文章编号:1672 - 058X(2012)03 - 0022 - 05

非凸全局最优化的一种凸化、凹化方法*

刘呈军
(重庆师范大学 数学学院,重庆 401331)

摘 要:给出了一种凸化、凹化变换,将一个严格单调函数转化为一个凸或凹的函数;给出了一种凸化和凹化变换,在约束函数都是单调递减时,将一个既不单调凸也不是单调凹的目标函数转化为一个凸和凹函数;最终,将原始问题转换成一个凹极小问题或反凸规划问题来求得其最优解.

关键词:全局最优化;凹极小;反凸规划;凸化;凹化

中图分类号:O221.2 **文献标志码**:A

全局最优化问题广泛应用于经济模型、金融、网络、交通、分子生物学、环境工程学等领域. 由于非凸全局最优化很可能存在多个不同于全局最优解的局部最优解,传统的方法只能求其局部最优解,所以不能顺利地应用其求解全局最优解.

但是,幸运的是,大多数非凸全局优化问题,凸性会在某些限制的或是“相反”的意义下展现出来. 比如,凹极小问题,反凸规划问题和 D. C. 规划问题在全局最优化中都有非常重要的地位. 最近,有许多文献[1-5]讨论了关于凹极小问题、反凸规划问题和 D. C. 规划问题的全局最优解. 因此,如果能够将一个非凸规划问题转化成一个等价的凹极小问题、反凸规划问题或 D. C. 规划问题,那么可以通过现有的算法(比如,外逼近法、分支定界法等等)来求得它的全局最优解或近似全局最优解. 所以,对一个非凸最优化问题进行凸化、凹化方法的研究是非常有意义和必要的一项工作. 此处提出一种凸化、凹化变换,将一个具有多个单调约束函数的非单调规划问题转化为一个等价的凹极小或一个反凸规划问题.

某些凸化、凹化变换能够将严格单调函数转化为凸或凹的函数,并且在约束函数都是单调递减时,借用这些变换能够将某些既不单调凸也不是单调凹的目标函数转化为凸和凹函数,意义在于利用这种方法将原始问题转换成一个凹极小问题或反凸规划问题来求得其最优解.

考虑如下形式的全局最优化问题:

$$\begin{aligned} &\min f(x) \\ \text{s.t.}\quad & g_i(x) \leqslant 0, i = 1,2,\cdots,m \\ & x \in X \end{aligned} \tag{1}$$

其中 $f, g_i: R^n \to R, i = 1, 2, \cdots, m$. 在 X 上二次连续可微,X 为一个箱子集,特别地,设:

收稿日期:2011 - 07 - 17;修回日期:2011 - 09 - 15.

*基金项目:重庆师范大学自然科学基金(08XLR022).

作者简介:刘呈军(1981 -),男,重庆云阳人,硕士研究生,从事最优化方法及其应用研究.

$$X = \{x \in R^n \mid 0 < l_i \leqslant x_i \leqslant u_{i,} i = 1,2,\cdots,n\} \tag{2}$$

$$\overline{X} = \{(x_1,\cdots,x_{n+1}) \mid l_i \leqslant x_i \leqslant u_i, i = 1,2,\cdots,n+1\} \tag{3}$$

其中 $0 < l_{n+1} \leqslant \min l_i, u_{n+1} \geqslant \max\limits_{1 \leqslant i \leqslant n} u_i + \ln m$.

给出几个定义:

定义 1　一个函数 $h: R^n \to R$ 称为在集合 $D \subset R^n$ 上是关于 x_i 单调增(单调减)的,如果对于 $x_i^1 < x_i^2$ 有 $h(x_1,\cdots,x_{i-1},x_i^1,x_{i+1},\cdots,x_n) \leqslant (\geqslant) h(x_1,\cdots,x_{i-1},x_i^2,x_{i+1},\cdots,x_n)$;一个函数 $h: R^n \to R$ 称为在集合 $D \subset R^n$ 上是关于 x_i 是严格单调增(严格单调减)的,如果对于 $x_i^1 < x_i^2$ 有 $h(x_1,\cdots,x_{i-1},x_i^1,x_{i+1},\cdots,x_n) < (>) h(x_1,\cdots,x_{i-1},x_i^2,x_{i+1},\cdots,x_n)$. 其中 $x_i^1, x_i^2 \in D_i = \{x_i \mid (x_1,\cdots,x_{i-1},x_i,x_{i+1},\cdots,x_n) \in D\}$.

定义 2　函数 $h: R^n \to R$ 称为集合 $D \subset R^n$ 上的单调增函数(单调减函数),如果对于 $\forall x, y \in D, h(x) \leqslant (\geqslant) h(y)$,其中 $x_i \leqslant y_i, i = 1,2,\cdots,n$.

定义 3　函数 $h: R^n \to R$ 称为集合 $D \subset R^n$ 上的严格单调增函数(严格单调减函数),如果对于 $\forall x, y \in D$, $h(x) < (>) h(y)$,其中 $x_i \leqslant y_i, i = 1,2,\cdots,n$ 且 $x \neq y$.

定义 4　式(1)称为单调规划问题,如果 $f, g_i, i = 1,2,\cdots,m$ 是单调的;式(1)称为严格单调规划问题,如果 $f, g_i, i = 1,2,\cdots,m$ 是严格单调的.

定理 1　设 $f(x) \in C^2(X)$ 满足以下条件:存在一个正数 $\delta_0 > 0$, s. t. $\dfrac{\partial f(x)}{\partial x_i} \leqslant -\delta_0 (\geqslant \delta_0)$, $\forall x \in X, i = 1,2,\cdots,n$. 设:

$$\Phi_q(y) = f\left(\frac{1}{q}\ln(1+y)\right) \tag{4}$$

$$Y_1 = \{y \in R^n \mid e^{ql_i} - 1 \leqslant y_i \leqslant e^{qu_i} - 1, i = 1,2,\cdots,n\} \tag{5}$$

一定存在一个 $q_1 > 0$, s. t. 当 $q > q_1$ 时,$\Phi_q(y)$ 在 y_1 上是凸(凹)的.

证明　设 $x = \dfrac{1}{q}\ln(1+y)$,对 $\forall y \in Y_1$,由式(4)有:

$$\nabla^2 \Phi_q(y) = \frac{1}{q^2(1+y_i)^2}\left(\nabla^2 f(x) - q\frac{\partial f(x)}{\partial x_i}\right), \forall y \in Y_1$$

设 $a_0(x) = \mathrm{diag}\left(\dfrac{1}{1+y_1}, \dfrac{1}{1+y_2}, \cdots, \dfrac{1}{1+y_n}\right)$, $a_1(x) = \mathrm{diag}\left(\dfrac{\partial f(x)}{\partial x_1}, \dfrac{\partial f(x)}{\partial x_2}, \cdots, \dfrac{\partial f(x)}{\partial x_n}\right)$. 设 $F(x)$ 和 $F_q(y)$ 分别是 $f(x)$ 在点 x 和 $\Phi_q(y)$ 在点 y 处的 hession 矩阵,则有 $F_q(y) = \dfrac{1}{q^2} a_0(x)(F(x) - q a_1(x)) a_0(x)$, $\forall y \in Y_1$. 因为 $a_0(x)$ 是非奇异的,所以 $F_q(y)$ 是正(负)定的当且仅当 $F(x) - q a_1(x)$ 是正(负)定的. 设 $\overline{\lambda} = \max\limits_{\substack{x \in X \\ d \in s^n}} d^{\mathrm{T}} F(x) d$, $\underline{\lambda} = \min\limits_{\substack{x \in X \\ d \in s^n}} d^{\mathrm{T}} F(x) d$,其中 S^n 是 R^n 中的单位球体,则如果 $\dfrac{\partial f(x)}{\partial x_i} \leqslant -\delta_0$, $\forall x \in X, i = 1,2,\cdots,n$,有 $d^{\mathrm{T}}(F(x) - q a_1(x)) d = d^{\mathrm{T}} F(x) d - q d^{\mathrm{T}} a_1(x) d \geqslant \underline{\lambda} + q\delta_0 \geqslant -|\underline{\lambda}| + q\delta_0$, $\dfrac{\partial f(x)}{\partial x_i} \geqslant \delta_0$, $\forall x \in X, i = 1,2,\cdots,n$,有 $d^{\mathrm{T}}(F(x) - q a_1(x)) d \leqslant \overline{\lambda} - q\delta_0 \leqslant |\overline{\lambda}| - q\delta_0$.

设 $q_1 = \left\{\dfrac{|\underline{\lambda}|}{\delta_0}, \dfrac{|\overline{\lambda}|}{\delta_0}\right\}$,当 $q > q_1$ 时,如果 $\dfrac{\partial f(x)}{\partial x_i} \leqslant -\delta_0 (\geqslant \delta_0)$, $\forall x \in X, i = 1,2,\cdots,n$,有 $d^{\mathrm{T}}(F(x) -$

$qa_1(x))d \geqslant -|\bar{\lambda}| + q\delta_0 > 0, (d^{\mathrm{T}}(F(x) - qa_1(x))d \leqslant |\bar{\lambda}| - q\delta_0 < 0), \forall x \in X, d \in S^n$. 所以,如果$\frac{\partial f(x)}{\partial x_i} \leqslant -\delta_0(\geqslant \delta_0), \forall x \in X, i = 1,2,\cdots,n, \forall x \in X, i = 1,2,\cdots,n$,当 $q > q_1$ 时,对 $\forall y \in Y_1, F_q(y)$ 是正(负)定的. 因此,如果$\frac{\partial f(x)}{\partial x_i} \leqslant -\delta_0(\geqslant \delta_0), \forall x \in X, i = 1,2,\cdots,n$,那么当 $q > q_1$ 时,$\Phi_q(y)$在 Y_1 上是凸(凹)的.

定理 2 设$f, g_i(x), i = 1,2,\cdots,m$ 在 X 上二次连续可微,满足下列条件:存在一个 $\delta_0 > 0$, s. t. $\frac{\partial g_k(x)}{\partial x_i} \leqslant -\delta_0, \forall x \in X, i = 1,2,\cdots,n, k = 1,2,\cdots,m$,设:

$$\varphi_q(x) = f\left(\frac{1}{q}\ln(1+y)\right) + qG_q\left(\frac{1}{q}\ln(1+y)\right) \tag{6}$$

$$\psi_q(x) = f\left(\frac{1}{q}\ln(1+y)\right) - qG_q\left(\frac{1}{q}\ln(1+y)\right) \tag{7}$$

其中 $G_q(x) \triangleq g_q(x) + x_{n+1}, g_q(x) \triangleq \frac{1}{q}\ln\sum_{i=1}^{m} e^{qg_i(x)}, q > 0$ 是文献[6]中定义的函数:

$$Y_2 = \{y \in R^{n+1} \mid e^{ql_i} - 1 \leqslant y_i \leqslant e^{qu_i} - 1, i = 1,2,\cdots,n+1\} \tag{8}$$

存在一个 $q_2 > 0$, s. t. 当 $q > q_2$ 时,$\varphi_q(y)$在 Y_2 上是凸的,$\psi_q(y)$在 Y_2 上是凹的.

证明 设 $x = \frac{1}{q}\ln(1+y), \forall x \in \bar{X}$,由式(6)有:

$$\frac{\partial \varphi_q(y)}{\partial y_i} = \left(\frac{\partial g_0(x)}{\partial x_i} + q\frac{\partial G_q(x)}{\partial x_i}\right)\frac{\partial x_i}{\partial y_i} = \frac{1}{q(1+y_i)}\left(\frac{\partial g_0(x)}{\partial x_i} + q\frac{\partial G_q(x)}{\partial x_i}\right)$$

$$\frac{\partial^2 \varphi_q(y)}{\partial y_i^2} = -\frac{1}{q(1+y_i)^2}\left(\frac{\partial g_0(x)}{\partial x_i} + q\frac{\partial G_q(x)}{\partial x_i}\right)\frac{\partial x_i}{\partial y_i} + \frac{1}{q(1+y_i)}\left(\frac{\partial^2 g_0(x)}{\partial x_i^2}\cdot\frac{1}{q(1+y_i)} + q\frac{\partial^2 G_q(x)}{\partial x_i^2}\cdot\frac{1}{q(1+y_i)}\right) = -\frac{1}{q(1+y_i)^2}\left(\frac{\partial g_0(x)}{\partial x_i} + q\frac{\partial G_q(x)}{\partial x_i} - \frac{1}{q}\frac{\partial^2 g_0(x)}{\partial x_i^2} - q\frac{\partial^2 G_q(x)}{\partial x_i^2}\right)$$

$$\frac{\partial^2 \varphi_q(y)}{\partial y_i \partial y_j} = \frac{1}{q(1+y_i)}\left(\frac{\partial^2 g_0(x)}{\partial x_i \partial x_j}\frac{1}{q(1+y_j)} + q\frac{\partial^2 G_q(x)}{\partial x_i \partial x_j}\cdot\frac{1}{q(1+y_j)}\right) = \frac{1}{q(1+y_i)(1+y_j)}\left(\frac{1}{q}\frac{\partial^2 g_0(x)}{\partial x_i \partial x_j} + \frac{\partial^2 G_q(x)}{\partial x_i \partial x_j}\right)d$$

设 $a(x) = \mathrm{diag}\left(\frac{\partial g_0(x)}{\partial x_1},\cdots,\frac{\partial g_0(x)}{\partial x_{n+1}}\right), b(x) = \mathrm{diag}\left(\frac{\partial G_q(x)}{\partial x_1},\cdots,\frac{\partial G_q(x)}{\partial x_{n+1}}\right), c(x) = \mathrm{diag}\left(\frac{1}{1+y_1},\cdots,\frac{1}{1+y_{n+1}}\right)$,所以$\nabla^2\varphi_q(y) = -\frac{1}{q}c(x)\left(a(x) + qb(x) - \frac{1}{q}\nabla^2 g_0(x) - \nabla^2 G_q(x)\right)c(x)$.

因为 $c(x)$是非奇异的,所以$\nabla^2\varphi_q(y)$是正定的当且仅当 $a(x) + qb(x) - \frac{1}{q}\nabla^2 g_0(x) - \nabla^2 G_q(x)$是负定的. 设 $D(x) = a(x) + qb(x) - \frac{1}{q}\nabla^2 g_0(x) - \nabla^2 G_q(x), \lambda_0 = \min_{\substack{x \in \bar{X}, d \in S^{n+1} \\ 0 \leqslant i \leqslant m}} d^{\mathrm{T}}\nabla^2 g_i(x)d$, S^{n+1}为 R^{n+1}中的单位球体,$\alpha_0 = \max_{x \in X, 1 \leqslant i \leqslant m}\frac{\partial g_0(x)}{\partial x_i}$,由文献[6]中的引理 2.3 有$\frac{\partial g_q(x)}{\partial x_i} \leqslant -\delta_0$,所以$\frac{\partial G_q(x)}{\partial x_i} = \frac{\partial g_q(x)}{\partial x_i} + \frac{\partial x_{n+1}}{\partial x_i} \leqslant -\delta_1, i = 1$,

$2,\cdots,m;\dfrac{\partial G_q(x)}{\partial x_i}=\dfrac{\partial g_q(x)}{\partial x_i}+\dfrac{\partial x_{n+1}}{\partial x_i}\geqslant\delta_1,i=n+1$,其中 $\delta_1=\min\{\delta_0,1\}$. 所以, $d^{\mathrm T}D(x)d=d^{\mathrm T}(a(x)+qb(x)-\dfrac{1}{q}\nabla^2 g_0(x)-\nabla^2 G_q(x))d\leqslant\alpha_0-q\delta_1-\dfrac{1}{q}\lambda_0-d^{\mathrm T}\nabla^2 G_q(x)d,x\in X$;由 $g_q(x)\triangleq\dfrac{1}{q}\ln\sum_{i=1}^m e^{qg_i(x)},q>0$ 有:

$$\forall x\in X,\nabla^2 g_q(x)=\sum_{j=1}^m r_j(x)\nabla^2 g_j(x)+q\sum_{j=1}^m r_j(x)\nabla g_j(x)\nabla g_j(x)^{\mathrm T}-$$

$$q\Big(\sum_{j=1}^m r_j(x)\nabla g_j(x)\Big)\Big(\sum_{j=1}^m r_j(x)\nabla g_j(x)\Big)^{\mathrm T}$$

其中 $r_j(x)=\dfrac{e^{qg_j(x)}}{\sum_{j=1}^m \mathrm{e}^{qg_j(x)}}$,显然有 $r_j(x)>0$ 且 $\sum_{j=1}^m r_j(x)=1$,由函数 $h(u)=u^2$ 的凸性,有 $\sum_{j=1}^m r_j(x)\delta_j^2(x)\geqslant\Big(\sum_{j=1}^m r_j(x)\delta_j(x)\Big)^2$,所以 $\forall x\in X,d\in R^n$ 有 $d^{\mathrm T}\nabla^2 g_q(x)d\geqslant\sum r_j(x)d^{\mathrm T}\nabla^2 g_j(x)d\geqslant\lambda_0\geqslant-|\lambda_0|$.

由文献[6]中的定理2.2的证明过程可知,$\nabla^2 G_q(x)=\begin{pmatrix}\nabla g_q(x) & 0\\ 0 & 0\end{pmatrix}$,所以 $d^{\mathrm T}\nabla^2 G_q(x)d\geqslant-|\lambda_0|,\forall x\in\overline{X},d\in S^{n+1}$,$d^{\mathrm T}D(x)d\leqslant\alpha_0-q\delta_1-\dfrac{1}{q}\lambda_0+|\lambda_0|,\forall x\in\overline{X},d\in S^{n+1}$.

设 $q_2=\max\left\{1,\dfrac{\alpha_0+2|\lambda_0|}{\delta_1}\right\}$,若 $1\geqslant\dfrac{\alpha_0+2|\lambda_0|}{\delta_1}$,即 $\alpha_0+2|\lambda_0|<\delta_1$ 时,当 $q>q_2$ 有:

$$d^{\mathrm T}D(x)d\leqslant\alpha_0-q\delta_1+2|\lambda_0|<\alpha_0-q_2\delta_1+2|\lambda_0|=\alpha_0+2|\lambda_0|-\delta_1<0$$

若 $1<\dfrac{\alpha_0+2|\lambda_0|}{\delta_1}$,即 $\alpha_0+2|\lambda_0|>\delta_1$,当 $q>q_2$ 有:

$$d^{\mathrm T}D(x)d\leqslant\alpha_0-q\delta_1+2|\lambda_0|<\alpha_0-q_2\delta_1+2|\lambda_0|=$$

$$\alpha_0-\frac{\alpha_0+2|\lambda_0|}{\delta_1}\cdot\delta_1+2|\lambda_0|=0$$

$\forall x\in\overline{X},d\in S^n$. 因此,当 $q>q_2$ 时,$D(x)$ 是负定的,$\forall x\in\overline{X}$. 所以当 $q>q_2$ 时,$\forall y\in Y_2$,$\nabla^2\varphi_q(y)$ 是正定的. 当 $q>q_2$ 时,$\varphi_q(y)$ 在 Y_2 上是凸的.

类似地可以证明,当 $q>q_2$ 时,$\psi_q(y)$ 在 Y_2 上是凹的.

参考文献:

[1] HOFFMAN K L A. A method for globally minimizing concave functions over convex set [J]. Math Program,1981,20:22-23

[2] BENSON H P. Deterministic algorithm for constrained concave minimization:A unified critical survey [J]. Naval Res Logist, 1996,43:765-795

[3] HORST R,PARDALOS P M,THOAI N V. Introduction to global optimization[M]. Kluwer Academic Publisher,Dordrecht, Netherland,1996

[4] TUY H. Convex Analysis and Global Optimization [M]. Kluwer Academic Publisher,Dordrecht/Boston/London,1998

[5] 申培萍. 全局优化方法[M]. 北京:科学出版社,2006

[6] WU Z Y,ZHANG L SH,LI SH L. The Global Optimization for Some Classes of Programming Problems[J]. OR Transactions,

2003,7(2):9-20

[7] WU Z Y,ZHANG L SH. The Approximate Global Optimal Solution for Some Constrained Programming Problem[J]. OR Transactions,2007,11(1):1-15

[8] SUN X L,MCKINNON K I M,LI D. A convexification method for a class of global optimization problems with applications to reliability optimization[J]. Journal of Global Optimization,2001,21:185-199

[9] LI D,SUN X L,BISWAL M P,et al. Convexification,concavification and monotonization in global optimizaton[J]. Annals of Operations Research,2001,105:213-226

[10] BAI F S,WU Z Y. A Novel Monotonization Transformation for Some Classes of Global Optimization Problems[J]. Asia-Pacific Journal of Operational Research,2006,23(3):371-392

[11] 吴至友. 全局最优化的几种确定性方法[D]. 上海:上海大学,2003

[12] 赵洁,陈林,赵克全. 一类多目标规划问题的混合型对偶[J]. 重庆工商大学学报:自然科学版,2011,28(2):145-146

A Convexification and Concavification Method for Non-convex Global Optimization

LIU Cheng-jun

(School of Mathematics,Chongqing Normal University,Chongqing 401331,China)

Abstract: A strictly monotone function had been converted into a convex or a concave function by the convexification or concavification transformations. If the constraint functions all are decreasing,then the objective function which is neither monotone convex nor concave can be converted into a convex and a concave function by the convexification and concavification transformations. Finally,the original problem was converted into concave minimization or reverse convex programming problem and its global optimal solution was obtained by studying the concave minimization or reverse convex programming problem.

Key words: global optimization;concave minimization;reverse convex programming;convexification;concavification

责任编辑:李翠薇