4.4. Rational Form

當 V 是 over F 的 vector space 且 $T: V \to V$ 為 over F 的 linear operator. 若 F 不是 algebraically closed, 則 T 的 characteristic polynomial 並不一定可以完全分解成 $F[x]$ 上的一次多項式的乘積. 我們將探討在這種情形之下, 如何找到合適的 V 的 ordered basis 使得 T 的 representative matrix 可為較簡單的形式.

要將 T 的 representative matrix 化為簡單的 form, 就必須將 V 寫成一些 T-invariant subspaces 的 direct sum. 寫成越多維度小的 T-invariant subspaces 的 direct sum, T 的 representative matrix 就可寫成越簡單的 form. 例如 T 是 diagonalizable 時, 就表示 V 可以寫成一些 1-dimensional T-invariant subspaces 的 direct sum. 接下來我們就是要討論若 $\mathbf{v} \in V$, 則包含 $\mathbf{v}$ 最小的 T-invariant subspace 為何.

假設 W 為包含 $\mathbf{v}$ 的 T-invariant subspace. 因為 $\mathbf{v} \in W$, 故由 W 為 T-invariant, 得 $T(\mathbf{v}) \in W$. 同理得 $T^{\circ 2}(\mathbf{v}) \in W$, $T^{\circ 3}(\mathbf{v}) \in W$, ..., 故由數學歸納法得 $T^{\circ i}(\mathbf{v}) \in W$, $\forall i \in \mathbb{N}$. 再由 W 為 over F 的 vector space, 可得對任意 $a_d, a_{d-1}, \ldots, a_1, a_0 \in F$ 皆有

$$a_d T^{\circ d}(\mathbf{v}) + a_{d-1} T^{\circ d-1}(\mathbf{v}) + \cdots + a_1 T(\mathbf{v}) + a_0 \mathbf{v} \in W.$$

換言之, 對於所有 $f(x) \in F[x]$ 皆有 $f(T)(\mathbf{v}) \in W$. 現考慮

$$C_{\mathbf{v}} = \{f(T)(\mathbf{v}) \mid f(x) \in F[x]\},$$

我們有 $C_{\mathbf{v}} \subseteq W$. 很容易檢查 $C_{\mathbf{v}}$ 為一個 vector space, 故 $C_{\mathbf{v}}$ 為 W 的 subspace. 又對於任意 $\mathbf{w} \in C_{\mathbf{v}}$, 我們有 $\mathbf{w} = a_d T^{\circ d}(\mathbf{v}) + \cdots + a_1 T(\mathbf{v}) + a_0 \mathbf{v}$, 其中 $a_d, \ldots, a_1, a_0 \in F$. 所以

$$T(\mathbf{w}) = a_d T^{\circ d+1}(\mathbf{v}) + \cdots + a_1 T^{\circ 2}(\mathbf{v}) + a_0 T(\mathbf{v}) = g(T)(\mathbf{v}) \in C_{\mathbf{v}},$$

其中 $g(x) = a_d x^{d+1} + \cdots + a_1 x^2 + a_0 x \in F[x]$. 這說明了 $C_{\mathbf{v}}$ 是一個 T-invariant subspace. 因此我們了解一個包含 $\mathbf{v}$ 的 T-invariant subspace, 一定包含 $C_{\mathbf{v}}$ 這一個 T-invariant subspace. 也就是說 $C_{\mathbf{v}}$ 是包含 $\mathbf{v}$ 最小的 T-invariant subspace. 我們有以下之定義.

Definition 4.4.1. 假設 V 是一個 F-space 且 $T: V \to V$ 是一個 F-linear operator. 給定 $\mathbf{v} \in V$, 考慮 $C_{\mathbf{v}} = \{f(T)(\mathbf{v}) \mid f(x) \in F[x]\}$. 我們稱 $C_{\mathbf{v}}$ 為一個 T-*cyclic subspace spanned* by $\mathbf{v}$. 其中 $\mathbf{v}$ 也稱作 $C_{\mathbf{v}}$ 的一個 *cyclic vector*.

要注意 $C_{\mathbf{v}}$ (the T-cyclic subspace spanned by $\mathbf{v}$) 和 T 有關, 由於我們不會探討不同 linear operator 間的關係, 所以為了符號簡便我們省略 $C_{\mathbf{v}}$ 中有關 T 的標記. 一般來說 $C_{\mathbf{v}}$ 並不是 the subspace spanned by $\mathbf{v}$, 而且一個 T-cyclic subspace 的 cyclic vector 並不唯一.

Question 4.14. *除了* $\mathbf{v}$ *以外, 你能找到另一個* $C_{\mathbf{v}}$ *的 cyclic vector 嗎?*

Question 4.15. *在甚麼情況之下* $C_{\mathbf{v}}$ *會等於* Span($\mathbf{v}$)*?*

接下來我們要更進一步來了解 $C_{\mathbf{v}}$ 這一個 T-cyclic subspace. 首先由於我們探討的 vector space V 是 finite dimensional, 所以 $C_{\mathbf{v}}$ 也是一個 finite dimensional vector space. 因此 $\{\mathbf{v}, T(\mathbf{v}), T^{\circ 2}(\mathbf{v}), \ldots, T^{\circ i}(\mathbf{v}), \ldots\}$ 為 linearly dependent. 也就是說存在 $k \in \mathbb{N}$, 以及

$a_0, a_1, \ldots, a_k \in F$ 不全為 0 使得 $a_k T^{\circ k}(\mathbf{v}) + \cdots + a_1 T(\mathbf{v}) + a_0 \mathbf{v} = \mathbf{O}_V$. 這告訴我們, 存在非零多項式 $f(x) \in F[x]$ 使得 $f(T)(\mathbf{v}) = \mathbf{O}_V$. 滿足這個性質的次數最小的 monic polynomial 對於我們了解 $C_{\mathbf{v}}$ 有很大的用處, 所以有以下之定義.

Definition 4.4.2. 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 $f(T)(\mathbf{v}) = \mathbf{O}_V$, 且次數最小的 monic polynomial 稱為 the T-*annihilator* of $\mathbf{v}$, 我們用 $\mu_{\mathbf{v}}(x)$ 來表示.

再次強調 $\mu_{\mathbf{v}}(x)$ 不只和 $\mathbf{v}$ 有關和 T 也有關, 不過由於我們只探討單一的 linear operator, 所以省略有關 T 的標記.

Question 4.16. 對於任意的 *linear operator* $T: V \to V$, 甚麼是 *the* T-*annihilator of* $\mathbf{O}_V$?

利用多項式的除法原理 (division algorithm), $\mathbf{v}$ 的 T-annihilator 和 Lemma 3.3.5 中有關於 T 的 minimal polynomial 有著類似的性質, 由於證明方法相同, 這裡就不再贅述.

Lemma 4.4.3. 假設 V 為一個 *finite dimensional* F-*space*, $\mathbf{v} \in V$ 且 $T: V \to V$ 為一個 *linear operator.* 則對於 $f(x) \in F[x]$, $f(T)(\mathbf{v}) = \mathbf{O}_V$ 若且唯若 $\mu_{\mathbf{v}}(x) \mid f(x)$.

利用 Lemma 4.4.3, 我們馬上知道 $\mu_{\mathbf{v}}(x) \mid \chi_T(x)$ 且 $\mu_{\mathbf{v}}(x) \mid \mu_T(x)$, 這是由於 $\chi_T(T) = \mu_T(T) = \mathbf{O}$, 所以對任意 $\mathbf{v} \in V$ 皆有 $\chi_T(T)(\mathbf{v}) = \mu_T(T)(\mathbf{v}) = \mathbf{O}_V$.

我們可以透過 $\mu_{\mathbf{v}}(x)$ 來了解 $C_{\mathbf{v}}$. 事實上我們有以下之結果.

Theorem 4.4.4. 設 V 為一個 *finite dimensional* F-*space,* $T: V \to V$ 為一個 *linear operator.* 若 $\mathbf{v} \in V$, 且其 T-*annihilator* 為

$$\mu_{\mathbf{v}}(x) = x^d + a_{d-1}x^{d-1} + \cdots + a_1 x + a_0,$$

則

$$\{\mathbf{v}, T(\mathbf{v}), \ldots, T^{\circ d-1}(\mathbf{v})\}$$

為 $C_{\mathbf{v}}$ 的一組 *basis.* 另外考慮 T 限制在 $C_{\mathbf{v}}$ 下的 *linear operator* $T|_{C_{\mathbf{v}}}: C_{\mathbf{v}} \to C_{\mathbf{v}}$, 我們有 $T|_{C_{\mathbf{v}}}$ 的 *characteristic polynomial* 和 *minimal polynomial* 皆等於 $\mathbf{v}$ 的 T-*annihilator,* 亦即 $\chi_{T|_{C_{\mathbf{v}}}}(x) = \mu_{T|_{C_{\mathbf{v}}}}(x) = \mu_{\mathbf{v}}(x)$.

Proof. 首先證明 $S = \{\mathbf{v}, T(\mathbf{v}), \ldots, T^{\circ d-1}(\mathbf{v})\}$ 為 linearly independent. 若 S 不是 linearly independent, 表示存在次數小於等於 $d-1$ 的 polynomial $f(x) \in F[x]$ 滿足 $f(T)(\mathbf{v}) = \mathbf{O}_V$. 此和 $\mu_{\mathbf{v}}(x)$ 為次數最小的定義相矛盾, 故得證 S 為 linearly independent.

接著證明 $C_{\mathbf{v}} = \mathrm{Span}(S)$. 首先很容易觀察

$$\mathrm{Span}(S) = \{g(T)(\mathbf{v}) | g(x) \in F[x], \deg(g(x)) \leq d-1\},$$

故知 $\mathrm{Span}(S) \subseteq C_{\mathbf{v}}$. 然而依定義對於任意 $\mathbf{w} \in C_{\mathbf{v}}$, 皆存在 $g(x) \in F[x]$ 使得 $g(T)(\mathbf{v}) = \mathbf{w}$. 現若 $\deg(g(x)) \leq d-1$, 則可得 $\mathbf{w} \in \mathrm{Span}(S)$. 而若 $\deg(g(x)) > d-1$, 則由除法原理, 存在 $h(x), r(x) \in F[x]$ 其中 $\deg(r(x)) \leq d-1$ 使得 $g(x) = h(x)\mu_{\mathbf{v}}(x) + r(x)$. 故由 $\mu_{\mathbf{v}}(T)(\mathbf{v}) = \mathbf{O}_V$, 知

$$\mathbf{w} = g(T)(\mathbf{v}) = h(T)(\mu_{\mathbf{v}}(T)(\mathbf{v})) + r(T)(\mathbf{v}) = h(T)(\mathbf{O}_V) + r(T)(\mathbf{v}) = r(T)(\mathbf{v}),$$

得證 $\mathbf{w} \in \mathrm{Span}(S)$. 此證得 $C_{\mathbf{v}} \subseteq \mathrm{Span}(S)$.

現考慮 $T|_{C_{\mathbf{v}}}$ 的 characteristic polynomial $\chi_{T|_{C_{\mathbf{v}}}}(x)$, 依定義 $\deg(\chi_{T|_{C_{\mathbf{v}}}}(x)) = \dim(C_{\mathbf{v}})$. 而由 S 為 $C_{\mathbf{v}}$ 的一組 basis, 得 $\dim(C_{\mathbf{v}}) = d$. 又 $\mathbf{v} \in C_{\mathbf{v}}$, 故依定義 $\chi_{T|_{C_{\mathbf{v}}}}(T)(\mathbf{v}) = \mathbf{O}_V$. 所以 Lemma 4.4.3 告訴我們 $\mu_{\mathbf{v}}(x) \mid \chi_{T|_{C_{\mathbf{v}}}}(x)$. 最後由 $\mu_{\mathbf{v}}(x)$ 以及 $\chi_{T|_{C_{\mathbf{v}}}}(x)$ 皆為 monic 以及它們的 degree 皆為 d 得證 $\chi_{T|_{C_{\mathbf{v}}}}(x) = \mu_{\mathbf{v}}(x)$. 同理 $\mu_{T|_{C_{\mathbf{v}}}}(T)(\mathbf{v}) = \mathbf{O}_V$, 故得 $\mu_{\mathbf{v}}(x) \mid \mu_{T|_{C_{\mathbf{v}}}}(x)$. 再由 $\deg(\mu_{T|_{C_{\mathbf{v}}}}(x)) \le \deg(\chi_{T|_{C_{\mathbf{v}}}}(x)) = \deg(\mu_{\mathbf{v}}(x))$ 得證 $\mu_{T|_{C_{\mathbf{v}}}}(x) = \mu_{\mathbf{v}}(x)$. $\square$

Question 4.17. 若 $\deg(\mu_{\mathbf{v}}(x)) = d$, 你能證明 $C_{\mathbf{v}}$ 中的元素皆可 "唯一" 寫成 $g(T)(\mathbf{v})$ 其中 $g(x) \in F[x]$ 且 $\deg(g(x)) \le d-1$ 嗎?

Theorem 4.4.4 中 $C_{\mathbf{v}}$ 的這組 basis 很重要, 我們有以下的定義.

Definition 4.4.5. 假設 V 為一個 finite dimensional F-space, $\mathbf{v} \in V$ 且 $T : V \to V$ 為一個 linear operator. 設 $\deg(\mu_{\mathbf{v}}(x)) = d$, 我們稱 $\{\mathbf{v}, T(\mathbf{v}), \ldots, T^{\circ d-1}(\mathbf{v})\}$ 為 $C_{\mathbf{v}}$ 的一組 *cyclic basis*.

當 $\deg(\mu_{\mathbf{v}}(x)) = d$ 時, 我們可以考慮 $T|_{C_{\mathbf{v}}} : C_{\mathbf{v}} \to C_{\mathbf{v}}$ 對於 $\beta = (\mathbf{v}, T(\mathbf{v}), \ldots, T^{\circ d-1}(\mathbf{v}))$ 這一個 cyclic basis 所形成的 ordered basis 的 representative matrix 為何. 由於 $T(\mathbf{v}) = 0\mathbf{v} + 1T(\mathbf{v}) + 0T^{\circ 2}(\mathbf{v}) + \cdots + 0T^{\circ d-1}(\mathbf{v})$, 我們知道此 matrix 的第一個 column 應為 $(0,1,0,\ldots,0)^{\mathrm{t}}$, 同理因 $T(T(\mathbf{v})) = 0\mathbf{v} + 0T(\mathbf{v}) + 1T^{\circ 2}(\mathbf{v}) + \cdots + 0T^{\circ d-1}(\mathbf{v})$, 我們知道此 matrix 第二個 column 應為 $(0,0,1,0,\ldots,0)^{\mathrm{t}}$, 這樣一直可得到前 $d-1$ 個 column. 至於最後一個 column, 由於 $T(T^{\circ d-1}(\mathbf{v})) = T^{\circ d}(\mathbf{v})$, 故若 $\mu_{\mathbf{v}}(x) = x^d + a_{d-1}x^{d-1} + \cdots + a_1x + a_0$, 由 $\mu_{\mathbf{v}}(T)(\mathbf{v}) = \mathbf{O}_V$, 得 $T^{\circ d}(\mathbf{v}) + a_{d-1}T^{\circ d-1}(\mathbf{v}) + \cdots + a_1T(\mathbf{v}) + a_0\mathbf{v} = \mathbf{O}_V$. 亦即

$$T^{\circ d}(\mathbf{v}) = -(a_0\mathbf{v} + a_1T(\mathbf{v}) + \cdots + a_{d-1}T^{\circ d-1}(\mathbf{v})),$$

因此得最後一個 column 為 $(-a_0, -a_1, \ldots, -a_{d-1})^{\mathrm{t}}$. 所以知 $T|_{C_{\mathbf{v}}}$ 對於 β 的 representative matrix 為

$$\begin{pmatrix} 0 & 0 & \cdots & 0 & -a_0 \\ 1 & 0 & \cdots & 0 & -a_1 \\ 0 & 1 & \ddots & 0 & -a_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & -a_{d-1} \end{pmatrix}. \tag{4.3}$$

令 $f(x) = x^d + a_{d-1}x^{d-1} + \cdots + a_1x + a_0$, 上述矩陣 (4.3) 稱為 the *companion matrix* of $f(x)$.

Question 4.18. 試計算

$$\det\begin{pmatrix} x & 0 & \cdots & 0 & a_0 \\ -1 & x & \cdots & 0 & a_1 \\ 0 & -1 & \ddots & 0 & a_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & -1 & x+a_{d-1} \end{pmatrix}.$$

Example 4.4.6. 考慮 $T : \mathbb{R}^3 \to \mathbb{R}^3$ 定義為 $T(x_1,x_2,x_3) = (2x_1, x_1+x_3, x_1-x_2)$. 考慮 $\mathbf{v} = (0,0,1)$, 則 $T(\mathbf{v}) = (0,1,0)$, $T^{\circ 2}(\mathbf{v}) = (0,0,-1) = -\mathbf{v}$. 由 $T^{\circ 3}(\mathbf{v}) = -T(\mathbf{v})$,... 很容易看出

$C_{\mathbf{v}} = \mathrm{Span}(\{\mathbf{v}, T(\mathbf{v})\}) = \mathrm{Span}(\{(0,0,1),(0,1,0)\})$. 由 $\{\mathbf{v}, T(\mathbf{v})\}$ 為 linearly independent 而 $\{\mathbf{v}, T(\mathbf{v}), T^{\circ 2}(\mathbf{v})\}$ 為 linearly dependent 知 $\mathbf{v}$ 的 T-annihilator 為 degree 2 的 polynomial. 事實上因 $T^{\circ 2}(\mathbf{v}) = -\mathbf{v}$, 即 $T^{\circ 2}(\mathbf{v}) + 0T(\mathbf{v}) + 1\mathbf{v} = \mathbf{O}_V$, 我們有 $\mu_{\mathbf{v}}(x) = x^2 + 1$. 很容易檢查 $\chi_T(x) = \mu_T(x) = (x-2)(x^2+1)$, 所以確實有 $\mu_{\mathbf{v}}(x) \mid \chi_T(x)$ 以及 $\mu_{\mathbf{v}}(x) \mid \mu_T(x)$. 另外考慮 $\beta = ((0,0,1),(0,1,0))$, 我們可以得到 $[T|_{C_{\mathbf{v}}}]_\beta = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$.

Question 4.19. 若 k 是最大的正整數滿足 $\{\mathbf{v}, T(\mathbf{v}), \ldots, T^{\circ k-1}(\mathbf{v}), T^{\circ k}(\mathbf{v})\}$ 為 *linearly independent,* 則 $\deg(\mu_{\mathbf{v}}(x))$ 為何?

在處理 finite dimensional vector space 的問題時, 我們通常會用 induction (數學歸納法). 也就是先探討 dimension 比較小的情況, 再將 dimension 比較大的情形化成比較小的情形. Quotient space 就是將 dimension 化成較小情形的一個方法 (回顧一下若 W 為 V 的 subspace, 則 $\dim(V/W) = \dim(V) - \dim(W)$). 現若 $T: V \to V$ 為 linear operator 且 $W \subseteq V$ 為 T-invariant subspace, 則我們可以定一個新的函數 $\overline{T}: V/W \to V/W$. 其定義為 $\overline{T}(\overline{\mathbf{v}}) = \overline{T(\mathbf{v})}$. 我們需說明 $\overline{T}$ 是 well-defined, 也就是說若 $\overline{\mathbf{v}} = \overline{\mathbf{u}} \in V/W$, 則 $\overline{T}(\overline{\mathbf{v}}) = \overline{T}(\overline{\mathbf{u}})$ in V/W. 然而 $\overline{\mathbf{v}} = \overline{\mathbf{u}}$ 表示 $\mathbf{v} - \mathbf{u} \in W$, 故利用 W 為 T-invariant 得 $T(\mathbf{v} - \mathbf{w}) \in W$, 即 $T(\mathbf{v}) - T(\mathbf{w}) \in W$. 這告訴我們 $\overline{T(\mathbf{v})} = \overline{T(\mathbf{u})}$ 故依 $\overline{T}$ 的定義得 $\overline{T}(\overline{\mathbf{v}}) = \overline{T}(\overline{\mathbf{u}})$. 依 $\overline{T}$ 的定義, 我們很容易得到 $\overline{T}: V/W \to V/W$ 為 linear operator. 我們稱 $\overline{T}: V/W \to V/W$ 為 linear operator *induced* by T on the quotient space V/W. 事實上 $\overline{T}$ 和 T 有許多的相關性, 我們有以下的性質.

Lemma 4.4.7. 設 $T: V \to V$ 為一個 *F-linear operator,* $W \subseteq V$ 為 *T-invariant subspace* 且令 $\overline{T}: V/W \to V/W$ 為 *linear operator induced by* T, 則對於任意 $g(x) \in F[x]$ 皆有 $g(\overline{T}) = \overline{g(T)}$.

Proof. 依定義我們知道 $g(\overline{T})$ 為 $V/W \to V/W$ 的 linear transformation 而且若 $g(x) = c_n x^n + \cdots + c_1 x + c_0$, 則對於任意 $\overline{\mathbf{v}} \in V/W$, $g(\overline{T})(\overline{\mathbf{v}}) = c_n \overline{T}^{\circ n}(\overline{\mathbf{v}}) + \cdots + c_1 \overline{T}(\overline{\mathbf{v}}) + c_0 \overline{\mathbf{v}}$. 又因 $\overline{T}^{\circ 2}(\overline{\mathbf{v}}) = \overline{T}(\overline{T}(\overline{\mathbf{v}})) = \overline{T}(\overline{T(\mathbf{v})}) = \overline{T(T(\mathbf{v}))} = \overline{T^{\circ 2}(\mathbf{v})}$, 利用數學歸納法可得 $\overline{T}^{\circ i}(\overline{\mathbf{v}}) = \overline{T^{\circ i}(\mathbf{v})}$, $\forall i \in \mathbb{N}$. 因此得

$$g(\overline{T})(\overline{\mathbf{v}}) = c_n \overline{T^{\circ n}(\mathbf{v})} + \cdots + c_1 \overline{T(\mathbf{v})} + c_0 \overline{\mathbf{v}}.$$

另一方面因 W 亦為 $g(T)$-invariant (Lemma 3.5.2), 故 $\overline{g(T)}$ 亦為 $V/W \to V/W$ 的 linear transformation 且

$$\overline{g(T)}(\overline{\mathbf{v}}) = \overline{g(T)(\mathbf{v})} = \overline{c_n T^{\circ n}(\mathbf{v}) + \cdots + c_1 T(\mathbf{v}) + c_0 \mathbf{v}}.$$

最後由 V/W 中元素運算的定義得 $g(\overline{T})(\overline{\mathbf{v}}) = \overline{g(T)}(\overline{\mathbf{v}})$, $\forall \overline{\mathbf{v}} \in V/W$. 故得證 $g(\overline{T}) = \overline{g(T)}$. □

利用 Lemma 4.4.7, 我們可以得到 T 和 $\overline{T}$ 的 minimal polynomial 之間的關係.

Corollary 4.4.8. 設 $T: V \to V$ 為一個 *F-linear operator,* $W \subseteq V$ 為 *T-invariant subspace* 且令 $\overline{T}: V/W \to V/W$ 為 *linear operator induced by* T, 則

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

另外給定 $\mathbf{v} \in V$, 令 $\mu_{\overline{\mathbf{v}}}(x)$ 為 *the $\overline{T}$-annihilator of* $\overline{v}$, 則我們有

$$\mu_{\overline{\mathbf{v}}}(x) \mid \mu_{\mathbf{v}}(x).$$