從一題矩陣的試題談起

林倉億

國立台南一中

1. 前言

「$A=\begin{pmatrix}1&4\\3&2\end{pmatrix}$，二階方陣 X 、Y 滿足 $X+Y=I$ 且 $XY=O$，其中 $I=\begin{pmatrix}1&0\\0&1\end{pmatrix}$、$O=\begin{pmatrix}0&0\\0&0\end{pmatrix}$。若存在實數 $a>b$ 使得 $A=aX+bY$，求 a 、b 之值。」

上述這個題目最近頗「紅」，不僅出現在友校本學期高二的段考試卷中（敘述略有出入），也出現在本學期高三的某份模擬考考卷中（敘述略有出入）。由於這題「殺很大」，難倒不少高二、高三學生，再加上幾位教師對試卷提供的解法都不甚滿意，認爲應該有更深層的意義與解法，所以，引起筆者的興趣想一探究竟。在嘗試不同的解法並查閱書籍之後，赫然發現從這個題目竟可以連結到線性代數中的重要定理，因而撰此文和大家分享。

2. 不同的解法

首先，出題老師提供的解法大致如解法一：

解法一：（努力計算）

$$\begin{cases}X+Y=I\\aX+bY=A\end{cases}\Rightarrow\ \text{解聯立得}\begin{cases}X=\dfrac{A-bI}{a-b}\\Y=\dfrac{A-aI}{b-a}\end{cases}\text{，因爲 } XY=O\text{，故}$$

$$\begin{pmatrix}0&0\\0&0\end{pmatrix}=\frac{1}{-(a-b)^2}\begin{pmatrix}1-b&4\\3&2-b\end{pmatrix}\begin{pmatrix}1-a&4\\3&2-a\end{pmatrix}=\frac{1}{-(a-b)^2}\begin{pmatrix}(1-b)(1-a)+12&4(3-a-b)\\3(3-a-b)&12+(2-b)(2-a)\end{pmatrix}$$

$$\Rightarrow\ b=3-a\text{ 代入 }(1-b)(1-a)+12=0\ \Rightarrow\ a^2-3a-10=0\ \Rightarrow\ (a,b)=(5,-2)\ or\ (-2,5)$$

又 $a>b$，故 $(a,b)=(5,-2)$。

上述解法就是將 X 、Y 用 A 表示後，再利用 $XY=O$ 解出 a 、b。解法一基本上都是在做計算，看不出此題背後的數學結構爲何，而這也正是幾位老師不滿意的原因。然而，這一題特別之處在於在 X 、Y 未知的情況下，竟可解出唯一一組 a、b 之值，再代回後可得到唯一的 $X=\begin{pmatrix}\frac{3}{7}&\frac{4}{7}\\\frac{3}{7}&\frac{4}{7}\end{pmatrix}$ 與

$Y=\begin{pmatrix}\frac{4}{7} & \frac{-4}{7}\\ \frac{-3}{7} & \frac{3}{7}\end{pmatrix}$，可見無論是$a$、$b$還是$X$、$Y$，對$A$來說，一定都有特別的意義！我們可以用另外一個方法解它，其背後的意義與結構就逐漸明朗了。

解法二：(凱萊—漢米爾頓定理)

$$A=aX+bY=aX+b(I-X)=(a-b)X+bI \Rightarrow (a-b)X=A-bI \cdots\cdots(1)$$

$$A=aX+bY=a(I-Y)+bY=aI+(b-a)Y \Rightarrow (b-a)Y=A-aI \cdots\cdots(2)$$

$(1)\times(2) \Rightarrow O=(A-bI)(A-aI)=A^2-(a+b)A+abI$，又由 Cayley-Hamilton Theorem 知

$A^2-(1+2)A+(1\cdot 2-3\cdot 4)I=O \Rightarrow A^2-3A-10I=O$，比較係數可得

$\begin{cases}a+b=3\\ ab=-10\end{cases}$，又$a>b$，故$\begin{cases}a=5\\ b=-2\end{cases}$。

利用凱萊—漢米爾頓定理 (Cayley-Hamilton Theorem)不僅可以更輕易地求出a、b，還可以看出a、b其實就是A的特徵值。事實上就此題來說，我們也可以在不使用凱萊—漢米爾頓定理的情況下，推導出a、b是A的特徵值。

解法三：（特徵值）

$$A=aX+bY=aX+b(I-X)=(a-b)X+bI \Rightarrow (a-b)X=A-bI \cdots\cdots(1)$$

$$A=aX+bY=a(I-Y)+bY=aI+(b-a)Y \Rightarrow (b-a)Y=A-aI \cdots\cdots(2)$$

$XY=O \Rightarrow$ 若X可逆，則$Y=X^{-1}O=O$，又$X+Y=I \Rightarrow X=I$，如此一來就不可能將A寫成$aX+bY=aI$，故X要不可逆。同理，Y也要不可逆。

X、Y均不可逆 $\Rightarrow \det X=\det Y=0$，故$\det(A-aI)=\det(A-bI)=0$

即a、b為A的特徵值 $\Rightarrow \begin{vmatrix}1-\lambda & 4\\ 3 & 2-\lambda\end{vmatrix}=\lambda^2-3\lambda-10=0 \Rightarrow \lambda=5\ or\ -2$，

又$a>b$，故$(a,b)=(5,-2)$。

至此，無論是如何求a、b之值，或是a、b的意義，可說是完全解決了。接下來就是探討$A=aX+bY$中X、Y的意義，畢竟要滿足$X+Y=I$與$XY=O$，又要能和特徵值配對在一起，這樣的X、Y一定是非同小可！讓我們看解法四：

解法四：（對角化）

$$\det(A-\lambda I)=\begin{vmatrix}1-\lambda & 4\\ 3 & 2-\lambda\end{vmatrix}=\lambda^2-3\lambda-10=0 \Rightarrow \lambda=5 \ or \ -2$$

當 $\lambda=5$ 時，可求得一特徵向量 $\begin{pmatrix}1\\1\end{pmatrix}$；當 $\lambda=-2$ 時，可求得一特徵向量 $\begin{pmatrix}-4\\3\end{pmatrix}$。

令 $P=\begin{pmatrix}1 & -4\\ 1 & 3\end{pmatrix} \Rightarrow P^{-1}=\begin{pmatrix}\frac{3}{7} & \frac{4}{7}\\ \frac{-1}{7} & \frac{1}{7}\end{pmatrix}$，則 $P^{-1}AP=\begin{pmatrix}5 & 0\\ 0 & -2\end{pmatrix}$，即 $A=P\begin{pmatrix}5 & 0\\ 0 & -2\end{pmatrix}P^{-1}$

$$\Rightarrow A=P\left[\begin{pmatrix}5 & 0\\ 0 & 0\end{pmatrix}+\begin{pmatrix}0 & 0\\ 0 & -2\end{pmatrix}\right]P^{-1}=P\begin{pmatrix}5 & 0\\ 0 & 0\end{pmatrix}P^{-1}+P\begin{pmatrix}0 & 0\\ 0 & -2\end{pmatrix}P^{-1}$$

$$=5\cdot P\begin{pmatrix}1 & 0\\ 0 & 0\end{pmatrix}P^{-1}+(-2)\cdot P\begin{pmatrix}0 & 0\\ 0 & 1\end{pmatrix}P^{-1}=5\begin{pmatrix}\frac{3}{7} & \frac{4}{7}\\ \frac{3}{7} & \frac{4}{7}\end{pmatrix}+(-2)\begin{pmatrix}\frac{4}{7} & \frac{-4}{7}\\ \frac{-3}{7} & \frac{3}{7}\end{pmatrix}$$

其中 $\begin{pmatrix}\frac{3}{7} & \frac{4}{7}\\ \frac{3}{7} & \frac{4}{7}\end{pmatrix}+\begin{pmatrix}\frac{4}{7} & \frac{-4}{7}\\ \frac{-3}{7} & \frac{3}{7}\end{pmatrix}=I$ 且 $\begin{pmatrix}\frac{3}{7} & \frac{4}{7}\\ \frac{3}{7} & \frac{4}{7}\end{pmatrix}\begin{pmatrix}\frac{4}{7} & \frac{-4}{7}\\ \frac{-3}{7} & \frac{3}{7}\end{pmatrix}=O$，故 $X=\begin{pmatrix}\frac{3}{7} & \frac{4}{7}\\ \frac{3}{7} & \frac{4}{7}\end{pmatrix}$ 與 $Y=\begin{pmatrix}\frac{4}{7} & \frac{-4}{7}\\ \frac{-3}{7} & \frac{3}{7}\end{pmatrix}$。

由解法四可以看出，依此方法解出的 X 、Y 一定會滿足 $X+Y=I$ 與 $XY=O$ ，更特別的，

$X^2=\left[P\begin{pmatrix}1 & 0\\ 0 & 0\end{pmatrix}P^{-1}\right]\cdot\left[P\begin{pmatrix}1 & 0\\ 0 & 0\end{pmatrix}P^{-1}\right]=X$ ，$Y^2=\left[P\begin{pmatrix}0 & 0\\ 0 & 1\end{pmatrix}P^{-1}\right]\cdot\left[P\begin{pmatrix}0 & 0\\ 0 & 1\end{pmatrix}P^{-1}\right]=Y$，也就是 X 、Y 都是冪等矩陣 (idempotent matrix)。這麼漂亮的結果，正是線性代數中的重要的定理「譜定理 (Spectral Theorem)」。

3. 譜定理 (Spectral Theorem)

在介紹譜定理前，先幫讀者複習幾個名詞。若一個向量空間 V 分解成兩個互補 (complementary)[1]的子空間 U 、W 的和，則稱 $U+W$ 是 V 的「直和 (direct sum)」，記作：

$V=U\oplus W$ ，即 $\forall v\in V$ ，$\exists u\in U,\ w\in W$ 使得 $v=u+w$ 。

此時存在唯一的線性算子 P (linear operator，不熟悉的讀者不妨先將它想成一個方陣)使得 $PV=U$ ，稱 P 將 V 沿著 W 投影到 U ，簡稱爲「投影 (projector)」。P 會滿足：

[1] U 、W 互補意指 $U+W=V$ 且 $U\cap W=0$ 。

(1) $P^2 = P$ ； (2) $I - P$ 也是投影，將 V 沿著 U 投影到 W ； (3) $P(I-P) = O$ 。

若有一組向量空間 V 上的投影 $P_1,\ P_2,\cdots\cdots,\ P_k$ 滿足：$P_i \cdot P_j = O,\ \forall i \neq j$ 且 $P_1 + P_2 + \cdots\cdots + P_k = I$ ，則稱 $\{P_1, P_2, \cdots\cdots, P_k\}$ 是 V 上的一組完整投影集 (complete set of projection)。

舉個簡單的例子，$V = R^3$，$U = xy$ 平面，$W = z$ 軸，則 $V = U \oplus W$ 。這是因爲 $\forall (a,b,c) \in R^3$ ，$(a,b,c) = (a,b,0) + (0,0,c)$ ，其中 $(a,b,0) \in U$ ，$(0,0,c) \in W$ 。令投影 $P_1 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}$ ，則

$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} a \\ b \\ c \end{pmatrix} = \begin{pmatrix} a \\ b \\ 0 \end{pmatrix}$ ，即 $P_1 V = U$ ，因此投影 P_1 的幾何意義就是將 $V = R^3$ 中的每個向量沿著 $W = z$ 軸投影到 $U = xy$ 平面。同樣地，令 $P_2 = I - P = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}$ 也是一個投影。$\{P_1, P_2\}$ 就是 $V = R^3$ 的一組完整投影集。熟悉這些名詞後，我們就可以來看譜定理了。

譜定理 (Spectral Theorem)

$\lambda_1,\ \lambda_2, \cdots\cdots,\ \lambda_k \in R$ 爲 n 階方陣 A 的相異特徵值，則：

A 可對角化 $\Leftrightarrow$ 存在一組完整投影集 $\{P_1, P_2, \cdots\cdots, P_k\}$ 使得 $A = \lambda_1 P_1 + \lambda_2 P_2 + \cdots\cdots + \lambda_k P_k$ 。

上述的特徵值 $\lambda_1,\ \lambda_2, \cdots\cdots,\ \lambda_k$ 亦可稱爲 A 的「譜 (spectrum)」，而 $A = \lambda_1 P_1 + \lambda_2 P_2 + \cdots\cdots + \lambda_k P_k$ 就稱爲 A 的「譜分解 (spectral decompositon)」。

回到本文一開始的題目，這下子終於真相大白了，原來該題背後的數學意義，就是將矩陣 $A = \begin{pmatrix} 1 & 4 \\ 3 & 2 \end{pmatrix}$ 作譜分解。譜定理的證明其實就是求 $P_1, P_2, \cdots\cdots, P_k$ 的過程，在此略去，僅再提供一個例子說明作法。相信各位讀者藉由解法四與下列例子的解法，應該就可以知道譜定理爲何成立了。

例：求$A=\begin{pmatrix}2&2&1\\1&3&1\\1&2&2\end{pmatrix}$的譜分解。

解：$\det(A-\lambda I)=\begin{vmatrix}2-\lambda&2&1\\1&3-\lambda&1\\1&2&2-\lambda\end{vmatrix}=(5-\lambda)(1-\lambda)^2=0 \Rightarrow \lambda=5\ \ or\ \ 1$

取$\lambda=5$的特徵向量$\begin{pmatrix}1\\1\\1\end{pmatrix}$，$\lambda=1$的特徵向量$\begin{pmatrix}2\\-1\\0\end{pmatrix}$與$\begin{pmatrix}1\\0\\-1\end{pmatrix}$，令$P=\begin{pmatrix}1&2&1\\1&-1&0\\1&0&-1\end{pmatrix}$

則$P^{-1}AP=\begin{pmatrix}5&0&0\\0&1&0\\0&0&1\end{pmatrix} \Rightarrow A=P\begin{pmatrix}5&0&0\\0&1&0\\0&0&1\end{pmatrix}P^{-1}=P\left(\begin{pmatrix}5&0&0\\0&0&0\\0&0&0\end{pmatrix}+\begin{pmatrix}0&0&0\\0&1&0\\0&0&1\end{pmatrix}\right)P^{-1}$

$$=P\begin{pmatrix}5&0&0\\0&0&0\\0&0&0\end{pmatrix}P^{-1}+P\begin{pmatrix}0&0&0\\0&1&0\\0&0&1\end{pmatrix}P^{-1}=5\begin{pmatrix}\frac{1}{4}&\frac{2}{4}&\frac{1}{4}\\\frac{1}{4}&\frac{2}{4}&\frac{1}{4}\\\frac{1}{4}&\frac{2}{4}&\frac{1}{4}\end{pmatrix}+\begin{pmatrix}\frac{3}{4}&\frac{-2}{4}&\frac{-1}{4}\\\frac{-1}{4}&\frac{2}{4}&\frac{-1}{4}\\\frac{-1}{4}&\frac{-2}{4}&\frac{3}{4}\end{pmatrix}$$

4. 後話

探究這一題，就如同登山一樣，一開始在迷霧森林之中，不知道東西南北，只有勞力的付出（努力計算）；等爬到了山腰上的涼亭，眼前開始出現令人驚豔的景色（特徵值）；登頂之後，環顧四面，才看到壯闊的全景（譜定理），教人不禁讚嘆數學的奧妙！從高觀點來看這個題目，看到的是漂亮的數學結構，但對學生來說，差不多就是陷在茫茫計算之中，不知所爲何事。這個題目是否適於出現在試卷之中，見仁見智。筆者並不清楚同出此題的兩位老師在出題時的考量，但個人會選擇避開它，因爲知道它背後有這樣美妙的數學，卻無法和學生分享，這是件何等痛苦的事啊！

參考資料

Meyer, Carl D. (2000). *Matrix Analysis and Applied Linear Algebra*. Philadelphia: Society for Industrial and Applied Mathematics..

廖亦德 (1990).《綜合線性代數》，台北市：巨德出版社。