Advances in Applied Mathematics 应用数学进展, 2022, 11(5), 3060-3068
Published Online May 2022 in Hans. http://www.hanspub.org/journal/aam
https://doi.org/10.12677/aam.2022.115326

Hans汉斯

最优化对数障碍法

郑陈轩

云南财经大学统计与数学学院，云南 昆明

收稿日期：2022年4月27日；录用日期：2022年5月21日；发布日期：2022年5月31日

摘 要

在求解不等式约束优化问题中对数障碍函数方法是非常流行的，众所周知，对数障碍函数在线性规划与线性半定规划的内点方法中起着重要的作用。本文主要介绍了对数障碍方法及其算法，并通过计算例子说明此方法的有效性。

关键词

对数障碍，有效性

Optimization Logarithmic Barrier Method

Chenxuan Zheng

School of Statistics and Mathematics, Yunnan University of Finance and Economics, Kunming Yunnan

Received: Apr. 27^{th}, 2022; accepted: May 21^{st}, 2022; published: May 31^{st}, 2022

Abstract

The method of solving the logarithmic barrier function method is very popular in solving the non-equality constraint optimization problem, it is well known that the

文章引用: 郑陈轩. 最优化对数障碍法[J]. 应用数学进展, 2022, 11(5): 3060-3068.
DOI: 10.12677/aam.2022.115326

logarithmic barrier function plays an important role in linear planning and linear semi-planning. This paper mainly introduces the logarithmic barrier method and its algorithm, and the effectiveness of this method is illustrated by calculating examples.

Keywords

Logarithmic Barrier, Validity

1. 引言

障碍函数法的基本思想是：对于接近可行域边界的可行点施加越来越大的惩罚，对边界上的点施加无穷大的惩罚，以使迭代点封闭在可行域内。如此便要求可行域的内点集合是非空的，否则当可行点全被加上无穷大的惩罚时，惩罚则变得毫无意义。所以，障碍函数仅适合于不等式约束，障碍函数法又叫内点法 [1]，同倒数障碍函数法相似 [2]。近几年，对数障碍函数应用在策略优化和统计学领域 [3] [4]。

考虑不等式约束最优化问题

$$\begin{aligned} &\underset{x}{min} f(x) \\ &s.t. c_i(x) \geqslant 0, i = 1, 2, \cdots, m \end{aligned} \tag{1}$$

其中$f(x), c_i(x)$是连续函数，可行域记作:$S = \{x|c_i(x) \geq 0, i = 1, \cdots, m\}$。假设可行域存在内点，即障碍函数的一般特征是:在$S^0$内，罚函数是光滑的;当可行点趋于$S^0$ 的边界时，罚函数的值趋于$+\infty$ 。下面我们着重讨论一类常用的障碍函数,对数障碍函数

$$F(x, \mu) = f(x) - \mu \sum_{i=1}^{m} \log c_i(x) \tag{2}$$

其中$\mu > 0$称为障碍参数 [5]。

考虑一个不等式约束的优化问题

$$\begin{aligned} &min \frac{1}{2}(x_1 + 1)^2 + x_2 \\ &s.t. x_1 \geqslant 0, x_2 \geqslant 0 \end{aligned} \tag{3}$$

首先构造对数障碍函数

$$P(x;\mu)=\frac{1}{2}(x_1+1)^2+x_2-u[\log(x_1-1)+\log x_2] \tag{4}$$

其中μ为很小的正数

令

$$\frac{\partial F(x_1,\mu)}{\partial x_1}=x_1+1-\frac{\mu}{x_1-1}=0 \tag{5}$$

$$\frac{\partial F(x_2,\mu)}{\partial x_2}=1-\frac{\mu}{x_2}=0 \tag{6}$$

解得$x(\mu)=\begin{pmatrix}\sqrt{1+\mu}\\ \mu\end{pmatrix}$同时当前的黑塞矩阵为

$$\nabla^2=\begin{pmatrix}1+\frac{\mu}{(\sqrt{1+\mu}-1)^2} & 0\\ 0 & \frac{1}{\mu}\end{pmatrix} \tag{7}$$

是正定阵，所以$x(\mu)$是$F(x,\mu)$是极小值

当$\mu\to 0$时，$x(\mu)\to x^*=\begin{pmatrix}\sqrt{1}\\ 0\end{pmatrix}$

$$f^*=1 \tag{8}$$

2. 算法

基于对数障碍函数的算法旨在识别$P(x;\mu)$的近似最小值，并以μ 的值为递减序列。该类型算法非常类似于二次罚函数框架，可以写成如下。

给出$\mu_0>0$,误差$\tau_0>0$,初始点x_0^s;

for　$k=0,1,2,\cdots$

　求$P(\cdot;\mu_k)$近似的最小值x_k,从x_k^s开始，当$\|\nabla P(x;\mu_k)\|\leq\tau_k$停止；

　　如果满足最终收敛检验

　　　停止近似解x_k;

　　选择新的障碍参数$\mu_{k+1}\in(0,\mu_k)$;

　　选择新的开始点x_{k+1}^s

end(for)

正如我们前面提到的，$\nabla_{xx}^2P(x;\mu)$对于小的μ使得牛顿法成为该框架每次迭代中唯一真正有效的寻找逼近最小化x_k的技术。甚至牛顿的方法也遇到了困难，因为它所基于的二次泰勒级数近似

DOI: 10.12677/aam.2022.115326

的不充分，但可以使用各种技术来确保牛顿法在几步之内进入$P(x;\mu_k)$的收敛域，并随后迅速收敛于收敛性$x(\mu_k)$。这些技术包括:

用一种确保全局收敛的技术来改进牛顿法，如线搜索或信赖域。它们必须保证迭代停留在$P(x;\mu)$的域内，即严格可行域S^0 内，还必须考虑到对数障函数的特性。

为了使$P(x;\mu_k)$最小化，一个好的起始点x_k^s可以沿前面的近似极小化$x_{k-1},x_{k-2},\cdots$所定义的路径外推得到。

通过沿前面的近似极小化$x_{k-1},x_{k-2},\cdots$所定义的路径外推，可以得到一个很好的$P(x;\mu_k)$最小化的起始点x_k^s或者，通过向$\nabla_x P(x;\mu)$计算向x_{k-1}点的路径$\{x(\mu)|\mu>0\}$的近似切向量得到

$$\nabla^2_{xx}P(x;\mu)\dot{x}+\frac{\partial}{\partial\mu}\nabla_x P(x;\mu)=0 \tag{9}$$

通过计算第二项得到

$$\nabla^2_{xx}P(x;\mu)\dot{x}-\sum_{i=1}^{m}\frac{1}{c_i(x)\nabla c_i(x)} \tag{10}$$

将$x=x_{k-1}$和$\mu=\mu_{k-1}$ 带入，我们可以得到近似的 $dotx$，然后用它来获得框架第k 步的起始点x_k^s：

$$x_k^s=x_{k-1}+(\mu_k-\mu_{k-1})\dot{x} \tag{11}$$

对于每一次迭代中新的障碍参数μ_{k+1}的选择，人们提出了各种各样的启发方法。一个常见的因素是，即用μ_{k+1}（$\mu_{k+1}=0.2\mu_k$或$\mu_{k+1}=0.1\mu_k$）如果$P(x;\mu_k)$ 的最小化问题不是太难解决，如果一个好的起始点x_{k+1}^s 可以有一定可信度地提出 [6]。

3. 性质

3.1. 定理1

设公式(1)中的f和$-c_i, i\in I$ 均为凸函数，且公式(2)中定义的严格可行域为非空。取$\{\mu_k\}$ 为递减序列使$\mu_k\downarrow 0$，并设解集M 为非空有界解集。那么下面的陈述是正确的。

（i）对于任意一个$\mu>0$,$P(x;\mu)$在S^0中是凸的，并且在S^0上得到最小的$x(\mu_k)$(不一定是唯一的)。任何局部极小值$x(\mu_k)$ 也是$P(x;\mu)$的全局极小值。

（ii）极小化$x(\mu_k)$的任意序列都有一个收敛的子序列，并且这些序列的所有可能的极限点都在M 中。

（iii）对于任意的极小化序列$\{x(\mu_k)\}$有$f(x(\mu_k))\to f^*$, $P(x(\mu_k);\mu_k)\to f^*$

对于M为空(考虑$\min x$服从$-x\geq 0$)或M是无界(考虑$\min\limits_{(x_1,x_2)} x_1$服从$x_1\geq 0$)的凸函数是可能的，

在这种情况下，定理的结论不适用于一般情况 [7]。

对于一般不等式约束问题(1)，相应的结果在本质上更局部。给定一个问题的局部解x^*(也就是说,一个满足定理二阶充分条件以及严格的互补性和约束条件),对数障碍函数$P(x;\mu)$ 有一个局部最小值接近x^*所有足够小值。当严格可行域S^0无界时，下面的对数函数$P(x;\mu)$可以无界。当$\mu_k \downarrow 0$时，$P(x;\mu_k)$的局部极小化序列也有可能收敛到(1)的非解。

$P(x;\mu)$的最小值与满足优化问题(1)KKT条件的点(x,λ) 之间有重要关系。在极小值$x(\mu)$时，$P(x;\mu)$对x的梯度为零，即，

$$\nabla_x P(x(\mu);\mu) = \nabla f(x(\mu)) - \sum_{i\in I} \frac{\mu}{c_i(x(\mu))} \nabla c_i(x(\mu)) = 0 \tag{12}$$

如果我们定义一个拉格朗日乘子估计

$$\lambda_i(\mu) \overset{\text{def}}{=} \frac{\mu}{c_i(x(\mu))}, i \in I \tag{13}$$

把（12）写成

$$\nabla f(x(\mu)) - \sum_{i\in I} \lambda_i(\mu) \nabla c_i(x(\mu)) = 0 \tag{14}$$

此条件与第一KKT条件$\nabla_x L(x,\lambda)$相同,0为问题(1) 的解，其中拉格朗日函数L由

$$L(x,\lambda) = f(x) - \sum_{i\in I} \lambda_i c_i(x) \tag{15}$$

对不等式约束问题(1)的其他KKT条件进行检验，如下:

$$\begin{aligned} c_i(x) \geq 0, i \in I \\ \lambda_i \geq 0, i \in I \\ \lambda_i c_i(x) \geq 0, i \in I \end{aligned} \tag{16}$$

非负性条件显然由$x = x(\mu), \lambda = \lambda(\mu)$; 事实上，对于所有$i \in I$，$c_i(x(\mu))$和$\lambda_i(\mu)$ 是严格正的。只有KKT条件不满足时，为互补条件。根据定义，我们得到了

$$\lambda_i(\mu) c_i(x(\mu)) = \mu, i \in I \tag{17}$$

其中，μ是严格正的

由上述观察可知，当$\mu_k \downarrow 0$时，$P(x;\mu)$的最小值$x(\mu)$及其拉格朗日乘子估计$\lambda(\mu)$ 越来越接近满足(1) 的KKT条件。事实上，如果在解(1)x^*处满足一些附加假设，我们可以证明$(x(\mu),\lambda(\mu))$接近最优原对偶解(x^*,λ^*) 当$\mu_k \downarrow 0$。

3.2. 定理2

假设S^0非空，x^*是(1)的局部解，在这个局部解中满足某λ^* 的KKT 条件。也假设线性独立约束条件(LICQ)，严格互补条件，二阶充分条件是(x^*,λ^*)。那么下面的陈述是正确的。

(i)有一个唯一的连续可微向量函数$x(\mu)$，对于所有足够小的向量函数$x(\mu)$在x^*的某个邻近区域是$P(x;\mu)$的局部极小值，从而$\lim\limits_{\mu\downarrow 0} x(\mu)=x^*$。

(ii)对于(i)中的函数$x(\mu)$，拉格朗日乘子估计(2)定义的$\lambda(\mu)$ 收敛于λ^*当$\mu\downarrow 0$。

(iii)黑塞矩阵$\nabla^2_{xx}P(x;\mu)$对于所有足够小的μ 都是正定的 [8]。

C_p定义

$$C_p \overset{\text{def}}{=} \{x(\mu)|\mu>0\} \tag{18}$$

4. 计算例子

4.1. 考虑一个不等式约束的优化问题

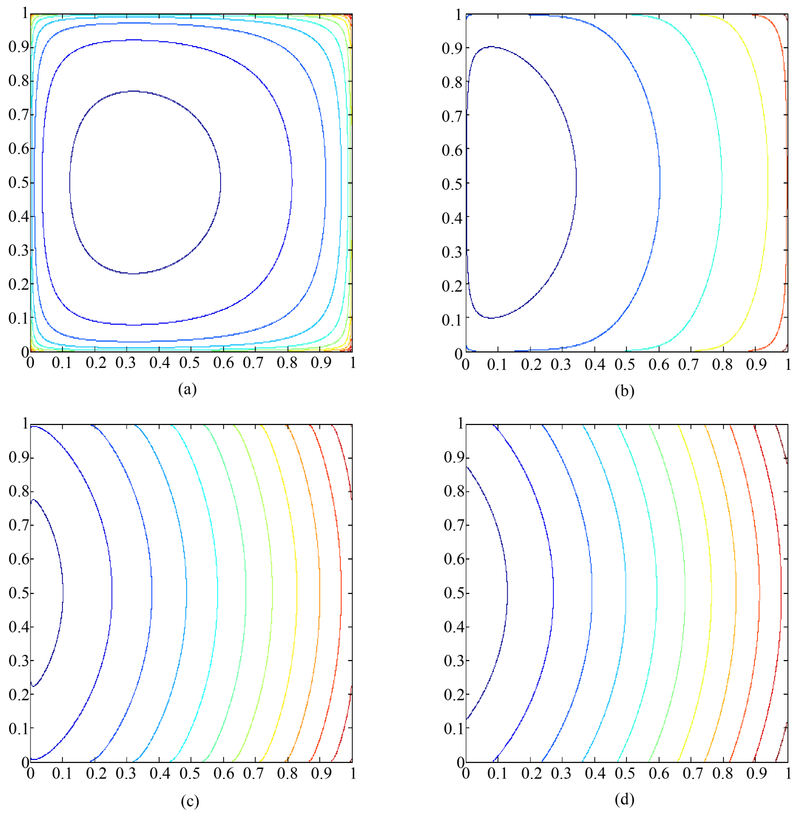

Figure 1. μ=1,0.1,0.01,0.001

图 1. μ=1,0.1,0.01,0.001

$$\min(x_1+0.5)^2+(x_2-0.5)^2 \quad s.t. \quad x_1\in[0,1],x_2\in[0,1] \tag{19}$$

首先构造对数障碍函数

$$P(x;u)=(x_1+0.5)^2+(x_2-0.5)^2-u[logx_1+log(1-x_1)+logx_2+log(1-x_2)] \tag{20}$$

这个函数的等高线μ=1,0.1,0.01,0.001被绘制在图1中。由图 1可知，除了可行区域边界附近外，其等高线都接近抛物线目标函数的等值线，即$\mu\downarrow 0$。前两幅图中最小值点周围的轮廓不是太偏心，距离椭圆也不是太远，这表明大多数无约束极小化算法都可以成功地应用于识别最小值点$x(\mu)$。

对于值$\mu=0.01$当最小的$x(\mu)$向对数障碍函数的“边界层”方向移动时，沿势垒函数的等高线会被拉长，不那么椭圆。在二次罚函数的情况下，等高线的伸长性质表明缩放效果不佳，这导致了拟牛顿、梯度下降和共轭梯度等无约束优化方法的性能不佳。牛顿的方法是不敏感的缩放，但非椭圆形的性质，等高线几乎是直线沿左边缘，而沿着圆形右边的边表明二次逼近，牛顿的方法是基于不能捕获真正的对数障碍函数的行为。因此,牛顿法也不能快速收敛到$x(\mu)$，除非在这个点的一个小邻域内。

4.2. 凸二次规划问题

凸二次规划是线性规划的推广，一些求解线性规划的内点算法已经成功的推广到了凸二次规划 [9]，下面给出凸二次规划问题。

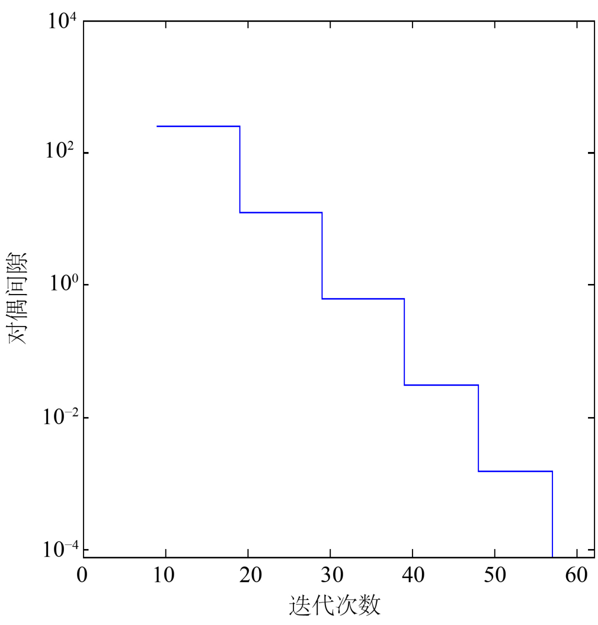

Figure 2. The relationship between the dual gap and the number of Newton iterations

图 2. 对偶间隙和Newton迭代次数关系

$$
\begin{aligned}
&\min f(x) = \frac{1}{2}x^T Qx + c^T x \\
&s.t. Ax \leq b,
\end{aligned} \tag{21}
$$

对数障碍函数

$$
P(x;\mu) = \frac{1}{2}x^T Qx + c^T x - \mu \sum_{i=1}^{n} \log x_i
$$

其中$\mu > 0$,μ是障碍因子，原问题与对偶问题的对偶间隙为：

$$
f(x) - P(x) = n\mu
$$

当$\mu \to 0$时，$f(x) - P(x) \to 0$。通过MATLAB软件进行编程，生成500行1000列的随机矩阵，采用Newton 回溯直线搜索方法，根据变化的μ寻找最优解的同时输出对偶间隙。

采用障碍法对偶间隙得出与Newton 迭代次数关系，如图 2所示。

从图 2中可以看出原问题与对偶问题的目标函数值之差即对偶间隙随着牛顿法的迭代次数的增加而越来越小。

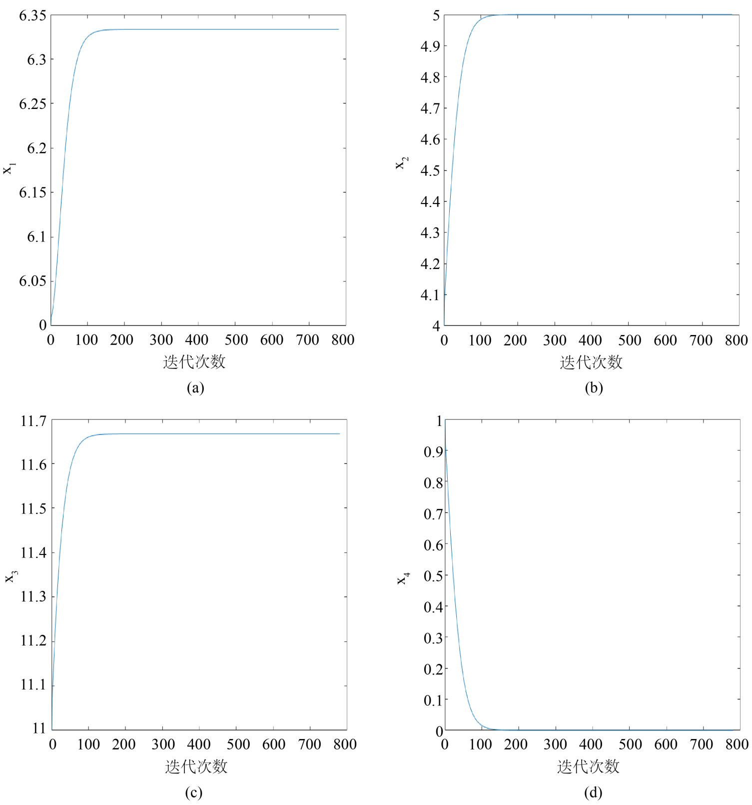

Figure 3. Iterative result

图 3. 迭代结果

考虑如下凸二次规划问题及其对偶问题：

$$\begin{aligned} &\min f(x) = \frac{1}{2}x_1^2 + x_2^2 + x_3^2 + \frac{1}{2}x_4^2 - 7x_1 - 5x_2 - 24x_3 + 9x_4 \\ &s.t. x_1 - x_2 + x_3 = 13 \\ &x_2 + x_4 = 5 \\ &x_1, x_2, x_3, x_4 \geq 0 \end{aligned} \tag{22}$$

从图 3中可以看出，利用原对偶对数障碍法求解凸二次规划问题，效果较好且收敛到极小值(6.3333,5.0000,11.6667,0.0000)。

5. 结论

根据计算实例结果得出，对数障碍法求解不等式约束优化问题和凸的二次规划问题非常的有效可行，但是对数障碍法有一定的局限性，初始点要取在可行域内，实际计算中较为麻烦，不同的问题对初始点有不同的要求。所以在以后的研究中，可以改进对数障碍法来求解最优化问题。

参考文献

[1] 袁亚湘, 孙文瑜. 最优化理论与方法[M]. 北京: 科学出版社, 1997.

[2] Polak, E., Higgins, J.E. and Mayne, D.Q. (1992) A Barrier Function Method for Minimax Problems. *Mathematical Programming*, **54**, 155-176. https://doi.org/10.1007/BF01586049

[3] Pourmohamad, T. and Lee, H.K.H. (2022) Bayesian Optimization via Barrier Functions. *Journal of Computational and Graphical Statistics*, **31**, 74-83. https://doi.org/10.1080/10618600.2021.1935270

[4] Zhang, R., Mei, J., Dai, B., Schuurmans, D. and Li, N. (2022) On the Effect of Log-Barrier Regularization in Decentralized Softmax Gradient Play in Multiagent Systems. arXiv preprint arXiv:2202.00872

[5] 陈宝林. 最优化理论与算法[M]. 北京: 清华大学出版社有限公司, 2005.

[6] Frisch, K.R. (1955) The Logarithmic Potential Method of Convex Programming. Memorandum, University Institute of Economics, Oslo, 5.

[7] Wright, M.H. (1992) Interior Methods for Constrained Optimization. *Acta Numerica*, **1**, 341-407. https://doi.org/10.1017/S0962492900002300

[8] Fiacco, A.V. and McCormick, G.P. (1990) Nonlinear Programming: Sequential Unconstrained Minimization Techniques. Society for Industrial and Applied Mathematics, Philadelphia, PA. https://doi.org/10.1137/1.9781611971316

[9] Fang, S.-C. and Puthenpura, S. (1993) Linear Optimization and Extensions: Theory and Algorithms. Prentice-Hall, Inc., Hoboken.