

Math-Net.Ru

Общероссийский математический портал

А. Ю. Попков, Б. С. Дарховский, Ю. С. Попков, Итерационный МК-алгоритм решения задач глобальной оптимизации,
Автомат. и телемех., 2017, выпуск 2, 82–98

<https://www.mathnet.ru/at14685>

Использование Общероссийского математического портала Math-Net.Ru подразумевает, что вы прочитали и согласны с пользовательским соглашением
<https://www.mathnet.ru/rus/agreement>

Параметры загрузки:

IP: 18.97.14.86

21 апреля 2025 г., 11:19:25



© 2017 г. А.Ю. ПОПКОВ, канд. техн. наук (apopkov@isa.ru),
(Институт системного анализа РАН, Москва,
Московский физико-технический институт),
Б.С. ДАРХОВСКИЙ, д-р физ.-мат. наук (darbor2004@mail.ru),
Ю.С. ПОПКОВ, д-р техн. наук (popkov@isa.ru)
(Институт системного анализа РАН,
Московский физико-технический институт,
Национальный исследовательский университет “Высшая школа экономики”,
Москва)

ИТЕРАЦИОННЫЙ МК-АЛГОРИТМ РЕШЕНИЯ ЗАДАЧ ГЛОБАЛЬНОЙ ОПТИМИЗАЦИИ¹

Предлагается новый метод решения задач глобальной минимизации гельдеровских функций на компактных множествах, описываемых непрерывными функциями. Метод основан на пакетных итерациях Монте-Карло, предназначенных для построения последовательностей значений “квази-глобальных” минимумов и их декрементов. Предложена количественная процедура формирования вероятностного правила остановки. Работоспособность метода подтверждена на многочисленных тестах и задаче с алгоритмически заданными функциями.

Ключевые слова: глобальная оптимизация, пакетные итерации Монте-Карло, константы Гельдера.

1. Введение

Проблема поиска глобального экстремума на компактном множестве традиционно привлекает интерес исследователей, обусловленный ростом количества прикладных задач, где требуется отыскивать глобальный экстремум функций на компактных множествах, информация о свойствах которых доступна лишь в алгоритмической форме.

Теории глобальной оптимизации посвящено огромное количество публикаций, ориентированных на определенные классы задач. Не претендуя на полноту классификации, можно выделить следующие четыре направления исследований.

Первым нужно упомянуть направление, в котором используются идеи ветвления (ограниченного перебора) и оценивания границ (верхней и нижней) оптимизируемой функции. Существует обширный круг задач, где удается построить указанные оценки [1, 3–6].

Второе направление базируется на конфигурации целевой функции и допустимого множества. Представителем этого направления является так называемая *ДС*-минимизация, когда целевая функция и функции, описывающие допустимое множество, строятся в виде разностей двух выпуклых функций. Полное представление о достижениях этого направления можно составить по [6, 7].

В рамках третьего направления изучаются задачи глобальной оптимизации на простых допустимых множествах (например, параллелепипедах) для

¹ Работа выполнена при поддержке Российского фонда фундаментальных исследований (проект № 15-07-03005).

целевых функций с известной константой Липшица. Прекрасный обзор результатов в этом направлении представлен в [8, 9]. Среди публикаций этого направления следует выделить [10, 11], посвященные редукции n -мерной задачи к одномерной с помощью кривых Пеано.

И наконец, четвертое направление базируется на случайном поиске (методе Монте-Карло) и его разнообразных интеллектуализированных эвристиках [12–15]. Одна из существенных проблем, возникающих при применении испытаний Монте-Карло — генерация равномерно распределенных случайных векторов в заданной области пространства поиска. По-видимому, базовыми в этой проблеме являются метод Hit-and-Run и его многочисленные модификации [16, 17], методы, использующие марковские цепи [18, 19] и энтропию Кульбака–Ляйблера [20, 21]. В [22] был предложен метод решения систем нелинейных уравнений и неравенств, основанный на простых испытаниях Монте-Карло. Исследованы вероятностные характеристики метода и доказана его сходимость с вероятностью единица и экспоненциальной скоростью.

В данной статье указанный метод развит для решения задач глобальной минимизации гельдеровских функций на компактных множествах, описываемых непрерывными функциями. Предложен новый алгоритм глобального поиска (GFS-алгоритм), основанный на итерационных процедурах с пакетами простых испытаний Монте-Карло и с последующей их фильтрацией и селекцией. Новизна и эффективность алгоритма состоит в том, что он:

- ориентирован на общие задачи глобальной оптимизации с *алгоритмически вычисляемыми функциями*; предполагается, что целевая функция гельдеровского типа, но с *неизвестными* константами;
- использует итерационные процедуры с пакетами *простых* испытаний Монте-Карло (т.е. генерацией независимых случайных векторов, равномерно распределенных в *единичном кубе*);
- осуществляет фильтрацию случайных векторов по принадлежности к *допустимому* множеству для формирования последовательности “квази-глобальных” минимумов;
- имеет оригинальное *вероятностное* правило остановки, основанное на вероятностных характеристиках (при конечном числе итераций) последовательности “квази-глобальных” минимумов.

Работоспособность и эффективность GFS-алгоритма подтверждена на тестовых задачах с аналитически заданными и алгоритмически вычисляемыми целевой функцией и функциями, описывающими допустимое множество.

2. Каноническая форма задачи глобальной оптимизации

Рассматривается задача глобальной минимизации в *канонической форме*, а именно:

$$(1) \quad F(\mathbf{z}) \Rightarrow \text{globmin}$$

при ограничениях:

$$\mathbf{z} \in \mathcal{K}_{\mathbf{z}} = \mathcal{G}_{\mathbf{z}} \cap Z_{+}^d,$$

$$Z_{+}^d = \{\mathbf{z} : \mathbf{0} \leq \mathbf{z} \leq \mathbf{1}\} \subset \mathbb{R}^d,$$

$$\mathcal{G}_{\mathbf{z}} = \{\mathbf{z} : \mathbf{w}(\mathbf{z}) \leq \mathbf{0}\}, \quad \mathbf{w} \in \mathbb{R}^r.$$

К такой форме всегда можно привести (путем замены переменных) общую задачу, если допустимое множество является компактом, задаваемым системой неравенств при помощи непрерывных функций.

Предполагается, что функции $\mathbf{w}(\mathbf{z})$ непрерывны, а функция $F(\mathbf{z})$ удовлетворяет условию Гельдера, но с неизвестными константами L и s . Для модуля непрерывности $\omega(h)$ функции $F(\mathbf{z})$ имеет место оценка:

$$(2) \quad \omega(h) = \max_{(\mathbf{v}, \mathbf{y}) \in \mathcal{K}_{\mathbf{z}} : \|\mathbf{v} - \mathbf{y}\| \leq h} |F(\mathbf{v}) - F(\mathbf{y})| \leq Lh^s.$$

Обозначим через \mathbb{Z}^* множество точек глобального минимума, а через $F(\mathbf{z}^*) = F^*$, $\mathbf{z}^* \in \mathbb{Z}^*$ — значение глобального минимума в задаче (1).

Функция $F(\mathbf{z})$ в силу непрерывности, очевидно, ограничена снизу на $\mathcal{K}_{\mathbf{z}}$, т.е. существует константа c , для которой

$$(3) \quad F^* \geq c.$$

Константа c в общем случае неизвестна, но в некоторых специальных задачах глобальной минимизации, где глобальный минимум существует, она равна нулю (при этом вопрос о его достижимости остается открытым). Это так, например, в задачах, где функция $F(\mathbf{z})$ имеет смысл невязки между левой и правой частями некоторой системы уравнений.

Цель статьи — построение сходящегося с вероятностью единица алгоритма (будем называть его *GFS-алгоритм* “generation, filtration, selection”) поиска приближенного значения глобального минимума функции $F(\mathbf{z})$ на множестве $\mathcal{K}_{\mathbf{z}}$ и соответствующего ему аргумента (точки) $\mathbf{z}^* \in \mathbb{Z}^*$ с использованием простых испытаний Монте-Карло. Приближенное значение глобального минимума сопровождается оценкой окрестности его точного значения и вероятностью попадания приближенного значения в эту окрестность.

3. Идея алгоритма GFS

На k -м шаге алгоритма ($k = 0, 1, 2, \dots$) генерируется пакет \mathcal{Z}_k из N_k независимых и равномерно распределенных на Z_+^d случайных векторов $\{\mathbf{z}^1, \dots, \mathbf{z}^{N_k}\}$ (*генерация* (G)).

Пакет формируется d датчиками, генерирующими параллельно по каждой координате вектора \mathbf{z} M_k независимых и равномерно распределенных на $[0, 1]$ случайных чисел. Из $(d \times M_k)$ -массива образуются $N_k = M_k$ случайных векторов, равномерно распределенных в единичном кубе Z_+^d .

Проверяется принадлежность (*фильтрация* (F)) векторов из пакета допустимому множеству $\mathcal{K}_{\mathbf{z}}$ и формируется пакет допустимых векторов $\tilde{\mathcal{Z}}_k$. Процедура фильтрации и формирования пакета $\tilde{\mathcal{Z}}_k$ описана в Приложении. Для векторов из этого пакета вычисляются значения функции $F(\mathbf{z}^j)$, $j = \overline{1, N_k}$, и производится их *селекция* (S), т.е. находится минимальное (“квази-глобальное”) значение F_k^* и соответствующий ему аргумент \mathbf{z}_k^* . В силу случайного характера простых МК-испытаний, “квази-глобальное” значение целевой функции F_k^* — случайное.

Переход к $(k + 1)$ -му шагу осуществляется путем формирования пакета \mathcal{Z}_{k+1} случайных векторов N_{k+1} по правилу:

$$(4) \quad N_{k+1} = \alpha N_k, \quad \alpha > 1.$$

При этом на шаге $(k + 1)$ выполняются указанные выше операции и находится новый “квази-глобальный” минимум F_{k+1}^* для пакета $\tilde{\mathcal{Z}}_{k+1}$.

Наряду с последовательностью “квази-глобальных” минимумов $\mathcal{F}^* = \{F_0^*, \dots, F_k^*, \dots\}$, формируется последовательность ее декрементов:

$$(5) \quad \mathcal{U} = \{u_1, \dots, u_k, \dots\},$$

где $u_k = |F_{(k+1)}^* - F_k^*|$, и последовательность логарифмов декрементов:

$$(6) \quad \mathcal{V} = \{v_1, \dots, v_k, \dots\},$$

где $v_k = \log u_k$.

В [22] на основе теоретического анализа предложена гипотеза о том, что для функций гельдеровского класса (2) при достаточно больших k справедлива следующая линейная регрессия:

$$(7) \quad v_k = A - B \log N_k + \zeta_k,$$

где коэффициенты A и B связаны с константами Гельдера, а $\{\zeta_k\}$ — последовательность случайных величин с нулевым математическим ожиданием и конечной дисперсией σ^2 . Гипотеза (7) подтверждена на практике. Последовательности \mathcal{V} (6) и \mathcal{N} (4) используются для вычисления оценок \hat{A}_k, \hat{B}_k , получаемых с помощью метода наименьших квадратов (МНК).

Алгоритм останавливается, если последовательность декрементов u_k в течение ρ шагов остается в заданной δ -окрестности нуля. Количество шагов до остановки равно

$$(8) \quad \tau = \min \left\{ k : \max_{k-\rho \leq w \leq k} u_w \leq \delta \right\}.$$

На момент остановки алгоритма по оценкам $\hat{A}_\tau, \hat{B}_\tau$ вычисляются оценки размера окрестности точного значения глобального минимума (см. [22], Приложение):

$$(9) \quad |F_\tau^* - F^*| \leq \hat{r}_\tau = a^{\hat{A}_\tau} \left(\frac{\sqrt{d}}{2} \right)^{\sqrt{2\hat{B}_\tau d}} N_\tau^{-\hat{B}_\tau},$$

где a — основание используемого логарифма, и нижняя оценка вероятности попадания найденного решения в эту окрестность:

$$(10) \quad \hat{P}_\tau = P \left\{ 0 \leq (F_\tau^* - F^*) \leq \hat{D}_\tau N^{-\hat{B}_\tau} \right\} \geq 1 - N_\tau^{\hat{s}_\tau/2} \exp \left(-N_\tau^{1-\hat{s}_\tau/2} \right),$$

где $\hat{D}_\tau = \log \hat{A}_\tau, \hat{s}_\tau = \sqrt{2\hat{B}_\tau d}$.

Таким образом, под *решением* задачи (1) понимается *набор*

$$(11) \quad S = \left\{ F_{\tau}^*, \mathbf{z}_{\tau}^*, \hat{r}_{\tau}, \hat{P}_{\tau} \right\}.$$

Вероятностные характеристики GFS-алгоритма и отдельных его процедур исследованы в [22].

4. Логическая структура алгоритма GFS

Алгоритм GFS ориентирован на поиск глобального экстремума для канонической формы задачи (1). Он состоит из двух процедур SeqLocMin и EstProb. Процедура SeqLocMin предназначена для формирования последовательности “квази-глобальных” и их декрементов и включает правило останова. Процедура EstProb предназначена для вычисления нижней оценки вероятности отклонения найденного значения “квази-глобального” минимума от точного глобального минимума.

Процедуры состоят из последовательности k шагов, внутри которых выполняется некоторое количество t операций. Поэтому этапы процедуры индексированы парой (k, t) .

Процедура SeqLocMin

Вычислительная схема k -го шага данной процедуры состоит в следующем:

1. *Генерация* — создание пакета случайных векторов;

$(k, 1)$ Генерируется последовательность из N_k случайных, независимых и равномерно распределенных в единичном кубе векторов;

$(k, 2)$ Формируется пакет случайных векторов $\mathcal{Z}_k = \{\mathbf{z}^1, \dots, \mathbf{z}^{N_k}\}$,

2. *Фильтрация* — отбор допустимых векторов;

$(k, 3)$ Из пакета \mathcal{Z}_k выбираются векторы, принадлежащие допустимому множеству $\mathcal{K}_{\mathbf{z}}$, и формируется допустимый пакет $\tilde{\mathcal{Z}}_k$ с объемом \tilde{N}_k (см. Приложение);

3. *Селекция* — определение “квази-глобального” минимума для пакета и вычисление декрементов;

$(k, 4)$ Вычисляются на допустимом пакете $\tilde{\mathcal{Z}}_k$ значения целевой функции $F(\tilde{\mathbf{z}}^i)$, $i = 1, \tilde{N}_k$;

$(k, 5)$ Выбирается минимальное $F_k^* = \min F(\tilde{\mathbf{z}}^i)$, $1 \leq i \leq \tilde{N}_k$;

$(k, 6)$ Вычисляются декремент $u_k = |F_{k+1}^* - F_k^*|$ и логарифмы декремента $v_k = \log u_k$ и $n_k = \log N_k$.

Таким образом, в результате работы процедуры формируются последовательности \mathcal{F} , \mathcal{U} , \mathcal{V} , \mathcal{N} . В случае если $v_k \leq \log \delta$, значение v_k включается в последовательность $\tilde{\mathcal{V}}$. Если в последовательности $\tilde{\mathcal{V}}_k$ накапливается ρ элементов, то $k = \tau$, алгоритм останавливается и осуществляется переход к процедуре EstProb.

Процедура EstProb

Вычислительная схема данной процедуры состоит в следующем:

1. *Вычисление оценок параметров регрессионной модели декрементов*;

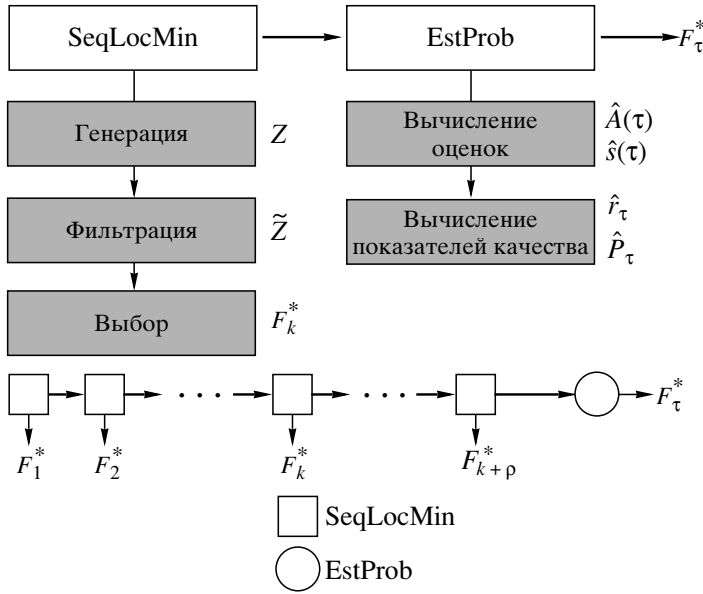


Рис. 1. Схема GFS-алгоритма.

- ($\tau, 1$) Используя последовательности \mathcal{N} и \mathcal{V} , вычисляются оценки параметров \hat{A}_τ , \hat{B}_τ методом наименьших квадратов;
2. *Вычисление оценок характеристик решения;*
- ($\tau, 2$) Вычисляется оценка размера окрестности глобального минимума \hat{r}_τ (9);
- ($\tau, 3$) Вычисляется нижняя оценка вероятности \hat{P}_τ (10).

На рис. 1 изображена логическая структурная схема GFS-алгоритма.

5. Тестирование алгоритма GFS

5.1. Аналитические тесты

Тестирование проводилось на некоторых задачах из [8, 23–25], в которых допустимая область представляла собой d -мерный параллелепипед $\Pi_{\mathbf{z}}$ со сторонами, указанными в Приложении.

Во всех расчетах $N_0 = 10$, $\alpha = 10$, $\delta = 0,01$, $a = 10$, $\rho = 2$, $\tau = 13$, $\varepsilon_\tau = |F^* - F_\tau^*|$. Результаты расчетов приведены в табл. 1, где t — время в секундах, затраченное на получение решения, а $\hat{r}_\tau, \hat{P}_\tau$ — оценки окрестности точного глобального минимума и его вероятности. Из табл. 1 следует, что приближенные решения во всех тестах принадлежат оценкам окрестности точного глобального минимума с вероятностью почти единица.

Принципиальной особенностью алгоритма является реализация в нем метода, использующего последовательность “квази-глобальных” минимумов. Результаты расчета по указанному методу иллюстрируют графики рис. 2, где показаны последовательности “квази-глобальных” минимумов F_k^* .

Таблица 1

Функция	Π_z	F_τ^*	z_τ^*	ε_τ	\hat{r}_τ	\hat{P}_τ	t, c
A1 $d = 2$	[0; 10]	-11,0225	4,0031 3,9983	0,0010	0,0265	$\approx 1,0$	1286
A1 $d = 4$	[0; 10]	-10,1242	4,0475 4,0160 4,0135 3,9590	0,4122	0,8810	$\approx 1,0$	2262
A2 $d = 2$	[-10; 10]	0,000016	0,9966 0,9930	0,000016	0,0221	$\approx 1,0$	2656
A2 $d = 5$	[-0,5; 1,5]	0,0288	0,9852 0,9843 0,9706 0,9357 0,8773	0,0288	3,3406	$\approx 1,0$	3186
A3 $d = 4$	[-0,5; 0,5]	0,000013	0,0405 -0,0043 0,0183 0,0181	0,000013	0,0479	$\approx 1,0$	1715
A4 $d = 2$	[0; 1]	1,0000	0,9000 0,9000	$2 \cdot 10^{-9}$	0,0000041	$\approx 1,0$	222
A4 $d = 5$	[0; 1]	1,0001	0,9002 0,8918 0,9007 0,8989 0,8923	0,0001	1,4970	$\approx 1,0$	599
A5 $d = 2$	[-5; 5]	0,0000	-0,0004 0,0000	$8 \cdot 10^{-8}$	0,00007	$\approx 1,0$	281
A5 $d = 5$	[-5; 5]	0,0018	-0,0294 -0,0140 -0,0206 -0,0036 -0,1094	0,0018	0,2533	$\approx 1,0$	584
A5 $d = 10$	[-5; 5]	0,1245	3,0566 -4,4010 -0,3254 -0,0586 0,2304 0,3659 0,1528 -0,9495 0,6693 0,1680	0,1245	0,1889	$\approx 1,0$	1594

Для проведения данного набора тестов алгоритм был реализован на языке C++ для использования на вычислительных системах с гибридной архитектурой с ускорителями NVIDIA. Расчеты проводились на вычислительной системе, состоящей из двух узлов, оснащенных двумя процессорами Intel Xeon X5650 и E5-2650, и двумя ускорителями Tesla M2050 и K20Xm.

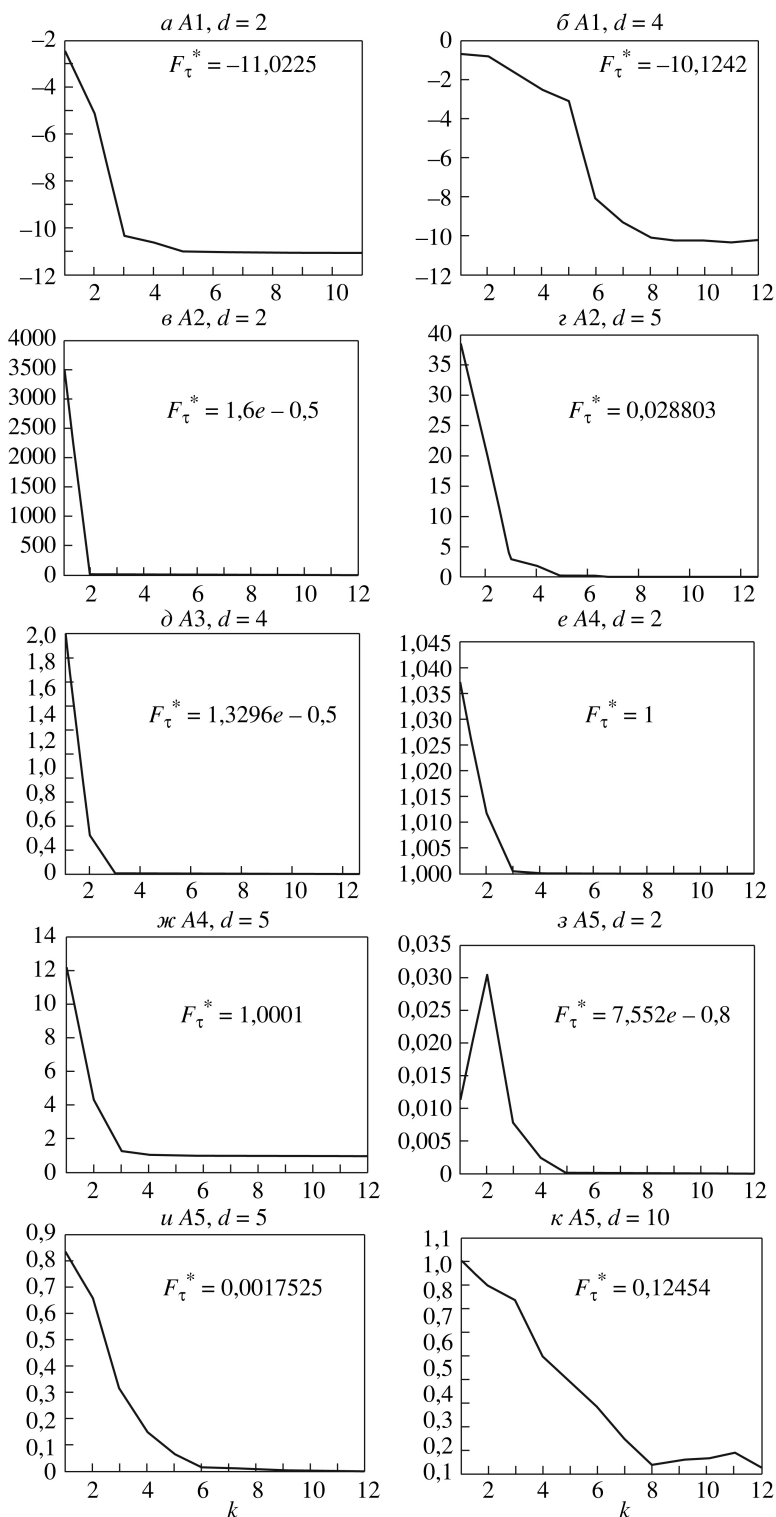


Рис. 2. F_τ^* .

5.2. Алгоритмический тест (энтропийное оценивание плотности распределения вероятностей параметров рандомизированных моделей)

Пусть имеется набор данных о входе и выходе некоторого динамического нелинейного объекта. Обозначим через $\mathbf{y} = \{y[m], \dots, y[m+s]\}$ — вектор (пакет) выходных данных и через $\mathbf{x} = \{x[0], \dots, x[m], \dots, x[m+s]\}$ — вектор (пакет) входных данных объекта. Предполагается, что вход измеряется точно, а выход — с ошибками, которые имитируются случайным вектором $\bar{\xi} = \{\xi[m], \dots, \xi[m+s]\}$ с независимыми компонентами. Компоненты этого вектора интервального типа, т.е.

$$(12) \quad \xi[j+m] \in \Xi_j = [\xi^-_j, \xi^+_j], \text{ для всех } j = \overline{0, s}.$$

На этих интервальных множествах существуют функции плотности распределения вероятностей (ПРВ) $q_j(\xi[j+m])$, $j = \overline{0, s}$.

Зависимость между пакетами входных и выходных данных, с учетом ошибок измерений выхода объекта, описывается дискретной динамической рандомизированной моделью (ДДРМ) со структурированными нелинейностями второй степени и с аддитивным случайным шумом:

$$(13) \quad v[j+1] = \sum_{h=1}^R \left(\sum_{n=0}^m w^{(h)}[n] x[j+m-n] \right)^h + \xi[j+m], \quad j = \overline{0, s},$$

где объем “памяти” $m = 1$, $s = 1$ и $R = 2$.

Дискретные случайные весовые функции (случайные импульсные характеристики) $w^{(1)}[n] = w^{(2)}[n] = 0$, $n > 1$, описывают динамические свойства линейного и нелинейного блоков модели. Введем обозначения:

$$(14) \quad \begin{aligned} a_0^{(1)} &= w^{(1)}[0], \quad a_1^{(1)} = w^{(1)}[1]; & \mathbf{a}^{(1)} &= \{a_0^{(1)}, a_1^{(1)}\}; \\ a_0^{(2)} &= w^{(2)}[0], \quad a_1^{(2)} = w^{(2)}[1]; & \mathbf{a}^{(2)} &= \{a_0^{(2)}, a_1^{(2)}\}; \\ \mathbf{x}[j] &= \{x[j+1], x[j]\}. \end{aligned}$$

Здесь $\mathbf{x}[0] = \{0,39; 0,19\}$ и $\mathbf{x}[1] = \{0,93; 0,39\}$.

Случайные параметры $a_n^{(h)}$ — интервального типа, т.е. они принимают значения в следующих интервальных множествах:

$$(15) \quad \mathcal{A}_n^{(h)} = [-A_n^{(h)}, A_n^{(h)}], \quad A_n^{(h)} = \beta_n^{(h)} \exp(-\alpha_n^{(h)} n).$$

Параметры указанных интервальных множеств указаны в табл. 2.

На интервальных множествах (15) определены функции ПРВ $P^{(1)}(\mathbf{a}^{(1)})$, $P^{(2)}(\mathbf{a}^{(2)})$.

Таблица 2

Параметры	$\alpha_0^{(1)}$	$\alpha_1^{(1)}$	$\beta_0^{(1)}$	$\beta_1^{(1)}$
Значения	0,5	1,0	1,0	2,0

Таблица 3

Вариант	Π_θ	J_τ^*	θ_τ^*	ε_τ	\hat{r}_τ	\hat{P}_τ	t, c
1	$\theta_0 \in [-10, 10]$ $\theta_1 \in [-10, 10]$	0,0025	-2,4457 -0,0537	0,0025	0,1200	$\approx 1,0$	1576
2	$\theta_0 \in [-3, -2]$ $\theta_1 \in [-0,5, 0,5]$	0,0001	-2,4042 -0,0634	0,0001	0,0051	$\approx 1,0$	1573

Согласно методу оценивания [26] энтропийно-оптимальные функции ПРВ параметров ДДРМ и шумов измерений будут иметь вид:

$$\tilde{P}^{(h)}(\mathbf{a}) = \frac{1}{\mathcal{R}^{(h)}(\theta)} \exp \left(- \sum_{j=0}^s \theta_j V_j^{(h)}(\mathbf{a}^{(h)}, \mathbf{x}[j]) \right), \quad h = \overline{1, R};$$

$$(16) \quad \tilde{q}_j(\xi) = \frac{1}{\mathcal{Q}_j(\theta_j)} \exp(-\theta_j \xi), \quad j = \overline{0, s}.$$

Здесь

$$\mathcal{R}^{(h)}(\theta) = \int_{\mathcal{A}^{(h)}} \exp \left(- \sum_{j=0}^s \theta_j V_j^{(h)}(\mathbf{a}^{(h)}, \mathbf{x}[j]) \right) d\mathbf{a}^{(h)},$$

$$(17) \quad \mathcal{Q}_j(\theta) = \int_{\Xi_j} \exp(-\theta_j \xi_j) d\xi_j, \quad j = \overline{0, s}.$$

$$V_j^{(1)}(\mathbf{a}, \mathbf{x}) = a_0^{(1)} x[j+1] + a_1^{(1)} x[j],$$

$$(18) \quad V_j^{(2)}(\mathbf{a}, \mathbf{x}) = \left(a^{(2)} x[j+1] \right)^2 + \left(a^{(2)} x[j] \right)^2 + 2a_0^{(2)} a_1^{(2)} x[j+1] x[j].$$

Множители Лагранжа θ_0, θ_1 определяются решением системы уравнений

$$(19) \quad \Phi_j(\theta) = \sum_{h=1}^R \frac{\tilde{\mathcal{R}}_j^{(h)}(\theta)}{\mathcal{R}^{(h)}(\theta)} + \frac{\tilde{\mathcal{B}}_j(\theta_j)}{\mathcal{B}_j(\theta_j)} - y[j+1] = 0, \quad j = \overline{0, s},$$

где интегральные компоненты имеют вид:

$$\tilde{\mathcal{R}}_j^{(h)}(\theta) = \int_{\mathcal{A}^{(h)}} V_j^{(h)}(\mathbf{a}^{(h)}, \mathbf{x}[j]) \exp \left(- \sum_{i=0}^1 \theta_i V_i^{(h)}(\mathbf{a}^{(h)}, \mathbf{x}[j]) \right) d\mathbf{a}^{(h)},$$

$$(20) \quad \tilde{\mathcal{B}}_j(\theta) = \int_{\Xi_j} \xi_j \exp(-\theta_j \xi_j) d\xi_j.$$

Измерения значений выхода объекта: $y[1] = 0,5; y[2] = 2,2$.

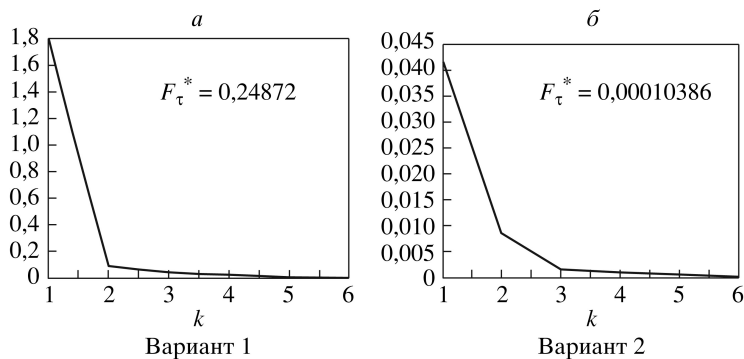


Рис. 3. $J_r^*(\theta)$.

Решение уравнений (19) сводится к минимизации невязки

$$(21) \quad J(\theta) = \|\Phi(\theta)\|_{L_2}$$

на кубе Π_θ .

Результаты расчетов представлены в табл. 3 и на рис. 3. Из-за наличия интегральных компонент (20) минимизируемая функция задана алгоритмически. Для проведения данного набора тестов алгоритм был реализован на платформе MATLAB с использованием компонента параллельных вычислений Parallel Computing Toolbox. Интегральные компоненты (20) вычисляются с помощью функции `integral2`, а поиск минимума осуществлялся GFS-алгоритмом. Расчеты проводились 32 рабочими процессами (`workers`) на той же вычислительной системе, что и в случае аналитических тестов. Определялось решение уравнений (21) для двух областей поиска Π_θ , представленных в табл. 3.

6. Заключение

Предложен новый метод поиска решений задач глобальной оптимизации с гельдеровской целевой функцией и непрерывными функциями в ограничениях, которые могут быть заданы только алгоритмически. Метод базируется на простых испытаниях Монте-Карло, с помощью которых формируются в результате пакетных итераций последовательности “квази-глобальных” минимумов и их декрементов. Проведено тестирование алгоритма на аналитических тестах, подтвердившее его работоспособность, и на алгоритмических тестах — эффективность.

ПРИЛОЖЕНИЕ

Вероятностные оценки

Рассмотрим процедуру генерации случайных, независимых и равномерно распределенных в d -мерном кубе векторов, осуществляемых параллельно функционирующими генераторами в количестве d штук. Принцип генерации иллюстрирует табл. 4. Строки таблицы заполняются d генераторами случайных независимых и равномерно распределенных на отрезке $[0, 1]$ чисел. В последней строке — случайные векторы с координатами, содержащимися в колонках таблицы.

Таблица 4

z_1	0,1	0,5	...	0,6
z_2	0,4	0,1	...	0,1
...
z_d	0,9	0,7	...	0,9
\mathbf{z}	\mathbf{z}^1	\mathbf{z}^2	...	\mathbf{z}^{N_k}

Напомним, что на k -м шаге алгоритма генерируется набор из N_k векторов.

Идея оценки состоит в том, чтобы найти такое *виртуальное разбиение* единичного d -мерного куба, при котором требуемая вероятность экспоненциально быстро стремится к нулю при росте числа шагов.

Будем искать такое разбиение в виде $\eta_k = N_k^{-s/2d}$, где η_k — сторона элементарного куба и s — показатель Гельдера, $0 < s \leq 1$.

Нетрудно показать (см. [22]), что вероятность того, что в конкретный элементарный куб не попадет ни один вектор из сгенерированного набора, равна

$$\left(1 - \eta_k^d\right)^{N_k},$$

а вероятность того, что такой элементарный куб найдется, т.е. $\mathbf{P}(N_k, d, s)$, не превосходит величины

$$\mathbf{P}(N_k, d, s) \leq \left(N_k^{s/2d} + 1\right)^d \left(1 - \eta_k^d\right)^{N_k} \stackrel{\text{def}}{=} \mathbf{Q}(N_k, d, s).$$

Теперь при $N_k \rightarrow \infty$ получаем

$$\mathbf{Q}(\cdot) \sim N_k^{s/2} \exp\left(-N_k^{1-s/2}\right).$$

Оценка отклонения от значения глобального минимума

Согласно лемме 2 из [22] с вероятностью большей, чем $1 - \mathbf{Q}(N_k, d, s)$, имеет место следующая оценка отклонения от значения глобального минимума для гильдеровских функций:

$$(II.1) \quad F_k^* - F^* \leq L \left(\frac{\sqrt{d}}{2} \eta_k\right)^s = L \left(\frac{\sqrt{d}}{2}\right)^s \frac{1}{N_k^{s^2/2d}}.$$

Таким образом, отсюда получаем, что существуют константы $D = L \left(\frac{\sqrt{d}}{2}\right)^s$ и $p = s^2/2d$ такие, что с вероятностью не менее $1 - \mathbf{Q}(\cdot)$ имеет место соотношение

$$(II.2) \quad 0 \leq F_k^* - F^* \leq DN_k^{-p} \stackrel{\text{def}}{=} r_k(\beta).$$

Так как $r_{k+1}(\beta) < r_k(\beta)$, то

$$(II.3) \quad u_k \stackrel{\text{def}}{=} |F_{(k+1)}^* - F_k^*| = |F_{(k+1)} - F^* + F^* - F_k^*| \leq \tilde{D}N_k^{-p},$$

где $\tilde{D} = D(1 + 1/\alpha^p)$ (напомним, что $N_{(k+1)} = \alpha N_k$, $\alpha > 1$).

Учитывая (П.3), естественно высказать гипотезу

$$u_k \approx \tilde{D} N_k^p$$

или, иными словами,

$$(П.4) \quad \log u_k = A - B \log N_k + \zeta_k,$$

где $A = \log \tilde{D}$, $B = p$ и ζ_k — последовательность независимых и одинаково распределенных случайных величин с нулевым математическим ожиданием и конечной дисперсией. Гипотеза (П.4) экспериментально подтверждена.

Генерация случайных векторов в допустимой области

Рассмотрим допустимую область $\mathcal{K}_{\mathbf{z}}$ (1) в единичном кубе. Будем предполагать, что в ней нет изолированных точек и ее d -мерная лебегова мера ν отлична от нуля, $0 < \nu_0 \leq \nu \leq 1$.

Предположим сначала, что величина ν известна. Тогда при генерации последовательности равномерно распределенных в кубе и независимых случайных векторов вероятность того, что конкретный вектор попадет в допустимую область, равна ν . Пусть на шаге k сгенерировано N_k независимых и равномерно распределенных на Z_+^d случайных векторов, из которых \hat{N}_k попали в допустимую область. Обозначим через κ_{N_k} частоту попадания сгенерированных векторов в допустимое множество $\mathcal{K}_{\mathbf{z}}$:

$$(П.5) \quad \kappa_{N_k} = \frac{\hat{N}_k}{N_k}.$$

Из закона больших чисел следует, что для любых (достаточно малых) $\epsilon > 0$, $\eta > 0$ и данного ν существует такое число $N_0(\epsilon, \eta, \nu)$, что при всех $N_k > N_0(\epsilon, \eta, \nu)$ с вероятностью не менее $(1 - \eta)$ отклонение $|\kappa_{N_k} - \nu|$ будет меньше, чем $\epsilon \nu$.

Таким образом, при $N_k > N_0(\epsilon, \eta, \nu)$ с вероятностью не менее $(1 - \eta)$

$$(П.6) \quad \kappa_{N_k} \geq (1 - \epsilon)\nu.$$

Следовательно, количество допустимых векторов \hat{N}_k в пакете из N_k векторов будет с вероятностью не менее, чем $(1 - \eta)$ удовлетворять неравенству

$$(П.7) \quad \hat{N}_k \geq N_k(1 - \epsilon)\nu.$$

Отсюда, в свою очередь, следует, что во всех формулах, где присутствует количество сгенерированных векторов N_k , надо это число (при $N_k > N_0(\cdot)$) заменить на $\hat{N}_k = (1 - \epsilon)\nu N_k$.

Приведем теперь оценку для числа $N_0(\epsilon, \eta, \nu)$. Для $\eta > 0$ определим функцию $c(\eta)$ из условия

$$(П.8) \quad \frac{1}{\sqrt{2\pi}} \int_{-c(\eta)}^{c(\eta)} \exp(-x^2/2) dx = 1 - \eta.$$

Тогда число $N_0(\epsilon, \eta, \nu)$ определяется из соотношения

$$(II.9) \quad N_0(\epsilon, \eta, \nu) \geq \frac{c^2(\eta)(1 - \nu)}{\epsilon^2 \nu}.$$

Теперь рассмотрим ситуацию, когда величина ν не известна, однако, известно число ν_0 такое, что $0 < \nu_0 \leq \nu$. Тогда, подставляя в (II.9) величину ν_0 вместо ν , получаем оценку для минимального числа сгенерированных векторов, при котором частота допустимых векторов с вероятностью не менее, чем $(1 - \eta)$ будет отличаться от неизвестного ν не более, чем на $\epsilon \nu$. В этом случае при $N > N_0(\epsilon, \eta, \nu)$ оценка $\hat{\nu}$ числа ν имеет вид

$$(II.10) \quad \hat{\nu} = \kappa_N,$$

причем с вероятностью не менее $(1 - \eta)$ будет выполнено соотношение $|\kappa_N - \nu| \leq \epsilon \nu$, т.е.

$$(II.11) \quad \frac{\kappa_N}{1 - \epsilon} \geq \nu \geq \frac{\kappa_N}{1 + \epsilon}.$$

После того как такая оценка будет получена, можно использовать ее нижнее значение вместо ν в формуле (II.9).

Описание тестовых задач

Приведем описания тестовых задач поиска глобальных минимумов непрерывных функций на параллелепипедах.

A1. *Shekel function*

$$(II.12) \quad F(\mathbf{x}) = - \sum_{i=1}^m \left(c_i + \sum_{j=1}^d (x_j - a_{ji})^2 \right)^{-1}, \quad m = 10,$$

$$\mathbf{x}^* = \mathbf{4}, \mathbf{0},$$

$$F^* = -10,5364, \quad d = 4,$$

$$F^* = -11,0216, \quad d = 2.$$

$$A = \begin{pmatrix} 4 & 1 & 8 & 6 & 3 & 2 & 5 & 8 & 6 & 7 \\ 4 & 1 & 8 & 6 & 7 & 9 & 3 & 1 & 2 & 3 \\ 4 & 1 & 8 & 6 & 3 & 2 & 5 & 8 & 6 & 7 \\ 4 & 1 & 8 & 6 & 7 & 9 & 3 & 1 & 2 & 3 \end{pmatrix},$$

$$c = \frac{1}{10}(1, 2, 2, 4, 4, 6, 3, 7, 5, 5).$$

A2. *Rosenbrock function*

$$(II.13) \quad F(\mathbf{x}) = \sum_{i=1}^d 100(x_{i+1} - x_i)^2 + (x_i - 1)^2,$$

$$\mathbf{x}^* = \mathbf{1}, \quad F^* = 0 \text{ для любого } d.$$

A3. Powell function

$$(П.14) \quad F(\mathbf{x}) = \sum_{i=1}^{d-2} (x_{i-1} + 10x_i)^2 + 5(x_{i+1} - x_{i+2})^2 + \\ + (x_i - 2x_{i+1})^4 + 10(x_{i-1} - x_{i+2})^4,$$

$\mathbf{x}^* = \mathbf{0}$, $F^* = 0$ для любого d .

A4. Trigonometric function

$$(П.15) \quad F(\mathbf{x}) = 1 + \sum_{i=1}^d 8 \sin^2 7(x_i - 0,9)^2 + 6 \sin^2 14(x_i - 0,9)^2 + x_i - 0,9, \\ \mathbf{x}^* = \mathbf{0,9}, \quad F^* = 1 \text{ для любого } d.$$

A5. Griewank function

$$(П.16) \quad F(\mathbf{x}) = 1 + \sum_{i=1}^d \frac{x_i^2}{4000} - \prod_{i=1}^d \cos \left(\frac{x_i}{\sqrt{i+1}} \right), \\ \mathbf{x}^* = \mathbf{0}, \quad F^* = 0 \text{ для любого } d.$$

Обозначения в логической структуре GFS

Процедура SeqLocMin

- d — размерность вектора \mathbf{z} ;
- M_k — длина набора значений случайных чисел на шаге k по каждой из координат вектора \mathbf{z} ;
- N_k — объем пакета \mathcal{Z}_k случайных векторов на шаге k ;
- $\mathcal{Z}_k = \{\mathbf{z}^1, \dots, \mathbf{z}^{N_k}\}$ — пакет случайных равномерно распределенных в Z_+^n векторов \mathbf{z} ;
- $\tilde{\mathcal{Z}}_k = \{\mathbf{z}^1, \dots, \mathbf{z}^{\tilde{N}_k}\}$, $\tilde{N}_k \leq N_k$ — пакет допустимых случайных векторов, принадлежащих допустимому множеству $\mathcal{K}(1)$;
- F_k^* — минимум целевой функции F на допустимом пакете $\tilde{\mathcal{Z}}_k$;
- $\mathcal{F} = \{F_0^*, F_1^*, \dots, F_k^*, \dots\}$ — последовательность “квази-глобальных” минимумов;
- $u_i = |F_i^* - F_{(i-1)}^*|$, $i = \overline{1, k}$, — декремент последовательности \mathcal{F}^k ;
- $\mathcal{U} = \{u_0, u_1, \dots, u_k, \dots\}$ — последовательность декрементов;
- $v_i = \log u_i$, $i = \overline{0, k}$;
- $\mathcal{V} = \{v_0, v_1, \dots, v_k, \dots\}$ — последовательность логарифмов декрементов;
- $\tilde{\mathcal{V}}(\rho) = \{\tilde{v}_{k-\rho}, \tilde{v}_{k-\rho+1}, \dots, \tilde{v}_k, \dots\}$ — последовательность логарифмов декрементов, меньших $\log \delta$;
- $n_i = \log N_i$, $i = \overline{0, k}$;
- $\mathcal{N} = \{n_0, n_1, \dots, n_k, \dots\}$ — последовательность логарифмов объемов пакетов;
- δ — допустимая погрешность вычисления значений декрементов;

- ρ — заданное количество элементов последовательности $\tilde{V}^k(\rho)$, которые должны оставаться внутри заданной трубки.

Процедура EstProb

- $\hat{A}_\tau, \hat{B}_\tau$ — оценки параметров аппроксимации после τ шагов;
- \hat{P}_τ — нижняя оценка вероятности (9);
- \hat{r}_τ — верхняя оценка расстояния до точного решения (10);

СПИСОК ЛИТЕРАТУРЫ

1. *Land A.N., Dorg A.G.* An Automatic Method of Solving Discrete Programming Problems // *Econometrics*. 1960. V. 28. No. 3. P. 497–520.
2. *Horst R., Tuy H.* Global Optimization. Deterministic Approaches. Springer, 1996.
3. *Хамисов О.В.* Глобальная оптимизация функций с вогнутой опорной минорантой // *Журн. вычислит. матем. и мат. физики*. 2004. Т. 44. № 9. С. 1552–1563.
4. *Евтушенко Ю.Г., Посыпкин М.А.* Варианты метода неравномерных покрытий для глобальной оптимизации частично целочисленных нелинейных задач // *ДАН*. 2011. Т. 437. № 2. С. 168–172.
5. *Колпаков Р.М., Посыпкин М.А.* Верхняя и нижняя оценки трудоемкости метода ветвей и границ // *Дискретная математика*. 2013. Т. 22. № 1. С. 58–73.
6. *Horst R., Pardalos P.M., Thoai N.V.* Introduction to Global Optimization. Kluwer Acad., 1996.
7. *Стрекаловский А.С.* Элементы невыпуклой оптимизации. Новосибирск: Наука, 2003.
8. *Strongin R.G., Sergeyev Ya.D.* Global Optimization with Non-Convex Constraints. Sequential and Parallel Algorithms. Dordrecht: Kluwer Acad. Publishers, 2000.
9. *Сергеев Я.Д., Квасов Д.Е.* Диагональные методы глобальной оптимизации. М.: Физматлит, 2008.
10. *Sergeev Y.D., Strongin R.G., Lera D.* Introduction to Global Optimization Exploiting Space-Filling Curves. SpringerBriefs in Optimisation, 2013.
11. *Тихонова М.В., Рябов В.В., Спивак Ц.И. и др.* Параллельная условная глобальная оптимизация для математического моделирования кинетики химических реакций // *Вычислит. методы и программирование*. 2013. Т. 14. С. 262–268.
12. *Caflich R.E.* Monte Carlo and Quasi-Monte Carlo methods // *Acta Numerica*. 1998. P. 1–49.
13. *Zhiglavsky A., Žilinskas A.* Stochastic Global Optimization. Springer, 2006.
14. *Rubinstein R.Y., Kroese D.P.* Simulation and the Monte Carlo Method. John Wiley & Sons, 2008.
15. *Стронгин Р.Г., Гергель В.П., Гришагин В.А. и др.* Параллельные вычисления в задачах глобальной оптимизации. М.: Изд-во Моск. ун-та, 2012.
16. *Zabinsky Z.B., Smith R.L., McDonald F., et al.* Improving Hit-and-Run for Global Optimization // *J. Global Optimization*. 1993. V. 3. P. 171–192.
17. *Polyak B., Gryazina E.* Hit-and-Run: New Design Technique for Stabilization, Robusness and Optimization of Linear Systems // *Proc. IFAC World Congress*. 2008. P. 376–380.
18. *Hastings W.K.* Monte Carlo Sampling Methods Using Markov Chains and Its Applications // *Biometrika*. 1970. V. 57. P. 92–109.
19. *Боровков К.А.* Об одном новом варианте метода Монте-Карло // *Теория вероятностей и ее применения*. 1991. Т. 36. Вып. 2. С. 342–346.

20. *Rubinstein R.Y., Kroese D.P.* The Cross-Entropy Methods. A Unified Approach to Combinatorial Optimization, Monte-Carlo Simulation and Machine Learning. Springer, Information Science and Statistics, 2004.
21. *Зорин А.В., Федоткин М.А.* Методы Монте-Карло для параллельных вычислений, М.: Изд-во Моск. ун-та, 2012.
22. *Дарховский Б.С., Попков А.Ю., Попков Ю.С.* Метод пакетных итераций Монте Карло для решения систем нелинейных уравнений и неравенств: вероятностные характеристики // *АиТ.* 2015. № 5. С. 87–98.
Darkhovskii B.S., Popkov Y.S., Popkov A.Y. Monte Carlo Method of Batch Iterations: Probabilistic Characteristics // *Autom. Remote Control.* 2015. V. 76. No. 5. P. 776–785.
23. *Fu M.C., Hu J., Marcus S.I.* Model-Based Randomized Methods for Global Optimization // *Proc. 17th Int. Sympos. on Mathematical Theory of Networks and Systems.* Kyoto, Japan, July 24–28, 2006.
24. Virtual Library of Simulation Experiments: Test Functions and Datasets [Электронный ресурс]. URL: <http://www.sfu.ca/~ssurjano/optimization.html> (Дата обращения: 18.09.2015).
25. Test functions for optimization [Электронный ресурс]. URL: http://en.wikipedia.org/wiki/Test_functions_for_optimization (Дата обращения: 18.09.2015).
26. *Popkov Y., Popkov A.* New Method of Entropy-Robust Estimation for Randomized Models under Limited Data // *Entropy.* 2014. V. 16. P. 675–698.

Статья представлена к публикации членом редколлегии П.С. Щербаковым.

Поступила в редакцию 27.09.2015