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 之文章爲藍本, 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 Altenative 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。

$$A\overrightarrow{x} = \overrightarrow{b} \qquad A : R^n \to R^m \quad (1)$$

其中

$$\begin{aligned} \overrightarrow{x} &= (x_1, x_2 \cdots x_n)^T \in R^n \\ \overrightarrow{b} &= (b_1, b_2 \cdots b_m)^T \in R^m \end{aligned}$$

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

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

首先我們將矩陣 A 視爲向量。(實際上矩陣是向量的推廣)

$$A = [A_1, \ldots, A_n] \qquad (3)$$
$$A_j = \begin{bmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{mj} \end{bmatrix} \qquad 1 \le j \le n \qquad (4)$$

有了上述之結果, 我們可將 (1) 式的左邊表爲向量的線性組合:

$$A\overrightarrow{x} = \begin{bmatrix} a_{11} & \ldots & a_{1n} \\ a_{21} & \ldots & a_{2n} \\ \vdots & \vdots & \vdots \\ a_{m1} & \ldots & a_{mn} \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ \vdots \\ x_n \end{bmatrix}$$

$$
\begin{aligned}
&= [A_1, \ldots, A_n]\begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \\
&= x_1A_1 + \cdots + x_nA_n = \overrightarrow{b} \qquad (5)
\end{aligned}
$$

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

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

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 $A\overrightarrow{x} = \overrightarrow{b}$ 的解, 相當於求所有 $A_1 \cdots A_n$ 的線性組合正好等於 $\overrightarrow{b}$, 即求 $(x_1 \cdots x_n) \in R^n$ 使得

$$x_1A_1 + \cdots + x_nA_n = \overrightarrow{b}$$

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爲線性變換 (linear transformation), 因此矩陣的結合律可視爲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証明矩陣的結合律

$$A(BC) = (AB)C。$$

由(5) 式知向量 $A\overrightarrow{x}$ 爲矩陣 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 爲矩陣 B 之第 i 行向量, 因此矩陣 A 與矩陣 B 之相乘可表爲

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

即矩陣 AB 的第 i 行向量 $(AB)_i$ 爲矩陣 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爲 B_i 爲一 n 維行向量, 即矩陣 B 爲一 $n \times k$ 矩陣

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

同理, 矩陣 BC 要有意義爲 C 爲一 $k \times l$ 矩陣。

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

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

$$A(BC) = A(B[C_1, \cdots, C_l])$$

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

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

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

此時 A 爲一 $n \times n$ 矩陣向量, $\overrightarrow{x}$, $\overrightarrow{b}$ 則視爲 $n \times 1$ 矩陣, 爲著簡便用符號 $[A \overset{i}{\longleftarrow} \overrightarrow{b}]$ 表示一 $n \times n$ 矩陣, 其中矩陣 $A = [A_1 \cdots A_n]$ 之第 i 行向量爲向量 $\overrightarrow{b}$ 所取代, 即

$$[A \overset{i}{\longleftarrow} \overrightarrow{b}] = [A_1 \cdots A_{i-1}, \overrightarrow{b}, A_{i+1}, \cdots, A_n]$$

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

$$
\begin{aligned}
& [A_1 \cdots A_{i-1}, A\overrightarrow{x}, A_{i+1}, \cdots A_n] \\
= \ & [A_1, \cdots A_{i-1}, x_1A_1 + \cdots + x_nA_n, A_{i+1}, \\
& \quad \cdots A_n] \\
= \ & [A_1, \cdots A_{i-1}, A_i, A_{i+1} \cdots A_n] \\
& \times \begin{bmatrix} 1 & 0 & x_1 & 0 & 0 \\ 0 & 1 & x_2 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & x_n & 0 & 1 \end{bmatrix} \\
= \ & A[I \overset{i}{\longleftarrow} \overrightarrow{x}]
\end{aligned}
$$

因此聯立方程 $A\overrightarrow{x} = \overrightarrow{b}$ 可改寫爲

$$A[I \overset{i}{\longleftarrow} \overrightarrow{x}] = [A \overset{i}{\longleftarrow} \overrightarrow{b}]$$

兩邊同時取行列式得

$$(\det A)(\det[I \overset{i}{\longleftarrow} \overrightarrow{x}]) = \det[A \overset{i}{\longleftarrow} \overrightarrow{b}]$$

由 Laplace 展開式或行列式的性質知

$$\det[I \overset{i}{\longleftarrow} \overrightarrow{x}] = x_i, \ \text{故}$$

$$x_i = \frac{\det[A \overset{i}{\longleftarrow} \overrightarrow{b}]}{\det A}$$

這就是 Cramer 公式。

基本定理:

習慣上, 我們將行空間 (column space) 記爲 $R(A)$, 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 vecter), 列空間 (row space), 記爲 $R(A^T)$, 另要提的子空間如下

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

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

底下我們將注意力都集中在這四個子空間, 當然讀者可能會問爲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 (colum space) 與列空間 (row space) 的維數是一樣的, 如此的描敍還是抽象了一些, 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

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

$\underbrace{\qquad\qquad}_{3}$

因此

$$\begin{aligned}\dim R(A)&=\dim R(A^T)=3\\ \dim N(A)&=5-3=2\\ \dim N(A^T)&=4-3=1\end{aligned}$$

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覘出端倪。

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

取各分量得

$\overrightarrow{B}_1\cdot\overrightarrow{x}=\cdots=\overrightarrow{B}_n\cdot\overrightarrow{x}=0$

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

$$\overrightarrow{x}\perp\overrightarrow{B}_i\ (1\le i\le m)$$

因此 $\overrightarrow{x}\perp\sum_{i=1}^{n}b_i\overrightarrow{B}_i$

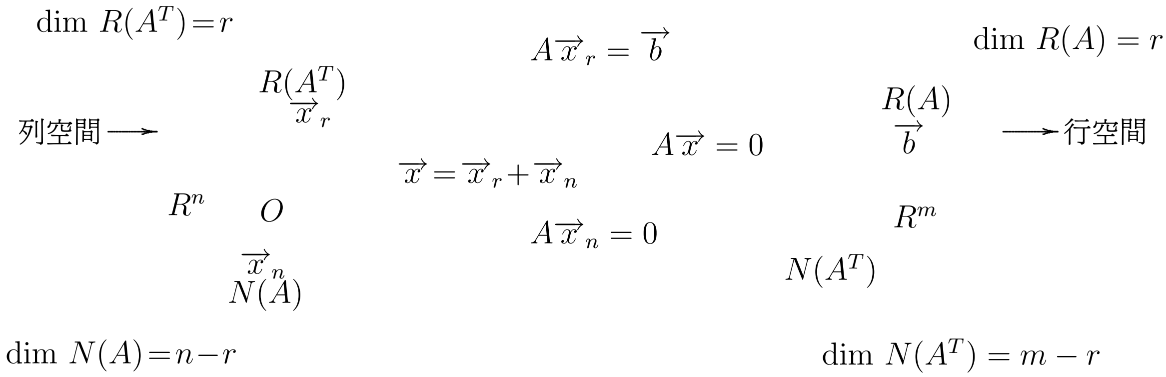
即 $\overrightarrow{x}\perp R(A^T)$, $\quad\forall\overrightarrow{x}\in N(A)$

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 $\overrightarrow{x}$ 可 分 解 爲 兩 個 互 相 垂 直 的 向 量。

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

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 $\overrightarrow{x}$ 帶 到 行 空 間 $R(A)$, 而 將 核 空 間 (null space) 皆 帶 到 0 向 量。

(c) 若 $\overrightarrow{b}$ 落 在 行 空 間 $R(A)$, 則 聯 立 方 程 組 $A\overrightarrow{x} = \overrightarrow{b}$ 是 可 解 的。換 言 之, 若 $\overrightarrow{b}$ 與 所 有 $N(A^T)$ 的 向 量 垂 直 ($\overrightarrow{b} \perp N(A^T)$) , 則 聯 立 方 程 組 $A\overrightarrow{x} = \overrightarrow{b}$ 是 可 解 的。而 其 解 $\overrightarrow{x}$ 可 分 解 爲 二 部 份

$$\begin{aligned} \overrightarrow{x} &= \overrightarrow{x}_p + \overrightarrow{x}_h \\ A\overrightarrow{x} &= A(\overrightarrow{x}_p + \overrightarrow{x}_h) = A\overrightarrow{x}_p + A\overrightarrow{x}_h \\ &= \overrightarrow{b} + 0 = \overrightarrow{b}。 \end{aligned}$$

$\overrightarrow{x}_p$ 爲 (1) 式 之 特 解 (particular solution) $\overrightarrow{x}_h$ 爲 (1) 式 之 均 勻 解 (homogeneous solution) 而 $\overrightarrow{x}_p + \overrightarrow{x}_h$ 爲 一 般 解 (general solution), $\overrightarrow{x}_p$ 稱 爲 特 解 是 因 爲

$$\overrightarrow{x}_p \perp \overrightarrow{x}_h\ , \qquad \forall \overrightarrow{x}_h \in N(A)。 \tag{10}$$

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)

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

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 爲 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。

如果值域爲一維, 即 $m=1$

$$A: R^n \to R \tag{11}$$

則A爲一有界線性泛函 (bounded linear functional)。由 Riesz 表現定理告訴我們, A 可表爲一內積之形式。

Riesz **表現定理**: A 爲一有界線性泛函從 R^n 映到 R。$A: R^n \to R$ 有界, 線性。則存在唯一 $\overrightarrow{y} \in R^n$ 使得

$$A(\overrightarrow{x}) = (\overrightarrow{x}, \overrightarrow{y}) \qquad \forall\ \overrightarrow{x} \in R^n。 \tag{12}$$

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 $A \neq 0$)

$$\begin{aligned} \forall\ \overrightarrow{x} &\in \mathbf{R}^n \quad \text{可表爲} \\ \overrightarrow{x} &= \overrightarrow{x}_r + \overrightarrow{x}_n, \overrightarrow{x}_r \in R(A^T), \overrightarrow{x}_n \in N(A) \end{aligned}$$

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

上式告訴我們 A 的值完全由在 $R(A^T)$ 上的值所決定。令 $\mathcal{P}$ 爲一 R^n 在 $R(A^T)$ 上之正交投影, 則我們有

$$\mathcal{P}(\overrightarrow{x}) = \mathcal{P}(\overrightarrow{x}_r + \overrightarrow{x}_n) = \overrightarrow{x}_r。$$

因此

$$A(\overrightarrow{x}) = A(\overrightarrow{x}_r) = A(\mathcal{P}(\overrightarrow{x}))。$$

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

而現在因爲 A 爲一線性泛函, $m=1$ 且 $A \neq 0$, 因此由定理1知

$$\dim R(A^T) = \dim R(A) = 1。$$

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

$$R(A^T) = <\overrightarrow{z}> = \{\alpha \overrightarrow{z} | \alpha \in R\}。$$

由正交投影知

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

令 $y \equiv \dfrac{A(\overrightarrow{z})}{\|\overrightarrow{z}\|^2} \overrightarrow{z}$ 即爲所求。

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 produt space) 皆可。

註2: 因爲用到投影的概念, 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 $\overrightarrow{b}$ 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)

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

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。因此我們需要有某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?

首先我們將問題表爲聯立方程組, 即

$$\begin{cases} C + Dt_1 = b_1 \\ \quad\vdots \\ C + Dt_m = b_m \end{cases}$$

或

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

乍見之下該聯立方程組爲二個未知數 (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爲

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

而 A 則爲一 $m \times 2$ 或 $m \times 3$ 矩陣, 因此 $A\overrightarrow{x} = \overrightarrow{b}$ 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 $\overrightarrow{b}$ 是落在矩陣 A 之行空間上。這種向量 $\overrightarrow{b}$ 是限制過大, 所以我們問問題的方式需略作改變, 即求向量 $\overrightarrow{x}$ 使得

$$A\overrightarrow{x} - \overrightarrow{b} \text{之長度爲最短} \tag{15}$$

或者是求向量 $\overrightarrow{x}$ 使得

$$\eta = (A\overrightarrow{x} - \overrightarrow{b}) \cdot (A\overrightarrow{x} - \overrightarrow{b}) = |A\overrightarrow{x} - \overrightarrow{b}|^2 \\ \text{之值爲最小} \tag{16}$$

由於 $A\overrightarrow{x}$ 始終是行向量, 因此上面之問題相當於是

$$\text{求向量}\overrightarrow{b}\text{至行空間之最短距離?} \tag{17}$$

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

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

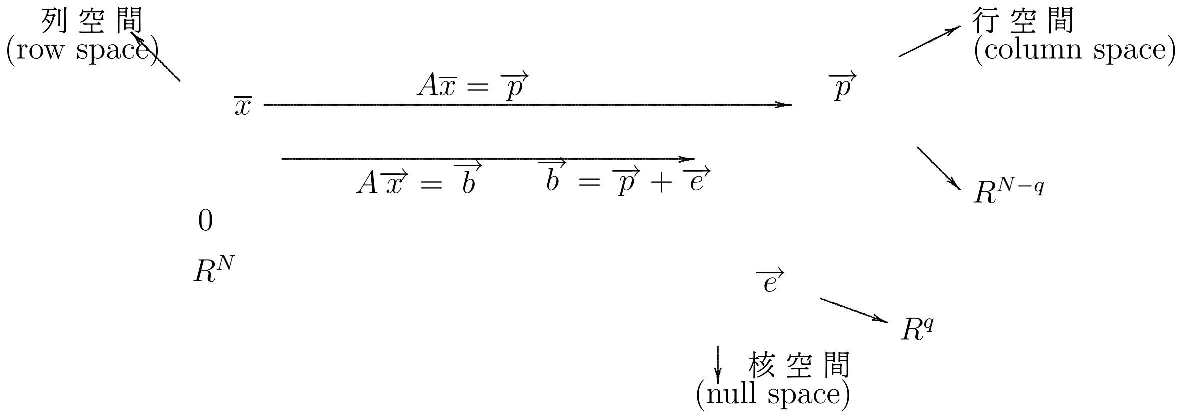
令

$$\overrightarrow{p} = A\overline{x} \tag{19}$$

因此 (18) 等於告訴我們

$$A^TA\overline{x} = A^T\overrightarrow{b} \tag{20}$$

(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)

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

則 $\dfrac{\partial\eta}{\partial x_i}=0 \quad i=1,\cdots N$

得 $A^TA\overrightarrow{x}=A^T\overrightarrow{b}$

為了方便, 令

$$\tilde{A}=A^TA \qquad \tilde{b}=A^T\overrightarrow{b} \qquad (22)$$

因此正則方程式可改寫為

$$\tilde{A}\overline{x}=\tilde{b} \qquad (23)$$

該方程式有利之處在不管原矩陣 A 是否為方陣, $\tilde{A}=A^TA$ 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為

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

不失一般性可設

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

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

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

D 為一 $(N-q)\times(N-q)$ 對角矩陣, 即

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

且 $detD\neq 0$, 我們現在決定矩陣 P。可設

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

$\overrightarrow{\varphi}_i,\ 1 \leq i \leq q$。$\overrightarrow{\xi}_j, q+1 \leq j \leq N$ 爲 行 向 量, 其 中 (可 由 (24) 式 看 出 來)

$$\text{span}\{\overrightarrow{\varphi}_1 \cdots \overrightarrow{\varphi}_q\} = Ker(A)$$

從 (24) 可 得

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

因 此 再 一 次 由 (24) 可 知

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

而 且

$$\begin{aligned} \left[Ker(A^T)\right]^\perp &= [A\overrightarrow{\xi}_{q+1}, \cdots, A\overrightarrow{\xi}_N] \\ &= [(AP)_{q+1}, \cdots, (AP)_N] \end{aligned} \qquad (30)$$

現 在 定 義

$$\overrightarrow{z} = P^T \overrightarrow{x} = (z_1 \cdots z_N)^T \qquad (31)$$

故 由 正 則 方 程 式 (28),(29) 知

$$\begin{aligned} &\begin{pmatrix} 0 & 0 \\ 0 & D \end{pmatrix} \overrightarrow{z} \qquad (32) \\ = \ & P^T \tilde{A} P \overrightarrow{z} = P^T A^T A P P^T \overrightarrow{x} \\ = \ & P^T (A^T A \overrightarrow{x}) = P^T A^T \overrightarrow{b} = (AP)^T \overrightarrow{b} \\ = \ & (0, \cdots, 0, A\xi_{q+1}, \cdots, A\xi_N)^T \overrightarrow{b} \\ = \ & (0, \cdots, 0, (A\xi_{q+1})^T \overrightarrow{b}, \cdots, (A\xi_N)^T \overrightarrow{b}) \end{aligned}$$

比 較 各 座 標 知

$$\begin{cases} \lambda_i z_i = A\overrightarrow{\xi}_i \cdot \overrightarrow{b} \quad q+1 \leq i \leq N \\ z_1 \cdots z_q \text{則 爲 任 意 數} \end{cases} \qquad (33)$$

故 待 求 之 $\overline{x}$ 可 表 爲

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

直 接 檢 驗 可 得

$$A^T A \overline{x} = A^T \overrightarrow{b}$$

如 果 取

$$\overrightarrow{b}_R = \overrightarrow{b} - \sum_{j=1}^{q} (\overrightarrow{b} \cdot \overrightarrow{\varphi}_i) \overrightarrow{\varphi}_i \qquad (35)$$

則 可 得

引理: $A\overline{x} = b_R$。

綜 合 上 面 之 討 論, 我 們 整 理 如 下:
方 程 式 $A\overrightarrow{x} = \overrightarrow{b}$ 有 解 之 充 分 必 要 條 件 爲

定理: (Fredholm Alternative)
方 程 組 $A\overrightarrow{x} = \overrightarrow{b}$ 有 解 之 充 分 必 要 條 件 爲 $\overrightarrow{b} \in Ker(A^T)^\perp$, 而 且 其 解 可 表 爲

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

其 中 $z_1 \cdots z_q$ 爲 任 意 常 數。如 果 $\overrightarrow{b} = \overrightarrow{b}_R$, 則 存 在 唯 一 $\overrightarrow{w}^* \in Ker(A^T)^\perp$ 使 得 $A\overrightarrow{w}^* = \overrightarrow{b}_R$。

參考資料

1. Gilbert Strang, *The Fundamental Theorm 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—