ISSN 1810-0198. Вестник Тамбовского университета. Серия: естественные и технические науки

Том 24, № 126 2019

DOI 10.20310/1810-0198-2019-24-126-150-165

УДК 519

Гибридная глобализация сходимости метода последовательного квадратичного программирования, стабилизированного вдоль подпространства

Николай Георгиевич ЖУРБЕНКО[1] **, Алексей Феридович ИЗМАИЛОВ**[2] **, Евгений Иванович УСКОВ**[3]

[1)] Институт кибернетики им. В. М. Глушкова НАН Украины
03187, Украина, г. Киев, проспект Академика Глушкова, 40
ORCID: https://orcid.org/0000-0001-5123-8607, e-mail: zhurbnick@gmail.com

[2)] ФГБОУ ВО «Московский государственный университет им. М. В. Ломоносова»
119992, ГСП-2, Российская Федерация, г. Москва, Ленинские горы, факультет ВМК
ORCID: https://orcid.org/0000-0001-9851-0524, e-mail: izmaf@ccas.ru

[3)] ФГБОУ ВО «Тамбовский государственный университет им. Г.Р. Державина»
392000, Российская Федерация, г. Тамбов, ул. Интернациональная, 33
ORCID: https://orcid.org/0000-0002-3639-0317, e-mail: euskov@cs.msu.ru

Hybrid globalization of convergence of subspace-stabilized sequential quadratic programming method

Nikolay G. ZHURBENKO[1] **, Alexey F. IZMAILOV**[2] **, Evgeniy I. USKOV**[3]

[1)] V. M. Glushkov Institute of Cybernetics of NAS of Ukraine
40 Akademika Glushkova Ave., Kiev 03187, Ukraine
ORCID: https://orcid.org/0000-0001-5123-8607, e-mail: zhurbnick@gmail.com

[2)] Lomonosov Moscow State University
VMK Faculty, Leninskiye Gory, Moscow 119991, Russian Federation
ORCID: https://orcid.org/0000-0001-9851-0524, e-mail: izmaf@ccas.ru

[3)] Derzhavin Tambov State University
33 Internatsionalnaya St., Tambov 392000, Russian Federation
ORCID: https://orcid.org/0000-0002-3639-0317, e-mail: euskov@cs.msu.ru

Аннотация. Локальная сверхлинейная сходимость стабилизированного метода последовательного квадратичного программирования устанавливается при очень слабых предположениях, не включающих в себя никакие условия регулярности ограничений. Однако, все попытки глобализации сходимости этого метода неминуемо сталкиваются с принципиальными трудностями, связанными с поведением этого метода при относительной удаленности текущей итерации от решений. А именно, стабилизированный метод последовательного квадратичного программирования имеет тенденцию генерировать длинные последовательности коротких шагов перед тем, как проявляется его сверхлинейная сходимость. В связи с этим был предложен метод последовательного квадратичного программирования, стабилизированный вдоль подпространства, обладающий лучшим «полулокальным» поведением, а значит, лучше приспособленный для разработки на его основе практических алгоритмов. В данной работе предлагаются два способа гибридной глобализации сходимости этого метода: алгоритм с возвратами и алгоритм с рекордами. Приводятся теоретические результаты о глобальной сходимости и скорости сходимости данных алгоритмов, а также результаты сравнительного численного тестирования.

Ключевые слова: последовательное квадратичное программирование; вырожденные решения; некритический множитель Лагранжа; двойственная стабилизация; глобализация сходимости

Благодарности: Работа выполнена при поддержке РФФИ (проекты № 17-01-00125_a и № 19-51-12003 ННИО_а) и фонда Volkswagen (грант 90306).
Авторы выражают признательность Дарию Панкову за помощь в проведении численного тестирования.

Для цитирования: *Журбенко Н.Г., Измаилов А.Ф., Усков Е.И.* Гибридная глобализация сходимости метода последовательного квадратичного программирования, стабилизированного вдоль подпространства // Вестник Тамбовского университета. Серия: естественные и технические науки. Тамбов, 2019. Т. 24. № 126. С. 150–165. DOI 10.20310/1810-0198-2019-24-126-150-165.

Abstract. Local superlinear convergence of the stabilized sequential quadratic programming method is established under very weak assumptions not involving any constraint qualification conditions. However, all attempts to globalize convergence of this method inevitably face principal difficulties related to the behavior of this method when the iterates are still relatively far from solutions. Specifically, the stabilized sequential quadratic programming method has a tendency to generate long sequences of short steps before its superlinear convergence shows up. To that end, the so-called subspace-stabilized sequential quadratic programming method has been proposed, demonstrating better "semi-local" behavior, and hence, more suitable for development of practical algorithms on its basis. In this work we propose two techniques for hybrid globalization of convergence of this method: algorithm with backups, and algorithm with records. We provide theoretical results on convergence and rate of convergence of these algorithms, as well as some results of their numerical testing.

Keywords: sequential quadratic programming; degenerate solutions; noncritical Lagrange multiplier; dual stabilization; globalization of convergence

Acknowledgements: The work is partially supported by the Russian Foundation for Basic Research (projects no. 17-01-00125_a and 19-51-12003 ННИО_а) and by the Volkswagen Foundation (grant 90306).
The authors are grateful to Dariy Pankov for assistance in conducting numerical testing.

For citation: Zhurbenko N.G., Izmailov A.F., Uskov E.I. Gibridnaya globalizatsiya skhodimosti metoda posledovatel'nogo kvadratichnogo programmirovaniya, stabilizirovannogo vdol' podprostranstva [Hybrid globalization of convergence of subspace-stabilized sequential quadratic programming method]. *Vestnik Tambovskogo universiteta. Seriya: estestvennye i tekhnicheskie nauki – Tambov University Reports. Series: Natural and Technical Sciences*, 2019, vol. 24, no. 126, pp. 150–165. DOI 10.20310/1810-0198-2019-24-126-150-165. (In Russian, Abstr. in Engl.)

Введение

Рассматривается задача оптимизации с ограничениями-равенствами

$$f(x) \to \min, \quad h(x) = 0, \tag{1}$$

где целевая функция $f : \mathbb{R} \to \mathbb{R}$ и отображение $h : \mathbb{R}^n \to \mathbb{R}^l$ по крайней мере дважды дифференцируемы.

Стационарные точки данной задачи и соответствующие им множители Лагранжа задаются системой оптимальности Лагранжа

$$\frac{\partial L}{\partial x}(x,\, \lambda) = 0, \quad h(x) = 0, \tag{2}$$

где $L : \mathbb{R}^n \times \mathbb{R}^l \to \mathbb{R}$,

$$L(x,\, \lambda) = f(x) + \langle \lambda,\, h(x) \rangle,$$

есть функция Лагранжа задачи (1). А именно, обозначим через $\mathcal{M}(\bar{x})$ множество множителей Лагранжа, отвечающих допустимой точке $\bar{x} \in \mathbb{R}^n$ задачи (1), то есть множество таких $\lambda \in \mathbb{R}^l$, для которых пара $(\bar{x},\, \lambda)$ удовлетворяет системе (2). Тогда говорят, что $\bar{x}$ является стационарной точкой данной задачи, если $\mathcal{M}(\bar{x}) \neq \emptyset$.

Как хорошо известно, стационарность является необходимым условием локальной оптимальности точки $\bar{x}$ в задаче (1) при выполнении условия регулярности ограничений

$$\operatorname{rank} h'(\bar{x}) = l, \tag{3}$$

которое к тому же является необходимым и достаточным условием единственности множителя Лагранжа, отвечающего стационарной точке $\bar{x}$.

В терминологии [1, разд. 1.3.3], множитель Лагранжа $\bar{\lambda} \in \mathcal{M}(\bar{x})$ называется критическим, если существует такой вектор $\xi \in \ker h'(\bar{x}) \backslash \{0\}$, что

$$\frac{\partial^2 L}{\partial x^2}(\bar{x},\, \bar{\lambda})\xi \in \operatorname{im}(h'(x))^{\mathrm{T}},$$

и некритическим иначе. Будем говорить, что в стационарной точке $\bar{x}$ задачи (1) для отвечающего ей множителя Лагранжа $\bar{\lambda} \in \mathcal{M}(\bar{x})$ выполнено достаточное условие второго порядка оптимальности, если

$$\left\langle \frac{\partial^2 L}{\partial x^2}(\bar{x},\, \bar{\lambda})\xi,\, \xi \right\rangle > 0 \quad \forall\, \xi \in \ker h'(\bar{x}) \setminus \{0\}.$$

При выполнении этого условия множитель $\bar{\lambda}$ всегда является некритическим.

Как показано в [1, предложение 1.43], некритичность $\bar{\lambda} \in \mathcal{M}(\bar{x})$ равносильна следующему свойству: существует константа $L > 0$ такая, что для всех $(x,\, \lambda) \in \mathbb{R}^n \times \mathbb{R}^l$, достаточно близких к $(\bar{x},\, \bar{\lambda})$, выполняется неравенство

$$\|x - \bar{x}\| + \operatorname{dist}(\lambda,\, \mathcal{M}(\bar{x})) \leq L\|\Phi(x,\, \lambda)\|, \tag{4}$$

где $\Phi : \mathbb{R}^n \times \mathbb{R}^l \to \mathbb{R}^n \times \mathbb{R}^l$,

$$\Phi(x,\,\lambda) = \left(\frac{\partial L}{\partial x}(x,\,\lambda),\; h(x)\right), \tag{5}$$

есть оператор системы Лагранжа (2).

Одним из наиболее эффективных методов общего назначения для задач условной оптимизации является метод последовательного квадратичного программирования (SQP, от «sequential quadratic programming») [1, разд. 4.2], [2, разд. 4.4], который для задачи (1) есть ни что иное, как метод Ньютона для системы уравнений Лагранжа (2). Локальная сверхлинейная сходимость метода SQP к решению $(\bar{x},\,\bar{\lambda})$ этой системы устанавливается при выполнении условия регулярности ограничений (3) и некритичности множителя $\bar{\lambda}$. О различных способах глобализации сходимости метода SQP см. [1, разд. 6.2], [2, разд. 5.4], а также разд. 1 ниже.

Стабилизированный метод последовательного квадратичного программирования (sSQP, от «stabilized SQP») был предложен в [3] для задачи оптимизации с ограничениями-неравенствами как средство восстановления сверхлинейной скорости сходимости метода SQP, которая обычно теряется при нарушении условий регулярности ограничений. Впоследствии метод был распространен и на задачи, в которых присутствуют и ограничения-равенства. В частности, в [4] (см. также разд. 2 ниже) для задачи (1) было показано, что при выполнении одного лишь условия некритичности $\bar{\lambda}$ метод sSQP локально сверхлинейно сходится к $(\bar{x},\,\lambda^*)$ с некоторым $\lambda^* \in \mathcal{M}(\bar{x})$, близким к $\bar{\lambda}$.

В работах [5]– [12] предпринимались различные попытки глобализации сходимости метода sSQP. В частности, в [5] были предложены два алгоритма гибридной глобализации данного метода, где в качестве алгоритма внешней фазы использовался глобализованный посредством одномерного поиска метод SQP. Однако, все эти попытки сталкиваются с принципиальными трудностями, связанными с поведением метода sSQP при относительной удаленности текущей итерации от решений. А именно, метод имеет тенденцию генерировать длинные последовательности коротких шагов перед тем, как проявляется его сверхлинейная сходимость. Подчеркнем, что это дефект самого метода sSQP, и он неминуемо будет проявляться при любых способах глобализации его сходимости. На сегодняшний день однозначно успешные способы такой глобализации авторам неизвестны.

В связи со сказанным в [13] было сделано предположение, что в модификации нуждаются не способы глобализации сходимости, а сам метод sSQP. Там же был разработан метод последовательного квадратичного программирования, стабилизированный вдоль подпространства (s-sSQP, от «subspace-stabilized SQP»), обладающий лучшим «полулокальным» поведением, а значит, лучше приспособленный для глобализации. Основные ингредиенты этой конструкции обсуждаются ниже в разд. 2, а в разд. 3 предлагаются два способа такой глобализации: алгоритм с возвратами и алгоритм с рекордами, а также соответствующие теоретические результаты о глобальной сходимости и скорости сходимости данных алгоритмов. Наконец, разд. 4 содержит некоторые результаты численного тестирования.

1. Глобализованный метод последовательного квадратичного программирования

Метод SQP состоит в следующем. Для текущего прямо-двойственного приближения $(x^k, \lambda^k) \in \mathbb{R}^n \times \mathbb{R}^l$ вычисляется направление (ξ^k, η^k) как решение системы линейных уравнений

$$H_k\xi + (h'(x^k))^{\mathrm{T}}\eta = -\frac{\partial L}{\partial x}(x^k, \lambda^k), \quad h'(x^k)\xi = -h(x^k), \tag{6}$$

с базовым выбором

$$H_k = \frac{\partial^2 L}{\partial x^2}(x^k, \lambda^k). \tag{7}$$

Следующее приближение определяется как $(x^{k+1}, \lambda^{k+1}) = (x^k + \xi^k, \lambda^k + \eta^k)$.

В разд. 3 в качестве метода внешней фазы для гибридного алгоритма будет использоваться следующий вариант метода SQP, глобализованный одномерным поиском для негладкой точной штрафной функции $\varphi_c : \mathbb{R}^n \to \mathbb{R}$,

$$\varphi_c(x) = f(x) + c\|F(x)\|_1, \tag{8}$$

при соответствующем выборе значений параметра штрафа $c > 0$; см. [1, разд. 6.2.1], [2, разд. 5.4.1]. Приводимый алгоритм снабжен так называемыми поправками второго порядка, позволяющими подавлять известный эффект Маратоса, замедляющий сходимость [1, разд. 6.2.2], [2, разд. 5.4.2].

Алгоритм 1. Выбираем параметры $\bar{c} > 0$, $\varepsilon, \theta \in (0, 1)$. Выбираем начальное приближение $(x^0, \lambda^0) \in \mathbb{R}^n \times \mathbb{R}^l$ и полагаем $k = 0$.

1. Выбираем симметричную матрицу $H_k \in \mathbb{R}^{n\times n}$ и вычисляем (ξ^k, η^k) как решение системы (6). Полагаем $\lambda^{k+1} = \lambda^k + \eta^k$.

2. Если $\xi^k = 0$, стоп.

3. Выбираем
$$c_k \geq \|\lambda^{k+1}\|_\infty + \bar{c},$$
и вычисляем
$$\Delta_k = \langle f'(x^k), \xi^k\rangle - c_k\|F(x^k)\|_1.$$

4. Если выполняется неравенство
$$\varphi_{c_k}(x^k + \xi^k) \leq \varphi_{c_k}(x^k) + \varepsilon\Delta_k,$$
где функция φ_{c_k} вводится согласно (8), то полагаем $x^{k+1} = x^k + \xi^k$, и переходим к п. 7.

5. Вычисляем $\widehat{\xi}^k$ как решение задачи оптимизации
$$\|\xi\| \to \min, \quad h(x_k + \xi_k) + h'(x^k)\xi = 0.$$

6. Полагаем $\alpha = 1$. Если выполняется неравенство
$$\varphi_{c_k}(x^k + \alpha\xi^k + \alpha^2\widehat{\xi}^k) \leq \varphi_{c_k}(x^k) + \varepsilon\alpha\Delta_k, \tag{9}$$
то полагаем $\alpha_k = \alpha$. В противном случае заменяем α на $\theta\alpha$ и проверяем (9), до тех пор, пока оно не будет выполнено. Полагаем $x^{k+1} = x^k + \alpha_k\xi^k + \alpha_k^2\widehat{\xi}^k$.

7. Увеличиваем k на 1 и переходим к п. 1.

Анализ свойств глобальной сходимости и скорости сходимости данного алгоритма приводится в [1, разд. 6.2], [2, разд. 5.4].

2. Стабилизированные методы последовательного квадратичного программирования

Метод sSQP получается заменой итерационной системы (6) на следующую:
$$H_k\xi + (h'(x^k))^{\mathrm{T}}\eta = -\frac{\partial L}{\partial x}(x^k,\, \lambda^k), \quad h'(x^k)\xi - \sigma(x^k,\, \lambda^k)\eta = -h(x^k), \tag{10}$$
с H_k выбираемым согласно (7), и с функцией $\sigma : \mathbb{R}^n \times \mathbb{R}^l \to \mathbb{R}_+$, задающей параметр стабилизации и традиционно определяемой как невязка системы Лагранжа (2):
$$\sigma(x,\, \lambda) = \|\Phi(x,\, \lambda)\|,$$
где Φ введено в (5).

Нарушение условия регулярности ограничений в точке $\bar{x}$ означает, что
$$(\operatorname{im} h'(\bar{x}))^{\perp} = \ker(h'(\bar{x}))^{\mathrm{T}}$$
является нетривиальным линейным подпространством в $\mathbb{R}^l$, которое в дальнейшем будем называть подпространством вырожденности. Когда подпространство вырожденности совпадает со всем пространством $\mathbb{R}^l$, т. е. когда $h'(\bar{x}) = 0$, метод sSQP и его упоминавшиеся выше глобализации обычно вполне успешны. Однако, при неполном вырождении, т. е. в тех случаях, когда подпространство вырожденности является нетривиальным, но *собственным* подпространством в $\mathbb{R}^l$, данный метод имеет тенденцию генерировать длинные последовательности коротких шагов. Такое поведение наблюдалось уже в некоторых примерах в [14]. Необходимость подавления этого эффекта является мотивировкой введения в [13] следующего класса методов.

Пусть задано отображение $P : \mathbb{R}^n \times \mathbb{R}^l \to \mathbb{R}^{l\times l}$ и функция $\sigma : \mathbb{R}^n \times \mathbb{R}^l \to \mathbb{R}$. Предполагается, что линейный оператор $P(x,\, \lambda)$ аппроксимирует проектор на подпространство вырожденности при $(x,\, \lambda) \to (\bar{x},\, \bar{\lambda})$. Итерационная система метода s-sSQP имеет вид
$$H_k\xi + (h'(x^k))^{\mathrm{T}}\eta = -\frac{\partial L}{\partial x}(x^k,\, \lambda^k), \quad h'(x^k)\xi - \sigma(x^k,\, \lambda^k)P(x^k,\, \lambda^k)\eta = -h(x^k), \tag{11}$$

т. е. отличается от итерационной системы (10) метода sSQP только присутствием линейного оператора $P(x^k,\, \lambda^k)$ во втором уравнении, а также, возможно, другим выбором функции σ.

В [13] были введены две группы методов s-sSQP. Методы s-sSQP с исчезающей стабилизацией характеризуются следующими требованиями на P и σ:

– отображение P непрерывно в точке $(\bar{x},\, \bar{\lambda})$, и $(\operatorname{im} h'(\bar{x}))^{\perp}$ является инвариантным подпространством $P(\bar{x},\, \bar{\lambda})$, т. е.

$$P(\bar{x},\, \bar{\lambda})\eta = \eta \quad \forall\, \eta \in (\operatorname{im} h'(\bar{x}))^{\perp};$$

– функция σ непрерывна в каждой точке множества $\{\bar{x}\} \times \mathcal{M}(\bar{x})$ из некоторой окрестности $(\bar{x},\, \bar{\lambda})$, $\sigma(\bar{x},\, \lambda) = 0$ для всех $\lambda \in \mathcal{M}(\bar{x})$ достаточно близких к $\bar{\lambda}$, $\sigma(x,\, \lambda) \neq 0$ для всех $(x,\, \lambda) \in (\mathbb{R}^n \times \mathbb{R}^l) \setminus (\{\bar{x}\} \times \mathcal{M}(\bar{x}))$ из некоторой окрестности $(\bar{x},\, \bar{\lambda})$, и

$$\|x - \bar{x}\| = O(|\sigma(x,\, \lambda)|)$$

при $(x,\, \lambda) \to (\bar{x},\, \bar{\lambda})$.

Заметим, что эти предположения позволяют взять, например, $P(\cdot) \equiv I$, тем самым покрывая метод sSQP. О других способах выбора P см. ниже. Кроме того, если $\bar{\lambda}$ — некритический множитель Лагранжа, то, согласно эквивалентности этого условия оценке (4), всем требованиям удовлетворяет выбор

$$\sigma(x,\, \lambda) = \|\Phi(x,\, \lambda)\|^{\beta}$$

с любым фиксированным показателем $\beta \in (0,\, 1]$, где Φ введен в (5).

Методы s-sSQP с неисчезающей стабилизацией предполагают следующие требования на P и σ:

– отображение P непрерывно в каждой точке множества $\{\bar{x}\} \times \mathcal{M}(\bar{x})$ из некоторой окрестности $(\bar{x},\, \bar{\lambda})$, $\operatorname{im} P(\bar{x},\, \lambda) \cap \operatorname{im} h'(\bar{x}) = \{0\}$ для всех $\lambda \in \mathcal{M}(\bar{x})$ из некоторой окрестности $\bar{\lambda}$, и $\ker P(\bar{x},\, \bar{\lambda}) \cap (\operatorname{im} h'(\bar{x}))^{\perp} = \{0\}$.

– функция σ непрерывна в точке $(\bar{x},\, \bar{\lambda})$, и $\sigma(\bar{x},\, \bar{\lambda}) \neq 0$.

Идея методов такого рода восходит к [15].

Как показано в [13], локальные свойства свойства указанных двух вариантов методов s-sSQP аналогичны соответствующим свойствам sSQP: локальная сверхлинейная сходимость гарантирована при одном лишь предположении некритичности множителя Лагранжа $\bar{\lambda}$. Там же предложены и практические способы задания подходящего отображения P, аппроксимирующего проектор на подпространство вырожденности.

3. Гибридная глобализация сходимости

В этом разделе будут описаны два способа гибридной глобализации метода s-sSQP: алгоритм с возвратами и алгоритм с рекордами. Общая идея этих техник глобализации изложена в [1, разд. 5.3].

Суть алгоритма с возвратами состоит в следующем. На каждой итерации из текущего приближения $(x^k,\, \lambda^k)$ делается шаг метода внутренней фазы (в данном случае s-sSQP), который принимается, если он приводит к линейному убыванию невязки системы Лагранжа, т. е. к ее уменьшению по крайней мере в заданное число раз, определяемое в алгоритме ниже параметром q. Если на какой-либо из последующих итераций шаг метода внутренней фазы не принимается, то происходит возврат к тому приближению, откуда был сделан первый в данной серии шагов метода внутренней фазы, и из этой точки делается шаг метода внешней фазы (в данном случае глобализованного метода SQP, т. е. алгоритма 1).

А л г о р и т м 2. Выбираем параметр $q \in (0,\, 1)$, а также параметры алгоритма 1. Выбираем начальное приближение $(x^0,\, \lambda^0) \in \mathbb{R}^n \times \mathbb{R}^l$ и полагаем $k = 0$.

1. Полагаем $\widehat{k} = k$ и $(\widehat{x},\, \widehat{\lambda}) = (x^k, \lambda^k)$.

2. Определяем H_k согласно (7). Определяем $P(x^k,\, \lambda^k)$ и $\sigma(x^k,\, \lambda^k)$ согласно выбранному варианту метода s-sSQP. Вычисляем $(\xi^k,\, \eta^k)$ как решение системы (11). Если решение найти не удается, то переходим к п. 4. В противном случае полагаем $(x^{k+1},\, \lambda^{k+1}) = (x^k + \xi^k,\, \lambda^k + \eta^k)$.

3. Если
$$\|\Phi(x^{k+1},\, \lambda^{k+1})\| \le q\|\Phi(x^k,\, \lambda^k)\|, \tag{12}$$
то увеличиваем k на 1 и переходим к п. 2. (Шаг метода внутренней фазы.)

4. Если $k > \widehat{k}$, полагаем $k = \widehat{k}$ и $(x^k,\, \lambda^k) = (\widehat{x},\, \widehat{\lambda})$. Вычисляем $(x^{k+1},\, \lambda^{k+1})$ с помощью итерации алгоритма 1, увеличиваем k на 1 и переходим к п. 1. (Шаг метода внешней фазы.)

Что касается свойств глобальной сходимости приведенного алгоритма, то возможны только два сценария: либо все итерации, начиная с некоторой, являются шагами метода внутренней фазы, либо все итерации являются шагами метода внешней фазы. В последнем случае алгоритм 2 наследует свойства глобальной сходимости алгоритма 1. В первом же случае из условия (12) вытекает следующее утверждение.

Теорема 1. *Пусть функция f и отображение h дважды непрерывно дифференцируемы на $\mathbb{R}^n$. Пусть последовательность $\{(x^k,\, \lambda^k)\}$ сгенерирована алгоритмом 2, и пусть все члены этой последовательности, начиная с некоторого, получены шагами метода внутренней фазы.*

Тогда

$$\Phi(x^k,\, \lambda^k) \to 0$$

при $k \to \infty$, *и, в частности, если* $(\bar{x},\, \bar{\lambda})$ *является предельной точкой последовательности* $\{(x^k,\, \lambda^k)\}$, *то* $\bar{x}$ *— стационарная точка задачи* (1), *а* $\bar{\lambda}$ *— отвечающий ей множитель Лагранжа.*

Скорость сходимости алгоритма 2 характеризуется следующим утверждением.

Теорема 2. *Пусть функция* f *и отображение* h *дважды непрерывно дифференцируемы в окрестности стационарной точки* $\bar{x}$ *задачи* (1), *и пусть* $\bar{\lambda}$ *— отвечающий* $\bar{x}$ *некритический множитель Лагранжа. Пусть точка* $(\bar{x},\, \bar{\lambda})$ *является предельной точкой последовательности* $\{(x^k,\, \lambda^k)\}$, *сгенерированной алгоритмом 2.*

Тогда вся последовательность $\{(x^k,\, \lambda^k)\}$ *сходится к* $(\bar{x},\, \bar{\lambda})$, *и скорость сходимости сверхлинейная.*

Д о к а з а т е л ь с т в о. Рассмотрим точку $(x^k,\, \lambda^k)$, достаточно близкую к $(\bar{x},\, \bar{\lambda})$. Пусть точка $(x^{k+1},\, \lambda^{k+1})$ получена шагом метода внутренней фазы.

Используя локальную липшицевость Φ в сделанных предположениях гладкости, а также (4) и [13, теоремы 1 и 2], получаем

$$\Phi(x^{k+1},\, \lambda^{k+1}) = O(\|x^{k+1} - \bar{x}\| + \operatorname{dist}(\lambda^{k+1},\, \mathcal{M}(\bar{x}))) =$$

$$= o(\|x^k - \bar{x}\| + \operatorname{dist}(\lambda^k,\, \mathcal{M}(\bar{x}))) = o(\|\Phi(x^k,\, \lambda^k)\|)$$

при $k \to \infty$. Из полученной оценки вытекает выполнение (12) для любого фиксированного $q \in (0,\, 1)$ при достаточной близости $(x^k,\, \lambda^k)$ к $(\bar{x},\, \bar{\lambda})$. Но тогда точка $(x^{k+1},\, \lambda^{k+1})$ принимается алгоритмом 2. Отсюда легко следует, что все точки последовательности $\{(x^k,\, \lambda^k)\}$ сгенерированы шагами метода внутренней фазы, а значит, алгоритм наследует сверхлинейную сходимость метода s-sSQP, установленную в [13, теоремы 1 и 2]. $\square$

Альтернативная стратегия гибридной глобализации, называемая алгоритмом с рекордами, состоит в следующем. Вместо того, чтобы сравнивать невязку системы Лагранжа в пробной точке с невязкой, полученной на предыдущей итерации, можно сравнивать ее с наименьшим достигнутым значением невязки по всем предыдущим итерациям (т. е. с рекордом), и принимать шаг метода внутренней фазы только в том случае, если наблюдается линейное убывание рекорда.

А л г о р и т м 3. Выбираем параметр $q \in (0,\, 1)$, а также параметры алгоритма 1. Выбираем начальное приближение $(x^0,\, \lambda^0) \in \mathbb{R}^n \times \mathbb{R}^l$ и полагаем $k = 0$. Полагаем $R = \|\Phi(x^0,\, \lambda^0)\|$.

1. Определяем H_k согласно (7). Определяем $P(x^k,\, \lambda^k)$ и $\sigma(x^k,\, \lambda^k)$ согласно выбранному варианту метода s-sSQP. Вычисляем $(\xi^k,\, \eta^k)$ как решение системы (11). Если решение найти не удается, то переходим к п. 3. В противном случае полагаем $(x^{k+1},\, \lambda^{k+1}) = (x^k + \xi^k,\, \lambda^k + \eta^k)$.

2. Если
$$\|\Phi(x^{k+1},\,\lambda^{k+1})\| \le qR, \tag{13}$$
то полагаем $R = \|\Phi(x^{k+1},\,\lambda^{k+1})\|$, увеличиваем k на 1 и переходим к п. 1. (Шаг метода внутренней фазы.)

3. Вычисляем $(x^{k+1},\,\lambda^{k+1})$ с помощью итерации алгоритма 1. Если
$$\|\Phi(x^{k+1},\,\lambda^{k+1})\| < R, \tag{14}$$
полагаем $R = \|\Phi(x^{k+1},\,\lambda^{k+1})\|$. Увеличиваем k на 1 и переходим к п. 1. (Шаг метода внешней фазы.)

Как и в случае алгоритма с возвратами, здесь возможны лишь два (правда, несколько иные) сценария: либо все итерации алгоритма 3, начиная с некоторой, являются шагами метода внешней фазы, либо бесконечное число итераций являются итерациями метода внутренней фазы. В первом случае свойства глобальной сходимости алгоритма 3 будут теми же, что и у метода внешней фазы. Во втором случае из условия (13) легко выводится следующий результат.

Теорема 3. *Пусть функция f и отображение h дважды непрерывно дифференцируемы на $\mathbb{R}^n$. Пусть последовательность $\{(x^k,\,\lambda^k)\}$ сгенерирована алгоритмом 3, и пусть члены $(x^{k_j},\,\lambda^{k_j})$ этой последовательности получены шагами метода внутренней фазы для бесконечного числа номеров j.*

Тогда
$$\Phi(x^{k_j},\,\lambda^{k_j}) \to 0$$
при $j \to \infty$, и, в частности, если $(\bar{x},\,\bar{\lambda})$ является предельной точкой последовательности $\{(x^{k_j},\,\lambda^{k_j})\}$, то $\bar{x}$ — стационарная точка задачи (1), а $\bar{\lambda}$ — отвечающий ей множитель Лагранжа.

Наконец, сверхлинейная скорость сходимости алгоритма 3 устанавливается в следующей теореме.

Теорема 4. *Пусть функция f и отображение h дважды непрерывно дифференцируемы в окрестности стационарной точки $\bar{x}$ задачи (1), и пусть $\bar{\lambda}$ — отвечающий $\bar{x}$ некритический множитель Лагранжа. Пусть последовательность $\{(x^k,\,\lambda^k)\}$, сгенерированная алгоритмом 3, сходится к точке $(\bar{x},\,\bar{\lambda})$.*

Тогда скорость сходимости сверхлинейная.

Д о к а з а т е л ь с т в о. Если значение рекорда R меняется бесконечное число раз, то можно рассмотреть точку $(x^k,\,\lambda^k)$, сколь угодно близкую к $(\bar{x},\,\bar{\lambda})$, и такую, что $R = \|\Phi(x^k,\,\lambda^k)\|$. Тогда аналогично доказательству теоремы 2 получаем, что точка $(x^{k+1},\,\lambda^{k+1})$, полученная шагом метода внутренней фазы, принимается алгоритмом 3, на последующих итерациях алгоритм работает идентично методу s-sSQP, и сверхлинейная скорость сходимости следует из [13, теоремы 1 и 2].

Остается рассмотреть случай, когда рекордное значение $R > 0$ остается неизменным начиная с некоторой итерации. Это означает, что, начиная с этой итерации, шаги внутренней фазы не принимаются алгоритмом 3. При этом последовательность $\{\Phi(x^k,\, \lambda^k)\}$ сходится к 0 (т. к. $\{(x^k,\, \lambda^k)\}$ сходится к $(\bar{x},\, \bar{\lambda})$), и поэтому (14) будет выполняться для достаточно большого k . Но тогда на соответствующей итерации рекордной значение R должно измениться, что дает противоречие. Значит, данный случай не может иметь места. □

4. Численные примеры

Численные эксперименты были проведены в среде Python 2.7. При реализации рассматриваемых алгоритмов использовалось расширение NumPy языка Python для поддержки многомерных матриц, а также функций для операций с этими матрицами. Приводимые графические изображения были получены с помощью библиотеки Matplotlib.

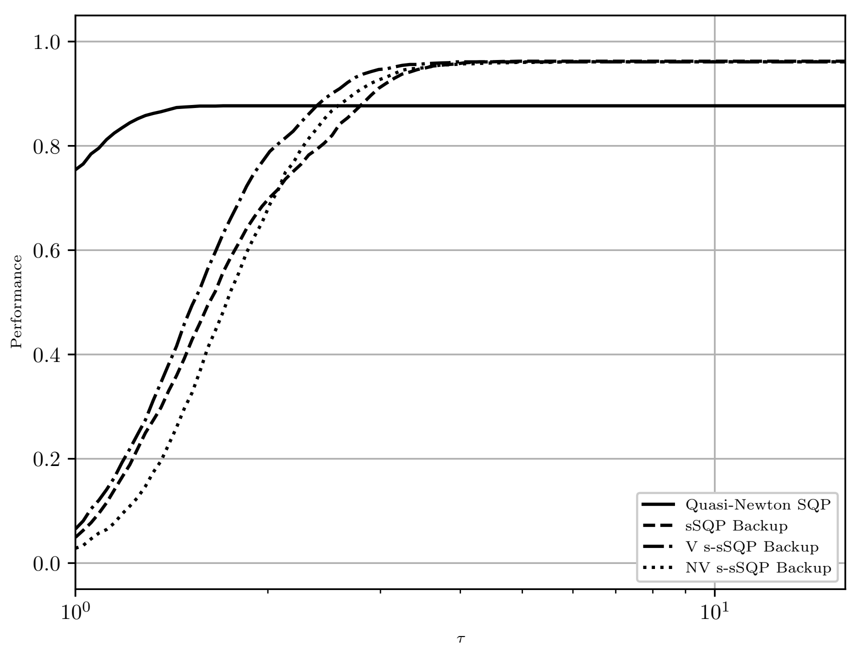

Рис. 1: Результаты для методов с возвратами.

Ниже используются следующие обозначения для тестируемых алгоритмов:

– Quasi-Newton SQP — алгоритм 1 с выбором матрицы H_k по правилу Бройдена–Флетчера–Голдфарба–Шанно с модификацией Пауэлла (см. [16, с. 536, 537]);

– sSQP Backup — алгоритм 1 из [5], т. е. алгоритм 2, в котором в качестве метода внутренней фазы вместо s-sSQP используется sSQP;

– sSQP Record — алгоритм 2 из [5], т. е. алгоритм 3, в котором в качестве метода внутренней фазы вместо s-sSQP используется sSQP;

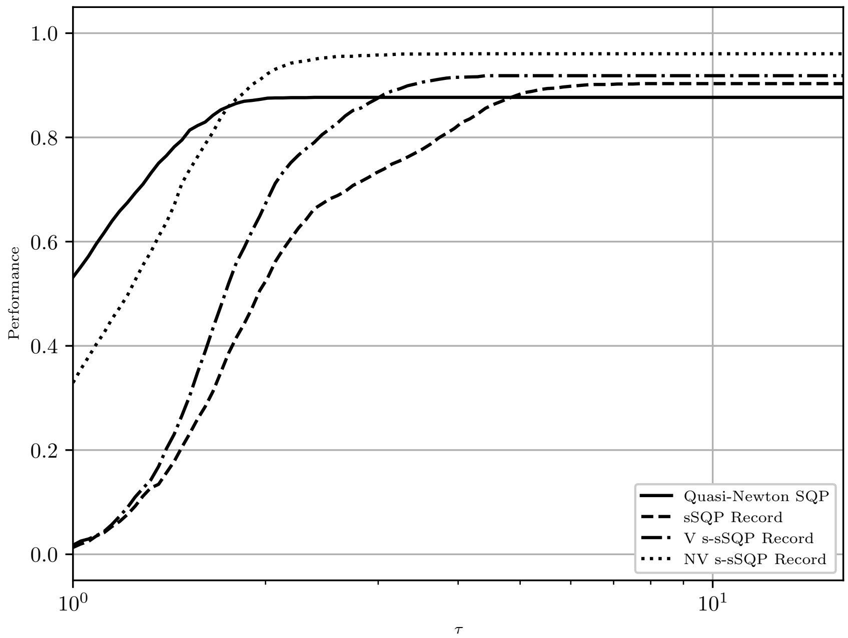

Рис. 2: Результаты для методов с рекордами.

– V s-sSQP Backup — алгоритм 2, использующий вариант метода s-sSQP с исчезающей стабилизацией;

– V s-sSQP Record — алгоритм 3, использующий вариант метода s-sSQP с исчезающей стабилизацией;

– NV s-sSQP Backup — алгоритм 2, использующий вариант метода s-sSQP с неисчезающей стабилизацией;

– MV s-sSQP Record — алгоритм 3, использующий вариант метода s-sSQP с неисчезающей стабилизацией.

Алгоритмы тестировались со следующими значениями параметров: $q = 0.9$, $\bar{c} = 1$, $\varepsilon = 0.1$, $\theta = 0.5$. Запуск считался успешным, если условие остановки

$$\|\Phi(x^k, \lambda^k)\| < 10^{-8}$$

выполнялось на некоторой из первых 500 итераций. В противном случае, а также когда на некоторой итерации не удавалось решить подзадачу метода внешней фазы, запуск прекращался и считался неудачным.

Тестирование проводилось на случайно сгенерированных задачах, для чего использовался генератор, описанный в [17], с параметром диапазона данных равным 1. А именно, для каждой рассматриваемой тройки (n, l, r) натуральных чисел генерировалось по 100 линейно-квадратичных задач вида (1), имеющих 0 своей стационарной точкой с

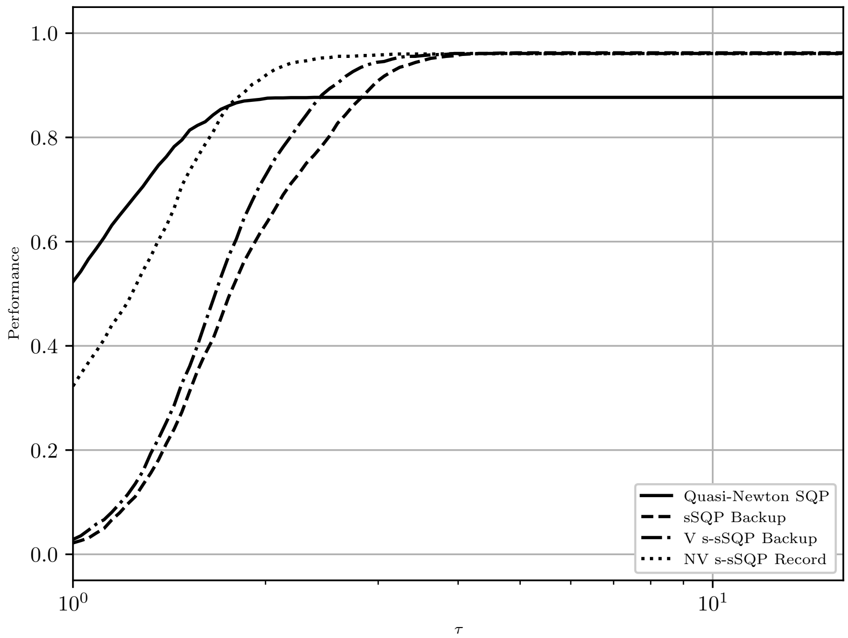

Рис. 3: Итоговые результаты.

известным отвечающим ей множителем Лагранжа $\bar{\lambda}$, и таких, что ранг матрицы Якоби ограничений в этой точке равен r. Ниже приводятся результаты, полученные с использованием всех троек (n, l, r), в которых $0 < r < l < n \leq 5$ (случай полного вырождения $r = 0$ не использовался, так как для этого случая характерно очень специальное поведение метода sSQP), а также троек вида $(n, n-1, n-2)$ и $(n, \lfloor n/2 \rfloor, \lfloor n/2 \rfloor - 1)$, для всех $n \in \{10, 15, 20, 25, 35\}$. Для каждой сгенерированной задачи выполнялось 10 запусков из случайных начальных точек в кубической окрестности $(0, \bar{\lambda})$, размер которой был равен 100.

Результаты представлены на рис. 1–3 в форме так называемых «performance profiles» [18]. Для каждого из алгоритмов значение функции, график которой изображен на рисунке, в точке $\tau \in [1, \infty)$ есть доля задач в наборе, на которых результат данного алгоритма был не более чем в τ раз хуже наилучшего результата для данной задачи среди всех сравниваемых алгоритмов. При этом считается, что результат неудачного запуска в бесконечное число раз хуже любого другого результата. В частности, значение этой функции при $\tau = 1$ соответствует доле запусков, на которых результат данного алгоритма был наилучшим. Значение функции при больших τ характеризует робастность алгоритма, т. е. долю его успешных запусков. Под «результатом» в данном случае понимается количество итераций, причем для алгоритмов с возвратами в это количество включаются отброшенные шаги sSQP и s-sSQP.

На рис. 1 и 2 представлены результаты тестирования алгоритмов с возвратами и с рекордами, соответственно. В первом случае результаты различаются мало: все алгоритмы показывают одинаковую робастность, более высокую, чем у Quasi-Newton SQP, и мало различаются по эффективности, существенно уступая Quasi-Newton SQP по этому показателю. Тем не менее, в качестве «победителей» были отобраны алгоритмы V s-sSQP Backup и sSQP Backup. Во втором же случае очевидным победителем является NV s-sSQP Record. Эта картина подтверждается и сравнительным тестированием отобранных алгоритмов, результаты которого приведены на рис. 3: NV s-sSQP Record демонстрирует ту же робастность, что и V s-sSQP Backup и sSQP Backup, и при этом существенно превосходит их по эффективности, хотя и несколько уступает по этому показателю Quasi-Newton SQP.

Таким образом, замена метода sSQP на s-sSQP на внутренней фазе гибридных алгоритмов повышает их эффективность и сохраняет робастность, более высокую, чем для метода SQP без стабилизаций. Вместе с тем, достичь эффективности последнего методам со стабилизацией по-прежнему не удается, а значит, эта проблема требует дальнейшего внимания.

Список литературы

[1] A. F. Izmailov, M. V. Solodov, *Newton-Type Methods for Optimization and Variational Problems*, Springer Series in Operations Research and Financial Engineering, Cham: Springer, 2014.

[2] А. Ф. Измаилов, М. В. Солодов, *Численные методы оптимизации*, 2-е изд., перераб. и доп., Физматлит, М., 2008.

[3] S. J. Wright, “Superlinear convergence of a stabilized SQP method to a degenerate solution”, *Comput. Optim. Appl.*, **11** (1998), 253–275.

[4] A. F. Izmailov, M. V. Solodov, “Stabilized SQP revisited”, *Math. Program.*, **133** (2012), 93–120.

[5] А. Ф. Измаилов, А. М. Крылова, Е. И. Усков, “Гибридная глобализация стабилизированного метода последовательного квадратичного программирования”, *Теоретические и прикладные задачи нелинейного анализа*, ВЦ РАН, М., 2011, 47–66.

[6] P. E. Gill, D. P. Robinson, “A primal-dual augmented Lagrangian”, *Comput. Optim. Appl.*, **51** (2012), 1–25.

[7] P. E. Gill, D. P. Robinson, “A globally convergent stabilized SQP method”, *SIAM J. Optim. Appl.*, **23** (2013), 1983–2010.

[8] D. Fernandez, E. A. Pilotta, G. A. Torres, “An inexact restoration strategy for the globalization of the sSQP method”, *Comput. Optim. Appl.*, **54** (2013), 595–617.

[9] A. F. Izmailov, M. V. Solodov, E. I. Uskov, “Combining stabilized SQP with the augmented Lagrangian algorithm”, *Comput. Optim. Appl.*, **62** (2015), 33–73.

[10] A. F. Izmailov, M. V. Solodov, E. I. Uskov, “Globalizing stabilized sequential quadratic programming method by smooth primal-dual exact penalty function”, *J. Optim. Theory. Appl.*, **69** (2016), 148–178.

[11] P. E. Gill, V. Kungurtsev, D. P. Robinson, “A stabilized SQP method: global convergence”, *IMA Journal of Numerical Analysis*, **37** (2017), 407–443.

[12] P. E. Gill, V. Kungurtsev, D. P. Robinson, “A stabilized SQP method: superlinear convergence”, *Math. Program.*, **163** (2017), 369–410.

[13] A. F. Izmailov, E. I. Uskov, “Subspace-stabilized sequential quadratic programming”, *Comput. Optim. Appl.*, **67** (2017), 129–154.

[14] E. M. E. Mostafa, L. N. Vicente, S. J. Wright, “Numerical behavior of a stabilized SQP method for degenerate NLP problems”, *Global Optimization and Constraint Satisfaction.*, Lecture Notes in Computer Science 2861., Springer, Berlin, 2003, 123–141.

[15] A. F. Izmailov, M. V. Solodov, “Newton-type methods for optimization problems without constraint qualifications”, *SIAM J. Optim.*, **15** (2004), 210–228.

[16] J. Nocedal, S. J. Wright, *Numerical Optimization*, Second ed., Springer, New York, 2006.

[17] A. F. Izmailov, M. V. Solodov, “On attraction of Newton-type iterates to multipliers violating second-order sufficiency conditions”, *Math. Program.*, **117** (2009), 271–304.

[18] E. D. Dolan, J. J. More, “Benchmarking optimization software with performance profiles”, *Math. Program.*, **91** (2002), 201–213.

References

[1] A. F. Izmailov, M. V. Solodov, *Newton-Type Methods for Optimization and Variational Problems*, Springer Series in Operations Research and Financial Engineering, Cham: Springer, 2014.

[2] A. F. Izmailov, M. V. Solodov, *Numerical Methods of Optimization*, Second ed., Fizmatlit, Moscow, 2008 (In Russian).

[3] S. J. Wright, “Superlinear convergence of a stabilized SQP method to a degenerate solution”, *Comput. Optim. Appl.*, **11** (1998), 253–275.

[4] A. F. Izmailov, M. V. Solodov, “Stabilized SQP revisited”, *Math. Program.*, **133** (2012), 93–120.

[5] A. F. Izmailov, A. M. Krylova, E. I. Uskov, “Hybrid globalization of stabilized sequential quadratic programming method”, *Theoretical and applied problems of nonlinear analysis*, Computing Center RAS, Moscow, 2011, 47–66 (In Russian).

[6] P. E. Gill, D. P. Robinson, “A primal-dual augmented Lagrangian”, *Comput. Optim. Appl.*, **51** (2012), 1–25.

[7] P. E. Gill, D. P. Robinson, “A globally convergent stabilized SQP method”, *SIAM J. Optim. Appl.*, **23** (2013), 1983–2010.

[8] D. Fernandez, E. A. Pilotta, G. A. Torres, “An inexact restoration strategy for the globalization of the sSQP method”, *Comput. Optim. Appl.*, **54** (2013), 595–617.

[9] A. F. Izmailov, M. V. Solodov, E. I. Uskov, “Combining stabilized SQP with the augmented Lagrangian algorithm”, *Comput. Optim. Appl.*, **62** (2015), 33–73.

[10] A. F. Izmailov, M. V. Solodov, E. I. Uskov, “Globalizing stabilized sequential quadratic programming method by smooth primal-dual exact penalty function”, *J. Optim. Theory. Appl.*, **69** (2016), 148–178.

[11] P. E. Gill, V. Kungurtsev, D. P. Robinson, “A stabilized SQP method: global convergence”, *IMA Journal of Numerical Analysis*, **37** (2017), 407–443.

[12] P. E. Gill, V. Kungurtsev, D. P. Robinson, “A stabilized SQP method: superlinear convergence”, *Math. Program.*, **163** (2017), 369–410.

[13] A. F. Izmailov, E. I. Uskov, “Subspace-stabilized sequential quadratic programming”, *Comput. Optim. Appl.*, **67** (2017), 129–154.

[14] E. M. E. Mostafa, L. N. Vicente, S. J. Wright, “Numerical behavior of a stabilized SQP method for degenerate NLP problems”, *Global Optimization and Constraint Satisfaction.*, Lecture Notes in Computer Science 2861., Springer, Berlin, 2003, 123–141.

[15] A. F. Izmailov, M. V. Solodov, “Newton-type methods for optimization problems without constraint qualifications”, *SIAM J. Optim.*, **15** (2004), 210–228.

[16] J. Nocedal, S. J. Wright, *Numerical Optimization*, Second ed., Springer, New York, 2006.

[17] A. F. Izmailov, M. V. Solodov, “On attraction of Newton-type iterates to multipliers violating second-order sufficiency conditions”, *Math. Program.*, **117** (2009), 271–304.

[18] E. D. Dolan, J. J. More, “Benchmarking optimization software with performance profiles”, *Math. Program.*, **91** (2002), 201–213.

Информация об авторах

Журбенко Николай Георгиевич, кандидат физико-математических наук, старший научный сотрудник. Институт кибернетики им. В. М. Глушкова НАН Украины, г. Киев, Украина. E-mail: zhurbnick@gmail.com
ORCID: https://orcid.org/0000-0001-5123-8607

Измаилов Алексей Феридович, доктор физико-математических наук, профессор кафедры исследования операций. Московский государственный университет им. М. В. Ломоносова, г. Москва, Российская Федерация. E-mail: izmaf@ccas.ru
ORCID: https://orcid.org/0000-0001-9851-0524

Усков Евгений Иванович, кандидат физико-математических наук, научный сотрудник. Тамбовский государственный университет им. Г. Р. Державина, г. Тамбов, Российская Федерация. E-mail: akurennoy@cs.msu.ru
ORCID: https://orcid.org/0000-0002-3639-0317

Конфликт интересов отсутствует.

Для контактов:
Измаилов Алексей Феридович
E-mail: izmaf@ccas.ru

Поступила в редакцию 26.01.2019 г.
Поступила после рецензирования 18.03.2019 г.
Принята к публикации 20.05.2019 г.

Information about the authors

Nikolay G. Zhurbenko, Candidate of Physics and Mathematics, Associate Professor of the Functional Analysis Department. V. M. Glushkov Institute of Cybernetics of NAS of Ukraine, Kiev, Ukraine. E-mail: zhurbnick@gmail.com
ORCID: https://orcid.org/0000-0001-5123-8607

Alexey F. Izmailov, Doctor of Physics and Mathematics, Professor of the Operations Research Department. Lomonosov Moscow State University, Moscow, the Russian Federation. E-mail: izmaf@ccas.ru
ORCID: https://orcid.org/0000-0001-9851-0524

Evgeniy I. Uskov, Candidate of Physics and Mathematics, Researcher. Derzhavin Tambov State University, Tambov, the Russian Federation. E-mail: akurennoy@cs.msu.ru
ORCID: https://orcid.org/0000-0002-3639-0317

There is no conflict of interests.

Corresponding author:
Alexey F. Izmailov
E-mail: izmaf@ccas.ru

Received 26 January 2019
Reviewed 18 March 2019
Accepted for press 20 May 2019