A. V. Gasnikov, Yu. E. Nesterov, V. G. Spokoiny, On the efficiency of a randomized mirror descent algorithm in online optimization problems, *Zh. Vychisl. Mat. Mat. Fiz.*, 2015, Volume 55, Number 4, 582–598

DOI: 10.7868/S0044466915040043

ЖУРНАЛ ВЫЧИСЛИТЕЛЬНОЙ МАТЕМАТИКИ И МАТЕМАТИЧЕСКОЙ ФИЗИКИ, 2015, том 55, № 4, с. 582–598

УДК 519.658

# ОБ ЭФФЕКТИВНОСТИ ОДНОГО МЕТОДА РАНДОМИЗАЦИИ ЗЕРКАЛЬНОГО СПУСКА В ЗАДАЧАХ ОНЛАЙН ОПТИМИЗАЦИИ[1)]

© 2015 г. **А. В. Гасников, Ю. Е. Нестеров, В. Г. Спокойный**

*(141000 Долгопрудный, М.о., Институтский пер., 9, МФТИ; 127051 Москва, Большой каретный пер., 19, стр. 1, ИППИ РАН; 101000 Москва, ул. Мясницкая, 20, НИУ ВШЭ)*

*e-mail: gasnikov@yandex.ru; yurii.nesterov@uclouvain.be; spokoiny@wias-berlin.de*

Поступила в редакцию 03.09.2014 г.
Переработанный вариант 29.10.2014 г.

Предлагается рандомизированная онлайн версия метода зеркального спуска. Отличие от имеющихся версий заключается в способе рандомизации. Рандомизация выводится не на этапе вычисления субградиента функции, как это повсеместно принято, а на этапе проектирования на единичный симплекс. В результате получается покомпонентный субградиентный спуск со случайным выбором компоненты, допускающий онлайн интерпретацию. Это наблюдение, например, позволило единообразно проинтерпретировать результаты о взвешивании экспертных решений и предложить наиболее эффективный способ поиска равновесия в антагонистической матричной игре с разреженной матрицей. Библ. 34.

**Ключевые слова**: метод зеркального спуска, метод двойственных усреднений, онлайн оптимизация, экспоненциальное взвешивание, многорукие бандиты, взвешивание экспертов, стохастическая оптимизация, рандомизация.

**DOI:** 10.7868/S0044466915040043

## 1. ВВЕДЕНИЕ

В конце 70-х годов А.С. Немировским и Д.Б. Юдиным был предложен итерационный метод решения негладких выпуклых задач оптимизации [1], который можно интерпретировать как разновидность метода проекции субградиента, когда проектирование понимается, например, в смысле расстояния Брэгмана (Кульбака–Лейблера) [2], или как прямодвойственный метод (см. [3]). Этот метод, получивший название *метода зеркального спуска* (МЗС), позволяет хорошо учитывать структуру множества, на котором происходит оптимизация (например, симплекса). Как и многие другие методы решения негладких выпуклых оптимизационных задач, этот метод требует $O(M^2R^2/\varepsilon^2)$ итераций, где $\varepsilon$ – точность найденного решения по функции, что соответствует нижним оценкам по $\varepsilon$ (см. [1]). Однако константа $M$, равномерно ограничивающая норму субградиента оптимизируемой функции, и размер множества $R$ зависят от выбора нормы в пространстве, в котором ведется оптимизация. Так, если мы выбрали норму в нашем пространстве $l_p$, $1 \le p \le \infty$, то $M$ есть сопряженная $l_q$-норма субградиента ($1/p + 1/q = 1$), а $R^2$ есть “размер” множества в “метрике”, сильно выпуклой относительно $l_p$, с константой сильной выпуклости $\alpha \ge 1$.

**Замечание 1.** Слово “размер” взято в кавычки, потому что в действительности то, что задается, мы интерпретируем в данном контексте как квадрат размера (физически правильнее квадратом размера называть $R^2/\alpha$), поскольку “метрика” сильно выпуклая относительно нашей “рабочей” нормы в этом пространстве. Слово “метрика” взято в кавычки, потому что может быть не выполнено одно из свойств метрики – нет симметричности, например, для расстояния (дивергенции) Брэгмана.

При таких предположениях говорят, что выбранная “метрика” порождает прокс-структуру на множестве. Например, когда множество, на котором происходит оптимизация, является сим-

[1)] Работа выполнена при финансовой поддержке РФФИ (код проекта 14-01-00722-а), Лаборатории структурных методов анализа данных в предсказательном моделировании МФТИ грант правительства РФ дог. 11.G34.31.0073, гранта Президента РФ № МК-5285.2013.9, исследование приложений МЗС (см. разд. 4) выполнено за счет гранта РНФ (проект № 14-50-00150).

плексом, то, как правило, выбирают $p = 1$, а “метрику” задают расстоянием Брэгмана (Кульбака–Лейблера). При этом “проекция” на симплекс согласно такому расстоянию считается по явным формулам (экспоненциальное взвешивание). В [4] была выдвинута гипотеза о том, что применительно к задачам стохастической оптимизации на единичном симплексе (не онлайн) такой выбор нормы и расстояния являются наилучшими с точки зрения зависимости $M^2R^2$ от размера пространства $n$ (в типичных приложениях эта зависимость $\sim\ln n$). Однако в определенных ситуациях (в задачах о многоруких бандитах, когда $M^2R^2 \sim n\ln n$) удается выиграть логарифмический по $n$ фактор (см. ниже пример 1), более подходящим образом выбирая расстояние (см. [5]). При этом теряется возможность явного вычисления проекции.

В работах [3]–[8] исследовались стохастические (рандомизированные) версии МЗС. В том числе и онлайн (см. [5]). При этом анализировалась ситуация, когда именно градиент функции выдается оракулом со случайными шумами, но несмещенным образом. Такая релаксация детерминированного МЗС оказалась весьма полезной применительно к задачам адаптивного агрегирования оценок [4], оптимизации в пространствах огромных размеров [7], [8], задачах о многоруких бандитах и т.п. (см. [5], [9]–[11]).

В [1] была также отмечена возможность онлайн интерпретации МЗС. У разных авторов можно найти заметки на эту тему (см. [3], [5], [9], [11]–[14]). Наблюдение состоит в том, что ничего не изменится с точки зрения изучения сходимости метода (и его стохастической версии), если на каждом шаге допускать, что функция меняется, причем, возможно, враждебным образом (при этом оставаясь в классе выпуклых функций с ограниченной нормой субградиента).

В данной работе приводятся две версии стохастического онлайн МЗС. Первая версия близка к классической. Приблизительно в таком же виде ее уже можно было встретить в литературе (см. [3]–[5], [7]–[14]). Точнее говоря, предложенная версия аккумулирует в себе в виде частных случаев многие известные ранее версии МЗС. Вторая неявно была предложена в [6] применительно к поиску равновесия в антагонистической матричной игре (онлайн модификация в [6] не была затронута, равно как и связь предложенного метода с МЗС). Согласно [6], мы рандомизируем не на этапе вычисления субградиента функции, как это общепринято (см. [7], [8]), а на этапе проектирования на симплекс. В результате получается покомпонентный субградиентный спуск со случайным выбором компоненты, который, как будет ниже показано, допускает онлайн интерпретацию. Получив такой метод, мы расширяем множество тех задач онлайн оптимизации, к которым можно применять МЗС.

## 2. ОНЛАЙН МЗС СО СТОХАСТИЧЕСКИМ СУБГРАДИЕНТОМ

Рассмотрим задачу стохастической онлайн оптимизации (запись $E_{\xi^k}[f_k(x;\xi^k)]$ означает, что математическое ожидание берется по $\xi^k$, т.е. $x$ и $f_k$ понимаются в такой записи не случайными)

$$\frac{1}{N}\sum_{k=1}^{N} E_{\xi^k}[f_k(x;\xi^k)] \to \min_{x\in S_n(1)}, \quad S_n(1) = \left\{x \geq 0 : \sum_{i=1}^{n} x_i = 1\right\}, \tag{1}$$

при следующих условиях.

**Условие 1.** $E_{\xi^k}[f_k(x;\xi^k)]$ – выпуклые функции (по $x$), для этого достаточно выпуклости по $x$ функций $f_k(x;\xi^k)$ и независимости распределения $\xi^k$ от $x$.

**Условие 2.** Существует такой вектор $\nabla_x f_k(x;\xi^k)$, который для компактности будем называть *субградиентом*, хотя последнее верно не всегда (см. ниже пример 1), что

$$E_{\xi^k}\left(\nabla_x f_k(x;\xi^k) - \nabla_x E_{\xi^k}[f_k(x;\xi^k)] \middle| \Xi^{k-1}\right) \equiv 0,$$

где $\Xi^{k-1}$ есть $\sigma$-алгебра, порожденная случайными величинами $\xi^1,\ldots,\xi^{k-1}$. Ниже везде будем использовать обозначения обычного градиента для векторов, которые мы назвали здесь субградиентами. В частности, если мы имеем дело с обычным субградиентом, то запись $\nabla_x f_k(x;\xi^k)$ в вычислительном контексте (например, в итерационной процедуре МЗС, описанной ниже) означает какой-то его элемент (не важно какой именно), а если в контексте проверки условий (например, ниже в условии 3), то $\nabla_x f_k(x;\xi^k)$ пробегает все элементы субградиента (говорят также, субдифференциала).

**Условие 3.** $\left\|\nabla_x f_k(x;\xi^k)\right\|_\infty \le M$ – (равномерно, с вероятностью 1) ограниченный субградиент. Для справедливости части утверждений достаточно требовать одно из следующих (более слабых) условий:

$$\text{а) } E_{\xi^k}\left[\left\|\nabla_x f_k(x;\xi^k)\right\|_\infty^2\right] \le M^2, \quad \text{б) } E_{\xi^k}\left[\exp\left(\frac{\left\|\nabla_x f_k(x;\xi^k)\right\|_\infty^2}{M^2}\right)\Xi^{k-1}\right] \le \exp(1).$$

Задача (1) является лишь компактной (и далеко не полной) записью настоящей постановки задачи стохастической онлайн оптимизации. В действительности, требуется подобрать последовательность $\{x^k\}$ (в разд. 2 $\{x^k\} \in S_n(1)$, а в разд. 3 и ряде примеров разд. 4 при дополнительном ограничении, что $\{x^k\}$ выбираются с возможными повторениями среди вершин симплекса $S_n(1)$) исходя из доступной исторической информации ($x^k$ может зависеть только от $\left\{x^1,\xi^1,f_1(\cdot);\ldots;x^{k-1},\xi^{k-1},f_{k-1}(\cdot)\right\}$) так, чтобы минимизировать псевдо регрет

$$\frac{1}{N}\sum_{k=1}^N E_{\xi^k}[f_k(x^k;\xi^k)] - \min_{x\in S_n(1)}\frac{1}{N}\sum_{k=1}^N E_{\xi^k}[f_k(x;\xi^k)]$$

или регрет

$$E_{\xi^1,\ldots,\xi^N}\left[\min_{x\in S_n(1)}\frac{1}{N}\sum_{k=1}^N\left(f_k(x^k;\xi^k) - f_k(x;\xi^k)\right)\right].$$

В данном разделе мы сосредоточимся на минимизации псевдо регрета на основе информации $\left\{\nabla f_1(x^1;\xi^1);\ldots;\nabla f_{k-1}(x^{k-1};\xi^{k-1})\right\}$ при расчете $x^k$. Онлайновость постановки задачи допускает, что на каждом шаге $k$ функция $f_k$ может подбираться из рассматриваемого класса функций враждебно по отношению к используемому нами методу генерации последовательности $\{x^k\}$. В частности, здесь $f_k$ может зависеть от $\left\{x^1,\xi^1,f_1(\cdot);\ldots;x^{k-1},\xi^{k-1},f_{k-1}(\cdot);x^k\right\}$.

В целом в данной статье мы ограничиваемся рассмотрением задач минимизации псевдо регрета, когда на каждом шаге мы можем получить независимую реализацию стохастического субградиента в одной указанной нами (допустимой) точке. Приведенные в статье результаты можно распространить на случай, когда градиент выдается не точно (с не случайной ошибкой), выдается не полностью (скажем, вместо градиента выдается производная по выбранному направлению) или вместо градиента выдается только значение функции (см. [14]). Немного об этом написано далее (см. ниже пример 1). Другим способом релаксации исходной постановки является возможность несколько раз обращаться на одном шаге за значением градиента функции и(или) значением самой функции и взвешивать $f_k(x,\xi^k)$ в (1) разными весами (см. [14], [15]). Результаты статьи также можно распространить и на случай, когда функции $E_{\xi^k}[f_k(x;\xi^k)]$ равномерно сильно выпуклые по $x$ с константой $\mu$ (см. [5], [11], [13]). При этом выбирается евклидова прокс-структура, поскольку в сильно выпуклом случае (в стохастическом и не стохастическом) игра на выборе прокс-структуры не может дать выгоды. Неулучшаемая и достижимая оценка в этом случае будет иметь следующий вид: $\varepsilon = O(M^2\ln N/(\mu N))$. В не онлайн стохастическом случае эта оценка (с точностью до фактора $\ln N$) также будет неулучшаемой. Отметим при этом, что (для рассмотренных выпуклых и сильно выпуклых задач) в отличие от не онлайн случая, в онлайн случае игра на гладкости функций $E_{\xi^k}[f_k(x;\xi^k)]$ и(или) отсутствии стохастичности ($f_k(x;\xi^k) \equiv f_k(x)$) не дает никаких дивидендов (выписанные нижние оценки сохранятся).

Выше мы исходили из того, что оптимизация ведется на единичном симплексе. Возникает вопрос: насколько все, что приведено в статье, обобщается на более общий случай? Собственно говоря, ответ на этот вопрос частично известен уже давно (см. [1]). Приведенные в разд. 2 рассуждения универсальны, т.е. если исходить из оптимизации на каком-нибудь другом выпуклом компакте (от условия компактности (ограниченности) множества можно отказаться, см. [3], поскольку, в действительности, в оценку числа итераций входит не размер множества, а “расстояние” от точки старта до решения), то, задав норму в прямом пространстве и расстояние (сильно выпуклое относительно этой нормы), согласно которому будет осуществляться проектирование субградиента на этот компакт, можно повторить аналогичные рассуждения (см. [16]).

Для решения задачи (1) воспользуемся адаптивным методом зеркального спуска (точнее двойственных усреднений) в форме [3], [4]. Положим $x_i^1 = 1/n$, $i = 1,\dots,n$. Пусть $t = 1,\dots,N-1$.

**Алгоритм МЗС1-адаптивный. Метод двойственных усреднений**

Пусть

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

Не сложно показать, что этот метод представим также в виде

$$x^{t+1} = \arg\min_{x\in S_n(1)}\left\{\sum_{k=1}^{t}\left\{f_k(x^k;\xi^k) + \left\langle\nabla f_k(x^k;\xi^k), x - x^k\right\rangle\right\} + \beta_{t+1}V(x)\right\}$$

или

$$\begin{gathered} y^k = y^{k-1} - \gamma_k\nabla_x f_k(x^k;\xi^k), \\ x^{k+1} = \nabla W_{\beta_{k+1}}(y^k),\quad y^0 = 0,\quad \gamma_k \equiv 1,\quad \beta_k = \frac{M}{\sqrt{\ln n}}\sqrt{k},\quad k = 1,\dots,N, \end{gathered}\tag{2}$$

где

$$W_\beta(y) = \sup_{x\in S_n(1)}\{\langle y,x\rangle - \beta V(x)\} = \beta\ln\left(\frac{1}{n}\sum_{i=1}^{n}\exp(y_i/\beta)\right),$$

$$V(x) = \ln n + \sum_{i=1}^{n}x_i\ln x_i.$$

Рассуждая подобно [3], [4], [7], можно получить следующий результат.

**Теорема 1.** *Пусть справедливы условия* 1, 2, 3а, *тогда верно*

$$\frac{1}{N}\sum_{k=1}^{N}E[f_k(x^k;\xi^k)] - \min_{x\in S_n(1)}\frac{1}{N}\sum_{k=1}^{N}E_{\xi^k}[f_k(x;\xi^k)] \le 2M\sqrt{\frac{\ln n}{N}}.$$

*Если* $f_k \equiv f$, *а* $\{\xi^k\}$ *независимы и одинаково распределены, как* $\xi$, *то имеем*

$$E\left[f\left(\frac{1}{N}\sum_{k=1}^{N}x^k;\xi\right)\right] - \min_{x\in S_n(1)}E_\xi[f(x;\xi)] \le 2M\sqrt{\frac{\ln n}{N}}.$$

*Пусть справедливы условия* 1, 2, 3, *тогда при* $\Omega \ge 0$ *получим*

$$P_{x^1,\dots,x^N}\left\{\frac{1}{N}\sum_{k=1}^{N}E_{\xi^k}[f_k(x^k;\xi^k)] - \min_{x\in S_n(1)}\frac{1}{N}\sum_{k=1}^{N}E_{\xi^k}[f_k(x;\xi^k)] \ge \frac{2M}{\sqrt{N}}\left(\sqrt{\ln n} + \sqrt{8\Omega}\right)\right\} \le \exp(-\Omega).$$

*Если* $f_k \equiv f$, *а* $\{\xi^k\}$ *независимы и одинаково распределены, как* $\xi$, *то верно*

$$P_{x^1,\dots,x^N}\left\{E_\xi\left[f\left(\frac{1}{N}\sum_{k=1}^{N}x^k;\xi\right)\right] - \min_{x\in S_n(1)}E_\xi[f(x;\xi)] \ge \frac{2M}{\sqrt{N}}\left(\sqrt{\ln n} + \sqrt{8\Omega}\right)\right\} \le \exp(-\Omega).$$

**Замечание 2.** Запись

$$“P_{x^1,\ldots,x^N}\left\{\frac{1}{N}\sum_{k=1}^{N}E_{\xi^k}[f_k(x^k;\xi^k)]-\ldots”\right.$$

означает, что под вероятностью мы считаем математическое ожидание по $\xi^k$, которое, вообще говоря, зависит и от $\xi^1,\ldots,\xi^{k-1}$ (мы не предполагаем независимости $\{\xi^k\}$), как бы “замораживая” (считая не случайными) $x^k$, т.е. забывая про то, что $x^k$ тоже зависит от $\xi^1,\ldots,\xi^{k-1}$. А вероятность берется как раз по $\{x^k\}$ с учетом того, что такая зависимость есть (см. определение алгоритма МЗС1).

**Доказательство теоремы 1.** Возьмем за основу обозначения МЗС1, приведенные в (2). Рассуждая далее аналогично [3], [4], [7], получаем (заметим, что $\beta_{k+1}\ge\beta_k>0$)

$$\begin{aligned}
W_{\beta_{k+1}}(y^k) &\le W_{\beta_k}(y^k) = W_{\beta_k}(y^{k-1}) + \int_0^1 (y^k-y^{k-1})^T\nabla W_{\beta_k}(\tau y^k+(1-\tau)y^{k-1})d\tau = \\
&= W_{\beta_k}(y^{k-1}) - \gamma_k\nabla_x f_k(x^k;\xi^k)\nabla W_{\beta_k}(y^{k-1}) - \\
&- \gamma_k\nabla_x f_k(x^k;\xi^k)^T\int_0^1\left(\nabla W_{\beta_k}\left(\tau y^k+(1-\tau)y^{k-1}\right)-\nabla W_{\beta_k}(y^{k-1})\right)d\tau \overset{(*)}{\le} \\
&\overset{(*)}{\le} W_{\beta_k}(y^{k-1}) - \gamma_k\nabla_x f_k(x^k;\xi^k)\nabla W_{\beta_k}(y^{k-1}) + \\
&+ \gamma_k\left\|\nabla_x f_k(x^k;\xi^k)\right\|_\infty\int_0^1\left\|\nabla W_{\beta_k}\left(\tau y^k+(1-\tau)y^{k-1}\right)-\nabla W_{\beta_k}(y^{k-1})\right\|_1 d\tau \le \\
&\le W_{\beta_k}(y^{k-1}) - \gamma_k\nabla_x f_k(x^k;\xi^k)\nabla W_{\beta_k}(y^{k-1}) + \frac{\gamma_k^2\left\|\nabla_x f_k\left(x^k;\xi^k\right)\right\|_\infty^2}{2\alpha\beta_k},
\end{aligned}$$

последнее неравенство следует из того, что (см. [3], [4], [7])

$$\left\|\nabla W_\beta(\tilde{y})-\nabla W_\beta(y)\right\|_1 \le \frac{1}{\alpha\beta}\|\tilde{y}-y\|_\infty,$$

где $\alpha=1$ — константа сильной выпуклости $V(x)$ в 1-норме. Мы специально выделили (*) неравенство, которое иногда (например, в задачах о многоруких бандитах) бывает довольно грубым. В работах [5], [14] указан способ уточнения этого неравенства.

Суммируя эти неравенства, учитывая, что $\nabla W_{\beta_k}(y^{k-1})=x^k$ и формулу (2), получаем

$$\sum_{k=1}^{N}\gamma_k(x^k-x)^{\mathrm{T}}\nabla_x f_k(x^k;\xi^k) \le W_{\beta_1}(y^0)+x^{\mathrm{T}}y^N-W_{\beta_{N+1}}(y^N)+\sum_{k=1}^{N}\frac{\gamma_k^2}{2\alpha\beta_k}\left\|\nabla_x f_k(x^k;\xi^k)\right\|_\infty^2,$$

где $x^{\mathrm{T}}$ — означает транспонирование вектора $x$, который мы выбираем так, чтобы он доставлял решение задачи (1). Поскольку (см. [4])

$$W_{\beta_1}(y^0)=W_{\beta_1}(0)=0 \quad\text{и}\quad \beta_{N+1}V(x)\ge x^{\mathrm{T}}y^N-W_{\beta_{N+1}}(y^N),$$

то

$$\sum_{k=1}^{N}\gamma_k(x^k-x)^{\mathrm{T}}\nabla_x f_k(x^k;\xi^k) \le \beta_{N+1}V(x)+\sum_{k=1}^{N}\frac{\gamma_k^2}{2\alpha\beta_k}\left\|\nabla_x f_k(x^k;\xi^k)\right\|_\infty^2.$$

Тогда, из выпуклости функции $E_{\xi^k}[f_k(x;\xi^k)]$ по $x$ (в виду условия 1) следует, что

$$\sum_{k=1}^{N}\gamma_k\left\{E_{\xi^k}[f_k(x^k;\xi^k)]-E_{\xi^k}[f_k(x;\xi^k)]\right\}\le\sum_{k=1}^{N}\gamma_k(x^k-x)^{\mathrm{T}}\nabla_x E_{\xi^k}[f_k(x^k;\xi^k)]\le\beta_{N+1}V(x)-$$
$$-\sum_{k=1}^{N}\gamma_k(x^k-x)^{\mathrm{T}}\left(\nabla_x f_k(x^k;\xi^k)-\nabla_x E_{\xi^k}[f_k(x^k;\xi^k)]\right)+\sum_{k=1}^{N}\frac{\gamma_k^2}{2\alpha\beta_k}\left\|\nabla_x f_k(x^k;\xi^k)\right\|_\infty^2. \tag{3}$$

Возьмем полное (т.е., в отличие от замечания 2, с учетом зависимости $x^k$ от $\{\xi^1,\ldots,\xi^{k-1}\}$) математическое ожидание (в два шага $E[\cdot]=E\left[E_{\xi^k}[\cdot|\Xi^{k-1}]\right]$ – для каждого слагаемого свое $k$) от обеих частей неравенства, учитывая условие 3а и то, что

$$E\left((x^k-x)^{\mathrm{T}}\left(\nabla_x f_k(x^k;\xi^k)-\nabla_x E_{\xi^k}[f_k(x^k;\xi^k)]\right)\right)=$$
$$=E\left[(x^k-x)^{\mathrm{T}}E_{\xi^k}\left(\nabla_x f_k(x^k;\xi^k)-\nabla_x E_{\xi^k}[f_k(x^k;\xi^k)]\middle|\Xi^{k-1}\right)\right]=0,$$

поскольку $x^k$ есть $\Xi^{k-1}$-измеримый вектор и внутреннее условное математическое ожидание в силу условия 2 равно 0, получаем

$$\sum_{k=1}^{N}\gamma_k\left\{E[f_k(x^k;\xi^k)]-E_{\xi^k}[f_k(x;\xi^k)]\right\}\le\beta_{N+1}V(x)+$$
$$+\sum_{k=1}^{N}\frac{\gamma_k^2}{2\alpha\beta_k}E\left[\left\|\nabla_x f_k(x^k;\xi^k)\right\|_\infty^2\right]\le\beta_{N+1}R^2+M^2\sum_{k=1}^{N}\frac{\gamma_k^2}{2\alpha\beta_k},$$

где $R^2=\max\limits_{x\in S_n(1)}V(x)=\ln n$. Подставляя $\gamma_k\equiv 1$ и минимизируя правую часть неравенства по неубывающим последовательностям с положительными элементами $\{\beta_k\}_{k=1}^{N+1}$, не допуская при этом зависимость $\{\beta_k\}_{k=1}^{N+1}$ от потенциально неизвестного $N$, получаем

$$\beta_k=\sqrt{\frac{M^2}{\alpha R^2}}\sqrt{k},$$
$$\frac{1}{N}\sum_{k=1}^{N}E[f_k(x^k;\xi^k)]-\min_{x\in S_n(1)}\frac{1}{N}\sum_{k=1}^{N}E_{\xi^k}[f_k(x;\xi^k)]\le 2\sqrt{\frac{M^2R^2}{\alpha N}}.$$

Для того чтобы доказать первую часть теоремы, осталось подставить $\alpha=1$ и $R^2=\ln n$.

**Замечание 3.** Если разрешать $\{\beta_k\}_{k=1}^{N+1}$ зависеть от $N$ (см., например ниже, алгоритм МЗС2-неадаптивный), то в последней формуле “2”-у можно занести под знак корня (см. [4]). Все это переносится и на последующие рассуждения с вероятностями больших отклонений.

**Замечание 4.** Строго говоря, в полученной оценке в знаменателе вместо $N$ нужно писать $N^2/(N+1)$. Считая, что $N\gg 1$, мы пренебрегаем этим для компактности записи.

Для доказательства второй части теоремы вернемся к формуле (3). Из условия 3 имеем

$$P\left(\sum_{k=1}^{N}\frac{\gamma_k^2}{2\alpha\beta_k}\left\|\nabla_x f_k(x^k;\xi^k)\right\|_\infty^2>M^2\sum_{k=1}^{N}\frac{\gamma_k^2}{2\alpha\beta_k}\right)=0. \tag{4}$$

Из неравенства Азума–Хеффдинга (см. [17]), подобно [7], [18], получаем для ограниченной мартингал–разности

$$\left|(x-x^k)^{\mathrm{T}}\left(\nabla_x f_k(x^k;\xi^k)-\nabla_x E_{\xi^k}[f_k(x^k;\xi^k)]\right)\right|\le 4M$$

следующее неравенство:

$$P\left(\sum_{k=1}^{N}\gamma_k(x-x^k)^{\mathrm{T}}\left(\nabla_x f_k(x^k;\xi^k)-\nabla_x E_{\xi^k}[f_k(x^k;\xi^k)]\right)\geq 4M\Lambda\sqrt{\sum_{k=1}^{N}\gamma_k^2}\right)\leq \exp(-\Lambda^2/2).$$

Подставляя $\gamma_k\equiv 1$, $\Lambda=\sqrt{2\Omega}$, получаем вторую часть теоремы.

**Замечание 5.** В [7] для задачи стохастической выпуклой оптимизации ($\{\xi^k\}$ – независимые случайные величины) приводится оценка вероятностей больших уклонений с точностью до констант, аналогичная оценке, приведенной в теореме. Аналогичная оценка приводится в [7] и для случая, когда вместо условия 3 предполагается условие 3б, при этом в правой части неравенства под вероятностью вместо $\sqrt{\Omega}$ необходимо писать $\Omega$ (с точностью до констант). При этом в [7] использовалось условие независимости $\{\xi^k\}$ (при установлении неравенства типа (4) в общем случае и в неравенстве Азума–Хеффдинга). Для онлайн оптимизации, как правило, независимость $\{\xi^k\}$ места не имеет. Тем не менее, условия 2, 3б обеспечивают выполнение этих неравенств в том же виде, как если бы независимость $\{\xi^k\}$ имела место. Приведем теперь оценки в случае тяжелых хвостов. Если $\|\nabla_x f(x,\xi)\|_\infty^2$ имеет степенной хвост $\alpha>2$

$$P\left(\frac{\|\nabla_x f(x,\xi)\|_\infty^2}{M^2}\geq t\right)=O\left(\frac{1}{t^\alpha}\right),$$

то существует такая константа $C_\alpha>0$, что с вероятностью $\geq 1-\sigma$ имеем

$$\frac{1}{N}\sum_{k=1}^{N}E_{\xi^k}[f_k(x^k;\xi^k)]-\min_{x\in S_n(1)}\frac{1}{N}\sum_{k=1}^{N}E_{\xi^k}[f_k(x;\xi^k)]\leq C_\alpha M\frac{\sqrt{N\ln n}+(N/\sigma)^{1/\alpha}}{N}.$$

Если мы не делаем никаких предположений относительно распределений случайных величин $\{\xi^k\}$, кроме 1, 2 и существования первых двух равномерно ограниченных моментов у $\nabla_x f_k\left(x^k;\xi^k\right)$, то из неравенства Маркова и первого неравенства в теореме 1 (на математические ожидания) имеем следующее: существует такая константа $C>0$, что с вероятностью $\geq 1-\sigma$ получим

$$\frac{1}{N}\sum_{k=1}^{N}E_{\xi^k}[f_k(x^k;\xi^k)]-\min_{x\in S_n(1)}\frac{1}{N}\sum_{k=1}^{N}E_{\xi^k}[f_k(x;\xi^k)]\leq \frac{CM}{\sigma}\sqrt{\frac{\ln n}{N}}.$$

Труднее обстоит дело, если мы хотим оценить псевдо регрет или

$$\frac{1}{N}\sum_{k=1}^{N}f_k(x^k;\xi^k)-\min_{x\in S_n(1)}\frac{1}{N}\sum_{k=1}^{N}f_k(x;\xi^k).$$

Тем не менее, при дополнительных оговорках и такие выражения можно вероятностно оценивать (см. [14], [18]).

Аналогичное замечание имеет место и для теоремы 2 ниже.

## 3. ОНЛАЙН МЗС СО СТОХАСТИЧЕСКОЙ ПРОЕКЦИЕЙ ГРАДИЕНТА

Снова рассмотрим постановку задачи стохастической онлайн оптимизации (1). На этот раз будем допускать, что метод генерирования последовательности $\{x^k\}$ может допускать (внешнюю, дополнительную) рандомизацию. Это допущение позволит частично перенести результаты разд. 2 на не выпуклые функции $E_{\xi^k}\left[f_k\left(x;\xi^k\right)\right]$ (см. ниже пример 4), на ситуации, когда по условию задачи $\{x^k\}$ должны выбираться среди вершин единичного симплекса (примеры 1 и 5). Также как и раньше онлайновость постановки задачи допускает, что на каждом шаге $k$ функция $f_k$ может подбираться из рассматриваемого класса функций враждебно по отношению к используемому нами методу генерации последовательности $\{x^k\}$. В частности, $f_k$ и $\xi^k$ могут зависеть от $\left\{x^1,\xi^1,f_1(\cdot);\ldots;x^{k-1},\xi^{k-1},f_{k-1}(\cdot)\right\}$, и даже от распределения вероятностей $p^k$ (многорукие бандиты), согласно которому осуществляется выбор $x^k$. Чтобы можно было работать с таким классом задач, нам придется наложить дополнительное

**Условие 4.** На каждом шаге генерирование случайной величины $x^k$ согласно распределению вероятностей $p^k$ осуществляется независимо ни от чего.

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

**Алгоритм МЗС2-адаптивный. Метод двойственных усреднений**

Согласно распределению вероятностей

$$p_i^{t+1} = \frac{\exp\left(-\frac{1}{\beta_{t+1}}\sum_{k=1}^{t}\frac{\partial f_k(x^k;\xi^k)}{\partial x_i}\right)}{\sum_{l=1}^{n}\exp\left(-\frac{1}{\beta_{t+1}}\sum_{k=1}^{t}\frac{\partial f_k(x^k;\xi^k)}{\partial x_l}\right)},\quad i = 1,\ldots,n,\quad \beta_t = \frac{M\sqrt{t}}{\sqrt{\ln n}},$$

получаем случайную величину $i(t+1)$, $x_{i(t+1)}^{t+1} = 1$, $x_j^{t+1} = 0$, $j \ne i(t+1)$.

**Алгоритм МЗС2-неадаптивный** (заранее известно $N$)

Согласно распределению вероятностей

$$p_i^{t+1} = \frac{\exp\left(-\frac{1}{\beta_{t+1}}\sum_{k=1}^{t}\gamma_k\frac{\partial f_k(x^k;\xi^k)}{\partial x_i}\right)}{\sum_{l=1}^{n}\exp\left(-\frac{1}{\beta_{t+1}}\sum_{k=1}^{t}\gamma_k\frac{\partial f_k(x^k;\xi^k)}{\partial x_l}\right)},\quad i = 1,\ldots,n,$$

$$\gamma_k \equiv M^{-1}\sqrt{2\ln n/N},\quad \beta_t \equiv 1,$$

получаем случайную величину $i(t+1)$, $x_{i(t+1)}^{t+1} = 1$, $x_j^{t+1} = 0$, $j \ne i(t+1)$.

**Мотивация** (ограничимся детерминированным случаем, $\gamma_k \equiv 1$). Аппроксимируя

$$\min_{x\in S_n(1)}\frac{1}{N}\sum_{k=1}^{t}f_k(x) \approx \min_{x\in S_n(1)}\frac{1}{N}\sum_{k=1}^{t}\left\{f_k(x^k) + \left\langle\nabla f_k(x^k), x - x^k\right\rangle\right\},$$

получаем

$$P\left(x_j^{t+1} = 1; x_i^{t+1} = 0, i \ne j\right) \overset{\text{def}}{=} P_\varsigma\left(j = \arg\max_{i=1,\ldots,n}\left\{\sum_{k=1}^{t}[-\nabla f_k(x^k)]_i + \varsigma_{t,i}\right\}\right),$$

где $\varsigma_{t,i}$ — независимые одинаково распределенные случайные величины по закону Гумбеля с параметром $\beta_{t+1}$, характеризующим среднеквадратичное отклонение $\varsigma_{t,i}$:

$$P(\varsigma_{t,i} < \tau) = \exp\{-e^{-\tau/\beta_{t+1}}\}.$$

**Замечание 6.** Поскольку случайные величины $\varsigma_{t,i}$ получаются в результате суммирования $t$ слагаемых (невязок в аппроксимации выпуклой функции линейными минорантами), то можно ожидать, что среднеквадратичное отклонение $\varsigma_{t,i}$ имеет порядок $\sqrt{t}$. Условие одинаковой распределенности $\varsigma_{t,i}$, которое не понятно за счет чего может иметь место, в действительности, не очень здесь и нужно. Достаточно, чтобы $E[\varsigma_{t,i} - \varsigma_{t,j}] = o(\sqrt{t})$ и $\mathrm{Var}[\varsigma_{t,i}] \sim \sqrt{t}$. Более того, если, с некоторой натяжкой, наряду с независимостью $\{\varsigma_{t,i}\}_{i=1}^n$ также считать, что при формировании $\varsigma_{t,i}$ суммируются независимые слагаемые, то можно ожидать, что $\varsigma_{t,i}$ нормальные случайные величины (центральная предельная теорема). Следовательно, в контексте последующих операций при $n \gg 1$, $\varsigma_{t,i}$ могут быть заменены на случайные величины, распределенные по закону Гумбеля [19], [20] с соответствующими среднеквадратичными отклонениями.

Тогда (см. [3], [9], [19]) имеем

$$E_{\varsigma}[x^{t+1}] = \nabla W_{\beta_{t+1}}\left(-\sum_{k=1}^{t}\nabla f_k(x^k)\right).$$

Распределение Гумбеля возникает именно в таком контексте совсем не случайно, и связано это с тем, что оно max-устойчиво (см. [20]). Пусть $\varsigma_1,\ldots,\varsigma_n$ – независимые одинаково распределенные случайные величины, и существуют такие константы $\alpha,\beta > 0$, что

$$\lim_{y\to\infty} e^{y/\beta}[1 - P(\varsigma_k < y)] = \alpha,$$

тогда (для $\alpha = 0$ при незначительных оговорках также будет сходимость к распределению Гумбеля, но формула будет немного другой)

$$\max\{\varsigma_1,\ldots,\varsigma_n\} - \beta\ln(\alpha n) \xrightarrow[n\to\infty]{d} \xi, \quad \text{где} \quad P(\xi < \tau) = \exp\{-e^{-\tau/\beta}\}.$$

Другими словами, при некоторых довольно общих предположениях (типа Крамера) $\varsigma_{t,i}$ могут быть распределенными произвольно, тем не менее, при $n \gg 1$ с хорошей точностью можно считать, что мы в итоге имеем дело с соответствующим распределением Гумбеля. Более того, если задаться вопросом: а какое распределение "наиболее подходит" для $\varsigma_{t,i}$, чтобы в случае "враждебной Природы" (т.е. в минимаксном смысле) иметь наилучшие оценки, то ответом будет: симметричное показательное распределение [9] (Лапласа), которое ведет себя в интересном для анализа диапазоне подобно распределению Гумбеля, но в отличие от распределения Лапласа, в случае Гумбеля мы явно можем посчитать интересующие нас вероятности. Как правило, такого рода задачи явно не решаются, и распределение Гумбеля является приятным исключением, для которого есть явные формулы.

Приведенная выше мотивация имеет одно интересное приложение в содержательной интерпретации равновесного распределения транспортных потоков. Немного об этом написано в [21]. Планируется посвятить этому приложению отдельную публикацию.

К сожалению, не делая относительно функций $f_k(x;\xi^k)$ дополнительно никаких предположений, не удается доказать для МЗС2 аналог теоремы 1. Чтобы можно было сформулировать такой аналог, мы вынуждены будем предполагать, что $f_k(x;\xi^k)$ – линейные функции по $x$ (можно обобщить и на сублинейные). С одной стороны, это существенно сужает класс задач, к которым применим МЗС2. С другой стороны, как будет продемонстрировано ниже, даже такой узкий класс функций за счет "онлайновости" позволяет применять МЗС2 к довольно широкому кругу задач. Для того чтобы лучше чувствовалась преемственность методов и доказательств их сходимости, далее мы по-прежнему будем использовать общие обозначения $f_k(x;\xi^k)$, не подчеркивая в формулах линейность.

**Теорема 2.** *Пусть справедливы условия* 1, 2, 3а, 4 *и* $f_k(x;\xi^k)$ *– линейные функции по* $x$*, тогда имеем*

$$\frac{1}{N}\sum_{k=1}^{N} E[f_k(x^k;\xi^k)] - \min_{x\in S_n(1)}\frac{1}{N}\sum_{k=1}^{N} E_{\xi^k}[f_k(x;\xi^k)] \le 2M\sqrt{\frac{\ln n}{N}}.$$

*Для неадаптивного метода* "2"*-у перед* $M$ *можно занести под знак корня.*

*Кроме того, если справедливы условия* 1, 2, 3, 4, *то при* $\Omega \ge 0$ *получаем*

$$P_{x^1,\ldots,x^N}\left\{\frac{1}{N}\sum_{k=1}^{N} E_{\xi^k}[f_k(x^k;\xi^k)] - \min_{x\in S_n(1)}\frac{1}{N}\sum_{k=1}^{N} E_{\xi^k}[f_k(x;\xi^k)] \ge \frac{2M}{\sqrt{N}}\left(\sqrt{\ln n} + \sqrt{18\Omega}\right)\right\} \le \exp(-\Omega).$$

*Если* $f_k(x;\xi^k) \equiv f_k(x)$*, то это неравенство можно уточнить:*

$$\frac{2M}{\sqrt{N}}\left(\sqrt{\ln n} + \sqrt{18\Omega}\right) \to \frac{2M}{\sqrt{N}}\left(\sqrt{\ln n} + \sqrt{2\Omega}\right),$$

*при этом же условии для неадаптивного метода можно еще больше уточнить:*

$$\frac{2M}{\sqrt{N}}\left(\sqrt{\ln n} + \sqrt{2\Omega}\right) \to \frac{\sqrt{2}M}{\sqrt{N}}\left(\sqrt{\ln n} + 2\sqrt{\Omega}\right).$$

**Схема доказательства теоремы 2.** Доказательство фактически дословно повторяет доказательство теоремы 1. Небольшая разница лишь в том, что основная формула (3) перепишется следующим образом:

$$\sum_{k=1}^{N}\gamma_k\left\{E_{\xi^k}[f_k(x^k;\xi^k)]-E_{\xi^k}[f_k(x;\xi^k)]\right\}\le\sum_{k=1}^{N}\gamma_k(x^k-x)^{\mathrm{T}}\nabla_x E_{\xi^k}[f_k(x^k;\xi^k)]\le$$

$$\le\beta_{N+1}V(x)+\sum_{k=1}^{N}\gamma_k(x^k-E_{x^k}[x^k])^{\mathrm{T}}\nabla_x f_k(x^k;\xi^k)-$$

$$-\sum_{k=1}^{N}\gamma_k(x^k-x)^{\mathrm{T}}\left(\nabla_x f_k(x^k;\xi^k)-\nabla_x E_{\xi^k}[f_k(x^k;\xi^k)]\right)+\sum_{k=1}^{N}\frac{\gamma_k^2}{2\alpha\beta_k}\left\|\nabla_x f_k(x^k;\xi^k)\right\|_\infty^2.$$

Здесь мы просто по-другому (по сравнению с доказательством теоремы 1) переписали неравенство

$$\sum_{k=1}^{N}\gamma_k(x^k-x)^{\mathrm{T}}f_k(x^k;\xi^k)\le\beta_{N+1}V(x)+\sum_{k=1}^{N}\frac{\gamma_k^2}{2\alpha\beta_k}\left\|\nabla_x f_k(x^k;\xi^k)\right\|_\infty^2,$$

использовав выпуклость функции $E_{\xi^k}[f_k(x;\xi^k)]$ по $x$ в виду условия 1.

Введем случайные величины

$$Z_k=(x^k-E_{x^k}[x^k])^{\mathrm{T}}\nabla_x f_k(x^k;\xi^k),\quad \tilde{Z}_k=(x^k-x)^{\mathrm{T}}\left(\nabla_x f_k(x^k;\xi^k)-\nabla_x E_{\xi^k}[f_k(x^k;\xi^k)]\right).$$

Поскольку $f_k(x^k;\xi^k)=l_k(\xi^k)^{\mathrm{T}}x^k$ – линейные функции, то по условиям 4, 2 имеем

$$E_{x^k}[Z_k|\Xi^{k-1}]\equiv 0,\quad E_{\xi^k}[\tilde{Z}_k|\Xi^{k-1}]\equiv 0.$$

Именно в этом месте и только в нем используется линейность $f_k(x;\xi^k)$. К сожалению, предложенный здесь способ рассуждения не позволяет хоть сколько-нибудь ослабить это условие.

Рассуждая далее также как в доказательстве теоремы 1, получаем теорему 2.

**Замечание 7.** Как уже отмечалось во введении, теорема 2 во многом мотивирована работой [6]. Собственно, форма, в который мы представили алгоритм МЗС2-неадаптивный, выбрана именно такой (альтернативным вариантом было положить $\gamma_k\equiv 1$, $\beta_k\equiv M\sqrt{N/(2\ln n)}$), чтобы была максимальная близость к алгоритму работы [6].

**Замечание 8.** Идея рандомизации (искусственного введения случайности), положенная в основу описанных алгоритмов, чрезвычайно продуктивна: против нас играет, возможно, враждебная “Природа”, которая, зная историю игры, и наши текущие намерения старается нам “предложить вариант похуже”. С этим можно “бороться” за счет случайного независимого осуществления своих намерений на каждом шаге, с реализацией неизвестной “Природе”. За счет этой случайности мы переходим от анализа по худшему случаю (роль которого в онлайн оптимизации играет враждебная “Природа”) к анализу “в среднем”. Такая рандомизация, как будет отмечено ниже, дает возможность получать оценки, которые в детерминированном случае получить невозможно. Причем, если в онлайн постановке такая рандомизация прописывается в “правилах игры”, то применительно к задачам обычной оптимизации все это возникает совершенно естественным образом, как желание с большой вероятностью обезопасить себя от “самых худших случаев” детерминированной версии метода. Отметим, что речь идет о, так называемых, массовых задачах, т.е., исследуя тот или иной метод, мы точно не знаем какой конкретно объект поступит на вход, поэтому, чтобы гарантированно что-то иметь, мы исходим из худшего (наименее благоприятного для данного метода) случая входных данных. Описанный МЗС2 естественно также понимать как покомпонентный метод (стохастического) субградиентного спуска со случайным выбором компоненты. Происходит рандомизация при проектировании (в смысле расстояния Брэгмана) на единичный симплекс. А именно, если спроектироваться в указанном выше смысле на единичный симплекс, то получится вектор, который можно проинтерпретировать как распределение вероятностей некоторой дискретной случайной величины, принимающей значения $1,\ldots,n$. Если выбрать вершину симплекса согласно этой дискретной случайной величине, и заменить проекцию этой вершиной, то получим случайную проекцию, математическое ожидание которой равно честной проекции. Как будет отмечено ниже, такой метод не только оптимален с точки зрения числа итераций, но и в некотором смысле с точки зрения затрат на выполнение одной итерации (см. также [22]).

**Замечание 9** (Ф.А. Федоренко). Если ограничение $x \in S_n(1)$ в задаче (1) заменить на ограничение

$$x = \left(z^1, \dots, z^m\right) \in \prod_{j=1}^m S_{n_j}\left(d_j\right),$$

типичное для популяционных игр (с несколькими популяциями), в частности, для транспортных приложений (см. [21]), то, взяв

$$V(x) = \sum_{j=1}^m V_j(z^j), \quad V_j(z^j) = d_j \ln n_j + \sum_{i=1}^{n_j} z_i^j \ln\left(\frac{z_i^j}{d_j}\right), \quad W_\beta(y) = \sup_{z_j \in S_{n_j}(d_j)} \left\{\sum_{j=1}^m \left\langle v^j, z^j\right\rangle - \beta V(x)\right\},$$

где $y = \left(v^1, \dots, v^m\right)$, можно получить аналогичные теореме 1 оценки с заменой (аналогично по теореме 2)

$$2M\sqrt{\frac{\ln n}{N}} \to 2M\sqrt{\frac{\left(\sum_{j=1}^m d_j\right)\left(\sum_{j=1}^m d_j \ln n_j\right)}{N}},$$

$$\frac{2M}{\sqrt{N}}\left(\sqrt{\ln n} + \sqrt{8\Omega}\right) \to \frac{2M\sum_{j=1}^m d_j}{\sqrt{N}}\left(\sqrt{\frac{\sum_{j=1}^m d_j \ln n_j}{\sum_{j=1}^m d_j}} + \sqrt{8\Omega}\right).$$

## 4. ПРИЛОЖЕНИЯ МЗС

В заключительном разделе мы постараемся продемонстрировать некоторые возможности и ограничения описанных выше методов. Мы не будем стремиться здесь к максимальной общности или рассмотрению всех основных приложений МЗС.

**Пример 1** (многорукие бандиты, см. [5], [9], [14]). Имеется $n$ различных ручек. Игра повторяется $N \gg 1$ раз (это число может быть заранее неизвестно). На каждом шаге $k$ мы должны выбрать ручку $i(k)$, которую “дергаем”. Дергание ручки приносит нам некоторые, вообще говоря, случайные потери $r_{i(k)}^k$ (считаем, для определенности, что всегда $r_{i(k)}^k \in [0,1]$), зависящие от номера шага, номера ручки и от того, какой стратегии мы придерживались до шага $k$ включительно. Наша стратегия на шаге $k$ описывается вектором распределения вероятностей $x^k \in S_n(1)$, согласно которому мы независимо ни от чего выбираем ручку, которую будем дергать. Все, чем мы располагаем на шаге $k$, — вектор

$$\left(\left(x^1, i(1), r_{i(1)}^1\right); \dots; \left(x^{k-1}, i(k-1), r_{i(k-1)}^{k-1}\right)\right).$$

Мы считаем, что потери на $k$-м шаге $r^k$ зависят от $x^k$ (но не от результата разыгрывания из распределения $x^k$), зависят от $(x^1, \dots, x^{k-1})$ и результатов соответствующих разыгрываний, а также зависят от $(r^1, \dots, r^{k-1})$. Целью является организация процедуры дергания ручек таким образом, чтобы ожидаемые суммарные потери были бы минимальны. Введем функцию ($r^k$ и результат разыгрывания, согласно распределению вероятностей, заданному вектором $x$, независимы; обе эти “случайности” мы обозначаем $\xi^k$)

$$f_k(x; \xi^k) = r_i^k \quad \text{с вероятностью } x_i, \quad i = 1, \dots, n,$$

и ее обобщенный (в смысле удовлетворения условию 2) стохастический градиент

$$\nabla_x f_k(x; \xi^k) = (\underbrace{0, \dots, r_i^k}_{i}/x_i, \dots, 0)^{\mathrm{T}} \quad \text{с вероятностью } x_i, \quad i = 1, \dots, n.$$

Тогда выполнены условия 1 и 2. Однако имеется проблема: константа $M$ в условии 3 получается слишком большой (например, в 3а $M = \sup_{x \in S_n(1)} \sqrt{\Sigma_{i=1}^n x_i^{-1}} = \infty$), т.е. теорема 1 ничего дать не может в том виде, в котором она была нами приведена. Возникает желание “что-то подкрутить” в доказательстве теоремы, чтобы можно было ей воспользоваться. Уже по ходу самого доказательства мы отмечали неравенство (*), которое может оказаться довольно грубым в определенных ситуациях. Многорукие бандиты дают пример как раз такой ситуации. Более аккуратный анализ (см. [5], [10], [14]), использующий специфику данной задачи, позволяет оценить (см. доказательство теоремы 1)

$$\gamma_k \nabla_x f_k(x^k;\xi^k)^{\mathrm{T}} \int_0^1 \left(\nabla W_{\beta_k}\left(\tau y^k + (1-\tau) y^{k-1}\right) - \nabla W_{\beta_k}(y^{k-1})\right) d\tau$$

точнее, что приводит в основной формуле (3) к замене слагаемых вида

$$\frac{\gamma_k^2}{2\alpha\beta_k}\left\|\nabla_x f_k(x^k;\xi^k)\right\|_\infty^2$$

на (в этом месте для наглядности мы намеренно несколько упрощаем и огрубляем)

$$\gamma_k^2 \frac{x_j^k(1-x_j^k)}{\alpha\beta_k}\left(\frac{r_j^k}{x_j^k}\right)^2,$$

где $j$ — номер ручки, выбранной алгоритмом на $k$-м шаге. В результате мы получаем, что теорема 1 остается верной с эффективной константой $M = \sqrt{2n}$. Таким образом, действуя согласно МЗС1, наши потери (псевдо регрет, см. [14]) будут следующими:

$$O\left(\sqrt{\frac{n\ln n}{N}}\right) \text{ в среднем; } \quad O\left(\sqrt{\frac{n\ln(n/\sigma)}{N}}\right) \text{ с вероятностью } \ge 1-\sigma,$$

что с точностью до логарифмического фактора соответствует нижним оценкам [5], [9], [14].

**Замечание 10.** Стоит обратить внимание, что если использовать более специальную прокс-структуру (см. [5], [14]), то для псевдо регрета можно получить оценки без логарифмического фактора $\ln n$ под корнем, что уже соответствует нижним оценкам. В частности, это обстоятельство означает, что выбирать “расстояние” Брэгмана для симплекса не всегда оптимально (но близко к оптимуму). Тем не менее, в последующих нескольких примерах мы убедимся, что для ряда других постановок оценки, полученные с помощью прокс-структуры, порожденной расстоянием Брэгмана, являются оптимальными. Кроме того, есть еще плата за избавление от фактора $\ln n$ под корнем — удорожание процедуры вычисления проекции на симплекс в смысле этой прокс-структуры. Другими словами, это ускорение оправдано только для онлайн постановок, в которых, как правило, стремятся минимизировать (псевдо) регрет, не учитывая общую вычислительную трудоемкость.

Труднее обстоит дело с оценкой регрета (в среднем и с вероятностями больших уклонений). Пример из лекции 6 (см. [11]) показывает ($n = 2$), что МЗС1 может давать регрет $\sim cN^{-1/4}$, что значительно хуже оценки псевдо регрета $\sim cN^{-1/2}$. По сути, речь идет о том, что написано в конце замечания 1. Здесь уже требуется некая игра “bias–variance trade off”: отказаться от несмещенности оценки градиента для уменьшения дисперсии этой оценки. Этот популярный трюк в математической статистике и машинном обучении позволяет с некоторыми оговорками распространить приведенные выше оценки псевдо регрета $O\left(\sqrt{n\ln(n/\sigma)/N}\right)$ и на случай оценок регрета. Кое-что на эту тему применительно к многоруким бандитам можно найти в обзоре [14].

Интересно заметить, что результаты, описанные в примере 1, можно получить (с аналогичными оговорками) с помощью МЗС2 и теоремы 2. Для этого нужно взять

$$f_k\left(x;\xi^k\right) = \left\langle r^k, x\right\rangle$$

и ее обобщенный (в смысле удовлетворения условию 2) стохастический градиент

$$\nabla_x f_k(x;\xi^k) = (\underbrace{0,\ldots,r_i^k}_{i}/p_i,\ldots,0)^{\mathrm{T}}, \quad \text{если } x = (\underbrace{0,\ldots,1}_{i},\ldots,0)^{\mathrm{T}},$$

$$\text{где } x = (\underbrace{0,\ldots,1}_{i},\ldots,0)^{\mathrm{T}} \text{ с вероятностью } p_i,\ i = 1,\ldots,n;$$

здесь $\xi^k$ отражает только случайность, сидящую в $r^k$. Это определение стохастического градиента (в отличие от рассмотренного выше) явно учитывает распределение вероятностей $p$, из которого генерируется номер единственной ненулевой (единичной) компоненты вектора $x$. Также как и раньше для выполнения условия 2 необходимо предполагать независимость $\xi^k$ и процедуры разыгрывания единичной компоненты вектора $x$ согласно распределению $p$.

**Пример 2** (взвешивание экспертных решений, линейные потери, см. [9], [23]). Рассмотрим задачу взвешивания экспертных решений, следуя [9], [23]. Имеется $n$ различных экспертов. Каждый эксперт играет на рынке. Игра повторяется $N \gg 1$ раз (это число может быть заранее неизвестно). Пусть $l_i^k$ – проигрыш эксперта $i$ на шаге $k$ ($\left|l_i^k\right| \le M$). На каждом шаге $k$ мы распределяем один доллар между экспертами, согласно вектору $x^k \in S_n(1)$. Потери, которые мы при этом несем, рассчитываются по потерям экспертов $\left\langle l^k, x^k \right\rangle$. Целью является организация процедуры распределения доллара на каждом шаге таким образом, чтобы наши суммарные потери были бы минимальны. Допускается, что потери экспертов $l^k$ могут зависеть еще и от текущего хода $x^k$. Легко проверить, что для данной постановки применима теорема 1 в детерминированном варианте с функциями

$$f_k\left(x;\xi^k\right) \equiv f_k(x) = \left\langle l^k, x \right\rangle.$$

При этом оценка, даваемая теоремой 1,

$$O\left(M\sqrt{\frac{\ln n}{N}}\right)$$

является оптимальной для данного класса задач (см. [9], [23]).

**Пример 3** (взвешивание экспертных решений, выпуклые потери, см. [9], [23]). В условиях предыдущего примера предположим, что на $k$-м шаге $i$-й эксперт использует стратегию $\zeta_i^k \in \Delta$ (множество $\Delta$ выпуклое), дающую потери $\lambda(\omega^k, \zeta_i^k)$, где $\omega^k$ – “ход”, возможно, враждебной “Природы”, знающей, в том числе, и нашу текущую стратегию. Функция $\lambda(\cdot)$ выпуклая по второму аргументу и $|\lambda(\cdot)| \le M$. На каждом шаге мы должны выбирать свою стратегию

$$x \overset{\text{def}}{=} \sum_{i=1}^{n} x_i \cdot \zeta_i^k \in \Delta,$$

дающую потери $\lambda(\omega^k, x)$ так, чтобы наши суммарные потери были минимальны. Для данной постановки также применима теорема в детерминированном варианте с

$$f_k(x;\xi^k) \equiv f_k(x) = \sum_{i=1}^{n} x_i \lambda(\omega^k, \zeta_i^k) \ge \lambda(\omega^k, x).$$

Чтобы применить теорему, осталось заметить, что функция $\lambda(\omega^k, \zeta)$ является выпуклой по $\zeta$ для любого $\omega^k$, поэтому

$$\sum_{k=1}^{N} \lambda(\omega^k, x^k) - \min_{i=1,\ldots,n} \sum_{k=1}^{N} \lambda(\omega^k, \zeta_i^k) \le \sum_{k=1}^{N} f_k(x^k) - \min_{x \in S_n(1)} \sum_{k=1}^{N} f_k(x).$$

При этом оценка, даваемая теоремой 1,

$$O\left(M\sqrt{\frac{\ln n}{N}}\right)$$

также является оптимальной для данного класса задач (см. [9], [23]).

**Пример 4** (взвешивание экспертных решений, невыпуклые потери, см. [9], [23]). Предположим, что в условиях примера 3 мы не можем гарантировать выпуклость $\lambda(\cdot)$ по второму аргументу. Тогда мы выбираем стратегию – распределение вероятностей на множестве стратегий экспертов, и разыгрываем случайную величину согласно этому распределению вероятностей. Другими словами, мы просто пользуемся МЗС2 с

$$f_k(x;\xi^k) \equiv f_k(x) = \sum_{i=1}^{n} x_i \lambda(\omega^k, \zeta_i^k),$$

применимость которого обосновывается теоремой 2, с оценками

$$O\left(M\sqrt{\frac{\ln n}{N}}\right) \text{ в среднем;} \quad O\left(M\sqrt{\frac{\ln(n/\sigma)}{N}}\right) \text{ с вероятностью } \geq 1-\sigma,$$

оптимальными для данного класса задач (см. [9], [23]). Ключевая разница в примерах 1 и 4, "сто́ящая" $\sim\sqrt{n}$ в оценке $M$ для многоруких бандитов (пример 1), заключается в том, что в многоруких бандитах мы имеем только свою историю дергания ручек (нам не известно, какие бы потери нам принесли другие ручки, если бы мы их выбрали), а в постановке взвешивания экспертных решений это все известно, и называется *потерями экспертов*.

Как будет видно из следующего примера, описанный только что подход вполне успешно работает (дает не улучшаемые результаты) и в случае выпуклой по второму аргументу функции $\lambda(\cdot)$.

**Пример 5** (антагонистические матричные игры, см. [6], [9], [22], [23]). Возьмем два игрока А и Б. Задана матрица игры $A = \|a_{ij}\|$, где $|a_{ij}| \leq M$, $a_{ij}$ – выигрыш игрока А (проигрыш игрока Б) в случае, когда игрок А выбрал стратегию $i$, а игрок Б – стратегию $j$. Отождествим себя с игроком Б, и предположим, что игра повторяется $N \gg 1$ раз (это число может быть заранее неизвестно). Мы находимся в условиях примера 4 с $\lambda(\omega^k, \zeta_j^k) = \Sigma_{i=1}^n \omega_i^k a_{ij}$, т.е.

$$f_k(x) = \langle \omega^k, Ax \rangle, \quad x \in S_n(1),$$

где $\omega^k$ – вектор (вообще говоря, зависящий от всей истории игры до текущего момента включительно, в частности, как-то зависящий и от текущей стратегии (не хода) игрока Б, заданной распределением вероятностей (результат текущего разыгрывания (ход Б) игроку А не известен)) со всеми компонентами, равными 0, кроме одной компоненты, соответствующей ходу А на шаге $k$, равной 1. Хотя функция $f_k(x)$ определена на единичном симплексе, по "правилам игры" вектор $x^k$ имеет ровно одну единичную компоненту, соответствующую ходу Б на шаге $k$, остальные компоненты равны нулю. Обозначим цену игры

$$C = \max_{\omega \in S_n(1)} \min_{x \in S_n(1)} \langle \omega, Ax \rangle = \min_{x \in S_n(1)} \max_{\omega \in S_n(1)} \langle \omega, Ax \rangle \quad \text{(теорема фон Неймана о минимаксе).}$$

**Замечание 11.** Отметим, что с помощью онлайн оптимизации и экспоненциального взвешивания можно похожим образом проинтерпретировать и вариант теоремы о минимаксе для векторнозначной функции выигрыша – теорему Блэкуэлла о достижимости (см. [9], [23]), которая используется, например, при построении калибруемых предсказаний.

Пару векторов $(\omega, x)$, доставляющих решение этой минимаксной задачи (т.е. седловую точку), назовем *равновесием* Нэша. По определению (это неравенство восходит к Ханнану [9], [23]):

$$\min_{x \in S_n(1)} \frac{1}{N} \sum_{k=1}^{N} f_k(x) \leq C.$$

Тогда если мы (игрок Б) будем придерживаться рандомизированной стратегии МЗС2, выбирая $\{x^k\}$, то по теореме 2 с вероятностью $\geq 1-\sigma$ (в случае когда $N$ заранее известно, оценку можно уточнить) получаем

$$\frac{1}{N}\sum_{k=1}^{N} f_k(x^k) - \min_{x \in S_n(1)} \frac{1}{N}\sum_{k=1}^{N} f_k(x) \leq \frac{2M}{\sqrt{N}}\left(\sqrt{\ln n} + \sqrt{2\ln\left(\sigma^{-1}\right)}\right),$$

т.е. с вероятностью $\geq 1-\sigma$ наши потери ограничены

$$\frac{1}{N}\sum_{k=1}^{N} f_k(x^k) \leq C + \frac{2M}{\sqrt{N}}\left(\sqrt{\ln n} + \sqrt{2\ln\left(\sigma^{-1}\right)}\right).$$

Самый плохой для нас случай (с точки зрения такой оценки) – это когда игрок А тоже "знает" теорему 2 и действует (выбирая $\{\omega^k\}$) согласно МЗС2 (точнее версии МЗС2 для максимизации вогнутых функций на симплексе). Очевидно, что если и А и Б будут придерживаться МЗС2, то они сойдутся к равновесию Нэша (седловой точке), причем чрезвычайно быстро (см. [22]):

$$\frac{8M\left(\ln n + 2\ln(\sigma^{-1})\right)}{\varepsilon^2} \text{ итераций;}$$

$$O\left(n + M\frac{s\ln n\left(\ln n + \ln(\sigma^{-1})\right)}{\varepsilon^2}\right) \text{ – общее число арифметических операций,}$$

где $s \le n$ – среднее число элементов в строках и столбцах матрицы $A$. Отсюда видно, что если $\varepsilon$ (зазор двойственности [3], [7]) не очень малое, то может случиться, что общее число арифметических операций будет много меньше числа элементов матрицы $A$, отличных от нуля, в то время как любой детерминированный способ поиска равновесия Нэша потребовал бы прочтения как минимум половины элементов матрицы $A$ (см. [6], [22]). Другой способ содержательной интерпретации описанного метода базируется на прямодвойственности МЗС (см. [3]).

**Замечание 12.** Если использовать методы работ [24], [25], то можно получить зависимость сложности от $\varepsilon$ вида $\varepsilon^{-1}$, но при этом число операций увеличится не менее чем в $n$ раз. Поскольку для таких задач вполне естественным является соотношение $\varepsilon \gg n^{-1}$, то выгоднее использовать описанный в этом примере алгоритм.

Есть основания полагать, что описанный здесь метод оптимален не только с точки зрения числа итераций, но и с точки зрения "стоимости" шага. Особенно ярко это проявляется, когда $s \ll n$ (см. [22]). Нам кажется, что описанная в этом примере методология может оказаться полезной в huge-scale optimization (не только для рассмотренной здесь задачи). В частности, при изучении того, как оптимальным образом можно учитывать разреженность задачи (см. [26]) и как оптимально организовывать покомпонентный спуск (см. [27]). Отметим в связи с вышесказанным другой пример задачи выпуклой оптимизации, когда удается получить ответ (с требуемой точностью), с большим запасом не просматривая весь объем имеющихся данных (общая проблема здесь: как понять, что просматривать, а что нет): метод наименьших квадратов с разреженной структурой (см. [28]). Нам представляется, что именно эта ветвь более общего и бурно развивающегося в последнее время направления huge-scale оптимизации является сейчас одной из наиболее интересных как с практической, так и с теоретической точек зрения, и основные результаты еще впереди. Уточним, что речь идет о задачах, приходящих из: машинного обучения (в частности, распознавания изображений), моделирования различных сетей огромных размеров типа сети Интернет или транспортных сетей, биоинформатики, численных методов (проектирование конструкций методом конечных элементов) и ряда других приложений. Все эти задачи отличают колоссальные размеры. Скажем, для задачи ранжирования web-страниц необходимо решать вспомогательную оптимизационную задачу в пространстве, размерность которого больше миллиарда (см. [22]). Помимо размеров, их отличают некоторые релаксированные требования к решению. Например, нет никакой необходимости ранжировать абсолютно все web-страницы по заданному запросу и делать это очень точно. Достаточно, чтобы качественно выдавались только первые сто наиболее значимых (больших) компонент ранжирующего вектора, причем важны не столько сами значения этих компонент, сколько их порядок. Также эти задачи довольно специальные, т.е. использовать концепцию черного ящика (см. [1]) для оценки числа требуемых итераций, как правило, не представляется возможным. Более того, оценки имеются, в основном, только на число шагов (итераций). Поскольку общее время работы алгоритма определяется произведением числа итераций на стоимость одной итерации, то возникает игра между числом итераций и стоимостью итерации. Описанный в этом примере метод как раз играет в эту игру. А именно, он увеличивает за счет рандомизации число итераций в $\ln(\sigma^{-1})$ раз, при этом итерация становится в $n/\ln n$ раз дешевле. Наконец, важной составляющей многих задач является разреженная структура. Тут имеется два варианта: разреженная структура решения и данных. В первом случае (частично уже затронутым выше на примере ранжирования web-страниц) часто речь идет о подмене исходной задачи, вычислительно более привлекательной задачей, по решению которой можно получить приближенное решение исходной. Пожалуй, наиболее ярким примером здесь является сжатие измерений (см. [30] и цитированную там литературу). В случае разреженных данных представляется перспективным использование специальных покомпонентных спусков и исследование вычислительных особенностей пересчета различных классов функций многих переменных в случае изменения лишь небольшого числа их аргументов (см. [26], [27]). В заключение отметим, что важными составляющими анализа эффективности методов (в виду специфики описываемых задач) является исследование возможности распараллеливания (см. [29]) (в рассматриваемом нами примере 5 это возможно сделать, см. [6], [22]) и вероятностный анализ в среднем (см. [30]) (или для почти всех входов). В отличие от Computer Science (см. [31], [32]), в численных методах выпуклой оптимизации такой анализ можно встретить не часто (все же кое-что есть, например, вероятностный анализ симплекс-метода Спилманом и Тэнгом, см. [33]). Хотя уже сейчас (в связи с бурным развитием идей концентрации меры, см. [32], [34]) становится все более и более ясно, что в пространствах огромных размеров такой анализ необходим, и может многое дать (см. [30]).

Отметим, что к рассмотренной в примере 5 антагонистической матричной игре (с $M = 1$) сводится Google problem: задача поиска вектора Фробениуса–Перрона стохастической матрицы $P$ огромных размеров [22]: $A = P^{\mathrm{T}} - I$. Разработанный в [22] и описанный в примере 5 метод позволяет учитывать разреженную структуру матрицы $P$, заменяя в алгоритме [8] $n$, которое с точностью до логарифмических факторов входит линейно в оценку общей трудоемкости метода (т.е. с учетом затрат на каждой итерации) на $s \ll n$.

Основные результаты этой статьи были получены в ходе визитов первого автора в CORE UCL (Лёвен–ЛаНёв, Бельгия) в декабре 2012 г. ко второму автору и в институт Вейерштрасса (Берлин) в мае 2013 г. к третьему автору. Статья была написана летом 2013 г. В 2014 г. Alexander Rakhlin и Karthik Sridharan прочитали курс лекций “Statistical Learning Theory and Sequential Prediction” (драфт этого курса свободно доступен в сети Интернет). Karthik защитил PhD диссертацию по теме “Learning from an Optimization viewpoint”. Все эти материалы нам представляются очень полезными в контексте перенесения результатов данной статьи на задачи статистической теории обучения. К сожалению, формат статьи и сроки не позволяют здесь об этом написать подробнее.

Авторы выражают благодарность В.В. Вьюгину, Г.К. Голубеву, О. Деволдеру, Ю.В. Дорну, П.Е. Двуреченскому, Г. Лугоши, А.В. Назину, А.С. Немировскому, Б.Т. Поляку, П. Рихтарику, У. Сэндхольму, С.П. Тарасову, И.О. Толстихину, А.Б. Юдицкому за ряд ценных замечаний.

## СПИСОК ЛИТЕРАТУРЫ

1. *Немировский А.С., Юдин Д.Б.* Сложность задач и эффективность методов оптимизации. М.: Наука, 1979. http://www2.isye.gatech.edu/~nemirovs/Lect_EMCO.pdf
2. *Beck A., Teboulle M.* Mirror descent and nonlinear projected subgradient methods for convex optimization // Oper. Res. Lett. 2003. V. 31. P. 167–175.
3. *Nesterov Y.* Primal-dual subgradient methods for convex problems // Math. Program. Ser. B. 2009. V. 120. № 1. P. 261–283.
4. *Юдицкий А.Б., Назин А.В., Цыбаков А.Б., Ваятис Н.* Рекуррентное агрегирование оценок методом зеркального спуска с усреднением // Пробл. передачи информации. 2005. Т. 41. № 4. С. 78–96.
5. *Bubeck S.* Introduction to online optimization. Princeton University: Lecture Notes, 2011. http://www.princeton.edu/~sbubeck/BubeckLectureNotes.pdf
6. *Grigoriadis M., Khachiyan L.* A sublinear-time randomized approximation algorithm for matrix games // Oper. Res. Lett. 1995. V. 18. № 2. P. 53–58.
7. *Juditsky A., Lan G., Nemirovski A., Shapiro A.* Stochastic approximation approach to stochastic programming // SIAM J. Optimizat. 2009. V. 19. № 4. P. 1574–1609. http://www2.isye.gatech.edu/~nemirovs/
8. *Назин А.В., Поляк Б.Т.* Рандомизированный алгоритм нахождения собственного вектора стохастической матрицы с применением к задаче PageRank // Автоматика и телемех. 2011. № 2. С. 131–141.
9. *Lugosi G., Cesa-Bianchi N.* Prediction, learning and games. New York: Cambridge University Press, 2006.
10. *Juditsky A., Nazin A.V., Tsybakov A.B., Vayatis N.* Gap-free bounds for stochastic multi-armed bandit // Proc. 17th World Congress IFAC, Seoul, Korea, July 6–11 2008. P. 11560–11563.
11. *Mansour Y.* Algorithmic game theory and machine learning. 2011. Lect. 1, 2, 6. http://www.tau.ac.il/~mansour/advanced-agt+ml/
12. *Rakhlin A.* Lecture notes on online learning, 2009. http://stat.wharton.upenn.edu/~rakhlin/papers/online_learning.pdf
13. *Hazan E.* The convex optimization approach to regret minimization. In: Optimiz. for Machine Learning. Eds. S. Sra, S. Nowozin, S. Wright. MIT Press, 2011. P. 287–303.

14. *Bubeck S., Cesa-Bianchi N.* Regret analysis of stochastic and nonstochastic multi-armed bandit problems // Foundation and Trends in Machine Learning. 2012. V. 5. № 1. P. 1–122. http://www.princeton.edu/~sbubeck/SurveyBCB12.pdf
15. *Shi Z., Liu R.* Online universal gradient method // e-print, 2013. arXiv:1311.3832v2
16. *Juditsky A., Nemirovski A.* First order methods for nonsmooth convex large-scale optimization, I, II. In: Optimizat. for Machine Learning. Eds. S. Sra, S. Nowozin, S. Wright. MIT Press, 2012.
17. *Boucheron S., Lugoshi G., Massart P.* Concentration inequalities: a nonasymptotic theory of independence. Oxford University Press, 2013.
18. *Lan G., Nemirovski A., Shapiro A.* Validation analysis of mirror descent stochastic approximation method // Math. Program. 2012. V. 134. № 2. P. 425–458.
19. *Andersen S.P., de Palma A., Thisse J.-F.* Discrete choice theory of product differentiation. MIT Press, Camb., 1992.
20. *Лидбеттер М., Линдгрен Г., Ротсен Х.* Экстремумы случайных последовательностей и процессов. М.: Мир, 1989.
21. *Гасников А.В., Дорн Ю.В., Нестеров Ю.Е., Шпирко С.В.* О трехстадийной версии модели стационарной динамики транспортных потоков // Матем. моделирование. 2014. Т. 26. № 6. С. 34–70. arXiv:1405.7630
22. *Гасников А.В., Дмитриев Д.Ю.* Об эффективных рандомизированных алгоритмах поиска вектора PageRank // Ж. вычисл. матем. и матем. физ. 2015. Т. 55. № 3. С. 355–371. arXiv:1410.3120
23. *Вьюгин В.В.* Математические основы теории машинного обучения и прогнозирования. М.: МЦНМО, 2013. http://www.iitp.ru/upload/publications/6256/vyugin1.pdf
24. *Nemirovski A.* Prox-method with rate of convergence $O(1/T)$ for variational inequalities with Lipschitz continuous monotone operators and smooth convex-concave saddle point problems // SIAM J. Optimizat. 2004. V. 15. P. 229–251.
25. *Nesterov Y.* Smooth minimization of non-smooth function // Math. Program. Ser. A. 2005. V. 103. № 1. P. 127–152.
26. *Nesterov Y.E.* Subgradient methods for huge-scale optimization problems // Math. Program. Ser. A. 2014 (in print); CORE Discussion Paper 2012/2. 2012.
27. *Nesterov Y.E.* Efficiency of coordinate descent methods on large scale optimization problem // SIAM J. Optimizat. 2012. V. 22. № 2. P. 341–362.
28. *Bubeck S.* Theory of convex optimization for machine learning // e-print, 2014. arXiv:1405.4980
29. *Fercoq O., Richtárik P.* Accelerated, parallel and proximal coordinate descent // e-print, 2013. arXiv:1312.5799
30. *Тао Т.* Структура и случайность. М.: МЦНМО, 2013.
31. *Motwani R., Raghavan P.* Randomized algorithms. Cambridge Univ. Press, 1995.
32. *Dubhashi D.P., Panconesi A.* Concentration of measure for the analysis of randomized algorithms. Cambridge University Press, 2009.
33. *Spielman D.A., Teng S.-H.* Smoothed analysis of algorithms: Why the simplex algorithm usually takes polynomial time // J. ACM. 2004. V. 51 P. 385–463.
34. *Ledoux M.* Concentration of measure phenomenon. Providence, RI, Amer. Math. Soc., 2001 (Math. Surveys Monogr. V. 89).