Cauchy-Schwarz 不等式

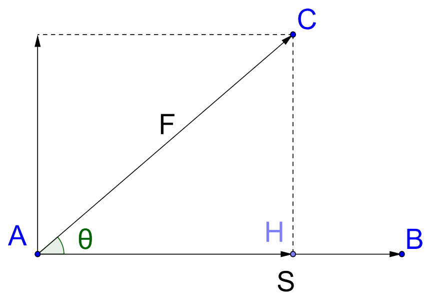
數學所碩二　黃名鉞

1　向量內積

在自然科學裡，很多物理量除了大小之外還帶有方向性，譬如說力、位移、速度及加速度等等，而在數學上相對應的就是向量的概念。我們已經定義了兩個向量相等的條件以及向量加減法的運算行為，在此我們想要討論的是兩個向量要如何相乘。首先，在物理上我們第一次遇到把兩個向量相乘，應該是在計算做功的時候。在直線運動中，做功 W 的公式為

$$W = F \times S$$

其中 F 是與運動方向同向的作用力大小，而 S 則是位移大小。若 F 與運動方向反向，則我們定義此時之做功為負。現在我們想要將作用力和位移的方向考慮進來，並將上面的公式推廣成對任意夾角的作用力和位移都能使用。考慮下圖，作用力向量 $\overrightarrow{F}$ 和位移向量 $\overrightarrow{S}$ 之間夾了一個 θ 角，其中 $0 \le \theta \le \pi/2$。



因為我們知道垂直位移的分力是不做功的，所以我們可以先將 $\overrightarrow{F}$ 分解成平行位移的分力 $\overrightarrow{F}_{/\!/}$ 以及垂直位移的分力 $\overrightarrow{F}_{\perp}$，此時我們的公式就變成了

$$W = \|\overrightarrow{F}_{/\!/}\| \times \|\overrightarrow{S}\| = \|\overrightarrow{F}\| \cos\theta \times \|\overrightarrow{S}\|$$

從數學的角度來看，這樣的向量乘法關心的其實是將其中一個向量「投影」到另一個向量的方向上之後的行為，於是我們可以定義

Definition 1. 給定兩非零向量 $\overrightarrow{u}, \overrightarrow{v} \in \mathbb{R}^3$，$\overrightarrow{u}$ 和 $\overrightarrow{v}$ 的內積定義為

$$\overrightarrow{u} \cdot \overrightarrow{v} = \|\overrightarrow{u}\|\,\|\overrightarrow{v}\| \cos\theta$$

其中 $0 \le \theta \le \pi$ 為 $\overrightarrow{u}$ 和 $\overrightarrow{v}$ 之間的夾角。此外，對任意向量 $\overrightarrow{w} \in \mathbb{R}^3$，我們定義

$$\overrightarrow{w} \cdot \overrightarrow{0} = \overrightarrow{0} \cdot \overrightarrow{w} = 0$$

在這裡我們要注意的是，兩個向量的內積是一個純量 (沒有方向性)，而且可以是任意正負實數 (正功與負功)。除此之外，當兩個非零向量的內積是 0 的時候，這兩個向量必然會垂直 ($\cos\theta = 0$)。

從上面的討論以及內積的定義可以發現，如果我們要讓做功的效果最大，最好的選擇就是 $\theta = 0$(最大正功) 以及 $\theta = \pi$(最大負功)，事實上在這兩個情況裡做功的絕對量都是兩個向量大小的乘積，只是差一個負號而已。而這個現象在數學上就被稱為 Cauchy-Schwarz 不等式。

Theorem 1. 給定 $\vec{u}, \vec{v} \in \mathbb{R}^3$，恆有

$$(\vec{u} \cdot \vec{v})^2 \leq \|\vec{u}\|^2 \|\vec{v}\|^2$$

其中等號成立若且唯若 $\vec{u}$ 和 $\vec{v}$ 平行，或者至少有一個 $\vec{u}, \vec{v}$ 是零向量。

Proof. 因為 $\cos^2\theta \leq 1$，故

$$(\vec{u} \cdot \vec{v})^2 = \|\vec{u}\|^2 \|\vec{v}\|^2 \cos^2\theta \leq \|\vec{u}\|^2 \|\vec{v}\|^2$$

其中等號成立若且唯若 $\cos^2\theta = 1$ 或者至少有一個 $\|\vec{u}\|, \|\vec{v}\|$ 是 0，也就是 $\theta = 0, \pi$ 或者至少有一個 $\vec{u}, \vec{v}$ 是零向量。 □

2 解析幾何

在上一節的討論中，我們已經知道了向量內積的概念。在這一節中，我們要引進座標系統，並看看在座標系統內向量內積的表現會是什麼。

Theorem 2. 給定 $(x_1, y_1, z_1), (x_2, y_2, z_2) \in \mathbb{R}^3$，恆有

$$(x_1x_2 + y_1y_2 + z_1z_2)^2 \leq (x_1^2 + y_1^2 + z_1^2)(x_2^2 + y_2^2 + z_2^2)$$

其中等號成立若且唯若

$$\frac{x_1}{x_2} = \frac{y_1}{y_2} = \frac{z_1}{z_2}$$

或者至少有一組 $x_i = y_i = z_i = 0$。

定理 2其實只是把定理 1用座標來表示而已，但是這卻是比較實用的版本。因為定理 1是對所有 $\mathbb{R}^3$ 中的向量都對，所以我們任意給定 6 個實數，就必須要滿足定理 2。

Exercise 1. 假設 x, y, z 為三個實數且滿足 $x + y + z = 1$，試求 $x^2 + y^2 + z^2$ 的最小值。

上面的這個練習同時可以讓我們看出一些 Cauchy-Schwarz 不等式的幾何解釋。給定空間中兩個非零向量 (a, b, c) 以及 (x, y, z)，如果我們把 x, y, z 滿足 $k = ax + by + cz$ 看成是一個動點 $P(x, y, z)$ 落在平面 $x + y + z = k$ 上，並把 $r^2 = x^2 + y^2 + z^2$ 看成是 P 和原點的距離平方，那麼根據點到平面的距離公式， r 必須要滿足

$$r \geq \frac{|-k|}{\sqrt{a^2 + b^2 + c^2}}$$

再將上式的兩邊同時平方即可得到 Cauchy-Schwarz 不等式。而等號成立勢必是發生在 (x, y, z) 和平面的法向量 (a, b, c) 平行 (也就是垂直平面) 的時候。另一個將 Cauchy-Schwarz 不等式與幾何物件連結的觀點是把 $x^2 + y^2 + z^2 = r^2$ 看成是一個球面方程式，因為 P 同時落在平面與球面上，所以球心 $(0, 0, 0)$ 到平面的距離必須要小於或等於球面的半徑 r，由此亦可導出上面 k 與 r 的大小關係。

3 歐氏空間

如果我們用代數的角度來看上面的結果，我們已經知道給定任意兩個長度是 2 或 3 的數列，就必須要滿足 Cauchy-Schwarz 不等式。那麼我們接下來的問題就是可不可以把這個不等式推廣到對任意兩個長度是 n 的數列都對？當 $n > 3$ 時我們就已經沒有具體的向量概念可以用了，因為我們根本就畫不出一個四維以上的空間。不過這個推廣其實是可以辦到的，而且只要用一點簡單的代數技巧就可以證明。

Theorem 3. 給定兩個實數序列 $(a_1, a_2, \cdots, a_n), (b_1, b_2, \cdots, b_n)$，恆有

$$\left(\sum_{i=1}^{n} a_i b_i\right)^2 \leq \left(\sum_{i=1}^{n} a_i^2\right)\left(\sum_{i=1}^{n} b_i^2\right)$$

其中等號成立若且唯若存在某個 c 使得 $a_i/b_i = c$ 對所有 i，或者至少有一個數列的每一項都是 0。

Proof. 考慮 $f(t) = \sum_{i=1}^{n}(a_i - tb_i)^2 = \sum_{i=1}^{n} a_i^2 - 2t\sum_{i=1}^{n} a_i b_i + t^2 \sum_{i=1}^{n} b_i^2$。因為 $f(t) \geq 0$，所以 $f(t) = 0$ 的判別式必定小於或等於 0，進而該不等式得證。其中等號成立若且唯若存在 c 使得 $f(c) = 0$，也就是 $a_i/b_i = c$ 對所有 i，或者至少有一個數列的每一項都是 0。 □

有了定理 3之後我們可以反過來問，是否可以將 $\mathbb{R}^3$ 中的向量性質全部推廣到 $\mathbb{R}^n$？首先，關於平移、伸縮以及加減法的性質基本上都可以直接類推 (各個分量比較與加減)。但是內積這個概念就比較麻煩，因為我們沒有夾角的概念，所以不能用直覺的投影來定義。但是定理 3告訴我們，如果我們模仿 $\mathbb{R}^3$ 中的情況定義

$$\|(a_1, a_2, \cdots, a_n)\| = \sqrt{\sum_{i=1}^{n} a_i^2}$$

$$(a_1, a_2, \cdots, a_n) \cdot (b_1, b_2, \cdots, b_n) = \sum_{i=1}^{n} a_i b_i$$

那麼我們就可以定義一個想像的餘弦值

$$\cos\theta = \frac{\sum_{i=1}^{n} a_i b_i}{\sqrt{\sum_{i=1}^{n} a_i^2}\sqrt{\sum_{i=1}^{n} b_i^2}}$$

因為 Cauchy-Schwarz 不等式保證了該比例的絕對值恆小於或等於 1，所以這個想像的夾角是合理的，也就是說如果我們想判斷 $\mathbb{R}^n$ 中的兩個向量貼得有多近，就只要看看兩者的內積和長度的比值就可以了。

有了這些概念之後，許多幾何物件也可以一起推廣到高維度的空間中，譬如說相對應 $\mathbb{R}^3$ 中的平面 $ax+by+cz=k$，在 $\mathbb{R}^n$ 中我們一樣可以考慮所謂的 affine hyperplane

$$E : a_1x_1 + a_2x_2 + \cdots + a_nx_n = k$$

這樣的 affine hyperplane 跟三維空間平面有許多一樣的性質，譬如說 $(a_1, a_2, \cdots, a_n)$ 是 E 的一個法向量，也就是說它和所有平面上兩點所形成的向量內積都是 0。此外，若給定 $\mathbb{R}^n$ 中一點 $P(p_1, p_2, \cdots, p_n)$，則 P 到 E 的最短距離就是

$$\frac{\left|\sum_{i=1}^n a_ip_i - k\right|}{\sqrt{\sum_{i=1}^n a_i^2}}$$

此外，若是我們考慮一個 $\mathbb{R}^n$ 中的 n 維球面

$$x_1^2 + x_2^2 + \cdots + x_n^2 = r^2$$

那麼利用 Cauchy-Schwarz 不等式可以證出，在 S 和 E 有交集的情況下， E 和球心的距離必然要小於或等於球面的半徑 r，而當等號成立的時候我們就稱 S 和 E 相切。