第二章 方阵的相似化简

§2.1 特征多项式和最小多项式

对于复数域上 n 阶方阵 $A=[a_{ij}]$，它的特征多项式

$$\det(\lambda I_n - A) = \begin{vmatrix} \lambda - a_{11} & -a_{12} & \cdots & -a_{1n} \\ -a_{21} & \lambda - a_{22} & \cdots & -a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ -a_{n1} & -a_{n2} & \cdots & \lambda - a_{nn} \end{vmatrix}$$

是λ的 n 次多项式

$$f(\lambda) \triangleq \det(\lambda I_n - A) = \lambda^n + b_1\lambda^{n-1} + \cdots + b_{n-1}\lambda + b_n .$$

这个多项式在复数域有 n 个根 $\lambda_1, \lambda_2, \cdots, \lambda_n$.

（有重根时重复出现），故 $f(\lambda)$ 又可表示为

$$f(\lambda) \triangleq \det(\lambda I_n - A) = (\lambda - \lambda_1)(\lambda - \lambda_2)\cdots(\lambda - \lambda_n).$$

注：(1) $-b_1 = \lambda_1 + \lambda_2 + \cdots + \lambda_n = a_{11} + a_{22} + \cdots + a_{nn} \triangleq \operatorname{tr} A;$ (2.1-1)

(2) $(-1)^n b_n = \lambda_1 \cdot \lambda_2 \cdots \lambda_n = \det A.$ (2.1-2)

(2.1-1)式说明,方阵A的迹tr$A$(即A的对角线元素之和)等于A的所有特征值(可以相重)的和.而(2.1-2)式表明,A的行列式等于所有特征值的乘积.

定义方阵 A 的多项式为

$$a_0A^m + a_1A^{m-1} + \cdots + a_mI,$$

并把它看作是多项式

$$g(t) = a_0t^m + a_1t^{m-1} + \cdots + a_m$$

中代换 t 为 A 的结果, 记为$g(A)$, 它是与 A 同阶的方阵.

例如, $g(t)=t^2-3t+4, h(t)=t^2-3t+2$

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

则有

$$g(A)=A^2-3A+4I=\begin{bmatrix}2 & 0\\ 0 & 2\end{bmatrix},$$

$$h(A)=A^2-3A+2I=\begin{bmatrix}0 & 0\\ 0 & 0\end{bmatrix}.$$

设 $p(t),q(t)$ 是两个多项式,它们的乘积

$$h(t)=p(t)q(t)$$

仍是一个多项式. 由于 $p(t)q(t)=q(t)p(t)$, 所以有

$$h(A)=p(A)q(A)=q(A)p(A),$$

即 $p(A)$与$q(A)$ 可以互换.

定理 设n阶方阵A的n个特征值是$\lambda_1,\lambda_2,\cdots,\lambda_n$,相应的特征向量分别为 $x_1,x_2,\cdots,x_n$,而$g(t)$是一多项式, 那么$g(A)$的n个特征值是 $g(\lambda_1),g(\lambda_2),\cdots,g(\lambda_n)$,且 $x_1,x_2,\cdots,x_n$ 分别是相应的特征向量.

证 不妨设

$$g(t)=a_0t^m+a_1t^{m-1}+\cdots+a_m$$

由于

$$Ax_i=\lambda_i x_i, A^2x_i=A(Ax_i)=\lambda_i(Ax_i)=\lambda_i^2x_i,\cdots,A^mx_i=\lambda_i^mx_i$$

故

$$g(A)x_i=(a_0A^m+a_1A^{m-1}+\cdots+a_mI)x_i$$

$$=(a_0\lambda_i^m+a_1\lambda_i^{m-1}+\cdots+a_m)x_i=g(\lambda_i)x_i, 1\le i\le n,$$

所以 $g(\lambda_i)$ 是 $g(A)$ 的特征值, x_i 是相应的特征向量.

方阵$A\neq 0$称为幂零矩阵,如果存在正整数k，使$A^k=O$.

容易验证 $A=\begin{bmatrix}0 & 1 & \\ & 0 & 1\\ & & 0\end{bmatrix}$ 是幂零矩阵,因为 $A^2=\begin{bmatrix}0 & 0 & 1\\ & 0 & 0\\ & & 0\end{bmatrix}, A^3=O.$

幂零矩阵的特征值必是0.事实上,设 λ_0 是幂零矩阵A的特征值, x_0是相应的特征向量,则由上述定理知

$$0=A^k x_0=\lambda_0^k x_0, x_0\neq 0$$

故 $\lambda_0^k=0$,即 $\lambda_0=0$.

定理 (Sylverster) 设A是$m\times n$矩阵, B是$n\times m$矩阵($m\geq n$), 方阵AB, BA 的特征多项式分别为 $f_{AB}(\lambda), f_{BA}(\lambda)$ ，则有

$$f_{AB}(\lambda)=\lambda^{m-n}f_{BA}(\lambda) \qquad (2.1\text{-}3)$$

证 设rankA＝r，则存在m阶可逆矩阵P和n阶可逆矩阵Q，使

$$PAQ=\begin{array}{cl}\left[\begin{array}{c|c} I_r & O \\ \hline O & O \end{array}\right] & \begin{array}{l}\}r^* \\ \}m-r^*\end{array} \\ \begin{array}{cc}\underbrace{}_{r} & \underbrace{}_{n-r}\end{array} & \end{array}$$

因而

$$PABP^{-1}=PAQQ^{-1}BP^{-1}$$

$$=\left[\begin{array}{c|c} I_r & O \\ \hline O & O \end{array}\right]\left[\begin{array}{c|c} G_{11} & G_{12} \\ \hline G_{21} & G_{22} \end{array}\right]=\begin{array}{cl}\left[\begin{array}{c|c} G_{11} & G_{12} \\ \hline O & O \end{array}\right] & \begin{array}{l}\}r \\ \}m-r,\end{array} \\ \begin{array}{cc}\underbrace{}_{r} & \underbrace{}_{m-r}\end{array} & \end{array}$$

其中

$$Q^{-1}BP^{-1} \triangleq \left[\begin{array}{c|c} G_{11} & G_{12} \\ \hline G_{21} & G_{22} \end{array}\right]\begin{array}{l}\}r \\ \}n-r.\end{array}$$

同理可得

$$Q^{-1}BAQ = Q^{-1}BP^{-1}PAQ$$

$$= \left[\begin{array}{c|c} G_{11} & G_{12} \\ \hline G_{21} & G_{22} \end{array}\right]\left[\begin{array}{c|c} I_r & O \\ \hline O & O \end{array}\right] = \left[\begin{array}{c|c} G_{11} & O \\ \hline G_{21} & O \end{array}\right]\begin{array}{l}\}r \\ \}n-r.\end{array}$$

$$\begin{array}{cc} \underbrace{\quad}_{r} & \underbrace{\quad}_{n-r} \end{array}$$

于是

$$f_{AB}(\lambda) = \det(\lambda I_m - AB)$$

$$= \det(\lambda I_m - PABP^{-1}) = \lambda^{m-r}\det(\lambda I_r - G_{11}),$$

$$f_{BA}(\lambda) = \det(\lambda I_n - BA)$$

$$= \det(\lambda I_n - Q^{-1}BAQ) = \lambda^{n-r}\det(\lambda I_r - G_{11}).$$

比较上述两式便得（2.1-3）式.

这个定理表明，m 阶方阵 AB 与 n 阶方阵 BA 的非零特征值是相同的，特别，若 $m=n$，则 AB 与 BA 的特征值相同.

例1 求镜像变换的Householder矩阵

$$H = I_n - 2\omega\omega^T$$

的特征值及它的迹和行列式.

解 令 $A = 2\omega, B = \omega^T$, 则 $AB = 2\omega\omega^T, BA = 2\omega^T\omega = 2$, 故

$$\det(\lambda I_n - H) = \det[(\lambda - 1)I_n + 2\omega\omega^T]$$
$$= (\lambda - 1)^{n-1}\det(\lambda - 1 + 2\omega^T\omega) = (\lambda - 1)^{n-1}(\lambda + 1).$$

因而 $\lambda = 1$ 是 H 的 $n-1$ 重特征值, $\lambda = -1$ 是单重特征值.

再由(2.1-1)和(2.1-2)式知

$$\mathrm{tr}H = n - 1 + (-1) = n - 2, \quad \det H = -1.$$

例2 若方阵A的所有特征值的模都小于1,则方阵$I-A$是可逆的.

证 设 $\lambda_1,\lambda_2,\cdots,\lambda_n$ 是A的所有特征值,且$|\lambda_i|<1, i=1,2,\cdots,n$,

则$1-\lambda_1,1-\lambda_2,\cdots,1-\lambda_n$是$I-A$的所有特征值.由于

$$|1-\lambda_i|\geq 1-|\lambda_i|>0, i=1,2,\cdots,n,$$

所以 $1-\lambda_i\neq 0$,故

$$\det(I-A)=(1-\lambda_1)(1-\lambda_2)\cdots(1-\lambda_n)\neq 0$$

从而$I-A$是可逆的.

定义 设A是一个n阶方阵,$g(t)$是一多项式,如果$g(A)=O$,则称$g(t)$是A的零化多项式.

如$h(t)=t^2-3t+2$就是 $A=\begin{bmatrix}1 & 1\\ 0 & 2\end{bmatrix}$ 的一个零化多项式.

如果$g(t)$是A的一个零化多项式,而$h(t)$是任一多项式,则$h(t)g(t)$是A的零化多项式,因此方阵的零化多项式不是唯一的.那么零化多项式是否存在呢?下述定理说明A的零化多项式是存在的.

定理(Cayley-Hamilton) 设n阶方阵A的特征多项式为

$$f(\lambda)=\det(\lambda I-A)=\lambda^n+b_1\lambda^{n-1}+\cdots+b_{n-1}\lambda+b_n$$

则$f(A)=O$,即A的特征多项式是A的一个零化多项式.

证 设$B(\lambda)$是$\lambda I-A$的伴随矩阵，则有

$$B(\lambda)(\lambda I-A)=f(\lambda)I=\lambda^n I+b_1\lambda^{n-1}I+\cdots+b_nI. \quad (2.1\text{-}4)$$

由于方阵$B(\lambda)$的元是$\lambda I-A$的 n-1 阶代数余子式，故均是λ的次数不大于n-1的多项式.从而由矩阵的运算性质，$B(\lambda)$可以表示成

$$B(\lambda)=\lambda^{n-1}B_0+\lambda^{n-2}B_1+\cdots+B_{n-1}, \quad (2.1\text{-}5)$$

其中 $B_0, B_1, \cdots, B_{n-1}$ 都是 *n* 阶数矩阵.于是

$$B(\lambda)(\lambda I - A) = (\lambda^{n-1} B_0 + \lambda^{n-2} B_1 + \cdots + B_{n-1})(\lambda I - A)$$

$$= \lambda^n B_0 + \lambda^{n-1}(B_1 - B_0 A) + \cdots + \lambda(B_{n-1} - B_{n-2}A) - B_{n-1}A. \quad (2.1\text{-}6)$$

比较（2.1-4）和（2.1-6）式中λ各次幂的系数，得

$$\left.\begin{array}{lll} \lambda^n & : & B_0 = I \\ \lambda^{n-1} & : & B_1 - B_0 A = b_1 I \\ \vdots & & \vdots \\ \lambda & : & B_{n-1} - B_{n-2}A = b_{n-1} I \\ \lambda^0 & : & -B_{n-1}A = b_n I \end{array}\right\} \quad (2.1\text{-}7)$$

以$A^n, A^{n-1}, \cdots, A, I$ 依次从右边乘（2.1-7）式中的第一式，第二式，…，第n式，第n+1式的两端，再把这n+1个式子加起来，则左端为零矩阵，右端即为f(A).因此 $f(A) = O.$

例3 已知

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

求 $g(A)=2A^5-3A^4-A^3+2A-I.$

解 A的特征多项式

$$f(\lambda)=\begin{vmatrix}\lambda-1 & -1 & 1\\ -1 & \lambda-1 & -1\\ 0 & 1 & \lambda-2\end{vmatrix}=(\lambda-1)^2(\lambda-2).$$

又 $g(\lambda)=2\lambda^5-3\lambda^4-\lambda^3+2\lambda-1$. 用 $f(\lambda)$ 除 $g(\lambda)$ 得

$$g(\lambda)=(2\lambda^2+5\lambda+9)f(\lambda)+15\lambda^2-33\lambda+17.$$

由于$f(A)=O$,故

$g(A)=15A^2-33A+17I$

$$=15\begin{bmatrix}2 & 3 & -2\\ 2 & 1 & 2\\ -1 & -3 & 3\end{bmatrix}-33\begin{bmatrix}1 & 1 & -1\\ 1 & 1 & 1\\ 0 & -1 & 2\end{bmatrix}+17\begin{bmatrix}1 & & \\ & 1 & \\ & & 1\end{bmatrix}$$

$$=\begin{bmatrix}14 & 12 & 3\\ -3 & -1 & -3\\ -15 & -12 & -4\end{bmatrix}.$$

如果不做多项式的除法，可采用待定系数法，用$f(\lambda)$除$g(\lambda)$一般得到下述表达式

$$g(\lambda)=h(\lambda)f(\lambda)+r(\lambda), \qquad (2.1\text{-}8)$$

其中余式$r(\lambda)$是一多项式，其次数$\deg r(\lambda)<\deg f(\lambda)=3$，因而可以设

$$r(\lambda)=a\lambda^2+b\lambda+c,$$

代入（2.1-8）式得

$$g(\lambda)=h(\lambda)f(\lambda)+a\lambda^2+b\lambda+c. \qquad (2.1\text{-}9)$$

将特征值$\lambda=1, \lambda=2$分别代入（2.1-9）式，则有

$$\begin{cases} a+b+c=g(1)=-1, \\ 4a+2b+c=g(2)=11. \end{cases}$$

为了得到a,b,c之间的第三个关系式，我们在（2.1-9）式两边对λ 求导，得到

$$g'(\lambda)=h'(\lambda)f(\lambda)+h(\lambda)f'(\lambda)+2a\lambda+b.$$

用$\lambda=1$代入，考虑到$\lambda=1$是$f(\lambda)=0$的二重根，故有

$f(1)=f'(1)=0,$ 于是

$$2a+b=g'(1)=-3.$$

解线性方程组

$$\begin{cases} a+b+c \quad =-1. \\ 4a+2b+c=11, \\ 2a+b \qquad =-3, \end{cases}$$

得$a=15,b=-33,c=17.$ 从而得到同样的结果

$$g(A)=r(A)=15A^2-33A+17I.$$

例4 就例3中的A，求逆矩阵A^{-1}。

解 由于A的特征多项式$f(\lambda)=\lambda^3-4\lambda^2+5\lambda-2$中常数项不是零，所以$\det tA\neq 0$，故$A$可逆。

又 $f(A)=A^3-4A^2+5A-2I=O$, 用A^{-1}乘之, 便得A^{-1}的表示式

$$A^{-1}=\frac{1}{2}\left(A^2-4A+5I\right)=\frac{1}{2}\begin{bmatrix} 3 & -1 & 2 \\ -2 & 2 & -2 \\ -1 & 1 & 0 \end{bmatrix}$$

定义　A的零化多项式中，次数最低的首一多项式称为A的**最小多项式**，记为 $m_A(\lambda)$ 。

定理　A的最小多项式 $m_A(\lambda)$ 可整除A的任何零化多项式 $g(\lambda)$ ，且 $m_A(\lambda)$ 是唯一的。

证　若$m_A(\lambda)$不能整除$g(\lambda)$，则有 $g(\lambda)=p(\lambda)m_A(\lambda)+r(\lambda)$, 其中 $\deg r(\lambda)<\deg m_A(\lambda)$. 由于$O=g(A)=p(A)m_A(A)+r(A)=r(A)$, 所以 $r(\lambda)$是 A 得一个零化多项式，且次数低于$m_A(\lambda)$ 的次数.这与 $m_A(\lambda)$ 的定义矛盾.因而$m_A(\lambda)$ 必能整除 $g(\lambda)$,记为 $m_A(\lambda)\mid g(\lambda)$.

再证唯一性.设 $m_1(\lambda), m_2(\lambda)$ 是A的两个最小多项式，则

$m_1(\lambda)|m_2(\lambda)$，且 $m_2(A)|m_1(A)$. 因而，$m_1(A)=bm_2(A), b\neq 0$ 是常数.

但 $m_1(A), m_2(A)$ 都是首一多项式，故有b=1，从而 $m_1(A)=m_2(A)$.

这个定理的一个推论是，A的最小多项式必能整除 A 的特征多项式。

定理 λ_0 是*A* 的特征值的充分必要条件是 λ_0 是*A* 的最小多项式 $m_A(\lambda)$ 的根。

证 设 λ_0 是*A*的特征值，x_0是相应的特征向量，则有

$$0 = m_A(A)x_0 = m_A(\lambda_0)x_0, \quad x_0 \neq 0,$$

故 $m_A(\lambda_0)=0$，即 λ_0 是 $m_A(\lambda)$ 的根。

反之，若 λ_0 是 $m_A(\lambda)$ 的根，那么由于 $m_A(\lambda)$ 可整除*A*的特征多项式 $f(\lambda)$，故 λ_0 必是特征多项式的根，即 λ_0 是*A* 的特征值。

这个定理给出了由特征多项式检验最小多项式的方法。

事实上，设*A*的所有不同的特征值是$\lambda_1, \lambda_2, \cdots, \lambda_s, A$ 的特征多项式为

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

且$n_1+n_2+\cdots+n_s=n$ ，则*A*的最小多项式一定有如下形式：

$$m_A(\lambda)=(\lambda-\lambda_1)^{k_1}(\lambda-\lambda_2)^{k_2}\cdots(\lambda-\lambda_s)^{k_s},$$

并且

$$1\le k_i\le n_i \quad (i=1,2,\cdots,s)$$

若*A* 的特征值的代数重数都为1，那么 $m_A(\lambda)=f(\lambda)$。

例5 求

$$A=\left[\begin{array}{cc|cc} 2 & 1 & & \\ 0 & 2 & & \\ \hline & & 1 & 1 \\ & & -2 & 4 \end{array}\right]$$

的最小多项式。

解 由于$\det(\lambda I-A)=(\lambda-3)(\lambda-2)^3$，所以$A$的最小多项式只能

有下列三种可能：

$$(\lambda-3)(\lambda-2);\ (\lambda-3)(\lambda-2)^2;\ (\lambda-3)(\lambda-2)^3$$

但

$$(A-3I)(A-2I)=\left[\begin{array}{cc|cc} -1 & 1 & & \\ 0 & -1 & & \\ \hline & & -2 & 1 \\ & & -2 & 1 \end{array}\right]\left[\begin{array}{cc|cc} 0 & 1 & & \\ 0 & 0 & & \\ \hline & & -1 & 1 \\ & & -2 & 2 \end{array}\right]$$

$$=\left[\begin{array}{cc|cc} 0 & -1 & & \\ 0 & 0 & & \\ \hline & & 0 & 0 \\ & & 0 & 0 \end{array}\right]\neq O,$$

而

$$(A-3I)(A-2I)^2=O, \text{故 } m_A(\lambda)=(\lambda-3)(\lambda-2)^2.$$

例6 求下述对角块矩阵A的最小多项式：

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

解 A的特征多项式是$(\lambda+2)^5(\lambda-3)^3$，故$A$的最小多项式的形式是 $(\lambda+2)^k(\lambda-3)^l$ 。由于

$$A+2I=\begin{pmatrix} 0 & 1 & & & & & & \\ & 0 & & & & & & \\ & & 0 & 1 & & & & \\ & & & 0 & 1 & & & \\ & & & & 0 & & & \\ & & & & & 5 & & \\ & & & & & & 5 & 1 \\ & & & & & & & 5 \end{pmatrix}$$

$$A-3I=\begin{pmatrix} -5 & 1 & & & & & & \\ & -5 & & & & & & \\ & & -5 & 1 & & & & \\ & & & -5 & 1 & & & \\ & & & & -5 & & & \\ & & & & & 0 & & \\ & & & & & & 0 & 1 \\ & & & & & & & 0 \end{pmatrix}$$

所以由分块矩阵乘法可知，要使 $(A+2I)^k(A-3I)^l=O$ ，只需注意使幂零矩阵的那部分变为零矩阵所需乘积的次数。

$A+2I$ 需自乘3次，即在 $(A+2I)^3$ 中原幂零矩阵的那部分变为零矩阵，而 $A-3I$ 只需2次便可达到目的。因此，A 的最小多项式是

$$m_A(\lambda)=(\lambda+2)^3(\lambda-3)^2$$

一般地，若 $A=diag(A_1,A_2,\cdots,A_m)$ ，且方阵 A_i 的最小多项式为 $m_{A_i}(\lambda),1\le i\le m$ ，则 A 的最小多项式 $m_A(\lambda)$ 是 $m_{A_1}(\lambda)$ $m_{A_2}(\lambda),\cdots,m_{A_m}(\lambda)$ 的最小公倍式。这是因为 A_i 的最小多项式 $m_{A_i}(\lambda)(1\le i\le m)$ 必可整除 A 的最小多项式 $m_A(\lambda)$ ，即

$$m_{A_i}(\lambda)\mid m_A(\lambda)$$

例如，例5中的 $A=\begin{bmatrix} A_1 & \\ & A_2 \end{bmatrix}$，且 $A_1=\begin{bmatrix} 2 & 1 \\ 0 & 2 \end{bmatrix}$，$A_2=\begin{bmatrix} 1 & 1 \\ -2 & 4 \end{bmatrix}$

A_1的最小多项式是 $(\lambda-2)^2$，A_2的最小多项式是$(\lambda-3)(\lambda-2)$ 。

而这两个多项式的最小公倍式是 $(\lambda-3)(\lambda-2)^2$，故

$$m_A(\lambda)=(\lambda-3)(\lambda-2)^2$$

在讨论如何由方阵A 的最小多项式判断A是否可对角化的问题之前，我们给出有关矩阵乘积的秩的一个重要不等式。

定理 设A是$m \times n$矩阵，B是$n \times s$矩阵，则

$$rankA + rankB - n \le rank(AB) \le \min\{rankA, rankB\} \qquad (2.1\text{-}10)$$

证 只证前一个不等式，后一个不等式自证.

设$B = [b_1\ b_2\ \cdots\ b_r \vdots b_{r+1}\ \cdots\ b_s] = [\underset{n\times r}{B_1} \vdots \underset{n\times(s-r)}{B_2}]$, 且 $rankB = rankB_1 = r$,

那么B_2中的任意一列$b_j (r+1 \le j \le s)$都可由$b_1, b_2, \cdots, b_r$线性表出，

从而存在$r\times(s-r)$矩阵G，使$B_2 = B_1 G$, 即 $B_2 - B_1 G = O$. 令

$$Q = \begin{bmatrix} I_r & -G \\ O & I_{s-r} \end{bmatrix},$$

显然，Q是 s 阶可逆阵，且有

$$BQ=[B_1 \mid B_2]\left[\begin{array}{c|c} I_r & -G \\ \hline O & I_{s-r} \end{array}\right]$$

$$=[B_1 \mid B_2-B_1G]=[b_1\ b_2 \cdots b_r \mid O],$$

其中向量组 $b_1, b_2, \cdots, b_r$ 是线性无关的.于是

$$ABQ=[Ab_1 \quad Ab_2 \quad \cdots \quad Ab_r \mid O],$$

且 $Ab_1, Ab_2, \cdots, Ab_r$ 中有rank(AB)个向量是线性无关的，因而在这些向量中至多有r-rank(AB)个零向量.但齐次线性方程组 $AX=0$ 有n-rankA个线性无关的解.因此

$$r-rank(AB)=rankB-rank(AB)\le n-rankA,$$

即前一个不等式成立.

例7　设 $A_i(i=1,2,\cdots,k)$ 都是 n 阶方阵。证明当 $A_1A_2\cdots A_k=O$

时有下述不等式：

$$\sum_{i=1}^{k}(n-rankA_i)\geq n.$$

即

$$\sum_{i=1}^{k}rankA_i\leq(k-1)n, \qquad (2.1\text{-}11)$$

证　由于这两个不等式是等价的，所以只需证明其中的一个即可。

根据（2.1-10）式，有

$$
\begin{aligned}
0 = rank(A_1 A_2 \cdots A_k) &\geq rank(A_1 A_2 \cdots A_{k-1}) + rankA_k - n \\
&\geq rank(A_1 A_2 \cdots A_{k-2}) + rankA_{k-1} + rankA_k - 2n \\
&\geq \cdots \geq \sum_{i-1}^{k} rankA_i - (k-1)n,
\end{aligned}
$$

即
$$\sum_{i-1}^{k} rankA_i \leq (k-1)n$$

类似于§1.2中线性变换的特征子空间，我们也可以定义方阵*A*的特征子空间。

设 λ_0 是n阶方阵A的特征值，则集

$$E_{\lambda_0} \triangleq \{x \mid (A-\lambda_0 I)x = 0\}$$

称为A关于 λ_0 的特征子空间。显然，特征子空间

$$E_{\lambda_0} = N(A-\lambda_0 I)$$

是由A关于 λ_0 的所有特征向量，再添加零向量所组成的。

$\dim E_{\lambda_0} = n - rank(A-\lambda_0 I)$ 称为 λ_0 的几何重数，且这个数就是A 关于 λ_0 的线性无关特征向量的最大个数，并且 λ_0 的几何重数不大于其代数重数。

定理 n阶方阵A可对角化（即相似于对角矩阵）的充分必要条件是，A的最小多项式没有重根。

证 必要性.由于对角矩阵可以看作是1阶方阵所组成的对角块矩阵，而1阶方阵的最小多项式是一次多项式，所以A的最小多项式没有重根.

充分性.设$m_A(\lambda)=(\lambda-\lambda_1)(\lambda-\lambda_2)\cdots(\lambda-\lambda_s)$，则有

$$(A-\lambda_1 I)(A-\lambda_2 I)\cdots(A-\lambda_s I)=O,$$

故有例7知$\sum_{i=1}^{s}[n-rank(A-\lambda_i I)]\geq n$，即

$$\dim E_{\lambda_1}+\dim E_{\lambda_2}+\cdots+E_{\lambda_s}\geq n.$$

但另一方面又有 $\dim E_{\lambda_i} \le n_i (1 \le i \le s)$，故

$$\dim E_{\lambda_1} + \dim E_{\lambda_2} + \cdots + \dim E_{\lambda_s} \le n_1 + n_2 + \cdots + n_s = n,$$

这里 n_i 是 λ_i 的代数重数.因此

$$\dim E_{\lambda_i} = n_i \quad (1 \le i \le s),$$

即A的任一特征值 λ_i 的几何重数等于 λ_i 的代数重数，故A可对角化.

例8 设n阶方阵A满足关系式

$$A^3-3A^2-A+3I=O$$

证明A必可对角化。

解 由题设可知 $g(\lambda)=\lambda^3-3\lambda^2-\lambda+3$ 是A的零化多项式，

又因 $g(\lambda)=(\lambda+1)(\lambda-1)(\lambda-3)$ 没有重根 $m_A(\lambda)\mid g(\lambda)$，

故 $m_A(\lambda)$ 没有重根。因此，A必可对角化。