

УДК 519.853.4

Разделение параметров в задачах глобальной оптимизации с помощью методов машинного обучения *

Баркалов К. А., Усова М. А.

Нижегородский государственный университет им. Н.И. Лобачевского

Аннотация: В работе представлены результаты исследования подхода к решению задач глобальной оптимизации с разным характером зависимости от разных групп параметров. Предложена схема выделения параметров задачи, которые оказывают локальное влияние на целевую функцию, что позволяет решать существенно многомерные задачи с использованием многошаговой схемы редукции размерности. При этом на разных уровнях рекурсии используются разные оптимизационные алгоритмы. Исследование работоспособности предложенного подхода проведено при решении тестовых и прикладных задач глобальной оптимизации.

Ключевые слова: глобальная оптимизация, локальная оптимизация, редукция размерности, разделение параметров

1. Постановка задачи

В настоящее время методы глобальной оптимизации используются в различных областях науки и техники, например, для идентификации значений параметров математических моделей, при которых результаты моделирования наиболее близки к результатам, полученным экспериментально. Число параметров, которые требуется определить подобным образом, например, для задач химической кинетики, может составлять десятки и сотни. В данном случае использование детерминированных методов глобальной оптимизации крайне ограничено в силу чрезвычайно больших вычислительных затрат на покрытие области поиска точками испытаний. Это остается справедливым даже в случае использования эффективных алгоритмов (например, [1, 2]), строящих существенно неравномерные покрытия. При этом частым явлением в подобных задачах является близкая к линейной зависимость по некоторой группе параметров, в то время как для остальных параметров наблюдается сложный многоэкстремальный характер. Заранее указать разделение на группы параметров с разным характером поведения целевой функции, как правило, не представляется возможным, т. к. целевая функция в обратных задачах задается как черный ящик.

Задача многомерной многоэкстремальной оптимизации может быть определена как проблема поиска наименьшего значения действительной многоэкстремальной функции $\varphi(y)$

$$\begin{aligned} \varphi(y^*) &= \min \{ \varphi(y) : y \in D \}, \\ D &= \{ y \in R^N : a_i \leq y_i \leq b_i, 1 \leq i \leq N \}, \end{aligned} \tag{1}$$

*Работа выполнена при поддержке Министерства науки и высшего образования РФ (проект № 0729-2020-0055) и научно-образовательного математического центра «Математика технологий будущего» (проект № 075-02-2021-1394).

где D есть область поиска, представляющая собой некоторый гиперпараллелепипед N -мерного евклидова пространства. Будем предполагать, что функция $\varphi(y)$ задана в виде «черного ящика» и удовлетворяет условию Липшица

$$|\varphi(y') - \varphi(y'')| \leq L \|y' - y''\|, \quad y', y'' \in D, \quad 0 < L < \infty,$$

с априори неизвестной константой L .

Целевые функции большинства практических задач оптимизации обладают данным свойством. Одновременно с этим будем предполагать, что зависимость целевой функции от N_1 параметров является многоэкстремальной, а от оставшихся $N_2 = N - N_1$ параметров зависимость является близкой к линейной или квадратичной. При этом разделение параметров является неизвестным.

2. Общее описание подхода к решению

Решение поставленной задачи можно организовать с помощью *многошаговой схемы редукции размерности* [3], согласно которой решение многомерной задачи оптимизации может быть получено посредством решения последовательности «вложенных» одномерных задач. Таким образом, решение исходной задачи (1) можно проводить следующим образом:

$$\varphi(y^*) = \min_{y \in D} \varphi(y) = \min_{y^1 \in D_1} f(y^1), \quad D_1 = \{y \in R^{N_1} : a_i \leq y_i \leq b_i, \quad 1 \leq i \leq N_1\}, \quad (2)$$

где

$$f(y^1) = \min_{y^2 \in D_2} \varphi(y^1, y^2), \quad D_2 = \{y \in R^{N_2} : a_i \leq y_i \leq b_i, \quad N_1 + 1 \leq i \leq N\}. \quad (3)$$

Данный подход широко используется для снижения сложности алгоритмов глобальной оптимизации и позволяет формировать эффективные покрытия области поиска. Решение многоэкстремальной подзадачи (2), для которой требуется использовать сложные оптимизационные алгоритмы, будет проводиться на верхнем уровне рекурсии с помощью *алгоритма глобального поиска* [4], принадлежащего классу характеристических алгоритмов, в сочетании со схемой редукции размерности на основе *кривых Пеано* [5]. Этот алгоритм зарекомендовал себя как мощный инструмент для решения сложных многоэкстремальных задач.

Унимодальные подзадачи (3) на нижнем уровне могут быть решены с помощью методов прямого поиска, которые не предполагают гладкости функции. При поиске минимума эти методы измеряют только значения функций, потому в отличие от методов типа BFGS, они не требуют вычисления градиента и хорошо работают в условиях наличия вычислительных ошибок в значениях функций, что является типичным при решении прикладных задач. Одним из наиболее популярных в практике методов этого класса является *метод Хука-Дживса*, использующийся в нашем алгоритме.

Описанный подход позволяет существенно снизить вычислительную сложность поиска оптимума. Однако его использование, вообще говоря, требует ручного распределения переменных по уровням, что, как уже отмечалось ранее, в практических задачах редко осуществимо. Потому возникла необходимость разработки схемы распределения параметров по уровням.

3. Разделение параметров задачи

Для разделения параметров задачи на локальные и глобальные предлагается следующая последовательность действий:

Этап 1. Зафиксировать опорную точку $\bar{y} = (\bar{y}_1, \bar{y}_2, \dots, \bar{y}_N)$ для целевой функции $\varphi(y)$ одним из следующих способов:

- точка фиксируется исходя из физического смысла решаемой прикладной задачи;
- в качестве фиксированной точки используется решение, найденное локальным методом за K итераций;
- в качестве фиксированной точки используется решение, найденное глобальным методом за K итераций.

Этап 2. Для всех компонент (переменных) y_i , $1 \leq i \leq N$ провести исследование на локальность:

1. Вычислить $P + 1$ значение целевой функции $\varphi(y)$ в точках

$$z_i^j = (\bar{y}_1, \dots, \bar{y}_{i-1}, y_i^j, \bar{y}_{i+1}, \dots, \bar{y}_N),$$

где $y_i^j = a_i + jh$, $h = (b_i - a_i)/P$, $0 \leq j \leq P$.

2. Сформировать множество Q , представляющее собой набор из $P + 1$ пары вида

$$\{y_i^j, \varphi(z_i^j)\}.$$

3. Построить регрессионную модель заданной степени deg на основе данных из Q .

4. Вычислить оценку R^2 и провести классификацию i -й переменной:

- если R^2 больше заданного порога T , то добавить переменную к множеству локальных переменных,
- иначе добавить переменную к множеству глобальных переменных.

Значения K , deg и T являются параметрами метода и устанавливаются исследователем в соответствии со свойствами решаемой задачей. В реализованном методе $deg = 1$ или $deg = 2$, что позволяет строить модель линейной (рис. 1) или квадратичной (рис. 2) регрессии соответственно.

В основе проводимой регрессии лежит *метод наименьших квадратов* (МНК). Оценка качества построенной модели производится с помощью скорректированного коэффициента детерминации R_a^2 , характеризующего схожесть построенной модели регрессии с целевой функцией в зафиксированном сечении.

Этап 3. Если локальные переменные были обнаружены, то продолжить решение задачи с использованием описанной ранее многошаговой схемы. В противном случае запустить алгоритм глобального поиска, считая все переменные глобальными.

4. Численные эксперименты

Предложенный подход показал свою работоспособность при решении нескольких серий тестовых задач, представляющих собой линейные комбинации подзадач с глобальными параметрами и подзадач с локальными параметрами – комбинации близких к линейным одномерных функций или функций с разным вкладом квадратичной составляющей.

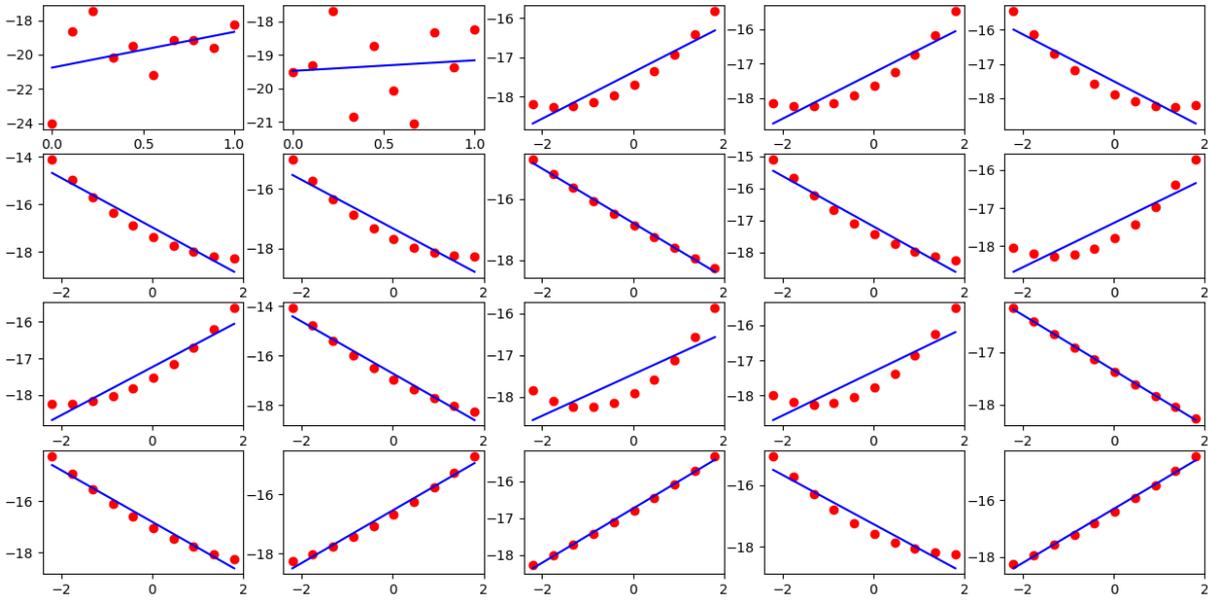


Рис. 1. Линейные регрессионные модели 20-мерной тестовой задачи (2 глобальных параметра).

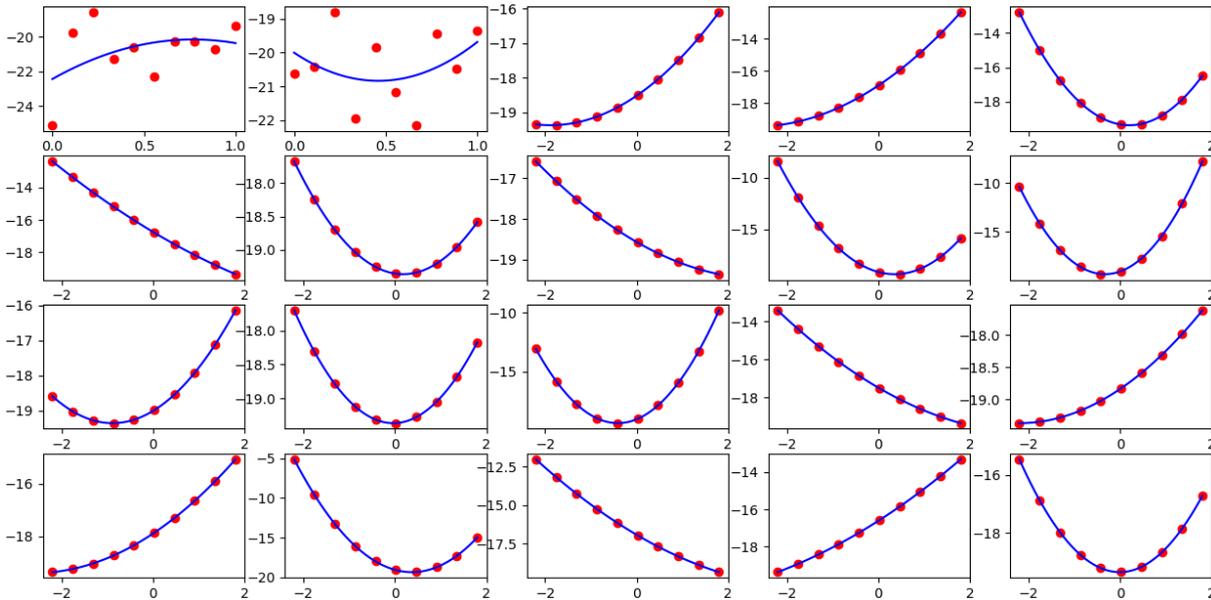


Рис. 2. Квадратичные регрессионные модели 20-мерной тестовой задачи (2 глобальных параметра).

4.1. Описание серий задач

В первом наборе тестовых задач (будем обозначать его GRIS-L) использовалось 20-мерные задачи комбинированного типа: локальная подзадача размерности 18, глобальная подзадача размерности 2. В качестве локальной составляющей использовались подзадачи, представляющие собой комбинацию близких к линейным одномерных функций:

$$\varphi_j^{loc}(y) = \sum_{i=1}^{18} (M_{ij}y_i + L_{ij}y_i^2), \quad (4)$$

где $-2.2 \leq y_i \leq 1.8$, а параметры $0.5 \leq |M_{ij}| \leq 1$ и $0.01 \leq L_{ij} \leq 0.3$ распределены случайно равномерно в соответствующих интервалах. В качестве глобальной составляющей задачи $\varphi_j^{glob}(y)$ использовались многоэкстремальные функции (рис. 3), предложенные Гришагиным [6]. Целевая функция $\varphi(y)$ во всех наборах представляла

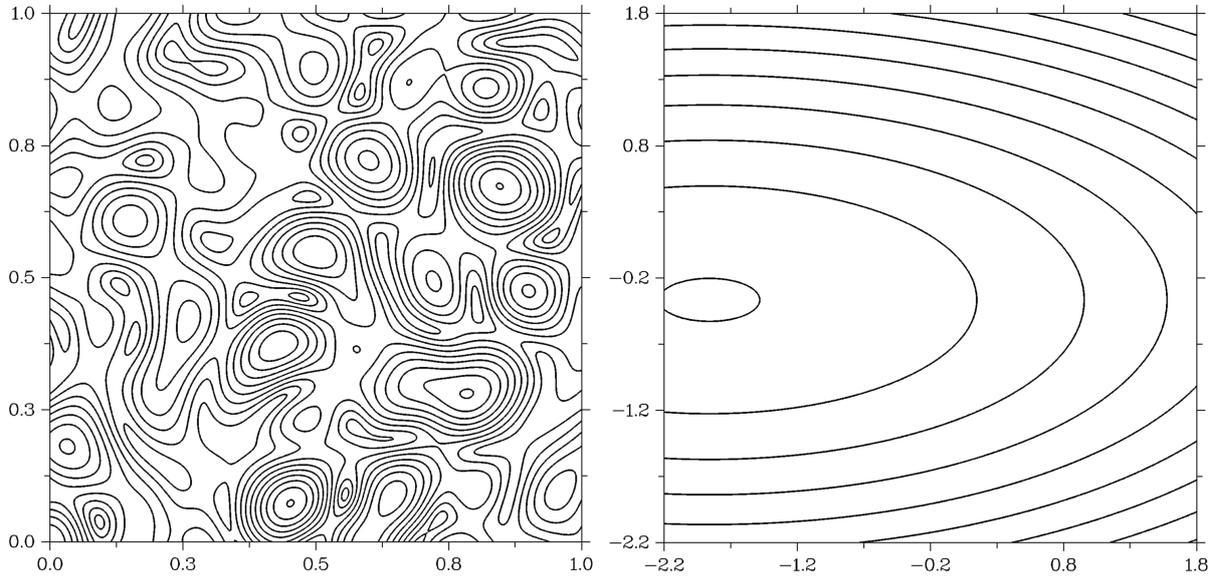


Рис. 3. Линии уровня многоэкстремальной подзадачи и двумерное сечение локальной подзадачи.

собой линейную комбинацию подзадач, т. е. $\varphi_j(y) = \alpha\varphi_j^{glob}(y) + \beta\varphi_j^{loc}(y)$. Последовательность параметров в каждой задаче задавалась случайным образом.

Во втором наборе задач (далее GRIS-LQ) параметры M_{ij} и L_{ij} локальной подзадачи (4) были устроены таким образом, чтобы в 60% случаев переменная имела несущественную квадратичную часть, в 30% вклад квадратичной составляющей был существенным (как минимум, соизмеримым со вкладом линейной составляющей), а оставшиеся 10% переменных были чисто линейными.

В третьей серии экспериментов (далее GKLS-L) в качестве многоэкстремальной части использовались 4-мерные функции из класса GKLS (функции Сергеева) [5] в качестве локальной составляющей использовалась локальные подзадачи серии GRIS-L.

4.2. Результаты экспериментов на сериях тестовых задач

По результатам экспериментов при решении серий задач с существенной многоэкстремальностью (GRIS-L и GRIS-LQ) было проведено корректное разделение на глобальные и локальные переменные, а затем задачи были успешно решены при помощи многошаговой схемы за приемлемое время и количество итераций.

Задачи с функциями GKLS для алгоритма являются более сложными в силу схожести части параметров функций с параболоидами, поэтому решить все задачи серии за ограниченное количество итераций метода не удалось.

Вычисления проводились на компьютере с CPU Intel Core™ i7 10750H, 2.6 GHz. Исследуемый алгоритм был реализован на C++; вычисление значений целевой функции было реализовано в Python 3.9. Для построения и анализа регрессионной модели использовался пакет scikit-learn из Python.

Численные результаты серий экспериментов приведены в Таблице 1. В ней пред-

ставлены данные о числе решенных задач S , среднем числе итераций глобального поиска на верхнем уровне рекурсии G_{av} , среднем числе испытаний на нижнем уровне L_{av} , среднем времени решения задачи t_{av} (в секундах).

Таблица 1. Результаты решения серий тестовых задач

	S	G_{av}	L_{av}	t_{av}
GRIS-L	20/20	705	378 545	1.3
GRIS-LQ	20/20	705	420 397	1.4
GKLS-L	18/20	9531	3 899 417	16.4

4.3. Решение прикладной задачи

В настоящее время наблюдается тенденция к улучшению экологических характеристик автомобильного топлива при сохранении высокого октанового числа. Алкилирование изобутана смешанными олефинами, катализируемое серной кислотой, позволяет получить высокооктановый компонент бензина с минимальным содержанием ароматических углеводородов. Для оптимизации описанного химического процесса в промышленности необходимо сначала разработать его модель, что в данном случае означает построение математической модели химического процесса, а затем его кинетической модели, т. е. численно рассчитать кинетические константы реакции.

Подробные описания моделей представлены в работе [7].

Аналитически узнать кинетические константы реакций, как правило, невозможно. В этом случае критерии качества найденного решения (целевая функция) не имеют явного аналитического описания, но допускают алгоритмическое представление и требуют значительных вычислительных ресурсов. Ранее использование параллельного асинхронного алгоритма, работающего на 8 узлах кластера, позволило значительно сократить время поиска оптимума. Для проведения указанного численного эксперимента использовался суперкомпьютер «Лобачевский» (операционная система – CentOS 7.2, на узле – два процессора Intel Sandy Bridge E5-2660 2.2 GHz, оперативная память – 64 Gb). Найденные оптимальные параметры 15-мерной модели позволили адекватно смоделировать описанный процесс (рис. 4). При параметре метода $r = 3$ было найдено лучшее значение целевой функции, равное 0.519793. Однако при использовании данного метода решения достигнуть заданной точности за обозримое время не представляется возможным (используется остановка по числу итераций или по ограничению времени решения).

С помощью исследуемого в данной работе метода за 15 часов было найдено улучшенное решение данной задачи – 0.497063. Размерность глобальной подзадачи – 4, число запусков глобального поиска – 427, число запусков локального метода – 427000.

5. Заключение

В данной работе исследовался подход к разделению параметров по уровням для использования известной многоуровневой схемы редукции размерности в задачах глобальной оптимизации, комбинирующей использование кривых Пеано и много-

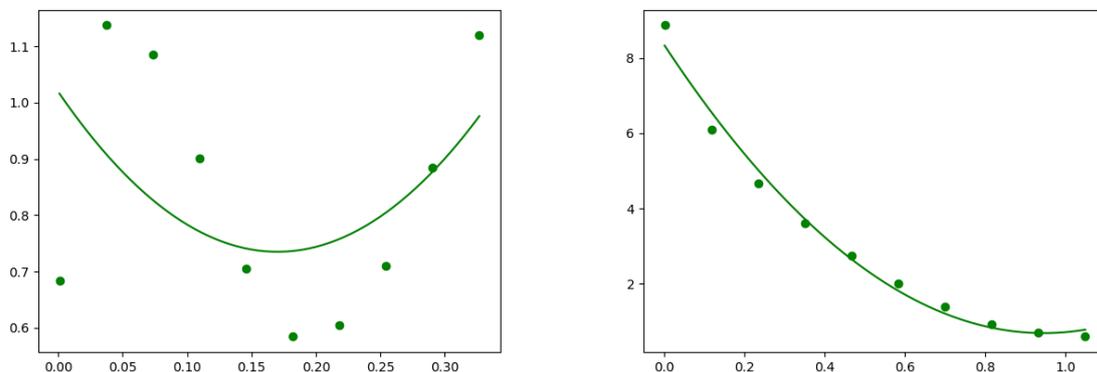


Рис. 4. Квадратичные регрессионные модели 15-мерной прикладной задачи глобального параметра (слева) и для локального параметра (справа).

шаговую схему. Для решения редуцированных глобальных подзадач использовался алгоритм глобального поиска, для локальных подзадач был выбран метод Хука-Дживса. Предложена схема разделения параметров с использованием построения регрессионной модели с помощью МНК. Проведены вычислительные эксперименты на серии тестовых задач большой размерности. Получено улучшенное решение прикладной задачи. Результаты экспериментов показывают, что предложенная схема позволяет корректно распределять параметры и эффективно находить оптимум задачи.

Литература

1. Евтушенко Ю.Г., Малкова В.У., Станевичюс А.-И. А. Параллельный поиск глобального экстремума функций многих переменных // Журн. вычисл. матем. и матем. физ. 2009. Т. 49, № 2. С. 255–269.
2. Paulavičius R., Žilinskas J. Advantages of Simplicial Partitioning for Lipschitz Optimization Problems with Linear Constraints // Optim. Lett. 2016. V. 10, No. 2. P. 237–246.
3. Городецкий С.Ю., Гришагин В.А. Нелинейное программирование и многоэкстремальная оптимизация. Н. Новгород: Изд-во ННГУ, 2007. 489 с.
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. Sergeyev Y.D., Strongin R.G., Lera D. Introduction to global optimization exploiting space-filling curves // Journal of Global Optimization, Springer. 2013. V. 60, No. 3. P. 595–596.
6. Sergeyev Y., Grishagin V. Parallel Asynchronous Global Search and the Nested Optimization Scheme // J. Comput. Anal. Appl. 2001. V. 3, No. 2. P. 123–145.
7. Gubaydullin I., Enikeeva L., Barkalov K., Lebedev I., Silenko D. Kinetic Modeling of Isobutane Alkylation with Mixed C4 Olefins and Sulfuric Acid as a Catalyst Using the Asynchronous Global Optimization Algorithm // Communications in Computer and Information Science. 2022. V. 1618. P. 293–306.

MSC 90C26

Separation of parameters using machine learning methods in global optimization problems

K. A. Barkalov, M. A. Usova

Lobachevsky State University of Nizhny Novgorod
Nizhny Novgorod, Russia

Abstract: The paper presents the findings from research into an approach to solving global optimization problems with a different nature of dependence on diverse groups of parameters. A scheme for choosing the problem parameters, which have a local effect on the objective function is proposed, which allows to solve essentially multidimensional problems using the nested optimization scheme. At the same time, different optimization algorithms are used at differing levels of recursion (nesting levels). The proposed approach has demonstrated its efficiency in solving several series of test problems. The effectiveness of the proposed approach was studied.

Keywords: global optimization, local optimization, nested optimization scheme, high-dimensional problems, parameter separation

References

1. Yu.G. Evtushenko, V.U. Malkova, A.A. Stanevichyus, Parallel global optimization of functions of several variables, *USSR Computational Mathematics and Mathematical Physics*, 2009, V. 49, No. 2, P. 255–269.
2. R. Paulavičius, J. Žilinskas, Advantages of Simplicial Partitioning for Lipschitz Optimization Problems with Linear Constraints, *Optim. Lett.*, 2016, V. 10, No. 2, P. 237–246.
3. S.Yu. Gorodetskii, V.A. Grishagin, Nonlinear Programming and Optimization Multiextremal, Nizhny Novgorod, UNN Publ, 2007, 489 p.
4. R.G. Strongin, V.P. Gergel, V.A. Grishagin, K.A. Barkalov, Parallel Computations in the Global Optimization Problems, Moscow, MSU Publ, 2012, 280 p.
5. Y.D. Sergeyev, R.G. Strongin, D. Lera, Introduction to global optimization exploiting space-filling curves, *Journal of Global Optimization*, Springer, 2013, V. 60, No. 3, P. 595–596.
6. Y. Sergeyev, V. Grishagin, Parallel Asynchronous Global Search and the Nested Optimization Scheme, *J. Comput. Anal. Appl.*, 2001, V. 3, No. 2, P. 123–145.
7. I. Gubaydullin, L. Enikeeva, K. Barkalov, I. Lebedev, D. Silenko, Kinetic Modeling of Isobutane Alkylation with Mixed C4 Olefins and Sulfuric Acid as a Catalyst Using the Asynchronous Global Optimization Algorithm, *Communications in Computer and Information Science*, 2022, V. 1618, P. 293–306.