凸函數與其應用實例

何景國

數播第六卷第二期中，王中烈教授曾指出：數學問題不是孤立存在的；數學學習者不可以專注重於問題本身的演練，及只講求問題的演算、推導、求證等等的技巧，而要着重問題的關連與發展，以及其方法的應用與改進。王教授且在文中提及凸函數與不等式的關係，本文即就就函數圖形之凹凸性談凸函數，進而就不等式的範圍例舉幾個凸函數在不等式上的應用例子。下面就是引發出凸函數概念的一個定義。

設 I 爲實數線上一任意區間，且 f 爲定義在 I 上的函數。

所謂凸函數就是指函數 $y=f(x)$ 的圖形是上凹的。亦即是這個函數 $y=f(x)$ 具有下列性質：

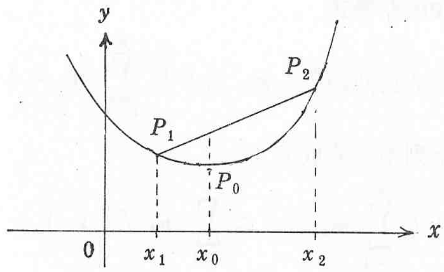
如果對任意 $x_1, x_2 \in I$，且 $x_1 < x_0 < x_2$ 時，則圖形 $y=f(x)$ 上三點 $P_i(x_i, f(x_i))$，$i=0, 1, 2$，恒滿足：點 P_0 會在 $\overline{P_1P_2}$ 線段之下方。（見右上圖）

自上述定義易知下面的凸函數的解析定義。

設 m_1，m_2 爲兩正實數，且 $m_1+m_2=1$ 若對任意 $x_1, x_2 \in I$，下列不等式成立：

$$f(m_1x_1+m_2x_2)$$
$$\leqslant m_1 f(x_1)+m_2 f(x_2)$$

則稱這個函數 f 爲在區間 I 上的凸函數。



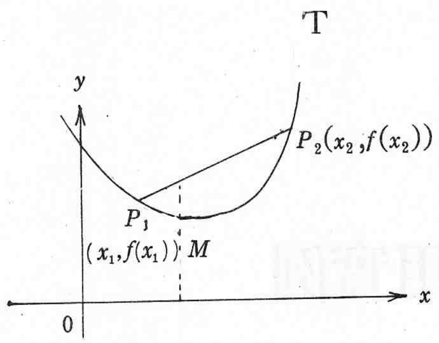
分析與證明：

在具有上凹性之函數圖形 Γ 上任取兩點 $P_1(x_1, f(x_1))$，$P_2(x_2, f(x_2))$ 與在 $\overline{P_1P_2}$ 線段上取一點 $P(x, f(x))$，且從 P 作 x 軸的一垂線交圖形 Γ 於點 M。得一參數向量式：

$$\overrightarrow{OP}=(1-t)\overrightarrow{OP_1}+\overrightarrow{OP_2}$$

其中 $t\in〔0, 1〕$

由函數圖形 Γ：爲上凹的，且 $|\overrightarrow{MP}|=|\overrightarrow{OP}$



$-OM|$ ，則：

$$(1-t)f(x_1)+tf(x_2)$$
$$-f((1-t)x_1+tx_2)\geqslant 0$$

令 $m_1=1-t$ ， $m_2=t$

則上式變成

$$f(m_1x_1+m_2x_2)$$
$$\leqslant m_1f(x_1)+m_2f(x_2)$$

上面式子就是用來當作凸函數的解析定義，如果上式中的不等號反向，則稱 f 為在區間 I 上的凹函數。

一般地，我們有下面的 Jensen 不等式。

Jensen 不等式

若 f 為凸函數，$m_i\geqslant 0$ 且 $\sum_{i=1}^{n} m_i=1$

則有下列不等式成立：

$$f\left(\sum_{i=1}^{n} m_ix_i\right)\leqslant\sum_{i=1}^{n} m_i f(x_i)$$

設 $n=2$ 時，則得最簡單情形：

$$f(m_1x_1+m_2x_2)$$
$$\leqslant m_1f(x_1)+m_2f(x_2)$$

此時，只不過是凸函數的定義。今假設 $n=k$ 時上式成立。

當 $m_i\geqslant 0$ ， $\sum_{i=1}^{n} m_i=1$ ，則有下列不等式:

$$f\left(\sum_{i=1}^{k+1} m_ix_i\right)$$
$$=f\left(\sum_{i=1}^{k-1} m_ix_i+m_kx_k+m_{k+1}x_{k+1}\right)$$

或

$$f\left(\sum_{i=1}^{k+1} m_ix_i\right)$$
$$=f\left(\sum_{i=1}^{k-1} m_ix_i+(m_k+m_{k+1})\left(\frac{m_k}{m_k+m_{k+1}}x_k+\frac{m_{k+1}}{m_k+m_{k+1}}x_{k+1}\right)\right)$$

由題設知 m_k ， m_{k+1} 中，至少有一不為零，因此得：

$$f\left(\sum_{i=1}^{k+1} m_ix_i\right)\leqslant\sum_{i=1}^{k-1} m_if(x_i)$$
$$+(m_k+m_{k+1})f\left(\frac{m_k}{m_k+m_{k+1}}x_k\right.$$
$$\left.+\frac{m_{k+1}}{m_k+m_{k+1}}x_{k+1}\right)$$

而其中

$$f\left(\frac{m_k}{m_k+m_{k+1}}x_k+\frac{m_{k+1}}{m_k+m_{k+1}}x_{k+1}\right)$$
$$\leqslant\frac{m_k}{m_k+m_{k+1}}f(x_k)+\frac{m_{k+1}}{m_k+m_{k+1}}f(x_{k+1})$$

故

$$f\left(\sum_{i=1}^{k+1} m_ix_i\right)\leqslant\sum_{i=1}^{k+1} m_if(x_i)$$

因此 Jensen 不等式對任何自然數 n 恒成立。

對於滿足上述的 m_1 ，m_2 ，… ，m_n 我們稱 $m_1x_1+m_2x_2+\cdots\cdots+m_nx_n$ 為 x_1， x_2，…， 與 x_n 的凸組合。由向量座標幾何知，凸組合與凸集合及質量中心的向量解釋有着重要的關係，因此可以推論到更多的性質。

定理 1 設 $P_i(m_i)$ ($i=1$ ，2 ，… ，n) 為點集合 Γ 中任意一組質點。若點 G 為該組質點的質量中心，則 Γ 為一凸集合的充要條件是點 G 亦屬於集合 Γ 。

分析與證明：

(1)充分條件：

在凸集合 Γ 中任取兩質點 P_1 與 P_2 ，且其負荷代數量為 m_1 ，m_2 若點 G 為兩質點 P_1,P_2

的質量中心則下式成立。

$$m_1\overrightarrow{GP_1}+m_2\overrightarrow{GP_2}=\vec{0}$$

（式中$m_1\geqslant 0$，$m_2\geqslant 0$，$m_1+m_2=1$）

或

$$\overrightarrow{P_1G}=\frac{m_2}{m_1+m_2}\overrightarrow{P_1P_2}$$

由於實數$\dfrac{m_2}{m_1+m_2}\in[0,1]$，因此得知點$G$落在$\overline{P_1P_2}$線段上。亦即有

$$G\in\Gamma$$

成立。

今假設有一組$n=k$個的質點$P_1, P_2, \cdots, P_k$的質量中心G'（其中$m_i\in R^+$，$i=1, 2, \cdots, k$）為集合Γ的一元素。

令 $$m_i\geqslant 0\ ,\quad \sum_{i=1}^{k+1}m_i=1$$

則兩質點P_{k+1}與G'的質量中心滿足下列式子：

$$\overrightarrow{G'G}=\frac{m_{k+1}}{(m_1+m_2+\cdots+m_k)+m_{k+1}}\overrightarrow{G'P_{k+1}}$$

且因為

$$\frac{m_k}{(m_1+m_2+\cdots+m_k)+m_{k+1}}\in[0,1]$$

故點G落在$\overline{G'P_{k+1}}$線段上。亦即有

$$G\in\Gamma$$

因此對任意一組n個質點的質量中心恒為點集合的一元素。

(2)必要條件：

設P_1, P_2為點集合Γ中任意的相異兩點，則線段$\overline{P_1P_2}$上任一點P會使下式成立：

$$(1-t)\overrightarrow{PP_1}+t\,\overrightarrow{PP_2}=\vec{0}$$

其中 $$0<t<1$$

若令 $$m_1=1-t\ ,\quad m_2=t$$

則有 $$m_1\overrightarrow{PP_1}+m_2\overrightarrow{PP_2}=\vec{0}$$

上述式子說明了兩質點P_1，P_2的質量中心為點P，而P亦是Γ的一元素，因此集合Γ為一凸集合。

故定理1得證。

定理2　設$f(x)$為定義在區間I上的函數，且點集合

$$\Gamma=\{(x,y)\mid x\in I\ ,\ y\geqslant f(x)\}$$

則函數$f(x)$為一凸函數的充要條件是集合Γ為一凸集合。

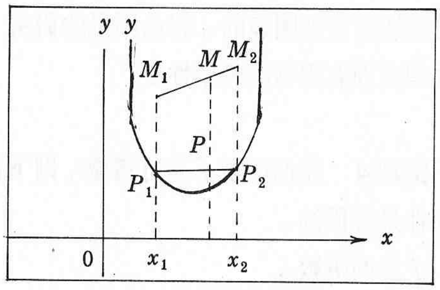
分析與證明：

設$M_1(x_1, y_1)$，$M_1(x_2, y_2)$為Γ中的兩點，且圖形$f(x)$上兩點$P_1(x_1,f(x_1))$，$P_2(x_2, f(x_2))$則線段P_1P_2上任意點P的座標$(x,f(x))$滿足下列等式：

$$\begin{cases}x=m_1x_1+m_2x_2\\ f(x)=m_1f(x_1)+m_2f(x_2)\end{cases}$$

其中$m_1, m_2>0$，且$m_1+m_2=1$

因此線段$\overline{M_1M_2}$上，點M的座標(x,y)會滿足下列等式：



$$y\geqslant m_1f(x_1)+m_2f(x_2)=f(x)$$

故

$$M(x,y)\in\Gamma$$

即點集合Γ為一凸集合。

反之，如果點集合

$$\Gamma=\{(x,y)\mid x\in I\ ,\ y\geqslant f(x)\}$$

為一凸集合則由定義易知函數$f(x)$為一凸函數。

故定理2得證。

定理3　設函數f為凸集合Γ上的凸函數則下列不等式成立。

$$f\left(\sum_{i=1}^{n}m_ix_i\right)\leqslant\sum_{i=1}^{n}m_if(x_i)$$

其中 $m_i\in R^+$，$i=1, 2, \cdots, n$

且　$$\sum_{i=1}^{n} m_i = 1$$

分析與證明：

設 $x_1, x_2, \cdots, x_n$ 爲函數 f 圖形上相異點 $P_1, P_2, \cdots, P_n$ 的橫座標，且一組 n 個正實數 $m_1, m_2, \cdots, m_n$ 爲各點 P_i 的負荷質量使得：

$$m_1 + m_2 + \cdots\cdots + m_n = 1$$

則由定理 1 知整個質點系 $P_1, P_2, \cdots, P_n$ 的質量中心必屬於集合 Γ，故下不等式成立。

$$f\left(\sum_{i=1}^{n} m_i x_i\right) \leqslant \sum_{i=1}^{n} m_i f(x_i)$$

其中　$m_i > 0$　且　$\sum_{i=1}^{n} m_i = 1$

根據上述分析，凸函數的解析定義可以確定函數的凸性，但是函數的凸性與函數的可導微性及可積性是密切關連的。對凸函數的研究，下面幾個定理有相當的應用性。

定理４　設凸函數 f 是可導微，則下面 5 個條件是等價的。

(1)　f 爲凸函數。

(2)　f 滿足下列不等式：

$$f\left(\sum_{i=1}^{n} m_i x_i\right) \leqslant \sum_{i=1}^{n} m_i f(x_i)$$

式中　$m_i > 0$　且　$\sum_{i=1}^{n} m_i = 1$，

$i = 1, 2, \cdots, n$

(3)　$f'' \geqslant 0$

(4)　f' 爲遞增函數

(5)　$y = f(x)$ 函數圖形恒在某切線的上方。

定理５

(1)　若兩函數 f 與 g 爲凸函數則函數 $(f+g)$ 及 (αf) 均爲凸函數，其中 α 爲一純量。

(2)　設函數 f 與 g 分別爲定義在區間 I 與 J 上的凸函數，且 f 的值域 R_f 包含於 J 內。若 g 爲遞增函數則函數 f 與 g 的合成函數 $(g \circ f)$ 爲定義在 I 上的凸函數。

分析與證明：

(1)　本定理容易驗證，故從略。

(2)　設 $x_1, x_2 \in I$，且 $m_1 m_2 > 0$，$m_1 + m_2 = 1$，則由下式不等式知 $(g \circ f)$ 爲定義在 I 上的凸函數。

$$\begin{aligned} & g(f(m_1 x_1 + m_2 x_2)) \\ \leqslant\ & g(m_1 f(x_1) + m_2 f(x_2)) \\ \leqslant\ & m_1 g(f(x_1)) + m_2 g(f(x_2)) \end{aligned}$$

定理６　設 f 與 g 均爲定義在同一區間 I 上的凸函數，若 f 與 g 同爲遞減函數（或遞增函數）則積函數 $h(x) = f(x) \cdot g(x)$ 亦爲凸函數。

分析與證明：

設　$x_1 < x_2$　且　$m_1 > 0$，$m_2 > 0$，$m_1 + m_2 = 1$　則依題意得：

$$[f(x_1) - f(x_2)] \cdot [g(x_2) - g(x_1)] \leqslant 0$$

或

$$\begin{aligned} & f(x_1) g(x_2) + f(x_2) g(x_1) \\ \leqslant\ & f(x_1) g(x_1) + f(x_2) g(x_2) \quad \cdots\cdots \text{①} \end{aligned}$$

又因爲

$$\begin{aligned} & h(m_1 x_1 + m_2 x_2) \\ =\ & f(m_1 x_1 + m_2 x_2) \\ & \cdot g(m_1 x_1 + m_2 x_2) \end{aligned}$$

所以

$$\begin{aligned} & h(m_1 x_1 + m_2 x_2) \\ \leqslant\ & [m_1 f(x_1) + m_2 f(x_2)] \\ & [m_1 g(x_1) + m_2 g(x_2)] \end{aligned}$$

或

$$\begin{aligned} \leqslant\ & m_1^2 f(x_1) g(x_1) \\ & + m_1 m_2 [f(x_1) g(x_2) + f(x_2) g(x_1)] \\ & + m_2^2 f(x_2) g(x_2) \end{aligned}$$

故由①式得

$$\begin{aligned} & h(m_1 x_1 + m_2 x_2) \\ \leqslant\ & m_1^2 f(x_1) g(x_1) \\ & + m_1 m_2 [f(x_1) g(x_1) + f(x_2) g(x_2)] \end{aligned}$$

$$+m_2^2 f(x_2)g(x_2)$$
$$\leqslant m_1 f(x_1)g(x_1)+m_2 f(x_2)g(x_2)$$

即有下列不等式：

$$h(m_1x_1+m_2x_2)$$
$$\leqslant m_1h(x_1)+m_2h(x_2)$$

因此知 $h(x)=f(x)\cdot g(x)$ 爲凸函數。

為了闡明凸函數與不等式的關係，我們舉數則不等式的例子如下：

論例 1 ── 高次平均不等式

設 $x_1, x_2, \cdots, x_n$ 爲 n 個正數，而 $m_1, m_2, \cdots, m_n$ 爲狹義凸性係數，即諸 $m_i>0$，$\sum m_i=1$，$i=1,2,\cdots,n$。

且

$$\begin{cases} \mu=m_1x_1+m_2x_2+\cdots\cdots+m_nx_n \\ \qquad\text{（稱“加權算術平均”）} \\ g=x_1^{m_1}\cdot x_2^{m_2}\cdot\cdots\cdots\cdot x_n^{m_n} \\ \qquad\text{（稱“加權幾何年均”）} \\ \dfrac{1}{h}=\dfrac{m_1}{x_1}+\dfrac{m_2}{x_2}+\cdots\cdots+\dfrac{m_n}{x_n} \\ \qquad\text{（稱“加權調和平均”）} \end{cases}$$

則

$$h\leqslant g\leqslant \mu$$

解：設 $f(x)=-\log x$

爲定義在 $I=(0,\infty)$ 上的函數

因爲 $$f''(x)=\frac{1}{x^2}\geqslant 0$$

所以

$$\log(m_1x_1+m_2x_2+\cdots+m_nx_n)$$
$$\geqslant m_1\log x_1+m_2\log x_2+\cdots+m_n\log x_n$$

又因 log 爲遞增函數故

$$m_1x_1+m_2x_2+\cdots\cdots+m_nx_n$$
$$\geqslant x_1^{m_1}\cdot x_2^{m_2}\cdot\cdots\cdot x_n^{m_n} \quad\cdots\cdots\cdots\cdots\cdots(1)$$

即 $$\mu\geqslant g$$

同時，考慮定義在 $I=(0,\infty)$ 上的函數

$$f(x)=\frac{1}{x}$$

因爲 $$f''(x)=\frac{2}{x^3}\geqslant 0$$

所以

$$\frac{1}{m_1x_1+m_2x_2+\cdots\cdots+m_nx_n}$$
$$\leqslant m_1\left(\frac{1}{x_1}\right)+m_2\left(\frac{1}{x^2}\right)+\cdots\cdots$$
$$+m_n\left(\frac{1}{x_n}\right)$$

再分別以 $\frac{1}{x_1}, \frac{1}{x_2}, \cdots\cdots, \frac{1}{x_n}$ 代入(1)式中 $x_1, x_2, \cdots, x_n$ 可得下列不等式。

$$m_1\left(\frac{1}{x_1}\right)+m_2\left(\frac{1}{x_2}\right)+\cdots+m_n\left(\frac{1}{x_n}\right)$$
$$\geqslant\frac{1}{x_1^{m_1}\cdot x_2^{m_2}\cdot\cdots\cdot x_n^{m_n}}$$

故得

$$h\leqslant g\leqslant\mu$$

註說：若諸 $m_i=\frac{1}{n}$，則 μ，g，h 分別稱爲諸正數 x_i 的算術平均數、幾何平均數及調和平均數。

論例 2 設 $x_1, x_2, \cdots, x_n$ 爲 n 個正銳角，則下列不等式成立：

$$\sin x_1\cdot\sin x_2\cdots\cdots\sin x_n$$
$$\leqslant\sin^n\left(\frac{x_1+x_2+\cdots+x_n}{n}\right)$$

解：選取兩個定義在區間 $I=\left(0,\frac{\pi}{2}\right)$ 上的函數：

$$\begin{cases} f(x)=\dfrac{1}{\sin x} \\ g(x)=-\log x \end{cases}$$

因 $$f''(x)=\frac{\cos x}{\sin^2 x}\geqslant 0, \ \forall x\in I$$

且 $$g''(x)=\frac{1}{x^2}>0$$

故 $(g\circ f)(x)=-\log\sin x$

又 $g(x)$ 在區間 I 上為一遞增函數，由定理 5-2 知

$$y=-\log\sin x$$

為一凸函數。

則

$$-\sin\left(\frac{1}{n}x_1+\frac{1}{n}x_2+\cdots+\frac{1}{n}x_n\right)$$

$$\leqslant-(\sin x_1\cdot\sin x_2\cdots\cdots\sin x_n)^{\frac{1}{n}}$$

或

$$\sin x_1\cdot\sin x_2\cdots\cdots\sin x_n$$

$$\leqslant\sin^n\left(\frac{x_1+x_2+\cdots+x_n}{n}\right)$$

論例3 設 $x_1,x_2>1$ 則下列不等式成立

$$\log\frac{x_1+x_2}{2}\geqslant\sqrt{\log x_1\cdot\log x_2}$$

解：考慮定義在 $I=(1,\infty)$ 上的函數

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

因為 $f(x)$ 的一階導微函數 $f'(x)=\dfrac{-1}{x\log x}$

及 $f(x)$ 的二階導函數 $f''(x)$

$$f''(x)=\frac{1+\log x}{x^2(\log x)^2}>0$$

所以 $f(x)=-\log\log x$ 為定義在 $(1,\infty)$ 上的凸函數。故得

$$-\log\log\left(\frac{x_1+x_2}{2}\right)$$

$$\leqslant-\frac{\log(\log x_1\cdot\log x_2)}{2}$$

或

$$\log\log\left(\frac{x_1+x_2}{2}\right)$$

$$\geqslant\frac{\log(\log x_1\cdot\log x_2)}{2}$$

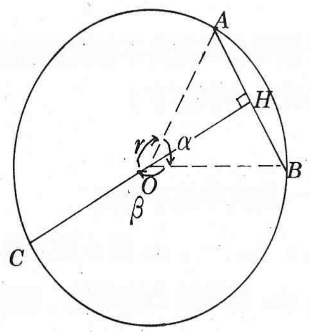
又因為在本區間 $I=(1,\infty)$ 上，log 為遞增函數所以得：

$$\log\left(\frac{x_1+x_2}{2}\right)\geqslant\sqrt{\log x_1\cdot\log x_2}$$

論例4 極值與凸函數

試證明：在所有圓內接三角形中周界最長者（或面積最大者）為正三角形。

解：設三角形 ΔABC 的外接圓的圓心為 O，半徑為 R（見下圖）



如上圖示，令 $CH\perp AB$ 於 H，且 $\angle AOB=\alpha$，$\angle BOC=\beta$，$\angle COA=r$，則得

$$\begin{cases}\overline{AB}=2R\sin\dfrac{\alpha}{2}\\[2ex]\overline{BC}=2R\sin\dfrac{\beta}{2}\qquad\text{其中}\alpha+\beta+r=2\pi\\[2ex]\overline{CR}=2R\sin\dfrac{r}{2}\end{cases}$$

因為

$$\sin\frac{\alpha}{2}=\sin\left(\frac{2\pi-\alpha-\beta}{2}\right)$$

$$=\sin\frac{\alpha+\beta}{2}$$

故 ΔABC 的周長為：

$$p=2R\left[\sin\frac{\alpha}{2}+\sin\frac{\beta}{2}+\sin\frac{\alpha+\beta}{2}\right]$$

設函數 $f(x)=\sin\dfrac{x}{2}+2\sin\dfrac{x}{4}$

則在區間 $I=[0,2\pi]$ 內易知 $f(x)$ 為一凸函數，故得下列不等式

$$m_1=m_2=\frac{1}{2};\quad x_1=\alpha,\ x_2=\beta$$

$$\frac{1}{2}\left(\sin\frac{\alpha}{2}+\sin\frac{\beta}{2}\right)\leqslant\sin\frac{\alpha+\beta}{4}$$

但

$$f(\alpha+\beta) = \sin\frac{\alpha+\beta}{2} + 2\sin\frac{\alpha+\beta}{4}$$

所以

$$p = 2R \cdot f(\alpha+\beta)$$

再因爲

$$f'(x) = \frac{1}{2}\left(\cos\frac{x}{2} + \cos\frac{x}{4}\right) = \cos\frac{x}{8} \cdot \cos\frac{3x}{8}$$

而且

當 $x = \frac{4}{3}\pi$ 時，$f'(x) = 0$

當 $0 \leqslant x < \frac{4\pi}{3}$ 時，$f'(x) > 0$

當 $\frac{4\pi}{3} \leqslant x < 2\pi$ 時，$f'(x) \leqslant 0$

我們可以列變化表如：

x	0		$\frac{4\pi}{3}$		2π
f'	1	$+$	0	$-$	
f	0	$\nearrow$	$\frac{3\sqrt{3}}{2}$	$\searrow$	2

因此當 $x = \frac{4\pi}{3}$ 時，取得極大值爲 $\frac{3\sqrt{3}}{2}$ 所以

$$f(\alpha+\beta) \leqslant \frac{3\sqrt{3}}{2}$$

故

$$p \leqslant 3R\sqrt{3}$$

上述不等式表示說：當 $\alpha+\beta = \frac{4\pi}{3}$ 且 $\alpha = \beta$ 時，所有圓內接三角形中周界最長者爲正三角形。

參考資料

1. 王中烈：「從半道高中數學競試題談起——兼淺談不等式與凸函數的關係」數學傳播第六卷第二期，71 年，頁 21 ～ 29 。
2. 王中烈：「你相信你的眼睛嗎」數學傳播第六卷第一期，71 年，頁 17 ～ 32 。
3. 夏文侯：「凸函數之初等證明法」數學傳播第四卷第三期，69 年，頁 35 ～ 38 。
4. 何景國：「高中數學科教材研究」數學傳播第三卷第二期，67 年，頁 116 ～ 120 。
5. A. WAYNE ROBERTS/DALE E. VARBERG :「CONVEX FUNCTIONS」 協進圖書 66年。
6. J. BASS :「Cours de mathematiques」Masson & Cie Paris 1969.

（本文作者現任敎於延平中學）