

凸函數及蕭爾定理

演講者：張福春教授
中山大學應用數學系
changfc@math.nsysu.edu.tw
日期：2004.06.21

1 凸函數與凹函數

在解析幾何中，我們常用 X, Y 兩個座標表示兩個數 x, y 的關係。通常我們稱 x 為變數 y 為 x 之函數。有時我們用 $y = f(x)$ 來表示 y 與 x 的關係。例如

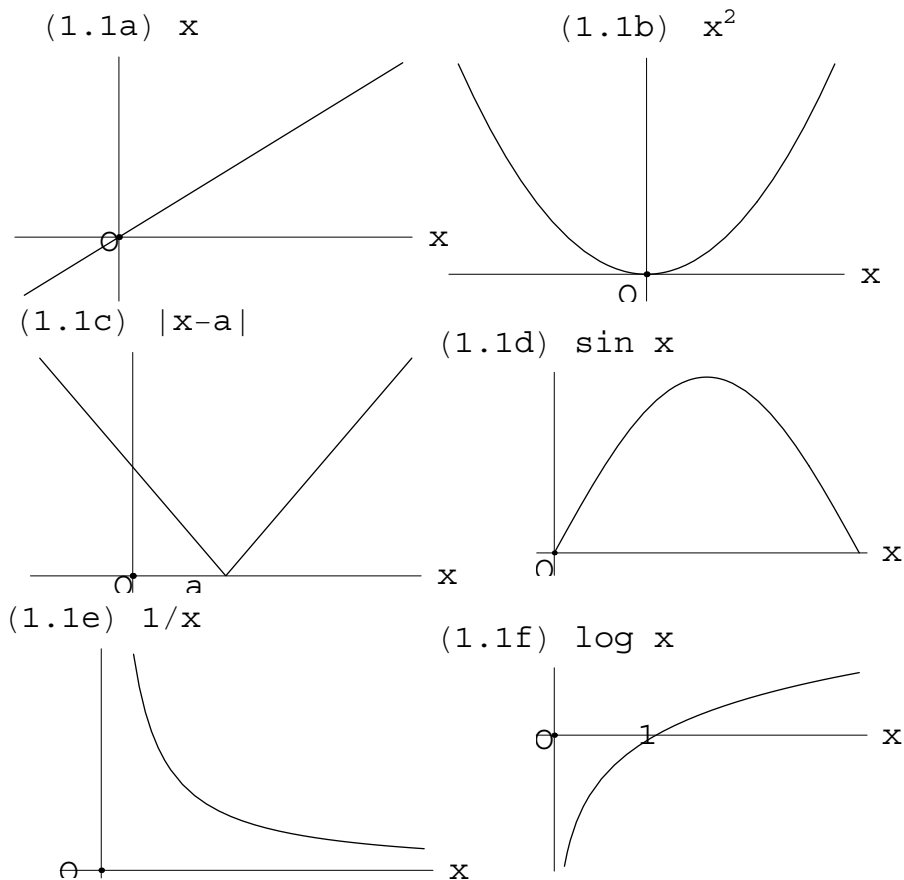
$$f(x) = x^2 + 2$$

$$f(x) = \sin x + \tan x$$

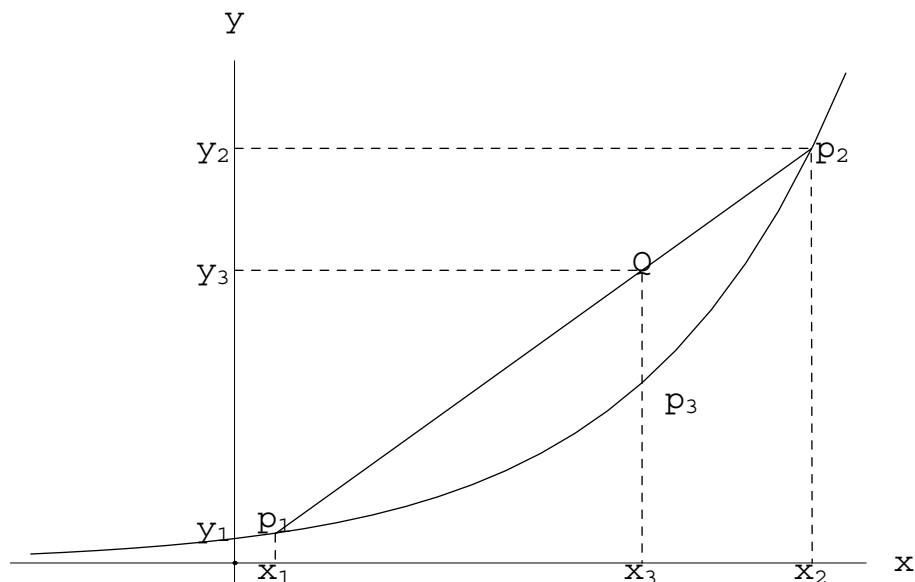
$$f(x) = x \log x$$

$$f(x) = 2^x$$

都是 x 的函數。我們要記住這裏的 x 在三角函數時代表的是弧度。(參看3.3)在圖1.1 a - f 中我們畫了一些常用函數的圖形。



從上面六個圖形我們可以看到有些 $f(x)$ 的圖形像一個向上放著的碗(如圖1.1b, c, e)，及有些圖形像一個蓋著的碗(例如1.1d, f)，我們把第一類函數稱之為凸函數(對 x 軸而言是凸的)，而第二類的函數稱為凹函數。其中圖1.1a， $f(x) = x$ 是這兩類函數的媒介，我們把它包含在凸函數之內，同時也包含在凹函數之內。



從圖1.2我們不難看出凸函數有這樣的性質。

定義1.1 一個函數 $y = f(x)$ 在 $x \in (a, b)$ 區間稱之為凸函數，如果對任何 $y = f(x)$ 上兩點 P_1 與 P_2 ，連接 P_1 與 P_2 的直線都高於在 P_1 與 P_2 間的曲線 $y = f(x)$ 。換言之，對任何 $0 \leq \lambda \leq 1$

$$f(\lambda x_1 + (1 - \lambda)x_2) \leq \lambda f(x_1) + (1 - \lambda)f(x_2) \quad (1)$$

註：在圖1.2中，若令 $x_3 = \lambda x_1 + (1 - \lambda)x_2$ ，則(1)表示

$$f(x_3) \leq \lambda y_1 + (1 - \lambda)y_2 = y_3,$$

及 Q 點高於 P 點。

同理我們定義 $f(x), x \in (a, b)$ 為一凹函數如果對所有的 $x_1, x_2 \in (a, b), 0 \leq \lambda \leq 1$ 而言：

$$f(\lambda x_1 + (1 - \lambda)x_2) \geq \lambda f(x_1) + (1 - \lambda)f(x_2) \quad (2)$$

都成立。

例1.1 證明 $f(x) = x^2$ 在 $x \in R$ 上為一凸函數。

證 根據(1)式，我們需要證明對任何 $x_1, x_2 \in R$ 及 $0 \leq \lambda \leq 1$

$$(\lambda x_1 + (1 - \lambda)x_2)^2 \leq \lambda x_1^2 + (1 - \lambda)x_2^2. \quad (3)$$

展開簡化(3)之後可知(3)與 $\lambda(1 - \lambda)(x_1 - x_2)^2 \geq 0$ 相同。故(3)式顯然成立，即 $f(x) = x^2$ 為一凸函數。

有時候要證明(1)並不容易，下面的定理可以簡化許多凸函數的證明。

定理1.1 若 $y = f(x)$, $x \in (a, b)$ 滿足下列的二個條件

- (i) 對任何 (a, b) 之間的三點 x_1, x_2, x_3 , 若 $x_1 < x_2 < x_3$, 則 $f(x_2)$ 不可能同時大於 $f(x_1)$ 及 $f(x_3)$, 並
- (ii) 對任何 $x_1, x_2 \in (a, b)$

$$f\left(\frac{x_1 + x_2}{2}\right) \leq \frac{1}{2}[f(x_1) + f(x_2)],$$

則 $f(x)$ 為一凸函數。

這個定理從圖型上很容易看出來，但要證明它必須用到連續的觀念，我們把它的證明從略。定理1.1的(i)往往很容易證明，如果一個函數是漸增，如 $f(x) = x^3, x \in (0, \infty)$ ；漸減，如 $f(x) = \frac{1}{x}, x \in (0, \infty)$ ，或先漸減而後增加，如 $f(x) = x^2, x \in (-\infty, \infty)$ ，(i)都成立。只是條件(ii)往往要費比較多的工夫去證明。

我們若要證明一個函數為凹函數，我們只要證明這個函數的負值是凸函數就可以了。

定理1.2 若 $f(x)$ 在 $x \in (a, b)$ 區間為一凹函數，則 $y = -f(x)$ 在同一區間為凸函數，反之亦然。

例1.2 試證 $f(x) = \sin x, x \in (0, \pi)$ 為一凹函數。

證 若 $x_1, x_2 \in (0, \pi)$ ，則 $-\pi < x_1 - x_2 < \pi, 0 < \cos \frac{x_1 - x_2}{2} < 1$ ，故

$$\frac{1}{2}(\sin x_1 + \sin x_2) = \sin \frac{x_1 + x_2}{2} \cos \frac{x_1 - x_2}{2} \leq \sin \frac{x_1 - x_2}{2} \quad (4)$$

又因 $\sin x$ 在 $x \in (0, \pi)$ 只有在 $x = \pi/2$ 時取極大值，故 $f(x) = -\sin x$ ，滿足定理1.1之(i)，又(4)乘負號之後滿足定理1.1之(ii)，故 $-\sin x$ 為一凸函數，即 $f(x) = \sin x$ 為一凹函數。

因凸凹函數有一個這樣簡單的關係，我們若知道凸，凹函數中一種的性質，必定也知道另一種了。在本書我們將只討論凸函數。

2 蕭爾(Schur)定理

在前面幾章我們看到一般不等式的證法都各有巧妙，似乎無一個規則可尋。蕭爾(Schur)在1923年發現了一個新的角度去看不等式，一下子歸納了並簡化了許多不等式的證明，要了解他的貢獻，我們必須下一點工夫去熟悉一個新的觀念。

我們知道如果一個不等式只包括一個變數，則其證法就相當容易。例如說 $f(x)$ 是一個漸增的函數，則 $x > y$ 就產生 $f(x) > f(y)$ 。譬如說

$$\sin \frac{\theta}{2} \leq \sin \theta \quad 0 < \theta < \frac{\pi}{2}; \quad 3 \log 3 \geq 2 \log 2$$

但一般我們想求證的不等式往往包含一個數列 x_1, x_2, \dots, x_n 。蕭爾想到了一個比較兩數列“大，小”的方法，使得許多不等式的證明就像證明 $x \leq y$ 則 $f(x) \leq f(y)$ 一樣的容易。

在本章裏我們用大寫的 X, Y, Z ，等表示數列，而小寫的 x, y, z 表示數列裡的元素，即

$$X = (x_1, x_2, \dots, x_n), \quad Y = (y_1, y_2, \dots, y_n), \quad Z = (z_1, z_2, \dots, z_n)$$

定義2.1 如果 $(x_{(1)}, x_{(2)}, \dots, x_{(n)})$ 是一個 x_1, x_2, \dots, x_n 的重組且 $x_{(1)} \geq x_{(2)} \geq \dots \geq x_{(n)}$ ，則 $(x_{(1)}, x_{(2)}, \dots, x_{(n)})$ 定義為實數列 $X = (x_1, x_2, \dots, x_n)$ 的一個漸減重組。

例如 $(1, -2, 4, 3, 6, 3)$ 的漸減重組是 $(6, 4, 3, 3, 1, -2)$ 。下面是蕭爾氏比較兩數列的方法。

定義2.2 $(Y \prec X)$

令 $x_{(1)} \geq x_{(2)} \geq \dots \geq x_{(n)}, y_{(1)} \geq y_{(2)} \geq \dots \geq y_{(n)}$ 分別表示兩實數 $X = (x_1, x_2, \dots, x_n), Y = (y_1, y_2, \dots, y_n)$ 的漸減重組，若

$$\begin{aligned} y_{(1)} &\leq x_{(1)} \\ y_{(1)} + y_{(2)} &\leq x_{(1)} + x_{(2)} \\ &\vdots \\ y_{(1)} + y_{(2)} + \dots + y_{(n-1)} &\leq x_{(1)} + \dots + x_{(n-1)} \\ y_{(1)} + y_{(2)} + \dots + y_{(n)} &= \sum_{i=1}^n y_i = \sum_{i=1}^n x_i = x_{(1)} + x_{(2)} + \dots + x_{(n)}. \end{aligned}$$

則我們定義 X 可以罩得住 Y ，並以 $Y \prec X$ 表示之。

例如 $X = (1, 2, 3), Y = (2.5, 1.2, 2.3)$ 兩數列，因

$$\begin{aligned} 3 &> 2.5 \\ 3 + 2 &> 2.5 + 2.3 \\ 3 + 2 + 1 &= 2.5 + 2.3 + 1.2 \end{aligned}$$

我們即說 X 可以罩住 Y ，亦得 $Y \prec X$ 。

例2.1 令 $X = (x_1, x_2, \dots, x_n)$ 為一實數列且定 $\bar{x} = (x_1 + x_2 + \dots + x_n)/n$ ，及 $Y = (\bar{x}, \bar{x}, \dots, \bar{x})$ 為一含 n 個 \bar{x} 的數列，則 $Y \prec X$ 。

證 很顯然 $\sum x_i = \sum y_i = n\bar{x}$ 。令 $x_{(1)} \geq x_{(2)} \geq \dots \geq x_{(n)}$ 。為 X 的漸減重組。則

$$\begin{aligned} x_{(1)} &= \frac{1}{n}(x_{(1)} + \dots + x_{(1)}) \geq \frac{1}{n}[x_{(1)} + x_{(2)} + \dots + x_{(n)}] = \bar{x} \\ x_{(1)} + x_{(2)} &= \frac{1}{n}[x_{(1)} + x_{(2)} + x_{(1)} + x_{(2)} + \dots + x_{(1)} + x_{(2)}] \\ &\geq \frac{1}{n}(x_{(1)} + x_{(2)} + x_{(1)} + x_{(2)} + 2x_{(3)} + \dots + 2x_{(n)}) \\ &= 2\bar{x} \\ &\vdots \end{aligned}$$

一般而言

$$\begin{aligned} &x_{(1)} + x_{(2)} + \dots + x_{(k)} \\ &= \frac{1}{n}[k(x_{(1)} + \dots + x_{(k)}) + (n-k)(x_{(1)} + x_{(2)} + \dots + x_{(k)})] \\ &\geq \frac{1}{n}[k(x_{(1)} + \dots + x_{(k)}) + k(x_{(k+1)} + \dots + x_{(n)})] = k\bar{x} \end{aligned}$$

故 $Y \prec X$ 。下面這個定理把單得住的觀念與凸函數的關係做了一個引線。

定理2.1 令 $X = (x_1, x_2, \dots, x_n), Y = (y_1, y_2, \dots, y_n)$ 為兩實數列，且 X 與 Y 只有兩個元素不相同，設其位置在 i 與 j 上。(即 $x_k = y_k$ 除非 $k = i$ 或 j)。則下列(i)與(ii)互為充要條件。

(i) $X \prec Y$

(ii) 有一個數 $0 < \lambda < 1$ ，且 $y_i = \lambda x_i + (1 - \lambda)x_j$

$$y_j = (1 - \lambda)x_i + \lambda x_j$$

證 因 i 與 j 可在任何位置，因此我們不妨假設 $x_1 \geq x_2 \geq \dots \geq x_i \geq \dots \geq x_j \geq \dots \geq x_n$ 。我們先由(i)推出(ii)。

因除了在 i, j 位置上的 x, y 皆相等又 $Y \prec X$ 則必須有

$$x_i > y_i \quad x_i + x_j = y_i + y_j$$

亦即 $x_i > y_i \geq y_j > x_j$ 。若令

$$\lambda = \frac{y_i - x_i}{x_i - x_j}$$

則 $1 > \lambda > 0$ 並很容易帶入而證得(ii)。

現若(ii)成立，則由於在 i 之前 x, y 皆相等，故對 $k < i$ 而言，

$$x_1 + x_2 + \dots + x_k = y_1 + y_2 + \dots + y_k$$

又在 $k = i$ 時

$$\begin{aligned} y_1 + y_2 + \dots + y_i &= x_1 + \dots + x_{i-1} + \lambda x_i + (1 - \lambda)x_j \\ &\leq x_1 + \dots + x_{i-1} + \lambda x_i + (1 - \lambda)x_i \\ &= x_1 + x_2 + \dots + x_i \end{aligned}$$

這種不等的情形一直到 $k \leq j$ 時又恢復了等式(因為 $y_i + y_j = x_i + x_j$)故本定理成立。

定理2.2 令實數列 $X = (x_1, x_2, \dots, x_n), Y = (y_1, y_2, \dots, y_n)$ 且 X 與 Y 不全相同並 $Y \prec X$ 。則我們可找到一個數列 $Z = (z_1, z_2, \dots, z_n)$ 具備下面三個性質。

(i) $Y \prec Z \prec X$ (即 $Y \prec Z$ 且 $Z \prec X$)

(ii) Z 與 Y 比 X 與 Y 至少多一個相同的元素

(iii) Z 與 X 最多只有兩個元素不同。

證 因本定理之證明與 X, Y 之排列法無關，不妨假定 X, Y 已成漸減排列，即 $X : x_1 \geq x_2 \geq \dots \geq x_n, Y : y_1 \geq y_2 \geq \dots \geq y_n$ 因 $Y \prec X$ ，故 $\sum x_i = \sum y_j$ 。又因 x_i 與 y_i 不全相同，

證 由定理2.2的系我們只需證明 Y 與 X 只有兩個元素不相同的情形就行了。(否則我們可由 Z_1, Z_2, \dots, Z_s 一直推下去)。令 x_i, x_j 與 y_i, y_j 不相同而其餘全相同，則(1)變成了

$$f(y_i) + f(y_j) \leq f(x_i) + f(x_j) \quad (2)$$

因 $Y \prec X$ ，由定理2.1知道我們可以找到一個 $0 < \lambda < 1$ ，且

$$y_i = \lambda x_i + (1 - \lambda)x_j$$

$$y_j = (1 - \lambda)x_i + \lambda x_j$$

由 $f(x)$ 之凸性質得

$$\begin{aligned} f(y_i) + f(y_j) &= f(\lambda x_i + (1 - \lambda)x_j) + f((1 - \lambda)x_i + \lambda x_j) \\ &\leq \lambda f(x_i) + (1 - \lambda)f(x_j) + (1 - \lambda)f(x_i) + \lambda f(x_j) \\ &= f(x_i) + f(x_j) \end{aligned}$$

故(2)得證。

現在我們將看定理2.3有什麼用處。

例2.2 因 $f(x) = \log x, 0 < x < \infty$ 是凹函數(本章習題1)，則 $g(x) = -f(x)$ 在 $x \in (0, \infty)$ 為一凸函數。由例2.1知 $(\bar{x}, \bar{x}, \dots, \bar{x}) \prec (x_1, x_2, \dots, x_n)$ ，故由定理2.3可得

$$\sum_{i=1}^n (-\log \bar{x}) \leq -\sum_{i=1}^n \log x_i \quad (x_i > 0, i = 1, 2, \dots, n)$$

亦即

$$\bar{x} \geq \left(\prod_{i=1}^n x_i \right)^{\frac{1}{n}}$$

這就是幾何與算術平均數的不等式。

例2.3 因 $f(x) = 1/x$ 為一凸函數(本章習題1)則

$$\sum_{i=1}^n \frac{1}{\bar{x}} \leq \sum_{i=1}^n \frac{1}{x_i} \quad x_i > 0$$

$$\text{即 } \sum_{i=1}^n \frac{1}{x_i} \geq \frac{n}{\bar{x}}$$

例2.4 有 n 個數 $0 \leq x_i < 1$ 且 $\sum_{i=1}^n x_i = 1$ ，試證

$$\sum_{i=1}^n \frac{x_i}{1 - x_i} \geq \frac{n}{1 - n}$$

證 我們先證 $f(x) = \frac{x}{1-x}$ ，在 $0 < x < 1$ 之間為凸函數。經過通分簡化可知

$$f\left(\frac{1}{2}(x_1 + x_2)\right) \leq \frac{1}{2}[f(x_1) + f(x_2)]$$

即

$$\frac{(x_1 + x_2)/2}{1 - (x_1 + x_2)/2} \leq \frac{1}{2} \left(\frac{x_1}{1 - x_1} + \frac{x_2}{1 - x_2} \right)$$

與 $(x_1 + x_2)^2 - 4x_1x_2 \geq 0$ 相同。既證得 $f(x)$ 為凸函數，又因 $\bar{x} = \frac{1}{n}$ ，故 $(\frac{1}{n}, \frac{1}{n}, \dots, \frac{1}{n}) \prec (x_1, x_2, \dots, x_n)$ ，由定理2.3可得

$$\sum_{i=1}^n \frac{x_i}{1 - x_i} \geq n \cdot \frac{1/n}{1 - (1/n)} = \frac{n}{1 - n}。$$

例2.5 設 $x_i, i = 1, 2, \dots, n$ 為 n 個實數， a 為一定數。令 $\bar{x} = \sum x_i/n$ ，試證

$$\sum_{i=1}^n |x_i - a| \geq n|\bar{x} - a|$$

證 很顯然，由於 $f(x) = |x - a|$ 是凸函數及 $(\bar{x}, \bar{x}, \dots, \bar{x}) \prec (x_1, x_2, \dots, x_n)$ 。

例2.6 由例2.1知 $\sin x$ 在 $x \in (0, \pi)$ 之間為凹函數，又在一三角形中，若以 A, B, C 表示其三頂角，則很容易求得

$$\left(\frac{\pi}{3}, \frac{\pi}{3}, \frac{\pi}{3}\right) \prec (A, B, C) \prec (\pi, 0, 0)$$

故 $0 \leq \sin A + \sin B + \sin C \leq 3 \sin \frac{\pi}{3} = \frac{3\sqrt{3}}{2}$ 。又 $1 \leq \sin \frac{A}{2} + \sin \frac{B}{2} + \sin \frac{C}{2} \leq 3 \sin \frac{\pi}{2} = \frac{3}{2}$ 。

例2.7 令 $f(x) = \log \sin x, 0 < x < \pi$ 。則 $f(x)$ 的凹函數性可由 $\sin x_1 \sin x_2 \leq \sin^2 \frac{x_1+x_2}{2}$ 即 $\cos(x_1 - x_2) \leq 1$ 證得。由於 $(\frac{\pi}{3}, \frac{\pi}{3}, \frac{\pi}{3})$ 可證得

$$\log \sin A + \log \sin B + \log \sin C \leq \log \left(\sin \frac{\pi}{3} \right)^3$$

即 $\sin A \sin B \sin C \leq \frac{3\sqrt{3}}{8}$ 。

蕭爾定理的用途可以推廣到很多超出本書範圍的不等式，一本很好的參考書是 Marshall 與 Olkin 所著 "Inequalities: Theory of Majorization and Its Applications" 1979 美國 Academic Press 出版。

1. 試鑑別下列各函數之凹凸性，

(i) $f(x) = x, -\infty < x < \infty$ 。

(ii) $f(x) = |x - a|, -\infty < x < \infty$ 。

(iii) $f(x) = 1/x, 0 < x < \infty$ 。

(iv) $f(x) = \log x, 0 < x < \infty$ 。

2. 設 a, b, c 為三正數，令 $x = b + c - a, y = x + a - b, z = a + b - c$ 。求證

$$abc(xy + yz + zx) \geq xyz(bx + ca + ab)$$

3. 設 $0 < x_i < 1, i = 1, 2, \dots, n$ 且 $\sum_{i=1}^n x_i = 1$ 。試證

(i) $\prod_{i=1}^n \left(\frac{1+x_i}{1-x_i} \right) \geq \left(\frac{n+1}{n-1} \right)^n$

(ii) $\prod_{i=1}^n \left(1 + \frac{1}{x_i} \right) \geq (n+1)^n$

(iii) $\prod_{i=1}^n \left(1 - \frac{1}{x_i} \right) \geq (n-1)^n$

[註:(i)在1975年被Klamkin發現，(ii)與(iii)在1970年發現]

4. 令 $(x_1, x_2, \dots, x_n), (y_1, y_2, \dots, y_n)$ 為二組實數列令 $x_{(1)} \geq x_{(2)} \geq \dots \geq x_{(n)}, y_{(1)} \geq y_{(2)} \geq \dots \geq y_{(n)}$ 分別表示它們的漸減重組。試證

$$\sum_{i=1}^n |x_i - y_i| \leq \sum_{i=1}^n |x_{(n)} - y_{(n-i+1)}|$$

5. 試證一定圓內接 n 邊形中，以正 n 邊形的面積最大。

6. 設 $x_i > 0, i = 1, 2, \dots, n$ ，令

$$\begin{aligned} y_1 &= \frac{x_1 + x_2}{2} \\ y_2 &= \frac{x_2 + x_3}{2} \\ &\vdots \\ y_{n-1} &= \frac{x_{n-1} + x_n}{2} \\ y_n &= \frac{x_n + x_1}{2} \end{aligned}$$

試證 $y_1 \cdot y_2 \cdots y_n \geq x_1 \cdot x_2 \cdots x_n$ (此不等式在1969被Daykin所發現。)

7. 設 a, b, c 與 a', b', c' 為同周長兩三角形之三邊。若 a', b', c' 均介於 $\min(a, b, c)$ 與 $\max(a, b, c)$ 之間。試證 a', b', c' 所圍成面積比 a, b, c 所圍成面積要大。

8. 試證在一銳角三角形 A, B, C 中

$$(i) \sqrt{2} \leq \sin \frac{A}{2} + \sin \frac{B}{2} + \sin \frac{C}{2} \leq \frac{3}{2}$$

$$(ii) 1 \leq \cos A + \cos B + \cos C \leq \frac{3}{2}$$

$$(iii) \frac{1}{2} \leq \cos \frac{A}{2} \cos \frac{B}{2} \cos \frac{C}{2} \leq \frac{3\sqrt{3}}{8}$$

$$(iv) 2 \leq \sin A + \sin B + \sin C \leq \frac{3\sqrt{3}}{2}。$$

在下面幾題(9-12)中， A, B, C 表一三角形之三角， a, b, c 表其對邊長， s 表示半周長， Δ 表面積，令 r 表內切圓半徑， r_a, r_b, r_c 表傍切圓之半徑，及 h_a, h_b, h_c 表三邊上的高。

9. 試證

$$(i) abc \leq \frac{1}{8}(a+b)(b+c)(c+a)$$

$$(ii) \frac{1}{s-a} + \frac{1}{s-b} + \frac{1}{s-c} \geq \frac{9}{s}$$

$$(iii) abc \geq 8(s-a)(s-b)(s-c)$$

10. 試證

$$h_a h_b + h_b h_c + h_c h_a \leq r_a r_b + r_b r_c + r_c r_a$$

(註:此不等式在1966年被Nasser所發現)

11. 試證 $\sqrt{s} \leq \sqrt{s-a} + \sqrt{s-b} + \sqrt{s-c} \leq \sqrt{3s}$

12. 試證

$$(i) h_a h_b h_c \geq 27r^3$$

$$(ii) h_a + h_b + h_c \leq 9r$$

$$(iii) \frac{1}{h_a-2r} + \frac{1}{h_b-2r} + \frac{1}{h_c-2r} \geq \frac{3}{r} \quad (1966 \text{ Bokov發現})$$

$$(iv) \frac{h_a+r}{h_a-r} + \frac{h_b+r}{h_b-r} + \frac{h_c+r}{h_c-r} \geq 6 \quad (1965 \text{ Cosnita 與Turtoiu發現})$$

文章出處：

楊重駿, 楊照崑(1982), 不等式(第六章)。東華書局, 台北。pp. 79-93.