Chapter 2

Systems of Linear Equations

這一章要探討的是多元一次的聯立方程組. 我們依然利用大家熟悉的加減消去法 (或高斯消去法) 來處理這類方程組. 不過我們不再只關心如何解特定的聯立方程組, 而會更著重於有系統地探討一般聯立方程組解的情況的理論. 我們會用矩陣來表示一個聯立方程組, 不過這裡的矩陣僅是為了方便起見而使用, 不會涉及矩陣的性質. 至於真正矩陣的運算及性質, 我們留待下一章再詳述.

2.1. 解一次聯立方程組

所謂 n 元一次的方程式就是有 n 個未知數 (variable) 的一次方程式 (linear equation). 例如 $2x_1+5x_2-x_3+x_4=1$ 就是一個 4 元一次的聯立方程組 (當然也可看成是 5 元或更高元). n 元一次的方程式抽象的表示法就是

$$a_1x_2+\cdots+a_nx_n=b,$$

其中這些 $a_1,\ldots,a_n$ 和 b 都是實數, 而這些 x_i 表未知數. 當我們有多個 n 元一次的方程式要討論它們的共同解時, 就稱為解一次聯立方程組 (system of linear equations). 一般抽象的表示法

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

表示有 m 個 n 元一次方程式所成的方程組. 這裡 $a_{11}x_1+a_{12}x_2+\cdots+a_{1n}x_n=b_1$ 表示第一個方程式, $a_{21}x_1+a_{22}x_2+\cdots+a_{2n}x_n=b_2$ 表示第二個方程式, 而當 $1\leq i\leq m$ 時, 第 i 個方程式就是 $a_{i1}x_1+a_{i2}x_2+\cdots+a_{in}x_n=b_i$, 所以最後一個 (即第 m 個) 方程式就是 $a_{m1}x_1+a_{m2}x_2+\cdots+a_{mn}x_n=b_m$. 這裡 a_{ij},b_i 皆為實數, 這些實數才是真正影響到聯立方程組

的因素, 所以我們也可特別把它們標明出來, 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. 矩陣 A 中的每一個 a_{ij} 稱為 A 的一個 *entry*. 因為 A 的每一個 entry 對應到聯立方程組中某個未知數的係數, 通常我們會稱矩陣 A 為此聯立方程式的係數矩陣. 一個矩陣的一個橫排稱為一個 row (列), 而一個豎排稱為一個 column (行). 我們算 row 時是從上而下來數的, 也就是說最上面的一個 row 稱為第一個 row, 下一個 row 稱為第二個 row, 依此類推. 而算 column 是由左而右來數的, 也就是說最左邊的一個 column 稱為第一個 column, 再往右一個 column 稱為第二個 column, 依此類推. 大家可以看出矩陣 A 的 row 對應的就是此聯立方程組的方程式, 第一個 row 對應到第一個方程式, 第二個 row 對應到第二個方程式, 依此類推. 而 column 對應到的是方程組的未知數, 第一個 column 對應到的是未知數 x_1 的係數, 第二個 column 對應到的是未知數 x_2 的係數, 依此類推. 因為這裡是由 m 個方程式而且每個方程式有 n 個未知數所組成的聯立方程組, 所以 A 共有 m 個 row 以及 n 個 column, 我們稱這樣的矩陣為 $m\times n$ matrix. 注意這裡 $\mathbf{x}$ 表示是一個未知的向量而且我們將向量 $\mathbf{x},\mathbf{b}$ 都寫成 column vector (行向量) 是為了配合將來矩陣乘法的寫法. 目前大家只要記住這也是聯立方程式的一種表示法即可.

例如解聯立方程組

$$\begin{array}{rcl} 3x_1-2x_2+9x_4 & = & 4 \\ 2x_1+2x_2-4x_4 & = & 6 \end{array} \tag{2.1}$$

我們就可以表成

$$\begin{bmatrix} 3 & -2 & 0 & 9 \\ 2 & 2 & 0 & -4 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix}=\begin{bmatrix} 4 \\ 6 \end{bmatrix}$$

注意這裡係數矩陣多出 $\begin{bmatrix} 0 \\ 0 \end{bmatrix}$ 這個 column 因為 x_3 的係數為 0.

為何要探討解多元一次聯立方程組呢? 事實上解多元一次聯立方程式和線性代數的許多問題息息相關. 例如在 1.3 節中我們提到 span 的概念. 若 $\mathbf{u}=(1,-1,2,2),\mathbf{v}=(3,1,-1,2)$ 我們要問 $\mathbf{w}=(1,0,1,0)$ 是否在 $\mathrm{Span}(\mathbf{u},\mathbf{v})$ 中就等同於問是否存在 $c_1,c_2\in\mathbb{R}$ 使得 $\mathbf{w}=c_1\mathbf{u}+c_2\mathbf{v}$. 利用向量相等的定義, 這表示 $(1,0,1,0)=c_1(1,-1,2,2)+c_2(3,1,-1,2)$. 亦即要解

$$\begin{array}{rcrcl} x_1 & + & 3x_2 & = & 1 \\ -x_1 & + & x_2 & = & 0 \\ 2x_1 & - & x_2 & = & 1 \\ 2x_1 & + & 2x_2 & = & 0 \end{array}$$

這一聯立方程組. 又如要問哪些向量 $\mathbf{x}=(x_1,x_2,x_3,x_4)$ 會同時滿足 $\mathbf{u}\cdot\mathbf{x}=0$ 且 $\mathbf{v}\cdot\mathbf{x}=0$, 就等同於要解聯立方程組

$$\begin{array}{rcrcrcrcl} x_1 & - & x_2 & + & 2x_3 & + & 2x_4 & = & 0 \\ 3x_1 & + & x_2 & - & x_3 & + & 2x_4 & = & 0. \end{array}$$

將來我們還會碰到許多和解聯立方程組有關的問題, 這裡我們就不再多談而將重點放在如何解一個多元一次聯立方程組.

過去學習解一次聯立方程組的方法不外加減消去法或高斯消去法, 它們的原理都是一樣的, 即利用以下三種基本方法:

(1) 變換式子的順序

(2) 將某一式乘上一非零實數

(3) 將某一式乘上一實數後加到另一式上

利用這三種基本方法將方程式的某些變數消去, 最後求出解來. 當然這裡有些問題是要探討的: 第一就是要消到甚麼地步才可確認可求出解來? 第二就是為什麼經過這些過程所得的解就會是原方程組的解? 在本節中我們將先介紹一個較有系統的方法解聯立方程組的步驟, 讓大家知道何時就可確認此方程組有解或無解, 且有解時如何求解. 下一節我們再說明為何每一個聯立方程組都可以利用這個方式找到其解集合.

當我們要解

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

這一個聯立方程組時, 先寫出如下的 augmented matrix (增廣矩陣)

$$\left[\begin{array}{cccc|c} a_{11} & a_{12} & \cdots & a_{1n} & b_1 \\ a_{21} & a_{22} & \cdots & a_{2n} & b_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} & b_m \end{array}\right]$$

例如式子 (2.1) 中的聯立方程組所對應的 augmented matrix 為

$$\left[\begin{array}{cccc|c} 3 & -2 & 0 & 9 & 4 \\ 2 & 2 & 0 & -4 & 6 \end{array}\right]$$

換言之, 若我們要解 $A\mathbf{x} = \mathbf{b}$ 這一個聯立方程組, 就要寫下 $[A \mid \mathbf{b}]$ 這一個 matrix. 反之一個 augmented matrix $[A \mid \mathbf{b}]$ 就對應到一個聯立方程組 $A\mathbf{x} = \mathbf{b}$.

接下來我們將如加減消去法的三種步驟, 利用所謂的 elementary row operation (基本列運算) 處理這個 augmented matrix. 所謂 elementary row operation 即表示對矩陣進行如下三種的列運算:

(1) 將矩陣的某兩個 row 對調

(2) 將矩陣的某一個 row 乘上一非零實數

(3) 將矩陣的某一個 row 乘上一實數後加到另一個 row.

大家應很容易看出一個 augmented matrix 經過以上這三種列運算後所得的 augmented matrix 所對應的聯立方程組就是前面所提加減消去法的三種步驟所得的方程組. 我們的目

的就是要將 augmented matrix $[A\,|\,\mathbf{b}]$ 中的係數矩陣 A 利用這三種 elementary row operation 化成所謂的 *echelon form*.

我們先解釋一下何謂 echelon form. 首先我們將矩陣每一個 row 從左到右來看第一個不為 0 的項稱為這個 row 的 *leading entry*. 因為係數矩陣中的每一個 entry 對應到聯立方程組中某個 variable 的係數, 所以 leading entry 若是 variable x_i 的係數, 我們就說這個 leading entry 發生在 x_i 的位置. 要注意, 這也等同於這個 leading entry 是位於從左到右算來第 i 個 column. 例如矩陣

$$\begin{bmatrix} 1 & 2 & 1 & 1 & 4 \\ 0 & 0 & 5 & 0 & 2 \\ 0 & 0 & 1 & -1 & 1 \end{bmatrix}$$

第一個 row 的 leading entry 為 1 不過因為第一個 row 還有其他位置 1, 所以我們特別要說明第一個 row 的 leading entry 發生在 x_1 的位置, 而第二個 row 和第三個 row 的 leading entry 分別為 5 和 1 且發生的位置皆在 x_3.

所謂一個矩陣是 echelon form 表示這個矩陣沒有 leading entry 的 row (即該 row 每一項皆為 0) 必需在最下方, 而有 leading entry 的 row 其 leading entry 所在位置從上到下來看是往右移的. 換言之, 若上一個 row 的 leading entry 所在的位置是 x_i, 而下一個 row 的 lading entry 是 x_j, 則必需 $i < j$. 例如上一個矩陣並非 echelon form, 因為第 3 個 row 和第 2 個 row 的 leading entry 的位置皆為 x_3, 並未右移. 另外矩陣

$$\begin{bmatrix} 1 & 2 & -1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 3 & 0 \end{bmatrix}, \quad \begin{bmatrix} 0 & 1 & 1 & 2 \\ 0 & 0 & 2 & -1 \\ 3 & 0 & 0 & 0 \end{bmatrix}$$

都不是 echelon form, 因為前一個矩陣全為 0 的 row 並未置於最下方, 而後一個矩陣第 3 個 row 的 leading entry 在第 2 個 row 的 leading entry 的左方. 至於矩陣

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

就是 echelon form. 當一個矩陣是 echelon form 時, 我們稱每一個 row 的 leading entry 為 *pivot*, 而 pivot 所在的位置我們稱為 *pivot variable.*

當我們將 augmented matrix $[A\,|\,\mathbf{b}]$ 利用 elementary row operation 將之化成 $[A'\,|\,\mathbf{b}']$ 且 A' 為 echelon form 後. A' 有兩種情形. 一種情形為 A' 每一個 row 皆不全為 0; 另一種為 A' 有些 row 全為 0. 我們分別依這兩種情形來討論聯立方程組的解.

(1) A' 每一個 row 皆不全為 0: 此時聯立方程組為 *consistent*, 即一定有解. 我們又可細分成兩種情況.

   (a) 第一種情況是每一個變數 (variable) x_i 皆為 pivot variable. 亦即 pivot 的個數等於方程組未知數的個數 (即係數矩陣 A 的 column 個數). 例如

$$\left[\begin{array}{ccc|c} 2 & 1 & 1 & 4 \\ 0 & 3 & 1 & 2 \\ 0 & 0 & -1 & 1 \end{array}\right]$$

此時 echelon form 的 pivot variable 分別為 x_1,x_2,x_3 恰就是聯立方程組的未知數 x_1,x_2,x_3. 在這種情況之下此聯立方程組會有唯一解, 而且我們可利用從下往上 "代回" 的方式求得解. 例如前面的 augmented matrix 所對應的聯立方程組為

$$\begin{array}{rrrcr} 2x_1 & +x_2 & +x_3 & = & 4 \\ & 3x_2 & +x_3 & = & 2 \\ & & -x_3 & = & 1 \end{array}$$

所以我們從最下面的 $-x_3=1$ 可得 $x_3=-1$. 再將 $x_3=-1$ 代入其上一式 $3x_2+x_3=2$, 得 $3x_2-1=2$, 即 $x_2=1$. 最後將 $x_3=-1,x_2=1$ 代入 $2x_1+x_2+x_3=4$, 得 $x_1=2$. 故得其解為 $x_1=2,x_2=1,x_3=-1$.

(b) 第二種情況是有些 variable x_i 不是 pivot variable. 也就是方程組未知數的個數多於 pivot 的個數. 例如

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

此時 echelon form 的 pivot variable 分別為 x_1,x_2,x_4 少於立方程組的未知數 x_1,x_2,x_3,x_4. 在此情形之下此聯立方程組會有無窮多解. 要得到這種方程組所有的解, 首先我們要找到 *free variables*. 所謂 free variable 指的是方程組不是 pivot variable 的 variable. 例如前面這個例子, x_3 就是 free variable. Free variable 意指它可以任意取值, 所以找到 free variables 後你可以給它們任意的參數, 然後再利用如上一情況中由下往上代回的方式找到聯立方程組所有的解. 例如上一個 augmented matrix 所對應的聯立方程組為

$$\begin{array}{rrrrcr} 2x_1 & +x_2 & +3x_3 & +x_4 & = & 4 \\ & 3x_2 & +3x_3 & +x_4 & = & 2 \\ & & & -x_4 & = & 1 \end{array}$$

首先令 free variable x_3 為一參數 t (表示它可以是任意實數 $t\in\mathbb{R}$). 接著我們從最下面的 $-x_4=1$ 可得 $x_4=-1$. 再將 $x_3=t,x_4=-1$ 代入其上一式 $3x_2+3x_3+x_4=2$, 得 $3x_2+3t-1=2$, 即 $x_2=1-t$. 最後將 $x_2=1-t,x_3=t,x_4=-1$ 代入 $2x_1+x_2+3x_3+x_4=4$, 得 $x_1=2-t$. 故得其解為 $x_1=2-t,x_2=1-t,x_3=t,x_4=-1$, 其中 t 為任意實數. 因為 t 可以是任意實數, 由此我們也知此方程組有無窮多解.

(2) A' 有些 row 全為 0: 此時聯立方程組可能無解, 我們分成兩種情況:

(a) A' 有一個 row 全為 0 但 $\mathbf{b}'$ 在該 row 不為 0. 例如

$$[A'\mid\mathbf{b}']=\left[\begin{array}{ccc|c} 2 & 1 & 1 & 4 \\ 0 & 3 & 1 & 2 \\ 0 & 0 & 0 & 1 \end{array}\right]$$

A' 最後一個 row 皆為 0, 但 $\mathbf{b}'$ 在該 row 的位置為 1. 在此情形之下聯立方程組為 *inconsistent*, 即無解. 例如上一個 augmented matrix 其最後一個 row 所對應的方程式為

$$0x_1+0x_2+0x_3=1$$

但不管 x_1,x_2,x_3 代任何的實數都無法滿足 $0x_1+0x_2+0x_3=1$, 所以此方程組無解.

(b) A' 全為 0 的 row, $\mathbf{b}'$ 在該 row 亦為 0. 例如

$$\left[\begin{array}{cc|c} 2 & 1 & 4 \\ 0 & 3 & 2 \\ 0 & 0 & 0 \end{array}\right], \quad \left[\begin{array}{cccc|c} 2 & 1 & 3 & 1 & 4 \\ 0 & 3 & 3 & 1 & 2 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right]$$

這兩個 augmented matrices 皆為這種情形. 在此情形之下聯立方程組一定是 consistent. 事實上在此情形我們可以忽略全為 0 的 row, 例如前兩個 augmented matrices 所對應的方程組和

$$\left[\begin{array}{cc|c} 2 & 1 & 4 \\ 0 & 3 & 2 \end{array}\right], \quad \left[\begin{array}{cccc|c} 2 & 1 & 3 & 1 & 4 \\ 0 & 3 & 3 & 1 & 2 \end{array}\right]$$

所對應的方程組一樣. 所以我們可依前面 (1) A' 每一個 row 皆不全為 0 的情況找出聯立方程組所有的解.

我們要強調, 絕不會有 pivot 的個數多於方程組 variables (未知數) 的個數的情形發生. 這是因為當係數矩陣 A 是 echelon form 時, 每一個 column 最多僅能有一個 pivot (因為不能有兩個 leading term 在同一個位置), 所以 pivot 的個數不能多於 column 的個數. 而 A 的 column 個數表示的就是此聯立方程組 variables 的個數, 因此 pivot 的個數不會多於 variables 的個數. 另一方面依定義每一個 row 最多僅能有一個 pivot, 所以 pivot 的個數也不會多於該方程組的方程式個數 (即係數矩陣 row 的個數).

Question 2.1. *考慮一個由 n 個 variables 的 m 個方成式所組成的聯立方程組. 試說明前面討論 (1)(a) 的情形只有在 $m=n$ 的時候才有可能發生; 而 (1)(b) 的情形只有在 $m<n$ 的情形才有可能發生;*

在這一節中我們介紹解一次聯立方程組 $A\mathbf{x}=\mathbf{b}$ 的步驟. 也就是先將 augmented matrix $[A\,|\,\mathbf{b}]$ 利用 elementary row operations 化成 $[A'\,|\,\mathbf{b}']$ 其中 A' 為 echelon form 的情形. 再利用上面談論的情況找出聯立方程組的解. 下一節中我們將說明為何可利用 elementary row operations 將 A 變成 echelon form A', 且要說明為何這樣所得的解便是原方程組的解.

2.2. Elementary Row Operations

前一節中我們知道要解一個聯立方程組 $A\mathbf{x}=\mathbf{b}$, 可以先將 augmented matrix $[A\,|\,\mathbf{b}]$ 經由一系列的 elementary row operations 化成 $[A'\,|\,\mathbf{b}']$ 其中 A' 為 echelon form 後再求解. 在本節中我們要說明為何經由 elementary row operations 我們可以將一個矩陣化為 echelon form 且解釋為何經由 elementary row operations 後所對應的聯立方程組與原方程組會有相同的解集合.

我們利用數學歸納法來說明為何一定可以將一個矩陣化為 echelon form. 或許有些人會對這裡數學歸納法處理的方式覺得奇怪, 不過若能仔細體會其真意, 會發現這是最好的處理方式. 我們是對矩陣的 row 的個數作數學歸納法. 先說明所有只有一個 row 的矩陣一定是 echelon form, 然後利用這件事實證明所有有兩個 row 的矩陣皆可利用 elementary row operations 化為 echelon form. 再利用兩個 row 的矩陣會成立的事實證明有 3 個 row 的矩陣也可利用 elementary row operations 化為 echelon form, 如此一直下去我們可證有 4,5,6,... 個 row 的矩陣會成立. 不過這樣的方法我們可以證得有特定個數的 row 的矩陣會成立 (例如 10 個 row), 但無法證得一般的情形 (即任意個數的 row). 此時數學歸納法是最好的論證工具了. 若我們能知道有 k 個 row 的矩陣一定能利用 elementary row operations 化為 echelon form 這個事實且利用這個事實證得有 $k+1$ 個 row 的矩陣一定能利用 elementary row operations 化為 echelon form, 這就表示當我們知道有一個 row 的矩陣能利用 elementary row operations 化為 echelon form 就能推得有兩個 row 的矩陣能利用 elementary row operations 化為 echelon form, 也進而推得有 3 個 row 的矩陣亦成立, 再進而推得有 4 個 row 的矩陣亦成立, 如此一直下去當然可知任意的矩陣皆能利用 elementary row operations 化為 echelon form.

由於這裡的論證不容易說明清楚, 我們先由一個例子來說明. 考慮一個有 3 個 row 的矩陣

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

要將之化為 echelon form 首先第一個 row 必需是 leading entry 的位置在最左邊, 所以我們利用將第一, 二兩個 row 交換的 elementary row operation 將此矩陣變換為

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

由於第三個 row 的 leading entry 位置與第一個 row 相同所以須將此位置的數消掉. 利用第一個 row 乘上 -2 加到第三個 row 的 elementary row operation 將此矩陣變換為

$$\begin{bmatrix} 0 & 1 & 2 & 1 & 3 \\ 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & -2 & -2 & -7 \end{bmatrix}$$

如此一來第一個 row 以下的各 row 的 leading entry 所在位置都在第一個 row 的 leading entry 所在位置的右方. 接下來我們可以不再管第一個 row 而處理第一個 row 以下的部份. 此部份是一個僅有兩個 row 的矩陣

$$\begin{bmatrix} 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & -2 & -2 & -7 \end{bmatrix}$$

若我們知道變換兩個 row 的矩陣成 echelon 的方法, 直接套用就可以完成了. 事實上我們就是將上面的方法再處理一遍, 將第一個 row 乘以 2 加到第二個 row 即可得

$$\begin{bmatrix} 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & -5 \end{bmatrix}$$

這一個 echelon form. 要注意我們故意忽略原來矩陣的第一個 row 的原因就是為了能套用數學歸納法的假設 (即將 row 的個數變少). 事實上若用原來的矩陣, 我們是將矩陣

$$\begin{bmatrix} 0 & 1 & 2 & 1 & 3 \\ 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & -2 & -2 & -7 \end{bmatrix}$$

的第二個 row 乘以 2 加到第三個 row 得

$$\begin{bmatrix} 0 & 1 & 2 & 1 & 3 \\ 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & -5 \end{bmatrix}$$

這一個 echelon form.

接下來我們處理一般的情形. 首先我們來看只有一個 row 的矩陣. 此時由於沒有任何的 row 在其下方所以依定義自然是 echelon form. 接著看有兩個 row 的矩陣. 首先注意依定義一個 echelon form 的第一個 row 其 leading entry (若有的話) 必在所有其他 row 的 leading entry 所在位置的左方. 所以我們在此有兩個 row 的矩陣挑出 leading entry 在最左方的一個 row (若兩個 row 的 leading entry 所在位置相同就任取一個 row) 利用 row 交換的 row operation 將之置於第一個 row. 接下來注意依定義下一個 row 的 leading entry 所在位置需在第一個 row 的 leading entry 的右方. 現若第二個 row 的 leading entry 所在位置和第一個 row 不同, 則因已知第一個 row 的 leading entry 所在位置在最左方, 第二個 row 的 leading entry 所在位置一定在第一個 row 的 leading entry 的右方, 故依定義此時已為 echelon form. 而若第二個 row 的 leading entry 所在位置和第一個 row 相同, 我們可將第一個 row 乘以 $-b/a$, 其中 a 為第一個 row 的 leading entry 而 b 為第二個 row 的 leading entry, 再加到第二個 row 上. 如此一來第二個 row 原本的 leading entry 所在位置變為 0, 故其 leading entry 所在位置往右移了, 依定義此時為 echelon form.

我們可以如法泡製處理有 3 個 row 的矩陣, 但由於要使用數學歸納法, 此時我們可直接假設我們已處理到有 k 個 row 的矩陣了, 亦即有 k 個 row 的矩陣皆可利用 elementary row operation 化為 echelon form. 現在我們要處理有 $k+1$ 個 row 的矩陣. 如前面的方法, 首先我們將 leading entry 的位置在最左邊的那個 row 利用兩 row 互換的 row operation 將之置於第一個 row. 假設此時第一個 row 的 leading entry 為 a. 接下來我們將 leading entry 的位置與第一個 row 的 leading entry 位置一樣的 row 挑出, 若該 row 的 leading entry 為 b, 我們便將第一個 row 乘上 $-b/a$ 後加到該 row 上. 如此一來該 row 的 leading entry 所在位置便往右移了. 一直重複此步驟, 直到第一個 row 以外的 row 其 leading entry 所在位置皆與第一個 row 的 leading entry 所在位置相異. 注意, 此時第一個 row 以下的各 row 其 leading entry 所在位置皆在第一個 row 的 leading entry 所在位置的右方. 若我們不看第一個 row, 所剩下的是一個有 k 個 row 的矩陣, 所以利用前面已知有 k 個 row 的矩陣皆可利用 elementary row operations 化為 echelon form, 我們可以利用 elementary row operations 將此矩陣第一個 row 以下的部份化為 echelon form. 但此時因各個 row 的 leading entry 所在位置皆在第一個 row 的 leading entry 所在位置的右方, 所以整個矩陣亦為 echelon form. 故得證所有矩陣皆可利用 elementary row operations 化為 echelon form. 大家或許注意到我們在化成 echelon form 的過程皆沒有用到將某個 row 乘上一非 0 實數這一個 elementary row operation. 事實上在化成 echelon form 的過程確實不需要這一種 row operation, 不過它在以後要談的化為 “reduced” echelon form 的過程是需要的, 留待以後再談.

既然每一個矩陣都能用 elementary row operations 化為 echelon form, 接下來我們要說明的是利用 elementary row operation 處理後的聯立方程組其解集合不會改變. 要注意這裡

指的是將 augmented matrix 用 elementary row operations 變換後的 augmented matrix 其對應的聯立方程組其解集合不會改變. 亦即聯立方程組 $A\mathbf{x}=\mathbf{b}$ 所對應的 augmented matrix $[A \mid \mathbf{b}]$ 若經一些 elementary row operations 後變換成 $[A' \mid \mathbf{b}']$, 那麼其對應的聯立方程組 $A'\mathbf{x}=\mathbf{b}'$ 和原方程組 $A\mathbf{x}=\mathbf{b}$ 有相同的解集合. 不要誤以為是將係數矩陣 A 利用 elementary row operations 變換成 A' 後聯立方程組 $A'\mathbf{x}=\mathbf{b}$ 的解集合和原方程組 $A\mathbf{x}=\mathbf{b}$ 的解集合相同.

首先觀察若將一聯立方程組 $A\mathbf{x}=\mathbf{b}$ 的 augmented matrix $[A \mid \mathbf{b}]$ 利用三種 elementary row operation 的任一種變換成 $[A' \mid \mathbf{b}']$ 表示將原方程組利用加減消去法的三個基本方法之一將之變成方程組 $A'\mathbf{x}=\mathbf{b}'$. 然而方程組 $A\mathbf{x}=\mathbf{b}$ 若利用加減消去法的三種方法 (即將兩式子對調順序或將某一式乘上某個非 0 實數或將一個式子乘上某個實數加到另一個式子) 變換成方程組 $A'\mathbf{x}=\mathbf{b}'$, 原來滿足 $A\mathbf{x}=\mathbf{b}$ 的一組解仍會滿足 $A'\mathbf{x}=\mathbf{b}'$. 換句話說 $A\mathbf{x}=\mathbf{b}$ 的解集合會包含於 $A'\mathbf{x}=\mathbf{b}'$ 的解集合. 不過加減消去法的三個基本方法是可以還原回去的, 也就是說方程組 $A'\mathbf{x}=\mathbf{b}'$ 也可以用加減消去法的三種方法還原回原方程組 $A\mathbf{x}=\mathbf{b}$. 因此 $A'\mathbf{x}=\mathbf{b}'$ 的解集合也會包含於 $A\mathbf{x}=\mathbf{b}$ 的解集合. 因此得證 $A\mathbf{x}=\mathbf{b}$ 和 $A'\mathbf{x}=\mathbf{b}'$ 會有相同的解集合. 我們證得了 $[A \mid \mathbf{b}]$ 若經由一個 elementary row operation 後得 $[A' \mid \mathbf{b}']$, 則它們所對應的聯立方程組會有相同的解集合. 因此若 $[A \mid \mathbf{b}]$ 經由好幾次的 elementary row operation 變換成 $[A' \mid \mathbf{b}']$, 它們所對應的聯立方程組當然也會有相同的解集合.

Example 2.2.1. Solve the linear system

$$\begin{array}{rrrcl} & x_2 & -3x_3 & = & -5 \\ 2x_1 & +3x_2 & -1x_3 & = & 7 \\ 4x_1 & +5x_2 & -2x_3 & = & 10. \end{array}$$

此聯立方程組的 augmented matrix 為

$$\left[\begin{array}{rrr|r} 0 & 1 & -3 & -5 \\ 2 & 3 & -1 & 7 \\ 4 & 5 & -2 & 10 \end{array}\right].$$

由於第二, 三 row 的 leading entry 在最左端. 但第二 row 的 leading entry 的值較小, 為了計算方便, 我們將之置於第一個 row, 即將一, 二 row 交換得

$$\left[\begin{array}{rrr|r} 2 & 3 & -1 & 7 \\ 0 & 1 & -3 & -5 \\ 4 & 5 & -2 & 10 \end{array}\right].$$

接下來由於第三 row 的 leading entry 也在 x_1 的位置需要消去, 所以將第一 row 乘上 -2 加到第三 row 得

$$\left[\begin{array}{rrr|r} 2 & 3 & -1 & 7 \\ 0 & 1 & -3 & -5 \\ 0 & -1 & 0 & -4 \end{array}\right].$$

此時係數矩陣仍不是 echelon form, 需將第三 row 的 x_2 位置的 entry 消去. 故將第二 row 加至第三 row 得

$$\left[\begin{array}{rrr|r} 2 & 3 & -1 & 7 \\ 0 & 1 & -3 & -5 \\ 0 & 0 & -3 & -9 \end{array}\right].$$

這是 echelon form. 由於係數矩陣沒有全為 0 的 row, 我們知此 linear system 為 consistent. 而又 pivot 的個數等於 variable 的個數, 故知此 linear system 的解唯一. 事實上, 最下面第三 row 表示 $-3x_3=-9$, 得 $x_3=3$. 代入第二 row 表示的 $x_2-3x_3=-5$, 得 $x_2=4$. 最後代入第一 row 表示的 $2x_1+3x_2-x_3=7$, 得 $x_1=-1$. 故知此 linear system 的解為 $(x_1,x_2,x_3)=(-1,4,3)$.

Example 2.2.2. Solve the linear system

$$\begin{array}{rrrrcr} x_1 & -1x_2 & +2x_3 & +3x_4 & = & 2 \\ 2x_1 & +1x_2 & +1x_3 & & = & 1 \\ x_1 & +2x_2 & -1x_3 & -3x_4 & = & 7. \end{array}$$

此聯立方程組的 augmented matrix 為

$$\left[\begin{array}{rrrr|r} 1 & -1 & 2 & 3 & 2 \\ 2 & 1 & 1 & 0 & 1 \\ 1 & 2 & -1 & -3 & 7 \end{array}\right].$$

第二, 三 row 的 leading entry 需被消去. 故將第一 row 分別乘上 $-2,-1$ 加到第二, 三 row 得

$$\left[\begin{array}{rrrr|r} 1 & -1 & 2 & 3 & 2 \\ 0 & 3 & -3 & -6 & -3 \\ 0 & 3 & -3 & -6 & 5 \end{array}\right].$$

接下來由於第三 row 的 leading entry 需要消去, 所以將第二 row 乘上 -1 加到第三 row 得

$$\left[\begin{array}{rrrr|r} 1 & -1 & 2 & 3 & 2 \\ 0 & 3 & -3 & -6 & -3 \\ 0 & 0 & 0 & 0 & 8 \end{array}\right].$$

這是 echelon form. 由於第三 row 表示 $0x_1+0x_2+0x_3=8$, 知此 linear system 為 inconsistent.

Example 2.2.3. Solve the linear system

$$\begin{array}{rrrrcr} x_1 & -2x_2 & +1x_3 & -1x_4 & = & 4 \\ 2x_1 & -3x_2 & +4x_3 & -3x_4 & = & -1 \\ 3x_1 & -5x_2 & +5x_3 & -4x_4 & = & 3 \\ -x_1 & +1x_2 & -3x_3 & +2x_4 & = & 5. \end{array}$$

此聯立方程組的 augmented matrix 為

$$\left[\begin{array}{rrrr|r} 1 & -2 & 1 & -1 & 4 \\ 2 & -3 & 4 & -3 & -1 \\ 3 & -5 & 5 & -4 & 3 \\ -1 & 1 & -3 & 2 & 5 \end{array}\right].$$

第二, 三, 四 row 的 leading entry 需被消去. 故將第一 row 分別乘上 $-2,-3,1$ 加到第二, 三, 四 row 得

$$\left[\begin{array}{rrrr|r} 1 & -2 & 1 & -1 & 4 \\ 0 & 1 & 2 & -1 & -9 \\ 0 & 1 & 2 & -1 & -9 \\ 0 & -1 & -2 & 1 & 9 \end{array}\right].$$

接下來第三, 四 row 的 leading entry 需要消去, 所以將第二 row 分別乘上 $-1,1$ 加到第三, 四 row 得

$$\left[\begin{array}{rrrr|r} 1 & -2 & 1 & -1 & 4 \\ 0 & 1 & 2 & -1 & -9 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right].$$

這是 echelon form. 由於係數矩陣全為 0 的第三, 四 row 全為 0, 知此 linear system 為 consistent.

事實上此 linear system 的 pivot variables 為 x_1,x_2, 而 free variables 為 x_3,x_4. 我們可以令 $x_4=r$, $x_3=s$, 代入第二 row 表示的 $x_2+2x_3-x_4=-9$, 得 $x_2=-9+r-2s$. 再代入第一 row 表示的 $x_1-2x_2+x_3-x_4=4$, 得 $x_1=-14+3r-5s$. 故知此 linear system 的解為

$$(x_1,x_2,x_3,x_4)=(-14+3r-5s,-9+r-2s,s,r), r,s\in\mathbb{R}.$$

通常我們習慣寫成 row vector 且將 r,s 提出. 故將解寫成

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} -14 \\ -9 \\ 0 \\ 0 \end{bmatrix} + r\begin{bmatrix} 3 \\ 1 \\ 0 \\ 1 \end{bmatrix} + s\begin{bmatrix} -5 \\ -2 \\ 1 \\ 0 \end{bmatrix}, r,s\in\mathbb{R}.$$

在解 linear system 的過程中還可以進一步將 echelon form 化為所謂的 *reduced echelon form*. Reduced echelon form 事實上仍為 echelon form, 不過再加上兩個限制. 第一個限制是每一個 pivot 需為 1. 另一個限制為 pivot 的位置上方全為 0. 要注意, 依定義 echelon form 的 pivot 位置下方已全為 0 所以 reduced echelon form 每一個 pivot 所在的 column, 除了自己需為 1 外其他部分皆為 0. 例如

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

都不是 reduced echelon form 但是

$$A'=\begin{bmatrix} 1 & 2 & 0 & 0 \\ 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \quad B'=\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 3 \\ 0 & 0 & 1 & -1 \end{bmatrix}$$

就是 reduced echelon form. 每一個 echelon form 皆可利用 elementary row operations 換為 reduced echelon form. 這是因為, 若有一個 row 的 pivot 為 a (注意依定義 $a\neq 0$), 我們只要將該 row 乘上 $1/a$, 則該 row 的 pivot 便是 1 了. 例如上面 A 這一個 echelon form 若將第二個 row 乘上 $1/3$, 就可得 A' 這一個 reduced echelon form. 當我們將每個 pivot 都變為 1 後, 就可利用將該 row 乘上某一實數加到另一個 row 的方法將 pivot 所在的 column 的其他部分化為 0. 例如上面 B 這一個 echelon form 若將第三個 row 分別乘上 -3, -1 加到第一個 row 和第二個 row, 得 $\begin{bmatrix} 1 & 1 & 0 & 3 \\ 0 & 1 & 0 & 3 \\ 0 & 0 & 1 & -1 \end{bmatrix}$. 再將第二個 row 乘上 -1 加到第一個 row, 就可得 B' 這一個 reduced echelon form. 注意一般我們都是從上而下將矩陣換成 echelon form, 不過得到 echelon form 後是從下而上將 echelon form 換成 reduced echelon form 較為方便.

既然每一個矩陣都可以經由 elementary row operations 化為 reduced echelon form, 所以我們可以利用這個方法求出聯立方程組的解. 化成 reduced echelon form 後由於每一個 row 除了該 row 的 pivot 外, 只剩 free variables (其他的 pivot variable 所在的 entry 皆為 0), 所以可以很快地看出解的形式. 例如方程組 $B'\mathbf{x}=\mathbf{O}$ 為

$$\begin{array}{cccccc} x_1 & & & & = & 0 \\ & x_2 & & +3x_4 & = & 0 \\ & & x_3 & -x_4 & = & 0 \end{array}$$

因僅 x_4 為 free variable, 令 $x_4=t$, 代入第三 row 得 $x_3=t$. 代入第二 row 得 $x_2=-3t$. 最後由第一 row 得 $x_1=0$. 故知解為 $(x_1,x_2,x_3,x_4)=(0,-3t,t,t)=t(0,-3,1,1),t\in\mathbb{R}$.

化成 reduced echelon form 雖然在最後可以很快地看出解的形式, 但一般來說化為 reduced echelon form 比僅化為 echelon form 所需的步驟多了許多, 所以利用 echelon form 來求解還是會比較快. 利用 echelon form 求解的方法一般稱為 *Gauss method*, 而用 reduced echelon form 求解一般稱為 *Gauss-Jordan method*.

Example 2.2.4. Example 2.2.1 的 linear system, 化成 echelon form 後為

$$\left[\begin{array}{ccc|c} 2 & 3 & -1 & 7 \\ 0 & 1 & -3 & -5 \\ 0 & 0 & -3 & -9 \end{array}\right].$$

將第三 row 乘以 $-1/3$ 得

$$\left[\begin{array}{ccc|c} 2 & 3 & -1 & 7 \\ 0 & 1 & -3 & -5 \\ 0 & 0 & 1 & 3 \end{array}\right].$$

再將第三 row 分別乘以 $3,1$ 加到第二, 第一 row 得

$$\left[\begin{array}{ccc|c} 2 & 3 & 0 & 10 \\ 0 & 1 & 0 & 4 \\ 0 & 0 & 1 & 3 \end{array}\right].$$

接著將第二 row 乘以 -3 加到第一 row 得

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

最後將第一 row 乘以 $1/2$ 得 reduced echelon form

$$\left[\begin{array}{ccc|c} 1 & 0 & 0 & -1 \\ 0 & 1 & 0 & 4 \\ 0 & 0 & 1 & 3 \end{array}\right],$$

且馬上看出解為 $(x_1,x_2,x_3)=(-1,4,3)$.

Example 2.2.3 的 linear system, 化成 echelon form 後為

$$\left[\begin{array}{cccc|c} 1 & -2 & 1 & -1 & 4 \\ 0 & 1 & 2 & -1 & -9 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right].$$

將第二 row 乘以 2 加到第一 row 得 reduced echelon form

$$\left[\begin{array}{cccc|c} 1 & 0 & 5 & -3 & -14 \\ 0 & 1 & 2 & -1 & -9 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right].$$

因 x_4, x_3 為 free variables, 令 $x_4 = r$, $x_3 = s$, 代入第二 row 得 $x_2 = -9 + r - 2s$. 再代入第一 row 得 $x_1 = -14 + 3r - 5s$.

2.3. Echelon Form

我們已知要探討聯立方程組 $A\mathbf{x} = \mathbf{b}$ 的解, 僅要考慮 A 為 echelon form 的情形. 這一節中我們就是要討論當 A 為 echelon form 時, 聯立方程組 $A\mathbf{x} = \mathbf{b}$ 解的特性. 事實上我們很容易理解利用 2.1 節中所提求解的方法所得的結果皆為方程組的一組解. 這裡要探討的是為何利用 2.1 節中所提求解的方法, 就可得所有的解. 接著我們將說明, 雖然一個矩陣利用 elementary row operations 化為 echelon form 的結果不唯一, 但是它們的 pivot variables 是唯一的. 我們也因此可推得化為 reduced echelon form 的結果會唯一.

如果我們得到 2.1 節 (2)(a) 的情形 (即 A 有一個 row 全為 0 但 $\mathbf{b}$ 在該 row 不為 0), 在該節已說明此時方程組無解. 所以我們只要探討有解的情形. 首先回顧一下在 2.1 節所提求解的方法: 首先我們要找到 *free variables*, 也就是是方程組除了 pivot variable 以外的 variable. 接著給這些 free variable 任意的參數值, 然後再利用由下往上代回的方式找到聯立方程組所有的解. 若無 free variable, 就直接由下往上一步一步求值即可.

由於可以忽略 augmented matrix 全為 0 的 row, 所以我們可假設係數矩陣 A 沒有一個 row 全為 0. 因為 A 為 echelon form, 這也表示 A 每一個 row 皆有 leading entry 且為 pivot. 現在我們回答當 A 是 echelon form 時, 2.1 節中所述解聯立方程組 $A\mathbf{x} = \mathbf{b}$ 的方法所求得的解就是所有的解. 也就是說給定 $x_1 = c_1, \ldots, x_n = c_n$ 為 $A\mathbf{x} = \mathbf{b}$ 的一組解, 我們要說明這組解確實可由 2.1 節所提的方法得到. 為了方便起見我們令 2.1 節所提的方法所得的解所成的集合為 S. 我們要說明 $(x_1, \ldots, x_n) = (c_1, \ldots, c_n)$ 確實為 S 中的元素. 現若 x_n 為 pivot variable, 則 x_n 的值是被唯一確定的. 所以 S 的所有解中 x_n 的取值一定也為 c_n. 若 x_n 為 free variable, 則因 S 的解中 x_n 可為任意值, 故 S 中一定有一組解其 x_n 的取值為 c_n. 也就是說不管 x_n 是否為 pivot variable, S 中必有一組解其 x_n 的取值為 c_n. 現若 x_{n-1} 為 pivot variable, 則由此 pivot 所在的 row 所對應的方程式可知 x_{n-1} 的取值會被 x_n 的取值所決定. 今已知 S 中必有一組解其 x_n 的取值為 c_n, 故此組解必滿足 $x_{n-1} = c_{n-1}, x_n = c_n$; 而若 x_{n-1} 為 free variable, 則因 S 的解中 x_{n-1} 可為任意值且其取值不影響到 x_n 的取值, 故知 S 中必有一組解其 x_{n-1}, x_n 的取值為 $x_{n-1} = c_{n-1}, x_n = c_n$. 如此一直下去我們知道 S 中必有一組解其 $x_1, \ldots, x_n$ 的取值為 $x_1 = c_1, \ldots, x_n = c_n$.

我們可以利用上面的概念, 推導出當 A 為 echelon form 時, pivot variables 和 free variables 對聯立方程組 $A\mathbf{x} = \mathbf{b}$ 解的影響. 首先看 pivot variable 對聯立方程組的解之影響.

Lemma 2.3.1. *假設 A 為一有 n 個 column 的 echelon form 且 $x_1 = c_1, \ldots, x_n = c_n$ 和 $x_1 = d_1, \ldots, x_n = d_n$ 皆為方程組 $A\mathbf{x} = \mathbf{b}$ 的一組解.*

(1) 假設 x_n 為 A 的一個 *pivot variable.* 則 $c_n = d_n$.

(2) 假設 x_k 為 A 的一個 *pivot variable,* 其中 $1 \le k \le n-1$. 若 $c_{k+1} = d_{k+1}, \ldots, c_n = d_n$, 則 $c_k = d_k$.

Proof. 假設聯立方程組為

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

其中

$$A = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ a_{21} & \cdots & a_{2n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix}$$

為 echelon form, 且不失一般性我們假設 A 的每一個 row, $a_{i1}, \ldots, a_{in}$ 皆不全為 0.

(1) 若 x_n 為 A 的一個 pivot variable, 表示 A 的最後一個 row 的 leading entry 所在位置為 x_n. 也就是說 $a_{m1} = a_{m2} = \cdots = a_{mn-1} = 0$ 且 $a_{mn} \neq 0$. 這表示此聯立方程組中最後一個式子為 $a_{mn}x_n = b_m$. 故由 $x_1 = c_1, \ldots, x_n = c_n$ 及 $x_1 = d_1, \ldots, x_n = d_n$ 皆為此聯立方程組的一組解知 $x_n = c_n$ 和 $x_n = d_n$ 皆需滿足 $a_{mn}x_n = b_m$, 亦即 $a_{mn}c_n = b_m$ 且 $a_{mn}d_n = b_m$. 故由 $a_{mn} \neq 0$ 得知 $c_n = d_n$.

(2) 若 x_k 為 A 的一個 pivot variable, 表示 A 有一個 row 的 leading entry 所在位置為 x_k. 也就是說若此 row 為 A 的第 i 個 row, 則 $a_{i1} = a_{i2} = \cdots = a_{ik-1} = 0$ 且 $a_{ik} \neq 0$. 此 row 所對應的式子為 $a_{ik}x_k + \cdots + a_{in}x_n = b_i$. 故由 $x_1 = c_1, \ldots, x_n = c_n$ 及 $x_1 = d_1, \ldots, x_n = d_n$ 皆為此聯立方程組的一組解知 $x_k = c_k, x_{k+1} = c_{k+1}, \ldots, x_n = c_n$ 和 $x_k = d_k, x_{k+1} = d_{k+1}, \ldots, x_n = d_n$ 皆需滿足 $a_{ik}x_k + \cdots + a_{in}x_n = b_i$. 因此由 $c_{k+1} = d_{k+1}, \ldots, c_n = d_n$ 的假設知

$$a_{ik}c_k = b_i - (a_{ik+1}c_{k+1} + \cdots + a_{in}c_n) = b_i - (a_{ik+1}d_{k+1} + \cdots + a_{in}d_n) = a_{ik}d_k.$$

再由 $a_{ik} \neq 0$ 得知 $c_k = d_k$.

□

相對於 pivot variable 我們知道對於 free variable 我們可以隨意取任何的實數而得到一組解, 所以我們有以下 free variable 對解的影響.

Lemma 2.3.2. 假設 A 為一有 n 個 *column* 的 *echelon form* 且沒有一個 *row* 全為 0.

(1) 假設 x_n 為 A 的一個 *free variable.* 則對任意的實數 r, 方程組 $A\mathbf{x} = \mathbf{b}$ 皆可找到一組解其 $x_n = r$.

(2) 假設 x_k 為 A 的一個 *free variable,* 其中 $1 \le k \le n-1$. 若 $x_1 = c_1, \ldots, x_n = c_n$ 為方程組 $A\mathbf{x} = \mathbf{b}$ 的一組解, 則對任意實數 r 方程組 $A\mathbf{x} = \mathbf{b}$ 皆可找到一組解其 $x_k = r$ 且 $x_{k+1} = c_{k+1}, \ldots, x_n = c_n$.

Proof. 在前面所提的求解過程中我們知道可將 free variable 定為任意的實數, 再一步一步由下往上代回得到一組解. 在這個過程中我們了解到若 x_i 是 free variable, 則它的取值可能會影響到的僅有 x_j, 其中 $j<i$ 這樣的變數, 而不會影響到其他變數 x_l, 其中 $i<l$ 的取值.

現若 x_n 是 free variable, 這表示我們可以設定 x_n 為任意實數, 再一步一步往上代求得聯立方程組的一組解, 所以對任意的實數 r, 方程組 $A\mathbf{x}=\mathbf{b}$ 皆可找到一組解其 x_n 為 r.

若 x_k 為 A 的一個 free variable, 其中 $1\leq k\leq n-1$ 且已知 $x_1=c_1,\ldots,x_n=c_n$ 為方程組 $A\mathbf{x}=\mathbf{b}$ 的一組解. 換言之, $x_{k+1}=c_{k+1},\ldots,x_n=c_n$ 皆滿足方程組 pivot 的位置在 x_k 右方的那些 row 所對應的那些方程式. 由於 x_k 可取任意的實數且不會影響 $x_{k+1},\ldots,x_n$ 的取值, 所以我們可令 $x_k=r$ 且 $x_{k+1}=c_{k+1},\ldots,x_n=c_n$ 一步一步代回求得聯立方程組的一組解. □

Lemma 2.3.1 和 Lemma 2.3.2 有許多應用. 例如當 A 是 echelon form 時若聯立方程組 $A\mathbf{x}=\mathbf{b}$ 已知有一個解 $x_1=c_1,\ldots,x_n=c_n$ 且 $x_1,\ldots,x_n$ 每一個都是 pivot variable, 則由 Lemma 2.3.1 知聯立方程組 $A\mathbf{x}=\mathbf{b}$ 的解僅能是 $x_1=c_1,\ldots,x_n=c_n$. 換句話說此方程組的解唯一. 另一方面, 若聯立方程組 $A\mathbf{x}=\mathbf{b}$ 已知有解且 $x_1,\ldots,x_n$ 中有 free variable, 則由 Lemma 2.3.2 知聯立方程組 $A\mathbf{x}=\mathbf{b}$ 會有無窮多解.

當我們給一個矩陣時, 有許多種方法將之化為 echelon form, 而且化成的 echelon form 很可能不一樣. 不過利用 Lemma 2.3.1 和 Lemma 2.3.2 我們可以得到這些 echelon form 雖然可能不一樣, 但他們 pivot 的所在位置都會一致. 由於我們只關心係數矩陣 A 化為 echelon form 後的情形, 所以我們可以考慮 $A\mathbf{x}=\mathbf{O}$ 這一種特殊形式的聯立方程組. 要注意這樣的聯立方程組都會有解, 因為 $x_1=0,\ldots,x_n=0$ 就是一組解. 我們特別稱這樣的聯立方程組為 *homogeneous system*.

Proposition 2.3.3. *給定一矩陣 A, 若 A_1,A_2 均為 A 利用 elementary row operations 化成的 echelon forms. 則 A_1 和 A_2 的 pivot 個數相同, 事實上他們的 pivot variables 是一致的.*

Proof. 我們考慮 $A\mathbf{x}=\mathbf{O}$ 這一組聯立方程組, 其中 A 有 n 個 column (即此方程組有 n 個變數). 因為 A 可利用 elementary row operation 化為 A_1 及 A_2, 這表示 augmented matrix $[A\mid\mathbf{O}]$ 可以利用 elementary row operation 化為 $[A_1\mid\mathbf{O}]$ 及 $[A_2\mid\mathbf{O}]$. 換句話說聯立方程組 $A_1\mathbf{x}=\mathbf{O}$ 和 $A_2\mathbf{x}=\mathbf{O}$ 皆與聯立方程組 $A\mathbf{x}=\mathbf{O}$ 有同樣的解. 再次強調這些聯立方程組都會有 $x_1=0,\ldots,x_n=0$ 這樣的一組解.

我們要用反證法處理. 假設 A_1 和 A_2 有 pivot variable 不一致, 不失一般性我們就假設對 A_1 來說 x_i 是 pivot variable 但對 A_2 來說 x_i 不是 pivot variable (即 free variable). 假設 $i=n$, 這表示方程組 $A_1\mathbf{x}=\mathbf{O}$ 的解中 x_n 的取值是唯一的 (Lemma 2.3.1), 事實上 x_n 一定為 0; 但 $A_2\mathbf{x}=\mathbf{O}$ 的解中 x_n 的取值卻可以是任意的實數 (Lemma 2.3.2). 這和此二方程組有相同的解相矛盾. 現若 $1\leq i\leq n-1$. 利用 $x_1=0,\ldots,x_n=0$ 已是這兩聯立方程組的解, 我們知道方程組 $A_1\mathbf{x}=\mathbf{O}$ 的解中一定找不到一組解其 $x_{i+1},\ldots,x_n$ 的取值皆為 0 但 x_i 的取值不是 0 (Lemma 2.3.1); 另一方面 Lemma 2.3.2 告訴我們 $A_2\mathbf{x}=\mathbf{O}$ 的解中一定可找到一組解其 $x_{i+1},\ldots,x_n$ 的取值皆為 0 但 x_i 的取值不是 0 (事實上 x_i 可以是任意實數). 這又和

$A_1\mathbf{x}=\mathbf{O}, A_2\mathbf{x}=\mathbf{O}$ 此二方程組有相同的解相矛盾. 故由反證法知 A_1 和 A_2 的 pivot variables 是一致的. □

由於一個矩陣化為 echelon form 其 pivot 的個數是固定的, 我們特別有以下的定義.

Definition 2.3.4. 假設 A 為一矩陣. 若 A 利用 elementary row operations 化為 echelon form 後其 pivot 的個數為 r, 我們稱 r 為 A 的 *rank*. 用 $\operatorname{rank}(A)=r$ 來表示.

Question 2.2. 假設矩陣 A 有 m 個 *row* 以及 n 個 *column*. 若 $\operatorname{rank}(A)=r$, 試說明 r 小於等於 m,n 的最小值.

我們可以利用 Proposition 2.3.3 證明一個矩陣利用 elementary row operations 化為 reduced echelon form 其結果是唯一的.

Proposition 2.3.5. 給定一矩陣 A, 若 A_1,A_2 均為 A 利用 *elementary row operations* 化成的 *reduced echelon forms,* 則 $A_1=A_2$.

Proof. 我們考慮 $A\mathbf{x}=\mathbf{O}$ 這一組聯立方程組, 依假設聯立方程組 $A_1\mathbf{x}=\mathbf{O}$ 和 $A_2\mathbf{x}=\mathbf{O}$ 皆與聯立方程組 $A\mathbf{x}=\mathbf{O}$ 有同樣的解.

利用反證法. 假設 $A_1\neq A_2$ 且假設從下往上, A_1,A_2 第一個發生相異的 row 其 pivot variable 為 x_k (注意由 Proposition 2.3.3, 我們知道 A_1,A_2 的 pivot variables 是一致的). 現假設 A_1,A_2 在此 row 所對應的方程式分別為

$$x_k+a_{k+1}x_{k+1}+\cdots+a_nx_n=0 \quad 與 \quad x_k+b_{k+1}x_{k+1}+\cdots+b_nx_n=0,$$

其中存在 l 滿足 $k+1\leq l\leq n$ 且 $a_l\neq b_l$. 由於 A_1,A_2 皆為 reduced echelon form, 若 $j>k$ 且 x_j 為 pivot variable, 則 $a_j=b_j=0$. 因此若 x_l 為 pivot variable, 則會導致 $a_l=b_l=0$ 之矛盾, 故知 x_l 必為 free variable. 我們知給定一組 free variables 的值, 可以用由下往上代回的方式得到聯立方程組的解. 現考慮除了 x_l 這一個 free variable 代 1, 其他 free variables 代 0 所得的解. 設依此所得 $A_1\mathbf{x}=\mathbf{O}$ 與 $A_2\mathbf{x}=\mathbf{O}$ 的解分別為 $x_1=c_1,\ldots,x_n=c_n$ 與 $x_1=d_1,\ldots,x_n=d_n$. 注意, 若 $j>k$ 且 x_j 為 pivot variable, 則 $a_j=b_j=0$, 所以此時 x_j 的取值不會影響到 x_k 的取值, 也就是說依此假設我們有 $c_l=d_l=1$ 以及 $c_k=-a_l$, $d_k=-b_l$. 又由於依假設 A_1,A_2 在 x_k 為 pivot 這一 row 以下的各 row 都一致, 我們知對所有 $k+1\leq i\leq n$ 皆有 $c_i=d_i$. 然而 $x_1=c_1,\ldots,x_n=c_n$ 與 $x_1=d_1,\ldots,x_n=d_n$ 皆為 $A\mathbf{x}=\mathbf{O}$ 的解且 x_k 為 pivot variable, 故由 Lemma 2.3.1 知 $c_k=d_k$. 可得 $a_l=-c_k=-d_k=b_l$. 此與 $a_l\neq b_l$ 的假設相矛盾, 故知 $A_1=A_2$. □

一般來說, 若矩陣 A 可經由 elementary row operations 化為矩陣 B, 我們就說 A is *row equivalent* to B, 用 $A\sim B$ 來表示. 若 A,B,C 為矩陣, 我們有以下的性質.

(1) $A\sim A$

(2) 若 $A\sim B$, 則 $B\sim A$.

(3) 若 $A\sim B$ 且 $B\sim C$, 則 $A\sim C$.

(1) 說的是 A 可由 elementary row operations 化為 A. 事實上只要先兩 row 互換再換回來即可. (3) 說的是若 A 可由 elementary row operations 化為 B 且 B 可由 elementary row operations 化為 C, 那麼我們可以把這兩組 row operations 連結一起, 直接將 A 化為 C. (2) 說的是前面提過每個 elementary raw operation 是可以還原回去的, 所以若 A 可由 elementary row operations 化為 B, 我們可以將每個步驟逆轉還原. 也就是說 B 也可由 elementary row operation 還原回 A.

假設 A,B 經由 elementary row operations 可化為 A',B' 其中 A',B' 皆為 echelon form. 若 $A'=B'$, 可由 $A\sim A'$ 且 $B\sim B'$ 利用前面所提 row equivalent 的性質推得 $A\sim B$. 但這不是有效的方法來判斷 A,B 是否為 row equivalent, 因為我們知道即使同一個矩陣, 化成 echelon form 並不唯一. 但若化成 reduced echelon form 問題便解決了 (這說明了唯一性的重要). 我們有以下的性質.

Proposition 2.3.6. 設 A,B 為矩陣且分別可由 *elementary row operations* 化為 *reduced echelon forms* A',B'. 則 $A\sim B$ 若且唯若 $A'=B'$.

Proof. ($\Leftarrow$) 若 $A'=B'$, 則 $A\sim A'=B'$, 故由 row equivalent 的性質 $A\sim B'$ 以及 $B'\sim B$ 得 $A\sim B$.

($\Rightarrow$) 若 $A\sim B$, 則由 $B\sim B'$ 得 $A\sim B'$. 亦即 A 可由 elementary row operations 化為 reduced echelon form B'. 然而 Proposition 2.3.5 A 化為 reduced echelon form 是唯一的, 故得證 $A'=B'$. □

由 Proposition 2.3.6, 我們只要將兩個矩陣化為 reduced echelon form, 就可以明確的知道它們是否為 row equivalent.

Question 2.3. 假設矩陣 A 的 *row* 和 *column* 的個數皆為 n. 試證明 $\mathrm{rank}(A)=n$ 若且唯若 A 和 $n\times n$ 的 *identity matrix (*單位矩陣*) (*即對角線位置為 1 其他位置為 0 形如 $\begin{bmatrix} 1 & & \mathbf{O} \\ & \ddots & \\ \mathbf{O} & & 1 \end{bmatrix}$ 的矩陣*)* 為 *row equivalent.*

Question 2.4. 若矩陣 A,B 有相同個數的 *row* 與相同個數的 *column.* 已知聯立方程組 $A\mathbf{x}=\mathbf{b}$ 和 $B\mathbf{x}=\mathbf{b}$, 皆有唯一解. 是否可得 $A\sim B$? 是否 $A\mathbf{x}=\mathbf{b}$ 和 $B\mathbf{x}=\mathbf{b}$ 會有同樣的解?

2.4. The Rank of a Matrix

在前一節中我們知道一個矩陣經由 elementary row operations 化為 echelon form 之後, 其中不全為 0 的 row 的個數 (即 pivot 的個數) 是固定的, 我們稱之為此矩陣的 rank. 其實一個一次聯立方程組其係數矩陣的 rank 可以提供我們有關此聯立方程組是否有解以及解是否唯一等訊息. 在本節中我們便是要探討這個課題.

考慮一次聯立方程組 $A\mathbf{x}=\mathbf{b}$, 其中 A 為 $m\times n$ matrix 且 $\mathrm{rank}(A)=r$. 現若 $r=m$, 表示將 A 利用 elementary row operations 化為 echelon form 後每一個 row 皆不全為 0, 此即 2.1 節 (1) 的情形. 因此此時不管 $\mathbf{b}$ 為何, $A\mathbf{x}=\mathbf{b}$ 為 consistent, 亦即聯立方程組必有解. 但若

$r < m$ (回顧一下 $r > m$ 不可能發生) 會怎樣呢? 此時 A 利用 elementary row operations 化為 echelon form A' 後, 有些 row 是全為 0 的. 因此有可能存在 $\mathbf{b}$ 使得增廣矩陣 $[A \mid \mathbf{b}]$ 利用 elementary row operations 化為 $[A' \mid \mathbf{b}']$ 後, $\mathbf{b}'$ 對應到 echelon form A' 中某個全為 0 的 row 的位置不是 0 (即 2.1 節 (2)(a) 的情形), 此時 $A\mathbf{x} = \mathbf{b}$ 為 inconsistent, 亦即聯立方程組 $A\mathbf{x} = \mathbf{b}$ 無解. 事實上只要任取一個 $\mathbf{b}' \in \mathbb{R}^m$ 使得 $\mathbf{b}'$ 對應到 echelon form A' 中某個全為 0 的 row 的位置不是 0, 然後利用 elementary row operations 逆推回去所得的 $\mathbf{b}$ 就會使得聯立方程組 $A\mathbf{x} = \mathbf{b}$ 無解. 也就是說當 $r < m$ 時一定存在 $\mathbf{b} \in \mathbb{R}^m$ 使得 $A\mathbf{x} = \mathbf{b}$ 無解. 所以我們有興趣於探討那些 $\mathbf{b} \in \mathbb{R}^m$ 可以使得 $A\mathbf{x} = \mathbf{b}$ 有解, 那些會使得 $A\mathbf{x} = \mathbf{b}$ 無解.

給定 $m \times n$ matrix A, 如何決定哪些 $\mathbf{b}$ 會使得 $A\mathbf{x} = \mathbf{b}$ 有解, 那些會使得 $A\mathbf{x} = \mathbf{b}$ 無解呢? 我們可以將 $\mathbf{b} = \begin{bmatrix} b_1 \\ \vdots \\ b_m \end{bmatrix}$ 視為 $\mathbb{R}^m$ 中的未知向量, 然後考慮 $[A \mid \mathbf{b}]$ 這一個增廣矩陣. 此時利用 elementary row operations 將 A 化成 echelon form A', 同時便把 $[A \mid \mathbf{b}]$ 化成了 $[A' \mid \mathbf{b}']$. 注意由於我們是利用 elementary row operations, 此時 $\mathbf{b}' = \begin{bmatrix} b'_1 \\ \vdots \\ b'_m \end{bmatrix}$ 上每一個位置 b'_i 都是以 $b_1, \ldots, b_m$ 為未知數的一次多項式. 由於假設 $\mathrm{rank}(A) = r$, echelon form A' 的前 r 個 row 皆不全為 0, 而 A' 最後 $m - r$ 個 row 皆全為 0. 因此要使得聯立方程組 $A\mathbf{x} = \mathbf{b}$ 有解, 由 2.1 節 (2) 的情形知, 等同於 $\mathbf{b}'$ 最後 $m - r$ 個位置 $b'_{r+1}, \ldots, b'_m$ 應為 0. 也就是說要使的 $A\mathbf{x} = \mathbf{b}$ 有解, 就等同於 $\mathbf{b} = \begin{bmatrix} b_1 \\ \vdots \\ b_m \end{bmatrix}$ 會滿足 $b'_{r+1}, \ldots, b'_m$ 所代表的 $m - r$ 個以 $b_1, \ldots, b_m$ 為未知數的一次多項式皆為 0. 換言之, 使得聯立方程組 $A\mathbf{x} = \mathbf{b}$ 有解的 $\mathbf{b}$ 必須滿足一組有 $m - r$ 個以 $b_1, \ldots, b_m$ 為未知數的 m 元一次方程組. 我們有以下的結論.

Proposition 2.4.1. *給定一 $m \times n$ matrix A 且假設 $\mathrm{rank}(A) = r$. 考慮以 A 為係數矩陣的聯立方程組.*

(1) *$r = m$ 若且唯若對任意 $\mathbf{b} \in \mathbb{R}^m$, 聯立方程組 $A\mathbf{x} = \mathbf{b}$ 一定有解.*

(2) *若 $r < m$, 則存在一個 $(m - r) \times m$ 的 matrix B, 使得 $A\mathbf{x} = \mathbf{b}$ 有解若且唯若 $\mathbf{y} = \mathbf{b}$ 為 $B\mathbf{y} = \mathbf{O}$ 的一組解.*

由於在 Proposition 2.4.1 (2) 中的聯立方程組 $B\mathbf{y} = \mathbf{O}$ 的一組解 $\mathbf{y} = \mathbf{b}$ 會使得聯立方程組 $A\mathbf{x} = \mathbf{b}$ 有解, 反之當 $A\mathbf{x} = \mathbf{b}$ 有解時, $\mathbf{y} = \mathbf{b}$ 就會是聯立方程組 $B\mathbf{y} = \mathbf{O}$ 的一組解. 因此我們稱此聯立方程組 $B\mathbf{y} = \mathbf{O}$ 為以 A 為係數矩陣的聯立方程組的 *constrain equations*. 要注意給定 A 其所對應的聯立方程組的 constrain equations 並不唯一, 但是不管是哪一組 constrain equations, 它們的解集合都會相同, 就是那些可以使的 $A\mathbf{x} = \mathbf{b}$ 為 consistent 的 $\mathbf{b}$ 所成的集合. 另外要注意 $r < m$ 時 Proposition 2.4.1 (2) 中的 constrain equations $B\mathbf{y} = \mathbf{O}$ 中 B 一定不是零矩陣 (即 B 的 entry 不全為 0), 否則會造成每個 $\mathbf{y} = \mathbf{b}$ 皆為 $B\mathbf{y} = \mathbf{O}$ 的一組解. 也就是說此時所有的 $\mathbf{b} \in \mathbb{R}^m$ 皆會使得聯立方程組 $A\mathbf{x} = \mathbf{b}$ 有解. 此與 $r < m$ 相矛盾.

Example 2.4.2. 考慮 $A=\begin{bmatrix}1&-1&1\\3&2&-1\\1&4&-3\\3&-3&3\end{bmatrix}$ 我們要找到以 A 為係數矩陣的聯立方程組的 constrain equations. 利用 elementary row operation, 我們可以得到 $\mathrm{rank}(A)=2$, 所以 A 的 constrain equations 是一個 $(4-2)\times 2$ 的矩陣所形成的聯立方程組.

首先考慮 augmented matrix $\left[\begin{array}{ccc|c}1&-1&1&b_1\\3&2&-1&b_2\\1&4&-3&b_3\\3&-3&3&b_4\end{array}\right]$. 將 1-st row 分別乘上 -3, -1, -3 加到 2-nd, 3-rd, 4-th rows 得 $\left[\begin{array}{ccc|c}1&-1&1&b_1\\0&5&-4&b_2-3b_1\\0&5&-4&b_3-b_1\\0&0&0&b_4-3b_1\end{array}\right]$. 最後將 2-nd row 乘上 -1 加到 3-rd row 得

$$\left[\begin{array}{ccc|c}1&-1&1&b_1\\0&5&-4&b_2-3b_1\\0&0&0&b_3-b_2+2b_1\\0&0&0&b_4-3b_1\end{array}\right].$$

由此知 $\mathbf{b}=\begin{bmatrix}b_1\\b_2\\b_3\\b_4\end{bmatrix}$ 會使得 $A\mathbf{x}=\mathbf{b}$ 為 consistent 若且唯若 b_1,b_2,b_3,b_4 會滿足 constrain equations

$$\begin{array}{rrrrcl}2b_1&-b_2&+b_3&&=&0\\-3b_1&&&+b_4&=&0\end{array}.$$

所以令 $B=\begin{bmatrix}2&-1&1&0\\-3&0&0&1\end{bmatrix}$, 此 2×4 matrix 所形成的聯立方程式 $B\mathbf{y}=\mathbf{O}$ 就是以 A 為係數矩陣的聯立方程組的 constrain equations.

我們可以再解 $B\mathbf{y}=\mathbf{O}$ 得其解集合為 $\begin{bmatrix}y_1\\y_2\\y_3\\y_4\end{bmatrix}=r\begin{bmatrix}0\\1\\1\\0\end{bmatrix}+s\begin{bmatrix}1\\2\\0\\3\end{bmatrix}$, $r,s\in\mathbb{R}$. 也就是說會使得 $A\mathbf{x}=\mathbf{b}$ 為 consistent 的 $\mathbf{b}$ 所成的集合為

$$\mathrm{Span}(\begin{bmatrix}0\\1\\1\\0\end{bmatrix},\begin{bmatrix}1\\2\\0\\3\end{bmatrix}).$$

Question 2.5. *假設 A 為 $m\times n$ matrix. 試說明當 $n<m$ 時, 一定存在 $\mathbf{b}\in\mathbb{R}^m$ 使得聯立方程組 $A\mathbf{x}=\mathbf{b}$ 為 inconsistent. 也就是當未知數個數少於方程式個數時, 一定存在 $\mathbf{b}$ 使得 $A\mathbf{x}=\mathbf{b}$ 無解.*

我們曾提過聯立方程組和 span 是有關的. 例如要探討 $\begin{bmatrix}3\\2\\1\end{bmatrix}$ 是否可以寫成 $\begin{bmatrix}1\\2\\3\end{bmatrix}$, $\begin{bmatrix}1\\0\\1\end{bmatrix}$ 的 linear combination (注意我們故意寫成 column vector)? 就是問是否存在 $x_1,x_2\in\mathbb{R}$ 使得

$\begin{bmatrix}3\\2\\1\end{bmatrix} = x_1\begin{bmatrix}1\\2\\3\end{bmatrix} + x_2\begin{bmatrix}1\\0\\1\end{bmatrix}$? 亦即是否聯立方程組

$$\begin{array}{rcrcl} x_1 & + & x_2 & = & 3 \\ 2x_1 & & & = & 2 \\ 3x_1 & + & x_2 & = & 1 \end{array}$$

有解? 所以若將 $m\times n$ 矩陣 A 的 n 個 columns 視為 $\mathbb{R}^m$ 的 n 個 column vectors, 且令其第一個 column 為 $\mathbf{v}_1$, 第二個 column 為 $\mathbf{v}_2$, ..., 第 n 個 column 為 $\mathbf{v}_n$, 則聯立方程組 $A\mathbf{x}=\mathbf{b}$ 有解就等同於 $\mathbf{b}$ 可以寫成 $\mathbf{v}_1,\ldots,\mathbf{v}_n$ 的 linear combination (亦即 $\mathbf{b}\in\mathrm{Span}(\mathbf{v}_1,\ldots,\mathbf{v}_n)$). 反之, 聯立方程組 $A\mathbf{x}=\mathbf{b}$ 無解就表示 $\mathbf{b}$ 不可以寫成 $\mathbf{v}_1,\ldots,\mathbf{v}_n$ 的 linear combination (關於這個事實, 等到介紹矩陣的乘法後, 就會更清楚). 因此若要測試許多 $\mathbb{R}^m$ 上的向量 $\mathbf{w}_1,\ldots,\mathbf{w}_k$ 哪些在 $\mathrm{Span}(\mathbf{v}_1,\ldots,\mathbf{v}_n)$ 中, 我們不必解 k 個 $A\mathbf{x}=\mathbf{w}_i$, $i=1,\ldots,k$ 聯立方程組. 只要將係數矩陣為 A 的 constrain equations 求出, 再代入這些 $\mathbf{w}_i$ 看那些符合即可.

Example 2.4.3. 令 $\mathbf{v}_1=\begin{bmatrix}1\\3\\1\\3\end{bmatrix}$, $\mathbf{v}_2=\begin{bmatrix}-1\\2\\4\\-3\end{bmatrix}$, $\mathbf{v}_3=\begin{bmatrix}1\\-1\\-3\\3\end{bmatrix}$. 我們要檢查 $\mathbf{w}_1=\begin{bmatrix}1\\-2\\-4\\6\end{bmatrix}$, $\mathbf{w}_2=\begin{bmatrix}2\\-1\\-2\\6\end{bmatrix}$, $\mathbf{w}_3=\begin{bmatrix}3\\0\\-6\\9\end{bmatrix}$ 那些在 $\mathrm{Span}(\mathbf{v}_1,\mathbf{v}_2,\mathbf{v}_3)$ 中. 事實上以 $\mathbf{v}_1,\mathbf{v}_2,\mathbf{v}_3$ 為 column vector 的矩陣就是 Example 2.4.2 中的矩陣 A. 故由 Example 2.4.2 的結果, 我們知以 A 為係數矩陣的聯立方程組的 constrain equations 為

$$\begin{array}{rrrrcl} 2y_1 & -y_2 & +y_3 & & = & 0 \\ -3y_1 & & & +y_4 & = & 0 \end{array}.$$

分別代 $\mathbf{w}_1,\mathbf{w}_2,\mathbf{w}_3$ 檢查, 僅 $\mathbf{w}_3$ 符合. 故僅有 $\mathbf{w}_3$ 屬於 $\mathrm{Span}(\mathbf{v}_1,\mathbf{v}_2,\mathbf{v}_3)$. 另一方面由於 $A\mathbf{x}=\mathbf{b}$ 為 consistent 等同於 $\mathbf{b}\in\mathrm{Span}(\mathbf{v}_1,\mathbf{v}_2,\mathbf{v}_3)$, 因此由 Example 2.4.2 知

$$\mathrm{Span}(\begin{bmatrix}1\\3\\1\\3\end{bmatrix},\begin{bmatrix}-1\\2\\4\\-3\end{bmatrix},\begin{bmatrix}1\\-1\\-3\\3\end{bmatrix}) = \mathrm{Span}(\begin{bmatrix}0\\1\\1\\0\end{bmatrix},\begin{bmatrix}1\\2\\0\\3\end{bmatrix}).$$

前面我們提及 parametric equations 換成 Cartesian equations 都是利用 normal vectors. 其實我們可以利用 constrain equations, 直接求出 Cartesian equation. 這是因為若 $Q(x_1,\ldots,x_m)$ 是 m 維坐標空間之一點, 它落在一個通過點 $P(p_1,\ldots,p_m)$ 且方向向量為 $\mathbf{v}_1,\ldots,\mathbf{v}_n$ 的集合中, 則 $\overrightarrow{PQ}=\begin{bmatrix}x_1-p_1\\\vdots\\x_m-p_m\end{bmatrix}$ 會落在 $\mathrm{Span}(\mathbf{v}_1,\ldots,\mathbf{v}_n)$ 中. 因此若考慮 A 是以

$\mathbf{v}_1,\ldots,\mathbf{v}_n$ 為 n 個 column 的 $m\times n$ matrix, 則 $\begin{bmatrix} x_1-p_1\\ \vdots \\ x_m-p_m\end{bmatrix}$ 應符合以 A 為係數矩陣的 constrain equations. 所以將 $\begin{bmatrix} x_1-p_1\\ \vdots \\ x_m-p_m\end{bmatrix}$ 代入 constrain equations, 便可得到此集合的 Cartesian equation. 我們看以下的例子.

Example 2.4.4. 在 Example 1.5.2 中我們考慮坐標空間上通過 $P(0,1,1)$ 且 direction vectors 為 $\mathbf{v}_1=\begin{bmatrix}2\\1\\0\end{bmatrix},\mathbf{v}_2=\begin{bmatrix}-5\\0\\1\end{bmatrix}$ 的平面 (注意我們故意將向量寫成 column vector). 若 $Q(t_1,t_2,t_3)$ 在此平面上, 則向量 $\overrightarrow{PQ}=\begin{bmatrix}t_1-0\\t_2-1\\t_3-1\end{bmatrix}$ 需落在 $\mathrm{Span}(\mathbf{v}_1,\mathbf{v}_2)$ 上. 因此考慮以 $\mathbf{v}_1,\mathbf{v}_2$ 為 column vector 的 3×2 matrix $A=\begin{bmatrix}2&-5\\1&0\\0&1\end{bmatrix}$. 由於 $\overrightarrow{PQ}$ 需使得聯立方程組 $A\mathbf{x}=\overrightarrow{PQ}$ 有解, 故考慮 augmented matrix

$$\left[\begin{array}{cc|c}2&-5&t_1\\1&0&t_2-1\\0&1&t_3-1\end{array}\right].$$

經過交換 row 的 elementary row operations 得

$$\left[\begin{array}{cc|c}1&0&t_2-1\\0&1&t_3-1\\2&-5&t_1\end{array}\right].$$

再分別將 1-st row, 2-nd row 乘上 -2, 5 加到 3-rd row 得

$$\left[\begin{array}{cc|c}1&0&t_2-1\\0&1&t_3-1\\0&0&t_1-2(t_2-1)+5(t_3-1)\end{array}\right].$$

故點 $Q(t_1,t_2,t_3)$ 需符合 $t_1-2t_2+5t_3=3$ 和 Example 1.5.2 利用 normal vector 所得的 Cartesian equation $x-2y+5z=3$ 結果一致.

接下來我們探討 rank 和解的個數的關係. 假設 A 為 $m\times n$ matrix 且 $\mathrm{rank}(A)=r$. 當 $r=n$ 時, 表示 $x_1\ldots,x_n$ 這 n 個 variable 皆為 pivot variable. 此時 $A\mathbf{x}=\mathbf{b}$ 是可能無解的 (即 2.1 節 (2)(a) 的情形); 不過若 $A\mathbf{x}=\mathbf{b}$ 有解, 則由於沒有 free variable, $x_1,\ldots,x_n$ 的取值都是固定的, 所以 $A\mathbf{x}=\mathbf{b}$ 的解便會是唯一. 反之當 $r<n$, 由於有 $n-r$ 個 free variables, 當 $A\mathbf{x}=\mathbf{b}$ 為 consistent 時, 這些 free variables 任意取值都能得到一組 $A\mathbf{x}=\mathbf{b}$ 的解, 故知有無窮多解. 由此我們可得以下的結果.

Proposition 2.4.5. *給定一 $m\times n$ matrix A. 假設 $\mathrm{rank}(A)=r$ 且假設聯立方程組 $A\mathbf{x}=\mathbf{b}$ 為 consistent.*

(1) *若 $r=n$, 則聯立方程組 $A\mathbf{x}=\mathbf{b}$ 有唯一解.*

(2) 若 $r<n$, 則聯立方程組 $A\mathbf{x}=\mathbf{b}$ 有無窮多解.

Question 2.6. 假設 A 為 $m\times n$ *matrix.* 試說明當 $n>m$ 時, 對於 $\mathbf{b}\in\mathbb{R}^m$ 聯立方程組 $A\mathbf{x}=\mathbf{b}$ 不可能僅有唯一解. 也就是當未知數個數多於方程式個數時, 聯立方程組 $A\mathbf{x}=\mathbf{b}$ 要不然無解; 要不然會有無窮多解.

特別的, 由於 homogeneous system $A\mathbf{x}=\mathbf{O}$ 一定是 consistent (因為 $x_1=\cdots=x_n=0$ 就是一組解), 我們有以下的結果.

Corollary 2.4.6. 給定一 $m\times n$ *matrix* A. 則 $A\mathbf{x}=\mathbf{O}$ 有唯一解 $x_1=\cdots=x_n=0$ 若且唯若 $\operatorname{rank}(A)=n$.

一個 homogeneous system $A\mathbf{x}=\mathbf{O}$ 一定有 $x_1=0,\ldots,x_n=0$ 這一組解, 我們稱這組解為 $A\mathbf{x}=\mathbf{O}$ 的 *trivial solution*. 而除了 $x_1=0,\ldots,x_n=0$ 以外 $A\mathbf{x}=\mathbf{O}$ 若還有其他的解, 我們便稱那些解為 $A\mathbf{x}=\mathbf{O}$ 的 *nontrivial solution*. 所以 $A\mathbf{x}=\mathbf{O}$ 的解唯一就等同於它沒有 nontrivial solution; 而解不唯一就等同於它有 nontrivial solution.

Question 2.7. 假設 A 為 $m\times n$ *matrix.* 試說明 $A\mathbf{x}=\mathbf{O}$ 有 *nontrivial solution* 若且唯若 $\operatorname{rank}(A)\neq n$.

事實上當 $A\mathbf{x}=\mathbf{b}$ 是 consistent 時, $A\mathbf{x}=\mathbf{b}$ 的解和 $A\mathbf{x}=\mathbf{O}$ 的解是息息相關的. 具體來說, 若 $x_1=c_1,\ldots,x_n=c_n$ 和 $x_1=c_1',\ldots,x_n=c_n'$ 皆為聯立方程組 $A\mathbf{x}=\mathbf{b}$ 的一組解. 則由

$$\left\{\begin{array}{ccccccc} a_{11}c_1 & + & \cdots & + & a_{1n}c_n & = & b_1\\ a_{21}c_1 & + & \cdots & + & a_{2n}c_n & = & b_2\\ & & \vdots & & & & \\ a_{m1}c_1 & + & \cdots & + & a_{mn}c_n & = & b_m \end{array}\right. \quad \text{以及} \quad \left\{\begin{array}{ccccccc} a_{11}c_1' & + & \cdots & + & a_{1n}c_n' & = & b_1\\ a_{21}c_1' & + & \cdots & + & a_{2n}c_n' & = & b_2\\ & & \vdots & & & & \\ a_{m1}c_1' & + & \cdots & + & a_{mn}c_n' & = & b_m \end{array}\right.$$

將相對應的式子相減得

$$\begin{array}{ccccccc} a_{11}(c_1-c_1') & + & \cdots & + & a_{1n}(c_n-c_n') & = & 0\\ a_{21}(c_1-c_1') & + & \cdots & + & a_{2n}(c_n-c_n') & = & 0\\ & & \vdots & & & & \\ a_{m1}(c_1-c_1') & + & \cdots & + & a_{mn}(c_n-c_n') & = & 0 \end{array}.$$

也就是說 $x_1=c_1-c_1',\ldots,x_n=c_n-c_n'$ 就會是 $A\mathbf{x}=\mathbf{O}$ 的一組解. 同樣的道理, 若 $x_1=c_1,\ldots,x_n=c_n$ 為 $A\mathbf{x}=\mathbf{b}$ 的一組解而 $x_1=u_1,\ldots,x_n=u_n$ 為 $A\mathbf{x}=\mathbf{O}$ 的一組解, 則 $x_1=c_1+u_1,\ldots,x_n=c_n+u_n$ 也會是 $A\mathbf{x}=\mathbf{b}$ 的一組解. 因此我們有以下的結論.

Proposition 2.4.7. 假設聯立方程組 $A\mathbf{x}=\mathbf{b}$ 為 *consistent* 且 $\mathbf{x}=\mathbf{c}$ 為其一組解. 則 $\mathbf{x}=\mathbf{c}'$ 為 $A\mathbf{x}=\mathbf{b}$ 的一組解若且唯若 $\mathbf{x}=\mathbf{c}-\mathbf{c}'$ 為 $A\mathbf{x}=\mathbf{O}$ 的一組解.

利用 Proposition 2.4.7, 我們馬上有以下之結果.

Corollary 2.4.8. 假設 A 為 $m\times n$ *matrix.* 下列敘述是等價的.

(1) 若聯立方程組 $A\mathbf{x}=\mathbf{b}$ 為 *consistent* 則 $A\mathbf{x}=\mathbf{b}$ 的解唯一.

(2) $A\mathbf{x} = \mathbf{O}$ 沒有 *nontrivial solution.*

(3) $\text{rank}(A) = n$.

Proof. 要證明一系列 (多於兩個) 的敘述是等價, 我們不必證明每兩個敘述皆為等價. 只要證明 $(1) \Rightarrow (2)$, $(2) \Rightarrow (3)$ 以及 $(3) \Rightarrow (1)$ 即可. 這並沒有少證到其他部分, 例如 $(2) \Rightarrow (1)$ 就可由 $(2) \Rightarrow (3)$ 以及 $(3) \Rightarrow (1)$ 推得. 這是一個相當簡便的方法來證明一系列的 statement 是等價的, 也是在數學證明中常用的技巧.

$(1) \Rightarrow (2)$: 利用反證法, 假設 $x_1 = u_1, \ldots, x_n = u_n$ 為 $A\mathbf{x} = \mathbf{O}$ 的一組 nontrivial solution 而 $x_1 = c_1, \ldots, x_n = c_n$ 為 $A\mathbf{x} = \mathbf{b}$ 的一組解, 則由 Proposition 2.4.7 知 $x_1 = c_1 + u_1, \ldots, x_n = c_n + u_n$ 會是 $A\mathbf{x} = \mathbf{b}$ 的另一組解. 此與 $A\mathbf{x} = \mathbf{b}$ 的解唯一相矛盾, 故知 $A\mathbf{x} = \mathbf{O}$ 沒有 nontrivial solution.

$(2) \Rightarrow (3)$: $A\mathbf{x} = \mathbf{O}$ 沒有 nontrivial solution, 表示 $A\mathbf{x} = \mathbf{O}$ 的解唯一, 故由 Corollary 2.4.6 知 $\text{rank}(A) = n$.

$(3) \Rightarrow (1)$: 此為 Proposition 2.4.5 (1). □

最後我們綜合上面的結果處理當 A 為 $n \times n$ 的方陣的情形.

Corollary 2.4.9. 假設 A 為 $n \times n$ *square matrix.* 下列敘述是等價的.

(1) 對於任意 $\mathbf{b} \in \mathbb{R}^n$, $A\mathbf{x} = \mathbf{b}$ 一定有解.

(2) 對於任意 $\mathbf{b} \in \mathbb{R}^n$, $A\mathbf{x} = \mathbf{b}$ 一定有解且解唯一.

(3) 存在 $\mathbf{b} \in \mathbb{R}^n$ 使得 $A\mathbf{x} = \mathbf{b}$ 有解且解唯一.

(4) $A\mathbf{x} = \mathbf{O}$ 沒有 *nontrivial solution.*

(5) $\text{rank}(A) = n$.

Proof. 要證明一系列的敘述是等價, 不一定要按照定理敘述的順序證明. 有時也可以挑選較易證明的順序來證明. 不過要注意證明的路徑一定要形成一個回圈這樣證明才完整. 這裡我們要證明的路徑為 $(1) \Rightarrow (5)$, $(3) \Leftrightarrow (4) \Leftrightarrow (5)$, $(5) \Rightarrow (2)$ 以及 $(2) \Rightarrow (1)$.

$(1) \Rightarrow (5)$: 由 Proposition 2.4.1 (1), 我們知對於任意 $\mathbf{b} \in \mathbb{R}^n$, $A\mathbf{x} = \mathbf{b}$ 一定有解表示 $\text{rank}(A)$ 等於 A 的 row 的個數, 故得 $\text{rank}(A) = n$.

$(3) \Leftrightarrow (4) \Leftrightarrow (5)$: 此為 Corollary 2.4.8 中三個皆等價的情形.

$(5) \Rightarrow (2)$: 當 $\text{rank}(A) = n$ 表示 $\text{rank}(A)$ 等於 A 的 row 的個數, 故由 Proposition 2.4.1 (1) 知對於任意 $\mathbf{b} \in \mathbb{R}^n$, $A\mathbf{x} = \mathbf{b}$ 一定有解. 再利用 Corollary 2.4.8 $(3) \Rightarrow (1)$ 得證其解唯一.

$(2) \Rightarrow (1)$: (2) 的條件比 (1) 強, 此敘述當然成立. □

一般我們會將符合 Corollary 2.4.9 的 $n \times n$ matrix (即 rank 為 n) 稱為 *nonsingular matrix*, 而 rank 小於 n 的 $n \times n$ matrix 稱為 *singular matrix.*

2.5. 結論

在本章中我們學習解聯立方程組的技巧. 利用 elementary row operations 將 augmented matrix 中的係數矩陣化為 echelon form 後, 我們很快的可以知道此聯立方程組是否有解, 而有解時也可利用此 echelon form 完整的得到此聯立方程組所有的解. 由 echelon form 的解法我們了解到 pivot 對聯立方程組是否有解以及解是否唯一有著重要的關連. 在下一章, 我們將介紹矩陣的運算, 這些運算都可以和解聯立方程組的問題相連結. 所以本章中有關聯立方程組的理論對後面理論的建立影響深遠, 千萬不要以為會解具體的聯立方程組就可以了, 而忽視這些理論.