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 4

Form Reduction

對於一個 linear operator, 我們希望能找到適當的 ordered basis, 使其 representative matrix 為特殊的形式 (form). 在 matrices 來說指的就是要找到有特別 form 的 similar matrices. 我們希望得到 form 有的是所謂的 canonical form (將矩陣化為 canonical form 能幫我們判斷哪些矩陣是 similar), 還有一些 form 在數學許多領域都有重要的應用. 不過在此我們不去談論這些應用 (大家在研讀相關領域時自然會學到), 而專注於如何將一個矩陣化為這些 forms.

前一章我們提到利用 Primary Decomposition Theorem, 我們可以將 linear operator 簡化成只要考慮 characteristic polynomial 為 $p(x)^c$ 這種形式的 linear operator, 其中 $p(x)$ 是 $F[x]$ 上的 irreducible polynomial. 我們將逐步由 $p(x)$ 的可能情形來得到各種 forms.

4.1. Diagonal From

這一節中, 我們將從最簡單的 T-invariant subspace 出發, 引進所謂的 eigenvalue 以及 eigenvector, 再說明哪些情形可以得到 diagonal form.

對於一個 linear operator $T : V \to V$, 除了 $\{\mathbf{O}\}$ 以外, 最簡單的 T-invariant subspace 自然是 dimension 為 1 的 T-invariant subspace. 現若 U 為 T-invariant subspace 且 $\dim(U) = 1$, 即存在 $\mathbf{v} \neq \mathbf{O}_V$ 使得 $U = \mathrm{Span}(\{\mathbf{v}\})$. 由 U 為 T-invariant 的假設, 我們得 $T(\mathbf{v}) \in U = \mathrm{Span}(\{\mathbf{v}\})$. 也就是說, 存在 $\lambda \in F$ 使得 $T(\mathbf{v}) = \lambda\mathbf{v}$. 我們有以下的定義.

Definition 4.1.1. 假設 $T : V \to V$ 為 linear operator, 若存在 $\lambda \in F$ 以及 $\mathbf{v} \in V$ 且 $\mathbf{v} \neq \mathbf{O}_V$ 使得 $T(\mathbf{v}) = \lambda\mathbf{v}$, 則稱 λ 為 T 的一個 *eigenvalue*, 而 $\mathbf{v}$ 為 T 的一個 *eigenvector*.

注意, 對於 eigenvector $\mathbf{v}$ 我們要求 $\mathbf{v} \neq \mathbf{O}_V$, 而對於 eigenvalue λ 我們並無要求 $\lambda \neq 0$. 也就是說 $\mathbf{O}_V$ 雖符合 $T(\mathbf{O}_V) = \lambda\mathbf{O}_V$, 但我們不考慮這種 trivial 的情形, 故不稱 $\mathbf{O}_V$ 為 eigenvector. 另一方面若 $\mathbf{v} \neq \mathbf{O}_V$ 滿足 $T(\mathbf{v}) = 0\mathbf{v} = \mathbf{O}_V$, 表示 $\mathbf{v}$ 為 $\mathrm{Ker}(T)$ 的元素. 所以若 0 為 T 的 eigenvalue, 表示 $\mathrm{Ker}(T) \neq \{\mathbf{O}_V\}$, 亦即 $T : V \to V$ 不是 one-to-one.

Question 4.1. 假設 V 為 *finite dimensional vector space* 且 $T:V\to V$ 為 *linear operator.* 下列哪些是等價的?

(1) T *is an isomorphism (2)* T *is one-to-one (3)* T *is onto (4)* 0 *is not an eigenvalue of* T.

要找到一個 linear operator 有哪些 eigenvalue 和 eigenvector, 程序上是先找 T 有哪些 eigenvalue, 再利用這些 eigenvalue 將其對應的 eigenvector 找出. 首先觀察若 λ 為 $T:V\to V$ 的 eigenvalue, 則必存在 $\mathbf{v}\neq\mathbf{O}_V$ 使得 $T(\mathbf{v})=\lambda\mathbf{v}$, 得 $\lambda\mathrm{id}(\mathbf{v})-T(\mathbf{v})=\mathbf{O}_V$. 也就是說 $\mathbf{v}\in\mathrm{Ker}(\lambda\mathrm{id}-T)$, 亦即 $\lambda\mathrm{id}-T$ 這一個 linear operator 不是 isomorphism, 利用 Lemma 3.1.4 知 $\det(\lambda\mathrm{id}-T)=0$. 如何求 $\det(\lambda\mathrm{id}-T)$? 回顧一下, 我們需先找 V 的一個 ordered basis β, 再求 $\lambda\mathrm{id}-T$ 對於 β 的 representative matrix $[\lambda\mathrm{id}-T]_\beta$. 依定義 $\det(\lambda\mathrm{id}-T)$ 就是 $\det([\lambda\mathrm{id}-T]_\beta)$. 然而若 $\dim(V)=n$, 則我們有

$$[\lambda\mathrm{id}-T]_\beta=[\lambda\mathrm{id}]_\beta-[T]_\beta=\lambda[\mathrm{id}]_\beta-[T]_\beta=\lambda I_n-[T]_\beta.$$

因此若 $\lambda\in F$ 是 T 的一個 eigenvalue, 則 $\det(\lambda I_n-[T]_\beta)=0$. 又 T 的 characteristic polynomial 為 $\chi_T(x)=\chi_{[T]_\beta}(x)=\det(xI_n-[T]_\beta)$. 得知, 若 $\lambda\in F$ 是 T 的一個 eigenvalue, 則 $\chi_T(\lambda)=0$. 反之, 若 $\lambda\in F$ 為 $\chi_T(x)=0$ 之一根, 則 $\det(\lambda\mathrm{id}-T)=0$. 表示 $\lambda\mathrm{id}-T$ 這一個 linear operator 不是 one-to-one, 亦即存在 $\mathbf{v}\in V$ 且 $\mathbf{v}\neq\mathbf{O}_V$ 使得 $T(\mathbf{v})=\lambda\mathbf{v}$. 因此我們有以下之結果.

Proposition 4.1.2. 假設 V 為 *finite dimensional vector space* 且 $T:V\to V$ 為 *linear operator.* 則 $\lambda\in F$ 為 T 的 *eigenvalue* 若且唯若 $\chi_T(\lambda)=0$.

當 $\dim(V)=n$ 時, 由於 $\chi_T(x)\in F[x]$ 是一個次數為 n 的多項式, 它在 F 中根的個數最多只有 n 個 (當然也可能沒有根), 所以 T 僅能有有限多個 eigenvalue. 若 $\lambda\in F$ 為 $\chi_T(x)$ 的一根, 則 $(x-\lambda)\mid\chi_T(x)$. 又 $x-\lambda$ 為 $F[x]$ 的 monic irreducible polynomial, 所以若將 $\chi_T(x)$ 分解成 monic irreducible polynomials 的乘積 $\chi_T(x)=p_1(x)^{c_1}\cdots p_k(x)^{c_k}$. 這些 $p_i(x)$ 中次數為一次的多項式就給我們一個 T 的 eigenvalue. 我們對 $x-\lambda$ 可整除 $\chi_T(x)$ 的最高次方有興趣, 因此有以下的定義.

Definition 4.1.3. 假設 V 為 finite dimensional F-space, $T:V\to V$ 為 linear operator 且 λ 為 T 的一個 eigenvalue. 我們稱 $x-\lambda$ 可整除 $\chi_T(x)$ 的最高次方為 λ 的 *algebraic multiplicity.*

依此定義, 若 $\chi_T(x)=p_1(x)^{c_1}\cdots p_k(x)^{c_k}$, 其中 $p_1(x),\ldots,p_k(x)$ 為相異的 monic irreducible polynomials 且 $p_1(x)=x-\lambda$, 則 c_1 是 λ 的 algebraic multiplicity.

Question 4.2. 若 $T:V\to V$ 為 *linear operator* 且 $\dim(V)=n$, 則 T 最多有多少個相異的 *eigenvalue?* 此時每個 *eigenvalue* 的 *algebraic multiplicity* 為多少?

找到 T 所有可能的 eigenvalue 後, 我們就可以決定這些 eigenvalue 所對應的 eigenvector 了. 若 λ 為 eigenvalue, 前面提過所有滿足 $\mathbf{v}\neq\mathbf{O}_V$ 以及 $T(\mathbf{v})-\lambda\mathbf{v}=\mathbf{O}_V$ 的元素 $\mathbf{v}$ 就是 eigenvalue 為 λ 的 eigenvector. 也就是說 eigenvalue 為 λ 的 eigenvector 就是 $\mathrm{Ker}(T-\lambda\mathrm{id})$ 中的非 $\mathbf{O}_V$ 元素. 我們很自然會考慮以下的 vector space.

Definition 4.1.4. 假設 $T: V \to V$ 為 linear operator 且 λ 為 T 的一個 eigenvalue. 令

$$E_\lambda(T) = \mathrm{Ker}(T - \lambda\mathrm{id}) = \{\mathbf{v} \in V \mid T(\mathbf{v}) = \lambda\mathbf{v}\}.$$

稱之為 T 對應於 λ 的 *eigenspace* 且 $\dim(E_\lambda(T))$ 稱為 λ 的 *geometric multiplicity.*

假設 $\mathbf{v} \in E_\lambda(T)$, 由於 $T(T(\mathbf{v})) = T(\lambda\mathbf{v}) = \lambda T(\mathbf{v})$, 我們得 $T(\mathbf{v}) \in E_\lambda(T)$. 得知 $E_\lambda(T)$ 是一個 T-invariant subspace.

Question 4.3. 你能用 *Lemma 3.5.3* 說明 $E_\lambda(T)$ 為 *T-invariant subspace* 嗎?

如何得到 $E_\lambda(T)$ 呢? 我們仍是利用 ordered basis β 得到 $[T - \lambda\mathrm{id}]_\beta = [T]_\beta - \lambda I_n$ 這一個 matrix, 再求 $[T]_\beta - \lambda I_n$ 的 null space $N([T]_\beta - \lambda I_n) = \{\mathbf{x} \in F^n \mid ([T]_\beta - \lambda I_n) \cdot \mathbf{x} = \mathbf{O}\}$. 再利用 β 將 $N([T]_\beta - \lambda I_n)$ 中的元素還原回 V 中的元素, 就是 $E_\lambda(T)$ 的元素, 而且 $\dim(N([T]_\beta - \lambda I_n)) = \dim(E_\lambda(T))$ 就是 λ 的 geometric multiplicity.

Example 4.1.5. 考慮 $T: M_2(F) \to M_2(F)$ 定義為 $T\begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} a & c \\ b & d \end{pmatrix}$. 考慮 $M_2(F)$ 上的 ordered basis $\beta = (\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix})$, 則 T 對於 β 的 representative matrix 為

$$[T]_\beta = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

求得 $\chi_T(x) = \chi_{[T]_\beta}(x) = (x-1)^3(x+1)$. 所以 1 和 -1 為 T 的 eigenvalue, 它們的 algebraic multiplicity 分別為 3 和 1.

要求 T 對於 1 的 eigenspace $E_1(T)$, 我們先考慮 $N([T]_\beta - I_4)$, 即解聯立方程組

$$\begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & -1 & 1 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \quad \text{也就是解} \begin{cases} 0 & = & 0 \\ -x_2 + x_3 & = & 0 \\ x_2 - x_3 & = & 0 \\ 0 & = & 0 \end{cases}$$

解得 $N([T]_\beta - I_4) = \{(x_1, x_2, x_2, x_4)^{\mathrm{t}} \mid x_1, x_2, x_4 \in F\}$. 知 1 的 geometric multiplicity 為 3 且

$$E_1(T) = \{\begin{pmatrix} x_1 & x_2 \\ x_2 & x_4 \end{pmatrix} \mid x_1, x_2, x_4 \in F\}.$$

同理, 對於 -1 的 eigenspace $E_{-1}(T)$, 我們先考慮 $N([T]_\beta - (-1)I_4)$, 即解聯立方程組

$$\begin{pmatrix} 2 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 2 \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \quad \text{也就是解} \begin{cases} 2x_1 & = & 0 \\ x_2 + x_3 & = & 0 \\ x_2 + x_3 & = & 0 \\ 2x_4 & = & 0 \end{cases}$$

解得 $N([T]_\beta - (-1)I_4) = \{(0, x_2, -x_2, 0)^{\mathrm{t}} \mid x_2 \in F\}$. 知 -1 的 geometric multiplicity 為 1 且

$$E_{-1}(T) = \{\begin{pmatrix} 0 & x_2 \\ -x_2 & 0 \end{pmatrix} \mid x_2 \in F\}.$$

Algebraic multiplicity 並不一定會等於 geometric multiplicity, 我們看一個簡單的例子.

Example 4.1.6. 考慮 $T:P_1(F)\to P_1(F)$ 定義為 $T(ax+b)=bx$, 並考慮 $P_1(F)$ 的 ordered basis $\beta=(x,1)$. 我們有 $[T]_\beta=\begin{pmatrix}0&1\\0&0\end{pmatrix}$, 得 $\chi_T(x)=x^2$. 所以 0 是 T 唯一的 eigenvalue 且其 algebraic multiplicity 為 2. 要求 $N([T]_\beta-0I_2)=N([T]_\beta)$ 即解 $\begin{pmatrix}0&1\\0&0\end{pmatrix}\cdot\begin{pmatrix}a\\b\end{pmatrix}=\begin{pmatrix}0\\0\end{pmatrix}$ 得 $b=0$, 即 $N([T]_\beta-0I_2)=\{(a,0)^{\mathrm{t}}\mid a\in F\}$. 故 0 的 geometric multplicity 為 1 且 $E_0(T)=\{ax\mid a\in F\}$.

雖然 T 的一個 eigenvalue λ 的 algebraic multiplicity 和 geometric multiplicity 有可能不同, 不過它們之間仍有著某種關係存在. 我們利用 Primary Decomposition Theorem 來說明. 利用 Theorem 3.5.8 的符號, 假設

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

其中 $p_1(x),\ldots,p_k(x)$ 為相異的 monic irreducible polynomials 且因為 λ 為 T 的 eigenvalue, 我們令 $p_1(x)=x-\lambda$. 若令 $V_i=\operatorname{Ker}(p_i(T)^{\circ m_i})$, for $i=1,\ldots,k$, 則 Primary Decomposition Theorem (Theorem 3.5.8) 告訴我們

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

且

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

因假設 $p_1(x)=x-\lambda$, 我們有

$$V_1=\operatorname{Ker}((T-\lambda\mathrm{id})^{\circ m_1})\supseteq\operatorname{Ker}(T-\lambda\mathrm{id})=E_\lambda(T).$$

由此知 λ 的 geometric multiplicity $\dim(E_\lambda(T))\leq\dim(V_1)$. 另一方面, 依定義 c_1 為 λ 的 algebraic multiplicity, 而又 $\chi_{T|_{V_1}}(x)=(x-\lambda)^{c_1}$, 知 $\deg(\chi_{T|_{V_1}}(x))=c_1$. 因為一個 linear operator 的 characteristic polynomial 的 degree 為此 operator 所在的 space 之 dimension, 故得 $\dim(V_1)=c_1$. 因此我們知 $\dim(E_\lambda(T))\leq c_1$, 得到以下的結果.

Lemma 4.1.7. 若 V 為 *finite dimensional F-space*, $T:V\to V$ 為 *linear operator* 且 λ 為 T 的一個 *eigenvalue,* 則 λ 的 *algebraic multiplicity* 大於等於其 *geometric multiplicity.*

當 λ 是 T 的 eigenvalue 時, 由於 $E_\lambda(T)$ 存在著非 $\mathbf{O}_V$ 的元素, 故知 $\dim(E_\lambda(T))\geq 1$, 也就是說 λ 的 geometric multiplicity 必大於等於 1. 此時若 λ 的 algebraic multiplicity 是 1, 則由 Lemma 4.1.7 知 λ 的 algebraic multiplicity 等於其 geometric multiplicity (即 λ 的 geometric multiplicity 等於 1). 在一般的情形, 什麼時候 λ 的 algebraic multiplicity 會等於其 geometric multiplicity 呢? 我們有以下的結果.

Proposition 4.1.8. 假設 V 為 *finite dimensional F-space*, $T:V\to V$ 為 *linear operator* 且 λ 為 T 的一個 *eigenvalue.* 則 λ 的 *algebraic multiplicity* 等於其 *geometric multiplicity* 若且唯若 $x-\lambda\mid\mu_T(x)$ 但 $(x-\lambda)^2\nmid\mu_T(x)$.

Proof. 我們用前面一樣的符號, 設 $\mu_T(x) = p_1(x)^{m_1} \cdots p_k(x)^{m_k}$ 以及 $\chi_T(x) = p_1(x)^{c_1} \cdots p_k(x)^{c_k}$, 其中 $p_1(x) = x - \lambda$. 又令 $V_1 = \operatorname{Ker}((T - \lambda \mathrm{id})^{\circ m_1})$. 若 $x - \lambda \mid \mu_T(x)$ 但 $(x-\lambda)^2 \nmid \mu_T(x)$, 此即表示 $m_1 = 1$, 故得 $V_1 = \operatorname{Ker}(T - \lambda \mathrm{id}) = E_\lambda(T)$. 前面已知 $\dim(V_1)$ 為 λ 的 algebraic multiplicity, 而依定義 $\dim(E_\lambda(T))$ 為 λ 的 geometric multiplicity, 故得證它們相等.

反過來, 若 λ 的 algebraic multiplicity 等於其 geometric multiplicity, 即表示 $\dim(V_1) = \dim(E_\lambda(T))$, 故得 $V_1 = \operatorname{Ker}(T - \lambda \mathrm{id})$. 換句話說, 對於任意 $\mathbf{v} \in V_1$, $T(\mathbf{v}) - \lambda \mathrm{id}(\mathbf{v}) = \mathbf{O}_V$. 這告訴我們 $T - \lambda \mathrm{id}$ 限制在 V_1 上是一個 zero mapping, 即 $(T - \lambda \mathrm{id})|_{V_1} = T|_{V_1} - \lambda \mathrm{id}|_{V_1} = \mathbf{O}$. 也就是說, 若令 $h(x) = x - \lambda$, 得 $h(T|_{V_1}) = \mathbf{O}$. 因此由 Lemma 3.3.5 知 $T|_{V_1}$ 的 minimal polynomial $\mu_{T|_{V_1}}(x)$ 整除 $h(x) = x - \lambda$. 然而 Theorem 3.5.8 告訴我們 $\mu_{T|_{V_1}}(x) = (x-\lambda)^{m_1}$, 故得證 $m_1 = 1$. □

特別的, 假設 $\chi_T(x)$ 可以完全分解成 $F[x]$ 中的一次多項式的乘積, 亦即 $\chi_T(x) = p_1(x)^{c_1} \cdots p_k(x)^{c_k}$, 其中每一個 $p_i(x)$ 皆為一次多項式 $x - \lambda_i$. 此時若每一個 λ_i 的 algebraic multiplicity 和 geometric multiplicity 皆相等, 則由 Proposition 4.1.8 知 $\mu_T(x) = p_1(x) \cdots p_k(x)$, 因此得 $V_i = \operatorname{Ker}(T - \lambda_i \mathrm{id}) = E_{\lambda_i}(T)$, $\forall i = 1, \ldots, k$. 因此由 Primary Decomposition Theorem 知

$$V = E_{\lambda_1}(T) \oplus \cdots \oplus E_{\lambda_k}(T).$$

也就是說此時 V 就會是 eigenspaces 的 (internal) direct sum. 因為每個 eigenspace 中的非 $\mathbf{O}_V$ 元素皆為 T 的 eigenvector, 所以 $E_{\lambda_i}(T)$ 中的任一組 basis S_i 皆由 T 的 eigenvector 所組成. 又因 $V = E_{\lambda_1}(T) \oplus \cdots \oplus E_{\lambda_k}(T)$, Proposition 3.4.6 告訴我們 $S_1 \cup \cdots \cup S_k$ 為 V 的一組 basis, 也就是說 V 有一組 basis 是由 T 的 eigenvector 所組成. 現假設 $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ 為 V 的一組 basis, 其中 $\mathbf{v}_i$ 為 T 的 eigenvector 且其對應的 eigenvalue 為 γ_i (這裡 γ_i 不一定相異), 此時考慮 V 的 ordered basis $\beta = (\mathbf{v}_1, \ldots, \mathbf{v}_n)$. 由於對所有 $i = 1, \ldots, n$, 皆有 $T(\mathbf{v}_i) = \gamma_i \mathbf{v}_i$, 我們得到

$$[T]_\beta = \begin{pmatrix} \gamma_1 & & \mathbf{O} \\ & \ddots & \\ \mathbf{O} & & \gamma_n \end{pmatrix}$$

為一個 diagonal matrix (對角矩陣). 因此有以下之定義.

Definition 4.1.9. 假設 V 為 finite dimensional F-space 且 $T : V \to V$ 為 linear operator. 若 V 存在一組 basis 是由 T 的 eigenvectors 所組成, 則稱 T 為一個 *diagonalizable linear operator.*

我們有以下等價的關係來判斷一個 linear operator 是否為 diagonalizable.

Theorem 4.1.10. 假設 V 為 *finite dimensional F-space* 且 $T : V \to V$ 為 *linear operator.* 則以下是等價的.

(1) T 是一個 *diagonalizable linear operator.*

(2) 存在 V 的 *ordered basis* β 使得 $[T]_\beta$ 為一個 *diagonal matrix.*

(3) T 的 *characteristic polynomial* $\chi_T(x)$ 可以完全分解成 $F[x]$ 中的一次多項式之乘積, 且 T 的每一個 *eigenvalue* 的 *algebraic multiplicity* 和 *geometric multiplicity* 相等.

(4) T 的 *minimal polynomial* $\mu_T(x)$ 可以完全分解成 $F[x]$ 中相異的 *monic* 一次多項式之乘積.

Proof. 前面我們已知 (3) $\Rightarrow$ (1) 且 (1) $\Rightarrow$ (2), 現要證明 (2) $\Rightarrow$ (4). 假設 $\dim(V)=n$ 且 β 為 V 的 ordered basis 使得

$$[T]_\beta = \begin{pmatrix} \gamma_1 & & \mathbf{O} \\ & \ddots & \\ \mathbf{O} & & \gamma_n \end{pmatrix}$$

為一個 diagonal matrix. 現假設 $\lambda_1,\ldots,\lambda_k$ 皆相異且 $\{\gamma_1,\ldots,\gamma_n\}=\{\lambda_1,\ldots,\lambda_k\}$. 亦即對任意 γ_i 皆存在 λ_j 使得 $\gamma_i=\lambda_j$. 依定義 $\chi_T(x)=\chi_{[T]_\beta}(x)=(x-\gamma_1)\cdots(x-\gamma_k)=(x-\lambda_1)^{c_1}\cdots(x-\lambda_k)^{c_k}$, 其中 $c_i\in\mathbb{N}$. 而且由 Theorem 3.3.7 (或 Theorem 3.3.9) 知 $\mu_T(x)=(x-\lambda_1)^{m_1}\cdots(x-\lambda_k)^{m_k}$, 其中 $m_i\in\mathbb{N}$. 現考慮 $h(x)=(x-\lambda_1)\cdots(x-\lambda_k)$, 由 Lemma 3.2.1 得

$$\begin{aligned} h([T]_\beta) &= ([T]_\beta-\lambda_1 I_n)\cdots([T]_\beta-\lambda_k I_n) \\ &= \begin{pmatrix} \gamma_1-\lambda_1 & & \mathbf{O} \\ & \ddots & \\ \mathbf{O} & & \gamma_n-\lambda_1 \end{pmatrix} \cdots \begin{pmatrix} \gamma_1-\lambda_k & & \mathbf{O} \\ & \ddots & \\ \mathbf{O} & & \gamma_n-\lambda_k \end{pmatrix} \\ &= \begin{pmatrix} (\gamma_1-\lambda_1)\cdots(\gamma_1-\lambda_k) & & \mathbf{O} \\ & \ddots & \\ \mathbf{O} & & (\gamma_n-\lambda_1)\cdots(\gamma_n-\lambda_k) \end{pmatrix} \end{aligned}$$

然而每個 γ_i 皆存在 λ_j, $j=1,\ldots,k$ 使得 $\gamma_i=\lambda_j$, 故得 $h([T]_\beta)=\mathbf{O}$, 亦即 $h(T)=\mathbf{O}$. 所以由 Lemma 3.3.5 得 $\mu_T(x)\mid h(x)$, 得證 $\mu_T(x)=h(x)=(x-\lambda_1)\cdots(x-\lambda_k)$, 亦即 T 的 minimal polynomial $\mu_T(x)$ 可以完全分解成 $F[x]$ 中相異的 monic 一次多項式之乘積.

最後我們要證明 (4) $\Rightarrow$ (3). 假設 $\mu_T(x)=(x-\lambda_1)\cdots(x-\lambda_k)$, 其中 $\lambda_i\in F$ 且皆相異. 由 Theorem 3.3.7, 知 $\chi_T(x)=(x-\lambda_1)^{c_1}\cdots(x-\lambda_k)^{c_k}$ 其中 $c_i\in\mathbb{N}$, 即 $\chi_T(x)$ 可以完全分解成 $F[x]$ 中的一次多項式之乘積. 然而 $\lambda_1,\ldots,\lambda_k$ 為 T 的所有 eigenvalues, 且對於每一個 $i=1,\ldots,k$ 皆有 $(x-\lambda_i)\mid\mu_T(x)$ 但 $(x-\lambda_i)^2\nmid\mu_T(x)$. 故 Proposition 4.1.8 告訴我們每個 λ_i 的 algebraic multiplicity 和 geometric multiplicity 皆相等. 得證本定理. □

Question 4.4. 假設 $\dim(V)=n$, $T:V\to V$ 為 *linear operator.* 若 T 有 n 個相異的 *eigenvalue,* 則 T 是否為 *diagonalizable?*

Question 4.5. *Example 4.1.5* 和 *Example 4.1.6* 中哪一個 T 是 *diagonalizable?*

雖然前面都是談 linear operator, 我們要強調這些性質對於 $n\times n$ 的方陣也有相對應的地方. 首先若 $A\in M_n(F)$, 我們也有所謂的 eigenvalue 以及 eigenvector.

Definition 4.1.11. 假設 $A\in M_n(F)$. 若存在 $\lambda\in F$ 以及 $\mathbf{x}\in F^n$ 且 $\mathbf{x}\neq\mathbf{O}$ 使得 $A\cdot\mathbf{x}=\lambda\mathbf{x}$, 則稱 λ 為 A 的一個 *eigenvalue*, 而 $\mathbf{x}$ 為 T 的一個 *eigenvector.*

接下來利用 A 的 characteristic polynomial $\chi_A(x)$ 來得到 A 的 eigenvalues λ 以及求 $N(A-\lambda I_n)$ 來得到 A 相對於 λ 的 eigenvector, 還有關於 eigenvalue 的 algebraic multiplicity 和 geometric multiplicity, ... 等性質, 我們就不再贅敘.

Question 4.6. 若 $A \in M_n(F)$, λ 為 A 的 *eigenvalue,* 你能定義 λ 的 *algebraic multiplicity* 和 *geometric multiplicity* 嗎? 你能寫出 A 相對於 *Lemma 4.1.7* 以及 *Proposition 4.1.6* 的定理嗎?

我們也可定義何謂 diagonalizable matrix 如下.

Definition 4.1.12. 假設 $A \in M_n(F)$. 若存在一組 F^n basis 是由 A 的 eigenvectors 所組成, 則稱 A 為一個 *diagonalizable matrix.*

我們也有如同 Theorem 4.1.10 判斷 A 是否為 diagonalizable 的等價方法. 因為證明就如同 linear operator 的情形, 我們就不再重複.

Theorem 4.1.13. 假設 $A \in M_n(F)$. 則以下是等價的.

(1) A 是一個 *diagonalizable matrix.*

(2) 存在 $P \in M_n(F)$ 為 *invertible* 使得 $P^{-1}\cdot A\cdot P$ 為一個 *diagonal matrix.*

(3) $\chi_A(x)$ 可以完全分解成 $F[x]$ 中的一次多項式之乘積, 且 A 的每一個 *eigenvalue* 的 *algebraic multiplicity* 和 *geometric multiplicity* 相等.

(4) $\mu_A(x)$ 可以完全分解成 $F[x]$ 中相異的 *monic* 一次多項式之乘積.

當 A 為 diagonalizable, Theorem 4.1.13 (2) 中 $P^{-1}\cdot A\cdot P$ 這一個 diagonal matrix 就稱為 A 的 *diagonal form.* 我們特別說明一下如何找到 P 將 A 化為 diagonal form. 假設

$$P^{-1}\cdot A\cdot P = D = \begin{pmatrix} \gamma_1 & & \mathbf{O} \\ & \ddots & \\ \mathbf{O} & & \gamma_n \end{pmatrix},$$

且令 $P_i \in F^n$ 為 P 的 i-th column. 前面提過求兩個矩陣相乘其 i-th column 的方法, 我們有 $A\cdot P$ 的 i-th column 為 $A\cdot P_i$, 而 $P\cdot D$ 的 i-th column 為 $\gamma_i P_i$, 所以利用 $A\cdot P = P\cdot D$ 得 $A\cdot P_i = \lambda P_i$, 也就是說 P 的 i-th column P_i 就是一個 eigenvalue 為 γ_i 的 eigenvector. 因此我們只要將一個 diagonalizable matrix A 的 eigenvectors 所組成 F^n 的一組 basis, 按照順序一個 column 一個 column 填入, 所得的 invertible matrix P, 就是可以將 A 對角化. 也就是說 $P^{-1}\cdot A\cdot P$ 為一個對角矩陣.

最後我們說明為何兩個 diagonalizable matrices, 將其化成 diagonal form 後就可以判斷其是否為 similar. 首先強調若 A 為 diagonalizable, 且 $B \sim A$, 則 B 必為 diagonalizable. 這是因為假設 P 為 invertible 且 $P^{-1}\cdot A\cdot P = D$ 為 diagonal matrix. 由存在 Q 為 invertible 使得 $B = O^{-1}\cdot A\cdot Q$, 得

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

又因 $Q^{-1}\cdot P$ 為 invertible 得證 B 為 diagonalizable.

另一方面若 A,B 皆為 diagonalizable, 若 $A\sim B$, 表示它們有相同的 characteristic polynomial, 因此有相同的 eigenvalues 且 A 和 B 同一個 eigenvalue 的 algebraic multiplicity 皆相等. 而每一個 eigenvalue 的 algebraic multiplicity 又等於其 geometric multiplicity, 所以將 A,B 化為 diagonal form 後同一個 eigenvalue 發生在的對角線上的次數會相同. 反之, 若將 A,B 化為 diagonal form 後同一個 eigenvalue 發生在的對角線上的次數相同, 表示將 diagonal form 對角線位置適當互換後, 兩個 diagonal form 會相等. 然而對角線位置互換只是將 eigenvector 所形成的 ordered basis 做適當重新排序 (例如將 (i,i)-th entry 和 (j,j)-th entry 互換只是將原來 P 的 i-th column 和 j-th column 互換), 所以得知 $A\sim B$.

4.2. Triangular Form

當 linear operator T 的 characteristic polynomial 可完全分解成一次的 monic polynomials 的乘積時, T 不一定是 diagonalizable. 這一節中我們將探討在這種情形時 T 可以化成怎樣的形式.

注意本節中我們仍假設 $\chi_T(x)$ 可以完全分解成一次的多項式的乘積 (即 $\chi_T(x)=(x-\lambda_1)^{c_1}\cdots(x-\lambda_k)^{c_k}$). 這個假設當 V over 的 field F 是 algebraically closed (例如 $F=\mathbb{C}$) 時自然會成立. 利用 Primary Decomposition Theorem, 我們假設 T 的 minimal polynomial 為 $\mu_T(x)=(x-\lambda)^m$. 也就是說 $(T-\lambda\mathrm{id})^{\circ m}=\mathbf{O}$.

當一個 linear operator $T:V\to V$ 滿足 $T^{\circ m}=\mathbf{O}$, 我們稱之為 *nilpotent*, 而最小的正整數 m 使得 $T^{\circ m}=\mathbf{O}$, 稱為這個 nilpotent operator 的 *index*. 因為我們假設 $T-\lambda\mathrm{id}$ 為 nilpotent 且 index 為 m. 我們來特別探討 nilpotent operator 的性質.

對於一個 linear operator $T:V\to V$. 若 $\mathbf{v}\in\mathrm{Im}(T^{\circ i})$, 表示存在 $\mathbf{u}\in V$ 使得 $\mathbf{v}=T^{\circ i}(\mathbf{u})$, 因此當 $i\geq 2$ 時, 我們有 $\mathbf{v}=T^{\circ i-1}(T(\mathbf{u}))\in\mathrm{Im}(T^{\circ i-1})$. 所以我們自然有以下的 chain of subspaces

$$V\supseteq\mathrm{Im}(T)\supseteq\mathrm{Im}(T^{\circ 2})\supseteq\cdots\supseteq\mathrm{Im}(T^{\circ i-1})\supseteq\mathrm{Im}(T^{\circ i})\supseteq\cdots.$$

特別的, 當 T 為 nilpotent of index m, 我們有以下情形.

Lemma 4.2.1. *假設 $\dim(V)>0$, 若 T 為 nilpotent operator of index m, 則我們有以下的 chain of subspaces.*

$$V\supsetneq\mathrm{Im}(T)\supsetneq\mathrm{Im}(T^{\circ 2})\supsetneq\cdots\supsetneq\mathrm{Im}(T^{\circ i-1})\supsetneq\mathrm{Im}(T^{\circ i})\supsetneq\cdots\supsetneq\mathrm{Im}(T^{\circ m-1})\supsetneq\mathrm{Im}(T^{\circ m})=\{\mathbf{O}_V\}.$$

Proof. 首先說明 $\mathrm{Im}(T^{\circ m-1})\supsetneq\mathrm{Im}(T^{\circ m})=\{\mathbf{O}_V\}$. 因為 $T^{\circ m}=\mathbf{O}$, 亦即對任意 $\mathbf{v}\in V, T^{\circ m}(\mathbf{v})=\mathbf{O}_V$, 所以 $\mathrm{Im}(T^{\circ m})=\{\mathbf{O}_V\}$. 另一方面, 若 $\mathrm{Im}(T^{\circ m-1})=\mathrm{Im}(T^{\circ m})=\{\mathbf{O}_V\}$, 則表示 $T^{\circ m-1}=\mathbf{O}$, 此與 m 為最小的正整數使得 $T^{\circ m}=\mathbf{O}$ 相矛盾, 故知 $\mathrm{Im}(T^{\circ m-1})\neq\mathrm{Im}(T^{\circ m})$.

接下來我們說明 $V\supsetneq\mathrm{Im}(T)$. 若 $\mathrm{Im}(T)=V$, 表示對任意 $\mathbf{v}\in V$ 皆存在 $\mathbf{v}_1\in V$ 使得 $\mathbf{v}=T(\mathbf{v}_1)$. 而 $\mathbf{v}_1\in V$, 故存在 $\mathbf{v}_2\in V$ 使得 $\mathbf{v}=T(\mathbf{v}_1)=T^{\circ 2}(\mathbf{v}_2)$, 得 $V=\mathrm{Im}(T^{\circ 2})$. 如此一直下去, 我們可證得 $V=\mathrm{Im}(T^{\circ i})$, $\forall i\in\mathbb{N}$. 因 $V\neq\{\mathbf{O}_V\}$, 此與 T 為 nilpotent 相矛盾, 故知 $V\neq\mathrm{Im}(T)$.

同理, 當 $1 \le i \le m-2$, 因對於所有 $\mathbf{v} \in \mathrm{Im}(T^{\circ i+1})$ 皆存在 $\mathbf{u} \in V$ 使得 $\mathbf{v} = T^{\circ i+1}(\mathbf{u}) = T(T^{\circ i}(\mathbf{u}))$. 現若 $\mathrm{Im}(T^{\circ i}) = \mathrm{Im}(T^{\circ i+1})$, 則由 $T^{\circ i}(\mathbf{u}) \in \mathrm{Im}(T^{\circ i}) = \mathrm{Im}(T^{\circ i+1})$ 知存在 $\mathbf{w} \in V$ 使得 $T^{\circ i}(\mathbf{u}) = T^{\circ i+1}(\mathbf{w})$. 亦即 $\mathbf{v} = T(T^{\circ i}(\mathbf{u})) = T^{\circ i+2}(\mathbf{w}) \in \mathrm{Im}(T^{\circ i+2})$, 得證 $\mathrm{Im}(T^{\circ i+1}) = \mathrm{Im}(T^{\circ i+2})$. 如此一直下去會推得 $\mathrm{Im}(T^{\circ m-1}) = \mathrm{Im}(T^{\circ m})$, 此與前面所得 $\mathrm{Im}(T^{\circ m-1}) \neq \mathrm{Im}(T^{\circ m})$ 相矛盾, 故知 $\mathrm{Im}(T^{\circ i}) \neq \mathrm{Im}(T^{\circ i+1})$, 得證本定理. □

接下來我們說明若 $\dim(V) = n$ 且 $T : V \to V$ 為 nilpotent operator of index m, 如何將其化為 triangular form. 首先選取 $\mathrm{Im}(T^{\circ m-1})$ 的 ordered basis $(\mathbf{v}_1, \ldots, \mathbf{v}_{k_1})$, 注意此時我們有 $T(\mathbf{v}_i) \in \mathrm{Im}(T^{\circ m}) = \{\mathbf{O}_V\}$, 故

$$T(\mathbf{v}_i) = \mathbf{O}_V, \forall i = 1, \ldots, k_1.$$

接著加入 $\{\mathbf{v}_{k_1+1}, \ldots, \mathbf{v}_{k_2}\}$ 使得 $(\mathbf{v}_1, \ldots, \mathbf{v}_{k_1}, \ldots, \mathbf{v}_{k_2})$ 為 $\mathrm{Im}(T^{\circ m-2})$ 的 ordered basis. 此時我們有

$$T(\mathbf{v}_i) \in \mathrm{Im}(T^{\circ m-1}) = \mathrm{Span}(\{\mathbf{v}_1, \ldots, \mathbf{v}_{k_1}\}), \forall i = k_1+1, \ldots, k_2,$$

而且利用 ordered basis $(\mathbf{v}_1, \ldots, \mathbf{v}_{k_1}, \ldots, \mathbf{v}_{k_2})$ 所得 $T|_{\mathrm{Im}(T^{\circ m-2})}$ 的 representative matrix 為

$$\begin{pmatrix} \mathbf{O}_{k_1,k_1} & * \\ \mathbf{O}_{k_2-k_1,k_1} & \mathbf{O}_{k_2-k_1,k_2-k_1} \end{pmatrix},$$

其中 $\mathbf{O}_{i,j}$ 表示為 $i \times j$ 階的零矩陣, 而右上角的 $*$ 為一個 $k_1 \times k_2 - k_1$ 階的非零矩陣. 接下來加入 $\{\mathbf{v}_{k_2+1}, \ldots, \mathbf{v}_{k_3}\}$ 使得 $(\mathbf{v}_1, \ldots, \mathbf{v}_{k_1}, \ldots, \mathbf{v}_{k_2}, \ldots, \mathbf{v}_{k_3})$ 為 $\mathrm{Im}(T^{\circ m-3})$ 的 ordered basis. 此時我們有

$$T(\mathbf{v}_i) \in \mathrm{Im}(T^{\circ m-2}) = \mathrm{Span}(\{\mathbf{v}_1, \ldots, \mathbf{v}_{k_1}, \ldots, \mathbf{v}_{k_2}\}), \forall i = k_2+1, \ldots, k_3,$$

而且利用 ordered basis $(\mathbf{v}_1, \ldots, \mathbf{v}_{k_1}, \ldots, \mathbf{v}_{k_2}, \ldots, \mathbf{v}_{k_3})$ 所得 $T|_{\mathrm{Im}(T^{\circ m-3})}$ 的 representative matrix 為

$$\begin{pmatrix} \mathbf{O}_{k_1,k_1} & * & * \\ \mathbf{O}_{k_2-k_1,k_1} & \mathbf{O}_{k_2-k_1,k_2-k_1} & * \\ \mathbf{O}_{k_3-k_2,k_1} & \mathbf{O}_{k_3-k_2,k_2-k_1} & \mathbf{O}_{k_3-k_2,k_3-k_2} \end{pmatrix}.$$

這樣一直下去我們可得到 $\mathrm{Im}(T)$ 的 ordered basis $(\mathbf{v}_1, \ldots, \mathbf{v}_{k_{m-1}})$, 其中對於 $j = 1, \ldots, m-1$, 皆有 $(\mathbf{v}_1, \ldots, \mathbf{v}_{k_j})$ 為 $\mathrm{Im}(T^{\circ m-j})$ 的 ordered basis 且

$$T(\mathbf{v}_i) \in \mathrm{Im}(T^{\circ m-(j-1)}) = \mathrm{Span}(\{\mathbf{v}_1, \ldots, \mathbf{v}_{k_{j-1}}\}), \forall i = k_{j-1}+1, \ldots, k_j.$$

最後加入 $\{\mathbf{v}_{k_{m-1}+1}, \ldots, \mathbf{v}_n\}$ 使得 $(\mathbf{v}_1, \ldots, \mathbf{v}_{k_{m-1}}, \ldots, \mathbf{v}_n)$ 為 V 的 ordered basis, 此時

$$T(\mathbf{v}_i) \in \mathrm{Im}(T) = \mathrm{Span}(\{\mathbf{v}_1, \ldots, \mathbf{v}_{k_{m-1}}\}), \forall i = k_{m-1}+1, \ldots, k_n,$$

而且利用 ordered basis $(\mathbf{v}_1, \ldots, \mathbf{v}_{k_{m-1}}, \ldots, \mathbf{v}_n)$ 所得 T 的 representative matrix 為

$$\begin{pmatrix} \mathbf{O} & * & * \\ \vdots & \ddots & * \\ \mathbf{O} & \mathbf{O} & \mathbf{O} \end{pmatrix}.$$

這一個矩陣是對角線皆為 0 的 upper triangular matrix (上三角矩陣), 所以我們有以下的結果.

Proposition 4.2.2. 假設 V 為 *finite dimensional* F*-space* 且 $T : V \to V$ 為 *linear operator.* 則 T 為 *nilpotent* 若且唯若存在 V 的 *ordered basis* β 使得 $[T]_\beta$ 為 *upper triangular matrix* 且 $[T]_\beta$ 的對角線皆為 0.

Proof. 由前面的討論我們知: 若 T 為 nilpotent, 則存在 V 的 ordered basis β 使得 $[T]_\beta$ 為 upper triangular matrix 且其對角線皆為 0. 反之, 若 $[T]_\beta$ 為 upper triangular matrix 且其對角線皆為 0, 我們知 $\chi_T(x) = \chi_{[T]_\beta}(x) = x^n$ (其中 $n = \dim(V)$), 故知 $T^{\circ n} = \mathbf{O}$, 得證 T 為 nilpotent. □

Question 4.7. 若 V 為 *finite dimensional* F*-space* 且 $T : V \to V$ 為 *nilpotent operator of index* m, 則 $\chi_T(x)$ 為何? 又 $\mu_T(x)$ 為何?

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

Example 4.2.3. 考慮 linear operator $T : P_2(\mathbb{R}) \to P_2(\mathbb{R})$, 定義為 $T(ax^2 + bx + c) = (c-a)x^2 + cx + (c-a)$. 若考慮 $P_2(\mathbb{R})$ 的 ordered basis $\beta = (x^2, x, 1)$, 我們有 $[T]_\beta = \begin{pmatrix} -1 & 0 & 1 \\ 0 & 0 & 1 \\ -1 & 0 & 1 \end{pmatrix}$, 得 $\chi_T(x) = x^3$. 又 $[T]_\beta^2 = \begin{pmatrix} 0 & 0 & 0 \\ -1 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}$ 知 $\mu_T(x) = x^3$, 即 T 為 nilpotent of index 3. 因 $[T]_\beta^2$ 的 column space 為 $\mathrm{Span}(\{(0,1,0)^t\})$, 我們得 $\mathrm{Im}(T^{\circ 2}) = \mathrm{Span}(\{x\})$. 同理由 $[T]_\beta$ 的 column space, 可得 $\mathrm{Im}(T) = \mathrm{Span}(\{x, x^2+1\})$. 最後因 $x^2 \notin \mathrm{Im}(T)$, 我們可以考慮 $P_2(\mathbb{R})$ 的 ordered basis $\beta' = (x, x^2+1, x^2)$. 因

$$T(x) = 0, T(x^2+1) = 1x + 0(x^2+1) + 0x^2, T(x^2) = 0x + (-1)(x^2+1) + 0x^2$$

得 $[T]_{\beta'} = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & -1 \\ 0 & 0 & 0 \end{pmatrix}$ 這一個 diagonal 皆為 0 的 upper triangular matrix.

現在我們回到 T 的 minimal polynomial 為 $\mu_T(x) = (x-\lambda)^m$ 的情形, 此時 $T - \lambda\,\mathrm{id}$ 為 nilpotent 所以由 Proposition 4.2.2 知存在 ordered basis β 使得 $[T - \lambda\,\mathrm{id}]_\beta = U$ 為一個 diagonal 皆為 0 的 upper triangular matrix

$$U = \begin{pmatrix} 0 & * & * \\ \vdots & \ddots & * \\ 0 & \cdots & 0 \end{pmatrix}.$$

然而若 $\dim(V) = n$, 因 $[T - \lambda\,\mathrm{id}]_\beta = [T]_\beta - \lambda I_n$, 故得 $[T]_\beta = \lambda I_n + U$, 為一個 diagonal 皆為 λ 的 upper triangular matrix

$$\lambda I_n + U = \begin{pmatrix} \lambda & * & * \\ & \ddots & * \\ \mathbf{O} & & \lambda \end{pmatrix}.$$

Theorem 4.2.4. 假設 V 為 *finite dimensional* F*-space.* 若 $T : V \to V$ 為 *linear operator* 其 *characteristic polynomial* 為

$$\chi_T(x) = (x-\lambda_1)^{c_1} \cdots (x-\lambda_k)^{c_k},$$

其中 $\lambda_1, \ldots, \lambda_k$ 為 F 中相異的元素, 則存在 V 的 *ordered basis* β 使得

$$[T]_\beta = \begin{pmatrix} A_1 & & \mathbf{O} \\ & \ddots & \\ \mathbf{O} & & A_k \end{pmatrix},$$

其中每個 A_i 為 $c_i \times c_i$ 階的 *upper triangular matrix*

$$\begin{pmatrix} \lambda_i & * & * \\ & \ddots & * \\ \mathbf{O} & & \lambda_i \end{pmatrix}.$$

Proof. 由 Theorem 3.3.9 知存在 $m_i \le c_i$ 使得 $\mu_T(x) = (x-\lambda_1)^{m_1} \cdots (x-\lambda_k)^{m_k}$, 故由 Primary Decomposition Theorem, 我們知 $V = V_1 \oplus \cdots \oplus V_k$, 其中 $V_i = \mathrm{Ker}((T - \lambda_i \mathrm{id})^{\circ m_i})$ 且 $\mu_{T|_{V_i}}(x) = (x-\lambda_i)^{m_i}$. 得 $T|_{V_i} - \lambda_i \mathrm{id}|_{V_i}$ 為 nilpotent, 故利用 Proposition 4.2.2, 我們知存在 β_i 為 V_i 的 ordered basis, 使得 $[T|_{V_i}]_{\beta_i}$ 為 A_i 這樣的 $c_i \times c_i$ 階的 upper triangular matrix. 故將 $\beta_1, \ldots, \beta_k$ 依序排列形成 V 的 ordered basis β, 可得 $[T]_\beta$ 為所要的 triangular matrix. □

Theorem 4.2.4 告訴我們當 T 的 characteristic polynomial 可完全分解成 $F[x]$ 中的一次多項式乘積, 雖然 T 可能不能化成 diagonal form 不過一定可以化成 triangular form.

若 T 是 diagonalizable, 我們可以利用對角化幫助我們求得 $T^{\circ i}$. 即利用 V 的 eigenvectors 所形成的 ordered basis β 得 $[T]_\beta = \begin{pmatrix} \gamma_1 & & \mathbf{O} \\ & \ddots & \\ \mathbf{O} & & \gamma_n \end{pmatrix}$, 故可得 $[T^{\circ i}]_\beta = \begin{pmatrix} \gamma_1^i & & \mathbf{O} \\ & \ddots & \\ \mathbf{O} & & \gamma_n^i \end{pmatrix}$. 當 T 不能化為 diagonal form 時, 我們可利用 trianbular form 來幫助計算 $T^{\circ i}$.

首先將 V 寫成 T-invariant subspaces 的 direct sum $V = V_1 \oplus \cdots \oplus V_k$. 由於任意 $\mathbf{v} \in V$, 都可以唯一寫成 $\mathbf{v} = \mathbf{v}_1 + \cdots + \mathbf{v}_k$, 其中 $\mathbf{v}_i \in V_i$ (Proposition 3.4.6). 對於所有的 $i = 1, \ldots, k$, 我們可定義一個 linear operator $\pi_i : V \to V$, 其定義為 $\pi_i(\mathbf{v}) = \mathbf{v}_i$. 此 linear operator 稱為 the *projection* to V_i with respect to the direct sum $V = V_1 \oplus \cdots \oplus V_k$. 依此定義我們知道對於所有 $\mathbf{v} \in V_i$, 皆有 $\pi_i(\mathbf{v}) = \mathbf{v}$. 另一方面由於 V_i 為 T-invariant, 對於 $\mathbf{v} \in V_i$, 我們有 $T(\mathbf{v}) \in V_i$. 因此對於任意 $\mathbf{v} \in V$, 將之寫成 $\mathbf{v} = \mathbf{v}_1 + \cdots + \mathbf{v}_k$, 其中 $\mathbf{v}_i \in V_i$, 則 $T(\pi_i(\mathbf{v})) = T(\mathbf{v}_i)$, 而 $\pi_i(T(\mathbf{v})) = \pi_i(T(\mathbf{v}_1) + \cdots + T(\mathbf{v}_k)) = T(\mathbf{v}_i)$. 得證

$$T \circ \pi_i = \pi_i \circ T, \forall i = 1, \ldots, k. \tag{4.1}$$

Theorem 4.2.5. 假設 V 為 *finite dimensional* F*-space.* 若 $T : V \to V$ 為 *linear operator* 其 *minimal polynomial* 為

$$\mu_T(x) = (x-\lambda_1)^{m_1} \cdots (x-\lambda_k)^{m_k}$$

其中 $\lambda_1, \ldots, \lambda_k$ 為 F 中相異的元素, 則 $T = T_D + T_N$ 其中 T_D 為 *diagonalizable,* T_N 為 *nilpotent of index* $m = \max\{m_1, \ldots, m_k\}$, 而且 $T_D \circ T_N = T_N \circ T_D$.

Proof. 考慮 Primary Decomposition $V=V_1\oplus\cdots\oplus V_k$, 其中 $V_i=\mathrm{Ker}((T-\lambda_i\,\mathrm{id})^{\circ m_i})$, 且令 $\pi_i:V\to V$ 為 the *projection* to V_i with respect to the direct sum $V=V_1\oplus\cdots\oplus V_k$. 考慮 V 的 linear operator $T_D=\lambda_1\pi_1+\cdots+\lambda_k\pi_k$. 因對任意 $\mathbf{v}_i\in V_i$, 皆有 $T_D(\mathbf{v}_i)=\lambda_i\mathbf{v}_i$, 所以每一組 V_i 的 basis, 皆由 T_D 的 eigenvectors 所組成. 故由 V 為 $V_1,\ldots,V_k$ 的 direct sum, 這些 V_i 的 basis 可組成 V 的 basis. 也就是說 V 有一組 basis 皆由 T_D 的 eigenvectors 所組成, 故 T_D 為 diagonalizable.

現令 $T_N=T-T_D$ 為 V 的 linear operator. 因對任意 $\mathbf{v}_i\in V_i$, $T_N(\mathbf{v}_i)=T(\mathbf{v}_i)-T_D(\mathbf{v}_i)=T(\mathbf{v}_i)-\lambda_i\mathbf{v}_i\in V_i$, 知 V_i 皆為 T_N-invariant. 又已知 $\mu_{T|_{V_i}}(x)=(x-\lambda_i)^{m_i}$, 即 m_i 是最小的正整數使得 $(T-\lambda_i\,\mathrm{id})^{\circ m_i}(\mathbf{v}_i)=\mathbf{O}_V$, $\forall\mathbf{v}_i\in V_i$, 故知 $\mu_{T_N|_{V_i}}(x)=x^{m_i}$. 利用 Lemma 3.5.6 知 $\mu_{T_N}(x)=\mathrm{lcm}(x^{m_1},\ldots,x^{m_k})=x^m$, 其中 $m=\max\{m_1,\ldots,m_k\}$, 得證 T_N 為 nilpotent of index m.

最後因為

$$T_D\circ T=(\lambda_1\pi_1+\cdots+\lambda_k\pi_k)\circ T=\lambda_1(\pi_1\circ T)+\cdots+\lambda_k(\pi_k\circ T),$$

由等式 (4.1) 得

$$T\circ T_D=\lambda_1(T\circ\pi_1)+\cdots+\lambda_k(T\circ\pi_k)=T_D\circ T.$$

因此得證

$$T_D\circ T_N=T_D\circ(T-T_D)=T_D\circ T-T_D\circ T_D=T\circ T_D-T_D\circ T_D=(T-T_D)\circ T_D=T_N\circ T_D.$$

□

Question 4.8. 考慮 *Theorem 4.2.4* 中的 *ordered basis* β, 若 $[T]_\beta$ 為 *upper triangular matrix,* 則 *Theorem 4.2.5* 中的 T_D,T_N 其對 β 的 *representative matrix* $[T_D]_\beta,[T_N]_\beta$ 應為何?

Question 4.9. 你能利用 *Theorem 4.2.5,* 證明若 T 的 *minimal polynomial* $\mu_T(x)$ 可以完全分解成 $F[x]$ 中相異的 *monic* 一次多項式之乘積, 則 T 為 *diagonalizable?*

由 Theorem 4.2.5, 我們便能利用 triangular form 來計算 $T^{\circ i}$ 了. 由 $T_D\circ T_N=T_N\circ T_D$, 得

$$T^{\circ 2}=(T_D+T_N)\circ(T_D+T_N)={T_D}^{\circ 2}+T_D\circ T_N+T_N\circ T_D+{T_N}^{\circ 2}={T_D}^{\circ 2}+2T_D\circ T_N+{T_N}^{\circ 2}.$$

故利用數學歸納法可得以下的二項式展開

$$T^{\circ i}=\sum_{j=0}^{i}\binom{i}{j}{T_D}^{\circ i-j}\circ{T_N}^{\circ j}.$$

由於 T_D 為 diagonalizable 我們很容易計算 ${T_D}^{\circ j}$, 而 T_N 為 nilpotent of index m, 我們知道當 $j\geq m$, ${T_N}^{\circ j}=\mathbf{O}$. 所以這是一個幫助我們計算 $T^{\circ i}$ 的方法.

最後我們來看 linear operator 相對應到 $n\times n$ matrix 的結論.

Theorem 4.2.6. 假設 $A\in M_n(F)$ 且其 *characteristic polynomial* 和 *minimal polynomial* 分別為

$$\chi_A(x)=(x-\lambda_1)^{c_1}\cdots(x-\lambda_k)^{c_k},\ \mu_A(x)=(x-\lambda_1)^{m_1}\cdots(x-\lambda_k)^{m_k}$$

其中 $\lambda_1,\dots,\lambda_k$ 為 F 中相異的元素. 則存在 *invertible matrix* P 使得

$$P^{-1}\cdot A\cdot P=\begin{pmatrix} A_1 & & \mathbf{O} \\ & \ddots & \\ \mathbf{O} & & A_k \end{pmatrix},$$

其中每個 A_i 為 $c_i\times c_i$ 階的 *upper triangular matrix*

$$\begin{pmatrix} \lambda_i & * & * \\ & \ddots & * \\ \mathbf{O} & & \lambda_i \end{pmatrix}.$$

此時我們有 $P^{-1}\cdot A\cdot P=D+N$, 其中 D 為 *diagonal matrix,* N 為 *nilpotent matrix* 滿足 $D\cdot N=N\cdot D$ 且 $N^m=\mathbf{O}$, $m=\max\{m_1,\dots,m_k\}$.

同樣的將 A 化成 triangular form $P^{-1}\cdot A\cdot P=D+N$, 由於 $D\cdot N=N\cdot D$, 我們有

$$P^{-1}\cdot A^i\cdot P=\sum_{j=0}^{i}\binom{i}{j}D^{i-j}\cdot N^j.$$

假設 $A\in M_n(F)$ 且 $\chi_A(X)=(x-\lambda_1)^{c_1}\cdots(x-\lambda_k)^{c_k}$, $\mu_A(x)=(x-\lambda_1)^{m_1}\cdots(x-\lambda_k)^{m_k}$. 我們說明如何找到 invertible matrix M 使得 $M^{-1}\cdot A\cdot M$ 為 upper triangular matrix. 首先我們利用 Chapter 3 primary decomposition 的方法找到 invertible matrix P 使得 $P^{-1}\cdot A\cdot P$ 為 block diagonal matrix

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

接著考慮每一個 $c_i\times c_i$ matrix A_i. 因為 $\mu_{A_i}(x)=(x-\lambda_i)^{m_i}$, $A_i-\lambda_i I_{c_i}$ 是 nilpotent of index m_i, 我們可以利用 Proposition 4.2.2 的方法首先找 $(A_i-\lambda_i I_{c_i})^{m_i-1}$ 的 column space 的一組 basis (此即相對於 Proposition 4.2.2 中 $\mathrm{Im}(T^{\circ m-1})$ 的 basis), 然後擴大成 $(A_i-\lambda_i I_{c_i})^{m_i-2}$ 的 column space 的一組 basis, 這樣一直下去直到擴大成 F^{c_i} 的一組 basis. 若令這組 basis 以 column by column 依序組成的 $c_i\times c_i$ 的 matrix 為 Q_i, 則我們有 $Q_i^{-1}\cdot A_i\cdot Q_i$ 為 upper triangular matrix. 最後將這些 Q_i 在 diagonal 的位置依序放入, 組成 $n\times n$ 的 invertible matrix

$$\begin{pmatrix} Q_1 & & \mathbf{O} \\ & \ddots & \\ \mathbf{O} & & Q_k \end{pmatrix},$$

就會使得

$$(P\cdot Q)^{-1}\cdot A\cdot(P\cdot Q)=Q^{-1}\cdot(P^{-1}\cdot A\cdot P)\cdot Q=\begin{pmatrix} Q_1^{-1}\cdot A_1\cdot Q_1 & & \mathbf{O} \\ & \ddots & \\ \mathbf{O} & & Q_k^{-1}\cdot A_k\cdot Q_k \end{pmatrix},$$

為 upper triangular matrix 了. 我們看以下的例子.

Example 4.2.7. 在 Example 3.5.10 中我們考慮 5×5 matrix

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

因為 $\chi_A(x)=\mu_A(x)=(x-1)^3(x-2)^2$ 在 $\mathbb{Q}[x]$ 中完全分解成一次多項式的乘積, 我們可找到 invertible matrix $M\in M_5(\mathbb{Q})$ 使得 $M^{-1}\cdot A\cdot M$ 為 upper triangular matrix.

在 Example 3.5.10 中我們已找到 $P\in M_5(\mathbb{Q})$ 將 A 化為 block diagonal matrix.

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

現在我們需將

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

化為 triangular forms. 因 $\mu_B(x)=(x-1)^3$, 考慮 $B-I_3$ 這一個 nilpotent matrix. 我們有

$$B-I_3=\begin{pmatrix} -1 & -1 & -1\\ 1 & 0 & 0\\ 0 & 1 & 1 \end{pmatrix},\quad (B-I_3)^2=\begin{pmatrix} 0 & 0 & 0\\ -1 & -1 & -1\\ 1 & 1 & 1 \end{pmatrix}.$$

由 Proposition 4.2.2 的方法首先選 $(B-I_3)^2$ 的 column space 的 basis, 我們選 $\mathbf{w}_1=(0,-1,1)^t$, 再加入 $B-I_3$ 的 column space 的元素 $\mathbf{w}_2$ 使得 $\{\mathbf{w}_1,\mathbf{w}_2\}$ 為 $B-I_3$ 的 column space 的 basis, 這裡我們選 $\mathbf{w}_2=(-1,1,0)^t$. 最後再加入 $\mathbf{w}_3\in\mathbb{Q}^3$ 使得 $\{\mathbf{w}_1,\mathbf{w}_2,\mathbf{w}_3\}$ 成為 $\mathbb{Q}^3$ 的 basis, 此處我們選 $\mathbf{w}_3=(0,0,1)^t$. 此時有 $B\mathbf{w}_1=\mathbf{w}_1, B\mathbf{w}_2=\mathbf{w}_1+\mathbf{w}_2, B\mathbf{w}_3=\mathbf{w}_1+\mathbf{w}_2+\mathbf{w}_3$, 故若令 $Q_1=\begin{pmatrix} 0 & -1 & 0\\ -1 & 1 & 0\\ 1 & 0 & 1 \end{pmatrix}$, 則 $Q_1^{-1}\cdot B\cdot Q_1=\begin{pmatrix} 1 & 1 & 1\\ 0 & 1 & 1\\ 0 & 0 & 1 \end{pmatrix}$ 為 upper triangular matrix.

另一方面因 $\mu_C(x)=(x-2)^2$, 我們考慮 $C-2I_2$ 這一個 nilpotent matrix. 因 $C-2I_2=\begin{pmatrix} -1 & -1\\ 1 & 1 \end{pmatrix}$, 我們選 $\mathbf{u}_1=\begin{pmatrix} -1\\ 1 \end{pmatrix}$ 為 $C-2I_2$ 的 basis, 再加上 $\mathbf{u}_2=\begin{pmatrix} 1\\ 0 \end{pmatrix}$ 使得 $\{\mathbf{u}_1,\mathbf{u}_2\}$ 為 $\mathbb{Q}^2$ 的 basis. 此時 $C\mathbf{u}_1=2\mathbf{u}_1$, $C\mathbf{u}_2=\mathbf{u}_1+2\mathbf{u}_2$, 故若令 $Q_2=\begin{pmatrix} -1 & 1\\ 1 & 0 \end{pmatrix}$, 則 $Q_2^{-1}\cdot C\cdot Q_2=\begin{pmatrix} 2 & 1\\ 0 & 2 \end{pmatrix}$ 為 upper triangular matrix. 最後將 Q_1,Q_2 合併為 5×5 的 invertible matrix

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

可得 upper triagular matrix

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

4.3. Jordan Form

將矩陣化為 Triangular form 並不容易讓我們判斷兩個矩陣是否為 similar. 我們將挑選更好的 ordered basis 將其化為所謂的 Jordan form. 本節中我們仍假設 $\chi_T(x)$ 可以完全分解成一次的多項式的乘積 (即 $\chi_T(x)=(x-\lambda_1)^{c_1}\cdots(x-\lambda_k)^{c_k}$). 同樣的我們先討論 nilpotent 的情形.

對於一個 linear operator $T:V\to V$. 這一次我們探討 $T,T^{\circ 2},\ldots$ 的 kernel 間的關係. 若 $\mathbf{v}\in\mathrm{Ker}(T^{\circ i})$, 表示 $T^{\circ i}(\mathbf{v})=\mathbf{O}_V$, 故得 $T^{\circ i+1}(\mathbf{v})=T(T^{\circ i}(\mathbf{v}))=\mathbf{O}_V$. 所以我們自然有以下的 chain of subspaces

$$\{\mathbf{O}_V\}\subseteq\mathrm{Ker}(T)\subseteq\mathrm{Ker}(T^{\circ 2})\subseteq\cdots\subseteq\mathrm{Ker}(T^{\circ i-1})\subseteq\mathrm{Ker}(T^{\circ i})\subseteq\cdots.$$

特別的, 當 T 為 nilpotent of index m, 我們有 $\mathrm{Ker}(T^{\circ i+1})\neq\mathrm{Ker}(T^{\circ i})$, $\forall i=1,\ldots,m-1$.

Lemma 4.3.1. 假設 $\dim(V)>0$, 若 T 為 *nilpotent operator of index m,* 則我們有以下的 *chain of subspaces.*

$$\{\mathbf{O}_V\}\subsetneq\mathrm{Ker}(T)\subsetneq\mathrm{Ker}(T^{\circ 2})\subsetneq\cdots\subsetneq\mathrm{Ker}(T^{\circ i-1})\subsetneq\mathrm{Ker}(T^{\circ i})\subsetneq\cdots\subsetneq\mathrm{Ker}(T^{\circ m-1})\subsetneq\mathrm{Ker}(T^{\circ m})=V.$$

Proof. 首先說明 $\mathrm{Ker}(T^{\circ m-1})\subsetneq\mathrm{Ker}(T^{\circ m})=V$. 因為 $T^{\circ m}=\mathbf{O}$, 亦即對任意 $\mathbf{v}\in V, T^{\circ m}(\mathbf{v})=\mathbf{O}_V$, 所以 $\mathrm{Ker}(T^{\circ m})=V$. 另一方面, 若 $\mathrm{Ker}(T^{\circ m-1})=\mathrm{Ker}(T^{\circ m})=V$, 則表示 $T^{\circ m-1}=\mathbf{O}$, 此與 m 為最小的正整數使得 $T^{\circ m}=\mathbf{O}$ 相矛盾, 故知 $\mathrm{Ker}(T^{\circ m-1})\neq\mathrm{Ker}(T^{\circ m})$.

接下來我們說明 $\{\mathbf{O}_V\}\subsetneq\mathrm{Ker}(T)$. 對任意 $\mathbf{v}\in V$ 因 $\mathbf{O}_V=T^{\circ m}(\mathbf{v})=T(T^{\circ m-1}(\mathbf{v}))$ 得 $T^{\circ m-1}(\mathbf{v})\in\mathrm{Ker}(T)$. 現若 $\mathrm{Ker}(T)=\{\mathbf{O}_V\}$, 則任意 $\mathbf{v}\in V$ 皆滿足 $T^{\circ m-1}(\mathbf{v})=\mathbf{O}_V$ 而得到 $T^{\circ m-1}=\mathbf{O}$ 之矛盾, 故知 $\mathrm{Ker}(T)\neq\{\mathbf{O}_V\}$.

同理, 當 $1\leq i\leq m-2$, 因對於所有 $\mathbf{v}\in V$ 皆有 $\mathbf{O}_V=T^{\circ m}(\mathbf{v})=T^{\circ i+1}(T^{\circ m-(i+1)}(\mathbf{v}))$, 即 $T^{\circ m-(i+1)}(\mathbf{v})\in\mathrm{Ker}(T^{\circ i+1})$. 現若 $\mathrm{Ker}(T^{\circ i})=\mathrm{Ker}(T^{\circ i+1})$, 則由 $T^{\circ m-(i+1)}(\mathbf{v})\in\mathrm{Ker}(T^{\circ i})$ 得 $\mathbf{O}_V=T^{\circ i}(T^{\circ m-(i+1)}(\mathbf{v}))=T^{\circ m-1}(\mathbf{v})$. 因為這是對任意 $\mathbf{v}\in V$ 皆成立, 故得到 $T^{\circ m-1}=\mathbf{O}$ 之矛盾, 得證 $\mathrm{Ker}(T^{\circ i})\neq\mathrm{Ker}(T^{\circ i+1})$. □

現假設 $i\geq 2$, 若 $\mathbf{v}_1,\ldots,\mathbf{v}_s\in\mathrm{Ker}(T^{\circ i+1})$ 為 linearly independent 且

$$\mathrm{Span}(\mathbf{v}_1,\ldots,\mathbf{v}_s)\cap\mathrm{Ker}(T^{\circ i})=\{\mathbf{O}_V\},$$

則 $T(\mathbf{v}_1),\ldots,T(\mathbf{v}_s)\in\mathrm{Ker}(T^{\circ i})$ 亦為 linearly independent. 事實上若

$$r_1T(\mathbf{v}_1)+\cdots+r_sT(\mathbf{v}_s)=\mathbf{O}_V,$$

則由 $T(r_1\mathbf{v}_1+\cdots+r_s\mathbf{v}_s)=\mathbf{O}_V$ 得 $r_1\mathbf{v}_1+\cdots+r_s\mathbf{v}_s\in\operatorname{Ker}(T)\subseteq\operatorname{Ker}(T^{\circ i})$. 故由 $\operatorname{Span}(\mathbf{v}_1,\ldots,\mathbf{v}_s)\cap\operatorname{Ker}(T^{\circ i})=\{\mathbf{O}_V\}$ 之假設得 $r_1\mathbf{v}_1+\cdots+r_s\mathbf{v}_s=\mathbf{O}_V$, 再由 $\mathbf{v}_1,\ldots,\mathbf{v}_s$ 為 linearly independent 得 $r_1=\cdots=r_s=0$, 得證 $T(\mathbf{v}_1),\ldots,T(\mathbf{v}_s)$ 為 linearly independent. 另外我們也可得

$$\operatorname{Span}(T(\mathbf{v}_1),\ldots,T(\mathbf{v}_s))\cap\operatorname{Ker}(T^{\circ i-1})=\{\mathbf{O}_V\}.$$

這是因為若 $\mathbf{v}=r_1T(\mathbf{v}_1)+\cdots+r_sT(\mathbf{v}_s)\in\operatorname{Ker}(T^{\circ i-1})$, 則

$$\mathbf{O}_V=T^{\circ i-1}(r_1T(\mathbf{v}_1)+\cdots+r_sT(\mathbf{v}_s))=T^{\circ i}(r_1\mathbf{v}_1+\cdots+r_s\mathbf{v}_s),$$

即 $r_1\mathbf{v}_1+\cdots+r_s\mathbf{v}_s\in\operatorname{Span}(\mathbf{v}_1,\ldots,\mathbf{v}_s)\cap\operatorname{Ker}(T^{\circ i})=\{\mathbf{O}_V\}$. 再由 $\mathbf{v}_1,\ldots,\mathbf{v}_s$ 為 linearly independent 得 $r_1=\cdots=r_s=0$, 得證 $\mathbf{v}=\mathbf{O}_V$. 我們有以下之結論.

Lemma 4.3.2. 假設 $T:V\to V$ 為 *linear operator.* 當 $i\geq 2$ 時, 若 $\mathbf{v}_1,\ldots,\mathbf{v}_s\in\operatorname{Ker}(T^{\circ i+1})$ 為 *linearly independent* 且 $\operatorname{Span}(\mathbf{v}_1,\ldots,\mathbf{v}_s)\cap\operatorname{Ker}(T^{\circ i})=\{\mathbf{O}_V\}$, 則 $T(\mathbf{v}_1),\ldots,T(\mathbf{v}_s)\in\operatorname{Ker}(T^{\circ i})$ 為 *linearly independent* 且 $\operatorname{Span}(T(\mathbf{v}_1),\ldots,T(\mathbf{v}_s))\cap\operatorname{Ker}(T^{\circ i-1})=\{\mathbf{O}_V\}$.

特別地, 若 V 為 *finite dimensional F-space,* 則

$$\dim(\operatorname{Ker}(T^{\circ i+1}))-\dim(\operatorname{Ker}(T^{\circ i}))\leq\dim(\operatorname{Ker}(T^{\circ i}))-\dim(\operatorname{Ker}(T^{\circ i-1})).\tag{4.2}$$

Proof. 我們僅剩要證明式子 (4.2). 假設 $\{\mathbf{u}_1,\ldots\mathbf{u}_t\}$ 為 $\operatorname{Ker}(T^{\circ i-1})$ 的一組 basis, 將之擴大成 $\{\mathbf{u}_1,\ldots\mathbf{u}_t,\mathbf{w}_1,\ldots,\mathbf{w}_l\}$ 使之為 $\operatorname{Ker}(T^{\circ i})$ 的一組 basis. 再擴大成 $\{\mathbf{u}_1,\ldots\mathbf{u}_t,\mathbf{w}_1,\ldots,\mathbf{w}_l,\mathbf{v}_1,\ldots,\mathbf{v}_s\}$ 使之為 $\operatorname{Ker}(T^{\circ i+1})$ 的一組 basis. 依此得 $\mathbf{v}_1,\ldots,\mathbf{v}_s\in\operatorname{Ker}(T^{\circ i+1})$ 為 linearly independent 且 $\operatorname{Span}(\mathbf{v}_1,\ldots,\mathbf{v}_s)\cap\operatorname{Ker}(T^{\circ i})=\{\mathbf{O}_V\}$. 由前面結果知 $T(\mathbf{v}_1),\ldots,T(\mathbf{v}_s)\in\operatorname{Ker}(T^{\circ i})$ 為 linearly independent 且 $\operatorname{Span}(T(\mathbf{v}_1),\ldots,T(\mathbf{v}_s))\cap\operatorname{Ker}(T^{\circ i-1})=\{\mathbf{O}_V\}$. 故 $\{\mathbf{u}_1,\ldots\mathbf{u}_t,T(\mathbf{v}_1),\ldots,T(\mathbf{v}_s)\}$ 為 $\operatorname{Ker}(T^{\circ i})$ 中的 linearly independent set. 得知 $t+s\leq\dim(\operatorname{Ker}(T^{\circ i}))=t+l$, 亦即

$$\dim(\operatorname{Ker}(T^{\circ i+1}))-\dim(\operatorname{Ker}(T^{\circ i}))=s\leq l=\dim(\operatorname{Ker}(T^{\circ i}))-\dim(\operatorname{Ker}(T^{\circ i-1})).$$

□

接下來我們先說明何謂 Jordan form, 然後再說明如何得到 Jordan form.

Definition 4.3.3. 給定 $\lambda\in F$, 對於 1×1 matrix (λ) 以及如下形式的更高階 square matrix

$$\begin{pmatrix}\lambda&0&0&\cdots&0&0\\1&\lambda&0&\cdots&0&0\\0&1&\lambda&0&\cdots&0\\\vdots&\ddots&\ddots&\ddots&\ddots&\vdots\\0&\cdots&0&1&\lambda&0\\0&0&\cdots&0&1&\lambda\end{pmatrix},$$

也就是說對角線 (i,i)-th entry 為 λ, 而對角線下方的位置即 $(i,i-1)$-th entry 為 1, 其他位置皆為 0 的矩陣, 我們稱為 *elementary Jordan matrix* associated with λ. 而由 associated with λ 的 elementary Jordan matrices 所組成的 block diagonal matrix, 即

$$\begin{pmatrix}J_1&&\mathbf{O}\\&\ddots&\\\mathbf{O}&&J_k\end{pmatrix},$$

其中每個 J_i 皆為 elementary Jordan matrix associated with λ, 稱為 *Jordan block matrix* associated with λ.

注意有些書本的 elementary Jordan matrix 的定義為 1 在對角線的上方 (即 $(i,i+1)$ 的位置), 不過只要將 ordered basis 順序前後對調, 不難發現這兩種矩陣為 similar.

Question 4.10. 假設 $T:V\to V$ 為 *linear operator* 且 $\beta=(\mathbf{v}_1,\mathbf{v}_2,\dots,,\mathbf{v}_{n-1},\mathbf{v}_n)$ 為 V 的 *ordered basis.* 若 $[T]_\beta$ 為 *elementary Jordan matrix associated with* λ, 考慮 *ordered basis* $\beta'=(\mathbf{v}_n,\mathbf{v}_{n-1},\dots,\mathbf{v}_2,\mathbf{v}_1)$, 則 $[T]_{\beta'}$ 為何種形式的 *matrix?*

接下來我們說明 nilpotent linear operator 皆可找到 ordered basis 使其 representative matrix 為 Jordan block matrix associated with 0.

Proposition 4.3.4. 假設 V 為 *finite dimensional F-space.* 若 $T:V\to V$ 是一個 *nilpotent linear operator of index m,* 則存在 V 的 *ordered basis* β 使得 $[T]_\beta$ 為 *Jordan block matrix associated with* 0.

Proof. 令 S_1 為 $\mathrm{Ker}(T)$ 的一組 basis, 將之擴大為 $\mathrm{Ker}(T^{\circ 2})$ 的一組 basis S_2, 一直下去直到得到 S_m 為 $\mathrm{Ker}(T^{\circ m})=V$ 的一組 basis. 也就是說當 $i=1,\dots,m$ 時 S_i 為 $\mathrm{Ker}(T^{\circ i})$ 的 basis (注意 Lemma 4.3.1 告訴我們當 $i=2,\dots,m$ 時 $S_{i-1}\subsetneq S_i$). 現考慮 $\{\mathbf{v}_1,\dots,\mathbf{v}_{k_1}\}=S_m\setminus S_{m-1}$ 這組 linear independent subset (它不是空集合). Corollary 1.4.4 告訴我們 $\mathrm{Span}(\{\mathbf{v}_1,\dots,\mathbf{v}_{k_1}\})$ 和 $\mathrm{Span}(S_{m-1})=\mathrm{Ker}(T^{\circ m-1})$ 的交集為 $\{\mathbf{O}_V\}$ 故由 Lemma 4.3.2 知 $\{T(\mathbf{v}_1),\dots,T(\mathbf{v}_{k_1})\}$ 為 $\mathrm{Ker}(T^{\circ m-1})$ 中的 linearly independent set 且 $\mathrm{Span}(\{T(\mathbf{v}_1),\dots,T(\mathbf{v}_{k_1})\})\cap\mathrm{Ker}(T^{\circ m-2})=\{\mathbf{O}_V\}$, 故利用 Corollary 1.4.4 知 $\{T(\mathbf{v}_1),\dots,T(\mathbf{v}_{k_1})\}\cup S_{m-2}$ 為 $\mathrm{Ker}(T^{\circ m-1})$ 中的 linearly independent set. 若 $\{T(\mathbf{v}_1),\dots,T(\mathbf{v}_{k_1})\}\cup S_{m-2}$ 亦為 $\mathrm{Ker}(T^{\circ m-1})$ 的 spanning set, 則它就是 $\mathrm{Ker}(T^{\circ m-1})$ 的一組 basis. 而若 $\{T(\mathbf{v}_1),\dots,T(\mathbf{v}_{k_1})\}\cup S_{m-2}$ 不是 $\mathrm{Ker}(T^{\circ m-1})$ 的 spanning set, 則我們可在 $\mathrm{Ker}(T^{\circ m-1})$ 中選取 $\mathbf{v}_{k_1+1},\dots,\mathbf{v}_{k_2}$ 使得

$$\{T(\mathbf{v}_1),\dots,T(\mathbf{v}_{k_1}),\mathbf{v}_{k_1+1},\dots,\mathbf{v}_{k_2}\}\cup S_{m-2}$$

為 $\mathrm{Ker}(T^{\circ m-1})$ 中的一組 basis. 也就是說我們將集合 $S_{m-1}\setminus S_{m-2}$ 用

$$\{T(\mathbf{v}_1),\dots,T(\mathbf{v}_{k_1}),\mathbf{v}_{k_1+1},\dots,\mathbf{v}_{k_2}\}$$

取代. 注意此時 $\{\mathbf{v}_1,\dots,\mathbf{v}_{k_1},T(\mathbf{v}_1),\dots,T(\mathbf{v}_{k_1}),\mathbf{v}_{k_1+1},\dots,\mathbf{v}_{k_2}\}\cup S_{m-2}$ 仍為 $\mathrm{Ker}(T^{\circ m})=V$ 的一組 basis.

接下來考慮取代 $S_{m-1}\setminus S_{m-2}$ 的集合 $\{T(\mathbf{v}_1),\dots,T(\mathbf{v}_{k_1}),\mathbf{v}_{k_1+1},\dots,\mathbf{v}_{k_2}\}$. 再次利用 Lemma 4.3.2, 我們知 $\{T^{\circ 2}(\mathbf{v}_1),\dots,T^{\circ 2}(\mathbf{v}_{k_1}),T(\mathbf{v}_{k_1+1}),\dots,T(\mathbf{v}_{k_2})\}\cup S_{m-3}$ 為 $\mathrm{Ker}(T^{\circ m-2})$ 中的 linearly independent set. 所以再加入 $\mathrm{Ker}(T^{\circ m-2})$ 中的元素 $\mathbf{v}_{k_2+1},\dots,\mathbf{v}_{k_3}$ 使得

$$\{T^{\circ 2}(\mathbf{v}_1),\dots,T^{\circ 2}(\mathbf{v}_{k_1}),T(\mathbf{v}_{k_1+1}),\dots,T(\mathbf{v}_{k_2}),\mathbf{v}_{k_2+1},\dots,\mathbf{v}_{k_3}\}\cup S_{m-3}$$

為 $\mathrm{Ker}(T^{\circ m-2})$ 的 basis. 也就是說我們將 $S_{m-2}\setminus S_{m-3}$ 集合用

$$\{T^{\circ 2}(\mathbf{v}_1),\dots,T^{\circ 2}(\mathbf{v}_{k_1}),T(\mathbf{v}_{k_1+1}),\dots,T(\mathbf{v}_{k_2}),\mathbf{v}_{k_2+1},\dots,\mathbf{v}_{k_3}\}$$

取代. 這樣一直下去, 簡單的說就是將取代 $S_{i+1} \setminus S_i$ 的集合 S'_{i+1} 代入 T 得到 $T(S'_{i+1})$ 這一組在 $\mathrm{Ker}(T^{\circ i})$ 的 linearly independent set, 再加入 $\mathrm{Ker}(T^{\circ i})$ 中的子集合 S''_i 使得 $T(S'_{i+1}) \cup S''_i \cup S_{i-1}$ 為 $\mathrm{Ker}(T^{\circ i})$ 的 basis. 接著就是將 S_i/S_{i-1} 用 $S'_i = T(S'_{i+1}) \cup S''_i$ 取代. 我們用以下圖表來表示:

$$\begin{array}{lllllll}
S_m \setminus S_{m-1} & \mathbf{v}_1, \dots \mathbf{v}_{k_1} & & & & & \\
S_{m-1} \setminus S_{m-2} \rightsquigarrow & T(\mathbf{v}_1), \dots, T(\mathbf{v}_{k_1}), & \mathbf{v}_{k_1+1}, \dots, \mathbf{v}_{k_2} & & & & \\
S_{m-2} \setminus S_{m-3} \rightsquigarrow & T^{\circ 2}(\mathbf{v}_1), \dots, T^{\circ 2}(\mathbf{v}_{k_1}), & T(\mathbf{v}_{k_1+1}), \dots, T(\mathbf{v}_{k_2}), & \mathbf{v}_{k_2+1}, \dots, \mathbf{v}_{k_3} & & & \\
\vdots & \vdots & \vdots & \vdots & & \vdots & \\
S_1 \rightsquigarrow & T^{\circ m-1}(\mathbf{v}_1), \dots, T^{\circ m-1}(\mathbf{v}_{k_1}), & T^{\circ m-2}(\mathbf{v}_{k_1+1}), \dots, T^{\circ m-2}(\mathbf{v}_{k_2}), & T^{\circ m-3}(\mathbf{v}_{k_2+1}), \dots, T^{\circ m-3}(\mathbf{v}_{k_3}), & \cdots & T(\mathbf{v}_{k_{m-2}+1}), \dots, T(\mathbf{v}_{k_{m-1}}), & \mathbf{v}_{k_{m-1}+1}, \dots, \mathbf{v}_{k_m}
\end{array}$$

最後一個步驟就是將取代 $S_2 \setminus S_1$ 的元素

$$T^{\circ m-2}(\mathbf{v}_1), \dots, T^{\circ m-2}(\mathbf{v}_{k_1}), T^{\circ m-3}(\mathbf{v}_{k_1+1}), \dots, T^{\circ m-3}(\mathbf{v}_{k_2}), \dots,$$
$$T(\mathbf{v}_{k_{m-3}+1}), \dots, T(\mathbf{v}_{k_{m-2}}), \mathbf{v}_{k_{m-2}+1}, \dots, \mathbf{v}_{k_{m-1}}$$

代入 T, 得到 $\mathrm{Ker}(T)$ 中的 linearly independent set

$$\{T^{\circ m-1}(\mathbf{v}_1), \dots, T^{\circ m-1}(\mathbf{v}_{k_1}), T^{\circ m-2}(\mathbf{v}_{k_1+1}), \dots, T^{\circ m-2}(\mathbf{v}_{k_2}), \dots,$$
$$T^{\circ 2}(\mathbf{v}_{k_{m-3}+1}), \dots, T^{\circ 2}(\mathbf{v}_{k_{m-2}}), T(\mathbf{v}_{k_{m-2}+1}), \dots, T(\mathbf{v}_{k_{m-1}})\}$$

再加入 $\mathrm{Ker}(T)$ 中的元素 $\mathbf{v}_{k_{m-1}+1}, \dots, \mathbf{v}_{k_m}$ 使其成為 $\mathrm{Ker}(T)$ 的 basis. 所以上面圖表的最後一個 row 的元素就是取代 S_1 的元素, 即 $\mathrm{Ker}(T)$ 的 basis. 將之與取代 $S_2 \setminus S_1$ 的集合聯集就是取代 S_2 的元素, 即 $\mathrm{Ker}(T^{\circ 2})$ 的 basis. 同理圖表中取代 $S_i \setminus S_{i-1}$ 的那一個 row 及其以下各 row 的元素即為組成 $\mathrm{Ker}(T^{\circ i})$ 的 basis. 也因此上表中的各元素即組成 $V = \mathrm{Ker}(T^{\circ m})$ 的 basis. 考慮將它們按順序一個 column 一個 column 由上往下排序所形成的 ordered basis β, 即 β 的第一個元素為 $\mathbf{v}_1$ 接著為 $T(\mathbf{v}_1)$, 一直到第 m 個為 $T^{\circ m-1}(\mathbf{v}_1)$ 接著放 $\mathbf{v}_2, T(\mathbf{v}_2), \dots$ 這樣一直下去最後依序為 $\mathbf{v}_{k_{m-1}}, T(\mathbf{v}_{k_{m-1}}), \mathbf{v}_{k_{m-1}+1}, \dots, \mathbf{v}_{k_m}$. 很容易看出 $[T]_\beta$ 便是一個 Jordan block matrix associated with 0. 例如 $[T]_\beta$ 對應 $(\mathbf{v}_1, T(\mathbf{v}_1), \dots, T^{\circ m-1}(\mathbf{v}_1))$ 的部分的 block 就是

$$\begin{pmatrix}
0 & 0 & 0 & \cdots & 0 & 0 \\
1 & 0 & 0 & \cdots & 0 & 0 \\
0 & 1 & 0 & 0 & \cdots & 0 \\
\vdots & \ddots & \ddots & \ddots & \ddots & \vdots \\
0 & \cdots & 0 & 1 & 0 & 0 \\
0 & 0 & \cdots & 0 & 1 & 0
\end{pmatrix}$$

這一個 associated with 0 的 $m \times m$ 階 elementary Jordan matrix. □

Question 4.11. 在上面證明中 $[T]_\beta$ 對應 $(\mathbf{v}_{k_1+1}, T(\mathbf{v}_{k_1+1}), \dots, T^{\circ m-2}(\mathbf{v}_1))$ 的部分的 *block* 是幾階的 *elementary Jordan matrix?* 對應到 $(\mathbf{v}_{k_{m-1}}, T(\mathbf{v}_{k_{m-1}}))$ 和 $(\mathbf{v}_{k_{m-1}+1}, \dots, \mathbf{v}_{k_m})$ 部分又是怎樣的 *matrices?*

Example 4.3.5. 考慮 linear operator $T : P_4(\mathbb{R}) \to P_4(\mathbb{R})$ 定義為 $T(ax^4 + bx^3 + cx^2 + dx + e) = 4ax^2 + 3bx + 2c$, 很容易判斷 T 為 nilpotent of index 3. 因 $\mathrm{Ker}(T) = \{dx + e \mid d, e \in \mathbb{R}\}$, 我們找到 $\mathrm{Ker}(T)$ 的一組 basis $S_1 = \{x, 1\}$. 而 $\mathrm{Ker}(T^{\circ 2}) = \{bx^3 + cx^2 + dx + e \mid b, c, d, e \in \mathbb{R}\}$, 我們

將 S_1 擴大為 $S_2 = \{x^3, x^2, x, 1\}$ 成為 $\mathrm{Ker}(T^{\circ 2})$ 的 basis. 最後將 S_2 擴大為 $S_3 = \{x^4, x^3, x^2, x, 1\}$ 使其為 $\mathrm{Ker}(T^{\circ 3}) = V$ 的 basis.

現 $S_3 \setminus S_2 = \{x^4\}$, 故考慮 $T(x^4) = 4x^2$. 其與 S_1 的聯集 $\{4x^2, x, 1\}$ 為 $\mathrm{Ker}(T^{\circ 2})$ 上的 linearly independent set, 可加入 x^3 使得 $\{4x^2, x^3\} \cup S_1 = \{4x^2, x^3, x, 1\}$ 為 $\mathrm{Ker}(T^{\circ 2})$ 的 basis. 此時用 $\{4x^2, x^3\}$ 取代 $S_2 \setminus S_1 = \{x^3, x^2\}$. 接著考慮 $T(4x^2) = 8, T(x^3) = 3x$, 因為 $\{8, 3x\}$ 已為 $\mathrm{Ker}(T)$ 的 basis 所以不需加入元素直接用 $\{8, 3x\}$ 取代 S_1. 我們有以下之圖表:

$$\begin{array}{rcll} S_3 \setminus S_2 = \{x^4\} & & x^4 & \\ S_2 \setminus S_1 = \{x^3, x^2\} & \rightsquigarrow & T(x^4) = 4x^2, & x^3 \\ S_1 = \{x, 1\} & \rightsquigarrow & T(T(x^4)) = 8, & T(x^3) = 3x \end{array}$$

所以考慮 ordered basis $\beta = (x^4, 4x^2, 8, x^3, 3x)$, 我們得到

$$[T]_\beta = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \end{pmatrix}.$$

其實在 Example 4.3.5, 我們不必真的找到 ordered basis 便能判斷 T 的 Jordan form 的可能形式. 利用 Proposition 4.3.4 的證明中所用的圖表, 我們知道 Jordan block matrix 中的 elementary Jordan matrix 的個數就是圖表中 column 的個數, 而 column 的個數就是圖表中最後一個 row 的元素個數, 即 $\#(S_1) = \dim(\mathrm{Ker}(T))$. 另外每一個 column 的元素個數就是它所對應的 elementary Jordan matrix 的階數. 例如第一個 column 就是來自 $S_m \setminus S_{m-1}$ 的元素 $\mathbf{v}_1$, 接著為 $T(\mathbf{v}_1), \ldots, T^{\circ m-1}(T)$ 共 m 個元素, 它所對應的就是一個 $m \times m$ 階的 elementary Jordan matrix. 這些 $m \times m$ 階的 elementary Jordan matrices 的個數就是那些有 m 個元數的 column 的個數, 由圖表中我們可以知道他們的個數就是 k_1, 即 $S_m \setminus S_{m-1}$ 的元素個數, 也就是 $\dim(\mathrm{Ker}(T^{\circ m})) - \dim(\mathrm{Ker}(T^{\circ m-1}))$. 由 Lemma 4.3.1, 我們知此數必大於 0. 至於 $(m-1) \times (m-1)$ 階的 elementary Jordan matrices 的個數, 就是 Proposition 4.3.4 中的 $k_2 - k_1$, 即 $\dim(\mathrm{Ker}(T^{\circ m-1})) - \dim(\mathrm{Ker}(T^{\circ m-2})) - (\dim(\mathrm{Ker}(T^{\circ m})) - \dim(\mathrm{Ker}(T^{\circ m-1})))$, 此數有可能為 0 (Lemma 4.2 僅告訴我們它大於等於 0). 同理, $i \times i$ 階的 elementary Jordan matrices 的個數為 $\dim(\mathrm{Ker}(T^{\circ i})) - \dim(\mathrm{Ker}(T^{\circ i-1})) - (\dim(\mathrm{Ker}(T^{\circ i+1})) - \dim(\mathrm{Ker}(T^{\circ i})))$.

在 Example 4.3.5 中的 T 為 nilpotent of index 3, 故其 Jordan block matrix 中一定有 3×3 階的 elementary Jordan matrix. 但 $\dim(V) = 5$, 故僅能有一個 (否則 Jordan block matrix 的階數會大於等於 6×6). 所以剩下可能是僅有一個 2×2 階的 elementary Jordan matrix, 或是有兩個 1×1 階的 elementary Jordan matrix. 不過 $\dim(\mathrm{Ker}(T)) = 2$, 所以僅能有 2 個 elementary Jordan matrices, 因此排除後者. 知 T 化成的 Jordan block matrix 一定是由一個 3×3 階和一個 2×2 階的 elementary Jordan matrices 所組成.

現在我們回到 T 的 minimal polynomial 為 $\mu_T(x) = (x - \lambda)^m$ 的情形, 此時 $T - \lambda\,\mathrm{id}$ 為 nilpotent 所以由 Proposition 4.3.4 知存在 ordered basis β 使得 $[T - \lambda\,\mathrm{id}]_\beta = J$ 為一個 diagonal 皆為 0 的 Jordan block matrix J. 然而若 $\dim(V) = n$, 因 $[T - \lambda\,\mathrm{id}]_\beta = [T]_\beta - \lambda I_n$, 故得 $[T]_\beta = \lambda I_n + J$, 為一個 diagonal 皆為 λ 的 Jordan block matrix.

Theorem 4.3.6. 假設 V 為 *finite dimensional* F*-space.* 若 $T: V \to V$ 為 *linear operator* 其 *characteristic polynomial* 和 *minimal polynomial* 分別為

$$\chi_T(x) = (x-\lambda_1)^{c_1}\cdots(x-\lambda_k)^{c_k},\ \mu_T(x) = (x-\lambda_1)^{m_1}\cdots(x-\lambda_k)^{m_k}$$

其中 $\lambda_1,\ldots,\lambda_k$ 為 F 中相異的元素, 則存在 V 的 *ordered basis* β 使得

$$[T]_\beta = \begin{pmatrix} J_1 & & \mathbf{O} \\ & \ddots & \\ \mathbf{O} & & J_k \end{pmatrix},$$

其中每個 J_i 為 $c_i \times c_i$ 階的 *Jordan block matrix associated with* λ_i, 而且組成 J_i 的 *elementary Jordan matrices* 的個數就是 λ_i 的 *geometric multiplicity,* 即 $\dim(\mathrm{Ker}(T-\lambda_i\,\mathrm{id}))$. 另外 J_i 中最高階的 *elementary Jordan matrix* 為 $m_i \times m_i$ 階.

Proof. 由 Primary Decomposition Theorem 得 $V = V_1 \oplus \cdots \oplus V_k$, 其中 $V_i = \mathrm{Ker}((T-\lambda_i\,\mathrm{id})^{\circ m_i})$ 且 $\mu_{T|_{V_i}}(x) = (x-\lambda_i)^{m_i}$. 得 $T|_{V_i} - \lambda_i\,\mathrm{id}|_{V_i}$ 為 nilpotent, 故利用 Proposition 4.3.4, 我們知存在 β_i 為 V_i 的 ordered basis, 使得 $[T|_{V_i}]_{\beta_i}$ 為 J_i 這樣的 $c_i \times c_i$ 階的 Jordan block matrix associated with λ_i. 故將 $\beta_1,\ldots,\beta_k$ 依序排列形成 V 的 ordered basis β, 可得 $[T]_\beta$ 為所要的 Jordan matrix. □

Question 4.12. 你能利用 *Theorem 4.3.6,* 證明若 T 的 *characteristic polynomial* $\chi_T(x)$ 可以完全分解成 $F[x]$ 中的一次多項式之乘積, 且 T 的每一個 *eigenvalue* 的 *algebraic multiplicity* 等於其 *geometric multiplicity,* 則 T 為 *diagonalizable?*

在 Theorem 4.3.6 中 T 的 representative matrix 就稱為 T 的 *Jordan form.* 前面提過有些情況我們可以用 $\chi_T(x), \mu_T(x)$ 來得到 T 的 Jordan form, 不過有時並不盡然, 我們用以下了例子來探討.

Example 4.3.7. 假設 $T: V \to V$ 為 linear operator 且 $\chi_T(x) = (x-\lambda_1)^3(x-\lambda_2)^4$, 其中 $\lambda_1 \neq \lambda_2$. 我們來探討 T 的 Jordan form 可能的形式. 首先由 $\deg(\chi_T(x)) = 7$, 知 $\dim(V) = 7$, 所以 T 的 Jordan form 一定是 7×7 matrix. 另外因 $\chi_T(x)$ 有兩個相異的一次因式, 所以它的 Jordan form, 一定是有兩個 Jordan block matrix 所組成, 其中一個是 3×3 且 associated with λ_1, 另一個是 4×4 且 associated with λ_2. 現在問題是這兩個 Jordan block matrix 是由哪些 elementary Jordan matrix 所組成.

首先看 3×3 的 block Jordan matrix associated with λ_1 的可能情況. 若 $\mu_T(x)$ 可被 $(x-\lambda_1)^3$ 整除, 則由 Theorem 4.3.6 知此 block 中一定有一個 3×3 的 elementary Jordan matrix, 所以此時這一個 Jordan block 就是一個 3×3 的 elementary Jordan matrix. 若 $\mu_T(x)$ 僅可被 $(x-\lambda_1)^2$ 整除, 不能被 $(x-\lambda_1)^3$ 整除, 則同樣由 Theorem 4.3.6 知此 block 中一定有一個 2×2 的 elementary Jordan matrix, 但此 block 為 3×3 matrix, 所以應該還有一個 1×1 的 elementary Jordan matrix. 所以這種情形這個 block 是由一個 2×2 和一個 1×1 matrix 所組成. 最後若 $(x-\lambda_1) \mid \mu_T(x)$ 但 $(x-\lambda_1)^2 \nmid \mu_T(x)$, 可知此 block 是由 3 個 1×1 的 elementary Jordan matrix 所組成. 我們將結果列為下表

$x-\lambda_1$ 的次數	3	2	1
block Jordan matrix	$\begin{pmatrix} \lambda_1 & 0 & 0 \\ 1 & \lambda_1 & 0 \\ 0 & 1 & \lambda_1 \end{pmatrix}$	$\begin{pmatrix} \lambda_1 & 0 & 0 \\ 1 & \lambda_1 & 0 \\ 0 & 0 & \lambda_1 \end{pmatrix}$	$\begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_1 & 0 \\ 0 & 0 & \lambda_1 \end{pmatrix}$

所以 Jordan block matrix associated with λ_1 皆可由 $\mu_T(x)$ 確定.

至於 4×4 的 block Jordan matrix associated with λ_2 的可能情況, 在 $x-\lambda_2$ 可整除 $\mu_T(x)$ 的最高次數為 $4,3,1$ 時, 和前面一樣的, 此時組成此 block 的 elementary Jordan matrix 是可以被確定的. 我們列出下表:

$x-\lambda_2$ 的次數	4	3	1
block Jordan matrix	$\begin{pmatrix} \lambda_2 & 0 & 0 & 0 \\ 1 & \lambda_2 & 0 & 0 \\ 0 & 1 & \lambda_2 & 0 \\ 0 & 0 & 1 & \lambda_2 \end{pmatrix}$	$\begin{pmatrix} \lambda_2 & 0 & 0 & 0 \\ 1 & \lambda_2 & 0 & 0 \\ 0 & 1 & \lambda_2 & 0 \\ 0 & 0 & 0 & \lambda_2 \end{pmatrix}$	$\begin{pmatrix} \lambda_2 & 0 & 0 & 0 \\ 0 & \lambda_2 & 0 & 0 \\ 0 & 0 & \lambda_2 & 0 \\ 0 & 0 & 0 & \lambda_2 \end{pmatrix}$

若 $x-\lambda_2$ 可整除 $\mu_T(x)$ 的最高次數為 2 時, 此時 associated with λ_2 的 4×4 的 block Jordan matrix 中雖知一定有一個 2×2 的 elementary Jordan matrix, 但其他的部分有可能是一個 2×2 的 elementary Jordan matrix 或是有兩個 1×1 的 elementary Jordan matrices 所組成. 也就是說此時這一個 block 有可能是以下兩種情況:

$$\begin{pmatrix} \lambda_2 & 0 & 0 & 0 \\ 1 & \lambda_2 & 0 & 0 \\ 0 & 0 & \lambda_2 & 0 \\ 0 & 0 & 1 & \lambda_2 \end{pmatrix}, \quad \begin{pmatrix} \lambda_2 & 0 & 0 & 0 \\ 1 & \lambda_2 & 0 & 0 \\ 0 & 0 & \lambda_2 & 0 \\ 0 & 0 & 0 & \lambda_2 \end{pmatrix}$$

當然了就得利用 λ_2 的 geometric multiplicity (即 $\dim(\mathrm{Ker}(T-\lambda_2\,\mathrm{id}))$) 來判斷了. 若 geometric multiplicity 為 2 則 block Jordan matrix 為前者 (即由兩個 2×2 elementary Jordan matrices 所組成), 而 geometric multiplicity 為 3 時就是後者 (即由一個 2×2 和兩個 1×1 elementary Jordan matrices 所組成).

將兩個 Jordan block matrices 確定後, 我們就能決定 T 的 Jordan form 了, 例如當 $\mu_T(x)=(x-\lambda_1)^2(x-\lambda_2)^3$ 的情形 T 的 Jordan form 就可確定為

$$\begin{pmatrix} \lambda_1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & \lambda_1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \lambda_1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \lambda_2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & \lambda_2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & \lambda_2 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & \lambda_2 \end{pmatrix}.$$

而若 $\mu_T(x) = (x-\lambda_1)^2(x-\lambda_2)^2$ 時 T 的 Jordan form 就有兩種可能, 當 λ_2 的 geometric multiplicity 為 2 和 3 時, T 的 Jordan form 分別為

$$\begin{pmatrix} \lambda_1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & \lambda_1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \lambda_1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \lambda_2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & \lambda_2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & \lambda_2 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & \lambda_2 \end{pmatrix}, \quad \begin{pmatrix} \lambda_1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & \lambda_1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \lambda_1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \lambda_2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & \lambda_2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & \lambda_2 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & \lambda_2 \end{pmatrix}.$$

Question 4.13. 在 *Example 4.3.7* 中 $\chi_T(x) = (x-\lambda_1)^3(x-\lambda_2)^4$. 當 $\mu_T(x) = (x-\lambda_1)^2(x-\lambda_2)^2$ 時, λ_2 的 *geometric multiplicity* 為什麼不可能是 1 或 4 呢? 又此時能確定 λ_1 的 *geometric multiplicity* 嗎?

對於 $n \times n$ 的矩陣我們也有相對應的定理, 也就是說若 $A \in M_n(F)$ 其 characteristic polynomial 可以完全分解成 $F[x]$ 中的一次多項式乘積 $\chi_A(x) = (x-\lambda_1)^{c_1} \cdots (x-\lambda_k)^{c_k}$, 則存在 invertible matrix $P \in M_n(F)$ 使得

$$P^{-1} \cdot A \cdot P = \begin{pmatrix} J_1 & & \mathbf{O} \\ & \ddots & \\ \mathbf{O} & & J_k \end{pmatrix},$$

其中每個 J_i 為 $c_i \times c_i$ 階的 Jordan block matrix associated with λ_i. 這個 matrix 我們稱為 A 的 Jordan form.

我們說明如何找到 invertible matrix P 使得 $P^{-1} \cdot A \cdot P$ 為 Jordan form. 首先注意我們不必如得到 triangular form 的情形先把 A 化為 block diagonal matrix. 這是因為若 $\mu_A(x) = (x-\lambda_1)^{m_1} \cdots (x-\lambda_k)^{m_k}$, 利用 Proposition 4.3.4 的方法, 對於每一個 $i = 1, \ldots, k$, 我們必須找到一組 $(A-\lambda_i I_n)^{m_i}$ 的 null space $N((A-\lambda_i I_n)^{m_i})$ 的 basis (此即相對於 Proposition 4.3.4 中 $\mathrm{Ker}(T^{\circ m})$ 的 basis), 而 $N((A-\lambda_i I_n)^{m_i})$ 剛好是 primary decomposition 中所考慮的 invariant subspace, 所以我們不必重複做變換 basis 的動作. 要找到 P 的步驟如下: 對於每一個 $i = 1, \ldots, k$, 先找到 $N(A-\lambda_i I_n)$ 的 basis S_1, 再將之擴大成 S_2 使之成為 $N((A-\lambda_i)^2)$ 的 basis. 這樣一直下去直到得到 $N((A-\lambda_i)^{m_i})$ 的 basis S_{m_i}. 接下來令 $S_{m_i} \setminus S_{m_1-1} = \{\mathbf{v}_1, \ldots, \mathbf{v}_{k_1}\}$, 然後得 $\{A\mathbf{v}_1, \ldots, A\mathbf{v}_{k_1}\}$. 將之擴大成 $\{A\mathbf{v}_1, \ldots, A\mathbf{v}_{k_1}, \mathbf{v}_{k_1+1}, \ldots, \mathbf{v}_{k_2}\}$ 使之與 S_{m_i-2} 的聯集成為 $N((A-\lambda_i)^{m_i-1})$ 的 basis 並取代 $S_{m_i-1} \setminus S_{m_i-2}$. 這樣一直下去直到將 $S_2 \setminus S_1$ 取代完畢. 最後依 Proposition 4.3.4 將這些 bases 排序得到 P. 我們看以下的例子.

Example 4.3.8. 考慮 5×5 matrix

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

雖然在 Example 3.5.10 中我們找到 invertible matrix P 使得 $P^{-1} \cdot A \cdot P$ 為 block diagonal matrix, 接著在 Example 4.2.7 中我們又找到 invertible matrix Q 使得 $(P \cdot Q)^{-1} \cdot A \cdot (P \cdot Q)$ 為

upper triangular matrix. 不過要將 A 化為 Jordan form 就不必先將 A 化為 block diagonal matrix 再化為 Jordan form 這麼麻煩, 我們直接用 Proposition 4.3.4 的方法處理.

因 $\chi_A(x)=\mu_A(x)=(x-1)^3(x-2)^2$, 首先觀察 $A-I_5$ 的 null space 為 dimension 1, 我們得 $S_1=\{(0,0,-1,1,0)^t\}$ 為其 basis. 接著擴大成 $S_2=\{(0,0,-1,1,0)^t,(1,0,-1,0,1)^t\}$ 為 $(A-I_5)^2$ 的 null space 的 basis. 然後擴大成 $S_3=\{(0,0,-1,1,0)^t,(1,0,-1,0,1)^t,(-1,0,0,0,1)^t\}$ 為 $(A-I_5)^3$ 的 null space 的 basis. 因為 $\mathbf{v}_1=(-1,0,0,0,1)^t\in S_3\setminus S_2$, 我們考慮 $\mathbf{v}_2=A\mathbf{v}_1-\mathbf{v}_1=(-1,0,0,1,-1)^t$, 且將之取代 $S_2\setminus S_1$ 的元素. 最後得 $\mathbf{v}_3=A\mathbf{v}_2-\mathbf{v}_2=(0,0,1,-1,0)$ 取代 S_1 的元素. 此時我們有 $\{\mathbf{v}_1,\mathbf{v}_2,\mathbf{v}_3\}$ 為 $(A-I_5)^3$ 的 null space 的 basis 且

$$A\mathbf{v}_1=\mathbf{v}_1+\mathbf{v}_2, A\mathbf{v}_2=\mathbf{v}_2+\mathbf{v}_3, A\mathbf{v}_3=\mathbf{v}_3.$$

同樣的我們有 $S_1'=\{(0,-1,1,0,0)^t\}$ 為 $A-2I_5$ 的 null space 的 basis. 將之擴大為 $S_2'=\{(0,-1,1,0,0)^t,(1,0,0,0,0)\}$ 為 $(A-2I_5)^2$ 的 null space 的 basis. 故考慮 $\mathbf{v}_1'=(1,0,0,0,0)$ 並用 $\mathbf{v}_2'=A\mathbf{v}_1'-2\mathbf{v}_1'=(0,1,-1,0,0)^t$ 取代 $(0,-1,1,0,0)^t$. 此時我們有 $\{bv_1',\mathbf{v}_2'\}$ 為 $(A-2I_5)^2$ 的 null space 的 basis 且

$$A\mathbf{v}_1'=2\mathbf{v}_1'+\mathbf{v}_2', A\mathbf{v}_2'=2\mathbf{v}_2'.$$

最後令

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

在前面 Section 4.1, 我們提到當 $A,B\in M_n(F)$ 為 diagonalizable 時, 我們可以將其對角線位置的 eigenvalue 適當的重排來判斷 A,B 是否為 similar. 同樣的若 A,B 的 characteristic polynomial 在 $F[x]$ 可完全分解成一次的 monic polynomials 的乘積, 我們可以將 A,B 化為 Jordan form 來判斷它們是否為 similar. 當然了先決條件是 $\chi_A(x)=\chi_B(x)$ 且 $\mu_A(x)=\mu_B(x)$. 若其中有一個不相等便可知 A,B 不為 similar, 而若皆相等就需藉由 A,B 的 Jordan form 來確定. 對於 A,B 的每一個 eigenvalue, 若將 A,B 對於 associated with λ 的 block Jordan matrix 中的 elementary Jordan matrices 做適當重排後會相同, 則知 $A\sim B$. 反之, 若 $A\sim B$, 我們可將 A,B 視為某一個 linear operator T 用不同 ordered bases 所得的 representative matrices. 由於 associated λ 的 elementary Jordan matrices 的各個階數的個數告訴我們

$$\ldots,\dim(\mathrm{Ker}((T-\lambda\mathrm{id})^{\circ i-1})),\dim(\mathrm{Ker}((T-\lambda\mathrm{id})^{\circ i})),\dim(\mathrm{Ker}((T-\lambda\mathrm{id})^{\circ i+1})),\ldots$$

這些 dimensions 之間的關係, 而這些關係和 ordered basis 的選取無關, 所以 A,B associated λ 的 elementary Jordan matrices 的各個階數的個數會相同, 也就是 A,B 可以化為相同的 Jordan form. 因為 Jordan form 可以用來判定兩個 matrixes 是否為 similar, 所以 Jordan form 可以視為一種 canonical form.

最後我們談論 Jordan form 一個重要的應用. 回顧一下, 若 $A\in M_n(F)$ 則 A 和 A 的 transpose A^t 有相同的 characteristic polynomial 和 minimal polynomial. 這表示 A 和 A^t 有可能為 similar. 事實上當 $\chi_A(x)$ 可以在 $F[x]$ 中完全分解成一次的 monic polynomials

的乘積, 我們可得 $A \sim A^{\mathrm{t}}$. 這是因為當 $A \in M_n(F)$ 時, $\dim(N(A)) + \dim(C(A)) = n$. 又因為 $\dim(C(A)) = \dim(C(A^{\mathrm{t}}))$, 所以我們得 $\dim(N(A)) = \dim(N(A^{\mathrm{t}}))$. 同理, 對於每一個 A 的 eigenvalue λ (也會是 A^{t} 的 eigenvalue), 我們有

$$\dim(N((A-\lambda I_n)^i)) = \dim(N(((A-\lambda I_n)^{\mathrm{t}})^i)) = \dim(N((A^{\mathrm{t}}-\lambda I_n)^i)).$$

所以將 A 化為 Jordan form, 每一個階數的 elementary Jordan matrix associated with λ 和 A^{t} 同階的 elementary Jordan matrix associated with λ 個數都相同, 也就是說 A 和 A^{t} 可以化成同樣的 Jordan form. 我們有以下之結果.

Theorem 4.3.9. 假設 $A \in M_n(F)$. 若 $\chi_A(x)$ 可以在 $F[x]$ 中完全分解成一次的 *monic polynomials* 的乘積, 則 A 的 *transpose* A^{t} 和 A 為 *similar.*

以後我們會提到若 $A, B \in M_n(F)$ 且存在一個比 F 大的 field $\tilde{F}$ 使得在 $M_n(\tilde{F})$ 中 $A \sim B$ (即存在 $\tilde{P} \in M_n(\tilde{F})$ invertible 使得 $B = \tilde{P}^{-1} \cdot A \cdot \tilde{P}$), 則在 $M_n(F)$ 中 $A \sim B$ (即存在 $P \in M_n(F)$ invertible 使得 $B = P^{-1} \cdot A \cdot P$). 所以事實上 Theorem 4.3.9 不需 $\chi_A(x)$ 可以在 $F[x]$ 中完全分解成一次的 monic polynomials 的乘積之假設, 我們仍可得 $A \sim A^{\mathrm{t}}$.

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 $a_d, a_{d-1}, \dots, a_1, a_0 \in F$ 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 $a_d, \dots, a_1, a_0 \in F$. 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 $\mathrm{Span}(\mathbf{v})$?

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

$$\mu_{\mathbf{v}}(x)=x^d+a_{d-1}x^{d-1}+\cdots+a_1x+a_0,$$

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 minimal 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 $\deg(\mu_{T|_{C_{\mathbf{v}}}}(x))\leq\deg(\chi_{T|_{C_{\mathbf{v}}}}(x))=\deg(\mu_{\mathbf{v}}(x))$ 得證 $\mu_{T|_{C_{\mathbf{v}}}}(x)=\mu_{\mathbf{v}}(x)$. □

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 $\deg(g(x))\leq d-1$ 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 $T^{\circ 3}(\mathbf{v}) = -T(\mathbf{v}), \ldots$ 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 $[T]_\beta = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$.

Question 4.19. 若 k 是最小的正整數滿足 $\{\mathbf{v}, T(\mathbf{v}), \ldots, T^{\circ k-1}(\mathbf{v}), T^{\circ k}(\mathbf{v})\}$ 為 *linearly dependent,* 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 $g(x)=c_nx^n+\cdots+c_1x+c_0$, 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

$$\overline{g(T)}(\overline{\mathbf{v}})=\overline{g(T)(\mathbf{v})}=\overline{c_nT^{\circ n}(\mathbf{v})+\cdots+c_1T(\mathbf{v})+c_0\mathbf{v}}.$$

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 $g(\overline{T})=\overline{g(T)}$. $\square$

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

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

Proof. 依 $\mu_T(x)$ 的定義, 對任意 $\mathbf{v}\in V$, 皆有 $\mu_T(T)(\mathbf{v})=\mathbf{O}_V\in W$, 得 $\overline{\mu_T(T)(\mathbf{v})}=\overline{\mathbf{O}_V}=\mathbf{O}_{V/W}$. 故由 Lemma 4.4.7 知

$$\mu_T(\overline{T})(\overline{\mathbf{v}})=\overline{\mu_T(T)}(\overline{\mathbf{v}})=\overline{\mu_T(T)(\mathbf{v})}=\mathbf{O}_{V/W}.$$

利用 Lemma 3.3.5 (套用在 $\overline{T}$) 得 $\mu_{\overline{T}}(x)\mid\mu_T(x)$.

同理, 因 $\mu_{\mathbf{v}}(T)(\mathbf{v})=\mathbf{O}_V$, 我們得 $\mu_{\mathbf{v}}(\overline{T})(\overline{\mathbf{v}})=\mathbf{O}_{V/W}$, 故由 Lemma 4.4.3 (套用在 $\overline{\mathbf{v}}$ 以及 $\overline{T}$) 得 $\mu_{\overline{\mathbf{v}}}(x)\mid\mu_{\mathbf{v}}(x)$. $\square$

事實上在某些情況之下有可能 $\mu_{\overline{\mathbf{v}}}(x)=\mu_{\mathbf{v}}(x)$, 例如以下的情況.

Lemma 4.4.9. 令 $T:V\to V$ 為一個 *F-linear operator.* 給定 $\mathbf{v}\in V$, 考慮 $\overline{T}:V/C_{\mathbf{v}}\to V/C_{\mathbf{v}}$ 為 *linear operator induced by* T *on* $V/C_{\mathbf{v}}$. 若 $\mathbf{w}\in V$ 滿足 $\mu_{\mathbf{w}}(x)\mid\mu_{\mathbf{v}}(x)$, 則存在 $\mathbf{u}\in V$ 滿足 $\overline{\mathbf{u}}=\overline{\mathbf{w}}\in V/C_{\mathbf{v}}$ 且 $\mu_{\mathbf{u}}(x)=\mu_{\overline{\mathbf{u}}}(x)=\mu_{\overline{\mathbf{w}}}(x)$.

Proof. 因 $\mu_{\overline{\mathbf{w}}}(\overline{T})(\overline{\mathbf{w}})=\mathbf{O}_{V/C_{\mathbf{v}}}$, 利用 Lemma 4.4.7 得 $\overline{\mu_{\overline{\mathbf{w}}}(T)(\mathbf{w})}=\overline{\mathbf{O}_V}$, 亦即 $\mu_{\overline{\mathbf{w}}}(T)(\mathbf{w})\in C_{\mathbf{v}}$. 換言之, 存在 $f(x)\in F[x]$ 使得

$$\mu_{\overline{\mathbf{w}}}(T)(\mathbf{w})=f(T)(\mathbf{v}).\tag{4.4}$$

依 $\mu_{\mathbf{w}}(x) \mid \mu_{\mathbf{v}}(x)$ 之假設, 以及由 Corollary 4.4.8 知 $\mu_{\overline{\mathbf{w}}}(x) \mid \mu_{\mathbf{w}}(x)$ 可得 $\mu_{\overline{\mathbf{w}}}(x) \mid \mu_{\mathbf{v}}(x)$, 亦即存在 $h(x) \in F[x]$ 使得 $\mu_{\mathbf{v}}(x) = h(x)\mu_{\overline{\mathbf{w}}}(x)$. 故由等式 (4.4) 得

$$\mu_{\mathbf{v}}(T)(\mathbf{w}) = h(T) \circ \mu_{\overline{\mathbf{w}}}(T)(\mathbf{w}) = h(T) \circ f(T)(\mathbf{v}). \tag{4.5}$$

然而 $\mu_{\mathbf{w}}(x) \mid \mu_{\mathbf{v}}(x)$, 故由 Lemma 4.4.3 與等式 (4.5) 知 $\mathbf{O}_V = \mu_{\mathbf{v}}(T)(\mathbf{w}) = h(T) \circ f(T)(\mathbf{v})$. 再次利用 Lemma 4.4.3 得 $\mu_{\mathbf{v}}(x) \mid h(x)f(x)$, 亦即 $h(x)\mu_{\overline{\mathbf{w}}}(x) \mid h(x)f(x)$. 由此知 $\mu_{\overline{\mathbf{w}}}(x) \mid f(x)$, 亦即存在 $g(x) \in F[x]$ 使得

$$f(x) = \mu_{\overline{\mathbf{w}}}(x)g(x). \tag{4.6}$$

現令 $\mathbf{u} = \mathbf{w} - g(T)(\mathbf{v})$. 因 $g(T)(\mathbf{v}) \in C_{\mathbf{v}}$, 我們有 $\overline{\mathbf{u}} = \overline{\mathbf{w}} \in V/C_{\mathbf{v}}$. 利用 $\mu_{\overline{\mathbf{w}}}(T)$ 為 linear operator 得

$$\mu_{\overline{\mathbf{w}}}(T)(\mathbf{u}) = \mu_{\overline{\mathbf{w}}}(T)(\mathbf{w} - g(T)(\mathbf{v})) = \mu_{\overline{\mathbf{w}}}(T)(\mathbf{w}) - \mu_{\overline{\mathbf{w}}}(T) \circ g(T)(\mathbf{v}),$$

所以由等式 (4.6) 以及等式 (4.4) 得

$$\mu_{\overline{\mathbf{w}}}(T)(\mathbf{u}) = \mathbf{O}_V.$$

再次利用 Lemma 4.4.3 得 $\mu_{\mathbf{u}}(x) \mid \mu_{\overline{\mathbf{w}}}(x)$. 然而 $\overline{\mathbf{u}} = \overline{\mathbf{w}}$, 故 $\mu_{\overline{\mathbf{w}}}(x) = \mu_{\overline{\mathbf{u}}}(x)$, 即 $\mu_{\mathbf{u}}(x) \mid \mu_{\overline{\mathbf{u}}}(x)$. 再加上 Lemma 4.4.8 告訴我們 $\mu_{\overline{\mathbf{u}}}(x) \mid \mu_{\mathbf{u}}(x)$, 得證 $\mu_{\mathbf{u}}(x) = \mu_{\overline{\mathbf{u}}}(x)$. □

一般來說若 $\deg(\mu_{\overline{\mathbf{w}}}(x)) = d$, 雖然 $\{\overline{\mathbf{w}}, \overline{T}(\overline{\mathbf{w}}), \ldots, \overline{T}^{\circ d-1}(\overline{\mathbf{w}})\}$ 會是 $C_{\overline{\mathbf{w}}}$ 的一組 basis, 不過 $\{\mathbf{w}, T(\mathbf{w}), \ldots, T^{\circ d-1}(\mathbf{w})\}$ 就未必會是 $C_{\mathbf{w}}$ 的一組 basis. 不過在 Lemma 4.4.9 的假設條件下我們可找到 $\mathbf{u}$ 滿足 $\overline{\mathbf{u}} = \overline{\mathbf{w}}$ 且 $\{\mathbf{u}, T(\mathbf{u}), \ldots, T^{\circ d-1}(\mathbf{u})\}$ 和 $\{\overline{\mathbf{u}}, \overline{T}(\overline{\mathbf{u}}), \ldots, \overline{T}^{\circ d-1}(\overline{\mathbf{u}})\}$ 分別會是 $C_{\mathbf{u}}$ 與 $C_{\overline{\mathbf{u}}} = C_{\overline{\mathbf{w}}}$ 的一組 basis.

現在我們可以利用 primary decomposition theorem 證得以下重要的定理.

Theorem 4.4.10 (Cyclic Decomposition Theorem)**.** 假設 V 為 *finite dimensional F-space* 且 $T: V \to V$ 為 *linear operator*. 則 V 可以寫成一些 *T-cyclic subspaces* 的 *direct sum*. 事實上, 若 $\mu_T(x) = p_1(x)^{m_1} \cdots p_k(x)^{m_k}$, 其中 $p_i(x) \in F[x]$ 為相異的 *monic irreducible polynomial*, 則 $V = W_1 \oplus \cdots \oplus W_k$, 其中

$$W_i = \operatorname{Ker}(p_i(T)^{\circ m_i}) = C_{\mathbf{v}_{i,1}} \oplus \cdots \oplus C_{\mathbf{v}_{i,n_i}},$$

而且每個 $\mathbf{v}_{i,j}$ 的 *T-annihilator* 為 $p_i(x)^{m_{i,j}}$ 滿足 $m_i = m_{i,1} \geq m_{i,2} \geq \cdots \geq m_{i,n_i} > 0$.

Proof. 由 primary decomposition theorem, 我們知道 $T|_{W_i}: W_i \to W_i$ 的 minimal polynomial 為 $p_i(x)^{m_i}$. 若能證得每一個 W_i 可以寫成定理所述的 T-cyclic subspaces 的 direct sum, 則由 Corollary 3.4.7 可得 V 可以寫成一些 T-cyclic subspaces 的 direct sum. 所以我們僅要證明當 $T: V \to V$ 是 F-linear operator 且 $\chi_T(x) = p(x)^m$ 其中 $p(x) \in F[x]$ 是 monic irreducible polynomial 的情形下, 存在 $\mathbf{v}_1, \ldots \mathbf{v}_n \in V$ 使得 $V = C_{\mathbf{v}_1} \oplus \cdots \oplus C_{\mathbf{v}_n}$ 且對於 $1 \leq i \leq n$, $\mu_{\mathbf{v}_i}(x) = p(x)^{m_i}$ 滿足 $m = m_1 \geq m_2 \geq \cdots \geq m_n$.

我們利用對 $\dim(V)$ 作數學歸納法證明. 當 $\dim(V) = 1$ 時, 很自然對於任意 $\mathbf{v} \neq \mathbf{O}_V$ in V, 我們有 $V = C_{\mathbf{v}}$ 且 $\mu_T(x) = \mu_{\mathbf{v}}(x)$, 所以定理成立. 現假設此定理在維度小於 $\dim(V)$ 的情形都成立, 此時依假設 $\mu_T(x) = p(x)^m$, 故存在 $\mathbf{v}_1 \in V$ 滿足 $p(T)^{\circ m-1}(\mathbf{v}_1) \neq \mathbf{O}_V$. 因

$\mu_{\mathbf{v}_1}(x) \mid \mu_T(x) = p(x)^m$ 以及 $p(x)$ 為 irreducible, 所以存在 $m_1 \leq m$ 使得 $\mu_{\mathbf{v}_1}(x) = p(x)^{m_1}$. 然而前面假設 $p(T)^{\circ m-1}(\mathbf{v}_1) \neq \mathbf{O}_V$, 所以 $m_1 > m-1$, 因此得 $m_1 = m$.

現考慮 $\overline{T} : W/C_{\mathbf{v}_1} \to V/C_{\mathbf{v}_1}$ induced by T on $V/C_{\mathbf{v}_1}$. 注意此時 $\mu_{\overline{T}}(x) \mid \mu_T(x)$ (Corollary 4.4.8), 故知 $\mu_{\overline{T}}(x) = p(x)^{m'}$, 其中 $m' \leq m$. 因此由 $\dim(V/C_{\mathbf{v}_1}) < \dim(V)$, 我們可以套用數學歸納法之假設, 即存在 $\mathbf{w}_2, \ldots, \mathbf{w}_n \in V$ 滿足

$$V/C_{\mathbf{v}_1} = C_{\overline{\mathbf{w}_2}} \oplus \cdots \oplus C_{\overline{\mathbf{w}_n}},$$

且對於 $2 \leq i \leq n$, $\mu_{\overline{\mathbf{w}_i}}(x) = p(x)^{m_i}$ 滿足 $m \geq m' = m_2 \geq \cdots \geq m_n$. 又由於 $m_i \leq m = m_1$, 即 $\mu_{\mathbf{w}_i}(x) \mid \mu_{\mathbf{v}_1}(x)$, 所以利用 Lemma 4.4.9 知, 存在 $\mathbf{v}_i \in V$ 使得 $\overline{\mathbf{v}_i} = \overline{\mathbf{w}_i}$ 且 $\mu_{\mathbf{v}_i}(x) = \mu_{\overline{\mathbf{v}_i}}(x) = p(x)^{m_i}$.

現若 $\deg(p(x)) = d$, 由 direct sum 的性質 (Proposition 3.4.6) 以及 Theorem 4.4.4 知

$$\{\overline{\mathbf{v}_2}, \overline{T}(\overline{\mathbf{v}_2}), \ldots, \overline{T}^{\circ dm_2-1}(\overline{\mathbf{v}_2}), \ldots, \overline{\mathbf{v}_n}, \overline{T}(\overline{\mathbf{v}_n}), \ldots, \overline{T}^{\circ dm_n-1}(\overline{\mathbf{v}_n})\}$$

為 $V/C_{\mathbf{v}_1} = C_{\overline{\mathbf{v}_2}} \oplus \cdots \oplus C_{\overline{\mathbf{v}_n}}$ 的一組 basis. 現因 $\overline{T}^{\circ j}(\overline{\mathbf{v}_i}) = \overline{T^{\circ j}(\mathbf{v}_i)}$ (Lemma 4.4.7), 以及 $\{\mathbf{v}_1, T(\mathbf{v}_1), \ldots, T^{\circ dm_1-1}(\mathbf{v}_1)\}$ 為 $C_{\mathbf{v}_1}$ 的一組 basis, 利用 Proposition 1.6.2 的證明所用的方法我們得

$$\{\mathbf{v}_1, T(\mathbf{v}_1), \ldots, T^{\circ dm_1-1}(\mathbf{v}_1), \mathbf{v}_2, T(\mathbf{v}_2), \ldots, T^{\circ dm_2-1}(\mathbf{v}_2), \ldots, \mathbf{v}_n, T(\mathbf{v}_n), \ldots, T^{\circ dm_n-1}(\mathbf{v}_n)\}$$

為 V 的一組 basis. 因為對所有 $1 \leq i \leq n$, $\{\mathbf{v}_i, T(\mathbf{v}_i), \ldots, T^{\circ dm_i-1}(\mathbf{v}_i)\}$ 為 $C_{\mathbf{v}_i}$ 的一組 basis, 故由 direct sum 的性質 (Proposition 3.4.6) 得證

$$V = C_{\mathbf{v}_1} \oplus C_{\mathbf{v}_2} \oplus \cdots \oplus C_{\mathbf{v}_n}.$$

□

Question 4.20. 在 *Theorem 4.4.10* 的證明中, 為何要將 $\mathbf{w}_2, \ldots, \mathbf{w}_n$ 改成 $\mathbf{v}_1, \ldots, \mathbf{v}_n$?

Question 4.21. 可以用 *cyclic decomposition theorem* 說明若 $\mu_T(x) = (x-\lambda_1)\cdots(x-\lambda_k)$, 其中 $\lambda_i \neq \lambda_j$ *for* $i \neq j$, 則 T 是 *diagonalizable* 嗎?

利用 primary decomposition theorem, 我們可以找到 V 的 ordered basis β, 使得 $[T]_\beta$ 為以下的 block diagonal matrix

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

其中每個 A_i 的 minimal polynomial 為 $p_i(x)^{m_i}$. 而 cyclic decomposition theorem 告訴我們, β 可以由一些 cyclic vectors 所形成的 cyclic bases 所組成, 此時每一個 A_i 可寫成

$$\begin{pmatrix} C_{i,1} & & \mathbf{O} \\ & \ddots & \\ \mathbf{O} & & C_{i,n_i} \end{pmatrix},$$

其中每個 $C_{i,j}$ 是 the companion matrix of $p_i(x)^{m_{i,j}}$. 這也告訴我們任何的方陣都會 similar to 這樣形式的方陣, 我們稱此為 *rational form*.

Example 4.4.11. 考慮 over $\mathbb{R}$, 我們要求出 A 的 rational form, 其中

$$A = \begin{pmatrix} 2 & -5 & -1 & 6 & 1 \\ 1 & -2 & 0 & 3 & 1 \\ 0 & 0 & 2 & -1 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}.$$

首先算出 $\chi_A(x) = (x^2+1)(x-1)^3$, 再求得 $\mu_A(x) = (x^2+1)(x-1)^2$.

首先考慮 primary decomposition, 求出 A^2+I_5 與 $(A-I_5)^2$ 的 null space W_1, W_2. 得 W_1, W_2 之一組 basis 分別為 $\{(1,0,0,0,0)^t, (0,1,0,0,0)^t\}$, $\{(-1,0,0,0,1)^t, (1,0,1,0,0)^t, (-1,1,0,1,0)^t\}$.

由 $\dim(W_1) = 2$ 可知 W_1 本身是一個 cyclic space. 事實上取 $\mathbf{w}_1 = (1,0,0,0,0)^t$, 則 $A\mathbf{w}_1 = (2,1,0,0,0)^t$ (注意 $A^2\mathbf{w}_1 = -\mathbf{w}_1$). 即 $W_1 = C_{\mathbf{w}_1}$.

至於要將 W_2 分解成 cyclic subspaces 的 direct sum, 我們需先選出 $\mathbf{w}_2$ 滿足 $(A-I_5)\mathbf{w}_2 \neq (0,0,0,0,0)^t$. 事實上若選 $\mathbf{w}_2 = (1,0,1,0,0)^t$, 則 $A\mathbf{w}_2 = (1,1,2,1,0)^t$ (注意 $A^2\mathbf{w}_2 = 2A\mathbf{w}_2 - \mathbf{w}_2$), 所以 $\dim(C_{\mathbf{w}_2}) = 2$. 由於 $\dim(W_2) = 3$, 我們知道 W_2 應為 $C_{\mathbf{w}_2}$ 和另一個 dimension 為 1 的 cyclic subspace 的 direct sum. 此 cyclic subspace 應為 eigenvalue 為 1 的 eigenvector $\mathbf{w}_3$ 所形成, 而且 $\mathbf{w}_3 \not\in C_{\mathbf{w}_2}$. 我們選取 $\mathbf{w}_3 = (-1,0,0,0,1)^t$, 所以若令

$$P = \begin{pmatrix} 1 & 2 & 1 & 1 & -1 \\ 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 2 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}, \text{ 則 } P^{-1}AP = \begin{pmatrix} 0 & -1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 & 0 \\ 0 & 0 & 1 & 2 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

為 A 的 rational form.

事實上在 Example 4.4.11 中, 我們可以很快的判斷出 A 的 rational form. 因為 $\chi_A(x) = (x^2+1)(x-1)^3$ 所以 A^2+I_5 的 null space 僅由一個 cyclic subspace 所組成, 且其 cyclic vector 的 annihilator 為 x^2+1. 而 $\mu_A(x) = (x^2+1)(x-1)^2$ 所以由 Theorem 4.4.10 知, $(A-I_5)^2$ 的 null space 中一定有一個 cyclic subspace 其 cyclic vector 的 annihilator 為 $(x-1)^2$. 也因而我們知僅剩的 cyclic subspace 其 cyclic vector 的 annihilator 為 $x-1$. 所以 A 的 rational form 為有三個 blocks 的 diagonal matrix 其中每個 block 分別為 x^2+1, x^2-2x+1 以及 $x-1$ 的 companion matrix. 一般來說一個 matrix 的 rational form 並不能僅由其 characteristic polynomial 和 minimal polynomial 就能確定. 不過我們可以列出其所有可能的情形. 另一方面 rational form 是一個 canonical form, 也就是說兩個矩陣是 similar 的若且唯若它們可以化成同樣的 rational form. 我們會在下一節討論完 classical form 之後再探討這些課題.

4.5. Classical Form

當一個 linear operator 的 minimal polynomial 可以完全分解成一次多項式的乘積時, 除非它沒有重根 (即 diagonalizable), 此 linear operator 的 rational form 並不是 Jordan form. 此節中我們將說明如何另外選取 cyclic subspace 的一組 basis, 將其化成所謂的 classical form. 我們很容易看出 classical form 就是 Jordan form 的推廣.

當 $T:V\to V$ 是 F-linear, 給定 $\mathbf{v}\in V$, 考慮 T-cyclic subspace $C_{\mathbf{v}}$. 如果 $\mathbf{v}$ 的 T-annihilator 可以寫成 $\mu_{\mathbf{v}}(x)=p(x)^m$ (這裡 $p(x)\in F[x]$ 不需假設為 irreducible), 回顧一下若 $\deg(p(x))=d$, 則 $\{\mathbf{v},T(\mathbf{v}),T^{\circ 2}(\mathbf{v}),\dots,T^{\circ md-1}(\mathbf{v})\}$ 為 $C_{\mathbf{v}}$ 的一組 basis, 稱為 cyclic basis. 我們可以考慮以下一組新的 basis.

Lemma 4.5.1. *假設 $T:V\to V$ 是 F-linear, 給定 $\mathbf{v}\in V$. 若 $\mu_{\mathbf{v}}(x)=p(x)^m$, 其中 $p(x)\in F[x]$ 且 $\deg(p(x))=d$, 則*

$$\begin{array}{cccc} \mathbf{v} & T(\mathbf{v}) & \dots & T^{\circ d-1}(\mathbf{v}) \\ p(T)(\mathbf{v}) & p(T)(T(\mathbf{v})) & \dots & p(T)(T^{\circ d-1}(\mathbf{v})) \\ \vdots & \vdots & & \vdots \\ p^{m-1}(T)(\mathbf{v}) & p^{m-1}(T)(T(\mathbf{v})) & \dots & p^{m-1}(T)(T^{\circ d-1}(\mathbf{v})) \end{array} \tag{4.7}$$

是 $C_{\mathbf{v}}$ 的一組 basis.

Proof. 由 $\mu_{\mathbf{v}}(x)=p(x)^m$ 知 $\dim(C_{\mathbf{v}})=dm$. 因為 (4.7) 中共有 dm 個元素, 若能證明它們為 linearly independent over F, 則它們便是 $C_{\mathbf{v}}$ 的一組 basis.

對於 $0\le i\le m-1$, $0\le j\le d-1$, 若令 $h_{i,j}(x)=p^i(x)x^j$, 則 $p^i(T)(T^{\circ j})(\mathbf{v})=h_{i,j}(T)(\mathbf{v})$. 因為 $\deg(h_{i,j}(x))=di+j$, 我們知若 $(i,j)\neq(i',j')$, 則 $\deg(h_{i,j}(x))\neq\deg(h_{i',j'}(x))$. 換言之, 若 $c_{0,0},\dots,c_{i,j},\dots,c_{m-1,d-1}\in F$ 不全為 0, 則 $\sum_{i,j}c_{i,j}h_{i,j}(x)$ 是 $F[x]$ 中一個 nonzero polynomial.

現若存在一組不全為 0 的 $\{c_{i,j}\}$ 使得 $\sum_{i,j}c_{i,j}p^i(T)(T^{\circ j}(\mathbf{v}))=\mathbf{O}_V$, 表示 $h(x)=\sum_{i,j}c_{i,j}h_{i,j}(x)$ 這一個 nonzero polynomial 會滿足 $h(T)(\mathbf{v})=\mathbf{O}_V$. 很顯然 $\deg(h(x))<dm=\deg(\mu_{\mathbf{v}}(x))$, 這和 annihilator 的定義相矛盾, 故得證 (4.7) 中的元素為 linearly independent. $\square$

Question 4.22. *試利用 $T^{\circ d}(\mathbf{v})\in\operatorname{Span}(\mathbf{v},T(\mathbf{v}),\dots,T^{\circ d-1}(\mathbf{v}),p(T)(\mathbf{v}))$ 來證明 (4.7) 為 $C_{\mathbf{v}}$ 的一組 basis.*

Example 4.5.2. 考慮 $A=\begin{pmatrix}0&-1&0&1\\1&0&-1&0\\0&1&0&0\\0&0&-1&0\end{pmatrix}$, 計算得 $\mu_A(x)=(x^2+1)^2$. 因 A^2+I_4 的 null space 為 $N(A^2+I_4)=\operatorname{Span}((0,1,0,1)^{\mathrm{t}},(0,0,1,0)^{\mathrm{t}})$, 我們得 $\mathbb{R}^4=C_{\mathbf{v}}$, 其中 $\mathbf{v}\notin N(A^2+I_4)$, 且 $\mu_{\mathbf{v}}(x)=(x^2+1)^2$. 所以取 $\mathbf{v}=(1,0,0,0)^{\mathrm{t}}$, 則由 Theorem 4.4.4 知

$$\{\mathbf{v},A\mathbf{v},A^2\mathbf{v},A^3\mathbf{v}\}=\left\{\begin{pmatrix}1\\0\\0\\0\end{pmatrix},\begin{pmatrix}0\\1\\0\\0\end{pmatrix},\begin{pmatrix}-1\\0\\1\\0\end{pmatrix},\begin{pmatrix}0\\-2\\0\\-1\end{pmatrix}\right\}$$

為 $\mathbb{R}^4=C_{\mathbf{v}}$ 的一組 basis. 然而 Lemma 4.5.1 告訴我們

$$\{\mathbf{v},A\mathbf{v},(A^2+I_4)\mathbf{v},(A^3+A)\mathbf{v}\}=\left\{\begin{pmatrix}1\\0\\0\\0\end{pmatrix},\begin{pmatrix}0\\1\\0\\0\end{pmatrix},\begin{pmatrix}0\\0\\1\\0\end{pmatrix},\begin{pmatrix}0\\-1\\0\\-1\end{pmatrix}\right\}$$

亦為 $\mathbb{R}^4=C_{\mathbf{v}}$ 的一組 basis.

另外考慮 $B=\begin{pmatrix} 0 & 1 & -1 \\ -2 & 3 & -2 \\ -1 & 1 & 0 \end{pmatrix}$, 得 $\mu_B(x)=(x-1)^2$. 因 $\mathrm{Span}((1,1,0)^{\mathrm{t}},(-1,0,1)^{\mathrm{t}})$ 為 $B-I_3$ 的 null space, 若令 $\mathbf{w}=(1,0,0)^{\mathrm{t}}$, 我們有 $C_{\mathbf{w}}$ 的 cyclic basis 為

$$\{\mathbf{w},B\mathbf{w}\}=\left\{\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix},\begin{pmatrix} 0 \\ -2 \\ -1 \end{pmatrix}\right\}.$$

然而 Lemma 4.5.1 告訴我們

$$\{\mathbf{w},(B-I_3)\mathbf{w}\}=\left\{\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix},\begin{pmatrix} -1 \\ -2 \\ -1 \end{pmatrix}\right\}$$

亦為 $C_{\mathbf{w}}$ 的一組 basis.

接下來我們要探討若利用 (4.7) 這一組 basis, 則 $T|_{C_{\mathbf{v}}}$ 的 representative matrix 為何. 對於 $0\leq i\leq m-1$, $0\leq j\leq d-1$, 令 $\mathbf{v}_{id+j+1}=p^i(T)(T^{\circ j}(\mathbf{v}))$. 我們就是要考慮 $\beta=(\mathbf{v}_1,\ldots,\mathbf{v}_{md})$ 這一個 $C_{\mathbf{v}}$ 的 ordered basis. 假設 $p(x)=x^d+a_{d-1}x^{d-1}+\cdots+a_1x+a_0$, 當 $1\leq k\leq d-1$ 時, 我們有 $T(\mathbf{v}_k)=T(T^{\circ k-1}(\mathbf{v}))=T^{\circ k}(\mathbf{v})=\mathbf{v}_{k+1}$. 而

$$\begin{aligned} T(\mathbf{v}_d)=T(T^{\circ d-1}(\mathbf{v}))=T^{\circ d}(\mathbf{v}) &= p(T)(\mathbf{v})-a_{d-1}T^{\circ d-1}(\mathbf{v})-\cdots-a_1T(\mathbf{v})-a_0\mathbf{v} \\ &= -a_0\mathbf{v}_1-a_1\mathbf{v}_2-\cdots-a_{d-1}\mathbf{v}_d+\mathbf{v}_{d+1}. \end{aligned}$$

也就是說 $[T|_{C_{\mathbf{v}}}]_\beta$ 這一個 matrix 的前 d 個 column 分別為

$$\begin{pmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix},\begin{pmatrix} 0 \\ 0 \\ 1 \\ \vdots \\ 0 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix},\ldots,\begin{pmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix},\begin{pmatrix} -a_0 \\ -a_1 \\ -a_2 \\ \vdots \\ -a_{d-1} \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}.$$

也就是說前 d 個 column 所形成的 matrix 為

$$\begin{pmatrix} C_{p(x)} \\ U \\ \mathbf{O} \end{pmatrix}$$

這樣的形式, 其中 $C_{p(x)}$ 為 $d\times d$ 的 companion matrix of $p(x)$, 而 U 為 $d\times d$ 的 matrix 其在最右上角為 1 其他位置皆為 0. 最後的 $\mathbf{O}$ 是 $(m-2)d\times d$ 的 zero matrix. 同理當 $id+1\leq k=id+j+1\leq(i+1)d-1$, 我們有 $T(\mathbf{v}_k)=T(p^i(T)(T^{\circ j}(\mathbf{v}))=p^i(T)(T^{\circ j+1}(\mathbf{v}))=\mathbf{v}_{k+1}$.

而當 $k=(i+1)d$ 時

$$\begin{aligned}
T(\mathbf{v}_{(i+1)d}) &= T(p^i(T)(T^{\circ d-1}(\mathbf{v}))) = p^i(T)(T^{\circ d}(\mathbf{v}))\\
&= p^{i+1}(T)(\mathbf{v}) - a_{d-1}p^i(T)(T^{\circ d-1}(\mathbf{v})) - \cdots - a_1 p^i(T)(T(\mathbf{v})) - a_0 p^i(T)(\mathbf{v})\\
&= \begin{cases} -a_0\mathbf{v}_{id+1} - a_1\mathbf{v}_{id+2} - \cdots - a_{d-1}\mathbf{v}_{(i+1)d} + \mathbf{v}_{(i+1)d+1}, & \text{if } i+1<m;\\ -a_0\mathbf{v}_{md+1-d} - a_1\mathbf{v}_{md+2-d} - \cdots - a_{d-1}\mathbf{v}_{md}, & \text{if } i+1=m.\end{cases}
\end{aligned}$$

故得到

$$[T|_{C_\mathbf{v}}]_\beta = \begin{pmatrix} C_{p(x)} & & & & \\ U & C_{p(x)} & & \mathbf{O} & \\ & U & \ddots & & \\ \mathbf{O} & & \ddots & C_{p(x)} & \\ & & & U & C_{p(x)} \end{pmatrix}.$$

這個 $md\times md$ 矩陣稱為 the *classical matrix associated with* $p(x)^m$.

Example 4.5.3. 我們探討在 Example 4.5.2 中, 選取不同的 basis 所得的 similar matrices. 利用 $\mathbf{v}$ 所形成的 cyclic basis, 考慮 $P_1=\begin{pmatrix}1&0&-1&0\\0&1&0&-2\\0&0&1&0\\0&0&0&-1\end{pmatrix}$, 則 $P_1^{-1}AP_1=\begin{pmatrix}0&0&0&-1\\1&0&0&0\\0&1&0&-2\\0&0&1&0\end{pmatrix}$ 為 $\mu_\mathbf{v}(x)=(x^2+1)^2=x^4+2x^2+1$ 的 companion matrix. 而若考慮 $P_2=\begin{pmatrix}1&0&0&0\\0&1&0&-1\\0&0&1&0\\0&0&0&-1\end{pmatrix}$, 則 $P_2^{-1}AP_2=\begin{pmatrix}0&-1&0&0\\1&0&0&0\\0&1&0&-1\\0&0&1&0\end{pmatrix}$ 為 classical matrix associated with $\mu_\mathbf{v}(x)=(x^2+1)^2$. 注意 $\begin{pmatrix}0&-1\\1&0\end{pmatrix}$ 為 x^2+1 的 companion matrix.

關於矩陣 B, 由於 $\mathbf{u}=(1,1,0)^{\mathrm{t}}\in N(B-I_3)$ 且 $\mathbf{u}\notin C_\mathbf{w}$, 我們得 $\mathbb{R}^3=C_\mathbf{w}\oplus C_\mathbf{u}$. 現考慮 $Q_1=\begin{pmatrix}1&0&1\\0&-2&1\\0&-1&0\end{pmatrix}$, 我們有 $Q_1^{-1}BQ_1=\begin{pmatrix}0&-1&0\\1&2&0\\0&0&1\end{pmatrix}$, 為 B 的 rational form. 而若考慮 $Q_2=\begin{pmatrix}1&-1&1\\0&-2&1\\0&-1&0\end{pmatrix}$, 我們有 $Q_2^{-1}BQ_2=\begin{pmatrix}1&0&0\\1&1&0\\0&0&1\end{pmatrix}$, 為 B 的 Jordan form.

Question 4.23. 試說明 *the classical matrix associated with* $(x-\lambda)^m$ 就是 $m\times m$ 的 *elementary Jordan block associated with* λ.

對於一般的情形, 若一個 F-linear operator $T:V\to V$ 的 minimal polynomial $\mu_T(x)=p_1(x)^{m_1}\cdots p_k(x)^{m_k}$, 利用 cyclic decomposition theorem (Theorem 4.4.10),

$$V=C_{\mathbf{v}_{1,1}}\oplus\cdots C_{\mathbf{v}_{1,n_1}}\oplus\cdots\oplus C_{\mathbf{v}_{k,1}}\oplus\cdots\oplus C_{\mathbf{v}_{k,n_k}},$$

其中每一個 $\mathbf{v}_{i,j}$ 的 T-annihilator 為 $\mu_{\mathbf{v}_{i,j}}(x) = p_i(x)^{m_{i,j}}$ 滿足 $m_i = m_{i,1} \geq m_{i,2} \geq \cdots \geq m_{i,n_i}$. 此時若對每個 $C_{\mathbf{v}_{i,j}}$，我們選取如 Lemma 4.5.1 中 (4.7) 這樣一組 ordered basis, 然後組合成 V 的一組 ordered basis β, 則 $[T]_\beta$ 為以下形式的 block diagonal matrix

$$[T]_\beta = \begin{pmatrix} A_{1,1} & & & & & & \\ & \ddots & & & & & \\ & & A_{1,n_1} & & & \mathbf{O} & \\ & \mathbf{O} & & \ddots & & & \\ & & & & A_{k,1} & & \\ & & & & & \ddots & \\ & & & & & & A_{k,n_k} \end{pmatrix},$$

其中每一個 $A_{i,j}$ 為 classical matrix associated with $\mu_{\mathbf{v}_{i,j}}(x) = p_i(x)^{m_{i,j}}$. 這也告訴我們任何的方陣都會 similar to 這樣形式的方陣, 我們稱此為 *classical form*.

Example 4.5.4. 我們延續 Example 4.4.11, 探討 A 的 classical form. 這裡為了方便起見, 我們沿用 Example 4.4.11 的符號. 因 $\mu_{\mathbf{w}_1}(x) = x^2+1$, $\mu_{\mathbf{w}_2}(x) = (x-1)^2$ 以及 $\mu_{\mathbf{w}_3}(x) = x-1$, 考慮 $C_{\mathbf{w}_1}, C_{\mathbf{w}_2}$ 和 $C_{\mathbf{w}_3}$ 如 Lemma 4.5.1 中 (4.7) 這樣的 ordered basis $(\mathbf{w}_1, A\mathbf{w}_1)$, $(\mathbf{w}_2, (A-I_5)\mathbf{w}_2)$ 以及 $(\mathbf{w}_3)$, 所形成 $\mathbb{R}^5$ 的 order basis. 即若令

$$Q = \begin{pmatrix} 1 & 2 & 1 & 0 & -1 \\ 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}, \text{ 則 } Q^{-1}AQ = \begin{pmatrix} 0 & -1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

為 A 的 rational form.

Question 4.24. 考慮 *square matrix A*, 設 $\mu_A(x) = p_1(x)^{m_1} \cdots p_k(x)^{m_k}$, 其中 $p_i(x)$ 為相異的 *monic irreducible polynomial.* 若 $m_1 = m_2 = \cdots = m_k = 1$, 試說明 A 的 *classical form* 就是 *rational form.* 另一方面, 若 $\deg(p_1(x)) = \deg(p_2(x)) = \cdots = \deg(p_k(x)) = 1$, 試說明 A 的 *classical form* 就是 *Jordan form.*

在求得 rational form 以及 classical form 的過程中, 其實只要知道每一個 cyclic vector $\mathbf{v}_{i,j}$ 的 annihilator $p_i(x)^{m_{i,j}}$, 就可以確定其 rational form 及 classical form. 這一組 annihilators 相當的重要, 我們有以下的定義.

Definition 4.5.5. 設 $T: V \to V$ 為 F-linear operator, 且

$$V = C_{\mathbf{v}_{1,1}} \oplus \cdots C_{\mathbf{v}_{1,n_1}} \oplus \cdots \oplus C_{\mathbf{v}_{k,1}} \oplus \cdots \oplus C_{\mathbf{v}_{k,n_k}},$$

其中每一個 $\mathbf{v}_{i,j}$ 的 T-annihilator 為 $\mu_{\mathbf{v}_{i,j}}(x) = p_i(x)^{m_{i,j}}$ 滿足 $m_{i,1} \geq m_{i,2} \geq \cdots \geq m_{i,n_i}$. 我們稱

$$(p_1(x)^{m_{1,1}}, \ldots, p_1(x)^{m_{1,n_1}}, \ldots, p_k(x)^{m_{k,1}}, \ldots, p_k(x)^{m_{k,n_k}})$$

這一組 polynomials 為 T 的 *elementary divisors.*

例如在 Example 4.4.11 中 A 的 elementary divisors 就是 $(x^2+1,(x-1)^2,x-1)$. 要注意 elementary divisors 指的是所有的 $\mathbf{v}_{i,j}$ 的 T-annihilators, 所以即使有可能 $p_i(x)^{m_{i,j}}=p_i(x)^{m_{i,j'}}$, 也要將它們一一列出. 例如一個 linear operator $T:V\to V$ 的 cyclic decomposition 為 $V=C_{\mathbf{v}}\oplus C_{\mathbf{w}}\oplus C_{\mathbf{u}}$ 其中 $\mu_{\mathbf{v}}(x)=(x+1)^2,\mu_{\mathbf{w}}(x)=\mu_{\mathbf{u}}(x)=x+1$, 則 T 的 elementary divisors 為 $((x+1)^2,x+1,x+1)$.

基本上, 我們需要利用 $\ker(p_i(T)^{\circ t})$, $\forall t\in\mathbb{N}$ 來確定 T 的 elementary divisors. 不過我們可以由 $\chi_T(x)$ 和 $\mu_T(x)$ 得到 T 的 elementary divisors 的可能情況. 首先我們需要以下有關 elementary divisors 的性質.

Lemma 4.5.6. 設 $T:V\to V$ 為 F-*linear operator* 且

$$(p_1(x)^{m_{1,1}},\ldots,p_1(x)^{m_{1,n_1}},\ldots,p_k(x)^{m_{k,1}},\ldots,p_k(x)^{m_{k,n_k}})$$

為 T 的 *elementary divisors,* 其中 $m_{i,1}\geq m_{i,2}\geq\cdots\geq m_{i,n_i}$, $\forall i\in\{1,\ldots,k\}$. 則

$$\chi_T(x)=p_1(x)^{m_{1,1}}\cdots p_1(x)^{m_{1,n_1}}\cdots p_k(x)^{m_{k,1}}\cdots p_k(x)^{m_{k,n_k}},$$

$$\mu_T(x)=p_1(x)^{m_{1,1}}p_2(x)^{m_{2,1}}\cdots p_k(x)^{m_{k,1}}.$$

Proof. 由 elementary divisors 的定義知存在 $\mathbf{v}_{i,j}\in V$ 使得

$$V=C_{\mathbf{v}_{1,1}}\oplus\cdots C_{\mathbf{v}_{1,n_1}}\oplus\cdots\oplus C_{\mathbf{v}_{k,1}}\oplus\cdots\oplus C_{\mathbf{v}_{k,n_k}},$$

其中每一個 $\mathbf{v}_{i,j}$ 的 T-annihilator 為 $\mu_{\mathbf{v}_{i,j}}(x)=p_i(x)^{m_{i,j}}$. 由 Theorem 4.4.4, 我們有 $\chi_{T|_{C_{\mathbf{v}_{i,j}}}}(x)=\mu_{\mathbf{v}_{i,j}}(x)=p_i(x)^{m_{i,j}}$, 故由 Lemma 3.5.5 得

$$\chi_T(x)=\prod_{i,j}\chi_{T|_{C_{\mathbf{v}_{i,j}}}}(x)=\prod_{i,j}p_i(x)^{m_{i,j}}.$$

另外由 Theorem 4.4.10, 我們已知若 $\mu_T(x)=p_1(x)^{m_1}p_2(x)^{m_2}\cdots p_k(x)^{m_k}$, 則 $m_i=m_{i,1}$, $\forall i\in\{1,\ldots,k\}$. 故得證 $\mu_T(x)=p_1(x)^{m_{1,1}}p_2(x)^{m_{2,1}}\cdots p_k(x)^{m_{k,1}}$. □

我們利用以下的例子說明判斷 elementary divisors 的方法.

Example 4.5.7. 設 $T:\mathbb{R}^{10}\to\mathbb{R}^{10}$ 為 $\mathbb{R}$-linear operator 且 $\chi_T(x)=(x^2+1)^3(x-1)^4$ 以及 $\mu_T(x)=(x^2+1)^2(x-1)^2$. 首先我們知道 $(x^2+1)^2$ 以及 $(x-1)^2$ 一定會出現在 T 的 elementary divisors 中. 不過 $(x^2+1)^2$ 不會出現兩次. 這是因為在 $\chi_T(x)$ 中 (x^2+1) 有三次方, 所以由 Lemma 4.5.6 知僅還有一個 x^2+1 會出現. 另一方面可能還有一個 $(x-1)^2$ 會出現在 elementary divisor 中, 要不然就是有兩個 $x-1$ 會出現. 這是因為 $\chi_T(x)$ 中 $x-1$ 有四次方. 所以 T 的 elementary divisors 會有兩種可能, 一個是 $((x^2+1)^2,x^2+1,(x-1)^2,(x-1)^2)$. 而另一個是 $((x^2+1)^2,x^2+1,(x-1)^2,x-1,x-1))$.

至於 Example 4.5.7 中 T 的 elementary divisors 到底是哪種可能, 就不能完全由 $\chi_T(x)$ 和 $\mu_T(x)$ 來決定了. 此時我們可以考慮 $\dim(\mathrm{Ker}(T-\mathrm{id}))$. 若 $\dim(\mathrm{Ker}(T-\mathrm{id}))=2$ 表示 $\mathrm{Ker}((T-\mathrm{id})^{\circ 2})$ 可以寫成兩個 T-cyclic subspaces 的 direct sum, 在這種情形我們有 $((x^2+1)^2,x^2+1,(x-1)^2,(x-1)^2)$ 為 T 的 elementary divisors. 而若 $\dim(\mathrm{Ker}(T-\mathrm{id}))=3$ 表示 $\mathrm{Ker}((T-\mathrm{id})^{\circ 2})$ 可以寫成三個 T-cyclic subspaces 的 direct sum, 在這種情形我們有

$((x^2+1)^2, x^2+1, (x-1)^2, x-1, x-1)$ 為 T 的 elementary divisors. 至於一般的情形, 我們就必須探討每一個 $\mathrm{Ker}(p_i^j(T))$ 的維度. 首先我們有以下的性質.

Lemma 4.5.8. 設 $T: V \to V$ 為 *F-linear operator*, $\mathbf{v} \in V$ 其 *T-annihilator* 為 $p(x)^m$, 其中 $p(x) \in F[x]$ 為 *monic irreducible* 且 $\deg(p(x)) = d$. 對於 *monic irreducible polynomial* $q(x) \in F[x]$, 我們有

$$\dim(\mathrm{Ker}(q^l(T)|_{C_\mathbf{v}})) = \begin{cases} ld, & \text{if } q(x) = p(x) \text{ and } 1 \le l \le m-1; \\ md, & \text{if } q(x) = p(x) \text{ and } l \ge m; \\ 0, & \text{if } q(x) \ne p(x). \end{cases}$$

Proof. 首先考慮 $p(x) = q(x)$ 的情況. 對於 $0 \le i \le m-1$, $0 \le j \le d-1$, 令 $\mathbf{v}_{id+j+1} = p^i(T)(T^{\circ j}(\mathbf{v}))$. 由 Lemma 4.5.1, 我們知 $\beta = \{\mathbf{v}_1, \ldots, \mathbf{v}_{md}\}$ 為 $C_\mathbf{v}$ 的一組 basis. 即若 $\mathbf{w} \in C_\mathbf{v}$, 則存在 $c_1, \ldots, c_{md} \in F$ 使得 $\mathbf{w} = \sum_{k=1}^{md} c_k \mathbf{v}_k$. 當 $1 \le l \le m-1$ 時

$$p^l(T)(\mathbf{w}) = \sum_{i=0}^{m-1} \sum_{j=0}^{d-1} c_{id+j+1} p^l(p^i(T)(T^{\circ j}(\mathbf{v}))) = \sum_{i=0}^{m-l-1} \sum_{j=0}^{d-1} c_{id+j+1} p^l(p^i(T)(T^{\circ j}(\mathbf{v}))).$$

故若 $\mathbf{w} \in \mathrm{Ker}(p^l(T)|_{C_\mathbf{v}}) = \mathrm{Ker}(p^l(T)) \cap C_\mathbf{v}$, 則由 β 為 linearly independent 知 $c_1 = c_2 = \cdots = c_{(m-l)d} = 0$. 得 $\mathbf{w} \in \mathrm{Span}(\{\mathbf{v}_{(m-l)d+1}, \ldots, \mathbf{v}_{md}\})$. 很容易看出 $\mathrm{Span}(\{\mathbf{v}_{(m-l)d+1}, \ldots, \mathbf{v}_{md}\}) \subseteq \mathrm{Ker}(p^l(T)) \cap C_\mathbf{v}$, 故得證 $\dim(\mathrm{Ker}(p^l(T)|_{C_\mathbf{v}})) = md - (m-l)d = ld$.

當 $l \ge m$ 時, 因 $p^l(T)(\mathbf{w}) = \mathbf{O}_V$, for all $\mathbf{w} \in C_\mathbf{v}$, 得 $\mathrm{Ker}(p^l(T)|_{C_\mathbf{v}}) = C_\mathbf{v}$. 故

$$\dim(\mathrm{Ker}(p^l(T)|_{C_\mathbf{v}})) = \dim(C_\mathbf{v}) = md.$$

又若 $p(x) \ne q(x)$, 由 $p(x), q(x)$ 皆為 monic irreducible 知 $p(x)$ 與 $q(x)$ 為互質. 現若 $\mathbf{w} \in \mathrm{Ker}(q^l(T)|_{C_\mathbf{v}}) = \mathrm{Ker}(q^l(T)) \cap C_\mathbf{v}$, 由 $\mathbf{w} \in C_\mathbf{v}$ 知存在 $f(x) \in F[x]$ 使得 $\mathbf{w} = f(T)(\mathbf{v})$, 再由 $\mathbf{w} \in \mathrm{Ker}(q^l(T))$ 得 $\mathbf{O}_V = q^l(T)(\mathbf{w}) = q^l(f(T))(\mathbf{v})$. 因此由 $\mu_\mathbf{v}(x) = p(x)^m$, 得 $p(x)^m \mid q(x)^l f(x)$. 然而 $p(x)$ 與 $q(x)$ 為互質, 故得 $p(x)^m \mid f(x)$, 即 $\mathbf{w} = \mathbf{O}_V$. 得證 $\dim(\mathrm{Ker}(q^l(T)|_{C_\mathbf{v}})) = 0$. □

當 $T: V \to V$ 為 linear operator 且 $V = W_1 \oplus \cdots W_k$, 其中 W_i 為 T-invariant subspace, 則 $\mathrm{Ker}(T) = \mathrm{Ker}(T|_{W_1}) \oplus \cdots \oplus \mathrm{Ker}(T|_{W_k})$. 這是因為若 $\mathbf{v} = \mathbf{w}_1 + \cdots + \mathbf{w}_k \in \mathrm{Ker}(T)$, 其中 $\mathbf{w}_i \in W_i$, 則 $\mathbf{O}_V = T(\mathbf{v}) = T(\mathbf{w}_1) + \cdots + T(\mathbf{w}_k)$. 由於 $T(\mathbf{w}_i) \in W_i$ 且 $V = W_1 \oplus \cdots \oplus W_k$ 為 inner direct sum, 由 Proposition 3.4.6 (2) 知 $T(\mathbf{w}_1) = \cdots = T(\mathbf{w}_k) = \mathbf{O}_V$. 也就是說 $\mathbf{w}_i \in \mathrm{Ker}(T) \cap W_i = \mathrm{Ker}(T|_{W_i})$, $\forall i = 1, \ldots, k$. 現若

$$V = C_{\mathbf{v}_{1,1}} \oplus \cdots C_{\mathbf{v}_{1,n_1}} \oplus \cdots \oplus C_{\mathbf{v}_{k,1}} \oplus \cdots \oplus C_{\mathbf{v}_{k,n_k}},$$

其中每一個 $\mathbf{v}_{i,j}$ 的 T-annihilator 為 $p_i(x)^{m_{i,j}}$, 則因為每個 $C_{\mathbf{v}_{i',j}}$ 為 $p_i(T)$-invariant 再由 Lemma 4.5.8 得 $\dim(\mathrm{Ker}(p_i(T)|_{C_{\mathbf{v}_{i,j}}})) = \deg(p_i(x))$, 故得

$$\dim(\mathrm{Ker}(p_i(T))) = \sum_{j=1}^{n_i} \dim(\mathrm{Ker}(p_i(T)|_{C_{\mathbf{v}_{i,j}}})) = n_i \deg(p_i(x)).$$

換言之, $\dim(\mathrm{Ker}(p_i(T)))/\deg(p_i(x))$ 告訴我們 $\mathrm{Ker}(p_i^{m_i}(T)$ 可以寫成多少個 T-cyclic subgroup 的 direct sum. 同理若 $m_{i,j} \ge 2$, 則 $\dim(\mathrm{Ker}(p_i^2(T)|_{C_{\mathbf{v}_{i,j}}})) = 2\deg(p_i(x))$. 而若 $m_{i,j} \le 2$, 則

$\dim(\mathrm{Ker}(p_i^2(T)|_{C_{\mathrm{v}_{i,j}}})) = \deg(p_i(x))$. 因此

$$\dim(\mathrm{Ker}(p_i^2(T))) = \sum_{j=1}^{n_i} \dim(\mathrm{Ker}(p_i^2(T)|_{C_{\mathrm{v}_{i,j}}})) = 2(n_i - s_1)\deg(p_i(x)) + s_1 \deg(p_i(x)),$$

其中 $s_1 = \#\{1 \le j \le n_i \mid m_{i,j} = 1\}$. 也就是說我們可由

$$\dim(\mathrm{Ker}(p_i^2(T))) - \dim(\mathrm{Ker}(p_i(T))) = (n_i - s_1)\deg(p_i(x))$$

得知 T 的 elementary divisors 中有多少個為 $p_i(x)^t$ 其中 $t > 1$ 這種形式. 而且我們知 T 的 elementary divisors 中有

$$s_1 = (2\dim(\mathrm{Ker}(p_i(T))) - \dim(\mathrm{Ker}(p_i^2(T))))/\deg(p_i(x))$$

個為 $p_i(x)$. 依此類推, 若令 $s_t = \#\{1 \le j \le n_i \mid m_{i,j} = t\}$, 則當 $1 \le l \le m_i$ 時,

$$\dim(\mathrm{Ker}(p_i^l(T))) = (l(n_i - (s_1 + s_2 + \cdots + s_{l-1})) + s_1 + 2s_2 + \cdots + (l-1)s_{l-1})\deg(p_i(x)).$$

故由

$$\dim(\mathrm{Ker}(p_i^l(T)) - \dim(\mathrm{Ker}(p_i^{l-1}(T)))) = (n_i - (s_1 + s_2 + \cdots + s_{l-1}))\deg(p_i(x)) \tag{4.8}$$

我們可以將 $s_1, s_2, \ldots, s_{m_i}$ 一一求出.

Proposition 4.5.9. 設 $T: V \to V$ 為 *F-linear operator* 且 $\mu_T(x) = p_1(x)^{m_1} \cdots p_k(x)^{m_k}$, 其中 $p_i(x) \in F[x]$ 為相異的 *monic irreducible polynomial.* 對於 $i \in \{1, \ldots, k\}$, 當 $1 \le l \le m_i$ 時, T 的 *elementary divisors* 中 $p_i(x)^l$ 出現的次數為

$$\frac{1}{\deg(p_i(x))}\Big(2\dim(\mathrm{Ker}(p_i^l(T))) - \dim(\mathrm{Ker}(p_i^{l-1}(T))) - \dim(\mathrm{Ker}(p_i^{l+1}(T)))\Big).$$

Proof. 利用前面的符號, T 的 elementary divisors 中 $p_i(x)^l$ 出現的次數為 s_l. 由式子 (4.8) 我們知當 $1 \le l \le m_i - 1$ 時

$$\begin{aligned}&\dim(\mathrm{Ker}(p_i^l(T))) - \dim(\mathrm{Ker}(p_i^{l-1}(T))) - \Big(\dim(\mathrm{Ker}(p_i^{l+1}(T))) - \dim(\mathrm{Ker}(p_i^l(T)))\Big)\\ &= (n_i - (s_1 + s_2 + \cdots + s_{l-1}) - (n_i - (s_1 + s_2 + \cdots + s_l)))\deg(p_i(x)) = s_l \deg(p_i(x)).\end{aligned}$$

另外當 $l = m_i$ 時 $\mathrm{Ker}(p_i^{m_i}(T)) = \mathrm{Ker}(p_i^{m_i+1}(T))$ 所以由式子 (4.8) 知

$$\begin{aligned}&2\dim(\mathrm{Ker}(p_i^{m_i}(T))) - \dim(\mathrm{Ker}(p_i^{m_i-1}(T))) - \dim(\mathrm{Ker}(p_i^{m_i+1}(T)))\\ =\ &\dim(\mathrm{Ker}(p_i^{m_i}(T)) - \dim(\mathrm{Ker}(p_i^{m_i-1}(T))))\\ =\ &(n_i - (s_1 + s_2 + \cdots + s_{m_i-1}))\deg(p_i(x)) = s_{m_i}\deg(p_i(x)),\end{aligned}$$

得證本定理. □

由 Proposition 4.5.9, 我們得知 T 的 elementary divisors 完全由 $\mathrm{Ker}(p_i^l(T))$ 來決定, 這和 V 的 basis 選取無關. 也就是說不管選取怎樣的 cyclic basis, 我們都會得到相同的 elementary divisors. 所以都可以化成相同的 rational form 和 classical form. 也就是說 rational form 和 classical form 都是 canonical form. 我們有以下之結論.

Theorem 4.5.10. 設 A, B 為 *$n \times n$ matrices.* 則 A 和 B 為 *similar* 若且唯若 A 和 B 可以化成相同的 *rational form* 也若且唯若 A 和 B 可以化成相同的 *classical form.*

在 Theorem 4.3.9 我們知道當 $A \in M_n(F)$ 且 $\chi_A(x)$ 可以在 $F[x]$ 中完全分解成一次的 monic polynomials 的乘積, 則 A 的 transpose A^{t} 和 A 為 similar. 當時我們也提到這個定理在一般的狀況也是對的, 現在我們可以證明這個更一般的結果.

Theorem 4.5.11. 設 A 為 $n \times n$ *matrix,* 則 A 的 *transpose* A^{t} 和 A 為 *similar.*

Proof. 因 $\mu_A(x) = \mu_{A^{\mathrm{t}}}(x)$. 若 $\mu_A(x) = p_1(x)^{m_1} \cdots p_k(x)^{m_k}$, 我們僅要討論 $\forall i \in \{1, \ldots, k\}$ 對於 $1 \leq l \leq m_i$, $p_i(x)^l$ 出現在 A 的 elementary divisors 的次數等於出現在 A^{t} 的 elementary divisors 的次數. 這表示 A 和 A^{t} 有相同的 elementary divisors, 所以他們為 similar. 然而 $p_i(x)^l$ 出現在 A 的 elementary divisors 的次數依 Proposition 4.5.9 知由 $\dim(\mathrm{Ker}(p_i^{l-1}(A))), \dim(\mathrm{Ker}(p_i^{l}(A)))$ 以及 $\dim(\mathrm{Ker}(p_i^{l+1}(A)))$ 所決定. 而對於任意 $j \in \mathbb{N}$, 我們有

$$\dim(\mathrm{Ker}(p_i^j(A))) = \dim(\mathrm{Ker}((p_i^j(A))^{\mathrm{t}})) = \dim(\mathrm{Ker}(p_i^j(A^{\mathrm{t}}))).$$

故得證本定理. □

以前我們曾提到若 $A, B \in M_n(F)$ 且存在一個比 F 大的 field $\tilde{F}$ 使得在 $M_n(\tilde{F})$ 中 $A \sim B$ (即存在 $\tilde{P} \in M_n(\tilde{F})$ invertible 使得 $B = \tilde{P}^{-1} \cdot A \cdot \tilde{P}$), 則在 $M_n(F)$ 中 $A \sim B$ (即存在 $P \in M_n(F)$ invertible 使得 $B = P^{-1} \cdot A \cdot P$). 這個事實是因為由 A, B 看成 $M_n(\tilde{F})$ 的 matrices 時它們的 elementary divisors 相同可以推得 A, B 看成 $M_n(F)$ 的 matrices 時它們的 elementary divisors 也相同. 證明的細節, 就留給大家當成習題了.