

一个新的连续可微的单参数填充函数

蔡珍珍 叶仲泉

(重庆大学数学与统计学院 重庆 401331)

摘要 填充函数法是求解非线性全局优化问题的有效方法。针对无约束优化问题,在目标函数及其梯度利普希兹连续的基础上,提出了一个新的连续可微的单参数填充函数,并研究了该填充函数的相关性质。最后,给出了一个填充函数算法,数值实验表明,该填充函数是有效的且算法是可行的。

关键词 全局优化,填充函数法,无约束

中图分类号 O224,TP301.6

文献标识码 A

DOI 10.11896/j.issn.1002-137X.2016.8.041

New Continuously Differentiable Filled Function with One Parameter

CAI Zhen-zhen YE Zhong-quan

(School of Mathematics and Statistics, Chongqing University, Chongqing 401331, China)

Abstract The filled function method is an effective approach to solve nonlinear global optimization problems. A new filled function with one parameter was proposed when the objective function has some certain conditions for unconstrained optimization problems, which is continuously differentiable. Then, theoretical properties of the filled function were investigated. At last, the paper gave several numerical experiments. The results show that the filled function is effective and the algorithm is feasible.

Keywords Global optimization, Filled function method, Nonlinear

1 引言

最优化是一门应用非常广泛的学科,随着科技的发展和社会的进步,最优化在工程设计、交通运输、生产管理、经济计划等方面都运用广泛。全局优化问题主要有两个困难:1)怎样从一个局部极小解出发找到更好的局部极小解;2)如何判断当前的极小解是全局最优解。全局优化主要分为两类算法:随机性算法和确定性算法。其中,填充函数法是一种有效的确定性算法,它由 Ge^[1]提出,主要解决第一种困难。其基本的思想是在已经得到的局部极小点处构造一个关于目标函数的复合函数,称为填充函数,利用该填充函数跳出当前的局部极小点,从而找到更好的局部极小点。填充函数法借助了目前发展较为成熟的局部极小化算法,受到了广大最优化研究者的欢迎,其关键在于填充函数的构造。

以下是文献[1-5]中的填充函数:

$$P(x, r, p) = \frac{1}{r + f(x)} \exp[-\|x - x^*\|^2 / p^2]$$

$$Q(x, A) = -[f(x) - f(x^*)] \exp[-A\|x - x^*\|^2], \\ A > 0$$

$$Z(x, A) = -\exp(\|x - x^*\| \arctan(A(f(x) - f(x^*))))$$

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

指数项的出现导致填充函数的梯度也含有指数项,使得在处理实际问题时可能会产生假的平稳点,使H函数缺失了 $f(x) < f(x^*) - 1$ 的点,导致一些好的极小点被忽略。

文献[6-9]的填充函数的构造克服了以上缺点,同时单参数也相对容易选择和处理,但是在 $f(x) \geq f(x^*)$ 时不出现目标函数的任何信息,因此对于某些实例,存在还未找到真正的局部极小点时算法就停止了的情况。本文将在无约束的情形下构造一个连续可微的填充函数,从而克服这一缺点。

2 基本知识

本文考虑以下无约束填充函数:

$$(UP) \begin{cases} \min f(x) \\ \text{s. t. } x \in R^n \end{cases}$$

假设1 $f(x)$ 函数是强制性的,即 $\|x\| \rightarrow \infty$ 时, $f(x) \rightarrow +\infty$ 。

那么存在一个有界闭集X,使得 $f(x)$ 的极小值点都在X的内部。

假设2 $f(x)$ 有有限个极小值。

假设3 $f(x)$ 和 $\nabla f(x)$ 连续可微,且在 R^n 上满足利普希兹连续的条件,即存在 $L > 0, L_{\nabla f} > 0$,使得任意的 $x, y \in R^n$,有下式成立。

$$|f(x) - f(y)| < L\|x - y\|$$

$$\|\nabla f(x) - \nabla f(y)\| < L_{\nabla f}\|x - y\|$$

定义1^[3] 函数 $F(x, x^*)$ 是关于 $f(x)$ 在局部极小点 x^* 的填充函数:

1) x^* 是 $F(x, x^*)$ 的一个严格的局部极大点;

2) 对于任意的 $x \in S_1, F(x, x^*)$ 的梯度不为0, $S_1 = \{x \in X | f(x) \geq f(x^*)\}$;

到稿日期:2015-06-11 返修日期:2015-08-14

蔡珍珍(1991-),女,硕士生,主要研究方向为全局优化理论与算法;叶仲泉(1961-),男,博士,教授,主要研究方向为最优化理论与算法、控制论、人工神经网络。

3) 如果 x^* 不是全局最优解, 那么 $F(x, x^*)$ 在 S_2 一定存在在局部极小点, $S_2 = \{x \in X | f(x) < f(x^*)\}$ 。

3 填充函数及其性质

文中提出的一个关于无约束问题的单参数填充函数为:

$$F(x, x^*, p) = -p \|x - x^*\|^3 \varphi(f(x) - f(x^*)) - \frac{1}{p} \{\max[0, f(x) - f(x^*)]\}^2$$

其中, φ 是连续可微的一元函数, 且满足 $\varphi(t) > 0, \varphi(t) =$

$$\frac{1}{p} (t > 0). \text{ 如 } \varphi(t) = \begin{cases} t^m + \frac{1}{p}, & t \leq 0 \\ \frac{1}{p}, & t > 0 \end{cases}, m \text{ 是偶数, 文中以 } m=2$$

为例。这里 x^* 是 $f(x)$ 当前的局部极小点, $p(p > 0)$ 是参数。

当参数 $p(p > 0)$ 充分大时, 以下定理表明 F 是一个满足定义 1 的填充函数。

定理 1 $F(x, x^*, p)$ 是连续可微的函数。

证明: 当 $f(x) > f(x^*)$ 时,

$$F(x, x^*, p) = -\|x - x^*\|^3 - \frac{1}{p} [f(x) - f(x^*)]^2$$

$$\nabla F(x, x^*, p) = -3\|x - x^*\|(x - x^*) - \frac{2}{p}(f(x) - f(x^*))\nabla f(x)$$

当 $f(x) \leq f(x^*)$ 时,

$$F(x, x^*, p) = -p\|x - x^*\|^3 \varphi(f(x) - f(x^*))$$

$$\nabla F(x, x^*, p) = -3p\|x - x^*\|(x - x^*)[(f(x) - f(x^*))^2 + \frac{1}{p}] + 2p\|x - x^*\|(f(x) - f(x^*))\nabla f(x)$$

由 $f(x)$ 连续可微, 易知 $F(x, x^*, p)$ 是连续可微的函数。

定理 2 $f(x)$ 当前的局部极小点 x^* 是 $F(x, x^*, p)$ 的一个严格的局部极大点。

证明: 由 x^* 是 $f(x)$ 的局部极小点和 $f(x)$ 的连续性知, 存在 x^* 的一个邻域 $U(x^*, \sigma), \sigma > 0$, 使得对于任意 $x \in U(x^*, \sigma)$ 且 $x \neq x^*$ 时, $f(x) \geq f(x^*)$, 那么

$$F(x, x^*, p) = -\|x - x^*\|^3 - \frac{1}{p} [f(x) - f(x^*)]^2 < 0 = F(x^*, x^*, p)$$

因此, x^* 是 $F(x, x^*, p)$ 的一个严格的局部极大点。

定理 3 假设 x^* 是 $f(x)$ 的局部极小点, 对任意 $x_1, x_2 \in X$, 若满足条件 $\|x_1 - x^*\| > \|x_2 - x^*\|, f(x^*) \leq f(x_2), f(x^*) \leq f(x_1)$, 那么在 p 充分大时, 下式成立:

$$F(x_1, x^*, p) < F(x_2, x^*, p) < 0 < F(x^*, x^*, p)$$

证明: 对任意的 $x \in S_1$,

$$F(x, x^*, p) = -\|x - x^*\|^3 - \frac{1}{p} [f(x) - f(x^*)]^2$$

当 $f(x^*) < f(x_2) < f(x_1)$ 时, 显然有 $F(x_1, x^*, p) < F(x_2, x^*, p)$ 成立。

当 $f(x^*) < f(x_1) < f(x_2)$ 时,

$$F(x_1, x^*, p) = -\|x_1 - x^*\|^3 - \frac{1}{p} [f(x_1) - f(x^*)]^2$$

$$F(x_2, x^*, p) = -\|x_2 - x^*\|^3 - \frac{1}{p} [f(x_2) - f(x^*)]^2$$

$$F(x_1, x^*, p) - F(x_2, x^*, p)$$

$$= -(\|x_1 - x^*\|^3 - \|x_2 - x^*\|^3) - \frac{1}{p} \{[f(x_1) - f(x^*)]^2 - [f(x_2) - f(x^*)]^2\} < 0$$

由 $f(x)$ 在 X 上是连续可微的, 可知 $[f(x_1) - f(x^*)]^2 - [f(x_2) - f(x^*)]^2$ 在有界闭集 X 上是有界的。因此, 当 p 充分大时, $F(x_1, x^*, p) < F(x_2, x^*, p)$ 成立。由 $f(x^*) \leq f(x_2)$ 易知 $F(x_2, x^*, p) < 0$ 。于是 $F(x_1, x^*, p) < F(x_2, x^*, p) < 0 < F(x^*, x^*, p)$ 。

定理 3 说明在 S_1 中, 当参数 p 充分大时, 离 $f(x)$ 的局部极小点 x^* 越远, 函数 $F(x, x^*, p)$ 的值越小, 于是在极小化填充函数的过程中, 要么找到更好的点使 $f(x) < f(x^*)$, 要么一直向 X 的边界移动。这样就消除了极小化填充函数的过程中出现来回跑动的情形。

定理 4 x^* 是 $f(x)$ 的局部极小点, 当参数 $p > 0$ 充分大时, 对于任意的 $x \in S_1, F(x^*, x^*, p)$ 的梯度不为 0。

证明: 对于任意的 $x_1 \in S_1$, 满足 $f(x_1) \geq f(x^*), x_1 \neq x^*, F(x_1, x^*, p) = -\|x_1 - x^*\|^3 - \frac{1}{p}(f(x_1) - f(x^*))^2$,

$$\nabla F(x_1, x^*, p) = -3\|x_1 - x^*\|(x_1 - x^*) - \frac{2}{p}(f(x_1) - f(x^*))\nabla f(x_1)。$$

当 $f(x_1) = f(x^*)$ 时, $\nabla F(x_1, x^*, p) = -3\|x_1 - x^*\|(x_1 - x^*) \neq 0$ 。

当 $f(x_1) > f(x^*)$ 时,

1) 若 $\nabla f(x_1) = 0$, 则

$$\nabla F(x_1, x^*, p) = -3\|x_1 - x^*\|(x_1 - x^*) \neq 0$$

2) 若 $\nabla f(x_1) \neq 0$, 则

$$\nabla F(x_1, x^*, p) = -3\|x_1 - x^*\|(x_1 - x^*) - \frac{2}{p}(f(x_1) - f(x^*))\nabla f(x_1)$$

$$\nabla F(x_1, x^*, p) = -3\|x_1 - x^*\|(x_1 - x^*) - \frac{2}{p}(f(x_1) - f(x^*))[\nabla f(x_1) - \nabla f(x^*)]$$

$$\nabla F(x_1, x^*, p)^T \frac{(x_1 - x^*)}{\|x_1 - x^*\|}$$

$$= -3\|x_1 - x^*\|^2 + \frac{2}{p} LL_{\nabla f} \|x_1 - x^*\|^2$$

$$= (\frac{2}{p} LL_{\nabla f} - 3) \|x_1 - x^*\|^2$$

$$\text{当 } p > \frac{2LL_{\nabla f}}{3} \text{ 充分大时, } \nabla F(x_1, x^*, p)^T \frac{(x_1 - x^*)}{\|x_1 - x^*\|} <$$

0, 那么 $\nabla F(x_1, x^*, p) \neq 0$ 。

定理 5 如果 x^* 不是全局最优解, 那么当 $p > 0$ 充分大时, 在 S_2 一定存在局部极小点。

证明: 记 $\Omega = \{x \in X | f(x) \leq f(x^-), x \in X\}, \partial\Omega = \{x \in X | f(x) = f(x^-), x \in X\}$, 因 $f(x)$ 是连续的, 所以 $\Omega, \partial\Omega$ 是有界闭集。那么对于任意的 $x \in \partial\Omega$, 有 $F(x, x^*, p) = -\|x - x^*\|^3$, 由于 $\partial\Omega$ 是紧集, 因此存在 $x' \in \partial\Omega$, 使得 $\min_{x \in \partial\Omega} F(x, x^*, p) = F(x', x^*, p) = -\|x' - x^*\|^3$ 。 x^* 不是全局最优解, 那么存在 $f(x)$ 的另一局部极小点 x^- , 满足 $f(x^-) < f(x^*), F(x^-, x^*, p) = -\|x^- - x^*\|^3 (p(f(x^-) - f(x^*))^2 + 1)$ 。当 $\|x' - x^*\|^3 \leq \|x^- - x^*\|^3$ 时, $F(x^-, x^*, p) < F(x', x^*, p)$ 显然成立; 当 $\|x' - x^*\|^3 > \|x^- - x^*\|^3$ 时, 如果 $p > \frac{\|x' - x^*\|^3 - \|x^- - x^*\|^3}{\|x^- - x^*\|^3 (f(x^-) - f(x^*))^2}$, $F(x^-, x^*, p) < F(x', x^*, p)$ 。由 Ω 是紧集, 存在 $x^0 \in \Omega$, 有 $\min_{x \in \Omega} F(x, x^*, p) = F(x^0, x^*, p)$, 那么 $\min_{x \in \Omega} F(x, x^*, p) = F(x^0, x^*, p) \leq F(x^-, x^*, p) < F(x', x^*, p) = \min_{x \in \partial\Omega} F(x, x^*, p)$, 于是 $x^0 \in \Omega \cap \partial\Omega \subset \text{int } X$ 是 $F(x, x^*, p)$ 的一个局部极小点。由于 $F(x, x^*, p)$ 是可微的,

因此 $\nabla F(x^0, x^*, p) = 0$, 于是由定理 4 可知 $x^0 \in S_2$ 。

定理 2—定理 5 说明文中第 3 节构造的关于目标函数的复合函数是填充函数。

4 算法和数值实验

4.1 算法

Step 1 令参数初始值 $p=100$, 上界为 10^8 ; 方向向量为 $d_j, j=1, 2, \dots, m(m \geq 2n)$; 运算次数为 $k, k=0$, 上界是 N, N 为充分大的正整数; ω 为搜索步长; 给定初始点 $x_k \in X$ 。

Step 2 以 x_k 为初始点, 极小化函数 $f(x)$ 得局部极小点 x_k^* 。

Step 3 在 $f(x)$ 的当前局部极小点 x_k^* 构造填充函数:

$$F(x, x^*, p) = -p \|x - x_k^*\|^3 \varphi(f(x) - f(x_k^*)) - \frac{1}{p} \{\max[0, f(x) - f(x_k^*)]\}^2$$

令 $i=1$ 。

Step 4 若 $i \leq m, x = x_k^* + \omega d_i$, 转 Step 5; 否则, 转 Step 7。

Step 5 从 x 出发, 用局部极小化方法极小化填充函数, 得到序列 $x_j (j=1, 2, \dots)$, 若存在 $x_{j_0} \notin X$, 停止填充函数的极小化, $i=i+1$, 转 Step 4; 否则, 得到填充函数的极小点 y , 转 Step 6。

Step 6 以 y 为初始点, 极小化函数 $f(x)$ 的局部极小点 x_{k+1}^* 。若 $f(x_{k+1}^*) < f(x_k^*)$, 令 $k=k+1$, 若 $k < N$, 转 Step 3, 否则算法终止, 称 $f(x)$ 的全局极小点为 x_k^* ; 若 $f(x_{k+1}^*) \geq f(x_k^*)$, 转 Step 7。

Step 7 若 $p \leq 10^8, p = p \times 10$, 转 Step 3; 否则, 算法终止, 称 $f(x)$ 的全局极小点为 x_k^* 。

注: 算法中的搜索步长 ω 可以根据需要调节; n 为所研究问题的维数; 在局部极小化函数时, 可以采用拟牛顿法、共轭梯度法等; 算法在 matlab7. 10. 0(R2010a) 的环境下进行。

4.2 演算实例

例 1 Rastrigin 函数

$$\begin{cases} \min f(x) = x_1^2 + x_2^2 - \cos(18x_1) - \cos(18x_2) \\ \text{s. t. } -1 \leq x_1, x_2 \leq 1 \end{cases}$$

该函数在给定可行域的局部极小点有多个, 如图 1 所示, 但是只有一个全局极小解 $x^* = (0, 0)^T$ 。

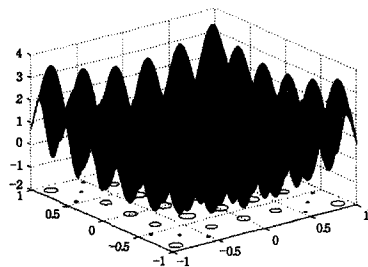


图 1 例 1 的函数图像

例 2 Two-dimensional 函数

$$\begin{cases} \min f(x) = [1 - 2x_2 + c \sin(4\pi x_2) - x_1]^2 - [x_2 - 0.5 \sin(2\pi x_1)]^2 \\ \text{s. t. } 0 \leq x_1 \leq 10, -10 \leq x_2 \leq 0 \\ c = 0.2, 0.5 \end{cases}$$

全局极小值都为 $f(x^*) = 0$ 。

例 3 Six-hump camel-back 函数

$$\begin{cases} \min f(x) = 4x_1^2 - 2.1x_1^4 + \frac{1}{3}x_1^6 - x_1x_2 - 4x_2^2 + 4x_2^4 \\ \text{s. t. } -3 \leq x_1, x_2 \leq 3 \end{cases}$$

该函数有两个全局极小解, 即 $x^* = (0.0898, 0.7127)^T$ 和 $x^* = (-0.0898, -0.7127)^T$; 该函数的全局极小值为 $f(x^*) = -1.0316$ 。

例 4 n-dimensional 函数

$$\begin{cases} \min f(x) = \frac{\pi}{n} [10 \sin^2 \pi x_1 + g(x) + (x_n - 1)^2] \\ \text{s. t. } -10 \leq x_j \leq 10, j=1, 2, \dots, n \end{cases}$$

其中 $g(x) = \sum_{i=1}^{n-1} [(x_i - 1)^2 (1 + 10 \sin^2 \pi x_{i+1})]$, 函数的全局最优解是 $x^* = (1, \dots, 1)^T$ 。

以上例题的具体演算结果如表 1 所列。

表 1 例题结果

例题	k	x^0	x_k^*	$f(x_k^*)$
1	3	(0.5; 0.5)	1.0e-07 * (0.0693; 0.1099)	-1.9999
2	2	(6; -2)	(2.7380; -0.7884)	0.0887
2	4	(8; -8)	(1.0000; 0.0000)	1.4500e-32
3	2	(-2; -1)	(-8.9946e-02; 0.7126)	-1.0316
4	2	(2; 2; 2; 2)	(1.0000; 1.0000; 1.0000; 1.0000; 1.0000)	0

结束语 在目标函数二次可微的条件下, 提出了一个新的连续可微单参数填充函数, 克服了文献[6-9]中填充函数在 $f(x) \geq f(x^*)$ 时不出现目标函数的任何信息的缺点, 并通过数值实验验证了其算法是可行的。

参考文献

- [1] Ge R P. A filled function method for finding a global minimizer of a function of several variables[J]. Mathematical Programming, 1990, 46: 191-204
- [2] Ge R P, Qin Y F. A class of filled functions for finding global minimizers of a function of several variables[J]. Journal of Optimization Theory and Applications, 1987, 54: 241-252
- [3] Yang Y J, Shang Y L. A new filled function Method for unconstrained global optimization[J]. Applied Mathematics and Computation, 2006, 173: 501-512
- [4] Yan Jun-hua, He Zhong-yue, Wang Yong-jun. A new filled function applied to global optimization[J]. Computer Science, 2009, 36(10): 250-252 (in Chinese)
颜俊华, 何中越, 王永军. 用于全局优化的一种新填充函数[J]. 计算机科学, 2009, 36(10): 250-252
- [5] Liu Xian. Finding global minima with a computable filled function[J]. Journal of Global Optimization, 2001, 19: 151-161
- [6] Gao C L, Yang Y J, Han B S. A new class of filled functions with one parameter for global optimization[J]. Computer & Mathematics Applications, 2011, 62(62): 2393-2403
- [7] Lin Hong-wei, Gao Yue-lin, Wang Yu-ping. A continuously differentiable filled function method for global optimization[J]. Numerical Algorithms, 2014, 66(3): 511-523
- [8] Wei Fei, Wang Yu-ping, Lin Hong-wei. A new filled function method with two parameter for global optimization[J]. Journal of Optimization Theory and Applications, 2014, 163: 510-527
- [9] Liang Yu-mei, Li Ming-ming, Chi Dong-xuan. A filled function method with one parameter for global optimization[J]. Journal of Operations Research, 2009, 13(4): 101-108 (in Chinese)
梁玉梅, 李铭明, 迟东璇. 全局优化问题的一种单参数填充函数方法[J]. 运筹学学报, 2009, 13(4): 101-108