現若 λ 是 linear operator $T: V \to V$ 的一個 eigenvalue, 我們考慮 $W = E_T(\lambda)$ 這個 T-invariant subspace，及 $T|_W : E_T(\lambda) \to E_T(\lambda)$ 的 characteristic polynomial. 假設 $\dim(E_A(\lambda)) = d$, 即 λ 的 geometric multiplicity 為 d. 令 $\beta = (\mathbf{v}_1, \ldots, \mathbf{v}_d)$ 為 $E_A(\lambda)$ 的一組 ordered basis，則 $[T|_W]_\beta = \lambda I_d$. 因此知 $T|_W$ 的 characteristic polynomial $g(t)$ 為 $(\lambda - t)^d$。而若 T 本身 λ 的 algebraic multiplicity 為 m，表示 T 的 characteristic polynomial $p_T(t)$ 中 $(t-\lambda)^m$ 恰整除 $p_T(t)$（即 m 為 $t-\lambda$ 可以整除 $p_T(t)$ 的最高次數），故由 Proposition 7.4.5 知 $(\lambda - t)^d$ 整除 $(\lambda - t)^m$，推得 $d \leq m$。這告訴我們 λ 的幾何重根數小於或等於其代數重根數。由於一個方陣可以視為一個 linear operator，所以這對方陣也是對的。我們有以下的結果：

Proposition 7.4.6. 對於一個 *linear operator* T 若 $\lambda \in \mathbb{F}$ 為 T 的一個 *eigenvalue* 且其 *geometric multiplicity* 為 d 以及 *algebraic multiplicity* 為 m, 則 $d \leq m$.

同理，對於方陣 $A \in \mathsf{M}_{n\times n}(\mathbb{F})$，若 $\lambda \in \mathbb{F}$ 為 A 的一個 *eigenvalue* 且其 *geometric multiplicity* 為 d 以及 *algebraic multiplicity* 為 m, 則 $d \leq m$.

要注意：會有幾何重根數小於代數重根數的情形發生，例如當 $A = \begin{bmatrix} 3 & 1 \\ 0 & 3 \end{bmatrix}$ 時，3 是其 eigenvalue，很容易由其 characteristic polynomial $P_A(t) = (3-t)^2$ 知其代數重根數為 2；但直接計算可得其幾何重根數為 1。所以一般來說，不能由代數重根數馬上決定其幾何重根數，必須依定義計算其 eigenspace 的維度。不過在一個特別的情況是可以由代數重根數確定幾何重根數的。由於 A 的 eigenvalue λ 的 geometric multiplicity 必大於 0（因對應 λ 的 eigenvector 必存在）且其值必小於或等於其 algebraic multiplicity（Proposition 7.4.6）. 因此當 λ 是 A 的 characteristic polynomial 的單根（即 λ 的 algebraic multiplicity 為 1）, 其 geometric multiplicity 一定等於 1. 這是唯一可以確定兩個重根數會一樣的情況。

7.5. Cayley-Hamilton Theorem

在這節中我們將介紹 Cayley-Hamilton Theorem. 這個定理有方陣的版本以及 linear operator 的版本。當然了，證明了其中任一版本，另一版本就自然成立。不過由於兩個版本的證明能讓我們更了解利用矩陣處理以及用 linear operator 處理其特點，所以這裡兩種方法我們都大略談論一下（不嚴格證明細節）。

Cayley-Hamilton Theorem，涉及將方陣或 linear operator 代入多項式，我們回顧一下其定義。考慮 over $\mathbb{F}$ 的多項式 $f(t) = c_m t^m + \cdots + c_1 t + c_0$, 其中 $c_i \in \mathbb{F}$。對任意 $A \in \mathsf{M}_{n\times n}(\mathbb{F})$, 我們定義 $f(A) = c_m A^m + \cdots + c_1 A + c_0 I_n$. 依此定義 $f(A)$ 也會是 n 階方陣。而對於 linear operator $T: V \to V$，我們也定義 $f(T) = c_m T^m + \cdots + c_1 T + c_0 \mathrm{id}_V$. 依此定義 $f(T)$ 也會是 V 到 V 的 linear operator。這樣代入多項式的方式會和多項式的乘法相吻合，也就說若 $f(t) = h(t)g(t)$，則 $f(A) = h(A)g(A)$ 且 $f(T) = h(T) \circ g(T)$。注意矩陣相乘與 linear operator 合成之間的關係。

Cayley-Hamilton Theorem 說的是：當 $A \in \mathsf{M}_{n\times n}(\mathbb{F})$ 且其 characteristic polynomial 為 $p_A(t) = \det(A - tI)$，則 $p_A(A)$ 為 n 階零方陣。相對的，當 $T : V \to V$ 為 linear operator 且其 characteristic polynomial 為 $p_T(t)$，則 $p_T(T)$ 為零函數。我們先處理矩陣的情形。

依定義 $p_A(t) = \det(A - tI_n)$, 常見錯誤的論證 Cayley-Hamilton Theorem 是直接將 A 代入 $\det(A - tI_n)$ 的 t 中，以至於誤以為 $p_A(A) = \det(A - AI_n) = \det(O) = 0$. 這個論證是不對的，左邊的 $p_A(A)$ 是一個 n 階方陣；而右邊的 $\det(A - AI_n)$ 是一個常數。錯誤之處在於當初定義 characteristic polynomial $\det(A - tI_n)$ 時, 是將 t 看待成 $\mathbb{F}$ 中的未知元素, 才能利用它解出在 $\mathbb{F}$ 中的 eigenvalue. 因此 $\det(A - tI_n)$ 可寫成 $\det\begin{bmatrix} a_{11}-t & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn}-t \end{bmatrix}$, 再將之展開得 t 的多項式. 因此在這一步驟當 t 代任何 $\mathbb{F}$ 中的元素進入上述的矩陣是行得通的, 但是 t 不能帶入矩陣 A. 因為 $\det\begin{bmatrix} a_{11}-A & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn}-A \end{bmatrix}$ 是個未定義的奇怪東西.

由於 A 的 characteristic polynomial 牽涉到 $A - tI_n$ 這樣的矩陣, 也就是說矩陣的 entry 中有 t 的多項式. 我們可以將一般在 entry 中有 t 的多項式的這類矩陣也同樣寫成 $t^dA_d + \cdots + tA_1 + A_0$, 其中 $A_i \in M_n(F)$ 這樣的型式. 例如我們可以有以下的表示法

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

由於我們是將 $\mathbb{F}$ 的元素代入 t, 所以 t^kA_k 就是常數 t^k 乘上矩陣 A_k. 這樣的寫法完全符合矩陣運算，不會有前面錯誤論述之虞。我們要強調若 $t^dA_d + \cdots + tA_1 + A_0 = t^dB_d + \cdots + tB_1 + B_0$, 其中 $A_i, B_i \in M_n(F)$, 則 $A_i = B_i$, $\forall i = 0, 1, \ldots, d$. 這是因為若有某個 $A_i \neq B_i$, 表示等式兩邊的矩陣有個 entry 其 t^i 的係數不相同, 造成矛盾. 了解了這些概念, 我們就可以處理 Cayley-Hamilton Theorem.

Cayley-Hamilton Theorem 談論的是 characteristic polynomial 和矩陣的運算。前者和行列式有關，因此我們必須將行列式的計算與矩陣運算相連結。矩陣的 adjugate（classical adjoint）提供了這方面的連結（參見 Theorem 5.5.12），也就是

$$\operatorname{adj}(A - tI_n) \cdot (A - tI_n) = \det(A - tI_n)I_n = p_A(t)I_n \qquad (7.2)$$

看到式子 (7.2) 或許會誤以為等式兩邊將 A 代入 t 後都是方陣，應該就可推論 $p_A(A)$ 為零矩陣了。事實不然，因為 tI_n 與任意方陣相乘皆可交換，但一般方陣 A 並無此性質。例如我們有 $(B - tI_n)(B + tI_n) = B^2 - t^2I_n$，但將 A 代入兩邊得 $(B - A)(B + A) = B^2 - A^2$ 未必正確（除非 $AB = BA$）。所以我們需摒除代入這樣的錯誤方式，改由前面所示：若 $t^dA_d + \cdots + tA_1 + A_0 = t^dB_d + \cdots + tB_1 + B_0$, 則 $A_i = B_i$, $\forall i = 0, 1, \ldots, d$；這樣的方式處理。

為了讓說明較易理解，接下來的論述我們僅討論 A 為 3 階方陣的情形，不難看出這個論述適用於一般的 n 階方陣。由於 $\operatorname{adj}(A - tI_3)$ 的每個 entry 皆是 $A - tI_3$ 的 minor 的行列式。而 $A - tI_3$ 的 minor 頂多 t 出現兩次，故知 $\operatorname{adj}(A - tI_3)$ 的每個 entry 最多只出現 t 的二次式。也因此我們可以將 $\operatorname{adj}(A - tI_3)$ 寫成 $t^2B_2 + tB_1 + B_0$，其中 B_2, B_1, B_0 皆為 3 階方陣。

若 A 的 characteristic polynomial 為 $p_A(t) = -t^3 + c_2t^2 + c_1t + c_0$，則由式子 (7.2) 我們有

$$(t^2B_2 + tB_1 + B_0)(A - tI_3) = t^3(-I_3) + t^2(c_2I_3) + t(c_1I_3) + c_0I_3. \tag{7.3}$$

將等式 (7.3) 左邊展開, 我們得

$$t^3(-B_2 \cdot I_3) + t^2(B_2 \cdot A - B_1 \cdot I_3) + t(B_1 \cdot A - B_0 \cdot I_3) + B_0 \cdot A$$

應該和等式 (7.3) 右邊相同, 故比較係數得

$$\begin{array}{rcl} B_0 \cdot A &=& c_0I_3 \\ B_1 \cdot A - B_0 \cdot I_3 &=& c_1I_3 \\ B_2 \cdot A - B_1 \cdot I_3 &=& c_2I_n \\ -B_2 \cdot I_3 &=& -I_3 \end{array}$$

將第一式不動, 第二式兩邊右乘 A, 第三式兩邊右乘 A^2, ..., 最後一式兩邊右乘 A^3, 我們得

$$\begin{array}{rcl} B_0 \cdot A &=& c_0I_n \\ B_1 \cdot A^2 - B_0 \cdot A &=& c_1A \\ B_2 \cdot A^3 - B_1 \cdot A^2 &=& c_2A^2 \\ -B_2 \cdot A^3 &=& -A^3 \end{array}$$

因為左邊全部加起來會等於右邊全部加起來, 得證 $O = -A^3 + c_2A^2 + \cdots + c_1A + c_0I_n = p_A(A)$. 我們將矩陣形式的 Cayley-Hamilton Theorem 敘述如下：

Theorem 7.5.1 (Cayley-Hamilton Theorem)**.** 若 $A \in \mathsf{M}_{n\times n}(\mathbb{F})$ 且 $p_A(t)$ 為 A 的 *characteristic polynomial*，則 $p_A(A)$ 為零矩陣。

論證 linear operator $T : V \to V$ 的 Cayley-Hamilton Theorem，我們的方法是證明對任意 $\mathbf{v} \in V$，皆會滿足 $p_T(T)(\mathbf{v}) = \mathbf{0}$，因此得證 $p_T(T)$ 為零函數。這裡的處理方式是針對每一個 $\mathbf{v} \in V$ 考慮包含 $\mathbf{v}$ 最小的 T-invariant subspace。假設 W 是一個包含 $\mathbf{v}$ 的 T-invariant subspace。由 $\mathbf{v} \in W$，我們自然要有 $T(\mathbf{v}) \in W$ 。再由 $T(\mathbf{v}) \in W$ 以及 W 是 T-invariant, 我們有 $T(T(\mathbf{v})) = T^2(\mathbf{v}) \in W$. 如此一直下去我們知 $T^m(\mathbf{v}) \in W, \forall m \in \mathbb{N}$. 由此我們知, 若 W 是包含 $\mathbf{v}$ 的 T-invariant subspace, 則 W 必須包含 $\{\mathbf{v}, T(\mathbf{v}), T^2(\mathbf{v}), \ldots, T^m(\mathbf{v}), \ldots\}$ 這個集合 (即 $\{T^i(\mathbf{v}) \mid i \in \mathbb{N}\}$) 中所有的元素. 不過 W 是 subspace, 所以也必須包含所有這些元素所展成的 subspace, 所以我們有以下的定義。

Definition 7.5.2. 假設 $T : V \to V$ 是 linear operator. 對任意 $\mathbf{v} \in V$, 令

$$C(T, \mathbf{v}) = \mathrm{Span}\{\mathbf{v}, T(\mathbf{v}), T^2(\mathbf{v}), \ldots, T^m(\mathbf{v}), \ldots\}.$$

我們稱 $C(T, \mathbf{v})$ 為 the T-cyclic space generated by $\mathbf{v}$.

要注意若令 $S = \{\mathbf{v}, T(\mathbf{v}), T^2(\mathbf{v}), \ldots, T^m(\mathbf{v}), \ldots\}$, 雖然 S 中可能有無窮多個元素, 不過依 span 的定義, $C(T, \mathbf{v}) = \mathrm{Span}(S)$ 中的元素是 S 中有限多個元素的線性組合. 因此若 $\mathbf{w} \in C(T, \mathbf{v}) = \mathrm{Span}(S)$, 表示存在 $c_0, c_1, \ldots, c_m \in \mathbb{F}$ 使得 $\mathbf{w} = c_0\mathbf{v} + c_1T(\mathbf{v}) + \cdots + c_mT^m(\mathbf{v})$ (其

中可能有些 $c_i=0$). 故得 $T(\mathbf{w})=c_0T(\mathbf{v})+c_1T^2(\mathbf{v})+\cdots+c_mT^{m+1}(\mathbf{v})\in \mathrm{Span}(S)=C(T,\mathbf{v})$. 也因此得證 $c(T,\mathbf{v})$ 是 T-invariant subspace. 前面提過包含 $\mathbf{v}$ 的 T-invariant subspace 必包含 S, 故我們得到下面的定理.

Proposition 7.5.3. 假設 $T:V\to V$ 是 *linear operator* 且 $\mathbf{v}\in V$. 則 $C(T,\mathbf{v})$ 是包含 $\mathbf{v}$ 最小的 T*-invariant subspace.*

Question 7.7. 假設 V 是 *vector space over* $\mathbb{F}$, $T:V\to V$ 是 *linear operator* 且 $\mathbf{v}\in V$. 證明對任意 $\mathbf{w}\in C(T,\mathbf{v})$, 皆存在係數在 $\mathbb{F}$ 的多項式 $f(x)$ 使得 $\mathbf{w}=f(T)(\mathbf{v})$. 依此得 $C(T,\mathbf{v})=\{f(T)(\mathbf{v})\mid f(x)\in\mathbb{F}[x]\}$.

當 V 是 finite dimensional vector space over $\mathbb{F}$, 因為 $C(T,\mathbf{v})$ 為其 subspace, 所以 $C(T,\mathbf{v})$ 也是 finite dimensional. 如何知道 $C(T,\mathbf{v})$ 的維度呢? 當然了, 若 $\mathbf{v}=\mathbf{0}$, 則 $T^i(\mathbf{v})=T^i(\mathbf{0})=\mathbf{0}$, $\forall i\in\mathbb{N}$, 故此時 $C(T,\mathbf{v})=\{\mathbf{0}\}$, 即 $\dim(C(T,\mathbf{v}))=0$. 因此我們僅考慮 $\mathbf{v}\neq\mathbf{0}$ 的情況. 首先我們考慮 $\mathbf{v},T(\mathbf{v})$ 是否為 linearly independent. 若 $\mathbf{v},T(\mathbf{v})$ 不是 independent, 由於 $\mathbf{v}\neq\mathbf{0}$, 故知存在 $c\in\mathbb{F}$ 使得 $T(\mathbf{v})=c\mathbf{v}$. 由此知 $T^i(\mathbf{v})=c^i\mathbf{v}\in\mathrm{Span}(\mathbf{v})$, $\forall i\in\mathbb{N}$. 因此得 $C(T,\mathbf{v})=\mathrm{Span}(\mathbf{v})$, 即 $\dim(C(T,\mathbf{v}))=1$. 而若 $\mathbf{v},T(\mathbf{v})$ 為 independent, 則我們考慮 $\mathbf{v},T(\mathbf{v}),T^2(\mathbf{v})$ 是否為 independent. 若它們不是 independent, 則由 $\mathbf{v},T(\mathbf{v})$ 為 independent 知 $T^2(\mathbf{v})\in\mathrm{Span}(\mathbf{v},T(\mathbf{v}))$ (Lemma 3.5.4). 因此存在 $c,d\in\mathbb{F}$ 使得 $T^2(\mathbf{v})=c\mathbf{v}+dT(\mathbf{v})$. 此時

$$T^3(\mathbf{v})=T(T^2(\mathbf{v}))=cT(\mathbf{v})+dT^2(\mathbf{v})=cT(\mathbf{v})+d(c\mathbf{v}+dT(\mathbf{v}))=dc\mathbf{v}+(c+d^2)T(\mathbf{v}).$$

因此得 $T^3(\mathbf{v})\in\mathrm{Span}(\mathbf{v},T(\mathbf{v}))$. 再利用數學歸納法, 我們可以證明 $T^i(\mathbf{v})\in\mathrm{Span}(\mathbf{v},T(\mathbf{v}))$, $\forall i\in\mathbb{N}$, 因此知此時 $C(T,\mathbf{v})=\mathrm{Span}(\mathbf{v},T(\mathbf{v}))$, 即 $\dim(C(T,\mathbf{v}))=2$. 我們可以一直這樣探討下去，利用數學歸納法得到以下的定理.

Proposition 7.5.4. 假設 V 是 *vector space over* $\mathbb{F}$, $T:V\to V$ 是 *linear operator* 且 $\mathbf{v}\in V$. 則 $\dim(C(T,\mathbf{v}))=m$ 若且唯若 m 是最大的 i 使得 $\mathbf{v},T(\mathbf{v}),\ldots,T^{i-1}(\mathbf{v})$ 為 *linear independent:* 也就是說 $\mathbf{v},T(\mathbf{v}),\ldots,T^{m-1}(\mathbf{v})$ 是 *linear independent* 但 $\mathbf{v},T(\mathbf{v}),\ldots,T^{m-1}(\mathbf{v}),T^m(\mathbf{v})$ 不是 *linearly independent.*

事實上若 $\dim(C(T,\mathbf{v}))=m$, 則 $\mathbf{v},T(\mathbf{v}),\ldots,T^{m-1}(\mathbf{v})$ 是 $C(T,\mathbf{v})$ 的一組 *basis.*

Question 7.8. 證明 $\dim(C(T,\mathbf{v}))=1$ 若且唯若 $\mathbf{v}$ 是 T 的 *eigenvector.*

既然 $C(T,\mathbf{v})$ 是 T-invariant, 我們知 $T|_{C(T,\mathbf{v})}:C(T,\mathbf{v})\to C(T,\mathbf{v})$ 是 linear operator. 那麼 $T|_{C(T,\mathbf{v})}$ 的 characteristic polynomial 會是甚麼呢? 要求 $T|_{C(T,\mathbf{v})}$ 的 characteristic polynomial, 我們要先找到 $T|_{C(T,\mathbf{v})}$ 的定義域 $C(T,\mathbf{v})$ 的一組 ordered basis, 在利用這組 ordered basis 得到 $T|_{C(T,\mathbf{v})}$ 的表現矩陣, 再求該矩陣的 characteristic polynomial. 根據 Proposition 7.5.4, 若 $\dim(C(T,\mathbf{v}))=m$, 我們很自然的選 $(\mathbf{v},T(\mathbf{v}),\ldots,T^{m-1}(\mathbf{v}))$ 這一組 $C(T,\mathbf{v})$ 的 ordered basis.

為了說明較簡明，這裡僅討論 $\dim(C(T,\mathbf{v}))=3$ 的情況，大家應該清楚這裡的論述對一般 $\dim(C(T,\mathbf{v}))=m$ 皆適用。因 $\dim(C(T,\mathbf{v}))=3$，故 $\beta=(\mathbf{v},T(\mathbf{v}),T^2(\mathbf{v}))$ 是 $C(T,\mathbf{v})$ ordered basis，且設 $T^3(\mathbf{v})=c_0\mathbf{v}+c_1T(\mathbf{v})+c_2T^2(\mathbf{v})$。接著我們來看 $T|_{C(T,\mathbf{v})}$ 用 β 這一組 ordered basis 其表現矩陣 $[T|_{C(T,\mathbf{v})}]_\beta$ 為何？首先 $[T|_{C(T,\mathbf{v})}]_\beta$ 的 1-st column 是 β 的第一個向量 (即 $\mathbf{v}$) 經由 T 映射後所得的向量 (即 $T(\mathbf{v})$) 用 ordered basis β 所得的坐標表示. 由於 $T(\mathbf{v})$ 恰好是 β 的第二個向量, 故其坐標表示為 $\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$. 同理得 $[T|_{C(T,\mathbf{v})}]_\beta$ 的 2-nd column 為 **c**001. 至於 $[T|_{C(T,\mathbf{v})}]_\beta$ 的最後一個 column, 應該是 β 的最後一個向量 (即 $T^2(\mathbf{v})$) 經由 T 映射後所得的向量 (即 $T^3(\mathbf{v})$) 用 ordered basis β 所得的坐標表示. 故由 $T^3(\mathbf{v}) = c_0\mathbf{v} + c_1T(\mathbf{v}) + c_2T^2(\mathbf{v})$ 的假設知 $[T|_{C(T,\mathbf{v})}]_\beta$ 的最後一個 column 就是 $\begin{bmatrix} c_0 \\ c_1 \\ c_2 \end{bmatrix}$. 因此得 $[T|_{C(T,\mathbf{v})}]_\beta = \begin{bmatrix} 0 & 0 & c_0 \\ 1 & 0 & c_1 \\ 0 & 1 & c_2 \end{bmatrix}$. 如何求這樣的矩陣的 characteristic polynomial $g(t)$ 呢? 因 $[T|_{C(T,\mathbf{v})}]_\beta - tI_3 = \begin{bmatrix} -t & 0 & c_0 \\ 1 & -t & c_1 \\ 0 & 1 & c_2 - t \end{bmatrix}$, 要求其 determinant, 由於我們想用數學歸納法處理一般情形, 所以不直接計算而是採用降階的方式處理. 對其 1-st row 降階求 determinant 得

$$g(t) = (-t)\det\begin{bmatrix} -t & c_1 \\ 1 & c_2 - t \end{bmatrix} + c_0 \det\begin{bmatrix} 1 & -t \\ 0 & 1 \end{bmatrix}.$$

其中第一個矩陣可得其 determinant 為 $t^2 - c_2t - c_1$, 而第二個矩陣式上三角矩陣故其 determinant 為 1. 因此可得

$$g(t) = (-t)(t^2 - c_2t - c_1) + c_0 = -(t^3 - c_2t^2 - c_1t - c_0).$$

在一般情況，利用數學歸納法我們可以得到以下之結果：

Proposition 7.5.5. 假設 V 是 *vector space over* $\mathbb{F}$, $T : V \to V$ 是 *linear operator* 且 $\mathbf{v} \in V$. 已知 $\dim(C(T,\mathbf{v})) = m$ 且 $T^m(\mathbf{v}) = c_0\mathbf{v} + c_1T(\mathbf{v}) + \cdots + c_{m-1}T^{m-1}(\mathbf{v})$。若考慮 $C(T,\mathbf{v})$ 的 *ordered basis* $\beta = (\mathbf{v}_1, T(\mathbf{v}), \ldots, T^{m-1}(\mathbf{v}))$，則

$$[T|_{C(T,\mathbf{v})}]_\beta = \begin{bmatrix} 0 & 0 & \cdots & 0 & c_0 \\ 1 & 0 & \cdots & 0 & c_1 \\ \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & \cdots & 0 & c_{m-2} \\ 0 & 0 & \cdots & 1 & c_{m-1} \end{bmatrix}$$

且 $T|_{C(T,\mathbf{v})}$ 的 *characteristic polynomial* 為

$$(-1)^m(t^m - c_{m-1}t^{m-1} - \cdots - c_1t - c_0).$$

利用 Proposition 7.5.5, 我們馬上可得以下結論.

Corollary 7.5.6. 假設 $T : V \to V$ 為 *linear operator* 且 $\mathbf{v} \in V$ 為非零向量. 令 $g(x)$ 為 $T|_{C(T,\mathbf{v})} : C(T,\mathbf{v}) \to C(T,\mathbf{v})$ 的 *characteristic polynomial.* 則 $g(T)(\mathbf{v}) = \mathbf{0}$.

Proof. 假設 $\dim(C(T,\mathbf{v})) = m$ 且 $T^m(\mathbf{v}) = c_0\mathbf{v} + c_1T(\mathbf{v}) + \cdots + c_{m-1}T^{m-1}(\mathbf{v})$. Proposition 7.5.5 告訴我們 $T|_{C(T,\mathbf{v})}$ 的 characteristic polynomial 為 $g(x) = (-1)^m(x^m - c_{m-1}x^{m-1} - \cdots - c_1x - c_0)$.

因此

$$\begin{aligned} g(T)(\mathbf{v}) &= (-1)^m(T^m - c_{m-1}T^{m-1} - \cdots - c_1T - c_0\mathrm{id}_{C(T,\mathbf{v})})(\mathbf{v}) \\ &= (-1)^m(T^m(\mathbf{v}) - c_{m-1}T^{m-1}(\mathbf{v}) - \cdots - c_1T(\mathbf{v}) - c_0\mathbf{v}) = \mathbf{0}. \end{aligned}$$

□

Question 7.9. 假設 $T : V \to V$ 為 *linear operator* 且 $\mathbf{v} \in V$ 為非零向量. 令 $g(x)$ 為 $T|_{C(T,\mathbf{v})}$ 的 *characteristic polynomial.* 證明 $g(T)|_{C(T,\mathbf{v})} = \mathbf{0}$.

接著利用 T 和 $T|_{C(T,\mathbf{v})}$ 的 characteristic polynomial 之間的關係，我們就可以證明 Cayley-Hamilton Theorem.

Theorem 7.5.7 (Cayley-Hamilton Theorem)**.** 假設 V 為 *vector space over* $\mathbb{F}$ 且 $T : V \to V$ 為 *linear operator.* 考慮 T 的 *characteristic polynomial* $p_T(x)$, 我們有 $p_T(T) : V \to V$ 為 *zero operator,* 亦即對任意 $\mathbf{v} \in V$, $p_T(T)(\mathbf{v}) = \mathbf{0}$.

Proof. 由於當 $\mathbf{v} = \mathbf{0}$, 時自然有 $p_T(T)(\mathbf{v}) = \mathbf{0}$ (因 $p_T(T)$ 是 linear transformation), 故僅考慮 $\mathbf{v} \in V$ 且 $\mathbf{v} \neq O$. 此時考慮 $\mathbf{v}$ 所產生的 T-cyclic space $C(T,\mathbf{v})$, 並令 $g(t)$ 為 $T|_{C(T,\mathbf{v})}$ 的 characteristic polynomial. 由 Proposition 7.4.5 知存在 $h(t)$ 係數在 $\mathbb{F}$ 的 polynomial 使得 $p_T(t) = h(t)g(t)$. 因此知 $p_T(T) = h(T) \circ g(T)$. 此時 $p_T(T)(\mathbf{v}) = (h(T) \circ g(T))(\mathbf{v}) = h(T)(g(T)(\mathbf{v}))$. 然而由 Corollary 7.5.6 知 $g(T)(\mathbf{v}) = \mathbf{0}$, 故得證 $p_T(T)(\mathbf{v}) = h(T)(g(T)(\mathbf{v})) = h(T)(\mathbf{0}) = \mathbf{0}$. □

最後我們看一個有關 Cayley-Hamilton Theorem 的應用。當 $A \in \mathsf{M}_{n\times n}(\mathbb{F})$，要計算 A^k 我們可以利用 Theorem 7.5.1 以及長除法將其次數降下來, 這樣也可很快算出 A^k, 我們看下面的 Example.

Example 7.5.8. 考慮 $A = \begin{bmatrix} 1 & 2 \\ -2 & 1 \end{bmatrix}$. 由於 $p_A(x) = x^2 - 2x + 5$ 故由 Cayley-Hamilton Theorem, 我們知 $A^2 - 2A + 5I_2 = O$. 我們可以利用此式計算 A 的高次方. 例如計算 A^5. 首先利用長除法, 將 x^5 除以 $p_A(x) = x^2 - 2x + 5$, 可得 $x^5 = (x^3 + 2x^2 - x - 12)(x^2 - 2x + 5) - 19x + 60$. 兩邊代入 A 得 $A^5 = (A^3 + 2A^2 - A - 12I_2)(A^2 - 2A + 5I_2) - 19A + 60I_2$. 由於 $A^2 - 2A + 5I_2 = O$, 故得 $A^5 = -19A + 60I_2 = \begin{bmatrix} 41 & -38 \\ 38 & 41 \end{bmatrix}$. ♯

Exercise 7.9. 假設 V 為 over $\mathbb{F}$ 的 vector space 且 $T: V \to V$ 為 linear operator. 已知 $\mathbf{v}_1, \mathbf{v}_2$ 為 linearly independent 且滿足 $T(\mathbf{v}_1) = a\mathbf{v}_1$, $T(\mathbf{v}_2) = b\mathbf{v}_1 + c\mathbf{v}_2$，其中 $a, b, c \in \mathbb{F}$. 令 $W = \mathrm{Span}(\mathbf{v}_1, \mathbf{v}_2)$. 試依是否 $a = c$ 以及 b 的取值討論 W 中所有可能 T 的 eigenvectors 及其對應的 eigenvalues.

Exercise 7.10. 考慮所有 3×3 且只有一個 eigenvalue $\lambda = 3$ 的實矩陣。請在下表空格中有符合條件的矩陣打 ✓，沒有符合的矩陣打 ✗。打 ✓的情形需舉例，打 ✗ 的情形需說明。其中 am 表示代數重根數、gm 表示幾何重根數。

$\lambda = 3$	$\mathrm{am} = 0$	$\mathrm{am} = 1$	$\mathrm{am} = 2$	$\mathrm{am} = 3$	$\mathrm{am} \geq 4$
$\mathrm{gm} = 0$					
$\mathrm{gm} = 1$					
$\mathrm{gm} = 2$					
$\mathrm{gm} = 3$					
$\mathrm{gm} \geq 4$					

Exercise 7.11. 假設 $T: V \to V$ 為 linear operator, $\mathbf{v} \in V$ 且 $g(x)$ 為非零多項式。令 W 為 $\mathbf{v}$ 所生成的 T-cyclic space 且令 N 為 linear operator $g(T)$ 的 nullspace（or kernel）.

(1) 假設 $\dim(W) = k$. 證明對任意 $\mathbf{w} \in W$, 皆存在唯一的一個次數小於 k 的多項式 $f(x)$ 使得 $\mathbf{w} = f(T)(\mathbf{v})$.

(2) 證明 N 為 T-invariant subspace.

(3) 證明 $\mathbf{v} \in N$ 若且唯若 $W \subseteq N$.

Exercise 7.12. 試求以下 linear operator T 限制在給定的 T-cyclic space $C(T, \mathbf{v})$，即 $T|_{C(T,\mathbf{v})}$ 的 characteristic polynomial.

- $T: \mathbb{R}^4 \to \mathbb{R}^4$ 定義為 $T(a,b,c,d) = (a+b, b-c, a+c, a+d)$；$\mathbf{v} = (1,0,0,0)$.
- $T: \mathrm{M}_2(\mathbb{R}) \to \mathrm{M}_2(\mathbb{R})$ 定義為 $T(A) = \begin{bmatrix} 1 & 1 \\ 2 & 2 \end{bmatrix} A$；$\mathbf{v} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$

Exercise 7.13. 考慮矩陣 $A = \begin{bmatrix} 1 & 2 & 1 \\ 0 & 2 & 3 \\ 0 & 0 & -1 \end{bmatrix}$.

(1) 試將 $\mathrm{adj}(A - tI_3)$ 寫成 $t^2 B_2 + tB_1 + B_0$, 其中 $B_i \in \mathrm{M}_3(\mathbb{R})$。

(2) 承 (1) 試展開 $(t^2B_2 + tB_1 + B_0)(A - tI_3)$ 寫成 $t^3C_3 + t^2C_2 + tC_1 + C_0$，其中 $C_i \in \mathrm{M}_3(\mathbb{R})$。

(3) 利用 Cayley-Hamilton Theorem 將 A^5 寫成 A^2, A, I_3 的線性組合, 並以此寫出 A^5.

(4) 利用 Cayley-Hamilton Theorem 將 A^{-1} 寫成 A^2, A, I_3 的線性組合, 並以此寫出 A^{-1}.