УДК 519.853.4

Алгоритм глобальной оптимизации с использованием численных оценок производной минимизируемой функции*

Баркалов К. А., Сысоев А. В., Ямщиков И. С.

Нижегородский государственный университет им. Н.И. Лобачевского

В работе рассматривается задача поиска глобального минимума x^* одномерной функции $\varphi(x)$ вида

$$\varphi(x^*) = \min \{\varphi(x) : x \in [a, b]\}. \tag{1}$$

Предполагается, что целевая функция является многоэкстремальной и задана в виде "черного ящика", т.е. некоторого алгоритма вычисления ее значений.

В задачах вида (1) получение гарантированных оценок глобального минимума возможно только при наличии тех или иных предположений о поведении минимизируемой функции $\varphi(x)$. Одним из наиболее часто используемых предположений является выполнимость условия Липщица

$$|\varphi(x_1) - \varphi(x_2)| \le L\,|x_1 - x_2|\,,\ x_1, x_2 \in [a, b],\ 0 < L < \infty.$$

Условие Липщица соответствует предположению ограниченности изменения значения функции при ограниченном изменении ее аргумента. Данное условие позволяет строить оценки возможного поведения функции $\varphi(x)$ на основе конечного множества ее значений, вычисленных в точках области поиска $[a, b]$.

Более сильным предположением о характере поведения функции является липшицевость ее первой производной

$$\left|\varphi'(x_1) - \varphi'(x_2)\right| \le L_1\,|x_1 - x_2|\,,\ x_1, x_2 \in [a, b],\ 0 < L_1 < \infty. \tag{2}$$

Выполнение условия (2) позволяет формировать более точные оценки возможных значений минимизируемой функции $\varphi(x)$, что обеспечивает возможность существенного повышения эффективности разрабатываемых алгоритмов. Однако при решении задач оптимизации с функциями вида "черный ящик" точные значения производных, как правило, недоступны, а их численная оценка (с помощью операторов численного дифференцирования) требует дополнительных вычислений значений функции. При этом даже одно вычисление значения функции (*испытание*) является трудоемкой операцией, т.к. связано с проведением численного моделирования.

Следовательно, перспективной является разработка методов глобальной оптимизации, в которых необходимые значения производных вычисляются на основе точек предшествующих итераций, без проведения дополнительных поисковых испытаний.

В работе [1] был сформулирован алгоритм глобального поиска с использованием производных для решения одномерных задач. Данный метод в процессе своей работы предполагает построение последовательности точек $x^k \in [a, b]$, в которых проводятся испытания.

*Исследование выполнено при финансовой поддержке РФФИ (проект 19-07-00242), научно-образовательного математического центра (075-02-2020-1483/1) и государственного задания (0729-2020-0055).

Множество троек $\{x^i, \varphi^i = \varphi(x^i), \varphi'^i = \varphi'(x^i)\}, 0 \le i \le k$, составляет поисковую информацию, накопленную методом после проведения k шагов.

Общая схема алгоритма может быть представлена в следующем виде. Первые два испытания проводятся в граничных точках области поиска, т.е. $x^0 = a, x^1 = b$. Для выбора точки $x^{k+1}, k > 1$, очередного испытания требуется выполнить следующие действия.

1. Перенумеровать точки $x^0, ..., x^k$ предшествующих испытаний нижним индексом в порядке возрастания координаты, т.е.

$$a = x_0 < x_1 < ... < x_k = b, \tag{3}$$

и сопоставить им значения $\varphi_i = \varphi(x_i), \varphi'_i = \varphi'(x_i)\}, 0 \le i \le k$, вычисленные в этих точках.

2. Для каждого интервала $(x_{i-1}, x_i), 1 \le i \le k$, вычислить его *характеристику* $R(i)$.

3. Определить интервал с максимальной характеристикой

$$R(t) = \max\{R(i), 1 \le i \le k\}.$$

4. Провести следующее испытание в точке интервала с максимальной характеристикой, т.е. $x^{k+1} \in (x_{t-1}, x_t)$.

5. Проверить условие остановки $x_t - x_{t-1} < \epsilon$, где t – номер интервала с максимальной характеристикой, а $\epsilon > 0$ – заданная точность.

Конкретные формулы для вычисления характеристики интервала и точки проведения нового испытания, а также соответствующие теоретические утверждения о сходимости алгоритма представлены в [1].

В рамках проведенного исследования был предложен алгоритм, в котором точные значения производных заменяются их численными оценками, вычисленными на основе ранее накопленной поисковой информации. В соответствии с нумерацией (3) точек испытаний в порядке возрастания координаты в качестве численных оценок значений производных $\overline{\varphi_i}' = \overline{\varphi}'(x_i) \approx \varphi'(x_i)$, $0 \le i \le k$ можно использовать трехточечную аппроксимацию вида

$$\begin{gathered}
\overline{\varphi_0}' = \frac{1}{H_1}\left(-(2+\delta_2)\varphi(x_0) + \frac{(\delta_2+1)^2}{\delta_2}\varphi(x_1) - \frac{1}{\delta_2}\varphi(x_2)\right), \\
\overline{\varphi_i}' = \frac{1}{H_i}\left(-\delta_{i+1}\varphi(x_{i-1}) - \frac{\delta_{i+1}^2 - 1}{\delta_{i+1}}\varphi(x_i) + \frac{1}{\delta_{i+1}}\varphi(x_{i+1})\right),\ 0 < i < k, \\
\overline{\varphi_k}' = \frac{1}{H_{k-1}}\left(\delta_k\varphi(x_{k-2}) - \frac{(\delta_k+1)^2}{\delta_k}\varphi(x_{k-1}) + \frac{2+\delta_k}{\delta_k}\varphi(x_k)\right),
\end{gathered} \tag{4}$$

где $H_i = h_i + h_{i+1}$, $\delta_{i+1} = \frac{h_{i+1}}{h_i}$, $h_i = x_i - x_{i-1}$.

Таким образом, новый алгоритм представляет собой дальнейшее развитие метода из [1], в котором значения первой производной минимизируемой функции заменяются их численные оценками вида (4). Отметим, что в алгоритме вместо априори неизвестного значения константы Липшица L_1 для $\varphi'(x)$ используется ее адаптивная оценка m, вычисляемая на основе полученной поисковой информации. Справедлива следующая теорема о сходимости алгоритма.

Теорема 1. *Пусть $\{x^k\}$ есть последовательность точек испытаний, порождаемой алгоритмом при минимизации функции $\varphi(x), x \in [a, b]$, производная которой удовлетворяет условию Липщица с константой L_1, и пусть на некотором шаге поиска k для адаптивной оценки константы Липшица m справедливо неравенство*

$$rm > \gamma L_1,\ \gamma = \left[4(b-a) + (b-a)^2 + 0.5\right]/\epsilon^2,$$

где $r > 1$ – параметр, а $\epsilon > 0$ – точность в условии остановки алгоритма. Тогда любая точка глобального минимума x^* является предельной точкой последовательности $\{x^k\}$ и, кроме того, любая предельная точка $\overline{x}$ этой последовательности является точкой глобального минимума функции $\varphi(x)$.

Для демонстрации эффективности предложенного алгоритма глобального поиска с использованием численных оценок производных (АГПЧП) было проведено сравнение со следующими методами глобальной оптимизации: метод Гальперина (МГ) [2], метод Пиявского (МП) [3], алгоритм глобального поиска Стронгина (АГП) [4], алгоритм глобального поиска с использованием производных (АГПП) [1]. Сравнение было проведено при решении серии из 20 многоэкстремальных задач, использованных для оценки эффективности методов глобальной оптимизации в [1].

В проведенных экспериментах точность поиска решения составила $\epsilon = 10^{-4}(b-a)$, где a и b – границы области поиска. Для методов Гальперина и Пиявского использовались точные значения константы Липшица минимизируемой функции, оцененные предварительно. Для алгоритма глобального поиска Стронгина использовалось значение параметра $r = 2$. Для методов с использованием производных использовался параметр $r = 1.1$. В таблице 1 отражено среднее число итераций (округленное до целых значений), потребовавшихся методам для решения всех задач серии.

Таблица 1. Результаты решения серии задач.

	МГ	МП	АГП	АГПП	АГПЧП
Среднее число итераций	720	314	244	26	24

Результаты экспериментов наглядно демонстрируют превосходство (по числу итераций) методов с использованием производных над методами, в которых производные не используются. Одновременно с этим видно, что методам АГПП и АГПЧП требуется примерно одинаковое число итераций. При этом каждая итерация АГПП подразумевает вычисление точного значения производной, тогда как в АГПЧП никакие дополнительные вычисления значений функции и ее производных не требуются.

Направлением дальнейших исследований будет являться исследование свойств предложенного алгоритма при решении многомерных задач многоэкстремальной оптимизации. Обобщение на многомерные задачи может быть выполнено, например, с использованием схемы диагонального разбиения области поиска [5] или же с использованием схемы рекурсивной (вложенной) оптимизации [6].

Литература

1. Гергель В. П. Об одном способе учета значений производных при минимизации многоэкстремальных функций // Ж. вычисл. матем. и матем. физ. 1996. Т. 36, № 5. С. 51-67.

2. Galperin E. A. The cubic algorithm. J. Math. Anal. Appl. 1985. vol. 112. pp. 635-640.

3. Пиявский С. А. Один алгоритм отыскания абсолютного экстремума функций // Журн. вычисл. матем. и матем. физ. 1972. Т. 12. № 4. С. 888-896.

4. Стронгин Р. Г., Гергель В. П., Гришагин В. А., Баркалов К. А. Параллельные вычисления в задачах глобальной оптимизации. М.: Изд-во МГУ, 2013. 280 с.

5. Сергеев Я. Д., Квасов Д. Е. Диагональные методы глобальной оптимизации. М.: Физматлит, 2008. 352 с.

6. Городецкий С. Ю., Гришагин В. А. Нелинейное программирование и многоэкстремальная оптимизация. Н. Новгород: Изд-во ННГУ, 2007. 489 с.

MSC2020 90C26

Global optimization algorithm using numerical estimates of the objective function derivative

K.A. Barkalov, A.V. Sysoyev, I.S. Yamshchikov

Lobachevsky State University of Nizhni Novgorod