§4 Hamilton-Cayley 定理

一、Hamilton-Cayley 定理

定理 设 $\boldsymbol{A}\in\mathbf{C}^{n\times n}$ 的特征多项式为 $\varphi(\lambda)=\det(\lambda\boldsymbol{I}-\boldsymbol{A})$，则 $\varphi(\boldsymbol{A})=\boldsymbol{O}$，即方阵 $\boldsymbol{A}$ 是其特征多项式的根。

证 设 $\boldsymbol{A}$ 的 n 个特征值为 $\lambda_1,\lambda_2,\cdots,\lambda_n$（可以有相同的），则有

$$\varphi(\lambda)=(\lambda-\lambda_1)(\lambda-\lambda_2)\cdots(\lambda-\lambda_n)$$

又存在 n 阶可逆阵 $\boldsymbol{P}$，使 $\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}=\boldsymbol{J}=\begin{pmatrix}\lambda_1 & * & & \\ & \lambda_2 & \ddots & \\ & & \ddots & * \\ & & & \lambda_n\end{pmatrix}$，其中*代表 1 或 0，

于是

$$\begin{aligned}\varphi(\boldsymbol{A})&=(\boldsymbol{A}-\lambda_1\boldsymbol{I})(\boldsymbol{A}-\lambda_2\boldsymbol{I})\cdots(\boldsymbol{A}-\lambda_n\boldsymbol{I})\\&=(\boldsymbol{PJP}^{-1}-\lambda_1\boldsymbol{I})(\boldsymbol{PJP}^{-1}-\lambda_2\boldsymbol{I})\cdots(\boldsymbol{PJP}^{-1}-\lambda_n\boldsymbol{I})\\&=\boldsymbol{P}(\boldsymbol{J}-\lambda_1\boldsymbol{I})(\boldsymbol{J}-\lambda_2\boldsymbol{I})\cdots(\boldsymbol{J}-\lambda_n\boldsymbol{I})\boldsymbol{P}^{-1}\end{aligned}$$

由于 $(\boldsymbol{J}-\lambda_1\boldsymbol{I})(\boldsymbol{J}-\lambda_2\boldsymbol{I})\cdots(\boldsymbol{J}-\lambda_n\boldsymbol{I})$

$$=\begin{pmatrix}0 & * & & \\ & \lambda_2-\lambda_1 & \ddots & \\ & & \ddots & * \\ & & & \lambda_n-\lambda_1\end{pmatrix}\begin{pmatrix}\lambda_1-\lambda_2 & * & & \\ & 0 & \ddots & \\ & & \ddots & * \\ & & & \lambda_n-\lambda_2\end{pmatrix}\cdots\begin{pmatrix}\lambda_1-\lambda_n & * & & \\ & \ddots & \ddots & \\ & & \lambda_{n-1}-\lambda_n & * \\ & & & 0\end{pmatrix}$$

$$=\begin{pmatrix}0 & 0 & * & \cdots & * \\ 0 & 0 & * & \cdots & * \\ \vdots & \vdots & & \ddots & \vdots \\ 0 & 0 & & & *\end{pmatrix}\cdots\begin{pmatrix}\lambda_1-\lambda_n & * & & \\ & \ddots & \ddots & \\ & & \lambda_{n-1}-\lambda_n & * \\ & & & 0\end{pmatrix}=\boldsymbol{O}$$

故 $\varphi(\boldsymbol{A})=\boldsymbol{O}$。**证毕**

二、Hamilton-Cayley 定理的应用

1．求矩阵多项式

例 已知 $\boldsymbol{A}=\begin{pmatrix}2 & -1 & -2\\ -1 & 2 & 2\\ 0 & 0 & 1\end{pmatrix}$，求

1）$g(\boldsymbol{A})=\boldsymbol{A}^8-9\boldsymbol{A}^6+\boldsymbol{A}^4-3\boldsymbol{A}^3+4\boldsymbol{A}^2+\boldsymbol{I}$； 2）$\boldsymbol{A}^{100}$。

解 1） $\boldsymbol{A}$ 的特征多项式为

$$\varphi(\lambda)=\det(\lambda\boldsymbol{I}-\boldsymbol{A})=(\lambda-1)^2(\lambda-3)=\lambda^3-5\lambda^2+7\lambda-3$$

设 $g(\lambda)=\lambda^8-9\lambda^6+\lambda^4-3\lambda^3+4\lambda^2+1$。 将 $g(\lambda)$ 用 $\varphi(\lambda)$ 除可得等式

$$g(\lambda)=(\lambda^5+5\lambda^4+9\lambda^3+13\lambda^2+18\lambda+23)\,\varphi(\lambda)+32\lambda^2-107\lambda+\cancel{68}\ 70$$

由于 $\varphi(\boldsymbol{A})=\boldsymbol{O}$， 于是

$$g(\boldsymbol{A})=32\boldsymbol{A}^2-107\boldsymbol{A}+68\boldsymbol{I}=\begin{pmatrix}14 & -21 & -42\\ -21 & 14 & 42\\ 0 & 0 & -7\end{pmatrix}\quad 22$$

2）用待定系数法

设 $$\lambda^{100}=q(\lambda)\varphi(\lambda)+b_2\lambda^2+b_1\lambda+b_0$$

需求出 b_2,b_1,b_0。注意 $\varphi(\lambda)$ 满足： $\varphi(3)=\varphi(1)=\varphi'(1)=0$ ；又对上式求导得

$$100\lambda^{99}=q'(\lambda)\varphi(\lambda)+q(\lambda)\varphi'(\lambda)+2b_2\lambda+b_1$$

于是有 $\begin{cases}3^{100}=9b_2+3b_1+b_0\\ 1=b_2+b_1+b_0\\ 100=2b_2+b_1\end{cases}$， 解得 $\begin{cases}b_0=\frac14(3^{100}-597)\\ b_1=\frac12(401-3^{100})\\ b_2=\frac14(3^{100}-201)\end{cases}$

故 $$\boldsymbol{A}^{100}=b_2\boldsymbol{A}^2+b_1\boldsymbol{A}+b_0\boldsymbol{I}=\begin{pmatrix}\frac12(3^{100}+1) & -\frac12(3^{100}-1) & -(3^{100}-1)\\ -\frac12(3^{100}-1) & \frac12(3^{100}+1) & 3^{100}-1\\ 0 & 0 & 1\end{pmatrix}$$

例 设 3 阶方阵 $\boldsymbol{A}$ 的特征值为 1，–1，2，试将 $\boldsymbol{A}^{2n}$ 表为 $\boldsymbol{A}$ 的二次多项式。

解 $\boldsymbol{A}$ 的特征多项式为 $\varphi(\lambda)=\det(\lambda\boldsymbol{I}-\boldsymbol{A})=(\lambda-1)(\lambda+1)(\lambda-2)$。

令 $\lambda^{2n}=q(\lambda)\varphi(\lambda)+b_2\lambda^2+b_1\lambda+b_0$ ， 将 λ=1，–1，2 依次代入得

$$\begin{cases}1=b_2+b_1+b_0\\ 1=b_2-b_1+b_0\\ 2^{2n}=4b_2+2b_1+b_0\end{cases}，\quad 解得\quad \begin{cases}b_0=\frac13(4-2^{2n})\\ b_1=0\\ b_2=\frac13(2^{2n}-1)\end{cases}$$

因此 $$\boldsymbol{A}^{2n}=\tfrac13(2^{2n}-1)\,\boldsymbol{A}^2+\tfrac13(4-2^{2n})\,\boldsymbol{I}$$

2．求方阵的逆

设 $\boldsymbol{A}\in\mathbf{C}^{n\times n}$ 的特征多项式为 $\varphi(\lambda)=\det(\lambda\boldsymbol{I}-\boldsymbol{A})=\lambda^n+a_{n-1}\lambda^{n-1}+\cdots+a_1\lambda+a_0$，

由于 $a_0=\varphi(0)=\det(-\boldsymbol{A})=(-1)^n\det\boldsymbol{A}$， 可见 $\det\boldsymbol{A}\neq0\Leftrightarrow a_0\neq0$。当 $\boldsymbol{A}$ 可逆时，

由 $$\boldsymbol{O}=\varphi(\boldsymbol{A})=\boldsymbol{A}^n+a_{n-1}\boldsymbol{A}^{n-1}+\cdots+a_1\boldsymbol{A}+a_0\boldsymbol{I}$$

得 $$\boldsymbol{A}[-\frac{1}{a_0}(\boldsymbol{A}^{n-1}+a_{n-1}\boldsymbol{A}^{n-2}+\cdots+a_1\boldsymbol{I})]=\boldsymbol{I}$$

故 $$\boldsymbol{A}^{-1}=-\frac{1}{a_0}(\boldsymbol{A}^{n-1}+a_{n-1}\boldsymbol{A}^{n-2}+\cdots+a_1\boldsymbol{I})$$

例 已知 $\boldsymbol{A}=\begin{pmatrix}-1 & 1 & 0\\ -4 & 3 & 0\\ 1 & 0 & 2\end{pmatrix}$，试求 $\boldsymbol{A}^{-1}$。

解 因为 $\varphi(\lambda)=\det(\lambda\boldsymbol{I}-\boldsymbol{A})=(\lambda-2)(\lambda-1)^2=\lambda^3-4\lambda^2+5\lambda-2$，

所以 $$\boldsymbol{A}^{-1}=-\frac{1}{-2}(\boldsymbol{A}^2-4\boldsymbol{A}+5\boldsymbol{I})=\frac{1}{2}\begin{pmatrix}6 & -2 & 0\\ 8 & -2 & 0\\ -3 & 1 & 1\end{pmatrix}$$

三、最小多项式

1．定义与性质

定义 设 $\boldsymbol{A}\in\mathbf{C}^{n\times n}$， $f(\lambda)$ 是多项式，如果有 $f(\boldsymbol{A})=\boldsymbol{O}$，则称 $f(\lambda)$ 是 $\boldsymbol{A}$ 的**零化多项式**。

问题：是否存在比 $\boldsymbol{A}$ 的特征多项式次数更低的零化多项式？

定义 在矩阵 $\boldsymbol{A}$ 的零化多项式中，次数最低的首一多项式称为 $\boldsymbol{A}$ 的**最小多项式**，记为 $m(\lambda)$（或 $m_{\boldsymbol{A}}(\lambda)$）。

性质 1 如果 $m_{\boldsymbol{A}}(\lambda)$ 是方阵 $\boldsymbol{A}$ 的最小多项式， $f(\lambda)$ 是 $\boldsymbol{A}$ 的任一零化多项式，则 $m_{\boldsymbol{A}}(\lambda)\mid f(\lambda)$，且 $m(\lambda)$ 是唯一的。

证 利用多项式带余除法知 $f(\lambda)=q(\lambda)m_{\boldsymbol{A}}(\lambda)+r(\lambda)$

其中余式 $r(\lambda)$ 的次数低于 $m_{\boldsymbol{A}}(\lambda)$ 的次数。如果 $r(\lambda)\neq 0$，则由

$$\boldsymbol{O}=f(\boldsymbol{A})=q(\boldsymbol{A})m_{\boldsymbol{A}}(\boldsymbol{A})+r(\boldsymbol{A})=r(\boldsymbol{A})$$

知 $r(\lambda)$ 是 $\boldsymbol{A}$ 的零化多项式，且其次数低于 $m_{\boldsymbol{A}}(\lambda)$，矛盾。故 $r(\lambda)\equiv 0$，即 $f(\lambda)=q(\lambda)m_{\boldsymbol{A}}(\lambda)$, 也即 $m_{\boldsymbol{A}}(\lambda)\mid f(\lambda)$。

再证唯一性。设 $m_1(\lambda)$ 和 $m_2(\lambda)$ 都是 $\boldsymbol{A}$ 的最小多项式，令

$g(\lambda)=m_1(\lambda)-m_2(\lambda)$，则$g(\lambda)$的次数低于$m_1(\lambda)$和$m_2(\lambda)$的次数。若 $g(\lambda)\neq 0$，则由

$$g(\boldsymbol{A})=m_1(\boldsymbol{A})-m_2(\boldsymbol{A})=\boldsymbol{O}$$

得出矛盾，故 $g(\lambda)\equiv 0$，即 $m_1(\lambda)=m_2(\lambda)$。**证毕**

性质 2 λ_0是$\boldsymbol{A}$的特征值的充要条件是，λ_0是$\boldsymbol{A}$的最小多项式$m_{\boldsymbol{A}}(\lambda)$的零点。

证 充分性 由性质 1 即得。

必要性 设λ_0对应的特征向量为$\boldsymbol{x}$，即$\boldsymbol{Ax}=\lambda_0\boldsymbol{x}$，则

$$\boldsymbol{0}=m_{\boldsymbol{A}}(\boldsymbol{A})\boldsymbol{x}=m_{\boldsymbol{A}}(\lambda_0)\boldsymbol{x}$$

由于 $\boldsymbol{x}\neq\boldsymbol{0}$，所以 $m_{\boldsymbol{A}}(\lambda_0)=0$，即 λ_0是$m_{\boldsymbol{A}}(\lambda)$的零点。**证毕**

推论 若n阶方阵$\boldsymbol{A}$的n个特征值互异，则$\boldsymbol{A}$的最小多项式就是特征多项式。

性质 3 相似矩阵有相同的最小多项式。

证 设$\boldsymbol{P}^{-1}\boldsymbol{AP}=\boldsymbol{B}$，又设 $m_{\boldsymbol{A}}(\lambda)$和$m_{\boldsymbol{B}}(\lambda)$分别是$\boldsymbol{A}$与$\boldsymbol{B}$的最小多项式，由

$\boldsymbol{P}^{-1}\boldsymbol{AP}=\boldsymbol{B}$得 $m_{\boldsymbol{B}}(\boldsymbol{B})=\boldsymbol{P}^{-1}m_{\boldsymbol{B}}(\boldsymbol{A})\boldsymbol{P}$，但 $m_{\boldsymbol{B}}(\boldsymbol{B})=\boldsymbol{O}$，所以 $m_{\boldsymbol{B}}(\boldsymbol{A})=\boldsymbol{O}$，

这表明$m_{\boldsymbol{B}}(\lambda)$是$\boldsymbol{A}$的零化多项式，从而 $m_{\boldsymbol{A}}(\lambda)\big|m_{\boldsymbol{B}}(\lambda)$（由性质 1）。

同理可证 $m_{\boldsymbol{A}}(\lambda)$是$\boldsymbol{B}$的零化多项式，从而 $m_{\boldsymbol{B}}(\lambda)\big|m_{\boldsymbol{A}}(\lambda)$。由它们的首项系数为 1 得 $m_{\boldsymbol{A}}(\lambda)=m_{\boldsymbol{B}}(\lambda)$。**证毕**

利用这一性质，可先对矩阵相似化简，再求其最小多项式。

2．最小多项式的求法

性质 1 和 2 给出了求最小多项式的一种方法——试探法。

例 试求下列矩阵的最小多项式

1） $\boldsymbol{A}=\begin{pmatrix}3 & -3 & 2\\ -1 & 5 & -2\\ -1 & 3 & 0\end{pmatrix}$；

解 $\boldsymbol{A}$的特征多项式为 $\varphi(\lambda)=\det(\lambda\boldsymbol{I}-\boldsymbol{A})=(\lambda-2)^2(\lambda-4)$。

$\varphi(\lambda)$的因式有： $\lambda-2$，$\lambda-4$，$(\lambda-2)^2$，（$\lambda-2$）（$\lambda-4$），$(\lambda-2)^2$（$\lambda-4$）

由性质 2，只需验证第 4 个因式。可知 $(\boldsymbol{A}-2\boldsymbol{I})(\boldsymbol{A}-4\boldsymbol{I})=\boldsymbol{O}$，故

$$m_A(\lambda)=(\lambda-2)(\lambda-4)$$

2) $\boldsymbol{B}=\begin{pmatrix}0&1&0&\cdots&0\\0&0&1&\cdots&0\\\vdots&\vdots&\vdots&&\vdots\\0&0&0&\cdots&1\\0&0&0&\cdots&0\end{pmatrix}_{n\times n}$。

解 $\boldsymbol{B}$ 的特征多项式为 $\varphi(\lambda)=\det(\lambda\boldsymbol{I}-\boldsymbol{B})=\lambda^n$，所以 $\varphi(\lambda)$ 的因式为

$$\lambda，\lambda^2，\cdots,\lambda^{n-1}，\lambda^n。$$

因为 $\boldsymbol{B}^{n-1}=\begin{pmatrix}0&\cdots&0&1\\&\ddots&&0\\&&\ddots&\vdots\\&&&0\end{pmatrix}\neq\boldsymbol{O}$，$\boldsymbol{B}^n=\boldsymbol{O}$，

故 $\boldsymbol{B}$ 的最小多项式为 $m_{\boldsymbol{B}}(\lambda)=\varphi(\lambda)=\lambda^n$

当矩阵 $\boldsymbol{A}$ 的阶数较高，且特征值的重数较大时，利用试探法计算的工作量较大。

定理 设 $\boldsymbol{A}\in\mathbf{C}^{n\times n}$，$\lambda_1,\lambda_2,\cdots,\lambda_s$ 是 $\boldsymbol{A}$ 的全部互异特征值，则

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

其中 n_i 是 $\boldsymbol{A}$ 的 Jordan 标准形中含特征值 λ_i 的 Jordan 块的最高阶数。

例 求下列矩阵的最小多项式

1) $\boldsymbol{A}=\begin{pmatrix}2&1&&&&&\\&2&&&&&\\&&2&1&&&\\&&&2&1&&\\&&&&2&&\\&&&&&3&&\\&&&&&&3&1\\&&&&&&&3\end{pmatrix}$； 2) $\boldsymbol{A}=\begin{pmatrix}3&-3&2\\-1&5&-2\\-1&3&0\end{pmatrix}$。

解 1) $m_A(\lambda)=(\lambda-2)^3(\lambda-3)^2$；

2) 因为 $\det(\lambda\boldsymbol{I}-\boldsymbol{A})=(\lambda-2)^2(\lambda-4)$，所以 $\boldsymbol{A}$ 的特征值为 $\lambda_1=\lambda_2=2,\ \lambda_3=4$。

对应$\lambda=2$有两个线性无关的特征向量 $(3,1,0)^{\mathrm{T}}$， $(-2,0,1)^{\mathrm{T}}$，从而$\boldsymbol{A}$的 Jordan 标准形为

$$\boldsymbol{J}=\begin{pmatrix}2 & & \\ & 2 & \\ & & 4\end{pmatrix}$$

故 $m_{\boldsymbol{A}}(\lambda)=(\lambda-2)(\lambda-4)$。

推论 n阶方阵$\boldsymbol{A}$与对角阵相似的充要条件是$\boldsymbol{A}$的最小多项式无重根。

定理 设$\boldsymbol{A}\in\mathbf{C}^{n\times n}$的特征多项式为$\varphi(\lambda)=\det(\lambda\boldsymbol{I}-\boldsymbol{A})$，又$D_{n-1}(\lambda)$是$\boldsymbol{A}$的$n-1$阶行列式因子，则 $m_{\boldsymbol{A}}(\lambda)=\dfrac{\varphi(\lambda)}{D_{n-1}(\lambda)}=d_n(\lambda)$。（证明略）

例 求下列矩阵的最小多项式

1） $\boldsymbol{A}=\begin{pmatrix}0 & 1 & 0 & \cdots & 0\\ 0 & 0 & 1 & \cdots & 0\\ \vdots & \vdots & \vdots & & \vdots\\ 0 & 0 & 0 & \cdots & 1\\ 0 & 0 & 0 & \cdots & 0\end{pmatrix}_{n\times n}$；

解 $\varphi(\lambda)=\det(\lambda\boldsymbol{I}-\boldsymbol{A})=\lambda^n$，但 $\lambda\boldsymbol{I}-\boldsymbol{A}=\begin{pmatrix}\lambda & -1 & & \\ & \lambda & \ddots & \\ & & \ddots & -1\\ & & & \lambda\end{pmatrix}$中右上角的$n-1$阶子式$\begin{vmatrix}-1 & & & \\ \lambda & \ddots & & \\ & \ddots & \ddots & \\ & & \lambda & -1\end{vmatrix}=(-1)^{n-1}\neq 0$， 故$D_{n-1}(\lambda)=1$，从而

$$m_{\boldsymbol{A}}(\lambda)=\frac{\varphi(\lambda)}{D_{n-1}(\lambda)}=\varphi(\lambda)=\lambda^n$$

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

解 $\varphi(\lambda)=\det(\lambda\boldsymbol{I}-\boldsymbol{A})=(\lambda-2)(\lambda-1)^2$，

但在 $\lambda\boldsymbol{I}-\boldsymbol{A}=\begin{pmatrix}\lambda+1 & -1 & 0\\ 4 & \lambda-3 & 0\\ -1 & 0 & \lambda-2\end{pmatrix}$ 中 1,3 行、1,2 列的二阶子式 $\begin{vmatrix}\lambda+1 & -1\\ -1 & 0\end{vmatrix}=-1$，

所以 $D_2(\lambda)=1$，从而 $m_{\boldsymbol{A}}(\lambda)=\varphi(\lambda)$。

这一方法的缺点是，求 $D_{n-1}(\lambda)$ 可能比较麻烦。