

線性代數的基本定理

林琦焜

前言:

最近在 American Mathematical Monthly 閱讀到 Gilbert Strang 探討線性代數之文章, 讀後收穫良多, 尤其幾個圖形實在有教學上之價值。在感動之餘想想何不動手, 以 Gilbert Strang 之文章為藍本, 同時把自己讀書與教學之心得將之整理後, 以與中文之讀者一起分享。

此文主要探討的是 Fredholm Alternative 定理, 要提醒的是雖然我們僅在有限的空間上討論, 但實際上都可推廣至無限維空間, 而這就是泛函分析 (functional analysis) 所研究的主題之一。

矩陣的本質:

要瞭解線性代數, 最直接且最有動機莫若於從求聯立方程組的解開始。

$$A\vec{x} = \vec{b} \quad A: R^n \rightarrow R^m \quad (1)$$

其中

$$\vec{x} = (x_1, x_2 \cdots x_n)^T \in R^n$$

$$\vec{b} = (b_1, b_2 \cdots b_m)^T \in R^m$$

在此向量都是以行向量來表示。其中 A 是一個 $m \times n$ 矩陣

$$A = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ a_{21} & \cdots & a_{2n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix} \quad a_{ij} \in R, \quad 1 \leq i \leq m, \quad 1 \leq j \leq n \quad (2)$$

首先我們將矩陣 A 視為向量。(實際上矩陣是向量的推廣)

$$A = [A_1, \dots, A_n] \quad (3)$$

$$A_j = \begin{bmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{mj} \end{bmatrix} \quad 1 \leq j \leq n \quad (4)$$

有了上述之結果, 我們可將 (1) 式的左邊表為向量的線性組合:

$$\begin{aligned} A\vec{x} &= \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ a_{21} & \cdots & a_{2n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \\ &= [A_1, \dots, A_n] \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \\ &= x_1 A_1 + \cdots + x_n A_n = \vec{b} \quad (5) \end{aligned}$$

註 (A): 如果我們將向量 $\begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$ 視為矩陣的話, 則 (5) 式同時也告訴我們, 矩陣乘在右手邊其運算為行運算。(同理乘在左手邊則為列運算。) 而其法則為

$$x_1 \times (\text{第一行}) + x_2 \times (\text{第二行}) + \cdots + x_n \times (\text{第 } n \text{ 行}) \quad (6)$$

註 (B): 由 (5) 式我們也可略窺 “行空間” (column space) 的雛形, 由此角度而言, 求 $A\vec{x} = \vec{b}$ 的解, 相當於求所有 $A_1 \cdots A_n$ 的線性組合正好等於 \vec{b} , 即求 $(x_1 \cdots x_n) \in R^n$ 使得

$$x_1 A_1 + \cdots + x_n A_n = \vec{b}$$

註 (C): (5) 式可幫助我們明白矩陣的結合律, 一般在線性代數的課本是將矩陣視為線性變換 (linear transformation), 因此矩陣的結合律可視為是函數之合成的結合律, 但這種作法, 對學生而言, 幫助並不大。在這裡我們希望藉由 (5) 及一些簡單的基本運算來證明矩陣的結合律

$$A(BC) = (AB)C.$$

由(5)式知向量 $A\vec{x}$ 為矩陣 A 之行向量的線性組合, 利用這個概念, 我們可以對矩陣的乘法有另一個角度的體會, 給定任一矩陣

$$B = [B_1, B_2, \cdots, B_k]$$

B_i 為矩陣 B 之第 i 行向量, 因此矩陣 A 與矩陣 B 之相乘可表為

$$\begin{aligned} AB &= A[B_1, B_2, \cdots, B_k] \\ &= [AB_1, AB_2, \cdots, AB_k] \end{aligned}$$

即矩陣 AB 的第 i 行向量 $(AB)_i$ 為矩陣 A 乘矩陣 B 的第 i 行向量 AB_i 。由 (5) 式知 AB_i 要有意義其先決條件為 B_i 為一 n 維行向量, 即矩陣 B 為一 $n \times k$ 矩陣

$$\begin{aligned} B &= [B_1, B_2, \cdots, B_k] \\ &= \begin{bmatrix} b_{11} & \cdots & b_{1k} \\ \vdots & & \vdots \\ b_{n1} & \cdots & b_{nk} \end{bmatrix} \end{aligned}$$

同理, 矩陣 BC 要有意義為 C 為一 $k \times l$ 矩陣。

$$C = [C_1, \cdots, C_l] = \begin{bmatrix} c_{11} & \cdots & b_{1l} \\ \vdots & & \vdots \\ c_{k1} & \cdots & c_{kl} \end{bmatrix}$$

我們現在考慮矩陣之結合律, 由 (5) 式知

$$\begin{aligned} A(BC_i) &= A(c_{1i}B_1 + \cdots + c_{ki}B_k) \\ &= c_{1i}AB_1 + \cdots + c_{ki}AB_k \\ &= [AB_1, \cdots, AB_k] \begin{bmatrix} c_{1i} \\ \vdots \\ c_{ki} \end{bmatrix} \\ &= [AB_1, \cdots, AB_k]C_i \end{aligned}$$

再次利用 (5) 式可得矩陣之結合律

$$\begin{aligned} A(BC) &= A(B[C_1, \cdots, C_l]) \\ &= A[BC_1, \cdots, BC_l] \\ &= [A(BC_1), \cdots, A(BC_l)] \end{aligned}$$

$$\begin{aligned}
&= [(AB)C_1, \dots, (AB)C_i] \\
&= (AB)[C_1, \dots, C_i] \\
&= (AB)C
\end{aligned}$$

註 (D): (5) 式告訴我們的還不僅如此。在中學階段就熟知 Cramer 公式, 亦可由此式再加點行列式的性質而得, 當然還是從解聯立方程組開始

$$A\vec{x} = \vec{b}$$

此時 A 為一 $n \times n$ 矩陣向量, \vec{x}, \vec{b} 則視為 $n \times 1$ 矩陣, 為著簡便用符號 $[A \leftarrow \vec{b}]$ 表示一 $n \times n$ 矩陣, 其中矩陣 $A = [A_1 \cdots A_n]$ 之第 i 行向量為向量 \vec{b} 所取代, 即

$$[A \leftarrow \vec{b}] = [A_1 \cdots A_{i-1}, \vec{b}, A_{i+1}, \dots, A_n]$$

但由 (5) 式並利用行運算基本上是矩陣乘在右手邊之原則得

$$\begin{aligned}
&[A_1 \cdots A_{i-1}, A\vec{x}, A_{i+1}, \dots, A_n] \\
&= [A_1, \dots, A_{i-1}, x_1 A_1 + \cdots + x_n A_n, A_{i+1}, \\
&\quad \cdots A_n] \\
&= [A_1, \dots, A_{i-1}, A_i, A_{i+1} \cdots A_n] \\
&\quad \times \begin{bmatrix} 1 & 0 & x_1 & 0 & 0 \\ 0 & 1 & x_2 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & x_n & 0 & 1 \end{bmatrix} \\
&= A[A \leftarrow \vec{x}]
\end{aligned}$$

因此聯立方程 $A\vec{x} = \vec{b}$ 可改寫為

$$A[A \leftarrow \vec{x}] = [A \leftarrow \vec{b}]$$

兩邊同時取行列式得

$$(\det A)(\det[A \leftarrow \vec{x}]) = \det[A \leftarrow \vec{b}]$$

由 Laplace 展開式或行列式的性質知

$$\det[A \leftarrow \vec{x}] = x_i, \text{ 故}$$

$$x_i = \frac{\det[A \leftarrow \vec{b}]}{\det A}$$

這就是 Cramer 公式。

基本定理:

習慣上, 我們將行空間 (column space) 記為 $R(A)$, 明顯地 $R(A) \subseteq R^m$ 。談了行向量, 行空間自然要提它的孿生兄弟列向量 (row vector), 列空間 (row space), 記為 $R(A^T)$, 另要提的子空間如下

$$N(A) \equiv \{\vec{x} \in R^n | A\vec{x} = 0\} \subseteq R^n \quad (7)$$

$$N(A^T) \equiv \{\vec{y} \in R^n | A^T \vec{y} = 0\} \subseteq R^m \quad (8)$$

底下我們將注意力都集中在這四個子空間, 當然讀者可能會問為何要探討這些子空間, 實際上所有線性代數上的運算與應用皆可經由子空間的瞭解而來, 例如在中學階段所學利用加減消去法, 代入消去法來求聯立方程組的解, 就是無形中已使用到子空間的某些特性, 其中最重要的一個就是維數 (dimension) 的不變性。維數在線性代數中扮演著極重要的角色。

定理 1: (i) $\dim R(A) = \dim R(A^T)$

(ii) $\dim R(A) + \dim N(A) = n$

(i) 式告訴我們行空間 (column space) 與列空間 (row space) 的維數是一樣的, 如此的描敘還是抽象了一些, 最好的方式還是以例子來明瞭定理的意義。其實學數學最好的方法即是從“例子”著手。

例1:

$$A = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 2 & 0 & 0 \\ 0 & 1 & 0 & 1 & 1 \\ 0 & 4 & 3 & 4 & 0 \end{pmatrix}_{4 \times 5}$$

經過化簡得

$$A \sim \left. \begin{pmatrix} 0 & 4 & 3 & 0 & 0 \\ 0 & 0 & 2 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} \right\} 3$$

因此

$$\dim R(A) = \dim R(A^T) = 3$$

$$\dim N(A) = 5 - 3 = 2$$

$$\dim N(A^T) = 4 - 3 = 1$$

對於一般 $m \times n$ 矩陣, 經過行運算 (或列運算) 後可容易判別上述的關係式

例2: $A = [A_1 \cdots A_n] \sim [B_1 \cdots B_r, 0 \cdots 0], \{B_1 \cdots B_r\}$ 是線性

獨立, 因此 $\dim R(A) = r$, 由定理1知 $\dim N(A) = n - r$ 。

線性代數另一個基本定理如下

定理2: $N(A) \perp R(A^T)$ 。

這定理告訴我們子空間的正交性 (orthogonality), 其意義與證明也可從聯立方程組的解來視出端倪。

$$\begin{aligned} A\vec{x} &= \begin{bmatrix} \vec{B}_1 \\ \vdots \\ \vec{B}_m \end{bmatrix} \vec{x} \\ &= \begin{bmatrix} \vec{B}_1 \cdot \vec{x} \\ \vdots \\ \vec{B}_m \cdot \vec{x} \end{bmatrix} = \begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix} \end{aligned}$$

取各分量得

$$\vec{B}_1 \cdot \vec{x} = \cdots = \vec{B}_m \cdot \vec{x} = 0$$

\vec{x} 與所有列向量 $\vec{B}_i, (1 \leq i \leq m)$ 垂直

$$\vec{x} \perp \vec{B}_i \quad (1 \leq i \leq m)$$

因此 $\vec{x} \perp \sum_{i=1}^m b_i \vec{B}_i$

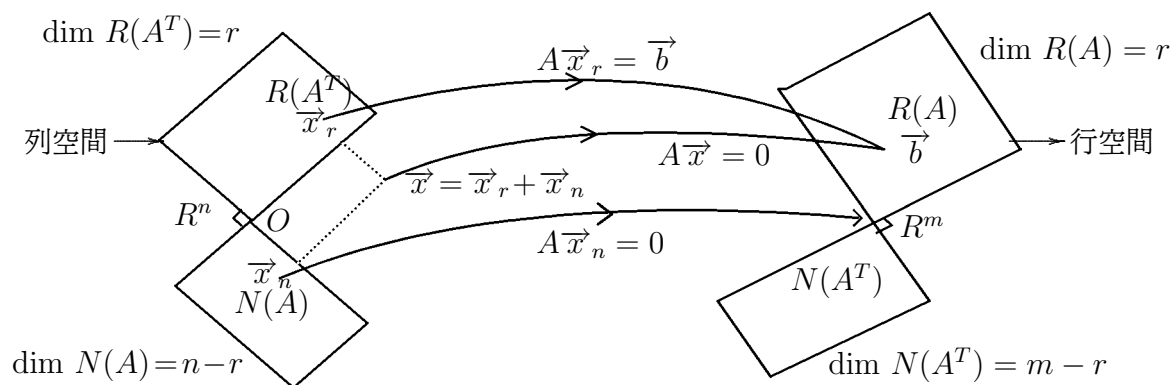
即 $\vec{x} \perp R(A^T), \quad \forall \vec{x} \in N(A)$

所以 $N(A) \perp R(A^T)$

同理, 取轉置 (transpose) 矩陣我們有

定理2': $N(A^T) \perp R(A)$ 。

茲以一個圖形來說明上面二個定理



這圖形說明了幾件事實：

(a) R^n 上的向量 \vec{x} 可分解為兩個互相垂直的向量。

$$\vec{x} = \vec{x}_r + \vec{x}_n, \quad \vec{x}_r \in R(A^T), \quad \vec{x}_n \in N(A) \quad (9)$$

或

$$R^n = R(A^T) \oplus N(A)$$

實際上此分解是唯一的。

(b) A 將 R^n 上的任一向量 \vec{x} 帶到行空間 $R(A)$ ，而將核空間 (null space) 皆帶到 0 向量。

(c) 若 \vec{b} 落在行空間 $R(A)$ ，則聯立方程組 $A\vec{x} = \vec{b}$ 是可解的。換言之，若 \vec{b} 與所有 $N(A^T)$ 的向量垂直 ($\vec{b} \perp N(A^T)$)，則聯立方程組 $A\vec{x} = \vec{b}$ 是可解的。而其解 \vec{x} 可分解為二部份

$$\begin{aligned} \vec{x} &= \vec{x}_p + \vec{x}_h \\ A\vec{x} &= A(\vec{x}_p + \vec{x}_h) = A\vec{x}_p + A\vec{x}_h \end{aligned}$$

$$= \vec{b} + 0 = \vec{b}.$$

\vec{x}_p 為 (1) 式之特解 (particular solution) \vec{x}_h 為 (1) 式之均勻解 (homogeneous solution) 而 $\vec{x}_p + \vec{x}_h$ 為一般解 (general solution), \vec{x}_p 稱為特解是因爲

$$\vec{x}_p \perp \vec{x}_h, \quad \forall \vec{x}_h \in N(A). \quad (10)$$

另外順便一提的是所有均勻解所成的空間正好是核空間 (null space)

$$N(A) = \{ \vec{x}_h \in R^n \mid A\vec{x}_h = 0 \}$$

註：上面 (c) 所談的在無窮維空間亦有相類似的結果，其實這就是泛函分析 (functional analysis) 或積分方程 (integral equation) 裡著名的 *Fredholm Alternative* 定理。

註：若將 A 視為一微分算子 (differential operator), 則 (c) 所言在微分

方程也有相對應的結果。讀者若有興趣可自行驗證。事實上，我們所提的線性代數的基本定理是可直接推廣至無窮維空間。

如果值域為一維，即 $m = 1$

$$A: R^n \rightarrow R \quad (11)$$

則 A 為一有界線性泛函 (bounded linear functional)。由 Riesz 表現定理告訴我們， A 可表為一內積之形式。

Riesz 表現定理: A 為一有界線性泛函從 R^n 映到 R 。 $A: R^n \rightarrow R$ 有界，線性。則存在唯一 $\vec{y} \in R^n$ 使得

$$A(\vec{x}) = (\vec{x}, \vec{y}) \quad \forall \vec{x} \in R^n. \quad (12)$$

利用上述之結果，再加上一點正交投影之概念可容易地證明且明白 Riesz 表現定理的幾何意義：

定理證明: (不失一般性可設 $A \neq 0$)

$\forall \vec{x} \in R^n$ 可表為

$$\vec{x} = \vec{x}_r + \vec{x}_n, \quad \vec{x}_r \in R(A^T), \quad \vec{x}_n \in N(A)$$

$$\begin{aligned} A(\vec{x}) &= A(\vec{x}_r + \vec{x}_n) = A(\vec{x}_r) + A(\vec{x}_n) \\ &= A(\vec{x}_r). \end{aligned}$$

上式告訴我們 A 的值完全由在 $R(A^T)$ 上的值所決定。令 \mathcal{P} 為一 R^n 在 $R(A^T)$ 上之正交投影，則我們有

$$\mathcal{P}(\vec{x}) = \mathcal{P}(\vec{x}_r + \vec{x}_n) = \vec{x}_r.$$

因此

$$A(\vec{x}) = A(\vec{x}_r) = A(\mathcal{P}(\vec{x})).$$

所以，如果我們能決定 $\mathcal{P}(\vec{x})$ 之長像，則 $A(\vec{x})$ 之形像也跟著決定。

而現在因為 A 為一線性泛函， $m = 1$ 且 $A \neq 0$ ，因此由定理 1 知

$$\dim R(A^T) = \dim R(A) = 1.$$

令 $\vec{z} \in R(A^T)$ ， $\vec{z} \neq 0$ ，則

$$R(A^T) = \langle \vec{z} \rangle = \{\alpha \vec{z} | \alpha \in R\}.$$

由正交投影知

$$\begin{aligned} \mathcal{P}(x) &= \frac{\langle \vec{x}, \vec{z} \rangle}{\|\vec{z}\|^2} \vec{z} \\ \text{所以 } A(\vec{x}) &= A(\vec{x}_r) = A(\mathcal{P}(\vec{x})) \\ &= A\left(\frac{\langle \vec{x}, \vec{z} \rangle}{\|\vec{z}\|^2} \cdot \vec{z}\right) \\ &= \frac{\langle \vec{x}, \vec{z} \rangle}{\|\vec{z}\|^2} \cdot A(\vec{z}) \\ &= \left\langle \vec{x}, \frac{A(\vec{z})}{\|\vec{z}\|^2} \cdot \vec{z} \right\rangle \end{aligned}$$

令 $y = \frac{A(\vec{z})}{\|\vec{z}\|^2} \vec{z}$ 即為所求。

註1: 這個證明方法並沒有維數的限制，對一般的內積空間 (inner product space) 皆可。

註2: 因為用到投影的概念，因此在無形中，我們已經將最短距離或變分學的概念注入這定理中。而事實上這定理本身已具有變分原理 (variational principle) 的內涵。在偏微分方程 (P.D.E) 這定理是弱解 (weak solution) 存在的最好證明工具呢！

最小二乘方:

關於 Fredholm Alternative 定理的另一個重要應用便是最小二乘方 (least-square)。由前面之理論知, 若 \vec{b} 不屬於行空間 (column space) 則聯立方程組 (矩陣 A 並沒有限制一定是方陣)

$$A\vec{x} = \vec{b}$$

無法求得其解。但在現實情形與應用, 期待一個非奇異方陣是不實際的。因此我們需要有一些方法以面對殘酷的現實。

我們的問題如下: 試求一直線: $b = C + Dt$ 或一拋物線 $b = C + Dt + Et^2$ 通過 $(t_1, b_1) \cdots (t_m, b_m)$ 這些點?

首先我們將問題表為聯立方程組, 即

$$\text{或 } \begin{cases} C + Dt_1 = b_1 \\ \vdots \\ C + Dt_m = b_m \\ C + Dt_1 + Et_1^2 = b_1 \\ \vdots \\ C + Dt_m + Et_m^2 = b_m \end{cases} \quad (13)$$

乍見之下該聯立方程組為二個未知數 (C, D) (或3個未知數 (C, D, E)) 卻要滿足 m 個方程式, 這顯然是要求過多。上面之聯立方程組可表為

$$\begin{cases} A\vec{x} = \vec{b} \\ \vec{x} = (C, D)^t \text{ 或 } \vec{x} = (C, D, E)^t \end{cases} \quad (14)$$

而 A 則為一 $m \times 2$ 或 $m \times 3$ 矩陣, 因此 $A\vec{x} = \vec{b}$ 要有解, 唯一可能的是這些點都

落在同一直線上 (或拋物線上), 即 \vec{b} 是落在矩陣 A 之行空間上。這種向量 \vec{b} 是限制過大, 所以我們問問題的方式需略作改變, 即求向量 \vec{x} 使得

$$A\vec{x} - \vec{b} \text{ 之長度為最短} \quad (15)$$

或者是求向量 \vec{x} 使得

$$\eta = (A\vec{x} - \vec{b}) \cdot (A\vec{x} - \vec{b}) = |A\vec{x} - \vec{b}|^2 \text{ 之值為最小} \quad (16)$$

由於 $A\vec{x}$ 始終是行向量, 因此上面之問題相當於是

$$\text{求向量 } \vec{b} \text{ 至行空間之最短距離?} \quad (17)$$

而眾所周知, 求最短距離當然是與投影 (projection) 有關聯。若 \vec{p} 為向量 \vec{b} 在行空間上之投影, 則向量 $\vec{e} = \vec{b} - \vec{p}$ 為所求, 而且 \vec{e} 與行空間垂直, 即 $\vec{e} \in \text{Ker}(A^T)$ 。

$$A^T \vec{e} = A^T (\vec{b} - \vec{p}) = 0 \quad (18)$$

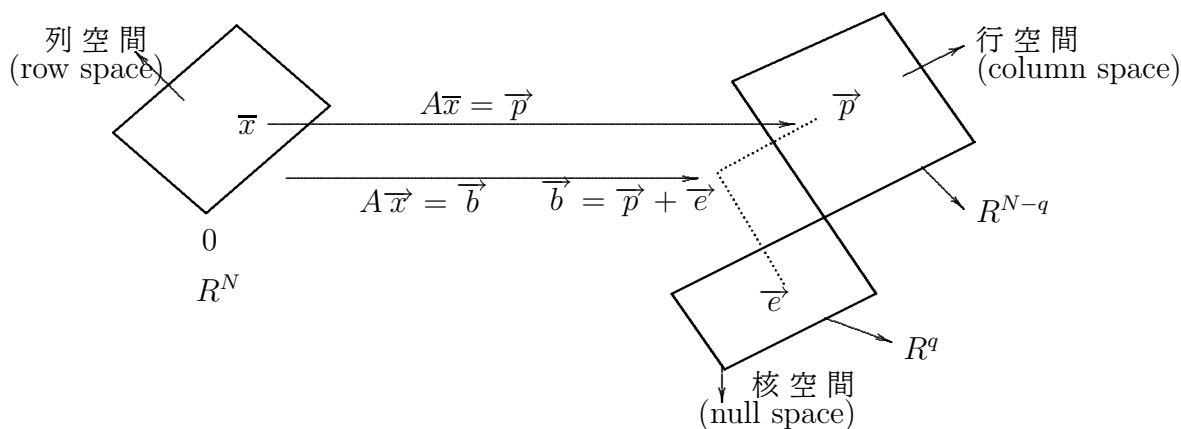
令

$$\vec{p} = A\vec{x} \quad (19)$$

因此 (18) 等於告訴我們

$$A^T A\vec{x} = A^T \vec{b} \quad (20)$$

(20) 式就是通常所說的正則方程式 (normal equation), 上面之結果可以圖來表示。



正則方程式亦可由微分而來
(最短距離當然是與微分有關)

$$\begin{aligned} \eta(\vec{x}) &= \eta(x_1 \cdots x_N) \\ &= (A\vec{x} - \vec{b}) \cdot (A\vec{x} - \vec{b}) \\ &= (A^T A \vec{x} - 2A^T \vec{b}) \cdot \vec{x} + \vec{b} \cdot \vec{b} \end{aligned} \quad (21)$$

則 $\frac{\partial \eta}{\partial x_i} = 0 \quad i = 1, \dots, N$

得 $A^T A \vec{x} = A^T \vec{b}$

爲了方便, 令

$$\tilde{A} = A^T A \quad \tilde{b} = A^T \vec{b} \quad (22)$$

因此正則方程式可改寫爲

$$\tilde{A} \vec{x} = \tilde{b} \quad (23)$$

該方程式有利之處在不管原矩陣 A 是否爲方陣, $\tilde{A} = A^T A$ 一定是個“對稱方陣”, 故前面的 Fredholm-Alternative 定理現在就可派上用場,

即 (23) 要有解其充分必要條件爲

$$\tilde{b} \perp N(\tilde{A}^T)$$

不失一般性可設

$$\begin{cases} \dim N(\tilde{A}) = \dim N(\tilde{A}^T) = q \\ \dim R(\tilde{A}) = N - q \end{cases} \quad (24)$$

因爲 \tilde{A} 爲對稱矩陣, 故存在正交矩陣 P , $P \cdot P^T = I$, 使得

$$\begin{aligned} P^T \tilde{A} P &= \begin{pmatrix} 0 & 0 \\ 0 & D \end{pmatrix} \\ &= P^T A^T (AP) = (AP)^T (AP) \end{aligned} \quad (25)$$

D 爲一 $(N - q) \times (N - q)$ 對角矩陣, 即

$$D = \begin{pmatrix} \lambda_{q+1} & & 0 \\ & \ddots & \\ 0 & & \lambda_N \end{pmatrix} \quad (26)$$

且 $\det D \neq 0$, 我們現在決定矩陣 P 。可設

$$P = [\vec{\varphi}_1, \dots, \vec{\varphi}_q, \vec{\xi}_{q+1}, \dots, \vec{\xi}_N] \quad (27)$$

$\vec{\varphi}_i, 1 \leq i \leq q, \vec{\xi}_j, q+1 \leq j \leq N$ 為行向量, 其中 (可由 (24) 式看出來)

$$\text{span}\{\vec{\varphi}_1 \dots \vec{\varphi}_q\} = \text{Ker}(A)$$

從 (24) 可得

$$\begin{aligned} AP &= [(AP)_1, \dots, (AP)_q, (AP)_{q+1}, \dots, (AP)_N] \\ &= [A\vec{\varphi}_1 \dots A\vec{\varphi}_q, A\vec{\xi}_{q+1}, \dots, A\vec{\xi}_N] \\ &= [0, \dots, 0, A\vec{\xi}_{q+1}, \dots, A\vec{\xi}_N] \quad (28) \end{aligned}$$

因此再一次由 (24) 可知

$$\begin{cases} (AP)_i \cdot (AP)_j = (AP)_i^T (AP)_j = 0 \\ \quad i \neq j \quad i = j \leq q \\ (AP)_j \cdot (AP)_j = (AP)_j^T (AP)_j = \lambda_j \\ \quad j > q \end{cases} \quad (29)$$

而且

$$\begin{aligned} [\text{Ker}(A^T)]^\perp &= [A\vec{\xi}_{q+1}, \dots, A\vec{\xi}_N] \\ &= [(AP)_{q+1}, \dots, (AP)_N] \quad (30) \end{aligned}$$

現在定義

$$\vec{z} = P^T \vec{x} = (z_1 \dots z_N)^T \quad (31)$$

故由正則方程式 (28), (29) 知

$$\begin{pmatrix} 0 & 0 \\ 0 & D \end{pmatrix} \vec{z} \quad (32)$$

$$\begin{aligned} &= P^T \tilde{A} P \vec{z} = P^T A^T A P P^T \vec{x} \\ &= P^T (A^T A \vec{x}) = P^T A^T \vec{b} = (AP)^T \vec{b} \\ &= (0, \dots, 0, A\xi_{q+1}, \dots, A\xi_N)^T \vec{b} \\ &= (0, \dots, 0, (A\xi_{q+1})^T \vec{b}, \dots, (A\xi_N)^T \vec{b}) \end{aligned}$$

比較各座標知

$$\begin{cases} \lambda_i z_i = A \vec{\xi}_i \cdot \vec{b} & q+1 \leq i \leq N \\ z_1 \dots z_q \text{ 則為任意數} \end{cases} \quad (33)$$

故待求之 \vec{x} 可表為

$$\vec{x} = \sum_{i=1}^q z_i \vec{\varphi}_i + \sum_{i=q+1}^N \frac{(A \vec{\xi}_i \cdot \vec{b})}{\lambda_i} \vec{\xi}_i \quad (34)$$

直接檢驗可得

$$A^T A \vec{x} = A^T \vec{b}$$

如果取

$$\vec{b}_R = \vec{b} - \sum_{j=1}^q (\vec{b} \cdot \vec{\varphi}_j) \vec{\varphi}_j \quad (35)$$

則可得

引理: $A\vec{x} = \vec{b}_R$.

綜合上面之討論, 我們整理如下:

方程式 $A\vec{x} = \vec{b}$ 有解之充分必要條件為

定理: (Fredholm Alternative)

方程組 $A\vec{x} = \vec{b}$ 有解之充分必要條件為 $\vec{b} \in \text{Ker}(A^T)^\perp$, 而且其解可表為

$$x = \sum_{i=1}^q z_i \vec{\varphi}_i + \sum_{i=q+1}^N \frac{(A \vec{\xi}_i \cdot \vec{b})}{\lambda_i} \vec{\xi}_i$$

其中 $z_1 \cdots z_q$ 為任意常數。如果 $\vec{b} = \vec{b}_R$ ，則存在唯一 $\vec{w}^* \in \text{Ker}(A^T)^\perp$ 使得 $A\vec{w}^* = \vec{b}_R$ 。

參考資料

1. Gilbert Strang, *The Fundamental Theorem of Linear Algebra American*, Mathematical Monthly, 100 (1993), 848

- 855.

2. Gilbert Strang, *Linear Algebra and Its Applications*, 3rd ed., Harcourt Brace Jovanovich (1988).
3. Gilbert Strang, *Introduction to Linear Algebra*, Wellesley - Cambridge Press (1993).

—本文作者任教於成功大學數學系—