[ИПМ им.М.В.Келдыша РАН](http://keldysh.ru/) • [Электронная библиотека](http://library.keldysh.ru/)

[Препринты ИПМ](http://library.keldysh.ru/preprints/) • [Препринт № 47 за 2016 г.](http://library.keldysh.ru/preprint.asp?id=2016-47)

ISSN 2071-2898 (Print)
ISSN 2071-2901 (Online)

Лагуновская А.А.

Алгоритмические модели поведения человека по Ю.Е. Нестерову

Рекомендуемая форма библиографической ссылки: Лагуновская А.А. Алгоритмические модели поведения человека по Ю.Е. Нестерову // Препринты ИПМ им. М.В.Келдыша. 2016. № 47. 12 с. doi:10.20948/prepr-2016-47
URL: http://library.keldysh.ru/preprint.asp?id=2016-47

О р д е н а   Л е н и н а
ИНСТИТУТ ПРИКЛАДНОЙ МАТЕМАТИКИ
имени М.В.Келдыша
Р о с с и й с к о й   а к а д е м и и   н а у к

А.А. Лагуновская

Алгоритмические модели поведения человека по Ю.Е. Нестерову

Москва — 2016

Лагуновская А.А.

Алгоритмические модели поведения человека по Ю.Е. Нестерову

В работе предполагается рассмотреть несколько сюжетов решения задач выпуклой стохастической оптимизации методами, которые соответствуют возможному поведению людей. Изложение построено в основном вокруг стохастического онлайн-варианта метода зеркального спуска. Оригинальности изложению добавляет то обстоятельство, что практически все изложение ведется в условиях наличия шумов в постановке задачи, также и не случайной природы. Одной из целей данного препринта является ответ на вопросы, поставленные во время заключительной лекции из цикла лекций, прочитанных Ю.Е. Нестеровым в апреле 2016 года в НИУ ВШЭ (https://www.youtube.com/user/PreMoLab).

Ключевые слова: выпуклая стохастическая оптимизация, метод зеркального спуска, онлайн оптимизация, алгоритмические модели человеческого поведения

Anastasia Aleksandrovna Lagunovskaya

Algorithmic models of human behavior following Yu.E.Nesterov

Given research considers several convex stochastic optimization problems. These problems are solved using methods that are very similar to the possible people behavior. The presentation is based mainly on the consideration of the online stochastic version of the mirror descent method. Novelty is made up by the fact that almost all the scenarios are considered in the case of noise, including those of nonrandom nature. One of the goals of this preprint is to answer the questions raised during the final lecture of the lecture series performed by Yu.E. Nesterov in April 2016 in the HSE (Moscow) (https://www.youtube.com/user/PreMoLab).

Key words: convex stochastic optimization, mirror descent method, online optimization, algorithmic models of human behavior

Работа выполнена при поддержке Российского фонда фундаментальных исследований, проекты 14-01-00722-а и 15-31-20571 мол-а-вед.

Оглавление

Постановка задачи

Далее будет рассматриваться задача выпуклой онлайн-минимизации [0] псевдо-регрета

$$\mathrm{Regret}_N\left(\left\{f_k(\cdot)\right\},\left\{x^k\right\}\right)=$$
$$=\frac{1}{N}\sum_{k=1}^{N}f_k\left(x^k\right)-\min_{x\in Q}\frac{1}{N}\sum_{k=1}^{N}f_k(x)\to\min_{\left\{x^k\left(\nabla_x\tilde{f}_1\left(x^1,\xi^1\right);\ldots;\nabla_x\tilde{f}_{k-1}\left(x^{k-1},\xi^{k-1}\right)\right)\in Q\right\}_{k=1}^{N}}. \qquad (1)$$

В данной работе мы, в основном, ограничимся случаем множества простой структуры, а именно симплексом:

$$Q=S_n(1)=\left\{x\geq 0:\ \sum_{i=1}^{n}x_i=1\right\},\ \|\cdot\|=\|\cdot\|_1,\ \|\cdot\|_*=\|\cdot\|_\infty.$$

Смысл приведенной постановки задачи в следующем. Пусть мы заранее знаем все выпуклые функции $f_k(\cdot)$, тогда оптимальное решение x_* определяется как решение задачи

$$\frac{1}{N}\sum_{k=1}^{N}f_k(x)\to\min_{x\in Q}.$$

Однако на самом деле нам предлагается минимизировать

$$\frac{1}{N}\sum_{k=1}^{N}f_k\left(x^k\right)\to\min_{\left\{x^k\left(\nabla_x\tilde{f}_1\left(x^1,\xi^1\right);\ldots;\nabla_x\tilde{f}_{k-1}\left(x^{k-1},\xi^{k-1}\right)\right)\in Q\right\}_{k=1}^{N}},$$

выбирая на каждом шаге $k=1,\ldots,N$ вектор $x^k\in Q$, исходя только из доступной на этом шаге информации

$$\nabla_x\tilde{f}_1\left(x^1,\xi^1\right);\ldots;\nabla_x\tilde{f}_{k-1}\left(x^{k-1},\xi^{k-1}\right),$$

где

$$E_{\xi^k}\left[\left\|\nabla_x\tilde{f}_k\left(x,\xi^k\right)-\nabla f_k(x)\right\|_*\right]\leq\delta/\tilde{R},\ \max_{x,y\in Q}\|x-y\|\leq\tilde{R}.$$

Случайные величины $\left\{\xi^k\right\}$ являются независимыми одинаково распределенными. Таким образом, $\mathrm{Regret}_N\left(\left\{f_k(\cdot)\right\},\left\{x^k\right\}\right)$ характеризует усредненные потери (приходящиеся на один шаг) за “не-знание будущего”.

Онлайновость постановки задачи допускает, что на каждом шаге k функция $f_k(\cdot)$ может выбираться из рассматриваемого класса функций враждебно по отношению к используемому нами методу генерации последовательности $\left\{x^k\right\}$. В частности, $f_k(\cdot)$ может зависеть от $\left\{x^1;\ldots;x^k\right\}$.

Чтобы строго сформулировать основной результат в нужной для последующих приложений (см. Пример 4 ниже) общности, предположим, что выполняются следующие **условия** [2]:

1. для любых $N \in \mathbb{N}$ (Ξ^{k-1} – сигма алгебра, порожденная ξ^1, …, ξ^{k-1})

$$\sup_{\left\{x^k = x^k\left(\xi^1,\dots,\xi^k\right)\right\}_{k=1}^N} E\left[\frac{1}{N}\sum_{k=1}^N \left\langle E_{\xi^k}\left[\nabla_x f_k\left(x^k,\xi^k\right) - \nabla_x \tilde{f}_k\left(x^k,\xi^k\right) \middle| \Xi^{k-1}\right], x^k - x_*\right\rangle\right] \le \delta$$

где x_* – решение задачи

$$\frac{1}{N}\sum_{k=1}^N f_k(x) \to \min_{x\in Q},$$

$$E_{\xi^k}\left[\nabla_x f_k\left(x,\xi^k\right)\right] = \nabla f_k(x).$$

Случайные величины $\left\{\xi^k\right\}$ являются независимыми одинаково распределенными. Онлайновость постановки задачи допускает, что на каждом шаге k функция $f_k(\cdot)$ может выбираться из рассматриваемого класса функций враждебно по отношению к используемому нами методу генерации последовательности $\left\{x^k\right\}$. В частности, $f_k(\cdot)$ может зависеть от $\left\{x^1;\dots;x^k\right\}$.

Введем **определение** δ**-оракула**, выдающего зашумленное значение градиента целевой функции $\nabla_x \tilde{f}_k(x,\xi)$: при получении на вход точки x, оракул выдает $\nabla_x \tilde{f}_k(x,\xi)$ при этом условие 1 должно выполняться с допустимым уровнем шума δ .

2. $\left\{f_k(\cdot)\right\}$ – выпуклые функции;

3. для любых $k = 1,\dots,N$, $x \in Q$

$$E_\xi\left[\left\|\nabla_x \tilde{f}_k(x,\xi)\right\|_\infty^2\right] \le M^2.$$

Далее опишем способ выбора

$$\left\{x^k\left(\nabla_x \tilde{f}_1\left(x^1,\xi^1\right);\dots;\nabla_x \tilde{f}_{k-1}\left(x^{k-1},\xi^{k-1}\right)\right) \in Q\right\}_{k=1}^N$$

и приведем оценку скорости его сходимости.

Описание метода

Для решения задачи (1) воспользуемся онлайн-вариантом метода двойственных усреднений Ю.Е. Нестерова [3].

Положим $x_i^1 = 1/n$, $i = 1,...,n$. Пусть $t = 1,...,N-1$.

$$x_i^{t+1} = \frac{\exp\left(-\frac{1}{\beta_{t+1}}\sum_{k=1}^{t}\frac{\partial \tilde{f}_k\left(x^k,\xi^k\right)}{\partial x_i}\right)}{\sum_{l=1}^{n}\exp\left(-\frac{1}{\beta_{t+1}}\sum_{k=1}^{t}\frac{\partial \tilde{f}_k\left(x^k,\xi^k\right)}{\partial x_l}\right)}, \quad i = 1,...,n, \quad \beta_t = \frac{M\sqrt{t}}{\sqrt{\ln n}}. \qquad (2)$$

Опишем подробнее общий случай (когда $Q \neq S_n(1)$) метода зеркального спуска для решения задачи (1). Введем норму $\| \ \|$ в прямом пространстве (соответствующая ей сопряженная норма $\| \ \|_*$) и прокс-функцию $d(x)$, сильно выпуклую относительно этой нормы, с константой сильной выпуклости не менее 1. Выберем точку старта $x^1 = \arg\min_{x\in Q} d(x)$, считаем, что $d\left(x^1\right) = 0$, $\nabla d\left(x^1\right) = 0$. Введем расстояние Брэгмана

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

Далее считаем, что $d(x) = V_{x^1}(x) \le R^2$ для всех $x \in Q$. Определим оператор проектирования согласно введенному расстоянию как

$$\text{Mirr}_{x^k}(g) = \arg\min_{y\in Q}\left\{\left\langle g, y - x^k\right\rangle + V_{x^k}(y)\right\}.$$

Метод зеркального спуска (МЗС) для задачи (1) будет иметь вид

$$x^{k+1} = \text{Mirr}_{x^k}\left(\alpha_k \nabla_x \tilde{f}_k\left(x^k,\xi^k\right)\right), k = 1,...,N.$$

Тогда при выполнении условия (2) для любого $u \in Q, k = 1,...,N$ имеет место неравенство

$$\alpha_k\left\langle \nabla_x \tilde{f}_k\left(x^k,\xi^k\right), x^k - u\right\rangle \le \frac{\alpha_k^2}{2}\left\|\nabla_x \tilde{f}_k\left(x^k,\xi^k\right)\right\|_*^2 + V_{x^k}(u) - V_{x^{k+1}}(u).$$

Проведя преобразования, получим:

$$N\cdot E\left[\text{Regret}_N\left(\left\{f_k(\cdot)\right\},\left\{x^k\right\}\right)\right] \le \frac{V_{x^1}(x_*)}{\alpha} - \frac{E\left[V_{x^{N+1}}(x_*)\right]}{\alpha} + \left(\frac{1}{2}M^2\alpha + \sigma\right)N \le$$

$$\leq \frac{R^2}{\alpha} + \left(\frac{1}{2}M^2\alpha + \sigma\right)N.$$

Выбираем $\alpha_k = \alpha = \frac{R}{M}\sqrt{\frac{2}{N}}$, тогда

$$E\left[\mathrm{Regret}_N\left(\left\{f_k(\cdot)\right\},\left\{x^k\right\}\right)\right] \leq MR\sqrt{\frac{2}{N}} + \sigma.$$

Для $Q = S_n(1)$ имеем $R^2 = \ln n$ [2], поэтому

$$E\left[\mathrm{Regret}_N\left(\left\{f_k(\cdot)\right\},\left\{x^k\right\}\right)\right] = \mathrm{O}\left(M\sqrt{\frac{\ln n}{N}} + \delta\right). \quad (3)$$

Выписанная оценка является неулучшаемой с точностью до мультипликативного числового множителя. Это остается верным и для детерминированного случая, когда нет шумов (т.е. $\sigma = 0$).

Сформулируем результат в виде теоремы.

Теорема 1. *Пусть ищется решение задачи стохастической оптимизации (1), при этом существует оракул, который выдает* $\nabla_x \tilde{f}_k(x,\xi)$. *Если выполняются условия 1, 2, 3, то при использовании МЗС, как описано выше, можно получить следующую оценку:*

$$E\left[\mathrm{Regret}_N\left(\left\{f_k(\cdot)\right\},\left\{x^k\right\}\right)\right] = \mathrm{O}\left(M\sqrt{\frac{\ln n}{N}} + \delta\right).$$

С помощью обобщения (3) результата работы [2] далее будет продемонстрировано, какие можно сделать уточнения упомянутой лекции Ю.Е. Нестерова.

Задачи оптимизации в теории принятия решений

В данном разделе мы рассмотрим несколько практических задач и их интерпретацию с точки зрения теории игр и поведения рационального индивида. Первые два примера являются задачами онлайн-оптимизации, решение которых будем находить с помощью описанного выше метода зеркального спуска.

Пример 1 (взвешивание экспертных решений при наличии шумов). Рассмотрим задачу взвешивания экспертных решений, следуя [0]. Имеется n

различных экспертов. Каждый эксперт “играет” на рынке. Процесс повторяется $N \gg 1$ раз. Пусть l_i^k – проигрыш (выигрыш со знаком минус) эксперта i на шаге k ($\left|l_i^k\right| \le M$). На каждом шаге k распределяется один доллар между экспертами согласно вектору $x^k \in S_n(1)$. Потери, которые мы при этом несем, рассчитываются по потерям экспертов $f_k(x) = \langle l^k, x \rangle$. Целью является таким образом организовать процедуру распределения доллара на каждом шаге, чтобы суммарные потери (за N шагов) были минимальны. Допускается, что потери экспертов l^k могут зависеть еще и от текущего «хода» x^k. Установленные далее в этом разделе результаты (формула (3)) позволяют утверждать, что если на каждом шаге можно наблюдать лишь зашумленные проигрыши всех экспертов

$$\frac{\partial \tilde{f}_k\left(x^k, \xi^k\right)}{\partial x_i} = l_i^k + \xi_i^k + \delta_i^k,$$

где $\left\{\xi_i^k\right\}$ – независимые одинаково распределенные случайные величины $\xi_i^k \in N(0,1)$, $\left|\delta_i^k\right| \le \delta$, то существует такой способ действий

$$x^k\left(\nabla_x \tilde{f}_1\left(x^1, \xi^1\right), \ldots, \nabla_x \tilde{f}_{k-1}\left(x^{k-1}, \xi^{k-1}\right)\right) \text{ (метод (2), Теорема 1),}$$

который позволяет после N шагов проиграть (в среднем) лучшему (на этом периоде $k = 1, \ldots, N$) эксперту не более

$$\mathrm{O}\left((M+1)\sqrt{N \ln n} + \delta N\right)$$

долларов. Средний проигрыш за один период получается равным

$$\mathrm{O}\left((M+1)\sqrt{\frac{\ln n}{N}} + \delta\right).$$

Эта оценка оптимальна для данного класса задач [0]. Рассмотренный пример демонстрирует решение классической задачи о взвешивании экспертов в условиях шума (в том числе неслучайной природы). Обычно рассматривается ситуация без шума.

Пример 2. Предположим, что “успешность” некоторого человека зависит от того, как он распоряжается своим временем. Пусть выпуклая функция $f(x)$ со свойствами (заданная в некоторой окрестности $S_n(1)$ с сохранением всех свойств)

$$\left|f(x)\right| \le B,\ \left|f(x)-f(y)\right| \le M_2\left\|x-y\right\|_2$$

правильно отражает реальное “положение дел”, т.е. минимум этой функции соответствует оптимальной для данного человека конфигурации. Задача человека заключается в том, чтобы (получая каждый день “обратную связь”), так организовать “процесс своего обучения” (на основе получаемой информации), чтобы как можно быстрее достичь такого состояния $\bar{x}^N \in S_n(1)$, что имеет место неравенство

$$E\left[f\left(\bar{x}^N\right)\right]-\min_{x\in S_n(1)} f(x) \le \varepsilon. \qquad (4)$$

Будем считать, что имеется n различных родов деятельности. В k-й день человек распоряжается своим временем согласно вектору $x^k \in S_n(1)$. Этот вектор отражает доли времени, уделенного соответствующим делам. В конце каждого дня человек получает “обратную связь” от “внешнего мира” вида

$$f\left(x^k+\mu e^k,\tilde{\xi}^k\right)=f\left(x^k+\mu e^k\right)\cdot\left(1+\tilde{\xi}^k\right),$$

где $\left\{e^k\right\}$ – независимые одинаково распределенные (на евклидовой сфере единичного радиуса) случайные векторы $e^k \in RS_2^n(1)$, $\mu \le \varepsilon\big/\left(2\sqrt{2}M_2\right)$ $\left\{\tilde{\xi}^k\right\}$ – независимые (между собой и от $\left\{e^k\right\} \in RS_2^n(1)$) одинаково распределенные случайные величины $\tilde{\xi}^k \in N(0,1)$. Далее используем метод зеркального спуска, описанный выше, заменяя в методе (1) истинный градиент целевой функции его сглаженным дискретным аналогом [2]:

$$\nabla_x \tilde{f}_k\left(x^k,\xi^k=\left(e^k,\tilde{\xi}^k\right)\right):=\frac{n}{\mu}f\left(x^k+\mu e^k,\tilde{\xi}^k\right)e^k,\ M^2=\frac{8B^2 n\ln n}{\mu^2}, \qquad (5)$$

для которого справедливы условия 1–3 с

$$\delta=\mathrm{O}(\varepsilon),\ M^2=\frac{8B^2 n\ln n}{\mu^2}.$$

Определяя

$$\bar{x}^N=\frac{1}{N}\sum_{k=1}^N x^k,$$

получим из (2) согласно Теореме 1 (см. также Теорему 1 из [2]) , что человек может достичь в среднем псевдорегрет ε за

$$N = O\left(\frac{B^2 M_2^2 n \ln^2 n}{\varepsilon^4}\right)$$

дней. Если у человек есть возможность получать каждый день информацию

$$\tilde{f}\left(x^k + \mu e^k, \xi^k\right),\ \tilde{f}\left(x^k, \xi^k\right)$$

и

$$\left\|\nabla f\left(x\right) - \nabla f\left(y\right)\right\|_2 \le L_2 \left\|x - y\right\|_2,$$

то тогда можно улучшить оценку до (см. [4])

$$N = O\left(\frac{M_2^2 \ln^2 n}{\varepsilon^2}\right)$$

дней. Для этого необходимо по-другому выбирать параметр сглаживания:

$$\mu \le \min\left\{\max\left\{\frac{\varepsilon}{2\sqrt{2}M_2}, \sqrt{\frac{\varepsilon}{L_2}}\right\}, \frac{M_2}{L_2}\sqrt{\frac{1}{3n}}\right\}$$

и переписать формулу (5) следующим образом:

$$\nabla_x \tilde{f}_k\left(x^k, \xi^k = e^k\right) := \frac{n}{\mu}\left(f\left(x^k + \mu e^k, \tilde{\xi}^k\right) - f\left(x^k, \tilde{\xi}^k\right)\right)e^k,\ M^2 = 10 M_2^2 \ln n.$$

Здесь $M^2 = 10 M_2^2 \ln n$ в условии 3.

Приведенные здесь оценки заметно улучшают оценки (в $\sim n$ раз) работы [5] для этой же постановки.

Пример 3. Пусть имеется $\bar{N}$ экспертов (в данном примере в термин «эксперт» вкладывается отличный от примера 1 смысл), каждый из которых решает одну и ту же задачу

$$f\left(x\right) \to \min_{x \in S_n(1)},$$

получая каждый день свою (независимую от внешних факторов) реализацию стохастического градиента $\nabla_x f\left(x, \xi\right)$. Предполагается, что

$$E\left[\nabla_x f\left(x, \xi\right)\right] = \nabla f\left(x, \xi\right) \text{ и } E_\xi\left[\left\|\nabla_x f\left(x, \xi\right)\right\|_\infty^2\right] \le M_\infty^2,\ 0 \le f\left(x\right) \le M_\infty.$$

Согласно (2) каждый эксперт $l=1,...,\bar{N}$, “прожив”

$$N = \mathrm{O}\left(\frac{M_\infty^2 \ln n}{\varepsilon^2}\right)$$

дней, гарантированно сможет найти такой $\bar{x}^{N,l}$, что

$$E\left[f\left(\bar{x}^{N,l}\right)\right] - \min_{x\in S_n(1)} f(x) \le \varepsilon,$$

т.е. по неравенству Маркова

$$P\left(f\left(\bar{x}^{N,l}\right) - \min_{x\in S_n(1)} f(x) \le 2\varepsilon\right) \ge \frac{1}{2}.$$

Возникает вопрос: сколько нужно таких экспертов $\bar{N}$, чтобы на основе их ответов $\left\{\bar{x}^{N,l}\right\}$ можно было бы найти такой $\bar{x}^{\bar{N}}$, что с вероятностью $\ge 1-\sigma$

$$f\left(\bar{x}^{\bar{N}}\right) - \min_{x\in S_n(1)} f(x) \le 2\varepsilon?$$

Если у нас есть доступ к соответствующим значениям функции $\left\{f\left(\bar{x}^{N,l}\right)\right\}$, то с помощью сравнения экспертов и определения «победителя» можно достичь требуемого условия, ограничившись $\log_2 \sigma^{-1}$ экспертами. Если имеет место неравенство $\left\|\nabla_x f(x,\xi)\right\|_\infty^2 \le M_\infty^2$, то аналогичную по порядку оценку можно получить с помощью всего лишь одного эксперта, дав ему чуть подольше «пожить» $N = \mathrm{O}\left(M_\infty^2 \ln(n/\sigma)/\varepsilon^2\right)$. Вопрос, который изучается в этом примере (в предположении, что значениям функции $\left\{f\left(\bar{x}^{N,l}\right)\right\}$ недоступны), – сколько экспертов в общем случае в состоянии заменить одного «мудреца» (долго живущего эксперта)? Пусть

$$P\left(\frac{\left\|\nabla_x f(x,\xi)\right\|_\infty^2}{M_\infty^2} \ge t\right) = \mathrm{O}\left(\frac{1}{t^\alpha}\right), \ \alpha > 2 \qquad (6)$$

и

$$\left(\frac{M_\infty \sqrt{\ln n}}{\varepsilon \sigma^{1/\alpha}}\right)^{\frac{\alpha}{\alpha-1}} \gg \frac{M_\infty^2 \ln n}{\varepsilon^2}.$$

$(2\varepsilon,\sigma)$**-мудрецом** называется такой «долгоживущий» эксперт, что на основе ответов мудреца с вероятностью не менее $1-\sigma$ исходную задачу удается решить с точностью 2ε по целевому функционалу. Согласно [6] $(2\varepsilon,\sigma)$-мудрецу потребуется

$$N = \mathrm{O}\left(\left(\frac{M_\infty\sqrt{\ln n}}{\varepsilon\sigma^{1/\alpha}}\right)^{\frac{\alpha}{\alpha-1}}\right) \tag{7}$$

дней. Если же взять (условие (4) не используется)

$$\bar{N} = \mathrm{O}\left(\frac{M_\infty^2 \ln \sigma^{-1}}{\varepsilon^2}\right)$$

экспертов, то по неравенству Хефдинга [0] (можно использовать (6), используя вместо неравенства Хефдинга более точное неравенство Бернштейна, – это немного уточнит приводимую выше оценку)

$$\bar{x}^{\bar{N}} = \frac{1}{\bar{N}}\sum_{l=1}^{\bar{N}} x^{N,l}$$

с вероятностью $\geq 1-\sigma$ удовлетворяет

$$f\left(\bar{x}^{\bar{N}}\right) - \min_{x\in S_n(1)} f(x) \leq 2\varepsilon .$$

При высоких доверительных уровнях (то есть когда σ мало) и при «терпимости» к неточности (ε не очень мало) можно сделать вывод (так же как и в [0]), что выгоднее положиться на экспертов, чем на мудреца, особенно если учесть, что их можно “обучать” параллельно. Более точные количественные выводы могут быть сделаны исходя из выписанных формул. В подходе работы [0] условие (5) не предполагалось, как следствие, использовалось достаточно грубое неравенство Маркова, приводящее вместо (7) к оценке

$$N = \mathrm{O}\left(\frac{M_\infty^2 \ln n}{\varepsilon^2\sigma^2}\right).$$

*   *   *

Было рассмотрено три сюжета из теории принятия решения. Данные сюжеты были сформулированы как задачи стохастической выпуклой оптимизации. Для решения данных задач был использован онлайн-вариант метода зеркального спуска, который схож с алгоритмом поведения рационального индивида в условиях неопределенности (неопределенность достигается путем введения в рассмотрение шумов, в том числе неслучайной природы). Данный препринт позволяет получить оценки количества итераций, необходимых для достижения решения с заданной точностью. В некоторых описанных случаях (при отсутствии шумов) полученные оценки являются неулучшаемыми.

Библиографический список

1. Lugosi G., Cesa-Bianchi N. Prediction, learning and games. New York: Cambridge University Press, 2006.
2. Гасников А.В., Крымова Е.А., Лагуновская А.А., Усманова И.Н., Федоренко Ф.А. Стохастическая онлайн оптимизация. Одноточечные и двухточечные нелинейные многорукие бандиты. Выпуклый и сильно выпуклый случаи // Автоматика и Телемеханика. 2016. (в печати) URL: http://arxiv.org/abs/1509.01679
3. Nesterov Y. Primal-dual subgradient methods for convex problems // Math. Program. Ser. B. 2009. V. 120(1). P. 261–283.
4. Гасников А.В., Лагуновская А.А., Усманова И.Н., Федоренко Ф.А. Безградиентные прокс-методы с неточным оракулом для негладких задач выпуклой стохастической оптимизации на симплексе // Автоматика и телемеханика. 2016. (в печати) URL: http://arxiv.org/abs/1412.3890
5. Nesterov Yu. Random gradient-free minimization of convex functions // CORE Discussion Paper 2011/1. 2011.
6. Гасников А.В., Двуреченский П.Е., Нестеров Ю.Е. Стохастические градиентные методы с неточным оракулом // Труды МФТИ. 2016. Т. 8. № 1. С. 41–91. URL: http://arxiv.org/abs/1411.4218
7. Nesterov Y., Vial J.-Ph. Confidence level solution for stochastic programming // Automatica. 2008. V. 44. No. 6. P. 1559–1568.