從等比級數談起

謹以此文紀念中央研究院數學所陳明博教授 (1941.6.23 - 1997.12.9)

林琦焜

單研究人, 獲得的是沒有靈魂的知識軀殼, 單研究書籍, 獲得的是沒有血肉的知識的魂。看見, 並且觀察; 閱讀, 並且反省, 便是走在通往知識的正途上, 但必須在探討別人的心時, 不忽略自己的心。

— Caleb Colton —

前言

山不在高, 有仙則名; 水不在深, 有龍則靈。數學也是如此, 常常我們只爲了尋求很難的問題而將周遭切身的眞理疏忽了, 以致淪爲耍雜技, 訓練有素的猴子, 或者最後一無所有。在此筆者要特別呼籲"好的數學"(good mathematics) 就是以越簡單的方法來解決困難的問題。如果要拔一根草需要開一部割草機或用大斧頭, 那實在是"有夠蠢"。就好比有人在小學, 先學了中學數學, 利用中學數學來解小學的題目, 無知的大衆都說這是天才, 但筆者則認爲這頂多是另一隻耍雜技的猴子。

我們將從等比數列談起, 而後想辦法將此觀念推廣, 爲了解決此問題, 自然而然矩陣的觀念就被引進, 除了介紹矩陣之外, 我們特別對 Fibonacci 數列作了深入之研究, 當然我們的方法並沒有限定在此數列對其他的數列, 讀者也可嘗試。

我們在此文特別要探討的是"固有值", 除了以數列的角度來看之外, 也將由極值與微分方程的角度來思考, 最後要說的是"固有值就在你身邊"。

1. 等比數列, Fibonacci 數列

相信諸位在國中就學過等比數列, 而且也專精於此, 跟大家談這個題目也並不是要欺負你們, 而實在是常常許多眞理在我們身邊, 卻不時爲我們所疏忽。在這個專題中我們要強調的是深奧的數學常常是寓含在熟悉的數學中。

回顧一下等比數列

$$a_0, a_1, a_2, \ldots, a_n, a_{n+1}, \ldots$$
$$\frac{a_{n+1}}{a_n} = \lambda, \qquad n = 0, 1, 2, \tag{1.1}$$

λ就是公比, 等比數列方便之處在於只需知道數列中任一項則每一項自然就尾隨而至, 換句話說, 描述等比數列僅需一個參數 (parameter) 即足夠

$$a_n = \lambda^n a_0, \qquad a_0 \in \mathbb{R} \tag{1.2}$$

或者我們可以這麼說, 這是一個一維的 (one dimensional) 問題, 但等比數列畢竟是少數的特例, 看個出名的 Fibonacci 數列

$$0, 1, 1, 2, 3, 5, 8, 13 \ldots \tag{1.3}$$

仔細觀察不難看出其遞推關係爲

$$a_0 = 0, a_1 = 1, \ldots, a_{n+2} = a_{n+1} + a_n,\ n \geq 0 \tag{1.4}$$

基本上這是一個差分方程 (difference equation)。現在的問題是如何求得第n項, 例如a_{100}? 對於第n項是否有相類似於等比數列 (1.2) 之公式呢? 首先我們先解決 Fibonacci 數列的問題, 典型的方法如下: 由遞迴關係

$$a_{n+2} = a_{n+1} + a_n$$

知處於中間媒介的a_{n+1}如果可分解成兩部份

$$a_{n+1} = (\alpha + \beta) a_{n+1}, \qquad \alpha + \beta = 1 \tag{1.5}$$

則原式可表爲

$$\begin{cases} a_{n+2} - \alpha a_{n+1} = \beta a_{n+1} + a_n \\ a_{n+2} - \beta a_{n+1} = \alpha a_{n+1} + a_n \end{cases} \tag{1.6}$$

由 (1.6) 式知如果a_n之係數可表爲

$$\alpha\beta = -1 \tag{1.7}$$

則 (1.6) 式爲

$$\begin{aligned} a_{n+2} - \alpha a_{n+1} &= \beta(a_{n+1} - \alpha a_n) \\ &= \cdots = \beta^{n+1}(a_1 - \alpha a_0) \\ a_{n+2} - \beta a_{n+1} &= \alpha(a_{n+1} - \beta a_n) \\ &= \cdots = \alpha^{n+1}(a_1 - \beta a_0) \end{aligned} \tag{1.8}$$

這是兩個等比數列! 換句話說, 如果可找到 α, β 同時滿足 (1.5), (1.7) 兩式, 則 Fibonacci 數列可化爲兩個等比數列。而α, β之意義呢?

$$\begin{aligned} a_{n+2} - a_{n+1} - a_n = 0 \\ \Longleftrightarrow x^2 - x - 1 = 0 \end{aligned} \tag{1.9}$$

由根與係數之關係知, 若 α, β爲一元二次方程式 $x^2 - x - 1 = 0$之根, 則 α, β滿足

$$\alpha + \beta = 1, \quad \alpha\beta = -1$$

這正是 (1.5), (1.7) 兩式告訴我們的。利用 (1.8) 式計算可得

$$\begin{aligned} a_{n+2} = \frac{1}{\alpha - \beta} \Big[& \alpha^{n+2}(a_1 - \beta a_0) \\ & - \beta^{n+2}(a_1 - \alpha a_0) \Big] \end{aligned} \tag{1.10}$$

有了公式當然很好, 但我們的目標並不在此, 數學並非以求得答案爲唯一目的。重要的是在學習過程中, 明白這其中之架構而後詢問是否可推廣, 有否一般性, 如果一個方法只能解決一類問題, 那麼這並非一"好數學"! 上述之方法對於更複雜的數列, 顯然有其困

難, 因此發展新的方法與工具乃是必然的結果。但記住我們的思考法則與概念還是建立在原有的基礎上, 而非憑空想像而來。如果學數學是依賴“靈感”的話, 那麼我們來接受教育便沒有意義!

對於 Fibonacci 數列而言: $a_{n+2} = a_{n+1} + a_n$ 每個數 a_{n+2} 都是由前面兩個數 a_{n+1}, a_n 所決定。換句話說, 必須有兩個參數 (parameters) 方能完整地描述 Fibonacci 數列, 因此我們引進向量的概念

$$\begin{aligned}\vec{x}_0 &= \begin{bmatrix}1\\0\end{bmatrix}, \vec{x}_1 = \begin{bmatrix}1\\1\end{bmatrix},\\ \vec{x}_2 &= \begin{bmatrix}2\\1\end{bmatrix}, \ldots, \vec{x}_n = \begin{bmatrix}a_{n+1}\\a_n\end{bmatrix}, \ldots\end{aligned} \tag{1.11}$$

則 Fibonacci 數列可表爲

$$\begin{cases} a_{n+2} = a_{n+1} + a_n \\ a_{n+1} = a_{n+1} \end{cases}$$

即

$$\begin{aligned}\vec{x}_{n+1} &= \begin{bmatrix}1 & 1\\1 & 0\end{bmatrix}\vec{x}_n \equiv A\vec{x}_n\\ &= \cdots = A^{n+1}\vec{x}_0\end{aligned} \tag{1.12}$$

此式正是等比數列 (1.2) 之推廣。我們先探討一下矩陣 $A = \begin{bmatrix}1 & 1\\1 & 0\end{bmatrix}$, 直接計算可得

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

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

即矩陣A滿足一元二次方程式$f(x) = x^2 - x - 1$

$$f(A) = A^2 - A - I = 0$$

這正是 (1.9) 式所告訴我們的, 在線性代數稱爲 **Cayley–Hamilton** 定理。其實這個多項式很特別, 通常稱之爲特徵多項式 (characteristic polynomial), 該多項式可由行列式而得

$$f(\lambda) = |\lambda I - A| = \det(\lambda I - A) \tag{1.13}$$

對於 Fibonacci 數列,$A = \begin{bmatrix}1 & 1\\1 & 0\end{bmatrix}$ 其特徵多項式爲

$$\begin{aligned}f(\lambda) &= \begin{vmatrix}\lambda - 1 & -1\\ -1 & \lambda\end{vmatrix}\\ &= (\lambda - 1)\lambda - 1 = \lambda^2 - \lambda - 1\end{aligned}$$

而 α, β爲$f(\lambda)$之兩個根, 在線性代數我們稱之爲 固有值(eigenvalue), 即固定不變的值, 這可由 (1.8) 式視出一些端倪。

在前面提到α, β之求法, 實際上就相當於根與係數之關係, 這在矩陣有相似的理論, 我們先介紹一個簡單但重要的概念 —— 跡 (trace)。

定義: 矩陣$A = (a_{ij})_{ij}$之跡 (trace) 爲

$$\mathrm{tr}A = \sum_{i=1}^{n} a_{ii} \quad (\text{對角線之和}) \tag{1.14}$$

對 Fibonacci 數列而言 $A = \begin{bmatrix}1 & 1\\1 & 0\end{bmatrix}$ 其跡 (trace) 爲

$$\mathrm{tr}A = 1 + 0 = 1$$

再回到特徵多項式可知

$$f(\lambda)=\lambda^2-\lambda-1=\lambda^2-(\mathrm{tr}A)\lambda+(\det A) \tag{1.15}$$

該式對任何2×2矩陣都對, 但我們並不打算證明, 只是提出這個事實。讓我們回到 (1.12) 式

$$\vec{x}_{n+1}=A\vec{x}_n=A^n\vec{x}_1$$

當n不大時A^n是不難計算, 但當n越來越大之時, 我們就不能憑匹夫之勇, 硬算算出來那是浪費生命的。認爲只要努力就有收穫, 那是不正確的, 一個人的奮鬥必須有正確的方向, 否則只是徒勞。某些特殊矩陣A, 其A^n是容易求, 例如:

$$A=\lambda I=\begin{bmatrix}\lambda & 0\\ 0 & \lambda\end{bmatrix} \implies A^n=\begin{bmatrix}\lambda^n & 0\\ 0 & \lambda^n\end{bmatrix}$$

其實這就是等比數列所對應的矩陣。另外

$$A=\begin{bmatrix}\lambda_1 & 0\\ 0 & \lambda_2\end{bmatrix} \implies A^n=\begin{bmatrix}\lambda_1^n & 0\\ 0 & \lambda_2^n\end{bmatrix}$$

對於其他矩陣就沒有這麼便宜, 如何求A^n?

這牽扯到矩陣對角化的問題。我們還是在回到 Fibonacci 數列, 由 (1.6) 式可令

$$\begin{cases} x_{n+2}\equiv a_{n+2}-\alpha a_{n+1}\\ y_{n+2}\equiv a_{n+2}-\beta a_{n+1}\end{cases}$$

$$\vec{y}_{n+2}=\begin{bmatrix}x_{n+2}\\ y_{n+2}\end{bmatrix} \tag{1.16}$$

即 (1.6) 式可改寫爲

$$\begin{cases} x_{n+2}=\beta x_{n+1}\\ y_{n+2}=\alpha y_{n+1}\end{cases} \Longleftrightarrow \vec{y}_{n+2}=\begin{bmatrix}\beta & 0\\ 0 & \alpha\end{bmatrix}\vec{y}_{n+1} \tag{1.17}$$

這告訴我們 Fibonacci 數列經由適當的變換, 可化爲兩個等比數列 (這個過程就是對角化 diagonalization) 其關係爲

$$\begin{aligned}\vec{x}_{n+2}&=\begin{bmatrix}1 & 1\\ 1 & 0\end{bmatrix}\vec{x}_{n+1}\\ \Longleftrightarrow \vec{y}_{n+2}&=\begin{bmatrix}\beta & 0\\ 0 & \alpha\end{bmatrix}\vec{y}_{n+1}\end{aligned} \tag{1.18}$$

至此我們遇到的第一個問題是兩個不同矩陣

$$A=\begin{bmatrix}1 & 1\\ 1 & 0\end{bmatrix}, \qquad \Lambda=\begin{bmatrix}\beta & 0\\ 0 & \alpha\end{bmatrix} \tag{1.19}$$

是如何關聯的? 這就是矩陣對角化的問題, 也是線性代數研究的主題之一。

由 (1.16) 式知

$$\vec{y}_{n+2}=\begin{bmatrix}1 & -\alpha\\ 1 & -\beta\end{bmatrix}\vec{x}_{n+2}\equiv P^{-1}\vec{x}_{n+2} \tag{1.20}$$

代入 (1.18) 式得

$$\begin{aligned}\vec{y}_{n+2}&=P^{-1}\vec{x}_{n+2}=P^{-1}A\vec{x}_{n+1}\\ &=\Lambda\vec{y}_{n+1}=\Lambda P^{-1}\vec{x}_{n+1}\end{aligned}$$

這相當於說

$$P^{-1}A=\Lambda P^{-1} \iff A=PAP^{-1} \tag{1.21}$$

這就是所謂對角化的過程, 好處何在? 首先看矩陣的相乘

$$\begin{aligned}A^2&=A\cdot A=(P\Lambda P^{-1})(P\Lambda P^{-1})\\ &=P\Lambda(P^{-1}P)\Lambda P^{-1}=P\Lambda^2P^{-1}\\ A^n&=A\cdots A=(P\Lambda P^{-1})(P\Lambda P^{-1})\\ &\quad\cdots(P\Lambda P^{-1})=P\Lambda^nP^{-1}\end{aligned}$$

而其中

$$\Lambda^n=\begin{bmatrix}\beta & 0\\ 0 & \alpha\end{bmatrix}^n=\begin{bmatrix}\beta^n & 0\\ 0 & \alpha^n\end{bmatrix}$$

這說明原先矩陣A^n是要經過 n個步驟, 但如今只需要三個步驟便可求得, 這就是對角化的好處之一。

關於矩陣P, 我們可以視爲高維度空間的變數變換, 但其意義卻不僅於此, 與固有值對應的就是 固有空間(eigenspace)

$$W_\alpha = \{\vec{x} \in \mathbb{R}^2 \mid A\vec{x} = \alpha\vec{x}\} \tag{1.22}$$

但

$$A\vec{x} = \alpha\vec{x} \iff (A - \alpha I)\vec{x} = 0 \tag{1.23}$$

因爲行列式 $\det(A - \alpha I) = |A - \alpha I| = 0$, 由解聯立方程組之經驗告訴我們 (1.23) 可能無解或無窮多解, 但顯然$\vec{x} = \vec{0}$爲(1.23)之一個解, 故 (1.23) 式有無窮多解, 這種解我們稱之爲 固有向量(eigenvector), 如何求固有向量呢? 當然還是要回到“聯立方程組”

$$(A - \alpha I)\vec{x} = 0 \iff \begin{cases} (1-\alpha)x + y = 0 \\ x - \alpha y = 0 \end{cases}$$

其解爲

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

即固有空間W_α爲

$$W_\alpha = \left\{ C_\alpha \begin{bmatrix} \alpha \\ 1 \end{bmatrix} \,\middle|\, C_\alpha \in \mathbb{R} \right\} \tag{1.24a}$$

同理可得另一固有空間

$$W_\beta = \left\{ C_\beta \begin{bmatrix} \beta \\ 1 \end{bmatrix} \,\middle|\, C_\beta \in \mathbb{R} \right\} \tag{1.24b}$$

將 W_α, W_β 表爲如此完全是爲了讀者之方便, 其表示法並沒有唯一。通常變數變換之矩陣 P 是將固有向量 $\begin{bmatrix} \beta \\ 1 \end{bmatrix}, \begin{bmatrix} \alpha \\ 1 \end{bmatrix}$ 按照順序排好而來, 若取特別的固有向量 $\frac{-1}{\alpha-\beta} \begin{bmatrix} \beta \\ 1 \end{bmatrix}, \frac{1}{\alpha-\beta} \begin{bmatrix} \alpha \\ 1 \end{bmatrix}$, 即

$$P = \frac{1}{\alpha - \beta} \begin{bmatrix} -\beta & \alpha \\ -1 & 1 \end{bmatrix} \tag{1.25}$$

反矩陣 P^{-1} 之求法可由解聯立方程組而來, 令

$$\begin{bmatrix} x \\ y \end{bmatrix} = \frac{1}{\alpha - \beta} \begin{bmatrix} -\beta & \alpha \\ -1 & 1 \end{bmatrix} = P \begin{bmatrix} a \\ b \end{bmatrix}$$

或

$$\begin{cases} x = \frac{1}{\alpha-\beta}(-\beta a + \alpha b) \\ y = \frac{1}{\alpha-\beta}(-a + b) \end{cases}$$

該聯立方程組之解爲

$$\begin{cases} a = x - \alpha y \\ b = x - \beta y \end{cases} \iff \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} 1 & -\alpha \\ 1 & -\beta \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}$$

這相當於說

$$P^{-1} = \begin{bmatrix} 1 & -\alpha \\ 1 & -\beta \end{bmatrix} \tag{1.25$'$}$$

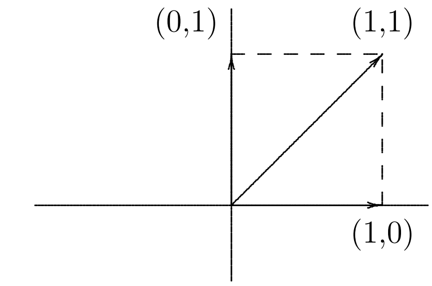
這正是 (1.20) 式所表示的反矩陣P^{-1}, 其實解聯立方程組之過程就是在求反矩陣! 其本質就是數學分析中出名的“反函數定理”。矩陣$P = \begin{bmatrix} \beta & \alpha \\ 1 & 1 \end{bmatrix}$之幾何意義爲

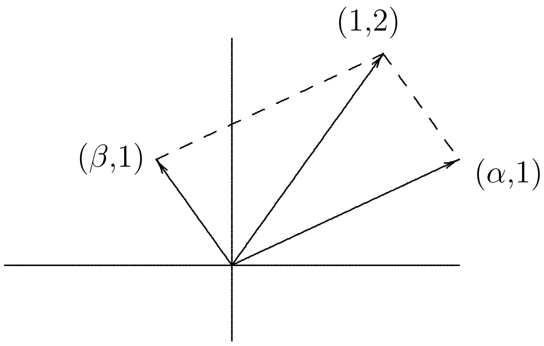
$$P \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} \beta & \alpha \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} \beta \\ 1 \end{bmatrix}$$

$$P \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} \beta & \alpha \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} \alpha \\ 1 \end{bmatrix}$$

$$P \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} \beta & \alpha \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} \beta + \alpha \\ 2 \end{bmatrix} = \begin{bmatrix} \beta \\ 1 \end{bmatrix} + \begin{bmatrix} \alpha \\ 1 \end{bmatrix}$$

$$= P\begin{bmatrix}1\\0\end{bmatrix} + P\begin{bmatrix}0\\1\end{bmatrix}$$





即將平行四邊形的邊帶到邊, 對角線帶到對角線, 換句話說, 就是保持直線與保持平行四邊形這個性質在數學上稱爲 線性(linear)。矩陣P的性質還不只如此, 兩個行向量 (即固有向量) 是互相垂直的, 這可由內積而得

$$\begin{bmatrix}\beta\\1\end{bmatrix}\cdot\begin{bmatrix}\alpha\\1\end{bmatrix} = (\beta,1)\cdot(\alpha,1) = 1+\alpha\beta = 0$$

這同時也說明了

$$\mathbb{R}^2 = W_\alpha \oplus W_\beta;$$
$$W_\beta = W_\alpha^\perp = \left\{\begin{bmatrix}x\\y\end{bmatrix}\in\mathbb{R}^2 \,\middle|\, \begin{bmatrix}x\\y\end{bmatrix}\cdot\begin{bmatrix}\alpha\\1\end{bmatrix}=0\right\}$$

要特別提醒的是矩陣P並沒有唯一, 原因是我們有無窮多解。矩陣P之求法並沒有限定由 (1.20) 式而得, 利用固有值, 固有向量以及矩陣之運算法則亦可得, 令 $\vec{x}_\alpha \in W_\alpha, \vec{x}_\beta \in W_\beta$, 爲矩陣$A$之固有向量

$$A\vec{x}_\alpha = \alpha\vec{x}_\alpha, \quad A\vec{x}_\beta = \beta\vec{x}_\beta,$$

則

$$\begin{aligned} A[\vec{x}_\beta, \vec{x}_\alpha] &= [A\vec{x}_\beta, A\vec{x}_\alpha] \\ &= [\beta\vec{x}_\beta, \alpha\vec{x}_\alpha] \\ &= [\vec{x}_\beta, \vec{x}_\alpha]\begin{bmatrix}\beta & 0\\0 & \alpha\end{bmatrix} \qquad (1.26) \end{aligned}$$

或

$$AQ = Q\Lambda, \quad Q = [\vec{x}_\beta, \vec{x}_\alpha] \qquad (1.26')$$

如果取特別的 $\vec{x}_\beta = \frac{1}{\alpha-\beta}\begin{bmatrix}-\beta\\-1\end{bmatrix}$, $\vec{x}_\alpha = \frac{1}{\alpha-\beta}\begin{bmatrix}\alpha\\1\end{bmatrix}$ 則 $Q = P$爲了方便底下我們默認 $\vec{x}_\beta = \frac{1}{\alpha-\beta}\begin{bmatrix}-\beta\\-1\end{bmatrix}$, $\vec{x}_\alpha = \frac{1}{\alpha-\beta}\begin{bmatrix}\alpha\\1\end{bmatrix}$。(1.26) 式的好處在於你僅需知道固有值、固有向量, 而不受維數之限制, 因此一般線性代數是從此處談起, 但依筆者之體會, 則建議由 (1.8) 式著手, 因爲該式隱含了固有值的原始意義。

α, β爲特徵多項式$\lambda^2 - \lambda - 1 = 0$之兩個根, 由公式知

$$\alpha = \frac{1+\sqrt{5}}{2} \approx 1.618,$$
$$\beta = \frac{1-\sqrt{5}}{2} \approx -0.618, \qquad (1.27)$$

α 就是鼎鼎大名的黃金分割。由 (1.12), (1.21) 式知

$$\begin{aligned}
&\vec{x}_{n+1}\\
&= A^{n+1}\vec{x}_0 = P\Lambda^{n+1}P^{-1}\vec{x}_0\\
&= \frac{1}{\alpha-\beta}\begin{bmatrix} -\beta & \alpha \\ -1 & 1 \end{bmatrix}\begin{bmatrix} -\beta^{n+1} & 0 \\ 0 & \alpha^{n+1} \end{bmatrix}\\
&\quad \cdot \begin{bmatrix} 1 & -\alpha \\ 1 & -\beta \end{bmatrix}\begin{bmatrix} a_1 \\ a_0 \end{bmatrix}\\
&= \frac{1}{\alpha-\beta}\\
&\quad \cdot \begin{bmatrix} \alpha^{n+2}-\beta^{n+2} & \alpha\beta^{n+2}-\alpha^{n+2}\beta \\ \alpha^{n+1}-\beta^{n+1} & \alpha\beta^{n+1}-\alpha^{n+1}\beta \end{bmatrix}\begin{bmatrix} a_1 \\ a_0 \end{bmatrix}
\end{aligned} \tag{1.28}$$

如果取第一項就是 (1.10) 式

$$\begin{aligned}
a_{n+2} &= \frac{1}{\alpha-\beta}\Big[a_1(\alpha^{n+2}-\beta^{n+2})\\
&\quad +a_0(\alpha\beta^{n+2}-\alpha^{n+2}\beta)\Big]\\
&= \frac{1}{\alpha-\beta}\Big[\alpha^{n+2}(a_1-\beta a_0)\\
&\quad -\beta^{n+2}(a_1-\alpha a_0)\Big]
\end{aligned} \tag{1.29}$$

當然, 若你只是在意答案, 那麼便沒有再談下去的必要。偉大的牛頓曾如此自我比喻 —我只不過是個小孩來到眞理的大海, 在此逗留, 偶而撿到幾顆貝殼吧! 許許多多的人來到了眞理的大海, 然而卻不曾“逗留”, 以致讓眞理從旁擦身而過。在此建議諸位略爲停留、思索, 難道我們的故事到此就結束了嗎? 如果學習數學只是要得到一個答案, 那麼你還不懂數學。許多時候過程比答案要來的重要。

將$(a_1,a_0)=(1,0)$, $\alpha=\frac{1+\sqrt{5}}{2}$, $\beta=\frac{1-\sqrt{5}}{2}$ 代入 (1.29) 式得, 例如,a_{100}爲

$$\begin{aligned}
a_{100} &= \frac{1}{\sqrt{5}}\Big[(\frac{1+\sqrt{5}}{2})^{100}-(\frac{1-\sqrt{5}}{2})^{100}\Big]\\
&\approx 3.54\times 10^{20}
\end{aligned} \tag{1.30}$$

諸位可能會懷疑 (1.30) 式這個數有根號又有分數, 實在很訝異, 但放心由 Fibonacci 數列之法則 $a_{n+2}=a_{n+1}+a_n$, $a_0=0$,$a_1=1$, 因此所有的 Fibonacci 數必定是整數, 另外

$$\begin{aligned}
&\frac{1-\sqrt{5}}{2} < \frac{1}{2}\\
&\Rightarrow (\frac{1-\sqrt{5}}{2})^{100} \ll 1
\end{aligned}$$

這告訴我們:

$$\begin{aligned}
&\text{第}k\text{個 Fibonacci 數}\\
&= \text{最靠近}\tfrac{1}{\sqrt{5}}(\tfrac{1+\sqrt{5}}{2})^k \text{ 之整數}
\end{aligned} \tag{1.31}$$

另外我們看一下 Fibonacci 數之比例

$$\frac{a_6}{a_5}=\frac{8}{5}=1.6$$

因此可相信的是 $\frac{a_{101}}{a_{100}}$更靠近 $\frac{1}{2}(1+\sqrt{5})$這個數就是前面提過古希臘人稱之爲黃金分割 (golden mean), 也因此爲何一個長方形之長寬比例爲1.618 : 1 看起來特別的美。

我們可以將 (1.28) 式表示得更漂亮, 首先將 $\vec{x}_0 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ 表爲固有向量 $\vec{x}_\beta = \frac{1}{\alpha-\beta}\begin{bmatrix} -\beta \\ -1 \end{bmatrix}$, $\vec{x}_\alpha = \frac{1}{\alpha-\beta}\begin{bmatrix} \alpha \\ 1 \end{bmatrix}$ 之線性組合 (linear combination)

$$\begin{aligned}
\vec{x}_0 &= \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \frac{1}{\alpha-\beta}\begin{bmatrix} -\beta \\ -1 \end{bmatrix} + \frac{1}{\alpha-\beta}\begin{bmatrix} \alpha \\ 1 \end{bmatrix}\\
&= \vec{x}_\beta + \vec{x}_\alpha
\end{aligned}$$

因此

$$\begin{aligned}
\vec{x}_1 &= A\vec{x}_0 = A(\vec{x}_\beta+\vec{x}_\alpha) = \beta\vec{x}_\beta+\alpha\vec{x}_\alpha\\
\vec{x}_{n+1} &= A^{n+1}\vec{x}_0 = \beta^{n+1}\vec{x}_\beta+\alpha^{n+1}\vec{x}_\alpha
\end{aligned}$$

固有向量$\vec{x}_\beta, \vec{x}_\alpha$原先是分開就一直是分開, 這情形就如在平面$x$–$y$座標,$x$座標歸$x$座標 y座標歸y座標, 而現在是推廣到$\vec{x}_\beta$–$\vec{x}_\alpha$座標吧! 座標是變換了, 但似乎有某些東西是不變的, 例如

(i) $A = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}$ 與 $\Lambda = \begin{bmatrix} \beta & 0 \\ 0 & \alpha \end{bmatrix}$ 有相同之固有值β, α

(ii) $\mathrm{tr}A = 1 + 0 = 1 = \beta + \alpha = \mathrm{tr}\Lambda$

(iii) $\det A = \begin{vmatrix} 1 & 1 \\ 1 & 0 \end{vmatrix} = -1 = \alpha\beta = \begin{vmatrix} \beta & 0 \\ 0 & \alpha \end{vmatrix} = \det\Lambda$

關於這一連串有趣的事, 我們就留給線性代數吧!

2. 從極值的角度

我們從求極大極小值來看固有值, 習慣上還是從題目開始

例題: 試求二次型$A(x, y) = 4x^2 + 8y^2$之極值, 其中x, y滿足$x^2 + y^2 = 1$?

解:

(法 I): 利用$x^2 + y^2 = 1$之關係可得

$$\begin{aligned} A(x, y) &= 4x^2 + 8y^2 \\ &= 4(x^2 + y^2) + 4y^2 \\ &= 4 + 4y^2 \end{aligned}$$

但$-1 \le y \le 1$, 因此$4 \le A(x, y) \le 8$即

$$\max_{x^2+y^2=1} A(x, y) = 8, \quad (y = \pm 1)$$
$$\min_{x^2+y^2=1} A(x, y) = 4, \quad (y = 0)$$

(法 II): 利用 Lagrange multiplier 此問題可化爲求

$$F(x, y, \lambda) = 4x^2 + 8y^2 - \lambda(x^2 + y^2 - 1) \tag{2.1}$$

之極值的問題, 分別對x, y微分得

$$\begin{cases} \frac{\partial F}{\partial x} = 8x - 2\lambda x = 0 \\ \frac{\partial F}{\partial y} = 16y - 2\lambda y = 0 \end{cases} \tag{2.2}$$

這相當於

$$\begin{bmatrix} 4-\lambda & 0 \\ 0 & 8-\lambda \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \tag{2.3}$$

但$(x, y)^t = (0, 0)^t$並無法滿足 $x^2 + y^2 = 1$, 因此上式除了 $(0, 0)^t$解之外應還有其他解, 故

$$\det\begin{bmatrix} 4-\lambda & 0 \\ 0 & 8-\lambda \end{bmatrix} = \begin{vmatrix} 4-\lambda & 0 \\ 0 & 8-\lambda \end{vmatrix} = (\lambda-4)(\lambda-8) = 0$$

這就是前面所提的特徵多項式, 而λ則爲固有值, 分別是

$$\lambda = 4 \quad 或 \quad \lambda = 8$$

此時其對應之固有向量爲

$$\lambda = 4 \quad \Longrightarrow \quad (x, y) = (1, 0)^t$$
$$\lambda = 8 \quad \Longrightarrow \quad (x, y) = (0, 1)^t$$

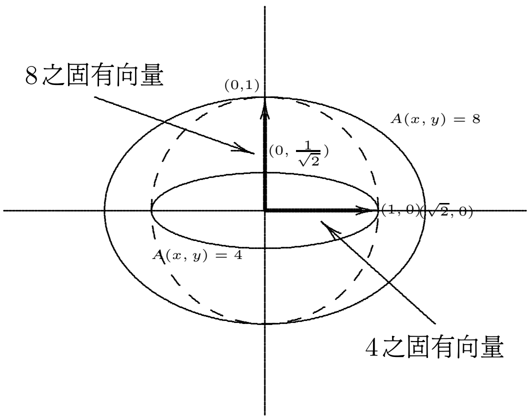
故

$$\max_{x^2+y^2=1} A(x, y) = 4 \cdot 0 + 8 \cdot 1 = 8$$
$$\min_{x^2+y^2=1} A(x, y) = 4 \cdot 1 + 8 \cdot 0 = 4$$

我們看看其幾何意義最大、最小值所對應之橢圓爲

$$\begin{cases} 4x^2 + 8y^2 = 8 \\ 4x^2 + 8y^2 = 4 \end{cases}$$

$$\Longleftrightarrow \begin{cases} (\frac{x}{\sqrt{2}})^2 + y^2 = 1 \\ x^2 + (\frac{y}{\frac{1}{\sqrt{2}}})^2 = 1 \end{cases} \tag{2.4}$$



爲何4, 8會是固有值呢? 我們將二次型表爲

$$4x^2 + 8y^2 = [x, y] \begin{bmatrix} 4 & 0 \\ 0 & 8 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}$$

$$\Leftrightarrow \begin{array}{c|cc} & \mathrm{x} & \mathrm{y} \\ \hline \mathrm{x} & 4 & 0 \\ \mathrm{y} & 0 & 8 \end{array} \Leftrightarrow \begin{bmatrix} 4 & 0 \\ 0 & 8 \end{bmatrix} \tag{2.5}$$

顯然4, 8爲矩陣 $\begin{bmatrix} 4 & 0 \\ 0 & 8 \end{bmatrix}$之固有值, 對於更一般的二次型則爲

$$ax^2+2bxy+cy^2=(x, y) \begin{bmatrix} a & b \\ b & c \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}$$

$$\Leftrightarrow \begin{array}{c|cc} & \mathrm{x} & \mathrm{y} \\ \hline \mathrm{x} & \mathrm{a} & \mathrm{b} \\ \mathrm{y} & \mathrm{b} & \mathrm{c} \end{array} \Leftrightarrow \begin{bmatrix} a & b \\ b & c \end{bmatrix} \tag{2.6}$$

這告訴我們研究二次型 $A(x, y) = ax^2 + 2bxy + cy^2$ 相當於探討對稱矩陣 $A = \begin{bmatrix} a & b \\ b & c \end{bmatrix}$, 我們同樣可問相同的問題。

例題: 試求二次型 $A(x, y) = ax^2 + 2bxy + cy^2$ 之極值, 其中 $x^2 + y^2 = 1$?

依照前面的經驗可猜測的是其極值爲固有值, 而$(x, y)^t$則爲固有向量。

解: 利用 Lagrange multiplier 將此問題化爲

$$F(x, y, \lambda)=ax^2+2bxy+cy^2-\lambda(x^2+y^2-1) \tag{2.7}$$

之極值問題分別微分得

$$\begin{cases} \frac{\partial F}{\partial x} = 2ax + 2by - 2\lambda x = 0 \\ \frac{\partial F}{\partial y} = 2bx + 2cy - 2\lambda y = 0 \end{cases} \tag{2.8}$$

或

$$\begin{bmatrix} a-\lambda & b \\ b & c-\lambda \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \tag{2.9}$$

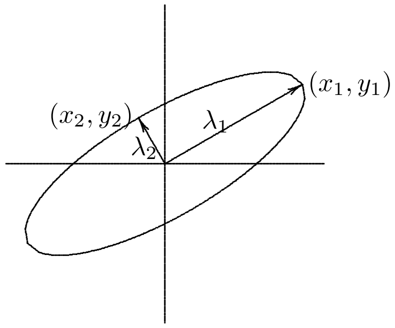
但$(x, y)^t = (0, 0)^t$並不滿足$x^2 + y^2 = 1$, 故行列式必爲0

$$\begin{aligned} & \det \begin{bmatrix} a-\lambda & b \\ b & c-\lambda \end{bmatrix} \\ = & \begin{vmatrix} a-\lambda & b \\ b & c-\lambda \end{vmatrix} \\ = & \lambda^2 - (a+c)\lambda + ac - b^2 = 0 \end{aligned}$$

設兩個固有值爲 $\lambda_1 < \lambda_2$, 而固有向量爲 $\begin{bmatrix} x_1 \\ y_1 \end{bmatrix}, \begin{bmatrix} x_2 \\ y_2 \end{bmatrix}$, 由標準式$\lambda_1 x^2 + \lambda_2 y^2$可知

$$\frac{\lambda_1 x^2 + \lambda_2 y^2}{x^2 + y^2} \geq \lambda_1$$
$$\Longrightarrow \min \frac{\lambda_1 x^2 + \lambda_2 y^2}{x^2 + y^2} = \lambda_1$$

其他細節我們就留給讀者。



例題: 試求$x^2 + y^2$之最大值、最小值, 其中x, y滿足$3x^2 + 4xy + 3y^2 = 1$。

解: 這個題目形式上與前一題雖然不一樣, 但本質上是相同的。既然我們已經習慣前一題之方法, 那麼就想辦法將問題轉化爲前一題之形式, 畢竟人是很難改變的。我們從複雜的一項開始

$$3x^2 + 4xy + 3y^2 = [x, y]\begin{bmatrix} 3 & 2 \\ 2 & 3 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix}$$

矩陣 $A = \begin{bmatrix} 3 & 2 \\ 2 & 3 \end{bmatrix}$ 之特徵多項式爲

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

固有值爲$1, 5$而固有向量爲

$$\frac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ -1 \end{bmatrix}, \quad \frac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

令

$$P = \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{-1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix}$$

這個矩陣P滿足

$$P^t = \begin{bmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix} = P^{-1}$$

$$AP = P\begin{bmatrix} 1 & 0 \\ 0 & 5 \end{bmatrix} \Longrightarrow A = P\begin{bmatrix} 1 & 0 \\ 0 & 5 \end{bmatrix}P^t$$

利用此式可將原二次型變形爲

$$\begin{aligned} &3x^2 + 4xy + 3y^2 \\ &= [x, y]\begin{bmatrix} 3 & 2 \\ 2 & 3 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} \\ &= [x, y]P\begin{bmatrix} 1 & 0 \\ 0 & 5 \end{bmatrix}P^t\begin{bmatrix} x \\ y \end{bmatrix} \end{aligned}$$

由上式知 (極自然) 可考慮變數變換

$$\begin{bmatrix} X \\ Y \end{bmatrix} = P^t\begin{bmatrix} x \\ y \end{bmatrix} = P^{-1}\begin{bmatrix} x \\ y \end{bmatrix},$$

$$\begin{bmatrix} x \\ y \end{bmatrix} = P\begin{bmatrix} X \\ Y \end{bmatrix} = \begin{bmatrix} \frac{1}{\sqrt{2}}X + \frac{1}{\sqrt{2}}Y \\ -\frac{1}{\sqrt{2}}X + \frac{1}{\sqrt{2}}Y \end{bmatrix}$$

取轉置矩陣得

$$[X, Y] = [x, y]P$$

因此

$$\begin{aligned} &3x^2 + 4xy + 3y^2 \\ &= [X, Y]\begin{bmatrix} 1 & 0 \\ 0 & 5 \end{bmatrix}\begin{bmatrix} X \\ Y \end{bmatrix} \\ &= X^2 + 5Y^2 = 1 \end{aligned}$$

而且

$$\begin{aligned} x^2 + y^2 &= (\frac{1}{\sqrt{2}}X + \frac{1}{\sqrt{2}}Y)^2 \\ &\quad + (-\frac{1}{\sqrt{2}}X + \frac{1}{\sqrt{2}}Y)^2 \\ &= X^2 + Y^2 \\ &= X^2 + \frac{1}{5}(1 - X^2) \\ &= \frac{1}{5} + \frac{4}{5}X^2 \end{aligned}$$

但X之範圍爲$-1 \le X \le 1$, 故

$$\frac{1}{5} \le X^2 + Y^2 \le \frac{1}{5} + \frac{4}{5} = 1$$

極小值爲 $\frac{1}{5}$產生在 $(X,Y)=(0,\pm\sqrt{\frac{1}{5}})$, 而極大值爲 1則產生在 $(X,Y)=(\pm 1,0)$。

3. 微分方程的角度

固有值的另一個看法就是微分方程。考慮平面上一運動系統, 其位置座標爲 $(x(t), y(t))$, 其速度 (就是對 t 之微分) $(\frac{dx}{dt},\frac{dy}{dt})$ 滿足微分方程組

$$\begin{cases}\frac{dx}{dt}=ax+by, & x(0)=x_0\\ \frac{dy}{dt}=cx+dy, & y(0)=y_0\end{cases} \tag{3.1}$$

如果引進向量的概念, 則 (3.1) 式可視爲一維的問題, 令

$$X(t)=\begin{bmatrix}x(t)\\ y(t)\end{bmatrix},$$
$$X(0)=\begin{bmatrix}x(0)\\ y(0)\end{bmatrix}=\begin{bmatrix}x_0\\ y_0\end{bmatrix} \tag{3.2}$$

則 (3.1) 可改寫爲

$$\frac{dX}{dt}=\begin{bmatrix}a & b\\ c & d\end{bmatrix}X\equiv AX,$$
$$X(0)=\begin{bmatrix}x_0\\ y_0\end{bmatrix}\equiv X_0 \tag{3.3}$$

對於n維空間, 我們也有與 (3.3) 式相似的結果。雖然 (3.3) 式很簡潔、清爽, 但重要的還是如何將解求出來, 否則只是畫餅充飢無濟於事。爲明瞭 (3.3) 式, 我們還是回到一維的情形, 數學的思維法則都是由簡而繁, 否則眞難想像一個人如何去思考。

$$\frac{dx}{dt}=at,\quad x(0)=x_0 \tag{3.4}$$

利用分離變數法, 而後積分得

$$\frac{dx}{x}=adt\Longrightarrow\int\frac{1}{x}\,dx=\int adt$$
$$\Longrightarrow x(t)=e^{at}x_0 \tag{3.5}$$

由 (3.5) 之結果, 我們也期待 (3.3) 式之解也能表爲 (3.5) 式之形式, 因此我們作如下之假設

$$\begin{cases}x(t)=e^{\lambda t}u\\ y(t)=e^{\lambda t}v\end{cases} \tag{3.6}$$

(u,v)待求。將 (3.6) 待入 (3.3) 得

$$\begin{cases}\lambda e^{\lambda t}u=ae^{\lambda t}u+be^{\lambda t}v\\ \lambda e^{\lambda t}v=ce^{\lambda t}u+de^{\lambda t}v\end{cases}$$

整理得

$$\begin{cases}au+bv=\lambda u\\ cu+dv=\lambda v,\end{cases} AU=\lambda U,\quad U=\begin{bmatrix}u\\ v\end{bmatrix} \tag{3.7}$$

這告訴我們解微分方程組 (3.3) 相當於解代數方程組 (3.7)。

微分方程 $\Longleftrightarrow$ 代數方程

這就是數學的精神, 以已知的知識(代數方程)來解決未知的問題(微分方程), 並將既有的知識推廣。

λ就是矩陣A的固有值, 而$U=\begin{bmatrix}u\\ v\end{bmatrix}$ 則是固有向量, 我們做個題目

例題: 試解微分方程

$$X'(t)=\begin{bmatrix}1 & 1\\ 4 & 1\end{bmatrix}X$$

解: 矩陣 $A=\begin{bmatrix}1 & 1\\ 4 & 1\end{bmatrix}$ 之特徵多項式爲

$$\begin{aligned}f(\lambda)&=\det(\lambda I-A)\\ &=\begin{vmatrix}\lambda-1 & -1\\ -4 & \lambda-1\end{vmatrix}\\ &=(\lambda-3)(\lambda+1)\end{aligned}$$

因此固有值與固有向量分別爲

$$\lambda = 3, -1, \quad U_1 = \begin{bmatrix} 1 \\ 2 \end{bmatrix}, \; U_2 = \begin{bmatrix} 1 \\ -2 \end{bmatrix},$$

故 $\begin{bmatrix} 1 \\ 2 \end{bmatrix} e^{3t}, \begin{bmatrix} 1 \\ -2 \end{bmatrix} e^{-t}$ 都是原微分方程之解, 再由重疊原理 (即線性) 知一般解爲

$$X(t) = c_1 \begin{bmatrix} 1 \\ 2 \end{bmatrix} e^{3t} + c_2 \begin{bmatrix} 1 \\ -2 \end{bmatrix} e^{-t}$$

參考資料:

1. 林琦焜: 線性代數的基本定理, 數學傳播 (中央研究院數學所), Vol. 74, p.47–54(1995).
2. Gilbert Strang, Introduction to Linear Algebra, Wellesley–Cambridge Press (1993), (美亞代理).

—本文作者任教於國立成功大學數學系—