

大學線性代數初步

李華介

國立台灣師範大學數學系

前言

本講義主要目的是針對大一學生介紹有關線性代數基本的理論。希望藉由較平易近人的方式介紹線性代數也慢慢引導熟悉數學的論證方式。本講義希望以淺顯易懂為主旨，而不是生動有趣。畢竟有些事情要說明清楚就會顯得囉唆，當然就不有趣了。

本講義是內容僅涵蓋實係數向量空間的基礎理論，希望大家對這些基本理論有了充分了解，在去談其他。因此內容並不足以應付研究所課程所需的進一步理論。另一方面我們的內容偏重於整個線性代數的理論架構，對於線性代數在其他領域的應用幾乎沒有著墨。我們依循一貫的原則就是基本理論清楚了，接下來的應用或推廣就不難了。所以對線性代數更進一步的理論或應用有興趣的同學應再參考其他的參考書籍。

本講義雖然主要以中文撰寫，不過當涉及定義或專有名詞時，為免翻譯的困擾將以英文取代。因此將以中英夾雜較不傳統的方式顯現，若有不便請見諒。

本講義編寫費時，編寫完後並未經過嚴謹的校對。疏漏在所難免，雖不至於有理論性上嚴重的錯誤，但讀者仍應注意不宜概括全收。若發現錯誤，歡迎提出寶貴的意見。

Part I

**Vectors, Systems of
Linear Equations,
Matrices**

在高中階段大家接受到與線性代數相關的數學，有向量 (vectors)，一次聯立方程式 (systems of linear equations) 以及矩陣 (matrices)。在本講義的第一部份，我們想利用這三個大家較為熟悉的題材慢慢引入線性代數的概念。在此要強調一下，千萬不要以為這三種題材在高中沒學好，線性代數就學不好。我們僅是利用這些大家可能較熟悉的題材當成動機，讓大家剛接受到抽象的線性代數理論時不會絕得太突兀摸不著頭緒。

Vectors

我們藉由大家熟悉的向量來介紹向量間運算的性質。要注意我們不再去定義何謂向量，而是著重於如何用這些已知的向量的性質，利用較抽象的方法來推導出一些幾何的性質。

1.1. 二維平面中的向量

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。也因此本節在細節的說明特別繁瑣。若對坐標平面的向量相當清楚的同學，可放心略過此節。

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 屬於實數，也就是說 a, b 皆為實數)。意思就是說如果你在坐標平面中任給一點 P ，然後從 P 點開始往水平方向走 a 單位 ($a > 0$ 時往右； $a < 0$ 時往左)，再沿鉛直方向走 b 單位 ($b > 0$ 時往上； $b < 0$ 時往下)，最後到達的點若記為 Q 。那麼從 P 點開始到 Q 點為止的這一個向量就可用 (a, b) 來表示，記為 $\overrightarrow{PQ} = (a, b)$ 。

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 $b = d$ ；從這觀點若用點來表示向量時就較麻煩，因為如果 P, P', Q, Q' 為平面中四個點 $\overrightarrow{PQ} = \overrightarrow{P'Q'}$ 並不代表 P 和 P' 為同一點且 Q 和 Q' 為同一點。不過若已知 $P = P'$ ，則可得 $Q = Q'$ 。反之若 $Q = Q'$ 則可得 $P = P'$ 。(我們說明一下這裡符號的使用：因為 P, P' 皆為“點”而不是“數”，所以這裡 $P = P'$ 這個等號表示同一“點”而不是同一“數”)。上面這個論述可用剛才定義 \overrightarrow{PQ} 的方法驗證，這裡就不再驗證。基於符號的方便性，當我們要用符號來表示一個向量時，除非已知此向量為特定兩點所決定的向量，通常會僅用 \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 $\mathbf{v} = (a, b)$ 。

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 1.1.1. 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

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

要注意這裡所謂的定義 (即 definition) 指的是規定, 也就是說我們規定向量必須這樣相加及乘上一個常數. 當然了你也可以自行規定一套向量加法的規則. 不過一般在數學上的定義都有其必要性, 通常一個定義對理論推導或實用上都會有相當的幫助. 例如這一個定義與我們直觀上認為若 P, Q, R 為坐標平面上三點, 則 $\overrightarrow{PQ} + \overrightarrow{QR} = \overrightarrow{PR}$ 相吻合. 這是因為: 假設 P, Q, R 三點的坐標分別為 $P(x_1, y_1), Q(x_2, y_2), R(x_3, y_3)$ (注意: 為了區分點和向量的坐標表示法, 當我們提到點的坐標時都會加上該點的代號), 則

$$\overrightarrow{PQ} = (x_2 - x_1, y_2 - y_1), \overrightarrow{QR} = (x_3 - x_2, y_3 - y_2), \overrightarrow{PR} = (x_3 - x_1, y_3 - y_1).$$

然而依前面向量加法的定義確實

$$\overrightarrow{PQ} + \overrightarrow{QR} = ((x_2 - x_1) + (x_3 - x_2), (y_2 - y_1) + (y_3 - y_2)) = (x_3 - x_1, y_3 - y_1).$$

故得

$$\overrightarrow{PQ} + \overrightarrow{QR} = \overrightarrow{PR}. \quad (1.1)$$

一個定義一定要清楚明確, 例如若 $\mathbf{u} = (1, 2), \mathbf{v} = (3, 4)$ 且 $r = 5$, 由 Definition 1.1.1 我們知 $\mathbf{u} + \mathbf{v} = (1 + 3, 2 + 4) = (4, 6)$ 且 $r\mathbf{u} = (5 \times 1, 5 \times 2) = (5, 10)$. 不過我們決不能用: 若 $\mathbf{u} = (1, 2), \mathbf{v} = (3, 4)$ 且 $r = 5$, 則定義 $\mathbf{u} + \mathbf{v} = (1 + 3, 2 + 4) = (4, 6)$ 且 $r\mathbf{u} = (5 \times 1, 5 \times 2) = (5, 10)$ 這樣的說法來定義向量的加法及係數積. 因為這樣的說法只是定義出兩個特殊的向量的加法以及和一特殊實數的乘積, 不能依這特例要求別人“依此類推”. 這就是當初定義時用 a_1, a_2, b_1, b_2, r 這些符號來代替具體數字的用意. 另外要注意在 Definition 1.1.1 中我們沒有提 $r\mathbf{v}$ 的定義, 不過當初既然 \mathbf{u} 已表成任意 \mathbb{R}^2 上的向量, 定義 $r\mathbf{u}$ 已足夠所以不必再去定 $r\mathbf{v}$. 最後要強調的是: 這裡我們並沒有定義兩向量的乘法, 而是定義實數和向量相乘, 所以我們稱為係數積 (scalar multiplication) 而不能說是向量的乘法.

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, 我們可依定義推導出一些性質, 利用這些性質簡化處理程序. 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 1.1.2. 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 $r, s \in \mathbb{R}$ 以及 $\mathbf{u} \in \mathbb{R}^2$, 皆有 $r(s\mathbf{u}) = (rs)\mathbf{u}$.
- (6) 對任意 $r, s \in \mathbb{R}$ 以及 $\mathbf{u} \in \mathbb{R}^2$, 皆有 $(r+s)\mathbf{u} = r\mathbf{u} + s\mathbf{u}$.
- (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 $\mathbf{u} \in \mathbb{R}^2$, 皆有 $1\mathbf{u} = \mathbf{u}$.

通常一個定理敘述完就要證明, 不過在這裡我們建議先緩一緩. 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. 這裡大家會覺得自然是因為大家對實數的加法運算很清楚, 不過數學上是存在許多“抽象”的數系它的加法是不能交換的. 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. 這裡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),(6),(7) 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 . 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.

(8) 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 (5) 的性質兩邊乘上 $1/2$, 得

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

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。現在我們就來看看這些證明。

Proof. (of Proposition 1.1.2) 這幾個性質其實很簡單，我們寫下證明是希望不熟悉寫抽象證明的同學利用這個簡單的證明學學看如何寫好證明。若自覺對這些性質的證明清楚的同學可跳過此證明。

- (1) 假設 $\mathbf{u} = (a_1, a_2), \mathbf{v} = (b_1, b_2)$ ，則依定義知

$$\mathbf{u} + \mathbf{v} = (a_1 + b_1, a_2 + b_2), \mathbf{v} + \mathbf{u} = (b_1 + a_1, b_2 + a_2).$$

故由實數加法的交換性 (即 $a + b = b + a$) 可得 $\mathbf{u} + \mathbf{v} = \mathbf{v} + \mathbf{u}$ 。

- (2) 假設 $\mathbf{u} = (a_1, a_2), \mathbf{v} = (b_1, b_2), \mathbf{w} = (c_1, c_2)$ ，則依定義知 $\mathbf{u} + \mathbf{v} = (a_1 + b_1, a_2 + b_2)$ 故得

$$(\mathbf{u} + \mathbf{v}) + \mathbf{w} = ((a_1 + b_1) + c_1, (a_2 + b_2) + c_2).$$

同理由 $\mathbf{v} + \mathbf{w} = (b_1 + c_1, b_2 + c_2)$ 可得

$$\mathbf{u} + (\mathbf{v} + \mathbf{w}) = (a_1 + (b_1 + c_1), a_2 + (b_2 + c_2)).$$

因此由實數加法的結合律 (即 $(a + b) + c = a + (b + c)$) 得知 $(\mathbf{u} + \mathbf{v}) + \mathbf{w} = \mathbf{v} + (\mathbf{u} + \mathbf{w})$ 。

- (3) 這是一個存在性的問題，也就是說要找到一個向量 $\mathbf{0}$ 滿足所求。這裡我們只要令 $\mathbf{0} = (0, 0)$ ，則對任意向量 $\mathbf{u} = (a_1, a_2)$ 皆有

$$\mathbf{0} + \mathbf{u} = (0 + a_1, 0 + a_2) = (a_1, a_2) = \mathbf{u}.$$

故知確實存在這樣的向量。

- (4) 上面我們已知可令 $\mathbf{0} = (0, 0)$ ，故此時對任意 $\mathbf{u} = (a_1, a_2)$ 我們只要考慮 $\mathbf{u}' = (-a_1, -a_2)$ ，則可得

$$\mathbf{u} + \mathbf{u}' = (a_1 + (-a_1), a_2 + (-a_2)) = (0, 0) = \mathbf{0}.$$

- (5) 假設 $\mathbf{u} = (a_1, a_2)$ ，依定義知 $s\mathbf{u} = (sa_1, sa_2)$ ，因此得

$$r(s\mathbf{u}) = (r(sa_1), r(sa_2)).$$

另一方面

$$(rs)\mathbf{u} = ((rs)a_1, (rs)a_2),$$

故由實數乘法結合律 (即 $r(sa) = (rs)a$) 可得 $r(s\mathbf{u}) = (rs)\mathbf{u}$ 。

- (6) 假設 $\mathbf{u} = (a_1, a_2)$ ，依定義知

$$(r + s)\mathbf{u} = ((r + s)a_1, (r + s)a_2).$$

另一方面 $r\mathbf{u} = (ra_1, ra_2), s\mathbf{u} = (sa_1, sa_2)$ ，可得

$$r\mathbf{u} + s\mathbf{u} = (ra_1 + sa_1, ra_2 + sa_2),$$

故由實數加法與乘法的分配律 (即 $(r + s)a = ra + sa$) 可得 $(r + s)\mathbf{u} = r\mathbf{u} + s\mathbf{u}$ 。

(7) 假設 $\mathbf{u} = (a_1, a_2), \mathbf{v} = (b_1, b_2)$, 依定義知

$$r(\mathbf{u} + \mathbf{v}) = (r(a_1 + b_1), r(a_2 + b_2)).$$

另一方面 $r\mathbf{u} = (ra_1, ra_2), r\mathbf{v} = (rb_1, rb_2)$, 可得

$$r\mathbf{u} + r\mathbf{v} = (ra_1 + rb_1, ra_2 + rb_2),$$

故由實數加法與乘法的分配律 (即 $r(a + b) = ra + rb$) 可得 $r(\mathbf{u} + \mathbf{v}) = r\mathbf{u} + r\mathbf{v}$.

(8) 假設 $\mathbf{u} = (a_1, a_2)$, 由於對任意實數 a 皆有 $1a = a$, 故由 $1\mathbf{u} = (1a_1, 1a_2)$ 可得 $1\mathbf{u} = \mathbf{u}$.

□

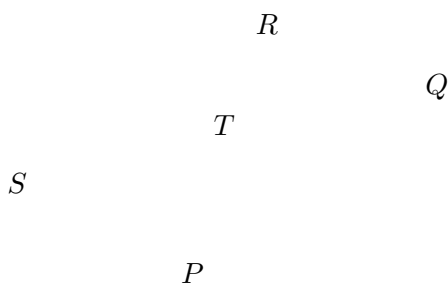
接下來我們將舉一個例子來說明我們可以用向量的性質 (Proposition 1.1.2) 來處理一些幾何的問題。由於這些性質以後我們不會用到, 所以在此不用定理的形式呈現而是用例題的形式呈現。

Example 1.1.3. 假設 P, Q, R, S 是平面中四個點, 其中 $\overrightarrow{PQ} = \overrightarrow{SR}$, 則我們有以下性質:

(1) $\overrightarrow{QR} = \overrightarrow{PS}$.

(2) 假設 T 為線段 \overline{PR} 和線段 \overline{SQ} 的交點, 則 $\overrightarrow{PT} = \overrightarrow{TR}$ 且 $\overrightarrow{ST} = \overrightarrow{TQ}$.

我們利用下圖來解釋這兩個性質會成立的原因:



(1) 我們要利用 $\overrightarrow{PQ} = \overrightarrow{SR}$ 來得到 $\overrightarrow{QR} = \overrightarrow{PS}$. 由前面式子 (1.1) 我們知道 $\overrightarrow{PQ} + \overrightarrow{QR} = \overrightarrow{PR}$ 且 $\overrightarrow{PS} + \overrightarrow{SR} = \overrightarrow{PR}$. 也就是說

$$\overrightarrow{PQ} + \overrightarrow{QR} = \overrightarrow{PS} + \overrightarrow{SR}. \quad (1.2)$$

由 Proposition 1.1.2 (4), 我們知存在 \mathbf{u} 使得 $\overrightarrow{PQ} + \mathbf{u} = \mathbf{0}$. 要注意由於 Proposition 1.1.2 (1) 以及 $\overrightarrow{PQ} = \overrightarrow{SR}$ 的假設, 這等於

$$\overrightarrow{PQ} + \mathbf{u} = \mathbf{u} + \overrightarrow{PQ} = \overrightarrow{SR} + \mathbf{u} = \mathbf{u} + \overrightarrow{SR} = \mathbf{0}$$

現將上面式子 (1.2) 兩邊加上 \mathbf{u} 可得

$$\mathbf{u} + (\overrightarrow{PQ} + \overrightarrow{QR}) = \mathbf{u} + (\overrightarrow{PS} + \overrightarrow{SR}) \quad (1.3)$$

然而利用 Proposition 1.1.2 (2), (3) 我們有

$$\mathbf{u} + (\overrightarrow{PQ} + \overrightarrow{QR}) = (\mathbf{u} + \overrightarrow{PQ}) + \overrightarrow{QR} = \mathbf{0} + \overrightarrow{QR} = \overrightarrow{QR}.$$

同理可得

$$\mathbf{u} + (\overrightarrow{PS} + \overrightarrow{SR}) = (\overrightarrow{PS} + \overrightarrow{SR}) + \mathbf{u} = \overrightarrow{PS} + (\overrightarrow{SR} + \mathbf{u}) = \overrightarrow{PS} + \mathbf{0} = \overrightarrow{PS}.$$

故由式子 (1.3) 可得 $\overrightarrow{QR} = \overrightarrow{PS}$.

- (2) 通常我們用坐標來處理問題時，碰到交點的問題可以用解方程式的方法處理，不過這裡我們要用抽象的向量處理，遇到交點問題就比較麻煩，需要用比較特殊的方法處理。這裡原要證明線段 \overline{PR} 和線段 \overline{SQ} 的交點會是 \overline{PR} 以及 \overline{SQ} 的中點。我們用反過來的看法處理，也就是要去說明 \overline{PR} 的中點和 \overline{SQ} 的中點會是同一點。如此一來這一點因同時在 \overline{PR} 和 \overline{SQ} 上，所以自然就是 \overline{PR} 和 \overline{SQ} 的交點。又因為 \overline{PR} 和 \overline{SQ} 不平行，僅能有一個交點，這就說明了線段 \overline{PR} 和線段 \overline{SQ} 的交點就是 \overline{PR} 以及 \overline{SQ} 的中點。具體來說我們假設 T', T'' 分別為 \overline{PR} 和 \overline{SQ} 的中點。然後說明 $T' = T''$ ，如此一來便得證 T' 為 \overline{PR} 和 \overline{SQ} 的交點，也就是說 $T = T' = T''$ ，故得證所求。

依假設 T', T'' 分別為 \overline{PR} 和 \overline{SQ} 的中點，故知

$$\overrightarrow{PT'} = \overrightarrow{T'R} = \frac{1}{2}\overrightarrow{PR} \quad (1.4)$$

且

$$\overrightarrow{ST''} = \overrightarrow{T''Q} = \frac{1}{2}\overrightarrow{SQ}. \quad (1.5)$$

若我們能證明 $\overrightarrow{PT'} = \overrightarrow{PT''}$ 便得證 $T' = T''$ 。然而 $\overrightarrow{PT''} = \overrightarrow{PS} + \overrightarrow{ST''}$ ，故由式子 (1.5) 可得

$$\overrightarrow{PT''} = \overrightarrow{PS} + \overrightarrow{ST''} = \overrightarrow{PS} + \frac{1}{2}\overrightarrow{SQ}.$$

又由於 (1) 告訴我們 $\overrightarrow{PS} = \overrightarrow{QR}$ 上式可改寫為

$$\overrightarrow{PT''} = \overrightarrow{QR} + \frac{1}{2}\overrightarrow{SQ} = \left(\frac{1}{2}\overrightarrow{QR} + \frac{1}{2}\overrightarrow{QR}\right) + \overrightarrow{SQ} = \frac{1}{2}\overrightarrow{QR} + \frac{1}{2}(\overrightarrow{QR} + \overrightarrow{SQ}) = \frac{1}{2}(\overrightarrow{QR} + \overrightarrow{SR})$$

再利用已知 $\overrightarrow{PQ} = \overrightarrow{SR}$ ，上式又可改寫為

$$\overrightarrow{PT''} = \frac{1}{2}(\overrightarrow{QR} + \overrightarrow{SR}) = \frac{1}{2}(\overrightarrow{QR} + \overrightarrow{PQ}) = \frac{1}{2}\overrightarrow{PR}$$

故由式子 (1.4) 得證 $\overrightarrow{PT'} = \overrightarrow{PT''}$ ，也因此得知 $T' = T''$ 就是 \overline{PR} 和 \overline{SQ} 的交點，也就是說 $T = T' = T''$ 。故由式子 (1.4, 1.5) 得證 $\overrightarrow{PT} = \overrightarrow{TR}$ 且 $\overrightarrow{ST} = \overrightarrow{TQ}$ 。

同學或許會想到用設定坐標的方式處理 Example 1.1.3 的問題。這裡我們故意不去設定坐標系，而僅用 Proposition 1.1.2 中列出的向量運算性質去處理，主要就是要強調這幾個運算性質就足以處理向量有關的性質。也就是說不一定需要架設坐標系，只要符合 Proposition 1.1.2 中列出的運算性質，都可享有 Example 1.1.3 的性質。另外由 Example 1.1.3 的處理過程中我們了解到，Proposition 1.1.2 的運算性質可幫助我們在處理向量有關的等式運算時可如處理實數的等式運算一樣（例如移項，消去...等）。這些等我們以後更進一步談向量空間時，大家就更能體會了。

1.2. \mathbb{R}^n 中的向量

在高中時我們也學坐標空間中的向量，也就是 \mathbb{R}^3 。很容易理解 \mathbb{R}^3 中的向量可以說是 \mathbb{R}^2 中的向量的推廣。同樣的我們也可將之推廣而定義 \mathbb{R}^n 中的向量，其中 $n \in \mathbb{N}$ 是任意的正整數。本節中我們將探討 \mathbb{R}^n 中的向量。或許同學們會覺得 \mathbb{R}^n 中的向量已看不到，而疑惑為何要探討它。事實上線性代數的應用有許多情況就是在這類抽象且看不到的狀況，這甚至可以說是線性代數發展的主因（如果僅為了 \mathbb{R}^2 或 \mathbb{R}^3 就沒必要發展這一套理論）。在這一節中我們會發現，雖然它看不到，但因為具有如 \mathbb{R}^2 中向量的運算性質，我們還是可以如處理 \mathbb{R}^2 中的向量一般的方式處理它們相關的性質。

\mathbb{R}^3 中的向量，即坐標空間中的向量我們可以用 (a_1, a_2, a_3) 其中 $a_1, a_2, a_3 \in \mathbb{R}$ 來表示，而且我們定義兩向量 $(a_1, a_2, a_3), (b_1, b_2, b_3)$ 相等表示 $a_1 = b_1, a_2 = b_2$ 且 $a_3 = b_3$ 。對於任意的 $n \in \mathbb{N}$ 我們也有如下的定義：

Definition 1.2.1. 給定任意 $n \in \mathbb{N}$ ，我們定義 \mathbb{R}^n 中的向量為 (a_1, \dots, a_n) ，其中 $a_1, \dots, a_n \in \mathbb{R}$ 。我們說兩向量 (a_1, \dots, a_n) 以及 (b_1, \dots, b_n) 相等若且唯若 $a_1 = b_1, \dots, a_n = b_n$ 。

我們多說明一下符號的表示法，這裡 (a_1, \dots, a_n) 表示有 n 個位置，每個位置我們依次填入實數其中第 1 個位置的元素我們用 a_1 來表示，第二個位置用 a_2 來表示，這樣一直下去直到第 n 個位置用 a_n 來表示。例如 $n = 4$ 時， \mathbb{R}^4 中的向量可以用 (a_1, a_2, a_3, a_4) 來表示。因為 n 可以是任意的正整數不能如 $n = 4$ 時將所有的 a_1, a_2, \dots 都列出來，所以我們就用 (a_1, \dots, a_n) 來表示。這裡我們也沿習這樣省略的方法說 $a_1, \dots, a_n \in \mathbb{R}$ ，表示 a_1 到 a_n 這 n 個元素都是實數。有的書用較嚴謹的說法，會用 $a_i \in \mathbb{R}, \forall 1 \leq i \leq n$ 來表示，此即對所有的 $1 \leq i \leq n$ 皆有 $a_i \in \mathbb{R}$ 的意思。這種說法的意義是： i 是任意 1 到 n 的整數，而對於這個 i 所對應的 a_i 會是實數。所以這等同於說 a_1 到 a_n 這 n 個數都是實數。另一方面談論向量的相等是必要的，這是因為在談論向量的運算時就如同實數的運算，我們必須明確規定等式的意義。實數的相等很明確，所以我們就利用實數的相等來定義向量的相等。這裡 $a_1 = b_1, \dots, a_n = b_n$ 也是一種省略的說法，一般也可以用 $a_i = b_i, \forall 1 \leq i \leq n$ 來表示。

接著我們沿用 \mathbb{R}^2 中向量的加法及係數積來定義 \mathbb{R}^n 中向量的加法 (*addition*) 以及係數積 (*scalar multiplication*)。

Definition 1.2.2. 令 $\mathbf{u} = (a_1, \dots, a_n), \mathbf{v} = (b_1, \dots, b_n) \in \mathbb{R}^n$ 以及 $r \in \mathbb{R}$ 。我們定義

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

依此定義，給定 \mathbb{R}^n 中的兩向量我們可以明確地計算它們之和，例如若 $\mathbf{u} = (1, 1, 2, 2)$ ， $\mathbf{v} = (5, 4, 3, 2) \in \mathbb{R}^4$ 則

$$\mathbf{u} + \mathbf{v} = (1 + 5, 1 + 4, 2 + 3, 2 + 2) = (6, 5, 5, 4).$$

這裡特別要注意的是必須是 \mathbf{u}, \mathbf{v} 皆在相同的 \mathbb{R}^n 中才能定義 $\mathbf{u} + \mathbf{v}$ 。例如若 \mathbf{u} 在 \mathbb{R}^3 而 \mathbf{v} 在 \mathbb{R}^4 是不能談 $\mathbf{u} + \mathbf{v}$ 的了！

如同 \mathbb{R}^2 的情形, \mathbb{R}^n 中向量加法及係數積有以下的性質. 由於這些性質的證明和 \mathbb{R}^2 的情形完全相同 (用到實數相對應的性質), 此處我們就不再證明了.

Proposition 1.2.3. 對於 \mathbb{R}^n 上的向量, 我們有以下的性質:

- (1) 對任意 $\mathbf{u}, \mathbf{v} \in \mathbb{R}^n$, 皆有 $\mathbf{u} + \mathbf{v} = \mathbf{v} + \mathbf{u}$.
- (2) 對任意 $\mathbf{u}, \mathbf{v}, \mathbf{w} \in \mathbb{R}^n$, 皆有 $(\mathbf{u} + \mathbf{v}) + \mathbf{w} = \mathbf{u} + (\mathbf{v} + \mathbf{w})$.
- (3) 存在一向量 $\mathbf{0} \in \mathbb{R}^n$ 滿足對任意 $\mathbf{u} \in \mathbb{R}^n$ 皆有 $\mathbf{0} + \mathbf{u} = \mathbf{u}$.
- (4) 對任意 $\mathbf{u} \in \mathbb{R}^n$ 皆可找到 $\mathbf{u}' \in \mathbb{R}^n$ 滿足 $\mathbf{u} + \mathbf{u}' = \mathbf{0}$.
- (5) 對任意 $r, s \in \mathbb{R}$ 以及 $\mathbf{u} \in \mathbb{R}^n$, 皆有 $r(s\mathbf{u}) = (rs)\mathbf{u}$.
- (6) 對任意 $r, s \in \mathbb{R}$ 以及 $\mathbf{u} \in \mathbb{R}^n$, 皆有 $(r + s)\mathbf{u} = r\mathbf{u} + s\mathbf{u}$.
- (7) 對任意 $r \in \mathbb{R}$ 以及 $\mathbf{u}, \mathbf{v} \in \mathbb{R}^n$ 皆有 $r(\mathbf{u} + \mathbf{v}) = r\mathbf{u} + r\mathbf{v}$.
- (8) 對任意 $\mathbf{u} \in \mathbb{R}^n$, 皆有 $1\mathbf{u} = \mathbf{u}$.

Proposition 1.2.3 (3) 所提的 $\mathbf{0}$ 其實就是向量 (a_1, \dots, a_n) , 其中 $a_i = 0, \forall 1 \leq i \leq n$, 我們也稱此 $\mathbf{0}$ 為零向量. 另外若 $\mathbf{u} = (a_1, \dots, a_n) \in \mathbb{R}^n$, 則 Proposition 1.2.3 (4) 所提的 \mathbf{u}' 其實就是 (b_1, \dots, b_n) , 其中 $b_i = -a_i, \forall 1 \leq i \leq n$. 由係數積的定義我們可得

$$\mathbf{u}' = (-a_1, \dots, -a_n) = (-1)(a_1, \dots, a_n) = (-1)\mathbf{u}.$$

因此我們又將此 \mathbf{u}' 簡記成 $-\mathbf{u}$. 有了這些性質, 我們可以將 \mathbb{R}^n 的向量如實數般運算. 例如, 若 $\mathbf{v} + \mathbf{u} = \mathbf{u}$, 則等式兩邊加上 $-\mathbf{u}$, 可得

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

從這裡我們知道只要 $\mathbf{v} + \mathbf{u} = \mathbf{u}$, 就能確保 $\mathbf{v} = \mathbf{0}$. 換言之, $\mathbf{0}$ 會是在 \mathbb{R}^n 中“唯一”的一個向量會滿足 Proposition 1.2.3 (3) 中所說 $\mathbf{0} + \mathbf{u} = \mathbf{u}$. 又例如若 $\mathbf{u} + \mathbf{v} = \mathbf{0}$, 則等式兩邊加上 $-\mathbf{u}$ 可得

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

也就是說給定 $\mathbf{u} \in \mathbb{R}^n$, $-\mathbf{u}$ 會是 \mathbb{R}^n 中唯一的向量會滿足 $\mathbf{u} + (-\mathbf{u}) = \mathbf{0}$. 綜合以上的結果, 我們有以下的性質.

Proposition 1.2.4. 假設 $\mathbf{u} = (a_1, \dots, a_n) \in \mathbb{R}^n$. 則我們有以下的性質:

- (1) $\mathbf{v} \in \mathbb{R}^n$ 滿足 $\mathbf{u} + \mathbf{v} = \mathbf{u}$ 若且唯若

$$\mathbf{v} = \mathbf{0} = (0, \dots, 0).$$

- (2) $\mathbf{w} \in \mathbb{R}^n$ 滿足 $\mathbf{u} + \mathbf{w} = \mathbf{0}$ 若且唯若

$$\mathbf{w} = -\mathbf{u} = (-1)(a_1, \dots, a_n) = (-a_1, \dots, -a_n).$$

Remark 1.2.5. Proposition 1.2.3 (3),(4) 其實談的是零向量和反向量的存在性而 Proposition 1.2.4 (1),(2) 談的是它們的唯一性. 同學或許會奇怪為何要將它們分開敘述呢? 這裡我們雖然用 \mathbb{R}^n 的加法及係數積的定義得到 Proposition 1.2.4, 事實上此 Proposition

1.2.4 是可以直接用 Proposition 1.2.3 得到的 (以後我們在抽象的向量空間中就會這麼處理). 就像我們在前一節所述, Proposition 1.2.3 中所述的這些性質就足以推導出一般向量有關的性質, 這些等到我們用抽象的方法定義向量空間時就會清楚, 這裡就不再詳談.

既然 \mathbb{R}^n 中向量的運算都可像實數一般處理, 我們可以如實數一樣引用“減法”的符號, 即將 $\mathbf{w} + (-\mathbf{v})$ 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 $\mathbf{u} = 1/2(\mathbf{w} - \mathbf{v})$.

1.3. Span of Vectors

在 \mathbb{R}^n 中有許多由 \mathbb{R}^n 中的向量展成的“子空間”. 這些子空間是這一節要探討的課題. 所謂子空間以後我們後有更正式的定義, 這裡我們僅暫時借用這個名詞, 大略指的是 \mathbb{R}^n 中一些向量所組成的特殊子集合 (subset).

我們先從熟悉的坐標平面開始. 在坐標平面中我們最常談的就是直線. 一個直線通常有很多方式來表示, 若用通俗一點的說法一個直線可以說是通過一特定點且沿著一特定方向前進(或後退)所得的點所成的集合. 如果這個特定點用 P 來表示, 而 Q 為此直線上另外一點, 則用向量的觀點來說就是此直線是通過 P 點且沿著 \overrightarrow{PQ} 方向的直線. 這個向量 \overrightarrow{PQ} 我們稱為此直線的“方向向量” (directional vector). 一個直線的方向向量並不唯一, 因為若 Q' 為此直線上 P, Q 以外的另一點, 則 $\overrightarrow{PQ'}$ 也會是此直線的一個方向向量. 不過因為 P, Q, Q' 皆在同一直線上, 依直線的定義我們知道會存在一實數 r 使得 $\overrightarrow{PQ'} = r\overrightarrow{PQ}$. 因此若已知一非零向量 \mathbf{v} 為一直線的方向向量, 則此直線上任兩點 P, Q 所成的向量都會存在一實數 r 使得 $\overrightarrow{PQ} = r\mathbf{v}$. 反之若已知一直線 L 通過 P 點且其方向向量為 \mathbf{v} , 若 Q 點滿足 $\overrightarrow{PQ} = r\mathbf{v}$, 其中 $r \in \mathbb{R}$, 則知 Q 點會在此直線 L 上. 因此我們可以用以下的集合來表示通過 P 點且方向向量為 \mathbf{v} 的直線 L 上的點, 即

$$L = \{Q \mid \overrightarrow{PQ} = r\mathbf{v}, r \in \mathbb{R}\}.$$

通常在集合的表示法中, 如果無法用列舉的方式一一列舉此集合的元素時, 我們會用上面的方法來表示. 也就是在“ $\{$ ”的左邊寫下此集合元素的形式, 右邊寫下這些元素所需符合的性質. 在這裡左邊的 Q 表示此集合所組成的元素是像 Q 這樣的點, 而右邊 $\overrightarrow{PQ} = r\mathbf{v}, r \in \mathbb{R}$ 表示 Q 點需滿足 $\overrightarrow{PQ} = r\mathbf{v}$, 其中 r 為實數.

從上面可知若 \mathbf{v} 是直線 L 的方向向量那麼對任意非零實數 r , 若 $\mathbf{w} = r\mathbf{v}$, 則 \mathbf{w} 也是 L 的方向向量. 為了方便起見我們就稱 \mathbf{w} 和 \mathbf{v} 為平行的向量. 若兩直線它們的方向向量是平行的, 我們就稱此二直線相平行.

我們可以發現, 上面這種用向量來描述一直線上的點的方法, 不僅在坐標平面中適用, 在坐標空間甚至更高維度的空間皆適用. 所以在 \mathbb{R}^n 中若 \mathbf{v} 是一個非零向量,

$$\{\mathbf{w} \in \mathbb{R}^n \mid \mathbf{w} = r\mathbf{v}, r \in \mathbb{R}\}$$

這一類由向量所組成的集合便顯得重要。提醒一下，這個集合表示法 $|$ 的左邊是“向量”，表示是由向量所成的集合，和前面直線 L 的集合表法 $|$ 左邊是“點”表示是由點所成的集合有所不同，大家應區分清楚。

在坐標空間中除了直線另一個大家常探討的便是平面。同樣的一個平面也有許多表示法。和向量有關的最常見的就是法向量的表示法，不過關於法向量的看法因牽涉內積我們留待以後再說明，這裡我們依然沿用剛才直線的看法。也就是說每一個平面，我們都可以在其上找到兩個不平行的直線 L_1, L_2 ，而整個平面就是沿著其中一條直線 L_1 畫出與 L_2 平行的直線而得。換言之，若 L_1, L_2 分別為此平面上不平行的兩條直線，其方向向量分別為 $\mathbf{v}_1, \mathbf{v}_2$ 且相交於 P 點，則對 L_1 上任一點 Q 皆可找到 $r_1 \in \mathbb{R}$ 滿足 $\overrightarrow{PQ} = r_1 \mathbf{v}_1$ 。而通過 Q 點且與 L_2 平行的直線會在此平面上，亦即若 R 點會在此與 L_2 平行的直線上，則可找到 $r_2 \in \mathbb{R}$ 使得 $\overrightarrow{QR} = r_2 \mathbf{v}_2$ 。利用 $\overrightarrow{PR} = \overrightarrow{PQ} + \overrightarrow{QR}$ 我們可得

$$\overrightarrow{PR} = r_1 \mathbf{v}_1 + r_2 \mathbf{v}_2,$$

也就是說對於平面上任一點 R 皆可找到 $r_1, r_2 \in \mathbb{R}$ 滿足 $\overrightarrow{PR} = r_1 \mathbf{v}_1 + r_2 \mathbf{v}_2$ 。反之若 R 滿足 $\overrightarrow{PR} = r_1 \mathbf{v}_1 + r_2 \mathbf{v}_2$ ，我們也可知 R 會落在一個過 L_1 且與 L_2 平行的直線上，也就是說 R 會在此平面上。所以若一平面 H 通過 P 點且 $\mathbf{v}_1, \mathbf{v}_2$ 為 H 上兩條不平行的直線的方向向量，我們也可用以下的集合表示平面 H 上的點，即

$$H = \{R \mid \overrightarrow{PR} = r_1 \mathbf{v}_1 + r_2 \mathbf{v}_2, r_1, r_2 \in \mathbb{R}\}.$$

要注意表示 H 的這兩個向量 v_1, v_2 並不唯一，至於要怎樣的兩個向量可以同樣描述 H 這個平面，這裡我們暫不探討，留待以後我們學習更多線性代數理論時再探討。這裡我們將專注於 $\{\mathbf{w} \in \mathbb{R}^n \mid \mathbf{w} = r\mathbf{v}, r \in \mathbb{R}\}$ 以及 $\{\mathbf{w} \in \mathbb{R}^n \mid \mathbf{w} = r_1 \mathbf{v}_1 + r_2 \mathbf{v}_2, r_1, r_2 \in \mathbb{R}\}$ 這一類由向量所組成的集合。

上述有關平面的向量表示法，不僅在坐標空間上適用，我們可以將上述概念推廣至 \mathbb{R}^n 上且考慮更一般的情形。在 \mathbb{R}^n 上我們不只可談直線和平面，還有許多和直線平面有類似的特性的事物值得探討。我們自然引進以下的定義。

Definition 1.3.1. 令 $\mathbf{v}_1, \dots, \mathbf{v}_m \in \mathbb{R}^n$ ，若存在 $r_1, \dots, r_m \in \mathbb{R}$ 使得

$$\mathbf{w} = r_1 \mathbf{v}_1 + \dots + r_m \mathbf{v}_m,$$

則稱 \mathbf{w} 為 $\mathbf{v}_1, \dots, \mathbf{v}_m$ 的一個 *linear combination* (線性組合)。所有 $\mathbf{v}_1, \dots, \mathbf{v}_m$ 的 linear combinations 所組成的集合稱為它們的 *span* (展成的空間)，我們用 $\text{Span}(\mathbf{v}_1, \dots, \mathbf{v}_m)$ 來表示這個集合，也就是說

$$\text{Span}(\mathbf{v}_1, \dots, \mathbf{v}_m) = \{\mathbf{w} \in \mathbb{R}^n \mid \mathbf{w} = r_1 \mathbf{v}_1 + \dots + r_m \mathbf{v}_m, r_1, \dots, r_m \in \mathbb{R}\}.$$

要注意因為 $\mathbf{v}_1, \dots, \mathbf{v}_m$ 在 \mathbb{R}^n 中，所以 $\mathbf{v}_1, \dots, \mathbf{v}_m$ 的一個線性組合仍在 \mathbb{R}^n 中。也就是說 $\text{Span}(\mathbf{v}_1, \dots, \mathbf{v}_m)$ 是 \mathbb{R}^n 的一個子集合。這個子集合具有一重要的性質，我們稱之為 \mathbb{R}^n 的「子空間」(以後會給正式定義)。這個子集合具有怎樣的特殊性質呢？比方說利用定義我們可以知零向量 $\mathbf{0}$ 一定會在 $\text{Span}(\mathbf{v}_1, \dots, \mathbf{v}_m)$ 中 (即每個 r_i 皆取為 0)，又若

$\mathbf{w} \in \text{Span}(\mathbf{v}_1, \dots, \mathbf{v}_m)$, 則因 $\mathbf{w} = r_1\mathbf{v}_1 + \dots + r_m\mathbf{v}_m$ 可知 $-\mathbf{w} = (-r_1)\mathbf{v}_1 + \dots + (-r_m)\mathbf{v}_m$ 也就是說 $-\mathbf{w} \in \text{Span}(\mathbf{v}_1, \dots, \mathbf{v}_m)$. 這些性質我們可以推廣到更一般的情況而得到以下的定理.

Proposition 1.3.2. 給定 $\mathbf{v}_1, \dots, \mathbf{v}_m \in \mathbb{R}^n$, 則對任意 $\mathbf{u}, \mathbf{w} \in \text{Span}(\mathbf{v}_1, \dots, \mathbf{v}_m)$ 以及 $s, t \in \mathbb{R}$ 皆有

$$s\mathbf{u} + t\mathbf{w} \in \text{Span}(\mathbf{v}_1, \dots, \mathbf{v}_m).$$

Proof. 因 $\mathbf{u}, \mathbf{w} \in \text{Span}(\mathbf{v}_1, \dots, \mathbf{v}_m)$ 由定義知存在 $r_1, \dots, r_m \in \mathbb{R}$ 以及 $r'_1, \dots, r'_m \in \mathbb{R}$ 使得 $\mathbf{u} = r_1\mathbf{v}_1 + \dots + r_m\mathbf{v}_m$ 以及 $\mathbf{w} = r'_1\mathbf{v}_1 + \dots + r'_m\mathbf{v}_m$. 因此得

$$s\mathbf{u} + t\mathbf{w} = (sr_1 + tr'_1)\mathbf{v}_1 + \dots + (sr_m + tr'_m)\mathbf{v}_m.$$

也就是說 $s\mathbf{u} + t\mathbf{w}$ 仍為 $\mathbf{v}_1, \dots, \mathbf{v}_m$ 的一個線性組合, 故得證 $s\mathbf{u} + t\mathbf{w} \in \text{Span}(\mathbf{v}_1, \dots, \mathbf{v}_m)$. \square

雖然在 Proposition 1.3.2 中我們僅提及 $\text{Span}(\mathbf{v}_1, \dots, \mathbf{v}_m)$ 中的兩個向量的線性組合仍在 $\text{Span}(\mathbf{v}_1, \dots, \mathbf{v}_m)$ 中, 不過利用數學歸納法, 我們可以證得任意有限多個 $\text{Span}(\mathbf{v}_1, \dots, \mathbf{v}_m)$ 中的向量的線性組合也會在 $\text{Span}(\mathbf{v}_1, \dots, \mathbf{v}_m)$ 中.

1.4. Inner Product

在 \mathbb{R}^2 和 \mathbb{R}^3 中大家熟悉內積的定義也可以推廣到一般的 \mathbb{R}^n . 將來我們會知道內積可以幫助我們定義出 \mathbb{R}^n 中許多重要的子空間, 在本節我們僅論及大家熟悉的內積性質在 \mathbb{R}^n 的情況.

首先我們回顧在 \mathbb{R}^2 和 \mathbb{R}^3 中內積的定義. 若在 \mathbb{R}^2 中 $\mathbf{u} = (a_1, a_2), \mathbf{v} = (b_1, b_2)$, 則 \mathbf{u}, \mathbf{v} 的內積 $\mathbf{u} \cdot \mathbf{v}$ 定義成 $\mathbf{u} \cdot \mathbf{v} = a_1b_1 + a_2b_2$. 而在 \mathbb{R}^3 中若 $\mathbf{u} = (a_1, a_2, a_3), \mathbf{v} = (b_1, b_2, b_3)$, 則 \mathbf{u}, \mathbf{v} 的內積 $\mathbf{u} \cdot \mathbf{v}$ 定義成 $\mathbf{u} \cdot \mathbf{v} = a_1b_1 + a_2b_2 + a_3b_3$. 由這定義我們很自然地可推廣到 \mathbb{R}^n 中向量的內積如下:

Definition 1.4.1. 假設 $\mathbf{u} = (a_1, \dots, a_n), \mathbf{v} = (b_1, \dots, b_n) \in \mathbb{R}^n$. 則定義 \mathbf{u}, \mathbf{v} 的 *inner product* (內積) 為

$$\mathbf{u} \cdot \mathbf{v} = a_1b_1 + \dots + a_nb_n = \sum_{i=1}^n a_ib_i.$$

向量的內積和向量的運算有一定的關係, 以下就是它們之間的關係

Proposition 1.4.2. 對任意 $\mathbf{u}, \mathbf{v}, \mathbf{w} \in \mathbb{R}^n$, 我們有以下的性質:

- (1) $\mathbf{u} \cdot \mathbf{v} = \mathbf{v} \cdot \mathbf{u}$.
- (2) $\mathbf{u} \cdot \mathbf{u} \geq 0$ 且 $\mathbf{u} \cdot \mathbf{u} = 0$ 若且唯若 $\mathbf{u} = \mathbf{0}$.
- (3) 對任意 $r \in \mathbb{R}$ 皆有 $(r\mathbf{u}) \cdot \mathbf{v} = \mathbf{u} \cdot (r\mathbf{v}) = r(\mathbf{u} \cdot \mathbf{v})$.
- (4) $\mathbf{u} \cdot (\mathbf{v} + \mathbf{w}) = \mathbf{u} \cdot \mathbf{v} + \mathbf{u} \cdot \mathbf{w}$.

Proof. 這些性質在 \mathbb{R}^2 和 \mathbb{R}^3 大家應都了解, 在 \mathbb{R}^n 上的證明其實也一樣, 不同的是 n 可以是任意自然數我們無法完整地寫下 \mathbb{R}^n 中的向量, 而需藉由符號的幫助.

假設 $\mathbf{u} = (a_1, \dots, a_n)$, $\mathbf{v} = (b_1, \dots, b_n)$, $\mathbf{w} = (c_1, \dots, c_n)$. 我們想利用 \sum (summation) 這個符號來處理內積, 讓大家習慣這個便利的符號.

(1) 依定義

$$\mathbf{u} \cdot \mathbf{v} = \sum_{i=1}^n a_i b_i$$

這表示 $\mathbf{u} \cdot \mathbf{v}$ 是這些 $a_i b_i$ 的和其中 i 是跑遍 1 到 n 的所有正整數. 由於這 n 項的每一項 $a_i b_i$ 皆等於 $b_i a_i$ (實數乘法交換率) 所以我們知道它們的和也相等, 也就是說

$$\sum_{i=1}^n a_i b_i = \sum_{i=1}^n b_i a_i = \mathbf{v} \cdot \mathbf{u}.$$

所以我們得 $\mathbf{u} \cdot \mathbf{v} = \mathbf{v} \cdot \mathbf{u}$.

(2) 依定義

$$\mathbf{u} \cdot \mathbf{u} = \sum_{i=1}^n a_i a_i = \sum_{i=1}^n a_i^2.$$

由於任一實數的平方皆大於等於 0, 即 $a_i^2 \geq 0$, 故有 $\sum_{i=1}^n a_i^2 \geq 0$, 而得證 $\mathbf{u} \cdot \mathbf{u} \geq 0$. 又上式中若 $\sum_{i=1}^n a_i^2 = 0$, 表示每一項 a_i^2 皆需等於 0, 故知對任意 $1 \leq i \leq n$ 皆需有 $a_i = 0$, 而得知

$$\mathbf{u} = (a_1, \dots, a_n) = (0, \dots, 0) = \mathbf{0}.$$

反之若 $\mathbf{u} = \mathbf{0}$ 表示對任意 $1 \leq i \leq n$ 皆有 $a_i = 0$, 故得

$$\mathbf{u} \cdot \mathbf{u} = \sum_{i=1}^n a_i a_i = 0.$$

(3) $(r\mathbf{u}) \cdot \mathbf{v}$ 這個符號表示 $r\mathbf{u}$ 這個向量與 \mathbf{v} 的內積, 因 $r\mathbf{u} = (ra_1, \dots, ra_n)$ 故由定義知

$$(r\mathbf{u}) \cdot \mathbf{v} = \sum_{i=1}^n (ra_i) b_i.$$

又對所有的 $1 \leq i \leq n$ 皆有 $(ra_i) b_i = r(a_i b_i)$ (實數乘法結合律) 故知

$$\sum_{i=1}^n (ra_i) b_i = \sum_{i=1}^n r(a_i b_i)$$

再加上 $\sum_{i=1}^n r(a_i b_i)$ 中每一項皆有 r 可提出, 故由實數加法與乘法的分配律可知

$$\sum_{i=1}^n r(a_i b_i) = r \sum_{i=1}^n a_i b_i = r(\mathbf{u} \cdot \mathbf{v})$$

而得證 $(r\mathbf{u}) \cdot \mathbf{v} = r(\mathbf{u} \cdot \mathbf{v})$. 我們也可用同樣方法證得 $\mathbf{u} \cdot (r\mathbf{v}) = r(\mathbf{u} \cdot \mathbf{v})$, 不過我們這裡可利用 (1) 知 $\mathbf{u} \cdot (r\mathbf{v}) = (r\mathbf{v}) \cdot \mathbf{u}$ 再利用剛才的結果得 $(r\mathbf{v}) \cdot \mathbf{u} = r(\mathbf{v} \cdot \mathbf{u})$, 再利用一次 (1) 得到 $r(\mathbf{v} \cdot \mathbf{u}) = r(\mathbf{u} \cdot \mathbf{v})$ 而得證 $\mathbf{u} \cdot (r\mathbf{v}) = r(\mathbf{u} \cdot \mathbf{v})$.

(4) $\mathbf{u} \cdot (\mathbf{v} + \mathbf{w})$ 這個符號表示 \mathbf{u} 這個向量與 $\mathbf{v} + \mathbf{w}$ 的內積, 因

$$\mathbf{v} + \mathbf{w} = (b_1 + c_1, \dots, b_n + c_n)$$

故由定義知

$$\mathbf{u} \cdot (\mathbf{v} + \mathbf{w}) = \sum_{i=1}^n a_i(b_i + c_i).$$

由實數加法與乘法的分配律知每一項 $a_i(b_i + c_i)$ 可表為 $a_i b_i + a_i c_i$, 也就是說

$$\sum_{i=1}^n a_i(b_i + c_i) = \sum_{i=1}^n (a_i b_i + a_i c_i)$$

因為實數加法有交換率, 我們可以先將 $a_i b_i$ 的部份先加在一起, 再將 $a_i c_i$ 的部份加在一起, 再求它們之和, 故知

$$\sum_{i=1}^n (a_i b_i + a_i c_i) = \sum_{i=1}^n a_i b_i + \sum_{i=1}^n a_i c_i = \mathbf{u} \cdot \mathbf{v} + \mathbf{u} \cdot \mathbf{w},$$

依此得證 $\mathbf{u} \cdot (\mathbf{v} + \mathbf{w}) = \mathbf{u} \cdot \mathbf{v} + \mathbf{u} \cdot \mathbf{w}$. □

Proposition 1.4.2 (2) 告訴我們除了零向量 $\mathbf{0}$ 以外其餘向量 \mathbf{v} 皆需符合 $\mathbf{v} \cdot \mathbf{v} > 0$, 所以很自然地我們可依此定義向量的長度.

Definition 1.4.3. 令 $\mathbf{v} = (a_1, \dots, a_n) \in \mathbb{R}^n$, 我們定義 \mathbf{v} 的長度 (length) 為

$$\|\mathbf{v}\| = \sqrt{\mathbf{v} \cdot \mathbf{v}} = \sqrt{a_1^2 + a_2^2 + \dots + a_n^2}.$$

我們可以利用 Proposition 1.4.2 的處理一些有關於內積的性質, 而不必涉及內積的定義.

Lemma 1.4.4. 假設 $\mathbf{u}, \mathbf{v} \in \mathbb{R}^n$, 則 $\|\mathbf{u} + \mathbf{v}\|^2 = \|\mathbf{u}\|^2 + 2\mathbf{u} \cdot \mathbf{v} + \|\mathbf{v}\|^2$.

Proof. 依定義 $\|\mathbf{u} + \mathbf{v}\|^2 = (\mathbf{u} + \mathbf{v}) \cdot (\mathbf{u} + \mathbf{v})$, 再依 Proposition 1.4.2 (4) 可得

$$(\mathbf{u} + \mathbf{v}) \cdot (\mathbf{u} + \mathbf{v}) = (\mathbf{u} + \mathbf{v}) \cdot \mathbf{u} + (\mathbf{u} + \mathbf{v}) \cdot \mathbf{v} = \mathbf{u} \cdot \mathbf{u} + \mathbf{v} \cdot \mathbf{u} + \mathbf{u} \cdot \mathbf{v} + \mathbf{v} \cdot \mathbf{v}.$$

最後再依 Proposition 1.4.2 (1) 的交換律知 $\mathbf{v} \cdot \mathbf{u} + \mathbf{u} \cdot \mathbf{v} = 2\mathbf{u} \cdot \mathbf{v}$ 而得證本定理. □

再次強調一次 Lemma 1.4.4 僅用到內積的性質, 所以在一般的情形若我們不是利用 Definition 1.4.1 的方法定義內積 (當然此時長度的定義也跟著改變) 但所定義的內積仍保有 Proposition 1.4.2 中的性質, 我們依然可得到 Lemma 1.4.4 中的性質. Lemma 1.4.4 最常見的就是可以幫助我們推得所謂的「柯希、舒瓦茲」不等式.

Proposition 1.4.5 (Cauchy-Schwarz inequality). 若 $\mathbf{u}, \mathbf{v} \in \mathbb{R}^n$, 則 $|\mathbf{u} \cdot \mathbf{v}| \leq \|\mathbf{u}\| \|\mathbf{v}\|$. 特別地當 \mathbf{u}, \mathbf{v} 皆不為零向量時, 等號成立若且唯若存在 $\lambda \in \mathbb{R}$ 使得 $\mathbf{v} = \lambda \mathbf{u}$.

Proof. 假設 \mathbf{u} 和 \mathbf{v} 中有一個為零向量, 即 $\mathbf{u} \cdot \mathbf{v} = 0$ 且 $\|\mathbf{u}\| \|\mathbf{v}\| = 0$, 故此不等式成立.

若 \mathbf{u}, \mathbf{v} 皆不為零向量, 考慮 $\mathbf{u}_0 = \mathbf{u}/\|\mathbf{u}\|$ 且 $\mathbf{v}_0 = \mathbf{v}/\|\mathbf{v}\|$. 此時

$$\|\mathbf{u}_0\|^2 = \mathbf{u}_0 \cdot \mathbf{u}_0 = \frac{1}{\|\mathbf{u}\|} \mathbf{u} \cdot \frac{1}{\|\mathbf{u}\|} \mathbf{u} = \frac{1}{\|\mathbf{u}\|^2} \mathbf{u} \cdot \mathbf{u} = 1.$$

同理得 $\|\mathbf{v}_0\|^2 = 1$, 故由 Lemma 1.4.4 得知

$$\|\mathbf{u}_0 + \mathbf{v}_0\|^2 = 2 + 2\mathbf{u}_0 \cdot \mathbf{v}_0, \quad (1.6)$$

$$\|\mathbf{u}_0 - \mathbf{v}_0\|^2 = 2 - 2\mathbf{u}_0 \cdot \mathbf{v}_0.$$

因為對任意的 $\mathbf{w} \in \mathbb{R}^n$ 皆有 $\|\mathbf{w}\|^2 \geq 0$, 故得 $-1 \leq \mathbf{u}_0 \cdot \mathbf{v}_0 \leq 1$. 換回 \mathbf{u}, \mathbf{v} 得

$$-\|\mathbf{u}\| \|\mathbf{v}\| \leq \mathbf{u} \cdot \mathbf{v} \leq \|\mathbf{u}\| \|\mathbf{v}\|.$$

亦即 $|\mathbf{u} \cdot \mathbf{v}| \leq \|\mathbf{u}\| \|\mathbf{v}\|$.

從上可知當 \mathbf{u}, \mathbf{v} 皆不為零向量時, 此不等式之等式會成立等同於 $\mathbf{u}_0 \cdot \mathbf{v}_0 = 1$ 或 $\mathbf{u}_0 \cdot \mathbf{v}_0 = -1$. 此時由式子 (1.6) 分別得 $\|\mathbf{u}_0 - \mathbf{v}_0\|^2 = 0$ 或 $\|\mathbf{u}_0 + \mathbf{v}_0\|^2 = 0$, 也就是說 $\mathbf{u}_0 = \mathbf{v}_0$ 或 $\mathbf{u}_0 = -\mathbf{v}_0$. 換回 \mathbf{u}, \mathbf{v} 我們得 $\mathbf{u}_0 = \mathbf{v}_0$ 或 $\mathbf{u}_0 = -\mathbf{v}_0$. 這分別表示

$$\mathbf{v} = \frac{\|\mathbf{v}\|}{\|\mathbf{u}\|} \mathbf{u} \text{ 或 } \mathbf{v} = -\frac{\|\mathbf{v}\|}{\|\mathbf{u}\|} \mathbf{u}.$$

故此時只要令 λ 分別為 $\|\mathbf{v}\|/\|\mathbf{u}\|$ 或 $-\|\mathbf{v}\|/\|\mathbf{u}\|$, 即可得 $\mathbf{v} = \lambda\mathbf{u}$.

反之若 $\mathbf{v} = \lambda\mathbf{u}$, 則由 Proposition 1.4.2 可得

$$|\mathbf{u} \cdot \mathbf{v}| = |\lambda| |\mathbf{u} \cdot \mathbf{u}| = |\lambda| \|\mathbf{u}\|^2 = \|\mathbf{u}\| \|\lambda\mathbf{u}\| = \|\mathbf{u}\| \|\mathbf{v}\|.$$

□

利用內積我們可以知道坐標平面或空間中向量之間的一些幾何關係. 例如若兩非零向量 \mathbf{u}, \mathbf{v} 的夾角為 θ , 因為 $\mathbf{u} \cdot \mathbf{v} = \|\mathbf{u}\| \|\mathbf{v}\| \cos \theta$, 所以我們可以利用內積得知此二非零向量所夾角度. 特別地當 $\mathbf{u} \cdot \mathbf{v} = 0$ 即表示 \mathbf{u} 和 \mathbf{v} 垂直. 我們也可將此幾何意義推廣到更一般的 \mathbb{R}^n . 雖然當 $n \geq 4$ 時, 我們無法“看到” \mathbb{R}^n 中的向量 (無法用幾何的方式來定義夾角), 此時我們可以沿襲 $\mathbb{R}^2, \mathbb{R}^3$ 上的結果定義兩非零向量 $\mathbf{u}, \mathbf{v} \in \mathbb{R}^n$ 的夾角為 θ , 其中 $0 \leq \theta \leq \pi$ 使得

$$\cos \theta = \frac{\mathbf{u} \cdot \mathbf{v}}{\|\mathbf{u}\| \|\mathbf{v}\|}.$$

當我們定義一個東西時要注意這個定義是否 “well-defined”. 也就是說要確認這樣定義出來的夾角 θ 是否可以找得到, 這是所謂「存在性」的問題. 我們都知道當 $0 \leq \theta \leq \pi$ 時, $|\cos \theta| \leq 1$. 所以這裡夾角 θ 的存在性就關係到 \mathbb{R}^n 中兩個非零向量 \mathbf{u}, \mathbf{v} 是否會滿足

$$\left| \frac{\mathbf{u} \cdot \mathbf{v}}{\|\mathbf{u}\| \|\mathbf{v}\|} \right| \leq 1.$$

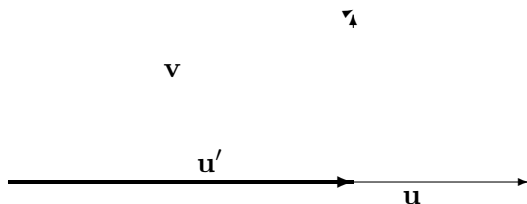
然而 Proposition 1.4.5 告訴我們這是一定對的, 所以這裡 θ 的存在性沒問題. 另一個要確認的問題是, 這樣定出來的夾角會不會有兩個或更多呢? 這是所謂「唯一性」的問題. 就是因為會有 $\theta' \neq \theta$ 但 $\cos \theta = \cos \theta'$ 的情形發生, 所以這裡我們要求 θ 要滿足 $0 \leq \theta \leq \pi$, 如此才能確保所得的夾角會是唯一的. 也就是說用這種方法定義兩非零向

量的夾角是沒有問題的，我們就稱這樣的定義是 well-defined. 利用夾角的定義我們進而定義出何謂「垂直」.

Definition 1.4.6. 令 $\mathbf{u}, \mathbf{v} \in \mathbb{R}^n$ 為非零向量，我們說 \mathbf{u} 和 \mathbf{v} 為 *orthogonal* 若且唯若 $\mathbf{u} \cdot \mathbf{v} = 0$.

注意這裡因在 \mathbb{R}^n 空間，習慣上垂直我們稱為 *orthogonal* 而不用大家一般幾何上常用的 *perpendicular*. 有了垂直概念後，我們也可以將 \mathbb{R}^2 或 \mathbb{R}^3 上的向量在另一向量上的投影 (projection) 之概念推廣至 \mathbb{R}^n .

我們先看 \mathbb{R}^2 的情況，給定一非零向量 $\mathbf{u} \in \mathbb{R}^2$ ，對任意 $\mathbf{v} \in \mathbb{R}^2$ ，若 \mathbf{u}' 為 \mathbf{v} 在 \mathbf{u} 上的投影，表示向量 $\mathbf{v} - \mathbf{u}'$ (參考下圖虛線表示的向量) 會和 \mathbf{u} 垂直，即 $(\mathbf{v} - \mathbf{u}') \cdot \mathbf{u} = 0$ ，也就是說 $\mathbf{v} \cdot \mathbf{u} = \mathbf{u}' \cdot \mathbf{u}$.



因為 \mathbf{u}' 和 \mathbf{u} 平行，我們可找到 $r \in \mathbb{R}$ 使得 $\mathbf{u}' = r\mathbf{u}$ ，代入前面式子得 $\mathbf{v} \cdot \mathbf{u} = r\mathbf{u} \cdot \mathbf{u} = r\|\mathbf{u}\|^2$. 也就是說，若令 $r = (\mathbf{v} \cdot \mathbf{u}) / \|\mathbf{u}\|^2$ (注意前面 \mathbf{u} 為非零向量的假設)，則 $r\mathbf{u}$ 就是 \mathbf{v} 在 \mathbf{u} 的投影. 我們可以將以上的概念推廣到 \mathbb{R}^n 的情形.

Proposition 1.4.7. 給定一非零向量 $\mathbf{u} \in \mathbb{R}^n$ ，對任意 $\mathbf{v} \in \mathbb{R}^n$ ，皆可寫成 $\mathbf{v} = \mathbf{u}' + \mathbf{v}'$ ，其中 $\mathbf{u}', \mathbf{v}' \in \mathbb{R}^n$ 滿足 $\mathbf{v}' \cdot \mathbf{u} = 0$ 且 $\mathbf{u}' = r\mathbf{u}$ ， $r \in \mathbb{R}$. 事實上這樣的寫法是唯一的，即

$$r = \frac{\mathbf{v} \cdot \mathbf{u}}{\|\mathbf{u}\|^2}.$$

Proof. 前面的論述在 \mathbb{R}^n 亦成立，亦即 $r = (\mathbf{v} \cdot \mathbf{u}) / \|\mathbf{u}\|^2$ 是唯一的實數會使得 $(\mathbf{v} - r\mathbf{u}) \cdot \mathbf{u} = 0$. 換言之，

$$\mathbf{u}' = \frac{\mathbf{v} \cdot \mathbf{u}}{\|\mathbf{u}\|^2} \mathbf{u}.$$

是唯一的向量會滿足 $\mathbf{u}' = r\mathbf{u}$ 且 $(\mathbf{v} - \mathbf{u}') \cdot \mathbf{u} = 0$. 既然 \mathbf{u}' 是唯一的，故而 \mathbf{v}' 要滿足 $\mathbf{v}' + \mathbf{u}' = \mathbf{v}$ ，即 $\mathbf{v}' = \mathbf{v} - \mathbf{u}'$ ，自然也就唯一確定了. \square

Proposition 1.4.7, 大致上是說給定一 \mathbb{R}^n 中的非零向量 \mathbf{u} 後，我們都可以將 \mathbb{R}^n 中任一向量 \mathbf{v} 分解成兩個向量之和，其中一個向量與 \mathbf{u} 平行 (即定理中的 \mathbf{u}') 而另一個與 \mathbf{u} 垂直 (即定理中的 \mathbf{v}')，且這個表法是唯一的. 我們稱與 \mathbf{u} 平行的那個向量

$$\frac{\mathbf{v} \cdot \mathbf{u}}{\|\mathbf{u}\|^2} \mathbf{u}$$

為 \mathbf{v} 在 \mathbf{u} 的 *projection* (投影).

1.5. 結論

在本章中我們將大家熟悉 $\mathbb{R}^2, \mathbb{R}^3$ 上的向量加法, 係數積, 內積等運算推廣至 \mathbb{R}^n 上. 這些推廣而得的運算與 $\mathbb{R}^2, \mathbb{R}^3$ 上的運算有共同的性質. 直接利用這些性質而不必利用這些運算的定義, 我們就可以得到許多和 $\mathbb{R}^2, \mathbb{R}^3$ 相類似的結果. 換言之, 只要有這些運算性質, 即使更抽象的空間, 有些事情我們依然可以如平面或空間一樣的情形“看到”. 希望大家能體會到這些運算性質的重要性, 將來我們就是要利用這些性質進入較抽象的線性代數世界.

Systems of Linear Equations

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

2.1. 解一次聯立方程組

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

$$a_1x_1 + \cdots + a_nx_n = b,$$

其中這些 a_1, \dots, 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

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

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

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

然後將上面的聯立方程式用 $A\mathbf{x} = \mathbf{b}$ 來表示。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式的方程式，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。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

$$\begin{aligned} 3x_1 - 2x_2 + 9x_4 &= 4 \\ 2x_1 + 2x_2 - 4x_4 &= 6 \end{aligned} \quad (2.1)$$

我們就可以表成

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

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

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

$$(1, 0, 1, 0) = c_1(1, -1, 2, 2) + c_2(3, 1, -1, 2).$$

亦即要解

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

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

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

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

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

- (1) 變換式子的順序
- (2) 將某一式乘上一非零實數
- (3) 將某一式乘上一實數後加到另一式上

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

當我們要解

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

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

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

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

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

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

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

- (1) 將矩陣的某兩個 row 對調
- (2) 將矩陣的某一個 row 乘上一非零實數
- (3) 將矩陣的某一個 row 乘上一實數後加到另一個 row。

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

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

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

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

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

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

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

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

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

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

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

(1) A' 每一個 row 皆不全為 0: 此時聯立方程組一定有解. 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“元”(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

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

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. 也就是方程組元的個數多於 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, 首先我們需要找到 *free variables*. 所謂 free variable 指的是方程組除了 pivot variable 以外的 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

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

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

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

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一定無解. 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有解. 事實上在此情形我們可以忽略全為 0 的 row, 例如前兩個 augmented matrices 所對應的方程組和

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

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

我們要強調, 絕不會有 pivot 的個數多於方程組 variables (元) 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.

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

2.2. Elementary Row Operations

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

我們利用數學歸納法來說明為何一定可以將一個矩陣化為 echelon form. 或許有些人會對這裡數學歸納法處理的方式覺得奇怪, 不過若能仔細體會其真意, 會發現這是最好的處理方式. 我們是對矩陣的 row 的個數作數學歸納法. 先說明所有只有一個 row

的矩陣一定是 echelon form, 然後利用這件事實證明所有有兩個 row 的矩陣皆可利用 elementary row operations 化為 echelon form. 再利用兩個 row 的矩陣會成立的事實證明有 3 個 row 的矩陣也可利用 elementary row operations 化為 echelon form, 如此一直下去我們可證有 4, 5, 6, ... 個 row 的矩陣會成立. 不過這樣的方法我們可以證得有特定個數的 row 的矩陣會成立 (例如 10 個 row), 但無法證得一般的情形 (即任意個數的 row). 此時數學歸納法是最好的論證工具了. 若我們能知道有 k 個 row 的矩陣一定能利用 elementary row operations 化為 echelon form 這個事實且利用這個事實證得有 $k+1$ 個 row 的矩陣一定能利用 elementary row operations 化為 echelon form, 這就表示當我們知道有一個 row 的矩陣能利用 elementary row operations 化為 echelon form 就能推得有兩個 row 的矩陣能利用 elementary row operations 化為 echelon form, 也進而推得有 3 個 row 的矩陣亦成立, 再進而推得有 4 個 row 的矩陣亦成立, 如此一直下去當然可知任意的矩陣皆能利用 elementary row operations 化為 echelon form.

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

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

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

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

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

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

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

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

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

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

這一個 echelon form. 要注意我們故意忽略原來矩陣的第一個 row 的原因就是餵了能套用歸納法的假設. 事實上前一步是將矩陣

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

的第二個 row 乘以 2 加到第三個 row 得

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

這一個 echelon form.

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

我們可以如法泡製處理有 3 個 row 的矩陣, 但由於要使用數學歸納法, 此時我們可直接假設我們以處理到有 k 個 row 的矩陣了, 亦即有 k 個 row 的矩陣皆可利用 elementary row operation 化為 echelon form. 現在我們要處理有 $k+1$ 個 row 的矩陣. 如欠面首先我們將 leading entry 的位置在最左邊的那個 row 利用兩 row 互換的 row operation 將之置於第一個 row. 假設此時第一個 row 的 leading entry 為 a . 接下來我們將 leading entry 的位置與第一個 row 的 leading entry 位置一樣的 row 挑出, 若該 row 的 leading entry 為 b , 我們便將第一個 row 乘上 $-b/a$ 後加到該 row 上. 如此一來該 row 的 leading entry 所在位置便往右移了. 一直重複此步驟, 直到第一個 row 以外的 row 其 leading entry 所在位置皆與第一個 row 的 leading entry 所在位置相異. 注意, 此時第一個 row 以下的各 row 其 leading entry 所在位置皆在第一個 row 的 leading entry 所在位置的右方. 若我們不看第一個 row, 所剩下的是一個有 k 個 row 的矩陣, 所以利用前面已知有 k 個 row 的矩陣皆可利用 elementary row operations 化為 echelon form, 我們可以利用 elementary row operations 將此矩陣第一個 row 以下的部份化為 echelon form. 但此時因各個 row 的 leading entry 所在位置皆在第一個 row 的

leading entry 所在位置的右方, 所以整個矩陣亦為 echelon form. 故得證所有矩陣皆可利用 elementary row operations 化為 echelon form. 大家或許注意到我們在化成 echelon form 的過程皆沒有用到將某個 row 乘上一非 0 實數這一個 elementary row operation. 事實上在化成 echelon form 的過程確實不需要這一種 row operation, 不過它在化為以後要談的 “reduced” echelon form 的過程是需要的, 留待以後再談.

既然每一個矩陣都能用 elementary row operations 化為 echelon form, 接下來我們要說明的是利用 elementary row operation 處理後的聯立方程組其解集合不會改變. 要注意這裡指的是將 augmented matrix 用 elementary row operations 變換後的 augmented matrix 其對應的聯立方程組其解集合不會改變. 亦即聯立方程組 $Ax = \mathbf{b}$ 所對應的 augmented matrix $[A | \mathbf{b}]$ 若經一些 elementary row operations 後變換成 $[A' | \mathbf{b}']$, 那麼其對應的聯立方程組 $A'\mathbf{x} = \mathbf{b}'$ 和原方程組 $Ax = \mathbf{b}$ 有相同的解集合. 不要誤以為是將係數矩陣 A 利用 elementary row operations 變換成 A' 後聯立方程組 $A'\mathbf{x} = \mathbf{b}$ 的解集合和原方程組 $Ax = \mathbf{b}$ 的解集合相同.

首先觀察若將一聯立方程組 $Ax = \mathbf{b}$ 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之一將之變成方程組 $A'\mathbf{x} = \mathbf{b}'$. 然而方程組 $Ax = \mathbf{b}$ 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 $Ax = \mathbf{b}$ 的一組解仍會滿足 $A'\mathbf{x} = \mathbf{b}'$. 換句話說 $Ax = \mathbf{b}$ 的解集合會包含於 $A'\mathbf{x} = \mathbf{b}'$ 的解集合. 不過 elementary row operations 是可以還原回去的, 也就是說方程組 $A'\mathbf{x} = \mathbf{b}'$ 也可以用 elementary row operations 還原回原方程組 $Ax = \mathbf{b}$. 因此 $A'\mathbf{x} = \mathbf{b}'$ 的解集合也會包含於 $Ax = \mathbf{b}$ 的解集合. 因此得證 $Ax = \mathbf{b}$ 和 $A'\mathbf{x} = \mathbf{b}'$ 會有相同的解集合. 我們證得了 $[A | \mathbf{b}]$ 若經由一個 elementary row operation 後得 $[A' | \mathbf{b}']$, 則它們所對應的聯立方程組會有相同的解集合. 因此若 $[A | \mathbf{b}]$ 經由好幾次的 elementary row operation 變換成 $[A' | \mathbf{b}']$, 它們所對應的聯立方程組當然也會有相同的解集合.

我們已經了解要解一個聯立方程組 $Ax = \mathbf{b}$ 經要將 augmented matrix $[A | \mathbf{b}]$ 利用 elementary row operations 換成 $[A' | \mathbf{b}']$ 其中 A' 為 echelon form 後, 再解聯立方程組 $A'\mathbf{x} = \mathbf{b}'$ 即可. 所以我們僅需探討 $A'\mathbf{x} = \mathbf{b}'$ 其中 A' 為 echelon form 這樣的聯立方程組的解即可. 在下節中我們要說明此時為何聯立方程組的解就是如上一節所述的情形.

2.3. Echelon Form

我們已知要探討聯立方程組 $Ax = \mathbf{b}$ 的解, 僅要考慮 A 為 echelon form 的情形. 這一節中我們就是要討論當 A 為 echelon form 時, 聯立方程組 $Ax = \mathbf{b}$ 的解集合. 事實上我們很容易理解利用 2.1 節中所提求解的方法所得的結果皆為方程組的一組解. 這裡要探討的是為何利用 2.1 節中所提求解的方法, 就可得所有的解.

如果我們得到 2.1 節 (2)(a) 的情形 (即 A 有一個 row 全為 0 但 \mathbf{b} 在該 row 不為 0), 在該節已說明此時方程組無解. 所以我們只要探討有解的情形. 首先回顧一下在 2.1

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。以下就是 pivot 對聯立方程組的解之重要性。

Lemma 2.3.1. 假設 A 為一有 n 個 column 的 echelon form 且 $x_1 = c_1, \dots, x_n = c_n$ 和 $x_1 = d_1, \dots, x_n = d_n$ 皆為方程組 $Ax = \mathbf{b}$ 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 $c_{k+1} = d_{k+1}, \dots, c_n = d_n$ ，則 $c_k = d_k$ 。

Proof. 假設聯立方程組為

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

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}, \dots, 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 $x_1 = c_1, \dots, x_n = c_n$ 及 $x_1 = d_1, \dots, x_n = d_n$ 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 $x_1 = c_1, \dots, x_n = c_n$ 及 $x_1 = d_1, \dots, x_n = d_n$ 皆為此聯立方程組的一組解知 $x_k = c_k, x_{k+1} = c_{k+1}, \dots, x_n = c_n$ 和 $x_k = d_k, x_{k+1} = d_{k+1}, \dots, x_n = d_n$ 皆需滿足 $a_{ik}x_k + \cdots + a_{in}x_n = b_i$ 。因此由 $c_{k+1} = d_{k+1}, \dots, c_n = d_n$ 的假設知

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

再由 $a_{ik} \neq 0$ 得知 $c_k = d_k$ 。

□

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

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

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

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 $l > i$ 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 $x_1 = c_1, \dots, x_n = c_n$ 為方程組 $A\mathbf{x} = \mathbf{b}$ 的一組解. 換言之, $x_{k+1} = c_{k+1}, \dots, x_n = c_n$ 皆滿足方程組 pivot 的位置在 x_k 右方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}, \dots, x_n 的取值, 所以我們可令 $x_k = r$ 且 $x_{k+1} = c_{k+1}, \dots, x_n = c_n$ 一步一步代回求得聯立方程組的一組解. \square

Lemma 2.3.1 和 Lemma 2.3.2 有許多應用. 例如當 A 是 echelon form 時若聯立方程組 $A\mathbf{x} = \mathbf{b}$ 已知有一個解 $x_1 = -1, \dots, x_n = c_n$ 且 x_1, \dots, x_n 每一個都是 pivot variable, 則由 Lemma 2.3.1 知聯立方程組 $A\mathbf{x} = \mathbf{b}$ 的解僅能是 $x_1 = c_1, \dots, x_n = c_n$. 換句話說此方程組的解唯一. 另一方面, 若若聯立方程組 $A\mathbf{x} = \mathbf{b}$ 已知有解且 x_1, \dots, x_n 中有 free variable, 則由 Lemma 2.3.2 知聯立方程組 $A\mathbf{x} = \mathbf{b}$ 會有無窮多解.

當我們給一個矩陣時, 有許多種方法將之化為 echelon form, 而且化成的 echelon form 很可能不一樣. 不過利用 Lemma 2.3.1 和 Lemma 2.3.2 我們可以得到這些 echelon form 雖然可能不一樣, 但他們 pivot 的所在位置都會一致.

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

Proof. 我們考慮 $A\mathbf{x} = \mathbf{0}$ 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. 要注意這些聯立方程組都會有解, 因為 $x_1 = 0, \dots, x_n = 0$ 就是一組解.

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 (故為 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 2.3.1), 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 2.3.2). 這和此二方程組有相同的解相矛盾. 現若 $1 \leq i \leq n-1$. 利用 $x_1 = 0, \dots, x_n = 0$ 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}, \dots, x_n 的取值皆為 0 但 x_i 的取值不是 0 (Lemma 2.3.1); 另一方面 Lemma 2.3.2 告訴我們 $A_2\mathbf{x} = \mathbf{0}$ 的解中一定可找到一組解其 x_{i+1}, \dots, 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. \square

現在我們回來說明當 A 是 echelon form 時, 前幾節所述解聯立方程組 $A\mathbf{x} = \mathbf{b}$ 的方法所求得的解就是所有的解. 也就是說若 $x_1 = c_1, \dots, x_n = c_n$ 為 $A\mathbf{x} = \mathbf{b}$ 的一組解, 我們要說明這組解確實可由前面 2.1 節所提的方法得到. 為了方便起見我們令前面 2.1 節所提的方法所得的解所成的集合為 S . 現若 x_n 為 pivot variable, 則由 Lemma 2.3.1 知, 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, 再由 Lemma 2.3.1 知這一組解 x_{n-1}, x_n 的取值必為 $x_{n-1} = c_{n-1}, x_n = c_n$; 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, \dots, x_n 的取值為 $x_1 = c_1, \dots, x_n = c_n$.

我們已完全了解整個解聯立方程組的步驟與原理了. 在求解的過程中還可以進一步將 echelon form 化為所謂的 reduced echelon form. Reduced echelon form 事實上仍為 echelon form, 不過再加上兩個限制, 第一個限制是每一個 pivot 需為 1. 另一個限制為 pivot 的位置上方全為 0. 要注意, 依定義 echelon form 的 pivot 位置下方已全為 0 所以 reduced echelon form 每一個 pivot 所在的 column, 除了自己需為 1 外其他部分皆為 0. 例如

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

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. 若有一個 row 的 pivot 為 a (注意依定義 $a \neq 0$) 我們只要將

該 row 乘上 $1/a$, 則該 row 的 pivot 便是 1 了. 例如上面 A 這一個 echelon form 若將第二個 row 乘上 $1/3$, 就可得 A' 這一個 reduced echelon form. 當我們將每個 pivot 都變為 1 後, 就可利用將該 row 乘上某一實數加到另一個 row 的方法將 pivot 所在的 column 的其他部分化為 0. 例如上面 B 這一個 echelon form 若將第三個 row 乘上 -1 加到第二個 row, 就可得 B' 這一個 reduced echelon form. 既然每一個矩陣都可以經由 elementary row operation 化為 reduced echelon form, 所以我們可以利用這個方法求出聯立方程組的解. 化成 reduced echelon form 後求解可以幫助我們很快地看出解的形式, 例如方程組 $B'\mathbf{x} = \mathbf{0}$ 為

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

它的解集合為 $x_1 = 0, x_2 = -3t, x_3 = t, x_4 = t$ 其中 t 為任意實數. 不過一般來說化為 reduced echelon form 比僅化為 echelon form 所需的步驟多了許多, 所以利用 echelon form 來求解還是比較簡潔.

最後我們說明一下, 雖然一個矩陣化為 echelon form 的情形不唯一, 不過我們可以利用 Proposition 2.3.3 的結果證明一個矩陣化為 reduced echelon form 的結果會唯一. 然而這個事實我們以後不會用到, 這裡就略去不證了.

2.4. Rank of a Matrix

從前幾節解聯立方成組的方法來看, 當一聯立方程組其係數矩陣化為 echelon form 後我們很容易判斷其是否有解, 解是否唯一. 其中 pivot 的個數就是影響到聯立方程組解的個數的重要因素. 本節中我們就是要探討 pivot 的個數和解的個數之關係.

Definition 2.4.1. 將一矩陣利用 elementary row operations 化為 echelon form 後, 其不全為 0 的 row 的個數 (即 pivot 的個數) 稱為此矩陣的 *rank*. 若 A 是一個矩陣, 則將 A 的 rank 記為 $\text{rank}(A)$.

我們曾經提醒過, 要給一個定義前需注意這個定義是否 well-defined. 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. 所以要確認 rank 是 well-defined 我們就必須要知道給定一個矩陣, 不管用怎樣的 elementary row operations 所得的 echelon form 它們的 pivot 個數都相同. 事實上 Proposition 2.3.3 已經告訴我們這個事實, 所以這裡 rank 的定義沒有問題.

若一個矩陣 A 有 m 個 row 和 n 個 column, 由於每個 row 致多僅有一個 pivot, 而每個 column 也至多僅有一個 pivot, 所以 pivot 的個數不能超過 row 的個數以及 column 的個數, 換言之我們有 $\text{rank}(A) \leq m$ 且 $\text{rank}(A) \leq n$. 如果 $m \geq n$, 則 $\text{rank}(A) \leq n$ 就足以表達這兩個不等式, 反之若 $n \geq m$, 則 $\text{rank}(A) \leq m$ 就足以表達這兩個不等式. 所以若我們令 $\min\{m, n\}$ 表示取 m, n 中較小的數, 就會有以下之結果.

Proposition 2.4.2. 若矩陣 A 有 m 個 row 和 n 個 column, 則

$$\text{rank}(A) \leq \min\{m, n\}.$$

當 proposition 2.4.2 中的不等式其等號成立時，方程組的解的情形便可以掌握。事實上我們有以下的結果。

Theorem 2.4.3. 假設矩陣 A 有 m 個 row 和 n 個 column.

- (1) 若 $\text{rank}(A) = m$ 則對任意 $\mathbf{b} \in \mathbb{R}^m$ ，方程組 $A\mathbf{x} = \mathbf{b}$ 皆有解；反之，若 $\text{rank}(A) \neq m$ 則存在 $\mathbf{b} \in \mathbb{R}^m$ 使得方程組 $A\mathbf{x} = \mathbf{b}$ 無解。
- (2) 若 $\text{rank}(A) = n$ 則當方程組 $A\mathbf{x} = \mathbf{b}$ 有解時其解唯一；反之，若 $\text{rank}(A) \neq n$ 則當方程組 $A\mathbf{x} = \mathbf{b}$ 有解時，會有無窮多解。

Proof. 首先我們說明一下為何這裡要考慮 $\mathbf{b} \in \mathbb{R}^m$ 。依假設係數矩陣 A 有 m 個 row 和 n 個 column 這表示 $A\mathbf{x} = \mathbf{b}$ 是一個有 n 個未知數和 m 個方程式的聯立方程組。而 \mathbf{b} 代表的是方程組每個式子的常數項，既然有 m 個方程式自然需要 m 個數字來表示這些常數項，所以 $\mathbf{b} \in \mathbb{R}^m$ 。

為了方便起見我們假設聯立方程組 $A\mathbf{x} = \mathbf{b}$ 的 augmented matrix $[A \mid \mathbf{b}]$ 利用 elementary row operations 可化為 $[A' \mid \mathbf{b}']$ 其中 A' 為 echelon form.

- (1) 若 $\text{rank}(A) = m$ ，表示解聯立方程式將係數矩陣 A 化為 echelon form 後，每一個 row 皆有 pivot，亦即每一個 row 不全為 0。故由 2.1 節的討論我們之此時聯立方程組一定都有解。

反之，若 $\text{rank}(A) \neq m$ ，因已知 $\text{rank}(A) \leq m$ ，這表示 $\text{rank}(A) < m$ 。也就是說 echelon form A' 的最後一個 row 必全為 0，因此此時若

$$\mathbf{b}' = \begin{bmatrix} b'_1 \\ \vdots \\ b'_m \end{bmatrix}$$

且 $b'_m \neq 0$ ，則 $[A' \mid \mathbf{b}']$ 所對應的聯立方程組無解，亦即原方程組無解。因此我們只要選一組 \mathbf{b}' 其中 $b'_m \neq 0$ ，再將 augmented matrix $[A' \mid \mathbf{b}']$ 利用 elementary row operation 還原為 $[A \mid \mathbf{b}]$ ，則此 $\mathbf{b} \in \mathbb{R}^m$ 會使得 $A\mathbf{x} = \mathbf{b}$ 無解。

- (2) 若 $\text{rank}(A) = n$ ，表示解聯立方程式將係數矩陣 A 化為 echelon form 後，每一個 column 皆有 pivot，亦即 x_1, \dots, x_n 每一個都是 pivot variable (也就是說沒有 free variable)。因此若聯立方程組 $A\mathbf{x} = \mathbf{b}$ 有解依 Lemma 2.3.1 知聯立方程組的解唯一。

反之，若 $\text{rank}(A) \neq n$ ，這也表示 $\text{rank}(A) < n$ ，故知 x_1, \dots, x_n 這 n 個 variable 中必有 free variable。因此若聯立方程組 $A\mathbf{x} = \mathbf{b}$ 有解依 Lemma 2.3.2 知聯立方程組有無窮多解。

□

要注意 Theorem 2.4.3 (1) 說的是若 $\text{rank}(A)$ 不等於 row 的個數，則聯立方程組有可能無解 (注意僅是有可能，所以也有可能解)。因此當聯立方程組 $A\mathbf{x} = \mathbf{b}$ 的方程式個數多於未知數的個數時，即 A 的 row 的個數 m 大於 A 的 column 的個數 n 。此時由

Proposition 2.4.2 知 $\text{rank}(A) \leq n < m$, 故由 Theorem 2.4.3 (1) 的結果知, 此聯立方程組有可能無解. 不過這並不表示當方程式個數小於或等於未知數的個數時方程組一定有解.

若未知數的個數多於方程式的個數時, 此時 $n > m$, 故由 Proposition 2.4.2 知 $\text{rank}(A) \leq m < n$. 因此由 Proposition 2.4.3 (2) 的結果知此聯立方程組若有解的話, 會有無窮多組解. 另一方面 Theorem 2.4.3 (2) 說的是若 $\text{rank}(A)$ 等於 column 的個數時, 則當聯立方程組有解時其解唯一. 注意這並不表示聯立方程組一定有解. 不過當方程組是 $Ax = \mathbf{0}$, 由於此方程組一定有解 (即 $x_1 = \cdots = x_n = 0$ 為一解), 所以我們有以下之結果.

Corollary 2.4.4. 假設矩陣 A 有 m 個 row 和 n 個 column. 則 $Ax = \mathbf{0}$ 有唯一解若且唯若 $\text{rank}(A) = n$.

在數學上一個理論的推廣, 若一個結果可由某定理直接推得我們就稱為 Corollary. 上一個結果由於已知 $Ax = \mathbf{0}$ 一定有解, 所以可以直接由 Theorem 2.4.3 推得, 因此我們用 Corollary 稱之.

當方程式個數等於未知數的個數時 (即 $m = n$), 我們有以下之結果.

Corollary 2.4.5. 假設矩陣 A 有 n 個 row 和 n 個 column. 則 $\text{rank}(A) = n$ 若且唯若對任意 $\mathbf{b} \in \mathbb{R}^n$, 聯立方程組 $Ax = \mathbf{b}$ 必有解且其解唯一.

Proof. 假設 $\text{rank}(A) = n$, 因為 $\text{rank}(A)$ 等於 row 的個數, 故由 Theorem 2.4.3 (1) 知對任意 $\mathbf{b} \in \mathbb{R}^n$, 聯立方程組 $Ax = \mathbf{b}$ 必有解. 另一方面, $\text{rank}(A)$ 等於 column 的個數, 故由 Theorem 2.4.3 (2) 此有解之聯立方程組的解唯一.

假設對任意 $\mathbf{b} \in \mathbb{R}^n$, 聯立方程組 $Ax = \mathbf{b}$ 必有解且其解唯一. 特別地當 $\mathbf{b} = \mathbf{0}$ 時 $Ax = \mathbf{0}$ 的解唯一. 故由 Corollary 2.4.4 知 $\text{rank}(A) = n$. \square

我們已經知道當 $\text{rank}(A)$ 等於 column 的個數或 row 的個數時, 聯立方程組 $Ax = \mathbf{b}$ 的解有一些重要的性質. 而當 $\text{rank}(A)$ 不等於 column 的個數時, 以後我們會知道若聯立方程組 $Ax = \mathbf{b}$ 有解時, 它們的解會和 $Ax = \mathbf{0}$ 的無窮多個解有關. 接下來我們要談的是當 $\text{rank}(A)$ 不等於 row 的個數時, 如何判斷哪些 $\mathbf{b} \in \mathbb{R}^m$ 會使得 $Ax = \mathbf{b}$ 有解, 哪些會無解. 首先我們利用一個例子來說明:

Example 2.4.6. 我們想找到所有可能的 $b_1, b_2, b_3, b_4 \in \mathbb{R}$ 使得以下的聯立方程組有解.

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

首先我們考慮 augmented matrix

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

將第一個 row 分別乘上 $-3, -1, -3$ 加到第二,三,四個 row 得

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

再將此 augmented matrix 的第二個 row 乘上 -1 加到第三個 row, 得

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

換言之, 我們經由 elementary row operations 將係數矩陣化為 echelon form 得到以下有相同解的聯立方程組

$$\begin{aligned} x_1 - x_2 + x_3 &= b_1 \\ 5x_2 - 4x_3 &= b_2 - 3b_1 \\ 0x_1 + 0x_2 + 0x_3 &= b_3 - b_2 + 2b_1 \\ 0x_1 + 0x_2 + 0x_3 &= b_4 - 3b_1 \end{aligned}$$

要注意, 這裡我們將係數矩陣全為 0 的 row (即第三,四 row) 所對應的方程式係數用 0 標出, 主要是呼應 2.1 節 (2)(a) 的情形 (即 A 有一個 row 全為 0 但 \mathbf{b} 在該 row 不為 0). 由此知, 若 $b_3 - b_2 + 2b_1 \neq 0$ 或 $b_4 - 3b_1 \neq 0$ 則不管 x_1, x_2, x_3 代任何實數皆無法滿足方程式 $0x_1 + 0x_2 + 0x_3 = b_3 - b_2 + 2b_1$ 以及 $0x_1 + 0x_2 + 0x_3 = b_4 - 3b_1$, 亦即此聯立方程組無解. 反之若 $b_3 - b_2 + 2b_1 = 0$ 且 $b_4 - 3b_1 = 0$, 此時原方程組的解等同於方程組

$$\begin{aligned} x_1 - x_2 + x_3 &= b_1 \\ 5x_2 - 4x_3 &= b_2 - 3b_1 \end{aligned}$$

的解, 由 2.1 節的討論之此時必有解. 也就是說原方程組有沒有解, 完全取決於 $b_3 - b_2 + 2b_1$ 以及 $b_4 - 3b_1$ 是否皆為 0. 若 $b_3 - b_2 + 2b_1 = 0$ 且 $b_4 - 3b_1 = 0$, 則聯立方程組必有解. 若 $b_3 - b_2 + 2b_1 \neq 0$ 或 $b_4 - 3b_1 \neq 0$ 則聯立方程組必無解. 例如當 $b_1 = 1, b_2 = 1, b_3 = -1, b_4 = 3$ 時我們得 b_1, b_2, b_3, b_4 皆滿足 $b_3 - b_2 + 2b_1 = 0$ 且 $b_4 - 3b_1 = 0$, 所以聯立方程組

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

有解. 但若 $b_1 = 1, b_2 = 1, b_3 = -1, b_4 = 2$, 此時 $b_4 - 3b_1 \neq 0$, 故知聯立方程組

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

無解.

由 Example 2.4.6 的說明, 我們可以理解一般來說若給定一個有 m 個 row 的係數矩陣 A , 我們要找到 \mathbf{b} 使得聯立方程組 $A\mathbf{x} = \mathbf{b}$ 有解. 首先我們將 \mathbf{b} 寫成

$$\mathbf{b} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix}$$

然後利用 elementary row operations 將 augmented matrix $[A | \mathbf{b}]$ 化成 $[A' | \mathbf{b}']$ 其中 A' 為 echelon form. 注意這裡因為 \mathbf{b} 是由 b_i 這樣的變數來表示, 所以 \mathbf{b}' 每一個位置是一些由 b_i 所組成的(一次)多項式. 挑出 A' 中全為 0 的 row, 令 \mathbf{b}' 在這些 row 所對應的多項式為 0 的方程式稱為此聯立方程組的 *constrain equations*. 使得 constrain equations 皆為 0 的 \mathbf{b} 就是使得聯立方程組有解的 \mathbf{b} . 例如在 Example 2.4.6 中,

$$b_3 - b_2 + 2b_1 = 0, b_4 - 3b_1 = 0$$

就是聯立方程組

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

的 constrain equations. 寫下 constrain equations 的意義就是使得聯立方程組 $A\mathbf{x} = \mathbf{b}$ 有解的那些 \mathbf{b} 的“限制”. 亦即, 滿足 constrain equations 的 \mathbf{b} 就會使得聯立方程組 $A\mathbf{x} = \mathbf{b}$ 有解; 而不滿足 constrain equations 的 \mathbf{b} 就會使得聯立方程組 $A\mathbf{x} = \mathbf{b}$ 無解. 特別地, 當 $\text{rank}(A)$ 等於 A 的 row 的個數 m 時, 由於沒有 constrain equation, 也就是說 \mathbf{b} 沒有限制, 所以對任意的 $\mathbf{b} \in \mathbb{R}^m$ 聯立方程組 $A\mathbf{x} = \mathbf{b}$ 皆有解. 這和 Theorem 2.4.3 (1) 的結果是一致的.

2.5. 結論

在本章中我們學習解聯立方程組的技巧. 利用 elementary row operations 將 augmented matrix 中的係數矩陣化為 echelon form 後, 我們很快的可以知道此聯立方程組是否有解, 而有解時也可利用此 echelon form 完整的得到此聯立方程組所有的解. 由 echelon form 的解法我們了解到 pivot 對聯立方程組是否有解以及解是否唯一有著重要的關連, 因而定義出矩陣的 rank. 矩陣的 rank 的概念, 以後我們會知道和許多概念都可以連結起來, 不過目前最重要的就是要了解它和聯立方程組解的關係. 在下一章, 我們將介紹矩陣的運算, 這些運算都可以和解聯立方程組的問題相連結. 所以本章中有關聯立

方程組的理論對後面理論的建立影響深遠，千萬不要以為會解具體的聯立方程組就可以了而忽視這些理論。