大學線性代數再探

李華介

國立台灣師範大學數學系

前言

本講義主要目的是針對大三學生介紹有關線性代數進一步的理論. 主要是著重於一個 linear operator 的結構問題. 先備知識在線性代數方面需了解矩陣的運算, 行列式的性質等. 至於向量空間以及 linear transformation 等基本性質, 在本講義會再次介紹. 另外代數方面需了解 field 的基本性質以及 over 一個 field 的 polynomial ring 的代數結構 (即多項式環的除法原理).

本講義雖然主要以中文撰寫, 不過當涉及定義或專有名詞時, 為免翻譯的困擾將以英文取代. 因此將以中英夾雜較不傳統的方式顯現, 若有不便請見諒.

本講義編寫費時, 編寫完後並未經過嚴謹的校對. 疏漏在所難免, 雖不至於有理論性上嚴重的錯誤, 但讀者仍應注意不宜概括全收. 若發現錯誤, 歡迎提出寶貴的意見.

Chapter 3

Linear Operator

當 V 是一個 vector space 時, 從 V 到 V 的 linear transformation, 就稱為是一個 *linear operator* on V. 當 $T:V\to V$ 是一個 linear operator 時, 我們很自然的可以考慮其合成 $T^{\circ 2}=T\circ T$, 以及 $T^{\circ 3}=T\circ T^{\circ 2},\dots$ 這樣一直下去對任意 $i\in\mathbb{N}$ 都可以定出 $T^{\circ i}=T\circ T^{\circ i-1}$ ($T^{\circ 0}=\mathrm{id}$). 如此一來賦予 V 一個很豐富的代數結構 (稱為 $F[T]$-module), 所以我們可以進一步去了解 T 和 V 的關係. 這就是我們這一章進一步談 linear operator 的原因. 由於大家可能對代數不是很熟悉, 所以我們會避免使用太多額外的代數語言, 用大家熟悉的方法 (藉由矩陣, 行列式) 來介紹相關的理論.

3.1. Basic Concept

一個 linear operator 就是一個 linear transformation 所以前一章的理論我們都可以利用. 由於定義域和對映域是同一個 vector space, 我們可以選相同的 ordered basis, 這會讓矩陣表示法變得較簡單. 也就是說若 $T:V\to V$ 是一個 linear operator, 要得到 T 的 representative matrix, 我們可以選定 V 的一個 ordered basis $\beta=(\mathbf{v}_1,\dots,\mathbf{v}_n)$, 兩邊都用 β, 得 ${}_\beta[T]_\beta$ 這一個 $n\times n$ matrix. 為了方便起見當兩邊選的 ordered basis 相同時, T 的 representative matrix, 我們就會用 $[T]_\beta$ 來表示, 也就是說

$$[T]_\beta=\big(\tau_\beta(T(\mathbf{v}_1)),\dots,\tau_\beta(T(\mathbf{v}_n))\big).$$

例如若 T_1,T_2 皆為 V 的 linear operator, 由 Chapter 2 的 Proposition 2.4.5 我們知

$$[T_2\circ T_1]_\beta=[T_2]_\beta\cdot[T_1]_\beta. \tag{3.1}$$

另外依此表法, 我們有 $[\mathrm{id}]_\beta=I_n$.

習慣上我們會把 V 的 linear operators 所成的 vector space $\mathscr{L}(V,V)$ 簡化成 $\mathscr{L}(V)$. 又因為這裡的矩陣皆為 $n\times n$ 的方陣, 所以我們用 $M_n(F)$ 來表示所有 over F 的 $n\times n$ matrices. 利用這些符號, 當固定一個 V 的 ordered basis β 時, Theorem 2.4.4 告訴我們可以得到一個 $\mathscr{L}(V)$ 到 $M_n(F)$ 的 isomorphism, 即 $\Phi:\mathscr{L}(V)\to M_n(F)$, $T\mapsto[T]_\beta$. 特別的由於 $[\mathrm{id}]_\beta=I_n$, 我們有 $[T]_\beta=I_n$ 若且唯若 $T=\mathrm{id}$. 同理 $[T]_\beta$ 是一個 zero matrix 若且唯若 $T:V\to V$ 是 zero

mapping, 即 $T(\mathbf{v}) = \mathbf{O}_V, \forall \mathbf{v} \in V$. 為了方便起見, 我們將 zero matrix 和 zero mapping 都用 $\mathbf{O}$ 表示. 所以我們有以下之結論.

Lemma 3.1.1. 假設 V 為 *finite dimensional vector space,* $\dim(V) = n$ 且 β 為 V 的一個 *ordered basis.* 設 $T : V \to V$ 為 *linear operator,* 我們有以下結果*:*

$$[T]_\beta = I_n \Leftrightarrow T = \mathrm{id} \quad \text{and} \quad [T]_\beta = \mathbf{O} \Leftrightarrow T = \mathbf{O}.$$

前面說過考慮 linear operator 時, 我們幾乎都會選定義域和對映域有相同的 ordered basis. 不過有一個例外, 就是 identity 這個 linear operator, $\mathrm{id} : V \to V$. 因為對同一個 linear operator 若換另外的一個 ordered basis 來處理, 它的 representative matrix 就可能不一樣了, 我們需了解這樣的 matrices 之間有何關係, 就得靠 ${}_{\beta'}[\mathrm{id}]_\beta$ 這樣的 change of basis matrix 來幫忙了. 利用 Proposition 2.4.6, 我們有以下之結果.

Lemma 3.1.2. 設 β, β' 為 V 的 *ordered bases,* $T : V \to V$ 為 *linear operator,* 則

$$[T]_{\beta'} = {}_\beta[\mathrm{id}]_{\beta'}^{-1} \cdot [T]_\beta \cdot {}_\beta[\mathrm{id}]_{\beta'}.$$

Proof. 利用 Proposition 2.4.6, 我們知 $[T]_{\beta'} = {}_{\beta'}[\mathrm{id}]_\beta \cdot [T]_\beta \cdot {}_\beta[\mathrm{id}]_{\beta'}$. 然而若 $\dim(V) = n$, 由式子 (2.6) 我們知

$${}_{\beta'}[\mathrm{id}]_\beta \cdot {}_\beta[\mathrm{id}]_{\beta'} = {}_\beta[\mathrm{id}]_{\beta'} \cdot {}_{\beta'}[\mathrm{id}]_\beta = I_n,$$

亦即 ${}_{\beta'}[\mathrm{id}]_\beta = {}_\beta[\mathrm{id}]_{\beta'}^{-1}$, 得證本定理. □

當 $A, B \in M_n(F)$, 而 P 為 $M_n(F)$ 中的 invertible matrix, 若 $B = P^{-1} \cdot A \cdot P$, 則稱 A, B 為 *similar matrix*, 用 $A \sim B$ 來表示. 此時因 $\det(P^{-1}) = \det(P)^{-1}$ 知

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

由 Lemma 3.1.2 我們知道 $[T]_\beta \sim [T]_{\beta'}$, 故得 $\det([T]_\beta) = \det([T]_{\beta'})$. 也就是說不管用哪一個 ordered basis, T 的 representative matrix 的 determinant 皆相同, 我們也因此定義這就是 T 的 determinant, 也就是說 $\det(T) = \det([T]_\beta)$.

Lemma 3.1.2 反過來是對嗎? 有就是說若 $A \sim [T]_\beta$, 是否可找到 V 的一個 ordered basis β' 使得 $A = [T]_{\beta'}$ 呢? 事實上, 若 P 是一個 invertible matrix 使得 $A = P^{-1} \cdot [T]_\beta \cdot P$, 則由 Proposition 2.4.7, 我們能找到 V 的一個 ordered basis β' 滿足 $P = {}_\beta[\mathrm{id}]_{\beta'}$, 故由 Lemma 3.1.2 知

$$A = P^{-1} \cdot [T]_\beta \cdot P = {}_\beta[\mathrm{id}]_{\beta'}^{-1} \cdot [T]_\beta \cdot {}_\beta[\mathrm{id}]_{\beta'} = [T]_{\beta'}.$$

因此我們有以下之結論.

Proposition 3.1.3. 假設 V 為 *finite dimensional vector space,* $\dim(V) = n$ 且 β 為 V 的一個 *ordered basis.* 設 $T : V \to V$ 為 *linear operator* 且 $A \in M_n(F)$, 則 $A \sim [T]_\beta$ 若且唯若存在 V 的一個 *ordered basis* β' 使得 $A = [T]_{\beta'}$.

當我們要探討一個 linear operator 的性質時, 我們可以固定一個 ordered basis 將之轉換成 square matrix 的問題, 而 Proposition 3.1.3 告訴我們這些性質應對於 similar matrices 應是不變的, 以後我們會看到許多例子和這事實相呼應. 我們先看一個簡單的情形.

Lemma 3.1.4. 假設 V 為 *finite dimensional vector space,* β 為 V 的一個 *ordered basis* 且 $T: V \to V$ 為 *linear operator,* 則下列是等價的*:*

(1) T 是一個 *isomorphism.*

(2) $[T]_\beta$ 是一個 *invertible matrix.*

(3) $\det(T) \neq 0$.

Proof. 我們知 $[T]_\beta$ 是一個 invertible matrix 若且唯若 $\det([T]_\beta) \neq 0$, 所以僅要證 (1) $\Leftrightarrow$ (2).

假設 $\dim(V) = n$, 由 T 是 isomorphism, 知 $T^{\circ -1}$ 存在且為 linear operator, 故由

$$[T^{\circ -1}]_\beta \cdot [T]_\beta = [\mathrm{id}]_\beta = [T]_\beta \cdot [T^{\circ -1}]_\beta \quad \text{以及} \quad [\mathrm{id}]_\beta = I_n$$

知 $[T]_\beta$ 為 invertible. 反之, 若 $A \cdot [T]_\beta = I_n$, 由 $\Phi: \mathscr{L}(V) \to M_n(F)$, 為 isomorphism, 知存在 $T': V \to V$ 使得 $\Phi(T') = [T']_\beta = A$. 故由 $[T' \circ T]_\beta = [T']_\beta \cdot [T]_\beta = I_n$ 以及 Lemma 3.1.1 得 $T' \circ T = \mathrm{id}$, 同理由 $[T]_\beta \cdot [T']_\beta = I_n$ 得 $T \circ T' = \mathrm{id}$, 得證 T 為 isomorphism. □

Question 3.1. 可否從 *Lemma 3.1.4* 知若 $A \sim B$ 則 A 是 *invertible* 若且唯若 B 是 *invertible.*

由這裡我們可以看出求一個談論 linear operator 的性質離不開 determinant, 我們在這裡複習一個求 determinant 的方法. 若 $A \in M_n(F)$, 令 $a_{ik} \in F$ 表示在 A 的 (i,k)-th entry (即在 A 的 i-th row 及 k-th column 位置的元素), 且令 $A_{ik} \in M_{n-1}(F)$ 為將 A 的 i-th row 和 k-th column 刪除後所得的 $(n-1) \times (n-1)$ matrix. 我們可以用降階的方法求 $\det(A)$ 即對 i-th row 降階, 得

$$\det(A) = \sum_{k=1}^{n} (-1)^{i+k} a_{ik} \det(A_{ik}),$$

也可對 j-th column 降階得

$$\det(A) = \sum_{k=1}^{n} (-1)^{k+j} a_{kj} \det(A_{kj}).$$

我們也可定義一個 $n \times n$ matrix 稱為 *adjoint matrix* of A, 用 $\mathrm{adj}(A)$ 來表示, 其定義為 $\mathrm{adj}(A)$ 的 (i,j)-th entry 為

$$\mathrm{adj}(A)_{ij} = (-1)^{i+j} \det(A_{ji}).$$

利用此矩陣我們有以下的結果.

Lemma 3.1.5. 假設 A 為 $n \times n$ *matrix,* 令 $\mathrm{adj}(A)$ 為 A 的 *adjoint matrix,* 則

$$A \cdot \mathrm{adj}(A) = \mathrm{adj}(A) \cdot A = \det(A) I_n.$$

Proof. $\det(A)I_n$ 是一個 diagonal matrix, 即在對角線的位置是 $\det(A)$ 而其他非對角線位置為 0. 先檢查 $A\cdot\mathrm{adj}(A)$ 的 (i,i)-th entry, 依矩陣乘法定義此即

$$\sum_{k=1}^{n} a_{ik}\,\mathrm{adj}(A)_{ki} = \sum_{k=1}^{n}(-1)^{k+i}a_{ik}\,\det(A_{ik}) = \det(A).$$

另一方面當 $i\neq j$, $A\cdot\mathrm{adj}(A)$ 的 (i,j)-th entry, 依矩陣乘法定義為

$$\sum_{k=1}^{n} a_{ik}\,\mathrm{adj}(A)_{kj} = \sum_{k=1}^{n}(-1)^{k+j}a_{ik}\,\det(A_{jk}).$$

若將矩陣 A 的 j-th row 用 i-th row 取代, 所得的矩陣用 A' 表示, 由於 A' 的 i-th row 和 j-th row 相同, 我們知 $\det(A')=0$. 然而利用 A' 的 (j,k)-th entry a'_{jk} 為 a_{ik} 以及 $A'_{jk}=A_{jk}$, 對 A' 的 j-th row 降階, 我們有

$$0=\det(A')=\sum_{k=1}^{n}(-1)^{j+k}a'_{jk}\,\det(A'_{jk})=\sum_{k=1}^{n}(-1)^{j+k}a_{ik}\,\det(A_{jk}),$$

故知當 $i\neq j$ 時 $A\cdot\mathrm{adj}(A)$ 的 (i,j)-th entry 為

$$\sum_{k=1}^{n} a_{ik}\,\mathrm{adj}(A)_{kj} = \sum_{k=1}^{n}(-1)^{k+j}a_{ik}\,\det(A_{jk}) = 0.$$

得證 $A\cdot\mathrm{adj}(A)=\det(A)I_n$. 同理, 利用對 column 降階求 determinant, 可得 $\mathrm{adj}(A)\cdot A=\det(A)I_n$. □

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 $f(x)=c_dx^d+\cdots+c_1x+c_0$ 以及一個 $n\times n$ matrix A, 我們定義

$$f(A)=c_dA^d+\cdots+c_1A+c_0I_n.$$

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

$$\begin{aligned}A^i\cdot f(A) &= A^i\cdot(c_dA^d+\cdots+c_1A+c_0I_n)\\ &= c_dA^{d+i}+\cdots+c_1A^{1+i}+c_0A^i=(c_dA^d+\cdots+c_1A+c_0I_n)\cdot A^i=f(A)\cdot A^i.\end{aligned}$$

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 $fA)\sim f(B)$ 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 $f(x)=c_dx^d+\cdots+c_1x+c_0$, 則

$$f(B)=c_dB^d+\cdots+c_1B+c_0I_n=c_d(P^{-1}\cdot A\cdot P)^d+\cdots+c_1(P^{-1}\cdot A\cdot P)+c_0I_n$$
$$=c_d(P^{-1}\cdot A^d\cdot P)+\cdots+c_1(P^{-1}\cdot A\cdot P)+c_0I_n=P^{-1}\cdot(c_dA^d+\cdots+c_1A+c_0I_n)\cdot P=P^{-1}\cdot f(A)\cdot P,$$

得證 $f(A)\sim f(B)$. □

我們也可把這概念推廣到 linear operator, 假設 $f(x)=c_dx^d+\cdots+c_1x+c_0\in F[x]$ 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

$$f(T)=c_dT^{\circ d}+\cdots+c_1T+c_0\,\mathrm{id},$$

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

$$[f(T)]_\beta=f([T]_\beta)=c_d[T]_\beta^d+\cdots+c_1[T]_\beta+c_0I_n.$$

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

$$[f(T)]_\beta=[c_dT^{\circ d}+\cdots+c_1T+c_0\,\mathrm{id}]_\beta=$$
$$c_d[T^{\circ d}]_\beta+\cdots+c_1[T]_\beta+c_0[\mathrm{id}]_\beta=c_d[T]_\beta^d+\cdots+c_1[T]_\beta+c_0I_n=f([T]_\beta).$$

□

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 $S=\{I_n,A,A^2,\dots,A^{n^2}\}$. 由於 $\#(S)=n^2+1>\dim(M_n(F))$, 我們知 S 為 linearly dependent. 亦即存在 $c_0,c_1,\dots,c_{n^2}\in F$ 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 $c_{n^2},\dots,c_1,c_0$ 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 $\deg(f(x))\le n^2$ 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

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

應該和等式 (3.2) 右式相同, 故比較係數得

$$\begin{aligned} -B_0 \cdot A &= c_0 I_n \\ B_0 \cdot I_n - B_1 \cdot A &= c_1 I_n \\ &\vdots \\ B_{n-2} \cdot I_n - B_{n-1} \cdot A &= c_{n-1} I_n \\ B_{n-1} \cdot I_n &= I_n \end{aligned}$$

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

$$\begin{aligned} -B_0 \cdot A &= c_0 I_n \\ B_0 \cdot A - B_1 \cdot A^2 &= c_1 A \\ &\vdots \\ B_{n-2} \cdot A^{n-1} - B_{n-1} \cdot A^n &= c_{n-1} A^{n-1} \\ B_{n-1} \cdot A^n &= A^n \end{aligned}$$

因未左邊全部加起來會等於右邊全部加起來, 得證

$$\mathbf{O} = A^n + c_{n-1}A^{n-1} + \cdots + c_1 A + c_0 I_n = \chi_A(A).$$

□

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

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

□

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*?*

3.4. Internal Direct Sum

給定一個 linear operator $T: V \to V$, 若選夠好的 ordered basis, T 的 representative matrix 可以是較好處理的 matrix. 不過這需要將 V 寫成所謂的 internal direct sum of T-invariant subspaces. 所以這一節我們先不談 linear operator, 先探討 internal direct sum 的性質.

我們在 Chapter 1 所介紹的 direct sum 其實是所謂的 *external direct sum*, 它是不管每個 vector space 之間的關係, 而造出的 vector space. 不過若每個 vector space 間有關係, 那麼我們便可以探討有關於 internal direct sum 的問題.

假設 U, W 皆為 V 的 subspace. 可以考慮函數 $T: U \oplus W \to U + W$, 定義為

$$T((\mathbf{u}, \mathbf{w})) = \mathbf{u} + \mathbf{w}, \quad \forall \mathbf{u} \in U, \mathbf{w} \in W.$$

依定義很容易得到 T 是 well-defined function, 且可得 T 是一個 onto 的 linear transformation. 接下來我們自然要問 $\mathrm{Ker}(T)$ 是什麼? 若 $(\mathbf{u}, \mathbf{w}) \in \mathrm{Ker}(T)$, 表示 $T((\mathbf{u}, \mathbf{w})) = \mathbf{u} + \mathbf{w} = \mathbf{O}_V$, 得 $\mathbf{u} = -\mathbf{w}$. 但 $\mathbf{u} \in U, \mathbf{w} \in W$, 故得 $\mathbf{u} = -\mathbf{w} \in U \cap W$. 反之, 若 $\mathbf{u} \in U \cap W$, 考慮 $(\mathbf{u}, -\mathbf{u}) \in U \oplus W$, 可得 $T((\mathbf{u}, -\mathbf{u})) = \mathbf{O}_V$. 得證 $\mathrm{Ker}(T) = \{(\mathbf{u}, -\mathbf{u}) \mid \mathbf{u} \in U \cap W\}$.

Question 3.10. 為何要得到 $T: U \oplus W \to U + W$ 這個函數需要 U, W 皆為 V 的 *subspace* 這個假設*?*

Question 3.11. 試證明 $\{(\mathbf{u}, -\mathbf{u}) \mid \mathbf{u} \in U \cap W\} \simeq U \cap W$. 利用 *the First Isomorphism Theorem,* 我們可不可以說 $(U \oplus W)/(U \cap W) \simeq U + W$*?*

特別地, 當 $U \cap W = \{\mathbf{O}_V\}$ 時, 因 $(\mathbf{O}_V, \mathbf{O}_V) = \mathbf{O}_{U \oplus W}$, 我們得 $\mathrm{Ker}(T) = \mathbf{O}_{U \oplus W}$. 亦即 T 為 one-to-one, 我們有以下之結果.

Proposition 3.4.1. 假設 U,W 皆為 V 的 *subspace,* 若 $U\cap W=\{\mathbf{O}_V\}$, 則

$$U\oplus W\simeq U+W.$$

就是因為這個原因, 當 U,W 皆為 V 的 subspace 且 $U\cap W=\{\mathbf{O}_V\}$ 時, 我們會將 $U+W$ 用 $U\oplus W$ 來表示. 要注意此時 $U\oplus W$ 指的是 V 的 subspace $U+W$, 不是以前定的那個新的 vector space. 這裡我們用 $U\oplus W$ 這個符號來強調 $U\cap W=\{\mathbf{O}_V\}$. 為了區分清楚, 我們會說這是 U,W 的 *internal direct sum.* 所以要注意, 若 U,W 皆為 V 的 subspace 且 $U\cap W\neq\{\mathbf{O}_V\}$ 時 $U\oplus W$ 這個符號絕對是代表 external direct sum. 若 U,W 皆為 V 的 subspace, 而我們強調 $U\oplus W\subseteq V$ 或說是 internal direct sum, 就表示 $U\cap W=\{\mathbf{O}_V\}$. 當然了若 U,W 沒有任何關聯, 那麼 $U\oplus W$ 指的是原本的 external direct sum.

當 V 為 finite dimensional vector space, 且 U 是 V 的 subspace. 我們可以找到另一個 V 的 subspace W 使得 $V=U\oplus W$. 事實上任取 U 的一組 basis $S=\{\mathbf{u}_1,\dots,\mathbf{u}_m\}$, 我們知可以將 S 擴大成 V 的一組 basis $\{\mathbf{u}_1,\dots,\mathbf{u}_m,\mathbf{w}_1,\dots,\mathbf{w}_n\}$. 此時若令 $W=\mathrm{Span}(\{\mathbf{w}_1,\dots,\mathbf{w}_n\})$, 由於 $\{\mathbf{u}_1,\dots,\mathbf{u}_m,\mathbf{w}_1,\dots,\mathbf{w}_n\}$ 為 linearly independent, 我們知 $U\cap W=\{\mathbf{O}_V\}$. 所以可得 $U\oplus W$ 這一個 U,W 的 internal direct sum. 又因為 $V=\mathrm{Span}(\{\mathbf{u}_1,\dots,\mathbf{u}_m,\mathbf{w}_1,\dots,\mathbf{w}_n\})$, 我們得 $U\oplus W=V$. 由於將一組 linearly independent 元素擴展成 basis 的方法並不唯一, 從這裡我們也了解到給定 V 的一個 subspace U, 可將 V 寫成 $U\oplus W$ 的 W 並不唯一.

Example 3.4.2. 考慮 $F^2=\{(x,y)\mid x,y\in F\}$, 若 $U=\{(x,0)\mid x\in F\}$, 則 $W_1=\{(0,y)\mid y\in F\}$ 和 $W_2=\{(y,y)\mid y\in F\}$ 都滿足 $F^2=U\oplus W_1$ 以及 $F^2=U\oplus W_2$.

將 V 寫成 internal direct sum $V=U\oplus W$ 的一個好處就是若 $\mathbf{v}\in V$, 則存在唯一的 $\mathbf{u}\in U$ 以及 $\mathbf{w}\in W$ 使得 $\mathbf{v}=\mathbf{u}+\mathbf{w}$. 我們將 V 寫成兩個 subspaces 的 internal direct sum 的性質列舉如下.

Proposition 3.4.3. 假設 U,W 為 V 的 *subspaces.* 下列是等價的

(1) $V=U\oplus W$.

(2) 若 $\mathbf{v}\in V$, 則存在唯一的 $\mathbf{u}\in U,\mathbf{w}\in W$ 使得 $\mathbf{v}=\mathbf{u}+\mathbf{w}$.

(3) 對任意 U,W 的 *basis* S_1,S_2, 我們有 $S_1\cap S_2=\emptyset$ 且 $S_1\cup S_2$ 為 V 的一組 *basis.*

Proof. (1) $\Rightarrow$ (2): 依定義 $V=U+W$, 故對任意 $\mathbf{v}\in V$, 必存在 $\mathbf{u}\in U,\mathbf{w}\in W$ 使得 $\mathbf{v}=\mathbf{u}+\mathbf{w}$. 現若 $\mathbf{u}'\in U,\mathbf{w}'\in W$ 使得 $\mathbf{v}+\mathbf{w}=\mathbf{u}'+\mathbf{w}'$, 則考慮 $\mathbf{u}-\mathbf{u}'=\mathbf{w}'-\mathbf{w}\in U\cap W=\{\mathbf{O}_V\}$, 得證 $\mathbf{u}=\mathbf{u}'$ 且 $\mathbf{w}=\mathbf{w}'$.

(2) $\Rightarrow$ (3): 假設 $\mathbf{v}\in S_1\cap S_2$, 表示 $\mathbf{v}\in U\cap W$. 考慮 $\mathbf{v}=\mathbf{v}+\mathbf{O}_V=\mathbf{O}_V+\mathbf{v}$ 其中第一個 $\mathbf{v}$ 看成在 U, 第二個 $\mathbf{v}$ 看成在 W 且第一個 $\mathbf{O}_V$ 看成在 W, 第二個 $\mathbf{O}_V$ 看成在 U, 則利用唯一性知 $\mathbf{v}=\mathbf{O}_V$. 但 $\mathbf{v}\in S_1$, 此和 S_1 為 linearly independent 相矛盾, 得知 $S_1\cap S_2=\emptyset$. 另外對任意 $\mathbf{v}\in V$, 知存在 $\mathbf{u}\in U,\mathbf{w}\in W$ 使得 $\mathbf{v}=\mathbf{u}+\mathbf{w}$, 然而因 S_1,S_2 分別為 U,W 的 basis, 知存在 $\mathbf{u}_1,\dots,\mathbf{u}_m\in S_1$, $\mathbf{w}_1,\dots,\mathbf{w}_n\in S_2$ 以及 $c_1,\dots,c_m,d_1,\dots,d_n\in F$ 使得 $\mathbf{u}=c_1\mathbf{u}_1+\cdots+c_m\mathbf{u}_m$, $\mathbf{w}=d_1\mathbf{w}_1+\cdots+d_n\mathbf{w}_n$. 因此得 $\mathbf{v}=c_1\mathbf{u}_1+\cdots+c_m\mathbf{u}_m+d_1\mathbf{w}_1+\cdots+d_n\mathbf{w}_n$, 得證 $S_1\cup S_2$ 為 V 的

spanning set. 另一方面若 $S_1 \cup S_2$ 不為 linearly independent, 利用 Corollary 1.4.4 知存在 $\mathbf{v} \neq \mathbf{O}_V$ 使得 $\mathbf{v} \in \text{Span}(S_1) \cap \text{Span}(S_2) = U \cap W$. 同前面證明 $S_1 \cap S_2 = \emptyset$ 的方法知, 此與 $\mathbf{v}$ 寫成 U, W 元素相加的唯一性相矛盾. 故知 $S_1 \cup S_2$ 為 linearly independent.

(3) $\Rightarrow$ (1): 由 $S_1 \cup S_2$ 為 V 的一組 basis, 知 $V = \text{Span}(S_1) + \text{Span}(S_2) = U + W$. 現僅需證 $U \cap W = \{\mathbf{O}_V\}$. 因 $S_1 \cap S_2 = \emptyset$ 我們有 $(S_1 \cup S_2) \setminus S_1 = S_2$, 故利用 Corollary 1.4.4 知 $S_1 \cup S_2$ 為 linearly independent 表示 $\text{Span}(S_1) \cap \text{Span}(S_2) = \{\mathbf{O}_V\}$, 亦即 $U \cap W = \{\mathbf{O}_V\}$. □

我們可以把兩個 subspaces 的 internal direct sum 推廣到更多 subspaces 的 internal direct sum. 例如 $V = U \oplus W$, 我們還可將 W 寫成兩個 W 的 subspaces W_1, W_2 的 direct sum, $W = W_1 \oplus W_2$, 而得 $V = U \oplus W_1 \oplus W_2$. 這裡因 $W = W_1 \oplus W_2$, 我們有 $W_1 \cap W_2 = \{\mathbf{O}_V\}$, 又因 $V = U \oplus W$, 我們也有 $U \cap W_1 \subseteq U \cap W = \{\mathbf{O}_V\}$, $U \cap W_2 \subseteq U \cap W = \{\mathbf{O}_V\}$. 不過這些條件 (即 $W_1 \cap W_2 = \{\mathbf{O}_V\}$, $U \cap W_1 = \{\mathbf{O}_V\}$ 和 $U \cap W_2 = \{\mathbf{O}_V\}$) 並不足以讓我們有類似 Proposition 3.4.3 的性質 (例如任意 $\mathbf{v}$ 有唯一的 $\mathbf{u} \in U, \mathbf{w}_1 \in W_1, \mathbf{w}_2 \in W_2$ 使得 $\mathbf{v} = \mathbf{u} + \mathbf{w}_1 + \mathbf{w}_2$), 我們看以下的例子.

Example 3.4.4. 在 Example 3.4.2 中 $U \cap W_1 = W_1 \cap W_2 = U \cap W_2 = \{(0,0)\}$, 不過任意 $(x,y) \in F^2$, 若 $y \neq 0$, 我們有 $(x,y) = (x,0) + (0,y) + (0,0) = (x-y,0) + (0,0) + (y,y)$, 其中 $((0,0) \in W_1$ 但 $(0,0) \neq (0,y) \in W_1$. 同樣的, $(0,0) \neq (y,y) \in W_2$. 所以 F^2 中的元素寫成 U, W_1, W_2 之和的方法不唯一.

到底要怎麼定義 internal direct sum 呢? 我們可以回到 external direct sum 的看法. 假設 V_1, V_2, V_3 為 V 的 subspace, 考慮從 external direct sum $V_1 \oplus V_2 \oplus V_3$ 到 $V_1 + V_2 + V_3$ 的 linear transformation T, 定義為 $T(\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3) = \mathbf{v}_1 + \mathbf{v}_2 + \mathbf{v}_3$. 依定義 T 為 onto. 若 T 為 one-to-one, 則需 $\text{Ker}(T) = \{(\mathbf{O}_V, \mathbf{O}_V, \mathbf{O}_V)\}$ 亦即若 $\mathbf{v}_1 \in V_1, \mathbf{v}_2 \in V_2, \mathbf{v}_3 \in V_3$ 且 $\mathbf{v}_1 + \mathbf{v}_2 + \mathbf{v}_3 = \mathbf{O}_V$, 則 $\mathbf{v}_1 = \mathbf{v}_2 = \mathbf{v}_3 = \mathbf{O}_V$. 然而 $\mathbf{v}_1 + \mathbf{v}_2 + \mathbf{v}_3 = \mathbf{O}_V$, 知 $\mathbf{v}_1 = -(\mathbf{v}_2 + \mathbf{v}_3) \in V_1 \cap (V_2 + V_3)$, 同理知 $\mathbf{v}_2 \in V_2 \cap (V_1 + V_3)$, $\mathbf{v}_3 \in V_3 \cap (V_1 + V_2)$. 因此若知 $V_1 \cap (V_2 + V_3) = V_2 \cap (V_1 + V_3) = V_3 \cap (V_1 + V_2) = \{\mathbf{O}_V\}$, 則可得 $\text{Ker}(T) = \{(\mathbf{O}_V, \mathbf{O}_V, \mathbf{O}_V)\}$. 反之, 若 $\mathbf{v}_1 \in V_1 \cap (V_2 + V_3)$, 則存在 $\mathbf{v}_1 \in V_2, \mathbf{v}_3 \in V_3$ 滿足 $\mathbf{v}_1 = \mathbf{v}_2 + \mathbf{v}_3$, 此時 $(\mathbf{v}_1, -\mathbf{v}_2, -\mathbf{v}_3) \in \text{Ker}(T)$. 因此若 $\text{Ker}(T) = \{(\mathbf{O}_V, \mathbf{O}_V, \mathbf{O}_V)\}$ 表示 $\mathbf{v}_1 = \mathbf{O}_V$, 故知 $V_1 \cap (V_2 + V_3) = \mathbf{O}_V$. 同理可得 $V_2 \cap (V_1 + V_3) = V_3 \cap (V_1 + V_2) = \mathbf{O}_V$. 將此推廣到任意有限多個 subspaces, 我們有以下之定義.

Definition 3.4.5. 假設 $V_1, \ldots, V_k$ 為 V 的 subspaces, 且

$$V_i \cap (\sum_{j \neq i} V_j) = \{\mathbf{O}_V\}, \forall\, i = 1, \ldots, k$$

則 V 的 subspace $V_1 + \cdots + V_k$ 稱為 $V_1, \ldots, V_k$ 的 *internal direct sum*, 用 $V_1 \oplus \cdots \oplus V_k$ 表示.

再次強調, 對於 V 的 subspaces $V_1, \ldots, V_k$, 我們都有 $V_1 + \cdots + V_k$ 這一個 subspace. 若我們寫成 $V_1 \oplus \cdots \oplus V_k \subseteq V$ 或強調為 internal direct sum, 便是說 $V_1, \ldots, V_k$ 滿足 $V_i \cap (\sum_{j \neq i} V_j) = \{\mathbf{O}_V\}, \forall\, i = 1, \ldots, k$ 這些條件. 另外, 以後我們要談的 decomposition theorem, 都是將一個 vector space 拆解成一些 subspaces 的 internal direct sum, 我們不會再去談 external direct sum, 所以我們就不再強調為 internal direct sum.

將 vector space 寫成多個 subspaces 的 direct sum, 和寫成兩個 subspaces 的 direct sum 有同樣的性質. 由於證明和 Proposition 3.4.3 相同, 我們就不再證明了.

Proposition 3.4.6. 假設 $V_1,\dots,V_k$ 為 V 的 *subspace.* 下列是等價的

(1) $V = V_1 \oplus \cdots \oplus V_k$.

(2) 若 $\mathbf{v} \in V$, 則對於所有 $i = 1,\dots,k$ 皆存在唯一的 $\mathbf{v}_i \in V_i$ 使得 $\mathbf{v} = \mathbf{v}_1 + \cdots + \mathbf{v}_k$.

(3) 對任意 V_i 的 *basis* S_i, 我們有 $S_1 \cap \cdots \cap S_k = \emptyset$ 且 $S_1 \cup \cdots \cup S_k$ 為 V 的一組 *basis.*

Question 3.12. 若 V 為 *finite dimensional vector space* 且 $V_1,\dots,V_k$ 為 V 的 *subspaces* 使得 $V = V_1 \oplus \cdots \oplus V_k$, 那麼可以知道 $\dim(V)$ 會等於 $\dim(V_1) + \cdots + \dim(V_k)$ 嗎*?*

當 U,W 為 V 的 subspaces 且 $V = U \oplus W$, 又 $W_1,\dots,W_k$ 為 W 的 subspaces 且 $W = W_1 \oplus \cdots \oplus W_k$, 那麼我們可以得 $V = U \oplus W_1 \oplus \cdots \oplus W_k$ 嗎? 答案是肯定的. 這是因為若 $\mathbf{v} \in V$, 由 $V = U \oplus W$ 知存在 $\mathbf{u} \in U, \mathbf{w} \in W$ 使得 $\mathbf{v} = \mathbf{u} + \mathbf{w}$. 另一方面由 $W = W_1 \oplus \cdots \oplus W_k$, 知存在 $\mathbf{w}_i \in W_i$, 使得 $\mathbf{w} = \mathbf{w}_1 + \cdots + \mathbf{w}_k$. 也就是說對任意 $\mathbf{v} \in V$, 皆存在 $\mathbf{u} \in U, \mathbf{w}_1 \in W_1,\dots,\mathbf{w}_k \in W_k$ 使得 $\mathbf{v} = \mathbf{u} + \mathbf{w}_1 + \cdots + \mathbf{w}_k$ (證得存在性). 又若 $\mathbf{u}' \in U, \mathbf{w}_1' \in W_1,\dots,\mathbf{w}_k \in W_k$ 使得 $\mathbf{v} = \mathbf{u}' + \mathbf{w}_1' + \cdots + \mathbf{w}_k'$, 則因 $\mathbf{u},\mathbf{u}' \in U$ 以及 $\mathbf{w}_1 + \cdots + \mathbf{w}_k, \mathbf{w}_1' + \cdots + \mathbf{w}_k' \in W$, 由 $V = U \oplus W$ 得 $\mathbf{u} = \mathbf{u}'$ 以及 $\mathbf{w}_1 + \cdots + \mathbf{w}_k = \mathbf{w}_1' + \cdots + \mathbf{w}_k'$. 又因 $\mathbf{w}_i, \mathbf{w}_i' \in W_i$, 由 $W = W_1 \oplus \cdots \oplus W_k$ 得 $\mathbf{w}_i = \mathbf{w}_i'$ (證得唯一性), 所以由 Proposition 3.4.6 我們有以下之結果.

Corollary 3.4.7. 若 U,W 為 V 的 *subspaces* 且 $V = U \oplus W$, 又若 $W_1,\dots,W_k$ 為 W 的 *subspaces* 且 $W = W_1 \oplus \cdots \oplus W_k$, 則 $V = U \oplus W_1 \oplus \cdots \oplus W_k$.

3.5. Primary Decomposition

讓我們回到 linear operator. 若 $T : V \to V$ 為 linear operator, 我們希望將 V 寫成一些 subspaces 的 direct sum, 使這些 subspaces 的 ordered basis 所組成 V 的 ordered basis 讓 T 的 representative matrix 有比較好的形式. 要達到這個目的, 我們希望 T 限制在這些 subspaces 上是不會跑掉的 (即希望它們仍為 linear operator), 所以我們有以下的定義.

Definition 3.5.1. 假設 $T : V \to V$ 是一個 linear operator. 若 W 為 V 的 subspace 且滿足 $T(W) \subseteq W$ (即對所有 $\mathbf{w} \in W$ 皆有 $T(\mathbf{w}) \in W$), 則稱 W 為 *T-invariant.*

Question 3.13. 假設 $T : V \to V$ 是一個 *linear operator.* 下列哪些 *subspaces* 是 *T-invariant?*

(1) V. (2) $\{\mathbf{O}_V\}$. (3) $\mathrm{Im}(T)$. (4) $\mathrm{Ker}(T)$.

回顧一下, 當 $T : V \to V$ 為 linear operator, 對於 $f(x) = a_d x^d + \cdots + a_1 x + a_0 \in F[x]$, 我們可定義一個 linear operator $f(T) = a_d T^{\circ d} + \cdots + a_1 T + a_0 \mathrm{id}$.

Lemma 3.5.2. 假設 V 為 *F-space,* $T : V \to V$ 為 *linear operator.* 若 W 為 *T-invariant,* 則對任意 $f(x) \in F[x]$, W 為 *$f(T)$-invariant*

Proof. 因 W 為 T-invariant, 對任意 $\mathbf{w}\in W$, 因為 $T(\mathbf{w})\in W$ 故得 $T^{\circ 2}(\mathbf{w})=T(T(\mathbf{w}))\in W$. 利用數學歸納法知 $T^{\circ i}(\mathbf{w})\in W$, $\forall i\in\mathbb{N}$. 現若 $f(x)=a_dx^d+\cdots+a_1x+a_0\in F[x]$, 因 W 為 subspace, 得 $f(T)(\mathbf{w})=a_dT^{\circ d}(\mathbf{w})+\cdots+a_1T(\mathbf{w})+a_0\mathbf{w}\in W$, $\forall\mathbf{w}\in W$. 得證 W 為 $f(T)$-invariant. □

很容易判斷 $\mathrm{Im}(T)$ 和 $\mathrm{Ker}(T)$ 皆為 T-invariant. 我們可以利用 $f(x)\in F[x]$ 得到更多 T-invariant subspaces.

Lemma 3.5.3. 假設 V 為 F-*space,* $T:V\to V$ 為 *linear operator* 且 $f(x)\in F[x]$. 則 $\mathrm{Im}(f(T))$ 和 $\mathrm{Ker}(f(T))$ 皆為 T-*invariant subspaces.*

Proof. 假設 $\mathbf{w}\in\mathrm{Im}(f(T))$, 即存在 $\mathbf{v}\in V$ 使得 $\mathbf{w}=f(T)(\mathbf{v})$. 由 Lemma 3.2.3 我們知 $T\circ f(T)=f(T)\circ T$, 因此

$$T(\mathbf{w})=T(f(T)(\mathbf{v}))=(T\circ f(T))(\mathbf{v})=(f(T)\circ T)(\mathbf{v})=f(T)(T(\mathbf{v}))\in\mathrm{Im}(f(T)),$$

得證 $\mathrm{Im}(f(T))$ 為 T-invariant.

假設 $\mathbf{v}\in\mathrm{Ker}(f(T))$, 亦即 $f(T)(\mathbf{v})=\mathbf{O}_V$. 此時 $f(T)(T(\mathbf{v}))=T(f(T)(\mathbf{v}))=T(\mathbf{O}_V)=\mathbf{O}_V$, 亦即 $T(\mathbf{v})\in\mathrm{Ker}(f(T))$, 得證 $\mathrm{Ker}(f(T))$ 為 T-invariant. □

給定一個 linear operator $T:V\to V$, 考慮 V 的一個 subspace W, 我們可以將 T 的定義域限制在 W 上, 即考慮 $T|_W:W\to V$, 定義為 $T|_W(\mathbf{w})=T(\mathbf{w}),\forall\mathbf{w}\in W$. 這是一個從 W 到 V 的 linear transformation, 我們稱為 the *restriction* on W. 當 W 為 T-invariant 時, 因 $T(\mathbf{w})\in W$, $\forall\mathbf{w}\in W$, 我們有 $T|_W:W\to W$, 為一個 W 上的 linear operator. 我們自然可以探討 $T|_W$ 和 T 的 minimal polynomial 之間的關係. 首先對於 $f(x)\in F[x]$, 因 W 亦為 $f(T)$-invariant (Lemma 3.5.2), 我們有興趣知道 $f(T)|_W$ 和 $f(T|_W)$ 這兩個 W 的 linear operator 之間的關係. 現對所有 $\mathbf{w}\in W$, 因

$$T^{\circ 2}|_W(\mathbf{w})=T^{\circ 2}(\mathbf{w})=T(T(\mathbf{w}))=T|_W(T|_W(\mathbf{w}))=T|_W{}^{\circ 2}(\mathbf{w}),$$

我們知 $T^{\circ 2}|_W$ 和 $T|_W{}^{\circ 2}$ 為 W 上相同的 linear operator. 利用數學歸納法可得 $T^{\circ i}|_W=T|_W{}^{\circ i},\forall i\in\mathbb{N}$. 現若 $f(x)=a_dx^d+\cdots+a_1x+a_0\in F[x]$, 則對於任意 $\mathbf{w}\in W$, 皆有

$$\begin{aligned}f(T)|_W(\mathbf{w})=f(T)(\mathbf{w})&=a_dT^{\circ d}|_W(\mathbf{w})+\cdots+a_1T|_W(\mathbf{w})+a_0\,\mathrm{id}|_W(\mathbf{w})\\&=a_dT|_W{}^{\circ d}(\mathbf{w})+\cdots+a_1T|_W(\mathbf{w})+a_0\,\mathrm{id}|_W(\mathbf{w})=f(T|_W)(\mathbf{w}).\end{aligned}$$

也就是說 $f(T)|_W$ 和 $f(T|_W)$ 是 W 上相同的 linear operator, 因此知

$$f(T)|_W=f(T|_W).\tag{3.3}$$

利用此結果, 我們有以下之 Lemma.

Lemma 3.5.4. 假設 $T:V\to V$ 為 *linear operator,* W 為 T-*invariant subspace,* 則 T 的 *restriction on* W, $T|_W:W\to W$ 為 W 上的 *linear operator,* 且其 *minimal polynomial* $\mu_{T|_W}(x)$ 滿足

$$\mu_{T|_W}(x)\mid\mu_T(x).$$

Proof. 已知 $\mu_T(T)=\mathbf{O}$ 為一個 zero mapping, 故 $\mu_T(T)|_W=\mathbf{O}$. 故由等式 (3.3) 知 $\mu_T(T|_W)=\mu_T(T)|_W=\mathbf{O}$, 再由 Lemma 3.3.5 得證 $\mu_{T|_W}(x)\mid\mu_T(x)$. □

假設 V 可以寫成兩個 T-invariant subspace U,W 的 (internal) direct sum $V=U\oplus W$, 分別選取 U,W 的一個 ordered basis $\beta_1=(\mathbf{u}_1,\dots,\mathbf{u}_l),\beta_2=(\mathbf{w}_1,\dots,\mathbf{w}_m)$, 則由 Proposition 3.4.3 知 $\beta=(\mathbf{u}_1,\dots,\mathbf{u}_l,\mathbf{w}_1,\dots,\mathbf{w}_m)$ 亦為 V 的 ordered basis. 此時由於 $T(\mathbf{u}_i)=T|_U(\mathbf{u}_i)\in U$, 我們知 $[T]_\beta$ 的前面 l 個 columns, 每個 column 的前 l 個 entry 都和 $[T|_U]_{\beta_1}$ 相同, 而且後面 m 個 entry 皆為 0. 同樣的, 由於 $T(\mathbf{w}_j)=T|_W(\mathbf{w}_j)\in W$, 我們知 $[T]_\beta$ 的後面 m 個 columns, 每個 column 的前 l 個 entry 都是 0 而後面 m 個 entry 皆和 $[T|_W]_{\beta_2}$ 相同. 也就是說 T 對於 β 的 representative matrix 為

$$[T]_\beta=\begin{pmatrix}[T|_U]_{\beta_1} & \mathbf{O}\\ \mathbf{O} & [T|_W]_{\beta_2}\end{pmatrix} \tag{3.4}$$

要探討 $T,T|_U,T|_W$ 的 characteristic polynomial 間的關係, 需了解等式 (3.4) 這類 block diagonal matrix 的 determinant 算法. 我們簡單回顧一下, 考慮 matrix

$$A=\begin{pmatrix}B & \mathbf{O}\\ \mathbf{O} & C\end{pmatrix}$$

其中 $A\in M_{l+m}(F),B\in M_l(F),C\in M_m(F)$ 皆為 square matrix. 我們可以用降階及數學歸納法證得 $\det(A)=\det(B)\det(C)$. 方法大致如下: 我們對第一個 row 作降階得 $\det(A)=\sum_{k=1}^{l+m}(-1)^{1+k}a_{1k}\det(A_{1k})$, 然而 A_{1k} 是將 A 的 first row 和 k-th column 刪除, 因此當 $1\le k\le l$ 時, $a_{1k}=b_{1k}$ 且 $A_{1k}=\begin{pmatrix}B_{1k} & \mathbf{O}\\ \mathbf{O} & C\end{pmatrix}$ 這樣的 block diagonal matrix. 所以依數學歸納法假設, 此時 $\det(A_{1k})=\det(B_{1k})\det(C)$. 又當 $l<k\le l+m$ 時, $a_{1k}=0$, 故得

$$\det(A)=\sum_{k=1}^{l+m}(-1)^{1+k}a_{1k}\det(A_{1k})=\sum_{k=1}^{l}(-1)^{1+k}b_{1k}\det(B_{1k})\det(C)=\det(B)\det(C).$$

利用這個結果我們就可以得到 characteristic polynomial 的關係了.

Lemma 3.5.5. 假設 V 為 *finite dimensional F-space,* $T:V\to V$ 為 *linear operator.* 若 $V=U\oplus W$, 其中 U,W 為 *T-invariant subspace,* 則

$$\chi_T(x)=\chi_{T|_U}(x)\chi_{T|_W}(x).$$

Proof. 選定 U,W 的 ordered basis $\beta_1=(\mathbf{u}_1,\dots,\mathbf{u}_l),\beta_2=(\mathbf{w}_1,\dots,\mathbf{w}_m)$, 可得 V 的 ordered basis $\beta=(\mathbf{u}_1,\dots,\mathbf{u}_l,\mathbf{w}_1,\dots,\mathbf{w}_m)$. 此時利用等式 (3.4) 我們有

$$xI_{l+m}-[T]_\beta=\begin{pmatrix}xI_l-[T|_U]_{\beta_1} & \mathbf{O}\\ \mathbf{O} & xI_m-[T|_W]_{\beta_2}\end{pmatrix}.$$

利用上面所述有關於 block diagonal matrix 的 determinant 算法得

$$\chi_T(x)=\det(xI_{l+m}-[T]_\beta)=\det(xI_l-[T|_U]_{\beta_1})\det(xI_m-[T|_W]_{\beta_2})=\chi_{T|_U}(x)\chi_{T|_W}(x).$$

□

至於 minimal polynomial, 我們需要在複習一下代數有關於 $F[x]$ 這一個 polynomial ring 的性質. 因為 F 是一個 field, $F[x]$ 上的元素有除法的性質, 即給定 $f(x),g(x)\in F[x]$, 若 $g(x)\neq 0$, 則存在 $h(x),r(x)\in F[x]$ 其中 $\deg(r(x))<\deg(g(x))$ 使得 $f(x)=g(x)h(x)+r(x)$. 這個性質使得 $F[x]$ 成為所謂的 Euclidean domain. 所以 $F[x]$ 會是一個 principle ideal domain, 也因此是一個 unique factorization domain. 換言之, 任取 $f(x)\in F[x]$, 我們都可以將 $f(x)$ 唯一寫成一些 irreducible polynomial 的乘積. 所以任取兩個 $F[x]$ 上的 polynomial $f(x),g(x)$, 我們可以定義它們的最高公因式 (用 $\gcd(f(x),g(x))$ 表示) 以及最低公倍式 (用 $\operatorname{lcm}(f(x),g(x))$ 表示). 注意, 這裡為了要有唯一性 $\gcd(f(x),g(x)),\operatorname{lcm}(f(x),g(x))$ 我們都選 monic polynomial. 若令 $l(x)=\operatorname{lcm}(f(x),g(x))$, 則我們有以下性質:

(1) $f(x)\mid l(x), g(x)\mid l(x)$.

(2) 若 $h(x)\in F[x]$ 則 $f(x)\mid h(x), g(x)\mid h(x)\Leftrightarrow l(x)\mid h(x)$.

利用這個性質我們可以得到以下有關 minimal polynomials 間的關係.

Lemma 3.5.6. 假設 V 為 *finite dimensional F-space,* $T:V\to V$ 為 *linear operator.* 若 $V=U\oplus W$, 其中 U,W 為 *T-invariant subspace,* 則

$$\mu_T(x)=\operatorname{lcm}(\mu_{T|_U}(x),\mu_{T|_W}(x)).$$

Proof. 令 $l(x)=\operatorname{lcm}(\mu_{T|_U}(x),\mu_{T|_W}(x))$. 由 Lemma 3.5.4 得 $\mu_{T|_U}(x)\mid\mu_T(x),\mu_{T|_W}(x)\mid\mu_T(x)$, 故知 $l(x)\mid\mu_T(x)$.

另一方面, 對於任意 $\mathbf{v}\in V$, 存在 $\mathbf{u}\in U,\mathbf{w}\in W$ 使得 $\mathbf{v}=\mathbf{u}+\mathbf{w}$, 故由等式 (3.3) 知

$$l(T)(\mathbf{v})=l(T)(\mathbf{u})+l(T)(\mathbf{w})=l(T)|_U(\mathbf{u})+l(T)|_W(\mathbf{w})=l(T|_U)(\mathbf{u})+l(T|_W)(\mathbf{w}).$$

然而 $\mu_{T|_U}(x)\mid l(x),\mu_{T|_W}(x)\mid l(x)$, 故知 $l(T|_U)=\mathbf{O},l(T|_W)=\mathbf{O}$, 亦即 $l(T|_U)(\mathbf{u})=\mathbf{O}_U=\mathbf{O}_V$ 且 $l(T|_W)(\mathbf{w})=\mathbf{O}_W=\mathbf{O}_V$. 由此知 $l(T)(\mathbf{v})=\mathbf{O}_V,\forall\mathbf{v}\in V$, 得證 $l(T)=\mathbf{O}$. 故由 Lemma 3.3.5 知 $\mu_T(x)\mid l(x)$. 因此由 $l(x)\mid\mu_T(x)$ 且 $\mu_T(x)\mid l(x)$ 以及 $\mu_T(x),l(x)$ 皆為 monic polynomial, 得證 $\mu_T(x)=l(x)=\operatorname{lcm}(\mu_{T|_U}(x),\mu_{T|_W}(x))$. □

現在我們來說明如何將 V 寫成 T-invariant subspaces 的 direct sum. 由於 $F[x]$ 是一個 principle ideal domain (P.I.D.), 給定 $f(x),g(x)\in F[x]$, 我們可以考慮 $f(x),g(x)$ 所生成的 ideal $(f(x),g(x))$, 這個 ideal 中的元素都是 $a(x)f(x)+b(x)g(x)$ (其中 $a(x),b(x)\in F[x]$) 這樣的形式. 因為 $F[x]$ 是 P.I.D. 所以存在 $d(x)\in F[x]$ 使得 $(f(x),g(x))=(d(x))$. 亦即 $(f(x),g(x))$ 中的元素, 都可以寫成 $h(x)d(x)$ 的形式. 因為 $f(x)\in(f(x),g(x))$, 所以 $d(x)\mid f(x)$, 同理 $d(x)\mid g(x)$. 另外又 $d(x)\in(f(x),g(x))$ 所以存在 $a(x),b(x)\in F[x]$ 使得 $d(x)=a(x)f(x)+b(x)g(x)$. 由此可知若 $h(x)\mid f(x),h(x)\mid g(x)$ 則 $h(x)\mid a(x)f(x)+b(x)g(x)$, 即 $h(x)\mid d(x)$. 可以看出其實 $d(x)$ 就是 $f(x),g(x)$ 的最高公因式, 即 $d(x)=\gcd(f(x),g(x))$. 我們將 $d(x)=\gcd(f(x),g(x))$ 的性質列出如下:

(1) $d(x)\mid f(x), d(x)\mid g(x)$.

(2) 若 $h(x)\in F[x]$ 則 $h(x)\mid f(x),h(x)\mid g(x)\Leftrightarrow h(x)\mid d(x)$.

(3) 存在 $a(x),b(x)\in F[x]$ 使得 $d(x)=a(x)f(x)+b(x)g(x)$.

特別地, 當 $f(x),g(x)$ 沒有共同的質因式時, 我們稱為 *relatively prime*, 此時 $\gcd(f(x),g(x))=1$, 故存在 $a(x),b(x)\in F[x]$ 使得 $a(x)f(x)+b(x)g(x)=1$.

Theorem 3.5.7. 假設 V 為 *finite dimensional F-space*, $T:V\to V$ 為 *linear operator* 且 $\mu_T(x)=f(x)g(x)$, 其中 $f(x),g(x)\in F[x]$ 為 *monic polynomials* 且 *relatively prime*. 若令 $U=\mathrm{Ker}(f(T))$, $W=\mathrm{Ker}(g(T))$, 則 V 可以寫成 *T-invariant subspaces* U,W 的 *internal direct sum*, 即 $V=U\oplus W$, 而且 $\mu_{T|_U}(x)=f(x)$ 以及 $\mu_{T|_W}(x)=g(x)$.

Proof. 我們已知 U,W 為 T-invariant subspaces. 現在要證明 $V=U+W$ 而且 $U\cap W=\{\mathbf{O}_V\}$. 首先因 $f(x),g(x)$ 為 relatively prime, 故存在 $a(x),b(x)\in F[x]$ 使得 $a(x)f(x)+b(x)g(x)=1$. 因此知 $a(T)\circ f(T)+b(T)\circ g(T)=\mathrm{id}$. 亦即對任意 $\mathbf{v}\in V$, 我們有

$$\mathbf{v}=a(T)\circ f(T)(\mathbf{v})+b(T)\circ g(T)(\mathbf{v}). \tag{3.5}$$

令 $\mathbf{w}=a(T)\circ f(T)(\mathbf{v}),\mathbf{u}=b(T)\circ g(T)(\mathbf{v})$, 此時利用 Lemma 3.2.3 得

$$f(T)(\mathbf{u})=f(T)\circ(b(T)\circ g(T))(\mathbf{v})=b(T)\circ(f(T)\circ g(T))(\mathbf{u})=b(T)\circ\mu_T(T)(\mathbf{v}).$$

然而 $\mu_T(T)=\mathbf{O}$, 故知 $f(T)(\mathbf{u})=\mathbf{O}_V$, 亦即 $\mathbf{u}\in\mathrm{Ker}(f(T))$. 同理可得 $\mathbf{w}\in\mathrm{Ker}(g(T))$. 得證 $V=\mathrm{Ker}(f(T))+\mathrm{Ker}(g(T))=U+W$.

現若 $\mathbf{v}\in U\cap W=\mathrm{Ker}(f(T))\cap\mathrm{Ker}(g(T))$, 表示 $f(T)(\mathbf{v})=g(T)(\mathbf{v})=\mathbf{O}_V$. 故由等式 (3.5) 得 $\mathbf{v}=a(T)(\mathbf{O}_V)+b(T)(\mathbf{O}_V)=\mathbf{O}_V$. 得證 $U\cap W=\{\mathbf{O}_V\}$.

現考慮 minimal polynomial. 由於 $U=\mathrm{Ker}(f(T))$, 故 $f(T)|_U=\mathbf{O}$. 因此由等式 (3.3) 得 $f(T|_U)=\mathbf{O}$. 再由 Lemma 3.3.5 得 $\mu_{T|_U}(x)\mid f(x)$. 同理得 $\mu_{T|_W}(x)\mid g(x)$. 但 $f(x),g(x)$ 為 relatively prime, 故知 $\mu_{T|_U}(x),\mu_{T|_W}(x)$ 亦為 relatively prime, 得

$$\mathrm{lcm}(\mu_{T|_U}(x),\mu_{T|_W}(x))=\mu_{T|_U}(x)\mu_{T|_W}(x).$$

因此由 Lemma 3.5.6 得

$$\mu_{T|_U}(x)\mu_{T|_W}(x)=\mu_T(x)=f(x)g(x).$$

故再由 $\mu_{T|_U}(x)\mid f(x)$ 以及 $\mu_{T|_W}(x)\mid g(x)$ 得證 $\mu_{T|_U}(x)=f(x)$ 以及 $\mu_{T|_W}(x)=g(x)$. □

$F[x]$ 是一個 unique factorization domain (U.F.D.), 表示 $F[x]$ 中的非常數多項式都可以唯一寫成一些 irreducible polynomials 的乘積. 因此對於 linear operator T 的 minimal polynomial, 我們可以找到相異的 monic irreducible polynomials $p_1(x),\dots,p_k(x)$ 使得 $\mu_T(x)=p_1(x)^{m_1}\cdots p_k(x)^{m_k}$, 其中 $m_1,\dots,m_k\in\mathbb{N}$. 由於 characteristic polynomial $\chi_T(x)$ 和 $\mu_T(x)$ 有相同的質因式 (Theorem 3.3.9) 且 $\mu_T(x)\mid\chi_T(x)$, 我們知道 $\chi_T(x)=p_1(x)^{c_1}\cdots p_k(x)^{c_k}$, 其中 $c_i\in\mathbb{N}$ 且 $c_i\geq m_i$.

Theorem 3.5.8 (Primary Decomposition Theorem)**.** 假設 V 是 *dimension* 為 n 的 *F-space*, $T:V\to V$ 為 *linear operator* 且

$$\mu_T(x)=p_1(x)^{m_1}\cdots p_k(x)^{m_k}\quad\text{and}\quad\chi_T(x)=p_1(x)^{c_1}\cdots p_k(x)^{c_k}$$

其中 $p_1(x),\dots,p_k(x)$ 為相異的 *monic irreducible polynomials.* 若令 $V_i=\mathrm{Ker}(p_i(T)^{\circ m_i})$, *for* $i=1,\dots,k$, 則

$$V=V_1\oplus\cdots\oplus V_k$$

且

$$\mu_{T|_{V_i}}(x)=p_i(x)^{m_i}\quad\text{and}\quad\chi_{T|_{V_i}}(x)=p_i(x)^{c_i},\forall i=1,\dots,k.$$

Proof. 我們對 $\mu_T(x)$ 的相異 moinic irreducible divisor (質因式) 的個數 k 作數學歸納法. 若 $k=1$, 表示 $\mu_T(x)=p_1(x)^{m_1}$, 此時因 $\mu_T(T)=p_1(T)^{\circ m_1}=\mathbf{O}$, 故知 $V=\mathrm{Ker}(p_1(T)^{\circ m_1})$, 因此在 $k=1$ 時定理成立. 現假設當 $\mu_T(x)$ 有 $k-1$ 個相異 monic irreducible divisor 時亦成立, 我們考慮 $\mu_T(x)=p_1(x)^{m_1}\cdots p_k(x)^{m_k}$ 的情形. 此時令 $f(x)=p_1(x)^{m_1}$, $g(x)=p_2(x)^{m_2}\cdots p_k(x)^{m_k}$, 因 $f(x),g(x)$ 為 relatively prime, 由 Theorem 3.5, 我們知 $V=U\oplus W$, 其中 $U=\mathrm{Ker}(f(T))=\mathrm{Ker}(p_1(T)^{\circ m_1})$, $W=\mathrm{Ker}(g(T))$ 而且 $\mu_{T|_U}(x)=p_1(x)^{m_1}$, $\mu_{T|_W}(x)=p_2(x)^{m_2}\cdots p_k(x)^{m_k}$. 現考慮 vector space W 以及 linear operator $T|_W:W\to W$, 套用 induction 在 $k-1$ 情形的假設知 $W=V_2\oplus\cdots\oplus V_k$ 其中 $V_i=\mathrm{Ker}(p_i(T|_W)^{\circ m_i})$ 且 $\mu_{(T|_W)|_{V_i}}(x)=p_i(x)^{m_i}$, $\forall i=2,\dots,k$. 然而當 $i=2,\dots,k$ 時 $p_i(x)^{m_i}\mid g(x)$, 故 $\mathrm{Ker}(p_i(T)^{\circ m_i})\subseteq\mathrm{Ker}(g(T))=W$, 因此

$$V_i=\mathrm{Ker}(p_i(T|_W)^{\circ m_i})=\mathrm{Ker}(p_i(T)^{\circ m_i})\cap W=\mathrm{Ker}(p_i(T)^{\circ m_i}).$$

另一方面, 因 $V_i\subseteq W$, 我們有 $(T|_W)|_{V_i}=T|_{V_i}$ 故得

$$\mu_{T|_{V_i}}(x)=\mu_{(T|_W)|_{V_i}}(x)=p_i(x)^{m_i}.$$

故令 $U=V_1$, 利用 Corollary 3.4.7, 得證 $V=U\oplus W=V_1\oplus V_2\oplus\cdots\oplus V_k$.

至於 characteristic polynomial, 利用 Theorem 3.3.9 我們知 $\chi_{T|_{V_i}}(x)=p_i(x)^{e_i}$ 其中 $e_i\geq m_i$. 因此由 Lemma 3.5.5 知

$$p_1(x)^{c_1}\cdots p_k(x)^{c_k}=\chi_T(x)=\chi_{T|_{V_1}}(x)\cdots\chi_{T|_{V_k}}(x)=p_1(x)^{e_1}\cdots p_k(x)^{e_k},$$

利用 $F[x]$ 為 U.F.D. 得證 $e_i=c_i$, 即 $\chi_{T|_{V_i}}(x)=p_i(x)^{c_i},\forall i=1,\dots,k.$ □

回顧一下, 對於 linear operator $T:V\to V$, 要找到 $\mathrm{Ker}(T)$, 我們可以利用 V 的 ordered basis β, 先得到 representative matrix $[T]_\beta$. 再求 $[T]_\beta$ 的 null space $N([T]_\beta)$ (我們用 $N(A)$ 表示矩陣 A 的 null space). 接著將 null space 的元素用 $\tau_\beta^{\circ -1}$ 還原成 V 的元素, 就得到 $\mathrm{Ker}(T)$ 的元素了. 我們看以下 primary decomposition 的例子.

Example 3.5.9. 考慮 Example 3.3.10 中的 linear operator $T:P_2(\mathbb{R})\to P_2(\mathbb{R})$. 我們要考慮它的 primary decomposition. 在 Example 3.3.10 中我們知道 T 的 minimal polynomial 為 $\mu_T(x)=(x+1)^2(x-2)$, 因此我們必須找出 $V_1=\mathrm{Ker}((T+\mathrm{id})^{\circ 2})$ 和 $V_2=\mathrm{Ker}(T-2\,\mathrm{id})$. 利用 representative matrix 可以幫助我們找到這兩個 T-invariant subspaces. 我們仍然沿用 ordered basis $\beta=(-x^2+x+1,x+1,1)$. 首先考慮 $([T]_\beta+I_3)^2$ 的 null space, 即解

$$\begin{pmatrix}-9&-9&0\\18&18&0\\0&0&0\end{pmatrix}\cdot\begin{pmatrix}x_1\\x_2\\x_3\end{pmatrix}=\begin{pmatrix}0\\0\\0\end{pmatrix},\quad i.e.\quad\begin{cases}-9x_1&-9x_2&=&0\\18x_1&+18x_2&=&0\end{cases}.$$

得知 $N(([T]_\beta+I_3)^2)=\mathrm{Span}((1,-1,0)^{\mathrm{t}},(0,0,1)^{\mathrm{t}})$ 故得 $V_1=\mathrm{Ker}((T+\mathrm{id})^{\circ 2})=\mathrm{Span}(x^2,1)$. 同理 $N([T]_\beta-2I_3)=\mathrm{Span}((1,-2,0)^{\mathrm{t}})$, 故得 $V_2=\mathrm{Ker}(T-2\,\mathrm{id})=\mathrm{Span}(x^2+x+1)$. 很容易驗證 V_1,V_2 皆為 T-invariant subspace 且 $V=V_1\oplus V_2$. 若令 $\beta'=(x^2,1,x^2+x+1)$, 則因

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

得

$$[T]_{\beta'}=\begin{pmatrix}-1&2&0\\0&-1&0\\0&0&2\end{pmatrix}.$$

考慮 $(x^2,1)$ 為 V_1 的 ordered basis, 則 $T|_{V_1}$ 的 representative matrix 為 $\begin{pmatrix}-1&2\\0&-1\end{pmatrix}$, 故得 $\chi_{T|_{V_1}}(x)=\mu_{T|_{V_1}}(x)=(x+1)^2$. 同理可得 $\chi_{T|_{V_2}}(x)=\mu_{T|_{V_2}}(x)=(x-2)$.

Primary Decomposition Theorem 告訴我們, 若 linear operator $T:V\to V$ 的 characteristic polynomial (或 minimal polynomial) 是 $p_1(x)^{c_1}\cdots p_k(x)^{c_k}$, 則我們可以找到 V 的 ordered basis β, 使得 $[T]_\beta$ 為以下的 block diagonal matrix

$$\begin{pmatrix}A_1&&\mathbf{O}\\&\ddots&\\\mathbf{O}&&A_k\end{pmatrix},\tag{3.6}$$

其中每個 A_i 的 characteristic polynomial 為 $\chi_{A_i}(x)=p_i(x)^{c_i}$. 因此以後我們只要個別探討 linear operator 其 characteristic polynomial 為 $p(x)^c$ (其中 $p(x)$ 為 monic irreducible, $c\in\mathbb{N}$) 這種情形就可以了.

對於 $n\times n$ 方陣 $A\in M_n(F)$, 我們也可以利用 linear operator 的 primary decomposition 的概念找到 invertible matrix $P\in M_n(F)$ 使得 $P^{-1}\cdot A\cdot P$ 為如 (3.6) 的 block diagonal matrix. 我們可以將 A 看成是 linear transformation $T:F^n\to F^n$ 其定義為 $T(\mathbf{x})=A\mathbf{x}$. 此時 A 便是 T 對於標準基底 ε 的 representative matrix $[T]_\varepsilon$. 利用 Primary Decomposition Theorem, 我們可以找到 F^n 的 ordered basis β 使得 $[T]_\beta$ 為 block diagonal matrix. 然而由 Proposition 2.4.6 知

$$[T]_\beta={}_\beta[\mathrm{id}]_\varepsilon\cdot[T]_\varepsilon\cdot{}_\varepsilon[\mathrm{id}]_\beta={}_\varepsilon[\mathrm{id}]_\beta^{-1}\cdot A\cdot{}_\varepsilon[\mathrm{id}]_\beta,$$

所以可以令 $P={}_\varepsilon[\mathrm{id}]_\beta$. 也就是說若將 ordered basis β 一個 column 一個 column (column by column) 的依序排成的 $n\times n$ matrix 就是我們想要的 P. 因此我們的步驟如下: 首先求得 $\mu_A(x)$ 並將之分解成相異的 monic irreducible polynomials 的乘積 $\mu_A(x)=p_1(x)^{m_1}\cdots p_k(x)^{m_k}$. 接下來便是求每一個 $p_i(A)^{m_i}$ 的 null space $N(p_i(A)^{m_i})$ (此即對應到 $\mathrm{Ker}(p_i(T)^{\circ m_i})$). 得到每個 null space 的 basis 後, 將之 column by column 的依序排成矩陣 P 即可. 我們看以下的例子.

Example 3.5.10. 考慮 5×5 matrix

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

我們要將之化為 block diagonal matrix. 首先求得 $\chi_A(x)=\mu_A(x)=(x-1)^3(x-2)^2$. 利用 Primary Decomposition Theorem 我們知道 A 可以化為有兩個 blocks 的 block diagonal matrix, 其中一個是 3×3 matrix 另一個是 2×2 matrix. 首先求得

$$N((A-I_5)^3)=\text{Span}((-1,0,0,0,1)^{\text{t}},(-2,0,0,1,0)^{\text{t}},(-2,0,1,0,0)^{\text{t}})$$

$$N(A-2I_5)^2=\text{Span}((1,-1,1,0,0)^{\text{t}},(1,0,0,0,0)^{\text{t}}).$$

若令

$$\mathbf{v}_1=\begin{pmatrix} -1\\ 0\\ 0\\ 0\\ 1 \end{pmatrix},\mathbf{v}_2=\begin{pmatrix} -2\\ 0\\ 0\\ 1\\ 0 \end{pmatrix},\mathbf{v}_3=\begin{pmatrix} -2\\ 0\\ 1\\ 0\\ 0 \end{pmatrix},\mathbf{v}_4=\begin{pmatrix} 1\\ -1\\ 1\\ 0\\ 0 \end{pmatrix},\mathbf{v}_5=\begin{pmatrix} 1\\ 0\\ 0\\ 0\\ 0 \end{pmatrix},$$

則因

$$A\mathbf{v}_1=\begin{pmatrix} -2\\ 0\\ 0\\ 1\\ 0 \end{pmatrix}=\mathbf{v}_2,\quad A\mathbf{v}_2=\begin{pmatrix} -3\\ 0\\ 1\\ 1\\ -1 \end{pmatrix}=-\mathbf{v}_1+\mathbf{v}_2+\mathbf{v}_3,\quad A\mathbf{v}_3=\begin{pmatrix} -3\\ 0\\ 2\\ 0\\ -1 \end{pmatrix}=-\mathbf{v}_1+2\mathbf{v}_3,$$

$$A\mathbf{v}_4=\begin{pmatrix} 2\\ -1\\ 1\\ 0\\ 0 \end{pmatrix}=\mathbf{v}_4+\mathbf{v}_5,\quad A\mathbf{v}_5=\begin{pmatrix} 2\\ 1\\ -1\\ 0\\ 0 \end{pmatrix}=-\mathbf{v}_4+3\mathbf{v}_5,$$

取

$$P=\begin{pmatrix} -1 & -2 & -2 & 1 & 1\\ 0 & 0 & 0 & -1 & 0\\ 0 & 0 & 1 & 1 & 0\\ 0 & 1 & 0 & 0 & 0\\ 1 & 0 & 0 & 0 & 0 \end{pmatrix}$$

可得 block diagonal matrix

$$P^{-1}\cdot A\cdot P=\begin{pmatrix} 0 & -1 & -1 & 0 & 0\\ 1 & 1 & 0 & 0 & 0\\ 0 & 1 & 2 & 0 & 0\\ 0 & 0 & 0 & 1 & -1\\ 0 & 0 & 0 & 1 & 3 \end{pmatrix}.$$

若令

$$B=\begin{pmatrix} 0 & -1 & -1 \\ 1 & 1 & 0 \\ 0 & 1 & 2 \end{pmatrix} \quad C=\begin{pmatrix} 1 & -1 \\ 1 & 3 \end{pmatrix},$$

我們有 $\chi_B(x)=\mu_B(x)=(x-1)^3$ 以及 $\chi_C(x)=\mu_C(x)=(x-2)^2$.