

大學線性代數初步

李華介

國立台灣師範大學數學系

前言

本講義主要目的是針對數學系大一學生介紹有關線性代數基本的理論，大致上僅談論實係數的向量空間。

一般經驗上大一學生會覺得線性代數學習起來會比微積分吃力，主要是因素可能是在高中時期學習有關線性代數部分主要著重於操作而較少論證。這一點在大學線性代數中就較不同了。不過經驗上也告訴我們當學生到三大四時（因準備研究所考試）再重新溫習線性代數時，會覺得它不再那麼難以親近。由此可知大學時期的線性代數其理論並不難懂，較大的障礙是要開始學習數學的論證。這個障礙或許到三大四一些數學思維較成熟時稍可解除，不過對一些同學來說可能為時已晚，畢竟線性代數的理論或概念與其他課程都脫不了關係。基於這個原因寫下這份講義，希望藉由較平易近人的方式介紹線性代數也慢慢引導熟悉數學的論證方式。本講義希望以淺顯易懂為主旨，而不是生動有趣。畢竟有些事情要說明清楚就會顯得囉唆，當然就不有趣了。

研讀本講義的同學要有心理準備，本講義是針對數學系學生而寫，自然偏重於整個線性代數的理論架構。對於線性代數在其他領域的應用幾乎沒有著墨。我們依循一貫的原則就是理論清楚了，接下來的應用或推廣就不難了。所以對應用有興趣的同學應再參考其他的參考書籍。另外本講義並未提供習題，不過在某些概念講述之後有時會提供一些問題（Question）。這些問題幾乎是檢視觀念是否正確或是對內容是否了解，大部分問題若觀念已清楚應可以立即回答。所以這些問題的份量仍不及一般習題，對熟習線性代數所給予的訓練。針對這一點，請欲學習好線性代數的同學務必參閱一般線性代數書籍的習題，自行磨練。

本講義雖然主要以中文撰寫，不過當涉及定義或專有名詞時，為免翻譯的困擾將以英文取代。因此將以中英夾雜較不傳統的方式顯現，若有不便請見諒。

本講義編寫費時，編寫完後並未經過嚴謹的校對。疏漏在所難免，雖不至於有理論性上嚴重的錯誤，但讀者仍應注意不宜概括全收。若發現錯誤，歡迎提出寶貴的意見。

Matrix

在上一章我們僅利用矩陣來表示一個聯立方程組，這種表示法不只有其方便性其實是有另一層的意義。在這一章中我們將介紹有關矩陣的運算，利用矩陣的運算我們對聯立方程組將有另一種看法。利用這新的看法，我們對聯立方程組的解可以有更進一步的了解。

3.1. 矩陣的運算

在本節中我們將簡單地回顧有關於矩陣的定義。一般來說一個矩陣是由數個（橫）列（row）以及（直）行（column）的數組成。若一矩陣由 m 個 row 和 n 個 column 的數所組成，我們便稱該矩陣為一個 $m \times n$ matrix。特別的，一個 $n \times n$ matrix（即 row 的個數等於 column 的個數），我們稱之為 *square matrix*。在本講義中，我們用 $\mathcal{M}_{m \times n}$ 來表示所有 $m \times n$ 的矩陣所成的集合。通常我們會用大寫的英文字母來表示一個矩陣。例如若

$$A = \begin{bmatrix} 1 & 0 & 2 & 3 \\ 0 & 1 & 5 & 8 \\ 2 & 1 & 1 & 0 \end{bmatrix}, \quad (3.1)$$

則 A 為一個 3×4 matrix，即 $A \in \mathcal{M}_{3 \times 4}$ 。當我們要抽象地描述一個矩陣時，我們也常用 $A = [a_{ij}]$ 這樣的方法來描述。這種表示法意指 A 中在第 i 個 row 和 j 個 column 的位置我們用 a_{ij} 來表示，並稱之為此矩陣的 (i, j) -th entry。因此當我們說 $A = [a_{ij}]$ 為 $m \times n$ 矩陣，這表示 $1 \leq i \leq m$ 且 $1 \leq j \leq n$ 。例如對於式子 (3.1) 中的矩陣 A ，若 $A = [a_{ij}]$ ，則

$$\begin{aligned} a_{11} &= 1, & a_{12} &= 0, & a_{13} &= 2, & a_{14} &= 3, \\ a_{21} &= 0, & a_{22} &= 1, & a_{23} &= 5, & a_{24} &= 8, \\ a_{31} &= 2, & a_{32} &= 1, & a_{33} &= 1, & a_{34} &= 0. \end{aligned}$$

另外為了方便起見，我們也會將矩陣 A 的每一個 row 和 column 用向量的方法來表示，這些稱為 A 的 row vectors 和 column vectors。在本講義我們會將矩陣 $A = [a_{ij}]$ 第 i 個 row 所成的 row vector 用 ${}_i\mathbf{a}$ 來表示，而第 j 個 column 所成的 column vector 用 \mathbf{a}_j 來表示。例如對於式子 (3.1) 中的矩陣 A ，我們有

$${}_1\mathbf{a} = [1 \ 0 \ 2 \ 3], \quad {}_2\mathbf{a} = [0 \ 1 \ 5 \ 8], \quad {}_3\mathbf{a} = [2 \ 1 \ 1 \ 0]$$

以及

$$\mathbf{a}_1 = \begin{bmatrix} 1 \\ 0 \\ 2 \end{bmatrix}, \quad \mathbf{a}_2 = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}, \quad \mathbf{a}_3 = \begin{bmatrix} 2 \\ 5 \\ 1 \end{bmatrix}, \quad \mathbf{a}_4 = \begin{bmatrix} 3 \\ 8 \\ 0 \end{bmatrix}.$$

注意由於我們也想將向量看成是一個矩陣，這裡的 row vector 和 column vector 們都用矩陣的形式呈現。

我們想給矩陣一個運算，既然要談運算就會牽涉相等的概念。所以我們要先定義何謂矩陣的相等（就如同我們曾定義 \mathbb{R}^n 中向量的相等）。

Definition 3.1.1. 假設 $A = [a_{ij}]$ 為一個 $m \times n$ matrix 且 $A' = [a'_{ij}]$ 為一個 $m' \times n'$ matrix. 我們定義 $A = A'$ 若且唯若 $m = m', n = n'$ 且對所有的 $1 \leq i \leq m$ 以及 $1 \leq j \leq n$ 皆有 $a_{ij} = a'_{ij}$.

很容易看出矩陣的相等的定義是向量相等的延伸。在向量中只有同在 \mathbb{R}^n 的向量我們才談是否相等，且兩個 \mathbb{R}^n 中的向量相等表示這兩個向量在每一個相同位置的數皆相等。同樣的只有同在 $\mathcal{M}_{m \times n}$ 的矩陣才談是否相等，且兩個 $\mathcal{M}_{m \times n}$ 中的矩陣相等表示這兩個矩陣在每一個相同位置的數皆相等。

我們也延伸向量加法與係數積的定義來定義矩陣的加法與係數積。也就是說只有同為 $\mathcal{M}_{m \times n}$ 的矩陣我們才定義它們之間的加法，且兩矩陣相加表示將這兩個矩陣在相同位置的數加起來。而一個實數乘上一個矩陣即為將該矩陣每一個位置上的數乘上該實數。具體來說我們有以下的定義。

Definition 3.1.2. 假設 $A = [a_{ij}], B = [b_{ij}]$ 皆為 $m \times n$ matrix. 定義 $A + B = [c_{ij}]$, 其中對所有的 $1 \leq i \leq m$ 以及 $1 \leq j \leq n$ 皆有 $c_{ij} = a_{ij} + b_{ij}$. 對任意實數 r , 我們定義 $rA = [d_{ij}]$ 其中對所有的 $1 \leq i \leq m$ 以及 $1 \leq j \leq n$ 皆有 $d_{ij} = ra_{ij}$.

Definition 3.1.2 告訴我們若

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}, \quad B = \begin{bmatrix} b_{11} & b_{12} & \cdots & b_{1n} \\ b_{21} & b_{22} & \cdots & b_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ b_{m1} & b_{m2} & \cdots & b_{mn} \end{bmatrix}$$

則

$$A + B = \begin{bmatrix} a_{11} + b_{11} & a_{12} + b_{12} & \cdots & a_{1n} + b_{1n} \\ a_{21} + b_{21} & a_{22} + b_{22} & \cdots & a_{2n} + b_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} + b_{m1} & a_{m2} + b_{m2} & \cdots & a_{mn} + b_{mn} \end{bmatrix}$$

且

$$rA = \begin{bmatrix} ra_{11} & ra_{12} & \cdots & ra_{1n} \\ ra_{21} & ra_{22} & \cdots & ra_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ ra_{m1} & ra_{m2} & \cdots & ra_{mn} \end{bmatrix}$$

既然矩陣的加法與係數積的定義是由向量的加法與係數積延伸而來，我們可以預期矩陣的加法與係數積應有和向量的加法與係數積相同的性質。事實上我們確有以下性質，這些性質的證明和向量的情形相同（用到實數相對應的性質），我們就不再重複了。

Proposition 3.1.3. 對於 $\mathcal{M}_{m \times n}$ 上的矩陣，我們有以下的性質：

- (1) 對任意 $A, B \in \mathcal{M}_{m \times n}$ ，皆有 $A + B = B + A$.
- (2) 對任意 $A, B, C \in \mathcal{M}_{m \times n}$ ，皆有 $(A + B) + C = A + (B + C)$.
- (3) 存在一矩陣 $O \in \mathcal{M}_{m \times n}$ 滿足對任意 $A \in \mathcal{M}_{m \times n}$ 皆有 $O + A = A$.
- (4) 對任意 $A \in \mathcal{M}_{m \times n}$ 皆可找到 $A' \in \mathcal{M}_{m \times n}$ 滿足 $A + A' = O$.
- (5) 對任意 $r, s \in \mathbb{R}$ 以及 $A \in \mathcal{M}_{m \times n}$ ，皆有 $r(sA) = (rs)A$.
- (6) 對任意 $r, s \in \mathbb{R}$ 以及 $A \in \mathcal{M}_{m \times n}$ ，皆有 $(r + s)A = rA + sA$.
- (7) 對任意 $r \in \mathbb{R}$ 以及 $A, B \in \mathcal{M}_{m \times n}$ 皆有 $r(A + B) = rA + rB$.
- (8) 對任意 $A \in \mathcal{M}_{m \times n}$ ，皆有 $1A = A$.

接著我們定義矩陣間的乘法。首先回顧在 Chapter 1 我們曾說明 $(1, 2, 3) \in \mathbb{R}^3$ 是否屬於 $\text{Span}((1, -1, 2), (2, 1, -2))$ 的問題，等同於是否存在 $x_1, x_2 \in \mathbb{R}$ 滿足 $(1, 2, 3) = x_1(1, -1, 2) + x_2(2, 1, -2)$ 。若將它們寫成 column vector 的形式，即

$$x_1 \begin{bmatrix} 1 \\ -1 \\ 2 \end{bmatrix} + x_2 \begin{bmatrix} 2 \\ 1 \\ -2 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix},$$

可看出此又等同於聯立方程組

$$\begin{aligned} x_1 + 2x_2 &= 1 \\ -x_1 + x_2 &= 2 \\ 2x_1 - 2x_2 &= 3 \end{aligned}$$

是否有解的問題。而這樣的聯立方程組，我們又把它寫成

$$\begin{bmatrix} 1 & 2 \\ -1 & 1 \\ 2 & -2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}.$$

所以若我們定義矩陣 $\begin{bmatrix} 1 & 2 \\ -1 & 1 \\ 2 & -2 \end{bmatrix}$ 和 $\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$ 的乘法為

$$\begin{bmatrix} 1 & 2 \\ -1 & 1 \\ 2 & -2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = x_1 \begin{bmatrix} 1 \\ -1 \\ 2 \end{bmatrix} + x_2 \begin{bmatrix} 2 \\ 1 \\ -2 \end{bmatrix}$$

那麼聯立方程組與矩陣的關係就不只是為了列式方便而已，聯立方程組和矩陣的運算產生了緊密的關係。

從這個角度出發，我們有以下定義。

Definition 3.1.4. 設 $A = [a_{ij}]$ 為 $m \times n$ matrix 以及 $\mathbf{b} = [b_j]$ 為 $n \times 1$ matrix (即 \mathbb{R}^n 中的 column vector)。若 \mathbf{a}_i 表示 A 的 i -th column 則定義

$$\mathbf{Ab} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix} = b_1 \mathbf{a}_1 + b_2 \mathbf{a}_2 + \cdots + b_n \mathbf{a}_n.$$

注意依此定義, 必需 $A \in \mathcal{M}_{m \times n}$ 的 column 的個數 n 等於 $\mathbf{b} \in \mathcal{M}_{n \times 1}$ 的 row 的個數 n , 才能定義 \mathbf{Ab} 且此時 \mathbf{Ab} 會是 $m \times 1$ matrix (即 \mathbb{R}^m 中的 column vector). 觀察此 column vector, 我們有

$$\mathbf{Ab} = b_1 \begin{bmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{bmatrix} + b_2 \begin{bmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{m2} \end{bmatrix} + \cdots + b_n \begin{bmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{mn} \end{bmatrix} = \begin{bmatrix} b_1 a_{11} + b_2 a_{12} + \cdots + b_n a_{1n} \\ b_1 a_{21} + b_2 a_{22} + \cdots + b_n a_{2n} \\ \vdots \\ b_1 a_{m1} + b_2 a_{m2} + \cdots + b_n a_{mn} \end{bmatrix}. \quad (3.2)$$

特別的, 當 $\mathbf{a} = [a_1 \ a_2 \ \cdots \ a_n]$ 為 $1 \times n$ matrix 而 $\mathbf{b} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix}$ 為 $n \times 1$ matrix, 依 Definition

3.1.4 的矩陣乘法定義

$$\mathbf{ab} = [a_1 \ a_2 \ \cdots \ a_n] \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix} = b_1 a_1 + b_2 a_2 + \cdots + b_n a_n. \quad (3.3)$$

也就是說此時若將 \mathbf{a}, \mathbf{b} 看成 \mathbb{R}^n 中的 vector, 則 \mathbf{ab} 就是 \mathbf{a}, \mathbf{b} 的內積. 依此看法, 由式子 (3.2), 我們可將 \mathbf{Ab} 寫成

$$\mathbf{Ab} = \begin{bmatrix} {}_1\mathbf{ab} \\ {}_2\mathbf{ab} \\ \vdots \\ {}_m\mathbf{ab} \end{bmatrix}. \quad (3.4)$$

也就是說 \mathbf{Ab} 這一個 $m \times 1$ matrix 的 i -th entry 為 ${}_i\mathbf{ab}$ 也就是 A 的 i -th row ${}_i\mathbf{a}$ 和 \mathbf{b} 的內積.

Remark 3.1.5. 注意式子 (3.3) 是說當 \mathbf{a}, \mathbf{b} 為 \mathbb{R}^n 上的向量, 若將 \mathbf{a} 寫成 row vector 的形式, \mathbf{b} 寫成 column vector 的形式, 那麼我們可以將 \mathbf{a}, \mathbf{b} 看成矩陣, 即 $\mathbf{a} \in \mathcal{M}_{1 \times n}$, $\mathbf{b} \in \mathcal{M}_{n \times 1}$. 此時由矩陣相乘的定義 \mathbf{ab} 就是一個 1×1 的矩陣 (即實數) 且其值就是 \mathbf{a}, \mathbf{b} 視為 \mathbb{R}^n 上的向量後取內積. 也就是說, 當看成向量的內積時, 我們不管兩向量是寫成 column vector 或 row vector. 但若看成矩陣的乘法時, 我們就必需說清楚哪一個是 column vector 哪一個是 row vector 了. 絕不能說由於向量內積是可交換的, 便認為 $\mathbf{b} \in \mathcal{M}_{n \times 1}$, $\mathbf{a} \in \mathcal{M}_{1 \times n}$ 看成矩陣相乘所得的 \mathbf{ba} 也會等於 \mathbf{a} 和 \mathbf{b} 看成 \mathbb{R}^n 上的向量後取內積. 下面我們定義更一般矩陣的乘法時, 便會知道看成矩陣的乘法時 \mathbf{ab} 和 \mathbf{ba} 是不同的 (\mathbf{ba} 會是一個 $n \times n$ matrix).

現在我們將矩陣乘法推廣到更一般的情況, 當 $A = [a_{ij}]$ 是一個 $m \times n$ matrix, $B = [b_{jk}]$ 是一個 $n \times l$ matrix. 由於對 B 的每一個 column vector $\mathbf{b}_k \in \mathcal{M}_{n \times 1}$, $1 \leq k \leq l$, 我們已定義了 \mathbf{Ab}_k 為何, 現在我們定義 AB 為 $m \times l$ matrix, 其中 AB 的 k -th column vector 為 \mathbf{Ab}_k . 我們大致上有以下的圖示.

$$A \begin{bmatrix} | & | & \cdots & | \\ \mathbf{b}_1 & \mathbf{b}_2 & \cdots & \mathbf{b}_l \\ | & | & \cdots & | \end{bmatrix} = \begin{bmatrix} | & | & \cdots & | \\ \mathbf{Ab}_1 & \mathbf{Ab}_2 & \cdots & \mathbf{Ab}_l \\ | & | & \cdots & | \end{bmatrix}$$

由於 \mathbf{Ab}_k 為 $m \times 1$ matrix, 依此定義確實 AB 為 $m \times l$ matrix. 現在我們來看正式的定義.

Definition 3.1.6. 設 $A = [a_{ij}]$ 為 $m \times n$ matrix 以及 $B = [b_{jk}]$ 為 $n \times l$ matrix, 則定義 $AB = C = [c_{ik}]$ 為 $m \times l$ matrix, 其中對於 $1 \leq k \leq l$, C 的 k -th column \mathbf{c}_k 為

$$\mathbf{c}_k = \mathbf{A}\mathbf{b}_k = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \begin{bmatrix} b_{1k} \\ b_{2k} \\ \vdots \\ b_{nk} \end{bmatrix} = b_{1k}\mathbf{a}_1 + b_{2k}\mathbf{a}_2 + \cdots + b_{nk}\mathbf{a}_n. \quad (3.5)$$

由此定義, 我們知對於 $1 \leq i \leq m$, $1 \leq k \leq l$, AB 的 (i, k) -th entry 應為其 k -th column (即 $\mathbf{A}\mathbf{b}_k$) 從上往下算的第 i 個 entry. 由式子 (3.4) 我們知此即 A 的 i -th row \mathbf{a}_i 和 B 的 k -th column \mathbf{b}_k 看成 \mathbb{R}^n 上的向量後取內積. 換言之, 若 $AB = [c_{ik}]$, 則 AB 的 (i, k) -th entry c_{ik} 為

$$c_{ik} = \mathbf{a}_i \mathbf{b}_k = a_{i1}b_{1k} + a_{i2}b_{2k} + \cdots + a_{in}b_{nk} = \sum_{j=1}^n a_{ij}b_{jk}. \quad (3.6)$$

再次強調一次, 並不是任取兩個矩陣都可以定義乘法, 必須是左邊矩陣的 column 個數和右邊矩陣的 row 個數相同才能相乘.

Example 3.1.7. 令

$$A = \begin{bmatrix} -2 & 4 \\ 3 & 6 \\ 2 & 2 \end{bmatrix}, B = \begin{bmatrix} 4 & -1 & 2 & 5 \\ 3 & 0 & 1 & 1 \end{bmatrix}$$

考慮矩陣乘法 AB . 依定義矩陣 AB 的 3-rd column 為

$$\mathbf{A}\mathbf{b}_3 = \begin{bmatrix} -2 & 4 \\ 3 & 6 \\ 2 & 2 \end{bmatrix} \begin{bmatrix} 2 \\ 1 \end{bmatrix} = 2\mathbf{a}_1 + 1\mathbf{a}_2 = 2 \begin{bmatrix} -2 \\ 3 \\ 2 \end{bmatrix} + 1 \begin{bmatrix} 4 \\ 6 \\ 2 \end{bmatrix} = \begin{bmatrix} 0 \\ 12 \\ 6 \end{bmatrix}.$$

所以 AB 的 $(2, 3)$ entry 為 12 等於 A 的 2-nd row 和 B 的 3-rd column 看成 \mathbb{R}^2 中的向量所得的內積, 即 $(3, 6) \cdot (2, 1) = 12$. 事實上我們有

$$AB = \begin{bmatrix} -2 & 4 \\ 3 & 6 \\ 2 & 2 \end{bmatrix} \begin{bmatrix} 4 & -1 & 2 & 5 \\ 3 & 0 & 1 & 1 \end{bmatrix} = \begin{bmatrix} 4 & 2 & 0 & -6 \\ 30 & -3 & 12 & 21 \\ 14 & -2 & 6 & 12 \end{bmatrix}.$$

Question 3.1. 設 \mathbf{a}, \mathbf{b} 為 \mathbb{R}^n 上的向量, 若將 \mathbf{a} 寫成 row vector 的形式, \mathbf{b} 寫成 column vector 的形式, 且將 \mathbf{a}, \mathbf{b} 看成矩陣, 即 $\mathbf{a} \in \mathcal{M}_{1 \times n}$, $\mathbf{b} \in \mathcal{M}_{n \times 1}$. 試問依矩陣乘法定義 $\mathbf{b}\mathbf{a}$ 應為何種矩陣? 它和 \mathbf{a}, \mathbf{b} 看成 \mathbb{R}^n 上的向量後取內積有關嗎?

大部分的書都會用式子 (3.6) 當成矩陣乘法的定義. 我們選用式子 (3.5) 的用意, 主要是它較能描繪當初矩陣乘法定義的用意. 另外它是由 column 來描繪矩陣的乘法, 在證明或推導有關矩陣乘法性質時, 有時比式子 (3.6) 利用 entry 來看方便多了. 例如我們有以下的性質.

Proposition 3.1.8. 假設 $A, A' \in \mathcal{M}_{m \times n}$, $B, B' \in \mathcal{M}_{n \times l}$. 我們有以下的性質.

- (1) $A(B + B') = AB + AB'$.
- (2) $(A + A')B = AB + A'B$.

Proof. 首先注意因 $A+A'$ 仍為 $m \times n$ 矩陣且 $B+B'$ 為 $n \times l$ 矩陣, 所以這些矩陣的階數是符合矩陣乘法的規定. 我們假設 $A = [a_{ij}]$, $A' = [a'_{ij}]$, $B = [b_{jk}]$, $B' = [b'_{jk}]$, 其中 $1 \leq i \leq m$, $1 \leq j \leq n$ 以及 $1 \leq k \leq l$.

(1) 我們證明當 $1 \leq k \leq l$ 時, $A(B+B')$ 的 k -th column 會等於 $AB+AB'$ 的 k -th column. 依定義 $A(B+B')$ 的 k -th column 為 A 的右邊乘上 $B+B'$ 的 k -th column. 然而由矩陣加法定義, $B+B'$ 的 k -th column 為 $\mathbf{b}_k + \mathbf{b}'_k$, 即 B 的 k -th column 加上 B' 的 k -th column. 因此我們有 $A(B+B')$ 的 k -th column 為

$$A(\mathbf{b}_k + \mathbf{b}'_k) = A \begin{bmatrix} b_{1k} + b'_{1k} \\ b_{2k} + b'_{2k} \\ \vdots \\ b_{nk} + b'_{nk} \end{bmatrix} = (b_{1k} + b'_{1k})\mathbf{a}_1 + (b_{2k} + b'_{2k})\mathbf{a}_2 + \cdots + (b_{nk} + b'_{nk})\mathbf{a}_n. \quad (3.7)$$

另一方面, $AB+AB'$ 的 k -th column 為 AB 的 k -th column 加上 AB' 的 k -th column, 因此 $AB+AB'$ 的 k -th column 為

$$A \begin{bmatrix} b_{1k} \\ b_{2k} \\ \vdots \\ b_{nk} \end{bmatrix} + A \begin{bmatrix} b'_{1k} \\ b'_{2k} \\ \vdots \\ b'_{nk} \end{bmatrix} = (b_{1k}\mathbf{a}_1 + b_{2k}\mathbf{a}_2 + \cdots + b_{nk}\mathbf{a}_n) + (b'_{1k}\mathbf{a}_1 + b'_{2k}\mathbf{a}_2 + \cdots + b'_{nk}\mathbf{a}_n). \quad (3.8)$$

由 Proposition 1.2.3 向量加法相關性質, 我們得證式子 (3.7) 和式子 (3.8) 相等.

(2) 我們證明當 $1 \leq k \leq l$ 時, $(A+A')B$ 的 k -th column 會等於 $AB+A'B$ 的 k -th column. 依定義 $(A+A')B$ 的 k -th column 為 $A+A'$ 的右邊乘上 B 的 k -th column. 然而由矩陣加法定義, 當 $1 \leq j \leq n$ 時, $A+A'$ 的 j -th column 為 $\mathbf{a}_j + \mathbf{a}'_j$, 即 A 的 j -th column 加上 A' 的 j -th column. 因此我們有 $(A+A')B$ 的 k -th column 為

$$(A+A')\mathbf{b}_k = (A+A') \begin{bmatrix} b_{1k} \\ b_{2k} \\ \vdots \\ b_{nk} \end{bmatrix} = b_{1k}(\mathbf{a}_1 + \mathbf{a}'_1) + b_{2k}(\mathbf{a}_2 + \mathbf{a}'_2) + \cdots + b_{nk}(\mathbf{a}_n + \mathbf{a}'_n). \quad (3.9)$$

另一方面, $AB+A'B$ 的 k -th column 為 AB 的 k -th column 加上 $A'B$ 的 k -th column, 因此 $AB+A'B$ 的 k -th column 為

$$A \begin{bmatrix} b_{1k} \\ b_{2k} \\ \vdots \\ b_{nk} \end{bmatrix} + A' \begin{bmatrix} b_{1k} \\ b_{2k} \\ \vdots \\ b_{nk} \end{bmatrix} = (b_{1k}\mathbf{a}_1 + b_{2k}\mathbf{a}_2 + \cdots + b_{nk}\mathbf{a}_n) + (b_{1k}\mathbf{a}'_1 + b_{2k}\mathbf{a}'_2 + \cdots + b_{nk}\mathbf{a}'_n). \quad (3.10)$$

再次由 Proposition 1.2.3 向量加法相關性質, 我們得證式子 (3.9) 和式子 (3.10) 相等. \square

矩陣乘法和 scalar multiplication (係數積) 也有以下關係

Proposition 3.1.9. 設 $r \in \mathbb{R}$, $A \in \mathcal{M}_{m \times n}$, $B \in \mathcal{M}_{n \times l}$. 則

$$r(AB) = (rA)B = A(rB).$$

Proof. 假設 $A = [a_{ij}]$, $B = [b_{jk}]$, 其中 $1 \leq i \leq m$, $1 \leq j \leq n$ 以及 $1 \leq k \leq l$. $r(AB)$ 的 k -th column 為 r 乘上 AB 的 k -th column, 故為

$$r(b_{1k}\mathbf{a}_1 + b_{2k}\mathbf{a}_2 + \cdots + b_{nk}\mathbf{a}_n). \quad (3.11)$$

而當 $1 \leq j \leq n$ 時, (rA) 的 j -th column 為 $r\mathbf{a}_j$. 故 $(rA)B$ 的 k -th column 為

$$b_{1k}(r\mathbf{a}_1) + b_{2k}(r\mathbf{a}_2) + \cdots + b_{nk}(r\mathbf{a}_n). \quad (3.12)$$

最後由於 rB 的 k -th column 為

$$\begin{bmatrix} rb_{1k} \\ rb_{2k} \\ \vdots \\ rb_{nk} \end{bmatrix},$$

故 $A(rB)$ 的 k -th column 為

$$(rb_{1k})\mathbf{a}_1 + (rb_{2k})\mathbf{a}_2 + \cdots + (rb_{nk})\mathbf{a}_n. \quad (3.13)$$

再次由 Proposition 1.2.3 向量相關性質, 我們得證 (3.11), (3.12), (3.13) 三個式子皆相等. \square

由 Proposition 3.1.8 和 Proposition 3.1.9 的證明我們可以看出, 有些矩陣乘法性質的推導可以簡化成右邊的矩陣是一個 column 的情形處理. 其實利用 row 來看矩陣的乘法也很很有用, 不過這個留待下一節介紹矩陣的 transpose (轉置) 後會更清楚.

利用矩陣乘法定義, 也可直接推得乘法具有結合律的性質 (即 $(AB)C = A(BC)$). 不過由於直接推導有點複雜, 這裡請大家自行驗證. 不過下一節我們學習了用 row 來處理矩陣乘法時, 我們會給一個較簡明的證明. 最後我們要強調的是矩陣乘法雖具有許多和實數乘法類似的性質, 但它卻沒有交換律. 事實上有可能 A 乘以 B 有定義, 但 B 卻不能乘以 A , 例如 $A \in \mathcal{M}_{2 \times 3}$, $B \in \mathcal{M}_{3 \times 4}$ 的情形. 也有可能即使 A 乘以 B 和 B 乘以 A 都有定義, 但由於乘了以後階數不同, 仍會使得 $AB \neq BA$, 例如 $A \in \mathcal{M}_{2 \times 3}$, $B \in \mathcal{M}_{3 \times 2}$ 的情形. 僅有在 A, B 為同階方陣時, 才有可能使得 AB 和 BA 的階數相同. 但此時仍有可能 $AB \neq BA$, 例如

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}, B = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}, AB = \begin{bmatrix} a & -b \\ c & -d \end{bmatrix}, BA = \begin{bmatrix} a & b \\ -c & -d \end{bmatrix},$$

這種情形只有在 $b = c = 0$ 時, 才會使得 $AB = BA$. 所以在處理矩陣乘法時要特別小心. 例如當 A, B 為同階方陣時由 Proposition 3.1.8 和 Proposition 3.1.9 可推得 $(A - B)(A + B) = A^2 - AB + BA - B^2$, 但由於可能 $AB \neq BA$, 我們不見得會有 $(A - B)(A + B) = A^2 - B^2$.

當然了, 仍然有許多方陣會和所有的同階方陣相乘是可交換的. 一個常見的就是 zero matrix (零矩陣) O (即 $O = [a_{i,j}]$ 滿足每一個 entry $a_{i,j} = 0$). 很容易驗證若 O 是一個 $n \times n$ square matrix, 則對任意 $A \in \mathcal{M}_{n \times n}$, 皆有 $OA = AO = O$. 另一個常見的便是所謂的 identity matrix. 通常 $n \times n$ 階的 identity matrix, 我們會用 I_n 來表示. I_n 的 i -th column 為 \mathbf{e}_i , 其中

$\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n$ 就是我們曾提過 \mathbb{R}^n 的 standard basis (寫成 column vector). 例如

$$I_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, I_4 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

利用矩陣乘法的定義, 很容易知道對任意 $A \in \mathcal{M}_{m \times n}, B \in \mathcal{M}_{n \times l}$ 我們皆有 $AI_n = A, I_n B = B$. 特別的, 當 A 為 $n \times n$ matrix, 我們有 $AI_n = I_n A = A$.

Question 3.2. 假設 $A \in \mathcal{M}_{n \times n}$, 是否 $(A - 2I_n)^2 = A^2 - 4A + 4I_n$ 為對?

Question 3.3. 試證明 I_n 是唯一的 $n \times n$ 滿足對任意 $A \in \mathcal{M}_{m \times n}$ 皆滿足 $AI_n = A$.

一個 $n \times n$ 的 square matrix 其 (i, i) -th entry 稱為 diagonal entry. 若除了 diagonal entries 以外, 其他的 entry 皆為 0, 我們便稱之為 *diagonal matrix*. Identity matrix 就是一個 diagonal matrix. 因為它的 diagonal entry 皆為 1, 其他的 entry 皆為 0. 另外, 對於任意 $r \in \mathbb{R}$, rI_n 亦為 diagonal matrix. 因為它 diagonal entry 皆為 r , 其他 entry 皆為 0. 對於任意 $A \in \mathcal{M}_{m \times n}, B \in \mathcal{M}_{n \times l}$ 我們很容易驗證 $rA = A(rI_n), rB = (rI_n)B$.

Question 3.4. 試利用 *Proposition 3.1.9* 驗證對任意 $n \times n$ square matrix A , 皆有 $(rI_n)A = A(rI_n)$.

要注意, 並不是所有 $n \times n$ 的 diagonal matrix 都會和 $n \times n$ 的 square matrix 相乘可交換. 前面曾給過例子 $\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$ 就不能和所有的 2×2 相乘可交換.

3.2. Transpose Operation

這一節中我們將介紹矩陣取 transpose (即轉置矩陣) 的概念, 並利用它來探討如何從 row 的角度來看矩陣相乘. 最後結合 column 和 row 的關係來處理矩陣相乘的結合律.

Definition 3.2.1. 對任意 $m \times n$ matrix $A = [a_{ij}]$. 我們定義 A 的 *transpose* 為一個 $n \times m$, 記為 A^T , 滿足對任意 $1 \leq i \leq n, 1 \leq j \leq m$, A^T 的 (i, j) -th entry 為 A 的 (j, i) -th entry.

首先我們看看此定義是否為 well-defined, 也就是說給定一個 $m \times n$ matrix A , 我們是否真的可以得到一個 $n \times m$ matrix A^T . 依定義對任意 $1 \leq i \leq n, 1 \leq j \leq m$, 在 A^T 的 i -th row 和 j -th column 位置的數應為 A 的 j -th row 和 i -th column 位置的數. 因為 A 的 row 和 column 的個數分別為 m 和 n , 而 j, i 又分別滿足 $1 \leq j \leq m, 1 \leq i \leq n$, 所以我們真的可以在 A 中找到此 entry 放到 A^T 的 i -th row 和 j -th column 的位置. 可知此定義是沒問題的. 我們看以下的例子.

Example 3.2.2. 令

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

依定義 A^T 應為 3×2 matrix. 設其為

$$A^T = \begin{bmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \\ x_{31} & x_{32} \end{bmatrix}.$$

依定義 A^T 的 (1,1)-th entry x_{11} 應為 A 的 (1,1)-th entry, 即 1. A^T 的 (2,1)-th entry x_{21} 應為 A 的 (1,2)-th entry, 即 2, 再來 A^T 的 (3,1)-th entry x_{31} 應為 A 的 (1,3)-th entry, 即 3. 同理可得 $x_{12} = -1$, $x_{22} = -2$, $x_{32} = -3$, 故知

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

可看出 A^T 的 1-st column 就是將 A 的 1-st row 寫成 column 而得, 且 A^T 的 2-nd column 就是將 A 的 2-nd row 寫成 column 而得. 而 A^T 的 1-st, 2-nd 和 3-rd row 就是將 A 的 1-st, 2-nd 和 3-rd column 寫成 row 而得.

由於我們希望整體的看一個矩陣的 row 和 column, 所以我們需要探討 A 和 A^T 它們的 row 與 column 之間的關係.

Lemma 3.2.3. 假設 A 為 $m \times n$ matrix. 對於 $1 \leq i \leq n$, A^T 的 i -th row 就是將 A 的 i -th column 寫成 row vector. 對於 $1 \leq j \leq m$, A^T 的 j -th column 就是將 A 的 j -th row 寫成 column vector.

Proof. 依定義 A^T 為 $n \times m$ matrix. 設 $A = [a_{kl}]$, $A^T = [a'_{ij}]$. 對於 $1 \leq i \leq n$, A^T 的 i -th row 為

$$[a'_{i1} \quad a'_{i2} \quad \cdots \quad a'_{im}].$$

依定義 $a'_{i1} = a_{1i}$, $a'_{i2} = a_{2i}$, \dots , $a'_{im} = a_{mi}$, 所以 A^T 的 i -th row 為

$$[a_{1i} \quad a_{2i} \quad \cdots \quad a_{mi}],$$

就是 A 的 i -th column

$$\begin{bmatrix} a_{1i} \\ a_{2i} \\ \vdots \\ a_{mi} \end{bmatrix}$$

寫成 row vector 的形式. 同理可得, 對於 $1 \leq j \leq m$, A^T 的 j -th column 就是將 A 的 j -th row 寫成 column vector. \square

現在我們來看矩陣取 transpose 的基本性質.

Proposition 3.2.4. 假設 A, B 為 $m \times n$ matrix, C 為 $n \times l$ matrix. 我們有以下之性質.

- (1) $(A^T)^T = A$.
- (2) $(A+B)^T = A^T + B^T$.
- (3) $(AC)^T = C^T A^T$.

Proof. 首先觀察 A^T 為 $n \times m$ matrix, 故 $(A^T)^T$ 為 $m \times n$ matrix, 與 A 階數相同. 同樣的, $A^T + B^T$ 為 $n \times m$ matrix 與 $(A+B)^T$ 的階數相同. 另一方面 C^T 為 $l \times n$ matrix, 故 $C^T A^T$ 為 $l \times m$ matrix. 而 AC 為 $m \times l$ matrix, 所以 $(AC)^T$ 為 $l \times m$ matrix 與 $C^T A^T$ 階數相同.

(1) 因 $(A^T)^T$ 與 A 皆為 $m \times n$ matrix, 對於 $1 \leq i \leq n$, 我們只要檢查 $(A^T)^T$ 的 i -th column 就是 A 的 i -th column. 然而 $(A^T)^T$ 的 i -th column 依 Lemma 3.2.3 知就是 A^T 的 i -th row 寫成 column vector, 而 A^T 的 i -th row 就是 A 的 i -th column. 故得證 $(A^T)^T = A$.

(2) 因 $A^T + B^T$ 與 $(A+B)^T$ 皆為 $n \times m$ matrix, 對於 $1 \leq i \leq m$, 我們只要檢查 $A^T + B^T$ 的 i -th column 就是 $(A+B)^T$ 的 i -th column. 依定義 $A^T + B^T$ 的 i -th column 就是 A^T 和 B^T 的 i -th column 之和. 依 Lemma 3.2.3 知它就是 A 和 B 的 i -th row 之和. 另一方面, $(A+B)^T$ 的 i -th column 就是 $A+B$ 的 i -th row, 也就是 A 和 B 的 i -th row 之和. 得證 $(A+B)^T = A^T + B^T$.

(3) 由於 $(AC)^T$ 的 column 是由 AC 的 row 所決定, 而我們尚未討論 A 和 C 相乘 row 之間的關係, 所以這裡我們利用 entry 相同來證明相等. 對於 $1 \leq i \leq l, 1 \leq j \leq m$, $(AC)^T$ 的 (i, j) -th entry 為 AC 的 (j, i) -th entry, 即 A 的 j -th row 和 C 的 i -th column (看成 \mathbb{R}^n 的向量) 的內積. 另一方面, $C^T A^T$ 的 (i, j) -th entry 為 C^T 的 i -th row 和 A^T 的 j -th column 的內積. 這也是 C 的 i -th column 和 A 的 j -th row 的內積, 故得證 $(AC)^T = C^T A^T$. \square

Question 3.5. 假設 A 為 $m \times n$ matrix, $r \in \mathbb{R}$. 試證明 $(rA)^T = rA^T$.

一個 $n \times n$ square matrix, 若滿足 $A^T = A$, 我們稱 A 為 *symmetric matrix*. 上一節介紹過的 diagonal matrix 就是 symmetric matrix. 以後我們會學到 symmetric matrix 的重要性質, 現在我們先看和 symmetric matrix 有關的幾個簡單情形.

Corollary 3.2.5. 假設 A 為 $n \times n$ square matrix, B 為 $m \times n$ matrix. 以下皆為 *symmetric matrix*.

$$A + A^T, BB^T, B^T B.$$

Proof. 由 Proposition 3.2.4, 我們有 $(A + A^T)^T = A^T + (A^T)^T = A^T + A$, 故知 $A + A^T$ 為 symmetric matrix. 另一方面, $(BB^T)^T = (B^T)^T B^T = BB^T$, 故得 BB^T 為 symmetric matrix. 同理可得 $B^T B$ 亦為 symmetric matrix. \square

利用 Proposition 3.2.4, 我們可以從 row 的角度處理矩陣的乘法. 首先我們看一個 $1 \times m$ matrix 乘上一個 $m \times n$ matrix 的情形. 假設 $A \in \mathcal{M}_{1 \times m}, B \in \mathcal{M}_{m \times n}$, 令

$$A = [a_1 \quad a_2 \quad \cdots \quad a_m], B = \begin{bmatrix} b_{11} & b_{12} & \cdots & b_{1n} \\ b_{21} & b_{22} & \cdots & b_{2n} \\ \vdots & \vdots & \cdots & \vdots \\ b_{m1} & b_{m2} & \cdots & b_{mn} \end{bmatrix}.$$

則由 $(AB)^T = B^T A^T$, 以及矩陣右邊乘 column vector 的定義得

$$(AB)^T = \begin{bmatrix} b_{11} & b_{21} & \cdots & b_{m1} \\ b_{12} & b_{22} & \cdots & b_{m2} \\ \vdots & \vdots & \cdots & \vdots \\ b_{1n} & b_{2n} & \cdots & b_{mn} \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_m \end{bmatrix} = a_1 \begin{bmatrix} b_{11} \\ b_{12} \\ \vdots \\ b_{1n} \end{bmatrix} + a_2 \begin{bmatrix} b_{21} \\ b_{22} \\ \vdots \\ b_{2n} \end{bmatrix} + \cdots + a_m \begin{bmatrix} b_{m1} \\ b_{m2} \\ \vdots \\ b_{mn} \end{bmatrix}.$$

亦即 $(AB)^T = a_1(\mathbf{1}\mathbf{b})^T + a_2(\mathbf{2}\mathbf{b})^T + \cdots + a_m(\mathbf{m}\mathbf{b})^T$, 這裡 $(i\mathbf{b})^T$ 指的是將 B 的 i -th row 取轉置 (寫成 column 的形式). 故利用 Proposition 3.2.4 將 $(AB)^T$ 再取轉置還原得

$$AB = a_1(\mathbf{1}\mathbf{b}) + a_2(\mathbf{2}\mathbf{b}) + \cdots + a_m(\mathbf{m}\mathbf{b}).$$

也就是說

$$[a_1 \ \cdots \ a_m] \begin{bmatrix} b_{11} & \cdots & b_{1n} \\ \vdots & \cdots & \vdots \\ b_{m1} & \cdots & b_{mn} \end{bmatrix} = a_1 [b_{11} \ \cdots \ b_{1n}] + \cdots + a_m [b_{m1} \ \cdots \ b_{mn}] \quad (3.14)$$

現在我們來看一般的情形, 設 $A = [a_{ij}]$ 為 $m \times n$ matrix 以及 $B = [b_{jk}]$ 為 $n \times l$ matrix. 考慮 $(AB)^T = B^T A^T$. 依定義 $B^T A^T$ 的 i -th column, 為 B^T 右邊乘上 A^T 的 i -th column. 然而 A^T 的 i -th column, 為 A 的 i -th row 取轉置, 即 $(i\mathbf{a})^T$. 也就是說 $(AB)^T$ 的 i -th column 為 $B^T(i\mathbf{a})^T$. 利用 Proposition 3.2.4 再取轉置還原得, AB 的 i -th row 為

$$(B^T(i\mathbf{a})^T)^T = ((i\mathbf{a})^T)^T (B^T)^T = {}_i\mathbf{a}B.$$

換言之, 我們有以下的圖示

$$\begin{bmatrix} - & \mathbf{1}\mathbf{a} & - \\ - & \mathbf{2}\mathbf{a} & - \\ & \vdots & \\ - & \mathbf{m}\mathbf{a} & - \end{bmatrix} B = \begin{bmatrix} - & \mathbf{1}\mathbf{a}B & - \\ - & \mathbf{2}\mathbf{a}B & - \\ & \vdots & \\ - & \mathbf{m}\mathbf{a}B & - \end{bmatrix}. \quad (3.15)$$

結合式子 (3.14), 我們有以下之結果.

Proposition 3.2.6. 設 $A = [a_{ij}]$ 為 $m \times n$ matrix 以及 $B = [b_{jk}]$ 為 $n \times l$ matrix, 則對於 $1 \leq i \leq m$, AB 的 i -th row 為

$${}_i\mathbf{a}B = [a_{i1} \ \cdots \ a_{in}] \begin{bmatrix} b_{11} & \cdots & b_{1l} \\ \vdots & \cdots & \vdots \\ b_{n1} & \cdots & b_{nl} \end{bmatrix} = a_{i1} [b_{11} \ \cdots \ b_{1l}] + \cdots + a_{in} [b_{n1} \ \cdots \ b_{nl}].$$

最後我們介紹矩陣乘法的結合律.

Proposition 3.2.7. 假設 $A \in \mathcal{M}_{m \times n}$, $B \in \mathcal{M}_{n \times l}$, $C \in \mathcal{M}_{l \times k}$, 則 $(AB)C = A(BC)$.

Proof. 依定義 $AB \in \mathcal{M}_{m \times l}$, 故 $(AB)C \in \mathcal{M}_{m \times k}$. 而 $BC \in \mathcal{M}_{n \times k}$, 故 $A(BC) \in \mathcal{M}_{m \times k}$ 與 $(AB)C$ 同階. 對於 $1 \leq i \leq m$, $1 \leq j \leq k$, 我們要證明 $(AB)C$ 和 $A(BC)$ 的 (i, j) -th entry 相等. 令 \mathbf{a} 為 A 的 i -th row, 且令 \mathbf{c}_j 為 C 的 j -th column. 依定義 $(AB)C$ 的 (i, j) -th entry 為 AB 的 i -th row 乘上 C 的 j -th column. 由 Proposition 3.2.6 知 AB 的 i -th row 為 ${}_i\mathbf{a}B$, 故 $(AB)C$ 的 (i, j) -th entry 為 $({}_i\mathbf{a}B)\mathbf{c}_j$. 另一方面, $A(BC)$ 的 (i, j) -th entry 為 A 的 i -th row 乘上 BC 的 j -th column. 由 Definition 3.1.6 知 BC 的 j -th row 為 $B\mathbf{c}_j$, 故 $A(BC)$ 的 (i, j) -th entry 為 $\mathbf{a}(B\mathbf{c}_j)$. 也就是說, 要證明 $(AB)C$ 和 $A(BC)$ 的 (i, j) -th entry 相等, 就是要證明 $({}_i\mathbf{a}B)\mathbf{c}_j = \mathbf{a}(B\mathbf{c}_j)$. 由

於 $i\mathbf{a}$ 和 \mathbf{c}_j 分別只有一個 row 和 column, 為了方便考量, 我們將 $i\mathbf{a}$ 和 \mathbf{c}_j 用單一足碼表達. 也就是說假設

$$i\mathbf{a} = [a_1 \ a_2 \ \cdots \ a_n], B = \begin{bmatrix} b_{11} & b_{12} & \cdots & b_{1l} \\ b_{21} & b_{22} & \cdots & b_{2l} \\ \vdots & \vdots & \cdots & \vdots \\ b_{n1} & b_{n2} & \cdots & b_{nl} \end{bmatrix}, \mathbf{c}_j = \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_l \end{bmatrix}$$

依矩陣乘法定義

$$i\mathbf{a}B = [a_1 \ a_2 \ \cdots \ a_n] \begin{bmatrix} b_{11} & b_{12} & \cdots & b_{1l} \\ b_{21} & b_{22} & \cdots & b_{2l} \\ \vdots & \vdots & \cdots & \vdots \\ b_{n1} & b_{n2} & \cdots & b_{nl} \end{bmatrix} = [i\mathbf{a}\mathbf{b}_1 \ i\mathbf{a}\mathbf{b}_2 \ \cdots \ i\mathbf{a}\mathbf{b}_l]$$

因此得

$$(i\mathbf{a}B)\mathbf{c}_j = [i\mathbf{a}\mathbf{b}_1 \ i\mathbf{a}\mathbf{b}_2 \ \cdots \ i\mathbf{a}\mathbf{b}_l] \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_l \end{bmatrix} = c_1(i\mathbf{a}\mathbf{b}_1) + c_2(i\mathbf{a}\mathbf{b}_2) + \cdots + c_l(i\mathbf{a}\mathbf{b}_l).$$

另一方面, 由矩陣乘法的定義 (Definition 3.1.4) 知

$$B\mathbf{c}_j = \begin{bmatrix} b_{11} & b_{12} & \cdots & b_{1l} \\ b_{21} & b_{22} & \cdots & b_{2l} \\ \vdots & \vdots & \cdots & \vdots \\ b_{n1} & b_{n2} & \cdots & b_{nl} \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_l \end{bmatrix} = c_1\mathbf{b}_1 + c_2\mathbf{b}_2 + \cdots + c_l\mathbf{b}_l,$$

故由內積性質 (Proposition 1.4.2), 得

$$i\mathbf{a}(B\mathbf{c}_j) = i\mathbf{a}(c_1\mathbf{b}_1 + c_2\mathbf{b}_2 + \cdots + c_l\mathbf{b}_l) = c_1(i\mathbf{a}\mathbf{b}_1) + c_2(i\mathbf{a}\mathbf{b}_2) + \cdots + c_l(i\mathbf{a}\mathbf{b}_l).$$

得證 $(i\mathbf{a}B)\mathbf{c}_j = i\mathbf{a}(B\mathbf{c}_j)$, 故知 $(AB)C = A(BC)$. \square

Question 3.6. 假設 $A \in \mathcal{M}_{m \times n}$, $r \in \mathbb{R}$, 試證明 $(rI_m)A = rA = A(rI_n)$ 並利用這個等式以及 Proposition 3.2.7 證明 Proposition 3.1.9.

有了矩陣乘法的結合律 (Proposition 3.2.7), 以後我們談多個矩陣相乘時, 為了方便起見, 我們會捨去括號例如直接用 ABC 表示. 特別的, 當 A 為方陣時, 既然 $(AA)A = A(AA)$, 我們就用 A^3 來表示. 同理, 當 n 個 A 相乘時, 我們就用 A^n 來表示.

3.3. Elementary Matrix

我們知道矩陣 A 左邊乘上另一矩陣 E , 可以視為 E 的 row 對矩陣 A 的作用. 事實上, 若將 A 做一個 elementary row operation, 會是將 A 的左邊乘上一個矩陣. 這樣的矩陣我們稱之為 elementary matrix.

設 $A = [a_{ij}]$ 為 $m \times n$ matrix. 首先觀察 identity matrix I_m 對 A 的作用. 由於 I_m 的 i -th row 為

$$[0 \ \cdots \ 1 \ \cdots \ 0] \\ \hat{i}$$

即 i -th entry 為 1, 其他 entry 皆為 0. 所以依 Proposition 3.2.6, $I_m A$ 的 i -th row 為

$$\begin{bmatrix} 0 & \cdots & 1 & \cdots & 0 \end{bmatrix} A = 0_1 \mathbf{a} + \cdots + 1_i \mathbf{a} + \cdots + 0_m \mathbf{a} = \mathbf{a}_i,$$

(就是將 1 乘上 A 的 i -th row, 而將 0 乘上 A 的其他 row 再加起來, 故為 A 的 i -th row.) 換言之, 將 I_m 乘在 A 的左邊, 會將 A 的每一個 row 都固定不變, 所以知 $I_m A = A$. 現若 $j \neq i$ 且 E 為將 I_m 的 i -th row 改為 j -th entry 為 1 其他 entry 為 0, 而 i -th row 以外的其他 row 不變. 從上面的看法知 EA 的 i -th row 會是 A 的 j -th row, 也就是說 EA 會是將 A 的 i -th row 換成 A 的 j -th row, 而其他的 row 不動的矩陣.

現若用 i -th row 和 j -th row 交換的 elementary row operation 將 I_m 轉換成矩陣 E , 即

$$E = \begin{bmatrix} 1 & & & & & & \\ & \ddots & & & & & \\ & & 0 & & 1 & & \\ & & & \ddots & & & \\ & & 1 & & 0 & & \\ & & & & & \ddots & \\ & & & & & & 1 \end{bmatrix} \quad (3.16)$$

則利用前述的說法, EA 的 i -th row 是 A 的 j -th row, 而 EA 的 j -th row 是 A 的 i -th row, 而其他的 row 都不變. 換言之, EA 就是將 A 利用 i -th row 和 j -th row 交換這樣的 elementary row operation 變換所得的矩陣.

同樣的若將 I_m 的 i -th row 乘上實數 r 加到 I_m 的 j -th row 所得的矩陣為 E , 即

$$E = \begin{bmatrix} 1 & & & & & & \\ & \ddots & & & & & \\ & & 1 & & & & \\ & & & \ddots & & & \\ & & r & & 1 & & \\ & & & & & \ddots & \\ & & & & & & 1 \end{bmatrix} \quad (3.17)$$

則因 E 的 j -th row 的 i -th entry 為 r , j -th entry 為 1. 故由 Proposition 3.2.6, EA 的 j -th row 就是將 r 乘上 A 的 i -th row 後再加上 A 的 j -th row, 而其他的 row 都不變. 換言之, EA 就是將 A 的 i -th row 乘上實數 r 加到 A 的 j -th row 這樣的 elementary row operation 變換所得的矩陣.

最後若將 I_m 的 i -th row 乘上非零實數 r 所得的矩陣為 E , 即

$$E = \begin{bmatrix} 1 & & & & & & \\ & \ddots & & & & & \\ & & 1 & & & & \\ & & & r & & & \\ & & & & 1 & & \\ & & & & & \ddots & \\ & & & & & & 1 \end{bmatrix} \quad (3.18)$$

則很容易看出 EA 的 i -th row 就是將 A 的 i -th row 乘上 r , 而其餘的 row 不變. 也就是說, EA 就是將 A 的 i -th row 乘上非零實數 r 這樣的 elementary row operation 變換所得的矩陣.

從上面的說明我們知道, 對一個 $m \times n$ matrix 做一個 elementary row operation, 事實上就是將此矩陣左邊乘上一個 $m \times m$ matrix. 這個 $m \times m$ matrix 就是將 $m \times m$ identity matrix I_m 做同樣的 elementary row operation 所得的矩陣. 這樣的矩陣我們稱之為 *elementary matrix*. (3.16), (3.17), (3.18) 就是 elementary matrix 的三種形式.

當我們對一個 $m \times n$ matrix A , 進行多次的 elementary row operations, 就是將 A 左邊逐次的乘上相對應的 elementary matrix. 比方說做兩次 elementary row operations, 就是將 A 的左邊乘上第一次 elementary row operation 所對應的 elementary matrix E_1 . 做第二次時就是將 E_1A 左邊再乘上第二次 elementary row operation 所對應的 elementary matrix E_2 . 故所得的矩陣 $E_2(E_1A)$ 就是將 A 做這兩次 elementary row operations 所得的矩陣. 又由於矩陣乘法的結合律, 我們又可以將 $E_2(E_1A)$ 寫成 $(E_2E_1)A$. 同理, 對一個矩陣 A 進行一連串的 elementary row operations, 就是將 A 左邊乘上一個矩陣, 而這個矩陣就是這一連串 elementary row operations 所對應的 elementary matrices 的乘積. 不過要注意, 這些 elementary matrices 乘在一起的順序很重要, 因為 elementary matrices 之間的乘法不一定可以交換.

Question 3.7. 試找出那些同階的 *elementary matrices* 其相乘是不可以交換的.

有時我們需知道一個矩陣經由一連串的 elementary row operations, 其左邊到底是乘上哪一個矩陣. 當然我們可以如前述將所對應的 elementary matrices 乘在一起即可, 但這樣做其實很麻煩費時. 接下來我們來看一個很 “clever” 的方法, 可以在做 elementary row operation 時便幫我們將這個矩陣記錄下來. 這個方法就是, 若要對一個 $m \times n$ matrix A 做 elementary row operations, 我們先寫下一個 augmented matrix $[A|I_m]$. 也就是一個 $m \times (n+m)$ 的增廣矩陣, 其左邊 n 個 columns (即前 n 個 columns) 為矩陣 A , 而右邊 m 個 columns (即後 m 個 column) 為 I_m . 現將 A 做第一次的 elementary row operation, 假設其對應的 elementary matrix 為 E_1 , 則對 $[A|I_m]$ 做相同的 elementary row operation 的話, 所得的結果為 $E_1[A|I_m]$. 然而依矩陣乘法的定義我們知

$$E_1[A|I_m] = [E_1A|E_1I_m] = [E_1A|E_1].$$

也就是說, 當我們對 $[A|I_m]$ 做同樣的 elementary row operation, 所得的增廣矩陣其左邊就是將 A 做此 elementary row operation 所得的矩陣, 而右邊就是此 elementary row operation 所對應的 elementary matrix. 接著當我們做下一個 elementary row operation, 假設此 elementary row operation 所對應的 elementary matrix 為 E_2 , 則此 elementary row operation 對 $[E_1A|E_1]$ 作用後所得的矩陣便是 $E_2[E_1A|E_1] = [E_2(E_1A)|E_2E_1]$. 這樣繼續下去, 當我們對增廣矩陣 $[A|I_m]$ 進行一連串的 elementary row operations 後, 所得的矩陣 $[A'|E]$, 其左邊 A' 就是 A 經由這一連串的 elementary row operations 作用後所得的矩陣, 而右邊的

E 就是這些 elementary row operations 所對應的 elementary matrices 依序從右到左相乘所得的結果, 因此 $EA = A'$. 我們有以下的結論.

Lemma 3.3.1. 假設 A 為 $m \times n$ matrix. 若將 A 經由一連串的 elementary row operations 轉換成 A' , 則存在一個 $m \times m$ matrix E 使得 $EA = A'$, 其中 E 為這一連串 elementary row operations 所對應的 elementary matrix 由右而左依序相乘的乘積. 事實上若將 augmented matrix $[A|I_m]$ 經由同樣的 elementary row operations 作用後, 所得的 augmented matrix 就是 $[A'|E]$.

Example 3.3.2. 將矩陣

$$A = \begin{bmatrix} 2 & -4 & 4 & -6 \\ 1 & -2 & 1 & -1 \\ 4 & -8 & 4 & -4 \end{bmatrix}$$

化為 reduced echelon form, 並找到 elementary matrices 的乘積 E 使得 EA 為此 reduced echelon form.

首先寫下 augmented matrix

$$[A|I_3] = \left[\begin{array}{cccc|ccc} 2 & -4 & 4 & -6 & 1 & 0 & 0 \\ 1 & -2 & 1 & -1 & 0 & 1 & 0 \\ 4 & -8 & 4 & -4 & 0 & 0 & 1 \end{array} \right]$$

並將此 augmented matrix 的 1-st 和 2-nd row 交換, 得

$$\left[\begin{array}{cccc|ccc} 1 & -2 & 1 & -1 & 0 & 1 & 0 \\ 2 & -4 & 4 & -6 & 1 & 0 & 0 \\ 4 & -8 & 4 & -4 & 0 & 0 & 1 \end{array} \right]$$

接著我們將 augmented matrix 的 1-st row 乘上 -2 加到 2-nd row 上, 得

$$\left[\begin{array}{cccc|ccc} 1 & -2 & 1 & -1 & 0 & 1 & 0 \\ 0 & 0 & 2 & -4 & 1 & -2 & 0 \\ 4 & -8 & 4 & -4 & 0 & 0 & 1 \end{array} \right].$$

然後將 augmented matrix 的 1-st row 乘上 -4 加到 3-rd row 得

$$\left[\begin{array}{cccc|ccc} 1 & -2 & 1 & -1 & 0 & 1 & 0 \\ 0 & 0 & 2 & -4 & 1 & -2 & 0 \\ 0 & 0 & 0 & 0 & 0 & -4 & 1 \end{array} \right].$$

繼續將 augmented matrix 的 2-nd row 乘上 $1/2$ 得

$$\left[\begin{array}{cccc|ccc} 1 & -2 & 1 & -1 & 0 & 1 & 0 \\ 0 & 0 & 1 & -2 & \frac{1}{2} & -1 & 0 \\ 0 & 0 & 0 & 0 & 0 & -4 & 1 \end{array} \right].$$

最後將 augmented matrix 的 2-nd row 乘上 -1 加到 1-st row 得

$$\left[\begin{array}{cccc|ccc} 1 & -2 & 0 & 1 & -\frac{1}{2} & 2 & 0 \\ 0 & 0 & 1 & -2 & \frac{1}{2} & -1 & 0 \\ 0 & 0 & 0 & 0 & 0 & -4 & 1 \end{array} \right].$$

令最後所得的 augmented matrix 為 $[A'|E]$, 我們檢查是否 A 的 reduced echelon form A' 就是 EA . 事實上, 我們確有

$$EA = \begin{bmatrix} -\frac{1}{2} & 2 & 0 \\ \frac{1}{2} & -1 & 0 \\ 0 & -4 & 1 \end{bmatrix} \begin{bmatrix} 2 & -4 & 4 & -6 \\ 1 & -2 & 1 & -1 \\ 4 & -8 & 4 & -4 \end{bmatrix} = \begin{bmatrix} 1 & -2 & 0 & 1 \\ 0 & 0 & 1 & -2 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

另外我們想確認 E 確為這五個 elementary row operations 所對應的 elementary matrices 的乘積. 因為 A 為 3×4 matrix, 所以第一個 elementary row operation 所對應的 elementary matrix E_1 就是將 3×3 的 identity matrix I_3 的 1-st 和 2-nd row 交換, 而第二個 elementary matrix E_2 為將 I_3 的 1-st row 乘上 -2 加到 2-nd row 上. 第三個 elementary matrix E_3 為將 I_3 的 1-st row 乘上 -4 加到 3-rd row 上. 接下來的 elementary matrix E_4 為將 I_3 的 2-nd row 乘上 $1/2$, 而最後一個 elementary matrix E_5 為將 I_3 的 2-nd row 乘上 -1 加到 1-st row 上. 也就是說, 我們有

$$E_1 = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}, E_2 = \begin{bmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, E_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -4 & 0 & 1 \end{bmatrix},$$

$$E_4 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{2} & 0 \\ 0 & 0 & 1 \end{bmatrix}, E_5 = \begin{bmatrix} 1 & -1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

將這五個 elementary matrices 由右而左依序相乘, 確實得

$$E_5 E_4 E_3 E_2 E_1 = \begin{bmatrix} -\frac{1}{2} & 2 & 0 \\ \frac{1}{2} & -1 & 0 \\ 0 & -4 & 1 \end{bmatrix} = E.$$

最後我們要再次強調 elementary matrices 之間相乘是沒有交換性的. 事實上一個矩陣若右邊乘上一個 elementary matrix, 是對此矩陣做所謂的 elementary column operation. 這用類似前面的看法, 改用 column 的觀點來處理矩陣乘法 (Definition 3.1.6) 便可看出. 也就是說若 E 為將 identity I_n 的 i -th row 和 j -th row 交換的 elementary matrix, 其實也可將 E 視為將 I_n 的 i -th column 和 j -th column 交換. 所以若將 E 乘在一個 $m \times n$ matrix A 的右邊, 所得的矩陣 AE , 事實上就是將 A 的 i -th column 和 j -th column 交換. 同樣的, 若 E 為將 identity I_n 的 i -th row 乘上一非零實數 r 的 elementary matrix, 我們也可將 E 視為將 I_n 的 i -th column 乘上 r . 所以若將 E 乘在一個 $m \times n$ matrix A 的右邊, 所得的矩陣 AE , 事實上就是將 A 的 i -th column 乘上 r . 不過要特別注意的是將 I_n 的 i -th row 乘上 r 加到 j -th row 的 elementary matrix E . 這個 elementary matrix 若視為 column operation, 是將 I_n 的 j -th column 乘上 r 加在 i -th column 上 (不是 i -th column 乘上 r 加在 j -th column 上). 所以若將 E 乘在一個 $m \times n$ matrix A 的右邊, 所得的矩陣 AE , 事實上就是將 A 的 j -th column 乘上 r 加在 i -th column 上.

Example 3.3.3. 考慮

$$E_1 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}, E_2 = \begin{bmatrix} 10 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, E_3 = \begin{bmatrix} 1 & 0 & 10 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, A = \begin{bmatrix} 1 & 2 & 3 \\ -1 & -2 & -3 \\ 11 & 22 & 33 \end{bmatrix}$$

E_1 可視為將 I_3 的 2-nd row 和 3-rd row 交換, 也可視為將 I_3 的 2-nd column 和 3-rd column 交換. 事實上我們有

$$E_1A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 2 & 3 \\ -1 & -2 & -3 \\ 11 & 22 & 33 \end{bmatrix} = \begin{bmatrix} 1 & 2 & 3 \\ 11 & 22 & 33 \\ -1 & -2 & -3 \end{bmatrix},$$

$$AE_1 = \begin{bmatrix} 1 & 2 & 3 \\ -1 & -2 & -3 \\ 11 & 22 & 33 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 3 & 2 \\ -1 & -3 & -2 \\ 11 & 33 & 22 \end{bmatrix}.$$

E_2 可視為將 I_3 的 1-st row 乘以 10, 也可視為將 I_3 的 1-st column 乘以 10. 事實上我們有

$$E_2A = \begin{bmatrix} 10 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 2 & 3 \\ -1 & -2 & -3 \\ 11 & 22 & 33 \end{bmatrix} = \begin{bmatrix} 10 & 20 & 30 \\ -1 & -2 & -3 \\ 11 & 22 & 33 \end{bmatrix},$$

$$AE_2 = \begin{bmatrix} 1 & 2 & 3 \\ -1 & -2 & -3 \\ 11 & 22 & 33 \end{bmatrix} \begin{bmatrix} 10 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 10 & 2 & 3 \\ -10 & -2 & -3 \\ 110 & 22 & 33 \end{bmatrix}.$$

E_3 可視為將 I_3 的 3-rd row 乘以 10 加到 1-st row, 也可視為將 I_3 的 1-st column 乘以 10 加到 3-rd column. 事實上我們有

$$E_3A = \begin{bmatrix} 1 & 0 & 10 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 2 & 3 \\ -1 & -2 & -3 \\ 11 & 22 & 33 \end{bmatrix} = \begin{bmatrix} 111 & 222 & 333 \\ -1 & -2 & -3 \\ 11 & 22 & 33 \end{bmatrix},$$

$$AE_3 = \begin{bmatrix} 1 & 2 & 3 \\ -1 & -2 & -3 \\ 11 & 22 & 33 \end{bmatrix} \begin{bmatrix} 1 & 0 & 10 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 2 & 13 \\ -1 & -2 & -13 \\ 11 & 22 & 143 \end{bmatrix}.$$

3.4. Matrix 和 System of Linear Equations 的連結

在上一章, 我們曾經利用矩陣的 rank 來探討其所對應的聯立方程組何時有解以及解是否唯一的問題. 現在我們又知道解一次聯立方程組的問題可以看成矩陣乘法的問題, 這一節中我們就是要用這個觀點進一步探討聯立方程組何時有解以及解是否唯一.

首先由於我們都要用矩陣的乘法來探討, 為了方便起見對於 \mathbb{R}^n 中的向量, 除非特別聲明為 row vector, 我們將一律用 column vector 來表示. 也就是說將它視為一個 $n \times 1$ matrix. 所以若 $\mathbf{u}, \mathbf{v} \in \mathbb{R}^n$, 我們用 $\mathbf{u}^T \mathbf{v}$ 表示 \mathbf{u}, \mathbf{v} 的內積 (而不用 $\mathbf{u} \cdot \mathbf{v}$). 另外回顧, 給定一次聯立方程組

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

我們令

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

然後將上面的聯立方程組用 $\mathbf{Ax} = \mathbf{b}$ 來表示. 現若 $x_1 = c_1, x_2 = c_2, \dots, x_n = c_n$, 為此聯立方程組的一組解, 我們便會用

$$\mathbf{c} = \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_m \end{bmatrix},$$

來表示這一組解, 而說 $\mathbf{x} = \mathbf{c} \in \mathbb{R}^n$ 為 $\mathbf{Ax} = \mathbf{b}$ 的一組解. 依矩陣乘法定義這等同於說 A 這一個 $m \times n$ matrix 乘以 \mathbf{c} 這一個 $n \times 1$ matrix 會等於 \mathbf{b} 這一個 $m \times 1$ matrix, 即 $\mathbf{Ac} = \mathbf{b}$.

3.4.1. 解的存在性. 我們再一次探討怎樣的 $m \times n$ matrix A 會滿足對任意的 $\mathbf{b} \in \mathbb{R}^m$, 聯立方程組 $\mathbf{Ax} = \mathbf{b}$ 皆有解.

首先假設 $\mathbf{b} \in \mathbb{R}^m$ 且 $\mathbf{Ax} = \mathbf{b}$ 有解. 令 $\mathbf{c} = \begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix}$ 為一解, 此即表示 $\mathbf{Ac} = \mathbf{b}$. 利用矩陣乘法定義得

$$c_1 \mathbf{a}_1 + \dots + c_n \mathbf{a}_n = \mathbf{b},$$

其中 $\mathbf{a}_1, \dots, \mathbf{a}_n$ 為 A 的 column vectors. 換句話說, \mathbf{b} 可以寫成 A 的 column vectors 的 linear combination. 用符號來表示就是 $\mathbf{b} \in \text{Span}(\mathbf{a}_1, \dots, \mathbf{a}_n)$. 反之, 若 $\mathbf{b} \in \text{Span}(\mathbf{a}_1, \dots, \mathbf{a}_n)$, 表示存在 $c_1, \dots, c_n \in \mathbb{R}$ 使得 $\mathbf{b} = c_1 \mathbf{a}_1 + \dots + c_n \mathbf{a}_n$. 故得 $x_1 = c_1, \dots, x_n = c_n$ 為 $\mathbf{Ax} = \mathbf{b}$ 的一組解. 我們證得了以下的性質.

Lemma 3.4.1. 假設 $A \in \mathcal{M}_{m \times n}$ 且 $\mathbf{b} \in \mathbb{R}^m$. 則 $\mathbf{Ax} = \mathbf{b}$ 有解若且唯若 $\mathbf{b} \in \text{Span}(\mathbf{a}_1, \dots, \mathbf{a}_n)$, 其中 $\mathbf{a}_1, \dots, \mathbf{a}_n$ 為 A 的 column vectors.

我們有興趣於知道怎樣的 $m \times n$ matrix A 會使得對任意的 $\mathbf{b} \in \mathbb{R}^m$, 聯立方程組 $\mathbf{Ax} = \mathbf{b}$ 皆有解. 我們利用過去學過的幾種不同觀點, 發現有許多和它等價的條件. 首先觀察由於 A 的 column vectors 皆在 \mathbb{R}^m 中, 所以自然有 $\text{Span}(\mathbf{a}_1, \dots, \mathbf{a}_n) \subseteq \mathbb{R}^m$. 然而由 Lemma 3.4.1 知, 若對於任意 $\mathbf{b} \in \mathbb{R}^m$ 皆會使得 $\mathbf{Ax} = \mathbf{b}$ 有解, 表示對任意 $\mathbf{b} \in \mathbb{R}^m$ 皆有 $\mathbf{b} \in \text{Span}(\mathbf{a}_1, \dots, \mathbf{a}_n)$. 故知此時 $\text{Span}(\mathbf{a}_1, \dots, \mathbf{a}_n) = \mathbb{R}^m$. 反之, 若 $\text{Span}(\mathbf{a}_1, \dots, \mathbf{a}_n) = \mathbb{R}^m$, 表示對任意 $\mathbf{b} \in \mathbb{R}^m$ 皆有 $\mathbf{b} \in \text{Span}(\mathbf{a}_1, \dots, \mathbf{a}_n)$. 同樣由 Lemma 3.4.1 知此即對於任意 $\mathbf{b} \in \mathbb{R}^m$ 皆會使得 $\mathbf{Ax} = \mathbf{b}$ 有解. 因此從這觀點來看, 對任意的 $\mathbf{b} \in \mathbb{R}^m$, 聯立方程組 $\mathbf{Ax} = \mathbf{b}$ 皆有解和 $\text{Span}(\mathbf{a}_1, \dots, \mathbf{a}_n) = \mathbb{R}^m$ 是等價的.

另外我們可以考慮聯立方程組 $\mathbf{Ax} = \mathbf{e}_i$, 其中 $\mathbf{e}_1, \dots, \mathbf{e}_m \in \mathbb{R}^m$ 為 \mathbb{R}^m 的 standard basis. 若已知對任意的 $\mathbf{b} \in \mathbb{R}^m$, 聯立方程組 $\mathbf{Ax} = \mathbf{b}$ 皆有解, 則對所有的 $i = 1, \dots, m$, 我們都可找到 $\mathbf{c}_i \in \mathbb{R}^n$ 使得 $\mathbf{x} = \mathbf{c}_i$ 為聯立方程組 $\mathbf{Ax} = \mathbf{e}_i$ 的一組解. 也就是說對所有的 $i = 1, \dots, m$ 皆有 $\mathbf{Ac}_i = \mathbf{e}_i$. 現考慮 $n \times m$ matrix C , 其 i -th column 就是 \mathbf{c}_i . 此時依矩陣乘法的定義我們有

$$AC = A \begin{bmatrix} | & | & \cdots & | \\ \mathbf{c}_1 & \mathbf{c}_2 & \cdots & \mathbf{c}_m \\ | & | & & | \end{bmatrix} = \begin{bmatrix} | & | & \cdots & | \\ \mathbf{Ac}_1 & \mathbf{Ac}_2 & \cdots & \mathbf{Ac}_m \\ | & | & & | \end{bmatrix} = \begin{bmatrix} | & | & \cdots & | \\ \mathbf{e}_1 & \mathbf{e}_2 & \cdots & \mathbf{e}_m \\ | & | & & | \end{bmatrix} = I_m.$$

也就是說，此時必存在 $n \times m$ matrix C 使得 $AC = I_m$ 。反之，若 C 為 $n \times m$ matrix 滿足 $AC = I_m$ ，則對任意 $\mathbf{b} \in \mathbb{R}^m$ ，我們考慮 $\mathbf{c} = C\mathbf{b} \in \mathbb{R}^n$ ，皆會有

$$A\mathbf{c} = A(C\mathbf{b}) = (AC)\mathbf{b} = I_m\mathbf{b} = \mathbf{b}.$$

也就是說此時對任意 $\mathbf{b} \in \mathbb{R}^m$ ，我們都可以找到 $\mathbf{c} = C\mathbf{b} \in \mathbb{R}^n$ ，使得 $\mathbf{x} = \mathbf{c}$ 是聯立方程組 $A\mathbf{x} = \mathbf{b}$ 的一組解。因此從這觀點來看，對任意的 $\mathbf{b} \in \mathbb{R}^m$ ，聯立方程組 $A\mathbf{x} = \mathbf{b}$ 皆有解和存在 $n \times m$ matrix C 使得 $AC = I_m$ 是等價的。

我們也曾經在 2.3 節中利用 rank 探討過這個問題 (參見 Proposition 2.3.2)。也就是說，若 A 經由 elementary row operations 化為 echelon form 後，其 pivot 的個數恰等於 A 的 row 的個數 m ，表示 A 的 echelon form 沒有一個 row 全為 0，故由 2.1 節的討論 (即 Case (1)) 知此時任意的 $\mathbf{b} \in \mathbb{R}^m$ ，聯立方程組 $A\mathbf{x} = \mathbf{b}$ 皆有解。反之，如果 pivot 的個數不等於 m ，表示 A 的 echelon form A' 中最後一個 row 必全為 0。此時我們一定可以找到 $\mathbf{b} \in \mathbb{R}^m$ 使得增廣矩陣 $[A|\mathbf{b}]$ 化為 echelon form $[A'|\mathbf{b}']$ 後， \mathbf{b}' 最後一個 entry 不為 0 (即 Case 2(a))。此時 $A\mathbf{x} = \mathbf{b}$ 會無解。因此從這觀點來看，對任意的 $\mathbf{b} \in \mathbb{R}^m$ ，聯立方程組 $A\mathbf{x} = \mathbf{b}$ 皆有解和 A 的 echelon form 的 pivot 的個數為 m (即 $\text{rank}(A) = m$) 是等價的。

綜合上面這幾種看法，我們證得了以下這個非常重要的定理。

Theorem 3.4.2. 假設 A 為 $m \times n$ matrix，令 $\mathbf{a}_1, \dots, \mathbf{a}_n \in \mathbb{R}^m$ 為 A 的 column vectors。以下各敘述是等價的。

- (1) 對任意的 $\mathbf{b} \in \mathbb{R}^m$ ，聯立方程組 $A\mathbf{x} = \mathbf{b}$ 皆有解。
- (2) $\text{Span}(\mathbf{a}_1, \dots, \mathbf{a}_n) = \mathbb{R}^m$ 。
- (3) $\text{rank}(A) = m$ 。
- (4) 存在 $n \times m$ matrix C 使得 $AC = I_m$ 。

特別提醒一下，Theorem 3.4.2，指的是對所有 $\mathbf{b} \in \mathbb{R}^m$ ，聯立方程組 $A\mathbf{x} = \mathbf{b}$ 皆有解的情況。所以若僅知單一的 \mathbf{b} 使得聯立方程組 $A\mathbf{x} = \mathbf{b}$ 有解，Theorem 3.4.2 並不適用 (不過 Lemma 3.4.1 是適用的)。

我們曾提及，當 A 為 $m \times n$ matrix，將 A 化為 echelon form 後其 pivot 的個數不能多於 row 和 column 的個數。也就是說 pivot 的個數應小於等於 $\min\{m, n\}$ (此指的是 m, n 中最小的那一個)。所以若 pivot 的個數為 m ，則表示 $n \geq m$ 。換言之，若 $n < m$ ，我們便知 pivot 的個數不可能等於 m ，所以 Theorem 3.4.2 中的情況不可能發生。我們有以下的結論。

Corollary 3.4.3. 假設 A 為 $m \times n$ matrix，其中 $n < m$ ，則必存在 $\mathbf{b} \in \mathbb{R}^m$ 使得聯立方程組 $A\mathbf{x} = \mathbf{b}$ 無解。而且此時，不會存在 $n \times m$ matrix C 使得 $AC = I_m$ 。

Proof. 由前所述，當 $n < m$ 時 A 化為 echelon form 後，其 pivot 的個數不可能為 m ，亦即 $\text{rank}(A) < m$ 。故由 Theorem 3.4.2 知不可能對任意的 $\mathbf{b} \in \mathbb{R}^m$ ，聯立方程組 $A\mathbf{x} = \mathbf{b}$ 皆有解。亦即存在 $\mathbf{b} \in \mathbb{R}^m$ 使得聯立方程組 $A\mathbf{x} = \mathbf{b}$ 無解。同理，由 Theorem 3.4.2 知不會存在 $n \times m$ matrix C 使得 $AC = I_m$ 。□

Question 3.8. 假設 A 為 $m \times n$ matrix, 其中 $m < n$. 是否存在 $n \times m$ matrix C 使得 $CA = I_n$?

前面提過 Theorem 3.4.2 是個很重要的定理, 它可以告訴我們一些解聯立方程組的訊息. 例如 Corollary 3.4.3 就是告訴我們當方程式的個數多於未知數的個數時, 會存在 $\mathbf{b} \in \mathbb{R}^m$ 使得聯立方程組 $A\mathbf{x} = \mathbf{b}$ 無解. 我們也可已將它運用於有關於 spanning set 的問題.

Corollary 3.4.4. 設 m, n 為正整數且 $n < m$. 任取 \mathbb{R}^m 中 n 個 vectors $\mathbf{a}_1, \dots, \mathbf{a}_n$, 則 $\text{Span}(\mathbf{a}_1, \dots, \mathbf{a}_n) \neq \mathbb{R}^m$.

Proof. 我們用反證法, 假設存在 $\mathbf{a}_1, \dots, \mathbf{a}_n \in \mathbb{R}^m$ 使得 $\text{Span}(\mathbf{a}_1, \dots, \mathbf{a}_n) = \mathbb{R}^m$. 令 A 為由 $\mathbf{a}_1, \dots, \mathbf{a}_n$ 為 column 所組成的 $m \times n$ matrix, 即對於 $i = 1, \dots, n$, A 的 i -th column 為 \mathbf{a}_i . 則由 Theorem 3.4.2 知 $\text{rank}(A) = m$. 但矩陣 A 的 column 的個數小於 row 的個數 (即 $n < m$), 得 $\text{rank}(A) \leq n < m$ 之矛盾. 由此矛盾得知 $\text{Span}(\mathbf{a}_1, \dots, \mathbf{a}_n) \neq \mathbb{R}^m$. \square

Corollary 3.4.4 告訴我們, 在 \mathbb{R}^m 中任取少於 m 個向量, 它們的線性組合是不能展成整個 \mathbb{R}^m 的.

3.4.2. 解的唯一性. 所謂聯立方程組解的唯一性, 指的是假設聯立方程組有解時, 探討其解是否唯一. 所以唯一性並不涉及解是否存在的問題.

給定 $A \in \mathcal{M}_{m \times n}$ 以及 $\mathbf{b} \in \mathbb{R}^m$. 我們曾說明過如果 $A\mathbf{x} = \mathbf{b}$ 有解, 則 $A\mathbf{x} = \mathbf{b}$ 的解和 $A\mathbf{x} = \mathbf{0}$ 的解 (這裡 $\mathbf{0}$ 是 \mathbb{R}^m 的零向量) 息息相關 (參見 Proposition 2.3.6). 這個事實現在由於可以利用矩陣的運算給予更簡明的證明, 所以我們再敘述及證明一次.

Lemma 3.4.5. 給定 $A \in \mathcal{M}_{m \times n}$ 以及 $\mathbf{b} \in \mathbb{R}^m$ 且假設 $\mathbf{x} = \mathbf{c} \in \mathbb{R}^n$ 是聯立方程組 $A\mathbf{x} = \mathbf{b}$ 的一組解. 則

- (1) 若 $\mathbf{x} = \mathbf{c}' \in \mathbb{R}^n$ 是 $A\mathbf{x} = \mathbf{b}$ 的一組解, 則 $\mathbf{x} = \mathbf{c}' - \mathbf{c}$ 為 $A\mathbf{x} = \mathbf{0}$ 的一組解.
- (2) 若 $\mathbf{x} = \mathbf{u} \in \mathbb{R}^n$ 為 $A\mathbf{x} = \mathbf{0}$ 的一組解, 則 $\mathbf{x} = \mathbf{c} + \mathbf{u}$ 是 $A\mathbf{x} = \mathbf{b}$ 的一組解.

Proof. (1) 假設 $\mathbf{x} = \mathbf{c}' \in \mathbb{R}^n$ 是 $A\mathbf{x} = \mathbf{b}$ 的一組解, 意即 $A\mathbf{c}' = \mathbf{b}$. 由已知 $A\mathbf{c} = \mathbf{b}$ 得

$$A(\mathbf{c}' - \mathbf{c}) = A\mathbf{c}' - A\mathbf{c} = \mathbf{b} - \mathbf{b} = \mathbf{0}.$$

因此 $\mathbf{x} = \mathbf{c}' - \mathbf{c}$ 會是 $A\mathbf{x} = \mathbf{0}$ 的一組解.

- (2) 若 $\mathbf{x} = \mathbf{u} \in \mathbb{R}^n$ 為 $A\mathbf{x} = \mathbf{0}$ 的一組解, 則

$$A(\mathbf{c} + \mathbf{u}) = A\mathbf{c} + A\mathbf{u} = \mathbf{b} + \mathbf{0} = \mathbf{b}.$$

得證 $\mathbf{x} = \mathbf{c} + \mathbf{u}$ 為聯立方程組 $A\mathbf{x} = \mathbf{b}$ 的一組解. \square

Lemma 3.4.5 告訴我們若已知 $\mathbf{x} = \mathbf{c}$ 為 $A\mathbf{x} = \mathbf{b}$ 的一組解, 且知道 $A\mathbf{x} = \mathbf{0}$ 所有的解, 就能利用 \mathbf{c} 以及 $A\mathbf{x} = \mathbf{0}$ 所有的解得到 $A\mathbf{x} = \mathbf{b}$ 所有的解. 所以了解 $A\mathbf{x} = \mathbf{0}$ 所有的解是很重要的課題 (以後我們會深入探討). 回顧一下 $A\mathbf{x} = \mathbf{0}$ 這樣的 linear system, 我們稱之為 homogeneous linear system. Homogeneous linear system 一定有解, 事實上當 $A \in \mathcal{M}_{m \times n}$ 時, $x_1 = 0, \dots, x_n = 0$ 就是 $A\mathbf{x} = \mathbf{0}$ 的一組解. 這組解 $\mathbf{x} = \mathbf{0} \in \mathbb{R}^n$ 因為不需任何計算就能得到,

我們稱之為 $A\mathbf{x} = \mathbf{0}$ 的 trivial solution. 注意 trivial solution $\mathbf{x} = \mathbf{0}$ 這裡的 $\mathbf{0}$ 是 \mathbb{R}^n 的零向量, 而 $A\mathbf{x} = \mathbf{0}$ 這裡的 $\mathbf{0}$ 是 \mathbb{R}^m 的零向量, 所以雖然我們用同樣的符號表示, 但當 $n \neq m$ 時它們是不同的, 大家需區分清楚. 當一個 homogeneous linear system $A\mathbf{x} = \mathbf{0}$ 除了 trivial solution 外還有其他的 solution (即解不唯一), 我們稱這些不為 $\mathbf{0}$ 的 solution 為 nontrivial solution.

從 Lemma 3.4.5 我們知, 若 $A\mathbf{x} = \mathbf{0}$ 沒有 nontrivial solution (即解唯一), 則對於 $\mathbf{b} \in \mathbb{R}^m$, 若 $A\mathbf{x} = \mathbf{b}$ 有解, 其解必唯一. 由這觀點, 我們可以得到以下關於聯立方程組解的唯一性的重要定理.

Theorem 3.4.6. 假設 A 為 $m \times n$ matrix. 以下各敘述是等價的.

- (1) 若 $\mathbf{b} \in \mathbb{R}^m$ 且聯立方程組 $A\mathbf{x} = \mathbf{b}$ 有解, 則解唯一.
- (2) Homogeneous system $A\mathbf{x} = \mathbf{0}$ 沒有 nontrivial solution.
- (3) $\text{rank}(A) = n$.
- (4) 存在 $n \times m$ matrix B 使得 $BA = I_n$.

Proof. 在 Corollary 2.3.7 中我們已證得 (1), (2), (3) 是等價的. 故僅需證明 (3) \Rightarrow (4) 以及 (4) \Rightarrow (1).

(3) \Rightarrow (4): 假設 $\text{rank}(A) = n$, 即 A 化為 echelon form 後, 其 pivot 的個數為 n . 考慮將 A 化為 reduced echelon form A' . 此時 A' 由於有 n 個 pivot, 所以每一個 pivot 必分別在 A' 前面 n 個 row 上. 而又 A' 為 $m \times n$ matrix, 有 n 個 column. 所以 A' 每一個 pivot 必落在 (i, i) -th entry, 其中 $1 \leq i \leq n$. 又因為 A' 為 reduced echelon form, 此 n 個 pivots 的值皆為 1. 然而 reduced echelon form 每一個 pivot 所在的 column, 除了 pivot 所在位置外, 其他位置應為 0, 所以我們知 A' 必為以下的 matrix $A' = \begin{bmatrix} I_n \\ \mathbf{0} \end{bmatrix}$, 即 A' 的前 n 個 row 就是 I_n . 由 Lemma 3.3.1, 我們知存在 $m \times m$ matrix E 使得 $EA = A'$. 現若令 E 的 i -th row 為 ${}_i\mathcal{E}$, 由 row 的觀點看矩陣乘法 (參見圖示 (3.15)), 我們有

$$EA = \begin{bmatrix} - & {}_1\mathcal{E} & - \\ & \vdots & \\ - & {}_n\mathcal{E} & - \\ & \vdots & \\ - & {}_m\mathcal{E} & - \end{bmatrix} A = \begin{bmatrix} - & {}_1\mathcal{E}A & - \\ & \vdots & \\ - & {}_n\mathcal{E}A & - \\ & \vdots & \\ - & {}_m\mathcal{E}A & - \end{bmatrix} = \begin{bmatrix} 1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & 1 \\ & & & \mathbf{0} \end{bmatrix}.$$

現若令 B 為 $n \times m$ matrix, 對於 $i = 1, \dots, n$, 其 i -th row 為 ${}_i\mathcal{E}$ (即 B 為截取 E 的前 n 個 row 的 $n \times m$ matrix), 則由前述的矩陣乘法性質知

$$BA = \begin{bmatrix} - & {}_1\mathcal{E} & - \\ & \vdots & \\ - & {}_n\mathcal{E} & - \end{bmatrix} A = \begin{bmatrix} - & {}_1\mathcal{E}A & - \\ & \vdots & \\ - & {}_n\mathcal{E}A & - \end{bmatrix} = \begin{bmatrix} 1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & 1 \end{bmatrix} = I_n.$$

(4) \Rightarrow (1): 我們利用反證法假設 $\mathbf{c} \neq \mathbf{c}' \in \mathbb{R}^n$ 且 $\mathbf{x} = \mathbf{c}, \mathbf{x} = \mathbf{c}'$ 皆為 $A\mathbf{x} = \mathbf{b}$ 的一組解. 亦即, $A\mathbf{c} = \mathbf{b}$ 且 $A\mathbf{c}' = \mathbf{b}$. 現已知存在 $B \in \mathcal{M}_{n \times m}$ 使得 $BA = I_n$, 故得 $B\mathbf{b} = B(A\mathbf{c}) = (BA)\mathbf{c} = \mathbf{c}$ 且

$Bb = B(Ac') = (BA)c' = c'$. 此結果 $c = Bb = c'$ 與當初假設 $c \neq c'$ 相矛盾, 故得證若 $Ax = b$ 有解, 則解必唯一. \square

再次提醒, Theorem 3.4.6, 並不能知道聯立方程組 $Ax = b$ 是否有解. 它告訴我們若已知 $Ax = \mathbf{0}$ 有 nontrivial solution, 則 $Ax = b$ 要不然無解, 要不然會有無窮多解.

Question 3.9. 假設 A 為 $m \times n$ matrix, $b, b' \in \mathbb{R}^m$. 若已知 $Ax = b$, $Ax = b'$ 皆有解, 且 $Ax = b$ 的解唯一. 是否 $Ax = b'$ 的也會唯一?

前面已提過, 當 A 為 $m \times n$ matrix, 將 A 化為 echelon form 後其 pivot 的個數應小於等於 $\min\{m, n\}$. 所以若 pivot 的個數為 n , 則表示 $m \geq n$. 換言之, 若 $m < n$, 我們便知 pivot 的個數不可能等於 n , 所以 Theorem 3.4.6 中的情況不可能發生. 我們有以下的結論.

Corollary 3.4.7. 假設 A 為 $m \times n$ matrix, 其中 $m < n$. 若 $b \in \mathbb{R}^m$ 且聯立方程組 $Ax = b$ 有解, 則必有無窮多解. 而且此時, 不會存在 $n \times m$ matrix B 使得 $BA = I_n$.

Proof. 由前所述, 當 $m < n$ 時 A 化為 echelon form 後, 其 pivot 的個數不可能為 n . 故由 Theorem 3.4.6 知 homogeneous linear system $Ax = \mathbf{0}$ 有 nontrivial solution. 亦即在 \mathbb{R}^n 存在非零向量 c 使得 $x = c$ 為 $Ax = \mathbf{0}$ 之一組解. 此時考慮任何時數 r , 由矩陣乘法性質 (Proposition 3.1.9) 知,

$$A(rc) = r(Ac) = r\mathbf{0} = \mathbf{0}.$$

亦即 $x = rc$ 也是 $Ax = \mathbf{0}$ 的一組解. 故知 $Ax = \mathbf{0}$ 有無窮多解, 所以 Lemma 3.4.5 告訴我們, 若 $Ax = b$ 有解, 則必有無窮多解.

另一方面 Theorem 3.4.6 也告訴我們若 pivot 的個數不是 n , 則不會存在 $n \times m$ matrix B 使得 $BA = I_n$. \square

Theorem 3.4.6 也和 Theorem 3.4.2 一樣是很重要的定理, 它可以告訴我們一些解聯立方程組的訊息. 例如 Corollary 3.4.7 就是告訴我們當方程式的個數少於未知數的個數時, 聯立方程組 $Ax = b$ 不可能解唯一. 我們也可已將它運用於有關於“線性相關”的問題.

Corollary 3.4.8. 設 m, n 為正整數且 $m < n$. 任取 \mathbb{R}^m 中 n 個 vectors $\mathbf{a}_1, \dots, \mathbf{a}_n$, 則存在 $c_1, \dots, c_n \in \mathbb{R}$ 不全為 0 使得 $c_1\mathbf{a}_1 + \dots + c_n\mathbf{a}_n = \mathbf{0}$.

Proof. 令 A 為由 $\mathbf{a}_1, \dots, \mathbf{a}_n$ 為 column 所組成的 $m \times n$ matrix. 由於矩陣 A 的 row 的個數小於 column 的個數 (即 $m < n$), 由 Corollary 3.4.7 知聯立方程組 $Ax = \mathbf{0}$ 有 nontrivial solution $x_1 = c_1, \dots, x_n = c_n$. 由矩陣乘法定義, 此即表示 c_1, \dots, c_n 不全為 0 且 $c_1\mathbf{a}_1 + \dots + c_n\mathbf{a}_n = \mathbf{0}$. \square

假設 c_1, \dots, c_n 不全為 0 且 $c_1\mathbf{a}_1 + \dots + c_n\mathbf{a}_n = \mathbf{0}$. 不失一般性我們可假設 $c_1 \neq 0$, 此時

$$\mathbf{a}_1 = \frac{-c_2}{c_1}\mathbf{a}_2 + \dots + \frac{-c_n}{c_1}\mathbf{a}_n.$$

這表示 \mathbf{a}_1 可以用其他的向量 $\mathbf{a}_2, \dots, \mathbf{a}_n$ 的線性組合來表達. 因此我們便會稱 $\{\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n\}$ 這一組向量為 *linearly dependent* (線性相關). 一組向量之間是否為 linearly dependent 也是

線性代數中很重要的課題，以後我們會再正式的給定義並深入的探討。這裡，Corollary 3.4.8 告訴我們，在 \mathbb{R}^m 中任取多於 m 個向量，它們必為 linearly dependent.

3.5. Invertible Matrix

所謂 invertible matrix 就是“可逆矩陣”。我們會發現只有 square matrix 才有可能 invertible matrix，但並不是所有的 square matrix 都是 invertible matrix。這一節中我們會探討有關 invertible matrix 的相關性質，並介紹判斷一個方陣是否為 invertible 且找出其反矩陣的方法。

當初我們將聯立方程組用矩陣乘法的方式 $A\mathbf{x} = \mathbf{b}$ 來表示，其中有一個很大的目的就是希望將解聯立方程式的問題此簡化成類似實數上解 $ax = b$ 的情形。在實數情況，當 $a \neq 0$ 時， $ax = b$ 的解就是很簡單的 $x = ba^{-1}$ 。但在矩陣的情形，我們沒有除法，所以只能借助乘法來幫忙。由於實數中 a^{-1} 有 $a^{-1}a = aa^{-1} = 1$ 的性質，所以推廣這個概念至矩陣，我們便希望找到矩陣滿足 BA 以及 AB 為 identity。不過當 A 為 $m \times n$ matrix 且 $m \neq n$ 時，由 Corollary 3.4.3 以及 Corollary 3.4.7，我們知這是不可能存在的。所以我們僅對 $m = n$ ，即 A 為 square matrix 時有以下的定義。

Definition 3.5.1. 假設 A 為 $n \times n$ square matrix，若存在 $n \times n$ matrix B 使得 $AB = BA = I_n$ ，則稱 A 為 invertible。反之，我們稱 A 為 non-invertible

再一次強調當 A 不是方陣時，我們知 A 絕對不是 invertible。因此當我們不知矩陣 A 的階數時，絕對不能用存在 B 滿足 BA 為 identity 來說 A 為 invertible，必須檢查另一邊 AB 亦為 identity 才可。不過當 A 為 $n \times n$ square matrix，確實檢查單邊就可以確定 A 為 invertible。我們有以下的性質。

Theorem 3.5.2. 假設 A 為 $n \times n$ square matrix。則下列是等價的。

- (1) A 為 invertible matrix.
- (2) 存在 $n \times n$ matrix B 使得 $BA = I_n$.
- (3) $\text{rank}(A) = n$.
- (4) 存在 $n \times n$ matrix C 使得 $AC = I_n$.

Proof. 依 A 為 invertible 的定義，我們知若 A 為 invertible，則存在 $B \in \mathcal{M}_{n \times n}$ 使得 $BA = I_n$ 。故 (1) \Rightarrow (2)。

由 A 為 $n \times n$ matrix 以及 Theorem 3.4.6 知存在 $B \in \mathcal{M}_{n \times n}$ 使得 $BA = I_n$ 若且唯若 A 化為 echelon form 後 pivot 的個數為 n 。故 (2) \Leftrightarrow (3)。

同理，由 A 為 $n \times n$ matrix 以及 Theorem 3.4.2 知存在 $C \in \mathcal{M}_{n \times n}$ 使得 $AC = I_n$ 若且唯若 A 化為 echelon form 後 pivot 的個數為 n 。故 (3) \Leftrightarrow (4)。

最後，由 A 化為 echelon form 後 pivot 的個數為 n 知存在 $B, C \in \mathcal{M}_{n \times n}$ 使得 $BA = I_n$ 以及 $AC = I_n$ 。若能證得 $C = B$ ，則由 $BA = AB = I_n$ 得證 A 為 invertible。然而由 $BA = I_n$ ，得 $(BA)C = I_n C = C$ 。又由 $(BA)C = B(AC) = BI_n = B$ ，得證 $B = C$ 。故 (3) \Rightarrow (1)。得證本定理。□

我們在上一章曾提過當一個 $n \times n$ matrix 的 rank 為 n 時, 我們稱之為 nonsingular matrix. 所以由 Theorem 3.5.2 我們知 invertible matrix 就是 nonsingular matrix. 反之, non-invertible matrix 就是 singular matrix. 不過為了讓大家不被這麼多名詞弄混. 以後我們僅用 invertible 和 non-invertible 而不用 nonsingular 和 singular.

由 Theorem 3.5.2 的證明我們知 A 為 $n \times n$ square matrix, 若存在 $B, C \in \mathcal{M}_{n \times n}$ 使得 $BA = I_n$ 且 $AC = I_n$, 則 $B = C$. 我們自然會問有沒有可能存在不同的 $B, B' \in \mathcal{M}_{n \times n}$ 皆滿足 $BA = I_n$ 以及 $B'A = I_n$. 下一個定理告訴我們這樣的方陣其實是唯一的.

Corollary 3.5.3. 假設 A 為 $n \times n$ square matrix 且 $B, B' \in \mathcal{M}_{n \times n}$ 滿足 $BA = I_n$ 以及 $B'A = I_n$. 則 $B = B'$.

Proof. 由 Theorem 3.5.2 我們知 A 為 invertible 且由其證明知 $BA = AB = I_n$ 以及 $B'A = AB' = I_n$. 故

$$B = I_n B = (B'A)B = B'(AB) = B'I_n = B'.$$

□

由 Corollary 3.5.3, 我們知道若 A 為 $n \times n$ invertible matrix, 則僅會存在唯一的一個 $n \times n$ matrix B 滿足 $BA = AB = I_n$. 它和 A 的關係如同在實數上非零實數的乘法的 inverse (乘法反元素), 所以我們給以下的定義.

Definition 3.5.4. 假設 A 為 $n \times n$ invertible matrix. 我們稱唯一滿足 $BA = AB = I_n$ 的 $n \times n$ matrix B 為 A 的 inverse (反矩陣), 且用 A^{-1} 表示.

給定一 $n \times n$ invertible matrix A 由於其反矩陣是唯一的, 所以若要確定 $B = A^{-1}$ 我們僅要檢查是否 $BA = I_n$ 或 $AB = I_n$ 即可. 我們有以下之性質

Proposition 3.5.5. 假設 $A, B \in \mathcal{M}_{n \times n}$. 我們有以下之性質

(1) 若 A 為 invertible, 則 A^{-1} 亦為 invertible 且

$$(A^{-1})^{-1} = A.$$

(2) A 為 invertible 若且唯若 A^T 為 invertible 且此時

$$(A^T)^{-1} = (A^{-1})^T.$$

(3) A, B 皆為 invertible 若且唯若 AB 為 invertible. 且此時

$$(AB)^{-1} = B^{-1}A^{-1}.$$

Proof. 由 Theorem 3.5.2, 我們要說一個 $n \times n$ matrix 為 invertible, 只要找到 $B \in \mathcal{M}_{n \times n}$ 使得 $BA = I_n$ 或 $AB = I_n$ 且此時由唯一性 (Corollary 3.5.3) 知 $B = A^{-1}$.

(1) 依定義 A^{-1} 亦為 $n \times n$ matrix 故 Theorem 3.5.2 適用. 利用 $A^{-1}A = I_n$, 得知 A^{-1} 亦為 invertible 且 $(A^{-1})^{-1} = A$.

(2) 依定義 A^T 亦為 $n \times n$ matrix 故 Theorem 3.5.2 適用. 由 $A^{-1}A = I_n$ 利用 Proposition 3.2.4 得

$$I_n = (A^{-1}A)^T = A^T(A^{-1})^T$$

故知 A^T 為 invertible 且 $(A^T)^{-1} = (A^{-1})^T$. 反之若 A^T 為 invertible, 由前知 $(A^T)^T$ 為 invertible, 故由利用 Proposition 3.2.4 $(A^T)^T = A$ 得證 A 為 invertible.

(3) 依定義 AB 為 $n \times n$ matrix 故 Theorem 3.5.2 適用. 現若 A, B 皆為 invertible, 則由

$$(AB)B^{-1}A^{-1} = A(BB^{-1})A^{-1} = AI_nA^{-1} = AA^{-1} = I_n,$$

得證 AB 為 invertible 且 $(AB)^{-1} = B^{-1}A^{-1}$. 反之, 若 AB 為 invertible, 且令 $C = (AB)^{-1}$. 此時由 $(AB)C = I_n$ 得 $A(BC) = I_n$, 故由假設 A 為 $n \times n$ matrix 以及 Theorem 3.5.2 得證 A 為 invertible. 同理, 由 $C(AB) = I_n$, 得 $(CA)B = I_n$, 得證 B 為 invertible. \square

要注意 Proposition 3.5.5 (3) 中由 AB invertible 推得 A, B 皆為 invertible 是需要用到 A, B 皆為 $n \times n$ matrix. 否則當 $m \neq n$ 時, 在 Theorem 3.4.2 中我們知道有可能 $A \in \mathcal{M}_{m \times n}, C \in \mathcal{M}_{n \times m}$ 滿足 $AC = I_m$. 此時 I_m 為 invertible, 但 A, C 皆為 non-invertible. 同樣的, 當 A, B 為方陣時, 因為由 AB 為 invertible 可推得 A, B 皆為 invertible, 故知 BA 亦為 invertible. 也就是說當 A, B 為方陣時 AB 為 invertible 和 BA 為 invertible 是等價的. 但在 A, B 不為方陣時, 若 AB 為 invertible 會導致 BA 不為 invertible.

Question 3.10. 試舉例 A, B 不為 invertible 但 AB 為 invertible. 同時也驗證此時 BA 為 non-invertible.

接下來我們探討如何判別一個具體的 $n \times n$ matrix 是否為 invertible, 且若為 invertible 如何找出其 inverse. 這個問題可藉由將方陣利用 elementary row operations 化為 reduced echelon form 來處理. 事實上, 當 A 為 $n \times n$ matrix, 由 Theorem 3.5.2 我們知道 A 為 invertible 若且唯若 A 化為 echelon form 後其 pivot 的個數等於 n . 因此我們只要將 A 化為 echelon form 後計算其 pivot 的個數, 便可以知道 A 是否為 invertible. 若 A 為 invertible, 即 pivot 的個數為 n , 此時由於 A 的 reduced echelon form 為 $n \times n$ matrix, 故得 A 的 reduced echelon form 為 I_n . 也就是說我們可以用 elementary row operations 將 A 化為 I_n . 故由 Lemma 3.3.1 我們知存在 $E \in \mathcal{M}_{n \times n}$ 為一些 elementary matrix 的乘積使得 $EA = I_n$. 事實上若將 augmented matrix $[A|I_n]$ 利用 elementary row operations 化為 $[I_n|E]$, 則 $EA = I_n$, 故此時 E 就是 A^{-1} . 我們看以下的例子.

Example 3.5.6. 考慮矩陣

$$A = \begin{bmatrix} 1 & 0 & 1 & -1 \\ 0 & -1 & -3 & 4 \\ 1 & 0 & -1 & 2 \\ -3 & 0 & 0 & -1 \end{bmatrix}$$

我們要決定是否 A 是否為 invertible. 若為 invertible, 要找出 A^{-1} .

我們直接考慮 augmented matrix $[A|I_4]$, 利用 elementary row operation 將 A 的部分轉換成 echelon form. 首先將 1-st row 分別乘上 $-1, 3$ 加至 3-rd, 4-th row, 即

$$\left[\begin{array}{cccc|cccc} 1 & 0 & 1 & -1 & 1 & 0 & 0 & 0 \\ 0 & -1 & -3 & 4 & 0 & 1 & 0 & 0 \\ 1 & 0 & -1 & 2 & 0 & 0 & 1 & 0 \\ -3 & 0 & 0 & -1 & 0 & 0 & 0 & 1 \end{array} \right] \sim \left[\begin{array}{cccc|cccc} 1 & 0 & 1 & -1 & 1 & 0 & 0 & 0 \\ 0 & -1 & -3 & 4 & 0 & 1 & 0 & 0 \\ 0 & 0 & -2 & 3 & -1 & 0 & 1 & 0 \\ 0 & 0 & 3 & -4 & 3 & 0 & 0 & 1 \end{array} \right].$$

接著將 3-rd row 乘上 $3/2$ 加至 4-th row 得

$$\left[\begin{array}{cccc|cccc} 1 & 0 & 1 & -1 & 1 & 0 & 0 & 0 \\ 0 & -1 & -3 & 4 & 0 & 1 & 0 & 0 \\ 0 & 0 & -2 & 3 & -1 & 0 & 1 & 0 \\ 0 & 0 & 0 & \frac{1}{2} & \frac{3}{2} & 0 & \frac{3}{2} & 1 \end{array} \right].$$

此時 augmented matrix 左半部為 echelon form, 其 pivot 的個數為 4, 故知 A 為 invertible. 我們繼續將左半部化為 reduced echelon form 便可得到 A^{-1} .

先將 4-th row 乘以 2, 然後將所得的 augmented matrix 的 4-th row 分別乘上 $-3, -4, 1$ 加至 3-rd, 2-nd 和 1-st row, 即

$$\left[\begin{array}{cccc|cccc} 1 & 0 & 1 & -1 & 1 & 0 & 0 & 0 \\ 0 & -1 & -3 & 4 & 0 & 1 & 0 & 0 \\ 0 & 0 & -2 & 3 & -1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 3 & 0 & 3 & 2 \end{array} \right] \sim \left[\begin{array}{cccc|cccc} 1 & 0 & 1 & 0 & 4 & 0 & 3 & 2 \\ 0 & -1 & -3 & 0 & -12 & 1 & -12 & -8 \\ 0 & 0 & -2 & 0 & -10 & 0 & -8 & -6 \\ 0 & 0 & 0 & 1 & 3 & 0 & 3 & 2 \end{array} \right].$$

接著將 3-rd row 乘以 $-1/2$, 然後將所得的 augmented matrix 的 3-rd row 分別乘上 $3, -1$ 加至 2-nd 和 1-st row, 即

$$\left[\begin{array}{cccc|cccc} 1 & 0 & 1 & 0 & 4 & 0 & 3 & 2 \\ 0 & -1 & -3 & 0 & -12 & 1 & -12 & -8 \\ 0 & 0 & 1 & 0 & 5 & 0 & 4 & 3 \\ 0 & 0 & 0 & 1 & 3 & 0 & 3 & 2 \end{array} \right] \sim \left[\begin{array}{cccc|cccc} 1 & 0 & 0 & 0 & -1 & 0 & -1 & -1 \\ 0 & -1 & 0 & 0 & 3 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 & 5 & 0 & 4 & 3 \\ 0 & 0 & 0 & 1 & 3 & 0 & 3 & 2 \end{array} \right].$$

最後將 augmented matrix 的 2-nd row 乘上 -1 . 此時所得 augmented matrix 左半部為 reduced echelon form (即 I_4), 故其右半部為 A^{-1} , 即

$$A^{-1} = \begin{bmatrix} -1 & 0 & -1 & -1 \\ -3 & -1 & 0 & -1 \\ 5 & 0 & 4 & 3 \\ 3 & 0 & 3 & 2 \end{bmatrix}.$$

由前面討論我們知當 $A \in \mathcal{M}_{n \times n}$ 為 invertible, 則存在 elementary matrices E_1, \dots, E_k 使得 $(E_k \cdots E_1)A = I_n$. 亦即 $A^{-1} = E_k \cdots E_1$, 由 Proposition 3.5.5 (3), 我們知 E_1, \dots, E_k 皆為 invertible, 且由 $(A^{-1})^{-1} = A$, 得 $A = E_1^{-1} \cdots E_k^{-1}$. 事實上這些 elementary matrix E_i 的 inverse 就是將 E_i 還原成 I_n 的 elementary row operation 所對應的 elementary matrix. 也就是說 E_i^{-1} 亦為 elementary matrix. 因此我們有以下的定理.

Proposition 3.5.7. A 為 invertible matrix 若且唯若 A 為一些 elementary matrices 的乘積.

Example 3.5.8. 考慮矩陣

$$A = \begin{bmatrix} 0 & 2 & 0 \\ 1 & -1 & 0 \\ 0 & 2 & 1 \end{bmatrix}.$$

在求 A 的 inverse 的過程中, 首先我們將 1-st row 和 2-nd row 交換. 令此 elementary row operation 所對應的 elementary matrix 為 E_1 . 因用相同的 elementary row operation 可將 E_1 還原成 I_3 , 故 $E_1 = E_1^{-1}$, 即

$$\left[\begin{array}{ccc|ccc} 0 & 2 & 0 & 1 & 0 & 0 \\ 1 & -1 & 0 & 0 & 1 & 0 \\ 0 & 2 & 1 & 0 & 0 & 1 \end{array} \right] \sim \left[\begin{array}{ccc|ccc} 1 & -1 & 0 & 0 & 1 & 0 \\ 0 & 2 & 0 & 1 & 0 & 0 \\ 0 & 2 & 1 & 0 & 0 & 1 \end{array} \right], E_1 = E_1^{-1} = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

接著將 augmented matrix 的 2-nd row 乘上 -1 加至 3-rd row. 令此 elementary row operation 所對應的 elementary matrix 為 E_2 . 因將 2-nd row 乘上 1 加至 3-rd row 的 elementary row operation 可將 E_2 還原成 I_3 , 故所得的 augmented matrix 及 E_2, E_2^{-1} 分別為

$$\left[\begin{array}{ccc|ccc} 1 & -1 & 0 & 0 & 1 & 0 \\ 0 & 2 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & -1 & 0 & 1 \end{array} \right], E_2 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -1 & 1 \end{bmatrix}, E_2^{-1} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix}.$$

然後將 augmented matrix 的 2-nd row 乘上 $1/2$. 令此 elementary row operation 所對應的 elementary matrix 為 E_3 . 因將 2-nd row 乘上 2 的 elementary row operation 可將 E_3 還原成 I_3 , 故所得的 augmented matrix 及 E_3, E_3^{-1} 分別為

$$\left[\begin{array}{ccc|ccc} 1 & -1 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & \frac{1}{2} & 0 & 0 \\ 0 & 0 & 1 & -1 & 0 & 1 \end{array} \right], E_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{2} & 0 \\ 0 & 0 & 1 \end{bmatrix}, E_3^{-1} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

最後將 augmented matrix 的 2-nd row 乘上 1 加至 1-st row. 令此 elementary row operation 所對應的 elementary matrix 為 E_4 . 因將 2-nd row 乘上 -1 加至 3-rd row 的 elementary row operation 可將 E_4 還原成 I_3 , 故所得的 augmented matrix 及 E_4, E_4^{-1} 分別為

$$\left[\begin{array}{ccc|ccc} 1 & 0 & 0 & \frac{1}{2} & 1 & 0 \\ 0 & 1 & 0 & \frac{1}{2} & 0 & 0 \\ 0 & 0 & 1 & -1 & 0 & 1 \end{array} \right], E_4 = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, E_4^{-1} = \begin{bmatrix} 1 & -1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

我們檢查可得

$$A^{-1} = E_4 E_3 E_2 E_1 = \begin{bmatrix} \frac{1}{2} & 1 & 0 \\ \frac{1}{2} & 0 & 0 \\ -1 & 0 & 1 \end{bmatrix}, A = E_1^{-1} E_2^{-1} E_3^{-1} E_4^{-1} = \begin{bmatrix} 0 & 2 & 0 \\ 1 & -1 & 0 \\ 0 & 2 & 1 \end{bmatrix}.$$

當 A 為 $n \times n$ invertible matrix, Theorem 3.5.2 告訴我們 A 化為 echelon 後其 pivot 的個數為 n . 由於 A 為方陣, 其 row 的個數和 column 的個數皆為 n , 此時很自然地可以將 Theorem 3.4.2 和 Theorem 3.4.6 相連結得到 invertible matrix 與其所對應的聯立方程組相關的性質. 由於以下性質直接套用上述兩個定理就可推得, 我們就不再證明了.

Theorem 3.5.9. 假設 A 為 $n \times n$ square matrix, 令 $\mathbf{a}_1, \dots, \mathbf{a}_n \in \mathbb{R}^n$ 為 A 的 column vectors. 則下列是等價的.

- (1) A 為 invertible matrix.

- (2) $\text{Span}(\mathbf{a}_1, \dots, \mathbf{a}_n) = \mathbb{R}^n$.
- (3) 對於任意 $\mathbf{b} \in \mathbb{R}^n$, 聯立方程組 $A\mathbf{x} = \mathbf{b}$ 皆有解.
- (4) 聯立方程組 $A\mathbf{x} = \mathbf{0}$ 沒有 *nontrivial solution*.
- (5) 對於任意 $\mathbf{b} \in \mathbb{R}^n$, 聯立方程組 $A\mathbf{x} = \mathbf{b}$ 皆有解且解唯一.

設 A 為 $n \times n$ invertible matrix, 則對任意 $\mathbf{b} \in \mathbb{R}^n$, $\mathbf{x} = A^{-1}\mathbf{b}$ 為聯立方程組 $A\mathbf{x} = \mathbf{b}$ 的一組解. 這是因為此時 $A\mathbf{x} = A(A^{-1}\mathbf{b}) = (AA^{-1})\mathbf{b} = \mathbf{b}$. 又由 Theorem 3.5.9 知此時 $A\mathbf{x} = \mathbf{b}$ 的解唯一. 故 $\mathbf{x} = A^{-1}\mathbf{b}$ 為聯立方程組 $A\mathbf{x} = \mathbf{b}$ 唯一的一組解.

Example 3.5.10. 考慮聯立方程組

$$\begin{array}{rcccc} x_1 & & +x_3 & -x_4 & = & b_1 \\ & -x_2 & -3x_3 & +4x_4 & = & b_2 \\ x_1 & & -x_3 & +2x_4 & = & b_3 \\ -3x_1 & & & -x_4 & = & b_4 \end{array}$$

其中 b_1, b_2, b_3, b_4 為任意實數. 由於此時聯立方程組為 $A\mathbf{x} = \mathbf{b}$, 其中 A 為 Example 3.5.6 中的 4×4 matrix A 且 $\mathbf{b} = [b_1 \ b_2 \ b_3 \ b_4]^T$. 因 A 為 invertible, 故由 Theorem 3.5.9 知, 對任意實數 b_1, b_2, b_3, b_4 , 聯立方程組 $A\mathbf{x} = \mathbf{b}$ 必有解且其解唯一. 事實上此唯一解為

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = A^{-1}\mathbf{b} = \begin{bmatrix} -1 & 0 & -1 & -1 \\ -3 & -1 & 0 & -1 \\ 5 & 0 & 4 & 3 \\ 3 & 0 & 3 & 2 \end{bmatrix} \begin{bmatrix} b_1 \\ b_2 \\ b_3 \\ b_4 \end{bmatrix} = \begin{bmatrix} -b_1 - b_3 - b_4 \\ -3b_1 - b_2 - b_4 \\ 5b_1 + 4b_3 + 3b_4 \\ 3b_1 + 3b_3 + 2b_4 \end{bmatrix}.$$

3.6. 結論

在本章中我們定義了矩陣之間的運算且與解聯立方程組相連結. 我們進而發現 elementary row operations 也都可以看成矩陣的乘法. 因而得到聯立方程組解的存在與唯一性與矩陣乘法的特殊關係. 特別當一個矩陣為方陣時, 存在性和唯一性便合而為一. 我們推得一個係數矩陣為 invertible matrix 的聯立方程組一定有解且解唯一. 我們也了解可以利用 elementary row operations 來判斷一個方陣是否為 invertible, 並求出 invertible matrix 的 inverse.