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數學系大一學生介紹有關線性代數基本的理論, 大致上僅談論實係數的向量空間.

一般經驗上大一學生會覺得線性代數學習起來會比微積分吃力, 主要是因素可能是在高中時期學習有關線性代數部分主要著重於操作而較少論證. 這一點在大學線性代數中就較不同了. 不過經驗上也告訴我們當學生到大三大四時 (因準備研究所考試) 再重新溫習線性代數時, 會覺得它不再那麼難以親近. 由此可知大學時期的線性代數其理論並不難懂, 較大的障礙是要開始學習數學的論證. 這個障礙或許到大三大四一些數學思維較成熟時稍可解除, 不過對一些同學來說可能為時已晚, 畢竟線性代數的理論或概念與其他課程都脫不了關係. 基於這個原因寫下這份講義, 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.

研讀本講義的同學要有心理準備, 本講義是針對數學系學生而寫, 自然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理論清楚了, 接下來的應用或推廣就不難了. 所以對應用有興趣的同學應再參考其他的參考書籍. 另外本講義並未提供習題, 不過在某些概念講述之後有時會提供一些問題 (Question). 這些問題幾乎是檢視觀念是否正確或是對內容是否了解, 大部分問題若觀念已清楚應可以立即回答. 所以這些問題的份量仍不及一般習題, 對熟習線性代數所給予的訓練. 針對這一點, 請欲學習好線性代數的同學務必參閱一般線性代數書籍的習題, 自行磨練.

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.

Chapter 1

Vectors in $\mathbb{R}^n$

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性質, 利用較抽象的方法來推導出一些幾何的性質.

1.1. 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. 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以常數. 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

$$\overrightarrow{PQ}+\overrightarrow{QR}=\overrightarrow{PR}. \tag{1.1}$$

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 $r\mathbf{u}=(5\times1,5\times2)=(5,10)$. 不過我們絕對不能用: 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 $r\mathbf{u}=(5\times1,5\times2)=(5,10)$ 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: 這裡我們並不是定義兩向量的乘法, 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 $\mathbf{O}\in\mathbb{R}^2$ 滿足對任意 $\mathbf{u}\in\mathbb{R}^2$ 皆有 $\mathbf{O}+\mathbf{u}=\mathbf{u}$.

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

(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 $\mathbf{u}+\mathbf{u}' = \mathbf{O}$. 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 $\mathbf{O}$ 滿足所求. 這裡我們只要令 $\mathbf{O}=(0,0)$, 則對任意向量 $\mathbf{u}=(a_1,a_2)$ 皆有

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

故知確實存在這樣的向量.

(4) 上面我們已知可令 $\mathbf{O}=(0,0)$, 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

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

(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 $1\mathbf{u}=\mathbf{u}$.

□

Question 1.1. 利用 $\mathbb{R}^2$ 向量加法的定義, 試證明以下性質:

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

(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 $\mathbf{u}+\mathbf{u}'=\mathbf{O}$.

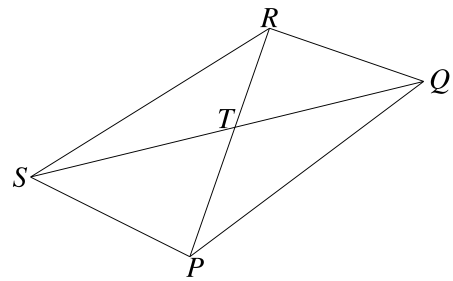
接下來我們將舉一個例子來說明我們可以用向量的性質 (Proposition 1.1.2) 來處理一些幾何的問題. 由於這些性質以後我們不會用到, 所以在此不用定理的形式呈現而是用例題的形式呈現.

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

$$\overrightarrow{PQ}+\overrightarrow{QR}=\overrightarrow{PS}+\overrightarrow{SR}. \tag{1.2}$$

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

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

現將上面式子 (1.2) 兩邊加上 $\mathbf{u}$ 可得

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

然而利用 Proposition 1.1.2 (2), (3) 我們有

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

同理可得

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

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 T 會是 $\overline{PR}$ 的中點, 也會是 $\overline{SQ}$ 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

$$\overrightarrow{PT'}=\overrightarrow{T'R}=\frac{1}{2}\overrightarrow{PR} \tag{1.4}$$

且

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

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

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

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 $(a_1,\ldots,a_n)$, 其中 $a_1,\ldots,a_n\in\mathbb{R}$. 我們說兩向量 $(a_1,\ldots,a_n)$ 以及 $(b_1,\ldots,b_n)$ 相等若且唯若 $a_1=b_1,\ldots,a_n=b_n$.

我們多說明一下符號的表示法, 這裡 $(a_1,\ldots,a_n)$ 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 $a_1,a_2,\ldots$ 都列出來, 所以我們就用 $(a_1,\ldots,a_n)$ 來表示. 這裡我們也沿習這樣省略的方法說 $a_1,\ldots,a_n\in\mathbb{R}$, 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 $a_1=b_1,\ldots,a_n=b_n$ 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 $\mathbf{u}=(a_1,\ldots,a_n),\mathbf{v}=(b_1,\ldots,b_n)\in\mathbb{R}^n$ 以及 $r\in\mathbb{R}$. 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, 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

$$\mathbf{u}=(1,1,2,2,3), \mathbf{v}=(5,4,3,2,1)\in\mathbb{R}^5$$

則

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

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 $\mathbf{O}\in\mathbb{R}^n$ 滿足對任意 $\mathbf{u}\in\mathbb{R}^n$ 皆有 $\mathbf{O}+\mathbf{u}=\mathbf{u}$.

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

(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 $\mathbf{O}$ 就是所謂 *additive identity*, 這裡其實就是向量 $(a_1,\ldots,a_n)$, 其中 $a_i = 0$, $\forall 1 \leq i \leq n$, 我們也稱此 $\mathbf{O}$ 為零向量. 另外若 $\mathbf{u} = (a_1,\ldots,a_n) \in \mathbb{R}^n$, 則 Proposition 1.2.3 (4) 所提的 $\mathbf{u}'$ 就是所謂 *additive inverse*, 這裡其實就是 $(b_1,\ldots,b_n)$, 其中 $b_i = -a_i$, $\forall 1 \leq i \leq n$. 如同 Question 1.1 所提, 利用 $\mathbb{R}^n$ 向量加法的定義, 我們知這裡 $\mathbf{O}$ 是 $\mathbb{R}^n$ 中唯一滿足對任意 $\mathbf{u} \in \mathbb{R}^n$ 皆有 $\mathbf{O}+\mathbf{u} = \mathbf{u}$ 的向量, 而給定 $\mathbf{u} \in \mathbb{R}^n$, $\mathbf{u}' = -1\mathbf{u}$ 是唯一滿足 $\mathbf{u}+\mathbf{u}' = \mathbf{O}$ 的向量. 事實上, 不必利用 $\mathbb{R}^n$ 向量加法的定義, 由 Proposition 1.2.3 (3) 的性質本身就可以確定 $\mathbf{O}$ 是唯一的. 這是因為若有另一個 $\mathbf{O}' \in \mathbb{R}^n$ 也是 additive identity (即滿足 Proposition 1.2.3 (3) 的性質), 則由 $\mathbf{O}$ 為 additive identity 可得 $\mathbf{O}+\mathbf{O}' = \mathbf{O}'$. 另一方面, 由 $\mathbf{O}'$ 為 additive identity 可得 $\mathbf{O}+\mathbf{O}' = \mathbf{O}$. 也就是說我們有 $\mathbf{O} = \mathbf{O}'$, 得到唯一性. 同樣的道理, 給定 $\mathbf{u} \in \mathbb{R}^n$, 若 $\mathbf{u}',\mathbf{u}''$ 皆滿足 Proposition 1.2.3 (3) 的性質, 即 $\mathbf{u}+\mathbf{u}' = \mathbf{u}+\mathbf{u}'' = \mathbf{O}$, 則得

$$\mathbf{u}' = \mathbf{u}' + \mathbf{O} = \mathbf{u}' + (\mathbf{u}+\mathbf{u}'') = (\mathbf{u}'+\mathbf{u}) + \mathbf{u}'' = \mathbf{O} + \mathbf{u}'' = \mathbf{u}''.$$

所以我們有以下之性質.

Corollary 1.2.4. *在 $\mathbb{R}^n$ 中存在唯一的向量 $\mathbf{O}$ 滿足對任意 $\mathbf{u} \in \mathbb{R}^n$ 皆有 $\mathbf{O}+\mathbf{u} = \mathbf{u}$. 另外, 給定 $\mathbf{u} \in \mathbb{R}^n$, 存在唯一的 $\mathbf{u}' \in \mathbb{R}^n$ 滿足 $\mathbf{u}+\mathbf{u}' = \mathbf{O}$.*

Remark 1.2.5. 在數學中, 一個定理若是由前一個定理直接得到 (不需引用新的性質), 通常會以 Corollary 稱之. Proposition 1.2.3 (3),(4) 其實談的是 additive identity 和 additive inverse 的存在性而 Corollary 1.2.4 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? 這也就是我們用 Corollary 稱之的用意, 就是要強調我們雖然可以用 $\mathbb{R}^n$ 的加法及係數積的定義得到 Corollary 1.2.4, 但這裡我們避開定義, 直接用 Proposition 1.2.3 的結果得到 Corollary 1.2.4 (以後我們在抽象的向量空間中就會這麼處理). 就如在前一節所述, 很多向量的運算性質, 不需回歸到定義, 利用 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.

有了唯一性, 以後我們將一律用 $\mathbf{O}$ 表示 $\mathbb{R}^n$ 中的 additive identity. 而給定 $\mathbf{u} \in \mathbb{R}^n$, 我們用 $-\mathbf{u}$ 表示 $\mathbf{u}$ 的 additive inverse.

我們利用 Proposition 1.2.3 (3) 有關 additive identity 的存在性證得 Corollary 1.2.4 中有關 additive identity 的唯一性. 證明過程中用了一個很重要的關鍵, 就是 Proposition 1.2.3 (3) 中要求 $\mathbf{O}$ 必須滿足對所有 $\mathbf{u} \in \mathbb{R}^n$ 皆滿足 $\mathbf{u}+\mathbf{O} = \mathbf{u}$. 如果當初僅只有特定 $\mathbf{u} \in \mathbb{R}^n$ 會滿足 $\mathbf{u}+\mathbf{O} = \mathbf{u}$, 則就無法證出 Corollary 1.2.4 中有關 additive identity 的唯一性了. 不過因為我們知道向量加法有 Proposition 1.2.3 這們完善的性質, 當我們要確認一個向量 $\mathbf{v}$ 是否為零向量時, 不必檢查所有的向量, 只要找到一個向量 $\mathbf{u}$ 滿足 $\mathbf{u}+\mathbf{v} = \mathbf{u}$, 就可以確認 $\mathbf{v}$ 是零向量了. 事實上我們有以下之結果.

Corollary 1.2.6. 假設 $\mathbf{v}\in\mathbb{R}^n$ 且存在 $\mathbf{u}\in\mathbb{R}^n$ 使得 $\mathbf{u}+\mathbf{v}=\mathbf{u}$, 則 $\mathbf{v}=\mathbf{O}$.

Proof. 由 $\mathbf{u}=\mathbf{u}+\mathbf{v}=\mathbf{v}+\mathbf{u}$, 將等式兩邊加上 $-\mathbf{u}$, 可得

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

□

這裡的邏輯順序一定要弄清楚. 我們是先知道 Proposition 1.2.3 是對的, 才能推得 Corollary 1.2.4 以及 Corollary 1.2.6. Corollary 1.2.6 告訴我們如果要確定一個 $\mathbb{R}^n$ 中的向量 $\mathbf{v}$ 是零向量, 我們只要在 $\mathbb{R}^n$ 中找到一個向量 $\mathbf{u}$, 使得 $\mathbf{v}+\mathbf{u}=\mathbf{u}$ 即可, 不必去試 $\mathbb{R}^n$ 上所有的向量. 所以我們馬上有以下的性質.

Corollary 1.2.7. 令 $\mathbf{v}\in\mathbb{R}^n$, 我們有以下之結果.

(1) $0\mathbf{v}=\mathbf{O}$.

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

Proof. 若直接用 $\mathbb{R}^n$ 上向量加法及係數積的定義, 馬上可得到此結果. 不過這裏我們利用 Proposition 1.2.3 以及其衍伸的 Corollaries 來證明.

(1) 利用 Corollary 1.2.6, 我們只要找到一個向量 $\mathbf{u}$ 滿足 $0\mathbf{v}+\mathbf{u}=\mathbf{u}$ 即可. 若考慮 $\mathbf{u}=\mathbf{v}$, 此時

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

得證 $0\mathbf{v}=\mathbf{O}$.

(2) 要證明 $(-1)\mathbf{v}=-\mathbf{v}$, 利用 Corollary 1.2.4 additive inverse 的唯一性我們只要證明 $(-1)\mathbf{v}+\mathbf{v}=\mathbf{O}$ 即可. 然而,

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

故由 (1) 的結果知 $(-1)\mathbf{v}+\mathbf{v}=\mathbf{O}$.

□

再次強調一次, Corollaries 1.2.4, 1.2.6, 1.2.7, 都可以很簡單的利用 $\mathbb{R}^n$ 上向量加法與係數積的定義得到. 它們很自然也很容易理解, 不必去死記. 這裡特別提到它們是要讓大家知道, 這些性質都可以直接由 Proposition 1.2.3 得到, 也讓大家更加確定 Proposition 1.2.3 中的性質足以使得 $\mathbb{R}^n$ 中向量的運算都可像實數一般處理. 同樣的, 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 $\mathbf{u}=\frac{1}{2}(\mathbf{w}-\mathbf{v})$.

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, 依向量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 *parallel* (平行), 用 $\mathbf{v} \parallel \mathbf{w}$ 來表示. 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 $\mathbf{v}_1, \mathbf{v}_2$ 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. 給定 $\mathbf{v}_1, \ldots, \mathbf{v}_m \in \mathbb{R}^n$, 以及 $r_1, \ldots, r_m \in \mathbb{R}$ 令

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

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

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

要注意因為 $\mathbf{v}_1, \ldots, \mathbf{v}_m$ 在 $\mathbb{R}^n$ 中, 所以 $\mathbf{v}_1, \ldots, \mathbf{v}_m$ 的一個線性組合仍在 $\mathbb{R}^n$ 中. 也就是說 $\mathrm{Span}(\mathbf{v}_1, \ldots, \mathbf{v}_m)$ 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 $\mathbf{O}$ 一定會在 $\mathrm{Span}(\mathbf{v}_1, \ldots, \mathbf{v}_m)$ 中 (即每個 r_i 皆取為 0), 又若 $\mathbf{w} \in \mathrm{Span}(\mathbf{v}_1, \ldots, \mathbf{v}_m)$, 則因 $\mathbf{w} = r_1\mathbf{v}_1 + \cdots + r_m\mathbf{v}_m$ 可知 $-\mathbf{w} = (-r_1)\mathbf{v}_1 + \cdots + (-r_m)\mathbf{v}_m$ 也就是說 $-\mathbf{w} \in \mathrm{Span}(\mathbf{v}_1, \ldots, \mathbf{v}_m)$. 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 $\mathbf{v}_1, \ldots, \mathbf{v}_m \in \mathbb{R}^n$, 則對任意 $\mathbf{u}, \mathbf{w} \in \mathrm{Span}(\mathbf{v}_1, \ldots, \mathbf{v}_m)$ 以及 $s, t \in \mathbb{R}$ 皆有

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

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

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

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

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

Question 1.2. 設 $\mathbf{v}_1,\ldots,\mathbf{v}_m\in\mathbb{R}^n$, 試證明若 $\mathbf{w}_1,\ldots,\mathbf{w}_k\in\mathrm{Span}(\mathbf{v}_1,\ldots,\mathbf{v}_m)$, 則

$$\mathrm{Span}(\mathbf{w}_1,\ldots,\mathbf{w}_k)\subseteq\mathrm{Span}(\mathbf{v}_1,\ldots,\mathbf{v}_m).$$

對於 span of vectors, 其中有一個基本的問題便是判斷哪些 vectors 會在特定 vectors 的 span 裡. 例如我們可以問 $(1,2,3)\in\mathbb{R}^3$ 是否屬於 $\mathrm{Span}((1,-1,2),(2,1,-2))$? 這樣的問題就等同於是否找得到 $r,s\in\mathbb{R}$ 使得 $(1,2,3)=r(1,-1,2)+s(2,1,-2)$. 比較向量各坐標, 我們得到一次聯立方程組

$$\left\{\begin{array}{rrcl} 1r & +2s & = & 1\\ -1r & +1s & = & 2\\ 2r & -2s & = & 3\end{array}\right..$$

所以知道這一類的問題會和解聯立方程組有關. 另外大家可以發現, 若將向量寫成直行 (即 column vector), $r(1,-1,2)+s(2,1,-2)=(1,2,3)$ 可寫成

$$r\begin{bmatrix}1\\-1\\2\end{bmatrix}+s\begin{bmatrix}2\\1\\-2\end{bmatrix}=\begin{bmatrix}1\\2\\3\end{bmatrix}.$$

這樣看成聯立方程組就更清楚方便了. 所以有的書都會將 vector 寫成 column vector 的樣子. 不過寫成 column vector 的缺點便是太占版面了, 本講義採 column vector 和 row vector (橫列) 混用的方法. 會視方便有時用 row vector 有時用 column vector. 這一點等介紹矩陣運算時我們會再釐清.

還有一個解決 column vector 占版面的方法就是利用所謂的 *standard basis*. 在 $\mathbb{R}^2$ 我們令 $\mathbf{i}=\begin{bmatrix}1\\0\end{bmatrix},\mathbf{j}=\begin{bmatrix}0\\1\end{bmatrix}$, 每一個 $\mathbb{R}^2$ 的 vector 都可以唯一寫成 $\mathbf{i},\mathbf{j}$ 的線性組合, 例如 $\begin{bmatrix}2\\3\end{bmatrix}$ 便可以寫成 $2\mathbf{i}+3\mathbf{j}$. 一般我們令 $\mathbb{R}^3$ 上的 standard basis vectors 為

$$\mathbf{i}=\begin{bmatrix}1\\0\\0\end{bmatrix},\mathbf{j}=\begin{bmatrix}0\\1\\0\end{bmatrix},\mathbf{k}=\begin{bmatrix}0\\0\\1\end{bmatrix}.$$

而 $\mathbb{R}^n$ 上的 standard basis vectors 定為

$$\mathbf{e}_1=\begin{bmatrix}1\\0\\0\\\vdots\\0\end{bmatrix},\mathbf{e}_2=\begin{bmatrix}0\\1\\0\\\vdots\\0\end{bmatrix},\ldots,\mathbf{e}_n=\begin{bmatrix}0\\0\\\vdots\\0\\1\end{bmatrix},$$

其中對任意 $1\leq i\leq n$, $\mathbf{e}_i$ 為第 i 個位置是 1, 其他位置為 0 的 column vector. 不難發現, 任意 $\mathbb{R}^n$ 的 vector 都可以唯一寫成 $\mathbf{e}_1,\ldots,\mathbf{e}_n$ 的 linear combination. 事實上 $\mathbb{R}^n$ 還有其他組和 standard basis vectors 有類似性質的 vectors, 我們留待以後再談.

1.4. Dot 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 $\mathbf{u}=(a_1,\ldots,a_n),\mathbf{v}=(b_1,\ldots,b_n)\in\mathbb{R}^n$. 則定義 $\mathbf{u},\mathbf{v}$ 的 *dot product* (*inner product*) 為

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

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 $\mathbf{u}=\mathbf{O}$.

(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 $\mathbf{u}=(a_1,\ldots,a_n),\mathbf{v}=(b_1,\ldots,b_n),\mathbf{w}=(c_1,\ldots,c_n)$. 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

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

這表示 $\mathbf{u}\cdot\mathbf{v}$ 是這些 a_ib_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_ib_i 皆等於 b_ia_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

$$\sum_{i=1}^{n}a_ib_i=\sum_{i=1}^{n}b_ia_i=\mathbf{v}\cdot\mathbf{u}.$$

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

(2) 依定義

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

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

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

反之若 $\mathbf{u} = (a_1,\ldots,a_n) = \mathbf{O}$ 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 $r\mathbf{u} = (ra_1,\ldots,ra_n)$ 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 $(ra_i)b_i = r(a_ib_i)$ (實數乘法結合律) 故知

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

再加上 $\sum_{i=1}^n r(a_ib_i)$ 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

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

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

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

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 $a_ib_i+a_ic_i$, 也就是說

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

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

$$\sum_{i=1}^n (a_ib_i+a_ic_i) = \sum_{i=1}^n a_ib_i + \sum_{i=1}^n a_ic_i = \mathbf{u}\cdot\mathbf{v}+\mathbf{u}\cdot\mathbf{w},$$

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 $\mathbf{O}$ 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 $\mathbf{v} = (a_1,\ldots,a_n) \in \mathbb{R}^n$, 我們定義 $\mathbf{v}$ 的長度 (*length*) 為

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

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. □

Question 1.3. 試證明平行四邊形定理 *(parallelogram relation):* 平行四邊形兩對角線長的平方和等於四邊長的平方和, 即對任意 $\mathbf{u},\mathbf{v}\in\mathbb{R}^n$ 皆有

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

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

$$\begin{aligned}\|\mathbf{u}_0+\mathbf{v}_0\|^2&=2+2\mathbf{u}_0\cdot\mathbf{v}_0,\\ \|\mathbf{u}_0-\mathbf{v}_0\|^2&=2-2\mathbf{u}_0\cdot\mathbf{v}_0.\end{aligned}\tag{1.6}$$

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

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

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

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

□

在證明 Proposition 1.4.5 時, 我們用了一個很特殊的技巧, 就是將 $\mathbf{u}$ 化成長度為 1 的向量 $\mathbf{u}_0 = \mathbf{u}/\|\mathbf{u}\|$. 一般來說一個長度為 1 的向量, 我們稱之為 *unit vector*. 任意的非零向量 $\mathbf{u}$ 都可以化成 unit vector, 即取 $\mathbf{u}_0 = \mathbf{u}/\|\mathbf{u}\|$. 這種化成 unit vector 的方法不只讓我們確定向量的長度且又保有原向量的方向性, 是線性代數處理內積有關的問題很好用的技巧.

利用 Proposition 1.4.5, 我們可以得到所謂的三角不等式.

Corollary 1.4.6 (Triangle inequality)**.** 若 $\mathbf{u},\mathbf{v} \in \mathbb{R}^n$, 則 $\|\mathbf{u}+\mathbf{v}\| \leq \|\mathbf{u}\|+\|\mathbf{v}\|$.

Proof. 由 Lemma 1.4.4 以及 Proposition 1.4.5, 我們有

$$\|\mathbf{u}+\mathbf{v}\|^2 = \|\mathbf{u}\|^2+2\mathbf{u}\cdot\mathbf{v}+\|\mathbf{v}\|^2 \leq \|\mathbf{u}\|^2+2\|\mathbf{u}\|\,\|\mathbf{v}\|+\|\mathbf{v}\|^2 = (\|\mathbf{u}\|+\|\mathbf{v}\|)^2.$$

不等式兩邊開根號得證 $\|\mathbf{u}+\mathbf{v}\| \leq \|\mathbf{u}\|+\|\mathbf{v}\|$. □

Question 1.4. 試找到充分必要條件使得 $\|\mathbf{u}+\mathbf{v}\| = \|\mathbf{u}\|+\|\mathbf{v}\|$.

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), 此時我們可以沿習 $\mathbb{R}^2,\mathbb{R}^3$ 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.

Example 1.4.7. 在 $\mathbb{R}^4$ 中設 $\mathbf{u} = (1,1,1,1), \mathbf{v} = (1,0,-2,-2)$ 的夾角為 θ, 則由

$$\cos\theta = \frac{\mathbf{u}\cdot\mathbf{v}}{\|\mathbf{u}\|\,\|\mathbf{v}\|} = \frac{-3}{2\times 3} = -\frac{1}{2},$$

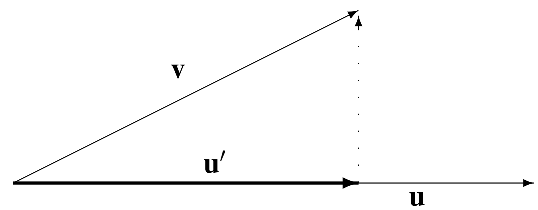
得知 $\theta = 120^\circ$.

利用夾角的定義我們進而定義出何謂「垂直」.

Definition 1.4.8. 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* 而較少用一般幾何上的 *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 $\mathbf{u}'$ 會落在 $\mathrm{Span}(\mathbf{u})$, 也就是說要找到 $r\in\mathbb{R}$ 使得 $\mathbf{u}'=r\mathbf{u}$, 且符合 $(\mathbf{v}-\mathbf{u}')\cdot\mathbf{u}=0$. 將 $\mathbf{u}'=r\mathbf{u}$ 代入得 $\mathbf{v}\cdot\mathbf{u}=r\mathbf{u}\cdot\mathbf{u}=r\|\mathbf{u}\|^2$, 亦即 $r=(\mathbf{v}\cdot\mathbf{u})/\|\mathbf{u}\|^2$. 也就是說, 若 $\mathbf{v}$ 在 $\mathbf{u}$ 的投影是存在的, 那們它的投影一定就是 $\frac{\mathbf{v}\cdot\mathbf{u}}{\|\mathbf{u}\|^2}\mathbf{u}$ (這說明了投影的唯一性). 然而若令 $r=(\mathbf{v}\cdot\mathbf{u})/\|\mathbf{u}\|^2$ (注意 $\mathbf{u}$ 為非零向量的假設), 則 $\mathbf{u}'=r\mathbf{u}$ 確實符合 $(\mathbf{v}-\mathbf{u}')\cdot\mathbf{u}=0$ (這說明了投影的存在性). 我們可以將以上的概念推廣到 $\mathbb{R}^n$ 的情形.

Proposition 1.4.9. 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. □

Question 1.5. 能否不用找到 r 的方法得到 *Proposition 1.4.9* 的唯一性?

Proposition 1.4.9, 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, 其中一個向量會落在 $\mathrm{Span}(\mathbf{u})$ (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. 我們稱落在 $\mathrm{Span}(\mathbf{u})$ 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).

Example 1.4.10. 在 $\mathbb{R}^4$ 中考慮 $\mathbf{u}=(1,1,1,1),\mathbf{v}=(1,0,-2,-2)$. 因 $\|\mathbf{u}\|=2$ 且 $\mathbf{v}\cdot\mathbf{u}=-3$, 得 $\mathbf{v}$ 在 $\mathbf{u}$ 的 projection 為

$$-\frac{3}{4}\mathbf{u}=-\frac{3}{4}(1,1,1,1).$$

又我們有

$$\mathbf{v} = (1,0,-2,-2) = -\frac{3}{4}(1,1,1,1) + (\frac{7}{4},\frac{3}{4},-\frac{5}{4},-\frac{5}{4}),$$

其中

$$-\frac{3}{4}(1,1,1,1) \in \mathrm{Span}((1,1,1,1)) \quad \text{and} \quad (\frac{7}{4},\frac{3}{4},-\frac{5}{4},-\frac{5}{4}) \cdot (1,1,1,1) = 0.$$

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.