數學示例：特徵值與譜

我們在《數學示例：線性算子》中介紹了線性算子的概念,現考慮 $L(\mathbb{C}^n, \mathbb{C}^n)$ 的成員，即把 $\mathbb{C}^n$ 映射到 $\mathbb{C}^n$ 內的線性算子。有一些線性算子的作用是把 $\mathbb{C}^n$ 的成員一律乘以某個倍數，例如如下定義的 $T_1 \in L(\mathbb{C}^2, \mathbb{C}^2)$：

$$T_1(x_1, x_2) = (2x_1, 2x_2) \qquad (1)$$

便把 $\mathbb{C}^2$ 的成員一律乘以 2。

雖然 $L(\mathbb{C}^n, \mathbb{C}^n)$ 的一般成員 T 並不把 $\mathbb{C}^n$ 的成員一律乘以某個倍數，但對於 $\mathbb{C}^n$ 中的某些成員來說，T 的作用卻是把那些成員乘以某個倍數。舉例說，考慮如下定義的 $T_2 \in L(\mathbb{C}^2, \mathbb{C}^2)$：

$$T_2(x_1, x_2) = (5x_1 + 4x_2, x_1 + 2x_2) \qquad (2)$$

如把 T_2 作用於 $(1, 1)$，所得結果是 $T_2(1, 1) = (9, 3)$，其中 $(9, 3)$ 並非 $(1, 1)$ 的倍數。但如果把 T_2 作用於 $(4, 1)$，所得結果是 $T_2(4, 1) = (24, 6)$，其中 $(24, 6)$ 是 $(4, 1)$ 的 6 倍。事實上，讀者可以驗證，如把 T_2 作用於 $\mathbb{C}^2$ 中具有 $(4x, x)$ (其中 $x \in \mathbb{C}$) 形式的成員，所得結果是 $T_2(4x, x) = (24x, 6x)$，即 $(4x, x)$ 的 6 倍。不僅如此，讀者還可驗證，如把 T_2 作用於 $\mathbb{C}^2$ 中具有 $(x, -x)$ (其中 $x \in \mathbb{C}$) 形式的成員，所得結果是 $T_2(x, -x) = (x, -x)$，即 $(x, -x)$ 的 1 倍。

一般地，對於 $L(\mathbb{C}^n, \mathbb{C}^n)$ 的成員 T，如有某個複數 λ 以及 $\mathbb{C}^n$ 的某個非零成員 $(x_1, \ldots, x_n)$ 使得以下等式成立：

$$T(x_1, \ldots, x_n) = \lambda(x_1, \ldots, x_n) \qquad (3)$$

我們便說 λ 是 T 的特徵值(eigenvalue)，$(x_1, \ldots, x_n)$ 則是對應於 λ 的特徵向量(eigenvector)。就特定的 T 而言，T 可能有不只一個特徵值，如把 T 的所有特徵值組成一個集合，這個集合稱為譜(spectrum)，以下記作 $\sigma(T)$；$\sigma(T)$ 以外的所有複數則組成另一個集合，稱為 T 的預解集(resolvent set)，以下記作 $\rho(T)$，即 $\rho(T) = \mathbb{C} - \sigma(T)$。此外，對應於某個特徵值的特徵向量也不只一個，事實上，可以證明這些特徵向量連同 $(0, \ldots, 0)$ 構成 $\mathbb{C}^n$ 的一

個子空間[1]，稱為特徵空間(eigenspace)。

以前述的 T_1 為例，根據前面的討論，可知 2 是 T_1 的特徵值，而且可以證明 2 是 T_1 唯一的特徵值，故有 $\sigma(T_1) = \{2\}$，並且 $\rho(T_1) = \mathbb{C} - \{2\}$。另外，從 (1) 可知，$\mathbb{C}^2$ 中的任何非零元素都是對應於 2 的特徵向量，因此對應於 2 的特徵空間就是 $\mathbb{C}^2$。

另外又如前述的 T_2，根據前面的討論，可知 T_2 有兩個特徵值，即 1 和 6，故有 $\sigma(T_2) = \{1, 6\}$，並且 $\rho(T_2) = \mathbb{C} - \{1, 6\}$。另外，對應於 1 的特徵向量是 $\mathbb{C}^2$ 中所有具有 $(x, -x)$ 形式的非零成員，而對應於 6 的特徵向量則是 $\mathbb{C}^2$ 中所有具有 $(4x, x)$ 形式的非零成員，因此對應於 1 的特徵空間是 $\{(x, -x) : x \in \mathbb{C}\}$，而對應於 6 的特徵空間則是 $\{(4x, x) : x \in \mathbb{C}\}$。

接著介紹求特徵值和特徵向量的方法。根據線性代數的知識，當取定 $\mathbb{C}^n$ 上的某個 (哈默爾) 基底後，可以把 $\mathbb{C}^n$ 上的線性算子 T 表示成 $n \times n$ 矩陣 $[T]$，因此我們可以把求 T 的特徵值和特徵向量的問題轉化為求 $[T]$ 的特徵值和特徵向量的問題。

在矩陣的語言下，等式 (3) 具有以下形式：

$$[T] \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = \lambda \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$$

為簡化推導，以下首先設定

$$[X] = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$$

這樣便可以把上式化簡為

$$[T][X] = \lambda[X] \qquad (4)$$

接下來的目標就是要找出滿足上式的複數 λ 和非零向量 $[X]$，從上式可以得到

$$[T][X] - \lambda[X] = [0] \qquad (5)$$

[1]請注意 $(0, \ldots, 0)$ 必然滿足 (3)，這是因為對任何線性算子 T 而言，必有 $T(0, \ldots, 0) = T((x_1, \ldots, x_n) + (-x_1, \ldots, -x_n)) = T(x_1, \ldots, x_n) + T(-x_1, \ldots, -x_n) = T(x_1, \ldots, x_n) - T(x_1, \ldots, x_n) = (0, \ldots, 0)$，而對任何複數 λ 而言，也必有 $\lambda(0, \ldots, 0) = (0, \ldots, 0)$。惟請注意，根據特徵向量的定義，$(0, \ldots, 0)$ 不是任何 T 的特徵向量。

其中 $[0]$ 代表由 n 個 0 組成的「零向量」。接著我們希望從上式左端中抽出 $[X]$ 這個因子，從而得到 $([T]-\lambda)[X]$，但我們不能這樣做，這是因為 $[T]$ 是矩陣，而 λ 卻是複數，兩者不能相減。為解決這個矛盾，一個方法是把上式中的複數 λ 改為矩陣 $\lambda[I_n]$，其中 $[I_n]$ 代表 n 階「單位矩陣」(identity matrix)，即以下矩陣：

$$[I_2]=\begin{bmatrix}1&0\\0&1\end{bmatrix},\qquad [I_3]=\begin{bmatrix}1&0&0\\0&1&0\\0&0&1\end{bmatrix},\qquad [I_4]=\begin{bmatrix}1&0&0&0\\0&1&0&0\\0&0&1&0\\0&0&0&1\end{bmatrix},\qquad \ldots$$

讀者可自行驗證，$\lambda[X]=\lambda[I_n][X]$。用 $\lambda[I_n]$ 取代 (5) 中的 λ 並抽出因子 $[X]$，可得

$$([T]-\lambda[I_n])[X]=[0] \qquad (6)$$

如果存在逆矩陣 $([T]-\lambda[I_n])^{-1}$，那麼把這個逆矩陣作用於上式左右兩端，便會得到 $[X]=[0]$。但我們現在的目標是求滿足上式的非零 $[X]$，為此必須假設 $([T]-\lambda[I_n])^{-1}$ 不存在。根據線性代數的知識，這等同於

$$\det([T]-\lambda[I_n])=0 \qquad (7)$$

其中 $\det([M])$ 代表矩陣 $[M]$ 的「行列式」(determinant)。上式就是求 $[T]$ 的特徵值的方程，稱為特徵方程(characteristic equation)。由於上式是以 λ 作為變項的多項式，根據「代數基本定理」(Fundamental Theorem of Algebra)，上式在複數域中必有解，因此任何 $[T]$ 都有至少一個特徵值。用上式求得特徵值後，便可利用 (4) 求相對應的特徵向量。

現以前述的 T_2 為例說明上述概念。根據線性代數的知識，為 $\mathbb{C}$ 取不同的基底，可把同一個線性算子表達成不同的矩陣，但可以證明這些矩陣有相同的特徵值，因此為計算 T_2 的特徵值，我們可以任意取一個基底。為方便起見，我們取標準基底。在此基底下，T_2 可以用以下矩陣代表：

$$[T_2]=\begin{bmatrix}5&4\\1&2\end{bmatrix}$$

接著求以下矩陣：

$$[T_2]-\lambda[I_2]=\begin{bmatrix}5-\lambda&4\\1&2-\lambda\end{bmatrix}$$

根據線性代數中行列式的公式，可求得

$$\det\left(\begin{bmatrix}5-\lambda&4\\1&2-\lambda\end{bmatrix}\right)=(5-\lambda)(2-\lambda)-4$$

把上述結果套用於 (7)，可得以下方程：

$$\lambda^2 - 7\lambda + 6 = 0$$

容易求得上述方程的兩個根為 $\lambda_1 = 1$ 和 $\lambda_2 = 6$，這兩個根就是 T_2 的特徵值。接著便可以利用 (4) 求對應用於這兩個特徵值的特徵向量，以 λ_2 為例，把有關數值代入 (4)，可得

$$\begin{bmatrix} 5 & 4 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = 6 \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

由此可得以下聯立方程：

$$\begin{cases} 5x_1 + 4x_2 = 6x_1 \\ x_1 + 2x_2 = 6x_2 \end{cases}$$

請注意上述聯立方程不是線性獨立的[2]，兩式都等於 $x_1 = 4x_2$，因此 T_2 對應於 6 的特徵向量是 $\mathbb{C}^2$ 中所有具有 $(4x, x)$ 形式的非零成員。讀者可自行驗證，利用 (4) 可求得 T_2 對應於 1 的特徵向量是 $\mathbb{C}^2$ 中所有具有 $(x, -x)$ 形式的非零成員。讀者也可自行驗證，用上述方法可同樣求得前面的 T_1 的唯一特徵值是 2，並且 $\mathbb{C}^2$ 中的所有非零成員都是對應於這個特徵值的特徵向量。

前面介紹的譜、預解集等是與 $\mathbb{C}^n$ 上的線性算子相關的概念，其中 $\mathbb{C}^n$ 是有限維賦範空間，我們也可以把這些概念推廣到無限維賦範空間，但情況會較為複雜。在進行推廣前，須先引入「稠密」的概念，這本是拓樸學上的概念，為免引入拓樸學中的複雜術語，這裡僅提供此概念的直觀意義。設 X 為距離空間，$S \subseteq X$，如果對 X 中任何元素 x 而言，要麼 $x \in S$，要麼與 x 相距任意小的範圍內，都總能找到一個 S 的元素，我們便說 S 在 X 中稠密(dense)。

舉例說，$\mathbb{Q}$ 在 $\mathbb{R}$ 中是稠密的，這是因為對任何實數 x 而言，要麼 x 是有理數，要麼在與 x 相距任意小的範圍內，都總能找到一個有理數。接著再提供一個不稠密集合的例子。考慮距離空間 l^2，這是由滿足 $\sum_{n=1}^{\infty} |x_n|^2 < \infty$ 的複數序列 $(x_n)_{n=1}^{\infty}$ 組成的空間，其距離函數為

$$d((x_n)_{n=1}^{\infty}, (y_n)_{n=1}^{\infty}) = \left(\sum_{n=1}^{\infty} |x_n - y_n|^2 \right)^{\frac{1}{2}} \qquad (8)$$

在這個空間中，以下集合不是稠密的：

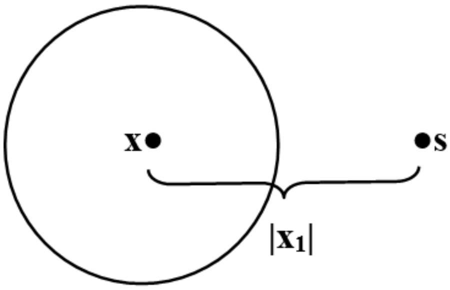
$$S = \left\{ (0, x_2, x_3, x_4, \ldots) : \sum_{n=2}^{\infty} |x_n|^2 < \infty \right\} \qquad (9)$$

[2]這是必然的，因為我們在前面設定了這組聯立方程的係數所組成的矩陣的行列式等於 0。

為證明這一點，任選 l^2 中的成員 $x=(x_1,x_2,x_3,x_4,\ldots)$，其中 $x_1\neq 0$，那麼與 x 距離最近的 S 中成員只能是 $s=(0,x_2,x_3,x_4,\ldots)$，而 x 與 s 的距離是

$$\begin{aligned} d(x,s) &= \left(|x_1-0|^2+|x_2-x_2|^2+|x_3-x_3|^2+|x_4-x_4|^2+\ldots\right)^{\frac{1}{2}} \\ &= |x_1| \end{aligned}$$

現把上述計算結果表示成下圖以幫助讀者了解：



上圖顯示 x 與 s 的距離是 $|x_1|$，圖中的圓圈代表與 x 相距小於 $|x_1|$ 的某個範圍，由於 s 是與 x 距離最近的 S 的成員，因此在這個圓圈範圍內不可能有任何 S 的成員，由此根據稠密的定義，可知 S 在 l^2 中不是稠密的。

具備上述概念後，便可以把前述的譜、預解集等概念推廣到一般賦範空間上的線性算子。設 $V_{\mathbb{C}}$ 為賦範空間，T 為 $L(V_{\mathbb{C}},V_{\mathbb{C}})$ 的成員，I 為恆等函數，λ 為複數，現考慮線性算子 $T-\lambda I$，如果 λ 同時滿足以下三個條件：

(i) $T-\lambda I$ 的值域[3]是 $V_{\mathbb{C}}$ 的稠密集；

(ii) 存在 $T-\lambda I$ 的逆算子 $(T-\lambda I)^{-1}$

(iii) $(T-\lambda I)^{-1}$ 是有界算子

則 λ 稱為 T 的正則值(regular value)。由 T 的所有正則值組成的集合稱為 T 的「預解集」，記作 $\rho(T)$。$\rho(T)$ 以外的所有複數則組成另一個集合，稱為 T 的「譜」，記作 $\sigma(T)$，即 $\sigma(T)=\mathbb{C}-\rho(T)$。

根據上述定義，$\sigma(T)$ 包含不滿足上面 (i) 或 (ii) 或 (iii) 的複數，因此可以把 $\sigma(T)$ 的成員細分為若干個子類，這裡只擬介紹其中一個子類－不滿

[3]某函數的「值域」(range) 是指該函數定義域中的元素所能產生的值組成的集合，請注意「值域」是「對應域」(codomain) 的子集。舉例說，設 f 為把實數映射到實數內的函數，其定義為 $f(x)=x^2$，那麼 f 的對應域是 $\mathbb{R}$，其值域卻是 $\mathbb{R}^+\cup\{0\}$，即由所有正實數和 0 組成的集合。請注意由於 f 的值只能是正實數或 0，所以其值域是 $\mathbb{R}$ 的真子集。

足 (ii) 的複數，這類複數稱為「特徵值」。換句話說，T 的特徵值就是使得下式成立的複數：

$$T-\lambda I \text{不可逆} \qquad (10)$$

請注意上式跟前面的 (7) 是一致的，這是因為在線性代數中，一個矩陣 $[M]$ 的行列式等於 0 當且僅當 $[M]$ 不可逆。若 $T-\lambda I$ 不可逆，則它必不是把 $V_{\mathbb{C}}$ 映射到其值域上的一一函數，即必存在 $V_{\mathbb{C}}$ 中的非零成員 x 使得下式成立[4]：

$$(T-\lambda I)(x)=0 \qquad (11)$$

這樣的 x 稱為對應於 λ 的特徵向量，請注意上式跟前面的 (6) 是一致的。總上所述，新定義下的特徵值和特徵向量跟舊定義是一致的。

由特徵值組成的集合稱為點譜(point spectrum)，記作 $\sigma_p(T)$，是 $\sigma(T)$ 的子集，而且這個子集有可能是 $\sigma(T)$ 的真子集，即 $\sigma(T)$ 可能存在一些並非特徵值的複數。但是如果 T 是有限維賦範空間上的線性算子，那麼 $\sigma(T)=\sigma_p(T)$，由此可見，泛函分析中的譜和預解集等概念確是線性代數中同一概念的推廣。

接下來用兩個例子說明上述概念，第一個例子是一般賦範空間 $V_{\mathbb{C}}$ 上的恆等函數 I，在 $V_{\mathbb{C}}$ 下，所有不等於 1 的複數都是 I 的正則值，即 $\rho(I)=\mathbb{C}-\{1\}$。為證明這一點，我們考慮線性算子 $I-\lambda I$ (其中 $\lambda \neq 1$)，這個算子等同於 $(1-\lambda)I$，其作用是把 $V_{\mathbb{C}}$ 中的任何成員乘以 $1-\lambda$ [5]。接著證明 λ 滿足上面有關正則值的三個條件。首先，對 $V_{\mathbb{C}}$ 中任何元素 x 而言，都有

$$\begin{aligned}(1-\lambda)I\left(\frac{x}{1-\lambda}\right) &= (1-\lambda)\frac{x}{1-\lambda} \\ &= x\end{aligned}$$

這即是說 $V_{\mathbb{C}}$ 中任何元素 x 都是把 $(1-\lambda)I$ 作用於另一個元素 $\frac{x}{1-\lambda}$ 所得的值，因此 $(1-\lambda)I$ 的值域就是 $V_{\mathbb{C}}$，而 $V_{\mathbb{C}}$ 當然在 $V_{\mathbb{C}}$ 中稠密 (這是因為 $V_{\mathbb{C}}$ 的任何成員都是 $V_{\mathbb{C}}$ 的成員，滿足前面有關「稠密」的定義)，故條件 (i) 成立。

其次，容易看到 $((1-\lambda)I)^{-1}$ 存在，這個逆算子就是 $(\frac{1}{1-\lambda})I$，這是因為對 $V_{\mathbb{C}}$ 中任何 x，都有 $(\frac{1}{1-\lambda})I\circ(1-\lambda)I(x)=x$，並且 $(1-\lambda)I\circ(\frac{1}{1-\lambda})I(x)=x$，故條件 (ii) 成立。

---

[4]根據集合論，函數 f「可逆」(invertible，即其逆函數存在) 當且僅當 f 是把其定義域映射到其值域上的一一函數。此外，還可以證明如果 f 是線性函數，則 f 是一一的當且僅當其定義域中沒有任何非零元素 x 使得 $f(x)=0$ (即只有 $f(0)=0$)。

[5]請注意 $V_{\mathbb{C}}$ 是複數域上的向量空間，所以它的任何成員都可以與複數 $1-\lambda$ 進行純量乘法。

最後，對 $V_{\mathbb{C}}$ 中任何 x，都有 $\|(\frac{1}{1-\lambda})I(x)\| = \frac{\|x\|}{|1-\lambda|}$。由於對任何不等於 1 的複數 λ 而言，必有正實數 k，使得 $\frac{1}{|1-\lambda|} \leq k$，故必有 $\|(\frac{1}{1-\lambda})I(x)\| \leq k\|x\|$，由此根據《數學示例：線性算子》中有界線性算子的定義，可知 $(\frac{1}{1-\lambda})I$ 有界，故條件 (iii) 成立。至此證得 $\rho(I) = \mathbb{C} - \{1\}$。

當 $\lambda = 1$ 時，$(1-\lambda)I$ 等於零算子，即把任何 x 都映射為 $V_{\mathbb{C}}$ 中的 0 的算子。這個零算子顯然不是一一函數，因此是不可逆的，即上面的 (10) 成立，由此可見 1 是 I 唯一的特徵值，而 $V_{\mathbb{C}}$ 中的任何非零成員都是對應於 1 的特徵向量。由於 $\rho(I)$ 和 $\{1\}$ 窮盡了所有複數，由此可知 $\sigma(I) = \sigma_p(I) = \{1\}$。

第二個例子是 l^2 中如下定義的算子 T_3：

$$T_3(x_1, x_2, x_3, x_4, \ldots) = (0, x_1, x_2, x_3, \ldots) \qquad (12)$$

這裡不能像上例那樣容易找出 $\rho(T_3)$ 和 $\sigma(T_3)$ 全體，所以以下只提供 $\sigma(T_3)$ 的一個成員－ 0。為證明 $0 \in \sigma(T_3)$，考慮線性算子 $T_3 - 0I$，即 T_3。容易看到 T_3 的值域就是前面討論過的集合 S (見 (9))，這是因為給定 S 中任何成員 $y = (0, x_2, x_3, x_4, \ldots)$，必有 l^2 中成員 $x = (x_2, x_3, x_4, \ldots)$ 使得 $T_3x = y$。由於我們在前面指出了 S 在 l^2 中不是稠密的，由此可見 0 並不滿足前面的條件 (i)，因此 $0 \notin \rho(T_3)$，即 $0 \in \sigma(T_3)$。

可是，0 不是 T_3 的特徵值，事實上，T_3 沒有特徵值。為證明這一點，假設 λ 是 T_3 的特徵值，那麼根據 (11) 和 (12)，必有 l^2 中的非零成員 $(x_1, x_2, x_3, x_4, \ldots)$ 使得

$$\begin{aligned}
(T_3 - \lambda I)(x_1, x_2, x_3, x_4, \ldots) &= (0, 0, 0, 0, \ldots) \\
(0, x_1, x_2, x_3, \ldots) - (\lambda x_1, \lambda x_2, \lambda x_3, \lambda x_4, \ldots) &= (0, 0, 0, 0, \ldots) \\
(-\lambda x_1, x_1 - \lambda x_2, x_2 - \lambda x_3, x_3 - \lambda x_4, \ldots) &= (0, 0, 0, 0, \ldots)
\end{aligned}$$

讀者可自行驗證，不論 λ 是否等於 0，從上式都可得出 $x_1 = 0$、$x_2 = 0$、$x_3 = 0$、$x_4 = 0$、...，即 $(x_1, x_2, x_3, x_4, \ldots)$ 不是非零成員。此一矛盾顯示 T_3 不可能有特徵值，至此我們看到並非所有線性算子的譜的所有成員都是特徵值。

連結至數學專題
連結至周家發網頁