

用于全局优化的一种新填充函数

颜俊华¹ 何中越¹ 王永军²

(西南石油大学计算机科学学院 成都 610500)¹ (东方地球物理公司研究院 涿州 072750)²

摘要 填充函数法是一种寻求多变量、多峰值函数的总体最优的优化方法。鉴于提出过的填充函数,给出了一种形式简单的单参填充函数,计算中无需考虑函数出现不连续点的情况,且函数不受指数项影响。对一些标准函数的仿真结果比较表明构造的填充函数是有效的。

关键词 填充函数法,全局优化,非线性优化

中图分类号 TP30, O29 **文献标识码** A

New Filled Function Applied to Global Optimization

YAN Jun-hua¹ HE Zhong-yue¹ WANG Yong-Jun²

(South West Petroleum University, Chengdu 610500, China)¹ (Bureau of Geophysical Prospecting(BGP), CNPC, Zhuozhou 072750, China)²

Abstract The filled function method is an effective approach to find the global optima of multivariable and multimodal functions. Inspired by the previous filled functions, we proposed a new filled function with simple form and one parameter. The case of discontinuous point in computation needn't be thought. Moreover, it overcame the disadvantages that some previous filled functions were affected by exponential terms. Numerical results on benchmark functions show that the proposed function is effective.

Keywords Filled function method, Global optimization, Nonlinear programming

现实中的许多问题都需要其总体最优值,在数学上表示为求目标函数的全局极小值。已存在的较为成熟的优化方法大多都是用来求函数的局部极值的,如牛顿法、拟牛顿法等。

葛人溥在1987年提出了一种用于全局优化的新方法,称之为填充函数法^[2,4]。该方法因设计简单,寻优方向明确,容易上机实现,对目标函数性态没有过多要求等优点而受到注意。所提出的一些较为有效的填充函数见文献^[1-3, 6, 9, 10]。

$$Q(x, A) = -[f(x) - f(x_1)] \exp[A||x - x_1||^2], \quad Q_1 - fun \quad (1)$$

$$Q(x, A) = A \exp(-||x - x_1||/\sqrt{A}) \operatorname{atan}(A[f(x) - f(x_1) + h]), \quad Q_2 - fun \quad (2)$$

$$P(x, A) = -(f(x) - f(x_1)) \exp(A \operatorname{atan}||x - x_1||^2), \quad Q_3 - fun \quad (3)$$

$$H(x, x_1) = 1/\ln(1 + f(x) - f_1(x_1)) - A||x - x_1||^2, \quad Q_4 - fun \quad (4)$$

这些函数或含双参数,或在等值线上不连续,或含对数式,或受指数式影响等均导致算法不收敛,出现计算溢出现象。

鉴于上述情况,在填充函数法的思想指导下,本文力求构造一个性能更优的填充函数。

1 填充函数的预备知识

填充函数法的主要思想是^[1]:利用局部极小点 x_1 和目标

函数 $f(x)$ 构造一个辅助函数——填充函数 $P(x)$ 。以 $P(x)$ 的稳定点 \hat{x} 为新的初始点,极小 $f(x)$ 并找到下一个极小点 x_2 使得 $f(x_2) < f(x_1)$ 。用 x_2 代替 x_1 ,构造新的填充函数,并找到 $f(x)$ 新的极小点 x_3 ,使得 $f(x_3) < f(x_2)$ 。重复以上过程,最终找到 $f(x)$ 的总体极小点。

定义 1^[1] 一个函数 $P(x)$ 称为 $f(x)$ 的填充函数,如果:

(1) x_1 是 $P(x)$ 的极大点,包含 x_1 的盆域^[1] S_1 是 $P(x)$ 的一个山丘,其中 x_1 是 $f(x)$ 的一个极小点。

(2) $P(x)$ 在比含 x_1 的域 S_1 高的盆域中没有稳定点。

(3) 若存在比 x_1 低的点 x_2 (即: $f(x_2) < f(x_1)$), 则 $P(x)$ 在比 x_1 低的 x_2 的盆中一定有稳定点。

2 填充函数 $Z(x, A)$ 的构造及性质分析

定义函数

$$Z(x, A) = -\exp(-||x - x_1||) \operatorname{atan}(A(f(x) - f(x_1) + h)) \quad (5)$$

其中, $A > 0$, h 是固定的非负数。为了简略,采用如下符号。

$f - f_1: f(x) - f(x_1)$, x_1 是 $f(x)$ 的极小点;

$L_1: L_1 = \max_{x \in \Omega} ||\nabla f(x)||$;

$L_2: L_2 = \min_{x \in \Omega} ||\nabla f(x)||$;

∇f 表示目标函数 $f(x)$ 的梯度,且 $\nabla f = \nabla f(x)$ 。

下面证明:当 $A > 0$ 适当大时, $Z(x, A)$ 是填充函数。

定理 1 如果 x_1 是 $f(x)$ 的一个局部极小点,那么 x_1 是

到稿日期:2008-12-02 返修日期:2009-03-31

颜俊华(1971-),女,讲师,主要研究方向为计算机应用等, E-mail: pengwen@vip. sina. com; 何中越(1979-),女,讲师,主要研究方向为计算机应用; 王永军(1970-),男,博士,主要研究方向为最优化等。

$Z(x, A)$ 的一个局部极大点。

证明:取方向 p ,不妨设 $\|p\|=1, Z(x, A)$ 在 x_1 的邻域中沿 p 的方向导数为

$$\begin{aligned} \frac{\partial}{\partial p} z(x, A) &= -\lim_{t \rightarrow 0^+} \frac{1}{t} [z(x_1 + tp, A) - z(x_1, A)] \\ &= -\lim_{t \rightarrow 0^+} \frac{1}{t} \{ \exp(t) \operatorname{atan}[A(f(x+tp) - f(x_1) + h)] \operatorname{atan}(Ah) \} \\ &= \operatorname{atan}(Ah) < 0 \text{ (用罗必达法则)} \end{aligned} \quad (6)$$

上述表明: $Z(x, A)$ 在 x_1 的盆中沿 $x-x_1$ 的任何方向都是下降的,因此 x_1 为其极大点。

定理 2 任给 $d \in R^n$,且 $f(x) > f(x_1)$,如果

$$\begin{cases} d^T \nabla f \geq 0 \\ d^T(x-x_1) > 0 \end{cases} \text{ 或 } \begin{cases} d^T \nabla f > 0 \\ d^T(x-x_1) \geq 0 \end{cases}$$

那么 d 是 $Z(x, A)$ 在 x_1 处的一个下降方向,特别沿 $x-x_1$ 方向 $Z(x, A)$ 是下降的。

据式(5)有

$$\begin{aligned} \nabla Z(x, A) &= -\exp\|x-x_1\| \left\{ \frac{(x-x_1)}{\|x-x_1\|} \operatorname{atan}[A(f-f_1+h)] + \frac{A \nabla f}{1+A^2(f-f_1+h)^2} \right\} \end{aligned} \quad (7)$$

$$\begin{aligned} d^T \nabla Z(x, A) &= -\exp\|x-x_1\| \left\{ \frac{d^T(x-x_1)}{\|x-x_1\|} \operatorname{atan}[A(f-f_1+h)] + \frac{Ad^T \nabla f}{1+A^2(f-f_1+h)^2} \right\} \end{aligned} \quad (8)$$

在给定的条件下不论 A 取何值,都有 $d^T \nabla f < 0$,特别在 $x-x_1$ 方向上成立。

定理 3 如果 $f(x) > f(x_1)$ 且 $d^T \nabla f < 0, d^T(x-x_1) > 0$,有

$$A > L_1 / (f-f_1)^2 \quad (9)$$

则 d 是 $Z(x, A)$ 于 x_1 处的下降方向,特别在 $x-x_1$ 方向上成立。

证明:只要证明在题设和式(9)的条件下 $d^T \nabla f < 0$ 即可。

由式(8)和本题设,要使 $d^T \nabla f < 0$,只要

$$\left\{ \frac{d^T(x-x_1)}{\|x-x_1\|} \operatorname{atan}[A(f-f_1+h)] + \frac{Ad^T \nabla f}{1+A^2(f-f_1+h)^2} \right\} > 0$$

即

$$\frac{1+A^2(f-f_1+h)^2}{A} \operatorname{atan}[A(f-f_1+h)] > -$$

$$\frac{\|x-x_1\| d^T \nabla f}{d^T(x-x_1)}$$

即可。又 $[1+A^2(f-f_1+h)^2]/A \operatorname{atan}[A(f-f_1+h)] > A(f-f_1)^2 \operatorname{atan}(Ah)$,及 A 充分大时, $\operatorname{atan}(Ah) \rightarrow \pi/2 > 1$,所以只要 $A(f-f_1)^2 > -\frac{\|x-x_1\| d^T \nabla f}{d^T(x-x_1)}$ 就可以了。

$$A > \frac{L_1}{(f-f_1)^2} = \left\| -\frac{\|x-x_1\| d^T L_1}{d^T(x-x_1)(f-f_1)^2} \right\| \geq \left\| -\frac{\|x-x_1\| d^T \nabla f}{d^T(x-x_1)(f-f_1)^2} \right\|$$

就有 $d^T \nabla f < 0$,显然对 $x-x_1$ 方向成立。

此定理说明:即使在 $f(x)$ 的一个极小点 x_2 盆中,只要 $f(x_2) > f(x_1)$,至少沿着方向 $x-x_1, Z(x, A)$ 将总是下降的。

因此不论在 $B_1(x_1) = \{x \in \Omega | f(x) > f(x_1)\}$,还是在 $B_2(x_1) = \{x \in \Omega | f(x) > f(x_1)\}, (f(x_1) > f(x_1), x_1$ 为一个极小点)中,只要满足式(9), $Z(x, A)$ 就没有任何极小点与鞍点。

定理 4 设 $f(x)$ 连续可微, Ω 是有界闭域,如果

$$\operatorname{atan}(Ah) > AL_1 / (1 + A^2 h^2) \quad (10)$$

则 $Z(x, A)$ 于 $B_1(x_1) = \{x \in \Omega | f(x) > f(x_1)\}$ 中无极点和鞍点。

证明:由式(7)知,若 $Z(x, A)$ 于 $B_1(x_1)$ 中有极小点或鞍点 x ,则 $\nabla Z(x, A) = 0$,从而

$$|\operatorname{atan}[A(f-f_1+h)]| = |A \nabla f / [1 + A^2(f-f_1+h)^2]| \quad (11)$$

因 x 在 $B_1(x_1) = \{x \in \Omega | F(x) > F(x_1)\}$ 中,从而 $f(x) > f(x_1)$,故

$$\operatorname{atan}[A(f-f_1+h)] > \operatorname{atan}[Ah] > \frac{AL_1}{1+A^2 h^2} \geq$$

$$\frac{A|\nabla f|}{1+A^2(f-f_1)^2}$$

当式(10)成立时,式(11)就不可能成立。

此定理说明:当 A 充分大时, $Z(x, A)$ 于 $B_1(x_1) = \{x \in \Omega | F(x) > F(x_1)\}$ 中无稳定点。

定理 5 如果 $f(x) > f(x_1)$ 且 $d^T \nabla f < 0, d^T(x-x_1) > 0$,那么

$$A < L_2 / [\operatorname{atan}[A(f-f_1+h)](1+f-f_1)] \quad (12)$$

则沿 d 的方向 $Z(x, A)$ 将上升。特别是 $Z(x, A)$ 在 $f(x)$ 的比 x_1 低的盆域中有一个极小点或鞍点。

证明:此题设与定理 3 相同。由式(8)得知要使 $d^T \nabla f > 0$,只要

$$\left\{ \frac{d^T(x-x_1)}{\|x-x_1\|} \operatorname{atan}[A(f-f_1+h)] + \frac{Ad^T \nabla f}{1+A^2(f-f_1+h)^2} \right\} < 0$$

即

$$\frac{1+A^2(f-f_1+h)^2}{A} \operatorname{atan}[A(f-f_1+h)] < -$$

$$\frac{\|x-x_1\| d^T \nabla f}{d^T(x-x_1)}$$

而

$$[1+A^2(f-f_1+h)^2]/A \operatorname{atan}[A(f-f_1+h)] < A(1+f-f_1+h) \operatorname{atan}[A(f-f_1+h)]$$

所以只要

$$A < \frac{\min|\nabla f|}{\operatorname{atan}[A(f-f_1+h)](1+f-f_1+h)}$$

$$= \frac{L_2}{\operatorname{atan}[A(f-f_1+h)](1+f-f_1)}$$

此时有 $d^T \nabla f > 0$,所以沿 d 方向 $Z(x, A)$ 将上升。显然当 $f(x) > f(x_1), h$ 值越小, $f(x) \rightarrow f_1(x)$ 时,尽管 A 很大,式(14)也能发生。因 Ω 是有界闭域, $Z(x, A)$ 必存在极小值,又极小点不可能发生在 $B_1(x_1)$ 中,所以只可能发生在比 x_1 所在的盆更低的盆中。

根据定理 1-定理 5,已经证明了 $Z(x, A)$ 符合 x_1 处填充函数的定义。记本文的 $Z(x, A)$ 为 $Z\text{-fun}$ 。

3 数值试验

3.1 填充函数中指数因子影响的避免

在式(1)右侧乘以 $\exp(-\lambda B)$,其中 λ 为大于 1 的整数, B

为固定的正常数。显然这样作用后的 $Z(x, A)$ 仍是填充函数(证略), 此时有

$$Z(x, A) = -\exp(|x - x_1| - \lambda B) \operatorname{atan}(A(f(x) - f(x_1) + h))$$

这样随着 $|x - x_1|$ 的增大, 只要 λB 大小适当, 就不会出现 $|x - x_1|$ 较大时函数 $Z(x, A)$ 及其梯度要么变化太快而导致溢出, 要么变化缓慢而影响收敛速度的问题。这样的 $Z(x, A)$ 克服了引言中所说的填充函数的缺点, 从而扩大了填充函数 $Z(x, A)$ 的使用范围。

3.2 数值实验

仿文献[1, 2], 本节根据函数的平均计算次数及找到给定精度解的成功率来比较本文填充函数与提出过的填充函数式(1)~(4)的有效性和可靠性。有效性比较如表1所列, 可靠性比较如表2所列。表中数值是10次平均结果。算法中的梯度值都是采用数值差分计算得到的。由于 z -fun 与 Q_1 -fun, Q_2 -fun 相比, 可靠性相当, 计算速度快, 因此表1只列出了它们的结果。 z -fun 与 Q_3 -fun, Q_4 -fun 相比, 成功率高, 计算速度相当, 所以图1只给出了 z -fun 与 Q_3 -fun, Q_4 -fun 可靠性比较结果。算法步骤请参考文献[1, 2]。

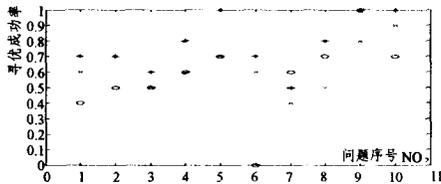


图1 Z -fun 和 Q_3 -fun, Q_4 -fun 寻优成功率比较

表1 数值试验结果表

No	n	Z-fun		Q ₁ -fun		Q ₂ -fun	
		nFf	nFq	nFf	nFq	nFf	nFq
1	2	146	33	348	72	341	60
	4	1417	97	1184	220	4525	395
	6	278	64	2985	639	2402	275
	20	3195	271	failed	failed	1895	271
2	2	170	36	170	36	589	67
	4	21479	4516	16852	3542	14481	2657
	6	1323332	27250	114941	23653	132332	27250
	8	503047	100718	274433	55257	260031	50485
	10	467518	96817	467518	96817	178006	34932
3	2	65	15	638	126	589	99
	4	634	34	1090	226	1985	374
	5	413	93	573	121	621	112
	6	5724	1063	failed	failed	failed	failed
	7	698	59	342	63	1267	62
	8	705	128	692	139	870	130
	9	25638	5539	25638	5539	25638	5539
	10	80276	16889	71857	15423	79855	16792
	15	50634	12467	54983	11707	59436	12467
10	4	107	21	107	21	107	21

本文的测试函数如表2所列, 详细情况见文献[3, 5, 7, 8]。图1和表1中的符号说明如下:

*: 表示 Z -fun 找到全局最优的成功率; O: 表示 Q_3 -fun 找到全局最优的成功率; X: 表示 Q_4 -fun 找到全局最优的成功率; No: 表示测试函数序号; n: 表示测试函数的维数; nFf: 表示第一次得到全局极值时目标函数的平均计算次数; nFq: 表示第一次得到全局极值时填充函数的平均计算次数; failed: 表示找不到全局最优点, 表中数字精确位数均为整数。

表2 测试函数

函数	x_i
$f(y) = \frac{\pi}{n} \left\{ \begin{aligned} &10\sin^2(\pi y_1) + y_1 = 1 + (x_{i-1})/4 \\ &\sum_{i=1}^{n-1} (y_i - 1)^2 [1 + 10\sin^2(\pi y_{i+1})] + (y_n - 1)^2 \end{aligned} \right\}$	$[-10, 10]$
$f(x) = \frac{\pi}{n} \left\{ \begin{aligned} &10\sin^2(\pi x_1) + \\ &\sum_{i=1}^{n-1} (x_i - 1)^2 [1 + 10\sin^2(\pi x_{i+1})] + (x_n - 1)^2 \end{aligned} \right\}$	$[-10, 10]$
$f(x) = [1 + (x_1 + x_2 + 1)^2 (19 - 14x_1 + 3x_1^2 - 14x_2 + 6x_1x_2 + 3x_2^2)]$	$[-2, 2]$
$f(x) = [30 + (2x_1 - 3x_2)^2 (18 - 32x_1 + 12x_1^2 + 48x_2 - 36x_1x_2 + 27x_2^2)]$	$[-10, 10]$
$f(x) = (x_2 - 5.1x_1^2/4\pi^2 + 5x_1/\pi - 6)^2 + 10(1 - 1/8\pi)\cos x_1 + 10$	$x_1 \in [-5, 10]$ $x_2 \in [0, 15]$
$f(x) = x_1^2 - 2x_1 + 10\cos(\pi x_1/2) + 8\sin(\pi x_1/2) - ((1/\sqrt{5})\exp(-(x_2 - 0.5)^2/2)) + 11$	$x_1 \in [-30, 30]$ $x_2 \in [-10, 10]$
$f(x) = 100(x_1^2 - x_2)^2 + (x_1 - 1)^2$	$[-50, 50]$
$f(x) = 10^n x_1^2 + x_2^2 - (x_1^2 + x_2^2)^2 + 10^m (x_1^2 + x_2^2)^4$	m, n, i : 自然数
$f(x) = \sum_{i=1}^{n-1} \{ (x_i - 1)^2 + 100(x_i^2 - x_{i+1}^2) \}$	$[-10^3, 10^3]$
$f(x) = 100(x_1^2 - x_2)^2 + (x_1 - 1)^2 + (x_3 - 1)^2 + 90(x_2^2 - x_4^2)^2 + 10.1((x_2 - 1)^2 + (x_4 - 1)^2) + 19.8(x_2 - 1)(x_4 - 1)$	$[-10, 10]$

从图1可以看出, Z -fun 比 Q_3 -fun 和 Q_4 -fun 对同一个问题的寻优成功率都要高(除 no7 外)。

由表1中的数字可以看出: Z -fun 比 Q_1 -fun 和 Q_2 -fun 在目标函数和填充函数运算次数上都要少, 从而提高了计算速度。

结束语 填充函数法是一种求解连续可微函数总体最优的有效方法。不同的填充函数对算法的计算精度和收敛速度影响很大。从理论分析和测试实例看, 本文提出的填充函数, 较好地解决了以往提出的填充函数存在的问题。另外, 在数值试验中, 本文提出的避免指数因子不利影响的措施同样可以应用到文献[2]中的填充函数。

参考文献

- [1] Ge R. 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, Qin Y. A glass of filled functions for finding global minimizers of several variables[J]. Journal of Optimization Theory and Applications, 1987, 54: 241-252
- [3] Ge R, Qin Y. The globally convexized filled functions for global optimization[J]. Applied Mathematics and Computation, 1990, 35: 131-158
- [4] Ge R. The theory of the filled function method for finding a global minimizer of a nonlinear constrained minimization problem[J]. Journal of Computational Mathematics, 1987, 5(1): 1-9
- [5] Goldstein A A, Price J F. On descent from a local minimum[J]. Mathematics of Computation, 1977, 25: 569-574
- [6] Liu Xian. Finding global minima with a computable filled function[J]. Journal of Global Optimization, 2001, 19: 151-161
- [7] Wolfe M A. Numerical Methods for Unconstrained Optimization [M]. New York: Van Nostrand Reinhold Company, 1978
- [8] Lucidi S, Piccioni M. Random tunneling by means of acceptance-rejection sampling for global optimization[J]. Journal of Optimization Theory and Applications, 1985, 62: 255-278
- [9] 粟塔山. 最优化计算原理与算法程序设计[M]. 长沙, 2001
- [10] 郑应根. 求总体极小值的一类填充函数[J]. 数学物理学报, 1994, 14(2): 184-189
- [11] 高娟, 徐裕, 等. 一类等式约束下的凸二次规划的共轭轨梯度算法[J]. 重庆工学院学报: 自然科学版, 2009, 23(3): 101-104