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 $\mathbf{b}_1,\ldots,\mathbf{b}_l$. 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), 我們得證它們皆相等. □

Question 2.3. 假設 $A,A' \in M_{m\times n}$, $B,B' \in M_{n\times l}$ 以及 $c,c' \in \mathbb{R}$. 可證明 $(cA+c'A)B = cAB + c'A'B$ 以及 $A(cB+c'B') = cAB + c'AB'$ 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 $i = 1,\ldots,l$, 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

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

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 $(AB)\mathbf{c}_j=A(B\mathbf{c}_j)$. □

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 $A\in M_{2\times3}, B\in M_{3\times4}$ 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 $A\in M_{2\times3}, B\in M_{3\times2}$ 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 $\mathbf{e}_1,\mathbf{e}_2,\ldots,\mathbf{e}_n$ 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 $AI_n=A, I_nB=B$. 特別的, 當 A 為 $n\times n$ matrix, 我們有 $AI_n=I_nA=A$.

Question 2.4. 假設 $A\in M_{n\times n}$, 是否 $(A-2I_n)^2=A^2-4A+4I_n$ 為對?

Question 2.5. 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 2.6. 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 2×2 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 $A^{\mathrm{t}}\in M_{n\times m}$, 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 3×2 matrix. 其中 A^{t} 的第一個 column 為 A 的第一個 row $\begin{bmatrix}1 & 2 & 3\end{bmatrix}$ 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 $\begin{bmatrix}-1 & -2 & -3\end{bmatrix}$ 寫成 column vector, 即 $\begin{bmatrix}-1\\-2\\-3\end{bmatrix}$. 故得

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

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 $\begin{bmatrix}a_{i1} & \cdots & a_{in}\end{bmatrix}$ 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_{i1}. 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 $A^{\mathrm{t}}=[a'_{sk}]$, 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 $\begin{bmatrix}a'_{j1} & \cdots & a'_{jm}\end{bmatrix}$ 即為 $\begin{bmatrix}a_{1j} & \cdots & a_{mj}\end{bmatrix}$, 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 $A^{\mathrm{t}}=[a'_{sk}]\in M_{n\times m}$. 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) $(A^{\mathrm{t}})^{\mathrm{t}}=A$.

(2) $(A+B)^{\mathrm{t}}=A^{\mathrm{t}}+B^{\mathrm{t}}$.

(3) $(AC)^{\mathrm{t}}=C^{\mathrm{t}}A^{\mathrm{t}}$.

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

(1) 因 $(A^{\mathrm{t}})^{\mathrm{t}}$ 與 A 皆為 $m\times n$ matrix, 對於 $1\le i\le n$, 我們只要檢查 $(A^{\mathrm{t}})^{\mathrm{t}}$ 的 i-th column 就是 A 的 i-th column. 然而 $(A^{\mathrm{t}})^{\mathrm{t}}$ 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 $(A^{\mathrm{t}})^{\mathrm{t}}=A$.

(2) 因 $A^{\mathrm{t}}+B^{\mathrm{t}}$ 與 $(A+B)^{\mathrm{t}}$ 皆為 $n\times m$ matrix, 對於 $1\le i\le m$, 我們只要檢查 $A^{\mathrm{t}}+B^{\mathrm{t}}$ 的 i-th column 就是 $(A+B)^{\mathrm{t}}$ 的 i-th column. 依定義 $A^{\mathrm{t}}+B^{\mathrm{t}}$ 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, $(A+B)^{\mathrm{t}}$ 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 $(A+B)^{\mathrm{t}}=A^{\mathrm{t}}+B^{\mathrm{t}}$.

(3) 由於 $(AC)^{\mathrm{t}}$ 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 $A=[a_{ij}]$, $A^{\mathrm{t}}=[a'_{ji}]$, $C=[c_{sk}]$, $C^{\mathrm{t}}=[c'_{ks}]$ 表示. 對於 $1\le t\le l$, $1\le i\le m$, $(AC)^{\mathrm{t}}$ 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, $C^{\mathrm{t}}A^{\mathrm{t}}$ 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 $(AC)^{\mathrm{t}}=C^{\mathrm{t}}A^{\mathrm{t}}$. □

Question 2.7. *假設 A 為 $m\times n$ matrix, $r\in\mathbb{R}$. 試證明 $(rA)^{\mathrm{t}}=rA^{\mathrm{t}}$.*

一個 $n\times n$ square matrix, 若滿足 $A^{\mathrm{t}}=A$, 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.*

$$A+A^{\mathrm{t}},BB^{\mathrm{t}},B^{\mathrm{t}}B.$$

Proof. 由 Proposition 2.2.4, 我們有 $(A+A^{\mathrm{t}})^{\mathrm{t}}=A^{\mathrm{t}}+(A^{\mathrm{t}})^{\mathrm{t}}=A^{\mathrm{t}}+A$, 故知 $A+A^{\mathrm{t}}$ 為 symmetric matrix. 另一方面, $(BB^{\mathrm{t}})^{\mathrm{t}}=(B^{\mathrm{t}})^{\mathrm{t}}B^{\mathrm{t}}=BB^{\mathrm{t}}$, 故得 BB^{t} 為 symmetric matrix. 同理可得 $B^{\mathrm{t}}B$ 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

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

則由 $(AB)^{\mathrm{t}}=B^{\mathrm{t}}A^{\mathrm{t}}$, 以及矩陣右邊乘 column vector 的定義得

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

亦即 $(AB)^{\mathrm{t}}=a_1\,({}_1\mathbf{b})^{\mathrm{t}}+a_2\,({}_2\mathbf{b})^{\mathrm{t}}+\cdots+a_m\,({}_m\mathbf{b})^{\mathrm{t}}$, 這裡 $({}_i\mathbf{b})^{\mathrm{t}}$ 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 $(AB)^{\mathrm{t}}$ 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

$$\begin{bmatrix}a_1 & \cdots & a_m\end{bmatrix}\begin{bmatrix}b_{11} & \cdots & b_{1n}\\ \vdots & \cdots & \vdots\\ b_{m1} & \cdots & b_{mn}\end{bmatrix}=a_1\begin{bmatrix}b_{11} & \cdots & b_{1n}\end{bmatrix}+\cdots+a_m\begin{bmatrix}b_{m1} & \cdots & b_{mn}\end{bmatrix} \tag{2.14}$$

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 $(AB)^{\mathrm{t}}=B^{\mathrm{t}}A^{\mathrm{t}}$. 依定義 $B^{\mathrm{t}}A^{\mathrm{t}}$ 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 $({}_i\mathbf{a})^{\mathrm{t}}$. 也就是說 $(AB)^{\mathrm{t}}$ 的 i-th column 為 $B^{\mathrm{t}}({}_i\mathbf{a})^{\mathrm{t}}$. 利用 Proposition 2.2.4 再取轉置還原得, AB 的 i-th row 為

$$(B^{\mathrm{t}}({}_i\mathbf{a})^{\mathrm{t}})^{\mathrm{t}}=(({}_i\mathbf{a})^{\mathrm{t}})^{\mathrm{t}}(B^{\mathrm{t}})^{\mathrm{t}}={}_i\,\mathbf{a}B.$$

換言之, 我們有以下的圖示

$$AB=\begin{bmatrix}— & {}_1\mathbf{a} & —\\ — & {}_2\mathbf{a} & —\\ & \vdots & \\ — & {}_m\mathbf{a} & —\end{bmatrix}B=\begin{bmatrix}— & {}_1\mathbf{a}B & —\\ — & {}_2\mathbf{a}B & —\\ & \vdots & \\ — & {}_m\mathbf{a}B & —\end{bmatrix}. \tag{2.15}$$

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 $1\le i\le m$, *AB* 的 *i-th row* 為

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

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

$$ E_1=\begin{bmatrix}1 &&&&&&\\ &\ddots&&&&&\\ &&0&&1&&\\ &&&\ddots&&&\\ &&1&&0&&\\ &&&&&\ddots&\\ &&&&&&1\end{bmatrix}. \tag{2.16} $$

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

$$ E_2=\begin{bmatrix}1 &&&&&&\\ &\ddots&&&&&\\ &&1&&&&\\ &&&r&&&\\ &&&&1&&\\ &&&&&\ddots&\\ &&&&&&1\end{bmatrix}. \tag{2.17} $$

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

$$ E_3=\begin{bmatrix}1 &&&&&&\\ &\ddots&&&&&\\ &&1&&&&\\ &&&\ddots&&&\\ &&r&&1&&\\ &&&&&\ddots&\\ &&&&&&1\end{bmatrix}. \tag{2.18} $$

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

$$\begin{bmatrix} 0 & \cdots & \underset{\hat{i}}{1} & \cdots & 0 \end{bmatrix}$$

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

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

(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 $I_mA=A$. 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_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 2.8. 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 3×4 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

$$E_1A=\begin{bmatrix}0&1&0\\1&0&0\\0&0&1\end{bmatrix}\begin{bmatrix}1&2&3&4\\2&1&-1&3\\4&0&1&2\end{bmatrix}=\begin{bmatrix}2&1&-1&3\\1&2&3&4\\4&0&1&2\end{bmatrix}=B.$$

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

$$E_2B=\begin{bmatrix}1&0&0\\0&2&0\\0&0&1\end{bmatrix}\begin{bmatrix}2&1&-1&3\\1&2&3&4\\4&0&1&2\end{bmatrix}=\begin{bmatrix}2&1&-1&3\\2&4&6&8\\4&0&1&2\end{bmatrix}=C.$$

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

$$E_3C=\begin{bmatrix}1&0&-3\\0&1&0\\0&0&1\end{bmatrix}\begin{bmatrix}2&1&-1&3\\2&4&6&8\\4&0&1&2\end{bmatrix}=\begin{bmatrix}-10&1&-4&-3\\2&4&6&8\\4&0&1&2\end{bmatrix}=D.$$

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 $E_1E_1=I_m$.

(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 $E_2E_2'=I_m$.

(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 $E_3E_3' = I_m$.

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

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

$$AE_1 = \begin{bmatrix} 1 & 2 & 3 \\ -1 & -2 & -3 \\ 11 & 22 & 33 \end{bmatrix} \begin{bmatrix} 1 & 0 & 10 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 2 & 13 \\ -1 & -2 & -13 \\ 11 & 22 & 143 \end{bmatrix}.$$

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.