

線性代數五講——

第三講 線性變換

龔 昇 · 張德健

3.1. 線性變換的矩陣表示

在第一、第二講中，我們提到線性代數是研究線性空間（向量空間），模和其上線性變換以及與之相關的問題的數學學科。在第二講中討論了向量空間和其上的線性泛函及對偶空間，其上的雙線性型式、二次型式及度量向量空間，正交幾何和辛幾何的分類，還有大家十分熟悉的內積空間。在這一講中將討論向量空間上線性變換以及與之相關的共軛算子及伴隨算子。

設 \mathcal{V} 及 \mathcal{W} 分別是體 \mathbb{F} 上的 n 維與 m 維向量空間。在這一節中要證明每個 $T \in \mathcal{L}(\mathcal{V}, \mathcal{W})$ 與 $\mathcal{M}_{m,n}(\mathbb{F})$ 中的一個矩陣對應。這就是 T 在 $\mathcal{M}_{m,n}(\mathbb{F})$ 中的矩陣表示。不但如此，還要證明 $\mathcal{L}(\mathcal{V}, \mathcal{W})$, $\mathcal{M}_{m,n}(\mathbb{F})$ 是同構的，所以對 $\mathcal{L}(\mathcal{V}, \mathcal{W})$ 的討論就是對 $\mathcal{M}_{m,n}(\mathbb{F})$ 的討論。

若 $T \in \mathcal{L}(\mathcal{V}, \mathcal{W})$, $\mathcal{B} = \{\vec{b}_1, \dots, \vec{b}_n\}$ 與 $\mathcal{C} = \{\vec{c}_1, \dots, \vec{c}_m\}$ 分別是 \mathcal{V} 與 \mathcal{W} 的基底，任給 $\vec{v} \in \mathcal{V}$ ，則 \vec{v} 在基底 \mathcal{B} 下的座標為

$$[\vec{v}]_{\mathcal{B}} = [v_1, \dots, v_n]^T,$$

$T(\vec{v})$ 在基底 \mathcal{C} 下的座標為

$$[T(\vec{v})]_{\mathcal{C}} = [s_1, \dots, s_m]^T.$$

於是對應於 T ，在 \mathbb{F}^n 與 \mathbb{F}^m 之間有一線性變換

$$T_A : [\vec{v}]_{\mathcal{B}} \rightarrow [T(\vec{v})]_{\mathcal{C}}, \quad T_A \in \mathcal{L}(\mathbb{F}^n, \mathbb{F}^m),$$

即對應於 T ，有一 $m \times n$ 的矩陣 A ，使得

$$[T(\vec{v})]_{\mathcal{C}} = A [\vec{v}]_{\mathcal{B}}.$$

現在來決定 A 。記 $A = [A_1, \dots, A_n]$ ，這裡 A_j , $j = 1, \dots, n$ 為列向量。取 $\vec{v} = \vec{b}_j$, $j = 1, \dots, n$ ，則立即得到 $A_j = [T(\vec{b}_j)]_{\mathcal{C}}$ ，故

$$A = [[T(\vec{b}_1)]_{\mathcal{C}}, \dots, [T(\vec{b}_n)]_{\mathcal{C}}].$$

記 A 為 $[T]_{\mathcal{B},\mathcal{C}}$, 於是

$$[T(\vec{v})]_{\mathcal{C}} = [T]_{\mathcal{B},\mathcal{C}} [\vec{v}]_{\mathcal{B}}.$$

$[T]_{\mathcal{B},\mathcal{C}}$ 即為當 \mathcal{V} 的基底為 \mathcal{B} , \mathcal{W} 的基底為 \mathcal{C} 時, T 的矩陣表示。於是可以定義映射

$$\Phi : \mathcal{L}(\mathcal{V}, \mathcal{W}) \rightarrow \mathcal{M}_{m,n}(\mathbb{F})$$

為 $\Phi(T) = [T]_{\mathcal{B},\mathcal{C}}$ 。我們現在來證明這是一個同構映射。先證 Φ 為線性映射: 若 $\alpha, \beta \in \mathbb{F}$, $T, S \in \mathcal{L}(\mathcal{V}, \mathcal{W})$, 則對 $j = 1, \dots, n$, 有

$$[(\alpha T + \beta S)(\vec{b}_j)]_{\mathcal{C}} = [(\alpha T)(\vec{b}_j) + (\beta S)(\vec{b}_j)]_{\mathcal{C}} = \alpha [T(\vec{b}_j)]_{\mathcal{C}} + \beta [S(\vec{b}_j)]_{\mathcal{C}}.$$

故

$$\Phi(\alpha T + \beta S) = [\alpha T + \beta S]_{\mathcal{B},\mathcal{C}} = \alpha [T]_{\mathcal{B},\mathcal{C}} + \beta [S]_{\mathcal{B},\mathcal{C}} = \alpha \Phi(T) + \beta \Phi(S)$$

即 Φ 為線性映射。

其次我們來證 Φ 為映成: 若 A 為一個 $n \times n$ 矩陣, 且可寫成 $[\vec{A}_1, \dots, \vec{A}_n]$, 這裡 \vec{A}_j , $j = 1, \dots, n$ 為列向量, 定義 $T : \mathcal{V} \rightarrow \mathcal{W}$, 使得 $[T(\vec{b}_j)]_{\mathcal{C}} = \vec{A}_j$, $j = 1, \dots, n$, 這是可以做到的, 故 Φ 為映成。

最後我們來證 Φ 為一對一: 由於 $[T]_{\mathcal{B},\mathcal{C}} = 0$, 導出 $[T(\vec{b}_j)]_{\mathcal{C}} = 0$, $j = 1, \dots, n$, 又導出 $T(\vec{b}_j) = \vec{0}$, $j = 1, \dots, n$, 故 $\Phi = 0$ 。因此, Φ 是一個同構映射。我們得到下面的定理:

定理 3.1.1: $\mathcal{L}(\mathcal{V}, \mathcal{W}) \approx \mathcal{M}_{m,n}(\mathbb{F})$ 。

由 T 的矩陣表示還可以導出: 若 $S : \mathcal{U} \rightarrow \mathcal{V}$ 及 $T : \mathcal{V} \rightarrow \mathcal{W}$, 且 \mathcal{B}, \mathcal{C} 與 \mathcal{D} 分別為 \mathcal{U}, \mathcal{V} 及 \mathcal{W} 的基底, 則

$$[T \circ S]_{\mathcal{B},\mathcal{D}} = [T]_{\mathcal{C},\mathcal{D}} [S]_{\mathcal{B},\mathcal{C}}.$$

因此, $T \circ S$ 的矩陣表示為 T 與 S 的矩陣表示之乘積。

驗證如下: 由於當 $\vec{u} \in \mathcal{U}$, $\vec{v} \in \mathcal{V}$ 時, 有

$$[S(\vec{u})]_{\mathcal{C}} = [S]_{\mathcal{B},\mathcal{C}} [\vec{u}]_{\mathcal{B}},$$

及

$$[T(\vec{v})]_{\mathcal{D}} = [T]_{\mathcal{C},\mathcal{D}} [\vec{v}]_{\mathcal{C}},$$

故

$$[T]_{\mathcal{C},\mathcal{D}} [S]_{\mathcal{B},\mathcal{C}} [\vec{v}]_{\mathcal{C}} = [T]_{\mathcal{C},\mathcal{D}} [S(\vec{u})]_{\mathcal{C}} = [T(S(\vec{u}))]_{\mathcal{D}} = [TS]_{\mathcal{B},\mathcal{D}} [\vec{u}]_{\mathcal{B}}.$$

假設 $T \in \mathcal{L}(\mathcal{V}, \mathcal{W})$ 。我們現在來討論當 \mathcal{V} 與 \mathcal{W} 的基底變換時, T 的相應的矩陣之間的關係。若 \mathcal{B} 與 \mathcal{C} 分別是 \mathcal{V} 與 \mathcal{W} 的基底, \mathcal{B}' 與 \mathcal{C}' 也分別是 \mathcal{V} 與 \mathcal{W} 的基底, T 對基底 \mathcal{B} 與 \mathcal{C} 及 \mathcal{B}' 與 \mathcal{C}' 分別有矩陣表示 $[T]_{\mathcal{B}, \mathcal{C}}$ 及 $[T]_{\mathcal{B}', \mathcal{C}'}$, 於是有: 對任意 $\vec{v} \in \mathcal{V}$, 下式成立:

$$[T(\vec{v})]_{\mathcal{C}} = [T]_{\mathcal{B}, \mathcal{C}}[\vec{v}]_{\mathcal{B}}, \quad [T(\vec{v})]_{\mathcal{C}'} = [T]_{\mathcal{B}', \mathcal{C}'}[\vec{v}]_{\mathcal{B}'}. \quad (3.1.1)$$

在第二講中已討論過, 對任意 $\vec{v} \in \mathcal{V}$, 我們有:

$$[\vec{v}]_{\mathcal{B}'} = M_{\mathcal{B}, \mathcal{B}'}[\vec{v}]_{\mathcal{B}}, \quad [T(\vec{v})]_{\mathcal{C}'} = M_{\mathcal{C}, \mathcal{C}'}[T(\vec{v})]_{\mathcal{C}}.$$

將這兩等式代入 (3.1.1), 我們便得到

$$[T(\vec{v})]_{\mathcal{C}'} = M_{\mathcal{C}, \mathcal{C}'}[T(\vec{v})]_{\mathcal{C}} = M_{\mathcal{C}, \mathcal{C}'}[T]_{\mathcal{B}, \mathcal{C}}[\vec{v}]_{\mathcal{B}}.$$

而

$$[T(\vec{v})]_{\mathcal{C}'} = [T]_{\mathcal{B}', \mathcal{C}'}[\vec{v}]_{\mathcal{B}'} = [T]_{\mathcal{B}', \mathcal{C}'}M_{\mathcal{B}, \mathcal{B}'}[T(\vec{v})]_{\mathcal{B}}.$$

由於 \vec{v} 為 \mathcal{V} 中任意的向量, 故

$$[T]_{\mathcal{B}', \mathcal{C}'}M_{\mathcal{B}, \mathcal{B}'} = M_{\mathcal{C}, \mathcal{C}'}[T]_{\mathcal{B}, \mathcal{C}},$$

也就是

$$[T]_{\mathcal{B}', \mathcal{C}'} = M_{\mathcal{C}, \mathcal{C}'}[T]_{\mathcal{B}, \mathcal{C}}M_{\mathcal{B}, \mathcal{B}'}^{-1}.$$

換句話說, $[T]_{\mathcal{B}', \mathcal{C}'}$ 與 $[T]_{\mathcal{B}, \mathcal{C}}$ 是等價的。特別當 $\mathcal{V} = \mathcal{W}$ 及 $T \in \mathcal{L}(\mathcal{V})$, 且 $\mathcal{B} = \mathcal{C}$ 與 $\mathcal{B}' = \mathcal{C}'$ 時, $[T]_{\mathcal{B}, \mathcal{B}} = [T]_{\mathcal{B}'}$, $[T]_{\mathcal{B}', \mathcal{B}'} = [T]_{\mathcal{B}'}$, 於是有

$$[T]_{\mathcal{B}'} = M_{\mathcal{B}, \mathcal{B}'}[T]_{\mathcal{B}}M_{\mathcal{B}, \mathcal{B}'}^{-1},$$

也就是說 $[T]_{\mathcal{B}}$ 與 $[T]_{\mathcal{B}'}$ 是相似的。在第五講中我們將討論 T 在相似意義下的分類。

3.2. 伴隨算子

由線性變換 $T \in \mathcal{L}(\mathcal{V}, \mathcal{W})$, 可以導出各種與之相關的線性變換來。在這一節中先來定義與討論在一般向量空間上的線性變換的伴隨算子。

若 $T \in \mathcal{L}(\mathcal{V}, \mathcal{W})$, 可定義 \mathcal{W} 的對偶空間 \mathcal{W}^* 到 \mathcal{V} 的對偶空間 \mathcal{V}^* 的的映射 $T^\times : \mathcal{W}^* \rightarrow \mathcal{V}^*$ 為

$$T^\times(f) = f \circ T = fT, \quad \forall f \in \mathcal{W}^*.$$

這是有意義的, 因為 $T : \mathcal{V} \rightarrow \mathcal{W}$, $f : \mathcal{W} \rightarrow \mathbb{F}$, 故 $fT : \mathcal{V} \rightarrow \mathbb{F}$, 於是屬於 \mathcal{V}^* , 即對任意 $\vec{v} \in \mathcal{V}$, 我們有:

$$T^\times(f)(\vec{v}) = f(T(\vec{v})).$$

T^\times 稱為 T 的伴隨算子。我們非常容易證明下面的命題。

命題3.2.1: (1). 對任意 $T, S \in \mathcal{L}(\mathcal{V}, \mathcal{W})$, 有

$$(T + S)^\times = T^\times + S^\times.$$

(2). 對任意 $\alpha \in \mathbb{F}$ 及 $T \in \mathcal{L}(\mathcal{V}, \mathcal{W})$, 有

$$(\alpha T)^\times = \alpha T^\times.$$

(3). 對任意 $T \in \mathcal{L}(\mathcal{V}, \mathcal{W})$ 及 $S \in \mathcal{L}(\mathcal{W}, \mathcal{U})$, 有

$$(S \circ T)^\times = T^\times \circ S^\times.$$

(4). 對任意可逆之 $T \in \mathcal{L}(\mathcal{V})$, 有

$$(T^{-1})^\times = (T^\times)^{-1}.$$

證明:(1) 與 (2) 是顯然成立的。對於 $f \in \mathcal{U}^*$,

$$(S \circ T)^\times(f) = f S T = T^\times(f S) = T^\times(S^\times(f)) = (T^\times \circ S^\times)(f),$$

故得 (3)。由 (3), 我們知

$$T^\times(T^{-1})^\times = (T^{-1}T)^\times = I^\times = I,$$

這裡 I 為恆等映射。同樣 $(T^{-1})^\times T^\times = I$, 故得 (4)。命題因而證畢。

命題3.2.2: 若 \mathcal{V} 為有限維向量空間, $T \in \mathcal{L}(\mathcal{V}, \mathcal{W})$, 且假設 \mathcal{V}^{**} 與 \mathcal{V} 等同, \mathcal{W}^{**} 與 \mathcal{W} 等同, 則 $T^{\times\times} = T$ 。

證明: 由定義, $T^{\times\times} : \mathcal{V}^{**} \rightarrow \mathcal{W}^{**}$ 。對任意 $f \in \mathcal{W}^*$, 我們有

$$T^{\times\times}(\vec{v}^{**})(f) = \vec{v}^{**}T^\times(f) = \vec{v}^{**}(fT) = fT(\vec{v}) = T(\vec{v})^{**}(f),$$

這裡 \vec{v}^{**} 由2.2節中定義: \vec{v}^{**} 由 \vec{v} 而來, $\vec{v}^{**} \in \mathcal{V}^{**}$, 定義為 $\vec{v}^{**}(g) = g(\vec{v})$, 這裡 $g \in \mathcal{V}^*$ 。

由上式即得

$$T^{\times\times}(\vec{v}^{**}) = T(\vec{v})^{**}.$$

如果 \mathcal{V}^{**} 與 \mathcal{V} 等同, \mathcal{W}^{**} 與 \mathcal{W} 等同, 則上式即為

$$T^{\times\times}(\vec{v}) = T(\vec{v})$$

對所有 $\vec{v} \in \mathcal{V}$ 都成立, 故 $T^{\times\times} = T$, 命題因而證畢。

此外, 伴隨算子與 2.2 節中定義的零化子還有以下一些結果。

命題 3.2.3: 若 $T \in \mathcal{L}(\mathcal{V}, \mathcal{W})$, 則

- (1) $\ker(T^\times) = \text{Im}(T)^\circ$;
- (2) $\text{Im}(T^\times)^\circ = \ker(T)$;
- (3) 當 \mathcal{V} 與 \mathcal{W} 均為有限維向量空間時, $\text{Im}(T^\times) = \ker(T)^\circ$ 。

證明: 由定義, $T : \mathcal{V} \rightarrow \mathcal{W}$, $T^\times : \mathcal{W}^* \rightarrow \mathcal{V}^*$, 故

$$\begin{aligned} f \in \ker(T^\times) &\Leftrightarrow T^\times(f) = 0 = fT \\ &\Leftrightarrow f(T(\vec{v})) = 0, \quad \forall \vec{v} \in \mathcal{V} \\ &\Leftrightarrow f(\text{Im}(T)) = 0 \\ &\Leftrightarrow f \in \text{Im}(T)^\circ. \end{aligned}$$

這就證明了 (1)。由於

$$\begin{aligned} \vec{v} \in \ker(T) &\Leftrightarrow T(\vec{v}) = \vec{0} \\ &\Leftrightarrow f(T(\vec{v})) = 0, \quad \forall f \in \mathcal{W}^* \\ &\Leftrightarrow T^\times(f)(\vec{v}) = 0 \\ &\Leftrightarrow \vec{v}^{**}(T^\times(f)) = 0, \quad \forall f \in \mathcal{W}^* \\ &\Leftrightarrow \vec{v}^{**} \in \text{Im}(T^\times)^\circ. \end{aligned}$$

若 \mathcal{V}^{**} 與 \mathcal{V} 等同, 我們便證明了 (2)。

最後來證明 (3)。對所有 $\vec{v} \in \ker(T)$, $f \in \mathcal{W}^*$, 有

$$T^\times(f)(\vec{v}) = f(T(\vec{v})) = 0.$$

所以 $T^\times(f)(\ker(T)) = 0$, 即 $T^\times(f) \in \ker(T)^\circ$, 這裡對所有的 $f \in \mathcal{W}^*$ 都成立, 故

$$\text{Im}(T^\times) \subset \ker(T)^\circ.$$

若向量空間為有限維, 則由命題 2.2.4 以及上述 (2), 我們得到:

$$\text{Im}(T^\times) \approx \text{Im}(T^\times)^\circ \approx \ker(T^\times)^\circ.$$

因此, $\text{Im}(T^\times) = \ker(T)^\circ$ 。命題因而證畢。

由此還可得到如下命題。

命題 3.2.4: 若 $T \in \mathcal{L}(\mathcal{V}, \mathcal{W})$, \mathcal{V} 與 \mathcal{W} 均為有限維向量空間, 則 T 與 T^\times 的秩滿足 $\text{rank}(T) = \text{rank}(T^\times)$ 。

證明: 由命題 2.2.6 中的 (1) 知道

$$\ker(T)^\circ \approx (\ker(T)^c)^*,$$

這裡 $\ker(T)^c$ 為 $\ker(T)$ 在 \mathcal{V} 中的餘集。另一方面, 由命題 3.2.3 中的 (3), $\text{Im}(T^\times) = \ker(T)^\circ$, 故

$$\begin{aligned} \dim(\text{Im}(T^\times)) &= \dim(\ker(T)^\circ) = \dim[(\ker(T)^c)^*] \\ &= \dim(\ker(T)^c) = \dim(\text{Im}(T)), \end{aligned}$$

這是因為 $\ker(T)^c \approx \text{Im}(T)$ 。於是 $\text{rank}(T) = \text{rank}(T^\times)$, 命題因而證畢。

若 \mathcal{V} 與 \mathcal{W} 均為有限維向量空間, $T \in \mathcal{L}(\mathcal{V}, \mathcal{W})$, $T^\times \in \mathcal{L}(\mathcal{W}^*, \mathcal{V}^*)$, $\mathcal{B} = \{\vec{b}_1, \dots, \vec{b}_n\}$ 與 $\mathcal{C} = \{\vec{c}_1, \dots, \vec{c}_n\}$ 分別為 \mathcal{V} 與 \mathcal{W} 的基底, 而 $\mathcal{B}^* = \{\vec{b}_1^*, \dots, \vec{b}_n^*\}$ 與 $\mathcal{C}^* = \{\vec{c}_1^*, \dots, \vec{c}_n^*\}$ 分別為對偶基底, 於是 T 有矩陣表示 $[T]_{\mathcal{B}, \mathcal{C}}$, T^\times 有矩陣表示 $[T^\times]_{\mathcal{C}^*, \mathcal{B}^*}$, 這兩個矩陣之間關係如何?

已知

$$[T]_{\mathcal{B}, \mathcal{C}} = [[T(\vec{b}_1)]_{\mathcal{C}}, \dots, [T(\vec{b}_n)]_{\mathcal{C}}],$$

及

$$[T^\times]_{\mathcal{C}^*, \mathcal{B}^*} = [[T(\vec{c}_1^*)]_{\mathcal{B}^*}, \dots, [T(\vec{c}_n^*)]_{\mathcal{B}^*}],$$

由於 $T(\vec{b}_j) \in \mathcal{W}$, 故在基底 \mathcal{C} 下, 這可表示為

$$T(\vec{b}_j) = \beta_1^{(j)} \vec{c}_1 + \beta_2^{(j)} \vec{c}_2 + \dots + \beta_n^{(j)} \vec{c}_n,$$

即 $[T(\vec{b}_j)]_{\mathcal{C}} = [\beta_1^{(j)} \dots \beta_n^{(j)}]^T$, $j = 1, \dots, n$, 於是

$$[T]_{\mathcal{B}, \mathcal{C}} = \begin{bmatrix} \beta_1^{(1)} & \dots & \beta_1^{(n)} \\ \beta_2^{(1)} & \dots & \beta_2^{(n)} \\ \vdots & \ddots & \vdots \\ \beta_n^{(1)} & \dots & \beta_n^{(n)} \end{bmatrix}.$$

由於 $T^\times(\vec{c}_j^*) \in \mathcal{V}^*$, 故在基底 \mathcal{B}^* 下, 這可表示為

$$T^\times(\vec{c}_j^*) = \alpha_1^{(j)}\vec{b}_1^* + \alpha_2^{(j)}\vec{b}_2^* + \cdots + \alpha_n^{(j)}\vec{b}_n^*,$$

即 $[T^\times(\vec{c}_j^*)]_{\mathcal{B}^*} = [\alpha_1^{(j)} \cdots \alpha_n^{(j)}]^T$, $j = 1, \dots, n$, 於是

$$[T^\times]_{\mathcal{C}^*, \mathcal{B}^*} = \begin{bmatrix} \alpha_1^{(1)} & \cdots & \alpha_1^{(n)} \\ \alpha_2^{(1)} & \cdots & \alpha_2^{(n)} \\ \vdots & \ddots & \vdots \\ \alpha_n^{(1)} & \cdots & \alpha_n^{(n)} \end{bmatrix}.$$

由 T^\times 的定義知

$$T^\times(\vec{c}_j^*)(\vec{b}_k) = c_j^*(T(\vec{b}_k)), \quad j, k = 1, \dots, n,$$

而這就是

$$\alpha_k^{(j)} = \beta_j^{(k)}.$$

於是我們得到下面的定理:

定理 3.2.1: $[T^\times]_{\mathcal{C}^*, \mathcal{B}^*} = ([T]_{\mathcal{B}, \mathcal{C}})^T$, 即 T 的伴隨算子 T^\times 所對應的矩陣表示是 T 所對應的矩陣表示的轉置。

3.3. 共軛算子

在上一節中對於一般的向量空間上的線性變換, 定義並討論了其伴隨算子。當向量空間是內積空間, 對其上的線性變換, 則可定義並討論其共軛算子。這是內積空間中十分重要的算子, 由此可以導出一系列的結果, 這是本節的內容。由內積空間的 Riesz 表示定理 (見 2.4 節), 若 \mathcal{V} 是有限維內積空間, $f \in \mathcal{V}^*$, 則存在唯一的 $\vec{x} \in \mathcal{V}$, 使得

$$f(\vec{v}) = \langle \vec{v}, \vec{x} \rangle, \quad \forall \vec{v} \in \mathcal{V}.$$

由此可以定義映射 $\phi : \mathcal{V}^* \rightarrow \mathcal{V}$ 為 $\phi(f) = \vec{x}$, 即 $\phi(f)$ 定義為

$$f(\vec{v}) = \langle \vec{v}, \phi(f) \rangle.$$

由於對 $\vec{v} \in \mathcal{V}$, $f, g \in \mathcal{V}^*$, $\alpha, \beta \in \mathbb{F}$, 我們有

$$\begin{aligned} \langle \vec{v}, \phi(\alpha f + \beta g) \rangle &= (\alpha f + \beta g)(\vec{v}) = \alpha f(\vec{v}) + \beta g(\vec{v}) \\ &= \langle \vec{v}, \alpha \phi(f) \rangle + \langle \vec{v}, \beta \phi(g) \rangle = \langle \vec{v}, \alpha \phi(f) + \beta \phi(g) \rangle, \end{aligned}$$

即

$$\phi(\alpha f + \beta g) = \alpha \phi(f) + \beta \phi(g),$$

於是 ϕ 為一線性變換。我們很容易看出 ϕ 為映成。另外, $\phi(f) = 0$ 導出 $f = 0$, 故 ϕ 也是一對一, 從而 $\phi : \mathcal{V}^* \rightarrow \mathcal{V}$ 為同構。令 $\mathbb{F} = \mathbb{C}$ 為複數體, 若 \mathcal{V} 與 \mathcal{W} 為兩個有限維的複內積空間, $T \in \mathcal{L}(\mathcal{V}, \mathcal{W})$, 對一固定的 $\vec{w} \in \mathcal{W}$, 考慮函數 $\psi_{\vec{w}} : \mathcal{V} \rightarrow \mathbb{F}$ 定義為

$$\psi_{\vec{w}}(\vec{v}) = \langle T(\vec{v}), \vec{w} \rangle.$$

則不難證出: $\psi_{\vec{w}}$ 是 \mathcal{V} 上的線性泛函。由 Riesz 表示定理, 存在唯一的 $\vec{x} \in \mathcal{V}$, 使得

$$\psi_{\vec{w}}(\vec{v}) = \langle T(\vec{v}), \vec{w} \rangle = \langle \vec{v}, \vec{x} \rangle, \quad \forall \vec{v} \in \mathcal{V}.$$

令 $T^*(\vec{w}) = \vec{x}$, 則

$$\langle T(\vec{v}), \vec{w} \rangle = \langle \vec{v}, T^*(\vec{w}) \rangle, \quad \forall \vec{v} \in \mathcal{V}.$$

故存在唯一的 T^* 使得上式成立。可證 T^* 是線性的: 對任意的 $\vec{v} \in \mathcal{V}$, $\vec{w}, \vec{w}' \in \mathcal{W}$,

$$\begin{aligned} \langle \vec{v}, T^*(\alpha\vec{w} + \beta\vec{w}') \rangle &= \langle T(\vec{v}), \alpha\vec{w} + \beta\vec{w}' \rangle = \alpha \langle T(\vec{v}), \vec{w} \rangle + \beta \langle T(\vec{v}), \vec{w}' \rangle \\ &= \alpha \langle \vec{v}, T^*(\vec{w}) \rangle + \beta \langle \vec{v}, T^*(\vec{w}') \rangle \\ &= \langle \vec{v}, \alpha T^*(\vec{w}) \rangle + \langle \vec{v}, \beta T^*(\vec{w}') \rangle \\ &= \langle \vec{v}, \alpha T^*(\vec{w}) + \beta T^*(\vec{w}') \rangle, \end{aligned}$$

故

$$T^*(\alpha\vec{w} + \beta\vec{w}') = \alpha T^*(\vec{w}) + \beta T^*(\vec{w}').$$

因此, $T^* \in \mathcal{L}(\mathcal{W}, \mathcal{V})$, 稱 T^* 為 T 的共軛算子 (adjoint transformation)。

我們現在來討論 T^* 與 T^\times 之間的關聯。若 \mathcal{V} 與 \mathcal{W} 為兩個複向量空間, $T \in \mathcal{L}(\mathcal{V}, \mathcal{W})$, 則

$$T^\times : \mathcal{W}^* \rightarrow \mathcal{V}^*, \quad T^* : \mathcal{W} \rightarrow \mathcal{V}.$$

於是由前面定義的 $\phi_1 : \mathcal{V}^* \rightarrow \mathcal{V}$, $\phi_2 : \mathcal{W}^* \rightarrow \mathcal{W}$, 我們知道:

$$\begin{array}{ccc} \mathcal{V}^* & \xleftarrow{T^\times} & \mathcal{W}^* \\ \phi_1 \downarrow & & \downarrow \phi_2 \\ \mathcal{V} & \xleftarrow{T^*} & \mathcal{W} \end{array}$$

定義 $S : \mathcal{W}^* \rightarrow \mathcal{V}^*$ 為

$$S = (\phi_1)^{-1} \circ T^* \phi_2,$$

這是線性映射。若 $\vec{x} \in \mathcal{V}$, $\phi_1^{-1}(\vec{x}) = g$, 則 $\phi_1(g) = \vec{x}$,

$$\phi_1^{-1}(\vec{x})(\vec{v}) = g(\vec{v}) = \langle \vec{v}, \phi_1(g) \rangle = \langle \vec{v}, \vec{x} \rangle.$$

因此對任意 $f \in \mathcal{W}^*$, $\vec{v} \in \mathcal{V}$, 我們有

$$\begin{aligned} [S(f)](\vec{v}) &= [(\phi_1)^{-1}T^*\phi_2(f)](\vec{v}) = (\phi_1)^{-1}[T^*\phi_2(f)](\vec{v}) \\ &= \langle \vec{v}, T^*\phi_2(f) \rangle = \langle T(\vec{v}), \phi_2(f) \rangle \\ &= f(T(\vec{v})) = T^\times(f)(\vec{v}), \end{aligned}$$

所以 $S = T^\times$ 。因此

$$T^\times = (\phi_1)^{-1}T^*\phi_2,$$

即上圖是交換圖。

若 $\mathcal{B} = \{\vec{b}_1, \vec{b}_2, \dots, \vec{b}_n\}$ 是 \mathcal{V} 的一組正規正交基底, 而 $\mathcal{C} = \{\vec{c}_1, \vec{c}_2, \dots, \vec{c}_m\}$ 是 \mathcal{W} 的一組正規正交基底, 已知

$$[T]_{\mathcal{B}, \mathcal{C}} = [[T(\vec{b}_1)]_{\mathcal{C}}, \dots, [T(\vec{b}_n)]_{\mathcal{C}}] = \begin{bmatrix} \beta_1^{(1)} & \beta_1^{(2)} & \dots & \beta_1^{(m)} \\ \beta_2^{(1)} & \beta_2^{(2)} & \dots & \beta_2^{(m)} \\ \vdots & \vdots & \ddots & \vdots \\ \beta_m^{(1)} & \beta_m^{(2)} & \dots & \beta_m^{(m)} \end{bmatrix}.$$

於是 $\beta_j^{(k)} = \langle T(\vec{b}_k), \vec{c}_j \rangle$ 。而

$$[T^*]_{\mathcal{C}, \mathcal{B}} = [[T^*(\vec{c}_1)]_{\mathcal{B}}, \dots, [T^*(\vec{c}_m)]_{\mathcal{B}}] = \begin{bmatrix} \gamma_1^{(1)} & \gamma_1^{(2)} & \dots & \gamma_1^{(m)} \\ \gamma_2^{(1)} & \gamma_2^{(2)} & \dots & \gamma_2^{(m)} \\ \vdots & \vdots & \ddots & \vdots \\ \gamma_n^{(1)} & \gamma_n^{(2)} & \dots & \gamma_n^{(m)} \end{bmatrix}.$$

於是 $\gamma_k^{(j)} = \langle T^*(\vec{c}_j), \vec{b}_k \rangle$ 。但是

$$\langle T^*(\vec{c}_j), \vec{b}_k \rangle = \overline{\langle \vec{b}_k, T^*(\vec{c}_j) \rangle} = \overline{\langle T(\vec{b}_k), \vec{c}_j \rangle} = \overline{\beta_j^{(k)}},$$

因此得到如下定理:

定理 3.3.1: $[T^*]_{\mathcal{C}, \mathcal{B}} = \overline{[T]_{\mathcal{B}, \mathcal{C}}} = [[T]_{\mathcal{B}, \mathcal{C}}]^*$, 即 T 的共軛算子 T^* 所對應矩陣為 T 所對應矩陣的共軛轉置, 也用 $*$ 來記之。

對於 T 的共軛算子, 易證以下的這些性質。

命題 3.3.1: 若 \mathcal{V} 與 \mathcal{W} 是兩個有限維內積空間, $T, S \in \mathcal{L}(\mathcal{V}, \mathcal{W})$, 則

- (1) $\langle T^*(\vec{w}), \vec{v} \rangle = \langle \vec{w}, T(\vec{v}) \rangle$;
- (2) $(T + S)^* = T^* + S^*$;
- (3) $(\alpha T)^* = \alpha T^*$, $\alpha \in \mathbb{F}$;
- (4) $T^{**} = T$;
- (5) 若 $\mathcal{V} = \mathcal{W}$, 則 $(S \circ T)^* = T^* \circ S^*$;
- (6) 若 T 為可逆, 則 $(T^{-1})^* = (T^*)^{-1}$ 。

進一步可給出如下的定義:

定義 3.3.1: 若 \mathcal{V} 是內積空間, $T \in \mathcal{L}(\mathcal{V})$, 則

- (1) 稱 T 為自共軛 (self-adjoint) 或 Hermitian, 若 $T^* = T$;
- (2) 稱 T 為酉 (unitary), 若 T 是一對一, 映成且 $T^* = T^{-1}$;
- (3) 稱 T 為正規 (normal), 若 $TT^* = T^*T$ 。

定義 3.3.2: 當 A 是一個 $n \times n$ 複矩陣, 則

- (1) 若 $A = A^*$, 稱 A 是 Hermitian;
- (2) 若 $A = -A^*$, 稱 A 是 skew-Hermitian;
- (3) 若 A 為可逆且 $A^* = A^{-1}$, 稱 A 是酉;
- (4) 若 $AA^* = A^*A$, 稱 A 是正規。

當 A 是一個 $n \times n$ 實矩陣, 則

- (a) 若 $A^T = A$, 稱 A 是對稱 (symmetric);
- (b) 若 $A^T = -A$, 稱 A 是斜對稱 (skew-symmetric);
- (c) 若 A 為可逆且 $A^T = A^{-1}$, 稱 A 是正交 (orthogonal)。

我們很容易證明 T 是正規、自共軛、酉若且唯若 T 在正規正交基底下對應的矩陣是正規、Hermitian、酉矩陣。下面我們來討論這三類的算子。

A. 自共軛算子 (self-adjoint operators)

由定義 $T \in \mathcal{L}(\mathcal{V})$ 是自共軛, 若對所有的 $\vec{v}, \vec{w} \in \mathcal{V}$, 有

$$\langle T(\vec{v}), \vec{w} \rangle = \langle \vec{v}, T(\vec{w}) \rangle,$$

即 $T^* = T$ 。不難證明下面的性質:

命題 3.3.2: 若 \mathcal{V} 是內積空間, $S, T \in \mathcal{L}(\mathcal{V})$,

- (1) 若 S, T 為自共軛, 則 $S + T$ 也是自共軛;

- (2) 若 T 為自共軛, α 為實數, 則 αT 也是自共軛;
- (3) 若 S, T 為自共軛, $S \circ T$ 也是自共軛若且唯若 $S \circ T = T \circ S$;
- (4) 若 T 為自共軛且可逆, 則 T^{-1} 也是自共軛;
- (5) 若 T 為自共軛, 則對所有的 $\vec{v} \in \mathcal{V}$, $\langle T(\vec{v}), \vec{v} \rangle$ 是實數;
- (6) 假設 \mathcal{V} 是酉空間, 若對所有的 $\vec{v} \in \mathcal{V}$, $\langle T(\vec{v}), \vec{v} \rangle$ 是實數, 則 T 是自共軛;
- (7) 若 T 為自共軛, 則對所有的 $\vec{v} \in \mathcal{V}$, $\langle T(\vec{v}), \vec{v} \rangle = 0$, 則 $T = 0$;
- (8) 若 T 為自共軛, 則對任意的 $k > 0$, 由 $T^k(\vec{v}) = \vec{0}$ 可導出 $T = 0$ 。

證明: 由於 T 為自共軛, 所以

$$\langle T(\vec{v}), \vec{v} \rangle = \langle \vec{v}, T(\vec{v}) \rangle = \overline{\langle T(\vec{v}), \vec{v} \rangle},$$

故 $\langle T(\vec{v}), \vec{v} \rangle$ 是實數, 這便證明了 (5)。為了證明 (6) 與 (7), 我們先證明: 假設 \mathcal{V} 是酉空間, 若對所有的 $\vec{v} \in \mathcal{V}$, $\langle T(\vec{v}), \vec{v} \rangle = 0$, 則 $T = 0$ 。

令 $\vec{v} = \alpha \vec{x} + \vec{y}$, 這裡 $\vec{x}, \vec{y} \in \mathcal{V}$, $\alpha \in \mathbb{C}$, 則

$$\begin{aligned} 0 &= \langle T(\alpha \vec{x} + \vec{y}), \alpha \vec{x} + \vec{y} \rangle \\ &= |\alpha|^2 \langle T(\vec{x}), \vec{x} \rangle + \langle T(\vec{y}), \vec{y} \rangle + \alpha \langle T(\vec{x}), \vec{y} \rangle + \bar{\alpha} \langle T(\vec{y}), \vec{x} \rangle \\ &= \alpha \langle T(\vec{x}), \vec{y} \rangle + \bar{\alpha} \langle T(\vec{y}), \vec{x} \rangle. \end{aligned}$$

取 $\alpha = 1$, 則

$$\langle T(\vec{x}), \vec{y} \rangle + \langle T(\vec{y}), \vec{x} \rangle = 0.$$

取 $\alpha = \sqrt{-1}$, 則

$$\langle T(\vec{x}), \vec{y} \rangle - \sqrt{-1} \langle T(\vec{y}), \vec{x} \rangle = 0.$$

故對所有 $\vec{x}, \vec{y} \in \mathcal{V}$, $\langle T(\vec{x}), \vec{y} \rangle = 0$ 都成立, 因此 $T = 0$ 。現在用這個來證明 (6),

$$\begin{aligned} \langle (T - T^*)(\vec{v}), \vec{v} \rangle &= \langle T(\vec{v}), \vec{v} \rangle - \langle T^*(\vec{v}), \vec{v} \rangle \\ &= \langle T(\vec{v}), \vec{v} \rangle - \langle \vec{v}, T(\vec{v}) \rangle \\ &= \langle T(\vec{v}), \vec{v} \rangle - \overline{\langle T(\vec{v}), \vec{v} \rangle} = 0. \end{aligned}$$

由上述結果知 $T - T^* = 0$, 即 T 是自共軛。

我們再來證明 (7)。當 $\mathbb{F} = \mathbb{C}$, 由上面的結果我們立即得到 (7)。所以只要證明當 $\mathbb{F} = \mathbb{R}$

的情形, 此時

$$\begin{aligned} 0 &= \langle T(\vec{x} + \vec{y}), \vec{x} + \vec{y} \rangle \\ &= \langle T(\vec{x}), \vec{x} \rangle + \langle T(\vec{y}), \vec{y} \rangle + \langle T(\vec{x}), \vec{y} \rangle + \langle T(\vec{y}), \vec{x} \rangle \\ &= \langle T(\vec{x}), \vec{y} \rangle + \langle T(\vec{y}), \vec{x} \rangle = \langle T(\vec{x}), \vec{y} \rangle + \langle \vec{x}, T(\vec{y}) \rangle \\ &= \langle T(\vec{x}), \vec{y} \rangle + \langle T(\vec{x}), \vec{y} \rangle = 2\langle T(\vec{x}), \vec{y} \rangle, \end{aligned}$$

故 $T = 0$ 。最後來證明 (8)。若 $T^k(\vec{v}) = \vec{0}$ 對所有 $\vec{v} \in \mathcal{V}$ 都成立, 則存在 m , 使得 $2^m \geq k$ 。於是, $T^{2^m}(\vec{v}) = \vec{0}$ 。因此

$$0 = \langle T^{2^m}(\vec{v}), \vec{v} \rangle = \langle T^{2^{m-1}}T^{2^{m-1}}(\vec{v}), \vec{v} \rangle = \langle T^{2^{m-1}}(\vec{v}), T^{2^{m-1}}(\vec{v}) \rangle.$$

故 $T^{2^{m-1}}(\vec{v}) = \vec{0}$ 。重覆這樣的步驟, 最後得到 $T = 0$, 命題因而證畢。

對於複數上的內積空間的任意一個線性變換 $T \in \mathcal{L}(\mathcal{V}, \mathcal{W})$, 均可寫成

$$T = T_1 + \sqrt{-1}T_2,$$

這裡 T_1 與 T_2 為自共軛變換。事實上, 令

$$T_1 = \frac{T + T^*}{2}, \quad T_2 = \frac{T - T^*}{2\sqrt{-1}}$$

即得上式。這與“任意一個複數 z 均可寫成 $x + iy$, 這裡 $x, y \in \mathbb{R}$ ”相類似。所以, 對於內積空間, 自共軛變換是十分基本的。

B. 酉算子 (unitary operators)

由定義, $T \in \mathcal{L}(\mathcal{V})$ 是酉算子, 若對所有的 $\vec{v}, \vec{w} \in \mathcal{V}$, 有

$$\langle T(\vec{v}), \vec{w} \rangle = \langle \vec{v}, T^{-1}(\vec{w}) \rangle.$$

易證如下的性質:

命題 3.3.3: 若 \mathcal{V} 是內積空間, $S, T \in \mathcal{L}(\mathcal{V})$,

- (1) 若 T 為一酉算子, 則 T^{-1} 也是酉算子;
- (2) 若 S, T 為酉算子, $S \circ T$ 也是酉算子;
- (3) T 為酉算子若且唯若 T 是一對一, 映成且保長。一個線性變換 $T \in \mathcal{L}(\mathcal{V})$ 稱為保長變換 (isometry), 若對所有 $\vec{v}, \vec{w} \in \mathcal{V}$, 有

$$\langle T(\vec{v}), T(\vec{w}) \rangle = \langle \vec{v}, \vec{w} \rangle;$$

(4) 若 \mathcal{V} 為有限維向量空間, 則 T 為酉算子若且唯若 T 將正規正交基底映為正規正交基底。

證明: 結論 (1) 與 (2) 是顯而易見的。現在來證明 (3)。

若 T 為一對一且映成, 則 T 為保長變換若且唯若對所有 $\vec{v}, \vec{w} \in \mathcal{V}$, 有

$$\langle T(\vec{v}), T(\vec{w}) \rangle = \langle \vec{v}, \vec{w} \rangle,$$

若且唯若對所有 $\vec{w} \in \mathcal{V}$, 有

$$T^* T(\vec{w}) = \vec{w},$$

若且唯若 $T^* T = \mathbf{I}$, 這裡 \mathbf{I} 是恆等映射。這便等價於 $T^* = T^{-1}$, 所以 T 是一個酉算子。

最後來證明 (4)。若 T 是一個酉算子, 且 $\mathcal{O} = \{\vec{u}_1, \dots, \vec{u}_n\}$ 是 \mathcal{V} 的一組正規正交基底, 於是

$$\langle T(\vec{u}_j), T(\vec{u}_k) \rangle = \langle \vec{u}_j, \vec{u}_k \rangle = \delta_{jk}.$$

故 $T(\mathcal{O})$ 也是 \mathcal{V} 的一組正規正交基底。反之, 若 \mathcal{O} 與 $T(\mathcal{O})$ 都是 \mathcal{V} 的正規正交基底, 則

$$\langle T(\vec{u}_j), T(\vec{u}_k) \rangle = \delta_{jk} = \langle \vec{u}_j, \vec{u}_k \rangle.$$

若 $\vec{v} = \sum_{j=1}^n \alpha_j \vec{u}_j$ 及 $\vec{w} = \sum_{j=1}^n \beta_j \vec{u}_j$, 則

$$\begin{aligned} \langle T(\vec{v}), T(\vec{w}) \rangle &= \left\langle \sum_{j=1}^n \alpha_j T(\vec{u}_j), \sum_{k=1}^n \beta_k T(\vec{u}_k) \right\rangle = \sum_{j,k=1}^n \alpha_j \bar{\beta}_k \langle T(\vec{u}_j), T(\vec{u}_k) \rangle \\ &= \sum_{j,k=1}^n \alpha_j \bar{\beta}_k \langle \vec{u}_j, \vec{u}_k \rangle = \left\langle \sum_{j=1}^n \alpha_j \vec{u}_j, \sum_{k=1}^n \beta_k \vec{u}_k \right\rangle \\ &= \langle \vec{v}, \vec{w} \rangle, \end{aligned}$$

故 T 是一個酉算子, 命題因此證畢。

與酉算子對應的是酉矩陣, 對酉矩陣顯然有以下性質。

命題 3.3.4: 若 A 為一 $n \times n$ 矩陣,

- (1) 矩陣 A 為酉矩陣若且唯若 A 的所有的列組成 \mathbb{C}^n 的一組正規正交基底;
- (2) 矩陣 A 為酉矩陣若且唯若 A 的所有的行組成 \mathbb{C}^n 的一組正規正交基底;
- (3) 矩陣 A 為酉矩陣, 則 $|\det(A)| = 1$, 即 $\det(A) = \pm 1$ 。

C. 正規算子 (regular operators)

由定義, $T \in \mathcal{L}(\mathcal{V})$ 是正規算子, 若 $T \circ T^* = T^* \circ T$ 。易證如下的性質

命題 3.3.5: 若 \mathcal{V} 是內積空間, T 是 \mathcal{V} 上的正規算子,

- (1) 若 $p(x) \in \mathbb{C}[x]$ 為對任意多項式, 則 $p(T)$ 也是 \mathcal{V} 上的正規算子;
 (2) 若 $T(\vec{v}) = \vec{0}$, 則 $T^*(\vec{v}) = \vec{0}$;
 (3) 對任意的 $k > 0$, 若 $T^k(\vec{v}) = \vec{0}$, 則 $T(\vec{v}) = \vec{0}$;
 (4) 對任意 $\lambda \in \mathbb{C}$, 若 $(T - \lambda)^k(\vec{v}) = \vec{0}$, 則 $(T - \lambda)(\vec{v}) = \vec{0}$;
 (5) 若 $T(\vec{v}) = \lambda \vec{v}$, 則 $T^*(\vec{v}) = \bar{\lambda} \vec{v}$.

證明: 結論 (1) 與 (2) 是顯而易見的。現在來證明 (3): 算子 $S = T \circ T^*$ 易見是自共軛的, 由於 T 是正規的, 故

$$S^k(\vec{v}) = (T^*)^k(T)^k(\vec{v}) = \vec{0}.$$

由命題3.3.2的 (8), 得知 $S(\vec{v}) = \vec{0}$, 即 $(T \circ T^*)(\vec{v}) = \vec{0}$, 但

$$0 = \langle (T^* \circ T)(\vec{v}), \vec{v} \rangle = \langle T(\vec{v}), T(\vec{v}) \rangle.$$

故 $T(\vec{v}) = \vec{0}$ 。

由 (1) 與 (3), 我們得到結論 (4)。最後來證明 (5)。若 $T(\vec{v}) = \lambda \vec{v}$, 這裡 $\vec{v} \neq \vec{0}$, 則 $(T - \lambda)(\vec{v}) = \vec{0}$ 。由 (2), $(T - \lambda)^*(\vec{v}) = \vec{0}$, 但 $(T - \lambda)^* = T^* - \bar{\lambda}$, 故得 (5), 命題因此證畢。

在最後一講中, 我們將對這三個算子作進一步深入的討論, 尤其是分解定理。

—本文作者龔昇任教於中國科技大學; 張德健任教於美國 Georgetown University 數學系—