Chapter 8 特徵值、特徵向量、對角線化

本章討論限於方陣，但是有可能使用到虛數 $\sqrt{-1} = \pm i$。
這是線性代數應用最廣的領域之一，然而時間的關係我們只能介紹基本性質。

8.1 特徵值（eigenvalue）與特徵向量（eigenvector）

給定 $n \times n$ 方陣 A，在介紹矩陣轉換時談過，$A\mathbf{x}$ 稱爲 $\mathbf{x}$ 的映像（image），
且現在 A 是方陣，所以 $\mathbf{x}$ 與 $A\mathbf{x}$ 都是 n 維向量。
本章的基本問題是：有那些向量能夠使得 $A\mathbf{x}$ 與 $\mathbf{x}$ 是平行的？

定義 A 是 $n \times n$ 方陣。若在 R^n 之中存在非零向量使得 $A\mathbf{x} = \lambda\mathbf{x}$，
則稱純量 λ 是 A 的**特徵值**（eigenvalue），
滿足 $A\mathbf{x} = \lambda\mathbf{x}$ 的向量 $\mathbf{x}$ 稱爲**對應於 λ 的特徵向量**，
（eigenvector of A associated with eigenvalue λ），簡稱特徵向量。

重要觀念 若 $\mathbf{x} = \mathbf{0}$ 則當然 $A\mathbf{x} = \lambda\mathbf{x} = \mathbf{0}$，但是依定義特徵向量不能是零向量。

重要觀念 λ 可以是實數或者複數，特徵向量 $\mathbf{x}$ 的元素也可能是複數。

例 1 若 A 是單位矩陣 I_n，則只有唯一的特徵值 $\lambda = 1$，

且 R^n 之中任意的非零向量都是對應 $\lambda = 1$ 的特徵向量。（why?）

例 2 驗證上述的定義，其中 $A = \begin{bmatrix} 0 & \frac{1}{2} \\ \frac{1}{2} & 0 \end{bmatrix}$， $\mathbf{x}_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$，$\mathbf{x}_2 = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$。

以實數的特徵值配合圖形來理解 λ 的大小與正負值與特徵向量之間的關係：

重要觀念 同一個 λ 可以對應許多的 $\mathbf{x}$，因爲 $\mathbf{x}$ 的任意倍數 $r\mathbf{x}$ 都能滿足

$A(r\mathbf{x}) = r(A\mathbf{x}) = r(\lambda\mathbf{x}) = \lambda(r\mathbf{x})$，可見只要平行的向量都是同一個 λ 的特徵向量。

重要觀念 同一個 λ 也可以對應不同方向的特徵向量，**必需求解之後才知道**。

定義 同一個 λ 對應的特徵向量是 R^n 的子空間，稱爲**特徵空間**（eigenspace）。

例 3 驗證上述定義，其中 $A=\begin{bmatrix}0 & 0\\ 0 & 1\end{bmatrix}$，$\mathbf{x}_1=\begin{bmatrix}1\\ 0\end{bmatrix}$，$\mathbf{x}_2=\begin{bmatrix}0\\ 1\end{bmatrix}$。

重要觀念 雖然零向量不能當作特徵向量，但是由 **例 3** 可知 0 可以是特徵值。

例 4 我們先使用定義來計算 A 的特徵值與特徵向量，其中 $A=\begin{bmatrix}1 & 1\\ -2 & 4\end{bmatrix}$。

（後面會介紹稍微簡潔一點的方法）

將上一題的過程整理，我們有以下的定義、定理與演算方式。

定義 若 A 是方陣，則行列式

$$f(\lambda)=\det(\lambda I-A)=\begin{vmatrix}\lambda-a_{11} & -a_{12} & \cdots & -a_{1n}\\ -a_{21} & \lambda-a_{22} & \cdots & -a_{2n}\\ \vdots & \vdots & & \vdots\\ -a_{n1} & -a_{n2} & \cdots & \lambda-a_{nn}\end{vmatrix}$$

稱爲矩陣 A 的**特徵多項式**（characteristic polynomial of A），
（將 $f(\lambda)$ 這個行列式乘開之後會得到一個 λ 的多項式，且我們要找非明顯解）
方程式 $f(\lambda)=0$ 稱爲矩陣 A 的**特徵方程式**（characteristic equation of A）。

定理 A 的特徵值是特徵多項式的根（roots of charateristic equation）。

重要觀念 將行列式乘開時，從主對角線知道會是 λ 的 n 次多項式，
一般式可以寫成 $f(\lambda) = \lambda^n + c_1\lambda^{n-1} + c_2\lambda^{n-2} + \cdots + c_{n-1}\lambda + c_n$，
假如將 $\lambda = 0$ 代入特徵方程式會得出 $\det(-A) = c_n$，即 $c_n = (-1)^n \det(A)$，
因此得知：若 $\lambda = 0$ 是 $f(\lambda)$ 的根，則 $\det(A) = 0$，定理如下。

定理 $n \times n$ 方陣 A 是奇異矩陣若且唯若 0 是 A 的特徵值。

下一個定理可以當作手算時的技巧，但是只能應用在實數根。別忘記，有可能有虛數根。

定理 若 $f(\lambda)$ 之中的係數 $c_1, c_2, \ldots, c_n$ 都是整數，則 $f(\lambda)$ 的實數根都是整數。

（所以先試試 c_n 能夠除盡的整數能不能當作多項式的根）

演算法 給定 A 求特徵值與特徵向量的方法

步驟 **1** 求 $f(\lambda)=\det(\lambda I-A)=0$ 的根，它們就是特徵值。

步驟 **2** 對每一個特徵值，求出齊次方程組 $(\lambda I-A)\mathbf{x}=\mathbf{0}$ 的非明顯解，以向量表達（可以乘除一下將係數變得簡潔）就是對應這個特徵值的特徵向量。

觀念 大一點的矩陣要手算特徵值與特徵向量是很困難的，必需借助軟體。
但是各位在使用軟體時要注意，通常商用軟體是用數值逼近的方法來算特徵向量，
因此可能會跟各位手算的結果有點不同。
設計一般化的 eigenvalue 與 eigenvector 的演算程式也是一門學問。

$\boxed{\textbf{例 5 與 6}}$ 求 A 的特徵值與特徵向量，其中 $A = \begin{bmatrix} 1 & 2 & -1 \\ 1 & 0 & 1 \\ 4 & -4 & 5 \end{bmatrix}$。

$\boxed{\textbf{例 7}}$ 求 A 的特徵值與特徵向量，其中 $A = \begin{bmatrix} 0 & 0 & 3 \\ 1 & 0 & -1 \\ 0 & 1 & 3 \end{bmatrix}$。

以下這個定理彙整了本課程關於非奇異矩的的性質。

$\boxed{\text{定理}}$ 若 A 是 $n \times n$ 的方陣，則以下的敘述是等價（equivalent）的：

1. A 是非奇異矩陣（即 A^{-1} 存在）
2. $A\mathbf{x} = \mathbf{0}$ 只有明顯解 $\mathbf{x} = \mathbf{0}$
3. A 列等價於 I
4. $A\mathbf{x} = \mathbf{b}$ 有唯一解 $\mathbf{x} = A^{-1}\mathbf{b}$
5. $\det(A) \neq 0$
6. A 的秩是 n（即 $\text{rank}(A) = n$）
7. A 的零核空間維數是 0
8. A 的 n 個列是 R^n 之中線性獨立的向量
9. A 的 n 個行是 R^n 之中線性獨立的向量
10. 0 不是 A 的特徵值

8.2 對角線化（diagonalization）

從 8.1 小節可以發現，矩陣的特徵值與特徵向量並不容易算出來。
本小節首先說明，有一類的矩陣，存在與它特徵值相同的其它矩陣，稱爲相似；
接下來，相似矩陣之中有一種比較容易找出特徵值的，它是對角線矩陣。
爲了簡化所有的說明，本小節與下一小節的矩陣與特徵值都是實數。

定義 給定 $n \times n$ 方陣 A。若存在一個非奇異矩陣 P 使得 $B = P^{-1}AP$，則稱 A 與 B 是**相似**（similar）矩陣。

定理（相似矩陣有遞移性）若 A 與 B 相似且 B 與 C 相似，則 A 與 C 相似。

例 1 驗證上述定義，其中（這些矩陣就是在 8.1 節 **例 4** 的矩陣）

$$A = \begin{bmatrix} 1 & 1 \\ -2 & 4 \end{bmatrix}, \quad P = \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix}, \quad P^{-1} = \begin{bmatrix} 2 & -1 \\ -1 & 1 \end{bmatrix}$$

定義 若 A 相似於一個對角線矩陣，則稱 A 是**可對角線化的矩陣**。

例 2 在 **例 1** 之中的矩陣 A 是可對角線化的矩陣。

注意 不要跟第一章的性質混淆：A 列運算之後與 I 列等價，則 A^{-1} 存在。

定理 相似矩陣的特徵值相同。

重要觀念 對角線矩陣的特徵值就是主對角線上的元素。
（因爲行列式乘開之後的特徵方程式是 $f(\lambda) = (\lambda - a_{11})(\lambda - a_{22}) \cdots (\lambda - a_{nn}) = 0$）

定理 $n \times n$ 方陣 A 可對角線化若且唯若 A 有 n 個線性獨立的特徵向量，
此時存在 P 使得 $P^{-1}AP = D$，其中 D 是對角線矩陣，主對角線元素是特徵值；
且 P 是由 A 的 n 個線性獨立的特徵向量當作行向量所組成。
（注意到 P 的行向量的順序與 D 的主對角線元素必定要一一對應）

例 3 續 **例 1**

例 4 請問 $A=\begin{bmatrix}1 & 1\\ 0 & 1\end{bmatrix}$ 能不能對角線化？

重要觀念（非常重要，常常有同學們觀念不正確而弄錯）

並不是 λ 有重根，A 就不能對角線化，

因爲定義是說有 n 的獨立的特徵向量就可以對角化，

λ 有重根還是有可能有 n 個獨立特徵向量，要計算才知道。

定理 若 $n\times n$ 方陣 A 的特徵多項式的根是 n 個不同數值的實數，
則 A 可以對角線化（此時 A 一定有 n 個線性獨立的特徵向量）。

重要觀念（容易出錯）這個定理並不是若且唯若的關係，

它是說有 n 個實數特徵值的方陣是好的性質，必然可以對角線化，

但是假如沒有 n 個實數特徵值，還是有可能可以對角線化的，算過才知道。

演算法 對角線化問題的演算過程

步驟 **1** 求出 $f(\lambda) = \det(\lambda I - A) = 0$ 的根

步驟 **2** 假如有重根，先計算這些重根有沒有足夠數量的線性獨立特徵向量

（例如 $(\lambda - 3)^2$ 則 $\lambda = 3$ 時必需要有兩個獨立的特徵向量）

假如任一個重根的特徵向量數目不足，則停止演算，A 不能對角線化；

假如可以求出 n 個獨立的特徵向量，則 A 可以對角線化，進入步驟 3

步驟 **3** 將特徵向量以行向量的方式寫成矩陣 P，對應的特徵值寫成對角線矩陣 D；

假如題目有問 P^{-1}，再另外計算。

例 5 請問 A 能不能對角線化？假如可以，求出 P、D、P^{-1} 使得 $P^{-1}AP = D$，其中

$$A = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{bmatrix}$$

例 6 請問 A 能不能對角線化？假如可以，求出 P、D、P^{-1} 使得 $P^{-1}AP = D$，其中

$$A = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix}$$

8.3 對稱矩陣的對角線化

從 8.2 小節，雖然我們知道如何求出 P 與 D 使得 $P^{-1}AP = D$，
但是要先判斷能不能對角線化，另外還有 P^{-1} 要算，並不容易計算。
但是有一種矩陣可以比較快而且容易知道對角線化所有的結果，對稱矩陣。

首先回憶一下，若 $A^T = A$，這種方陣稱爲對稱矩陣，它有許多非常好的性質，首先：
（對稱矩陣在工程與管理的應用問題之中，並不是非常罕見的矩陣）

定理 對稱矩陣的特徵方程式的根一定都是實數。

定理 對稱矩陣之中對應不同特徵值的特徵向量一定互相垂直。

例 1 驗證上述定理，$A = \begin{bmatrix} 0 & 0 & -2 \\ 0 & -2 & 0 \\ -2 & 0 & 3 \end{bmatrix}$

第六章學過，把垂直向量正規化，互相垂直的單位向量（orthonormal）有更好的性質。

假設給定對稱矩陣 A，已經求出 n 個互相垂直的特徵向量。將特徵向量正規化之後，
（注意，若 $\mathbf{x}$ 是特徵向量則 $r\mathbf{x}$ 也是同一個方向的特徵向量，性質並不影響）
把這些正規化正交向量當成行向量寫成矩陣 P，則（注意 $\mathbf{x}_i$ 是行向量，$\mathbf{x}_i^T$ 是列向量）

$$P = \begin{bmatrix} \mathbf{x}_1 & \mathbf{x}_2 & \cdots & \mathbf{x}_n \end{bmatrix}$$ ，因此

$$P^T = \begin{bmatrix} \mathbf{x}_1^T \\ \mathbf{x}_2^T \\ \vdots \\ \mathbf{x}_n^T \end{bmatrix}$$

各位可以很快檢驗得知，$P^TP = I$ 且 $P^T = P^{-1}$。
因此假如要算這種矩陣的反矩陣實在太方便了，轉置就行了。

定義 若 $A^{-1}=A^T$ 則稱 A 是**正交矩陣**（orthogonal matrix），此時 $A^TA=I$。

注意 正交矩陣（orthogonal matrix）這個名詞初學者容易誤會，
事實上裡面的向量不止是正交，而且是正規化正交向量。

重要觀念 要驗證是不是正交矩陣，檢查 $A^{-1}=A^T$ 比較慢，算 $A^TA=I$ 比較快。

例 2 驗證 $A=\begin{bmatrix}\frac{2}{3} & -\frac{2}{3} & \frac{1}{3}\\ \frac{2}{3} & \frac{1}{3} & -\frac{2}{3}\\ \frac{1}{3} & \frac{2}{3} & \frac{2}{3}\end{bmatrix}$ 是正交矩陣。

例 3 續 **例 1**，求出 P、D、P^{-1} 使得 $P^{-1}AP=D$。

定理 A 是正交矩陣若且唯若 A 的行（或列）是 R^n 的正規化正交基底。

最後這個性質告訴我們，**對稱矩陣不管 λ 有沒有重根**，**都可以對角線化**。

定理 若 A 是 $n \times n$ 的對稱矩陣，則存在正交矩陣 P 使得

$P^{-1}AP = D$。（當然手算的時候 $P^{-1} = P^T$ 比較容易）

例 4 求出 P、D、P^{-1} 使得 $P^{-1}AP = D$，其中 $A = \begin{bmatrix} 0 & 2 & 2 \\ 2 & 0 & 2 \\ 2 & 2 & 0 \end{bmatrix}$