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 $c_i T^{\circ i} \circ f(T)^* = c_i \sum_{j=0}^n \overline{c_j} T^{\circ i} \circ (T^*)^{\circ j}$, 可得

$$f(T) \circ f(T)^* = \sum_{i,j} c_i \overline{c_j} T^{\circ i} \circ (T^*)^{\circ j}.$$

同理

$$f(T)^* \circ f(T) = \sum_{i,j} \overline{c_j} c_i (T^*)^{\circ j} \circ T^{\circ i}.$$

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 $f(T) \circ f(T)^* = f(T)^* \circ f(T)$. $\square$

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

$$\mathrm{Ker}(T) \cap \mathrm{Im}(T) = \mathrm{Im}(T)^\perp \cap \mathrm{Im}(T) = \{\mathbf{O}_V\}.$$

$\square$

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

$$\langle \mathbf{v}, \mathbf{w} \rangle = \langle g(T) \circ \mu_{\mathbf{w}}(T)(\mathbf{v}), \mathbf{w} \rangle = \langle \mathbf{v}, (g(T) \circ \mu_{\mathbf{w}}(T))^*(\mathbf{w}) \rangle = \langle \mathbf{v}, \mathbf{O}_V \rangle = 0.$$

$\square$

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 $V = W_1 \boxplus \cdots \boxplus W_k$. $\square$

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.