

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。AI 需要龐大的矩陣運算，尤其是矩陣乘法。我們將介紹四種處理矩陣乘法的方法。充分了解這些方法之間的關聯，對於有效處理矩陣乘法有很大的幫助。

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 3×4 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。這些分別稱為 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 $\mathbf{a}_i$ 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，稱為“係數積”。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 $1A = A$ 。

一般來說一個集合及其運算，若符合以上 8 個性質，便稱它形成一個“向量空間”（vector space）。所以 Proposition 2.1.3 告訴我們 $m \times n$ 矩陣利用其加法與係數積，會形成一個向量空間。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 $c_{ij} = 0, \forall i, j$ 。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

$$\mathbf{Ab} = \begin{bmatrix} {}_1\mathbf{ab} \\ {}_2\mathbf{ab} \\ \vdots \\ {}_m\mathbf{ab} \end{bmatrix}. \quad (2.4)$$

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。

我們來看 $m \times n$ 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) $(A + A')\mathbf{b} = \mathbf{Ab} + \mathbf{A}'\mathbf{b}$.

Proof. 令 A 的 column vectors 依序為 $\mathbf{a}_1, \dots, \mathbf{a}_n$ 且 A' 的 column vectors 依序為 $\mathbf{a}'_1, \dots, \mathbf{a}'_n$ 。

(1) 依定義

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

而 $A\mathbf{b} + A\mathbf{b}'$ 為

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

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

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

而 cA 的 column vectors 依序為 $c\mathbf{a}_1, \dots, c\mathbf{a}_n$ 。故 $(cA)\mathbf{b}$ 為

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

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

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

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 依序為 $\mathbf{a}_1 + \mathbf{a}'_1, \dots, \mathbf{a}_n + \mathbf{a}'_n$ ，所以

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

另一方面， $A\mathbf{b} + A'\mathbf{b}$ 為

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

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*。考慮 $c_1, \dots, c_k \in \mathbb{R}$ ，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。設 $A = [a_{ij}]$ 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。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 $\mathbf{a}_i$ 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

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

再強調一次，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

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

‡

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) 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 依序為 $\mathbf{b}_1, \dots, \mathbf{b}_l$ 且 B' 的 column vectors 依序為 $\mathbf{b}'_1, \dots, \mathbf{b}'_l$ 。

(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。□

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 依序為 $\mathbf{b}_1, \dots, \mathbf{b}_l$ 。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 $c(A\mathbf{b}_k)$ 。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), 得證它們皆相等。□

Question 2.3. 假設 $A, A' \in M_{m \times n}$, $B, B' \in M_{n \times l}$ 以及 $c, c' \in \mathbb{R}$ 。證明 $(cA + c'A')B = cAB + c'A'B$ 以及 $A(cB + c'B') = cAB + c'AB'$ 。

Exercise 2.1. 假設 $A = [a_{ij}] \in M_{3 \times 4}$, $B = [b_{ij}] \in M_{4 \times 3}$ 其中 $a_{ij} = i + j$, $b_{ij} = i \cdot j$ 。

- (1) 試完整寫下矩陣 A 和 B 。
- (2) 試利用 B 的 3rd column 計算 AB 的 3rd column 並依所求結果寫下 AB 的 (2,3)-th entry。
- (3) 直接利用高中內積方法計算 AB 的 (2,3)-th entry。

Exercise 2.2. 一個 $n \times n$ 的方陣 (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*。假設 A, B 皆為 $n \times n$ 的 diagonal matrix 且設 A, B 的 (i, i) -th entry 分別為 a_i, b_i 。

- (1) 試證明 $A + B$ 也是 diagonal matrix 且寫下 $A + B$ 的 (i, i) -th entry。
- (2) 試證明 AB 也是 diagonal matrix 且寫下 AB 的 (i, i) -th entry。並依此說明 A 和 B 是否為 commutative (即是否乘法可交換 $AB = BA$)?

Exercise 2.3. 假設 $A = [a_{ij}] \in M_{m \times n}$, $B = [b_{ij}] \in M_{n \times \ell}$ 且令 $AB = [c_{ij}]$ 。令 $\mathbf{b}_k, \mathbf{c}_k$ 分別表示 B 和 AB 的 k -th column。已知 $\mathbf{b}_4 = r\mathbf{b}_1 + s\mathbf{b}_2 + t\mathbf{b}_3$ 其中 $r, s, t \in \mathbb{R}$ 。證明 $\mathbf{c}_4 = r\mathbf{c}_1 + s\mathbf{c}_2 + t\mathbf{c}_3$ 。

Exercise 2.4. 令 A 為 4×6 matrix 且可利用 elementary row operation 化為 reduced echelon form B 。令 $\mathbf{a}_k, \mathbf{b}_k$ 分別表示 A, B 的 k -th column。

- (1) 假設 $\mathbf{a}_4 = 4\mathbf{a}_1 + 3\mathbf{a}_3$, 試說明 homogeneous system $\mathbf{Ax} = \mathbf{0}$ 有一組解為

$$(x_1, x_2, x_3, x_4, x_5, x_6) = (4, 0, 3, -1, 0, 0).$$

- (2) 證明 $\mathbf{a}_4 = 4\mathbf{a}_1 + 3\mathbf{a}_3$ 若且唯若 $\mathbf{b}_4 = 4\mathbf{b}_1 + 3\mathbf{b}_3$ 。(Hint: 利用 $\mathbf{Ax} = \mathbf{0}$ 和 $\mathbf{Bx} = \mathbf{0}$ 有相同的解集合)

- (3) 已知

$$\mathbf{a}_1 = \begin{bmatrix} 1 \\ -2 \\ -1 \\ 3 \end{bmatrix}; \quad \mathbf{a}_3 = \begin{bmatrix} -1 \\ 1 \\ 2 \\ -4 \end{bmatrix}; \quad \mathbf{a}_6 = \begin{bmatrix} 3 \\ -9 \\ 2 \\ 5 \end{bmatrix}; \quad B = \begin{bmatrix} 1 & -3 & 0 & 4 & 0 & 5 \\ 0 & 0 & 1 & 3 & 0 & 2 \\ 0 & 0 & 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

試寫下矩陣 A 。