Inner Product Space

在 $\mathbb{R}^2$ 和 $\mathbb{R}^3$ 中大家熟悉內積 (dot product) 的定義可以推廣到一般的 $\mathbb{R}^n$. 甚至可推廣到更一般的 vector space. 在一般的 vector space, 我們就不再用 dot product 而用所謂 *inner product* 來稱之. Inner product 可以幫助我們定義出 vector space 中許多重要的 subspaces. 具有 inner product 的 vector space 就稱為 *inner product space*. 在這一章中我們將介紹 inner product space 的性質.

4.1. Dot Product

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 4.1.1. 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 4.1.2. 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. 我們假設 $\mathbf{u}=(a_1,\ldots,a_n),\mathbf{v}=(b_1,\ldots,b_n),\mathbf{w}=(c_1,\ldots,c_n)$.

(1) 依定義 $\mathbf{u}\cdot\mathbf{v}=\sum_{i=1}^n a_ib_i$ 由於每一項 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 $\mathbf{u}=(a_1,\ldots,a_n)=(0,\ldots,0)=\mathbf{0}$. 反之若 $\mathbf{u}=(a_1,\ldots,a_n)=\mathbf{0}$ 表示對任意 $1\leq i\leq n$ 皆有 $a_i=0$, 故得 $\mathbf{u}\cdot\mathbf{u}=\sum_{i=1}^n a_ia_i=0$.

(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 4.1.2 (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 4.1.3. 令 $\mathbf{v}=(a_1,\ldots,a_n)\in\mathbb{R}^n$, 我們定義 $\mathbf{v}$ 的 *norm* (or *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 4.1.2 的處理一些有關於內積的性質, 而不必涉及內積的定義.

Lemma 4.1.4. 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 4.1.2 (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 4.1.2 (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 4.1. 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 4.1.4 僅用到內積的性質, 所以在一般的情形若我們不是利用 Definition 4.1.1 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 4.1.2 中的性質, 我們依然可得到 Lemma 4.1.4 中的性質. Lemma 4.1.4 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 4.1.5 (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 4.1.4 得知

$$\begin{aligned}\|\mathbf{u}_0+\mathbf{v}_0\|^2&=2+2\mathbf{u}_0\cdot\mathbf{v}_0, \qquad (4.1)\\ \|\mathbf{u}_0-\mathbf{v}_0\|^2&=2-2\mathbf{u}_0\cdot\mathbf{v}_0.\end{aligned}$$

因為 $\|\mathbf{u}_0+\mathbf{v}_0\|^2\geq 0$ 且 $\|\mathbf{u}_0-\mathbf{v}_0\|^2\geq 0$, 故由式子 (4.1) 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 (4.1) 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 4.1.2 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 4.1.5 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 4.1.5, 我們可以得到所謂的三角不等式.

Corollary 4.1.6 (Triangle inequality)**.** *若 $\mathbf{u},\mathbf{v}\in\mathbb{R}^n$, 則 $\|\mathbf{u}+\mathbf{v}\|\leq\|\mathbf{u}\|+\|\mathbf{v}\|$.*

Proof. 由 Lemma 4.1.4 以及 Proposition 4.1.5, 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 4.2. 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 4.1.5 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 4.1.7. 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

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

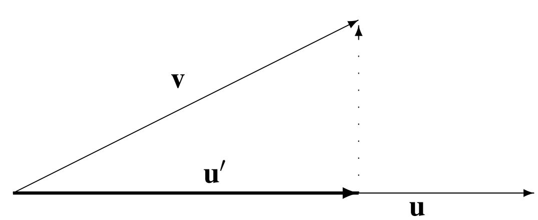
得知 $\theta = 120^\circ$.

利用夾角的定義我們進而定義出何謂「垂直」.

Definition 4.1.8. 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 4.1.9. 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 4.3. 能否不用找到 r 的方法得到 *Proposition 4.1.9* 的唯一性?

Proposition 4.1.9, 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 4.1.10. 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

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

4.2. Inner Product

在 $\mathbb{R}^m$ 中有關於 dot product 的性質, 可以推廣到一般 over $\mathbb{R}$ 的 vector space. 在一般的 vector space, 我們就不再用 dot product 而用所謂 *inner product* 來稱之. 由於我們要談的是一班的 inner product, 為了和原本 $\mathbb{R}^n$ 的 dot product 區分, 對於 $\mathbf{v}, \mathbf{w} \in V$, 原來我們用 $\langle \mathbf{v}, \mathbf{w} \rangle$ 來表示 $\mathbf{v}, \mathbf{w}$ 的 inner product.

考慮 over $\mathbb{R}$ 的 vector space V. 若對於任意 $\mathbf{v}, \mathbf{w} \in V$ 皆有 $\langle \mathbf{v}, \mathbf{w} \rangle \in \mathbb{R}$ 滿足 Proposition 4.1.2 有關內積的四個性質, 即

(1) $\langle \mathbf{v},\mathbf{w}\rangle = \langle \mathbf{w},\mathbf{v}\rangle$, $\forall \mathbf{v},\mathbf{w}\in V$.

(2) 對任意 $r\in\mathbb{R}$ 皆有 $\langle r\mathbf{v},\mathbf{v}\rangle = \langle \mathbf{v},r\mathbf{w}\rangle = r\langle \mathbf{v},\mathbf{w}\rangle$.

(3) $\langle \mathbf{u},\mathbf{v}+\mathbf{w}\rangle = \langle \mathbf{u},\mathbf{v}\rangle + \langle \mathbf{u},\mathbf{w}\rangle$, $\forall \mathbf{u},\mathbf{v},\mathbf{w}\in V$.

(4) $\langle \mathbf{v},\mathbf{v}\rangle \geq 0$ 且 $\langle \mathbf{v},\mathbf{v}\rangle = 0$ 若且唯若 $\mathbf{v}=\mathbf{0}$.

我們便稱 $\langle\, ,\, \rangle$ 為 V 上的一個 *inner product* 而稱 V 為 *inner product space.*

簡單的來說, 當我們稱 V 是一個 inner product space 時, 表示我們已給定 V 中的一個 inner product $\langle\, ,\, \rangle$. 例如在 $\mathbb{R}^n$ 上, 若我們定義 $\langle \mathbf{v},\mathbf{w}\rangle = \mathbf{v}\cdot\mathbf{w}$, $\forall \mathbf{v},\mathbf{w}\in\mathbb{R}^n$ (原先的 dot product), 那麼我們就說 $\mathbb{R}^n$ 是一個以 $\langle\, ,\, \rangle$ 為內積的 inner product space. 另外要注意的是, 依我們的定義, 我們考慮的 inner product 是實數, 所以只有在 V 是 vector space over $\mathbb{R}$ 時, 才會談論是否為 inner product space. 事實上還有定義在複數 $\mathbb{C}$ 上的 inner product, 不過由於本課程並不需要用到這種情況, 我們就略過不談.

接下來我們介紹一個定義在 $\mathsf{M}_{m\times n}(\mathbb{R})$ 這一個 vector space 的 inner product. 其實當 $M\in\mathsf{M}_{m\times n}(\mathbb{R})$, 我們可以將 M 視為一個在 $\mathbb{R}^{mn}$ 上的向量, 也就是說先寫第一個 column 再接著串接第二個 column, 這樣一直下去將 n 個 column 寫成一個長長的 column. 所以我們可以將兩個 $\mathsf{M}_{m\times n}(\mathbb{R})$ 上的矩陣看成是兩個 $\mathbb{R}^{mn}$ 的向量, 因此再利用 $\mathbb{R}^{mn}$ 上的 dot product 就可以定義 $\mathsf{M}_{m\times n}(\mathbb{R})$ 上的 inner product 了. 也就是說若 $A,B\in\mathsf{M}_{m\times n}$ 其中 $\mathbf{v}_1,\ldots,\mathbf{v}_n$ 依序為 A 的 column vectors 而 $\mathbf{w}_1,\ldots,\mathbf{w}_n$ 依序為 B 的 column vectors, 則可定義 $\langle A,B\rangle = \mathbf{v}_1\cdot\mathbf{w}_1+\cdots+\mathbf{v}_n\cdot\mathbf{w}_n$. 由於 dot product 符合 inner product 的性質, 所以我們知道這個方法確實給了 $\mathsf{M}_{m\times n}(\mathbb{R})$ 上的一個 inner product. 事實上若從矩陣的乘法來看, 我們可以將 $\mathbb{R}^n$ 上的 column vector 看成 $n\times 1$ matrix, 所以當 $\mathbf{v},\mathbf{w}\in\mathbb{R}^n$, 將 $\mathbf{v},\mathbf{w}$ 視為 $\mathsf{M}_{n\times 1}(\mathbb{R})$ 中的矩陣, 則 $\mathbf{v}\cdot\mathbf{w} = \mathbf{w}^{\mathrm{t}}\mathbf{v}$. 從這個角度來看若 $A,B\in\mathsf{M}_{m\times n}$ 其中 $\mathbf{v}_1,\ldots,\mathbf{v}_n$ 依序為 A 的 column vectors 而 $\mathbf{w}_1,\ldots,\mathbf{w}_n$ 依序為 B 的 column vectors, 則 $\mathbf{v}_i\cdot\mathbf{w}_i$ 就是 $B^{\mathrm{t}}A$ 這個矩陣對角線上 (i,i)-th entry. 我們曾定義過一個方陣 M 其對角線上的 entries 之和稱為 M 的 *trace*, 記為 $\mathrm{tr}(M)$. 因此上面定義的 $\langle A,B\rangle = \mathbf{v}_1\cdot\mathbf{w}_1+\cdots+\mathbf{v}_n\cdot\mathbf{w}_n$ 也等於 $\mathrm{tr}(B^tA)$. 也就是說對於 $A,B\in\mathsf{M}_{m\times n}(\mathbb{R})$, 定義 $\langle A,B\rangle = \mathrm{tr}(B^{\mathrm{t}}A)$ 就給了 $\mathsf{M}_{m\times n}(\mathbb{R})$ 上的一個 inner product, 在此 inner product 之下 $\mathsf{M}_{m\times n}(\mathbb{R})$ 就形成一個 inner product space. 有些同學或許會疑惑為何不定義 $\langle A,B\rangle = \mathrm{tr}(A^{\mathrm{t}}B)$ 呢? 事實上對於任意方陣 M, 皆有 $\mathrm{tr}(M) = \mathrm{tr}(M^{\mathrm{t}})$, 所以 $\mathrm{tr}(A^{\mathrm{t}}B) = \mathrm{tr}((A^{\mathrm{t}}B)^{\mathrm{t}}) = \mathrm{tr}(B^{\mathrm{t}}A)$. 不過以後當我們介紹更多 inner product 和矩陣關係時, $B^{\mathrm{t}}A$ 有其方便性, 所以一般來說我們會用 $\langle A,B\rangle = B^{\mathrm{t}}A$ 來表示這一個 inner product.

Question 4.4. *對任意 $A,B\in\mathsf{M}_{m\times n}(\mathbb{R})$, 證明 $\mathrm{tr}(AB) = \mathrm{tr}(BA)$. 並利用此證明 $\langle A,B\rangle = \mathrm{tr}(B^{\mathrm{t}}A), \forall A.B\in\mathsf{M}_{m\times n}(\mathbb{R})$ 是 $\mathsf{M}_{m\times n}(\mathbb{R})$ 上的 inner product.*

Question 4.5. *試說明 $\langle A,B\rangle = \det(B^{\mathrm{t}}A)$, $\forall A,B\in\mathsf{M}_{m\times n}(\mathbb{R})$ 是否為 $\mathsf{M}_{m\times n}(\mathbb{R})$ 的 inner product? 又若對任意 $M,M'\in\mathsf{M}_{n\times n}(\mathbb{R})$, 定義 $\langle M,M'\rangle = \mathrm{tr}(MM')$, 是否為 $\mathsf{M}_{n\times n}(\mathbb{R})$ 上的 inner product?*

對於次數小於等於 n 的實係數多項式所形成的向量空間 $\mathsf{P}_n(\mathbb{R})$, 也有一個有趣的 inner product, 我們看以下的例子.

Example 4.2.1. 選取 $n+1$ 個相異的實數 $c_1,\ldots,c_{n+1}$, 定義

$$\langle f,g\rangle = f(c_1)g(c_1)+\cdots+f(c_{n+1})g(c_{n+1}) = \sum_{i=1}^{n+1} f(c_i)g(c_i), \quad \forall f,g\in \mathsf{P}_n(\mathbb{R}).$$

我們說明在此定義之下 $\langle,\rangle$ 是 $\mathsf{P}_n(\mathbb{R})$ 的 inner product.

考慮 $f,g,h\in \mathsf{P}_n(\mathbb{R})$ 以及 $r\in\mathbb{R}$, 關於 $\langle f,g\rangle=\langle g,f\rangle$, $\langle f,g+h\rangle=\langle f,g\rangle+\langle f,h\rangle$ 以及 $\langle f,rg\rangle=r\langle f,g\rangle$ 等性質很容易由實數的乘法加法性質推得, 這裡就不再證明. 我們僅證明 $\langle f,f\rangle\geq 0$ 以及 $\langle f,f\rangle=0$ 若且唯若 $f=0$ 這個性質. 依定義 $\langle f,f\rangle=f(c_1)^2+\cdots+f(c_{n+1})^2$. 由於 $f(c_i)$ 是實數, 故知 $f(c_i)^2\geq 0, \forall i=1,\ldots,n+1$. 由此得證 $\langle f,f\rangle\geq 0$. 又若 $\langle f,f\rangle=0$, 表示 $f(c_i)^2=0, \forall i=1,\ldots,n+1$. 亦即 $c_1,\ldots,c_{n+1}$ 是 $f(x)=0$ 的 $n+1$ 個相異實根. 因此由 f 是次數小於等於 n 的多項式知, 唯一的可能就是 f 是零多項式, 即 $f=0$.

Question 4.6. 在 *Example 4.2.1* 中, 若僅選 n 個相異實數, 是否可定出 $\mathsf{P}_n(\mathbb{R})$ 的 *inner product?* 而若選 $n+2$ 個相異實數, 是否可定出 $\mathsf{P}_n(\mathbb{R})$ 的 *inner product?*

依照 inner product 的定義可以推導出一些 inner product 性質, 以下我們列出幾個常用的性質以利以後操作. 要注意以下的性質是可由 inner product 定義推得的, 它們並不屬於 inner product 的定義. 也就是說以後當我們要說明一個定義出來的是否為 inner product, 我們僅要驗證它是否符合 inner product 的四個要求. 若符合, 自然地它便會符合以下的性質.

Proposition 4.2.2. 假設 V 是以 $\langle,\rangle$ 為 *inner product* 的 *inner product space.* 則有以下的性質.

(1) $\langle \mathbf{v},\mathbf{0}\rangle=0$, $\forall \mathbf{v}\in V$.

(2) 若 $\mathbf{v},\mathbf{w}\in V$ 且 $\langle \mathbf{v},\mathbf{u}\rangle=\langle \mathbf{w},\mathbf{u}\rangle$, $\forall \mathbf{u}\in V$, 則 $\mathbf{v}=\mathbf{w}$.

Proof. (1) 任取 $\mathbf{v}\in V$, 由 $\langle \mathbf{v},\mathbf{0}\rangle=\langle \mathbf{v},\mathbf{0}+\mathbf{0}\rangle=\langle \mathbf{v},\mathbf{0}\rangle+\langle \mathbf{v},\mathbf{0}\rangle$, 得 $\langle \mathbf{v},\mathbf{0}\rangle=0$.

(2) 由 $\langle \mathbf{v},\mathbf{u}\rangle=\langle \mathbf{w},\mathbf{u}\rangle$ 得 $\langle \mathbf{v}-\mathbf{w},\mathbf{u}\rangle=0$. 由於此等式依假設是對所有的 $\mathbf{u}\in V$ 皆成立, 故令 $\mathbf{u}=\mathbf{v}-\mathbf{w}$, 可得 $\langle \mathbf{v}-\mathbf{w},\mathbf{v}-\mathbf{w}\rangle=0$. 故依 innner product 的定義知 $\mathbf{v}-\mathbf{w}=\mathbf{0}$, 得證 $\mathbf{v}=\mathbf{w}$. □

從 Proposition 4.2.2 (2) 的證明中, 我們了解到 inner product 的定義中 $\langle \mathbf{v},\mathbf{v}\rangle=0$ 若且唯若 $\mathbf{v}=\mathbf{0}$ 是相當重要的性質. 再加上 $\langle \mathbf{v},\mathbf{v}\rangle\geq 0$, 相當符合我們對長度的要求. 因此我們很自然地可以定義一個 inner product space 中的長度概念.

Definition 4.2.3. 假設 V 是以 $\langle,\rangle$ 為 inner product 的 inner product space. 對任意 $\mathbf{v}\in V$, 定義 $\mathbf{v}$ 的 *norm* (or *length*) 為

$$\|\mathbf{v}\|=\sqrt{\langle \mathbf{v},\mathbf{v}\rangle}.$$

稱之為 norm 自然會是符合長度定義的要求, 即三角不等式. 在介紹 dot product 時, 我們完整的推導出 norm 的性質, 當初證明這些性質因為僅用到 inner product 定義的基本性質, 這裡我們就不再證明, 僅列出其結果.

Proposition 4.2.4. 假設 V 是以 $\langle,\rangle$ 為 *inner product* 的 *inner product space* 且 $\|\,\|$ 為以 $\langle,\rangle$ 定義的 *norm,* 則對任意 $r\in\mathbb{R}$, $\mathbf{v},\mathbf{w}\in V$ 皆有以下的性質.

(1) $\|r\mathbf{v}\| = |r|\,\|\mathbf{v}\|$.

(2) $\|\mathbf{v}\|\geq 0$ 且 $\|\mathbf{v}\|=0$ 若且唯若 $\mathbf{v}=\mathbf{0}$.

(3) $\|\mathbf{v}+\mathbf{w}\|^2=\|\mathbf{v}\|^2+2\langle\mathbf{v},\mathbf{w}\rangle+\|\mathbf{w}\|^2$.

(4) $\|\mathbf{v}+\mathbf{w}\|^2+\|\mathbf{v}-\mathbf{w}\|^2=2\|\mathbf{v}\|^2+2\|\mathbf{w}\|^2$.

(5) $|\langle\mathbf{v},\mathbf{w}\rangle|\leq\|\mathbf{v}\|\,\|\mathbf{w}\|$. 特別地當 $\mathbf{v},\mathbf{w}$ 皆不為零向量時, $|\langle\mathbf{v},\mathbf{w}\rangle|=\|\mathbf{v}\|\,\|\mathbf{w}\|$ 若且唯若存在 $\lambda\in\mathbb{R}$ 使得 $\mathbf{w}=\lambda\mathbf{v}$.

(6) $\|\mathbf{v}+\mathbf{w}\|\leq\|\mathbf{v}\|+\|\mathbf{w}\|$.

將來用到 norm 的時候, 我們都會用到 Proposition 4.2.4 上的性質. 例如當 $\mathbf{v}\neq\mathbf{0}$ 時, 令 $\mathbf{u}=(1/\|\mathbf{v}\|)\mathbf{v}$, 則由 (1) 知
$$\|\mathbf{u}\|=\left\|\frac{1}{\|\mathbf{v}\|}\mathbf{v}\right\|=\left|\frac{1}{\|\mathbf{v}\|}\right|\|\mathbf{v}\|=1.$$
這種符合 $\|\mathbf{u}\|=1$ 的 $\mathbf{u}$, 就稱為在此 norm 之下的 *unit vector.* Proposition 4.2.4 的性質 (4) 稱為 parallelogram relation; (5) 就是 Cauchy-Schwarz inequality; (6) 就是 triangle inequality.

Question 4.7. 考慮 $\mathsf{P}_2(\mathbb{R})$ 上利用 $-1,0,1$ 三相異實數所訂出的 *inner product (*參見 *Example 4.2.1).* 試利用此 *inner product* 求 $\|x\|$ 並找到 $f\in\mathrm{Span}(x)$ 滿足 $\|f\|=1$.