

《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



哈尔滨工业大学  
HARBIN INSTITUTE OF TECHNOLOGY

《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



哈尔滨工业大学出版社  
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

---

(如因印装质量问题影响阅读,我社负责调换)

◎  
目  
录

第〇章 引言 //1

一个闭区间内取值的凸函数最值定理的两个应用 //1

参考文献 //4

第一章 什么是凸函数 //5

§1 Jensen 凸函数的定义 //5

§2 Jensen 凸函数的连续性 //10

§3 凸函数 //12

§4 凸函数的连续性和可微性 //15

§5 对数性凸函数 //17

§6 凸函数概念的一些推广 //19

§7 凸性的谱系 //22

参考文献 //24

第二章 特殊类的凸函数 //31

§1  $N$ -函数 //31

§2 余  $N$ -函数 //43

§3  $N$ -函数的比较 //47

§4  $\Delta_2$ -条件 //57

§5  $\Delta'$ -条件 //65

§6 较幂函数增加得快的  $N$ -函数 //72

§7 关于一类  $N$ -函数 //93

第三章  $p$ -凸函数与几类不等式 //101

§1 引言 //101

§ 2	$p$ -凸函数的性质与判别准则	//101
§ 3	$p$ -凸函数的几类不等式	//105
	参考文献	//110
<b>第四章</b>	<b>凸函数与凸规划</b>	//112
§ 1	单变量凸函数	//112
§ 2	线性空间上的凸函数	//121
§ 3	次线性函数和 Minkowski 函数	//130
<b>第五章</b>	<b>极小问题和变分不等式:凸性、单调性和不动点</b>	//137
§ 1	直接形式	//139
§ 2	弱形式	//140
§ 3	线性化形式	//142
§ 4	不动点形式	//142
§ 5	上图形式	//145
§ 6	赋范空间中的极小问题	//146
§ 7	单调算子和变分不等式:线性化引理	//150
§ 8	变分不等式和不动点	//153
§ 9	不可微泛函的极小化和混合变分不等式	//158
<b>第六章</b>	<b>HILBERT 空间凸规划最优解的可移性</b>	//162
§ 1	最优解与平稳点的关系	//162
§ 2	不动点与问题 $P$ 的关系	//170
§ 3	最优解与鞍点	//172
	参考文献	//176
<b>第七章</b>	<b>凸函数和凸映射</b>	//178
§ 1	凸函数及有关性质	//178
§ 2	凸函数的连续性	//196
<b>第八章</b>	<b>线性约束凸规划的既约变尺度法</b>	//198
§ 1	引言	//198
§ 2	问题、假设及记号	//199
§ 3	既约变尺度法	//203

§ 4	既约变尺度法的收敛性	//206
	参考文献	//213
附录 I	赋范空间中凸泛函 Lipschitz 连续性与函数有下界的关系	//214
附录 II	凸函数的一些新性质	//220
附录 III	多元函数凹凸性的定义和判别法	//230
附录 IV	关于 $(\alpha, m)$ - 预不变凸函数的 Ostrowski 型不等式	//240
	编辑手记	//249



# 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 $x_1, x_2, \dots, x_n \in [a, b]$ ,且满足

$$x_1 + x_2 + \dots + x_n = s \quad (na \leq s \leq nb)$$

下列表达式

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

$F$ 达到最大值,当且仅当数组 $(x_1, x_2, \dots, x_n)$ 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 $2a \leq x+y \leq 2b$ ,则

$$f(x) + f(y) \leq \begin{cases} f(a) + f(s-a), & s \leq a+b \\ f(b) + f(s-b), & s \geq a+b \end{cases}$$

实际上,假设 $s \leq a+b$ ,则 $s-a \leq b$ .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 $f(x) \leq tf(a) +$

### 凸函数最值定理

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

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

在 $s \geq a+b$ 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

$$-6 \leq x_i \leq 10 (i=1, 2, \dots, 10), \quad \sum_{i=1}^{10} x_i = 50$$

当 $\sum_{i=1}^{10} x_i^2$ 取得最大值时, 在 $x_1, x_2, \dots, x_{10}$ 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

$\sum_{i=1}^{10} x_i^2$ 取得最大值时,  $x_1, x_2, \dots, x_{10}$ 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  $-6 \leq 16m - 40 \leq 10$ , 即  $2.875 \leq m \leq 3.125$ , 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  $x_1, x_2, \dots, x_{1997}$  满足如下两个条件:

$$(1) -\frac{1}{\sqrt{3}} \leq x_i \leq \sqrt{3} (i=1, 2, \dots, 1997);$$

$$(2) x_1 + x_2 + \dots + x_{1997} = -318\sqrt{3}.$$

试求:  $x_1^{12} + x_2^{12} + \dots + x_{1997}^{12}$  的最大值, 并说明理由.



解 由于  $f(x) = x^{12}$  是下凸函数, 从而  $\sum_{i=1}^{1997} x_i^{12}$  取得最大值时,  $x_1, x_2, \dots, x_{1997}$  中至少有 1996 个等于  $-\frac{1}{\sqrt{3}}$  或  $\sqrt{3}$ .

设其中有  $t$  个  $-\frac{1}{\sqrt{3}}$ ,  $1996-t$  个  $\sqrt{3}$ , 则余下一个为

$$-318\sqrt{3} - \left| -\frac{1}{\sqrt{3}}t + (1996-t)\sqrt{3} \right| = \frac{4}{\sqrt{3}}t - 2314\sqrt{3}$$

由已知  $-\frac{1}{\sqrt{3}} \leq \frac{4}{\sqrt{3}}t - 2314\sqrt{3} \leq \sqrt{3}$ , 解之得

$$1735.25 \leq t \leq 1736.25$$

注意到  $t \in \mathbf{N}$ , 故  $t = 1736$ .

进一步  $1996 - t = 260$ .

所以由定理可知: 当  $x_1, x_2, \dots, x_{1997}$  中有 1736 个取  $-\frac{1}{\sqrt{3}}$ , 260 个取  $\sqrt{3}$ , 一个取  $\frac{2\sqrt{3}}{3}$  时,  $\sum_{i=1}^{1997} x_i^{12}$  取最大值.

其最大值为

$$\left(-\frac{1}{\sqrt{3}}\right)^{12} \times 1736 + (\sqrt{3})^{12} \times 260 + \left(\frac{2\sqrt{3}}{3}\right)^{12} = 189548$$

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

$$k_1 + k_2 + \dots + k_r = k \quad (k_i \in \mathbf{N}, 1 \leq r \leq k)$$

下, 求

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

的最大值.

解 由定理可知, 当且仅当  $k_1, k_2, \dots, k_r$  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.



# 什么是凸函数

## 第

## 一

## 章

### § 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

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

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.

### 凸函数最值定理

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  $a = (a_1, \dots, a_n)$  称为凸序列,如果对  $k = 1, \dots, n-1$ , 有

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

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

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

成立, 由此得出结论(1).

对  $x, y \in \bar{I}$ , 我们有

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

这就证明了(2).

$f, g$  在  $\bar{I}$  上非减的假设蕴涵着

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

即

$$f(x)g(y) + f(y)g(x) \leq f(x)g(x) + f(y)g(y) \quad (1.1.2)$$

若我们把下列不等式相乘

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

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

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

我们得到

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

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 $x_1, \dots, x_n \in \bar{I}$ 以及任意非负有理数 $r_1, \dots, r_n$ (满足 $r_1 + \dots + r_n = 1$ ),我们有

$$f\left(\sum_{i=1}^n r_i x_i\right) \leq \sum_{i=1}^n r_i f(x_i) \quad (1.1.3)$$

**证明 情形 1**  $r_i = \frac{1}{n} (i=1, \dots, n)$ . 此时式(1.1.

3)即变成

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

### 凸函数最值定理

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  $x_1, \dots, x_m \in \bar{I}$  和  $m=2^{k+1}=2 \cdot 2^k=2n$ , 我们有

$$\begin{aligned} f\left(\frac{x_1 + \dots + x_m}{m}\right) &= f\left(\frac{\frac{1}{n} \sum_{m=1}^n x_m + \frac{1}{n} \sum_{m=1}^n x_{m+n}}{2}\right) \\ &\leq \frac{f\left(\frac{1}{n} \sum_{m=1}^n x_m\right) + f\left(\frac{1}{n} \sum_{m=1}^n x_{m+n}\right)}{2} \\ &\leq \frac{\sum_{m=1}^n f(x_m) + \sum_{m=1}^n f(x_{m+n})}{2n} = \frac{\sum_{m=1}^{2n} f(x_m)}{2n} \end{aligned}$$

因此对于每个自然数  $n \in \{2, 2^2, 2^3, \dots\}$ , 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  $x_1, \dots, x_{n-1} \in \bar{I}$ . 对于数  $x_1, \dots, x_{n-1}$  和  $x_n = \frac{1}{n-1}(x_1 + \dots + x_{n-1})$ . 式(1.1.4)成立, 即

$$\begin{aligned} & f\left(\frac{x_1 + \dots + x_{n-1} + \frac{x_1 + \dots + x_{n-1}}{n-1}}{n}\right) \\ & \leq \frac{f(x_1) + \dots + f(x_{n-1}) + f\left(\frac{x_1 + \dots + x_{n-1}}{n-1}\right)}{n} \quad (1.1.5) \end{aligned}$$

这个不等式的左边经过整理变为

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

所以式(1.1.5)变为

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

由此得

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

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

$$\begin{aligned} & f\left(\frac{(x_1 + \cdots + x_1) + \cdots + (x_n + \cdots + x_n)}{m}\right) \\ & \leq \frac{(f(x_1) + \cdots + f(x_1)) + \cdots + (f(x_n) + \cdots + f(x_n))}{m} \end{aligned} \quad (1.1.6)$$

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

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

**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  $f$

## 凸函数最值定理

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<sup>[3]</sup>证明了在  $f$ (在  $I$  上)两次可微,  $f''(x) \geq 0$  (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<sup>[4]</sup>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<sup>[6]</sup>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

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

(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

$$f(\lambda x + (1 - \lambda)y) \leq \lambda f(x) + (1 - \lambda)f(y) \quad (1.2.1)$$

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<sup>[8]</sup>证明了这种解的存在性; 这个函数方程的通解见 J. Aczél<sup>[9]</sup>). 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

$$\begin{aligned} f(x) &= x \quad (-1 < x < 1) \\ f(-1) &= f(1) = 2 \end{aligned}$$

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<sup>[1]</sup>以及 F. Bernstein 和 G. Doetsch<sup>[10]</sup>):

**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<sup>[11]</sup>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<sup>[12]</sup>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<sup>[13]</sup>研究了若  $J$  凸函数  $f$  在  $T$  上有界, 则