda) & & & \\ & & & & 0 & & \\ & & & & & \ddots & \\ & & & & & & 0 \end{pmatrix},$$

其中 $r \le \min\{m,n\}$，$d_i(\lambda)$ $(i=1,2,\cdots,r)$ 都是首一多项式，且有

$$d_i(\lambda)|d_{i+1}(\lambda),\quad (i=1,2,\cdots,r-1)。$$

上面定理中的标准型称为矩阵 $A(\lambda)$ 的法式。如果 $A(\lambda)$ 是 n 阶可逆λ-矩阵，仍有上述的标准型成立。这时，按初等变换矩阵的行列式性质，知有 $|A(\lambda)| = c\prod_{i=1}^{r} d_i(\lambda) \times \underbrace{0\times\cdots\times 0}_{n-r\text{个}}$，其中 c 为非零数，而 $|A(\lambda)|$ 为一个非零数，于是得到 $r=n$ 且 $d_i(\lambda)=1$。

命题

1. 可逆的 λ-矩阵必是初等矩阵的积；
2. 若 A 为 n 阶数字矩阵，则 $|\lambda I_n - A| = \prod_{i=1}^{n} d_i(\lambda)$，其中 $d_1(\lambda) \mid d_2(\lambda) \mid \cdots \mid d_n(\lambda)$ 是 $\lambda I_n - A$ 的法式中的对角元。

下面我们给个具体的例子。

例

求矩阵 $\lambda I_3 - A$ 的法式，这里 $A = \begin{pmatrix} -3 & 12 & 16 \\ 1 & -2 & -4 \\ -2 & 6 & 9 \end{pmatrix}$。

$$\lambda I_3 - A = \begin{pmatrix} \lambda+3 & -12 & -16 \\ -1 & \lambda+2 & 4 \\ 2 & -6 & \lambda-9 \end{pmatrix} \Rightarrow \begin{pmatrix} 1 & 0 & 0 \\ 0 & (\lambda+6)(\lambda-1) & 4\lambda-4 \\ 0 & 2\lambda-2 & \lambda-1 \end{pmatrix}$$

$$\Rightarrow \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2\lambda-2 & \lambda-1 \\ 0 & (\lambda-1)(\lambda+6) & 4\lambda-4 \end{pmatrix} \Rightarrow \begin{pmatrix} 1 & 0 & 0 \\ 0 & \lambda-1 & 0 \\ 0 & 0 & \lambda^2-3\lambda+2 \end{pmatrix}$$

这一节，我们将证明 λ-矩阵的法式是唯一的，并给出计算法式的别一个途径。

定义（行列式因子）

设 $A(\lambda)$ 是一个 n 阶 λ-矩阵，$1 \le k \le n$，则 $A(\lambda)$ 的所有 k 阶子式的最大公因式称为 $A(\lambda)$ 的 k 阶行列式因子，记首一的 k 阶行列式因子为 $D_k(\lambda)$。若所有的 k 阶子式均为零，则约定 k 阶行列式因子 $D_k(\lambda)$ 为零。

命题

对任一 n 阶 λ-矩阵 $A(\lambda)$，必有 $D_1(\lambda) \mid D_2(\lambda) \mid \cdots \mid D_{n-1}(\lambda) \mid D_n(\lambda)$。

下面我们将通过说明行列式因子在初等变换下是不变的来说明 λ-矩阵的法式是唯一的。

命题

相抵的 λ-矩阵有相同的行列式因子。

定理

设 $A(\lambda)$ 为 n 阶 λ-矩阵，若 $A(\lambda)$ 的法式

为 $$\begin{pmatrix} d_1(\lambda) & & & & & & \\ & d_2(\lambda) & & & & & \\ & & \ddots & & & & \\ & & & d_r(\lambda) & & & \\ & & & & 0 & & \\ & & & & & \ddots & \\ & & & & & & 0 \end{pmatrix}$$ ，则必有

$$\begin{aligned} d_1(\lambda) &= D_1(\lambda),\ d_2(\lambda) = D_2(\lambda)/D_1(\lambda),\ \cdots, \\ d_r(\lambda) &= D_r(\lambda)/D_{r-1}(\lambda),\ D_{r+1}(\lambda) = \cdots = D_n(\lambda) = 0; \end{aligned}$$

因此 $A(\lambda)$ 的法式是唯一的。

定义

n 阶 λ-矩阵 $A(\lambda)$ 的法式中的 $d_1(\lambda), d_2(\lambda), \ldots, d_r(\lambda)$ 称为 $A(\lambda)$ 的不变因子组。

结合前面给出的矩阵的相似与其 λ-矩阵相抵等价的命题，下列结论是自然的。

定理

设 A 和 B 是数域 K 上同阶方阵，则 A 与 B 相似的充分必要条件是 $\lambda I_n - A$ 与 $\lambda I_n - B$ 有相同的不变因子组或是有相同的初等因子组。

由于数字矩阵的 λ-矩阵的相抵标准型可能通过其中项经过若干次带余除法得到，而每一个带余除法的商和余式都是可以在原来的数域中进行，所以最后得到的法式还是原来数域上的多项式矩阵。所以有

推论（相似与扩域的无关性）

一个数域上的两个数字矩阵，若在该域的某个扩域上相似，则必在原来的域上相似。

有理标准型（ Frobenius 标准型）

这节中，我们可以根据一个数字矩阵 A 所对应的 λ-矩阵 $\lambda I - A$ 的法式，构造出最'粗略'的标准型。下列引理在标准型结论的证明中很有用。

引理

设 $f(\lambda), g(\lambda) \in K[\lambda]$，则

1. 矩阵 $\begin{pmatrix} f(\lambda) & \\ & g(\lambda) \end{pmatrix}$ 与矩阵 $\begin{pmatrix} g(\lambda) & \\ & f(\lambda) \end{pmatrix}$ 相抵；
2. 若 $(f(\lambda), g(\lambda)) = 1$，则 $\begin{pmatrix} f(\lambda) & \\ & g(\lambda) \end{pmatrix}$ 还与矩阵 $\begin{pmatrix} 1 & \\ & f(\lambda)g(\lambda) \end{pmatrix}$ 相抵。
3. 设 $g(\lambda) = \lambda^n + a_1\lambda^{n-1} + \cdots + a_{n-1}\lambda + a_n \in K[\lambda]$，定义
$$F_g = \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \\ -a_n & -a_{n-1} & -a_{n-2} & \cdots & -a_1 \end{pmatrix}_{n\times n},$$
则 $\lambda I_n - F_g$ 的法式为 $\operatorname{diag}(1, \ldots, 1, g(\lambda))$。

上面的引理是本节有理标准型及初等因子、Jordan 标准型理论的依据。
下面构造一个矩阵的最‘粗略’的标准型。
对一个数字矩阵 $A_{n\times n}$，其法式为 $\mathrm{diag}\left(1,\cdots,1,d_1(\lambda),d_2(\lambda),\ldots,d_k(\lambda)\right)$，
其中 $1\neq d_1(\lambda)\mid d_2(\lambda)\mid\cdots\mid d_k(\lambda)$，且 $n-k$ 个 1。如果
记 $\deg d_i(\lambda)=n_i$，则有 $|\lambda I_n-A|=d_1(\lambda)d_2(\lambda)\cdots d_k(\lambda)$，于是
有 $n=\sum_{i=1}^k n_i$，且 $\sum_{i=1}^k(n_i-1)=n-k$。由上引理，在对角矩阵中交换对角
元的位置，是一个相抵变换，$\lambda I_n-A\approx$

$$\begin{pmatrix}1&&&&&\\&\ddots&&&&\\&&1&&&\\&&&d_1(\lambda)&&\\&&&&\ddots&\\&&&&&d_k(\lambda)\end{pmatrix}\approx\begin{pmatrix}1&&&&&&&&\\&\ddots&&&&&&&\\&&1&&&&&&\\&&&d_1(\lambda)&&&&&\\&&&&\ddots&&&&\\&&&&&1&&&\\&&&&&&\ddots&&\\&&&&&&&1&\\&&&&&&&&d_k(\lambda)\end{pmatrix}$$

式中每个 $d_i(\lambda)$ 的左上角 1 恰有 n_i-1 个。

所以

$$\lambda I_n - A \approx\approx \begin{pmatrix} \lambda I_{n_1} - F_{d_1} & & \\ & \ddots & \\ & & \lambda I_{n_1} - F_{d_k} \end{pmatrix} = \lambda I_n - \begin{pmatrix} F_{d_1} & & \\ & \ddots & \\ & & F_{d_k} \end{pmatrix},$$

得到

$$\lambda I_n - A \approx \lambda I_n - \begin{pmatrix} F_{d_1} & & \\ & \ddots & \\ & & F_{d_k} \end{pmatrix}。$$

定理

设 n 阶方阵 A 的不变因子组（略去 1 后）为 $d_1(\lambda) \mid d_2(\lambda) \mid \cdots \mid d_k(\lambda)$，则 A 必相似于矩阵 $F = \begin{pmatrix} F_{d_1} & & \\ & \ddots & \\ & & F_{d_k} \end{pmatrix}$ 相似，且 A 的最小多项式恰为 $d_k(\lambda)$。F 为 A 的有理标准型（也称为 A 的Frobenius 标准型）。

例

求矩阵 A、B 的有理标准型，这里 $A=\begin{pmatrix}-3&12&16\\1&-2&-4\\-2&6&9\end{pmatrix}$，$B=\begin{pmatrix}1&\\&2\end{pmatrix}$。

解： 前面已经说明过，$\lambda I_3-A\approx\begin{pmatrix}1&0&0\\0&\lambda-1&0\\0&0&\lambda^2-3\lambda+2\end{pmatrix}$，所以 A 的有理标准型为

$$\begin{pmatrix}F_{\lambda-1}&\\&F_{\lambda^2-3\lambda+2}\end{pmatrix}=\begin{pmatrix}1&&\\&0&1\\&-2&3\end{pmatrix};$$

$\lambda I_2-\begin{pmatrix}1&\\&2\end{pmatrix}=\begin{pmatrix}\lambda-1&\\&\lambda-2\end{pmatrix}\approx\begin{pmatrix}1&\\&(\lambda-1)(\lambda-2)\end{pmatrix}$，所以 $\begin{pmatrix}1&\\&2\end{pmatrix}$ 的有理标准型为 $\begin{pmatrix}0&1\\-2&3\end{pmatrix}$。 □

前面给出了 Frobenius 标准型的理论。但由上面的例子可以看出，有时本来一个简单的矩阵，它的 Frobenius 标准型反而变得更加复杂了。我们复述前面的引理并将第三条做些改变：

引理

设 $f(\lambda), g(\lambda) \in K[\lambda]$，则

1. 矩阵 $\begin{pmatrix} f(\lambda) & \\ & g(\lambda) \end{pmatrix}$ 与矩阵 $\begin{pmatrix} g(\lambda) & \\ & f(\lambda) \end{pmatrix}$ 相抵；
2. 若 $(f(\lambda), g(\lambda)) = 1$，则 $\begin{pmatrix} f(\lambda) & \\ & g(\lambda) \end{pmatrix}$ 还与矩阵 $\begin{pmatrix} 1 & \\ & f(\lambda)g(\lambda) \end{pmatrix}$ 相抵。
3. 对 $\forall a \in K$，$n \in \mathbb{Z}^+$，定义 $J_n(a) = \begin{pmatrix} a & 1 & 0 & \cdots & 0 \\ 0 & a & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & 0 & a & 1 \\ 0 & 0 & \cdots & 0 & a \end{pmatrix}_{n\times n}$，
   则 $\lambda I_n - J_n(a)$ 的法式为 $\mathrm{diag}\left(1, \ldots, 1, (\lambda - a)^n\right)$。

遇到复杂的多项式，我们还可以使用引理中的 2，将其分解为简单的情形。注意，分解中的条件，$(f(\lambda), g(\lambda)) = 1$，也就是说，将之分解成互素的部分！

所以，我们有下面的初等因子的定义。

定义

设 A 为 n 阶方阵，A 的不等于 1 的不变因子在数域 K 分解成互素的不可约多项式方幂的乘积，这些不可约多项式方幂（一个方幂可能多次出现，每次出现都应该计算在内）称为矩阵 A 的初等因子。

由于不变因子组 $d_1(\lambda) \mid d_2(\lambda) \mid \cdots \mid d_n(\lambda)$，所以从初等因子组可以重建不变因子组，即若设初等因子组中所有涉及到的不可约因式为 $p_1(\lambda), \ldots, p_r(\lambda)$，其中一个不可约因式的幂在初等因子中出现次数最多，设为 k 次，然后把初等因子按从左到右依次整除（指标依次增加）填入下表（不够时填入 1 ）：

$$\begin{array}{cccc} & p_1^{n_{1,1}}(\lambda) & \cdots & p_1^{n_{1,n}}(\lambda) \\ & \cdots & \cdots & \cdots \\ \times & p_r^{n_{r,1}}(\lambda) & \cdots & p_r^{n_{r,n}}(\lambda) \\ & d_1(\lambda) & \cdots & d_n(\lambda) \end{array}，\text{则取 } d_i(\lambda) = \prod_{j=1}^{r} p_j^{n_{j,i}}(\lambda)，\text{得}$$

到 $d_1(\lambda) \mid d_2(\lambda) \mid \cdots \mid d_n(\lambda)$ 则为 A 的不变因子组。

所以，由初等因子组可决定不变因子组，而前面已说明，有相同的不变因子组的矩阵相似，于是有：

定理

设 K 是数域，K 上两个 n 阶矩阵 A、B 相似当且仅当它们的初等因子组相同。

当我们取数域为 $\mathbb{C}$ 时，由于 $\mathbb{C}$ 上的不可约多项式只能是一次式，所以其上每个方阵的初等因子组必全部由一次式的幂 $(\lambda-a)^k$ 等构成，按照有理标准型的证明方式，只需找到 k 阶方阵 B，使得 $\lambda I_k - B \approx \operatorname{diag}\left(1,\ldots,1,(\lambda-a)^k\right)$。前面的引理给出了这样一个矩阵。所以，我们有

定理

设矩阵 A 在 $\mathbb{C}$ 上的初等因子组为 $(\lambda-\lambda_1)^{n_1},\ldots,(\lambda-\lambda_k)^{n_k}$，则 A 必相似于矩阵 $\begin{pmatrix} J_{n_1}(\lambda_1) & & \\ & \ddots & \\ & & J_{n_k}(\lambda_k) \end{pmatrix}$。

上述定理中的矩阵称为 A 的 Jordan 标准型，形如上述标准型形式的矩阵称为 Jordan 矩阵。从初等因子组的内容知道，矩阵 A 的 Jordan 标准型在不计每块在对角线的位置时是唯一的。

推论

复数域的线性空间上的任一变换 φ，均存在一组基，满足它在这组基下的矩阵为 Jordan 矩阵。

推论

设 A 是数域 K 上的 n 阶方阵，则下列说法等价：

1. A 在数域 K 上可对角化；
2. A 的初等因子组均由一次式构成；
3. A 的不变因子组中，$d_n(\lambda)$ 的根均在 K 中且没有重根；
4. A 的极小多项式的根均在 K 中且无重根。

当 $K=\mathbb{C}$ 时，上述结论总是正确。

例

讨论矩阵 $A=\begin{pmatrix}1 & -3 & 0 & 3\\ -2 & -6 & 0 & 13\\ 0 & -3 & 1 & 2-a\\ -1 & -4 & 0 & 8\end{pmatrix}$ 的 Jordan 标准型，并在 A 的 Jordan 标准型中 Jordan 块数大于 1 时，求可逆矩阵 P，使 $P^{-1}AP$ 为 Jordan 标准型。

解： $\det(\lambda I_4-A)=(\lambda-1)^4$，$A$ 的 Jordan 块数等于特征向量组的秩。现 A 只有一个特征值，$I_4-A=\begin{pmatrix}0 & 3 & 0 & -3\\ 2 & 7 & 0 & -13\\ 0 & 3 & 0 & a-2\\ 1 & 4 & 0 & -7\end{pmatrix}\Rightarrow\begin{pmatrix}1 & 4 & 0 & -7\\ 0 & 1 & 0 & -1\\ 0 & 0 & 0 & a+1\\ 0 & 0 & 0 & 0\end{pmatrix}$，$a\neq-1$ 时只能找到一个无关解，Jordan 块只能有一块。故 $a=-1$。此时得解 $v_1=\begin{pmatrix}0 & 0 & 1 & 0\end{pmatrix}'$，$v_2=\begin{pmatrix}3 & 1 & 0 & 1\end{pmatrix}'$。本题接下来，对 $(A-I_4)x=v_1$ 和 $(A-I_4)x=v_2$ 都找不到解，但 $(A-I_4)x=3v_1+v_2$ 能找到解 $v_3=\begin{pmatrix}3-3x_4 & -x_4-1 & -x_3 & -x_4\end{pmatrix}'$，同理判断当 $x_3-3x_4+3=0$ 时可找到 $(A-I_4)x=v_3$ 的解 $v_4=\begin{pmatrix}3x_4-4 & 1-x_4 & 0 & 0\end{pmatrix}'$，取 $P=\begin{pmatrix}v_1 & v_2 & v_3 & v_4\end{pmatrix}$，即有 $P^{-1}AP=\operatorname{diag}(J_1(1),J_3(1))$。 □

Jordan 标准型的进一步讨论和应用举例

设 $\varphi \in \mathcal{L}(V)$ 在 V 的一组基下的矩阵为 Jordan 矩阵 $J = \operatorname{diag}(J_{n_1}(\lambda_1), J_{n_2}(\lambda_2), \ldots, J_{n_k}(\lambda_k))$，记此组基为 $v_1, \ldots, v_{n_1}; v_{n_1+1}, \ldots, v_{n_1+n_2}; \cdots; v_{n_1+\cdots+n_{k-1}+1}, \ldots, v_n$；则有

$$\begin{cases} \varphi(v_1) = \lambda_1 v_1 \\ \varphi(v_2) = v_1 + \lambda_1 v_2 \\ \cdots \\ \varphi(v_{n_1-1}) = v_{n_1-2} + \lambda_1(v_{n_1-1}) \\ \varphi(v_{n_1}) = v_{n_1-1} + \lambda_1(v_{n_1}) \end{cases}, \text{或是：} \begin{cases} (\varphi - \lambda_1 I)(v_1) = 0 \\ (\varphi - \lambda_1 I)(v_2) = v_1 \\ \cdots \\ (\varphi - \lambda_1 I)(v_{n_1-1}) = v_{n_1-2} \\ (\varphi - \lambda_1 I)(v_{n_1}) = v_{n_1-1} \end{cases}$$

于是，我们有 φ 的一个不变子空间 V_1：它由 $v_1, \ldots, v_{n_1}$ 线性张成，以另一种方式看待，该空间是由 $v_{n_1}, (\varphi - \lambda_1 I)(v_{n_1}), (\varphi - \lambda_1 I)^2(v_{n_1}), \ldots,$ $(\varphi - \lambda_1 I)^{n_1-1}(v_{n_1})$ 张成且 $(\varphi - \lambda_1 I)^{n_1}(v_{n_1}) = 0$；对于其它的对角 Jordan 块，均可以找到如此的不变线性子空间。

定义

设 $\psi \in \mathcal{L}(V)$，$\alpha \in V$ 且 $\psi^{k-1}(\alpha) \neq 0$ 而 $\psi^k(\alpha) = 0$，则由 $\alpha, \psi(\alpha), \ldots, \psi^{k-1}(\alpha)$ 生成的 V 的子空间是 ψ 的一个不变子空间，称为 ψ 的一个循环子空间。

这时，容易证明，$\alpha, \psi(\alpha), \ldots, \psi^{k-1}(\alpha)$ 是该循环子空间的一组基。

Jordan 标准型的进一步讨论和应用举例

所以有

定理

设 $\varphi \in \mathcal{L}(V)$ 在 V 的一组基下的矩阵为 Jordan 矩阵 $J = \mathrm{diag}\left(J_{n_1}(\lambda_1), J_{n_2}(\lambda_2), \ldots, J_{n_k}(\lambda_k)\right)$，则存在 V 的一个直和分解

$$V = \bigoplus_{i=1}^{k} V_i，其中V_i是\psi - \lambda_i I的一个n_i维循环子空间。$$

这个分解并不是唯一的（即使不计次序），读者可以以维数大于 1 的线性空间上的恒等变换为例自己说明。不唯一性主要是同一个特征值可能对应的Jordan 块不只一个。但如果我们在 J 中把同一个特征值对应的 Jordan 块相应的循环子空间的都和在一起，这时就得到了 V 的另一种直和分解（每一个特征值恰对应到一个子空间），这个分解是唯一的。

Jordan 标准型的进一步讨论和应用举例

注意到 Jordan 块的性质：$J_m(0)^{m-1} \neq 0$，而 $J_m(0)^m = 0$，对应于特征值 μ 的上述空间可由 $\ker(\varphi - \mu I)^n$ 而得到。我们称这样的子空间为属于特征值 μ 的根子空间。

定理

设 V 是 $\mathbb{C}$ 上的 n 维线性空间，$\varphi \in \mathcal{L}(V)$ 的不同特征值为 $\lambda_1, \ldots, \lambda_k$，记 $R(\lambda_i) = \ker(\varphi - \lambda_i I)^n$ 为根子空间，则

$$V = \bigoplus_{i=1}^{k} R(\lambda_i)$$

称为 V 的关于 φ 的根子空间分解。

例

$$\begin{pmatrix} \lambda & 1 & & & \\ & \lambda & 1 & & \\ & & \ddots & \ddots & \\ & & & \lambda & 1 \\ & & & & \lambda \end{pmatrix} = \begin{pmatrix} & & & 1 & \lambda \\ & & 1 & \lambda & \\ & \iddots & \iddots & & \\ 1 & \iddots & & & \\ \lambda & & & & \end{pmatrix} \begin{pmatrix} & & & 1 \\ & & 1 & \\ & \iddots & & \\ 1 & & & \end{pmatrix},$$

所以一个 Jordan 块是两个对称矩阵之积，由此得 $\mathbb{C}$ 上每个方阵也是两个对称矩阵之积。原因：$A = P^{-1}JP = P^{-1}BCP = \underline{P^{-1}B\left(P^{-1}\right)'}\ \underline{P'CP}$ 为两个对称矩阵之积。

对于 Jordan 块的幂，有如下的结论，记 $J_0 = J_n(0), J_\lambda = J_n(\lambda)$：

1. 则 J_0A 的结果为把 A 的每行向上移一位，最后一行补零；AJ_0 的结果是将 A 的每一列向右移一位，第一列补零；
2. J_0^k 的结果为把 J_0 中的 1 向右上角移 k 位（$k < n$），$J_0^n = 0$；
3. 当 $\lambda \neq 0$ 时，J_λ^k 在距主对角线 l 位的对角平行线上，值为 $\dfrac{1}{l!}\dfrac{d^l}{d\lambda^l}\lambda^k$；所以有 $f(J_\lambda)$ 在距主对角线 l 位的对角平行线上的值为 $\dfrac{1}{l!}\dfrac{d^l}{d\lambda^l}f(\lambda)$，$f(x) \in \mathbb{C}[x]$。

下列的定理在代数中是一个很重根的定理。

定理 (Jordan-Chevalley 定理)

设 A 是 $\mathbb{C}$ 上 n 阶方阵，则 A 必可分解为两个矩阵之和 $A = B + C$，其中

1. B 相似于一个对角阵；
2. C 是幂零阵；
3. B 与 C 乘法可交换。
   上述分解是唯一的，且分解中的矩阵 B、C 都是矩阵 A 的多项式。

这个分解称为 Jordan-Chevalley 分解。

对于一个方阵，我们可以代入到一个多项式，这样得到一个同阶方阵。是不是可以把一个方阵代入到一个更加一般点的函数呢？在讨论复数上的函数时，通常是把这个函数看成幂级数—即多项式的极限函数。

如 $e^z = 1 + z + \frac{z^2}{2!} + \cdots + \frac{z^k}{k!} + \cdots$，等。如同实数域上，复级数、复幂级数也有收敛性、收敛半径的概念，许多结论和实数时类似。

定义（矩阵列的收敛）

设有 n 阶复方阵构成的序列 $\{A_p\}_{p=0}^{+\infty}$，其中 $A_p = \left(a_{ij}^{(p)}\right)_{n\times n}$，若每个固定位置的项 $\left\{a_{ij}^{(p)}\right\}$ 在 $p \to +\infty$ 时都收敛（$1 \le i,j \le n$），则称矩阵列 $\{A_p\}_{p=0}^{+\infty}$ 收敛于 $\left(\lim_{p\to+\infty} a_{ij}^{(p)}\right)_{n\times n} = A$，记为 $\lim_{p\to+\infty} A_p = A$。

一个复（幂）级数，若其部分和（函数）序列收敛，则称该（幂）级数收敛，其和即为收敛的极限（函数）。

定理

设 $f(z)=\sum_{i=0}^{\infty}a_iz^i$ 是一个复幂级数，则

1. 如果 A 是一个方阵，$f(A)$ 收敛，则对任一个同阶可逆方阵 P，必有 $f\left(P^{-1}AP\right)$ 也收敛，且 $f\left(P^{-1}AP\right)=P^{-1}f(A)P$；
2. 若 $A=\operatorname{diag}(A_1,\ldots,A_k)$，$f(A)$ 收敛当且仅当 $f(A_i)$ 都收敛（$1\le i\le k$），此时有 $f(A)=\operatorname{diag}(f(A_1),\ldots,f(A_k))$；
3. 如果 $A=J_n(\lambda)$，且 $f(z)$ 的收敛半径为 R，则当 $|\lambda|<R$ 时，$f(A)$ 收敛

且 $$f(A)\begin{pmatrix} f(\lambda) & f'(\lambda) & \dfrac{f^{(2)}(\lambda)}{2!} & \cdots & \dfrac{f^{(n-1)}(\lambda)}{(n-1)!} \\ & f(\lambda) & f'(\lambda) & \cdots & \dfrac{f^{(n-2)}(\lambda)}{(n-2)!} \\ & & \ddots & \ddots & \vdots \\ & & & \ddots & f'(\lambda) \\ & & & & f(\lambda) \end{pmatrix}$$

下面使用矩阵的 Jordan 标准型及上面的定理得到一般矩阵的幂级数收敛性。

定理

设 $f(z)$ 是收敛半径为 R 的复幂级数，A 为 n 阶复方阵，其特征值为 $\lambda_1,\ldots,\lambda_n$。则

1. 当 $\max_{1\le i\le n}|\lambda_i|<R$ 时，$f(A)$ 收敛；
2. 当 $\max_{1\le i\le n}|\lambda_i|>R$ 时，$f(A)$ 发散；
3. 当 $\max_{1\le i\le n}|\lambda_i|=R$ 时，$f(A)$ 收敛当且仅当对 $|\lambda_i|=R$ 的每个特征值 λ_i，级数 $f(\lambda_i),f'(\lambda_i),,\ldots,f^{(n_i-1)}(\lambda_i)$ 都收敛，这里 n_i 为 A 的初等因子中，$(\lambda-\lambda_i)$ 出现的最高幂次；
4. 当 $f(A)$ 收敛时，$f(A)$ 的特征值恰为

$$f(\lambda_1),f(\lambda_2),\ldots,f(\lambda_n)\text{。}$$

下面是几个常见的幂级数：
$e^z=\sum_{i=0}^{\infty}\frac{z^i}{i!}$，$\sin z=\sum_{i=0}^{\infty}\frac{(-1)^i z^{2i+1}}{(2i+1)!}$，$\cos z=\sum_{i=0}^{\infty}\frac{(-1)^i z^{2i}}{(2i)!}$，$\ln(1+z)=\sum_{i=1}^{\infty}\frac{(-1)^{i+1} z^i}{i}$，等，对方阵，可求相应的幂级数的和。
当 A 为幂零阵时，上述级数和均为有限和，且相应的特征值分别为 1、0、1、0。且得到的矩阵分别与 $I+A$、A、$I+A$、A 相似。
可以证明（练习）：当 A 为幂零矩阵时，e^A 与 $I+A$ 相似。