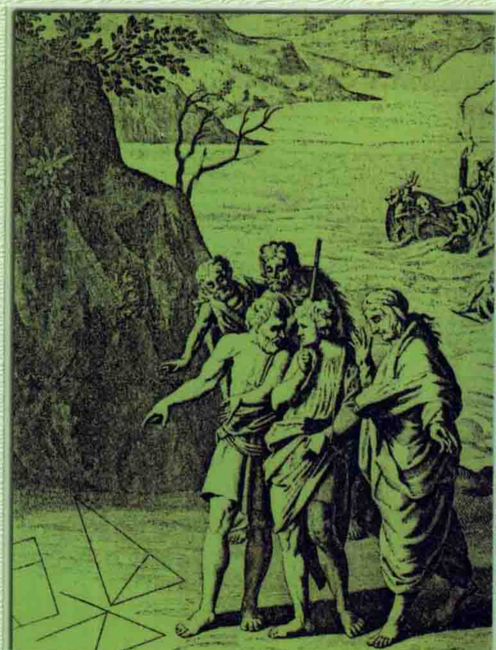
《数学中的小问题大定理》丛书（第四辑）

凸函数最值定理

——从一道华约自主招生题的解法谈起

佩捷 编著



◎凸函数的连续性和可微性

◎特殊类的凸函数

◎凸函数与凸规划

◎凸函数和凸映射

◎线性约束凸规划的

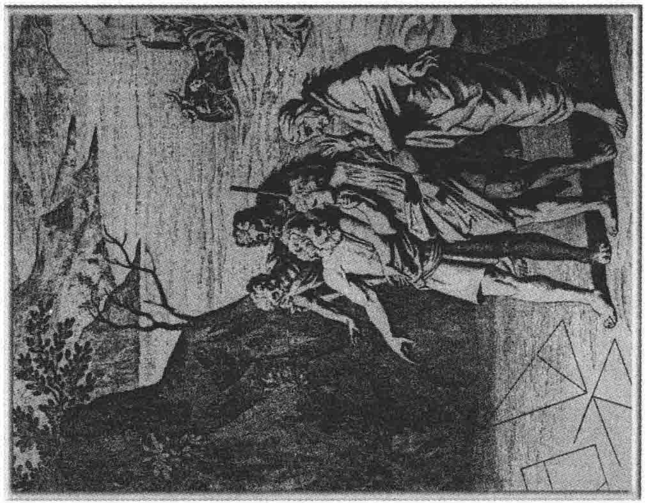
HITP
哈爾濱工業大學出版社
HARBIN INSTITUTE OF TECHNOLOGY PRESS

《数学中的小问题大定理》丛书（第四辑）

凸函数最值定理

——从一道华约自主招生题的解法谈起

佩捷 编著



◎ 凸函数的连续性和可微性

◎ 特殊类的凸函数

◎ 凸函数与凸规划

◎ 凸函数和凸映射

◎ 线性约束凸规划的既约变尺度法

哈爾濱工業大學出版社
HARBIN INSTITUTE OF TECHNOLOGY PRESS

内容简介

本书从一道华约自主招生题解法中所应用的凸函数最值定理谈起，详细地介绍了凸函数及凸函数的众多性质.

本书适合广大数学爱好者阅读参考.

图书在版编目(CIP)数据

凸函数最值定理：从一道华约自主招生题的解法谈起／佩捷编著. —哈尔滨：哈尔滨工业大学出版社，2014.10

ISBN 978 - 7 - 5603 - 4948 - 0

Ⅰ.①凸… Ⅱ.①佩… Ⅲ.①凸函数 - 研究 Ⅳ.①O174.13

中国版本图书馆 CIP 数据核字(2014)第 223284 号

策划编辑　刘培杰　张永芹
责任编辑　张永芹　刘春雷
封面设计　孙茵艾
出版发行　哈尔滨工业大学出版社
社　　址　哈尔滨市南岗区复华四道街 10 号　邮编 150006
传　　真　0451 - 86414749
网　　址　http://hitpress.hit.edu.cn
印　　刷　哈尔滨工业大学印刷厂
开　　本　787mm×960mm　1/16　印张 17.25　字数 175 千字
版　　次　2014 年 10 月第 1 版　2014 年 10 月第 1 次印刷
书　　号　ISBN 978 - 7 - 5603 - 4948 - 0
定　　价　28.00 元

◎ 目录

引　言

第〇章

一个闭区间内取值的凸函数最值定理的两个应用

定理　$f(x)$为定义在$[a,b]\in\mathbf{R}$上的实的凸函数,实数$x_1,x_2,\cdots,x_n\in[a,b]$,且满足

$$x_1+x_2+\cdots+x_n=s\quad(na\leqslant s\leqslant nb)$$

下列表达式

$$F=f(x_1)+f(x_2)+\cdots+f(x_n)$$

F达到最大值,当且仅当数组$(x_1,x_2,\cdots,x_n)$中至少有$n-1$个元素等于a或者b.

证明　这个定理直接从$n=2$的情况可以得到. 事实上,只需要证明:如果$x,y\in[a,b]$且$2a\leqslant x+y\leqslant 2b$,则

$$f(x)+f(y)\leqslant\begin{cases}f(a)+f(s-a),s\leqslant a+b\\f(b)+f(s-b),s\geqslant a+b\end{cases}$$

实际上,假设$s\leqslant a+b$,则$s-a\leqslant b$. 由于$x\in[a,s-a]$,则存在数$t\in[0,1]$使得$x=ta+(1-t)(s-a)$,$y=(1-t)a+t(s-a)$. 由凸函数的定义我们有$f(x)\leqslant tf(a)+$

$(1-t)f(s-a)$以及$f(y)\leqslant(1-t)f(a)+tf(s-a)$，两式相加，我们得到

$$f(x)+f(y)\leqslant f(a)+f(s-a)$$

在$s\geqslant a+b$的情况下，定理类似可证.

下面我们应用这个定理解决几个自主招生和中国数学奥林匹克中的试题：

例1 （2012年"华约"自主招生数学选择压轴题）已知

$$-6\leqslant x_i\leqslant 10(i=1,2,\cdots,10),\ \sum_{i=1}^{10} x_i=50$$

当$\sum\limits_{i=1}^{10} x_i^2$取得最大值时，在$x_1,x_2,\cdots,x_{10}$这10个数中等于$-6$的数共有（　　）.

A. 1个　　B. 2个　　C. 3个　　D. 4个

解 由于$f(x)=x^2$是下凸函数，从而由定理可知$\sum\limits_{i=1}^{10} x_i^2$取得最大值时，$x_1,x_2,\cdots,x_{10}$中至少有9个等于$-6$或10.

设其中有m个-6，$9-m$个10，则余下一个为$50-(-6m+90-10m)=16m-40$，注意到$-6\leqslant 16m-40\leqslant 10$，即$2.875\leqslant m\leqslant 3.125$，故整数$m$只能取3，此时$16m-40=8\neq -6$，从而本题选C.

例2 （1997年中国数学奥林匹克）设实数$x_1,x_2,\cdots,x_{1\,997}$满足如下两个条件：

(1) $-\frac{1}{\sqrt{3}}\leqslant x_i\leqslant\sqrt{3}(i=1,2,\cdots,1\,997)$；

(2) $x_1+x_2+\cdots+x_{1\,997}=-318\sqrt{3}$.

试求：$x_1^{12}+x_2^{12}+\cdots+x_{1\,997}^{12}$的最大值，并说明理由.

解　由于$f(x)=x^{12}$是下凸函数，从而$\sum\limits_{i=1}^{1\,997}x_i^{12}$取得最大值时，$x_1,x_2,\cdots,x_{1\,997}$中至少有1 996个等于$-\frac{1}{\sqrt{3}}$或$\sqrt{3}$.

设其中有t个$-\frac{1}{\sqrt{3}}$，1 996 $-t$个$\sqrt{3}$，则余下一个为

$$-318\sqrt{3}-\left|-\frac{1}{\sqrt{3}}t+(1\,996-t)\sqrt{3}\right|=\frac{4}{\sqrt{3}}t-2\,314\sqrt{3}$$

由已知$-\frac{1}{\sqrt{3}}\leqslant\frac{4}{\sqrt{3}}t-2\,314\sqrt{3}\leqslant\sqrt{3}$，解之得

$$1\,735.25\leqslant t\leqslant 1\,736.25$$

注意到$t\in\mathbf{N}$，故$t=1\,736$.

进一步$1\,996-t=260$.

所以由定理可知：当$x_1,x_2,\cdots,x_{1\,997}$中有1 736个取$-\frac{1}{\sqrt{3}}$，260个取$\sqrt{3}$，一个取$\frac{2\sqrt{3}}{3}$时，$\sum\limits_{i=1}^{1\,997}x_i^{12}$取最大值.其最大值为

$$\left(-\frac{1}{\sqrt{3}}\right)^{12}\times 1\,736+(\sqrt{3})^{12}\times 260+\left(\frac{2\sqrt{3}}{3}\right)^{12}=189\,548$$

例3　(1993年中国数学奥林匹克(第八届数学冬令营))给定$k\in\mathbf{N}$及实数$a>0$，在条件

$$k_1+k_2+\cdots+k_r=k\quad(k_1\in\mathbf{N},1\leqslant r\leqslant k)$$

下，求

$$a^{k_1}+a^{k_2}+\cdots+a^{k_r}$$

的最大值.

解　由定理可知，当且仅当$k_1,k_2,\cdots,k_r$中有$r-1$个1，一个为$k-r+1$时取最大值，最大值为

凸函数最值定理

$$(r-1)a+a^{k-r+1}$$

参考文献

庄燕文. 最新竞赛试题选编及解析[M]. 北京:首都师范大学出版社,2011:19,138-139.

第一章 什么是凸函数

§1 Jensen 凸函数的定义

在本节中,I 表示区间(a,b),$\bar{I}$表示线段$[a,b]$,f 表示定义在 I 或$\bar{I}$上的实函数.

定义 1.1.1 函数f 称为 Jensen 意义下的凸函数,或称为 J 凸函数,如果对任意两点 $x,y\in\bar{I}$,f 满足不等式

$$f\left(\frac{x+y}{2}\right)\leqslant\frac{f(x)+f(y)}{2}\quad(1.1.1)$$

定义 1.1.2 一个在 I 上的 J 凸函数f 称为严格 J 凸函数,如果对每一对点 x,$y\in\bar{I}$,$x\neq y$,不等式(1.1.1)中的严格不等式成立.

定义 1.1.3 函数f 称为$\bar{I}$ 上的 J 凹函数(严格 J 凹函数),如果函数 $x\mapsto -f(x)$ 是$\bar{I}$ 上的 J 凸函数(严格 J 凸函数).

注 有的文献中,满足定义 1.1.1 的条件的函数称为 J 非凹函数,定义为 J 凹的那些函数称为 J 非凸函数,因此严格 J 凸和严格 J 凹函数分别称为 J 凸和 J 凹函数.

与 J 凸函数类似,我们定义凸序列如下:

定义 1.1.4 实数序列 $a=(a_1,\cdots,a_n)$ 称为凸序列,如果对 $k=1,\cdots,n-1$,有

$$2a_k\leqslant a_{k-1}+a_{k+1}$$

当序列是无限时,也可用类似的定义.

定理 1.1.1 设 f 是 $\bar{I}'$ 上的 J 凸函数,g 是 $\bar{I}''$ 上的 J 凸函数,并设 $\bar{I}=\bar{I}'\cap\bar{I}''$ 至少有两点,则

(1) $x\mapsto\max(f(x),g(x))$ 是 $\bar{I}$ 上的 J 凸函数.

(2) $x\mapsto h(x)=f(x)+g(x)$ 是 $\bar{I}$ 上的 J 凸函数.

(3) 假定 f,g 都是 $\bar{I}$ 上的正的非减函数,则 $x\mapsto f(x)\cdot g(x)$ 是 $\bar{I}$ 上的 J 凸函数.

(4) 假定 g 是 $\bar{I}''$ 上的非减函数,$[f(a),f(b)]\subset\bar{I}''$,则 $x\mapsto g(f(x))$ 是 $\bar{I}'$ 上的 J 凸函数.

证明 对于实数 x,y,u,v,不等式

$$\max(x+y,u+v)\leqslant\max(x,u)+\max(y,v)$$

成立,由此得出结论(1).

对 $x,y\in\bar{I}$,我们有

$$\begin{aligned}h\left(\frac{x+y}{2}\right)&=f\left(\frac{x+y}{2}\right)+g\left(\frac{x+y}{2}\right)\\&\leqslant\frac{f(x)+f(y)}{2}+\frac{g(x)+g(y)}{2}\\&=\frac{h(x)+h(y)}{2}\end{aligned}$$

这就证明了(2).

f,g 在 $\bar{I}$ 上非减的假设蕴涵着

$$(f(x)-f(y))(g(y)-g(x))\leqslant 0\quad(x,y\in\bar{I})$$

即

$$f(x)g(y)+f(y)g(x)\leqslant f(x)g(x)+f(y)g(y) \tag{1.1.2}$$

若我们把下列不等式相乘

$$f\left(\frac{x+y}{2}\right)\leqslant\frac{f(x)+f(y)}{2},g\left(\frac{x+y}{2}\right)\leqslant\frac{g(x)+g(y)}{2}$$

根据假设,其中 f,g 是正的,则应用(2),我们得到所要求的结论(3).

因为 g 在 $\bar{I}''$ 上非减,由

$$f\left(\frac{x+y}{2}\right)\leqslant\frac{f(x)+f(y)}{2}\quad (x,y\in\bar{I}')$$

我们得到

$$\begin{aligned}h\left(\frac{x+y}{2}\right)&=g\left(f\left(\frac{x+y}{2}\right)\right)\leqslant g\left(\frac{f(x)+f(y)}{2}\right)\\&\leqslant\frac{g(f(x))+g(f(y))}{2}=\frac{h(x)+h(y)}{2}\end{aligned}$$

这就证明了(4).

下面,我们证明 J 凸函数的一个有趣的不等式,在某种意义上,J 凸函数是最好的一类凸函数.

定理 1.1.2　假定 f 是 $\bar{I}$ 上的 J 凸函数,对于任意的点 $x_1,\cdots,x_n\in\bar{I}$ 以及任意非负有理数 $r_1,\cdots,r_n$(满足 $r_1+\cdots+r_n=1$),我们有

$$f\left(\sum_{i=1}^{n}r_ix_i\right)\leqslant\sum_{i=1}^{n}r_if(x_i) \tag{1.1.3}$$

证明　**情形** 1　$r_i=\dfrac{1}{n}(i=1,\cdots,n)$. 此时式(1.1.3)即变成

$$f\left(\frac{1}{n}\sum_{i=1}^{n}x_i\right)\leqslant\frac{1}{n}\sum_{i=1}^{n}f(x_i) \tag{1.1.4}$$

式(1.1.4)的证明是用对 n 的归纳法. 当 $n=2$ 时,因为它与定义 1.1.1 一致,所以式(1.1.4)成立. 假定对于$n=2^k$(k 是自然数),定理 1.1.2 成立. 对于 $x_1,\cdots,x_m\in\bar{I}$ 和 $m=2^{k+1}=2\cdot 2^k=2n$,我们有

$$f\left(\frac{x_1+\cdots+x_m}{m}\right)=f\left(\frac{\frac{1}{n}\sum_{m=1}^{n}x_m+\frac{1}{n}\sum_{m=1}^{n}x_{m+n}}{2}\right)$$

$$\leqslant\frac{f\left(\frac{1}{n}\sum_{m=1}^{n}x_m\right)+f\left(\frac{1}{n}\sum_{m=1}^{n}x_{m+n}\right)}{2}$$

$$\leqslant\frac{\sum_{m=1}^{n}f(x_m)+\sum_{m=1}^{n}f(x_{m+n})}{2n}=\frac{\sum_{m=1}^{2n}f(x_m)}{2n}$$

因此对于每个自然数 $n\in\{2,2^2,2^3,\cdots\}$,式(1.1.4)成立.

若我们证明对于 $n(n>2)$,式(1.1.4)成立,则对于$n-1$,式(1.1.4)也成立. 设 $x_1,\cdots,x_{n-1}\in\bar{I}$. 对于数 $x_1,\cdots,x_{n-1}$和 $x_n=\frac{1}{n-1}(x_1+\cdots+x_{n-1})$. 式(1.1.4)成立,即

$$f\left(\frac{x_1+\cdots+x_{n-1}+\frac{x_1+\cdots+x_{n-1}}{n-1}}{n}\right)$$

$$\leqslant\frac{f(x_1)+\cdots+f(x_{n-1})+f\left(\frac{x_1+\cdots+x_{n-1}}{n-1}\right)}{n}\qquad(1.1.5)$$

这个不等式的左边经过整理变为

$$f\left(\frac{x_1+\cdots+x_{n-1}}{n-1}\right)$$

所以式(1.1.5)变为

$$f\left(\frac{x_1+\cdots+x_{n-1}}{n-1}\right)\leqslant\frac{1}{n}(f(x_1)+\cdots+f(x_{n-1}))+\frac{1}{n}f\left(\frac{x_1+\cdots+x_{n-1}}{n-1}\right)$$

由此得

$$f\left(\frac{x_1+\cdots+x_{n-1}}{n-1}\right)\leqslant\frac{f(x_1)+\cdots+f(x_{n-1})}{n-1}$$

因此,若对 $n(n>2)$,式(1.1.4)成立,则对 $n-1$,式(1.1.4)也成立.

这就完成了定理1.1.2在上述情形的证明.

情形2　因为 $r_1,\cdots,r_n$ 是非负的有理实数,因此存在一个自然数 m 和非负整数 $p_1,\cdots,p_n$,使 $m=p_1+\cdots+p_n$ 且 $r_i=\frac{p_i}{m}(i=1,\cdots,n)$,现在,根据情形1,我们有

$$f\left(\frac{(x_1+\cdots+x_1)+\cdots+(x_n+\cdots+x_n)}{m}\right)\leqslant\frac{(f(x_1)+\cdots+f(x_1))+\cdots+(f(x_n)+\cdots+f(x_n))}{m}\tag{1.1.6}$$

其中第一个括号里有 p_1 项,……,第 n 个括号里有 p_n 项,于是式(1.1.6)变成

$$f\left(\frac{1}{m}\sum_{i=1}^{n}p_ix_i\right)\leqslant\sum_{i=1}^{n}\frac{p_i}{m}f(x_i)$$

注　假定 n 维欧氏空间的子集 $\bar{I}$ 有中点性质,即由 $x,y\in\bar{I}$ 可推出 $\frac{x+y}{2}\in\bar{I}$,则 J 凸函数的定义,定理1.1.1和定理1.1.2对于在 $\bar{I}$ 上定义的任意实值函数 f

也适用.

J. L. W. V. Jensen(见文献[1]和[2])是第一个用不等式(1.1.1)定义凸函数并对它们的重要性引起注意的人,他还证明了定理1.1.2. 我们在这里引用他的已完全被证明是正确的话:"我觉得凸函数的概念和正函数,增函数一样也是基本的. 如果这一点我没有弄错的话,这个概念应当在初等的实变函数理论陈述中占有自己的位置"(见文献[2]的p.191).

但是,甚至在Jensen以前,已有关于凸函数的结果,例如,在1889年,O. Hölder[3]证明了在f(在I上)两次可微,$f''(x)\geqslant 0$(即f是凸的,虽然在文章中没有明确说明)的条件下,不等式(1.1.3)成立. 让我们还提一下1893年由O. Stolz[4]提出的关于满足不等式(1.1.1)的连续函数的左、右导数存在的结果. 这个结果将在§4的定理1.4.1中正式叙述,我们同意T. Popoviciu(见文献[5]的p.48)的说法:看来是Stolz首先通过证明刚才提到的结果而引入凸函数的,1893年,Hadamard[6]的结果也是属于比Jensen的论文发表更早的时期,这个结果是:若函数f是可微的,且它的导数是$\bar{I}$上的递增函数,则对于所有的$x_1,x_2\in I(x_1\neq x_2)$,有

$$f\left(\frac{x_1+x_2}{2}\right)<\frac{1}{x_2-x_1}\int_{x_1}^{x_2}f(x)\,\mathrm{d}x$$

(见文献[7]的p.441).

§2 Jensen凸函数的连续性

若f是$\bar{I}$上的J凸函数,则根据§1中的定理1.1.2,

对所有 $x,y\in\bar{I}$ 和所有有理数 $\lambda\in[0,1]$，不等式

$$f(\lambda x+(1-\lambda)y)\leqslant\lambda f(x)+(1-\lambda)f(y) \tag{1.2.1}$$

成立. 因此，若 f 是连续的，则对所有实数 $\lambda\in[0,1]$，式(1.2.1)成立.

但是，在 $\bar{I}$ 上的 J 凸函数不一定在 $\bar{I}$ 上连续. 我们给出一个这样的函数的例子(见文献[7]). 设 f 是 Cauchy 函数方程

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

的任意不连续解(G. Hamel[8]证明了这种解的存在性；这个函数方程的通解见 J. Aczél[9]). 由 $g(x)=\max(x^2,f(x)+x^2)$ 定义的函数是 J 凸且不连续的.

此外，由

$$f(x)=x\quad(-1<x<1)$$
$$f(-1)=f(1)=2$$

定义的函数在 $[-1,1]$ 上是 J 凸的，在 $(-1,1)$ 上连续，但在区间的端点不连续.

有许多结果保证 J 凸函数在各种条件下的连续性，我们只介绍其中一些，最重要的一个结果是由不等式(1.2.1)得到的，即(见 Jensen[1] 以及 F. Bernstein 和 G. Doetsch[10])：

定理 1.2.1　若一个 J 凸函数在 I 上有定义并且上有界，则它在 I 上连续.

W. Sierpiński[11]推广了这个结果. 他还证明了以一个可测函数(在 Lebesgue 意义下)为上界的 J 凸函数 f 是连续的. A. Ostrowski[12]证明了在一个正测度的集合上有界的 J 凸函数也是连续的，从而推广了那个结果. S. Kurepa[13]研究了若 J 凸函数 f 在 T 上有界，则