文章编号: 1000_0887(2006) 07_0874_09

一种求约束总极值的水平值估计方法*

邬冬华[1,2], 俞武扬[1], 田蔚文[1], 张连生[1]

(1. 上海大学 数学系, 上海 200444;
2. 南京大学 数学系, 南京 210093)

(刘曾荣推荐)

摘要: 给出了一种求约束总极值的水平值估计方法, 说明了修正的方差方程的根与原始问题的最优值之间的等价性, 给出了一种基于牛顿法的水平值估计算法并证明了实现算法的收敛性• 初步的计算例子表明所给算法是有效的•

关 键 词: 约束总极值; 水平值估计; 一致分布

中图分类号: O221.2 文献标识码: A

引 言

考虑如下有约束的全局最优化问题

$$(\mathrm{P}) \quad \min_{x \in D} f(x),$$

其中 D 是一个有界闭箱, f 是 D 上的连续函数•

积分方法[1,2]和值估计方法[3]是作为解全局最优化问题提出来的• 这些工作中, 郑权等[1,2]发展了一种理论算法, 该算法能够产生一列下降的序列 $\{\alpha_k\}$, 此序列是在水平集 $\{H_k\}$ 上的积分, 序列 $\{\alpha_k\}$ 是收敛的而且 $\{H_k\}$ 收敛到 H_0• Phu 和 Hoffmann[3]给出了计算可求和函数的本质上确界与寻找一个一维凸函数的根之间的等价性, 并且提出了一种使算法加速的方法• 所有这些算法都是用 Monte Carlo 模型实现的, 其实现算法的收敛性至今尚未解决•

本文将首先介绍数论中的一致分布概念, 并且提出一种修正的积分水平集方法• 然后我们提出水平值估计算法和它的一些主要性质• 最后我们建立了一个求解修正的方差方程根的算法, 并且证明了所给实现算法的收敛性• 初步的计算例子表明所给算法是有效的•

1 数论中的一致分布与修正的积分水平集方法

不失一般性, 不妨设 D 是一个单位闭箱, 以下均认为 $D = G_n$, $G_n = [0, 1]^n$•

若(σ)为 G_n 的任意分割

* 收稿日期: 2005_01_25; 修订日期: 2006_04_11
基金项目: 国家自然科学基金资助项目(10871053)
作者简介: 邬冬华(1963—), 男, 上海人, 教授, 博士(联系人. Tel: + 86_21_66132413(o); + 86_21_69190204(h); E_mail: dhwu@staff.shu.edu.cn)•

$$
(\sigma)\quad \begin{cases} 0 = x_1^0 < x_1^1 < \cdots < x_1^{l_1} = 1 \\ 0 = x_2^0 < x_2^1 < \cdots < x_2^{l_2} = 1 \\ \vdots \\ 0 = x_n^0 < x_n^1 < \cdots < x_n^{l_n} = 1 \end{cases} \tag{1}
$$

设 $f(x)$ 在 G_n 上有定义, 且

$$
\Delta_{k_1} f(x_1, \cdots, x_{k_1}^{i_{k_1}}, \cdots, x_n) = f(x_1, \cdots, x_{k_1}^{i_{k_1}}, \cdots, x_n) - f(x_1, \cdots, x_{k_1}^{i_{k_1}+1}, \cdots, x_n),
$$

$$
1 \leqslant k_1 \leqslant n,
$$

$$
\begin{aligned}
\Delta_{k_1, k_2} f(x_1, \cdots, x_{k_1}^{i_{k_1}}, \cdots, x_{k_2}^{i_{k_2}}, \cdots, x_n) = {} & \\
f(x_1, \cdots, x_{k_1}^{i_{k_1}}, \cdots, x_{k_2}^{i_{k_2}}, \cdots, x_n) - f(x_1, \cdots, x_{k_1}^{i_{k_1}+1}, \cdots, x_{k_2}^{i_{k_2}}, \cdots, x_n) - {} & \\
f(x_1, \cdots, x_{k_1}^{i_{k_1}}, \cdots, x_{k_2}^{i_{k_2}+1}, \cdots, x_n) + f(x_1, \cdots, x_{k_1}^{i_{k_1}+1}, \cdots, x_{k_2}^{i_{k_2}+1}, \cdots, x_n), & \\
1 \leqslant k_1 \leqslant k_2 \leqslant n, &
\end{aligned}
$$

$$\vdots$$

$$
\begin{aligned}
\Delta_{k_1, k_2, \cdots, k_m} f(x_1, \cdots, x_{k_1}^{i_{k_1}}, \cdots, x_{k_2}^{i_{k_2}}, \cdots, x_{k_m}^{i_{k_m}}, \cdots, x_n) = {} & \\
f(x_1, \cdots, x_{k_1}^{i_{k_1}}, \cdots, x_{k_2}^{i_{k_2}}, \cdots, x_{k_m}^{i_{k_m}}, \cdots, x_n) - f(x_1, \cdots, x_{k_1}^{i_{k_1}+1}, \cdots, x_{k_2}^{i_{k_2}}, \cdots, x_{k_m}^{i_{k_m}}, \cdots, x_n) - {} & \\
f(x_1, \cdots, x_{k_1}^{i_{k_1}}, \cdots, x_{k_2}^{i_{k_2}+1}, \cdots, x_{k_m}^{i_{k_m}}, \cdots, x_n) + {} & \\
f(x_1, \cdots, x_{k_1}^{i_{k_1}+1}, \cdots, x_{k_2}^{i_{k_2}+1}, \cdots, x_{k_m}^{i_{k_m}}, \cdots, x_n) + \cdots + {} & \\
(-1)^{m-1} f(x_1, \cdots, x_{k_1}^{i_{k_1}}, \cdots, x_{k_2}^{i_{k_2}+1}, \cdots, x_{k_m}^{i_{k_m}+1}, \cdots, x_n) + {} & \\
(-1)^m f(x_1, \cdots, x_{k_1}^{i_{k_1}+1}, \cdots, x_{k_2}^{i_{k_2}+1}, \cdots, x_{k_m}^{i_{k_m}+1}, \cdots, x_n), & \\
1 \leqslant k_1 < k_2 < \cdots < k_m \leqslant n, &
\end{aligned}
$$

$$
\begin{aligned}
\Delta_{1, 2, \cdots, n} f(x_1^{i_1}, x_2^{i_2}, \cdots, x_n^{i_n}) = f(x_1^{i_1}, x_2^{i_2}, \cdots, x_n^{i_n}) - f(x_1^{i_1+1}, x_2^{i_2}, \cdots, x_n^{i_n}) - {} & \\
f(x_1^{i_1}, x_2^{i_2+1}, \cdots, x_n^{i_n}) + f(x_1^{i_1+1}, x_2^{i_2+1}, \cdots, x_n^{i_n}) + \cdots + {} & \\
(-1)^{n-1} f(x_1^{i_1}, x_2^{i_2+1}, \cdots, x_n^{i_n+1}) + (-1)^n f(x_1^{i_1+1}, x_2^{i_2+1}, \cdots, x_n^{i_n+1}), &
\end{aligned}
$$

设

$$
\begin{aligned}
V_\sigma = {} & \sum_{i_1=0}^{l_1-1} \sum_{i_2=0}^{l_2-1} \cdots \sum_{i_n=0}^{l_n-1} \left| \Delta_{1, 2, \cdots, n} f(x_1^{i_1}, x_2^{i_2}, \cdots, x_n^{i_n}) \right| + \cdots + \\
& \sum_{1 \leqslant k_1 < k_2 < \cdots < k_m \leqslant n} \sum_{i_{k_1}=0}^{l_{k_1}-1} \sum_{i_{k_2}=0}^{l_{k_2}-1} \cdots \sum_{i_{k_m}=0}^{l_{k_m}-1} \Big| \Delta_{k_1, k_2, \cdots, k_m} \times \\
& f(1, \cdots, 1, x_{k_1}^{i_{k_1}}, 1, \cdots, 1, x_{k_2}^{i_{k_2}}, \cdots, x_{k_m}^{i_{k_m}}, 1, \cdots, 1) \Big| + \cdots + \\
& \sum_{1 \leqslant k_1 \leqslant k_2 \leqslant n} \sum_{i_{k_1}=0}^{l_{k_1}-1} \sum_{i_{k_2}=0}^{l_{k_2}-1} \left| \Delta_{k_1, k_2} f(1, \cdots, 1, x_{k_1}^{i_{k_1}}, 1, \cdots, 1, x_{k_2}^{i_{k_2}}, 1, \cdots, 1) \right| + \\
& \sum_{1 \leqslant k_1 \leqslant n} \sum_{i_{k_1}=0}^{l_{k_1}-1} \left| \Delta_{k_1} f(1, \cdots, 1, x_{k_1}^{i_{k_1}}, 1, \cdots, 1) \right| .
\end{aligned} \tag{2}
$$

定义 1.1　若变差 V_σ 有一个与分割(σ)无关的上界, 则称 $f(x)$ 为在 Hardy_Kraus 意义下的

有界变差函数。V_0 的上确界称为$f(x)$ 的全变差，记为 $V(f)$，其全体记为 B_n。

定义 1.2 若

$$0 \leqslant x_k^{i_k} < x_k^{i_k+1} \leqslant 1, \qquad k = 1, 2, \cdots, n,$$

且存在常数 L 使得

$$\begin{cases}
1)\ \left|\Delta_{k_1} f(1, \cdots, 1, x_{k_1}^{i_{k_1}}, 1, \cdots, 1)\right| \leqslant L(x_{k_1}^{i_{k_1}+1} - x_{k_1}^{i_{k_1}}), \qquad 1 \leqslant k_1 \leqslant n, \\
2)\ \left|\Delta_{k_1, k_2} f(1, \cdots, 1, x_{k_1}^{i_{k_1}}, 1, \cdots, 1, x_{k_2}^{i_{k_2}}, 1, \cdots, 1)\right| \leqslant L(x_{k_1}^{i_{k_1}+1} - x_{k_1}^{i_{k_1}})(x_{k_2}^{i_{k_2}+1} - x_{k_2}^{i_{k_2}}), \\
\qquad\qquad 1 \leqslant k_1 \leqslant k_2 \leqslant n, \\
\vdots \\
m)\ \Delta_{k_1, k_2, \cdots, k_m} f(1, \cdots, 1, x_{k_1}^{i_{k_1}}, 1, \cdots, 1, x_{k_2}^{i_{k_2}}, \cdots, x_{k_m}^{i_{k_m}}, 1, \cdots, 1)\,| \leqslant \\
\qquad L(x_{k_1}^{i_{k_1}+1} - x_{k_1}^{i_{k_1}})(x_{k_2}^{i_{k_2}+1} - x_{k_2}^{i_{k_2}}) \cdots (x_{k_m}^{i_{k_m}+1} - x_{k_m}^{i_{k_m}}), \\
\qquad\qquad 1 \leqslant k_1 \leqslant k_2 < \cdots < k_m \leqslant n, \\
\vdots \\
n)\ \Delta_{1, 2, \cdots, n} f(x_1^{i_1}, x_n^{i_2}, \cdots, x_n^{i_n})\,| \leqslant L(x_1^{i_1+1} - x_1^{i_1})(x_2^{i_2+1} - x_2^{i_2}) \cdots (x_n^{i_1+1} - x_n^{i_n}),
\end{cases} \tag{3}$$

则称$f(x)$ 满足广义局部 Lipschitz 条件，此种函数全体记为 L_n。 事实上 $L_n \subset B_n$。

若$f(x)$ 满足

$$|f(x)| \leqslant L, \quad \left|\frac{\partial f(x)}{\partial x_i}\right| \leqslant L, \qquad 1 \leqslant i \leqslant n,$$

$$\left|\frac{\partial^2 f(x)}{\partial x_i \partial x_j}\right| \leqslant L, \ 1 \leqslant i \leqslant j \leqslant n, \cdots, \qquad \left|\frac{\partial^n f(x)}{\partial x_1 \partial x_2 \cdots \partial x_n}\right| \leqslant L,$$

则$f(x) \in L_n$。 显然 B_n 是一类范围很广的函数类。

对任意整数 l，$P_l(k) = (x_1^{(l)}(k), \cdots, x_n^{(l)}(k))$，$1 \leqslant k \leqslant l$ 为 G_n 中的一个点集。

定义 1.3 对于任意 $\gamma = (\gamma_1, \gamma_2, \cdots, \gamma_n) \in G_n$，设 $N_l(\gamma) = N_l(\gamma_1, \cdots, \gamma_n)$ 表示 $P_l(k)(1 \leqslant k \leqslant l)$ 中适合不等式 $0 \leqslant x_1^{(l)}(k) < \gamma_1, \cdots, 0 \leqslant x_n^{(l)}(k) < \gamma_n$ 的个数。 若

$$\sup_{\gamma \in G_n} \left|\frac{N_l(\gamma)}{l} - |\gamma|\right| = \varphi(l), \qquad |\gamma| = \gamma_1, \cdots, \gamma_n, \tag{4}$$

则称点集 $P_l(k)(1 \leqslant k \leqslant l)$ 有偏差 $\varphi(l)$。

定义 1.4 设$\{l_s\}$为一整数序列满足 $1 < l_1 < l_2 < \cdots$，对于每一个 l_s，有 G_n 中的一个点集$P_{l_s}(k)(1 \leqslant k \leqslant l_s)$ 与之对应，且有偏差 $\varphi(l_s)$。 若 $\varphi(l_s) = o(1)$，则称点集 $P_{l_s}(k)(l_1 < l_2 < \cdots)$ 为一致分布且有偏差 $\varphi(l_s)$。

定理 1.1 若$f \in B_n$ 且$P_l(k)(1 \leqslant k \leqslant l)$ 为存在偏差 $\varphi(l)$ 的点集，则

$$\left|\int_{G_n} f(x)\,\mathrm{d}\mu - \frac{1}{l}\sum_{k=1}^{l} f(P_l(k))\right| \leqslant V(f)\,\varphi(l). \tag{5}$$

证明 参见文献[4]，p103。

定义 1.5 设 $\gamma \in G_n$，若形为

$$P(k) = (\{\gamma_1(k)\}, \cdots, \{\gamma_n(k)\}), \qquad k = 1, 2, \cdots, l \tag{6}$$

的点集存在不大于 $c(\gamma, \epsilon)\, l^{-1+\epsilon}$ 的偏差，则称点集(4) 为佳点集，而 γ 称为佳点。

华罗庚、王元[4]介绍了如何构造一致分布佳点集列并且得到了以下定理

定理 1.2　设 $P(k) = (\{\gamma_1(k)\}, \ldots, \{\gamma_n(k)\}), 1 \leqslant k \leqslant l$ 为存在偏差 $\varphi(l) = c(\gamma, \epsilon) l^{-1+\epsilon}$ 的点集, 则

$$\left| \int_{G_n} f(x)\mathrm{d}\mu - \frac{1}{l}\sum_{k=1}^{l} f(\{\gamma_1(k)\}, \ldots, \{\gamma_n(k)\}) \right| \leqslant V(f)\, c(\gamma, \epsilon)\, l^{-1+\epsilon}. \tag{7}$$

为了提高实现算法的可操作性及数值计算效率, 我们提出了修正的积分水平集方法, 参见文献[5]和文献[6]. 现表述如下

步1　给定一个充分小的正数 ϵ_0, 取 $x_0 \in D$. 设

$$c_0 = f(x_0), H_{c_0} = \{x \mid x \in D, f(x) \leqslant c_0\}, k := 0.$$

步2　如果 $\mu(H_{c_k}) = 0$, 那么 c_k 是全局最优值, H_{c_k} 是全局最优值点集, 终止; 否则转步 3.

步3　构造 $f_{c_k}(x)$, 满足

$$f_{c_k}(x) = \begin{cases} c_k, & x \in D \setminus H_{c_k}, \\ f(x), & x \in H_{c_k}, \end{cases}$$

计算均值 $c_{k+1} = \dfrac{1}{\mu(D)}\displaystyle\int_D f_{c_k}(x)\mathrm{d}\mu$. 令

$$H_{c_{k+1}} = \{x \mid x \in D, f_{c_k}(x) \leqslant c_{k+1}\} = \{x \mid x \in D, f(x) \leqslant c_{k+1}\}.$$

若 $c_{k+1} = c_k$, 则 c_{k+1} 是全局最优值, $H_{c_{k+1}}$ 是全局最优值点集, 终止; 否则转步 4.

步4　计算均方差

$$V_F = \frac{1}{\mu(D)}\int_D (f_{c_k}(x) - c_k)^2 \mathrm{d}\mu.$$

如果 $V_F \geqslant \epsilon$, 令 $k := k+1$, 转步 3; 否则转步 5.

步 5　令 $f^* = c_{k+1}$ 且 $H^* = H_{c_{k+1}}$, f^* 为近似全局最优值, H^* 是近似全局最优值点集, 终止.

从步 4 我们可以发现方程 $V_F = 0$ 的根是原始问题(P)的最优值, 在下一节我们将给出一种求全局最优值的水平值估计方法.

2　水平值估计方法与若干性质

本节中, 我们引进函数 $V_f: R \to R$ 及函数 $M_f: R \to R$

$$V_f(c) = \int_D (c - f_c(x))^2 \mathrm{d}\mu, \quad M_f(c) = \int_D (c - f_c(x))\mathrm{d}\mu, \tag{8}$$

其中

$$f_c(x) = \begin{cases} c, & x \in D \setminus H_c, \\ f(x), & x \in H_c, \end{cases} \tag{9}$$

关于 V_f 与 M_f, 我们证明了一些值得关注的性质.

定理 2.1　函数 $V_f(c)$ 和 $M_f(c)$ 是非负, 单调不减的凸函数, 有一阶与二阶导数为

$$1)\ V_f'(c) = 2\int_D (c - f_c(x))\mathrm{d}\mu, \quad 2)\ V_f''(c) = 2\int_D \chi_{H_c}(x)\mathrm{d}\mu, \tag{10}$$

其中

$$\chi_{H_c}(x) = \begin{cases} 1, & x \in H_c, \\ 0, & x \in D \setminus H_c. \end{cases}$$

证明 由 $V_f(c)$ 和 $M_f(c)$ 的定义根据导数定义容易证得本定理•

令 $c^* = \min_D f(x)$, 利用函数 $V_f(c)$ 和 $M_f(c)$ 我们可以用另外的方法来刻划 c^*•

定理 2.2 设 $f(x)$ 是 D 上的连续函数, 则 c^* 是 $f(x)$ 在 D 上的全局最优值当且仅当

$$V_f(c^*) = 0 \text{ 或 } M_f(c^*) = 0, \tag{11}$$

其中 $H_{c^*} = \{x \mid f(x) \leqslant c^*, x \in D\} \neq \varnothing$, D 是一个紧集•

证明(充分性) 由 $f(x)$ 在 D 上的连续性可知 $f_c(x)$ 在 D 上连续• 根据定理 2.2 的条件, 我们有

$$(c^* - f_{c^*}(x))^2 = 0 \text{ 或者 } c^* - f_{c^*}(x) = 0, \tag{12}$$

这是因为 $c^* - f_{c^*}(x) \geqslant 0$, $\forall x \in D$•

因此 $c^* - f_{c^*}(x) = 0$ 对于所有 $x \in D$ 都成立• 由 $f_{c^*}(x)$ 的定义, 存在 $c^* \in R$ 使得

$$c^* - f_{c^*}(x) = \begin{cases} 0, & x \in D \setminus H_{c^*}, \\ c^* - f(x), & x \in H_{c^*}• \end{cases} \tag{13}$$

从(13)可以推出

$$f(x) = c^*, \qquad \forall x \in H_{c^*}• \tag{14}$$

进一步有集合 $H_{c^*} = \{x \mid f(x) \leqslant c^*, x \in D\}$ 与 $\{x \mid f(x) = c^*, x \in D\}$ 等价• 即 c^* 为 D 上的全局最优值•

(必要性) 设 c^* 是 $f(x)$ 在 D 上的全局最优值, 则

$$H_{c^*} = \{x \mid f(x) \leqslant c^*, x \in D\} \neq \varnothing• \tag{15}$$

以下证明 $\mathrm{int}H_{c^*}$ 为空集, 事实上, 若 $\mathrm{int}H_{c^*} \neq \varnothing$, 则存在点 $x_0 \in \mathrm{int}H_{c^*}$, 使得 $f(x_0) < c^*$ 与定理 2.2 的条件矛盾• 因此, $H_{c^*} = \partial H_{c^*} = \{x \mid f(x) = c^*, x \in D\}$•

根据(11)可得 $c^* - f_{c^*} = 0$, $\forall x \in D$• 因此 $V_f(c^*) = 0$ 或 $M_f(c^*) = 0$•

定理 2.3 若 c^* 是问题(P)的全局最优值, 则 $V_f(c)$ 在 $[c^*, +\infty)$ 上严调单调增加•

证明 由定理 2.2 可知, 我们只需证明对于 $c > c^*$ 是有 $V_f'(c) > 0$• 反证, 假设当 $c > c^*$ 时有 $V_f'(c) = 0$• 根据水平集的定义, 有 $H_{c^*} \subset H_c$•

令 $c = (c + c^*)/2$, 可得 $H_{c^*} \subset H_c \subset H_c$• 进而由 $f(x)$ 的连续性可知, 存在一点 $x \in H_c$ 使得

$$f(x) = (c + c^*)/2, \quad x \in \mathrm{int}H• \tag{16}$$

另一方面, 对于任意的 ϵ, 存在邻域 $U_\epsilon(x)$ 使得

$$|f(x) - f(c)| = |f(x) - (c + c^*)/2| < \epsilon, \qquad \forall x \in U_\epsilon(x)• \tag{17}$$

取 $\epsilon = (c + c^*)/4$, 则有

$$3(c - c^*)/4 > c - f(x) > (c - c^*)/4, \qquad \forall x \in U_{(c-c^*)/4}(x)• \tag{18}$$

基于 $V_f(c)$ 的定义以及积分的性质, 我们有

$$V_f(c) = \int_{H_c} (c - f(x))^2 \mathrm{d}\mu \geqslant \int_{U_{(c-c^*)/4}} (c - f(x))^2 \mathrm{d}\mu \geqslant$$
$$(c - c^*)^2 \mu(U_{(c-c^*)/4})/16 > 0, \tag{19}$$

这与假设矛盾•

下面我们给出 $V_f(c) = 0$ 的最大根与全局最优值之间的等价性关系•

定理 2.4 若 c^* 为(P)的全局最优值, 则

$$c^* = \max\{c \mid V_f(c) = 0\}. \tag{20}$$

证明 因为 c^* 为(P)的全局最优值，则定理2.2可知 $V_f(c^*) = 0$. 进而由定理2.3可得

$$V_f(c) \geqslant V_f(c^*),\ 当\ c \geqslant c^*\ 时, \tag{21}$$

所以 $V_f(c) \geqslant 0$，当 $c \geqslant c^*$ 时.

由 $V_f(c)$ 在 $(-\infty, +\infty)$ 上的定义及 $V_f(c)$ 在 $(c^*, +\infty)$ 上是连续增加的，我们有

$$\sup\{c \mid V_f(c) = 0\} \leqslant c^*. \tag{22}$$

由于 $V_f(c^*) = 0$，从而 $c^* \in \sup\{c \mid V_f(c) = 0\}$. 所以有

$$c^* = \sup\{\alpha \mid V_f(\alpha) = 0\} = \max\{\alpha \mid V_f(\alpha) = 0\}. \tag{23}$$

3 水平值估计算法及其收敛性

利用以上性质，我们建立了如下基于牛顿法[7]的水平值估计算法

算法 A

步0 选择 $\epsilon > 0$ 以及 $c_0 \in R$ 满足 $M_f(c_0) > 0$，令 $k := 0$.

步1 计算

$$\alpha_k = \frac{V_f(c_k)}{M_f(c_k)}. \tag{24}$$

步2 若 $\alpha_k < \epsilon$，则转步3；否则令 $c_{k+1} = c_k - \alpha_k$，$k := k + 1$ 转步1.

步3 令 $c_k = \min_{x \in D} f(x)$，终止.

显然算法 A 在每一次迭代过程中都需要计算积分 $V_f(c)$ 和 $M_f(c)$. 当然 V_f 与 M_f 都是近似结果. 我们给出了一种基于数论中的一致分布技术的实现方法以近似计算 V_f 与 M_f. 从而建立了如下确定性的实现算法

算法 B

步0 选择 $\delta > 0$ 及 $x_0 \in D$，令 $\hat{c}_0 = f(x_0)$，$H_{\hat{c}_0} = \{x \mid x \in D, f(x) \leqslant \hat{c}_0\}$，$k := 0$.

步1 构造 $f_{\hat{c}_k}(x)$ 满足

$$f_{\hat{c}_k}(x) = \min(f(x), \hat{c}_k) = \begin{cases} \hat{c}_k, & x \in D \setminus H_{\hat{c}_k}, \\ f(x), & x \in H_{\hat{c}_k}. \end{cases} \tag{25}$$

利用一致分布数值积分计算水平

$$V_f(k) = \frac{1}{l_k}\sum_{s=1}^{l_k}(\hat{c}_{k-1} - f_{\hat{c}_{k-1}}(P(s)))^2,\quad M_f(k) = \frac{1}{l_k}\sum_{s=1}^{l_k}(\hat{c}_{k-1} - f_{\hat{c}_{k-1}}(P(s))). \tag{26}$$

其中序列 $\{l_k\}$ 是严格增加的.

设离散水平集为

$$\hat{H}_{\hat{c}_k} = \{x \mid x = P(s) \in D, f(P(s)) \leqslant \hat{c}_k,\ s = 1, 2, \cdots, l_k;\ x \in \hat{H}_{\hat{c}_{k-1}}, f(x) \leqslant \hat{c}_k\} \tag{27}$$

以及水平集为 $H_{\hat{c}_k} = \{x \mid x \in D, f(x) \leqslant \hat{c}_k\}$. 这里 $\{P(s)\}_{s=1}^{l_k}$ 为一致分布佳格子点集.

步2 若 $M_f(k) \leqslant \delta$，转步3；否则计算

$$\hat{\alpha}_k = \frac{V_f(k)}{M_f(k)}. \tag{28}$$

令 $\hat{c}_k = \hat{c}_{k-1} - \hat{\alpha}_k$，转1.

步3 若 $l_k \geqslant M(|f|)/\delta$，则转步4；否则令 $l_k = [M(|f|)/\delta] + 1$，转步1，其中 $M(|f|)$

是 $|f(x)|$ 在 D 中的一个上界•

步4 令 $c^* = \hat{c}_k$ 及 $H = H_{\hat{c}_k}$, H 是一个近似的全局最优点集, 终止•

定理 3.1 设 $f(x)$ 在有界闭箱 D 上连续• c^* 是 $f(x)$ 在 D 上的全局最优值, 且 $\{c_k\}$ 是由算法 A 产生的序列, 则

$$\lim_{k\to\infty} c_k = c^*•$$

证明 设算法 A 迭代无穷多次, 则由算法 A 产生的序列 $\{c_k\}$ 是单调减的• 它必定满足以下情况之一: 1) $\lim_{k\to\infty} c_k = -\infty$, 2) $\lim_{k\to\infty} c_k = c$•

首先, 我们说明情况 1) 不成立• 否则, 假设 $\lim_{k\to\infty} c_k = -\infty$, 则存在一个正整数 k 使得 $c_k < c^*$• 根据算法 A 以及函数 $V_f(c)$ 与 $M_f(c)$ 的定义, 我们有

$$V_f(c) = 0,\quad M_f(c) = 0, \tag{29}$$

表示算法 A 在有限步迭代后终止, 与假设矛盾• 因此只有情况 2) 成立且有 $c_k \geqslant c^*$ $(k = 1, 2, \dots)$•

其次, 我们证明 $c = c^*$• 否则, 假设 $c > c^*$, 则有 $V_f(c) > V_f(c^*) = 0$, $M_f(c) > M_f(c^*) = 0$, 即 $V_f(c)/M_f(c) > 0$, 因为 $V_f(c)$ 与 $M_f(c)$ 在 $(c^*, +\infty)$ 上严格单调增加• 由于 $\lim_{k\to\infty} c_k = c$, 根据算法 A, 我们有

$$\lim_{k\to\infty} \frac{V_f(c_k)}{M_f(c_k)} = 0• \tag{30}$$

然而, 因为 $V_f(c)$ 与 $M_f(c)$ 在实数轴上是连续的, 所以

$$\lim_{k\to\infty} \frac{V_f(c_k)}{M_f(c_k)} = \frac{V_f(c)}{M_f(c)} > 0, \tag{31}$$

与 (30) 矛盾, 从而 $c = c^*$•

定理 3.2 设 $f(x)$ 是有界闭箱 D 上的一个 Lipschitz 函数, c^* 是 $f(x)$ 在 D 上的全局最优值, 且 $\{\hat{c}_k\}$ 是算法 B 产生的序列, 则有

$$\lim_{k\to\infty} \hat{c}_k = c^*•$$

证明 设算法 B 迭代无穷多次, 则由算法 B 产生的序列 $\{\hat{c}_k\}$ 是单调减的• 它必定满足以下情形之一: 1) $\lim_{k\to\infty} \hat{c}_k = -\infty$, 2) $\lim_{k\to\infty} \hat{c}_k = \hat{c}$•

首先, 我们说明情况 1) 不成立• 否则, 假设 $\lim_{k\to\infty} c_k = -\infty$, 则存在一个正整数 k 使得 $\hat{c}_k < c^*$• 根据算法 A 以及函数 $V_f(c)$ 与 $M_f(c)$ 的定义, 我们有 $V_f(\hat{c}_k) = 0$, $M_f(\hat{c}_k) = 0$, 表示算法 B 在有限步迭代后终止, 与假设矛盾• 因此只有情况 2) 成立, 且 $\hat{c}_k \geqslant c^*$ $(k = 1, 2, \dots)$•

其次, 我们证明 $\hat{c} = c^*$• 由算法 B 及 $\lim_{k\to\infty} \hat{c}_k = \hat{c}$ 可知

$$\lim_{k\to\infty} \frac{V_f(\hat{c}_k)}{M_f(\hat{c}_k)} = 0• \tag{32}$$

反证, 假设 $\hat{c} > c^*$• 因为 $M_f(c)$ 与 $V_f(c)$ 关于 $c = \hat{c}$ 连续, 且 $\lim_{k\to\infty} \hat{c}_k = \hat{c}$, 对于给定的 $M_f(\hat{c})$ 与 $V_f(\hat{c})$, 存在一个充分大的正整数 K, 当 $k > K$ 时, 有 $|M_f(\hat{c}_k) - M_f(\hat{c})| < M_f(\hat{c})$ 及 $|V_f(\hat{c}_k) - V_f(\hat{c})| < V_f(\hat{c})/2$, 即当 $k > K$ 时, 有 $M_f(\hat{c}_k) < M_f(\hat{c}_k)$ 及 $V_f(\hat{c}_k) > V_f(\hat{c}_k)/2$• 取 $\delta < \min\left\{(M_f(\hat{c})/V_1(f))^{1/(1-\varepsilon_1)}, (V_f(\hat{c})/4V_2(f))^{1/(1-\varepsilon_2)}\right\}$, 并且 $l_k > M(f)/\delta$, 由定理 1.2, 可得

$$|M(\hat{c}_k) - M_f(\hat{c}_k)| < V_1(f)\, l_k^{-1+\varepsilon_1} \leqslant V_1(f)\left(\frac{\delta}{M(f)}\right)^{1-\varepsilon_1}, \tag{33}$$

从而

$$M(\hat{c}_k) \leqslant M_f(\hat{c}_k) + V_1(f)\left(\frac{\delta}{M(f)}\right)^{1-\varepsilon_1} \leqslant$$
$$2M_f(\hat{c}) + V_1(f)\left(\frac{\delta}{M(f)}\right)^{1-\varepsilon_1} \leqslant 3M_f(\hat{c}), \tag{34}$$

类似地，由定理 1.2 可得 $|V(\hat{c}_k) - V_f(\hat{c}_k)| < V_2(f)\, l_k^{-1+\varepsilon_2} \leqslant V_2(f)\left((\delta/M(f))\right)^{1-\varepsilon_2}$，从而

$$V(\hat{c}_k) \geqslant V_f(\hat{c}_k) - V_2(f)\left(\frac{\delta}{M(f)}\right)^{1-\varepsilon_2} \geqslant$$
$$\frac{1}{2}V_f(\hat{c}) - V_2(f)\left(\frac{\delta}{M(f)}\right)^{1-\varepsilon_2} \geqslant \frac{1}{4}V_f(\hat{c})\cdot \tag{35}$$

由于 $V_f(c)$ 与 $M_f(c)$ 在 $(c^*, +\infty)$ 上严格单调增加，因此 $V_f(\hat{c}) > V_f(c^*) = 0$，$M_f(\hat{c}) > M_f(c^*) = 0$，但是

$$\lim_{k\to\infty}\frac{V_f(\hat{c}_k)}{M_f(\hat{c}_k)} \geqslant \lim_{k\to\infty}\frac{(1/4)V_f(\hat{c}_k)}{3M_f(\hat{c}_k)} = \frac{V_f(\hat{c})}{12M_f(\hat{c})} > 0\cdot \tag{36}$$

这与(32)矛盾，所以 $\hat{c} = c^*$•

4　数值结果

利用算法 B 我们对两个例子作了数值测试•

例 1　问题如下

$$\min f(x) = 100(x_2 - x_1^2)^2 + (x_1 - 1)^2,$$
$$\text{s.t.}\ -1.5 \leqslant x_1 \leqslant 1.5,\ -0.5 \leqslant x_2 \leqslant 2.5,$$

最小值为 $c^* = 0$•　利用算法 B 迭代 17 步得到表 1 所得结果•

表 1　例 1 计算结果

k	$\hat{c}_k$	α_k
1	1.645 27	4.854 729
2	0.242 472 2	1.402 799
3	0.028 753 25	0.213 718 9
4	0.003 585 828	0.025 167 42
5	0.000 968 027 9	0.002 617 801
6	0.000 302 347 3	0.000 665 680 6
7	0.000 104 384 9	0.000 197 962 3
8	$3.321\ 568\times10^{-5}$	$7.116\ 927\times10^{-5}$
9	$1.398\ 668\times10^{-5}$	$1.922\ 900\times10^{-5}$
10	$3.647\ 967\times10^{-6}$	$1.033\ 871\times10^{-5}$
11	$1.370\ 875\times10^{-6}$	$2.277\ 092\times10^{-6}$
12	$3.279\ 73\times10^{-7}$	$1.042\ 902\times10^{-6}$
13	$6.625\ 285\times10^{-8}$	$2.617\ 201\times10^{-7}$
14	$1.424\ 192\times10^{-8}$	$5.201\ 094\times10^{-8}$
15	$4.892\ 658\times10^{-9}$	$9.349\ 258\times10^{-9}$
16	$1.508\ 707\times10^{-9}$	$3.383\ 951\times10^{-9}$
17	$5.766\ 464\times10^{-10}$	$9.320\ 605\times10^{-10}$

表 2　例 2 计算结果

k	$\hat{c}_k$	α_k
1	52.445 86	147.554 1
2	16.467 53	35.978 33
3	4.894 969	11.572 56
4	1.321 45	3.573 494
5	0.363 753 2	0.957 721 8
6	0.083 701 01	0.280 052 2
7	0.019 050 20	0.064 650 82
8	0.003 256 429	0.015 793 77
9	0.000 607 323 8	0.002 649 105
10	0.000 113 371 8	0.000 493 952 1
11	$5.014\ 073\times10^{-6}$	$1.083\ 577\times10^{-4}$
12	$1.343\ 109\times10^{-6}$	$3.670\ 965\times10^{-6}$
13	$4.464\ 178\times10^{-7}$	$8.966\ 91\times10^{-7}$
14	$1.493\ 976\times10^{-7}$	$2.970\ 203\times10^{-7}$
15	$4.991\ 695\times10^{-8}$	$9.948\ 06\times10^{-8}$
16	$1.649\ 75\times10^{-8}$	$3.336\ 720\times10^{-8}$
17	$5.498\ 648\times10^{-9}$	$1.105\ 11\times10^{-8}$
18	$1.839\ 578\times10^{-9}$	$3.659\ 07\times10^{-9}$
19	$6.083\ 076\times10^{-10}$	$1.231\ 27\times10^{-9}$
20	$2.018\ 226\times10^{-10}$	$4.064\ 85\times10^{-10}$
21	$6.647\ 595\times10^{-11}$	$1.353\ 467\times10^{-10}$

例 2 问题如下

$$\min f(x) = 2(x_1^2 + x_2^2) - [x_1^2 + x_2^2],$$

$$\text{s.t.} \; -10 \leqslant x_1 \leqslant 10, \; -10 \leqslant x_2 \leqslant 10,$$

这里 $[z]$ 为 z 的整数部分。最优值为 $c^* = 0$。利用算法 B 迭代 21 步所得结果如表 2 所示。

[参 考 文 献]

[1] 郑权, 蒋百川, 庄松林. 一个求总极值的方法[J]. 应用数学学报, 1978, **2**(1): 164—174.

[2] CHEW Soo_hong, ZHENG Quan. Integral Global Optimization Theory, Implementation and Application [M]. Lecture Notes in Economics and Mathematical Systems, No. 298. Berlin: Springer_Verlag, 1988.

[3] Phu H X, Hoffmann A. Essential supremum and supremum of summable functions[J]. Numerical Fucntional Analysis and Optimization, 1996, **17**(1/2): 167—180.

[4] 华罗庚, 王元. 数论在近似分析中的应用[M]. 北京: 科学出版社, 1978.

[5] 邬冬华, 田蔚文, 张连生. 一个求总极值的实现算法及其收敛性[J]. 运筹学学报, 1999, **3**(2): 82—89.

[6] WU Dong_hua, TIAN Wei_wen, ZHANG Lian_sheng. Optimality Condition for Solving Global Optimization[J]. OR Transactions, 2000, **4**(1): 33—42.

[7] 袁亚湘. 非线性规划数值方法[M]. 上海: 上海科学技术出版社, 1993.

A Level_Value Estimation Method for Solving Global Optimization

WU Dong_hua[1,2], YU Wu_yang[1], TIAN Wei_wen[1], ZHANG Lian_sheng[1]

(1. Department of Mathematics, Shanghai University, Shanghai 200444, P.R. China;

2. Department of Mathematics, Nanjing University, Nanjing 210093, P.R. China)

Abstract: A level_value estimation method was illustrated for solving the constrained global optimization problem. The equivalence between the root of a modified variance equation and the optimal value of the original optimization problem is shown. An alternate algorithm based on the Newton's method is presented and its implementable approach's convergence is proved. Preliminary numerical results indicate that the method is effective.

Key words: global optimization; level_value estimation; uniform distribution