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.

Contents

Chapter 1

Vector Spaces

在本章我們要介紹 linear algebra 所要探討的主要對象 “Vector Space”. 接著我們將探討與 vector space 息息相關的 basis 以及 dimension.

1.1. Definition and Basic Properties

組成 vector space 的元素, 並不一定要是我們很熟悉的向量 (vector). 不過它的元素間需要有如向量一樣的運算性質. 一般向量中有所謂的加法, 所以我們也用 “+” 來表示一個 vector space 中元素的運算. 也就是說要有一個 vector space 首先要有的是一個非空的集合 V (也就是說 V 裡面一定要有元素), 再來元素間要有一個運算 “+”. 這個運算必須有封閉性, 亦即對所有 $\mathbf{v},\mathbf{w}\in V$ 皆使得 $\mathbf{v}+\mathbf{w}\in V$. 另外這個運算有我們熟悉的以下性質

VS1: 對任意 $\mathbf{u},\mathbf{v}\in V$, 皆有 $\mathbf{u}+\mathbf{v}=\mathbf{v}+\mathbf{u}$.

VS2: 對任意 $\mathbf{u},\mathbf{v},\mathbf{w}\in V$, 皆有 $(\mathbf{u}+\mathbf{v})+\mathbf{w}=\mathbf{u}+(\mathbf{v}+\mathbf{w})$.

VS3: 存在一元素 $\mathbf{O}\in V$ 滿足對任意 $\mathbf{u}\in V$ 皆有 $\mathbf{O}+\mathbf{u}=\mathbf{u}$.

VS4: 對任意 $\mathbf{u}\in V$ 皆可找到 $\mathbf{u}'\in V$ 滿足 $\mathbf{u}+\mathbf{u}'=\mathbf{O}$.

大家可以看出來, V 在此 + 的運算下, 形成一個 abelian group. 不過 vector space 不只是一個 group, 它的元素還可以和一個 filed 的元素 “作用” (action), 這是所謂的 *scaler multiplication*. 這個作用也必須有封閉性. 也就是說對於一個 vector space V, 還必須有一個 field F, 且對任意的 $r\in F$ 以及 $\mathbf{v}\in V$, r 和 $\mathbf{v}$ 作用之下的元素, (在此我們記作 $r\mathbf{v}$), 仍必須在 V 中.

當然了這個作用和 + 之間仍需保有一定的關係才有意義, 它們之間有我們熟悉的以下性質

VS5: 對任意 $r,s\in F$ 以及 $\mathbf{u}\in V$, 皆有 $r(s\mathbf{u})=(rs)\mathbf{u}$.

VS6: 對任意 $r,s\in F$ 以及 $\mathbf{u}\in V$, 皆有 $(r+s)\mathbf{u}=r\mathbf{u}+s\mathbf{u}$.

VS7: 對任意 $r\in F$ 以及 $\mathbf{u},\mathbf{v}\in V$ 皆有 $r(\mathbf{u}+\mathbf{v})=r\mathbf{u}+r\mathbf{v}$.

VS8: 對任意 $\mathbf{u} \in V$, 皆有 $1\mathbf{u} = \mathbf{u}$.

為什麼要有這些運算性質呢? 有了以上 VS1 $\sim$ VS8 這 8 個性質, 我們就能對一個 vector space 中的元素像處理數字一樣來作運算. 注意這 8 個性質缺一不可, 另一方面它們又足以讓我們推導出許多其他的性質. 等一下我們會看一些例子, 在此我們先對一些符號加以說明.

首先我們會用粗體字來表示一個向量空間 V 中的元素, 而用細字表示 field F 中的元素. 例如 V 中的加法單位元素我們用 $\mathbf{O}$ 來表示, 而 F 中的加法單位元素用 0 來表示. 另外 V 和 F 中都有加法, 一般來說這兩個加法是不一樣的 (除非 $V = F$), 不過我們都用 $+$ 來表示. 這是因為除非 $V = F$, 要不然 V 中的元素是不可以和 F 中的元素相加的, 所以不至於造成混淆. 最後我們還要再強調, 要形成一個 vector space 一定要有一個 abelian group V 及一個 filed F. 這中間不只含有 V, F 本身的代數結構也牽涉到 V, F 之間作用的關係. 在具體例子裡要給一個 vector space 都必須明確說明這些運算關係. 不過在一般抽象的情形我們都會直接說 V 是一個 over F 的 vector space. 有時更會省略稱 V 為 F-space.

Example 1.1.1. 我們介紹一些常見的 vector space. 以下例子中 F 為一個 filed.

(1) 令 $F^n = \{(a_1,\ldots,a_n) \mid a_i \in F\}$. 定義加法為: 若 $(a_1,\ldots,a_n),(b_1,\ldots,b_n) \in F^n$, 則 $(a_1,\ldots,a_n)+(b_1,\ldots,b_n) = (a_1+b_1,\ldots,a_n+b_n)$. 依此定義很容易檢驗 F^n 在此加法之下是封閉的且 VS1 $\sim$ VS4 是符合的, 其中 $\mathbf{O} = (0,\ldots,0)$. 令一方面我們定義 F 和 F^n 的作用為: 若 $r \in F$, $(a_1,\ldots,a_n) \in F^n$, 則 $r(a_1,\ldots,a_n) = (a_1,\ldots,a_n)$. 也很容易看出此作用是封閉的且符合 VS5 $\sim$ VS8. 在這運算之下我們得 F^n 是一個 over F 的 vector space. 特別的是 F 本身是 over F 的 vector space. 我們也可將此推廣到 $M_{m\times n}(F)$ 是所有 entry 在 F 的 $m \times n$ 矩陣所成的集合. 利用類似上述運算方法 (即一般矩陣的運算方法), 我們得 $M_{m\times n}(F)$ 是一個 over F 的 vector space.

(2) 考慮所有係數在 F 且次數為 n 的多項式, 我們定義加法為: 若 $f(x) = a_nx^n + \cdots + a_0, g(x) = b_nx^n + \cdots + b_0$, 則 $f(x)+g(x) = (a_n+b_n)x^n + \cdots + (a_0+b_0)$. 首先要注意所有係數在 F 且次數為 n 的多項式在此加法的定義之下並不是封閉的. (兩個次數相同的多項式相加有可能次數變小). 不過若我們考慮所有次數小於或等於 n 的多項式, 則在此加法之下就是封閉的了, 而且它們符合 VS1 $\sim$ VS4. 若 $r \in F$, $f(x) = a_nx^n + \cdots + a_0$, 令 $rf(x) = ra_nx^n + \cdots + ra_0$, 則在此作用之下我們得所有係數在 F 且次數小於或等於 n 的多項式是一個 vector space over F. 我們用 $P_n(F)$ 來表示這一個 vector space. 零多項式是 $P_n(F)$ 中的 $\mathbf{O}$ (加法單位元素).

(3) 給定一集合 S. 考慮所有從 S 映射到 F 的函數所成的集合 F^S. 若 $r \in F$, $f,g \in F^S$, 我們定義 $f+g, rf \in F^S$ 為 $f+g : s \mapsto f(s)+g(s)$ 且 $rf : s \mapsto rf(s)$. 依此所定的運算我們可得 F^S 是一個 over F 的 vector space 且 $f \in F^S$ 滿足 $f(s) = 0, \forall s \in S$ 為 F^S 中的 $\mathbf{O}$.

Question 1.1. 你能看出來 *Example 1.1.1* 中 *(1)* 是 *(3)* 的一個特例嗎? *(1)* 和 *(2)* 有關連嗎?

Question 1.2. 若 V 是一個 *vector space over* F, V' 是 V 的一個 *subgroup* 且 F' 是 F 的一個 *subfield.* 利用 V, F 的運算, 是否 V' 是一個 *over* F 的 *vector space?* 是否 V 是一個 *over* F' 的 *vector space?*

接下來我們要談論 vector space 的一些基本性質. 若 V 是一個 F-space, 首先 V 本身由 abelian group 的結構所得的性質, 在這裡我們就略而不深談. 不過我們要特別提醒: VS3 中所提存在 $\mathbf{O}\in V$ 使得對所有 $\mathbf{v}\in V$ 皆有 $\mathbf{v}+\mathbf{O}=\mathbf{v}$. 這裡的 $\mathbf{O}$ 其實是唯一的, 就是所謂的加法單位元素. 另外 VS4 所提對所有 $\mathbf{v}\in V$ 皆存在 $\mathbf{v}'\in V$ 使得 $\mathbf{v}+\mathbf{v}'=\mathbf{O}$, 這裡的 $\mathbf{v}'$ 也會隨著 $\mathbf{v}$ 而唯一確定, 就是所謂的加法反元素, 在此情形之下我們用 $-\mathbf{v}$ 來表示 $\mathbf{v}'$. 最後我們要強調利用加法反元素存在, 對於 V 中的一個元素 $\mathbf{w}$, 只要存在一個 $\mathbf{v}\in V$ 使得 $\mathbf{v}+\mathbf{w}=\mathbf{v}$, 那麼我們便可以確認 $\mathbf{w}=\mathbf{O}$.

至於 VS5 $\sim$ VS8, 這些有關 V 與 F 相互間的作用關係, 可幫我們得到以下性質.

Proposition 1.1.2. 假設 V 為一個 *over* F 的 *vector space,* 我們有以下之性質*:*

(1) 設 $r\in F$, $\mathbf{v}\in V$, 則 $r\mathbf{v}=\mathbf{O}$ 若且唯若 $r=0$ 或 $\mathbf{v}=\mathbf{O}$.

(2) 對任意 $\mathbf{v}\in V$, $\mathbf{v}+(-1)\mathbf{v}=\mathbf{O}$. 換言之, $-\mathbf{v}=(-1)\mathbf{v}$.

Proof. (1) ($\Leftarrow$) 若 $r=0$, 要證明 $r\mathbf{v}=\mathbf{O}$, 由上面所討論的結果, 我們僅要證明 $0\mathbf{v}+\mathbf{v}=\mathbf{v}$ 即可. 然而由 VS6 及 VS8 知

$$0\mathbf{v}+\mathbf{v}=0\mathbf{v}+1\mathbf{v}=(0+1)\mathbf{v}=1\mathbf{v}=\mathbf{v}.$$

故得證. 另一方面, 若 $\mathbf{v}=\mathbf{O}$, 要證明 $r\mathbf{v}=\mathbf{O}$ 我們僅要證明 $r\mathbf{O}+r\mathbf{O}=r\mathbf{O}$ 即可. 然而由 VS3 及 VS7 知

$$r\mathbf{O}+r\mathbf{O}=r(\mathbf{O}+\mathbf{O})=r\mathbf{O}.$$

故得證.

($\Rightarrow$) 若 $r\mathbf{v}=\mathbf{O}$ 且 $r\neq 0$, 則由 F 是一個 field 知存在 $r^{-1}\in F$ 使得 $r^{-1}r=1$. 故由 VS5, VS8 及前面證得結果知

$$\mathbf{v}=1\mathbf{v}=r^{-1}(r\mathbf{v})=r^{-1}\mathbf{O}=\mathbf{O}.$$

故得證.

(2) 利用 VS8, VS6 及 (1) 所證得之結果, 可知

$$\mathbf{v}+(-1)\mathbf{v}=1\mathbf{v}+(-1)\mathbf{v}=(1-1)\mathbf{v}=0\mathbf{v}=\mathbf{O}.$$

故由加法反元素的唯一性得 $-\mathbf{v}=(-1)\mathbf{v}$.

□

Question 1.3. 設 V 是一個 F*-space.* 利用 *Proposition 1.1.2* 你能證明以下性質嗎*?*

(1) 若 $r,r'\in F$, $\mathbf{u},\mathbf{v}\in V$ 且 $r\neq r'$, $\mathbf{u}\neq\mathbf{O}$, 則 $r\mathbf{u}+\mathbf{v}\neq r'\mathbf{u}+\mathbf{v}$.

(2) 若 $r\in F$, $\mathbf{v}\in V$, 則 $-(r\mathbf{v})=(-r)\mathbf{v}=r(-\mathbf{v})$.

1.2. Subspaces

若 V 是一個 over F 的 vector space, U 是 V 的一個非空子集 (nonempty subset), 且在原先 V, F 的運算之下, U 是一個 over F 的 vector space, 則稱 U 是 V 的一個 *subspace*. 有時我們會用 U 是 V 的一個 F-subspace 這樣的說法來強調是 over F 的 subspace. 另外為了方便我們用 $U \leq V$ 來表示 U 是 V 的一個 subspace.

在 V 中任選一個非空的 subset S 會不會就是 V 的一個 subspace 呢? 答案顯然是不一定會, 因為要成為 subspace, S 本身一定要是一個 abelian group, 也就是說 S 需為 V 的 subgroup. 不過即使 S 是 V 的 subgroup, 在前面 Question 1.2 中我們也知道 S 仍不一定是一個 F-space. 問題出現在 S 和 F 作用不一定有封閉性. 事實上只要 S 在 V 的運算之下是封閉的且和 F 作用是封閉的就可以確認 S 是 V 的 subspace. 我們有以下之結果

Proposition 1.2.1. 設 V 是一個 *over* F 的 *vector space* 且 S 為 V 的一個 *subset.* 則 S 是 V 的一個 F*-subspace* 若且唯若 S 有以下之性質:

(1) $\mathbf{O} \in S$.

(2) 對於所有 $r,s \in F, \mathbf{u},\mathbf{v} \in S$ 皆有 $r\mathbf{u}+s\mathbf{v} \in S$.

Proof. ($\Rightarrow$) 若 S 是 V 的 subspace, 因 S 是非空, 故存在 $\mathbf{v}$ 在 S 中. 由 subspace 的定義知 S 是一個 F-space, 故由 S, F 作用的封閉性及 Proposition 1.1.2 (1) 得 $0\mathbf{v} = \mathbf{O} \in S$. 另一方面若 $r,s \in F, \mathbf{u},\mathbf{v} \in S$, 由 S, F 作用的封閉性得 $r\mathbf{u}, s\mathbf{v} \in S$, 再由 S 加法的封閉性得 $r\mathbf{u}+s\mathbf{v} \in S$.

($\Leftarrow$) 由 (1) $\mathbf{O} \in S$ 知 S 是一個非空集合. 要證明 S 是一個 F-space, 我們需驗證 S 加法封閉性及 S, F 作用的封閉性, 再驗證 VS1 $\sim$ VS8 成立. 首先說明對於 $\mathbf{u},\mathbf{v} \in S$ 因為 $S \subseteq V$, 故 $\mathbf{u},\mathbf{v} \in V$ 再由 V 是 F-space, 得 $1\mathbf{u} = \mathbf{u}, 1\mathbf{v} = \mathbf{v}$. 令 $r = 1, s = 1$ 由 (2) 得 $\mathbf{u}+\mathbf{v} = 1\mathbf{u}+1\mathbf{v} \in S$, 即得證 S 加法封閉性. 另外對任意 $r \in F, \mathbf{u} \in S$, 因 $\mathbf{O} \in S$, 令 $s = 1, \mathbf{v} = \mathbf{O}$ 由 (2) 得 $r\mathbf{u}+1\mathbf{O} \in S$. 但 $\mathbf{u},\mathbf{O}$ 皆為 V 中元素, 故由 V 為 F-space 之假設得 $r\mathbf{u}+1\mathbf{O} = r\mathbf{u}$, 即得證 S, F 作用的封閉性. 最後我們要驗證 VS1 $\sim$ VS8 對於 S 皆成立. 由於 $S \subseteq V$, VS1, VS2 以及 VS5 $\sim$ VS8 對於所有 V 中元素都成立, 自然對 S 中元素亦成立. 然而 VS3, 為 (1) 之假設. VS4 由前 S, F 作用的封閉性以及 Proposition 1.1.2 知對任意 $\mathbf{v} \in S$, $-\mathbf{v} = (-1)\mathbf{v} \in S$. 得證 S 為 V 的一個 F-subspace. $\square$

Question 1.4. 若 $\mathbf{u} \in S$ 令 $r = 1, s = -1$ 以及 $\mathbf{v} = \mathbf{u}$, 則利用 *Proposition 1.2.1* 的 *(2)* 可得 $\mathbf{O} = 1\mathbf{u}+(-1)\mathbf{u} \in S$. 為什麼還需要 *(1)* 的假設呢?

Question 1.5. 設 V 是一個 F*-space,* 且 S 是 V 的一個非空子集. 試證明若 S 利用 V 的加法運算是封閉的且利用 V, F 的作用也是封閉的, 則 S 是 V 的一個 F*-subspace.* 一般來講一個 *abelian group* 它的非空子集若在原運算之下是封閉的並不一定會是這個 *abelian group* 的 *subgroup.* 然而在 *vector space* 前述的情形為何 S 會是 V 的 *subgroup* 呢?

利用 Proposition 1.2.1, 我們很容易檢驗一個 vector space 的非空子集是否為其 subspace. 通常要檢查一個集合及其運算是否為 vector space, 要檢查 VS1 $\sim$ VS8 這些項目, 不過若已知它是包含於某個 vector space, 那麼所需檢查的項目就簡單多了.

Example 1.2.2. 以下我們舉出幾個在 Example 1.1.1 中的 vector space 它們的 subspace.

(1) 給定 $c_1,\ldots,c_n \in F$. $E = \{(a_1,\ldots,a_n) \in F^n \mid c_1a_1 + \cdots + c_na_n = 0\}$, 是 F^n 的一個 subspace. E 可用 $c_1x_1 + \cdots + c_nx_n = 0$ 來表示. 在幾何中通常在 F^n 中滿足 $c_1x_1 + \cdots + c_nx_n = b$ 的元素所成的集合稱為 F^n 的 hyperplane. 不過僅有 $b = 0$ 時的 hyperplane 是一個 F-space.

(2) $P_n(F)$ 中考慮次數小於等於 $n-1$ 的多項式所成的集合, 即 $P_{n-1}(F)$, 是一個 subspace. 另外給定 $\lambda \in F$, 集合 $\Lambda = \{f(x) \in P_n(F) \mid f(\lambda) = 0\}$ 也是 $P_n(F)$ 的 subspace. 特別在 $P_n(F)$ 中常數項為 0 的多項式所成的集合 (即當 $\lambda = 0$) 是一個 subspace.

(3) 給定 $T \subseteq S$, F^S 中的子集合 $\{f \in F^S \mid f(t) = 0, \forall t \in T\}$ 是一個 subspace.

給定一個 over F 的 vector space V, 很容易得知 V 和 $\{\mathbf{O}\}$ 皆為其 subspace. 這兩個 subspace 稱為 V 的 trivial subspace. 若 V 中有其他的 F-subspace, 我們有興趣是否能利用這些 subspace 得到更多的 F-subspace. 首先有以下之結果.

Proposition 1.2.3. 若 V 是一個 *vector space over* F 且 U,W 為 V 的 *subspace,* 則 $U \cap W$ 也是 V 的 *subspace.*

Proof. 我們利用 Proposition 1.2.1 來證明. 首先因為 U,W 皆為 V 的 subspace, 我們有 $\mathbf{O} \in U$ 且 $\mathbf{O} \in W$, 故得 $\mathbf{O} \in U \cap W$. 另外若 $r,s \in F$ 且 $\mathbf{u},\mathbf{v} \in U \cap W$, 則由 $\mathbf{u},\mathbf{v} \in U$ 得 $r\mathbf{u} + s\mathbf{v} \in U$. 同理 $r\mathbf{u} + s\mathbf{v} \in W$ 故知 $r\mathbf{u} + s\mathbf{v} \in U \cap W$. □

Question 1.6. V 是一個 *vector space over* F. 設 I 為一個 *index set,* 且對於任意 $i \in I$, V_i 皆為 V 的 *subspace.* 是否 $\bigcap_{i\in I} V_i$ 也是 V 的 *subspace?*

我們很自然會問是否 V,W 為 V 的 F-subspace, 也會使得 $U \cup W$ 亦為 V 的 F-subspace. 一般來說這是不對的. 例如在 $\mathbb{R}^2$ 中 $L_1 = \{(r,0) \mid r \in \mathbb{R}\}$, $L_2 = \{(0,s) \mid s \in \mathbb{R}\}$ 皆為 $\mathbb{R}^2$ 的 subspace. 因為 $(1,0) \in L_1$, $(0,1) \in L_2$ 所以 $(1,0),(0,1) \in L_1 \cup L_2$, 但是 $(1,1) = (1,0) + (0,1) \notin L_1 \cup L_2$, 故知 $L_1 \cup L_2$ 不是 $\mathbb{R}^2$ 的 subspace. 事實上當 F 是一個 infinite field 時, 我們有以下之結果.

Theorem 1.2.4. 設 F 是一個 *infinite field* 且 V 是一個 F*-space.* 若 $V_1,\ldots,V_n$ 為 V 的 *nontrivial* F*-subspaces,* 則 $V_1 \cup \cdots \cup V_n \neq V$.

Proof. 我們利用反證法, 假設 $V = V_1 \cup \cdots \cup V_n$. 若 $V_1 \subseteq V_2 \cup \cdots \cup V_n$, 則可將 V_1 拿掉仍得 $V = V_2 \cup \cdots \cup V_n$. 故不失一般性我們可假設 $V_1 \not\subseteq V_2 \cup \cdots \cup V_n$. 因此存在 $\mathbf{u} \in V_1 \setminus V_2 \cup \cdots \cup V_n$ (即 $\mathbf{u} \in V_1$ 但 $\mathbf{u} \notin V_2 \cup \cdots \cup V_n$). 又 $V_1 \subsetneq V$, 故存在 $\mathbf{v} \in V \setminus V_1$. 考慮集合

$$S = \{r\mathbf{u} + \mathbf{v} \mid r \in F\}.$$

注意若 $r \neq r'$ 則 $r\mathbf{u} + \mathbf{v} \neq r'\mathbf{u} + \mathbf{v}$ (否則會導致 $(r - r')\mathbf{u} = \mathbf{O}$, 而得到 $\mathbf{u} = \mathbf{O}$ 之矛盾). 因此由 F 是 infinite field 知 S 有無窮多個元素.

接下來我們看看 S 和每一個 V_i 個別會有多少個共同元素. 若 $r\mathbf{u}+\mathbf{v}\in V_1$, 因 $r\mathbf{u}\in V_1$, 由 V_1 是 F-subspace 可得 $\mathbf{v}\in V_1$ 和當初 $\mathbf{v}$ 的選取矛盾, 故知 $S\cap V_1=\emptyset$. 另一方面, 當 $2\leq i\leq n$, 若存在 $r\neq r'$ 使得 $r\mathbf{u}+\mathbf{v}\in V_i$ 且 $r'\mathbf{u}+\mathbf{v}\in V_i$, 則由 V_i 是 F-subspace 得 $(r-r')\mathbf{u}\in V_i$, 再推得 $\mathbf{u}\in V_i\subseteq V_2\cup\cdots\cup V_n$. 此與 $\mathbf{u}$ 的選取相矛盾, 故知 S 和每個 V_i 最多僅有一個共同元素. 也就是說集合 $S\cap(V_1\cup\cdots\cup V_n)$ 最多僅有 $n-1$ 個元素. 但因 V 是一個 F-space, 由 $\mathbf{u},\mathbf{v}\in V$ 可得 $S\subseteq V$. 換言之, $V=V_1\cup\cdots\cup V_n$ 的假設告訴我們 $S\cap(V_1\cup\cdots\cup V_n)=S\cap V=S$ 應有無窮多個元素. 此與剛才結論相矛盾故知 $V=V_1\cup\cdots\cup V_n$ 不可能成立. □

要注意 Theorem 1.2.4 若 F 是一個 finite field, 則不一定成立. 另外由此定理可知若 F 是一個 infinite field, 則一個 over F 的 vector space 中兩個無包含關係的 subspaces 的聯集一並不會是一個 F-space (參見以下 Questions).

Question 1.7. 當 F 是一個 *finite field,* 試找出一個 *Theorem 1.2.4* 的反例.

Question 1.8. 設 V 是一個 *over infinite field* F 的 *vector space* 且 $V_1,\ldots,V_n$ 為 V 的 F-*subspaces.* 你可以看出 *Theorem 1.2.4* 告訴我們 $V_1\cup\cdots\cup V_n$ 不會是一個 F-*space* 除非存在 $i\in\{1,\ldots,n\}$ 滿足 $V_j\subseteq V_i$, $\forall j\in\{1,\ldots,n\}$.

一般來說雖然一個 vector space 中的一些 subspaces 所成的聯集並不一定是一個 vector space, 但是我們仍能造出一個包含這些 subspaces 的 vector space. 我們需要以下的定義.

Definition 1.2.5. 假設 V 為一個 over F 的 vector space 且 $V_1,\ldots,V_n$ 為 V 的 F-subspaces. 定義

$$V_1+\cdots+V_n=\{\mathbf{v}_1+\cdots+\mathbf{v}_n\mid \mathbf{v}_i\in V_i,\ 1\leq i\leq n\}.$$

要注意這只是為了以後方便使用所定的符號, 並不是要對這些 subspaces 定義加法.

Question 1.9. 當 V 是一個 F-*space* 而 W 是 V 的 *subspace,* 依定義 $W+W$ 會是什麼?

依照定義, 我們可得到以下的性質.

Proposition 1.2.6. 假設 V 為一個 *over* F 的 *vector space* 且 $V_1,\ldots,V_n$ 為 V 的 *subspaces.* 則 $V_1+\cdots+V_n$ 是 V 的 *subspace.*

Proof. 因 $\mathbf{O}\in V_i$ for $1\leq i\leq n$, 故 $\mathbf{O}\in V_1+\cdots+V_n$. 另一方面, 若 $\mathbf{u}_i,\mathbf{v}_i\in V_i$, $r,s\in F$ 因 V_i 是 F-subspace, 故 $r\mathbf{u}_i+s\mathbf{v}_i\in V_i$, 亦即

$$r(\mathbf{u}_1+\cdots+\mathbf{u}_n)+s(\mathbf{v}_1+\cdots+\mathbf{v}_n)=(r\mathbf{u}_1+s\mathbf{v}_1)+\cdots+(r\mathbf{u}_n+s\mathbf{v}_n)\in V_1+\cdots+V_n.$$

故由 Proposition 1.2.1 得知 $V_1+\cdots+V_n$ 是 V 的 subspace. □

Question 1.10. 假設 V 為一個 *over* F 的 *vector space* 且 U,W 為 V 的 *subspaces.* 很容易知道 $U\cap W$ 是 V 中包含於 U 且包含於 W 最大的 *subspace.* 而什麼是 V 中包含 U 且包含 W 最小的 *subspace?*

1.3. Spanning Sets

我們介紹 linear combination 的概念, 進而引進另一種得到 subspace 的方法.

Definition 1.3.1. 令 V 是一個 over F 的 vector space 且令 S 為 V 的一個非空子集. 若 $\mathbf{v}\in V$ 滿足 $\mathbf{v}=r_1\mathbf{v}_1+\cdots+r_n\mathbf{v}_n$, 其中 $r_i\in F$ 且 $\mathbf{v}_i\in S$, 則稱 $\mathbf{v}$ 是 S 的一個 *linear combination*. 我們用 $\mathrm{Span}(S)$ 來表示所有 S 的 linear combination 所成的集合.

要注意, 即使 S 是一個 infinite set, 每一個 S linear combination 僅牽涉到有限多個 S 中有限多個元素. 有的書會將 S 的一個 linear combination 用 $\mathbf{v}=\sum_{\mathbf{u}\in S} r_{\mathbf{u}}\mathbf{u}$ 這樣來表示, 不過都會標明其中每個 $r_{\mathbf{u}}\in F$ 且僅有有限多個 $r_{\mathbf{u}}$ 不等於 0. 另外若 $\mathbf{v}=\sum_{\mathbf{u}\in S} r_{\mathbf{u}}\mathbf{u}=\sum_{\mathbf{u}\in S} s_{\mathbf{u}}\mathbf{u}$, 其中有某個 $\mathbf{u}\in S$ 有 $r_{\mathbf{u}}\neq s_{\mathbf{u}}$, 則稱 $\mathbf{v}$ 寫成 S 的 linear combination 其表法“不唯一”.

Question 1.11. 由定義能看出若 $S'\subseteq S\subseteq V$, 則 $\mathrm{Span}(S')\subseteq \mathrm{Span}(S)$ 嗎? 一般來說將 S 拿掉一些元素有可能會使得 $\mathrm{Span}(S)$ 變小. 想想看在 S 中拿掉哪些元素才不會使得 $\mathrm{Span}(S)$ 變小呢?

若 S 是非空的, 任取 $\mathbf{w}\in S$, 因 $0\mathbf{w}=\mathbf{O}$ 知 $\mathbf{O}$ 是一個 S 的 linear combination. 故 $\mathbf{O}\in \mathrm{Span}(S)$. 另外若 $\mathbf{u}=r_1\mathbf{u}_1+\cdots+r_n\mathbf{u}_n$, $\mathbf{v}=s_1\mathbf{v}_1+\cdots+s_m\mathbf{v}_m$ 是 S 的 linear combination (即 $r_i,s_j\in F$ 且 $\mathbf{u}_i,\mathbf{v}_j\in S$), 很明顯的對任意 $r,s\in F$,

$$\begin{aligned} r\mathbf{u}+s\mathbf{v} &= r(r_1\mathbf{u}_1+\cdots+r_n\mathbf{u}_n)+s(s_1\mathbf{v}_1+\cdots+s_m\mathbf{v}_m) \\ &= (rr_1)\mathbf{u}_1+\cdots(rr_n)\mathbf{u}_n+(ss_1)\mathbf{v}_1+\cdots+(ss_m)\mathbf{v}_m \end{aligned}$$

也是 S 的 linear combination. 所以利用 Proposition 1.2.1 我們有以下之結果.

Lemma 1.3.2. *若 V 是一個 over F 的 vector space 且 S 為 V 的一個非空子集, 則 $\mathrm{Span}(S)$ 為 V 的一個 subspace.*

既然 $\mathrm{Span}(S)$ 是一個 F-subspace 我們便稱之為 the *subspace spanned by S*. 事實上 $\mathrm{Span}(S)$ 是 V 中包含 S 最小的 subspace.

Proposition 1.3.3. *若 V 是一個 over F 的 vector space 且 S 為 V 的一個非空子集, 則*

$$\mathrm{Span}(S)=\bigcap_{S\subseteq W,\, W\leq V} W,$$

這裡 W 是考慮 V 中所有包含 S 的 subspaces.

Proof. 首先由 Lemma 1.3.2 知 $\mathrm{Span}(S)$ 就是一個包含 S 的 subspace, 所以自然有

$$\mathrm{Span}(S)\supseteq \bigcap_{S\subseteq W,\, W\leq V} W.$$

另一方面若 W 是 V 的 subspace 且 $S\subseteq W$ 則任取 $\mathbf{v}\in \mathrm{Span}(S)$, 因 $\mathbf{v}=r_1\mathbf{v}_1+\cdots+r_n\mathbf{v}_n$, 其中 $r_1\in F$, $\mathbf{v}_i\in S\subseteq W$, 故由 W 為 subspace 得 $\mathbf{v}\in W$, 亦即 $\mathrm{Span}(S)\subseteq W$ (此即表示 $\mathrm{Span}(S)$ 是 V 中包含 S 最小的 subspace). 因此

$$\mathrm{Span}(S)\subseteq \bigcap_{S\subseteq W,\, W\leq V} W,$$

故得證. □

當 $S=\emptyset$, 因為所有集合皆包含空集合, Proposition 1.3.3 中的交集部分就是所有 V 的 subspaces 的交集, 也就是 $\{\mathbf{O}\}$. 所以當 S 是空集合時, 我們也定義 $\operatorname{Span}(S)=\{\mathbf{O}\}$.

Question 1.12. 假設 V 為一個 *over* F 的 *vector space* 且 $V_1,\dots,V_n$ 為 V 的 *subspaces.* 很容易看出 $\operatorname{Span}(V_i)=V_i$. 你能知道 $\operatorname{Span}(V_1\cup\cdots\cup V_n)$ 是什麼嗎?

前面 Question 1.11 中我們提到一般來說將 S 拿掉一些元素有可能會使得 $\operatorname{Span}(S)$ 變小. 以下我們回答在 S 中拿掉哪些元素才不會使得 $\operatorname{Span}(S)$ 變小.

Corollary 1.3.4. 令 V 為一個 *vector space over* F 且 $S'\subseteq S\subseteq V$. 則 $\operatorname{Span}(S)=\operatorname{Span}(S')$ 若且唯若 $S\setminus S'\subseteq\operatorname{Span}(S')$.

Proof. ($\Rightarrow$) 因 $S\setminus S'\subseteq S$ 依定義自然有 $S\setminus S'\subseteq\operatorname{Span}(S)$. 故由前提 $\operatorname{Span}(S)=\operatorname{Span}(S')$ 可得 $S\setminus S'\subseteq\operatorname{Span}(S')$.

($\Leftarrow$) 由 $S'\subseteq S$ 可得 $\operatorname{Span}(S')\subseteq\operatorname{Span}(S)$, 因此我們僅要證明 $\operatorname{Span}(S)\subseteq\operatorname{Span}(S')$. 事實上, 我們只要證明 $S\subseteq\operatorname{Span}(S')$ 即可. 這是因為 Lemma 1.3.2 告訴我們 $\operatorname{Span}(S')$ 是一個 subspace of V, 故若 S 包含於 $\operatorname{Span}(S')$ 則由 Proposition 1.3.3 可得 $\operatorname{Span}(S)\subseteq\operatorname{Span}(S')$. 然而對任意 $\mathbf{v}\in S$, 我們有 $\mathbf{v}\in S'$ 或 $\mathbf{v}\in S\setminus S'$. 若 $\mathbf{v}\in S'$ 依定義自然有 $\mathbf{v}\in\operatorname{Span}(S')$; 而若 $\mathbf{v}\in S\setminus S'$ 依假設也有 $\mathbf{v}\in\operatorname{Span}(S')$. 故知 $S\subseteq\operatorname{Span}(S')$, 因而得證 $\operatorname{Span}(S)=\operatorname{Span}(S')$. □

特別地, 當 V 是一個 F-space, 而 S 是 V 的子集滿足 $\operatorname{Span}(S)=V$, 則稱 S 是 V 的一個 *spanning set.* 依此定義很容易知道若 $S\subseteq S'\subseteq V$, 且 S 是 V 的 spanning set, 則 S' 也是 V 的 spanning set.

Example 1.3.5. 我們藉由 Example 1.1.1 的 vector spaces, 舉出它們幾個基本的 spanning sets.

(1) 在 F^n 中考慮 $\mathbf{e}_i=(0,\dots,1,\dots,0)$, 其中 1 是在第 i 個位置, 其他位置都是 0. 則 $\{\mathbf{e}_1,\dots,\mathbf{e}_n\}$ 是 F^n 的一個 spanning set.

(2) 在 $P_n(F)$ 中因為所有元素皆可用 $a_nx^n+\cdots+a_1x+a_0$ 來表示, 其中 $a_i\in F$, 所以 $\{1,x,\dots,x^n\}$ 是 $P_n(F)$ 的一個 spanning set.

(3) 在 F^S 中, 任取 $\lambda\in S$ 定義 $f_\lambda\in F^S$, 為

$$f_\lambda(s)=\begin{cases}1, & s=\lambda;\\ 0, & s\neq\lambda.\end{cases}$$

當 S 是一個 finite set 時, $\{f_\lambda\mid\lambda\in S\}$ 是 F^S 的一個 spanning set. 不過當 S 是 infinite set 時, 這就不對了. 這是因為在 vector space 中我們僅考慮有限多個元素相加 (無窮多個元素相加會有收斂發散的問題, 這會牽涉到 "Topology", 不是一般線性代數所談的範疇).

Question 1.13. 在上述 F^S 的情形, 若 S 是一個 *infinite set,* $\operatorname{Span}(\{f_\lambda\mid\lambda\in S\})$ 是什麼?

1.4. Linear Independence

我們介紹 linear independence 的概念, 不過為了避免大家邏輯上可能發生的謬誤, 我們由 linear dependence 的概念出發. 事實上 Linear independence 和 spanning set 之間有許多相呼應的性質, 希望大家研習此節一些結果時能常常和前一節的結果相對照.

Definition 1.4.1. 令 V 是一個 over F 的 vector space 且令 S 為 V 的一個非空子集. 若存在 $\mathbf{v} \in S$ 滿足 $\mathbf{v} \in \mathrm{Span}(\{\mathbf{w} \in S \mid \mathbf{w} \neq \mathbf{v}\})$, 則稱 S 為 *linearly dependent.* 反之, 則稱 S 為 *linearly independent.*

依此定義我們很容易得知, 若 $\mathbf{O} \in S$, 則 S 一定是 linearly dependent. 因為 $\mathbf{O}$ 一定會在任何的 subspace 中.

Question 1.14. *依此定義你能看出若 $S \subseteq S'' \subseteq V$, 而 S'' 是 linearly independent, 則 S 是 linearly independent 嗎? (or 若 S 是 linearly dependent, 則 S'' 是 linearly dependent) 不過若 S 是 linearly independent, S'' 不一定是 linearly independent. 也就是說一個 linearly independent 的集合多加了元素後有可能變成 linearly dependent. 要加入怎樣的元素才能保持 linearly independent 呢?*

Linear dependence 有以下等價的看法.

Proposition 1.4.2. *令 V 是一個 over F 的 vector space 且 S 為 V 的一個非空子集. 則 S 是 linearly dependent 若且唯若存在 $\mathbf{v}_1, \ldots, \mathbf{v}_n \in S$ 以及 $r_1, \ldots, r_n \in F$, 其中這些 $\mathbf{v}_i$ 皆相異而 r_i 全不為 0 使得 $r_1\mathbf{v}_1 + \cdots + r_n\mathbf{v}_n = \mathbf{O}$.*

Proof. ($\Rightarrow$) 由 S 是 linearly dependent 知存在 $\mathbf{v}_1 \in S$ 滿足 $\mathbf{v}_1 \in \mathrm{Span}(\{\mathbf{w} \in S \mid \mathbf{w} \neq \mathbf{v}_1\})$, 亦即存在 $\mathbf{v}_2, \ldots, \mathbf{v}_n \in S$ 皆相異且不等於 $\mathbf{v}_1$ 以及 $r_2, \ldots, r_n \in F$ 皆不為 0 使得 $\mathbf{v}_1 = r_2\mathbf{v}_2 + \cdots + r_n\mathbf{v}_n$. 故得 $(-1)\mathbf{v}_1 + r_2\mathbf{v}_2 + \cdots + r_n\mathbf{v}_n = \mathbf{O}$.

($\Leftarrow$) 假設對於 $i \in \{1, \ldots, n\}$, $\mathbf{v}_i \in S$ 皆相異且 $r_i \in F$ 全不為 0 使得 $r_1\mathbf{v}_1 + \cdots + r_n\mathbf{v}_n = \mathbf{O}$, 則 $\mathbf{v}_1 = (-r_2r_1^{-1})\mathbf{v}_2 + \cdots + (-r_nr_1^{-1})\mathbf{v}_n \in \mathrm{Span}(\{\mathbf{w} \in S \mid \mathbf{w} \neq \mathbf{v}_1\})$, 亦即 S 為 linearly dependent. □

提醒一下 Proposition 1.4.2 中若 $\mathbf{v}_1, \ldots, \mathbf{v}_n$ 中沒有 $\mathbf{O}$, 則會有 $n \geq 2$, 否則若僅有 $r_1 \neq 0$, 會導致 $r_1\mathbf{v}_1 = \mathbf{O}$ 而推得 $\mathbf{v}_1 = \mathbf{O}$.

另外 linear independence 也有以下等價的看法.

Proposition 1.4.3. *令 V 是一個 over F 的 vector space 且 S 為 V 的一個非空子集. 則 S 是 linearly independent 若且唯若 $\mathrm{Span}(S)$ 中任意的元素皆僅有唯一的方法寫成 S 中元素的 linear combination.*

Proof. ($\Rightarrow$) 我們利用反證法證明表法唯一. 首先若 $\mathbf{v} = \mathbf{O}$ 有兩種表示法, 即存在 $r_1, \ldots, r_n \in F$ 全不為 0 及 $\mathbf{v}_1, \ldots, \mathbf{v}_n \in S$ 使得 $r_1\mathbf{v}_1 + \cdots + r_n\mathbf{v}_n = \mathbf{O}$. 由 Proposition 1.4.2 知此與 S 為 linearly independent 相矛盾. 另一方面若 $\mathbf{v} \in \mathrm{Span}(S)$ 且 $\mathbf{v} \neq \mathbf{O}$ 有兩種寫法, 即存在

$r_1,\dots,r_n,s_1,\dots,s_m \in F$ (其中這些 r_i,s_j 皆不為 0) 以及 $\mathbf{v}_1,\dots,\mathbf{v}_n,\mathbf{w}_1,\dots,\mathbf{w}_m \in S$ (其中這些 $\mathbf{v}_1,\dots,\mathbf{v}_n$ 皆相異且 $\mathbf{w}_1,\dots,\mathbf{w}_m$ 皆相異) 使得

$$\mathbf{v} = r_1\mathbf{v}_1 + \cdots + r_n\mathbf{v}_n = s_1\mathbf{w}_1 + \cdots + s_m\mathbf{w}_m.$$

經適當排序後我們假設: $\mathbf{v}_1 = \mathbf{w}_1,\dots,\mathbf{v}_k = \mathbf{w}_k$ 且其他的 $\mathbf{v}_i,\mathbf{w}_j$ 皆相異. 而得

$$\mathbf{O} = (r_1 - s_1)\mathbf{v}_1 + \cdots + (r_k - s_k)\mathbf{v}_k + r_{k+1}\mathbf{v}_{k+1} + \cdots + r_n\mathbf{v}_n + s_{k+1}\mathbf{w}_{k+1} + \cdots + s_m\mathbf{w}_m.$$

依表法不同之假設, 若 $k = n = m$, 則必有某個 $r_i \neq s_i$; 而其他情形必有 $k < n$ (此時 $r_{k+1} \neq 0$); 或 $k < m$ (此時 $s_{k+1} \neq 0$). 因此由 Proposition 1.4.2 知此與 S 為 linearly independent 相矛盾, 故得證.

($\Leftarrow$) 依假設對任意 $\mathbf{v} \in S$ 因 $\mathbf{v} \in \mathrm{Span}(S)$ 且 $\mathbf{v} = 1\mathbf{v}$ 故由 $\mathbf{v}$ 寫成 S 中元素的 linear combination 的寫法僅一種, 得 $\mathbf{v} \notin \mathrm{Span}(\{\mathbf{w} \in S \mid \mathbf{w} \neq \mathbf{v}\})$, 依定義此即表示 S 為 linearly independent.

□

要注意, linearly dependent 和 linearly independent 是兩個互補的關係, 我們只是為了敘述方便將 Propositions 1.4.2, 1.4.3 分開, 其實它們是相關的. 例如要證明一個集合是 linearly independent, 你可以用反證法先假設它是 linearly dependent, 然後利用 Proposition 1.4.2 得到矛盾. 另一方面若有一個集合已知是 linearly independent, 你就可以利用 Proposition 1.4.3 表法的唯一性來推導其他相關的性質.

Question 1.15. 分別利用 *Proposition 1.4.2* 和 *Proposition 1.4.3* 來說明 $S = \{\mathbf{v}_1,\dots,\mathbf{v}_n\}$ 是 *linearly independent* 等價於

$$r_1\mathbf{v}_1 + \cdots + r_n\mathbf{v}_n = \mathbf{O} \Rightarrow r_1 = \cdots = r_n = 0.$$

Question 1.16. 若 S 是 *linearly dependent* 且 $\mathbf{O} \notin S$, *Proposition 1.4.3* 告訴我們存在一個 *Span*(S) 中的元素會有兩種 (或更多) 方法寫成 S 中元素的 *linear combination.* 你看得出在這種情形, 其實每一個 $\mathrm{Span}(S)$ 中的元素都會有兩種 (或更多) 方法寫成 S 中元素的 *linear combination* 嗎? 甚至當 F 是一個 *infinite field* 時, 每一個 $\mathrm{Span}(S)$ 中的元素都會有無窮多種方法寫成 S 中元素的 *linear combination.*

前面 Question 1.14 中我們提到一般來說將一個 linearly independent set S 多加一些元素有可能會使得 S 變成 linearly dependent. 以下我們回答在一個 linearly independent set 中多加哪些元素仍會保持 linearly independent.

Corollary 1.4.4. 令 V 為一個 *vector space over* F 且 $S \subseteq S'' \subseteq V$. 則 S'' 是 *linearly independent* 若且唯若 S 和 $S'' \setminus S$ 皆為 *linearly independent* 且 $\mathrm{Span}(S) \cap \mathrm{Span}(S'' \setminus S) = \{\mathbf{O}\}$.

Proof. ($\Rightarrow$) 因 $S \subseteq S'' \subseteq V$ 依定義若 S'' 是 linearly independent 自然 S 和 $S'' \setminus S$ 皆為 linearly independent. 今若存在 $\mathbf{v} \in \mathrm{Span}(S) \cap \mathrm{Span}(S'' \setminus S)$ 且 $\mathbf{v} \neq \mathbf{O}$, 即表示存在 $r_1,\dots,r_n,s_1,\dots,s_m \in F$ 皆不為 0 以及 $\mathbf{v}_1,\dots,\mathbf{v}_n \in S$, $\mathbf{w}_1,\dots,\mathbf{w}_m \in S'' \setminus S$ 使得 $\mathbf{v} = \sum_{i=1}^n r_i\mathbf{v}_i = \sum_{j=1}^m s_j\mathbf{w}_j$. 亦即

$\mathbf{v} \in \mathrm{Span}(S'')$ 但有兩種寫成 S'' 中元素的 linear combination 的寫法, 由 Proposition 1.4.3 知此與 S'' 為 linearly independent 相矛盾. 故得證 $\mathrm{Span}(S) \cap \mathrm{Span}(S'' \setminus S) = \{\mathbf{O}\}$.

($\Leftarrow$) 利用反證法, 若 S'' 是 linearly dependent, 由 Proposition 1.4.2 知存在 $r_1,\ldots,r_n \in F$ 全不為 0 以及 $\mathbf{v}_1,\ldots,\mathbf{v}_n \in S''$ 使得 $r_1\mathbf{v}_1 + \cdots + r_n\mathbf{v}_n = \mathbf{O}$. 因 S 以及 $S'' \setminus S$ 皆為 linearly independent, Proposition 1.4.2 告訴我們這些 $\mathbf{v}_i$ 不可能全落在 S 中也不可能全落在 $S'' \setminus S$ 中. 適當排序後, 我們假設 $\mathbf{v}_1,\ldots,\mathbf{v}_m \in S$ 且 $\mathbf{v}_{m+1},\ldots,\mathbf{v}_n \in S'' \setminus S$. 換言之, $r_1\mathbf{v}_1 + \cdots + r_m\mathbf{v}_m = (-r_{m+1})\mathbf{v}_{m+1} + \cdots + (-r_n)\mathbf{v}_n \in \mathrm{Span}(S) \cap \mathrm{Span}(S'' \setminus S)$. 然而 $r_1,\ldots,r_m$ 皆不為 0 且 $\mathbf{v}_1,\ldots,\mathbf{v}_m$ 為 linearly independent (因 S 為 linearly independent), 由 Proposition 1.4.3 知 $r_1\mathbf{v}_1 + \cdots + r_m\mathbf{v}_m \neq \mathbf{O}$, 亦即 $\mathrm{Span}(S) \cap \mathrm{Span}(S'' \setminus S) \neq \{\mathbf{O}\}$. 此矛盾告訴我們 S'' 是 linearly independent. $\square$

關於 linearly independent 的例子, 我們很容易驗證在 Example 1.3.5 中介紹的 spanning set 都是 linearly independent. 不過請不要誤以為 spanning set 一定就是 linearly independent. 例如在 $\mathbb{R}^n$ 的情況 $\{\mathbf{e}_1,\mathbf{e}_2,\ldots,\mathbf{e}_n,\mathbf{e}_1+\mathbf{e}_2\}$ 也是 $\mathbb{R}^n$ 的 spanning set, 不過就不再是 linearly independent 了.

1.5. Basis and Dimension

當 vector space V 中的一個子集合 S 太小時, 可能 S 無法成為 V 的 spanning set, 不過若 S 太大時, 又可能不是 linearly independent. Basis 就是保持平衡的最佳狀況. 我們有以下之定義.

Definition 1.5.1. 令 V 是一個 vector space over F 且 $S \subseteq V$. 當 $\mathrm{Span}(S) = V$ 且 S 是 linearly independent 時, 我們稱 S 為 V 的一組 basis.

依此定義, 我們可以知道在 Example 1.3.5 中介紹的 $\{\mathbf{e}_1,\ldots,\mathbf{e}_n\}$ 就是 F^n 的一組 basis; 而 $\{1,x,\ldots,x^n\}$ 就是 $P_n(F)$ 的一組 basis; 而當 S 是一個 finite set 時, $\{f_\lambda \mid \lambda \in S\}$ 就是 F^S 的一組 basis.

Question 1.17. 依此定義利用前一節的 *Proposition,* 你能看出 S 是 V 的一組 *basis* 等價於任何 V 中的元素都可以唯一寫成 S 中元素的 *linear combination* 嗎?

Question 1.18. 若 $S' \subsetneq S \subsetneq S''$ 且 S 是 V 的一組 *basis,* 那麼 S', S'' 有可能是 V 的一組 *basis* 嗎? 你能看出此時 $\mathrm{Span}(S') \neq V$ 而 S'' 一定是 *linearly dependent* 嗎?

由上一個 Question 我們知道若 S 是 V 的一組 basis, 則不會有其他的 spanning set 會包含於 S; 另一方面也知不會有其他的 linearly independent set 會包含 S. 這個論述的反向也是正確的, 事實上我們有以下之結果.

Proposition 1.5.2. 令 V 為一個 *vector space over* F 且 $S \subseteq V$. 我們有以下的等價關係.

(1) S 是 V 的一組 *basis.*

(2) S 是 V 的一個 *spanning set,* 且對任意 $S' \subsetneq S$ 皆有 $\mathrm{Span}(S') \neq V$.

(3) S 是 *linearly independent* 且對任意 $S'' \supsetneq S$ 皆有 S'' 為 *linearly dependent.*

Proof. 我們證明 (1)(2) 是等價的, 再證明 (1)(3) 為等價.

((1) $\Rightarrow$ (2)) 因 S 是一組 basis 由定義知 S 是 V 的一組 spanning set. 任取 $S' \subsetneq S$, 利用反證法假設 $\mathrm{Span}(S') = V$. 此時取 $\mathbf{v} \in S \setminus S'$ 可得 $\mathbf{v} \in V = \mathrm{Span}(S') \subseteq \mathrm{Span}(\{\mathbf{w} \in S \mid \mathbf{w} \neq \mathbf{v}\}$. 此與 S 為 linearly independent 相矛盾, 故得 $\mathrm{Span}(S') \neq V$.

((2) $\Rightarrow$ (1)) 因已知 S 為 spanning set, 依定義我們僅要證明 S 為 linearly independent. 利用反證法, 假設 S 為 linearly dependent, 亦即存在 $\mathbf{v} \in S$ 滿足 $\mathbf{v} \in \mathrm{Span}(S \setminus \{\mathbf{v}\})$. 故考慮 $S' = S \setminus \{\mathbf{v}\}$. 因 $S \setminus S' = \{\mathbf{v}\}$, 由 Corollary 1.3.4 知 $\mathrm{Span}(S') = \mathrm{Span}(S) = V$, 但因 $S' \subsetneq S$, 此與 (2) 的前提相矛盾, 故得證 S 為 linearly independent.

((1) $\Rightarrow$ (3)) 因 S 是一組 basis 由定義知 S 為 linearly independent. 任取 $S'' \supsetneq S$, 利用反證法假設 S'' 為 linearly independent. 此時取 $\mathbf{v} \in S'' \setminus S$ 可得 $\mathbf{v} \notin \mathrm{Span}(S)$. 此與 S 為 V 的 spanning set 相矛盾, 故得 S'' 為 linearly dependent.

((3) $\Rightarrow$ (1)) 因已知 S 為 linearly independent, 依定義我們僅要證明 S 為 V 的 spanning set. 利用反證法, 假設 $\mathrm{Span}(S) \neq V$ 為, 亦即存在 $\mathbf{v} \in V$ 滿足 $\mathbf{v} \notin \mathrm{Span}(S)$ (亦即 $\mathrm{Span}(\{\mathbf{v}\}) \cup \mathrm{Span}(S) = \{\mathbf{O}\}$). 故考慮 $S'' = S \cup \{\mathbf{v}\}$. 因 $S'' \setminus S = \{\mathbf{v}\}$, 由 Corollary 1.4.4 知 S'' 為 linearly independent, 但因 $S'' \supsetneq S$, 此與 (3) 的前提相矛盾, 故得證 S 為 V 的 spanning set.

□

在本講義中我們將專注於有一組 finite set 為 basis 的 vector space, 我們有以下之定義.

Definition 1.5.3. 假設 V 是一個 vector space over F. 若 $V = \{\mathbf{O}\}$ 或是存在一個 finite set $S \subseteq F$ 是 V 的一組 basis, 則稱 V 為一個 *finite dimensional* F-space.

注意在此我們特別將 $V = \{\mathbf{O}\}$ 這個情況列出, 主要是依定義 $\mathrm{Span}(\emptyset) = \{\mathbf{O}\}$, 所以我們定義 $\emptyset$ 為 $\{\mathbf{O}\}$ 的 basis. 在前面提過的例子 F^n, $P_n(F)$ 以及當 S 是 finite set 的 F^S 皆為 finite dimensional vector space over F. 不過要特別注意若 F' 是 F 的一個 subfield, 前面提過一個 F-space 可看成為 F'-space, 因此在此情況一定要強調是 over 哪一個 filed 為 finite dimensional, 因為有可能一個 finite dimensional F-space 不是 finite dimensional F'-space. 例如 $\mathbb{R}^n$ 是 finite dimensional $\mathbb{R}$-space 卻不是 finite dimensional $\mathbb{Q}$-space.

我們在 Definition 1.5.1 中定義了何謂 vector space 的 basis, 但我們還沒有去探討是否一個 vector space 一定會有一組 basis. 所以依目前的定義一個 vector space 若不是 finite dimensional 則有可能它沒有 basis 或是有 basis 但是所有的 basis 皆為 infinite set. 事實前者不會發生, 也就是說我們可以證明所有的 vector space 都會有 basis, 不過它的證明牽涉到所謂 Zorn's lemma. 由於以後我們將只專注於 finite dimensional vector space, 再加上有些同學可能覺得 Zorn's lemma 不是很容易了解. 以下我們的探討將分為兩段: 第一段專注於 finite dimensional 的情形; 第二段談的是一般的情形. 第一段的內容相當重要, 而第二段談的就是所有的 vector space 皆會有 basis. (大家可以直接接受這個事實而略過第二段).

1.5.1. Finite Dimensional Case. 當 V 是 finite dimensional vector space, 依定義是有一個 finite set 為 V 的一組 basis. 會不會有一個 infinite set 也是 V 的一組 basis 呢? 答案是不會的, 事實上這時候每一個組成 V 的 basis 的集合它們的元素個數是相同的. 以下我們便是要處理這一個問題. 為了方便起見, 給定一個 finite set S, 我們用 $\#(S)$ 來表示 S 中元素的個數.

Lemma 1.5.4. 令 V 是一個 *vector space over* F 且假設 $S \subseteq V$ 是一個 *finite set* 滿足 $\mathrm{Span}(S) = V$. 若 $S' \subseteq V$ 且 $\#(S') > \#(S)$, 則 S' 為 *linearly dependent.*

Proof. 令 $S = \{\mathbf{v}_1, \dots, \mathbf{v}_n\}$, $S' = \{\mathbf{u}_1, \dots, \mathbf{u}_m\}$ 其中 $m > n$. 利用反證法, 我們假設 S' 為 linearly independent.

由於 $\mathrm{Span}(S) = V$, 故存在 $r_1, \dots, r_n \in F$ 使得 $\mathbf{u}_1 = r_1\mathbf{v}_1 + \cdots + r_n\mathbf{v}_n$. 因假設 S' 為 linearly independent, 我們知 $r_1, \dots, r_n$ 不全為 0. (否則 $r_1 = \cdots = r_n = 0$ 會導致 $\mathbf{O} = \mathbf{u}_1 \in S'$, 那麼 S' 就不會是 linearly independent 了.) 因此不失一般性, 我們假設 $r_1 \neq 0$, 此時

$$\mathbf{v}_1 = r_1^{-1}(\mathbf{u}_1 - r_2\mathbf{u}_2 - \cdots - r_n\mathbf{v}_n) \in \mathrm{Span}(\{\mathbf{u}_1, \mathbf{v}_2, \dots, \mathbf{v}_n\}).$$

故利用 Corollary 1.3.4 知

$$\mathrm{Span}(\{\mathbf{u}_1, \mathbf{v}_2, \dots, \mathbf{v}_n\}) = \mathrm{Span}(\{\mathbf{u}_1, \mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n\}) = V.$$

接著利用 $\mathbf{u}_2 \in \mathrm{Span}(\{\mathbf{u}_1, \mathbf{v}_2, \dots, \mathbf{v}_n\})$ 知存在 $s_1, \dots, s_n \in F$ 使得 $\mathbf{u}_2 = s_1\mathbf{v}_1 + s_2\mathbf{u}_2 + \cdots + s_n\mathbf{u}_n$. 同理利用 S' 為 linearly independent 的假設, 得 $s_2, \dots, s_n$ 不全為 0. (否則若 $s_2 = \cdots = s_n = 0$, 則得 $\mathbf{u}_2 = s_1\mathbf{u}_1 \in \mathrm{Span}(\{\mathbf{u}_1\})$, 那麼 S' 就不會是 linearly independent 了.) 故不失一般性, 我們假設 $s_2 \neq 0$, 此時

$$\mathbf{v}_2 = s_2^{-1}(\mathbf{u}_2 - s_1\mathbf{u}_1 - s_3\mathbf{v}_3 - \cdots - s_n\mathbf{v}_n) \in \mathrm{Span}(\{\mathbf{u}_1, \mathbf{u}_2, \mathbf{v}_3, \dots, \mathbf{v}_n\}).$$

故利用 Corollary 1.3.4 知

$$\mathrm{Span}(\{\mathbf{u}_1, \mathbf{u}_2, \mathbf{v}_3, \dots, \mathbf{v}_n\}) = \mathrm{Span}(\{\mathbf{u}_1, \mathbf{u}_2, \mathbf{v}_2, \dots, \mathbf{v}_n\}) = V.$$

簡單來說, 我們將 $\mathbf{v}_1, \dots, \mathbf{v}_n$ 做適當排序後, 可以用 $\mathbf{u}_1$ 取代 $\mathbf{v}_1$; $\mathbf{u}_2$ 取代 $\mathbf{v}_2$, ... 如此一直下去. 利用數學歸納法, 我們假設 $k < n$ 且 $\mathrm{Span}(\{\mathbf{u}_1, \dots, \mathbf{u}_k, \mathbf{v}_{k+1}, \dots, \mathbf{v}_n\}) = V$, 想要用 $\mathbf{u}_{k+1}$ 來取代剩下的某個 $\mathbf{v}_i$. 如同前面的做法, 我們知道存在 $t_1, \dots, t_n \in F$ 使得 $\mathbf{u}_{k+1} = t_1\mathbf{u}_1 + \cdots + t_k\mathbf{u}_k + t_{k+1}\mathbf{v}_{k+1} + \cdots t_n\mathbf{v}_n$. 利用 S' 為 linearly independent 的假設, 得 $t_{k+1}, \dots, t_n$ 不全為 0. 故不失一般性, 我們假設 $t_{k+1} \neq 0$, 此時

$$\mathbf{v}_{k+1} = t_{k+1}^{-1}(\mathbf{u}_{k+1} - t_1\mathbf{u}_1 - \cdots - t_k\mathbf{u}_k - t_{k+2}\mathbf{v}_{k+2} \cdots - t_n\mathbf{v}_n) \in \mathrm{Span}(\{\mathbf{u}_1, \dots, \mathbf{u}_{k+1}, \mathbf{v}_{k+2}, \dots, \mathbf{v}_n\}).$$

故利用 Corollary 1.3.4 知

$$\mathrm{Span}(\{\mathbf{u}_1, \dots, \mathbf{u}_{k+1}, \mathbf{v}_{k+2}, \dots, \mathbf{v}_n\}) = \mathrm{Span}(\{\mathbf{u}_1, \dots, \mathbf{u}_k, \mathbf{u}_{k+1}, \mathbf{v}_{k+1}, \dots, \mathbf{v}_n\}) = V.$$

數學歸納法告訴我們, 可以如此一直下去直到將所有的 $\mathbf{v}_1, \dots, \mathbf{v}_n$ 置換完畢, 亦即得 $\mathrm{Span}(\{\mathbf{u}_1, \dots, \mathbf{u}_n\}) = V$. 不過如此一來會造成 $\mathbf{u}_{n+1} \in \mathrm{Span}(\{\mathbf{u}_1, \dots, \mathbf{u}_n\})$ 而與 S' 為 linearly independent 相矛盾, 故 S' 必為 linearly dependent. $\square$

Question 1.19. 在 *Lemma 1.5.4* 中若將 S 為 *spanning set* 改為 *linearly dependent,* 當 $S' \subseteq V$, 你覺得 S' 相對應的有關於 *spanning set* 的性質應該是甚麼?

Lemma 1.5.4 是有關 finite dimensional vector space 相當重要的定理, 它告訴我們一個 linearly independent set 的元素個數不可以有多於一個 spanning set 的元素個數. 我們有以下幾個重要應用.

Theorem 1.5.5. 假設 V 是一個 *finite dimensional vector space.* 若 S 為 V 的一組 *basis,* 則 S 是一個 *finite set.* 另外若 S' 亦為 V 的一組 *basis,* 則 $\#(S) = \#(S')$.

Proof. 依 V 為 finite dimensional vector space 之假設, 存在 $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ 為 V 的一組 basis. 故我們有 $\mathrm{Span}(\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}) = V$.

今若 S 為 V 的一組 basis 且為 infinite set, 由於 S 為 linearly independent, S 的任何 subset 亦為 linearly independent. 故任取 $\mathbf{u}_1, \ldots, \mathbf{u}_{n+1} \in S$, $\{\mathbf{u}_1, \ldots, \mathbf{u}_{n+1}\} \subseteq S$ 亦為 linearly independent, 此明顯與 Lemma 1.5.4 相矛盾, 故知 S 必為 finite set.

現若 S, S' 皆為 V 的一組 basis, 由於 $\mathrm{Span}(S) = V$ 且 S' 為 linearly independent, Lemma 1.5.4 告訴我們 $\#(S) \geq \#(S')$. 同理由 $\mathrm{Span}(S') = V$ 以及 S 為 linearly independent, 得 $\#(S) \leq \#(S')$. 故得證 $\#(S) = \#(S')$. □

由 Theorem 1.5.5 知一個 finite dimension vector space 它的任一組 basis 的元素個數是固定的. 因此我們有以下之定義.

Definition 1.5.6. 令 V 為一個 finite dimensional vector space over F. 若 S 為 V 的一組 basis 且 $\#(S) = n$, 則稱 V over F 的 *dimension* 為 n, 記做 $\dim(V) = n$.

依此定義由於我們視 $\emptyset$ 為 $\{\mathbf{O}\}$ 的 basis, 故 $\dim(\{\mathbf{O}\}) = 0$.

注意, 我們談過當 V 看成 over 不同 field 的 vector space 時, 其 dimension 就可能不同了. 當這種情形發生時, 我們特別用 $\dim_F(V)$ 來強調是 over F 的 dimension. 例如複數 $\mathbb{C}$ 可看成是 over $\mathbb{C}$ 或 over $\mathbb{R}$ 的 vector space, 而我們有 $\dim_{\mathbb{C}}(\mathbb{C}) = 1$ 及 $\dim_{\mathbb{R}}(\mathbb{C}) = 2$.

Question 1.20. 你知道 $\dim_F(F^n)$, $\dim_F(P_n(F))$ 以及 $\dim_F(F^S)$ *(S 為 finite set)* 為何嗎? 又若 S 為 *infinite set,* 如何說明 F^S 不是 *finite dimensional F-space?*

Question 1.21. 設 V 為 *finite dimensional F-space* 且令 $\dim(V) = n$. 若 $S' \subseteq V$ 為 *linearly independent* 及 $S'' \subseteq V$ 為 V 的一個 *spanning set,* 則 $\#(S')$ 和 $\#(S'')$ 與 n 的關係為何?

利用 dimension 我們可將 Proposition 1.5.2 化為以下形式.

Corollary 1.5.7. 令 V 為一個 *finite dimensional vector space over F* 且 $S' \subseteq V$. 我們有以下的等價關係.

(1) S 是 V 的一組 *basis.*

(2) S 是 V 的一個 *spanning set* 且 $\#(S) = \dim(V)$.

(3) S 是 *linearly independent* 且 $\#(S) = dim(V)$.

Proof. 令 $\dim(V) = n$. 若 S 是 V 的一組 basis, 依定義 S 為 V 的 spanning set 且 S 為 linearly independent. 又由 dimension 的定義知 $\#(S) = \dim(V)$. 故 $(1) \Rightarrow (2)$ 且 $(1) \Rightarrow (3)$.

$(2) \Rightarrow (1)$: 對任意 $S' \subsetneq S$, 我們有 $\#(S') < \#(S) = n$. 若 $\mathrm{Span}(S') = V$, 由 Lemma 1.5.4 知任意 n 個元素所成的集合皆不可能為 linearly independent, 此與 $\dim(V) = n$ 相矛盾. 故知 $\mathrm{Span}(S') \neq V$ 因此由 Proposition 1.5.2 ($(2) \Rightarrow (1)$) 得證 S 為 V 的一組 basis.

$(3) \Rightarrow (1)$: 對任意 $S'' \supsetneq S$, 我們有 $\#(S'') > \#(S) = n$. 因 $\dim(V) = n$ 表示 V 中存在一個有 n 個元素的 spanning set, 由 Lemma 1.5.4 知 S'' 必為 linearly dependent. 因此由 Proposition 1.5.2 ($(3) \Rightarrow (1)$) 得證 S 為 V 的一組 basis. □

若 V 是一個 finite dimensional F-space, 而 W 是 V 的 nontrivial F-subspace, 是否 W 也是 finite dimensional F-space 呢? 要注意, 我們不能直接套用 Lemma 1.5.4 來回答這個問題, 因為我們不知是否 W 有 basis, 不過我們可以利用 Proposition 1.5.2 的想法幫助我們解決這個問題.

我們說明 W 的 basis 是存在的. 考慮 $S_1 = \{\mathbf{v}_1\}$, 其中 $\mathbf{v}_1 \in W$ 且 $\mathbf{v}_1 \neq \mathbf{O}$. 依定義 S_1 是 linearly independent. 若 $\mathrm{Span}(S_1) = W$, 則得到 S_1 為 W 的一組 basis; 而若 $\mathrm{Span}(S_1) \neq W$, 則任取 $\mathbf{v}_2 \in W \setminus \mathrm{Span}(S_1)$. 令 $S_2 = \{\mathbf{v}_1, \mathbf{v}_2\}$, 依定義 S_2 亦為 linearly independent. 若 $\mathrm{Span}(S_2) = W$, 則得到 S_2 為 W 的一組 basis; 而若 $\mathrm{Span}(S_2) \neq W$, 我們再加 $\mathbf{v}_3 \in W \setminus \mathrm{Span}(S_2)$ 得 $S_3 = \{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$, 由 Corollary 1.4.4 知 S_3 亦為 linearly independent. 如此一直下去得 $S_3, S_4, \ldots$ 其中 S_i 皆為 linearly independent. 若這個步驟無法停止 (即對所有 $n \in \mathbb{N}$ 皆無法得 $\mathrm{Span}(S_n) = W$) 表示對任意 $n \in \mathbb{N}$, 在 V 中皆存在有 n 個元素的集合 S_n 是 linearly independent. 但 V 是 finite dimensional vector space, 若 $n > \dim(V)$, 依 Lemma 1.5.4 知 S_n 不可能是 linearly independent. 因此這個步驟一定要停止, 亦即存在一個 m 使得 $\mathrm{Span}(S_m) = W$ 也就是說 S_m 是 W 的一組 basis, 故 W 亦為 finite dimensional F-space. 我們有以下之結果.

Theorem 1.5.8. 若 V 為一個 *finite dimensional F-space* 且 W 為 V 的一個 *nontrivial F-subspace,* 則 W 亦為 *finite dimensional F-space,* 且 $\dim(W) < \dim(V)$.

Proof. 前面已證得 W 為 finite dimensional F-space. 現假設 S 為 W 的一組 basis, 因 S 為 linearly independent, 由 Lemma 1.5.4 知 $\dim(W) = \#(S) \leq \dim(V)$. 若 $\#(S) = \dim(V)$, 則由 Corollary 1.5.7 ($(3) \Rightarrow (1)$) 知 S 亦為 V 的一組 basis, 得 $W = \mathrm{Span}(S) = V$. 此與 W 為 nontrivial subspace 相矛盾, 故知 $\dim(W) < dim(V)$. □

在 Theorem 1.5.8 的證明過程中, 我們事實上是將一個 finite dimensional vector space 中的一組 linearly independent 的子集合設法加入元素且保持 linearly independent, 這樣一直擴大到無法再擴大為止, 然後就可以由 Proposition 1.5.2 得知此為一組 basis. 相對應的, 一個 spanning set 我們也可以設法在其中取出元素且保持仍為 spanning set 直到無法再縮小為止, 此時也可由 Proposition 1.5.2 得知其為一組 basis. 我們有以下之結果.

Theorem 1.5.9. 若 V 為一個 *finite dimensional* F*-space* 且 $S' \subseteq S'' \subseteq V$, 其中 S' 為 *linearly independent* 而 S'' 為 V 的 *spanning set,* 則 V 有一組 *basis* S 滿足 $S' \subseteq S \subseteq S''$.

Proof. 我們在 S' 中逐一加入 $S'' \setminus S'$ 的元素使其仍保持 linearly independent. 由於 V 是 finite dimensional vector space, 由 Lemma 1.5.4 知不可能一直如此將 S' 擴大下去. 我們令加入元素到最後不能再擴大的集合為 S, 也就是說 $S'' \setminus S$ 中的元素都在 $\mathrm{Span}(S)$ 中. 依假設 S 為 linearly independent, 所以我們要證明 $\mathrm{Span}(S) = V$, 亦即 $\mathrm{Span}(S) = \mathrm{Span}(S'')$. 然而依 S 的選取方式, $S'' \setminus S \subseteq \mathrm{Span}(S)$, 故 Corollary 1.3.4 告訴我們這是成立的. □

我們再強調 Theorem 1.5.9 中 S 的選取並不一定是唯一的, 但是 S 中元素個數是固定的, 即為 $\dim(V)$.

Question 1.22. 你能看出 *Theorem 1.5.9* 告訴我們一個 *finite dimensional vector space* 中的一個 *linearly independent set* 皆可擴大成為一組 *basis;* 而一個 *spanning set* 皆可縮小成為一組 *basis?*

下一段我們談論一般 vector space 時, 就是利用和 Theorem 1.5.9 類似的結果 (除去 finite dimensional 的假設) 來證明所有的 vector space 皆存在一組 basis.

1.5.2. General Case. 在這一小段中我們要說明所有的 vector space 皆存在一組 basis. 由於這個證明需用到 Zorn's Lemma 而且以後我們不會用到這個結果, 所以若可以接受這個事實的同學, 可將此當成已知, 略過這一小段不讀.

要證明所有的 vector space 皆存在一組 basis, 我們可以沿用前面的想法, 一個一個加入此 vector space 中的元素且保持 linearly independent, 直到不能再加為止. 但問題是我們無法得知這樣的步驟一定會停止 (除非事先知道此 vector space 為 finite dimensional). 所以要利用 Proposition 1.5.2 得到一組 basis 我們必須用到 Zorn's Lemma. 首先我們將簡單的介紹一下這個 Lemma.

通常在一個集合中給定了比較大小 (order) 的關係後, 有可能這個集合中每個元素相互間都可比較大小 (例如實數上的大於, 小於關係), 我們稱此為 totally ordered; 也有可能兩個元素間並不一定能比較大小 (例如集合間的包含關係), 我們稱此為 partially ordered. 在 totally ordered 的情形, 當我們談到 maximal element 時, 指的是此元素比其他的元素大; 不過若在 partially ordered 的情形這樣的定義並不合理, 因為並不是所有的元素都可以比大小. 所以在此時 maximal element 指的是沒有其他的元素比該元素大. 同樣的道理, minimal element 指的便是沒有其他的元素比該元素小. 利用這種說法一個 vector space 的一組 basis 利用 Proposition 1.5.2 便可以說成就是該 vector space 中所有 spanning set 的 minimal element; 也可以說是該 vector space 中所有 linearly independent set 的 maximal element. 所以要說明一個 vector space 的 basis 是存在的, 我們必須說明此 vector space 中所有 linearly independent set 的 maximal element 是存在的 (或所有 spanning set 的 minimal element 存在). 這就是我們需要 Zorn's Lemma 的地方.

Zorn's Lemma 是一個檢查 partially ordered set 的 maximal element 存在的方法. 它說若一個非空的 partially ordered set $\mathscr{P}$ 中任取一個遞增的序列 (ascending chain) 皆可在 $\mathscr{P}$ 中找到一元素比此序列中的所有元素都大, 則 $\mathscr{P}$ 的 maximal element 便會存在. 所以目前的情形, 要使用 Zorn's Lemma, 我們可以考慮一個 vector space V 中所有的 linearly independent set 所成的集合 $\mathscr{P}$. 也就是說 $\mathscr{P}$ 中的元素都是 V 的一組 linearly independent set. 我們考慮 $\mathscr{P}$ 中元素的包含關係所形成的 partial order. 此時 $\mathscr{P}$ 中任意的 ascending chain, 即為 $S_1 \subseteq S_2 \subseteq \cdots \subseteq S_n \subseteq \cdots$ 這種遞增的 linearly independent sets (其中 S_i 皆在 $\mathscr{P}$ 所以都是 V 的 linearly independent set). 若我們在 $\mathscr{P}$ 中能找到 S (即 S 為 V 的一組 linearly independent set) 滿足 $S_i \subseteq S, \forall i \in \mathbb{N}$, 則由 Zorn's Lemma 得知 $\mathscr{P}$ 中會有 maximal element. 事實上我們有以下有關於 Theorem 1.5.9 的推廣.

Theorem 1.5.10. 若 V 為一個 *vector space* 且 $S' \subseteq S'' \subseteq V$, 其中 S' 為 *linearly independent* 而 S'' 為 V 的 *spanning set,* 則 V 有一組 *basis* S 滿足 $S' \subseteq S \subseteq S''$.

Proof. 考慮 $\mathscr{P} = \{S \subseteq V \mid S \text{ is linearly independent}, S' \subseteq S \subseteq S''\}$. 注意, 因 $S' \in \mathscr{P}$, 故 $\mathscr{P}$ 為 nonempty. 今任取 $S_1 \subseteq S_2 \subseteq \cdots$, 其中對所有 $i \in \mathbb{N}$, S_i 皆在 $\mathscr{P}$. 首先令 $T = \cup_{i\in\mathbb{N}} S_i$. 我們要說明 T 在 $\mathscr{P}$ 中, 亦即 T 為 linearly independent 且 $S' \subseteq T \subseteq S''$. 因 S_i 皆滿足 $S' \subseteq S_i \subseteq S''$, 很自然的 $S' \subseteq T \subseteq S''$. 現假設 T 不為 linearly independent, 依定義此即表示存在 $\mathbf{v} \in T$ 且 $\mathbf{v} \in \mathrm{Span}(T \setminus \{\mathbf{v}\})$. 也就是說存在 $\mathbf{v}_1, \ldots, \mathbf{v}_n \in T$ 且這些 $\mathbf{v}_i$ 皆不等於 $\mathbf{v}$ 使得 $\mathbf{v} \in \mathrm{Span}(\{\mathbf{v}_1, \ldots, \mathbf{v}_n\})$. 因 $\mathbf{v} \in T$ 故由 $T = \cup_{i\in\mathbb{N}} S_i$ 知 $\mathbf{v}$ 會落在某個 S_k 中 (也因此會落在 $S_{k+1}, S_{k+2}, \ldots$), 同理每個 $\mathbf{v}_i$ 也會落在某個 S_{k_i} 中. 因此我們僅要取 $m = \max\{k, k_1, \ldots, k_n\}$, 則可得 $\mathbf{v}, \mathbf{v}_1, \ldots, \mathbf{v}_n$ 皆在 S_m 中. 故由 $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\} \subseteq S_m \setminus \{\mathbf{v}\}$, 得 $\mathbf{v} \in \mathrm{Span}(S_m \setminus \{\mathbf{v}\})$, 此與 S_m 為 linearly independent (因假設 S_m 在 $\mathscr{P}$ 中) 相矛盾故得證 T 為 linearly independent. 因此知 T 為 $\mathscr{P}$ 的一個元素且滿足 $S_i \subseteq T, \forall i \in \mathbb{N}$, 故由 Zorn's Lemma 得知 $\mathscr{P}$ 中存在著 maximal element, 我們令 S 為 $\mathscr{P}$ 的一個 maximal element.

接下來我們要證明 S 確為 V 的一組 basis, 且滿足 $S' \subseteq S \subseteq S''$. 首先 S 為 $\mathscr{P}$ 裡的元素, 自然有 S 為 linearly independent 且 $S' \subseteq S \subseteq S''$, 所以我們僅剩要證明 $\mathrm{Span}(S) = V$. 現假設 $\mathrm{Span}(S) \neq V = \mathrm{Span}(S'')$, 由 Corollary 1.3.4 知存在 $\mathbf{w} \in S''$ 但 $\mathbf{w} \notin \mathrm{Span}(S)$. 此時考慮 $S^+ = S \cup \{\mathbf{w}\}$. 很明顯的 $S' \subseteq S^+ \subseteq S''$, 然而 Corollary 1.4.4 告訴我們 S^+ 仍為 linearly independent, 因此 S^+ 亦為 $\mathscr{P}$ 的一個元素. 但是 $S \subsetneq S^+$, 此與 S 為 $\mathscr{P}$ 中的 maximal element 相矛盾, 故得證 $\mathrm{Span}(S) = V$. □

1.6. Direct Sum and Quotient Space

前面介紹 subspace 時, 我們談到幾種建構 subspace 的方法, 那些方法所得的 vector space 都是在原先的 vector space 中. 在這節中, 我們將介紹兩個建構“全新”的 vector space 的方法.

1.6.1. Direct Sum. 給定兩個 over F 的 vector spaces U,W (這裡不需假設 U,W 是某個 vector space 的 subspaces) 我們考慮一個新的集合

$$U \oplus W = \{(\mathbf{u},\mathbf{w}) \mid \mathbf{u} \in U, \mathbf{w} \in W\}.$$

稱此集合為 the (*external*) *direct sum* of U and W. 要注意 $U \oplus W$ 是一個新的集合, 所以我們要說明這個集合裡的元素怎樣才會相等 (這就好像當我們介紹多項式時要說明何謂多項式的相等). 在這裡我們要求若 $(\mathbf{u}_1,\mathbf{w}_1) = (\mathbf{u}_2,\mathbf{w}_2)$, 則 $\mathbf{u}_1 = \mathbf{u}_2$ 且 $\mathbf{w}_1 = \mathbf{w}_2$.

我們可以利用 U,W 本身為 vector space 的性質定義在 $U \oplus W$ 中的運算及 F 的作用. 給定 $(\mathbf{u}_1,\mathbf{w}_1),(\mathbf{u}_2,\mathbf{w}_2) \in U \oplus W$ 及 $r \in F$, 我們定義

$$\begin{aligned}(\mathbf{u}_1,\mathbf{w}_1)+(\mathbf{u}_2,\mathbf{w}_2) &= (\mathbf{u}_1+\mathbf{u}_2,\mathbf{w}_1+\mathbf{w}_2)\\ r(\mathbf{u}_1,\mathbf{w}_1) &= (r\mathbf{u}_1,r\mathbf{w}_1)\end{aligned}$$

在此定義之下, 很容易檢查 $U \oplus W$ 為一個 vector space over F.

Question 1.23. 請檢查 $U \oplus W$ 為一個 *vector space over* F*.* 你知道為什麼不能僅檢查封閉性呢? $U \oplus W$ 的 $\mathbf{O}$ (加法單位元素) 應該長什麼樣子呢?

Question 1.24. 若 U',W' 分別為 U,W 的 *subspaces,* 試證明 $U' \oplus W'$ 是 $U \oplus W$ 的 *subspace.* 反過來若 V 為 $U \oplus W$ 的 *subspace,* 是否可找到 U,W 的 *subspaces* U',W' 使得 $V = U' \oplus W'$?

若 U,W 為 finite dimensional F-spaces, 我們自然會問是否 $U \oplus W$ 亦為 finite dimensional F-space, 且其 dimension 為何?

Proposition 1.6.1. 假設 U,W 為 *finite dimensional* F*-spaces,* 則 $U \oplus W$ 亦為 *finite dimensional* F*-space,* 且

$$\dim(U \oplus W) = \dim(U) + \dim(W).$$

Proof. 設 $\{\mathbf{u}_1,\ldots,\mathbf{u}_m\}$ 為 U 的一組 basis 且 $\{\mathbf{w}_1,\ldots,\mathbf{w}_n\}$ 為 W 的一組 basis. 我們僅要證明 $S = \{(\mathbf{u}_1,\mathbf{O}_W),\ldots,(\mathbf{u}_m,\mathbf{O}_W),(\mathbf{O}_U,\mathbf{w}_1),\ldots,(\mathbf{O}_U,\mathbf{w}_n)\}$ 為 $U \oplus W$ 的一組 basis (其中 $\mathbf{O}_U,\mathbf{O}_W$ 分別表 U,W 中的加法單位元素).

首先我們證明 S 為 $U \oplus W$ 的一組 spanning set. 對任意 $(\mathbf{u},\mathbf{w}) \in U \oplus W$, 由於 $\mathbf{u} \in U$ 且 $\{\mathbf{u}_1,\ldots,\mathbf{u}_m\}$ 為 U 的一組 basis, 故存在 $c_1,\ldots,c_m \in F$ 使得 $\mathbf{u} = c_1\mathbf{u}_1 + \cdots + c_m\mathbf{u}_m$. 同理存在 $d_1,\ldots,d_n \in F$ 使得 $\mathbf{w} = d_1\mathbf{w}_1 + \cdots + d_n\mathbf{w}_n$. 因此得

$$(\mathbf{u},\mathbf{w}) = c_1(\mathbf{u}_1,\mathbf{O}_W) + \cdots + c_m(\mathbf{u}_m,\mathbf{O}_W) + d_1(\mathbf{O}_U,\mathbf{w}_1) + \cdots + d_n(\mathbf{O}_U,\mathbf{w}_n),$$

故得證 S 為 $U \oplus W$ 的一組 spanning set.

最後我們要證明 S 為 linearly independent. 用反證法, 假設 S 為 linearly dependent, 由 Proposition 1.4.2 知存在 $c_1,\ldots,c_m,d_1,\ldots,d_n$ 不全為 0 使得

$$(\mathbf{O}_U,\mathbf{O}_W) = c_1(\mathbf{u}_1,\mathbf{O}_W) + \cdots + c_m(\mathbf{u}_m,\mathbf{O}_W) + d_1(\mathbf{O}_U,\mathbf{w}_1) + \cdots + d_n(\mathbf{O}_U,\mathbf{w}_n),$$

依 $U \oplus W$ 中元素相等的定義此即 $\mathbf{O}_U = c_1\mathbf{u}_1 + \cdots + c_m\mathbf{u}_m$ 且 $\mathbf{O}_W = d_1\mathbf{w}_1 + \cdots + d_n\mathbf{w}_n$. 由於 $\{\mathbf{u}_1, \ldots, \mathbf{u}_m\}$ 和 $\{\mathbf{w}_1, \ldots, \mathbf{w}_n\}$ 皆為 linearly independent 可得 $c_1, \ldots, c_m$ 和 $d_1, \ldots, d_n$ 皆為 0, 此與之前假設相矛盾. 故得證 S 為 linearly independent. □

最後我們要強調要 over 相同的 field 的 vector spaces 才可以談論其 direct sum. 另外我們可以將兩個 F-spaces 的 direct sum 的定義推廣到任意有限多個 F-spaces 的 direct sum.

Question 1.25. *假設 $U_1, \ldots, U_n$ 為 F-spaces, 你認為 $U_1 \oplus \cdots \oplus U_n$ 的定義應為何? 又若 $U_1, \ldots, U_n$ 皆為 finite dimensional F-spaces, 則 $\dim(U_1 \oplus \cdots \oplus U_n)$ 為何?*

1.6.2. Quotient Space. 給定 vector space V 及其 subspace W, 我們可以利用 W 在 V 中定義一個 equivalent relation, 其定義為對任意 $\mathbf{v}_1, \mathbf{v}_2 \in V$, $\mathbf{v}_1 \sim \mathbf{v}_2$ 若且唯若 $\mathbf{v}_1 - \mathbf{v}_2 \in W$. 我們來說明一下這是一個 equivalent relation.

(1) 對任意 $\mathbf{v} \in V$ 皆有 $\mathbf{v} \sim \mathbf{v}$: 這是因為 $\mathbf{O} \in W$, 故 $\mathbf{v} - \mathbf{v} \in W$.
(2) 若 $\mathbf{v}_1 \sim \mathbf{v}_2$, 則 $\mathbf{v}_2 \sim \mathbf{v}_1$: 這是因為 $\mathbf{v}_1 \sim \mathbf{v}_2$ 表示 $\mathbf{v}_1 - \mathbf{v}_2 \in W$, 故由 W 為 vector space 知 $\mathbf{v}_2 - \mathbf{v}_1 = -(\mathbf{v}_1 - \mathbf{v}_2) \in W$, 亦即 $\mathbf{v}_2 \sim \mathbf{v}_1$.
(3) 若 $\mathbf{v}_1 \sim \mathbf{v}_2$ 且 $\mathbf{v}_2 \sim \mathbf{v}_3$, 則 $\mathbf{v}_1 \sim \mathbf{v}_3$: 這是因為由 $\mathbf{v}_1 - \mathbf{v}_2 \in W$ 以及 $\mathbf{v}_2 - \mathbf{v}_3 \in W$, 可得 $\mathbf{v}_1 - \mathbf{v}_3 = (\mathbf{v}_1 - \mathbf{v}_2) + (\mathbf{v}_2 - \mathbf{v}_3) \in W$.

由於這一個 equivalent relation, 我們定義一個新的集合

$$V/W = \{\overline{\mathbf{v}} \mid v \in V\}.$$

當然了我們要說明 V/W 上的元素怎樣才會相等, 在這裡我們要求 $\overline{\mathbf{u}} = \overline{\mathbf{v}}$ 若且唯若 $\mathbf{u} \sim \mathbf{v}$ (亦即 $\mathbf{u} - \mathbf{v} \in W$).

Question 1.26. *前面為什麼要去說明 $\sim$ 是一個 equivalent relation 才能定義 V/W 呢?*

若學過代數 group 的同學可以看出因為 V 在加法上是一個 abelian group, 而 W 為 V 的 (normal) subgroup, 所以我們可以定出 V/W 上的運算使其成為一個 abelian group. 事實上我們還可以定出 F 對 V/W 上元素的作用使得 V/W 為一個 vector space over F. 其定義為對任意的 $\overline{\mathbf{u}}, \overline{\mathbf{v}} \in V/W$ 且 $r \in F$, 我們定

$$\begin{aligned} \overline{\mathbf{u}} + \overline{\mathbf{v}} &= \overline{\mathbf{u} + \mathbf{v}} \\ r\overline{\mathbf{v}} &= \overline{r\mathbf{v}}. \end{aligned}$$

因 W 為 V 的 subspace, 很容易驗證這個 V/W 的運算與作用皆為 well-defined, 而且我們可得 V/W 為一個 vector space over F, 稱之為 the *quotient space* of V *modulo* W.

Question 1.27. *請驗證這個運算是 well-defined 且 V/W 是一個 vector space over F. 什麼是 V/W 上的加法單位元素呢?*

Question 1.28. *若 U 為 V 的 subspace 且 $W \subseteq U$, 則 U/W 為 V/W 的 subspace 嗎? $W \subseteq U$ 這個假設是需要的嗎?*

若 V,W 為 finite dimensional F-spaces, 我們自然會問是否 V/W 亦為 finite dimensional F-space, 且其 dimension 為何?

Proposition 1.6.2. 假設 V 為 *finite dimensional F-spaces* 且 W 為 V 的一個 *F-subspace,* 則 V/W 亦為 *finite dimensional F-space,* 且

$$\dim(V/W) = \dim(V) - \dim(W).$$

Proof. 由 Theorem 1.5.8 我們知 W 亦為 finite dimensional F-space. 設 $\{\mathbf{w}_1,\ldots,\mathbf{w}_m\}$ 為 W 的一組 basis, 由 Theorem 1.5.9 知存在 $\mathbf{v}_1,\ldots,\mathbf{v}_n \in V$ 使得 $\{\mathbf{w}_1,\ldots,\mathbf{w}_m,\mathbf{v}_1,\ldots,\mathbf{v}_n\}$ 為 V 的一組 basis. 我們要證明 $S = \{\overline{\mathbf{v}_1},\ldots,\overline{\mathbf{v}_n}\}$ 為 V/W 的一組 basis.

首先我們證明 S 為 V/W 的一組 spanning set. 對任意 $\overline{\mathbf{v}} \in V/W$, 由於 $\mathbf{v} \in V$ 且 $\{\mathbf{w}_1,\ldots,\mathbf{w}_m,\mathbf{v}_1,\ldots,\mathbf{v}_n\}$ 為 V 的一組 basis, 故存在 $c_1,\ldots,c_m,d_1,\ldots,d_n \in F$ 使得

$$\mathbf{v} = c_1\mathbf{w}_1 + \cdots + c_m\mathbf{w}_m + d_1\mathbf{v}_1 + \cdots + d_n\mathbf{v}_n.$$

因此依定義得

$$\overline{\mathbf{v}} = c_1\overline{\mathbf{w}_1} + \cdots + c_m\overline{\mathbf{w}_m} + d_1\overline{\mathbf{v}_1} + \cdots + d_n\overline{\mathbf{v}_n}.$$

然而對所有 $\overline{\mathbf{w}_i}$ 因 $\mathbf{w}_i \in W$, 我們有 $\overline{\mathbf{w}_i} = \overline{\mathbf{O}}$, 因此

$$\overline{\mathbf{v}} = d_1\overline{\mathbf{v}_1} + \cdots + d_n\overline{\mathbf{v}_n}.$$

故得證 S 為 V/W 的一組 spanning set.

最後我們要證明 S 為 linearly independent. 用反證法, 假設 S 為 linearly dependent, 由 Proposition 1.4.2 知存在 $d_1,\ldots,d_n \in F$ 不全為 0 使得 $\overline{\mathbf{O}} = d_1\overline{\mathbf{v}_1} + \cdots + d_n\overline{\mathbf{v}_n}$. 依定義此即 $d_1\mathbf{v}_1 + \cdots + d_n\mathbf{v}_n \in W = \mathrm{Span}(\{\mathbf{w}_1,\ldots,\mathbf{w}_m\})$, 故知

$$d_1\mathbf{v}_1 + \cdots + d_n\mathbf{v}_n \in \mathrm{Span}(\{\mathbf{v}_1,\ldots,\mathbf{v}_n\}) \cap \mathrm{Span}(\{\mathbf{w}_1,\ldots,\mathbf{w}_m\}).$$

然而 S 為 linearly independent, 由 Corollary 1.4.4 知

$$\mathrm{Span}(\{\mathbf{v}_1,\ldots,\mathbf{v}_n\}) \cap \mathrm{Span}(\{\mathbf{w}_1,\ldots,\mathbf{w}_m\}) = \{\mathbf{O}\},$$

故得 $d_1\mathbf{v}_1 + \cdots + d_n\mathbf{v}_n = \mathbf{O}$. 由於 $d_1,\ldots,d_n$ 不全為 0, 此與 $\{\mathbf{v}_1,\ldots,\mathbf{v}_n\}$ 為 linearly independent 相矛盾. 故得證 S 為 linearly independent. $\square$

Linear Transformations

在學習數學的過程中大家應該體會到函數的重要性. 在不同課程中我們討論的函數對象都不同, 例如在微積分中我們有興趣於連續函數、可微函數; 而在群與環中我們有興趣於 group homomorphisms 及 ring homomorphisms. 在線性代數中我們有興趣的函數是希望能保持 vector spaces 中的運算與作用, 也就是所謂的 linear transformations.

2.1. Definition and Basic Properties

Definition 2.1.1. 設 V,W 皆為 over F 的 vector spaces. 給定一個從 V 到 W 的函數 $T:V\to W$, 若對所有 $\mathbf{v}_1,\mathbf{v}_2\in V$ 以及 $r\in F$ 皆有 $T(r\mathbf{v}_1+\mathbf{v}_2)=rT(\mathbf{v}_1)+T(\mathbf{v}_2)$, 則稱 T 為 *linear transformation* (或 *linear mapping*) from V to W.

有時候我們會簡稱為 T is F-linear. 另外我們用 $\mathscr{L}(V,W)$ 表示所有從 V 到 W 的 linear transformations 所成之集合.

Question 2.1. 你看得出來 T *is* F*-linear* 等價於對所有 $\mathbf{v},\mathbf{v}'\in V$ 以及 $r\in F$ 皆有 $T(\mathbf{v}+\mathbf{v}')=T(\mathbf{v})+T(\mathbf{v}')$ 以及 $T(r\mathbf{v})=rT(\mathbf{v})$ 嗎?

接著我們將介紹一些有關於 linear transformation 的基本性質, 由於 linear transformation 可能牽涉到不同 vector spaces, 我們用 $\mathbf{O}_V$ 來表示 V 的加法單位元素.

Proposition 2.1.2. 若 $T:V\to W$ 為一個 *linear transformation,* 則

(1) $T(\mathbf{O}_V)=\mathbf{O}_W$

(2) 對所有 $\mathbf{v}\in V$ 皆有 $T(-\mathbf{v})=-T(\mathbf{v})$.

Proof.

(1) 由 $T(\mathbf{O}_V)=T(\mathbf{O}_V+\mathbf{O}_V)=T(\mathbf{O}_V)+T(\mathbf{O}_V)$, 可得 $T(\mathbf{O}_V)=\mathbf{O}_W$.

(2) 若 $\mathbf{v}\in V$, 則由 $\mathbf{O}_W=T(\mathbf{v}+(-\mathbf{v}))=T(\mathbf{v})+T(-\mathbf{v})$ 得證 $T(-\mathbf{v})=-T(\mathbf{v})$.

□

我們可以利用一些 linear transformations 創造新的 linear transformation. 若 T, T' 皆為 V 到 W 的 linear transformations, 我們定義一個新的 V 到 W 的函數 $T+T': V \to W$ 為對任意 $\mathbf{v} \in V$, $(T+T')(\mathbf{v}) = T(\mathbf{v}) + T'(\mathbf{v})$. 給定 $r \in F$ ，我們也可定義一個新的 V 到 W 的函數 $rT: V \to W$ 為對任意 $\mathbf{v} \in V$, $(rT)(\mathbf{v}) = rT(\mathbf{v})$. 事實上這樣建構的新函數仍為 linear transformation.

Proposition 2.1.3. *若 T, T' 皆為 V 到 W 的 F-linear transformations 且 $r \in F$ ，則 $T+T'$ 以及 rT 皆為 V 到 W 的 F-linear transformations.*

Proof. 對於任意 $\mathbf{v}_1, \mathbf{v}_2 \in V$ 以及 $s \in F$, 皆有 $(T+T')(s\mathbf{v}_1 + \mathbf{v}_2) = T(s\mathbf{v}_1 1 + \mathbf{v}_2) + T'(s\mathbf{v}_1 + \mathbf{v}_2)$ 由於 T, T' 為 F-linear, 故有 $T(s\mathbf{v}_1 + \mathbf{v}_2) + T'(s\mathbf{v}_1 + \mathbf{v}_2) = sT(\mathbf{v}_1) + T(\mathbf{v}_2) + sT'(\mathbf{v}_1) + T'(\mathbf{v}_2) = s(T(\mathbf{v}_1) + T'(\mathbf{v}_1)) + (T(\mathbf{v}_2) + T'(\mathbf{v}_2))$. 亦即 $(T+T')(s\mathbf{v}_1 + \mathbf{v}_2) = s(T+T')(\mathbf{v}_1) + (T+T')(\mathbf{v}_2)$. 同理 $(rT)(s\mathbf{v}_1 + \mathbf{v}_2) = rT(s\mathbf{v}_1 + \mathbf{v}_2) = rsT(\mathbf{v}_1) + rT(\mathbf{v}_2) = s(rT(\mathbf{v}_1)) + rT(\mathbf{v}_2) = s(rT)(\mathbf{v}_1) + (rT)(\mathbf{v}_2)$. □

Question 2.2. *考慮所有從 V 到 W 的 linear transformations 所成之集合 $\mathscr{L}(V, W)$, Proposition 2.1.3 是不是告訴我們 $\mathscr{L}(V, W)$ 是一個 vector space over F?*

當一個函數的對應域剛好是另一個函數的定義域時, 我們可以將之合成為一個新的函數. 下一個 Proposition 告訴我們 linear transformations 的合成仍為 linear transformation.

Proposition 2.1.4. *若 $T_1: V \to W$, $T_2: W \to U$ 皆為 F-linear, 則 $T_2 \circ T_1: V \to U$ 亦為 F-linear.*

Proof. 對於任意 $\mathbf{v}, \mathbf{v}' \in V$ 以及 $r \in F$, 考慮 $T_2 \circ T_1(r\mathbf{v} + \mathbf{v}') = T_2(T_1(r\mathbf{v} + \mathbf{v}'))$. 因 T_1 為 F-linear, 故知 $T_1(r\mathbf{v} + \mathbf{v}') = rT_1(\mathbf{v}) + T_1(\mathbf{v}')$ 再由 T_2 為 F-linear 得 $T_2 \circ T_1(r\mathbf{v} + \mathbf{v}') = T+2(rT_1(\mathbf{v}) + T_1(\mathbf{v}')) = rT_2(T_1(\mathbf{v})) + T_2(T_1(\mathbf{v}')) = rT_2 \circ T_1(\mathbf{v}) + T_2 \circ T_1(\mathbf{v}')$. □

Question 2.3. *設 T, T' 皆為 V 到 W 的 F-linear transformations, T'' 為 W 到 U 的 F-linear transformation. 是否 $T'' \circ (T+T') = T'' \circ T + T'' \circ T'$? 又對任意 $r \in F$ 是否 $r(T'' \circ T) = (rT'') \circ T = T'' \circ (rT)$?*

其實給定兩個 F-spaces V, W, 我們很容易建構出一個從 V 到 W 的 linear transformation. 下一個 Theorem 說的是所有 V 到 W 的 linear transformations 我們都可以完全掌握.

Theorem 2.1.5. *假設 $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ 是 V 的一組 basis, 給定任意 $\mathbf{w}_1, \ldots, \mathbf{w}_n \in W$, 存在一個唯一的 F-linear transformation $T: V \to W$ 滿足 $T(\mathbf{v}_i) = \mathbf{w}_i$, $\forall i \in \{1, \ldots, n\}$.*

Proof. 證明存在性: 也就是說我們要找到一個 $T \in \mathscr{L}(V, W)$ 滿足 $T(\mathbf{v}_i) = \mathbf{w}_i$. 定義 $T: V \to W$, 滿足對所有 $\mathbf{v} = c_1\mathbf{v}_1 + \cdots + c_n\mathbf{v}_n \in V$, $T(\mathbf{v}) = c_1\mathbf{w}_1 + \cdots + c_n\mathbf{w}_n$. 由於 $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ 是 V 的一組 basis, T 是一個從 V 到 W 的 well-defined function. 又 T 滿足 $T(\mathbf{v}_i) = \mathbf{w}_i$, 所以我們僅剩下證明 T 為 F-linear. 對任意 $\mathbf{v} = \sum_{i=1}^n c_i\mathbf{v}_i$, $\mathbf{v}' = \sum_{i=1}^n d_i\mathbf{v}_i \in V$ 以及 $r \in F$, 我們有 $T(r\mathbf{v} + \mathbf{v}') =$

$T(\sum_{i=1}^{n}(rc_i+d_i)\mathbf{v}_i)=\sum_{i=1}^{n}(rc_i+d_i)\mathbf{w}_i$; 另一方面 $rT(\mathbf{v})+T(\mathbf{v}')=rT(\sum_{i=1}^{n}c_i\mathbf{v}_i)+T(\sum_{i=1}^{n}d_i\mathbf{v}_i)=r\sum_{i=1}^{n}c_i\mathbf{w}_i+\sum_{i=1}^{n}d_i\mathbf{w}_i$, 利用 vector space 的運算性質, 我們得 $T(r\mathbf{v}+\mathbf{v}')=rT(\mathbf{v})+T(\mathbf{v}')$.

證明唯一性: 我們要證明若 $T':V\to W$ 亦為 F-linear 且滿足 $T'(\mathbf{v}_i)=\mathbf{w}_i$, $\forall i\in\{1,\ldots,n\}$, 則 $T=T'$. 亦即證明 $T(\mathbf{v})=T'(\mathbf{v}),\forall\mathbf{v}\in V$. 然而對任意 $\mathbf{v}=\sum_{i=1}^{n}c_i\mathbf{v}_i\in V$, 依 T 的定義 $T(\mathbf{v})=\sum_{i=1}^{n}c_i\mathbf{w}_i$, 而依 T' 是 F-linear 可得 $T'(\mathbf{v})$ 亦為 $\sum_{i=1}^{n}c_iT(\mathbf{v}_i)=\sum_{i=1}^{n}c_i\mathbf{w}_i$, 故得證 $T=T'$. □

Theorem 2.1.5 告訴我們, 給定一個 linear transformation $T:V\to W$, 若能找到一組 basis S 讓我們知道對所有的 $\mathbf{u}\in S$, $T(\mathbf{u})$ 為何, 則對所有 $\mathbf{v}\in V$, 便可知 $T(\mathbf{v})$ 為何了! 例如 $T_\theta:\mathbb{R}^2\to\mathbb{R}^2$ 為將 $\mathbb{R}^2$ 上任一組向量 (x,y) 以原點 $(0,0)$ 為圓心逆時針繞 θ 角所得的向量. $T_\theta((x,y))$ 是甚麼向量呢? 由於將向量 $(1,0)=(\cos 0,\sin 0)$ 以 $(0,0)$ 為圓心繞 θ 角後依定義所得向量為 $(\cos\theta,\sin\theta)$, 所以我們得到 $T_\theta((1,0))=(\cos\theta,\sin\theta)$. 同理 $(0,1)=(\cos(\pi/2),\sin(\pi/2))$, 故得 $T_\theta((0,1))=(\cos((\pi/2)+\theta),\sin((\pi/2)+\theta))=(-\sin\theta,\cos\theta)$. 故知 $T_\theta((x,y))=T_\theta(x(1,0)+y(0,1))=xT_\theta((1,0))+yT_\theta((0,1))=x(\cos\theta,\sin\theta)+y(-\sin\theta,\cos\theta)=(x\cos\theta-y\sin\theta,x\sin\theta+y\cos\theta)$. 不過要注意, 這個方法僅對 linear transformation 才成立, 所以要用這個方法處理函數的取值, 須先檢驗這個函數是 linear transformation 才行. 也就是說在上面的例子, 我們要先知道 T_θ 是 linear transformation (請自行驗證), 才可利用此法得到 $T_\theta((x,y))=(x\cos\theta-y\sin\theta,x\sin\theta+y\cos\theta)$.

Question 2.4. 若 $T:\mathbb{R}^2\to\mathbb{R}^2$ 是一個 *linear transformation,* 滿足 $T((1,2))=(2,1)$, $T((2,4))=(4,2)$ 是否可得 $T((x,y))=(y,x)$? 又若 $T':\mathbb{R}^2\to\mathbb{R}^2$ 是一個 *linear transformation,* 滿足 $T'((1,2))=(2,1)$, $T'((2,1))=(1,2)$ 是否可得 $T'((x,y))=(y,x)$?

2.2. Image and Kernel

Linear transformation 既然保持了 vector spaces 的運算, 可以理解它應該也會“保持” subspaces. 首先我們定義一些符號, 給定一函數 $f:S_1\to S_2$. 若 $S_1'\subseteq S_1$, 我們定

$$f(S_1')=\{f(s)\mid s\in S_1'\}.$$

注意 $f(S_1')$ 會是 S_2 的一個 subset, 稱之為 the *image* of S_1' under f; 另一方面若 $S_2'\subseteq S_2$, 令

$$f^{-1}(S_2')=\{s\in S_1\mid f(s)\in S_2'\}.$$

注意 $f^{-1}(S_2')$ 會是 S_1 的一個 subset, 稱之為 the *preimage* of S_2' under f.

Question 2.5. *Image* 和 *preimage* 是否為 *inclusion-preserving?* 也就是說一個函數 $f:S_1\to S_2$, 若 $S_1''\subseteq S_1'\subseteq S_1$ 是否可得 $f(S_1'')\subseteq f(S_1')$? 若 $S_2''\subseteq S_2'\subseteq S_2$, 是否可得 $f^{-1}(S_2'')\subseteq f^{-1}(S_2')$?

Question 2.6. 假設 $f:S_1\to S_2$ 為一個函數, 且 $S_1',S_1''\subseteq S_1$ 以及 $S_2',S_2''\subseteq S_2$. 下列哪些是正確的?

(1) $f(S_1'\cap S_1'')=f(S_1')\cap f(S_1'')$.

(2) $f(S_1'\cup S_1'')=f(S_1')\cup f(S_1'')$.

(3) $f^{-1}(S_2'\cap S_2'')=f^{-1}(S_2')\cap f^{-1}(S_2'')$.

(4) $f^{-1}(S_2' \cup S_2'') = f^{-1}(S_2') \cup f^{-1}(S_2'')$.

下面我們便是說 linear transformation 確實保持 subspace 的性質.

Lemma 2.2.1. 設 $T : V \to W$ 為一個 *linear transformation.*

(1) 若 V' 為 V 的 *subspace,* 則 $T(V')$ 為 W 的 *subspace.*

(2) 若 W' 為 W 的 *subspace,* 則 $T^{-1}(W')$ 為 V 的 *subspace.*

Proof. 依定義已知 $T(V') \subseteq W$ 且 $T^{-1}(W') \subseteq V$, 所以我們可利用 Proposition 1.2.1 來證明.

(1) 首先 $\mathbf{O}_V \in V'$ (因 V' 是 subspace), 故由 Proposition 2.1.2 (1) 知 $\mathbf{O}_W = T(\mathbf{O}_V) \in T(V')$. 再來對所有的 $\mathbf{w}_1, \mathbf{w}_2 \in T(V')$ 及 $r, s \in F$, 依定義存在 $\mathbf{v}_1, \mathbf{v}_2 \in V'$ 使得 $\mathbf{w}_1 = T(\mathbf{v}_1), \mathbf{w}_2 = T(\mathbf{v}_2)$. 故考慮 $\mathbf{v} = r\mathbf{v}_1 + s\mathbf{v}_2 \in V'$, 可得 $r\mathbf{w}_1 + s\mathbf{w}_2 = T(\mathbf{v}) \in T(V')$, 得證 $T(V')$ 是 W 的 subspace.

(2) 因 $\mathbf{O}_W \in W'$ 故由 $T(\mathbf{O}_V) = \mathbf{O}_W \in W'$ 得 $\mathbf{O}_V \in T^{-1}(W')$. 再來對所有的 $\mathbf{v}_1, \mathbf{v}_2 \in T^{-1}(W')$ 及 $r, s \in F$, 依定義 $T(\mathbf{v}_1) \in W'$ 且 $T(\mathbf{v}_2) \in W'$. 故由 W' 是 W 的 subspace 得 $T(r\mathbf{v}_1 + s\mathbf{v}_2) = rT(\mathbf{v}_1) + sT(\mathbf{v}_2) \in W'$, 亦即 $r\mathbf{v}_1 + s\mathbf{v}_2 \in T^{-1}(W')$, 得證 $T^{-1}(W')$ 是 V 的 subspace.

□

特別的, 我們對 $V' = V$ 以及 $W' = \{\mathbf{O}_W\}$ 的情形有興趣, 因此特別給予以下定義.

Definition 2.2.2. 設 $T : V \to W$ 為一個 linear transformation.

(1) $T(V)$ 稱為 the *image* (or *range*) of T , 我們用 $\mathrm{Im}(T)$ 來表示.

(2) $T^{-1}(\{\mathbf{O}_W\})$ 稱為 the *kernel* (or *null-space*) of T, 我們用 $\mathrm{Ker}(T)$ 來表示.

由 Lemma 2.2.1 知 $\mathrm{Im}(T)$ 是 W 的 subspace, 而 $\mathrm{Ker}(T)$ 為 V 的 subspace.

Question 2.7. 由 *image* 和 *preimage* 為 *inclusion-preserving* 我們知 $\mathrm{Im}(T) = T(V)$ 是所有 *subspaces* 的 *image* 中最大的 *subspace,* 而 $\mathrm{Ker}(T) = T^{-1}(\{\mathbf{O}_W\})$ 是所有 *subspaces* 的 *preimage* 中最小的. 為何不去探討 $T(\{\mathbf{O}_V\})$ 以及 $T^{-1}(W)$ 呢?

給定一個函數, 我們有興趣的是這個函數是否為映成 (onto) 或是一對一 (one-to-one). $\mathrm{Im}(T)$ 和 $\mathrm{Ker}(T)$ 之所以重要是因為在 T 為 linear transformation 的情形, 我們可以用 $\mathrm{Im}(T)$ 和 $\mathrm{Ker}(T)$ 來判斷 T 是否為映成或一對一, 事實上我們有以下之結果.

Proposition 2.2.3. 假設 $T : V \to W$ 是一個 *linear transformation.*

(1) T 為映成若且唯若 $\mathrm{Im}(T) = W$.

(2) T 為一對一若且唯若 $\mathrm{Ker}(T) = \{\mathbf{O}_V\}$.

Proof. 我們已知 $\mathrm{Im}(T) \subseteq W$ 以及 $\{\mathbf{O}_V\} \subseteq \mathrm{Ker}(T)$, 所以事實上 (1) 我們僅需考慮 $W \subseteq \mathrm{Im}(T)$ 的部分, 同理 (2) 我們僅需考慮 $\mathrm{Ker}(T) \subseteq \{\mathbf{O}_V\}$ 的部分.

(1) T 為 onto, 表示對所有 $\mathbf{w} \in W$, 皆存在 $\mathbf{v} \in V$ 使得 $T(\mathbf{v}) = \mathbf{w}$, 亦即 $\mathbf{w} \in T(V) = \mathrm{Im}(T)$. 得證 $W \subseteq \mathrm{Im}(T)$, 故得 $W = \mathrm{Im}(T)$. 反之, 由 $W \subseteq \mathrm{Im}(T)$ 可知每一個 $\mathbf{w} \in W$ 皆在 $\mathrm{Im}(T)$ 中, 亦即存在 $\mathbf{v} \in V$ 使得 $\mathbf{w} = T(\mathbf{v})$, 故知 T 為 onto.

(2) T 為 one-to-one, 表示 V 中唯一滿足 $T(\mathbf{v}) = \mathbf{O}_W$ 的 $\mathbf{v}$ 應為 $\mathbf{O}_V$ (因已知 $T(\mathbf{O}_V) = \mathbf{O}_W$). 故若 $\mathbf{v} \in \mathrm{Ker}(T)$, 表示 $T(\mathbf{v}) = \mathbf{O}_W$, 得 $\mathbf{v} = \mathbf{O}_V$. 證得 $\mathrm{Ker}(T) = \{\mathbf{O}_V\}$. 反之, 假設 $\mathrm{Ker}(T) = \{\mathbf{O}_V\}$. 若 $\mathbf{v}_1, \mathbf{v}_2 \in V$ 滿足 $T(\mathbf{v}_1) = T(\mathbf{v}_2)$, 則由 T 為 linear 得 $T(\mathbf{v}_1 - \mathbf{v}_2) = \mathbf{O}_W$, 亦即 $\mathbf{v}_1 - \mathbf{v}_2 \in \mathrm{Ker}(T) = \{\mathbf{O}_V\}$. 故得 $\mathbf{v}_1 = \mathbf{v}_2$, 得證 T 為 one-to-one.

□

從 Proposition 2.2.3 的證明可以看出, 並不需用到 T 是 linear 的假設來證明 T 是 onto 和 $\mathrm{Im}(T) = W$ 為等價的; 不過 T 是 one-to-one 和 $\mathrm{Ker}(T) = \{\mathbf{O}_V\}$ 為等價的就需要 T 為 linear 的假設了. 所以對一般函數 f, 除非先知道 f 是 linear 不能用 $f^{-1}(\{0\}) = \{0\}$ 來判斷 f 是否為 one-to-one.

既然 image 和 kernel 這麼重要, 我們當然要去了解它們. 由下一個 Lemma 我們了解到 image 和 spanning set 有關而 kernel 就和 linear independency 相關.

Lemma 2.2.4. 假設 $T: V \to W$ 是一個 *linear transformation,* 且 S, S' 為 V 的 *subsets.*

(1) 若 S 是 V 的 *spanning set,* 則 $T(S)$ 是 $\mathrm{Im}(T)$ 的 *spanning set.*

(2) 若 S' 為 *linearly independent* 且 $\mathrm{Span}(S') \cap \mathrm{Ker}(T) = \{\mathbf{O}_V\}$, 則 $T(S')$ 亦為 *linearly independent.*

Proof.

(1) 依定義若 $\mathbf{w} \in \mathrm{Im}(T)$ 表示存在 $v \in V$ 使得 $\mathbf{w} = T(\mathbf{v})$, 但 $V = \mathrm{Span}(S)$ 故存在 $c_1, \ldots, c_n \in F$ 以及 $\mathbf{v}_1, \ldots, \mathbf{v}_n \in S$ 使得 $\mathbf{v} = c_1\mathbf{v}_1 + \cdots + c_n\mathbf{v}_n$. 依定義 $T(\mathbf{v}_1), \ldots, T(\mathbf{v}_n) \in T(S)$, 因此由 T 為 linear 知 $\mathbf{w} = T(c_1\mathbf{v}_1 + \cdots + c_n\mathbf{v}_n) = c_1T(\mathbf{v}_1) + \cdots + c_nT(\mathbf{v}_n) \in \mathrm{Span}(T(S))$. 故得 $\mathrm{Im}(T) \subseteq \mathrm{Span}(T(S))$. 另一方面, 若 $\mathbf{w} \in \mathrm{Span}(T(S))$, 表示存在 $c_1, \ldots, c_n \in F$ 以及 $\mathbf{w}_1, \ldots, \mathbf{w}_n \in T(S)$ 使得 $\mathbf{w} = \sum_{i=1}^{n} c_i\mathbf{w}_i$. 但 $\mathbf{w}_i \in T(S)$ 依定義表示存在 $\mathbf{v}_i \in S$ 使得 $\mathbf{w}_i = T(\mathbf{v}_i)$, 故知 $\mathbf{w} = \sum_{i=1}^{n} c_iT(\mathbf{v}_i) = \sum_{i=1}^{n} T(c_i\mathbf{v}_i) \in T(\mathrm{Span}(S)) = T(V) = \mathrm{Im}(T)$, 得證 $\mathrm{Span}(T(S)) \subseteq \mathrm{Im}(T)$.

(2) 首先我們說明依假設 $T(S')$ 中的元素皆相異. 否則若存在 $\mathbf{v}, \mathbf{v}' \in S'$ 且 $\mathbf{v}, \mathbf{v}'$ 相異使得 $T(\mathbf{v}) = T(\mathbf{v}')$, 可得 $T(\mathbf{v} - \mathbf{v}') = \mathbf{O}_W$, 亦即 $\mathbf{v} - \mathbf{v}' \in \mathrm{Ker}(T)$. 然而 $\mathbf{v} - \mathbf{v}' \in \mathrm{Span}(S')$, 故由假設知 $\mathbf{v} - \mathbf{v}' \in \mathrm{Span}(S') \cap \mathrm{Ker}(T) = \{\mathbf{O}_V\}$ 推得 $\mathbf{v} = \mathbf{v}'$ 之矛盾. 故知 $T(S')$ 中的元素必相異.

現利用反證法設 $T(S')$ 為 linearly dependent, 表示存在 $\mathbf{v}_1, \ldots, \mathbf{v}_n \in S'$ 皆相異且 $c_1, ..., c_n \in F$ 皆不為 0 使得 $c_1T(\mathbf{v}_1) + \cdots + c_nT(\mathbf{v}_n) = \mathbf{O}_W$. 利用 T 為 linear 得 $T(c_1\mathbf{v}_1 + \cdots + c_n\mathbf{v}_n) = \mathbf{O}_W$, 亦即 $c_1\mathbf{v}_1 + \cdots + c_n\mathbf{v}_n \in \mathrm{Ker}(T)$. 然而 $c_1\mathbf{v}_1 + \cdots + c_n\mathbf{v}_n \in \mathrm{Span}(S')$, 故由 $\mathrm{Span}(S') \cap \mathrm{Ker}(T) = \{\mathbf{O}_V\}$, 得 $c_1\mathbf{v}_1 + \cdots + c_n\mathbf{v}_n = \mathbf{O}_V$. 此與 S' 為 linearly independent 相矛盾, 故得 $T(S')$ 為 linearly independent.

□

Lemma 2.2.4 (1) 告訴我們當 $T : V \to W$ 是一個 linear transformation, 若 V 為 finite dimensional F-space, 則 $\mathrm{Im}(T)$ 亦為 finite dimensional F-space.

Question 2.8. 假設 $T : V \to W$ 是一個 *linear transformation.* 若 V 為 *finite dimensional F-space,* 可否得 $\dim(\mathrm{Im}(T)) \le \dim(V)$? 又若已知 $\dim(W) > \dim(V)$ 是否可確定 T 是否為映成的?

以下我們將利用 V 的一組 basis 得到 $\mathrm{Im}(T)$ 的一組 basis.

Theorem 2.2.5. 假設 $T : V \to W$ 是一個 *linear transformation.* 若 S_0 為 $\mathrm{Ker}(T)$ 的一組 *basis* 且 $S_0 \cup S$ 為 V 的一組 *basis,* 則 $T(S)$ 為 $\mathrm{Im}(T)$ 的一組 *basis.*

Proof. 證明 $T(S)$ 為 $\mathrm{Im}(T)$ 的 spanning set: 因 $S_0 \cup S$ 為 V 的 spanning set, 由 Lemma 2.2.4 (1) 知 $T(S_0 \cup S)$ 為 $\mathrm{Im}(T)$ 的 spanning set. 由於 $S_0 \in \mathrm{Ker}(T)$ 故得 $T(S_0) = \{\mathbf{O}_W\}$. 但由於 $T(S_0 \cup S) = T(S_0) \cup T(S) = \{\mathbf{O}_W\} \cup T(S)$, 故由 $\{\mathbf{O}_W\} \subseteq \mathrm{Span}(T(S)$ (利用 Corollary 1.3.4) 知 $\mathrm{Span}(T(S)) = \mathrm{Span}(\{\mathbf{O}_W\} \cup T(S)) = \mathrm{Span}(T(S_0 \cup S)) = \mathrm{Im}(T)$.

證明 $T(S)$ 為 linearly independent: 因 $S_0 \cup S$ 為 linearly independent 且 S_0 為 linearly independent, 由 Corollary 1.4.4 知 S 為 linearly independent 且 $\mathrm{Span}(S) \cap \mathrm{Ker}(T) = \mathrm{Span}(S) \cap \mathrm{Span}(S_0) = \{\mathbf{O}_V\}$. 故由 Lemma 2.2.4 (2) 知 $T(S)$ 為 linearly independent. □

特別的, 當 V 為 finite dimensional vector space, 我們可計算 V, $\mathrm{Ker}(T)$ 以及 $\mathrm{Im}(T)$ 之間 dimension 的關係.

Corollary 2.2.6 (Dimension Theorem)**.** 若 V 為一個 *finite dimensional F-space* 且 $T : V \to W$ 是一個 *linear transformation,* 則

$$\dim(V) = \dim(\mathrm{Ker}(T)) + \dim(\mathrm{Im}(T)).$$

Proof. 首先利用 Theorem 1.5.8 找到 $\mathrm{Ker}(T)$ 的一組 basis $S_0 = \{\mathbf{v}_1, \ldots, \mathbf{v}_m\}$, 再利用 Theorem 1.5.9 找到 $S = \{\mathbf{v}_{m+1}, \ldots, \mathbf{v}_n\}$ 使得 $S_0 \cup S = \{\mathbf{v}_1, \ldots, \mathbf{v}_m, \mathbf{v}_{m+1}, \ldots, \mathbf{v}_n\}$ 為 V 的一組 basis. Theorem 2.2.5 告訴我們 $\{T(\mathbf{v}_{m+1}), \ldots, T(\mathbf{v}_n)\}$ 為 $\mathrm{Im}(T)$ 的一組 basis, 故知 $n - m = \dim(\mathrm{Im}(T))$, 得證 $\dim(V) = n = m + (n-m) = \dim(\mathrm{Ker}(T)) + \dim(\mathrm{Im}(T))$. □

當 V 為 finite dimensional vector space 且 $T : V \to W$ 是一個 linear transformation, $\dim(\mathrm{Im}(T))$ 一般也稱為 the *rank* of T, 而 $\dim(\mathrm{Ker}(T))$ 也稱為 the *nullity* of T. 所以 Dimension Theorem 也有人稱為 Rank Theorem: rank of T + nullity of T = $\dim(V)$.

Question 2.9. 假設 $T : V \to W$ 是一個 *linear transformation.* 若 V 為 *finite dimensional F-space,* 若已知 $\dim(W) < \dim(V)$ 是否可確定 T 是否為一對一的?

2.3. Isomorphism

Linear transformation 將兩個 vector spaces 關連起來. 如果兩個 vector spaces 間可找到一個一對一且映成的 linear transformation, 我們便認為這兩個 vector spaces 有相同的結構, 稱它們為 isomorphic. 這一節中我們主要便是要探討幾個有關 isomorphism 的重要性質.

Definition 2.3.1. 假設 $T:V\to W$ 是一個 linear transformation. 若 T 為 one-to-one and onto, 則稱 T 為一個 *isomorphism.* 此時我們稱 V 和 W 為 *isomorphic* 且用 $V\simeq W$ 來表示.

由 Proposition 2.2.3 我們知 $T:V\to W$ 為 isomorphism 若且唯若 $\mathrm{Im}(T)=W$ and $\mathrm{Ker}(T)=\{\mathbf{O}_V\}$. 所以當 T 為 isomorphism 時, 若 S 為 V 的一組 basis, 由 T 為 one-to-one 知 $T(S)$ 中的元素皆相異 (即若 $\mathbf{v},\mathbf{v}'\in S$ 且 $\mathbf{v}\neq\mathbf{v}'$, 則 $T(\mathbf{v})\neq T(\mathbf{v}')$), 再由 Lemma 2.2.4 知 $T(S)$ 亦為 W 的一組 basis. 反之, 若 $T(S)$ 中的元素皆相異且 $T(S)$ 為 W 的一組 basis, 由 $\mathrm{Span}(T(S))=W$, 得 $\mathrm{Im}(T)=W$. 另一方面, 若 $\mathbf{v}\in\mathrm{Ker}(T)$, 由於 S 為 V 的一組 basis, 存在 $\mathbf{v}_1,\dots,\mathbf{v}_n\in S$ 皆相異, $c_1,\dots,c_n\in F$ 使得 $\mathbf{v}=c_1\mathbf{v}_1+\cdots+c_n\mathbf{v}_n$. 由此得 $\mathbf{O}_W=T(c_1\mathbf{v}_1+\cdots+c_n\mathbf{v}_n)=c_1T(\mathbf{v}_1)+\cdots+c_n(T\mathbf{v}_n)$. 但 $T(S)$ 為一組 basis 且 $T(\mathbf{v}_1),\dots,T(\mathbf{v}_n)\in T(S)$ 皆相異得 $c_1,\dots,c_n$ 皆為 0, 故得 $\mathbf{v}=\mathbf{O}_V$, 即 $\mathrm{Ker}(T)=\{\mathbf{O}_V\}$. 綜合以上, 我們有以下之結論.

Proposition 2.3.2. 假設 $T:V\to W$ 是一個 *linear transformation* 且 S 為 V 的一組 *basis.* 則 T 是 *isomorphism* 若且唯若 $T(S)$ 中的元素皆相異且 $T(S)$ 為 W 的一組 *basis.*

Question 2.10. *Proposition 2.3.2* 中為何要強調 $T(S)$ 中元素皆相異?

當 V 為 finite dimensional vector space 時, 我們有以下很好之結果.

Corollary 2.3.3. 假設 V,W 皆為 F*-spaces* 且 V 為 *finite dimensional* F*-space.* 則 $V\simeq W$ 若且唯若 $\dim(V)=\dim(W)$.

Proof. $V\simeq W$ 表示存在一個 isomorphism $T:V\to W$, 因此由 Proposition 2.3.2 知 W 亦為 finite dimensional F-space 且 $\dim(W)=\dim(V)$. 反之, 若 W 為 finite dimensional F-space 且 $\dim(W)=\dim(V)=n$, 可任取 V 的一組 basis $\{\mathbf{v}_1,\dots,\mathbf{v}_n\}$ 以及 W 的一組 basis $\{\mathbf{w}_1,\dots,\mathbf{w}_n\}$, 利用 Theorem 2.1.5 知存在一個 linear transformation $T:V\to W$ 滿足 $T(\mathbf{v}_i)=\mathbf{w}_i$, $\forall i\in\{1,\dots,n\}$. 由於 $\{T(\mathbf{v}_1)\dots,T(\mathbf{v}_n)\}=\{\mathbf{w}_1,\dots,\mathbf{w}_n\}$ 為 W 的一組 basis, 由 Proposition 2.3.2 知 T 為 isomorphism, 得證 $V\simeq W$. □

Corollary 2.3.3 告訴我們考慮 finite dimensional vector spaces 間是否 isomorphic 是個簡單的問題, 只要看它們的 dimension 是否相同即可. 底下我們介紹幾個 isomorphism 的定理, 事實上在 finite dimension 的情形確實可以很簡單的利用 dimension 來證得. 不過這些定理在一般狀況之下也成立, 雖然我們不會去談 infinite dimensional vector space, 不過利用 linear transformation 的性質來證明這些 isomorphism 更能讓我們了解這些 vector spaces 間的關係, 所以我們還是不假設 finite dimensional 的情形, 選擇不用 dimension 的方法來證明這些定理 (不過大家可以利用 dimension 來驗證這些定理的正確性).

當一個函數 f 是一對一且映成時, 我們知其反函數 $f^{\circ-1}$ 是存在的, 但當 f 是 linear transformation 時其反函數 $f^{\circ-1}$ 是否也是 linear transformation? 以下我們便是回答這個問題. 要注意, 在這裡因為要避免和 preimage 的符號相混淆, 我們用 $f^{\circ-1}$ 來表示 f 的反函數.

Proposition 2.3.4. 假設 $T:V\to W$ 為一個 *isomorphism,* 則 T 的 *inverse* (反函數) $T^{\circ-1}:W\to V$ 亦為 *isomorphism.*

Proof. T 為 one-to-one and onto, 故其 inverse $T^{\circ-1}$ 亦為 one-to-one and onto. 所以我們僅要證明 $T^{\circ-1}$ 為 W 到 V 的 linear transformation 即可. 任取 $\mathbf{w},\mathbf{w}'\in W$, 由 T 為 isomorphism 知存在唯一的 $\mathbf{v},\mathbf{v}'\in V$ 滿足 $T(\mathbf{v})=\mathbf{w}$ 且 $T(\mathbf{v}')=\mathbf{w}'$. 依 inverse 的定義 $T^{\circ-1}(\mathbf{w})=\mathbf{v}$ 且 $T^{\circ-1}(\mathbf{w}')=\mathbf{v}'$. 對任意 $r\in F$, 要證明 $T^{\circ-1}$ 為 linear transformation, 即要證明 $T^{\circ-1}(r\mathbf{w}+\mathbf{w}')=rT^{\circ-1}(\mathbf{w})+T^{\circ-1}(\mathbf{w}')=r\mathbf{v}+\mathbf{v}'$. 然而依 T 為 linear, 可得 $T(r\mathbf{v}+\mathbf{v}')=rT(\mathbf{v})+T(\mathbf{v}')=r\mathbf{w}+\mathbf{w}'$. 故再依 inverse 之定義得 $T^{\circ-1}(r\mathbf{w}+\mathbf{w}')=r\mathbf{v}+\mathbf{v}'$, 得證 $T^{\circ-1}:W\to V$ 為 linear transformation, 故為 isomorphism. $\square$

Question 2.11. *Vector spaces* 之間的 *isomorphic* 是否為一個 *equivalent relation?*

Question 2.12. 當 V 是一個 *finite dimensional vector space,* 你能利用 *Theorem 2.1.5* 證明 *Proposition 2.3.4* 嗎?

接下來我們將介紹大家在學習代數時已接觸過的幾個 Isomorphism Theorems. 首先我們先看利用一個已知的 linear transformation 得到新的 linear transformation 的方法. 假設 $T:V\to W$ 是一個 linear transformation 且 U 是一個 $\mathrm{Ker}(T)$ 的 subspace, 定義 $\overline{T}:V/U\to\mathrm{Im}(T)$ 為 $\overline{T}(\overline{\mathbf{v}})=T(\mathbf{v}),\forall\overline{\mathbf{v}}\in V/U$. 首先我們先說明 $\overline{T}$ 是一個 well-defined function. 亦即, 若 $\overline{\mathbf{v}_1}=\overline{\mathbf{v}_2}$ in V/U, 則需驗證 $\overline{T}(\overline{\mathbf{v}_1})=\overline{T}(\overline{\mathbf{v}_2})$. 然而 $\overline{\mathbf{v}_1}=\overline{\mathbf{v}_2}$, 表示 $\mathbf{v}_1-\mathbf{v}_2\in U\subseteq\mathrm{Ker}(T)$, 故得 $T(\mathbf{v}_1-\mathbf{v}_2)=\mathbf{O}_W$, 亦即 $\overline{T}(\overline{\mathbf{v}_1})=T(\mathbf{v}_1)=T(\mathbf{v}_2)=\overline{T}(\overline{\mathbf{v}_2})$. 另一方面, 對任意 $\overline{\mathbf{v}_1},\overline{\mathbf{v}_2}\in V/U$ 以及 $r\in F$,

$$\overline{T}(r\overline{\mathbf{v}_1}+\overline{\mathbf{v}_2})=\overline{T}(\overline{r\mathbf{v}_1+\mathbf{v}_2})=T(r\mathbf{v}_1+\mathbf{v}_2)=rT(\mathbf{v}_1)+T(\mathbf{v}_2)=r\overline{T}(\overline{\mathbf{v}_1})+\overline{T}(\overline{\mathbf{v}_2}),$$

得知 $\overline{T}$ 是一個 linear transformation.

Theorem 2.3.5. 假設 $T:V\to W$ 是一個 *linear transformation* 且 U 是一個 $\mathrm{Ker}(T)$ 的 *subspace,* 則函數 $\overline{T}:V/U\to\mathrm{Im}(T)$ 定義為 $\overline{T}(\overline{\mathbf{v}})=T(\mathbf{v}),\forall\overline{\mathbf{v}}\in V/U$ 是一個 *linear transformation* 且 $\mathrm{Ker}(\overline{T})=\mathrm{Ker}(T)/U$ 以及 $\mathrm{Im}(\overline{T})=\mathrm{Im}(T)$.

Proof. 前面已知, $\overline{T}$ 是一個 linear transformation. 現若 $\overline{\mathbf{v}}\in\mathrm{Ker}(\overline{T})$, 依定義表示 $\mathbf{O}_W=\overline{T}(\overline{\mathbf{v}})=T(\mathbf{v})$, 亦即 $\mathbf{v}\in\mathrm{Ker}(T)$, 得證 $\overline{\mathbf{v}}\in\mathrm{Ker}(T)/U$. 另一方面, 若 $\overline{\mathbf{v}}\in\mathrm{Ker}(T)/U$, 表示 $\mathbf{v}\in\mathrm{Ker}(T)$, 故 $\overline{T}(\overline{\mathbf{v}})=T(\mathbf{v})=\mathbf{O}_W$, 得證 $\overline{\mathbf{v}}\in\mathrm{Ker}(\overline{T})$. 故得 $\mathrm{Ker}(\overline{T})=\mathrm{Ker}(T)/U$. 最後依定義, 我們知 $\mathbf{w}\in\mathrm{Im}(\overline{T})$ 若且唯若存在 $\overline{\mathbf{v}}\in V/U$ 使得 $\mathbf{w}=\overline{T}(\overline{\mathbf{v}})=T(\mathbf{v})$ 若且唯若 $\mathbf{w}\in\mathrm{Im}(T)$. 得證 $\mathrm{Im}(\overline{T})=\mathrm{Im}(T)$. $\square$

特別地當 $U=\mathrm{Ker}(T)$, 我們有以下的定理.

Corollary 2.3.6 (The First Isomorphism Theorem)**.** 假設 $T:V\to W$ 是一個 *linear transformation* 且 $\overline{T}:V/\mathrm{Ker}(T)\to\mathrm{Im}(T)$ 定義為 $\overline{T}(\overline{\mathbf{v}})=T(\mathbf{v})$, 則 $\overline{T}$ 是一個 *isomorphism,* 即得

$$V/\mathrm{Ker}(T)\simeq\mathrm{Im}(T).$$

Proof. 因 $\overline{T}:V/\mathrm{Ker}(T)\to\mathrm{Im}(T)$ 為 linear transformation 滿足 $\mathrm{Ker}(\overline{T})=\mathrm{Ker}(T)/\mathrm{Ker}(T)=\{\overline{\mathbf{O}_V}\}$ 且 $\mathrm{Im}(\overline{T})=\mathrm{Im}(T)$, 故知 $\overline{T}$ 為 isomorphism 得證 $V/\mathrm{Ker}(T)\simeq\mathrm{Im}(T)$. □

Question 2.13. 當 V 為 *finite dimensional vector space,* 你能利用 *quotient space* 的 *dimension* 性質以及 *Dimension Theorem* 證明 $V/\mathrm{Ker}(T)\simeq\mathrm{Im}(T)$ 嗎?

Theorem 2.3.6 之所以稱為 The First Isomorphism Theorem, 是因為可以利用它證得其他的 Isomorphism Theorems.

Corollary 2.3.7 (The Second Isomorphism Theorem)**.** 假設 V 為一個 *vector space* 且 U,W 為 V 的 *subspaces.* 則

$$(U+W)/U\simeq W/U\cap W.$$

Proof. 首先注意 U 為 $U+W$ 的 subspace, 所以 $(U+W)/U$ 為一個 vector space. 考慮 $T:W\to(U+W)/U$ 定義為 $T(\mathbf{w})=\overline{\mathbf{w}}$, $\forall\mathbf{w}\in W$. 很容易驗證 T 為一個 linear transformation. 接著我們要證明 $\mathrm{Im}(T)=(U+W)/U$ 以及 $\mathrm{Ker}(T)=U\cap W$, 便可由 Corollary 2.3.6 得證 $(U+W)/U\simeq W/U\cap W$.

由於 $\mathrm{Im}(T)\subseteq(U+W)/U$, 我們僅要證明 $\mathrm{Im}(T)\supseteq(U+W)/U$. 對於任意 $\overline{\mathbf{v}}\in(U+W)/U$, 由定義知存在 $\mathbf{u}\in U,\mathbf{w}\in W$ 使得 $\overline{\mathbf{v}}=\overline{\mathbf{u}+\mathbf{w}}$. 此時考慮 $T(\mathbf{u})=\overline{\mathbf{u}}$, 我們要說明 $\overline{\mathbf{u}}=\overline{\mathbf{v}}$. 然而 $\mathbf{v}-(\mathbf{u}+\mathbf{w})\in W$ (因 $\overline{\mathbf{v}}=\overline{\mathbf{u}+\mathbf{w}}$) 以及 $\mathbf{w}\in W$, 我們有 $\mathbf{v}-\mathbf{u}=\mathbf{v}-(\mathbf{u}+\mathbf{w})+\mathbf{w}\in W$, 故得 $\overline{\mathbf{v}}=\overline{\mathbf{u}}=T(\mathbf{u})\in\mathrm{Im}(T)$. 得證 $\mathrm{Im}(T)=(U+W)/U$.

現若 $\mathbf{u}\in\mathrm{Ker}(T)$, 由於 T 的定義域為 U, 可知 $\mathbf{u}\in U$. 又因 $(U+W)/U$ 的加法單位元素為 $\overline{\mathbf{O}}$, 其中 $\mathbf{O}$ 為 V 的加法單位元素, 得 $\overline{\mathbf{O}}=T(\mathbf{u})=\overline{\mathbf{u}}$. 亦即 $\mathbf{u}=\mathbf{u}-\mathbf{O}\in W$, 故 $\mathbf{u}\in U\cap W$, 得證 $\mathrm{Ker}(T)\subseteq U\cap W$. 反之, 若 $\mathbf{u}\in U\cap W$ 則因 $\mathbf{u}\in W$, 得 $\overline{\mathbf{u}}=\overline{\mathbf{O}}$. 故 $T(\mathbf{u})=\overline{\mathbf{u}}=\overline{\mathbf{O}}$, 即 $\mathbf{u}\in\mathrm{Ker}(T)$, 得證 $\mathrm{Ker}(T)=U\cap W$. □

從這裡我們可以看出, 只要定出一個好的 linear transformation 就可利用 the First Isomorphism Theorem, 得到好的 isomorphic 性質.

Corollary 2.3.8 (The Third Isomorphism Theorem)**.** 假設 V 為一個 *vector space* 且 U,W 為 V 的 *subspaces* 滿足 $U\subseteq W$. 則

$$(V/U)/(W/U)\simeq V/W.$$

Proof. 首先注意, 由於 $U\subseteq W\subseteq V$ 且皆為 vector spaces, 故 V/U 以及 V/W 皆為 vector spaces. 現考慮 $T:V\to V/W$, 定義為 $T(\mathbf{v})=\overline{\mathbf{v}}$, $\forall\mathbf{v}\in V$. 很容易驗證 $\mathrm{Ker}(T)=W$ 且 $\mathrm{Im}(T)=V/W$. 故由 $U\subseteq\mathrm{Ker}(T)=W$, 利用 Theorem 2.3.5, 知 $\overline{T}:V/U\to V/W$ 為一個

linear transformation, 且 $\mathrm{Im}(\overline{T}) = \mathrm{Im}(T) = V/W$ 以及 $\mathrm{Ker}(\overline{T}) = \mathrm{Ker}(T)/U = W/U$. 因此利用 Corollary 2.3.6 得 $(V/U)/\mathrm{Ker}(\overline{T}) \simeq \mathrm{Im}(\overline{T})$, 即 $(V/U)/(W/U) \simeq V/W$. □

Question 2.14. 當 V 為 *finite dimensional vector space,* 你能利用 *dimension* 來證明 *the Second and Third Isomorphism Theorems* 嗎?

Question 2.15. 假設 V 為一個 *finite dimensional vector space* 且 U,W 為 V 的 *subspaces* 滿足 $U \subseteq W \subseteq V$. 我們知 $\dim(V/U) = \dim(V) - \dim(U), \dim(V/W) = \dim(V) - \dim(W)$ 且 $\dim(W/U) = \dim(W) - \dim(U)$. 也就是說 $\dim(V/U) - \dim(V/W) = \dim(W/U)$, 那我們可不可以說 $(V/U)/(V/W) \simeq W/U$?

2.4. The Matrix Connection

一個 vector space V, 若給定一組 basis 後, 由於 V 中的元素用這組 basis 都只有唯一的表法, 所以我們可以將 V 中的元素用熟悉的向量坐標來表示. 另一方面, 一個 linear transformation, 只要給定 vector space 的 basis, 也可以被唯一確定. 所以我們很自然的會將 linear transformation 和 matrix 相連結. 雖然大家以前可能僅接觸過 over $\mathbb{R}$ 或 over $\mathbb{C}$ 的 matrix. 不過關於 matrix 的運算性質的證明, 其實和 over 哪一個 field 是無關的, 所以這裡我們假設大家已熟悉這些性質, 不會再去證明它們.

若 V 是一個 finite dimensional vector space 且給定 V 的一組 basis $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$, 我們要先將這些 $\mathbf{v}_i$ 的順序排定, 用 $\beta = (\mathbf{v}_1, \ldots, \mathbf{v}_n)$ 來表示這一組排好順序的 basis, 稱為 V 的一組 *ordered basis*. 所以要注意, 一組 basis 若將其元素重新排列, 那麼所排出的 ordered basis, 雖然看成集合的話元素皆相同, 但是因為順序不同, 我們將它們視為不同的 ordered basis. 也就是說若 $\beta = (\mathbf{v}_1, \ldots, \mathbf{v}_n)$, $\beta' = (\mathbf{v}'_1, \ldots, \mathbf{v}'_n)$ 為 V 的 ordered bases, 則 $\beta = \beta'$ 表示 $\mathbf{v}_i = \mathbf{v}'_i$, $\forall i = 1, \ldots, n$.

假設 $\dim(V) = n$, 給定 V 的一個 ordered basis β 以後, 我們很自然的定出一個 V 到 $\mathbb{F}^n$ 的 linear transformation $\tau_\beta : V \to F^n$, 其中對所有 $\mathbf{v} \in V$, 利用 β 將 $\mathbf{v}$ 寫成 $\mathbf{v} = c_1\mathbf{v}_1 + \cdots + c_n\mathbf{v}_n$, 定義

$$\tau_\beta(\mathbf{v}) = \begin{pmatrix} c_1 \\ \vdots \\ c_n \end{pmatrix}.$$

這裡我們將 F^n 裡的元素寫成 column vector, 因為版面的關係有時會寫成 $(c_1, \ldots, c_n)^{\mathrm{t}}$ (即將 row vector $(c_1, \ldots, c_n)$ 取轉置). 將 F^n 裡的元素寫成 column vector 的原因是習慣性的問題, 主要是符合以後矩陣乘法的運算. 因為 β 為 ordered basis, 很容易看出 τ_β 是 well-defined, 且是一個 isomorphism. 換言之, 對於 V 中的元素, 我們可以利用 τ_β 將之置換成 F^n 中的一個 column vector. 同樣的對於 F^n 中的 column vector, 我們可利用 $\tau_\beta^{\circ -1}$ 將之還原成 V 中的元素. 這就是我們要選取 ordered basis 的主要目的, 可以利用一組 ordered basis 將 V 中的元素和 F^n 中的 column vector 作一個一對一的置換且保持 vector space 中的運算. 這裡要強調的是, 在一般的情形要將 V 中的元素轉換成 F^n 的元素 $\tau_\beta(\mathbf{v})$ 過程較麻煩 (可

能牽涉到解聯立方程組), 不過將 F^n 中的 column vector $(c_1,\ldots,c_n)^{\mathrm{t}}$ 轉換成 V 中的元素 $\tau_\beta^{\circ-1}((c_1,\ldots,c_n)^{\mathrm{t}})$ 就簡單多了, 它就是 $c_1\mathbf{v}_1+\cdots+c_n\mathbf{v}_n$.

Example 2.4.1. 考慮 $P_2(\mathbb{R})=\{ax^2+bx+c \mid a,b,c\in\mathbb{R}\}$ 這一個 $\mathbb{R}$-space, 以及它的一個 ordered basis $\beta=(x^2,x+1,-1)$. 因為 $ax^2+bx+c=a(x^2)+b(x+1)+(b-c)(-1)$, 我們可得 $\tau_\beta(ax^2+bx+c)=(a,b,b-c)^{\mathrm{t}}$. 例如 $\tau_\beta(x^2+x+1)=(1,1,0)^{\mathrm{t}}$, 而我們也可馬上知 $\tau_\beta^{\circ-1}((1,1,0)^{\mathrm{t}})=1(x^2)+1(x+1)+0(-1)=x^2+x+1$.

另外考慮 $P_1(\mathbb{R})=\{ax+b \mid a,b\in\mathbb{R}\}$ 這一個 $\mathbb{R}$-space, 以及它的一個 ordered basis $\beta'=(x-1,x+1)$. 解 $ax+b=r(x-1)+s(x+1)$, 我們可得 $r=(a-b)/2, s=(a+b)/2$, 故 $\tau_{\beta'}(ax+b)=((a-b)/2,(a+b)/2)^{\mathrm{t}}$.

Question 2.16. 在 *Example 2.4.1* 中若將 β,β' 改為 $\beta=(-1,x+1,x^2),\beta'=(x+1,x-1)$, 那麼 $\tau_\beta(ax^2+bx+c),\tau_{\beta'}(ax+b)$ 會是什麼?

現給定一 linear transformation $T:V\to W$, 我們分別選定 V,W 上的 ordered basis $\beta=(\mathbf{v}_1,\ldots,\mathbf{v}_n)$, $\beta'=(\mathbf{w}_1,\ldots,\mathbf{w}_m)$. 利用 β,β', 我們可將 T 用一個 over F 的 $m\times n$ (m 個 row, n 個 column) 的矩陣來表示. 這矩陣的每個 column 是用以下的方法定的: 第 i 個 column 為 $\tau_{\beta'}(T(\mathbf{v}_i))$. 也就是說若 $T(\mathbf{v}_i)$ 利用 β' 這個 ordered basis 可表示為 $T(\mathbf{v}_i)=c_1\mathbf{w}_1+\cdots+c_m\mathbf{w}_m$, 則這個矩陣的 i-th column 為 $\begin{pmatrix} c_1 \\ \vdots \\ c_m \end{pmatrix}$. 這個矩陣和 T 有關也和 β,β' 有關, 我們就用 ${}_{\beta'}[T]_\beta$ 來表示, 亦即

$$ {}_{\beta'}[T]_\beta=\Big(\tau_{\beta'}(T(\mathbf{v}_1)),\ldots,\tau_{\beta'}(T(\mathbf{v}_n))\Big), $$

注意每一個 $\tau_{\beta'}(T(\mathbf{v}_i))\in F^m,\forall i=1,\ldots,n$ 是一個 $m\times 1$ 的 column vector, 所以 ${}_{\beta'}[T]_\beta$ 是一個 $m\times n$ 的 over F 的 matrix.

Example 2.4.2. 同 Example 2.4.1, 考慮 $P_2(\mathbb{R})$ 以及其 ordered basis $\beta=(x^2,x+1,-1)$, 和 $P_1(\mathbb{R})$ 以及其 ordered basis $\beta'=(x-1,x+1)$. 若 $T:P_2(\mathbb{R})\to P_1(\mathbb{R})$ 定義為

$$ T(ax^2+bx+c)=2ax+b, $$

很容易驗證 T 為 linear transformation. 因為 $\dim(P_1(\mathbb{R}))=2$, $\dim(P_2(\mathbb{R}))=3$ 我們可定出 ${}_{\beta'}[T]_\beta$ 這一個 2×3 的 matrix. 事實上由於 $T(x^2)=2x, T(x+1)=1, T(-1)=0$, 利用 Example 2.4.1 的結果我們得

$$ \tau_{\beta'}(T(x^2))=\begin{pmatrix} 1 \\ 1 \end{pmatrix}, \tau_{\beta'}(T(x+1))=\begin{pmatrix} \frac{-1}{2} \\ \frac{1}{2} \end{pmatrix}, \tau_{\beta'}(T(-1))=\begin{pmatrix} 0 \\ 0 \end{pmatrix}, $$

故知

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

Question 2.17. 在 *Example 2.4.2* 中若將 β,β' 改為 $\beta=(-1,x+1,x^2),\beta'=(x+1,x-1)$, 那麼 ${}_{\beta'}[T]_\beta$ 會是什麼?

${}_{\beta'}[T]_\beta$ 這一個矩陣有什麼用呢? 它稱作 the *representative matrix* of T with respect to β,β'. 意思是說矩陣 ${}_{\beta'}[T]_\beta$ 足以代表 T 這一個 linear transformation. 回顧一下, 給定一個 $m\times n$ over F 的 matrix A, 我們可以定義一個從 F^n 到 F^m 的函數 $m_A : F^n \to F^m$, 其定義為將任意 F^n 的 column vector $\mathbf{x}$ 乘上 A 這一個 $m\times n$ matrix, 得到 $A\cdot\mathbf{x}$ 這個 F^m 的 column vector, 即 $m_A(\mathbf{x}) = A\cdot\mathbf{x}$. 利用矩陣運算的性質 $A\cdot(r\mathbf{x}+\mathbf{x}') = rA\cdot\mathbf{x}+A\cdot\mathbf{x}'$, 我們知 $m_A : F^n \to F^m$ 是一個 linear transformation. 所以有了 ${}_{\beta'}[T]_\beta$, 我們可以得到一個從 F^n 到 F^m 的 linear transformation $m_{{}_{\beta'}[T]_\beta} : F^n \to F^m$, 它和 $T : V \to W$ 有著密切的關係, 我們用以下的圖示來說明:

$$\begin{tikzcd}
V \arrow[rrr, "T"] \arrow[ddd, "\tau_\beta"'] & & & W \\
& \mathbf{v} \arrow[r, mapsto] \arrow[d, mapsto] & T(\mathbf{v}) & \\
& \tau_\beta(\mathbf{v}) \arrow[r, mapsto] & {}_{\beta'}[T]_\beta\cdot\tau_\beta(\mathbf{v}) \arrow[u, mapsto] & \\
F^n \arrow[rrr, "m_{{}_{\beta'}[T]_\beta}"'] & & & F^m \arrow[uuu, "\tau_{\beta'}^{\circ-1}"']
\end{tikzcd}$$

首先我們將任意 $\mathbf{v}\in V$ 利用 τ_β 將它轉換成 F^n 中的 column vector $\tau_\beta(\mathbf{v})$, 然後將 $\tau_\beta(\mathbf{v})$ 左邊乘上 $m\times n$ 的 matrix ${}_{\beta'}[T]_\beta$, 得到 ${}_{\beta'}[T]_\beta\cdot\tau_\beta(\mathbf{v})$ 這一個 F^m 中的 column vector. 最後我們利用 $\tau_{\beta'} : W \to F^m$ 的反函數 $\tau_{\beta'}^{\circ-1} : F^m \to W$ 將 ${}_{\beta'}[T]_\beta\cdot\tau_\beta(\mathbf{v})$ 轉換成 W 上的元素

$$\tau_{\beta'}^{\circ-1}({}_{\beta'}[T]_\beta\cdot\tau_\beta(\mathbf{v})).$$

我們希望這個元素就是 $T(\mathbf{v})$. 也就是說, 我們要說明 T 和 $\tau_{\beta'}^{\circ-1}\circ m_{{}_{\beta'}[T]_\beta}\circ\tau_\beta$ 是相同的函數, 若真是如此, 將來我們求 $T(\mathbf{v})$ 之值的問題, 就可轉換成簡單的矩陣乘法問題.

Example 2.4.3. matmul 延續 Examples 2.4.1 和 2.4.2, 我們檢查上述矩陣乘法的方法所得的元素是否就是 $T(ax^2+bx+c) = 2ax+b$. 首先將 $P_2(\mathbb{R})$ 中的任一元素 ax^2+bx+c 轉為 $\mathbb{R}^3$ 的元素 $\tau_\beta(ax^2+bx+c) = (a,b,b-c)^{\mathrm{t}}$, 再將之左邊乘上矩陣 ${}_{\beta'}[T]_\beta$ 得

$$\begin{pmatrix} 1 & \frac{-1}{2} & 0 \\ 1 & \frac{1}{2} & 0 \end{pmatrix}\cdot\begin{pmatrix} a \\ b \\ b-c \end{pmatrix} = \begin{pmatrix} a-\frac{1}{2}b \\ a+\frac{1}{2}b \end{pmatrix}.$$

最後將此 $\mathbb{R}^2$ 的元素轉回 $P_1(\mathbb{R})$ 的元素得 $(a-(b/2))(x-1)+(a+(b/2))(x+1) = 2ax+b$ 確實和 $T(ax^2+bx+c)$ 相等.

在回答為何 $T = \tau_{\beta'}^{\circ-1}\circ m_{{}_{\beta'}[T]_\beta}\circ\tau_\beta$ 之前, 我們先回顧一個矩陣乘法的重要看法. 當一個 $m\times n$ matrix A, 乘上 F^n 的一個 column vector $\mathbf{x}$, 我們知道 $A\cdot\mathbf{x}$ 會是 F^m 的 column vector. 事實上, 若 $\mathbf{A}_1,\dots,\mathbf{A}_n$ 為 A 的第 1 到第 n 個 column (別忘了 A 有 n 個 column 且每個

column 是 F^m 的一個 column vector), 而 $\mathbf{x} = (x_1, \ldots, x_n)^t$, 則

$$A \cdot \mathbf{x} = x_1\mathbf{A}_1 + \cdots + x_n\mathbf{A}_n.$$

現已知 T 和 $\tau_{\beta'}^{\circ -1} \circ m_{_{\beta'}[T]_\beta} \circ \tau_\beta$ 皆為 $V \to W$ 的 linear transformation, 由 Theorem 2.1.5, 我們知道要說它們相等, 只要將 β 這一個 basis 內的每個元素 $\mathbf{v}_i$, 分別代入檢查是否相同即可. 然而依定義 $\tau_\beta(\mathbf{v}_i) = (0, \ldots, 1, \ldots, 0)^t$, 其中 $(0, \ldots, 1, \ldots, 0)$ 只有在第 i 個位置是 1 其他位置皆為 0. 所以依前面所提矩陣的乘法知 ${}_{\beta'}[T]_\beta \cdot \tau_\beta(\mathbf{v}_i) = {}_{\beta'}[T]_\beta \cdot (0, \ldots, 1, \ldots, 0)^t$ 為矩陣 ${}_{\beta'}[T]_\beta$ 的第 i 個 column, 依前面定義知此 column 為 $\tau_{\beta'}(T(\mathbf{v}_i))$, 因此

$$\tau_{\beta'}^{\circ -1} \circ m_{_{\beta'}[T]_\beta} \circ \tau_\beta(\mathbf{v}_i) = \tau_{\beta'}^{\circ -1}({}_{\beta'}[T]_\beta \cdot \tau_\beta(\mathbf{v}_i)) = \tau_{\beta'}^{\circ -1}(\tau_{\beta'}(T(\mathbf{v}_i))) = T(\mathbf{v}_i), \forall i = 1, \ldots, n$$

得證

$$T = \tau_{\beta'}^{\circ -1} \circ m_{_{\beta'}[T]_\beta} \circ \tau_\beta. \tag{2.1}$$

當 V, W 為 vector spaces, 我們曾介紹所有 V 到 W 的 linear transformation 形成一個 vector space, 用 $\mathscr{L}(V, W)$ 來表示. 現令 $M_{m\times n}(F)$ 表示所有 over F 的 $m \times n$ matrices. 依矩陣的運算性質, 很容易檢查 $M_{m\times n}(F)$ 亦為一個 vector space. 今若固定 V 的一組 ordered basis $\beta = (\mathbf{v}_1, \ldots, \mathbf{v}_n)$ 和 W 的一組 ordered basis $\beta' = (\mathbf{w}_1, \ldots, \mathbf{w}_m)$, 前面將 linear transformation 轉換成 matrix 的方法, 給了我們一個從 $\mathscr{L}(V, W)$ 到 $M_{m\times n}(F)$ 的函數 Φ, 其定義為將任意 linear transformation $T : V \to W$ 送到 ${}_{\beta'}[T]_\beta$ 這一個 $m \times n$ matrix, 亦即 $\Phi(T) = {}_{\beta'}[T]_\beta, \forall T \in \mathscr{L}(V, W)$. 我們需要說明 Φ 是一個 linear transformation, 即說明對任意 $T_1, T_2 \in \mathscr{L}(V, W)$ 以及 $r \in F$, $\Phi(rT_1 + T_2) = {}_{\beta'}[rT_1 + T_2]_\beta$ 會和 $r\Phi(T_1) + \Phi(T_2) = r({}_{\beta'}[T_1]_\beta) + {}_{\beta'}[T_2]_\beta$ 相等. 然而依定義, 矩陣 ${}_{\beta'}[rT_1 + T_2]_\beta$ 的 i-th column 為 $\tau_{\beta'}((rT_1 + T_2)(\mathbf{v}_i)) = \tau_{\beta'}(rT_1(\mathbf{v}_i) + T_2(\mathbf{v}_i))$, 但因 $\tau_{\beta'}$ 為 linear, 此即 $r\tau_{\beta'}(T_1(\mathbf{v}_i)) + \tau_{\beta'}(T_2(\mathbf{v}_i))$. 另一方面矩陣 ${}_{\beta'}[T_1]_\beta, {}_{\beta'}[T_2]_\beta$ 的 i-th column 分別為 $\tau_{\beta'}(T_1(\mathbf{v}_i)), \tau_{\beta'}(T_2(\mathbf{v}_i))$, 故依矩陣加法及係數積的定義矩陣 $r({}_{\beta'}[T_1]_\beta) + {}_{\beta'}[T_2]_\beta$ 的 i-th column 為 $r\tau_{\beta'}(T_1(\mathbf{v}_i)) + \tau_{\beta'}(T_2(\mathbf{v}_i))$. 得證 ${}_{\beta'}[rT_1 + T_2]_\beta$ 和 $r({}_{\beta'}[T_1]_\beta) + {}_{\beta'}[T_2]_\beta$ 為相同的矩陣, 故知 Φ 為 linear transformation.

接著我們要說明 $\Phi : \mathscr{L}(V, W) \to M_{m\times n}(F)$ 為 isomorphism (即 one-to-one and onto). 給定任意矩陣 $A \in M_{m\times n}(F)$, 若 A 的 i-th column 為 $\mathbf{A}_i$, 考慮 $\tau_{\beta'}^{\circ -1}(\mathbf{A}_i) \in W$. 由 Theorem 2.1.5 知存在唯一的 linear transformation $T : V \to W$ 滿足 $T(\mathbf{v}_i) = \tau_{\beta'}^{\circ -1}(\mathbf{A}_i), \forall i = 1, \ldots, n$. 依定義, 此時 ${}_{\beta'}[T]_\beta$ 的 i-th row 為

$$\tau_{\beta'}(T(\mathbf{v}_i)) = \tau_{\beta'}(\tau_{\beta'}^{\circ -1}(\mathbf{A}_i)) = \mathbf{A}_i,$$

故知 ${}_{\beta'}[T]_\beta = A$. 亦即 T 為 $\mathscr{L}(V, W)$ 中唯一滿足 $\Phi(T) = {}_{\beta'}[T]_\beta = A$ 的 linear transformation, 得證 Φ 為 isomorphism. 我們將此結果整理如下.

Theorem 2.4.4. 假設 V, W 為 *vector spaces* 且 $\dim(V) = n, \dim(W) = m$. 給定 V, W 的 *ordered basis* β, β'. 若令 $\Phi : \mathscr{L}(V, W) \to M_{m\times n}(F)$ 滿足 $\Phi(T) = {}_{\beta'}[T]_\beta, \forall T \in \mathscr{L}(V, W)$, 則 Φ 為一個 *isomorphism*, 即

$$\mathscr{L}(V, W) \simeq M_{m\times n}(F).$$

Question 2.18. 若 $\dim(V)=n, \dim(W)=m$, *Theorem 2.4.4* 告訴我們 $\dim(\mathscr{L}(V,W)) = \dim(M_{m\times n}(F)) = mn$, 你能利用 $M_{m\times n}(F)$ 的標準基底, 找到 $\mathscr{L}_(V,W)$ 的 *basis?*

Question 2.19. 給定 $A \in M_{m\times n}(F)$. 令 $m_A : F^n \to F^m$ 為 $m_A(\mathbf{x}) = A\cdot\mathbf{x}, \forall \mathbf{x} \in F^n$. 回顧 $\operatorname{Im}(m_A) = \{A\cdot\mathbf{x} \in F^m \mid \mathbf{x} \in F^n\}$ 稱為 A 的 column space, 用 $C(A)$ 來表示, 且 $\dim(C(A))$ 稱為 *the* rank *of* A. 而 $\operatorname{Ker}(m_A) = \{\mathbf{x} \in F^n \mid A\cdot\mathbf{x} = \mathbf{O}\}$ 稱為 A 的 null space, 用 $N(A)$ 來表示, 且 $\dim(N(A))$ 稱為 *the* nullity *of* A. 若 $T : V \to W$ 為 *linear transformation* 且 $\Phi(T) = A$, 利用 $\tau_\beta, \tau_{\beta'}$ 為 *isomorphism,* 可得 $\operatorname{Ker}(T) \simeq N(A)$ 且 $\operatorname{Im}(T) \simeq C(A)$. 你能看出

$$\text{rank of } A + \text{ nullity of } A = n?$$

在前面我們曾用以下圖示說明 linear transformation T 和 matrix ${}_{\beta'}[T]_\beta$ 之間關係.

$$\begin{tikzcd}
V \arrow[r, "T"] \arrow[d, shift right, "\tau_\beta"'] & W \arrow[d, shift right, "\tau_{\beta'}"'] \\
F^n \arrow[r, "m_{{}_{\beta'}[T]_\beta}"] \arrow[u, shift right, "\tau_\beta^{\circ-1}"'] & F^m \arrow[u, shift right, "\tau_{\beta'}^{\circ-1}"']
\end{tikzcd}$$

因為我們證得了 $T = \tau_{\beta'}^{\circ-1} \circ m_{{}_{\beta'}[T]_\beta} \circ \tau_\beta$, 這個圖示便可稱為是 *commutative diagram.* 在 Commutative diagram 中, 任兩個端點若有不同路徑可連結, 則這兩個路徑所對應的函數會相同. 例如 V 到 W 有兩個路徑: 一個是直接利用 T; 另一個是由 V 先利用 τ_β 到達 F^n, 再利用 $m_{{}_{\beta'}[T]_\beta}$ 到 F^m, 最後經 $\tau_{\beta'}^{\circ-1}$ 到達 W. 所以說這個圖示為 commutative diagram 就是表達 $T = \tau_{\beta'}^{\circ-1} \circ m_{{}_{\beta'}[T]_\beta} \circ \tau_\beta$. 不過這裡要注意路徑的方向性, 例如從 F^m 到 W 的 $\tau_{\beta'}^{\circ-1}$ 是 isomorphism 所以也有一個反向的 W 到 F^m 路徑可行, 即其反函數 $\tau_{\beta'}$. 所以從 V 到 F^m 我們也有兩個路徑: 一個先利用 T 從 V 到 W, 再接 $\tau_{\beta'}$ 到達 F^m; 另一個是利用 τ_β 由 V 到 F^n, 再經 $m_{{}_{\beta'}[T]_\beta}$ 到達 F^m. 所以我們有 $\tau_{\beta'} \circ T = m_{{}_{\beta'}[T]_\beta} \circ \tau_\beta$. 事實上

$$\tau_{\beta'} \circ T = \tau_{\beta'} \circ (\tau_{\beta'}^{\circ-1} \circ m_{{}_{\beta'}[T]_\beta} \circ \tau_\beta) = (\tau_{\beta'} \circ \tau_{\beta'}^{\circ-1}) \circ m_{{}_{\beta'}[T]_\beta} \circ \tau_\beta = m_{{}_{\beta'}[T]_\beta} \circ \tau_\beta,$$

所以 commutative diagram 是一個很方便判斷兩函數是否相同的工具.

另一方面 V 到 F^n 就不能說有兩個路徑了, 主要是可利用 T 從 V 到 W 然後經 $\tau_{\beta'}$ 到 F^m 但無法保證能由 F^m 到 F^n 了 (除非知矩陣 ${}_{\beta'}[T]_\beta$ 為 invertible).

Question 2.20. 上圖示中 F^n 到 F^m 是否有兩個路徑? 它們代表哪些函數間的關係?

接著我們便是要用 commutative diagram 來看合成函數與矩陣乘法的關係. 設 V, W, U 為 vector spaces, 其中 $\dim(V) = n, \dim(W) = m, \dim(U) = q$ 且 $\beta, \beta'\beta''$ 分別為它們的 ordered bases. 若 $T_1 : V \to W$, $T_2 : V \to W$ 為 linear transformations 我們有以下的 commutative diagram:

$$\begin{tikzcd}
V \arrow[r, "T_1"] \arrow[d, "\tau_\beta"'] & W \arrow[r, "T_2"] \arrow[d, shift left, "\tau_{\beta'}"] & U \\
F^n \arrow[r, "m_{{}_{\beta'}[T_1]_\beta}"] & F^m \arrow[r, "m_{{}_{\beta''}[T_2]_{\beta'}}"] \arrow[u, shift left, "\tau_{\beta'}^{\circ-1}"] & F^q \arrow[u, "\tau_{\beta''}^{\circ-1}"']
\end{tikzcd}$$

事實上

$$T_2 \circ T_1 = (\tau_{\beta''}^{\circ-1} \circ m_{{}_{\beta''}[T_2]_{\beta'}} \circ \tau_{\beta'}) \circ (\tau_{\beta'}^{\circ-1} \circ m_{{}_{\beta'}[T_1]_\beta} \circ \tau_\beta) = \tau_{\beta''}^{\circ-1} \circ m_{{}_{\beta''}[T_2]_{\beta'}} \circ m_{{}_{\beta'}[T_1]_\beta} \circ \tau_\beta.$$

回顧一下, 當 A 是一個 $m \times n$ matrix, B 是一個 $q \times m$ matrix, 則 $m_A : F^n \to F^m$ 和 $m_B : F^m \to F^q$ 的合成 $m_B \circ m_A : F^n \to F^q$ 就是 $m_{B \cdot A} : F^n \to F^q$. 這是因為對任意 $\mathbf{x} \in F^n$, 由矩陣乘法的結合率可得

$$m_B \circ m_A(\mathbf{x}) = m_B(m_A(\mathbf{x})) = m_B(A \cdot \mathbf{x}) = B \cdot (A \cdot \mathbf{x}) = (B \cdot A) \cdot \mathbf{x} = m_{B \cdot A}(\mathbf{x}).$$

所以利用 $m_{{}_{\beta''}[T_2]_{\beta'}} \circ m_{{}_{\beta'}[T_1]_\beta} = m_{{}_{\beta''}[T_2]_{\beta'} \cdot {}_{\beta'}[T_1]_\beta}$ 可得

$$T_2 \circ T_1 = \tau_{\beta''}^{\circ-1} \circ m_{{}_{\beta''}[T_2]_{\beta'} \cdot {}_{\beta'}[T_1]_\beta} \circ \tau_\beta. \tag{2.2}$$

我們有以下之結果.

Proposition 2.4.5. 假設 V, W, U 為 *finite dimensional vector spaces,* 且 $\beta, \beta'\beta''$ 分別為它們的 *ordered bases.* 若 $T_1 : V \to W$, $T_2 : V \to W$ 為 *linear transformations,* 則

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

Proof. 首先回顧, 若 A, A' 皆為 $q \times n$ matrices, 且 $m_A : F^n \to F^q$ 和 $m_{A'} : F^n \to F^q$ 相同, 則考慮 $\mathbf{x}_i = (0, \ldots, 1, \ldots, 0)^t$, 其中 $(0, \ldots, 1, \ldots, 0)$ 表示第 i 位置為 1 其他位置為 0, 可由 $m_A(\mathbf{x}_i) = m_{A'}(\mathbf{x}_i), \forall i = 1, \ldots, n$ 得 A 和 A' 每一個 column 皆相同. 故得 $A = A'$. 現由等式 (2.1) 和等式 (2.2) 我們知

$$\tau_{\beta''}^{\circ-1} \circ m_{{}_{\beta''}[T_2 \circ T_1]_\beta} \circ \tau_\beta = T_1 \circ T_2 = \tau_{\beta''}^{\circ-1} \circ m_{{}_{\beta''}[T_2]_{\beta'} \cdot {}_{\beta'}[T_1]_\beta} \circ \tau_\beta,$$

故由 $\tau_\beta, \tau_{\beta''}^{\circ-1}$ 皆為 isomorphism 知

$$m_{{}_{\beta''}[T_2 \circ T_1]_\beta} = m_{{}_{\beta''}[T_2]_{\beta'} \cdot {}_{\beta'}[T_1]_\beta},$$

得證 ${}_{\beta''}[T_2 \circ T_1]_\beta = {}_{\beta''}[T_2]_{\beta'} \cdot {}_{\beta'}[T_1]_\beta$. □

最後我們要談換了一組 ordered basis 對 representative matrix 的影響. 假設 $\dim(V) = n, \dim(W) = m$ 且 β_1, β_2 為 V 的 ordered bases, β_1', β_2' 為 W 的 ordered bases. $T : V \to W$ 為 linear transformation, 依定義 T 分別利用 V, W 的 ordered bases β_1, β_1' 所得的 representative matrix 為 ${}_{\beta_1'}[T]_{\beta_1}$, 而 T 用 ordered bases β_2, β_2' 所得的 representative matrix 為 ${}_{\beta_2'}[T]_{\beta_2}$. 我

們要探討這兩個矩陣的關係為何, 首先我們有以下的 commutative diagram.

$$\begin{tikzcd}
F^n \arrow[r, "m_{\beta_2'}[T]_{\beta_2}"] & F^m \arrow[d, "\tau_{\beta_2'}^{\circ-1}"] \\
V \arrow[u, "\tau_{\beta_2}"] \arrow[r, "T"] \arrow[d, "\tau_{\beta_1}"'] & W \\
F^n \arrow[r, "m_{\beta_1'}[T]_{\beta_1}"] & F^m \arrow[u, "\tau_{\beta_1'}^{\circ-1}"']
\end{tikzcd}$$

最上層的 $m_{\beta_2'}[T]_{\beta_2} : F^n \to F^m$, 可用下面的路徑 (別忘了 $\tau_{\beta_2}, \tau_{\beta_2'}$ 為 isomorphism), 所以我們有

$$m_{\beta_2'}[T]_{\beta_2} = (\tau_{\beta_2'} \circ \tau_{\beta_1'}^{\circ-1}) \circ m_{\beta_1'}[T]_{\beta_1} \circ (\tau_{\beta_1} \circ \tau_{\beta_2}^{\circ-1}). \tag{2.3}$$

這裡 $\tau_{\beta_1} \circ \tau_{\beta_2}^{\circ-1}$ 和 $\tau_{\beta_2'} \circ \tau_{\beta_1'}^{\circ-1}$ 皆為 linear transformations, 所以我們也可以探討它們的矩陣表示法. 考慮 V 到 V 的 identity map, $\mathrm{id} : V \to V$, 滿足 $\mathrm{id}(\mathbf{v}) = \mathbf{v}$, $\forall \mathbf{v} \in V$. 很自然的這是一個 linear transformation, 所以對定義域的 V 使用 ordered basis β_2, 而對映域的 V 使用 ordered basis β_1, 我們有以下的 commutative diagram.

$$\begin{tikzcd}
V \arrow[r, "\mathrm{id}"] \arrow[d, "\tau_{\beta_2}"'] & V \\
F^n \arrow[r, "m_{\beta_1}[\mathrm{id}]_{\beta_2}"] & F^n \arrow[u, "\tau_{\beta_1}^{\circ-1}"']
\end{tikzcd}$$

考慮下層的 $m_{\beta_1}[\mathrm{id}]_{\beta_2} : F^n \to F^n$, 可用往上的路徑 $\tau_{\beta_2}^{\circ-1}$ 從 F^n 到 V 接 identity map id, 再接 τ_{β_1} 到 F^n. 利用 identity map 和函數合成不變我們有以下的等式

$$m_{\beta_1}[\mathrm{id}]_{\beta_2} = \tau_{\beta_1} \circ \mathrm{id} \circ \tau_{\beta_2}^{\circ-1} = \tau_{\beta_1} \circ \tau_{\beta_2}^{\circ-1}. \tag{2.4}$$

同理我們有

$$m_{\beta_2'}[\mathrm{id}]_{\beta_1'} = \tau_{\beta_2'} \circ \tau_{\beta_1'}^{\circ-1}. \tag{2.5}$$

結合等式 (2.3, 2.4, 2.5), 得

$$m_{\beta_2'}[T]_{\beta_2} = m_{\beta_2'}[\mathrm{id}]_{\beta_1'} \circ m_{\beta_1'}[T]_{\beta_1} \circ m_{\beta_1}[\mathrm{id}]_{\beta_2},$$

因此我們有以下之結果.

Proposition 2.4.6. 設 β_1, β_2 為 V 的 *ordered bases,* β_1', β_2' 為 W 的 *ordered bases.* $T : V \to W$ 為 *linear transformation,* 則

$${}_{\beta_2'}[T]_{\beta_2} = {}_{\beta_2'}[\mathrm{id}]_{\beta_1'} \cdot {}_{\beta_1'}[T]_{\beta_1} \cdot {}_{\beta_1}[\mathrm{id}]_{\beta_2}.$$

要注意, 這裡第一個 ${}_{\beta_1}[\mathrm{id}]_{\beta_2}$ 的 identity map 是 V 到 V 的 linear transformation, 所以若 $\dim(V)=n$, 則 ${}_{\beta_1}[\mathrm{id}]_{\beta_2}$ 是一個 $n\times n$ 的 matrix. 而第二個 ${}_{\beta_2'}[\mathrm{id}]_{\beta_1'}$ 的 identity map 是 W 到 W 的 linear transformation, 所以若 $\dim(W)=m$, 則 ${}_{\beta_2'}[\mathrm{id}]_{\beta_1'}$ 是一個 $m\times m$ 的 matrix.

Question 2.21. 你可以利用 *Proposition 2.4.5* 證明 *Proposition 2.4.6* 嗎?

當 $\beta_1=(\mathbf{v}_1,\dots,\mathbf{v}_n)$, $\beta_2=(\mathbf{v}_1',\dots,\mathbf{v}_n')$ 為 V 的 ordered bases, 利用 representative matrix 的造法, 我們知道矩陣 ${}_{\beta_1}[\mathrm{id}]_{\beta_2}$ 的第 i 個 column, 就是將 $\mathbf{v}_i'$ 用 β_1 所得的坐標, 即 $\tau_{\beta_1}(\mathbf{v}_i')$. 因此可得

$$ {}_{\beta_1}[\mathrm{id}]_{\beta_2}=\big(\tau_{\beta_1}(\mathbf{v}_1'),\dots,\tau_{\beta_1}(\mathbf{v}_n')\big). $$

特別的, 當 $\beta_1=\beta_2$ 時, 我們有 ${}_{\beta_1}[\mathrm{id}]_{\beta_1}$ 就是 $n\times n$ 的 identity matrix I_n.

Question 2.22. 考慮 $P_2(\mathbb{R})$ 以及其 *ordered bases* $\beta_1=(x^2,x+1,-1)$, $\beta_2=(x+1,-1,x^2)$. 試求 ${}_{\beta_1}[\mathrm{id}]_{\beta_2}$ 和 ${}_{\beta_2}[\mathrm{id}]_{\beta_1}$. 什麼是 ${}_{\beta_1}[\mathrm{id}]_{\beta_2}\cdot{}_{\beta_2}[\mathrm{id}]_{\beta_1}$ 和 ${}_{\beta_2}[\mathrm{id}]_{\beta_1}\cdot{}_{\beta_1}[\mathrm{id}]_{\beta_2}$ 呢?

${}_{\beta_1}[\mathrm{id}]_{\beta_2}$ 稱為 the *change of basis matrix* from β_2 to β_1. 它指的是 V 中的元素 $\mathbf{v}$ 用 β_2 所得的坐標 $\tau_{\beta_2}(\mathbf{v})$ 轉換成用 β_1 所得的坐標 $\tau_{\beta_1}(\mathbf{v})$ 所需乘上的矩陣. 也就是說若 $\mathbf{v}=c_1\mathbf{v}_1'+\cdots+c_n\mathbf{v}_n'$ (即 $\tau_{\beta_2}(\mathbf{v})=(c_1,\dots,c_n)^{\mathrm{t}}$), 且 ${}_{\beta_1}[\mathrm{id}]_{\beta_2}\cdot(c_1,\dots,c_n)^{\mathrm{t}}=(d_1,\dots,d_n)^{\mathrm{t}}$, 則可得 $\mathbf{v}=d_1\mathbf{v}_1+\cdots+d_n\mathbf{v}_n$. 很容易理解將這個動作反向操作便可還原成原坐標, 事實上由 Proposition 2.4.5, 我們知

$$ {}_{\beta_1}[\mathrm{id}]_{\beta_2}\cdot{}_{\beta_2}[\mathrm{id}]_{\beta_1}={}_{\beta_1}[\mathrm{id}]_{\beta_1}=I_n \quad \text{且} \quad {}_{\beta_2}[\mathrm{id}]_{\beta_1}\cdot{}_{\beta_1}[\mathrm{id}]_{\beta_2}={}_{\beta_2}[\mathrm{id}]_{\beta_2}=I_n. \tag{2.6} $$

所以 ${}_{\beta_1}[\mathrm{id}]_{\beta_2}$ 和 ${}_{\beta_2}[\mathrm{id}]_{\beta_1}$ 皆為 invertible matrices, 且它們互為 inverse.

我們已知 change of basis matrix 必為 invertible matrix, 那給定一個 invertible matrix, 是否也會是一個 change of basis matrix 呢? 我們有以下之結果.

Proposition 2.4.7. 假設 V 為 *finite dimensional vector space* 且 $\dim(V)=n$. 若 β_1 為 V 的一個 *ordered basis* 且 P 為一個 $n\times n$ 的 *invertible matrix,* 則可找到 V 的一個 *ordered basis* β_2 使得 ${}_{\beta_1}[\mathrm{id}]_{\beta_2}=P$, 也可找到 V 的一個 *ordered basis* β_3 使得 ${}_{\beta_3}[\mathrm{id}]_{\beta_1}=P$.

Proof. 令 $\beta_1=(\mathbf{v}_1,\dots,\mathbf{v}_n)$, 對於 $i=1,\dots,n$ 若 P_i 為 P 的 i-th column, 考慮 $\mathbf{v}_i'=\tau_{\beta_1}^{\circ-1}(P_i)$ (即若 $P_i=(r_1,\dots,r_n)^{\mathrm{t}}$, 則令 $\mathbf{v}_i'=r_1\mathbf{v}_1+\cdots+r_n\mathbf{v}_n$). 因 P 為 invertible, $\{P_1,\dots,P_n\}$ 為 F^n 中的 linearly independent column vectors, 故由 τ_{β_1} 為 isomorphism 知 $\{\mathbf{v}_1',\dots,\mathbf{v}_n'\}$ 亦在 V 中為 linearly independent, 得知 $\{\mathbf{v}_1',\dots,\mathbf{v}_n'\}$ 為 V 中的一組 basis. 令 $\beta_2=(\mathbf{v}_1',\dots,\mathbf{v}_n')$, 則依定義 ${}_{\beta_1}[\mathrm{id}]_{\beta_2}$ 的 i-th column 為 $\tau_{\beta_1}(\mathbf{v}_i')=\tau_{\beta_1}(\tau_{\beta_1}^{\circ-1}(P_i))=P_i$, 得證 $P={}_{\beta_1}[\mathrm{id}]_{\beta_2}$. 同理, 可得 V 的一個 ordered basis β_3 使得 ${}_{\beta_1}[\mathrm{id}]_{\beta_3}=P^{-1}$, 則此時 $P={}_{\beta_1}[\mathrm{id}]_{\beta_3}^{-1}={}_{\beta_3}[\mathrm{id}]_{\beta_1}$. $\square$

Question 2.23. 假設 V,W 為 *vector spaces,* $\dim(V)=n,\dim(W)=m$ 且 β_1,β_1' 分別為 V,W 的 *ordered basis.* 設 $T:V\to W$ 為 *linear transformation,* 若 $P\in M_{n\times n}(F)$ 且 $Q\in M_{m\times m}(F)$ 皆為 *invertible matrices,* 是否可找到 β_2,β_2' 分別為 V,W 的 *ordered basis* 使得

$$ {}_{\beta_2'}[T]_{\beta_2}=Q\cdot{}_{\beta_1'}[T]_{\beta_1}\cdot P? $$

Chapter 3

Linear Operator

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

3.1. Basic Concept

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

因此我們有以下之結論.

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

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

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

(1) T 是一個 *isomorphism.*

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

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

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

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

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

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

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

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

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

也可對 j-th column 降階得

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

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

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

利用此矩陣我們有以下的結果.

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

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

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

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

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

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

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

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

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

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

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

3.2. Characteristic Polynomial

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

給定一個係數在 F 的 polynomial $f(x) = c_dx^d + \cdots + c_1x + c_0$ 以及一個 $n\times n$ matrix A, 我們定義

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

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

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

因此加上利用矩陣加法乘法的分配律, 我們可以得到以下的結果.

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

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

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

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

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

利用數學歸納法可得

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

我們有以下結果.

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

Proof. 由 $A\sim B$ 知存在 P 為 invertible 使得 $B=P^{-1}\cdot A\cdot P$. 若 $f(x)=c_dx^d+\cdots+c_1x+c_0$, 則

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

得證 $f(A)\sim f(B)$. □

我們也可把這概念推廣到 linear operator, 假設 $f(x)=c_dx^d+\cdots+c_1x+c_0\in F[x]$ 以及 $T:V\to V$ 是一個 linear operator, 由於 linear operators 之間的合成和矩陣之間的相乘相對應 (參見式子 (3.1)), 我們定義

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

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

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

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

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

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

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

由此我們有以下之結果.

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

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

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

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

□

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

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

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

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

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

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

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

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

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

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

則

$$\chi_{A_1}=\det\begin{pmatrix}x-1&1\\-1&x+1\end{pmatrix}=(x-1)(x+1)+1=x^2,$$
$$\chi_{A_2}=\det\begin{pmatrix}x-1&1\\0&x+1\end{pmatrix}=(x-1)(x+1)=x^2-1,$$
$$\chi_{A_3}=\det\begin{pmatrix}x-1&1\\-2&x+1\end{pmatrix}=(x-1)(x+1)+2=x^2+1.$$

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

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

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

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

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

得證

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

□

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

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

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

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

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

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

例如因矩陣加法乘法有分配律, 我們有

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

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

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

Example 3.2.9. 考慮

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

以及

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

直接相乘我們有

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

而另一邊如多項式相乘展開有

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

所以兩種算法結果是相等的.

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

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

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

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

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

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

將等式 (3.2) 左邊展開, 我們得

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

應該和等式 (3.2) 右式相同, 故比較係數得

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

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

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

因未左邊全部加起來會等於右邊全部加起來, 得證

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

□

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

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

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

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

3.3. Minimal Polynomial

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

Question 3.5. *你會證明 Lemma 3.3.5 嗎?*

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

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

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

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

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

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

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

Theorem 3.3.7.

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

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

Proof.

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

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

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

□

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

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

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

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

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

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

Theorem 3.3.9.

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

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

Proof.

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

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

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

□

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

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

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

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

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

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

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

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

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

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

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

3.4. Internal Direct Sum

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

3.5. Primary Decomposition

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

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

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

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

回顧一下, 當 $T:V\to V$ 為 linear operator, 對於 $f(x)=a_dx^d+\cdots+a_1x+a_0\in F[x]$, 我們可定義一個 linear operator $f(T)=a_dT^{\circ d}+\cdots+a_1T+a_0\,\mathrm{id}$.

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

Proof. 因 W 為 T-invariant, 對任意 $\mathbf{w} \in W$, 因為 $T(\mathbf{w}) \in W$ 故得 $T^{\circ 2}(\mathbf{w}) = T(T(\mathbf{w})) \in W$. 利用數學歸納法知 $T^{\circ i}(\mathbf{w}) \in W$, $\forall i \in \mathbb{N}$. 現若 $f(x) = a_d x^d + \cdots + a_1 x + a_0 \in F[x]$, 因 W 為 subspace, 得 $f(T)(\mathbf{w}) = a_d T^{\circ d}(\mathbf{w}) + \cdots + a_1 T(\mathbf{w}) + a_0 \mathbf{w} \in W$, $\forall \mathbf{w} \in W$. 得證 W 為 $f(T)$-invariant. □

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

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

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

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

得證 $\mathrm{Im}(f(T))$ 為 T-invariant.

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

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

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

我們知 $T^{\circ 2}|_W$ 和 $T|_W{}^{\circ 2}$ 為 W 上相同的 linear operator. 利用數學歸納法可得 $T^{\circ i}|_W = T|_W{}^{\circ i}, \forall i \in \mathbb{N}$. 現若 $f(x) = a_d x^d + \cdots + a_1 x + a_0 \in F[x]$, 則對於任意 $\mathbf{w} \in W$, 皆有

$$\begin{aligned} f(T)|_W(\mathbf{w}) = f(T)(\mathbf{w}) &= a_d\, T^{\circ d}|_W(\mathbf{w}) + \cdots + a_1\, T|_W(\mathbf{w}) + a_0\, \mathrm{id}|_W(\mathbf{w}) \\ &= a_d\, T|_W{}^{\circ d}(\mathbf{w}) + \cdots + a_1\, T|_W(\mathbf{w}) + a_0\, \mathrm{id}|_W(\mathbf{w}) = f(T|_W)(\mathbf{w}). \end{aligned}$$

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

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

利用此結果, 我們有以下之 Lemma.

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

□

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

因此由 Lemma 3.5.6 得

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

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

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

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

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

其中 $p_1(x),\ldots,p_k(x)$ 為相異的 *monic irreducible polynomials.* 若令 $V_i=\operatorname{Ker}(p_i(T)^{\circ m_i})$, *for* $i=1,\ldots,k$, 則
$$V=V_1\oplus\cdots\oplus V_k$$
且
$$\mu_{T|_{V_i}}(x)=p_i(x)^{m_i}\quad\text{and}\quad\chi_{T|_{V_i}}(x)=p_i(x)^{c_i},\forall i=1,\ldots,k.$$

Proof. 我們對 $\mu_T(x)$ 的相異 moinic irreducible divisor (質因式) 的個數 k 作數學歸納法. 若 $k=1$, 表示 $\mu_T(x)=p_1(x)^{m_1}$, 此時因 $\mu_T(T)=p_1(T)^{\circ m_1}=\mathbf{O}$, 故知 $V=\operatorname{Ker}(p_1(T)^{\circ m_1})$, 因此在 $k=1$ 時定理成立. 現假設當 $\mu_T(x)$ 有 $k-1$ 個相異 monic irreducible divisor 時亦成立, 我們考慮 $\mu_T(x)=p_1(x)^{m_1}\cdots p_k(x)^{m_k}$ 的情形. 此時令 $f(x)=p_1(x)^{m_1}$, $g(x)=p_2(x)^{m_2}\cdots p_k(x)^{m_k}$, 因 $f(x),g(x)$ 為 relatively prime, 由 Theorem 3.5, 我們知 $V=U\oplus W$, 其中 $U=\operatorname{Ker}(f(T))=\operatorname{Ker}(p_1(T)^{\circ m_1})$, $W=\operatorname{Ker}(g(T))$ 而且 $\mu_{T|_U}(x)=p_1(x)^{m_1}$, $\mu_{T|_W}(x)=p_2(x)^{m_2}\cdots p_k(x)^{m_k}$. 現考慮 vector space W 以及 linear operator $T|_W:W\to W$, 套用 induction 在 $k-1$ 情形的假設知 $W=V_2\oplus\cdots\oplus V_k$ 其中 $V_i=\operatorname{Ker}(p_i(T|_W)^{\circ m_i})$ 且 $\mu_{(T|_W)|_{V_i}}(x)=p_i(x)^{m_i}$, $\forall i=2,\ldots,k$. 然而當 $i=2,\ldots,k$ 時 $p_i(x)^{m_i}\mid g(x)$, 故 $\operatorname{Ker}(p_i(T)^{\circ m_i})\subseteq\operatorname{Ker}(g(T))=W$, 因此
$$V_i=\operatorname{Ker}(p_i(T|_W)^{\circ m_i})=\operatorname{Ker}(p_i(T)^{\circ m_i})\cap W=\operatorname{Ker}(p_i(T)^{\circ m_i}).$$
另一方面, 因 $V_i\subseteq W$, 我們有 $(T|_W)|_{V_i}=T|_{V_i}$ 故得
$$\mu_{T|_{V_i}}(x)=\mu_{(T|_W)|_{V_i}}(x)=p_i(x)^{m_i}.$$
故令 $U=V_1$, 利用 Corollary 3.4.7, 得證 $V=U\oplus W=V_1\oplus V_2\oplus\cdots\oplus V_k$.

至於 characteristic polynomial, 利用 Theorem 3.3.9 我們知 $\chi_{T|_{V_i}}(x)=p_i(x)^{e_i}$ 其中 $e_i\geq m_i$. 因此由 Lemma 3.5.5 知
$$p_1(x)^{c_1}\cdots p_k(x)^{c_k}=\chi_T(x)=\chi_{T|_{V_1}}(x)\cdots\chi_{T|_{V_k}}(x)=p_1(x)^{e_1}\cdots p_k(x)^{e_k},$$
利用 $F[x]$ 為 U.F.D. 得證 $e_i=c_i$, 即 $\chi_{T|_{V_i}}(x)=p_i(x)^{c_i},\forall i=1,\ldots,k.$ □

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

Example 3.5.9. 考慮 Example 3.3.10 中的 linear operator $T:P_2(\mathbb{R})\to P_2(\mathbb{R})$. 我們要考慮它的 primary decomposition. 在 Example 3.3.10 中我們知道 T 的 minimal polynomial 為 $\mu_T(x)=(x+1)^2(x-2)$, 因此我們必須找出 $V_1=\operatorname{Ker}((T+\mathrm{id})^{\circ 2})$ 和 $V_2=\operatorname{Ker}(T-2\,\mathrm{id})$. 利用 representative matrix 可以幫助我們找到這兩個 T-invariant subspaces. 我們仍然沿用 ordered basis $\beta=(-x^2+x+1,x+1,1)$. 首先考慮 $([T]_\beta+I_3)^2$ 的 null space, 即解
$$\begin{pmatrix}-9&-9&0\\18&18&0\\0&0&0\end{pmatrix}\cdot\begin{pmatrix}x_1\\x_2\\x_3\end{pmatrix}=\begin{pmatrix}0\\0\\0\end{pmatrix},\quad i.e.\quad\left\{\begin{array}{rrcl}-9x_1&-9x_2&=&0\\18x_1&+18x_2&=&0\end{array}\right..$$

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

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

得

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

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

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

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

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

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

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

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

Example 3.5.10. 考慮 5×5 matrix

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

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

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

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

若令

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

則因

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

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

取

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

可得 block diagonal matrix

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

若令

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

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

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 $N([T]_\beta-I_4)=\{(x_1,x_2,x_2,x_4)^t\mid x_1,x_2,x_4\in F\}$. 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 $N([T]_\beta-(-1)I_4)=\{(0,x_2,-x_2,0)^t\mid x_2\in F\}$. 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 $V_i = \mathrm{Ker}(p_i(T)^{\circ m_i})$, 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

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

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 $V_1=\mathrm{Ker}((T-\lambda\mathrm{id})^{\circ m_1})$. 若 $x-\lambda \mid \mu_T(x)$ 但 $(x-\lambda)^2\nmid \mu_T(x)$, 此即表示 $m_1=1$, 故得 $V_1=\mathrm{Ker}(T-\lambda\mathrm{id})=E_\lambda(T)$. 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 $V_1=\mathrm{Ker}(T-\lambda\mathrm{id})$. 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 $V_i=\mathrm{Ker}(T-\lambda_i\mathrm{id})=E_{\lambda_i}(T)$, $\forall i=1,\ldots,k$. 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 γ_iP_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 $\mathrm{Span}(\{(0,1,0)^{\mathrm{t}}\})$, 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 $V_i = \operatorname{Ker}((T-\lambda_i \mathrm{id})^{\circ m_i})$ 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 $V_1,\dots,V_k$ 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 $\mu_{T_N}(x) = \mathrm{lcm}(x^{m_1},\dots,x^{m_k}) = x^m$, 其中 $m = \max\{m_1,\dots,m_k\}$, 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 $\lambda_1,\ldots,\lambda_k$ 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 $N^m=\mathbf{O}$, $m=\max\{m_1,\ldots,m_k\}$.

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 $\mathbf{w}_1=(0,-1,1)^{\mathrm{t}}$, 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 $\mathbf{w}_2=(-1,1,0)^{\mathrm{t}}$. 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 $\mathbf{w}_3=(0,0,1)^{\mathrm{t}}$. 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 $T,T^{\circ 2},\dots$ 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 $\mathrm{Ker}(T^{\circ i+1})\neq\mathrm{Ker}(T^{\circ i})$, $\forall i=1,\dots,m-1$.

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 $\mathbf{v}_1,\dots,\mathbf{v}_s\in\mathrm{Ker}(T^{\circ i+1})$ 為 linearly independent 且

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

則 $T(\mathbf{v}_1),\dots,T(\mathbf{v}_s)\in\mathrm{Ker}(T^{\circ i})$ 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 $r_1\mathbf{v}_1+\cdots+r_s\mathbf{v}_s\in\mathrm{Ker}(T)\subseteq\mathrm{Ker}(T^{\circ i})$. 故由 $\mathrm{Span}(\mathbf{v}_1,\ldots,\mathbf{v}_s)\cap\mathrm{Ker}(T^{\circ i})=\{\mathbf{O}_V\}$ 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

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

這是因為若 $\mathbf{v}=r_1T(\mathbf{v}_1)+\cdots+r_sT(\mathbf{v}_s)\in\mathrm{Ker}(T^{\circ i-1})$, 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 $r_1\mathbf{v}_1+\cdots+r_s\mathbf{v}_s\in\mathrm{Span}(\mathbf{v}_1,\ldots,\mathbf{v}_s)\cap\mathrm{Ker}(T^{\circ i})=\{\mathbf{O}_V\}$. 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 $\mathbf{v}_1,\ldots,\mathbf{v}_s\in\mathrm{Ker}(T^{\circ i+1})$ 為 *linearly independent* 且 $\mathrm{Span}(\mathbf{v}_1,\ldots,\mathbf{v}_s)\cap\mathrm{Ker}(T^{\circ i})=\{\mathbf{O}_V\}$, 則 $T(\mathbf{v}_1),\ldots,T(\mathbf{v}_s)\in\mathrm{Ker}(T^{\circ i})$ 為 *linearly independent* 且 $\mathrm{Span}(T(\mathbf{v}_1),\ldots,T(\mathbf{v}_s))\cap\mathrm{Ker}(T^{\circ i-1})=\{\mathbf{O}_V\}$.

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

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

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

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

□

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 $\beta=(\mathbf{v}_1,\mathbf{v}_2,\ldots,,\mathbf{v}_{n-1},\mathbf{v}_n)$ 為 V 的 *ordered basis.* 若 $[T]_\beta$ 為 *elementary Jordan matrix associated with* λ, 考慮 *ordered basis* $\beta'=(\mathbf{v}_n,\mathbf{v}_{n-1},\ldots,\mathbf{v}_2,\mathbf{v}_1)$, 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 $i=1,\ldots,m$ 時 S_i 為 $\mathrm{Ker}(T^{\circ i})$ 的 basis (注意 Lemma 4.3.1 告訴我們當 $i=2,\ldots,m$ 時 $S_{i-1}\subsetneq S_i$). 現考慮 $\{\mathbf{v}_1,\ldots,\mathbf{v}_{k_1}\}=S_m\setminus S_{m-1}$ 這組 linear independent subset (它不是空集合). Corollary 1.4.4 告訴我們 $\mathrm{Span}(\{\mathbf{v}_1,\ldots,\mathbf{v}_{k_1}\})$ 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 $\{T(\mathbf{v}_1),\ldots,T(\mathbf{v}_{k_1})\}$ 為 $\mathrm{Ker}(T^{\circ m-1})$ 中的 linearly independent set 且 $\mathrm{Span}(\{T(\mathbf{v}_1),\ldots,T(\mathbf{v}_{k_1})\})\cap\mathrm{Ker}(T^{\circ m-2})=\{\mathbf{O}_V\}$, 故利用 Corollary 1.4.4 知 $\{T(\mathbf{v}_1),\ldots,T(\mathbf{v}_{k_1})\}\cup S_{m-2}$ 為 $\mathrm{Ker}(T^{\circ m-1})$ 中的 linearly independent set. 若 $\{T(\mathbf{v}_1),\ldots,T(\mathbf{v}_{k_1})\}\cup S_{m-2}$ 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 $\{T(\mathbf{v}_1),\ldots,T(\mathbf{v}_{k_1})\}\cup S_{m-2}$ 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 $\mathbf{v}_{k_1+1},\ldots,\mathbf{v}_{k_2}$ 使得

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

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

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

取代. 注意此時 $\{\mathbf{v}_1,\ldots,\mathbf{v}_{k_1},T(\mathbf{v}_1),\ldots,T(\mathbf{v}_{k_1}),\mathbf{v}_{k_1+1},\ldots,\mathbf{v}_{k_2}\}\cup S_{m-2}$ 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 $\{T(\mathbf{v}_1),\ldots,T(\mathbf{v}_{k_1}),\mathbf{v}_{k_1+1},\ldots,\mathbf{v}_{k_2}\}$. 再次利用 Lemma 4.3.2, 我們知 $\{T^{\circ 2}(\mathbf{v}_1),\ldots,T^{\circ 2}(\mathbf{v}_{k_1}),T(\mathbf{v}_{k_1+1}),\ldots,T(\mathbf{v}_{k_2})\}\cup S_{m-3}$ 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 $\mathbf{v}_{k_2+1},\ldots,\mathbf{v}_{k_3}$ 使得

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

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

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

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:

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

最後一個步驟就是將取代 $S_2\setminus S_1$ 的元素

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

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

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

再加入 $\mathrm{Ker}(T)$ 中的元素 $\mathbf{v}_{k_{m-1}+1},\ldots,\mathbf{v}_{k_m}$ 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 $\mathbf{v}_2,T(\mathbf{v}_2),\ldots$ 這樣一直下去最後依序為 $\mathbf{v}_{k_{m-1}},T(\mathbf{v}_{k_{m-1}}),\mathbf{v}_{k_{m-1}+1},\ldots,\mathbf{v}_{k_m}$. 很容易看出 $[T]_\beta$ 便是一個 Jordan block matrix associated with 0. 例如 $[T]_\beta$ 對應 $(\mathbf{v}_1,T(\mathbf{v}_1),\ldots,T^{\circ m-1}(\mathbf{v}_1))$ 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 $(\mathbf{v}_{k_1+1},T(\mathbf{v}_{k_1+1}),\ldots,T^{\circ m-2}(\mathbf{v}_1))$ 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 $(\mathbf{v}_{k_{m-1}+1},\ldots,\mathbf{v}_{k_m})$ 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 $(A-\lambda_iI_n)^{m_i}$ 的 null space $N((A-\lambda_iI_n)^{m_i})$ 的 basis (此即相對於 Proposition 4.3.4 中 $\mathrm{Ker}(T^{\circ m})$ 的 basis), 而 $N((A-\lambda_iI_n)^{m_i})$ 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 $N(A-\lambda_iI_n)$ 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 $S_1=\{(0,0,-1,1,0)^{\mathrm{t}}\}$ 為其 basis. 接著擴大成 $S_2=\{(0,0,-1,1,0)^{\mathrm{t}},(1,0,-1,0,1)^{\mathrm{t}}\}$ 為 $(A-I_5)^2$ 的 null space 的 basis. 然後擴大成 $S_3=\{(0,0,-1,1,0)^{\mathrm{t}},(1,0,-1,0,1)^{\mathrm{t}},(-1,0,0,0,1)^{\mathrm{t}}\}$ 為 $(A-I_5)^3$ 的 null space 的 basis. 因為 $\mathbf{v}_1=(-1,0,0,0,1)^{\mathrm{t}}\in S_3\setminus S_2$, 我們考慮 $\mathbf{v}_2=A\mathbf{v}_1-\mathbf{v}_1=(-1,0,0,1,-1)^{\mathrm{t}}$, 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 $S_1'=\{(0,-1,1,0,0)^{\mathrm{t}}\}$ 為 $A-2I_5$ 的 null space 的 basis. 將之擴大為 $S_2'=\{(0,-1,1,0,0)^{\mathrm{t}},(1,0,0,0,0)\}$ 為 $(A-2I_5)^2$ 的 null space 的 basis. 故考慮 $\mathbf{v}_1'=(1,0,0,0,0)$ 並用 $\mathbf{v}_2'=A\mathbf{v}_1'-2\mathbf{v}_1'=(0,1,-1,0,0)^{\mathrm{t}}$ 取代 $(0,-1,1,0,0)^{\mathrm{t}}$. 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

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

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

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

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 $S=\{\mathbf{v},T(\mathbf{v}),\dots,T^{\circ d-1}(\mathbf{v})\}$ 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

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

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 $\deg(g(x))\le d-1$, 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 $\deg(r(x))\le d-1$ 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 $\deg(\mu_{T|_{C_{\mathbf{v}}}}(x))\le\deg(\chi_{T|_{C_{\mathbf{v}}}}(x))=\deg(\mu_{\mathbf{v}}(x))$ 得證 $\mu_{T|_{C_{\mathbf{v}}}}(x)=\mu_{\mathbf{v}}(x)$. □

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

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

Definition 4.4.5. 假設 V 為一個 finite dimensional F-space, $\mathbf{v}\in V$ 且 $T:V\to V$ 為一個 linear operator. 設 $\deg(\mu_{\mathbf{v}}(x))=d$, 我們稱 $\{\mathbf{v},T(\mathbf{v}),\dots,T^{\circ d-1}(\mathbf{v})\}$ 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 $\beta=(\mathbf{v},T(\mathbf{v}),\dots,T^{\circ d-1}(\mathbf{v}))$ 這一個 cyclic basis 所形成的 ordered basis 的 representative matrix 為何. 由於 $T(\mathbf{v})=0\,\mathbf{v}+1\,T(\mathbf{v})+0\,T^{\circ 2}(\mathbf{v})+\cdots+0\,T^{\circ d-1}(\mathbf{v})$, 我們知道此 matrix 的第一個 column 應為 $(0,1,0,\dots,0)^{\mathrm{t}}$, 同理因 $T(T(\mathbf{v}))=0\,\mathbf{v}+0\,T(\mathbf{v})+1\,T^{\circ 2}(\mathbf{v})+\cdots+0\,T^{\circ d-1}(\mathbf{v})$, 我們知道此 matrix 第二個 column 應為 $(0,0,1,0,\dots,0)^{\mathrm{t}}$, 這樣一直可得到前 $d-1$ 個 column. 至於最後一個 column, 由

於 $T(T^{\circ d-1}(\mathbf{v}))=T^{\circ d}(\mathbf{v})$, 故若 $\mu_{\mathbf{v}}(x)=x^d+a_{d-1}x^{d-1}+\cdots+a_1x+a_0$, 由 $\mu_{\mathbf{v}}(T)(\mathbf{v})=\mathbf{O}_V$, 得 $T^{\circ d}(\mathbf{v})+a_{d-1}T^{\circ d-1}(\mathbf{v})+\cdots+a_1T(\mathbf{v})+a_0\mathbf{v}=\mathbf{O}_V$. 亦即

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

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

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

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

Question 4.18. 試計算

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

Example 4.4.6. 考慮 $T:\mathbb{R}^3\to\mathbb{R}^3$ 定義為 $T(x_1,x_2,x_3)=(2x_1,x_1+x_3,x_1-x_2)$. 考慮 $\mathbf{v}=(0,0,1)$, 則 $T(\mathbf{v})=(0,1,0)$, $T^{\circ 2}(\mathbf{v})=(0,0,-1)=-\mathbf{v}$. 由 $T^{\circ 3}(\mathbf{v})=-T(\mathbf{v})$,... 很容易看出 $C_{\mathbf{v}}=\mathrm{Span}(\{\mathbf{v},T(\mathbf{v})\})=\mathrm{Span}(\{(0,0,1),(0,1,0)\})$. 由 $\{\mathbf{v},T(\mathbf{v})\}$ 為 linearly independent 而 $\{\mathbf{v},T(\mathbf{v}),T^{\circ 2}(\mathbf{v})\}$ 為 linearly dependent 知 $\mathbf{v}$ 的 T-annihilator 為 degree 2 的 polynomial. 事實上因 $T^{\circ 2}(\mathbf{v})=-\mathbf{v}$, 即 $T^{\circ 2}(\mathbf{v})+0T(\mathbf{v})+1\mathbf{v}=\mathbf{O}_V$, 我們有 $\mu_{\mathbf{v}}(x)=x^2+1$. 很容易檢查 $\chi_T(x)=\mu_T(x)=(x-2)(x^2+1)$, 所以確實有 $\mu_{\mathbf{v}}(x)\mid\chi_T(x)$ 以及 $\mu_{\mathbf{v}}(x)\mid\mu_T(x)$. 另外考慮 $\beta=((0,0,1),(0,1,0))$, 我們可以得到 $[T]_\beta=\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$.

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 $g(x) = c_n x^n + \cdots + c_1 x + c_0$, 則對於任意 $\overline{\mathbf{v}} \in V/W$, $g(\overline{T})(\overline{\mathbf{v}}) = c_n \overline{T}^{\circ n}(\overline{\mathbf{v}}) + \cdots + c_1 \overline{T}(\overline{\mathbf{v}}) + c_0 \overline{\mathbf{v}}$. 又因 $\overline{T}^{\circ 2}(\overline{\mathbf{v}}) = \overline{T}(\overline{T}(\overline{\mathbf{v}})) = \overline{T}(\overline{T(\mathbf{v})}) = \overline{T(T(\mathbf{v}))} = \overline{T^{\circ 2}(\mathbf{v})}$, 利用數學歸納法可得 $\overline{T}^{\circ i}(\overline{\mathbf{v}}) = \overline{T^{\circ i}(\mathbf{v})}$, $\forall i \in \mathbb{N}$. 因此得

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

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

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

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

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 $\mu_{\mathbf{u}}(x) = \mu_{\overline{\mathbf{u}}}(x)$. $\square$

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 $m_1 \le m$ 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 $m' \le m$. 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 $2 \le i \le n$, $\mu_{\overline{\mathbf{w}_i}}(x) = p(x)^{m_i}$ 滿足 $m \ge m' = m_2 \ge \cdots \ge m_n$. 又由於 $m_i \le m = m_1$, 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 $1 \le i \le n$, $\{\mathbf{v}_i, T(\mathbf{v}_i), \ldots, T^{\circ dm_i-1}(\mathbf{v}_i)\}$ 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 $\{(1,0,0,0,0)^{\mathrm{t}},(0,1,0,0,0)^{\mathrm{t}}\}$, $\{(-1,0,0,0,1)^{\mathrm{t}},(1,0,1,0,0)^{\mathrm{t}},(-1,1,0,1,0)^{\mathrm{t}}\}$.

由 $\dim(W_1)=2$ 可知 W_1 本身是一個 cyclic space. 事實上取 $\mathbf{w}_1=(1,0,0,0,0)^{\mathrm{t}}$, 則 $A\mathbf{w}_1=(2,1,0,0,0)^{\mathrm{t}}$ (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 $(A-I_5)\mathbf{w}_2\neq(0,0,0,0,0)^{\mathrm{t}}$. 事實上若選 $\mathbf{w}_2=(1,0,1,0,0)^{\mathrm{t}}$, 則 $A\mathbf{w}_2=(1,1,2,1,0)^{\mathrm{t}}$ (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 $\mathbf{w}_3=(-1,0,0,0,1)^{\mathrm{t}}$, 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 $\{\mathbf{v},T(\mathbf{v}),T^{\circ 2}(\mathbf{v}),\ldots,T^{\circ md-1}(\mathbf{v})\}$ 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*

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

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 $c_{0,0},\ldots,c_{i,j},\ldots,c_{m-1,d-1}\in F$ 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. ☐

Question 4.22. *試利用 $T^{\circ d}(\mathbf{v})\in\mathrm{Span}(\mathbf{v},T(\mathbf{v}),\ldots,T^{\circ d-1}(\mathbf{v}),p(T)(\mathbf{v}))$ 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 $N(A^2+I_4)=\mathrm{Span}((0,1,0,1)^{\mathrm{t}},(0,0,1,0)^{\mathrm{t}})$, 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 $\mathrm{Span}((1,1,0)^t,(-1,0,1)^t)$ 為 $B-I_3$ 的 null space, 若令 $\mathbf{w}=(1,0,0)^t$, 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 $\dim(\mathrm{Ker}(q^l(T)|_{C_\mathbf{v}})) = 0$. $\square$

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

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

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

其中 $s_1 = \#\{1 \leq j \leq n_i \mid m_{i,j} = 1\}$. 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 $s_t = \#\{1 \leq j \leq n_i \mid m_{i,j} = t\}$, 則當 $1 \leq l \leq m_i$ 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 $1 \leq l \leq m_i$ 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 $1 \leq l \leq m_i - 1$ 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.

Chapter 5

Operators on Inner Product Spaces

在這一章中我們談論在 inner product spaces 中的 linear operators 的性質. 由於 inner product spaces 比一般 vector spaces 有更豐富的結構, 所以我們可以更深入的探討這情況下的 linear operators. 我們談論的是常用的 inner product spaces, 所以這一章中的 vector spaces 皆為 vector space over $\mathbb{C}$ 或是 $\mathbb{R}$.

5.1. Inner Product Spaces

在這一節中, 我們簡單介紹 inner product space 的定義及基本性質. 首先回顧熟悉的 Real inner product space.

Definition 5.1.1. 令 V 為一個 vector space over $\mathbb{R}$. 若函數 $\langle , \rangle : V \times V \to \mathbb{R}$ 滿足以下的性質, 便稱為 V 的一個 *inner product.*

(1) $\langle \mathbf{v}, \mathbf{w} \rangle = \langle \mathbf{w}, \mathbf{v} \rangle$, $\forall \mathbf{v}, \mathbf{w} \in V$.

(2) $\langle r\mathbf{v} + s\mathbf{w}, \mathbf{u} \rangle = r\langle \mathbf{v}, \mathbf{u} \rangle + s\langle \mathbf{w}, \mathbf{u} \rangle$, $\forall \mathbf{u}, \mathbf{v}, \mathbf{w} \in V$ and $r, s \in \mathbb{R}$.

(3) $\langle \mathbf{v}, \mathbf{v} \rangle \geq 0$, $\forall \mathbf{v} \in V$. 而且 $\langle \mathbf{v}, \mathbf{v} \rangle = 0$ 若且唯若 $\mathbf{v} = \mathbf{O}_V$.

此時我們稱 V 為 *real inner product space.*

至於 complex 的情形, 首先回顧, 若 $z \in \mathbb{C}$, 我們用 $\overline{z}$ 表示 z 的 conjugate (共軛複數).

Definition 5.1.2. 令 V 為一個 vector space over $\mathbb{C}$. 若函數 $\langle , \rangle : V \times V \to \mathbb{C}$ 滿足以下的性質, 便稱為 V 的一個 *inner product.*

(1) $\langle \mathbf{v}, \mathbf{w} \rangle = \overline{\langle \mathbf{w}, \mathbf{v} \rangle}$, $\forall \mathbf{v}, \mathbf{w} \in V$.

(2) $\langle r\mathbf{v} + s\mathbf{w}, \mathbf{u} \rangle = r\langle \mathbf{v}, \mathbf{u} \rangle + s\langle \mathbf{w}, \mathbf{u} \rangle$, $\forall \mathbf{u}, \mathbf{v}, \mathbf{w} \in V$ and $r, s \in \mathbb{C}$.

(3) $\langle \mathbf{v}, \mathbf{v} \rangle \geq 0$, $\forall \mathbf{v} \in V$. 而且 $\langle \mathbf{v}, \mathbf{v} \rangle = 0$ 若且唯若 $\mathbf{v} = \mathbf{O}_V$.

此時我們稱 V 為 *complex inner product space.*

注意, 一個 vector space 可以有不同的 inner product. 當我們說 V 是一個 inner product space, 表示我們已給定某一個 inner product.

Example 5.1.3. 在 $\mathbb{R}^n$ 中我們定義

$$\langle (x_1,\ldots,x_n),(y_1,\ldots,y_n)\rangle = x_1y_1+\cdots+x_ny_n,$$

此為 $\mathbb{R}^n$ 的 *standard inner product.* 在此 inner product 之下, 我們稱 $\mathbb{R}^n$ 為 n-dimensional *Euclidean space.*

在 $\mathbb{C}^n$ 中我們定義

$$\langle (x_1,\ldots,x_n),(y_1,\ldots,y_n)\rangle = x_1\overline{y_1}+\cdots+x_n\overline{y_n},$$

此為 $\mathbb{C}^n$ 的 *standard inner product.* 在此 inner product 之下, 我們稱 $\mathbb{C}^n$ 為 n-dimensional *unitary space.*

Question 5.1. 假設 V 是一個 *over* $\mathbb{C}$ 的 *inner product space.* 若將 V 看成是 *vector space over* $\mathbb{R}$, 是否 V 為 *over* $\mathbb{R}$ 的 *inner product space?*

注意在 real inner product space 的情形, 由於 (1) 的對稱性, 利用 (2) 對於任意的 $\mathbf{u},\mathbf{v},\mathbf{w}\in V$ 以及 $r,s\in\mathbb{R}$ 我們有

$$\langle \mathbf{u}, r\mathbf{v}+s\mathbf{w}\rangle = r\langle \mathbf{u},\mathbf{v}\rangle + s\langle \mathbf{u},\mathbf{w}\rangle.$$

因此對任意的 $\mathbf{v},\mathbf{v}',\mathbf{w},\mathbf{w}'\in V$ 以及 $r,r',s,s'\in\mathbb{R}$ 我們有

$$\begin{aligned}\langle r\mathbf{v}+r'\mathbf{v}', s\mathbf{w}+s'\mathbf{w}'\rangle &= r\langle \mathbf{v}, s\mathbf{w}+s'\mathbf{w}'\rangle + r'\langle \mathbf{v}', s\mathbf{w}+s'\mathbf{w}'\rangle \\ &= rs\langle \mathbf{v},\mathbf{w}\rangle + rs'\langle \mathbf{v},\mathbf{w}'\rangle + r's\langle \mathbf{v}',\mathbf{w}\rangle + r's'\langle \mathbf{v}',\mathbf{w}'\rangle \end{aligned}\tag{5.1}$$

不過在 complex 的情形, 則由 (1), (2) 對於任意的 $\mathbf{u},\mathbf{v},\mathbf{w}\in V$ 以及 $r,s\in\mathbb{C}$ 我們有

$$\langle \mathbf{u}, r\mathbf{v}+s\mathbf{w}\rangle = \overline{\langle r\mathbf{v}+s\mathbf{w},\mathbf{u}\rangle} = \overline{r}\overline{\langle \mathbf{v},\mathbf{u}\rangle} + \overline{s}\overline{\langle \mathbf{w},\mathbf{u}\rangle} = \overline{r}\langle \mathbf{u},\mathbf{v}\rangle + \overline{s}\langle \mathbf{u},\mathbf{w}\rangle.$$

因此對於任意的 $\mathbf{v},\mathbf{v}',\mathbf{w},\mathbf{w}'\in V$ 以及 $r,r',s,s'\in\mathbb{C}$ 我們有

$$\begin{aligned}\langle r\mathbf{v}+r'\mathbf{v}', s\mathbf{w}+s'\mathbf{w}'\rangle &= r\langle \mathbf{v}, s\mathbf{w}+s'\mathbf{w}'\rangle + r'\langle \mathbf{v}', s\mathbf{w}+s'\mathbf{w}'\rangle \\ &= r\overline{s}\langle \mathbf{v},\mathbf{w}\rangle + r\overline{s'}\langle \mathbf{v},\mathbf{w}'\rangle + r'\overline{s}\langle \mathbf{v}',\mathbf{w}\rangle + r'\overline{s'}\langle \mathbf{v}',\mathbf{w}'\rangle \end{aligned}\tag{5.2}$$

在 inner product 的定義中, $\langle \mathbf{v},\mathbf{v}\rangle = 0$ 若且唯若 $\mathbf{v}=\mathbf{O}_V$, 這一個性質稱為 *non-degenerate.* 它可以確保我們有以下之性質.

Lemma 5.1.4. 若 V 是一個 *inner product space* 且 $\mathbf{v}\in V$ 滿足 $\langle \mathbf{v},\mathbf{w}\rangle = 0$, $\forall \mathbf{w}\in V$, 則 $\mathbf{v}=\mathbf{O}_V$.

Proof. 只要選取 $\mathbf{w}=\mathbf{v}$, 則有 $\langle \mathbf{v},\mathbf{v}\rangle = 0$. 故由 inner product 的定義知 $\mathbf{v}=\mathbf{O}_V$. □

Lemma 5.1.4 告訴我們一個判定 V 中元素是否為 $\mathbf{O}_V$ 的方法. 現若 $\mathbf{v},\mathbf{u}\in V$, 滿足 $\langle \mathbf{v},\mathbf{w}\rangle = \langle \mathbf{u},\mathbf{w}\rangle$, $\forall \mathbf{w}\in V$, 則由

$$\langle \mathbf{v}-\mathbf{u},\mathbf{w}\rangle = \langle \mathbf{v},\mathbf{w}\rangle - \langle \mathbf{u},\mathbf{w}\rangle = 0$$

得知 $\mathbf{v}=\mathbf{u}$. 所以我們有以下簡單但有用的性質.

Corollary 5.1.5. 設 V 是一個 *inner product space.* 若 $\mathbf{v},\mathbf{u}\in V$ 滿足 $\langle \mathbf{v},\mathbf{w}\rangle = \langle \mathbf{u},\mathbf{w}\rangle$, $\forall \mathbf{w}\in V$, 則 $\mathbf{v}=\mathbf{u}$.

Lemma 5.1.4 也可以幫助我們了解 linear operator, 以下的性質將來會很有用. 要注意在 real 和 complex 的不同.

Proposition 5.1.6. 設 V 是一個 *inner product space* 且 $T:V\to V$ 為 *linear operator.*

(1) 當 V 是一個 *real inner product space,* 若 $\langle T(\mathbf{v}),\mathbf{w}\rangle = 0$, $\forall \mathbf{v},\mathbf{w}\in V$ 則 T 為 *zero mapping.*

(2) 當 V 是一個 *complex inner product space,* 若 $\langle T(\mathbf{v}),\mathbf{v}\rangle = 0$, $\forall \mathbf{v}\in V$ 則 T 為 *zero mapping.*

Proof. 給定任意 $\mathbf{v}\in V$, 因 $\langle T(\mathbf{v}),\mathbf{w}\rangle = 0$, $\forall \mathbf{w}\in V$, 故由 Lemma 5.1.4 知 $T(\mathbf{v})=\mathbf{O}_V$. 因為對任意 $\mathbf{v}\in V$ 皆成立, 故 $T=\mathbf{O}$.

至於 complex 的情形, 利用等式 (5.2) 對於任意的 $\mathbf{v},\mathbf{w}\in V$ 以及 $r\in\mathbb{C}$ 我們有

$$\begin{aligned} 0 &= \langle T(r\mathbf{v}+\mathbf{w}), r\mathbf{v}+\mathbf{w}\rangle \\ &= \langle rT(\mathbf{v})+T(\mathbf{w}), r\mathbf{v}+\mathbf{w}\rangle \\ &= r\bar{r}\langle T(\mathbf{v}),\mathbf{v}\rangle + r\langle T(\mathbf{v}),\mathbf{w}\rangle + \bar{r}\langle T(\mathbf{w}),\mathbf{v}\rangle + \langle T(\mathbf{w}),\mathbf{w}\rangle \\ &= r\langle T(\mathbf{v}),\mathbf{w}\rangle + \bar{r}\langle T(\mathbf{w}),\mathbf{v}\rangle \end{aligned}$$

當我們分別代 $r=1$ 和 $r=\sqrt{-1}$, 可得 $\langle T(\mathbf{v}),\mathbf{w}\rangle + \langle T(\mathbf{w}),\mathbf{v}\rangle = 0$ 和 $\langle T(\mathbf{v}),\mathbf{w}\rangle - \langle T(\mathbf{w}),\mathbf{v}\rangle = 0$. 依此得 $\langle T(\mathbf{v}),\mathbf{w}\rangle$, $\forall \mathbf{v},\mathbf{w}\in V$. 故由前面結論知 $T=\mathbf{O}$. □

當 V 是一個 inner product space, 不管是 over $\mathbb{R}$ 或是 over $\mathbb{C}$, 我們都會有以下的 Cauchy-Schwarz inequality.

Lemma 5.1.7. 假設 V 是一個 *inner product space over* F, 其中 $F=\mathbb{R}$ 或 $\mathbb{C}$. 若對於任意 $\mathbf{v}\in V$ 我們定義 $\|\mathbf{v}\| = \sqrt{\langle \mathbf{v},\mathbf{v}\rangle}$, 則對於任意 $\mathbf{v},\mathbf{w}\in V$, 皆有

$$|\langle \mathbf{v},\mathbf{w}\rangle| \le \|\mathbf{v}\|\|\mathbf{w}\|.$$

而且 $|\langle \mathbf{v},\mathbf{w}\rangle| = \|\mathbf{v}\|\,\|\mathbf{w}\|$ 若且唯若 $\mathbf{v},\mathbf{w}$ 其中有一個為 $\mathbf{O}_V$ 或是存在 $r\in F$ 使得 $\mathbf{v}=r\mathbf{w}$.

Proof. 當 $\mathbf{v},\mathbf{w}$ 其中有一個為 $\mathbf{O}_V$, 很容易知道等式成立. 所以我們假設 $\mathbf{v},\mathbf{w}$ 皆不為 $\mathbf{O}_V$. 對於任意 $r\in F$, 當 F 為 $\mathbb{R}$ 時利用式子 (5.1) 我們有

$$0 \le \langle \mathbf{v}-r\mathbf{w}, \mathbf{v}-r\mathbf{w}\rangle = \langle \mathbf{v},\mathbf{v}\rangle - 2r\langle \mathbf{v},\mathbf{w}\rangle + r^2\langle \mathbf{w},\mathbf{w}\rangle.$$

此時令 $r=\langle\mathbf{v},\mathbf{w}\rangle/\langle\mathbf{w},\mathbf{w}\rangle$, 我們有
$$\frac{\langle\mathbf{v},\mathbf{w}\rangle^2}{\langle\mathbf{w},\mathbf{w}\rangle}\leq\langle\mathbf{v},\mathbf{v}\rangle.$$
當 F 為 $\mathbb{C}$ 時利用式子 (5.2) 我們有
$$0\leq\langle\mathbf{v}-r\mathbf{w},\mathbf{v}-r\mathbf{w}\rangle=\langle\mathbf{v},\mathbf{v}\rangle-r\langle\mathbf{w},\mathbf{v}\rangle-\bar{r}\langle\mathbf{v},\mathbf{w}\rangle+r\bar{r}\langle\mathbf{w},\mathbf{w}\rangle.$$
此時令 $r=\langle\mathbf{v},\mathbf{w}\rangle/\langle\mathbf{w},\mathbf{w}\rangle$ (注意 $\bar{r}=\overline{\langle\mathbf{v},\mathbf{w}\rangle}/\langle\mathbf{w},\mathbf{w}\rangle$, 因為 $\langle\mathbf{w},\mathbf{w}\rangle\in\mathbb{R}$), 我們有
$$\frac{\langle\mathbf{v},\mathbf{w}\rangle\overline{\langle\mathbf{v},\mathbf{w}\rangle}}{\langle\mathbf{w},\mathbf{w}\rangle}\leq\langle\mathbf{v},\mathbf{v}\rangle.$$
因為在 $F=\mathbb{C}$ 時 $\langle\mathbf{v},\mathbf{w}\rangle\overline{\langle\mathbf{v},\mathbf{w}\rangle}=|\langle\mathbf{v},\mathbf{w}\rangle|^2$ 故得證此 inequality.

最後若等式成立表示存在 $r\in F$ 會使得 $\langle\mathbf{v}-r\mathbf{w},\mathbf{v}-r\mathbf{w}\rangle=0$ 此即 $\mathbf{v}=r\mathbf{w}$. □

給定一個 inner product 之後, 我們就可以定義所謂的 norm. 這是因為我們有以下的性質:

Proposition 5.1.8. 假設 V 是一個 *inner product space over* F *(*$F=\mathbb{R}$ 或 $\mathbb{C}$*).* 若對於任意 $\mathbf{v}\in V$ 我們定義 $\|\mathbf{v}\|=\sqrt{\langle\mathbf{v},\mathbf{v}\rangle}$, 則我們有以下的性質*:*

(1) $\|\mathbf{v}\|\geq 0$ 而且 $\|\mathbf{v}\|=0$ 若且唯若 $\mathbf{v}=\mathbf{O}_V$*.*

(2) 對於任意 $r\in F$ 以及 $\mathbf{v}\in V$*,* 皆有 $\|r\mathbf{v}\|=|r|\|\mathbf{v}\|$*.*

(3) 對於任意 $\mathbf{v},\mathbf{w}\in V$*,* 皆有 $\|\mathbf{v}+\mathbf{w}\|\leq\|\mathbf{v}\|+\|\mathbf{w}\|$*.*

Proof. (1) 直接由 inner product 的性質 (3) 可得, 而 (2) 由式子 (5.1), (5.2) 可得, 所以我們僅證明 (3). 由 $\langle\mathbf{v}+\mathbf{w},\mathbf{v}+\mathbf{w}\rangle=\langle\mathbf{v},\mathbf{v}\rangle+2\langle\mathbf{v},\mathbf{w}\rangle+\langle\mathbf{w},\mathbf{w}\rangle$, 以及 Lemma 5.1.7 我們知
$$\|\mathbf{v}+\mathbf{w}\|^2=\langle\mathbf{v}+\mathbf{w},\mathbf{v}+\mathbf{w}\rangle\leq\|\mathbf{v}\|^2+2\|\mathbf{v}\|\,\|\mathbf{w}\|+\|\mathbf{w}\|^2=(\|\mathbf{v}\|+\|\mathbf{w}\|)^2,$$
得證 $\|\mathbf{v}+\mathbf{w}\|\leq\|\mathbf{v}\|+\|\mathbf{w}\|$. □

Proposition 5.1.8 (3) 的性質就是所謂的三角不等式 (triangle inequality). 一個 vector space V, 若存在一個函數 $\|\;\|:V\to\mathbb{R}$ 滿足 Proposition 5.1.8 的三個性質便稱為 *normed linear space*, 而函數 $\|\;\|$ 便稱作是一個 *norm*. 所以一個 inner product space 可以利用 Proposition 5.1.8 所定出的 norm 使之成為一個 normed linear space. 另外有了 norm 對於任意 $\mathbf{v},\mathbf{w}\in V$, 我們可以定出 $\mathbf{v},\mathbf{w}$ 的距離 (distance) $d(\mathbf{v},\mathbf{w})=\|\mathbf{v}-\mathbf{w}\|$. 而存有 distance 的 vector space 我們稱為 *metric space*. 所以 inner product space 也會是 metric space. 在 metric space 中有了距離, 我們便可定義 sequence 的收斂與發散. 不過這不屬於本講義所要談論的課題, 我們就不多談.

Question 5.2. 假設 V 是一個 *inner product space.* 試利用 *Proposition 5.1.8* 所定出的 *norm* 證明 parallelogram law*,* 即對於任意 $\mathbf{v},\mathbf{w}\in V$ 皆有
$$\|\mathbf{v}+\mathbf{w}\|^2+\|\mathbf{v}-\mathbf{w}\|^2=2\|\mathbf{v}\|^2+2\|\mathbf{w}\|^2.$$

Inner product space 不只讓我們得到 metric space, 另一個重要的性質就是垂直 (orthogonal) 的概念. 我們有以下的定義.

Definition 5.1.9. 假設 V 是一個 inner product space. 若 $\mathbf{v},\mathbf{w}\in V$ 滿足 $\langle\mathbf{v},\mathbf{w}\rangle=0$, 則稱 $\mathbf{v},\mathbf{w}$ 為 *orthogonal*, 用 $\mathbf{v}\perp\mathbf{w}$ 來表示.

若 $\{\mathbf{v}_1,\ldots,\mathbf{v}_n\}$ 為 V 的一組 basis 且對任意 $i\neq j$ 皆有 $\mathbf{v}_i\perp\mathbf{v}_j$, 則我們稱 $\{\mathbf{v}_1,\ldots,\mathbf{v}_n\}$ 為 V 的一組 *orthogonal basis*. 若 orthogonal basis 中任意的 $\mathbf{v}_i$ 皆滿足 $\|\mathbf{v}_i\|=1$, 我們稱之為 V 的一組 *orthonormal basis*.

注意若 $\langle\mathbf{v},\mathbf{w}\rangle=0$, 則 $\langle\mathbf{w},\mathbf{v}\rangle=0$, 所以 $\mathbf{v}\perp\mathbf{w}$ 亦等同於 $\mathbf{w}\perp\mathbf{v}$.

Question 5.3. $\perp$ 是否為一個 *equivalent relation?* 它符合哪些 *equivalent relation* 的條件?

依定義若 $\{\mathbf{w}_1,\ldots,\mathbf{w}_n\}$ 為 V 的一組 orthogonal basis, 令 $\mathbf{v}_i=\frac{1}{\|\mathbf{w}_i\|}\mathbf{w}_i$, 則 $\{\mathbf{v}_1,\ldots,\mathbf{v}_n\}$ 就是 V 的一組 orthonormal basis.

有一組 orthogonal basis (或 orthonormal basis) $\{\mathbf{v}_1,\ldots,\mathbf{v}_n\}$ 的好處便是對於任意 $\mathbf{v}\in V$ 我們可以很快的將 $\mathbf{v}$ 寫成 $\{\mathbf{v}_1,\ldots,\mathbf{v}_n\}$ 的 linear combination. 這是因為若 $\mathbf{v}=c_1\mathbf{v}_1+\cdots+c_n\mathbf{v}_n$, 則利用

$$\langle\mathbf{v},\mathbf{v}_i\rangle=c_1\langle\mathbf{v}_1,\mathbf{v}_i\rangle+\cdots+c_n\langle\mathbf{v}_n,\mathbf{v}_i\rangle=c_i\langle\mathbf{v}_i,\mathbf{v}_i\rangle,$$

可得

$$c_i=\frac{\langle\mathbf{v},\mathbf{v}_i\rangle}{\langle\mathbf{v}_i,\mathbf{v}_i\rangle}.$$

當 V 為 finite dimensional 時, 我們可以利用 *Gram-Schmidt orthogonalization process* 找到 V 的一組 orthogonal basis (或 orthonormal basis). 在此僅略述一下這個 process.

首先任取 $\mathbf{w}_1\in V\setminus\{\mathbf{O}_V\}$ 且令 $\mathbf{v}_1=\mathbf{w}_1$. 接著選取 $\mathbf{w}_2\in V\setminus\mathrm{Span}(\{\mathbf{w}_1\})$, 且令

$$\mathbf{v}_2=\mathbf{w}_2-\frac{\langle\mathbf{w}_2,\mathbf{v}_1\rangle}{\langle\mathbf{v}_1,\mathbf{v}_1\rangle}\mathbf{v}_1.$$

注意此時 $\langle\mathbf{v}_1,\mathbf{v}_2\rangle=0$ 且 $\mathrm{Span}(\{\mathbf{v}_1,\mathbf{v}_2\})=\mathrm{Span}(\{\mathbf{w}_1,\mathbf{w}_2\})$. 接下來如果 $\mathrm{Span}(\{\mathbf{v}_1,\mathbf{v}_2\})=V$, 則 $\{\mathbf{v}_1,\mathbf{v}_2\}$ 就是 V 的一組 orthogonal basis. 否則再找到 $\mathbf{w}_3\in V\setminus\mathrm{Span}(\{\mathbf{w}_1,\mathbf{w}_2\})$, 然後令

$$\mathbf{v}_3=\mathbf{w}_3-\left(\frac{\langle\mathbf{w}_3,\mathbf{v}_1\rangle}{\langle\mathbf{v}_1,\mathbf{v}_1\rangle}\mathbf{v}_1+\frac{\langle\mathbf{w}_3,\mathbf{v}_2\rangle}{\langle\mathbf{v}_2,\mathbf{v}_2\rangle}\mathbf{v}_2\right).$$

注意此時 $\langle\mathbf{v}_1,\mathbf{v}_3\rangle=\langle\mathbf{v}_2,\mathbf{v}_3\rangle=0$ 且 $\mathrm{Span}(\{\mathbf{v}_1,\mathbf{v}_2,\mathbf{v}_3\})=\mathrm{Span}(\{\mathbf{w}_1,\mathbf{w}_2,\mathbf{w}_3\})$. 如此一直下去, 也就是說找到 $\mathbf{w}_i\in V\setminus\mathrm{Span}(\{\mathbf{w}_1,\ldots,\mathbf{w}_{i-1}\})$, 然後令

$$\mathbf{v}_i=\mathbf{w}_i-\left(\frac{\langle\mathbf{w}_i,\mathbf{v}_1\rangle}{\langle\mathbf{v}_1,\mathbf{v}_1\rangle}\mathbf{v}_1+\cdots+\frac{\langle\mathbf{w}_i,\mathbf{v}_{i-1}\rangle}{\langle\mathbf{v}_{i-1},\mathbf{v}_{i-1}\rangle}\mathbf{v}_{i-1}\right).$$

注意此時 $\langle\mathbf{v}_1,\mathbf{v}_i\rangle=\cdots=\langle\mathbf{v}_{i-1},\mathbf{v}_i\rangle=0$ 且 $\mathrm{Span}(\{\mathbf{v}_1,\ldots,\mathbf{v}_i\})=\mathrm{Span}(\{\mathbf{w}_1,\ldots,\mathbf{w}_i\})$. 由於 V 是 finite dimensional, 這個程序一定會停止. 亦即找到 $\{\mathbf{v}_1,\ldots,\mathbf{v}_n\}$ 為 V 的一組 orthogonal basis. 再次強調, 在前面過程中如果我們將每個 $\mathbf{v}_i$ 乘上 $\|\mathbf{v}_i\|^{-1}$ 就得到 orthonormal basis. 另一方面, 若原本已有 V 的一組 basis $\{\mathbf{w}_1,\ldots,\mathbf{w}_n\}$, 則由於 $\mathbf{w}_i\in V\setminus\mathrm{Span}(\{\mathbf{w}_1,\ldots,\mathbf{w}_{i-1}\})$, 所以直接套用上面的 process, 就可以得到 V 的一組 orthogonal basis.

當 W 是 V 的 subspace 時, 我們知道可以找到 W' 為 V 的 subspace 使得 $V = W \oplus W'$, 不過符合這個條件的 W' 並不唯一. 在 inner product space 中, 我們可以加上條件使得 W' 為唯一. 我們需要以下的定義.

Definition 5.1.10. 假設 V 為 inner product space, S 為 V 的一個 nonempty subset. 令

$$S^{\perp} = \{\mathbf{v} \in V \mid \langle \mathbf{v}, \mathbf{w} \rangle = 0, \forall \mathbf{w} \in S\}.$$

我們稱 $S^{\perp}$ 為 the *orthogonal complement* of S in V.

Question 5.4. 什麼是 $\{\mathbf{O}_V\}^{\perp}$? 什麼是 $V^{\perp}$?

關於 $S^{\perp}$ 我們有以下幾個簡單的性質.

Lemma 5.1.11. 假設 V 為 *inner product space.*

(1) 若 S 為 V 的 *nonempty subset,* 則 $S^{\perp}$ 為 V 的 *subspace.*

(2) 若 S_1, S_2 為 V 的 *nonempty subsets* 滿足 $S_1 \subseteq S_2$, 則 $S_2^{\perp} \subseteq S_1^{\perp}$.

(3) 若 S 為 V 的 *nonempty subset,* 則 $S^{\perp} = \mathrm{Span}(S)^{\perp}$.

(4) 若 W 為 V 的 *subspace,* 則 $W \cap W^{\perp} = \{\mathbf{O}_V\}$.

Proof. 令 V 為 inner product space over F (即 $F = \mathbb{C}$ 或 $F = \mathbb{R}$).

(1) 首先依定義 $\mathbf{O}_V \in S^{\perp}$. 又若 $\mathbf{v}_1, \mathbf{v}_2 \in S^{\perp}$, 則對於任意 $r, s \in F$ 皆有 $\langle r\mathbf{v}_1 + s\mathbf{v}_2, \mathbf{w} \rangle = r\langle \mathbf{v}_1, \mathbf{w} \rangle + s\langle \mathbf{v}_2, \mathbf{w} \rangle = 0$, $\forall \mathbf{w} \in S$. 亦即 $r\mathbf{v}_1 + s\mathbf{v}_2 \in S^{\perp}$. 得證 $S^{\perp}$ 為 V 的 subspace.

(2) 若 $\mathbf{v} \in S_2^{\perp}$, 表示對任意 $\mathbf{w} \in S_2$ 皆有 $\langle \mathbf{v}, \mathbf{w} \rangle = 0$. 故對任意 $\mathbf{w} \in S_1$ 因 $S_1 \subseteq S_2$, 知 $\mathbf{w} \in S_2$, 得證 $\mathbf{v} \in S_1^{\perp}$, 即 $S_2^{\perp} \subseteq S_1^{\perp}$.

(3) 因 $S \subseteq \mathrm{Span}(S)$, 故由 (2) 知 $\mathrm{Span}(S)^{\perp} \subseteq S^{\perp}$. 另一方面, 若 $\mathbf{v} \in S^{\perp}$, 則對任意 $\mathbf{w} \in \mathrm{Span}(S)$, 因存在 $c_1, \ldots, c_n \in F$ 以及 $\mathbf{w}_1 \ldots, \mathbf{w}_n \in S$ 使得 $\mathbf{w} = c_1\mathbf{w}_1 + \cdots + c_n\mathbf{w}_n$, 我們有 $\langle \mathbf{w}, \mathbf{v} \rangle = c_1\langle \mathbf{w}_1, \mathbf{v} \rangle + \cdots + c_n\langle \mathbf{w}_n, \mathbf{v} \rangle = 0$. 亦即 $\mathbf{v} \in \mathrm{Span}(S)^{\perp}$, 得證 $S^{\perp} \subseteq \mathrm{Span}(S)^{\perp}$, 故 $S^{\perp} = \mathrm{Span}(S)^{\perp}$.

(4) 若 $\mathbf{v} \in W \cap W^{\perp}$, 表示 $\mathbf{v} \perp \mathbf{v}$, 即 $\langle \mathbf{v}, \mathbf{v} \rangle = 0$. 故由 inner product 的性質知 $\mathbf{v} = \mathbf{O}_V$.

□

Question 5.5. 在 *Lemma 5.1.11 (1)* 中 S 不需假設為 V 的 *subspace,* $S^{\perp}$ 仍為 V 的 *subspace.* 為何在 *(4)* 中需假設 W 為 V 的 *subspace?*

Question 5.6. 試證明若 W_1, W_2 為 V 的 *subspaces,* 則 $(W_1 + W_2)^{\perp} = W_1^{\perp} \cap W_2^{\perp}$.

當 W 是 V 的 subspace 我們可以利用 Gram-Schmidt process 找到 W 的一組 orthogonal basis $S = \{\mathbf{w}_1, \ldots, \mathbf{w}_k\}$. 此時對任意 $\mathbf{v} \in V$ 若令

$$\tilde{\mathbf{v}} = \frac{\langle \mathbf{v}, \mathbf{w}_1 \rangle}{\langle \mathbf{w}_1, \mathbf{w}_1 \rangle}\mathbf{w}_1 + \cdots + \frac{\langle \mathbf{v}, \mathbf{w}_k \rangle}{\langle \mathbf{w}_k, \mathbf{w}_k \rangle}\mathbf{w}_k,$$

我們有 $\tilde{\mathbf{v}} \in W$ 而且對任意 $\mathbf{w}_i$ 皆有 $\langle \mathbf{v} - \tilde{\mathbf{v}}, \mathbf{w}_i \rangle = 0$. 因此由 Lemma 5.1.11 (3) 我們得 $\mathbf{v} - \tilde{\mathbf{v}} \in S^{\perp} = \mathrm{Span}(S)^{\perp} = W^{\perp}$. 因此我們有以下的性質.

Proposition 5.1.12. 假設 V 為 *inner product space* 且 W 為 V 的 *finite dimensional subspace.* 對任意 $\mathbf{v}\in V$, 存在 $\tilde{\mathbf{v}}\in W$ 滿足以下性質.

(1) $\mathbf{v}-\tilde{\mathbf{v}}\in W^{\perp}$

(2) 對於任意 $\mathbf{w}\in W\setminus\{\tilde{\mathbf{v}}\}$, 皆有 $\|\mathbf{v}-\tilde{\mathbf{v}}\|<\|\mathbf{v}-\mathbf{w}\|$.

(3) $\|\tilde{\mathbf{v}}\|\leq\|\mathbf{v}\|$.

Proof. 選定 W 的一組 orthogonal basis $S=\{\mathbf{w}_1,\ldots,\mathbf{w}_k\}$, 對任意 $\mathbf{v}\in V$, 利用前面所述我們有 $\mathbf{v}-\tilde{\mathbf{v}}\in W^{\perp}$. 得證 (1).

現任取 $\mathbf{w}\in W$, 由 $\tilde{\mathbf{v}}-\mathbf{w}\in W$, 我們有

$$\langle\mathbf{v}-\mathbf{w},\mathbf{v}-\mathbf{w}\rangle=\langle\mathbf{v}-\tilde{\mathbf{v}}+\tilde{\mathbf{v}}-\mathbf{w},\mathbf{v}-\tilde{\mathbf{v}}+\tilde{\mathbf{v}}-\mathbf{w}\rangle=\langle\mathbf{v}-\tilde{\mathbf{v}},\mathbf{v}-\tilde{\mathbf{v}}\rangle+\langle\tilde{\mathbf{v}}-\mathbf{w},\tilde{\mathbf{v}}-\mathbf{w}\rangle.$$

故得 $\|\mathbf{v}-\mathbf{w}\|\geq\|\mathbf{v}-\tilde{\mathbf{v}}\|$ 且等號成立若且唯若 $\langle\tilde{\mathbf{v}}-\mathbf{w},\tilde{\mathbf{v}}-\mathbf{w}\rangle=0$, 即 $\mathbf{w}=\tilde{\mathbf{v}}$. 得證 (2).

最後因 $\mathbf{v}-\tilde{\mathbf{v}}\in W^{\perp}$ 且 $\tilde{\mathbf{v}}\in W$, 得

$$\langle\mathbf{v},\mathbf{v}\rangle=\langle\mathbf{v}-\tilde{\mathbf{v}}+\tilde{\mathbf{v}},\mathbf{v}-\tilde{\mathbf{v}}+\tilde{\mathbf{v}}\rangle=\langle\mathbf{v}-\tilde{\mathbf{v}},\mathbf{v}-\tilde{\mathbf{v}}\rangle+\langle\tilde{\mathbf{v}},\tilde{\mathbf{v}}\rangle.$$

故知 $\|\tilde{\mathbf{v}}\|\leq\|\mathbf{v}\|$. 得證 (3). ☐

由 Proposition 5.1.12 (2) 我們知 $\tilde{\mathbf{v}}$ 是 W 中距離 $\mathbf{v}$ 最近的, 所以 $\tilde{\mathbf{v}}$ 是唯一的, 即不會因 W 的 orthogonal basis 的選取不同而有所不同. 我們通常用 $\mathrm{proj}_W(\mathbf{v})$ 來表示 $\tilde{\mathbf{v}}$ 且稱之為 the *projection* of $\mathbf{v}$ on W. 另外我們要強調即使 V 不是 finite dimensional, 只要 W 是 V 的 finite dimensional subspace, Proposition 5.1.12 仍然成立. 不過若 W 不是 finite dimensional, 則 Proposition 5.1.12 就不一定成立了.

Question 5.7. 試證明 *Proposition 5.1.12* 中 $\|\tilde{\mathbf{v}}\|=\|\mathbf{v}\|$ 的充要條件為 $\mathbf{v}\in W$.

現對於任意 $\mathbf{v}\in V$, 我們可寫成 $\mathbf{v}=\mathbf{v}-\mathrm{proj}_W(\mathbf{v})+\mathrm{proj}_W(\mathbf{v})$. 由於 $\mathbf{v}-\mathrm{proj}_W(\mathbf{v})\in W^{\perp}$ 且 $\mathrm{proj}_W(\mathbf{v})\in W$, 我們得 $V=W+W^{\perp}$. 又由 Lemma 5.1.11 (4), $W\cap W^{\perp}=\{\mathbf{O}_V\}$, 我們得證以下定理.

Theorem 5.1.13. 令 V 為 *inner product space* 且 W 為 V 的 *finite dimensional subspace.* 則

$$V=W\oplus W^{\perp}.$$

特別的當 V 本身是 finite dimensional, 任何的 subspace 也是 finite dimensional, 所以 Theorem 5.1.13 對於 V 的任意的 subspace 皆成立. 現若考慮 $W^{\perp}$ 這個 subspace, 我們有 $V=W^{\perp}\oplus(W^{\perp})^{\perp}$. 我們自然會問是否 $W=(W^{\perp})^{\perp}$?

Corollary 5.1.14. 令 V 為 *inner product space* 且 W 為 V 的 *finite dimensional subspace.* 則

$$(W^{\perp})^{\perp}=W.$$

Proof. 若 $\mathbf{w}\in W$, 則對任意 $\mathbf{v}\in W^{\perp}$, 因 $\langle \mathbf{w},\mathbf{v}\rangle = 0$, 得 $\mathbf{w}\in (W^{\perp})^{\perp}$. 得證 $W\subseteq (W^{\perp})^{\perp}$. 另一方面若 $\mathbf{v}\in (W^{\perp})^{\perp}$, 首先利用 Theorem 5.1.13, 我們可將 $\mathbf{v}$ 寫成 $\mathbf{v}=\mathbf{w}+\mathbf{w}'$, 其中 $\mathbf{w}\in W$ 且 $\mathbf{w}'\in W^{\perp}$. 由於 $\mathbf{v}\in (W^{\perp})^{\perp}$, 我們有 $\langle \mathbf{v},\mathbf{w}'\rangle = 0$, 故得

$$0=\langle \mathbf{v},\mathbf{w}'\rangle = \langle \mathbf{w},\mathbf{w}'\rangle + \langle \mathbf{w}',\mathbf{w}'\rangle = \langle \mathbf{w}',\mathbf{w}'\rangle.$$

此即表示 $\mathbf{w}'=\mathbf{O}_V$, 故知 $\mathbf{v}=\mathbf{w}\in W$, 即證得 $(W^{\perp})^{\perp}\subseteq W$. □

Question 5.8. 試證明對於一般 *inner product space* 的 *subspace* W *(*不需假設 *finite dimensional)* 皆有 $W^{\perp}=\big((W^{\perp})^{\perp}\big)^{\perp}$.

Question 5.9. 設 V 為 *finite dimensional inner product space,* S 為 V 的 *subset.* 試問 $(S^{\perp})^{\perp}$ 會是甚麼? 又若 W_1,W_2 為 V 的 *subspace,* 試證 $(W_1\cap W_2)^{\perp}=W_1^{\perp}+W_2^{\perp}$

對於 V 的兩個 subsets S,S', 若對於任意 $\mathbf{v}\in S,\mathbf{v}'\in S'$ 皆有 $\langle \mathbf{v},\mathbf{v}'\rangle=0$, 則用 $S\perp S'$ 來表示. 特別的, 若 W,W' 為 V 的 subspaces 且 $W\perp W'$, 則 $W'\subseteq W^{\perp}$, 故利用 Lemma 5.1.11 可得 $W\cap W'=\{\mathbf{O}_V\}$. 現若 $W_1,\ldots,W_k$ 為 V 的 subspaces 滿足 $V=W_1+\cdots+W_k$, 且對於任意 $i\neq j$, 皆有 $W_i\perp W_j$, 則 V 為 $W_1,\ldots,W_k$ 的 direct sum. 我們特別稱此為 V 的一種 *orthogonal direct sum* 且將這種 direct sum 用

$$V=W_1\boxplus\cdots\boxplus W_k$$

來表示. 例如當 W 為 V 的 finite dimensional subspace, Theorem 5.1.13 告訴我們

$$V=W\boxplus W^{\perp}.$$

5.2. Dual Spaces

Dual space 的概念和 inner product space 的概念有許多相關性, 而且有許多 inner product space 的性質用 dual space 來描述較為清楚. 所以在這一節中我們特別介紹 dual space.

Definition 5.2.1. 假設 V 是一個 vector space over F, 若 $f:V\to F$ 為一個 F-linear transformation, 則稱 f 為一個 *linear functional* on V. 所有的 linear functional on V 形成一個 vector space over F, 我們稱之為 V 的 *dual space*, 用 V^* 來表示.

Question 5.10. 考慮 $n\times n$ 矩陣的 *determinant* 函數 $\det: M_n(F)\to F$ 以及 *trace* 函數 $\mathrm{tr}: M_n(F)\to F$. 哪一個是 *linear functional on* $M_n(F)$*?*

回顧一下給定 V 的一組 basis, $\{\mathbf{v}_1,\ldots,\mathbf{v}_n\}$ 任意選取 $\mathbf{w}_1,\ldots,\mathbf{w}_n\in W$, 我們可找到唯一的 linear transformation $T:V\to W$ 滿足 $T(\mathbf{v}_i)=\mathbf{w}_i$, $\forall i=1,\ldots,n$. 現對於任意 $i=1,\ldots,n$ 我們考慮 $\mathbf{v}_i^*:V\to F$ 為唯一的 linear function on V, 滿足

$$\mathbf{v}_i^*(\mathbf{v}_j)=\begin{cases}1, & \text{if } j=i;\\ 0, & \text{if } j\neq i.\end{cases}$$

我們有以下性質.

Theorem 5.2.2. 假設 $\{\mathbf{v}_1,\ldots,\mathbf{v}_n\}$ 為 V 的一組 *basis,* 則 $\{\mathbf{v}_1^*,\ldots,\mathbf{v}_n^*\}$ 為 V^* 的一組 *basis.* 特別的, 我們有 $\dim(V)=\dim(V^*)$.

Proof. 首先證明 $\mathrm{Span}(\{\mathbf{v}_1^*,\ldots,\mathbf{v}_n^*\})=V^*$. 也就是說對於任意 $f\in V^*$, 皆存在 $c_1,\ldots,c_n\in F$ 滿足 $f=c_1\mathbf{v}_1^*+\cdots+c_n\mathbf{v}_n^*$. 由前述 linear transformation 可以由 basis 所唯一確定的性質, 我們僅要找到 $c_1,\ldots,c_n\in F$ 使得 f 和 $c_1\mathbf{v}_1^*+\cdots+c_n\mathbf{v}_n^*$ 在每一個 $\mathbf{v}_i$ 的取值皆相同. 然而對每個 $\mathbf{v}_i$, 我們有

$$(c_1\mathbf{v}_1^*+\cdots+c_n\mathbf{v}_n^*)(\mathbf{v}_i)=c_1\mathbf{v}_1^*(\mathbf{v}_i)+\cdots+c_n\mathbf{v}_n^*(\mathbf{v}_i)=c_i\mathbf{v}_i^*(\mathbf{v}_i)=c_i. \tag{5.3}$$

所以若令 $c_i=f(\mathbf{v}_i)$, 便可得 $f=c_1\mathbf{v}_1^*+\cdots+c_n\mathbf{v}_n^*$.

接著證明 $\{\mathbf{v}_1^*,\ldots,\mathbf{v}_n^*\}$ 為 linearly independent. 假設 $c_1\mathbf{v}_1^*+\cdots+c_n\mathbf{v}_n^*=\mathbf{O}$ 為 zero mapping. 亦即對任意 $\mathbf{v}_i$, 皆有 $(c_1\mathbf{v}_1^*+\cdots+c_n\mathbf{v}_n^*)(\mathbf{v}_i)=0$, 故由式子 (5.3) 得知 $c_i=0$, $\forall i=1,\ldots,n$. □

給定 V 的一組 basis $\{\mathbf{v}_1,\ldots,\mathbf{v}_n\}$, 我們稱 $\{\mathbf{v}_1^*,\ldots,\mathbf{v}_n^*\}$ 為對應 $\{\mathbf{v}_1,\ldots,\mathbf{v}_n\}$ 的 *dual basis*.

Question 5.11. 假設 $\{\mathbf{v}_1,\ldots,\mathbf{v}_n\}$ 為 V 的一組 *basis,* 對於 $\mathbf{v}\in V$, $\mathbf{v}_i^*(\mathbf{v})$ 為何?

既然 V^* 亦為 F-space, 我們自然會問 V^* 的 dual space 為何? 即 $(V^*)^*$ (稱為 V 的 *double dual space*). 依定義 $(V^*)^*$ 中的元素為 linear functional on V^*. 也就是說若 $\sigma\in(V^*)^*$, 則 $\sigma:V^*\to F$ 為一個 linear transformation 將任意的 $f\in V^*$ 送到一個 F 值. 特別的, 若 $\mathbf{v}\in V$, 我們可以考慮 $\hat{\mathbf{v}}:V^*\to F$ 其定義為對任意 $f\in V^*$, $\hat{\mathbf{v}}(f)=f(\mathbf{v})$. 要說明 $\hat{\mathbf{v}}\in(V^*)^*$, 我們必須說明 $\hat{\mathbf{v}}$ 為 linear functional, 即對於任意 $f,g\in V^*$ 以及 $r,s\in F$, 我們有

$$\hat{\mathbf{v}}(rf+sg)=(rf+sg)(\mathbf{v})=rf(\mathbf{v})+sg(\mathbf{v})=r\hat{\mathbf{v}}(f)+s\hat{\mathbf{v}}(g).$$

Theorem 5.2.2 告訴我們當 V 是 finite dimensional 時, $\dim(V)=\dim(V^*)$, 所以亦得 $\dim(V^*)=\dim((V^*)^*)$. 也就是說 $\dim(V)=\dim((V^*)^*)$. 當然此時 V 和 $(V^*)^*$ 為 isomorphic, 事實上我們可以利用 $\hat{\mathbf{v}}$ 建構出 V 和 $(V^*)^*$ 間的一個重要 isomorphism, 我們稱為 V 和 $(V^*)^*$ 的 *canonical map*

Proposition 5.2.3. 考慮 $\tau:V\to(V^*)^*$ 定義為 $\tau(\mathbf{v})=\hat{\mathbf{v}}$, $\forall\mathbf{v}\in V$. 則 τ 為一個 *one-to-one* 的 *linear transformation.* 特別當 V 為 *finite dimensional* 時, τ 為一個 *isomorphism.*

Proof. 首先證明 τ 為 linear transformation. 這是因為對任意 $\mathbf{v},\mathbf{w}\in V$ 以及 $r,s\in F$ 我們有

$$\tau(r\mathbf{v}+s\mathbf{w})(f)=f(r\mathbf{v}+s\mathbf{w})=rf(\mathbf{v})+sf(\mathbf{w})=(r\tau(\mathbf{v})+s\tau(\mathbf{w}))(f),\quad\forall f\in V^*,$$

此即表示 $\tau(r\mathbf{v}+s\mathbf{w})$ 和 $r\tau(\mathbf{v})+s\tau(\mathbf{w})$ 為定義在 V^* 上的同樣函數, 故知 $\tau(r\mathbf{v}+s\mathbf{w})=r\tau(\mathbf{v})+s\tau(\mathbf{w})$.

接著證明 τ 為 one-to-one, 即證明 $\mathrm{Ker}(\tau)=\mathbf{O}_V$. 現假設 $\mathbf{v}\in\mathrm{Ker}(\tau)$, 即 $\tau(\mathbf{v})=\hat{\mathbf{v}}$ 為定義在 V^* 上的 zero mapping. 亦即對任意 $f\in V^*$ 皆有 $0=\hat{\mathbf{v}}(f)=f(\mathbf{v})$. 然而若 $\mathbf{v}\neq\mathbf{O}_V$, 我們一定可以找到 $f\in V^*$ 使得 $f(\mathbf{v})\neq0$. 由此矛盾得證 $\mathbf{v}=\mathbf{O}_V$.

最後若 V 為 finite dimensional, 由 $\tau: V \to (V^*)^*$ 為 one-to-one 以及 $\dim(V) = \dim((V^*)^*)$, 得證 τ 為 onto, 亦即 τ 為 isomorphism. □

在 dual space 中有一個和 orthogonal complement 類似的概念, 我們介紹如下.

Definition 5.2.4. 假設 V 為 vector space, S 為 V 的一個 nonempty subset. 令

$$S^0 = \{f \in V^* \mid f(\mathbf{v}) = 0, \forall \mathbf{v} \in S\}.$$

我們稱 S^0 為 the *annihilator* of S.

Question 5.12. 什麼是 $\{\mathbf{O}_V\}^0$? 什麼是 V^0?

關於 S^0 我們有以下幾個類似 Lemma 5.1.11 的性質.

Lemma 5.2.5. 假設 V 為 *vector space over* F.

(1) 若 S 為 V 的 *nonempty subset,* 則 S^0 為 V^* 的 *subspace.*
(2) 若 S_1, S_2 為 V 的 *nonempty subsets* 滿足 $S_1 \subseteq S_2$, 則 $S_2^0 \subseteq S_1^0$.
(3) 若 S 為 V 的 *nonempty subset,* 則 $S^0 = \mathrm{Span}(S)^0$.

Proof.

(1) 首先依定義 V^* 中的 zero mapping $\mathbf{O}$ 在 S^0. 又若 $f, g \in S^0$, 則對於任意 $r, s \in F$ 皆有 $rf + sg(\mathbf{v}) = rf(\mathbf{v}) + sg(\mathbf{v}) = 0$, $\forall \mathbf{v} \in S$. 亦即 $rf + sg \in S^0$. 得證 S^0 為 V^* 的 subspace.
(2) 若 $f \in S_2^0$, 表示對任意 $\mathbf{w} \in S_2$ 皆有 $f(\mathbf{w}) = 0$. 故對任意 $\mathbf{w} \in S_1$ 因 $S_1 \subseteq S_2$, 知 $\mathbf{w} \in S_2$, 得證 $f \in S_1^0$, 即 $S_2^0 \subseteq S_1^0$.
(3) 因 $S \subseteq \mathrm{Span}(S)$, 故由 (2) 知 $\mathrm{Span}(S)^0 \subseteq S^0$. 另一方面, 若 $f \in S^0$, 則對任意 $\mathbf{w} \in \mathrm{Span}(S)$, 因存在 $c_1, \ldots, c_n \in F$ 以及 $\mathbf{w}_1 \ldots, \mathbf{w}_n \in S$ 使得 $\mathbf{w} = c_1\mathbf{w}_1 + \cdots + c_n\mathbf{w}_n$, 我們有 $f(\mathbf{w}) = c_1 f(\mathbf{w}_1) + \cdots + c_n f(\mathbf{w}_n) = 0$. 亦即 $f \in \mathrm{Span}(S)^0$, 得證 $S^0 \subseteq \mathrm{Span}(S)^0$, 故 $S^0 = \mathrm{Span}(S)^0$.

□

當 V 是 finite dimensional inner product space, 若 W 為 V 的 subspace, 則 Theorem 5.1.13 告訴我們 $\dim(W^\perp) = \dim(V) - \dim(W)$. 關於 annihilator 我們也有相同的性質.

Proposition 5.2.6. 假設 $V = W \oplus U$. 則存在一個 *isomorphism* $\phi: U^* \to W^0$. 特別當 V 為 *finite dimensional,* 則對任意的 *subspace* W 皆有 $\dim(W^0) = \dim(V) - \dim(W)$.

Proof. 由 direct sum 的性質, 對任意 $\mathbf{v} \in V$, 存在唯一的 $\mathbf{w} \in W, \mathbf{u} \in U$ 使得 $\mathbf{v} = \mathbf{w} + \mathbf{u}$. 故對任意 $f \in U^*$, 我們定義 $\phi(f): V \to F$ 為 $\phi(f)(\mathbf{w} + \mathbf{u}) = f(\mathbf{u})$. 因 f 是 linear functional 很容易檢查 $\phi(f)$ 亦為 linear functional, 亦即 $\phi(f) \in V^*$. 另依此定義對於所有 $\mathbf{w} \in W$, 因 $\mathbf{w} = \mathbf{w} + \mathbf{O}_U$, 故得 $\phi(f)(\mathbf{w}) = f(\mathbf{O}_U) = 0$, 亦即 $\phi(f) \in W^0$. 所以 $\phi: U^* \to W^0$ 是一個

well-defined 的 function. 很容易驗證 ϕ 為 linear transformation. 又若 $f,g \in U^*$ 且 $f \neq g$, 表示存在 $\mathbf{u} \in U$ 使得 $f(\mathbf{u}) \neq g(\mathbf{u})$, 因此 $\phi(f)(\mathbf{u}) \neq \phi(g)(\mathbf{u})$, 得證 $\phi(f) \neq \phi(g)$, 亦即 ϕ 為 one-to-one. 至於 ϕ 是 onto 的原因是因為對任意 $f \in W^0$, 我們皆有 $f(\mathbf{v}) = f(\mathbf{w}+\mathbf{u}) = f(\mathbf{u})$, 其中 $\mathbf{w} \in W, \mathbf{u} \in U$ 滿足 $\mathbf{v} = \mathbf{w}+\mathbf{u}$. 所以若考慮 $f|_U \in U^*$, 我們有 $\phi(f|_U) = f$, 得證 ϕ 為 onto.

現若 V 為 finite dimensional, 對於任意 subspace W, 我們都可將 W 的一組 basis $\{\mathbf{w}_1,\ldots,\mathbf{w}_k\}$ 擴大成 V 的一組 basis $\{\mathbf{w}_1,\ldots,\mathbf{w}_k,\mathbf{u}_1,\ldots,\mathbf{u}_l\}$. 所以我們有 $V = W \oplus U$, 其中 $U = \mathrm{Span}(\{\mathbf{u}_1,\ldots,\mathbf{u}_l\})$. 由 Theorem 5.2.2 以及套用上面的結果, 我們有

$$\dim(W^0) = \dim(U^*) = \dim(U) = \dim(V) - \dim(W).$$

$\square$

Question 5.13. 試證明若 W_1, W_2 為 V 的 *subspaces,* 則 $(W_1+W_2)^0 = W_1^0 \cap W_2^0$. 又若 V 為 *finite dimensional,* 則 $(W_1 \cap W_2)^0 = W_1^0 + W_2^0$.

既然 W^0 為 V^* 的 subspace, 我們也可以問 W^0 的 annihilator, 即 $(W^0)^0$. 我們有如以下類似 Corollary 5.1.14 的性質.

Corollary 5.2.7. 假設 V 是一個 *finite dimensional vector space,* 且 W 為 V 的 *subspace.* 考慮 *canonical map* $\tau: V \to (V^*)^*$ 定義為 $\tau(\mathbf{v}) = \hat{\mathbf{v}}$, $\forall \mathbf{v} \in V$, 則 $\tau(W) = (W^0)^0$.

Proof. 設 $\mathbf{w} \in W$, 則對於任意 $f \in W^0$, 我們有 $\hat{\mathbf{w}}(f) = f(\mathbf{w}) = 0$, 亦即 $\tau(\mathbf{w}) = \hat{\mathbf{w}} \in (W^0)^0$. 得知 $\tau(W) \subseteq (W^0)^0$. 由 Proposition 5.2.3, 我們知道 $\tau: V \to (V^*)^*$ 為 one-to-one, 故 $\dim(\tau(W)) = \dim(W)$. 另一方面 Proposition 5.2.6 告訴我們

$$\dim((W^0)^0) = \dim(V^*) - \dim(W^0) = \dim(V) - (\dim(V) - \dim(W)) = \dim(W).$$

故知 $\dim(\tau(W)) = \dim((W^0)^0)$, 得證 $\tau(W) = (W^0)^0$. $\square$

最後我們來探討 dual space 和 inner product space 間的關係. 首先我們需要以下的定義.

Definition 5.2.8. 令 V, W 為 vector spaces over F, 其中 F 為 $\mathbb{C}$ 或 $\mathbb{R}$. 若 $T: V \to W$ 符合對所有 $\mathbf{v}, \mathbf{v}' \in V$, $r \in F$ 皆有 $T(\mathbf{v}+\mathbf{v}') = T(\mathbf{v}) + T(\mathbf{v}')$ 以及 $T(r\mathbf{v}) = \overline{r}T(\mathbf{v})$, 則稱 T 為一個 *conjugate transformation.* 又若 T 是 one-to-one and onto, 則稱之為 *conjugate isomorphism.*

注意當 $F = \mathbb{R}$ 時, conjugate transformation 就是 linear transformation. 其實當 $F = \mathbb{C}$ 時 conjugate transformation 和 linear transformation 有許多相似之處, 我們介紹幾個有關 conjugate transformation 的性質.

Lemma 5.2.9. 假設 V, W, U 皆為 *vector spaces over* $\mathbb{C}$. $T_1: V \to W$, $T_2: W \to U$.

(1) 若 T_1, T_2 中有一個是 *linear transformation* 另一個是 *conjugate transformation,* 則 $T_2 \circ T_1: V \to U$ 為 *conjugate transformation.*

(2) 若 T_1, T_2 皆為 *conjugate transformation,* 則 $T_2 \circ T_1: V \to U$ 為 *linear transformation.*

(3) 若 T_1 為 *conjugate isomorphism,* 則 $T_1^{-1}: W \to V$ 亦為 *conjugate isomorphism.*

Proof.

(1) 我們僅證明 T_1 為 conjugate transformation 且 T_2 為 linear transformation 的情況, 另一情況類似, 請自行驗證. 我們需證明對任意 $\mathbf{v}, \mathbf{v}' \in V$ 以及 $r, s \in \mathbb{C}$ 皆有 $T_2 \circ T_1(r\mathbf{v} + s\mathbf{v}') = \bar{r}T_2 \circ T_1(\mathbf{v}) + \bar{s}T_2 \circ T_1(\mathbf{v}')$. 事實上

$$T_2 \circ T_1(r\mathbf{v} + s\mathbf{v}') = T_2(\bar{r}T_1(\mathbf{v}) + \bar{s}T_1(\mathbf{v}')) = \bar{r}T_2(T_1(\mathbf{v})) + \bar{s}T_2(T_1(\mathbf{v}')) = \bar{r}T_2 \circ T_1(\mathbf{v}) + \bar{s}T_2 \circ T_1(\mathbf{v}').$$

故 $T_2 \circ T_1$ 為 conjugate transformation.

(2) 若 T_1, T_2 皆為 conjugate transformation, 我們證明對任意 $\mathbf{v}, \mathbf{v}' \in V$ 以及 $r, s \in \mathbb{C}$ 皆有 $T_2 \circ T_1(r\mathbf{v} + s\mathbf{v}') = rT_2 \circ T_1(\mathbf{v}) + sT_2 \circ T_1(\mathbf{v}')$. 事實上

$$T_2 \circ T_1(r\mathbf{v} + s\mathbf{v}') = T_2(\bar{r}T_1(\mathbf{v}) + \bar{s}T_1(\mathbf{v}')) = \overline{(\bar{r})}T_2(T_1(\mathbf{v})) + \overline{(\bar{s})}T_2(T_1(\mathbf{v}')) = rT_2 \circ T_1(\mathbf{v}) + sT_2 \circ T_1(\mathbf{v}').$$

故 $T_2 \circ T_1 : V \to U$ 為 linear transformation.

(3) 若 T_1 為 conjugate isomorphism, 因 T_1 為 one-to-one and onto, $T_1^{-1}: W \to V$ 存在且滿足 $T_1(T_1^{-1}(\mathbf{w})) = \mathbf{w}$, $\forall \mathbf{w} \in W$. 現對於任意 $\mathbf{w}, \mathbf{w}' \in W$ 以及 $r, s \in \mathbb{C}$, 我們有 $T_1(T_1^{-1}(r\mathbf{w} + s\mathbf{w}')) = r\mathbf{w} + s\mathbf{w}'$ 以及

$$T_1(\bar{r}T^{-1}(\mathbf{w}) + \bar{s}T^{-1}(\mathbf{w}')) = \overline{(\bar{r})}T_1(T^{-1}(\mathbf{w})) + \overline{(\bar{s})}T_1(T^{-1}(\mathbf{w}')) = r\mathbf{w} + s\mathbf{w}',$$

故得 $T_1(T^{-1}(r\mathbf{w} + s\mathbf{w}')) = T_1(\bar{r}T^{-1}(\mathbf{w}) + \bar{s}T^{-1}(\mathbf{w}'))$. 因為 T_1 為 one-to-one, 此即表示

$$T^{-1}(r\mathbf{w} + s\mathbf{w}') = \bar{r}T^{-1}(\mathbf{w}) + \bar{s}T^{-1}(\mathbf{w}').$$

故得證 T_1^{-1} 亦為 conjugate isomorphism. □

現若 V 為 inner product space over F, 給定 $\mathbf{v} \in V$, 考慮以下函數 $\rho(\mathbf{v}) : V \to F$ 定義為 $\mathbf{w} \mapsto \langle \mathbf{w}, \mathbf{v} \rangle$, (即 $\rho(\mathbf{v}) = \langle \cdot, \mathbf{v} \rangle$) 則 $\rho(\mathbf{v})$ 為 linear functional on V, 亦即 $\rho(\mathbf{v}) \in V^*$. 所以 ρ 給了我們一個從 V 到 V^* 的 mapping. 現對於任意 $\mathbf{v}, \mathbf{v}' \in V$, 我們有

$$\rho(\mathbf{v} + \mathbf{v}')(\mathbf{w}) = \langle \mathbf{w}, \mathbf{v} + \mathbf{v}' \rangle = \langle \mathbf{w}, \mathbf{v} \rangle + \langle \mathbf{w}, \mathbf{v}' \rangle = \rho(\mathbf{v})(\mathbf{w}) + \rho(\mathbf{v}')(\mathbf{w}), \quad \forall \mathbf{w} \in V.$$

得知 $\rho(\mathbf{v} + \mathbf{v}') = \rho(\mathbf{v}) + \rho(\mathbf{v}')$ in V^*. 另外若 $r \in F$, 則

$$\rho(r\mathbf{v})(\mathbf{w}) = \langle \mathbf{w}, r\mathbf{v} \rangle = \bar{r}\langle \mathbf{w}, \mathbf{v} \rangle = \bar{r}\rho(\mathbf{v})(\mathbf{w}), \quad \forall \mathbf{w} \in V.$$

得知 $\rho(r\mathbf{v}) = \bar{r}\rho(\mathbf{v})$ in V^*. 我們有以下的定理.

Proposition 5.2.10. 設 V 為 *finite dimensional inner product space.* 考慮 $\rho : V \to V^*$ 定義為 $\rho(\mathbf{v}) = \langle \cdot, \mathbf{v} \rangle$, 則 ρ 為 *conjugate isomorphism.*

Proof. 我們已知 ρ 為 conjugate transformation, 所以僅要證明 ρ 為 one-to-one and onto.

首先證明 ρ 為 one-to-one. 假設 $\mathbf{v}, \mathbf{v}' \in V$ 滿足 $\rho(\mathbf{v}) = \rho(\mathbf{v}')$, 此即表示 $\langle \mathbf{w}, \mathbf{v} \rangle = \langle \mathbf{w}, \mathbf{v}' \rangle$, $\forall \mathbf{w} \in V$. 然而 Corollary 5.1.5 告訴我們此即表示 $\mathbf{v} = \mathbf{v}'$, 得證 ρ 為 one-to-one.

至於 ρ 是 onto, 我們首先選定 V 的一組 orthonormal basis $\{\mathbf{v}_1,\ldots,\mathbf{v}_n\}$. 對任意 $f\in V^*$, 考慮 $\mathbf{v}=\overline{f(\mathbf{v}_1)}\mathbf{v}_1+\cdots+\overline{f(\mathbf{v}_n)}\mathbf{v}_n$. 則

$$\rho(\mathbf{v})(\mathbf{v}_i)=\langle\mathbf{v}_i,\overline{f(\mathbf{v}_1)}\mathbf{v}_1+\cdots+\overline{f(\mathbf{v}_n)}\mathbf{v}_n\rangle=\langle\mathbf{v}_i,\overline{f(\mathbf{v}_i)}\mathbf{v}_i\rangle=f(\mathbf{v}_i),\quad\forall i\in\{1,\ldots,n\}.$$

也就是說 f 和 $\rho(\mathbf{v})$ 在 $\{\mathbf{v}_1,\ldots,\mathbf{v}_n\}$ 這一組 basis 的取值皆相同, 亦即 $f=\rho(\mathbf{v})$. 得證 ρ 為 onto. □

注意因為 ρ 為 conjugate isomorphism, 由 Lemma 5.2.9 知 $\rho^{-1}:V^*\to V$ 亦為 conjugate isomorphism.

Question 5.14. 在 *Proposition 5.2.10* 中可以利用 $\rho:V\to V^*$ 為 *one-to-one* 以及 $\dim(V)=\dim(V^*)$ 來說明 $\rho:V\to V^*$ 為 *onto* 嗎?

Example 5.2.11. 考慮 $\mathbb{R}^n$ 上的 standard inner product, 令 $\{\mathbf{e}_1,\ldots,\mathbf{e}_n\}$ 為 $\mathbb{R}^n$ 的 standard basis. 若 $\mathbf{v}=x_1\mathbf{e}_1+\cdots+x_n\mathbf{e}_n\in\mathbb{R}^n$, 則 $\rho(\mathbf{v})(\mathbf{e}_i)=\langle\mathbf{e}_i,\mathbf{v}\rangle=x_i$. 也就是說若令 $\rho(\mathbf{v})=f\in(\mathbb{R}^n)^*$, 則 $f(\mathbf{e}_i)=x_i$, $\forall i\in\{1,\ldots,n\}$. 反過來, 若 $f\in(\mathbb{R}^n)^*$, 則 $\rho^{-1}(f)=f(\mathbf{e}_1)\mathbf{e}_1+\cdots+f(\mathbf{e}_n)\mathbf{e}_n$.

考慮 $\mathbb{C}^n$ 上的 standard inner product, 令 $\{\mathbf{e}_1,\ldots,\mathbf{e}_n\}$ 為 $\mathbb{C}^n$ 的 standard basis. 若 $\mathbf{v}=z_1\mathbf{e}_1+\cdots+z_n\mathbf{e}_n\in\mathbb{C}^n$, 則 $\rho(\mathbf{v})(\mathbf{e}_i)=\langle\mathbf{e}_i,\mathbf{v}\rangle=\overline{z_i}$. 也就是說若令 $\rho(\mathbf{v})=f\in(\mathbb{C}^n)^*$, 則 $f(\mathbf{e}_i)=\overline{z_i}$, $\forall i\in\{1,\ldots,n\}$. 反過來, 若 $f\in(\mathbb{C}^n)^*$, 則 $\rho^{-1}(f)=\overline{f(\mathbf{e}_1)}\mathbf{e}_1+\cdots+\overline{f(\mathbf{e}_n)}\mathbf{e}_n$.

5.3. Transpose and Adjoint

雖然以後我們關心的是 linear operator, 不過有關於 transpose 和 adjoint 的性質, 用一般的 linear transformation 較容易看出. 所以這一節中我們考慮的是一般的 linear transformation. 給定一個 linear transformation $T:V\to W$, 在 dual spaces W^*,V^* 上我們可以定義 T 的 transpose, 而若 V,W 為 inner product space, 我們也可以定義 T 的 adjoint. 我們將探討 transpose 和 adjoint 之間的關係.

首先我們探討一個 linear transformation $T:V\to W$ 的 transpose, 其中 V,W 皆為 vector space over F. 再次強調有關 T 的 transpose 不需假設 V,W 為 inner product space 而且也不需假設為 finite dimensional. 現對於任意 $f\in W^*$, 由於 T,f 皆為 linear transformation, 所以 $f\circ T:V\to F$ 亦為 linear transformation. 也就是說 $f\circ T\in V^*$. 所以我們可定義一個 mapping $T^{\mathrm{t}}:W^*\to V^*$, 其定義為 $T^{\mathrm{t}}(f)=f\circ T$, $\forall f\in W^*$. 依此定義, 對於任意 $f,g\in W^*$, $r,s\in F$ 我們有

$$T^{\mathrm{t}}(rf+sg)=(rf+sg)\circ T=r(f\circ T)+s(g\circ T)=rT^{\mathrm{t}}(f)+sT^{\mathrm{t}}(g),$$

故知 $T^{\mathrm{t}}:W^*\to V^*$ 為 linear transformation. 我們稱 T^{t} 為 T 的 *transpose*. 依定義 T 和 T^{t} 的關係為, 對任意 $\mathbf{v}\in V$, $f\in W^*$ 我們有

$$f(T(\mathbf{v}))=T^t(f)(\mathbf{v}).$$

接下來我們列出 transpose 的基本性質.

Proposition 5.3.1. 設 V,W,U 皆為 *vector space over* F.

(1) 若 $r,s\in F$ 且 $T_1:V\to W$, $T_2:V\to W$ 皆為 *linear transformations,* 則

$$(rT_1+sT_2)^{\mathrm{t}}=rT_1^{\mathrm{t}}+sT_2^{\mathrm{t}}.$$

(2) 若 $T_1:V\to W$, $T_2:W\to U$ 皆為 *linear transformations,* 則

$$(T_2\circ T_1)^{\mathrm{t}}=T_1^{\mathrm{t}}\circ T_2^{\mathrm{t}}.$$

(3) $(\mathrm{id}_V)^{\mathrm{t}}=\mathrm{id}_{V^*}$, 特別地, 若 $T:V\to W$ 為 *isomorphism,* 則 $T^{\mathrm{t}}:W^*\to V^*$ 亦為 *isomorphism* 且

$$(T^{\mathrm{t}})^{-1}=(T^{-1})^{\mathrm{t}}.$$

Proof.

(1) 對於任意 $f\in W^*$, 因為 f 為 linear, 我們有

$$(rT_1+sT_2)^{\mathrm{t}}(f)=f\circ(rT_1+sT_2)=r(f\circ T_1)+s(f\circ T_2)=rT_1^{\mathrm{t}}(f)+sT_2^{\mathrm{t}}(f)=(rT_1^{\mathrm{t}}+sT_2^{\mathrm{t}})(f).$$

得證 $(rT_1+sT_2)^{\mathrm{t}}=rT_1^{\mathrm{t}}+sT_2^{\mathrm{t}}$.

(2) 因為 $T_2\circ T_1$ 為從 V 到 U 的 linear transformation, 所以 $(T_2\circ T_1)^{\mathrm{t}}$ 為從 U^* 到 V^* 的 linear transformation. 現對任意 $f\in U^*$, 我們有

$$(T_2\circ T_1)^{\mathrm{t}}(f)=f\circ(T_2\circ T_1)=(f\circ T_2)\circ T_1=T_1^{\mathrm{t}}(f\circ T_2)=T_1^{\mathrm{t}}(T_2^{\mathrm{t}}(f))=T_1^{\mathrm{t}}\circ T_2^{\mathrm{t}}(f).$$

得證 $(T_2\circ T_1)^{\mathrm{t}}=T_1^{\mathrm{t}}\circ T_2^{\mathrm{t}}$.

(3) 對於任意 $f\in V^*$, 我們有 $(\mathrm{id}_V)^{\mathrm{t}}(f)=f\circ\mathrm{id}_V=f$, 所以 $(\mathrm{id}_V)^{\mathrm{t}}=\mathrm{id}_{V^*}$. 現若 $T:V\to W$ 為 isomorphism, 我們有 $T^{-1}\circ T=\mathrm{id}_V$ 且 $T\circ T^{-1}=\mathrm{id}_W$. 故由 (2) 得

$$\mathrm{id}_{V^*}=(\mathrm{id}_V)^{\mathrm{t}}=T^{\mathrm{t}}\circ(T^{-1})^{\mathrm{t}},\quad \mathrm{id}_{W^*}=(\mathrm{id}_W)^{\mathrm{t}}=(T^{-1})^{\mathrm{t}}\circ T^{\mathrm{t}},$$

得證 T^{t} 為 isomorphism 且 $(T^{\mathrm{t}})^{-1}=(T^{-1})^{\mathrm{t}}$. □

以下結果告訴我們一個 linear transformation 其 kernel 和 image 和其 transpose 的 kernel 與 image 之間的關係. 這個結果對任意的 vector space 皆成立, 不過我們僅探討 finite dimensional 的情形.

Proposition 5.3.2. 假設 V,W 皆為 *finite dimensional vector space,* $T:V\to W$ 為 *linear transformation,* $T^{\mathrm{t}}:W^*\to V^*$ 為其 *transpose.* 則

$$\mathrm{Ker}(T^{\mathrm{t}})=\mathrm{Im}(T)^0,\quad \mathrm{Im}(T^{\mathrm{t}})=\mathrm{Ker}(T)^0.$$

Proof. 假設 $f\in\mathrm{Ker}(T^{\mathrm{t}})$, 即 $T^{\mathrm{t}}(f)=f\circ T=\mathbf{O}$ in V^*. 這表示對任意 $\mathbf{v}\in V$, 皆有 $T^{\mathrm{t}}(f)(\mathbf{v})=f(T(\mathbf{v}))=0$. 得 $f\in\{T(\mathbf{v})\mid\mathbf{v}\in V\}^0=\mathrm{Im}(T)^0$. 反之, 若 $f\in\mathrm{Im}(T)^0$, 表示對任意 $\mathbf{v}\in V$, 皆有 $f(T(\mathbf{v}))=0$, 亦即 $T^{\mathrm{t}}(f)=\mathbf{O}$ in V^*, 得證 $f\in\mathrm{Ker}(T^{\mathrm{t}})$.

另一方面, 假設 $f\in\mathrm{Im}(T^{\mathrm{t}})$, 即存在 $g\in W^*$ 使得 $f=T^{\mathrm{t}}(g)=g\circ T$. 故對任意 $\mathbf{v}\in\mathrm{Ker}(T)$, 我們有 $f(\mathbf{v})=g(T(\mathbf{v}))=g(\mathbf{O}_W)=0$. 也就是說 $f\in\mathrm{Ker}(T)^0$, 得證 $\mathrm{Im}(T^{\mathrm{t}})\subseteq\mathrm{Ker}(T)^0$. 現因 $T^{\mathrm{t}}:W^*\to V^*$ 為 linear transformation, 利用 Theorem 5.2.2 我們有

$$\dim(\mathrm{Im}(T^{\mathrm{t}}))=\dim(W^*)-\dim(\mathrm{Ker}(T^{\mathrm{t}}))=\dim(W)-\dim(\mathrm{Ker}(T^{\mathrm{t}})).$$

再由 $\mathrm{Ker}(T^{\mathrm{t}})=\mathrm{Im}(T)^0$ 以及 Proposition 5.2.6 知

$$\dim(\mathrm{Ker}(T^{\mathrm{t}}))=\dim(\mathrm{Im}(T)^0)=\dim(W)-\dim(\mathrm{Im}(T)),$$

故得

$$\dim(\mathrm{Im}(T^{\mathrm{t}}))=\dim(\mathrm{Im}(T)).$$

而因 $T:V\to W$ 為 linear transformation, 故知 $\dim(\mathrm{Ker}(T))=\dim(V)-\dim(\mathrm{Im}(T))$, 再由 Proposition 5.2.6 知

$$\dim(\mathrm{Ker}(T)^0)=\dim(V)-\dim(\mathrm{Ker}(T))=\dim(\mathrm{Im}(T)).$$

故得 $\dim(\mathrm{Im}(T^{\mathrm{t}}))=\dim(\mathrm{Ker}(T)^0)$, 得證 $\mathrm{Im}(T^{\mathrm{t}})=\mathrm{Ker}(T)^0$. □

Question 5.15. 在 *Proposition 5.3.2* 的證明中我們知道 $\dim(\mathrm{Im}(T^{\mathrm{t}}))=\dim(\mathrm{Im}(T))$. 是否 $\dim(\mathrm{Ker}(T^{\mathrm{t}}))=\dim(\mathrm{Ker}(T))$?

既然 $T^{\mathrm{t}}:W^*\to V^*$ 也是 linear transformation, 我們自然也會去考慮 T^{t} 的 transpose, 也就是 $(T^{\mathrm{t}})^{\mathrm{t}}:(V^*)^*\to(W^*)^*$. 回顧一下當 V 為 finite dimensional, 我們有一個 isomorphism $\tau_V:V\to(V^*)^*$, 定義為對任意 $\mathbf{v}\in V$, $\tau_V(\mathbf{v})=\hat{\mathbf{v}}$ 其中 $\hat{\mathbf{v}}(f)=f(\mathbf{v}), \forall f\in V^*$. 我們有以下的圖示:

$$\begin{array}{ccc} V & \xrightarrow{\quad T\quad} & W \\ \downarrow{\scriptstyle \tau_V} & & \downarrow{\scriptstyle \tau_W} \\ (V^*)^* & \xrightarrow{\ (T^{\mathrm{t}})^{\mathrm{t}}\ } & (W^*)^* \end{array}$$

現對於任意 $\mathbf{v}\in V$, 我們有 $(T^{\mathrm{t}})^{\mathrm{t}}(\tau_V(\mathbf{v}))=\hat{\mathbf{v}}\circ T^{\mathrm{t}}\in(W^*)^*$. 也就是說對於任意 $f\in W^*$, 我們有

$$(T^{\mathrm{t}})^{\mathrm{t}}(\tau_V(\mathbf{v}))(f)=\hat{\mathbf{v}}\circ T^{\mathrm{t}}(f)=\hat{\mathbf{v}}(f\circ T)=f\circ T(\mathbf{v})=f(T(\mathbf{v})).$$

另一方面 $T(\mathbf{v})\in W$, 所以 $\tau_W(T(\mathbf{v}))=\widehat{T(\mathbf{v})}\in(W^*)^*$. 也就是說對於任意 $f\in W^*$, 我們有

$$\tau_W(T(\mathbf{v}))(f)=\widehat{T(\mathbf{v})}(f)=f(T(\mathbf{v})).$$

得證 $(T^{\mathrm{t}})^{\mathrm{t}}(\tau_V(\mathbf{v}))=\tau_W(T(\mathbf{v})), \forall\mathbf{v}\in V$, 亦即

$$(T^{\mathrm{t}})^{\mathrm{t}}\circ\tau_V=\tau_W\circ T.$$

我們證明了上面那個圖示為一個 *commutative diagram.* 再加上 τ_V 為 isomorphism, 故 $\tau_V^{-1}:(V^*)^*\to V$ 存在 (亦為 isomorphism). 我們有以下之結果.

Proposition 5.3.3. 假設 V,W 皆為 *finite dimensional vector space,* $T:V\to W$ 為 *linear transformation.* 令 τ_V 以及 τ_W 分別為 $V,(V^*)^*$ 以及 $W,(W^*)^*$ 之間的 *canonical map.* 則

$$(T^{\mathrm{t}})^{\mathrm{t}}=\tau_W\circ T\circ\tau_V^{-1}.$$

我們要了解一個 Linear transformation $T:V\to W$ 和 T 的 transpose $T^{\mathrm{t}}:W^*\to V^*$ 分別對應到 V,W 的 ordered basis 及其 dual basis 的 representative matrix 之間的關係.

Proposition 5.3.4. 令 $\beta=(\mathbf{v}_1,\ldots,\mathbf{v}_n),\gamma=(\mathbf{w}_1,\ldots,\mathbf{w}_m)$ 分別為 V,W 的 *ordered basis* 且令 $\beta^*=(\mathbf{v}_1^*,\ldots,\mathbf{v}_n^*),\gamma^*=(\mathbf{w}_1^*,\ldots,\mathbf{w}_m^*)$ 分別為相對應的 *dual basis* 所組成 V^*,W^* 的 *ordered basis.* 給定一個 *Linear transformation* $T:V\to W$, 我們有

$$_{\beta^*}[T^{\mathrm{t}}]_{\gamma^*}={}_{\gamma}[T]_{\beta}^{\mathrm{t}}.$$

Proof. 因為 $T^{\mathrm{t}}(\mathbf{w}_i^*)=\mathbf{w}_i^*\circ T\in V^*$, 若 $\mathbf{w}_i^*\circ T=c_{1,i}\mathbf{v}_1^*+\cdots+c_{n,i}\mathbf{v}_n^*$, 則依定義

$$c_{j,i}=\mathbf{w}_i^*\circ T(\mathbf{v}_j)=\mathbf{w}_i^*(T(\mathbf{v}_j)).$$

故若 $T(\mathbf{v}_j)=d_{1,j}\mathbf{w}_1+\cdots+d_{m,j}\mathbf{w}_m$, 則 $c_{j,i}=d_{i,j}$. 然而此處 $c_{j,i}$ 為 $_{\beta^*}[T^{\mathrm{t}}]_{\gamma^*}$ 的 (j,i)-th entry, 而 $d_{i,j}$ 為 $_{\gamma}[T]_{\beta}$ 的 (i,j)-th entry, 得證 $_{\beta^*}[T^{\mathrm{t}}]_{\gamma^*}={}_{\gamma}[T]_{\beta}^{\mathrm{t}}$. $\square$

由 Proposition 5.3.4, 我們可將一個 linear transformation 的 transpose 和一個 matrix 的 transpose 相連結. 我們可以將 Proposition 5.3.1, 5.3.2, 5.3.3 換成有關 matrix 的 transpose 的性質.

Question 5.16. 令 $\beta=(\mathbf{v}_1,\ldots,\mathbf{v}_n),\gamma=(\mathbf{w}_1,\ldots,\mathbf{w}_m)$ 分別為 V,W 的 *ordered basis* 且令 $\hat{\beta}=(\widehat{\mathbf{v}_1},\ldots,\widehat{\mathbf{v}_n}),\hat{\gamma}=(\widehat{\mathbf{w}_1},\ldots,\widehat{\mathbf{w}_m})$ 分別為 $(V^*)^*,(W^*)^*$ 的 *ordered basis,* 其中 $\widehat{\mathbf{v}_i}=\tau_V(\mathbf{v}_i),\widehat{\mathbf{w}_j}=\tau_W(\mathbf{w}_j)$ *(*$\tau_V:V\to(V^*)^*$, $\tau_W:W\to(W^*)^*$ 為 *canonical maps.)* 請利用 *Proposition 5.3.3, 5.3.4* 證明

$$({}_{\gamma}[T]_{\beta}^{\mathrm{t}})^{\mathrm{t}}={}_{\hat{\gamma}}[(T^{\mathrm{t}})^{\mathrm{t}}]_{\hat{\beta}}={}_{\gamma}[T]_{\beta}.$$

接下來我們要探討的是一個 linear transformation $T:V\to W$ 的 adjoint. 再次強調 adjoint 的定義是需要 V,W 皆為 finite dimensional inner product spaces. 回顧一下, 在此時我們有 $\rho_V:V\to V^*$, $\rho_W:W\to W^*$ 定義為 $\rho_V(\mathbf{v})=\langle\cdot,\mathbf{v}\rangle,\forall\mathbf{v}\in V$ 以及 $\rho_W(\mathbf{w})=\langle\cdot,\mathbf{w}\rangle,\forall\mathbf{w}\in W$ 且 ρ_V,ρ_W 皆為 conjugate isomorphism. 我們定義 T 的 *adjoint* $T^*:W\to V$ 為

$$T^*=\rho_V^{-1}\circ T^{\mathrm{t}}\circ\rho_W.$$

換言之對任意 $\mathbf{w}\in W$, 我們有

$$T^*(\mathbf{w})=\rho_V^{-1}\circ\rho_W(\mathbf{w})\circ T=\rho_V^{-1}(\langle T(\cdot),\mathbf{w}\rangle).$$

由 Lemma 5.2.9 (1) 我們知 $T^{\mathrm{t}}\circ\rho_W:W\to V^*$ 為 conjugate transformation, 故再由 Lemma 5.2.9 (2) 得 $T^*:W\to V$ 為 linear transformation. 我們有以下的 commutative diagram.

$$\begin{tikzcd} V \arrow[d, "\rho_V"] & W \arrow[l, "T^*"'] \arrow[d, "\rho_W"] \\ V^* & W^* \arrow[l, "T^{\mathrm{t}}"'] \end{tikzcd}$$

Theorem 5.3.5. 假設 V,W 皆為 *finite dimensional inner product space,* $T:V\to W$ 為 *linear transformation.* 則 T 的 *transpose* $T^*:W\to V$, 為唯一的 *linear transformation* 滿足

$$\langle T(\mathbf{v}),\mathbf{w}\rangle=\langle\mathbf{v},T^*(\mathbf{w})\rangle,\forall\mathbf{v}\in V,\mathbf{w}\in W.$$

Proof. 我們已知 T^* 為 linear transformation. 現對任意 $\mathbf{w}\in W$, 我們有 $\rho_V(T^*(\mathbf{w}))\in V^*$ 滿足對任意 $\mathbf{v}\in V$, 皆有 $\rho_V(T^*(\mathbf{w}))(\mathbf{v})=\langle\mathbf{v},T^*(\mathbf{w})\rangle$. 另一方面 $\rho_V(T^*(\mathbf{w}))=T^{\mathrm{t}}\circ\rho_W(\mathbf{w})$, 而對任意 $\mathbf{v}\in V$, 皆有 $T^{\mathrm{t}}\circ\rho_W(\mathbf{w})(\mathbf{v})=\langle T(\mathbf{v}),\mathbf{w}\rangle$. 得證 $\langle T(\mathbf{v}),\mathbf{w}\rangle=\langle\mathbf{v},T^*(\mathbf{w})\rangle$

至於唯一性, 假設 $T':W\to V$ 亦滿足 $\langle T(\mathbf{v}),\mathbf{w}\rangle=\langle\mathbf{v},T'(\mathbf{w})\rangle$, $\forall\mathbf{v}\in V,\mathbf{w}\in W$, 則知對任意 $\mathbf{v}\in V,\mathbf{w}\in W$ 皆有 $\langle\mathbf{v},T^*(\mathbf{w})\rangle=\langle\mathbf{v},T'(\mathbf{w})\rangle$. 由 Proposition 5.1.6 得證 $T'=T^*$. □

要注意 Theorem 5.3.5 中左式的 inner product 是 W 的 inner product, 右式為 V 的 inner product. 當一個 linear transformation $T:V\to W$ 保持 inner product, 亦即

$$\langle T(\mathbf{v}),T(\mathbf{v}')\rangle=\langle\mathbf{v},\mathbf{v}'\rangle,\quad\forall\mathbf{v},\mathbf{v}'\in V.$$

我們很自然有 T 為 one-to-one. 這是因為若 $\mathbf{v}\in\mathrm{Ker}(T)$, 則 $\langle\mathbf{v},\mathbf{v}\rangle=\langle T(\mathbf{v}),T(\mathbf{v})\rangle=0$, 得證 $\mathbf{v}=\mathbf{O}_V$. 故若 $\dim(W)=\dim(V)$, 則 T 為 isomorphism, 我們稱這樣的 isomorphism 為 *inner product isomorphism*. 此時 linear transformation $T^{-1}:W\to V$ 也會是 inner product isomorphism. 這是因為

$$\langle T^{-1}(\mathbf{w}),T^{-1}(\mathbf{w}')\rangle=\langle T(T^{-1}(\mathbf{w})),T(T^{-1}(\mathbf{w}'))\rangle=\langle\mathbf{w}.\mathbf{w}'\rangle,\quad\forall\mathbf{w},\mathbf{w}'\in W.$$

我們有以下判斷 inner product isomorphism 的方法.

Corollary 5.3.6. 假設 V,W 為 *finite dimensional inner product spaces,* $T:V\to W$ 為 *linear transformation* 且 $T^*W\to V$ 為其 *adjoint.* 則以下敘述為等價的 (*equivalent*).

(1) T 為 *inner product isomorphism.*

(2) T 為 *isomorphism* 且 $T^{-1}=T^*$.

Proof. (1) $\Rightarrow$ (2): 對於任意 $\mathbf{v}\in V,\mathbf{w}\in W$, 我們有

$$\langle T(\mathbf{v}),\mathbf{w}\rangle=\langle T(\mathbf{v}),T(T^{-1}(\mathbf{w}))\rangle=\langle\mathbf{v},T^{-1}(\mathbf{w})\rangle.$$

故利用 Theorem 5.3.5 有關 adjoint 的唯一性得證 $T^{-1}=T^*$.

(2) $\Rightarrow$ (1): 對於任意 $\mathbf{v},\mathbf{v}'\in V$, 皆有

$$\langle T(\mathbf{v}),T(\mathbf{v}')\rangle=\langle\mathbf{v},T^*(T(\mathbf{v}'))\rangle=\langle\mathbf{v},T^{-1}(T(\mathbf{v}'))\rangle=\langle\mathbf{v},\mathbf{v}'\rangle.$$

故 T 為 inner product isomorphism. □

接下來我們列出 adjoint 的基本性質.

Proposition 5.3.7. 設 V,W,U 皆為 *finite dimensional inner product space over* F.

(1) 若 $r,s\in F$ 且 $T_1:V\to W$, $T_2:V\to W$ 皆為 *linear transformations,* 則

$$(rT_1+sT_2)^*=\bar{r}T_1^*+\bar{s}T_2^*.$$

(2) 若 $T_1:V\to W$, $T_2:W\to U$ 皆為 *linear transformations,* 則

$$(T_2\circ T_1)^*=T_1^*\circ T_2^*.$$

(3) $\mathrm{id}_V^* = \mathrm{id}_V$, 特別地, 若 $T: V \to W$ 為 *isomorphism*, 則 $T^*: W \to V$ 亦為 *isomorphism* 且
$$(T^*)^{-1} = (T^{-1})^*.$$

(4) 若 $T: V \to W$ 為 *linear transformation*, 則
$$(T^*)^* = T.$$

Proof.

(1) 對於任意 $\mathbf{v} \in V, \mathbf{w} \in W$, 利用 inner product 的性質我們有

$\langle (rT_1 + sT_2)(\mathbf{v}), \mathbf{w}\rangle = r\langle T_1(\mathbf{v}), \mathbf{w}\rangle + s\langle T_2(\mathbf{v}), \mathbf{w}\rangle = r\langle \mathbf{v}, T_1^*(\mathbf{w})\rangle + s\langle \mathbf{v}, T_2^*(\mathbf{w})\rangle = \langle \mathbf{v}, (\bar{r}T_1^* + \bar{s}T_2^*)(\mathbf{w})\rangle$.

故由 adjoint 的唯一性得證 $(rT_1 + sT_2)^* = \bar{r}T_1^* + \bar{s}T_2^*$.

(2) 因為 $T_2 \circ T_1$ 為從 V 到 U 的 linear transformation, 所以 $(T_2 \circ T_1)^*$ 為從 U 到 V 的 linear transformation. 現對任意 $\mathbf{v} \in V, \mathbf{u} \in U$, 我們有
$$\langle (T_2 \circ T_1)(\mathbf{v}), \mathbf{u}\rangle = \langle T_2(T_1(\mathbf{v})), \mathbf{u}\rangle = \langle T_1(\mathbf{v}), T_2^*(\mathbf{u})\rangle = \langle \mathbf{v}, T_1^*(T_2^*(\mathbf{u}))\rangle = \langle \mathbf{v}, T_1^* \circ T_2^*(\mathbf{u})\rangle.$$
故由 adjoint 的唯一性得證 $(T_2 \circ T_1)^* = T_1^* \circ T_2^*$.

(3) 對於任意 $\mathbf{v}, \mathbf{v}' \in V$, 我們有 $\langle \mathbf{v}, \mathbf{v}'\rangle = \langle \mathrm{id}_V(\mathbf{v}), \mathbf{v}'\rangle = \langle \mathbf{v}, \mathrm{id}_V^*(\mathbf{v}')\rangle$, 所以利用 Corollary 5.1.5 得證 $\mathrm{id}_V^*(\mathbf{v}') = \mathbf{v}'$, $\forall \mathbf{v}' \in V$. 亦即 $\mathrm{id}_V^* = \mathrm{id}_V$. 現若 $T: V \to W$ 為 isomorphism, 我們有 $T^{-1} \circ T = \mathrm{id}_V$ 且 $T \circ T^{-1} = \mathrm{id}_W$. 故由 (2) 得
$$\mathrm{id}_V = \mathrm{id}_V^* = T^* \circ (T^{-1})^*, \quad \mathrm{id}_W = \mathrm{id}_W^* = (T^{-1})^* \circ T^*,$$
得證 T^* 為 isomorphism 且 $(T^*)^{-1} = (T^{-1})^*$.

(4) 因 $T^*: W \to V$ 為 linear transformation, 對任意 $\mathbf{v} \in V, \mathbf{w} \in W$ 我們有
$$\langle T^*(\mathbf{w}), \mathbf{v}\rangle = \langle \mathbf{w}, (T^*)^*(\mathbf{v})\rangle.$$
另一方面
$$\langle \mathbf{w}, T(\mathbf{v})\rangle = \overline{\langle T(\mathbf{v}), \mathbf{w}\rangle} = \overline{\langle \mathbf{v}, T^*(\mathbf{w})\rangle} = \langle T^*(\mathbf{w}), \mathbf{v}\rangle,$$
我們得知 $(T^*)^*(\mathbf{v}) = T(\mathbf{v})$, $\forall \mathbf{v} \in V$. 得證 $(T^*)^* = T$. □

Question 5.17. 試利用 *adjoint* 的定義證明 *Proposition 5.3.7.*

接下來, 我們來看一個 linear transformation 和其 adjoint 它們的 kernel 和 image 之間的關係.

Proposition 5.3.8. 假設 V, W 為 *finite dimensional inner product spaces*, $T: V \to W$ 為 *linear transformation* 且 $T^*: W \to V$ 為其 *adjoint.* 則

(1) $\mathrm{Ker}(T^*) = \mathrm{Im}(T)^\perp$ 且 $\mathrm{Im}(T^*) = \mathrm{Ker}(T)^\perp$.

(2) T *is one-to-one* 若且唯若 T^* *is onto.* T *is onto* 若且唯若 T^* *is one-to-one.*

(3) $\mathrm{Ker}(T^* \circ T) = \mathrm{Ker}(T)$ 且 $\mathrm{Ker}(T \circ T^*) = \mathrm{Ker}(T^*)$.

(4) $\mathrm{Im}(T^* \circ T) = \mathrm{Im}(T^*)$ 且 $\mathrm{Im}(T \circ T^*) = \mathrm{Im}(T)$.

Proof.

(1) 設 $\mathbf{w}\in \mathrm{Ker}(T^*)$, 即 $T^*(\mathbf{w})=\mathbf{O}_V$, 此時對任意 $\mathbf{v}\in V$, 我們有

$$\langle T(\mathbf{v}),\mathbf{w}\rangle = \langle \mathbf{v},T^*(\mathbf{w})\rangle = \langle \mathbf{v},\mathbf{O}_V\rangle = 0,$$

得證 $\mathbf{w}\in \mathrm{Im}(T)^{\perp}$. 反之, 若 $\mathbf{w}\in \mathrm{Im}(T)^{\perp}$, 則對任意 $\mathbf{v}\in V$ 皆有 $0=\langle T(\mathbf{v}),\mathbf{w}\rangle = \langle \mathbf{v},T^*(\mathbf{w})\rangle$. 故由 Lemma 5.1.4 得證 $\mathbf{w}\in \mathrm{Ker}(T^*)$. 另一方面, 若 $\mathbf{v}\in \mathrm{Im}(T^*)$, 表示存在 $\mathbf{w}\in W$ 使得 $\mathbf{v}=T^*(\mathbf{w})$. 故對任意 $\mathbf{v}'\in \mathrm{Ker}(T)$, 皆有 $\langle \mathbf{v}',\mathbf{v}\rangle = \langle \mathbf{v}',T^*(\mathbf{w})\rangle = \langle T(\mathbf{v}'),\mathbf{w}\rangle = \langle \mathbf{O}_W,\mathbf{w}\rangle = 0$, 即 $\mathbf{v}\in \mathrm{Ker}(T)^{\perp}$. 得證 $\mathrm{Im}(T^*)\subseteq \mathrm{Ker}(T)^{\perp}$. 最後利用 $\mathrm{Ker}(T^*)=\mathrm{Im}(T)^{\perp}$ 得

$$\begin{aligned}\dim(\mathrm{Im}(T^*)) &= \dim(W)-\dim(\mathrm{Ker}(T^*)) = \dim(W)-\dim(\mathrm{Im}(T)^{\perp})\\ &= \dim(W)-(\dim(W)-\dim(\mathrm{Im}(T))) = \dim(\mathrm{Im}(T))\end{aligned}$$

又因 $\dim(\mathrm{Ker}(T)^{\perp})=\dim(V)-\dim(\mathrm{Ker}(T))=\dim(\mathrm{Im}(T))$ 得 $\dim(\mathrm{Im}(T^*))=\dim(\mathrm{Ker}(T)^{\perp})$ 故得證 $\mathrm{Im}(T^*)=\mathrm{Ker}(T)^{\perp}$.

(2) 若 T is one-to-one, 則 $\mathrm{Ker}(T)=\{\mathbf{O}_V\}$. 故由 (1) 知 $\mathrm{Im}(T^*)=\mathrm{Ker}(T)^{\perp}=\{\mathbf{O}_V\}^{\perp}=V$, 得證 T^* 為 onto. 反之, 若 T^* 為 onto, 得 $\mathrm{Ker}(T)^{\perp}=V$, 由 Corollary 5.1.14 知

$$\mathrm{Ker}(T)=(\mathrm{Ker}(T)^{\perp})^{\perp}=V^{\perp}=\{\mathbf{O}_V\}.$$

得證 T 為 one-to-one. 利用此結果將 T 用 T^* 取代, 得知 T^* 為 one-to-one 若且唯若 $(T^*)^*=T$ 為 onto.

(3) 很明顯我們有 $\mathrm{Ker}(T)\subseteq \mathrm{Ker}(T^*\circ T)$. 現若 $\mathbf{v}\in \mathrm{Ker}(T^*\circ T)$, 則

$$\langle T(\mathbf{v}),T(\mathbf{v})\rangle = \langle \mathbf{v},T^*(T(\mathbf{v}))\rangle = \langle \mathbf{v},\mathbf{O}_V\rangle = 0.$$

得證 $T(\mathbf{v})=\mathbf{O}_W$, 即 $\mathbf{v}\in \mathrm{Ker}(T)$. 利用此結果將 T 用 T^* 取代, 再利用 $(T^*)^*$ 得 $\mathrm{Ker}(T\circ T^*)=\mathrm{Ker}(T^*)$.

(4) 很明顯我們有 $\mathrm{Im}(T^*\circ T)\subseteq \mathrm{Im}(T^*)$. 然而由 (3) 我們有

$$\dim(\mathrm{Im}(T^*\circ T))=\dim(V)-\dim(\mathrm{Ker}(T^*\circ T))=\dim(V)-\dim(\mathrm{Ker}(T))=\dim(\mathrm{Ker}(T)^{\perp}),$$

再由 (1) $\mathrm{Im}(T^*)=\mathrm{Ker}(T)^{\perp}$ 得知 $\dim(\mathrm{Im}(T^*\circ T))=\dim(\mathrm{Im}(T^*))$, 證得 $\mathrm{Im}(T^*\circ T)=\mathrm{Im}(T^*)$. 利用此結果將 T 用 T^* 取代, 得證 $\mathrm{Im}(T\circ T^*)=\mathrm{Im}(T)$. □

Question 5.18. 試證明 $\mathrm{Ker}(T)=\mathrm{Im}(T^*)^{\perp}$ 且 $\mathrm{Im}(T)=\mathrm{Ker}(T^*)^{\perp}$.

對於 $T:V\to W$ 和它的 adjoint $T^*:W\to V$ 的 representative matrix 當然有一定的關係, 由於 adjoint 和 inner product 有關, 所以 V,W 所選的 ordered basis 應也要和 inner product 有關. 事實上若分別取 V 和 W 的 orthonormal basis 所組成的 ordered basis $\beta=(\mathbf{v}_1,\dots,\mathbf{v}_n)$ 和 $\gamma=(\mathbf{w}_1,\dots,\mathbf{w}_m)$, 且 $T(\mathbf{v}_i)=c_{1,i}\mathbf{w}_1+\cdots+c_{m,i}\mathbf{w}_m$, 則依定義 ${}_\gamma[T]_\beta$ 的 (j,i)-th entry 為 $c_{j,i}$. 然而因 $\{\mathbf{w}_1,\dots,\mathbf{w}_m\}$ 為 W 的 orthonormal basis, 我們有 $c_{j,i}=\langle T(\mathbf{v}_i),\mathbf{w}_j\rangle = \langle \mathbf{v}_i,T^*(\mathbf{w}_j)\rangle$. 另一方面若 $T^*(\mathbf{w}_j)=d_{1,j}\mathbf{v}_1+\cdots+d_{n,j}\mathbf{v}_n$, 則 ${}_\beta[T^*]_\gamma$ 的 (i,j)-th entry 為

$$d_{i,j}=\langle T^*(\mathbf{w}_j),\mathbf{v}_i\rangle = \overline{\langle \mathbf{v}_i,T^*\mathbf{w}_j\rangle} = \overline{c_{j,i}}.$$

也就是說 ${}_{\beta}[T^*]_{\gamma}$ 是將 ${}_{\gamma}[T]_{\beta}$ 先取 transpose 再將每個 entry 取其 conjugate 而得. 我們有以下的定義.

Definition 5.3.9. 假設 $A \in M_{m\times n}(\mathbb{C})$ 且 $B \in M_{n\times m}(\mathbb{C})$. 對所有 $1 \leq i \leq n$, $1 \leq j \leq m$, 令 $a_{j,i}$ 為 A 的 (j,i)-th entry 且 $b_{i,j}$ 為 B 的 (i,j)-th entry. 若 $b_{i,j} = \overline{a_{j,i}}$, $\forall i \in \{1,\ldots,n\}, j \in \{1,\ldots,m\}$, 則稱 B 為 A 的 *adjoint* 且以 A^* 來表示.

要注意, 依此定義 $(A^*)^* = A$, 且若 $A \in M_{m\times n}(\mathbb{R})$, 則 $A^* = A^{\mathrm{t}}$. 又依此定義, 我們有以下之結果.

Proposition 5.3.10. 假設 V,W 為 *finite dimensional inner product space,* 且分別選取 V 和 W 的 *orthonormal basis* 所組成的 *ordered basis* $\beta = (\mathbf{v}_1,\ldots,\mathbf{v}_n)$ 和 $\gamma = (\mathbf{w}_1,\ldots,\mathbf{w}_m)$. 若 $T : V \to W$ 為 *linear transformation* 且 $T^* : W \to V$ 為 T 的 *adjoint,* 則

$${}_{\beta}[T^*]_{\gamma} = ({}_{\gamma}[T]_{\beta})^*.$$

由 Proposition 5.3.10, 我們可將一個 linear transformation 的 adjoint 和一個 matrix 的 adjoint 相連結. 我們可以將 Proposition 5.3.7, 5.3.8 換成有關 matrix 的 adjoint 的性質.

另外要注意的是在 Theorem 5.3.5 中我們強調一個 linear transformation 的 adjoint 是唯一的, 這是在給定一個 inner product 的條件之下. 在不同的 inner product 之下, 一個 linear transformation 會有不同的 adjoint, 我們看以下的例子.

Example 5.3.11. 考慮 $T : \mathbb{R}^3 \to \mathbb{R}^3$ 定義為

$$T(x_1,x_2,x_3) = (4x_1 + x_2 + 2x_3, 4x_1 + 4x_2 + 4x_3, 2x_1 + x_2 + 4x_3).$$

考慮 $\mathbb{R}^3$ 上兩個 inner products $\langle\, , \rangle$ 以及 $\langle\langle\, , \rangle\rangle$, 其中 $\langle\, , \rangle$ 為 standard inner product, 即

$$\langle (x_1,x_2,x_3),(y_1,y_2,y_3)\rangle = x_1y_1 + x_2y_2 + x_3y_3.$$

而 $\langle\langle\, , \rangle\rangle$ 的定義為

$$\langle\langle (x_1,x_2,x_3),(y_1,y_2,y_3)\rangle\rangle = x_1y_1 + \frac{1}{4}x_2y_2 + x_3y_3.$$

若考慮 $\mathbb{R}^3$ 為 standard inner product space, 則 $\varepsilon = \{(1,0,0),(0,1,0),(0,0,1)\}$ 為 $\mathbb{R}^3$ 的一組 ordered orthonormal basis. 此時 $[T]_{\varepsilon} = \begin{pmatrix} 4 & 1 & 2 \\ 4 & 4 & 4 \\ 2 & 1 & 4 \end{pmatrix}$, 故由 Proposition 5.3.10 得 $[T^*]_{\varepsilon} = \begin{pmatrix} 4 & 4 & 2 \\ 1 & 4 & 1 \\ 2 & 4 & 4 \end{pmatrix}$, 即此時 $T^*(x_1,x_2,x_3) = (4x_1 + 4x_2 + 2x_3, x_1 + 4x_2 + x_3, 2x_1 + 4x_2 + 4x_3)$.

若考慮 $\mathbb{R}^3$ 為以 $\langle\langle\, , \rangle\rangle$ 為 inner product 的 inner product space, 則可取

$$\beta = \{(1,0,0),(0,2,0),(0,0,1)\}$$

為 $\mathbb{R}^3$ 的一組 ordered orthonormal basis. 此時 $[T]_\beta = \begin{pmatrix} 4 & 2 & 2 \\ 2 & 4 & 2 \\ 2 & 2 & 4 \end{pmatrix}$, 由 Proposition 5.3.10 得 $[T^*]_\beta = \begin{pmatrix} 4 & 2 & 2 \\ 2 & 4 & 2 \\ 2 & 2 & 4 \end{pmatrix}$, 故知此時 $T^* = T$, 即

$$T^*(x_1,x_2,x_3) = (4x_1+x_2+2x_3, 4x_1+4x_2+4x_3, 2x_1+x_2+4x_3).$$

5.4. The Adjoint of Linear Operators

這一節中我們特別要探討一個 linear operator $T : V \to V$ 的 adjoint. 我們也會探討一些特別的 operator 的 adjoint. 所以本節中的 vector space V 永遠是一個 finite dimensional inner product space over F, 其中 $F = \mathbb{C}$ 或 $F = \mathbb{R}$.

當 $T : V \to V$ 為 linear operator, 則其 adjoint $T^* : V \to V$ 亦為 linear operator. 所以對任意 $f(x) \in F[x]$, $f(T^*)$ 亦有定義且為 linear operator. 我們對任意 $f(x) \in F[x]$, $f(T)$ 的 adjoint 為何有興趣. 首先若 $f(x) = c_nx^n + \cdots + c_1x + c_0$, 我們定義 $\overline{f}(x) = \overline{c_n}x^n + \cdots + \overline{c_1}x + \overline{c_0}$, 則有以下結果.

Lemma 5.4.1. 若 $T : V \to V$ 為 *linear operator,* 則對任意 $f(x) \in F[x]$ 皆有

$$(f(T))^* = \overline{f}(T^*).$$

Proof. 首先由 Proposition 5.3.7(2) 知 $(T^{\circ n})^* = (T^*)^{\circ n}$. 故若 $f(x) = c_nx^n + \cdots + c_1x + c_0$, 則再由 Proposition 5.3.7(1)(3) 得證

$$\begin{aligned} (f(T))^* &= (c_nT^{\circ n} + \cdots + c_1T + c_0\mathrm{id}_V)^* = \overline{c_n}\,(T^{\circ n})^* + \cdots + \overline{c_1}\,T^* + \overline{c_0}\,\mathrm{id}_V^* \\ &= \overline{c_n}\,(T^*)^{\circ n} + \cdots + \overline{c_1}\,T^* + \overline{c_0}\,\mathrm{id}_V = \overline{f}(T^*). \end{aligned}$$

□

由 Proposition 5.3.10 以及 Lemma 5.4.1, 我們馬上可以得到 T 和 T^* 的 characteristic polynomials $\chi_T(x)$ 和 $\chi_{T^*}(x)$ 之間的關係, 以及它們的 minimal polynomials $\mu_T(x)$ 和 $\mu_{T^*}(x)$ 之間的關係.

Lemma 5.4.2. 若 $T : V \to V$ 為 *linear operator,* 則

$$\chi_{T^*}(x) = \overline{\chi_T}(x) \quad \textit{and} \quad \mu_{T^*}(x) = \overline{\mu_T}(x).$$

Proof. 由於 characteristic polynomial 與 ordered basis 的選取無關, 所以我們特別選取 V 的一組 orthonormal basis 所組成的 ordered basis β. 由 Proposition 5.3.10 知 $[T^*]_\beta = ([T]_\beta)^*$, 故得 $\chi_{T^*}(x) = \overline{\chi_T}(x)$.

由 minimal polynomial 定義知 $\mu_T(T) = \mathbf{O}$, 故由 Lemma 5.4.1, 得 $\overline{\mu_T}(T^*) = (\mu_T(T))^* = \mathbf{O}$, 故知 $\mu_{T^*}(x) \mid \overline{\mu_T}(x)$. 同理由 $(T^*)^* = T$, 得 $\mu_T(x) \mid \overline{\mu_{T^*}}(x)$, 再將兩邊多項式的係數取 conjugate 得 $\overline{\mu_T}(x) \mid \mu_{T^*}(x)$. 得證 $\mu_{T^*}(x) = \overline{\mu_T}(x)$. □

有關 linear operator 的 decomposition, 最重要的便是 T-invariant subspace. 現若 W 為 T-invariant subspace, 我們自然會問 W 是否為 T^*-invariant subspace. 一般來說這不一定對, 但我們有以下之結果.

Lemma 5.4.3. 若 $T: V \to V$ 為 *linear operator,* 則 $W \subseteq V$ 為 T*-invariant subspace* 若且唯若 $W^\perp \subseteq V$ 為 T^**-invariant subspace.*

Proof. 假設 W 為 T-invariant, 要說明 $W^\perp$ 為 T^*-invariant, 就是要說明對任意 $\mathbf{w}' \in W^\perp$ 皆有 $T^*(\mathbf{w}') \in W^\perp$. 然而若 $\mathbf{w} \in W$, 則由 $T(\mathbf{w}) \in W$ 以及 $\mathbf{w}' \in W^\perp$, 得 $\langle \mathbf{w}, T^*(\mathbf{w}') \rangle = \langle T(\mathbf{w}), \mathbf{w}' \rangle = 0$. 此即證明 $T^*(\mathbf{w}') \in W^\perp$, 故 $W^\perp$ 為 T^*-invariant.

反之, 若 $W^\perp$ 為 T^*-invariant, 則由上面所證 $(W^\perp)^\perp = W$ 為 $(T^*)^*$-invariant, 即 T-invariant. $\square$

給定一個 inner product space V, 及其 subspace W, 最直接的 decomposition 為 $V = W \boxplus W^\perp$. 所以當 W 為 T-invariant 時, 我們自然會問是否 $W^\perp$ 亦為 T-invariant. 利用 Lemma 5.4.3, 這剛好回答了何時 W 亦為 T^*-invariant.

Corollary 5.4.4. 假設 $T: V \to V$ 為 *linear operator* 且 $W \subseteq V$ 為 T*-invariant subspace.* 則 $W^\perp$ 為 T*-invariant subspace* 若且唯若 W 為 T^**-invariant subspace.* 另外若 W 為 T*-invariant* 和 T^**-invariant,* 則

$$(T|_W)^* = T^*|_W.$$

Proof. 由 Lemma 5.4.3, 我們知 $W^\perp$ 為 T-invariant 等價於 $W = (W^\perp)^\perp$ 為 T^*-invariant. 此時對任意的 $\mathbf{w}, \mathbf{w}' \in W$ 皆有

$$\langle T|_W(\mathbf{w}), \mathbf{w}' \rangle = \langle T(\mathbf{w}), \mathbf{w}' \rangle = \langle \mathbf{w}, T^*(\mathbf{w}') \rangle = \langle \mathbf{w}, T^*|_W(\mathbf{w}') \rangle,$$

得證 $(T|_W)^* = T^*|_W$. $\square$

接下來我們探討幾個特殊的 linear operator 及其 adjoint 間的關係. 首先回顧若 $V = W_1 \oplus W_2$, 則對任意 $\mathbf{v} \in V$, 皆存在唯一的 $\mathbf{w}_1 \in W_1$, $\mathbf{w}_2 \in W_2$ 使得 $\mathbf{v} = \mathbf{w}_1 + \mathbf{w}_2$. 此時我們定義 $\pi_{W_1,W_2}: V \to V$, 為 $\pi_{W_1,W_2}(\mathbf{v}) = \mathbf{w}_1$. 我們稱 π_{W_1,W_2} 為 *projection* on W_1 along W_2. 要注意, 若 $V = W_1 \oplus W_2'$, 其中 $W_2 \neq W_2'$, 則 $\pi_{W_1,W_2} \neq \pi_{W_1,W_2'}$.

Question 5.19. 如上所述, $\pi_{W_2,W_1}(\mathbf{v})$ 為何? 並說明若 $V = W_1 \oplus W_2 = W_1 \oplus W_2'$ 但 $W_2 \neq W_2'$ 則 $\pi_{W_1,W_2} \neq \pi_{W_1,W_2'}$.

當 V 是 finite dimensional inner product space, 我們可由 $V = W_1 \oplus W_2$ 推得 $V = W_1^\perp \oplus W_2^\perp$. 這是因為若 $\mathbf{w} \in W_1^\perp \cap W_2^\perp$, 則對任意 $\mathbf{v} \in V$, 我們有 $\mathbf{v} = \mathbf{w}_1 + \mathbf{w}_2$, 其中 $\mathbf{w}_1 \in W_1, \mathbf{w}_2 \in W_2$, 故得 $\langle \mathbf{v}, \mathbf{w} \rangle = \langle \mathbf{w}_1, \mathbf{w} \rangle + \langle \mathbf{w}_2, \mathbf{w} \rangle = 0$. 由 Lemma 5.1.4 得知 $\mathbf{w} = \mathbf{O}_V$. 再由

$$\dim(W_1^\perp + W_2^\perp) = \dim(W_1^\perp) + \dim(W_2^\perp) = \dim(W_2) + \dim(W_1) = \dim(V),$$

得證 $V = W_1^\perp \oplus W_2^\perp$. 利用此我們可以得到 π_{W_1,W_2} 的 adjoint $\pi^*_{W_1,W_2}$.

Proposition 5.4.5. 若 $V = W_1 \oplus W_2$, 則 $V = W_1^{\perp} \oplus W_2^{\perp}$ 且

$$\pi_{W_1,W_2}^* = \pi_{W_2^{\perp},W_1^{\perp}}.$$

Proof. 前面已知 $V = W_1^{\perp} \oplus W_2^{\perp}$, 故對任意 $\mathbf{v},\mathbf{v}' \in V$, 我們知存在 $\mathbf{w}_1 \in W_1, \mathbf{w}_2 \in W_2$ 以及 $\mathbf{w}_1' \in W_1^{\perp}, \mathbf{w}_2' \in W_2^{\perp}$ 滿足 $\mathbf{v} = \mathbf{w}_1 + \mathbf{w}_2, \mathbf{v}' = \mathbf{w}_1' + \mathbf{w}_2'$. 此時

$$\langle \pi_{W_1,W_2}(\mathbf{v}), \mathbf{v}' \rangle = \langle \mathbf{w}_1, \mathbf{w}_1' + \mathbf{w}_2' \rangle = \langle \mathbf{w}_1, \mathbf{w}_2' \rangle.$$

另一方面

$$\langle \mathbf{v}, \pi_{W_2^{\perp},W_1^{\perp}}(\mathbf{v}') \rangle = \langle \mathbf{w}_1 + \mathbf{w}_2, \mathbf{w}_2' \rangle = \langle \mathbf{w}_1, \mathbf{w}_2' \rangle.$$

得證 $\langle \pi_{W_1,W_2}(\mathbf{v}), \mathbf{v}' \rangle = \langle \mathbf{v}, \pi_{W_2^{\perp},W_1^{\perp}}(\mathbf{v}') \rangle, \forall \mathbf{v},\mathbf{v}' \in V$, 即 $\pi_{W_1,W_2}^* = \pi_{W_2^{\perp},W_1^{\perp}}$. □

我們知道在 inner product space 中最常用的 decomposition 就是 $V = W \boxplus W^{\perp}$. 此時我們稱 projection $\pi_{W,W^{\perp}}$ 為 *orthogonal projection* on W. 為了方便, 以後我們都會把 orthogonal projection on W 用 π_W 來表示 (因為它僅和 W 有關). 利用 Proposition 5.4.5, 我們馬上有以下之結果.

Corollary 5.4.6. 設 V 是 *finite dimensional inner product space* 且 W 為其 *subspace.* 令 $\pi_W : V \to V$ 為 *orthogonal projection on* W, 則 $\pi_W = \pi_W^{\circ 2}$ 且 $\pi_W^* = \pi_W$.

Proof. 對任意 $\mathbf{v} \in V$, 我們有 $\mathbf{v} = \mathbf{w} + \mathbf{w}'$, 其中 $\mathbf{w} \in W, \mathbf{w}' \in W^{\perp}$, 故

$$\pi_W^{\circ 2}(\mathbf{v}) = \pi_W(\pi_W(\mathbf{v})) = \pi_W(\pi_W(\mathbf{w} + \mathbf{w}')) = \pi_W(\mathbf{w}) = \pi_W(\mathbf{v}), \forall \mathbf{v} \in V.$$

得證 $\pi_W = \pi_W^{\circ 2}$. 另一方面, 利用 Proposition 5.4.5 以及 $(W^{\perp})^{\perp} = W$, 我們有

$$\pi_W^* = \pi_{W,W^{\perp}}^* = \pi_{(W^{\perp})^{\perp},W^{\perp}} = \pi_{W,W^{\perp}} = \pi_W.$$

□

Question 5.20. 試證明對任意 $\mathbf{v} \in V$, π_W 就是 *Proposition 5.1.12* 中提的 *projection* proj_W.

一個 linear operator $T : V \to V$ 若滿足 $T^{\circ 2} = T$, 則稱為 *idempotent.* 由 Corollary 5.4.6 的證明我們知道任何的 projection 皆為 idempotent (不需 orthogonal projection 之假設). 至於 $T^* = T$ 的性質, 我們稱為 *self-adjoint.* 一般的 projection 不會是 self-adjoint, 除非它是 orthogonal projection, 這是由於我們有下結果.

Proposition 5.4.7. 假設 $T : V \to V$ 為 *linear transformation* 滿足 $T^{\circ 2} = T$, 則下列敘述為等價:

(1) T 為 *orthogonal projection.*

(2) $T^* = T$.

(3) $\mathrm{Ker}(T) = \mathrm{Im}(T)^{\perp}$.

(4) $\mathrm{Im}(T) = \mathrm{Ker}(T)^{\perp}$.

Proof. 首先我們說明若 T 為 idempotent (即 $T^{\circ 2}=T$), 則 $T=\pi_{\mathrm{Im}(T),\mathrm{Ker}(T)}$. 這需先證明 $V=\mathrm{Im}(T)\oplus\mathrm{Ker}(T)$. 事實上 $\dim(V)=\dim(\mathrm{Im}(T))+\dim(\mathrm{Ker}(T))$, 所以只要證明

$$\mathrm{Im}(T)\cap\mathrm{Ker}(T)=\{\mathbf{O}_V\},$$

則由 $\mathrm{Im}(T)+\mathrm{Ker}(T)\subseteq V$, 可得 $V=\mathrm{Im}(T)\oplus\mathrm{Ker}(T)$. 然而若 $\mathbf{v}\in\mathrm{Im}(T)$, 表示存在 $\mathbf{w}\in V$ 使得 $\mathbf{v}=T(\mathbf{w})$, 故由 $T^{\circ 2}=T$, 得 $T(\mathbf{v})=T^{\circ 2}(\mathbf{w})=T(\mathbf{w})=\mathbf{v}$. 若再加上 $\mathbf{v}\in\mathrm{Ker}(T)$ 可得 $\mathbf{v}=T(\mathbf{v})=\mathbf{O}_V$. 現已知 $V=\mathrm{Im}(T)\oplus\mathrm{Ker}(T)$, 故對任意 $\mathbf{v}$, 皆存在 $\mathbf{w}\in\mathrm{Im}(T),\mathbf{w}'\in\mathrm{Ker}(T)$ 使得 $\mathbf{v}=\mathbf{w}+\mathbf{w}'$. 可得 $T(\mathbf{v})=T(\mathbf{w})+T(\mathbf{w}')=T(\mathbf{w})$. 但前面已知若 $\mathbf{w}\in\mathrm{Im}(T)$, 則 $T(\mathbf{w})=\mathbf{w}$, 故知 $T(\mathbf{v})=\mathbf{w}$. 此和 projection on $\mathrm{Im}(T)$ along $\mathrm{Ker}(T)$ 即 $\pi_{\mathrm{Im}(T),\mathrm{Ker}(T)}$ 對 $\mathbf{v}$ 的作用相同, 故得證 $T=\pi_{\mathrm{Im}(T),\mathrm{Ker}(T)}$.

現由 Corollary 5.1.14, 我們知道 $\mathrm{Ker}(T)=\mathrm{Im}(T)^{\perp}$ 若且唯若 $\mathrm{Im}(T)=\mathrm{Ker}(T)^{\perp}$ (即 (3) $\Leftrightarrow$ (4)). 而又若 $\mathrm{Ker}(T)=\mathrm{Im}(T)^{\perp}$ 則 $\pi_{\mathrm{Im}(T),\mathrm{Ker}(T)}=\pi_{\mathrm{Im}(T),\mathrm{Im}(T)^{\perp}}$, 得證 $T=\pi_{\mathrm{Im}(T),\mathrm{Ker}(T)}$ 為 orthogonal projection (即 (3) $\Rightarrow$ (1)). 又若 T 為 orthogonal projection, 則由 Corollary 5.4.6 知 $T^*=T$ (即 (1) $\Rightarrow$ (2)). 最後利用 Proposition 5.3.8(1) 知若 $T^*=T$, 則 $\mathrm{Im}(T)=\mathrm{Ker}(T)^{\perp}$ (且 $\mathrm{Ker}(T)=\mathrm{Im}(T)^{\perp}$) (即 (2) $\Rightarrow$ (3)(4)) 得證本定理. □

一般來說, 我們可以將一個 finite dimensional inner product space 寫成 orthogonal direct sum $V=W_1\boxplus\cdots\boxplus W_k$, 此時 $W_i^{\perp}=\oplus_{j\neq i}W_j$. 若令 π_i 為 orthogonal projection on W_i (即 $\pi_i=\pi_{W_i}$), 則很容易看出

$$\pi_1+\cdots+\pi_k=\mathrm{id}_V.$$

我們稱此為 *orthogonal resolution of the identity.* 注意此時我們有 $\pi_i^{\circ 2}=\pi_i$, $\pi_i^*=\pi_i$ 以及當 $i\neq j$ 時, $\pi_i\circ\pi_j=\mathbf{O}$. 事實上, 我們有以下之結果.

Lemma 5.4.8. 假設 V 為 *finite dimensional inner product space* 且 W_1,W_2 為 V 的 *subspace.* 令 π_1,π_2 分別為 *orthogonal projection on* W_1,W_2. 則下列為等價:

(1) $W_1\perp W_2$.

(2) $\pi_1\circ\pi_2=\mathbf{O}$.

(3) $\pi_2\circ\pi_1=\mathbf{O}$.

Proof. 對於任意 $\mathbf{v}\in V$, 由於 $\pi_1(\mathbf{v})\in W_1$, 故若 $W_1\perp W_2$, 則 $\pi_1(\mathbf{v})\in W_2^{\perp}$. 得證 $\pi_2(\pi_1(\mathbf{v}))=\mathbf{O}_V$, $\forall\mathbf{v}\in V$ 即 $\pi_1\circ\pi_2=\mathbf{O}$. 同理可得 $\pi_2\circ\pi_1=\mathbf{O}$.

現對任意 $\mathbf{w}_1\in W_1,\mathbf{w}_2\in W_2$, 我們有 $\langle\mathbf{w}_1,\mathbf{w}_2\rangle=\langle\pi_1(\mathbf{w}_1),\pi_2(\mathbf{w}_2)\rangle=\langle\mathbf{w}_1,\pi_1^*(\pi_2(\mathbf{w}_2))\rangle$. 因 π_1 為 orthogonal projection, 由 Corollary 5.4.6 知 $\pi_1^*=\pi_1$. 故若 $\pi_1\circ\pi_2=\mathbf{O}$, 則 $\langle\mathbf{w}_1,\mathbf{w}_2\rangle=\langle\mathbf{w}_1,\mathbf{O}_V\rangle=0$, 得證 $W_1\perp W_2$. 同理若 $\pi_2\circ\pi_1=\mathbf{O}$ 亦可得 $W_1\perp W_2$. □

回顧過去我們提過, 若對於 linear operator $T:V\to V$, 我們可以找到 V 的一組 basis 是由 T 的 eigenvectors 所組成, 則表示 T 是 diagonalizable. 特別的, 若存在一組 T 的 eigenvectors 形成 V 的一組 orthogonal basis (或 orthonormal basis), 則我們有以下的定義.

Definition 5.4.9. 假設 V 是一個 finite dimensional inner product space over F 且 $T: V \to V$ 為 linear operator. 若存在一組 T 的 eigenvectors 形成 V 的一組 orthogonal basis (或 orthonormal basis), 則我們稱 T 為 *unitary diagonalizable* (當 $F = \mathbb{R}$ 時, 有的書會特別稱 T 為 *orthogonal diagonalizable*).

當 A 是一個 $n \times n$ matrix over F, 若考慮 F^n 上的 standard inner product, 存在一組 A 的 eigenvectors 為 F^n 的一組 orthogonal basis (或 orthonormal basis), 則我們稱 A 為 *unitary diagonalizable* (當 $F = \mathbb{R}$ 時, 有的書會特別稱 A 為 *orthogonal diagonalizable*).

在此特別說明, 當一個 matrix A 為 unitary diagonalizable 時, 在 $F = \mathbb{C}$ 的情況即表示存在一個 matrix P 滿足 $P^* = P^{-1}$ 使得 P^*AP 為一個 diagonal matrix. 這樣的矩陣 P 是由那些 eigenvectors 所成的 orthonormal basis 所組成的, 我們稱為 *unitary matrix.* 而在 $F = \mathbb{R}$ 的情況, 表示存在一個 matrix P 滿足 $P^{\mathrm{t}} = P^{-1}$ 使得 $P^{\mathrm{t}}AP$ 為一個 diagonal matrix. 這樣的矩陣 P 是由那些 eigenvectors 所成的 orthonormal basis 所組成的, 我們稱為 *orthogonal matrix.* 這就是 unitary diagonalizable (或 orthogonal diagonalizable) 名稱的由來. 為了方便起見, 我們不去區分 unitary diagonalizable 或 orthogonal diagonalizable, 一律用 unitary diagonalizable 來稱之. 接下來我們來探討 unitary diagonalizable linear operator 的特性.

Proposition 5.4.10. 假設 V 是一個 *finite dimensional inner product space over* F 且 $T: V \to V$ 為 *linear operator.* 則下列是等價的.

(1) T 為 *unitary diagonalizable.*

(2) 存在相異 $\lambda_1, \ldots, \lambda_k \in F$ 使得 $V = E_{\lambda_1} \boxplus \cdots \boxplus E_{\lambda_k}$, 其中 E_{λ_i}, 為 λ_i 的 *eigenspace,* 即 $E_{\lambda_i} = \{\mathbf{v} \in V \mid T(\mathbf{v}) = \lambda_i \mathbf{v}\}$.

(3) 存在 $\lambda_1, \ldots, \lambda_k \in F$ 使得 $T = \lambda_1 \pi_1 + \cdots + \lambda_k \pi_k$, 其中 π_i 為 *orthogonal projection* 且滿足 $\pi_i \circ \pi_j = \mathbf{O}$, $\forall i \neq j$.

Proof. 假設 T 為 unitary diagonalizable, 假設 $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ 為 T 的一組 eigenvectors 所組成的 orthogonal basis. 將其適當重排使得相同 eigenvalue 的 eigenvectors 擺在一起, 我們假設 $\mathbf{v}_1, \ldots, \mathbf{v}_{l_1}$ 為 eigenvalue 為 λ_1 的 eigenvectors, $\mathbf{v}_{l_1+1}, \ldots, \mathbf{v}_{l_2}$ 為 eigenvalue 為 λ_2 的 eigenvectors, 依此類推 (其中每個 λ_i 皆相異). 我們要說明 $W_i = \mathrm{Span}(\{\mathbf{v}_{l_{i-1}+1}, \ldots, \mathbf{v}_{l_i}\})$ 等於 E_{λ_i}. 事實上 $W_i \subseteq E_{\lambda_i}$, 又由 T 為 diagonalizable 知 $V = E_{\lambda_1} \oplus \cdots \oplus E_{\lambda_k}$. 我們有

$$\dim(V) = \dim(W_1) + \cdots + \dim(W_k) \leq \dim(E_{\lambda_1}) + \cdots + \dim(E_{\lambda_k}) = \dim(V),$$

得證 $\dim(W_i) = \dim(E_{\lambda_i})$, 即 $W_i = E_{\lambda_i}$. 很自然的, 對於 $i \neq j$ 我們有 $W_i \perp W_j$, 故得證 $V = E_{\lambda_1} \boxplus \cdots \boxplus E_{\lambda_k}$. 反之, 若 $V = E_{\lambda_1} \boxplus \cdots \boxplus E_{\lambda_k}$, 則考慮每個 E_{λ_i} 上的一組 orthogonal basis, 由於這些 E_{λ_i} 上的 orthogonal basis 可組成 V 的一組 orthogonal basis, 且皆為 T 的 eigenvectors. 得證 (1) $\Leftrightarrow$ (2).

現若 $V = E_{\lambda_1} \boxplus \cdots \boxplus E_{\lambda_k}$, 則對任意 $\mathbf{v} \in V$, 我們有 $\mathbf{v} = \mathbf{v}_1 + \cdots + \mathbf{v}_k$, 其中 $\mathbf{v}_i \in E_{\lambda_i}$. 此時令 π_i 為 orthogonal projection on E_{λ_i}, 則由 Lemma 5.4.8 知 $\pi_i \circ \pi_j = \mathbf{O}$, $\forall i \neq j$ 且

$$(\lambda_1 \pi_1 + \cdots + \lambda_k \pi_k)(\mathbf{v}) = \lambda_1 \mathbf{v}_1 + \cdots \lambda_k \mathbf{v}_k = T(\mathbf{v}_1) + \cdots + T(\mathbf{v}_k) = T(\mathbf{v}_1 + \cdots + \mathbf{v}_k) = T(\mathbf{v}).$$

得證 $T=\lambda_1\pi_1+\cdots+\lambda_k\pi_k$, 故 (2) $\Rightarrow$ (3).

反之, 若 $T=\lambda_1\pi_1+\cdots+\lambda_k\pi_k$, 令 $\mathrm{Im}(\pi_i)=W_i$, 則由 $\pi_i\circ\pi_j=\mathbf{O}$, $\forall i\neq j$ 知 $W_i\perp W_j$, $\forall i\neq j$. 考慮 β_i 為 W_i 的一組 orthogonal basis, 且令 $W=W_1\boxplus\cdots\boxplus W_k$. 若 $W=V$, 則 $\beta=\cup_{i=1}^k\beta_i$ 為 V 的一組 orthogonal basis, 且對任意 $\mathbf{v}\in\beta_i$, 由於 $\mathbf{v}\in W_i$ 當 $j\neq i$, 我們有 $\pi_j(\mathbf{v})=\pi_j(\pi_i(\mathbf{v}))=\mathbf{O}_V$. 故 $T(\mathbf{v})=\lambda_i\mathbf{v}$, 亦即 $\mathbf{v}$ 為 T 的 eigenvector. 而若 $V\neq W$, 令 $W_{k+1}=W^{\perp}$, 此時 $V=W_1\boxplus\cdots\boxplus W_k\boxplus W_{k+1}$. 若再取 β_{k+1} 為 W_{k+1} 的一組 orthogonal basis, 則 $\beta=\cup_{i=1}^{k+1}\beta_i$ 為 V 的一組 orthogonal basis. 又對於 $\mathbf{v}\in\beta_{k+1}$ 我們有 $\mathbf{v}\in W_i^{\perp}$, $\forall i\in\{1,\ldots,k\}$, 故 $\pi_i(\mathbf{v})=0$ 得 $T(\mathbf{v})=\mathbf{O}_V=0\mathbf{v}$, 即 $\mathbf{v}$ 為 T 的 eigenvector. 得證 β 為 V 的一組 orthogonal basis 且 β 中的元素皆為 T 的 eigenvector, 即 T 為 unitary diagonalizable. 得證 (3) $\Rightarrow$ (1). $\square$

由 Proposition 5.4.10 我們知道若 $T:V\to V$ 為 unitary diagonalizable, 則存在 $\lambda_1,\ldots,\lambda_k\in F$ 以及 orthogonal projections $\pi_1,\ldots,\pi_k$ 使得 $T=\lambda_1\pi_1+\cdots+\lambda_k\pi_k$, 這稱為 T 的 *spectral resolution.* 此時由 Proposition 5.3.7 得

$$T^*=\overline{\lambda_1}\pi_1^*+\cdots+\overline{\lambda_k}\pi_k^*=\overline{\lambda_1}\pi_1+\cdots+\overline{\lambda_k}\pi_k.$$

再由 $\pi_i\circ\pi_j=\mathbf{O}$, $\forall i\neq j$ 且 $\pi_i^{\circ 2}=\pi_i$ 得

$$T^*\circ T=\overline{\lambda_1}\lambda_1\pi_1+\cdots+\overline{\lambda_k}\lambda_k\pi_k=T\circ T^*.$$

符合 $T^*\circ T=T\circ T^*$ 的 linear operator T 稱為 *normal operator*, 這是我們下一節所要探討的課題.

Question 5.21. 試證明若 $n\times n$ *matrix* A 為 *unitary diagonalizable,* 則 $A^*A=AA^*$.

5.5. Normal Operators

我們將探討 normal operator 的性質. 由於我們探討的 normal operator 為定義在 over F ($F=\mathbb{C}$ 或 $F=\mathbb{R}$) 的 finite dimensional inner product space. 所以本節中的 vector space 一定是 finite dimensional inner product space 且其 over 的 field 為 $\mathbb{C}$ 或 $\mathbb{R}$, 我們就不再贅敘.

Definition 5.5.1. 設 $T:V\to V$ 為 linear operator, 若 T 滿足 $T^*\circ T=T\circ T^*$, 則稱 T 為一個 *normal operator.*

設 A 為一個 $n\times n$ matrix, 若 A 滿足 $A^*A=AA^*$, 則稱 A 為一個 *normal matrix.*

Question 5.22. 一個 *orthogonal projection* 是否為 *normal?*

首先我們觀察可以由一個 normal operator 得到更多的 normal operator.

Lemma 5.5.2. 假設 $T:V\to V$ 為 *normal operator.*

(1) T^* 亦為 *normal operator.*

(2) 若 $W,W^{\perp}$ 皆為 T*-invariant,* 則 $T|_W$ 亦為 *normal operator.*

(3) 若 T 為 *isomorphism,* 則 T^{-1} 亦為 *normal operator.*

(4) 對任意 $f(x) \in F[x]$, $f(T)$ 亦為 *normal operator.*

Proof. (1) 由 $(T^*)^* = T$, 可得 T^* 亦為 normal.

(2) 由 Corollary 5.4.4, 我們知此時 W 亦為 T^*-invariant 且 $(T|_W)^* = T^*|_W$. 故

$$(T|_W)^* \circ T|_W = T^*|_W \circ T|_W = (T^* \circ T)|_W = (T \circ T^*)|_W = T|_W \circ (T|_W)^*.$$

(3) 由 Proposition 5.3.7, 我們知此時 $(T^{-1})^* = (T^*)^{-1}$, 故

$$(T^{-1})^* \circ T^{-1} = (T^*)^{-1} \circ T^{-1} = (T \circ T^*)^{-1} = (T^* \circ T)^{-1} = T^{-1} \circ (T^{-1})^*.$$

(4) 若 $f(x) = c_n x^n + \cdots + c_1 x + c_0$, 則 $f(T) = c_n T^{\circ n} + \cdots + c_1 T + c_0 \mathrm{id}_V$ 且由 Lemma 5.4.1 知 $f(T)^* = \overline{c_n}(T^*)^{\circ n} + \cdots + \overline{c_1} T^* + c_0 \mathrm{id}_V$. 現因 $c_i T^{\circ i} \circ f(T)^* = c_i \sum_{j=0}^{n} c_j T^{\circ i} \circ (T^*)^{\circ j}$, 可得

$$f(T) \circ f(T)^* = \sum_{i,j} c_i c_j T^{\circ i} \circ (T^*)^{\circ j}.$$

同理

$$f(T)^* \circ f(T) = \sum_{i,j} c_j c_i (T^*)^{\circ j} \circ T^{\circ i}.$$

故依假設 $T \circ T^* = T^* \circ T$, 知 $T^{\circ i} \circ (T^*)^{\circ j} = (T^*)^{\circ j} \circ T^{\circ i}$, 得證 $f(T) \circ f(T)^* = f(T)^* \circ f(T)$. □

接下來我們探討 normal operator 的性質, 在這一節我們先探討一些在 $F = \mathbb{C}$ 和 $F = \mathbb{R}$ 時都會對的性質, 下一節我們在分別針對 $F = \mathbb{C}$ 和 $F = \mathbb{R}$ 來探討個別的性質.

首先由 normal 的定義 (即 $T \circ T^* = T^* \circ T$), 對任意 $\mathbf{v}, \mathbf{w} \in V$, 我們有

$$\langle T(\mathbf{v}), T(\mathbf{w}) \rangle = \langle \mathbf{v}, T^*(T(\mathbf{w})) \rangle = \langle \mathbf{v}, T(T^*(\mathbf{w})) \rangle = \langle T^*(\mathbf{v}), T^*(\mathbf{w}) \rangle.$$

因此對任意 $\mathbf{v} \in V$ 可得

$$\|T(\mathbf{v})\|^2 = \langle T(\mathbf{v}), T(\mathbf{v}) \rangle = \langle T^*(\mathbf{v}), T^*(\mathbf{v}) \rangle = \|T^*(\mathbf{v})\|^2.$$

得證以下性質.

Lemma 5.5.3. 假設 $T : V \to V$ 為 *normal operator,* 對任意 $\mathbf{v}, \mathbf{w} \in V$ 我們有以下的性質.

$$\langle T(\mathbf{v}), T(\mathbf{w}) \rangle = \langle T^*(\mathbf{v}), T^*(\mathbf{w}) \rangle \quad \text{and} \quad \|T(\mathbf{v})\| = \|T^*(\mathbf{v})\|.$$

現若 $\mathbf{v} \in \mathrm{Ker}(T)$, 則因 $\|T(\mathbf{v})\| = \|\mathbf{O}_V\| = 0$, 利用 Lemma 5.5.3 可得 $\|T^*(\mathbf{v})\| = 0$, 亦即 $T^*(\mathbf{v}) = \mathbf{O}_V$. 也就是說 $\mathbf{v} \in \mathrm{Ker}(T^*)$, 得證 $\mathrm{Ker}(T) \subseteq \mathrm{Ker}(T^*)$. 同理可得 $\mathrm{Ker}(T^*) \subseteq \mathrm{Ker}(T)$, 故知 $\mathrm{Ker}(T) = \mathrm{Ker}(T^*)$. 事實上由 Proposition 5.3.8, 我們可得下列性質.

Lemma 5.5.4. 假設 $T : V \to V$ 為 *normal operator,* 則我們有以下的性質.

$$\mathrm{Ker}(T) = \mathrm{Ker}(T^*), \quad \mathrm{Im}(T) = \mathrm{Im}(T^*), \quad \text{and} \quad \mathrm{Ker}(T) \cap \mathrm{Im}(T) = \{\mathbf{O}_V\}.$$

Proof. 因 $T^* \circ T = T \circ T^*$, 由 Proposition 5.3.8 知

$$\mathrm{Ker}(T) = \mathrm{Ker}(T^* \circ T) = \mathrm{Ker}(T \circ T^*) = \mathrm{Ker}(T^*),$$

$$\mathrm{Im}(T) = \mathrm{Im}(T \circ T^*) = \mathrm{Im}(T^* \circ T) = \mathrm{Im}(T^*).$$

由此再利用 Proposition 5.3.8 所知的 $\mathrm{Ker}(T^*) = \mathrm{Im}(T)^{\perp}$, 得

$$\mathrm{Ker}(T) = \mathrm{Ker}(T^*) = \mathrm{Im}(T)^{\perp}.$$

得證

$$\mathrm{Ker}(T) \cap \mathrm{Im}(T) = \mathrm{Im}(T)^{\perp} \cap \mathrm{Im}(T) = \{\mathbf{O}_V\}.$$

□

利用 Lemma 5.5.4, 我們馬上得到關於 normal operator 非常重要的性質.

Corollary 5.5.5. *假設 $T : V \to V$ 為 normal operator, 則我們有以下的性質.*

(1) *$\mathbf{v} \in V$ 是 T 的 eigenvector 且其 eigenvalue 為 λ 若且唯若 $\mathbf{v} \in V$ 是 T^* 的 eigenvector 且其 eigenvalue 為 $\overline{\lambda}$.*

(2) *對任意 $m \in \mathbb{N}$, $\mathrm{Ker}(T^{\circ m}) = \mathrm{Ker}(T)$.*

(3) *若 $\mathbf{v}, \mathbf{w} \in V$ 且 $\mu_{\mathbf{v}}(x)$ 和 $\mu_{\mathbf{w}}(x)$ 為 relatively prime (互質), 則 $\mathbf{v} \perp \mathbf{w}$.*

Proof. (1) 我們知道 T 的 eigenvalue 為 λ 的 eigenvector 為 $\mathrm{Ker}(T - \lambda \mathrm{id}_V)$ 中的 nonzero element, 而 T^* 的 eigenvalue 為 $\overline{\lambda}$ 的 eigenvector 為 $\mathrm{Ker}(T^* - \overline{\lambda} \mathrm{id}_V)$ 中的 nonzero element. 由 Lemma 5.5.2(4) 我們知 $T - \lambda \mathrm{id}_V$ 亦為 normal operator, 故由 Lemma 5.5.4 知

$$\mathrm{Ker}(T - \lambda \mathrm{id}_V) = \mathrm{Ker}((T - \lambda \mathrm{id}_V)^*) = \mathrm{Ker}(T^* - \overline{\lambda} \mathrm{id}_V^*) = \mathrm{Ker}(T^* - \overline{\lambda} \mathrm{id}_V).$$

證得 eigenvalue 為 λ 的 T 的 eigenvector 就是 eigenvalue 為 $\overline{\lambda}$ 的 T^* 的 eigenvector.

(2) 首先證明 $\mathrm{Ker}(T^{\circ 2}) = \mathrm{Ker}(T)$. 很明顯的 $\mathrm{Ker}(T) \subseteq \mathrm{Ker}(T^{\circ 2})$. 現若 $\mathbf{v} \in \mathrm{Ker}(T^{\circ 2})$, 則因 $T(\mathbf{v}) \in \mathrm{Ker}(T)$ 且 $T(\mathbf{v}) \in \mathrm{Im}(T)$. 由 Lemma 5.5.4 可得 $T(\mathbf{v}) = \mathbf{O}_V$, 故知 $\mathbf{v} \in \mathrm{Ker}(T)$. 即 $\mathrm{Ker}(T^{\circ 2}) \subseteq \mathrm{Ker}(T)$, 得證 $\mathrm{Ker}(T^{\circ 2}) = \mathrm{Ker}(T)$.

接著利用數學歸納法, 若 $\mathbf{v} \in \mathrm{Ker}(T^{\circ m})$, 則由 $T^{\circ m-1}(\mathbf{v}) \in \mathrm{Ker}(T) \cap \mathrm{Im}(T)$, 得 $T^{\circ m-1}(\mathbf{v}) = \mathbf{O}_V$, 即 $\mathbf{v} \in \mathrm{Ker}(T^{\circ m-1})$. 依歸納假設, 得 $\mathbf{v} \in \mathrm{Ker}(T^{\circ m-1}) = \mathrm{Ker}(T)$. 得證 $\mathrm{Ker}(T^{\circ m}) = \mathrm{Ker}(T)$.

(3) 由 $\mu_{\mathbf{v}}(x)$ 和 $\mu_{\mathbf{w}}(x)$ 互質知存在 $f(x), g(x) \in F[x]$ 使得 $f(x)\mu_{\mathbf{v}}(x) + g(x)\mu_{\mathbf{w}}(x) = 1$. 由此得

$$\mathbf{v} = f(T) \circ \mu_{\mathbf{v}}(T)(\mathbf{v}) + g(T) \circ \mu_{\mathbf{w}}(T)(\mathbf{v}) = g(T) \circ \mu_{\mathbf{w}}(T)(\mathbf{v}).$$

另一方面, 由 $g(T) \circ \mu_{\mathbf{w}}(T)(\mathbf{w}) = \mathbf{O}_V$ 知 $\mathbf{w} \in \mathrm{Ker}(g(T) \circ \mu_{\mathbf{w}}(T))$, 又由 Lemma 5.5.2(4) 知 $g(T) \circ \mu_{\mathbf{w}}(T)$ 亦為 normal operator, 故再由 Lemma 5.5.4 知 $\mathbf{w} \in \mathrm{Ker}((g(T) \circ \mu_{\mathbf{w}}(T))^*)$, 亦即 $(g(T) \circ \mu_{\mathbf{w}}(T))^*(\mathbf{w}) = \mathbf{O}_V$. 最後由 adjoint 的性質知

$$\langle \mathbf{v}, \mathbf{w} \rangle = \langle g(T) \circ \mu_{\mathbf{w}}(T)(\mathbf{v}), \mathbf{w} \rangle = \langle \mathbf{v}, (g(T) \circ \mu_{\mathbf{w}}(T))^*(\mathbf{w}) \rangle = \langle \mathbf{v}, \mathbf{O}_V \rangle = 0.$$

□

回顧當 $T : V \to V$ 是 linear operator, 我們有所謂的 primary decomposition theorem, 也就是說若 $\mu_T(x) = p_1(x)^{m_1} \cdots p_k(x)^{m_k}$, 其中 $m_1, \ldots, m_k \in \mathbb{N}$ 且 $p_1(x), \ldots, p_k(x) \in F[x]$ 為相異的 monic irreducible polynomial, 則 $V = W_1 \oplus \cdots \oplus W_k$, 其中 $W_i = \mathrm{Ker}(p_i(T)^{\circ m_i})$. 當 T 為 normal operator 時, 其 primary decomposition 有以下特殊的形式.

Proposition 5.5.6. 假設 $T:V\to V$ 為 *normal operator,* 則 $\mu_T(x)=p_1(x)\cdots p_k(x)$, 其中 $p_1(x),\ldots,p_k(x)\in F[x]$ 為相異的 *monic irreducible polynomial.* 若對所有 $i=1,\ldots,n$ 令 $W_i=\mathrm{Ker}(p_i(T))$, 則

$$V=W_1\boxplus\cdots\boxplus W_k.$$

Proof. 假設 $\mu_T(x)=p_1(x)^{m_1}\cdots p_k(x)^{m_k}$. 利用 Lemma 5.5.2(4) 我們知 $p_i(T)$ 亦為 normal operator, 故由 Corollary 5.5.5(2) 知 $W_i=\mathrm{Ker}(p_i(T)^{\circ m_i})=\mathrm{Ker}(p_i(T))$. 換言之對任意 $i=1,\ldots,k$, 皆有 $\mu_{T|_{W_i}}(x)=p_i(x)$, 得證

$$\mu_T(x)=\mu_{T|_{W_i}}(x)\cdots\mu_{T|_{W_k}}(x)=p_1(x)\cdots p_k(x).$$

現若 $i\neq j$, 因 $\mu_{T|_{W_i}}(x)=p_i(x)$ 和 $\mu_{T|_{W_j}}(x)=p_j(x)$ 為 relatively prime, 對任意 $\mathbf{w}_i\in W_i$, $\mathbf{w}_j\in W_j$ 因 $\mu_{\mathbf{w}_i}(x)\mid p_i(x)$ 和 $\mu_{\mathbf{w}_j}(x)\mid p_j(x)$, 知 $\mu_{\mathbf{w}_i}(x)$ 和 $\mu_{\mathbf{w}_j}(x)$ 為 relatively prime. 利用 Corollary 5.5.5(3), 知 $\mathbf{w}_i\perp\mathbf{w}_j$, 得證 $W_i\perp W_j$, 故知 $V=W_1\boxplus\cdots\boxplus W_k$. □

接下來我們介紹兩種特別的 normal operator.

Definition 5.5.7. 假設 $T:V\to V$ 為 linear operator. 若 $T^*=T$, 則稱 T 為 *self-adjoint* operator. 若 $T^*=-T$, 則稱 T 為 *skew-adjoint* operator.

若一個 $n\times n$ matrix A 滿足 $A=A^*$, 則稱 A 為 *self-adjoint* matrix. 若 $A=-A^*$, 則稱 A 為 *skew-adjoint* matrix.

注意在 $A\in M_n(\mathbb{R})$ 的情形, 當 A 為 self-adjoint matrix (即 $A^{\mathrm{t}}=A$) 一般也稱為 *symmetric* matrix. 而當 A 為 skew-adjoint matrix (即 $A^{\mathrm{t}}=-A$) 一般也稱為 *skew-symmetric* matrix. 在 $A\in M_n(\mathbb{C})$ 的情形, 有的書將 self-adjoint 稱為 *Hermitian* 而將 skew-adjoint 稱為 *skew-Hermitian.*

很明顯的 self-adjoint operator 和 skew-adjoint operator 皆為 normal operator. 我們曾在上一節介紹過 self-adjoint operator. 事實上特別介紹這兩種 normal operator 是因為有以下重要的性質.

Proposition 5.5.8. 假設 $T:V\to V$ 為 *linear operator.* 存在唯一的 *self-adjoint operator* $T_1:V\to V$ 以及唯一的 *skew-adjoint operator* $T_2:V\to V$ 滿足 $T=T_1+T_2$. 特別的, 我們有 T 為 *normal operator* 若且唯若 $T_1\circ T_2=T_2\circ T_1$.

Proof. 若 T_1 為 self-adjoint, T_2 為 skew-adjoint 且 $T_1+T_2=T$. 則由 Proposition 5.3.7(1) 知 $T^*=T_1^*+T_2^*=T_1-T_2$, 得 $T_1=\frac{1}{2}(T+T^*)$, $T_2=\frac{1}{2}(T-T^*)$. 故得證唯一性. 另一方面若 $T_1=\frac{1}{2}(T+T^*)$, $T_2=\frac{1}{2}(T-T^*)$, 則由 Proposition 5.3.7(1) 得 $T_1^*=T_1$, $T_2^*=-T_2$. 亦即 T_1 為 self-adjoint, T_2 為 skew-adjoint 且 $T_1+T_2=T$. 故得證存在性.

現由

$$T\circ T^*=(T_1+T_2)\circ(T_1-T_2)=T_1^{\circ 2}-T_1\circ T_2+T_2\circ T_1-T_2^{\circ 2},$$
$$T^*\circ T=(T_1-T_2)\circ(T_1+T_2)=T_1^{\circ 2}+T_1\circ T_2-T_2\circ T_1-T_2^{\circ 2}.$$

得 $T\circ T^*=T^*\circ T$ 若且唯若 $T_1\circ T_2=T_2\circ T_1$. □

和 normal operator 的情況一樣, 我們來看看如何可得到更多的 self-adjoint operator.

Lemma 5.5.9. 假設 $T:V\to V$, $T_1:V\to V$ 為 *self-adjoint operator.*

(1) $T+T_1$ 亦為 *self-adjoint operator.*

(2) 若 W 為 T*-invariant,* 則 $T|_W$ 亦為 *self-adjoint operator.*

(3) 若 T 為 *isomorphism,* 則 T^{-1} 亦為 *self-adjoint operator.*

(4) 對任意 $f(x)\in\mathbb{R}[x]$, $f(T)$ 亦為 *self-adjoint operator.*

Proof. (1) 因 $(T+T_1)^*=T^*+T_1^*=T+T_1$, 故 $T+T_1$ 為 self-adjoint operator.

(2) 由 $T=T^*$ 得 W 亦為 T^*-invariant, 故由 Corollary 5.4.4, 得 $(T|_W)^*=T^*|_W=T|_W$.

(3) 由 Proposition 5.3.7(3) 知 $(T^{-1})^*=(T^*)^{-1}=T^{-1}$, 故 T^{-1} 為 self-adjoint.

(4) 由 Lemma 5.4.1 知 $(f(T))^*=\overline{f}(T^*)=\overline{f}(T)$. 但 $f(x)\in\mathbb{R}[x]$, 故有 $\overline{f}(x)=f(x)$, 得證 $f(T)$ 為 self-adjoint operator. □

Question 5.23. 若 $T:V\to V$, $T_1:V\to V$ 為 *normal operator,* 是否 $T+T_1$ 亦為 *normal operator?*

Question 5.24. 假設 $T:V\to V$ 為 *self-adjoint operator* 且 $f(x)\in F[x]$. 試證明 $f(T)$ 為 *self-adjoint operator* 若且唯若 $\mu_T(x)\mid f(x)-\overline{f}(x)$.

關於 self-adjoint operator 我們有以下重要的性質.

Lemma 5.5.10. 假設 $T:V\to V$ 為 *self-adjoint operator.* 若 λ 為 T 的 *eigenvalue,* 則 $\lambda\in\mathbb{R}$.

Proof. 因 T 為 normal, 由 Corollary 5.5.5(1) 知, 若 $\mathbf{v}\in V$ 為相對於 λ 的 T 的 eigenvector, 即 $T(\mathbf{v})=\lambda\mathbf{v}$, 則 $\mathbf{v}$ 亦為相對於 $\overline{\lambda}$ 的 T^* 的 eigenvector, 即 $T^*(\mathbf{v})=\overline{\lambda}\mathbf{v}$. 但由 T 為 self-adjoint 之假設, 我們有 $T(\mathbf{v})=T^*(\mathbf{v})$, 亦即 $\lambda\mathbf{v}=\overline{\lambda}\mathbf{v}$. 故由 $\mathbf{v}\neq\mathbf{O}_V$, 得知 $\lambda=\overline{\lambda}$, 即 $\lambda\in\mathbb{R}$. □

至於 skew-adjoint operator, 我們也有和 self-adjoint 相類似的性質. 以下因證明和 Lemma 5.5.9 相似, 我們就不再證明.

Lemma 5.5.11. 假設 $T:V\to V$, $T_1:V\to V$ 為 *skew-adjoint operator.*

(1) $T+T_1$ 亦為 *skew-adjoint operator.*

(2) 若 W 為 T*-invariant,* 則 $T|_W$ 亦為 *skew-adjoint operator.*

(3) 若 T 為 *isomorphism,* 則 T^{-1} 亦為 *skew-adjoint operator.*

Question 5.25. 假設 $T:V\to V$ 為 *skew-adjoint operator.* 是否對任意 $f(x)\in\mathbb{R}[x]$, $f(T)$ 亦為 *skew-adjoint operator?*

關於 skew-adjoint operator 我們有以下重要的性質.

Lemma 5.5.12. 假設 $T : V \to V$ 為 *skew-adjoint operator.* 若 λ 為 T 的 *eigenvalue,* 則 λ 的 *real part* (實部) 為 0.

Proof. 因 T 為 normal, 由 Corollary 5.5.5(1) 知, 若 $\mathbf{v} \in V$ 為相對於 λ 的 T 的 eigenvector, 即 $T(\mathbf{v}) = \lambda\mathbf{v}$, 則 $\mathbf{v}$ 亦為相對於 $\overline{\lambda}$ 的 T^* 的 eigenvector, 即 $T^*(\mathbf{v}) = \overline{\lambda}\mathbf{v}$. 但由 T 為 skew-adjoint 之假設, 我們有 $T^*(\mathbf{v}) = -T(\mathbf{v})$, 亦即 $-\lambda\mathbf{v} = \overline{\lambda}\mathbf{v}$. 故由 $\mathbf{v} \neq \mathbf{O}_V$, 得知 $\lambda + \overline{\lambda} = 0$, 即 λ 之實部為 0. □

上面所提有關 normal operator, self-adjoint operator 以及 skew-adjoint operator 的性質都可以套用在相對應的 normal matrix, self-adjoint matrix 以及 skew-adjoint matrix. 例如任意的 $n \times n$ matrix 都可以唯一寫成一個 self-adjoint matrix 和一個 skew-adjoint matrix 之和. 這些請大家自行轉換, 我們就不再贅述了.

5.6. The Spectral Theorem

這一節中我們將 over $\mathbb{C}$ 和 over $\mathbb{R}$ 的情況分開討論, 進而推導出 normal operator 的 spectral theorem.

5.6.1. Normal Operator over $\mathbb{C}$. 首先我們考慮 V 為 finite dimensional inner product space over $\mathbb{C}$ 且 $T : V \to V$ 為 normal operator 的情形. 此時因 $\mathbb{C}$ 為 algebraically closed, 我們可以將 T 的 minimal polynomial 完全分解. 亦即存在 $\lambda_1, \ldots, \lambda_k \in \mathbb{C}$ 以及 $m_1, \ldots, m_k \in \mathbb{N}$ 使得 $\mu_T(x) = (x-\lambda_1)^{m_1} \cdots (x-\lambda_k)^{m_k}$. 依此, 利用 Proposition 5.5.6 我們可以得到 over $\mathbb{C}$ 的 normal operator 的 spectral theorem.

Theorem 5.6.1 (Spectral Theorem: complex case)**.** 假設 V 為 *finite dimensional inner product space over* $\mathbb{C}$ 且 $T : V \to V$ 為 *linear operator.* 則 T 為 *unitary diagonalizable* 若且唯若 T 為 *normal operator.*

Proof. 由 Proposition 5.4.10 我們知道若 T 為 unitary diagonalizable 則 T 為 normal operator.

現假設 T 為 normal operator 由 Proposition 5.5.6, 得 $\mu_T(x) = (x-\lambda_1)\cdots(x-\lambda_k)$, 其中 $\lambda_1, \ldots, \lambda_k \in \mathbb{C}$ 為相異, 且 $V = W_1 \boxplus \cdots \boxplus W_k$, 其中 $W_i = \mathrm{Ker}((T-\lambda_i \mathrm{id}_V)) = E_{\lambda_i}$ 為 λ_i 的 eigenspace, 得

$$V = E_{\lambda_1} \boxplus \cdots \boxplus E_{\lambda_k},$$

故由 Proposition 5.4.10 知 T 為 unitary diagonalizable. □

考慮 $\mathbb{C}^n$ 上的 standard inner product, 利用 Theorem 5.6.1, 我們馬上得到矩陣形式的 spectral theorem.

Corollary 5.6.2. 假設 $A \in M_n(\mathbb{C})$. 則 A 為 *normal matrix* (即 $AA^* = A^*A$) 若且唯若存在 *unitary matrix* $P \in M_n(\mathbb{C})$ (即 $P^{-1} = P^*$) 使得 P^*AP 為 *diagonal matrix.*

我們馬上得到有關 self-adjoint operator 以及 skew-adjoint operator 的 spectral theorem.

Corollary 5.6.3. 假設 V 為 *finite dimensional inner product space over* $\mathbb{C}$ 且 $T:V\to V$ 為 *linear operator.* 我們有以下性質.

(1) T 為 *self-adjoint operator* 若且唯若 T 為 *unitary diagonalizable* 且 T 的 *eigenvalue* 皆為實數.

(2) T 為 *skew-adjoint operator* 若且唯若 T 為 *unitary diagonalizable* 且 T 的 *eigenvalue* 的實部為 0.

Proof. 利用 Theorem 5.6.1, Lemma 5.5.10 我們得若 T 為 self-adjoint operator 則 T 為 unitary diagonalizable 且 T 的 eigenvalue 皆為實數. 反之若 T 為 unitary diagonalizable 且 T 的 eigenvalue 皆為實數, 則由 Proposition 5.4.10 知存在 $\lambda_1,\dots,\lambda_k\in\mathbb{R}$ 以及 orthogonal projections π_i 使得 $T=\lambda_1\pi_1+\cdots+\lambda_k\pi_k$. 此時 $T^*=\lambda_1\pi_1^*+\cdots+\lambda_k\pi_k^*$ 且因 $\pi_i^*=\pi_i$, 得證 $T^*=T$. 同理可證得 skew-adjoint 的情形, 在此略過. □

Question 5.26. 若 $T:V\to V$ 為 *self-adjoint,* 則 $\chi_T(x),\mu_T(x)$ 為何種形式? 若 $T:V\to V$ 為 *skew-adjoint,* 則 $\chi_T(x),\mu_T(x)$ 為何種形式?

同樣的我們也有 self-adjoint matrix 和 skew-adjoint 相對應的結果.

Corollary 5.6.4. 假設 $A\in M_n(\mathbb{C})$. 我們有以下性質.

(1) $A^*=A$ 若且唯若存在 *unitary matrix* $P\in M_n(\mathbb{C})$ 使得

$$P^*AP=\begin{pmatrix}\gamma_1 & & \mathbf{O}\\ & \ddots & \\ \mathbf{O} & & \gamma_n\end{pmatrix},$$

其中 $\gamma_1,\dots,\gamma_n\in\mathbb{R}$.

(2) $A^*=-A$ 若且唯若存在 *unitary matrix* $P\in M_n(\mathbb{C})$ 使得

$$P^*AP=\begin{pmatrix}\gamma_1 & & \mathbf{O}\\ & \ddots & \\ \mathbf{O} & & \gamma_n\end{pmatrix},$$

其中 $\gamma_1,\dots,\gamma_n$ 的實部為 0.

注意 Corollary 5.6.4 中 $\gamma_1,\dots,\gamma_n$ 不必相異. 另外在 skew-adjoint 的情形 γ_i 的實部為 0 一般稱為 purely imaginary number (純虛數), 不過這裡 γ_i 有可能為 0.

5.6.2. Normal operator over $\mathbb{R}$. 在 over $\mathbb{C}$ 的情況我們用到 $\mathbb{C}$ 為 algebraically closed, 由於 $\mathbb{R}[x]$ 中的 irreducible polynomial, 有可能是二次的, 因此我們需要用其他的方法處理 over $\mathbb{R}$ 的情形. 假設 V 為 over $\mathbb{R}$ 的 finite dimensional inner product space, 首先注意在此情形 $T:V\to V$ 為 normal operator 並不一定為 unitary (orthogonal) diagonalizable. 事實上由 Proposition 5.4.10 我們知道若 T 為 unitary diagonalizable, 則存在 $\lambda_1,\dots\lambda_k\in\mathbb{R}$ 且 $\pi_1,\dots,\pi_k$ 為 orthogonal projection 使得 $T=\lambda_1\pi_1+\cdots+\lambda_k\pi_k$. 此時因

$$T^*=\overline{\lambda_1}\pi_1^*+\cdots+\overline{\lambda_k}\pi_k^*=\lambda_1\pi_1+\cdots+\lambda_k\pi_k=T,$$

得知 T 必為 self-adjoint operator. 事實上我們有以下之結果.

Theorem 5.6.5. 假設 V 為 *finite dimensional inner product space over* $\mathbb{R}$ 且 $T:V\to V$ 為 *linear operator,* 則 T 為 *unitary diagonalizable* 若且唯若 T 為 *self-adjoint operator.*

Proof. 前面已知若 T 為 unitary diagonalizable 則 T 為 self-adjoint operator. 現假設 T 為 self-adjoint operator, 因 T 為 normal, 故由 Proposition 5.5.6 知 $\mu_T(x)=p_1(x)\cdots p_k(x)$, 其中 $p_1(x),\ldots,p_k(x)$ 為相異的 monic irreducible polynomial in $\mathbb{R}[x]$, 且 $V=W_1\boxplus\cdots\boxplus W_k$, 其中 $W_i=\mathrm{Ker}(p_i(T))$. 我們欲說明對所有 $i=1,\ldots,k$, $p_i(x)$ 皆為一次的 polynomial, 即存在 $\lambda_i\in\mathbb{R}$ 使得 $p_i(x)=x-\lambda_i$. 此即得證 $W_i=E_{\lambda_i}$ 為 λ_i 的 eigenspace, 即 $V=E_{\lambda_1}\boxplus\cdots\boxplus E_{\lambda_k}$, 故由 Proposition 5.4.10 知 T 為 unitary diagonalizable.

現任取 V 的一組 orthonormal basis 所組成的 ordered basis $\beta=(\mathbf{v}_1,\ldots,\mathbf{v}_n)$. 由 $T=T^*$ 以及 Proposition 5.3.10 知 $[T]_\beta^*=[T^*]_\beta=[T]_\beta$, 亦即 $[T]_\beta$ 為 self-adjoint matrix in $M_n(\mathbb{R})$. 當然 $[T]_\beta$ 看成 $M_n(\mathbb{C})$ 的 matrix 亦為 self-adjoint, 故套用 Corollary 5.6.4(1), 我們知存在 $\gamma_1,\ldots,\gamma_n\in\mathbb{R}$ 使得

$$\chi_{[T]_\beta}(x)=\chi_T(x)=(x-\gamma_1)\cdots(x-\gamma_n).$$

亦即 $\chi_T(x)$ 可以完全分解成 $\mathbb{R}[x]$ 中的一次多項式的乘積, 故由 $\mu_T(x)\mid\chi_T(x)$, 得證存在 $\lambda_i\in\mathbb{R}$ 使得 $p_i(x)=x-\lambda_i$. □

我們馬上得到有關於 symmetric matrix 相對應的性質.

Corollary 5.6.6. 假設 $A\in M_n(\mathbb{R})$, 則 $A^{\mathrm{t}}=A$ 若且唯若存在 *orthogonal matrix* $P\in M_n(\mathbb{R})$ (即 $P^{-1}=P^{\mathrm{t}}$) 使得 $P^{\mathrm{t}}AP$ 為 $M_n(\mathbb{R})$ 中的 *diagonal matrix.*

接下來我們探討 over $\mathbb{R}$ 的 skew-adjoint operator.

Lemma 5.6.7. 假設 V 為 *finite dimensional inner product space over* $\mathbb{R}$ 且 $T:V\to V$ 為 *linear operator,* 則 T 為 *skew-adjoint* 若且唯若對任意 $\mathbf{v}\in V$ 皆有 $\langle T(\mathbf{v}),\mathbf{v}\rangle=0$.

Proof. 首先強調 over $\mathbb{R}$ 的 inner product 有對稱性, 即 $\langle\mathbf{v},\mathbf{w}\rangle=\langle\mathbf{w},\mathbf{v}\rangle,\forall\mathbf{v},\mathbf{w}\in V$. 現假設 $T^*=-T$, 則對任意 $\mathbf{v}\in V$ 皆有

$$\langle T(\mathbf{v}),\mathbf{v}\rangle=\langle\mathbf{v},T^*(\mathbf{v})\rangle=\langle\mathbf{v},-T(\mathbf{v})\rangle=-\langle T(\mathbf{v}),\mathbf{v}\rangle,$$

得證 $\langle T(\mathbf{v}),\mathbf{v}\rangle=0$.

反之, 若對任意 $\mathbf{v}\in V$ 皆有 $\langle T(\mathbf{v}),\mathbf{v}\rangle=0$, 則對任意 $\mathbf{v},\mathbf{w}\in V$, 我們有

$$0=\langle T(\mathbf{v}+\mathbf{w}),\mathbf{v}+\mathbf{w}\rangle=\langle T(\mathbf{v}),\mathbf{w}\rangle+\langle T(\mathbf{w}),\mathbf{v}\rangle.$$

得知對任意 $\mathbf{v},\mathbf{w}\in V$ 皆有

$$\langle T(\mathbf{v}),\mathbf{w}\rangle=-\langle T(\mathbf{w}),\mathbf{v}\rangle=\langle\mathbf{v},-T(\mathbf{w})\rangle,$$

利用 adjoint 的唯一性 (Theorem 5.3.5) 得證 $T^*=-T$. □

當 T 為 skew-adjoint, 假設 $x-\lambda \in \mathbb{R}[x]$ 是 $\mu_T(x)$ 的一次因式, 即 $\lambda \in \mathbb{R}$ 為 T 的一個 eigenvalue. 令 $\mathbf{v} \in V$ 為相對於 λ 的 eigenvector, 即 $T(\mathbf{v}) = \lambda\mathbf{v}$. 由 Lemma 5.6.7 知 $0 = \langle T(\mathbf{v}), \mathbf{v}\rangle = \langle \lambda\mathbf{v}, \mathbf{v}\rangle = \lambda\langle \mathbf{v}, \mathbf{v}\rangle$. 由於 $\mathbf{v} \neq \mathbf{O}_V$, 得知 $\lambda = 0$. 也就是說 x 是 $\mu_T(x)$ 唯一可能的實係數一次因式. 事實上我們有以下結果.

Corollary 5.6.8. 假設 V 為 *finite dimensional inner product space over* $\mathbb{R}$ 且 $T: V \to V$ 為 *nonzero skew-adjoint operator.* 若 T 為 *isomorphism,* 則存在 $b_1, \ldots, b_k \in \mathbb{R}$ 皆相異且不為 0 使得 $\mu_T(x) = (x^2+b_1^2)\cdots(x^2+b_k^2)$. 若 T 不是 *isomorphism,* 則存在 $b_1, \ldots, b_k \in \mathbb{R}$ 皆相異且不為 0 使得 $\mu_T(x) = x(x^2+b_1^2)\cdots(x^2+b_k^2)$.

Proof. 我們已知 $\mu_T(x)$ 在 $\mathbb{R}[x]$ 中可能的一次因式為 x. 現在考慮 $\mu_T(x)$ 在 $\mathbb{R}[x]$ 中可能的二次質因式. 假設 $x^2+rx+s \in \mathbb{R}[x]$, 為 $\mu_T(x)$ 的一個二次質因式. 則由 cyclic decomposition theorem 知存在 $\mathbf{v} \in V$ 滿足 $\mu_{\mathbf{v}}(x) = x^2+rx+s$. 由 Lemma 5.6.7, 我們知 $\langle T^{\circ 2}(\mathbf{v}), T(\mathbf{v})\rangle = 0$. 又 $T^{\circ 2}(\mathbf{v}) = -rT(\mathbf{v}) - s\mathbf{v}$, 故得

$$0 = \langle -rT(\mathbf{v}) - s\mathbf{v}, T(\mathbf{v})\rangle = -r\langle T(\mathbf{v}), T(\mathbf{v})\rangle - s\langle \mathbf{v}, T(\mathbf{v})\rangle = -r\langle T(\mathbf{v}), T(\mathbf{v})\rangle.$$

然而 $T(\mathbf{v}) \neq 0$ (否則 $\mu_{\mathbf{v}}(x) \mid x$), 故得 $r = 0$. 而又若 $s \leq 0$, 則 x^2+s 不為 irreducible polynomial, 故知 $s > 0$. 得知 $\mu_T(x)$ 唯一可能的二次質因式為 x^2+b^2 其中 $b \neq 0$ 這樣的形式.

現若 T 為 isomorphism, 即 $\mathrm{Ker}(T) = \{\mathbf{O}_V\}$, 得知 x 不可能為 $\mu_T(x)$ 的因式. 故 $\mu_T(x)$ 只有二次的質因式, 又由 Proposition 5.5.6 知 $\mu_T(x)$ 的質因式不會重複出現, 證得存在 $b_1, \ldots, b_k \in \mathbb{R}$ 皆相異且不為 0 使得 $\mu_T(x) = (x^2+b_1^2)\cdots(x^2+b_k^2)$. 而若 T 不是 isomorphism, 即 $\dim(\mathrm{Ker}(T)) > 0$, 得知 x 必為 $\mu_T(x)$ 的因式, 證得存在 $b_1, \ldots, b_k \in \mathbb{R}$ 皆相異且不為 0 使得 $\mu_T(x) = x(x^2+b_1^2)\cdots(x^2+b_k^2)$. $\square$

若 T 為 skew-adjoint operator 且 $p(x)$ 是 $\mu_T(x)$ 的質因式, 令 $W = \mathrm{Ker}(p(T))$ 由 primary decomposition theorem, 我們需要了解 $T|_W$. 當 $p(x) = x$ 時, $W = \mathrm{Ker}(T)$ 此時 $T|_W$ 為 zero mapping. 故僅要考慮 $p(x) = x^2+b^2$, $b \neq 0$ 的情形. 此時 $W = \mathrm{Ker}(T^{\circ 2} + b^2\mathrm{id}_V)$, 為 T-invariant, 故由 Lemma 5.5.11(2), $T|_W$ 亦為 skew-adjoint. 又此時 $\mu_{T|_W}(x) = x^2+b^2$, 所以我們僅要考慮以下之情形.

Lemma 5.6.9. 假設 V 為 *finite dimensional inner product space over* $\mathbb{R}$ 且 $T: V \to V$ 為 *skew-adjoint operator.* 假設 $\mu_T(x) = x^2+b^2$ 其中 $b \neq 0$, 則存在 V 的一組 *ordered orthonormal basis* β 使得 $[T]_\beta$ 為以下的 *block diagonal matrix*

$$[T]_\beta = \begin{pmatrix} M & & \mathbf{O} \\ & \ddots & \\ \mathbf{O} & & M \end{pmatrix}, \text{其中 } M = \begin{pmatrix} 0 & -b \\ b & 0 \end{pmatrix}. \tag{5.4}$$

Proof. 任取 $\mathbf{v} \in V$ 且 $\|\mathbf{v}\| = 1$. 因 $T^{\circ 2} + b^2\mathrm{id}_V = \mathbf{O}$, 我們有 $T(T(\mathbf{v})) = -b^2\mathbf{v}$. 考慮 T-cyclic subspace $W = \mathrm{Span}(\{\mathbf{v}, T(\mathbf{v})\})$. 因

$$\|T(\mathbf{v})\|^2 = \langle T(\mathbf{v}), T(\mathbf{v})\rangle = \langle \mathbf{v}, T^*(T(\mathbf{v}))\rangle = \langle \mathbf{v}, -T(T(\mathbf{v}))\rangle = \langle \mathbf{v}, b^2\mathbf{v}\rangle = b^2.$$

令 $\mathbf{v}' = \frac{1}{b}T(\mathbf{v})$. 由 Lemma 5.6.7 知 $\langle \mathbf{v}', \mathbf{v} \rangle = \frac{1}{b}\langle T(\mathbf{v}), \mathbf{v} \rangle = 0$. 故 $\beta_1 = (\mathbf{v}, \mathbf{v}')$ 為 W 的一個 ordered orthonormal basis. 因

$$T(\mathbf{v}) = b\mathbf{v}', T(\mathbf{v}') = \frac{1}{b}T(T(\mathbf{v})) = -b\mathbf{v},$$

我們得

$$[T|_W]_{\beta_1} = \begin{pmatrix} 0 & -b \\ b & 0 \end{pmatrix}.$$

現考慮 $V = W \boxplus W^{\perp}$. 因 W 為 T-invariant, 由 Lemma 5.4.3 知 $W^{\perp}$ 為 T^*-invariant. 但 $T^* = -T$, 故知 $W^{\perp}$ 亦為 T-invariant. 再由 Lemma 5.5.11 知 $T|_{W^{\perp}}$ 亦為 skew-adjoint. 現因 $\mu_{T|_{W^{\perp}}}(x) \mid \mu_T(x)$, 得 $\mu_{T|_{W^{\perp}}}(x) = 1$ 或 $\mu_{T|_{W^{\perp}}}(x) = x^2 + b^2$. 若 $\mu_{T|_{W^{\perp}}}(x) = 1$, 表示 $W^{\perp} = \{\mathbf{O}_V\}$, 亦即表示 $V = W$, 得證本定理. 若 $\mu_{T|_{W^{\perp}}}(x) = x^2 + b^2$, 則我們可套用數學歸納法找到 $W^{\perp}$ 的一組 ordered orthonormal basis β_2 使得 $[T|_{W^{\perp}}]_{\beta_2}$ 為如 (5.4) 的 block diagonal matrix. 此時將 β_1, β_2 合組成 V 的一組 ordered orthonormal basis β, 則 $[T]_\beta$ 即為所求. □

Example 5.6.10. 假設 V 為 finite dimensional inner product space over $\mathbb{R}$ 且 $T : V \to V$ 為 skew-adjoint operator. 若 $\chi_T(x) = (x^2 + 2)^2$, 則由 Proposition 5.5.6 知 $\mu_T(x) = x^2 + 2$. 故由 Lemma 5.6.9 知存在 V 的一組 ordered orthonormal basis β 使得

$$[T]_\beta = \begin{pmatrix} 0 & -\sqrt{2} & 0 & 0 \\ \sqrt{2} & 0 & 0 & 0 \\ 0 & 0 & 0 & -\sqrt{2} \\ 0 & 0 & \sqrt{2} & 0 \end{pmatrix}.$$

Question 5.27. 假設 V 為 *finite dimensional inner product space over* $\mathbb{R}$ 且 $T : V \to V$ 為 *skew-adjoint operator.* 若已知 $\mu_T(x) = x^2 + b^2$ 其中 $b \neq 0$, 試說明 $\dim(V)$ 必為偶數.

對於一般的 over $\mathbb{R}$ 的 skew-adjoint operator $T : V \to V$, 由 Proposition 5.5.6 知 $V = W_1 \boxplus \cdots \boxplus W_k$, 其中 $W_i = \mathrm{Ker}(T)$ 或 $W_i = \mathrm{Ker}(T^{\circ 2} + b^2 \mathrm{id}_V)$. 若 $W_i = \mathrm{Ker}(T)$, 因 $T|_{W_i}$ 為 zero mapping, 任取 W_i 的一組 ordered orthonormal basis β_i, 皆會使得 $[T|_{W_i}]_{\beta_i}$ 為 zero matrix. 而若 $W_i = \mathrm{Ker}(T^{\circ 2} + b^2 \mathrm{id}_V)$, 由 Lemma 5.6.9 我們可以找到 W_i 的一組 ordered orthonormal basis β_i, 使得 $[T|_{W_i}]_{\beta_i}$ 為如 (5.4) 的 block diagonal matrix. 由於這些 $\beta_1, \ldots, \beta_k$ 可以組成 V 的一組 orthonormal ordered basis, 我們有以下之結果.

Theorem 5.6.11. 假設 V 為 *finite dimensional inner product space over* $\mathbb{R}$. 則 $T : V \to V$ 為 *skew-adjoint operator* 若且唯若存在 V 的一組 *ordered orthonormal basis* β 使得 $[T]_\beta$ 為以下的 *block diagonal matrix*

$$[T]_\beta = \begin{pmatrix} M_1 & & \mathbf{O} \\ & \ddots & \\ \mathbf{O} & & M_l \end{pmatrix}, \text{其中 } M_i = \begin{pmatrix} 0 & -b_i \\ b_i & 0 \end{pmatrix}, b_i \neq 0 \text{ 或 } M_i = 0. \tag{5.5}$$

Proof. 若 $T : V \to V$ 為 skew-adjoint operator 由前面的說明我們知存在 V 的一組 ordered orthonormal basis β 使得 $[T]_\beta$ 為如 (5.5) 的 block diagonal matrix.

反之, 若存在 V 的一組 ordered orthonormal basis β 使得 $[T]_\beta$ 為如 (5.5) 的 block diagonal matrix, 則由 Proposition 5.3.10 知

$$[T^*]_\beta = [T]_\beta^* = \begin{pmatrix} M_1^* & & \mathbf{O} \\ & \ddots & \\ \mathbf{O} & & M_l^* \end{pmatrix},$$

其中 $M_i^* = \begin{pmatrix} 0 & b_i \\ -b_i & 0 \end{pmatrix} = -M_i$ 或 $M_i^* = 0 = -M_i$. 換言之, $[T^*]_\beta = -[T]_\beta = [-T]_\beta$, 得證 $T^* = -T$. □

由 Theorem 5.6.11, 我們可得到有關於 skew-adjoint matrix 的性質. 也就是說 $A \in M_n(\mathbb{R})$ 滿足 $A^t = -A$ 若且唯若存在 orthogonal matrix $P \in M_n(\mathbb{R})$ 使得 P^tAP 為如 (5.5) 的 block diagonal matrix.

Question 5.28. 假設 $A \in M_n(\mathbb{R})$ 且 $A^t = -A$. 若已知 $\chi_A(x) = x^3(x^2+1)(x^2+2)^2$, 試寫下如 *(5.5)* 的 *block diagonal matrix* B 使得 $A \sim B$. 試寫下 A 的 *rational form (*或 *classical form)*. 為何 A 的 *rational form* 不是 *skew-symmetric?*

最後我們來探討一般 over $\mathbb{R}$ 的 normal operator. 若 $T: V \to V$ 為 normal operator, 則由 Proposition 5.5.6 我們有 $V = W_1 \boxplus \cdots \boxplus W_k$, 其中 $W_i = \operatorname{Ker}(p_i(T))$ 且 $p_i(x) \in \mathbb{R}[x]$ 為 monic irreducible polynomial. 我們知 $\mathbb{R}[x]$ 中的 monic irreducible polynomial 僅有以下兩種形式, 即 $x - \lambda, \lambda \in \mathbb{R}$, 以及 $(x-a)^2 + b^2$, $a, b \in \mathbb{R}$ 且 $b \neq 0$. 若 $p_i(x) = x - \lambda$, 則 $W_i = \operatorname{Ker}(p_i(T))$ 為 eigenvalue 為 λ 的 eigenspace. 所以我們僅要探討 $p_i(x) = (x-a)^2 + b^2$ 的情形. 注意由於 $W_i^\perp = \boxplus_{j \neq i} W_j$ 且 W_j 皆為 T-invariant, 得 $W_i^\perp$ 亦為 T-invariant. 故由 Lemma 5.5.2(2) 知 $T|_{W_i}$ 亦為 normal operator. 又此時 $\mu_{T|_{W_i}}(x) = (x-a)^2 + b^2$, 所以我們僅要考慮以下之情形.

Lemma 5.6.12. 假設 V 為 *finite dimensional inner product space over* $\mathbb{R}$, $T: V \to V$ 為 *normal operator* 且 $\mu_T(x) = (x-a)^2 + b^2$ 其中 $b \neq 0$. 令 $T = T_1 + T_2$, 其中 $T_1: V \to V$ 為 *self-adjoint operator,* $T_2: V \to V$ 為 *skew-adjoint operator.* 則

(1) T_2 為 *isomorphism.*

(2) $\mu_{T_1}(x) = x - a.$

(3) $\mu_{T_2}(x) = x^2 + b^2.$

Proof. Proposition 5.5.8 告訴我們 T_1, T_2 必存在且唯一. 又 T 為 normal 故 $T_1 \circ T_2 = T_2 \circ T_1$.

(1) 我們只要證明 $\operatorname{Ker}(T_2) = \{\mathbf{O}_V\}$, 便可得 T_2 為 isomorphism. 現假設 $\operatorname{Ker}(T_2) \neq \{\mathbf{O}_V\}$ 且令 $W = \operatorname{Ker}(T_2)$. 由 $T_1 \circ T_2 = T_2 \circ T_1$ 知對任意 $\mathbf{w} \in W$, $T_2(T_1(\mathbf{w})) = T_1(T_2(\mathbf{w})) = T_1(\mathbf{O}_V) = \mathbf{O}_V$. 亦即 $T_1(\mathbf{w}) \in \operatorname{Ker}(T_2) = W$, 得證 W 為 T_1-invariant. 又對任意 $\mathbf{w} \in W$, 皆有

$$T(\mathbf{w}) = T_1(\mathbf{w}) + T_2(\mathbf{w}) = T_1(\mathbf{w}), \text{ 即 } T|_W = T_1|_W.$$

由 Lemma 5.5.9 知 $T_1|_W$ 為 self-adjoint, 故 $T|_W$ 為 self-adjoint. 然而 $\mu_{T|_W}(x) \mid \mu_T(x)$ 且 $\mu_{T|_W}(x) \neq 1$ (否則 $W = \{\mathbf{O}_V\}$), 由 $(x-a)^2 + b^2$ $(b \neq 0)$ 在 $\mathbb{R}[x]$ 中為 irreducible 得 $\mu_{T|_W}(x) =$

$\mu_T(x) = (x-a)^2 + b^2$. 另一方面由 $T|_W$ 為 self-adjoint, Theorem 5.6.5 告訴我們 $T|_W$ 為 unitary-diagonalizable, 此即表示 $\mu_{T|_W}(x)$ 可以分解成 $\mathbb{R}[x]$ 中一次多項式的乘積. 由此矛盾得證 $W = \mathrm{Ker}(T_2) = \{\mathbf{O}_V\}$.

(2) 由 $T_1 \circ T_2 = T_2 \circ T_1$ 得

$$T^{\circ 2} = (T_1 + T_2) \circ (T_1 + T_2) = T_1{}^{\circ 2} + 2T_1 \circ T_2 + T_2{}^{\circ 2}.$$

同理由 $T^* = T_1^* + T_2^* = T_1 - T_2$, 得 $(T^*)^{\circ 2} = T_1{}^{\circ 2} - 2T_1 \circ T_2 + T_2{}^{\circ 2}$, 與上式相減得

$$T^{\circ 2} - (T^*)^{\circ 2} = 4T_1 \circ T_2. \tag{5.6}$$

另一方面, 由 Lemma 5.4.2 知 $\mu_{T^*}(x) = \overline{\mu_T}(x) = (x-a)^2 + b^2$, 故知

$$T^{\circ 2} = 2aT - (a^2+b^2)\mathrm{id}_V, \quad (T^*)^{\circ 2} = 2aT^* - (a^2+b^2)\mathrm{id}_V.$$

相減得

$$T^{\circ 2} - (T^*)^{\circ 2} = 2a(T - T^*) = 4aT_2. \tag{5.7}$$

由等式 (5.6), (5.7) 得 $4T_1 \circ T_2 = 4aT_2$, 由 (1) 知 T_2^{-1} 存在, 故將前面等式兩邊與 T_2^{-1} 合成得 $T_1 = a\,\mathrm{id}_V$. 即 $T_1 - a\,\mathrm{id}_V = \mathbf{O}$, 證得 $\mu_{T|_1}(x) = x - a$.

(3) 由 (2) 知 $T_2 = T - a\,\mathrm{id}_V$, 故

$$T_2{}^{\circ 2} = (T - a\,\mathrm{id}_V)^{\circ 2} = T^{\circ 2} - 2aT + a^2\mathrm{id}_V = (2aT - (a^2+b^2)\mathrm{id}_V) - 2aT + a^2\mathrm{id}_V = -b^2\mathrm{id}_V,$$

即 $T_2{}^{\circ 2} + b^2\mathrm{id}_V = \mathbf{O}$. 由 $b \neq 0$, $x^2 + b^2$ 在 $\mathbb{R}[x]$ 中為 irreducible, 得證 $\mu_{T_2}(x) = x^2 + b^2$. □

在 Lemma 5.6.12 中, 因 T_2 為 skew-adjoint 且 $\mu_{T_2}(x) = x^2 + b^2$, 由 Lemma 5.6.9 知存在 V 的一組 ordered orthonormal basis β 使得 $[T_2]_\beta$ 為如 (5.4) 的 block diagonal matrix. 而 $T_1 = a\,\mathrm{id}_V$, 故若 $\dim(V) = n$, 則 $[T_1]_\beta = aI_n$. 故由 $[T]_\beta = [T_1 + T_2]_\beta = [T_1]_\beta + [T_2]_\beta$ 我們有以下之結果.

Corollary 5.6.13. 假設 V 為 *finite dimensional inner product space over* $\mathbb{R}$, $T : V \to V$ 為 *normal operator*. 若 $\mu_T(x) = (x-a)^2 + b^2$ 其中 $b \neq 0$ 則存在 V 的一組 *ordered orthonormal basis* β 使得 $[T]_\beta$ 為以下的 *block diagonal matrix*

$$[T]_\beta = \begin{pmatrix} M & & \mathbf{O} \\ & \ddots & \\ \mathbf{O} & & M \end{pmatrix}, \text{其中 } M = \begin{pmatrix} a & -b \\ b & a \end{pmatrix}. \tag{5.8}$$

Example 5.6.14. 假設 V 為 finite dimensional inner product space over $\mathbb{R}$ 且 $T : V \to V$ 為 normal operator. 若 $\chi_T(x) = (x^2 + 4x + 6)^2$, 則由 Proposition 5.5.6 知 $\mu_T(x) = x^2 + 4x + 6 = (x+2)^2 + 2$. 故由 Corollary 5.6.13 知存在 V 的一組 ordered orthonormal basis β 使得

$$[T]_\beta = \begin{pmatrix} -2 & -\sqrt{2} & 0 & 0 \\ \sqrt{2} & -2 & 0 & 0 \\ 0 & 0 & -2 & -\sqrt{2} \\ 0 & 0 & \sqrt{2} & -2 \end{pmatrix}.$$

對於一般的 over $\mathbb{R}$ 的 normal operator $T:V\to V$, 由 Proposition 5.5.6 知 $\mu_T(x)=p_1(x)\cdots p_k(x)q_1(x)\cdots q_l(x)$, 其中 $p_1(x),\ldots,p_k(x)\in\mathbb{R}[x]$ 為相異的 monic irreducible polynomial of degree 2, $q_1(x),\ldots,q_l(x)\in\mathbb{R}[x]$ 為相異的 monic polynomial of degree 1. 令 $W_i=\mathrm{Ker}(p_i(T))$, $E_j=\mathrm{Ker}(q_j(T))$, 則

$$V=W_1\boxplus\cdots\boxplus W_k\boxplus E_1\boxplus\cdots\boxplus E_l.$$

由於 $E_j=\mathrm{Ker}(q_j(T))$ 為 T 的一個 eigenspace, 任取 E_j 的一組 ordered orthonormal basis β'_j, 皆會使得 $[T|_{E_j}]_{\beta'_j}$ 為 diagonal matrix. 而因 $p_i(x)$ 為 $(x-a)^2+b^2$ 的形式且 $T|_{W_j}$ 仍為 normal, 又 $\mu_{T|_{W_i}}(x)=p_i(x)$, 由 Corollary 5.6.13 我們可以找到 W_i 的一組 orthonormal ordered basis β_i, 使得 $[T|_{W_i}]_{\beta_i}$ 為如 (5.8) 的 block diagonal matrix. 由於這些 $\beta_1,\ldots,\beta_k,\beta'_1,\ldots,\beta'_l$ 可以組成 V 的一組 ordered orthonormal basis, 我們有以下有關 over $\mathbb{R}$ 的 spectral theorem.

Theorem 5.6.15 (Spectral Theorem: real case)**.** 假設 V 為 *finite dimensional inner product space over* $\mathbb{R}$. 則 $T:V\to V$ 為 *normal operator* 若且唯若存在 V 的一組 *ordered orthonormal basis* β 使得 $[T]_\beta$ 為以下的 *block diagonal matrix*

$$[T]_\beta=\begin{pmatrix} M_1 & & \mathbf{O}\\ & \ddots & \\ \mathbf{O} & & M_l\end{pmatrix}, \text{其中 } M_i=\begin{pmatrix} a_i & -b_i\\ b_i & a_i\end{pmatrix}, b_i\neq 0 \text{ 或 } M_i=\lambda_i. \tag{5.9}$$

Proof. 若 $T:V\to V$ 為 normal operator 由前面的說明我們知存在 V 的一組 ordered orthonormal basis β 使得 $[T]_\beta$ 為如 (5.9) 的 block diagonal matrix.

反之, 若存在 V 的一組 ordered orthonormal basis β 使得 $[T]_\beta$ 為如 (5.9) 的 block diagonal matrix, 則由 Proposition 5.3.10 知

$$[T^*]_\beta=[T]_\beta^*=\begin{pmatrix} M_1^* & & \mathbf{O}\\ & \ddots & \\ \mathbf{O} & & M_l^*\end{pmatrix},$$

其中 $M_i^*=\begin{pmatrix} a_i & b_i\\ -b_i & a_i\end{pmatrix}$ 或 $M_i^*=\lambda_i=M_i$. 由於 $M_iM_i^*=M_i^*M_i$, 可得 $[T]_\beta[T^*]_\beta=[T^*]_\beta[T]_\beta$, 得證 $T\circ T^*=T^*\circ T$. □

由 Theorem 5.6.15, 我們可得到有關於 normal matrix 的性質. 也就是說 $A\in M_n(\mathbb{R})$ 滿足 $A^{\mathrm{t}}A=AA^{\mathrm{t}}$ 若且唯若存在 orthogonal matrix $P\in M_n(\mathbb{R})$ 使得 $P^{\mathrm{t}}AP$ 為如 (5.9) 的 block diagonal matrix.

Question 5.29. 假設 $A\in M_n(\mathbb{R})$ 且 $A^{\mathrm{t}}A=AA^{\mathrm{t}}$. 若已知

$$\chi_A(x)=(x^2+1)(x^2+4x+6)^2(x+1)^2(x^2-2),$$

試寫下如 *(5.9)* 的 *block diagonal matrix* B 使得 $A\sim B$. 試寫下 A 的 *rational form.* 為何 A 的 *rational form* 不是 *normal?*

最後我們特別做以下的提醒. 一個 linear operator $T:V\to V$ 是否為 diagonalizable, 和 V 是否為 inner product space 無關. 但 T 是否為 unitary diagonalizable 就和 inner product 有關了. 也就是說 V 用不同的 inner product, 不會改變 T 是否為 diagonalizable 的性質, 但有可能改變其是否為 unitary diagonalizable 的性質. 我們有以下的例子.

Example 5.6.16. 在 Example 5.3.11 中我們考慮 $T:\mathbb{R}^3\to\mathbb{R}^3$ 定義為

$$T(x_1,x_2,x_3)=(4x_1+x_2+2x_3,4x_1+4x_2+4x_3,2x_1+x_2+4x_3).$$

以及 $\mathbb{R}^3$ 上兩個 inner products $\langle\,,\rangle$ 以及 $\langle\langle\,,\rangle\rangle$. 簡單的計算我們得 T 的 eigenvalues 為 2 和 8 且其對應的 eigenspaces 分別為 $E_2=\mathrm{Span}(\{(1,-2,0),(1,0,-1)\})$ 以及 $E_8=\mathrm{Span}(\{(1,2,1)\})$.

若考慮 $\mathbb{R}^3$ 為 standard inner product space, 此時 $T\neq T^*$ 故 T 不是 self-adjoint operator. 由 Theorem 5.6.5 知 T 在 standard inner product $\langle\,,\rangle$ 之下不是 unitary diagonalizable. 事實上 $\langle(1,-2,0),(1,2,1)\rangle=-3\neq 0$, 知 $E_2\not\subseteq E_8^{\perp}$. 但因 T 仍為 diagonalizable, 亦即 T 的 minimal polynomial $\mu_T(x)$ 可寫成一次 monic polynomial 的乘積 (事實上我們有 $\mu_T(x)=(x-2)(x-8)$). 依此我們可利用 Theorem 5.6.15 得知 T 在 standard inner product 之下不是 normal operator. 事實上由 Example 5.3.11 知此時

$$[T^*\circ T]_\varepsilon=[T^*]_\varepsilon[T]_\varepsilon=\begin{pmatrix}4&4&2\\1&4&1\\2&4&4\end{pmatrix}\begin{pmatrix}4&1&2\\4&4&4\\2&1&4\end{pmatrix}=\begin{pmatrix}36&22&32\\22&18&22\\32&22&36\end{pmatrix},$$

而

$$[T\circ T^*]_\varepsilon=[T]_\varepsilon[T^*]_\varepsilon=\begin{pmatrix}4&1&2\\4&4&4\\2&1&4\end{pmatrix}\begin{pmatrix}4&4&2\\1&4&1\\2&4&4\end{pmatrix}=\begin{pmatrix}21&28&17\\28&48&28\\17&28&21\end{pmatrix},$$

考慮 $\mathbb{R}^3$ 為以 $\langle\langle\,,\rangle\rangle$ 為 inner product 的 inner product space 時, 我們得 $T^*=T$, 亦即 T 為 self-adjoint operator, 故由 Theorem 5.6.5 知 T 在此 inner product 之下為 unitary diagonalizable. 事實上由 $\langle\langle(1,-2,0),(1,2,1)\rangle\rangle=0$ 以及 $\langle\langle(1,0,-1),(1,2,1)\rangle\rangle=0$ 得此時 $E_2\perp E_8$. 利用 Gram-Schmidt's process 找到 E_2 的一組 orthogonal basis, 其過程如下: 令 $\mathbf{v}_1=(1,-2,0),\mathbf{w}_2=(1,0,-1)$, 則取

$$\mathbf{v}_2=\mathbf{w}_2-\frac{\langle\langle\mathbf{w}_2,\mathbf{v}_1\rangle\rangle}{\langle\langle\mathbf{v}_1,\mathbf{v}_1\rangle\rangle}\mathbf{v}_1=(1,0,-1)-\frac{1}{2}(1,-2,0)=(\frac{1}{2},1,-1),$$

可得 $\langle\langle\mathbf{v}_1,\mathbf{v}_2\rangle\rangle=0$, 故 $\{(1,-2,0),(\frac{1}{2},1,-1),(1,2,1)\}$ 為 T 的 eigenvectors 且為 $\mathbb{R}^3$ 在 $\langle\langle\,,\rangle\rangle$ 之下的一組 orthogonal basis.