單元 37: 線性映射，特徵向量與特徵值
(課本 §9.3)

爲方便起見, 以 2×2 矩陣說明, 但此節中的結果可推廣至 $n \times n$ 矩陣

定義. 設給定一 2×2 矩陣 A, 稱

$$\text{映射 (map): } \mathbf{x} \to A\mathbf{x}$$

爲一線性映射 (linear map), 其中 $\mathbf{x}$ 爲一 2×1 行向量, 且得 $A\mathbf{x}$ 也爲一 2×1 行向量. 稱爲線性乃因爲此映射使得下列二式成立:

(i) 對任意二向量 $\mathbf{x}$ 與 $\mathbf{y}$,

$$A(\mathbf{x}+\mathbf{y}) = A\mathbf{x} + A\mathbf{y} \text{ (根據分配律)}$$

(ii) 對任一純量 λ,

$$A(\lambda\mathbf{x}) = \lambda(A\mathbf{x})$$

舉例, A 爲 §9.2 中的 Leslie 矩陣 L, $\mathbf{x}$ 爲在 0 時的族群大小, $A\mathbf{x}$ 爲在 1 時的族群大小.

一. 向量的圖形表示法

直角坐標系統下, 行向量

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

相當於 (對應於) 平面上的一個以原點爲起點的有向線段, 如圖示. 一些相關的定義如下:

(1) $\mathbf{x}$ 的長度

$$|\mathbf{x}| = \sqrt{x_1^2 + x_2^2}$$

(2) 與橫軸形成的夾角的角度 $\alpha \in [0, 2\pi)$ 滿足

$$\tan\alpha = \frac{x_2}{x_1}$$

(3) 極坐標系統 (polar coordinate system): 令 $r = |\mathbf{x}|$, 則

$$x_1 = r\cos\alpha,\ x_2 = r\sin\alpha$$

所以, 行向量 $\mathbf{x}$ 的極坐標表示法爲

$$\mathbf{x} = \begin{bmatrix} r\cos\alpha \\ r\sin\alpha \end{bmatrix}$$

(4) 加法與純量乘法: 令 $\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$, $\mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}$ 且 a 爲一純量 (scalar), 則

$$\mathbf{x} + \mathbf{y} = \begin{bmatrix} x_1 + y_1 \\ x_2 + y_2 \end{bmatrix}$$

所對應的平面上的運算如圖示. 所以, 向量加法律又稱作平行四邊形法則 (parallelogram law).

又

$$a\mathbf{x} = \begin{bmatrix} ax_1 \\ ax_2 \end{bmatrix}$$

所對應的平面上的運算如圖示. 顯示出, 當 $a > 0$ 時, $a\mathbf{x}$ 表示將 $\mathbf{x}$ 放大 a 倍; 當 $a < 0$ 時, $a\mathbf{x}$ 表示將 $\mathbf{x}$ 反向放大 a 倍.

(5) 單位映射 (identity map):

$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \to \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

所對應的平面效應如圖示.

(6) 分量展縮映射 (coordinate stretching or contracting map):

$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \to \begin{bmatrix} a & 0 \\ 0 & b \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} ax_1 \\ bx_2 \end{bmatrix}$$

如,

$$\begin{bmatrix} 1 \\ 2 \end{bmatrix} \to \begin{bmatrix} 2 & 0 \\ 0 & -\frac{1}{2} \end{bmatrix} \begin{bmatrix} 1 \\ 2 \end{bmatrix} = \begin{bmatrix} 2 \\ -1 \end{bmatrix}$$

乃相當於將第一個分量同向伸展 2 倍, 同時將第二個分量反向壓縮 $\frac{1}{2}$ 倍. 所對應的平面向量如圖示. 又

$$\begin{bmatrix} 2 \\ 1 \end{bmatrix} \to \begin{bmatrix} -\frac{1}{2} & 0 \\ 0 & 2 \end{bmatrix} \begin{bmatrix} 2 \\ 1 \end{bmatrix} = \begin{bmatrix} -1 \\ 2 \end{bmatrix}$$

乃相當於將第一個分量反向壓縮 $\frac{1}{2}$ 倍, 且將第二個分量同向伸展 2 倍. 所對應的平面向量如圖示.

(7) 旋轉映射 (rotation map):

$$\mathbf{x} = \begin{bmatrix} r\cos\alpha \\ r\sin\alpha \end{bmatrix} \to \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \cdot \begin{bmatrix} r\cos\alpha \\ r\sin\alpha \end{bmatrix}$$

將箭頭右邊的矩陣乘開後, 得

$$\begin{aligned} \begin{bmatrix} r\cos\alpha \\ r\sin\alpha \end{bmatrix} & \\ \to & \begin{bmatrix} r(\cos\theta\cos\alpha - \sin\theta\sin\alpha) \\ r(\sin\theta\cos\alpha + \cos\theta\sin\alpha) \end{bmatrix} \\ = & \begin{bmatrix} r\cos(\theta+\alpha) \\ r\sin(\theta+\alpha) \end{bmatrix} \end{aligned}$$

也就是說, 將與橫軸的夾角由 α 多旋轉 θ, 至 $\alpha+\theta$. 對應的平面向量如圖示.

一般的矩陣 A 通常就沒有以上簡單的幾何意義上的解釋, 如, 令

$$A=\begin{bmatrix} 1 & 2 \\ 3 & 2 \end{bmatrix}$$

若

$$\mathbf{x}=\begin{bmatrix} 3 \\ -1 \end{bmatrix}$$

則

$$A\mathbf{x}=\begin{bmatrix} 1 \\ 7 \end{bmatrix}$$

且所對應的平面向量如圖示.

若

$$\mathbf{x}=\begin{bmatrix} 2 \\ 3 \end{bmatrix}$$

則

$$\begin{aligned} A\mathbf{x} &= \begin{bmatrix} 1 & 2 \\ 3 & 2 \end{bmatrix}\begin{bmatrix} 2 \\ 3 \end{bmatrix} \\ &= \begin{bmatrix} 8 \\ 12 \end{bmatrix}=4\begin{bmatrix} 2 \\ 3 \end{bmatrix} \end{aligned}$$

亦即, 將 $\mathbf{x}$ 的分量都放大 4 倍. 所對應的平面向量如圖示.

因為將 A 作用在不同的兩個行向量上, 會得到兩個不同的效應, 所以 A 既不是展縮亦不是旋轉.

但有興趣的是: 給一線性映射

$$\mathbf{x} \to A\mathbf{x}$$

而試圖找出 $\mathbf{x}$ 與 λ 使得

$$A\mathbf{x} = \lambda\mathbf{x}$$

如上例中的

$$\mathbf{x} = \begin{bmatrix} 2 \\ 3 \end{bmatrix} \text{ 與 } \lambda = 4$$

稱此種展縮因子 λ 為特徵值 (eigenvalue), 行向量 $\mathbf{x}$ 為特徵向量 (eigenvector).

二. 特徵值與特徵向量

定義. 設 A 為一方陣. 若一非零向量 $\mathbf{x}$ (亦即, $\mathbf{x} \neq \mathbf{0}$) 使得

$$A\mathbf{x} = \lambda\mathbf{x}$$

則稱 $\mathbf{x}$ 為 A 的特徵向量, λ 為 A 的特徵值.

註 1. 幾何意義: 若 $\mathbf{x}$ 為特徵向量且 λ 為一實數特徵值, 則 $A\mathbf{x} = \lambda\mathbf{x}$ 是落在通過 $\mathbf{x}$ 的直線上. 此乃因為

$$\lambda\mathbf{x} \text{ 的斜率 } = \frac{\lambda x_2}{\lambda x_1} = \frac{x_2}{x_1} = \mathbf{x} \text{ 的斜率}$$

若斜率存在的話, 否則均在縱軸上. 此外, 對任意在通過 $\mathbf{x}$ 的直線上的向量 $\mathbf{u}$, 都會有一常數 c 使得

$$\mathbf{u} = c\mathbf{x}$$

此乃因為 $\mathbf{u}$ 與 $\mathbf{x}$ 要有相同的斜率, 若斜率存在的話 (或 $\mathbf{u}$ 與 $\mathbf{x}$ 均在縱軸上), 故 $\mathbf{u}$ 是 $\mathbf{x}$ 的 某個 c 倍. 由此可得出

$$A\mathbf{u} = Ac\mathbf{x} = cA\mathbf{x} = c\lambda\mathbf{x} = \lambda c\mathbf{x} = \lambda\mathbf{u}$$

亦在同一直線上且 $\mathbf{u}$ 也是 A 的一個特徵向量. 因此, 可推導出特徵向量不是唯一的且差異僅是一個常數倍數 (亦即, 存在一常數 c 使得 $\mathbf{u} = c\mathbf{x}$), 如圖示.

註 2. 二階方陣有 2 個特徵值 (相同或相異). 本課程僅探討 2 個相異的情況.

註 3. 如何求? 以例說明如下.

例 1. 試求

$$A = \begin{bmatrix} 1 & 2 \\ 3 & 2 \end{bmatrix}$$

的特徵值與特徵向量.

<解> 根據特徵值與特徵向量的定義, 需找 λ 與 $\mathbf{x} \neq \mathbf{0}$ 使得

$$A\mathbf{x} = \lambda\mathbf{x}$$

此乃相當於

$$A\mathbf{x} - \lambda\mathbf{x} = \mathbf{0}$$

亦相當於

$$A\mathbf{x} - \lambda I_2\mathbf{x} = \mathbf{0}$$

根據分配律, 由上式得

$$(A - \lambda I)\mathbf{x} = \mathbf{0}$$

其中, 爲簡單計, 令 $I = I_2$. 因爲 $\mathbf{x} \neq \mathbf{0}$, 上式相當於 $A - \lambda I$ 爲奇異的 (singular), 也就是說,

$$\det(A - \lambda I) = 0$$

因此, 求特徵值與特徵向量的方法爲

(1) 找 λ 使得 $\det(A - \lambda I) = 0$:

首先,

$$\begin{aligned} A-\lambda I &= \begin{bmatrix} 1 & 2 \\ 3 & 2 \end{bmatrix} - \begin{bmatrix} \lambda & 0 \\ 0 & \lambda \end{bmatrix} \\ &= \begin{bmatrix} 1-\lambda & 2 \\ 3 & 2-\lambda \end{bmatrix} \end{aligned}$$

接著, 根據二階矩陣的行列式公式, 得

$$\begin{aligned} \det(A-\lambda I) &= (1-\lambda)(2-\lambda)-2\cdot 3 \\ &= 2-3\lambda+\lambda^2-6 \\ &= \lambda^2-3\lambda-4 \\ &= (\lambda-4)(\lambda+1) \end{aligned}$$

所以, 由

$$\det(A-\lambda I)=0$$

可得出兩個特徵值

$$\lambda_1=-1 \text{ 與 } \lambda_2=4$$

(2) 分別求各特徵值所對應的特徵向量:

$\lambda_1=-1$: 找 $\mathbf{x}=\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$ 使得

$$A\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}=(-1)\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

此乃相當於解線性方程式系統

$$\begin{aligned} x_1 + 2x_2 &= -x_1 \\ 3x_1 + 2x_2 &= -x_2 \end{aligned}$$

移項整理, 得

$$\begin{aligned} 2x_1 + 2x_2 &= 0 \\ 3x_1 + 3x_2 &= 0 \end{aligned}$$

亦相當於

$$x_1 + x_2 = 0$$

即,

$$x_2 = -x_1 \tag{1}$$

所以,

$$\mathbf{x} = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$$

是 $\lambda_1 = -1$ 所相對應的一特徵向量. 事實上, 由 (1) 式知, 對任意的 $a \neq 0$,

$$a \begin{bmatrix} 1 \\ -1 \end{bmatrix}$$

都是 $\lambda_1 = -1$ 所對應的特徵向量. 因此, 有無窮多個特徵向量且均在通過 $\begin{bmatrix} 1 \\ -1 \end{bmatrix}$ 的同一直線 (亦即, $x_1 + x_2 = 0$) 上, 如圖示.

$\lambda_2 = 4$: 解 $A\mathbf{x} = 4\mathbf{x}$ 相當於解線性系統

$$\begin{aligned} x_1 + 2x_2 &= 4x_1 \\ 3x_1 + 2x_2 &= 4x_2 \end{aligned}$$

亦相當於解

$$\begin{aligned} -3x_1 + 2x_2 &= 0 \\ 3x_1 - 2x_2 &= 0 \end{aligned}$$

由此得

$$3x_1 - 2x_2 = 0$$

亦即,

$$x_2 = \frac{3}{2}x_1$$

因此,

$$\mathbf{x} = \begin{bmatrix} 2 \\ 3 \end{bmatrix}$$

是 $\lambda_2 = 4$ 所對應的一特徵向量. 事實上, 所有在直線 $3x_1 - 2x_2 = 0$ 上的向量均是 $\lambda_2 = 4$ 所對應的特徵向量, 如圖示.

註. 直線 $l_1 : x_1 + x_2 = 0$ 與 $l_2 : 3x_1 - 2x_2 = 0$ 在映射

$$\mathbf{x} \to A\mathbf{x}$$

的作用下, 都是不變的 (invariant), 也就是說, 對任意的向量

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \in l_1 \text{ 或 } l_2$$

$A\mathbf{x}$ 也會在 l_1 或 l_2 上. 此乃因爲若 $\mathbf{x} \in l_1$, 則

$$x_1 + x_2 = 0$$

亦即, $\mathbf{x}$ 爲 λ_1 所對應的一特徵向量. 所以,

$$A\mathbf{x} = \lambda_1 \mathbf{x} = \begin{bmatrix} \lambda_1 x_1 \\ \lambda_1 x_2 \end{bmatrix}$$

而且

$$\lambda_1 x_1 + \lambda_1 x_2 = \lambda_1(x_1 + x_2) = 0$$

因此, $A\mathbf{x} \in l_1$. 同理可證 l_2 的情況.

例 2. 令

$$A = \begin{bmatrix} \cos 30^\circ & -\sin 30^\circ \\ \sin 30^\circ & \cos 30^\circ \end{bmatrix}$$

試描述 A 作用在

$$\mathbf{x} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

上的結果以及求 A 的特徵值.

<解> (a) 根據旋轉映射的定義, 矩陣 A 為旋轉 30^o 的映射. 所以,

$$\begin{aligned} A\begin{bmatrix} 1 \\ 0 \end{bmatrix} &= A\begin{bmatrix} \cos 0^\circ \\ \sin 0^\circ \end{bmatrix} \\ &= \begin{bmatrix} \cos 30^\circ \\ \sin 30^\circ \end{bmatrix} \\ &= \begin{bmatrix} \frac{\sqrt{3}}{2} \\ \frac{1}{2} \end{bmatrix} \\ &= \text{將} \begin{bmatrix} 1 \\ 0 \end{bmatrix} \text{逆時針轉 } 30^\circ \end{aligned}$$

如圖示.

(b) 根據前例所得出的求特徵值的方法, 需要求 λ 使得

$$\det(A-\lambda I)=0$$

此乃相當於

$$\det\begin{bmatrix} \frac{\sqrt{3}}{2}-\lambda & -\frac{1}{2} \\ \frac{1}{2} & \frac{\sqrt{3}}{2}-\lambda \end{bmatrix}=0$$

根據行列式的公式, 得

$$\left(\frac{\sqrt{3}}{2}-\lambda\right)^2+\frac{1}{4}=0$$

將上式展開後, 得

$$\frac{3}{4} - \sqrt{3}\lambda + \lambda^2 + \frac{1}{4} = 0$$

亦相當於

$$\lambda^2 - \sqrt{3}\lambda + 1 = 0$$

因此, 根據二次式的公式解, 二特徵值爲

$$\begin{aligned} \lambda_{1,2} &= \frac{\sqrt{3} \pm \sqrt{3-4}}{2} \\ &= \frac{1}{2}(\sqrt{3} \pm i),\ i = \sqrt{-1} \end{aligned}$$

故, 特徵值可爲複數. 對應的特徵向量的求法相當複雜, 超出本書範圍, 故略.

特徵值的一個應用: 判斷微分方程式系統的平衡點的穩定性 (第 11 章的內容) 時, 需要考慮特徵值實部的正負性. 而特徵值實部的正負性可由下列二量決定: 設矩陣

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

(1) $\det A = ad - bc$

(2) tr$A \overset{\text{def}}{=} a + d$, 稱作 trace of A, 爲對角線上的元素和.

註. 使用此二量的好處乃在於, 不需要求出特徵值, 僅由矩陣 A 就可判斷出特徵值實部的正負性. 爲何如此? 怎麼用? 根據下述的二結論, 可得知.

(i) 設矩陣

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

的特徵值爲 λ_1 與 λ_2, 則

$$\text{tr}A = \lambda_1 + \lambda_2$$

且

$$\det A = \lambda_1 \lambda_2$$

<證> 首先, 根據特徵值的定義所導出的求法, λ_1 與 λ_2 是

$$\begin{aligned} \det(A - \lambda I) &= \det \begin{bmatrix} a - \lambda & b \\ c & d - \lambda \end{bmatrix} \\ &= (a - \lambda)(d - \lambda) - bc \\ &= \lambda^2 - (a + d)\lambda + (ad - bc) \\ &= 0 \end{aligned}$$

的 2 根. 因此, 二根之和

$$\lambda_1 + \lambda_2 = a + d = \mathrm{tr}A$$

且二根之積

$$\lambda_1\lambda_2 = ad - bc = \det A$$

(ii) 設 2 階方陣 A 的特徵值爲 λ_1 與 λ_2, 則 λ_1 與 λ_2 的實部均爲負若且爲若

$$\mathrm{tr}A < 0 \text{ 且 } \det A > 0$$

<證> 情況 1. λ_1, λ_2 均爲實數.

"$\Rightarrow$": 假設 $\lambda_1 < 0$, $\lambda_2 < 0$. 則明顯地,

$$\mathrm{tr}A = \lambda_1 + \lambda_2 = (-) + (-) < 0$$

且

$$\det A = \lambda_1\lambda_2 = (-)(-) > 0$$

"$\Leftarrow$": 由假設

$$\det A = \lambda_1\lambda_2 > 0$$

可得 λ_1 與 λ_2 同號. 再根據另一假設

$$\mathrm{tr}A = \lambda_1 + \lambda_2 < 0$$

可導出 λ_1 與 λ_2 均爲負.

情況 2. λ_1 與 λ_2 爲二共軛複數, 亦即,

$$\lambda_1 = x + iy,\ \lambda_2 = x - iy$$

"$\Rightarrow$": 設實部 $x < 0$, 則

$$\mathrm{tr}A = (x + iy) + (x - iy) = 2x < 0$$

且

$$\begin{aligned}\det A &= (x + iy)(x - iy)\\ &= x^2 - i^2y^2\\ &= x^2 + y^2 > 0\end{aligned}$$

"$\Leftarrow$": 由假設

$$\mathrm{tr}A = 2x < 0$$

可導出 $x < 0$. 所以, λ_1 與 λ_2 的實部 x 均爲負.

例 3. 設

$$A = \begin{bmatrix} -1 & -3 \\ 1 & -2 \end{bmatrix}$$

試證 A 的特徵值的實部均爲負.

<證> 因爲

$$\text{tr}A = (-1) + (-2) = -3 < 0$$

且

$$\det A = (-1)(-2) - (-3)(1) = 5 > 0$$

故可導出特徵值的實部均爲負.

註. 不需要實際求出特徵值, 即可判斷出特徵值實部的正負性.

驗證: 求 λ_1 與 λ_2 相當於解

$$\det(A - \lambda I) = 0$$

亦即,

$$\det \begin{bmatrix} -1-\lambda & -3 \\ 1 & -2-\lambda \end{bmatrix} = 0$$

接者根據行列式的公式, 上式相當於

$$\begin{aligned} \det(A - \lambda I) &= (-1-\lambda)(-2-\lambda) + 3 \\ &= \lambda^2 + 3\lambda + 5 = 0 \end{aligned}$$

所以, 根據二次式的公式解,

$$\lambda_{1,2} = \frac{-3 \pm \sqrt{9-20}}{2} = -\frac{3}{2} \pm \frac{1}{2}\sqrt{11}i$$

因此, 二特徵值的實部均爲 $-\frac{3}{2} < 0$.

三. 迭代映射 (iterated maps)

考慮連續重複的映射

$$\begin{aligned} \mathbf{x} \;\to\; & A\mathbf{x} \to A(A\mathbf{x}) = A^2\mathbf{x} \to \\ & A(A^2\mathbf{x}) = A^3\mathbf{x} \to \cdots \to A^n\mathbf{x} \to \cdots \end{aligned}$$

其中的 $A^n\mathbf{x}$ 稱作第 n 個迭代映射:

$$\mathbf{x} \to A^n\mathbf{x}$$

此迭代映射亦相當於

(i) 先將 A 自乘 n 次得 A^n (一個繁複計算的過程).

(ii) 算 $A^n\mathbf{x}$

問. 有無簡單方法? 有, 如下所示.

定義. 二非零向量 $\mathbf{x}_1$ 與 $\mathbf{x}_2$ 稱作線性獨立 (linearly independent), 若不存在 a 使得

$$\mathbf{x}_1 = a\mathbf{x}_2$$

(亦即, 線性獨立 $\Leftrightarrow$ 不在同一直線上).

兩個事實:

(i) 設 A 爲一二階方陣且有特徵值 λ_1 與 λ_2 以及對應的特徵向量 $\mathbf{u}_1$ 與 $\mathbf{u}_2$. 若 $\lambda_1 \neq \lambda_2$ 則 $\mathbf{u}_1$ 與 $\mathbf{u}_2$ 爲線性獨立.

<證> 反證法: 設對於某個 a,

$$\mathbf{u}_2 = a\mathbf{u}_1$$

則根據特徵向量的性質及如上的假設,

$$A\mathbf{u}_2 = \lambda_2\mathbf{u}_2 = \lambda_2 a\mathbf{u}_1 = a\lambda_2\mathbf{u}_1$$

同理,

$$A\mathbf{u}_2 = Aa\mathbf{u}_1 = aA\mathbf{u}_1 = a\lambda_1\mathbf{u}_1$$

因此, 比較以上二式, 得

$$\lambda_1 = \lambda_2$$

此乃與命題的假設

$$\lambda_1 \neq \lambda_2$$

矛盾. 故, 不存在此種 a, 也就是說, 此二特徵向量爲線性獨立.

(ii) 若 $\mathbf{u}_1$ 與 $\mathbf{u}_2$ 線性獨立, 則對任一 2×1 行向量 $\mathbf{x}$ 都存在唯一的 a_1 與 a_2 使得

$$\mathbf{x} = a_1\mathbf{u}_1 + a_2\mathbf{u}_2$$

亦即, 任意的 $\mathbf{x}$ 都可寫成 $\mathbf{u}_1$ 與 $\mathbf{u}_2$ 的線性組合.

結論: 若 $\mathbf{x} = a_1\mathbf{u}_1 + a_2\mathbf{u}_2$, 則

$$\begin{aligned} A\mathbf{x} &= A(a_1\mathbf{u}_1 + a_2\mathbf{u}_2) \\ &= a_1A\mathbf{u}_1 + a_2A\mathbf{u}_2 \\ &= a_1\lambda_1\mathbf{u}_1 + a_2\lambda_2\mathbf{u}_2 \end{aligned}$$

$$\begin{aligned} A^2\mathbf{x} &= A(A\mathbf{x}) \\ &= A(a_1\lambda_1\mathbf{u}_1 + a_2\lambda_2\mathbf{u}_2) \\ &= a_1\lambda_1A\mathbf{u}_1 + a_2\lambda_2A\mathbf{u}_2 \\ &= a_1\lambda_1\lambda_1\mathbf{u}_1 + a_2\lambda_2\lambda_2\mathbf{u}_2 \\ &= a_1\lambda_1^2\mathbf{u}_1 + a_2\lambda_2^2\mathbf{u}_2 \end{aligned}$$

$$\vdots$$

$$A^n\mathbf{x} = a_1\lambda_1^n\mathbf{u}_1 + a_2\lambda_2^n\mathbf{u}_2 \tag{2}$$

也就是說, 不需要計算 A^n, 僅需計算 λ_1^n 與 λ_2^n 即可由 (2) 式求出 $A^n\mathbf{x}$.

例 4. 令

$$A = \begin{bmatrix} 1 & 2 \\ 3 & 2 \end{bmatrix}$$

且

$$\mathbf{x} = \begin{bmatrix} 4 \\ 1 \end{bmatrix}$$

試求 $A^{10}\mathbf{x}$.

<解> (i) 找特徵值 λ_1 與 λ_2 以及對應的特徵向量 $\mathbf{u}_1$ 與 $\mathbf{u}_2$: 根據特徵值的定義, 需要解

$$\begin{aligned} \det(A - \lambda I) &= \det \begin{bmatrix} 1-\lambda & 2 \\ 3 & 2-\lambda \end{bmatrix} \\ &= (1-\lambda)(2-\lambda) - 6 = 0 \end{aligned}$$

亦相當於解

$$\lambda^2 - 3\lambda - 4 = (\lambda - 4)(\lambda + 1) = 0$$

故, 得特徵值

$$\lambda_1 = -1 \text{ 與 } \lambda_2 = 4$$

接著, 求各特徵值所對應的特徵向量如下.

$\lambda_1 = -1$: 根據特徵向量的定義, 需要求

$$A \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = -1 \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

將等號左邊乘開後, 此乃相當於解線性方程式系統

$$\begin{aligned} x_1 + 2x_2 &= -x_1 \\ 3x_1 + 2x_2 &= -x_2 \end{aligned}$$

亦相當於解

$$\begin{aligned} 2x_1 + 2x_2 &= 0 \\ 3x_1 + 3x_2 &= 0 \end{aligned}$$

由此得

$$x_1 + x_2 = 0$$

所以, 可取

$$\mathbf{u}_1 = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$$

$\lambda_2 = 4$: 需要求

$$A \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = 4 \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

亦即, 解

$$\begin{aligned} x_1 + 2x_2 &= 4x_1 \\ 3x_1 + 2x_2 &= 4x_2 \end{aligned}$$

亦相當於解

$$\begin{aligned} -3x_1 + 2x_2 &= 0 \\ 3x_1 - 2x_2 &= 0 \end{aligned}$$

由此得

$$3x_1 - 2x_2 = 0$$

因此, 可取

$$\mathbf{u}_2 = \begin{bmatrix} 2 \\ 3 \end{bmatrix}$$

(ii) 找

$$\mathbf{x} = \begin{bmatrix} 4 \\ 1 \end{bmatrix}$$

的線性組合表示法: 找 a_1 與 a_2 使得

$$a_1 \begin{bmatrix} 1 \\ -1 \end{bmatrix} + a_2 \begin{bmatrix} 2 \\ 3 \end{bmatrix} = \begin{bmatrix} 4 \\ 1 \end{bmatrix}$$

此乃相當於解線性方程式系統

$$\begin{aligned} a_1 + 2a_2 &= 4 \\ -a_1 + 3a_2 &= 1 \end{aligned}$$

將上兩式相加後, 得

$$5a_2 = 5$$

所以,

$$a_2 = 1 \text{ 以及 } a_1 = 2$$

因此,

$$\begin{bmatrix} 4 \\ 1 \end{bmatrix} = 2\begin{bmatrix} 1 \\ -1 \end{bmatrix} + 1\begin{bmatrix} 2 \\ 3 \end{bmatrix}$$

(iii) 最後, 根據 (2) 式,

$$\begin{aligned} A^{10}\mathbf{x} &= a_1\lambda_1^{10}\mathbf{u}_1 + a_2\lambda_2^{10}\mathbf{u}_2 \\ &= 2\cdot(-1)^{10}\begin{bmatrix} 1 \\ -1 \end{bmatrix} + 1\cdot(4)^{10}\begin{bmatrix} 2 \\ 3 \end{bmatrix} \\ &= \begin{bmatrix} 2 \\ -2 \end{bmatrix} + (1024)(1024)\begin{bmatrix} 2 \\ 3 \end{bmatrix} \\ &= \begin{bmatrix} 2097154 \\ 3145726 \end{bmatrix} \end{aligned}$$

Leslie 矩陣的再探討:

一個事實: 設一 2×2 Leslie 矩陣有 2 個特徵值 λ_1 與 λ_2, 則

(i) 大的特徵值 = 族群成長率

(ii) 大的特徵值所對應的特徵向量 = 穩定年齡層分布

以例說明如下:

在 §9.2 例 8 中, 若 Leslie 矩陣

$$L = \begin{bmatrix} 1.5 & 2 \\ 0.08 & 0 \end{bmatrix}$$

則以

$$N(0) = \begin{bmatrix} 100 \\ 100 \end{bmatrix}$$

開始, 當 t 夠大時, 得

(1) $N(t) = 1.6N(t-1)$, 亦即,

$$\text{族群成長率} = 1.6$$

(大的特徵值).

(2) 穩定年齡層分布爲

$$\begin{bmatrix} 2000 \\ 100 \end{bmatrix}$$

(大的特徵值 1.6 所對應的特徵向量).

此二結果可由上述的事實獲得, 如下述.

(i) 求特徵值 λ_1 與 λ_2: 需解

$$\det(L - \lambda I) = 0$$

亦即, 解

$$\det \begin{bmatrix} 1.5-\lambda & 2 \\ 0.08 & -\lambda \end{bmatrix} = 0$$

計算行列式後, 得

$$(1.5-\lambda)(-\lambda) - 2(0.08) = 0$$

亦相當於

$$\lambda^2 - 1.5\lambda - 0.16 = 0$$

經由因式分解, 得

$$(\lambda - 1.6)(\lambda + 0.1) = 0$$

所以, 兩個特徵值為

$$\lambda_1 = 1.6 \text{ 與 } \lambda_2 = -0.1$$

其中大的特徵值 1.6 = 族群成長率.

(ii) 求 $\lambda_1 = 1.6$ 所對應的特徵向量: 需要求

$$L \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = 1.6 \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

此乃相當於解

$$\begin{aligned} 1.5x_1 + 2x_2 &= 1.6x_1 \\ 0.08x_1 + 0x_2 &= 1.6x_2 \end{aligned}$$

亦相當於解

$$\begin{aligned} -0.1x_1 + 2x_2 &= 0 \\ 0.08x_1 - 1.6x_2 &= 0 \end{aligned}$$

第一式 $\times(-10)$, 第二式 $\div 0.08$ 後, 得等價的線性方程式系統

$$\begin{aligned} x_1 - 20x_2 &= 0 \\ x_1 - 20x_2 &= 0 \end{aligned}$$

由此得

$$x_1 = 20x_2$$

因此, 可取

$$\mathbf{u}_1 = \begin{bmatrix} 20 \\ 1 \end{bmatrix}, \begin{bmatrix} 40 \\ 2 \end{bmatrix}, \ldots, \begin{bmatrix} 200 \\ 10 \end{bmatrix}, \begin{bmatrix} 2000 \\ 100 \end{bmatrix}, \ldots$$

這些都是穩定年齡層分布.

問. 爲何會有上述結果?

答. 以例說明, 設

$$N(0) = \begin{bmatrix} 105 \\ 1 \end{bmatrix}$$

首先根據前述的結果, 特徵值 $\lambda_1 = 1.6$ 所對應的特徵向量

$$\mathbf{u}_1 = \begin{bmatrix} 20 \\ 1 \end{bmatrix}$$

接著, 求特徵值 $\lambda_2 = -0.1$ 所對應的特徵向量. 根據特徵向量的定義, 此乃相當於解矩陣方程式

$$L \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = -0.1 \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

將上式乘開後, 得一線性方程式系統

$$\begin{aligned} 1.5x_1 + 2x_2 &= -0.1x_1 \\ 0.08x_1 &= -0.1x_2 \end{aligned}$$

亦相當於

$$\begin{aligned} 1.6x_1 + 2x_2 &= 0 \\ 0.08x_1 + 0.1x_2 &= 0 \end{aligned}$$

將上式中的 "第一式 $\div 2$" 以及 "第二式 $\times 10$", 可得等價的方程式

$$0.8x_1 + x_2 = 0$$

因此, 可取

$$\mathbf{u}_2 = \begin{bmatrix} 5 \\ -4 \end{bmatrix}$$

又, 很明顯地可將

$$N(0) = \begin{bmatrix} 105 \\ 1 \end{bmatrix}$$

表示成特徵向量 $\mathbf{u}_1$ 與 $\mathbf{u}_2$ 的線性組合, 如下所示:

$$\begin{bmatrix} 105 \\ 1 \end{bmatrix} = 5 \begin{bmatrix} 20 \\ 1 \end{bmatrix} + \begin{bmatrix} 5 \\ -4 \end{bmatrix} = 5\mathbf{u}_1 + \mathbf{u}_2$$

因此, 根據迭代映射的結論,

$$\begin{aligned} N(t) &= L^t \begin{bmatrix} 105 \\ 1 \end{bmatrix} \\ &= L^t(5\mathbf{u}_1 + \mathbf{u}_2) \\ &= 5(1.6)^t \begin{bmatrix} 20 \\ 1 \end{bmatrix} + (-0.1)^t \begin{bmatrix} 5 \\ -4 \end{bmatrix} \end{aligned}$$

因爲

$$\lim_{t \to \infty} (-0.1)^t = 0$$

故, 當 t 夠大時,

$$N(t) \approx 5(1.6)^t \begin{bmatrix} 20 \\ 1 \end{bmatrix} \tag{3}$$

因此, 根據 (3) 式, 當 t 夠大時,

$$\begin{aligned} N(t) &\approx (1.6)5(1.6)^{t-1}\begin{bmatrix} 20 \\ 1 \end{bmatrix} \\ &\approx 1.6N(t-1) \end{aligned}$$

所以, 由最後一式可得如下的二點結論:

(1) 大的特徵值 1.6 = 族群成長率

(2) 若

$$N(0) = \begin{bmatrix} 20 \\ 1 \end{bmatrix} = 1\cdot\mathbf{u}_1 + 0\cdot\mathbf{u}_2$$

則, 根據迭代映射的結論,

$$\begin{aligned} N(t) &= L^t\begin{bmatrix} 20 \\ 1 \end{bmatrix} \\ &= 1\cdot(1.6)^t\begin{bmatrix} 20 \\ 1 \end{bmatrix} + 0\cdot(-0.1)^t\begin{bmatrix} 5 \\ -4 \end{bmatrix} \\ &= (1.6)^t\begin{bmatrix} 20 \\ 1 \end{bmatrix} \end{aligned}$$

所以, 在任意的 t 時, 年齡 0 所占的比率

$$q_0(t) = \frac{(1.6)^t 20}{(1.6)^t 20 + (1.6)^t} = \frac{20}{20+1} = q_0(0)$$

剛好就是在 0 時, 年齡 0 所占的比率. 因此, 根據定義, 最大特徵值所對應的特徵向量

$$\mathbf{u}_1 = \begin{bmatrix} 20 \\ 1 \end{bmatrix}$$

是一穩定年齡層分布. 事實上, 對任意的 $a > 0$,

$$a \begin{bmatrix} 20 \\ 1 \end{bmatrix}$$

均是穩定年齡層分布.