Chapter 1

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.

1.1. 一次聯立方程組及基本列運算

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

$$\begin{aligned} 3x_1-2x_2+9x_4 &= 4 \\ 2x_1+2x_2-4x_4 &= 6 \end{aligned} \tag{1.1}$$

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.

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. 我們將介紹一個有系統的方法來解聯立方程組, 把這三種基本方法看成是對矩陣的運算.

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 (1.1) 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.

為了方便起見, 我們將上面 (1), (2), (3) 三種 elementary row operation 分別稱為 *type 1*, *type 2* 以及 *type 3* 的 elementary row operation.

Example 1.1.1. 考慮矩陣

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

將 A 的第一、第二兩個 row 互換 (即做一個 type 1 的 elementary row operation), 可得

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

而將 B 的第二個 row 乘上 2 (即做一個 type 2 的 elementary row operation), 可得

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

而將 C 的第三個 row 乘上 -3 加到第一個 row (即做一個 type 3 的 elementary row operation), 可得

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

注意, 若一個矩陣 P 經由一個 elementary row operation 轉換成矩陣 Q, 我們也可以對 Q 藉由同樣 type 的 elementary row operation 將之轉換回 P. 例如前面 Example 1.1.1 中, 我們可以將 B 的第一、第二兩個 row 互換而得到 A. 我們也可將 C 的第二個 row 乘上 $1/2$ 而得到 B. 另外我們也可將 D 的第三個 row 乘上 3 加到第一個 row 而轉換回 C.

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

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

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

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

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

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

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.

當 A 是一個 $m \times n$ matrix. 要對 A 做一個 type 1 的 elementary row operation 得到矩陣 B, 我們可以先可慮 type 1 elementary matrix E_1, 其中 E_1 是將 $m \times m$ 的 identity matrix I_m 做同樣的 type 1 elementary row operation 所得的 elementary matrix. 將來當我們更深入介紹矩陣乘法後可以驗證 B 就會是 $E_1 A$. 同理對 A 做 type 2, type 3 的 elementary row operation 就是將 A 左邊乘上其所對應的 elementary matrix.

Example 1.1.2. 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

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

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

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

可得

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

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

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

可得

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

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

(1) 設 E_1 是將 I_m 的 i-th row 和 j-th row 互換的 type 1 elementary matrix. 我們將 E_1 的 i-th row 和 j-th row 再互換就可轉換回 identity matrix I_m. 所以我們有 $E_1 E_1 = I_m$.

(2) 設 E_2 是將 I_m 的 i-th row 乘上非零實數 r 的 type 2 elementary matrix. 我們將 E_2 的 i-th row 乘上 r^{-1} 就可轉換回 I_m. 所以若令 E_2' 為將 I_m 的 i-th row 乘上 r^{-1} 的 type 2 elementary matrix, 我們有 $E_2'E_2 = I_m$. 同理可得 $E_2E_2' = I_m$.

(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 1.1.3. *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.*

在這一節的最後我們要說明一下, 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. 這個部分, 等到以後介紹矩陣乘法時, 我們會有更進一步的說明. 目前也請自行找例子驗證. 為了方便起見, 我們綜合成以下的結論.

Theorem 1.1.4. *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.*

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.

1.2. 解聯立方程組

大家應很容易看出一個 augmented matrix 經過上節所提的 elementary row operation 後所得的 augmented matrix 所對應的聯立方程組就是大家熟悉的加減消去法的三種步驟所得的方程組. 利用加減消去法最常遇到的問題就是 (尤其在處理未知數很多的方程組時), 常常做了幾次後, 混亂到不知道那些式子是消過了以及那些式子還可以再進一步消減. 還有就是, 到底要將方程組的式子消到哪種地步時, 才可以解出方程組. 關於第一個問題, 我們可以理解用矩陣的表示法就可以把這消去的過程記錄下來. 而接下來我們要探討的就是第二個問題, 也就是將矩陣化成某種特定的形式就可以解出方程式來.

我們的目的就是要將 augmented matrix $[A \mid \mathbf{b}]$ 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 (未知數) 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 還有其他位置也是 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 $[A \mid \mathbf{b}]$ 利用 elementary row operation 將之化成 $[A' \mid \mathbf{b}']$ 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

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

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

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

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.

我們總結一下, 當 A 是 echelon form 時, 如何找出 $A\mathbf{x} = \mathbf{b}$ 所有的解的方法. 很容易看出當 A 有一個 row 全為 0 但 $\mathbf{b}$ 在該 row 不為 0 時該方程組無解, 我們僅討論其他有解的情況. 此時我們先挑出 free variable (即非 pivot variable). 由於 free variable 可以任意取值, 一般來說我們會用一些參數表示之 (注意不同的 free variable 要用不同的參數代號). 接著, 我們由下而上, 從最大編號的 pivot variable 開始, 利用 free variables 的那些參數將它的值寫下來, 再依序直到寫出所有 pivot variables 的值.

另外, 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 1.1. *考慮一個由 n 個 variables 的 m 個方程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;*

我們看以下幾個解聯立方程組的例子.

Example 1.2.1. 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 1.2.2. Solve the linear system

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

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 1.2.3. Solve the linear system

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

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 column vector 且將 r, s 提出. 故將解寫成

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

了解到解聯立方程組的方法及步驟後, 有幾件事必須要說明: (1) 為何經由 elementary row operations 我們可以將一個矩陣化為 echelon form? (2) 為何利用這個 echelon form, 便可得到與原方程組相同的解集合? (3) 為什麼用前面介紹 (pivot variables, free variables) 的方法就可以把係數矩陣是 echelon form 的聯立方程組所有的解找出來? 在下一節, 我們將詳細介紹有關 echelon form 的特性, 然後一一回答這些問題. 不過再次提醒大家務必先熟悉這節介紹解聯立方程組的方法及步驟.

1.3. Echelon Form

這一節中我們將說明前面提到有關 echelon form 的三個問題. 首先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. 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 $4, 5, 6, \ldots$ 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.

我們先看只有一個 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 b 所在位置和第一個 row 的 leading entry a 相同, 我們可將第一個 row 乘以 $-b/a$, 再加到第二個 row 上. 如此一來第二個 row 原本的 leading entry 變為 0, 故其 leading entry 所在位置往右移了, 依定義此時為 echelon form.

接著我們使用數學歸納法的假設, 亦即任何有 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. 現假設此時第一個 row 的 leading entry 為 a. 接下來我們挑出其他 row 中 leading entry 的位置與第一個 row 的 leading entry 位置一樣的 row. 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 type 2 的 elementary row operation. 事實上在化成 echelon form 的過程確實只需要用到 type 1,3 這兩種 elementary row operations, 至於 type 2 的 elementary row operation 會在以後我們會介紹化為 “reduced” echelon form 的過程是需要的, 留待以後再談.

接下來我們說明為何將 augmented matrix $[A\,|\,\mathbf{b}]$ 利用 elementary row operations 化成 echelon form $[A'\,|\,\mathbf{b}']$, 則其對應的聯立方程組 $A'\mathbf{x}=\mathbf{b}'$ 會和原方程組 $A\mathbf{x}=\mathbf{b}$ 有相同的解集合. 首先觀察若將一聯立方程組 $A\mathbf{x}=\mathbf{b}$ 的 augmented matrix $[A\,|\,\mathbf{b}]$ 利用三種 elementary row operation 的任一種變換成 $[A'\,|\,\mathbf{b}']$ 表示將原方程組利用加減消去法的三個基本方法將之變成方程組 $A'\mathbf{x}=\mathbf{b}'$. 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 $A'\mathbf{x}=\mathbf{b}'$. 也就是說 $A\mathbf{x}=\mathbf{b}$ 的解就會是 $A'\mathbf{x}=\mathbf{b}'$ 的解. 不過這不表示它們會有相同的解集合, 我們還要說明 $A'\mathbf{x}=\mathbf{b}'$ 的解也會是 $A\mathbf{x}=\mathbf{b}$ 的解才行. 然而我們前面提及 elementary row operations 是可以還原的. 換句話說 $[A'\,|\,\mathbf{b}']$ 也可經由 elementary row operations 變換成 $[A\,|\,\mathbf{b}]$. 所以套用剛才的理由, 我們也知 $A'\mathbf{x}=\mathbf{b}'$ 的解就會是 $A\mathbf{x}=\mathbf{b}$ 的解. 因此得證 $A\mathbf{x}=\mathbf{b}$ 和 $A'\mathbf{x}=\mathbf{b}'$ 會有相同的解集合.

當連立方程組 $A\mathbf{x}=\mathbf{b}$ 和 $A'\mathbf{x}=\mathbf{b}'$ 它們的解集合相同, 這表示兩組方程組是有很特別的關係的. 我們有以下的定義.

Definition 1.3.1. 假設 linear systems $A\mathbf{x}=\mathbf{b}$ 和 $A'\mathbf{x}=\mathbf{b}'$ 的解集合相同, 則稱 $A\mathbf{x}=\mathbf{b}$ 和 $A'\mathbf{x}=\mathbf{b}'$ 為 *equivalent linear systems*

從上面的探討我們知道 augmented matrix $[A\mid\mathbf{b}]$ 若利用 elementary row operations 化成 $[A'\mid\mathbf{b}']$, 則 $A\mathbf{x}=\mathbf{b}$ 和 $A'\mathbf{x}=\mathbf{b}'$ 為 equivalent linear systems.

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. 接下來我們就是要討論當 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 1.2 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 1.2 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.

如果我們得到 1.2 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 1.2 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, 1.2 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 1.2 節所提的方法得到. 為了方便起見我們令 1.2 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 1.3.2. *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 $1\leq k\leq n-1$. 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 1.3.3. 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 $1\leq k\leq n-1$. 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 1.3.2 和 Lemma 1.3.3 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 1.3.2 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 1.3.3 知聯立方程組 $A\mathbf{x}=\mathbf{b}$ 會有無窮多解 (除非我們討論的數系僅有有限多個元素).

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 1.3.2 和 Lemma 1.3.3 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 $A\mathbf{x}=\mathbf{0}$ 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 1.3.4. 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 $A\mathbf{x}=\mathbf{0}$ 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 $[A\mid\mathbf{0}]$ 可以利用 elementary row operation 化為 $[A_1\mid\mathbf{0}]$ 及 $[A_2\mid\mathbf{0}]$. 換句話說聯立方程組 $A_1\mathbf{x}=\mathbf{0}$ 和 $A_2\mathbf{x}=\mathbf{0}$ 皆與聯立方程組 $A\mathbf{x}=\mathbf{0}$ 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 $A_1\mathbf{x}=\mathbf{0}$ 的解中 x_n 的取值是唯一的 (Lemma 1.3.2), 事實上 x_n 一定為 0; 但 $A_2\mathbf{x}=\mathbf{0}$ 的解中 x_n 的取值卻可以是任意的實數 (Lemma 1.3.3). 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 $A_1\mathbf{x}=\mathbf{0}$ 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 1.3.2); 另一方面 Lemma 1.3.3 告訴我們 $A_2\mathbf{x}=\mathbf{0}$ 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 $A_1\mathbf{x}=\mathbf{0},A_2\mathbf{x}=\mathbf{0}$ 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 1.3.5. 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 $\mathrm{rank}(A)=r$ 來表示.

Question 1.2. *假設矩陣 A 有 m 個 row 以及 n 個 column. 若 $\mathrm{rank}(A)=r$, 試說明 $r\leq\min\{m,n\}$.*

在解聯立方程組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 entry 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 entry 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 entry 便是 1 了 (這就是需要 type 2 的 elementary row operation 的地方). 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.

化為 reduced echelon form 後, 我們就可以利用前面由 echelon form 求解的方法求出聯立方程組的解. 由於 reduced echelon form 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 $B'\mathbf{x}=\mathbf{0}$ 為

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

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 $(x_1,x_2,x_3,x_4)=(0,-3t,t,t)=t(0,-3,1,1),t\in\mathbb{R}$.

我們知道每個矩陣皆可經由 elementary row operations 化為 echelon form. 而我們又知每個 echelon form 也可利用 elementary row operations 化為 reduced echelon form. 因此每

個矩陣皆可利用 elementary row operations 化為 reduced echelon form. 另外我們也知道將聯立方程組的 augmented matrix 做 elementary row operations 後所對應的聯立方程組會是 equivalent, 所以化為 reduced echelon form 所得的解集合也會和原方程組的解集合相同.

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* 或 *Gaussian elimination*, 而用 reduced echelon form 求解一般稱為 *Gauss-Jordan method*.

Example 1.3.6. Example 1.2.1 的 linear system, 化成 echelon form 後為

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

將第三 row 乘以 $-1/3$ 得

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

再將第三 row 分別乘以 $3,1$ 加到第二, 第一 row 得

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

接著將第二 row 乘以 -3 加到第一 row 得

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

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

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

且馬上看出解為 $(x_1,x_2,x_3)=(-1,4,3)$.

Example 1.2.3 的 linear system, 化成 echelon form 後為

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

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

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

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 $x_1=-14+3r-5s$.

我們可以套用 Proposition 1.3.4 的證明方法, 證明一個矩陣利用 elementary row operations 化為 reduced echelon form 其結果是唯一的.

Theorem 1.3.7. 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 $A\mathbf{x} = \mathbf{0}$ 這一組聯立方程組, 依假設聯立方程組 $A_1\mathbf{x} = \mathbf{0}$ 和 $A_2\mathbf{x} = \mathbf{0}$ 皆與聯立方程組 $A\mathbf{x} = \mathbf{0}$ 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 1.3.4, 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

$$x_k + a_{k+1}x_{k+1} + \cdots + a_n x_n = 0 \quad \text{與} \quad x_k + b_{k+1}x_{k+1} + \cdots + b_n x_n = 0 \tag{1.2}$$

其中存在 l 滿足 $k+1 \leq l \leq n$ 且 $a_l \neq b_l$. 若 x_l 為 pivot variable, 由於 A_1,A_2 皆為 reduced echelon form, 在這個 row 中, 其他的 pivot variable 對應的係數應為 0, 而導致 $a_l = b_l = 0$ 之矛盾, 故知 x_l 應為 free variable. 我們已知給定一組 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 $A_1\mathbf{x} = \mathbf{0}$ 與 $A_2\mathbf{x} = \mathbf{0}$ 的解分別為 $(x_1,\ldots,x_n) = (c_1,\ldots,c_n)$ 與 $(x_1,\ldots,x_n) = (d_1,\ldots,d_n)$. 注意, 若 x_j 為 pivot variable, 其中 $j > k$, 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, 也就是說式子 (1.2) 帶入這兩組解後分別為

$$c_k + a_l = 0 \quad \text{與} \quad d_k + b_l = 0.$$

又由於依假設 A_1,A_2 在 x_k 為 pivot 這一個 row 以下的每一個 row 都一致, 我們知對所有 $k+1 \leq i \leq n$ 皆有 $c_i = d_i$. 然而這兩組解皆為 $A\mathbf{x} = \mathbf{0}$ 的解且 x_k 為 pivot variable, 故由 Lemma 1.3.2 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 $A_1 = A_2$. □

利用化成 reduced echelon form 來解 linear system, 雖然步驟較多, 不過仍然有它的好處. 例如因為化成 echelon form 並不唯一, 所以有可能同一組 linear system 因化成 echelon form 求解, 寫下來的解集合的元素的表現“形式”會不同 (只是形式不同, 解集合是一樣的). 若化成 reduced echelon form 就不會這樣了, 因為它是唯一的, 所以大家寫下來的解集合的元素表示的“形式”是一致的. 另外若要判斷兩個矩陣是否可以用一些 elementary row operations 將其中一個換成另一個, 將這兩個矩陣化成 reduced echelon form 就可以了. 若它們化成 reduced echelon form 是一致的, 那當然表示它們是可以用一些 elementary row operations 將其中一個換成另一個, 而若不一致, 則由唯一性可知它們不可能用一些 elementary row operations 將其中一個換成另一個.

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 variables 和 free variables 對聯立方程組是否有解以及解是否唯一有著重要的關連. 本章中有關聯立方程組的理論對後面線性代數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.

Chapter 2

Vector Spaces

在這一章中, 我們利用大家熟悉的坐標平面中的向量, 將之推廣到所謂的 vector space (向量空間). Vector space 一種有特定代數結構的系統, 是線性代數中主要的探討對象.

2.1. 坐標平面中的向量

本節針對對抽象數學論述不熟悉的同學, 想利用大家熟悉坐標平面的向量慢慢引導進入狀況.

在平面中的向量我們可以用幾何的方式規定向量的加法及其倍數關係. 相信大家對這種定法已相當熟悉, 在這裡我們不再重複. 我們可以將平面坐標化, 這就是所謂的坐標平面. 這種在坐標平面中的向量, 我們都可用 (a,b) 來表示, 其中 $a,b \in \mathbb{R}$ (我們用 $\mathbb{R}$ 來表示所有實數所成的集合, 所以 $a,b \in \mathbb{R}$ 表示 a,b 皆為實數).

用坐標來表示一個向量 (即用 (a,b) 這種方法) 有許多好處, 例如大家很容易理解: 當兩個向量 (a,b) 和 (c,d) 相等時 (即 $(a,b) = (c,d)$), 這表示 $a = c$ 且 $b = d$; 坐標表示法的另一個好處是很容易幫助我們定義向量的加法 (*addition*) 以及係數積 (*scalar multiplication*).

Definition 2.1.1. 令 $\mathbf{u} = (a_1,a_2), \mathbf{v} = (b_1,b_2) \in \mathbb{R}^2$ 以及 $r \in \mathbb{R}$. 我們定義

$$\mathbf{u}+\mathbf{v} = (a_1+b_1, a_2+b_2) \quad \text{and} \quad r\mathbf{u} = (ra_1, ra_2).$$

這裡我們要強調, Definition 2.1.1 中所定義的加法及係數積, 和前面所述用幾何的方式定義所得的結果是一致的. 基於符號的方便性, 當我們要用符號來表示一個向量時, 會用 $\mathbf{u},\mathbf{v}$ 這類的粗體字符號來表示. 一般來說我們用 $\mathbb{R}^2$ 來表示坐標平面上的向量所成的集合, 所以若我們說 $\mathbf{v} \in \mathbb{R}^2$, 就表示 $\mathbf{v}$ 是坐標平面上的一個向量, 也就是說可以找到 $a,b \in \mathbb{R}$ 使得 $\mathbf{v} = (a,b)$.

一般來說有了定義之後, 我們就需依定義處理相關問題, 但通常直接依定義處理較繁複, 我們可已先依定義推導出一些性質, 利用這些性質簡化處理程序, 再處理更進一步的問題. 例如在微積分, 我們定義出一個函數在某一點的極限後, 若每次都得依定義處理極限問題論證起來很複雜; 但當我們利用定義推導出一些極限的性質後, 用這些性質處理極限問題就簡

單方便多了. 所以在定義之後我們會有一些定理 (Proposition 或 Theorem) 來論證一些依定義可得的性質, 以方便我們處理更進一步的問題. 以下就是要談向量加法及係數積有關的性質.

Proposition 2.1.2. 對於 $\mathbb{R}^2$ 上的向量, 我們有以下的性質:

(1) 對任意 $\mathbf{u},\mathbf{v}\in\mathbb{R}^2$, 皆有 $\mathbf{u}+\mathbf{v}=\mathbf{v}+\mathbf{u}$.

(2) 對任意 $\mathbf{u},\mathbf{v},\mathbf{w}\in\mathbb{R}^2$, 皆有 $(\mathbf{u}+\mathbf{v})+\mathbf{w}=\mathbf{u}+(\mathbf{v}+\mathbf{w})$.

(3) 存在一向量 $\mathbf{0}\in\mathbb{R}^2$ 滿足對任意 $\mathbf{u}\in\mathbb{R}^2$ 皆有 $\mathbf{0}+\mathbf{u}=\mathbf{u}$.

(4) 對任意 $\mathbf{u}\in\mathbb{R}^2$ 皆可找到 $\mathbf{u}'\in\mathbb{R}^2$ 滿足 $\mathbf{u}+\mathbf{u}'=\mathbf{0}$.

(5) 對任意 $\mathbf{u}\in\mathbb{R}^2$, 皆有 $1\mathbf{u}=\mathbf{u}$.

(6) 對任意 $r,s\in\mathbb{R}$ 以及 $\mathbf{u}\in\mathbb{R}^2$, 皆有 $r(s\mathbf{u})=(rs)\mathbf{u}$.

(7) 對任意 $r\in\mathbb{R}$ 以及 $\mathbf{u},\mathbf{v}\in\mathbb{R}^2$ 皆有 $r(\mathbf{u}+\mathbf{v})=r\mathbf{u}+r\mathbf{v}$.

(8) 對任意 $r,s\in\mathbb{R}$ 以及 $\mathbf{u}\in\mathbb{R}^2$, 皆有 $(r+s)\mathbf{u}=r\mathbf{u}+s\mathbf{u}$.

通常一個定理敘述完就要證明, 不過這幾項的證明都僅是一般制式的代數操作, 相信大家都很熟悉, 這裡就不再證明了. 對同學來說了解定理說些什麼比起證明來得重要. 在這裡我們就一一說明一下這個定理說些什麼.

(1) 敘述的是所謂向量加法的交換性. 它告訴我們在處理向量加法時可以依方便交換順序. 或許同學覺得這個很自然為何還要證明. 事實上只要是定義未提的事情都要證明, 不能因為覺得自然而不去處理. 在證明時會發現這個性質會成立主要是實數加法有交換性. 不過數學上是存在許多“抽象”的數系它的運算是不能交換的. 所以經由證明不只讓我們確認事情是對的, 也能幫助我們釐清事情是對的其背後的主要因素.

(2) 說的就是所謂的結合律, 它依然是因為實數加法的性質而成立. 這裡 $(\mathbf{u}+\mathbf{v})+\mathbf{w}$ 是說先將 $\mathbf{u}$ 和 $\mathbf{v}$ 相加後所得的向量再和 $\mathbf{w}$ 相加. 這樣所得的向量和先將 $\mathbf{v}$ 和 $\mathbf{w}$ 相加後再和 $\mathbf{u}$ 相加會是同樣的向量. 因為向量的加法是定義兩個向量的加法, 所以兩個以上的向量相加結合律就顯得重要了. 有了結合律, 我們就不必擔心哪兩個向量先加. 結合律雖然也是談向量加法的順序問題, 不過和 (1) 所談的順序是兩回事, 大家應該要分清楚.

(3) 談的就是所謂的零向量, 零向量的特點就是加上任何向量都不動. 為什麼要特別談零向量的存在性? 這就好比在實數上若沒有零的概念就沒有減法一樣, 在向量的運算上是相當重要的. 尤其以後要用抽象的方式談向量系統時零向量的存在性更不容忽視.

(4) 談的就是所謂的反向量, 要注意需有零向量的存在才能談反向量. 而且要區分清楚這裡的敘述是給了 $\mathbf{u}$ 後可找到 $\mathbf{u}'$ 使得 $\mathbf{u}+\mathbf{u}'=\mathbf{0}$. 這裡 $\mathbf{u}'$ 是會隨著 $\mathbf{u}$ 而改變, 而不是一個固定的向量和所有的向量加起來會是零向量. 數學的敘述要弄清楚否則差之毫釐失之千里.

(5) 指的是所有向量乘上 1 後仍不動. 這裡特別提出來其實和零向量意義很像, 唯有 1 的引入以後才能談係數的除法. 例如已知 $2\mathbf{u}=\mathbf{v}$, 就可利用 (6) 的性質兩邊乘上 $1/2$, 得

$$\mathbf{u}=1\mathbf{u}=\frac{1}{2}(2\mathbf{u})=\frac{1}{2}\mathbf{v}.$$

(6),(7),(8) 談的是係數積的性質, 例如 $r(s\mathbf{u})$ 表示是先將 $\mathbf{u}$ 乘上 s 倍後所得的向量再乘上 r, 而 $(rs)\mathbf{u}$ 是表示先將 r,s 乘在一起得 rs 再乘上 $\mathbf{u}$. 這幾個性質也都和實數乘法性質息息相關, 雖然看起來不顯眼但在處理向量的運算時非常重要.

最後要強調一下: 這裡將這些性質列出, 並不是要求大家將這幾個性質背下來. 一來我們希望大家知道有些性質不能覺得理所當然就不去證明, 另一方面也讓大家知道以後在處理向量運算時可以放心且自然的使用這幾個性質. 這些性質也讓坐標平面上向量的系統享有許多豐富的性質.

Question 2.1. 利用 $\mathbb{R}^2$ 向量加法的定義, 試證明以下性質:

(1) $\mathbf{0}=(0,0)$ 是 $\mathbb{R}^2$ 中唯一的向量滿足對任意 $\mathbf{u}\in\mathbb{R}^2$ 皆有 $\mathbf{0}+\mathbf{u}=\mathbf{u}$.

(2) 給定 $\mathbf{u}=(a,b)\in\mathbb{R}^2$, 試證明 $\mathbf{u}'=(-a,-b)$ 是 $\mathbb{R}^2$ 中唯一的向量滿足 $\mathbf{u}+\mathbf{u}'=\mathbf{0}$.

坐標平面上向量的運算也可推廣到坐標空間, 即 $\mathbb{R}^3$. 同樣的概念也可推廣到更一般的 $\mathbb{R}^n$, $n\in\mathbb{N}$. 這些系統中的運算都享有 Proposition 2.1.2 的 8 項性質, 而這些性質讓這些系統有著豐富的性質. 所以接下來我們將專注於有這 8 項性質的系統, 稱之為 vector space (向量空間).

2.2. Vector Space 的定義及其基本性質

我們曾經提過像 $\mathbb{R}^2$ 這樣, 裡面任意兩個向量相加仍在 $\mathbb{R}^2$ 中且向量乘上任意的實數後也仍在 $\mathbb{R}^2$, 而向量的運算又符合 Proposition 2.1.2 的 8 項規則, 我們便稱之為 vector space. 在這一節中我們將正式定義 vector space 並探討 vector space 相關性質.

給定一非空集合 V, 我們說 V 中有加法運算 (sddition) $+$, 表示對於任意 V 中兩個元素 $\mathbf{u},\mathbf{v}\in V$, 經由這個運算所得的結果 $\mathbf{u}+\mathbf{v}$ 仍然是 V 中的元素 (此為加法封閉性). 至於係數積 (scaler multiplication) 我們要注意的是, 可以乘在向量上的數所在的數系必須像實數一樣有加法與乘法, 且加法, 乘法都是有交換律及結合律, 還有加法乘法之間要有分配律. 更重要的是這個數系裡非 0 的元素都有乘法反元素. 這樣的數系我們稱之為 *field* (體). 例如實數 $\mathbb{R}$ 和有理數 $\mathbb{Q}$ 在我們一般熟悉的加法, 乘法運算下都是 field, 但是整數 $\mathbb{Z}$ 就不是 field, 因為除了 ±1 以外其他的非 0 整數在 $\mathbb{Z}$ 中就無法找到乘法反元素. 由於以後我們談的向量空間, 向量前所乘的係數所在的數系只要是 field, 則我們所要探討的性質都會成立. 所以係數積我們都不會強調是哪一個 field, 而用 $\mathbb{F}$ 來表示. 不過由於我們給的例子大多是係數積為 $\mathbb{R}$ 的情況, 所以若對 field 的概念覺得陌生, 不妨就用 $\mathbb{F}=\mathbb{R}$ 的情況來思考即可. 現若 $\mathbb{F}$ 是一個 field, 我們說 $\mathbb{F}$ 對 V 有係數積表示對任意 $c\in\mathbb{F}$ 以及 $\mathbf{v}\in V$, 皆有 c 對 $\mathbf{v}$ 所得係數積 $c\mathbf{v}$ 仍然在 V 中 (此為係數積封閉性). 當一個集合 V 上有加法運算, 且 field $\mathbb{F}$ 對其有係數積, 則我們可以探討其是否為 vector space, 也就是說探討它是否符合以下之定義.

Definition 2.2.1. 假設非空集合 V 中有加法運算 $+$, 以及 field $\mathbb{F}$ 對 V 的係數積. 若這兩種運算符合以下 8 項性質, 則稱 V 為一個 *vector space over* $\mathbb{F}$.

(1) 對任意 $\mathbf{u},\mathbf{v}\in V$, 皆有 $\mathbf{u}+\mathbf{v}=\mathbf{v}+\mathbf{u}$.

(2) 對任意 $\mathbf{u},\mathbf{v},\mathbf{w}\in V$, 皆有 $(\mathbf{u}+\mathbf{v})+\mathbf{w}=\mathbf{u}+(\mathbf{v}+\mathbf{w})$.

(3) 存在一向量 $\mathbf{0}\in V$ 滿足對任意 $\mathbf{u}\in V$ 皆有 $\mathbf{0}+\mathbf{u}=\mathbf{u}$.

(4) 對任意 $\mathbf{u}\in V$ 皆可找到 $\mathbf{u}'\in V$ 滿足 $\mathbf{u}+\mathbf{u}'=\mathbf{0}$.

(5) 對任意 $\mathbf{u}\in V$, 皆有 $1\mathbf{u}=\mathbf{u}$.

(6) 對任意 $r,s\in\mathbb{F}$ 以及 $\mathbf{u}\in V$, 皆有 $r(s\mathbf{u})=(rs)\mathbf{u}$.

(7) 對任意 $r\in\mathbb{F}$ 以及 $\mathbf{u},\mathbf{v}\in V$ 皆有 $r(\mathbf{u}+\mathbf{v})=r\mathbf{u}+r\mathbf{v}$.

(8) 對任意 $r,s\in\mathbb{F}$ 以及 $\mathbf{u}\in V$, 皆有 $(r+s)\mathbf{u}=r\mathbf{u}+s\mathbf{u}$.

在此要說明一下, 一般來說我們不能說一個集合是 vector space, 一定要附帶說明它的加法及係數積為何. 不過當我們談到一般抽象的 vector space 時, 我們說 V 是一個 vector space over $\mathbb{F}$ 時就隱含其中有加法運算且直接用 $+$ 表示, 同時也隱含 $\mathbb{F}$ 是一個 field 且 V 中有 $\mathbb{F}$ 的係數積, 而不再去強調其中有加法及係數積. 同樣的對於常見的 vector space, 例如 $\mathbb{R}^n$, 由於我們已經有常用的加法及係數積, 所以不會再次強調其加法及係數積為何. 不過當我們要介紹一個新的具體的 vector space 時, 就一定要說明如何定出其加法及係數積. 尤其要注意, 我們必須明確說是 over 哪一個 field 的 vector space (以後我們會看到例子, 同樣的集合看成 over 不同的 field 的 vector space 影響會很大).

或許很多同學會疑惑, 為何在上一節中這 8 個性質是定理要證明, 而這一節中卻是定義不必證明呢? 要回答這個問題就要回歸到整個過程的演變. 上一節中我們定義了坐標平面向量加法及係數積, 然後驗證它們符合 Proposition 2.1.2 這 8 個性質. 而後由這些性質得到許多運算上很方便且豐富的性質. 事實上這些豐富的性質成立的原因, 主因並不是這些平面向量的加法和係數積是如何定義的, 而是由於它們符合 Proposition 2.1.2 這 8 個性質. 注意到這一點後, 我們專注於符合這 8 個性質的系統稱之為向量空間. 希望以後探討向量空間的問題, 可以不必用到它們真正的運算僅利用這 8 個性質就能得到向量空間所有的性質. 所以當我們想推導向量空間的性質時, 我們就可以直接套用定義中這 8 性質去推導, 這樣推導出來的結果便適用於所有的向量空間. 反過來說, 當我們遇到一個系統有加法, 有係數積, 只要我們能利用該系統的運算 "證明" 它符合這基本的 8 個性質, 那它就是向量空間. 因而所有向量空間的性質它都會符合了, 而不必再用該系統的運算一一去推導.

2018

例如, 在 Question 2.1 中我們利用 $\mathbb{R}^2$ 加法的定義說明 vector space 性質 (3) 中 $\mathbf{0}$ 是唯一的而 (4) 中若給定 $\mathbf{u}$, 則反向量 $\mathbf{u}'$ 也是唯一的. 這兩個唯一性, 事實上不需要知道向量的加法如何定義, 直接用 vector space 的性質就可以證明, 也就是說這個結果對一般的 vector space 都成立. 首先我們看以下之定理.

Proposition 2.2.2. 假設 V 為 *vector space,* 且 $\mathbf{u},\mathbf{v},\mathbf{w}\in V$. 若 $\mathbf{u}+\mathbf{w}=\mathbf{v}+\mathbf{w}$, 則 $\mathbf{u}=\mathbf{v}$.

Proof. 利用 vector space 的性質 (4), 我們知道存在 $\mathbf{w}'\in V$ 滿足 $\mathbf{w}+\mathbf{w}'=\mathbf{0}$. 然而由 $\mathbf{u}+\mathbf{w}=\mathbf{v}+\mathbf{w}$ 知 $(\mathbf{u}+\mathbf{w})+\mathbf{w}'=(\mathbf{v}+\mathbf{w})+\mathbf{w}'$. 左式由性質 (2),(3) 可得 $(\mathbf{u}+\mathbf{w})+\mathbf{w}'=\mathbf{u}+(\mathbf{w}+\mathbf{w}')=\mathbf{u}+\mathbf{0}=\mathbf{u}$. 同理右式可得 $(\mathbf{v}+\mathbf{w})+\mathbf{w}'=\mathbf{v}+(\mathbf{w}+\mathbf{w}')=\mathbf{v}+\mathbf{0}=\mathbf{v}$, 得證 $\mathbf{u}=\mathbf{v}$. □

Proposition 2.2.2 告訴我們以後在 vector space 處理向量加法問題時, 可以自然地像處理實數一樣使用消去法. 利用這個結果, 我們可以用來處理上述零向量以及反向量的唯一性.

Corollary 2.2.3. 假設 V 為 *vector space,* 則在 V 中存在唯一的向量 $\mathbf{0}$ 滿足對任意 $\mathbf{u}\in V$ 皆有 $\mathbf{0}+\mathbf{u}=\mathbf{u}$. 另外, 給定 $\mathbf{u}\in V$, 存在唯一的 $\mathbf{u}'\in V$ 滿足 $\mathbf{u}+\mathbf{u}'=\mathbf{0}$.

Proof. 首先證明 $\mathbf{0}$ 是唯一的. 假設 $\mathbf{0}'$ 也滿足 (3) 的性質, 即對任意 $\mathbf{u}\in V$ 皆有 $\mathbf{0}'+\mathbf{u}=\mathbf{u}$. 因此任取 $\mathbf{u}\in V$, 我們有 $\mathbf{0}+\mathbf{u}=\mathbf{u}=\mathbf{0}'+\mathbf{u}$, 故由 Proposition 2.2.2 知 $\mathbf{0}=\mathbf{0}'$, 得證唯一性.

另一方面, 給定 $\mathbf{u}\in V$, 若 $\mathbf{u}',\mathbf{u}''$ 皆滿足 (4) 的性質, 即 $\mathbf{u}+\mathbf{u}'=\mathbf{u}+\mathbf{u}''=\mathbf{0}$. 故由 Proposition 2.2.2 知 $\mathbf{u}''=\mathbf{u}'$, 得證唯一性. □

既然 $\mathbf{0}$ 是唯一的, 以後就用 $\mathbf{0}$ 這個專屬的符號來表示 V 中唯一符合 $\mathbf{0}+\mathbf{u}=\mathbf{u}$, $\forall\mathbf{u}\in V$ 的這個元素, 且稱之為 V 的 *additive identity* 或依慣例稱之為 *zero vector.* 又給定 $\mathbf{u}\in V$, 存在唯一的 $\mathbf{u}'$ 使得 $\mathbf{u}+\mathbf{u}'=\mathbf{0}$, 依慣例我們以後就用 $-\mathbf{u}$ 來表示這一個唯一的 $\mathbf{u}'$, 且稱之為 $\mathbf{u}$ 的 *additive inverse.* 而 $\mathbf{u}$ 和 $\mathbf{v}$ 的 additive inverse $-\mathbf{v}$ 相加, 即 $\mathbf{u}+(-\mathbf{v})$, 我們就用 $\mathbf{u}-\mathbf{v}$ 來表示.

Question 2.2. 假設 V 為 *vector space.* 試證明對任意 $\mathbf{v}\in V$ 皆有 $-(-\mathbf{v})=\mathbf{v}$.

接著我們要再強調的是, 雖然在性質 (3) 中提到 $\mathbf{0}$ 是必須滿足對所有 $\mathbf{u}\in V$ 皆有 $\mathbf{0}+\mathbf{u}=\mathbf{u}$ 才可以 (事實上 Proposition 2.2.2 的證明需用到對所有 $\mathbf{u}$ 皆對才可以), 也就是說依定義要驗證對所有 $\mathbf{u}\in V$ 皆有 $\mathbf{w}+\mathbf{u}=\mathbf{u}$, 才能確定 $\mathbf{w}$ 是零向量. 不過當我們確定 V 是 vector space 之後, 就可利用 Corollary 2.2.3, 知道只要有一個 $\mathbf{u}\in V$, 會使得 $\mathbf{w}+\mathbf{u}=\mathbf{u}$, 就可以認定 $\mathbf{w}=\mathbf{0}$ 了. 利用這一個唯一性, 我們可以推得許多有關於 $\mathbf{0}$ 的性質. 為了方便起見, 我們列出以下的結果.

Proposition 2.2.4. 假設 V 為 *vector space over* $\mathbb{F}$, 我們有以下之結果.

(1) 若對於 $\mathbf{w}\in V$ 存在 $\mathbf{u}\in V$ 使得 $\mathbf{w}+\mathbf{u}=\mathbf{u}$, 則 $\mathbf{w}=\mathbf{0}$.

(2) 對任意 $\mathbf{v}\in V$ 皆有 $0\mathbf{v}=\mathbf{0}$.

(3) 對任意 $r\in\mathbb{F}$ 皆有 $r\mathbf{0}=\mathbf{0}$.

(4) 對任意 $r \in \mathbb{F}, \mathbf{v} \in V$ 皆有 $(-1)(r\mathbf{v}) = -(r\mathbf{v}) = r(-\mathbf{v})$.

Proof. (1) 由於 $\mathbf{w}$ 和 $\mathbf{0}$ 皆滿足 $\mathbf{w}+\mathbf{u}=\mathbf{u}=\mathbf{0}+\mathbf{u}$, 故利用 Proposition 2.2.2 推得 $\mathbf{w}=\mathbf{0}$.

(2) 要證明 $0\mathbf{v}=\mathbf{0}$, 利用 (1), 我們只要檢查是否存在 $\mathbf{u}\in V$ 使得 $0\mathbf{v}+\mathbf{u}=\mathbf{u}$. 事實上, 若考慮 $\mathbf{u}=\mathbf{v}$ 的情形, 此時因 $\mathbf{v}=1\mathbf{v}$, 故 $0\mathbf{v}+\mathbf{v}=0\mathbf{v}+1\mathbf{v}=(0+1)\mathbf{v}=1\mathbf{v}=\mathbf{v}$. 故得證 $0\mathbf{v}=\mathbf{0}$.

(3) 同理, 利用 (1), 我們考慮 $\mathbf{u}=r\mathbf{0}$ 的情況, 此時 $r\mathbf{0}+\mathbf{u}=r\mathbf{0}+r\mathbf{0}=r(\mathbf{0}+\mathbf{0})=r\mathbf{0}=\mathbf{u}$, 得證 $r\mathbf{0}=\mathbf{0}$.

(4) 這裡要證明的是 $(-1)(r\mathbf{v})$ 和 $r(-\mathbf{v})$ 都是 $r\mathbf{v}$ 的 additive inverse (反向量). 由 Corollary 2.2.3 我們知到只要驗證它們是否加上 $r\mathbf{v}$ 都會是 $\mathbf{0}$ 即可. 然而由 (2) 我們有

$$(-1)(r\mathbf{v})+r\mathbf{v}=(-1)(r\mathbf{v})+1(r\mathbf{v})=(-1+1)(r\mathbf{v})=0(r\mathbf{v})=\mathbf{0}.$$

由 (3) 我們有

$$r(-\mathbf{v})+r\mathbf{v}=r(-\mathbf{v})+r(\mathbf{v})=r(-\mathbf{v}+\mathbf{v})=r(\mathbf{0})=\mathbf{0}.$$

因此得證 $(-1)(r\mathbf{v})=-(r\mathbf{v})=r(-\mathbf{v})$. □

由 Proposition 2.2.4 我們知當 $r=0$ 或 $\mathbf{v}=\mathbf{0}$ 時會有 $r\mathbf{v}=\mathbf{0}$, 但若 $r\neq 0$ 且 $\mathbf{v}\neq\mathbf{0}$, 是否有可能 $r\mathbf{v}=\mathbf{0}$ 呢? 答案是不可能. 這是因為若 $r\mathbf{v}=\mathbf{0}$, 首先由 $r\neq 0$ 且 $\mathbb{F}$ 是一個 field, 我們知道存在 $r'\in\mathbb{F}$ 滿足 $r'r=1$. 因此可以考慮 r' 乘上 $r\mathbf{v}$ 由 Proposition 2.2.4 (3) 得到 $r'(r\mathbf{v})=r'\mathbf{0}=\mathbf{0}$. 然而由 vector space 運算性質 (5), (6) 我們有 $r'(r\mathbf{v})=(rr')\mathbf{v}=1\mathbf{v}=\mathbf{v}$. 也就是說 $\mathbf{v}=\mathbf{0}$, 此與假設 $\mathbf{b}\neq\mathbf{0}$ 相矛盾, 故知 $r\mathbf{v}$ 絕對不會是 $\mathbf{0}$.

Question 2.3. 假設 V 為 *vector space over* $\mathbb{F}$.

(1) 已知 $\mathbf{v}\in V$ 且 $\mathbf{v}\neq\mathbf{0}$. 試證明若 $r,s\in\mathbb{F}$ 且 $r\mathbf{v}=s\mathbf{v}$, 則 $r=s$.

(2) 已知 $r\in\mathbb{F}$ 且 $r\neq 0$. 試證明若 $\mathbf{u},\mathbf{v}\in V$ 且 $r\mathbf{u}=r\mathbf{v}$, 則 $\mathbf{u}=\mathbf{v}$.

利用以上結果證明若 V 是 *vector space over* $\mathbb{R}$ 且 V 中有非零元素, 則 V 有無窮多個元素.

總而言之, vector space 中所要求加法及係數積的 8 項性質, 就是要確保一個 vector space 中的元素運算都可像實數一般處理. 例如, 我們可以如實數一樣引用“減法”的符號, 也就是說將 $\mathbf{w}+(-\mathbf{v})$ 寫成 $\mathbf{w}-\mathbf{v}$. 如此一來以後我們在一些等式的推演時就直接沿用大家習慣的「移項」的說法. 例如 $2\mathbf{u}+\mathbf{v}=\mathbf{w}$, 我們就直接移項且乘以 $1/2$ 得 $\mathbf{u}=\frac{1}{2}(\mathbf{w}-\mathbf{v})$.

接下來我們看一些有關 vector space 的例子.

Example 2.2.5. (A) 考慮 S 為所有次數等於 2 的有理係數多項式所成的集合. 對 S 中兩多項式 $f(x)=ax^2+bx+c$, $g(x)=a'x^2+b'x+c'$ 我們定義

$$f(x)+g(x)=(a+a')x^2+(b+b')x+(c+c').$$

S 在這加法定義下並無封閉性. 例如 $f(x)=x^2+2x+1\in S$ 且 $g(x)=-x^2\in S$, 但 $f(x)+g(x)=2x+1\notin S$. 所以在此加法下 S 不是 vector space. 現考慮 $P_2(\mathbb{Q})$ 為次數小於等於 2 的有理係數多項式所成的集合. 利用剛才的加法定義, 我們可得這個加法對 $P_2(\mathbb{Q})$ 有封閉性. 另外若對任意實數 $r\in\mathbb{R}$, 我們定義 r 對 $f(x)=ax^2+bx+c\in P_2(\mathbb{Q})$ 的係數積為

$r\cdot f(x)=(ra)x^2+(rb)x+(rc)$. 在此定義之下實數對 $P_2(\mathbb{Q})$ 的係數積並無封閉性, 例如 $\sqrt{2}$ 乘上 $P_2(\mathbb{Q})$ 的元素 x^2+x+1 會是 $\sqrt{2}x^2+\sqrt{2}x+\sqrt{2}$ 就不再是 $P_2(\mathbb{Q})$ 的元素, 所以在此定義之下 $P_2(\mathbb{Q})$ 也不是 over $\mathbb{R}$ 的 vector space. 不過若同樣的定義考慮有理數 $\mathbb{Q}$ 對 $P_2(\mathbb{Q})$ 的係數積, 則會符合封閉性. 所以我們可以考慮 $P_2(\mathbb{Q})$ 是否是 vector space over $\mathbb{Q}$. 事實上我們很容易驗證此時的加法與係數積會符合 vector space 的 8 個性質, 所以在此定義之下 $P_2(\mathbb{Q})$ 確實是 vector space over $\mathbb{Q}$. 同樣的若 n 是正整數, 若令 $P_n(\mathbb{F})$ 為次數小於 n 且係數在 $\mathbb{F}$ 的多項式所成的集合, 利用上述加法及係數積的定義, 我們可得 $P_n(\mathbb{F})$ 是一個 over $\mathbb{F}$ 的 vector space.

(B) 對任意的 field $\mathbb{F}$, 考慮 $P(\mathbb{F})$ 為所有以 $\mathbb{F}$ 的元素為係數的多項式所成的集合. 利用如 (A) 中定義多項式的加法與係數積, 我們可以證明 $P(\mathbb{F})$ 為 vector space. 首先若 $f(x)=a_nx^n+\cdots+a_1x+a_0, g(x)=b_mx^m+\cdots+b_1x+b_0\in P(\mathbb{F})$, 其中 $m\leq n$, 則我們可以將 $g(x)$ 寫成 $g(x)=b_nx^n+\cdots+b_mx^m+\cdots+b_1x+b_0$, 其中 $b_n=b_{n-1}=\cdots=b_{m+1}=0$. 為了方便起見雖然多項式的次數可能不同, 以後我們都用這種方式將它們補成相同次數再相加. 所以我們可以將 $f(x)+g(x)$ 的定義寫成 $f(x)+g(x)=\sum_{i=0}^{n}(a_i+b_i)x^i$. 而對於 $r\in\mathbb{F}$, 係數積 $rf(x)$ 的定義為 $rf(x)=\sum_{i=0}^{n}(ra_i)x^i$. 利用這個定義以及 $\mathbb{F}$ 是 field 的假設, 我們知道在此定義之下加法和係數積確為 $P(\mathbb{F})$ 中的運算 (有封閉性). 接著我們要一一檢查是否符合 vectors space 的 8 項運算規則.

(1) 對任意 $f(x)=\sum_{i=0}^{n}a_ix^i$, $g(x)=\sum_{i=0}^{n}b_ix^i\in P(\mathbb{F})$ 我們有
$$f(x)+g(x)=\sum_{i=0}^{n}(a_i+b_i)x^i=\sum_{i=0}^{n}(b_i+a_i)x^i=g(x)+f(x).$$

(2) 對任意 $f(x)=\sum_{i=0}^{n}a_ix^i$, $g(x)=\sum_{i=0}^{n}b_ix^i, h(x)=\sum_{i=0}^{n}c_ix^i\in P(\mathbb{F})$ 我們有
$$(f(x)+g(x))+h(x)=\sum_{i=0}^{n}(a_i+b_i)x^i+\sum_{i=0}^{n}c_ix^i=\sum_{i=0}^{n}((a_i+b_i)+c_i)x^i,$$
$$f(x)+(g(x)+h(x))=\sum_{i=0}^{n}a_ix^i+\sum_{i=0}^{n}(b_i+c_i)x^i=\sum_{i=0}^{n}(a_i+(b_i+c_i))x^i.$$
由於 $(a_i+b_i)+c_i=a_i+(b_i+c_i)$, 故得證 $(f(x)+g(x))+h(x)=f(x)+(g(x)+h(x))$.

(3) 考慮零多項式 $g(x)=0=\sum_{i=0}^{n}b_ix^i\in P(\mathbb{F})$, 其中 $b_i=0$, $\forall i=0,1,\ldots,n$. 此時對任意 $f(x)=\sum_{i=0}^{n}a_ix^i\in P(\mathbb{F})$, 可得
$$f(x)+g(x)=\sum_{i=0}^{n}(a_i+b_i)x^i=\sum_{i=0}^{n}a_ix^i=f(x).$$

(4) 給定 $f(x)=\sum_{i=0}^{n}a_ix^i\in P(\mathbb{F})$, 我們考慮 $h(x)=\sum_{i=0}^{n}(-a_i)x^i\in P(\mathbb{F})$, 則
$$f(x)+h(x)=\sum_{i=0}^{n}(a_i-a_i)x^i=\sum_{i=0}^{n}0x^i=0.$$

(5) 對任意 $f(x)=\sum_{i=0}^{n}a_ix^i\in P(\mathbb{F})$, 皆有
$$1f(x)=\sum_{i=0}^{n}(1a_i)x^i=\sum_{i=0}^{n}a_ix^i=f(x).$$

(6) 對任意 $r,s\in\mathbb{R}$ 以及 $f(x)=\sum_{i=0}^{n}a_ix^i\in P(\mathbb{F})$, 我們有

$$r(sf(x))=r(\sum_{i=0}^{n}(sa_i)x^i)=\sum_{i=0}^{n}(r(sa_i))x^i=\sum_{i=0}^{n}((rs)a_i)x^i=(rs)\sum_{i=0}^{n}a_ix^i=(rs)f(x).$$

(7) 對任意 $r,s\in\mathbb{R}$ 以及 $f(x)=\sum_{i=0}^{n}a_ix^i\in P(\mathbb{F})$, 皆有

$$(r+s)f(x)=\sum_{i=0}^{n}((r+s)a_i)x^i=\sum_{i=0}^{n}(ra_i+sa_i)x^i=\sum_{i=0}^{n}(ra_i)x^i+\sum_{i=0}^{n}(sa_i)x^i=rf(x)+sf(x).$$

(8) 對任意 $r\in\mathbb{R}$ 以及 $f(x)=\sum_{i=0}^{n}a_ix^i$, $g(x)=\sum_{i=0}^{n}b_ix^i\in P(\mathbb{F})$ 皆有

$$r(f(x)+g(x))=r(\sum_{i=0}^{n}(a_i+b_i)x^i)=\sum_{i=0}^{n}(r(a_i+b_i))x^i=\sum_{i=0}^{n}(ra_i+rb_i)x^i=rf(x)+rg(x).$$

因為 $P(\mathbb{F})$ 的加法與 $\mathbb{F}$ 的係數積符合 vector space 的 8 項運算規則, 所以在這個加法與係數積的運算之下 $P(\mathbb{F})$ 是一個 vector space over $\mathbb{F}$. 例如所有有理係數多項式所成的集合 $P(\mathbb{Q})$ 就是一個 over $\mathbb{Q}$ 的 vector space, 而實係數多項式所成的集合 $P(\mathbb{R})$ 就是一個 vector space over $\mathbb{R}$. 有趣的是 $P(\mathbb{R})$ 也是一個 vector space over $\mathbb{Q}$, 大家想想看為什麼.

(C) 給定任意 $n\in\mathbb{N}$, 以及一個 field $\mathbb{F}$. 我們令 $\mathbb{F}^n=\{(a_1,\ldots,a_n):a_i\in\mathbb{F}\}$. 我們沿用 $\mathbb{R}^2$ 中向量的加法及係數積來定義 $\mathbb{F}^n$ 中向量的加法以及係數積. 令 $\mathbf{u}=(a_1,\ldots,a_n),\mathbf{v}=(b_1,\ldots,b_n)\in\mathbb{F}^n$ 以及 $r\in\mathbb{F}$. 我們定義

$$\mathbf{u}+\mathbf{v}=(a_1+b_1,\ldots,a_n+b_n)\quad\text{and}\quad r\mathbf{u}=(ra_1,\ldots,ra_n).$$

依此定義我們很容易驗證此加法和係數積運算在 $\mathbb{F}^n$ 是封閉的, 而且符合 vector space 的 8 項運算規則, 所以在這個加法與係數積的運算之下 $\mathbb{F}^n$ 是一個 vector space over $\mathbb{F}$. 同樣的考慮所有 entry 皆為 $\mathbb{F}$ 中元素的 $m\times n$ 矩陣所成的集合 $M_{m,n}(\mathbb{F})$ 利用一般矩陣的加法及係數積, 我們也可驗證 $M_{m,n}(\mathbb{F})$ 是一個 vector space over $\mathbb{F}$.

(D) 給定一非空集合 S 以及 field $\mathbb{F}$, 我們令 $F(S,\mathbb{F})$ 表示所有定義域是 S 且對應域是 $\mathbb{F}$ 的函數所成的集合. 現若 $f,g\in F(S,\mathbb{F})$, 表示對任意 $s\in S$, $f(s),g(s)$ 都會是 $\mathbb{F}$ 中的元素. 所以我們定義 $f+g$ 為定義域是 S 且對應域是 $\mathbb{F}$ 的函數, 其定義為 $f+g$ 在任意 $s\in S$ 的取值為 $f(s)+g(s)$. 亦即 $(f+g)(s)=f(s)+g(s),\forall s\in S$. 對任意 $c\in\mathbb{F}$ 我們也定義 cf 為定義域是 S 且對應域是 $\mathbb{F}$ 的函數, 其定義為 cf 在任意 $s\in S$ 的取值為 $c\cdot f(s)$. 亦即 $(cf)(s)=c\cdot f(s),\forall s\in S$. 由於在此定義之下 $f+g$ 和 cf 仍然在 $F(S,\mathbb{F})$ 中, 所以此加法和係數積運算在 $F(S,\mathbb{F})$ 是封閉的. 我們可以驗證這兩種運算也符合 vector space 的 8 項運算規則, 所以在這個加法與係數積的運算之下 $F(S,\mathbb{F})$ 是一個 vector space over $\mathbb{F}$.

Question 2.4. *依照 Example 2.2.5 (D) 的定義, 試說明* $F(\mathbb{R},\mathbb{R})$*,* $F(\mathbb{R},\mathbb{Q})$*,* $F(\mathbb{Q},\mathbb{Q})$ *以及* $F(\mathbb{Q},\mathbb{R})$ *中那些是 vector space over* $\mathbb{R}$*? 哪些是 vector space over* $\mathbb{Q}$*?*

2.3. Subspaces

在這一節, 我們介紹 subspace 的概念, 簡單地說, 對於一個 vector space 的非空子集合, 如果在此 vector space 的加法及係數積運算之下這個子集合亦為 vector space, 則稱為此 vector space 的 *subspace*.

Definition 2.3.1. 假設 V 為 vector space over $\mathbb{F}$ 且 W 為 V 的 nonempty subset. 若對 W 的元素利用原先 V 的加法及 $\mathbb{F}$ 係數積運算之下 W 亦為 vector space, 則稱 W 為 V 的 *subspace*.

雖然一個 vector space 的 subspace 仍為 vector space, 但要檢查是否為 subspace 不必像檢查 vector space 一樣要去檢查 8 項的運算規則. 這是因為原本 vector space 的 8 項運算規則中除了 (3)(4) 兩項會和所在的集合有關外, 其他各項僅是元素間的運算規則, 和所在的集合無關. 以下的定理告訴我們要辨認是否為 subspace, 只要檢查封閉性即可.

Proposition 2.3.2. 假設 V 為 *vector space over* $\mathbb{F}$ 且 W 為 V 的非空子集合. 則 W 為 V 的 *subspace* 若且唯若對任意 $\mathbf{u},\mathbf{v}\in W$ 且 $r\in\mathbb{F}$ 皆有 $\mathbf{u}+\mathbf{v}\in W$ 以及 $r\mathbf{u}\in W$.

Proof. 首先假設 W 為 V 的 subspace. 由於 vector space 的首要條件就是加法與係數積的封閉性, 因此依 subspace 的定義 W 在 V 的加法以及係數積之下應有封閉性, 亦即對任意 $\mathbf{u},\mathbf{v}\in W$ 且 $r\in\mathbb{F}$ 皆有 $\mathbf{u}+\mathbf{v}\in W$ 以及 $r\mathbf{u}\in W$.

反之, 若 W 在 V 的加法以及係數積之下應有封閉性, 則依定義若此加法及係數積符合 vector space 的 8 項性質, 則 W 就是 over $\mathbb{F}$ 的 vector space, 因此依定義就是 V 的 subspace. 然而這 8 項性質中除了 (3), (4) 兩項外, 其餘了性質由於在 V 中的元素皆成立, 所以當然限制在 W 上依然成立. 因此我們僅要驗證 (3), (4) 兩項即可.

性質 (3) 要求的是在 W 中存在一元素 $\mathbf{w}\in W$ 滿足對任意 $\mathbf{u}\in W$ 皆有 $\mathbf{w}+\mathbf{u}=\mathbf{u}$. 然而由於這些元素皆在 V 中, 而 Proposition 2.2.4 (1) 告訴我們此 $\mathbf{w}$ 就是 V 的 zero vector $\mathbf{0}$. 所以我們要檢查的是 $\mathbf{0}\in W$. 現因 W 不是空集合, 所以必定存在 $\mathbf{u}\in W$, 此時因 $0\in\mathbb{F}$ 且由封閉性 $0\mathbf{u}\in W$, 因此由 Proposition 2.2.4 (2) 得證 $\mathbf{0}=0\mathbf{u}\in W$.

性質 (4) 要求的是對任意 W 中的元素 $\mathbf{u}\in W$ 皆存在 $\mathbf{u}'\in W$ 滿足 $\mathbf{u}+\mathbf{u}'=\mathbf{0}$. 由於 $W\subseteq V$, $\mathbf{u}$ 亦在 V 中, 故由 additive inverse 的唯一性 (Proposition 2.2.3知 $\mathbf{u}'=-\mathbf{u}$ 再由 Proposition 2.2.4 (4) 知 $-\mathbf{u}=(-1)\mathbf{u}$. 因此由 $-1\in\mathbb{F}$ 以及係數積的封閉性知 $-\mathbf{u}\in W$. □

注意在這個證明裡, 我們僅利用係數積的封閉性證明 (3), (4) 成立, 不過在驗證 subspace 時一定還要驗證加法的封閉性, 否則無加法封閉性根本沒資格成為 vector space.

一個 vector space V 中有兩個 trivial subspace, 即 V 和 $\{\mathbf{0}\}$. 其中 $\{\mathbf{0}\}$ 稱為 *zero subspace* of V, 以後我們用 $\mathbf{O}$ 來表示. 例外要注意 subspace 不能是空集合, 又因為 $\mathbf{0}$ 一定在其中, 所以以後檢查 V 中的子集合是否為 subspace, 我們可以先檢查 $\mathbf{0}$ 是否在其中. 一來可以知道它是不是空集合, 而且若 $\mathbf{0}$ 不在其中就可以斷定它不是 subspace, 真是一舉兩得啊! 以下我們寫下一個檢查是否為 subspace 更簡明的方法.

Corollary 2.3.3. 假設 V 為 *vector space over* $\mathbb{F}$ 且 W 為 V 的子集合. 則 W 為 V 的 *subspace* 若且唯若 $\mathbf{0}\in W$ 且對任意 $\mathbf{u},\mathbf{v}\in W$, $r\in\mathbb{R}$ 皆有 $\mathbf{u}+r\mathbf{v}\in W$.

Proof. $(\Rightarrow)$: 依 subspace 的定義, 加法及係數積皆有封閉性, 故對任意 $\mathbf{u},\mathbf{v}\in W$, $r\in\mathbb{F}$ 由係數積的封閉性得 $r\mathbf{v}\in W$ 再由加法封閉性得 $\mathbf{u}+r\mathbf{v}\in W$. 又依定義 W 為非空集合, 故必

存在一向量 $\mathbf{w} \in W$. 現考慮 $0\mathbf{w}$, 依封閉性 $0\mathbf{w} \in W$. 又因 V 為 vector space, 我們知 $0\mathbf{w} = \mathbf{0}$ (Proposition 2.2.4(2)). 故得證 $\mathbf{0} = 0\mathbf{w} \in W$.

($\Leftarrow$) : 由 $\mathbf{0} \in W$, 我們知 W 為 V 的非空子集合. 故由 Proposition 2.3.2, 我們僅要證明 W 有加法和係數積的封閉性. 我們要用對任意 $\mathbf{u}, \mathbf{v} \in W$, $r \in \mathbb{F}$ 皆有 $\mathbf{u} + r\mathbf{v} \in W$ 這個假設證明封閉性. 因 $1 \in \mathbb{F}$, 故對任意 $\mathbf{u}, \mathbf{v} \in W$ 考慮 $r = 1$ 的情形可得 $\mathbf{u} + \mathbf{v} \in W$, 證得加法封閉性. 又若 $\mathbf{v} \in W$ 以及 $r \in \mathbb{F}$, 因為已知 $\mathbf{0} \in W$, 故考慮 $\mathbf{u} = \mathbf{0}$ 的情形得證 $r\mathbf{v} = \mathbf{0} + r\mathbf{y} \in W$. □

由 Corollary 2.3.2, 我們知道要檢查一個 vector space V 中的子集合 W 是否為 V 的 subspace, 我們僅要檢查

(1) $\mathbf{0} \in W$

(2) $\mathbf{u}, \mathbf{v} \in W$, $r \in \mathbb{F} \Rightarrow \mathbf{u} + r\mathbf{v} \in W$.

是否成立即可. 我們看以下的例子.

Example 2.3.4. (A) 考慮 $M_{m\times n}(\mathbb{F})$, 即所有 entries 在 $\mathbb{F}$ 中的 $m \times n$ matrices 所成的 vector space. 所謂 $M_{m\times n}(\mathbb{F})$ 的 *upper triangular* matrix 表示當 $i > j$ 時該矩陣第 (i, j)-th entry 為 0. 我們要證明 $M_{m\times n}(\mathbb{F})$ 中所有的 upper triangular matrix 所成的集合是 $M_{m\times n}(\mathbb{F})$ 的 subspace. 首先觀察 $m \times n$ 階零矩陣, 由於其任意 entry 皆為 0, 當然 (i, j)-th entry 當 $i > j$ 時亦為 0, 因此零矩陣是 upper triangular. 現考慮 $A, B \in M_{m\times n}(\mathbb{F})$ 皆為 upper triangular, 設 a_{ij}, b_{ij} 分別表示 A, B 的 (i, j)-th entry. 對任意 $r \in \mathbb{F}$, 我們有 $A + rB$ 的 (i, j)-th entry 為 $a_{ij} + rb_{ij}$. 現當 $i > j$ 時 $a_{ij} = b_{ij} = 0$, 故得 $a_{ij} + rb_{ij} = 0$. 證得 $A + rB$ 亦為 upper triangular. 因此得證 $M_{m\times n}(\mathbb{F})$ 中所有的 upper triangular matrix 所成的集合是 $M_{m\times n}(\mathbb{F})$ 的 subspace.

(B) 考慮 $M_{n\times n}(\mathbb{F})$, 即所有 entries 在 $\mathbb{F}$ 中的 $n \times n$ 方陣所成的 vector space. 我們想知道 $M_{n\times n}(\mathbb{F})$ 中所有的 symmetric matrices (對稱矩陣) 所成的集合是否為 $M_{n\times n}(\mathbb{F})$ 的 subspace. 首先回顧, 對一個 $m \times n$ matrix A, 我們定義 A 的 *transpose* 為一個 $n \times m$ matrix, 記為 A^{t}, 滿足對任意 $1 \leq i \leq n$, $1 \leq j \leq m$, A^{t} 的 (i, j)-th entry 為 A 的 (j, i)-th entry. 利用矩陣加法及係數積運算, 我們很容易驗證對任意 $m \times n$ 矩陣 A, B 以及 $r \in \mathbb{F}$ 皆滿足 $(A + rB)^{\mathrm{t}} = A^{\mathrm{t}} + (rB)^{\mathrm{t}} = A + rB$. 現回到對稱矩陣的定義, 對於 $A \in M_{n\times n}(\mathbb{F})$ 我們稱 A 為 symmetric matrix, 表示 $A^{\mathrm{t}} = A$. 很明顯的 $n \times n$ 階零矩陣 $\mathbf{0}$ 為 symmetric matrix. 而若 $A, B \in \mathscr{M}_{n\times n}$ 滿足 $A^{\mathrm{t}} = A, B^{\mathrm{t}} = B$, 則對任意 $r \in \mathbb{R}$, 利用 $A^{\mathrm{t}} = A, B^{\mathrm{t}} = B$, 我們有 $(A + rB)^{\mathrm{t}} = A^{\mathrm{t}} + (rB)^{\mathrm{t}} = A + rB = A + rB$. 亦即 $A + rB$ 亦為 symmetric matrix, 得證 $M_{n\times n}(\mathbb{F})$ 中所有的 symmetric matrices 所成的集合為 $M_{n\times n}(\mathbb{F})$ 的 subspace.

$M_{n\times n}(\mathbb{F})$ 中所有的 invertible matrices (可逆矩陣) 所成的集合是否為 $M_{n\times n}(\mathbb{F})$ 的 subspace 呢? 答案是否定的. 很明顯的零矩陣 $\mathbf{0}$ 就不是 invertible, 所以由 $\mathbf{0}$ 不在其中就可得 $\mathscr{M}_{n\times n}$ 中所有的 invertible matrices 所成的集合不是 $M_{n\times n}(\mathbb{F})$ 的 subspace. 其實即使我們考慮 invertible matrices 所成的集合與 $\{\mathbf{0}\}$ 的聯集, 仍不會是 $M_{n\times n}(\mathbb{F})$ 的 subspace. 因為即使此時 $\mathbf{0}$ 在其中, 但仍有可能兩個 invertible matrices 相加後就不是 invertible. 例如在 2×2 的情形, $\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ 和 $\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$ 皆為 invertible, 但是 $\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$ 不是 invertible.

(C) 考慮 $P(\mathbb{F})$, 即所有以 $\mathbb{F}$ 的元素為係數的多項式所成的 vector space. 給定一自然數 $n \in \mathbb{N}$, 我們說明所有次數小於等於 n 的多項式所成的集合 $P_n(\mathbb{F})$ 是 $P(\mathbb{F})$ 的 subspace. 首先我們可以將 $P_n(\mathbb{F})$ 寫成 $P_n(\mathbb{F}) = \{\sum_{i=0}^{n} a_i x^i \mid a_i \in \mathbb{F}\}$. 很明顯的零多項式屬於 $P_n(\mathbb{R})$ (注意一般數學上定義零多項式的次數為 $-\infty$, 而不是 0. 這個部分以後代數課程會去談論, 這裡就不多談). 又若 $f(x) = \sum_{i=0}^{n} a_i x^i, g(x) = \sum_{i=0}^{n} b_i x^i \in P_n(\mathbb{R})$, 則對任意 $r \in \mathbb{R}$, 我們有 $f(x) + rg(x) = \sum_{i=0}^{n}(a_i + rb_i)x^i \in P_n(\mathbb{R})$. 故知 $P_n(\mathbb{F})$ 為 $P(\mathbb{F})$ 的 subspace. 要注意, 若僅考慮次數等於 n 的多項式所成的集合, 那麼就不會是 $P(\mathbb{F})$ 的 subspace 了. 很明顯的零多項式就不會在裡面. 又即使加入零多項式, 但仍有可能兩個次數為 n 的多項式相加之後其次數變小了, 例如 $(x^2 + x + 1) + (-x^2 + x + 1) = 2x + 2$. 也就是說在這情況之下加法是不封閉的, 所以無法成為一個 vector space.

(D) 對一非空集合 S 以及 field $\mathbb{F}$, 考慮 $F(S,\mathbb{F})$ 為所有從 S 映射到 $\mathbb{F}$ 的函數所成的 vector space. 現假設 T 是 S 的一個非空子集合, 考慮 $N_T = \{f \in F(S,\mathbb{F}) : f(t) = 0, \forall t \in T\}$, 亦即 N_T 為 S 到 $\mathbb{F}$ 的函數, 但將 T 中的元素都映射到 0. 我們要說明 N_T 是 $F(S,\mathbb{F})$ 的 subspace. 首先 $F(S,\mathbb{F})$ 中的 zero vector $\mathbf{0}$ 就是零函數, 也就是把 S 中的元素都映射到 0 的函數. 現由於 $T \subseteq S$, 所以此零函數當然把 T 中的元素都映射到 0. 得證 $\mathbf{0} \in N_T$. 現若 $f, g \in N_T$ 且 $r \in \mathbb{F}$. 依定義 $(f + rg)(s) = f(s) + rg(s)$, $\forall s \in S$, 故由 $f, g \in N_T$ 的假設知對任意 $t \in T$, $(f + rg)(t) = f(t) + rg(t) = 0 + 0 = 0$, 得證 $f + rg \in N_T$, 也因此證明了 N_T 是 $F(S,\mathbb{F})$ 的 subspace.

Question 2.5. *在 $M_{m\times n}(\mathbb{F})$ 這個 vector space 中, 考慮在固定的特定位置 (例如對角線位置) 為 0 的矩陣所成的集合, (例如對角線位置皆為 0 的矩陣所成的集合). 試問這樣的集合是否為 $M_{m\times n}(\mathbb{F})$ 的 subspace? 又若考慮在固定的特定位置皆不是 0 的矩陣所成的集合, 是否為 $M_{m\times n}(\mathbb{F})$ 的 subspace?*

Subspace 是 vector space 中特殊的子集合, 所以我們當然希望能利用已知的 subspace "製造" 出新的 subspace. 這裡我們介紹兩種常見的方法.

Proposition 2.3.5. *假設 V 為 vector space over $\mathbb{F}$ 且 W_1, W_2 為 V 的 subspace, 則 $W_1 \cap W_2$ 亦為 V 的 subspace.*

Proof. 首先 W_1, W_2 為 subspace, 故 $\mathbf{0} \in W_1$ 且 $\mathbf{0} \in W_2$, 故得 $\mathbf{0} \in W_1 \cap W_2$. 現若 $\mathbf{u}, \mathbf{v} \in W_1 \cap W_2$ 且 $r \in \mathbb{F}$, 由於 $\mathbf{u}, \mathbf{v}$ 皆屬於 W_1 且 W_1 是 subspace, 故知 $\mathbf{u} + r\mathbf{v} \in W_1$, 同理可得 $\mathbf{u} + r\mathbf{v} \in W_2$, 故得證 $\mathbf{u} + r\mathbf{v} \in W_1 \cap W_2$. □

Question 2.6. *證明任意多個 V 的 subspace 的交集依然是 V 的 subspace. (注意不要用數學歸納法, 數學歸納法僅能證明有限多個的情況)*

雖然兩個 subspaces 的交集仍為 subspace, 但他們的聯集就未必是 subspace 了. 當然了, 當 W_1, W_2 為 V 的 subspace, 若 $W_1 \subseteq W_2$, 則 $W_1 \cup W_2 = W_2$ 當然就是 V 的 subspace. 同樣的, 若 $W_2 \subseteq W_1$, 則 $W_1 \cup W_2 = W_1$ 當然也是 V 的 subspace. 下一個定理就是告訴我們, 除了這兩個明顯的情況外, 兩個 subspaces 的聯集不會是 subspace.

Proposition 2.3.6. 假設 V 為 *vector space over* $\mathbb{F}$ 且 W_1, W_2 為 V 的 *subspace.* 若 $W_1 \not\subseteq W_2$ 且 $W_2 \not\subseteq W_1$, 則 $W_1 \cup W_2$ 不是 V 的 *subspace.*

Proof. 依假設 $W_1 \not\subseteq W_2$ 表示存在一個元素在 W_1 中但不在 W_2, 我們假設 $\mathbf{w}_1 \in W_1$ 但 $\mathbf{w}_1 \notin W_2$, 同理由 $W_2 \not\subseteq W_1$, 我們假設 $\mathbf{w}_2 \in W_2$ 但 $\mathbf{w}_2 \notin W_1$. 當然了, 依定義我們有 $\mathbf{w}_1, \mathbf{w}_2 \in W_1 \cup W_2$, 我們要利用 $\mathbf{w}_1, \mathbf{w}_2$ 這兩個 $W_1 \cup W_2$ 中的元素說明 $W_1 \cup W_2$ 在加法之下沒有封閉性, 因此得證 $W_1 \cup W_2$ 不是 V 的 subspace. □

從 Proposition 2.3.6 的證明中, 我們發現 $W_1 \cup W_2$ 不是 vector space 最主要的原因是沒有加法封閉性 (它有係數積的封閉性), 我們可以考慮以下的集合故意收集加法所得的元素讓它有加法封閉性.

Definition 2.3.7. 假設 V 為 vector space W_1, W_2 為其 subspace, 定義集合

$$W_1 + W_2 = \{\mathbf{w}_1 + \mathbf{w}_2 \mid \mathbf{w}_1 \in W_1, \mathbf{w}_2 \in W_2\}$$

並稱之為 the *sum* of W_1 and W_2.

對任意 $\mathbf{w}_1 \in W_1$, 由於 $\mathbf{w}_1 = \mathbf{w}_1 + \mathbf{0}$, 且 $\mathbf{0} \in W_2$ (因 W_2 為 subspace), 我們有 $\mathbf{w}_1 \in W_1 + W_2$, 亦即 $W_1 \subseteq W_1 + W_2$. 同理可得 $W_2 \subseteq W_1 + W_2$. 以下定理告訴我們 $W_1 + W_2$ 也會是 V 的 subspace, 事實上它是包含 W_1 和 W_2 最小的 subspace.

Proposition 2.3.8. 假設 V 為 *vector space over* $\mathbb{F}$ 且 W_1, W_2 為 V 的 *subspace,* 則 $W_1 + W_2$ 也是 V 的 *subspace.* 特別若 W 是 V 的 *subspace* 且滿足 $W_1 \subseteq W$ 以及 $W_2 \subseteq W$, 則 $W_1 + W_2 \subseteq W$.

Proof. 首先因 $\mathbf{0} \in W_1$ 且 $\mathbf{0} \in W_2$, 故由 $\mathbf{0} = \mathbf{0} + \mathbf{0}$ 可得 $\mathbf{0} \in W_1 + W_2$. 現若 $\mathbf{u}, \mathbf{v} \in W_1 + W_2$ 且 $r \in \mathbb{F}$, 此時由 $\mathbf{u} \in W_1 + W_2$, 知存在 $\mathbf{u}_1 \in W_1$ 以及 $\mathbf{u}_2 \in W_2$ 使得 $\mathbf{u} = \mathbf{u}_1 + \mathbf{u}_2$, 同理存在 $\mathbf{v}_1 \in W_1$ 以及 $\mathbf{v}_2 \in W_2$ 使得 $\mathbf{v} = \mathbf{v}_1 + \mathbf{v}_2$. 因此 $\mathbf{u} + r\mathbf{v} = (\mathbf{u}_1 + \mathbf{u}_2) + r(\mathbf{v}_1 + \mathbf{v}_2) = (\mathbf{u}_1 + r\mathbf{v}_1) + (\mathbf{u}_2 + r\mathbf{v}_2)$. 然而 W_1 是 subspace, 由 $\mathbf{u}_1, \mathbf{v}_1 \in W_1$ 以及 $r \in \mathbb{F}$ 知 $\mathbf{u}_1 + r\mathbf{v}_1 \in W_1$. 同理知 $\mathbf{u}_2 + r\mathbf{v}_2 \in W_2$, 故得 $W_1 + W_2$ 是 V 的 subspace.

現對任意 $\mathbf{u} \in W_1 + W_2$, 因為存在 $\mathbf{w}_1 \in W_1, \mathbf{w}_2 \in W_2$ 滿足 $\mathbf{u} = \mathbf{w}_1 + \mathbf{w}_2$, 故由 $W_1 \subseteq W$ 以及 $W_2 \subseteq W$ 知 $\mathbf{w}_1, \mathbf{w}_2 \in W$, 因此由 W 是 subspace 知 $\mathbf{u} = \mathbf{w}_1 + \mathbf{w}_2 \in W$, 得證 $W_1 + W_2 \subseteq W$. □

Question 2.7. 假設 $n \geq 3$, $W_1, W_2, \ldots, W_n$ 皆為 *vector space* V 的 *subspaces.* 學習 *Definition 2.3.7* 的定義方法, 你覺得要如何定義 $W_1 + W_2 + \cdots + W_n$ 才能讓它為包含 $W_1, W_2, \ldots, W_n$ 最小的 *subspace* 呢?

2.4. Linear Combination and Span of Vectors

在這一節中, 我們將介紹線性組合的概念.

當 V 是 vector space over $\mathbb{F}$, 要如何得到 V 的 subspace 呢? 我們可以在 V 中先找到一個 $\mathbf{v} \in V$ 然後找包含 $\mathbf{v}$ 的集合使其為包含 $\mathbf{v}$ 最小的 subspace. 首先這個集合必須包含所有 $\mathbb{F}$ 中的元素與 $\mathbf{v}$ 的係數積, 如此方可保證係數積的封閉性. 所以我們考慮集合 $\{r\mathbf{v} \mid r \in \mathbb{F}\}$. 這個集合不只對 $\mathbb{F}$ 的係數積有封閉性, 而且有加法的封閉性, 事實上它就是

包含 $\mathbf{v}$ 最小的 subspace 了. 我們用 $\mathrm{Span}(\mathbf{v})$ 來表示它, 意即 $\mathbf{v}$ 所 *span* (展成) 的向量空間. 我們來驗證 $\mathrm{Span}(\mathbf{v})$ 確為 V 的 subspace. 首先由於 $\mathbf{0}=0\mathbf{v}$, 所以的確有 $\mathbf{0}\in\mathrm{Span}(\mathbf{v})$. 接著, 若 $\mathbf{u},\mathbf{w}\in\mathrm{Span}(V)$, 表示存在 $s,t\in\mathbb{F}$ 滿足 $\mathbf{u}=s\mathbf{v}$ 且 $\mathbf{w}=t\mathbf{v}$, 因此對任意 $r\in\mathbb{F}$, 我們有 $\mathbf{u}+r\mathbf{w}=(s\mathbf{v})+r(t\mathbf{v})=(s+rt)\mathbf{v}$. 由於 $s+rt\in\mathbb{F}$, 我們有 $(s+rt)\mathbf{v}\in\mathrm{Span}(\mathbf{v})$, 亦即 $\mathbf{u}+r\mathbf{w}\in\mathrm{Span}(\mathbf{v})$. 得證 $\mathrm{Span}(\mathbf{v})$ 為 V 的 subspace.

現若 $\mathbf{u},\mathbf{v}\in V$, 既然 $\mathrm{Span}(\mathbf{u}),\mathrm{Span}(\mathbf{v})$ 為 V 的 subspace, 由上一節 subspace 的 sum 的概念, 我們知 $\mathrm{Span}(\mathbf{u})+\mathrm{Span}(\mathbf{v})$ 亦為 V 的 subspace. 依定義

$$\mathrm{Span}(\mathbf{u})+\mathrm{Span}(\mathbf{v})=\{r\mathbf{u}+s\mathbf{v}\mid r,s\in\mathbb{F}\}.$$

由於它是包含 $\mathbf{u},\mathbf{v}$ 最小的 subspace, 我們視之為由 $\mathbf{u},\mathbf{v}$ 所展成的 subspace, 故一般用 $\mathrm{Span}(\mathbf{u},\mathbf{v})$ 來表示. 而 $\mathrm{Span}(\mathbf{u},\mathbf{v})$ 中的元素, $r\mathbf{u}+s\mathbf{v}$ 就稱之為 $\mathbf{u},\mathbf{v}$ 的 *linear combination* (線性組合). 這個概念可以推廣到一般有限多個向量的情況, 我們有以下的定義.

Definition 2.4.1. 假設 V 為 vector space over $\mathbb{F}$, 且 $\mathbf{v}_1,\dots,\mathbf{v}_n\in V$. 對於任意 $c_1,\dots,c_n\in\mathbb{F}$, 我們稱 $c_1\mathbf{v}_1+\cdots+c_n\mathbf{v}_n$ 為 $\mathbf{v}_1,\dots,\mathbf{v}_n$ 的 *linear combination*. 所有 $\mathbf{v}_1,\dots,\mathbf{v}_n$ 的 linear combination 所成的集合, 我們用 $\mathrm{Span}(\mathbf{v}_1,\dots,\mathbf{v}_n)$ 來表示, 亦即

$$\mathrm{Span}(\mathbf{v}_1,\dots,\mathbf{v}_n)=\{\sum_{i=1}^{n}c_i\mathbf{v}_i\mid c_1,\dots,c_n\in\mathbb{F}\}.$$

我們可以直接驗證 $\mathrm{Span}(\mathbf{v}_1,\dots,\mathbf{v}_n)$ 會是 V 的 subspace (或是利用 sum 的概念, 參見 Question 2.7). 事實上它是包含 $\mathbf{v}_1,\dots,\mathbf{v}_n$ 最小的 subspace. 這是因為若 W 是 V 的 subspace 且 $\mathbf{v}_1,\dots,\mathbf{v}_n\in W$, 則由 W 的加法與係數積的封閉性得 $\mathrm{Span}(\mathbf{v}_1,\dots,\mathbf{v}_n)\subseteq W$.

其實我們不只談論有限多個 V 中的向量所展成的 subspace, 我們也可談論 V 中任意的非空子集合所展成的 subspace. 不過這裡要注意的是, 我們每次只能處理有限多個向量的加法, 所以線性組合也僅能是有限多個向量的線性組合. 因此當 V 中的子集合 S 有無窮多個元素時, S 的 span 是由 S 中有限多個向量的線性組合所組成的. 我們有以下的定義.

Definition 2.4.2. 假設 V 為 vector space over $\mathbb{F}$ 且 S 是 V 的非空子集. 則定義

$$\mathrm{Span}(S)=\{\sum_{i=1}^{n}c_i\mathbf{v}_i\mid n\in\mathbb{N},\mathbf{v}_1,\dots,\mathbf{v}_n\in S,c_1,\dots,c_n\in\mathbb{F}\}.$$

要提醒大家注意, 在 Definition 2.4.1 中 $\mathrm{Span}(\mathbf{v}_1,\dots,\mathbf{v}_n)$ 的定義, 由於涉及 $\mathbf{v}_1,\dots,\mathbf{v}_n$ 這 n 個給定的向量, 所以在此我們用集合表示法說明 $\mathrm{Span}(\mathbf{v}_1,\dots,\mathbf{v}_n)$ 的元素時不必提及 n 和 $\mathbf{v}_1,\dots,\mathbf{v}_n$ 是什麼. 然而在 Definition 2.4.2 中當我們用集合表示法說明 $\mathrm{Span}(S)$ 的元素時, 我們是在 S 中任選 n 個元素 $\mathbf{v}_1,\dots,\mathbf{v}_n$, 這 n 和 $\mathbf{v}_1,\dots,\mathbf{v}_n$ 是變動的, 所以必須寫下表示 n 是任意可能的正整數, 而 $\mathbf{v}_1,\dots,\mathbf{v}_n$ 是 S 中任意可能的向量.

事實上 $\mathrm{Span}(S)$ 也會是 V 的 subspace. 首先 S 不是空集合, 所以存在 $\mathbf{v}\in S$, 此時考慮 $0\mathbf{v}=\mathbf{0}$, 由 $\mathrm{Span}(S)$ 的定義 (取 $n=1$, $\mathbf{v}_1=\mathbf{v}$, $c_1=0$), 知 $\mathbf{0}\in\mathrm{Span}(S)$. 現若 $\mathbf{u},\mathbf{v}\in\mathrm{Span}(S)$ 且 $r\in\mathbb{F}$, 則由於 $\mathbf{u}=c_1\mathbf{u}_1+\cdots+c_n\mathbf{u}_n$, 其中 $n\in\mathbb{N}$, $\mathbf{u}_1,\dots,\mathbf{u}_n\in S$ 且 $c_1,\dots,c_n\in\mathbb{F}$ 以及 $\mathbf{v}=c'_1\mathbf{v}_1+\cdots+c'_m\mathbf{v}_m$, 其中 $m\in\mathbb{N}$, $\mathbf{v}_1,\dots,\mathbf{v}_m\in S$ 且 $c'_1,\dots,c'_m\in\mathbb{F}$, 我們有

$$\mathbf{u}+r\mathbf{v}=c_1\mathbf{u}_1+\cdots+c_n\mathbf{u}_n+rc'_1\mathbf{v}_1+\cdots+rc'_m\mathbf{v}_m,$$

仍符合 $\mathrm{Span}(S)$ 中元素之定義, 故知 $\mathbf{u}+r\mathbf{v}\in\mathrm{Span}(S)$ 得證 $\mathrm{Span}(S)$ 為 V 的 subspace.

在我們寫下一些向量的線性組合時, 前面乘的係數若是 0, 通常我們的省略不寫. 例如 $2\mathbf{v}_1+3\mathbf{v}_2+0\mathbf{v}_3$ 是 $\mathbf{v}_1,\mathbf{v}_2,\mathbf{v}_3$ 的一個線性組合, 不過我們通常寫成 $2\mathbf{v}_1+3\mathbf{v}_2$. 基於這個原因再加上我們希望任何集合的 span 皆為 vector space, 因此當 S 是空集合 (用 $\emptyset$ 表示), 我們定義 $\mathrm{Span}(S)=\mathrm{Span}(\emptyset)=\{\mathbf{0}\}$ 這一個 zero subspace.

給定一個 vector space V, 若能找到一個 subset S 使得 $\mathrm{Span}(S)=V$, 這當然很好, 表示我們可以用較小的集合 S 就能描述 V. 特別的, 若 S 是有限個元素的集合那就更好, 表示僅需有限多個元素就能 "掌握" V 中所有的元素.

Definition 2.4.3. 假設 V 為 vector space over $\mathbb{F}$ 且 $S\subseteq V$. 若 $\mathrm{Span}(S)=V$, 則稱 S 為 V 的一組 *spanning set*. 此時我們說 S generates (或 spans V). 特別的, 若能找到 finite set (即僅有有限個元素的集合) S 滿足 $\mathrm{Span}(S)=V$, 此時我們稱 V 為 *finitely generated vector space*.

為何特別稱之為 finitely generated vector space 呢? 這表示此 vector space 中所有的元素都可表示成固定有限多個元素的線性組合. 一般的 vector space, 有可能不是 finitely generated, 所以要區分出來, 我們看以下的例子.

Example 2.4.4. 我們討論前面提的一些 vector space 哪些是 finitely generated vector space.

(A) $M_{m\times n}(\mathbb{F})$ 是 finitely generated. 因為考慮 $E_{ij}\in M_{m\times n}(\mathbb{F})$, 表示 (i,j)-th entry 為 1, 其他 entry 為 0 的 $m\times n$ matrix. 很容易看出所有的 $m\times n$ matrix 皆可寫成 E_{ij} 其中 $1\le i\le m$, $1\le j\le n$ 的 linear combination. 所以 $\{E_{ij}\mid 1\le i\le m, 1\le j\le n\}$ 是 $M_{m\times n}(\mathbb{F})$ 的 spanning set, 也因此 $M_{m\times n}(\mathbb{F})$ 是 finitely generated vector space.

(B) $P(\mathbb{F})$ 不是 finitely generated vector space. 這是因為如果 $\{f_1(x),\dots,f_n(x)\}$ 為 $P(\mathbb{F})$ 的 spanning set, 假設 $f_1(x),\dots,f_n(x)$ 的最高次為 m, 則任何 $f_1(x),\dots,f_n(x)$ 的 linear combination $c_1f_1(x)+\cdots+c_nf_n(x)$ 的次數皆不可能大於 m. 也就是說 $\mathrm{Span}(f_1(x),\dots,f_n(x))$ 不可能包含次數大於 m 的多項式. 此與 $\{f_1(x),\dots,f_n(x)\}$ 為 $P(\mathbb{F})$ 的 spanning set 明顯不合, 故知 $P(\mathbb{R})$ 不可能是 finitely generated. 不過次數小於等於 n 的多項式所成的集合 $P_n(\mathbb{F})$ 就是 finitely generated vector space. 很容易看出 $\{x^n,\dots,x,1\}$ 就是 $P_n(\mathbb{F})$ 的 spanning set.

大家或許直覺會認為 finite generated vector space 的 subspace 一定也是 finitely generated. 這是對的, 不過證明卻不是如直覺那麼簡單 (大家不妨現在試著證明看看). 它的證明等到介紹完 linearly independence 的概念, 我們就可以處理.

談到 $\mathrm{Span}(S)$, 我們自然會需要探討那些元素會是 $\mathrm{Span}(S)$ 的元素. 我們看以下的例子.

Example 2.4.5. 在 $P_3(\mathbb{R})$ 中考慮 $\mathbf{u}=x^3-2x^2-5x-3$ and $\mathbf{v}=3x^3-5x^2-4x-9$. 我們要檢查 $2x^3-2x^2+12x-6$ 和 $3x^3-2x^2+7x+8$ 是否屬於 $\mathrm{Span}(\mathbf{u},\mathbf{v})$. 首先檢查是否存在 $a,b\in\mathbb{R}$ 使得 $2x^3-2x^2+12x-6=a\mathbf{u}+b\mathbf{v}=a(x^3-2x^2-5x-3)+b(3x^3-5x^2-4x-9)$ 比較係數後我

們發現 a,b 需滿足聯立方程組

$$\begin{cases} a+3b &= 2 \\ -2a-5b &= -2 \\ -5a-4b &= 12 \\ -3a-9b &= -6 \end{cases}$$

利用上一章解聯立方程組的方法, 我們解得 $a=-4, b=2$, 因此得到 $2x^3-2x^2+12x-6 \in \mathrm{Span}(\mathbf{u},\mathbf{v})$. 同樣的我們要檢查是否存在 $a,b \in \mathbb{R}$ 使得 $3x^3-2x^2+7x+8 = a\mathbf{u}+b\mathbf{v} = a(x^3-2x^2-5x-3)+b(3x^3-5x^2-4x-9)$ 比較係數後我們發現 a,b 需滿足聯立方程組

$$\begin{cases} a+3b &= 3 \\ -2a-5b &= -2 \\ -5a-4b &= 7 \\ -3a-9b &= 8 \end{cases}$$

結果發現此聯立方程組無解, 因此知 $3x^3-2x^2+7x+8 \notin \mathrm{Span}(\mathbf{u},\mathbf{v})$.

我們可以討論更一般的情形, 即探討何時 $f(x)=c_1x^3+c_2x^2+c_3x+c_4$ 會屬於 $\mathrm{Span}(\mathbf{u},\mathbf{v})$. 依定義, 這表示存在 $a,b \in \mathbb{R}$ 使得 $c_1x^3+c_2x^2+c_3x+c_4 = a(x^3-2x^2-5x-3)+b(3x^3-5x^2-4x-9)$ 比較係數後我們發現 a,b 需滿足聯立方程組

$$\begin{cases} a+3b &= c_1 \\ -2a-5b &= c_2 \\ -5a-4b &= c_3 \\ -3a-9b &= c_4 \end{cases}$$

利用 elementary row operation 我們將 augmented matrix

$$\left[\begin{array}{rr|l} 1 & 3 & c_1 \\ -2 & -5 & c_2 \\ -5 & -4 & c_3 \\ -3 & -9 & c_4 \end{array}\right].$$

轉換得 reduced echelon form

$$\left[\begin{array}{cc|r} 1 & 0 & -5c_1-3c_2 \\ 0 & 1 & 2c_1+c_2 \\ 0 & 0 & -17c_1-11c_2+c_3 \\ 0 & 0 & 3c_1+c_4 \end{array}\right].$$

這告訴我們此聯立方程組有解若且唯若 $-17c_1-11c_2+c_3=0$ 且 $3c_1+c_4=0$ 且有解時其解為 $a=-5c_1-3c_2, b=2c_1+c_2$. 也就是說當 $-17c_1-11c_2+c_3=0$ 且 $3c_1+c_4=0$ 時, 多項式 $f(x)=c_1x^3+c_2x^2+c_3x+c_4$ 會屬於 $\mathrm{Span}(\mathbf{u},\mathbf{v})$ 且此時 $f(x)=(-5c_1-3c_2)\mathbf{u}+(2c_1+c_2)\mathbf{v}$.

在這一節最後, 我們列下一些有關 $\mathrm{Span}(S)$ 的性質.

Lemma 2.4.6. *假設 V 為 vector space over $\mathbb{F}$ 且 $S \subseteq V$, 則 $\mathrm{Span}(S)$ 是 V 中包含 S 最小的 subspace. 換句話說, 若 W 是 V 的 subspace 且 $S \subseteq W$, 則 $\mathrm{Span}(S) \subseteq W$.*

Proof. 依定義 $S \subseteq \mathrm{Span}(S)$ 且我們已知 $\mathrm{Span}(S)$ 是 V 的 subspace. 現假設 W 是 V 的 subspace 且 $S \subseteq W$, 我們要說明 $\mathrm{Span}(S) \subseteq W$. 對任意 $\mathbf{v} \in \mathrm{Span}(S)$, 我們知存在 $\mathbf{v}_1,\ldots,\mathbf{v}_n \in S$ 以及 $c_1,\ldots,c_n \in \mathbb{F}$ 使得 $\mathbf{v}=\sum_{i=1}^n c_i\mathbf{v}_i$. 現因 $S \subseteq W$, 我們有 $\mathbf{v}_1,\ldots,\mathbf{v}_n \in W$, 故由 W 是 subspace, 得證 $\mathbf{v}=\sum_{i=1}^n c_i\mathbf{v}_i \in W$. □

利用 Lemma 2.4.6, 我們馬上可知若 $S_1 \subseteq S_2$, 則 $\mathrm{Span}(S_1) \subseteq \mathrm{Span}(S_2)$. 這當然可直接用定義來證明, 不過利用 Lemma 2.4.6, 我們就可以直接套用, 而省去許多繁瑣的論證. 這是因為 $\mathrm{Span}(S_2)$ 是 V 的 subspace 又 $S_1 \subseteq S_2 \subseteq \mathrm{Span}(S_2)$, 故套用 Lemma 2.4.6 (考慮 $S = S_1$, $W = \mathrm{Span}(S_2)$ 的情形) 得證 $\mathrm{Span}(S_1) \subseteq \mathrm{Span}(S_2)$. 利用這個概念我們也可以很快的處理 Span 對兩集合交集及聯集的影響.

當 S_1, S_2 是 V 的 subsets, 我們可以考慮 $\mathrm{Span}(S_1 \cap S_2)$ 和 $\mathrm{Span}(S_1) \cap \mathrm{Span}(S_2)$ 的關係. 由於 $S_1 \cap S_2 \subseteq S_1$, 我們有 $\mathrm{Span}(S_1 \cap S_2) \subseteq \mathrm{Span}(S_1)$. 同理 $\mathrm{Span}(S_1 \cap S_2) \subseteq \mathrm{Span}(S_2)$. 由於 $\mathrm{Span}(S_1 \cap S_2)$ 同時包含於 $\mathrm{Span}(S_1)$ 和 $\mathrm{Span}(S_2)$ 可推得 $\mathrm{Span}(S_1 \cap S_2) \subseteq \mathrm{Span}(S_1) \cap \mathrm{Span}(S_2)$. 不過反向 $\mathrm{Span}(S_1 \cap S_2) \supseteq \mathrm{Span}(S_1) \cap \mathrm{Span}(S_2)$ 就不一定成立, 主要原因是取集合的交集遠比 Span 後再取交集小的多. 例如在 $\mathbb{R}^2$ 上考慮 $S_1 = \{(1,1)\}$, $S_2 = \{(2,2)\}$. 我們有 $S_1 \cap S_2 = \emptyset$ 所以 $\mathrm{Span}(S_1 \cap S_2) = \mathrm{Span}(\emptyset) = \{\mathbf{0}\}$; 不過 $\mathrm{Span}(S_1) = \{(r,r) \mid r \in \mathbb{R}\} = \mathrm{Span}(S_2)$, 所以 $\mathrm{Span}(S_1 \cap S_2) = \{\mathbf{0}\} \subsetneq \mathrm{Span}(S_1) \cap \mathrm{Span}(S_2)$.

接下來我們看聯集的情況. 因為 $S_1 \subseteq S_1 \cup S_2$ 且 $S_2 \subseteq S_1 \cup S_2$, 我們有 $\mathrm{Span}(S_1) \subseteq \mathrm{Span}(S_1 \cup S_2)$ 以及 $\mathrm{Span}(S_2) \subseteq \mathrm{Span}(S_1 \cup S_2)$, 所以當然 $\mathrm{Span}(S_1) \cup \mathrm{Span}(S_2) \subseteq \mathrm{Span}(S_1 \cup S_2)$. 不過 $\mathrm{Span}(S_1) \cup \mathrm{Span}(S_2)$ 比起 $\mathrm{Span}(S_1 \cup S_2)$ 太小了, 事實上我們知道 $\mathrm{Span}(S_1) \cup \mathrm{Span}(S_2)$ 在大多數的情況甚至不是 subspace (Proposition 2.3.6). 不過 Proposition 2.3.8 告訴我們 $\mathrm{Span}(S_1) + \mathrm{Span}(S_2)$ 是包含 $\mathrm{Span}(S_1)$ 和 $\mathrm{Span}(S_2)$ 最小的 subspace, 因此由 $\mathrm{Span}(S_1 \cup S_2)$ 包含 $\mathrm{Span}(S_1)$ 和 $\mathrm{Span}(S_2)$ 且是 subspace 得 $\mathrm{Span}(S_1) + \mathrm{Span}(S_2) \subseteq \mathrm{Span}(S_1 \cup S_2)$. 另一方面 $\mathrm{Span}(S_1 \cup S_2)$ 是包含 $S_1 \cup S_2$ 最小的 subspace, 然而 $S_1 \subseteq \mathrm{Span}(S_1) + \mathrm{Span}(S_2)$ 且 $S_2 \subseteq \mathrm{Span}(S_1) + \mathrm{Span}(S_2)$, 再加上 $\mathrm{Span}(S_1) + \mathrm{Span}(S_2)$ 是 subspace, 所以我們也有 $\mathrm{Span}(S_1 \cup S_2) \subseteq \mathrm{Span}(S_1) + \mathrm{Span}(S_2)$. 因此證得 $\mathrm{Span}(S_1 \cup S_2) = \mathrm{Span}(S_1) + \mathrm{Span}(S_2)$. 我們總結以上的結果如下.

Proposition 2.4.7. 假設 V 為 *vector space over* $\mathbb{F}$ 且 S_1, S_2 皆為 V 的 *subsets.*

(1) 若 $S_1 \subseteq S_2$, 則 $\mathrm{Span}(S_1) \subseteq \mathrm{Span}(S_2)$.

(2) $\mathrm{Span}(S_1 \cap S_2) \subseteq \mathrm{Span}(S_1) \cap \mathrm{Span}(S_2)$.

(3) $\mathrm{Span}(S_1 \cup S_2) = \mathrm{Span}(S_1) + \mathrm{Span}(S_2)$.

2.5. Linearly Dependence and Independence

Spanning set 的概念是處理 linear combination 的存在性, 也就是了解一個向量可不可以寫成一些特定向量的線性組合; 而 linear independence 的概念便是處理 linear combination 的唯一性, 也就是說寫成線性組合的方法是否唯一. 這一節中便是探討 linearly independent 的概念.

考慮 over $\mathbb{F}$ 的 vector space V. 若 $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3 \in V$, 我們知道 $\mathrm{Span}(\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3)$ 為 V 中包含 $\mathbf{v}_1, \mathbf{v}_2$ 以及 $\mathbf{v}_3$ 最小的 subspace. 這個 subspace 很容易掌握, 因為每個元素都是 $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$ 的線性組合, 我們只要了解 $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$ 就可以瞭解這個 subspace. 當然了, 若能用較少的元素就能掌握整個 subspace 就更好, 所以我們自然會問這裡 $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$ 有沒有多餘, 不必要的呢? 這是有可能的, 例如若 $\mathbf{v}_3 \in \mathrm{Span}(\mathbf{v}_1, \mathbf{v}_2)$, 則 $\mathrm{Span}(\mathbf{v}_1, \mathbf{v}_2) = \mathrm{Span}(\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3)$. 所以此時僅要了

解 $\mathbf{v}_1,\mathbf{v}_2$ 就夠了. 當 $\mathbf{v}_3 \in \mathrm{Span}(\mathbf{v}_1,\mathbf{v}_2)$ 表示存在 $r_1,r_2 \in \mathbb{F}$ 使得 $\mathbf{v}_3 = r_1\mathbf{v}_1 + r_2\mathbf{v}_2$, 此時表示 $\mathbf{v}_3$ 和 $\mathbf{v}_1,\mathbf{v}_2$, 是有關係的, 我們稱 $\mathbf{v}_1,\mathbf{v}_2,\mathbf{v}_3$ 是 *linearly dependent* (線性相依或線性相關). 不過要注意 $\mathbf{v}_3 \in \mathrm{Span}(\mathbf{v}_1,\mathbf{v}_2)$ 並不表示 $\mathbf{v}_1 \in \mathrm{Span}(\mathbf{v}_2,\mathbf{v}_3)$. 我們看以下的例子.

Example 2.5.1. 在 $\mathbb{R}^2$ 中考慮 $\mathbf{v}_1 = (1,0)$, $\mathbf{v}_2 = (1,1)$, $\mathbf{v}_3 = (2,2)$. 我們有 $\mathbf{v}_3 \in \mathrm{Span}(\mathbf{v}_1,\mathbf{v}_2)$, 因為 $\mathbf{v}_3 = (2,2) = 0(1,0)+2(1,1) = 0\mathbf{v}_1 + 2\mathbf{v}_2 \in \mathrm{Span}(\mathbf{v}_1,\mathbf{v}_2)$. 不過 $\mathbf{v}_1 \notin \mathrm{Span}(\mathbf{v}_2,\mathbf{v}_3)$, 因為 $(1,0)$ 無法寫成 $(1,1)$ 和 $(2,2)$ 的線性組合.

因此當我們要檢查 $\mathbf{v}_1,\mathbf{v}_2,\mathbf{v}_3$ 之間是否有線性關係, 不能只檢查是否 $\mathbf{v}_3 \in \mathrm{Span}(\mathbf{v}_1,\mathbf{v}_2)$ 還要檢查是否 $\mathbf{v}_1 \in \mathrm{Span}(\mathbf{v}_2,\mathbf{v}_3)$ 以及 $\mathbf{v}_2 \in \mathrm{Span}(\mathbf{v}_1,\mathbf{v}_3)$. 不過分開檢查這三種情況有點麻煩. 我們再深入看一下這三種情況代表甚麼. 當 $\mathbf{v}_1 \in \mathrm{Span}(\mathbf{v}_2,\mathbf{v}_3)$, 表示存在 $r,s \in \mathbb{F}$ 使得 $\mathbf{v}_1 = r\mathbf{v}_2 + s\mathbf{v}_3$, 也就是說 $1\mathbf{v}_1 + (-r)\mathbf{v}_2 + (-s)\mathbf{v}_3 = \mathbf{0}$. 同理, $\mathbf{v}_2 \in \mathrm{Span}(\mathbf{v}_1,\mathbf{v}_3)$ 和 $\mathbf{v}_3 \in \mathrm{Span}(\mathbf{v}_1,\mathbf{v}_2)$ 分別表示存在 $r',s' \in \mathbb{F}$ 和 $r'',s'' \in \mathbb{F}$ 分別使得 $\mathbf{v}_2 = r'\mathbf{v}_1 + s'\mathbf{v}_3$ 和 $\mathbf{v}_3 = r''\mathbf{v}_1 + s''\mathbf{v}_2$. 也就是說當 (1) $\mathbf{v}_1 \in \mathrm{Span}(\mathbf{v}_2,\mathbf{v}_3)$, (2) $\mathbf{v}_2 \in \mathrm{Span}(\mathbf{v}_1,\mathbf{v}_3)$ 或 (3) $\mathbf{v}_3 \in \mathrm{Span}(\mathbf{v}_1,\mathbf{v}_2)$ 其中有一個發生時, 我們會有相對應的

$$\begin{array}{rcrcrcll}
1\mathbf{v}_1 & + & (-r)\mathbf{v}_2 & + & (-s)\mathbf{v}_3 & = & \mathbf{0} & \qquad (1)\\
(-r')\mathbf{v}_1 & + & 1\mathbf{v}_2 & + & (-s')\mathbf{v}_3 & = & \mathbf{0} & \qquad (2)\\
(-r'')\mathbf{v}_1 & + & (-s'')\mathbf{v}_2 & + & 1\mathbf{v}_3 & = & \mathbf{0} & \qquad (3)
\end{array}$$

這三種可能的情況發生, 其中 $r,s,r',s',r'',s'' \in \mathbb{F}$. 不過不管 (1), (2), (3) 哪一種情況發生, 總有一個 $\mathbf{v}_i$ 其前面的係數是 1, 不為 0. 也就是說我們可找到不全為 0 的 c_1,c_2,c_3 使得 $c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + c_3\mathbf{v}_3 = \mathbf{0}$. 反之若能找到 c_1,c_2,c_3 不全為 0 使得 $c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + c_3\mathbf{v}_3 = \mathbf{0}$, 我們就可將前面係數 c_i 不等於 0 的 $\mathbf{v}_i$ 寫成另兩個向量的線性組合. 例如若 $c_1 \neq 0$, 則可得 $\mathbf{v}_1 = (-c_1'c_2)\mathbf{v}_2 + (-c_1'c_3)\mathbf{v}_3$, 其中 c_1' 為 c_1 的乘法反元素 (因 $c_1 \neq 0$). 由此可知, 存在 c_1,c_2,c_3 不全為 0 使得 $c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + c_3\mathbf{v}_3 = \mathbf{0}$ 和前面所提 (1), (2), (3) 三種情況是等價的, 因此我們用此方法來定義 $\mathbf{v}_1,\mathbf{v}_2,\mathbf{v}_3$ 為 linearly dependent.

讓我們把以上概念推廣到任意多個向量的情況. 我們稱一組向量為 linearly dependent, 指的是這一組向量之間有關係, 也就是說其中有一個向量是其他向量的線性組合. 例如假設 $\mathbf{v}_1,\dots,\mathbf{v}_n$ 為 linearly dependent, 就表示其中有一個 $\mathbf{v}_i$ 可以寫成 $\mathbf{v}_1,\dots,\mathbf{v}_{i-1},\mathbf{v}_{i+1},\dots,\mathbf{v}_n$ 的線性組合. 每次要提有一個 $\mathbf{v}_i$ 是其他向量的線性組合有點麻煩. 不過若我們更進一步觀察, 此時 $\mathbf{v}_i = r_1\mathbf{v}_1 + \cdots + r_{n-1}\mathbf{v}_{n-1} + r_{n+1}\mathbf{v}_{n+1} + \cdots + r_n\mathbf{v}_n$, 其中這些 r_j 皆為實數. 所以我們得

$$r_1\mathbf{v}_1 + \cdots + r_{n-1}\mathbf{v}_{n-1} + (-1)\mathbf{v}_i + r_{n+1}\mathbf{v}_{n+1} + \cdots + r_n\mathbf{v}_n = \mathbf{0}.$$

也就是說我們找到一組不全為 0 的實數 $c_1,\dots,c_n$ 使得 $c_1\mathbf{v}_1 + \cdots + c_n\mathbf{v}_n = \mathbf{0}$. 反之, 若存在一組不全為 0 的實數 $c_1,\dots,c_n$ 使得 $c_1\mathbf{v}_1 + \cdots + c_n\mathbf{v}_n = \mathbf{0}$. 我們假設 $c_i \neq 0$, 此時令 c_i' 為 c_i 的乘法反元素 (即 $c_ic_i' = 1$), 可得

$$\mathbf{v}_i = (-c_1c_i')\mathbf{v}_1 + \cdots + (-c_{i-1}c_i')\mathbf{v}_{i-1} + (-c_{i+1}c_i')\mathbf{v}_{i+1} + \cdots + (-c_nc_i')\mathbf{v}_n,$$

也就是說 $\mathbf{v}_i$ 可以寫成 $\mathbf{v}_1,\dots,\mathbf{v}_{i-1},\mathbf{v}_{i+1},\dots,\mathbf{v}_n$ 的線性組合. 由此可知, 存在一組不全為 0 的實數 $c_1,\dots,c_n$ 使得 $c_1\mathbf{v}_1 + \cdots + c_n\mathbf{v}_n = \mathbf{0}$ 就等同於 $\mathbf{v}_1,\dots,\mathbf{v}_n$ 這一組向量之間有關係. 由於這

個方式來表達線性相關不必敘述其中哪一個向量是其他向量的線性組合, 較為簡潔. 一般就以這個方式來定義線性相關.

Definition 2.5.2. 假設 V 是一個 vector space over $\mathbb{F}$, 且 $\mathbf{v}_1,\ldots,\mathbf{v}_n \in V$, 若存在一組不全為 0 的 $c_1,\ldots,c_n \in \mathbb{F}$ 使得
$$c_1\mathbf{v}_1+\cdots+c_n\mathbf{v}_n=\mathbf{0},$$
則稱 $\mathbf{v}_1,\ldots,\mathbf{v}_n$ 為 *linearly dependent* (線性相依或線性相關). 反之, 若 $\mathbf{v}_1,\ldots,\mathbf{v}_n$ 不是 linearly dependent, 則稱為 *linearly independent* (線性獨立).

這個定義可以推廣到 V 的任意子集合. 若 $S\subseteq V$, 且 S 中存在相異 $\mathbf{v}_1,\ldots,\mathbf{v}_n \in S$ 使得 $\mathbf{v}_1,\ldots,\mathbf{v}_n$ 為 linearly dependent, 則稱集合 S 為 linearly dependent; 反之, 若 S 不是 linearly dependent, 表示任意的一組相異向量 $\mathbf{v}_1,\ldots,\mathbf{v}_n \in S$ 皆為 linearly independent, 則稱 S 為 linearly independent. 在此定義之下空集合 $\emptyset$ 是 linearly independent, 因為我們無法在 $\emptyset$ 中找到任何 $\mathbf{v}_1,\ldots,\mathbf{v}_n$ 是 linearly dependent. 另一方面, 若 $\mathbf{0}\in S$, 則 S 一定是 linearly dependent. 因為 $\mathbf{0}$ 本身一個元素就是 linearly dependent. 原因是 $1\mathbf{0}=\mathbf{0}$, 所以我們找到一個係數不為 0 的 $\mathbf{0}$ 的線性組合. 因此依定義 S 為 linearly dependent.

一般來說, 除了上述 $S=\emptyset$ 和 $\mathbf{0}\in S$ 這兩種情況外, 要說明 S 是否為 linearly independent 並不是馬上能看出來的. 通常當我們要證明一組向量 $\mathbf{v}_1,\ldots,\mathbf{v}_n$ 為 linearly independent, 我們有以下兩個方法: 第一個方法就是先設 $c_1\mathbf{v}_1+\cdots+c_n\mathbf{v}_n=\mathbf{0}$, 再證明此時 $c_1,\ldots,c_n$ 必全為 0. 第二種方法, 就是所謂的反證法, 亦即先假設 $\mathbf{v}_1,\ldots,\mathbf{v}_n$ 為 linearly dependent (也就是說假設存在不全為 0 的 $c_1,\ldots,c_n\in\mathbb{F}$ 使得 $c_1\mathbf{v}_1+\cdots+c_n\mathbf{v}_n=\mathbf{0}$), 再推得矛盾. 第一個方法通常在有具體的向量時使用, 而處理抽象的情形大多使用第二種方法, 如下面的例子.

Example 2.5.3. 假設 V 為 vector space over $\mathbb{F}$ 且 $\mathbf{v}_1,\ldots,\mathbf{v}_n \in V$ 為 linearly independent. 這表示 $\mathbf{v}_1,\ldots,\mathbf{v}_n$ 之間沒有線性關係. 因此可以理解若我們在 $\mathbf{v}_1,\ldots,\mathbf{v}_n$ 中移除 $\mathbf{v}_n$, 則 $\mathbf{v}_1,\ldots,\mathbf{v}_{n-1}$ 這一組向量應仍為 linearly independent. 要證明這一個事實, 若我們用第一個方法, 很難由 $c_1\mathbf{v}_1+\cdots+c_{n-1}\mathbf{v}_{n-1}=\mathbf{0}$ 推得 $c_1,\ldots,c_{n-1}$ 必全為 0. 然而若利用第二個方法, 即假設存在不全為 0 的實數 $c_1,\ldots,c_{n-1}$ 使得 $c_1\mathbf{v}_1+\cdots+c_{n-1}\mathbf{v}_{n-1}=\mathbf{0}$. 此時令 $c_n=0$, 我們得到一組不全為 0 的實數 $c_1,\ldots,c_n$ 使得
$$c_1\mathbf{v}_1+\cdots+c_{n-1}\mathbf{v}_{n-1}+c_n\mathbf{v}_n=c_1\mathbf{v}_1+\cdots+c_{n-1}\mathbf{v}_{n-1}=\mathbf{0}.$$
此與 $\mathbf{v}_1,\ldots,\mathbf{v}_n\in\mathbb{R}^m$ 為 linearly independent 的假設相矛盾, 故得證 $\mathbf{v}_1,\ldots,\mathbf{v}_{n-1}$ 為 linearly independent. 大家應可以看出, 我們其實是證明了當 $\mathbf{v}_1,\ldots,\mathbf{v}_{n-1}\in\mathbb{R}^m$ 為 linearly dependent 時, 加入任意的 $\mathbf{v}_n\in\mathbb{R}^n$ 後, $\mathbf{v}_1,\ldots,\mathbf{v}_{n-1},\mathbf{v}_n$ 也是 linearly dependent.

Question 2.8. 假設 V 為 *vector space over* $\mathbb{F}$ 且 $S\subseteq S'\subseteq V$. 試說明以下的對錯.

(1) 若 S 為 *linearly independent,* 則 S' 為 *linearly independent.*

(2) 若 S 為 *linearly dependent,* 則 S' 為 *linearly dependent.*

(3) 若 S' 為 *linearly independent,* 則 S 為 *linearly independent.*

(4) 若 S' 為 *linearly dependent,* 則 S 為 *linearly dependent.*

在 Example 2.5.3 中, 我們知道當 $\mathbf{v}_1,\ldots,\mathbf{v}_n \in V$ 這一組向量為 linearly independent 時, 在這一組向量中移除一些向量, 仍不會改變其 linearly independent 的性質. 但若加入新的向量情況可能改變. 下一個定理就是告訴我們何時加入新的向量仍會保持 linearly independent.

Lemma 2.5.4. 假設 V 為 *vector space over* $\mathbb{F}$ 且 $\mathbf{v}_1,\ldots,\mathbf{v}_n,\mathbf{v}_{n+1} \in V$. 若已知 $\mathbf{v}_1,\ldots,\mathbf{v}_n$ 為 *linearly independent,* 則 $\mathbf{v}_1,\ldots,\mathbf{v}_n,\mathbf{v}_{n+1}$ 為 *linearly independent* 若且唯若 $\mathbf{v}_{n+1} \notin \mathrm{Span}(\mathbf{v}_1,\ldots,\mathbf{v}_n)$.

Proof. 如果 $\mathbf{v}_{n+1} \in \mathrm{Span}(\mathbf{v}_1,\ldots,\mathbf{v}_n)$, 表示 $\mathbf{v}_1,\ldots,\mathbf{v}_n,\mathbf{v}_{n+1}$ 之間有線性關係, 即為 linearly dependent. 故知若 $\mathbf{v}_1,\ldots,\mathbf{v}_n,\mathbf{v}_{n+1}$ 為 linearly independent, 不可能會有 $\mathbf{v}_{n+1} \in \mathrm{Span}(\mathbf{v}_1,\ldots,\mathbf{v}_n)$ 的情形發生. 得證 $\mathbf{v}_{n+1} \notin \mathrm{Span}(\mathbf{v}_1,\ldots,\mathbf{v}_n)$.

反之, 假設 $\mathbf{v}_{n+1} \notin \mathrm{Span}(\mathbf{v}_1,\ldots,\mathbf{v}_n)$ 我們要證明 $\mathbf{v}_1,\ldots,\mathbf{v}_n,\mathbf{v}_{n+1}$ 為 linearly independent. 利用反證法, 即設 $\mathbf{v}_1,\ldots,\mathbf{v}_n,\mathbf{v}_{n+1}$ 為 linearly dependent, 也就是說存在一組不全為 0 的實數 $c_1,\ldots,c_n,c_{n+1}$ 使得 $c_1\mathbf{v}_1+\cdots+c_n\mathbf{v}_n+c_{n+1}\mathbf{v}_{n+1}=\mathbf{0}$. 我們知此時 c_{n+1} 必為 0, 否則由 $c_{n+1}\neq 0$ 知存在 $c'_{n+1}\in\mathbb{F}$ 使得 $c_{n+1}c'_{n+1}=1$ 此時

$$\mathbf{v}_{n+1}=(-c_1c'_{n+1})\mathbf{v}_1+\cdots+(-c_nc'_{n+1})\mathbf{v}_n,$$

會得到 $\mathbf{v}_{n+1} \in \mathrm{Span}(\mathbf{v}_1,\ldots,\mathbf{v}_n)$ 之矛盾. 因此由 $c_{n+1}=0$, 得 $c_1,\ldots,c_n$ 不全為 0 且使得

$$c_1\mathbf{v}_1+\cdots+c_n\mathbf{v}_n=\mathbf{0},$$

亦即 $\mathbf{v}_1,\ldots,\mathbf{v}_n$ 為 linearly dependent. 這和已知的假設 $\mathbf{v}_1,\ldots,\mathbf{v}_n$ 為 linearly independent 相矛盾, 故得證 $\mathbf{v}_1,\ldots,\mathbf{v}_n,\mathbf{v}_{n+1}$ 為 linearly independent. □

我們常常看到一個 vector space 中, 若一個集合中向量的個數太多時, 就不會是 linearly independent 了. 例如在 $\mathbb{R}^2$ 中任意 3 個向量 $\mathbf{v}_1,\mathbf{v}_2,\mathbf{v}_3$ 就一定會 linearly dependent. 這事實的原因便是我們在找 $c_1,c_2,c_3\in\mathbb{R}$ 使得 $c_1\mathbf{v}_1+c_2\mathbf{v}_2+c_3\mathbf{v}_3=\mathbf{0}$, 就等同於解聯立方程組 $x_1\mathbf{v}_1+x_2\mathbf{v}_2+x_3\mathbf{v}_3=(0,0)$. 這是有三個未知數 x_1,x_2,x_3 但僅有兩個方程式的齊次聯立方程組, 我們知道一定有無窮多解, 也就是說存在 $c_1,c_2,c_3\in\mathbb{R}$ 不全為 0 使得 $c_1\mathbf{v}_1+c_2\mathbf{v}_2+c_3\mathbf{v}_3=\mathbf{0}$. 因此 $\mathbf{v}_1,\mathbf{v}_2,\mathbf{v}_3$ 一定是 linearly dependent. 我們可以利用這個概念證明以下著個重要的定理.

Lemma 2.5.5. 設 V 為 *vector space over* $\mathbb{F}$ 且 $\mathbf{v}_1,\ldots,\mathbf{v}_n\in V$. 若 $\mathbf{w}_1,\ldots,\mathbf{w}_m\in\mathrm{Span}(\mathbf{v}_1,\ldots,\mathbf{v}_n)$ 且 $m>n$, 則 $\mathbf{w}_1,\ldots,\mathbf{w}_m$ 為 *linearly dependent.*

Proof. 由於 $\mathbf{w}_1,\ldots,\mathbf{w}_m\in\mathrm{Span}(\mathbf{v}_1,\ldots,\mathbf{v}_n)$, 因此對任意 $j=1,\ldots,m$, $\mathbf{w}_j$ 都可以寫成 $\mathbf{v}_1,\ldots,\mathbf{v}_n$ 的 linear combination. 也就是說, 存在 $a_{1,j},\ldots,a_{i,j},\ldots,a_{n,j}\in\mathbb{R}$ 使得

$$\mathbf{w}_j=a_{1,j}\mathbf{v}_1+\cdots+a_{i,j}\mathbf{v}_i+\cdots+a_{n,j}\mathbf{v}_n.$$

現在我們要找到 $c_1,\ldots,c_m\in\mathbb{R}$ 不全為 0 使得 $c_1\mathbf{w}_1+\cdots+c_m\mathbf{w}_m=\mathbf{0}$, 便證得 $\mathbf{w}_1,\ldots,\mathbf{w}_m$ 為 linearly dependent. 現將 $c_1\mathbf{w}_1+\cdots+c_m\mathbf{w}_m$ 中每一個 $\mathbf{w}_j$ 換成 $\mathbf{v}_1,\ldots,\mathbf{v}_n$ 的 linear combination 後會等於

$$(c_1a_{1,1}+\cdots+c_ma_{1,m})\mathbf{v}_1+\cdots+(c_1a_{i,1}+\cdots+c_ma_{i,m})\mathbf{v}_i+\cdots+(c_1a_{m,1}+\cdots+c_ma_{m,m})\mathbf{v}_m. \quad (2.1)$$

因此若我們能找到 $c_1,\ldots,c_m \in \mathbb{F}$ 使得式子 (2.1) 中每個 $\mathbf{v}_i$ 的係數等於 0, 便可得到 $c_1\mathbf{w}_1+\cdots+c_m\mathbf{w}_m=\mathbf{0}$. 因此我們只要找到聯立方程組

$$\begin{cases} a_{1,1}x_1+\cdots+a_{1,m}x_m &= 0 \\ \qquad\vdots & \\ a_{i,1}x_1+\cdots+a_{i,m}x_m &= 0 \\ \qquad\vdots & \\ a_{n,1}x_1+\cdots+a_{n,m}x_m &= 0 \end{cases}$$

的一組解 $x_1=c_1,\ldots,x_m=c_m$, 就可以使得 $c_1\mathbf{w}_1+\cdots+c_m\mathbf{w}_m=\mathbf{0}$. 然而這個 homogeneous linear system 的方程式個數 n 少於未知數個數 m. 也就是說將它對應的矩陣化為 echelon form 時, 其 pivot 的個數 (小於等於 n) 必少於 variables 的個數 m, 也就是存在著 free variables, 因此由此方程組有 $x_1=\cdots=x_n=0$ 這一組解知此方程組有其他解 (參考 Lemma 1.3.3), 即存在不全為 0 的 $c_1,\ldots,c_m \in \mathbb{F}$ 使得 $x_1=c_1,\ldots,x_m=c_m$ 為其一組解. 故得證 $\mathbf{w}_1,\ldots,\mathbf{w}_m$ 為 linearly dependent. □

Question 2.9. 假設 $\mathbf{v}_1,\ldots,\mathbf{v}_n$ 為 *linearly independent.* 若 $\mathbf{w}_1,\ldots,\mathbf{w}_k \in \mathrm{Span}(\mathbf{v}_1,\ldots,\mathbf{v}_n)$ 且 $k<n$, 試證明 $\mathrm{Span}(\mathbf{w}_1,\ldots,\mathbf{w}_k) \neq \mathrm{Span}(\mathbf{v}_1,\ldots,\mathbf{v}_n)$.

假設 W 為 V 的 subspace, 且 $\mathbf{w}_1,\ldots,\mathbf{w}_n \in W$ 為 linearly independent. 如果 $\mathbf{w}_1,\ldots,\mathbf{w}_n$ 不是 W 的 spanning vectors (即 $\mathrm{Span}(\mathbf{w}_1,\ldots,\mathbf{w}_n) \subsetneq W$), 則我們可以在 W 中選取 $\mathbf{w}_{n+1}$ 滿足 $\mathbf{w}_{n+1} \not\in \mathrm{Span}(\mathbf{w}_1,\ldots,\mathbf{w}_n)$. 此時 Lemma 2.5.4 告訴我們 $\mathbf{w}_1,\ldots,\mathbf{w}_n,\mathbf{w}_{n+1}$ 仍保持 linearly independent. 利用這個概念我們可以回答 finitely generated vector space 的 subspace 也是 finitely generated.

Proposition 2.5.6. 假設 V 為 *finitely generated vector space.* 若 W 為 V 的 *subspace,* 則 W 為 *finitely generated vector space.*

Proof. 依 V 為 finitely generated 的假設, 存在 $\mathbf{v}_1,\ldots,\mathbf{v}_n \in V$ 滿足 $\mathrm{Span}(\mathbf{v}_1,\ldots,\mathbf{v}_n)=V$. 由於 $\{\mathbf{0}\}=\mathrm{Span}(\mathbf{0})$ 為 finitely generated, 我們僅需要考慮 $W \neq \{\mathbf{0}\}$ 的情況. 我們用反證法, 假設 W 不是 finitely generated. 現任取 $\mathbf{w}_1 \in W$ 其中 $\mathbf{w}_1 \neq \mathbf{0}$. 由於 W 不是 finitely generated, 我們知 $\mathrm{Span}(\mathbf{w}_1) \neq W$, 亦即存在 $\mathbf{w}_2 \in W$ 且 $\mathbf{w}_2 \not\in \mathrm{Span}(\mathbf{w}_1)$. 由 Lemma 2.5.5 知 $\mathbf{w}_1,\mathbf{w}_2$ 為 linearly independent. 同理, 因 W 不是 finitely generated, 我們知 $\mathrm{Span}(\mathbf{w}_1,\mathbf{w}_2) \neq W$ 亦即存在 $\mathbf{w}_3 \in W$ 且 $\mathbf{w}_3 \not\in \mathrm{Span}(\mathbf{w}_1,\mathbf{w}_2)$. 由 Lemma 2.5.4 知 $\mathbf{w}_1,\mathbf{w}_2,\mathbf{w}_3$ 為 linearly independent. 這樣一直下去, 利用數學歸納法假設我們得 $\mathbf{w}_1,\ldots,\mathbf{w}_k \in W$ 為 linearly independent. 由於 $\mathrm{Span}(\mathbf{w}_1,\ldots,\mathbf{w}_k) \neq W$, 存在 $\mathbf{w}_{k+1} \in W$ 且 $\mathbf{w}_{k+1} \not\in \mathrm{Span}(\mathbf{w}_1,\ldots,\mathbf{w}_k)$. 因此再由 Lemma 2.5.4 知 $\mathbf{w}_1,\ldots,\mathbf{w}_k,\mathbf{w}_{k+}$ 為 linearly independent. 我們利用數學歸納法證明了, 若 W 不是 finitely generated, 則對任意 $m \in \mathbb{N}$, 皆存在 $\mathbf{w}_1,\ldots,\mathbf{w}_m \in W$ 為 linearly independent. 然而這在 $m>n$ 是會造成矛盾的. 因為此時由於 $\mathbf{w}_1,\ldots,\mathbf{w}_m \in W \subseteq V = \mathrm{Span}(\mathbf{v}_1,\ldots,\mathbf{v}_n)$, Lemma 2.5.5 告訴我們 $\mathbf{w}_1,\ldots,\mathbf{w}_m$ 必為 linearly dependent. 因此知 W 為 finitely generated vector space. □

Question 2.10. 利用在 $P(\mathbb{F})$ 中對於任意 $n \in \mathbb{N}$, $x^n,x^{n-1},\ldots,x,1$ 為 *linearly independent,* 證明 $P(\mathbb{F})$ 不是 *finitely generated vector space.*

前面提過若 $\mathbf{v}_1,\ldots,\mathbf{v}_n \in V$, 要知道 $\mathbf{v}\in V$ 是否屬於 $\mathrm{Span}(\mathbf{v}_1,\ldots,\mathbf{v}_n)$ 這是一個存在性的問題, 也就是問是否存在 $c_1,\ldots,c_n\in\mathbb{F}$ 使得 $\mathbf{v}=c_1\mathbf{v}_1+\cdots+c_n\mathbf{v}_n$. 但若已知 $\mathbf{v}=c_1\mathbf{v}_1+\cdots+c_n\mathbf{v}_n\in \mathrm{Span}(\mathbf{v}_1,\ldots,\mathbf{v}_n)$, 是否會存在另一組 $c'_1,\ldots,c'_n\in\mathbb{F}$ 使得 $\mathbf{v}=c'_1\mathbf{v}_1+\cdots+c'_n\mathbf{v}_n$ 呢? 這個唯一性的問題就會和 $\mathbf{v}_1,\ldots,\mathbf{v}_n$ 是否為 linearly independent 有關了. 我們有以下的結果.

Proposition 2.5.7. 假設 V 為 *vector space over* $\mathbb{F}$ 且 $\mathbf{v}_1,\ldots,\mathbf{v}_n\in V$.

(1) 若 $\mathbf{v}_1,\ldots,\mathbf{v}_n$ 為 *linearly dependent,* 則對任意 $\mathbf{v}\in\mathrm{Span}(\mathbf{v}_1,\ldots,\mathbf{v}_n)$, 將 $\mathbf{v}$ 寫成 $\mathbf{v}_1,\ldots,\mathbf{v}_n$ 的 *linear combination* 的寫法不唯一. 也就是說存在 $c_1,\ldots,c_n\in\mathbb{F}$ 以及 $c'_1,\ldots,c'_n\in\mathbb{F}$ 其中某個 $c_i\neq c'_i$ 使得 $\mathbf{v}=c_1\mathbf{v}_1+\cdots+c_n\mathbf{v}_n=c'_1\mathbf{v}_1+\cdots+c'_n\mathbf{v}_n$.

(2) 若 $\mathbf{v}_1,\ldots,\mathbf{v}_n$ 為 *linearly independent,* 則對任意 $\mathbf{v}\in\mathrm{Span}(\mathbf{v}_1,\ldots,\mathbf{v}_n)$, 將 $\mathbf{v}$ 寫成 $\mathbf{v}_1,\ldots,\mathbf{v}_n$ 的 *linear combination* 的寫法唯一. 也就是說若 $\mathbf{v}=c_1\mathbf{v}_1+\cdots+c_n\mathbf{v}_n$ 則對任意 $c'_1,\ldots,c'_n\in\mathbb{F}$ 其中某個 $c_i\neq c'_i$, 皆有 $\mathbf{v}\neq c'_1\mathbf{v}_1+\cdots+c'_n\mathbf{v}_n$.

Proof. (1) 因為 $\mathbf{v}_1,\ldots,\mathbf{v}_n$ 為 linearly dependent, 故存在 $d_1,\ldots,d_n\in\mathbb{F}$ 且某個 $d_i\neq 0$ 使得 $d_1\mathbf{v}_1+\cdots+d_n\mathbf{v}_n=\mathbf{0}$. 現若 $\mathbf{v}=c_1\mathbf{v}_1+\cdots+c_n\mathbf{v}_n$, 考慮 $c'_j=c_j+d_i\in\mathbb{F}$, for $j=1,\ldots,n$. 此時因 $d_i\neq 0$, 故知 $c_i\neq c'_i$, 但我們仍有

$$\begin{aligned}&c'_1\mathbf{v}_1+\cdots+c'_n\mathbf{v}_n=(c_1+d_1)\mathbf{v}_1+\cdots+(c_n+d_n)\mathbf{v}_n=\\&\quad(c_1\mathbf{v}_1+\cdots+c_n\mathbf{v}_n)+(d_1\mathbf{v}_1+\cdots+d_n\mathbf{v}_n)=c_1\mathbf{v}_1+\cdots+c_n\mathbf{v}_n+\mathbf{0}=c_1\mathbf{v}_1+\cdots+c_n\mathbf{v}_n.\end{aligned}$$

(2) 現假設 $\mathbf{v}_1,\ldots,\mathbf{v}_n$ 為 linearly independent, 對任意 $\mathbf{v}\in\mathrm{Span}(\mathbf{v}_1,\ldots,\mathbf{v}_n)$, 假設存在 $c_1,\ldots,c_n\in\mathbb{F}$ 以及 $c'_1,\ldots,c'_n\in\mathbb{F}$ 滿足 $\mathbf{v}=c_1\mathbf{v}_1+\cdots+c_n\mathbf{v}_n$ 以及 $\mathbf{v}=c'_1\mathbf{v}_1+\cdots+c'_n\mathbf{v}_n$, 此時 $(c_1-c'_1)\mathbf{v}_1+\cdots+(c_n-c'_n)\mathbf{v}_n=\mathbf{0}$, 故由 $\mathbf{v}_1,\ldots,\mathbf{v}_n$ 為 linearly independent 知 $c_1-c'_1=\cdots=c_n-c'_n=0$, 即 $c_1=c'_1,\ldots,c_n=c'_n$, 得證 $\mathbf{v}$ 寫成 $\mathbf{v}_1,\ldots,\mathbf{v}_n$ 的 linear combination 的寫法唯一. □

2.6. Basis and Dimension

對於一個 vector space V, 我們希望能找到一個集合 S 使得 $V=\mathrm{Span}(S)$. 這樣我們要了解 V 便只要了解 S 即可. 當然了 S 越大越容易展成 V, 但是我們又希望其越小越好, 這樣就可以用小一點的集合來了解 V. 如何找到這樣大小合宜的 S 便是 basis 的概念.

假設 V 是 over $\mathbb{F}$ 的 vector space 且 $V\neq\{\mathbf{0}\}$. 我們可以在 V 中任取非零向量 $\mathbf{v}_1$, 考慮 $\mathrm{Span}(\mathbf{v}_1)$. 若 $\mathrm{Span}(\mathbf{v}_1)=V$, 則我們找到集合 $S_1=\{\mathbf{v}_1\}$ 使得 $\mathrm{Span}(S_1)=V$, 且由於 $\mathbf{v}_1\neq\mathbf{0}$, 依定義 S_1 是 linearly independent. 若 $\mathrm{Span}(\mathbf{v}_1)\subsetneq V$, 表示存在 $\mathbf{v}_2\in V$ 且 $\mathbf{v}_1\not\in\mathrm{Span}(\mathbf{v}_1)$, 此時考慮 $S_2=\{\mathbf{v}_1,\mathbf{v}_2\}$. 由於 $\mathbf{v}_2\not\in\mathrm{Span}(\mathbf{v}_1)$, 我們知 $\mathbf{v}_1,\mathbf{v}_2$ 是 linearly independent, 故若 $\mathrm{Span}(S_2)=V$, 我們找到了 S_2 是 V 的 spanning set 且 S_2 為 linearly independent. 這樣一直下去假設我們找到 $\{\mathbf{v}_1,\ldots,\mathbf{v}_n\}$ 為 V 的 spanning set 且為 linearly independent. 此時有了 spanning set 的存在性以及 linearly independent 的唯一性, 則 V 中的元素都可以唯一寫成 $\mathbf{v}_1,\ldots,\mathbf{v}_n$ 的 linear combination, 也因此我們有以下的定義.

Definition 2.6.1. 假設 V 為 vector space. 若 $\{\mathbf{v}_1,\dots,\mathbf{v}_n\}$ 為 V 的 spanning set 且為 linearly independent, 則稱 $\mathbf{v}_1,\dots,\mathbf{v}_n$ 為 V 的一組 *basis.*

首先要注意 spanning set 未必是 linearly independent. 例如在 $\mathbb{R}^2$ 中 $\{\mathbf{v}_1=(1,0),\mathbf{v}_2=(0,1),\mathbf{v}_3=(1,1)\}$ 是 $\mathbb{R}^2$ 的 spanning set, 但不是 linearly independent (因 $\mathbf{v}_3=\mathbf{v}_1+\mathbf{v}_2$). 同樣的 linearly independent 的元素未必形成 spanning set. 例如在 $\mathbb{R}^3$ 中 $\mathbf{w}_1=(1,0,0),\mathbf{w}_2=(0,1,0)$ 就是 linearly independent, 但 $\mathrm{Span}(\mathbf{w}_1,\mathbf{w}_2)\neq\mathbb{R}^3$. 因此要說明 $\mathbf{v}_1,\dots,\mathbf{v}_n$ 為 V 的一組 basis, 必須說明 $\mathrm{Span}(\mathbf{v}_1,\dots,\mathbf{v}_n)=V$ 以及 $\mathbf{v}_1,\dots,\mathbf{v}_n$ 為 linearly independent, 缺一不可. 當然了, 若 $\{\mathbf{v}_1,\dots,\mathbf{v}_n\}$ 為 V 的 spanning set 或 $\{\mathbf{v}_1,\dots,\mathbf{v}_n\}$ 為 linearly independent 其中只要任一項不滿足, 則知 $\{\mathbf{v}_1,\dots,\mathbf{v}_n\}$ 不是 V 的 basis. 例如前面舉的例子 $\mathbf{v}_1,\mathbf{v}_2,\mathbf{v}_3$ 就不是 $\mathbb{R}^2$ 的 basis, 而 $\mathbf{w}_1,\mathbf{w}_2$ 也不是 $\mathbb{R}^3$ 的 basis. 接下來我們看幾個常見的 vector space 的 basis.

Example 2.6.2. 假設 $\mathbb{F}$ 是一個 field, 我們考慮以下常見 over $\mathbb{F}$ 的 vector spaces.

(A) 在 $\mathbb{F}^n$ 中考慮 $\mathbf{e}_1=(1,0,\dots,0),\mathbf{e}_2=(0,1,0,\dots,o),\dots,\mathbf{e}_n=(0,0,\dots,o,1)$ (即 $\mathbf{e}_i$ 為 i-th entry 為 1, 其他 entry 為 0 的向量), 由於對任意 $(c_1,\dots,c_n)\in\mathbb{F}^n$, 我們有 $(c_1,\dots,c_n)=c_1\mathbf{e}_1+\cdots+c_n\mathbf{e}_n$, 知 $\mathrm{Span}(\mathbf{e}_1,\dots,b\mathbf{e}_n)=\mathbb{F}^n$. 又若 $c_1\mathbf{e}_1+\cdots+c_n\mathbf{e}_n=\mathbf{0}$, 表示 $(c_1,\dots,c_n)=(0,\dots,0)$, 亦即 $c_1=\cdots=c_n=0$, 故知 $\mathbf{e}_1,\dots,b\mathbf{e}_n$ 為 linearly independent. 因此 $\mathbf{e}_1,\dots,\mathbf{e}_n$ 為 $\mathbb{F}^n$ 的一組 basis. 這一組 basis 是 $\mathbb{F}^n$ 中最直接且最常用的 basis 所以我們又稱之為 $\mathbb{F}^n$ 的 *standard basis.*

(B) 在 $M_{m\times n}(\mathbb{F})$ 中, 考慮 E_{ij}, 為 (i,j)-th entry 為 1, 其他 entry 為 0 的 $m\times n$ matrix. 則利用和 $\mathbb{F}^n$ 上類似的證法, 我們可以知 $\{E_{ij}:1\leq i\leq m,1\leq j\leq n\}$ 是 $M_{m\times n}(\mathbb{F})$ 的 spanning set 且為 linearly independent, 故為 $M_{m\times n}(\mathbb{F})$ 的一組 basis.

(C) 在 $P_n(\mathbb{F})$ 中 $\{1,x,x^2,\dots,x^n\}$ 可展成 $P_n(\mathbb{F})$ 且為 linearly independent, 所以是 $P_n(\mathbb{F})$ 的 basis. 這組 basis 也稱為 $P_n(\mathbb{F})$ 的 *standard basis.*

在 Definition 2.6.1 中 V 可由有限多個元素所展成, 所以此時的 V 依定義是 finitely generated vector space (Definition 2.4.3), 不過我們提過一般的 vector space 未必會是 finitely generated, 所以對於一般的 vector space, 我們有以下 basis 的定義.

Definition 2.6.3. 假設 V 為 vector space 且 $S\subseteq V$. 若 S 為 V 的 spanning set 且為 linearly independent, 則稱 S 為 V 的一組 *basis.*

由於我們已定義 $\mathrm{Span}(\emptyset)=\{\mathbf{0}\}$ 且 $\emptyset$ 為 linearly independent, 所以依照 Definition 2.6.3, 我們說空集合 $\emptyset$ 是 zero vector space $\{\mathbf{0}\}$ 的 basis. 另外在 Example 2.4.4 中我們知道 $P(\mathbb{F})$ 不是 finitely generated. 然而很容易看出 $\{1,x,x^2,\dots\}$ 是 $P(\mathbb{F})$ 的 spanning set 且為 linearly independent, 所以 $\{1,x,x^2,\dots\}$ 是 $P(\mathbb{F})$ 的 basis.

我們碰到的第一個問題便是 basis 的存在性問題. 也就是說, 對於一般的 vector space 是不是都會有 basis. 接下來我們要說明非零的 finitely generated vector space 都會有 basis. 其實這對於不是 finitely generated 的 vector space 也對, 不過由於這會牽涉到較抽象的邏

輯概念而且我們以後談論的 vector space 都是 finitely generated, 所以我們不去談論它. 在本講義中我們僅探討 finitely generated vector space.

Theorem 2.6.4. 假設 $V \neq \{\mathbf{0}\}$ 為 *finitely generated vector space over* $\mathbb{F}$. 則存在 $\mathbf{v}_1, \ldots, \mathbf{v}_n \in V$ 為 V 的一組 *basis.*

Proof. 這裡我們利用 finitely generated 的性質, 用數學歸納法處理.

我們對 vector space 的 spanning set 的元素個數做數學歸納法. 首先假設 V 可由一個元素展成. 也就是說 $V = \mathrm{Span}(\mathbf{u})$. 此時由於 $V \neq \{\mathbf{0}\}$, 知 $\mathbf{u} \neq \mathbf{0}$. 故 $\mathbf{u}$ 本身是 linearly independent 且 $\{\mathbf{u}\}$ 是 V 的 spanning set 故 $\mathbf{u}$ 是 V 的 basis. 利用歸納法假設當一個 vector space 可由 k 個元素展成時, basis 是存在的. 現假設 V 是一個可由 $k+1$ 個元素展成的 vector space, 我們假設 $V = \mathrm{Span}(\mathbf{u}_1, \ldots, \mathbf{u}_k, \mathbf{u}_{k+1})$. 現考慮 $W = \mathrm{Span}(\mathbf{u}_1, \ldots, \mathbf{u}_k)$. 由於 W 是一個可由 k 個元素展成的 vector space, 依歸納法假設存在 $\mathbf{v}_1, \ldots, \mathbf{v}_n \in W$ 為 W 的一組 basis. 亦即 $\mathrm{Span}(\mathbf{v}_1, \ldots, \mathbf{v}_n) = W$ 且 $\mathbf{v}_1, \ldots, \mathbf{v}_n$ 為 linearly independent. 現若 $W = V$, 當然 $\mathbf{v}_1, \ldots, \mathbf{v}_n$ 就是 V 的一組 basis. 而若 $W \neq V$, 表示 $\mathbf{u}_{k+1} \notin W$ (否則會造成 $V = \mathrm{Span}(\mathbf{u}_1, \ldots, \mathbf{u}_k, \mathbf{u}_{k+1}) \subseteq W$ 之矛盾). 因此由 $\mathbf{u}_{k+1} \notin W = \mathrm{Span}(\mathbf{v}_1, \ldots, \mathbf{v}_n)$ 以及 Lemma 2.5.4 知 $\mathbf{v}_1, \ldots, \mathbf{v}_n, \mathbf{u}_{k+1}$ 為 linearly independent. 又因為 $V = \mathrm{Span}(\mathbf{u}_1, \ldots, \mathbf{u}_k, \mathbf{u}_{k+1})$ 而 $\mathrm{Span}(\mathbf{u}_1, \ldots, \mathbf{u}_k) = W = \mathrm{Span}(\mathbf{v}_1, \ldots, \mathbf{v}_n)$ 故得 $V = \mathrm{Span}(\mathbf{v}_1, \ldots, \mathbf{v}_n, \mathbf{u}_{k+1})$. 得證 $\mathbf{v}_1, \ldots, \mathbf{v}_n, \mathbf{u}_{k+1}$ 為 V 的一組 basis. □

在 Theorem 2.6.4 的證明中, 看起來是將 $V = \mathrm{Span}(\mathbf{u}_1, \ldots, \mathbf{u}_k, \mathbf{u}_{k+1})$ 找到其他的 $\mathbf{v}_1, \ldots, \mathbf{v}_n$ 取代 $\mathbf{u}_1, \ldots, \mathbf{u}_k$, 得到 V 的 basis $\mathbf{v}_1, \ldots, \mathbf{v}_n, \mathbf{u}_{k+1}$. 事實上我們是可以在 $\{\mathbf{u}_1, \ldots, \mathbf{u}_k, \mathbf{u}_{k+1}\}$ 中挑選出 V 的 basis. 主要的原因是, 如果 $S = \{\mathbf{u}_1, \ldots, \mathbf{u}_k, \mathbf{u}_{k+1}\}$ 本身是 linearly independent, 則依定義 S 就是一組 basis. 而若 $\{\mathbf{u}_1, \ldots, \mathbf{u}_k, \mathbf{u}_{k+1}\}$ 是 linearly dependent, 表示存在 $\mathbf{u}_i$ 是其他的 $S' = S \setminus \{\mathbf{u}_i\}$ 的 linear combination. 注意此時 $\S(S') = \S(S) = V$ 因此我們可以考慮 S' 是否為 linearly independent. 若是 linearly independent, 則 S' 就是 V 的一組 basis, 否則就如前繼續下去直到其為 linearly independent 為止. 因此利用數學歸納法, 我們有以下的結果.

Proposition 2.6.5. 假設 $V \neq \{\mathbf{0}\}$ 為 *vector space over* $\mathbb{F}$, 且 $S \subseteq V$ 滿足 $V = \mathrm{Span}(S)$. 則存在 $S' \subseteq S$ 為 V 的一組 *basis.*

Proof. 我們對 S 的元素個數 n 做數學歸納法. 假設 $n = 1$, 即 $V = \mathrm{Span}(\mathbf{v}_1)$, 因 $V \neq \{\mathbf{0}\}$, 知 $\mathbf{v}_1 \neq \mathbf{0}$. 故由 $\mathbf{v}_1$ 本身是 linearly independent, 得證 $S = \{\mathbf{v}_1\}$ 是 V 的 basis. 假設 $n = k$ 時成立, 亦即對所有有 k 元素的集合 S, 若 $V = \mathrm{Span}(S)$, 則存在 $S' \subseteq S$ 為 V 的 basis. 我們要證明當 $n = k+1$ 時亦成立. 現假設 $S = \{\mathbf{v}_1, \ldots, \mathbf{v}_{k+1}\}$ 且 $\mathrm{Span}(S) = V$, 若 S 為 linearly independent, 則依定義令 $S' = S$, 即為 V 的一組 basis. 而若 S 不是 linearly independent, 則不失一般性, 我們假設 $\mathbf{v}_1 \in \mathrm{Span}(\mathbf{v}_2, \ldots, \mathbf{v}_n)$. 此時令 $\tilde{S} = S \setminus \{\mathbf{v}_1\} = \{\mathbf{v}_2, \ldots, \mathbf{v}_1\}$, 因 $V = \mathrm{Span}(\tilde{S})$ 且 $\tilde{S}$ 的元素個數為 k, 故由歸納假設知存在 $S' \subseteq \tilde{S} \subseteq S$ 為 V 的一組 basis. 得證本定理. □

Question 2.11. 試利用 *Proposition 2.6.5* 的結果證明 *Theorem 2.6.4.*

Proposition 2.6.5 說的是當 $\mathbf{v}_1, \ldots, \mathbf{v}_n$ 是 V 的 spanning set, 則我們可以在 $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ 中捨去一些元素使其為 linearly independent 且仍為 V 的 spanning set, 故可成為 V 的一組

basis. 反之, 若 $\{\mathbf{v}_1,\ldots,\mathbf{v}_n\}$ 為 V 的一組 linearly independent set, 則我們可以加入一些元素擴大 $\{\mathbf{v}_1,\ldots,\mathbf{v}_n\}$, 使其成為 V 的 spanning set 且仍保持 linearly independent, 故可成為 V 的一組 basis. 我們有以下的定理.

Proposition 2.6.6. 假設 $V \neq \{\mathbf{0}\}$ 為 *vector space over* $\mathbb{F}$, 且 $V = \mathrm{Span}(\mathbf{u}_1,\ldots,\mathbf{u}_m)$. 若 $S = \{\mathbf{v}_1,\ldots,\mathbf{v}_n\} \subseteq V$ 為 *linearly independent,* 則存在 $S' \subseteq \{\mathbf{u}_1,\ldots,\mathbf{u}_m\}$ 使得 $S \cup S'$ 為 V 的 *basis.*

Proof. 我們依然對 S 的元素個數做歸納法, 不過這次是做反向的歸納法. 注意因 $V = \mathrm{Span}(\mathbf{u}_1,\ldots,\mathbf{u}_m)$ 且 $\mathbf{v}_1,\ldots,\mathbf{v}_n$ 為 linearly independent, 故由 Lemma 2.5.5 知 $n \leq m$, 所以我們可以假設 $n = m-t$, 其中 $0 \leq t \leq m-1$. 我們要對 t 做 mathematical induction. 當 $t = 0$ 時表示 $m = n$, 此時我們要說明 $S = \{bv_1,\ldots,\mathbf{v}_n\}$ 為 V 的 spanning set. 這是因為若 $\mathrm{Span}(S) \neq V$, 表示存在 $\mathbf{w} \in V$ 且 $\mathbf{w} \notin \mathrm{Span}(S)$, 故由 Lemma 2.5.4 知 $\{\mathbf{v}_1,\ldots,\mathbf{v}_n,\mathbf{w}\}$ 為 linearly independent, 但此時 $\{\mathbf{v}_1,\ldots,\mathbf{v}_n.\mathbf{w}\}$ 有 $n+1 = m+1$ 個元素, 多於 $\mathrm{Span}(\mathbf{u}_1,\ldots,\mathbf{u}_m)$ 的 m 個元素, 與 Lemma 2.5.5 相矛盾, 故知 S 為 V 的 spanning set. 再利用已知 S 為 linearly independent, 故令 $S' = \emptyset$ 得證 $S = S \cup S'$ 為 V 的 basis. 現假設 $t = k$ 時成立, 即若 $S = \{\mathbf{v}_1,\ldots,\mathbf{v}_{m-k}\}$ 為 linearly independent, 則存在 $S' \subseteq \{\mathbf{u}_1,\ldots,\mathbf{u}_m\}$ 使得 $S \cup S'$ 為 V 的 basis. 現考慮 $t = k+1$ 的情形, 即 $S = \{\mathbf{v}_1,\ldots,\mathbf{v}_{m-k-1}\}$ 為 linearly independent. 首先考慮 S 是否為 V 的 spanning set. 若 $V = \mathrm{Span}(S)$, 則依定義知 S 為 V 的一組 basis, 故取 $S' = \emptyset$, 則 $S' \subseteq \{\mathbf{u}_1,\ldots,\mathbf{u}_m\}$, 且 $S \cup S' = S$ 為 V 的 basis. 而若 $\mathrm{Span}(S) \neq V$, 表示 $\{\mathbf{u}_1,\ldots,\mathbf{u}_m\}$ 中必有一元素 $\mathbf{u}_i \notin \mathrm{Span}(S)$, 否則若 $\mathbf{u}_1,\ldots,\mathbf{u}_m$ 皆屬於 $\mathrm{Span}(S)$, 會造成 $V = \mathrm{Span}(\mathbf{u}_1,\ldots,\mathbf{u}_m) \subseteq \mathrm{Span}(S)$ 之矛盾. 現考慮 $\tilde{S} = S \cup \{\mathbf{u}_i\}$, 因 $\mathbf{u}_i \notin \mathrm{Span}(S)$ 由 Lemma 2.5.4 知 $\tilde{S}$ 為 linearly independent, 再由 $\tilde{S}$ 的元素個數為 $m-k$, 故由歸納假設知存在 $\tilde{S}' \subseteq \{\mathbf{u}_1,\ldots,\mathbf{u}_m\}$ 使得 $\tilde{S} \cup \tilde{S}'$ 為 V 的 basis. 故令 $S' = \{\mathbf{u}_i\} \cup \tilde{S}'$, 我們有 $S' \subseteq \{\mathbf{u}_1,\ldots,\mathbf{u}_m\}$ 且 $S \cup S' = S \cup \{\mathbf{u}_i\} \cup \tilde{S}' = \tilde{S} \cup \tilde{S}'$ 為 V 的 basis. □

我們已經知道 finitely generated vector space 的 basis 是存在的, 不過它並不唯一. 例如在 $\mathbb{R}^2$ 中除了 standard basis $\{(1,0),(0,1)\}$ 外, 我們很容易看出 $\{(1,1),(0,1)\}$ 也可以是 $\mathbb{R}^2$ 的 basis. 不過 basis 雖然不唯一, 不過在 finitely generated vector space 中組成 basis 的元素個數是固定的. 我們有以下的定理.

Theorem 2.6.7. 假設 V 為 *vector space over* $\mathbb{F}$, 且 $\{\mathbf{v}_1,\ldots,\mathbf{v}_n\}$ 和 $\{\mathbf{u}_1,\ldots,\mathbf{u}_m\}$ 皆為 V 的 *basis,* 則 $n = m$.

Proof. 我們用反證法, 假設 $n \neq m$, 不失一般性我們就假設 $m > n$. 因 $V = \mathrm{Span}(\mathbf{v}_1,\ldots,\mathbf{v}_n)$, 故 $\mathbf{u}_i \in \mathrm{Span}(\mathbf{v}_1,\ldots,\mathbf{v}_n)$, $\forall i = 1,\ldots,m$. 因此由 Lemma 2.5.5 知 $\{\mathbf{u}_1,\ldots,\mathbf{u}_m\}$ 是 linearly dependent. 此與 $\{\mathbf{u}_1,\ldots,\mathbf{u}_m\}$ 是 basis 的假設相矛盾, 故得證 $m = n$. □

Theorem 2.6.7 告訴我們組成 V 的一組 basis 的元素個數是固定的. 也就是說若找到 n 個元素形成 V 的 basis, 則 V 其他的 basis 一定也會是由 n 個元素所組成. 由於這個重要的結果我們有以下的定義.

Definition 2.6.8. 假設 V 是一個 finitely generated vector space over $\mathbb{F}$. 組成 V 的一組 basis 的元素個數稱為 V over $\mathbb{F}$ 的 *dimension* (維度), 用 $\dim_{\mathbb{F}}(V)$ 來表示.

由於組成 finitely generated vector space 的一組 basis 的元素個數是有限的, 所以以後我們稱 finitely generated vector space 為 *finite dimensional vector space.*

Example 2.6.9. 我們探討在 Example 2.6.2 中的 finite dimensional vector space 的維度為多少.

(A) 考慮 $\mathbb{F}^n$ 中的 standard basis $\{\mathbf{e}_i \mid 1 \le i \le n\}$, 因為共有 n 個元素所以 $\dim_{\mathbb{F}}(\mathbb{F}^n) = n$.

(B) 我們知道 $\{E_{ij} \in \mathscr{M}_{m\times n} \mid 1 \le i \le m, 1 \le j \le n\}$ 是 $\mathscr{M}_{m\times n}(\mathbb{F})$ 的一組 basis. 因此 $\dim \mathbb{F}(\mathscr{M}_{m\times n}) = m \times n$.

(C) 我們知道 $\{x^n, \ldots, x, 1\}$ 是 $P_n(\mathbb{F})$ 的 spanning set 且為 linearly independent. 故知 $\{x^n, \ldots, x, 1\}$ 為 $P_n(\mathbb{F})$ 的一組 basis, 因此 $\dim_{\mathbb{F}}(P_n(\mathbb{F})) = n+1$.

大家或許注意到, 我們在表示維度的 dim 符號的下標特別標上 $\mathbb{F}$, 即 $\dim_{\mathbb{F}}$. 這個原因是強調我們將 vector space 看成 over $\mathbb{F}$ 的 vector space 所得的 dimension. 我們曾經提過, 同樣的集合我們有可能看成 over 不同的 field 的 vector space. 在此情況之下它們的 basis 也就會不同, 也因此我們要標示用哪一個 field. 我們看以下的例子.

Example 2.6.10. 我們用 $\mathbb{C}$ 表示 complex numbers (複數) 所成的 field, 而用 $\mathbb{R}$ 表示 real numbers (實數) 所成的 field. 現考慮集合 $\mathbb{C}^2 = \{(z_1, z_2) \mid z_1, z_2 \in \mathbb{C}\}$. 很容易檢查用一般的加法及係數積, $\mathbb{C}^2$ 是 vector space over $\mathbb{C}$, 也會是 vector space over $\mathbb{R}$. 首先我們知道 $\{(1,0),(0,1)\}$ 是 $\mathbb{C}^2$ 看成 over $\mathbb{C}$ 的 vector space 的 basis (Example 2.6.2 (A) $n = 2$, $\mathbb{F} = \mathbb{C}$ 的情況), 所以我們得 $\dim_{\mathbb{C}}(\mathbb{C}^2) = 2$. 不過 $\{(1,0),(0,1)\}$ 在 over $\mathbb{R}$ 之下就不是 basis 了. 很容易看出來任何 $(1,0),(0,1)$ over $\mathbb{R}$ 的 linear combination 都無法表示 $(i,0)$ 這一個 $\mathbb{C}^2$ 的元素 (這裡 i 是滿足 $i^2 = -1$ 的純虛數). 不過 $\{(1,0),(i,0),(0,1),(0,i)\}$ 就是 $\mathbb{C}^2$ over $\mathbb{R}$ 的 spanning set. 這是因為任意 $\mathbb{C}^2$ 的元素都可以寫成 $(a+bi, c+di)$ 其中 $a,b,c,d \in \mathbb{R}$, 所以可得 $(a+bi, c+di) = a(1,0)+b(i,0)+c(0,1)+d(0,i)$. 又 $\{(1,0),(i,0),(0,1),(0,i)\}$ over $\mathbb{R}$ 是 linearly independent. 這是因為若 a,b,c,d 是實數滿足 $a(1,0)+b(i,0)+c(0,1)+d(0,i) = (0,0)$, 即表示 $a+bi = 0$ 且 $c+di = 0$, 故得 $a = b = c = d = 0$. 由此知 $\{(1,0),(i,0),(0,1),(0,i)\}$ 是 $\mathbb{C}^2$ over $\mathbb{R}$ 的 basis, 所以我們有 $\dim_{\mathbb{R}}(\mathbb{C}^2) = 4$.

由 Example 2.6.10, 我們知道要說明一個 vector space 的 dimension 為何, 一定要說明其 over 的 field 是甚麼. 不過一般情形, 當我們很明確知道 over 的 field 是甚麼而沒有如 Example 2.6.10 這種模稜兩可的情形, 我們便會省略直接用 $\dim(V)$ 來表示其 dimension.

對於 finite dimensional vector space 有關於 dimension 的性質, 我們匯集如下. 再次強調, 由於這裡我們只考慮 over 一個固定的 field $\mathbb{F}$, 所以我們僅用 $\dim(V)$ 來表示其 dimension.

Proposition 2.6.11. 假設 V 為 *finite dimensional vector space over* $\mathbb{F}$.

(1) 若 $\{\mathbf{v}_1,\ldots,\mathbf{v}_n\}$ 為 V 的 *spanning set*, 則 $\dim(V)\le n$. 特別的, 若此時 $\mathbf{v}_1,\ldots,\mathbf{v}_n$ 為 *linearly dependent*, 則 $\dim(V)<n$.

(2) 若 $\{\mathbf{v}_1,\ldots,\mathbf{v}_n\}$ 為 *linearly independent*, 則 $\dim(V)\ge n$. 特別的, 若此時 $\{\mathbf{v}_1,\ldots,\mathbf{v}_n\}$ 不是 V 的 *spanning set*, 則 $\dim(V)>n$.

(3) 假設 $\mathbf{v}_1,\ldots,\mathbf{v}_n\in V$. 下列的敘述為等價.
  (a) $\mathbf{v}_1,\ldots,\mathbf{v}_n$ 為 V 的一組 *basis*.
  (b) $\dim(V)=n$ 且 $\{\mathbf{v}_1,\ldots,\mathbf{v}_n\}$ 為 V 的 *spanning set*.
  (c) $\dim(V)=n$ 且 $\{\mathbf{v}_1,\ldots,\mathbf{v}_n\}$ 為 *linearly independent*.

(4) 若 W 為 V 的 *subspace*, 則 $\dim(W)\le\dim(V)$. 特別的, 若 $\dim(W)=\dim(V)$ 則 $W=V$.

Proof. 為了方便起見, 我們令 $S=\{\mathbf{v}_1,\ldots,\mathbf{v}_n\}$.

(1) 依假設 $V=\mathrm{Span}(S)$, 故利用 Proposition 2.6.5 知存在 $S'\subseteq S$ 為 V 的 basis. 也就是說 S' 的元素個數就是 V 的 dimension. 然而 S' 是 S 的 subset, 所以其元素個數小於等於 S 的元素個數 n, 故得證 $\dim(V)\le n$. 現若 S 為 linearly dependent, 即表示存在 $\mathbf{v}_i$ 可寫成 S 中其他元素的線性組合, 因此考慮 $\tilde{S}=S\setminus\{\mathbf{v}_i\}$, 我們仍有 $\mathrm{Span}(\tilde{S})=V$. 此時 $\tilde{S}$ 的元素個數為 $n-1$, 所以再套用前面所證的可得 $\dim(V)\le n-1<n$.

(2) 依假設 S 是 linearly independent, 故利用 Proposition 2.6.6 知存在某個有限集合 S' 使得 $S\cup S'$ 為 V 的 basis. 也就是說 $S\cup S'$ 的元素個數就是 V 的 dimension. 然而 $S\subseteq S\cup S'$, 所以 $S\cup S'$ 的元素個數大於等於 S 的元素個數 n, 故得證 $\dim(V)\ge n$. 現若 S 不是 V 的 spanning set, 表示存在 $\mathbf{w}\in V$ 且 $\mathbf{w}\not\in\mathrm{Span}(S)$, 因此考慮 $\tilde{S}=S\cup\{\mathbf{w}\}$, 我們仍有 $\tilde{S}$ 為 linearly independent (Lemma 2.5.4). 此時 $\tilde{S}$ 的元素個數為 $n+1$, 所以再套用前面所證的可得 $\dim(V)\ge n+1>n$.

(3) 我們要證明 (a) 可推得 (b), (b) 可推得 (c) 以及 (c) 可推得 (a). 因此知 (a),(b),(c) 是等價的.

(a) ⇒ (b): 假設 S 是 V 的 basis, 當然 S 是 V 的 spanning set. 又由於 S 的元素個數為 n, 依定義 $\dim(V)=n$.

(b) ⇒ (c): 由於 S 是 V 的 spanning set, 由前面 (1) 的結果, 若 S 是 linearly dependent, 則 $\dim(V)<n$. 此與 $\dim(V)=n$ 假設相矛盾, 故推得 S 是 linearly independent.

(c) ⇒ (a): 由於 S 是 linearly independent, 由前面 (2) 的結果, 若 S 不是 V 的 spanning set, 則 $\dim(V)>n$. 此與 $\dim(V)=n$ 假設相矛盾, 故推得 S 是 V 的 spanning set. 因此得證 S 是 V 的 basis.

(4) 因 W 是 V 的 subspace, 故由 Proposition 2.5.6 知 W 亦為 finite dimensional vector space, 我們假設 $S=\{\mathbf{v}_1,\ldots,\mathbf{v}_n\}$ 為 W 的 basis. 由於 $\mathbf{v}_1,\ldots,\mathbf{v}_n\in V$ 且為 linearly independent, 故由 (2) 的結果知 $\dim(W)=n\le\dim(V)$. 而若 $\dim(V)=n$, 則由 S 是 linearly independent 利用 (3)((c) ⇒ (a)) 知 S 也是 V 的 basis, 故得證 $W=V$. □

強調一下, Proposition 2.6.11 告訴我們知道 V 的 dimension 的好處. 若我們知道 $\dim(V)$ 恰好是 n, 則 (3) 告訴我們當要檢查 $\mathbf{v}_1,\ldots,\mathbf{v}_n$ 是否為 V 的一組 basis 時, 則僅要檢查它們是否為 spanning set 或 linearly independent 其中一項就可. 所以我們只要選擇檢查哪一個較好處理即可. 另外若已知 W 為 V 的 subspace, 要檢查 W 是否為 V, 我們不必再像以前檢查是否每個 V 中的元素都在 W, 而只要算出 $\dim(W)$ 是否等於 n 即可.

Example 2.6.12. 很容易看出在 $P(\mathbb{R})$ 中 $x^n, x^{n-1},\ldots,x,1$ 為 linearly independent. 這是因為如果 $c_n,\ldots,c_1,c_0\in\mathbb{R}$ 使得 $c_nx^n+\cdots+c_1x+c_01$ 為零多項式, 則依零多項式的定義, $c_n,\ldots,c_1,c_0$ 必全為 0. 現在我們介紹 $P(\mathbb{R})$ 中另一種重要的 linearly independent 的多項式的建構方法, 稱為 *Lagrange interpolation polynomials*. 我們僅舉出一個例子, 一般狀況請大家自行推廣.

給定 a,b,c 三相異實數, 我們希望找到三個二次多項式 $p_1(x),p_2(x),p_3(x)$ 滿足

$$p_1(a)=1, p_1(b)=p_1(c)=0, \quad p_2(b)=1, p_2(a)=p_2(c)=0 \quad \text{and} \quad p_3(c)=1, p_3(a)=p_3(b)=0.$$

由於 $p_1(b)=p_1(c)=0$, 我們知 $p_1(x)$ 應為 $(x-b)(x-c)$ 的倍式, 也就是存在實數 r 使得 $p_1(x)=r(x-b)(x-c)$. 但又要求 $p_1(a)=1$, 故代入 $x=a$ 得 $r=1/(a-b)(a-c)$. 同理可求出 $p_2(x),p_3(x)$ 因此我們有

$$p_1(x)=\frac{(x-b)(x-c)}{(a-b)(a-c)}, \quad p_2(x)=\frac{(x-a)(x-c)}{(b-a)(b-c)} \quad \text{and} \quad p_3(x)=\frac{(x-a)(x-b)}{(c-a)(c-b)}.$$

我們要說明 $p_1(x),p_2(x),p_3(x)$ 為 linearly independent. 首先觀察, 若 $f(x)=c_1p_1(x)+c_2p_2(x)+c_3p_3(x)$, 則代入 $x=a$ 時可由 $p_1(a)=1, p_2(a)=p_3(a)=0$, 得 $f(a)=c_1$. 同理知 $f(b)=c_2, f(c)=c_3$. 因此現若 $f(x)$ 為零多項式, 由 $f(a)=f(b)=f(c)=0$, 可得 $c_1=c_2=c_3=0$. 也就是說只有當 $c_1=c_2=c_3=0$ 時才會使得 $c_1p_1(x)+c_2p_2(x)+c_3p_3(x)$ 為零多項式, 得證 $p_1(x),p_2(x),p_3(x)$ 為 linearly independent.

我們知道了 $p_1(x),p_2(x),p_3(x)\in P_2(\mathbb{R})$ 為 linearly independent, 事實上 $p_1(x),p_2(x),p_3(x)$ 會是 $P_2(\mathbb{R})$ 的 spanning set. 不過要證明這一點, 需用到次數小於 3 的實係數非零多項式不會有 3 個相異實根這個事實, 說明起來有點麻煩. 不過由於我們已知 $\dim(P_2(\mathbb{R}))=3$, 故由 Proposition 2.6.11 (3) 知 $p_1(x),p_2(x),p_3(x)$ 為 $P_2(\mathbb{R})$ 的一組 basis. 也就是說任何實係數的次數小於 3 的多項式 $f(x)$ 都可以都可以找到唯一的一組 $c_1,c_2,c_3\in\mathbb{R}$ 使得 $f(x)=c_1p_1(x)+c_2p_2(x)+c_3p_3(x)$. 事實上代入 $x=a,b,c$, 我們知道 $c_1=f(a)$, $c_2=f(b)$, $c_3=f(c)$ 就是這唯一的一組.

同理對於任意 n 個相異實數 $a_1,\ldots,a_n$, 我們有 n 個 $n-1$ 次的多項式 $p_1(x),\ldots,p_n(x)$ 滿足 $p_i(a_i)=1$ 且當 $j\neq i$ 時, $p_i(a_j)=0$. 由於 $p_1(x),\ldots,p_n(x)\in P_{n-1}(\mathbb{R})$ 且為 linearly independent, 故由 $\dim(P_{n-1}(\mathbb{R}))=n$ 知 $p_1(x),\ldots,p_n(x)$ 為 $P_{n-1}(\mathbb{R})$ 的一組 basis.

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.

3.1. 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 $M_{m\times n}(\mathbb{F})$ 來表示所有係數在 $\mathbb{F}$ 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. 例如考慮係數為實數 $\mathbb{R}$, 若

$$A=\begin{bmatrix} 1 & 0 & 2 & 3\\ 0 & 1 & 5 & 8\\ 2 & 1 & 1 & 0 \end{bmatrix}, \tag{3.1}$$

則 A 為一個 3×4 matrix, 即 $A\in M_{3\times 4}(\mathbb{R})$. 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 (3.1) 中的矩陣 A, 若 $A=[a_{ij}]$, 則

$$\begin{array}{llll} a_{11}=1, & a_{12}=0, & a_{13}=2, & a_{14}=3,\\ a_{21}=0, & a_{22}=1, & a_{23}=5, & a_{24}=8,\\ a_{31}=2, & a_{32}=1, & a_{33}=1, & a_{34}=0. \end{array}$$

另外為了方便起見, 我們也會將矩陣 A 的每一個 row 和 column 用向量的方法來表示, 這些稱為 A 的 row vectors 和 column vectors. 在本講義我們會將矩陣 $A=[a_{ij}]$ 第 i 個 row 所成的 row vector 用 ${}_i\mathbf{a}$ 來表示, 而第 j 個 column 所成的 column vector 用 $\mathbf{a}_j$ 來表示. 例如對於式子 (3.1) 中的矩陣 A, 我們有

$${}_1\mathbf{a}=\begin{bmatrix}1 & 0 & 2 & 3\end{bmatrix},\quad {}_2\mathbf{a}=\begin{bmatrix}0 & 1 & 5 & 8\end{bmatrix},\quad {}_3\mathbf{a}=\begin{bmatrix}2 & 1 & 1 & 0\end{bmatrix}$$

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 (就如同我們曾定義 $\mathbb{R}^n$ 中向量的相等).

Definition 3.1.1. 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 $M_{m\times n}(\mathbb{F})$ 的矩陣才談是否相等, 且兩個 $M_{m\times n}(\mathbb{F})$ 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 $M_{m\times n}(\mathbb{F})$ 的矩陣我們才定義它們之間的加法, 且兩矩陣相加表示將這兩個矩陣在相同位置的數加起來. 而一個實數乘上一個矩陣即為將該矩陣每一個位置上的數乘上該實數. 具體來說我們有以下的定義.

Definition 3.1.2. 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 3.1.2 告訴我們若

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

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

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

既然矩陣的加法與係數積的定義是由 $\mathbb{R}^n$ 的加法與係數積延伸而來, 我們可以預期矩陣的加法與係數積應有和 $\mathbb{R}^n$ 的加法與係數積相同的性質. 事實上我們確有 $M_{m\times n}(\mathbb{F})$ 為

vector space over $\mathbb{F}$, 這些性質的證明和 $\mathbb{R}^n$ 的情形相同 (用到 $\mathbb{F}$ 為 field 相對應的性質), 我們就不再重複了.

Proposition 3.1.3. 對於 $M_{m\times n}(\mathbb{F})$ 上的矩陣, 我們有以下的性質:

(1) 對任意 $A,B\in M_{m\times n}(\mathbb{F})$, 皆有 $A+B=B+A$.

(2) 對任意 $A,B,C\in M_{m\times n}(\mathbb{F})$, 皆有 $(A+B)+C=A+(B+C)$.

(3) 存在一矩陣 $O\in M_{m\times n}(\mathbb{F})$ 滿足對任意 $A\in M_{m\times n}(\mathbb{F})$ 皆有 $O+A=A$.

(4) 對任意 $A\in M_{m\times n}(\mathbb{F})$ 皆可找到 $A'\in M_{m\times n}(\mathbb{F})$ 滿足 $A+A'=O$.

(5) 對任意 $r,s\in\mathbb{F}$ 以及 $A\in M_{m\times n}(\mathbb{F})$, 皆有 $r(sA)=(rs)A$.

(6) 對任意 $r,s\in\mathbb{F}$ 以及 $A\in M_{m\times n}(\mathbb{F})$, 皆有 $(r+s)A=rA+sA$.

(7) 對任意 $r\in\mathbb{F}$ 以及 $A,B\in M_{m\times n}(\mathbb{F})$ 皆有 $r(A+B)=rA+rB$.

(8) 對任意 $A\in M_{m\times n}(\mathbb{F})$, 皆有 $1A=A$.

接著我們定義矩陣間的乘法. 首先回顧當我們要說明 $(1,2,3)\in\mathbb{R}^3$ 是否屬於 $\mathrm{Span}((1,-1,2),(2,1,-2))$ 的問題, 等同於是否存在 $x_1,x_2\in\mathbb{R}$ 滿足 $(1,2,3)=x_1(1,-1,2)+x_2(2,1,-2)$. 若將它們寫成 column vector 的形式, 即

$$x_1\begin{bmatrix}1\\-1\\2\end{bmatrix}+x_2\begin{bmatrix}2\\1\\-2\end{bmatrix}=\begin{bmatrix}1\\2\\3\end{bmatrix},$$

可看出此又等同於聯立方程組

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

是否有解的問題. 而這樣的聯立方程組, 我們又把它寫成

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

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 3.1.4. 設 $A=[a_{ij}]$ 為 $m\times n$ matrix 以及 $\mathbf{b}=[b_j]$ 為 $n\times 1$ matrix (即 $\mathbb{F}^n$ 中的 column vector). 若 $\mathbf{a}_i$ 表示 A 的 i-th column 則定義

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

注意依此定義, 必需 $A\in M_{m\times n}(\mathbb{F})$ 的 column 的個數 n 等於 $\mathbf{b}\in M_{n\times 1}(\mathbb{F})$ 的 row 的個數 n, 才能定義 $A\mathbf{b}$ 且此時 $A\mathbf{b}$ 會是 $m\times 1$ matrix (即 $\mathbb{F}^m$ 中的 column vector). 觀察此 column vector, 我們有

$$A\mathbf{b}=b_1\begin{bmatrix} a_{11}\\ a_{21}\\ \vdots\\ a_{m1}\end{bmatrix}+b_2\begin{bmatrix} a_{12}\\ a_{22}\\ \vdots\\ a_{m2}\end{bmatrix}+\cdots+b_n\begin{bmatrix} a_{1n}\\ a_{2n}\\ \vdots\\ a_{mn}\end{bmatrix}=\begin{bmatrix} b_1a_{11}+b_2a_{12}+\cdots+b_na_{1n}\\ b_1a_{21}+b_2a_{22}+\cdots+b_na_{2n}\\ \vdots\\ b_1a_{m1}+b_2a_{m2}+\cdots+b_na_{mn}\end{bmatrix}. \tag{3.2}$$

特別的, 當 $\mathbf{a}=\begin{bmatrix} a_1 & a_2 & \cdots & a_n\end{bmatrix}$ 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 3.1.4 的矩陣乘法定義

$$\mathbf{a}\,\mathbf{b}=\begin{bmatrix} a_1 & a_2 & \cdots & a_n\end{bmatrix}\begin{bmatrix} b_1\\ b_2\\ \vdots\\ b_n\end{bmatrix}=b_1a_1+b_2a_2+\cdots+b_na_n. \tag{3.3}$$

(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 (3.2), 我們可將 $A\mathbf{b}$ 寫成

$$A\mathbf{b}=\begin{bmatrix} {}_1\mathbf{a}\,\mathbf{b}\\ {}_2\mathbf{a}\,\mathbf{b}\\ \vdots\\ {}_m\mathbf{a}\,\mathbf{b}\end{bmatrix}. \tag{3.4}$$

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

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

Lemma 3.1.5. *假設 $A=[a_{ij}]$, $A'=[a'_{ij}]$ 為係數在 $\mathbb{F}$ 的 $m\times n$ matrices 以及 $\mathbf{b}=[b_j]$, $\mathbf{b}'=[b'_j]$ 為 $\mathbb{F}^n$ 中的 column vectors (即 $n\times 1$ matrices) 以及 $c\in\mathbb{F}$. 我們有以下的性質.*

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

(2) $A(c\mathbf{b})=c(A\mathbf{b})=(cA)\mathbf{b}$.

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

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

(1) 依定義

$$A(\mathbf{b}+\mathbf{b}') = \begin{bmatrix} | & & | \\ \mathbf{a}_1 & \cdots & \mathbf{a}_n \\ | & & | \end{bmatrix} \begin{bmatrix} b_1+b_1' \\ \vdots \\ b_n+b_n' \end{bmatrix} = (b_1+b_1')\mathbf{a}_1+\cdots+(b_n+b_n')\mathbf{a}_n. \tag{3.5}$$

而 $A\mathbf{b}+A\mathbf{b}'$ 為

$$\begin{bmatrix} | & & | \\ \mathbf{a}_1 & \cdots & \mathbf{a}_n \\ | & & | \end{bmatrix} \begin{bmatrix} b_1 \\ \vdots \\ b_n \end{bmatrix} + \begin{bmatrix} | & & | \\ \mathbf{a}_1 & \cdots & \mathbf{a}_n \\ | & & | \end{bmatrix} \begin{bmatrix} b_1' \\ \vdots \\ b_n' \end{bmatrix} = (b_1\mathbf{a}_1+\cdots+b_n\mathbf{a}_n)+(b_1'\mathbf{a}_1+\cdots+b_n'\mathbf{a}_n). \tag{3.6}$$

由 $\mathbb{F}^n$ 向量加法和係數積的分配律 (vector space 的性質 (7),(8)), 我們得證式子 (3.5) 和式子 (3.6) 相等.

(2) 依定義 $c(A\mathbf{b})$ 為

$$c\left(\begin{bmatrix} | & & | \\ \mathbf{a}_1 & \cdots & \mathbf{a}_n \\ | & & | \end{bmatrix} \begin{bmatrix} b_1 \\ \vdots \\ b_n \end{bmatrix}\right) = c(b_1\mathbf{a}_1+b_2\mathbf{a}_2+\cdots+b_n\mathbf{a}_n). \tag{3.7}$$

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

$$\begin{bmatrix} | & & | \\ c\mathbf{a}_1 & \cdots & c\mathbf{a}_n \\ | & & | \end{bmatrix} \begin{bmatrix} b_1 \\ \vdots \\ b_n \end{bmatrix} = b_1(c\mathbf{a}_1)+\cdots+b_n(c\mathbf{a}_n). \tag{3.8}$$

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

$$\begin{bmatrix} | & & | \\ \mathbf{a}_1 & \cdots & \mathbf{a}_n \\ | & & | \end{bmatrix} \begin{bmatrix} cb_1 \\ \vdots \\ cb_n \end{bmatrix} = (cb_1)\mathbf{a}_1+\cdots+(cb_n)\mathbf{a}_n. \tag{3.9}$$

由 $\mathbb{F}^n$ 向量加法和係數積的結合律 (vector space 的性質 (6)), 我們得證 (3.7), (3.8), (3.9) 三個式子皆相等.

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

$$(A+A')\mathbf{b} = \begin{bmatrix} | & & | \\ \mathbf{a}_1+\mathbf{a}_1' & \cdots & \mathbf{a}_n+\mathbf{a}_n' \\ | & & | \end{bmatrix} \begin{bmatrix} b_1 \\ \vdots \\ b_n \end{bmatrix} = b_1(\mathbf{a}_1+\mathbf{a}_1')+\cdots+b_n(\mathbf{a}_n+\mathbf{a}_n'). \tag{3.10}$$

另一方面, $A\mathbf{b}+A'\mathbf{b}$ 為

$$\begin{bmatrix} | & & | \\ \mathbf{a}_1 & \cdots & \mathbf{a}_n \\ | & & | \end{bmatrix} \begin{bmatrix} b_1 \\ \vdots \\ b_n \end{bmatrix} + \begin{bmatrix} | & & | \\ \mathbf{a}_1' & \cdots & \mathbf{a}_n' \\ | & & | \end{bmatrix} \begin{bmatrix} b_1 \\ \vdots \\ b_n \end{bmatrix} = (b_1\mathbf{a}_1+\cdots+b_n\mathbf{a}_n)+(b_1\mathbf{a}_1'+\cdots+b_n\mathbf{a}_n'). \tag{3.11}$$

再次由 $\mathbb{F}^n$ 向量加法和係數積的分配律, 我們得證式子 (3.10) 和式子 (3.11) 相等. □

Lemma 3.1.5 (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 3.1.6. 假設 $A \in M_{m\times n}(\mathbb{F})$, 且 $\mathbf{b},\mathbf{b}'$ 為 $\mathbb{F}^n$ 中的 *column vectors,* 以及 $c,c' \in \mathbb{F}$. 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 $\mathbb{F}^n$ 中的 column vectors, 由 Lemma 3.1.5 (1) 知 $A(c\mathbf{b}+c'\mathbf{b}') = A(c\mathbf{b})+A(c'\mathbf{b}')$. 再由 Lemma 3.1.5 (2) 知 $A(c\mathbf{b}) = c(A\mathbf{b}), A(c'\mathbf{b}') = c'(A\mathbf{b}')$, 故得證本定理. □

Question 3.1. 假設 $A \in M_{m\times n}(\mathbb{F})$, 且 $\mathbf{b}_1,\dots,\mathbf{b}_k$ 為 $\mathbb{F}^n$ 中的 *column vectors,* $c_1,\dots,c_k \in \mathbb{F}$. 試利用數學歸納法證明

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

現在我們將矩陣乘法推廣到更一般的情況, 當 $A=[a_{ij}]$ 是一個 $m\times n$ matrix, $B=[b_{jk}]$ 是一個 $n\times l$ matrix. 由於對 B 的每一個 column vector $\mathbf{b}_k \in M_{n\times 1}(\mathbb{F})$, $1\le k\le l$, 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.

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

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 3.1.7. 設 $A=[a_{ij}]$ 為 $m\times n$ matrix 以及 $B=[b_{jk}]$ 為 $n\times l$ matrix, 則定義 $AB=C=[c_{ik}]$ 為 $m\times l$ matrix, 其中對於 $1\le k\le l$, C 的 k-th column $\mathbf{c}_k$ 為

$$\mathbf{c}_k = A\mathbf{b}_k = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}\begin{bmatrix} b_{1k} \\ b_{2k} \\ \vdots \\ b_{nk} \end{bmatrix} = b_{1k}\mathbf{a}_1 + b_{2k}\mathbf{a}_2 + \cdots + b_{nk}\mathbf{a}_n. \tag{3.12}$$

由此定義, 我們知對於 $1\le i\le m$, $1\le k\le l$, 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 (3.4) 我們知此即 A 的 i-th row ${}_i\mathbf{a}$ 和 B 的 k-th column $\mathbf{b}_k$ 看成向量後取內積. 換言之, 若 $AB=[c_{ik}]$, 則 AB 的 (i,k)-th entry c_{ik} 為

$$c_{ik} = {}_i\mathbf{a}\,\mathbf{b}_k = a_{i1}b_{1k}+a_{i2}b_{2k}+\cdots+a_{in}b_{nk} = \sum_{j=1}^{n} a_{ij}b_{jk}. \tag{3.13}$$

再次強調一次, 並不是任取兩個矩陣都可以定義乘法, 必須是左邊矩陣的 column 個數和右邊矩陣的 row 個數相同才能相乘.

Example 3.1.8. 令

$$A = \begin{bmatrix} -2 & 4 \\ 3 & 6 \\ 2 & 2 \end{bmatrix}, B = \begin{bmatrix} 4 & -1 & 2 & 5 \\ 3 & 0 & 1 & 1 \end{bmatrix}$$

考慮矩陣乘法 AB. 依定義矩陣 AB 的 3-rd column 為

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

所以 AB 的 $(2,3)$ entry 為 12 等於 A 的 2-nd row 和 B 的 3-rd column 看成 $\mathbb{R}^2$ 中的向量所得的內積, 即 $(3,6)\cdot(2,1) = 12$. 事實上我們有

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

Question 3.2. 設 $\mathbf{a},\mathbf{b}$ 為 $\mathbb{R}^n$ 上的向量, 若將 $\mathbf{a}$ 寫成 *row vector* 的形式, $\mathbf{b}$ 寫成 *column vector* 的形式, 且將 $\mathbf{a},\mathbf{b}$ 看成矩陣, 即 $\mathbf{a} \in M_{1\times n}(\mathbb{R})$, $\mathbf{b} \in M_{n\times 1}(\mathbb{R})$. 試問依矩陣乘法定義 $\mathbf{ba}$ 應為何種矩陣? 它和 $\mathbf{a},\mathbf{b}$ 看成 $\mathbb{R}^n$ 上的向量後取內積 $\langle \mathbf{b},\mathbf{a}\rangle$ 有關嗎?

大部分的書都會用式子 (3.13) 當成矩陣乘法的定義. 我們選用式子 (3.12) 的用意, 主要是它較能描繪當初矩陣乘法定義的用意. 另外它是由 column 來描繪矩陣的乘法, 在證明或推導有關矩陣乘法性質時, 有時比式子 (3.13) 利用 entry 來看方便多了. 例如我們有以下的性質.

Proposition 3.1.9. *假設 $A,A' \in M_{m\times n}(\mathbb{F})$, $B,B' \in M_{n\times l}(\mathbb{F})$. 我們有以下的性質.*

(1) $A(B+B') = AB+AB'$.

(2) $(A+A')B = AB+A'B$.

Proof. 首先注意因 $A+A'$ 仍為 $m\times n$ 矩陣且 $B+B'$ 為 $n\times l$ 矩陣, 所以這些矩陣的階數是符合矩陣乘法的規定. 我們假設 B 的 column vectors 依次為 $\mathbf{b}_1,\ldots,\mathbf{b}_l$ 且 B' 的 column vectors 依次為 $\mathbf{b}'_1,\ldots,\mathbf{b}'_l$.

(1) 我們證明當 $1\le k\le l$ 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 3.1.5 (1), 我們得證它們相等.

(2) 我們證明當 $1\le k\le l$ 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 3.1.5 (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 3.1.10. *設 $c\in\mathbb{F}$, $A \in M_{m\times n}(\mathbb{F}), B \in M_{n\times l}(\mathbb{F})$. 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 r 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 3.1.5 (2), 我們得證它們皆相等. □

由 Proposition 3.1.9 和 Proposition 3.1.10 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 3.1.11. 假設 $A\in M_{m\times n}(\mathbb{F}),B\in M_{n\times l}(\mathbb{F}),C\in M_{l\times k}(\mathbb{F})$, 則 $(AB)C=A(BC)$.

Proof. 依定義 $AB\in M_{m\times l}(\mathbb{F})$, 故 $(AB)C\in M_{m\times k}(\mathbb{F})$. 而 $BC\in M_{n\times k}(\mathbb{F})$, 故 $A(BC)\in M_{m\times k}(\mathbb{F})$ 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 $\mathbb{F}^n$ 中的 column vector, 故套用 Proposition 3.1.6 (或 Question 3.1) 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 3.1.11), 以後我們談多個矩陣相乘時, 為了方便起見, 我們會捨去括號例如直接用 ABC 表示. 特別的, 當 A 為方陣時, 既然 $(AA)A=A(AA)$, 我們就用 A^3 來表示. 同理, 當 n 個 A 相乘時, 我們就用 A^n 來表示.

最後我們要強調的是矩陣乘法雖具有許多和實數乘法類似的性質, 但它卻沒有交換律. 事實上有可能 A 乘以 B 有定義, 但 B 卻不能乘以 A, 例如 $A \in M_{2\times 3}(\mathbb{F})$, $B \in M_{3\times 4}(\mathbb{F})$ 的情形. 也有可能即使 A 乘以 B 和 B 乘以 A 都有定義, 但由於乘了以後階數不同, 仍會使得 $AB \neq BA$, 例如 $A \in M_{2\times 3}(\mathbb{F}), B \in M_{3\times 2}(\mathbb{F})$ 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 3.1.9 和 Proposition 3.1.10 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 $A \in M_{n\times n}(\mathbb{F})$, 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 $\mathbf{e}_i$, 其中 $\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n$ 就是我門曾提過 $\mathbb{R}^n$ 的 standard basis (寫成 column vector). 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 $A \in M_{m\times n}(\mathbb{F}), B \in M_{n\times l}(\mathbb{F})$ 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 3.3. 假設 $A \in M_{n\times n}(\mathbb{F})$, 是否 $(A - 2I_n)^2 = A^2 - 4A + 4I_n$ 為對?

Question 3.4. 試證明 I_n 是唯一的 $n \times n$ 滿足對任意 $A \in M_{m\times n}(\mathbb{F})$ 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 $r \in \mathbb{F}$, rI_n 亦為 diagonal matrix. 因為它 diagonal entry 皆為 r, 其他 entry 皆為 0. 對於任意 $A \in M_{m\times n}(\mathbb{F}), B \in M_{n\times l}(\mathbb{F})$ 我們很容易驗證 $rA = A(rI_n), rB = (rI_n)B$.

Question 3.5. 試利用 *Proposition 3.1.10* 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.

3.2. 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 3.2.1. 給定 $A\in M_{m\times n}(\mathbb{F})$. 定義 $A^{\mathrm{t}}\in M_{n\times m}(\mathbb{F})$, 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 3.2.2. 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 $A\in M_{m\times n}(\mathbb{F})$, 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 $A^{\mathrm{t}}=[a'_{st}]$, 則 $1\leq s\leq n$, $1\leq t\leq m$, 且 $a'_{st}=a_{ts}$. 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 $\begin{bmatrix}a_{1j} & a_{2i} & \cdots & a_{mj}\end{bmatrix}$ 寫成 row vector. 我們將以上的討論寫成以下的結論.

Lemma 3.2.3. 假設 $A=[a_{ij}]\in M_{m\times n}(\mathbb{F})$ 且 $A^{\mathrm{t}}=[a'_{st}]\in M_{n\times m}(\mathbb{F})$. 則對於 $1\leq s\leq n$, $1\leq t\leq m$, $a'_{st}=a_{ts}$ 且 A^{t} 的 *s-th row* 就是將 A 的 *s-th column* 寫成 *row vector.*

由 Lemma 3.2.3, 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 3.2.4. 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 3.2.3 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 $A=[a_{ij}]$, $A^{\mathrm{t}}=[a'_{ji}]$, $C=[c_{st}]$, $C^{\mathrm{t}}=[c'_{ts}]$ 表示. 對於 $1\le t\le l$, $1\le i\le m$, $(AC)^{\mathrm{t}}$ 的 (t,i)-th entry 為 AC 的 (i,t)-th entry, 由式子 (3.13) 知應為

$$a_{i1}c_{1t}+a_{i2}c_{2t}+\cdots+a_{in}c_{nt}.$$

另一方面, $C^{\mathrm{t}}A^{\mathrm{t}}$ 的 (t,i)-th entry 為

$$c'_{t1}a'_{1i}+c'_{t2}a'_{2i}\cdots+c'_{tn}a'_{ni}.$$

利用 Lemma 3.2.3 知此即為

$$c_{1t}a_{i1}+c_{2t}a_{i2}\cdots+c_{nt}a_{in}.$$

故得證 $(AC)^{\mathrm{t}}=C^{\mathrm{t}}A^{\mathrm{t}}$. □

Question 3.6. 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 3.2.5. 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 3.2.4, 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 3.2.4, 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 $A\in M_{1\times m}(\mathbb{F}), B\in M_{m\times n}(\mathbb{F})$, 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 3.2.4 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

$$\begin{bmatrix}a_1 & \cdots & a_m\end{bmatrix}\begin{bmatrix}b_{11} & \cdots & b_{1n}\\ \vdots & \cdots & \vdots\\ b_{m1} & \cdots & b_{mn}\end{bmatrix}=a_1\begin{bmatrix}b_{11} & \cdots & b_{1n}\end{bmatrix}+\cdots+a_m\begin{bmatrix}b_{m1} & \cdots & b_{mn}\end{bmatrix} \tag{3.14}$$

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 3.2.4 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

$$AB=\begin{bmatrix}- & {}_1\mathbf{a} & -\\ - & {}_2\mathbf{a} & -\\ & \vdots & \\ - & {}_m\mathbf{a} & -\end{bmatrix}B=\begin{bmatrix}- & {}_1\mathbf{a}B & -\\ - & {}_2\mathbf{a}B & -\\ & \vdots & \\ - & {}_m\mathbf{a}B & -\end{bmatrix}. \tag{3.15}$$

結合式子 (3.14), 我們有以下之結果.

Proposition 3.2.6. *設 $A=[a_{ij}]$ 為 $m\times n$ matrix 以及 $B=[b_{jk}]$ 為 $n\times l$ matrix, 則對於 $1\leq i\leq m$, AB 的 i-th row 為*

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

3.3. 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. 我們曾經提過, 若將 A 做一個 elementary row operation, 就等同將 A 的左邊乘上相對應的 elementary matrix. 這一節中我們將更深入探討 elementary row operation 和 elementary matrix 的關係.

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 3.2.6, 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, 即

$$E=\begin{bmatrix} 1 &&&&&& \\ & \ddots &&&&& \\ && 0 && 1 && \\ &&& \ddots &&& \\ && 1 && 0 && \\ &&&&& \ddots & \\ &&&&&& 1 \end{bmatrix} \tag{3.16}$$

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, 即

$$E=\begin{bmatrix} 1 &&&&&& \\ & \ddots &&&&& \\ && 1 &&&& \\ &&& r &&& \\ &&&& 1 && \\ &&&&& \ddots & \\ &&&&&& 1 \end{bmatrix} \tag{3.17}$$

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, 即

$$E=\begin{bmatrix} 1 & & & & & & \\ & \ddots & & & & & \\ & & 1 & & & & \\ & & & \ddots & & & \\ & & r & & 1 & & \\ & & & & & \ddots & \\ & & & & & & 1 \end{bmatrix} \tag{3.18}$$

則因 E 的 j-th row 的 i-th entry 為 r, j-th entry 為 1. 故由 Proposition 3.2.6, 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. (3.16), (3.17), (3.18) 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 3.7. 試找出那些同階的 *elementary matrices* 其相乘是不可以交換的.

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 3.3.1. *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 3.3.2. 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

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

最後我們要再次強調 elementary matrices 之間相乘是沒有交換性的. 事實上一個矩陣若右邊乘上一個 elementary matrix, 是對此矩陣做所謂的 elementary column operation. 這用類似前面的看法, 改用 column 的觀點來處理矩陣乘法 (Definition 3.1.7) 便可看出, 這裡就不再多說明了.

Example 3.3.3. 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

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

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

3.4. 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 $\mathbb{F}^n$ 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 $\mathbf{x}=\mathbf{c}\in\mathbb{F}^n$ 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 $A\mathbf{c}=\mathbf{b}$.

3.4.1. 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 $\mathbf{b}\in\mathbb{F}^m$, 聯立方程組 $A\mathbf{x}=\mathbf{b}$ 皆有解.

首先假設 $\mathbf{b}\in\mathbb{F}^m$ 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. 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 $c_1,\ldots,c_n\in\mathbb{F}$ 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 3.4.1. 假設 $A\in M_{m\times n}(\mathbb{F})$ 且 $\mathbf{b}\in\mathbb{F}^m$. 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 $\mathbf{b}\in\mathbb{F}^m$, 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 $\mathbb{F}^m$ 中, 所以自然有 $\mathrm{Span}(\mathbf{a}_1,\ldots,\mathbf{a}_n)\subseteq\mathbb{F}^m$. 然而由 Lemma 3.4.1 知, 若對於任意 $\mathbf{b}\in\mathbb{F}^m$ 皆會使得 $A\mathbf{x}=\mathbf{b}$ 有解, 表示對任意 $\mathbf{b}\in\mathbb{F}^m$ 皆有 $\mathbf{b}\in\mathrm{Span}(\mathbf{a}_1,\ldots,\mathbf{a}_n)$. 故知此時 $\mathrm{Span}(\mathbf{a}_1,\ldots,\mathbf{a}_n)=\mathbb{F}^m$. 反之, 若 $\mathrm{Span}(\mathbf{a}_1,\ldots,\mathbf{a}_n)=\mathbb{F}^m$, 表示對任意 $\mathbf{b}\in\mathbb{F}^m$ 皆有 $\mathbf{b}\in\mathrm{Span}(\mathbf{a}_1,\ldots,\mathbf{a}_n)$. 同樣由 Lemma 3.4.1 知此即對於任意 $\mathbf{b}\in\mathbb{F}^m$ 皆會使得 $A\mathbf{x}=\mathbf{b}$ 有解. 因此從這觀點來看, 對任意的 $\mathbf{b}\in\mathbb{F}^m$, 聯立方程組 $A\mathbf{x}=\mathbf{b}$ 皆有解和 $\mathrm{Span}(\mathbf{a}_1,\ldots,\mathbf{a}_n)=\mathbb{F}^m$ 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 $\mathbf{e}_1,\ldots,\mathbf{e}_m\in\mathbb{F}^m$ 為 $\mathbb{F}^m$ 的 standard basis. 若已知對任意的 $\mathbf{b}\in\mathbb{F}^m$, 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 $\mathbf{c}_i\in\mathbb{F}^n$ 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 $\mathbf{b}\in\mathbb{F}^m$, 我們考慮 $\mathbf{c}=C\mathbf{b}\in\mathbb{F}^n$, 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 $\mathbf{b}\in\mathbb{F}^m$, 我們都可以找到 $\mathbf{c}=C\mathbf{b}\in\mathbb{F}^n$, 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 $\mathbf{b}\in\mathbb{F}^m$, 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 $\mathbf{b}\in\mathbb{F}^m$, 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 $\mathbf{b}\in\mathbb{F}^m$ 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 $\mathbf{b}\in\mathbb{F}^m$, 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 3.4.2. *假設 $A\in M_{m\times n}(\mathbb{F})$, 令 $\mathbf{a}_1,\ldots,\mathbf{a}_n\in\mathbb{F}^m$ 為 A 的 column vectors. 以下各敘述是等價的.*

(1) *對任意的 $\mathbf{b}\in\mathbb{F}^m$, 聯立方程組 $A\mathbf{x}=\mathbf{b}$ 皆有解.*

(2) $\mathrm{Span}(\mathbf{a}_1,\ldots,\mathbf{a}_n)=\mathbb{F}^m$.

(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 3.4.2, 指的是對所有 $\mathbf{b}\in\mathbb{F}^m$, 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 3.4.2 並不適用 (不過 Lemma 3.4.1 是適用的).

我們曾提及, 當 $A\in M_{m\times n}(\mathbb{F})$, 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 3.4.2 中的情況不可能發生. 我們有以下的結論.

Corollary 3.4.3. *假設 $A\in M_{m\times n}(\mathbb{F})$, 其中 $n<m$, 則必存在 $\mathbf{b}\in\mathbb{F}^m$ 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 3.4.2 知不可能對任意的 $\mathbf{b}\in\mathbb{F}^m$, 聯立方程組 $A\mathbf{x}=\mathbf{b}$ 皆有解.

亦即存在 $\mathbf{b}\in\mathbb{F}^m$ 使得聯立方程組 $A\mathbf{x}=\mathbf{b}$ 無解. 同理, 由 Theorem 3.4.2 知不會存在 $n\times m$ matrix C 使得 $AC=I_m$. □

Question 3.8. 假設 $A\in M_{m\times n}(\mathbb{F})$, 其中 $m<n$. 是否存在 $n\times m$ *matrix* C 使得 $CA=I_n$?

前面提過 Theorem 3.4.2 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 3.4.3 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 $\mathbf{b}\in\mathbb{F}^m$ 使得聯立方程組 $A\mathbf{x}=\mathbf{b}$ 無解.

3.4.2. 解的唯一性. 所謂聯立方程組解的唯一性, 指的是假設聯立方程組有解時, 探討其解是否唯一. 所以唯一性並不涉及解是否存在的問題.

給定 $A\in M_{m\times n}(\mathbb{F})$ 以及 $\mathbf{b}\in\mathbb{F}^m$. 如果 $A\mathbf{x}=\mathbf{b}$ 有解, 則 $A\mathbf{x}=\mathbf{b}$ 的解和 $A\mathbf{x}=\mathbf{0}$ 的解 (這裡 $\mathbf{0}$ 是 $\mathbb{F}^m$ 的零向量) 息息相關, 我們有以下之定理.

Lemma 3.4.4. 給定 $A\in M_{m\times n}(\mathbb{F})$ 以及 $\mathbf{b}\in\mathbb{F}^m$ 且假設 $\mathbf{x}=\mathbf{c}\in\mathbb{F}^n$ 是聯立方程組 $A\mathbf{x}=\mathbf{b}$ 的一組解. 則

(1) 若 $\mathbf{x}=\mathbf{c}'\in\mathbb{F}^n$ 是 $A\mathbf{x}=\mathbf{b}$ 的一組解, 則 $\mathbf{x}=\mathbf{c}'-\mathbf{c}$ 為 $A\mathbf{x}=\mathbf{0}$ 的一組解.

(2) 若 $\mathbf{x}=\mathbf{u}\in\mathbb{F}^n$ 為 $A\mathbf{x}=\mathbf{0}$ 的一組解, 則 $\mathbf{x}=\mathbf{c}+\mathbf{u}$ 是 $A\mathbf{x}=\mathbf{b}$ 的一組解.

Proof. (1) 假設 $\mathbf{x}=\mathbf{c}'\in\mathbb{F}^n$ 是 $A\mathbf{x}=\mathbf{b}$ 的一組解, 意即 $A\mathbf{c}'=\mathbf{b}$. 由已知 $A\mathbf{c}=\mathbf{b}$ 得

$$A(\mathbf{c}'-\mathbf{c})=A\mathbf{c}'-A\mathbf{c}=\mathbf{b}-\mathbf{b}=\mathbf{0}.$$

因此 $\mathbf{x}=\mathbf{c}'-\mathbf{c}$ 會是 $A\mathbf{x}=\mathbf{0}$ 的一組解.

(2) 若 $\mathbf{x}=\mathbf{u}\in\mathbb{F}^n$ 為 $A\mathbf{x}=\mathbf{0}$ 的一組解, 則

$$A(\mathbf{c}+\mathbf{u})=A\mathbf{c}+A\mathbf{u}=\mathbf{b}+\mathbf{0}=\mathbf{b}.$$

得證 $\mathbf{x}=\mathbf{c}+\mathbf{u}$ 為聯立方程組 $A\mathbf{x}=\mathbf{b}$ 的一組解. □

Lemma 3.4.4 告訴我們若已知 $\mathbf{x}=\mathbf{c}$ 為 $A\mathbf{x}=\mathbf{b}$ 的一組解, 且知道 $A\mathbf{x}=\mathbf{0}$ 所有的解, 就能利用 $\mathbf{c}$ 以及 $A\mathbf{x}=\mathbf{0}$ 所有的解得到 $A\mathbf{x}=\mathbf{b}$ 所有的解. 所以了解 $A\mathbf{x}=\mathbf{0}$ 所有的解是很重要的課題 (以後我們會深入探討). 回顧一下 $A\mathbf{x}=\mathbf{0}$ 這樣的 linear system, 我們稱之為 homogeneous linear system. Homogeneous linear system 一定有解, 事實上當 $A\in M_{m\times n}(\mathbb{F})$ 時, $x_1=0,\ldots,x_n=0$ 就是 $A\mathbf{x}=\mathbf{0}$ 的一組解. 這組解 $\mathbf{x}=\mathbf{0}\in\mathbb{F}^n$ 因為不需任何計算就能得到, 我們稱之為 $A\mathbf{x}=\mathbf{0}$ 的 trivial solution. 注意 trivial solution $\mathbf{x}=\mathbf{0}$ 這裡的 $\mathbf{0}$ 是 $\mathbb{F}^n$ 的零向量, 而 $A\mathbf{x}=\mathbf{0}$ 這裡的 $\mathbf{0}$ 是 $\mathbb{F}^m$ 的零向量, 所以雖然我們用同樣的符號表示, 但當 $n\neq m$ 時它們是不同的, 大家需區分清楚. 當一個 homogeneous linear system $A\mathbf{x}=\mathbf{0}$ 除了 trivial solution 外還有其他的 solution (即解不唯一), 我們稱這些不為 $\mathbf{0}$ 的 solution 為 nontrivial solution.

從 Lemma 3.4.4 我們知, 若 $A\mathbf{x}=\mathbf{0}$ 沒有 nontrivial solution (即解唯一), 則對於 $\mathbf{b}\in\mathbb{F}^m$, 若 $A\mathbf{x}=\mathbf{b}$ 有解, 其解必唯一. 由這觀點, 我們可以得到以下關於聯立方程組解的唯一性的重要定理.

Theorem 3.4.5. 假設 $A\in M_{m\times n}(\mathbb{F})$. 以下各敘述是等價的.

(1) 若 $\mathbf{b} \in \mathbb{F}^m$ 且聯立方程組 $A\mathbf{x} = \mathbf{b}$ 有解, 則解唯一.

(2) *Homogeneous system* $A\mathbf{x} = \mathbf{0}$ 沒有 *nontrivial solution.*

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

(4) 存在 $n \times m$ *matrix* B 使得 $BA = I_n$.

Proof. (1) $\Rightarrow$ (2): 利用反證法, 假設 $\mathbf{x} = \mathbf{u}$ 為 $A\mathbf{x} = \mathbf{0}$ 的一組 nontrivial solution 而 $\mathbf{x} = \mathbf{c}$ 為 $A\mathbf{x} = \mathbf{b}$ 的一組解, 則由 Lemma 3.4.4 知 $\mathbf{x} = \mathbf{c} + \mathbf{u} \neq \mathbf{c}$ 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 $A\mathbf{x} = \mathbf{0}$ 沒有 nontrivial solution.

(2) $\Rightarrow$ (3): $A\mathbf{x} = \mathbf{0}$ 沒有 nontrivial solution, 表示 A 化成 echelon form 後沒有 free variables. 也就是說所有的 variables 皆為 pivot variables. 因此 pivot 的個數就是未知數的個數 n, 故得 $\operatorname{rank}(A) = n$.

(3) $\Rightarrow$ (4): 假設 $\operatorname{rank}(A) = n$, 即 A 化為 echelon form 後, 其 pivot 的個數為 n. 考慮將 A 化為 reduced echelon form A'. 此時 A' 由於有 n 個 pivot, 所以每一個 pivot 必分別在 A' 前面 n 個 row 上. 而又 A' 為 $m \times n$ matrix, 有 n 個 column. 所以 A' 每一個 pivot 必落在 (i,i)-th entry, 其中 $1 \leq i \leq n$. 又因為 A' 為 reduced echelon form, 此 n 個 pivots 的值皆為 1. 然而 reduced echelon form 每一個 pivot 所在的 column, 除了 pivot 所在位置外, 其他位置應為 0, 所以我們知 A' 必為以下的 matrix $A' = \begin{bmatrix} I_n \\ \mathbf{0} \end{bmatrix}$, 即 A' 的前 n 個 row 就是 I_n. 由 Lemma 3.3.1, 我們知存在 $m \times m$ matrix E 使得 $EA = A'$. 現若令 E 的 i-th row 為 ${}_i\varepsilon$, 由 row 的觀點看矩陣乘法 (參見圖示 (3.15)), 我們有

$$EA = \begin{bmatrix} — & {}_1\varepsilon & — \\ & \vdots & \\ — & {}_n\varepsilon & — \\ & \vdots & \\ — & {}_m\varepsilon & — \end{bmatrix} A = \begin{bmatrix} — & {}_1\varepsilon A & — \\ & \vdots & \\ — & {}_n\varepsilon A & — \\ & \vdots & \\ — & {}_m\varepsilon A & — \end{bmatrix} = \begin{bmatrix} 1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & 1 \\ & \mathbf{0} & \end{bmatrix}.$$

現若令 B 為 $n \times m$ matrix, 對於 $i = 1, \ldots, n$, 其 i-th row 為 ${}_i\varepsilon$ (即 B 為截取 E 的前 n 個 row 的 $n \times m$ matrix), 則由前述的矩陣乘法性質知

$$BA = \begin{bmatrix} — & {}_1\varepsilon & — \\ & \vdots & \\ — & {}_n\varepsilon & — \end{bmatrix} A = \begin{bmatrix} — & {}_1\varepsilon A & — \\ & \vdots & \\ — & {}_n\varepsilon A & — \end{bmatrix} = \begin{bmatrix} 1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & 1 \end{bmatrix} = I_n.$$

(4) $\Rightarrow$ (1): 我們利用反證法假設 $\mathbf{c} \neq \mathbf{c}' \in \mathbb{F}^n$ 且 $\mathbf{x} = \mathbf{c}, \mathbf{x} = \mathbf{c}'$ 皆為 $A\mathbf{x} = \mathbf{b}$ 的一組解. 亦即, $A\mathbf{c} = \mathbf{b}$ 且 $A\mathbf{c}' = \mathbf{b}$. 現已知存在 $B \in M_{n \times m}(\mathbb{F})$ 使得 $BA = I_n$, 故得 $B\mathbf{b} = B(A\mathbf{c}) = (BA)\mathbf{c} = \mathbf{c}$ 且 $B\mathbf{b} = B(A\mathbf{c}') = (BA)\mathbf{c}' = \mathbf{c}'$. 此結果 $\mathbf{c} = B\mathbf{b} = \mathbf{c}'$ 與當初假設 $\mathbf{c} \neq \mathbf{c}'$ 相矛盾, 故得證若 $A\mathbf{x} = \mathbf{b}$ 有解, 則解必唯一. □

再次提醒, Theorem 3.4.5, 並不能知道聯立方程組 $A\mathbf{x} = \mathbf{b}$ 是否有解. 它告訴我們若已知 $A\mathbf{x} = \mathbf{0}$ 有 nontrivial solution, 則 $A\mathbf{x} = \mathbf{b}$ 要不然無解, 要不然會有無窮多解.

Question 3.9. 假設 $A \in M_{m\times n}(\mathbb{F})$, $\mathbf{b}, \mathbf{b}' \in \mathbb{F}^m$. 若已知 $A\mathbf{x} = \mathbf{b}$, $A\mathbf{x} = \mathbf{b}'$ 皆有解, 且 $A\mathbf{x} = \mathbf{b}$ 的解唯一. 是否 $A\mathbf{x} = \mathbf{b}'$ 的也會唯一?

前面已提過, 當 $A \in M_{m\times n}(\mathbb{F})$, 將 A 化為 echelon form 後其 pivot 的個數應小於等於 $\min\{m,n\}$. 所以若 pivot 的個數為 n, 則表示 $m \geq n$. 換言之, 若 $m < n$, 我們便知 pivot 的個數不可能等於 n, 所以 Theorem 3.4.5 中的情況不可能發生. 我們有以下的結論.

Corollary 3.4.6. *假設 $A \in M_{m\times n}(\mathbb{F})$, 其中 $m < n$. 若 $\mathbf{b} \in \mathbb{F}^m$ 且聯立方程組 $A\mathbf{x} = \mathbf{b}$ 有解, 則解不唯一 (即必有兩個以上的解). 而且此時, 不會存在 $n \times m$ matrix B 使得 $BA = I_n$.*

Proof. 由前所述, 當 $m < n$ 時 A 化為 echelon form 後, 其 pivot 的個數不可能為 n. 故由 Theorem 3.4.5 知 homogeneous linear system $A\mathbf{x} = \mathbf{0}$ 有 nontrivial solution. 亦即在 $\mathbb{F}^n$ 存在非零向量 $\mathbf{c}$ 使得 $\mathbf{x} = \mathbf{c}$ 為 $A\mathbf{x} = \mathbf{0}$ 之一組解. 所以 Lemma 3.4.4 告訴我們, 若 $A\mathbf{x} = \mathbf{b}$ 有解, 則解不唯一.

另一方面 Theorem 3.4.5 也告訴我們若 pivot 的個數不是 n, 則不會存在 $n \times m$ matrix B 使得 $BA = I_n$. □

Theorem 3.4.5 也和 Theorem 3.4.2 一樣是很重要的定理, 它可以告訴我們一些解聯立方程組的訊息. 例如 Corollary 3.4.6 就是告訴我們當方程式的個數少於未知數的個數時, 聯立方程組 $A\mathbf{x} = \mathbf{b}$ 不可能有唯一解.

3.5. Invertible Matrix

所謂 invertible matrix 就是“可逆矩陣”. 我們會發現只有 square matrix 才有可能是 invertible matrix, 但並不是所有的 square matrix 都是 invertible matrix. 這一節中我們會探討有關 invertible matrix 的相關性質, 並介紹判斷一個方陣是否為 invertible 且找出其反矩陣的方法.

當初我們將聯立方程組用矩陣乘法的方式 $A\mathbf{x} = \mathbf{b}$ 來表示, 其中有一個很大的目的就是希望將解聯立方程式的問題此簡化成類似實數上解 $ax = b$ 的情形. 在實數情況, 當 $a \neq 0$ 時, $ax = b$ 的解就是很簡單的 $x = ba^{-1}$. 但在矩陣的情形, 我們沒有除法, 所以只能借助乘法來幫忙. 由於實數中 a^{-1} 有 $a^{-1}a = aa^{-1} = 1$ 的性質, 所以推廣這個概念至矩陣, 我們便希望找到矩陣 B 滿足 BA 以及 AB 為 identity. 不過當 $A \in M_{m\times n}(\mathbb{F})$ 且 $m \neq n$ 時, 由 Corollary 3.4.3 以及 Corollary 3.4.6, 我們知道不可能存在 B 同時滿足 BA 和 AB 皆為 identity matrix (因為 $\mathrm{rank}(A)$ 不可能同時為 m 和 n). 所以我們僅對 $m = n$, 即 A 為 square matrix 時有以下的定義.

Definition 3.5.1. 假設 $A \in M_{n\times n}(\mathbb{F})$ 為 n 階 square matrix, 若存在 $B \in M_{n\times n}(\mathbb{F})$ 使得 $AB = BA = I_n$, 則稱 A 為 *invertible*. 反之, 我們稱 A 為 *non-invertible*

再一次強調當 A 不是方陣時, 我們知 A 絕對不是 invertible. 因此當我們不知矩陣 A 的階數時, 絕對不能用存在 B 滿足 BA 為 identity 來說 A 為 invertible, 必須檢查另一邊

AB 亦為 identity 才可. 不過當 A 為 $n\times n$ square matrix, 確實檢查單邊就可以確定 A 為 invertible. 我們有以下的性質.

Theorem 3.5.2. 假設 $A\in M_{n\times n}(\mathbb{F})$ 為 n 階 *square matrix.* 則下列是等價的.

(1) A 為 *invertible matrix.*

(2) 存在 $B\in M_{n\times n}(\mathbb{F})$ 使得 $BA=I_n$.

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

(4) 存在 $C\in M_{n\times n}(\mathbb{F})$ 使得 $AC=I_n$.

Proof. 依 A 為 invertible 的定義, 我們知若 A 為 invertible, 則存在 $B\in M_{n\times n}(\mathbb{F})$ 使得 $BA=I_n$. 故 $(1)\Rightarrow(2)$.

由 A 為 $n\times n$ matrix 以及 Theorem 3.4.5 知存在 $B\in M_{n\times n}(\mathbb{F})$ 使得 $BA=I_n$ 若且唯若 A 化為 echelon form 後 pivot 的個數為 n. 故 $(2)\Leftrightarrow(3)$.

同理, 由 A 為 $n\times n$ matrix 以及 Theorem 3.4.2 知存在 $C\in M_{n\times n}(\mathbb{F})$ 使得 $AC=I_n$ 若且唯若 A 化為 echelon form 後 pivot 的個數為 n. 故 $(3)\Leftrightarrow(4)$.

最後, 由 A 化為 echelon form 後 pivot 的個數為 n 知存在 $B,C\in M_{n\times n}(\mathbb{F})$ 使得 $BA=I_n$ 以及 $AC=I_n$. 若能證得 $C=B$, 則由 $BA=AB=I_n$ 得證 A 為 invertible. 然而由 $BA=I_n$, 得 $(BA)C=I_nC=C$. 又由 $(BA)C=B(AC)=BI_n=B$, 得證 $B=C$. 故 $(3)\Rightarrow(1)$. 得證本定理. □

當一個 $n\times n$ matrix 的 rank 為 n 時, 有的書為了強調這個 rank 和階數相等的特殊情況, 特別稱之為 *nonsingular matrix.* 所以由 Theorem 3.5.2 我們知 invertible matrix 就是 nonsingular matrix. 反之, non-invertible matrix 就是 singular matrix. 不過為了讓大家不被這麼多名詞弄混. 以後我們一律採用 invertible 和 non-invertible 這樣的說法, 而不用 nonsingular 和 singular 這樣的說法.

由 Theorem 3.5.2 的證明我們知若 $A\in M_{n\times n}(\mathbb{F})$ 且存在 $B,C\in M_{n\times n}(\mathbb{F})$ 使得 $BA=I_n$ 且 $AC=I_n$, 則 $B=C$. 我們自然會問有沒有可能存在不同的 $B,B'\in M_{n\times n}(\mathbb{F})$ 皆滿足 $BA=I_n$ 以及 $B'A=I_n$. 下一個定理告訴我們這樣的方陣其實是唯一的.

Corollary 3.5.3. 假設 $A\in M_{n\times n}(\mathbb{F})$ 且 $B,B'\in M_{n\times n}(\mathbb{F})$ 滿足 $BA=I_n$ 以及 $B'A=I_n$. 則 $B=B'$.

Proof. 由 Theorem 3.5.2 我們知 A 為 invertible 且由其證明知 $BA=AB=I_n$ 以及 $B'A=AB'=I_n$. 故

$$B=I_nB=(B'A)B=B'(AB)=B'I_n=B'.$$

□

由 Corollary 3.5.3, 我們知道若 A 為 $n\times n$ invertible matrix, 則僅會存在唯一的一個 $n\times n$ matrix B 滿足 $BA=AB=I_n$. 它和 A 的關係如同在實數上非零實數的乘法的 inverse (乘法反元素), 所以我們給以下的定義.

Definition 3.5.4. 假設 $A \in M_{n\times n}(\mathbb{F})$ 為 invertible matrix. 我們稱唯一滿足 $BA = AB = I_n$ 的 $n\times n$ matrix B 為 A 的 *inverse* (反矩陣), 且用 A^{-1} 表示.

給定一 $n\times n$ invertible matrix A 由於其反矩陣是唯一的, 所以若要確定 $B = A^{-1}$ 我們僅要檢查是否 $BA = I_n$ 或 $AB = I_n$ 即可. 我們有以下之性質

Proposition 3.5.5. 假設 $A, B \in M_{n\times n}(\mathbb{F})$. 我們有以下之性質

(1) 若 A 為 *invertible,* 則 A^{-1} 亦為 *invertible* 且

$$(A^{-1})^{-1} = A.$$

(2) A 為 *invertible* 若且唯若 A^{t} 為 *invertible* 且此時

$$(A^{\mathrm{t}})^{-1} = (A^{-1})^{\mathrm{t}}.$$

(3) A, B 皆為 *invertible* 若且唯若 AB 為 *invertible.* 且此時

$$(AB)^{-1} = B^{-1}A^{-1}.$$

Proof. 由 Theorem 3.5.2, 我們要說一個 $n\times n$ matrix 為 invertible, 只要找到 $B \in M_{n\times n}(\mathbb{F})$ 使得 $BA = I_n$ 或 $AB = I_n$ 且此時由唯一性 (Corollary 3.5.3) 知 $B = A^{-1}$.

(1) 依定義 A^{-1} 亦為 $n\times n$ matrix 故 Theorem 3.5.2 適用. 利用 $A^{-1}A = I_n$, 得知 A^{-1} 亦為 invertible 且 $(A^{-1})^{-1} = A$.

(2) 依定義 A^{t} 亦為 $n\times n$ matrix 故 Theorem 3.5.2 適用. 由 $A^{-1}A = I_n$ 利用 Proposition 3.2.4 得

$$I_n = (A^{-1}A)^{\mathrm{t}} = A^{\mathrm{t}}(A^{-1})^{\mathrm{t}}$$

故知 A^{t} 為 invertible 且 $(A^{\mathrm{t}})^{-1} = (A^{-1})^{\mathrm{t}}$. 反之若 A^{t} 為 invertible, 由前知 $(A^{\mathrm{t}})^{\mathrm{t}}$ 為 invertible, 故由利用 Proposition 3.2.4 $(A^{\mathrm{t}})^{\mathrm{t}} = A$ 得證 A 為 invertible.

(3) 依定義 AB 為 $n\times n$ matrix 故 Theorem 3.5.2 適用. 現若 A, B 皆為 invertible, 則由

$$(AB)B^{-1}A^{-1} = A(BB^{-1})A^{-1} = AI_nA^{-1} = AA^{-1} = I_n,$$

得證 AB 為 invertible 且 $(AB)^{-1} = B^{-1}A^{-1}$. 反之, 若 AB 為 invertible, 且令 $C = (AB)^{-1}$. 此時由 $(AB)C = I_n$ 得 $A(BC) = I_n$, 故由假設 A 為 $n\times n$ matrix 以及 Theorem 3.5.2 得證 A 為 invertible. 同理, 由 $C(AB) = I_n$, 得 $(CA)B = I_n$, 得證 B 為 invertible. □

要注意 Proposition 3.5.5 (3) 中由 AB invertible 推得 A, B 皆為 invertible 是需要用到 A, B 皆為 $n\times n$ matrix. 否則當 $m \neq n$ 時, 在 Theorem 3.4.2 中我們知道有可能 $A \in M_{m\times n}(\mathbb{F}), C \in M_{n\times m}(\mathbb{F})$ 滿足 $AC = I_m$. 此時 I_m 為 invertible, 但 A, C 皆為 non-invertible. 同樣的, 當 A, B 為方陣時, 因為由 AB 為 invertible 可推得 A, B 皆為 invertible, 故知 BA 亦為 invertible. 也就是說當 A, B 為方陣時 AB 為 invertible 和 BA 為 invertible 是等價的. 但在 A, B 不為方陣時, 若 AB 為 invertible 會導致 BA 不為 invertible.

Question 3.10. 試舉例 A, B 不為 *invertible* 但 AB 為 *invertible.* 同時也驗證此時 BA 為 *non-invertible.*

接下來我們探討如何判別一個具體的 $n\times n$ matrix 是否為 invertible, 且若為 invertible 如何找出其 inverse. 這個問題可藉由將方陣利用 elementary row operations 化為 reduced echelon form 來處理. 事實上, 當 A 為 $n\times n$ matrix, 由 Theorem 3.5.2 我們知道 A 為 invertible 若且唯若 A 化為 echelon form 後其 pivot 的個數等於 n. 因此我們只要將 A 化為 echelon form 後計算其 pivot 的個數, 便可以知道 A 是否為 invertible. 若 A 為 invertible, 即 pivot 的個數為 n, 此時由於 A 的 reduced echelon form 為 $n\times n$ matrix, 故得 A 的 reduced echelon form 為 I_n. 也就是說我們可以用 elementary row operations 將 A 化為 I_n. 故由 Lemma 3.3.1 我們知存在 $E\in M_{n\times n}(\mathbb{F})$ 為一些 elementary matrix 的乘積使得 $EA=I_n$. 事實上若將 augmented matrix $[A|I_n]$ 利用 elementary row operations 化為 $[I_n|E]$, 則 $EA=I_n$, 故此時 E 就是 A^{-1}. 我們看以下的例子.

Example 3.5.6. 考慮矩陣

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

我們要決定是否 A 是否為 invertible. 若為 invertible, 要找出 A^{-1}.

我們直接考慮 augmented matrix $[A|I_4]$, 利用 elementary row operation 將 A 的部分轉換成 echelon form. 首先將 1-st row 分別乘上 -1, 3 加至 3-rd, 4-th row, 即

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

接著將 3-rd row 乘上 3/2 加至 4-th row 得

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

此時 augmented matrix 左半部為 echelon form, 其 pivot 的個數為 4, 故知 A 為 invertible. 我們繼續將左半部化為 reduced echelon form 便可得到 A^{-1}.

先將 4-th row 乘以 2, 然後將所得的 augmented matrix 的 4-th row 分別乘上 -3, -4, 1 加至 3-rd, 2-nd 和 1-st row, 即

$$\left[\begin{array}{cccc|cccc}1&0&1&-1&1&0&0&0\\0&-1&-3&4&0&1&0&0\\0&0&-2&3&-1&0&1&0\\0&0&0&1&3&0&3&2\end{array}\right]\rightsquigarrow\left[\begin{array}{cccc|cccc}1&0&1&0&4&0&3&2\\0&-1&-3&0&-12&1&-12&-8\\0&0&-2&0&-10&0&-8&-6\\0&0&0&1&3&0&3&2\end{array}\right].$$

接著將 3-rd row 乘以 $-1/2$, 然後將所得的 augmented matrix 的 3-rd row 分別乘上 3, -1 加至 2-nd 和 1-st row, 即

$$\left[\begin{array}{cccc|cccc} 1 & 0 & 1 & 0 & 4 & 0 & 3 & 2 \\ 0 & -1 & -3 & 0 & -12 & 1 & -12 & -8 \\ 0 & 0 & 1 & 0 & 5 & 0 & 4 & 3 \\ 0 & 0 & 0 & 1 & 3 & 0 & 3 & 2 \end{array}\right] \rightsquigarrow \left[\begin{array}{cccc|cccc} 1 & 0 & 0 & 0 & -1 & 0 & -1 & -1 \\ 0 & -1 & 0 & 0 & 3 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 & 5 & 0 & 4 & 3 \\ 0 & 0 & 0 & 1 & 3 & 0 & 3 & 2 \end{array}\right].$$

最後將 augmented matrix 的 2-nd row 乘上 -1. 此時所得 augmented matrix 左半部為 reduced echelon form (即 I_4), 故其右半部為 A^{-1}, 即

$$A^{-1} = \begin{bmatrix} -1 & 0 & -1 & -1 \\ -3 & -1 & 0 & -1 \\ 5 & 0 & 4 & 3 \\ 3 & 0 & 3 & 2 \end{bmatrix}.$$

由前面討論我們知當 $A \in M_{n\times n}(\mathbb{F})$ 為 invertible, 則存在 elementary matrices $E_1,\ldots,E_k$ 使得 $(E_k\cdots E_1)A = I_n$. 亦即 $A^{-1} = E_k\cdots E_1$, 由 Proposition 3.5.5 (3), 我們知 $E_1,\ldots,E_k$ 皆為 invertible, 且由 $(A^{-1})^{-1} = A$, 得 $A = E_1^{-1}\cdots E_k^{-1}$. 事實上這些 elementary matrix E_i 的 inverse 就是將 E_i 還原成 I_n 的 elementary row operation 所對應的 elementary matrix. 也就是說 E_i^{-1} 亦為 elementary matrix. 因此我們有以下的定理.

Proposition 3.5.7. *A 為 invertible matrix 若且唯若 A 為一些 elementary matrices 的乘積.*

Example 3.5.8. 考慮矩陣

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

在求 A 的 inverse 的過程中, 首先我們將 1-st row 和 2-nd row 交換. 令此 elementary row operation 所對應的 elementary matrix 為 E_1. 因用相同的 elementary row operation 可將 E_1 還原成 I_3, 故 $E_1 = E_1^{-1}$, 即

$$\left[\begin{array}{ccc|ccc} 0 & 2 & 0 & 1 & 0 & 0 \\ 1 & -1 & 0 & 0 & 1 & 0 \\ 0 & 2 & 1 & 0 & 0 & 1 \end{array}\right] \rightsquigarrow \left[\begin{array}{ccc|ccc} 1 & -1 & 0 & 0 & 1 & 0 \\ 0 & 2 & 0 & 1 & 0 & 0 \\ 0 & 2 & 1 & 0 & 0 & 1 \end{array}\right], E_1 = E_1^{-1} = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

接著將 augmented matrix 的 2-nd row 乘上 -1 加至 3-rd row. 令此 elementary row operation 所對應的 elementary matrix 為 E_2. 因將 2-nd row 乘上 1 加至 3-rd row 的 elementary row operation 可將 E_2 還原成 I_3, 故所得的 augmented matrix 及 E_2, E_2^{-1} 分別為

$$\left[\begin{array}{ccc|ccc} 1 & -1 & 0 & 0 & 1 & 0 \\ 0 & 2 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & -1 & 0 & 1 \end{array}\right], E_2 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -1 & 1 \end{bmatrix}, E_2^{-1} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix}.$$

然後將 augmented matrix 的 2-nd row 乘上 $1/2$. 令此 elementary row operation 所對應的 elementary matrix 為 E_3. 因將 2-nd row 乘上 2 的 elementary row operation 可將 E_3 還原

成 I_3, 故所得的 augmented matrix 及 E_3, E_3^{-1} 分別為

$$\left[\begin{array}{ccc|ccc} 1 & -1 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & \frac{1}{2} & 0 & 0 \\ 0 & 0 & 1 & -1 & 0 & 1 \end{array}\right], E_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{2} & 0 \\ 0 & 0 & 1 \end{bmatrix}, E_3^{-1} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

最後將 augmented matrix 的 2-nd row 乘上 1 加至 1-st row. 令此 elementary row operation 所對應的 elementary matrix 為 E_4. 因將 2-nd row 乘上 -1 加至 3-rd row 的 elementary row operation 可將 E_4 還原成 I_3, 故所得的 augmented matrix 及 E_4, E_4^{-1} 分別為

$$\left[\begin{array}{ccc|ccc} 1 & 0 & 0 & \frac{1}{2} & 1 & 0 \\ 0 & 1 & 0 & \frac{1}{2} & 0 & 0 \\ 0 & 0 & 1 & -1 & 0 & 1 \end{array}\right], E_4 = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, E_4^{-1} = \begin{bmatrix} 1 & -1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

我們檢查可得

$$A^{-1} = E_4E_3E_2E_1 = \begin{bmatrix} \frac{1}{2} & 1 & 0 \\ \frac{1}{2} & 0 & 0 \\ -1 & 0 & 1 \end{bmatrix}, A = E_1^{-1}E_2^{-1}E_3^{-1}E_4^{-1} = \begin{bmatrix} 0 & 2 & 0 \\ 1 & -1 & 0 \\ 0 & 2 & 1 \end{bmatrix}.$$

最後讓我們回到解聯立方程組的問題. 怎樣的 $A \in M_{m\times n}(\mathbb{F})$ 會使得對任意 $\mathbf{b} \in \mathbb{F}^m$, 聯立方程組 $A\mathbf{x} = \mathbf{b}$ 皆有解且解唯一呢? 由 Theorem 3.4.2 和 Theorem 3.4.5 知此時 $\operatorname{rank}(A) = m$ 且 $\operatorname{rank}(A) = n$, 即 $m = n$. 也就是說 A 必須是 $n \times n$ 且 $\operatorname{rank}(A) = n$. 因此由 Theorem 3.5.2 知 A 為 $n \times n$ invertible matrix. 事實上我們有以下的等價關係. 由於它們直接套用 Theorem 3.4.2 和 Theorem 3.4.5 就可推得, 我們就不再證明了.

Theorem 3.5.9. *假設 $A \in M_{n\times n}(\mathbb{F})$, 令 $\mathbf{a}_1, \dots, \mathbf{a}_n \in \mathbb{F}^n$ 為 A 的 column vectors. 則下列是等價的.*

(1) *A 為 invertible matrix.*

(2) *$\operatorname{Span}(\mathbf{a}_1, \dots, \mathbf{a}_n) = \mathbb{F}^n$.*

(3) *對於任意 $\mathbf{b} \in \mathbb{F}^n$, 聯立方程組 $A\mathbf{x} = \mathbf{b}$ 皆有解.*

(4) *聯立方程組 $A\mathbf{x} = \mathbf{0}$ 沒有 nontrivial solution.*

(5) *對於任意 $\mathbf{b} \in \mathbb{F}^n$, 聯立方程組 $A\mathbf{x} = \mathbf{b}$ 皆有解且解唯一.*

設 $A \in M_{n\times n}(\mathbb{F})$ 為 invertible matrix, 則對任意 $\mathbf{b} \in \mathbb{F}^n$, 我們可以利用 A 的反矩陣 A^{-1} 得到聯立方程組 $A\mathbf{x} = \mathbf{b}$ 的唯一解. 事實上若令 $\mathbf{x} = A^{-1}\mathbf{b}$, 此時 $A\mathbf{x} = A(A^{-1}\mathbf{b}) = (AA^{-1})\mathbf{b} = \mathbf{b}$. 又由 Theorem 3.5.9 知此時 $A\mathbf{x} = \mathbf{b}$ 的解唯一. 故 $\mathbf{x} = A^{-1}\mathbf{b}$ 為聯立方程組 $A\mathbf{x} = \mathbf{b}$ 唯一的一組解.

Example 3.5.10. 考慮聯立方程組

$$\begin{array}{rrrrcl} x_1 & & +x_3 & -x_4 & = & b_1 \\ & -x_2 & -3x_3 & +4x_4 & = & b_2 \\ x_1 & & -x_3 & +2x_4 & = & b_3 \\ -3x_1 & & & -x_4 & = & b_4 \end{array}$$

其中 b_1,b_2,b_3,b_4 為任意實數. 由於此時聯立方程組為 $A\mathbf{x}=\mathbf{b}$, 其中 A 為 Example 3.5.6 中的 4×4 matrix A 且 $\mathbf{b}=[b_1\ b_2\ b_3\ b_4]^{\mathrm{t}}$. 因 A 為 invertible, 故由 Theorem 3.5.9 知, 對任意實數 b_1,b_2,b_3,b_4, 聯立方程組 $A\mathbf{x}=\mathbf{b}$ 必有解且其解唯一. 事實上此唯一解為

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = A^{-1}\mathbf{b} = \begin{bmatrix} -1 & 0 & -1 & -1 \\ -3 & -1 & 0 & -1 \\ 5 & 0 & 4 & 3 \\ 3 & 0 & 3 & 2 \end{bmatrix} \begin{bmatrix} b_1 \\ b_2 \\ b_3 \\ b_4 \end{bmatrix} = \begin{bmatrix} -b_1-b_3-b_4 \\ -3b_1-b_2-b_4 \\ 5b_1+4b_3+3b_4 \\ 3b_1+3b_3+2b_4 \end{bmatrix}.$$

3.6. Column Space and Nullspace

我們將介紹一個矩陣的 column space, row space 以及 nullspace 並探討如何找到它們的 basis. 我們會發現 column space 和 row space 的 dimension 皆相同且等於矩陣的 rank. 最後我們探討如何得到一般 subspace 的 basis.

給定一個矩陣, 它的 column space 和 nullspace 和以該矩陣為係數矩陣所形成的聯立方程組是否有解以及解是否唯一息息相關. 由於 column space 和 nullspace 的重要性, 我們將之正式定義如下:

Definition 3.6.1. 假設 $A=\begin{bmatrix} | & | & & | \\ \mathbf{a}_1 & \mathbf{a}_2 & \cdots & \mathbf{a}_n \\ | & | & & | \end{bmatrix}$ 為以 $\mathbb{F}^m$ 中的向量 $\mathbf{a}_1,\ldots,\mathbf{a}_n$ 為 column vectors 的 $m\times n$ matrix.

(1) 我們稱 $\operatorname{Span}(\mathbf{a}_1,\ldots,\mathbf{a}_n)$ 為 A 的 *column space*, 且用 $\operatorname{Col}(A)$ 來表示 A 的 column space.

(2) 我們稱 homogeneous linear system $A\mathbf{x}=\mathbf{0}$ 所有解所成的集合為 A 的 *nullspace* 且用 $N(A)$ 表示 A 的 nullspace. 即 $N(A)=\{\mathbf{u}\in\mathbb{F}^n \mid A\mathbf{u}=\mathbf{0}\}$.

要注意當 $A\in M_{m\times n}(\mathbb{F})$, 則 A 的 column space $\operatorname{Col}(A)$ 會是 $\mathbb{F}^m$ 的 subspace, 而 A 的 nullspace $N(A)$ 會是 $\mathbb{F}^n$ 的 subspace. 利用 Lemma 3.4.1 以及 Theorem 3.4.5 我們馬上有以下的結果.

Proposition 3.6.2. *假設 A 為 $m\times n$ matrix 且 $\mathbf{b}\in\mathbb{F}^m$, 考慮聯立方程組 $A\mathbf{x}=\mathbf{b}$.*

(1) $A\mathbf{x}=\mathbf{b}$ 有解若且唯若 $\mathbf{b}\in\operatorname{Col}(A)$.

(2) 假設 $A\mathbf{x}=\mathbf{b}$ 有解則其解唯一若且唯若 $N(A)=\{\mathbf{0}\}$.

接下來我們就是要找到一個矩陣的 column space 以及 nullspace 這兩個重要的 subspaces 的 basis. 一般來說要找到 $\mathbb{F}^m$ 的 subspace V 的一組 basis, 我們會先找 V 的一組 spanning vectors. 然後在其中再挑出仍保持為 spanning vectors 且為 linearly independent 的一組向量. 當只有兩個向量時, 我們可以馬上由它們是否為平行來判斷是否為 linearly independent. 不過通常有三個以上的向量時, 並不容易直接看出哪些向量會 linearly independent, 除非如以下的例子.

Example 3.6.3. 考慮 $\mathbb{R}^3$ 中

$$\mathbf{v}_1=\begin{bmatrix}0\\3\\0\\1\end{bmatrix},\mathbf{v}_2=\begin{bmatrix}2\\-5\\0\\0\end{bmatrix},\mathbf{v}_3=\begin{bmatrix}0\\7\\-1\\0\end{bmatrix}.$$

要說明 $\mathbf{v}_1,\mathbf{v}_2,\mathbf{v}_3$ 為 linearly independent, 我們必須說明只有當 $c_1=c_2=c_3=0$ 時, 才會使得 $c_1\mathbf{v}_1+c_2\mathbf{v}_2+c_3\mathbf{v}_3=\mathbf{0}$. 然而

$$c_1\mathbf{v}_1+c_2\mathbf{v}_2+c_3\mathbf{v}_3=c_1\begin{bmatrix}0\\3\\0\\1\end{bmatrix}+c_2\begin{bmatrix}2\\-5\\0\\0\end{bmatrix}+c_3\begin{bmatrix}0\\7\\-1\\0\end{bmatrix}=\begin{bmatrix}2c_2\\3c_1-5c_2+7c_3\\-c_3\\c_1\end{bmatrix}.$$

所以要使得 $c_1\mathbf{v}_1+c_2\mathbf{v}_2+c_3\mathbf{v}_3=\mathbf{0}$, 就必須讓 $c_1\mathbf{v}_1+c_2\mathbf{v}_2+c_3\mathbf{v}_3$ 的 1-st entry $2c_2$, 3-rd entry $-c_3$ 以及 4-th entry c_1 皆為 0, 即 $c_1=c_2=c_3=0$. 得證只有當 $c_1=c_2=c_3=0$ 時, 才會使得 $c_1\mathbf{v}_1+c_2\mathbf{v}_2+c_3\mathbf{v}_3=\mathbf{0}$, 故知 $\mathbf{v}_1,\mathbf{v}_2,\mathbf{v}_3$ 為 linearly independent.

從 Example 3.6.3 我們可以看出來, 當 $\mathbf{v}_1,\dots,\mathbf{v}_n$ 中每一個向量 $\mathbf{v}_i$ 都可以找到一個 entry 不為 0, 而其他向量在該 entry 皆為 0, 則 $\mathbf{v}_1,\dots,\mathbf{v}_n$ 為 linearly independent. (例如 Example 3.6.3 中 $\mathbf{v}_1$ 的 4-th entry 為 1, 而 $\mathbf{v}_2,\mathbf{v}_3$ 的 4-th entry 為 0; $\mathbf{v}_2$ 的 1-st entry 為 2, 而 $\mathbf{v}_1,\mathbf{v}_3$ 的 1-st entry 為 0; $\mathbf{v}_3$ 的 3-rd entry 為 -1, 而 $\mathbf{v}_1,\mathbf{v}_2$ 的 3-th entry 為 0, 就符合這個條件). 此時假設每個 $\mathbf{v}_i$ 的那個非 0 的特殊 entry 為 a_i, 由於 $c_1\mathbf{v}_1+\cdots+c_n\mathbf{v}_n$ 在該位置的 entry 為 c_ia_i, 所以若 $c_1\mathbf{v}_1+\cdots+c_n\mathbf{v}_n=\mathbf{0}$, 則必 $c_ia_i=0$, 得每一個 c_i 皆為 0. 因此 $\mathbf{v}_1,\dots,\mathbf{v}_n$ 為 linearly independent.

當 A 為 $m\times n$ matrix, A 的 nullspace $N(A)$ 就是 homogeneous linear system $A\mathbf{x}=\mathbf{0}$ 的所有解所成的集合. 由於我們已經知道如何找到 $A\mathbf{x}=\mathbf{0}$ 的解, 所以我們就從如何找 nullspace 的 basis 開始.

回顧我們找 $A\mathbf{x}=\mathbf{0}$ 的解集合的方法為, 利用 elementary row operations 將 A 化為 echelon form (或 reduced echelon form) A'. 此時 $A'\mathbf{x}=\mathbf{0}$ 的解集合就是 $A\mathbf{x}=\mathbf{0}$ 的解集合, 也就是說 A 和 A' 有相同的 nullspace. 接著我們找出 free variable, 再將每個 free variable 代入任意的實數, 從下往上推得出一組解. 注意在這個過程中, pivot variable 的值會由 free variables 的值所決定, 所以只要定出 free variable 的值, 就可以得到一組解. 現假設 free variables 為 $x_{i_1},\dots,x_{i_k}$. 對每一個 $j=1,\dots,k$, 我們考慮 $x_{i_j}=1$, 其他 free variable 為 0 的情形, 令這樣推得出來的解為 $\mathbf{v}_j$. 由於 $\mathbf{v}_j$ 的 i_j-th entry 為 1, 而其他 $\mathbf{v}_{i_{j'}}$ 的 i_j-th entry 為 0, 由前討論知 $\mathbf{v}_1,\dots,\mathbf{v}_k$ 為 linearly independent. 而對於任意 $r_1,\dots,r_k\in\mathbb{F}$, $r_1\mathbf{v}_1+\cdots+r_k\mathbf{v}_k$ 就等同於是將每個 free variables $x_{i_1},\dots,x_{i_k}$ 分別代 $x_{i_1}=r_1,\dots,x_{i_k}=r_k$ 所得的解. 換言之每個解都可以寫成 $r_1\mathbf{v}_1+\cdots+r_k\mathbf{v}_k$ 的形式, 也就是說 $\mathbf{v}_1,\dots,\mathbf{v}_k$ 是 A 的 nullspace 的一組 spanning vectors. 我們證明了 $\mathbf{v}_1,\dots,\mathbf{v}_k$ 就是 A 的 nullspace 的一組 basis, 也因此得知 A 的 nullspace 的 dimension 為 free variables 的個數, 亦即 A 的 column 的個數減去 pivot 的個數, 因此有以下之結果.

Proposition 3.6.4. *假設 A 為 $m\times n$ matrix. 若利用 row operations 將 A 化為 echelon form A' 後, A' 的 pivot 個數為 r, 則 A 的 nullspace 的 dimension 為 $n-r$. 假設 $A'\mathbf{x}=\mathbf{0}$ 的 free variables 為 $x_{i_1},\dots,x_{i_k}$. 對每一個 $j=1,\dots,k$, 我們取 $x_{i_j}=1$, 其他 free variable 為 0, 令這樣推得出來的解為 $\mathbf{v}_j$. 則 $\mathbf{v}_1,\dots,\mathbf{v}_k$ 為 A 的 nullspace 的一組 basis.*

由於一個矩陣的 nullspace 不會因為其化為 echelon form 的不同而改變, 而且 nullspace 的 dimension 是固定的, 所以 Proposition 3.6.4 也說明了“不管一個矩陣利用 elementary row operations 所化得的 echelon form 為何, 其 pivot 的個數必相同”, 也就是這個矩陣的 rank.

Example 3.6.5. 考慮 A 的 nullspace, 其中

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

將 A 的 2-nd row 分別乘上 -2, -1, -1 加至 1-st, 3-rd 和 4-th row, 然後再將 1-st, 2-nd rows 交換得

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

接著將 2-nd row 分別乘上 $-1,-2$ 加至 3-rd 和 4-th row 得

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

最後將 3-rd row 乘上 -1 加至 4-th row, 得 echelon form

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

我們就是要找到 homogeneous linear system

$$\begin{array}{cccccccc} x_1 & & & +x_4 & & & = & 0\\ & x_2 & +x_3 & -2x_4 & & & = & 0\\ & & & +x_4 & +x_5 & +2x_6 & = & 0\end{array}$$

所有的解. 由 echelon form 看出 x_1,x_2,x_4 為 pivot variable, x_3,x_5,x_6 為 free variable. 現令 $x_6=1,x_5=0,x_3=0$, 解出 $x_4=-2$, $x_2=-4$, $x_1=2$, 而令 $x_6=0,x_5=1,x_3=0$ 解出 $x_4=-1$,

$x_2=-2$, $x_1=1$, 最後令 $x_6=0, x_5=0, x_3=1$ 解出 $x_4=0$, $x_2=-1$, $x_1=0$. 故得

$$\mathbf{v}_1=\begin{bmatrix}2\\-4\\0\\-2\\0\\1\end{bmatrix}, \mathbf{v}_2=\begin{bmatrix}1\\-2\\0\\-1\\1\\0\end{bmatrix}, \mathbf{v}_3=\begin{bmatrix}0\\-1\\1\\0\\0\\0\end{bmatrix}$$

為 A 的 nullspace 的一組 basis. 事實上, 若令 x_6,x_5,x_3 分別為任意的實數 r,s,t, 則可得 $x_4=-2r-s$, $x_2=-4r-2s-t$, $x_1=2r+s$. 也就是說 A 的 nullspace 中的向量都可以寫成

$$\begin{bmatrix}2r+s\\-4r-2s-t\\t\\-2r-s\\s\\r\end{bmatrix}=r\begin{bmatrix}2\\-4\\0\\-2\\0\\1\end{bmatrix}+s\begin{bmatrix}1\\-2\\0\\-1\\1\\0\end{bmatrix}+t\begin{bmatrix}0\\-1\\1\\0\\0\\0\end{bmatrix}=r\mathbf{v}_1+s\mathbf{v}_2+t\mathbf{v}_3.$$

故知 $\mathbf{v}_1,\mathbf{v}_2,\mathbf{v}_3$ 為 A 的 nullspace 的 spanning vectors, 又很容易看出 $\mathbf{v}_1,\mathbf{v}_2,\mathbf{v}_3$ 為 linearly independent, 得證 $\mathbf{v}_1,\mathbf{v}_2,\mathbf{v}_3$ 為 $N(A)$ 的一組 basis.

Question 3.11. *試將 Example 3.6.5 中的 A 化為 reduced echelon form. 是否更容易看出 $N(A)$ 的一組 basis 呢?*

接下來我們來看如何找 matrix A 的 column space $\mathrm{Col}(A)$ 的 basis. 首先一個直接的想法就是 A 的 column space, 就是使得聯立方程組 $A\mathbf{x}=\mathbf{v}$ 有解的 $\mathbf{v}$ 所成的集合. 所以我們只要找出這些 $\mathbf{v}$, 就可以得到 A 的 column space. 我們看以下的例子.

Example 3.6.6. 考慮 Example 3.6.5 中的 4×6 matrix A. 我們要找出 A 的 column vectors 的一組 basis. 假設 $\mathbf{b}$ 為 A 的 column space 裡的一個向量, 我們知道此時 $A\mathbf{x}=\mathbf{b}$ 有解, 因此令

$$\mathbf{b}=\begin{bmatrix}b_1\\b_2\\b_3\\b_4\end{bmatrix},$$

我們要找到 b_1,b_2,b_3,b_4 的條件使得以下聯立方程組有解.

$$\begin{array}{ccccccccc}2x_1 & +x_2 & +x_3 & & & & = & b_1\\ x_1 & & & +x_4 & & & = & b_2\\ x_1 & +x_2 & +x_3 & & +x_5 & +2x_6 & = & b_3\\ x_1 & +2x_2 & +2x_3 & -2x_4 & +x_5 & +2x_6 & = & b_4\end{array}$$

考慮 augmented matrix $[A\,|\,\mathbf{b}]$, 利用 Example 3.6.5 相同的 elementary row operations 我們得

$$\left[\begin{array}{cccccc|c}2&1&1&0&0&0&b_1\\1&0&0&1&0&0&b_2\\1&1&1&0&1&2&b_3\\1&2&2&-2&1&2&b_4\end{array}\right]\rightsquigarrow\left[\begin{array}{cccccc|c}1&0&0&1&0&0&b_2\\0&1&1&-2&0&0&b_1-2b_2\\0&1&1&-1&1&2&b_3-b_2\\0&2&2&-3&1&2&b_4-b_2\end{array}\right]\rightsquigarrow$$

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

由解聯立方程組的方法 (即 1.2 節 (2)(a)(b) 的情形) 知, 聯立方程組 $A\mathbf{x}=\mathbf{b}$ 有解若且唯若 $b_4-b_3+2b_2-b_1=0$. 換言之, 由所有 $b_1-2b_2+b_3-b_4=0$ 的解, 所得的 $\mathbf{b}$ 所成的集合便是 A 的 column space. 所以我們回到求矩陣 $B=\begin{bmatrix} 1 & -2 & 1 & -1 \end{bmatrix}$ 的 nullspace. 由於 x_1 為 pivot variable, x_2,x_3,x_4 為 free variable. 利用前面求 nullspace 的 basis 的方法, 令 $x_4=1,x_3=0,x_2=0$ 解出 $x_1=1$, 而令 $x_4=0,x_3=1,x_2=0$ 解出 $x_1=-1$, 最後令 $x_4=0,x_3=0,x_2=1$ 解出 $x_1=2$. 故得

$$\mathbf{v}_1=\begin{bmatrix}1\\0\\0\\1\end{bmatrix}, \mathbf{v}_2=\begin{bmatrix}-1\\0\\1\\0\end{bmatrix}, \mathbf{v}_3=\begin{bmatrix}2\\1\\0\\0\end{bmatrix}$$

為 B 的 nullspace 的一組 basis, 也就是 A 的 column space 的一組 basis..

注意用這個方法, 若 $m\times n$ matrix A 化成 echelon form 後沒有一個 row 全為 0, 就表示所有的 $\mathbf{b}\in\mathbb{F}^m$ 皆會使聯立方程組有解, 故此時 A 的 column space 為 $\mathbb{F}^m$.

Example 3.6.6 找 column space 所用的方法缺點就是還要再求另一個矩陣的 nullspace 才能找到 column space 的 basis. 接下來我們介紹一個更簡捷的方法.

首先注意當我們利用 elementary row operations 將 A 化為 echelon form A' 後, homogeneous linear system $A\mathbf{x}=\mathbf{0}$ 和 $A'\mathbf{x}=\mathbf{0}$ 有相同的解集合. 現假設 $\mathbf{a}_1,\ldots,\mathbf{a}_n$ 為 A 的 column vectors, 而 $\mathbf{a}'_1,\ldots,\mathbf{a}'_n$ 為 A' 的 column vectors. 若 $x_1=c_1,\ldots,x_n=c_n$ 為 $A'\mathbf{x}=\mathbf{0}$ 的一組解, 表示 $c_1\mathbf{a}'_1+\cdots+c_n\mathbf{a}'_n=\mathbf{0}$, 此時由於 $x_1=c_1,\ldots,x_n=c_n$ 亦為 $A\mathbf{x}=\mathbf{0}$ 的一組解故我們亦有 $c_1\mathbf{a}_1+\cdots+c_n\mathbf{a}_n=\mathbf{0}$. 同理若 $c_1,\cdots,c_n\in\mathbb{F}$ 使得 $c_1\mathbf{a}_1+\cdots+c_n\mathbf{a}_n=\mathbf{0}$, 我們亦會有 $c_1\mathbf{a}'_1+\cdots+c_n\mathbf{a}'_n=\mathbf{0}$. 這告訴我們存在不全為 0 的 c_i 使得 $c_1\mathbf{a}_1+\cdots+c_n\mathbf{a}_n=\mathbf{0}$ 若且唯若存在不全為 0 的 c_i 使得 $c_1\mathbf{a}'_1+\cdots+c_n\mathbf{a}'_n=\mathbf{0}$. 換言之, $\mathbf{a}_1,\ldots,\mathbf{a}_n$ 為 linearly dependent 若且唯若 $\mathbf{a}'_1,\ldots,\mathbf{a}'_n$ 為 linearly dependent. 這也等價於 $\mathbf{a}_1,\ldots,\mathbf{a}_n$ 為 linearly independent 若且唯若 $\mathbf{a}'_1,\ldots,\mathbf{a}'_n$ 為 linearly independent. 簡單來說當我們利用 elementary row operations 將一個矩陣變換到另一個矩陣, 兩個矩陣 column vectors 之間的線性關係是會被保持的. 我們看以下的例子.

Example 3.6.7. 考慮 Example 3.6.5 中的 4×6 matrix A, 且利用 elementary row operation 將之化為 reduced echelon form A'. 也就是將 Example 3.6.5 中的 echelon form 的 3-rd row 乘上 2 加到 echelon form 的 2-nd row, 再將 echelon form 的 3-rd row 乘上 -1 加到 echelon form 的 1-st row 得

$$A=\begin{bmatrix} 2 & 1 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 & 0 \\ 1 & 1 & 1 & 0 & 1 & 2 \\ 1 & 2 & 2 & -2 & 1 & 2 \end{bmatrix} \rightsquigarrow \begin{bmatrix} 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 1 & -2 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 2 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \rightsquigarrow A'=\begin{bmatrix} 1 & 0 & 0 & 0 & -1 & -2 \\ 0 & 1 & 1 & 0 & 2 & 4 \\ 0 & 0 & 0 & 1 & 1 & 2 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

我們很容易看出 A' 的 3 個 pivot 所在的 column vectors $\mathbf{a}'_1,\mathbf{a}'_2,\mathbf{a}'_4$ 為 linearly independent. 事實上 $\mathbf{a}'_1,\mathbf{a}'_2,\mathbf{a}'_4$ 每一個都符合有一個非 0 entry (即 pivot 之 entry) 而其他向量在該 entry 為 0. 我們考慮相對應到 A 的 column vectors $\mathbf{a}_1,\mathbf{a}_2,\mathbf{a}_4$. 它們也會是 linearly independent. 這是因為若我們考慮新的 4×3 matrix $[\mathbf{a}_1\quad \mathbf{a}_2\quad \mathbf{a}_4]$ 經由將 A 換成 A' 一樣步驟的 elementary row operation 我們會得到 $[\mathbf{a}'_1\quad \mathbf{a}'_2\quad \mathbf{a}'_4]$. 所以依前面的討論知, 因為 $\mathbf{a}'_1,\mathbf{a}'_2,\mathbf{a}'_4$ 為 linearly independent, 所以 $\mathbf{a}_1,\mathbf{a}_2,\mathbf{a}_4$ 也會是 linearly independent. 另一方面, 在 A' 中我們很容易看出 $\mathbf{a}'_3=\mathbf{a}'_2$, $\mathbf{a}'_5=-\mathbf{a}'_1+2\mathbf{a}'_2+\mathbf{a}'_4$ 以及 $\mathbf{a}'_6=-2\mathbf{a}'_1+4\mathbf{a}'_2+2\mathbf{a}'_4$. 所以和剛才同樣理由, 依 elementary row operations 保持線性關係的性質, 我們有 $\mathbf{a}_3=\mathbf{a}_2$, $\mathbf{a}_5=-\mathbf{a}_1+2\mathbf{a}_2+\mathbf{a}_4$ 以及 $\mathbf{a}_6=-2\mathbf{a}_1+4\mathbf{a}_2+2\mathbf{a}_4$. 事實上直接檢查得

$$\mathbf{a}_2=\mathbf{a}_3=\begin{bmatrix}1\\0\\1\\2\end{bmatrix},\quad -\mathbf{a}_1+2\mathbf{a}_2+\mathbf{a}_4=-\begin{bmatrix}2\\1\\1\\1\end{bmatrix}+2\begin{bmatrix}1\\0\\1\\2\end{bmatrix}+\begin{bmatrix}0\\1\\0\\-2\end{bmatrix}=\begin{bmatrix}0\\0\\1\\1\end{bmatrix}=\mathbf{a}_5,$$

$$-2\mathbf{a}_1+4\mathbf{a}_2+2\mathbf{a}_4=-2\begin{bmatrix}2\\1\\1\\1\end{bmatrix}+4\begin{bmatrix}1\\0\\1\\2\end{bmatrix}+2\begin{bmatrix}0\\1\\0\\-2\end{bmatrix}=\begin{bmatrix}0\\0\\2\\2\end{bmatrix}=\mathbf{a}_6.$$

換言之, $\mathbf{a}_3,\mathbf{a}_5,\mathbf{a}_6\in\mathrm{Span}(\mathbf{a}_1,\mathbf{a}_2,\mathbf{a}_4)$. 故知 A 的 column space 為

$$\mathrm{Span}(\mathbf{a}_1,\mathbf{a}_2,\mathbf{a}_3,\mathbf{a}_4,\mathbf{a}_5,\mathbf{a}_6)=\mathrm{Span}(\mathbf{a}_1,\mathbf{a}_2,\mathbf{a}_4).$$

再加上 $\mathbf{a}_1,\mathbf{a}_2,\mathbf{a}_4$ 為 linearly independent, 得證 $\mathbf{a}_1,\mathbf{a}_2,\mathbf{a}_4$ 是 A 的 column space 的一組 basis.

注意在 Example 3.6.7 中我們為了方便說明將 A 化為 reduced echelon form. 事實上既然我們知道 column space 的 basis 是由對應到 pivot 所在位置 A 的 column vectors 所組成, 所以化成 echelon form 知道 pivot 在那些 column 就可以找到 basis 了. 因此除非我們想要將 A 的其他 column vectors 用這組 basis 來表示, 一般是不需要進一步化成 reduced echelon form. 另外我們要強調的是 column space 的 basis 是要回到 A 的 column vectors 所組成, 而不是由 A 的 echelon form (或 reduced echelon form) A' 的 pivot 所在的 column vectors 所組成. 這是因為一般我們在做 elementary row operations 已將 column vectors 各個 entry 做了調動, 所以 echelon form A' 的 column space 已不再是原來 A 的 column space 了.

我們將這個求 column space 的 basis 的方法做一個總結. 首先將 $m\times n$ matrix A 利用 elementary row operation 化為 echelon form A'. 假設 A' 的 pivot variables 為 $x_{i_1},\dots,x_{i_r}$, 則由於 A' 的 pivot 所在的 column vectors $\mathbf{a}'_{i_1},\dots,\mathbf{a}'_{i_r}$ 為 linearly independent 且 elementary row operations 會保持各 column vectors 之間的線性關係, 我們知對應到 A 的 column vectors $\mathbf{a}_{i_1},\dots,\mathbf{a}_{i_r}$ 亦為 linearly independent. 同理, 由於 A' 的其他 column vectors $\mathbf{a}'_j$ 皆符合 $\mathbf{a}'_j\in\mathrm{Span}(\mathbf{a}'_{i_1},\dots,\mathbf{a}'_{i_r})$, 我們得 A 的其他 column vectors $\mathbf{a}_j$ 也符合 $\mathbf{a}_j\in\mathrm{Span}(\mathbf{a}_{i_1},\dots,\mathbf{a}_{i_r})$. 因此得 $\mathrm{Span}(\mathbf{a}_1,\dots,\mathbf{a}_n)=\mathrm{Span}(\mathbf{a}_{i_1},\dots,\mathbf{a}_{i_r})$. 我們證得了 $\mathbf{a}_{i_1},\dots,\mathbf{a}_{i_r}$ 為 A 的 column space 的 spanning vectors 且為 linearly independent, 故 $\mathbf{a}_{i_1},\dots,\mathbf{a}_{i_r}$ 為 A 的 column space 的一組 basis. 我們有以下的定理.

Proposition 3.6.8. 假設 $A \in M_{m\times n}(\mathbb{F})$ 且 $\mathbf{a}_1,\ldots,\mathbf{a}_n \in \mathbb{F}^m$ 為 A 的 *column vectors.* 若利用 *elementary row operations* 將 A 化為 *echelon form* A' 後, A' 的 *pivot* 個數為 r, 則 A 的 *column space* 的 *dimension* 為 r. 假設 A' 的 *pivot variables* 為 $x_{i_1},\ldots,x_{i_r}$, 則 $\mathbf{a}_{i_1},\ldots,\mathbf{a}_{i_r}$ 為 A 的 *column space* 的一組 *basis.*

相對於矩陣的 column space, 我們也可考慮矩陣的 row space. 我們有以下的定義.

Definition 3.6.9. 假設 $A = \begin{bmatrix} - & {}_1\mathbf{a} & - \\ & \vdots & \\ - & {}_m\mathbf{a} & - \end{bmatrix}$ 為以 $\mathbb{F}^n$ 中的向量 ${}_1\mathbf{a}_1,\ldots,{}_m\mathbf{a}$ 為 row vectors 的 $m\times n$ matrix. 則 A 的 row space 為 $\mathrm{Span}({}_1\mathbf{a},\ldots,{}_m\mathbf{a})$, 且用 $\mathrm{Row}(A)$ 來表示.

如何求 A 的 row space 的 basis 呢? 我們可以考慮 A 的 transpose A^{t}. 因為 A^{t} 的 column vectors 就是 A 的 row vectors, 求出 A^{t} 的 column space 的 basis 就等同於求 A 的 row space 的 basis. 所以我們可以用求 column space 的 basis 方法求出 A^{t} 的 column space 的 basis, 便得到 A 的 row space 的 basis. 不過這個方法有個缺點, 因為我們是換了一個矩陣 A^{t} 做 row operations, 因此就無法得到和原來 A 的 column space 之間的關係了. 以下介紹的方法, 便是直接對 A 做 elementary row operations 來求得 A 的 row space 的 basis, 所以我們可以得到 A 的 row space 和 column space 之間的關係.

這個方法的主要概念是 A 經過 elementary row operations 變換成 A' 後, A 和 A' 的 row space 是相同的. 這是因為若 ${}_1\mathbf{a},\ldots,{}_m\mathbf{a}$ 為 A 的 row vectors, ${}_1\mathbf{a}',\ldots,{}_m\mathbf{a}'$ 為 A' 的 row vectors, 則每個 ${}_i\mathbf{a}'$ 其實是 ${}_1\mathbf{a},\ldots,{}_m\mathbf{a}$ 中的向量互相交換, 或是乘上某個非 0 實數, 或是乘上某個實數後加到另一個向量. 也就是說每個 ${}_i\mathbf{a}'$ 其實是 ${}_1\mathbf{a},\ldots,{}_m\mathbf{a}$ 的線性組合, 所以對所有 $i=1,\ldots,m$ 皆有 ${}_i\mathbf{a}' \in \mathrm{Span}({}_1\mathbf{a},\ldots,{}_m\mathbf{a})$. 因此由 $\mathrm{Span}({}_1\mathbf{a},\ldots,{}_m\mathbf{a})$ 是 $\mathbb{F}^n$ 的 subspace 知 $\mathrm{Span}({}_1\mathbf{a}',\ldots,{}_m\mathbf{a}') \subseteq \mathrm{Span}({}_1\mathbf{a},\ldots,{}_m\mathbf{a})$. 同理, 因 elementary row operation 是可以還原的, A' 也可經由 elementary row operations 換成 A, 所以我們也有 $\mathrm{Span}({}_1\mathbf{a},\ldots,{}_m\mathbf{a}) \subseteq \mathrm{Span}({}_1\mathbf{a}',\ldots,{}_m\mathbf{a}')$. 得證 $\mathrm{Span}({}_1\mathbf{a},\ldots,{}_m\mathbf{a}) = \mathrm{Span}({}_1\mathbf{a}',\ldots,{}_m\mathbf{a}')$, 亦即 A 和 A' 有相同的 row space. 我們看以下的例子.

Example 3.6.10. 考慮 Example 3.6.5 中的 4×6 matrix A, 且利用 elementary row operation 將之化為 reduced echelon form A' (參見 Example 3.6.7), 令 ${}_1\mathbf{a},{}_2\mathbf{a},{}_3\mathbf{a},{}_4\mathbf{a}$ 為 A 的 row vectors, ${}_1\mathbf{a}',{}_2\mathbf{a}',{}_3\mathbf{a}',{}_4\mathbf{a}'$ 為 A' 的 row vectors. 亦即

$$ {}_1\mathbf{a} = [2\ 1\ 1\ 0\ 0\ 0], {}_2\mathbf{a} = [1\ 0\ 0\ 1\ 0\ 0], {}_3\mathbf{a} = [1\ 1\ 1\ 0\ 1\ 2], {}_4\mathbf{a} = [1\ 2\ 2\ -2\ 1\ 2], $$

$$ {}_1\mathbf{a}' = [1\ 0\ 0\ 0\ -1\ -2], {}_2\mathbf{a}' = [0\ 1\ 1\ 0\ 2\ 4], {}_3\mathbf{a}' = [0\ 0\ 0\ 1\ 1\ 2], {}_4\mathbf{a}' = [0\ 0\ 0\ 0\ 0\ 0]. $$

利用 Example 3.6.5 的 elementary row operations, 我們知 A' 的 3-rd row ${}_3\mathbf{a}'$ 是由 A 的 3-rd row 減去 A 的 2-nd row 後再減去 A 的 2-nd row 乘上 -2 加到 1-st row 的向量, 亦即

$$ ({}_3\mathbf{a} - {}_2\mathbf{a}) - ({}_1\mathbf{a} - 2\,{}_2\mathbf{a}) = {}_3\mathbf{a} + {}_2\mathbf{a} - {}_1\mathbf{a} = [1\ 1\ 1\ 0\ 1\ 2] + [1\ 0\ 0\ 1\ 0\ 0] - [2\ 1\ 1\ 0\ 0\ 0] = {}_3\mathbf{a}'. $$

而利用 Example 3.6.7 的 elementary row operations, A' 的 2-rd row ${}_2\mathbf{a}'$ 是由 A 的 2-nd row 乘上 -2 加到 A 的 1-st row 後再加上 2 倍的 A' 的 3-rd row 的向量, 亦即

$$({}_1\mathbf{a}-2\,{}_2\mathbf{a})+2({}_3\mathbf{a}+{}_2\mathbf{a}-{}_1\mathbf{a})=2\,{}_3\mathbf{a}-{}_1\mathbf{a}=[2\ 2\ 2\ 0\ 2\ 4]-[2\ 1\ 1\ 0\ 0\ 0]={}_2\mathbf{a}'.$$

而 A' 的 1-st row ${}_1\mathbf{a}'$ 是由 A 的 2-nd row 減去 A' 的 3-rd row 的向量, 亦即

$${}_2\mathbf{a}-({}_3\mathbf{a}+{}_2\mathbf{a}-{}_1\mathbf{a})={}_1\mathbf{a}-{}_3\mathbf{a}=[2\ 1\ 1\ 0\ 0\ 0]-[1\ 1\ 1\ 0\ 1\ 2]={}_1\mathbf{a}'.$$

從這裡我們得 $\mathrm{Span}({}_1\mathbf{a}',{}_2\mathbf{a}',{}_3\mathbf{a}',{}_4\mathbf{a}')\subseteq\mathrm{Span}({}_1\mathbf{a},{}_2\mathbf{a},{}_3\mathbf{a},{}_4\mathbf{a})$. 同理得 $\mathrm{Span}({}_1\mathbf{a},{}_2\mathbf{a},{}_3\mathbf{a},{}_4\mathbf{a})\subseteq\mathrm{Span}({}_1\mathbf{a}',{}_2\mathbf{a}',{}_3\mathbf{a}',{}_4\mathbf{a}')$ (此處略去不檢查了). 故得 $\mathrm{Span}({}_1\mathbf{a},{}_2\mathbf{a},{}_3\mathbf{a},{}_4\mathbf{a})=\mathrm{Span}({}_1\mathbf{a}',{}_2\mathbf{a}',{}_3\mathbf{a}',{}_4\mathbf{a}')$, 亦即 ${}_1\mathbf{a}',{}_2\mathbf{a}',{}_3\mathbf{a}',{}_4\mathbf{a}'$ 為 A 的 row space 的 spanning vectors. 在 echelon form 中, 沒有 pivot 的 row 必為零向量. 現 A' 的 pivot 個數為 3, 即 pivot 發生於前 3 個 row ${}_1\mathbf{a}',{}_2\mathbf{a}',{}_3\mathbf{a}'$, 而 ${}_4\mathbf{a}'$ 為零向量, 所以僅 pivot 所在的 row ${}_1\mathbf{a}',{}_2\mathbf{a}',{}_3\mathbf{a}'$ 就可以成為 A 的 row space 的 spanning vectors. 現又由於 A' 為 reduced echelon form, 每一個 row 中 pivot 所在的位置其他的 row 在該位置皆為 0, 所以 ${}_1\mathbf{a}',{}_2\mathbf{a}',{}_3\mathbf{a}'$ 為 linearly independent. 得證 ${}_1\mathbf{a}',{}_2\mathbf{a}',{}_3\mathbf{a}'$ 為 A 的 row space 的一組 basis.

注意在 Example 3.6.10 中我們為了方便說明將 A 化為 reduced echelon form. 事實上既然 A 的 echelon form 和 reduced echelon form 有相同的 row space, 而它們 pivot 的個數又相同, 所以由 dimension 的性質, 知 echelon form 中 pivot 所在的 row vectors 也會是 A 的 row space 的一組 basis. 化成 reduced echelon form 的好處是比較容易讓我們將 row space 中的向量寫成這組 basis 的線性組合. 因此因此除非我們想要將 A 的 row space 中的 vectors 用這組 basis 來表示, 若僅想找到 row space 的 basis 一般是不需要進一步化成 reduced echelon form. 另外我們要強調的是 row space 的 basis 不可以回到 A 的 row vectors 去找. 這是因為一般我們在做 elementary row operations 已將 row vectors 所在的位置做了調動, 所以 row operation 並沒有保持 row vectors 之間的線性關係.

我們將這個求 row space 的 basis 的方法做一個總結. 首先將 $m\times n$ matrix A 利用 elementary row operation 化為 echelon form A'. 假設 A' 的 pivot 個數為 r, 則由於 A' 為 echelon form, A' 前 r 個 row vectors ${}_1\mathbf{a}',\dots,{}_r\mathbf{a}'$ 為 nonzero vectors. A' 其餘的 row vectors 皆為 zero vectors. 由於 elementary row operations 會保持 row space, 我們得 ${}_1\mathbf{a}',\dots,{}_r\mathbf{a}'$ 為 A 的 row space 的 spanning vectors. 又由化為 reduced echelon form 的情形我們知 A 的 row space 的 dimension 為 r, 故由 Proposition 2.6.11 知 ${}_1\mathbf{a}',\dots,{}_r\mathbf{a}'$ 為 A 的 row space 的一組 basis. 我們有以下的定理.

Proposition 3.6.11. 假設 A 為 $m\times n$ *matrix.* 若利用 *elementary row operations* 將 A 化為 *echelon form* A' 後*,* A' 的 *pivot* 個數為 r*,* 則 A 的 *row space* 的 *dimension* 為 r 且 A' 的前 r 個 *row vectors* ${}_1\mathbf{a}',\dots,{}_r\mathbf{a}'$ (即 A' 中的 *nonzero row vectors*) 為 A 的 *row space* 的一組 *basis.*

我們可以利用找 column space 和 row space 的 basis 的方法找一般 $\mathbb{F}^m$ 的 subspace V 的 basis. 首先我們先找出 V 的一組 spanning vectors $\mathbf{v}_1,\dots,\mathbf{v}_n$, 然後造一個以 $\mathbf{v}_1,\dots,\mathbf{v}_n$

為 column vectors 的矩陣 $A=\begin{bmatrix} | & & | \\ \mathbf{v}_1 & \cdots & \mathbf{v}_n \\ | & & | \end{bmatrix}$. 然後再利用找 A 的 column space 的 basis 的方法得到 V 的一組 basis. 我們也可造一個以 $\mathbf{v}_1^t,\ldots,\mathbf{v}_n^t$ 為 row vectors 的 $n\times m$ matrix $B=\begin{bmatrix} — & \mathbf{v}_1^t & — \\ & \vdots & \\ — & \mathbf{v}_n^t & — \end{bmatrix}$ 然後再利用找 B 的 row space 的 basis 的方法得到 V 的一組 basis. 兩種方法都有它們的好處. 利用 column vectors 的方法, 由於最後找出的 basis 是從原來的 spanning vectors $\mathbf{v}_1,\ldots,\mathbf{v}_n$ 中的向量所組成, 所以適合處理希望 basis 的 vectors 是由原來 spanning vectors 中選出的問題. 而利用 row vectors 的方法, 由於可以化為 reduced echelon form, 而 basis 是由此 reduced echelon form 中的 nonzero vectors 所組成, 所以雖然和來的 spanning vectors 無關, 不過很適合拿來判斷哪些向量在此 subspace 以及處理將 subspace 中的向量用此 basis 表示的問題. 例如以下的例子.

Example 3.6.12. 考慮 $\mathbb{R}^6$ 中的向量

$$\mathbf{v}_1=\begin{bmatrix}2\\1\\1\\0\\0\\0\end{bmatrix},\mathbf{v}_2=\begin{bmatrix}1\\0\\0\\1\\0\\0\end{bmatrix},\mathbf{v}_3=\begin{bmatrix}1\\1\\1\\0\\1\\2\end{bmatrix},\mathbf{v}_4=\begin{bmatrix}1\\2\\2\\-2\\1\\2\end{bmatrix}\quad\text{and}\quad\mathbf{w}=\begin{bmatrix}1\\-2\\-2\\3\\-2\\4\end{bmatrix}.$$

令 $V=\mathrm{Span}(\mathbf{v}_1,\mathbf{v}_2,\mathbf{v}_3,\mathbf{v}_4)$, 試找出 V 的一組 basis, 並用之判斷 $\mathbf{w}$ 是否在 V 中.

依前面的討論我們知此問題適合用 row space 的方式處理, 所以考慮以 $\mathbf{v}_1,\mathbf{v}_2,\mathbf{v}_3,\mathbf{v}_4$ 為 row vectors 的矩陣 A. 此時 A 就是 Example 3.6.10 中的矩陣 A. 利用 Example 3.6.10 的結果, 我們得

$$\mathbf{u}_1=\begin{bmatrix}1\\0\\0\\0\\-1\\-2\end{bmatrix},\mathbf{u}_2=\begin{bmatrix}0\\1\\1\\0\\2\\4\end{bmatrix},\mathbf{u}_3=\begin{bmatrix}0\\0\\0\\1\\1\\2\end{bmatrix}$$

為 V 的一組 basis. 注意因為原來 V 中的向量為 column vector 的形式, 為了一致性這裡我們將 A 的 row space 的 basis 寫回成 column vectors. 依此我們知 $\mathbf{w}\in V$ 若且唯若存在 $c_1,c_2,c_3\in\mathbb{R}$ 使得 $\mathbf{w}=c_1\mathbf{u}_1+c_2\mathbf{u}_2+c_3\mathbf{u}_3$. 然而

$$c_1\mathbf{u}_1+c_2\mathbf{u}_2+c_3\mathbf{u}_3=\begin{bmatrix}c_1\\c_2\\c_2\\c_3\\-c_1+2c_2+c_3\\-2c_1+4c_2+2c_3\end{bmatrix},$$

我們發現要使得 $\mathbf{w}=c_1\mathbf{u}_1+c_2\mathbf{u}_2+c_3\mathbf{u}_3$, 在 1-st, 2-nd 和 3-rd entry 的地方 (這些地方就是 reduced echelon form 的 pivot 的位置) 需有 $c_1=1,c_2=-2,c_3=3$. 故將 $c_1=1,c_2=-2,c_3=3$ 代入, 發現每個 entry 皆吻合, 故有 $\mathbf{w}=\mathbf{u}_1-2\mathbf{u}_2+3\mathbf{u}_3$, 得知 $\mathbf{w}\in V$.

當我們對要找的 basis 沒有特殊要求時, 我們可以選擇兩種方法中可以使得矩陣的 row 的個數較少的那一種方法處理. 因為如此所需的 elementary row operations 相對起來會較少, 可以較快找到一組 basis. 例如前一個例子 Example 3.6.12 中的 $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \mathbf{v}_4$ 若考慮成 column vectors, 所得的矩陣為 6×4 matrix, 而考慮成 row vectors, 所得的矩陣為 4×6 matrix. 所以此時若僅想找出 $\mathrm{Span}(\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \mathbf{v}_4)$ 的一組 basis, 用 row vectors 的情況處理會比較快.

給定一個矩陣 A, 從 Proposition 3.6.8 和 Proposition 3.6.11 我們知道 A 的 column space 和 row space 有相同的 dimension, 亦即 A 利用 elementary row operations 化為 echelon form 後其 pivot 的個數, 即 A 的 rank. A 的 column space 的維度就是 A 的 rank, 至於 A 的 nullspace 的維度, 我們也給予一特殊的名稱, 即以下的定義.

Definition 3.6.13. 假設 A 為 $m \times n$ matrix. A 的 null space 的 dimension 稱為 A 的 *nullity*, 記為 $\mathrm{nullity}(A)$, 亦即 $\mathrm{nullity}(A) = \dim(N(A))$.

依此定義, 由 Proposition 3.6.8 和 Proposition 3.6.11 我們知道 $\mathrm{rank}(A)$ 即為 A 利用 elementary row operations 化為 echelon form 後其 pivot 的個數, 而由 Proposition 3.6.4 我們知道 $\mathrm{nullity}(A)$ 就是 homogeneous linear system $A\mathbf{x} = \mathbf{0}$ 的 free variables 的個數, 即 A 的 column 個數減去 echelon form 的 pivot 個數, 因此我們有以下的 *Dimension Theorem* (或稱為 *rank equation*).

Theorem 3.6.14 (Dimension Theorem)**.** 假設 A 為 $m \times n$ *matrix.* 則

$$\mathrm{rank}(A) + \mathrm{nullity}(A) = n.$$

Question 3.12. 假設 A 為 $n \times n$ *invertible matrix.* 試求 $\mathrm{rank}(A)$ 以及 $\mathrm{nullity}(A)$.

Question 3.13. 假設 $A \in M_{m\times n}(\mathbb{F})$.

(1) 若對於任意 $\mathbf{v} \in \mathbb{F}^m$ 聯立方程組 $A\mathbf{x} = \mathbf{v}$ 皆有解, 試求 $\mathrm{rank}(A)$ 以及 $\mathrm{nullity}(A)$.

(2) 若存在 $\mathbf{v} \in \mathbb{F}^m$ 使得聯立方程組 $A\mathbf{x} = \mathbf{v}$ 有唯一解, 試求 $\mathrm{rank}(A)$ 以及 $\mathrm{nullity}(A)$.

Proposition 3.6.8 告訴我們 A 的 column space 的維度就是 A 的 rank, 亦即 $\dim(\mathrm{Col}(A)) = \mathrm{rank}(A)$, 而 Proposition 3.6.11 告訴我們 $\dim(\mathrm{Row}(A)) = \mathrm{rank}(A)$, 因此得 $\dim(\mathrm{Col}(A)) = \dim(\mathrm{Row}(A))$. 也就是說一個矩陣的 column space 和 row space 有相同的維度. 現考慮矩陣 A 的 transpose A^{t}. 由於 A 的 column space 就是 A^{t} 的 row space (且 A 的 row space 就是 A^{t} 的 column space), 所以我們有 $\mathrm{rank}(A) = \dim(\mathrm{Col}(A)) = \dim(\mathrm{Row}(A^{\mathrm{t}})) = \mathrm{rank}(A^{\mathrm{t}})$. 得證以下的性質.

Proposition 3.6.15. 假設 $A \in M_{m\times n}(\mathbb{F})$. 則

$$\dim(\mathrm{Col}(A)) = \dim(\mathrm{Row}(A)) = \mathrm{rank}(A) = \mathrm{rank}(A^{\mathrm{t}}).$$

亦即利用 *elementary row operations* 將 A 以及 A^{t} 化為 *echelon form* 後, 它們的 *pivot* 個數相同.

Question 3.14. 假設 A 為 $m\times n$ *matrix.* 試證明 $\operatorname{nullity}(A^{t})=m-\operatorname{rank}(A)$.

注意, 對於 Proposition 3.6.15 的證明, 若想要直接證明利用 elementary row operations 將 A 以及 A^{t} 化為 echelon form 後, 它們的 pivot 個數相同, 會有相當的困難度. 但將這個問題轉為 column space 以及 row space 的 dimension 問題, 就很容易解決. 其實許多數學問題都是這樣, 有時只要換個角度看問題就能迎刃而解. 這也是當我們介紹一個新的概念後常常會去探討這個概念還有甚麼等價的條件的原因. 最後我們利用類似的概念來處理兩個矩陣相乘後 rank 的變化關係.

Proposition 3.6.16. 假設 $A\in M_{m\times n}(\mathbb{F})$, $B\in M_{n\times l}(\mathbb{F})$. 則

(1) $\operatorname{Col}(AB)\subseteq\operatorname{Col}(A)$ 且 $\operatorname{rank}(AB)\leq\operatorname{rank}(A)$.

(2) $\operatorname{Row}(AB)\subseteq\operatorname{Row}(B)$ 且 $\operatorname{rank}(AB)\leq\operatorname{rank}(B)$.

(3) 若 $E\in M_{n\times n}(\mathbb{F})$ 為 *invertible,* 則 $\operatorname{Col}(AE)=\operatorname{Col}(A)$ 且 $\operatorname{rank}(AE)=\operatorname{rank}(A)$.

(4) 若 $H\in M_{m\times m}(\mathbb{F})$ 為 *invertible,* 則 $\operatorname{Row}(HA)=\operatorname{Row}(A)$ 且 $\operatorname{rank}(HA)=\operatorname{rank}(A)$.

Proof. 令 A 和 B 的 column vectors 依序為 $\mathbf{a}_1,\ldots,\mathbf{a}_n$ 和 $\mathbf{b}_1,\ldots,\mathbf{b}_l$ 且令 A 和 B 的 row vectors 依序為 ${}_1\mathbf{a},\ldots,{}_m\mathbf{a}$ 和 ${}_1\mathbf{b},\ldots,{}_n\mathbf{b}$.

(1) 依定義 $\operatorname{Col}(AB)=\operatorname{Span}(A\mathbf{b}_1,\ldots,A\mathbf{b}_l)$, 而對任意 $\mathbf{b}=\begin{bmatrix} b_1\\ \vdots\\ b_n\end{bmatrix}$, 我們有

$$A\mathbf{b}=\begin{bmatrix} | & & | \\ \mathbf{a}_1 & \cdots & \mathbf{a}_n \\ | & & | \end{bmatrix}\begin{bmatrix} b_1\\ \vdots\\ b_n\end{bmatrix}=b_1\mathbf{a}_1+\cdots+b_n\mathbf{a}_n\in\operatorname{Span}(\mathbf{a}_1,\ldots,\mathbf{a}_n)=\operatorname{Col}(A).$$

因此 $A\mathbf{b}_i\in\operatorname{Col}(A)$, $\forall 1\leq i\leq l$. 因此得 $\operatorname{Col}(AB)=\operatorname{Span}(A\mathbf{b}_1,\ldots,A\mathbf{b}_l)\subseteq\operatorname{Col}(A)$. 換言之 $\operatorname{Col}(AB)$ 為 $\operatorname{Col}(A)$ 的 subspace, 故由 Proposition 2.6.11 (4) 知 $\operatorname{rank}(AB)=\dim(\operatorname{Col}(AB))\leq\dim(\operatorname{Col}(A))=\operatorname{rank}(A)$.

(2) 依定義 $\operatorname{Row}(AB)=\operatorname{Span}({}_1\mathbf{a}B,\ldots,{}_m\mathbf{a}B)$, 而對任意 $\mathbf{a}=\begin{bmatrix} a_1 & \cdots & a_n\end{bmatrix}$, 我們有

$$\mathbf{a}B=\begin{bmatrix} a_1 & \cdots & a_n\end{bmatrix}\begin{bmatrix} — & {}_1\mathbf{b} & — \\ & \vdots & \\ — & {}_n\mathbf{b} & — \end{bmatrix}=a_1({}_1\mathbf{b})+\cdots+a_n({}_n\mathbf{b})\in\operatorname{Span}({}_1\mathbf{b},\ldots,{}_n\mathbf{b})=\operatorname{Row}(B).$$

因此 ${}_i\mathbf{a}B\in\operatorname{Row}(B)$, $\forall 1\leq i\leq m$. 因此得 $\operatorname{Row}(AB)=\operatorname{Span}({}_1\mathbf{a}B,\ldots,{}_m\mathbf{a}B)\subseteq\operatorname{Row}(B)$. 換言之 $\operatorname{Row}(AB)$ 為 $\operatorname{Row}(B)$ 的 subspace, 故由 Proposition 2.6.11 (4) 知 $\operatorname{rank}(AB)=\dim(\operatorname{Row}(AB))\leq\dim(\operatorname{Row}(B))=\operatorname{rank}(B)$.

(3) 利用前面 (1) 的結果我們知 $\operatorname{Col}(AE)\subseteq\operatorname{Col}(A)$. 因 E 為 invertible, 考慮 $(AE)E^{-1}=A(EE^{-1})=A$. 再利用 (1) 知 $\operatorname{Col}(A)=\operatorname{Col}((AE)E^{-1})\subseteq\operatorname{Col}(AE)$. 因此得證 $\operatorname{Col}(AE)=\operatorname{Col}(A)$ 且取 dimension 得 $\operatorname{rank}(AE)=\dim(\operatorname{Col}(AE))=\dim(\operatorname{Col}(A))=\operatorname{rank}(A)$.

(4) 利用前面 (2) 的結果我們知 $\mathrm{Row}(HA) \subseteq \mathrm{Row}(A)$. 因 H 為 invertible, 考慮 $H^{-1}(HA) = (H^{-1}H)A = A$. 再利用 (2) 知 $\mathrm{Row}(A) = \mathrm{Row}(H^{-1}(HA)) \subseteq \mathrm{Row}(HA)$. 因此得證 $\mathrm{Row}(HA) = \mathrm{Row}(A)$ 且取 dimension 得 $\mathrm{rank}(HA) = \dim(\mathrm{Row}(HA)) = \dim(\mathrm{Row}(A)) = \mathrm{rank}(A)$. $\square$

注意, 在 Proposition 3.6.16 (3) 中我們知道當 E 為 invertible 時 $\mathrm{rank}(AE) = \mathrm{rank}(A)$, 也因此知 $\dim(\mathrm{Row}(AE)) = \mathrm{rank}(AE) = \mathrm{rank}(A) = \dim(Row(A))$. 不過這並不代表 $\mathrm{Row}(AE) = \mathrm{Row}(A)$. 這是因為 $\mathrm{Row}(AE)$ 和 $\mathrm{Row}(A)$ 未必有包含關係, 因此即使 $\dim(\mathrm{Row}(AE)) = \dim(\mathrm{Row}(A))$, 也無法推得 $\mathrm{Row}(AE) = \mathrm{Row}(A)$. 同理在 Proposition 3.6.16 (4) 中我們知道當 H 為 invertible 時 $\dim(\mathrm{Col}(HA)) = \dim(\mathrm{Col}(A))$, 但 $\mathrm{Col}(HA)$ 也未必等於 $\mathrm{Col}(A)$.

Question 3.15. 試找到例子 $A, E \in M_{3\times 3}(\mathbb{R})$ 其中 E 為 *invertible* 使得 $\mathrm{Row}(AE) \neq \mathrm{Row}(A)$, $\mathrm{Col}(EA) \neq \mathrm{Col}(A)$. *(Hint: 考慮 E 為 elementary matrix.)*

Question 3.16. 試用取轉置矩陣的方法利用 *Proposition 3.6.16 (1)* 證明 *(2)*, 且利用 *(3)* 證明 *(4)*.

Question 3.17. 試利用 *Proposition 3.5.7*, 即 *invertible matrix* 皆可寫成 *elementary matrices* 的乘積, 證明 *Proposition 3.6.16 (3), (4)*.