

# Matrix

在第一章我們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。

## 2.1. 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  $M_{m \times n}$  來表示所有係數在  $\mathbb{R}$  的  $m \times n$  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。例如考慮

$$A = \begin{bmatrix} 1 & 0 & 2 & 3 \\ 0 & 1 & 5 & 8 \\ 2 & 1 & 1 & 0 \end{bmatrix}, \quad (2.1)$$

則  $A$  為一個  $3 \times 4$  matrix，即  $A \in M_{3 \times 4}$ 。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 (2.1) 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 (2.1) 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 vectors 和 column vectors 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。

**Definition 2.1.1.** 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  $M_{m \times n}$  的矩陣才談是否相等，且兩個  $M_{m \times n}$  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  $M_{m \times n}$  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 2.1.2.** 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 2.1.2 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

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

矩陣的加法與係數積的定義可以說是由大家熟悉的  $\mathbb{R}^2, \mathbb{R}^3$  的向量（甚至  $\mathbb{R}^n$ ）的加法與係數積延伸而來，我們有以下這些性質。

**Proposition 2.1.3.** 對於  $M_{m \times n}$  上的矩陣，我們有以下的性質：

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

Proposition 2.1.3 的證明用到實數  $\mathbb{R}$  相對應的性質。例如 (1) 談的是矩陣加法的交換性質，用到的是  $\mathbb{R}$  的加法交換性。事實上若令  $A = [a_{ij}]$ ,  $B = [b_{ij}]$  以及  $A + B = [c_{ij}]$ ,  $B + A = [d_{ij}]$ 。依定義，我們有  $c_{ij} = a_{ij} + b_{ij}$  以及  $d_{ij} = b_{ij} + a_{ij}$ 。由於實數的加法交換性  $a_{ij} + b_{ij} = b_{ij} + a_{ij}$ ，因此得  $c_{ij} = d_{ij}$ ，故依矩陣相等的定義，得證  $A + B = B + A$ 。其他各項都可用一樣的方法證明，這裡就不證明了。不過 (3), (4) 談的是存在性的問題。這類問題就必須說明存在的是什麼。例如  $O$  就是零矩陣，也就是說  $O = [c_{ij}]$  其中  $c_{ij} = 0, \forall ij$ 。而 (4) 的  $A'$  稱為  $A$  的加法反元素，它是跟著  $A$  而變的。如果  $A = [a_{ij}]$  則令  $A' = [a'_{ij}]$  其中  $a'_{ij} = -a_{ij}$ ，就會滿足  $A + A' = O$  了。一般來說，給定  $A$  我們會將  $A$  的加法反元素用  $-A$  來表示。

接著我們定義矩陣間的乘法。首先回顧當我們要解聯立方程組

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

我們會把它寫成

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

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 2.1.4.** 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

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

注意依此定義, 必需  $A \in M_{m \times n}$  的 column 的個數  $n$  等於  $\mathbf{b} \in M_{n \times 1}$  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

$$\mathbf{Ab} = b_1 \begin{bmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{bmatrix} + b_2 \begin{bmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{m2} \end{bmatrix} + \cdots + b_n \begin{bmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{mn} \end{bmatrix} = \begin{bmatrix} b_1 a_{11} + b_2 a_{12} + \cdots + b_n a_{1n} \\ b_1 a_{21} + b_2 a_{22} + \cdots + b_n a_{2n} \\ \vdots \\ b_1 a_{m1} + b_2 a_{m2} + \cdots + b_n a_{mn} \end{bmatrix}. \quad (2.2)$$

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

2.1.4 的矩陣乘法定義

$$\mathbf{ab} = [a_1 \ a_2 \ \cdots \ a_n] \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix} = b_1 a_1 + b_2 a_2 + \cdots + b_n a_n. \quad (2.3)$$

(注意, 若  $\mathbf{a}, \mathbf{b}$  是  $\mathbb{R}^n$  中的向量, 則  $\mathbf{ab}$  就是我們熟悉  $\mathbf{a}, \mathbf{b}$  的內積  $\langle \mathbf{a}, \mathbf{b} \rangle$ .) 依此看法, 由式子 (2.2), 我們可將  $\mathbf{Ab}$  寫成

$$\mathbf{Ab} = \begin{bmatrix} \mathbf{1ab} \\ \mathbf{2ab} \\ \vdots \\ \mathbf{mab} \end{bmatrix}. \quad (2.4)$$

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

我們來看一個  $m \times n$  matrix 以及  $\mathbb{R}^n$  的 column vector 在此乘法的定義之下的基本性質.

**Lemma 2.1.5.** 假設  $A = [a_{ij}]$ ,  $A' = [a'_{ij}]$  為  $m \times n$  matrices 以及  $\mathbf{b} = [b_j]$ ,  $\mathbf{b}' = [b'_j]$  為  $\mathbb{R}^n$  中的 column vectors (即  $n \times 1$  matrices) 以及  $c \in \mathbb{R}$ . 我們有以下的性質.

- (1)  $A(\mathbf{b} + \mathbf{b}') = \mathbf{Ab} + \mathbf{Ab}'$ .
- (2)  $A(c\mathbf{b}) = c(\mathbf{Ab}) = (cA)\mathbf{b}$ .
- (3)  $(A + A')\mathbf{b} = \mathbf{Ab} + A'\mathbf{b}$ .

**Proof.** 令  $A$  的 column vectors 依次為  $\mathbf{a}_1, \dots, \mathbf{a}_n$  且  $A'$  的 column vectors 依次為  $\mathbf{a}'_1, \dots, \mathbf{a}'_n$ .

(1) 依定義

$$A(\mathbf{b} + \mathbf{b}') = \left[ \begin{array}{c|c|c} | & & | \\ \mathbf{a}_1 & \cdots & \mathbf{a}_n \\ | & & | \end{array} \right] \begin{bmatrix} b_1 + b'_1 \\ \vdots \\ b_n + b'_n \end{bmatrix} = (b_1 + b'_1)\mathbf{a}_1 + \cdots + (b_n + b'_n)\mathbf{a}_n. \quad (2.5)$$

而  $A\mathbf{b} + A\mathbf{b}'$  為

$$\left[ \begin{array}{c|c|c} | & & | \\ \mathbf{a}_1 & \cdots & \mathbf{a}_n \\ | & & | \end{array} \right] \begin{bmatrix} b_1 \\ \vdots \\ b_n \end{bmatrix} + \left[ \begin{array}{c|c|c} | & & | \\ \mathbf{a}_1 & \cdots & \mathbf{a}_n \\ | & & | \end{array} \right] \begin{bmatrix} b'_1 \\ \vdots \\ b'_n \end{bmatrix} = (b_1\mathbf{a}_1 + \cdots + b_n\mathbf{a}_n) + (b'_1\mathbf{a}_1 + \cdots + b'_n\mathbf{a}_n). \quad (2.6)$$

由矩陣加法和係數積的分配律 (Proposition 2.1.3 的性質 (7),(8)), 我們得證式子 (2.5) 和式子 (2.6) 相等.

(2) 依定義  $c(A\mathbf{b})$  為

$$c\left(\left[ \begin{array}{c|c|c} | & & | \\ \mathbf{a}_1 & \cdots & \mathbf{a}_n \\ | & & | \end{array} \right] \begin{bmatrix} b_1 \\ \vdots \\ b_n \end{bmatrix}\right) = c(b_1\mathbf{a}_1 + b_2\mathbf{a}_2 + \cdots + b_n\mathbf{a}_n). \quad (2.7)$$

而  $cA$  的 column vectors 依次為  $c\mathbf{a}_1, \dots, c\mathbf{a}_n$ . 故  $(cA)\mathbf{b}$  為

$$\left[ \begin{array}{c|c|c} | & & | \\ c\mathbf{a}_1 & \cdots & c\mathbf{a}_n \\ | & & | \end{array} \right] \begin{bmatrix} b_1 \\ \vdots \\ b_n \end{bmatrix} = b_1(c\mathbf{a}_1) + \cdots + b_n(c\mathbf{a}_n). \quad (2.8)$$

最後依定義  $c\mathbf{b}$  為  $\begin{bmatrix} cb_1 \\ \vdots \\ cb_n \end{bmatrix}$ , 故  $A(c\mathbf{b})$  為

$$\left[ \begin{array}{c|c|c} | & & | \\ \mathbf{a}_1 & \cdots & \mathbf{a}_n \\ | & & | \end{array} \right] \begin{bmatrix} cb_1 \\ \vdots \\ cb_n \end{bmatrix} = (cb_1)\mathbf{a}_1 + \cdots + (cb_n)\mathbf{a}_n. \quad (2.9)$$

由矩陣加法和係數積的結合律 (Proposition 2.1.3 的性質 (6)), 我們得證 (2.7), (2.8), (2.9) 三個式子皆相等.

(3) 依定義  $A + A'$  的 column vectors 依次為  $\mathbf{a}_1 + \mathbf{a}'_1, \dots, \mathbf{a}_n + \mathbf{a}'_n$ , 所以

$$(A + A')\mathbf{b} = \left[ \begin{array}{c|c|c} | & & | \\ \mathbf{a}_1 + \mathbf{a}'_1 & \cdots & \mathbf{a}_n + \mathbf{a}'_n \\ | & & | \end{array} \right] \begin{bmatrix} b_1 \\ \vdots \\ b_n \end{bmatrix} = b_1(\mathbf{a}_1 + \mathbf{a}'_1) + \cdots + b_n(\mathbf{a}_n + \mathbf{a}'_n). \quad (2.10)$$

另一方面,  $A\mathbf{b} + A'\mathbf{b}$  為

$$\left[ \begin{array}{c|c|c} | & & | \\ \mathbf{a}_1 & \cdots & \mathbf{a}_n \\ | & & | \end{array} \right] \begin{bmatrix} b_1 \\ \vdots \\ b_n \end{bmatrix} + \left[ \begin{array}{c|c|c} | & & | \\ \mathbf{a}'_1 & \cdots & \mathbf{a}'_n \\ | & & | \end{array} \right] \begin{bmatrix} b_1 \\ \vdots \\ b_n \end{bmatrix} = (b_1\mathbf{a}_1 + \cdots + b_n\mathbf{a}_n) + (b_1\mathbf{a}'_1 + \cdots + b_n\mathbf{a}'_n). \quad (2.11)$$

再次由矩陣加法和係數積的分配律, 我們得證式子 (2.10) 和式子 (2.11) 相等.  $\square$

Lemma 2.1.5 (1),(2) 告訴我們矩陣對向量的乘法有類似分配律的性質, 這個性質很重要 (以後我們會再提到並稱之為 *linear* 的性質), 我們特別用以下定理表示.

**Proposition 2.1.6.** 假設  $A \in M_{m \times n}$ , 且  $\mathbf{b}, \mathbf{b}'$  為  $\mathbb{R}^n$  中的 column vectors, 以及  $c, c' \in \mathbb{R}$ . 則

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

**Proof.** 因  $c\mathbf{b}, c'\mathbf{b}'$  皆為  $\mathbb{R}^n$  中的 column vectors, 由 Lemma 2.1.5 (1) 知  $A(c\mathbf{b} + c'\mathbf{b}') = A(c\mathbf{b}) + A(c'\mathbf{b}')$ . 再由 Lemma 2.1.5 (2) 知  $A(c\mathbf{b}) = c(A\mathbf{b}), A(c'\mathbf{b}') = c'(A\mathbf{b}')$ , 故得證本定理.  $\square$

**Question 2.1.** 假設  $A \in M_{m \times n}$ , 且  $\mathbf{b}_1, \dots, \mathbf{b}_k$  為  $\mathbb{R}^n$  中的 column vectors,  $c_1, \dots, c_k \in \mathbb{R}$ . 試利用數學歸納法證明

$$A\left(\sum_{i=1}^k c_i \mathbf{b}_i\right) = \sum_{i=1}^k c_i (A\mathbf{b}_i).$$

**Question 2.2.** 假設  $A, A' \in M_{m \times n}$ , 且  $\mathbf{b}$  為  $\mathbb{R}^n$  中的 column vectors, 以及  $c, c' \in \mathbb{R}$ . 是否  $(cA + c'A')\mathbf{b} = c(A\mathbf{b}) + c'(A'\mathbf{b})$  ?

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  $\mathbf{b}_k \in M_{n \times 1}$ ,  $1 \leq k \leq l$ , 我們已定義了  $A\mathbf{b}_k$  為何, 現在我們定義  $AB$  為  $m \times l$  matrix, 其中  $AB$  的  $k$ -th column vector 為  $A\mathbf{b}_k$ . 我們大致上有以下的圖示.

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

由於  $A\mathbf{b}_k$  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 2.1.7.** 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

$$\mathbf{c}_k = A\mathbf{b}_k = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \begin{bmatrix} b_{1k} \\ b_{2k} \\ \vdots \\ b_{nk} \end{bmatrix} = b_{1k}\mathbf{a}_1 + b_{2k}\mathbf{a}_2 + \cdots + b_{nk}\mathbf{a}_n. \quad (2.12)$$

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  $A\mathbf{b}_k$ ) 從上往下算的第  $i$  個 entry. 由式子 (2.4) 我們知此即  $A$  的  $i$ -th row  $\mathbf{a}$  和  $B$  的  $k$ -th column  $\mathbf{b}_k$  看成向量後取內積. 換言之, 若  $AB = [c_{ik}]$ , 則  $AB$  的  $(i, k)$ -th entry  $c_{ik}$  為

$$c_{ik} = \mathbf{a}\mathbf{b}_k = a_{i1}b_{1k} + a_{i2}b_{2k} + \cdots + a_{in}b_{nk} = \sum_{j=1}^n a_{ij}b_{jk}. \quad (2.13)$$

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 2.1.8.** 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

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

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 2.3.** 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  $\mathbf{a} \in M_{1 \times n}(\mathbb{R})$ ,  $\mathbf{b} \in M_{n \times 1}(\mathbb{R})$ . 試問依矩陣乘法定義  $\mathbf{ba}$  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  $\langle \mathbf{b}, \mathbf{a} \rangle$  有關嗎?

大部分的書都會用式子 (2.13) 當成矩陣乘法的定義. 我們選用式子 (2.12) 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 (2.13) 利用 entry 來看方便多了. 例如我們有以下的性質.

**Proposition 2.1.9.** 假設  $A, A' \in M_{m \times n}$ ,  $B, B' \in M_{n \times l}$ . 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  $B$  的 column vectors 依次為  $\mathbf{b}_1, \dots, \mathbf{b}_l$  且  $B'$  的 column vectors 依次為  $\mathbf{b}'_1, \dots, \mathbf{b}'_l$ .

(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  $A(\mathbf{b}_k + \mathbf{b}'_k)$ . 另一方面,  $AB+AB'$  的  $k$ -th column 為  $AB$  的  $k$ -th column  $A\mathbf{b}_k$  加上  $AB'$  的  $k$ -th column  $A\mathbf{b}'_k$ , 因此  $AB+AB'$  的  $k$ -th column 為  $A\mathbf{b}_k + A\mathbf{b}'_k$ . 由 Lemma 2.1.5 (1), 我們得證它們相等.

(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, 即  $(A+A')\mathbf{b}_k$ . 另一方面,  $AB+A'B$  的  $k$ -th column 為  $AB$  的  $k$ -th column  $A\mathbf{b}_k$  加上  $A'B$  的  $k$ -th column  $A'\mathbf{b}_k$ , 因此  $AB+A'B$  的  $k$ -th column 為  $A\mathbf{b}_k + A'\mathbf{b}_k$ . 因此由 Lemma 2.1.5 (3), 我們得證它們相等.  $\square$

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

**Proposition 2.1.10.** 設  $c \in \mathbb{R}$ ,  $A \in M_{m \times n}$ ,  $B \in M_{n \times l}$ . 則

$$c(AB) = (cA)B = A(cB).$$

**Proof.** 假設  $B$  的 column vectors 依次為  $\mathbf{b}_1, \dots, \mathbf{b}_l$ . 首先注意  $c(AB)$ ,  $(cA)B$  和  $A(cB)$  皆為  $m \times l$  matrix, 我們僅要證明當  $1 \leq k \leq l$  時,  $c(AB)$ ,  $(cA)B$  和  $A(cB)$  的  $k$ -th column 皆相等.  $c(AB)$  的  $k$ -th column 為  $c$  乘上  $AB$  的  $k$ -th column, 故為  $c(\mathbf{Ab}_k)$ . 而  $(cA)B$  的  $k$ -th column 為  $cA$  右邊乘上  $B$  的  $k$ -th column, 故為  $(cA)\mathbf{b}_k$ . 最後由於  $cB$  的  $k$ -th column 為  $c\mathbf{b}_k$ , 故  $A(cB)$  的  $k$ -th column 為  $A(c\mathbf{b}_k)$ . 因此由 Lemma 2.1.5 (2), 我們得證它們皆相等.  $\square$

**Question 2.4.** 假設  $A, A' \in M_{m \times n}$ ,  $B, B' \in M_{n \times l}$  以及  $c, c' \in \mathbb{R}$ . 可證明  $(cA + c'A)B = cAB + c'A'B$  以及  $A(cB + c'B') = cAB + c'A'B'$  嗎?

由 Proposition 2.1.9 和 Proposition 2.1.10 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, 也可推得乘法具有結合律的性質 (即  $(AB)C = A(BC)$ ). 這裡要注意  $A, B, C$  的階數必須要有限制  $(AB)C$  和  $A(BC)$  才會有意義.

**Proposition 2.1.11.** 假設  $A \in M_{m \times n}$ ,  $B \in M_{n \times l}$ ,  $C \in M_{l \times k}$ , 則  $(AB)C = A(BC)$ .

**Proof.** 依定義  $AB \in M_{m \times l}$ , 故  $(AB)C \in M_{m \times k}$ . 而  $BC \in M_{n \times k}$ , 故  $A(BC) \in M_{m \times k}$  與  $(AB)C$  同階.

對於  $1 \leq j \leq k$ , 我們要證明  $(AB)C$  和  $A(BC)$  的  $j$ -th column 相等. 令  $\mathbf{c}_j$  為  $C$  的  $j$ -th column 依定義  $(AB)C$  的  $j$ -th column 為  $(AB)\mathbf{c}_j$ . 至於  $A(BC)$  的  $j$ -th column, 依定義為  $A$  右邊乘上  $(BC)$  的  $j$ -th column (即  $B\mathbf{c}_j$ ). 所以我們僅要說明  $(AB)\mathbf{c}_j = A(B\mathbf{c}_j)$ , 就可證得結合律.

由於  $\mathbf{c}_j$  只有一個 column, 為了方便考量, 我們將  $\mathbf{c}_j$  用單一足碼表達, 即令  $\mathbf{c}_j = \begin{bmatrix} c_1 \\ \vdots \\ c_l \end{bmatrix}$ .

現對任意  $i = 1, \dots, l$ , 令  $AB$  的  $i$ -th column 為  $\mathbf{p}_i$ , 則

$$(AB)\mathbf{c}_j = \begin{bmatrix} | & & | \\ \mathbf{p}_1 & \cdots & \mathbf{p}_l \\ | & & | \end{bmatrix} \begin{bmatrix} c_1 \\ \vdots \\ c_l \end{bmatrix} = c_1\mathbf{p}_1 + \cdots + c_l\mathbf{p}_l.$$

然而若  $\mathbf{b}_i$  為  $B$  的  $i$ -th column, 依定義  $\mathbf{p}_i$  為  $AB$  的  $i$ -th column 故得  $\mathbf{p}_i = A\mathbf{b}_i$ . 因此我們得

$$(AB)\mathbf{c}_j = c_1(A\mathbf{b}_1) + \cdots + c_l(A\mathbf{b}_l).$$

另一方面

$$A(B\mathbf{c}_j) = A \left( \begin{bmatrix} | & & | \\ \mathbf{b}_1 & \cdots & \mathbf{b}_l \\ | & & | \end{bmatrix} \begin{bmatrix} c_1 \\ \vdots \\ c_l \end{bmatrix} \right) = A(c_1\mathbf{b}_1 + \cdots + c_l\mathbf{b}_l).$$

注意這裡我們將  $\mathbf{b}_i$  視為  $\mathbb{R}^n$  中的 column vector, 故套用 Proposition 2.1.6 (或 Question 2.1) 可得  $A(c_1\mathbf{b}_1 + \cdots + c_l\mathbf{b}_l) = c_1(A\mathbf{b}_1) + \cdots + c_l(A\mathbf{b}_l)$ . 所以得證  $(AB)\mathbf{c}_j = A(B\mathbf{c}_j)$ .  $\square$



有了矩陣乘法的結合律 (Proposition 2.1.11), 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.

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  $A \in M_{2 \times 3}, B \in M_{3 \times 4}$  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  $A \in M_{2 \times 3}, B \in M_{3 \times 2}$  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 2.1.9 和 Proposition 2.1.10 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  $A \in M_{n \times n}$ , 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  $\mathbf{e}_i$ , 即  $\mathbb{R}^n$  的 column vector, 其中  $i$ -th entry 為 1, 其他 entry 為 0. 事實上  $\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n$  就是我們熟悉的  $\mathbb{R}^n$  的 standard basis (標準基底). 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  $A \in M_{m \times n}, B \in M_{n \times l}$  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 2.5.** 假設  $A \in M_{n \times n}$ , 是否  $(A - 2I_n)^2 = A^2 - 4A + 4I_n$  為對?

**Question 2.6.** 試證明  $I_n$  是唯一的  $n \times n$  滿足對任意  $A \in M_{m \times n}$  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  $A \in M_{m \times n}, B \in M_{n \times l}$  我們很容易驗證  $rA = A(rI_n), rB = (rI_n)B$ .

**Question 2.7.** 試利用 Proposition 2.1.10 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  $2 \times 2$  相乘可交換.

## 2.2. Transpose Operation

這一節中我們將介紹矩陣取 transpose (即轉置矩陣) 的概念, 即其相關性質. 最後利用它來探討如何從 row 的角度來看矩陣相乘.

對於一個  $m \times n$  matrix, 簡單來說其轉置矩陣就是將此矩陣的 row 與 column 的腳色互換, 也就是說將 row vectors 依序換成 column vectors. 我們有以下的定義.

**Definition 2.2.1.** 給定  $A \in M_{m \times n}$ . 定義  $A^t \in M_{n \times m}$ , 其中對於  $1 \leq i \leq m$ ,  $A^t$  的  $i$ -th column 就是將  $A$  的  $i$ -th row 寫成 column vector. 我們稱  $A^t$  為  $A$  的 *transpose*.

**Example 2.2.2.** 令

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

依定義  $A^t$  應為  $3 \times 2$  matrix. 其中  $A^t$  的第一個 column 為  $A$  的第一個 row  $[1 \ 2 \ 3]$  寫成

column vector, 即  $\begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$ . 同理  $A^t$  的第二個 column 為  $A$  的第二個 row  $[-1 \ -2 \ -3]$  寫成 column vector, 即  $\begin{bmatrix} -1 \\ -2 \\ -3 \end{bmatrix}$ . 故得

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

注意,  $A^t$  的 1-st, 2-nd 和 3-rd row 也恰為  $A$  的 1-st, 2-nd 和 3-rd column 寫成 row 而得.

由上面的例子我們看到, 當  $A \in M_{m \times n}$ , 對於  $1 \leq j \leq n$ ,  $A^t$  的  $j$ -th row 就是將  $A$  的  $j$ -th column 寫成 row vector. 事實上若將  $A$  寫成  $A = [a_{ij}]$ . 對於  $1 \leq i \leq m$ ,  $A^t$  的  $i$ -th column 就

是將  $A$  的  $i$ -th row  $[a_{i1} \ \cdots \ a_{in}]$  寫成 column vector  $\begin{bmatrix} a_{i1} \\ \vdots \\ a_{im} \end{bmatrix}$ . 因此  $A^t$  的  $(1, i)$ -th entry 就

是  $A$  的  $(i, 1)$ -th entry  $a_{i1}$ , 而  $A^t$  的  $(2, i)$ -th entry 就是  $A$  的  $(i, 2)$ -th entry  $a_{i2}$ . 依此類推我們可以得到對於  $1 \leq j \leq n$ ,  $A^t$  的  $(j, i)$ -th entry 就是  $A$  的  $(i, j)$ -th entry  $a_{ij}$ . 也就是說若我們將  $A^t$  寫成  $A^t = [a'_{sk}]$ , 則  $1 \leq s \leq n$ ,  $1 \leq k \leq m$ , 且  $a'_{sk} = a_{ks}$ . 因此對於  $1 \leq j \leq n$ ,  $A^t$  的  $j$ -th row  $[a'_{j1} \ \cdots \ a'_{jm}]$  即為  $[a_{1j} \ \cdots \ a_{mj}]$ , 恰為  $A$  的  $j$ -th column 寫成 row vector. 我們將以上的討論寫成以下的結論.

**Lemma 2.2.3.** 假設  $A = [a_{ij}] \in M_{m \times n}$  且  $A^t = [a'_{sk}] \in M_{n \times m}$ . 則對於  $1 \leq s \leq n$ ,  $1 \leq k \leq m$ ,  $a'_{sk} = a_{ks}$  且  $A^t$  的  $s$ -th row 就是將  $A$  的  $s$ -th column 寫成 row vector.

由 Lemma 2.2.3, 以後要談論  $A$  和  $A^t$  間的關係, 我們可以用 row 換成 column, column 換成 row 以及  $(i, j)$ -th entry 換成  $(j, i)$ -th entry 三種看法處理. 現在我們來看矩陣取 transpose 的基本性質.

**Proposition 2.2.4.** 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 依定義知就是  $A^t$  的  $i$ -th row 寫成 column vector, 而  $A^t$  的  $i$ -th row 依 Lemma 2.2.3 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 transpose 定義知它就是  $A$  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. 我們將這些矩陣分別用  $A = [a_{ij}]$ ,  $A^t = [a'_{ji}]$ ,  $C = [c_{sk}]$ ,  $C^t = [c'_{ks}]$  表示. 對於  $1 \leq t \leq l$ ,  $1 \leq i \leq m$ ,  $(AC)^t$  的  $(k, i)$ -th entry 為  $AC$  的  $(i, k)$ -th entry, 由式子 (2.13) 知應為

$$a_{i1}c_{1k} + a_{i2}c_{2k} + \cdots + a_{in}c_{nk}.$$

另一方面,  $C^t A^t$  的  $(k, i)$ -th entry 為

$$c'_{k1}a'_{1i} + c'_{k2}a'_{2i} \cdots + c'_{kn}a'_{ni}.$$

利用 Lemma 2.2.3 知此即為

$$c_{1k}a_{i1} + c_{2k}a_{i2} \cdots + c_{nk}a_{in}.$$

故得證  $(AC)^t = C^t A^t$ . □

**Question 2.8.** 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 2.2.5.** 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 2.2.4, 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. □

利用 Proposition 2.2.4, 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  $A \in M_{1 \times m}, B \in M_{m \times n}$ , 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  $(AB)^t = a_1 ({}_1\mathbf{b})^t + a_2 ({}_2\mathbf{b})^t + \cdots + a_m ({}_m\mathbf{b})^t$ , 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 2.2.4 將  $(AB)^t$  再取轉置還原得

$$AB = a_1 ({}_1\mathbf{b}) + a_2 ({}_2\mathbf{b}) + \cdots + a_m ({}_m\mathbf{b}).$$

也就是說

$$[a_1 \quad \cdots \quad a_m] \begin{bmatrix} b_{11} & \cdots & b_{1n} \\ \vdots & \cdots & \vdots \\ b_{m1} & \cdots & b_{mn} \end{bmatrix} = a_1 [b_{11} \quad \cdots \quad b_{1n}] + \cdots + a_m [b_{m1} \quad \cdots \quad b_{mn}] \quad (2.14)$$

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 2.2.4 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

$$AB = \begin{bmatrix} - & {}_1\mathbf{a} & - \\ - & {}_2\mathbf{a} & - \\ & \vdots & \\ - & {}_m\mathbf{a} & - \end{bmatrix} B = \begin{bmatrix} - & {}_1\mathbf{a}B & - \\ - & {}_2\mathbf{a}B & - \\ & \vdots & \\ - & {}_m\mathbf{a}B & - \end{bmatrix}. \quad (2.15)$$

結合式子 (2.14), 我們有以下之結果.

**Proposition 2.2.6.** 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

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

### 2.3. Elementary Matrix

前面提過，我們將聯立方乘式用矩陣  $A\mathbf{x} = \mathbf{b}$  來表示，是想利用矩陣的乘法來處理聯立方程式。事實上 elementary row operation 已可以看成是矩陣的乘法運算。首先考慮  $n \times n$  的單位矩陣  $I_n$  (即  $I_n$  的對角線位置皆為 1, 其他位置為 0). 若用  $i$ -th row 和  $j$ -th row 交換的 type 1 elementary row operation 將  $I_n$  轉換成矩陣  $E_1$ , 可得

$$E_1 = \begin{bmatrix} 1 & & & & & & & & \\ & \ddots & & & & & & & \\ & & 0 & & 1 & & & & \\ & & & \ddots & & & & & \\ & & & & & 1 & & & \\ & & & & & & \ddots & & \\ & & & & & & & & 1 \end{bmatrix}. \quad (2.16)$$

同樣的若使用 type 2 elementary row operation 將  $I_n$  的  $i$ -th row 乘上非零實數  $r$  轉換成矩陣  $E_2$ , 可得

$$E_2 = \begin{bmatrix} 1 & & & & & & & & \\ & \ddots & & & & & & & \\ & & 1 & & & & & & \\ & & & r & & & & & \\ & & & & 1 & & & & \\ & & & & & \ddots & & & \\ & & & & & & & & 1 \end{bmatrix}. \quad (2.17)$$

最後若使用 type 3 elementary row operation 將  $I_n$  的  $i$ -th row 乘上實數  $r$  加到  $I_m$  的  $j$ -th row 所得的矩陣為  $E_3$ , 可得

$$E_3 = \begin{bmatrix} 1 & & & & & & & & \\ & \ddots & & & & & & & \\ & & 1 & & & & & & \\ & & & \ddots & & & & & \\ & & & & r & & & & \\ & & & & & 1 & & & \\ & & & & & & \ddots & & \\ & & & & & & & & 1 \end{bmatrix}. \quad (2.18)$$

這樣的矩陣我們稱之為 *elementary matrix*. 而我們分別稱  $E_1, E_2, E_3$  為 *type 1, type 2* 以及 *type 3* 的 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. 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

$$\left[ 0 \quad \cdots \quad 1 \quad \cdots \quad 0 \right]$$

$\hat{i}$

即  $i$ -th entry 為 1, 其他 entry 皆為 0. 所以依 Proposition 2.2.6,  $I_m A$  的  $i$ -th row 為

$$\left[ 0 \quad \cdots \quad 1 \quad \cdots \quad 0 \right] A = 0_1 \mathbf{a} + \cdots + 1_i \mathbf{a} + \cdots + 0_m \mathbf{a} = i \mathbf{a},$$

(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 type 1 elementary row operation 將  $I_m$  轉換成矩陣  $E$ , 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 type 1 elementary row operation 變換所得的矩陣.

同樣的若將  $I_m$  的  $i$ -th row 乘上非零實數  $r$  所得的 type 2 elementary matrix 為  $E$ , 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 type 2 elementary row operation 變換所得的矩陣.

最後若將  $I_m$  的  $i$ -th row 乘上實數  $r$  加到  $I_m$  的  $j$ -th row 所得的 type 3 elementary matrix 為  $E$ , 則因  $E$  的  $j$ -th row 的  $i$ -th entry 為  $r$ ,  $j$ -th entry 為 1. 故由 Proposition 2.2.6,  $EA$  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 type 3 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相對應的 elementary matrix. (2.16), (2.17), (2.18) 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  $E_1 A$  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  $E_2(E_1 A)$  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  $E_2(E_1 A)$  寫成  $(E_2 E_1)A$ . 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 2.9.** 試找出那些同階的 elementary matrices 其相乘是不可以交換的.

**Example 2.3.1.** 考慮 Example 1.1.1 的情形.  $A$  是  $3 \times 4$  matrix 且  $B$  是將  $A$  的第一個 row 和第二個 row 交換所得. 考慮將  $I_3$  的第一個 row 和第二個 row 交換所得的 elementary

matrix

$$E_1 = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

可得

$$E_1 A = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 2 & 3 & 4 \\ 2 & 1 & -1 & 3 \\ 4 & 0 & 1 & 2 \end{bmatrix} = \begin{bmatrix} 2 & 1 & -1 & 3 \\ 1 & 2 & 3 & 4 \\ 4 & 0 & 1 & 2 \end{bmatrix} = B.$$

同樣的  $C$  是由  $B$  的第二個 row 乘上 2 所得, 所以我們考慮將  $I_3$  的第二個 row 乘上 2 所得的 elementary matrix

$$E_2 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

可得

$$E_2 B = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 2 & 1 & -1 & 3 \\ 1 & 2 & 3 & 4 \\ 4 & 0 & 1 & 2 \end{bmatrix} = \begin{bmatrix} 2 & 1 & -1 & 3 \\ 2 & 4 & 6 & 8 \\ 4 & 0 & 1 & 2 \end{bmatrix} = C.$$

最後  $D$  是將  $C$  的第三個 row 乘上  $-3$  加到第一個 row, 所以我們考慮將  $I_3$  的第三個 row 乘上  $-3$  加到第一個 row 所得的 elementary matrix

$$E_3 = \begin{bmatrix} 1 & 0 & -3 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

可得

$$E_3 C = \begin{bmatrix} 1 & 0 & -3 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 2 & 1 & -1 & 3 \\ 2 & 4 & 6 & 8 \\ 4 & 0 & 1 & 2 \end{bmatrix} = \begin{bmatrix} -10 & 1 & -4 & -3 \\ 2 & 4 & 6 & 8 \\ 4 & 0 & 1 & 2 \end{bmatrix} = D.$$

上面所提這種將一個矩陣做 elementary row operation 可視為將此矩陣左邊乘上其對應的 elementary matrix 的看法, 將來對我們探討矩陣的性質是很有幫助的. 這種看法的互換, 也時能讓我們得到有趣的結果. 例如前面提過一個矩陣經由一個 elementary row operation 轉變成另一個矩陣後, 我們可以再用相同 type 的 elementary row operation 將其轉換回原來的矩陣. 這個事實用 elementary matrices 的角度來看, 可以有以下的看法:

- (1) 設  $E_1$  是將  $I_m$  的  $i$ -th row 和  $j$ -th row 互換的 type 1 elementary matrix. 我們將  $E_1$  的  $i$ -th row 和  $j$ -th row 再互換就可轉換回 identity matrix  $I_m$ . 所以我們有  $E_1 E_1 = I_m$ .
- (2) 設  $E_2$  是將  $I_m$  的  $i$ -th row 乘上非零實數  $r$  的 type 2 elementary matrix. 我們將  $E_2$  的  $i$ -th row 乘上  $r^{-1}$  就可轉換回  $I_m$ . 所以若令  $E'_2$  為將  $I_m$  的  $i$ -th row 乘上  $r^{-1}$  的 type 2 elementary matrix, 我們有  $E'_2 E_2 = I_m$ . 同理可得  $E_2 E'_2 = I_m$ .
- (3) 設  $E_3$  是將  $I_m$  的  $i$ -th row 乘上實數  $r$  加到  $j$ -th row 所得的矩陣的 type 3 elementary matrix. 我們將  $E_3$  的  $i$ -th row 乘上  $-r$  再加到  $j$ -th row 就可轉換回  $I_m$ . 所以若令  $E'_3$  為將  $I_m$  的  $i$ -th row 乘上  $-r$  的 type 3 elementary matrix, 我們有  $E'_3 E_3 = I_m$ . 同理可得  $E_3 E'_3 = I_m$ .

我們知道當一個  $m \times m$  的矩陣  $A$  若可找到矩陣  $B$  使得  $BA = AB = I_m$ , 則稱  $A$  為一個 *invertible matrix* (可逆矩陣), 且  $B$  為  $A$  的 *inverse* (反矩陣). 從上面的探討我們有以下之結論.

**Proposition 2.3.2.** 假設  $E$  是一個 *elementary matrix*, 則  $E$  為 *invertible* 且  $E$  的 *inverse* 是和  $E$  相同 *type* 的 *elementary matrix*.

既然有所謂的 elementary row operations 當然也會有 elementary column operations. 它的概念只是將 row operation 對 row 的動作改為對 column 的動作. 我們將一個矩陣的  $i$ -th column 和  $j$ -th column 對調, 這一個動作及稱為 type 1 的 elementary column operation. 若將矩陣的  $i$ -th column 上的數皆乘上非零實數  $r$ , 則稱 type 2 的 elementary column operation. 至於 type 3 的 elementary column operation 就是把矩陣的  $i$ -th column 乘上  $r$  後加到其  $j$ -th column. 由於 column operations 並未用在解聯立方乘組的問題, 所以這裡我們僅約略介紹其相關的概念, 不再像前面依樣詳述. 事實上 column operations 的概念和 row operations 是相呼應的, 大家可以用前面探討的方式檢驗.

將 identity matrix  $I_m$  做 elementary column operation 後會得到甚麼樣的矩陣呢? 結果也會是前面提到的 elementary matrix (這也是 elementary matrix 沒有區分 row 和 column 的原因). 例如將  $I_m$  的  $i$ -th column 和  $j$ -th column 互換所得的矩陣就是將  $I_m$  的  $i$ -th row 和  $j$ -th row 互換的 type 1 elementary matrix. 而將  $I_m$  的  $i$ -th column 乘上非零實數  $r$  的矩陣, 就是將  $I_m$  的  $i$ -th row 乘上  $r$  的 type 2 elementary matrix. 不過要注意, 將  $I_m$  的  $i$ -th column 乘上實數  $r$  加到  $j$ -column 所得的矩陣不是將  $I_m$  的  $i$ -th row 乘上實數  $r$  加到  $j$ -th row 所得的 elementary matrix, 而是將  $I_m$  的  $j$ -th row 乘上實數  $r$  加到  $i$ -th row 所得的矩陣的 type 3 elementary matrix. 這一部分請務必檢驗, 就能了解其中原因.

既然一個 elementary matrix 同時可對應到 elementary row operation 也可對應到 elementary column operation, 那要如何區分呢? 別忘了, 矩陣的乘法是沒有交換性的. 前面我們知道, 當一個 elementary matrix  $E$  乘在一個矩陣  $A$  的左邊時, 所得的矩陣  $EA$  會是對  $A$  做  $E$  所對應的 elementary row operation. 而若將  $E$  乘在矩陣  $B$  的右邊, 則所得的矩陣  $BE$  會是對  $B$  做  $E$  所對應的 elementary column operation. 為了方便起見, 我們綜合成以下的結論.

**Theorem 2.3.3.** 假設  $A$  是一個  $m \times n$  matrix. 若  $E$  是對  $I_m$  做 elementary row operation 所得的 elementary matrix, 則  $EA$  就會是對  $A$  作相對應的 elementary row operation 所得的矩陣. 若  $E'$  是對  $I_n$  做 elementary column operation 所得的 elementary matrix, 則  $AE'$  就會是對  $A$  作相對應的 elementary column operation 所得的矩陣.

**Example 2.3.4.** 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

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

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

這裡我們再說明一下, 當  $A$  是一個  $m \times n$  matrix, 因為  $A$  有  $m$  個 row, 所以乘在左邊的 elementary matrix (對應到 elementary row operation) 必須是一個  $m$  階方陣. 同樣的, 因為  $A$  有  $n$  個 column, 所以乘在右邊的 elementary matrix (對應到 elementary column operation) 必須是一個  $n$  階方陣.

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  $A$  被轉換為  $E_1A$ . 現若對  $[A|I_m]$  做相同的 elementary row operation 的話, 所得的結果會是  $E_1[A|I_m]$ . 然而此時原先  $A$  的部分會變成  $E_1A$ , 而  $I_m$  的部分經同樣的 elementary row operation, 所以  $I_m$  這部分會是  $E_1I_m$ . 因此我們知

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

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 2.3.5.** 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 2.3.6.** 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  $3 \times 4$  matrix, 所以第一個 elementary row operation 所對應的 elementary matrix  $E_1$  就是將  $3 \times 3$  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

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