

簡介蓋、舒爾凸函數與 Karamata 不等式

陳柏宇 · 張福春

摘要: 本文主要介紹 Karamata 不等式, 它與不等式理論中最多產的蓋及舒爾凸函數相關, 並考慮它在基本對稱函數、樣本變異數、熵及生日問題上的應用。

1. 前言

本文介紹的 Karamata 不等式是延森不等式 (Jensen's Inequality) 的一種推廣, 又稱做蓋不等式或 Hardy-Littlewood 不等式。延森不等式通常只能提供凸函數或凹函數的其中一個極值(極大值或極小值), 而 Karamata 不等式則能在某些情況下, 同時給出兩個極值。如同延森不等式, Karamata 不等式涵蓋的層面很廣, 尤其在不等式及極值問題的領域上更是運用廣泛的技巧。

早期, 在不同的領域裡, 蓋理論常被以不同的名稱重複的介紹, 例如, 在經濟學上稱之為羅倫斯優勢 (Lorenz dominance)。這不僅使得研究者在研究相關議題時難以確認蓋理論既存的結果, 也使得很多的研究學者並未發現蓋理論在各領域中被廣泛運用的程度。直到 Marshall and Olkin (1979) 將 Karamata 不等式做了系統化的整理, 才正式的慢慢將蓋理論推展開來。有興趣的讀者, 也可參考 Steele (2004), 書中對於 Karamata 不等式有著詳盡的介紹。另外在 Arnold (2007) 中也提到了在 Marshall and Olkin (1979) 的書出版後將近三十年間, 此書及 Karamata 不等式對學術界的影響, 除此之外也介紹了很多 Karamata 不等式在各個領域的應用, 最新版的書是 Marshall, Olkin and Arnold (2009)。

本文的安排如下: 第 2 節說明蓋、雙重隨機矩陣及凸包之間的等價關係, 第 3 節討論舒爾

美國數學會 2010 年分類索引主要 26D.

關鍵詞: 伯克霍夫定理、凸包、雙重隨機矩陣、羅倫斯曲線、蓋、舒爾凸函數、舒爾準則、Karamata 不等式、相異代表系。

凸函數及舒爾準則，並考慮在基本對稱函數、樣本變異數、熵上的應用。第 4 節介紹 Karamata 不等式，並以生日問題為例子作說明。

2. 蓋

蓋的起源來自於人們從以前就很好奇有關於兩個向量間，有沒有所謂其中某一個向量較為分散，或是較為平均的現象。如果有的話，該如何做比較呢？這樣子的議題，在各個領域中被研究著。

在二十世紀初，經濟學家開始對衡量收入或財富的不等式感興趣，爲了要去衡量這件事情，希望能夠說明在收入或財富的分佈上，何謂一個收入（財富）的分佈較另一個平均。最早針對此議題所發展出的理論爲 Lorenz（1905）所提出的羅倫斯曲線（Lorenz curve），其他還有許多不同解釋的角度，最著名的便是蓋的觀念了。

首先，定義一個新的符號：

定義 1: 給定 n 維向量 $a = (a_1, a_2, \dots, a_n)$ ，定義 $a_{[j]}$, $1 \leq j \leq n$ 爲此 n 維向量中對各分量由大至小做排序後的第 j 項，亦即 $a_{[1]} \geq a_{[2]} \geq \dots \geq a_{[n]}$ ，並定義 $a_{\downarrow} = (a_{[1]}, a_{[2]}, \dots, a_{[n]})$ 。

定義 2 (蓋 (majorization)): 給定兩組向量 $\alpha = (\alpha_1, \alpha_2, \dots, \alpha_n)$, $\beta = (\beta_1, \beta_2, \dots, \beta_n) \in \mathbb{R}^n$ ，若滿足條件

$$(i) \alpha_{[1]} + \alpha_{[2]} + \dots + \alpha_{[j]} \leq \beta_{[1]} + \beta_{[2]} + \dots + \beta_{[j]}, \quad 1 \leq j < n$$

$$(ii) \alpha_{[1]} + \alpha_{[2]} + \dots + \alpha_{[n]} = \beta_{[1]} + \beta_{[2]} + \dots + \beta_{[n]}$$

則稱 α 被 β 蓋住，記做 $\alpha \prec \beta$ 或 $\beta \succ \alpha$ 。

爲了更瞭解關於“蓋”的定義，以下舉一個簡單的例子：

$$(1, 1, 1, 1) \prec (2, 1, 1, 0) \prec (3, 1, 0, 0) \prec (4, 0, 0, 0) \quad (1)$$

式 (1) 可經由簡單的計算得到。又因爲 $\alpha \prec \beta$ 的關係只與 α 與 β 排序後的向量有關，故 (1) 也可寫成

$$(1, 1, 1, 1) \prec (0, 1, 1, 2) \prec (0, 1, 3, 0) \prec (0, 4, 0, 0)$$

定理 3(蓋的遞移性): 設 $\alpha, \beta, \gamma \in \mathbb{R}^n$ ，若有 $\alpha \prec \beta$, $\beta \prec \gamma$ ，則有 $\alpha \prec \gamma$ 。

證明: 因爲

$$(i) \alpha_{[1]} + \alpha_{[2]} + \cdots + \alpha_{[j]} \leq \beta_{[1]} + \beta_{[2]} + \cdots + \beta_{[j]} \leq \gamma_{[1]} + \gamma_{[2]} + \cdots + \gamma_{[j]}, 1 \leq j < n$$

$$(ii) \alpha_{[1]} + \alpha_{[2]} + \cdots + \alpha_{[n]} = \beta_{[1]} + \beta_{[2]} + \cdots + \beta_{[n]} = \gamma_{[1]} + \gamma_{[2]} + \cdots + \gamma_{[n]}$$

故得證。 □

了解蓋的定義後，底下介紹一個在蓋觀念中十分重要的等價關係。首先，介紹凸包的先備知識。

定義 4: 在一個實數向量空間 V 中，對於給定集合 X ，所有包含 X 的凸集合的交集

$$S := \bigcap_{X \subseteq K \subseteq V} K, \quad K \text{ 是凸集合}$$

被稱為 X 的凸包 (convex hull)。

例如在二維空間中，任意三點所形成的凸包，即為連接此三點所成的三角形，任意 n 點所形成的凸包，則為能包含此 n 點的最小 k 多邊形 ($k \leq n$)。

給定任意 n 維向量 $\alpha = (\alpha_1, \alpha_2, \dots, \alpha_n)$ 及 $\beta = (\beta_1, \beta_2, \dots, \beta_n)$ ，考慮 α 落在由 $\{(\beta_{\tau(1)}, \beta_{\tau(2)}, \dots, \beta_{\tau(n)}) \mid \tau \in \mathcal{S}_n\}$ 所組成的凸包 $H(\beta)$ 中，記作 $\alpha \in H(\beta)$ ，其中 \mathcal{S}_n 代表集合 $\{1, 2, \dots, n\}$ 中元素的 $n!$ 種排列所成的集合， $\tau(i), i = 1, 2, \dots, n$ 表示集合 \mathcal{S}_n 之中元素的第 i 個分量。例如在 $n = 3$ 的情況下，

$$\mathcal{S}_3 = \{(1, 2, 3), (1, 3, 2), (2, 1, 3), (2, 3, 1), (3, 1, 2), (3, 2, 1)\}$$

即為集合 $\{1, 2, 3\}$ 的所有可能排序，考慮其中元素 $(2, 3, 1)$ ，其 $\tau(1) = 2, \tau(2) = 3, \tau(3) = 1$ 。

接著說明對於兩向量 $\alpha, \beta \in \mathbb{R}^n$ 中， $\alpha \in H(\beta)$ 與 $\alpha \prec \beta$ 間的關係，以進一步瞭解蓋的觀念。

定理 5: 設向量 $\alpha, \beta \in \mathbb{R}^n$ ，且 $\alpha \in H(\beta)$ ，則 $\alpha \prec \beta$ 。

證明: 對於 $\alpha \in H(\beta)$ ， α 可表達為

$$(\alpha_1, \alpha_2, \dots, \alpha_n) = \sum_{\tau \in \mathcal{S}_n} p_\tau (\beta_{\tau(1)}, \beta_{\tau(2)}, \dots, \beta_{\tau(n)})$$

其中 $\sum_{\tau \in \mathcal{S}_n} p_\tau = 1, p_\tau \geq 0$ 。若只考慮第 j 個分量，則有以下等式

$$\alpha_j = \sum_{\tau \in \mathcal{S}_n} p_\tau \beta_{\tau(j)} = \sum_{k=1}^n \left\{ \sum_{\tau: \tau(j)=k} p_\tau \right\} \beta_k = \sum_{k=1}^n d_{jk} \beta_k \tag{2}$$

式 (2) 中, 爲了簡化式子, 令

$$d_{jk} = \sum_{\tau: \tau(j)=k} p_{\tau}$$

此處, 可發現 $d_{jk} \geq 0$, 且因爲對 d_{jk} 的任意下標求和皆相當於 \mathcal{S}_n 中所有 p_{τ} 的總和, 故有

$$\sum_{j=1}^n d_{jk} = 1, \quad \sum_{k=1}^n d_{jk} = 1 \quad (3)$$

對一個非負實係數的矩陣 $D = \{d_{jk}\}$ 若能滿足 (3), 則稱此矩陣爲雙重隨機矩陣 (doubly stochastic matrix)。因此, 若將 α, β 視爲行向量, 則由 (2) 有

$$\alpha \in H(\beta) \Rightarrow \alpha = D\beta$$

因此, 接下來只需證明 $\alpha = D\beta \Rightarrow \alpha \prec \beta$ 。

因爲 $\alpha \in H(\beta)$ 及 $\alpha \prec \beta$ 皆不受 α, β 向量內排序的影響, 因此不失一般性可假設 $\alpha_1 \geq \alpha_2 \geq \cdots \geq \alpha_n$ 且 $\beta_1 \geq \beta_2 \geq \cdots \geq \beta_n$ 。接著, 將 (2) 對 j 做加總, 可得

$$\sum_{j=1}^k \alpha_j = \sum_{j=1}^k \sum_{t=1}^n d_{jt} \beta_t = \sum_{t=1}^n c_t \beta_t \quad \text{其中 } c_t \stackrel{\text{定義}}{=} \sum_{j=1}^k d_{jt} \quad (4)$$

因爲 c_t 爲 D 中第 t 行的前 k 個元素的加總, 且 D 爲雙重隨機矩陣, 所以

$$0 \leq c_t \leq 1, \quad 1 \leq t \leq n \quad \text{且} \quad c_1 + c_2 + \cdots + c_n = k$$

最後, 根據蓋的定義, 觀察以下等式

$$\begin{aligned} \Delta_k &= \sum_{j=1}^k \alpha_j - \sum_{j=1}^k \beta_j = \sum_{j=1}^n c_j \beta_j - \sum_{j=1}^k \beta_j + \beta_k \left(k - \sum_{j=1}^n c_j \right) \\ &= \sum_{j=1}^k (\beta_k - \beta_j) (1 - c_j) + \sum_{j=k+1}^n c_j (\beta_j - \beta_k) \end{aligned}$$

因爲當 $1 \leq j \leq k$ 時, $\beta_j \geq \beta_k$, 而當 $k < j \leq n$ 時, $\beta_j \leq \beta_k$, 所以有 $\Delta_k \leq 0$, $1 \leq k < n$, 且由 (4) 可知 $\Delta_n = 0$, 滿足蓋的定義, 故 $\alpha \prec \beta$ 得證。 \square

如同之前提到蓋的起源與經濟學上的關係, 一個有趣的例子是假設觀察甲與乙兩個國家的人民收入狀況, 設定 α_1 爲甲國前 10% 收入者的收入占甲國全國總收入的比例, α_2 則爲接下來 10% 收入者的收入占甲國全國總收入的比例, 依此類推, α_{10} 爲甲國最後 10% 收入者的收入占甲國全國總收入的比例。接著以 $\beta_1, \dots, \beta_{10}$ 對乙國做相同的定義。則 $\alpha \prec \beta$ 在經濟學上可被視爲乙國的人民收入分佈比甲國來得不平均。

很直覺的，可聯想到是否 $\alpha \prec \beta$ 也可逆推回 $\alpha = D\beta$ 使得乙國人民的收入分佈能藉由雙重隨機矩陣的轉換慢慢的與甲國類似。也有人將這樣的觀念聯想到羅賓漢-劫富濟貧的故事。

定理 6: 設向量 $\alpha, \beta \in \mathbb{R}^n$ ，且 $\alpha \prec \beta$ 則存在一雙重隨機矩陣 D 使得 $\alpha = D\beta$ 。

證明: 首先，考慮最簡單的情況，當 $n = 2$ 時，令 $\alpha = (\alpha_1, \alpha_2) = (\rho + \sigma, \rho - \sigma)$ 且 $\beta = (\beta_1, \beta_2) = (\rho + \tau, \rho - \tau)$ 。不失一般性，可假設 $\alpha_1 \geq \alpha_2, \beta_1 \geq \beta_2$ ，且 $\alpha_1 + \alpha_2 = \beta_1 + \beta_2$ 。因此，若 $\sigma \leq \tau$ ，則相當於說明 $\alpha \prec \beta$ ，因此，剩下的工作只需找出一雙重隨機矩陣 D 使得 $\alpha = D\beta$ ，利用分點公式求 D 可得

$$D\beta = \begin{bmatrix} \frac{\tau+\sigma}{2\tau} & \frac{\tau-\sigma}{2\tau} \\ \frac{\tau-\sigma}{2\tau} & \frac{\tau+\sigma}{2\tau} \end{bmatrix} \begin{bmatrix} \rho + \tau \\ \rho - \tau \end{bmatrix} = \begin{bmatrix} \rho + \sigma \\ \rho - \sigma \end{bmatrix} = \alpha \quad (5)$$

故此定理在 $n = 2$ 的情況成立。

雖然上述的情況很簡單，但可以利用上述的結果，推廣到 n 維的情況，並且說明 $n \times n$ 的雙重隨機矩陣 D 為由有限個每次改變兩個分量的轉換累積相乘而得。

考慮向量 α, β 有 $\alpha \prec \beta$ ，且 $\alpha_1 \geq \alpha_2 \geq \dots \geq \alpha_n, \beta_1 \geq \beta_2 \geq \dots \geq \beta_n$ ，已知其中有 N 個分量不相等。考慮 $N \geq 2$ ，則根據蓋的定義可知，必定存在整數 $1 \leq j < k \leq n$ ，使得

$$\beta_j > \alpha_j, \quad \beta_k < \alpha_k, \quad \text{且} \quad \beta_s = \alpha_s, \quad \forall j < s < k \quad (6)$$

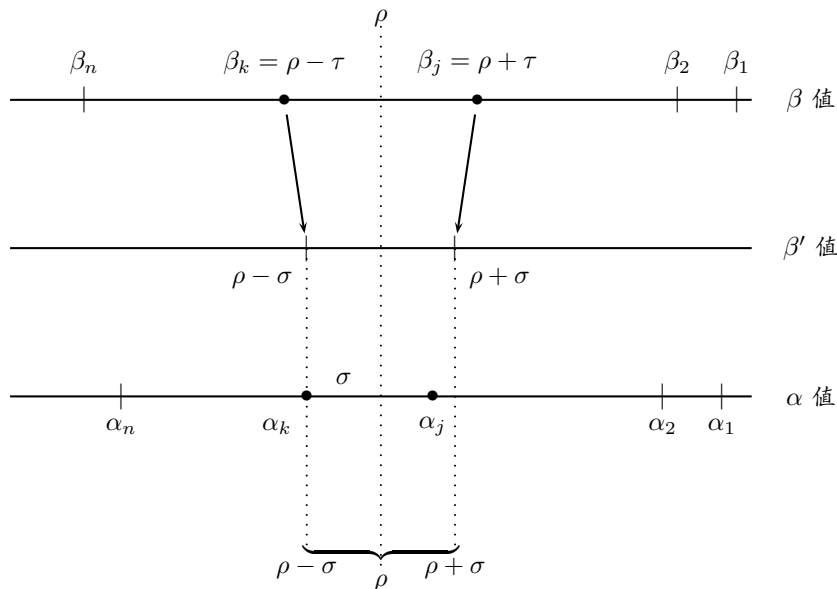


圖1: 做一次轉換為 β'

考慮類似 $n = 2$ 的情況，取 $\rho = (\beta_j + \beta_k) / 2, \tau = \beta_j - \rho$ ，故 $\beta_j = \rho + \tau, \beta_k = \rho - \tau$ ，

接著取 $\sigma = \max\{|\alpha_k - \rho|, |\alpha_j - \rho|\}$, 如圖 1。然後令 T 為能將 $\beta = (\beta_1, \beta_2, \dots, \beta_n)$ 送到 $\beta' = (\beta'_1, \beta'_2, \dots, \beta'_n)$ 的 $n \times n$ 的雙重隨機矩陣, 其中 β' 滿足

$$\beta'_k = \beta_k + \tau - \sigma, \quad \beta'_j = \beta_j - \tau + \sigma, \quad \beta'_t = \beta_t, \quad \forall t \neq j, t \neq k$$

T 的表示式可將 $n = 2$ 時所推導出的雙重隨機矩陣的係數分別對應放置於矩陣 T 的 (j, j) , (j, k) , (k, j) , (k, k) 四個點上, 接著將對角線上的空位補滿 1, 其他空位則補滿 0, 如 (7)。值得注意的是, 轉換後的 β' 滿足 $\beta' \prec \beta$ 。

$$\begin{bmatrix} 1 & & & & & & \\ & \ddots & & & & & \\ & & 1 & & & & \\ & & & \frac{\tau+\sigma}{2\tau} & \dots & \frac{\tau-\sigma}{2\tau} & \\ & & & \vdots & & \vdots & \\ & & & \frac{\tau-\sigma}{2\tau} & \dots & \frac{\tau+\sigma}{2\tau} & \\ & & & & & & 1 & & & \\ & & & & & & & \ddots & & \\ & & & & & & & & & 1 \end{bmatrix} \quad (7)$$

利用簡單的觀察及計算可以檢查 $\alpha \prec \beta'$, 因此, 藉由矩陣 T 將 β 轉換成 β' , 可將 α 與 β' 間不相等的分量縮小為至多 $N - 1$ 個。最後, 利用歸納法, 假設存在雙重隨機矩陣 D' 使得 $\alpha = D'\beta'$, 因為 $\beta' = T\beta$, 故有 $\alpha = D'(T\beta) = (D'T)\beta$, 且知兩個雙重隨機矩陣的乘積依然為雙重隨機矩陣, 故得證。□

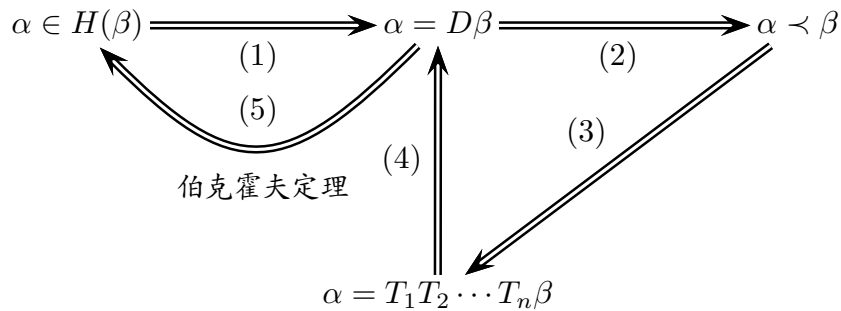


圖 2: 蓋與凸包的關係

觀察圖 2, 其中 (1) 與 (2) 已經在定理 5 說明, 而 (3) 跟 (4) 的證明則可由定理 6 得到, 若能證明最後一事情, 即 $\alpha = D\beta \Rightarrow \alpha \in H(\beta)$, 則可瞭解 $\alpha \prec \beta$ 與 $\alpha \in H(\beta)$ 兩件事為等價的。要證明 (5), 需要證明伯克霍夫定理, 這個定理又被稱為是雙重隨機矩陣的基本定理。此處採用一個較有趣的證明方式, 更詳細的介紹可參考 Steele (2004)。

定理 7 (結婚問題): 設 $S_1, S_2, \dots, S_n \subset S$, 則對一集合 $R = \{x_1, x_2, \dots, x_n\} \subset S$, 若其中元素皆不相同且 $x_k \in S_k, k = 1, 2, \dots, n$, 則稱 R 為一相異代表系 (system of distinct representatives) (又稱 SDR)。證明 SDR 存在若且唯若滿足以下條件

$$|A| \leq \left| \bigcup_{j \in A} S_j \right| \quad \text{其中 } A \subset \{1, 2, \dots, n\} \quad (8)$$

此處符號 $|C|$ 代表集合 C 中的元素個數。

這個問題來自 Weyl (1949) 的文章。原本的題目為考慮一群男孩及一群女孩, 若每個女孩只能嫁給自己認識的男孩, 則順利將所有女孩嫁出若且唯若可得任意 k 個女孩至少認識 k 個男孩。

證明: 明顯可知, 若 SDR 存在, 則條件必定滿足。考慮另一個方向的證明, 利用 Weyl 的題意, 即第 j 個女孩認識的男孩為集合 S_j , 所以給定女孩的集合 A , 則集合 $\bigcup_{j \in A} S_j$ 中的男孩必定有某個 A 中的女孩認識。考慮下列兩種狀況:

情形一: 假設 (8) 中的不等號為嚴格小於 (即無等號成立的情況發生), 且已知 $|A| < n$ 。將第 n 個女孩配對給她所認識的任一個男孩 b 。因為條件 (8) 依然在集合 $A \subset \{1, 2, \dots, n-1\}$ 及每個 $S_j, 1 \leq j \leq n-1$ 皆被替換成 $S_j \setminus \{b\}$ 的情況下成立 (須再重新判斷其為情況一或情況二), 因此其餘的女孩可依相同的方法配對給其餘的男孩們。

情形二: 假設對某個集合 A_0 考慮條件 (8), 發現等號成立, 且已知 $|A_0| < n$ 。令

$$B = \bigcup_{j \in A_0} S_j \quad \text{且} \quad S'_j = S_j \setminus B \quad \text{對所有 } j \in A_0^c$$

則根據歸納法, A_0 中的女孩必可配對給 B 中的男孩, 因此只需證明在 A_0^c 中的女孩也可適當的配對給 B^c 中的男孩。取任意的 $A \subset A_0^c$ 可發現

$$\left| \bigcup_{j \in A_0 \cup A} S_j \right| \geq |A_0 \cup A| = |A| + |A_0|$$

且下列等式恆成立

$$\left| \bigcup_{j \in A_0 \cup A} S_j \right| = \left| \left\{ \bigcup_{j \in A_0} S_j \right\} \cup \left\{ \bigcup_{j \in A} S'_j \right\} \right| = |A_0| + \left| \bigcup_{j \in A} S'_j \right|$$

因此, 可發現對所有 $A \subset A_0^c$ 皆可導出

$$\left| \bigcup_{j \in A} S'_j \right| \geq |A|$$

即對任意 A_0^c 中取 k 個女孩的集合皆至少認識 k 個 B^c 中的男孩。根據數學歸納法, 可知 A_0^c 中的女孩必可適當的配對給 B^c 中的男孩。□

定理 8 (伯克霍夫定理 (Birkhoff Theorem)): 給定排列 $\sigma \in \mathcal{S}_n$, 對應 σ 的置換矩陣為一 $n \times n$ 的矩陣 $P_\sigma = (P_\sigma(j, k))$, $1 \leq j, k \leq n$ 其中元素為

$$P_\sigma(j, k) = \begin{cases} 1 & \text{若 } \sigma(j) = k \\ 0 & \text{其他} \end{cases}$$

證明若 D 為一 $n \times n$ 的雙重隨機矩陣, 則存在非負權重 $\{w_\sigma : \sigma \in \mathcal{S}_n\}$ 使得

$$\sum_{\sigma \in \mathcal{S}_n} w_\sigma = 1 \quad \text{且} \quad \sum_{\sigma \in \mathcal{S}_n} w_\sigma P_\sigma = D$$

即所有的雙重隨機矩陣皆可表示為置換矩陣的加權平均。

證明: 此處, 利用定理 7 結婚問題的結果推導伯克霍夫定理。給定雙重隨機矩陣 D , 考慮 $1 \leq j \leq n$, 令 S_j 為能使得 $d_{jk} > 0$ 的所有 k 所成的集合, 則對任意集合 $A \subset \{1, 2, \dots, n\}$ 必有

$$|A| = \sum_{j \in A} \sum_{k \in S_j} d_{jk} \leq \sum_{k \in \bigcup_{j \in A} S_j} \sum_{1 \leq j \leq n} d_{jk} = \left| \bigcup_{j \in A} S_j \right|$$

根據定理 7, 必存在 $\{S_1, S_2, \dots, S_n\}$ 的 SDR, 所以我們可以定義排列 σ 藉由設定 $\sigma(j)$ 為 S_j 中的代表值 (representative), 以 Weyl 的例子說明則相當於將向量中的第 j 個分量取為第 j 個女生所配對到的男生。接下來, 令 P_σ 為 σ 的置換矩陣並令 $\alpha = \min d_{j\sigma(j)} > 0$ 。可知, 若 $\alpha = 1$ 則 D 為一置換矩陣, 故得證。但若 $\alpha < 1$, 則考慮定義一個新的矩陣 $D' = (1 - \alpha)^{-1}(D - \alpha P_\sigma)$, 可改寫為

$$D = \alpha P_\sigma + (1 - \alpha) D'$$

觀察可知 D' 依然為一雙重隨機矩陣, 且矩陣中含有比 D 中更多的 0。最後, 可利用數學歸納法完成此定理的證明。 \square

定理 9 (蓋的稠密性): 設向量 $\alpha, \beta \in \mathbb{R}^n$, $\alpha \prec \beta$, 且 $\alpha_\downarrow \neq \beta_\downarrow$, 則必定存在 $\gamma \in \mathbb{R}^n$, 使得 $\alpha \prec \gamma \prec \beta$ 。

證明: 不失一般性考慮向量 α, β 有 $\alpha \prec \beta$, 且 $\alpha_1 \geq \alpha_2 \geq \dots \geq \alpha_n$, $\beta_1 \geq \beta_2 \geq \dots \geq \beta_n$, 假設其中有 N 個分量不相等。考慮 $N \geq 2$, 則根據蓋的定義可知, 必定存在整數 $1 \leq j < k \leq n$, 使得

$$\beta_j > \alpha_j, \quad \beta_k < \alpha_k, \quad \text{且} \quad \beta_s = \alpha_s, \quad \forall j < s < k \quad (9)$$

考慮類似定理 6 的證明方式, 取 $\rho = (\beta_j + \beta_k)/2$, $\tau = \beta_j - \rho$, 故 $\beta_j = \rho + \tau$,

$\beta_k = \rho - \tau$, 接著取 $\sigma = \max\{|\alpha_k - \rho|, |\alpha_j - \rho|\}$, 如圖 1。接著取 γ 滿足

$$\gamma_k = \beta_k + c(\tau - \sigma), \quad \gamma_j = \beta_j - c(\tau - \sigma), \quad \gamma_t = \beta_t, \quad t \neq j, \quad t \neq k$$

其中 $c \in (0, 1)$, 故 $\gamma \neq \alpha, \gamma \neq \beta$ 。接著驗證此 γ 是否符合我們的需求

$$\begin{aligned} \sum_{i=1}^t \alpha_i &\leq \sum_{i=1}^t \gamma_i = \sum_{i=1}^t \beta_i, & 1 \leq t < j \\ \sum_{i=1}^t \alpha_i &\leq \sum_{i=1}^t \gamma_i \leq \sum_{i=1}^t \beta_i, & j \leq t < k \\ \sum_{i=1}^t \alpha_i &\leq \sum_{i=1}^t \gamma_i = \sum_{i=1}^t \beta_i, & k \leq t < n \\ \sum_{i=1}^n \alpha_i &= \sum_{i=1}^n \gamma_i = \sum_{i=1}^n \beta_i \end{aligned}$$

故有 $\alpha \prec \gamma \prec \beta$, 得證。 □

3. 舒爾凸函數

要清楚的了解 Karamata 不等式, 除了熟悉蓋的觀念外, 由於 Karamata 不等式為舒爾凸函數的一個重要特例, 因此, 底下從較為廣義的舒爾凸函數談起。

定義 10: 設 $f : \mathcal{D} \rightarrow \mathbb{R}, \mathcal{D} \subseteq \mathbb{R}^n, \alpha, \beta \in \mathcal{D}$, 且 $\alpha \prec \beta$, 若函數 f 滿足

- (i) $f(\alpha) \leq f(\beta)$, 稱 f 為在 \mathcal{D} 上的舒爾凸函數 Schur convex function。
- (ii) $f(\alpha) \geq f(\beta)$, 稱 f 為在 \mathcal{D} 上的舒爾凹函數 Schur concave function。

首先介紹一個可用舒爾凸函數得到的不等式。

例 1 (1971 American Mathematical Monthly problem E2284): 設 a, b, c 為正實數, $x = b + c - a, y = a + c - b, z = a + b - c$, 且 $x, y, z \geq 0$, 試證

$$\frac{1}{a} + \frac{1}{b} + \frac{1}{c} \leq \frac{1}{x} + \frac{1}{y} + \frac{1}{z}$$

證明: 因為對稱性, 故不失一般性, 假設 $a \leq b \leq c$, 則可觀察到有 $a \leq b + c - a = x, a + b \leq 2c = x + y$, 且 $a + b + c = x + y + z$, 所以可以得到 $(a, b, c) \prec (x, y, z)$ 。接著, 若可以得到 $f(t_1, t_2, t_3) = \frac{1}{t_1} + \frac{1}{t_2} + \frac{1}{t_3}$ 為一舒爾凸函數, 則欲求的不等式便可利用定義 10 輕鬆得到。但問題是, 要如何說明一給定函數為舒爾凸函數呢? □

以下介紹判斷一函數是否為舒爾凸函數最常用的一個方法—舒爾準則。這個方法是由 Schur (1923) 提出, 但直至今日, 它依舊是判斷舒爾凸函數最常用的工具。

定理 11 (舒爾準則 (Schur Criterion)): 給定一對稱且連續可微的函數 $f : (a, b)^n \rightarrow \mathbb{R}$, 此函數 f 為舒爾凸函數若且唯若對於 $1 \leq j < k \leq n$ 及 $\mathbf{x} \in (a, b)^n$ 滿足

$$0 \leq (x_j - x_k) \left(\frac{\partial f(\mathbf{x})}{\partial x_j} - \frac{\partial f(\mathbf{x})}{\partial x_k} \right)$$

若滿足的不等式為此不等式的變號, 則稱函數 f 為舒爾凹函數。

在證明舒爾準則之前, 先介紹一個符號, 以簡化證明的過程。

定義 12: 設 $\mathbf{w} = (w_1, w_2, \dots, w_n)$ 為 n 維向量, 且 $w_1 \geq w_2 \geq \dots \geq w_n$, 定義

$$\tilde{w}_j = w_1 + w_2 + \dots + w_j, \quad j = 1, 2, \dots, n$$

並定義

$$\tilde{\mathbf{w}} = (\tilde{w}_1, \tilde{w}_2, \dots, \tilde{w}_n)$$

從上述的定義可觀察到, 對任意兩組 n 維向量 $\mathbf{x} \prec \mathbf{y}$ 成立, 若且唯若 $\tilde{x}_j \leq \tilde{y}_j$, $1 \leq j < n$, 且 $\tilde{x}_n = \tilde{y}_n$ 。

證明: 根據假設 f 為對稱函數, 可知 f 在 $(a, b)^n$ 上為舒爾凸函數若且唯若其在 $\mathcal{B} = (a, b)^n \cap \mathcal{D}$ 也為舒爾凸函數, 其中 $\mathcal{D} = \{(x_1, x_2, \dots, x_n) \mid x_1 \geq x_2 \geq \dots \geq x_n\}$ 。設集合 $\tilde{\mathcal{B}} = \{\tilde{\mathbf{x}} : \mathbf{x} \in \mathcal{B}\}$, 則可定義新的函數 $\tilde{f} : \tilde{\mathcal{B}} \rightarrow \mathbb{R}$, 其中對於所有的 $\tilde{\mathbf{x}} \in \tilde{\mathcal{B}}$ 有 $\tilde{f}(\tilde{\mathbf{x}}) = f(\mathbf{x})$ 。

注意對 $\mathbf{x}, \mathbf{y} \in \mathcal{B}$, 當 $\mathbf{x} \prec \mathbf{y}$ 時, $f(\mathbf{x}) \leq f(\mathbf{y})$ 成立若且唯若對 $\tilde{\mathbf{x}}, \tilde{\mathbf{y}} \in \tilde{\mathcal{B}}$ 有 $\tilde{f}(\tilde{\mathbf{x}}) \leq \tilde{f}(\tilde{\mathbf{y}})$ 。換句話說, f 在 \mathcal{B} 上為舒爾凸函數若且唯若 \tilde{f} 的前 $n-1$ 個分量在 $\tilde{\mathcal{B}}$ 上為一非遞減函數。又因為 f 為連續可微函數, 所以 f 為舒爾凸函數若且唯若在 $\tilde{\mathcal{B}}$ 內的每一個 $\tilde{\mathbf{x}}$ 皆有

$$0 \leq \frac{\partial \tilde{f}(\tilde{\mathbf{x}})}{\partial \tilde{x}_j}, \quad 1 \leq j < n \quad (10)$$

更明確的說, 因為 $\tilde{f}(\tilde{\mathbf{x}}) = f(\tilde{x}_1, \tilde{x}_2 - \tilde{x}_1, \dots, \tilde{x}_n - \tilde{x}_{n-1})$, 可根據鏈鎖律 (chain rule) 將 (10) 改寫為

$$0 \leq \frac{\partial \tilde{f}(\tilde{\mathbf{x}})}{\partial \tilde{x}_j} = \frac{\partial f(\mathbf{x})}{\partial x_j} - \frac{\partial f(\mathbf{x})}{\partial x_{j+1}}, \quad 1 \leq j < n \quad (11)$$

所以, 若將 (11) 中的下標 j 由 $1 \leq j < k \leq n$ 做加總, 則可得到

$$0 \leq \frac{\partial f(\mathbf{x})}{\partial x_j} - \frac{\partial f(\mathbf{x})}{\partial x_k}, \quad \mathbf{x} \in \mathcal{B} \quad (12)$$

最後，根據 f 在 $(a, b)^n$ 上的對稱性，(12) 的條件等價於

$$0 \leq (x_j - x_k) \left(\frac{\partial f(\mathbf{x})}{\partial x_j} - \frac{\partial f(\mathbf{x})}{\partial x_k} \right), \quad \mathbf{x} \in (a, b)^n \quad (13)$$

故得證。 \square

有了舒爾準則，再回頭討論例 1，取函數 $f(t_1, t_2, t_3) = \frac{1}{t_1} + \frac{1}{t_2} + \frac{1}{t_3}$ 。對此函數，考慮舒爾準則，可得對 $1 \leq j < k \leq n$

$$(t_j - t_k) \left(\frac{\partial f(\mathbf{t})}{\partial t_j} - \frac{\partial f(\mathbf{t})}{\partial t_k} \right) = (t_j - t_k) (-t_j^{-2} + t_k^{-2}) \geq 0$$

所以 $f(t_1, t_2, t_3) = \frac{1}{t_1} + \frac{1}{t_2} + \frac{1}{t_3}$ 為一舒爾凸函數，故例 1 得證。 \square

例 2: 設 x, y, z 為正實數，試證

$$xyz \leq \left(\frac{x}{2} + \frac{y}{3} + \frac{z}{6} \right) \left(\frac{x}{3} + \frac{2y}{3} \right) \left(\frac{x}{6} + \frac{5z}{6} \right)$$

證明: 觀察欲證之不等式，首先令 (a, b, c) 滿足下述等式

$$\begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} \frac{1}{2} & \frac{1}{3} & \frac{1}{6} \\ \frac{1}{3} & \frac{2}{3} & 0 \\ \frac{1}{6} & 0 & \frac{5}{6} \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix}$$

根據定理 5 可知 $(a, b, c) \prec (x, y, z)$ 。取函數 $f(t_1, t_2, t_3) = t_1 t_2 t_3$, $t_1, t_2, t_3 > 0$ ，考慮舒爾準則，可得對 $1 \leq j < k \leq 3$

$$\begin{aligned} (t_j - t_k) \left(\frac{\partial f(\mathbf{t})}{\partial t_j} - \frac{\partial f(\mathbf{t})}{\partial t_k} \right) &= (t_j - t_k) \left(\prod_{i=1,2,3;i \neq j} t_i - \prod_{i=1,2,3;i \neq k} t_i \right) \\ &= (t_j - t_k) \left(\frac{t_1 t_2 t_3}{t_j} - \frac{t_1 t_2 t_3}{t_k} \right) \leq 0 \end{aligned}$$

所以當 $t_1, t_2, t_3 > 0$ 時， $f(t_1, t_2, t_3) = t_1 t_2 t_3$ 為一舒爾凹函數，最後將 (a, b, c) , (x, y, z) 代入函數 f

$$\begin{aligned} f(x, y, z) &\leq f(a, b, c) \\ \Rightarrow xyz &\leq \left(\frac{x}{2} + \frac{y}{3} + \frac{z}{6} \right) \left(\frac{x}{3} + \frac{2y}{3} \right) \left(\frac{x}{6} + \frac{5z}{6} \right) \end{aligned}$$

故得證。 \square

例 3: 設 x, y, z 為正實數, 試證

$$\left(\frac{2}{y+z}\right)^5 + \left(\frac{6}{3x+y+2z}\right)^5 + \left(\frac{6}{3x+2y+z}\right)^5 \leq \frac{1}{x^5} + \frac{1}{y^5} + \frac{1}{z^5}$$

證明: 觀察欲證之不等式, 首先令 (a, b, c) 滿足下述等式

$$\begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} 0 & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{6} & \frac{1}{3} \\ \frac{1}{2} & \frac{1}{3} & \frac{1}{6} \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix}$$

根據定理 5 可知 $(a, b, c) \prec (x, y, z)$ 。取函數 $f(t_1, t_2, t_3) = \frac{1}{t_1^5} + \frac{1}{t_2^5} + \frac{1}{t_3^5}$, $t_1, t_2, t_3 > 0$, 考慮舒爾準則, 可得對 $1 \leq j < k \leq 3$

$$(t_j - t_k) \left(\frac{\partial f(\mathbf{t})}{\partial t_j} - \frac{\partial f(\mathbf{t})}{\partial t_k} \right) = (t_j - t_k) (5t_k^{-6} - 5t_j^{-6}) \geq 0$$

所以當 $t_1, t_2, t_3 > 0$ 時, $f(t_1, t_2, t_3) = \frac{1}{t_1^5} + \frac{1}{t_2^5} + \frac{1}{t_3^5}$ 為一舒爾凸函數, 最後將 (a, b, c) , (x, y, z) 代入函數 f 即可得證。□

在介紹下一個例子前, 需要先定義一個新的函數。

定義 13: 對任意實數 x_1, x_2, \dots, x_n , 考慮多項式

$$P(t) = (t-x_1)(t-x_2)\cdots(t-x_n) = t^n - e_1(\mathbf{x})t^{n-1} + \cdots + (-1)^k e_k(\mathbf{x})t^{n-k} + \cdots + (-1)^n e_n(\mathbf{x})$$

稱 e_i , $i = 1, 2, \dots, n$ 為 n 變數 x_1, x_2, \dots, x_n 的 i 階基本對稱函數 (elementary symmetric functions), 其中

$$e_k(\mathbf{x}) = e_k(x_1, x_2, \dots, x_n) = \sum_{1 \leq i_1 < i_2 < \cdots < i_k \leq n} x_{i_1} x_{i_2} \cdots x_{i_k}$$

例如:

$$e_1(x_1, x_2, x_3, x_4) = x_1 + x_2 + x_3 + x_4$$

$$e_2(x_1, x_2, x_3, x_4) = x_1 x_2 + x_1 x_3 + x_1 x_4 + x_2 x_3 + x_2 x_4 + x_3 x_4$$

$$e_3(x_1, x_2, x_3, x_4) = x_1 x_2 x_3 + x_1 x_2 x_4 + x_1 x_3 x_4 + x_2 x_3 x_4$$

$$e_4(x_1, x_2, x_3, x_4) = x_1 x_2 x_3 x_4$$

例 4: 證明 $e_k(\mathbf{x})$ 對 $\mathbf{x} \in [0, \infty)^n$ 為一舒爾凹函數。

證明: 觀察基本對稱函數, 已知其微分可寫成

$$\frac{\partial e_k(\mathbf{x})}{\partial x_s} = e_{k-1}(x_1, x_2, \dots, x_{s-1}, x_{s+1}, \dots, x_n) \quad (14)$$

接著觀察舒爾準則

$$\begin{aligned} & (x_s - x_t) \left(\frac{\partial e_k(\mathbf{x})}{\partial x_s} - \frac{\partial e_k(\mathbf{x})}{\partial x_t} \right) \\ &= (x_s - x_t) (e_{k-1}(x_1, \dots, x_{s-1}, x_{s+1}, \dots, x_n) - e_{k-1}(x_1, \dots, x_{t-1}, x_{t+1}, \dots, x_n)) \\ &= -(x_s - x_t)^2 e_{k-2}(x_1, \dots, x_{s-1}, x_{s+1}, \dots, x_{t-1}, x_{t+1}, \dots, x_n) \leq 0 \end{aligned}$$

可發現 $e_k(\mathbf{x})$ 在 $\mathbf{x} \in [0, \infty)^n$ 為舒爾凹函數。 □

例 5: 用來測量資料離散程度的方式有很多, 舉例來說, 在統計上通常使用樣本變異數

$$s(\mathbf{x}) = \frac{1}{n-1} \sum_{j=1}^n (x_j - \bar{x})^2$$

此處 $\bar{x} = (x_1 + x_2 + \dots + x_n)/n$, $\mathbf{x} \in \mathbb{R}^n$, $n \geq 2$ 。而在資訊理論中, 使用熵 (entropy)

$$h(\mathbf{p}) = - \sum_{k=1}^n p_k \log p_k$$

來測量機率分佈 (p_1, p_2, \dots, p_n) 的離散程度, 此處 $p_k \geq 0$ 且 $p_1 + p_2 + \dots + p_n = 1$ 。試證樣本變異數 $s(\mathbf{x})$ 為舒爾凸函數而熵 $h(\mathbf{p})$ 為舒爾凹函數。

證明: 對樣本變異數 $s(\mathbf{x})$ 觀察舒爾準則

$$(x_j - x_k) \left(\frac{\partial s(\mathbf{x})}{\partial x_j} - \frac{\partial s(\mathbf{x})}{\partial x_k} \right) = 2(x_j - x_k)^2 / (n-1) \geq 0$$

可知樣本變異數 $s(\mathbf{x})$ 為舒爾凸函數。接著對函數熵 $h(\mathbf{p})$ 觀察舒爾準則, 有

$$(p_j - p_k) (h_{p_j}(\mathbf{p}) - h_{p_k}(\mathbf{p})) = (p_j - p_k) (\log p_k - \log p_j) \leq 0$$

故知熵 $h(\mathbf{p})$ 為舒爾凹函數, 故得證。 □

4. Karamata 不等式

舒爾凸函數在不等式的題目裡, 有相當廣泛的應用, 但其中最著名的一個特例, 即為 Karamata 不等式。

定理 14: (Karamata 不等式 (Karamata Inequality)) 對函數

$\phi : (a, b) \rightarrow \mathbb{R}$, 定義函數 $f : (a, b)^n \rightarrow \mathbb{R}$ 為

$$f(x_1, x_2, \dots, x_n) = \sum_{k=1}^n \phi(x_k)$$

則對任意 $\alpha, \beta \in (a, b)^n$ 且 $\alpha \prec \beta$ 有

$$\sum_{k=1}^n \phi(\alpha_k) \leq \sum_{k=1}^n \phi(\beta_k), \quad \text{當 } \phi : (a, b) \rightarrow \mathbb{R} \text{ 為凸函數}$$

$$\sum_{k=1}^n \phi(\alpha_k) \geq \sum_{k=1}^n \phi(\beta_k), \quad \text{當 } \phi : (a, b) \rightarrow \mathbb{R} \text{ 為凹函數}$$

若 ϕ 為嚴格凸(凹)函數, 則等號成立若且唯若 $\alpha_{\downarrow} = \beta_{\downarrow}$ 。

證明: 此處證明當 ϕ 為凸函數的情況, 凹函數同理可證。首先考慮函數 ϕ 的可微性。

情形一: ϕ 為可微函數

觀察函數 $f(\mathbf{x})$, 利用舒爾準則可得到

$$(x_j - x_k) \left(\frac{\partial f(\mathbf{x})}{\partial x_j} - \frac{\partial f(\mathbf{x})}{\partial x_k} \right) = (x_j - x_k) (\phi'(x_j) - \phi'(x_k)) \geq 0 \quad (\text{因為 } \phi \text{ 為凸函數})$$

故 $f(\mathbf{x})$ 為一舒爾凸函數, 又 $\alpha \prec \beta$, 所以有

$$\sum_{k=1}^n \phi(\alpha_k) = f(\alpha) \leq f(\beta) = \sum_{k=1}^n \phi(\beta_k)$$

若 ϕ 為嚴格凸函數, 則等號成立若且唯若 $\alpha_{\downarrow} = \beta_{\downarrow}$ 。 □

情形二: ϕ 不為可微函數

由於 $\alpha \prec \beta$, 故根據定理 6 可知存在一雙重隨機矩陣 $D = \{d_{jk}\}$, 使得 $\alpha = D\beta$, 或寫成

$$\alpha_j = \sum_{k=1}^n d_{jk} \beta_k$$

接著應用延森不等式在凸函數 ϕ 上, 得

$$\phi(\alpha_j) \leq \sum_{k=1}^n d_{jk} \phi(\beta_k) \tag{15}$$

最後, 對 (15) 中的下標 j 做加總, 發現

$$\sum_{j=1}^n \phi(\alpha_j) \leq \sum_{j=1}^n \sum_{k=1}^n d_{jk} \phi(\beta_k) = \sum_{k=1}^n \left\{ \phi(\beta_k) \sum_{j=1}^n d_{jk} \right\} = \sum_{k=1}^n \phi(\beta_k)$$

即為 Karamata 不等式的結果。若 ϕ 為嚴格凸函數，則等號成立若且唯若 $\alpha_{\downarrow} = \beta_{\downarrow}$ 。 \square

之前提到過，延森不等式通常只能提供關於凸函數或凹函數的其中一個極值(極大值或極小值)，而 Karamata 不等式則能在某些情況下，同時給出兩個極值。觀察 Karamata 不等式的定理，可發現 Karamata 不等式的大小關係與給定的兩向量 α, β 彼此之間的蓋 " \prec " 有十分密切的關係。

觀察下面的式子，雖然利用蓋的定義可以輕易得到，但卻是在 Karamata 不等式中十分重要的結果。

$$\begin{aligned} \left(\frac{1}{n}, \frac{1}{n}, \dots, \frac{1}{n}\right) &\prec \left(\frac{1}{n-1}, \dots, \frac{1}{n-1}, 0\right) \prec \dots \\ &\prec \left(\frac{1}{2}, \frac{1}{2}, 0, \dots, 0\right) \prec (1, 0, \dots, 0) \end{aligned}$$

或者，更廣義的來說，給定一組 n 維向量 $\alpha = (\alpha_1, \alpha_2, \dots, \alpha_n)$ ，若 $\alpha_1 + \alpha_2 + \dots + \alpha_n = s$ ， $\alpha_i > 0, i = 1, 2, \dots, n$ ，則必有

$$\left(\frac{s}{n}, \frac{s}{n}, \dots, \frac{s}{n}\right) \prec (\alpha_1, \alpha_2, \dots, \alpha_n) \prec (s, 0, \dots, 0)$$

因此在總量固定，且各分量皆不為負數的情況下，所能造出最集中的向量即為將總量集中在某一分量上的向量，而最分散的向量，為將各分量取平均的向量。套用在 Karamata 不等式上，即為 Karamata 不等式中所能求得的最大值及最小值。

最後要介紹延森不等式，延森不等式為 Karamata 不等式的一個特例。若取 $\alpha = (\bar{x}, \bar{x}, \dots, \bar{x})$ ， $\beta = (x_1, x_2, \dots, x_n)$ ，其中 $\bar{x} = \frac{x_1 + x_2 + \dots + x_n}{n}$ ，則因為 $(\bar{x}, \bar{x}, \dots, \bar{x}) \prec (x_1, x_2, \dots, x_n)$ ，故此時的 Karamata 不等式可寫成

$$n\phi(\bar{x}) \leq \sum_{i=1}^n \phi(x_i)$$

即為延森不等式。

例 6: 對銳角三角形 ABC ，試證

$$1 < \cos A + \cos B + \cos C \leq \frac{3}{2}$$

等號何時成立？

證明: 不失一般性，假設 $A \geq B \geq C$ ，則有 $A \geq \frac{\pi}{3}$ 且 $C \leq \frac{\pi}{3}$ 。因為 $\frac{\pi}{2} \geq A \geq \frac{\pi}{3}$ 且 $\pi \geq A + B = \pi - C \geq \frac{2\pi}{3}$ ，所以 $(\frac{\pi}{2}, \frac{\pi}{2}, 0) \succ (A, B, C) \succ (\frac{\pi}{3}, \frac{\pi}{3}, \frac{\pi}{3})$ 。取函數 $f(x) = \cos x$ ，

因為 $f''(x) = -\cos x \leq 0$, $x \in [0, \frac{\pi}{2}]$, 故在區間 $[0, \frac{\pi}{2}]$ 中 $f(x) = \cos x$ 為凹函數, 故根據 Karamata 不等式, 有

$$\begin{aligned} 1 &= \cos \frac{\pi}{2} + \cos \frac{\pi}{2} + \cos 0 \leq \cos A + \cos B + \cos C \\ &\leq \cos \frac{\pi}{3} + \cos \frac{\pi}{3} + \cos \frac{\pi}{3} = \frac{3}{2} \end{aligned}$$

考慮第一個不等式, 等號不會成立, 因為三角形中, 不會同時出現兩個直角。第二個不等式中, 等號成立若且唯若此三角形為正三角形。 \square

例 7: 設 $x, y, z \in (0, 1)$ 且 $\max(x, y, z) \leq \frac{x+y+z}{2} \leq 1$, 試證

$$\left(\frac{1+x}{1-x}\right) \left(\frac{1+y}{1-y}\right) \left(\frac{1+z}{1-z}\right) \leq \left(\frac{1+\frac{1}{2}(x+y+z)}{1-\frac{1}{2}(x+y+z)}\right)^2$$

證明: 令 $s = \frac{x+y+z}{2}$, 則根據題意有 $(x, y, z) \prec (s, s, 0)$ 。接著取

$$\phi(t) = \log \left(\frac{1+t}{1-t} \right)$$

且因 $\phi''(t) = \frac{4t}{(t^2-1)^2} > 0$, $0 < t < 1$, 故知 $\phi(t)$ 在 $(0, 1)$ 區間中為凸函數。最後, 利用 Karamata 不等式

$$\begin{aligned} \log \left(\frac{1+x}{1-x} \right) + \log \left(\frac{1+y}{1-y} \right) + \log \left(\frac{1+z}{1-z} \right) &\leq \log \left(\frac{1+s}{1-s} \right) + \log \left(\frac{1+s}{1-s} \right) \\ \log \left(\frac{1+x}{1-x} \cdot \frac{1+y}{1-y} \cdot \frac{1+z}{1-z} \right) &\leq \log \left(\frac{1+s}{1-s} \cdot \frac{1+s}{1-s} \right) \\ \left(\frac{1+x}{1-x} \right) \left(\frac{1+y}{1-y} \right) \left(\frac{1+z}{1-z} \right) &\leq \left(\frac{1+s}{1-s} \right)^2 = \left(\frac{1+\frac{x+y+z}{2}}{1-\frac{x+y+z}{2}} \right)^2 \end{aligned}$$

故得證。 \square

例 8 (生日問題): 隨機給定 n 個人, 且其生日視為在集合 $\{1, 2, \dots, 365\}$ 上獨立的均勻分佈, 試證明若 $n \geq 23$, 則其中有兩個或兩個以上的人生日相同的機率至少為 $1/2$ 。

考慮更廣義的情況, 若在沒有均勻假設的條件下, 試證明若 $n \geq 23$, 則其中有兩個或兩個以上的人生日相同的機率依舊至少為 $1/2$ 。

證明: 首先, 假設此問題中考慮的一年皆為 365 天, 不考慮閏年。在均勻分配的假設下, 給定任意 23 人, 則有兩個或兩個以上的人生日相同的機率為 $1 - (1 - 1/365) \cdot (1 - 2/365) \cdots (1 - 22/365) \approx 0.5079$ 。考慮更一般的情況, 若沒有均勻分佈的假設, 令 p_k 為一隨機抽取的人生

日落在第 k 天的機率, 且 $p_1 + p_2 + \cdots + p_{365} = 1$ 。考慮機率 $1 - e_{23}(p_1, p_2, \dots, p_{365})$, 又 $(\frac{1}{365}, \frac{1}{365}, \dots, \frac{1}{365}) \prec (p_1, p_2, \dots, p_{365})$, 根據例 4 基本對稱函數為一舒爾凹函數, 故有

$$e_{23}(p_1, p_2, \dots, p_{365}) \leq e_{23}\left(\frac{1}{365}, \frac{1}{365}, \dots, \frac{1}{365}\right)$$

所以

$$1 - e_{23}(p_1, p_2, \dots, p_{365}) \geq 1 - e_{23}\left(\frac{1}{365}, \frac{1}{365}, \dots, \frac{1}{365}\right)$$

故得證。關於此問題, 詳細的討論可參考 Clevenston and Watkins (1991) 和 Joag-Dev and Proschan (1992)。□

參考文獻

- Arnold, B.C. (2007). Majorization: Here, there and everywhere. *Statistical Science* **22** (3): 407.
- Clevenston, M.L. and Watkins, W. (1991). Majorization and the birthday inequality. *Mathematics Magazine* **64** (3): 183-188.
- Joag-Dev, K. and Proschan, F. (1992). Birthday problem with unlike probabilities. *Amer. Math. Monthly* **99** (1): 10-12.
- Lorenz, M.O. (1905). Methods of measuring the concentration of wealth. *Publications of the American Statistical Association* **9** (70): 209-219.
- Marshall, A.W. and Olkin, I. (1979). *Inequalities: Theory of Majorization and Its Applications*. New York: Academic Press.
- Marshall, A.W., Olkin, I. and Arnold, B. (2009). *Inequalities: Theory of Majorization and Its Applications*, 2nd edition. Springer, New York.
- Schur, I. (1923). Über eine Klasse von Mittelbildungen mit Anwendungen die Determinanten-Theorie Sitzungsber. Berlin. Math. Gesellschaft 22, 9-20. Issai Schur Collected Works (A. Brauer and H. Rohrbach, eds.) Vol. II. pp. 416-427. Berlin: Springer-Verlag, 1973.
- Steele, J.M. (2004). *The Cauchy-Schwarz Master Class: An Introduction to the Art of Mathematical Inequalities*. Cambridge: Cambridge University Press.
- Weyl, H. (1949). Almost periodic invariant vector sets in a metric vector space. *Amer. J. Math.* **71**, 178-205.

—本文作者陳柏宇任教新北市三多國中, 張福春任教中山大學應用數學系—

台北表現理論研討會IV

日期: 2013年12月20日(星期五) ~ 2013年12月23日(星期一)

地點: 台北市大安區羅斯福路四段1號 天文數學館

詳見中研院數學所網頁 <http://www.math.sinica.edu.tw>