УДК 519.853.4

ЭФФЕКТИВНЫЙ АЛГОРИТМ РЕШЕНИЯ ВСПОМОГАТЕЛЬНОЙ ЗАДАЧИ В ОДНОРОДНЫХ АЛГОРИТМАХ ГЛОБАЛЬНОЙ ОПТИМИЗАЦИИ

С. М. Елсаков

Исследуется вспомогательная задача в однородных алгоритмах глобальной оптимизации. Найдены локальные максимумы в целевой функции вспомогательной задачи, оценены размеры окрестностей вокруг локальных максимумов, где поиск малоперспективен. Предложен эффективный алгоритм поиска перспективных начальных точек для запуска процедуры локальной оптимизации. Найдено условие при котором локальный поиск позволит решить вспомогательную задачу.

Ключевые слова: многоэкстремальная оптимизация; однородные алгоритмы; вспомогательная задача.

Рассматриваются алгоритмы решения задачи:

$$\underset{x \in X}{absmin}\, f(x), \tag{1}$$

где $f(x): R^d \to R$ – функция удовлетворяющая условию Липшица $|f(x_1) - f(x_2)| \le L\|x_1 - x_2\|$, а $X \subset R^d$ – допустимое множество, которое является компактом. К особенностям задачи (1) можно отнести длительное время вычисления значения целевой функции, а также, как правило, отсутствие аналитического описания функции $f(x)$. Задачи вида (1) часто встречаются на практике [1, 2, 3].

Для решения задачи (1) рассматриваются однородные алгоритмы глобальной оптимизации [2]. Пусть $\{x_i\}_{i=1}^k$ – последовательность точек допустимого множества, в которых осуществляется вычисление значений целевой функции. Будем говорить, что алгоритм относится к классу однородных алгоритмов глобальной оптимизации, если одновременно выполнены три условия:

1. Алгоритм является однородным, т.е. для любых двух целевых функций различающихся на константу последовательности точек испытаний совпадают.

2. Результаты уже проведенных вычислений представляются не в виде таблицы наборов $\{(x_i; f(x_1))\}_{i=1}^k$, а в виде пары функций $m_k(x) = m(x, x_1, f(x_1), ..., x_k, f(x_k))$ и $s_k(x) = s(x, x_1, f(x_1), ..., x_k, f(x_k))$, которые удовлетворяют следующим условиям:

А1. $m_k(x_i) = f(x_i),\ i = \overline{1,k}$;

А2. $s_k(x_i) = 0,\ i = \overline{1,k}$;

А3. $s_k(x) > 0, x \neq x_i, i = \overline{1,k}$;

А4. $m_k(x)$, $s_k(x)$ – липшицевы.

3. Потребуем, чтобы алгоритм мог быть представлен в виде следующего алгоритма:

Шаг 1. Выбрать начальные точки $x_i \in \mathbb{R}^d, i = \overline{1,M}$. Положить $k = M$.

Шаг 2. Вычислить:

$$x_{k+1} = \underset{x \in X}{argabsmin}\, P\left(m_k(x), s_k(x)\right), \tag{2}$$

где $P\left(m_k(x), s_k(x)\right): R^2 \to R$ – функция выбора, которая единственна для алгоритма, $m_k(x), s_k(x)$ – функции, которые перестраиваются после каждой итерации алгоритма и удовлетворяют условиям А1–А4.

Шаг 3. Вычислить значения $f(x_{k+1})$, положить $k = k+1$.

Шаг 4. Если выполнено условие выхода ϕ, то выйти, иначе вернуться на Шаг 2.

Для однородных алгоритмов справедлива теорема [2]:

Теорема 1. *Если для функции* $s_k(x)$ *выполняются условия*

А5. $m\left(x, x_1, f(x_1)+c, ..., x_k, f(x_k)+c\right) = m\left(x, x_1, f(x_1), ..., f(x_k)\right)+c,\ k \in \mathbb{N}, c \in \mathbb{R}$;

А6. $s\left(x, x_1, f(x_1)+c, ..., x_k, f(x_k)+c\right) = s\left(x, x_1, f(x_1), ..., x_k, f(x_k)\right),\ k \in \mathbb{N}, c \in \mathbb{R}$,

А7. $s_k(x) \geq \min\limits_{i=\overline{1,k}} \|x - x_i\|$,

то множество предельных точек порождаемых алгоритмом с критерием:

$$P\left(s_k(x), m_k(x)\right) = m_k(x) - 2Ks_k(x), \tag{3}$$

где $K > L + L_m$, *будет совпадать с множеством глобальных минимумов липшицевой функции* $f(x)$ *с константой Липшица* L.

При практической реализации однородных алгоритмов возникают три сложности: требуется алгоритм для построения функций $m_k(x)$ и $s_k(x)$, требуется достаточно быстро решать задачу (2), требуется каким-либо образом выбирать коэфициент K.

Быстрый алгоритм построения функций $m_k(x)$ и $s_k(x)$ можно построить следующим образом. Пусть допустимое множество X разбито на подмножества $X = X_1 \bigcup X_2 \bigcup \ldots \bigcup X_M$, таким образом, что $X_i \bigcap X_j \neq \varnothing$, $i, j = \overline{1,M}$ и количество точек испытаний внутри каждого подмножества отлично от нуля. Обозначим границу множества X как ∂X, а множеств X_i как ∂X_i. Введем расстояние до внутренней границы подмножества как $\rho_i(x) = \min\limits_{y \in \partial X_i \backslash \partial X} \|x - y\|$.

Теорема 2. *Пусть существует алгоритм построения функций* $m_k(x)$ *и* $s_k(x)$*, удовлетворяющих условиям А1)–А7) и разбиение таково, что для любой точки* $x \in X$*, которая принадлежит более чем одному подмножеству* X_i *выполняется неравенство* $\sum\limits_{x \in X_i} \rho_i(x) > \delta > 0$*, тогда функции:*

$$m_k(x)' = \sum_{x \in X_i} \rho_i(x) m_k^i(x) / \sum_{x \in X_i} \rho_i(x),$$
$$s_k(x)' = \sum_{x \in X_i} \rho_i(x) s_k^i(x) / \sum_{x \in X_i} \rho_i(x),$$

где $m_k^i(x)$, $s_k^i(x)$ *– функции* $m_k(x)$ *и* $s_k(x)$ *построенные для испытаний, которые располагаются тлько внутри подмножества* X_i, *также удовлетворяют условиям А1)–А7).*

Теорема 3. *Для того чтобы решить задачу глобальной оптимизации (1) с точностью* ε *по значению целевой функции с помощью однородного алгоритма (для функций* $m_k(x)$ *и* $s_k(x)$ *выполняются условия А1–А7) достаточно на каждом шаге решать вспомогательную задачу оптимизации (2) таким образом, чтобы выполнялось условие:*

$$P_k(x_{k+1}) \le \min_{i=\overline{1,k}} f(x_i) - \varepsilon. \tag{4}$$

Теорема 3 позволяет на каждом шаге искать следующую точку с помощью алгоритмов локальной оптимизации, не решая задачу (2) точно. И только при выходе требуется решить задачу (2) точно.

Для подбора коэфициента K известно несколько подходов [2]. Нами предлагается подход основанный на подходе адаптивного оценивания коэфициента K. Предлагается ввести некий коэфициент достоверности r с начальным значением меньше 1, например 0.5. Затем решить задачу (1) с $K = r * \max \frac{|f(x_j) - f(x_i)|}{\|x_j - x_i\|}$, после чего значение r увеличивается допустим на 0,1. Это позволяет найти глобальный минимум при минимальном r, как правило, за наименьшее количество операций.

Рассмотрим задачу (2):

$$m_k(x) - 2Ks_k(x) \to \min_{x \in X}, \tag{5}$$

где значения целевой функции в точках $\{x_i\}_{i=1}^k$ равны, соответственно, $\{f(x_i)\}_{i=1}^k$. В соответствии с теоремой 3 нам не требуется на каждом шаге находить точной решение, требуется только найти значение удовлетворяющее неравенству $P_k(x_{k+1}) \le \min_{i=\overline{1,k}} f(x_i) - \varepsilon$.

Исследуем эту задачу более подробно.

Лемма. *Пусть* $x \in S_r(x_i)$, *где* $S_r(x_i)$ *– замкнутый шар с центром в точке* x_i *и радиусом:*

$$r = 0.5 * \min_{\substack{j=\overline{1,k} \\ j \ne i}} \|x_i - x_j\|. \tag{6}$$

Тогда внутри шара:

$$\|x - x_i\| = \min_{i=\overline{1,k}} \|x - x_i\|. \tag{7}$$

Доказательство:

Для каждой точки x существует некий номер j, такой что $\|x-x_j\| = \min\limits_{i=\overline{1,k}} \|x-x_i\|$. Поскольку точка x принадлежит шару, то $\|x-x_i\| \le r$. Покажем, что для любой точки x_j отличной от x_i выполняется неравенство:

$$\|x-x_i\| \le \|x-x_j\|. \tag{8}$$

Для этого рассмотрим расстояние $\|x_j - x_i\|$. По условиям леммы:

$$2\|x-x_i\| \le 2r = \min_{\substack{j=\overline{1,k}\\ j\ne i}} \|x_i - x_j\| \le \|x_j - x_i\|. \tag{9}$$

Запишем неравенство треугольника:

$$\|x_j - x_i\| \le \|x_j - x\| + \|x - x_i\|. \tag{10}$$

Из неравеств (9) и (10) следует, что:

$$2\|x-x_i\| \le \|x_j - x\| + \|x - x_i\|, \tag{11}$$
$$\|x-x_i\| \le \|x_j - x\|.$$

Следовательно, (7) верно.

Теорема 4. *Функция* $m_k(x) - 2Ks_k(x)$ *имеет локальные максимумы в точках* $x_i, i = \overline{1,k}$.

Доказательство

Рассмотрим разность $m_k(x) - m_k(x_i)$, поскольку функция $m_k(x)$ липшицева, то выполняется неравенство:

$$m_k(x) - m_k(x_i) \le L_m\|x-x_i\|. \tag{12}$$

С учетом того, что $K > L + L_m$, можно записать:

$$L_m x - x_i \le 2K\|x-x_i\|. \tag{13}$$

Пусть $x \in S_r(x_i)$, где $S_r(x_i)$ – замкнутый шар с центром в точке x_i и радиусом $r = 0.5 * \min\limits_{\substack{j=\overline{1,k}\\ j\ne i}} \|x_i - x_j\|$. Тогда, на основе Леммы, в пределах шара будет выполняться соотношение $\|x-x_i\| = \min\limits_{i=\overline{1,k}} \|x-x_i\|$ и, воспользуясь неравенством, (A7) получим:

$$2K\|x-x_i\| = 2K \min_{i=\overline{1,k}} \|x-x_i\| \le 2Ks_k(x). \tag{14}$$

Из неравенств (12)–(14) следует:

$$m_k(x) - m_k(x_i) \le 2Ks_k(x). \tag{15}$$

С учетом того, что $s_k(x_i) = 0$, можно преобразовать неравенство (15):

$$m_k(x) - 2Ks_k(x) \le m_k(x_i) - 2Ks_k(x_i). \tag{16}$$

Неравенство (16) свидетельствует, о существовании окрестности вокруг точки x_i радиусом r внутри которой выполняется условие локального максимума.

Покажем, что значения функции $m_k(x) - 2Ks_k(x)$ достаточно большие в окрестностях точек x_i.

Рассмотрим шар $S_r(x_i)$ – замкнутый шар с центром в точке x_i и радиусом $r = 0.5 * \min\limits_{\substack{j=\overline{1,k} \\ j \neq i}} \left\| x_i - x_j \right\|$. На границе этого шара будет выполняться неравенство:

$$m_k(x) - 2Ks_k(x) - m_k(x_i) \leq -2Kr + L_m r, \tag{17}$$

где L_m – константа Липшица для функции $m_k(x)$. Учитывая, что $K > L + L_m$, запишем:

$$-2Kr + L_m r < r(-2L - 2L_m + L_m) = -r(2L + L_m). \tag{18}$$

Таким образом на границе шара значение функции $m_k(x) - 2Ks_k(x)$ меньше, чем в центре шара на $r(2L + L_m)$. Поскольку функция $m_k(x) - 2Ks_k(x)$ является липшицевой, то в каждой точке x_i существует шар меньшего радиуса, значения в котором заведомо больше, чем $m_k(x_i) - r(2L + L_m)$. Радиус этого шара R может быть найден из условия Липшица:

$$r(2L + L_m) \leq \left| m_k(x) - 2Ks_k(x) - m_k(x_i) \right| \leq, \tag{19}$$

запишем условие Липшица:

$$\leq (L_m + 2KL_s)\left\| x - x_i \right\| = (L_m + 2KL_s)R. \tag{20}$$

Выразим радиус R:

$$R \geq r \frac{2L + L_m}{L_m + 2KL_s}. \tag{21}$$

С учетом, того, что $K > L + L_m$ и $L_s \geq 1$, то правая часть неравенства (21) может изменяться в пределах:

$$0 \leq r \frac{2L + L_m}{L_m + 2KL_s} \leq r. \tag{22}$$

Таким образом для решения задачи (2) требуется искать локальные минимумы функции $m_k(x) - 2Ks_k(x)$, которые лежат как можно дальше от точек $x_i, i = \overline{1,k}$. Точное решение этой задачи представляется достаточно сложной задачей, однако, в нашем случае нет необходимости искать точное решение, достаточно указать любую точку из «области притяжения» локального минимума и из этой точки спустить к локальному минимуму. Рассмотрим возможный подход к приблизительному решению этой задачи.

Пусть в качестве допустимого множества X выбрано множество $X = [0;1]^d$. Построим 2^d-арное дерево следующим образом. Корню дерева сопоставим всё множество X. Дальнейшее сопоставление описывается рекурсивной формулой. Пусть узлу сопоставлено некоторое множество $[a;b]^d$, где $a,b \in [0;1]^d$. Тогда потомку этого узла с индексом $i \in \{0,1,\ldots 2^d - 1\}$ сопоставим множество $[u;v]^d$, где:

$$u_j = \begin{cases} a_j, если\, i\, and\, 2^j = 0; \\ \dfrac{a_j + b_j}{2}, если\, i\, and\, 2^j = 2^j; \end{cases} \qquad (23)$$
$$v_j = \begin{cases} \dfrac{a_j + b_j}{2}, если\, i\, and\, 2^j = 0; \\ b_j, если\, i\, and\, 2^j = 2^j. \end{cases}$$

Высота дерева H является параметром алгоритма и задается пользователем. Каждому узлу сопоставим целое число (индекс узла) – количество листьев в поддереве максимальной высоты h этого узла, не содержащем ни одной точки проведенного испытания. Описанная структура данных позволяет эффективно добавлять информацию о проведенных вычислениях в дерево и эффективно определять поддерево максимальной высоты свободное от точек испытаний. Время выполнения этих операций фиксировано и составляет $O(H)$. Дерево высоты H содержит:

$$\frac{2^{Hd} - 1}{2^d - 1} \qquad (24)$$

узлов.

Теорема 5. Если найдено поддерево высоты $h \ge \log_2\left(\left(\max\limits_{i=\overline{1,k}} f(x_i) - \min\limits_{i=\overline{1,k}} f(x_i) + \varepsilon\right) / (2L + L_m)\right) + H + 1$, то вспомогательная задача решена с точностью ε.

Доказательство

Пусть высота максимального поддерева h. Тогда можно вписать шар диаметром 2^{h-H} с центром в точке $x_ш$ таким образом, чтобы внутри шара не оказалось ни одной точки уже проведенного испытания. Т.е. найдена точка:

$$\min_{i=\overline{1,k}} \|x_ш - x_i\| \ge 2^{h-H-1}. \qquad (25)$$

Для того, чтобы задача была решена требуется:

$$m_k(x_ш) - 2Ks_k(x_ш) \le \min_{i=\overline{1,k}} f(x_i) - \varepsilon. \qquad (26)$$

Воспользуемся оценкой из предыдущей теоремы:

$$m_k(x_ш) - 2Ks_k(x_ш) - m_k(x_i) \le -\min_{i=\overline{1,k}} \|x_ш - x_i\| (2L + L_m) \le \qquad (27)$$

с учетом неравенства (25):

$$\le -2^{h-H-1}\left(2L+L_m\right). \tag{28}$$

Объединяя неравенства (27) и (28):

$$m_k\left(x_ш\right)-2Ks_k\left(x_ш\right)-m_k\left(x_i\right)\le -2^{h-H-1}\left(2L+L_m\right). \tag{29}$$

Выражая значение функции $m_k\left(x\right)-2Ks_k\left(x\right)$ в центре шара:

$$m_k\left(x_ш\right)-2Ks_k\left(x_ш\right)\le m_k\left(x_i\right)-2^{h-H-1}\left(2L+L_m\right)\le \tag{30}$$

$$m_k\left(x_i\right)-2^{h-H-1}\left(2L+L_m\right)\le \max_{i=\overline{1,k}} f\left(x_i\right)-2^{h-H-1}\left(2L+L_m\right). \tag{31}$$

Из неравенст (30) и (31) следует:

$$m_k\left(x_ш\right)-2Ks_k\left(x_ш\right)\le \max_{i=\overline{1,k}} f\left(x_i\right)-2^{h-H-1}\left(2L+L_m\right). \tag{32}$$

Задача будет решена, если будет выполнено следующее неравенство:

$$\max_{i=\overline{1,k}} f\left(x_i\right)-2^{h-H-1}\left(2L+L_m\right)\le \min_{i=\overline{1,k}} f\left(x_i\right)-\varepsilon. \tag{33}$$

Преобразуем:

$$\max_{i=\overline{1,k}} f\left(x_i\right)-\min_{i=\overline{1,k}} f\left(x_i\right)+\varepsilon\le 2^{h-H-1}\left(2L+L_m\right),$$

$$\frac{\max_{i=\overline{1,k}} f\left(x_i\right)-\min_{i=\overline{1,k}} f\left(x_i\right)+\varepsilon}{2L+L_m}\le 2^{h-H-1},$$

$$\log_2\left(\frac{\max_{i=\overline{1,k}} f\left(x_i\right)-\min_{i=\overline{1,k}} f\left(x_i\right)+\varepsilon}{2L+L_m}\right)\le h-H-1.$$

Выражая высоту найденного поддерева:

$$h\ge \log_2\left(\frac{\max_{i=\overline{1,k}} f\left(x_i\right)-\min_{i=\overline{1,k}} f\left(x_i\right)+\varepsilon}{2L+L_m}\right)+H+1. \tag{34}$$

Таким образом выполнение неравенства (34) влечет выполнение неравенства (26), что означает возможность решения вспомогательной задачи с точностью ε.

Библиографический список

1. Елсаков, С.М. Однородные алгоритмы многоэкстремальной оптимизации / С.М. Елсаков, В.И. Ширяев // Ж. вычисл. матем. и матем. физ. – 2010. – № 10. – С. 1727–1740.

2. Елсаков, С.М. Однородные алгоритмы многоэкстремальной оптимизации для целевых функций со значительным временем вычисления значения / С.М. Елсаков, В.И. Ширяев // Вычислительные методы и программирование. – 2011. – № 1. – С. 52–73.

3. Elsakov, S.M. Linear Homogeneous Algorithms of Global Optimization / S.M. Elsakov, V.I. Shiryaev // Global Optimization: Theory, Methods & Applications. Series Lecture notes in decision science. Vol. 12. – Hong Kong, London, Tokyo: Global-Link Publisher, 2009. – Pp. 241–247.