2.4. 另類的矩陣乘法看法

我們前面介紹過矩陣乘法的三種看法：inner product（每個 entry 來看）、multiplication by column（每個 column 來看）、multiplication by row（每個 row 來看）。這三種看法，在以後談論解聯立方程、向量空間等概念非常重要。然而在 AI 的應用方面，需要處理階數很大的矩陣乘法運算，這時候另一種乘法看法，有的稱之為 *outer product*（注意不是我們稱為的向量外積 *cross product*），就非常有幫助。

我們知道矩陣的加法、乘法有分配律（Proposition 2.1.9）。當 A 為 $m\times n$ 矩陣，我們可將 A 依序寫成 m 個 $m\times n$ 矩陣相加，其中第一個為留下 1-st row，其他 row 改為 $\mathbf{0}$ 的 $m\times n$ 矩陣；第二個加為留下 2-nd row，其他 row 為 $\mathbf{0}$ 的 $m\times n$ 矩陣；其餘依此類推。例如將 $\begin{bmatrix} a_{11} & a_{12} & a_{13}\\ a_{21} & a_{22} & a_{23}\end{bmatrix}$ 寫成

$$\begin{bmatrix} a_{11} & a_{12} & a_{13}\\ a_{21} & a_{22} & a_{23}\end{bmatrix}=\begin{bmatrix} a_{11} & a_{12} & a_{13}\\ 0 & 0 & 0\end{bmatrix}+\begin{bmatrix} 0 & 0 & 0\\ a_{21} & a_{22} & a_{23}\end{bmatrix}.$$

我們也可將 A 依序寫成 n 個 $m\times n$ 矩陣相加，其中第一個為留下 1-st column，其他 column 改為 $\mathbf{0}$ 的 $m\times n$ 矩陣；第二個加為留下 2-nd column，其他 column 為 $\mathbf{0}$ 的 $m\times n$ 矩陣；其餘依此類推。例如將 $\begin{bmatrix} a_{11} & a_{12} & a_{13}\\ a_{21} & a_{22} & a_{23}\end{bmatrix}$ 寫成

$$\begin{bmatrix} a_{11} & a_{12} & a_{13}\\ a_{21} & a_{22} & a_{23}\end{bmatrix}=\begin{bmatrix} a_{11} & 0 & 0\\ a_{21} & 0 & 0\end{bmatrix}+\begin{bmatrix} 0 & a_{12} & 0\\ 0 & a_{22} & 0\end{bmatrix}+\begin{bmatrix} 0 & 0 & a_{13}\\ 0 & 0 & a_{23}\end{bmatrix}.$$

同樣的，若 B 為 $n\times l$ 矩陣，我們也可將之寫成如前這些簡單的矩陣相加。此時若要計算矩陣乘法 AB，我們就可利用加法、乘法的分配律將之展開來計算。我們看以下的例子：

Example 2.4.1. 考慮 $A=\begin{bmatrix} a_{11} & a_{12} & a_{13}\\ a_{21} & a_{22} & a_{23}\end{bmatrix}$ 以及 $B=\begin{bmatrix} b_{11} & b_{12}\\ b_{21} & b_{22}\\ b_{31} & b_{32}\end{bmatrix}$。我們將 A,B 寫成簡單形式的矩陣加法，再利用分配律計算 AB。共有以下四種拆解方法：

(1) $\left(\begin{bmatrix} a_{11} & a_{12} & a_{13}\\ 0 & 0 & 0\end{bmatrix}+\begin{bmatrix} 0 & 0 & 0\\ a_{21} & a_{22} & a_{23}\end{bmatrix}\right)\left(\begin{bmatrix} b_{11} & 0\\ b_{21} & 0\\ b_{31} & 0\end{bmatrix}+\begin{bmatrix} 0 & b_{12}\\ 0 & b_{22}\\ 0 & b_{32}\end{bmatrix}\right)$：由分配律，我們首先看 $\begin{bmatrix} a_{11} & a_{12} & a_{13}\\ 0 & 0 & 0\end{bmatrix}\begin{bmatrix} b_{11} & 0\\ b_{21} & 0\\ b_{31} & 0\end{bmatrix}$ 這部分。由於左邊矩陣第二個 row 全為 0 可知乘出來的矩陣第二個 row 全為 0；而右邊矩陣第二個 column 全為 0 可知乘出來的矩陣第二個 column 也是全為 0。而僅剩的 (1,1)-entry 就是內積 ${}_1\mathbf{a}\cdot\mathbf{b}_1=a_{11}b_{11}+a_{12}b_{21}+a_{13}b_{31}$。同樣的，$\begin{bmatrix} a_{11} & a_{12} & a_{13}\\ 0 & 0 & 0\end{bmatrix}\begin{bmatrix} 0 & b_{12}\\ 0 & b_{22}\\ 0 & b_{32}\end{bmatrix}$ 乘出來的矩陣其第二個 row 和第一個 column 皆為 0。而僅剩的 (1,2)-entry 就是內積 ${}_1\mathbf{a}\cdot\mathbf{b}_2=a_{11}b_{12}+a_{12}b_{22}+a_{13}b_{32}$。同樣的，對其他矩陣展開相乘，可得 AB 等於以下四個矩陣之和 $\begin{bmatrix} {}_1\mathbf{a}\cdot\mathbf{b}_1 & 0\\ 0 & 0\end{bmatrix}+\begin{bmatrix} 0 & {}_1\mathbf{a}\cdot\mathbf{b}_2\\ 0 & 0\end{bmatrix}+\begin{bmatrix} 0 & 0\\ {}_2\mathbf{a}\cdot\mathbf{b}_1 & 0\end{bmatrix}+\begin{bmatrix} 0 & 0\\ 0 & {}_2\mathbf{a}\cdot\mathbf{b}_2\end{bmatrix}$。這樣的拆解和矩陣乘法的內積看法是一樣的。

(2) $\left(\begin{bmatrix}a_{11}&0&0\\a_{21}&0&0\end{bmatrix}+\begin{bmatrix}0&a_{12}&0\\0&a_{22}&0\end{bmatrix}+\begin{bmatrix}0&0&a_{13}\\0&0&a_{23}\end{bmatrix}\right)\left(\begin{bmatrix}b_{11}&0\\b_{21}&0\\b_{31}&0\end{bmatrix}+\begin{bmatrix}0&b_{12}\\0&b_{22}\\0&b_{32}\end{bmatrix}\right)$：由分配律，我們首先看 $\begin{bmatrix}a_{11}&0&0\\a_{21}&0&0\end{bmatrix}\begin{bmatrix}b_{11}&0\\b_{21}&0\\b_{31}&0\end{bmatrix}$ 這部分。由於右邊矩陣第二個 column 全為 0 可知乘出來的矩陣第二個 column 也是全為 0。而僅剩的第一個 column，由矩陣乘法的 column 看法就是 $b_{11}\mathbf{a}_1=b_{11}\begin{bmatrix}a_{11}\\a_{21}\end{bmatrix}$。同樣的，$\begin{bmatrix}0&a_{12}&0\\0&a_{22}&0\end{bmatrix}\begin{bmatrix}b_{11}&0\\b_{21}&0\\b_{31}&0\end{bmatrix}$、$\begin{bmatrix}0&0&a_{13}\\0&0&a_{23}\end{bmatrix}\begin{bmatrix}b_{11}&0\\b_{21}&0\\b_{31}&0\end{bmatrix}$ 分別乘出來的矩陣其第二個 column 皆為 0。而它們的第一個 column，分別為 $b_{21}\mathbf{a}_2=b_{21}\begin{bmatrix}a_{12}\\a_{22}\end{bmatrix}$、$b_{31}\mathbf{a}_3=b_{31}\begin{bmatrix}a_{13}\\a_{23}\end{bmatrix}$。同樣的，對其他矩陣展開相乘，可得 AB 等於以下六個矩陣之和 $b_{11}\begin{bmatrix}a_{11}&0\\a_{21}&0\end{bmatrix}+b_{21}\begin{bmatrix}a_{12}&0\\a_{22}&0\end{bmatrix}+b_{31}\begin{bmatrix}a_{13}&0\\a_{23}&0\end{bmatrix}+b_{12}\begin{bmatrix}0&a_{11}\\0&a_{21}\end{bmatrix}+b_{22}\begin{bmatrix}0&a_{12}\\0&a_{22}\end{bmatrix}+b_{32}\begin{bmatrix}0&a_{13}\\0&a_{23}\end{bmatrix}$。這樣的拆解和矩陣乘法的 column 看法是一樣的。

(3) $\left(\begin{bmatrix}a_{11}&a_{12}&a_{13}\\0&0&0\end{bmatrix}+\begin{bmatrix}0&0&0\\a_{21}&a_{22}&a_{23}\end{bmatrix}\right)\left(\begin{bmatrix}b_{11}&b_{12}\\0&0\\0&0\end{bmatrix}+\begin{bmatrix}0&0\\b_{21}&b_{22}\\0&0\end{bmatrix}+\begin{bmatrix}0&0\\0&0\\b_{31}&b_{32}\end{bmatrix}\right)$：由分配律，我們首先看 $\begin{bmatrix}a_{11}&a_{12}&a_{13}\\0&0&0\end{bmatrix}\begin{bmatrix}b_{11}&b_{12}\\0&0\\0&0\end{bmatrix}$ 這部分。由於左邊矩陣第二個 row 全為 0 可知乘出來的矩陣第二個 row 也是全為 0。而僅剩的第一個 row，由矩陣乘法的 row 看法就是 $a_{11}\begin{bmatrix}b_{11}&b_{12}\end{bmatrix}$。同樣的，$\begin{bmatrix}a_{11}&a_{12}&a_{13}\\0&0&0\end{bmatrix}\begin{bmatrix}0&0\\b_{21}&b_{22}\\0&0\end{bmatrix}$、$\begin{bmatrix}a_{11}&a_{12}&a_{13}\\0&0&0\end{bmatrix}\begin{bmatrix}0&0\\0&0\\b_{31}&b_{32}\end{bmatrix}$ 分別乘出來的矩陣其第二個 row 皆為 0。而它們的第一個 row，分別為 $a_{12}\begin{bmatrix}b_{21}&b_{22}\end{bmatrix}$ 以及 $a_{13}\begin{bmatrix}b_{31}&b_{32}\end{bmatrix}$。同樣的，對其他矩陣展開相乘，可得 AB 等於以下六個矩陣之和 $a_{11}\begin{bmatrix}b_{11}&b_{12}\\0&0\end{bmatrix}+a_{12}\begin{bmatrix}b_{21}&b_{22}\\0&0\end{bmatrix}+a_{13}\begin{bmatrix}b_{31}&b_{32}\\0&0\end{bmatrix}+a_{21}\begin{bmatrix}0&0\\b_{11}&b_{12}\end{bmatrix}+a_{22}\begin{bmatrix}0&0\\b_{21}&b_{22}\end{bmatrix}+a_{23}\begin{bmatrix}0&0\\b_{31}&b_{32}\end{bmatrix}$ 這樣的拆解和矩陣乘法的 row 看法是一樣的。

(4)

$$\left(\begin{bmatrix}a_{11}&0&0\\a_{21}&0&0\end{bmatrix}+\begin{bmatrix}0&a_{12}&0\\0&a_{22}&0\end{bmatrix}+\begin{bmatrix}0&0&a_{13}\\0&0&a_{23}\end{bmatrix}\right)\left(\begin{bmatrix}b_{11}&b_{12}\\0&0\\0&0\end{bmatrix}+\begin{bmatrix}0&0\\b_{21}&b_{22}\\0&0\end{bmatrix}+\begin{bmatrix}0&0\\0&0\\b_{31}&b_{32}\end{bmatrix}\right):$$

由分配律，我們首先看 $\begin{bmatrix}a_{11}&0&0\\a_{21}&0&0\end{bmatrix}\begin{bmatrix}b_{11}&b_{12}\\0&0\\0&0\end{bmatrix}$ 這部分。由 row 看法，我們知乘出來的矩陣第一個 row 為 $a_{11}\begin{bmatrix}b_{11}&b_{12}\end{bmatrix}$，而第二個 row 為 $a_{12}\begin{bmatrix}b_{11}&b_{12}\end{bmatrix}$。不過，$\begin{bmatrix}a_{11}&0&0\\a_{21}&0&0\end{bmatrix}\begin{bmatrix}0&0\\b_{21}&b_{22}\\0&0\end{bmatrix}$、$\begin{bmatrix}a_{11}&0&0\\a_{21}&0&0\end{bmatrix}\begin{bmatrix}0&0\\0&0\\b_{31}&b_{32}\end{bmatrix}$ 分別乘出來的矩陣皆為零矩陣。同樣的，$\begin{bmatrix}0&a_{12}&0\\0&a_{22}&0\end{bmatrix}$ 除了乘上

$\begin{bmatrix} 0 & 0 \\ b_{21} & b_{22} \\ 0 & 0 \end{bmatrix}$ 外，其餘皆為零矩陣；而 $\begin{bmatrix} 0 & 0 & a_{13} \\ 0 & 0 & a_{23} \end{bmatrix}$ 除了乘上 $\begin{bmatrix} 0 & 0 \\ 0 & 0 \\ b_{31} & b_{32} \end{bmatrix}$ 外，其餘皆為零矩陣。由此可得 AB 等於以下三個矩陣之和

$$\begin{bmatrix} a_{11}b_{11} & a_{11}b_{12} \\ a_{21}b_{11} & a_{21}b_{12} \end{bmatrix} + \begin{bmatrix} a_{12}b_{21} & a_{12}b_{22} \\ a_{22}b_{21} & a_{22}b_{22} \end{bmatrix} + \begin{bmatrix} a_{13}b_{31} & a_{13}b_{32} \\ a_{23}b_{31} & a_{23}b_{32} \end{bmatrix}$$

這樣的拆解就是我們要談的矩陣乘法的 outer product 看法。注意，第一個矩陣兩個 row 分別為 $\begin{bmatrix} b_{11} & b_{12} \end{bmatrix}$ 乘上 a_{11}, a_{21}，所以利用 elementary row operation 可知其 rank 最多為 1。其餘兩個矩陣的 rank 也是同樣情況。 $\sharp$

從這個例子可以看出兩矩陣 A,B 相乘 AB 其 outer product 的看法就是將 A 依 column 寫成一些只有一個 column 不為 0 的矩陣相加，而將 B 依 row 寫成一些只有一個 row 不為 0 的矩陣相加，再依分配律展開。其實這個看法並沒有想象的複雜。若 A,B 分別為 $m\times n$ 和 $n\times l$ 矩陣，最後會是 n 個 rank 至多是 1 的矩陣相加。事實上會是以下的 n 個矩陣相加

$$AB = \begin{bmatrix} | & & | \\ \mathbf{a}_1 & \cdots & \mathbf{a}_n \\ | & & | \end{bmatrix} \begin{bmatrix} - & {}_1\mathbf{b} & - \\ & \vdots & \\ - & {}_n\mathbf{b} & - \end{bmatrix} = \begin{bmatrix} | \\ \mathbf{a}_1 \\ | \end{bmatrix} \begin{bmatrix} - & {}_1\mathbf{b} & - \end{bmatrix} + \cdots + \begin{bmatrix} | \\ \mathbf{a}_n \\ | \end{bmatrix} \begin{bmatrix} - & {}_n\mathbf{b} & - \end{bmatrix} \tag{2.19}$$

我們先看實際的例子。

Example 2.4.2. 我們將 Example 2.1.8 中的矩陣乘法 AB 利用 outer product 寫成兩個 rank 1 的矩陣相加。亦即將 AB 寫成

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

其中 $\begin{bmatrix} -2 \\ 3 \\ 2 \end{bmatrix} \begin{bmatrix} 4 & -1 & 2 & 5 \end{bmatrix}$ 這個乘法很簡單，利用 row 的看法其三個 row 依序為 $-2, 3, 2$ 乘上 $\begin{bmatrix} 4 & -1 & 2 & 5 \end{bmatrix}$。亦即 $\begin{bmatrix} -2 \\ 3 \\ 2 \end{bmatrix} \begin{bmatrix} 4 & -1 & 2 & 5 \end{bmatrix} = \begin{bmatrix} -8 & 2 & -4 & -10 \\ 12 & -3 & 6 & 15 \\ 8 & -2 & 4 & 10 \end{bmatrix}$，同理可得

$$AB = \begin{bmatrix} -2 \\ 3 \\ 2 \end{bmatrix} \begin{bmatrix} 4 & -1 & 2 & 5 \end{bmatrix} + \begin{bmatrix} 4 \\ 6 \\ 2 \end{bmatrix} \begin{bmatrix} 3 & 0 & 1 & 1 \end{bmatrix} = \begin{bmatrix} -8 & 2 & -4 & -10 \\ 12 & -3 & 6 & 15 \\ 8 & -2 & 4 & 10 \end{bmatrix} + \begin{bmatrix} 12 & 0 & 4 & 4 \\ 18 & 0 & 6 & 6 \\ 6 & 0 & 2 & 2 \end{bmatrix} \quad \sharp$$

接下來我們說明一下原因。考慮矩陣 A 為 $m\times n$ 矩陣，其 i-th column 為 $\begin{bmatrix} a_1 \\ \vdots \\ a_m \end{bmatrix}$，其他 column 皆為 $\mathbf{0}$；也考慮矩陣 B 為 $n\times l$ 矩陣，其 j-th row 為 $\begin{bmatrix} b_1 & \cdots & b_l \end{bmatrix}$，其他 row 皆為 $\mathbf{0}$。利用 row 的看法，因為 A 的第一個 row 只有在第 i 個位置為 a_1 其他位置為 0，所以 AB 的第一個 row 應為 a_1 乘上 B 的 i-th row。然而若 $i \neq j$，則 B 的 i-th row 全為 0，故 AB 的第一個 row 全為 0；而若 $i = j$ 則 B 的 i-th row 為 $\begin{bmatrix} b_1 & \cdots & b_l \end{bmatrix}$，故 AB 的第一個 row 就是 $a_1\begin{bmatrix} b_1 & \cdots & b_l \end{bmatrix}$。同理，當 $i \neq j$ 時 AB 的任意一個 row 皆為 0；而當 $i = j$

時，AB 的第 k 個 row 為 $a_k\begin{bmatrix} b_1 & \cdots & b_l\end{bmatrix}$。所以若 $i\neq j$ 則 AB 為零矩陣；而若 $i=j$ 則 $AB=\begin{bmatrix} a_1\\ \vdots\\ a_m\end{bmatrix}\begin{bmatrix} b_1 & \cdots & b_l\end{bmatrix}$。

現在回到一般的情況，當要計算 AB 時，我們先將 A 依 column 寫成一些只有一個 column 不為 0 的矩陣相加，而將 B 依 row 寫成一些只有一個 row 不為 0 的矩陣相加，然後用分配律展開。由前可知，此時只要不為 0 的 column 所在的位置和不為 0 的 row 所在的位置不符，相乘起來會是零矩陣；只有不為 0 的 column 所在的位置和不為 0 的 row 所在的位置相符的 “n 對” 矩陣相乘需要考慮。所以我們才會有式子 (2.19) 這樣的結果。

前面提過利用 outer product 將矩陣乘法寫成一些 rank 為 1 的矩陣相加，在 IA 需處理大量資料用到階數很大的矩陣乘法，相當重要。另外以後我們談到矩陣的分解、對角化以及 Singular Value Decomposition 這個看法也很重要。

2.5. Matrix 和 System of Linear Equations 的連結

我們曾經利用 elementary row operations 將增廣矩陣化為 echelon form 來探討其所對應的聯立方程組何時有解以及解是否唯一的問題. 現在我們又知道解一次聯立方程組的問題可以看成矩陣乘法的問題, 這一節中我們就是要用這個觀點進一步探討聯立方程組何時有解以及解是否唯一.

首先由於我們都要用矩陣的乘法來探討, 為了方便起見對於 $\mathbb{R}^n$ 中的向量, 除非特別聲明為 row vector, 我們將一律用 column vector 來表示. 也就是說將它視為一個 $n\times 1$ matrix. 另外回顧, 給定一次聯立方程組

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

我們令

$$A=\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n}\\ a_{21} & a_{22} & \cdots & a_{2n}\\ \vdots & \vdots & \vdots & \vdots\\ a_{m1} & a_{m2} & \cdots & a_{mn}\end{bmatrix}, \mathbf{x}=\begin{bmatrix} x_1\\ x_2\\ \vdots\\ x_n\end{bmatrix}, \mathbf{b}=\begin{bmatrix} b_1\\ b_2\\ \vdots\\ b_m\end{bmatrix},$$

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

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

來表示這一組解, 而說 $\mathbf{x}=\mathbf{c}\in\mathbb{R}^n$ 為 $A\mathbf{x}=\mathbf{b}$ 的一組解. 依矩陣乘法定義這等同於說 A 這一個 $m\times n$ matrix 乘以 $\mathbf{c}$ 這一個 $n\times 1$ matrix 會等於 $\mathbf{b}$ 這一個 $m\times 1$ matrix, 即 $A\mathbf{c}=\mathbf{b}$.

2.5.1. 解的存在性. 我們再一次探討怎樣的 $m \times n$ matrix A 會滿足對任意的 $\mathbf{b} \in \mathbb{R}^m$, 聯立方程組 $A\mathbf{x} = \mathbf{b}$ 皆有解.

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

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

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

Lemma 2.5.1. *假設 $A \in M_{m\times n}$ 且 $\mathbf{b} \in \mathbb{R}^m$. 則 $A\mathbf{x} = \mathbf{b}$ 有解若且唯若 $\mathbf{b} \in \mathrm{Span}(\mathbf{a}_1, \ldots, \mathbf{a}_n)$, 其中 $\mathbf{a}_1, \ldots, \mathbf{a}_n$ 為 A 的 column vectors.*

我們有興趣於知道怎樣的 $m \times n$ matrix A 會使得對任意的 $\mathbf{b} \in \mathbb{R}^m$, 聯立方程組 $A\mathbf{x} = \mathbf{b}$ 皆有解. 我們利用過去學過的幾種不同觀點, 發現有許多和它等價的條件. 首先觀察由於 A 的 column vectors 皆在 $\mathbb{R}^m$ 中, 所以自然有 $\mathrm{Span}(\mathbf{a}_1, \ldots, \mathbf{a}_n) \subseteq \mathbb{R}^m$. 然而由 Lemma 2.5.1 知, 若對於任意 $\mathbf{b} \in \mathbb{R}^m$ 皆會使得 $A\mathbf{x} = \mathbf{b}$ 有解, 表示對任意 $\mathbf{b} \in \mathbb{R}^m$ 皆有 $\mathbf{b} \in \mathrm{Span}(\mathbf{a}_1, \ldots, \mathbf{a}_n)$. 故知此時 $\mathrm{Span}(\mathbf{a}_1, \ldots, \mathbf{a}_n) = \mathbb{R}^m$. 反之, 若 $\mathrm{Span}(\mathbf{a}_1, \ldots, \mathbf{a}_n) = \mathbb{R}^m$, 表示對任意 $\mathbf{b} \in \mathbb{R}^m$ 皆有 $\mathbf{b} \in \mathrm{Span}(\mathbf{a}_1, \ldots, \mathbf{a}_n)$. 同樣由 Lemma 2.5.1 知此即對於任意 $\mathbf{b} \in \mathbb{R}^m$ 皆會使得 $A\mathbf{x} = \mathbf{b}$ 有解. 因此從這觀點來看, 對任意的 $\mathbf{b} \in \mathbb{R}^m$, 聯立方程組 $A\mathbf{x} = \mathbf{b}$ 皆有解和 $\mathrm{Span}(\mathbf{a}_1, \ldots, \mathbf{a}_n) = \mathbb{R}^m$ 是等價的.

另外我們可以考慮聯立方程組 $A\mathbf{x} = \mathbf{e}_i$, 其中 $\mathbf{e}_1, \ldots, \mathbf{e}_m \in \mathbb{R}^m$ 為 $\mathbb{R}^m$ 的 standard basis. 若已知對任意的 $\mathbf{b} \in \mathbb{R}^m$, 聯立方程組 $A\mathbf{x} = \mathbf{b}$ 皆有解, 則對所有的 $i = 1, \ldots, m$, 我們都可找到 $\mathbf{c}_i \in \mathbb{R}^n$ 使得 $\mathbf{x} = \mathbf{c}_i$ 為聯立方程組 $A\mathbf{x} = \mathbf{e}_i$ 的一組解. 也就是說對所有的 $i = 1, \ldots, m$ 皆有 $A\mathbf{c}_i = \mathbf{e}_i$. 現考慮 $n \times m$ matrix C, 其 i-th column 就是 $\mathbf{c}_i$. 此時依矩陣乘法的定義我們有

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

也就是說, 此時必存在 $n \times m$ matrix C 使得 $AC = I_m$. 反之, 若 C 為 $n \times m$ matrix 滿足 $AC = I_m$, 則對任意 $\mathbf{b} \in \mathbb{R}^m$, 我們考慮 $\mathbf{c} = C\mathbf{b} \in \mathbb{R}^n$, 皆會有

$$A\mathbf{c} = A(C\mathbf{b}) = (AC)\mathbf{b} = I_m\mathbf{b} = \mathbf{b}.$$

也就是說此時對任意 $\mathbf{b} \in \mathbb{R}^m$, 我們都可以找到 $\mathbf{c} = C\mathbf{b} \in \mathbb{R}^n$, 使得 $\mathbf{x} = \mathbf{c}$ 是聯立方程組 $A\mathbf{x} = \mathbf{b}$ 的一組解. 因此從這觀點來看, 對任意的 $\mathbf{b} \in \mathbb{R}^m$, 聯立方程組 $A\mathbf{x} = \mathbf{b}$ 皆有解和存在 $n \times m$ matrix C 使得 $AC = I_m$ 是等價的.

我們曾探討過, 若 A 經由 elementary row operations 化為 echelon form 後, 其 pivot 的個數恰等於 A 的 row 的個數 m, 表示 A 的 echelon form 沒有一個 row 全為 0, 故由 1.2 節的討論 (即 Case (1)) 知此時任意的 $\mathbf{b} \in \mathbb{R}^m$, 聯立方程組 $A\mathbf{x} = \mathbf{b}$ 皆有解. 反之, 如果 pivot

的個數不等於 m, 表示 A 的 echelon form A' 中最後一個 row 必全為 0. 此時我們一定可以找到 $\mathbf{b} \in \mathbb{R}^m$ 使得增廣矩陣 $[A|\mathbf{b}]$ 化為 echelon form $[A'|\mathbf{b}']$ 後, $\mathbf{b}'$ 最後一個 entry 不為 0 (即 Case 2(a)). 此時 $A\mathbf{x} = \mathbf{b}$ 會無解. 因此從這觀點來看, 對任意的 $\mathbf{b} \in \mathbb{R}^m$, 聯立方程組 $A\mathbf{x} = \mathbf{b}$ 皆有解和 A 的 echelon form 的 pivot 的個數為 m (即 $\text{rank}(A) = m$) 是等價的.

綜合上面這幾種看法, 我們證得了以下這個非常重要的定理.

Theorem 2.5.2. *假設 $A \in M_{m\times n}$, 令 $\mathbf{a}_1, \ldots, \mathbf{a}_n \in \mathbb{R}^m$ 為 A 的 column vectors. 以下各敘述是等價的.*

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

(2) *$\text{Span}(\mathbf{a}_1, \ldots, \mathbf{a}_n) = \mathbb{R}^m$.*

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

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

特別提醒一下, Theorem 2.5.2, 指的是對所有 $\mathbf{b} \in \mathbb{R}^m$, 聯立方程組 $A\mathbf{x} = \mathbf{b}$ 皆有解的情況. 所以若僅知單一的 $\mathbf{b}$ 使得聯立方程組 $A\mathbf{x} = \mathbf{b}$ 有解, Theorem 2.5.2 並不適用 (不過 Lemma 2.5.1 是適用的).

我們曾提及, 當 $A \in M_{m\times n}$, 將 A 化為 echelon form 後其 pivot 的個數不可能多於 row 和 column 的個數. 也就是說 pivot 的個數應小於等於 $\min\{m,n\}$ (此指的是 m,n 中最小的那一個). 所以若 pivot 的個數為 m, 則表示 $n \geq m$. 換言之, 若 $n < m$, 我們便知 pivot 的個數不可能等於 m, 所以 Theorem 2.5.2 中的情況不可能發生. 我們有以下的結論.

Corollary 2.5.3. *假設 $A \in M_{m\times n}$, 其中 $n < m$, 則必存在 $\mathbf{b} \in \mathbb{R}^m$ 使得聯立方程組 $A\mathbf{x} = \mathbf{b}$ 無解. 而且此時, 不會存在 $n \times m$ matrix C 使得 $AC = I_m$.*

Proof. 由前所述, 當 $n < m$ 時 A 化為 echelon form 後, 其 pivot 的個數不可能為 m, 亦即 $\text{rank}(A) < m$. 故由 Theorem 2.5.2 知不可能對任意的 $\mathbf{b} \in \mathbb{R}^m$, 聯立方程組 $A\mathbf{x} = \mathbf{b}$ 皆有解. 亦即存在 $\mathbf{b} \in \mathbb{R}^m$ 使得聯立方程組 $A\mathbf{x} = \mathbf{b}$ 無解. 同理, 由 Theorem 2.5.2 知不會存在 $n \times m$ matrix C 使得 $AC = I_m$. □

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

前面提過 Theorem 2.5.2 是個很重要的定理, 它可以告訴我們一些解聯立方程組的訊息. 例如 Corollary 2.5.3 就是告訴我們當方程式的個數多於未知數的個數時, 會存在 $\mathbf{b} \in \mathbb{R}^m$ 使得聯立方程組 $A\mathbf{x} = \mathbf{b}$ 無解.

Exercise 2.12. 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 $\mathbf{ba}$、$\mathbf{ab}$ 分別應為幾階矩陣? 它們和 $\mathbf{a},\mathbf{b}$ 看成 $\mathbb{R}^n$ 上的向量後取內積有關嗎?(注意 $\mathbf{ba}$ 的 diagonal entries)

Exercise 2.13. 考慮 $A = [a_{ij}] \in M_{3\times 4}$, $B = [b_{ij}] \in M_{4\times 3}$ 其中 $a_{ij} = i + j$, $b_{ij} = i \cdot j$。(參見 Exercise 2.1, Exercise 2.6) 請利用 outer product 的方法將 AB 和 BA 分別寫成一些 rank 1 的矩陣相加。

Exercise 2.14. 令 $A = \begin{bmatrix} 1 & 0 & 1 & 0 & 1 \\ 1 & 1 & 0 & 0 & 2 \\ 3 & 1 & 1 & 1 & 1 \\ 0 & 1 & 2 & 1 & 2 \end{bmatrix}$ 且令 A' 為 A 的 reduced echelon form. (參見 Exercise 2.11)

(1) 說明對任意 $\mathbf{b} \in \mathbb{R}^k$ 聯立方程組 $A\mathbf{x} = \mathbf{b}$ 皆在 $\mathbf{x} \in \mathbb{R}^\ell$ 有解, 其中 k,ℓ 為何?

(2) 試找到無窮多個 C 滿足 $AC = I_4$。並說明這些 C 皆滿足 $A'C = E$, 其中 E 為 Exercise 2.11 (1) 小題的 E, 並依此說明這個 E 是唯一的。

(3) 試說明有無窮多個 matrix C' 使得 $C'A^{\mathrm{t}} = I_4$.

Exercise 2.15. 給定一 $m \times n$ matrix A, 其中 $m \neq n$. 假設 B 為 $n \times m$ matrix 滿足 $AB = I_m$.

(1) 說明 $m < n$, 並依此說明對任意 $\mathbf{b} \in \mathbb{R}^m$, 聯立方程組 $A\mathbf{x} = \mathbf{b}$ 皆存在無窮多組解。

(2) 試說明對任意 $k \in \mathbb{N}$, 皆存在無窮多 $n \times k$ 的矩陣 C 使得 AC 為 $m \times k$ 的零矩陣。

(3) 試說明對任意 $k \in \mathbb{N}$ 僅有唯一 $k \times m$ 的矩陣 D 使得 DA 為 $k \times n$ 的零矩陣。

(4) 試說明 $\mathrm{rank}(B^{\mathrm{t}})$ 為何?