3

基本矩陣運算與線性聯立方程式

數學上我們經常需要處理解線性聯立方程式的問題, 這一章的內容將圍繞在關於線性聯立方程式的理論而鋪陳。單元 3.1 一開始先從解三元一次聯立方程式為例, 從解題中重新利用矩陣註記的方式將解題的過程精煉出來, 並且把解聯立方程式的操作形態分類而得到三種類型的基本列運算並抽象化引出三類基本行運算; 進一步地, 在分別引進三類基本矩陣後, 我們可以把基本列運算與基本行運算和左乘與右乘基本矩陣聯繫, 這麼一來, 只要對基本矩陣充分認識後就能對線性聯立方程式的數學理論研究清楚。

單元 3.2 是介紹矩陣的秩, 並探索線性變換的秩與矩陣的秩之間的關聯。一方面, 我們會證明將矩陣左乘或是右乘一個可逆矩陣的話仍然保秩, 於是將此結果應用到基本矩陣再解讀回線性聯立方程式的求解就能推得所有解聯立的操作也都保秩。至於一個矩陣的秩要如何解讀? 又要如何快速看出矩陣的秩? 在這個單元中會提出一些觀點回答這個問題, 例如可以研究矩陣行向量生成的子空間維度, 也可以研究列向量生成的子空間維度, 或是將矩陣轉化成一種特殊形式後判讀左上角區塊單位矩陣的大小, 也可以從觀察轉置矩陣的秩得知原矩陣的秩。至於這一個單元還會再回答另外兩個數學問題: 一個是驗證線性變換的秩與其矩陣表現的秩相同, 與有序基底的選取無關; 另一個是給出若矩陣左乘或右乘一個不見得是可逆的矩陣時秩的關係之一般結果。

對於一個方陣是否可逆是一個重要的議題, 一個可逆矩陣要如何找到它的反矩陣在實務上也非常重要, 單元 3.3 將從基本列運算與基本行運算的原理提供一個如何判定矩陣是否可逆以及如何找到反矩陣的方法。日後在其它章節也會提供尋找反矩陣的方法, 屆時可以比較各種方法的優劣性。單元 3.4 則是要給予線性聯立方程式解的存在性與唯一性的數學理論。一方面, 非齊次線性聯立方程式的解集合與齊次線性聯立方程式的解空間有代數結構上的關係; 此外, 不論是線性聯立方程式有解或無解, 有唯一解或是有無限多解, 我們也可以從係數矩陣還有增廣矩陣的秩的關係進行判讀, 所以這將更加突顯秩的重要性。當中也會討論幾個特別的情形, 例如齊次線性聯立方程式的解空間之判讀, 還有方程式個數與未知數個數一樣多的線性聯立方程式有存在唯一解的充分必要條件。

單元 3.5 要介紹高斯消去法, 它是一套處理線性聯立方程式求解的流程, 將線性聯立方程式所對應的增廣矩陣經過有限步的基本列運算後可以轉化為簡約列梯形式, 然後就可以順勢地解讀這個線性聯立方程式是否有解, 如果有解的話, 解的一般式也可以確實表達。這個單元也會討論幾個簡約列梯形式的性質。而這些結果將結合這一章所有的數學理論以及線性聯立方程式求解的實作以得到完整地說明。

3.1 基本矩陣運算與基本矩陣

各位在中學階段一定遇過解三元一次聯立方程式的問題, 這裡舉一個例子重現當時怎麼處理這類問題並得到答案, 考慮以下聯立方程式:

$$\begin{cases} x+y+z=1 \\ 2x+3y-z=-6 \\ x-2y+z=7。\end{cases}$$

步驟一 將第一式乘以 -2 加到第二式, 得到

$$\begin{cases} x+y+z=1 \\ y-3z=-8 \\ x-2y+z=7。\end{cases}$$

步驟二 將第一式乘以 -1 加到第三式, 得到

$$\begin{cases} x+y+z=1 \\ y-3z=-8 \\ -3y=6。\end{cases}$$

步驟三 將第三式乘以 $-\frac{1}{3}$ 後得到

$$\begin{cases} x+y+z=1 \\ y-3z=-8 \\ y=-2。\end{cases}$$

步驟四 將第二式與第三式互換, 得到

$$\begin{cases} x+y+z=1 \\ y=-2 \\ y-3z=-8。\end{cases}$$

步驟五 將第二式乘以 -1 加到第一式, 得到

$$\begin{cases} x+z=3 \\ y=-2 \\ y-3z=-8。\end{cases}$$

步驟六 將第二式乘以 -1 加到第三式, 得到

$$\begin{cases} x+z=3 \\ y=-2 \\ -3z=-6。\end{cases}$$

步驟七 將第三式乘以 $-\frac{1}{3}$ 後得到

$$\begin{cases} x \quad + z = 3 \\ \quad y \quad = -2 \\ \quad\quad z = 2。\end{cases}$$

步驟八 將第三式乘以 -1 加到第一式, 得到

$$\begin{cases} x \quad\quad = 1 \\ \quad y \quad = -2 \\ \quad\quad z = 2。\end{cases}$$

步驟九 透過上述八個步驟, 我們得到聯立方程式的解為 $x = 1, y = -2, z = 2$。

這個例子相對來說算是單純, 因為題目經過設計使得係數都是整數而且就算經過四則運算後仍為整數, 各位應該有處理方程式的係數更醜導致計算變得更複雜的經驗, 然後若沒有解題策略或目標的話常常就會鬼打牆, 始終沒有辦法化簡成功, 也很容易計算錯誤。除此以外, 光是上面的聯立方程式就要用到一頁以上的篇幅處理, 倘偌面對更複雜的問題那計算量不就爆增嗎?

於是我們必須想出一個辦法把上面的解題過程適當地簡化並尋求當中的數學理論。首先回顧解聯立方程式的過程, 我們可以歸結出這些操作可以分成以下三種類型:

(I) 將兩個方程式條列的順序互換, 例如 **步驟四**。

(II) 將某一個方程式乘上非零的倍數, 例如 **步驟三**, **七**。

(III) 將某一個方程式乘上非零的倍數後加到另外一個方程式, 例如 **步驟一**, **二**, **五**, **六**, **八**。

再來我們引進矩陣處理問題, 其實一言以蔽之, 就是考慮分離係數法, 只要把未知數前的每一個係數排排站順序不要弄混亂即可, 例如矩陣的第 $1, 2, 3$ 列分別記錄第 $1, 2, 3$ 個方程式, 而第 $1, 2, 3$ 行分別對應到未知數 x, y, z 前的係數, 第 4 行則是記錄等式右邊不帶未知數的數字。所以下面的矩陣註記方式大大地簡化剛才的解題過程:

$$\left[\begin{array}{ccc|c} 1 & 1 & 1 & 1 \\ 2 & 3 & -1 & -6 \\ 1 & -2 & 1 & 7 \end{array}\right] \xrightarrow{(\mathrm{III})} \left[\begin{array}{ccc|c} 1 & 1 & 1 & 1 \\ 0 & 1 & -3 & -8 \\ 1 & -2 & 1 & 7 \end{array}\right] \xrightarrow{(\mathrm{III})} \left[\begin{array}{ccc|c} 1 & 1 & 1 & 1 \\ 0 & 1 & -3 & -8 \\ 0 & -3 & 0 & 6 \end{array}\right]$$

$$\xrightarrow{(\mathrm{II})} \left[\begin{array}{ccc|c} 1 & 1 & 1 & 1 \\ 0 & 1 & -3 & -8 \\ 0 & 1 & 0 & -2 \end{array}\right] \xrightarrow{(\mathrm{I})} \left[\begin{array}{ccc|c} 1 & 1 & 1 & 1 \\ 0 & 1 & 0 & -2 \\ 0 & 1 & -3 & -8 \end{array}\right] \xrightarrow{(\mathrm{III})} \left[\begin{array}{ccc|c} 1 & 0 & 1 & 3 \\ 0 & 1 & 0 & -2 \\ 0 & 1 & -3 & -8 \end{array}\right]$$

$$\xrightarrow{(\mathrm{III})} \left[\begin{array}{ccc|c} 1 & 0 & 1 & 3 \\ 0 & 1 & 0 & -2 \\ 0 & 0 & -3 & -6 \end{array}\right] \xrightarrow{(\mathrm{II})} \left[\begin{array}{ccc|c} 1 & 0 & 1 & 3 \\ 0 & 1 & 0 & -2 \\ 0 & 0 & 1 & 2 \end{array}\right] \xrightarrow{(\mathrm{III})} \left[\begin{array}{ccc|c} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & -2 \\ 0 & 0 & 1 & 2 \end{array}\right]$$

利用上述寫法解題過程變得較為清爽, 並且可以直接解讀聯立方程式的解為 $x = 1, y = -2, z = 2$。

前一頁的矩陣運算中, 每個矩陣的第三行與第四行之間多了一條直線, 這條直線是有意義的, 在直線的左邊所成的矩陣我們稱為係數矩陣 (coefficient matrix), 也就是聯立方程式各未知數前面的係數, 而將係數矩陣再追加一行 (等式右邊的行向量) 所成的矩陣稱為增廣矩陣 (augmented matrix)。我們會在單元 3.3 與單元 3.4 分別給出增廣矩陣與係數矩陣的定義, 此外還會探討兩個矩陣與聯立方程式是否可解的關係。

若將解聯立的過程轉化為矩陣的註記充其量也只是略去所有 x, y, z 的符號而已, 只要相對位置放得正確就可以還原回帶有 x, y, z 的聯立方程式, 而我們更希望得到的是: 這樣的註記方法能否帶給我們更實值的效益, 特別是能否從中深刻理解背後的數學理論。

於是接下來將思考的問題是: 雖然我們從解題過程中歸結出三種類型的操作, 而這些操作要如何對應到矩陣的乘法? 這裡先把每一種操作給出數學名詞。

定義 **1** (第 148 頁)**.** 給定 $A \in M_{m\times n}(\mathbb{F})$, 以下三種操作的任何一種都稱為 基本列 [行] 運算 (elementary row [column] operation):

(I) 將 A 的某兩列 [行] 交換。

(II) 將 A 的某一列 [行] 乘上一個非零的常數。

(III) 將 A 的某一列 [行] 乘上一個非零的常數後加到另一列 [行]。

這些基本運算分別稱為 第 I 型 (type I)、第 II 型 (type II) 與 第 III 型 (type III)。

接著我們引進基本矩陣的概念。

定義 **2** (第 149 頁)**.** 將一個單位矩陣 I_n 進行一次基本運算後所得的矩陣稱為 基本矩陣 (elementary matrix)。同樣地, 我們也用第 I 型、第 II 型與第 III 型分類。

例 **3.** 以 $n = 3$ 為例, 第 I 型的基本矩陣有以下三種:

$$\begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} \quad \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix},$$

第 II 型的基本矩陣有以下三種:

$$\begin{bmatrix} k & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad \begin{bmatrix} 1 & 0 & 0 \\ 0 & k & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & k \end{bmatrix},$$

其中 $k \in \mathbb{F}, k \neq 0$。至於第 III 型的基本矩陣有以下六種:

$$\begin{bmatrix} 1 & 0 & 0 \\ k & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ k & 0 & 1 \end{bmatrix} \quad \begin{bmatrix} 1 & k & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & k & 1 \end{bmatrix} \quad \begin{bmatrix} 1 & 0 & k \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & k \\ 0 & 0 & 1 \end{bmatrix},$$

其中 $k \in \mathbb{F}, k \neq 0$。

引進基本矩陣之後, 我們就可以順利地將「矩陣進行基本運算」和「矩陣乘上基本矩陣」這兩件事建立完整地對應。白話來說, 以下定理欲說明的是: 將一個矩陣進行一次基本列 [行] 運算等同於將一個矩陣乘上一個基本矩陣。

定理 4 (第 149 頁)**.** 令 $A \in M_{m\times n}(\mathbb{F})$, 假設矩陣 B 是將矩陣 A 進行一次基本列 [行] 運算, 則存在一個 $m \times m$ $[n \times n]$ 基本矩陣 E 使得 $B = EA$ $[B = AE]$。事實上, E 是將 I_m $[I_n]$ 操作和由 A 得 B 相同的基本列 [行] 運算所得之矩陣。反之, 若 E 是基本 $m \times m$ $[n \times n]$ 矩陣, 則 EA $[AE]$ 是將 A 操作由 I_m $[I_n]$ 得 E 相同的基本列 [行] 運算所得之矩陣。

證明: 首先討論第 I 型基本列運算與第 I 型基本矩陣的關係, 將 E 用列向量的方式表達, 而 A 用行向量的方式表達:

$$E = \left[\begin{array}{c} \mathbf{e}_1 \\ \hline \vdots \\ \hline \mathbf{e}_q \\ \hline \vdots \\ \hline \mathbf{e}_p \\ \hline \vdots \\ \hline \mathbf{e}_m \end{array}\right] \qquad A = \left[\begin{array}{c|c|c|c|c|c|c} \tilde{\mathbf{a}}_1 & \cdots & \tilde{\mathbf{a}}_p & \cdots & \tilde{\mathbf{a}}_q & \cdots & \tilde{\mathbf{a}}_n \end{array}\right],$$

(這裡將行向量加上波浪號主要是與下面的列向量做區別) 得到

$$EA = \begin{bmatrix} \mathbf{e}_1\cdot\tilde{\mathbf{a}}_1 & \mathbf{e}_1\cdot\tilde{\mathbf{a}}_2 & \cdots & \mathbf{e}_1\cdot\tilde{\mathbf{a}}_n \\ \cdots & \cdots & \cdots & \cdots \\ \mathbf{e}_q\cdot\tilde{\mathbf{a}}_1 & \mathbf{e}_q\cdot\tilde{\mathbf{a}}_2 & \cdots & \mathbf{e}_q\cdot\tilde{\mathbf{a}}_n \\ \cdots & \cdots & \cdots & \cdots \\ \mathbf{e}_p\cdot\tilde{\mathbf{a}}_1 & \mathbf{e}_p\cdot\tilde{\mathbf{a}}_2 & \cdots & \mathbf{e}_p\cdot\tilde{\mathbf{a}}_n \\ \cdots & \cdots & \cdots & \cdots \\ \mathbf{e}_m\cdot\tilde{\mathbf{a}}_1 & \mathbf{e}_m\cdot\tilde{\mathbf{a}}_2 & \cdots & \mathbf{e}_m\cdot\tilde{\mathbf{a}}_n \end{bmatrix} = \begin{bmatrix} A_{11} & A_{12} & \cdots & A_{1n} \\ \cdots & \cdots & \cdots & \cdots \\ A_{q1} & A_{q2} & \cdots & A_{qn} \\ \cdots & \cdots & \cdots & \cdots \\ A_{p1} & A_{p2} & \cdots & A_{pn} \\ \cdots & \cdots & \cdots & \cdots \\ A_{m1} & A_{m2} & \cdots & A_{mn} \end{bmatrix},$$

上述使用了兩向量內積 (inner product) 的記號, 定義為: 給定 $\mathbf{u} = (u_1, u_2, \ldots, u_n), \mathbf{v} = (v_1, v_2, \ldots, v_n)$, 則 $\mathbf{u}\cdot\mathbf{v} = u_1v_1 + u_2v_2 + \cdots + u_nv_n$。於是對 EA 而言, 從列的觀點來看, 則為

$$EA = \left[\begin{array}{c} \mathbf{a}_1 \\ \hline \vdots \\ \hline \mathbf{a}_q \\ \hline \vdots \\ \hline \mathbf{a}_p \\ \hline \vdots \\ \hline \mathbf{a}_m \end{array}\right],$$

也就是呈現 A 的第 p 列與第 q 列互換的結果。

再看第 II 型基本列運算與第 II 型基本矩陣的關係, 將 E 用列的方式表達, 而 A 用行的方式表示:

$$E = \left[\begin{array}{c} \mathbf{e}_1 \\ \hline \vdots \\ \hline k\cdot\mathbf{e}_p \\ \hline \vdots \\ \hline \mathbf{e}_m \\ \hline \end{array}\right] \qquad A = \left[\begin{array}{c|c|c|c|c} \tilde{\mathbf{a}}_1 & \cdots & \tilde{\mathbf{a}}_p & \cdots & \tilde{\mathbf{a}}_n \end{array}\right],$$

其中 $k \in \mathbb{F}, k \neq 0$, 得到

$$EA = \begin{bmatrix} \mathbf{e}_1\cdot\tilde{\mathbf{a}}_1 & \mathbf{e}_1\cdot\tilde{\mathbf{a}}_2 & \cdots & \mathbf{e}_1\cdot\tilde{\mathbf{a}}_n \\ \cdots & \cdots & \cdots & \cdots \\ k\cdot\mathbf{e}_p\cdot\tilde{\mathbf{a}}_1 & k\cdot\mathbf{e}_p\cdot\tilde{\mathbf{a}}_2 & \cdots & k\cdot\mathbf{e}_p\cdot\tilde{\mathbf{a}}_n \\ \cdots & \cdots & \cdots & \cdots \\ \mathbf{e}_m\cdot\tilde{\mathbf{a}}_1 & \mathbf{e}_m\cdot\tilde{\mathbf{a}}_2 & \cdots & \mathbf{e}_m\cdot\tilde{\mathbf{a}}_n \end{bmatrix} = \begin{bmatrix} A_{11} & A_{12} & \cdots & A_{1n} \\ \cdots & \cdots & \cdots & \cdots \\ k\cdot A_{p1} & k\cdot A_{p2} & \cdots & k\cdot A_{pn} \\ \cdots & \cdots & \cdots & \cdots \\ A_{m1} & A_{m2} & \cdots & A_{mn} \end{bmatrix},$$

於是對 EA 而言, 從列的觀點來看, 則是將矩陣 A 的第 p 列乘上 k 之後的結果。

再看第 III 型基本列運算與第 III 型基本矩陣的關係, 將 E 用列的方式表達, 而 A 用行的方式表示:

$$E = \left[\begin{array}{c} \mathbf{e}_1 \\ \hline \vdots \\ \hline \mathbf{e}_p + k\cdot\mathbf{e}_q \\ \hline \vdots \\ \hline \mathbf{e}_m \\ \hline \end{array}\right] \qquad A = \left[\begin{array}{c|c|c|c|c} \tilde{\mathbf{a}}_1 & \cdots & \tilde{\mathbf{a}}_p & \cdots & \tilde{\mathbf{a}}_n \end{array}\right],$$

其中 $k \in \mathbb{F}, k \neq 0$, 得到

$$\begin{aligned} EA &= \begin{bmatrix} \mathbf{e}_1\cdot\tilde{\mathbf{a}}_1 & \mathbf{e}_1\cdot\tilde{\mathbf{a}}_2 & \cdots & \mathbf{e}_1\cdot\tilde{\mathbf{a}}_n \\ \cdots & \cdots & \cdots & \cdots \\ (\mathbf{e}_p + k\cdot\mathbf{e}_q)\cdot\tilde{\mathbf{a}}_1 & (\mathbf{e}_p + k\cdot\mathbf{e}_q)\cdot\tilde{\mathbf{a}}_2 & \cdots & (\mathbf{e}_p + k\cdot\mathbf{e}_q)\cdot\tilde{\mathbf{a}}_n \\ \cdots & \cdots & \cdots & \cdots \\ \mathbf{e}_m\cdot\tilde{\mathbf{a}}_1 & \mathbf{e}_m\cdot\tilde{\mathbf{a}}_2 & \cdots & \mathbf{e}_m\cdot\tilde{\mathbf{a}}_n \end{bmatrix} \\ &= \begin{bmatrix} A_{11} & A_{12} & \cdots & A_{1n} \\ \cdots & \cdots & \cdots & \cdots \\ A_{p1} + k\cdot A_{q1} & A_{p2} + k\cdot A_{q2} & \cdots & A_{pn} + k\cdot A_{qn} \\ \cdots & \cdots & \cdots & \cdots \\ A_{m1} & A_{m2} & \cdots & A_{mn} \end{bmatrix}, \end{aligned}$$

於是對 EA 而言, 從列的觀點來看, 則是將 A 的第 q 列乘以 k 後加到第 p 列之結果。

再來我們要觀察的是基本行運算的情況。同樣地, 先探討第 I 型基本行運算與第 I 型基本矩陣的關係。將 A 用列向量的方式表達, 而 E 用行向量的方式表達:

$$A = \left[\begin{array}{c} \mathbf{a}_1 \\ \hline \vdots \\ \hline \mathbf{a}_p \\ \hline \vdots \\ \hline \mathbf{a}_q \\ \hline \vdots \\ \hline \mathbf{a}_m \end{array}\right] \qquad E = \left[\begin{array}{c|c|c|c|c|c|c} \tilde{\mathbf{e}}_1 & \cdots & \tilde{\mathbf{e}}_q & \cdots & \tilde{\mathbf{e}}_p & \cdots & \tilde{\mathbf{e}}_n \end{array}\right],$$

得到

$$\begin{aligned} AE &= \begin{bmatrix} \mathbf{a}_1\cdot\tilde{\mathbf{e}}_1 & \cdots & \mathbf{a}_1\cdot\tilde{\mathbf{e}}_q & \cdots & \mathbf{a}_1\cdot\tilde{\mathbf{e}}_p & \cdots & \mathbf{a}_1\cdot\tilde{\mathbf{e}}_n \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ \mathbf{a}_p\cdot\tilde{\mathbf{e}}_1 & \cdots & \mathbf{a}_p\cdot\tilde{\mathbf{e}}_q & \cdots & \mathbf{a}_p\cdot\tilde{\mathbf{e}}_p & \cdots & \mathbf{a}_p\cdot\tilde{\mathbf{e}}_n \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ \mathbf{a}_q\cdot\tilde{\mathbf{e}}_1 & \cdots & \mathbf{a}_q\cdot\tilde{\mathbf{e}}_q & \cdots & \mathbf{a}_q\cdot\tilde{\mathbf{e}}_p & \cdots & \mathbf{a}_q\cdot\tilde{\mathbf{e}}_n \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ \mathbf{a}_m\cdot\tilde{\mathbf{e}}_1 & \cdots & \mathbf{a}_m\cdot\tilde{\mathbf{e}}_q & \cdots & \mathbf{a}_m\cdot\tilde{\mathbf{e}}_p & \cdots & \mathbf{a}_m\cdot\tilde{\mathbf{e}}_n \end{bmatrix} \\ &= \begin{bmatrix} A_{11} & \cdots & A_{1q} & \cdots & A_{1p} & \cdots & A_{1n} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ A_{p1} & \cdots & A_{pq} & \cdots & A_{pp} & \cdots & A_{pn} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ A_{q1} & \cdots & A_{qq} & \cdots & A_{qp} & \cdots & A_{qn} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ A_{m1} & \cdots & A_{mq} & \cdots & A_{mp} & \cdots & A_{mn} \end{bmatrix}, \end{aligned}$$

於是對 AE 而言, 從行的觀點來看, 則為

$$AE = \left[\begin{array}{c|c|c|c|c|c|c} \tilde{\mathbf{a}}_1 & \cdots & \tilde{\mathbf{a}}_q & \cdots & \tilde{\mathbf{a}}_p & \cdots & \tilde{\mathbf{a}}_n \end{array}\right],$$

也就是呈現 A 的第 p 行與第 q 行互換的結果。

再看第 II 型基本行運算與第 II 型基本矩陣的關係, 將 A 用列的方式表達, 而 E 用行的方式表示:

$$A = \left[\begin{array}{c} \mathbf{a}_1 \\ \hline \vdots \\ \hline \mathbf{a}_p \\ \hline \vdots \\ \hline \mathbf{a}_m \end{array}\right] \qquad E = \left[\begin{array}{c|c|c|c|c} \tilde{\mathbf{e}}_1 & \cdots & k\cdot\tilde{\mathbf{e}}_p & \cdots & \tilde{\mathbf{e}}_n \end{array}\right],$$

其中 $k \in \mathbb{F}, k \neq 0$, 得到

$$AE = \begin{bmatrix} \mathbf{a}_1 \cdot \tilde{\mathbf{e}}_1 & \cdots & \mathbf{a}_1 \cdot k \cdot \tilde{\mathbf{e}}_p & \cdots & \mathbf{a}_1 \cdot \tilde{\mathbf{e}}_n \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ \mathbf{a}_p \cdot \tilde{\mathbf{e}}_1 & \cdots & \mathbf{a}_p \cdot k \cdot \tilde{\mathbf{e}}_p & \cdots & \mathbf{a}_p \cdot \tilde{\mathbf{e}}_n \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ \mathbf{a}_m \cdot \tilde{\mathbf{e}}_1 & \cdots & \mathbf{a}_m \cdot k \cdot \tilde{\mathbf{e}}_p & \cdots & \mathbf{a}_m \cdot \tilde{\mathbf{e}}_n \end{bmatrix} = \begin{bmatrix} A_{11} & \cdots & k \cdot A_{1p} & \cdots & A_{1n} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ A_{p1} & \cdots & k \cdot A_{pp} & \cdots & A_{pn} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ A_{m1} & \cdots & k \cdot A_{mp} & \cdots & A_{mn} \end{bmatrix},$$

於是對 AE 而言, 從行的觀點來看, 則為

$$AE = \left[\begin{array}{c|c|c|c|c} \tilde{\mathbf{a}}_1 & \cdots & k \cdot \tilde{\mathbf{a}}_p & \cdots & \tilde{\mathbf{a}}_n \end{array} \right],$$

也就是是將矩陣 A 的第 p 行乘上 k 之後的結果。

最後要討論的是第 III 型基本行運算與第 III 型基本矩陣的關係, 將 A 用列的方式表達, 而 E 用行的方式表示:

$$A = \left[\begin{array}{c} \mathbf{a}_1 \\ \hline \vdots \\ \hline \mathbf{a}_p \\ \hline \vdots \\ \hline \mathbf{a}_m \end{array} \right] \qquad E = \left[\begin{array}{c|c|c|c|c} \tilde{\mathbf{e}}_1 & \cdots & \tilde{\mathbf{e}}_p + k \cdot \tilde{\mathbf{e}}_q & \cdots & \tilde{\mathbf{e}}_n \end{array} \right],$$

其中 $k \in \mathbb{F}, k \neq 0$, 得到

$$\begin{aligned} AE &= \begin{bmatrix} \mathbf{a}_1 \cdot \tilde{\mathbf{e}}_1 & \cdots & \mathbf{a}_1 \cdot (\tilde{\mathbf{e}}_p + k \cdot \tilde{\mathbf{e}}_q) & \cdots & \mathbf{a}_1 \cdot \tilde{\mathbf{e}}_n \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ \mathbf{a}_p \cdot \tilde{\mathbf{e}}_1 & \cdots & \mathbf{a}_p \cdot (\tilde{\mathbf{e}}_p + k \cdot \tilde{\mathbf{e}}_q) & \cdots & \mathbf{a}_p \cdot \tilde{\mathbf{e}}_n \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ \mathbf{a}_m \cdot \tilde{\mathbf{e}}_1 & \cdots & \mathbf{a}_m \cdot (\tilde{\mathbf{e}}_p + k \cdot \tilde{\mathbf{e}}_q) & \cdots & \mathbf{a}_m \cdot \tilde{\mathbf{e}}_n \end{bmatrix} \\ &= \begin{bmatrix} A_{11} & \cdots & A_{1p} + k \cdot A_{1q} & \cdots & A_{1n} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ A_{p1} & \cdots & A_{pp} + k \cdot A_{pq} & \cdots & A_{pn} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ A_{m1} & \cdots & A_{mp} + k \cdot A_{mq} & \cdots & A_{mn} \end{bmatrix}, \end{aligned}$$

於是對 AE 而言, 從行的觀點來看, 則為

$$AE = \left[\begin{array}{c|c|c|c|c} \tilde{\mathbf{a}}_1 & \cdots & \tilde{\mathbf{a}}_p + k \cdot \tilde{\mathbf{a}}_q & \cdots & \tilde{\mathbf{a}}_n \end{array} \right],$$

也就是是將矩陣 A 的第 q 行乘上 k 後加到第 p 行的結果。 $\square$

介紹完基本矩陣與矩陣操作的關係後, 以下定理要給出基本矩陣的性質。

定理 5 (第 150 頁)**.** 所有的基本矩陣皆為可逆矩陣, 其反矩陣亦為基本矩陣, 且具有相同的類型。

證明:

- 假設 E 是由單位矩陣 I_n 的第 p 列與第 q 列 $(p \neq q)$ 互換而得, 則 $E^2 = EE = I_n$, 所以 E 是可逆矩陣, 並且 $E^{-1} = E$。此時, E 和 E^{-1} 都是第 I 型基本矩陣。

- 假設 E 是由單位矩陣 I_n 的第 p 列乘以一個非零常數 k 而得的矩陣, 因為 $k \neq 0$, 故 k 的乘法反元素 k^{-1} 存在唯一, 考慮 $\bar{E}$ 是由單位矩陣 I_n 的第 p 列乘以 k^{-1} 而得的矩陣, 則 $E\bar{E} = \bar{E}E = I_n$, 所以 E 是可逆矩陣, 並且 $E^{-1} = \bar{E}$。此時, E 和 E^{-1} 都是第 II 型基本矩陣。

- 假設 E 是由單位矩陣 I_n 的第 q 列乘以 k 倍加至第 p 列而得的矩陣, 其中 $p \neq q$ 且 $k \in \mathbb{F}, k \neq 0$, 考慮 $\bar{E}$ 是由 I_n 的第 q 列乘以 $-k$ 倍加至第 p 列而得的矩陣, 則 $E\bar{E} = \bar{E}E = I_n$, 所以 E 是可逆矩陣, 並且 $E^{-1} = \bar{E}$。此時, E 和 E^{-1} 都是第 III 型基本矩陣。

$\square$

各位是否對於 定義 **1** 與 定理 **4** 產生一個疑問: 定義與定理中提及基本的行運算還有相關結果, 而我們在處理解聯立方程式的時候, 為了要把未知數消除, 所以三種類型的列運算都有實質的對應, 那麼行運算又代表什麼意義呢?

若要將行運算的操作和解聯立方程式這件事產生聯繫, 第 I 型行運算是很好解釋的, 其實就只是把兩類型的未知數的擺放順序對調而已, 當然, 在最後要呼應聯立方程式的解的話, 就要注意你曾經把未知數換了位置。然而, 第 II 型與第 III 型的行運算你可能就會有一種百思不得其解的感覺, 甚至會覺得好像沒有什麼道理可言。實際上, 第 II 型與第 III 型的行運算的確不容易解釋, 這兩類的行運算與解聯立方程式的關係, 可以想成是一種變數變換 (change of variables)。這裡舉一個例子說明這個變數變換的關係:

$$\left\{\begin{array}{l} x+y=1 \\ x-y=3 \end{array}\right. \Leftrightarrow \left[\begin{array}{cc|c} 1 & 1 & 1 \\ 1 & -1 & 3 \end{array}\right] \xrightarrow{(\mathrm{III})} \left[\begin{array}{cc|c} 2 & 1 & 1 \\ 0 & -1 & 3 \end{array}\right] \Leftrightarrow \left\{\begin{array}{r} 2\tilde{x}+\tilde{y}=1 \\ -\tilde{y}=3, \end{array}\right.$$

其中 $\tilde{x} = x, \tilde{y} = y - x$。接著,

$$\left\{\begin{array}{r} 2\tilde{x}+\tilde{y}=1 \\ -\tilde{y}=3 \end{array}\right. \Leftrightarrow \left[\begin{array}{cc|c} 2 & 1 & 1 \\ 0 & -1 & 3 \end{array}\right] \xrightarrow{(\mathrm{II})} \left[\begin{array}{cc|c} 1 & 1 & 1 \\ 0 & -1 & 3 \end{array}\right] \Leftrightarrow \left\{\begin{array}{r} u+v=1 \\ -v=3, \end{array}\right.$$

其中 $u = 2\tilde{x} = 2x, v = \tilde{y} = y - x$。然後,

$$\left\{\begin{array}{r} u+v=1 \\ -v=3, \end{array}\right. \Leftrightarrow \left[\begin{array}{cc|c} 1 & 1 & 1 \\ 0 & -1 & 3 \end{array}\right] \xrightarrow{(\mathrm{III})} \left[\begin{array}{cc|c} 1 & 0 & 1 \\ 0 & -1 & 3 \end{array}\right] \Leftrightarrow \left\{\begin{array}{l} \tilde{u} \qquad =1 \\ \quad -\tilde{v}=3, \end{array}\right.$$

其中 $\tilde{u} = u + v = 2\tilde{x} + \tilde{y} = 2x + y - x = x + y, \tilde{v} = v = \tilde{y} = y - x$。最後,

$$\left\{\begin{array}{l} \tilde{u} \qquad =1 \\ \quad -\tilde{v}=3 \end{array}\right. \Leftrightarrow \left[\begin{array}{cc|c} 1 & 0 & 1 \\ 0 & -1 & 3 \end{array}\right] \xrightarrow{(\mathrm{III})} \left[\begin{array}{cc|c} 1 & 0 & 1 \\ 0 & 1 & 3 \end{array}\right] \Leftrightarrow \left\{\begin{array}{l} z \qquad =1 \\ \qquad w=3, \end{array}\right.$$

其中 $z = \tilde{u} = u + v = 2\tilde{x} + \tilde{y} = 2x + y - x = x + y, w = -\tilde{v} = -v = -\tilde{y} = x - y$。

當你看完上述的行運算與變數變換的所有過程後, 或許會感到這件事顯得相當荒謬, 因為做了各式各樣的行運算, 然後把每一個行運算所對應到的變數變換確實寫下, 結果最後的 $z = x + y$ 以及 $w = x - y$ 只不過是把聯立方程式的左式逐一用新的符號包裝而已, 對於想要真正找到聯立方程式的解 (也就是到底哪個 x 與哪個 y 滿足聯立方程式) 一點幫助也沒有。所以你可能會對於將矩陣進行行運算這件事嗤之以鼻。

若你有上述的想法, 又或者曾經心中浮現出行運算的操作感到些許奇怪的想法, 應該給予正向的鼓勵, 因為我覺得到了此時此刻各位對於數學的學習不應該只是照單全收, 在講義或是課堂上沒有或是還沒有提及的一些內容或問題若能自發性提問並找尋答案, 然後對學習數學的情緒再強烈一些, 感受到數學的厲害又或者是當你覺得這件事並非那麼理想時, 又要怎麼改進或是換個思維去看待它。這是需要一些時間培養並嘗試這麼做, 日後就會對數學有不同的體會。

現在回到原問題, 難到行運算就這麼沒用嗎? 如果只是像上面所述將符號替換這層意義的話, 那為什麼還要介紹行運算呢? 針對這個問題, 我想要下幾個註解。首先, 數學本來就只會在某些情況下與實際的問題有完全地對應。比方說我舉各位一定遇過的雞兎同籠的問題為例 (雖然我對這個應用問題也很有意見, 但這裡暫且用這個例子), 雞兎同籠, 共有 10 隻動物, 而腳的總數是 24, 請問雞、兎各幾隻? 而各位會很下意識地就假設雞有 x 隻, 兎有 y 隻, 所以根據線索, 得到 $x + y = 10$ 以及 $2x + 4y = 24$, 到目前為止你都可以把符號與題意的關係完全地對應, 但是當你把第一式乘以 2 然後與第二式相減得到 $-2y = -4$, 這時你就無法 (或不需要) 解釋它在原應用問題的確切意義了, 我們是藉助著數學運算系統得到 $y = 2$ 與 $x = 8$, 而最後的結果才會再呼應原問題。

這裡再舉一個例子, 各位在微積分課程中一定學過羅必達法則 (L' Hôpital's Rule) 然後對它愛不釋手 (同樣地我對羅必達法則也很有意見, 但這裡還是暫且用了這個例子), 而羅必達法則的確切意義是: 對於一個極限問題, 在使用羅必達法則之前的極限和將函數的分子與分母各自求導後再去問極限, 這是兩個完全不同的極限問題, 只是羅必達法則是在說: 在某些條件下, 這兩個極限值會相同。所以羅必達法則是將兩個表象完全不同的極限問題只是在極限值相同的意義下用等號串起來, 等號前的極限與等號後面的極限會有各別的詮釋。

於是, 若你對前面幾段文字去解釋行向量與聯立方程式的關係感到很不滿意的話也不打緊, 或許可以不必過於堅持一定要講出個什麼道理, 而是從數學的抽象思維去看待它, 比方說, 給定一個矩陣, 我們可以對於一個矩陣進行列的運算, 自然在操作面上我們也可以考慮行的運算, 這樣的想法就比起硬是要和聯立方程式有對應的想法合理許多。在這個概念下, 我們應該要把焦點放在另一個問題: 當我們做了一些行運算之後, 這麼做能帶給我們什麼新的啟發? 各位可以在後面的單元及章節就會發現到, 若我們允許將矩陣進行行向量的操作, 讓數學運算系統變得更靈活而豐富 (不侷限於只能列的層面考量事情), 會幫助我們更快得到某些核心的結果 (例如行運算的保秩, 或是矩陣行運算的化簡與行列式的關係)。所以我們應專注的點是: 如果我們額外做了一些事情, 那麼哪些東西仍然保持不變, 或者是哪些東西做了哪些改變。

最後我想補充的是: 關於行運算的確切意涵, 如果你對於第二章曾經提及的對偶空間 (dual space) 熟悉的話, 那裡從中得到一些體悟。列運算與行運算會和向量空間與其對偶空間是有關係的這件事也很合理, 因為我們曾經證明過一個矩陣的轉置矩陣是要從對偶空間的觀點得到對偶基底的矩陣表現, 所以矩陣的行運算可以想成是它的轉置矩陣的列運算, 所以在對偶空間看事情才會對行運算有更為貼切的解釋。

3.2 矩陣的秩與保秩變換

這一個單元我們將從線性變換的秩還有矩陣的秩證明矩陣在基本運算下具有保秩的結構。而保秩的結構會對於解聯立方程式的理論有一些幫助, 這會在下一個單元繼續介紹。首先定義矩陣的秩。

定義 **1** (第 152 頁)**.** 給定 $A \in M_{m\times n}(\mathbb{F})$, 定義矩陣 A 的 秩 (rank) 為左乘變換 $L_A : \mathbb{F}^n \to \mathbb{F}^m$ 的秩, 記為 $\mathrm{rank}(A)$。

以下直接給出關於矩陣的秩的基本結果。

定理 **2** (第 153 頁)**.** 假設 $A \in M_{m\times n}(\mathbb{F})$, 而 $P \in M_{m\times m}(\mathbb{F}), Q \in M_{n\times n}(\mathbb{F})$ 是可逆矩陣, 則

(A) $\mathrm{rank}(AQ) = \mathrm{rank}(A)$。

(B) $\mathrm{rank}(PA) = \mathrm{rank}(A)$。

(C) $\mathrm{rank}(PAQ) = \mathrm{rank}(A)$。

證明:

(A) 因為 Q 是可逆矩陣, 所以 $L_Q : \mathbb{F}^n \to \mathbb{F}^n$ 映成, 於是

$$R(L_{AQ}) = R(L_A L_Q) = L_A L_Q(\mathbb{F}^n) = L_A(L_Q(\mathbb{F}^n)) = L_A(\mathbb{F}^n) = R(L_A),$$

所以 $\mathrm{rank}(AQ) = \dim(R(L_{AQ})) = \dim(R(L_A)) = \mathrm{rank}(A)$。

(B) 因為一個可逆線性變換 $T : V \to W$ 的特性是: 若 U 是 V 中的 k 維子空間, 則 $T(U)$ 是 W 中的 k 維子空間。由此, 考慮 $L_P : \mathbb{F}^m \to \mathbb{F}^m$, 則 L_P 是可逆的線性變換, 因為 $R(L_{PA}) = R(L_P L_A) = L_P(L_A(\mathbb{F}^n))$, 而 $\mathrm{rank}(A)$ 是子空間 $L_A(\mathbb{F}^n)$ 的維度, 所以子空間 $L_P(L_A(\mathbb{F}^n))$ 的維度與 $L_A(\mathbb{F}^n)$ 的維度相同, 因此 $\mathrm{rank}(PA) = \mathrm{rank}(A)$。

以下補充證明可逆線性變換 $T : V \to W$ 的特性: 若 U 是 V 中的 k 維子空間, 則 $T(U)$ 是 W 中的 k 維子空間。給定 $\beta = \{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k\}$ 是 V 中的子空間 U 的基底, 考慮方程式 $c_1 \cdot T(\mathbf{v}_1) + c_2 \cdot T(\mathbf{v}_2) + \cdots + c_k \cdot T(\mathbf{v}_k) = \mathbf{0}$, 其中 $c_1, c_2, \dots, c_k \in \mathbb{F}$ 待定。因為 T 是線性, 所以 $T(c_1 \cdot \mathbf{v}_1 + c_2 \cdot \mathbf{v}_2 + \cdots + c_k \cdot \mathbf{v}_k) = \mathbf{0}$, 而 T 可逆, 所以 T 是一對一, 於是 $c_1 \cdot \mathbf{v}_1 + c_2 \cdot \mathbf{v}_2 + \cdots + c_k \cdot \mathbf{v}_k = \mathbf{0}$, 因為 β 是 U 的基底, 所以 $c_1 = 0, c_2 = 0, \dots, c_n = 0$, 故 $T(\beta) \overset{\text{記}}{=} \{T(\mathbf{v}_1), T(\mathbf{v}_2), \dots, T(\mathbf{v}_k)\}$ 是線性獨立。

任給 $\mathbf{w} \in T(U)$, 則存在 $\mathbf{u}$ 使得 $\mathbf{w} = T(\mathbf{u})$, 因為 β 是 U 的基底, 所以存在 $c_1, c_2, \dots, c_k \in \mathbb{F}$ 使得 $\mathbf{u} = c_1 \cdot \mathbf{v}_1 + c_2 \cdot \mathbf{v}_2 + \cdots + c_k \cdot \mathbf{v}_k$, 於是

$$T(\mathbf{u}) = T(c_1 \cdot \mathbf{v}_1 + c_2 \cdot \mathbf{v}_2 + \cdots + c_k \cdot \mathbf{v}_k) = c_1 \cdot T(\mathbf{v}_1) + c_2 \cdot T(\mathbf{v}_2) + \cdots + c_k \cdot T(\mathbf{v}_k),$$

得到 $\mathbf{w} = T(\mathbf{u}) \in \mathrm{span}(T(\beta))$, 於是 $T(U) \subset \mathrm{span}(T(\beta))$, 又 $T(U)$ 是子空間, 且 $T(\beta) \subset T(U)$, 所以 $\mathrm{span}(T(\beta)) \subset T(U)$, 因此 $T(U) = \mathrm{span}(T(\beta))$。

由上討論可得 $T(U)$ 的維度是 k。

(C) 由 (A) 和 (B) 得知: $\operatorname{rank}(PAQ) = \operatorname{rank}((PA)Q) = \operatorname{rank}(PA) = \operatorname{rank}(A)$。

□

關於上面的定理, 直觀上也能接受不論是將矩陣左乘或是右乘可逆矩陣的話都會保秩, 這是因為可逆的矩陣對應於一個不退化的線性變換, 也就是定義域與值域大小 (維度) 一致的線性變換, 而左乘矩陣對應到的是在後面進行合成, 右乘矩陣則是在源頭先進行合成, 所以合成之後的結果將保秩。

由上定理就可以直接得到基本列運算以及基本行運算皆保秩。

推論 **3** (第 153 頁)**.** 對矩陣作基本列運算或是基本行運算皆為保秩的運算。

證明: 假設矩陣 B 是將矩陣 A 經過某一種基本列 [行] 運算而得的矩陣, 由單元 3.1 的 定理 **4**, 存在一個基本矩陣使得 $B = EA$ $[B = AE]$, 因為單元 3.1 的 定理 **5** 知 E 是一個可逆矩陣, 由 定理 **2** 得到

$$\operatorname{rank}(B) = \operatorname{rank}(EA) = \operatorname{rank}(A) \quad [\operatorname{rank}(B) = \operatorname{rank}(AE) = \operatorname{rank}(A)]。$$

□

目前為止, 我們已經知道一大類型的保秩運算, 接下來我們將關心的問題是: 任何一個矩陣的秩要如何判讀? 以下要介紹的幾個定理最終將提供如何解讀一個矩陣的秩。

定理 **4** (第 153 頁)**.** *任何一個矩陣的秩是由其行向量生成之子空間的維度; 換言之, 任何一個矩陣的秩等於其最大線性獨立行向量的數目。*

證明: 給定 $A \in M_{m\times n}(\mathbb{F})$, 首先將 A 用行向量的方式表達, 即

$$A = \left[\begin{array}{c|c|c|c} \mathbf{a}_1 & \mathbf{a}_2 & \cdots & \mathbf{a}_n \end{array}\right],$$

而 $\operatorname{rank}(A) = \operatorname{rank}(L_A) = \dim(R(L_A))$。令 $\beta = \{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n\}$ 是 $\mathbb{F}^n$ 的標準有序基底, 則

$$R(L_A) = \operatorname{span}(\{L_A(\mathbf{e}_1), L_A(\mathbf{e}_2), \ldots, L_A(\mathbf{e}_n)\}),$$

對於 $j = 1, 2, \ldots, n$, 因為 $L_A(\mathbf{e}_j)$ 是 A 的第 j 行, 也就是 $\mathbf{a}_j$, 所以

$$R(L_A) = \operatorname{span}(\{\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n\}),$$

因此

$$\operatorname{rank}(A) = \dim(R(L_A)) = \dim(\operatorname{span}\{\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n\})。$$

□

這個定理基本上只是將矩陣的秩之定義進一步地詮釋, 而下面要給出的結果就可以顯示出如何判讀矩陣的秩。

定理 5 (第 155 頁)**.** 若 $A \in M_{m\times n}(\mathbb{F})$, $\operatorname{rank}(A) = r$, 則 $r \le \min(m,n)$, 而且 A 經過有限次基本列運算與基本行運算, 可以使 A 轉化成矩陣

$$D = \begin{bmatrix} I_r & O_1 \\ O_2 & O_3 \end{bmatrix},$$

其中 O_1, O_2 及 O_3 皆為零矩陣; 也就是說, 對所有 $i = 1,2,\ldots,r$, $D_{ii} = 1$ 而其它項 $D_{ij} = 0$。

證明: 若 A 是零矩陣, 則這個矩陣的每一個行向量都是 $\mathbf{0}$, 得到 $R(L_A) = \operatorname{span}(\{\mathbf{0}\}) = \mathbf{0}$, 所以 $r = \operatorname{rank}(A) = \dim(R(L_A)) = 0$。這個情況 $D = A$ 即為所求。

現假設 $A \ne O$, 則 A 至少包含一個線性獨立的行向量, 得到 $r = \operatorname{rank}(A) = \dim(R(L_A)) > 0$, 以下對於 A 的列的個數用數學歸納法證明定理結果:

- 當 $m = 1$, 因為 $r > 0$, 在 $1,2,\ldots,n$ 中存在 j 使得 $A_{1j} \ne 0$, 利用至多一次第 I 型基本行運算及至多一次第 II 型基本行運算將這個非零的項互換並單位化, 得到 A 可轉化成矩陣 $B \in M_{1\times n}(\mathbb{F})$, 其中 $B_{11} = 1$, 再利用至多 $n-1$ 次第 III 型行運算, 得到 B 可以轉化成矩陣

  $$D = \begin{bmatrix} 1 & 0 & \cdots & 0 \end{bmatrix}。$$

  因為 D 中只有一個線性獨立的行向量, 由 定理 **4** 知 $\operatorname{rank}(D) = 1$, 又由 推論 **3** 知 $\operatorname{rank}(A) = \operatorname{rank}(D) = 1 \le \min(m,n)$。於是這個定理在 $m = 1$ 的情況成立。

- 假設定理的結果在 $m = k, k \in \mathbb{N}$ 的情況下成立。

  當 $m = k+1$, 給定 $A \in M_{(k+1)\times n}(\mathbb{F})$, 首先證明當 $n = 1$ 的時候定理成立: 因為 $r > 0$, 在 $1,2,\ldots,k+1$ 中存在 i 使得 $A_{i1} \ne 0$, 利用至多一次第 I 型基本列運算及至多一次第 II 型基本列運算將這個非零的項互換並單位化, 得到 A 可轉化成矩陣 $B \in M_{(k+1)\times 1}(\mathbb{F})$, 其中 $B_{11} = 1$, 再利用至多 k 次第 III 型基本列運算, 得到 B 可以轉化成矩陣

  $$D = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}。$$

  因為 D 中只有一個線性獨立的行向量, 由 定理 **4** 知 $\operatorname{rank}(D) = 1$, 又由 推論 **3** 知 $\operatorname{rank}(A) = \operatorname{rank}(D) = 1 \le \min(m,n)$。於是這個定理在 $m = k+1$, $n = 1$ 的情況成立。

  以下將證明當 $n > 1$ 的結果。由於 $A \ne O$, 故存在 i 及 j 使得 $A_{ij} \ne 0$, 其中 i 是 $1,2,\ldots,k+1$ 中之一數, 而 j 是 $1,2,\ldots,n$ 中之一數。利用至多一次基本列運算及至多一次基本行運算 (皆為第 I 型), A 可以轉化成 $B' \in M_{(k+1)\times n}(\mathbb{F})$ 使得 $B'_{11} \ne 0$; 利用至多一次第 II 型基本行運算, 可將 B' 轉化為 B'' 使得 $B''_{11} = 1$。再利用至多 k 次第 III 型基本列運算及 $n-1$ 次第

III 型基本行運算, 我們可消去第一列以及第一行中其它非零項。總結來說, 經過有限次的基本列運算與基本行運算, A 可轉化成矩陣

$$B''' = \left[\begin{array}{c|ccc} 1 & 0 & \cdots & 0 \\ \hline 0 & & & \\ \vdots & & \tilde{B} & \\ 0 & & & \end{array}\right],$$

其中 $\tilde{B} \in M_{k\times(n-1)}(\mathbb{F})$。因為 $\tilde{B}$ 的秩比 B''' 的秩小 1, 而且 $\mathrm{rank}(A) = \mathrm{rank}(B''') = r$, 故 $\mathrm{rank}(\tilde{B}) = r-1$, 由數學歸納法的假設知道 $r-1 \le k$ 且 $r-1 \le n-1$, 於是 $r \le k+1$ 且 $r \le n$, 得到 $r \le \min(k+1, n) = \min(m, n)$。

以下再看如何將 A 轉化為 D 的樣式。由數學歸納法假設, 關於矩陣 $\tilde{B}$ 可利用有限次基本列運算及基本行運算轉化成矩陣 $\tilde{D} \in M_{k\times(n-1)}(\mathbb{F})$, 其中

$$\tilde{D} = \begin{bmatrix} I_{r-1} & O_4 \\ O_5 & O_6 \end{bmatrix}$$

其中 O_4, O_5, O_6 皆為零矩陣, 以及 $\tilde{D}_{ii} = 1, i = 1, 2, \ldots, r-1$, 而其它項 $\tilde{D}_{ij} = 0$。令

$$D = \left[\begin{array}{c|ccc} 1 & 0 & \cdots & 0 \\ \hline 0 & & & \\ \vdots & & \tilde{D} & \\ 0 & & & \end{array}\right],$$

因為我們可將 $\tilde{B}$ 經過有限次基本列運算及基本行運算轉化成 $\tilde{D}$, 得到 $\begin{bmatrix} O & \tilde{B} \end{bmatrix}$ 經過有限次的基本列運算及行運算而得 $\begin{bmatrix} O & \tilde{D} \end{bmatrix}$; 同樣地,

$$\begin{bmatrix} O \\ \tilde{B} \end{bmatrix}$$

可經過有限次的基本列運算及行運算轉化成

$$\begin{bmatrix} O \\ \tilde{D} \end{bmatrix},$$

又矩陣 D 的第一列與第一行在上述操作下使終保持不變, 所以我們可將 B''' 經有限次基本列運算及基本行運算而得 D。

由於 A 可轉化成 B''' 且 B''' 可轉化成 D, 而且每一個過程都是經有限次基本運算轉化而成, 故 A 經有限次基本運算可轉化成 D。

最後, 因矩陣 $\tilde{D}$ 滿足 $\tilde{D}_{ii} = 1$, 其中 $i = 1, 2, \ldots, r-1$, 而且其它項 $\tilde{D}_{ij} = 0$, 所以矩陣 D 滿足 $D_{ii} = 1$, 其中 $i = 1, 2, \ldots, r$, 而且其它項 $D_{ij} = 0$。

□

推論 6 (第 158 頁)**.** 令 $A \in M_{m\times n}(\mathbb{F})$, 其中 $\mathrm{rank}(A) = r$, 則存在可逆矩陣 $B \in M_{m\times m}(\mathbb{F})$ 與 $C \in M_{n\times n}(\mathbb{F})$ 使得 $D = BAC$, 其中

$$D = \begin{bmatrix} I_r & O_1 \\ O_2 & O_3 \end{bmatrix}$$

是一個 $m \times n$ 的矩陣, 在這個矩陣區塊的表示下, I_r 是 $r \times r$ 的單位矩陣, 而 O_1, O_2, O_3 皆為零矩陣, 其大小分別為 $r \times (n-r), (m-r) \times r$ 與 $(m-r) \times (n-r)$。

證明: 由 定理 **5**, 矩陣 A 可以經由有限次基本列運算以及基本行運算轉化成矩陣 D, 其中 D 符合命題中的所有條件; 換言之, 存在 $E_1, E_2, \ldots, E_p$ 皆為 $m \times m$ 的可逆矩陣以及 $G_1, G_2, \ldots, G_q$ 皆為 $n \times n$ 的可逆矩陣使得

$$D = E_p E_{p-1} \cdots E_2 E_1 A G_1 G_2 \cdots G_{q-1} G_q。$$

令 $B = E_p E_{p-1} \cdots E_2 E_1$ 與 $C = G_1 G_2 \cdots G_{q-1} G_q$, 則 B, C 皆為可逆矩陣, 而且 $D = BAC$。 □

關於矩陣的秩有一些常用的推論, 這些也給出相應的結果及其證明。

推論 7 (第 158 頁)**.**

(A) 任何矩陣的秩與其轉置矩陣的秩相同。

(B) 任何矩陣的秩等於由列向量生成的子空間之維度; 換言之, 矩陣的秩等於最大線性獨立列向量的數目。

(C) 任何矩陣由列向量所生成的子空間之維度與由行向量所生成的子空之維度相同, 而且它們的維度亦為矩陣的秩。

證明:

(A) 給定 $A \in M_{m\times n}(\mathbb{F})$, 假設 $\mathrm{rank}(A) = r$, 由 推論 **6**, 存在可逆矩陣 B 與 C 使得 $D = BAC$, 其中 D 的矩陣分塊左上角為單位矩陣 I_r, 其它矩陣區塊為零矩陣。因為 $D^t = (BAC)^t = C^t A^t B^t$, 而 B 與 C 可逆得知 B^t 與 C^t 可逆, 由 定理 **2**, $\mathrm{rank}(A^t) = \mathrm{rank}(C^t A^t B^t) = \mathrm{rank}(D^t)$。對於 $D^t \in M_{n\times m}(\mathbb{F})$, 因為 D^t 的矩陣分塊左上角為單位矩陣 I_r, 其它矩陣區塊為零矩陣, 所以 $\mathrm{rank}(D^t) = r$, 因此 $\mathrm{rank}(A^t) = \mathrm{rank}(D^t) = r = \mathrm{rank}(A)$。

(B) 給定 $A \in M_{m\times n}(\mathbb{F})$, 假設 $\mathrm{rank}(A) = r$, 存在可逆矩陣 B 與 C 使得 $D = BAC$, 其中 D 的矩陣分塊左上角為單位矩陣 I_r, 其它矩陣區塊為零矩陣。由最大線性獨立的定義知: $\mathrm{rank}(D)$ 也是它的最大線性獨立的列向量數, 因為 B, C 可逆, 且 $A = B^{-1} D C^{-1}$, 所以 D 在進行有限次基本列運算與基本行運算後可以轉化成 A, 所以矩陣 A 的秩也等於由列向量生成的子空間之維度; 換言之, 矩陣的秩等於最大線性獨立列向量的數目。

(C) 由 (B) 知道行向量與列向量的線性獨立數相同, 故兩者產生的子空間的維度亦相同。

□

推論 8 (第 159 頁)**.** 任何可逆矩陣必可表示為基本矩陣之乘積。

證明: 若 $A \in M_{n\times n}(\mathbb{F})$ 是可逆矩陣, 則 $\mathrm{rank}(A)=n$, 由 推論 **6** 得知, 存在可逆矩陣 B 及 C 使得 $D=BAC$, 其中 $D_{ij}=\delta_{ij}$, 於是 $D=I_n$; 亦即, $I_n=BAC$。

注意到 B 和 C 可以分別表示為基本矩陣的乘積, 即 $B=E_pE_{p-1}\cdots E_1$ 及 $C=G_1G_2\cdots G_q$, 其中 $E_1,E_2,\ldots,E_p$ 及 $G_1,G_2,\ldots,G_q$ 皆為基本矩陣, 於是 $A=B^{-1}I_nC^{-1}=B^{-1}C^{-1}$ 使得

$$A=E_1^{-1}E_2^{-1}\cdots E_p^{-1}G_q^{-1}G_{q-1}^{-1}\cdots G_1^{-1},$$

又基本矩陣的反矩陣亦為基本矩陣, 於是 A 是基本矩陣的乘積。 □

這個單元的最後想討論一個數學上必須加以確定的問題: 我們對於一個矩陣 A 的秩的定義是去探討它對應的左乘變換 $L_A:\mathbb{F}^n\to\mathbb{F}^m$ 的秩; 然而, 任何有限維向量空間之間的線性變換 $T:V\to W$, 只要我們分別選定好 V 與 W 的有序基底 β 與 γ, 那麼這個線性變換就有它對應的矩陣表現 $[T]_\beta^\gamma$, 而這個矩陣表現 $[T]_\beta^\gamma \overset{記}{=} A$ 又必須透過它的左乘變換 L_A 定義這個矩陣的秩, 那麼, 線性變換的秩和矩陣的秩會相同嗎? 此外, 一個線性變換的矩陣表現 $[T]_\beta^\gamma$ 會隨著有序基底的選取而異, 那麼對於一個線性變換 T, 在基底改變之下所得到的矩陣表現的秩也會一樣嗎? 以下定理將給出這個問題的結果。

定理 9 (第 152 頁)**.** *假設 $T:V\to W$ 是有限維向量空間之間的一種線性變換, 令 β 與 γ 分別為 V 與 W 的有序基底, 則 $\mathrm{rank}(T)=\mathrm{rank}([T]_\beta^\gamma)$。*

證明: 首先回顧以下線性變換與其矩陣表現所對應到的左乘變換之關係圖:

$$\begin{tikzcd} V \arrow[r, "T"] \arrow[d, "\phi_\beta"'] & W \arrow[d, "\phi_\gamma"] \\ \mathbb{F}^n \arrow[r, "L_A"] & \mathbb{F}^m \end{tikzcd}$$

圖 3.1: 與其矩陣表現所對應到的左乘變換之關係。

考慮 $\phi_\beta:V\to\mathbb{F}^n$, 其中 $\phi_\beta(\mathbf{v})=[\mathbf{v}]_\beta$ 是 V 相對於有序基底 β 的標準表現, 而 $\phi_\gamma:W\to\mathbb{F}^m$, 其中 $\phi_\gamma(\mathbf{w})=[\mathbf{w}]_\gamma$ 是 W 相對於有序基底 γ 的標準表現, 則 ϕ_β 與 ϕ_γ 都是同構變換, 而且記 $A=[T]_\beta^\gamma$, 則 $L_A(\phi_\beta(\mathbf{v}))=\phi_\gamma(T(\mathbf{v}))$。

假設 $\mathrm{rank}(T)=\dim(R(T))=r$, 因為 ϕ_γ 是可逆的, 而且 $R(T)$ 是 W 的子空間, 所以 $\dim(\phi_\gamma(R(T)))=r$, 以下討論將證明 $\phi_\gamma(R(T))=R(L_A)$, 這麼一來就有

$$\mathrm{rank}([T]_\beta^\gamma)=\mathrm{rank}(A)=\mathrm{rank}(L_A)=\dim(R(L_A))=\dim(\phi_\gamma(R(T)))=r=\mathrm{rank}(T)。$$

欲證 $\phi_\gamma(R(T))=R(L_A)$, 首先證明 $\phi_\gamma(R(T))\subset R(L_A)$: 對任何 $[\mathbf{w}]_\gamma\in\phi_\gamma(R(T))$, 則 $\mathbf{w}\in R(T)$, 存在 $\mathbf{v}\in V$ 使得 $\mathbf{w}=T(\mathbf{v})$, 且 $[\mathbf{v}]_\beta=\phi_\beta(\mathbf{v})\in\mathbb{F}^n$, 則

$$L_A([\mathbf{v}]_\beta)=[T]_\beta^\gamma[\mathbf{v}]_\gamma=[T(\mathbf{v})]_\gamma=[\mathbf{w}]_\gamma,$$

所以 $[\mathbf{w}]\in R(L_A)$。因此 $\phi_\gamma(R(T))\subset R(L_A)$。

再證 $R(L_A) \subset \phi_\gamma(R(T))$: 對任何 $\bar{\mathbf{w}} = (b_1, b_2, \ldots, b_m) \in R(L_A)$, 其中 $A = [T]_\beta^\gamma$, 存在 $\bar{\mathbf{v}} = (a_1, a_2, \ldots, a_n) \in \mathbb{F}^n$ 使得 $L_A(\bar{\mathbf{v}}) = A\bar{\mathbf{v}} = \bar{\mathbf{w}}$, 將 $A\bar{\mathbf{v}} = \bar{\mathbf{w}}$ 寫下, 則為 $b_i = \sum\limits_{j=1}^{n} a_j A_{ij}$, 其中 $i = 1, 2, \ldots, m$。記 $\beta = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ 與 $\gamma = \{\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_m\}$ 分別為 V 與 W 的有序基底, 考慮 $\mathbf{v} = \sum\limits_{j=1}^{n} a_j \mathbf{v}_j \in V$, 則

$$T(\mathbf{v}) = T\left(\sum_{j=1}^{n} a_j \mathbf{v}_j\right) = \sum_{j=1}^{n} a_j T(\mathbf{v}_j) = \sum_{j=1}^{n} a_j A_{ij} \mathbf{w}_i,$$

於是

$$\phi_\gamma(T(\mathbf{v})) = \phi_\gamma\left(\sum_{j=1}^{n} a_j A_{ij} \mathbf{w}_i\right) = \left(\sum_{j=1}^{n} a_j A_{1j}, \sum_{j=1}^{n} a_j A_{2j}, \ldots, \sum_{j=1}^{n} a_j A_{mj}\right) = \bar{\mathbf{w}},$$

得到 $\bar{\mathbf{w}} \in \phi_\gamma(R(T))$。因此 $R(L_A) \subset \phi_\gamma(R(T))$。 □

這個定理的結果也可以順勢回答前段文字的另一個提問: 對於線性變換 $T : V \to W$, 若在 V 與 W 上分別選用了不同的有序基底 β' 與 γ' 以得到這個線性變換的矩陣表現 $[T]_{\beta'}^{\gamma'}$, 由 定理 **9** 知: $\operatorname{rank}([T]_{\beta'}^{\gamma'}) = \operatorname{rank}(T)$, 所以這些矩陣的秩與基底選取無關。

由 定理 **2**, 我們知道一個矩陣若左乘或是右乘一個可逆矩陣的話會保秩, 那麼對於一般的情況下會有什麼結果呢? 以下命題將探討兩個矩陣相乘後的秩和各別矩陣的秩之間的關係。討論這些內容的其中一個用意是為了在第四章探討行列式的理論時會需要用到。

命題 **10.** 令 $T : V \to W$ 及 $S : W \to Z$ 是有限維向量空間 V, W, Z 上的線性變換, 令矩陣 A, B 滿足 AB 有意義, 則

(A) $\operatorname{rank}(ST) \leq \operatorname{rank}(S)$。

(B) $\operatorname{rank}(ST) \leq \operatorname{rank}(T)$。

(C) $\operatorname{rank}(AB) \leq \operatorname{rank}(A)$。

(D) $\operatorname{rank}(AB) \leq \operatorname{rank}(B)$。

證明: 以下證明順序為 (A) (C) (D) (B)。注意到這樣的證明順序只是一種將這些結果快速而簡便地鋪陳完畢, 它並非是必要的證明順序, 我們可以各自獨立地證明每個結果。

(A) 因為 $R(T) \subset W$, 所以 $R(ST) = ST(V) = S(R(T)) \subset S(W) = R(S)$, 因此

$$\operatorname{rank}(ST) = \dim(R(ST)) \leq \dim(R(S)) = \operatorname{rank}(S)。$$

(C) 由 (A) 知 $\operatorname{rank}(AB) = \operatorname{rank}(L_{AB}) = \operatorname{rank}(L_A L_B) \leq \operatorname{rank}(L_A) = \operatorname{rank}(A)$。

(D) 由 (C) 以及 推論 **7** 得到 $\operatorname{rank}(AB) = \operatorname{rank}((AB)^t) = \operatorname{rank}(B^t A^t) = \operatorname{rank}(B^t) = \operatorname{rank}(B)$。

(B) 令 α, β, γ 分別為 V, W, Z 的有序基底, 且令 $A' = [S]_\beta^\gamma$ 以及 $B' = [T]_\alpha^\beta$, 因為 $A'B' = [ST]_\alpha^\gamma$, 所以 $\operatorname{rank}(ST) = \operatorname{rank}(A'B') \leq \operatorname{rank}(B') = \operatorname{rank}(T)$。

□

3.3 矩陣的反矩陣

在很多情況下, 我們會需要知道一個矩陣是否可逆, 若得知矩陣可逆, 還要得到可逆矩陣的反矩陣表達式。這個單元提供一種確定矩陣是否可逆以及找到可逆矩陣的反矩陣方法, 主要是透過基本列運算的方式逐步求得反矩陣。為此, 我們要先提出增廣矩陣的概念。

定義 **1** (第 161 頁)**.** 假設 $A \in M_{m\times n}(\mathbb{F})$ 及 $B \in M_{m\times p}(\mathbb{F})$, 定義 $[A|B] \in M_{m\times(n+p)}(\mathbb{F})$, 它的前 n 行即為矩陣 A 的 n 個行向量, 後 p 行為 B 的 p 個行向量; 也就是說,

$$\left[\begin{array}{c|c} A & B \end{array}\right] = \left[\begin{array}{c|c|c|c|c|c|c|c} \mathbf{a}_1 & \mathbf{a}_2 & \cdots & \mathbf{a}_n & \mathbf{b}_1 & \mathbf{b}_2 & \cdots & \mathbf{b}_p \end{array}\right],$$

其中 $\mathbf{a}_i, i = 1, 2, \ldots, n$ 及 $\mathbf{b}_j, j = 1, 2, \ldots, p$ 分別表示 A 的第 i 行及 B 的第 j 行。我們稱矩陣 $[A|B]$ 為 A 和 B 的 *增廣矩陣* (augmented matrix)。

以下說明如何對於一個可逆矩陣找到其反矩陣的表示法: 假設 A 是一個可逆的 $n \times n$ 的矩陣, 考慮 A 和 I_n 的增廣矩陣 $[A|I_n] \in M_{n\times(2n)}(\mathbb{F})$, 因為

$$A^{-1}\left[\begin{array}{c|c} A & I_n \end{array}\right] = \left[\begin{array}{c|c} A^{-1}A & A^{-1}I_n \end{array}\right] = \left[\begin{array}{c|c} I_n & A^{-1} \end{array}\right],$$

又由單元 3.2 的 *推論* **8** 知道: 可逆矩陣 A^{-1} 必能表示為基本矩陣的乘積, 記 $A^{-1} = E_pE_{p-1}\cdots E_1$, 所以透過上述的增廣矩陣, 左邊試圖從 A 出發透過有限步驟的基本列運算與基本行運算將矩陣轉變為單位矩陣, 而右邊逐一記錄每一次基本矩陣的操作, 最終即得 A 的反矩陣。

注意到從單元 3.2 的 *推論* **6**, 若將增廣矩陣 $[A|I_n]$ 逐步透過基本列運算與基本行運算的方式得到左邊形成一個秩為 $r, r < n$ 的矩陣時 (你可能不需要化簡完全, 比方說只要發現到有某一列或某一行全為零的時候, 就可以確定矩陣的秩不可能是 n), 那麼就可以得到 A 是不可逆這個結論。

在給出實例示範之前, 我們想要定義一些簡便的記號以註記基本列運算與基本行運算:

- $R_{p,q}$ 代表將矩陣的第 p 列與第 q 列互換。
- $R_p^{(k)}$ 代表將矩陣的第 p 列乘以 k 倍, 其中 $k \in \mathbb{F}, k \neq 0$。
- $R_{q,p}^{(k)}$ 代表將矩陣的第 q 列乘以 k 倍加到第 p 列, 其中 $k \in \mathbb{F}, k \neq 0$。
- $C_{p,q}$ 代表將矩陣的第 p 行與第 q 行互換。
- $C_p^{(k)}$ 代表將矩陣的第 p 行乘以 k 倍, 其中 $k \in \mathbb{F}, k \neq 0$。
- $C_{q,p}^{(k)}$ 代表將矩陣的第 q 行乘以 k 倍加到第 p 行, 其中 $k \in \mathbb{F}, k \neq 0$。

引進上述記號後, 我們可以把第 I 型基本矩陣寫成 $E(R_{p,q;n}) = R_{p,q}(I_n)$ 或 $E(C_{p,q;n}) = C_{p,q}(I_n)$, 第 II 型基本矩陣寫成 $E(R_{p;n}^{(k)}) = R_p^{(k)}(I_n)$ 或 $E(C_{p;n}^{(k)}) = C_p^{(k)}(I_n)$, 而第 III 型基本矩陣寫成 $E(R_{q,p;n}^{(k)}) = R_{q,p}^{(k)}(I_n)$ 或 $E(C_{q,p;n}^{(k)}) = C_{q,p}^{(k)}(I_n)$。注意到這些記號並非完全公認的符號, 你可能會在其它的文獻或書籍中看到基本列運算、基本行運算與基本矩陣用別的方式註記。

例 **2.** 試決定矩陣

$$A = \begin{bmatrix} 1 & 1 & 1 \\ 2 & 3 & -1 \\ 1 & -2 & 1 \end{bmatrix}$$

是否可逆, 若可逆, 反矩陣為何?

解**.** 從增廣矩陣 $[A|I_3]$ 出發, 目標是將 $[A|I_3]$ 看看能否轉換為 $[I_3|B]$, 若可以的話, 那麼 $B = A^{-1}$。於是

$$\left[\begin{array}{ccc|ccc} 1 & 1 & 1 & 1 & 0 & 0 \\ 2 & 3 & -1 & 0 & 1 & 0 \\ 1 & -2 & 1 & 0 & 0 & 1 \end{array}\right] \xrightarrow{R_{1,2}^{(-2)}} \left[\begin{array}{ccc|ccc} 1 & 1 & 1 & 1 & 0 & 0 \\ 0 & 1 & -3 & -2 & 1 & 0 \\ 1 & -2 & 1 & 0 & 0 & 1 \end{array}\right]$$

$$\xrightarrow{R_{1,3}^{(-1)}} \left[\begin{array}{ccc|ccc} 1 & 1 & 1 & 1 & 0 & 0 \\ 0 & 1 & -3 & -2 & 1 & 0 \\ 0 & -3 & 0 & -1 & 0 & 1 \end{array}\right] \xrightarrow{R_{3}^{(-\frac{1}{3})}} \left[\begin{array}{ccc|ccc} 1 & 1 & 1 & 1 & 0 & 0 \\ 0 & 1 & -3 & -2 & 1 & 0 \\ 0 & 1 & 0 & \frac{1}{3} & 0 & -\frac{1}{3} \end{array}\right]$$

$$\xrightarrow{R_{2,3}} \left[\begin{array}{ccc|ccc} 1 & 1 & 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & \frac{1}{3} & 0 & -\frac{1}{3} \\ 0 & 1 & -3 & -2 & 1 & 0 \end{array}\right] \xrightarrow{R_{2,1}^{(-1)}} \left[\begin{array}{ccc|ccc} 1 & 0 & 1 & \frac{2}{3} & 0 & \frac{1}{3} \\ 0 & 1 & 0 & \frac{1}{3} & 0 & -\frac{1}{3} \\ 0 & 1 & -3 & -2 & 1 & 0 \end{array}\right]$$

$$\xrightarrow{R_{2,3}^{(-1)}} \left[\begin{array}{ccc|ccc} 1 & 0 & 1 & \frac{2}{3} & 0 & \frac{1}{3} \\ 0 & 1 & 0 & \frac{1}{3} & 0 & -\frac{1}{3} \\ 0 & 0 & -3 & -\frac{7}{3} & 1 & \frac{1}{3} \end{array}\right] \xrightarrow{R_{3}^{(-\frac{1}{3})}} \left[\begin{array}{ccc|ccc} 1 & 0 & 1 & \frac{2}{3} & 0 & \frac{1}{3} \\ 0 & 1 & 0 & \frac{1}{3} & 0 & -\frac{1}{3} \\ 0 & 0 & 1 & \frac{7}{9} & -\frac{1}{3} & -\frac{1}{9} \end{array}\right]$$

$$\xrightarrow{R_{3,1}^{(-1)}} \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & -\frac{1}{9} & \frac{1}{3} & \frac{4}{9} \\ 0 & 1 & 0 & \frac{1}{3} & 0 & -\frac{1}{3} \\ 0 & 0 & 1 & \frac{7}{9} & -\frac{1}{3} & -\frac{1}{9} \end{array}\right],$$

因為增廣矩陣的左邊可以順利地轉化成單位矩陣, 所以矩陣 A 是可逆矩陣, 而且從這些操作的過程中可以解讀出

$$A^{-1} = \begin{bmatrix} -\frac{1}{9} & \frac{1}{3} & \frac{4}{9} \\ \frac{1}{3} & 0 & -\frac{1}{3} \\ \frac{7}{9} & -\frac{1}{3} & -\frac{1}{9} \end{bmatrix}。$$

從上述的例子, 我們知道這是一個尋求逆矩陣的方法, 它的特色是可以把每一個步驟看得清楚, 這對於若未來想要把這些過程實現在寫電腦程式來說是好的呈現方式, 但是這在進行人工求解的時候並不是一個好的方法, 特別是當數字變得更不友善、矩陣變大甚至是矩陣的項帶有變數的時候都會造成更大的困擾, 所以這不是一個最快速求得答案的方法。在之後的章節會針對尋求可逆矩陣的反矩陣這個問題給出其它的理論, 並給出不一樣的尋求反矩陣的方法。而各位到時候也可以再比較看看各種方法的優缺點。

3.4 線性聯立方程式解的存在性與唯一性

給定 $a_{ij}, b_i \in \mathbb{F}$, 其中 $1 \le i \le m, 1 \le j \le n$, 考慮帶有 n 個未知數 $x_1, x_2, \ldots, x_n$ 的 m 個線性聯立方程式

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2 \\ \qquad\qquad\qquad \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = b_m, \end{cases}$$

透過矩陣的運算, 線性聯立方程式可以寫成 $A\mathbf{x} = \mathbf{b}$, 其中

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}, \quad \mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}, \quad \mathbf{b} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix}。$$

這裡我們稱矩陣 A 為 *係數矩陣* (coefficient matrix)。

我們感興趣的問題是線性聯立方程式 $A\mathbf{x} = \mathbf{b}$ 是否有解? 如果有解的話, 解是否唯一? 這個單元就是要給出線性聯立方程式解的存在性與唯一性的理論性探討及其證明。

定義 1 (第 171 頁)**.** 給定帶有 n 個未知數的 m 個線性聯立方程式 $A\mathbf{x} = \mathbf{b}$, 若 $\mathbf{b} = \mathbf{0}$, 我們稱聯立方程式是 *齊次的* (homogeneous); 若 $\mathbf{b} \neq \mathbf{0}$, 則稱聯立方程式是 *非齊次的* (nonhomogeneous)。

首先我們探討齊次線性聯立方程式 $A\mathbf{x} = \mathbf{0}$ 的情況。因為 $\mathbf{x} = \mathbf{0}$ 滿足 $A\mathbf{x} = \mathbf{0}$, 所以齊次線性聯立方程式至少有一個解。至於它只是唯一解或是有無限多解, 以下定理及推論將回答這個問題。

定理 2 (第 171 頁)**.** *考慮齊次線性聯立方程式 $A\mathbf{x} = \mathbf{0}$, 令 K_h 是齊次線性聯立方程式的解集合, 則 $K_h = N(L_A)$; 換言之, 解集合 K_h 是 $\mathbb{F}^n$ 的子空間, 而且 K_h 的維度是 $n - \operatorname{rank}(L_A) = n - \operatorname{rank}(A)$。*

證明: 根據定義, $K_h = \{\mathbf{v} \in \mathbb{F}^n | A\mathbf{v} = \mathbf{0}\} = N(L_A)$。由維度定理 (Dimension Theorem), 因為 $\operatorname{nullity}(L_A) + \operatorname{rank}(L_A) = \dim(\mathbb{F}^n) = n$, 所以

$$\dim(K_h) = \dim(N(L_A)) = \operatorname{nullity}(L_A) = n - \operatorname{rank}(L_A) = n - \operatorname{rank}(A)。$$

□

推論 3 (第 171 頁)**.** *給定矩陣 $A \in M_{m\times n}(\mathbb{F})$, 若 $m < n$, 則齊次線性聯立方程式 $A\mathbf{x} = \mathbf{0}$ 有非零解。*

證明: 假設 $m < n$, 則 $\operatorname{rank}(A) = \operatorname{rank}(L_A) \le \min(m, n) \le m$, 由 **定理 2** 得知

$$\dim(K_h) = n - \operatorname{rank}(L_A) \ge n - m > 0,$$

其中 $K_h = N(L_A)$ 是齊次線性聯立方程式的解集合。因為 $\dim(K_h) > 0$, 所以 $K_h \neq \{\mathbf{0}\}$, 因此存在非零向量 $\mathbf{x}_h \in K_h$, 即 $\mathbf{x}_h$ 是聯立方程式 $A\mathbf{x} = \mathbf{0}$ 的非零解。 □

關於 推論 **3**, 有幾件事情是值得討論的:

(1) 探討線性方程式通常會和自由度 (degrees of freedom) 的概念一同對照, 因為向量空間的一個維度代表一個自由選取的方向, 即一個自由度, 當我們設立了一個方程式, 因為這個方程式當中的某個變數可以被這個方程式的其它變數表示, 所以一個方程式會限制一個自由度。

從這個原理來看, 在討論線性聯立方程式的時候, 就要再顧慮到這些方程式之間是不是給出不同的限制。在解聯立的過程中, 如果透過加減消去法可以把某一個方程式全部消掉, 那麼這意味著這些限制式當中會有一個限制是多餘的。所以我們可以保守一點地說: m 個方程式聯立最多會限制 m 個未知數的任意性。

從上述的觀點再重新思考 推論 **3** 的話就會覺得很自然, 因為原本有 n 個變數, 但因為只有 m 條方程式, 所以最多只會限制 m 個變數, 在 $m < n$ 的情況下, 自然還會有變數是自由的, 所以這樣對應到的齊次線性聯立方程式有無限多解。

(2) 這裡應特別留意 推論 **3** 的命題當中「若 P 則 Q」的語句, 這裡只是回答在 $m < n$ 的情況下齊次線性聯立方程式會有無限多組解, 並沒有提到在 $m \geq n$ 的時候發生了什麼事。實際上, 若 $m \geq n$, 齊次線性聯立方程式可能有唯一解, 也可能有無限多組解。我們同樣從自由度的原理去想這個問題的話也會覺得這件事是合理的, 因為有 n 個變數, 但是有 m 個方程式, 看似給出了 m 個限制, 然而這些限制之間有些可能是重覆的, 把那些重覆的限制刪除後如果個數與 n 一致, 那麼就有唯一解, 在限制條件少於 n 時就會有無限多解。

再來要討論的是非齊次線性聯立方程式 $A\mathbf{x} = \mathbf{b}$ 的解。下面的定理要描述的是 $A\mathbf{x} = \mathbf{b}$ 的解集合與它對應的齊次線性聯立方程式 $A\mathbf{x} = \mathbf{0}$ 的解空間之間有關聯。

定理 4 (第 172 頁)**.** *若 $\mathbf{x}_p$ 是線性聯立方程式 $A\mathbf{x} = \mathbf{b}$ 的一個解, 記 K 是 $A\mathbf{x} = \mathbf{b}$ 的解集合, 而 K_h 是對應的齊次線性聯立方程式 $A\mathbf{x} = \mathbf{0}$ 的解空間, 則*

$$K = \{\mathbf{x}_p\} + K_h \overset{\text{定義}}{=\!=} \{\mathbf{x}_p + \mathbf{x}_h : \mathbf{x}_h \in K_h\}。$$

證明: 令 $\mathbf{x}_p$ 是 $A\mathbf{x} = \mathbf{b}$ 的一個解。若有向量 $\mathbf{x}_0 \in K$, 則 $A\mathbf{x}_0 = \mathbf{b}$。因為

$$A(\mathbf{x}_0 - \mathbf{x}_p) = A\mathbf{x}_0 - A\mathbf{x}_p = \mathbf{b} - \mathbf{b} = \mathbf{0},$$

所以 $\mathbf{x}_0 - \mathbf{x}_p \in K_h$, 故存在 $\mathbf{x}_h \in K_h$ 使得 $\mathbf{x}_0 - \mathbf{x}_p = \mathbf{x}_h$, 於是 $\mathbf{x}_0 = \mathbf{x}_p + \mathbf{x}_h \in \{\mathbf{x}_p\} + K_h$。因此 $K \subset \{\mathbf{x}_p\} + K_h$。

假設 $\mathbf{x}_0 \in \{\mathbf{x}_p\} + K_h$, 則 $\mathbf{x}_0 = \mathbf{x}_p + \mathbf{x}_h$, 其中 $\mathbf{x}_h \in K_h$。於是

$$A\mathbf{x}_0 = A(\mathbf{x}_p + \mathbf{x}_h) = A\mathbf{x}_p + A\mathbf{x}_h = \mathbf{b} + \mathbf{0} = \mathbf{b},$$

故 $\mathbf{x}_0 \in K$, 因此 $\{\mathbf{x}_p\} + K_h \subset K$。

綜合上述討論, 我們得到 $K = \{\mathbf{x}_p\} + K_h$。 □

注意到上面的定理敘述是「假設有一個解 $\mathbf{x}_p$」的情況下, 則 $A\mathbf{x} = \mathbf{b}$ 的解集合有代數結構上的分解。而線性聯立方程式 $A\mathbf{x} = \mathbf{b}$ 有可能無解, 這個情況會是 $K = \emptyset$。

這裡想要提供一個視覺化的觀點說明 定理 4 的結果。如圖 3.2, 因為線性聯立方程式有 n 個未知數, 所以先設立 $\mathbb{R}^n$ 這個向量空間, 我們知道齊次線性聯立方程式 $A\mathbf{x} = \mathbf{0}$ 的解集合是一個向量子空間 K_h, 所以這個集合會通過坐標原點 $\mathbf{0}$, 而非齊次線性聯立方程式 $A\mathbf{x} = \mathbf{b}$ 的解集合是 K, 定理的結果是在說明兩個解集合之間彼此平行。

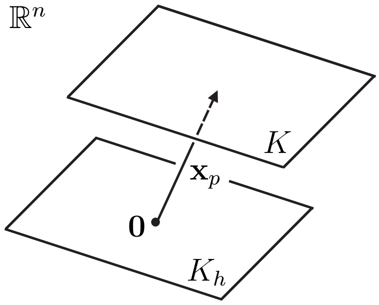

圖 3.2: 非齊次線性聯立方程式與對應的齊次線性聯立方程式解集合之間的關係。

定理 5 (第 174 頁)**.** 考慮 n 個未知數的 n 條線性聯立方程式 $A\mathbf{x} = \mathbf{b}$。則這個線性聯立方程式存在唯一解若且唯若 $A \in M_{n\times n}(\mathbb{F})$ 是一個可逆矩陣。

證明: ($\Rightarrow$) 假設 n 個未知數的 n 條線性聯立方程式 $A\mathbf{x} = \mathbf{b}$ 存在唯一解 $\mathbf{x}_0$, 令 K_h 表示對應的齊次線性聯立方程式 $A\mathbf{x} = \mathbf{0}$ 的解空間, 由 定理 4 知道 $\{\mathbf{x}_0\} = \{\mathbf{x}_0\} + K_h$, 此時等式只有在 $K_h = \{\mathbf{0}\}$ 的時候才成立, 於是 $N(L_A) = K_h = \{\mathbf{0}\}$, 得知 A 是可逆矩陣。

若 A 是可逆矩陣, 考慮 $\mathbf{x}_0 = A^{-1}\mathbf{b}$, 代入聯立方程式得到 $A(A^{-1}\mathbf{b}) = (AA^{-1})\mathbf{b} = I_n\mathbf{b} = \mathbf{b}$, 所以 $\mathbf{x}_0 = A^{-1}\mathbf{b}$ 為一個解。若 $\mathbf{x}_0$ 滿足 $A\mathbf{x}_0 = \mathbf{b}$, 將方程式兩邊同乘 A^{-1} 後得到 $A^{-1}A\mathbf{x}_0 = A^{-1}\mathbf{b}$, 即 $\mathbf{x}_0 = A^{-1}\mathbf{b}$, 故線性聯立方程式 $A\mathbf{x} = \mathbf{b}$ 的解有唯一性。 □

這裡應注意的是上述定理只是在觀察 n 個未知數的 n 條線性聯立方程式的情況。對於 n 個未知數的 m 條線性聯立方程式 $A\mathbf{x} = \mathbf{b}$ 來說, 方程式可能無解, 也可能有解, 而且有解的情況也可能是有唯一解又或者是有無限多解。下面的定理將給出決定線性聯立方程式至少有一解的判別法。這個判別法涉及聯立方程式 $A\mathbf{x} = \mathbf{b}$ 係數矩陣 A 的秩及增廣矩陣 $[A|\mathbf{b}]$ 的秩。

定理 6 (第 174 頁)**.** 考慮線性聯立方程式 $A\mathbf{x} = \mathbf{b}$, 則方程式有解若且唯若 $\operatorname{rank}(A) = \operatorname{rank}([A|\mathbf{b}])$。

證明: 首先, 我們知道 $A\mathbf{x} = \mathbf{b}$ 有一解等同於 $\mathbf{b} \in R(L_A)$。此外, 我們又有

$$R(L_A) = \operatorname{span}(\{\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n\}),$$

即 A 的行向量生成集。於是 $A\mathbf{x} = \mathbf{b}$ 有一解若且唯若 $\mathbf{b} \in \operatorname{span}(\{\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n\})$, 但 $\mathbf{b} \in \operatorname{span}(\{\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n\})$ 若且唯若 $\operatorname{span}(\{\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n\}) = \operatorname{span}(\{\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n, \mathbf{b}\})$, 而這件事等同於

$$\dim(\operatorname{span}(\{\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n\})) = \dim(\operatorname{span}(\{\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n, \mathbf{b}\})),$$

故 $\operatorname{rank}(A) = \operatorname{rank}([A|\mathbf{b}])$。 □

例 **7.** 考慮線性聯立方程式

$$\begin{cases} x_1 + x_2 + x_3 = 3 \\ x_1 - x_2 + x_3 = 3 \\ x_1 \qquad\ + x_3 = 2, \end{cases}$$

記

$$A = \begin{bmatrix} 1 & 1 & 1 \\ 1 & -1 & 1 \\ 1 & 0 & 1 \end{bmatrix}, \quad \mathbf{b} = \begin{bmatrix} 3 \\ 3 \\ 2 \end{bmatrix},$$

首先計算矩陣 A 的秩:

$$\begin{bmatrix} 1 & 1 & 1 \\ 1 & -1 & 1 \\ 1 & 0 & 1 \end{bmatrix} \xrightarrow{R_{1,2}^{(-1)}} \begin{bmatrix} 1 & 1 & 1 \\ 0 & -2 & 0 \\ 1 & 0 & 1 \end{bmatrix} \xrightarrow{R_{1,3}^{(-1)}} \begin{bmatrix} 1 & 1 & 1 \\ 0 & -2 & 0 \\ 0 & -1 & 0 \end{bmatrix}$$

$$\xrightarrow{R_2^{(-\frac{1}{2})}} \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 0 \\ 0 & -1 & 0 \end{bmatrix} \xrightarrow{R_{2,1}^{(-1)}} \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & -1 & 0 \end{bmatrix} \xrightarrow{R_{2,3}^{(1)}} \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

$$\xrightarrow{C_{1,3}^{(-1)}} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix},$$

所以 $\operatorname{rank}(A) = 2$。再看增廣矩陣 $[A|\mathbf{b}]$ 的秩:

$$\begin{bmatrix} 1 & 1 & 1 & 3 \\ 1 & -1 & 1 & 3 \\ 1 & 0 & 1 & 2 \end{bmatrix} \xrightarrow{R_{1,2}^{(-1)}} \begin{bmatrix} 1 & 1 & 1 & 3 \\ 0 & -2 & 0 & 0 \\ 1 & 0 & 1 & 2 \end{bmatrix} \xrightarrow{R_{1,3}^{(-1)}} \begin{bmatrix} 1 & 1 & 1 & 3 \\ 0 & -2 & 0 & 0 \\ 0 & -1 & 0 & -1 \end{bmatrix}$$

$$\xrightarrow{R_2^{(-\frac{1}{2})}} \begin{bmatrix} 1 & 1 & 1 & 3 \\ 0 & 1 & 0 & 0 \\ 0 & -1 & 0 & -1 \end{bmatrix} \xrightarrow{R_{2,1}^{(-1)}} \begin{bmatrix} 1 & 0 & 1 & 3 \\ 0 & 1 & 0 & 0 \\ 0 & -1 & 0 & -1 \end{bmatrix} \xrightarrow{R_{2,3}^{(1)}} \begin{bmatrix} 1 & 0 & 1 & 3 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix}$$

$$\xrightarrow{C_{1,3}^{(-1)}} \begin{bmatrix} 1 & 0 & 0 & 3 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix} \xrightarrow{C_{3,4}} \begin{bmatrix} 1 & 0 & 3 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \end{bmatrix} \xrightarrow{R_3^{(-1)}} \begin{bmatrix} 1 & 0 & 3 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}$$

$$\xrightarrow{R_{3,1}^{(-3)}} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix},$$

得到 $\operatorname{rank}([A|\mathbf{b}]) = 3$。

因為 $\operatorname{rank}(A) = 2$ 以及 $\operatorname{rank}([A|\mathbf{b}]) = 3$, 所以線性聯立方程式無解。

3.5 高斯消去法

這一個單元想要透過單元 3.4 介紹的線性聯立方程式理論將單元 3.1 處理線性聯立方程式的求解過程解釋清楚。這個求解的過程數學上稱為 *高斯消去法* (Gaussian elimination)。

在說明高斯消去法之前, 我們先給出關於聯立方程式的求解意義。

定義 **1** (第 182 頁)**.** 給定兩個含有 n 個未知數的 m 條線性聯立方程式, 若它們有相同的解集合, 則稱這兩個線性聯立方程式 *等價* (equivalent)。

定理 **2** (第 182 頁)**.** *假設* $A\mathbf{x} = \mathbf{b}$ *是有* n *個未知數的* m *條線性聯立方程式, 若* $C \in M_{m\times m}(\mathbb{F})$ *是一個可逆矩陣, 則線性聯立方程式* $(CA)\mathbf{x} = C\mathbf{b}$ *等價於* $A\mathbf{x} = \mathbf{b}$。

證明: 令 K 與 $\bar{K}$ 分別為 $A\mathbf{x} = \mathbf{b}$ 與 $(CA)\mathbf{x} = C\mathbf{b}$ 的解集合, 若 $\mathbf{x}_0 \in K$, 則 $A\mathbf{x}_0 = \mathbf{b}$, 得到 $(CA)\mathbf{x}_0 = C(A\mathbf{x}_0) = C\mathbf{b}$, 於是 $\mathbf{x}_0 \in \bar{K}$。因此 $K \subset \bar{K}$。

若 $\mathbf{x}_0 \in \bar{K}$, 則 $(CA)\mathbf{x}_0 = C\mathbf{b}$, 於是

$$A\mathbf{x}_0 = I_m A\mathbf{x}_0 = (C^{-1}C)A\mathbf{x}_0 = C^{-1}(CA)\mathbf{x}_0 = C^{-1}C\mathbf{b} = (C^{-1}C)\mathbf{b} = I_m\mathbf{b} = \mathbf{b},$$

所以 $\mathbf{x}_0 \in K$。因此 $\bar{K} \subset K$。

由上討論得知 $K = \bar{K}$, 即線性聯立方程式 $(CA)\mathbf{x} = C\mathbf{b}$ 等價於 $A\mathbf{x} = \mathbf{b}$。 □

推論 **3** (第 182 頁)**.** *假設* $A\mathbf{x} = \mathbf{b}$ *表示含有* n *個未知數的* m *條線性聯立方程式, 若矩陣* $[A'|\mathbf{b}']$ *是由增廣矩陣* $[A|\mathbf{b}]$ *經有限次基本列運算而得, 則* $A'\mathbf{x} = \mathbf{b}'$ *等價於* $A\mathbf{x} = \mathbf{b}$。

證明: 因 $[A'|\mathbf{b}']$ 是由 $[A|\mathbf{b}]$ 經有限次基本列運算而得, 則存在基本矩陣 $E_1, E_2, \ldots, E_p \in M_{m\times m}(\mathbb{F})$ 使得

$$E_p \cdots E_2 E_1 \left[\begin{array}{c|c} A & \mathbf{b} \end{array}\right] = \left[\begin{array}{c|c} E_p \cdots E_2 E_1 A & E_p \cdots E_2 E_1 \mathbf{b} \end{array}\right] = \left[\begin{array}{c|c} A' & \mathbf{b}' \end{array}\right]。$$

令 $C = E_p \cdots E_2 E_1$, 因為 $E_1, E_2, \ldots, E_p$ 皆為可逆矩陣, 所以 C 是可逆矩陣。因為 $A' = CA$ 以及 $\mathbf{b}' = C\mathbf{b}$, 所以由 *定理* **2** 得知 $A'\mathbf{x} = \mathbf{b}'$ 等價於 $A\mathbf{x} = \mathbf{b}$。 □

回顧單元 3.1 的最後曾經討論過基本列運算與基本行運算和解聯立方程式之間的關聯, 只有列運算才與求解的過程有一致性, 若是將矩陣操作行運算則會造成未知數的位置互換或是改用新的未知數重新包裝原問題, 對於我們想要知道原始未知數的解沒有什麼幫助。引進行運算的用意是在幫助我們確定一個矩陣的秩, 然後再從秩去解讀解的存在性或是唯一性, 這個概念已經跳脫出實質求解的層次。於是接下來要介紹的高斯消去法之過程中, 我們完全避免行運算的任何操作以便到時候能直接得到聯立方程式的解。

在欠缺矩陣的行運算操作下, 我們自然不能預期一個矩陣可以轉化成像是單元 3.2 的 *定理* **5** 中矩陣 D 的樣式; 也就是說, 我們不能期待一個矩陣最終能轉化成左上角矩陣區塊是一個 $r \times r$ 的單位矩陣而其它的矩陣區塊都是零矩陣的樣式, 只能退而求其次地盡可能地化簡。於是我們先給出簡約列梯形式的定義, 然後再解釋如何利用高斯消去法以得到一個矩陣對應的簡約列梯形式。

定義 4 (第 185 頁)**.** 一個矩陣稱為 *簡約列梯形式* (reduced row echelon form) 意指該矩陣滿足以下三個條件:

(1) 含有非零項的任一列排在全為零項之任一列的上面。

(2) 各列中第一個非零項是其所在的行當中唯一的非零項。

(3) 各列中第一個非零項是 1, 而且它出現在前一列中第一個非零項的右邊的某一行。

例 **5.** 試說明為何以下矩陣不是簡約列梯形式:

$$A=\begin{bmatrix}0&0&0\\1&1&0\\0&0&1\end{bmatrix},\quad B=\begin{bmatrix}1&1&0\\0&1&0\\1&0&1\end{bmatrix},\quad C=\begin{bmatrix}0&1&0&2\\1&0&0&1\\0&0&1&1\end{bmatrix},\quad D=\begin{bmatrix}2&0&0\\0&1&0\end{bmatrix}。$$

解**.**

- 關於矩陣 A, 全為零項的列排在非零項的列的上面, 不符合簡約列梯形式的條件 (1)。
- 關於矩陣 B, 因為第一行含有一個以上的非零項, 不符合簡約列梯形式的條件 (2)。
- 關於矩陣 C, 因為第二列的第一個非零項不在第一列的第一個非零項的右邊, 不符合簡約列梯形式的條件 (3)。
- 關於矩陣 D, 因為第一列的非零項不是 1, 不符合簡約列梯形式的條件 (3)。

高斯消去法是提供一種流程使得一個線性聯立方程式所對應的增廣矩陣都能轉化為簡約列梯形式, 該流程主要分成兩個部分:

(A) 前向過程 (forward pass): 將增廣矩陣轉化成上三角矩陣, 並且各列的第一個非零項為 1; 此外, 它出現在前一列中第一非零項的右邊某一行。

- 前向過程的處理方式是先將矩陣以列的方式看待, 這些列當中找到第一行不是 0 的列 (如果第一行全部都是 0, 那就往後看下一行直到非零的行出現), 利用第 I 型基本列運算將這一列換到第一列, 用第 II 型基本列運算將第一列的第一個非零項為 1, 再用第 III 型基本列運算將下面的列在那一行全部轉化成 0。接著第一列固定, 從第二列到第 m 列尋找之後的某一行第一次出現非零項的列, 依剛才的流程 (互換、單位化、消去後面的列) 即可完成前向過程。此時由前向過程化簡後的矩陣符合簡約列梯形式的條件 (1) 及 (3)。

(B) 倒回過程 (backward pass): 將各列的第一個非零項成為所在行中唯一的非零項。

- 在前向過程中只能確定每一列第一次出現非零的行當中之後的列為 0, 而倒回過程的目的是要從最後一個不全為零的列出發, 利用第 III 型基本列運算得到之前的列的那一行為 0, 再看倒數第二個非零列、倒數第三個非零列... 依剛才的步驟讓它前面的列的那一行都變為 0。以符合簡約列梯形式的條件 (2)。

以下透過例子說明如何利用高斯消去法將增廣矩陣化成簡約列梯形式。

例 **6** (第 187 頁). 欲解線性聯立方程式

$$\begin{cases} 2x_1+3x_2+x_3+4x_4-9x_5=17 \\ x_1+x_2+x_3+x_4-3x_5=6 \\ x_1+x_2+x_3+2x_4-5x_5=8 \\ 2x_1+2x_2+2x_3+3x_4-8x_5=14, \end{cases}$$

考慮這個線性聯立方程式對應的增廣矩陣並進行高斯消去法:

$$\left[\begin{array}{ccccc|c} 2 & 3 & 1 & 4 & -9 & 17 \\ 1 & 1 & 1 & 1 & -3 & 6 \\ 1 & 1 & 1 & 2 & -5 & 8 \\ 2 & 2 & 2 & 3 & -8 & 14 \end{array}\right] \xrightarrow{R_{1,2}} \left[\begin{array}{ccccc|c} 1 & 1 & 1 & 1 & -3 & 6 \\ 2 & 3 & 1 & 4 & -9 & 17 \\ 1 & 1 & 1 & 2 & -5 & 8 \\ 2 & 2 & 2 & 3 & -8 & 14 \end{array}\right]$$

$$\xrightarrow{R_{1,2}^{(-2)},\ R_{1,3}^{(-1)},\ R_{1,4}^{(-2)}} \left[\begin{array}{ccccc|c} 1 & 1 & 1 & 1 & -3 & 6 \\ 0 & 1 & -1 & 2 & -3 & 5 \\ 0 & 0 & 0 & 1 & -2 & 2 \\ 0 & 0 & 0 & 1 & -2 & 2 \end{array}\right] \xrightarrow{R_{3,4}^{(-1)}} \left[\begin{array}{ccccc|c} 1 & 1 & 1 & 1 & -3 & 6 \\ 0 & 1 & -1 & 2 & -3 & 5 \\ 0 & 0 & 0 & 1 & -2 & 2 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right]$$

$$\xrightarrow{R_{3,1}^{(-1)},\ R_{3,2}^{(-2)}} \left[\begin{array}{ccccc|c} 1 & 1 & 1 & 0 & -1 & 4 \\ 0 & 1 & -1 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 & -2 & 2 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right] \xrightarrow{R_{2,1}^{(-1)}} \left[\begin{array}{ccccc|c} 1 & 0 & 2 & 0 & -2 & 3 \\ 0 & 1 & -1 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 & -2 & 2 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right],$$

再將矩陣改寫回線性聯立方程式的樣子:

$$\begin{cases} x_1+2x_3-2x_5=3 \\ x_2-x_3+x_5=1 \\ x_4-2x_5=2。 \end{cases}$$

再來, 我們將未知數 x_1,x_2,x_3,x_4,x_5 分成兩類, 第一類是最後的聯立方程式當中出現在最左邊的未知數, 以這個例子來說是 x_1,x_2,x_4, 第二類是其餘的未知數, 以上例而言是 x_3,x_5。現指定第二類的未知數為參數 (parameter), 即 $x_3=t_1, x_5=t_2$, 其中 $t_1,t_2\in\mathbb{R}$, 然後再將第一類的未知數完全用第二類的參數重新表達:

$$\begin{aligned} x_1 &= -2x_3+2x_5+3 = -2t_1+2t_2+3 \\ x_2 &= x_3-x_5+1 = t_1-t_2+1 \\ x_4 &= 2x_5+2 = 2t_2+2, \end{aligned}$$

於是聯立方程式的解為

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{bmatrix} = \begin{bmatrix} -2t_1+2t_2+3 \\ t_1-t_2+1 \\ t_1 \\ 2t_2+2 \\ t_2 \end{bmatrix} = \begin{bmatrix} 3 \\ 1 \\ 0 \\ 2 \\ 0 \end{bmatrix} + t_1 \begin{bmatrix} -2 \\ 1 \\ 1 \\ 0 \\ 0 \end{bmatrix} + t_2 \begin{bmatrix} 2 \\ -1 \\ 0 \\ 2 \\ 1 \end{bmatrix}。$$

若和單元 3.4 的 定理 **4** 對照, 則

$$\beta = \left\{ \begin{bmatrix} -2 \\ 1 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 2 \\ -1 \\ 0 \\ 2 \\ 1 \end{bmatrix} \right\}$$

是對應到的齊次線性聯立方程式的解空間之一組基底, 而

$$\mathbf{x}_p = \begin{bmatrix} 3 \\ 1 \\ 0 \\ 2 \\ 0 \end{bmatrix}$$

是原線性聯立方程式的一個特解。

在轉化增廣矩陣為簡約列梯形式的過程中, 我們不僅可以找到原線性聯立方程式的特解, 也可以求出對應的齊次線性聯立方程式的一組基底。一般來說, 高斯消去法也可以確定什麼時候線性聯立方程式無解; 也就是說, 如果得到一列使得唯一非零項是位在最後一行的時候, 則聯立方程式無解。

總結來說, 若要解有 n 個未知數的 m 條線性聯立方程式, 我們可用高斯消去法將增廣矩陣 $[A|\mathbf{b}]$ 轉化成簡約列梯形式 $[A'|\mathbf{b}']$,

- 若 $[A'|\mathbf{b}']$ 有一列使其唯一非零項在最後一行中, 則原聯立方程式無解。
- 若是帶有非零的列 (簡約列梯形式得知該列的第一個非零的數字是 1) 出現在倒數第二行或是再更之前的行, 則由 $[A'|\mathbf{b}']$ 刪去所有全為零的列, 並改寫回線性聯立方程式的樣子, 接著解這個聯立方程式, 將未知數分成兩類, 第一類是每一列的第一個未知數, 第二類是其它的未知數, 對於第二類的未知數, 逐一設定參數, 再將第一類的未知數透過線性聯立方程式逐一用第二類的參數重新表達, 這麼一來就可得到下列形式的任意解

$$\mathbf{x} = \mathbf{x}_p + t_1\mathbf{u}_1 + t_2\mathbf{u}_2 + \cdots + t_{n-r}\mathbf{u}_{n-r},$$

  其中 r 是 A' 中非零列的數目。上式將表示線性聯立方程式 $A\mathbf{x} = \mathbf{b}$ 所有的解, 它可以用 $n-r$ 個參數表示。以下定理將確實證明 $\beta = \{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_{n-r}\}$ 是對應的齊次線性聯立方程式的解空間之基底。

定理 7 (第 197 頁). 考慮 $A\mathbf{x} = \mathbf{b}$ 是含有 n 個未知數的 r 個非零的線性聯立方程式。假設 $\text{rank}(A) = \text{rank}([A|\mathbf{b}])$, 而且假設 $[A|\mathbf{b}]$ 是簡約列梯形式, 則

(A) $\text{rank}(A) = r$。

(B) 若由前面所述的方法所得的通解具有下列形式

$$\mathbf{x} = \mathbf{x}_p + t_1\mathbf{u}_1 + t_2\mathbf{u}_2 + \cdots + t_{n-r}\mathbf{u}_{n-r},$$

則 $\beta = \{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_{n-r}\}$ 是對應齊次線性聯立方程式的解空間之一組基底, 而且 $\mathbf{x}_p$ 是原非齊次線性聯立方程式的一個解。

證明: 因為 $[A|\mathbf{b}]$ 是簡約列梯形式, 則 $[A|\mathbf{b}]$ 必有 r 個非零列。由簡約列梯形式的定義可知這 r 列為線性獨立, 所以 $\text{rank}([A|\mathbf{b}]) = r$。因此 $\text{rank}(A) = r$。

令 K 表示 $A\mathbf{x} = \mathbf{b}$ 的解集合, 記 K_h 為 $A\mathbf{x} = \mathbf{0}$ 的解空間。假設 $t_1 = t_2 = \cdots = t_{n-r} = 0$, 則 $\mathbf{x} = \mathbf{x}_p \in K$, 因為 $K = \{\mathbf{x}_p\} + K_h$, 所以

$$K_h = \{-\mathbf{x}_p\} + K = \text{span}(\{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_{n-r}\})。$$

因為 $\text{rank}(A) = r$, 所以 $\dim(K_h) = \dim(N(L_A)) = n - r$。由於 $\dim(K_h) = n - r$ 且 K_h 是由 $\beta = \{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_{n-r}\}$ 生成, 所以 β 是 K_h 的一組基底。 □

以下繼續給出簡約列梯形式的結果。

定理 8. 假設 $A \in M_{m\times n}(\mathbb{F})$, 而 $\text{rank}(A) = r > 0$, 令 B 是 A 的簡約列梯形式, 則以下敘述成立:

(A) B 的非零列數為 r。

(B) 存在 B 的某 r 行, 記為 $\mathbf{b}_{j_i}$, 其中 $i = 1, 2, \ldots, r$ 使得 $\mathbf{b}_{j_i} = \mathbf{e}_i$。

(C) 分別被標為 $j_1, j_2, \ldots, j_r$ 的 A 之行是線性獨立的。

(D) 對於 $k = 1, 2, \ldots, n$, 假如 B 的第 k 行是 $d_1\mathbf{e}_1 + d_2\mathbf{e}_2 + \cdots + d_r\mathbf{e}_r$, 則 A 的第 k 行是 $d_1\mathbf{a}_{j_1} + d_2\mathbf{a}_{j_2} + \cdots + d_r\mathbf{a}_{j_r}$。

證明:

(A) 令 $A \in M_{m\times n}(\mathbb{F})$, 而 A 的每一行分別為 $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n$, 令 B 是 A 的簡約列梯形式。將 B 的所有行分別表示為 $\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_n$, 若 A 的秩為 r, 因為由 A 轉化為 B 的過程中都是透過基本列運算而完成, 所以矩陣 B 的秩亦為 r。因為 B 具簡約列梯形式, B 沒有一個非零列可為其它列的線性組合。因此 B 恰有 r 個非零列。

(B) 假如 $r \geq 1$, 向量 $\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_r$ 必出現在 B 的行裡。對於 $i = 1, 2, \ldots, r$, 令 j_i 表示 B 的行數滿足 $\mathbf{b}_{j_i} = \mathbf{e}_i$ 即為所求。

(C) 考慮方程式

$$c_1\mathbf{a}_{j_1} + c_2\mathbf{a}_{j_2} + \ldots + c_r\mathbf{a}_{j_r} = \mathbf{0},$$

其中 $c_1, c_2, \ldots, c_r \in \mathbb{F}$ 待定, 因為矩陣 B 是由矩陣 A 經過有限次基本列運算轉化而得, 存在一個可逆的 $m \times m$ 矩陣 $M = E_pE_{p-1}\cdots E_2E_1$, 其中 $E_1, E_2, \ldots, E_p$ 皆為基本矩陣, 使得 $MA = B$, 將上面方程式兩邊同乘 M 得到

$$c_1M\mathbf{a}_{j_1} + c_2M\mathbf{a}_{j_2} + \ldots + c_rM\mathbf{a}_{j_r} = \mathbf{0}。$$

因為 $M\mathbf{a}_{j_i} = \mathbf{b}_{j_i} = \mathbf{e}_i$, 所以 $c_1\mathbf{e}_1 + c_2\mathbf{e}_2 + \ldots + c_r\mathbf{e}_r = \mathbf{0}$, 因為 $\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n$ 線性獨立, 所以 $c_1 = c_2 = \cdots = c_r = 0$, 故 $\mathbf{a}_{j_1}, \mathbf{a}_{j_2}, \ldots, \mathbf{a}_{j_r}$ 線性獨立。

(D) 因為矩陣 B 只有 r 個非零列, 所以 B 的每一行都具有下面的形式

$$\begin{bmatrix} d_1 \\ d_2 \\ \vdots \\ d_r \\ 0 \\ \vdots \\ 0 \end{bmatrix},$$

其中 $d_1, d_2, \ldots, d_r \in \mathbb{F}$, 於是矩陣 A 相對應的行為

$$\begin{aligned} M^{-1}(d_1\mathbf{e}_1 + d_2\mathbf{e}_2 + \cdots + d_r\mathbf{e}_r) &= d_1M^{-1}\mathbf{e}_1 + d_2M^{-1}\mathbf{e}_2 + \cdots + d_rM^{-1}\mathbf{e}_r \\ &= d_1M^{-1}\mathbf{b}_{j_1} + d_2M^{-1}\mathbf{b}_{j_2} + \cdots + d_rM^{-1}\mathbf{b}_{j_r} \\ &= d_1\mathbf{a}_{j_1} + d_2\mathbf{a}_{j_2} + \cdots + d_r\mathbf{a}_{j_r}。 \end{aligned}$$

□

給定一個矩陣, 每個人採用的高斯消去法的方式可能不同, 那麼透過各種高斯消去法所得到的簡約列梯形式會一樣嗎? 數學上我們必須加以確定。

定理 9 (第 191 頁)**.** 一個矩陣的簡約列梯形式是唯一的。

證明: 給定矩陣 $A \in M_{m\times n}(\mathbb{F})$, 假設透過不同流程的高斯消去法得到簡約列梯形式 B 與 C。首先注意到 A, B, C 之間彼此都是透過基本列運算轉化而得; 換句話說, A 的每一列都可以用 B 中的所有列進行線性組合表示, 同樣地, A 的每一列都可以用 C 中的所有列進行線性組合表示, 所以 B 和 C 中的列也都可以互相以線性組合的方式表示。

假設 $B \neq C$, 首先, 從第一行開始逐一比較 B 與 C, 將第一次出現不同的行挑出來 (例如 $\mathbf{b}_j \neq \mathbf{c}_j$), 然後針對那一行由下往回找到第一個不為零的列 (假設是第 k 列), 分別在 $\mathbf{b}_j$ 與 $\mathbf{c}_j$ 這兩個行向

量的左邊逐一補上行向量 $\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_k$ 後分別得到矩陣 B' 與 C'。因為 B 和 C 之間可以用列運算進行轉化而得, 所以 B' 與 C' 之間也可以用列運算進行轉化而得。

現將 B' 與 C' 視為增廣矩陣 $[I_k|\mathbf{b}_j]$ 與 $[I_k|\mathbf{c}_j]$ 再對應回線性聯立方程式時, 它們分別得到存在唯一解 $\mathbf{b}'_j$ 與 $\mathbf{c}'_j$, 一方面, 因為 B' 與 C' 轉換成齊次線性聯立方程式來說是等價的, 另一方面, $\mathbf{b}'_j \neq \mathbf{c}'_j$, 這兩個結果是互相矛盾的, 故 $B = C$。 $\square$