平 均 論

李春億

一. 平均數的基本特性

設 $n \geq 2$, 某實數集 D 中的 n 個數 $x_1, x_2, \ldots, x_n$ 的平均數有下列數種:

1. 算術平均數 $\frac{x_1+x_2+\cdots+x_n}{n}$。$D$ 爲所有實數的集合。
2. 幾何平均數 $\sqrt[n]{x_1x_2\cdots x_n}$。$D$ 爲所有正實數的集合。
3. 調和平均數 $\frac{n}{\frac{1}{x_1}+\frac{1}{x_2}+\cdots+\frac{1}{x_n}}$。$D$ 可以選爲正實數的集合。
4. 方均根 $\sqrt{\frac{x_1^2+x_2^2+\cdots+x_n^2}{n}}$。$D$ 可以選爲正實數的集合。
5. 反調和平均數 $\frac{x_1^2+x_2^2+\cdots+x_n^2}{x_1+x_2+\cdots+x_n}$。$D$ 可以選爲正實數的集合。

此外尙有兩個正數的希羅平均數 $\frac{a+\sqrt{ab}+b}{3}$。這希羅平均數似乎沒有推廣到 n 個數的理想辦法。

現在數學只是定義出如上幾個特定的平均數, 並討論其個別特性及彼此間的關係, 而缺乏完整的理論。本文將試著定義出較完整的平均數的理論, 以創立一個平均論 (Average Theory)。

若 $x_1, x_2, \cdots, x_n$ 均爲實數的集合 D 中的數, $a = m(x_1, x_2, \cdots, x_n)$ 是類似上述各種平均數的一項函數。我們稱得 m 爲平均法 (averaging method), a 爲平均數 (average) 或平均值 (average value), n 爲項次, D 爲平均數的定義域。我們至少對平均法作以下的要求:

1、確定性: 對任意正整數 $n \geq 2$, m 均可視爲一個 n 個變數的函數。換言之, a 可由 $x_1, x_2, \ldots, x_n$ 完全決定。

2、封閉性: $a \in D$.

3、居中性: $\min(x_1, x_2, \ldots, x_n) \leq a \leq \max(x_1, x_2, \ldots, x_n)$。

4、對稱性: 若$i_1, i_2, \ldots, i_n$ 爲 $1, 2, \ldots, n$ 的一個排列, 則

$$m(x_{i_1}, x_{i_2}, \ldots, x_{i_n}) = m(x_1, x_2, \ldots, x_n).$$

5、次方遞迴性: 若 $a_1 = m(x_{1,1}, x_{1,2}, \ldots, x_{1,n^k})$, $a_2 = m(x_{2,1}, x_{2,2}, \ldots, x_{2,n^k})$,..., $a_n =$

$m(x_{n,1}, x_{n,2}, \ldots, x_{n,n^k})$, 則

$$m(a_1, a_2, \ldots, a_n) = m(x_{1,1}, \ldots, x_{1,n^k}, x_{2,1}, \ldots, x_{2,n^k}, \ldots, x_{n,1}, \ldots, x_{n,n^k}).$$

對稱性基本上就是交換律的延伸。次方遞迴性類似結合律。以上定義的種種平均數都有確定性、對稱性、定義域的選擇也使它們有封閉性。但反調和平均數沒有居中性和性。我們願意把它排除在要討論的平均論之外。

二. 運算平均數

設 D 爲實數的集合, $p(x,y)$ 爲 $x \in D$, $y \in D$ 的函數。若 $p(x,y) \in D$, 且對 $x, y, z \in D$ 滿足

交換律: $p(x,y) = p(y,x)$,

結合律: $p(p(x,y),z) = p(x,p(y,z))$,

則我們稱 p 爲一個運算。我們可以用遞迴的方式定義多於兩個運算元的運算如下:

$$p(x_1, x_2, \ldots, x_n) = p(p(x_1, x_2, \ldots, x_{n-1}), x_n).$$

這樣定義的運算具有確定性、對稱性、封閉性和組合性。

設 p 爲集合 D 上的一運算, $x_1, x_2, \ldots, x_n \in D$。若有一數 a 使

$$p(\underbrace{a, a, \ldots, a}_{n\text{項}}) = p(x_1, x_2, \ldots, x_n),$$

我們則暫時把 a 叫作由 p 所定義的 $x_1, x_2, \ldots, x_n$ 的運算平均數。

例如: 若 $p(x,y) = x+y$, 則由 p 所定義的運算平均數爲算術平均數。

若 $p(x,y) = xy$, 則由 p 所定義的運算平均數爲幾何平均數。

若 $p(x,y) = xy$, 則由 p 所定義的運算平均數爲幾何平均數。

若 $p(x,y) = \frac{1}{x} + \frac{1}{y}$, 則由 p 所定義的運算平均數爲調和平均數。

若 $p(x,y) = x^2 + y^2$, 則由 p 所定義的運算平均數爲方均根。

若 $p(x,y) = \max(x,y)$, 則由p所定義的運算平均數爲最大值。

若 $p(x,y) = \min(x,y)$, 則由p所定義的運算平均數爲最小值。

但希羅平均數和反調和平均數似乎都不是運算平均數。

定理1: 運算平均數具次方遞迴性。

證明: 設 m 爲由運算 p 所得的運算平均數, $a_i = m(x_{i,1}, x_{i,2}, \ldots, x_{i,n^k})$, $i = 1, 2, \ldots, n$, $a = m(a_1, a_2, \ldots, a_n)$, 則 $p(x_{i,1}, x_{i,2}, \ldots, x_{i,n^k})$ 等於 n^k 個 a_i 經過運算 p 所運算出來的結果。由結合律, $p(x_{1,1}, \ldots, x_{1,n^k}, x_{2,1}, \ldots, x_{2,n^k}, \ldots, x_{n,1}, \ldots, x_{n,n^k})$ 等於 $a_1, a_2, \ldots, a_n$ 諸數每個出現 n^k 次在 p 下計算的結果。但每組 $a_1, a_2, \ldots, a_n$ 在 p 下的運算相當於 n 個 a 在 p 下的運算, 所以 $p(x_{1,1}, \ldots, x_{1,n^k}, x_{2,1}, \ldots, x_{2,n^k}, \ldots, x_{n,1}, \ldots, x_{n,n^k})$ 等於 $n \cdot n^k = n^{k+1}$ 個 a 在 p 下運算的結果。這就是說

$$a = m(x_{1,1}, \ldots, x_{1,n^k}, x_{2,1}, \ldots, x_{2,n^k}, \ldots, x_{n,1}, \ldots, x_{n,n^k}).$$

利用和定理1相似的證明, 我們也可以證明。

定理2: 設m爲運算平均法,

$$\begin{aligned}
a &= m(x_1, x_2, \ldots, x_n),\\
a_1 &= m(a, x_2, \ldots, x_n),\\
a_2 &= m(x_1, a, \ldots, x_n),\\
&\vdots\\
a_n &= m(x_1, x_2, \ldots, a),\\
a_0 &= m(a_1, a_2, \ldots, a_n).
\end{aligned}$$

則 $a_0 = a$。

定理2是運算平均數的一個必要條件, 當它不成立時, 可以證明這種平均數不是運算平均數。另一方面, 雖然由定理1知運算平均數具有次方遞迴性, 但在別的性質上就不那麼理想了。首先, 若 $12 \in D$, 而令 $p(x, y) = 12$, 則 p 是 D 上的一個運算, 但這時可以得到任意 D 中的數都是 D 中任意多個數在 p 下的運算平均數, 所以在這個 p 下沒有確定性。再者, 令 D 爲正整數的集合, 若 $p(x, y) = \text{g.c.d}(x, y)$, 則 p 是一個好的運算, 但 $p(4, 6) = 2 < min(4, 6)$。所以這運算平均數沒有居中性。同理 l.c.m. 運算也沒有居中性。由於這些問題, 我們必須縮減運算的範圍, 以達到適當的平均數的理論。

三. 函數平均數

在本節中我們把運算平均數特殊化, 選擇呈下形的運算:

$$p(x, y) = f(x) + f(y),$$

式中 f 爲一特定函數, 其定義域爲一區間 (u, v)。我們對 f 作以下的要求:

1. $f:(u,v)\to(u,v)$ 爲二階可微的單調函數。
2. $\dfrac{f''(x)}{f'(x)}$ 連續。

這時我們稱由運算 $f(x)+f(y)$ 所定義的平均數爲函數平均數

$$m(x_1,x_2,\ldots,x_n)=f^{-1}\Big[\frac{1}{n}\sum_{i=1}^{n}f(x_i)\Big].$$

函數 f 叫作平均化函數 (averaging function)。

函數平均數顯然具有確定性、封閉性、居中性、對稱性。由定理1知它也具有次方遞迴性。所以在研究平均數時, 它是很好的研究對象。本文以下以討論函數平均數爲主。

設 $f(x)$ 爲平均化函數, 則 $f'(x)\neq 0$。由不定積分知 $\ln|f'(x)|$ 連續。但 $f(x)$ 是單調函數, 所以 $f'(x)$ 不改變符號, 因而連續。再從條件2知 $f''(x)$ 連續。

定理 3: 設 $f(x)$ 和 $g(x)$ 均爲區間 (u,v) 上的平均化函數, 則下列四條件同值:

(1) $\dfrac{f''(x)}{f'(x)}=\dfrac{g''(x)}{g'(x)}$;

(2) $f(x)$ 和 $g(x)$ 互相可以表爲一次式. 即有常數 c_0 和 $c_1\neq 0$ 使

$$f(x)=c_0+c_1g(x).$$

(3) 若 $x_1,x_2,\ldots,x_n\in(u,v)$, 則

$$f^{-1}\Big(\frac{1}{n}\sum_{i=1}^{n}f(x_i)\Big)=g^{-1}\Big(\frac{1}{n}\sum_{i=1}^{n}g(x_i)\Big).$$

(4) 若 $x,y\in(u,v)$, 則

$$f^{-1}\Big(\frac{1}{2}[f(x)+f(y)]\Big)=g^{-1}\Big(\frac{1}{2}[g(x)+g(y)]\Big).$$

證明: (1) $\Rightarrow$ (2). 從 (1) 得

$$\Big[\frac{f'(x)}{g'(x)}\Big]'=\frac{g'(x)f''(x)-f'(x)g''(x)}{g'(x)^2}=0.$$

所以有常數 c_1 使 $f'(x)=c_1g'(x)$。再積分便得 (2)。

(2) $\Rightarrow$ (1). 從 (2), $f'(x)=c_1g'(x)$。再微分得 $f''(x)=c_1g''(x)$。相除即得 (1) 式。

(2) $\Rightarrow$ (3). 令 $f(x)=y$。$g(x)=\frac{1}{c_1}(y-c_0)$; $f^{-1}(y)=x=g^{-1}(\frac{1}{c_1}(y-c_0))$。於是

$$\begin{aligned}f^{-1}(\frac{1}{n}\sum_{i=1}^{n}f(x_i))&=f^{-1}(\frac{1}{n}\sum_{i=1}^{n}[c_0+c_1g(x_i)])\\&=f^{-1}(c_0+c_1\frac{1}{n}\sum_{i=1}^{n}g(x_i))=g^{-1}(\frac{1}{n}\sum_{i=1}^{n}g(x_i))\end{aligned}$$

(3) $\Rightarrow$ (4). (4) 其實就是 (3) 中 $n=2$ 的特殊情形。

(4) $\Rightarrow$ (2). 設 $fg^{-1}=h$, $g(x)=\xi$, $g(y)=\eta$。(4) 式可以看成

$$\frac{h(\xi)+h(\eta)}{2}=h(\frac{\xi+\eta}{2}).$$

將 ξ 和 η 看成獨立變數。兩邊關於 ξ 作偏微分, 得 $\frac{1}{2}h'(\xi)=\frac{1}{2}h'(\frac{\xi+\eta}{2})$。再關於法無 η 作偏微分, 得 $\frac{1}{4}h''(\frac{\xi+\eta}{2})=0$。因 $\frac{\xi+\eta}{2}$ 可以是 g 的值域中的任意數, 故得 $h''(\xi)=0$, 即 $h(\xi)=c_0+c_1+\xi$, 其中 c_0 和 c_1 是常數。但這就 (2)。

以上討論的平均化函數 $f(x)$ 的定義域均爲一開區間 (u,v)。若 $\lim_{x\to u^+}f(x)$存在, 我們可以把 u 加到 $f(x)$ 的定義域裡去。同理若 $\lim_{x\to v^-}f(x)$ 存在, 我們也可以把 v 加到 $f(x)$ 的定義域裡去。下表列出幾個平均化函數:

$f(x)$	$\frac{f''(x)}{f(x)}$	定義域
x^t, $t>0$	$\frac{t-1}{x}$	$[0,\infty)$
x^t, $t<0$	$\frac{t-1}{x}$	$(0,\infty)$
e^x	1	$(-\infty,\infty)$
$\ln x$	$-\frac{1}{x}$	$(0,\infty)$
$\cos x$	$\cot x$	$[0,\pi]$
$\sin x$	$-\tan x$	$[-\frac{\pi}{2},\frac{\pi}{2}]$
$\tan x$	$2\tan x$	$(-\frac{\pi}{2},\frac{\pi}{2})$

設 $f(x)$ 爲 $f(x)$ 的反函數。若 $f(x)$ 爲單調, $F(x)$ 亦爲單調。若 $\frac{f''(x)}{f'(x)}$ 爲連續, 則

$$\frac{F''(x)}{F'(x)}=-\frac{f''(F(x))}{[f'(F(x))]^2}=\frac{f''(F(x))}{f'(x)}\frac{-1}{f'(F(x))}$$

亦爲連續。因此, 一個平均化函數的反函數亦爲平均化函數。

四. 函數平均數間的不等式

算術平均數的平均化函數爲 $f(x)=x$。此時 $\frac{f''(x)}{f'(x)}=0$。幾何平均數的平均化函數爲 $g(x)=\ln x$, $(x>0)$。此時 $\frac{g''(x)}{g'(x)}=-\frac{1}{x}$。我們知道算術平均數大於幾何平均數, 同時也得到 $0\geq-\frac{1}{x}$。其實這現象是普遍的, 即

定理4: 設 $f(x)$ 和 $g(x)$ 均爲區間 (u,v) 上的平均化函數, 則下列三條件同值:

(1) $\dfrac{f''(x)}{f'(x)} \geq \dfrac{g''(x)}{g'(x)}$;
(2) 若$x_1, x_2, \ldots, x_n \in (u, v)$, 則

$$f^{-1}\Big(\frac{1}{n}\sum_{k=1}^{n} f(x_k)\Big) \geq g^{-1}\Big(\frac{1}{n}\sum_{k=1}^{n} g(x_k)\Big).$$

(3) 若 $x_1, x_2, \ldots, x_n \in (u, v)$, $w_1, w_2, \ldots, w_n$ 爲正實數, $w_1 + w_2 + \cdots + w_n = 1$, 則

$$f^{-1}\Big(\sum_{k=1}^{n} w_k f(x_k)\Big) \geq g^{-1}\Big(\sum_{k=1}^{n} w_k g(x_k)\Big).$$

註: (3) 中左右二算式叫加權平均數。

證: (1) $\Rightarrow$ (2). 因 $f'(x) \neq 0$ 連續, 故 $f'(x)$ 在 (u, v) 上符號不變。以下就 $f'(x) > 0$ 的情形立。令 $\xi = g(x)$, $h(\xi) = f(g^{-1}(\xi))$, $\xi_k = g(x_k)$。選擇 $\bar{x} \in (u, v)$ 使 $g(\bar{x}) = \frac{1}{n}[g(x_1) + g(x_2) + \cdots + g(x_n)]$。令 $\bar{\xi} = g(\bar{x})$。$\bar{\xi} = \frac{1}{n}(\xi_1 + \xi_2 + \cdots + \xi_n)$。現在計算 $h(\xi)$ 的導函數:

$$\begin{aligned}
h'(\xi) &= \frac{f'(g^1(\xi))}{g'(g^{-1}(\xi))} = \frac{f'(x)}{g'(x)},\\
h''(x) &= \frac{1}{g'(x)}\frac{d}{dx}\frac{f'(x)}{g'(x)}\\
&= \frac{g'(x)f''(x) - f'(x)g''(x)}{g'(x)^3} = \frac{f'(x)}{g'(x)}\Big[\frac{f''(x)}{f'(x)} - \frac{g''(x)}{g'(x)}\Big].
\end{aligned}$$

注意 $g'(x)$ 在 (u, v) 中的符號也不變。若$g'(x) > 0$, $h'(\xi)$ 爲增函數。此時 $h''(\xi) > 0$。當 $x_k < \bar{x}$ 時, $\xi_k < \bar{\xi}$, 由微分學的均值定理可得

$$\frac{h(\xi_k) - h(\bar{\xi})}{\xi_k - \bar{\xi}} \leq h'(\bar{\xi}),$$

$$f(x_k) - f(\bar{x}) \geq h'(\bar{\xi})[g(x_k) - g(\bar{x})],$$

即

$$f(\bar{x}) \leq f(x_k) - h'(\bar{\xi})[g(x_k) - g(\bar{x})].$$

仿此當 $x_k > \bar{x}$ 時亦有此式。當 $x_k = \bar{x}$ 時, 此不等式爲等式。將此式關於法無k求算術平均數, 乃得 $f(\bar{x}) \leq \frac{1}{n}\sum f(x_k)$, 即

$$g^{-1}\Big[\frac{1}{n}\sum g(x_k)\Big] = g^{-1}g(\bar{x}) = \bar{x} \leq f^{-1}\Big[\frac{1}{n}\sum f(x_k)\Big].$$

若 $g'(x) < 0$, $h'(\xi)$ 爲減函數, 我們將上述的討論作適當的修改, 仍可以得到相同的結論。

$(2) \Rightarrow (3)$. 先設 w_k 均爲有理數。經通分後可令 $w_k = m_k/m$。$m_1+m_2+\cdots+m_n = m$ (3) 式的兩邊分別是 m_1 個 x_1, m_2 個 x_2,$\ldots$, m_n 個 x_n 在平均化函數 f 和 g 下的函數平均數, 所以從 (2) 知 (3) 成立。利用任意實數都可以用一系列的實數來逼近的原理, 我們可以把這結果推展至 w_k 爲實數的情形。

$(3) \Rightarrow (1)$. 我們就 $f'(x) > 0$ 的情形立論, $f'(x) < 0$ 的情形仿此不贅。

取 $x, y, z \in (u, v)$。令 $\xi = g(x)$, $\eta = g(y)$, $\zeta = g(z)$。我們假定 $\xi < \eta < \zeta$, 則 $\frac{\eta-\xi}{\zeta-\xi}$, $\frac{\zeta-\eta}{\zeta-\xi}$ 均爲正, 且其和爲1。從 (2) 得

$$f^{-1}\Big(\frac{\eta-\xi}{\zeta-\xi}f(z)+\frac{\zeta-\eta}{\zeta-\xi}f(x)\Big) \ge g^{-1}\Big(\frac{\eta-\xi}{\zeta-\xi}\zeta+\frac{\zeta-\eta}{\zeta-\xi}\xi\Big) = g^{-1}(\eta).$$

定義函數 h 如下: $h(\xi) = f(g^{-1}(\xi))$, 則上式可以寫作

$$\frac{\eta-\xi}{\zeta-\xi}h(\zeta)+\frac{\zeta-\eta}{\zeta-\xi}h(\xi) \ge h(\eta).$$

於是有

$$h(\zeta)-h(\eta) \ge \frac{\zeta-\eta}{\zeta-\xi}h(\zeta)-\frac{\zeta-\eta}{\zeta-\xi}h(\xi) \ge h(\eta).$$

故

$$\frac{h(\zeta)-h(\eta)}{\zeta-\eta} \ge \frac{h(\zeta)-h(\xi)}{\zeta-\xi}.$$

令 $\eta \to \zeta$, 則得

$$h'(\zeta) \ge \frac{h(\zeta)-h(\xi)}{\zeta-\xi}.$$

仿此, $\frac{h(\eta)-h(\xi)}{\eta-\xi} \le \frac{h(\zeta)-h(\xi)}{\zeta-\xi}$。令 $\eta \to \zeta$, 則得

$$h'(\xi) \le \frac{h(\zeta)-h(\xi)}{\zeta-\xi}.$$

所以 $h'(\zeta) \ge h'(\xi)$, 即 $h'(\xi)$ 是 ξ 的增函數, 所以 $h''(\xi) \ge 0$。

以下計算 $h''(\xi)$。從 $h(\xi)$ 的定義知

$$h'(\xi) = f'(g^{-1}(\xi))(g^{-1})'(\xi) = \frac{f'(x)}{g'(x)}.$$

遂有

$$0 \le h''(\xi) = \frac{g'(x)f''(x)-f'(x)g''(x)}{g'(x)^2}\frac{dx}{d\xi} = \Big[\frac{f''(x)}{f'(x)}-\frac{g''(x)}{g'(x)}\Big]\frac{f'(x)}{g'(x)^2}.$$

由假定, $f'(x) > 0$。故得

$$\frac{f''(x)}{f'(x)} \ge \frac{g''(x)}{g'(x)}.$$

五. 算幾不等式

算幾不等式 (算術平均數大於或等於幾何不均數) 是函數平均數不等式的一個例子。算術平均數的平均化函數爲 $f(x)=x$, 幾何平均數的平均化函數爲 $g(x)=\ln x$。二者定義域的交界爲所有正實數的集合。因爲

$$\frac{f''(x)}{f'(x)}=0<-\frac{1}{x}=\frac{g''(x)}{g'(x)},$$

所以有算幾不等式。由於這不等式的重要性, 以下我們另給它一個簡短的證明:

定理5: 設 $x_1,x_2,\ldots,x_n$ 爲 n 個正數,

$$a=\frac{x_1+x_2+\cdots+x_n}{n}, \qquad g=\sqrt[n]{x_1x_2\cdots x_n}$$

分別爲它們的算術平均數和幾何平均數, 則 $a\geq g$。其中等號當且僅當所有 x_k 都等於法無 a 時成立。

證: 令 $f(x)=\ln x$。$f'(a)=\frac{1}{a}$, $f''(x)=-\frac{1}{x^2}$. 由 Taylor 定理可得對每個 x_k 均有在 a 和 x_k 間的一數 ξ_k, 使

$$\ln x_k=\ln a+\frac{1}{a}(x_k-a)-\frac{1}{\xi_k^2}(x_k-a)^2.$$

因此

$$\begin{aligned}\ln g&=\frac{1}{n}\sum\ln x_k=\ln a-\frac{1}{na}(\sum x_k-a)-\frac{1}{n}\sum\frac{1}{\xi_k^2}(x_k-a)^2\\&=\frac{1}{n}\sum\ln x_k=\ln a-\frac{1}{n}\sum\frac{1}{\xi_k^2}(x_k-a)^2\\&\leq\ln a.\end{aligned}$$

遂有 $a\geq g$。因 $\frac{1}{\xi_k^2}>0$, 所以等號成立的充要條件是每個 x_k-a 都等於0。

參考資料

1. 數學史概論, 著者: Howard Eves, 譯者: 歐陽絳. 曉園出版社.
2. 凸分析導論, 著者: (荷蘭) J. D. 蒂爾 (van Tiel). 亞東書局.

—本文作者任職於瑞昱半導體—