ТРУДЫ ИНСТИТУТА МАТЕМАТИКИ И МЕХАНИКИ УрО РАН

Том 24 № 2 2018

УДК 519.85

АДАПТИВНЫЕ АЛГОРИТМЫ ЗЕРКАЛЬНОГО СПУСКА В ЗАДАЧАХ ВЫПУКЛОГО ПРОГРАММИРОВАНИЯ С ЛИПШИЦЕВЫМИ ОГРАНИЧЕНИЯМИ[1]

Ф. С. Стонякин, М. С. Алкуса, А. Н. Степанов, М. А. Баринов

Работа посвящена новым модификациям недавно предложенных адаптивных методов зеркального спуска для задач выпуклой минимизации в случае нескольких выпуклых функциональных ограничений. Предложены адаптивные методы зеркального спуска для задач двух типов. Первый тип — задачи с липшицевым (вообще говоря, негладким) целевым функционалом. Второй тип — задачи с липшицевым градиентом целевого функционала. Рассматривается также случай негладкого целевого функционала, равного максимуму гладких функционалов с липшицевым градиентом. Во всех случаях функциональные ограничения считаются выпуклыми, липшицевыми и, вообще говоря, негладкими. Предлагаемые методы позволяют сэкономить время работы алгоритма за счет рассмотрения не всех функциональных ограничений на непродуктивных шагах. Получены оценки на скорость сходимости рассматриваемых методов. Эти оценки демонстрируют оптимальность методов с точки зрения нижних оракульных оценок. Приведены результаты численных экспериментов, иллюстрирующие преимущества предлагаемой методики для некоторых примеров.

Ключевые слова: адаптивный метод зеркального спуска, липшицев функционал, липшицев градиент, продуктивный шаг, непродуктивный шаг.

F. S. Stonyakin, M. S. Alkousa, A. N. Stepanov, M. A. Barinov. Adaptive mirror descent algorithms in convex programming problems with Lipschitz constraints.

The paper is devoted to new modifications of recently proposed adaptive Mirror Descent methods for convex minimization problems in the case of several convex functional constraints. Methods for problems of two types are considered. In problems of the first type, the objective functional is Lipschitz (generally, nonsmooth). In problems of the second type, the gradient of the objective functional is Lipschitz. We also consider the case of a nonsmooth objective functional equal to the maximum of smooth functionals with Lipschitz gradient. In all the cases, the functional constraints are assumed to be Lipschitz and, generally, nonsmooth. The proposed modifications make it possible to reduce the running time of the algorithm due to skipping some of the functional constraints at nonproductive steps. We derive bounds for the convergence rate, which show that the methods under consideration are optimal from the viewpoint of lower oracle estimates. The results of numerical experiments illustrating the advantages of the proposed procedure for some examples are presented.

Keywords: adaptive Mirror Descent, Lipschitz functional, Lipschitz gradient, productive step, nonproductive step.

MSC: 90C25, 90C06, 49J52

DOI: 10.21538/0134-4889-2018-24-2-266-279

Введение

Задачи минимизации негладкого функционала с ограничениями возникают в широком классе проблем современной large-scale оптимизации и ее приложений [1; 2]. Для таких задач имеется множество методов, среди которых можно отметить метод зеркального спуска (МЗС) [3; 4] для задач выпуклой минимизации.

Отметим, что в случае негладкого целевого функционала или функциональных ограничений естественно использовать субградиентные методы, восходящие к хорошо известным

[1]Исследование Ф. С. Стонякина (доказательство теоремы 2) выполнено при поддержке Российского фонда фундаментальных исследований, проект 18-31-00219. Исследование Ф .С. Стонякина и А. Н. Степанова в части доказательства следствий 1 и 2, проведения численных экспериментов и заключительных замечаний выполнены при частичной поддержке гранта Президента Российской Федерации для государственной поддержки молодых российских ученых — кандидатов наук, код МК-176.2017.1 .

работам [5; 6]. Метод зеркального спуска возник для безусловных задач в [4; 7] как аналог стандартного субградиентного метода с неевклидовым проектированием. Для условных задач аналог этого метода был предложен в [4] (см. также [8]). Проблема адаптивного выбора шага без использования констант Липшица рассмотрена в [9] для задач без ограничений, а также в [8] для задач с функциональными ограничениями. Однако критерии остановки для указанных методов зеркального спуска оставались неадаптивными.

Недавно в [10] были предложены алгоритмы зеркального спуска как с адаптивным выбором шага, так и с адаптивным критерием остановки. При этом помимо случая липшицевых целевого функционала и функционального ограничения был предложен оптимальный метод для класса условных задач выпуклой минимизации со специальными условиями роста (будем называть их нестандартными) целевого функционала. Например, в задачах с квадратичными функционалами мы сталкиваемся с ситуацией, когда такой функционал не удовлетворяет обычному свойству Липшица (или константа Липшица достаточно большая), но градиент удовлетворяет условию Липшица. Для задач такого типа в (см. [10, п. 3.3]) был предложен адаптивный алгоритм зеркального спуска на базе идеологии [11; 12].

В этой статье предложены новые модификации адаптивных методов зеркального спуска из [10, пп. 3.1 и 3.3] для задач выпуклой минимизации в случае, когда имеется несколько функциональных ограничений. Отдельно рассмотрены методы для случаев липшицевости целевого функционала (разд. 2), а также для случая нестандартных условий роста целевого функционала (разд. 3). Предлагаемые модификации позволяют сэкономить время работы алгоритма за счет рассмотрения не всех функциональных ограничений на непродуктивных шагах. Получены оценки на скорость сходимости рассматриваемых методов (теоремы 1 и 2, следствия 1 и 2). Оценки необходимого числа итераций, представленные в разд. 2 и 3, указывают на оптимальность предлагаемых методов с точки зрения нижних оракульных оценок [4]. В разд. 4 приведены результаты численных экспериментов, иллюстрирующие преимущества предлагаемых нами методов для некоторых примеров.

1. Постановка задачи и необходимые вспомогательные понятия

Пусть $(E, \|\cdot\|)$ — конечномерное нормированное векторное пространство и E^* — сопряженное пространство к E со стандартной нормой

$$||y||_* = \max_x\{\langle y, x\rangle, ||x|| \leq 1\},$$

где $\langle y, x\rangle$ — значение линейного непрерывного функционала y в точке $x \in E$.

Пусть $X \subset E$ — замкнутое выпуклое множество. Рассмотрим набор выпуклых субдифференцируемых функционалов f и $g_m : X \to \mathbb{R}$ для всякого $m = \overline{1, M}$. Также предположим, что все функционалы g_m удовлетворяют условию Липшица с некоторой константой M_g

$$|g_m(x) - g_m(y)| \leqslant M_g||x - y|| \;\; \forall\, x, y \in X, \quad m = \overline{1, M}. \tag{1.1}$$

Будем рассматривать следующий тип задач выпуклой оптимизации:

$$f(x) \to \min_{x \in X}, \tag{1.2}$$

где

$$g_m(x) \leqslant 0 \;\; \forall\, m = \overline{1, M}. \tag{1.3}$$

Сделаем предположение о разрешимости задачи (1.2), (1.3).

Для дальнейших рассуждений нам потребуются вспомогательные понятия (см., например, [9]). Введем так называемую *прокс-функцию* $d : X \to \mathbb{R}$, обладающую свойством непрерывной дифференцируемости и 1-сильной выпуклости относительно нормы $\|\cdot\|$, т. е.

$$\langle \nabla d(x) - \nabla d(y), x - y\rangle \geqslant \|x - y\|^2 \;\; \forall\, x, y \in X,$$

и предположим, что $\min\limits_{x\in X} d(x) = d(0)$. Будем полагать, что существует такая константа $\Theta_0 > 0$, что $d(x_*) \leqslant \Theta_0^2$, где x_* — точное решение (1.2), (1.3). Если имеется множество решений X_*, то мы предполагаем, что для константы Θ_0 $\min\limits_{x_*\in X_*} d(x_*) \leqslant \Theta_0^2$. Для всех $x, y \in X$ рассмотрим соответствующую дивергенцию Брегмана

$$V(x,y) = d(y) - d(x) - \langle \nabla d(x), y - x\rangle.$$

В зависимости от постановки конкретной задачи возможны различные подходы к определению прокс-структуры задачи и соответствующей дивергенции Брегмана (см., например, [9]). Стандартно определим оператор проектирования

$$\mathrm{Mirr}_x(p) = \arg\min_{u\in X}\big\{\langle p, u\rangle + V(x,u)\big\} \ \text{ для всяких } x \in X \text{ и } p \in E^*.$$

Сделаем предположение о том, что оператор $\mathrm{Mirr}_x(p)$ легко вычислим.

Ясно, что вместо набора выпуклых функциональных ограничений $\{g_m(\cdot)\}_{m=1}^M$ можно рассмотреть одно ограничение $g : X \to \mathbb{R}$

$$g(x) = \max_{m=\overline{1,M}} g_m(x), \quad |g(x) - g(y)| \leqslant M_g||x - y|| \ \ \forall\, x, y \in X.$$

Для условных задач с целевым функционалом f с одним выпуклым субдифференцируемым функциональным ограничением в работе [10] предложено два метода зеркального спуска. Сходимость первого из них доказана для случая липшицевости целевого функционала (см. [10, п. 3.1]), сходимость же второго обоснована, в частности, в предположении липшицевости градиента f (см. [10, п. 3.3]). Напомним эти алгоритмы (см. алгоритмы 1 и 2).

Алгоритм 1. Адаптивный зеркальный спуск (липшицев целевой функционал) [10, алгоритм 1]

Require: $\varepsilon > 0$, $\Theta_0 : d(x_*) \leqslant \Theta_0^2$
$x^0 = \mathrm{argmin}_{x\in X}\, d(x)$
$I =: \varnothing$
$N \leftarrow 0$
repeat
  if $g(x^N) \leqslant \varepsilon$ **then**
    $M_N = ||\nabla f(x^N)||_*$, $h_N = \dfrac{\varepsilon}{M_N^2}$
    $x^{N+1} = \mathrm{Mirr}_{x^N}(h_N \nabla f(x^N))$ // *"продуктивные шаги"*
    $N \to I$
  else
    $M_N = ||\nabla g(x^N)||_*$, $h_N = \dfrac{\varepsilon}{M_N^2}$
    $x^{N+1} = \mathrm{Mirr}_{x^N}(h_N \nabla g(x^N))$ // *"непродуктивные шаги"*
  end if
  $N \leftarrow N + 1$
until $\sum\limits_{j=0}^{N-1} \dfrac{1}{M_j^2} \geqslant 2\dfrac{\Theta_0^2}{\varepsilon^2}$
Ensure: $\bar{x}^N := \dfrac{\sum\limits_{k\in I} x^k h_k}{\sum\limits_{k\in I} h_k}$

А л г о р и т м 2. Адаптивный зеркальный спуск (нестандартные условия роста) [10, алгоритм 3]

Require: $\varepsilon > 0, \quad \Theta_0 : d(x_*) \leqslant \Theta_0^2$

$x^0 = \operatorname{argmin}_{x \in X} d(x)$
$I =: \varnothing$
$N \leftarrow 0$
repeat
  if $g(x^N) \leqslant \varepsilon$ **then**
    $h_N \leftarrow \dfrac{\varepsilon}{||\nabla f(x^N)||_*}$
    $x^{N+1} \leftarrow \mathrm{Mirr}_{x^N}(h_N \nabla f(x^N))$ // *"продуктивные шаги"*
    $N \to I$
  else
    // $(g(x^N) > \varepsilon)$
    $h_N \leftarrow \dfrac{\varepsilon}{||\nabla g(x^N)||_*^2}$
    $x^{N+1} \leftarrow \mathrm{Mirr}_{x^N}(h_N \nabla g(x^N))$ // *"непродуктивные шаги"*
  end if
  $N \leftarrow N + 1$
until $\Theta_0^2 \leqslant \dfrac{\varepsilon^2}{2}\Big(|I| + \sum\limits_{k \notin I} \dfrac{1}{||\nabla g(x^k)||_*^2}\Big)$

Ensure: $\bar{x}^N := \operatorname{argmin}_{x^k,\, k \in I} f(x^k)$

З а м е ч а н и е 1. В обоих методах в ходе работы необходимо делить на нормы субградиентов целевого функционала $||\nabla f(x^k)||_*^2$ или ограничения $||\nabla g(x^k)||_*^2$. В связи с этим прокомментируем ситуацию, когда $\nabla f(x^k) = 0$ или $\nabla g(x^k) = 0$. Если верно $\nabla f(x^k) = 0$, то ясно, что x^k — точное решение задачи минимизации f вне зависимости от ограничений и в этом случае работу алгоритма нужно останавливать. Условимся не оговаривать это обстоятельство отдельно в листингах алгоритмов. Если же $\nabla g(x^k) = 0$, то x^k — точка глобального минимума g. В сочетании с условием $g(x^k) > \varepsilon$ это означает, что $g(x) > \varepsilon > 0$ на всей области определения и тогда поставленная задача просто неразрешима. Ввиду изложенных обстоятельств условимся здесь и всюду далее полагать, что $\nabla f(x^k) \neq 0$ и $\nabla g(x^k) \neq 0$.

В настоящей работе мы покажем, что на “непродуктивных” шагах ($k \notin I$) можно вместо субградиента ограничения max-типа $g(x) = \max\limits_{m=\overline{1,M}} g_m(x)$ использовать субградиент любого из функционалов g_m, для которого верно $g_m(x^k) > \varepsilon$.

2. Модификация алгоритма зеркального спуска в случае липшицевости целевого функционала

Будем предполагать, что целевой функционал $f : X \to \mathbb{R}$ удовлетворяет условию Липшица

$$|f(x) - f(y)| \leqslant M_f ||x - y|| \quad \forall\, x, y \in X. \tag{2.1}$$

При таком соглашении мы введем в рассмотрение алгоритм 3 для задачи (1.2), (1.3).

Напомним один известный факт (см., например, [9]). Пусть $f : X \to \mathbb{R}$ — выпуклый субдифференцируемый функционал. Для произвольного $y \in X$, некоторого $h > 0$ и произвольного $x \in X$ справедливо неравенство

$$h\langle \nabla f(y), y - x\rangle \leqslant \frac{h^2}{2}||\nabla f(y)||_*^2 + V(y, x) - V(z, x), \quad \text{где} \;\; z = \mathrm{Mirr}_y(h \nabla f(y)). \tag{2.2}$$

А л г о р и т м 3. Модифицированный адаптивный зеркальный спуск (липшицев целевой функционал)

Require: $\varepsilon > 0, \quad \Theta_0 : d(x_*) \leqslant \Theta_0^2$

$x^0 = \operatorname{argmin}_{x \in X} d(x)$

$I =: \varnothing$

$N \leftarrow 0$

repeat

  if $g(x^N) \leqslant \varepsilon$ **then**

    $M_N = ||\nabla f(x^N)||_*$

    $h_N = \dfrac{\varepsilon}{M_N^2}, \ x^{N+1} = \mathrm{Mirr}_{x^N}(h_N \nabla f(x^N))$ *// "продуктивные шаги"*

    $N \to I$

  else

    // $(g_{m(N)}(x^N) > \varepsilon)$ *для некоторого* $m(N) \in \{1, \ldots, M\}$

    $M_N = ||\nabla g_{m(N)}(x^N)||_*, \ h_N = \dfrac{\varepsilon}{M_N^2}$

    $x^{N+1} = \mathrm{Mirr}_{x^N}(h_N \nabla g_{m(N)}(x^N))$ *// "непродуктивные шаги"*

  end if

  $N \leftarrow N + 1$

until $\sum\limits_{j=0}^{N-1} \dfrac{1}{M_j^2} \geqslant 2\dfrac{\Theta_0^2}{\varepsilon^2}$

Ensure: $\bar{x}^N := \dfrac{\sum\limits_{k \in I} x^k h_k}{\sum\limits_{k \in I} h_k}$

Теорема 1. *Пусть* $\varepsilon > 0$ *— фиксированное число и выполнен критерий остановки алгоритма* 3. *Тогда* $\bar{x}^N$ *есть* ε*-решение задачи* (1.2), (1.3):

$$f(\bar{x}^N) - f^* \leqslant \varepsilon, \quad g(\bar{x}^N) \leqslant \varepsilon.$$

При этом алгоритм 3 *работает не более*

$$N = \left\lceil \frac{2\max\{M_f^2, M_g^2\}\Theta_0^2}{\varepsilon^2} \right\rceil \tag{2.3}$$

итераций.

Д о к а з а т е л ь с т в о. Мы будем отталкиваться от [10, п. 3.1] с изменениями, соответствующими случаю нескольких ограничений. Согласно определению $\bar{x}^N$ (см. алгоритм 3) и ввиду выпуклости целевого функционала f

$$\sum_{k \in I} h_k f(\bar{x}^N) - f(x_*) \leqslant \sum_{k \in I} h_k \big(f(x^k) - f(x_*)\big). \tag{2.4}$$

Далее, по (2.2)

$$h_k\big(f(x^k) - f(x)\big) \leqslant \frac{h_k^2}{2}\|\nabla f(x^k)\|_*^2 + V(x^k, x) - V(x^{k+1}, x) \quad \forall\, k \in I,$$

$$h_k\big(g_{m(k)}(x^k) - g_{m(k)}(x)\big) \leqslant \frac{h_k^2}{2}\|\nabla g_{m(k)}(x^k)\|_*^2 + V(x^k, x) - V(x^{k+1}, x) \quad \forall\, k \notin I,$$

и ввиду выбора величины шагов h_k в алгоритме 3 имеем

$$\sum_{k\in I} h_k\big(f(x^k)-f(x_*)\big)+\sum_{k\notin I} h_k\big(g_{m(k)}(x^k)-g_{m(k)}(x_*)\big)$$

$$\leqslant \sum_{k\in I}\frac{h_k^2||\nabla f(x^k)||_*^2}{2}+\sum_{k\notin I}\frac{h_k^2||\nabla g_{m(k)}(x^k)||_*^2}{2}+\sum_{k=0}^{N-1}\big(V(x^k,x_*)-V(x^{k+1},x_*)\big)\leqslant \frac{\varepsilon}{2}\sum_{k=0}^{N-1}h_k+\Theta_0^2,$$

где на непродуктивном шаге под $g_{m(k)}$ мы понимаем любое из ограничений, для которого верно неравенство $g_{m(k)}(x^k)>\varepsilon$. Поскольку для всякого $k\notin I$ $g_{m(k)}(x^k)-g_{m(k)}(x_*)\geqslant g_{m(k)}(x^k)>\varepsilon$, то с учетом (2.4) мы имеем

$$\sum_{k\in I} h_k\big(f(\bar{x}^N)-f(x_*)\big)<\frac{\varepsilon}{2}\sum_{k=0}^{N-1}h_k-\varepsilon\sum_{k\notin I}h_k+\Theta_0^2$$

$$=\varepsilon\sum_{k\in I}h_k-\frac{\varepsilon^2}{2}\sum_{k\in I}\frac{1}{||\nabla f(x^k)||_*^2}-\frac{\varepsilon^2}{2}\sum_{k\notin I}\frac{1}{||\nabla g_{m(k)}(x^k)||_*^2}+\Theta_0^2\leqslant\varepsilon\sum_{k\in I}h_k.$$

Ввиду строгости последнего неравенства мы имеем, что множество продуктивных шагов I непусто. Ясно, что для всякого $k\in I$ справедливо неравенство $g(x^k)\leqslant\varepsilon$. Тогда ввиду выпуклости g верно

$$\sum_{k\in I}h_k g(\bar{x}^N)\leqslant\sum_{k\in I}h_k g(x^k)\leqslant\varepsilon\sum_{k\in I}h_k.$$

Отметим также, что ввиду условий Липшица для целевого функционала и функционального ограничения (см. (1.1) и (2.1)) на любой итерации работы алгоритма 3 справедливы неравенства $||\nabla f(x^k)||_*\leqslant M_f$ и $||\nabla g_{m(k)}(x^k)||_*\leqslant M_g$. Поэтому критерий остановки алгоритма 3 будет заведомо выполнен не более чем после (2.3) итераций работы, что завершает доказательство теоремы.

З а м е ч а н и е 2. Ввиду липшицевости и, вообще говоря, негладкости целевого функционала и ограничений оценка (2.3) на число итераций означает, что предложенный метод оптимален с точки зрения оракульных оценок [9]: для достижения требуемой точности ε решения задачи (1.2), (1.3) достаточно $O\Big(\dfrac{1}{\varepsilon^2}\Big)$ итераций алгоритма 3.

3. Модифицированный алгоритм зеркального спуска в случае нестандартных условий роста целевого функционала

Теперь рассмотрим метод для класса условных задач выпуклой минимизации с нестандартными условиями роста целевого функционала. Например, в задачах с квадратичными целевыми функционалами мы сталкиваемся с ситуацией, когда такой функционал не удовлетворяет обычному свойству Липшица (или константа Липшица достаточно большая), но градиент удовлетворяет условию Липшица. Если целевой функционал f не удовлетворяет свойству Липшица, то критерии остановки алгоритмов 1 и 3 не позволяют получить оценку вида (2.3) и обосновать оптимальность метода с точки зрения нижних оракульных оценок. Такие проблемы возникают и для более широкого класса уже негладких целевых функционалов

$$f(x)=\max_{1\leqslant i\leqslant m}f_i(x), \tag{3.1}$$

где

$$f_i(x)=\frac{1}{2}\langle A_i x,x\rangle-\langle b_i,x\rangle+\alpha_i,\quad i=1,\dots,m, \tag{3.2}$$

в случае, когда A_i $(i=1,\dots,m)$ — положительно определенные матрицы $x^TA_ix\geqslant 0\;\forall\,x\in X$.

Отметим, что функционалы вида (3.1), (3.2) возникают в задачах проектирования механических конструкций Truss Topology Design со взвешенными балками (см., например, презентацию [12]).

Для задач (1.2), (1.3) с целевыми функционалами вида (3.1), (3.2) в [10, п. 3.3] был предложен адаптивный алгоритм типа зеркального спуска на базе идеологии [11; 12]. Мы введем в рассмотрение модификацию этого метода в случае наличия нескольких функциональных ограничений. Аналогично [11, п. 3.2.2] для всякого субградиента $\nabla f(x)$ целевого функционала f в точке $y \in X$ введем вспомогательную величину

$$v_f(x,y) = \begin{cases} \left\langle \dfrac{\nabla f(x)}{\|\nabla f(x)\|_*}, x - y \right\rangle, & \nabla f(x) \neq 0, \\ 0, & \nabla f(x) = 0, \end{cases} \quad x \in X. \tag{3.3}$$

Рассмотрим следующий алгоритм адаптивного зеркального спуска для задач (1.2), (1.3).

А л г о р и т м 4. Модифицированный адаптивный зеркальный спуск (липшицев целевой функционал)

Require: $\varepsilon > 0, \quad \Theta_0 : d(x_*) \leqslant \Theta_0^2$

$x^0 = \mathrm{argmin}_{x \in X}\, d(x)$

$I =: \varnothing$

$N \leftarrow 0$

repeat

  if $g(x^N) \leqslant \varepsilon$ **then**

    $h_N \leftarrow \dfrac{\varepsilon}{||\nabla f(x^N)||_*}$

    $x^{N+1} \leftarrow \mathrm{Mirr}_{x^N}(h_N \nabla f(x^N))$ // *"продуктивные шаги"*

    $N \rightarrow I$

  else

    // $(g_{m(N)}(x^N) > \varepsilon)$ *для некоторого* $m(N) \in \{1, \ldots, M\}$

    $h_N \leftarrow \dfrac{\varepsilon}{||\nabla g_{m(N)}(x^N)||_*^2}$

    $x^{N+1} \leftarrow \mathrm{Mirr}_{x^N}(h_N \nabla g_{m(N)}(x^N))$ // *"непродуктивные шаги"*

  end if

  $N \leftarrow N + 1$

until $\Theta_0^2 \leqslant \dfrac{\varepsilon^2}{2}\Big(|I| + \sum\limits_{k \notin I} \dfrac{1}{||\nabla g_{m(k)}(x^k)||_*^2}\Big)$

Ensure: $\bar{x}^N := \mathrm{argmin}_{x^k,\, k \in I}\, f(x^k)$

Справедлива следующая

Теорема 2. *Пусть* $\varepsilon > 0$ — *фиксированное число и выполнен критерий остановки алгоритма* 4. *Тогда*

$$\min_{k \in I} v_f(x^k, x_*) < \varepsilon. \tag{3.4}$$

Отметим, что алгоритм 4 *работает не более*

$$N = \left\lceil \frac{2\max\{1, M_g^2\}\Theta_0^2}{\varepsilon^2} \right\rceil \tag{3.5}$$

итераций.

Д о к а з а т е л ь с т в о. Мы будем отталкиваться от [10, п. 3.3] с изменениями, соответствующими случаю нескольких ограничений. Пусть $[N] = \{k \in \overline{0, N-1}\}$, $J = [N] \setminus I$ — набор номеров непродуктивных шагов.

Для продуктивных шагов ввиду (2.2) имеем, что

$$h_k\langle\nabla f(x^k), x^k - x\rangle \leqslant \frac{h_k^2}{2}||\nabla f(x^k)||_*^2 + V(x^k, x) - V(x^{k+1}, x).$$

Примем во внимание $\frac{h_k^2}{2}||\nabla f(x^k)||_*^2 = \frac{\varepsilon^2}{2}$, имеем

$$h_k\langle\nabla f(x^k), x^k - x\rangle = \varepsilon\Big\langle\frac{\nabla f(x^k)}{||\nabla f(x^k)||_*},\, x^k - x\Big\rangle = \varepsilon v_f(x^k, x) \leqslant \frac{\varepsilon^2}{2} + V(x^k, x_*) - V(x^{k+1}, x_*). \quad (3.6)$$

Аналогично для “непродуктивных” шагов $k \in J$ (под $g_m(k)$ мы понимаем любое ограничение, для которого $g_m(k) > \varepsilon$) по (2.2)

$$h_k\big(g_{m(k)}(x^k) - g_{m(k)}(x_*)\big) \leqslant \frac{h_k^2}{2}||\nabla g_{m(k)}(x^k)||_*^2 + V(x^k, x_*) - V(x^{k+1}, x_*)$$

$$= \frac{\varepsilon^2}{2||\nabla g_{m(k)}(x^k)||_*^2} + V(x^k, x_*) - V(x^{k+1}, x_*).$$

Из (1.1) и (1.2) при $x = x_*$ имеем

$$\varepsilon\sum_{k\in I} v_f(x^k, x_*) + \sum_{k\in J}\frac{\varepsilon^2(g_{m(k)}(x^k) - g_{m(k)}(x_*))}{2||\nabla g_{m(k)}(x^k)||_*^2} \leqslant \frac{\varepsilon^2}{2}\cdot|I| + \sum_{k=0}^{N-1}(V(x^k, x_*) - V(x^{k+1}, x_*)). \quad (3.7)$$

Заметим, что для любого $k \in J$

$$g_{m(k)}(x^k) - g_{m(k)}(x_*) \geqslant g_{m(k)}(x^k) > \varepsilon \quad (3.8)$$

и ввиду $\sum_{k=0}^{N-1}(V(x^k, x_*) - V(x^{k+1}, x_*)) \leqslant \Theta_0^2$ неравенство (3.7) может быть преобразовано следующим образом:

$$\varepsilon\sum_{k\in I} v_f(x^k, x_*) \leqslant |I|\cdot\frac{\varepsilon^2}{2} + \Theta_0^2 - \sum_{k\in J}\frac{\varepsilon^2}{2||\nabla g_{m(k)}(x^k)||_*^2}, \quad \sum_{k\in I} v_f(x^k, x_*) \geqslant |I|\min_{k\in I} v_f(x^k, x_*).$$

Таким образом,

$$\varepsilon\cdot\min_{k\in I} v_f(x^k, x_*)\cdot|I| \leqslant \frac{\varepsilon^2}{2}\cdot|I| + \Theta_0^2 - \sum_{k\in J}\frac{\varepsilon^2}{2||\nabla g_{m(k)}(x^k)||_*^2} \leqslant \varepsilon^2\cdot|I|,$$

откуда

$$\min_{k\in I} v_f(x^k, x_*) \leqslant \varepsilon. \quad (3.9)$$

В завершение покажем, что $|I| \neq 0$. Предположим обратное: $|I| = 0 \Rightarrow |J| = N$, т. е. все шаги непродуктивны. Тогда с учетом (3.8) получаем, что для всякого $k \in \{\overline{0, N-1}\}$

$$h_k(g_{m(k)}(x^k) - g_{m(k)}(x_*)) > \frac{\varepsilon^2}{\|\nabla g_{m(k)}(x^k)\|_*^2}$$

и

$$\sum_{k=0}^{N-1} h_k(g_{m(k)}(x^k) - g_{m(k)}(x_*)) \leqslant \sum_{k=0}^{N-1}\frac{\varepsilon^2}{2\|\nabla g_{m(k)}(x^k)\|_*^2} + \Theta_0^2 \leqslant \sum_{k=0}^{N-1}\frac{\varepsilon^2}{\|\nabla g_{m(k)}(x^k)\|_*^2}.$$

Итак, получили противоречие. Это означает, что $|I| \neq 0$.

Отметим, что ввиду условия Липшица для функционального ограничения на любой итерации работы алгоритма 4 справедливо неравенство $||\nabla g_{m(k)}(x^k)||_* \leqslant M_g$. Поэтому критерий остановки алгоритма 4 будет заведомо выполнен не более чем после (3.5) итераций работы. Теорема доказана.

Теперь покажем, как можно оценить скорость сходимости предлагаемого метода. Напомним, что x_* — решение задачи (1.2), (1.3).

Введем следующую функцию (см. [11, лемма 3.2.1]):

$$\omega(\tau) = \max_{x \in X}\{f(x) - f(x_*) : ||x - x_*|| \leqslant \tau\}, \tag{3.10}$$

где τ — положительное число. Тогда для всякого $y \in X$

$$f(y) - f(x_*) \leqslant \omega(v_f(y, x_*)). \tag{3.11}$$

На базе (3.10), (3.11) и теоремы 2 можно оценить скорость сходимости алгоритма 4 для дифференцируемого целевого функционала f с градиентом, удовлетворяющим условию Липшица

$$||\nabla f(x) - \nabla f(y)||_* \leqslant L||x - y|| \quad \forall\, x, y \in X. \tag{3.12}$$

Используя известное неравенство (см., например, [11])

$$f(x) \leqslant f(x_*) + ||\nabla f(x_*)||_*||x - x_*|| + \frac{1}{2}L||x - x_*||^2,$$

мы можем получить, что

$$\min_{k \in I} f(x^k) - f(x_*) \leqslant \min_{k \in I}\Big\{||\nabla f(x_*)||_*||x^k - x_*|| + \frac{1}{2}L||x^k - x_*||^2\Big\}.$$

Далее, верны оценки $f(x) - f(x_*) \leqslant \varepsilon \cdot ||\nabla f(x_*)||_* + \dfrac{1}{2}L\varepsilon^2$.

Поэтому справедливо

Следствие 1. *Пусть f дифференцируема на X и верно* (3.12). *Тогда после остановки алгоритма* 4 *верна оценка*

$$\min_{1 \leqslant k \leqslant N} f(x^k) - f(x_*) \leqslant \varepsilon_f + \frac{L\varepsilon^2}{2},$$

где $\varepsilon_f = \varepsilon \cdot ||\nabla f(x_*)||_*$.

Также можно рассмотреть специальный класс негладких целевых функционалов.

Следствие 2. *Пусть $f(x) = \max\limits_{i=\overline{1,m}} f_i(x)$, где f_i дифференцируема для всякого $x \in X$ и*

$$||\nabla f_i(x) - \nabla f_i(y)||_* \leqslant L_i||x - y|| \quad \forall\, x, y \in X.$$

Тогда после остановки алгоритма 4 *верна оценка*

$$\min_{0 \leqslant k \leqslant N} f(x^k) - f(x_*) \leqslant \varepsilon_f + \frac{L\varepsilon^2}{2},$$

где $\varepsilon_f = \varepsilon \cdot ||\nabla f(x_*)||_*$, $L = \max\limits_{i=\overline{1,m}} L_i$.

З а м е ч а н и е 3. Ввиду липшицевости и, вообще говоря, негладкости функциональных ограничений оценка (3.5) на число итераций означает, что предложенный метод оптимален с точки зрения оракульных оценок [9]: для достижения требуемой точности ε решения задачи (1.2), (1.3) для рассмотренного в данном разделе статьи класса целевых функционалов достаточно $O\Big(\dfrac{1}{\varepsilon^2}\Big)$ итераций алгоритма 4. Отметим также, что к рассмотренным классам задач (1.2), (1.3) с целевыми функционалами вида (3.1), (3.2) применимы и рассмотренные ранее алгоритмы 1 и 3. Однако невыполнение, вообще говоря, условия Липшица для f не позволяет обосновать оптимальность алгоритмов 1 и 3 для целевых функционалов вида (3.1), (3.2). Точнее говоря, возможны ситуации, когда на продуктивных шагах нормы (суб)градиентов целевого функционала $\|\nabla f(x^k)\|_*$ будут достаточно большими и это будет мешать быстрому достижению критерия остановки алгоритмов 1 и 3. В таком случае алгоритмы 2 и 4 могут работать быстрее, что показано ниже на некоторых конкретных примерах в численных экспериментах.

4. Численные эксперименты

Для демонстрации преимуществ алгоритмов 3 и 4 по сравнению с алгоритмами 1 и 2 соответственно был проведен ряд тестов. Рассматриваемые целевые функционалы $f(x)$ и ограничения $g_m(x)$ $(m = 1, 2, 3, \ldots, 10)$ указаны в табл. 1 (более простые целевые функционалы) и 3 (негладкие целевые функционалы max-типа). Отметим, что мы рассматриваем функционалы 10 переменных, стандартную евклидову прокс-структуру, начальную точку $x^0 = (1, 1, 1, \ldots, 1)$ и $\Theta_0 = 3$ (можно проверить, что для всех примеров одной из оптимальных точек будет $x_* = (0, 0, 0, \ldots, 0)$), точность $\varepsilon = 0.05$.

Результаты выполнения алгоритмов представлены в сравнительных табл. 2 и 4. Приводится количество итераций и время (указано в секундах) работы алгоритмов 1 и 2, а также соответствующих модификаций. Все вычисления были произведены с помощью с помощью программного обеспечения CPython 3.6.4 [13] на компьютере с 3-ядерным процессором AMD Athlon II X3 450 с тактовой частотой 803,5 МГц на каждое ядро. ОЗУ компьютера составляла 8 Гб.

Как видим, предлагаемые нами модификации (алгоритм 3 — модификация алгоритма 1 и алгоритм 4 — модификация алгоритма 2) могут существенно сокращать как необходимое для достижения нужного качества решения количество итераций, так и время работы алгоритмов для задач с указанными ограничениями. Особенно наглядно это видно в примерах 1, 2 и 4 для алгоритмов 1 и 3. Также заметные преимущества дает алгоритм 4 перед алгоритмом 2 в примерах 3, 5 и 6.

Таблица 1

Входные данные

	Алгоритмы 1–4
	$f(x)$
Пример 1	$\sqrt{0.1\left(\sum_{i=1}^{10} x_i^2 + \sum_{i=1}^{9} x_i x_{i+1}\right)}$
Пример 2	$\sum_{i=1}^{10} x_i^2 - x_1x_2 + x_3 - x_8 + x_9x_{10}$
Пример 3	$\sum_{i=1}^{10} 5^i x_i^2$
	$g_m(x),\ m = \overline{1,10}$
	$x_1 + 20x_2 + 30x_3 + 40x_4 + 50x_5 + 60x_6 + 70x_7 + 80x_8 + 90x_9 + 100x_{10}$, $x_1 + 120x_2 + 130x_3 + 140x_4 + 150x_5 + 160x_6 + 170x_7 + 180x_8 + 190x_9 + 200x_{10}$, $x_1 + 220x_2 + 230x_3 + 240x_4 + 250x_5 + 260x_6 + 270x_7 + 280x_8 + 290x_9 + 300x_{10}$, $x_1 + 320x_2 + 330x_3 + 340x_4 + 350x_5 + 360x_6 + 370x_7 + 380x_8 + 390x_9 + 400x_{10}$, $x_1 + 420x_2 + 430x_3 + 440x_4 + 450x_5 + 460x_6 + 470x_7 + 480x_8 + 490x_9 + 500x_{10}$, $x_1 + 520x_2 + 530x_3 + 540x_4 + 550x_5 + 560x_6 + 570x_7 + 580x_8 + 590x_9 + 600x_{10}$, $x_1 + 620x_2 + 630x_3 + 640x_4 + 650x_5 + 660x_6 + 670x_7 + 680x_8 + 690x_9 + 700x_{10}$, $x_1 + 720x_2 + 730x_3 + 740x_4 + 750x_5 + 760x_6 + 770x_7 + 780x_8 + 790x_9 + 800x_{10}$, $x_1 + 820x_2 + 830x_3 + 840x_4 + 850x_5 + 860x_6 + 870x_7 + 880x_8 + 890x_9 + 900x_{10}$, $x_1 + 920x_2 + 930x_3 + 940x_4 + 950x_5 + 960x_6 + 970x_7 + 980x_8 + 990x_9 + 1\,000x_{10}$

Таблица 2

Сравнение результатов работы алгоритмов

	Итерации	Время	Итерации	Время
	Алгоритм 1		Алгоритм 3	
Пример 1	730 829	133	261 800	40
Пример 2	1 638 946	262	453 580	30
Пример 3	$>10^7$	>500	$>10^7$	>500
	Алгоритм 2		Алгоритм 4	
Пример 1	–	–	–	–
Пример 2	1 584 616	300	1 434 006	156
Пример 3	184 706	124	89 940	110

Таблица 3

Входные данные

	Алгоритмы 1–4
	$f(x)$
Пример 4	$\max\{0.1\lvert x_1+x_2+x_3\rvert+1,\ 0.01\lvert x_4+2x_5+x_6\rvert+2,\ 0.001\lvert x_7+3x_8+4x_9+10x_{10}\rvert+5\}$
Пример 5	$\max\{x_1^2,\ 10x_2^2,\ 50x_3^2,\ 100x_4^2,\ 200x_5^2,\ 400x_6^2,\ 800x_7^2,\ 1\,000x_8^2,\ 5\,000x_9^2,\ 10\,000x_{10}^2\}$
Пример 6	$\max\{x_1+2x_2+3x_3,\ x_3+4x_4+6x_5,\ x_4+3x_5+6x_6+7x_7,\ 5x_7+8x_8+9x_9,\ x_1+10x_{10}\}$
	$g_m(x),\ m=\overline{1,10}$
	$x_1+20x_2+30x_3+40x_4+50x_5+60x_6+70x_7+80x_8+90x_9+100x_{10},$ $x_1+120x_2+130x_3+140x_4+150x_5+160x_6+170x_7+180x_8+190x_9+200x_{10},$ $x_1+220x_2+230x_3+240x_4+250x_5+260x_6+270x_7+280x_8+290x_9+300x_{10},$ $x_1+320x_2+330x_3+340x_4+350x_5+360x_6+370x_7+380x_8+390x_9+400x_{10},$ $x_1+420x_2+430x_3+440x_4+450x_5+460x_6+470x_7+480x_8+490x_9+500x_{10},$ $x_1+520x_2+530x_3+540x_4+550x_5+560x_6+570x_7+580x_8+590x_9+600x_{10},$ $x_1+620x_2+630x_3+640x_4+650x_5+660x_6+670x_7+680x_8+690x_9+700x_{10},$ $x_1+720x_2+730x_3+740x_4+750x_5+760x_6+770x_7+780x_8+790x_9+800x_{10},$ $x_1+820x_2+830x_3+840x_4+850x_5+860x_6+870x_7+880x_8+890x_9+900x_{10},$ $x_1+920x_2+930x_3+940x_4+950x_5+960x_6+970x_7+980x_8+990x_9+1\,000x_{10}$

Таблица 4

Сравнение результатов работы алгоритмов

	Итерации	Время	Итерации	Время
	Алгоритм 1		Алгоритм 3	
Пример 4	172 821	24	17 255	1
Пример 5	$>10^6$	>500	$>10^6$	>500
Пример 6	$>10^6$	>500	$>10^6$	>500
	Алгоритм 2		Алгоритм 4	
Пример 4	–	–	–	–
Пример 5	182 993	106	66 095	79
Пример 6	180 020	101	24 454	78

Это связано с тем, что адаптивные критерии остановки алгоритмов 1–4, как легко видеть, определяют зависимость остановки алгоритмов от величины норм (суб)градиентов ограничений на непродуктивных шагах. Чем меньше эти нормы, тем скорее будут выполнены критерии остановки. Точнее говоря, замена (суб)градиента max-ограничения $g(\cdot)=\max_{m\in\overline{1,M}} g_m(\cdot)$ (если применять алгоритм 1 или 2) на какой-либо (суб)градиент $g_m(\cdot)$ (если применять алгоритм 3 или 4) может ускорить выполнение критерия остановки метода.

Показано также, что в случае целевых функционалов f вида (3.1), (3.2) ввиду того, что величины $\|\nabla f(x^k)\|_*$ могут быть достаточно большими, алгоритмы 2 и 4 могут работать существенно быстрее алгоритмов 1 и 3. Это показывают примеры 3, 5 и 6. Однако, если нормы $\|\nabla f(x^k)\|_*$ не особо велики, то алгоритмы 1 и 3 могут в конкретных примерах работать быстрее алгоритмов 2 и 4 и для квадратичных функционалов (см. пример 2).

5. Заключительные замечания

В заключение отметим, что основные результаты работы (сходимость и оптимальность предложенных методов с точки зрения нижних оракульных оценок) не зависят от выбора на непродуктивном шаге ограничения $g_m(\cdot)$, на котором $g_m(x^k) > \varepsilon$.

Однако замена (суб)градиента max-ограничения $g(\cdot) = \max\limits_{m \in \overline{1,M}} g_m(\cdot)$ (если применять алгоритм 1 или 2) на какой-либо (суб)градиент $g_m(\cdot)$ (если применять алгоритм 3 или 4) может привести к увеличению нормы рассматриваемого на непродуктивном шаге (суб)градиента. Чтобы этого избежать, нужно на каждом непродуктивном шаге k среди ограничений $g_m(\cdot)$, для которых $g_m(x^k) > \varepsilon$, находить ограничение с наименьшей нормой (суб)градиента $\|\nabla g_m(x^k)\|$. Ясно, что операция минимизации $\|\nabla g_m(x^k)\|$ по всем подходящим ограничениям на каждой итерации может привести к увеличению времени работы алгоритма, и в результате будут не очевидны преимущества перед исходными алгоритмами 1 и 2.

Однако отметим, что предлагаемые нами модификации алгоритмов 1 и 2 выгодны в случае, если возможно априорно независимо от выбора точки упорядочить по неубыванию нормы (суб)градиентов функциональных ограничений. Это легко сделать, например, если задача имеет аффинные функциональные ограничения. Ясно, что существуют и другие примеры подходящих задач с ограничениями. Если такая сортировка возможна, то в алгоритмах 1 и 2 на непродуктивных шагах достаточно использовать первое ограничение $g_m(\cdot)$, для которого $g_m(x^k) > \varepsilon$. Представляется, что дальнейшая разработка предложенных методов для условных задач с различной структурой — интересная задача на будущее.

Авторы выражают огромную признательность Александру Владимировичу Гасникову и Павлу Евгеньевичу Двуреченскому, а также рецензенту за полезные обсуждения и пожелания.

СПИСОК ЛИТЕРАТУРЫ

1. **Ben-Tal A. and Nemirovski A.** Robust truss topology design via semidefinite programming // SIAM J. Optim. 1997. Vol. 7, no. 4, pp. 991–1016. doi: 10.1137/S1052623495291951 .
2. **Shpirko S., Nesterov Yu.** Primal-dual subgradient methods for huge-scale linear conic problem // SIAM J. Optim. 2014. Vol. 24, no. 3, pp. 1444–1457. doi: 10.1137/130929345 .
3. **Beck A., Teboulle M.** Mirror descent and nonlinear projected subgradient methods for convex optimization // Operations Research Letters. 2003. Vol. 31, no. 3, pp. 167–175. doi: 10.1016/S0167-6377(02)00231-6 .
4. **Nemirovsky A., Yudin D.** Problem complexity and method efficiency. N Y: Wiley & Sons, 1983. 404 p. ISBN: 978-0471103455 .
5. **Поляк Б. Т.** Один общий метод решения экстремальных задач // Докл. АН СССР. 1967. Т. 8, № 3. С. 593–597.
6. **Шор Н. З.** Применение обобщенного градиентного спуска в блочном программировании // Кибернетика. 1967. № 3. С. 53–55.
7. **Немировский А. С., Юдин Д. Б.** Эффективные методы решения задач выпуклого программирования большой размерности // Экономика и математические методы. 1979. № 2. С. 135–152.
8. The comirror algorithm for solving nonsmooth constrained convex problems / A. Beck, A. Ben-Tal, N. Guttmann-Beck, L. Tetruashvili // Operations Research Letters. 2010. Vol. 38, no. 6. P. 493–498. doi: 10.1016/j.orl.2010.08.005 .

9. **Ben-Tal A., Nemirovski A.** Lectures on modern convex optimization. Philadelphia: Society for Industrial and Appl. Math., 2001. 590 p. ISBN: 0-89871-491-5 .
10. Mirror descent and convex optimization problems with non-smooth inequality constraints [e-resource] / A. Bayandina, P. Dvurechensky, A. Gasnikov, F. Stonyakin, A. Titov. 2018. 30 p. URL: arxiv.org/abs/1710.06612 .
11. **Nesterov Y.** Introductory lectures on convex optimization: a basic course. Boston: Kluwer Acad. Publ., 2004. 236 p. doi: 10.1007/978-1-4419-8853-9 .
12. **Nesterov Y.** Subgradient methods for convex functions with nonstandard growth properties [e-resource, slides]. URL: www.mathnet.ru:8080/PresentFiles/16179/growthbm_nesterov.pdf [Online; accessed 30-March-2018].
13. CPython [site]. URL: https://www.python.org/ .

Поступила 30.03.2018

Стонякин Федор Сергеевич
канд. физ.-мат. наук, доцент
Крымский федеральный университет им. В. И. Вернадского,
г. Симферополь
e-mail: fedyor@mail.ru

Mohammad S. Alkousa
аспирант
Московский физико-технический институт (государственный университет),
г. Москва
e-mail: mohammad.alkousa@phystech.edu

Степанов Алексей Николаевич
соисполнитель работ по гранту Президента РФ
для молодых кандидатов наук, код МК-176.2017.1
Крымский федеральный университет им. В. И. Вернадского,
г. Симферополь
e-mail: stepanov.student@gmail.com

Баринов Максим Александрович
студент
Крымский федеральный университет им. В. И. Вернадского
г. Симферополь
e-mail: ice8scream@gmail.com

REFERENCES

1. Ben-Tal A., Nemirovski A. Robust truss topology design via semidefinite programming. *SIAM J. Optim.*, 1997, vol. 7, no. 4, pp. 991–1016. doi: 10.1137/S1052623495291951 .
2. Shpirko S., Nesterov Yu. Primal-dual subgradient methods for huge-scale linear conic problem. *SIAM J. Optim.*, 2014, vol. 24, no. 3, pp. 1444–1457. doi: 10.1137/130929345 .
3. Beck A., Teboulle M. Mirror descent and nonlinear projected subgradient methods for convex optimization. *Operations Research Letters*, 2003, vol. 31, no. 3, pp. 167–175. doi: 10.1016/S0167-6377(02)00231-6 .
4. Nemirovsky A., Yudin D. *Problem complexity and method efficiency in optimization.* N Y: J. Wiley & Sons, 1983, 404 p. ISBN: 978-0471103455 .
5. Polyak B. A general method of solving extremum problems. *Soviet Math. Dokl.*, 1967, vol. 8, no. 3, pp. 593–597 (in Russian).
6. Shor N.Z. Generalized gradient descent with application to block programming. *Cybernetics*, 1967, vol. 3, no. 3, pp. 43–45 (in Russian).
7. Nemirovskii A. Efficient methods for solving convex programming problems of high dimension. *Ekonomika i Matematicheskie Metody*, 1979, vol. 15, no. 1 (in Russian).

8. Beck A., Ben-Tal A., Guttmann-Beck N., Tetruashvili L. The comirror algorithm for solving nonsmooth constrained convex problems. *Operations Research Letters*, 2010, vol. 38, no. 6, pp. 493–498. doi: 10.1016/j.orl.2010.08.005 .
9. Ben-Tal A., Nemirovski A. *Lectures on Modern Convex Optimization.* Philadelphia: Society for Industrial and Appl. Math., 2001, 488 p. ISBN: 0-89871-491-5 .
10. Bayandina A., Dvurechensky P., Gasnikov A., Stonyakin F., Titov A. Mirror descent and convex optimization problems with non-smooth inequality constraints [e-resource]. 2018. 30 p. At available: https://arxiv.org/abs/1710.06612 .
11. Nesterov Y. *Introductory lectures on convex optimization: a basic course.* Boston: Kluwer Acad. Publ., 2004, 236 p. doi: 10.1007/978-1-4419-8853-9 .
12. Nesterov Y. *Subgradient methods for convex functions with nonstandard growth properties* [e-resource, slides]. At available: https://www.mathnet.ru:8080/PresentFiles/16179/growthbm_nesterov.pdf [Online; accessed 15-April-2018] .
13. CPython [site]. At available: https://www.python.org/ .

The paper was received by the Editorial Office on March 30, 2018

Fedor Sergeevich Stonyakin, Cand. Sci. (Phys.-Math.), V.I. Vernadsky Crimean Federal University, Simferopol, Republic of Crimea, 295007 Russia, e-mail: fedyor@mail.ru .

Mohammad S. Alkousa, doctoral student, Moscow Institute of Physics and Technologies, Moscow, 141701 Russia, e-mail: mohammad.alkousa@phystech.edu .

Aleksei Nikolaevich Stepanov, co-executor of research work by grant the President of Russian Federation for young candidates of sciences, project no. MK-176.2017.1, V.I. Vernadsky Crimean Federal University, Simferopol, Republic of Crimea, 295007 Russia,
e-mail: stepanov.student@gmail.com .

Maksim Aleksandrovich Barinov, undergraduate student, V.I. Vernadsky Crimean Federal University, Simferopol, Republic of Crimea, 295007 Russia, e-mail: ice8scream@gmail.com .