

全局最优化问题的一个无参数的填充函数算法

李博, 鲁殿军

(青岛科技大学数理学院, 山东 青岛 266061)

摘要: 本文研究了全局最优化问题. 利用构造填充函数的方法, 提出了一个新的无参数填充函数, 它是目标函数的一个明确表达式. 得到了一个新的无参数填充函数算法, 数值试验结果表明该填充函数算法是有效的, 从而推广了填充函数算法在求解全局最优化问题方面的应用.

关键词: 非线性规划; 全局最优化; 确定性算法; 填充函数

MR(2010) 主题分类号: 90C30 中图分类号: O221.2

文献标识码: A 文章编号: 0255-7797(2014)04-0773-06

1 引言

填充函数法由葛仁溥教授在 1990 年提出, 是求解全局最优化问题的一种确定性算法. 实践表明它是求解全局最优化问题的一种有效方法. 在社会生产中遇到的许多实际问题都可以转化为全局最优化问题. 因此, 对它的研究有着重要的理论和实际应用意义.

文献 [1] 中给出了一个带两个参数的填充函数

$$p(x_1^*, x, r, \rho) = \frac{1}{r + f(x)} \exp\left(-\frac{\|x - x_1^*\|^2}{\rho^2}\right). \quad (1.1)$$

数值试验表明, 该填充函数和算法是有效的, 但存在如下的缺陷 [2]: 首先, 该算法的运算效率很大程度上依赖于两个参数 r 和 ρ . 如果参数 r 较大而 ρ^2 与 $r + f(x)$ 的比较小, 那么 $f(x)$ 的全局极小点会在计算中被遗漏; 其次文献 [3] 中, 由于函数 (1.1) 中存在指数项 $\exp(-\frac{\|x - x_1^*\|^2}{\rho^2})$, 在 $\|x - x_1^*\|^2$ 过大或 ρ 过小时, $p(x_1^*, x, r, \rho)$ 和 $\|\nabla p(x_1^*, x, r, \rho)\|$ 将变得很小, 从而产生假的平衡点.

为解决以上问题, 以后的学者构造了一些新的填充函数, 其中文献 [3] 给出了一个只含有一个参数的填充函数

$$H(x, x_1^*, a) = \frac{1}{\ln[1 + f(x) - f(x_1^*)]} - a \|x - x_1^*\|^2, \quad (1.2)$$

其中参数 a 应充分大. 该填充函数比 (1.1) 在效率上有了一定提高, 但仍存在参数的选取问题. 文献 [4] 给出了一无参数填充函数

$$F(x, x_1^*) = \frac{1}{(\|x - x_1^*\| + 1)} \psi(f(x) - f(x_1^*)), \quad (1.3)$$

*收稿日期: 2013-05-15

接收日期: 2013-09-03

作者简介: 李博 (1957-), 男, 山东桓台县, 教授, 主要研究方向: 最优化理论与应用.

其中

$$\psi(t) = \begin{cases} 1, & t \geq 0, \\ 0, & t < 0. \end{cases}$$

数值试验表明填充函数 (1.3) 是有效的. 由于不需要调整参数, 提高了运算效率. 但 $F(x, x_1^*)$ 不是关于目标函数 $f(x)$ 的明确表达式.

本文研究了以上填充函数, 为克服其存在的缺陷, 构造了一个无参数的填充函数, 该填充函数具有关于目标函数 $f(x)$ 的明确表达式. 进行了数值试验.

2 全局优化模型和填充函数定义

关于填充函数的盆地和山头的定义详见参考文献 [1].

本文考虑如下全局无约束最优化问题

$$\min_{x \in R^n} f(x), \quad (2.1)$$

其中 $f(x)$ 是 $R^n \rightarrow R^1$ 的目标函数, x 是 n 维向量.

设目标函数 $f(x)$ 满足一下假设 [1]:

- (1) 目标函数 $f(x)$ 在 R^n 上连续可微;
- (2) 当 $\|x\| \rightarrow +\infty$ 时, $f(x) \rightarrow +\infty$. 在该假设的条件下, 意味着存在一个有界闭区域 Ω , 使得在它的内部包含了 $f(x)$ 的所有全局极小点;
- (3) 目标函数 $f(x)$ 在 Ω 上只有有限个局部极小点. 所以每个极小点都是孤立的, 而且对于 $f(x)$ 的每一个极小点 x^* 处的盆地 B^* 内的任意点 $x \neq x^*$, 都有 $f(x) > f(x^*)$.

定义 1 [1] 设 x^* 是 $f(x)$ 的一个已知的局部极小点, 如果 $P(x, x^*)$ 满足一下条件:

- (1) x^* 是 $P(x, x^*)$ 的局部极大点, 并且 $f(x)$ 的整个盆地 B^* 变成了 $P(x, x^*)$ 的山头的一部分;
- (2) $P(x, x^*)$ 在 $f(x)$ 的任何比 B^* 高的盆地里没有极小点或者鞍点;
- (3) 如果 $f(x)$ 在 x_0^* 处有比 B^* 低的盆地 B_0^* , 则 B_0^* 中存在一个点 x' , 它在 x^* 和 $\overline{x_0^*}$ 的直线上且最小化 $P(x, x^*)$, 其中 $\overline{x_0^*}$ 是 x_0^* 的某一邻域中的任意一点, 则 $P(x, x^*)$ 称为 $f(x)$ 在点 x^* 处的一个填充函数.

3 无参数填充函数

构造填充函数:

$$W(x, x^*) = -[\arctan(e^{(f(x)-f(x^*))}) - 1] \|x - x^*\|^2,$$

其中 x^* 是 $f(x)$ 的一个已知的局部极小点.

定理 1 $W(x, x^*)$ 是一个连续函数.

证 由假设条件 (1) 易得 $W(x, x^*)$ 是一个连续函数.

定理 2 设 x^* 是 $f(x)$ 的一个已知的局部极小点, 则点 x^* 是 $W(x, x^*)$ 的一个严格局部极大点.

证 因为 x^* 是 $f(x)$ 的一个局部极小点, 所以在 x^* 的盆地 B^* 内, 任意 $x \in B^*$ 且 $x \neq x^*$ 时, $f(x) > f(x^*)$. 所以对任意的 $x \in B^*$ 且 $x \neq x^*$,

$$W(x, x^*) = -[\arctan(e^{f(x)-f(x^*)}) - 1] \|x - x^*\|^2 < 0,$$

又因为 $W(x^*, x^*) = 0$, 所以 $W(x, x^*) < W(x^*, x^*)$. 所以点 x^* 是 $W(x, x^*)$ 的一个严格局部极大点.

定理 3 设 $f(x) > f(x^*)$ 且 $x \neq x^*$, 则 $\nabla W(x, x^*) \neq 0$.

证 因为 $W(x, x^*) = -[\arctan(e^{f(x)-f(x^*)}) - 1] \|x - x^*\|^2$, 则

$$\nabla W(x, x^*) = -\frac{e^{f(x)-f(x^*)}}{1 + [e^{f(x)-f(x^*)} - 1]^2} \|x - x^*\|^2 \nabla f(x) - 2[\arctan(e^{f(x)-f(x^*)}) - 1](x - x^*),$$

由 $x \neq x^*$ 且 $f(x) > f(x^*)$ 可知

$$\frac{e^{f(x)-f(x^*)}}{1 + [e^{f(x)-f(x^*)} - 1]^2} \|x - x^*\|^2 > 0, \arctan(e^{f(x)-f(x^*)}) - 1 > 0.$$

因为 $f(x)$ 在 R^n 上连续可微, 所以由泰勒公式展开得

$$f(x^*) = f(x) + (x^* - x)^T \nabla f(x) + o(\|x^* - x\|),$$

所以 $f(x^*) - f(x) = (x^* - x)^T \nabla f(x) + o(\|x^* - x\|)$. 因为 $f(x) > f(x^*)$ 得 $f(x^*) - f(x) < 0$. 所以

$$(x^* - x)^T \nabla f(x) + o(\|x^* - x\|) < 0.$$

即 $(x^* - x)^T \nabla f(x) < 0$. 所以, $(x - x^*)^T \nabla f(x) = -(x^* - x)^T \nabla f(x) > 0$, 所以

$$\begin{aligned} (x - x^*)^T \nabla W(x, x^*) &= -\frac{e^{f(x)-f(x^*)}}{1 + [e^{f(x)-f(x^*)} - 1]^2} \|x - x^*\|^2 (x - x^*)^T \nabla f(x) \\ &\quad - 2[\arctan(e^{f(x)-f(x^*)}) - 1] \|x - x^*\|^2. \end{aligned}$$

所以 $(x - x^*)^T \nabla W(x, x^*) < 0$.

此定理证明了: 当 $f(x) > f(x^*)$ 时, $\nabla W(x, x^*)$ 没有鞍点, 且可得 $d = x - x^*$ 为 $W(x, x^*)$ 的一个下降方向.

定理 4 设 x^* 不是 $f(x)$ 的一个全局极小点, 而 x_0^* 是 $f(x)$ 邻近 x^* 的另一个局部极小点, 使得 $f(x_0^*) < f(x^*)$, 则存在一个点 $x' \in B_0^*$, \bar{x}_0^* 为 x_0^* 的某一领域内的任意一点, x' 在 x^* 和 \bar{x}_0^* 的直线上并且使得 $W(x, x^*)$ 最小.

证 因为 x^* 是的一个局部极小点, 则对于 x^* 处的盆地 B^* 内的任意 $x \neq x^*$ 点, 有 $f(x) > f(x^*)$, 则 $W(x, x^*) = -[\arctan(e^{f(x)-f(x^*)}) - 1] \|x - x^*\|^2 < 0$.

x_0^* 是 $f(x)$ 邻近 x^* 的另一个局部极小点, 则存在 x_0^* 的某一领域 $O(x_0^*, \delta_0^*)$, 使得任意 $\bar{x}_0^* \in O(x_0^*, \delta_0^*)$, 有 $f(x^*) > f(\bar{x}_0^*) > f(x_0^*)$, 则

$$W(\bar{x}_0^*, x^*) = -[\arctan(e^{f(\bar{x}_0^*)-f(x^*)}) - 1] \|\bar{x}_0^* - x^*\|^2 > 0.$$

易得 $W(x^*, x^*) = 0$, 当 x 远离 x^* 时, $W(x, x^*) < 0$; 当 x 接近 \bar{x}_0^* 时, $W(x, x^*) > 0$, 可得, 当 $f(x) > f(x^*)$ 时, $W(x, x^*)$ 递增; 当 $f(x) < f(x^*)$ 时, $W(x, x^*)$ 递减. 又因为 $W(x, x^*)$

的连续性, 可得在 x_0^* 的某一领域内的点 $\overline{x_0^*}$, 在 x^* 和 $\overline{x_0^*}$ 的直线上有点 x' 使得 $W(x, x^*)$ 最小.

定理 5 若 x^* 是 $f(x)$ 的一个全局极小点, 则对任意的 $x \neq x^*$, 都有 $W(x, x^*) \leq 0$ 成立; 否则存在一个点 x_0 和它的一个邻域 $O(x_0, \delta_0)$, 其中 $\delta_0 > 0$, 并且 $f(x_0) < f(x^*)$, 使得对任意的 $x \in O(x_0, \delta_0)$, 有 $W(x, x^*) > 0$ 成立.

证 因为 x^* 是 $f(x)$ 的一个全局极小点, 则对任意的 $x \neq x^*$, 都有 $f(x) \geq f(x^*)$, 易得 $W(x, x^*) \leq 0$.

如果 x^* 不是 $f(x)$ 的一个全局极小点, 则存在 x_0 使得 $f(x_0) < f(x^*)$, 由 $f(x)$ 的连续性可知: 存在 x_0 的一个邻域 $O(x_0, \delta_0)$, 其中 $\delta_0 > 0$, 使得对于任意的 $x \in O(x_0, \delta_0)$, 都有 $f(x) < f(x^*)$, 所以 $W(x, x^*) > 0$. 由以上证明可得 $W(x, x^*)$ 满足填充函数的定义.

4 算法和数值试验

本文对文献 [2] 的算法加以改进.

步 1 设定搜索步长 $\delta > 0$, 选取任一初始点 $x_0 \in \Omega$, 令 $k = 1, i = 1, K = 2 \cdot \min(3^n - 1, 2n)$, 其中 n 为变量的个数.

步 2 以 x_0 为初始点, 用一局部极小化找到得目标函数 $f(x)$ 的一个局部极小点 x^* , 并计算出局部极小值 $f(x^*)$.

步 3 如果 $i < K$, 转到步 4; 否则 x^* 作为全局极小点, 停止运算.

步 4 以 x^* 为初始点, 寻找到 $W(x, x^*)$ 的局部极小点 x_1^* .

步 5 令 $x = x_1^*$.

步 6 如 $W(x, x^*) > 0$ 或者 $f(x) < f(x^*)$, 则令 $k = k + 1, x_0 = x$, 转步 2; 否则转步 7.

步 7 如果 $j < 3^n - 1$, 令 $x = x + \delta e_j$, 转步 6, 否则, $j = j + 1$; 若 $x \notin X$, 则 $i = i + 1$, 转步 3; 否则, 转步 7.

用 Matlab 编程对此算法加以实现, 成功找到下列函数的全局极小点:

(1) $f(x) = \sin x + \sin(10x/3) + \ln x - 0.84x$ s.t. $1 \leq x \leq 6$, 全局极小点 $x^* = 5.1998$, 全局最小值 $f^* = -4.6013$.

表 1: 计算结果

N	x_0	x^*	$f(x^*)$	k
1	1	5.1998	-4.6013	3
2	3	5.1998	-4.6013	2
3	5	5.1998	-4.6013	1

(2) $f(x) = \sin x + \sin(3x)$ s.t. $0 \leq x \leq 4$, 全局极小点 $x^* = 3.7571$, 全局最小值 $f^* = -1.5396$.

表 2: 计算结果

N	x_0	x^*	$f(x^*)$	k
1	1	3.7571	-1.53963	2
2	2	3.7571	-1.53963	2
3	3	3.7571	-1.53963	1

(3) 3-hump back camel function:

$$f(x) = 2x_1^2 - 1.05x_1^4 + x_1^6/6 - x_1x_2 + x_2^2 \quad \text{s.t.} \quad -3 \leq x_1, x_2 \leq 3,$$

全局极小点 $x^* = (0, 0)$, 全局最小值 $f^* = 0$.

表 3: 计算结果

N	x_0	x^*	$f(x^*)$	k
1	(-2,0.5)	(-0.2766e-008,0.1852 e-008)	2.39e-17	2
2	(2,0.5)	(0.1428e-008,0.2224e-008)	5.85e-14	8
3	(0.8,0.8)	(-0.2298e-007, -0.0449e-007)	9.74e-16	1

(4) Simplified Rosenbrock's function:

$$f(x) = 0.5(x_1^2 - x_2)^2 + (x_1 - 1)^2 \quad \text{s.t.} \quad -3 \leq x_1, x_2 \leq 3,$$

全局极小点 $x^* = (1.0000, 1.0000)$, 全局最小值 $f^* = 0$.

表 4: 计算结果

N	x_0	x^*	$f(x^*)$	k
1	(-3,3)	(1,1)	2.56e-10	2
2	(-2,2)	(1,1)	1.23e-10	2
3	(2,2)	(1,1)	2.02e-10	1

5 结语

本文构造了一个连续的无参数填充函数 $W(x, x^*)$, 它克服了含参数的填充函数在应用时调节参数的麻烦, 提高了实用性, 同时也使的算法的效率得到了提高. 将新的填充函数用 Matlab 编程对一些目标函数进行数值试验, 结果表明该方法对求解全局最优化问题是有效的.

该填充函数算法仍存在缺点, 它要求目标函数只能有有限个局部极小点. 其次, 算法的效率还需进一步提高, 下一步工作将在此基础上提高该算法对目标函数的适应性, 构造更高效率的填充函数.

参 考 文 献

- [1] Ge R P. A filled function method for finding a global minimizer of a function of several variables [J]. Math. Prog., 1990, 46: 191-204.
- [2] 王鹏, 李博, 王攀. 全局最优化问题的无参数填充函数法 [J]. 青岛科技大学学报 (自然科学版), 2008, 29(6): 553-556.
- [3] Liu Xian. Finding global minima with a computable filled function[J]. J. Global Optimization, 2001, 19: 151-161.

- [4] 茅嘉, 杨勇建. 一个无参数的填充函数算法 [J]. 应用数学与计算数学学报, 2010, 24(1): 35–44.
- [5] Yang Yongjian, Shang Youlin. A new filled function method for unconstrained global optimization [J]. Appl. Math. Comput., 2006, 173: 501–512.
- [6] Yang Yongjian, Liang Yumei. A new discrete filled function algorithm for discrete global optimization[J]. J. Comput. Appl. Math., 2007, 202: 280–291.

A PARAMETER-FREE FILLED FUNCTION METHOD FOR GLOBAL OPTIMIZATION

LI Bo, LU Dian-jun

(*College of Math. and Physics, Qingdao University of Science and Technology,
Shandong 266061, China*)

Abstract: In this article, we study the global optimization problem. By using the method of constructing filled function, we propose a new parameter-free filled function that is a clear expression of the objective function, and get a new non-parameter filled function algorithm. Numerical experiments show that the filled function algorithm is effective, which generalizes filled function algorithm in solving global optimization problems in the application.

Keywords: nonlinear programming; global optimization; deterministic method; filled function

2010 MR Subject Classification: 90C30