ДОКЛАДЫ РОССИЙСКОЙ АКАДЕМИИ НАУК. МАТЕМАТИКА, ИНФОРМАТИКА, ПРОЦЕССЫ УПРАВЛЕНИЯ, 2023, том 512, с. 78–80

МАТЕМАТИКА

УДК 519.615

МЕТОД ЛОКАЛИЗАЦИИ ФИКТИВНЫХ ЭКСТРЕМУМОВ В ЗАДАЧЕ ГЛОБАЛЬНОЙ ОПТИМИЗАЦИИ

Академик Ю. Г. Евтушенко[1,2,*], А. А. Третьяков[1,3,]**

Поступило 19.04.2023 г.
После доработки 04.07.2023 г.
Принято к публикации 13.07.2023 г.

Рассматривается задача поиска глобального экстремума неотрицательной функции на положительном параллелепипеде в n-мерном евклидовом пространстве. Предложен метод локализации фиктивных экстремумов в ограниченной области вблизи начала координат, что позволяет отделить точку глобального экстремума от фиктивных экстремумов путем отбрасывания его на существенное расстояние от множества локализации фиктивных минимумов. При этом за счет выбора начальной точки в методе градиентного спуска удается обосновать сходимость итерационной последовательности к глобальному экстремуму минимизируемой функции.

Ключевые слова: глобальный экстремум, локальный минимум, градиентный метод, сходимость

DOI: 10.31857/S2686954323600222, **EDN:** PNRHVU

Рассматривается задача

$$\min_{x\in\Pi}\varphi(x), \tag{1}$$

где $\Pi = \{x\in\mathbb{R}^n \,|\, a^i \le x^i \le b^i, i = 1,\dots,m\}$. При этом считаем, что в точке минимума $x^* = \arg\min_{x\in\Pi}\varphi(x)$, $\varphi(x^*) = 0$ и $a^i > 0$, $i = 1,\dots,m$.

Данные предположения не являются принципиальными, но облегчают описание предложенного подхода решения таких задач. Пусть $\{x_j^*\}$, $j\in P$, множество локальных минимумов для задачи (1), $P = \{1,\dots,p\}$, и глобальный минимум достигается при $j = j_0$ в точке $x_{j_0}^* = x^*$. В общем случае данная задача является многоэкстремальной. Один из способов ее решения – это классический градиентный метод [1], схема которого имеет следующую структуру

$$x_{k+1} = x_k - \alpha_k\varphi'(x_k),\quad x_0\in\mathbb{R}^n, \tag{2}$$

и при определенных условиях на выбор шага α_k (например, [1, 2]),

$$\begin{gathered}\varphi(x_k) - \varphi(x_{k+1}) \ge q_k\alpha_k\|\varphi'(x_k)\|^2,\\ q_k\in\left(\frac{1}{2},1\right),\quad k = 0,1,\dots,\end{gathered} \tag{3}$$

последовательность x_k сходится к некоторому локальному минимуму $x_j^*(x_0)$ (см. [1]). Проблема состоит в том, что локальных минимумов может быть много ($j\in P$) и, в зависимости от начальной точки x_0, мы получим тот или иной локальный экстремум $x_j^*(x_0) \ne x^*$. Учитывая свойства градиентного метода, естественно предположить, что

– *итерационная последовательность* (2)*, определенная начальной точкой* x_0*, будет сходиться к ближайшему локальному минимуму*, *т.е.*

$$x_j^*(x_0) = \arg\min_{j\in P}\|x_0 - x_j^*\|. \tag{4}$$

Предположение (4) весьма естественно, и мы примем его в качестве отправного.

Отметим, что функция $\varphi(x)$ определена только в прямоугольнике Π, но мы можем ее доопределить гладким образом за границы прямоугольника (параллелепипеда) Π так, что

$$\varphi(x) \ge \varphi(\mathrm{Pr}_\Pi(x)) + \|x - \mathrm{Pr}_\Pi(x)\|^2,$$

где $\mathrm{Pr}_\Pi(\cdot)$ – оператор проектирования на Π, т.е. $\varphi(x)$ будет квадратично возрастать вне прямоугольника Π и заведомо вне Π не будет локальных экстремумов у доопределенной функции $\varphi(x)$:

$$\varphi'(x) \ne 0,\quad x\in\mathbb{R}^n\backslash\Pi, \tag{5}$$

[1] *Федеральный исследовательский центр “Информатика и управление” Российской академии наук, Москва, Россия*
[2] *Московский физико-технический институт (национальный исследовательский университет), Долгопрудный, Московская обл., Россия*
[3] *Siedlce University, Faculty of Sciences, Siedlce, Poland*
**E-mail: yuri-evtushenko@yandex.ru*
***E-mail: prof.tretyakov@gmail.com*

и значения $\varphi(x)$ при $x \not\in \Pi$ будут больше, чем внутри прямоугольника Π.

Новизна метода состоит в том, что делается замена координат

$$y = \frac{x}{\varphi(x) + \varepsilon} \tag{6}$$

и минимизируется функция $\overline{\varphi}(y) = \varphi(x(y))$ во всем пространстве $\mathbb{R}^n$, где $x(y)$ – решение системы (6). Здесь $\varepsilon > 0$ достаточно малое число, при котором фиктивные (локальные) минимумы $y_j^* = \dfrac{x_j^*}{\varphi(x_j^*) + \varepsilon}$, $j \in P\backslash\{j_0\}$, функции $\overline{\varphi}(y)$ остаются в ограниченном параллелепипеде

$$\overline{\Pi} = \{y \in \mathbb{R}^n \,|\, \overline{a}^i \le y^i \le \overline{b}^i, i = 1,\ldots,m\},$$

где $\overline{a}^i = \dfrac{a^i}{M + \varepsilon}$, $\overline{b}^i = \dfrac{b^i}{m + \varepsilon}$, $M = \max\limits_{j \in P\backslash\{j_0\}} \varphi(x_j^*)$, $m = \min\limits_{j \in P\backslash\{j_0\}} \varphi(x_j^*)$, а глобальный минимум $y_j^* = y^* = \dfrac{x^*}{\varphi(x^*) + \varepsilon} = \dfrac{x^*}{\varepsilon}$ отбрасывается от области $\overline{\Pi}$, в которой локализуются фиктивные экстремумы y_j^*, $j \in P\backslash\{j_0\}$, т.е.

$$\|y^*\| \ge N = \frac{1}{\varepsilon},$$

где N достаточно велико и $N \to \infty$ при $\varepsilon \to 0$. Причем в новых переменных функция $\overline{\varphi}(y)$ достигает минимума в точке $y^* = \dfrac{x^*}{\varepsilon}$, т.е. $\overline{\varphi}'(y^*) = 0$. В дальнейшем, при использовании наших построений, потребуется решать одномерное уравнение (см. ниже (8)) на каждом шаге итерационного процесса, что является платой за возможность получения (вычисления) глобального минимума в пространстве $\mathbb{R}^n$ и, на наш взгляд, является несравненно более элементарной процедурой при решении n-мерной задачи. Поэтому мы будем использовать задачу (8) как вспомогательную и несравнимо более простую, не требующую особых вычислительных затрат. При этом для всех фиктивных экстремумов y_j^*, $j \in P\backslash\{j_0\}$, будет выполнено неравенство $\|y_j^*\| \le C$, где константа C не зависит от N, а $\|y^*\| = \|y_{j_0}^*\| \ge N = \dfrac{1}{\varepsilon}$ при N достаточно больших. Если взять начальную точку $y_0 = (2N,\ldots,2N)^\top$, то будет выполнено соотношение (4), так как

$$\|y_0 - y^*\| \sim N \cdot \sqrt{n - \frac{3}{4}} < 2N\sqrt{n} \sim \|y_0 - y_j^*\|, \quad j \in P\backslash\{j_0\}$$

при N достаточно больших. Здесь мы учитываем, что $y^* > 0$, ибо $x^* > 0$. Таким образом, градиентный метод

$$y_{k+1} = y_k - \alpha_k \overline{\varphi}'(x(y_k)), \quad k = 0,1,\ldots, \tag{7}$$

$y_0 = (2N,\ldots,2N)^\top$ с выбором шага по формуле (3), согласно предположению (4), будет сходиться к точке глобального минимума y^* функции $\overline{\varphi}(y) = \varphi(x(y))$, а значит к x^*.

При этом $\overline{\varphi}'(y) = \varphi_x'(x(y)) \cdot x'(y)$, где $x(y)$ берем как решение уравнения (6) и по теореме о неявной функции

$$x'(y) = -\begin{pmatrix} \dfrac{\varphi(x) + \varepsilon - x^1 \dfrac{\partial\varphi(x)}{\partial x^1}}{(\varphi(x)+\varepsilon)^2} & \dfrac{-x^1 \dfrac{\partial\varphi(x)}{\partial x^2}}{(\varphi(x)+\varepsilon)^2} & \cdots & \dfrac{-x^1 \dfrac{\partial\varphi(x)}{\partial x^n}}{(\varphi(x)+\varepsilon)^2} \\ \dfrac{-x^2 \dfrac{\partial\varphi(x)}{\partial x^1}}{(\varphi(x)+\varepsilon)^2} & \dfrac{\varphi(x) + \varepsilon - x^2 \dfrac{\partial\varphi(x)}{\partial x^2}}{(\varphi(x)+\varepsilon)^2} & \cdots & \dfrac{-x^2 \dfrac{\partial\varphi(x)}{\partial x^n}}{(\varphi(x)+\varepsilon)^2} \\ \cdots\cdots & \cdots\cdots & \cdots & \cdots\cdots \\ \dfrac{-x^n \dfrac{\partial\varphi(x)}{\partial x^1}}{(\varphi(x)+\varepsilon)^2} & \cdots & \dfrac{-x^n \dfrac{\partial\varphi(x)}{\partial x^{n-1}}}{(\varphi(x)+\varepsilon)^2} & \dfrac{\varphi(x) + \varepsilon - x^n \dfrac{\partial\varphi(x)}{\partial x^n}}{(\varphi(x)+\varepsilon)^2} \end{pmatrix}^{-1}.$$

Очевидно, что для существования $\overline{\varphi}'(y)$ нужно требовать невырожденность матрицы $x'(y)$ во всех точках $y \in \mathbb{R}^n$.

Все вышеизложенное можно сформулировать в виде следующего результата:

Теорема 1. *Пусть* $\varphi \in C^2(\mathbb{R}^n)$, $\|\varphi''(x)\| \le C$, $x \in \mathbb{R}^n$, *выполняется предположение* (4) *и матрица* $x'(y)$ *невырождена для всех* $y \in \mathbb{R}^n$.

Тогда существует достаточно большое $N_0 > 0$, *такое, что при* $N \geq N_0$ *и* $y_0 = (2N, \ldots, 2N)^\top$ *выполняется неравенство*

$$\|y_0 - y^*\| < \|y_0 - y_j^*\|, \quad j \in P \backslash \{j_0\}$$

и метод (7) *сходится к* y^* *при* $k \to \infty$, *где* α_k *выбирается из условия* (3). (*Или* $x_k \to x^*$ *при* $k \to \infty$.)

Доказательство. Достаточно показать, что вне области $\bar{\Pi}$ не будет локальных экстремумов, кроме y^*, т.е. $\bar{\varphi}'(y) \neq 0$, $y \in \mathbb{R}^n \backslash \bar{\Pi}$.

Действительно, если для некоторого $\tilde{y} \in \mathbb{R}^n \backslash \bar{\Pi}$ будет $\bar{\varphi}'(\tilde{y}) = 0$, то либо $\varphi_x'(x(\tilde{y})) = 0$, либо матрица $x'(\tilde{y})$ вырождена. Второе очевидно невозможно в силу условий теоремы. Если же $\varphi_x'(x(\tilde{y})) = 0$, то либо $\tilde{y} \in \bar{\Pi}$, что и требовалось, либо соответствующий $x(\tilde{y}) \not\in \bar{\Pi}$. Но в последнем случае градиент $\varphi_x'(x(\tilde{y}))$ не может равняться нулю вне области Π, согласно (5). Теорема доказана. □

Подчеркнем тот факт, что по формуле

$$\bar{\varphi}'(y_k) = \varphi_x'(x(y_k)) \cdot x_y'(y_k)$$

$x(y_k)$ ищется как решение одномерного уравнения относительно неизвестного x:

$$y_k = \frac{x}{\varphi(x) + \varepsilon} = \frac{t \cdot y_k}{\varphi(t \cdot y_k) + \varepsilon}, \qquad (8)$$
$$\text{откуда} \quad x_k = x(y_k),$$

здесь $x = t \cdot y_k$, $t \in R$, что является более простой вычислительной задачей, по сравнению с решением системы n-мерных уравнений. Тогда последовательность $\{y_k\}$ сходится к y^* при $k \to \infty$, $x^* = \varepsilon \cdot y^*$.

Другой подход к решению задачи глобальной оптимизации можно найти, например, в [3].

ИСТОЧНИК ФИНАНСИРОВАНИЯ

Работа выполнена при финансовой поддержке Российского научного фонда (проект №~21-71-30005).

СПИСОК ЛИТЕРАТУРЫ

1. *Евтушенко Ю.Г.* Методы решения экстремальных задач и их применение в системах оптимизации. М.: Наука, 1982.
2. *Карманов В.Г.* Математическое программирование. М.: Наука, 1986.
3. *Grishagin V., Israfilov R., Sergeyev Y.* Convergence conditions and numerical comparison of global optimization methods based on dimensionality reduction schemes // Applied Mathematics and Computation. 2018. V. 318. P. 270–280.

THE METHOD OF FICTITIOUS EXTREMA LOCALIZATION IN THE PROBLEM OF GLOBAL OPTIMIZATION

Academician of the RAS Yu. G. Evtushenko[a,b] and A. A. Tret'yakov[a,c]

[a] *Federal Research Center "Informatics and Control" of the Russian Academy of Sciences, Moscow, Russian Federation*
[b] *Moscow Institute of Physics and Technology (National Research University), Dolgoprudny, Moscow olast, Russian Federation*
[c] *Siedlce University, Faculty of Sciences, Siedlce, Poland*

The problem of finding the global extremum of a non-negative function on a positive parallelepiped in n-dimensional Euclidean space is considered. A method of fictitious extrema localization in a bounded area near the origin is proposed, which allows to separate the global extremum point from fictitious extrema by discarding it at a significant distance from the localization set of fictitious minima. At the same time, due to the choice of the starting point in the gradient descent method, it is possible to justify the convergence of the iterative sequence to the global extremum of the minimized function.

Keywords: global extremum, local minimum, gradient descent method, convergence