

文章编号:1000-1735(2013)01-0011-04

doi:10.11679/lxxblk2013010011

凸函数的共轭函数的几何解释

董玉林¹, 邵福波², 夏尊铨³

(1. 山东科技大学 信息科学与工程学院, 山东 青岛 266510; 2. 青岛滨海学院 信息工程学院, 山东 青岛 266555;
3. 大连理工大学 数学科学学院, 辽宁 大连 116024)

摘 要:在凸分析和凸优化中,凸函数的共轭函数是一个非常重要的概念. 给定一个凸函数,要求出该函数的共轭函数,并对共轭函数的图像有一个直观清晰的了解并不容易. 而一个函数的图像对理解该函数的性质非常重要. 基于凸分析中凸函数及其共轭函数的概念及性质,给出了共轭函数的几何解释. 这种解释有助于对共轭函数的概念及其性质的理解.

关键词:共轭函数;凸函数;凸优化

中图分类号:O221 **文献标志码:**A

在凸函数及凸优化中,凸函数和凸函数的共轭函数以及他们之间的关系理论是一个非常重要的内容,对他们的详细介绍可参见文献[1-3]. 共轭函数被广泛地应用到科学研究与应用领域,如金融优化问题,网络优化问题,分布式优化等领域中^[4].

但是对给定的一个凸函数,通常情况下并不容易直观的想象到它的共轭函数的图像. 如果能对凸共轭函数的图像有一个直观的了解,就能更好地理解这个函数的一些性质,譬如:函数的凸性,函数的极值性质等,尤其是有助于理解凸函数及其共轭函数之间的关系. 笔者基于对凸函数及其共轭函数概念及其性质的研究,给出了凸函数的共轭函数的几何解释.

1 共轭函数的几何解释

对给定的凸函数 $f(x)$,给出该函数的共轭函数 $f^*(x^*)$ 的几何解释. 首先给出相关的一些基本概念.

定义 1 若 $f:R^n \rightarrow R$ 是定义在 R^n 上的凸函数,则 $f(x)$ 的共轭函数 $f^*:R^n \rightarrow R$ 定义为

$$f^*(x^*) = \sup_{x \in R^n} \{ \langle x, x^* \rangle - f(x) \}.$$

例 1 考虑凸函数 $f(x) = e^x, x \in R$,根据定义 1, $f(x)$ 的共轭函数 $f^*(x^*)$ 为

$$f^*(x^*) = \begin{cases} x^* \log x^* - x^* & \text{若 } x^* > 0, \\ 0 & \text{若 } x^* = 0, \\ +\infty & \text{若 } x^* < 0. \end{cases}$$

从定义 1,可以得到

收稿日期:2012-11-22

基金项目:国家自然科学基金项目(11171049);山东科技大学春蕾计划资助(2008AZZ049)

作者简介:董玉林(1968-),男,山东济宁人,山东科技大学副教授,博士;

夏尊铨(1937-),男,辽宁大连人,大连理工大学教授,博士生导师.

$$\begin{aligned}
 f^*(x^*) &= \sup_{x \in R^n} \{ \langle x, x^* \rangle - f(x) \} = \\
 &= \inf_{\mu^* \in R} \{ \mu^* \mid \mu^* \geq \langle x, x^* \rangle - f(x), \forall x \in R^n \} = \\
 &= \inf_{\mu^* \in R} \{ \mu^* \mid f(x) \geq \langle x, x^* \rangle - \mu^*, \forall x \in R^n \}
 \end{aligned} \quad (1)$$

根据式(1),可以给出凸函数的共轭函数的其中一种几何解释. 共轭函数 $f^*(x^*)$ 在点 x^* 处函数值的负数(即 $-f^*(x^*)$)可以理解为包含函数 $f(x)$ 上图的某些超平面的截距的上确界. 图 1 以凸函数 $f(x) = e^x, x \in R$ 为例,给出了详细的解释.

图 1 首先给出了 $f(x)$ 的图像,其中,把 x^* 理解为空间 R^2 中直线的斜率,考虑函数 $f(x)$ 图像下方,且以 x^* 为斜率的所有直线,则 $-f^*(x^*)$ 就是这些直线的截距的上确界. 根据图示可以看出,以 $-f^*(x^*)$ 为截距的直线,必定是函数 $f(x) = e^x$ 的图像的一条切线.

若 $x^* > 0$,通过计算得 $f(x)$ 的以 x^* 为斜率的切线方程为 $y = x^*x - x^* \log x^* + x^*$,则 $f^*(x^*) = x^* \log x^* - x^*$. 例如令 $x^* = \frac{1}{e}$,则以 $\frac{1}{e}$ 为斜率的直线和 $f(x)$ 相切于点 $(-1, \frac{1}{e})$,且该直线的截距为 $\frac{2}{e}$,所以 $f^*(\frac{1}{e}) = -\frac{2}{e}$; 若 $x^* = 0$,以 0 为斜率的直线,若在函数 $f(x)$ 的图像下方且具有最大的截距,必为 $y = 0$,即 x 轴,该截距为 0,即 $f^*(0) = 0$; 若 $x^* < 0$,则以 x^* 为斜率的任何直线,必与函数 $f(x) = e^x$ 相交,即不存在以 x^* 为斜率的直线在 $f(x) = e^x$ 的下方,则定义 $f^*(x^*) = +\infty$.

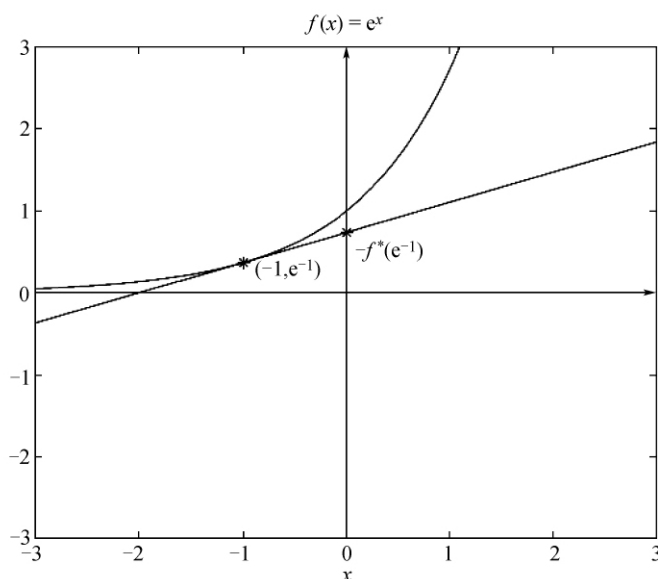


图 1 凸函数 $f(x)$ 及其共轭函数 $f^*(x^*)$ 之间的关系

Fig. 1 The relation of convex function $f(x)$ and its conjugate function $f^*(x^*)$

虽然图 1 能反映出凸函数 $f(x)$ 及其共轭数 $f^*(x^*)$ 的一些关系,但是不能直观地看到共轭函数 $f^*(x^*)$ 的图像. 下面给出凸函数及其共轭函数的另外一种几何解释,通过该几何图示,可以同时看到凸函数 $f(x)$ 及其共轭函数 $f^*(x^*)$ 的图像. 在给出几何图示之前,需要先给出一些相关基本概念及其性质.

定义 2 任给非空凸锥 K ,则锥 K 的极锥定义为

$$K^\circ = \{ x^* \mid \forall x \in K, \langle x, x^* \rangle \leq 0 \}.$$

例 2 考虑二维欧氏空间 R^2 中的凸锥

$$K = \{ (x_1, x_2) \mid x_2 \geq |x_1|, x_1 \in R, x_2 \in R \},$$

根据定义, 锥 K 的极锥为:

$$K^{\circ} = \{ (x_1, x_2) \mid x_2 \leq -|x_1| \}.$$

定义 3 任给非空凸集 C , 则由凸集 C 生成的凸锥记作 $\text{cone } C$, 定义为

$$\text{cone } C = \{ \lambda x \mid \lambda \geq 0, x \in C \}.$$

下面我们给出能刻画凸函数及其共轭函数关系的一个定理.

定理 1^[3] 若 $f: R^n \rightarrow R$ 是定义在 R^n 上的正常闭凸函数, $f^*: R^n \rightarrow R$ 是 $f(x)$ 的共轭函数, K 是由凸集

$$C_1 = \{ (1, x, \mu) \in R^{n+2} \mid \mu \geq f(x) \} \quad (2)$$

生成的锥. 而 K^* 是由凸集

$$C_2 = \{ (1, x^*, \mu^*) \in R^{n+2} \mid \mu^* \geq f^*(x^*) \} \quad (3)$$

生成的锥. 则下列结论成立.

$$\text{cl } K^* = \{ (\lambda^*, x^*, \mu^*) \mid (-\mu^*, x^*, -\lambda^*) \in K^{\circ} \}. \quad (4)$$

证明 略, 见文献[3].

给定凸函数 f , 定理 1 刻画了锥 K , 锥 K^* 及极锥 K° 之间的关系. 锥 K^* 与锥 K° 之间的关系可由下面的映射得到:

$$\phi: \text{cl } K^* \rightarrow K^{\circ},$$

其中

$$\phi((\lambda^*, x^*, \mu^*)) = (-\mu^*, x^*, -\lambda^*). \quad (5)$$

因此给定凸函数 $f(x)$, 可以由下列步骤得到 $f(x)$ 的共轭函数 $f^*(x^*)$ 的图像.

I 根据式(2), 得到凸集 C_1 .

II 根据定义 3, 由凸集 C_1 生成凸锥 K , 即:

$$\begin{aligned} K &= \text{cone } C_1 = \\ &= \{ \lambda C_1 \mid \lambda \geq 0, x \in C_1 \} = \\ &= \{ (\lambda, \lambda x, \lambda \mu) \in R^{n+2} \mid \lambda \geq 0, \mu \geq f(x) \}. \end{aligned}$$

III 由极锥的定义 2, 得到锥 K 的极锥 K° .

IV 再根据映射(5)的逆映射 $\phi^{-1}: K^{\circ} \rightarrow \text{cl } K^*$, 即:

$$\phi^{-1}((\mu^*, x^*, \lambda^*)) = (-\lambda^*, x^*, -\mu^*),$$

得到凸锥 $\text{cl } K^*$.

V 锥 $\text{cl } K^*$ 被超平面 $\mu = -1$ (即: $\lambda^* = 1$) 切割, 得到凸集 C_2 .

VI 由 C_2 作为函数 $f^*(x^*)$ 的上图, 得到 $f(x)$ 的共轭函数 $f^*(x^*)$.

$f(x)$ 及其共轭函数 $f^*(x^*)$ 的关系也可以用下式表示.

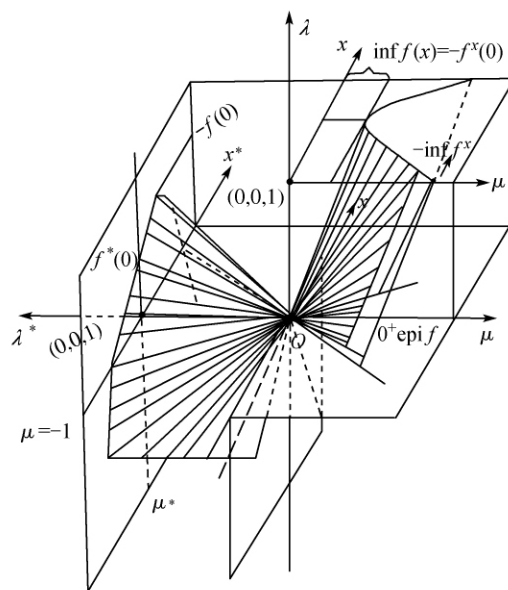
$$f(x) \rightarrow C_1 \rightarrow K \rightarrow K^{\circ} \rightarrow \text{cl } K^* \rightarrow C_2 \rightarrow f^*(x^*).$$

图 2 给出了凸函数 $f(x)$ 及其共轭函数 $f^*(x^*)$ 之间关系的一个示例. 注意到 ϕ^{-1} 的实质就是把 μ^* 及 λ^* 所代表的坐标轴互换位置, 并且这两个坐标轴的负向改成相反的正方向. 因此我们在同一个坐标系下给出了凸函数 $f(x)$ 及其共轭函数 $f^*(x^*)$ 的图像.

图 2 同时给出了凸函数 $f(x)$ 的回收锥及其共轭函数 $f^*(x^*)$ 的回收锥, 它们的极小值之间的关系, 如 $\inf f(x) = -f^*(0)$ 等.

2 结束语

给出了凸函数的共轭函数的几何解释, 该几何解释有助于我们对共轭函数概念的理解, 加深对凸函数及其共轭函数之间关系的理解.

图 2 凸函数 $f(x)$ 及其共轭函数 $f^*(x^*)$ Fig. 2 Convex function $f(x)$ and its conjugate function $f^*(x^*)$

参考文献:

- [1] FENCHEL W. Conjugate convex functions[J]. Can J Math, 1949, 1: 73-77.
- [2] MANFRED W. Theory of Duality in Mathematical Programming[M]. Jena, German: Democratic Re-public, 1989: 41-54.
- [3] ROCKAFELLAR R T. Convex Analysis[M]. New Jersey: Princeton University Press, 1970: 95-150.
- [4] STEPHEN Boyd, NEAL Parikh, ERIC Chu, et al. Distributed optimization and statistical learning via the alternating direction method of multiplier[J]. Foundations and Trends in Machine Learning, 2010, 3: 1-122.

Geometric interpretation of conjugates of convex functions

DONG Yulin¹, SHAO Fubo², XIA Zunquan³

(1. College of Information Science and Technology, Shandong University of Science and Technology, Qingdao 266510, China;

2. Institute of Information Engineering, Qingdao Binhai University, Qingdao 266555, China;

3. School of Mathematical Sciences, Dalian University of Science and Technology, Dalian 116024, China)

Abstract: Conjugates of convex functions is an important concept in convex analysis and convex optimization. Given an convex function, it is not easy to obtain the conjugate of the function, and have clear imagination of the graph of the conjugate function. But the graph of the function can help us to understand the properties of the function. This paper, based on the properties of the convex function and its conjugates function, presents the geometric interpretation of the conjugates of convex function. This interpretation can help us to understand the concept of the conjugate.

Key words: conjugates of convex functions; convex function; convex optimization