SS20. 凱雷-漢米爾頓定理

一般的 LTI 系統的數學模式除了可以利用 ODE-CC 來描述輸出-輸入的關係外，也可以使用狀態空間描述法(state-space description)，這種描述法已經在設計濾波器時介紹過，在數學上已經證明也可以將狀態空間描述法轉換為 ODE-CC，所採用的是凱雷-漢米爾頓原理(Cayley-Hamilton theorem)。

凱雷-漢米爾頓定理之敘述如下：

若方陣 $\boldsymbol{A}\in R^{n\times n}$ 之特徵方程式為

(1) $$\left|\lambda\boldsymbol{I}-\boldsymbol{A}\right|=\lambda^n+\alpha_{n-1}\lambda^{n-1}+\cdots+\alpha_1\lambda+\alpha_0=0$$

或者是

(2) $$\left|\lambda\boldsymbol{I}-\boldsymbol{A}\right|=\left(\lambda-\lambda_1\right)\left(\lambda-\lambda_2\right)\cdots\left(\lambda-\lambda_n\right)=0$$

其中 $\alpha_i\in R$ 為實數，i=1,2,…,n–1， $\lambda_k\in C$ 為特徵值，k=1,2,…,n，則下列之矩陣方程式必須成立：

(3) $$\boldsymbol{A}^n+\alpha_{n-1}\boldsymbol{A}^{n-1}+\cdots+\alpha_2\boldsymbol{A}^2+\alpha_1\boldsymbol{A}+\alpha_0\boldsymbol{I}=\boldsymbol{0}$$

或者是

(4) $$\left(\boldsymbol{A}-\lambda_1\boldsymbol{I}\right)\left(\boldsymbol{A}-\lambda_2\boldsymbol{I}\right)\cdots\left(\boldsymbol{A}-\lambda_n\boldsymbol{I}\right)=\boldsymbol{0}$$

這就是數學上所稱的凱雷-漢米爾頓定理。

為了易於了解，首先考慮最單純的情況：$\boldsymbol{A}\in R^{n\times n}$ 具有 n 個兩兩相互獨立之特徵向量。例如 $\boldsymbol{A}\in R^{4\times 4}$ 具有 4 個特徵值，可能是單根或重根，也可能是實數或共軛複數，不論這些特徵值是何種情形，若存在 4 個獨立的特徵向量 $\boldsymbol{v}_k$ 滿足

(5) $$\boldsymbol{A}\boldsymbol{v}_k=\lambda_k\boldsymbol{v}_k,\quad k=1,2,3,4$$

亦即

(6) $$\left(\boldsymbol{A}-\lambda_k\boldsymbol{I}\right)\boldsymbol{v}_k=\boldsymbol{0},\quad k=1,2,3,4$$

其中 $\boldsymbol{v}_k\in C^4$ 所對應的特徵值為 $\lambda_k\in C$ ，則 $V=\left\{\boldsymbol{v}_1\ \boldsymbol{v}_2\ \boldsymbol{v}_3\ \boldsymbol{v}_4\right\}$ 可構成的一組基底，並組合成特徵向量矩陣 $\boldsymbol{V}=\left[\boldsymbol{v}_1\ \boldsymbol{v}_2\ \boldsymbol{v}_3\ \boldsymbol{v}_4\right]$，其反矩陣 $\boldsymbol{V}^{-1}$ 亦存在，再根據(1)與(2)兩式可知

(7) $$\lambda^4+\alpha_3\lambda^3+\alpha_2\lambda^2+\alpha_1\lambda+\alpha_0=\left(\lambda-\lambda_1\right)\left(\lambda-\lambda_2\right)\left(\lambda-\lambda_3\right)\left(\lambda-\lambda_4\right)$$

若將此式之λ以 $\boldsymbol{A}$ 取代，則可得

(8) $$\begin{aligned}&\boldsymbol{A}^4+\alpha_3\boldsymbol{A}^3+\alpha_2\boldsymbol{A}^2+\alpha_1\boldsymbol{A}+\alpha_0\boldsymbol{I}\\&=(\boldsymbol{A}-\lambda_1\boldsymbol{I})(\boldsymbol{A}-\lambda_2\boldsymbol{I})(\boldsymbol{A}-\lambda_3\boldsymbol{I})(\boldsymbol{A}-\lambda_4\boldsymbol{I})\end{aligned}$$

接著乘上$\boldsymbol{v}_k$，k=1,2,3,4，其表示式如下：

(9) $$\begin{aligned}&\left(\boldsymbol{A}^4+\alpha_3\boldsymbol{A}^3+\alpha_2\boldsymbol{A}^2+\alpha_1\boldsymbol{A}+\alpha_0\boldsymbol{I}\right)\boldsymbol{v}_k\\&=(\boldsymbol{A}-\lambda_1\boldsymbol{I})(\boldsymbol{A}-\lambda_2\boldsymbol{I})(\boldsymbol{A}-\lambda_3\boldsymbol{I})(\boldsymbol{A}-\lambda_4\boldsymbol{I})\boldsymbol{v}_k\end{aligned}$$

在此式中，若 k=4，則

(10) $$\begin{aligned}&\left(\boldsymbol{A}^4+\alpha_3\boldsymbol{A}^3+\alpha_2\boldsymbol{A}^2+\alpha_1\boldsymbol{A}+\alpha_0\boldsymbol{I}\right)\boldsymbol{v}_4\\&=(\boldsymbol{A}-\lambda_1\boldsymbol{I})(\boldsymbol{A}-\lambda_2\boldsymbol{I})(\boldsymbol{A}-\lambda_3\boldsymbol{I})\underline{(\boldsymbol{A}-\lambda_4\boldsymbol{I})\boldsymbol{v}_4}\end{aligned}$$

再由(6)式可知$(\boldsymbol{A}-\lambda_4\boldsymbol{I})\boldsymbol{v}_4=\boldsymbol{0}$，故(10)式可化為

(11) $$\left(\boldsymbol{A}^4+\alpha_3\boldsymbol{A}^3+\alpha_2\boldsymbol{A}^2+\alpha_1\boldsymbol{A}+\alpha_0\boldsymbol{I}\right)\boldsymbol{v}_4=\boldsymbol{0}$$

同樣在(9)式中，若 k=3，則

(12) $$\begin{aligned}&\left(\boldsymbol{A}^4+\alpha_3\boldsymbol{A}^3+\alpha_2\boldsymbol{A}^2+\alpha_1\boldsymbol{A}+\alpha_0\boldsymbol{I}\right)\boldsymbol{v}_3\\&=(\boldsymbol{A}-\lambda_1\boldsymbol{I})(\boldsymbol{A}-\lambda_2\boldsymbol{I})(\boldsymbol{A}-\lambda_3\boldsymbol{I})(\boldsymbol{A}-\lambda_4\boldsymbol{I})\boldsymbol{v}_3\\&=(\boldsymbol{A}-\lambda_1\boldsymbol{I})(\boldsymbol{A}-\lambda_2\boldsymbol{I})(\boldsymbol{A}-\lambda_4\boldsymbol{I})\underline{(\boldsymbol{A}-\lambda_3\boldsymbol{I})\boldsymbol{v}_3}\end{aligned}$$

此式利用到$(\boldsymbol{A}-\lambda_3\boldsymbol{I})(\boldsymbol{A}-\lambda_4\boldsymbol{I})=(\boldsymbol{A}-\lambda_4\boldsymbol{I})(\boldsymbol{A}-\lambda_3\boldsymbol{I})$之事實，再由(6)式可將(12)式化為

(13) $$\left(\boldsymbol{A}^4+\alpha_3\boldsymbol{A}^3+\alpha_2\boldsymbol{A}^2+\alpha_1\boldsymbol{A}+\alpha_0\boldsymbol{I}\right)\boldsymbol{v}_3=\boldsymbol{0}$$

依此類推，最後可得

(14) $$\left(\boldsymbol{A}^4+\alpha_3\boldsymbol{A}^3+\alpha_2\boldsymbol{A}^2+\alpha_1\boldsymbol{A}+\alpha_0\boldsymbol{I}\right)\boldsymbol{v}_k=\boldsymbol{0},\quad k=1,2,3,4$$

亦即

(15) $$\left(\boldsymbol{A}^4+\alpha_3\boldsymbol{A}^3+\alpha_2\boldsymbol{A}^2+\alpha_1\boldsymbol{A}+\alpha_0\boldsymbol{I}\right)\underbrace{[\boldsymbol{v}_1\ \boldsymbol{v}_2\ \boldsymbol{v}_3\ \boldsymbol{v}_4]}_{\boldsymbol{V}}=\boldsymbol{0}$$

此式乘上$\boldsymbol{V}^{-1}$後成為

(16) $$\boldsymbol{A}^4+\alpha_3\boldsymbol{A}^3+\alpha_2\boldsymbol{A}^2+\alpha_1\boldsymbol{A}+\alpha_0\boldsymbol{I}=\boldsymbol{0}$$

故符合(3)式之凱雷-漢米爾頓定理。

另外還有一個重要的觀念與凱雷-漢米爾頓定理有關，即最小多項式(minimal polynomial)，為了方便說明，還是考慮$\boldsymbol{A} \in R^{4\times 4}$，所不同的是此方陣只存在 2 個特徵值$\lambda_1$與$\lambda_2$，其中$\lambda_2$為三重根，在此情況下，可能存在下列三種情況：

(A) 三重根λ_2對應到三個特徵向量
(B) 三重根λ_2對應到兩個特徵向量
(C) 三重根λ_2對應到一個特徵向量

首先探討(A)之情況，滿足$\boldsymbol{A}\boldsymbol{v}_1 = \lambda_1 \boldsymbol{v}_1$，$\boldsymbol{A}\boldsymbol{v}_i = \lambda_2 \boldsymbol{v}_i$，$i$ =2,3,4，亦即

(17) $$(\boldsymbol{A} - \lambda_1 \boldsymbol{I})\boldsymbol{v}_1 = \boldsymbol{0}$$
(18) $$(\boldsymbol{A} - \lambda_2 \boldsymbol{I})\boldsymbol{v}_i = \boldsymbol{0}, \quad i = 2,3,4$$

其中λ_2對應到三個特徵向量$\boldsymbol{v}_2$、$\boldsymbol{v}_3$與$\boldsymbol{v}_4$，令特徵向量矩陣為$\boldsymbol{V} = [\boldsymbol{v}_1\ \boldsymbol{v}_2\ \boldsymbol{v}_3\ \boldsymbol{v}_4]$，且存在反矩陣$\boldsymbol{V}^{-1}$。由(17)與(18)兩式可知

(19) $$(\boldsymbol{A} - \lambda_1 \boldsymbol{I})(\boldsymbol{A} - \lambda_2 \boldsymbol{I})\boldsymbol{v}_i = (\boldsymbol{A}^2 + \alpha_1' \boldsymbol{A} + \alpha_0' \boldsymbol{I})\boldsymbol{v}_i = \boldsymbol{0}, \quad i = 1,2,3,4$$

亦即

(20) $$(\boldsymbol{A}^2 + \alpha_1' \boldsymbol{A} + \alpha_0' \boldsymbol{I})[\boldsymbol{v}_1\ \boldsymbol{v}_2\ \boldsymbol{v}_3\ \boldsymbol{v}_4] = (\boldsymbol{A}^2 + \alpha_1' \boldsymbol{A} + \alpha_0' \boldsymbol{I})\boldsymbol{V} = \boldsymbol{0}$$

其中$\alpha_1' = -(\lambda_1 + \lambda_2)$，$\alpha_0' = \lambda_1 \lambda_2$，此式乘上$\boldsymbol{V}^{-1}$後成為

(21) $$(\boldsymbol{A} - \lambda_1 \boldsymbol{I})(\boldsymbol{A} - \lambda_2 \boldsymbol{I}) = \boldsymbol{A}^2 + \alpha_1' \boldsymbol{A} + \alpha_0' \boldsymbol{I} = \boldsymbol{0}$$

故

(22) $$\begin{aligned}&(\boldsymbol{A}^2 + \alpha_1' \boldsymbol{A} + \alpha_0' \boldsymbol{I})(\boldsymbol{A} - \lambda_2 \boldsymbol{I})(\boldsymbol{A} - \lambda_2 \boldsymbol{I}) \\ &= \boldsymbol{A}^4 + \alpha_3 \boldsymbol{A}^3 + \alpha_2 \boldsymbol{A}^2 + \alpha_1 \boldsymbol{A} + \alpha_0 \boldsymbol{I} = \boldsymbol{0}\end{aligned}$$

滿足凱雷-漢米爾頓定理。

接著探討(B)之情況，在此情況下可求得 3 個特徵向量$\boldsymbol{v}_1$、$\boldsymbol{v}_2$、$\boldsymbol{v}_4$，滿足

(23) $$(\boldsymbol{A} - \lambda_1 \boldsymbol{I})\boldsymbol{v}_1 = \boldsymbol{0}$$
(24) $$(\boldsymbol{A} - \lambda_2 \boldsymbol{I})\boldsymbol{v}_2 = \boldsymbol{0}$$
(25) $$(\boldsymbol{A} - \lambda_2 \boldsymbol{I})\boldsymbol{v}_4 = \boldsymbol{0}$$

另外在數學上已經證明可找出$\bar{\boldsymbol{v}}_3$滿足下列條件：

(26) $$(\boldsymbol{A} - \lambda_2 \boldsymbol{I})\bar{\boldsymbol{v}}_3 = \boldsymbol{v}_4$$

由(25)與(26)兩式可知

(27) $\left(\boldsymbol{A}-\lambda_2\boldsymbol{I}\right)^2\bar{\boldsymbol{v}}_3=\left(\boldsymbol{A}-\lambda_2\boldsymbol{I}\right)\boldsymbol{v}_4=\boldsymbol{0}$

應注意的是，雖然$\bar{\boldsymbol{v}}_3$不是特徵向量，但$\boldsymbol{V}=\left[\boldsymbol{v}_1\ \boldsymbol{v}_2\ \bar{\boldsymbol{v}}_3\ \boldsymbol{v}_4\right]$確實存在反矩陣$\boldsymbol{V}^{-1}$，由(23)~(25)與(27)式可知

(28) $\left(\boldsymbol{A}-\lambda_1\boldsymbol{I}\right)\left(\boldsymbol{A}-\lambda_2\boldsymbol{I}\right)^2\bar{\boldsymbol{v}}_3=\boldsymbol{0}$

(29) $$\begin{aligned}\left(\boldsymbol{A}-\lambda_1\boldsymbol{I}\right)\left(\boldsymbol{A}-\lambda_2\boldsymbol{I}\right)^2\boldsymbol{v}_i&=\left(\boldsymbol{A}-\lambda_2\boldsymbol{I}\right)^2\left(\boldsymbol{A}-\lambda_1\boldsymbol{I}\right)\boldsymbol{v}_i\\&=\left(\boldsymbol{A}^3+\alpha_2'\boldsymbol{A}^2+\alpha_1'\boldsymbol{A}+\alpha_0'\boldsymbol{I}\right)\boldsymbol{v}_i=\boldsymbol{0},\quad i=1,2,4\end{aligned}$$

亦即

(30) $$\begin{aligned}&\left(\boldsymbol{A}^3+\alpha_2'\boldsymbol{A}^2+\alpha_1'\boldsymbol{A}+\alpha_0'\boldsymbol{I}\right)\left[\boldsymbol{v}_1\ \boldsymbol{v}_2\ \bar{\boldsymbol{v}}_3\ \boldsymbol{v}_4\right]\\&=\left(\boldsymbol{A}^3+\alpha_2'\boldsymbol{A}^2+\alpha_1'\boldsymbol{A}+\alpha_0'\boldsymbol{I}\right)\boldsymbol{V}=\boldsymbol{0}\end{aligned}$$

其中$\alpha_2'=-\left(\lambda_1+2\lambda_2\right)$，$\alpha_1'=2\lambda_1\lambda_2+\lambda_2^2$，$\alpha_0'=-\lambda_1\lambda_2^2$，此式乘上$\boldsymbol{V}^{-1}$後可得

(31) $\left(\boldsymbol{A}-\lambda_1\boldsymbol{I}\right)\left(\boldsymbol{A}-\lambda_2\boldsymbol{I}\right)^2=\boldsymbol{A}^3+\alpha_2'\boldsymbol{A}^2+\alpha_1'\boldsymbol{A}+\alpha_0'\boldsymbol{I}=\boldsymbol{0}$

故

(32) $$\begin{aligned}&\left(\boldsymbol{A}^3+\alpha_2'\boldsymbol{A}^2+\alpha_1'\boldsymbol{A}+\alpha_0'\boldsymbol{I}\right)\left(\boldsymbol{A}-\lambda_2\boldsymbol{I}\right)\\&=\boldsymbol{A}^4+\alpha_3\boldsymbol{A}^3+\alpha_2\boldsymbol{A}^2+\alpha_1\boldsymbol{A}+\alpha_0\boldsymbol{I}=\boldsymbol{0}\end{aligned}$$

仍然滿足凱雷-漢米爾頓定理。

最後探討(C)之情況，在此情況下可求得 2 個特徵向量$\boldsymbol{v}_1$與$\boldsymbol{v}_4$，滿足

(33) $\left(\boldsymbol{A}-\lambda_1\boldsymbol{I}\right)\boldsymbol{v}_1=\boldsymbol{0}$

(34) $\left(\boldsymbol{A}-\lambda_2\boldsymbol{I}\right)\boldsymbol{v}_4=\boldsymbol{0}$

並且可找出$\bar{\boldsymbol{v}}_2$與$\bar{\boldsymbol{v}}_3$滿足下列條件：

(35) $\left(\boldsymbol{A}-\lambda_2\boldsymbol{I}\right)\bar{\boldsymbol{v}}_3=\boldsymbol{v}_4$

(36) $\left(\boldsymbol{A}-\lambda_2\boldsymbol{I}\right)\bar{\boldsymbol{v}}_2=\bar{\boldsymbol{v}}_3$

由(34)~(36)三式可知

(37) $\left(\boldsymbol{A}-\lambda_2\boldsymbol{I}\right)^3\bar{\boldsymbol{v}}_2=\left(\boldsymbol{A}-\lambda_2\boldsymbol{I}\right)^2\bar{\boldsymbol{v}}_3=\left(\boldsymbol{A}-\lambda_2\boldsymbol{I}\right)\boldsymbol{v}_4=\boldsymbol{0}$

同樣地，$\boldsymbol{V}=\left[\boldsymbol{v}_1\ \bar{\boldsymbol{v}}_2\ \bar{\boldsymbol{v}}_3\ \boldsymbol{v}_4\right]$存在反矩陣$\boldsymbol{V}^{-1}$，由(33)與(37)兩式可知

(38) $$\begin{aligned}&\left(\boldsymbol{A}-\lambda_1\boldsymbol{I}\right)\left(\boldsymbol{A}-\lambda_2\boldsymbol{I}\right)^3\left[\boldsymbol{v}_1\ \bar{\boldsymbol{v}}_2\ \bar{\boldsymbol{v}}_3\ \boldsymbol{v}_4\right]\\&=\left(\boldsymbol{A}^4+\alpha_3\boldsymbol{A}^3+\alpha_2\boldsymbol{A}^2+\alpha_1\boldsymbol{A}+\alpha_0\boldsymbol{I}\right)\boldsymbol{V}=\boldsymbol{0}\end{aligned}$$

此式乘上$\boldsymbol{V}^{-1}$後成為

(39) $$\boldsymbol{A}^4+\alpha_3\boldsymbol{A}^3+\alpha_2\boldsymbol{A}^2+\alpha_1\boldsymbol{A}+\alpha_0\boldsymbol{I}=\boldsymbol{0}$$

亦滿足凱雷-漢米爾頓定理。

根據以上之分析可知，方陣$\boldsymbol{A}\in\Re^{n\times n}$具有 n 個特徵值，這些特徵值可能有重根，所求得的特徵向量之個數也不相同，但特徵方程式都可表為

(40) $$\left|\lambda\boldsymbol{I}-\boldsymbol{A}\right|=\lambda^n+\alpha_{n-1}\lambda^{n-1}+\cdots+\alpha_2\lambda^2+\alpha_1\lambda+\alpha_0=0$$

若是令 $p(\lambda)=\left|\lambda\boldsymbol{I}-\boldsymbol{A}\right|$，則稱 $p(\lambda)$ 為特徵多項式(characteristic polynomial)，所相對應的矩陣多項式為

(51) $$p(\boldsymbol{A})=\boldsymbol{A}^n+\alpha_{n-1}\boldsymbol{A}^{n-1}+\cdots+\alpha_2\boldsymbol{A}^2+\alpha_1\boldsymbol{A}+\alpha_0\boldsymbol{I}$$

根據凱雷-漢米爾頓定理可表為

(52) $$p(\boldsymbol{A})=\boldsymbol{A}^n+\alpha_{n-1}\boldsymbol{A}^{n-1}+\cdots+\alpha_2\boldsymbol{A}^2+\alpha_1\boldsymbol{A}+\alpha_0\boldsymbol{I}=\boldsymbol{0}$$

若特徵值λ是 k 重根，$k\geq 2$，且可找到的相對應特徵向量個數大於 1，則存在最小次數的多項式

(53) $$p_{min}(\lambda)=\lambda^m+\alpha'_{n-1}\lambda^{m-1}+\cdots+\alpha'_2\lambda^2+\alpha'_1\lambda+\alpha'_0$$

其中 $m<n$，使得

(54) $$p_{min}(\boldsymbol{A})=\boldsymbol{A}^m+\alpha'_{n-1}\boldsymbol{A}^{m-1}+\cdots+\alpha'_2\boldsymbol{A}^2+\alpha'_1\boldsymbol{A}+\alpha'_0\boldsymbol{I}=\boldsymbol{0}$$

在數學上稱 $p_{min}(\lambda)$ 為 $\boldsymbol{A}\in\Re^{n\times n}$ 之最小多項式。例如情況(A)與(B)中，三重根λ_2所對應到特徵向量分別為 3 與 2，兩者之個數都大於 1，因此都具有最小多項式，分別為

(55) $$p_{min}=(\lambda-\lambda_1)(\lambda-\lambda_2)=\lambda^2+\alpha'_1\lambda+\alpha'_0$$

(56) $$p_{min}=(\lambda-\lambda_1)(\lambda-\lambda_2)^2=\lambda^3+\alpha'_2\lambda^2+\alpha'_1\lambda+\alpha'_0$$

至於情況(C)，三重根λ_2只對應到 1 個特徵向量，可知 $p_{min}(\lambda)$ 即特徵方程式，亦即 $p_{min}(\lambda)=\left|\lambda\boldsymbol{I}-\boldsymbol{A}\right|$。

例題： 若 $\boldsymbol{A}=\begin{bmatrix} 3 & 1 & 1 \\ 3 & 1 & 3 \\ -1 & -1 & 1 \end{bmatrix}$，則 $\boldsymbol{A}$ 的最小多項式為何？

解： 先求取特徵方程式如下：

$$p(\lambda)=|\lambda \boldsymbol{I}-\boldsymbol{A}|=\begin{vmatrix} \lambda-3 & -1 & -1 \\ -3 & \lambda-1 & -3 \\ 1 & 1 & \lambda-1 \end{vmatrix}=0$$

亦即

$$p(\lambda)=\lambda^3-5\lambda^2+8\lambda-4=(\lambda-1)(\lambda-2)^2=0$$

令 λ_2=2的特徵向量 $\boldsymbol{v}$，則 $(2\boldsymbol{I}-\boldsymbol{A})\boldsymbol{v}=\boldsymbol{0}$ ，亦即

$$\begin{bmatrix} -1 & -1 & -1 \\ -3 & 1 & -3 \\ 1 & 1 & 1 \end{bmatrix}\begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix}=\begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

其簡約列階梯型式為

$$\begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}\begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix}=\begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

故 $\boldsymbol{v}=[v_1 \quad v_2 \quad v_3]^T=[1 \quad 0 \quad -1]^T$ ，只找到 1 個特徵向量 $\boldsymbol{v}$，所以最小多項式即特徵方程式，如下所示：

$$p_{min}(\lambda)=p(\lambda)=\lambda^3-5\lambda^2+8\lambda-4$$