GZH007Y 40/2

高等工程数学

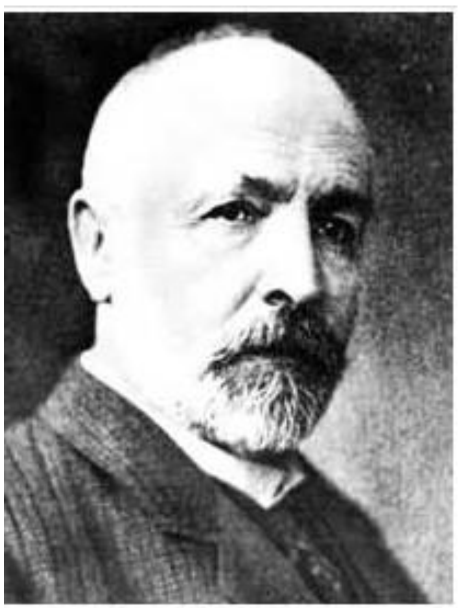
Advanced Engineering Mathematics

王 泳

中国科学院大学

2017.09.23

在数学的领域中，提出问题的艺术比解答问题的艺术更为重要



—格奥尔格·康托尔 (Canter 1845-1918)

德国数学家，集合论和超穷数理论的创始人

矩阵理论

矩阵理论在自然科学、工程技术、控制理论和社会经济学等领域的应用日趋深广，应用矩阵的理论和方法来解决工程技术和社会经济领域中的实际问题也越来越普遍。

线性空间与线性变换

- 线性空间
- 基变换与坐标变换
- 子空间与维数定理
- 线性空间的同构
- 线性变换的概念
- 线性变换的矩阵表示

内积空间

- 欧氏空间
- 正交基及子空间的正交关系
- 内积空间的同构
- 正交变换
- 复内积空间（酉空间）
- 正规矩阵

矩阵的标准形

- 矩阵的相似对角形
- 矩阵的约当标准形
- 最小多项式

矩阵函数及其应用

- 向量范数
- 矩阵范数
- 向量和矩阵的极限
- 矩阵幂级数
- 矩阵函数
- 矩阵的微分与积分
- 常用矩阵函数的性质

第三章：矩阵的标准形

1 矩阵的相似对角形

2 矩阵的约当标准形

3 最小多项式

矩阵的相似对角形

1. 矩阵的相似对角形
2. 矩阵的约当标准形
3. 最小多项式

设 $e_1, e_2, \cdots, e_n$ 和 $e_1', e_2', \cdots, e_n'$ 是数域 P 上 n 维线性空间 V 的两组基，从前一组基到后一组基的过渡矩阵是 C 。又设 T 是 V 的一个线性变换，它在前后两组基下的矩阵分别是 A 与 B，则有 $B = C^{-1}AC$

根据定理 1.13 可以得到以下结论：

设 V 是复数域 C 上 n 维线性空间，T 是 V 的一个线性变换，又 $e_1, e_2, \cdots, e_n$ 和 $\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n$ 是 V 的两组基，从第一组基到第二组基的过渡矩阵是 P 。则线性变换 T 在这两组基下的矩阵 A 与 B 相似，即

$$B = P^{-1}AP$$

这自然会引出一个问题：矩阵 A 是否可以相似于一个对角形矩阵？换言之，是否可以适当地选择第二组基 $\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n$，使得线性变换 T 在这组基下的矩阵 B 是个对角形矩阵呢？本节就是要研究这个问题，即什么是一个 n 阶矩阵能够相似于对角形矩阵的重要条件。

■ 定理3.1

设 $A\in \mathbf{C}^{n\times n}$，则 A 能与对角形矩阵相似的充要条件是 A 有 n 个线性无关的特征向量

证明：

（必要性）设矩阵 A 能相似于一个对角形矩阵，即设

$$P^{-1}AP=\begin{bmatrix}\lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n\end{bmatrix} \tag{3.1}$$

因而有

$$AP=P\begin{bmatrix}\lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n\end{bmatrix} \tag{3.2}$$

$$AP = P\begin{bmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{bmatrix} \qquad (3.2)$$

若把 P 写成分块矩阵的形式

$$P = (X_1, X_2, \cdots, X_n)$$

这里 $X_1, X_2, \cdots, X_n$ 代表 P 的 n 个列向量，应用分块矩阵的乘法规则，得到

$$AP = (AX_1, AX_2, \cdots, AX_n)$$

$$P\begin{bmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{bmatrix} = (\lambda_1 X_1, \lambda_2 X_2, \cdots, \lambda_n X_n)$$

故由（3.2）式可得

$$AX_i = \lambda_i X_i \quad (i = 1, 2, \cdots, n) \qquad (3.3)$$

或

$$(\lambda_i E - A) X_i = \theta \quad (i = 1, 2, \cdots, n)$$

这说明，若 A 能相似于一个对角形矩阵（见（3. 1）式），则可逆矩阵 $P=(X_1,X_2,\cdots,X_n)$ 的每个列向量（非零向量）X_i 都满足（3. 3）式。

（若 A 为 n 阶矩阵，X 是 n 维列向量，则满足方程 $AX=\lambda X$ 的非零向量 X 被称为矩阵 A 的对应于（或属于）特征值 λ 的特征向量）

所以，根据矩阵的特征值与特征向量的定义可知：

当 A 与对角形矩阵相似时，这对角形矩阵主对角线上的元素 λ_i（$i=1,2,\cdots,n$）都是 A 的特征值，且矩阵 P 的 n 个列向量 $X_1,X_2,\cdots,X_n$ 是 A 的 n 个线性无关的特征向量（因 P 可逆，则 $|P|\neq 0$，故 P 的 n 个列向量线性无关）。

（充分性）若 A 有 n 个线性无关的特征向量 $X_1, X_2, \cdots, X_n$，并且它们相应的特征值分别为 $\lambda_1, \lambda_2, \cdots, \lambda_n$，即

$$AX_i = \lambda_i X_i \quad (i = 1, 2, \cdots, n)$$

则取 $P = (X_1, X_2, \cdots, X_n)$。显然，$P$ 为满秩矩阵。

又因为 $AP = (AX_1, AX_2, \cdots, AX_n) = (\lambda_1 X_1, \lambda_2 X_2, \cdots, \lambda_n X_n)$

$$= (X_1, X_2, \cdots, X_n)\begin{bmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{bmatrix} = P\begin{bmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{bmatrix}$$

可得

$$P^{-1}AP=\begin{bmatrix}\lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n\end{bmatrix}$$

即 A 与对角形矩阵相似。

命题得证

- **定理3.2**

（充分条件） 若 n 阶矩阵 A 有 n 个不同的特征值，则 A 可与对角形矩阵相似。

属于 A 的不同特征值的特征向量线性无关

■ **例3.1**

研究矩阵

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

能否与对角形矩阵相似。

现在是否可以判断A不相似于对角形矩阵呢?

解： A 的特征多项式为 $|\lambda E-A|=(\lambda-1)^2(\lambda+2)$，因此其特征值为 $\lambda_1=\lambda_2=1$，$\lambda_3=-2$。

对应于特征值 1 的一切特征向量为

$$X=k\begin{pmatrix}3\\-6\\20\end{pmatrix},\quad k\neq 0$$

又对应于特征值−2 的一切特征向量为

$$X=k_1\begin{pmatrix}0\\0\\1\end{pmatrix},\quad k_1\neq 0$$

所以，不存在三个线性无关的特征向量，从而 A 不能与对角形矩阵相似。

■ **例3.2**

设

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

求 A 的相似对角形及 A^{100}。

解： 由 $|\lambda E-A|=(\lambda-1)^2(\lambda+2)$，得 $\lambda_1=-2$，$\lambda_2=1$（二重根），求得对应于 λ_1 的一个特征向量 $X_1=(-1,1,1)^T$，及对应于二重根 λ_2 的两个线性无关特征向量为

$$X_2=(-2,1,0)^T,\quad X_3=(0,0,1)^T$$

因此可得

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

因而有

$$P^{-1}AP=\begin{bmatrix}-2 & & \\ & 1 & \\ & & 1\end{bmatrix} \tag{3.5}$$

等式右边就是 A 的相似对角形。

由（3.5）式得

$$A=P\begin{bmatrix}-2 & & \\ & 1 & \\ & & 1\end{bmatrix}P^{-1}$$

因此易知

$$A^{100}=P\begin{bmatrix}-2 & & \\ & 1 & \\ & & 1\end{bmatrix}^{100}P^{-1}=\begin{bmatrix}-1 & -2 & 0\\ 1 & 1 & 0\\ 1 & 0 & 1\end{bmatrix}\begin{bmatrix}2^{100} & & \\ & 1 & \\ & & 1\end{bmatrix}\begin{bmatrix}1 & 2 & 0\\ -1 & -1 & 0\\ -1 & -2 & 1\end{bmatrix}$$

$$=\begin{bmatrix}-2^{100}+2 & -2^{101}+2 & 0\\ 2^{100}-1 & 2^{101}-1 & \\ 2^{100}-1 & 2^{101}-2 & 1\end{bmatrix}$$

注意:

$\lambda_1=-2$，$\lambda_2=1$（二重根）

$$P=\begin{bmatrix}-1&-2&0\\1&1&0\\1&0&1\end{bmatrix}$$

上题中若取

$$P_1=\begin{bmatrix}-1&0&-2\\1&0&1\\1&1&0\end{bmatrix}$$

则亦有

$$P_1^{-1}AP_1=\begin{bmatrix}-2&&\\&1&\\&&1\end{bmatrix}$$

因此可见 P 不是唯一的。

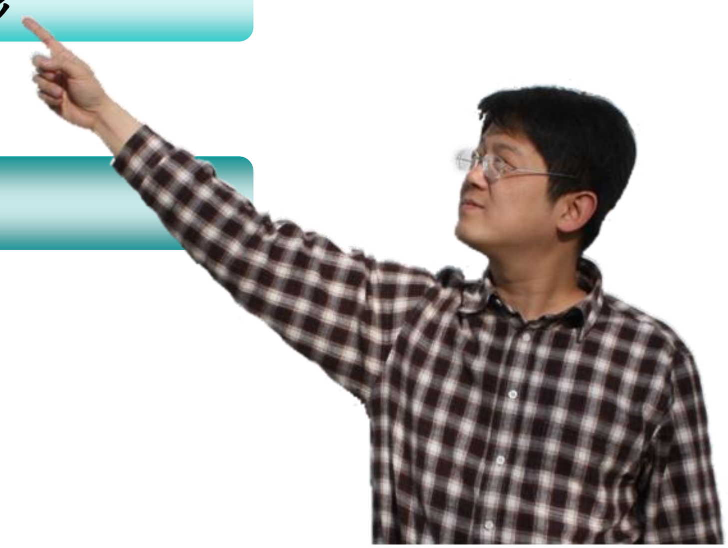
思考题 （15分）

下列矩阵能否与对角形矩阵相似？若 A 能与对角形矩阵相似，则求出可逆矩阵 P，使得 $P^{-1}AP$ 为对角形矩阵：

（1） $A=\begin{bmatrix} 3 & 4 \\ 5 & 2 \end{bmatrix}$；　　（2） $A=\begin{bmatrix} 5 & -3 & 2 \\ 6 & -4 & 4 \\ 4 & -4 & 5 \end{bmatrix}$

矩阵的约当标准形

1. 矩阵的相似对角形
2. 矩阵的约当标准形
3. 最小多项式



由上节例 3.1 可以看出，并非每个矩阵都可以相似于对角形矩阵。当矩阵 $A=\left(a_{ij}\right)\in \mathbf{C}^{n\times n}$ 不能和对角形矩阵相似时，能否找到一个构造比较简单的分块对角形矩阵和它相似呢？当在复数域 $\mathbf{C}$ 内考虑这个问题时，这样的矩阵确实是存在的，这就是约当（Jordan）形矩阵，它被称为矩阵 A 的约当标准形。

■ 定义3.1 多项式的最大公因式

设 $f(\lambda)$， $g(\lambda)$ 是数域 P 上的多项式，如果存在 P 上的多项式 $d(\lambda)$ 满足：

（1） $d(\lambda)\mid f(\lambda)$， $d(\lambda)\mid g(\lambda)$；

$d(\lambda)\mid f(\lambda)$ 表示 $d(\lambda)$ 能整除 $f(\lambda)$

（2）若有 P 上的多项式 $d_1(\lambda)$ 满足 $d_1(\lambda)\mid f(\lambda)$， $d_1(\lambda)\mid g(\lambda)$，则有 $d_1(\lambda)\mid d(\lambda)$。

则称 $d(\lambda)$ 是 $f(\lambda)$ 和 $g(\lambda)$ 的一个最大公因式，且以 $\left(f(\lambda),g(\lambda)\right)$ 表示首项系数为 1 的最大公因式。

三个多项式 $f(\lambda)$，$g(\lambda)$，$h(\lambda)$ 的最大公因式 $(f(\lambda),g(\lambda),h(\lambda))$ 可定义为：

$$((f(\lambda),g(\lambda)),h(\lambda))$$

类似地可以定义 n 个多项式 $f_1(\lambda),\cdots,f_n(\lambda)$ 的最大公因式：

$$(f_1(\lambda),\cdots,f_n(\lambda))=((f_1(\lambda),\cdots,f_{n-1}(\lambda)),f_n(\lambda))$$

由定义容易验证，若 c 为非 0 常数，$f(\lambda)$ 是首项系数为 1 的多项式，则有：

$$(f(\lambda),c)=1$$

$$(f(\lambda),0)=f(\lambda)$$

零次多项式

零多项式

事实上，在复数域上求若干个多项式的最大公因式时，先把每个多项式分解成一次因式的方幂的乘积形式，然后取含有公共一次因式的最低方幂的乘积，即为所求的最大公因式。

■ **例**

$$f(\lambda)=(\lambda-1)^3(\lambda+2)^2(\lambda+5)$$

$$g(\lambda)=(\lambda-1)^2(\lambda+2)^3(\lambda+3)$$

$$h(\lambda)=(\lambda-1)^2(\lambda+2)(\lambda+3)^5$$

则有

$$(f(\lambda),g(\lambda),h(\lambda))=(\lambda-1)^2(\lambda+2)$$

设 $A=(a_{ij})\in \mathbf{C}^{n\times n}$，则 A 的特征矩阵 $\lambda E-A$ 可以用 $A(\lambda)$ 来记。

■ **定义3.2 特征矩阵的行列式因子**

$A(\lambda)$ 中所有非零的 k 阶子式的首项（最高次项）系数为 1 的最大公因式 $D_k(\lambda)$ 被称为 $A(\lambda)$ 的 k 级行列式因子，$k=1,2,\cdots,n$。

由定义可知：$D_n(\lambda)=|\lambda E-A|$。又因为 $D_{k-1}(\lambda)$ 能整除每个 $k-1$ 级子式，从而可以整除每个 k 级子式（将 k 级子式按一行或一列展开即可），因此 $D_{k-1}(\lambda)$ 能整除 $D_k(\lambda)$，并记为 $D_{k-1}(\lambda)\mid D_k(\lambda)$（$k=1,2,\cdots,n$）。

- **例3.3**

求下列矩阵的特征矩阵的行列式因子

（1） $A=\begin{bmatrix}-1 & & \\ & 1 & \\ & & 2\end{bmatrix}$；（2） $A=\begin{bmatrix}a & & & \\ 1 & \ddots & & \\ & \ddots & \ddots & \\ & & 1 & a\end{bmatrix}_{n\times n}$

解： （1） $\lambda E-A=\begin{bmatrix}\lambda+1 & & \\ & \lambda-1 & \\ & & \lambda-2\end{bmatrix}$

$$D_1(\lambda)=(\lambda+1,\lambda-1,\lambda-2)=1$$

$$D_2(\lambda)=((\lambda+1)(\lambda-1),(\lambda+1)(\lambda-2),(\lambda-1)(\lambda-2))=1$$

$$D_3(\lambda)=(\lambda+1)(\lambda-1)(\lambda-2)$$

求下列矩阵的特征矩阵的行列式因子

■ **例3.3**

（1）$A=\begin{bmatrix}-1 & & \\ & 1 & \\ & & 2\end{bmatrix}$；（2）$A=\begin{bmatrix}a & & & \\ 1 & \ddots & & \\ & \ddots & \ddots & \\ & & 1 & a\end{bmatrix}_{n\times n}$

解： （2）$\lambda E-A=\begin{bmatrix}\lambda-a & & & \\ -1 & \ddots & & \\ & \ddots & \ddots & \\ & & -1 & \lambda-a\end{bmatrix}_{n\times n}$

因为存在一个$n-1$级子式$=\begin{vmatrix}-1 & \lambda-a & & \\ & \ddots & \ddots & \\ & & \ddots & \lambda-a \\ & & & -1\end{vmatrix}_{(n-1)\times(n-1)}=(-1)^{n-1}$为非零常数，故

$D_{n-1}(\lambda)=1$，从而$D_1(\lambda)=\cdots=D_{n-1}(\lambda)=1$，$D_n(\lambda)=(\lambda-a)^n$。

■ 定义3.3 特征矩阵的不变因式和初级因子

下列的 n 个多项式

$$d_1(\lambda)=D_1(\lambda),\ d_2(\lambda)=\frac{D_2(\lambda)}{D_1(\lambda)},\ \cdots,\ d_k(\lambda)=\frac{D_k(\lambda)}{D_{k-1}(\lambda)},\ \cdots,\ d_n(\lambda)=\frac{D_n(\lambda)}{D_{n-1}(\lambda)}$$ 被称

为 $\boldsymbol{A}(\lambda)$ 的不变因式。

把每个次数大于零的不变因式分解为互不相同的一次因式的方幂的乘积(在复数域内这样的分解是可能的)，所有这些一次因式的方幂（相同的必须按出现次数计算)，被称为 $\boldsymbol{A}(\lambda)$ 的初级因子。

因为这里的 $\boldsymbol{A}(\lambda)=\lambda\boldsymbol{E}-\boldsymbol{A}$ 完全由 $\boldsymbol{A}$ 决定，所以这里 $\boldsymbol{A}(\lambda)$ 的不变因式及初级因子，也常被称为矩阵 $\boldsymbol{A}$ 的不变因式及初级因子。

■ **例3.4** 求下列矩阵的不变因式及初级因子

（1） $A=\begin{bmatrix}-1 & & & \\ & -2 & & \\ & & 1 & \\ & & & 2\end{bmatrix}$；（2） $A=\begin{bmatrix}1 & 2 & 0 \\ 0 & 2 & 0 \\ -2 & -2 & -1\end{bmatrix}$

解：（1）因为 A 的特征矩阵为

$$A(\lambda)=\lambda E-A=\begin{bmatrix}\lambda+1 & & & \\ & \lambda+2 & & \\ & & \lambda-1 & \\ & & & \lambda-2\end{bmatrix}$$

所以 $A(\lambda)$ 的行列式因子为

$$D_4(\lambda)=|\lambda E-A|=(\lambda+1)(\lambda-1)(\lambda+2)(\lambda-2)$$

$$D_3(\lambda)=D_2(\lambda)=D_1(\lambda)=1$$

不变因式为

$$d_1(\lambda)=D_1(\lambda)=1，\quad d_2(\lambda)=d_3(\lambda)=1$$

$$d_4(\lambda)=D_4(\lambda)/D_3(\lambda)=(\lambda+1)(\lambda-1)(\lambda+2)(\lambda-2)$$

而次数大于零的不变因式只有$d_4(\lambda)$，故由定义可见 A 的全部初级因子为

$$\lambda+1，\quad \lambda-1，\quad \lambda+2，\quad \lambda-2$$

■ **例3.4** 求下列矩阵的不变因式及初级因子

（1） $A=\begin{bmatrix}-1 & & & \\ & -2 & & \\ & & 1 & \\ & & & 2\end{bmatrix}$；（2） $A=\begin{bmatrix}1 & 2 & 0\\ 0 & 2 & 0\\ -2 & -2 & -1\end{bmatrix}$

解：（2）因为

$$A(\lambda)=\lambda E-A=\begin{bmatrix}\lambda-1 & -2 & 0\\ 0 & \lambda-2 & 0\\ 2 & 2 & \lambda+1\end{bmatrix}$$

所以 $D_3(\lambda)=|\lambda E-A|=(\lambda-1)(\lambda+1)(\lambda-2)$。又易得 $D_2(\lambda)=1$，从而 $D_1(\lambda)=1$。

于是，不变因式为：

$$d_1(\lambda)=d_2(\lambda)=1，\ d_3(\lambda)=(\lambda-1)(\lambda+1)(\lambda-2)$$

而初级因子为： $\lambda-1，\ \lambda+1，\ \lambda-2$

■ **例3.5**

设

$$A=\begin{bmatrix} a & -b_1 & & \\ & a & \ddots & \\ & & \ddots & -b_{n-1} \\ & & & a \end{bmatrix}_{n\times n}$$

，（各个 $b_i \neq 0$ ）

求 A 的初级因子。

解：

因为

$$A(\lambda)=\lambda E-A=\begin{bmatrix} \lambda-a & b_1 & & \\ & \lambda-a & \ddots & \\ & & \ddots & b_{n-1} \\ & & & \lambda-a \end{bmatrix}_{n\times n}$$

所以 $D_n(\lambda)=|\lambda E-A|=(\lambda-a)^n$。

$$A(\lambda)=\lambda E-A=\begin{bmatrix} \lambda-a & b_1 & & \\ & \lambda-a & \ddots & \\ & & \ddots & b_{n-1} \\ & & & \lambda-a \end{bmatrix}_{n\times n}$$

又因为它有一个 $n-1$ 级子式

$$\begin{bmatrix} b_1 & & & \\ \lambda-a & b_2 & & \\ & \ddots & \ddots & \\ & & \lambda-a & b_{n-1} \end{bmatrix}_{(n-1)\times(n-1)} = b_1 b_2 \cdots b_{n-1} \neq 0$$

故 $D_{n-1}(\lambda)=1$，从而 $D_{n-2}(\lambda)=\cdots=D_1(\lambda)=1$，于是，不变因式为

$$d_1(\lambda)=d_2(\lambda)=\cdots=d_{n-1}(\lambda)=1,\quad d_n(\lambda)=(\lambda-a)^n$$

因此，初级因子只有一个：$(\lambda-a)^n$。

■ **定义： 约当标准形**

设矩阵 $A=(a_{ij})\in C^{n\times n}$ 的全部初级因子为： $(\lambda-\lambda_1)^{k_1}$，$(\lambda-\lambda_2)^{k_2}$，…，$(\lambda-\lambda_s)^{k_s}$

其中 λ_i（$i=1,2,\cdots,s$）可能有相同的，指数 k_i（$i=1,2,\cdots,s$）也可能有相同的。对每个初级因子 $(\lambda-\lambda_i)^{k_i}$ 构造一个 k_i 阶矩阵（约当块）：

$$J_i=\begin{bmatrix}\lambda_i & & & \\ 1 & \lambda_i & & \\ & \ddots & \ddots & \\ & & 1 & \lambda_i\end{bmatrix}_{k_i\times k_i}，（i=1,2,\cdots,s）$$

由所有这些约当块构成的分块对角矩阵

$$J=\begin{bmatrix}J_1 & & & \\ & J_2 & & \\ & & \ddots & \\ & & & J_s\end{bmatrix}_{n\times n}$$

被称为矩阵 A 的约当形矩阵，或 A 的约当标准形。

定理3.3

每个 n 阶复数矩阵 A 都与一个约当形矩阵 J 相似：

$$P^{-1}AP = J$$

除去约当块的排列次序外，约当形矩阵 J 是被矩阵 A 唯一决定的。

推论

复数矩阵 A 与对角形矩阵相似的充要条件是 A 的初级因子全为一次式。

■ 例

例 3.4 中（1）的矩阵 A 的约当标准形是

$$A=\begin{bmatrix}-1 & & & \\ & -2 & & \\ & & 1 & \\ & & & 2\end{bmatrix}$$

$\lambda+1$，$\lambda-1$，$\lambda+2$，$\lambda-2$

$$J=\begin{bmatrix}-1 & & & \\ & 1 & & \\ & & -2 & \\ & & & 2\end{bmatrix}$$

例 3.4 中（2）的矩阵 A 的约当标准形是

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

$\lambda-1$，$\lambda+1$，$\lambda-2$

$$J=\begin{bmatrix}1 & & \\ & -1 & \\ & & 2\end{bmatrix}$$

矩阵的约当标准形

例 3.3 中（2）的矩阵 A 的约当标准形是

A 的初级因子不全为一次式

$$A=\begin{bmatrix} a & & & \\ 1 & \ddots & & \\ & \ddots & \ddots & \\ & & 1 & a \end{bmatrix}_{n\times n} \xrightarrow{(\lambda-a)^n} J=\begin{bmatrix} a & & & \\ 1 & a & & \\ & \ddots & \ddots & \\ & & 1 & a \end{bmatrix}_{n\times n}$$

虽然由
$$|\lambda E-A|=|\lambda E-J|=|\lambda E_1-J_1|\bullet|\lambda E_2-J_2|\cdots|\lambda E_s-J_s|$$
$$=(\lambda-\lambda_1)^{k_1}(\lambda-\lambda_2)^{k_2}\cdots(\lambda-\lambda_s)^{k_s}$$

可知，约当形矩阵 J 的主对角线上的元素 λ_i（$i=1,2,\cdots,s$）全为 A 的特征值，并且 $\sum_{i=1}^{s}k_i=n$，但由于 $i\neq j$ 时可能有 $\lambda_i=\lambda_j$，所以 λ_i 不一定就是 A 的 k_i 重特征根，故一般由矩阵的特征值多项式是不能写出矩阵的约当形矩阵的。

■ 例3.6

求线性微分方程组

$$\begin{cases} \dfrac{\mathrm{d}x_1}{\mathrm{d}t} = -x_1 + x_2 \\ \dfrac{\mathrm{d}x_2}{\mathrm{d}t} = -4x_1 + 3x_2 \\ \dfrac{\mathrm{d}x_3}{\mathrm{d}t} = x_1 + 2x_3 \end{cases}$$

的通解。这里 x_1， x_2， x_3 都是 t 的未知函数。

解： 将这组方程写成矩阵形式为

$$\frac{\mathrm{d}X}{\mathrm{d}t} = AX$$

其中

$$A = \begin{bmatrix} -1 & 1 & 0 \\ -4 & 3 & 0 \\ 1 & 0 & 2 \end{bmatrix},\ X = (x_1, x_2, x_3)^T$$

矩阵的约当标准形

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

不难求得矩阵 A 的初级因子为

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

于是存在可逆矩阵 $P=(X_1,X_2,X_3)$，使得

$$P^{-1}AP=J=\begin{bmatrix}2&&\\&1&\\&1&1\end{bmatrix}$$

所以 $AP=PJ$

又因为 $$AP=(AX_1,AX_2,AX_3)$$

$$PJ=(X_1,X_2,X_3)\begin{bmatrix}2&&\\&1&\\&1&1\end{bmatrix}=(2X_1,X_2+X_3,X_3)$$

矩阵的约当标准形

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

所以

$$\begin{cases}AX_1=2X_1\\ AX_2=X_2+X_3\\ AX_3=X_3\end{cases}$$

求得

$$X_1=(0,0,1)^T,\quad X_2=(0,1,-1)^T,\quad X_3=(1,2,-1)^T$$

所以有

$$P=\begin{bmatrix}0 & 0 & 1\\ 0 & 1 & 2\\ 1 & -1 & -1\end{bmatrix}$$

作满秩线性变换 $X = PY$，其中 $Y = (y_1, y_2, y_3)^T$

则 $\dfrac{\mathrm{d}X}{\mathrm{d}t} = AX$ 可以转化为

$$\frac{\mathrm{d}(PY)}{\mathrm{d}t} = APY$$

$$\Rightarrow P\frac{\mathrm{d}Y}{\mathrm{d}t} = APY$$

$$\Rightarrow \frac{\mathrm{d}Y}{\mathrm{d}t} = P^{-1}APY = JY$$

$$\Rightarrow \frac{\mathrm{d}Y}{\mathrm{d}t} = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix} Y$$

即

$$\begin{cases} \dfrac{\mathrm{d}y_1}{\mathrm{d}t} = 2y_1 \\ \dfrac{\mathrm{d}y_2}{\mathrm{d}t} = y_2 \\ \dfrac{\mathrm{d}y_3}{\mathrm{d}t} = y_2 + y_3 \end{cases}$$

$$\begin{cases} \dfrac{\mathrm{d}x_1}{\mathrm{d}t} = -x_1 + x_2 \\ \dfrac{\mathrm{d}x_2}{\mathrm{d}t} = -4x_1 + 3x_2 \\ \dfrac{\mathrm{d}x_3}{\mathrm{d}t} = x_1 + 2x_3 \end{cases}$$

对方程进行积分可得

$$\boxed{y_1 = k_1 e^{2t}, \quad y_2 = k_2 e^{t}, \quad y_3 = (k_2 t + k_3) e^{t}}$$

故微分方程组的通解为

$$X = PY = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 2 \\ 1 & -1 & -1 \end{bmatrix} \begin{bmatrix} k_1 e^{2t} \\ k_2 e^{t} \\ (k_2 t + k_3) e^{t} \end{bmatrix}$$

$$= \begin{bmatrix} (k_2 t + k_3) e^{t} \\ (2k_2 t + k_2 + 2k_3) e^{t} \\ k_1 e^{2t} - (k_2 t + k_2 + k_3) e^{t} \end{bmatrix}$$

这里 k_1， k_2， k_3 是任意常数。

思考题

（15分）

思考题 （15分）

在复数域上，求下列矩阵的约当标准形：

$$（1）\ A=\begin{bmatrix} 3 & 0 & 8 \\ 3 & -1 & 6 \\ -2 & 0 & -5 \end{bmatrix};\qquad （2）\ A=\begin{bmatrix} 4 & 5 & -2 \\ -2 & -2 & 1 \\ -1 & -1 & 1 \end{bmatrix}$$

最小多项式

1 矩阵的相似对角形

2 矩阵的约当标准形

3 最小多项式

■ 定理3.4

哈密顿一凯莱（Hamilton-Cayley）定理：每个 n 阶矩阵 A 都是它的特征多项式的根，即

$$A^n + a_1 A^{n-1} + \cdots + a_{n-1} A + a_n E = 0 \text{（零矩阵）} \qquad (3.6)$$

因为

$$f(\lambda) = |\lambda E - A| = \lambda^n + a_1 \lambda^{n-1} + \cdots + a_{n-1}\lambda + a_n$$

所以（3.6）式常写成 $f(A) = 0$（零矩阵）。

■ **例3.8** 设 $A=\begin{bmatrix}1&0&2\\0&-1&1\\0&1&0\end{bmatrix}$，试计算 $\varphi(A)=2A^8-3A^5+A^4+A^2-4E$。

解： 因为 A 的特征多项式为

$$f(\lambda)=|\lambda E-A|=\lambda^3-2\lambda+1$$

再取多项式

$$\varphi(\lambda)=2\lambda^8-3\lambda^5+\lambda^4+\lambda^2-4$$

以 $f(\lambda)$ 去除 $\varphi(\lambda)$ 可得

$$\varphi(\lambda)=(2\lambda^5+4\lambda^3-5\lambda^2+9\lambda-14)f(\lambda)+r(\lambda)$$

这里余式 $r(\lambda)=24\lambda^2-37\lambda+10$。

由哈密顿—开莱定理，$f(A)=0$（零矩阵），所以

$$\varphi(A)=r(A)=24A^2-37A+10E$$

$$=\begin{bmatrix}-3&48&-26\\0&95&-61\\0&-61&34\end{bmatrix}$$

这个例子说明，一个n阶矩阵的多项式，如其次数高于n次，则应用定理3.4可将它转化为次数小于n的多项式来计算。

■ **定义：零化多项式**

若$\varphi(\lambda)$是一个多项式，$\boldsymbol{A}$是一个方阵，如果有$\varphi(\boldsymbol{A})=\mathbf{0}$（零矩阵），则称$\varphi(\lambda)$是矩阵$\boldsymbol{A}$的零化多项式。

注意：

每个方阵都有零化多项式，因为它的特征多项式就是一个，但其零化多项式并不唯一。

■ **定义3.4 最小多项式**

设$\boldsymbol{A}$是n阶矩阵，则$\boldsymbol{A}$的首项系数为1的次数最小的零化多项式$m(\lambda)$被称为$\boldsymbol{A}$的最小多项式。

首项（最高次项）系数为1的多项式，以后简称为首一多项式。

■ **定理3.5** 矩阵 A 的任何零化多项式都被其最小多项式所整除

证明：（反证法）

$$\boxed{\text{整除} \Leftrightarrow r(\lambda)=0}$$

设 $\varphi(\lambda)$ 是 A 的任一零化多项式，又 $m(\lambda)$ 是 A 的最小多项式，以 $m(\lambda)$ 除 $\varphi(\lambda)$

即得

$$\varphi(\lambda)=q(\lambda)m(\lambda)+r(\lambda)$$

这里 $r(\lambda)$ 如不为零时则其次数小于 $m(\lambda)$ 的次数。

于是有 $$\varphi(A)=q(A)m(A)+r(A)$$

因为 $\varphi(A)=m(A)=0$（零矩阵），所以有 $r(A)=0$（零矩阵），即 $r(\lambda)$ 也是

A 的零化多项式。

如果 $r(\lambda)\neq 0$，则 $r(\lambda)$ 的次数小于 $m(\lambda)$ 的次数，这与 $m(\lambda)$ 是最小多项式矛盾。

所以，只能 $r(\lambda)\equiv 0$，故 $m(\lambda)$ 被 $\varphi(\lambda)$ 整除（$m(\lambda)\mid\varphi(\lambda)$）。

命题得证

■ **定理3.6** 矩阵 A 的最小多项式是唯一的

证明:

（反证法）若 $m(\lambda)$ 与 $n(\lambda)$ 均为 A 的最小多项式，那么每一个都可以被另一个所整除，因此两者只有常数因子的差别。又因为两者都是首一多项式，所以这常数因子必定等于 1，故 $m(\lambda)=n(\lambda)$。

命题得证

■ **定理3.7**

矩阵 A 的最小多项式的根必定是 A 的特征根；反之，A 的特征根也必定是 A 的最小多项式的根。

证明：

因为 A 的特征多项式 $f(\lambda)=|\lambda E-A|$ 是 A 的零化多项式，故由定理 3.5 可知，$f(\lambda)$ 可被 A 的最小多项式 $m(\lambda)$ 所整除，即 $m(\lambda)$ 是 $f(\lambda)$ 的因式，所以 $m(\lambda)$ 的根都是 $f(\lambda)$ 的根。

反之，若 λ_0 是 A 的一个特征根，且

$$AX=\lambda_0 X,（X\neq\boldsymbol{\theta}）$$

又设 A 的最小多项式

$$m(\lambda)=\lambda^k+\beta_1\lambda^{k-1}+\cdots+\beta_{k-1}\lambda+\beta_k$$

则

$$\begin{aligned}m(A)X&=A^kX+\beta_1A^{k-1}X+\cdots+\beta_{k-1}AX+\beta_kX\\&=\lambda_0^{\,k}X+\beta_1\lambda_0^{\,k-1}X+\cdots+\beta_{k-1}\lambda_0X+\beta_kX\\&=m(\lambda_0)X\end{aligned}$$

命题得证

由于 $m(A)=\mathbf{0}$（零矩阵），又 $X\neq\boldsymbol{\theta}$，所以 $m(\lambda_0)=0$，亦即 λ_0 是 $m(\lambda)$ 的根。

例3.9 求矩阵 $A=\begin{bmatrix}3 & -3 & 2\\ -1 & 5 & -2\\ -1 & 3 & 0\end{bmatrix}$ 的最小多项式 $m(\lambda)$

解：

A 的特征多项式为 $$f(\lambda)=|\lambda E-A|=(\lambda-2)^2(\lambda-4)$$

故根据定理 3. 5～定理 3. 7 可知 A 的最小多项式只能是

$$m(\lambda)=(\lambda-2)(\lambda-4) \text{ 或 } m(\lambda)=f(\lambda)$$

但由 $m(A)=(A-2E)(A-4E)=0$ （零矩阵），可知 A 的最小多项式不是 $f(\lambda)$，而应是 $m(\lambda)=(\lambda-2)(\lambda-4)$。

最小多项式

■ 定理3.8

设 $\boldsymbol{A}$ 是 n 阶矩阵，$D_{n-1}(\lambda)$ 是特征矩阵 $\lambda\boldsymbol{E}-\boldsymbol{A}$ 的 $n-1$ 阶行列式因子，则 $\boldsymbol{A}$ 的最小多项式为

$$m(\lambda)=\frac{|\lambda\boldsymbol{E}-\boldsymbol{A}|}{D_{n-1}(\lambda)}=\frac{D_n(\lambda)}{D_{n-1}(\lambda)}=d_n(\lambda)$$

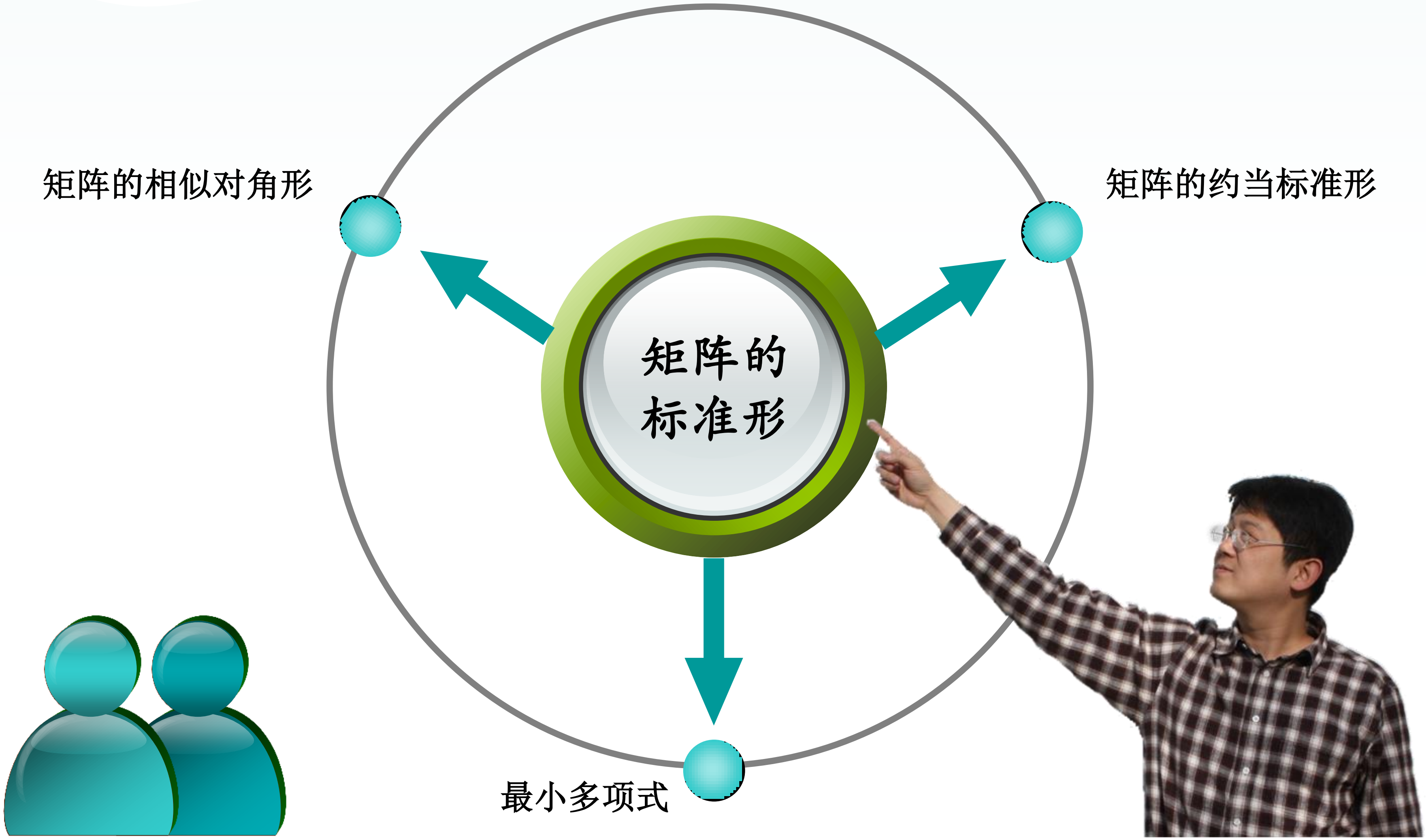
这里的 $d_n(\lambda)$ 就是 $\lambda\boldsymbol{E}-\boldsymbol{A}$ 的第 n 个不变因式。

思考题 （**15分**） 设

$$\boldsymbol{A}=\begin{bmatrix}1&0&0\\1&0&1\\0&1&0\end{bmatrix}$$

证明：当 $n\geq 3$ 时，$\boldsymbol{A}^n=\boldsymbol{A}^{n-2}+\boldsymbol{A}^2-\boldsymbol{E}$，并求 $\boldsymbol{A}^{100}$

总结



人民语录

政治上常常是少数服从多数，科学上常常是多数服从少数，因为政治服从的是利益，科学服从的是真理

To be continued…

wangyong@ucas.ac.cn
http://people.ucas.ac.cn/~wangyong

Thank you !