淺談特徵值與特徵向量

國立台中師範學院教育測驗統計研究所

楊明宗

一、前言

在多變項分析中，因素分析(factor analysis)是一種被廣泛使用的統計方法，在我們實際使用時，有幾種不同的方法可用來抽取共同因素。過去電腦未被廣泛使用之前，一般多使用「形心法」，但電腦逐漸普及後，「主軸法」則是最常被用來抽取因素的方法。在「主軸法」的各項運算中涉及兩個很重要的基本概念，即「特徵值」(eigen value)與「特徵向量」(eigen vector)，這部份看起來十分抽象，有很多學生在學習時，都只知道強記算法或公式，而不思考概念所蘊涵的真義，因此，常有不知所云或一知半解的現象。若是學習者能先對這兩個概念的意義有所理解並加以引申運用，則在學習因素分析的原理時，許多疑難雜症必能迎刃而解。基於此點，本文擬從線性代數的觀點著手，舉例說明並配合座標圖的呈現，分析矩陣在數學上的意義，而後再進一步探討特徵值與特徵向量的定義、計算方法及其作用。

二、矩陣的意義

我們先從解下列的二元一次聯立方程組出發

$$a_1x + b_1y = c_1$$

$$a_2x + b_2y = c_2$$

它可以下列之矩陣形式來表達

$$\begin{bmatrix} a_1 & b_1 \\ a_2 & b_2 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} c_1 \\ c_2 \end{bmatrix}$$

即為 $AX=C$，其中

$$A=\begin{bmatrix} a_1 & b_1 \\ a_2 & b_2 \end{bmatrix}，X=\begin{bmatrix} x \\ y \end{bmatrix}，C=\begin{bmatrix} c_1 \\ c_2 \end{bmatrix}$$

又 $AA^{-1}X=A^{-1}C$，即 $X=A^{-1}C$

因此等式無常數項，所以 X 向量必通過原點，如圖 2-1。

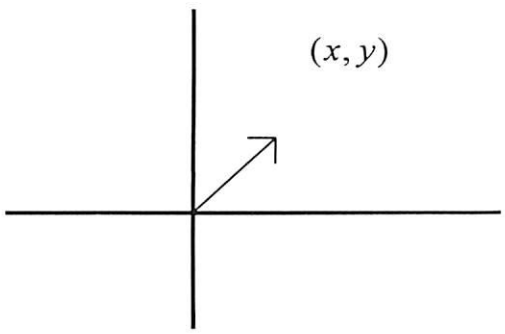


圖 2-1

由上面的計算可知二元一次聯立方程組的求解，可以表達成 $AX=C$ 的求解，而 $AX=C$ 的求解則很類似解一元一次方程式 $ax=c$，由於轉換成矩陣形式，讓我們更能清楚看出矩陣的線性意義(即比例倍數上的意義)。

以下將舉幾個例子，可讓我們充分掌握線性意義，同時亦讓我們了解到二元一次方程組與一元一次方程式，透過矩陣表達後，在概念上沒有兩樣。

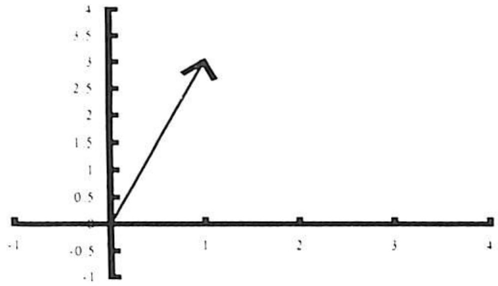
例 2-1：

令 $X=\begin{bmatrix} 1 \\ 3 \end{bmatrix}$，$X'=\begin{bmatrix} 2 \\ 0 \end{bmatrix}$，$A=\begin{bmatrix} 2 & 0 \\ 0 & 0 \end{bmatrix}$

因 $\begin{bmatrix} 2 & 0 \\ 0 & 0 \end{bmatrix}\begin{bmatrix} 1 \\ 3 \end{bmatrix}=\begin{bmatrix} 2 \\ 0 \end{bmatrix}$，

即 $AX=X'$ (2-1)

由式 2-1，X 向量經由 A 矩陣的作用後轉換為 X' 矩陣，其在幾何上的意義如圖 2-2 所示。



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

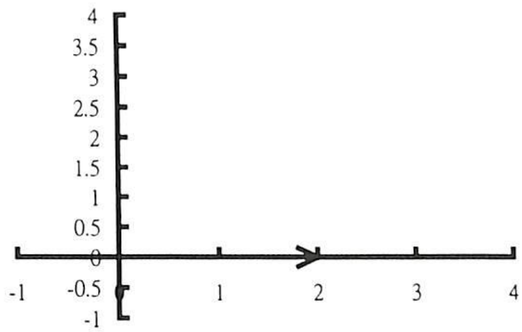


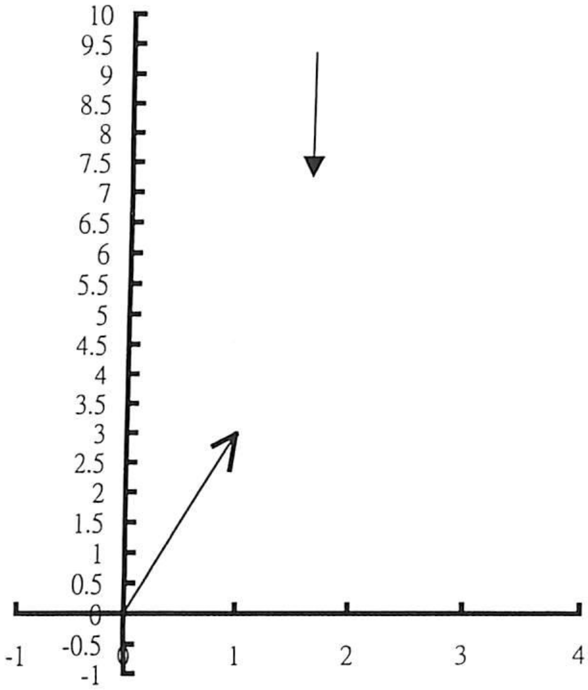
圖 2-2

由圖 2-2，我們可知，X 向量經 A 矩陣作用後，只對 X 軸分量加倍。

例 2-2：改變 A 矩陣，作法同上。

即 $\begin{bmatrix}0 & 0\\ 0 & 3\end{bmatrix}\begin{bmatrix}1\\ 3\end{bmatrix}=\begin{bmatrix}0\\ 9\end{bmatrix}$，其在幾何上的意義如圖 2-3 所示。

由圖 2-3 可知，X 向量經 A 矩陣作用後，只在 Y 軸分量變 3 倍。從上面兩個例子，我們不難發現，若將 A 矩陣改為對稱矩陣，結果將會非常有趣，請看以下例子。



$$A=\begin{bmatrix} 0 & 0 \\ 0 & 3 \end{bmatrix}$$

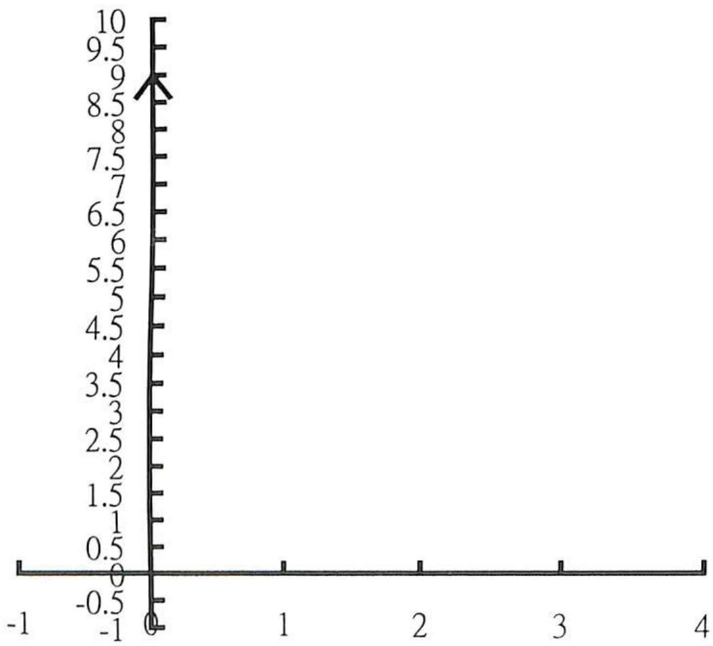


圖 2-3

由圖 2-3 可知，X 向量經 A 矩陣作用後，只在 Y 軸分量變 3 倍。從上面兩個例子，我們不難發現，若將 A 矩陣改為對稱矩陣，結果將會非常有趣，請看以下例子。

例 2-3：將$\begin{bmatrix}2 & 0\\0 & 3\end{bmatrix}\begin{bmatrix}1\\3\end{bmatrix}=\begin{bmatrix}2\\9\end{bmatrix}$以圖 2-4 表示二向量之間的變化如下

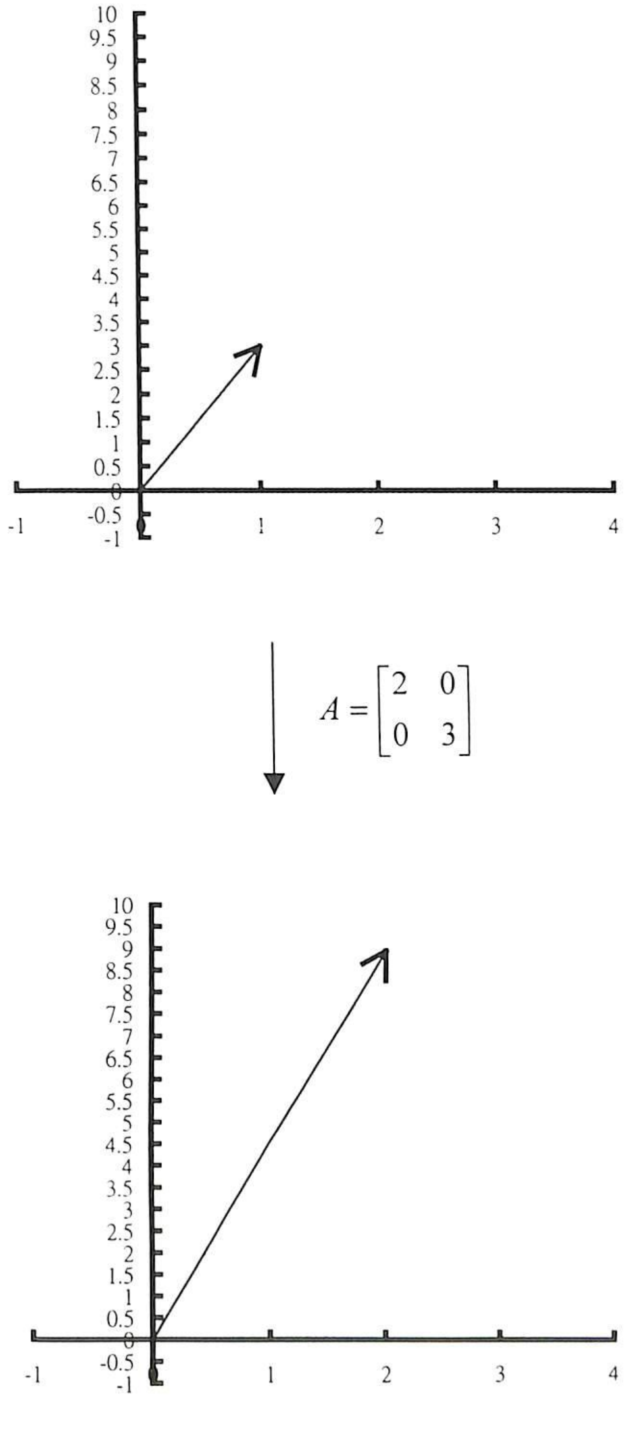
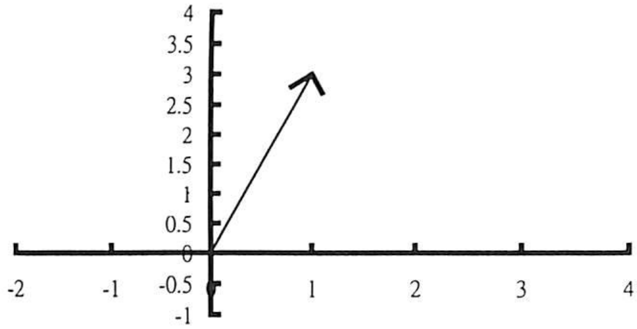


圖 2-4

由圖 2-4 可知，X 向量經 A 矩陣作後，會在 X 軸分量加倍，在 Y 軸分量變 3 倍。綜合上述三例，我們可看出矩陣的第一種意義。

結論一： 矩陣具有縮小放大之意義，即任一非零向量經矩陣作用後可被縮小或放大。

例 2-4：$\begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} 1 \\ 3 \end{bmatrix}=\begin{bmatrix} -1 \\ 3 \end{bmatrix}$ 以圖 2-5 表示二向量之間的變化如下



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

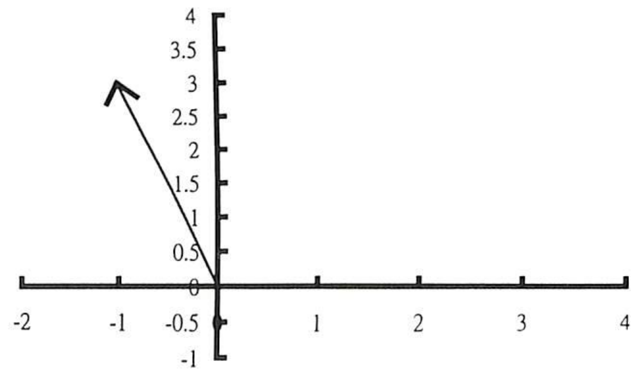


圖 2-5

由圖 2-5 可知 X 向量和 X' 向量之間具有鏡射之關係，因此我們可以推出矩陣的第二種意義。

結論二：矩陣具有鏡射(線對稱)之意義，即任一非零向量經矩陣作用後可產生對稱之向量。

例 2-5：$\begin{bmatrix} \frac{1}{2} & -\frac{\sqrt{3}}{2} \\ \frac{\sqrt{3}}{2} & \frac{1}{2} \end{bmatrix}\begin{bmatrix} 1 \\ 0 \end{bmatrix}=\begin{bmatrix} \frac{1}{2} \\ \frac{\sqrt{3}}{2} \end{bmatrix}$

因為 A 矩陣$=\begin{bmatrix} \frac{1}{2} & -\frac{\sqrt{3}}{2} \\ \frac{\sqrt{3}}{2} & \frac{1}{2} \end{bmatrix}=\begin{bmatrix} \cos\frac{\pi}{3} & \sin\frac{\pi}{3} \\ \sin\frac{\pi}{3} & \cos\frac{\pi}{3} \end{bmatrix}$

所以，X 向經 A 矩陣作用後，會旋轉 $\frac{\pi}{3}$ 的角度，而 X' 為向量，如圖 2-6。

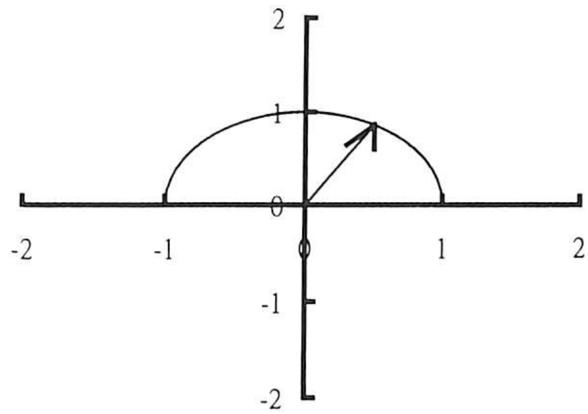


圖 2-6

若將 X 向量改為 $\begin{bmatrix} 0 \\ 1 \end{bmatrix}$ 則旋轉的情況如圖 2-7。

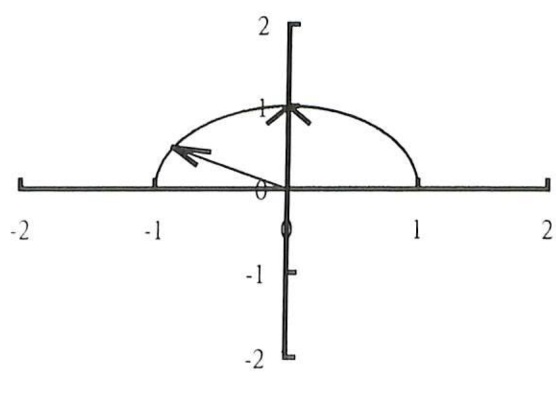


圖 2-7

以上類似此種具有旋轉功能的矩陣，我們稱之為「旋轉矩陣」。由上面的例子，提供了旋轉矩陣的二個條件：

(1) $A=\begin{bmatrix} a & -b \\ b & a \end{bmatrix} \quad a,b\in R$

(2) $a^2+b^2=1$

從幾何的觀點來看，我們亦可輕易看出經旋轉矩陣的作用後，向量 X 並不增加其長度，只會改變其方向。

設 $A=\begin{bmatrix} a & -b \\ b & a \end{bmatrix}$，$X=\begin{bmatrix} x \\ y \end{bmatrix}$，$X^{'}=\begin{bmatrix} ax-by \\ bx+ay \end{bmatrix}$

$\because \begin{bmatrix} a & -b \\ b & a \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix}=\begin{bmatrix} ax-by \\ bx+ay \end{bmatrix}$

$X^{'}$ 向量之長度為

$$\|X'\|=\sqrt{(ax-by)^2+(bx+ay)^2}=\sqrt{(a^2+b^2)x^2+(a^2+b^2)y^2}=\sqrt{x^2+y^2}=\|X\|$$

綜合以上，我們可推出矩陣的第三種意義。

結論三：矩陣具有旋轉的作用，即任何非零向量經矩陣作用後可被旋轉。

三、特徵值與特徵向量

由上述對矩陣意義的說明，我們不難發現，就抽象如「矩陣」者，也是可以根據它所作用的向量，及作用後的向量來加以掌握。而特徵值與特徵向量便是基於此種特性而產生，以下說明特徵值與特徵向量的定義，並配合實例加以說明其計算之方法，以期讀者能一目了然。

【定義 3.1】設 A 為一 m×n 矩陣，在 R^n 中，若存在一非零向量，滿足 AX=λX 則稱實數λ為 A 之特徵值，任意非零向量 X 稱為 A 之對應於特徵值λ之特徵向量。

由上面的定義可知，矩陣 A 的特徵值意義就是指存在一個向量 X 經矩陣 A 的作用後，會往 X 方向延伸 λ 倍。為了便於說明，我們先從特徵值與特徵向量的計算著手，再將計算的結果應用在向量與矩陣的交互作用。

例 3-1：設 $A=\begin{bmatrix}7 & 5\\ 4 & 2\end{bmatrix}$，求 A 的特徵值與其對應之特徵向量。

（解法）

由定義 3.1 我們知道，存在一非零向量 X 使得 AX=λX

即 $\begin{bmatrix}7 & 2\\ 4 & 5\end{bmatrix}\begin{bmatrix}x\\ y\end{bmatrix}=\lambda\begin{bmatrix}x\\ y\end{bmatrix} \longrightarrow \begin{bmatrix}7 & 2\\ 4 & 5\end{bmatrix}\begin{bmatrix}x\\ y\end{bmatrix}=\lambda\begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix}\begin{bmatrix}x\\ y\end{bmatrix}$

$$\begin{bmatrix}7 & 2\\ 4 & 5\end{bmatrix}\begin{bmatrix}x\\ y\end{bmatrix}=\begin{bmatrix}\lambda & 0\\ 0 & \lambda\end{bmatrix}\begin{bmatrix}x\\ y\end{bmatrix} \longrightarrow \begin{bmatrix}7 & 2\\ 4 & 5\end{bmatrix}\begin{bmatrix}x\\ y\end{bmatrix}-\begin{bmatrix}\lambda & 0\\ 0 & \lambda\end{bmatrix}\begin{bmatrix}x\\ y\end{bmatrix}=0$$

$$\begin{bmatrix}7-\lambda & 2\\ 4 & 5-\lambda\end{bmatrix}\begin{bmatrix}x\\ y\end{bmatrix}=0 \qquad (1)$$

$$\begin{vmatrix}7-\lambda & 2\\ 4 & 5-\lambda\end{vmatrix}=0$$

$$(7-\lambda)(5-\lambda)-8=0$$

$$\lambda^2-12\lambda+27=0$$

$$(\lambda-3)(\lambda-9)=0$$

$\lambda=3$ 或 9

當 λ=3 時代入式(1)

$$\begin{bmatrix} 7-\lambda & 2 \\ 4 & 5-\lambda \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix}=0 \rightarrow \begin{bmatrix} 4 & 4 \\ 2 & 2 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix}=0$$

$$\begin{bmatrix} 4x+2y \\ 4x+2y \end{bmatrix}=0$$

$$4x+2y=0$$

$$y=-2x$$

$$\begin{bmatrix} x \\ y \end{bmatrix}=\begin{bmatrix} x \\ -2x \end{bmatrix}=x\begin{bmatrix} 1 \\ -2 \end{bmatrix}$$

$X\lambda_1=\begin{bmatrix} 1 \\ -2 \end{bmatrix}$ (即 λ_1=3 時之特徵向量)

當 λ=9 時代入式(1)

$$\begin{bmatrix} 7-\lambda & 2 \\ 4 & 5-\lambda \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix}=0 \rightarrow \begin{bmatrix} -2 & 2 \\ 4 & -4 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix}=0$$

$$\begin{bmatrix} -2x+2y \\ 4x-4y \end{bmatrix}=0$$

$-2x+2y=0$ 且 $4x-4y=0$

$$x=y$$

$$\begin{bmatrix} x \\ y \end{bmatrix}=\begin{bmatrix} y \\ y \end{bmatrix}=y\begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

$$X\lambda_2=\begin{bmatrix}1\\1\end{bmatrix}$$ (即 $\lambda_2=9$ 時之特徵向量)

由上解得　$\lambda_1=3$ 且 $X\lambda_1=\begin{bmatrix}1\\-2\end{bmatrix}$

$\lambda_2=9$ 且 $X\lambda_2=\begin{bmatrix}1\\1\end{bmatrix}$

四、特徵值與特徵向量之應用

例 4-1：設 $A=\begin{bmatrix}7 & 2\\4 & 5\end{bmatrix}$，$X=\begin{bmatrix}3\\-3\end{bmatrix}$，求 AX

我們可直接從矩陣的運算得

$$AX=\begin{bmatrix}7 & 2\\4 & 5\end{bmatrix}\begin{bmatrix}3\\-3\end{bmatrix}=\begin{bmatrix}15\\-3\end{bmatrix}$$

但考慮將 $X=\begin{bmatrix}3\\-3\end{bmatrix}$ 表成 $X=\begin{bmatrix}3\\-3\end{bmatrix}=2\begin{bmatrix}1\\-2\end{bmatrix}+1\begin{bmatrix}1\\1\end{bmatrix}$　(因 $X\lambda_1$ 與 $X\lambda_2$ 為線性獨立)

即 $X=2X\lambda_1+1X\lambda_2$　($X\lambda_1$，$X\lambda_2$ 為上例對應於 A 之特徵向量)

再將 A 作用於 X 上，

$$AX=A(2X\lambda_1+1X\lambda_2)$$

$$AX=2AX\lambda_1+AX\lambda_2$$

$AX=2\lambda_1X\lambda_1+\lambda_2X\lambda_2=6X\lambda_1+9X\lambda_2$ (由定義 3.1 可得 $AX\lambda_1=\lambda_1X\lambda_1$，$AX\lambda_2=\lambda_2X\lambda_2$)

五、結論

由上述之過程我們可知向量 X 經過矩陣 A 之作用，可由複雜的計算轉換成簡單的相對於 A 之特徵值與特徵向量之計算。其中必須先將向量 X 之歐氏基底轉換成以 $X\lambda_1$ 與 $X\lambda_2$ 為基底，再行計算，用圖形可表為如圖 5-1。

$X=\begin{bmatrix} 3 \\ -3 \end{bmatrix}$，經過 $A=\begin{bmatrix} 7 & 2 \\ 4 & 5 \end{bmatrix}$ 的作用後如下圖

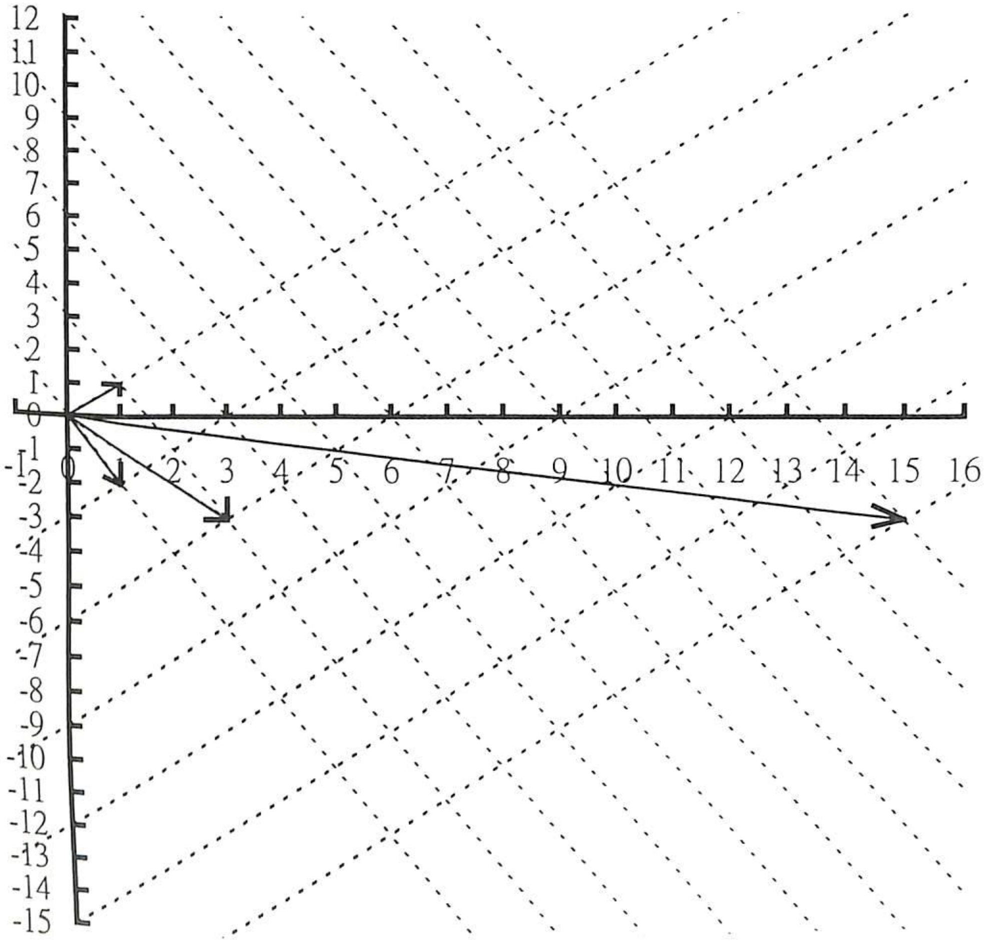


圖 5-1

從圖形中我們知道 X=$\begin{bmatrix} 3 \\ -3 \end{bmatrix}$=2$\begin{bmatrix} 1 \\ -2 \end{bmatrix}$+1$\begin{bmatrix} 1 \\ 1 \end{bmatrix}$已被轉換成 X=2X$\lambda_1$+1X$\lambda_2$ 再將 AX 用 2λ_1Xλ_1 與λ_2Xλ_2來表示即 6Xλ_1與 9Xλ_2表示，如圖可看出 AX 已經用 6 倍 Xλ_1的與 9 倍的 Xλ_2來表示了。

從圖形中可看出向量 X，經過 A 的作用後，不但在方向上改變，其長度亦被放大了。正說明了矩陣具有放大或縮小與改變方向(旋轉)的性質。

由此我們可知，任一非零向量 X 經過一 m×n 矩陣 A 作用後，可表成

$$AX=A(a_1X\lambda_1+a_2X\lambda_2+\cdots+a_nX\lambda_n)$$

即 $AX=a_1AX\lambda_1+a_2AX\lambda_2+\cdots+a_nAX\lambda_n$

$$AX=a_1\lambda_1X\lambda_1+a_2\lambda_2X\lambda_2+\cdots+a_n\lambda_nX\lambda_n$$

其中 $X=a_1X\lambda_1+a_2X\lambda_2+\cdots+a_nX\lambda_n$ 為 X 轉換成以 Xλ_1，Xλ_2，…，Xλ_n 當基底表示。Xλ_1，Xλ_2，…，Xλ_n 為 A 之對應於特徵值λ_n之特徵向量。

參考文獻

許天維、施淑娟(1995)。特徵值與特徵向量。台中師院數理教育系刊，14-21。