第 28 卷 第 1 期　　应 用 数 学 学 报　　Vol. 28 No. 1
2005 年 1 月　　ACTA MATHEMATICAE APPLICATAE SINICA　　Jan., 2005

一类非光滑全局优化问题的区间展开方法*

申培萍

(河南师范大学数学与信息科学学院，新乡 453007)

(E-mail: shenpeiping@163.com)

张可村

(西安交通大学理学院，西安 710049)

摘　要　本文利用区间展开的特点，对一类全局优化问题提出一新的区间求解方法，该方法能处理多元函数的全局优化问题．数值试验表明提出的方法是可行和有效的．

关键词　全局优化；区间展开；区间删除准则

MR(2000) 主题分类　65K05; 90C30

中图分类　O242.26

1　引言

在解决最优化问题

$$\min f(x), \quad x \in X^0 = \{x \mid f(x) \text{ 的定义域}\} \subset R^n \tag{1}$$

的研究中，国内外众多学者做了大量工作，提出了很多有效算法．但这些算法大多是在求解局部极小方面，而很多现实问题，如分子生物学、经济金融、化学工程设计等方面的问题非常需要的却是获得函数在其定义域上的全局极小，这比求解局部极小要困难得多．目前对这一问题的研究一般分为随机性和确定性两类．前者有模拟退火、遗传算法等，这类方法对 f 的性能要求低，从而应用广泛，但从理论上很难保证获得一个全局最优解．对于后者，一般是通过将 f 变换成一个新的成本函数或者是将定义域进行连续分割等技术手段建立全局优化算法，如区间方法、填充函数法等，这类方法能充分利用 f 的性能，可靠性好，但算法复杂．

本文假定 $X^0 = \{x = (x_j)_{n\times 1} \in R^n \mid a_j^0 \le x_j \le b_j^0,\ j = 1, \cdots, n\}$，$f: R^n \to R^1$ 是 X^0 上连续的多峰函数．记 f^* 为 f 在 X^0 上的全局极小值，X^* 为 f 在 X^0 上所有全局极小点集，这里不妨设 X^* 仅由有限个点组成．在 [1] 中，作者在一维 (即 $n = 1$) 情况下利用斜率给出了求解问题 (1) 的全局优化算法．本文在区间分析的基础上，提出了求多元函数区间展开的计算方法 (参看定理 1)，建立了求解问题 (1) 的不含全局极小点的区间

本文 2003 年 10 月 21 日收到．2004 年 3 月 29 日收到修改稿．
* 国家自然科学基金 (69874010) 和河南省教育厅自然科学基金 (2004110007) 资助项目．

删除判定准则，给出了求解问题 (1) 的全局优化区间算法. 提出的方法具有以下特点: (i) 给出的区间删除准则相当于光滑问题中区间方法常用的单调性准则，因不要求问题的光滑性，从而比单调性准则具有更广的应用范围 (由定理 3 知). (ii) Ratz[1] 的方法是本文方法的特殊情况 ($n=1$)，即本文方法是解决多元函数全局优化问题，而 Ratz [1] 的方法是解决一元函数的全局优化问题. (iii) 数值计算表明当 f 光滑和非光滑时，提出的方法都是可行和有效的.

2 区间删除判定准则和全局优化算法

我们以 $I(X^0)$ 表示 $X^0\subset R^n$ 上所有区间向量的集合. 以 $\mathrm{mid}\,(X), w(X)$ 分别表示 $X\in I(X^0)$ 的中点和宽度. 与本文有关的区间数学的其它概念和内容，读者可参阅 [2] 等.

定义 1[3] 给定区间值映射 $F: D\subset R^1\to I(R^1)$ 和区间向量 $\widetilde{X}$, $X\in I(D)$. 设 $F_c, F_r, F_s\in I(R^1)$, 如果三重关系 (F_c, F_r, F_s) 满足:

(1) $\forall\widetilde{x}\in\widetilde{X}$, 有 $F(\widetilde{x})\subseteq F_c$, 且 $\forall x\in X$, 有 $F(x)\subseteq F_r$;

(2) $\forall\widetilde{x}\in\widetilde{X}$, $\forall x\in X$, $\forall\widetilde{y}\in F(\widetilde{x})$, $\forall y\in F(x)$, 存在 $\widetilde{F_s}\in F_s$, 使得 $y-\widetilde{y}=(x-\widetilde{x})\widetilde{F_s}$. 则称三重关系 (F_c, F_r, F_s) 为 F 在 X 上关于 $\widetilde{X}$ 的区间展开. 进而 F 关于 X 和 $\widetilde{X}$ 的斜率定义为 $\mathrm{sl}\,(F,\widetilde{X},X):=\left\{\frac{y-\widetilde{y}}{x-\widetilde{x}}\middle| y\in F(X), \widetilde{y}\in F(\widetilde{X}),\ x\in X,\ \widetilde{x}\in\widetilde{X},\ x\neq\widetilde{x}\right\}$.

定义 2 给定 $f: R^n\to R^1$, 及 $X=(X_j)_{n\times 1}\in I(R^n)$, $\widetilde{X}=(\widetilde{X}_j)_{n\times 1}\in I(R^n)$. 对任意 j ($j=1,\cdots,n$), 称区间值函数 $f^j: R^1\to I(R^1)$ 为 f 在第 j 个方向上关于 X 和 $\widetilde{X}$ 的分量函数，即对任意 $y\in R^1$, $f^j(y)=f(X_1,\cdots,X_{j-1},y,\widetilde{X}_{j+1},\cdots,\widetilde{X}_n)$.

由定义 1, 2 知 $f^j(\widetilde{X}_j)=f^{j-1}(X_{j-1})\subseteq f_c^j$, $f^j(X_j)\subseteq f_r^j$, $\mathrm{sl}\,(f^j,X_j,\widetilde{X}_j)\subseteq f_s^j$, 且三重关系 $(f^j(X_j), f^j(\widetilde{X}_j), \mathrm{sl}\,(f^j,X_j,\widetilde{X}_j))$ 是 f^j 在 X_j 上关于 $\widetilde{X}_j$ 的区间展开，进而

$$f(x_1,\cdots,x_j,\widetilde{x}_{j+1},\cdots,\widetilde{x}_n)\in f^j(X_j)\subseteq f(\widetilde{x})+\sum_{i=1}^{j}\mathrm{sl}\,(f^i,\widetilde{X}_i,X_i)(X_i-\widetilde{X}_i). \tag{2}$$

显然，当 $j=n$ 时 (2) 式右端即是 f 在 X 上的区间扩张 (关于区间扩张的定义参看 [2]). 特别地，对任意 $x\in X$, 若 $f(x)=\lambda$ (λ 为常数), 则 $(\lambda,\lambda,0)$ 为 f^j 在 X_j 上关于 $\widetilde{X_j}$ 的区间展开; 若 $f(x)=x_i$ (x_i 是 x 的某个分量), 则三重关系 (f_c^j, f_r^j, f_s^j) 中的每一项分别为

$$f_c^j=\begin{cases}\widetilde{X_i}, & 若\ j\le i,\\ X_i, & 若\ j>i,\end{cases}\qquad f_r^j=\begin{cases}\widetilde{X_i}, & 若\ j<i,\\ X_i, & 若\ j\ge i,\end{cases}\qquad f_s^j=\begin{cases}1, & 若\ j=i,\\ 0, & 若\ j\ne i.\end{cases}$$

对一般的 n 元函数 f, 下面定理 1 表明三重关系 (f_c^j, f_r^j, f_s^j) 的计算方法.

定理 1 对给定的 $f,g: D\subset R^n\to R^1$ 和 $\widetilde{X}\subseteq D$, $X\subseteq D$, 设 f^j 和 g^j 分别是 f 和 g 的分量函数. 记 $f_c^j=f^j(\widetilde{X_j})$, $f_r^j=f^j(X_j)$, $f_s^j=\mathrm{sl}\,(f^j,\widetilde{X_j},X_j)$, $g_c^j=g^j(\widetilde{X_j})$, $g_r^j=g^j(X_j)$, $g_s^j=\mathrm{sl}\,(g^j,\widetilde{X_j},X_j)$, $j=1,2,\cdots,n$, 则有如下结论:

(1) 若 (f_c^j, f_r^j, f_s^j) 和 (g_c^j, g_r^j, g_s^j) 分别是 f^j 和 g^j 在 X_j 上关于 $\widetilde{X_j}$ 的区间展开，则 (h_c^j, h_r^j, h_s^j) 是由 f 和 g 经四则运算后所得的函数 $h=f\circ g$ ($\circ\in\{+,-,\cdot,/\}$) 的分量函数 h^j 在 X_j 上关于 $\widetilde{X_j}$ 的区间展开，其中

$$h_c^1=f_c^1\circ g_c^1,\quad h_c^j=h_r^{j-1},\quad j\neq 1,\qquad 若\ \circ\in\{+,-,\cdot,/\},$$

$$h_r^j=(f_r^j\circ g_r^j)\cap(h_c^1+\sum_{i=1}^{j}(X_i-\widetilde{X_i})\,h_s^j),\qquad 若\ \circ\in\{+,-,\cdot,/\},$$

$$h_s^j=\begin{cases} f_s^j\circ g_s^j, & 若\ \circ\in\{+,-\},\\ (f_s^j g_r^j+f_r^{j-1}g_s^j)\cap(f_r^j g_s^j+f_s^j g_r^{j-1}), & 若\ \circ=\cdot,\quad 这里\ g_r^0=g_c^1,\ f_r^0=f_c^1,\\ (f_s^j-h_c^j\,g_s^j)/g_r^j,\quad g_r^j\neq 0, & 若\ \circ=/. \end{cases}$$

(2) 若 $\varphi:\overline{D}\subseteq R^1\to R^1$ 且 $\varphi(f_c^n\cup f_r^n)\subseteq\overline{D}$, 则对复合函数 $h(x)=\varphi(f(x))$, 有

$$h_c^1=\varphi(f_c^1),\quad h_c^j=h_r^{j-1},\quad 其中\ j\neq 1,\qquad h_s^j=sl(\varphi,f_c^j,f_s^j)\,f_s^j,$$

$$h_r^j=\varphi(f_r^j)\cap\Big(h_c^1+\sum_{i=1}^{j}(X_i-\widetilde{X_i})\,h_s^i\Big).$$

定理 1 的证明与 [3] 中定理 2 类似，在此略去.

注意定理 1 中所有运算除了计算 $\mathrm{sl}(\varphi,f_c^j,f_s^j)$ 时可能仅出现点运算外 (看下面定理 2), 其余运算均为区间运算. 另外, 对定理 1 中的函数 φ, 由 [4], 当 φ 可微时, $\mathrm{sl}(\varphi,f_c^j,f_r^j)$ 可被代替为 $\varphi'(f_r^j\underline{\cup}f_c^j)$, 这里 $\varphi'(f_r^j\underline{\cup}f_c^j)$ 表示 φ 的导函数 φ' 在 $f_r^j\underline{\cup}f_c^j$ 上的区间扩张, $\underline{\cup}$ 表示凸包. 一般来说, $\mathrm{sl}(\varphi,f_c^j,f_r^j)\subseteq\varphi'(f_r^j\underline{\cup}f_c^j)$, 但在特殊情况下, $\mathrm{sl}(\varphi,f_c^j,f_r^j)$ 可显式给出，且产生更好的 (即更小的) 区间扩张，看下面定理 2.

定理 2 给定 $f:D\subset R^n\to R^1$, $g:\widehat{D}\subset R^1\to R^1$, 及 $\widetilde{X}$, $X\in I(D)$. 设 (f_c^j,f_r^j,f_s^j) 是 f 的分量函数 f^j 在 X_j 上关于 $\widetilde{X_j}$ 的区间展开, $f_c^j\cup f_r^j$ 的凸包 $f_c^j\underline{\cup}f_r^j$ 满足 $f_c^j\underline{\cup}f_r^j\subseteq\widehat{D}$, $j=1,\cdots,n$. 又设 $h:D\to R^1$ 是 f 和 g 的复合函数 $h=g(f)$. 记 $f_r^j=[\underline{f}_r^j,\overline{f}_r^j]$, $f_c^j=[\underline{f}_c^j,\overline{f}_c^j]$, $h_c^j=g(f_c^j)$, $h_r^j=g(f_r^j)$, $\mathrm{sl}(g,f_c^j,f_r^j)=[\underline{s}^j,\overline{s}^j]$. 如果 g 在 $f_c^j\underline{\cup}f_r^j$ 上是凸函数，那么 $h_s^j=[\underline{s}_1^j,\overline{s}_1^j]\,f_s^j$, 其中

$$\underline{s}_1^j=\begin{cases} \dfrac{g(\underline{f}_r^j)-g(\underline{f}_c^j)}{\underline{f}_r^j-\underline{f}_c^j}, & 若\ \underline{f}_r^j\neq\underline{f}_c^j,\\ g'(\underline{f}_r^j), & 若\ \underline{f}_r^j=\underline{f}_c^j\ 且\ g\ 在\ \widehat{D}\ 上是可微的\\ \underline{\lim}_{\varepsilon\to 0^+}\dfrac{g(\underline{f}_r^j+\varepsilon)-g(\underline{f}_r^j)}{\varepsilon}, & 若\ \underline{f}_r^j=\underline{f}_c^j\ 且\ g\ 在\ \widehat{D}\ 上是不可微的, \end{cases}$$

$$\overline{s}_1^j=\begin{cases} \dfrac{g(\overline{f}_r^j)-g(\overline{f}_c^j)}{\overline{f}_r^j-\overline{f}_c^j}, & 若\ \overline{f}_r^j\neq\overline{f}_c^j,\\ g'(\overline{f}_r^j), & 若\ \overline{f}_r^j=\overline{f}_c^j\ 且\ g\ 在\ \widehat{D}\ 上是可微的,\\ \overline{\lim}_{\varepsilon\to 0^-}\frac{g(\overline{f}_r^j+\varepsilon)-g(\overline{f}_r^j)}{\varepsilon}, & 若\ \overline{f}_r^j=\overline{f}_c^j\ 且\ g\ 在\ \widehat{D}\ 上是不可微的. \end{cases}$$

如果 g 在 $f_c^j\underline{\cup}f_r^j$ 上是凹函数，那么 $h_s^j=[\underline{s}_2^j,\overline{s}_2^j]\,f_s^j$, 其中

$$\underline{s}_2^j=\begin{cases} \dfrac{g(\overline{f}_r^j)-g(\overline{f}_c^j)}{\overline{f}_r^j-\overline{f}_c^j}, & 若\ \overline{f}_r^j\neq\overline{f}_c^j,\\ g'(\overline{f}_r^j), & 若\ \overline{f}_r^j=\overline{f}_c^j\ 且\ g\ 在\ \widehat{D}\ 上是可微的,\\ \overline{\lim}_{\varepsilon\to 0^+}\dfrac{g(\overline{f}_r^j+\varepsilon)-g(\overline{f}_r^j)}{\varepsilon}, & 若\ \overline{f}_r^j=\overline{f}_c^j\ 且\ g\ 在\ \widehat{D}\ 上是不可微的, \end{cases}$$

$$\overline{s}_2^j=\begin{cases}\dfrac{g(\underline{f}_r^j)-g(\underline{f}_c^j)}{\underline{f}_r^j-\underline{f}_c^j}, & \text{若 } \underline{f}_r^j\neq\underline{f}_c^j,\\ g'(\underline{f}_r^j), & \text{若 } \underline{f}_r^j=\underline{f}_c^j \text{ 且 } g \text{ 在 } \widehat{D} \text{ 上是可微的},\\ \lim\limits_{\varepsilon\to 0^-}\dfrac{g(\underline{f}_r^j+\varepsilon)-g(\underline{f}_r^j)}{\varepsilon}, & \text{若 } \underline{f}_r^j=\underline{f}_c^j \text{ 且 } g \text{ 在 } \widehat{D} \text{ 上是不可微的}.\end{cases}$$

关于定理 2 的详细证明过程与 [3] 中的定理 3 的证明类似，此处从略.

通过上述讨论，对目标函数 f, 总可以通过定理 1 中给出的运算规则最终确定 f 在某一区间向量 X 上的区间扩张，再根据 f 在某个方向上的分量函数 f^j 的区间斜率 f_s^j, 我们可以构造不含全局极小点的区间删除判定准则，这将由下面的定理 3 表明.

设 $f: R^n\to R^1$, $\forall X=(X_j)_{n\times 1}\subseteq X^0\in I(R^n)$, $X_j=[a_j,b_j]$, 选取 $c=(c_j)_{n\times 1}\in X$. 又设 $\overline{f}$ 是 f^* 的一个上界，即 $\overline{f}\geq f^*=\min\limits_{x\in X^0}f(x)$. 为了下面的定理 3, 对任意 i $(i=1,\cdots,n)$, 记

$$f^i(y)=f(X_1,\cdots,X_{i-1},y,c_{i+1},\cdots,c_n),\qquad f_s^i=sl(f^i,c_i,X_i)=[\underline{f}_s^i,\overline{f}_s^i],$$

$$\eta_i=\sum_{j=1,\,j\neq i}^{n}\min((b_j-c_j)\underline{f}_s^j,(a_j-c_j)\overline{f}_s^j),\qquad \delta_i=\overline{f}-f(c)-\eta_i.$$

定理 3 对任意给定的 $X=(X_j)_{n\times 1}\subseteq X^0\in I(R^n)$, 如果存在某个指标 i $(i=1,\cdots,n)$, 使得区间向量 $Y=(Y_j)_{n\times 1}\in I(X^0)$ 满足：对任意 $j=1,\cdots,n$, 当 $j\neq i$ 时 $Y=X_j$; 当 $j=i$ 时 $Y_j=Z_i$, 则 $Y\subseteq X$, 且 f 在 Y 上不存在全局极小点，即 $\min\limits_{y\in Y}f(y)>f^*$, 其中

$$Z_i=\begin{cases}\left(c_i+\dfrac{\delta_i}{\overline{f}_s^i},b_i\right]\cap X_i, & \text{若 } \underline{f}_s^i>0,\ \delta_i\leq 0,\ \text{或者 } \underline{f}_s^i=0,\ \overline{f}_s^i>0,\ \delta_i<0,\\ \left[a_i,c_i+\dfrac{\delta_i}{\underline{f}_s^i}\right)\cap X_i, & \text{若 } \overline{f}_s^i<0,\ \delta_i\leq 0,\ \text{或者 } \underline{f}_s^i<0,\ \overline{f}_s^i=0,\ \delta_i<0,\\ \left(c_i+\dfrac{\delta_i}{\overline{f}_s^i},c_i+\dfrac{\delta_i}{\underline{f}_s^i}\right)\cap X_i, & \text{若 } \underline{f}_s^i<0<\overline{f}_s^i,\ \delta_i\leq 0,\\ \left[a_i,c_i+\dfrac{\delta_i}{\overline{f}_s^i}\right), & \text{若 } \overline{f}_s^i<0,\ 0<\delta_i\leq(a_i-c_i)\overline{f}_s^i,\\ \left(c_i+\dfrac{\delta_i}{\underline{f}_s^i},b_i\right], & \text{若 } \underline{f}_s^i>0,\ 0<\delta_i\leq(b_i-c_i)\underline{f}_s^i,\\ [a_i,b_i], & \text{若 } \underline{f}_s^i=\overline{f}_s^i=0,\ \delta_i<0.\end{cases}$$

证 $Y\subseteq X$ 是显然的. 不失一般性，对任意 $y=(y_j)_{n\times 1}\in Y$, 取 $c=(c_j)_{n\times 1}\in Y$, 有

$$\begin{aligned}f(y)\in& f(c)+\sum_{j=1}^{n}sl(f^j,Y_j,c_j)(y_j-c_j)\\ \subseteq& f(c)+sl(f^i,Y_i,c_i)(y_i-c_i)+\sum_{j=1,\,j\neq i}^{n}\mathrm{sl}\,(f^j,Y_j,c_j)(Y_j-c_j)\end{aligned}$$

$$=f(c)+[\underline{f}_s^i,\overline{f}_s^i](y_i-c_i)+\sum_{j=1,\,j\neq i}^{n}[\underline{f}_s^j,\overline{f}_s^j](Y_j-c_j).$$

下面分别就四种情况，即 (1) $Z_i=\left(c_i+\frac{\delta_i}{\overline{f}_s^i},b_i\right]\cap X_i$, (2) $Z_i=\left[a_i,c_i+\frac{\delta_i}{\overline{f}_s^i}\right)$, (3) $Z_i=\left(c_i+\frac{\delta_i}{\overline{f}_s^i},c_i+\frac{\delta_i}{\underline{f}_s^i}\right)\cap X_i$, (4) $Z_i=[a_i,b_i]$ 进行证明，其余情况可类似证明.

(1) 当 $Z_i=\left(c_i+\frac{\delta_i}{\overline{f}_s^i},b_i\right]\cap X_i$ 时，考虑下面两种情况: (i) 若 $\underline{f}_s^i>0$ 且 $\delta_i\leq 0$, 则当 $y_i\geq c_i$ 时，由 $y_i>c_i+\delta_i/\overline{f}_s^i\geq c_i+\delta_i/\underline{f}_s^i$ 知

$$\begin{aligned}f(y)\geq&f(c)+(y_i-c_i)\underline{f}_s^i+\sum_{j=1,\,j\neq i}^{n}\min\left((a_j-c_j)\underline{f}_s^j,(b_j-c_j)\underline{f}_s^j,(a_j-c_j)\overline{f}_s^j,(b_j-c_j)\overline{f}_s^j\right)\\=&f(c)+(y_i-c_i)\underline{f}_s^i+\eta_i=(y_i-c_i)\underline{f}_s^i+\overline{f}-\delta_i>\overline{f}\geq f^*.\end{aligned}$$

当 $y_i<c_i$ 时，由 $c_i+\delta_i/\overline{f}_s^i<y_i<c_i$ 可得

$$f(y)\geq f(c)+(y_i-c_i)\overline{f}_s^i+\eta_i=(y_i-c_i)\overline{f}_s^i+\overline{f}-\delta_i>\overline{f}\geq f^*.$$

(ii) 若 $\underline{f}_s^i=0$, $\overline{f}_s^i>0$, 且 $\delta_i<0$, 则当 $y_i\geq c_i$ 时，有

$$\begin{aligned}f(y)\geq&f(c)+\sum_{j=1,\,j\neq i}^{n}\min\left((a_j-c_j)\underline{f}_s^j,(b_j-c_j)\underline{f}_s^j,(a_j-c_j)\overline{f}_s^j,(b_j-c_j)\overline{f}_s^j\right)\\=&f(c)+\eta_i>\overline{f}\geq f^*.\end{aligned}$$

当 $y_i<c_i$ 时，由 $c_i+\delta_i/\overline{f}_s^i<y_i<c_i$ 可得

$$f(y)\geq f(c)+(y_i-c_i)\overline{f}_s^i+\eta_i=(y_i-c_i)\overline{f}_s^i+\overline{f}-\delta_i>\overline{f}\geq f^*.$$

由此可得 $Z_i=\left(c_i+\frac{\delta_i}{\overline{f}_s^i},b_i\right]\cap X_i$ 时的结论，同理可证 $Z_i=[a_i,c_i+\frac{\delta_i}{\underline{f}_s^i})\cap X_i$ 时的结论.

(2) 当 $Z_i=[a_i,c_i+\frac{\delta_i}{\overline{f}_s^i})$ 时, 根据 $0<\delta_i\leq(a_i-c_i)\overline{f}_s^i$ 和 $\overline{f}_s^i<0$, 我们有 $a_i\leq c_i+\frac{\delta_i}{\overline{f}_s^i}<c_i$, 又 $y_i<c_i+\frac{\delta_i}{\overline{f}_s^i}<c_i$. 因此 $f(y)\geq f(c)+(y_i-c_i)\overline{f}_s^i+\eta_i=(y_i-c_i)\overline{f}_s^i+\overline{f}-\delta_i>\overline{f}\geq f^*$. 类似可得 $Z_i=\left(c_i+\frac{\delta_i}{\underline{f}_s^i},b_i\right]$ 时的结论.

(3) 当 $Z_i=\left(c_i+\frac{\delta_i}{\overline{f}_s^i},c_i+\frac{\delta_i}{\underline{f}_s^i}\right)\cap X_i$ 时，根据 $\delta_i\leq 0$ 和 $\underline{f}_s^i<0<\overline{f}_s^i$, 我们有 $c_i+\frac{\delta_i}{\overline{f}_s^i}\leq c_i\leq c_i+\frac{\delta_i}{\underline{f}_s^i}$. 再由 $y_i\in\left(c_i+\frac{\delta_i}{\overline{f}_s^i},c_i+\frac{\delta_i}{\underline{f}_s^i}\right)$, 若 $y_i>c_i$, 则 $f(y)\geq f(c)+(y_i-c_i)\underline{f}_s^i+\eta_i=(y_i-c_i)\underline{f}_s^i+\overline{f}-\delta_i>\overline{f}\geq f^*$; 否则，有 $y_i\leq c_i$, 则 $f(y)\geq f(c)+(y_i-c_i)\overline{f}_s^i+\eta_i=(y_i-c_i)\overline{f}_s^i+\overline{f}-\delta_i>\overline{f}\geq f^*$.

(4) 当 $Z_i=[a_i,b_i]$ 时，由 $\underline{f}_s^i=\overline{f}_s^i=0$, $\delta_i<0$, 和 $y_i\in[a_i,b_i]$, 我们有

$$f(y)\geq f(c)+\eta_i=\overline{f}-\delta_i>\overline{f}\geq f^*.$$

综上所述，对任意 $y\in Y$, 有 $\min\limits_{y\in Y}f(y)>f^*$, 定理 3 得证.

根据上述定理 3, 对于给定的 $X=(X_j)_{n\times 1}\in I(X^0)$, $X_j=[a_j,b_j]$, 我们可以建立删除 X 中不含全局极小点的区间判定准则. 下面给出具体的删除过程:

首先对每个 i ($i=1,\cdots,n$), 计算 $\delta_i := \overline{f}-f(c)-\sum\limits_{j=1,\,j\neq i}^{n}\min\left((b_j-c_j)\underline{f}_s^j,(a_j-c_j)\overline{f}_s^j\right)$, $p_i := c_i+\delta_i/\overline{f}_s^i$ 和 $q_i := c_i+\delta_i/\underline{f}_s^i$. 然后逐个判定定理 3 的诸条件中的哪一个成立, 根据不同的条件, 相应地进行不同的区间删除操作. 不妨设存在 k $(1\le k\le n)$, 使得 $\underline{f}_s^k<0<\overline{f}_s^k$ 和 $\delta_k\le 0$ 成立, 此时 $Z_k=\left(c_k+\frac{\delta_k}{\overline{f}_s^k},c_k+\frac{\delta_k}{\underline{f}_s^k}\right)\cap X_k$, 则用新的 $\widetilde{X}=(\widetilde{X}_j)_{n\times 1}$ 代替 X, 其中, 当 $j\neq k$ 时 $\widetilde{X}_j=X_j$, $j=k$ 时

$$\widetilde{X}_k=\begin{cases}[a_k,p_k]\cup[q_k,b_k], & \text{如果 } a_k\le p_k \text{ 且 } q_k\le b_k,\\ [a_k,p_k], & \text{如果 } a_k\le p_k \text{ 且 } q_k> b_k,\\ [q_k,b_k], & \text{如果 } a_k> p_k \text{ 且 } q_k\le b_k,\\ \emptyset, & \text{其它}.\end{cases}$$

从而删除 X 中不含解的一大部分区域或全部区域 (当 $\widetilde{X}_k=\emptyset$ 时), X 中剩余部分归入工作队列, 等待以后进行进一步的删除检验. 对于定理 3 中的其它假定条件成立的情形, 可仿上述过程进行类似的删除操作. 此外, 若对每个 i, $i=1,\cdots,n$, 定理 3 中的所有假定条件都不满足, 此时我们将 X 沿最大宽度方向进行二等分, 得到的两个新区间归入工作队列.

注 本文给出的区间删除准则相当于光滑问题中所常用的单调性准则, 即若 $0\notin F_i'(X)$, 则可将 X 缩减为 $X=X_1\times\cdots\times X_{i-1}\times\overline{X}_i\times X_{i+1}\times\cdots\times X_n$, 其中 $\overline{X}_i=a_i$ 或 b_i, $\nabla F_i'(X)$ 是偏导数 $\frac{\partial f}{\partial x_i}$ 在 X 上的区间扩张.

二者相比, (i) 单调性准则只能用于光滑问题, 而本文给出的删除准则既可用于光滑问题, 也可处理非光滑性问题 (由定理 3 知); (ii) 当 $0\in F_i'(X)$, 即 f 在 X 上存在多个鞍点或多个局部极小点时单调性准则不能进行任何区间删除, 而我们的删除准则不依赖于 f 在 X 上是否存在鞍点或局部极小点, 只要给定的条件满足, 仍可进行删除处理. 因此对全局优化算法而言, 本文的删除判定准则对删除可行域中不含全局解的区域的有效执行几率高于单调性准则, 具有更广的应用范围, 能作为全局优化区间方法重要的加速工具.

根据构造的区间删除判定准则, 我们可以建立求解问题 (1) 的区间算法.

假定 $\forall X^m=(X_j^m)_{n\times 1}\in I(X^0)$, $X_j^m=[a_j^m,b_j^m]$, 记 $a^m=(a_j^m)_{n\times 1}\in R^n$, $b^m=(b_j^m)_{n\times 1}\in R^n$, $m=0,1,2,\cdots$. 又设满足精度要求的最优区间队列记为 L_ε, 工作队列记为 L_w. $\forall X\in L_w$, 设 $F(X)=\left[\underline{F}(X),\overline{F}(X)\right]$ 为 f 在 X 上的区间扩张, $\overline{f}$ 和 $\overline{z}$ 分别为 f 在 X 上的当前极小值和所有当前极小点构成的集合, k 表示算法的迭代次数. 算法的具体步骤如下:

步 1 初始化: $L_w:=\emptyset$, $L_\varepsilon:=\emptyset$, 给定 X^0 及精度 ε, 令 $k:=0$.

步 2 对 X^0 进行处理: 令 $X=(X_j)_{n\times 1}:=X^0$, $c=(c_j)_{n\times 1}=\mathrm{mid}\,(X^0)$, 确定 X 的最大宽度方向, 计算 $F(X),w(X)$, $\mathrm{sl}\,(f^j,X_j,c_j)=[\underline{f}_s^j,\overline{f}_s^j]$, $j=1,\cdots,n$, $\overline{f}=\min\left(f(a^0),f(b^0)\right)$, $\overline{z}=\left\{x\mid f(x)=\overline{f},\ x=a^0 \text{ 或 } b^0\right\}$.

步 3 分裂盒子或删除盒子的一部分: 令 $k:=k+1$, $c=(c_j)_{n\times 1}=\mathrm{mid}\,(X)$.

3.1 若 $f(c)=\overline{f}$, 则 $\overline{z}=\{c\}\cup\overline{z}$, 沿 X 的最大宽度方向将其二分为 X^1,X^2.

3.2 若 $f(c)>\overline{f}$, 则对 X 应用定理 3 进行区间删除, 此时 X 的剩余部分记为 X^1,X^2.

3.3 若 $f(c)<\overline{f}$, 则令 $\overline{f}:=f(c)$, $\overline{z}:=\{c\}$. 更新 c 点, 以保证新选取的 c 点满足 $f(c)>\overline{f}$, 然后应用定理 3 进行区间删除, 此时 X 的剩余部分记为 X^1,X^2.

步 4 循环处理盒子 X^1,X^2: 令 $l=1$,

4.1 若 X^l 为空, 则跳出本次循环; 否则执行 4.2.

4.2 计算 $F(X^l), w(X^l)$, 和 $\mathrm{sl}(f^j, c_j, X^l_j) = [\underline{f}^j_s, \overline{f}^j_s]$, $j = 1, \cdots, n$, 这里 $X^l = (X^l_j)_{n\times 1}$.

若 $\underline{F}(X^l) > \overline{f}$, 则删除该盒子 (这一判断过程即为中点考查准则), 跳出本次循环; 否则, 更新当前极小值, 即令 $\overline{f} = \min\left(f(a^l), f(b^l), f(c)\right)$, $\overline{z} = \{x \mid f(x) = \overline{f},\ x = a^l$, 或 b^l, 或 $c\}$.

4.3 对 X^l 进行归属判别, 若 $w(X^l) < \varepsilon$ 或 $\overline{F}(X^l) - \underline{F}(X^l) < \varepsilon$, 则将 X^l 归入 L_ε; 否则将 X^l 归入 L_w, $l := l + 1$. 若 $l > 2$, 则执行步 5, 否则返回 4.1.

步 5 若 $L_w = \emptyset$, 则转入步 7, 否则删除 L_w 中所有满足 $\underline{F}(X^l) > \overline{f}$ 的盒子 X^l ($X^l \in L_w$), 转入步 6.

步 6 若 $L_w = \emptyset$, 则转入步 7, 否则从 L_w 中选取具有最小 $\underline{F}(X^l)$ 的盒子 X^l, 令 $X := X^l$, 转入步 3.

步 7 若 $L_\varepsilon \neq \emptyset$, 则删除 L_ε 中所有满足 $\underline{F}(X^l) > \overline{f}$ 的盒子 X^l (中点考查).

步 8 输出 $\varepsilon, \overline{z}, L_\varepsilon, k$, 及 $F^* = [\underline{F}, \overline{f}]$ $\left(这里 \underline{F} = \min\limits_{X^l \in L_\varepsilon} \underline{F}(X^l)\right)$, 停机.

下面给出上述算法的性质:

定理 4 设 $f: X^0 \subseteq R^n \to R$, 给定精度 $\varepsilon > 0$, 对任意 $X \in I(X^0)$ 执行上述算法. 设 X^1 和 X^2 是由算法步 3 产生的, 则 (1) $X^1 \cup X^2 \subseteq X$. (2) f 在 X 上的全局极小点 x^* 满足 $x^* \in X^1 \cup X^2$. (3) 若 $X^1 \cup X^2 = \emptyset$, 则 f 在 X 上无全局极小点. (4) $f^* \in F^* = [\underline{F}, \overline{f}]$, 且 $X^* \subseteq \bigcap\limits_{X \in L_\varepsilon} X$.

证 由 X^1 和 X^2 的定义得性质 (1), 由定理 3 得出性质 (2), 由 (2) 可得 (3). 下证 (4) 成立. 对任意 Y^1, $Y^2 \in L_w$, 如果 $Y^1 \subseteq Y^2$, 那么 $\underline{F}(Y^1) \geq \underline{F}(Y^2)$. 从而由 L_ε 及性质 (2) 知 $\underline{F}$ 是 f^* 的一个下界. 另外 $\overline{f}$ 是 f^* 的一个上界. 因此 $f^* \in F^*$. 再由性质 (2) 可得 $X^* \subseteq \bigcap\limits_{X \in L_\varepsilon} X$.

3 数值实验

下面对一些典型的光滑和非光滑测试问题在 Pentium Ⅲ (733 MHz) 微机上进行数值实验. 对所有测试问题, 本文算法 (记为方法 1) 均能在给定精度内求出优化问题的所有全局最优解. 在计算过程中, 对于光滑问题, 将本文算法与利用单调性准则进行区间删除的算法 (记为方法 2)(参见 [1, 5–7]) 进行比较, 计算结果列于表 1 中. 由表 1 知对所有测试问题方法 1 比方法 2 在总的计算时间上改进 53%, 总的工作队列最大长度改进 14%, 总的函数计算次数改进 71%. 对于非光滑问题, 目前区间方法只能用中点考查作为区间删除工具 (参见 [1, 5–7]), 且大多数数值实验仅对光滑问题进行. 为了说明本文算法对非光滑问题的执行效率, 我们将方法 1 和只用中点考查准则作为区间删除的算法 (记为方法 3) 进行比较, 结果表明方法 3 对一维非光滑问题尚可行, 但对高维问题算法的性能很差, 甚至在一般微机上不能求解, 而本文算法对高维非光滑问题不仅能求出非光滑问题的所有全局最优解, 且在精度要求, 运行时间和存储空间上都优于方法 3(参见表 2).

所用的测试问题分别为: Goldstein-Price (记为 GP), Levy-2 (L2), Levy-5 (L5), Levy-1 (L1), Branin RCOS (BR), Schwefel-3.7 (Sch37), Schwefel-1.2 (Sch12), Six-Hump-Camel-Back (SHCB), Test problem 2 (TP2). 本文对上述问题的搜索区域与 [7–9] 相同, 其余测试问题 (记为 Pr.) 如下:

Pr.1 $f(x) = \sin\frac{1}{x}$, $0.02 \leq x \leq 1$, $X^* = \left\{\frac{2}{(4k-1)\pi} \mid k = 1, \cdots, 8\right\}$, $f^* = -1$.

Pr.2 $f(x) = x_1^6/6 - 0.75x_1^4 + 2x_1^2 - x_1^2 x_2 + 10$, $-5 \leq x_1$, $x_2 \leq 53$,
$X^* = \left\{(-1.999\cdots, 3.999\cdots), (1.999\cdots, 3.999\cdots)\right\}$, $f^* = -8.66666\cdots$.

Pr.3　$f(x)=\sum_{i=1}^{n}|x_i-0.5|,\ -5\le x_i\le 5\ (i=1,\cdots,n),\ f^*=0,\ X^*=\{(0.5)_{n\times 1}\}.$

Pr.4　$f(x)=-20\exp\Big(-0.2\sqrt{\frac{1}{n}\sum_{i=1}^{n}|x_i|}\Big)-\exp\Big(\frac{1}{n}\sum_{i=1}^{n}\cos(2\pi x_i)\Big)+20,$

$-20\le x_i\le 30\ (i=1,\cdots,n),\ f^*=-2.7182818\cdots,\ X^*=\{0\}.$

Pr.5　$f(x)=\sum_{j=1}^{5}j\big|\cos((j+1)x+j)\big|+5,\ -10\le x\le 10,\ f^*=6.699793\cdots,$

$X^*=\{5.169\cdots,-4.255\cdots,-1.114\cdots,2.0274\cdots,\ -7.397\cdots,8.3106\cdots\}.$

Pr.6　$f(x)=\big|(x-1)/4\big|+\big|\sin\big(\pi(1+(x-4)/4)\big)\big|+1,\ -10\le x\le 10,$

$f^*=1,\ X^*=\{1.\}.$

Pr.7　$f(x)=\Big\{\sum_{k=1}^{5}k\big|\cos\big((k+1)x_1+k\big)\big|+5\Big\}\Big\{\sum_{k=1}^{5}k\big|\cos\big((k+1)x_2+k\big)\big|+5\Big\},$

$x_1,x_2\in[-10,10],$

$f^*=44.887236639\cdots,\ X^*=\{(5.169\cdots,2.027\cdots),\cdots\}$, X^* 中共有 36 个元素.

表 1　用方法 1(记为 I) 和方法 2(记为 II) 计算各光滑问题的数值结果

Problem			NFE			MLL			Time		
Name	Dim	Ref	II	I	$\frac{I}{II}$%	II	I	$\frac{I}{II}$%	II	I	$\frac{I}{II}$%
L5	2	[2, 8, 9]	1839	870	47	59	65	110	0.22	0.11	50
BR	2	[9]	2247	1141	51	60	54	90	0.05	0.05	100
GP	2	[8, 9]	3543	2617	74	103	82	80	0.16	0.11	69
Sch37	7	[2]	27645	7423	27	128	128	100	11.86	5.83	49
Sch37	10	[2]	251901	63487	25	1024	1024	100	221.35	102.21	47
Sch12	3	[2]	15273	7447	49	602	326	54	0.49	0.1	20
L2	2	[2, 8, 9]	8901	6956	78	314	256	82	1.05	0.93	89
TP2	1	[7]	391	161	41	17	17	100	0.05	0.0	0
Pr 1	1		634	206	32	17	16	94	0.0	0.0	100
SCHB	2	[8]	1341	1219	91	48	50	104	0.05	0.05	100
Pr.2	2		2565	1387	54	48	75	156	0.05	0.05	100
L1	1	[2]	374	158	42	16	6	38	0.0	0.0	100
$\sum$			316663	93070	29	2436	2099	86	235.33	109.44	47
AoP					51			92			69

表 2　用方法 1(I) 和方法 3(记为 III) 计算各非光滑问题的数值结果

Problem		NFE		MLL		Time		ε	
Name	Dim	I	III	I	III	I	III	I	III
Pr.3	15/2	45591	26265	2559	20	19.72	21.26	10^{-10}	10^{-4}
Pr.4	4/3	636	52461	285	21	0.27	97.05	10^{-10}	10^{-3}
Pr.5	1	208	269	33	18	0.0	0.0	10^{-10}	10^{-10}
Pr.6	1	25	75	8	9	0.0	0.0	10^{-10}	10^{-10}
Pr.7	2	14545	56001	945	384	2.42	119.63	10^{-10}	10^{-3}

上面表 1, 2 中的记号分别为: Name: 问题的名字; Dim: 问题的维数; Ref: 参考文献; Time: 算法的运行时间 (秒); MLL: 工作队列最大长度; NFE: 函数计算次数; $\frac{I}{II}$%: 方法 1 和方法 2 对应各量的百分比; $\sum$: 所有问题相应量的总和; AoP: 相应量的平均值. 表 1 中各题精度除问题 Sch37 取 $\varepsilon=10^{-8}$ 外, 其余均为 $\varepsilon=10^{-10}$. 表 2 中第三行的 Dim=15/2 表示分别用方法 1 和方法 3 对算例 Pr.3 计算时, 相应的维数分别取 15 和 2, 类似可知 Dim=4/3 的意义. 计算过程中我们观察到在对高维非光滑函数用方法 3 计算时, 必须降低问题的精度和维数才能使算法停止, 正如表 2 中所选择的参数 ε 和 Dim 那样, 否则算法在我们实验的微机上不能正常输出结果.

致谢 作者十分感谢评审人对本文提供的有价值的建议！

参　考　文　献

1 Ratz D. A Nonsmooth Global Optimization Technique Using Slopes: One-dimensional Case. *Journal of Global Optimization*, 1999, 4: 365–393

2 Hansen E R. Global Optimization Using Interval Analysis. New York: Marcel Dekker, 1993, 113–190

3 Rump S M. Expansion and Estimation of the Range of Nonlinear Functions. *Mathematics of Computation*, 1996, 65: 1503–1512

4 Shen Zuhe, Wolfe M A. On Interval Enclosures Using Slope Arithmetic. *Applied Mathematics and Computation*, 1990, 39: 89–105

5 Andras Erik Csallner, Tibor Csendes, Csaba Markot Mihaly. Multisection in Interval Branch-and-bound Methods for Global Optimization I: Theoretical Results. *Journal of Global Optimization*, 2000, 16: 371–392

6 Kolev Lubomir V. Use of Interval Slopes for the Irrational Part of Factorable Functions. *Reliable Computing*, 1997, 3: 83–93

7 Hansen P, Jaumard B, Xiong J. Cord-slope Form of Taylor's Expansion in Univariate Global Optimization. *Journal Of Optimization Theory and Applications*, 1994, 80(3): 441–464

8 Tibor Csendes. New Subinterval Selection Criteria for Interval Global Optimization. *Journal of Global Optimization*, 2001, 19: 307–327

9 Casado L G, Martinez J A, Garcia I. Experiments with a New Selection Criterion in a Fast Interval Optimization Algorithm. *Journal of Global Optimization*, 2001, 19: 247–264

AN INTERVAL EXPANSION METHOD OF A KIND OF NONSMOOTH GLOBAL OPTIMIZATION

SHEN PEIPING

(*College of Mathematics and Information Science, Henan Normal University, Xinxiang* 453007)

(*E-mail: shenpeiping@163.com*)

ZHANG KECUN

(*Faculty of Science, Xi'an Jiaotong University, Xi'an* 710049)

Abstract In the paper, a new interval method is given for a kind of global optimization problem using the feature of interval expansion, and it can solve global optimization problems with several variables. Numerical experiments show that the proposed method is feasible and effective.

Key words global optimization; interval expansion; interval elimination rule

MR(2000) Subject Classification 65K05; 90C30

Chinese Library Classification O242.26