

### 6.5. Direct Sum

當  $V$  是一個 vector space 且  $W_1, W_2$  為  $V$  的 subspaces, 我們曾介紹過  $W_1 + W_2$  的概念. 事實上  $W_1 + W_2$  是由  $W_1, W_2$  中的向量所展成的向量空間. 換言之,  $W_1$  的 basis 和  $W_2$  的 basis 聯合起來可以展成  $W_1 + W_2$ . 不過它們聯合起來未必會是 linearly independent. 所以它們不一定是  $W_1 + W_2$  的 basis. 特別地當這兩組 basis 一起會是 linearly independent, 我們就說  $W_1 + W_2$  是  $W_1, W_2$  的 *direct sum* 且用  $W_1 \oplus W_2$  表示.

要檢查  $V$  的 subspace  $W$  是否為  $W_1, W_2$  的 direct sum  $W_1 \oplus W_2$ , 當然首先要符合  $W = W_1 + W_2$ . 這個部分關係到  $W$  和  $W_1, W_2$  也就是說我們要檢查  $W$  中的元素是否都可寫成  $\mathbf{w}_1 + \mathbf{w}_2$ , 其中  $\mathbf{w}_1 \in W_1, \mathbf{w}_2 \in W_2$  (這表示  $W \subseteq W_1 + W_2$ ). 另一方面, 要確認對任意  $\mathbf{w}_1 \in W_1, \mathbf{w}_2 \in W_2$  皆會滿足  $\mathbf{w}_1 + \mathbf{w}_2 \in W$  (這表示  $W \supseteq W_1 + W_2$ ). 接著要檢查的部分就和  $W$  無關了, 而是要確認  $W_1, W_2$  元素間有沒有線性關係. 簡單來說就是要檢查  $W_1 \cap W_2 = \{\mathbf{0}\}$ . 這是因為如果  $(\mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_r)$  為  $W_1$  的一組 ordered basis,  $(\mathbf{w}'_1, \dots, \mathbf{w}'_s)$  為  $W_2$  的一組 ordered basis. 若  $(\mathbf{w}_1, \dots, \mathbf{w}_r, \mathbf{w}'_1, \dots, \mathbf{w}'_s)$  為 linearly dependent, 這表示存在  $c_1, \dots, c_r, d_1, \dots, d_s \in \mathbb{F}$  不全為 0 使得  $c_1\mathbf{w}_1 + \dots + c_r\mathbf{w}_r + d_1\mathbf{w}'_1 + \dots + d_s\mathbf{w}'_s = \mathbf{0}$ , 亦即  $c_1\mathbf{w}_1 + \dots + c_r\mathbf{w}_r = -(d_1\mathbf{w}'_1 + \dots + d_s\mathbf{w}'_s) \in W_1 \cap W_2$ . 因此若  $W_1 \cap W_2 = \{\mathbf{0}\}$ , 可得  $c_1\mathbf{w}_1 + \dots + c_r\mathbf{w}_r = d_1\mathbf{w}'_1 + \dots + d_s\mathbf{w}'_s = \mathbf{0}$ . 再加上  $\{\mathbf{w}_1, \dots, \mathbf{w}_r\}$  為 linearly independent 以及  $\{\mathbf{w}'_1, \dots, \mathbf{w}'_s\}$  亦為 linearly independent, 可得  $c_1 = \dots = c_r = d_1 = \dots = d_s = 0$ . 此與當初假設  $c_1, \dots, c_r, d_1, \dots, d_s$  不全為 0 相矛盾, 故得  $(\mathbf{w}_1, \dots, \mathbf{w}_r, \mathbf{w}'_1, \dots, \mathbf{w}'_s)$  為 linearly independent. 反過來, 當  $(\mathbf{w}_1, \dots, \mathbf{w}_r, \mathbf{w}'_1, \dots, \mathbf{w}'_s)$  為 linearly independent, 現若  $W_1 \cap W_2 \neq \{\mathbf{0}\}$ , 表示存在  $\mathbf{v} \in W_1 \cap W_2$  且  $\mathbf{v} \neq \mathbf{0}$ . 然而因  $\mathbf{v} \in W_1$  且  $\{\mathbf{w}_1, \dots, \mathbf{w}_r\}$  為  $W_1$  的一組 basis, 故存在  $c_1, \dots, c_r \in \mathbb{F}$  使得  $\mathbf{v} = c_1\mathbf{w}_1 + \dots + c_r\mathbf{w}_r$ . 又因  $\mathbf{v} \neq \mathbf{0}$ , 所以  $c_1, \dots, c_r$  不全為 0. 同理可得存在  $d_1, \dots, d_s \in \mathbb{F}$  不全為 0 使得  $\mathbf{v} = d_1\mathbf{w}'_1 + \dots + d_s\mathbf{w}'_s$ . 然而  $c_1\mathbf{w}_1 + \dots + c_r\mathbf{w}_r = d_1\mathbf{w}'_1 + \dots + d_s\mathbf{w}'_s$  與  $(\mathbf{w}_1, \dots, \mathbf{w}_r, \mathbf{w}'_1, \dots, \mathbf{w}'_s)$  為 linearly independent 相矛盾故得證  $W_1 \cap W_2 = \{\mathbf{0}\}$ .

我們可以將兩個 subspaces 的 direct sum 推廣到更多個 subspaces 的 direct sum. 例如  $W_1, \dots, W_k$  為  $V$  的 subspaces, 我們可以考慮

$$W_1 + \dots + W_k = \{\mathbf{w}_1 + \dots + \mathbf{w}_k \mid \mathbf{w}_i \in W_i, \forall i = 1, \dots, k\}$$

這一個 subspace (稱為  $W_1, \dots, W_k$  的 *sum*). 和前面一樣的, 假設對所有的  $i = 1, \dots, k$ , 我們有  $\beta_i$  為  $W_i$  的 basis. 我們希望將這些  $\beta_1, \dots, \beta_k$  放在一起就會是  $W_1 + \dots + W_k$  的 basis. 當然和兩個 subspace 的情況一樣, 這一般不一定會成立. 問題發生在它們會不會是 linearly independent. 如果它們會是 linearly independent, 我們就稱  $W_1 + \dots + W_k$  是  $W_1, \dots, W_k$  的 *direct sum*, 用  $W_1 \oplus \dots \oplus W_k$  來表示. 因為  $\beta_1, \dots, \beta_k$  放在一起其向量的個數會是每個  $\beta_i$  中向量個數 (即  $\dim(W_i)$ ) 加在一起 (即  $\dim(W_1) + \dots + \dim(W_k)$ ). 所以若  $\beta_1, \dots, \beta_k$  放在一起會是  $W_1 + \dots + W_k$  的 basis, 則可得

$$\dim(W_1 + \dots + W_k) = \dim(W_1) + \dots + \dim(W_k).$$

這裡就用這個等式作為 direct sum 的定義.

**Definition 6.5.1.** 假設  $V$  為 vector space 且  $W_1, \dots, W_k$  為  $V$  的 subspaces. 若

$$\dim(W_1 + \dots + W_k) = \dim(W_1) + \dots + \dim(W_k),$$

則稱  $W_1 + \dots + W_k$  是  $W_1, \dots, W_k$  的 *direct sum*, 用  $W_1 \oplus \dots \oplus W_k$  來表示.

要檢查兩個以上的 subspaces 是否可形成 direct sum 並不像檢查兩個 subspaces 的情況只要檢查它們的交集即可. 例如檢查 subspaces  $W_1, W_2, W_3$  是否可形成 direct sum 不能僅檢查是否  $W_1 \cap W_2 \cap W_3 = \{\mathbf{0}\}$ . 因為由  $W_1 \cap W_2 \cap W_3 = \{\mathbf{0}\}$  未必可得  $W_1 \cap W_2 = \{\mathbf{0}\}$ , 也就是說  $W_1, W_2$  就可能無法形成 direct sum 就更遑論  $W_1, W_2, W_3$  了. 或許大家認為檢查  $W_1 \cap W_2 = \{\mathbf{0}\}$ ,  $W_1 \cap W_3 = \{\mathbf{0}\}$  和  $W_3 \cap W_2 = \{\mathbf{0}\}$  就可以了, 其實不然. 例如  $V = \mathbb{R}^2$  的情形, 若令  $\mathbf{e}_1 = (1, 0)$ ,  $\mathbf{e}_2 = (0, 1)$ , 考慮  $W_1 = \text{Span}(\mathbf{e}_1)$ ,  $W_2 = \text{Span}(\mathbf{e}_2)$  以及  $W_3 = \text{Span}(\mathbf{e}_1 + \mathbf{e}_2)$ . 很容易檢查當  $i \neq j$  時  $W_i \cap W_j = \{\mathbf{0}\}$ . 但這僅表示  $W_i, W_j$  可形成 direct sum 即  $\mathbb{R}^2 = W_i + W_j = W_i \oplus W_j$ , 但無法表示  $W_1, W_2, W_3$  可形成 direct sum. 事實上  $\mathbb{R}^2 = W_1 + W_2 + W_3$  但  $\dim(\mathbb{R}^2) = 2 \neq \dim(W_1) + \dim(W_2) + \dim(W_3) = 1 + 1 + 1 = 3$ . 其實在一般的情況, 若令  $W = W_1 + W_2$ , 則由  $W_1 \cap W_2 = \{\mathbf{0}\}$  可知  $W = W_1 \oplus W_2$ , 但若要達到  $W_1 + W_2 + W_3 = W \oplus W_3$ , 則需要檢查  $W \cap W_3 = (W_1 + W_2) \cap W_3 = \{\mathbf{0}\}$ , 這和單純檢查  $W_1 \cap W_3 = \{\mathbf{0}\}$  和  $W_2 \cap W_3 = \{\mathbf{0}\}$  (即  $(W_1 \cup W_2) \cap W_3 = \{\mathbf{0}\}$ ) 是不一樣的. 由此可知檢查更多 subspaces 是否可形成 direct sum 就更加複雜了, 底下的定理給我們一個較有效的方法處理.

**Proposition 6.5.2.** 假設  $V$  為 vector space 且  $W_1, \dots, W_k$  為  $V$  的 subspaces. 則  $W_1 + \dots + W_k = W_1 \oplus \dots \oplus W_k$  若且唯若任取  $W_1, \dots, W_k$  中的非零向量皆為 *linearly independent*.

**Proof.** 首先假設  $W_1 + \dots + W_k = W_1 \oplus \dots \oplus W_k$ . 若存在  $\mathbf{w}_1, \dots, \mathbf{w}_k$  為 linearly dependent 其中  $\mathbf{w}_i \in W_i$  且  $\mathbf{w}_i \neq \mathbf{0}, \forall i \in \{1, \dots, k\}$ , 這表示由  $W_1, \dots, W_k$  的 basis 所組成的集合是 linearly dependent. 這和  $\dim(W_1 + \dots + W_k) = \dim(W_1) + \dots + \dim(W_k)$  之假設相矛盾, 故知  $\mathbf{w}_1, \dots, \mathbf{w}_k$  為 linearly independent.

反之, 要證明  $W_1 + \dots + W_k = W_1 \oplus \dots \oplus W_k$ , 我們僅要證明當  $\mathbf{w}_{i,1}, \dots, \mathbf{w}_{i,m_i}$  為  $W_i$  的一組 basis. 將  $W_1, \dots, W_k$  的這  $k$  組 basis 收集在一起後它們  $\mathbf{w}_{1,1}, \dots, \mathbf{w}_{1,m_1}, \dots, \mathbf{w}_{k,1}, \dots, \mathbf{w}_{k,m_k}$  仍是 linearly independent. 照慣例, 我們先假設  $\mathbf{w}_{1,1}, \dots, \mathbf{w}_{1,m_1}, \dots, \mathbf{w}_{k,1}, \dots, \mathbf{w}_{k,m_k}$  是 linearly dependent. 亦即存在不全為 0 的  $c_{1,1}, \dots, c_{1,m_1}, \dots, c_{k,1}, \dots, c_{k,m_k} \in \mathbb{F}$  使得

$$c_{1,1}\mathbf{w}_{1,1} + \dots + c_{1,m_1}\mathbf{w}_{1,m_1} + \dots + c_{k,1}\mathbf{w}_{k,1} + \dots + c_{k,m_k}\mathbf{w}_{k,m_k} = \mathbf{0}.$$

此時對任意  $i \in \{1, \dots, k\}$ , 我們令  $\mathbf{w}_i = c_{i,1}\mathbf{w}_{i,1} + \dots + c_{i,m_i}\mathbf{w}_{i,m_i}$ . 因此由於  $\mathbf{w}_{i,1}, \dots, \mathbf{w}_{i,m_i}$  為 linearly independent, 如果  $c_{i,1}, \dots, c_{i,m_i}$  不全為 0, 可得  $\mathbf{w}_i \neq \mathbf{0}$ . 但由於  $\mathbf{w}_i \in W_i$ . 這表示  $W_1, \dots, W_i$  存在一些  $\mathbf{w}_i \in W_i$  不為  $\mathbf{0}$  但會是 linearly dependent. 此和我們的假設不符, 故得證本定理.  $\square$

在 Proposition 6.4.1 我們知道對於一個矩陣 (或 linear operator) 其 eigenvalue 相異的 eigenspaces 之間任選非零向量 (即 eigenvector) 都會是 linearly independent, 因此由 Proposition 6.5.2 我們知 eigenvalue 相異的 eigenspaces 可形成 direct sum. 現假設  $\lambda_1, \dots, \lambda_k \in \mathbb{F}$  為  $A \in M_{n \times n}$  所有的 eigenvalues, 則我們可以考慮這些 eigenvalue 所對應的

eigenspace 的 direct sum  $E_A(\lambda_1) \oplus \cdots \oplus E_A(\lambda_k)$ . 因為任意  $A$  的 eigenvector 都會在某個  $E_A(\lambda_i)$  中, 所以  $A$  的所有 eigenvectors 所展成的向量空間會包含在  $E_A(\lambda_1) \oplus \cdots \oplus E_A(\lambda_k)$ . 另一方面由於  $E_A(\lambda_1) \oplus \cdots \oplus E_A(\lambda_k)$  中的非零向量都是  $A$  的一些 eigenvectors 之和, 所以我們也知  $E_A(\lambda_1) \oplus \cdots \oplus E_A(\lambda_k)$  會包含於  $A$  的所有 eigenvectors 所展成的向量空間. 因此我們有以下的結果.

**Corollary 6.5.3.** 假設  $A \in M_{n \times n}$  且  $\lambda_1, \dots, \lambda_k \in \mathbb{F}$  為  $A$  所有的 eigenvalues. 則  $A$  的所有 eigenvectors 所展成的向量空間等於  $E_A(\lambda_1) \oplus \cdots \oplus E_A(\lambda_k)$ . 特別的, 若  $A$  為 diagonalizable, 則  $\mathbb{F}^n = E_A(\lambda_1) \oplus \cdots \oplus E_A(\lambda_k)$ .

**Proof.** 我們僅剩證明  $A$  為 diagonalizable 的情況. 此時依定義  $\mathbb{F}^n$  中可找到一組由  $A$  的 eigenvectors 所形成的 basis. 換言之  $\mathbb{F}^n$  中的非零向量都可寫成  $A$  的一些 eigenvectors 之和, 亦即  $\mathbb{F}^n$  就是  $A$  的所有 eigenvectors 所展成的向量空間. 故得證  $\mathbb{F}^n = E_A(\lambda_1) \oplus \cdots \oplus E_A(\lambda_k)$ .  $\square$

**Question 6.7.** 設  $V$  為 vector space over  $\mathbb{F}$  且  $T: V \rightarrow V$  為 diagonalizable linear operator. 令  $\lambda_1, \dots, \lambda_k \in \mathbb{F}$  為  $T$  所有的 eigenvalues 且  $E_T(\lambda_i) = \{\mathbf{v} \in V \mid T(\mathbf{v}) = \lambda_i \mathbf{v}\}$  為  $\lambda_i$  所對應的 eigenspace. 試證明  $V = E_T(\lambda_1) \oplus \cdots \oplus E_T(\lambda_k)$ .

當  $\mathbf{v}_1, \dots, \mathbf{v}_n$  是  $V$  的一組 basis 時, 由  $\mathbf{v}_1, \dots, \mathbf{v}_n$  可展成  $V$ , 知  $V$  中的元素都可以寫成  $\mathbf{v}_1, \dots, \mathbf{v}_n$  的線性組合. 又由  $\mathbf{v}_1, \dots, \mathbf{v}_n$  是 linearly independent 知  $V$  中的元素寫成  $\mathbf{v}_1, \dots, \mathbf{v}_n$  的線性組合的寫法是唯一的. 而當  $V$  可以寫成 subspaces  $W_1, \dots, W_k$  的 direct sum 時, 由  $V = W_1 + \cdots + W_k$ , 我們可得  $V$  中的元素都可以寫成  $W_1, \dots, W_k$  中的元素之和. 而  $W_1, \dots, W_k$  中的非零向量是 linearly independent 這個性質也可推導出  $V$  中的元素寫成  $W_1, \dots, W_k$  中的元素之和的寫法也是唯一的. 簡單來說 direct sum 的概念可以說是將 basis 的概念由 vectors 推廣到 subspace. 我們有以下之定理.

**Proposition 6.5.4.** 假設  $V$  為 vector space 且  $W_1, \dots, W_k$  為  $V$  的 subspaces. 則  $V = W_1 \oplus \cdots \oplus W_k$  若且唯若對任意  $\mathbf{v} \in V$  皆存在唯一的  $\mathbf{w}_i \in W_i, i = 1, \dots, k$  使得  $\mathbf{v} = \mathbf{w}_1 + \cdots + \mathbf{w}_k$ .

**Proof.** 首先假設  $V = W_1 \oplus \cdots \oplus W_k$ . 由此假設知  $V = W_1 + \cdots + W_k$ , 因此對任意  $\mathbf{v} \in V$  皆存在  $\mathbf{w}_i \in W_i$  使得  $\mathbf{v} = \mathbf{w}_1 + \cdots + \mathbf{w}_k$ . 現證明唯一性, 即假設存在另一組  $\mathbf{w}'_i \in W_i$  使得  $\mathbf{v} = \mathbf{w}'_1 + \cdots + \mathbf{w}'_k$  我們要推得矛盾. 此時可得  $(\mathbf{w}_1 - \mathbf{w}'_1) + \cdots + (\mathbf{w}_k - \mathbf{w}'_k) = \mathbf{0}$ . 然而  $\mathbf{w}_i - \mathbf{w}'_i \in W_i$  且依假設存在  $j \in \{1, \dots, k\}$  使得  $\mathbf{w}_j - \mathbf{w}'_j \neq \mathbf{0}$ , 這會造成與 Proposition 6.5.2 所述  $W_1, \dots, W_k$  中的非零向量會 linearly independent 相矛盾. 因此得證唯一性.

反之, 假設對任意  $\mathbf{v} \in V$  皆存在唯一的  $\mathbf{w}_i \in W_i, i = 1, \dots, k$  使得  $\mathbf{v} = \mathbf{w}_1 + \cdots + \mathbf{w}_k$ . 則由存在性我們知  $V = W_1 + \cdots + W_k$ , 故現僅要證明  $W_1, \dots, W_k$  可形成 direct sum. 利用 Proposition 6.5.2, 我們要證明  $W_1, \dots, W_k$  中的非零向量會是 linearly independent. 我們仍然用反證法, 假設存在一組  $W_1, \dots, W_k$  中的非零向量是 linearly dependent. 這表示這組非零向量中存在一個  $\mathbf{w}_j \in W_j$  會是其他  $W_j$  以外的向量的線性組合. 換言之  $\mathbf{w}_j$  會有兩種表示法 (一種是  $\mathbf{w}_j$  另一種是  $W_j$  以外的向量之和), 此語表法唯一之假設相違, 故得證  $W_1, \dots, W_k$  可形成 direct sum.  $\square$

當一個 linear operator  $T:V \rightarrow V$  是 diagonalizable 時, 在 Question 6.7 我們知道  $V$  可以寫成  $T$  的 eigenspace 的 direct sum. 因此當給定  $\mathbf{v} \in V$ , 由 Proposition 6.5.4 我們可以將  $\mathbf{v}$  寫成  $\mathbf{v}_1 + \cdots + \mathbf{v}_k$ , 其中  $\mathbf{v}_i \in E_T(\lambda_i)$ . 此時

$$T(\mathbf{v}) = T(\mathbf{v}_1 + \cdots + \mathbf{v}_k) = T(\mathbf{v}_1) + \cdots + T(\mathbf{v}_k) = \lambda_1 \mathbf{v}_1 + \cdots + \lambda_k \mathbf{v}_k.$$

由於  $T:V \rightarrow V$  是 linear operator, 其定義域等於對應域, 因此我們可以考慮  $T$  的合成. 通常我們會用  $T^2$  來表示  $T$  和  $T$  的合成函數, 即  $T^2 = T \circ T$ . 也就是說  $T^2(\mathbf{v}) = T(T(\mathbf{v}))$ . 因此套用上面的式子, 再利用  $T$  是 linear 的性質, 可得  $T^2(\mathbf{v}) = \lambda_1^2 \mathbf{v}_1 + \cdots + \lambda_k^2 \mathbf{v}_k$ . 同樣的, 對於  $m \in \mathbb{N}$ , 我們令  $T^m$  表示由  $m$  個  $T$  合成起來的 linear operator, 所以我們可得

$$T^m(\mathbf{v}) = \lambda_1^m \mathbf{v}_1 + \cdots + \lambda_k^m \mathbf{v}_k, \forall m \in \mathbb{N}. \quad (6.6)$$

回顧過去, 我們提及當  $T_1, T_2$  是從  $V$  到  $W$  的 linear transformation, 其中  $V, W$  為 over  $\mathbb{F}$  的 vector spaces, 則我們可定義  $T_1, T_2$  的線性組合使其仍為  $V$  到  $W$  的 linear transformation. 也就是說, 若  $c_1, c_2 \in \mathbb{F}$ , 對於任意  $\mathbf{v} \in V$ , 我們定義  $(c_1 T_1 + c_2 T_2)(\mathbf{v}) = c_1 T_1(\mathbf{v}) + c_2 T_2(\mathbf{v})$ . 當  $T:V \rightarrow V$  是 linear operator, 我們知  $T^2, T^3, \dots, T^m$  以及  $\text{id}_V$  皆為定義在  $V$  上的 linear operators. 所以我們也可以考慮它們的線性組合, 亦即給定  $c_m, c_{m-1}, \dots, c_1, c_0 \in \mathbb{F}$ , 我們考慮  $c_m T^m + c_{m-1} T^{m-1} + \cdots + c_1 T + c_0 \text{id}_V$  這個 linear operator. 由於這個符號有點複雜, 我們一般會先寫下一個係數在  $\mathbb{F}$  的多項式  $f(x) = c_m x^m + c_{m-1} x^{m-1} + \cdots + c_1 x + c_0$ , 然後將上面提到的 linear operator 寫成

$$f(T) = c_m T^m + c_{m-1} T^{m-1} + \cdots + c_1 T + c_0 \text{id}_V.$$

若  $\mathbf{v} = \mathbf{v}_1 + \cdots + \mathbf{v}_k$ , 其中  $\mathbf{v}_i \in E_T(\lambda_i)$  (即  $T(\mathbf{v}_i) = \lambda_i \mathbf{v}_i$ ), 則  $f(T)(\mathbf{v}) = c_m T^m(\mathbf{v}) + \cdots + c_1 T(\mathbf{v}) + c_0 \text{id}_V(\mathbf{v})$ . 利用式子 6.6 以及  $c_0 \text{id}_V(\mathbf{v}) = c_0 \mathbf{v} = c_0(\mathbf{v}_1 + \cdots + \mathbf{v}_k)$ , 將每個  $\mathbf{v}_i$  的係數合併可得

$$f(T)\mathbf{v} = (c_m \lambda_1^m + \cdots + c_1 \lambda_1 + c_0) \mathbf{v}_1 + \cdots + (c_m \lambda_k^m + \cdots + c_1 \lambda_k + c_0) \mathbf{v}_k = f(\lambda_1) \mathbf{v}_1 + \cdots + f(\lambda_k) \mathbf{v}_k.$$

這個結果在  $f(x)$  是  $T$  的 characteristic polynomial 時特別有趣. 我們有以下的結果.

**Proposition 6.5.5.** 假設  $V$  為 vector space 且  $T:V \rightarrow V$  為 diagonalizable linear operator. 考慮  $T$  的 characteristic polynomial  $p_T(x)$ , 我們有  $p_T(T):V \rightarrow V$  為 zero operator, 亦即對任意  $\mathbf{v} \in V$ ,  $p_T(T)(\mathbf{v}) = \mathbf{0}$ .

**Proof.** 因為  $T:V \rightarrow V$  為 diagonalizable, 假設  $\lambda_1, \dots, \lambda_k \in \mathbb{F}$  為  $T$  的所有 eigenvalue, 對任意  $\mathbf{v} \in V$ , 我們知存在  $\mathbf{v}_i \in E_T(\lambda_i)$  使得  $\mathbf{v} = \mathbf{v}_1 + \cdots + \mathbf{v}_k$ . 由前面所提結果我們知  $p_T(T)(\mathbf{v}) = p_T(\lambda_1) \mathbf{v}_1 + \cdots + p_T(\lambda_k) \mathbf{v}_k$ . 然而每個  $T$  的 eigenvalue 都會是  $T$  的 characteristic polynomial  $p_T(x)$  的一個根, 因此  $p_T(\lambda_i) = 0, \forall i = 1, \dots, k$ . 得證  $p_T(T)(\mathbf{v}) = 0 \mathbf{v}_1 + \cdots + 0 \mathbf{v}_k = \mathbf{0}$ . 因為  $p_T(T)$  這個 linear operator 將所有  $\mathbf{v} \in V$  皆映射到  $\mathbf{0}$ , 故其為 zero operator.  $\square$

雖然 Proposition 6.5.5 我們假設  $T$  是 diagonalizable linear operator, 事實上它對一般的 linear operator 也會成立, 這就是所謂的 Cayley-Hamilton Theorem. 我們留待下一節再探討這更一般的情形. 目前我們先將剛才探討 diagonalizable linear operator 的性質轉換成 diagonalizable matrix 的情況討論.

假設  $A \in M_{n \times n}(\mathbb{F})$  為 diagonalizable, 我們知存在 invertible matrix  $Q$  使得  $Q^{-1}AQ$  為 diagonal matrix  $D$ . 換言之, 我們可以將  $A$  寫成  $A = QDQ^{-1}$ . 也因此我們可得

$$A^2 = (QDQ^{-1})(QDQ^{-1}) = QD^2Q^{-1}.$$

同理對任意  $m \in \mathbb{N}$ , 我們有  $A^m = QD^mQ^{-1}$ . 寫成這樣有什麼好處呢? 因為  $D$  為對角矩陣  $\begin{bmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{bmatrix}$ , 我們很容易算出  $D^m$ , 即  $\begin{bmatrix} \lambda_1^m & & 0 \\ & \ddots & \\ 0 & & \lambda_n^m \end{bmatrix}$ . 因此只要知道  $Q$  和  $Q^{-1}$ , 我們就可以很輕易算出  $A^m$  (即  $QD^mQ^{-1}$ ), 而不必真正將  $A$  乘到  $m$  次方了.

**Example 6.5.6.** 考慮實矩陣  $B = \begin{bmatrix} -1 & 4 & 2 \\ -1 & 3 & 1 \\ -1 & 2 & 2 \end{bmatrix}$ . 在 Example 6.4.4 我們算出  $Q^{-1}BQ = D$ ,

其中  $Q = \begin{bmatrix} 2 & 1 & 2 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{bmatrix}$  以及  $D = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{bmatrix}$ . 由於  $Q^{-1} = \begin{bmatrix} 1 & -1 & -1 \\ 1 & -2 & 0 \\ -1 & 2 & 1 \end{bmatrix}$ , 我們得

$$B^5 = QD^5Q^{-1} = \begin{bmatrix} 2 & 1 & 2 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 32 \end{bmatrix} \begin{bmatrix} 1 & -1 & -1 \\ 1 & -2 & 0 \\ -1 & 2 & 1 \end{bmatrix} = \begin{bmatrix} -61 & 124 & 62 \\ -31 & 63 & 31 \\ -31 & 62 & 32 \end{bmatrix}.$$

對於  $A \in M_{n \times n}(\mathbb{F})$ , 給定一係數在  $\mathbb{F}$  的多項式  $f(x) = c_mx^m + c_{m-1}x^{m-1} + \cdots + c_1x + c_0$ , 我們可定義  $f(A) = c_mA^m + c_{m-1}A^{m-1} + \cdots + c_1A + c_0I_n$ . 也就是在  $c_jx^j$  處用  $c_jA^j$  取代, 不過注意常數項  $c_0$  需用  $c_0I_n$  取代. 如此一來  $f(A)$  會是一個  $n \times n$  matrix. 現若  $A = QDQ^{-1}$ , 則依剛才的計算我們有

$$f(A) = c_mQD^mQ^{-1} + c_{m-1}QD^{m-1}Q^{-1} + \cdots + c_1QDQ^{-1} + c_0I_n.$$

在利用矩陣加法與乘法的分配律可得

$$f(A) = Q(c_mD^m + c_{m-1}D^{m-1} + \cdots + c_1D + c_0I_n)Q^{-1} = Qf(D)Q^{-1}.$$

然而  $D = \begin{bmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{bmatrix}$ , 為在對角線  $(i, i)$ -th entry 為  $\lambda_i$  的對角矩陣. 我們很快推得  $f(D) = \begin{bmatrix} f(\lambda_1) & & 0 \\ & \ddots & \\ 0 & & f(\lambda_n) \end{bmatrix}$ , 即在對角線  $(i, i)$ -th entry 為  $f(\lambda_i)$  的對角矩陣. 特別地, 當

$f(x)$  為  $A$  的 characteristic polynomial  $p_A(x)$ , 由於  $A$  的 eigenvalue  $\lambda_i$  皆滿足  $p_A(\lambda_i) = 0$ , 因此我們可得  $p_A(D) = \mathbf{0}$  (即零矩陣). 因而  $p_A(A) = Q(p_A(D))Q^{-1} = \mathbf{0}$ . 我們推得以下的定理.

**Proposition 6.5.7.** 假設  $A \in M_{n \times n}(\mathbb{F})$  為 diagonalizable 且  $p_A(x)$  為  $A$  的 characteristic polynomial. 則  $p_A(A) = \mathbf{0}$ .

同樣的對於矩陣也有所謂的 Cayley-Hamilton Theorem, 亦即對任意  $A \in M_{n \times n}(\mathbb{F})$ , 若  $p_A(x)$  為  $A$  的 characteristic polynomial. 則  $p_A(A) = \mathbf{0}$  (不需  $A$  為 diagonalizable 的假設). 我們留待下一節再探討.

## 6.6. Cayley-Hamilton Theorem

在這節中我們將介紹 Cayley-Hamilton Theorem. 首先我們先介紹 linear operator 的 invariant subspace, 再利用 invariant subspace 的概念證明 linear operator 的 Cayley-Hamilton Theorem, 再因此推得矩陣的 Cayley-Hamilton Theorem.

在探討函數的理論時, 通常當定義域很大時, 我們可以透過所謂的 restriction 將函數限制在較小的範圍來了解該函數. 給定一個函數  $f: X \rightarrow Y$ , 以及  $X$  中的子集合  $S$ , 所謂  $f$  的 restriction on  $S$ , 用  $f|_S$  表示, 就是將  $f$  的定義域縮小到  $S$ , 其他對於  $f$  的映射方式都沒有改變. 也就是說  $f|_S$  是一個定義域為  $S$  的函數  $f|_S: S \rightarrow Y$ , 且對於任意  $s \in S$ ,  $f|_S(s) = f(s)$ , 不過若  $x \in X$  但  $x \notin S$ , 則  $f|_S(x)$  是無定義的. 現若  $T: V \rightarrow V$  是 linear operator,  $W$  為  $V$  的 subspace, 則  $T|_W$  依然會是一個 linear transformation (只是定義域在  $W$  上). 不過  $T|_W$  未必會是一個 linear operator, 因為  $T$  未必會將  $W$  中的元素映射到  $W$ . 如此一來, 我們就不能將過去探討 linear operator 的理論運用在  $T|_W$  上了. 為了達到  $T|_W$  仍為 linear operator 的目的, 我們必須選有特殊性質的  $W$  (即  $T$  會將  $W$  的元素映射到  $W$ ), 這樣就能套用 linear operator 的理論了. 因此我們有以下的定義.

**Definition 6.6.1.** 假設  $V$  是一個 vector space over  $\mathbb{F}$  且  $T: V \rightarrow V$  是 linear operator. 若  $W$  是  $V$  的 subspace 且滿足  $T(W) \subseteq W$  (即  $T(\mathbf{w}) \in W, \forall \mathbf{w} \in W$ ), 則稱  $W$  為一個  $T$ -invariant subspace.

要注意當  $T: V \rightarrow V$  是 linear operator,, Definition 6.6.1, 告訴我們  $W$  是  $T$ -invariant, 表示  $T(W) \subseteq W$ , 並不是說  $T(W) = W$ , 也不是說  $T(\mathbf{w}) = \mathbf{w}, \forall \mathbf{w} \in W$ . 請大家不要誤解. 也就是說要檢查  $V$  的 subspace  $W$  是否為  $T$ -invariant subspace, 我們僅要檢查是否所有  $W$  的元素  $\mathbf{w}$  經由  $T$  的映射 (即  $T(\mathbf{w})$ ) 依然在  $W$  中. 當然了, 因  $T$  為 linear operator, 對任意  $\mathbf{v} \in V$ , 皆有  $T(\mathbf{v}) \in V$ , 故  $V$  本身是  $T$ -invariant. 還有因為  $T$  是 linear transformation, 我們知道  $T(\mathbf{0}) = \mathbf{0}$ , 所以 zero space  $\{\mathbf{0}\}$  也是  $T$ -invariant. 另外若  $\lambda \in \mathbb{F}$  是  $T$  的 eigenvector, 則  $\lambda$  所對應的 eigenspace  $E_T(\lambda) = \{\mathbf{v} \in V \mid T(\mathbf{v}) = \lambda\mathbf{v}\}$  也會是  $T$ -invariant subspace. 這是因為若  $\mathbf{v} \in E_T(\lambda)$ , 則  $T(\mathbf{v}) = \lambda\mathbf{v}$ . 由於  $E_T(\lambda)$  是  $V$  的 subspace 且  $\mathbf{v} \in E_T(\lambda)$ , 自然有  $\lambda\mathbf{v} \in E_T(\lambda)$ , 亦即  $T(\mathbf{v}) \in E_T(\lambda)$ . 故  $E_T(\lambda)$  亦為  $T$ -invariant. 另外我們過去熟悉的  $T$  的 range  $R(T)$ , 也是  $T$ -invariant, 這是因為對任意  $\mathbf{v} \in V$ , 自然有  $T(\mathbf{v}) \in T(V) = R(T)$ . 當然當  $\mathbf{v} \in R(T)$ , 由於  $T: V \rightarrow V$  是 linear operator, 故  $R(T) \subseteq V$ , 因此我們依然有  $T(\mathbf{v}) \in R(T)$ .  $T$  的 null space  $N(T)$  也是  $T$ -invariant. 這是因為對任意  $\mathbf{v} \in N(T)$ , 由於  $T(\mathbf{v}) = \mathbf{0}$  且  $\mathbf{0} \in N(T)$  (別忘了  $T(\mathbf{0}) = \mathbf{0}$ ), 故  $T(\mathbf{v}) \in N(T)$ .

除了前面舉的幾個例子外, 還有哪些  $T$ -invariant subspace 呢? 前面提過考慮  $T$ -invariant subspace 就是想將  $T$  的定義域縮小. 所以給定  $\mathbf{v} \in V$ , 我們很想知道甚麼是包含  $\mathbf{v}$  最小的  $T$ -invariant subspace. 假設  $W$  是包含  $\mathbf{v}$  的  $T$ -invariant subspace. 當然了, 我們有  $\mathbf{v} \in W$ . 不過由  $W$  是  $T$ -invariant, 我們自然要有  $T(\mathbf{v}) \in W$  (因  $\mathbf{v} \in W$ ). 再由  $T(\mathbf{v}) \in W$  以及  $W$  是  $T$ -invariant, 我們有  $T(T(\mathbf{v})) = T^2(\mathbf{v}) \in W$ . 如此一直下去我們知  $T^m(\mathbf{v}) \in W, \forall m \in \mathbb{N}$ . 由此我們知, 若  $W$  是包含  $\mathbf{v}$  的  $T$ -invariant subspace, 則  $W$  必須包含  $\{\mathbf{v}, T(\mathbf{v}), T^2(\mathbf{v}), \dots, T^m(\mathbf{v}), \dots\}$

這個集合 (即  $\{T^i(\mathbf{v}) \mid i \in \mathbb{N}\}$ ) 中所有的元素. 不過  $W$  是 subspace, 所以也必須包含所有這些元素所 span 的 subspace, 所以我們有以下的定義.

**Definition 6.6.2.** 假設  $T: V \rightarrow V$  是 linear operator. 對任意  $\mathbf{v} \in V$ , 令

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

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

要注意若令  $S = \{\mathbf{v}, T(\mathbf{v}), T^2(\mathbf{v}), \dots, T^m(\mathbf{v}), \dots\}$ , 雖然  $S$  中可能有無窮多個元素, 不過依 span 的定義,  $C(T, \mathbf{v}) = \text{Span}(S)$  中的元素是  $S$  中有限多個元素的線性組合. 因此若  $\mathbf{w} \in C(T, \mathbf{v}) = \text{Span}(S)$ , 表示存在  $c_0, c_1, \dots, c_m \in \mathbb{F}$  使得  $\mathbf{w} = c_0\mathbf{v} + c_1T(\mathbf{v}) + \dots + c_mT^m(\mathbf{v})$  (其中可能有些  $c_i = 0$ ). 故得  $T(\mathbf{w}) = c_0T(\mathbf{v}) + c_1T^2(\mathbf{v}) + \dots + c_mT^{m+1}(\mathbf{v}) \in \text{Span}(S) = C(T, \mathbf{v})$ . 也因此得證  $C(T, \mathbf{v})$  是  $T$ -invariant subspace. 前面提過包含  $\mathbf{v}$  的  $T$ -invariant subspace 必包含  $S$ , 故我們得到下面的定理.

**Proposition 6.6.3.** 假設  $T: V \rightarrow V$  是 linear operator 且  $\mathbf{v} \in V$ . 則  $C(T, \mathbf{v})$  是包含  $\mathbf{v}$  最小的  $T$ -invariant subspace.

**Question 6.8.** 假設  $V$  是 vector space over  $\mathbb{F}$ ,  $T: V \rightarrow V$  是 linear operator 且  $\mathbf{v} \in V$ . 證明對任意  $\mathbf{w} \in C(T, \mathbf{v})$ , 皆存在係數在  $\mathbb{F}$  的多項式  $f(x)$  使得  $\mathbf{w} = f(T)(\mathbf{v})$ . 依此得  $C(T, \mathbf{v}) = \{f(T)(\mathbf{v}) \mid f(x) \in \mathbb{F}[x]\}$ .

當  $V$  是 finite dimensional vector space over  $\mathbb{F}$ ,  $C(T, \mathbf{v})$  為其 subspace, 故  $C(T, \mathbf{v})$  也是 finite dimensional. 如何知道  $C(T, \mathbf{v})$  的維度呢? 當然了, 若  $\mathbf{v} = \mathbf{0}$ , 則  $T^i(\mathbf{v}) = T^i(\mathbf{0}) = \mathbf{0}$ ,  $\forall i \in \mathbb{N}$ , 故此時  $C(T, \mathbf{v}) = \{\mathbf{0}\}$ , 即  $\dim(C(T, \mathbf{v})) = 0$ . 因此我們僅考慮  $\mathbf{v} \neq \mathbf{0}$  的情況. 首先我們考慮  $\mathbf{v}, T(\mathbf{v})$  是否為 linearly independent. 若  $\mathbf{v}, T(\mathbf{v})$  不是 independent, 由於  $\mathbf{v} \neq \mathbf{0}$ , 故知存在  $c \in \mathbb{F}$  使得  $T(\mathbf{v}) = c\mathbf{v}$ . 由此知  $T^i(\mathbf{v}) = c^i\mathbf{v} \in \text{Span}(\mathbf{v})$ ,  $\forall i \in \mathbb{N}$ . 因此得  $C(T, \mathbf{v}) = \text{Span}(\mathbf{v})$ , 即  $\dim(C(T, \mathbf{v})) = 1$ . 而若  $\mathbf{v}, T(\mathbf{v})$  為 independent, 則我們考慮  $\mathbf{v}, T(\mathbf{v}), T^2(\mathbf{v})$  是否為 independent. 若它們不是 independent, 則由  $\mathbf{v}, T(\mathbf{v})$  為 independent 知  $T^2(\mathbf{v}) \in \text{Span}(\mathbf{v}, T(\mathbf{v}))$  (Lemma 2.5.4). 因此存在  $c, d \in \mathbb{F}$  使得  $T^2(\mathbf{v}) = c\mathbf{v} + dT(\mathbf{v})$ . 此時

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

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