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 3×4 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

$$E_5E_4E_3E_2E_1=\begin{bmatrix} -\frac{1}{2} & 2 & 0 \\ \frac{1}{2} & -1 & 0 \\ 0 & -4 & 1 \end{bmatrix}=E.$$

Question 2.9. 假設 A 為 $m\times n$ *matrix.* 若要記錄對 A 所做的 *column operations* 所對應的 *elementary matrices* 的乘積，增廣矩陣應如何擺放？

2.4. Matrix 和 System of Linear Equations 的連結

我們曾經利用 elementary row operations 將增廣矩陣化為 echelon form 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. 另外回顧, 給定一次聯立方程組

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

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 $A\mathbf{x}=\mathbf{b}$ 來表示. 現若 $x_1=c_1,x_2=c_2,\ldots,x_n=c_n$, 為此聯立方程組的一組解, 我們便會用

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

來表示這一組解, 而說 $\mathbf{x}=\mathbf{c}\in\mathbb{R}^n$ 為 $A\mathbf{x}=\mathbf{b}$ 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 $A\mathbf{c}=\mathbf{b}$.

2.4.1. 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 $A\mathbf{x}=\mathbf{b}$ 皆有解.

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

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

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

Lemma 2.4.1. 假設 $A\in M_{m\times n}$ 且 $\mathbf{b}\in\mathbb{R}^m$. 則 $A\mathbf{x}=\mathbf{b}$ 有解若且唯若 $\mathbf{b}\in\mathrm{Span}(\mathbf{a}_1,\ldots,\mathbf{a}_n)$, 其中 $\mathbf{a}_1,\ldots,\mathbf{a}_n$ 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 $A\mathbf{x}=\mathbf{b}$ 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 $\mathrm{Span}(\mathbf{a}_1,\ldots,\mathbf{a}_n)\subseteq\mathbb{R}^m$. 然而由 Lemma 2.4.1 知, 若對於任意 $\mathbf{b}\in\mathbb{R}^m$ 皆會使得 $A\mathbf{x}=\mathbf{b}$ 有解, 表示對任意 $\mathbf{b}\in\mathbb{R}^m$ 皆有 $\mathbf{b}\in\mathrm{Span}(\mathbf{a}_1,\ldots,\mathbf{a}_n)$. 故知此時 $\mathrm{Span}(\mathbf{a}_1,\ldots,\mathbf{a}_n)=\mathbb{R}^m$. 反之, 若 $\mathrm{Span}(\mathbf{a}_1,\ldots,\mathbf{a}_n)=\mathbb{R}^m$, 表示對任意 $\mathbf{b}\in\mathbb{R}^m$ 皆有 $\mathbf{b}\in\mathrm{Span}(\mathbf{a}_1,\ldots,\mathbf{a}_n)$. 同樣由 Lemma 2.4.1 知此即對於任意 $\mathbf{b}\in\mathbb{R}^m$ 皆會使得 $A\mathbf{x}=\mathbf{b}$ 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 $A\mathbf{x}=\mathbf{b}$ 皆有解和 $\mathrm{Span}(\mathbf{a}_1,\ldots,\mathbf{a}_n)=\mathbb{R}^m$ 是等價的.

另外我們可以考慮聯立方程組 $A\mathbf{x}=\mathbf{e}_i$, 其中 $\mathbf{e}_1,\ldots,\mathbf{e}_m\in\mathbb{R}^m$ 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 $A\mathbf{x}=\mathbf{b}$ 皆有解, 則對所有的 $i=1,\ldots,m$, 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 $A\mathbf{x}=\mathbf{e}_i$ 的一組解. 也就是說對所有的 $i=1,\ldots,m$ 皆有 $A\mathbf{c}_i=\mathbf{e}_i$. 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

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

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.

我們曾探討過, 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 1.2 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 $\mathrm{rank}(A)=m$) 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 2.4.2. *假設 $A \in M_{m\times n}$, 令 $\mathbf{a}_1,\ldots,\mathbf{a}_n \in \mathbb{R}^m$ 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) $\mathrm{Span}(\mathbf{a}_1,\ldots,\mathbf{a}_n) = \mathbb{R}^m$.

(3) $\mathrm{rank}(A) = m$.

(4) *存在 $n \times m$ matrix C 使得 $AC = I_m$.*

特別提醒一下, Theorem 2.4.2, 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 2.4.2 並不適用 (不過 Lemma 2.4.1 是適用的).

我們曾提及, 當 $A \in M_{m\times n}$, 將 A 化為 echelon form 後其 pivot 的個數不可能多於 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 2.4.2 中的情況不可能發生. 我們有以下的結論.

Corollary 2.4.3. *假設 $A \in M_{m\times n}$, 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 $\mathrm{rank}(A) < m$. 故由 Theorem 2.4.2 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 2.4.2 知不會存在 $n \times m$ matrix C 使得 $AC = I_m$. □

Question 2.10. *假設 $A \in M_{m\times n}$, 其中 $m < n$. 是否存在 $n \times m$ matrix C 使得 $CA = I_n$?*

前面提過 Theorem 2.4.2 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 2.4.3 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.