СИБИРСКИЙ ЖУРНАЛ ИНДУСТРИАЛЬНОЙ МАТЕМАТИКИ. 2024. Т. 27, № 2. С. 80–99

УДК 519.86

О ДВОЙСТВЕННОМ МЕТОДЕ ГРАДИЕНТНОГО СПУСКА ДЛЯ ЗАДАЧИ О РАСПРЕДЕЛЕНИИ РЕСУРСОВ В МНОГОАГЕНТНЫХ СИСТЕМАХ

© 2024 Д. Б. Рохлин

Южный федеральный университет,
Институт математики, механики и компьютерных наук,
Региональный научно-образовательный математический центр,
ул. Мильчакова, 8а, г. Ростов-на-Дону 344090, Россия

E-mail: dbrohlin@sfedu.ru

Поступила в редакцию 13.01.2024 г.; после доработки 09.03.2024 г.;
принята к публикации 17.04.2024 г.

Рассматривается последовательность блочно-сепарабельных задач выпуклого программирования, описывающих распределение ресурсов в многоагентных системах. Построено несколько итерационных алгоритмов назначения цен на ресурсы. При различных предположениях о функциях полезности и ресурсных ограничениях получены оценки для среднего отклонения целевой функции от оптимального значения (сожаления) и величины невязки в ограничениях. Среднее здесь понимается как математическое ожидание для независимых одинаково распределённых данных, и как временно́е среднее в задаче онлайн оптимизации. Анализ задачи проводится на основе методов онлайн оптимизации и теории двойственности. Рассмотренные алгоритмы основаны на информации о разности между суммарным спросом и предложением, которая порождается реакциями агентов на цены и соответствует невязке в ограничениях.

Ключевые слова: онлайн оптимизация, адаптивный градиентный спуск, двойственность, сожаление, выявленные предпочтения.

DOI: 10.33048/SIBJIM.2024.27.206

1. ПОСТАНОВКА ЗАДАЧИ

Для задачи блочно-сепарабельного выпуклого программирования двойственная по Лагранжу распадается на несвязанные между собой подзадачи. В задаче о распределении ресурсов между активными агентами эта декомпозиция имеет следующий экономический смысл: каждая из указанных подзадач описывает реакцию агента на цены ресурсов, представленные вектором множителей Лагранжа. Вместо ресурсов речь может идти также о продуктах или товарах. Использование цен как координационного механизма в таких задачах хорошо известно [1], гл. 6, [2] и связано с механизмом Вальраса [3], гл. 10.

В данной работе мы считаем, что цены назначаются лидером (например, регулятором рынка электричества), и целью их назначения является приближённая максимизация кооперативного критерия качества (суммарной полезности агентов) при приближённом выполнении ресурсных ограничений. При этом оптимальное решение двойственной задачи стимулирует точное оптимальное решение прямой задачи, т. е. кооперативно оптимальное решение, удовлетворяющее ресурсным ограничениям.

Нас интересуют алгоритмы назначения цен, использующие лишь информацию о невязках в ограничениях, порождаемых реакциями агентов. Эти невязки соответствуют разности между суммарным спросом со стороны агентов (или, например, потоком в сети) и имеющимися ресурсами. Такие алгоритмы не используют личную или труднодоступную и ненадёжную информацию о функциях затрат агентов.

Базовым методом решения поставленной задачи можно считать двойственный градиентный спуск, который непосредственно приводит к методу корректировки цен на основе разности между спросом и предложением. Этот подход многократно использовался в различных задачах, касающихся регулирования транспортных потоков [4], распределения ресурсов в компьютерных сетях [5], отыскания равновесия типа Вальраса [6], производства электроэнергии [7], и др.

В данной работе мы считаем, что агенты являются неоднородными и их поведение может меняться с течением времени. Формально говоря, в общем случае, рассматриваемая задача относится к классу задач выпуклой онлайн оптимизации с долгосрочными ограничениями [8]. Её исследование проводится обычно основе понятия сожаления, которое равно разности между накопленными потерями алгоритма и накопленными потерями ретроспективно лучшего фиксированного действия (статическое сожаление), либо оптимальной последовательности действий (динамическое сожаление) [9, 10]. Если целевые и функции ограничения представляют собой независимые одинаково распределённые величины, то получаем задачу стохастической оптимизации. В этом случае среднее понимается как математическое ожидание и рассматриваются оценки ошибки оценивания (estimation error). Их можно получить с помощью известных методов конвертирования онлайн алгоритмов в пакетные [11–13].

Для задач выпуклой онлайн оптимизации с долгосрочными ограничениями разработано достаточно много алгоритмов, которые при определённых условиях обеспечивают одновременно малое в среднем (по времени) сожаление и малую невязку в ограничениях. Известны теоретические гарантии как для статического, так и для динамического сожаления [14–16]. Полученные в указанных этих работах оценки отличаются от представленных ниже. В частности, полученная в теореме 1 оценка динамического сожаления не зависит от суммы длин приращений оптимальной последовательности действий. Это связано со спецификой постановки задачи: мы считаем, что агенты решают локальные задачи с полной информацией, хотя центру при назначении цен она неизвестна.

Подчеркнём также, что в рассматриваемой постановке задачи действия агентов определяются текущей ценой и их функциями полезности. Таким образом, стратегии агентов не могут быть навязаны извне. В связи с этим, с точки зрения допустимых алгоритмов решения поставленной задачи, данная работа ближе к [4–6, 17] и литературе, посвящённой оптимизации на основе выявленных предпочтений, т. е. непосредственной реакции агентов на действия лидера [18–20].

Сформулируем математическую постановку задачи. Пусть $x = (x_1, \ldots, x_n) \in \prod_{j=1}^{n} \mathbb{R}^{k_j}$. Рассмотрим последовательность блочно-сепарабельных задач выпуклого программирования

$$f_t(x) := \sum_{j=1}^{n} f_{t,j}(x_j) \to \max_{x \in S}, \tag{1}$$

$$S : g_{t,i}(x) := \sum_{j=1}^{n} g_{t,ij}(x_j) \leqslant b_{t,i},\ i = 1, \ldots, m; \quad x \in B := \prod_{i=1}^{n} B_i, \tag{2}$$

где $B_j \subset \mathbb{R}^{k_j}$ — выпуклые компактные множества, $f_{t,j}$ — вогнутые функции, $g_{t,ij}$ — выпуклые функции, которые для простоты будем считать определёнными на $\mathbb{R}^{k_j}$, и $t \in \{1, \ldots, T\}$. Рассмотрим последовательность функций Лагранжа

$$L_t(x, \lambda) = f_t(x) + \langle \lambda, b_t - g_t(x) \rangle : B \times \mathbb{R}_+^m \to \mathbb{R},$$

$b_t = (b_{t,1}, \dots, b_{t,m})$, $g_t = (g_{t,1}, \dots, g_{t,m})$, и двойственных оптимизационных задач

$$d_t(\lambda) := \sup_{x \in B} L_t(x, \lambda) = \langle \lambda, b_t \rangle + \sum_{j=1}^n \sup_{x_j \in B_j} \left(f_{t,j}(x_j) - \langle \lambda, g_{t,\cdot j}(x_j) \rangle \right) \to \min_{\lambda \in \mathbb{R}_+^m}. \tag{3}$$

Здесь $\langle a, b \rangle = \sum_{i=1}^m a_i b_i$ — стандартное скалярное произведение, и

$$g_{t,\cdot j} = (g_{t,1j}, \dots, g_{t,mj}).$$

Ясно, что оптимальное решение $x_t^* = (x_{t,1}^*, \dots, x_{t,n}^*)$ задачи (1), (2) существует. Хорошо известен следующий результат [21], §28, [22], предложение 6.4.4. Пусть выполнено слабое условие Слейтера, т. е. существует точка y из относительной внутренности B: $y \in \operatorname{ri} B = \operatorname{ri} B_1 \times \dots \times \operatorname{ri} B_n$ такая, что

$$g_{t,i}(y) < b_{t,i}, \ i \in I; \quad g_{t,i}(y) \leqslant b_{t,i}, \ i \notin I,$$

где I — множество индексов, для которых $g_{t,i}$ не является аффинной. Тогда существует оптимальное решение λ_t^* двойственной задачи и нет скачка двойственности:

$$f_t(x_t^*) = d_t(\lambda_t^*).$$

Если интерпретировать $f_{t,j}$ как функцию полезности j-го агента, а компоненты вектора множителей Лагранжа λ как цены ресурсов, то, как видно из формулы (3), после назначения цен каждый агент выбирает оптимальное действие

$$\widetilde{x}_{t,j}(\lambda) \in \operatorname*{argmax}_{x_j \in B_j} \left(f_{t,j}(x_j) - \langle \lambda, g_{t,\cdot j}(x_j) \rangle \right). \tag{4}$$

Оптимальная реакция $\widetilde{x}_t(\lambda)$ называется выявленным предпочтением [18, 19]. Заметим, что в общем случае необязательно $\widetilde{x}_t(\lambda_t^*)$ является оптимальным решением x_t^* прямой задачи. Но для сильно выпуклых функций $f_{t,j}$ это верно. В разделе 3 это будет показано в более общей ситуации.

Из хорошо известных формул для субдифференциала суммы и максимума [23], теоремы 3.40, 3.55, вытекает, что $b_t - g_t(\widetilde{x}_t(\lambda))$ принадлежит субдифференциалу выпуклой функции d_t в точке λ:

$$b_t - g_t(\widetilde{x}_t(\lambda)) \in \partial d_t(\lambda), \quad \lambda \in \mathbb{R}_+^m.$$

Наша цель состоит в построении последовательности векторов цен λ_t, обеспечивающих для соответствующих реакций (4) агентов малое (растущее медленнее T) динамическое сожаление

$$\sum_{t=1}^T f_t(x_t^*) - \sum_{t=1}^T f_t(\widetilde{x}_t(\lambda_t)) \tag{5}$$

для последовательности прямых задач, и малую суммарную невязку в ограничениях

$$\left(\sum_{t=1}^T (g_{t,i}(\widetilde{x}_t(\lambda_t)) - b_{t,i}) \right)^+, \tag{6}$$

$a^+ = \max\{0, a\}$. Использование величин (5), (6) для оценки качества алгоритмов является стандартным в теории онлайн оптимизации [10, 12] и, в частности, в задаче с долгосрочными ограничениями, понимаемыми в среднем [8].

Для построения последовательности λ_t мы применяем адаптивный метод градиентного спуска. При отсутствии вероятностных предположений оценки величин (5), (6) получены в теореме 1. В предположении, что f_t, g_t, b_t являются независимыми одинаково распределёнными случайными величинами, мы строим вектор цен $\overline{\lambda}_T$ после T наблюдений. В этом случае математические ожидания величин (5), (6) соответствуют ошибкам оценивания и ожидаемой невязке в ограничениях. Три варианта оценок этих величин, соответствующие различным способам выбора $\overline{\lambda}_T$, представлены в теоремах 2–4. Отметим, что вычислительная сложность итераций предложенных алгоритмов имеет порядок $O(m)$, где m — число ограничений (размерность вектора множителей Лагранжа). Это вытекает из явных формул.

2. ЗАДАЧА ОНЛАЙН ОПТИМИЗАЦИИ

Всюду далее будем считать выполненными следующие условия

$$0 \in B, \quad f_t(0) = g_{t,i}(0) = 0, \tag{7}$$

$$|f_t(x)| \leqslant \overline{f}, \quad |g_{t,i}(x)| \leqslant \overline{g}_i,\ x \in B, \quad |b_{t,i}| \leqslant \overline{b}_i. \tag{8}$$

Первого из них можно всегда добиться с помощью переобозначений, а условие (8) равномерной (по t) ограниченности является стандартным.

Базовым методом решения задач выпуклой онлайн оптимизации является онлайн градиентный спуск [9,10,12]. Применим его с евклидовой проекцией на неотрицательный ортант $\mathbb{R}^m_+$ к последовательности двойственных задач (3):

$$\lambda_{t+1,i} = (\lambda_{t,i} - \eta_{t,i}\nabla_{t,i})^+, \quad \nabla_t := b_t - g_t(\widetilde{x}_t(\lambda_t)). \tag{9}$$

Проекция вектора a на $\mathbb{R}^m_+$ сводится к поэлементному вычислению неотрицательной части: a^+. Вектор ∇_t является субградиентом функции d_t. Величина шага $\eta_{t,i}$ может выбираться отдельно для каждой координаты. Далее предполагается, что векторная последовательность шагов η_t является поэлементно невозрастающей:

$$\eta_{t+1,i} \leqslant \eta_{t,i}.$$

Заметим, что последовательность (9) строится на основе информации о разности между суммарным спросом $g_t(\widetilde{x}_t(\lambda_t))$, соответствующем выявленным предпочтениям, и предложением b_t.

Основной результат сформулирован в теореме 1, которой мы предпошлём несколько лемм.

Лемма 1. *Для любого $\lambda \in \mathbb{R}^m_+$ верно неравенство*

$$\sum_{t=1}^{T}(\lambda_{t,i} - \lambda_i)\nabla_{t,i} \leqslant \frac{1}{2\eta_{T,i}} \max_{1\leqslant t\leqslant T}(\lambda_{t,i} - \lambda_i)^2 + \sum_{t=1}^{T}\frac{\eta_{t,i}}{2}\nabla^2_{t,i}.$$

Доказательство. Легко видеть, что

$$(a^+ - z)^2 \leqslant (a - z)^2$$

для любых $a \in \mathbb{R}$, $z \geqslant 0$. Поэтому

$$\frac{1}{2}(\lambda_{t+1,i} - \lambda_i)^2 - \frac{1}{2}(\lambda_{t,i} - \lambda_i)^2 \leqslant \frac{1}{2}(\lambda_{t,i} - \eta_{t,i}\nabla_{t,i} - \lambda_i)^2 - \frac{1}{2}(\lambda_{t,i} - \lambda_i)^2$$
$$= -(\lambda_{t,i} - \lambda_i)\eta_{t,i}\nabla_{t,i} + \frac{1}{2}\eta^2_{t,i}\nabla^2_{t,i}.$$

Следовательно,

$$\sum_{t=1}^{T}(\lambda_{t,i}-\lambda_i)\nabla_{t,i} \leqslant \sum_{t=1}^{T}\frac{(\lambda_{t,i}-\lambda)_i^2}{2\eta_{t,i}} - \sum_{t=1}^{T}\frac{(\lambda_{t+1,i}-\lambda_i)^2}{2\eta_{t,i}} + \sum_{t=1}^{T}\frac{\eta_{t,i}}{2}\nabla_{t,i}^2$$
$$= \frac{(\lambda_{1,i}-\lambda_i)^2}{2\eta_{1,i}} + \sum_{t=1}^{T-1}\left(\frac{(\lambda_{t+1,i}-\lambda_i)^2}{2\eta_{t+1,i}} - \frac{(\lambda_{t+1,i}-\lambda_i)^2}{2\eta_{t,i}}\right) - \frac{(\lambda_{T+1,i}-\lambda_i)^2}{2\eta_{T,i}} + \sum_{t=1}^{T}\frac{\eta_{t,i}}{2}\nabla_{t,i}^2$$
$$\leqslant \frac{(\lambda_{1,i}-\lambda_i)^2}{2\eta_{1,i}} + \frac{1}{2}\max_{2\leqslant t\leqslant T}(\lambda_{t,i}-\lambda_i)^2\left(\frac{1}{\eta_{T,i}}-\frac{1}{\eta_{1,i}}\right) + \sum_{t=1}^{T}\frac{\eta_{t,i}}{2}\nabla_{t,i}^2.$$

В последнем неравенстве использована монотонность последовательности η_t. Из него очевидным образом вытекает утверждение леммы. □

Лемма 2. *Справедливо неравенство*

$$\sum_{t=1}^{T} f_t(x_t^*) - \sum_{t=1}^{T} f_t(\widetilde{x}_t(\lambda_t)) \leqslant \sum_{t=1}^{T}\langle\lambda_t, \nabla_t\rangle.$$

Доказательство. Воспользуемся определением двойственной целевой функции и сильной теоремой двойственности:

$$d_t(\lambda) = f_t(\widetilde{x}_t(\lambda)) + \langle\lambda, b_t - g_t(\widetilde{x}_t(\lambda))\rangle, \quad f_t(x_t^*) = d_t(\lambda_t^*).$$

Отсюда следует, что

$$f_t(x_t^*) - f_t(\widetilde{x}_t(\lambda_t)) = d_t(\lambda_t^*) - d_t(\lambda_t) + \langle\lambda_t, b_t - g_t(\widetilde{x}_t(\lambda_t))\rangle \leqslant \langle\lambda_t, \nabla_t\rangle,$$

так как в точке λ_t^* достигается минимум функции d_t. □

Не ограничивая общности будем считать, что $\nabla_1 \neq 0$. В противном случае $\lambda_1 = \lambda_1^*$, $\widetilde{x}_1(\lambda_1) = x_1^*$ и первые слагаемые в суммах (5), (6) будут нулевыми. Поскольку в этом случае $\lambda_2 = \lambda_1$, то последовательность λ_t не сдвинется с места и будет порождать оптимальное решение $\widetilde{x}_t(\lambda_t)$ прямой задачи до тех пор пока ∇_t не станет отличным от нуля.

Лемма 3. *Для метода градиентного спуска (9) с шагами типа AdaGrad (см., например, [24], гл. 4)*

$$\eta_{t,i} = \frac{\gamma_i}{\sqrt{\sum_{s=1}^{t}\nabla_{s,i}^2}} \tag{10}$$

справедливы оценки

$$\sum_{t=1}^{T} f_t(x_t^*) - \sum_{t=1}^{T} f_t(\widetilde{x}_t(\lambda_t)) \leqslant \sum_{i=1}^{m}\left(\frac{1}{2\gamma_i}\max_{1\leqslant t\leqslant T}\lambda_{t,i}^2 + \gamma_i\right)\sqrt{\sum_{t=1}^{T}\nabla_{t,i}^2} \tag{11}$$

$$\sum_{t=1}^{T}(g_{t,i}(\widetilde{x}_t(\lambda_t)) - b_{t,i}) \leqslant \frac{\lambda_{T+1,i}}{\gamma_i}\sqrt{\sum_{t=1}^{T+1}\nabla_{t,i}^2}. \tag{12}$$

Доказательство. Как известно, для любых чисел $a_1 > 0, a_2 \geqslant 0, \dots, a_T \geqslant 0$ справедливо неравенство [25], лемма 3.5,

$$\sum_{t=1}^{T} \frac{a_t}{\sqrt{\sum_{s=1}^{t} a_s}} \leqslant 2\sqrt{\sum_{t=1}^{T} a_t}.$$

Следовательно,

$$\sum_{t=1}^{T} \frac{\eta_{t,i}}{2} \nabla_{t,i}^2 = \frac{\gamma_i}{2} \sum_{t=1}^{T} \frac{\nabla_{t,i}^2}{\sqrt{\sum_{k=1}^{t} \nabla_{k,i}^2}} \leqslant \gamma_i \sqrt{\sum_{t=1}^{T} \nabla_{t,i}^2}. \tag{13}$$

Чтобы получить неравенство (11), воспользуемся леммами 2 и 1 (с $\lambda = 0$):

$$\begin{aligned}
\sum_{t=1}^{T} f_t(x_t^*) - \sum_{t=1}^{T} f_t(\widetilde{x}_t(\lambda_t)) &\leqslant \sum_{t=1}^{T} \langle \lambda_t, \nabla_t \rangle \\
&\leqslant \sum_{i=1}^{m} \left(\frac{1}{2\eta_{T,i}} \max_{1 \leqslant t \leqslant T} \lambda_{t,i}^2 + \sum_{t=1}^{T} \frac{\eta_{t,i}}{2} \nabla_{t,i}^2 \right) \\
&\leqslant \sum_{i=1}^{m} \left(\frac{1}{2\gamma_i} \max_{1 \leqslant t \leqslant T} \lambda_{t,i}^2 + \gamma_i \right) \sqrt{\sum_{t=1}^{T} \nabla_{t,i}^2}.
\end{aligned}$$

В последнем неравенстве использована оценка (13) и определение $\eta_{T,i}$.

Далее, поскольку

$$\lambda_{t+1,i} \geqslant \lambda_{t,i} - \eta_{t,i} \nabla_{t,i},$$

то

$$\begin{aligned}
\sum_{t=1}^{T} (g_{t,i}(\widetilde{x}_t(\lambda_t)) - b_{t,i}) = -\sum_{t=1}^{T} \nabla_{t,i} &\leqslant \sum_{t=1}^{T} \left(\frac{\lambda_{t+1,i}}{\eta_{t,i}} - \frac{\lambda_{t,i}}{\eta_{t,i}} \right) \\
&\leqslant \sum_{t=1}^{T} \left(\frac{\lambda_{t+1,i}}{\eta_{t+1,i}} - \frac{\lambda_{t,i}}{\eta_{t,i}} \right) \leqslant \frac{\lambda_{T+1,i}}{\eta_{T+1,i}},
\end{aligned}$$

что соответствует (12). □

Чтобы получить из леммы 3 оценки сожаления и суммарной невязки в ограничениях, нужно оценить λ_t. В следующей лемме установлена ограниченность последовательности λ_t при дополнительных предположениях.

Лемма 4. *Пусть выполнены условия*

$$g_{t,i}(x) \geqslant 0, \quad x \in B, \tag{14}$$

$$b_{t,i} \geqslant \underline{b}_i > 0. \tag{15}$$

Тогда если неравенство

$$0 \leqslant \lambda_{t,i} \leqslant C_i + \eta_{1,i} \overline{g}_i, \quad C_i = \frac{\overline{f}}{\underline{b}_i}$$

верно при $t = 0$, то оно верно при всех t. В частности, при $\lambda_1 = 0$ последовательность λ_t равномерно по t ограничена.

Доказательство. Пусть $\lambda_{t,i}\leqslant C_i+\eta_{1,i}\overline{g}_i$. Если $\lambda_{t,i}\leqslant C_i$, то либо

$$\lambda_{t,i}+\eta_{t,i}(g_{t,i}(\widetilde{x}_t(\lambda_t))-b_{t,i})\leqslant 0,$$

и тогда $\lambda_{t+1,i}=0$, либо

$$\lambda_{t+1,i}=\lambda_{t,i}+\eta_{t,i}(g_{t,i}(\widetilde{x}_t(\lambda_t))-b_{t,i})\leqslant C_i+\eta_{1,i}\overline{g}_i$$

в силу (8), (15). Если $C_i<\lambda_{t,i}\leqslant C_i+\eta_{1,i}\overline{g}_i$, то применяя последовательно условие (7), определение $\widetilde{x}_t(\lambda)$, условия (8) и (14), находим

$$\begin{aligned}0\leqslant \sup_{x\in B}(f_t(x)-\langle\lambda,g_t(x)) &= f_t(\widetilde{x}_t(\lambda))-\langle\lambda,g_t(\widetilde{x}_t(\lambda))\\ &\leqslant \overline{f}-\lambda_i g_{t,i}(\widetilde{x}_t(\lambda))\leqslant C_i\underline{b}_i-C_i g_{t,i}(\widetilde{x}_t(\lambda)).\end{aligned}$$

С учётом (15) отсюда следует, что $\nabla_{t,i}=b_{t,i}-g_{t,i}(\widetilde{x}_t(\lambda_t))\geqslant \underline{b}_i-g_{t,i}(\widetilde{x}_t(\lambda_t))\geqslant 0$, и значит

$$\lambda_{t+1,i}\leqslant\lambda_{t,i}\leqslant C_i+\eta_{1,i}\overline{g}_i.$$

Лемма доказана. □

Условие (14) достаточно естественно, если интерпретировать $g_{t,i}$ как количество запрашиваемых ресурсов. Условие (15) можно рассматривать как равномерное условие Слейтера. В целом, смысл приведённого доказательства состоит в том, что при достаточно больших ценах ресурсов спрос не превосходит предложения, что приводит к понижению цен.

Теорема 1. *Пусть выполнены условия (14), (15). Тогда для метода градиентного спуска (9) шагами (10) и* $\lambda_1=0$ *справедливы оценки*

$$\frac{1}{T}\left(\sum_{t=1}^T f_t(x_t^*)-\sum_{t=1}^T f_t(\widetilde{x}_t(\lambda_t))\right)\leqslant\frac{A}{\sqrt{T}},\tag{16}$$

$$\frac{1}{T}\sum_{t=1}^T(g_{t,i}(\widetilde{x}_t(\lambda_t))-b_{t,i})\leqslant\frac{A_i}{\sqrt{T}},\tag{17}$$

$$A=\sum_{i=1}^m\left(\frac{1}{2\gamma_i}\left(\frac{\overline{f}}{\underline{b}_i}+\eta_{1,i}\overline{g}_i\right)^2+\gamma_i\right)|\overline{b}_i+\overline{g}_i|,\quad A_i=\frac{\sqrt{2}}{\gamma_i}\left(\frac{\overline{f}}{\underline{b}_i}+\eta_{1,i}\overline{g}_i\right)|\overline{b}_i+\overline{g}_i|.$$

Доказательство. Доказательство сразу следует из оценок леммы 3 с учётом установленного в лемме 4 неравенства

$$\lambda_{t,i}\leqslant\frac{\overline{f}}{\underline{b}_i}+\eta_{1,i}\overline{g}_i$$

и очевидных неравенств

$$|\nabla_{t,i}|=|b_{t,i}-g_{t,i}(\widetilde{x}_t(\lambda_t))|\leqslant|\overline{b}_i+\overline{g}_i|,\quad\sqrt{T+1}\leqslant\sqrt{2}\sqrt{T}.$$

Теорема доказана. □

Оценки, аналогичные (16), (17), установлены в работе [4] для задачи о регулировании транспортных потоков с помощью назначения платы за проезд по дорогам. Рассмотрение в указанной работе ведётся в рамках задач линейного программирования, а в предложенном алгоритме градиентного спуска использован постоянный шаг. Тем не менее, постановка задачи и методы близки к тем, которые использованы в настоящей работе.

3. ЗАДАЧА С НЕЗАВИСИМЫМИ ОДИНАКОВО РАСПРЕДЕЛЁННЫМИ ДАННЫМИ

Пусть f_t, $g_{t,i}$, $b_{t,i}$ имеют вид

$$f_{t,j}(x_j) = f_j(x_j, \xi_t), \quad g_{t,ij}(x_j) = g_{ij}(x_j, \xi_t), \quad b_{t,i} = b_i(\xi_t), \tag{18}$$

где $-f_j, g_{ij} : B_j \times \mathbb{R}^q \to \mathbb{R}$ — выпуклы по первому аргументу и измеримы по второму, $b_i : \mathbb{R}^q \to \mathbb{R}$ — измеримы, и ξ_t — независимые одинаково распределённые случайные величины со значениями в $\mathbb{R}^q$. Условия (7), (8) по-прежнему считаются выполненными.

Пусть $\sigma(\xi_t)$ — σ-алгебра, порождённая случайной величиной ξ_t, и $L^0(\sigma(\xi_t), B)$ — множество $\sigma(\xi_t)$-измеримых случайных величин со значениями в B. Рассмотрим кооперативную оптимизационную задачу о распределении ресурсов, в которой агентам известны значения случайных величин ξ_t, и ресурсное ограничение понимается в среднем:

$$\mathfrak{F}(x) := \mathsf{E} f_t(x) = \mathsf{E} f(x; \xi_t) = \mathsf{E} \sum_{j=1}^{n} f_j(x_j; \xi_t) \to \max_{x \in \mathfrak{S}}, \tag{19}$$

$$\mathfrak{S} : \mathsf{E}(g_{t,i} - b_{t,i}) = \mathsf{E}(g_i(x, \xi_t) - b_i(\xi_t)) \leqslant 0, \quad i = 1, \ldots, m, \quad x \in L^0(\sigma(\xi_t), B). \tag{20}$$

Оптимальное решение задачи (19), (20) обозначим через $x_t^* \in L^0(\sigma(\xi_t), B)$. Чтобы доказать его существование рассмотрим случайные оптимизационные задачи

$$f(x, \xi_t) = \sum_{j=1}^{n} f_j(x_j; \xi_t) \to \max_{x \in S_t}, \tag{21}$$

$$S_t : g_i(x, \xi_t) - b_i(\xi_t) \leqslant 0, \quad i = 1, \ldots, m, \quad x \in B \tag{22}$$

и воспользуемся известными результатами об измеримых многозначных отображениях [26], гл. 18. Многозначные отображения $\omega \mapsto S_t(\omega)$ являются компактнозначными и измеримыми относительно $\sigma(\xi_t)$ [27], теорема 14.36. По теореме об измеримом максимуме [26], теорема 18.19, существует измеримая функция x_t^* со значениями в S_t такая, что $x_t^*(\omega)$ является оптимальным решением задачи (21), (22) при каждом $\omega \in \Omega$, где Ω — множество элементарных исходов. Легко видеть, что функция x_t^* является решением задачи (19), (20). Кроме того, поскольку ξ_t одинаково распределены, то можно считать, что x_t^* не зависит от t.

Будем разыскивать вектор $\overline{\lambda}_T$ цен ресурсов (а не последовательность цен λ_t), стимулирующий приближенно оптимальное решение задачи (19), (20). Наиболее прямой путь состоит в том, чтобы построить $\overline{\lambda}_T$ исходя из последовательности (9) с помощью процедуры рандомизации. С этой целью рассмотрим равномерно распределённую на $\{1, \ldots, T\}$ случайную величину τ, не зависящую от $(\xi_t)_{t=1}^T$, и положим

$$\lambda_\tau = \sum_{s=1}^{T} \lambda_t I_{\{\tau = t\}}, \quad I_{\{\tau = t\}} = \begin{cases} 1, & \tau = t, \\ 0, & \tau \neq t. \end{cases}$$

Теорема 2. *В условиях теоремы 1 для функций вида (18) справедливы неравенства*

$$\mathsf{E} f_T(x_T^*) - \mathsf{E} f_T(\widetilde{x}_T(\lambda_\tau)) \leqslant \frac{A}{\sqrt{T}}, \tag{23}$$

$$\mathsf{E}(g_{T,i}(\widetilde{x}_T(\lambda_\tau)) - b_{T,i}) \leqslant \frac{A_i}{\sqrt{T}}, \tag{24}$$

где A, A_i определены в теореме 1.

Доказательство. Легко видеть, что

$$\mathsf{E}f(\widetilde{x}_T(\lambda_\tau),\xi_T)=\mathsf{E}\sum_{t=1}^T I_{\{\tau=t\}}f(\widetilde{x}_T(\lambda_t),\xi_T))\overset{(a)}{=}\sum_{t=1}^T \mathsf{P}(\tau=t)\mathsf{E}f(\widetilde{x}_T(\lambda_t),\xi_T))$$
$$\overset{(b)}{=}\frac{1}{T}\sum_{t=1}^T \mathsf{E}f(\widetilde{x}_t(\lambda_t),\xi_t).$$

В равенстве (a) использована независимость τ и $(\xi_t)_{t=1}^T$, а в равенстве (b) — тот факт, что при фиксированном λ_t случайные величины $(\widetilde{x}_t(\lambda_t),\xi_t)$ и $(\widetilde{x}_T(\lambda_t),\xi_T)$ одинаково распределены. Учитывая также, что

$$\mathsf{E}f(\widetilde{x}_T^*,\xi_T)=\frac{1}{T}\sum_{t=1}^T \mathsf{E}f(\widetilde{x}_t^*,\xi_t),$$

из оценки (16) теоремы 1 получаем оценку (23). Оценка (24) выводится аналогичным образом из (17). □

В машинном обучении величины вида (23) называются ошибками оценивания [28], гл. 5. Средние невязки (24) имеют аналогичный смысл.

Заметим, что рандомизация, фактически, означает усреднение прямой последовательности $\widetilde{x}_t(\lambda_t)$ при сохранении реакций агентов как выявленных предпочтений. С помощью детерминированного усреднения прямой последовательности в работе [29] также получены оценки $O(T^{-1/2})$ для детерминированной задачи. Однако, при таком усреднении действия агентов не являются прямой реакцией на цены.

Вычисление λ_τ требует хранения в памяти всей последовательности $(\lambda_t)_{t=1}^T\in\mathbb{R}^{mT}$. Рассмотрим алгоритмы, не предъявляющие таких требований к памяти и использующие временно́е усреднение последовательности λ_t, которое может быть вычислено по рекуррентной формуле. Для их обоснования нам потребуются оценки ошибки по функции и невязки в ограничениях в прямой задаче через ошибку по функции в двойственной задаче. Такие оценки приведены в лемме 7 после некоторой предварительной работы.

Для задачи (19), (20) составим функцию Лагранжа

$$\mathscr{L}(x,\lambda)=\mathsf{E}\left(f(x;\xi_t)+\langle\lambda,g(x,\xi_t)-b_t\rangle\right),\quad (x,\lambda)\in L^0(\sigma(\xi_t),B)\times\mathbb{R}_+^m$$

и двойственную целевую функцию:

$$D(\lambda)=\sup_{x\in L^0(\sigma(\xi_t),B)}\mathscr{L}(x,\lambda)=\mathsf{E}d_t(\lambda),\tag{25}$$

где d_t является мгновенной двойственной целевой функцией, соответствующей случайной реализации прямой задачи:

$$d_t(\lambda)=d(\lambda,\xi_t):=\langle b_t,\lambda\rangle+\sum_{j=1}^n\sup_{x_j\in B_j}(f_{t,j}(x_j)-\langle g_{t,\cdot j}(x),\lambda\rangle).$$

Второе равенство (25) вытекает из принципа перестановки минимизации и математического ожидания [27], теорема 14.60, [30], теорема 7.80.

Дополнительно к уже наложенным условиям потребуем, чтобы функции $f_{t,j}$ были μ_j-сильно вогнутыми на B_j, т. е. чтобы функции $f_{t,j}+\frac{\mu_j}{2}\|\cdot\|^2$ были вогнутыми. Здесь и всюду далее $\|a\|=\sqrt{\langle a,a\rangle}$ — стандартная евклидова норма. Напомним, что дифференцируемая функция $h:\mathbb{R}^m\to\mathbb{R}$ называется β-гладкой, если

$$\|\nabla h(x)-\nabla h(y)\|\leqslant\beta\|x-y\|,\quad x,y\in\mathbb{R}^m.$$

Лемма 5. *Пусть* $f_{t,j}$ *являются* μ_j*-сильно вогнутыми, а* $g_{t,i}$ *удовлетворяют условию Липшица*

$$|g_{t,i}(x) - g_{t,i}(y)| \leqslant L_i \|x - y\|, \quad L_i > 0.$$

Тогда мгновенные двойственные целевые функции d_t *являются* β*-гладкими с*

$$\beta = \frac{\sum_{i=1}^m L_i^2}{\min_{1 \leqslant j \leqslant n} \mu_j}.$$

Доказательство. Запишем мгновенную двойственную целевую функцию в виде

$$d_t(\lambda) = \langle b, \lambda \rangle + \sup_{x \in B} \left(f_t(x) - \sum_{i=1}^m \lambda_i g_{t,i}(x) \right).$$

Очевидно, что функция $f_t(x) = f_{t,1}(x_1) + \dots + f_{t,n}(x_n)$ является μ-сильно вогнутой с $\mu = \min_{1 \leqslant j \leqslant n} \mu_j$. Из сильной вогнутости $f_{t,j}$ вытекает единственность точки максимума (4). Следовательно, субградиент функции d_t состоит из единственной точки $b_t - g_t(\widetilde{x}_t(\lambda))$, и функция d_t дифференцируема (на $\mathbb{R}^m$) [23], теорема 3.33.

Положим для краткости $z_1 = \widetilde{x}_t(\lambda_1)$, $z_2 = \widetilde{x}_t(\lambda_2)$. Согласно критерию оптимальности для элементарной задачи (21), (22) при некоторых $\nabla f_t(z_k) \in -\partial(-f_t(z_k))$, $\nabla g_{t,i}(z_k) \in \partial g_{t,i}(z_k)$ имеют место неравенства

$$\langle \nabla f_t(z_1) - \sum_{i=1}^m \lambda_{1,i} \nabla g_{t,i}(z_1), z_2 - z_1 \rangle \leqslant 0, \quad \langle \nabla f_t(z_2) - \sum_{i=1}^m \lambda_{2,i} \nabla g_{t,i}(z_2), z_1 - z_2 \rangle \leqslant 0.$$

Сложим эти неравенства

$$\langle \nabla f_t(z_1) - \nabla f_t(z_2), z_2 - z_1 \rangle - \langle \sum_{i=1}^m \lambda_{1,i} \nabla g_{t,i}(z_1) - \sum_{i=1}^m \lambda_{2,i} \nabla g_{t,i}(z_2), z_2 - z_1 \rangle \leqslant 0.$$

Отсюда вытекает, что

$$\begin{aligned}
\mu \|z_2 - z_1\|^2 &\leqslant \langle \nabla f_t(z_1) - \nabla f_t(z_2), z_2 - z_1 \rangle \\
&\leqslant \langle \sum_{i=1}^m \lambda_{1,i} \nabla g_{t,i}(z_1) - \sum_{i=1}^m \lambda_{2,i} \nabla g_{t,i}(z_2), z_2 - z_1 \rangle \\
&= \sum_{i=1}^m \lambda_{1,i} \langle \nabla g_{t,i}(z_1) - \nabla g_{t,i}(z_2), z_2 - z_1 \rangle + \sum_{i=1}^m (\lambda_{1,i} - \lambda_{2,i}) \langle \nabla g_{t,i}(z_2), z_2 - z_1 \rangle \\
&\leqslant \sum_{i=1}^m (\lambda_{1,i} - \lambda_{2,i}) \langle \nabla g_{t,i}(z_2), z_2 - z_1 \rangle,
\end{aligned}$$

где первое неравенство следует из сильной выпуклости $-f_t$ [23], теорема 5.24, а последнее — из выпуклости $g_{t,i}$. Используя оценку норм субградиентов липшицевых функций через константу Липшица [23], теорема 3.61, из последнего неравенства выводим, что

$$\mu \|z_2 - z_1\|^2 \leqslant \sum_{i=1}^m |\lambda_{1,i} - \lambda_{2,i}| L_i \|z_2 - z_1\| \leqslant L \|\lambda_1 - \lambda_2\| \|z_2 - z_1\|, \quad L = \sqrt{\sum_{i=1}^m L_i^2}.$$

Утверждение леммы 5 вытекает из данного неравенства и липшицевости $g_{t,i}$:

$$\|\nabla d_t(\lambda_2) - \nabla d_t(\lambda_2)\| = \|g_t(z_2) - g_t(z_1)\| = \sqrt{\sum_{i=1}^{m}(g_{t,i}(z_2) - g_{t,i}(z_1))^2}$$
$$\leqslant \sqrt{\sum_{i=1}^{m} L_i^2\|z_2 - z_1\|^2} = L\|z_2 - z_1\| \leqslant \frac{L^2}{\mu}\|\lambda_2 - \lambda_1\|.$$

Лемма доказана. □

Лемма 6. *В условиях леммы 5 решение λ^* задачи*

$$D(\lambda) \to \max_{\lambda \in \mathbb{R}_+^m} \tag{26}$$

существует и удовлетворяет неравенству

$$\mathsf{E}\langle \nabla d_t(\lambda^*), \lambda - \lambda^*\rangle \geqslant 0, \quad \lambda \in \mathbb{R}_+^m. \tag{27}$$

Доказательство. Поскольку

$$D(\lambda) = \mathsf{E} d_t(\lambda) \geqslant \mathscr{L}(0, \lambda) = \langle \mathsf{E} b_t, \lambda\rangle \to +\infty, \quad \|\lambda\| \to \infty,\ \lambda \in \mathbb{R}_+^m,$$

то существует точка глобального минимума λ^* непрерывной функции D на $\mathbb{R}_+^m$. Функция D дифференцируема и справедлива формула перестановки дифференцирования по λ и операции математического ожидания [31]:

$$\nabla D(\lambda) = \mathsf{E}\nabla d_t(\lambda).$$

Неравенство (27) равносильно критерию оптимальности λ^*:

$$\langle \nabla D(\lambda^*), \lambda - \lambda^*\rangle = \langle \mathsf{E}\nabla d_t(\lambda^*), \lambda - \lambda^*\rangle \geqslant 0, \quad \lambda \in \mathbb{R}_+^d.$$

Лемма доказана. □

Лемма 7. *В условиях леммы 5 для любого допустимого решения x_t задачи (19), (20) справедливы неравенства*

$$\mathsf{E} f_t(x_t) - \mathsf{E} f_t(\widetilde{x}_t(\lambda)) \leqslant \sqrt{2\beta}\|\lambda\|\sqrt{D(\lambda) - D(\lambda^*)}, \tag{28}$$
$$\mathsf{E}(g_{t,i}(\widetilde{x}_t(\lambda)) - b_{t,i}) \leqslant \sqrt{2\beta}\sqrt{D(\lambda) - D(\lambda^*)}. \tag{29}$$

В частности, $\widetilde{x}_t(\lambda^*)$ является оптимальным решением указанной задачи.

Доказательство. Неравенство

$$\mathfrak{F}(x_t) = \mathsf{E} f_t(x_t) \leqslant \mathsf{E}\left(f(x_t; \xi_t) + \langle b(\xi_t) - g(x_t, \xi_t), \lambda\rangle\right)$$
$$\leqslant \mathsf{E}\sup_{x \in B}\left(f(x; \xi_t) + \langle b(\xi_t) - g(x, \xi_t), \lambda\rangle\right) = \mathsf{E} d_t(\lambda) = D(\lambda)$$

соответствует слабой теореме двойственности. С другой стороны, по определению двойственной целевой функции,

$$d_t(\lambda) = f_t(\widetilde{x}_t(\lambda)) + \langle \nabla d_t(\lambda), \lambda\rangle,$$

так как $\nabla d_t(\lambda) = b_t - g_t(\widetilde{x}_t(\lambda))$. Таким образом,

$$\begin{aligned} \mathsf{E} f_t(x_t) - \mathsf{E} f_t(\widetilde{x}_t(\lambda)) &\leqslant \mathsf{E}(d_t(\lambda^*) - d_t(\lambda)) + \mathsf{E}\langle \nabla d_t(\lambda), \lambda\rangle \\ &\leqslant \mathsf{E}\langle \nabla d_t(\lambda^*), \lambda^* - \lambda\rangle + \mathsf{E}\langle \nabla d_t(\lambda), \lambda\rangle \\ &\leqslant \mathsf{E}\langle \nabla d_t(\lambda) - \nabla d_t(\lambda^*), \lambda\rangle \leqslant \|\lambda\| \mathsf{E}\|\nabla d_t(\lambda) - \nabla d_t(\lambda^*)\|, \end{aligned} \tag{30}$$

так как $\mathsf{E}\langle d_t(\lambda^*), \lambda^*\rangle \leqslant 0$ по лемме 6.

По лемме 5 выпуклая функция d_t является β-гладкой. Применим к ней известное неравенство [23], теорема 5.8(iii):

$$d_t(\lambda) \geqslant d_t(\lambda^*) + \langle \nabla d_t(\lambda^*), \lambda - \lambda^*\rangle + \frac{1}{2\beta}\|\nabla d_t(\lambda) - \nabla d_t(\lambda^*)\|^2.$$

Из данного неравенства следует, что

$$\begin{aligned} \mathsf{E}\|\nabla d_t(\lambda) - \nabla d_t(\lambda^*)\| &\leqslant \sqrt{2\beta}\mathsf{E}\sqrt{d_t(\lambda) - d_t(\lambda^*) - \langle \nabla d_t(\lambda^*), \lambda - \lambda^*\rangle} \\ &\leqslant \sqrt{2\beta}\sqrt{\mathsf{E}(d_t(\lambda) - d_t(\lambda^*))} \end{aligned} \tag{31}$$

в силу неравенства Иенсена, так как $\mathsf{E}\langle \nabla d_t(\lambda^*), \lambda - \lambda^*\rangle \geqslant 0$ по лемме 6. Из неравенств (30), (31) вытекает (28):

$$\mathsf{E} f_t(x_t) - \mathsf{E} f_t(\widetilde{x}_t(\lambda)) \leqslant \sqrt{2\beta}\|\lambda\|\sqrt{\mathsf{E}(d_t(\lambda) - d_t(\lambda^*))}.$$

Далее, точка $\widetilde{x}_t(\lambda^*)$ удовлетворяет ограничениям (20):

$$\mathsf{E}(g_t(\widetilde{x}_t(\lambda^*)) - b_t) = -\mathsf{E}\nabla d_t(\lambda^*) \leqslant 0.$$

Последнее неравенство вытекает из (27). С учётом этого факта неравенство (29) выводится следующим образом:

$$\begin{aligned} \mathsf{E}(g_{t,i}(\widetilde{x}_t(\lambda)) - b_{t,i}) &\leqslant \mathsf{E}(g_{t,i}(\widetilde{x}_t(\lambda) - b_{t,i}) - \mathsf{E}(g_{t,i}(\widetilde{x}_t(\lambda^*)) - b_{t,i}) \\ &\leqslant \mathsf{E}\|\nabla d_t(\lambda) - \nabla d_t(\lambda^*)\| \leqslant \sqrt{2\beta}\sqrt{\mathsf{E}(d_t(\lambda^*) - d_t(\lambda))}. \end{aligned}$$

Лемма доказана. □

Лемма 8. *Для последовательности (9) с шагом (10) справедлива оценка*

$$R_T(\lambda) := \sum_{t=1}^T d_t(\lambda_t) - \sum_{t=1}^T d_t(\lambda) \leqslant \sum_{i=1}^m \left(\frac{1}{2\gamma_i}\max_{1\leqslant t\leqslant T}(\lambda_{t,i} - \lambda_i)^2 + \gamma_i\right)\sqrt{\sum_{t=1}^T \nabla_{t,i}^2}.$$

Доказательство. В силу выпуклости d_t

$$d_t(\lambda_t) - d_t(\lambda) \leqslant \langle \nabla d_t(\lambda_t), \lambda_t - \lambda\rangle = \sum_{i=1}^m (\lambda_{t,i} - \lambda_i)\nabla_{t,i}.$$

Используя лемму 1, неравенство (13) и определение $\eta_{T,i}$, находим

$$\begin{aligned} \sum_{t=1}^T (\lambda_i - \lambda_{t,i})\nabla_{t,i} &\leqslant \frac{1}{2\eta_{T,i}}\max_{1\leqslant t\leqslant T}(\lambda_{t,i} - \lambda)^2 + \sum_{t=1}^T \frac{\eta_{t,i}}{2}\nabla_{t,i}^2 \\ &\leqslant \sum_{i=1}^m \left(\frac{1}{2\gamma_i}\max_{1\leqslant t\leqslant T}(\lambda_{t,i} - \lambda)^2 + \gamma_i\right)\sqrt{\sum_{t=1}^T \nabla_{t,i}^2}. \end{aligned}$$

Лемма доказана. □

Теорема 3. *Пусть выполнены условия лемм 4 и 5. Тогда для последовательности* $(\lambda_t)_{t=1}^T$, *построенной по методу градиентного спуска (9) с шагом (10) справедливы оценки*

$$\mathsf{E} f_T(x_T^*) - \mathsf{E} f_T(\widetilde{x}_T(\overline{\lambda}_T)) \leqslant \frac{\sqrt{2\beta}}{\sqrt{T}} \sqrt{\mathsf{E}\|\overline{\lambda}_T\|} \sqrt{\mathsf{E} R_T(\lambda^*)} = O(T^{-1/4}), \tag{32}$$

$$\mathsf{E}(g_{T,i}(\widetilde{x}_T(\overline{\lambda}_T)) - b_{T,i}) \leqslant \frac{\sqrt{2\beta}}{\sqrt{T}} \sqrt{\mathsf{E} R_T(\lambda^*)} = O(T^{-1/4}), \tag{33}$$

где

$$\overline{\lambda}_T = \frac{1}{T} \sum_{t=1}^T \lambda_t.$$

Доказательство. Выражение $f_T(\widetilde{x}_T(\lambda))$ является $\sigma(\xi_T)$-измеримым. Обозначим через $\mathscr{F}_t$ σ-алгебру, порождённую $\xi_1, \dots, \xi_t$. Поскольку $\overline{\lambda}_T$ является $\mathscr{F}_{T-1}$-измеримым, и ξ_T не зависит от $\mathscr{F}_{T-1}$, то

$$\mathsf{E} f_T(\widetilde{x}_T(\overline{\lambda}_T)) = \mathsf{E}\mathsf{E}(f_T(\widetilde{x}_T(\overline{\lambda}_T)) | \mathscr{F}_{T-1}) = \mathsf{E}\mathsf{E}\left(f_T(\widetilde{x}_T(\lambda))\big|_{\lambda=\overline{\lambda}_T} \right),$$

где в последнем равенстве использована лемма о независимости [32], лемма 2.3.4. Подставляя в (28)

$$t = T, \quad \lambda = \overline{\lambda}_T, \quad x_T = x_T^*$$

и вычисляя математическое ожидание, находим

$$\begin{aligned} \mathsf{E} f_T(x_T^*) - \mathsf{E} f_T(\widetilde{x}_T(\overline{\lambda}_T)) &\leqslant \sqrt{2\beta}\mathsf{E}\left(\|\overline{\lambda}_T\| \sqrt{D(\overline{\lambda}_T) - D(\lambda^*)}, \right) \\ &\leqslant \sqrt{2\beta} \sqrt{\mathsf{E}\|\overline{\lambda}_T\|^2} \sqrt{\mathsf{E} D(\overline{\lambda}_T) - D(\lambda^*)}. \end{aligned} \tag{34}$$

Аналогичным образом, из (29), выводим, что

$$\mathsf{E}(g_{t,i}(\widetilde{x}_t(\overline{\lambda}_T)) - b_{t,i}) \leqslant \sqrt{2\beta} \sqrt{\mathsf{E} D(\overline{\lambda}_T) - D(\lambda^*)}. \tag{35}$$

Из выпуклости D и одинаковой распределенности $d_t(\lambda^*)$ вытекает, что

$$D(\overline{\lambda}_T) \leqslant \frac{1}{T} \sum_{t=1}^T D(\lambda_t), \quad D(\lambda^*) = \frac{1}{T} \mathsf{E} \sum_{t=1}^T d_t(\lambda^*).$$

Снова используя лемму о независимости, находим, что

$$D(\lambda_t) = \mathsf{E}(d_t(\lambda_t) | \mathscr{F}_{t-1}),$$

и значит

$$\mathsf{E} D(\overline{\lambda}_T) \leqslant \frac{1}{T} \sum_{t=1}^T \mathsf{E} D(\lambda_t) = \frac{1}{T} \sum_{t=1}^T \mathsf{E} d_t(\lambda_t).$$

Таким образом,

$$\mathsf{E} D(\overline{\lambda}_T) - D(\lambda^*) \leqslant \frac{1}{T} \mathsf{E} \sum_{t=1}^T (d_t(\lambda_t) - d_t(\lambda^*)) = \frac{\mathsf{E} R_T(\lambda^*)}{T}, \tag{36}$$

и неравенства (32), (33) вытекают из (34), (35), (36). Тот факт, что правые части (32), (33) представляют собой величины порядка $T^{-1/4}$, вытекает из леммы 8, так как градиенты ∇_t равномерно ограничены и последовательность λ_t равномерно ограничена по лемме 4. □

Такие же по порядку T оценки получены в задаче о распределении ресурсов в сетях связи в работе [33]. Заметим, что в теореме 3 использовано простое усреднение двойственной последовательности λ_t, полученной методом онлайн градиентного спуска. Оценки (32), (33) хуже, чем указанные в теореме 2 оценки для рандомизированного алгоритма, хотя в теореме 3 наложено дополнительное условие сильной вогнутости функций $f_{t,j}$. В связи с этим рассмотрим ускоренный алгоритм для двойственной задачи, который эффективно использует гладкость двойственных целевых функций d_t.

Дополнительно предположим, что для оптимального решения двойственной задачи (26) известна оценка $\lambda_i^* \leqslant R$ компонент оптимального решения λ^*. К соответствующей задаче, ограниченной на гиперкуб $\Lambda = [0, R]^m$:

$$D(\lambda) = \mathsf{E} d_t(\lambda) \to \min_{\lambda \in \Lambda}, \tag{37}$$

применим алгоритм AO-FTRL (adaptive optimistic follow the regularized leader) с усреднением [34] и шагами типа AdaGrad:

$$\lambda_{t+1} \in \arg\min_{\lambda \in \Lambda} \left\{ \langle c_t, \lambda \rangle + \sum_{s=1}^{t} r_s(\lambda) \right\}, \quad t \geqslant 1, \tag{38}$$

$$c_t = -\sum_{s=1}^{t} \alpha_s \nabla d_s(\overline{\lambda}_s) - \alpha_{t+1} \nabla_t d_t(\overline{\lambda}_t), \quad \overline{\lambda}_t = \frac{1}{\alpha_1 + \cdots + \alpha_t} \sum_{s=1}^{t} \alpha_s \lambda_s, \tag{39}$$

$$r_t(\lambda) = \gamma \sum_{i=1}^{m} \frac{\eta_{t,i} - \eta_{t-1,i}}{2} (\lambda_i - \lambda_{t,i})^2,$$

$$\eta_{t,j} = \sqrt{\sum_{s=1}^{t} \alpha_s^2 (\nabla_j d_s(\overline{\lambda}_s) - \nabla_j d_{s-1}(\overline{\lambda}_{s-1}))^2}, \ t > 0, \quad \eta_0 = 0, \tag{40}$$

где α_t, γ — положительные числа, и $\lambda_1 \in \Lambda$ выбирается произвольным образом. Данный алгоритм относится к семейству алгоритмов следования за регуляризованным лидером [10], гл. 5, [24], гл. 7, [35]. Выражение $\sum_{s=1}^{t} \left\langle -\alpha_s \nabla d_s(\overline{\lambda}_s), \lambda \right\rangle$ соответствует взвешенным суммарным потерям действия λ для линеаризованной функции потерь на текущей истории. Использование градиента $\nabla d_t(\overline{\lambda}_t)$ в средней точке $\overline{\lambda}_t$ приводит к стабилизации итераций. Слагаемое $\langle -\alpha_{t+1} \nabla_t d_t(\overline{\lambda}_t), \lambda \rangle$ можно считать предсказанием потерь на следующей итерации. Понятие «оптимизма» связано с надеждой на то, что это предсказание является достаточно точным. Использование предсказаний в сочетании со стабилизацией приводит при определённых условиях к ускорению алгоритма [13, 34].

Запишем формулу (38) в более явном виде:

$$\begin{aligned}
\lambda_{t+1} &\in \arg\min_{\lambda \in \Lambda} \left\{ \sum_{i=1}^{m} c_{t,i} \lambda_i + \gamma \sum_{s=1}^{t} \sum_{i=1}^{m} \frac{\eta_{s,i} - \eta_{s-1,i}}{2} (\lambda_i - \lambda_{s,i})^2 \right\} \\
&\overset{(a)}{=} \arg\min_{\lambda \in \Lambda} \sum_{i=1}^{m} \gamma \frac{\eta_{t,i}}{2} \left(\lambda_i^2 + \frac{2}{\gamma \eta_{t,i}} \left(c_{t,i} - \gamma \sum_{s=1}^{t} (\eta_{s,i} - \eta_{s-1,i}) \lambda_{i,s} \right) \lambda_i \right) \\
&\overset{(b)}{=} \arg\min_{\lambda \in \Lambda} \sum_{i=1}^{d} \gamma \frac{\eta_{t,i}}{2} \left(\lambda_i + \frac{c_{t,i}}{\gamma \eta_{t,i}} - \sum_{s=1}^{t} \frac{\eta_{s,i} - \eta_{s-1,i}}{\eta_{t,i}} \lambda_{i,s} \right)^2 .
\end{aligned} \tag{41}$$

В равенстве (a) после раскрытия скобок отброшены слагаемые, не содержащие λ. В равенстве (b) выделен полный квадрат и так же отброшены слагаемые, не содержащие λ. Фактически, оптимизация в (41) проводится по каждой компоненте отдельно:

$$\lambda_{t+1,i} \in \arg\min_{\lambda_i\in[0,R]} \left(\lambda_i + \frac{c_{t,i}}{\gamma\eta_{t,i}} - \sum_{s=1}^{t} \frac{\eta_{s,i}-\eta_{s-1,i}}{\eta_{t,i}}\lambda_{i,s}\right)^2.$$

Таким образом,

$$\lambda_{t+1,i} = \left(-\frac{c_{t,i}}{\gamma\eta_{t,i}} + \sum_{s=1}^{t} \frac{\eta_{s,i}-\eta_{s-1,i}}{\eta_{t,i}}\lambda_{i,s}\right)^+ \wedge R, \tag{42}$$

где $a\wedge b := \min\{a,b\}$.

Положим

$$\sigma_t = \nabla d_t(\overline{\lambda}_t) - \nabla D(\overline{\lambda}_t).$$

Следующий результат является непосредственно вытекает из теоремы 5 работы [34] и леммы 7.

Теорема 4. *Пусть $\lambda_i^* \leqslant R$, $\mathsf{E}\sigma_{t,j}^2 \leqslant \overline{\sigma}_j^2$ и выполнены условия леммы 5. Тогда для алгоритма (42), (39), (40) с $\alpha_t = t$, $\gamma = 2/R$ справедливы оценки*

$$D(\lambda^*) - \mathsf{E}D(\overline{\lambda}_T) \leqslant \nu_T = C\left(\frac{\beta m R^2 + \Delta R}{T^2} + \frac{mR\max_j\overline{\sigma}_j}{\sqrt{T}}\right), \tag{43}$$

$$\mathsf{E}f_T(\widetilde{x}_T(\overline{\lambda}_T)) - \mathsf{E}f_T(x_T^*) \leqslant \sqrt{2\beta}R\sqrt{m}\sqrt{\nu_T}, \tag{44}$$

$$\mathsf{E}\left(g_{T,i}(\widetilde{x}_T(\overline{\lambda}_T)) - b_{T,i}\right) \leqslant \sqrt{2\beta}\sqrt{\nu_T}, \tag{45}$$

где $\Delta = \sum_{i=1}^m |\nabla_j D(\lambda_1)|$, и C — абсолютная константа.

Доказательство. Оценка (43) представляет собой утверждение теоремы 5(ii) работы [34], применённое к задаче (37). С учётом (43) неравенства (44), (45) выводятся из леммы 7 совершенно аналогично тому, как это сделано в доказательстве теоремы 3. □

При наличии шума ($\mathsf{E}\sigma_{t,j}^2 > 0$) оценки (44), (45) для прямой задачи имеют порядок $O(T^{-1/4})$, как и оценки теоремы 3. Улучшение может происходить при малом шуме. Наша гипотеза состоит в том, что по порядку T эта оценка неулучшаема, если действия агентов являются непосредственной реакцией на *нерандомизированные* цены, т. е. определяются формулой (4). Без этого условия, улучшение возможно: см., например, [36], теорема 2, где в близкой задаче также получены оценки порядка $O(T^{-1/2})$ для метода двойственного стохастического градиентного спуска со специальным усреднением последовательности в прямой задаче. Аналогичная оценка получена выше в теореме 2 с помощью рандомизации цен.

В детерминированном случае оценка для двойственной задачи (43) имеет порядок $O(T^{-2})$, как и для быстрого алгоритма Нестерова [37]. В общем случае эта оценка неулучшаема в классе градиентных методов [38], теорема 2.1.7. В прямой задаче из (43), (44), (45) получаем оценки

$$f(\widetilde{x}_T(\overline{\lambda}_T)) - f(x_T^*) = O(T^{-1}), \qquad g_i(\widetilde{x}_T(\overline{\lambda}_T)) - b_i = O(T^{-1})$$

такого же порядка, как и в работе [5], где был использован метод [37] для двойственной задачи. В отличие от данного метода, использованный выше алгоритм [34] не требует информации о константе гладкости β (как и, например, алгоритм [39]).

Было бы интересно установить ограниченность итераций вида (42) без проекции на гиперкуб Λ, т. е. без операции $\wedge R$. Это позволило бы избежать предположения, что $\lambda_i^* \leqslant R$.

БЛАГОДАРНОСТИ

Автор благодарен рецензентам за полезные замечания.

ФИНАНСИРОВАНИЕ РАБОТЫ

Работа выполнена при поддержке Регионального научно-образовательного математического центра Южного федерального университета, соглашение Минобрнауки РФ 075-02-2024-1427. Других источников финансирования проведения или руководства данным конкретным исследованием не было.

КОНФЛИКТ ИНТЕРЕСОВ

Автор данной работы заявляет, что у него нет конфликта интересов.

ЛИТЕРАТУРА

1. *Bertsekas D. P.* Nonlinear Programming. Belmont: Athena Scientific, 2016.
2. *Rockafellar R. T.* Problem decomposition in block-separable convex optimization: Ideas old and new // J. Nonlinear Convex Anal. 2018. V. 19, N 9. P. 1459–1474.
3. *Обен Ж.-П.* Нелинейный анализ и его экономические приложения. М.: Мир, 1988.
4. *Jalota D., Gopalakrishnan K., Azizan N., Johari R., Pavone M.* Online learning for traffic routing under unknown preferences // Proc. 26th Int. Conf. Artif. Intell. Stat. 2023. V. 206. P. 3210–3229.
5. *Beck A., Nedić A., Ozdaglar A., Teboulle M.* An $O(1/k)$ gradient method for network resource allocation problems // IEEE Trans. Control. Netw. Syst. 2014. V. 1, N 1. P. 64–73; DOI: 10.1109/TCNS.2014.2309751
6. *Воронцова Е. А., Гасников А. В., Иванова А. С., Нурминский Е. А.* Поиск равновесия по Вальрасу и централизованная распределённая оптимизация с точки зрения современных численных методов выпуклой оптимизации на примере задачи распределения ресурсов // Сиб. журн. вычисл. матем. 2019. V. 22, № 4. С. 415–436; DOI: 10.1134/S1995423919040037
7. *Dengl R.* Smart grid and demand side management // Handbook of real-time computing. 2022. P. 681-703; DOI: 10.1007/978-981-287-251-7_43
8. *Mahdavi M., Jin R., Yang T.* Trading regret for efficiency: online convex optimization with long term constraints // J. Mach. Learn. Res. 2012. V. 13, N 1. P. 2503–2528.
9. *Zinkevich M.* Online convex programming and generalized infinitesimal gradient ascent // Proc. Twentieth Int. Conf. Mach. Learn. 2003. P. 928–936.
10. *Hazan E.* Introduction to online convex optimization. Cambridge—Massachusetts: MIT Press, 2022.
11. *Cesa-Bianchi N., Conconi A., Gentile C.* On the generalization ability of on-line learning algorithms // IEEE Trans. Inf. Theory. 2004. V. 50, N 9. P. 2050–2057; DOI: 10.1109/TIT.2004.833339
12. *Shalev-Shwartz S.* Online learning and online convex optimization // Found. Trends Mach. Learn. 2012. V. 4, N 2. P. 107–194; DOI: 10.1561/2200000018
13. *Cutkosky A.* Anytime online-to-batch, optimism and acceleration // Proc. 36th Int. Conf. Mach. Learn. 2019. V. 97. P. 1446–1454.
14. *Chen T., Ling Q., Giannakis G. B.* An online convex optimization approach to proactive network resource allocation // IEEE Trans. Signal Process. 2017. V. 65, N 24. P. 6350–6364; DOI: 10.1109/TSP.2017.2750109
15. *Yu H., Neely M. J.* A low complexity algorithm with $O(\sqrt{T})$ regret and $O(1)$ constraint violations for online convex optimization with long term constraints // J. Mach. Learn. Res. 2020. V. 21, N 1. P. 1–24.
16. *Yi X., Li X., Yang T., Xie L., Johansson K.* Regret and cumulative constraint violation analysis for online convex optimization with long term constraints // Proc. 38th Int. Conf. Mach. Learn. 2021. V 139, P. 11998–12008.

17. *Jalota D., Sun H., Azizan N.* Online learning for equilibrium pricing in markets under incomplete information // arXiv. 2023; DOI: 10.48550/arXiv.2303.11522
18. *Roth A., Ullman J., Wu Z. S.* Watch and learn: optimizing from revealed preferences feedback // Proc. forty-eighth Annual ACM Symp. Theory Comput. 2016. P. 949–962.
19. *Ji Z., Mehta R., Telgarsky M.* Social welfare and profit maximization from revealed preferences // Lect. Notes Comput. Sci. 2018. V. 11316. P. 264–281; DOI: 10.1007/978-3-030-04612-5_18
20. *Roth A., Slivkins A., Ullman J., Wu Z. S.* Multidimensional dynamic pricing for welfare maximization // ACM Trans. Econ. Comput. 2020. V. 8, N 1. P.1–35; DOI: 10.1145/3381527
21. *Рокафеллар Р.* Выпуклый анализ. М.: Мир, 1973.
22. *Bertsekas D. P., Nedic A., Ozdaglar A. E.* Convex analysis and optimization. Belmont: Athena Scientific, 2003.
23. *Beck A.* First-order methods in optimization. Philadelphia: SIAM, 2017; DOI: 10.1137/1.9781611974997
24. *Orabona F.* A Modern Introduction to Online Learning // arXiv. 2019; DOI: 10.48550/arXiv.1912.13213
25. *Auer P., Cesa-Bianchi N., Gentile C.* Adaptive and self-confident on-line learning algorithms // J. Comput. Syst. Sci. 2002. V. 64, N 1. P. 48–75; DOI: 10.1006/jcss.2001.1795
26. *Aliprantis C. D., Border K. C.* Infinite dimensional analysis: a hitchhiker's guide. Berlin: Springer, 2006; DOI: 10.1007/3-540-29587-9
27. *Rockafellar R. T., Wets R. J.-B.* Variational analysis. Berlin: Springer, 2009.
28. *Шалев-Шварц Ш., Бен-Давид Ш.* Идеи машинного обучения: от теории к алгоритмам. М.: ДМК Пресс, 2019.
29. *Nesterov Y., Shikhman V.* Dual subgradient method with averaging for optimal resource allocation // Eur. J. Oper. Res. 2018. V. 270, N 3. P. 907–916; DOI: 10.1016/j.ejor.2017.09.043
30. *Shapiro A., Dentcheva D., Ruszczynski A.* Lectures on stochastic programming: modeling and theory. Philadelphia: SIAM, 2009; DOI: 10.1137/1.9780898718751
31. *Bertsekas D.P.* Stochastic optimization problems with nondifferentiable cost functionals // J. Optim. Theory Appl. 1973. V. 12, N 2. P. 218–231; DOI: 10.1007/bf00934819
32. *Shreve S. E.* Stochastic calculus for finance II, continuous-time models. N. Y.: Springer, 2004.
33. *Рохлин Д. Б.* Распределение ресурсов в сетях связи с большим числом пользователей: двойственный стохастический градиентный метод // Теория вероятн. и её примен. 2021. Т. 66, № 1, Р. 129–148; DOI: 10.1137/S0040585X97T990289
34. *Joulani P., Raj A., György A.,Szepesvári C.* A simpler approach to accelerated optimization: iterative averaging meets optimism // Proc. 37th Int. Conf. Mach. Learn. 2020 V. 119. P. 4984–4993.
35. *Joulani P., György P., Szepesvári C.* A modular analysis of adaptive (non-)convex optimization: Optimism, composite objectives, variance reduction, and variational bounds // Theor. Comput. Sci. 2020. V. 808, N 2. P. 108–138; DOI: 10.1016/j.tcs.2019.11.015
36. *Воронцова Е. А., Гасников А. В., Двуреченский П. Е., Иванова А. С., Пасечнюк Д. А.* Численные методы для задачи распределения ресурсов в компьютерной сети // Журн. вычисл. матем. и матем. физ. 2021. Т. 61, N 2. Р. 312–344; DOI: 10.31857/s0044466921020149
37. *Нестеров Ю. Е.* Метод решения задач выпуклого программирования с трудоёмкостью $O(1/k^2)$ // Доклады АН СССР. 1982. Т. 269, № 3. С. 543–547.
38. *Нестеров Ю. Е.* Введение в выпуклую оптимизацию. М.: МЦНМО, 2010.
39. *Nesterov Y.* Universal gradient methods for convex optimization problems // Math. Program. 2015. V. 152, N 1–2. P. 381–404.

SIBERIAN JOURNAL OF INDUSTRIAL MATHEMATICS

UDC 519.86

ON THE DUAL GRADIENT DESCENT METHOD FOR THE RESOURCE ALLOCATION PROBLEM IN MULTIAGENT SYSTEMS

© 2024 **D. B. Rokhlin**

Institute of Mathematics, Mechanics, and Computer Sciences and Regional Scientific and Educational
Mathematical Center of the Southern Federal University, Rostov-on-Don, 344090 Russia

E-mail: dbrohlin@sfedu.ru

Received 13.01.2024, revised 09.03.2024, accepted 17.04.2024

Abstract. We consider a sequence of block-separable convex programming problems describing the resource allocation in multiagent systems. We construct several iterative algorithms for setting the resource prices. Under various assumptions about the utility functions and resource constraints, we obtain estimates for the average deviation (regret) of the objective function from the optimal value and the constraint residuals. Here the average is understood as the expectation for independent identically distributed data and as the time average in the online optimization problem. The analysis of the problem is carried out by online optimization methods and duality theory. The algorithms considered use the information concerning the difference between the total demand and supply that is generated by agents' reactions to prices and corresponds to the constraint residuals.

Keywords: online optimization, adaptive gradient descent, duality, regret, revealed preferences.

DOI: 10.33048/SIBJIM.2024.27.206

REFERENCES

1. D. P. Bertsekas, *Nonlinear Programming* (Athena Sci., Belmont, 2016).
2. R. T. Rockafellar, "Problem decomposition in block-separable convex optimization: Ideas old and new," J. Nonlinear Convex Anal. **19** (9), 1459–1474 (2018).
3. J.-P. Aubin, *L'analyse non linéaire et ses motivations économiques* (Paris—New York—Barcelone, Masson, 1984; Mir, Moscow, 1988).
4. D. Jalota, K. Gopalakrishnan, N. Azizan, R. Johari, and M. Pavone, "Online learning for traffic routing under unknown preferences," *Proc. 26th Int. Conf. Artif. Intell. Stat.* (2023), **206**, 3210–3229.
5. A. Beck, A. NediBrc, A. Ozdaglar, and M. Teboulle, "An $O(1/k)$ gradient method for network resource allocation problems," IEEE Trans. Control. Network Syst. **1** (1), 64–73 (2014). https://doi.org/10.1109/TCNS.2014.2309751
6. E. A. Vorontsova, A. V. Gasnikov, A. S. Ivanova, and E. A. Nurminsky, "The Walrasian equilibrium and centralized distributed optimization in terms of modern convex optimization methods by an example of the resource allocation problem," Numer. Anal. Appl. **12** (4), 338–358 (2019). https://doi.org/10.1134/S1995423919040037
7. R. Dengl, "Smart grid and demand side management," *Handbook of Real-Time Computing* (2022), 681–703. https://doi.org/10.1007/978-981-287-251-7_43
8. M. Mahdavi, R. Jin, and T. Yang, "Trading regret for efficiency: Online convex optimization with long term constraints," J. Mach. Learn. Res. **13** (1), 2503–2528 (2012).

English translation is published in Journal of Applied and Industrial Mathematics, 2024, Vol. 18, No. 2, pp. 316–332.

9. M. Zinkevich, “Online convex programming and generalized infinitesimal gradient ascent,” *Proc. Twentieth Int. Conf. Mach. Learn.* (2003), 928–936.
10. E. Hazan, *Introduction to Online Convex Optimization* (MIT Press, Cambridge, MA, 2022).
11. N. Cesa-Bianchi, A. Conconi, and C. Gentile, “On the generalization ability of on-line learning algorithms,” IEEE Trans. Inf. Theory **50** (9), 2050–2057 (2004). https://doi.org/10.1109/TIT.2004.833339
12. S. Shalev-Shwartz, “Online learning and online convex optimization,” Found. Trends Mach. Learn. **4** (2), 107–194 (2012). https://doi.org/10.1561/2200000018
13. A. Cutkosky, “Anytime online-to-batch, optimism and acceleration,” *Proc. 36th Int. Conf. Mach. Learn.* (2019), **97**, 1446–1454.
14. T. Chen, Q. Ling, and G. B. Giannakis, “An online convex optimization approach to proactive network resource allocation,” IEEE Trans. Signal Process. **65** (24), 6350–6364 (2017). https://doi.org/10.1109/TSP.2017.2750109
15. H. Yu and M. J. Neely, “A low complexity algorithm with $O(\sqrt{T})$ regret and $O(1)$ constraint violations for online convex optimization with long term constraints,” J. Mach. Learn. Res. **21** (1), 1–24 (2020).
16. X. Yi, X. Li, T. Yang, L. Xie, and K. Johansson, “Regret and cumulative constraint violation analysis for online convex optimization with long term constraints,” *Proc. 38th Int. Conf. Mach. Learn.* (2021), **139**, 11998–12008.
17. D. Jalota, H. Sun, and N. Azizan, “Online learning for equilibrium pricing in markets under incomplete information,” arXiv, 2023. https://doi.org/10.48550/arXiv.2303.11522
18. A. Roth, J. Ullman, and Z. S. Wu, “Watch and learn: Optimizing from revealed preferences feedback,” *Proc. Forty-Eighth Annu. ACM Symp. Theory Comput.* (2016), 949–962.
19. Z. Ji, R. Mehta, and M. Telgarsky, “Social welfare and profit maximization from revealed preferences,” Lect. Notes Comput. Sci. **11316**, 264–281 (2018). https://doi.org/10.1007/978-3-030-04612-5_18
20. A. Roth, A. Slivkins, J. Ullman, and Z. S. Wu, “Multidimensional dynamic pricing for welfare maximization,” ACM Trans. Econ. Comput. **8** (1), 1–35 (2020). https://doi.org/10.1145/3381527
21. R. T. Rockafellar, *Convex Analysis* (Princeton Univ. Press, Princeton, 1970; Mir, Moscow, 1973).
22. D. P. Bertsekas, A. Nedic, and A. E. Ozdaglar, *Convex Analysis and Optimization* (Athena Sci., Belmont, 2003).
23. A. Beck, *First-Order Methods in Optimization* (SIAM, Philadelphia, 2017). https://doi.org/10.1137/1.9781611974997
24. F. Orabona, “A modern introduction to online learning,” arXiv, 2019. https://doi.org/10.48550/arXiv.1912.13213
25. P. Auer, N. Cesa-Bianchi, and C. Gentile, “Adaptive and self-confident on-line learning algorithms,” J. Comput. Syst. Sci. **64** (1), 48–75 (2002). https://doi.org/10.1006/jcss.2001.1795
26. C. D. Aliprantis and K. C. Border, *Infinite Dimensional Analysis: A Hitchhiker’s Guide* (Springer, Berlin, 2006). https://doi.org/10.1007/3-540-29587-9
27. R. T. Rockafellar and R. J.-B.Wets, *Variational Analysis* (Springer, Berlin, 2009).
28. Sh. Shalev-Shwartz and Sh. Ben-David, *Understanding Machine Learning: From Theory to Algorithms* (Cambridge Univ. Press, Cambridge, 2014; DMK Press, Moscow, 2019).
29. Y. Nesterov and V. Shikhman, “Dual subgradient method with averaging for optimal resource allocation,” Eur. J. Oper. Res. **270** (3), 907–916 (2018). https://doi.org/10.1016/j.ejor.2017.09.043
30. A. Shapiro, D. Dentcheva, and A. Ruszczynski, *Lectures on Stochastic Programming: Modeling and Theory* (SIAM, Philadelphia, 2009). https://doi.org/10.1137/1.9780898718751
31. D. P. Bertsekas, “Stochastic optimization problems with nondifferentiable cost functionals,” J. Optim. Theory Appl. **12** (2), 218–231 (1973). https://doi.org/10.1007/bf00934819
32. S. E. Shreve, *Stochastic Calculus for Finance II, Continuous-Time Models* (Springer, New York, 2004).
33. D. B. Rokhlin, “Resource allocation in communication networks with large number of users: The dual stochastic gradient method,” Theory Probab. Appl. **66** (1), 105–120 (2021). https://doi.org/10.1137/S0040585X97T990289

34. P. Joulani, A. Raj, A. György, and C. SzepesvB́ŕari, “A simpler approach to accelerated optimization: Iterative averaging meets optimism,” *Proc. 37th Int. Conf. Mach. Learn.* (2020), **119**, 4984–4993.

35. P. Joulani, P. György, and C. Szepesvári, “A modular analysis of adaptive (non-)convex optimization: Optimism, composite objectives, variance reduction, and variational bounds,” Theor. Comput. Sci. **808** (2), 108–138 (2020). https://doi.org/10.1016/j.tcs.2019.11.015

36. E. A. Vorontsova, A. V. Gasnikov, P. E. Dvurechensky, A. S. Ivanova, and D. A. Pasechnyuk, “Numerical methods for the resource allocation problem in a computer network,” Comput. Math. Math. Phys. **61** (2), 297–328 (2021). https://doi.org/10.1134/S0965542521020135

37. Yu. E. Nesterov, “A method for solving convex programming problems with complexity $O(1/k^2)$,” Dokl. Akad. Nauk SSSR **269** (3), 543–547 (1982) [in Russian].

38. Yu. E. Nesterov, *Introduction to Convex Optimization* (MTsNMO, Moscow, 2010) [in Russian].

39. Y. Nesterov, “Universal gradient methods for convex optimization problems,” Math. Program. **152** (1–2), 381–404 (2015).