Pure Mathematics 理论数学, 2024, 14(4), 184-189
Published Online April 2024 in Hans. https://www.hanspub.org/journal/pm
https://doi.org/10.12677/pm.2024.144125

Hans汉斯

数学分析中凸函数的性质及其应用

马　雷

上海理工大学理学院，上海

收稿日期：2024年3月9日；录用日期：2024年3月31日；发布日期：2024年4月25日

摘　要

凸函数是数学分析中的一个重要内容，同时也是教学的重点和难点。本文从凸函数的定义出发，介绍了凸函数三个常见的等价定义，并且仔细研究了凸函数的性质。最后又从凸函数的角度讨论了几个常用不等式的证明。

关键词

数学分析，凸函数，不等式

The Properties of Convex Functions in Mathematical Analysis and Its Applications

Lei Ma

College of Science, University of Shanghai for Science and Technology, Shanghai

Received: Mar. 9th, 2024; accepted: Mar. 31st, 2024; published: Apr. 25th, 2024

Abstract

Convex functions are an important content of mathematical analysis, and it is also the focus and difficulty part of teaching. This paper introduces three equivalent definitions of the convex function, and gives some properties of convex functions. Finally, several useful inequalities are also discussed from the perspective of convex functions.

Keywords

Mathematical Analysis, Convex Functions, Inequalities

文章引用: 马雷. 数学分析中凸函数的性质及其应用[J]. 理论数学, 2024, 14(4): 184-189.
DOI: 10.12677/pm.2024.144125

Copyright © 2024 by author(s) and Hans Publishers Inc.
This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).
http://creativecommons.org/licenses/by/4.0/
Open Access

1. 引言

在数学分析中，凸函数的定义、性质及其应用一直是教学的重点和难点。凸函数是一类特殊的函数，在数学、经济学、物理学等领域中都有广泛的应用[1]。在纯数学领域，凸函数由于其独特的性质，如单调性、二阶导数的非负性等，被广泛应用于优化问题、微积分和不等式的研究中。在凸函数理论中，凸集、凸包络和凸优化等问题都占据了重要地位，这些问题的解决有助于我们更深入地理解数学中的一些基本问题。本文先从定义出发，介绍了什么是凸函数，同时这个定义也是凸函数最基础、最重要的性质之一。

凸函数有很多重要的应用场景，例如它的图像是一个凸集，它的局部最小值就是全局最小值等等。这些性质为我们研究凸函数提供了重要的工具。在优化问题中，我们经常需要找到函数的最小值或最大值，如果函数具有凸性这个性质，这可以帮助我们更好地解决这个问题。接着本文又介绍了凸函数的性质，并且说明了当函数具有更高的正则性时，凸性这一条件可以为研究函数的其他性质提供更多帮助。同时凸函数的另一个重要应用是不等式的证明，有些复杂的不等式证明往往可以通过凸函数的办法把证明过程简单化，从而达到通俗易懂的目的。另外凸函数还有很多其他的应用，例如[2]与[3]都是对凸函数做的推广并且证明了相关不等式。

2. 凸函数的常见定义及等价条件

凸函数有多种定义。这里将介绍三种常见的定义，并给出这三种定义在一定条件下可以相互等价。

2.1. 凸函数的定义

定义 1 设函数 $\varphi(x)$ 是定义在区间 I 上的实函数。如果对任意的 $x_1, x_2 \in I$，以及任意常数 $t \in (0,1)$，均有

$$\varphi(tx_1 + (1-t)x_2) \le t\varphi(x_1) + (1-t)\varphi(x_2), \tag{1}$$

那么称 $\varphi(x)$ 是下凸的函数，简称凸函数。若将上式中“≤”改为“<”，则称 $\varphi(x)$ 是严格下凸的函数，简称严格凸函数。

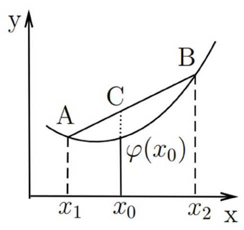

Figure 1. The definition of the convex function
图 1. 凸函数示意图

几何意义：当 $x_1 < x_0 < x_2$ 时，点 $C = (x_0, \varphi(x_0))$ 位于两点 $A = (x_1, \varphi(x_1))$ 和 $B = (x_2, \varphi(x_2))$ 连线的上方(见图 1)。例如：指数函数在实数 $\mathbb{R}$ 上是凸的。

定义 2 设函数 $\varphi(x)$ 是定义在区间 I 上的实函数。如果对任意的 $x \in I$， $y \in I$，有

$$\varphi\left(\frac{x_1+x_2}{2}\right) \le \frac{\varphi(x_1)+\varphi(x_2)}{2}, \tag{2}$$

那么称 $\varphi(x)$ 凸函数。若将上式中“$\le$”改为“$<$”，则称 $\varphi(x)$ 是严格凸函数。

定义 3 设函数 $\varphi(x)$ 是定义区间 I 上的实函数。如果对任意的 $\forall x_1, x_2, \cdots, x_n \in I$，有

$$\varphi\left(\frac{x_1+x_2+\cdots+x_n}{n}\right) \le \frac{\varphi(x_1)+\varphi(x_2)+\cdots+\varphi(x_n)}{n}, \tag{3}$$

那么称 $\varphi(x)$ 凸函数。若将上式中“$\le$”改为“$<$”即是严格凸的定义。

根据上述的定义，常见的凸函数：$\mathrm{e}^x (x \in \mathbb{R})$， $x^2 (x>0)$， $-\ln x (x>0)$。

2.2. 等价证明

这一节中，我们将给出 2.1 节中凸函数三种定义相互等价的证明，首先证明定义 2 与定义 3 相互等价。接着证明定义 1 与定义 2、3 在函数连续的情况下二者等价。具体将用两个定理来表示。

定理 1 定义 2 与定义 3 等价。

证明：首先(2)式是(3)式的一种特殊情况，即 $n=2$ 的情形。所以定义 3 可以推出定义 2。

另一方面，由(2)式可知

$$\varphi\left(\frac{x_1+x_2+x_3+x_4}{4}\right) = \varphi\left(\frac{\frac{x_1+x_2}{2}+\frac{x_3+x_4}{2}}{2}\right) \le \frac{\varphi\left(\frac{x_1+x_2}{2}\right)+\varphi\left(\frac{x_3+x_4}{2}\right)}{2} \le \frac{\varphi(x_1)+\varphi(x_2)+\varphi(x_3)+\varphi(x_4)}{4}。$$

依此类推，可得对任意的正整数 k 有

$$\varphi\left(\frac{x_1+x_2+\cdots+x_{2^k}}{4}\right) \le \frac{\varphi(x_1)+\varphi(x_2)+\cdots+\varphi(x_{2^k})}{2^k}。$$

所以(3)式对一切 $n=2^k$ 成立。现在我们用反向数学归纳法。假设(3)式对 $n=k+1$ 成立，证明(3)式对 $n=k$ 成立。令 $T=\frac{x_1+x_2+\cdots+x_k}{k}$，所以 $kT = x_1+x_2+\cdots+x_k$，即

$$T = \frac{x_1+x_2+\cdots+x_k+T}{k+1}。$$

由假设(3)式对 $n=k+1$ 成立，有

$$\varphi(T) = \varphi\left(\frac{x_1+x_2+\cdots+x_k+T}{k+1}\right) \le \frac{\varphi(x_1)+\varphi(x_2)+\cdots+\varphi(x_k)+\varphi(T)}{k+1},$$

将 T 的定义代入上式，经整理可得(3)式。证毕。

定理 2 若函数 $\varphi(x)$ 在开区间 $I=(a,b)$ 上连续，则定义 1，2，3 等价。

证明：只需要证明定义 1 与定义 2 等价即可。首先当 $t=\frac{1}{2}$ 时，(2)式是(1)式的特殊形式，所以定义 1 可以推出定义 2。接着我们来证定义 2 推出定义 1。先证 t 为有理数的情况，再由函数的连续性，通过有理数来逼近无理数最终证明(1)式。由于 t 为有理数，因此 $t=\frac{m}{n}$，其中 m,n 均为正整数并且 $m<n$。通过

计算可得

$$\begin{aligned}\varphi\left(tx_1+(1-t)x_2\right)&=\varphi\left(\frac{mx_1+(n-m)x_2}{n}\right)\\&\le\frac{m\varphi(x_1)+(n-m)\varphi(x_2)}{n}\\&=t\varphi(x_1)+(1-t)\varphi(x_2)\end{aligned}。$$

上式的证明用到了定理 1 的结论。因此$t>0$为有理数时，(1)式成立。当$t>0$为无理数时，可知存在有理数列$\{t_n\}$使得$\lim\limits_{n\to+\infty}t_n\to t$。因此有函数$\varphi(x)$的连续性可以得出(1)式对任意的无理数$t>0$也成立。证毕。

通过上述内容可知，定理 2 的证明过程用到了函数$\varphi(x)$的连续性。因此定义 1 要强于定义 2 和定义 3，因此通常我们用定义 1 的方式来描述凸函数。本文的后续内容均用定义 1 的方式描述凸函数。

3. 凸函数的性质

定义 1 中，凸函数的几何意义是线段 AB 一直在曲线 AB 的上方。这一节中，我们将对凸函数进行更为细致的研究，并且得出几个常见的凸函数的性质。

性质 1 设函数$\varphi(x)$是定义在区间I上的凸函数则对任意的$x_1<x_2<x_3$，有

$$\frac{\varphi(x_2)-\varphi(x_1)}{x_2-x_1}\le\frac{\varphi(x_3)-\varphi(x_1)}{x_3-x_1}\le\frac{\varphi(x_3)-\varphi(x_2)}{x_3-x_2}。\tag{4}$$

证明： 令$t=\dfrac{x_3-x_2}{x_3-x_1}$。若$\varphi(x)$是凸函数，将$x_2=tx_1+(1-t)x_3$代入(1)式则有

$$\varphi(x_2)\le t\varphi(x_1)+(1-t)\varphi(x_3),$$

同时将t的定义代入上式并且经过整理可得

$$\frac{\varphi(x_2)-\varphi(x_1)}{x_2-x_1}\le\frac{\varphi(x_3)-\varphi(x_1)}{x_3-x_1}。$$

类似地，通过定义$t=\dfrac{x_2-x_1}{x_3-x_1}$可得(4)式中第二个不等式。证毕。

性质 2 函数$\varphi(x)$是定义在开区间$I=(a,b)$上的凸函数，则$\varphi(x)$必是连续函数。

证明： 设$x\in(a,b)$，并有$s<x<t$。利用性质 1 可得

$$\frac{\varphi(x)-\varphi(s)}{x-s}\le\frac{\varphi(t)-\varphi(x)}{t-x},$$

因此，当$s,t\to x$时，根据单调有界定理可知$\varphi'_-(x)$，$\varphi'_+(x)$均存在，并且$\varphi'_-(x)\le\varphi'_+(x)$。所以$\varphi(x)$在$x$处左右连续。证毕。

性质 3 函数$\varphi(x)$在区间I上可微，若$\varphi(x)$在区间I上是凸函数，则$\varphi'(x)$在区间I上单调递增。

证明： 设有$x_1<x_2<x_3<x_4$，连续应用性质 1 可得

$$\frac{\varphi(x_2)-\varphi(x_1)}{x_2-x_1}\le\frac{\varphi(x_3)-\varphi(x_2)}{x_3-x_2}\le\frac{\varphi(x_4)-\varphi(x_3)}{x_4-x_3},$$

所以对内点x均有$\varphi'_-(x)$，$\varphi'_+(x)$递增，而又$\varphi(x)$可微，所以$\varphi'(x)$在区间I上单调递增。证毕。

如果$\varphi(x)$二阶可导，通过性质 3 可知$\varphi''(x)\ge0$。这也是判断函数凸性的一个条件。凸函数的性质还

有很多具体参见[4]和[5]。

4. 凸函数的应用

当前利用凸函数证明不等式已经是一个很普遍的办法，[6]列举了一些例题。这一节中，本文也将利用凸函数证明一些常用不等式。

(詹森不等式)设 $\varphi(x)$ 是区间 I 上二阶可导的凸函数，那么对任意的 $x_1,x_2,\cdots,x_n\in I$，与 $t_1+t_2+\cdots+t_n=1\quad(t_i>0,i=1,2,\cdots,n)$，有下式成立

$$\varphi(t_1x_1+t_2x_2+\cdots+t_nx_n)\le t_1\varphi(x_1)+t_2\varphi(x_2)+\cdots+t_n\varphi(x_n)。$$

证明：令 $\overline{x}=t_1x_1+t_2x_2+\cdots+t_nx_n$，对每一点 $x_i\,(i=1,2,\cdots,n)$ 在 $\overline{x}$ 处进行泰勒展开，即

$$\varphi(x_i)=\varphi(\overline{x})+\varphi'(\overline{x})(x_i-\overline{x})+\frac{\varphi''(\xi_i)}{2}(x_i-\overline{x})^2\,(i=1,2,\cdots,n),$$

其中 ξ_i 是 $(x_i,\overline{x})$ 或 $(\overline{x},x_i)$ 中一点。因此

$$\varphi(x_i)\ge\varphi(\overline{x})+\varphi'(\overline{x})(x_i-\overline{x}),$$

对上式分别乘上 t_i 再相加，经整理可得

$$\sum_{i=1}^{n}t_i\varphi(x_i)\ge\sum_{i=1}^{n}t_i\varphi(\overline{x})+\varphi'(\overline{x})\sum_{i=1}^{n}(x_it_i-t_i\overline{x})=\varphi(\overline{x})。$$

将 $\overline{x}=t_1x_1+t_2x_2+\cdots+t_nx_n$ 代入上式，即证。

(詹森不等式的积分形式)设 μ 是定义在集合 $\Omega\subset\mathbb{R}^n$ 中 σ 代数的一个正测度，并且满足 $\mu(\Omega)=1$。若 $f(x)\in L^1(\mu)$，满足 $a<f(x)<b,\forall x\in\Omega$，并且有 $\varphi(x)$ 在区间 (a,b) 上是凸函数，那么有

$$\varphi\left(\int_\Omega f\,\mathrm{d}\mu\right)\le\int_\Omega\varphi(f(x))\mathrm{d}\mu。$$

证明：令 $t=\int_\Omega f\,\mathrm{d}\mu$，并且 $\beta=\sup\limits_{a\le s\le t}\dfrac{\varphi(t)-\varphi(s)}{t-s}$。由于 $\mu(\Omega)=1$，所以 $a<t<b$。又由 β 的定义可知 $\varphi(s)\ge\varphi(t)+\beta(s-t)$。

将 $s=f(x)$ 和 t 的值代入上式可得

$$\varphi(f(x))\ge\varphi(t)+\beta(f(x)-t)。$$

两边积分有

$$\int_\Omega\varphi(f(x))\mathrm{d}\mu\ge\mu(\Omega)\varphi\left(\int_\Omega f(x)\mathrm{d}\mu\right)+\beta\left(\int_\Omega f(x)\mathrm{d}\mu+\mu(\Omega)\int_\Omega f(x)\mathrm{d}\mu\right)=\varphi\left(\int_\Omega f(x)\mathrm{d}\mu\right)。$$

即证。

(霍尔德不等式)设常数 p,q 满足 $\dfrac{1}{p}+\dfrac{1}{q}=1$，$1<p,q<1$。$\Omega\subset\mathbb{R}^n$。若 $u\in L^p(\Omega)$，$v\in L^q(\Omega)$ 则有

$$\int_\Omega|uv|\mathrm{d}x\le\|u\|_{L^p(\Omega)}\|v\|_{L^q(\Omega)}。$$

证明：首先证明 $\|u\|_{L^p(\Omega)}=\|v\|_{L^q(\Omega)}=1$ 时的情形。由函数 e^x 是凸函数可知对 $a>0,b>0$ 有

$$ab=\mathrm{e}^{\ln a+\ln b}=\mathrm{e}^{\frac{1}{p}\ln a^p+\frac{1}{q}\ln b^q}\le\frac{1}{p}\mathrm{e}^{\ln a^p}+\frac{1}{q}\mathrm{e}^{\ln b^q}=\frac{a^p}{p}+\frac{b^q}{q}$$

将 $a=|u|$，$b=|v|$ 代入上式可得

$$|uv| \le \frac{|u|^p}{p} + \frac{|v|^q}{q},$$

两边积分

$$\int_\Omega |uv|\,\mathrm{d}x \le \frac{1}{p} + \frac{1}{q} = 1 = \|u\|_{L^p(\Omega)} \|v\|_{L^q(\Omega)}。 \tag{5}$$

当 $\|u\|_{L^p(\Omega)}$，$\|v\|_{L^q(\Omega)}$ 不为 1 时，设

$$\|u\|_{L^p(\Omega)} = K_1, \|v\|_{L^q(\Omega)} = K_2$$

此时 $\frac{u}{K_1}$ 和 $\frac{v}{K_2}$ 满足

$$\left\|\frac{u}{K_1}\right\|_{L^p(\Omega)} = \left\|\frac{v}{K_2}\right\|_{L^q(\Omega)} = 1,$$

因此利用(5)式可得

$$\int_\Omega \left|\frac{uv}{K_1 K_2}\right| \mathrm{d}x \le 1。$$

证毕。

参考文献

[1] 陈秋涵. 凸函数在微观经济学中的应用[J]. 科技通报, 2014(5): 48-50.

[2] 梁亦孔. 函数凸性在证明级数发散中的应用[J]. 上海工程技术大学学报, 2021, 35(1): 94-96.

[3] 詹婉荣, 于海. 强凸函数的一个不等式性质及其应用[J]. 大学数学, 2022, 38(3): 93-96.

[4] 裴礼文. 数学分析中的典型问题与方法[M]. 第 2 版. 北京: 高等教育出版社, 2017: 334-347.

[5] 谢惠民, 恽自求, 易法槐, 钱定边. 数学分析习题课讲义(上册) [M]. 北京: 高等教育出版社, 2003: 243-245.

[6] 晏忠红. 凸函数的应用[J]. 湖南工业职业技术学院学报, 2003, 3(4): 86-87.