3.2. Characteristic Polynomial

前面提過一個 linear operator 的問題, 我們可以轉化成有關於 square matrix 的問題, 所以我們會先探討一般 $n \times n$ matrix, 然後再將之轉化成 linear operator 的情形.

給定一個係數在 F 的 polynomial $f(x) = c_d x^d + \cdots + c_1 x + c_0$ 以及一個 $n \times n$ matrix A, 我們定義

$$f(A) = c_d A^d + \cdots + c_1 A + c_0 I_n.$$

很明顯的, $f(A)$ 仍然是一個 $n \times n$ matrix. 一般來說矩陣相乘是不可交換的, 不過 A^i 和 $f(A)$ 相乘是可以交換的. 事實上

$$\begin{aligned} A^i \cdot f(A) &= A^i \cdot (c_d A^d + \cdots + c_1 A + c_0 I_n) \\ &= c_d A^{d+i} + \cdots + c_1 A^{1+i} + c_0 A^i = (c_d A^d + \cdots + c_1 A + c_0 I_n) \cdot A^i = f(A) \cdot A^i. \end{aligned}$$

因此加上利用矩陣加法乘法的分配律, 我們可以得到以下的結果.

Lemma 3.2.1. 假設 $f(x), g(x), h(x) \in F[x]$ 且 $f(x) = g(x)h(x)$. 若 $A \in M_n(F)$, 則

$$g(A) \cdot h(A) = h(A) \cdot g(A) = f(A).$$

再次強調這裡都是和 A 相關的矩陣相乘才會成立, 一般來說若 $g(x), h(x) \in F[x]$ 以及 $A, B \in M_n(F)$, 不一定會有 $g(A) \cdot h(B) = h(B) \cdot g(A)$.

接下來我們有興趣的是若 $A \sim B$, 是否 $f(A) \sim f(B)$ 呢? 首先觀察若 P 為 invertible, 則

$$(P^{-1} \cdot A \cdot P)^2 = (P^{-1} \cdot A \cdot P) \cdot (P^{-1} \cdot A \cdot P) = P^{-1} \cdot A^2 \cdot P.$$

利用數學歸納法可得

$$(P^{-1} \cdot A \cdot P)^i = P^{-1} \cdot A^i \cdot P.$$

我們有以下結果.

Lemma 3.2.2. 假設 $f(x) \in F[x]$ 且 $A, B \in M_n(F)$. 若 $A \sim B$, 則 $f(A) \sim f(B)$.

Proof. 由 $A \sim B$ 知存在 P 為 invertible 使得 $B = P^{-1} \cdot A \cdot P$. 若 $f(x) = c_d x^d + \cdots + c_1 x + c_0$, 則

$$\begin{aligned} & f(B) = c_d B^d + \cdots + c_1 B + c_0 I_n = c_d (P^{-1} \cdot A \cdot P)^d + \cdots + c_1 (P^{-1} \cdot A \cdot P) + c_0 I_n \\ &= c_d (P^{-1} \cdot A^d \cdot P) + \cdots + c_1 (P^{-1} \cdot A \cdot P) + c_0 I_n = P^{-1} \cdot (c_d A^d + \cdots + c_1 A + c_0 I_n) \cdot P = P^{-1} \cdot f(A) \cdot P, \end{aligned}$$

得證 $f(A) \sim f(B)$. $\square$

我們也可把這概念推廣到 linear operator, 假設 $f(x) = c_d x^d + \cdots + c_1 x + c_0 \in F[x]$ 以及 $T : V \to V$ 是一個 linear operator, 由於 linear operators 之間的合成和矩陣之間的相乘相對應 (參見式子 (3.1)), 我們定義

$$f(T) = c_d T^{\circ d} + \cdots + c_1 T + c_0 \mathrm{id},$$

很明顯的 $f(T)$ 仍然是 V 到 V 的 linear operator. 我們可以檢查 $T^{\circ i}\circ f(T)=f(T)\circ T^{\circ i}$, 所以一樣有以下結果.

Lemma 3.2.3. 假設 $f(x),g(x),h(x)\in F[x]$ 且 $f(x)=g(x)\cdot h(x)$. 若 $T\in\mathscr{L}(V)$, 則

$$g(T)\circ h(T)=h(T)\circ g(T)=f(T).$$

這裡要強調一下當 $f(x)=g(x)\cdot h(x)$ 時 $f(T)=g(T)\circ h(T)$ 而不是等於 $g(h(T))$. 也就是說將 $g(T)$ 和 $h(T)$ 這兩個 linear operator 合成會得到 $f(T)$ 這個 operator, 但並不是將 $h(T)$ 這個 linear operator 代入 $g(x)$ 這個多項式.

給定 V 的一個 ordered basis β 我們自然要問 $F(T)$ 的 representative matrix 是否和 T 的 representative matrix 有關. 事實上再次利用等式 3.1, 我們有 $[T^{\circ 2}]_\beta=[T]_\beta^2$, 利用數學歸納法可得

$$[T^{\circ i}]_\beta=[T\circ T^{\circ i-1}]_\beta=[T]_\beta\cdot[T]_\beta^{i-1}=[T]_\beta^i,$$

由此我們有以下之結果.

Lemma 3.2.4. 假設 V 是一個 *finite dimensional F-space,* β 為 V 的一個 *ordered basis* 且 $T:V\to V$ 是一個 *linear operator.* 若 $f(x)=c_dx^d+\cdots+c_1x+c_0\in F[x]$, 則

$$[f(T)]_\beta=f([T]_\beta)=c_d\,[T]_\beta^d+\cdots+c_1\,[T]_\beta+c_0\,I_n.$$

Proof. 依定義 $[f(T)]_\beta$ 是 $f(T)$ 的 representative matrix, 利用 Φ 是 linear transformation, 我們知

$$\begin{aligned}[f(T)]_\beta&=[c_d\,T^{\circ d}+\cdots+c_1\,T+c_0\,\mathrm{id}]_\beta=\\ &c_d\,[T^{\circ d}]_\beta+\cdots+c_1\,[T]_\beta+c_0\,[\mathrm{id}]_\beta=c_d\,[T]_\beta^d+\cdots+c_1\,[T]_\beta+c_0\,I_n=f([T]_\beta).\end{aligned}$$

□

現在回到 $n\times n$ matrix 的情形. 我們知 $\dim(M_n(F))=n^2$, 現若 $A\in M_n(F)$, 考慮 $S=\{I_n,A,A^2,\ldots,A^{n^2}\}$. 由於 $\#(S)=n^2+1>\dim(M_n(F))$, 我們知 S 為 linearly dependent. 亦即存在 $c_0,c_1,\ldots,c_{n^2}\in F$ 不全為 0 使得

$$c_{n^2}A^{n^2}+\cdots+c_1A+c_0I_n=\mathbf{O}.$$

若令 $f(x)=c_{n^2}x^{n^2}+\cdots+c_1x+c_0$, 則得 $f(A)=\mathbf{O}$. 因此我們可以說: 對任意的 $n\times n$ matrix A, 皆存在一個次數不大於 n^2 的非零多項式 $f(x)\in F[x]$ 使得 $f(A)$ 為 $n\times n$ 的 zero matrix $\mathbf{O}$. 注意這裡 c_{n^2} 有可能是 0 所以我們不能說 $\deg(f(x))=n^2$, 另外 $c_{n^2},\ldots,c_1,c_0$ 不全為 0, 所以 $f(x)$ 不是零多項式.

Question 3.2. 若 $A\sim B$ 且 $f(x)\in F[x]$ 滿足 $f(A)=\mathbf{O}$, 是否可得 $f(B)=\mathbf{O}$?

Question 3.3. 若 $\dim(V)=n$ 且 $T:V\to V$ 是一個 *linear operator,* 是否可找到一個 *nonzero polynomial* $f(x)\in F[x]$ 且 $\deg(f(x))\leq n^2$ 使得 $f(T)=\mathbf{O}$?

事實上我們可以找到次數為 n 的多項式 $f(x)$ 使得 $f(A)=\mathbf{O}$, 就是所謂的 characteristic polynomial.

Definition 3.2.5. 假設 $A\in M_n(F)$, 考慮 $\chi_A(x)=\det(xI_n-A)\in F[x]$, 稱為 A 的 *characteristic polynomial.*

注意有的書定義 $\det(A-xI_n)$ 為 A 的 characteristic polynomial, 我們用 $\det(xI_n-A)$ 主要是讓 $\chi_A(x)$ 是一個 monic polynomial (最高次項係數為 1). 利用降階求 determinant 的方法以及數學歸納法, 我們可以知當 A 為 $n\times n$ matrix 時, $\chi_A(x)$ 的次數為 n 且最高次項係數為 1. 也可更進一步得到 $\chi_A(x)$ 的次高項 (即 x^{n-1} 項) 係數為 $-\mathrm{tr}(A)$ (註: $\mathrm{tr}(A)$ 為 A 的 trace, 即對角線之和). 另外將 $x=0$ 代入 $\chi_A(x)$ 可得 $\chi_A(x)$ 的常數項為 $\chi_A(0)=\det(-A)=(-1)^n\det(A)$.

Example 3.2.6. 由於 $xI_n-I_n=(x-1)I_n$, 我們可得 $\chi_{I_n}(x)=\det((x-1)I_n)=(x-1)^n$. 我們計算幾個 2×2 matrix 的 characteristic polynomial. 考慮

$$A_1=\begin{pmatrix}1&-1\\1&-1\end{pmatrix}, A_2=\begin{pmatrix}1&-1\\0&-1\end{pmatrix}, A_3=\begin{pmatrix}1&-1\\2&-1\end{pmatrix},$$

則

$$\chi_{A_1}=\det\begin{pmatrix}x-1&1\\-1&x+1\end{pmatrix}=(x-1)(x+1)+1=x^2,$$

$$\chi_{A_2}=\det\begin{pmatrix}x-1&1\\0&x+1\end{pmatrix}=(x-1)(x+1)=x^2-1,$$

$$\chi_{A_3}=\det\begin{pmatrix}x-1&1\\-2&x+1\end{pmatrix}=(x-1)(x+1)+2=x^2+1.$$

Question 3.4. 試檢查看看 $\chi_{I_2}(I_2)$, $\chi_{A_1}(A_1)$, $\chi_{A_2}(A_2)$, $\chi_{A_3}(A_3)$ 是哪些矩陣.

接下來我們來看看 similar matrices 它們的 characteristic polynomial 有什麼關係.

Proposition 3.2.7. *若 $A,B\in M_n(F)$ 且 $A\sim B$, 則 $\chi_A(x)=\chi_B(x)$.*

Proof. 由 $A\sim B$ 知存在 invertible matrix P 使得 $B=P^{-1}\cdot A\cdot P$. 因 xI_n 為 diagonal matrix, 我們知 $xI_n\cdot P=P\cdot xI_n$, 故有 $P^{-1}\cdot xI_n\cdot P=xI_n$. 因此

$$xI_n-B=xI_n-P^{-1}\cdot A\cdot P=P^{-1}\cdot xI_n\cdot P-P^{-1}\cdot A\cdot P=P^{-1}\cdot(xI_n-A)\cdot P.$$

得證

$$\chi_B(x)=\det(xI_n-B)=\det(P^{-1}\cdot(xI_n-A)\cdot P)=\det(P)^{-1}\det(xI_n-A)\det(P)=\chi_A(x).$$

□

特別的, 當 $T:V\to V$ 是一個 linear operator, β,β' 為 V 的 ordered bases, 由於 $[T]_\beta\sim[T]_{\beta'}$, Proposition 3.2.7 告訴我們 $\chi_{[T]_\beta}(x)=\chi_{[T]_{\beta'}}(x)$. 因此我們可以定義 linear operator 的 characteristic polynomial.

Definition 3.2.8. 假設 V 為 finite dimensional F-space. 對於 V 的 linear operator $T:V\to V$, 任取 V 的一個 ordered basis β, 定義 T 的 characteristic polynomial 為 $\chi_{[T]_\beta}(x)$, 且以 $\chi_T(x)$ 來表示.

由於 A 的 characteristic polynomial 牽涉到 xI_n-A 這樣的矩陣, 也就是說矩陣的 entry 中有多項式, 現在我們來探討這一類的矩陣. 首先, 我們可以將這一類的矩陣寫成 $x^dA_d+\cdots+xA_1+A_0$, 其中 $A_i\in M_n(F)$ 這樣的型式. 例如我們可以有以下的表示法

$$\begin{pmatrix} 5x^2+3 & 4x-1 \\ 7 & x^3-2x^2+x \end{pmatrix} = x^3\begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} + x^2\begin{pmatrix} 5 & 0 \\ 0 & -2 \end{pmatrix} + x\begin{pmatrix} 0 & 4 \\ 0 & 1 \end{pmatrix} + \begin{pmatrix} 3 & -1 \\ 7 & 0 \end{pmatrix}.$$

由於我們是將 F 的元素代入 x, 所以我們可將 xA 視為常數 x 乘上矩陣 A. 因此當 $A,B\in M_n(F)$, 由矩陣乘法 $(rA)\cdot(sB)=(rs)A\cdot B$, 我們有

$$(x^iA)\cdot(x^jB)=x^{i+j}A\cdot B.$$

例如因矩陣加法乘法有分配律, 我們有

$$(A+xB)^2=(A+xB)\cdot(A+xB)=A^2+A\cdot(xB)+xB\cdot A+(xB)^2=A^2+x(A\cdot B+B\cdot A)+x^2B^2,$$

不過要注意因矩陣乘法沒有交換律, $(A+xB)^2$ 不一定等於 $A^2+2x(A\cdot B)+x^2B^2$.

當兩個 entry 中有多項式的 square matrices 相乘時, 我們可以它們如同一般的矩陣來相乘. 也可利用上面的方法將它們有 x 的部分提出, 然後像多項式相乘一樣展開. 由於這樣處理仍依循著矩陣乘法的規律, 所以得到的結果會相同. 我們看一個例子.

Example 3.2.9. 考慮

$$\begin{pmatrix} 5x^2+3 & 4x-1 \\ 7 & x \end{pmatrix} = x^2\begin{pmatrix} 5 & 0 \\ 0 & 0 \end{pmatrix} + x\begin{pmatrix} 0 & 4 \\ 0 & 1 \end{pmatrix} + \begin{pmatrix} 3 & -1 \\ 7 & 0 \end{pmatrix}.$$

以及

$$\begin{pmatrix} x-1 & 1 \\ -x & x+2 \end{pmatrix} = x\begin{pmatrix} 1 & 0 \\ -1 & 1 \end{pmatrix} + \begin{pmatrix} -1 & 1 \\ 0 & 2 \end{pmatrix}.$$

直接相乘我們有

$$\begin{pmatrix} 5x^2+3 & 4x-1 \\ 7 & x \end{pmatrix}\cdot\begin{pmatrix} x-1 & 1 \\ -x & x+2 \end{pmatrix} = \begin{pmatrix} 5x^3-9x^2+4x-3 & 9x^2+7x+1 \\ -x^2+7x-7 & x^2+2x+7 \end{pmatrix},$$

而另一邊如多項式相乘展開有

$$\begin{aligned}
&\left(x^2\begin{pmatrix} 5 & 0 \\ 0 & 0 \end{pmatrix} + x\begin{pmatrix} 0 & 4 \\ 0 & 1 \end{pmatrix} + \begin{pmatrix} 3 & -1 \\ 7 & 0 \end{pmatrix}\right)\cdot\left(x\begin{pmatrix} 1 & 0 \\ -1 & 1 \end{pmatrix} + \begin{pmatrix} -1 & 1 \\ 0 & 2 \end{pmatrix}\right) \\
&\quad = x^3\begin{pmatrix} 5 & 0 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ -1 & 1 \end{pmatrix} + x^2\left(\begin{pmatrix} 5 & 0 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} -1 & 1 \\ 0 & 2 \end{pmatrix} + \begin{pmatrix} 0 & 4 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ -1 & 1 \end{pmatrix}\right) \\
&\quad + x\left(\begin{pmatrix} 0 & 4 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} -1 & 1 \\ 0 & 2 \end{pmatrix} + \begin{pmatrix} 3 & -1 \\ 7 & 0 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ -1 & 1 \end{pmatrix}\right) + \begin{pmatrix} 3 & -1 \\ 7 & 0 \end{pmatrix}\begin{pmatrix} -1 & 1 \\ 0 & 2 \end{pmatrix} \\
&\qquad\qquad = x^3\begin{pmatrix} 5 & 0 \\ 0 & 0 \end{pmatrix} + x^2\begin{pmatrix} -9 & 9 \\ -1 & 1 \end{pmatrix} + x\begin{pmatrix} 4 & 7 \\ 7 & 2 \end{pmatrix} + \begin{pmatrix} -3 & 1 \\ -7 & 7 \end{pmatrix}
\end{aligned}$$

所以兩種算法結果是相等的.

接著我們要強調若 $x^dA_d+\cdots+xA_1+A_0=x^dB_d+\cdots+xB_1+B_0$, 其中 $A_i,B_i\in M_n(F)$, 則 $A_i=B_i$, $\forall i=0,1,\ldots,d$. 這是因為若有某個 $A_i\neq B_i$, 表示等式兩邊的矩陣有個 entry 其 x^i 的係數不相同, 造成矛盾. 了解了這些概念, 我們就可以處理 characteristic polynomial 的重要性質.

Theorem 3.2.10 (Cayley-Hamilton Theorem)**.** 若 $A\in M_n(F)$, $\chi_A(x)$ 為 A 的 *characteristic polynomial,* 則 $\chi_A(A)=\mathbf{O}$.

Proof. 令 $\chi_A(x)=x^n+\cdots+c_1x+c_0$. 利用 xI_n-A 的 adjoint matrix, 由 Lemma 3.1.5 我們有

$$\operatorname{adj}(xI_n-A)\cdot(xI_n-A)=\det(xI_n-A)I_n=\chi_A(x)I_n=x^nI_n+\cdots+xc_1I_n+c_0I_n.$$

若將 xI_n-A 的 i-th row 和 k-th column 移除, 所得的 $(n-1)\times(n-1)$ matrix 其 determinant 為次數小於 n 的多項式, 所以依 adjoint matrix 的定義 $\operatorname{adj}(A-xI_n)$ 的每個 entry 皆為次數小於 n 的多項式, 故假設 $\operatorname{adj}(A-xI_n)=x^{n-1}B_{n-1}+\cdots+xB_1+B_0$, 其中 $B_i\in M_n(F)$. 因此我們有以下的等式

$$(x^{n-1}B_{n-1}+x^{n-2}B_{n-2}+\cdots+xB_1+B_0)\cdot(xI_n-A)=x^nI_n+x^{n-1}c_{n-1}I_n+\cdots+xc_1I_n+c_0I_n \quad (3.2)$$

將等式 (3.2) 左邊展開, 我們得

$$\begin{aligned}&(x^{n-1}B_{n-1}+x^{n-2}B_{n-2}+\cdots+xB_1+B_0)\cdot(xI_n-A)\\&\qquad=x^n(B_{n-1}\cdot I_n)+x^{n-1}(B_{n-2}\cdot I_n-B_{n-1}\cdot A)+\cdots+x(B_0\cdot I_n-B_1\cdot A)-B_0\cdot A\end{aligned}$$

應該和等式 (3.2) 右式相同, 故比較係數得

$$\begin{aligned}-B_0\cdot A&=c_0I_n\\B_0\cdot I_n-B_1\cdot A&=c_1I_n\\&\vdots\\B_{n-2}\cdot I_n-B_{n-1}\cdot A&=c_{n-1}I_n\\B_{n-1}\cdot I_n&=I_n\end{aligned}$$

將第一式不動, 第二式兩邊右乘 A, 第三式兩邊右乘 A^2, $\ldots$, 最後一式兩邊右乘 A^n, 我們得

$$\begin{aligned}-B_0\cdot A&=c_0I_n\\B_0\cdot A-B_1\cdot A^2&=c_1A\\&\vdots\\B_{n-2}\cdot A^{n-1}-B_{n-1}\cdot A^n&=c_{n-1}A^{n-1}\\B_{n-1}\cdot A^n&=A^n\end{aligned}$$

因未左邊全部加起來會等於右邊全部加起來, 得證

$$\mathbf{O}=A^n+c_{n-1}A^{n-1}+\cdots+c_1A+c_0I_n=\chi_A(A).$$

□

當 β 為 V 的一個 ordered basis, $T : V \to V$ 為 linear operator, 我們定義 $\chi_T(x) = \chi_{[T]_\beta}(x)$. 此時 $\chi_T(T)$ 為 linear operator, 其對 β 的 representative matrix , 依 Lemma 3.2.4 知為

$$[\chi_{[T]_\beta}(T)]_\beta = \chi_{[T]_\beta}([T]_\beta).$$

故由 Theorem 3.2.10 知 $[\chi_T(T)]_\beta = \mathbf{O}$, 因此利用 Lemma 3.1.1 得知 $\chi_T(T) = \mathbf{O}$. 這就是 linear operator 版本的 Cayley-Hamilton Theorem.

Corollary 3.2.11 (Cayley-Hamilton Theorem)**.** 若 V 為 *finite dimensional F-space,* $T : V \to V$ 為 *linear operator,* 則 $\chi_T(T) = \mathbf{O}$.

3.3. Minimal Polynomial

若 A 是 $n \times n$ matrix, 利用 A 的 characteristic polynomial, 我們知道存在次數為 n 的多項式 $f(X) \in F[x]$ 使得 $f(A) = \mathbf{O}$. 會不會有次數更小的多項式可以達到這個目的呢? 這是有可能的, 例如 $A = I_n$ 時, $\chi_{I_n}(x) = (x-1)^n$, 但考慮 $f(x) = x - 1$, 我們有 $f(I_n) = I_n - I_n = \mathbf{O}$. 所以我們想要找到次數最小的非零多項式 $f(X) \in F[x]$ 使得 $f(A) = \mathbf{O}$.

Definition 3.3.1. 設 $A \in M_n(F)$, 在所有非零多項式 $f(x) \in F[x]$ 中滿足 $f(A) = \mathbf{O}$, 且次數最小的 monic polynomial (即最高次項係數為 1) 稱為 A 的 *minimal polynomial*, 用 $\mu_A(x)$ 來表示.

我們知道一定存在次數最小的非零多項式 $f(x) \in F[x]$ 使得 $f(A) = \mathbf{O}$, 而這裡要求 monic 就是要求唯一性. 事實上若 $f(x), g(x) \in F[x]$ 為次數最小的非零 monic polynomial 使得 $f(A) = g(A) = \mathbf{O}$, 因皆為次數最小故必有 $\deg(f) = \deg(g)$, 又要求 $f(x), g(x)$ 為 monic, 故知 $\deg(f(x) - g(x)) < \deg(f(x))$. 但此時 $f(A) - g(A) = \mathbf{O} - \mathbf{O} = \mathbf{O}$, 故由次數最小的要求知 $f(x) - g(x)$ 必為零多項式, 即 $f(x) = g(x)$, 所以 minimal polynomial $\mu_A(x)$ 是唯一的.

接下來我們要問若 $A \sim B$, 那麼它們的 minimal polynomial $\mu_A(x), \mu_B(x)$ 是否相等. 首先來看一個 minimal polynomial 最基本的性質.

Lemma 3.3.2. 假設 $A \in M_n(F)$ 且 $f(x) \in F[x]$. 則 $f(A) = \mathbf{O}$ 若且唯若 $\mu_A(x) \mid f(x)$.

Proof. 假設 $f(x) \mid \mu_A(x)$, 表示存在 $h(x) \in F[x]$ 使得 $f(x) = \mu_A(x)h(x)$, 因 $\mu_A(A) = \mathbf{O}$, 利用 Lemma 3.2.1 知 $f(A) = \mu_A(A) \cdot h(A) = \mathbf{O} \cdot h(A)$. 因為零矩陣乘以任何同階的矩陣亦為零矩陣, 故得 $f(A) = \mathbf{O}$.

另一方面, 因 F 是一個 field, 考慮多項式的除法 $f(x) = \mu_A(x)h(x) + r(x)$, 其中 $h(x), r(x) \in F[x]$ 且 $\deg(r(x)) < \deg(\mu_A(x))$. 由 $f(A) = \mathbf{O}$ 的假設我們得

$$\mathbf{O} = f(A) = \mu_A(A) \cdot h(A) + r(A) = \mathbf{O} \cdot h(A) + r(A) = r(A).$$

亦即 $r(x) \in F[x]$ 是一個次數比 $\mu_A(x)$ 小卻滿足 $r(A) = \mathbf{O}$ 的多項式. 依 $\mu_A(x)$ 是 A 的 minimal polynomial 之定義得 $r(x)$ 為零多項式, 得證 $f(x)$ 是 $\mu_A(x)$ 的倍式, 即 $\mu_A(x) \mid f(x)$. □

現若 $A \sim B$, 利用 Lemma 3.2.2 知 $\mu_A(B) \sim \mu_A(A) = \mathbf{O}$, 然而和零矩陣 similar 的矩陣必為零矩陣 (因對任意 invertible matrix P, $P^{-1} \cdot \mathbf{O} \cdot P = \mathbf{O}$), 故得 $\mu_A(B) = \mathbf{O}$. 由 Lemma 3.3.2 知 $\mu_B(x) \mid \mu_A(x)$. 同理利用 $\mu_B(A) \sim \mu_B(B) = \mathbf{O}$, 得 $\mu_A(x) \mid \mu_B(x)$. 然而 $\mu_A(x), \mu_B(x)$ 皆為 monic, 故得 $\mu_A(x) = \mu_B(x)$. 證得以下之結果.

Proposition 3.3.3. *若 $A, B \in M_n(F)$ 且 $A \sim B$, 則 $\mu_A(x) = \mu_B(x)$.*

我們也可以定一個 linear operator 的 minimal polynomial.

Definition 3.3.4. 設 V 為一個 finite dimensional F-space, $T : V \to V$ 為一個 linear operator. 在所有非零多項式 $f(x) \in F[x]$ 中滿足 $f(T) = \mathbf{O}$, 且次數最小的 monic polynomial 稱為 T 的 *minimal polynomial*, 用 $\mu_T(x)$ 來表示.

同 matrix 的情形, T 的 minimal polynomial 必存在且唯一. 利用 Lemma 3.3.2 相同的證明方法 (需用到零函數和任何函數合成仍為零函數) 我們會有以下結果.

Lemma 3.3.5. *假設 V 為一個 finite dimensional F-space, $T : V \to V$ 為一個 linear operator. 則 $f(T) = \mathbf{O}$ 若且唯若 $\mu_T(x) \mid f(x)$.*

Question 3.5. *你會證明 Lemma 3.3.5 嗎?*

當 β 為 V 的 ordered basis, T 的 characteristic polynomial $\chi_T(x)$ 是由 T 的 representative matrix $[T]_\beta$ 的 characteristic polynomial $\chi_{[T]_\beta}(x)$ 定義而得. 不過 T 的 minimal polynomial $\mu_T(x)$ 並不是由 $\mu_{[T]_\beta}$ 定義得到, 所以我們要探討它們是否相同.

Proposition 3.3.6. *設 V 為一個 finite dimensional F-space, β 為 V 的一個 ordered basis 且 $T : V \to V$ 為 linear operator. 則*

$$\mu_T(x) = \mu_{[T]_\beta}(x).$$

Proof. 首先注意, 若 $f(x) \in F[x]$, 則利用 Lemma 3.2.4 以及 Lemma 3.1.1 我們有

$$f(T) = \mathbf{O} \Leftrightarrow [f(T)]_\beta = \mathbf{O} \Leftrightarrow f([T]_\beta) = \mathbf{O}.$$

所以由 $\mu_T(T) = \mathbf{O}$ 可得 $\mu_T([T]_\beta) = \mathbf{O}$, 故由 Lemma 3.3.2 知 $\mu_{[T]_\beta}(x) \mid \mu_T(x)$. 同樣的由 $\mu_{[T]_\beta}([T]_\beta) = \mathbf{O}$, 可得 $\mu_{[T]_\beta}(T) = \mathbf{O}$, 故知 $\mu_T(x) \mid \mu_{[T]_\beta}(x)$. 又因 $\mu_T(x), \mu_{[T]_\beta}(x)$ 皆為 monic polynomial, 得證 $\mu_T(x) = \mu_{[T]_\beta}(x)$. □

最後我們來探討 minimal polynomial 和 characteristic polynomial 之間的關係.

Theorem 3.3.7.

(1) *假設 $A \in M_n(F)$, 則 $\mu_A(x) \mid \chi_A(x)$. 而且 $\lambda \in F$ 滿足 $\chi_A(\lambda) = 0$ 若且唯若 $\mu_A(\lambda) = 0$.*

(2) *假設 V 為 finite dimensional F-space, $T : V \to V$ 為 linear operator, 則 $\mu_T(x) \mid \chi_T(x)$. 而且 $\lambda \in F$ 滿足 $\chi_T(\lambda) = 0$ 若且唯若 $\mu_T(\lambda) = 0$.*

Proof.

(1) 因 $\chi_A(A)=\mathbf{O}$, 由 Lemma 3.3.2 知 $\mu_A(x)\mid\chi_A(x)$. 由此可得若 $\mu_A(\lambda)=0$ 則 $\chi_A(\lambda)=0$. 反之, 若 $\chi_A(\lambda)=0$, 則表示 $\det(\lambda I_n-A)=0$ 亦即 λI_n-A 不是 invertible matrix. 現考慮 $\mu_A(x)$ 除以 $x-\lambda$, 得 $\mu_A(x)=(x-\lambda)h(x)+r$, 其中 $h(x)\in F[x]$ 且 $r\in F$. 代入 A, 得 $\mathbf{O}=\mu_A(A)=(A-\lambda I_n)\cdot h(A)+rI_n$. 若 $r\neq 0$, 由 $(\lambda I_n-A)\cdot h(A)=rI_n$ 得 $(\lambda I_n-A)\cdot r^{-1}h(A)=I_n$. 此代表 $r^{-1}h(A)$ 為 λI_n-A 的 inverse, 與 λI_n-A 不是 invertible matrix 相矛盾, 得知 $r=0$, 亦即 $x-\lambda\mid\mu_A(x)$. 得證 $\mu_A(\lambda)=0$.

(2) 對於 linear operator $T:V\to V$, 選定 V 的一個 ordered basis β, 由於 $\chi_T(x)=\chi_{[T]_\beta}(x)$ 以及 $\mu_T(x)=\mu_{[T]_\beta}(x)$. 套用 (1) 的結果於 $[T]_\beta$, 我們得證 $\mu_T(x)\mid\chi_T(x)$ 且

$$\chi_T(\lambda)=0\Leftrightarrow\mu_T(\lambda)=0.$$

$\square$

Example 3.3.8. 我們利用前面 Example 3.2.6 所得的 characteristic polynomial 來求它們的 minimal polynomial. 因 $\chi_{A_1}(x)=x^2$, 依 Theorem 3.3.7 知 $\mu_{A_1}(x)$ 應為 x 或 x^2. 但 $A_1\neq\mathbf{O}$, 知 A_1 的 minimal polynomial 不可能為 x, 得知 $\mu_{A_1}(x)=x^2$.

因 $\chi_{A_2}(x)=x^2-1$, 依 Theorem 3.3.7 知 $x-1$ 和 $x+1$ 都是 $\mu_{A_2}(x)$ 的因式, 又 $\mu_{A_2}(x)\mid x^2-1$ 得知 $\mu_{A_2}(x)=x^2-1$.

因 $\chi_{A_3}(x)=x^2+1$, 依 Theorem 3.3.7 知 $\mu_{A_3}(x)\mid x^2+1$. 若 $F=\mathbb{R}$, x^2+1 的 monic factor (因式) 僅有 1 和 x^2+1, 又 minimal polynomial 不能是常數多項式, 得證 $\mu_{A_3}(x)\mid x^2+1$. 又若 $F=\mathbb{C}$, 因 $i,-i$ 皆為 $x^2+1=0$ 的根, 依 Theorem 3.3.7 知 $\mu_{A_3}(x)=x^2+1$.

Question 3.6. 你能找到 $A\in M_2(\mathbb{R})$, 使得 $\mu_A(x)\neq\chi_A(x)$ 嗎?

Question 3.7. 若 $A\in M_n(F)$ 且 $\chi_A(x)=(x-\lambda_1)\cdots(x-\lambda_n)$ 其中 $\lambda_i\in F$ 且 $\lambda_i\neq\lambda_j$ *for* $i\neq j$, 則 $\mu_A(x)$ 是什麼?

我們可以將 Theorem 3.3.7 做進一步的推廣, 這需要複習一下學過的代數. 假設 $p(x)\in F[x]$ 是一個 irreducible polynomial, 我們可以找到 F 的一個 finite extension $\tilde{F}$, 使得 $p(x)=0$ 在 $\tilde{F}$ 中有根. 假設 $\lambda\in\tilde{F}$ 為一根 (即 $p(\lambda)=0$), 則對於任意 $f(x)\in F[x]$, 滿足 $f(\lambda)=0$, 因 $p(x)$ 為 irreducible, 我們知 $p(x)\mid f(x)$. 現若 $A\in M_n(F)$, A 也可視為在 $M_n(\tilde{F})$ 中. A 的 characteristic polynomial 不管將 A 視為哪裡的矩陣, 其定義皆為 $\det(xI_n-A)$, 此和將 A 視為 $M_n(F)$ 或 $M_n(\tilde{F})$ 中的 matrix 無關. 但 minimal polynomial 的定義就和哪一個 field 有關了. 若將 A 視為 $M_n(\tilde{F})$ 的矩陣, 其 minimal polynomial (在此用 $\tilde{\mu}_A(x)$ 表示), 其定義為在 $\tilde{F}[x]$ 中次數最小的 monic polynomial $f(x)$ 使得 $f(A)=\mathbf{O}$. 所以因為 $\mu_A(x)\in F[x]\subseteq\tilde{F}[x]$, 利用 Lemma 3.3.2 我們知 $\tilde{\mu}_A(x)\mid\mu_A(x)$. 了解了這一層關係, 我們便有以下之重要定理.

Theorem 3.3.9.

(1) 假設 $A\in M_n(F)$ 且 $p(x)\in F[x]$ 是一個 *irreducible polynomial.* 則 $p(x)\mid\chi_A(x)$ 若且唯若 $p(x)\mid\mu_A(x)$.

(2) 假設 V 為 *finite dimensional F-space,* $T:V\to V$ 為 *linear operator,* 且 $p(x)\in F[x]$ 是一個 *irreducible polynomial.* 則 $p(x)\mid\chi_T(x)$ 若且唯若 $p(x)\mid\mu_T(x)$.

Proof.

(1) 由 Theorem 3.3.7 我們知 $\mu_A(x) \mid \chi_A(x)$, 故若 $p(x) \mid \mu_A(x)$ 則得 $p(x) \mid \chi_A(x)$. 另一方面, 若 $p(x) \in F[x]$ 為 irreducible 且 $p(x) \mid \chi_A(x)$. 考慮 $\tilde{F}$ 為 F 的 finite extension, 使得 $p(x) = 0$ 在 $\tilde{F}$ 中有一根 λ. 將 A 視為在 $M_n(\tilde{F})$ 的矩陣且令 $\tilde{\mu}_A(x) \in \tilde{F}[x]$ 為 $A \in M_n(\tilde{F})$ 在 $\tilde{F}[x]$ 的 minimal polynomial. 此時由於 $p(x) \mid \chi_A(x)$, 我們有 $\chi_A(\lambda) = 0$. 利用 Theorem 3.3.7 套用在 $\tilde{F}$ 的情形, 得 $\tilde{\mu}_A(\lambda) = 0$. 然而已知 $\tilde{\mu}_A(x) \mid \mu_A(x)$, 得 $\mu_A(\lambda) = 0$. 現因 $\mu_A(x) \in F[x]$ 且 $p(x) \in F[x]$ 為 irreducible, 得證 $p(x) \mid \mu_A(x)$.

(2) 對於 linear operator $T : V \to V$, 選定 V 的一個 ordered basis β, 由於 $\chi_T(x) = \chi_{[T]_\beta}(x)$ 以及 $\mu_T(x) = \mu_{[T]_\beta}(x)$. 套用 (1) 的結果於 $[T]_\beta$, 我們得證

$$p(x) \mid \chi_T(x) \Leftrightarrow p(x) \mid \chi_{[T]_\beta}(x) \Leftrightarrow p(x) \mid \mu_{[T]_\beta}(x) \Leftrightarrow p(x) \mid \mu_T(x).$$

□

Question 3.8. 若 $A \in M_n(F)$ 且 $\chi_A(x) = p_1^{c_1}(x) \cdots p_k^{c_k}(x)$ 其中 $c_i \in \mathbb{N}$, $p_i(x) \in F[x]$ 為 *monic irreducible polynomial* 且 $p_i(x) \neq p_j(x)$ *for* $i \neq j$, 則 $\mu_A(x)$ 會是怎樣的形式?

Example 3.3.10. 考慮 linear operator $T : P_2(\mathbb{R}) \to P_2(\mathbb{R})$ 滿足

$$T(1) = 2x^2 - 1, T(x+1) = 3x^2 + 2x + 2, T(-x^2 + x + 1) = 4x^2 + 2x + 2.$$

我們想找出 T 的 minimal polynomial $\mu_T(x)$.

首先考慮 $P_2(\mathbb{R})$ 的 ordered basis $\beta = (-x^2 + x + 1, x + 1, 1)$. 因

$$\begin{aligned} T(-x^2+x+1) &= (-4)(-x^2+x+1) + 6(x+1) \\ T(x+1) &= (-3)(-x^2+x+1) + 5(x+1) \\ T(1) &= (-2)(-x^2+x+1) + 2(x+1) + (-1)1 \end{aligned}$$

得 $[T]_\beta = \begin{pmatrix} -4 & -3 & -2 \\ 6 & 5 & 2 \\ 0 & 0 & -1 \end{pmatrix}$. 計算得 $\chi_T(x) = \chi_{[T]_\beta}(x) = (x+1)^2(x-2)$. 又

$$([T]_\beta + I_3) \cdot ([T]_\beta - 2I_3) = \begin{pmatrix} -3 & -3 & -2 \\ 6 & 6 & 2 \\ 0 & 0 & 0 \end{pmatrix} \cdot \begin{pmatrix} -6 & -3 & -2 \\ 6 & 3 & 2 \\ 0 & 0 & -3 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 6 \\ 0 & 0 & -6 \\ 0 & 0 & 0 \end{pmatrix},$$

知 $\mu_T(x) = \mu_{[T]_\beta}(x) \neq (x+1)(x-2)$, 而得 $\mu_T(x) = \mu_{[T]_\beta}(x) = (x+1)^2(x-2)$. 事實上

$$([T]_\beta + I_3)^2 \cdot ([T]_\beta - 2I_3) = \begin{pmatrix} -3 & -3 & -2 \\ 6 & 6 & 2 \\ 0 & 0 & 0 \end{pmatrix} \cdot \begin{pmatrix} 0 & 0 & 6 \\ 0 & 0 & -6 \\ 0 & 0 & 0 \end{pmatrix} = \mathbf{O}.$$

Question 3.9. 試利用 *ordered basis* $(x^2, x, 1)$ 處理 *Question 3.3.10.* 會不會有一樣結果?

Exercise 3.2. Determinant the characteristic and minimal polynomials of each of the following matrices:

$$\begin{pmatrix} 1 & 2 & 3 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 0 & 3 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 0 & 1 \\ 0 & 2 & 0 \\ 1 & 0 & 1 \end{pmatrix}, \begin{pmatrix} 1 & -1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{pmatrix}, \begin{pmatrix} 0 & 0 & 2 \\ 1 & 0 & -1 \\ 0 & 1 & 1 \end{pmatrix}.$$

Exercise 3.3. Suppose that $T \in \mathscr{L}(V)$ and $p(x)$ is an irreducible polynomial in $F[x]$ such that $p(T)$ is not one-to-one. Prove that $p(x) \mid \chi_T(x)$ and $p(x) \mid \mu_T(x)$.