同济大学

高等代数与解析几何

教师： 蒋志洪

理学院数学系

§7.4 可对角化条件

定理. 设V是n维线性空间,T是V上的线性变换. 则T 可以对角化的充分必要条件T 有n 个线性无关的特征向量.

证明. (必要性) 如果T可以对角化，则V 中有一个有序基$(\beta_1, \beta_2, \cdots, \beta_n)$ 使得T在这个有序基下的矩阵是对角阵, 即,

$$T(\beta_1, \beta_2, \cdots, \beta_n) = (\beta_1, \beta_2, \cdots, \beta_n) \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix}$$

于是$T(\beta_i) = \lambda_i \beta_i, i = 1, 2, \cdots, n$. 由于基向量$\beta_i$ 总是非零向量, 这说明λ_i是T的特征值, β_i是T 的属于特征值λ_i 的特征向量, 所以T有n 个线性无关的特征向量.

(充分性) 如果T 有n 个线性无关的特征向量$\alpha_1, \alpha_2, \cdots, \alpha_n$, 则$(\alpha_1, \alpha_2, \cdots, \alpha_n)$可以构成$V$的一个有序基, 由于$\alpha_1, \alpha_2, \cdots, \alpha_n$是特征向量, 所以存在$\lambda_i, (i = 1, 2, \cdots, n)$使得$T(\alpha_i) = \lambda_i \alpha_i$, 由此推出$T$ 在有序基$(\alpha_1, \alpha_2, \cdots, \alpha_n)$下的矩阵为

$$\begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix}$$

说明. V上的一个线性变换T可以对角化, 则V中存在一个有序基使得T在这个有序基下的矩阵是对角阵.从这个定理的证明可以看出: 这个有序基由T的特征向量组成, 这个对角阵的对角线上的元素由T的特征值组成, 而且对应的位置一致.

这个定理的矩阵形式

定理. n阶方阵A可以相似于对角阵充分必要条件A 有n 个线性无关的特征向量. 而且进一步如果$X_1, X_2, \cdots, X_n$是A的n 个线性无关的特征向量使得$AX_i = \lambda_i X_i$, $i = 1, 2, \cdots, n$, 则令P 是以$X_1, X_2, \cdots, X_n$ 为列向量的n 阶方阵, 即, $P = (X_1, X_2, \cdots, X_n)$. 则有

$$P^{-1}AP = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix},$$

$\lambda_1, \lambda_2, \cdots, \lambda_n$ 是A 的全部特征值(重根按重数计).

例. 设V是F上的三维线性空间, $(\alpha_1,\alpha_2,\alpha_3)$是$V$的一个有序基, T是V上的线性空间, 它在有序基$(\alpha_1,\alpha_2,\alpha_3)$下的矩阵为:

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

我们已经知道: T有特征值为0和3(二重根). $\beta_1=\alpha_1+\alpha_2+\alpha_3$是属于特征值0的特征向量, $\beta_2=-\alpha_1+\alpha_2, \beta_3=-\alpha_1+\alpha_3$是属于特征值3的特征向量, 由于$\beta_1,\beta_2,\beta_3$线性无关, 所以线性变换$T$可以对角化, 而且$T$在有序基$(\beta_1,\beta_2,\beta_3)$下的矩阵是

$$\begin{pmatrix}0&0&0\\0&3&0\\0&0&3\end{pmatrix},\quad \text{或者矩阵形式:}T(\beta_1,\beta_2,\beta_3)=(\beta_1,\beta_2,\beta_3)\begin{pmatrix}0&0&0\\0&3&0\\0&0&3\end{pmatrix}$$

用矩阵的语言: 由于A有线性无关的特征向量$\eta_1=\begin{pmatrix}1\\1\\1\end{pmatrix}$(属于特征值0), $\eta_2=\begin{pmatrix}-1\\1\\0\end{pmatrix}, \eta_2=\begin{pmatrix}-1\\0\\1\end{pmatrix}$(属于特征值3), 所以$A$相似对角阵, 而且

$$P^{-1}AP=\begin{pmatrix}0&0&0\\0&3&0\\0&0&3\end{pmatrix},\quad \text{其中}P=\begin{pmatrix}1&-1&-1\\1&1&0\\1&0&1\end{pmatrix}$$

定义. 设V 是数域$\mathbb{F}$ 上n 维线性空间, T 是V上的线性变换, $\Delta_T(\lambda)$是T 的特征多项式. 如果$\Delta_T(\lambda)$ 在域$\mathbb{F}$ 上可分解成一次因式的乘积, 设$\Delta_T(\lambda)$的标准分解式为

$$\Delta_T(\lambda) = (\lambda - \lambda_1)^{c_1}(\lambda - \lambda_2)^{c_2}\cdots(\lambda - \lambda_s)^{c_s}$$

则c_i 称为特征值λ_i 的代数重数.

设V_{λ_i} 是特征值λ_i 的特征子空间, 即

$$V_{\lambda_i} = \{\alpha \in V | T(\alpha) = \lambda_i\alpha\}, \quad i = 1, 2, \cdots, s.$$

记$d_i = \dim V_{\lambda_i}$, 则称d_i为特征值λ_i 的几何重数.

说明. 由定义可以看出特征值λ_i的代数重数就是λ_i作为代数方程$\Delta_T(\lambda) = 0$的根的重数, 特征值λ_i的几何重数就是齐次线性方程组$(\lambda_i E - A)X = 0$的基础解系所含向量的个数(这里A是T在一个有序基下的矩阵).

引理. 线性变换T的任一特征值λ_i的代数重数大于等于几何重数, 即$c_i \geq d_i$.

证明. 设$\alpha_1, \alpha_2. \cdots, \alpha_{d_i}$是$V_{\lambda_i}$的一个基, 将它扩充$V$的一个基$\alpha_1, \alpha_2. \cdots, \alpha_{d_i}, \alpha_{d_i+1}, \cdots, \alpha_n$. 由于$V_{\lambda_i}$中非零向量都是属于特征值$\lambda_i$的特征向量, 即有$T(\alpha_j) = \lambda_i \alpha_j (1 \leq j \leq d_i)$, 因此,

$$T(\alpha_1, \cdots, \alpha_{d_i}, \alpha_{d_i+1}, \cdots, \alpha_n) = (\alpha_1, \cdots, \alpha_{d_i}, \alpha_{d_i+1}, \cdots, \alpha_n) \begin{pmatrix} A_{11} & A_{12} \\ \mathbf{0} & A_{22} \end{pmatrix},$$

其中

$$A_{11} = \begin{pmatrix} \lambda_i & 0 & \cdots & 0 \\ 0 & \lambda_i & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_i \end{pmatrix}_{d_i \times d_i},$$

A_{22}是$n - d_i$阶方阵, A_{12}是$d_i \times (n - d_i)$矩阵.于是T的特征多项式为

$$\begin{aligned} \Delta_T(\lambda) &= \begin{vmatrix} \lambda E_{d_i} - A_{11} & -A_{12} \\ 0 & \lambda E_{n-d_i} - A_{22} \end{vmatrix} \\ &= |\lambda E_{d_i} - A_{11}||\lambda E_{n-d_i} - A_{22}| \\ &= (\lambda - \lambda_i)^{d_i} |\lambda E_{n-d_i} - A_{22}| \end{aligned}$$

这说明, 特征值λ_i在$\Delta_T(\lambda)$中的重数c_i至少为d_i, 即$c_i \geq d_i$.

引理. 如果$\lambda_1, \lambda_2 \cdots, \lambda_s$ 是线性变换T 的全部两两不同特征值, $\alpha_i \in V_{\lambda_i}(i = 1, 2, \cdots, s)$, 则当$\alpha_1 + \alpha_2 + \cdots + \alpha_s = \mathbf{0}$ 时, 必有$\alpha_1 = \alpha_2 = \cdots = \alpha_s = \mathbf{0}$.

证明. 由于$T(\alpha_i) = \lambda_i\alpha_i(1 \le i \le s)$ 和$\alpha_1 + \alpha_2 + \cdots + \alpha_s = \mathbf{0}$, 所以对于$i = 0, 1, 2, \cdots, s$

$$\begin{aligned}\mathbf{0} = T^i(\mathbf{0}) &= T^i(\alpha_1 + \alpha_2 + \cdots + \alpha_s) \\ &= T^i(\alpha_1) + T^i(\alpha_2) + \cdots + T^i(\alpha_s) \\ &= \lambda_1^i\alpha_1 + \lambda_2^i\alpha_2 + \cdots + \lambda_s^i\alpha_s\end{aligned}$$

于是有下面一组等式

$$\begin{cases}\alpha_1 + \alpha_2 + \cdots + \alpha_s = \mathbf{0} \\ \lambda_1\alpha_1 + \lambda_2\alpha_2 + \cdots + \lambda_s\alpha_s = \mathbf{0} \\ \lambda_1^2\alpha_1 + \lambda_2^2\alpha_2 + \cdots + \lambda_s^2\alpha_s = \mathbf{0} \\ \cdots\cdots \\ \lambda_1^{s-1}\alpha_1 + \lambda_2^{s-1}\alpha_2 + \cdots + \lambda_s^{s-1}\alpha_s = \mathbf{0}\end{cases}$$

将这组方程写成矩阵形式,

$$(\alpha_1, \alpha_2, \cdots, \alpha_s)B = (\mathbf{0}, \mathbf{0}, \cdots, \mathbf{0})$$

这里

$$B = \begin{pmatrix} 1 & \lambda_1 & \cdots & \lambda_1^{s-1} \\ 1 & \lambda_2 & \cdots & \lambda_2^{s-1} \\ \cdots & \cdots & & \\ 1 & \lambda_s & \cdots & \lambda_s^{s-1} \end{pmatrix}$$

由于B的行列式是范德蒙行列式, 而且$\lambda_1, \lambda_2 \cdots, \lambda_s$ 两两不同, 所以B 是可逆矩阵. 于是

$$\begin{aligned}(\alpha_1, \alpha_2, \cdots, \alpha_s) &= (\alpha_1, \alpha_2, \cdots, \alpha_s)BB^{-1} \\ &= (\mathbf{0}, \mathbf{0}, \cdots, \mathbf{0})B^{-1} = (\mathbf{0}, \mathbf{0}, \cdots, \mathbf{0})\end{aligned}$$

即$\alpha_1 = \alpha_2 = \cdots = \alpha_s = \mathbf{0}$.

引理. 如果$\lambda_1, \lambda_2 \cdots, \lambda_s$ 是线性变换T 的两两不同特征值,

$$B_i = \{\alpha_{i1}, \alpha_{i2}, \cdots, \alpha_{ir_i}\}(i = 1, 2, \cdots, s)$$

是T 的属于特征值λ_i 的线性无关的特征向量组. 则向量组$B_1 \cup B_2 \cup \cdots \cup B_s$ 线性无关.

证明. 设V_{λ_i}是特征值λ_i的特征子空间, 如果

$$\sum_{j=1}^{r_1} a_{1j}\alpha_{1j} + \sum_{j=1}^{r_2} a_{2j}\alpha_{2j} + \cdots + \sum_{j=1}^{r_s} a_{sj}\alpha_{sj} = \mathbf{0},$$

则令$\beta_i = \sum_{j=1}^{r_i} a_{ij}\alpha_{ij}, (i = 1, 2, \cdots, s)$, 由于$B_i \subseteq V_{\lambda_i}$, 所以$\beta_i \in V_{\lambda_i}(i = 1, 2, \cdots, s)$ 于是由$\beta_1 + \beta_2 + \cdots + \beta_s = \mathbf{0}$推出$\beta_1 = \beta_2 = \cdots = \beta_s = \mathbf{0}$ 从而$\beta_i = \sum_{j=1}^{r_i} a_{ij}\alpha_{ij} = \mathbf{0}$. 但是向量组$B_i$是线性无关, 因此$a_{i1} = a_{i2} = \cdots = a_{ir_i} = 0(i = 1, 2, \cdots s)$, 即向量组$B_1 \cup B_2 \cup \cdots \cup B_s$ 线性无关.

命题. 如果n 维线性空间V 的线性变换T 的特征多项式$\Delta_T(\lambda)$ 有n 个不同的根(即T 有n 个不同特征值), 则T 必可对角化.

证明. 设$\lambda_i(i = 1, 2, \cdots n)$是$T$ 有n 个不同特征值, α_i是属于特征值λ_i的特征向量, 则$\{\lambda_i\}$线性无关, 由此推出

$$\bigcup_{i=1}^{n}\{\alpha_i\} = \{\alpha_1, \alpha_2, \cdots, \alpha_n\}$$

线性无关, 所以T n 个特征向量线性无关, 因此T 可对角化.

定理. 设n 维线性空间V 的线性变换T 的不同特征值为$\lambda_1, \lambda_2, \cdots, \lambda_s$, λ_i 的代数重数为c_i, 几何重数为d_i, 则T 可对角化的充分必要条件$c_i = d_i (i = 1, 2, \cdots, s)$

证明. (必要性) 设T 可对角化, 则T 有n 个线性无关的特征向量, 将这n 个线性无关特征向量按特征值进行分类, 所以对于$(k = 1, 2, \cdots, s)$可设$\alpha_{k1}, \alpha_{k2}, \cdots, \alpha_{ki_k}$ 是属于特征值λ_k 的特征向量, 而且$i_1 + i_2 + \cdots + i_s = n$. 于是

$$n = i_1 + i_2 + \cdots + i_s \leq d_1 + d_2 + \cdots + d_s \leq c_1 + c_2 + \cdots + c_s = n$$

再由于$d_k \leq c_k (k = 1, 2, \cdots, s)$ 所以$c_k = d_k (k = 1, 2, \cdots, s)$.

(充分性) 由于$\dim V_{\lambda_k} = d_k = c_k$, 分别在每个特征子空间$V_{\lambda_k}$ 中取基$B_k = \{\alpha_{k1}, \alpha_{k2}, \cdots, \alpha_{kd_k}\} (k = 1, 2, \cdots, s)$, 则向量组

$$\bigcup_{k=1}^{s} B_k$$

线性无关, 并且它恰好包含$d_1 + d_2 + \cdots + d_s = c_1 + c_2 + \cdots + c_s = n$ 个向量. 所以T 的n 个线性无关的特征向量. 因此T 可对角化.

例. 设

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

判断A是否可对角化?

解: 因为

$$\Delta_A(\lambda)=\begin{vmatrix}\lambda-5&6&6\\1&\lambda-4&-2\\-3&6&\lambda+4\end{vmatrix}=(\lambda-2)^2(\lambda-1),$$

所以A有2个不同特征值$\lambda_1=2,\lambda_2=1.$
V_{λ_1}是齐次线性方程组$(2E-A)X=\mathbf{0}$的解空间,

$$2E-A=\begin{pmatrix}-3&6&6\\1&-2&-2\\-3&6&6\end{pmatrix}\leftrightarrow\begin{pmatrix}1&-2&-2\\0&0&0\\0&0&0\end{pmatrix}$$

由于$R(2E-A)=1$, 因而$\dim V_{\lambda_1}=3-1=2=$特征值$\lambda_1$的代数重数. 又由于$1\leq\dim V_{\lambda_2}\leq$特征值$\lambda_2$的代数重数$=1$, 所以$\dim V_{\lambda_2}=1$, 由此推出$A$可以对角化.

例. 设

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

试问A是否可以对角化?

解. A的特征多项式为

$$\Delta_A(\lambda)=|\lambda E-A|=\begin{vmatrix}\lambda+1 & -1 & 0\\ 4 & \lambda-3 & 0\\ -1 & 0 & \lambda-2\end{vmatrix}=(\lambda-2)(\lambda-1)^2$$

所以A的特征值为$\lambda_1=1$和$\lambda_2=2$.
V_{λ_1} 是齐次线性方程组$(1E-A)X=\mathbf{0}$ 的解空间,

$$(1E-A)=\begin{pmatrix}2 & -1 & 0\\ 4 & -2 & 0\\ -1 & 0 & -1\end{pmatrix}\leftrightarrow\begin{pmatrix}1 & 0 & 1\\ 0 & 1 & 2\\ 0 & 0 & 0\end{pmatrix}$$

所以$R(1E-A)=2$, 于是$\dim V_{\lambda_1}=3-2=1\neq$特征值$\lambda_1$的代数重数$=2$. 所以$A$不可以对角化.

练习. 习题7.4: 1,2,3.

命题.[汉米尔顿-凯莱Hamilton-Cayley]　设域$\mathbb{F}$ 上n 阶方阵A 的特征多项式为

$$\Delta_A(\lambda) = |\lambda E - A| = \lambda^n + c_1\lambda^{n-1} + \cdots + c_{n-1}\lambda + c_n,$$

则

$$\Delta_A(A) = A^n + c_1 A^{n-1} + \cdots + c_{n-1}A + c_n E = \mathbf{0}.$$

这个命题的证明将留在下后面. 这里用一个例子来说明.

例.　设

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

则

$$\begin{aligned} \Delta_A(\lambda) &= \begin{vmatrix} \lambda - 2 & -1 \\ 3 & \lambda + 1 \end{vmatrix} = \lambda^2 - \lambda + 1, \\ \Delta_A(A) &= A^2 - A + E \\ &= \begin{pmatrix} 1 & 1 \\ -3 & -2 \end{pmatrix} - \begin{pmatrix} 2 & 1 \\ -3 & -1 \end{pmatrix} + \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \\ &= \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}. \end{aligned}$$

定义. 设A 是n 阶方阵, $f(\lambda) = a_0\lambda^n + a_1\lambda^{n-1} + \cdots + a_{n-1}\lambda + a_n$, 如果$f(A) = \mathbf{0}$, 则称多项式$f(\lambda)$ 为A 的零化多项式.

注意. 对于n 阶方阵A, A的特征多项式$\Delta_A(\lambda)$是A 的一个零化多项式. 所以对于n 阶方阵A, 始终存在一个n 次多项式, 它零化A.

定义. 对给定的n 阶方阵A, 一个多项式如果它是A 的零化多项式中次数最低而且首项系数为1 多项式, 则称它为A 的最小多项式.

例. 求

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

的最小多项式.

解. 因为

$$A^2 = \begin{pmatrix} 1 & 0 & 0 \\ -1 & 0 & 0 \\ 1 & 1 & 1 \end{pmatrix} = A,$$

所以$\lambda^2 - \lambda$ 是A 的零化多项式. 由于$A \neq d_0E$, 因而$\lambda^2 - \lambda$ 是A 的一个最小多项式.

命题. 对于n 阶方阵A, A 的最小多项式唯一存在.

证明. (存在性)在线性空间$M_n(\mathbb{F})$, 求出最小s使得$E, A, \cdots A^{s-1}, A^s$线性相关, 而$E, A, \cdots A^{s-1}, A^{s-1}$线性无关. 由于$M_n(\mathbb{F})$是有限维的, 所以这样的$s$存在. 于是$A^s$可以由$E, A, \cdots A^{s-1}, A^{s-1}$线性表示, 即存在$a_0, a_1, \cdots, a_{s-1}$使得

$$A^s = a_0E + a_1A + \cdots + a_{s-1}A^{s-1}$$

所以多项式$f(\lambda) = -a_0 - a_1\lambda - \cdots - a_{s-1}\lambda^{s-1} + \lambda^s$零化$A$, 由于$E, A, \cdots A^{s-1}, A^{s-1}$线性无关, 所以次数小于$s$的多项式不可能零化$A$. 即$f(\lambda)$是$A$的最小多项式.

(唯一性) 如果$m_1(\lambda), m_2(\lambda)$ 都是A 的最小多项式, 则它们有相同的次数(设为r), 而且首项系数都是1. 这样, 多项式

$$g(\lambda) = m_1(\lambda) - m_2(\lambda) \neq 0$$

而且$g(\lambda)$ 是次数$< r$ 的多项式.

$$g(A) = m_1(A) - m_2(A) = 0,$$

这矛盾于A 的零化多项式的次数至少为r 的假设.

注意. 由于n 阶方阵A的最小多项式唯一存在, 所以下面我们将用$m_A(\lambda)$表示A的最小多项式.

命题. 多项式$f(\lambda)$ 是A 的零化多项式的充分必要条件为A 的最小多项式$m_A(\lambda)$ 整除$f(\lambda)$, 特别$m_A(\lambda)|\Delta_A(\lambda)$.

证明. (必要性) 设$f(\lambda)$ 是A 的零化多项式. 由多项式带余除法得

$$f(\lambda) = q(\lambda)m_A(\lambda) + r(\lambda),$$

其中$r(\lambda)$ 的次数小于$m_A(\lambda)$ 的次数. 如果$r(\lambda) \neq 0$, 则

$$f(A) = q(A)m_A(A) + r(A).$$

由$f(A) = 0$ 和$m_A(A) = 0$推出$r(A) = 0$. 这与$m_A(\lambda)$是最小多项式矛盾. 所以$r(\lambda) = 0$, 即$m_A(\lambda)|f(\lambda)$.

(充分性) 如果$m_A(\lambda)|f(\lambda)$. 则$f(\lambda) = q(\lambda)m_A(\lambda)$. 所以$f(A) = q(A)m_A(A) = 0$.

例. 设n阶非零方阵A满足$A^2 = A$, 而且$A \neq E$. 求最小多项式$m_A(\lambda)$

解. 由$A^2 = A$, $f(\lambda) = \lambda^2 - \lambda$ 是A 的一个零化多项式. 这样, $m_A(\lambda)$ 应是$f(\lambda)$ 的因子, 即$m_A(\lambda)$ 有三种可能:

(1) $m_A(\lambda) = \lambda$, 这时$A = m_A(A) = \mathbf{0}$, 不合题意;

(2) $m_A(\lambda) = \lambda - 1$, 这时$A - E = m_A(A) = \mathbf{0}$, 即$A = E$, 也不合题意;

(3) 这时, 仅有的可能是$m_A(\lambda) = \lambda(\lambda - 1)$.

推论. 相似矩阵有相同的最小多项式.

证明. 设$B = P^{-1}AP$, 则

$$m_A(B) = m_A(P^{-1}AP) = P^{-1}m_A(A)P = P^{-1}\mathbf{0}P = \mathbf{0}$$

所以$m_A(\lambda)$ 是B 的零化多项式, 因此$m_B(\lambda)|m_A(\lambda)$. 同理可证$m_A(\lambda)|m_B(\lambda)$. 两个多项式的首项系数都是1, 所以$m_A(\lambda) = m_B(\lambda)$.

定义. T的最小多项式为T 在V 的某一个基下对应的矩阵A 的最小多项式, 用$m_T(\lambda)$表示.

说明. 由于线性空间V 上线性变换T在不同基下的矩阵是相似的, 而相似不改变最小多项式, 所以这个定义是有明确意义的.

练习. 习题7.4:4.

命题. 如果A的特征多项式为

$$\Delta_A(\lambda) = (\lambda - \lambda_1)^{c_1}(\lambda - \lambda_2)^{c_2} \cdots (\lambda - \lambda_s)^{c_s},$$

则A的最小多项式形如

$$m_A(\lambda) = (\lambda - \lambda_1)^{e_1}(\lambda - \lambda_2)^{e_2} \cdots (\lambda - \lambda_s)^{e_s},$$

其中$1 \leq e_i \leq c_i \ (i = 1, 2, \cdots, s)$.

证明. 由$m_A(\lambda)|\Delta_A(\lambda)$, 推出$e_i \leq c_i \ (i = 1, 2, \cdots, s)$. 所以我们只要证$e_i \geq 1 (i = 1, 2, \cdots, s)$.
如果有某个$e_j = 0$, 为书写简便起见, 不妨设$e_1 = 0$, 则

$$m_A(\lambda) = (\lambda - \lambda_2)^{e_2}(\lambda - \lambda_3)^{e_3} \cdots (\lambda - \lambda_s)^{e_s}$$

设α_1是属于特征值λ_1的特征向量. 则

$$\begin{aligned} \mathbf{0} = \mathbf{0}\alpha &= m_A(A)\alpha_1 = (A - \lambda_2 E)^{e_2}(A - \lambda_3 E)^{e_3} \cdots (A - \lambda_s E)^{e_s}\alpha_1 \\ &= (\lambda_1 - \lambda_2)^{e_2}(\lambda_1 - \lambda_3)^{e_3} \cdots (\lambda_1 - \lambda_s)^{e_s}\alpha_1 \end{aligned}$$

由于当$i \geq 2$时, $\lambda_i \neq \lambda_1$, 这迫使$\alpha_1 = \mathbf{0}$, 这是一个矛盾, 因此$1 \leq e_i \leq d_i \ (i = 1, 2, \cdots, s)$.

说明. 命题告诉我们A的最小多项式的根包含了A的所有特征值.

定理. n 阶方阵A 可对角化的充分必要条件是A 的最小多项式$m_A(\lambda)$ 无重根.

证明. (必要性) A 相似于对角阵

$$D=\mathrm{diag}(D_1,D_2,\cdots,D_s),\ \text{其中}D_i=\mathrm{diag}(\underbrace{\lambda_i,\cdots,\lambda_i}_{c_i\text{个}}),\ i=1,2,\cdots,s$$

则$m_A(\lambda)=m_D(\lambda)$并且

$$\Delta_A(\lambda)=\Delta_D(\lambda)=(\lambda-\lambda_1)^{c_1}(\lambda-\lambda_2)^{c_2}\cdots(\lambda-\lambda_s)^{c_s}.$$

设

$$f(\lambda)=(\lambda-\lambda_1)(\lambda-\lambda_2)\cdots(\lambda-\lambda_s)$$

由于$m_D(\lambda)$包含D的所有特征值, 所以$f(\lambda)|m_D(\lambda)$. 另一方面由于

$$f(D)=\mathrm{diag}(f(D_1),f(D_2),\cdots,f(D_s))=\mathbf{0}$$

所以$m_D(\lambda)|f(\lambda)$. 由此推出$m_A(\lambda)=m_D(\lambda)=f(\lambda)$, 于是$m_A(\lambda)$ 无重根.

(充分性) 设

$$m_A(\lambda) = (\lambda - \lambda_1)(\lambda - \lambda_2)\cdots(\lambda - \lambda_s)$$

则$\lambda_1, \lambda_2, \cdots, \lambda_s$ 是A 的全部不同的特征值. 对于$(i = 1, 2, \cdots, s)$, 令

$$f_i(\lambda) := \frac{m_A(\lambda)}{\lambda - \lambda_i}$$

对于任意$X \in \mathbf{F}^n$,

$$(\lambda_i E - A) f_i(A) X = -m_A(A) X = \mathbf{0} X = 0$$

这说明$f_i(A)X \in V_{\lambda_i}$. 由于多项式$f_1(\lambda), f_2(\lambda), \cdots, f_s(\lambda)$没有次数大于零公因式, 所以存在多项式$u_1(\lambda), u_2(\lambda), \cdots, u_s(\lambda)$使得

$$u_1(\lambda) f_1(\lambda) + u_2(\lambda) f_2(\lambda) + \cdots + u_s(\lambda) f_s(\lambda) = 1$$

由此得

$$X = f_1(A) u_1(A) X + f_2(A) u_2(A) X + \cdots + f_s(A) u_s(A) X$$

这说明

$$\mathbf{F}^n = V_{\lambda_1} + V_{\lambda_2} + \cdots + V_{\lambda_s}$$

由于零向量的分解式是唯一的, 所以上面的和是直和, 因此

$$\dim V_{\lambda_1} + \dim V_{\lambda_2} + \cdots + \dim V_{\lambda_s} = n$$

所以A有n个线性无关的特征向量.

例. 设T 是幂零线性变换, 即存在$k \geq 2$, 使$T^{k-1} \neq \mathcal{O}$, 但是$T^k = \mathcal{O}$, 则T 不可对角化.

证明. 由$T^k = \mathbf{0}$, 知$f(\lambda) = \lambda^k$ 是T 的零化多项式. 于是T 的最小多项式形如$m_(\lambda) = \lambda^r$, 其中$1 \leq r \leq k$. 但$r = 1$ 意味着$T = \mathcal{O}$, 矛盾于假设. 因此$r \geq 2$, 即T 的最小多项式有重根, 从而幂零线性变换T 不可对角化.

例. 设

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

试问A可对角化吗?

解. (1) 求A 的特征多项式

$$|\lambda E - A| = \begin{vmatrix} \lambda - 5 & 6 & 6 \\ 1 & \lambda - 4 & -2 \\ -3 & 6 & \lambda + 4 \end{vmatrix} = (\lambda - 2)^2(\lambda - 1).$$

(2) 验算$M = (A - E)(A - 2E)$ 是否为零

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

因而$m_A(\lambda) = (\lambda - 1)(\lambda - 2)$ 无重根, A 可对角化.

例. 设

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

请问A可对角化吗?

解. (1) 求A 的特征多项式

$$|\lambda E-A|=\begin{vmatrix} \lambda+1 & -1 & 0 \\ 4 & \lambda-3 & 0 \\ -1 & 0 & \lambda-2 \end{vmatrix}=(\lambda-1)^2(\lambda-2).$$

(2) 验算$M=(A-E)(A-2E)$ 是否为零

$$(A-E)(A-2E)=\begin{pmatrix} -2 & 1 & 0 \\ -4 & 2 & 0 \\ 1 & 0 & 1 \end{pmatrix}\begin{pmatrix} -3 & 1 & 0 \\ -4 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}\neq \mathbf{0}.$$

因而$m_A(\lambda)=\Delta_A(\lambda)=(\lambda-1)^2(\lambda-2)$ 有重根, A 不可对角化.

练习1. 设$A\in M_n(\mathbb{F})$,

$$V:=\{f(A)|f(x)\in\mathbb{F}[x]\}$$

证明$M_n(\mathbb{F})$的子空间V的维数等于A的最小多项式的次数.

练习2. 习题7.4:5(2),6,7,10.