第45卷 第4期
2006年7月
厦门大学学报(自然科学版)
Journal of Xiamen University (Natural Science)
Vol. 45 No. 4
Jul. 2006

结合区间和符号方法的多变元多项式的混合全局最优化方法

冯亚丽[1], 沈喜生[2], 薛继伟[1], 伊三泉[3]

(1. 大庆石油学院 计算机与信息技术学院, 黑龙江 大庆 163318;
2. 厦门大学数学科学学院, 福建 厦门 361005; 3. 大庆石油学院教务处, 黑龙江 大庆 163318)

摘要: 提出了一种求解带边界约束的多变元多项式全局最优解的混合方法. 混合是指在优化的过程中结合了区间方法、符号方法和数值方法. 一方面通过区间方法在舍入误差存在的情况下得到包含最优解且满足要求的任意小区间; 另一方面通过符号方法解决当Jacobi矩阵在区间内某点奇异时区间牛顿法无法验证驻点的存在性与惟一性的问题; 同时, 利用数值优化方法(如BFGS方法)来有效克服区间方法运算速度慢的缺点. 此外, 文中的算法非常有利于并行化, 因此可以进一步提高算法效率.

关键词: 全局优化; 区间分析; Gröbner基; 特征值

中图分类号: TP 391.4　　文献标识码: A　　文章编号: 0438-0479(2006)04-0470-05

全局优化研究的是非线性函数在某个区域上全局最优点的特征和计算方法. 在科学、经济和工程中, 许多新的进展都依赖于相应优化问题全局最优解的计算. 全局优化问题的来源相当广泛, 包括经济建模、网络、运输、数据库、芯片设计、图像处理、化学工程设计、控制及分子生物学等. 与局部优化相比, 全局优化在理论上还不是很成熟, 仍有许多问题亟待解决.

区间算法(Interval Arithmetic)是Moore R E于20世纪60年代提出的, 是科学计算和可靠计算的一个热门研究领域, 它由定义在一系列区间上的算法组成. 区间优化方法可以保证找到全局最优解, 并给出包含最优解的满足要求的任意小区间. 算法中通常采用区间Newton Gauss-Seidel方法来验证给定区间内是否存在极值点. 虽然区间优化方法解决了许多具有挑战性的全局优化问题, 但区间方法存在以下缺点:

(i) 时间复杂度和空间复杂度较高;

(ii) 判断某个区间内是否存在最优点比较困难.

本文设计了一种严格的混合全局优化方法, 即: 以区间优化方法为基本框架, 尽可能地利用数值优化速度快的优点提高算法的性能, 并利用Pedersen等提出的符号方法解决区间方法中判断驻点的不严密性. 主要要点如下:

(i) 在调用区间算法之前, 利用BFGS变尺度法和多随机点法得到问题的尽可能优的解;

(ii) 当区间Newton Gauss-Seidel方法验证了某区间内只有一个极值点, 但区间的宽度还不满足需求时, 首先调用BFGS方法, 然后再利用区间方法;

(iii) 当区间的宽度已经满足需求, 但区间Newton Gauss-Seidel方法无法验证极值点的存在性与惟一性时, 利用符号方法进行验证.

文中小写黑体字母(如: $\boldsymbol{x}, \boldsymbol{y}$)表示区间、小写字母(如: x, y)表示标量、大写字母(如: A, B)表示向量或矩阵、大写黑体字母(如: $\boldsymbol{A}, \boldsymbol{B}$)表示区间向量或区间矩阵. $\overline{\boldsymbol{x}}$ 表示区间的上界、$\underline{\boldsymbol{x}}$ 表示区间的下界、$m(\boldsymbol{x})$ 表示区间 $\boldsymbol{x}$ 的中点、R表示实数域. ∇f 和 $\nabla^2 f$ 分别表示函数 f 的梯度和Hessian矩阵.

1 区间优化方法的描述及其缺点

下面简要给出利用区间技术进行优化的算法描述[1,2], 有关区间运算的详细内容请参阅文献[3]. 令 f 表示目标函数, $\boldsymbol{x}_0$ 表示初始区间, L 表示等待处理的区间列表(为了提高算法的效率, L 中的区间一般是按照目标函数 f 在此区间上的下界的升序进行排列), C 是可以确定包含最优点的区间列表, U 中包含那些宽度已经满足要求但无法确定是否包含最优点的区间.

算法1(区间优化方法的一般结构)

IntGlobalOpt($f, \boldsymbol{x}_0, \varepsilon$)

收稿日期: 2005-12-09
基金项目: 国家"973"计划项目(2004CB318003), 黑龙江省自然科学基金(F01-21)资助

作者简介: 冯亚丽(1958-), 女, 副教授.
Email: dqfyl2008@sina.com

Begin

(i) 初始化列表 $L = \boldsymbol{x}_0$;

(ii) Do While $L \neq \emptyset$

1) 从 L 中取出第一个区间 $\boldsymbol{x}$;

2) 对 $\boldsymbol{x}$ 进行处理，可能得到如下结果之一：

(a) 舍弃 $\boldsymbol{x}$;

(b) 减小 $\boldsymbol{x}$;

(c) 如果可以确定 $\boldsymbol{x}$ 只包含一个极值点，则继续处理直到满足精度要求；

(d) 把区间 $\boldsymbol{x}$ 划分成多个小的区间，以便于处理.

3) 对(ii)，2) 得到的区间进行如下处理：

(a) 如果 $\boldsymbol{x}$ 被划分成多个区间，则留下一个区间继续处理，其余的区间插入到列表 L 中；

(b) 如果已经证明区间 $\boldsymbol{x}$ 包含惟一一个极值点，则把 $\boldsymbol{x}$ 加入到列表 C 中；

(c) 如果 $\boldsymbol{x}$ 的宽度已经满足用户要求，但还无法证明是否包含极值点，则把其加入到列表 U 中.

End Do

End

上述算法中的 2)-(b) 是区间优化方法的主要步骤，并主要影响算法的效率. 在实际使用中通常采用 Interval Newton Step、Midpoint Test、Monotonicity Test 和 Concavity Test 等技术对区间进行处理以提高算法的性能.

Interval Newton Step(区间牛顿法)：在区间优化方法中，通常采用 Interval Newton Gauss-Seidel 方法来求解下面的非线性方程组

$$\nabla f(Y) = 0,\ Y \in \boldsymbol{Y},$$

也就是说，通过求解下面的线性区间方程组

$$\nabla^2 f(\boldsymbol{Y}) \circ (m(\boldsymbol{Y}) - \boldsymbol{Y}) = \nabla f(m(\boldsymbol{Y}))$$

来计算

$$\boldsymbol{Y} = (m(\boldsymbol{Y}) - (\nabla^2 f(\boldsymbol{Y}))^{-1} \circ \nabla f(m(\boldsymbol{Y}))) \cap \boldsymbol{Y} \tag{1}$$

下面的定理给出了利用区间 Newton Gauss-Seidel 方法检验区间 $\boldsymbol{Y}$ 内 f 的极值点 Y(或 ∇f 的零点 Y)的存在性与惟一性的标准.

定理 1　设 $\nabla f: \boldsymbol{Y} \subseteq \mathrm{R}^n \to \mathrm{R}^n$ 是 Lipschitz 连续的，$\boldsymbol{Y}_N$ 是 ∇f 在区间 $\boldsymbol{Y}$ 上应用一次 Interval Newton 方法后得到的结果，则有如下的结论：

(i) 对于 ∇f 的任意零点 Y^*，若 $Y^* \in \boldsymbol{Y}$ 则 $Y^* \in \boldsymbol{Y}_N$;

(ii) 如果 $\boldsymbol{Y}_N \cap \boldsymbol{Y} = \emptyset$，则区间 $\boldsymbol{Y}$ 内不包含 ∇f 的零点；

(iii) 如果 $\boldsymbol{Y}_N \subseteq \mathrm{int}(\boldsymbol{Y})$(int 表示区间 $\boldsymbol{Y}$ 的内部)，则区间 $\boldsymbol{Y}$ 包含 ∇f 的零点，并且只包含一个零点.

虽然上述区间 Newton Gauss-Seidel 方法可以验证解的存在性与惟一性，但是，此验证条件是苛刻的，一般情况下很难满足. 因为利用式(1) 计算新区间时，如果出现 $0 \in (\nabla^2 f(\boldsymbol{Y}))^{-1}$，将使用区间扩展除法进行运算，运算的结果再与原区间取交(取交后的区间的某个端点一定包含原区间的某个端点)，因此不会满足解惟一性的标准 $\boldsymbol{Y}_N \subseteq \mathrm{int}(\boldsymbol{Y})$. 下面通过例子说明区间方法的这一缺点.

例 1　验证方程组

$$\begin{cases} f_1 = x_1^2 - x_2 = 0 \\ f_2 = x_1^2 + x_2 = 0 \end{cases}$$

在区间 $\boldsymbol{X} = (x_1, x_2)^T = ([-0.001, 0.001], [-0.001, 0.001])^T$ 上解的存在性与惟一性.

显然，方程组 $F = (f_1, f_2)^T = 0$ 在区间 $\boldsymbol{X}$ 内具有惟一解 $X^* = (0, 0)^T$. 然而，Jacobi 矩阵 $F'(X^*)$ 是奇异的. 所以区间 Newton Gauss-Seidel 方法无法验证解的惟一性. 事实上，F 的区间 Jacobi 矩阵

$$\boldsymbol{F}'(\boldsymbol{X}) = \begin{bmatrix} [-0.002, 0.002] & -1 \\ [-0.002, 0.002] & 1 \end{bmatrix}$$

的中点矩阵

$$Y = m(\boldsymbol{F}'(\boldsymbol{X})) = \begin{bmatrix} 0 & -1 \\ 0 & 1 \end{bmatrix}.$$

此时，$0 \in Y_1 \boldsymbol{F}'(\boldsymbol{X})_1 = [-0.002, 0.002]$，从而导致区间 Newton Gauss-Seidel 方法验证失败.

为了克服这一缺点，本文提出利用 Gröbner 基和特征值的方法来验证区间内解的存在性与惟一性.

2　解的符号验证方法

众所周知，通过 Buchberger 算法计算 Gröbner 基的效率主要依赖于所选择的序. 尤其是对于有些问题，计算字典序的 Gröbner 基需要大量的时间或者输出结果太大. 一般情况下，分级逆字典序的 Gröbner 基的计算时间相对较少，如果没有特殊要求，在计算 Gröbner 基时应首选分级逆字典序. 本文中的算法在计算 Gröbner 基时，只需选择分级逆字典序，因此可以减少计算量.

在文中我们只考虑零维理想，并假设读者熟悉理想和 Gröbner 基的相关理论，只给出一些必要的概念和表示[4~7].

2.1　理想的商代数

设 $G = \{g_1, g_2, \cdots, g_t\}$ 是理想 $I \subset K[x_1, x_2, \cdots,$

$x_n]$ 的 Gröbner 基. 对任意的 $f \in K[x_1, x_2, \cdots, x_n]$, 由 Gröbner 基的性质和多项式的除法算法可知:

$$f = h_1 g_1 + h_2 g_2 + \cdots + h_t g_t + \overline{f}^G,$$

其中余式 $\overline{f}^G$ 是单项式 $X^\alpha \notin \langle \mathrm{LT}(I) \rangle$ ($\langle \mathrm{LT}(I) \rangle$ 表示理想 I 的 Gröbner 基的首项) 的线性组合且由 f 惟一确定, 并且:

(i) $\overline{f}^G = 0 \Leftrightarrow f \in I$,

(ii) $\overline{f}^G = \overline{g}^G \Leftrightarrow f - g \in I$,

(iii) $\overline{f}^G + \overline{g}^G = \overline{f + g}^G$,

(iv) $\overline{\overline{f}^G \circ \overline{g}^G}^G = \overline{fg}^G$,

此除法算法与商环 $K[x_1, x_2, \cdots, x_n]/I$ 有紧密的联系. 给定 $f \in K[x_1, x_2, \cdots, x_n]$, 其陪集定义为 $[f] = f + I = \{f + h : h \in I\}$. 商环 $K[x_1, x_2, \cdots, x_n]/I$ 是由所有 $f \in K[x_1, x_2, \cdots, x_n]$ 的陪集 $[f]$ 组成的. 由上可知 $\overline{f}^G \in [f]$, 并且余式和陪集之间存在一一对应关系, 即: $\overline{f}^G \leftrightarrow [f]$. 因而, 可以把余式 $\overline{f}^G$ 作为它的陪集 $[f] \in K[x_1, x_2, \cdots, x_n]/I$ 的标准表示, 且易验证余式算法恰好是商环 $K[x_1, x_2, \cdots, x_n]/I$ 上的算法, 即:

$$\overline{f}^G + \overline{g}^G = [f] + [g], \quad \overline{\overline{f}^G \circ \overline{g}^G}^G = [f] \circ [g].$$

由于商环 $K[x_1, x_2, \cdots, x_n]/I$ 满足域 K 上的数乘运算, 因此商环 $K[x_1, x_2, \cdots, x_n]/I$ 也是域 K 上的向量空间, 因而, 商环 $K[x_1, x_2, \cdots, x_n]/I$ 为一个代数, 称为商代数, 用 A 表示. 因为 A 中元素的标准表示是单项式 $X^\alpha \notin \langle \mathrm{LT}(I) \rangle$ 的线性组合, 且这些单项式在 A 中是线性无关的, 因而可将它们组成的集合看成 A 的基. 即: 单项式集合 $B = \{X^\alpha : X^\alpha \notin \langle \mathrm{LT}(I) \rangle\}$ 构成了代数 A 的一组基. 把 B 中的元素称为标准单项式.

定理 2(有限维定理)　令 $K \subset \mathbb{C}$ 为一个域, $I \subset K[x_1, x_2, \cdots, x_n]$ 为多项式环 $K[x_1, x_2, \cdots, x_n]$ 的一个理想, 则下列条件等价:

(i) 代数 $A = K[x_1, x_2, \cdots, x_n]/I$ 是有限维的;

(ii) 仿射簇 $V(I) \subset \mathbb{C}^n$ 是有限集;

(iii) 如果 G 是 I 的 Gröbner 基, 则对每个 $i (1 \leqslant i \leqslant n)$, 存在 $m_i \geqslant 0$, 使得对某一 $g \in G$ 有 $x_i^{m_i} = \mathrm{LT}(g)$.

满足上述任一条件的理想称为零维理想.

定义 1　任给多项式 $f \in K[x_1, x_2, \cdots, x_n]$, 其中 K 为代数闭域(如复数域 $\mathbb{C}$), 用乘法定义代数 $A = K[x_1, x_2, \cdots, x_n]/I$ 到自身的映射 $m_f : A \to A$, 即

$$m_f([g]) = [f] \circ [g] = [fg] \in A, \quad \forall [g] \in A.$$

令 I 为零维理想, B 是代数 $A = K[x_1, x_2, \cdots, x_n]/I$ 在某种单项式序下的基. 由于 I 为零维理想, 由定理 2 可知 A 是有限维的, 因此 B 中只含有有限个单项式. 计算 B 中每一元素在映射 m_f 下的像, 从而得到映射 m_f 的矩阵(仍记为 m_f). 矩阵 m_f 具有如下性质:

命题 1　设 f, g 是代数 A 中的元素, 则有

(i) $m_{f+g} = m_f + m_g$,

(ii) $m_{f \circ g} = m_f \circ m_g$.

定义 2　设 $I \subset K[x_1, x_2, \cdots, x_n]$ 为零维理想, B 为代数 $A = K[x_1, x_2, \cdots, x_n]/I$ 在某种单项式序下的基, $f, g \in A$. 用迹(矩阵的迹 Tr 是它的对角线元素之和) 定义一个对称双线性形式 $Q : A \times A \to K$ 如下:

$$Q(f, g) = \mathrm{Tr}(m_f \circ m_g) = \mathrm{Tr}(m_{f \circ g}).$$

则在 $B = \{X^{\alpha(i)}\}$(代数 A 的基) 上定义的对称双线性形式 Q 的矩阵如下:

$$M = (Q(X^{\alpha(i)}, X^{\alpha(j)})) = \mathrm{Tr}(m_{X^{\alpha(i)} \circ X^{\alpha(j)}}).$$

给定非零多项式 $h \in K[x_1, x_2, \cdots, x_n]$, 定义对称双线性形式

$$Q_h(f, g) = \mathrm{Tr}(m_{hf} \circ m_g) = \mathrm{Tr}(m_{hfg}),$$

如此定义的 Q_h 也具有与 Q 一样的性质. 用 M_h 表示 Q_h 相对于 B 的矩阵.

假设 $K \subset \mathbb{R}$(实数域), 则矩阵 M 和 M_h 是实对称矩阵, 从而由线性代数的主轴定理知 M 和 M_h 的所有特征值均为实数. 对称双线性形式具有如下性质:

定理 3　令 I 是由 $K[x_1, x_2, \cdots, x_n]$ ($K \subset \mathbb{R}$) 中的多项式 $(f_1, f_2, \cdots, f_s)$ 生成的零维理想, 则对 $h \in K[x_1, x_2, \cdots, x_n]$, Q_h 的符号差和秩满足:

$$\sigma(Q_h) = \#\{\alpha \in V(I) \cap \mathbb{R}^n : h(\alpha) > 0\} - \#\{\alpha \in V(I) \cap \mathbb{R}^n : h(\alpha) < 0\},$$

$$\rho(Q_h) = \#\{\alpha \in V(I) : h(\alpha) \neq 0\},$$

其中 σ 和 ρ 分别表示 Q_h 的符号差和秩.

在实际计算过程中, 我们只关心特征值的符号, 因此不必求出具体的特征值. 可利用下面的命题计算正特征值的个数:

命题 2　令 M_h 是 Q_h 的矩阵, 并且设 $P_h(t) = \det(M_h - tI)$ 是它的特征多项式, 则 Q_h 正特征值的个数等于 $P_h(t)$ 的系数序列中的变号数(在计算变号数时, 零系数不计).

根据 Descartes 符号法则, 负特征值的个数等于 $P_h(-t)$ 的系数序列中的变号数.

定义 3　令

$$a_{+1}(P; h) = \#\{x \in \mathbb{R}^n : P(x) = 0, h(x) > 0\},$$

$$a_{-1}(P; h) = \#\{x \in \mathbb{R}^n : P(x) = 0, h(x) < 0\},$$

$$a_0(P; h) = \{x \in \mathbb{R}^n : P(x) = 0, h(x) = 0\}.$$

其中 P 为多项式组, h 为多项式.

由定义 3 可得到如下命题.

命题 3

$$\begin{bmatrix}1 & 1 & 1\\ 0 & 1 & -1\\ 0 & 1 & 1\end{bmatrix}\circ\begin{bmatrix}c_{[0]}(P;h)\\ c_{[+]}(P;h)\\ c_{[-]}(P;h)\end{bmatrix}=\begin{bmatrix}\sigma(Q_1)\\ \sigma(Q_h)\\ \sigma(Q_{h^2})\end{bmatrix}\qquad(2)$$

因此，通过计算符号差 $\sigma(Q_1)$、$\sigma(Q_h)$ 和 $\sigma(Q_{h^2})$，并求解(2)，可以得到 $c_{[0]}(P;h)$、$c_{[+]}(P;h)$ 和 $c_{[-]}(P;h)$ 的取值情况.

2.2 计算给定区域内实根个数

给定多项式系统 $P\subset K[x_1,x_2,\cdots,x_k]$，可通过以下方法判断满足 $a<x_1<b$ 的多项式组 P 的实根个数：

(i) 令 $h_1=(b-x_1)(x_1-a)$；

(ii) 按照前面介绍的方法计算 $c_{[+]}(P;h_1)$.

与一维区间类似，多维区间内实根个数的判断方法如下：

(i) 求多项式组所生成理想的根理想 I 的分级逆字典序 Gröbner 基；

(ii) 求代数 $A=\mathrm{R}[x_1,x_2,\cdots,x_n]/I$ 的基 B；

(iii) 对每个变元 $x_i(1\leqslant i\leqslant k)$ 构造区间条件多项式 $h_i(x)=(x_i-a)(b-x_i)$；

(iv) 求 I 的实零点中满足 $h_1(x)>0,h_2(x)>0,\cdots,h_k(x)>0$ 的零点的个数.

令 $S=\{x\in V_k(I):h_1(x)>0,h_2(x)>0,\cdots,h_k(x)>0\}(V_k(I)=V(I)\cap\mathrm{R}^k)$，则我们的最终目标是求 S 中元素的个数，即 $\#(S)$. S 的特征函数如下[8]：

$$\chi_S(x)=2^{-k}\prod_{i=1}^{k}(1+\mathrm{sgn}(h_i(x))),$$

此式可以展开成

$$\chi_S(x)=2^{-k}\sum_{J\subset[1\cdots k]}\prod_{i\in J}\mathrm{sgn}(h_i(x)),$$

若记 $h_J(x)=\prod_{i\in J}h_i(x)$，则

$$\#(S)=\sum_{\alpha\in V_k(I)}\chi_S(\alpha)=$$
$$2^{-k}\sum_{J\subset[1\cdots k]}\sum_{\alpha\in V_k(I)}\mathrm{sgn}(h_J(\alpha)).$$

对于任意的 $J\subset[1\cdots k]$，

$$\sum_{\alpha\in V_k(I)}\mathrm{sgn}(h_J(\alpha))=\#\{\alpha\in V_k(I):h_J(\alpha)>0\}-$$
$$\#\{\alpha\in V_k(I):h_J(\alpha)<0\}=\sigma(Q_{h_J}).$$

因此

$$\#(S)=2^{-k}\sum_{J\subset[1\cdots k]}\sigma(Q_{h_J})\qquad(3)$$

通过计算式(3)的值，可以判断在给定的多维区间内多项式组实零点的存在情况. 并且式(3)的值可以根据前面介绍的方法求得. 但此方法需要计算 2^k 个符号差，效率较低. 实际使用时一般使用 Ben-or, Kozen 和 Reif 提出的算法[4].

3 算 例

例 2 用前面介绍的方法验证例 1 中的问题.

(i) 求多项式组 $\{f_1=x_1^2-x_2,f_2=x_1^2+x_2\}$ 的根理想 $I=\{x_1^2-x_2,x_1^2+x_2,x_1,x_2\}$；

(ii) 构造区域 B 的约束多项式：

$h_i=(x_i+0.001)(0.001-x_i),i=1,2;$

(iii) 求式(3)中的所有符号差. 此例中 $k=2$，所以共有 $2^2=4$ 种不同的条件组合：

(a) 0 个条件：$J=\varnothing,h_J=1,\sigma(Q_{h_J})=1$；

(b) 1 个条件：$J=\{1\},h_J=h_1,\sigma(Q_{h_J})=1$；

$J=\{2\},h_J=h_2,\sigma(Q_{h_J})=1$；

(c) 2 个条件：$J=\{1,2\},h_J=h_1h_2,\sigma(Q_{h_J})=1$.

(iv) 计算式(3). $\#(S)=2^{-2}(1+1+1+1)=1$，因此在区间 $[-0.001,0.001]^2$ 内，目标函数有 1 个零点.

例 3(Schwefel 2.14)

$$\min f=(x_1+10x_2)^2+5(x_3-x_4)^2+(x_2-2x_3)^4+10(x_1-x_4)^4.$$

初始区间 $X=[-4,5]^4$，$\varepsilon=10^{-6}$.

Ibraev 在其博士论文[2]中给出如下结论：

(i) 最小值

$f^*\in[0.000000000000,4.475425668718\times10^{-8}]$；

(ii) f 的最小值点在以下 4 个区间中：

(a) [[−0.009392075036, 0.009317319146],
[−0.000907897950, 0.000939203138],
[−0.002899169922, 0.003141232908],
[−0.002899169922, 0.003141401623]]；

(b) [[0.002319335937, 0.003033963163],
[−0.000289916993, −0.000221252441],
[−0.003233245939, −0.003173828125],
[−0.003233442004, −0.003173828125]]；

(c) [[−0.004477072180, −0.004272460937],
[0.000396728515, 0.000447702853],
[0.002868652343, 0.003184367790],
[0.003143310546, 0.003184542425]]；

(d) [[−0.004516072739, −0.004272460937],
[0.000396728515, 0.000451600668],
[0.003417968750, 0.003612312397],
[0.003417968750, 0.003612576659]].

事实上，在区间 $[-4,5]^4$ 内，目标函数只有一个驻点，上面给出的(b)～(d)区间根本不包含极值点，而产生这些区间的原因就是区间 Newton Gauss-Seidel

方法无法验证这些区间内是否有驻点. 下面用本文提出的符号验证方法验证目标函数在区间$[-4,5]^4$内实驻点的存在惟一性:

(i) 目标函数的一阶偏导数构成的多项式组为:

$$\begin{cases} f_1 = 2x_1 + 20x_2 + 40x_1^3 - 120x_1^2x_4 + \\ \quad 120x_1x_4^2 - 40x_4^3 \\ f_2 = 20x_1 + 200x_2 + 4x_2^3 - 24x_2^2x_3 + \\ \quad 48x_2x_3^2 - 32x_3^3 \\ f_3 = 10x_3 - 10x_4 - 8x_2^3 + 48x_2^2x_3 - \\ \quad 96x_2x_3^2 + 64x_3^3 \\ f_4 = -10x_3 + 10x_4 - 40x_1^3 + 120x_1^2x_4 - \\ \quad 120x_1x_4^2 + 40x_4^3 \end{cases}$$

(ii) 构造区域$[-4,5]^4$的约束多项式:

$h_1 = (x_1+4)(5-x_1)$; $h_2 = (x_2+4)(5-x_2)$;

$h_3 = (x_3+4)(5-x_3)$; $h_4 = (x_4+4)(5-x_4)$.

(iii) 计算式(3)中的所有符号差. 此例中$k=4$, 所以共有$2^4 = 16$种不同的条件组合.

(iv) 计算式(3). $\#(S) = 2^{-4}\sum_{i=1}^{16} 1 = 1$, 因此在区间$[-4,5]^4$内, 目标函数$f$有惟一驻点.

通过此例我们更加看清了区间优化方法的缺点, 从而也显示出了改进方法的必要性.

4 总 结

本文给出了一种求解带边界约束的多变元多项式全局最优解的混合方法. 一方面通过区间算法和二分法克服经典数值优化方法和符号优化方法的缺点; 另一方面通过符号方法解决 Jacobi 矩阵在区间内某点奇异而无法验证驻点的存在性与惟一性的问题; 同时, 通过利用经典数值优化方法(如 BFGS 方法)来有效克服区间优化方法运算速度慢的缺点. 另外, 文中的算法非常有利于并行化, 例如, 我们可以从两方面进行并行化处理: (i)对区间算法得到的不同区间进行并行化符号运算以确定包含最优解的区间; (ii)利用符号方法对某个多维区间的实根个数判断时, 对各个维可用并行计算. 因此, 通过并行化可进一步提高算法效率.

参考文献:

[1] Casado L G, Martínez J A, García I, et al. New interval analysis support functions using gradient information in a global minimization algorithm[J]. Journal of Global Optimization, 2003, 25(4): 345—362.

[2] Ibraev S. A new parallel method for verified global optimization[D]. Genehmigte: Universität Gesamthochschule Wuppertal, 2001.

[3] Hentenryck P V, McAllester D, Kapur D. Solving polynomial systems using a branch and prune approach[J]. SIAM Journal on Numerical Analysis, 1997, 34(2): 797—827.

[4] Ben-Or M, Kozen D, Reif J. The complexity of elementary algebra and geometry[J]. Journal of Computation and Systems, 1986, 32: 251—264.

[5] Cox D, Little J, O′Shea D. Using algebraic geometry[M]. New York: Springer-Verlag, 1998.

[6] Gonalez-Vega L, Rouillier F, Roy M F. Symbolic recipes for polynomial system solving[M]//Some Tapas of Computer Algebra. Berlin Heidelberg: Springer-Verlag, 1999.

[7] Gonalez-Vega L, Rouillier F, Roy M F, et al. Symbolic recipes for real solutions[M]//Some Tapas of Computer Algebra. Berlin Heidelberg: Springer-Verlag, 1999.

[8] Pedrsen P, Roy M F, Szpirglas A. Counting real zeros in the multivariate case[C]//Computational Algebraic Geometry. Eyssette F, Galligo A, eds. Boston: BirkhÄauser, 1993: 203—224.

Hybrid Global Optimization for Multivariate Polynomials by Interval and Symbolic Methods

FENG Ya-li[1], SHENG Xi-sheng[2], XUE Ji-wei[1], YI San-quan[3]

(1. College of Computer and Information, Daqing Petroleum Institute, Daqing 163318, China;

2. School of Mathematical Sciences, Xiamen University, Xiamen 361005, China;

3. Academic Affair Office, Daqing Petroleum Institute, Daqing 163318, China)

Abstract: A hybrid algorithm for global optimization of multivariate polynomials with bound constraints is proposed. Small intervals containing global optima in the presence of rounding errors are obtained by interval methods. Symbolic methods are employed when the Jacobi matrix is singular. Numerical methods are also adopted to improve the efficiency since the interval methods are usually very slow. Furthermore, the performance of our algorithm can be improved by parallelization.

Key words: Global optimization; interval analysis; Gräbner basis; eigenvalue