



Российская Академия Наук

**А
Т
АВТОМАТИКА
И ДАЛЕКОМЕРИКА
AUTOMATION AND REMOTE CONTROL**

Журнал основан в 1936 году
Выходит 12 раз в год

8

АВГУСТ

Москва

2025

Учредители журнала:

Отделение энергетики, машиностроения, механики и процессов управления РАН,
Институт проблем управления им. В.А. Трапезникова РАН (ИПУ РАН),
Институт проблем передачи информации им. А.А. Харкевича РАН (ИППИ РАН)

Главный редактор:

Галяев А.А.

Заместители главного редактора:

Рубинович Е.Я., Хлебников М.В.

Ответственный секретарь:

Самохин А.С.

Редакционный совет:

Васильев С.Н., Желтов С.Ю., Каляев И.А., Кузнецов О.П., Кулешов А.П.,
Куржанский А.Б., Мартынюк А.А. (Украина), Пешехонов В.Г., Попков Ю.С.,
Федосов Е.А., Черноусько Ф.Л.

Редакционная коллегия:

Алескеров Ф.Т., Арутюнов А.В., Бахтадзе Н.Н., Бобцов А.А., Виноградов Д.В.,
Вишневский В.М., Воронцов К.В., Граничин О.Н., Жилякова Н.Ю.,
Каравай М.Ф., Кибзун А.И., Краснова С.А., Крищенко А.П., Кузнецов Н.В.,
Кушнер А.Г., Лазарев А.А., Ляхов А.И., Маликов А.И., Матасов А.И.,
Меерков С.М. (США), Мещеряков Р.В., Миллер Б.М., Михальский А.И.,
Моржин О.В., Мунасыпов Р.А., Назин А.В., Немировский А.С. (США),
Новиков Д.А., Олейников А.Я., Пакшин П.В., Пальчунов Д.Е.,
Поляков А.Е. (Франция), Протасов В.Ю., Рапопорт Л.Б., Родионов И.В.,
Сельвесюк Н.И., Соболевский А.Н., Степанов О.А., Филимонюк Л.Ю.,
Фрадков А.Л., Хоров Е.М., Цыбаков А.Б. (Франция), Чеботарев П.Ю.,
Чхартишвили А.Г., Щербаков П.С.

Адрес редакции: 117997, Москва, Профсоюзная ул., 65

Тел./факс: 8 (495) 198-17-20, доб. 1443

Электронная почта: redacsia@ipu.ru, ait@pran.ru

Сайт: <https://aitras.ru>

Зав. редакцией *Е.А. Мартехина*

Москва

ФГБУ «Издательство «Наука»

К 90-летию Б.Т. Поляка



(04.05.1935–03.02.2023)

В этом году исполняется 90 лет со дня рождения блестящего математика и замечательного человека, Бориса Теодоровича Поляка, ушедшего из жизни два года назад.

Пионер теории и методов оптимизации в СССР, получивший всемирное признание уже в середине 1970-х гг., он более полувека проработал в Институте проблем управления им. В.А. Трапезникова. Круг научных интересов Бориса Теодоровича не ограничивался задачами оптимизации. Он с увлечением работал над такими, на первый взгляд, далекими друг от друга проблемами, как управление хаосом, стабилизация космического корабля в точках Лагранжа, эффект всплеска решений дифференциальных и разностных уравнений, синхронизация осцилляторов, рандомизированные версии задачи PageRank, устойчивость энергетических сетей.

Работы Поляка в области математической оптимизации и теории управления оказали огромное влияние на развитие этих разделов науки и нашли применения, порой неожиданные, в практических задачах. Из обилия заме-

чательных результатов, полученных Борисом Теодоровичем, выделим лишь самые яркие:

- Метод тяжелого шарика
- Метод стохастической аппроксимации с усреднением
- Метод сопряженных градиентов
- Алгоритмы невыпуклой оптимизации при условии Поляка–Лоясиевича
- Адаптивный выбор длины шага в субградиентных методах
- Метод последовательного проектирования
- Метод условного градиента
- Метод Ньютона с кубической регуляризацией
- Сертификаты выпуклости/невыпуклости квадратичных отображений
- Результаты по параметрической робастности линейных систем (в том числе годограф Цыпкина–Поляка)
- Новые версии методов случайных блужданий
- Результаты по эллипсоидальному оцениванию и методу инвариантных эллипсоидов.

Широта научных взглядов Бориса Поляка иллюстрируется и разнообразием журналов, в которых ему доводилось печататься: «Журнал вычислительной математики и математической физики», «Математический сборник», «Journal of Optimization Theory and Applications», «Mathematical Programming», «SIAM Journal on Control and Optimization», «Теория вероятностей и ее применения», «Автоматика и телемеханика», «International Journal of Robust and Nonlinear Control», «IEEE Transactions on Automatic Control», «Automatica», «European Journal of Control», «Systems and Control Letters» и многие другие.

Личные качества Бориса Теодоровича – оптимизм, человеколюбие и острый добрый юмор – всегда притягивали к нему коллег и молодежь. Десятки его учеников и их последователей активно и плодотворно работают в университетах и исследовательских центрах во многих уголках планеты, а созданная им Традиционная молодежная научная школа «Управление, информация и оптимизация» с неизменным успехом проводится вот уже более 15 лет.

В честь юбилейной даты было решено подготовить специальный выпуск журнала «Автоматика и телемеханика», в который вошли статьи учеников и коллег Бориса Теодоровича из разных городов России (Москва, Санкт-Петербург, Нижний Новгород, Сыктывкар), а также из Англии, Испании, Италии, Франции и США. Тематика статей выпуска необычайно широка: от классических методов оптимизации до вероятностных аппроксимаций множеств, от оптимизации потоков в сетях до робастного оценивания и робастного управления линейными системами. Несмотря на такой широкий спектр тем, все они так или иначе связаны с научными интересами Б.Т. Поляка.

С глубокой признательностью привожу список авторов, любезно согласившихся предоставить свои работы для настоящего выпуска, и рецензентов, чей экспертный вклад трудно переоценить:

С.С. Аблаев, И.А. Акинфиев, Т. Аламо, М.С. Алкуса, А. Ахаван, М.В. Балашов, Я. Бекри, Р.С. Бирюков, С. Бойд, А.В. Гасников, О.Н. Граничин, Ф. Даббене, П.Е. Двуреченский, М.М. Коган, А.Л. Лукашевич, М. Маммарелла, Т. Шикаке Мапунгвана, Н.Е. Машалов, В. Мирасьерра, С.А. Назин, А.С. Немировский, С.Э. Парсегов, О.С. Савчук, В.Ф. Соколов, Ф.С. Стонякин, Е.Ю. Тарасова, М.Н. Федотов, Р. Хильдебранд, А.Б. Цыбаков, Ф. Чжан, А.Б. Юдицкий, Д.В. Ярмошик.

Отдельной благодарности заслуживает А.Ю. Мазуров, приложивший немалые усилия к переводу статей с/на английский язык.

Как и в случае с прошлогодним спецвыпуском, посвященным Борису Теодоровичу, ограничение на объем одного номера журнала «Автоматика и телемеханика» вынуждает нас публиковать три статьи, принятые в нынешний выпуск, в следующем номере.

Ответственный за выпуск

П.С. Щербаков

© 2025 г. И.А. АКИНФИЕВ (i@iakinfiev.ru)
(Санкт-Петербургский государственный университет),
О.Н. ГРАНИЧИН, д-р физ.-мат. наук (o.granichin@spbu.ru)
(Санкт-Петербургский государственный университет;
Институт проблем машиноведения РАН, Санкт-Петербург),
Е.Ю. ТАРАСОВА (elizaveta.tarasova@spbu.ru)
(Санкт-Петербургский государственный университет)

ПОИСКОВЫЙ МЕТОД СТОХАСТИЧЕСКОЙ НЕСТАЦИОНАРНОЙ ОПТИМИЗАЦИИ ФУНКЦИИ С ГЕЛЬДЕРОВСКИМ ГРАДИЕНТОМ¹

В статье рассматривается поисковый метод стохастической оптимизации с возмущением на входе, предназначенный для отслеживания изменений точки минимума функции (трекинга) с гельдеровским градиентом в условиях наблюдений при почти произвольных неизвестных ограниченных помехах (unknown-but-bounded noise). Подобные методы используются в задачах адаптивного управления (энергетика, логистика, робототехника, трекинг целей), оптимизации зашумленных систем (биомоделирование, физические эксперименты) и онлайн-обучения с дрейфом параметров данных (финансы, потоковая аналитика).

В качестве апробации алгоритма исследуется эффективность его работы в условиях, имитирующих отслеживание эволюции человеческих ожиданий в задачах обучения с подкреплением на основе обратной связи от человека и при отслеживании центра кластера задач в системах массового обслуживания. Поисковые методы с возмущениями на входе активно развивались в работах Б.Т. Поляка с 1990 г.

Ключевые слова: трекинг, возмущение на входе, рандомизация, стохастическая оптимизация, безградиентные методы, обучение с подкреплением на основе обратной связи от человека, системы массового обслуживания, неизвестные, но ограниченные помехи.

DOI: 10.31857/S0005231025080013, EDN: USSKZZ

1. Введение

Задача минимизации функции (или функционала) $f(x)$ лежит в основе решения множества практических задач от управления техническими системами до машинного обучения. Аналитические решения часто недоступны из-за высокой размерности, нелинейностей или отсутствия явного вида функции. Даже при аналитическом задании функции их практическая применимость ограничена вычислительными ресурсами, неточностями измерений или ошибками округления. Традиционные итерационные градиентные

¹ Теоретическая часть работы, разделы 1–4, выполнены при финансовой поддержке РФФ в ИПМАШ РАН, (проект № 23-41-00060), практическая часть работы, разделы 5–6, выполнены при финансовой поддержке СПбГУ, шифр проекта 121061000159-6.

методы эффективны для поиска минимума гладких или дифференцируемых функций. Однако в реальных задачах часто возникают ситуации, когда вычисление градиента затруднено или невозможно. Обычно целевая функция подвержена стохастическим возмущениям, либо ее явный вид неизвестен. На практике оптимизируемая функция часто задается некоторым оракулом, обращаясь к которому с запросами (аргументами функции) можно получить возможные реализации. Доступность измерений непосредственно градиента возможна при реализации специальных измерительных устройств для той или иной задачи или за счет конечно-разностных аппроксимаций, которые неработоспособны при высоком уровне помех в получаемых измерениях. В таких случаях требуются альтернативные подходы, не зависящие от информации о градиенте.

Значительный вклад в развитие теории и методов стохастической оптимизации внес Б.Т. Поляк и вся научная группа, с которой он работал. Их исследования охватывают широкий спектр вопросов, включая градиентные методы [1], псевдоградиентные алгоритмы адаптации и обучения [2–4] и методы ускорения сходимости [5–7]. Две статьи [8, 9] дают и сейчас исчерпывающие ответы при анализе сходимости итеративных стохастических алгоритмов в общем случае в терминах среднеквадратических отклонений, а также в линейном случае в терминах матриц ковариаций ошибки. Предложенный в 1990 г. новый поисковый метод стохастической аппроксимации [10] не только развивает общее направление алгоритмов случайного поиска [11], но и существенно продвигает вперед всю общую теорию итеративных алгоритмов оптимизации: в статье показано, что при наблюдении значений оптимизируемой функции с помехами предложенный алгоритм имеет асимптотически оптимальную скорость сходимости в том смысле, что невозможно найти более быстрый алгоритм среди всевозможных итеративных алгоритмов оптимизации для достаточно широкого класса функций. Ранее похожий алгоритм был предложен в [12], и для него была обоснована состоятельность оценок при наблюдении на фоне почти произвольных помех. В англоязычной литературе похожие методы получили название SPSA (Simultaneous Perturbation Stochastic Approximation) [13, 14]. Важной особенностью этих безградиентных методов является то, что вне зависимости от размерности задачи на каждой итерации надо вызывать оракул всего один или два раза с аргументами, выбираемыми на случайной прямой (рандомизация алгоритма), проходящей через точку очередной оценки. Подробный анализ истории развития, а также свойств оценок поисковых алгоритмов стохастической аппроксимации с возмущением на входе, дан в [15–17].

Важным ограничением классических итерационных методов стохастической оптимизации нулевого порядка (не использующих значения градиента), таких как процедура Кифера–Вольфовица [18] в многомерном случае, является необходимость многократного вычисления функции на каждой итерации. Это становится особенно непрактичным в динамических средах, где целевая функция $f_n(x)$ изменяется со временем. Подобная ситуация возника-

ет, например, в задачах оптимизации систем реального времени. Оказалось, что методы типа предложенных ранее поисковых алгоритмов стохастической оптимизации с возмущением на входе в этой ситуации продолжают быть работоспособными при замене уменьшающихся со временем размеров шагов на постоянные [19, 20]. Позже удалось сформулировать и обосновать свойства распределенного алгоритма такого типа, совмещенного с алгоритмом консенсуса [20].

На практике [21, 22] часто сталкиваются со статистическими неопределенностями, которые не имеют второго статистического момента. Например, стабильные распределения, в частности, Леви–Парето лучше описывают цены на акции и товары, чем нормальные распределения [23]. В [24] исследованы свойства оценок алгоритма типа SPSSA в таких условиях. В статье эти исследования распространены на случай оптимизации нестационарного функционала среднего риска.

2. Постановка задачи

Пусть время является дискретным и определяется номером шага (итерации) $n = 0, 1, \dots$, $\{F_n(\cdot, \cdot) : \mathbb{R}^d \times \mathbb{R}^q \rightarrow \mathbb{R}\}$ – набор функций от двух векторных переменных, дифференцируемых по первому аргументу, на каждом шаге n в известных (выбираемых) точках x_n (план эксперимента) производятся наблюдения (измеряются значения)

$$(1) \quad y_n = F_n(x_n, w_n) + v_n,$$

где w_n – неконтролируемые возмущения, заданные на некотором вероятностном пространстве Ω и имеющие одинаковое, неизвестное распределение $P_w(\cdot)$, v_n – помехи в наблюдениях (возможно и неслучайные).

Обозначим \mathcal{F}_{n-1} – σ -алгебру всех вероятностных событий, которые реализовались до момента n , \mathbb{E} – символ математического ожидания, $\mathbb{E}_{\mathcal{F}_{n-1}}$ – символ условного математического ожидания при условии σ -алгебры \mathcal{F}_{n-1} .

Рассмотрим задачу минимизации нестационарного функционала среднего риска:

$$(2) \quad f_n(x) = \mathbb{E}_{\mathcal{F}_{n-1}} F_n(x, w) = \int_{\mathbb{R}^q} F_n(x, w) P_w(dw) \rightarrow \min_x.$$

Требуется оценить точку минимума θ_n функции $f_n(x)$, т.е. найти

$$\theta_n = \arg \min_x f_n(x).$$

Точность оценки x точек θ_n характеризуется с помощью скалярных функций Ляпунова

$$V_n(x) = \|x - \theta_n\|^{\rho+1} = \sum_{i=1}^n |x^{(i)} - \theta_n^{(i)}|^{\rho+1},$$

где θ_n – искомые векторы, $\rho \in (0, 1]$ – показатель Гельдера градиентов функций $V_n(x)$. Далее в статье будут использоваться обозначения $\|\cdot\|_{\rho+1}$ для нормы $l_{\rho+1}$ и $\langle \cdot, \cdot \rangle$ для скалярного произведения в \mathbb{R}^d .

Для характеристики поведения оценок точек минимума нестационарного функционала (2) введем два определения.

Определение 1. Последовательность оценок $\hat{\theta}_n$ точек минимума θ_n называется $l_{\rho+1}$ -стабилизированной, если существует такое $C > 0$, что

$$\mathbb{E}V_n(\hat{\theta}_n) \leq \text{const}, \quad \forall n.$$

Определение 2. Число L называется асимптотической верхней границей ошибок оценивания по $l_{\rho+1}$ -норме, если для последовательности оценок $\{\hat{\theta}_n\}$ точек минимума θ_n выполняется:

$$\overline{\lim}_{n \rightarrow \infty} \mathbb{E}V_n(\hat{\theta}_n) \leq L < \infty.$$

Далее будем рассматривать задачу о построении последовательности стабилизирующихся оценок $\{\hat{\theta}_n\}$, в рамках определения 2, при следующих условиях, выполняющихся при любом $n > 0$.

(А) Функции $f_n(\cdot)$ сильно выпуклые по первому аргументу:

$$\langle \nabla V_n(x), \nabla f_n(x) \rangle \geq \mu V_n(x).$$

(В) При любом w градиенты функций $\nabla F_n(\cdot, w)$ удовлетворяют условию:

$$\|\nabla F_n(x, w) - \nabla F_n(y, w)\|_1 \leq M \|x - y\|_\rho^p$$

с некоторой константой M .

(С) Локальное свойство Лебега: $\forall x \in \mathbb{R}^d \exists$ окрестность U_x точки x и функция $\Phi_x(w)$, такие что: $\mathbb{E}\Phi_x(w) < \infty$ и $\|\nabla F_n(x', w)\|_2 \leq \Phi_x(w) \quad \forall x' \in U_x$.

(D) Скорость дрейфа точки минимума ограничена следующими условиями

$$a: \quad \|\theta_n - \theta_{n-1}\|_1 \leq A,$$

либо, если $\{\theta_n\}$ представляют собой последовательность случайных величин, то

$$\mathbb{E}_{\mathcal{F}_{n-1}} \|\theta_n - \theta_{n-1}\|_{\rho+1}^{\rho+1} \leq A^{\rho+1},$$

а также

$$b: \quad \mathbb{E}_{\mathcal{F}_{n-1}} \|\nabla_x F_n(x, w) - \nabla_x F_{n-1}(x, w)\|_1 \leq B \|x - \theta_{n-1}\|_1^\rho,$$

$$c: \quad \mathbb{E}_{\mathcal{F}_{n-1}} \|\nabla_x F_n(\theta_n, w_n)\|_{\rho+1}^{\rho+1} \leq C,$$

$$d: \quad \mathbb{E}_{\mathcal{F}_{2n-2}} |F_{2n}(x, w_{2n}) - F_{2n-1}(x, w_{2n-1})|^{\rho+1} \leq DV_{2n-2}(x) + E.$$

(E) Для помех наблюдения v_n выполнены условия:

$$|v_{2n} - v_{2n-1}| \leq \sigma_v,$$

либо, если они представляют собой последовательность случайных величин, то

$$\mathbb{E}_{\mathcal{F}_{2n-2}} \{|v_{2n} - v_{2n-1}|^{\rho+1}\} \leq \sigma_v^{\rho+1}.$$

Заметим, что последнему условию удовлетворяют любые детерминированные ограниченные последовательности $\{v_n\}$. Условие (С) обеспечивает возможность перестановки операций интегрирования и дифференцирования при обосновании стабилизируемости оценок. Ограничения типа (D) включают как дрейф типа случайных блужданий, так и направленный дрейф в определенную сторону. Например, в [1] приводится ограничение на основе (D):

$$\theta_n = \theta_{n-1} + a + \xi_n,$$

где ξ_n – центрированная случайная величина и a – тренд. Стабилизируемость оценок алгоритма поиска минимума в условиях (D) означает применимость его к широкому классу различных задач.

3. Поисковый рандомизированный алгоритм оценивания

Пусть $\{\Delta_n\}$ – последовательность пробных одновременных возмущений, подаваемых на вход алгоритма оценивания, является некоторой реализацией последовательности независимых бернуллиевских случайных векторов из \mathbb{R}^d , у которых каждая компонента независимо принимает с вероятностью $\frac{1}{2}$ значения, равные $\pm \frac{1}{\sqrt{d}}$. Выберем некоторый начальный вектор $\theta_0 \in \mathbb{R}^d$. Будем оценивать последовательность точек минимума $\{\theta_n\}$ последовательностью $\{\hat{\theta}_n\}$, определяемой алгоритмом стохастической оптимизации с пробным одновременным возмущением на входе, который имеет следующий вид:

$$(3) \quad \begin{cases} \hat{\theta}_{2n-1} = \hat{\theta}_{2n-2}, \\ x_{2n} = \hat{\theta}_{2n-2} + \beta \Delta_n, \quad x_{2n-1} = \hat{\theta}_{2n-2} - \beta \Delta_n, \\ \hat{\theta}_{2n} = \hat{\theta}_{2n-2} - \frac{\alpha}{2\beta} \Delta_n (y_{2n} - y_{2n-1}), \end{cases}$$

где α и β – параметры размеров шага алгоритма. Для обоснования стабилизации оценок алгоритма (3) будем считать, что

(F) случайные векторы Δ_n и w_{2n}, w_{2n-1} не зависят между собой, а также от \mathcal{F}_{n-1} . Если $\{v_n\}$ предполагаются случайной природы, то Δ_n не зависят от v_{2n}, v_{2n-1} .

4. Стабилизация оценок

Обозначим $H = A + \alpha\beta M$, где A и M – константы ограничения скорости дрейфа точки и ограничение изменения градиентов соответственно.

Теорема 1. Пусть выполнены условия (A)–(F) и параметры алгоритма α, β выбраны такими, чтобы константа $K > 0$, определенная далее при доказательстве теоремы, была меньше единицы. Тогда для любого начального приближения $\hat{\theta}_0$ с $E\|\hat{\theta}_0 - \theta_0\|^{\rho+1} < \infty$ оценки алгоритма (4) стабили-

зируются в следующем смысле:

$$\overline{\lim}_{n \rightarrow \infty} \mathbb{E} \|\hat{\theta}_n - \theta_n\|_{\rho+1} \leq \left(\frac{L}{K} \right)^{\frac{1}{\rho+1}},$$

где L также определена в конце доказательства теоремы.

Условия (A)–(C), (E)–(F) являются стандартными для доказательства состоятельности оценок алгоритмов стохастической оптимизации с возмущением на входе [18]. Ранее факт среднеквадратичной стабилизации оценок алгоритма (3) был доказан в [19] при более жестких ограничениях.

Доказательство теоремы 1 и определение констант K и L приводится в Приложении 1.

5. Моделирования в RLHF-сценарии

В задачах обучения с подкреплением на основе обратной связи от человека (Reinforcement Learning from Human Feedback, RLHF) ключевой вызов заключается в работе с зашумленными и нестабильными данными [26, 27]. Человеческие оценки часто содержат случайные ошибки и могут изменяться со временем, что усложняет процесс оптимизации. В частности, в задачах, связанных с тонкой настройкой языковых моделей (Large Language Models, LLM), RLHF используется для улучшения качества генерации текста, согласования с предпочтениями пользователей и минимизации нежелательного поведения моделей. Однако субъективность и изменчивость человеческих оценок создают значительные трудности для традиционных методов оптимизации [28–30].

5.1. Модель

В моделировании исследуется эффективность поискового алгоритма в условиях, приближенных к реальным: при наличии шума с тяжелыми хвостами (распределение Парето) и дрейфа предпочтений [28, 30], имитирующего эволюцию человеческих ожиданий. Рассматриваются три сценария: умеренный дрейф, почти стационарные предпочтения и стационарные предпочтения с асимметричным шумом. Это позволяет оценить устойчивость и адаптивность поискового алгоритма в условиях, характерных для RLHF, и определить его применимость для задач, связанных с обучением LLM и других систем, где человеческая обратная связь играет ключевую роль.

Цель моделирования – проверить способность RLHF-агента адаптироваться к модели вознаграждения, которая формируется на основе зашумленных и изменяющихся человеческих оценок. Далее:

- моделируется шум с тяжелыми хвостами (Pareto-распределение), отражающий неопределенность и редкие, но значительные отклонения в оценках;
- вводится модель дрейфа предпочтений, имитирующий постепенное изменение человеческих ожиданий;
- все функции и параметры вводятся в условиях (A)–(F) части 2.

Агент должен минимизировать расхождение между своей оценкой параметра и истинным значением, задаваемым моделью вознаграждения, несмотря на шум и динамику целевых предпочтений, для минимизации используется поисковый алгоритм.

Модель вознаграждения, основанная на RLHF, задается следующим образом:

$$(4) \quad F_n(\mathbf{x}) = - \sum_{i=1}^m (x_i - x_n^*)^{1,35},$$

где x_n^* – целевой параметр, который дрейфует со временем n

$$x_n^* = x_{n-1}^* + \delta, \quad x_0^* = 5,$$

отражая изменение предпочтений.

При выборе x_n на основе обратной связи получаем

$$y_n = F_n(\mathbf{x}) + v_n,$$

где v_n – шум, моделирующий неопределенность в обратной связи. Для моделирования использовались два вида шума:

- симметричный шум: $v_i = Z_i \times \text{sign}_i$, где $Z_i \sim \text{Pareto}(\beta, \sigma)$, $\text{sign}_i \sim \text{Uniform}(\{-1, 1\})$;
- асимметричный шум: $v_i = Z_i$, где $Z_i \sim \text{Pareto}(\beta, \sigma)$, что может отражать склонность к завышению оценок.

В табл. 1 представлены основные параметры, использованные при проведении численного моделирования. Они охватывают структуру эксперимента, настройки алгоритма (гиперпараметры), а также характеристики шума и сценарии дрейфа, имитирующие условия нестабильной обратной связи.

Таблица 1. Параметры моделирования

Параметр	Описание	Значение
Начальная оценка агента	Начальная точка для обучения	$\hat{\theta}_0 = 0$
Число итераций	Количество шагов адаптации	$N = 1000$
Число запусков	Количество независимых экспериментов	$m = 1000$
Гиперпараметры		
Шаг обучения	Консервативный шаг для устойчивости	$\gamma = 0,05$
Размер возмущения	Амплитуда для оценки градиента	$c = 0,1$
Характеристики шума		
Параметр формы	Определяет тяжесть хвостов	$\beta = 1,6$
Масштаб	Интенсивность отклонений	$\sigma = 2,0$
Скорость дрейфа	Умеренный дрейф параметров	$\delta = 0,01$
	Почти стационарный режим	$\delta = 0,0001$
Тип шума	Случайные отклонения	Симметричный
	Систематическое смещение	Асимметричный

5.2. Сценарии моделирования

Для анализа поисковой адаптивности рассматриваются три сценария:

1. Умеренный дрейф предпочтений ($\delta = 0,01$) и симметричный шум (здесь и далее: шум с симметричным распределением). Имитирует постепенное изменение целевых значений при наличии случайных ошибок в оценках.
2. Почти стационарные предпочтения ($\delta = 0,0001$) и симметричный шум. Проверка точности настройки в условиях, близких к стабильным.
3. Стационарные предпочтения с асимметричным шумом ($\delta = 0,0001$) и асимметричный шум (здесь и далее: шум с несимметричным распределением). Отражает систематическое искажение обратной связи, например, постоянное завышение оценок.

5.3. Процесс адаптации агента

Агент обновляет свою оценку параметра $\hat{\theta}$ на основе наблюдаемых значений y (вознаграждений), полученных из модели. Алгоритм следует итерационной схеме, описанной в формуле (3):

На каждой четной итерации $k = 2n$ (где $n = 1, 2, \dots$):

1. Используется оценка с предыдущего четного шага, $\hat{\theta}_{2n-2}$ (для $n = 1$ используется $\hat{\theta}_0$).
2. Создается случайный вектор возмущений Δ_n , где каждая компонента независимо принимает значение $+1$ или -1 с вероятностью $0,5$.
3. Формируются две точки согласно формуле (3):

$$\begin{aligned}x_{2n} &= \hat{\theta}_{2n-2} + \beta \Delta_n, \\x_{2n-1} &= \hat{\theta}_{2n-2} - \beta \Delta_n.\end{aligned}$$

4. Наблюдаются значения (вознаграждения) в возмущенных точках: y_{2n} (соответствующее x_{2n}) и y_{2n-1} (соответствующее x_{2n-1}). Эти y включают как истинное значение функции, так и шум ($y_n = F_n(x_n, w_n) + v_n$ в терминах статьи).
5. Обновление оценки: Оценка $\hat{\theta}$ обновляется по формуле, аналогичной третьей строке системы (3), но со знаком “+”, так происходит максимизация:

$$\hat{\theta}_{2n} \leftarrow \hat{\theta}_{2n-2} + \frac{\alpha}{2\beta} \Delta_n (y_{2n} - y_{2n-1}).$$

На каждой нечетной итерации $k = 2n - 1$:

1. Копирование оценки: $\hat{\theta}_{2n-1} \leftarrow \hat{\theta}_{2n-2}$.

5.4. Проверка условий (A)–(F) для моделирования в RLHF-сценарии

(A) Сильная выпуклость функции $f_n(\mathbf{x})$.

$$\nabla f_n(\mathbf{x}) = -\nabla F_n(\mathbf{x}) = -[1,35(x_1 - x_n^*)^{0,35}, \dots, 1,35(x_m - x_n^*)^{0,35}]^\top,$$

$$\nabla V_n(\mathbf{x}) = [(\rho + 1)\text{sign}(x_1 - x_n^*)|x_1 - x_n^*|^\rho, \dots, (\rho + 1)\text{sign}(x_m - x_n^*)|x_m - x_n^*|^\rho]^\top,$$

$$\langle \nabla V_n(\mathbf{x}), \nabla f_n(\mathbf{x}) \rangle = -1,35(\rho + 1) \sum_{i=1}^m |x_i - x_n^*|^{\rho+0,35}.$$

Используя неравенство $|x_i - x_n^*|^{\rho+0,35} \geq |x_i - x_n^*|^{\rho+1} a^{-0,65}$, где $a \leq |x_i - x_n^*|$, получаем:

$$\sum_{i=1}^m |x_i - x_n^*|^{\rho+0,35} \geq a^{-0,65} \sum_{i=1}^m |x_i - x_n^*|^{\rho+1} = a^{-0,65} V_n(\mathbf{x}).$$

Таким образом,

$$\langle \nabla V_n(\mathbf{x}), \nabla f_n(\mathbf{x}) \rangle \leq -1,35(\rho + 1) a^{-0,65} V_n(\mathbf{x}),$$

т.е. условие вида $\langle \nabla V_n(\mathbf{x}), \nabla f_n(\mathbf{x}) \rangle \geq \mu V_n(\mathbf{x})$ выполнено при $\mu = -1,35(\rho + 1) a^{-0,65} < 0$. При минимизации $f_n(\mathbf{x})$ условие сильной выпуклости в смысле данного скалярного неравенства выполняется при $\mu < 0$.

(B) Гельдерова непрерывность градиента.

Для функцию вознаграждения $F_n(x)$ градиент имеет вид:

$$\nabla F_n(x) = -1,35 \cdot [(x_1 - x_n^*)^{0,35}, \dots, (x_m - x_n^*)^{0,35}]^\top.$$

Для любого i компонентная разность градиентов оценивается как:

$$|\partial_i F_n(x) - \partial_i F_n(y)| = 1,35 |(x_i - x_n^*)^{0,35} - (y_i - x_n^*)^{0,35}| \leq 1,35 M' |x_i - y_i|^{0,35},$$

где M' – константа Гельдера, зависящая от ограниченной области, на которой определены x_i и x_n^* .

Суммируя по всем координатам, получаем:

$$\|\nabla F_n(x) - \nabla F_n(y)\|_1 \leq 1,35 M' \sum_{i=1}^m |x_i - y_i|^{0,35} = M \|x - y\|_{0,35}^{0,35},$$

где $M = 1,35 M'$.

Заметим, что M' можно оценить явно. M' ограничена сверху на отрезке $s \in [\varepsilon, R]$ при $\varepsilon > 0$. Например, если предполагается, что $|x_i - x_n^*| \geq 1$, то

$$M' = \max_{s \in [1, R]} 0,35 s^{-0,65} = 0,35, \quad \Rightarrow \quad M = 1,35 \cdot 0,35 \approx 0,4725.$$

(C) Локальное условие Лебега.

Фиксируем точку x и рассмотрим ее окрестность $U_x = B(x, \varepsilon)$ для некоторого $\varepsilon > 0$. Тогда для любой точки $x' \in U_x$:

$$\|\nabla F_n(x', w)\|_2^2 = 1,35^2 \sum_{i=1}^m |x'_i - x_n^*|^{0,7} \leq 1,35^2 m R^{0,7},$$

где $R = \sup_{x' \in U_x} \max_i |x'_i - x_n^*| < \infty$ – конечен по построению окрестности U_x .

Тогда можно положить $\Phi_x(w) = 1,35\sqrt{m}R^{0,35}$, которая не зависит от w и, следовательно, $\mathbb{E}\Phi_x(w) = \Phi_x(w) < \infty$. Условие (C) выполнено.

(D) Ограниченность скорости дрейфа точки минимума.

(D-a) *Ограниченность изменения минимума.*

Так как $\theta_n = x_n^* \mathbf{1}$, а $x_n^* = x_{n-1}^* + \delta$, имеем $\|\theta_n - \theta_{n-1}\|_1 = \|\delta \mathbf{1}\|_1 = m\delta$. Следовательно, условие (D-a) выполнено при $A = m\delta$.

(D-b) *Ограниченность изменения градиента.*

Пусть $r_i = x_i - x_{n-1}^*$, тогда:

$$|\partial_i F_n(x) - \partial_i F_{n-1}(x)| \leq 1,35M'|\delta|^{0,35},$$

где M' – гельдеровская константа функции $s^{0,35}$ на допустимом компакте.

Суммируя по i , получаем:

$$\|\nabla_x F_n(x) - \nabla_x F_{n-1}(x)\|_1 \leq 1,35M'm|\delta|^{0,35}.$$

Обозначим $R = \inf_{x \neq \theta_{n-1}} \|x - \theta_{n-1}\|_1 > 0$, тогда $\|x - \theta_{n-1}\|_1^\rho \geq R^\rho$, и условие (D-b) выполнено при

$$B = \frac{1,35M'm\delta^{0,35}}{R^{0,35}}.$$

(D-c) *Ограниченность градиента в точке минимума.*

Так как $\theta_n = x_n^* \mathbf{1}$, получаем $\nabla_x F_n(\theta_n) = \mathbf{0}$, следовательно,

$$\|\nabla_x F_n(\theta_n, w_n)\|_{\rho+1}^{\rho+1} = 0.$$

Таким образом, условие выполнено при $C = 0$.

(D-d) *Ограниченность изменения функции.*

$$|F_n(x) - F_{n-1}(x)| \leq \sum_{i=1}^m |(x_i - x_n^*)^{1,35} - (x_i - x_{n-1}^*)^{1,35}| \leq mM''\delta^{1,35},$$

где M'' – гельдеровская константа.

Шум v_n подчиняется распределению Парето с параметром $\beta = 1,6 > 1,35$, следовательно:

$$\mathbb{E}|v_n - v_{n-1}|^{1,35} \leq \tilde{E} < \infty.$$

Таким образом, условие (D-d) выполнено при $D = 0$ и

$$E = (mM'' \cdot \delta^{1,35} + \tilde{E})^{1,35}.$$

(E) Ограниченность изменения наблюдаемого шума.

Рассмотрим помехи наблюдений v_n , определяемые через шум Парето:

$$v_n = \begin{cases} Z_n \text{sign}_n, & \text{симметричный шум,} \\ Z_n, & \text{асимметричный шум,} \end{cases}$$

где $Z_n \sim \text{Pareto}(\beta = 1,6, \sigma = 2,0)$, а $\text{sign}_n \sim \text{Uniform}\{-1, 1\}$.

Условие (E) требует выполнения неравенства:

$$\mathbb{E}_{\mathcal{F}_{2n-2}} |v_{2n} - v_{2n-1}|^{\rho+1} \leq \sigma_v^{\rho+1},$$

где $\rho + 1 = 1,5 < \beta$, т.е. момент порядка 1,5 существует.

Поскольку v_{2n} и v_{2n-1} независимы, разность $v_{2n} - v_{2n-1}$ также является случайной величиной с конечным моментом порядка $\rho + 1$. Для симметричного случая (с переменными знаками) численное моделирование на 10^6 реализациях дает $\mathbb{E} |v_{2n} - v_{2n-1}|^{1,5} \approx 53,73$, что позволяет положить $\sigma_v^{1,5} = 53,73$. Следовательно, условие (E) выполнено с явно определенной константой $\sigma_v^{\rho+1} = 53,73$.

(F) Независимость возмущений Δ_n .

По построению поискового алгоритма и экспериментов в RLHF-модели, векторы Δ_n генерируются независимо от всех внешних факторов. Шум v_n добавляется постфактум и не зависит от выбранного направления возмущения.

5.5. Метрики оценки и результаты моделирования

Для количественной оценки поведения алгоритма в условиях дрейфа оптимума и воздействия шума с тяжелыми хвостами используется система эмпирических метрик, отражающих как точность и устойчивость оценки, так и динамику адаптации к изменяющимся условиям. Метрики подбираются таким образом, чтобы охватывать как установившиеся характеристики алгоритма, так и его поведение на всем протяжении оптимизации. Это позволяет выявить сильные и слабые стороны метода в различных сценариях: от стационарных до резко меняющихся и зашумленных.

Оценка средней точности отслеживания дрейфующего параметра на поздних стадиях работы производится через среднюю абсолютную ошибку по последним итерациям. Стабильность поведения алгоритма при этом определяется по стандартному отклонению этих ошибок. Диапазон колебаний в рамках одного запуска характеризуется через средние минимальные и максимальные ошибки по запускам, что позволяет оценить как достижимый потенциал, так и наихудшие случаи.

Динамические характеристики алгоритма отражаются в метрике среднего времени достижения заданного уровня точности – это дает представление о скорости адаптации при ограничениях на ошибку. Связь с теоретическими определениями устойчивости обеспечивается через два момента ошибки: момент порядка, оценивающий сходимость в среднем, и соответствующую ему асимптотическую границу, которая нормирует ошибку согласно выбранному порядку момента. Используемый порядок выбирается в зависимости от параметров шума, чтобы обеспечить существование соответствующих математических ожиданий.

Формулы метрик представлены в табл. 2, сравнение результатов работы всех метрик представлены в табл. 3, а их динамика – на графиках 1–2.

Таблица 2. Основные метрики алгоритма

Метрика	Формула
Среднее абсолютное отклонение за последние 100 итераций	$\mu_{\text{last100}} = \frac{1}{100 \cdot m} \sum_{n=N-100}^{N-1} \sum_{i=1}^m x_{n,i} - x_n^* $
Стандартное отклонение ошибки за последние 100 итераций	$\sigma_{\text{last100}} = \sqrt{\frac{1}{100 \cdot m - 1} \sum_{n=N-100}^{N-1} \sum_{i=1}^m (x_{n,i} - x_n^* - \mu_{\text{last100}})^2}$
Среднее минимальное отклонение по запускам	$\bar{D}_{\min} = \frac{1}{m} \sum_{i=1}^m \min_{0 \leq n < N} x_{n,i} - x_n^* $
Среднее максимальное отклонение по запускам	$\bar{D}_{\max} = \frac{1}{m} \sum_{i=1}^m \max_{0 \leq n < N} x_{n,i} - x_n^* $
Среднее время сходимости до порога ϵ	$\bar{T}_\epsilon = \frac{1}{m} \sum_{i=1}^m T_{i,\epsilon}$ $T_{i,\epsilon} = \min\{\{n \mid 0 \leq n < N, x_{n,i} - x_n^* < \epsilon\} \cup \{N\}\}$
$l_{\rho+1}$ метрика оценки ошибки	$\mu_{\text{def2,last100}} = \frac{1}{100} \sum_{n=N-100}^{N-1} \frac{1}{m} \sum_{i=1}^m (x_{n,i} - x_n^* ^{\rho+1})^{1/2}$

Таблица 3. Сравнение результатов для разных условий дрейфа и шума

Метрика	Умеренный дрейф $\delta = 0,01$ (symm)	Почти стационарный $\delta = 0,0001$ (symm)	Несимметричный шум $\delta = 0,0001$ (asymm)
$l_{\rho+1}$ метрика оценки ошибки	0,4219	0,1954	0,1206
Среднее расстояние $E[x - x^*]$	0,3012	0,0520	0,0356
Ст. отклонение оценки	0,1682	0,2776	0,0989
Минимальное отклонение	0,0002	0,0000	0,0000
Максимальное отклонение	19,7897	57,1922	10,3462
Время сходимости ($<1,0$)	20	18	18

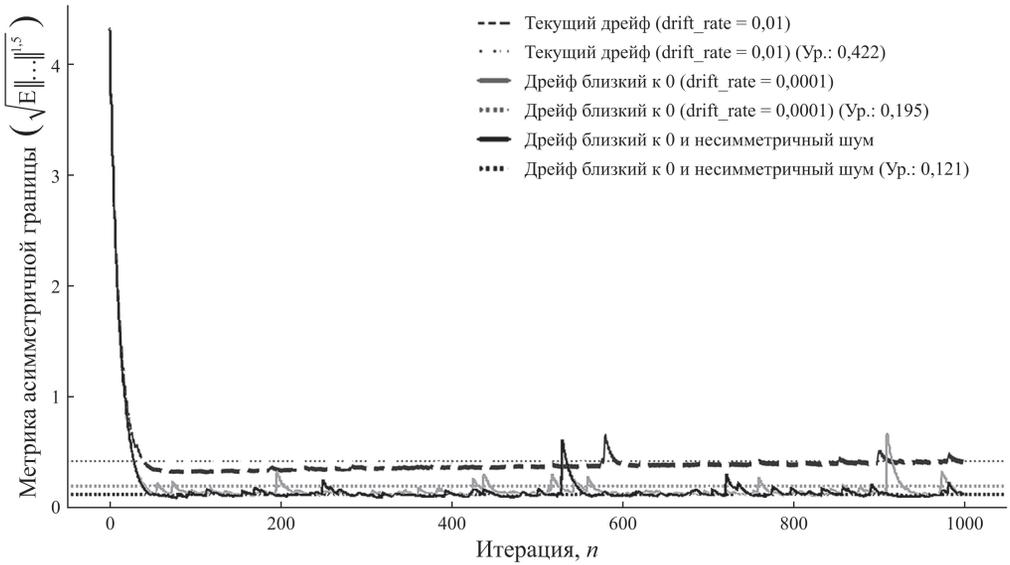


Рис. 1. График зависимости $l_{\rho+1}$ оценки ошибки от номера итерации n .

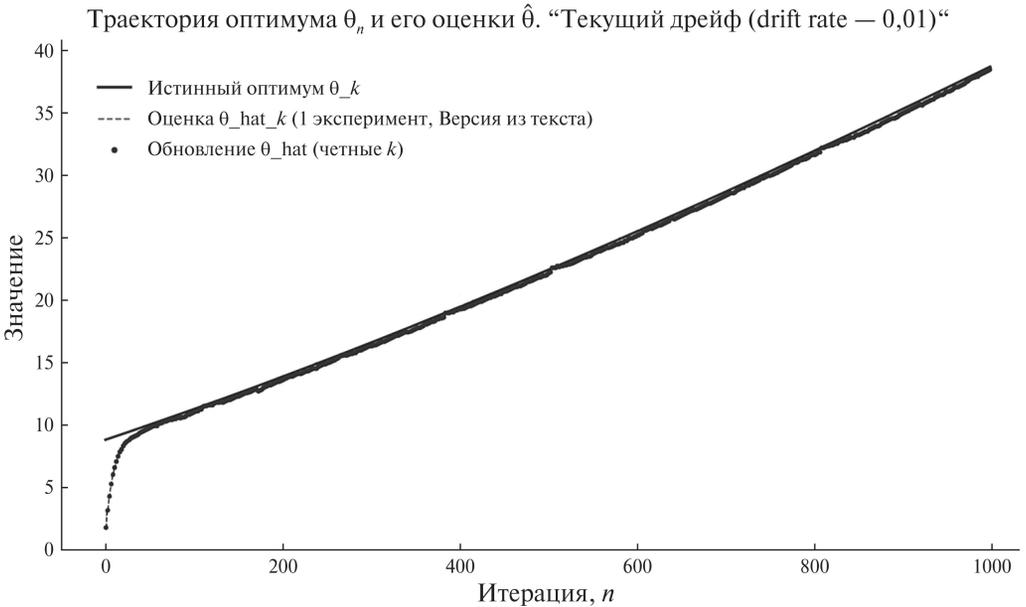


Рис. 2. График истинной траектории оптимума x_n^* и траектории оценки x_{n,i_0} .

Результаты моделирования (табл. 3) демонстрируют влияние параметров среды на поведение алгоритма (на основе 1000 экспериментов, $\rho = 0,50$, статистика по последним 100 итерациям). При умеренном дрейфе ($\delta = 0,01$) и симметричном шуме агент отслеживает цель, однако с заметной средней ошибкой (0,3012), умеренной устойчивостью ($\text{Std} = 0,1682$) и редкими, но значи-

тельными выбросами (максимум – 19,79). Метрики высокого порядка составляют 0,1788 и 0,4219, сходимость достигается за 20 итераций.

Снижение дрейфа до $\delta = 0,0001$ (почти стационарная среда) улучшает среднюю ошибку (0,0520), но одновременно увеличивает нестабильность: стандартное отклонение достигает 0,2776, а максимальная ошибка – 57,19. Это указывает на рост чувствительности к шуму с тяжелыми хвостами при ослабленном дрейфе.

Наилучшие результаты достигнуты при асимметричном шуме в условиях слабого дрейфа. Ошибка снижается до 0,0356, вариативность ограничена ($\text{Std} = 0,0989$), а максимальные отклонения существенно ниже (10,35). Метрики стабильности (0,0153, 0,1206) и время сходимости (18 итераций) также улучшаются.

Таким образом, снижение дрейфа повышает точность, но устойчивость к шуму зависит от его характера. Асимметричный шум демонстрирует лучший контроль над экстремальными ошибками, вероятно, за счет специфики градиентной оценки. Эффект требует дополнительного анализа.

6. Моделирование системы распределения задач в задачах массового обслуживания

Системы массового обслуживания (СМО), такие как современные колл-центры, характеризуются входящим потоком задач, время выполнения которых часто подчиняется распределениям с «тяжелыми хвостами» [31]. Это означает наличие статистически значимой доли задач, требующих несопоставимо большого времени на обработку, что отличает их от систем, описываемых классическими экспоненциальными или нормальными распределениями. Распределение Парето является подходящей моделью для описания таких явлений [32], позволяя учесть влияние редких, но длительных операций на общую производительность системы [33].

Для эффективного управления подобными СМО необходимо адаптивно оценивать характеристики потока и времени обслуживания. Далее описывается пример применения поискового алгоритма (3) стохастической оптимизации для модели динамической подстройки оценок ожидаемого времени обслуживания различных типов задач, подробно описанной в [34]. Метод применяется для итеративной оптимизации параметров $\theta_k, \hat{\theta}_m$, представляющих собой адаптивные оценки времени обслуживания для каждого кластера задач m и для системы (k) в целом.

В исследовании [34] представлена симуляционная модель колл-центра. Время обслуживания задач в модели генерируется из распределения Парето, параметры которого для каждого кластера калибруются на основе характеристик логнормальных распределений, аппроксимирующих исторические данные. Поисковый алгоритм (3) используется для уточнения оценок $\hat{\theta}_k, \hat{\theta}_m$, которые, в свою очередь, применяются в механизме назначения поступающих задач агентам. Проведенное моделирование показывает работоспособ-

ность метода для рассмотренной задачи в условиях стохастичности и тяжеловостого характера времени обработки задач.

6.1. Модель

Рассматривается система агентов с одинаковыми ресурсами и производительностью. Нагрузка агента i , обозначаемая как q^i , соответствует числу задач в его очереди. Каждая задача x_k характеризуется типом m и предсказанным временем выполнения, вычисляемым по формуле:

$$x_{km} = \alpha \hat{\theta}_k^i + (1 - \alpha) \hat{\theta}_m^i, \quad \alpha = \frac{\chi |\lambda_m|}{N_m + 1},$$

$$\lambda_m(\hat{\theta}_m) = \frac{1}{N_m} \sum_{k \in N_m} \omega_k \frac{\hat{\theta}_m - t_{km}}{\hat{\theta}_m} \rightarrow \min,$$

где $\hat{\theta}_k^i$ – индивидуальный прогноз агента i для задачи k , $\hat{\theta}_m^i$ – среднее предсказанное время выполнения задач типа m (с учетом локальной истории), α – весовой коэффициент, определяющий вклад индивидуального прогноза и агрегированной статистики, а χ – коэффициент сходимости. Величина λ_m отражает точность предсказания модели для задач типа m и корректируется при поступлении новых наблюдений; здесь N_m – число завершенных задач типа m , ω_k – вес соответствующей ошибки, а t_{km} – фактическое время выполнения k -й задачи типа m .

Такой механизм расчета предсказания и точности позволяет адаптировать модель к текущему качеству прогнозов, снижая влияние недостоверных данных и усиливая вклад накопленной статистики при высокой уверенности.

На каждом шаге k , при поступлении новой задачи x_k , проводится назначение x_k такому агенту i_k для балансировки нагрузки агентов:

$$(5) \quad i_k = \arg \min_i \sum_j \left| \frac{q_k^i + x_{km} - q_k^j}{d_{ij} + 1} \right|,$$

где q^j – нагрузка агента j , d_{ij} – «расстояние» между агентами (например, на основе нагрузки или физического расположения). Агенты соединены в полностью связную топологию, где каждый агент взаимодействует со всеми остальными. Это обеспечивает глобальную коммуникацию при варьирующемся влиянии агентов в зависимости от их относительной близости.

6.2. Описание набора данных и первичный анализ

Для демонстрации эффективности разработанного метода было проведено моделирование системы распределения нагрузки на основе реальных данных операторского колл-центра за сентябрь 2023 г. (более 2,3 млн вызовов). Для каждого обращения регистрировались момент поступления, время ожидания

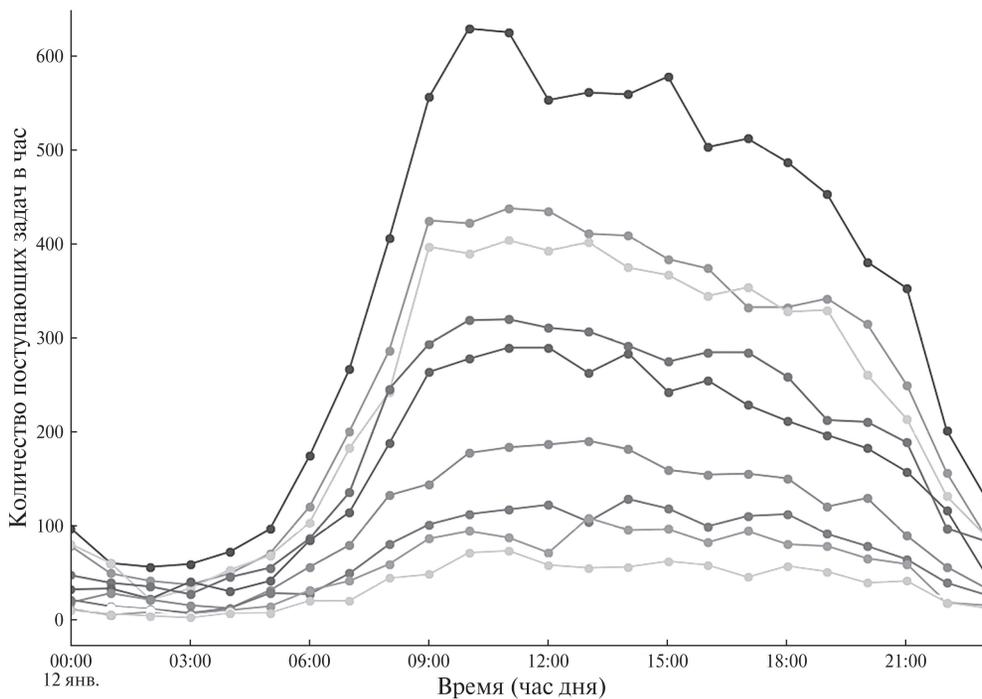


Рис. 3. Пример почасовой интенсивности различных задач, топ 10 кластеров.

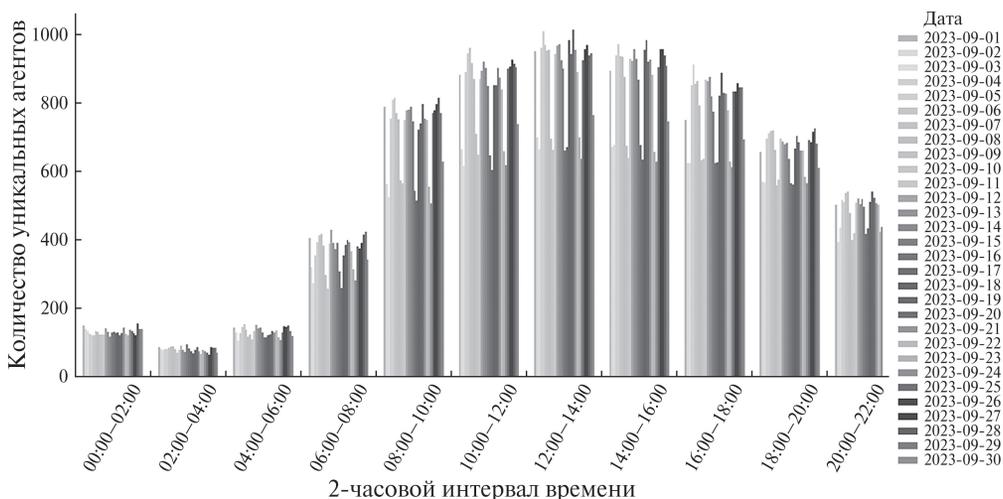


Рис. 4. Число агентов, работающих в одно время, по 2-часовым интервалам.

ответа, фактическая продолжительность разговора (ACD Time) и сегмент клиента.

На рис. 3-4 представлены две взаимодополняющие визуализации, раскрывающие ключевые характеристики входящего потока некоторых кластеров и кадрового потенциала колл-центра. График 3 демонстрирует почасовую ин-

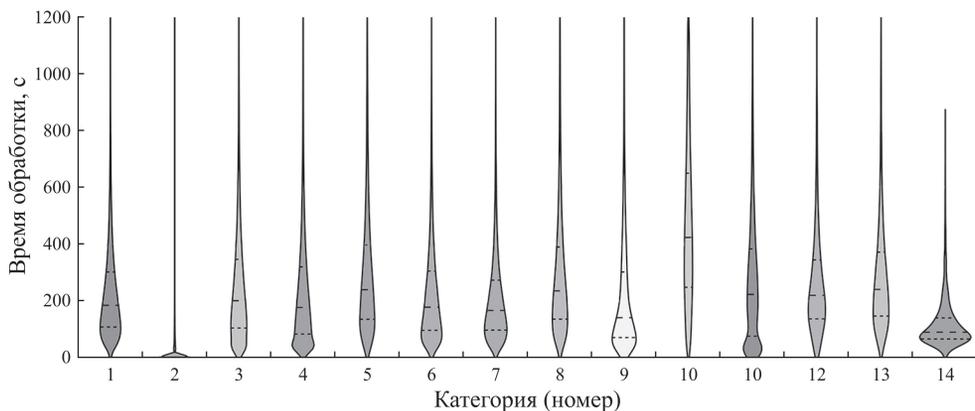


Рис. 5. Длительность разговоров по топ-14 кластерам (макс. ACD = 1200 с).

тенсивность задач по десяти крупнейшим кластерам, при этом пик нагрузки наблюдается в кластере «Young» в интервале 11–12 ч. График 4 показывает распределение активных операторов (принявших более 50 вызовов за двухчасовой интервал), максимальные значения приходятся на период с 8 до 16 ч. При этом кадровые ресурсы не всегда успевают за резкими колебаниями входящего трафика. Симуляционные задержки, агрегированные по времени суток, в целом воспроизводят динамику реального ожидания, включая утренний рост около 8–9 ч и вечерний пик после 17 ч.

Диаграммы на рис. 5 отображают распределение длительности разговоров для 14 клиентских сегментов. В целях обезличивания все сегменты были переименованы в числовые идентификаторы от 1 до 14 (см. табл. 4). Наибольшую вариативность и протяженные хвосты распределения наблюдают у сегментов 11 и 13, тогда как сегмент 2 характеризуется исключительно коротким диапазоном длительности обработки. Сегменты с номерами 3 и 14 также демонстрируют относительно узкое распределение с короткими медианами.

6.3. Результаты моделирования

Для оценки качества предложенного метода была проведена симуляция работы колл-центра на реальных данных. Результаты позволили оценить как динамику времени ожидания задач в течение суток, так и стабильность распределения нагрузки. На рис. 6–7 приведены результаты одной симуляционной сессии.

График 6 демонстрирует среднее время ожидания задач в 20-минутных интервалах, отражая характерный пик в дневные часы, связанный с высокой нагрузкой. Модель эффективно адаптируется к изменяющимся условиям: после резкого роста задержек около 12:00 среднее время ожидания быстро снижается за счет перераспределения задач.

Среднее время ожидания задачи по ходу симуляции (20 мин интервалы)

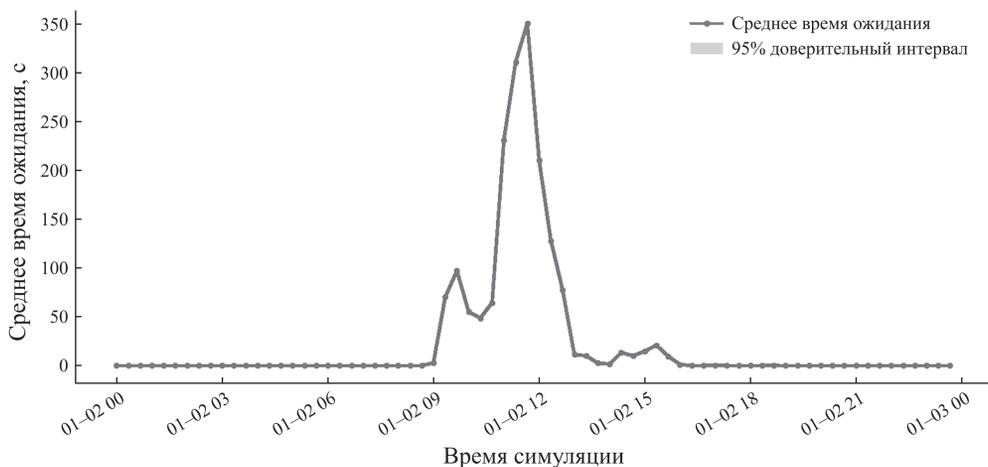


Рис. 6. Среднее время ожидания (20-мин интервалы).

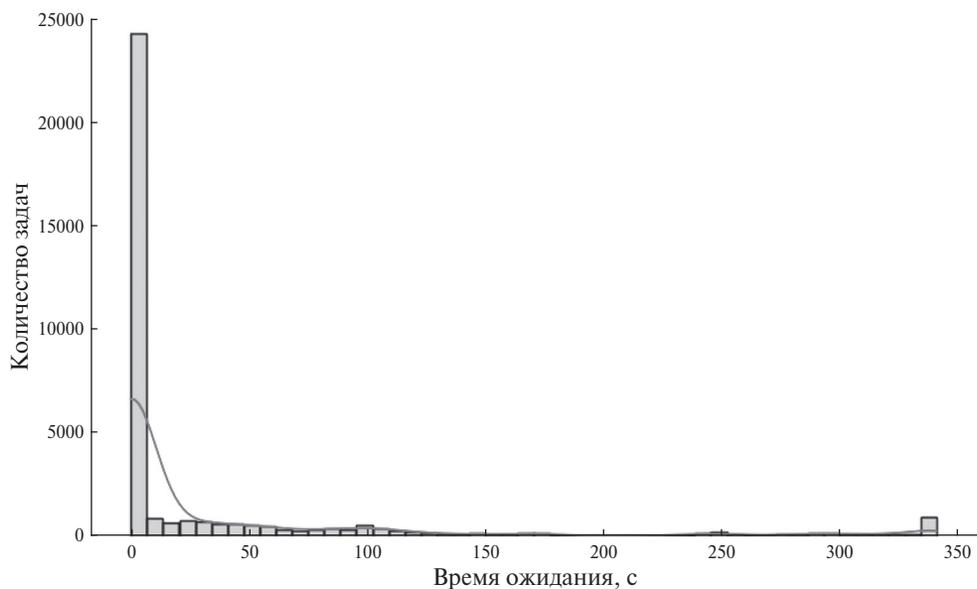


Рис. 7. Распределение времени ожидания (по 98-му перцентилю).

На гистограмме 7 представлено распределение времени ожидания задач, усеченное по 98-му перцентилю. Большинство задач были обслужены менее чем за 50 с, что соответствует целевым SLA-показателям для типичных сценариев.

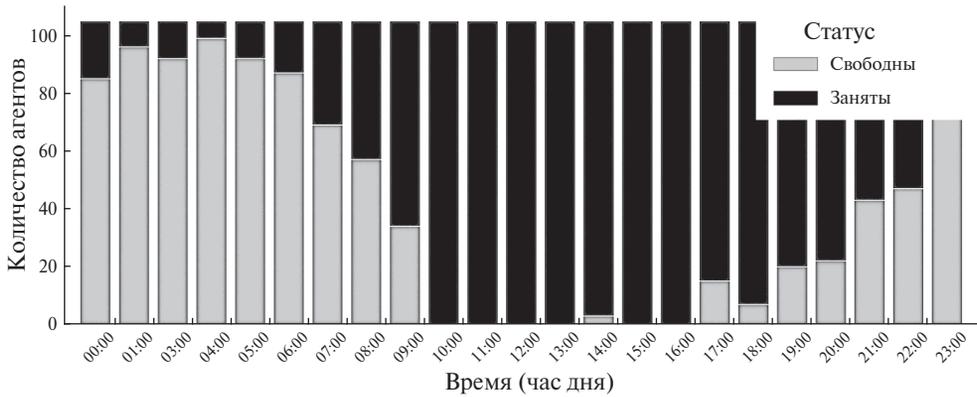


Рис. 8. Почасовая загрузка агентов: сравнение свободных и занятых ресурсов.

Для ключевых кластеров в табл. 4 представлены значения предсказанного времени выполнения задач z , количества завершенных обращений k , среднего фактического времени t_{avg} и максимальной длительности t_{max} . Кластеры с номерами от 1 до 14 соответствуют тем, что отображены на рис. 5, а кластер с номером 0 объединяет все прочие сегменты, не вошедшие в топ-14. Как видно, величины z хорошо соответствуют эмпирическим средним, несмотря на разную статистику по каждому кластеру. Это подтверждает устойчивость адаптивного предсказания на основе поискового алгоритма.

Таблица 4. Параметры категорий

	0	1	2	3	4	5	6	7
z	143,76	163,14	3,73	174,75	147,62	196,05	159,75	151,75
k	36 858	8180	6166	5764	4461	3857	2523	1707
t_{avg}	124,34	158,94	3,71	163,51	135,06	174,91	161,16	157,93
t_{max}	200	200	200	200	200	200	200	200
	8	9	10	11	12	13	14	
z	219,63	151,74	321,53	144,92	147,55	191,31	49,54	
k	1329	943	906	484	265	223	44	
t_{avg}	233,06	153,30	330,74	158,08	156,68	198,90	87,34	
t_{max}	200	200	200	200	200	200	44	

График на рис. 8 иллюстрирует почасовую загрузку операторов в ходе симуляции. В ночное и утреннее время (до 08:00) значительная доля агентов остается свободной, однако начиная с 09:00 и до 15:00 наблюдается полная загрузка всех ресурсов: количество свободных агентов опускается до нуля. Это совпадает с пиком входящего потока задач и подчеркивает необходимость точного предсказания длительности обработки. Вечером и ночью нагрузка постепенно снижается, а система возвращается к сбалансированному состоянию.

Таким образом, модель демонстрирует способность корректно адаптироваться к нагрузке, обеспечивая сдерживание времени ожидания и равномерное распределение задач в течение суток. Предложенный подход позволяет эффективно использовать ресурсы в условиях высокой вариативности обращений и может быть рекомендован для внедрения в распределенные системы поддержки с интенсивной и нерегулярной нагрузкой.

7. Заключение

В представленной работе предложен и исследован метод оценивания минимума функционала, изменяющегося во времени, в условиях, когда измерения подвержены помехам. Этот метод, основанный на псевдоградиентном подходе с рандомизацией, не требует знания градиента целевой функции и использует небольшое число измерений на каждой итерации. Сделано предположение об ограниченности скорости изменения (дрейфа) экстремума функционала. Доказано, что асимптотическая ошибка оценивания ограничена величиной $\frac{L}{K}$, где L и K определяются свойствами целевой функции, характеристиками шумов и параметрами алгоритма. Справедливость теоретических выводов была подтверждена результатами численного моделирования. Моделирование адаптации RLHF-агента к зашумленной и динамической обратной связи, в частности, с использованием шума с тяжелыми хвостами и различными скоростями дрейфа предпочтений, продемонстрировало, что поисковый алгоритм обеспечивает сходимость оценки к области целевого значения. Наблюдаемая в моделировании установившаяся ошибка и колебания оценки, обусловленные шумом и дрейфом, согласуются с теоретическими предсказаниями об ограниченности асимптотической ошибки. Кроме того, предложенный метод был протестирован на моделировании, основанном на реальных данных операторского колл-центра. Использование эмпирических характеристик потока обращений и времени обработки задач позволило продемонстрировать применимость алгоритма в задачах динамического распределения нагрузки и предсказания параметров обслуживания в реальных сервисных системах.

ПРИЛОЖЕНИЕ

Доказательство теоремы 1.

Обозначим ошибку оценивания $\text{err}_n = \hat{\theta}_n - \theta_n$.

Шаг 1. Рекуррентное соотношение для ошибки оценивания

В силу алгоритма (3) имеем

$$\hat{\theta}_{2n} = \hat{\theta}_{2n-2} - \frac{\alpha}{2\beta} \Delta_{2n}(y_{2n} - y_{2n-1}).$$

Следовательно,

$$\text{err}_{2n} = \text{err}_{2n-2} - \underbrace{(\theta_{2n} - \theta_{2n-2})}_{\text{dreif}_n} - \underbrace{\frac{\alpha}{2\beta} \Delta_{2n}(y_{2n} - y_{2n-1})}_{\text{step}_n}.$$

Шаг 2. Рекуррентное соотношение для оценки функции Ляпунова $V(x)$

Для векторов $a = \hat{\theta}_{2n-2}$ и $b = \text{dreif}_n + \text{step}_n$ по определению имеем

$$V_{2n}(\hat{\theta}_{2n}) = V_{2n-2}(\hat{\theta}_{2n} - \text{dreif}_n) = V_{2n-2}(a - b) = \|a - b - \theta_{2n-2}\|_{\rho+1}^{\rho+1}.$$

Для функции $V_{2n-2}(a - b)$ применим разложение Тейлора около точки a в направлении $-b$:

$$(II.1) \quad V_{2n-2}(a - b) = V_{2n-2}(a) - \langle \nabla V_{2n-2}(a - \delta b), b \rangle, \quad \delta \in [0, 1],$$

где градиент $\nabla V_{2n-2}(a - \delta b)$ вычисляется по формуле:

$$\nabla V_{2n-2}(a - \delta b) = (\rho + 1) \text{sign}(\delta) \odot |a - \theta_{2n-2} - \delta b|^\rho,$$

где $\text{sign}_n^{(i)}(\delta) = 0$ или ± 1 в зависимости от знака выражения i -й компоненты вектора $a - \theta_{2n-2} - \delta b$, $|a - \theta_{2n-2} - \delta b|^\rho$ – вектор из абсолютных величин компонент вектора $a - \theta_{2n-2} - \delta b$ в степени ρ , \odot – символ покомпонентного умножения. Второй член в (II.1) можно оценить следующим образом:

$$\begin{aligned} -\langle \nabla V_{2n-2}(a - \delta b), b \rangle &\leq -\langle (\rho + 1) \text{sign}(0) \odot |a - \theta_{2n-2}|^\rho, b \rangle + 2^{1-\rho} \delta^\rho \|b\|_{\rho+1}^{\rho+1} \leq \\ &\leq -\langle \nabla V_{2n-2}(a), b \rangle + 2^{1-\rho} \|b\|_{\rho+1}^{\rho+1} \end{aligned}$$

(см. доказательство теоремы 1 в [24], с. 93).

Учитывая вышеизложенное и условие (D.a):

$$(II.2) \quad \begin{aligned} V_{2n}(\hat{\theta}_{2n}) &\leq V_{2n-2}(\hat{\theta}_{2n-2}) - \langle \nabla V_{2n-2}(\hat{\theta}_{2n-2}), \text{dreif}_n + \text{step}_n \rangle + \\ &\quad + 2(A^{\rho+1} + \|\text{step}_n\|_{\rho+1}^{\rho+1}). \end{aligned}$$

Шаг 3. Разложение корректирующего шага

Из модели наблюдений разложим step_n на два слагаемых

$$\text{step}_n = \underbrace{\frac{\alpha}{2\beta} \Delta_n(F_{2n}(x_{2n}, w_{2n}) - F_{2n-1}(x_{2n-1}, w_{2n-1}))}_{\text{Почти псевдоградиентный член}} + \underbrace{\frac{\alpha}{2\beta} \Delta_n(v_{2n} - v_{2n-1})}_{\text{Шум}}.$$

а. Почти псевдоградиентный член.

Обозначим $n^\pm = 2n - \frac{1}{2} \pm \frac{1}{2}$.

Воспользовавшись формулой Тейлора, добавив и вычтя сначала $\sum_{n^\pm} \langle \nabla_x F_{n^\pm}(\hat{\theta}_{2n-2}, w_{n^\pm}), \beta \Delta_n \rangle$, а потом $\langle \nabla_x F_{2n-2}(\hat{\theta}_{2n-2}, w_{n^\pm}), \beta \Delta_n \rangle$ и еще раз $\langle \nabla_x F_{2n-2}(\theta_{2n-2}, w_{n^\pm}), \beta \Delta_n \rangle$, выводим

$$\begin{aligned}
& \sum_{n^\pm} \pm F_{n^\pm}(x_{n^\pm}, w_{n^\pm}) = \\
& = \sum_{n^\pm} \pm F_{n^\pm}(\hat{\theta}_{2n-2}, w_{n^\pm}) + \langle \nabla_x F_{n^\pm}(\hat{\theta}_{2n-2} \pm \delta_{n^\pm} \beta \Delta_n, w_{n^\pm}), \beta \Delta_n \rangle = \\
& = \sum_{n^\pm} \pm F_{n^\pm}(\hat{\theta}_{2n-2}, w_{n^\pm}) + \langle \nabla_x F_{n^\pm}(\hat{\theta}_{2n-2}, w_{n^\pm}), \beta \Delta_n \rangle + \\
& + \langle \nabla_x F_{n^\pm}(\hat{\theta}_{2n-2} \pm \delta_{n^\pm} \beta \Delta_n, w_{n^\pm}), \beta \Delta_n \rangle - \langle \nabla_x F_{n^\pm}(\hat{\theta}_{2n-2}, w_{n^\pm}), \beta \Delta_n \rangle = \\
& = \sum_{n^\pm} \pm F_{n^\pm}(\hat{\theta}_{2n-2}, w_{n^\pm}) + \langle \nabla_x F_{2n-2}(\theta_{2n-2}, w_{n^\pm}), \beta \Delta_n \rangle + \\
& + \langle \nabla_x F_{2n-2}(\hat{\theta}_{2n-2}, w_{n^\pm}), \beta \Delta_n \rangle - \langle \nabla_x F_{2n-2}(\theta_{2n-2}, w_{n^\pm}), \beta \Delta_n \rangle + \\
& + \langle \nabla_x F_{n^\pm}(\hat{\theta}_{2n-2}, w_{n^\pm}), \beta \Delta_n \rangle - \langle \nabla_x F_{2n-2}(\hat{\theta}_{2n-2}, w_{n^\pm}), \beta \Delta_n \rangle + \\
& + \langle \nabla_x F_{n^\pm}(\hat{\theta}_{2n-2} \pm \delta_{n^\pm} \beta \Delta_n, w_{n^\pm}), \beta \Delta_n \rangle - \langle \nabla_x F_{n^\pm}(\hat{\theta}_{2n-2}, w_{n^\pm}), \beta \Delta_n \rangle,
\end{aligned}$$

где $\delta_{n^\pm} \in [0, 1]$.

Применим операцию условного математического ожидания относительно σ -алгебры \mathcal{F}_{2n-2} . Учитывая независимость в силу условия (F) векторов Δ_n от w_{n^\pm} и σ -алгебры \mathcal{F}_{2n-2} , получаем

$$\frac{\alpha}{2\beta} \mathbb{E}_{\mathcal{F}_{2n-2}} \left\{ \Delta_n \sum_{n^\pm} \pm F_{n^\pm}(\hat{\theta}_{2n-2}, w_{n^\pm}) \right\} = 0$$

с силу центрированности Δ_n и

$$\frac{\alpha}{2\beta} \mathbb{E}_{\mathcal{F}_{2n-2}} \left\{ \Delta_n \sum_{n^\pm} \langle \nabla_x F_{2n-2}(\theta_{2n-2}, w_{n^\pm}), \beta \Delta_n \rangle \right\} = 0,$$

так как в силу условия (C) имеем $\mathbb{E}_{\mathcal{F}_{2n-2}} \{ \nabla_x F_{2n-2}(\theta_{2n-2}, w_{n^\pm}) \} = \nabla_x f_{2n-2}(\theta_{2n-2})$ и градиент функции $f_{2n-2}(\cdot)$ в точке минимума θ_{2n-2} равен нулю.

В итоге в силу условия (C) выводим для почти псевдоградиентного члена

$$\mathbb{E}_{\mathcal{F}_{2n-2}} \left\{ \frac{\alpha}{2\beta} \Delta_n \sum_{n^\pm} \pm F_{n^\pm}(x_{n^\pm}, w_{n^\pm}) \right\} = \frac{\alpha}{d} \nabla f_{2n}(\hat{\theta}_{2n-2}) + \frac{\alpha}{2\beta} \mathbb{E}_{\mathcal{F}_{2n-2}} \text{corr}_n,$$

где

$$\begin{aligned}
\text{corr}_n = & \sum_{n^\pm} \langle \nabla_x F_{n^\pm}(\hat{\theta}_{2n-2} \pm \delta_{n^\pm} \beta \Delta_n, w_{n^\pm}), \beta \Delta_n \rangle - \langle \nabla_x F_{n^\pm}(\hat{\theta}_{2n-2}, w_{n^\pm}), \beta \Delta_n \rangle + \\
& + \langle \nabla_x F_{2n-2}(\hat{\theta}_{2n-2}, w_{n^\pm}), \beta \Delta_n \rangle - \langle \nabla_x F_{2n-2}(\theta_{2n-2}, w_{n^\pm}), \beta \Delta_n \rangle + \\
& + \langle \nabla_x F_{n^\pm}(\hat{\theta}_{2n-2}, w_{n^\pm}), \beta \Delta_n \rangle - \langle \nabla_x F_{2n-2}(\hat{\theta}_{2n-2}, w_{n^\pm}), \beta \Delta_n \rangle.
\end{aligned}$$

В силу условий (B) и (D.b) имеем оценку

$$\begin{aligned} \|\text{corr}_n\| &\leq M\beta^\rho \|\Delta_n\| \left(2\|\Delta_n\|^\rho + 2\|\hat{\theta}_{2n-2} - \theta_{2n-2}\|^\rho \right) + 3B\|\hat{\theta}_{2n-2} - \theta_{2n-2}\|^\rho = \\ &= 2M\beta^\rho + (2 + 3B)\|\hat{\theta}_{2n-2} - \theta_{2n-2}\|^\rho. \end{aligned}$$

б. Шум. Применим операцию условного математического ожидания относительно σ -алгебры \mathcal{F}_{2n-2} . Учитывая независимость Δ_n от v_{2n} , v_{2n-1} и \mathcal{F}_{2n-2} , получаем

$$\mathbb{E}_{\mathcal{F}_{2n-2}} \left\{ \frac{\alpha}{2\beta} \Delta_n (v_{2n} - v_{2n-1}) \right\} = 0.$$

с. Итоговая оценка для второго слагаемого в правой части неравенства (П.2). В силу условия сильной выпуклости (A) получаем

$$\begin{aligned} -\mathbb{E}_{\mathcal{F}_{2n-2}} \{ \langle \nabla V_{2n-2}(\hat{\theta}_{2n-2}), \text{dreif}_n + \text{step}_n \rangle \} &\leq -\frac{\mu\alpha}{d} V_{2n-2}(\hat{\theta}_{2n-2}) - \\ -\frac{\alpha}{2\beta} \mathbb{E}_{\mathcal{F}_{2n-2}} \langle \nabla V_{2n-2}(\hat{\theta}_{2n-2}), \text{dreif}_n + \text{corr}_n \rangle &\leq -\frac{\mu\alpha}{d} V_{2n-2}(\hat{\theta}_{2n-2}) + \\ + 2(A + \alpha M\beta^{\rho-1})^2 + \left(2 + \frac{\alpha}{2\beta}(2 + 3B) \right) &\sum_{i=1}^d |\hat{\theta}_{2n-2}^i - \theta_{2n-2}^i|^{2\rho} \leq \\ \leq -\frac{\mu\alpha}{d} V_{2n-2}(\hat{\theta}_{2n-2}) + \varepsilon V_{2n-2}(\hat{\theta}_{2n-2}) + c_1, \end{aligned}$$

где $\varepsilon > 0$ и

$$c_1 = 2(A + \alpha M\beta^{\rho-1})^2 + \varepsilon^{\rho-1} \left(2 + \frac{\alpha}{2\beta}(2 + 3B) \right)^{\frac{1-\rho}{\rho+1}}.$$

Шаг 4. Оценка третьего слагаемого в правой части неравенства (П.2)

По аналогии с выкладками на Шаге 4 step_n можно представить в виде

$$\text{step}_n = \frac{\alpha}{2\beta} \Delta_n \sum_{i=1}^8 a_i,$$

где

$$\begin{aligned} a_1 &= \sum_{n^\pm} \pm F_{n^\pm}(\hat{\theta}_{2n-2}, w_{n^\pm}), \\ a_{2,3} &= \langle \nabla_x F_{n^\pm}(\hat{\theta}_{2n-2} \pm \delta_{n^\pm} \beta \Delta_n, w_{n^\pm}), \beta \Delta_n \rangle - \langle \nabla_x F_{n^\pm}(\hat{\theta}_{2n-2}, w_{n^\pm}), \beta \Delta_n \rangle, \\ a_{4,5} &= \langle \nabla_x F_{n^\pm}(\hat{\theta}_{2n-2}, w_{n^\pm}), \beta \Delta_n \rangle - \langle \nabla_x F_{n^\pm}(\theta_{n^\pm}, w_{n^\pm}), \beta \Delta_n \rangle, \\ a_{6,7} &= \langle \nabla_x F_{n^\pm}(\theta_{n^\pm}, w_{n^\pm}), \beta \Delta_n \rangle, \\ a_8 &= v_{2n} - v_{2n-1}. \end{aligned}$$

Для a_1 в силу условия (D.d) имеем $\mathbb{E}_{\mathcal{F}_{2n-2}} |a_1|^{\rho+1} \leq DV_{2n-2}(\hat{\theta}_{2n-2}) + E$, для a_2, a_3 в силу условия (B) имеем $\mathbb{E}_{\mathcal{F}_{2n-2}} |a_i|^{\rho+1} \leq M^{\rho+1} \beta^{2\rho+2}$, $i = 2, 3$, для a_4, a_5 в силу условия (B) и неравенства Йенсена выводим $\mathbb{E}_{\mathcal{F}_{2n-2}} |a_i|^{\rho+1} \leq (M\beta\|\hat{\theta}_{2n-2} - \theta_{n^\pm}\|_2^\rho)^{\rho+1} \leq M^{\rho+1} \beta^{\rho+1} d^{\frac{\rho-1}{2}} V_{n^\pm}(\hat{\theta}_{2n-2})$, $i = 4, 5$,

для a_6, a_7 в силу условия (D.c) имеем $\mathbb{E}_{\mathcal{F}_{2n-2}} |a_i|^{\rho+1} \leq C$, $i = 6, 7$,
 для a_8 в силу условия (E) имеем $\mathbb{E}_{\mathcal{F}_{2n-2}} |a_8|^{\rho+1} \leq \sigma_v^{\rho+1}$.
 В силу неравенства Йенсена имеем

$$\left(\frac{\sum_{i=1}^8 |a_i|}{8} \right)^{\rho+1} \leq \frac{1}{8} \sum_{i=1}^8 |a_i|^{\rho+1}$$

и, следовательно, выводим

$$\begin{aligned} 2A^{\rho+1} + 2\mathbb{E}_{\mathcal{F}_{2n-2}} \|\text{step}_n\|_{\rho+1}^{\rho+1} &\leq 2A^{\rho+1} + 2.8^\rho \left(\frac{\alpha}{2\beta} \right)^{\rho+1} \sum_{i=1}^7 |a_i|^{\rho+1} \leq 2A^{\rho+1} + \\ &+ 2^{2\rho} \alpha^{\rho+1} \left(2M^{\rho+1} \left(\beta^{\rho+1} + d^{\frac{\rho-1}{2}} \sum_{n^\pm} V_{n^\pm}(\hat{\theta}_{2n-2}) \right) + \right. \\ &\quad \left. + \frac{2C + DV_{2n-2}(\hat{\theta}_{2n-2}) + E + \sigma_v^{\rho+1}}{\beta^{\rho+1}} \right) \leq \\ &\leq c_2 \alpha^{\rho+1} V_{2n-2}(\hat{\theta}_{2n-2}) + c_3, \end{aligned}$$

где

$$c_2 = 2^{3\rho+1} M^{\rho+1} \left(d^{\frac{\rho-1}{2}} + \frac{D}{\beta^{\rho+1}} \right)$$

и

$$c_3 = 2A^{\rho+1} + 2^{2\rho} \alpha^{\rho+1} \left(2M^{\rho+1} \left(\beta^{\rho+1} + 3,2^\rho d^{\frac{\rho-1}{2}} \right) + \frac{E + 2C + \sigma_v^{\rho+1}}{\beta^{\rho+1}} \right).$$

Шаг 5. Формирование рекуррентного неравенства

Собирая все оценки, получаем

$$V_{2n} \leq V_{2n-2} - (\mu\alpha d^{-1} - \varepsilon - c_2 \alpha^{\rho+1}) V_{2n-2} + c_1 + c_3.$$

Вводя обозначения

$$K = 1 - \mu\alpha d^{-1} + \varepsilon + c_2 \alpha^{\rho+1}, \quad L = c_1 + c_3,$$

получаем

$$V_{2n} \leq (1 - K)V_{2n-2} + L.$$

Выбрав достаточно малые α и ε , можем обеспечить неравенство $K < 1$, при выполнении которого справедливо заключение теоремы 1. \square

СПИСОК ЛИТЕРАТУРЫ

1. Поляк Б.Т. Введение в оптимизацию. М.: Наука, 1983. 384 с.
2. Поляк Б.Т., Цыпкин Я.З. Псевдоградиентные алгоритмы адаптации и обучения // *АиТ*. 1973. № 3. С. 45–68.
3. Поляк Б.Т., Цыпкин Я.З. Адаптивные алгоритмы оценивания (сходимость, оптимальность, устойчивость) // *АиТ*. 1979. № 3. С. 71–84.
4. Поляк Б.Т., Цыпкин Я.З. Оптимальные псевдоградиентные алгоритмы адаптации // *АиТ*. 1980. № 8. С. 74–84.
5. Поляк Б.Т. О некоторых способах ускорения сходимости итерационных методов // *Журн. вычисл. мат. и мат. физики*. 1964. V. 4. № 5. С.791–803.
6. Поляк Б.Т. Новый метод типа стохастической аппроксимации // *АиТ*. 1990. № 7. С. 98–108.
7. Polyak B.T., Yuditskij A.V. Acceleration of stochastic approximation procedures by averaging // *SIAM J. Contr. Optim.* 1992. V. 30. No. 4. P. 838–855.
8. Поляк Б.Т. Сходимость и скорость сходимости итеративных стохастических алгоритмов. I // *АиТ*. 1976. № 12. С. 83–94.
9. Поляк Б.Т. Сходимость и скорость сходимости итеративных стохастических алгоритмов. II // *АиТ*. 1977. № 4 С. 101–107.
10. Поляк Б.Т., Цыбаков А.Б. Оптимальные порядки точности поисковых алгоритмов стохастической оптимизации // *Проблемы передачи информации*. 1990. № 26. 2. С. 45–53.
11. Растрюгин Л.А. Статистические методы поиска. М.: Наука, 1968. 376 с.
12. Граничин О.Н. Стохастическая аппроксимация с возмущением на входе при зависимых помехах наблюдения // *Вестн. ЛГУ*. 1989. С. 27–31.
13. Spall J.C. Multivariate stochastic approximation using a simultaneous perturbation gradient approximation // *IEEE Transact. Autom. Control*. 1992. 37(3). С. 332–341.
14. Spall J.C. A one-measurement form of simultaneous perturbation stochastic approximation // *Automatica*. 1997. 33(1). P. 109–112.
15. Граничин О.Н., Поляк Б.Т. Рандомизированные алгоритмы оценивания и оптимизации при почти произвольных помехах. М.: Наука, 2003. 291 с.
16. Granichin O., Volkovich V., Toledano-Kitai D. Randomized algorithms in automatic control and data mining. Berlin Heldenberg: Springer, 2015. 251 p.
17. Попков А.Ю. Градиентные методы для нестационарных задач безусловной оптимизации // *АиТ*. 2005. № 6. С. 38–46.
18. Kiefer J., Wolfowitz J. Stochastic estimation of the maximum of a regression function // *The Annals of Mathematical Statistics*. 1952. 23(3). P. 462–466.
19. Вахитов А.Т., Граничин О.Н., Гуревич Л.С. Алгоритм стохастической аппроксимации с пробным возмущением на входе в нестационарной задаче оптимизации // *АиТ*. 2009. № 11. С. 70–79.
20. Granichin O., Amelina N. Simultaneous perturbation stochastic approximation for tracking under unknown but bounded disturbances // *IEEE Transact. Autom. Control*. 2015. V. 60. No. 6. P. 1653–1658.
21. Шиббаев И.А. Безградиентные методы оптимизации для функций с гельдеровым градиентом // *Дисс. ... канд. физ.-мат. наук. Долгопрудный: МФТИ, 2024.*

22. *Shibaev I., Dvurechensky P., Gasnikov A.* Zeroth-order methods for noisy Holder-gradient functions // *Optimization Letters*. 2022. V. 16. P. 2123–2143.
23. *Mandelbrot B.* New methods in statistical economics // *Journal of Political Economy*. 1963. V. 71. No. 5. P. 421–440.
24. *Вахитов А.Т., Граничин О.Н., Сысоев С.С.* Точность оценивания рандомизированного алгоритма стохастической оптимизации // *АиТ*. 2006. № 4. С. 86–96.
25. *Граничин О.Н.* Поисковые алгоритмы стохастической аппроксимации с рандомизацией на входе // *АиТ*. 2015. № 5. С. 43–59.
26. *Min T. et al.* Understanding Impact of Human Feedback via Influence Functions. arXiv preprint arXiv:2501.05790. 2025.
27. *Shen W. et al.* Loose lips sink ships: mitigating length bias in reinforcement learning from human feedback // *Findings of the Association for Computational Linguistics: EMNLP 2023*, 2023. P. 2859–2873.
28. *Christiano P.F. et al.* Deep reinforcement learning from human preferences // *Advances in Neural Information Processing Systems*. 2017. V. 30. P. 1–9.
29. *Stiennon N. et al.* Learning to summarize with human feedback // *Advances in Neural Information Processing Systems*. 2020. V. 33. P. 3008–3021.
30. *Ouyang L. et al.* Training language models to follow instructions with human feedback // *Advances in Neural Information Processing Systems*. 2022. V. 35. P. 27730–27744.
31. *Gans N., Koole G., Mandelbaum A.* Telephone call centers: Tutorial, review, and research prospects // *Manufacturing and Service Operations Management*. 2003. V. 5. No. 2. P. 79–141.
32. *Anderson. C.* *The Long Tail: Why the Future of Business is Selling Less of More*, NY.: Hyperion, 2006. 256 p.
33. *Goel S., Broder A., Gabrilovich E., Pang. B.* Anatomy of the long tail: Ordinary People With Extraordinary Tastes // *Proceedings of the Third ACM International Conference on Web Search and Data Mining (WSDM'10)*, ACM, New York, NY, USA, 2010. P. 201–210.
34. *Akinfiyev I., Tarasova E.* Cluster-Aware LVP: Enhancing Task Allocation with Growth Dynamics // *15th IFAC Workshop on Adaptive and Learning Control Systems (ALCOS)*, 2025.

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

Поступила в редакцию 23.06.2025

После доработки 30.06.2025

Принята к публикации 04.07.2025

© 2025 г. Я. БЕКРИ, д-р философии (yannisbekri@hotmail.com)
(Университет Гренобль-Альпы, Сен-Мартен-д'Эр, Франция),
А. НЕМИРОВСКИЙ, д-р физ.-мат. наук (nemirovs@isy.e.gatech.edu)
(Технологический институт Джорджии, Атланта, США),
А. ЮДИЦКИЙ, канд. техн. наук (anatoli.juditsky@univ-grenoble-alpes.fr)
(Университет Гренобль-Альпы, Сен-Мартен-д'Эр, Франция)

О НАДЕЖНОМ ВОССТАНОВЛЕНИИ СИГНАЛОВ ПО НЕПРЯМЫМ НАБЛЮДЕНИЯМ¹

Рассматривается следующая *линейная обратная задача с неопределенностью*. При имеющемся наблюдении $\omega = Ax_* + \zeta$, где $A \in \mathbf{R}^{m \times p}$ и $\zeta \in \mathbf{R}^m$ – шум наблюдения, требуется восстановить неизвестный сигнал x_* , который принадлежит заданному выпуклому множеству $\mathcal{X} \subset \mathbf{R}^n$. В отличие от стандартной постановки такой задачи предполагается, что модельный шум ζ является “искаженным”, т.е. содержит детерминированную, плотную или сингулярную, компоненту. В частности, ζ допускает разложение $\zeta = N\nu_* + \xi$, где ξ – случайный шум, а $N\nu_*$ – “адверсарияльное загрязнение” с известным множеством $\mathcal{N} \subset \mathbf{R}^n$, т.е. $\nu_* \in \mathcal{N}$ и $N \in \mathbf{R}^{m \times n}$. Рассматриваются две постановки, в которых \mathcal{N} является либо выпуклым ограниченным множеством, либо множеством разреженных векторов (с не более чем s ненулевыми элементами). Исследуются свойства “устойчивых” по отношению к загрязнению *полиэдральных оценок* – специального класса нелинейных оценок, представленных в [1, 2]. Показывается, как оценки такого рода с гарантированным качеством могут быть построены в ситуации, когда множество сигналов является *элитопом* (по существу, симметричным выпуклым множеством, ограниченным квадратичными поверхностями), с помощью эффективных процедур выпуклой оптимизации.

Ключевые слова: робастное оценивание, линейные обратные задачи с загрязненными наблюдениями, оценивание сигнала в сингулярном шуме.

DOI: 10.31857/S0005231025080025, EDN: UTATKM

1. Постановка задачи и цели исследования

1.1. Введение

Проблема “робастного оценивания” находится в центре внимания классической статистической литературы с момента появления термина в 1950-х гг. Невозможно дать исчерпывающий обзор существующих публикаций по робастному оцениванию, и авторы не пытаются сделать это ниже; чтобы получить базовое представление о “классической” литературе, читатель может

¹ Работа выполнена при финансовой поддержке Мультидисциплинарного института искусственного интеллекта (МИАИ), @ Гренобль-Альпы (грант ANR-19-P3IA-0003).

обратиться к ранним ссылкам в [3], основополагающей монографии [4] или недавнему обзору [5].²

Настоящая статья посвящена робастному оцениванию сигнала по “непрямым” линейным наблюдениям. В частности, целью является восстановление линейного образа $w_* = Bx_*$ неизвестного сигнала x_* , который принадлежит заданному выпуклому множеству $\mathcal{X} \subset \mathbf{R}^p$, при известных $B \in \mathbf{R}^{q \times p}$, $A \in \mathbf{R}^{m \times p}$ и имеющемуся зашумленному наблюдению

$$(1) \quad \omega = Ax_* + \eta_* + \xi \in \mathbf{R}^m$$

сигнала x_* , возмущенного смешанным (случайным и детерминированным) шумом $\xi + \eta_*$. Здесь ξ – случайная составляющая шума, а η_* – “адверсарийный” детерминированный шум. В последнее время эта проблема привлекает большое внимание в контексте робастного восстановления разреженного сигнала x_* (с не более чем $s \ll p$ ненулевыми элементами). В частности, разреженная регрессия с акцентом на помеху с загрязнением исследована в [12–16]; методы, основанные на введении штрафов для вектора выбросов, изучены в [17, 18], см. также [19, 20]. Современное состояние этой области описано в монографии [21].

В данной работе акцент сделан на существенно отличных от стандартных предположениях о структуре неизвестного сигнала x_* и о загрязнении η_* , что исключает прямое сравнение с приведенными результатами. А именно, предполагается, что множество \mathcal{X} сигналов является *эллитопом* – выпуклым, компактным и симметричным относительно начала координат подмножеством \mathbf{R}^p , ограниченным квадратичными поверхностями.³ Интерес авторов к эллитопам обусловлен тем, что эти множества сигналов “удобны” для решения задачи оценки неизвестного сигнала x_* по наблюдениям (1) в гауссовском шуме (при отсутствии загрязнения) ($\eta_* = 0$, $\xi \sim \mathcal{N}(0, \sigma^2 I_m)$). В частности, рассмотрим линейную оценку $\hat{w}_{\text{lin}}(\omega) = G_{\text{lin}}^T \omega$ и *полиэдральную оценку* $\hat{w}_{\text{poly}}(\omega) = B\hat{x}_{\text{poly}}(\omega)$, где

$$\hat{x}_{\text{poly}}(\omega) \in \underset{x \in \mathcal{X}}{\text{Argmin}} \|G_{\text{poly}}^T (Ax - \omega)\|_{\infty},$$

для сигнала w_* . Пусть \mathcal{X} – эллитоп, а ошибка оценки измеряется в коэллитопической норме $\|\cdot\|$ (т.е. такой, что единичный шар \mathcal{B}_* нормы $\|\cdot\|_*$, сопряженной к $\|\cdot\|$, является эллитопом). В этой ситуации можно указать (ср. [1, 2, 22]) эффективно вычислимые *контрастные матрицы* $G_{\text{lin}} \in \mathbf{R}^{m \times q}$ и $G_{\text{poly}} \in \mathbf{R}^{m \times m}$, при которых оценки $\hat{w}_{\text{lin}}(\cdot)$ и $\hat{w}_{\text{poly}}(\cdot)$ становятся почти минимаксными.

² Важным вкладом в развитие робастной статистики стали работы Б. Поляка и Я. Цыпкина по робастным алгоритмам стохастической оптимизации и идентификации систем, см. [6–11].

³ См. [2, раздел 4.2.1] или раздел 1.3 ниже; простейшим примером эллитопа является пересечение конечного семейства эллипсоидов/эллиптических цилиндров с общим центром.

Предположим, что возмущение η_* имеет особую структуру $\eta_* = N\nu_*$, где $\nu_* \in \mathcal{N}$, “множество помех” $\mathcal{N} \subset \mathbf{R}^n$ известно, а матрица N размерности $m \times n$ задана. Рассмотрим два типа предположений относительно \mathcal{N} : \mathcal{N} является либо 1) (непустым) компактным и выпуклым множеством, либо 2) множеством разреженных возмущений (с не более чем $s \leq n$ ненулевыми элементами). Внимание авторов сосредоточено на построении и анализе эффективности полиэдральной оценки $\hat{w}(\omega)$ для $w_* = Bx_*$, робастной по отношению к загрязняющему сигналу, при этом решение задачи в первой постановке позволяет построить нетривиальное решение для второй.

Хотелось бы подчеркнуть главную особенность предлагаемого авторами подхода: в данной работе A , B и N – произвольные матрицы соответствующей размерности, а \mathcal{X} и \mathcal{N} – достаточно общие множества. Как следствие, здесь используется “операционный” подход⁴, предложенный в [23] и развитый в [1, 2, 24, 25]. В рамках этого подхода и оценки, и их риски являются результатом эффективного вычисления, а не явных аналитических выкладок, которые попросту невозможны в данных обстоятельствах. При этом термин “эффективный” является существенным, а также обуславливает принципиальные ограничения последующих результатов. Прежде всего, он накладывает ограничения на структуру интересующего авторов множества сигналов и на норму, в которой измеряются ошибки оценивания. Как показано в [2], максимум квадратичной формы по эллитопу допускает достаточно точную эффективно вычислимую верхнюю границу, что приводит к качественным границам для риска линейных и полиэдральных оценок, когда множество сигналов является эллитопом. Более того, если в случае выпуклого компактного множества \mathcal{N} загрязнений построение границ риска для полиэдральной оценки $\hat{w}_G(\omega)$, связанной с *заданной* матрицей контрастов G , возможно при довольно слабых предположениях о множестве помех \mathcal{N} (по существу, эффективной вычислимости⁵ этого множества достаточно), то фундаментальная задача *контрастного синтеза* (минимизация этих границ по матрицам контрастов) допускает эффективно вычислимое решение только тогда, когда \mathcal{N} является либо эллитопом, либо “коэллитопом” (полярной к эллитопу).

Здесь хотелось бы упомянуть еще одно направление исследований проблемы оценки сигнала x_* по наблюдениям (1) при детерминированном возмущении (случай $\xi = 0$): стандартную задачу оптимального восстановления [27, 28] и гарантированное оценивание в динамических системах при случайных ограниченных возмущениях [29–35]. Настоящая работа может рассматриваться как попытка распространить соответствующие методы на случай, когда присутствуют как детерминированные, так и случайные шумы наблюдения.

⁴ В отличие от классического “описательного” подхода к решению задачи оценивания путем выведения аналитического выражения для риска оценки-кандидата и последующей оптимизации ее параметров.

⁵ Для большинства практических целей эффективная вычислимость означает, что ограничения на множество можно моделировать с помощью CVX [26]. Что это значит на самом деле? Краткий ответ читатель найдет в [2, приложение А].

Структура статьи. В разделе 1.2 даны точная постановка рассматриваемой задачи оценивания и необходимые сведения для последующего анализа. В разделе 2 представлены анализ и построение полиэдральной оценки для случая ограниченного загрязнения. Затем в разделе 3 описано применение предложенного подхода в случае (неограниченного) сингулярного загрязнения с использованием разреженной модели вектора помех. В Приложении 2 приведены некоторые результаты по ℓ_1 -восстановлению, использованные в статье.

Обозначения. В дальнейшем отношения порядка между векторами понимаются поэлементно; например, запись $t \geq t'$ для $t, t' \in \mathbf{R}^n$ означает, что вектор $t - t'$ имеет неотрицательные элементы. Записи $[A; B]$ и $[A, B]$ означают вертикальную и горизонтальную конкатенацию матриц A и B соответствующей размерности. Обозначим через \mathbf{S}^m пространство симметричных матриц размерности $m \times m$, а через \mathbf{S}_+^m – конус неотрицательно определенных матриц в \mathbf{S}^m ; запись $A \succeq B$ ($A \succ B$) означает, что матрица $A - B$ неотрицательно определена (соответственно положительно определена). В дальнейшем

$$\phi_{\mathcal{Z}}(\zeta) := \max_{z \in \mathcal{Z}} \zeta^T z$$

представляет собой опорную функцию непустого компактного множества $\mathcal{Z} \subset \mathbf{R}^N$; $\|G\|$ обозначает максимальную евклидову норму столбцов матрицы G .

1.2. Постановка задачи

Интересующая авторов задача оценивания состоит в следующем.

Считается доступным наблюдение (ср. (1))

$$(2) \quad \omega = Ax_* + N\nu_* + \xi \in \mathbf{R}^m,$$

где:

- $N \in \mathbf{R}^{m \times n}$ и $A \in \mathbf{R}^{m \times p}$ – заданные матрицы,
- $\nu_* \in \mathbf{R}^n$ – неизвестный сигнал помехи, $\nu_* \in \mathcal{N}$ (известное подмножество \mathbf{R}^n),
- x_* – неизвестный искомый сигнал, принадлежащий заданному выпуклому компактному множеству $\mathcal{X} \subset \mathbf{R}^p$, симметричному относительно начала координат,
- $\xi \sim P_\xi$ – случайный шум наблюдения.

По наблюдению ω требуется восстановить линейный образ $w_* = Bx_*$ сигнала x_* , где B – заданная матрица размерности $q \times p$.

При заданном $\epsilon \in (0, 1)$ будем оценивать качество восстановления $\hat{w}(\cdot)$ его ϵ -риском⁶

$$\text{Risk}_\epsilon[\hat{w}] = \inf \left\{ \rho : \text{Prob}_{\xi \sim P_\xi} \{ \xi : \|Bx_* - \hat{w}(\omega)\| > \rho \} \leq \epsilon \quad \forall (\nu_* \in \mathcal{N} \text{ и } x_* \in \mathcal{X}) \right\},$$

где $\|\cdot\|$ – заданная норма.

⁶ Помимо ϵ и самой оценки, ϵ -риск оценки зависит от “параметров” $\|\cdot\|$, \mathcal{X} и \mathcal{N} ; эти элементы ясны из контекста и поэтому не упоминаются в обозначении $\text{Risk}_\epsilon[\cdot]$.

Предположение о шуме наблюдения. Шум наблюдения ξ предполагается субгауссовским с параметрами $(0, \sigma^2 I_m)$, т.е.

$$\mathbf{E} \left\{ e^{h^T \xi} \right\} \leq \exp \left(\frac{1}{2} \sigma^2 \|h\|_2^2 \right).$$

1.3. Эллиптопы

Геометрия множества сигналов \mathcal{X} и нормы $\|\cdot\|$ играет важную роль в анализе риска полиэдральной оценки. В дальнейшем ограничимся случаем, когда \mathcal{X} и поляр \mathcal{B}_* единичного шара нормы $\|\cdot\|$ являются эллиптопами.

По определению [2, 22] базовый эллитоп в \mathbf{R}^n – это множество вида

$$(3) \quad \mathcal{X} = \{x \in \mathbf{R}^n : \exists, t \in \mathcal{T} : x^T T_\ell x \leq t_\ell, \ell \leq L\},$$

где $T_\ell \in \mathbf{S}^k$, $T_\ell \succeq 0$, $\sum_\ell T_\ell \succ 0$, а $\mathcal{T} \subset \mathbf{R}_+^L$ – выпуклое компактное множество с непустой внутренностью, обладающее свойством монотонности: при $0 \leq t' \leq t \in \mathcal{T}$ имеем $t' \in \mathcal{T}$. Любой эллитоп – это образ базового эллитопа при некотором линейном отображении. Будем называть L эллитопической размерностью \mathcal{X} .

Очевидно, что каждый эллитоп является выпуклым компактным множеством, симметричным относительно начала координат; кроме того, базовый эллитоп имеет непустую внутренность.

Примеры.

А. Ограниченное пересечение \mathcal{X} L эллипсоидов/эллиптических цилиндров с центром в начале координат $\{x \in \mathbf{R}^n : x^T T_\ell x \leq 1\}$ [$T_\ell \succeq 0$] является базовым эллитопом:

$$\mathcal{X} = \{x \in \mathbf{R}^n : \exists t \in \mathcal{T} := [0, 1]^L : x^T T_\ell x \leq t_\ell, \ell \leq L\}.$$

В частности, единичный гиперкуб $\{x \in \mathbf{R}^n : \|x\|_\infty \leq 1\}$ является базовым эллитопом.

Б. Единичный шар $\|\cdot\|_p$ -нормы, где $p \in [2, \infty]$, является базовым эллитопом в \mathbf{R}^n :

$$\{x \in \mathbf{R}^n : \|x\|_p \leq 1\} = \left\{ x : \exists t \in \mathcal{T} = \{t \in \mathbf{R}_+^n, \|t\|_{p/2} \leq 1\} : \underbrace{x_\ell^2}_{x^T T_\ell x} \leq t_\ell, \ell \leq n \right\}.$$

Эллиптопы допускают полностью алгоритмическое “исчисление”: это семейство замкнуто относительно основных операций, сохраняющих выпуклость и симметрию относительно начала координат, например операций конечных пересечений, взятия линейного образа, обратного образа при линейном отображении, прямых произведений и арифметического суммирования (подробнее см. [2, раздел 4.6]).

Основное предположение. В дальнейшем считаем, что множество сигналов \mathcal{X} и поляра \mathcal{B}_* единичного шара нормы $\|\cdot\|$ являются *базовыми эллитопами*:⁷

$$(4a) \quad \mathcal{X} = \{x \in \mathbf{R}^n : \exists t \in \mathcal{T} : x^T T_k x \leq t_k, k \leq K\},$$

$$(4b) \quad \mathcal{B}_* = \{y \in \mathbf{R}^q : \exists s \in \mathcal{S} : y^T S_\ell y \leq s_\ell, \ell \leq L\},$$

где $\mathcal{T} \subset \mathbf{R}_+^K$, $\mathcal{S} \subset \mathbf{R}_+^L$ – монотонные, выпуклые и компактные множества с непустыми внутренностями, $T_k \succeq 0$, $\sum_k T_k \succ 0$, $S_\ell \succeq 0$ и $\sum_\ell S_\ell \succ 0$.

2. Случай ограниченной помехи

В этом разделе рассмотрим случай *ограниченной* неопределенности. В частности, считаем, что $\mathcal{N} \subset \mathbf{R}^n$ – выпуклое и компактное множество, симметричное относительно начала координат. Определим $\pi(\cdot)$ как полунорму в \mathbf{R}^m вида

$$\pi(h) = \sup_u \{(Nu)^T h : u \in \mathcal{N}\}.$$

2.1. Вычисление границ ϵ -риска полиэдральной оценки

В этом разделе полиэдральная оценка задается матрицей контрастов G размерности $m \times I$ следующим образом: при имеющемся наблюдении ω (см. (2)) найдем оптимальное решение $\hat{x}_G(\omega)$ (очевидно разрешимой) задачи оптимизации

$$(5) \quad \min_{x, \nu} \{\|G^T(Ax + N\nu - \omega)\|_\infty : x \in \mathcal{X}, \nu \in \mathcal{N}\}.$$

При заданной матрице $G = [g_1, \dots, g_I]$ размерности $m \times I$ пусть $\Xi_\epsilon[G]$ – множество всех реализаций ξ , удовлетворяющих условию

$$(6) \quad |[g_i^T \xi]_i| \leq \underbrace{\sigma \sqrt{2 \ln [2I/\epsilon]}}_{=: \kappa(\epsilon)} \|g_i\|_2, \quad \forall i \leq I.$$

Заметим, что

$$(7) \quad \text{Prob}_{\xi \sim \mathcal{SG}(0, \sigma^2 I_m)} \{\xi \notin \Xi_\epsilon(G)\} \leq \epsilon.$$

Действительно, $\mathbf{E}_\xi \left\{ e^{\gamma g^T \xi} \right\} \leq e^{\frac{1}{2} \gamma^2 \|g\|_2^2 \sigma^2}$; значит, для всех $a \geq 0$

$$\text{Prob}\{g^T \xi > a\} \leq \inf_{\gamma > 0} \exp \left\{ \frac{1}{2} \gamma^2 \|g\|_2^2 \sigma^2 - \gamma a \right\} = \exp \left\{ -\frac{1}{2} a^2 \sigma^2 \|g\|_2^{-2} \right\},$$

⁷ Последующие результаты прямо распространяются на случай, когда \mathcal{X} и \mathcal{B}_* являются “общими” эллитопами; полагаем их базовыми для компактных обозначений.

и в результате имеем

$$\text{Prob} \left\{ \exists i \leq I : |g_i^T \xi| > \varkappa(\epsilon) \|g_i\|_2 \right\} \leq 2I \exp \left\{ -\frac{\varkappa^2(\epsilon)}{2\sigma^2} \right\} \leq \epsilon.$$

При заданной матрице контрастов $G = [g_1, \dots, g_I]$ размерности $m \times I$ рассмотрим задачу оптимизации

$$(8) \quad \text{Opt}[G] = \min_{\lambda, \mu, \gamma} \left\{ f_G(\lambda, \mu, \gamma) : \lambda \geq 0, \mu \geq 0, \gamma \geq 0, \right. \\ \left. \left[\begin{array}{c|c} \sum_{\ell} \lambda_{\ell} S_{\ell} & \frac{1}{2} B \\ \hline \frac{1}{2} B^T & \sum_k \mu_k T_k + A^T \left[\sum_i \gamma_i g_i g_i^T \right] A \end{array} \right] \succeq 0 \right\},$$

где

$$f_G(\lambda, \mu, \gamma) = \phi_S(\lambda) + 4\phi_T(\mu) + 4\psi^2[G] \sum_i \gamma_i$$

с

$$\psi[G] = \max_i \pi(g_i) + \varkappa(\epsilon) n[G].$$

Предложение 1. Пусть (λ, μ, γ) – допустимое решение задачи (8). Тогда

$$\text{Risk}_{\epsilon}[\hat{w}_G] \leq f_G(\lambda, \mu, \gamma),$$

т.е. $f_G(\lambda, \mu, \gamma)$ – верхняя граница ϵ -риска оценки \hat{w}_G .

Доказательство. Зафиксируем $\xi \in \Xi_{\epsilon}[G]$, $x_* \in \mathcal{X}$ и $\eta_* \in \mathcal{N}$. Кроме того, пусть $\hat{x} = \hat{x}_G(\omega)$ является x -компонентой некоторого оптимального решения $[\hat{x}; \hat{\nu}]$, $\hat{\nu} \in \mathcal{N}$, задачи (5) и пусть $\Delta = \hat{x} - x_*$. Очевидно, что пара $[x, \nu] = [x_*, \nu_*]$ допустима для (5) и $\|G^T[Ax_* + N\nu_* - \omega]\|_{\infty} = \|G^T \xi\|_{\infty} \leq \varkappa(\epsilon) n[G]$, из чего следует, что также $\|G^T[A\hat{x} + N\hat{\nu} - \omega]\|_{\infty} \leq \varkappa(\epsilon) n[G]$. Поэтому

$$\|G^T A \Delta\|_{\infty} \leq 2\varkappa(\epsilon) n[G] + \|G^T N[\hat{\nu} - \nu_*]\|_{\infty}.$$

Учитывая, что $\hat{\nu}, \nu_* \in \mathcal{N}$, имеем $\|G^T N[\hat{\nu} - \nu_*]\|_{\infty} \leq 2 \max_i \pi(g_i)$ и приходим к

$$(9) \quad |g_i^T A \Delta| \leq 2\psi[G], \quad i = 1, \dots, I.$$

Теперь пусть $\Delta \in 2\mathcal{X}$, т.е. для некоторого $t \in \mathcal{T}$ и всех k выполняется $\Delta^T T_k \Delta \leq 4t_k$; кроме того, пусть $v \in \mathcal{B}_*$ таким образом, что для некоторого $s \in \mathcal{S}$ и всех ℓ справедливо неравенство $v^T S_{\ell} v \leq s_{\ell}$. Ограничение неотрицательной определенности задачи (8) означает, что

$$v^T B \Delta \leq v^T \left[\sum_{\ell} \lambda_{\ell} S_{\ell} \right] v + \Delta^T \left[\sum_k \mu_k T_k \right] \Delta + [A \Delta]^T \sum_i \gamma_i g_i g_i^T A \Delta \leq \\ \leq \sum_{\ell} \lambda_{\ell} s_{\ell} + 4 \sum_k \mu_k t_k + \sum_i \gamma_i (g_i^T A \Delta)^2 \leq \phi_S(\lambda) + 4\phi_T(\mu) + \sum_i \gamma_i (g_i^T A \Delta)^2 \leq \\ \leq \phi_S(\lambda) + 4\phi_T(\mu) + 4\psi^2[G] \sum_i \gamma_i.$$

Максимизируя левую часть полученного неравенства по $v \in \mathcal{B}_*$, в итоге получаем $\|B\Delta\| \leq f_G(\lambda, \mu, \gamma)$.

Заметим, что задача в правой части (8) (очевидно) является выпуклой задачей оптимизации и $\text{Opt}[G]$ эффективно вычислима при условии, что таковы функции ϕ_S, ϕ_T и π . Таким образом, предложение 1 дает эффективно вычисляемую верхнюю границу для ϵ -риска полиэдральной оценки при заданной матрице контрастов G и, по сути, удобный вычислительный инструмент для *анализа* эффективности полиэдральной оценки. К сожалению, этот инструмент не позволяет *строить* такую оценку с помощью оптимизации матрицы контрастов: попытка сделать G переменной, а не параметром, в правой части (8), приводит к невыпуклой, а значит трудно решаемой, задаче оптимизации. Теперь рассмотрим две ситуации, в которых эту трудность можно преодолеть.

2.2. Синтез матриц контрастов для оценок гарантированного качества

Рассмотрим два типа предположений о множестве \mathcal{N} помех, которые позволяют эффективно вычислять “вполне хорошие” матрицы контрастов. А именно:

- 1) “эллиптический случай”, когда \mathcal{N} является базовым эллипсоидом;
- 2) “коэллиптический случай”, когда множество $N\mathcal{N} = \{N\nu : \nu \in \mathcal{N}\}$ является *полярной* эллипсоидом

$$\begin{aligned} \mathcal{N}_* &= \{w \in \mathbf{R}^m : \exists \bar{r} \in \bar{\mathcal{R}} : w^T \bar{R}_j w \leq \bar{r}_j, j \leq \bar{J}\}, \\ &\left[\bar{R}_j \succeq 0, \sum_j \bar{R}_j \succ 0; \right. \\ &\left. \bar{\mathcal{R}} \subset \mathbf{R}_+^{\bar{J}}, \text{int} \bar{\mathcal{R}} \neq \emptyset - \text{—монотонный выпуклый компакт} \right]. \end{aligned}$$

Укажем, что \mathcal{N}_* есть единичный шар нормы $\pi(g) = \max_{\nu \in \mathcal{N}} g^T N\nu$.

2.2.1. Эллиптический случай

Отметим, что *эллиптический случай может быть непосредственно сведен к случаю отсутствия помех*. Действительно, когда \mathcal{N} является эллипсоидом, прямое произведение $\bar{\mathcal{X}} = \mathcal{X} \times \mathcal{N}$ — также эллипсоид. Таким образом, при задании $\bar{A}[x; \nu] = Ax + N\nu$ и $\bar{B}[x; \nu] = Bx$ наблюдение (2) принимает вид

$$\omega = \bar{A}\bar{x}_* + \xi, \quad [\bar{x}_* = [x_*; \nu_*] \in \bar{\mathcal{X}}],$$

и задача заключается в восстановлении по такому наблюдению линейного образа $w_* = \bar{B}\bar{x}_*$ нового сигнала \bar{x}_* . Построение близких к минимаксным, когда $\xi \sim \mathcal{N}(0, \sigma^2 I_m)$, полиэдральных оценок в этой постановке рассмотрено в [2]. Кратко опишем идею построения, поскольку она будет использоваться ниже.

Рассмотрим случай, когда $\mathcal{N} = \{0\}$, и пусть множество сигналов и норма $\|\cdot\|$ по-прежнему заданы (4). В этой ситуации задача (8) сводится к

$$(10) \quad \text{Opt}[G] = \min_{\lambda, \mu, \gamma} \left\{ \phi_S(\lambda) + 4\phi_T(\mu) + 4\kappa^2(\epsilon)n^2[G] \sum_i \gamma_i : \lambda \geq 0, \mu \geq 0, \gamma \geq 0, \right. \\ \left. \left[\begin{array}{c|c} \sum_\ell \lambda_\ell S_\ell & \frac{1}{2}B \\ \hline \frac{1}{2}B^T & \sum_k \mu_k T_k + A^T \left[\sum_i \gamma_i g_i g_i^T \right] A \end{array} \right] \succeq 0 \right\}.$$

Заметим, что $\text{Opt}[G] = \text{Opt}[\theta G]$ при $\theta > 0$. Действительно, (λ, μ, γ) является допустимым решением задачи, определяющей $\text{Opt}[G]$, тогда и только тогда, когда $(\lambda, \mu, \theta^2 \gamma)$ является допустимым решением задачи, определяющей $\text{Opt}[\theta G]$, причем значения соответствующих целевых функций при этих решениях совпадают. Отсюда следует, что при оптимизации $\text{Opt}[G]$ по G можно без потери качества ограничиться матрицами контрастов G с $\kappa(\epsilon)n[G] = 1$. Другими словами, задавая

$$(11) \quad \theta(g) = \kappa(\epsilon)\|g\|_2$$

и дополняя переменные λ, μ и γ в (10) переменными $g_i, \theta(g_i) \leq 1, i = 1, \dots, I$, приходим к задаче

$$(12) \quad \text{Opt} = \min_{\lambda, \mu, \gamma, \{g_i\}, \rho} \left\{ \phi_S(\lambda) + 4\phi_T(\mu) + 4\rho : \lambda \geq 0, \mu \geq 0, \gamma \geq 0, \right. \\ \left. \theta(g_i) \leq 1, \sum_i \gamma_i \leq \rho, \left[\begin{array}{c|c} \sum_\ell \lambda_\ell S_\ell & \frac{1}{2}B \\ \hline \frac{1}{2}B^T & \sum_k \mu_k T_k + A^T \left[\sum_i \gamma_i g_i g_i^T \right] A \end{array} \right] \succeq 0 \right\}.$$

(Напомним, что параметр G рассматривается как переменная для минимизации $\text{Opt}[G]$ по G .) Теперь, объединяя переменные γ и g_1, \dots, g_I в матрицу $\Theta = \sum_i \gamma_i g_i g_i^T$, обозначим через \mathfrak{T} множество пар $(\Theta \in \mathbf{S}_+^m, \rho)$, для которых существует разложение $\Theta = \sum_{i \leq I} \gamma_i g_i g_i^T$ с $\theta(g_i) \leq 1$ и $\gamma_i \geq 0, \sum_i \gamma_i \leq \rho$. Тогда задача (12) может быть записана как задача оптимизации

$$(13) \quad \text{Opt} = \min_{\lambda, \mu, \Theta, \rho} \left\{ \phi_S(\lambda) + 4\phi_T(\mu) + 4\rho : \lambda \geq 0, \mu \geq 0, \right. \\ \left. (\Theta, \rho) \in \mathfrak{T}, \left[\begin{array}{c|c} \sum_\ell \lambda_\ell S_\ell & \frac{1}{2}B \\ \hline \frac{1}{2}B^T & \sum_k \mu_k T_k + A^T \Theta A \end{array} \right] \succeq 0 \right\}.$$

Укажем, что если $I \geq m$, то \mathfrak{T} — простой выпуклый конус:

$$\mathfrak{T} = \{(\Theta, \rho) : \Theta \succeq 0, \rho \geq \kappa^2(\epsilon)\text{Tr}(\Theta)\}.$$

Следовательно, (13) является (очевидно разрешимой) задачей выпуклой оптимизации. Чтобы преобразовать оптимальное решение $(\lambda^*, \mu^*, \Theta^*, \rho^*)$ для (13) в оптимальное решение для (12), достаточно разложить Θ^* по собственным значениям, т.е. $\Theta^* = \sum_{i=1}^I v_i e_i e_i^T$ с $\|e_i\|_2 = 1$ и $v_i \geq 0, i \in \{1, \dots, m\}$, и $e_i = 0, v_i = 0, i \in \{m, \dots, I\}$, а также задать $g_i^* = \kappa^{-1}(\epsilon)e_i, \gamma_i^* = \kappa^2(\epsilon)v_i$. В результате приходим к оптимальному решению $(\lambda^*, \mu^*, \{g_i^*, \gamma_i^*\}_{i \leq I}, \rho^*)$ задачи (12).

2.2.2. Коэллитопический случай

Вышеизложенный подход, позволяющий свести невыпуклую задачу (12) (синтез наилучшей матрицы контрастов G в смысле $\text{Opt}[G]$) к (очевидно) выпуклой задаче оптимизации в значительной степени использует тот факт, что единичный шар нормы $\theta(\cdot)$ (ср. (11)) является простым эллитопом, кратным единичному евклидову шару. В силу данного обстоятельства \mathfrak{T} – эффективно вычислимый выпуклый конус. Дальнейшие построения основаны на том, что когда единичный шар нормы $\theta(\cdot)$ является базовым эллитопом, происходит нечто подобное: связанное с ним множество \mathfrak{T} , хотя и не обязательно выпуклое и допускающее эффективное вычисление, может быть хорошо аппроксимировано эффективно вычислимым выпуклым конусом. Основной результат ([36, Предложение 3.2], с точностью до принятых обозначений) состоит в следующем:

Предложение 2. Пусть $I \geq m$ и $\mathcal{W} \subset \mathbf{R}^m$ – базовый эллитоп:

$$\mathcal{W} = \{w \in \mathbf{R}^m : \exists r \in \mathcal{R} : w^T R_j w \leq r_j, j \leq J\},$$

$$\left[R_j \succeq 0, \sum_j R_j \succ 0; \mathcal{R} \subset \mathbf{R}_+^J, \text{int} \mathcal{R} \neq \emptyset \text{ — монотонный выпуклый компакт} \right].$$

Свяжем с \mathcal{W} замкнутый выпуклый конус⁸

$$\mathbf{K} = \{(\Theta, \rho) : \exists r \in \mathcal{R} : \text{Tr}(\Theta R_j) \leq \rho r_j, 1 \leq j \leq J, \Theta \succeq 0, \rho \geq 0\}.$$

Если матрица $\Theta \in \mathbf{S}_+^m$ представима в виде $\sum_i \gamma_i w_i w_i^T$ с $\gamma_i \geq 0$ и $w_i \in \mathcal{W}$, то $(\Theta, \sum_{i=1}^I \gamma_i) \in \mathbf{K}$, и “почти наоборот”: при $(\Theta, \rho) \in \mathbf{K}$ можно эффективно (с помощью рандомизированного алгоритма) найти векторы $w_i \in \mathcal{W}$, а также числа $\gamma_i \geq 0$, $i \leq I$ такие, что $\Theta = \sum_i \gamma_i w_i w_i^T$ и

$$\sum_i \gamma_i \leq 2\sqrt{2} \ln(4m^2 J) \rho.$$

Теперь построим “вполне хороший” контраст в коэллитопическом случае. Положим $R_j = \frac{1}{4} \bar{R}_j$, $j \leq \bar{J}$, $R_{\bar{J}+1} = \frac{\varkappa^2(\epsilon)}{4} I_m$ и рассмотрим эллитоп

$$\begin{aligned} \mathcal{W} &= 2 \left[\mathcal{N}_* \cap \{w : \varkappa(\epsilon) \|w\|_2 \leq 1\} \right] = \\ (14) \quad &= \{w \in \mathbf{R}^m : \exists r \in \mathcal{R} = \bar{\mathcal{R}} \times [0, 1] : w^T R_j w \leq r_j, j \leq J = \bar{J} + 1\}. \end{aligned}$$

Пусть $\theta(\cdot)$ – норма в \mathbf{R}^n с единичным шаром \mathcal{W} . Заметим, что $\theta(\cdot) = 2 \max[\pi(\cdot), \varkappa(\epsilon) \|\cdot\|_2]$ и для каждого $G = [g_1, \dots, g_I]$ величина $\psi[G]$ (8) ограничена сверху $\max_i \theta(g_i)$; эта граница точна с точностью до множителя 2.

⁸ \mathbf{K} – это действительно замкнутый выпуклый конус – коническая оболочка выпуклого компактного множества $\{\Theta \succeq 0 : \exists r \in \mathcal{R} : \text{Tr}(\Theta R_j) \leq r_j, 1 \leq j \leq J\} \times \{1\}$.

Следовательно, согласно предложению 1 ϵ -риск полиэдральной оценки с матрицей контрастов G ограничен сверху величиной

$$(15) \quad \overline{\text{Opt}}[G] = \min_{\lambda, \mu, \gamma} \left\{ \phi_{\mathcal{S}}(\lambda) + 4\phi_{\mathcal{T}}(\mu) + 4 \left[\max_i \theta(g_i) \right]^2 \sum_i \gamma_i : \lambda \geq 0, \mu \geq 0, \gamma \geq 0, \right. \\ \left. \left[\begin{array}{c|c} \sum_{\ell} \lambda_{\ell} S_{\ell} & \frac{1}{2} B \\ \hline \frac{1}{2} B^T & \sum_k \mu_k T_k + A^T \left[\sum_i \gamma_i g_i g_i^T \right] A \end{array} \right] \succeq 0 \right\}$$

и $\text{Opt}[G] \leq \overline{\text{Opt}}[G] \leq 2\text{Opt}[G]$. Как и в предыдущем разделе, задача минимизации $\overline{\text{Opt}}[G]$ по G может быть переформулирована в виде (13). Следующий результат дает вычислительно эффективный способ построить эффективную аппроксимацию оптимального решения этой задачи.

Пусть $I \geq m$, $\alpha = 2\sqrt{2} \ln(4m^2 J)$ и

$$\mathbf{K} = \{(\Theta, \rho) : \exists r \in \mathcal{R} : \text{Tr}(\Theta R_j) \leq \rho r_j, 1 \leq j \leq J, \Theta \succeq 0, \rho \geq 0\}$$

(см. (14)). Рассмотрим выпуклую задачу оптимизации

$$(16) \quad \text{Opt}_* = \min_{\lambda, \mu, \gamma, \Theta, \rho} \left\{ \phi_{\mathcal{S}}(\lambda) + 4\phi_{\mathcal{T}}(\mu) + 4\alpha\rho : \lambda \geq 0, \mu \geq 0, \right. \\ \left. (\Theta, \rho) \in \mathbf{K}, \left[\begin{array}{c|c} \sum_{\ell} \lambda_{\ell} S_{\ell} & \frac{1}{2} B \\ \hline \frac{1}{2} B^T & \sum_k \mu_k T_k + A^T \Theta A \end{array} \right] \succeq 0 \right\}.$$

Теорема 1. Θ -компонента Θ^* оптимального решения (разрешимой) задачи (16) может быть эффективным способом преобразована в матрицу контрастов G^* , удовлетворяющую условию

$$\overline{\text{Opt}}[G^*] \leq \sqrt{\alpha} \min_G \overline{\text{Opt}}[G] \leq 2\sqrt{\alpha} \min_G \text{Opt}[G].$$

В частности, ϵ -риск полиэдральной оценки с матрицей контрастов G^* (этот риск ограничен сверху $\overline{\text{Opt}}[G^*]$) не превышает $2\sqrt{\alpha} \min_G \text{Opt}[G]$.

Доказательство. Повторяя рассуждения предыдущего раздела, приходим к выводу, что $\overline{\text{Opt}} := \inf_G \overline{\text{Opt}}[G]$ равно

$$\inf_{g_1, \dots, g_I} \left\{ \overline{\text{Opt}} / ([g_1, \dots, g_I]) : \max_i \theta(g_i) = 1 \right\}.$$

Последний \inf очевидно достигается на определенном наборе g_1^+, \dots, g_I^+ с $\max_i \theta(g_i^+) = 1$. Пусть $G^+ = [g_1^+, \dots, g_I^+]$ и $\lambda^+, \mu^+, \gamma_i^+, i \leq I$, – оптимальное решение задачи в правой части (15), связанное с $g_i = g_i^+, i \leq I$, и пусть $\Theta^+ = \sum_i \gamma_i^+ [g_i^+] [g_i^+]^T$ и $\rho^+ = \sum_i \gamma_i^+$. Ясно, что

$$\overline{\text{Opt}} = \overline{\text{Opt}}[G^+] = \phi_{\mathcal{S}}(\lambda^+) + 4\phi_{\mathcal{T}}(\mu^+) + 4\rho^+.$$

Кроме того, в рассматриваемом случае неравенство $\theta(g) \leq 1$ эквивалентно $g \in \mathcal{W}$, а значит, $(\Theta^+, \rho^+) \in \mathbf{K}$ по первому утверждению предложения 2.

Отсюда следует, что $(\lambda^+, \mu^+, \Theta^+, \rho^+)$ – допустимое решение задачи оптимизации в (16). В силу структуры последней задачи при $\kappa > 0$ набор $(\kappa^{-1}\lambda^+, \kappa\mu^+, \kappa\Theta^+, \kappa\rho^+)$ допустим для (16) с соответствующим значением целевой функции $\kappa^{-1}\phi_S(\lambda^+) + \kappa[\phi_T(\mu^+) + 4\alpha\rho^+]$. Поэтому имеем

$$\begin{aligned} \text{Opt}_* &\leq \inf_{\kappa > 0} [\kappa^{-1}\phi_S(\lambda^+) + \kappa[\phi_T(\mu^+) + 4\alpha\rho^+]] = \\ &= 2(\phi_S(\lambda^+) \underbrace{[4\phi_T(\mu^+) + 4\alpha\rho^+] }_{\leq \alpha[4\phi_T(\mu^+) + 4\rho^+]})^{1/2} \leq 2\sqrt{\phi_S(\lambda^+)[4\phi_T(\mu^+) + 4\rho^+]}\sqrt{\alpha} \leq \\ &\leq \sqrt{\alpha}[\phi_S(\lambda^+) + 4\phi_T(\mu^+) + 4\rho^+] = \sqrt{\alpha}\overline{\text{Opt}}. \end{aligned}$$

Наконец, пусть $\bar{\lambda}, \bar{\mu}, \bar{\Theta}, \bar{\rho}$ – оптимальное решение (16). Поскольку $(\bar{\Theta}, \bar{\rho}) \in \mathbf{K}$, по второму утверждению предложения 2 существует (и может быть эффективно найдено) разложение $\bar{\Theta} = \sum_i \bar{\gamma}_i \bar{g}_i [\bar{g}_i]^T$ с $\bar{g}_i \in \mathcal{W}$ (т.е. $\theta(\bar{g}_i) \leq 1$), $i \leq I$, $\bar{\gamma}_i \geq 0$, и $\sum_i \bar{\gamma}_i \leq \alpha\bar{\rho}$. Тогда ϵ -риск полиэдральной оценки с матрицей контрастов $\bar{G} = [\bar{g}_1, \dots, \bar{g}_I]$ имеет верхнюю границу $\overline{\text{Opt}}[\bar{G}]$. Однако $\bar{\lambda}, \bar{\mu}$ и $\{\bar{\gamma}_i\}$ образуют допустимое решение задачи, определяющей $\overline{\text{Opt}}[\bar{G}]$, и значение целевой функции на этом решении ограничено сверху величиной

$$\phi_S(\bar{\lambda}) + 4\phi_T(\bar{\mu}) + 4 \left[\max_i \theta(\bar{g}_i) \right] \sum_i \bar{\gamma}_i \leq \phi_S(\bar{\lambda}) + 4\phi_T(\bar{\mu}) + 4\alpha\bar{\rho} = \text{Opt}_*.$$

Таким образом, ϵ -риск полиэдральной оценки с матрицей контрастов \bar{G} не превышает

$$\text{Opt}_* \leq \sqrt{\alpha}\overline{\text{Opt}} \leq 2\sqrt{\alpha} \min_G \text{Opt}[G]. \quad \square$$

3. Наблюдения с выбросами

Рассмотрим задачу оценивания, поставленную в разделе 1.2, в ситуации разреженной помехи ν_* в (2), т.е. когда она имеет не более чем s ненулевых элементов.

Построение оценки. Пусть $\epsilon \in (0, 1)$ – заданный допуск надежности. Рассмотрим следующую полиэдральную оценку, определенную двумя матрицами контрастов $H = [h_1, \dots, h_n] \in \mathbf{R}^{m \times n}$ и $G = [g_1, \dots, g_I] \in \mathbf{R}^{n \times I}$. При имеющемся наблюдении ω (см. (2)) решим задачу оптимизации

$$(17) \quad \min_{\nu, x} \left\{ \|\nu\|_1 : x \in \mathcal{X}, \begin{array}{l} |h_k^T [N\nu + Ax - \omega]| \leq \bar{\alpha}(\epsilon) \|h_k\|_2, \quad k = 1, \dots, n, \\ |g_i^T [N\nu + Ax - \omega]| \leq \bar{\alpha}(\epsilon) \|g_i\|_2, \quad i = 1, \dots, I \end{array} \right\},$$

где

$$\bar{\alpha}(\epsilon) = \sigma \sqrt{2 \ln[2(n+I)/\epsilon]}.$$

Пусть $(\hat{\nu}, \hat{x}) = (\hat{\nu}(\omega), \hat{x}(\omega))$ – оптимальное решение задачи, если она разрешима; в противном случае положим $(\hat{\nu}, \hat{x}) = (0, 0)$. В качестве оценки $w_* = Bx_*$ будем использовать вектор

$$\hat{w}_{G,H}(\omega) = B\hat{x}(\omega).$$

3.1. Анализ рисков

Обозначим через $\Xi_\epsilon(G, H)$ множество реализаций ξ , удовлетворяющих условию

$$(18) \quad \begin{aligned} |h_k^T \xi| &\leq \overline{\alpha}(\epsilon) \|h_k\|_2, \quad k = 1, \dots, n, \\ |g_i^T \xi| &\leq \overline{\alpha}(\epsilon) \|g_i\|_2, \quad i = 1, \dots, I, \end{aligned} \quad \forall \xi \in \Xi_\epsilon(G, H).$$

На том же основании, что и в (7), получаем

$$\text{Prob}_{\xi \sim \mathcal{S}G(0, \sigma^2 I_m)}(\Xi_\epsilon(G, H)) \geq 1 - \epsilon.$$

Теперь зафиксируем $x_* \in \mathcal{X}$, s -разреженный сигнал ν_* и $\xi \in \Xi_\epsilon(G, H)$; таким образом, наблюдение принимает вид $\omega = Ax_* + N\nu_* + \xi$.

А. В силу (18) имеем $|h_k^T \xi| \leq \overline{\alpha}(\epsilon) \|h_k\|_2$ и $|g_i^T \xi| \leq \overline{\alpha}(\epsilon) \|g_i\|_2$ для всех $k \leq n$ и $i \leq I$. Тогда задача (17) сводится к

$$(19) \quad \min_{\nu, x} \left\{ \|\nu\|_1 : x \in \mathcal{X}, \right. \\ \left. \begin{aligned} |h_k^T [N[\nu - \nu_*] + A[x - x_*] - \xi]| &\leq \overline{\alpha}(\epsilon) \|h_k\|_2, \quad k = 1, \dots, n, \\ |g_i^T [N[\nu - \nu_*] + A[x - x_*] - \xi]| &\leq \overline{\alpha}(\epsilon) \|g_i\|_2, \quad i = 1, \dots, I \end{aligned} \right\}.$$

Заключаем, что $(\nu, x) = (\nu_*; x_*)$ – допустимое решение для (17). Таким образом, в рассматриваемом случае $\widehat{\nu}$ и \widehat{x} являются допустимыми для (19) и

$$\|\widehat{\nu}\|_1 \leq \|\nu_*\|_1.$$

Б. В дальнейшем предполагается, что $(H, \|\cdot\|_\infty)$ удовлетворяет условию $Q_\infty(s, \kappa)$ из раздела 2 с $\kappa < \frac{1}{2}$, т.е.⁹

$$(20) \quad \|w\|_\infty \leq \|H^T N w\|_\infty + \frac{\kappa}{s} \|w\|_1 \quad \forall w \in \mathbf{R}^n.$$

Поскольку $\widehat{\nu}$ и \widehat{x} допустимы для (19), имеем

$$|h_k^T [N[\widehat{\nu} - \nu_*] + A[\widehat{x} - x_*] - \xi]| \leq \overline{\alpha}(\epsilon) \|h_k\|_2, \quad \forall k \leq n.$$

Задействуя (18) и $A[\widehat{x} - x_*] \in 2A\mathcal{X}$ (в силу симметрии \mathcal{X} относительно начала координат), заключаем, что

$$\|H^T N[\widehat{\nu} - \nu_*]\|_\infty \leq \max_k \left[\overline{\alpha}(\epsilon) \|h_k\|_2 + 2 \max_{x \in \mathcal{X}} |h_k^T A x| \right];$$

⁹ Условие $Q_\infty(s, \kappa)$ является простейшим (и наиболее ограничительным) в семействе условий $Q_q(s, \kappa)$, $q \in [1, \infty]$, используемых для установления свойств ℓ_1 -восстановления разреженных сигналов. Это условие может быть эффективно проверено, что очень важно здесь. В [2, раздел 1.3] рассмотрены эффективно проверяемые условия в разреженном восстановлении и их связь с другими используемыми условиями (свойство ограниченной изометрии (Restricted Isometry Property, RIP) [37], ограниченное собственное значение (Restricted Eigenvalue, RE) [38], взаимная некогерентность (Mutual Incoherence, MI) [39], и совместимость (Compatibility) [40]).

кроме того, вектор ν_* является s -разреженным и $\|\widehat{\nu}\|_1 \leq \|\nu_*\|_1$. Теперь предложение 5 с ν_* в роли ν дает

$$(21) \quad \|\widehat{\nu} - \nu_*\|_q \leq \frac{(2s)^{\frac{1}{q}}}{1 - 2\kappa} \max_k \left[\overline{\varkappa}(\epsilon) \|h_k\|_2 + 2 \max_{x \in \mathcal{X}} |h_k^T A x| \right], \quad 1 \leq q \leq \infty;$$

в частности,

$$(22a) \quad \|\widehat{\nu} - \nu_*\|_\infty \leq \frac{1}{1 - 2\kappa} \max_k \left[\overline{\varkappa}(\epsilon) \|h_k\|_2 + 2 \max_{x \in \mathcal{X}} |h_k^T A x| \right] =: \rho_H,$$

$$(22b) \quad \|\widehat{\nu} - \nu_*\|_1 \leq \frac{2s}{1 - 2\kappa} \max_k \left[\overline{\varkappa}(\epsilon) \|h_k\|_2 + 2 \max_{x \in \mathcal{X}} |h_k^T A x| \right] = 2s\rho_H.$$

Более того, согласно [2, Предложение 1.10] множество \mathcal{H} пар (H, κ) с матрицами H размерности $m \times n$, удовлетворяющими условию $Q_\infty(s, \kappa)$, является эффективно вычислимым выпуклым множеством

$$(23) \quad \mathcal{H} = \left\{ (H, \kappa) \in \mathbf{R}^{m \times n} \times \mathbf{R} : |[I_n - N^T H]_{ij}| \leq s^{-1} \kappa, 1 \leq i, j \leq n \right\}.$$

В. Поскольку $\widehat{\nu}$ и \widehat{x} допустимы для (19), имеем

$$|g_i^T (N[\widehat{\nu} - \nu_*] + A[\widehat{x} - x_*] - \xi)| \leq \overline{\varkappa}(\epsilon) \|g_i\|_2, \quad i = 1, \dots, I,$$

а $|g_i^T \xi| \leq \overline{\varkappa}(\epsilon) \|g_i\|_2 \forall i$ в силу $\xi \in \Xi_\epsilon(G, H)$. Заключаем, что

$$(24) \quad |g_i^T A[\widehat{x} - x_*]| \leq 2\overline{\varkappa}(\epsilon) \|g_i\|_2 + |g_i^T N[\widehat{\nu} - \nu_*]|, \quad i \leq I.$$

Пусть $\|z\|_{k,1}$, $z \in \mathbf{R}^n$, – сумма $\min[k, n]$ наибольших величин элементов в z ; отметим, что $\|\cdot\|_{k,1}$ является нормой, сопряженной к норме с единичным шаром $\{u : \|u\|_\infty \leq 1, \|u\|_1 \leq k\}$. Следовательно, из (22) получаем, что

$$(25) \quad |g_i^T N[\widehat{\nu} - \nu_*]| \leq \rho_H \|N^T g_i\|_{2s,1}$$

и на основании (24)

$$(26) \quad |g_i^T A[\widehat{x} - x_*]| \leq \psi_H[G], \quad \psi_H[G] = \max_i \left[2\overline{\varkappa}(\epsilon) \|g_i\|_2 + \rho_H \|N^T g_i\|_{2s,1} \right].$$

Пусть

$$f_{G,H}(\lambda, \mu, \gamma) = \phi_S(\lambda) + 4\phi_T(\mu) + \psi_H^2[G] \sum_i \gamma_i;$$

рассмотрим задачу оптимизации (ср. (8))

$$(27) \quad \text{Opt}[G, H] = \min_{\lambda, \mu, \gamma} \left\{ f_{G,H}(\lambda, \mu, \gamma) : \lambda \geq 0, \mu \geq 0, \gamma \geq 0, \left[\frac{\sum_\ell \lambda_\ell S_\ell}{\frac{1}{2} B^T} \mid \frac{\frac{1}{2} B}{\sum_k \mu_k T_k + A^T [\sum_i \gamma_i g_i g_i^T] A} \right] \succeq 0 \right\}.$$

Применяя те же рассуждения, что и при доказательстве предложения (1), с (26) в роли (9), устанавливаем следующий результат.

Предложение 3. В условиях текущего раздела при заданных $\kappa \in (0, 1/2)$ и матрице H размерности $m \times n$, удовлетворяющей $(H, \kappa) \in \mathcal{H}$ (23), пусть (λ, μ, γ) – допустимое решение (27). Тогда

$$\text{Risk}_\epsilon[\widehat{w}_{G,H}] \leq f_{G,H}(\lambda, \mu, \gamma).$$

3.2. Синтез матриц контрастов

Цель авторов – получить матрицы контрастов H и G с небольшой верхней границей $\text{Opt}[G, H]$ для ϵ -риска оценки $\widehat{w}_{G,H}$.

Г. Построение матрицы контрастов $H \in \mathbf{R}^{m \times n}$ не вызывает затруднений: граница $\text{Opt}[G, H]$ риска зависит от $H = [h_1, \dots, h_n]$ исключительно через величину

$$\rho_H = \frac{1}{1 - 2\kappa} \max_{k \leq n} \left[\overline{\alpha}(\epsilon) \|h_k\|_2 + 2 \max_{x \in \mathcal{X}} \|h_k^T A x\|_\infty \right]$$

и будет тем меньше, чем ниже ρ_H . Для фиксированного числа $\kappa \in (0, 1/2)$ вполне хороший выбор $H = [h_1, \dots, h_n]$ задается оптимальными решениями n задач оптимизации

$$(28) \quad h_k = \underset{h}{\text{argmin}} \left\{ \overline{\alpha}(\epsilon) \|h\|_2 + 2 \max_{x \in \mathcal{X}} \|h^T A x\|_\infty : h \in \mathbf{R}^m, \right. \\ \left. \|\text{Col}_i[I_n - N^T h]\|_\infty \leq s^{-1} \kappa \right\}.$$

В силу формы \mathcal{X} и конической двойственности эти задачи эквивалентны задачам

$$h_k = \underset{h, v, \chi}{\text{argmin}} \left\{ \overline{\alpha}(\epsilon) \|h\|_2 + v + \phi_{\mathcal{T}}(\chi) : h \in \mathbf{R}^m, \chi \geq 0, \right. \\ \left. \left[\begin{array}{c|c} v & h^T A \\ \hline A^T h & \sum_k \chi_k T_k \end{array} \right] \succeq 0, \|\text{Col}_i[I_n - N^T h]\|_\infty \leq s^{-1} \kappa \right\}, \quad 1 \leq k \leq n.$$

Д. Предлагаемое построение G менее простое. Действуем следующим образом. Пусть $G = [G_1, G_2]$, где $G_2, G_1 \in \mathbf{R}^{m \times m}$ (таким образом, $I = 2m$).

Д.1 Заметим, что поскольку $\xi \in \Xi_\epsilon(G, H)$, задача (19) разрешима, а $(\widehat{x}, \widehat{v})$ – ее допустимое решение. Для столбца g из G согласно ограничениям задачи имеем

$$(29) \quad |g^T A[\widehat{x} - x_*]| \leq 2\overline{\alpha}(\epsilon) \|g\|_2 + |g^T N[\widehat{v} - \nu_*]| \leq 2\overline{\alpha}(\epsilon) \|g\|_2 + 2s\rho_H \|N^T g\|_\infty$$

(здесь использованы (24) и (25)). Отсюда следует, что

$$(30) \quad (g^T A[\widehat{x} - x_*])^2 \leq 2(4\overline{\alpha}^2(\epsilon) \|g\|_2^2 + 4s^2 \rho_H^2 \|N^T g\|_\infty^2), \quad i = 1, \dots, m.$$

Укажем, что множество

$$\mathcal{M} = \{g \in \mathbf{R}^m : 8\overline{\alpha}^2(\epsilon) \|g\|_2^2 + 8s^2 \rho_H^2 \|N^T g\|_\infty^2 \leq 1\}$$

представляет собой эллипс: обозначая $N = [n_1, \dots, n_n]$, приходим к

$$\mathcal{M} = \left\{ g \in \mathbf{R}^m : \exists r \in [0, 1]^n : g^T \underbrace{(8\bar{\varkappa}^2(\epsilon)I_m + 8s^2\rho_H^2 n_j n_j^T)}_{M_j} g \leq r_j, j = 1, \dots, n \right\}.$$

Д.2 Далее заметим, что если $\xi \in \Xi_\epsilon(G, H)$, на основании (21) получаем

$$\|\hat{\nu} - \nu_*\|_2 \leq \frac{\sqrt{2s}}{1 - 2\kappa} \max_{k \leq n} \left[\bar{\varkappa}(\epsilon) \|h_k\|_2 + \max_{x \in \mathcal{X}} |h_k^T A x| \right] = \sqrt{2s} \rho_H.$$

Тогда согласно (29) для столбца g матрицы G имеет место

$$\begin{aligned} (g^T A[\hat{x} - x_*])^2 &\leq (2\bar{\varkappa}(\epsilon) \|g\|_2 + |g^T N[\hat{\nu} - \nu_*]|)^2 \leq \\ &\leq \left(2\bar{\varkappa}(\epsilon) \|g\|_2 + \sqrt{2s} \rho_H \|N^T g\|_2 \right)^2 \leq \\ (31) \quad &\leq g^T (8\bar{\varkappa}^2(\epsilon)I_m + 4s\rho_H^2 N N^T) g. \end{aligned}$$

Теперь примем

$$(32) \quad Q = (8\bar{\varkappa}^2(\epsilon)I_m + 4s\rho_H^2 N N^T)^{-1/2}$$

и рассмотрим задачу оптимизации

$$(33a) \quad \text{Opt} = \min_{\lambda, \mu, \Theta_1, \Theta_2, \rho} \left\{ f_H(\lambda, \mu, \Theta_1, \Theta_2, \rho) : \lambda \geq 0, \mu \geq 0, \Theta_1 \succeq 0, \Theta_2 \succeq 0, \right. \\ \left. \text{Tr}(M_j \Theta_1) \leq \rho, j = 1, \dots, n, \right. \\ \left. \left[\begin{array}{c|c} \sum_\ell \lambda_\ell S_\ell & \frac{1}{2} B \\ \hline \frac{1}{2} B^T & \sum_k \mu_k T_k + A^T (\Theta_1 + Q \Theta_2 Q^T) A \end{array} \right] \succeq 0 \right\},$$

где

$$(33b) \quad f_H(\lambda, \mu, \Theta_1, \Theta_2, \rho) = \phi_S(\lambda) + 4\phi_T(\mu) + \text{Tr}(\Theta_2) + 2\sqrt{2} \ln(4m^2 n) \rho.$$

Заметим, что ограничения на Θ_1 и ρ задачи (33a) говорят именно о принадлежности (Θ_1, ρ) конусу \mathbf{K} , связанному (см. Предложение 2) с эллипсом \mathcal{M} в роли \mathcal{W} .

Теорема 2. При имеющемся допустимом решении $(\lambda, \mu, \tau, \Theta_1, \Theta_2)$ задачи (33) построим матрицы контрастов G_1 и G_2 размерности $m \times m$ следующим образом.

- Для построения G_1 применим вторую часть Предложения 2 к $\Theta_1, \rho, \mathcal{M}$ в роли $\Theta, \rho, \mathcal{W}$, чтобы получить (вычислительно эффективным образом) разложение $\Theta_1 = \sum_{i=1}^m \gamma_i g_{1,i} g_{1,i}^T$ с $g_{1,i} \in \mathcal{M}$ и $\gamma_i \geq 0$, $\sum_i \gamma_i \leq 2\sqrt{2} \ln(4m^2 n) \rho$. Положим $G_1 = [g_{1,1}, \dots, g_{1,m}]$.
- Для построения G_2 разложим Θ_2 по собственным значениям, т.е. $\Theta_2 = \Gamma \text{Diag}\{\chi\} \Gamma^T$, и зададим $G_2 = [g_{2,1}, \dots, g_{2,m}] = Q \Gamma$.

Заметим, что $\Theta_1 + Q \Theta_2 Q = \sum_i \gamma_i g_{1,i} g_{1,i}^T + \sum_i \chi_i g_{2,i} g_{2,i}^T$.

Для полученной полиэдральной оценки $\widehat{w}_{G,H}$ и для всех $x_* \in \mathcal{X}$, s -разреженных ν_* и $\xi \in \Xi_\epsilon(G, H)$ выполняется

$$(34) \quad \|\widehat{w}_{G,H}(Ax_* + N\nu_* + \xi) - Bx_*\| \leq f_H(\lambda, \mu, \Theta_1, \Theta_2, \rho).$$

Следовательно, ϵ -риск оценки ограничен сверху величиной $f_H(\lambda, \mu, \Theta_1, \Theta_2, \rho)$ (в силу $\xi \in \Xi_\epsilon(G, H)$ с вероятностью не менее $1 - \epsilon$).

Доказательство. Зафиксируем $x_* \in \mathcal{X}$, s -разреженный ν_* и $\xi \in \Xi_\epsilon(G, H)$, и пусть $w = Ax_* + N\nu_* + \xi$. На основании п. **A** задача (17) разрешима, $(\widehat{x}, \widehat{\nu}) = (\widehat{x}(w), \widehat{\nu}(w))$ – ее оптимальное решение, а $\widehat{w} = B\widehat{x}$ – оценка $\widehat{w}_{G,H}(w)$. Пусть $\Delta = \widehat{x} - x_*$, и пусть e_1, \dots, e_m – столбцы ортонормированной матрицы Γ . По построению G_2 имеем для всех $j \leq m$ (см. (31))

$$\begin{aligned} (g_{2,j}^T A \Delta)^2 &\leq g_{2,j}^T (8\overline{\mathcal{Z}}^2(\epsilon)I_m + 4s\rho_H^2 NN^T) g_{2,j} = \\ &= e_j^T [Q (8\overline{\mathcal{Z}}^2(\epsilon)I_m + 4s\rho_H^2 NN^T) Q] e_j = e_j^T e_j = 1. \end{aligned}$$

Более того, в силу $g_{1,i} \in \mathcal{M}$ имеем (см. (30))

$$(g_{1,i}^T A \Delta)^2 \leq 8\overline{\mathcal{Z}}^2(\epsilon) \|g\|_2^2 + 8s^2 \rho_H^2 \|N^T g\|_\infty^2 \leq 1 \quad \forall i \leq m.$$

Теперь согласно ограничению неотрицательной определенности в (33а) и $\Theta_1 + Q\Theta_2Q = \sum_i \gamma_i g_{1,i} g_{1,i}^T + \sum_i \chi_i g_{2,i} g_{2,i}^T$ для каждого $v \in \mathcal{B}_*$ получаем

$$\begin{aligned} v^T B \Delta &\leq v^T \left[\sum_\ell \lambda_\ell S_\ell \right] v + \Delta^T \left[\sum_k \mu_k T_k \right] \Delta + \\ &+ [A\Delta]^T \left[\sum_i \gamma_i g_{1,i} g_{1,i}^T + \sum_i \chi_i g_{2,i} g_{2,i}^T \right] A \Delta \leq \\ &\leq \phi_{\mathcal{S}}(\lambda) + 4\phi_{\mathcal{T}}(\mu) + \sum_i \chi_i (g_{1,i}^T A \Delta)^2 + \sum_j \gamma_j (g_{2,j}^T A \Delta)^2 \leq \\ &= [v^T S_1 v; \dots; v^T S_L v] \in \mathcal{S} \text{ в силу } v \in \mathcal{B}_* \text{ и} \\ &[\Delta^T T_1 \Delta; \dots; \Delta^T T_L \Delta] \in 4\mathcal{T} \text{ в силу } \Delta \in 2\mathcal{X} \leq \\ &\leq \phi_{\mathcal{S}}(\lambda) + 4\phi_{\mathcal{T}}(\mu) + \sum_i \chi_i + \sum_j \gamma_j \leq f_H(\lambda, \mu, \tau, \Theta_1, \Theta_2) \end{aligned}$$

благодаря $\sum_i \gamma_i \leq 2\sqrt{2} \ln(4m^2 n) \rho$ и $\sum_i \chi_i = \text{Tr}(\Theta_2)$. Вычислив в полученном неравенстве супремум по $v \in \mathcal{B}_*$, в итоге приходим к (34). \square

3.3. Альтернативное построение

Цель раздела – уточнить границы риска (27) и (33а) и построить улучшенные матрицы контраста. Действуем следующим образом. Во-первых, чтобы

расширить возможный выбор H -контрастов, ответственных за восстановление возмущений, улучшаем границы (22) для точности разреженного восстановления, в частности, позволяя использовать контрасты, не удовлетворяющие условию $Q_\infty(s, \kappa)$. Во-вторых, улучшаем границу для риска оценки $\hat{w}(\omega)$ благодаря учету вклада H -компоненты “полной” матрицы контраста $[H, G]$ при оптимизации G -компоненты контраста.

В дальнейшем рассматриваем оценку, описанную в начале раздела 3, с той лишь разницей, что размерности матриц контрастов G и H отличаются: теперь $H = [h_1, \dots, h_M] \in \mathbf{R}^{m \times M}$, а $G = [g_1, \dots, g_{2m}]$. Таким образом, в текущей постановке при имеющемся наблюдении ω решаем задачу оптимизации

$$(35) \quad \min_{\nu, x} \left\{ \|\nu\|_1 : x \in \mathcal{X}, \quad \begin{array}{l} |h_k^T(N\nu + Ax - \omega)| \leq \bar{\varkappa}(\epsilon) \|h_k\|_2, \quad k = 1, \dots, M, \\ |g_i^T(N\nu + Ax - \omega)| \leq \bar{\varkappa}(\epsilon) \|g_{2,i}\|_2, \quad i = 1, \dots, 2m \end{array} \right\}$$

с

$$\bar{\varkappa}(\epsilon) = \sigma \sqrt{2 \ln[(2M + 4m)/\epsilon]}$$

и указываем $\hat{x}(\omega), \hat{\nu}(\omega)$ в качестве оптимального решения задачи, если задача разрешима; в противном случае задаем $(\hat{x}(\omega), \hat{\nu}(\omega)) = (0, 0)$ и берем $\hat{w}_{G,H}(\omega) = B\hat{x}(\omega)$ в качестве оценки Bx_* .

3.3.1. Анализ риска

Приведенную выше задачу можно эквивалентно записать в виде

$$(36) \quad \min_{\nu, x} \left\{ \|\nu\|_1 : x \in \mathcal{X}, \quad \begin{array}{l} |h_k^T(N[\nu - \nu_*] + A[x - x_*] - \xi)| \leq \bar{\varkappa}(\epsilon) \|h_k\|_2, \quad k = 1, \dots, M, \\ |g_i^T(N[\nu - \nu_*] + A[x - x_*] - \xi)| \leq \bar{\varkappa}(\epsilon) \|g_i\|_2, \quad i = 1, \dots, 2m, \end{array} \right\}.$$

Задавая

$$(37) \quad \Xi_\epsilon(G, H) := \left\{ \xi \in \mathbf{R}^m : \begin{array}{l} |h_k^T \xi| \leq \bar{\varkappa}(\epsilon) \|h_k\|_2, \quad k = 1, \dots, M, \\ |g_i^T \xi| \leq \bar{\varkappa}(\epsilon) \|g_i\|_2, \quad i = 1, \dots, 2m \end{array} \right\},$$

получаем

$$\text{Prob}_{\xi \sim \mathcal{S}\mathcal{G}(0, \sigma^2 I_m)}(\Xi_\epsilon(G, H)) \geq 1 - \epsilon.$$

Зафиксируем $\xi \in \Xi_\epsilon(G, H)$, и пусть $\omega = Ax_* + N\nu_* + \xi$. Поскольку пара $(\hat{\nu}, \hat{x})$ является допустимой для (36), $\hat{x} := \hat{x}(\omega)$, $\hat{\nu} := \hat{\nu}(\omega)$ также допустима, $\|\hat{\nu}\|_1 \leq \|\nu_*\|_1$. Таким образом, по аналогии с доказательством Предложения 5, для $z = \hat{\nu} - \nu_*$ имеет место

$$\|z\|_1 \leq 2\|z\|_{s,1}$$

и, следовательно,

$$(38) \quad \|z\|_1 \leq 2s\|z\|_\infty, \quad \|z\|_2 \leq \sqrt{2s}\|z\|_\infty.$$

Теперь введем обозначение $\Delta = \widehat{x} - x_*$ и рассмотрим n пар выпуклых задач оптимизации

$$(P_2[i]) \quad \text{Opt}_2[i] = \max_{v,t,w} \left\{ \sqrt{w_i t} : v \in 2\mathcal{X}, \right. \\ \left. \begin{aligned} \|w\|_\infty \leq w_i, \|w\|_1 \leq t, t \leq 2sw_i, \\ |h_k^T(Nw + Av)| \leq 2\overline{\pi}(\epsilon)\|h_k\|_2, k = 1, \dots, M \end{aligned} \right\},$$

$$(P_\infty[i]) \quad \text{Opt}_\infty[i] = \max_{v,w} \left\{ w_i : v \in 2\mathcal{X}, \right. \\ \left. \begin{aligned} \|w\|_\infty \leq w_i, \|w\|_1 \leq 2sw_i, \\ |h_k^T(Nw + Av)| \leq 2\overline{\pi}(\epsilon)\|h_k\|_2, k = 1, \dots, M \end{aligned} \right\}.$$

Заметим, что допустимое решение (v, t, w) для $(P_2[i])$ удовлетворяет $\|w\|_\infty \leq w_i$ и $\|w\|_1 \leq t$, откуда

$$(39) \quad \|w\|_2 \leq \sqrt{w_i t} \leq \text{Opt}_2[i].$$

Пусть $\iota = \iota_z$ – индекс наибольшего по абсолютной величине элемента в z . Поскольку $\xi \in \Xi_\epsilon(G, H)$ и $\Delta \in 2\mathcal{X}$, заключаем, что при условии $z_\iota \geq 0$ тройка $(v, t, w) = (\Delta, \|z\|_1, z)$ допустима для $(P_2[\iota])$, а пара $(v, w) = (\Delta, z)$ допустима для $(P_\infty[\iota])$; если $z_\iota < 0$, то же самое справедливо для $(v, t, w) = (-\Delta, \|z\|_1, -z)$ и $(v, w) = (-\Delta, -z)$. Действительно, в первом случае $v = \Delta \in \mathcal{X}$, $|h_k^T[A\widehat{x} + N\widehat{v} - \omega]| \leq \overline{\pi}(\epsilon)\|h_k\|_2$ и $|h_k^T[Ax_* + N\nu_* - \omega]| \leq \overline{\pi}(\epsilon)\|h_k\|_2$, так как обе пары $(\widehat{x}, \widehat{v})$ и (x_*, ν_*) допустимы для (35), откуда следуют ограничения из второй строки $(P_2[i])$. Заметим, что в рассматриваемом случае $z_\iota = \|z\|_\infty$, т.е. ограничения из первой строки $(P_2[i])$ выполняются для $w = z$ в силу (38). Таким образом, тройка $(\Delta, \|z\|_1, z)$ действительно допустима для $(P_2[i])$. В качестве сопутствующего результата этих рассуждений получаем, что тройка $(\Delta, \|z\|_1, z)$ допустима для $(P_\infty[i])$. Во втором случае аргументация полностью аналогична.

Теперь положим

$$(40) \quad \text{Opt}_2 = \max_i \text{Opt}_2[i], \quad \text{Opt}_\infty = \max_i \text{Opt}_\infty[i]$$

и напомним, что тройка $(\Delta, \|z\|_1, z)$ или $(-\Delta, \|z\|_1, -z)$ допустима для некоторых задач $(P_2[i])$, а пара (Δ, z) или $(-\Delta, -z)$ допустима для некоторых задач $(P_\infty[i])$. Используя (39), для всех $\xi \in \Xi_\epsilon(G, H)$ получаем

$$\|z\|_\infty \leq \text{Opt}_\infty, \quad \|z\|_2 \leq \text{Opt}_2, \quad \|z\|_1 \leq 2s\text{Opt}_\infty.$$

Как следствие, для всех $d \in \mathbf{R}^m$ имеет место

$$(41) \quad |d^T Nz| \leq \max_z \{d^T Nz : \|z\|_\infty \leq \text{Opt}_\infty, \|z\|_2 \leq \text{Opt}_2, \|z\|_1 \leq 2s\text{Opt}_\infty\} = \\ = \underbrace{\min_{u,v,w} \{ \|u\|_1 \text{Opt}_\infty + \|v\|_2 \text{Opt}_2 + 2s\|w\|_\infty \text{Opt}_\infty, u + v + w = N^T d \}}_{=:\overline{\pi}(N^T d)}.$$

Поскольку пара $\widehat{x}, \widehat{\nu}$ допустима для (36) и, кроме того, $\xi \in \Xi_\epsilon(G, H)$, заключаем, что столбцы $d_i, i = 1, \dots, M+2m$, “агрегированной” матрицы контрастов $D = [G, H]$ удовлетворяют

$$|d_i^T A \Delta| \leq |d_i^T N z| + |d_i^T \xi| + \overline{\alpha}(\epsilon) \|g\|_2,$$

откуда

$$(42) \quad |d_i^T A \Delta| \leq \underbrace{\overline{\pi}(N^T d_i) + 2\overline{\alpha}(\epsilon) \|d_i\|_2}_{=: \psi_H(d_i)}, \quad i \leq M+2m.$$

Далее, пусть

$$\bar{f}_{G,H}(\lambda, \mu, \gamma) = \phi_S(\lambda) + 4\phi_T(\mu) + \sum_i \gamma_i \psi_H^2(d_i),$$

и рассмотрим задачу оптимизации (ср. (27))

$$(43) \quad \text{Opt}[G, H] = \min_{\lambda, \mu, \gamma} \left\{ \bar{f}_{G,H}(\lambda, \mu, \gamma) : \lambda \geq 0, \mu \geq 0, \gamma \geq 0, \right. \\ \left. \left[\begin{array}{c|c} \sum_\ell \lambda_\ell S_\ell & \frac{1}{2} B \\ \hline \frac{1}{2} B^T & \sum_k \mu_k T_k + A^T \left[\sum_i \gamma_i d_i d_i^T \right] A \end{array} \right] \succeq 0 \right\}.$$

Используя те же рассуждения, что и при доказательстве Предложения 1, с (42) в роли (9), устанавливаем следующий результат.

Предложение 4. В условиях текущего раздела пусть (λ, μ, γ) – допустимое решение для (43). Тогда

$$\text{Risk}_\epsilon[\widehat{w}_{G,H} | \mathcal{X}, \mathcal{N}] \leq \bar{f}_{G,H}(\lambda, \mu, \gamma).$$

3.3.2. Синтез матриц контрастов

Продолжаем анализ оценки $\widehat{w}_{G,H}$ при имеющемся наблюдении $\omega = Ax_* + N\nu_* + \xi$ с $\xi \in \Xi_\epsilon(G, H)$, см. (37). Согласно (41) для $z = \widehat{\nu} - \nu_*$ и всех $g \in \mathbf{R}^m$ имеем

$$|g^T N z| \leq \min \left\{ \|N^T g\|_2 \text{Opt}_2, \sqrt{2s} \|N^T g\|_2 \text{Opt}_\infty, 2s \|N^T g\|_\infty \text{Opt}_\infty \right\}.$$

Отсюда следует (ср. (42)), что для всех $i \leq 2m$ справедливо неравенство

$$(44) \quad |g_i^T A \Delta| \leq 2\overline{\alpha}(\epsilon) \|g_i\|_2 + \\ + \min \left\{ \|N^T g_i\|_2 \text{Opt}_2, \sqrt{2s} \|N^T g_i\|_2 \text{Opt}_\infty, 2s \|N^T g_i\|_\infty \text{Opt}_\infty \right\}.$$

Заметим, что правая часть в (44) невыпукла по g , поэтому описанные выше методы синтеза матриц контрастов не применимы. Чтобы преодолеть эту трудность, используем следующее важное свойство полиэдральных оценок:

“агрегирование” оценок позволяет получить новую оценку с риском, незначительно отличающимся от наименьшего из рисков всех агрегируемых оценок.

В рассматриваемой постановке данный подход работает следующим образом. Разбиваем контраст G размерности $m \times 2m$ на два блока $G_\chi = [g_{\chi,1}, \dots, g_{\chi,m}]$, $\chi = 1, 2$, размерности $m \times m$ и строим блоки, используя соответствующие неравенства, унаследованные от (44), а именно:

$$\begin{aligned} |g_{1,i}^T A \Delta| &\leq 2\bar{\varkappa}(\epsilon) \|g_{1,i}\|_2 + 2s \|N^T g_{1,i}\|_\infty \text{Opt}_\infty, \\ |g_{2,i}^T A \Delta| &\leq 2\bar{\varkappa}(\epsilon) \|g_{2,i}\|_2 + \|N^T g_{2,i}\|_2 \underbrace{\min\{\text{Opt}_2, \sqrt{2s} \text{Opt}_\infty\}}_{=: \varrho_{2,H}}, \end{aligned}$$

где $\Delta = \hat{x} - x_*$. Ослабим эти неравенства до

$$\begin{aligned} |g_{1,i}^T A \Delta|^2 &\leq \pi_1^2(g_{1,i}), \quad \pi_1(g) = \sqrt{8\bar{\varkappa}^2(\epsilon) \|g\|_2^2 + 8s^2 \text{Opt}_\infty^2 \|N^T g\|_\infty^2}, \\ |g_{2,i}^T A \Delta|^2 &\leq \pi_2^2(g_{2,i}), \quad \pi_2(g) = \sqrt{8\bar{\varkappa}^2(\epsilon) \|g\|_2^2 + 2\varrho_{2,H}^2 \|N^T g\|_2^2}. \end{aligned}$$

Заметим, что нормы π_χ , $\chi = 1, 2$, являются эллитопическими; поэтому в текущей ситуации можно использовать процедуру синтеза из раздела 3.2, придя к аналогу теоремы 2. Для этого обозначим через n_1, \dots, n_n столбцы N и зададим

$$\bar{M}_j = 8\bar{\varkappa}^2(\epsilon) I_m + 8s^2 \text{Opt}_\infty^2 n_j n_j^T, \quad j \leq m,$$

и

$$\bar{Q} = (8\bar{\varkappa}^2(\epsilon) I_m + 2\varrho_{2,H}^2 N N^T)^{-1/2}.$$

Далее заметим, что единичный шар $\pi_1(\cdot)$ является эллитопом

$$\bar{M} = \{w \in \mathbf{R}^m : \exists r \in [0, 1]^M : w^T \bar{M}_j w \leq \rho_j, j = 1, \dots, M\},$$

а единичный шар π_2 – эллипсоидом $w^T \bar{Q}^{-2} w \leq 1$. Теперь рассмотрим задачу оптимизации

$$(45a) \quad \text{Opt} = \min_{\lambda, \mu, \tau, \Theta_1, \Theta_2, \rho} \left\{ \bar{f}_H(\lambda, \mu, \tau, \Theta_1, \Theta_2, \rho) : \lambda \geq 0, \mu \geq 0, \tau \geq 0, \right. \\ \left. \begin{aligned} &\Theta_1 \succeq 0, \Theta_2 \succeq 0, \text{Tr}(\bar{M}_j \Theta_1) \leq \rho, j = 1, \dots, n, \\ &\left[\frac{\sum_\ell \lambda_\ell S_\ell}{\frac{1}{2} B^T \sum_k \mu_k T_k + A^T \left[\sum_i \tau_i h_i h_i^T \right] A + A^T (\Theta_1 + Q \Theta_2 Q^T) A} \middle| \frac{1}{2} B \right] \succeq 0 \end{aligned} \right\},$$

где

$$(45b) \quad \begin{aligned} &\bar{f}_H(\lambda, \mu, \tau, \Theta_1, \Theta_2, \rho) = \\ &= \phi_S(\lambda) + 4\phi_T(\mu) + \sum_i \tau_i \psi_H^2(h_i) + \text{Tr}(\Theta_2) + 2\sqrt{2} \ln(4m^2 n) \rho. \end{aligned}$$

Обращаем внимание, что в силу ограничений на Θ_1 и ρ в этой задаче, (Θ_1, ρ) принадлежит конусу \mathbf{K} , связанному с эллитопом \bar{M} в роли \mathcal{W} (Предложение 2).

Теорема 3. При заданном допустимом решении $(\lambda, \mu, \tau, \Theta_1, \Theta_2, \rho)$ задачи (45) построим матрицы контрастов G_1 и G_2 размерности $t \times t$ следующим образом.

- Для построения G_1 применим вторую часть Предложения 2 к $(\Theta_1, \rho, \overline{\mathcal{M}})$ в роли $(\Theta, \rho, \mathcal{W})$, чтобы эффективно вычислить разложение $\Theta_1 = \sum_{i=1}^m \gamma_i g_{1,i} g_{1,i}^T$ с $g_{1,i} \in \overline{\mathcal{M}}$ и $\gamma_i \geq 0$, $\sum_i \gamma_i \leq 2\sqrt{2} \ln(4m^2 n) \rho$. Зададим $G_1 = [g_{1,1}, \dots, g_{1,m}]$.
- Для построения G_2 разложим Θ_2 по собственным значениям, т.е. для $\Theta_2 = \Gamma \text{Diag}\{\chi\} \Gamma^T$ зададим $G_2 = [g_{2,1}, \dots, g_{2,m}] = Q\Gamma$.

Заметим, что $\Theta_1 + Q\Theta_2Q = \sum_i \gamma_i g_{1,i} g_{1,i}^T + \sum_i \chi_i g_{2,i} g_{2,i}^T$.

Для полученной полиэдральной оценки $\widehat{w}_{G,H}$ и всех $x_* \in \mathcal{X}$, s -разреженных ν_* и $\xi \in \Xi_\epsilon(G, H)$ выполняется

$$(46) \quad \|\widehat{w}_{G,H}(Ax_* + N\nu_* + \xi) - Bx_*\| \leq \overline{f}_H(\lambda, \mu, \tau, \Theta_1, \Theta_2, \rho).$$

Следовательно, ϵ -риск оценки ограничен сверху величиной $f_H(\lambda, \mu, \tau, \Theta_1, \Theta_2, \rho)$ (в силу $\xi \in \Xi_\epsilon(G, H)$ с вероятностью не менее $1 - \epsilon$).

Доказательство аналогично доказательству теоремы 2 и опускается.

3.4. Итоговый результат

Наконец, можно “собрать” полиэдральные оценки из разделов 3.2 и 3.3 в следующей конструкции (ср. [2, раздел 5.1.6]):

Положим

$$\overline{\varkappa}(\epsilon) = \sigma(2 \ln[(2n + 2M + 8m)/\epsilon])^{1/2},$$

и пусть $\widetilde{H} = [\widetilde{h}_1, \dots, \widetilde{h}_n] \in \mathbf{R}^{m \times n}$, $\widetilde{G} = [\widetilde{g}_1, \dots, \widetilde{g}_{2m}] \in \mathbf{R}^{m \times 2m}$ и $\overline{H} = [\overline{h}_1, \dots, \overline{h}_M] \in \mathbf{R}^{m \times M}$, $\overline{G} = [\overline{g}_1, \dots, \overline{g}_{2m}] \in \mathbf{R}^{m \times 2m}$ – матрицы контрастов, полученные с помощью процедуры синтеза из разделов 3.2 и 3.3 соответственно. Определим агрегированную оценку \widehat{w} для w_* как $\widehat{w}(\omega) = B\widehat{x}(\omega)$, где $\widehat{x}(\omega)$ – x -компонента вектора

$$(\widehat{x}(\omega), \widehat{\nu}(\omega)) \in \underset{x, \nu}{\text{Argmin}} \left\{ \|\nu\|_1 : x \in \mathcal{X}, \right. \\ \left. \begin{array}{l} \widetilde{h}_k^T [N\nu + Ax - \omega] \leq \overline{\varkappa}(\epsilon) \|\widetilde{h}_k\|_2, \quad k = 1, \dots, n, \\ \overline{h}_k^T [N\nu + Ax - \omega] \leq \overline{\varkappa}(\epsilon) \|\overline{h}_k\|_2, \quad k = 1, \dots, M, \\ \widetilde{g}_i^T [N\nu + Ax - \omega]_\infty \leq \overline{\varkappa}(\epsilon) \|\widetilde{g}_i\|_2, \quad i = 1, \dots, 2m, \\ \overline{g}_i^T [N\nu + Ax - \omega]_\infty \leq \overline{\varkappa}(\epsilon) \|\overline{g}_i\|_2, \quad i = 1, \dots, 2m \end{array} \right\},$$

когда задача разрешима, и $\widehat{x}(\omega) = 0$ в противном случае.

Введем обозначение $G = [\widetilde{G}, \overline{G}] \in \mathbf{R}^{m \times 4m}$; кроме того, пусть $(\widetilde{\lambda}, \widetilde{\mu}, \widetilde{\gamma})$ – допустимое решение задачи (27) при $H = \widetilde{H}$, а $(\overline{\lambda}, \overline{\mu}, \overline{\gamma})$ – допустимое решение задачи (43) при $H = \overline{H}$. Пусть $f_{G,H}$ и $\overline{f}_{G,H}$ заданы формулами (27) и (43) соответственно. Из Предложений 3 и 4 непосредственно следует, что для каждого

s -разреженного ν_* и каждого $x_* \in \mathcal{X}$ справедлива верхняя граница риска

$$(47) \quad \text{Risk}_\epsilon[\widehat{w}(\cdot)|\mathcal{X}, \mathcal{N}] \leq \min[f_{G, \widetilde{H}}(\widetilde{\lambda}, \widetilde{\mu}, \widetilde{\gamma}), \bar{f}_{G, \overline{H}}(\overline{\lambda}, \overline{\mu}, \overline{\gamma})].$$

Заметим, что полученную оценку можно эффективно оптимизировать по всем параметрам, кроме \overline{H} , задав:

- \widetilde{H} как (почти) минимум для $\rho[H]$ по $H \in \mathcal{H}$ (23);
- \widetilde{G} как результат разложения (Θ_1, Θ_2) -компоненты (почти) оптимального решения задачи (33a), (33b) (см. теорему 2), связанной с \widetilde{H} ;
- \overline{G} как результат разложения (Θ_1, Θ_2) -компоненты (почти) оптимального решения задачи (45) (см. теорему 3), связанной с \overline{H} .

3.5. Иллюстративный пример

В данном эксперименте для подтверждения предложенной концепции сравним три оценки x_* : 1) оценку \widehat{x}_{HG} с матрицей контрастов $[H, G]$, вычисленной с помощью процедуры из раздела 3.2, 2) оценку \widehat{x}_{IG} с контрастом $[\overline{H}, \overline{G}] = [I_m, \overline{G}]$ с \overline{G} , полученной с помощью процедуры синтеза из раздела 3.3.2, и 3) “агрегированную” оценку \widehat{x}_{HIG} с комбинированным контрастом $[H, I_m, G, \overline{G}]$. Решаем задачи оптимизации (28) и (33a), (33b) для вычисления контрастов H и G оценки \widehat{x}_{HG} , а также решаем (45) для построения контраста \overline{G} оценки \widehat{x}_{IG} . Например, при вычислении контраста \overline{G} зададим $\overline{\alpha}(\epsilon) = \sqrt{2}\sigma \text{erfcinv}(\frac{\epsilon}{2n})$, где $\text{erfcinv}(\cdot)$ – обратная дополнительная гауссова функция ошибок. При численном решении задачи (45) Θ_1 была задана 0; полученная задача может быть записана в виде

$$(48) \quad \text{Opt} = \min_{\lambda, \mu, \gamma, \Theta} \left\{ \lambda + 4 \sum_{k=1}^p \mu_k + \sum_{j=1}^n \gamma_j + \text{Tr}(\Theta) : \lambda \geq 0, \mu \geq 0, \gamma \geq 0, \Theta \succeq 0, \right. \\ \left. \left[\frac{\lambda I_p}{\frac{1}{2} I_p} \middle| \frac{\frac{1}{2} I_p}{\overline{A}^T \Theta \overline{A} + P^T \text{Diag}\{\mu\} P + \overline{A}^T \text{Diag}\{\gamma\} \overline{A}} \right] \succeq 0 \right\},$$

где $\overline{A} = A/\rho_2$ с $\rho_2 = 2\overline{\alpha}(\epsilon) + \varrho_{2, \overline{H}}$, последующие элементы в Pz равны $z_1, [z_2 - z_1]/h, [z_{i-2} - 2z_{i-1} + z_i]/h^2, 3 \leq i \leq p, \text{ с } h = 2\pi/p$. Соответствующие границы риска получены путем вычисления решений (43). Задачи оптимизации решены с помощью коммерческого решателя Mosek [41] с использованием CVX [26].¹⁰

В иллюстративном примере:

- $m = n = 256, q = p = 32, N = I_n, B = I_p$, а A – случайная матрица размерности $n \times p$ с гауссовыми элементами, удовлетворяющая соотношению $A^T A = I_p$;

¹⁰ Код MATLAB для этого эксперимента доступен в репозитории GitHub по адресу: <https://github.com/ai1-fr/poly-robust>.

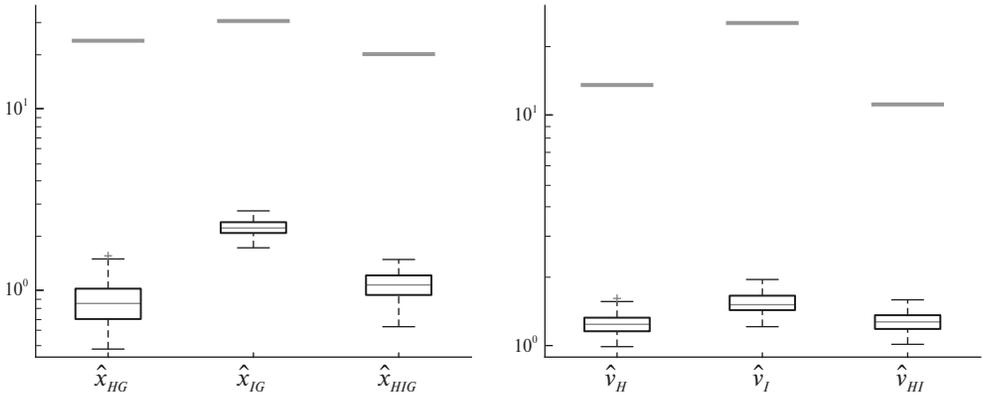


Рис. 1. Левый график: распределения $\|\cdot\|_2$ -ошибок восстановления x_* и теоретические верхние границы для $\text{Risk}_{0,05}$ (жирные горизонтальные линии); правый график: распределения $\|\cdot\|_2$ -ошибок восстановления ν_* и теоретические верхние границы для $\text{Risk}_{0,05}$.

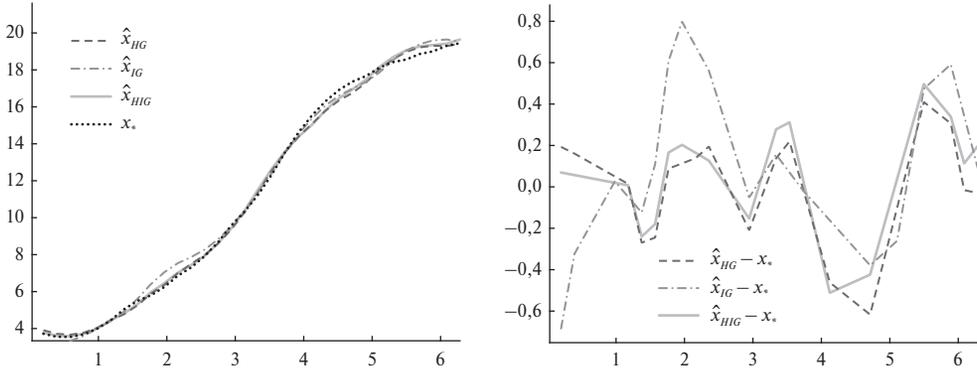


Рис. 2. Типичная реализация сигнала, оценок и ошибок восстановления.

- \mathcal{X} – ограничение на равномерную p -точечную сетку на отрезке $\Delta = [0, 2\pi]$ функций f , удовлетворяющих $|f(0)| \leq 4$, $|f'(0)| \leq 1$ и $|f''(t)| \leq 4$, $t \in \Delta$;
- ошибка восстановления определяется стандартной евклидовой нормой $\|\cdot\|$ на \mathbf{R}^p ;
- $\xi \sim \mathcal{N}(0, \sigma^2 I_m)$ с $\sigma = 0,1$, $\epsilon = 0,05$ и $s = 8$.

На рис. 1 показаны результаты вычислений. В каждом эксперименте построены $n_S = 100$ восстановлений \hat{x}_{HG} , \hat{x}_{IG} и \hat{x}_{HIG} случайно выбранных сигналов $x_* \in \mathcal{X}$ со сгенерированными случайным образом разреженными помехами ν_* . Результаты представлены на левом графике рис. 1. На правом графике даны ошибки восстановления помех ν_* вместе с верхней границей риска Opt_2 (40). На рис. 2 показана типовая реализация сигнала и ошибок восстановления; значения $\|\cdot\|_2$ -ошибок восстановления составляют $\|\hat{x}_{HG} - x_*\|_2 = 1,48\dots$, $\|\hat{x}_{IG} - x_*\|_2 = 2,02\dots$ и $\|\hat{x}_{HIG} - x_*\|_2 = 1,43\dots$, а соответствующая норма сигнала равна $\|x_*\|_2 = 72,2\dots$.

4. Заключение

В работе проанализированы различные постановки задачи восстановления сигналов по непрямым наблюдениям со случайными и адверсариальными шумами. Рассмотрены процедуры оценки для работы с двумя типами загрязнения – детерминированным возмущением, которое принадлежит заданному (известному) выпуклому множеству, и разреженным (сингулярным) возмущением. Важной особенностью описанных выше *полиэдральных оценок* является их “операционная природа”: они дают “вполне хорошие” оценки сигнала при очень общих предположениях относительно параметров задачи – линейных отображениях и задействованных классах сигналов. При этом полиэдральные оценки не задаются аналитическими выражениями и не имеют аналитических формул эффективности. Напротив, соответствующие границы риска и сами оценки доступны численно, как эффективно вычисляемые решения задач выпуклой оптимизации.

ПРИЛОЖЕНИЕ. Граница ошибки для ℓ_1 -восстановления

Условие $\mathbf{Q}_\infty(s, \kappa)$

При заданной наблюдательной матрице N размерности $m \times n$, положительном целом числе $s \leq n$ и $\kappa \in (0, 1/2)$ будем говорить, что матрица H размерности $m \times r$ удовлетворяет условию $\mathbf{Q}_\infty(s, \kappa)$, если

$$(П.1) \quad \|w\|_\infty \leq \|H^T N w\|_\infty + \frac{\kappa}{s} \|w\|_1 \quad \forall w \in \mathbf{R}^n.$$

Обозначим через y^s вектор, полученный из вектора $y \in \mathbf{R}^n$ обнулением всех элементов, кроме s наибольших по абсолютной величине.

Предложение 5. При заданных N и целом числе $s > 0$ предположим, что H удовлетворяет условию $\mathbf{Q}_\infty(s, \kappa)$ с $\kappa < \frac{1}{2}$. Тогда для всех $\nu, \hat{\nu} \in \mathbf{R}^n$ таких, что $\|\hat{\nu}\|_1 \leq \|\nu\|_1$, справедливо неравенство

$$(П.2) \quad \|\hat{\nu} - \nu\|_q \leq \frac{(2s)^{\frac{1}{q}}}{1 - 2\kappa} \left[\|H^T N [\hat{\nu} - \nu]\|_\infty + \frac{\|\nu - \nu^s\|_1}{s} \right], \quad 1 \leq q \leq \infty.$$

Доказательство. Введем обозначение $\rho = \|H^T N [\hat{\nu} - \nu]\|_\infty$ и положим $z = \hat{\nu} - \nu$.

1°. Пусть $I \subset \{1, \dots, n\}$ с не более чем s элементами, и пусть \bar{I} – его дополнение в $\{1, \dots, n\}$. Обозначим через x_I вектор, полученный из вектора x обнулением всех элементов с индексами, не принадлежащими I . Тогда

$$\|\hat{\nu}_{\bar{I}}\|_1 = \|\hat{\nu}\|_1 - \|\hat{\nu}_I\|_1 \leq \|\nu\|_1 - \|\hat{\nu}_I\|_1 = \|\nu_I\|_1 + \|\nu_{\bar{I}}\|_1 - \|\hat{\nu}_I\|_1 \leq \|z_I\|_1 + \|\nu_{\bar{I}}\|_1,$$

и поэтому

$$\|z_{\bar{I}}\|_1 \leq \|\hat{\nu}_{\bar{I}}\|_1 + \|\nu_{\bar{I}}\|_1 \leq \|z_I\|_1 + 2\|\nu_{\bar{I}}\|_1.$$

Отсюда следует, что

$$(II.3) \quad \|z\|_1 = \|z_I\|_1 + \|z_{\bar{I}}\|_1 \leq 2\|z_I\|_1 + 2\|\nu_{\bar{I}}\|_1.$$

Кроме того, по определению ρ имеем

$$(II.4) \quad \|H^T N z\|_\infty \leq \rho.$$

2°. Поскольку H удовлетворяет $\mathbf{Q}_\infty(s, \kappa)$, получаем

$$\|z\|_{s,1} \leq s\|H^T N z\|_\infty + \kappa\|z\|_1,$$

где $\|z\|_{s,1}$ — ℓ_1 -норма s -мерного вектора, состоящего из s наибольших по абсолютной величине элементов z . В силу (II.4) отсюда следует, что $\|z\|_{s,1} \leq s\rho + \kappa\|z\|_1$. Вместе с очевидным неравенством $\|z_I\| \leq \|z\|_{s,1}$ (напомним, что $\text{Card}(I) = s$) и (II.3) это дает

$$\|z\|_1 \leq 2\|z_I\|_1 + 2\|\nu_{\bar{I}}\|_1 \leq 2s\rho + 2\kappa\|z\|_1 + 2\|\nu_{\bar{I}}\|_1.$$

Поэтому (напомним, что $\kappa \leq \frac{1}{2}$) можно записать

$$(II.5) \quad \|z\|_1 \leq \frac{2s\rho + 2\|\nu_{\bar{I}}\|_1}{1 - 2\kappa}.$$

С другой стороны, поскольку H удовлетворяет условию $\mathbf{Q}_\infty(s, \kappa)$, также имеем

$$\|z\|_\infty \leq \|H^T N z\|_\infty + \frac{\kappa}{s}\|z\|_1.$$

На основании (II.5) и (II.4) отсюда следует, что

$$(II.6) \quad \|z\|_\infty \leq \rho + \frac{\kappa}{s} \frac{2s\rho + 2\|\nu_{\bar{I}}\|_1}{1 - 2\kappa} = (1 - 2\kappa)^{-1} \left[\rho + \frac{\|\nu_{\bar{I}}\|_1}{s} \right].$$

Закключаем, что для всех $1 \leq q \leq \infty$ справедливо неравенство

$$\|z\|_p \leq \|z\|_\infty^{\frac{q-1}{q}} \|z\|_1^{\frac{1}{q}} \leq \frac{(2s)^{\frac{1}{q}}}{1 - 2\kappa} \left[\rho + \frac{\|\nu_{\bar{I}}\|_1}{s} \right]. \quad \square$$

СПИСОК ЛИТЕРАТУРЫ

1. *Juditsky A., Nemirovski A.* On polyhedral estimation of signals via indirect observations // *Electronic Journal of Statistics*. 2020. V. 14. No. 1. P. 458–502.
2. *Juditsky A., Nemirovski A.* *Statistical Inference via Convex Optimization*. Princeton: Princeton University Press, 2020.
3. *Tukey J.W.* A survey of sampling from contaminated distributions / *Contributions to Probability and Statistics*. 1960. P. 448–485.
4. *Huber P.J.* *Robust Statistics*. New York: Wiley, 1981.

5. *Yu C., Yao W.* Robust linear regression: A review and comparison // *Communications in Statistics – Simulation and Computation*. 2017. V. 46. No. 8. P. 6261–6282.
6. *Поляк Б.Т., Цыпкин Я.З.* Адаптивные алгоритмы оценивания (сходимость, оптимальность, стабильность) // *АиТ*. 1979. № 3. С. 71–84.
7. *Polyak B.T., Tsyppkin Ya.Z.* Robust identification // *Automatica*. 1980. V. 16. No. 1. P. 53–63.
8. *Поляк Б.Т., Цыпкин Я.З.* Робастные псевдоградиентные алгоритмы адаптации // *АиТ*. 1980. № 10. С. 91–97.
9. *Polyak B.T., Tsyppkin Ya.Z.* Optimal and robust methods for unconditional optimization // *IFAC Proceedings Volumes*. 1981. V. 14. No. 2. P. 519–523.
10. *Поляк Б.Т., Цыпкин Я.З.* Критериальные алгоритмы стохастической оптимизации // *АиТ*. 1984. № 6. С. 95–104.
11. *Polyak B.T., Tsyppkin Ya.Z.* Optimal recurrent algorithms for identification of non-stationary plants // *Computers & Electrical Engineering*. 1992. V. 18. No. 5. P. 365–371.
12. *Chen Y., Caramanis C., Mannor Sh.* Robust sparse regression under adversarial corruption // *International Conference on Machine Learning*. PMLR, 2013. P. 774–782.
13. *Balakrishnan S., Du Simon S., Li J., Singh A.* Computationally efficient robust sparse estimation in high dimensions // *Conference on Learning Theory*. PMLR, 2017. P. 169–212.
14. *Diakonikolas I., Kong W., Stewart A.* Efficient algorithms and lower bounds for robust linear regression // *Proceedings of the Thirtieth Annual ACM–SIAM Symposium on Discrete Algorithms*. SIAM, 2019. P. 2745–2754.
15. *Liu L., Shen Y., Li T., Caramanis C.* High dimensional robust sparse regression // *International Conference on Artificial Intelligence and Statistics*. PMLR, 2020. P. 411–421.
16. *Minsker S., Ndaoud M., Wang L.* Robust and tuning-free sparse linear regression via square-root slope // *SIAM J. Math. Data Sci.* 2024. V. 6. No. 2. P. 428–453
17. *Foygel R., Mackey L.* Corrupted sensing: Novel guarantees for separating structured signals // *IEEE Transactions on Information Theory*. 2014. V. 60. No. 2. P. 1223–1247.
18. *Dalalyan A., Thompson Ph.* Outlier-robust estimation of a sparse linear model using l_1 -penalized Huber’s M -estimator // *Advances in Neural Information Processing Systems*. 2019. V. 32.
19. *Bruce A.G., Donoho D.L., Gao H.-Y., Martin R.D.* Denoising and robust nonlinear wavelet analysis // *Wavelet Applications*. SPIE, 1994. V. 2242. P. 325–336.
20. *Sardy S., Tseng P., Bruce A.* Robust wavelet denoising // *IEEE Transactions on Signal Processing*. 2001. V. 49. No. 6. P. 1146–1152.
21. *Diakonikolas I., Kane D.M.* Algorithmic High-Dimensional Robust Statistics. Cambridge: Cambridge University Press, 2023.
22. *Juditsky A., Nemirovski A.* Near-optimality of linear recovery from indirect observations // *Mathematical Statistics and Learning*. 2018. V. 1. No. 2. P. 171–225.
23. *Donoho D.L.* Statistical estimation and optimal recovery // *The Annals of Statistics*. 1994. V. 22. No. 1. P. 238–270.
24. *Juditsky A.B., Nemirovski A.S.* Nonparametric estimation by convex programming // *The Annals of Statistics*. 2009. V. 37. No. 5A. P. 2278–2300.

25. *Juditsky A., Nemirovski A.* Near-optimality of linear recovery in Gaussian observation scheme under $\|\cdot\|_2^2$ -loss // *The Annals of Statistics*. 2018. V. 46. No. 4. P. 1603–1629.
26. *Grant M., Boyd S.* *The CVX Users' Guide*. Release 2.1, 2014. <https://web.cvxr.com/cvx/doc/CVX.pdf>
27. *Micchelli C.A., Rivlin T.J.* A Survey of optimal recovery / *Optimal Estimation in Approximation Theory*. Micchelli C.A., Rivlin T.J. (Eds.). Boston, MA: Springer, 1977. P. 1–54.
28. *Micchelli C.A., Rivlin T.J.* Lectures on optimal recovery / *Numerical Analysis Lancaster 1984. Lecture Notes in Mathematics*. Turner P.R. (Ed.). Berlin–Heidelberg: Springer, 1985. V. 1129. P. 21–93.
29. *Черноусько Ф.Л.* Оценивание фазового состояния динамических систем. М.: Наука, 1988.
30. *Fogel E., Huang Y.-F.* On the value of information in system identification-bounded noise case // *Automatica*. 1982. V. 18. No. 2. P. 229–238.
31. *Граничин О.Н., Поляк Б.Т.* Рандомизированные алгоритмы оценивания и оптимизации при почти произвольных помехах. М.: Наука, 2003.
32. *Kurzanski A.B.* Identification – a theory of guaranteed estimates // *From Data to Model*. Willems J.C. (Ed.). Berlin–Heidelberg: Springer, 1989. P. 135–214.
33. *Kurzanski A., Vályi I.* *Ellipsoidal Calculus for Estimation and Control*. Boston, MA: Birkhäuser, 1997.
34. *Milanese M., Vicino A.* Optimal estimation theory for dynamic systems with set membership uncertainty: An overview // *Automatica*. 1991. V. 27. No. 6. P. 997–1009.
35. *Schweppe F.C.* *Uncertain Dynamic Systems*. Englewood Cliffs, NJ: Prentice-Hall, 1973.
36. *Juditsky A., Nemirovski A.* On design of polyhedral estimates in linear inverse problems // *SIAM Journal on Mathematics of Data Science*. 2024. V. 6. No. 1. P. 76–96.
37. *Candes E.J., Tao T.* Decoding by linear programming // *IEEE Transactions on Information Theory*. 2005. V. 51. No. 12. P. 4203–4215.
38. *Bickel P.J., Ritov Ya., Tsybakov A.B.* Simultaneous analysis of Lasso and Dantzig selector // *The Annals of Statistics*. 2009. V. 37. No. 4. P. 1705–1732.
39. *Donoho D.L., Huo X.* Uncertainty principles and ideal atomic decomposition // *IEEE Transactions on Information Theory*. 2001. V. 47. No. 7. P. 2845–2862.
40. *van de Geer S.* *Estimation and Testing under Sparsity*. Cham: Springer, 2016.
41. *Mosek ApS.* *The MOSEK optimization toolbox for MATLAB manual*. Version 8.0, 2015. <http://docs.mosek.com/8.0/toolbox/>

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

Поступила в редакцию 03.03.2025

После доработки 20.05.2025

Принята к публикации 27.06.2025

© 2025 г. В. МИРАСЬЕРРА, д-р философии (vmirasierra@us.es)
(Университет Севильи, Испания),
М. МАММАРЕЛЛА, д-р философии (martina.mammarella@cnr.it),
Ф. ДАББЕНЕ, д-р философии (fabrizio.dabbene@cnr.it)
(Политехнический университет Турина, Италия),
Т. АЛАМО, д-р философии (talamo@us.es)
(Университет Севильи, Испания)

ТОЧНЫЕ АППРОКСИМАЦИИ МНОЖЕСТВ С ВЕРОЯТНОСТНЫМИ ОГРАНИЧЕНИЯМИ С ПОМОЩЬЮ ПАКЕТНОГО ВЕРОЯТНОСТНОГО МАСШТАБИРОВАНИЯ

Вычисление “надежных” в вероятностном смысле областей остается актуальной проблемой в стохастических постановках задач теории систем. Точно вычислить такие области можно для некоторых конкретных задач; однако, как правило, результаты вычислений слишком сложны (например, невыпуклые или несвязные множества) и неприменимы для приложений реального времени. В данной работе представлена основанная на случайной выборке процедура для получения “точных” внутренних аппроксимаций вероятностно-надежной области. Предлагаемый подход не требует каких-либо предположений о распределении вероятностей, а внутренняя аппроксимация может быть найдена в автономном режиме (офлайн). В отличие от аналогичных подходов предлагаемое пакетное вероятностное масштабирование включает в себя усиленное ограничение на точность, позволяющее настраивать степень консерватизма полученной аппроксимации.

Ключевые слова: рандомизированные алгоритмы, вероятностная робастность, неопределенные системы, статистическая теория обучения.

DOI: 10.31857/S0005231025080031, EDN: UTBLUS

1. Введение

Реальные системы часто не являются детерминированными и включают неопределенность, требуя разработки робастных и стохастических стратегий управления. При робастном управлении [1–3] неопределенность предполагается неизвестной, но ограниченной некоторой компактной областью, а регулятор строится исходя из гарантированного выполнения ограничений для всех допустимых значений неопределенности. При стохастическом управлении [4–6], напротив, используются вероятностные соображения и вводится понятие вероятностных ограничений [7]. В отличие от жестких ограничений для вероятностных ограничений допускается нарушение, но при условии, что вероятность их выполнения остается выше заданного порога.

Ослабление ограничений и учет вероятностей делают стохастические схемы менее консервативными по сравнению с их робастными аналогами. Более того, стохастические схемы позволяют работать с неопределенностями с бесконечным носителем. В свою очередь, процедура построения в результате оказывается гораздо более сложной по двум основным причинам. Во-первых, очень трудно проверить, являются ли решения задач с вероятностными ограничениями допустимыми; во-вторых, вероятностные ограничения обычно не являются выпуклыми (например, см. [8, рис. 1] и [9, рис. 1]).

В последнее десятилетие основанные на случайной выборке схемы стали использоваться в качестве эффективного инструмента для решения стохастических задач. Ключевую роль в этой области исследований сыграл, в частности, Б.Т. Поляк, одним из первых признав потенциал рандомизированных методов при решении оптимизационных задач со стохастической неопределенностью; например, см. [10–12]. Эти работы проложили путь к последующим результатам, сочетающим выборку и оптимизацию; одним из примеров может служить сценарный подход, предложенный в [13]. Обзор соответствующих методов можно найти в [14, 15].

Вероятностно-надежная область, также называемая *множеством с вероятностными ограничениями* (*chance constrained set, CCS*), состоит из всех точек, удовлетворяющих вероятностным ограничениям. В общем случае точное вычисление CCS громоздко и требуется, чтобы неопределенность имела некоторое распределение [16, 17]. Кроме того, в силу геометрической сложности CCS может быть непригодным для приложений реального времени [18]. В связи с этим, актуальным является решение проблемы аппроксимации вероятностно-надежных областей с помощью множеств настраиваемой сложности.

Для задачи стохастического управления было предложено несколько релаксаций, опирающихся на вычислительно эффективные аппроксимации множества с вероятностными ограничениями. Эти релаксации могут быть либо основаны на некоторых неравенствах концентрации, например использующих априорные знания о структуре неопределенности [19], либо построены с помощью методов случайной выборки [20, 21].

Настоящая работа развивает результаты исследований [9, 22], где была представлена основанная на случайной выборке методология внутренней аппроксимации множеств CCS, названная *вероятностным масштабированием* (*probabilistic scaling, PS*). В рамках этого подхода сначала находится простое аппроксимирующее множество, которое затем масштабируется для выполнения требуемых вероятностных гарантий. Все эти операции выполняются в режиме *офлайн*, а компромисс между необходимым размером выборки и требованием к точности аппроксимации может быть настроен пользователем.

В данной статье обсуждается и расширяется метод пакетного вероятностного масштабирования (доложенный в предварительной форме на конференции [22]) за счет определения новой меры точности аппроксимирующего мно-

жества. Затем показывается, как построить аппроксимирующее множество для выполнения требуемых вероятностных гарантий при соблюдении указанного ограничения на точность. Таким образом, пользователь получает возможность одновременно настраивать сложность и приемлемость полученного аппроксимирующего множества, достигая компромисса между требуемым размером выборки и вычислительной сложностью задачи аппроксимации, которая проводится в автономном режиме (офлайн).

Статья организована следующим образом. В разделе 2 приводятся постановка задачи аппроксимации множества с вероятностными ограничениями, а также численный пример для сравнения различных подходов. В разделе 3 рассматриваются решения этой задачи с помощью статистической теории обучения: сначала классический вероятностный подход к масштабированию (раздел 3.1), а затем его расширение до пакетной структуры (раздел 3.2). Раздел 4 посвящен пакетному вероятностному масштабированию с “плотным” погружением (*tight immersed pack-based probabilistic scaling, TI-PBPS*), являющемуся основным научным результатом статьи. В разделе 5 проводится сравнительный анализ различных подходов с точки зрения консерватизма.

Обозначения: $\mathbb{N}_{\geq 0}$ – множество натуральных чисел, включающее 0. Символ \oplus соответствует сумме Минковского для множеств. При заданном наборе из N скаляров $\{x_1, \dots, x_N\}$, скаляр $x_{1:N}$ обозначает наименьший из них, $x_{2:N}$ – второй наименьший и т.д. до $x_{N:N}$, который является наибольшим скаляром. При заданном $r \geq 0$ величина $x_{1+r:N}$ соответствует тому, что не более, чем r элементов из набора $\{x_1, \dots, x_N\}$ строго меньше $x_{1+r:N}$. Биномиальное распределение задано в виде

$$B(s; N, \varepsilon) = \sum_{i=0}^s \binom{N}{i} \varepsilon^i (1 - \varepsilon)^{N-i}.$$

2. Аппроксимация множеств с вероятностными ограничениями

Рассмотрим задачу робастности, в которой настройки регулятора и вспомогательные переменные параметризованы с помощью вектора $\theta \in \Theta \subseteq \mathbb{R}^{n_\theta}$ (называемого далее *настраиваемым параметром*). Вектор неопределенности w представляет собой одну из допустимых реализаций случайного вектора с заданным распределением вероятности $\text{Pr}_{\mathcal{W}}$ с (возможно, неограниченным) носителем \mathcal{W} . Тогда типичное “выборочное” ограничение может быть задано как

$$(1) \quad g(\theta, w) \leq 0,$$

где функция $g : \mathbb{R}^{\Theta \times \mathcal{W}} \rightarrow \mathbb{R}$ отражает требование к θ при заданном w . В частности, при робастной постановке ограничение (1) должно выполняться для всех возможных значений w . Очевидно, что ограничения в робастном виде бессмысленны в некоторых ситуациях, например при неограниченном носителе w [23]. В таких случаях естественно допустить, что ограничение (1) может

быть нарушено для некоторых элементов \mathcal{W} . Эта концепция строго формализована в определении вероятностных ограничений.

Определение 1 [множество ε -CCS [9]]. *Рассмотрим вероятностную меру $\text{Pr}_{\mathcal{W}}$ на \mathcal{W} . При заданном уровне нарушения $\varepsilon \in (0; 1)$ определим ε -вероятностное ограничивающее множество ε -CCS следующим образом:*

$$\Omega(\varepsilon) = \{\theta \in \Theta : \text{Pr}_{\mathcal{W}}\{g(\theta, w) > 0\} \leq \varepsilon\}.$$

Недавно было предложено несколько подходов к построению вероятностно гарантированной аппроксимации множества с вероятностными ограничениями. Эти подходы основаны на результатах с использованием выборки (например, см. [21, 24, 25]). При заданном \mathcal{W} рассмотрим выборку $\mathbf{z} = \{w_1, \dots, w_N\}$ из N независимых одинаково распределенных (н.о.р.) случайных элементов из \mathcal{W} . В этом случае будем говорить, что \mathbf{z} принадлежит декартову произведению $\mathcal{W}^N \doteq \mathcal{W} \times \dots \times \mathcal{W}$ (N раз), а \mathbf{z} выбирается в соответствии с вероятностной мерой произведения $\text{Pr}_{\mathcal{W}^N}$. Введем понятие индикаторной функции, которое впоследствии будет использовано для переопределения множества с вероятностными ограничениями.

Определение 2 [индикаторная функция для g]. *При заданных $\theta \in \Theta$ и $w \in \mathcal{W}$ индикаторная функция $I^g : \Theta \times \mathcal{W} \rightarrow \{0, 1\}$ ограничения (1) имеет вид*

$$I^g(\theta, w) \doteq \begin{cases} 0, & \text{если } g(\theta, w) \leq 0, \\ 1 & \text{в противном случае.} \end{cases}$$

В контексте статистической теории обучения можно найти аппроксимации для множества ε -CCS с помощью ограничения на эмпирическое среднее, т.е. величину

$$\frac{1}{N} \sum_{i=1}^N I^g(\theta, w_i).$$

А именно, при $\mathbf{z} = \{w_1, \dots, w_N\} \in \mathcal{W}^N$ и так называемом уровне отбрасывания $r \geq 0$ величина $\rho = \frac{r}{N}$ ограничивает эмпирическое среднее таким образом, что множество

$$(2) \quad \Phi_{\rho_N}(\mathbf{z}) \doteq \left\{ \theta \in \Theta : \frac{1}{N} \sum_{i=1}^N I^g(\theta, w_i) \leq \rho \right\}$$

аппроксимирует множество $\Omega(\varepsilon)$. Обращаем внимание, неравенство $\frac{1}{N} \sum_{i=1}^N I^g(\theta, w_i) \leq \frac{r}{N}$ означает, что ограничение $g(\theta, w_i) \leq 0$ нарушается не более чем r элементами \mathbf{z} .

Замечание 1. При заданном ε множество $\Omega(\varepsilon)$ является фиксированным. С другой стороны, когда множество ε -CCS аппроксимируется с использованием случайной выборки (например, см. [9, 26]), соответствующее множество-аппроксимация имеет случайный характер, поскольку оно построено по случайной выборке $\mathbf{z} \in \mathcal{W}^N$.

Предположим, что индикаторная функция I^g имеет конечную размерность Вапника–Червоненкиса (VC-размерность) [27] и что $\rho < \varepsilon$. Тогда вероятность того, что $\Phi_{\rho_N}(\mathbf{z})$ является внутренней аппроксимацией для $\Omega(\varepsilon)$, т.е.

$$\Pr_{\mathcal{W}^N} \{ \Phi_{\rho_N}(\mathbf{z}) \subseteq \Omega(\varepsilon) \},$$

сходится к единице с ростом размера выборки N . В [28] в явном виде вычислены оценки для размера N выборки, которые гарантируют, что $\Phi_{\rho_N}(\mathbf{z})$ вложено в $\Omega(\varepsilon)$ с заданной доверительной вероятностью $\delta \in (0; 1)$, т.е. $\Pr_{\mathcal{W}^N} \{ \Phi_{\rho_N}(\mathbf{z}) \subseteq \Omega(\varepsilon) \} \geq 1 - \delta$.

Размер выборки линейно растет с VC-размерностью I^g , умноженной на некоторый коэффициент, больший чем $\frac{1}{\varepsilon}$. Однако, как показано в [9], эта аппроксимация может быть очень консервативной. Кроме того, в случае невыпуклой функции g полученная аппроксимация, как правило, является невыпуклым (и часто несвязным) множеством. Это обстоятельство может существенно затруднить практическое применение аппроксимации, делая ее непригодной для решения задач реального времени.

Исходя из этих рассуждений, в [9, 22] авторы предложили идею вероятностного масштабирования. В рамках данного подхода вводится исходное простое аппроксимирующее множество (*simple approximating set, SAS*) $\theta_c \oplus \Omega_0$, которое обладает двумя основными свойствами: (i) оно достаточно хорошо отражает геометрию вероятностного множества $\Omega(\varepsilon)$ и одновременно (ii) является достаточно простым. От исходного множества SAS не требуется каких-либо гарантий вероятностного характера, оно лишь должно передавать геометрию множества ε -CCS.

В [9] показано, как *масштабировать* это множество вокруг его центра θ_c для получения масштабируемого множества SAS

$$\Omega(\gamma) = \theta_c \oplus \gamma\Omega_0.$$

Кроме того, введена основанная на случайной выборке процедура для построения вероятно осмысленной аппроксимации множества ε -CCS. В частности, при заданной геометрии Ω_0 и центре масштабирования θ_c цель вероятностного масштабирования – найти наибольший масштабный коэффициент $\bar{\gamma}$ такой, что

$$(3) \quad \Pr_{\mathcal{W}} \{ \theta_c \oplus \bar{\gamma}\Omega_0 \subseteq \Omega(\varepsilon) \} \geq 1 - \delta.$$

Следовательно, выборочное ограничение

$$(4) \quad \Pr_{\mathcal{W}} \{ g(\theta, w) \leq 0 \} \geq 1 - \varepsilon$$

будет выполнено с вероятностью, не меньшей $1 - \delta$.

Процедура построения такой аппроксимации подробно рассмотрена в [9]; см. краткое формальное описание в разделе 3.1.

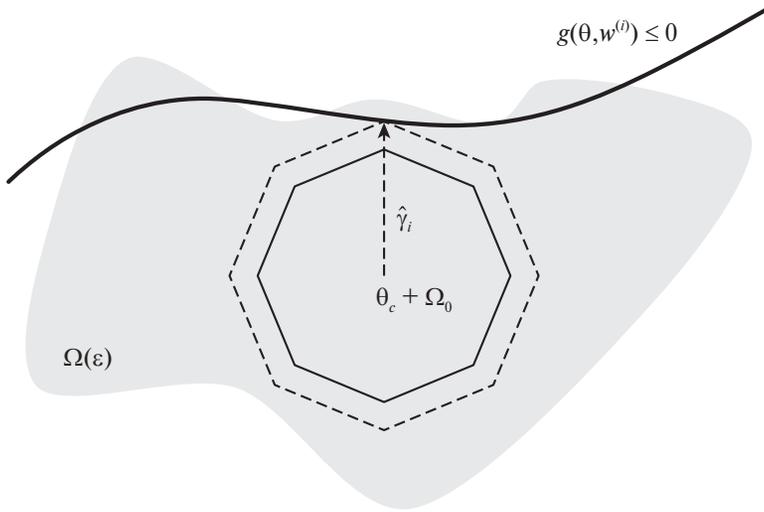


Рис. 1. Схема вероятностного масштабирования.

На рис. 1 приведена простая иллюстрация данного подхода. Затененная область соответствует множеству ε -CCS, в общем случае (как видно) это может быть невыпуклое множество. Тогда необходимо:

- 1) выбрать аппроксимирующее множество–“кандидата” $\theta_c + \Omega_0$ (черный многоугольник);
- 2) сгенерировать выборку $\mathbf{z} = \{w_1, \dots, w_N\} \in \mathcal{W}^N$ для определения оптимального масштабного коэффициента $\bar{\gamma}$;
- 3) для каждой случайной реализации w_i вычислить максимальный масштабный коэффициент γ_i таким образом, чтобы масштабированное множество (пунктирный многоугольник) не нарушало ограничений, соответствующих w_i ;
- 4) выбрать оптимальный масштабный коэффициент в виде $\bar{\gamma} = \gamma_{1+r:N}$, т.е. как r -ое наименьшее значение среди γ_i .

Тогда для $B(r; N, \varepsilon) \leq \delta$ будет справедливо (3).

Несмотря на неоспоримые преимущества *вероятностного масштабирования* (особенно в расширенной версии, где вычислительная сложность дополнительно снижается с помощью описанных выше множеств SAS [9]), масштабированное решение может оказаться консервативным. Эта проблема наглядно продемонстрирована на следующем примере из [22].

Пример 1. Рассмотрим задачу с отдельными вероятностными ограничениями, где каждое ограничение задается касательной к единичному кругу заданного размера в случайной точке с равномерным распределением. В этом случае очевидно, что единичный шар является вероятностно-надежной областью с вероятностью 1, в то время как множество ε -CCS всегда является немного большей масштабированной версией единичного шара по мере уве-

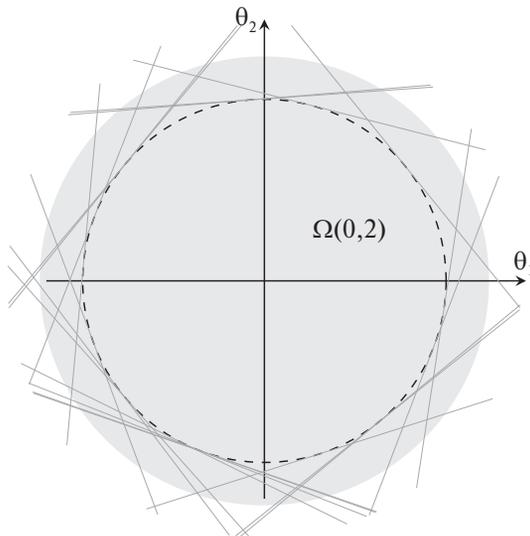


Рис. 2. Затененный круг – $\Omega(0,2)$, пунктирный круг – SAS (единичный круг), отрезки касательных – выборочные ограничения.

личения ε . В частности, легко показать, что точный радиус, соответствующий области с вероятностными ограничениями $\Omega(\varepsilon)$, может быть найден с помощью некоторых трансцендентных функций. На рис. 2 дана иллюстрация этого примера в \mathbb{R}^2 : пунктирная линия соответствует единичному кругу в \mathbb{R}^2 , а внешний затененный круг – множеству с вероятностными ограничениями $\Omega(\varepsilon)$ для конкретного значения $\varepsilon = 0,20$.

Предположим, что необходимо аппроксимировать множество ε -CCS с помощью эмпирического среднего $\Phi_{\rho_N}(\mathbf{z})$ (2). Для этого генерируем N случайных линейных ограничений в виде касательных к единичному шару в точках, равномерно распределенных на его поверхности, и строим аппроксимацию как их пересечение (возможно, отбрасывая самые “плохие” из них). Ясно, что такая аппроксимация не сможет охватить затененный круг.

Кроме того, используя вероятностный подход к масштабированию, выберем единичный шар в качестве начальной аппроксимации $\theta_c \oplus \Omega_0$ множества с вероятностными ограничениями $\Omega(\varepsilon)$. Затем, применяя описанную выше процедуру, будем масштабировать эту начальную геометрию вокруг ее центра (начала координат), чтобы получить внутреннюю аппроксимацию $\Omega(\varepsilon)$ с заданной доверительной вероятностью $\delta \in (0; 1)$. Очевидно, однако, что схема масштабирования всегда дает единичную гиперсферу, поскольку каждое выборочное ограничение является касательной к ней – значит, все масштабные коэффициенты будут равны единице. Следовательно, радиус истинного множества $\Omega(\varepsilon)$ не может быть найден с помощью простых процедур, основанных на случайной выборке. Обращаем внимание, что этот радиус может быть значительно больше единицы, особенно при увеличении n_θ (рис. 3).

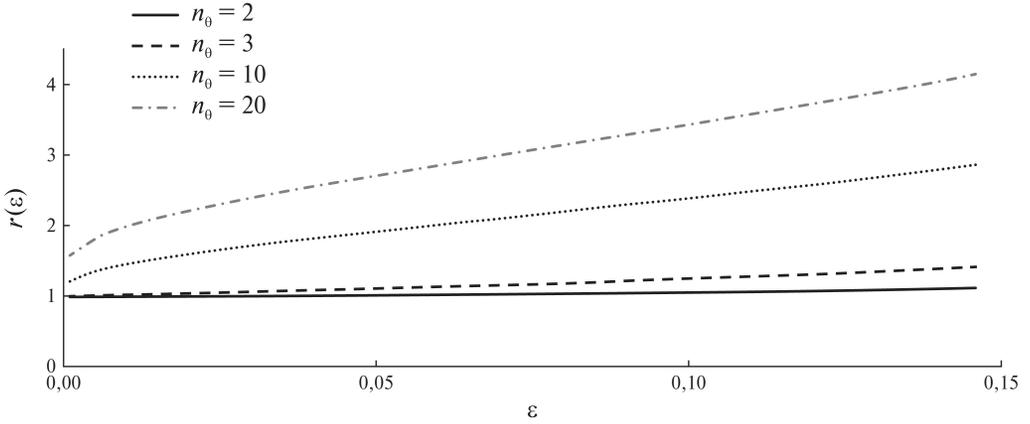


Рис. 3. Радиус r множества $\Omega(\varepsilon)$ как функция ε для различных размерностей задачи n_θ .

С другой стороны, в текущем примере большие масштабные коэффициенты можно получить масштабированием единичной окружности, рассматривая только те области, в которых нарушено более заданного количества ограничений. В данной работе прибегаем к *пакетной стратегии*, успешно применяемой в контексте статистической теории обучения [28] и в выпуклом случае [29], чтобы получить менее консервативные размеры выборки и гарантировать включение масштабированного множества во множество с вероятностными ограничениями с заданной доверительной вероятностью. В частности, цель авторов – расширить пакетную стратегию, впервые предложенную в [22], для получения выборочных аппроксимаций $\Omega(\varepsilon)$ с настраиваемой сложностью, не требующих каких-либо априорных знаний о задаче, например симметрии. Способность предлагаемого подхода к уменьшению степени консерватизма будет позже продемонстрирована на примере 1.

3. Предварительные сведения

В этом разделе напомним некоторые сведения о пакетной стратегии, необходимые для описания основных результатов данной работы. В следующем разделе представим *пакетное вероятностное масштабирование* (*pack-based probabilistic scaling, PBPS*) [22] и расширим этот подход для дальнейшего уменьшения степени консерватизма аппроксимирующего множества. Начнем с понятия *пакета выборок*.

Определение 3 [пакет из L выборок]. При заданном целом числе L набор $\mathbf{z} = \{w_1, \dots, w_L\} \in \mathcal{W}^L$ называется *пакетом размера L* .

Расширим понятие индикаторной функции (определение 2) на пакетный случай.

Определение 4 [пакетная индикаторная функция]. При заданных целых числах s и L , $0 \leq s < L$, и пакете $\mathbf{z} \in \mathcal{W}^L$ размера L *пакетная индикаторная*

торная функция $I_s^g : \Theta \times \mathcal{W}^L \rightarrow \{0, 1\}$ имеет вид

$$(5) \quad I_s^g(\theta, \mathbf{z}) \doteq \begin{cases} 0, & \text{если } \sum_{\ell=1}^L I^g(\theta, w_\ell) \leq s, \\ 1 & \text{в противном случае,} \end{cases}$$

где $I_s^g(\theta, \mathbf{z})$ показывает, нарушает ли точка θ более чем s ограничений, связанных с реализациями неопределенности пакета \mathbf{z} .

Определение 5 [пакетная вероятностно-надежная область]. *Пакетная вероятностно-надежная область $\Phi_s(\mathbf{z})$ – множество точек, которые нарушают не более s ограничений, связанных с реализацией неопределенности \mathbf{z} :*

$$\Phi_s(\mathbf{z}) \doteq \{\theta \in \Theta \mid I_s^g(\theta, \mathbf{z}) = 0\}.$$

В следующем разделе обобщим результаты по вероятностному масштабированию применительно к пакетной стратегии. В частности, опишем процедуру масштабирования множества $\theta_c \oplus \Omega_0$ вокруг его центра θ_c , гарантирующую включение масштабированного (основанного на пакетах) множества в $\Omega(\varepsilon)$ с доверительной вероятностью $\delta \in (0; 1)$.

3.1. Обобщенное вероятностное масштабирование

Сначала введем определение масштабного коэффициента в рамках пакетного подхода.

Определение 6 [коэффициент масштабирования пакета]. *При заданном масштабируемом множестве $SAS \Omega(\gamma)$, определяемом центром масштабирования $\theta_c \in \Theta$ и геометрией Ω_0 , и заданном пакете $\mathbf{z} \in \mathcal{W}^L$ масштабный коэффициент пакета для $\Omega(\gamma)$, соответствующий реализации вероятностных ограничений $g(\theta, w_i) \leq 0, \forall w_i \in \mathbf{z}$, имеет вид*

$$(6) \quad \gamma^s(\theta_c, \Omega_0, \mathbf{z}) \doteq \begin{cases} 0, & \text{если } \theta_c \notin \Phi_s(\mathbf{z}), \\ \max_{\theta_c \oplus \gamma \Omega_0 \subseteq \Phi_s(\mathbf{z})} \gamma & \text{в противном случае.} \end{cases}$$

Теперь формализуем обобщенную задачу вероятностного масштабирования, рассматривая M н.о.р. пакетов \mathbf{z}_i размера L каждый. Заметим, что вероятностное масштабирование [9] является ее частным случаем при $M = N$ и $L = 1$ (N пакетов размера 1).

Свойство 1 [обобщенное вероятностное масштабирование]. *При заданных параметре точности $\varepsilon \in (0; 1)$ и доверительной вероятности $\delta \in (0; 1)$ рассмотрим целочисленный уровень отбрасывания $r \geq 0$ и предположим, что M выбрано исходя из условия*

$$(7) \quad B(r; M, \varepsilon) \leq \delta.$$

Сгенерируем M н.о.р. пакетов $\mathbf{z}_i \in \mathcal{W}^L$, $i = 1, \dots, M$, размера L . Для каждого пакета \mathbf{z}_i вычислим соответствующий ему масштабный коэффициент γ_i как

$$\gamma_i \doteq \gamma^s(\theta_c, \Omega_0, \mathbf{z}_i)$$

(формула (6)) и определим $\bar{\gamma} \doteq \gamma_{1+r:M} > 0$. Тогда с вероятностью, не меньшей $1 - \delta$, справедливо

$$\Pr_{\mathcal{W}^L} \{ \theta_c \oplus \bar{\gamma} \Omega_0 \not\subseteq \Phi_s(\mathbf{z}) \} \leq \varepsilon.$$

Доказательство. Это свойство можно получить из результатов выпуклого сценарного подхода [21, 26] в случае скалярного θ . Другой способ – использовать свойства обобщенной функции \max [30, свойство 3]. Рассмотрим задачу оптимизации

$$(8) \quad \max_{\gamma} \quad \gamma$$

при ограничениях $\theta_c \oplus \gamma \Omega_0 \subseteq \Phi_s(\mathbf{z}_i)$, $i = 1, \dots, M$.

При наличии допустимого решения, используя определение $\gamma^s(\cdot)$, эту задачу можно записать в виде

$$(9) \quad \max_{\gamma} \quad \gamma$$

при ограничениях $\gamma \leq \gamma^s(\theta_c, \Omega_0, \mathbf{z}_i)$, $i = 1, \dots, M$.

Согласно [21, 26] при отбрасывании не более r ограничений в выпуклой задаче с M вероятностными ограничениями вероятность нарушения ограничений решением, полученным из случайной выпуклой задачи, не превышает ε с вероятностью, не меньшей $1 - \delta$, где

$$\delta = \binom{d+r-1}{d-1} B(d+r-1; M, \varepsilon),$$

а d – количество переменных. Сначала заметим, что задача (9) является выпуклой и имеет единственную скалярную переменную γ , т.е. $d = 1$. Также легко проверяются предположения, необходимые для применения результатов [21, 26]. В частности, невырожденность следует из скалярности задачи, а единственность может быть гарантирована введением правила тай-брейка (tie-break). Значит, при допущении r нарушений в вышеприведенной задаче минимизации с вероятностью, не меньшей $1 - \delta$, $\delta = B(r; M, \varepsilon)$, оптимальное решение $\bar{\gamma}$ задачи (9) с не более чем r отброшенными ограничениями удовлетворяет неравенству

$$\Pr_{\mathcal{W}^L} \{ \bar{\gamma} > \gamma^s(\theta_c, \Omega_0, \mathbf{z}) \} \leq \varepsilon.$$

Следовательно, с вероятностью, не меньшей $1 - \delta$, справедливо

$$\Pr_{\mathcal{W}^L} \{ \theta_c \oplus \bar{\gamma} \Omega_0 \not\subseteq \Phi_s(\mathbf{z}) \} \leq \varepsilon.$$

Заметим, что задача (9) с отбрасыванием ограничений может быть непосредственно решена упорядочиванием значений $\gamma_i = \gamma^s(\theta_c, \Omega_0, \mathbf{z}_i)$. Очевидно, если допускается $r \geq 0$ нарушений, оптимальное значение γ будет равно $\bar{\gamma} = \gamma_{1+r:N}$. Меньшие значения γ будут удовлетворять ограничению на включение, но не будут оптимальными, а большие значения γ уже не будут удовлетворять ограничению на включение. \square

Как уже говорилось, результат [22, свойство 1] может быть получен из свойства 2, если задать $M = N$ и $L = 1$. См. следствие ниже.

Следствие 1 [классическое вероятностное масштабирование]. *Предположим, что N выбрано исходя из условия*

$$B(r; N, \varepsilon) \leq \delta$$

и $\mathbf{z} \in \mathcal{W}^N$. Для каждого ограничения $i = 1, \dots, N$ определим $\gamma_i = \gamma^r(\theta_c, \Omega_0, \mathbf{z}_i)$ и допустим, что $\bar{\gamma} = \gamma_{1+r:N} > 0$. Тогда с вероятностью, не меньшей $1 - \delta$, справедливо

$$\Pr_{\mathcal{W}}\{\theta_c \oplus \bar{\gamma}\Omega_0 \not\subseteq \Omega(\varepsilon)\} \leq \varepsilon.$$

Доказательство простое и следует непосредственно из определения 1. Данное следствие показывает, что вероятностный подход к масштабированию [9] можно рассматривать как частный случай более общей пакетной схемы.

Найти аппроксимации множества ε -CCS с помощью классического вероятностного масштабирования в целом легко; при этом не требуется никаких предположений об используемых вероятностях (таких как конечная VC-размерность), обеспечиваются вероятностные гарантии для масштабируемой области, а эффективность метода доказана в [9]. Несмотря на все свои преимущества, классическое вероятностное масштабирование все же может приводить к очень консервативным решениям, см. пример 1. В этом случае, если $\gamma_i = 1$ для всех $i = 1, \dots, N$ и все ограничения учитываются независимо, отбрасывание некоторых из них никак не повлияет на масштабированное аппроксимирующее множество.

В следующем разделе опишем пакетное вероятностное масштабирование, впервые предложенное в [22]. Для тех же исходных множеств SAS этот вариант классического вероятностного масштабирования, применяемый в рамках пакетной стратегии, может приводить к менее консервативным результатам за счет (потенциально) более высоких вычислительных затрат.

Важно подчеркнуть, что уровень отбрасывания r задается пользователем исходя из того, что большие значения r делают результирующее множество менее чувствительным к экстремальным значениям за счет большего размера выборки N . С другой стороны, выпуклость масштабированного аппроксимирующего множества не зависит от уровня отбрасывания r и зависит только от выбора геометрии множества SAS.

Замечание 2. Свойство 1 можно конкретизировать и на случай $r = 0$. Пусть M таково, что $(1 - \varepsilon)^M \leq \delta$. Сгенерируем M н.о.р. пакетов $\mathbf{z}_i \in \mathcal{W}^L$

размера L и определим $\gamma_i \doteq \gamma^s(\theta_c, \Omega_0, \mathbf{z}_i)$. Допустим, что $\bar{\gamma} = \gamma_{1:M} > 0$. Тогда с вероятностью, не меньшей $1 - \delta$, справедливо $\text{Pr}_{\mathcal{W}^L} \{\theta_c \oplus \bar{\gamma}\Omega_0 \not\subseteq \Phi_s(\mathbf{z})\} \leq \varepsilon$.

3.2. Пакетное вероятностное масштабирование

Основная идея пакетного вероятностного масштабирования заключается в том, чтобы разделить выборки неопределенности на пакеты и допустить некоторые нарушения ограничений внутри каждого пакета. В отличие от обычного вероятностного масштабирования, где масштабный коэффициент вычисляется *независимо* для каждого ограничения, в рамках пакетного подхода ограничения внутри каждого пакета учитываются *все вместе*. В конечном счете, это может привести к более точным аппроксимациям множества ε -CCS и уменьшить размер выборки.

Пусть N сгенерированных ограничений разбиты на M пакетов по L ограничений в каждом, т.е. $\mathbf{z} = \{\mathbf{z}_1, \dots, \mathbf{z}_M\} = \{w_1, \dots, w_N\}$, причем $\mathbf{z} \in \mathcal{W}^N$ и $\mathbf{z}_i \in \mathcal{W}^L$ для $i = 1, \dots, M$. Следующая теорема позволяет определить подходящий масштабный коэффициент, используя пакетный подход, чтобы масштабированное множество SAS было полностью включено во множество ε -CCS с заданной доверительной вероятностью δ .

Теорема 1 [пакетное вероятностное масштабирование]. *Рассмотрим геометрию Ω_0 , центр масштабирования θ_c , параметр точности $\varepsilon \in (0; 1)$, доверительную вероятность $\delta \in (0; 1)$, а также неотрицательные целые числа M, L, s , $L > s$, удовлетворяющие условию*

$$(10) \quad B(s; L, \varepsilon)^M \leq \delta.$$

Для каждого пакета ограничений $i = 1, \dots, M$ пусть $\mathbf{z}_i \in \mathcal{W}^L$ и $\gamma_i \doteq \gamma^s(\theta_c, \Omega_0, \mathbf{z}_i)$ задано формулой (6). Предположим, что $\bar{\gamma} = \gamma_{1:M} > 0$. Тогда с вероятностью, не меньшей $1 - \delta$, справедливо

$$\theta_c \oplus \bar{\gamma}\Omega_0 \subseteq \Omega(\varepsilon).$$

Доказательство. Положим $p = 1 - B(s; L, \varepsilon)$. Согласно замечанию 2, выбрав M исходя из условия $B(s; L, \varepsilon)^M \leq \delta$ и $\bar{\gamma} = \gamma_{1:M} > 0$, с вероятностью, не меньшей $1 - \delta$, получаем

$$\text{Pr}_{\mathcal{W}^L} \{\theta_c \oplus \bar{\gamma}\Omega_0 \not\subseteq \Phi_s(\mathbf{z})\} \leq p.$$

Это условие эквивалентно $\text{Pr}_{\mathcal{W}^L} \{I_s^g(\theta, \mathbf{z}) = 1, \forall \theta \in \theta_c \oplus \bar{\gamma}\Omega_0\} \leq p$. Более того, из свойства 3 (см. Приложение 3) следует, что

$$(11) \quad \text{Pr}_{\mathcal{W}^L} \{I_s^g(\theta, \mathbf{z}) = 1\} \leq p \iff \text{Pr}_{\mathcal{W}} \{I^g(\theta, w) = 1\} \leq \varepsilon.$$

Таким образом, получаем $\text{Pr}_{\mathcal{W}^L} \{I_s^g(\theta, w) = 1\} \leq \varepsilon, \forall \theta \in \theta_c \oplus \bar{\gamma}\Omega_0$, что эквивалентно $\theta_c \oplus \bar{\gamma}\Omega_0 \subseteq \Omega(\varepsilon)$. \square

В отличие от обычного вероятностного масштабирования, размер выборки в PBPS задается двумя параметрами: количеством пакетов M и размером каждого пакета L . Размер выборки вычисляется по формуле $N = ML$. Следовательно, условие (10) определяется тремя настраиваемыми параметрами: M , L и s . Подобно уровню отбрасывания r обычного вероятностного масштабирования, большие значения уровня отбрасывания каждого пакета s делают аппроксимирующее множество менее чувствительным к экстремальным значениям. Что касается M и L , то их можно выбрать по любому критерию, например минимизировать размер выборки N . Более подробную информацию можно найти в [22].

В следующем разделе подход PBPS будет расширен за счет введения схемы “ужесточения” ограничений, а именно, *плотного погружения* (*tight immersion*), чтобы получить более точную аппроксимацию множества ε -CCS. Это расширение дает очевидный способ настройки параметров пакета (см. раздел 4.1).

4. Плотное погружение

Введем понятие плотного погружения.

Определение 7 [τ -плотное погружение]. Множество \mathcal{S} является τ -плотно погруженным во множество ε -CCS $\Omega(\varepsilon)$, если

$$(12a) \quad \mathcal{S} \subseteq \Omega(\varepsilon),$$

$$(12b) \quad \mathcal{S} \not\subseteq \Omega(\tau\varepsilon),$$

где $\tau \in [0, 1)$ – показатель плотности.

Замечание 3. Если множество ε -CCS $\Omega(\varepsilon)$ строго возрастает по ε , т.е. $\forall \tau_1, \tau_2 \in [0, 1)$, $\tau_1 < \tau_2$, то выполнено $\Omega(\tau_1\varepsilon) \subset \Omega(\tau_2\varepsilon)$. Следовательно, чем больше τ , тем больше будет $\Omega(\tau\varepsilon)$.

Плотное погружение гарантирует, что аппроксимирующее множество будет находиться *внутри* множества ε -CCS (12a) и, более того, *не внутри* консервативного множества, характеризуемого τ (12b). Поэтому оно накладывает более строгое условие, чем обычная *внутренняя аппроксимация*. Тем не менее, плотное погружение не должно использоваться для сравнения качества двух различных геометрий. Действительно (см. рис. 4), для одной и той же геометрии множество с наибольшим значением τ лучше аппроксимирует множество ε -CCS (рис. 4,а). Напротив, для разных геометрий плотное погружение само по себе не означает хорошей аппроксимации (рис. 4,б).

Следующее свойство дополняет определение 7 плотного погружения.

Свойство 2. Если аппроксимирующее множество $\underline{\Omega}(\varepsilon)$ является τ -плотно погруженным во множество $\Omega(\varepsilon)$, то оно также является $\tilde{\tau}$ -плотно погруженным в него, где $\tilde{\tau} \in [0, \tau)$.

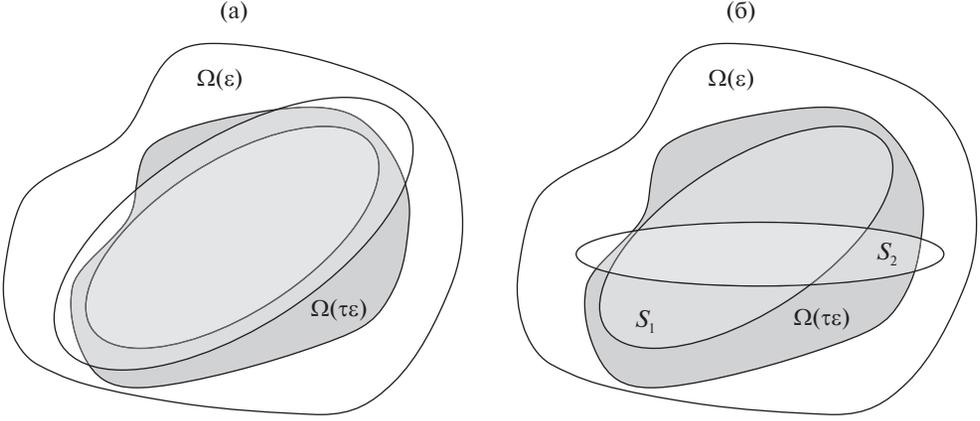


Рис. 4. Концепция плотного погружения.

Доказательство. По определению τ -плотное погружение множества $\underline{\Omega}(\varepsilon)$ во множество $\Omega(\varepsilon)$ означает $\underline{\Omega}(\varepsilon) \not\subseteq \Omega(\tau\varepsilon)$. Тогда для любого $\tilde{\tau} \in [0, \tau)$ имеем $\Omega(\tilde{\tau}\varepsilon) \subseteq \Omega(\tau\varepsilon)$. Следовательно, условие $\underline{\Omega}(\varepsilon) \not\subseteq \Omega(\tilde{\tau}\varepsilon)$ (12b) выполняется для любого $\tilde{\tau} \in [0, \tau)$ и доказательство завершено. \square

Далее выберем подходящие параметры пакета (M, L, s) , чтобы при пакетном вероятностном масштабировании условие (12b) выполнялось с доверительной вероятностью $1 - \bar{\delta}$, $\bar{\delta} \in (0, 1)$. Таким образом, при заданном множестве SAS Ω_0 с центром θ_c определим оптимальный масштабный коэффициент $\bar{\gamma}$, при котором масштабируемое множество $\mathcal{S} = \theta_c \oplus \bar{\gamma}\Omega_0$ будет плотно погружено в $\Omega(\varepsilon)$.

Теорема 2 [пакетное вероятностное масштабирование с плотным погружением]. *Рассмотрим множество SAS с геометрией Ω_0 и центром масштабирования θ_c , параметр точности $\varepsilon \in (0, 1)$, доверительную вероятность $\bar{\delta} \in (0, 1)$, показатель плотности $\tau \in [0, 1)$, а также неотрицательные целые числа M, L, s , $L > s$, удовлетворяющие условию*

$$(13) \quad B(s; L, \tau\varepsilon)^M \geq 1 - \bar{\delta}.$$

Сгенерируем M н.о.р. мультивыборок $\mathbf{z}_i \in \mathcal{W}^L$, $i = 1, \dots, M$, и для каждого i -го пакета вероятностных ограничений зададим масштабный коэффициент пакета в виде (6), т.е. $\gamma_i = \gamma^s(\theta_c, \Omega_0, \mathbf{z}_i)$. Допустим, что $\bar{\gamma} = \gamma_{1:M} > 0$. Тогда с вероятностью, не меньшей $1 - \bar{\delta}$, справедливо

$$\theta_c \oplus \bar{\gamma}\Omega_0 \subseteq \Omega(\varepsilon), \quad \theta_c \oplus \bar{\gamma}\Omega_0 \not\subseteq \Omega(\tau\varepsilon).$$

Доказательство. Пусть $p = B(s; L, \tau\varepsilon)$. Согласно свойству 5 из Приложения 5, выбрав параметры (M, L, s) исходя из (13), получаем оптимальный масштабный коэффициент $\bar{\gamma} = \gamma_{1:M} > 0$, удовлетворяющий с вероятностью, не меньшей $1 - \bar{\delta}$, неравенству $\text{Pr}_{\mathcal{W}^L} \{\theta_c \oplus \bar{\gamma}\Omega_0 \subseteq \Phi_s(\mathbf{z})\} \leq p$. Это условие можно эквивалентно записать как

$$\text{Pr}_{\mathcal{W}^L} \{I_s^g(\theta, \mathbf{z}) = 0, \quad \forall \theta \in \theta_c \oplus \bar{\gamma}\Omega_0\} \leq p.$$

Из свойства 4 (Приложение 2) следует, что

$$(14) \quad \Pr_{\mathcal{W}L} \{I_s^g(\theta, \mathbf{z}) = 0\} \leq p \iff \Pr_{\mathcal{W}} \{I^g(\theta, w) = 0\} \leq 1 - \tau\varepsilon.$$

Таким образом, можно заключить, что

$$\Pr_{\mathcal{W}} \{I^g(\theta, w) = 0\} \leq 1 - \tau\varepsilon, \quad \forall \theta \in \theta_c \oplus \bar{\gamma}\Omega_0,$$

т.е. $\theta_c \oplus \bar{\gamma}\Omega_0 \not\subseteq \Omega(\tau\varepsilon)$ с вероятностью, не меньшей $1 - \bar{\delta}$. \square

Замечание 4. Отметим, что в теореме 2 используется “уточняющая” доверительная вероятность $1 - \bar{\delta}$ вместо исходной $1 - \delta$. Она определяется пользователем и может быть выбрана ниже исходной доверительной вероятности для уменьшения размера выборки.

4.1. Настройка параметров пакета

В этом разделе покажем, как выбрать подходящие параметры (M, L, s) пакетного подхода для плотного погружения с доверительными вероятностями δ и $\bar{\delta}$ соответственно. В силу свойства 1 и теоремы 2 условия (12a) и (12b) будут выполнены при выборе параметров пакета (M, L, s) следующим образом:

$$(15a) \quad M \ln B(s; L, \varepsilon) \leq \ln \delta,$$

$$(15b) \quad M \ln B(s; L, \tau\varepsilon) \geq \ln(1 - \bar{\delta}).$$

Заметим, что неравенство (15a) дает вероятностные гарантии, а (15b) используется только для уточнения решения. Поскольку $B(s; L, \varepsilon)$ – отрицательная величина, можно разделить (15a) на $\ln B(s; L, \varepsilon)$, получив

$$M \geq \frac{\ln \delta}{\ln B(s; L, \varepsilon)}.$$

Следовательно, для справедливости (15a) достаточно выбрать

$$(16) \quad M = \left\lceil \frac{\ln \delta}{\ln B(s; L, \varepsilon)} \right\rceil.$$

Аналогично, для выполнения (15b) M должно быть выбрано исходя из

$$(17) \quad M \leq \frac{\ln(1 - \bar{\delta})}{\ln B(s; L, \tau\varepsilon)}.$$

Для заданного набора параметров $(\varepsilon, \delta, \tau, \bar{\delta})$ существует множество комбинаций (M, L, s) , удовлетворяющих (16) и (17). В данной работе предлагаются два различных критерия ζ : (i) минимизировать количество возможных комбинаций из $s + 1$ ограничений, т.е. $\zeta = M \binom{L}{s+1}$, (ii) минимизировать общий

размер выборки, т.е. $\zeta = ML$. Тогда параметры пакета (M, L, s) являются решением следующей задачи оптимизации:

$$(18) \quad \begin{aligned} (M^o, L^o, s^o) &= \underset{M, L, s \in \mathbb{N}_{\geq 0}}{\operatorname{argmin}} \quad \zeta \\ \text{при ограничениях} \quad M &\leq \frac{\ln(1 - \bar{\delta})}{\ln B(s; L, \tau \varepsilon)}, \\ M &= \left\lceil \frac{\ln \delta}{\ln B(s; L, \varepsilon)} \right\rceil, \\ L &\geq s + 1. \end{aligned}$$

Для решения задачи (18) используем полный перебор [31], чтобы найти подходящую комбинацию параметров пакета (M, L, s) (см. пример ниже).

Пример 2. При заданных $\varepsilon = 0,05, \delta = 0,001, \bar{\delta} = 0,1$ для каждого $s = [1, 30]$ определим M по формуле (16). Затем проверим условие (17) для пар (L, s) , где $L = [s + 1, \dots, s + 300]$. Наконец, среди всех пар, удовлетворяющих (17), выберем ту, что доставляет минимум критерию ζ (либо $\zeta = M \binom{L}{s+1}$, либо $\zeta = ML$). В табл. 1 приведены параметры пакета (M, L, s) , полученные при решении задачи (18) с помощью предложенного подхода (оба критерия ζ). Согласно табл. 1 для каждого рассматриваемого критерия увеличение показателя плотности τ приводит к росту как необходимого размера выборки (N), так и комбинаторной сложности ($M \binom{L}{s+1}$). Если количество доступных наблюдений w_i ограничено, можно добиться плотного погружения, используя всего 390 наблюдений, но за счет более высокой комбинаторной сложности. Более того, можно ограничить сверху размер выборки $N = ML$ в задаче оптимизации (18) и получить параметры пакета, минимизирующие комбинаторную сложность при выполнении ограничения на размер выборки. Обращаем внимание, что плотное погружение обычно находится в автономном режиме (офлайн), поэтому сложность вычисления аппроксимации CCS не мешает работе контуров управления в режиме онлайн.

Таблица 1. Параметры пакета, размер выборки и количество возможных комбинаций TI-RBPS для $\tau = [0,2, 0,3, 0,4, 0,5]$, доставляющих минимум двум критериям ζ

τ	Критерий: $\min M \binom{L}{s+1}$					Критерий: $\min N$				
	M	L	s	N	$M \binom{L}{s+1}$	M	L	s	N	$M \binom{L}{s+1}$
0,2	43	27	2	1,16e+03	1,25e+05	2	195	4	3,90e+02	4,46e+09
0,3	155	27	3	4,19e+03	2,72e+06	2	303	8	6,06e+02	1,05e+17
0,4	2681	20	4	5,36e+04	4,16e+07	4	278	10	1,11e+03	6,28e+19
0,5	15 033	29	6	4,36e+05	2,35e+10	8	309	14	2,47e+03	9,68e+25

5. Результаты

Сравним аппроксимации множества $0,05\text{-CCS}$ (см. пример 1), полученные с помощью обычного вероятностного масштабирования (PS) и пакетного вероятностного масштабирования с плотным погружением (TI-PBPS) для различных размерностей n_θ . В качестве начальной SAS-аппроксимации $\theta_c \oplus \Omega_0$ выберем единичный шар с центром в начале координат. Таким образом, получаемые аппроксимирующие множества являются шарами с центром в начале координат и радиусом α . Для каждой из рассматриваемых размерностей используем процедуру Монте-Карло для равномерного генерирования 10^7 линейных ограничений в виде касательных к единичной сфере. Пользуясь симметрией задачи, вычислим точки пересечения вероятностных ограничений с фиксированной координатной осью и используем их для вычисления точного значения радиуса для множества $0,05\text{-CCS}$. Затем найдем радиусы аппроксимирующих множеств, полученных с помощью PS и TI-PBPS для различных значений показателя плотности τ . Для снижения вариативности результатов радиусы TI-PBPS соответствуют медианному радиусу трех отдельных экспериментов, каждый из которых содержит различные реализации ограничений.

Согласно рис. 5 для данной конкретной задачи техника TI-PBPS дает существенно лучшие результаты, чем PS (самая нижняя пунктирная линия): полученные аппроксимирующие радиусы ближе к реальным (самая верхняя пунктирная линия). Как и ожидалось, с увеличением показателя плотности τ качество аппроксимации, полученной с помощью TI-PBPS, повышается.

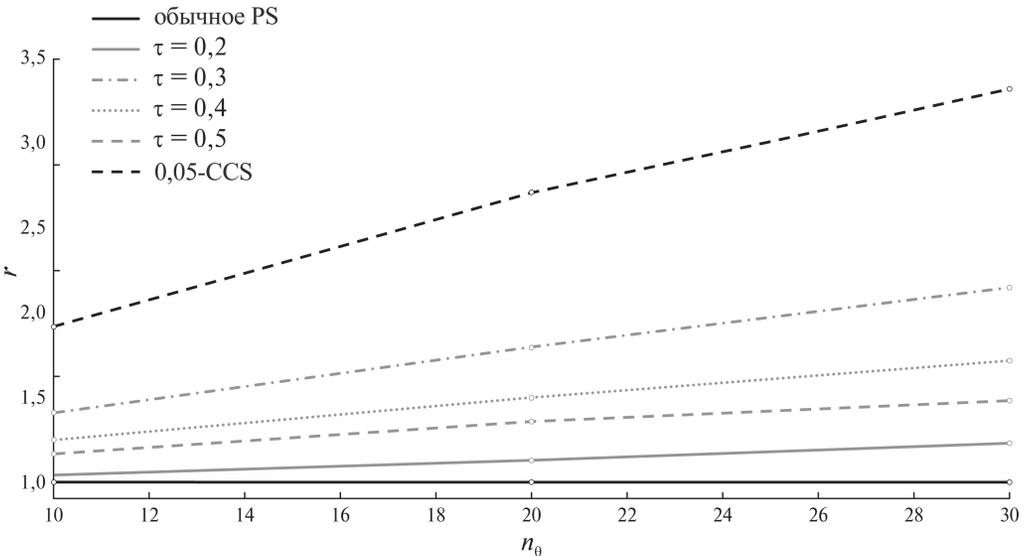


Рис. 5. Сравнение радиусов α аппроксимирующего множества для различных размерностей задачи n_θ , полученных с помощью PS и TI-PBPS при $\varepsilon = 0,05$, $\delta = 0,001$ и $\bar{\delta} = 0,1$.

6. Заключение

В данной работе рассмотрен подход вероятностного масштабирования для вычисления аппроксимаций множеств с вероятностными ограничениями на основе случайных выборок. Этот подход позволяет пользователю сначала выбрать любое множество, а затем применить линейное преобразование для аппроксимации вероятностно-надежной области с желаемыми вероятностными гарантиями. В результате сложность аппроксимации настраивается априори. Кроме того, предложен пакетный вариант вероятностного масштабирования с плотным погружением, позволяющий сделать решение менее консервативным. Компромисс между размером выборки, сложностью задачи и степенью консерватизма этого подхода может быть настроен пользователем. Будущие исследования могут быть направлены на улучшение предложенного решения, например, за счет применения схем выборки по важности [32].

ПРИЛОЖЕНИЕ 1

Свойство 3. Рассмотрим целочисленные параметры $L > s \geq 0$, пакет $\mathbf{z} \in \mathcal{W}^L$ и допустимый уровень нарушения $\varepsilon \in (0; 1)$. Для $w \in \mathcal{W}$ справедливо неравенство

$$(П.1.1) \quad \Pr_{\mathcal{W}^L} \{I_s^g(\theta, \mathbf{z}) = 1\} \leq 1 - B(s; L, \varepsilon) \iff \Pr_{\mathcal{W}} \{I^g(\theta, w) = 1\} \leq \varepsilon.$$

Доказательство. Определим $E(\theta) = \Pr_{\mathcal{W}} \{I^g(\theta, w) = 1\}$. Тогда

$$(П.1.2) \quad \Pr_{\mathcal{W}^L} \{I_s^g(\theta, \mathbf{z}) = 0\} = \sum_{i=0}^s \binom{L}{i} E(\theta)^i (1 - E(\theta))^{L-i} = B(s; L, E(\theta)).$$

Введем обозначение $p = 1 - B(s; L, \varepsilon)$. Поскольку $B(s; L, \varepsilon)$ – строго убывающая функция ε (см. [23, свойство 4]), имеем

$$(П.1.3) \quad B(s; L, E(\theta)) \geq B(s; L, \varepsilon) = 1 - p \iff E(\theta) \leq \varepsilon.$$

Следовательно,

$$(П.1.4) \quad \Pr_{\mathcal{W}^L} \{I_s^g(\theta, \mathbf{z}) = 1\} \leq p \iff \Pr_{\mathcal{W}} \{I^g(\theta, w) = 1\} \leq \varepsilon.$$

Доказательство завершено. \square

ПРИЛОЖЕНИЕ 2

Свойство 4. Рассмотрим целочисленные параметры $L > s \geq 0$, пакет $\mathbf{z} \in \mathcal{W}^L$, выборку $w \in \mathcal{W}$ и допустимый уровень нарушения $\varepsilon \in (0; 1)$. Справедливо неравенство

$$\Pr_{\mathcal{W}^L} \{I_s^g(\theta, \mathbf{z}) = 0\} \leq B(s; L, \tau\varepsilon) \iff \Pr_{\mathcal{W}} \{I^g(\theta, w) = 0\} \leq 1 - \tau\varepsilon.$$

Доказательство. Вспомнив определение $E(\theta) = \Pr_{\mathcal{W}} \{I^g(\theta, w) = 1\}$, имеем

$$\Pr_{\mathcal{W}^L} \{I_s^g(\theta, \mathbf{z}) = 0\} = \sum_{i=0}^s \binom{L}{i} E(\theta)^i (1 - E(\theta))^{L-i} = B(s; L, E(\theta)).$$

Поскольку $B(s; L, \tau\varepsilon)$ строго убывает по $\tau\varepsilon$ (см. [23, свойство 4]), получаем

$$\begin{aligned} \Pr_{\mathcal{W}^L} \{I_s^g(\theta, \mathbf{z}) = 0\} = B(s; L, E(\theta)) \leq B(s; L, \tau\varepsilon) &\iff \\ \iff \Pr_{\mathcal{W}} \{I^g(\theta, w) = 1\} = E(\theta) \geq \tau\varepsilon. \end{aligned}$$

Таким образом,

$$\Pr_{\mathcal{W}^L} \{I_s^g(\theta, \mathbf{z}) = 0\} \leq B(s; L, \tau\varepsilon) \iff \Pr_{\mathcal{W}} \{I^g(\theta, w) = 0\} \leq 1 - \tau\varepsilon. \quad \square$$

ПРИЛОЖЕНИЕ 3

Свойство 5. При заданных параметре точности $p \in (0, 1)$ и доверительной вероятности $\bar{\delta} \in (0, 1)$ предположим, что количество пакетов M выбрано исходя из условия

$$1 - p^M \leq \bar{\delta}.$$

Для каждого пакета ограничений $i = 1, \dots, M$ сгенерируем M н.о.р. мультивыборок $\mathbf{z} \sim \Pr_{\mathcal{W}^L}$ и определим $\gamma_i = \gamma^s(c, \Omega_0, \mathbf{z}_i)$. Предположим, что $\bar{\gamma} = \gamma_{1:M} > 0$. Тогда с вероятностью, не меньшей $1 - \bar{\delta}$, справедливо

$$\Pr_{\mathcal{W}^L}^M \{c \oplus \bar{\gamma} \Omega_0 \subseteq \Phi_s^g(\mathbf{z})\} \leq p.$$

Доказательство. Этот результат доказывается по аналогии со свойством 1. Рассмотрим следующую задачу оптимизации:

$$\begin{aligned} \text{(П.3.1)} \quad & \min_{\gamma} \quad \gamma \\ & \text{при ограничениях} \quad c \oplus \gamma \Omega_0 \not\subseteq \Phi_s^g(\mathbf{z}), \quad i = 1, \dots, M. \end{aligned}$$

Если задача (П.3.1) имеет допустимое решение, то в силу (6) можно записать (П.3.1) в виде

$$\begin{aligned} \text{(П.3.2)} \quad & \min_{\gamma} \quad \gamma \\ & \text{при ограничениях} \quad \gamma > \gamma^s(c, \Omega_0, \mathbf{z}), \quad i = 1, \dots, M. \end{aligned}$$

Согласно подходу “генерирование/отбрасывание” [21, 26] при отбрасывании не более $M - 1$ ограничений и $d = 1$ (одна переменная), вероятность нарушения масштабированного аппроксимирующего множества ограничений не

превышает $p \in (0; 1)$ с вероятностью, не меньшей $1 - \bar{\delta}$, где доверительная вероятность $\bar{\delta}$ определяется следующим образом:

$$\begin{aligned}\bar{\delta} &= \binom{M-1}{M-1} B(M-1; M, p) = \sum_{i=0}^{M-1} \binom{M}{i} p^i (1-p)^{M-i} = \\ &= 1 - \sum_{i=M}^M \binom{M}{i} p^i (1-p)^{M-i} = 1 - p^M.\end{aligned}$$

Если отбросить не более $M - 1$ ограничений, то оптимальное решение задачи (П.3.2) будет иметь вид $\bar{\gamma} = \gamma_{1:M}$, причем $\gamma_i = \gamma^s(c, \Omega_0, \mathbf{z})$. Соответственно, получаем $\Pr_{\mathcal{W}^L}^M \{\bar{\gamma} \leq \gamma(c, \Omega_0, \mathbf{z})\} \leq p$; следовательно, с вероятностью, не меньшей $1 - \bar{\delta}$, справедливо

$$\Pr_{\mathcal{W}^L}^M \{\theta_c \oplus \bar{\gamma} \Omega_0 \subseteq \Phi_s^g(\mathbf{z})\} \leq p. \quad \square$$

СПИСОК ЛИТЕРАТУРЫ

1. *Bemporad A., Morari, M.* Robust model predictive control: A survey / Robustness in Identification and Control. Lecture Notes in Control and Information Sciences. Garulli A., Tesi A. (Eds.). London: Springer, 1999. V. 245. <https://doi.org/10.1007/BFb0109870>
2. *Mayne D., Seron M., Raković S.* Robust model predictive control of constrained linear systems with bounded disturbances // Automatica. 2005. V. 41. No. 2. P. 219–224.
3. *Mayne D., Raković S., Findeisen R., Allgöwer F.* Robust output feedback model predictive control of constrained linear systems // Automatica. 2006. V. 42. No. 7. P. 1217–1222.
4. *Lorenzen M., Dabbene F., Tempo R., Allgöwer F.* Constraint-tightening and stability in stochastic model predictive control // IEEE Transactions on Automatic Control. 2016. V. 62. No. 7. P. 3165–3177.
5. *Farina M., Giullioni L., Scattolini R.* Stochastic linear model predictive control with chance constraints – a review // J. Process Control. 2016. V. 44. P. 53–67.
6. *Mesbah A.* Stochastic model predictive control: An overview and perspectives for future research // IEEE Control Systems Magazine. 2016. V. 36. No. 6. P. 30–44.
7. *Charnes A., Cooper W.* Chance constraints and normal deviates // J. Amer. Stat. Associat. 1962. V. 57. No. 297. P. 134–148.
8. *Küçükavuz S., Jiang, R.* Chance-constrained optimization under limited distributional information: A review of reformulations based on sampling and distributional robustness // EURO J. Comput. Optim. 2022. V. 10. Art. No. 100030.
9. *Mammarella M., Mirasierra V., Lorenzen M., Alamo T., Dabbene F.* Chance-constrained sets approximation: A probabilistic scaling approach // Automatica. 2022. V. 137. Art. No. 110108.
10. *Polyak B., Tempo R.* Probabilistic robust design with linear quadratic regulators // Syst. Control Lett. 2001. V. 43. No. 5. P. 343–353.

11. *Calafiore G., Polyak B.* Stochastic algorithms for exact and approximate feasibility of robust LMIs // IEEE Transactions on Automatic Control. 2001. V. 46. No. 11. P. 1755–1759.
12. *Dabbene F., Gay P., Polyak B.* Recursive algorithms for inner ellipsoidal approximation of convex polytopes // Automatica. 2003. V. 39. P. 1773–1781.
13. *Calafiore G., Campi M.* The scenario approach to robust control design // IEEE Transactions on Automatic Control. 2006. V. 51. No. 5. P. 742–753.
14. *Tempo R., Calafiore G., Dabbene F.* Randomized Algorithms for Analysis and Control of Uncertain Systems, with Applications. London: Springer, 2013.
15. *Polyak B., Shcherbakov P.* Randomization in robustness, estimation, and optimization / Uncertainty in Complex Networked Systems: In Honor of Roberto Tempo. Basar T. (Ed.). Cham: Springer, 2018. P. 181–208.
16. *Kataoka S.* A stochastic programming model // Econometrica: J. Econom. Soc. 1963. P. 181–196.
17. *Prékopa A.* Logarithmic concave measures with application to stochastic programming // Acta Scientiarum Mathematicarum. 1971. V. 32. P. 301–316.
18. *Geng X., Xie L.* Data-driven decision making in power systems with probabilistic guarantees: Theory and applications of chance-constrained optimization // Ann. Rev. Control. 2019. V. 47. P. 341–363.
19. *Soudjani S., Majumdar R.* Concentration of measure for chance-constrained optimization // IFAC-PapersOnLine. 2018. V. 51. No. 16. P. 277–282.
20. *Lejeune M., Prékopa A.* Relaxations for probabilistically constrained stochastic programming problems: Review and extensions // Annals of Operations Research. 2018. P. 1–22.
21. *Campi M., Garatti S.* A sampling-and-discarding approach to chance-constrained optimization: Feasibility and optimality // J. Optim. Theory Appl. 2011. V. 148. No. 2. P. 257–280.
22. *Alamo T., Mirasierra V., Dabbene F., Lorenzen M.* Safe approximations of chance constrained sets by probabilistic scaling // 2019 18th European Control Conference (ECC). IEEE, 2019. P. 1380–1385.
23. *Alamo T., Tempo R., Luque A., Ramirez D.* Randomized methods for design of uncertain systems: Sample complexity and sequential algorithms // Automatica. 2015. V. 52. P. 160–172.
24. *Lorenzen M., Dabbene F., Tempo R., Allgöwer F.* Stochastic MPC with offline uncertainty sampling // Automatica. 2017. V. 81. P. 176–183.
25. *Mammarella M., Lorenzen M., Capello E., Park H., Dabbene F., Guglieri G., Romano M., Allgöwer F.* An offline-sampling SMPC framework with application to autonomous space maneuvers // IEEE Transactions on Control Systems Technology. 2018. V. 28. No. 2. P. 388–402.
26. *Calafiore G.* Random convex programs // SIAM J. Optim. 2010. V. 20. No. 6. P. 3427–3464.
27. *Vapnik V.* The Nature of Statistical Learning Theory. New York: Springer Science & Business Media, 1999.
28. *Alamo T., Tempo R., Camacho E.* Randomized strategies for probabilistic solutions of uncertain feasibility and optimization problems // IEEE Transactions on Automatic Control. 2009. V. 54. No. 11. P. 2545–2559.

29. *Alamo T., Tempo R., Camacho E.* Improved sample size bounds for probabilistic robust control design: A pack-based strategy // 2007 46th IEEE Conference on Decision and Control. IEEE, 2007. P. 6178–6183.
30. *Alamo T., Manzano J., Camacho E.* Robust design through probabilistic maximization / Uncertainty in Complex Networked Systems. Springer, 2018. P. 247–274.
31. *Nievergelt J.* Exhaustive search, combinatorial optimization and enumeration: Exploring the potential of raw computing power // International Conference on Current Trends in Theory and Practice of Computer Science. Springer, 2000. P. 18–35.
32. *Lukashevich A., Gorchakov V., Vorobev P., Deka D., Maximov Y.* Importance sampling approach to chance-constrained DC optimal power flow // IEEE Transactions on Control of Network Systems. 2023. V. 11. No. 2. P. 928–937.

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

Поступила в редакцию 03.03.2025

После доработки 20.05.2025

Принята к публикации 27.06.2025

© 2025 г. В.Ф. СОКОЛОВ, д-р физ.-мат. наук (sokolov@ipm.komisc.ru)
(Коми научный центр УрО РАН, Сыктывкар)

ОПТИМАЛЬНОЕ РОБАСТНОЕ СЛЕЖЕНИЕ ДЛЯ ДИСКРЕТНОГО МИНИМАЛЬНО-ФАЗОВОГО ОБЪЕКТА ПРИ НЕИЗВЕСТНЫХ СМЕЩЕНИИ И НОРМЕ ВНЕШНЕГО ВОЗМУЩЕНИЯ И НОРМАХ НЕОПРЕДЕЛЕННОСТЕЙ¹

Для дискретного минимально-фазового объекта управления с известной приближенной номинальной моделью, ограниченным смещенным внешним возмущением и операторными возмущениями по выходу и управлению рассмотрена задача оптимального отслеживания заданного ограниченного сигнала. Смещение и норма внешнего возмущения и нормы операторных возмущений предполагаются неизвестными. Показателем качества слежения служит зависящая от названных неизвестных параметров и задающего сигнала наихудшая асимптотическая ошибка отслеживания заданного ограниченного сигнала в классе рассматриваемых возмущений. Решение задачи оптимального слежения с заданной точностью базируется на оптимальной квантификации возмущений в рамках ℓ_1 -теории робастного управления, полиэдральном оценивании неизвестных параметров и использовании показателя качества задачи управления как идентификационного критерия.

Ключевые слова: робастное управление, оптимальное управление, ограниченное возмущение, неопределенность, квантификация возмущений, множественное оценивание.

DOI: 10.31857/S0005231025080041, EDN: UTCMAP

1. Введение

В настоящей работе рассматривается задача оптимального слежения для линейного дискретного динамического объекта с заданной и тестируемой передаточной функцией. Предполагается, что на объект воздействуют ограниченное внешнее возмущение с неизвестными смещением и границами и операторные возмущения (неопределенности) по выходу и управлению с неизвестными нормами (коэффициентами усиления). Задача рассматривается в рамках ℓ_1 -теории робастного управления, заложенной в работах [1, 2] и соответствующей сигнальному пространству ℓ_∞ ограниченных вещественных последовательностей. Сложность задачи заключается в том, что для минимизации показателя качества в виде наихудшей в классе допустимых возмущений асимптотической ошибки слежения необходима компенсация неизвестного смещения и обоснованная оптимальная оценка показателя качества в условиях неидентифицируемости всех указанных выше неизвестных параметров.

¹ Работа выполнена в рамках госзадания ФМИ ФИЦ Коми НЦ УрО РАН, проект № 125031203621-2.

Решение описанной оптимальной задачи слежения базируется на оптимальной квантификации возмущений с помощью множественного оценивания неизвестных параметров (set-membership approach) и использовании показателя качества задачи управления как идентификационного критерия. Множественное оценивание в идентификации систем, изначально базировавшееся на предположении известных верхних границ детерминированных возмущений, приобрело широкую популярность в конце 1980-х и сводилось к разработке вычислимых верхних и нижних аппроксимаций (эллипсоидов, параллелотопов и т.п.) множеств параметров, согласованных с данными измерений (ограничимся ссылками на первые специальные выпуски двух ведущих журналов по теории управления [3, 4]). Описания применений этих аппроксимаций к задачам управления редки и сопровождаются различными дополнительными предположениями типа априори известного стабилизирующего управления. Сторонниками стохастических моделей возмущений этот подход критикуется за консерватизм, порождаемый априорными предположениями об известных верхних границах возмущений. Параллельно с начала 1990-х начались активные исследования в области идентификации для робастного управления и квантификации неопределенностей. Полтора десятилетия спустя в обзорной статье [5] отмечалось, что деятельность по оценке множеств неопределенности часто ошибочно относилась к идентификации для управления, поскольку в большинстве этих работ цель управления не учитывалась при идентификации. Проблемы верификации модели и оценки неопределенностей до настоящего времени остаются актуальными [6, 7], но по-прежнему рассматриваются в основном вне контекста задач управления и с искусственными критериями, мотивированными целями собственно идентификации.

В настоящей работе оценка смещения и квантификация возмущений базируются на множественном оценивании и прямом использовании показателя качества задачи управления как идеального идентификационного критерия. Потенциальные возможности применения такого подхода возникают благодаря двум обстоятельствам. Во-первых, в ℓ_1 -теории робастного управления получены явные представления для асимптотических показателей качества через индуцированные нормы передаточных функций замкнутой системы управления и нормы всех возмущений и неопределенностей [8–11]. Во-вторых, модель ограниченных возмущений допускает прямое использование текущих данных измерений для онлайн верификации модели [12]. В общем случае такой подход к идентификации для управления вычислительно не реализуем ввиду сложности вычисления текущих оптимальных оценок. Но он реализуем в случае линейных или дробно-линейных относительно оцениваемых параметров показателей качества [13]. В рассматриваемой задаче показатель качества является невыпуклой дробно-квадратичной функцией неизвестных параметров (см. представление (9)). При известном смещении показатель качества становится дробно-линейным, и задача квантификации возмущений для этого случая решена в [14], где также сформулирована идея оценивания неизвестного смещения с использованием сетки тестируемых значений. В на-

стоящей статье дается строгое обоснование этой идеи и доказывается строгий результат о решении поставленной задачи асимптотически оптимального слежения с заданной точностью. Результаты моделирования и комментарии к ним иллюстрируют эффективность предложенного решения.

Обозначения:

$|\varphi|$ – евклидова норма вектора $\varphi \in \mathbb{R}^n$;

$x_s^t = (x_s, x_{s+1}, \dots, x_t)$ для вещественной последовательности $x = (\dots, x_{-1}, x_0, x_1, \dots)$;

$|x_s^t| = \max_{s \leq k \leq t} |x_k|$;

$\|x\|_{ss} = \limsup_{t \rightarrow +\infty} |x_t|$;

$\|x\|_\infty = \sup_t |x_t|$ – норма в пространстве ограниченных последовательностей ℓ_∞ ;

$\|x\|_1 = \sum_{k=0}^{+\infty} |x_k|$ – норма в пространстве абсолютно суммируемых последовательностей ℓ_1 ;

$\|G\| = \sum_{k=0}^{+\infty} |g_k| = \|g\|_1$ – индуцированная норма устойчивой линейной стационарной причинной системы $G : \ell_\infty \rightarrow \ell_\infty$ с передаточной функцией $G(\lambda) = \sum_{k=0}^{+\infty} g_k \lambda^k$.

2. Модель управляемого объекта и содержательная постановка задачи

Модель управляемого объекта описывается уравнением

$$(1) \quad a(q^{-1})y_t = b(q^{-1})u_t + v_t, \quad t = 1, 2, 3, \dots,$$

где $y_t \in \mathbb{R}$ – измеряемый выход объекта в момент времени t , $u_t \in \mathbb{R}$ – управление, $v_t \in \mathbb{R}$ суммарное возмущение, q^{-1} – оператор сдвига назад ($q^{-1}y_t = y_{t-1}$). Начальные данные $y_{1-n}^0 = (y_{1-n}, \dots, y_0)$ произвольны, $u_t = 0$ при $t \leq 0$. Полиномы

$$a(\lambda) = 1 + a_1\lambda + \dots + a_n\lambda^n, \quad b(\lambda) = b_1\lambda + \dots + b_m\lambda^m$$

характеризуют **номинальную модель** объекта, т.е. модель без возмущения v . Суммарное возмущение v имеет вид

$$(2) \quad v_t = c^w + \delta^w w_t + \delta^y \Delta^1(y)_t + \delta^u \Delta^2(u)_t, \quad \|w\|_\infty \leq 1, \quad \delta^w > 0, \quad \delta^y > 0, \quad \delta^u > 0.$$

Параметры c^w и δ^w в (2) характеризуют соответственно смещение внешнего возмущения $c^w + \delta^w w_t$ и верхнюю границу несмещенного возмущения $\delta^w w$. Числа $\delta^y > 0$ и $\delta^u > 0$ – коэффициенты усиления (индуцированные нормы) операторных возмущений по выходу и управлению соответственно и

$$(3) \quad |\Delta^1(y)_t| \leq p_t^y = \max_{t-\mu \leq k \leq t-1} |y_k|, \quad |\Delta^2(u)_t| \leq p_t^u = \max_{t-\mu \leq k \leq t-1} |u_k|.$$

В ℓ_1 -теории робастного управления эти возмущения называются неопределенностями с ограниченной памятью μ , которая гарантирует независимость

асимптотической динамики замкнутой системы управления от начальных данных. Память неопределенностей μ выбирается конструктором сколь угодно большой, но не бесконечной, без ущерба для гарантируемого качества управления. Описание возмущений в виде (2), (3) эквивалентно (см. [10, 11]) неравенствам

$$(4) \quad |v_t - c^w| \leq \delta^w + \delta^y p_t^y + \delta^u p_t^u \quad \forall t.$$

Априорная информация об управляемом объекте содержится в следующих предположениях.

П1. Полиномы $a(\lambda)$ и $b(\lambda)$ номинального объекта управления известны, $b_1 \neq 0$.

П2. Корни полинома $\frac{b(\lambda)}{\lambda}$ лежат вне замкнутого единичного круга комплексной плоскости.

П3. Столбец параметров $\delta = (\delta^w, \delta^y, \delta^u)^T$ неизвестен, и неизвестно смещение $c^w \in [c_{\min}^w, c_{\max}^w]$ с известными c_{\min}^w, c_{\max}^w .

П4. Известна асимптотическая верхняя граница $\|r\|_{ss}$ задающего сигнала r или ее оценка сверху.

Предположение П1 подразумевает и тот случай, когда “истинная” номинальная модель неизвестна и для тестирования предоставлена ее оценка, полученная каким-либо методом ее идентификации. Предположение П2 гарантирует ограниченность управления u , если выход объекта y ограничен (такой объект называется минимально-фазовым). Предположение П4 будет прокомментировано после строгой постановки задачи в конце раздела 3. Еще одно обязательное предположение об ограничениях на нормы операторных возмущений будет сформулировано в разделе 3 после теоремы 1.

Содержательная постановка задачи заключается в построении управления, минимизирующего наихудшую асимптотическую ошибку отслеживания заданного ограниченного сигнала для множества возмущений, удовлетворяющих неравенствам (4).

Для решение оптимальной задачи с заданной точностью необходимы онлайн квантификация возмущений (т.е. оценивание их неизвестных параметров δ) для оценки качества слежения и компенсации неизвестного смещения c^w .

3. Качество слежения оптимального регулятора при известном смещении c^w и постановка задачи

Пусть $r = (r_1, r_2, r_3, \dots)$ – заданный ограниченный сигнал ($r \in \ell_\infty$), и показатель качества задачи слежения имеет вид

$$(5) \quad J_\mu(c^w, \delta) = \sup_{v \in V} \|y - r\|_{ss}, \quad \|y - r\|_{ss} := \limsup_{t \rightarrow +\infty} |y_t - r_t|,$$

где V – множество возмущений v , удовлетворяющих неравенствам (4).

Рассмотрим регулятор, описываемый уравнением

$$(6) \quad b(q^{-1})u_t = (a(q^{-1}) - 1)y_t + r_t - c^w.$$

Заметим, что уравнение (6) задает значение u_{t-1} , а не u_t , не входящее в это уравнение. Тогда для выхода замкнутой системы управления (1), (6) получаем

$$(7) \quad y_t - r_t = v_t - c^w = \delta^w w_t + \delta^y \Delta^1(y)_t + \delta^u \Delta^2(u)_t.$$

В силу произвольности и непредсказуемости правой части в (7) регулятор (6) является **оптимальным** для показателя качества (5).

Определение 1. Замкнутая система (1), (6) называется робастной устойчивой в классе возмущений V , если $J_\mu(c^w, \delta) < +\infty$.

Для формулировки теоремы о качестве оптимального регулятора (6) введем обозначения для его передаточных функций от y и r к управлению u :

$$G_{uy}(\lambda) = \frac{a(\lambda) - 1}{b(\lambda)}, \quad G_{ur}(\lambda) = \frac{1}{b(\lambda)}.$$

Теорема 1. Пусть выполнены предположения П1, П2. Тогда справедливы следующие утверждения.

1) Замкнутая система (1), (6) робастно устойчива в классе V при памяти возмущений $\mu = +\infty$ если и только если

$$(8) \quad \delta^y + \delta^u \|G_{uy}\| < 1.$$

Для системы с нулевыми начальными данными y_{1-n}^0 и $\mu = +\infty$

$$(9) \quad J(c^w, \delta) := J_{+\infty}(c^w, \delta) = \frac{\delta^w + \delta^y \|r\|_{ss} + \delta^u (|c^w| + \|r\|_{ss}) \|1/b(q^{-1})\|}{1 - \delta^y - \delta^u \|G_{uy}\|}.$$

2) Для системы с произвольными начальными данными y_{1-n}^0 и $\mu < +\infty$

$$(10) \quad J_\mu(c^w, \delta) \leq J(c^w, \delta) \quad \forall \mu > 0,$$

и если последовательность $|r|$ равномерно часто попадает в окрестности верхнего предела $\|r\|_{ss}$ (см. определение в [10]), то при любых начальных данных

$$(11) \quad J_\mu(c^w, \delta) \nearrow J(c^w, \delta) \quad (\mu \rightarrow +\infty),$$

где знак \nearrow означает монотонную сходимость снизу при $\mu \rightarrow +\infty$.

Доказательство теоремы 1 приведено в [14].

Заключительное предположение об ограничении на нормы операторных возмущений П5 вытекает из теоремы 1.

П5. Известно число $\bar{\delta}$ такое, что

$$(12) \quad \delta^y + \delta^u \|G_{uy}\| \leq \bar{\delta} < 1.$$

Предположение П5 не ограничительно по сравнению с условием робастной устойчивости (8). По смыслу поставленной задачи параметр $\bar{\delta}$ выбирается конструктором и может быть выбран сколь угодно близким к 1. Но для близких к 1 значений $\delta^y + \delta^u \|G_{uy}\|$ показатель качества $J_\mu(c^w, \delta)$ становится слишком большим и номинальную модель с заданными тестируемыми полиномами $a(\lambda)$ и $b(\lambda)$ или объект с такими операторными возмущениями можно считать неприемлемыми.

Постановка задачи. Требуется при априорной информации П1–П5 и заданном сигнале слежения r построить обратную связь вида $u_t = U_t(y_1^t, u_1^{t-1})$ (и с конечной памятью), гарантирующую с заданной точностью неравенство

$$(13) \quad \|y - r\|_{ss} \leq J(c^w, \delta).$$

Главная сложность поставленной задачи заключается в обеспечении неравенства (13) в условиях неидентифицируемости c^w и δ (см. раздел 4.1).

Используемый далее как идентификационный критерий показатель (9) зависит от $\|r\|_{ss}$. Если это значение априори не известно, вместо него можно использовать рекуррентно вычисляемые неубывающие оценки $R_t = \max_{1 \leq k \leq t} |r_k| \leq \|r\|_\infty$, чтобы гарантировать принципиально теоретически наилучшую гарантию качества слежения с заменой $\|r\|_{ss}$ на $\|r\|_\infty$.

4. Оптимальное слежение

Решение поставленной задачи базируется на оптимальной квантификации возмущений для тестируемой номинальной модели.

4.1. Оптимальная квантификация возмущений при известном смещении c^w

Из уравнения объекта (1) и неравенств (4) следует, что при известном смещении c^w полная информация о неизвестном δ в момент времени t содержится в априорном предположении П5 и включении

$$(14) \quad \delta \in D_t = \left\{ \hat{\delta} \geq 0 \mid |a(q^{-1})y_k - b(q^{-1})u_k - c^w| \leq \hat{\delta}^w + \hat{\delta}^y p_k^y + \hat{\delta}^u p_k^u \quad \forall k \leq t \right\},$$

где $\hat{\delta} = (\hat{\delta}^w, \hat{\delta}^y, \hat{\delta}^u)^T$. Система неравенств в (14) эквивалентна описанию системы (1)–(4) на промежутке $[1, t]$ при любом управлении u_0^{t-1} . Тогда наилучшая относительно показателя качества J согласованная с измерениями y_0^t, u_0^{t-1} оценка параметра δ имеет вид

$$(15) \quad \delta_t = \underset{\hat{\delta} \in D_t}{\operatorname{argmin}} J(c^w, \hat{\delta}).$$

Оптимальная задача (15) является задачей дробно-линейного программирования относительно неизвестной строки $\hat{\delta}$ и стандартным образом сводится к задаче линейного программирования с помощью введения дополнительной вещественной переменной [15]. Число линейных относительно $\hat{\delta}$ неравенств в описании множеств D_t может неограниченно возрастать с ростом t . Для обеспечения ограниченности числа неравенств и сходимости полиэдральных и векторных оценок неизвестного столбца δ за конечное время выберем параметр $\varepsilon_1 > 0$, задающий размер мертвой зоны при обновлении оценок. Начальная полиэдральная оценка δ имеет вид

$$P_0 = \left\{ \hat{\delta} = (\hat{\delta}^w, \hat{\delta}^y, \hat{\delta}^u)^T \mid \hat{\delta} \geq 0, \hat{\delta}^y + \hat{\delta}^u \|G_{uy}\| \leq \bar{\delta} < 1 \right\}, \quad \delta_0 = (0, 0, 0)^T.$$

Введем обозначения

$$(16) \quad \nu_{t+1} = |a(q^{-1})y_{t+1} - b(q^{-1})u_{t+1} - c^w|, \quad \phi_{t+1} = (1, p_{t+1}^y, p_{t+1}^u)^T$$

и запишем новое неравенство в описании D_{t+1} в виде

$$(17) \quad \delta \in \Omega_{t+1} = \left\{ \hat{\delta} \mid \nu_{t+1} \leq \hat{\delta}^T \phi_{t+1} \right\}.$$

Пусть P_t и δ_t – полиэдральная и векторная оценки δ в момент t . Положим

$$(18) \quad P_{t+1} = \begin{cases} P_t, & \text{если } \nu_{t+1} \leq \delta_t^T \phi_{t+1} + \varepsilon_1 |\phi_{t+1}|, \\ P_t \cap \Omega_{t+1}, & \text{в противном случае,} \end{cases}$$

$$(19) \quad \delta_{t+1} = \underset{\hat{\delta} \in P_{t+1}}{\operatorname{argmin}} J(c^w, \hat{\delta}).$$

Согласно (18), полиэдральная оценка P_{t+1} обновляется добавлением нового неравенства только если расстояние от δ_t до полупространства $\Omega_{t+1} \subset \mathbb{R}^3$ больше ε_1 . Заметим, что все множественные оценки P_t не ограничены в направлении роста переменной $\hat{\delta}^w$.

4.2. Оптимальное слежение при неизвестном смещении c^w

Для компенсации неизвестного смещения $c^w \in [c_{\min}^w, c_{\max}^w]$ оно будет оцениваться с помощью сетки

$$(20) \quad c_k^w = c_{\min}^w + k\varepsilon_2, \quad k = 0, 1, \dots, N, \quad \varepsilon_2 = \frac{c_{\max}^w - c_{\min}^w}{N} > 0,$$

гарантирующей оценку смещения c^w с желаемой точностью $\varepsilon_2/2$ за счет выбора достаточно большого N . Для каждого смещения c_k^w в каждый момент времени t будут вычисляться полиэдральные оценки $P_{k,t}$ и векторные оценки $\delta_{k,t}$ неизвестного вектора δ . Номер наилучшей оценки k_t вектора δ в момент времени t определим формулой

$$(21) \quad k_t = \underset{k}{\operatorname{argmin}} J(c_k^w, \delta_{k,t}).$$

Управление u_t в момент t определяется соответствующим этой оценке *адаптивным регулятором*

$$(22) \quad b(q^{-1})u_{t+1} = (a(q^{-1}) - 1)y_{t+1} + r_{t+1} - c_{k_t}^w.$$

После измерения выхода y_{t+1} невязки

$$\nu_{k,t+1} = |a(q^{-1})y_{t+1} - b(q^{-1})u_{t+1} - c_k^w|$$

и оценки $P_{k,t+1}$ и $\delta_{k,t+1}$ для $k = 0, 1, \dots, N$ вычисляются согласно (16)–(19) (соответствующие формулы с индексом k в каждой из них для краткости не приводятся).

Теорема 2. Пусть выполнены предположения П1–П5 и управление объектом (1) осуществляется адаптивным регулятором (22) с алгоритмом оценивания (16)–(19), (21) и параметром мертвой зоны ε_1 в (18), удовлетворяющим неравенству

$$(23) \quad 0 < \varepsilon_1 < \frac{1 - \bar{\delta}}{1 + \|G_{uy}\|}.$$

Тогда число обновлений полиэдральных оценок $P_{k,t}$ и векторных оценок $\delta_{k,t}$ конечно для всех $k \in \{0, 1, \dots, N\}$, и ошибка слежения удовлетворяет неравенству

$$(24) \quad \|y - r\|_{ss} \leq J(c_{k_\infty}^w, \delta_\infty + \varepsilon_1(1, 1, 1)^T) = J(c_{k_\infty}^w, \delta_\infty) + O(\varepsilon_1) \quad (\varepsilon_1 \rightarrow 0),$$

где k_∞ – финальное значение номера (21) наилучшей оценки неизвестного δ , δ_∞ – финальное значение $\delta_{k_\infty,t}$, и

$$(25) \quad \begin{aligned} J(c_{k_\infty}^w, \delta_\infty) &\leq J\left(c^w, \delta + \left(\frac{\varepsilon_2}{2} + \varepsilon_1, \varepsilon_1, \varepsilon_1\right)\right) = \\ &= J(c^w, \delta) + O(\varepsilon_1 + \varepsilon_2) \quad (\varepsilon_1 + \varepsilon_2 \rightarrow 0). \end{aligned}$$

Доказательство. При любом управлении u_t для любого $k \in \{0, 1, \dots, N\}$ из уравнения объекта (1) и представления суммарного возмущения v в виде (4) следует

$$(26) \quad |a(q^{-1})y_{t+1} - b(q^{-1})u_{t+1} - c_k^w| \leq |c^w - c_k^w| + \delta^w + \delta^y p_{t+1}^y + \delta^u p_{t+1}^u \quad \forall t.$$

В силу представления (4) неравенства (26) позволяют рассматривать управляемый объект как виртуальный объект вида (1), в котором виртуальное внешнее возмущение имеет смещение c_k^w и норма несмещенного внешнего возмущения не превосходит

$$(27) \quad \bar{\delta}_k^w = |c^w - c_k^w| + \delta^w.$$

Докажем, что число обновлений оценок $P_{k,t}$ и $\delta_{k,t}$ конечно для всех k . При каждом обновлении оценок согласно (18)

$$\varepsilon_1 |\phi_{t+1}| < \nu_{t+1} - \delta_t^T \phi_{t+1}.$$

Тогда для любого $\hat{\delta} \in \Omega_{t+1}$ из (17) имеем

$$\varepsilon_1 |\phi_{t+1}| < |(\hat{\delta} - \delta_t)^\top \phi_{t+1}| \leq |\hat{\delta} - \delta_t| |\phi_{t+1}|$$

и, следовательно, $|\hat{\delta} - \delta_t| > \varepsilon_1$. Тогда для всех k расстояние от оценки $\delta_{k,t}$ до полупространства $\Omega_{k,t+1}$ больше ε_1 . Так как $P_{k,t+1} \subset \Omega_{k,t+1}$, то расстояние от $\delta_{k,t}$ до $P_{k,t+1}$ также больше ε_1 . Полиэдральные оценки $P_{k,t}$ монотонно убывают по времени ввиду добавления новых неравенств. При этом шары радиуса $\varepsilon_1/2$ с центрами $\delta_{k,t}$ не пересекаются с аналогичными шарами с центрами в будущих обновленных оценках $\delta_{k,s}$ (при $s > t$) и, следовательно, со всеми шарами $\delta_{k,s}$ при $s \neq t$. Из этого следует конечность числа возможных обновлений оценок $\delta_{k,t}$ для всех k , так как все они находятся в соответствующих ограниченных множествах $\{\hat{\delta}_k \mid J(c_k^w, \hat{\delta}_k) \leq J(c_k^w, (\bar{\delta}_k^w, \delta^y, \delta^u)^\top)\}$, где $\bar{\delta}_k^w$ из (27).

Обозначим через $\delta_{k,\infty} = (\delta_{k,\infty}^w, \delta_{k,\infty}^y, \delta_{k,\infty}^u)^\top$ финальные значения, т.е. достигаемые за конечное время $t_{k,\infty}$, предельные значения оценок $\delta_{k,t}$ и положим $t_\infty = \max_k t_{k,\infty}$. Тогда $\delta_{k,t} = \delta_{k,\infty}$ для всех $t \geq t_\infty$ и всех k .

Пусть k_∞ – установившийся номер наилучшей оценки вектора δ :

$$k_\infty = \underset{k}{\operatorname{argmin}} J(c_k^w, \delta_{k,\infty}).$$

Тогда в силу (21)

$$(28) \quad J(c_{k_\infty}^w, \delta_{k_\infty}) \leq J(c_k^w, \delta_{k,\infty}) \quad \forall k.$$

Для всех $t \geq t_\infty$ для невязок (26) в замкнутой адаптивной системе с установившимся регулятором в силу (18) имеем

$$(29) \quad \nu_{k_\infty,t} \leq \delta_{k_\infty}^\top \phi_t + \varepsilon_1 |\phi_t| \leq (\delta_{k_\infty}^\top + \varepsilon_1(1, 1, 1)) \phi_t,$$

откуда по теореме 1 следует (24).

Введем обозначение для номера наиболее близкой к c^w оценки $c_{k_*}^w$:

$$k_* = \underset{k}{\operatorname{argmin}} |c^w - c_{k_*}^w|.$$

Тогда $|c^w - c_{k_*}^w| \leq \varepsilon_2/2$ и для всех $t \geq t_\infty$ с учетом (18)

$$(30) \quad \begin{aligned} |a(q^{-1})y_{t+1} - b(q^{-1})u_{t+1} - c_{k_*,t}^w| &\leq \frac{\varepsilon_2}{2} + \delta^\top \phi_{t+1} + \varepsilon_1 |\phi_t| \leq \\ &\leq \delta + \left(\frac{\varepsilon_2}{2} + \varepsilon_1, \varepsilon_1, \varepsilon_1\right)^\top \phi_{t+1}. \end{aligned}$$

Из этого неравенства по теореме 1 следует

$$(31) \quad J(c_{k_*}^w, \delta_{k_*,\infty}) \leq J\left(c^w, \delta + \left(\frac{\varepsilon_2}{2} + \varepsilon_1, \varepsilon_1, \varepsilon_1\right)\right).$$

Используя (28) при $k = k_*$ и (31) получаем неравенство (25). Наконец, слагаемое $O(\varepsilon_1 + \varepsilon_2)$ в (25) следует из того, что $J(c^w, \delta)$ является дробно-рациональной функцией δ и ее знаменатель отделен от 0 в силу предположения П5. Теорема 2 доказана.

Замечание 1. Неравенства (24) и (25) означают субоптимальность решения поставленной задачи слежения (13). Гарантированная неравенством (25) оценка $O(\varepsilon_1 + \varepsilon_2)$ точности решения поставленной оптимальной задачи является только качественной и ею нельзя воспользоваться для вычислений, поскольку c^w и δ неизвестны. Наилучшей вычислимой оценкой точности решения задачи оптимального слежения для конкретной реализации возмущений является текущая разность

$$(32) \quad (\Delta J)_t = J(c_{k_t}^w, \delta_{k_t, t} + \varepsilon_1(1, 1, 1)^T) - J(c_{k_t}^w, \delta_{k_t, t}),$$

которая согласована с данными измерений y_{1-n}^t, u_1^t и гарантирована после сходимости оценок за конечное время. Хотя время сходимости оценок к финальному значению неизвестно, длительный период неизменных оценок фактически гарантирует справедливость этой оценки в силу теоремы 2. Если текущая оценка точности решения неудовлетворительна, можно в любой момент уменьшить выбранный параметр мертвой зоны ε_1 для улучшения точности. При этом число обновлений оценок $P_{k,t}$ и $\delta_{k,t}$ может возрасти. Влияние шага сетки ε_2 на точность оптимизации более прозрачно (слагаемое $\varepsilon_2/2$ добавляется к оценкам $\delta_{k,t}^w$) и его можно выбирать априори, учитывая при этом, что уменьшение шага сетки ε_2 влечет увеличение числа параллельно вычисляемых полиэдральных оценок $P_{k,t}$ и векторных оценок $\delta_{k,t}$.

5. Моделирование

Пусть объект управления описывается уравнением (1) с неизвестными параметрами номинальной модели

$$(33) \quad \theta^* = [a_1^*, a_2^*, b_1^*, b_2^*, b_3^*] = [-2,7; 1,8; 2; -3,36; 1,4],$$

а для тестирования дана номинальная модель с полюсами (корнями $a(\lambda)$) 0,7 и 0,8, нулями (корнями $b(\lambda)$) 1,1 и 1,3 и коэффициентом $b_1 = 2$. Этой номинальной модели соответствует неустойчивый минимально-фазовый объект (1) с параметрами

$$(34) \quad \theta = [a_1, a_2, b_1, b_2, b_3] = [-2,6786; 1,7857; 2; 3,3566; 1,3986],$$

немного отличающимися от параметров (33). Пусть управление объектом с параметрами (33) осуществляется адаптивным регулятором (22), оптимальным для тестируемого объекта с параметрами (34). Корни характеристического полинома этой замкнутой системы с точностью 10^{-4} равны $0,7512 \pm 8,9242i, 1,3032, 1,0945$, по модулю больше 1, так что замкнутая система без операторных возмущений устойчива. Динамика этой замкнутой системы может рассматриваться как динамика системы с объектом с номинальными параметрами θ , в котором присутствуют дополнительные относительно небольшие операторные возмущения

$$\Delta(y_{t-2}^{t-1}, u_{t-2}^{t-1}) = (a_1 - a_1^*)y_{t-1} + (a_2 - a_2^*)y_{t-2} + (b_1^* - b_1)u_{t-1} + (b_1^* - b_1)u_{t-2},$$

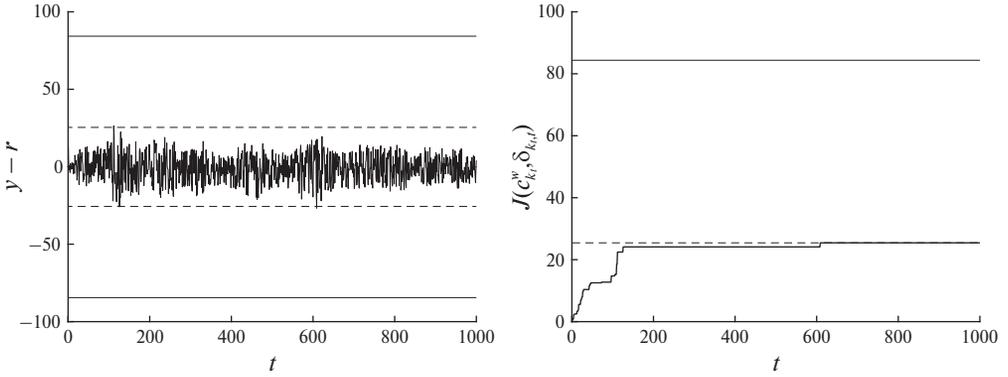


Рис. 1. Графики $y - r$ (слева) и $J(c_{k_t}^w, \delta_{k_t, t})$ (справа); $\pm J(c^w, \delta)$ – сплошные прямые, $\pm J(c_{k_\infty}^w, \delta_{k_\infty})$ – пунктирные прямые.

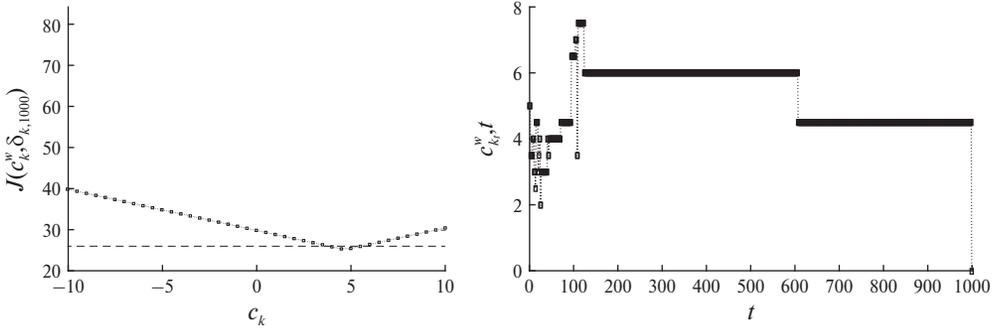


Рис. 2. Слева – значения $J(c_k^w, \delta_{k, 1000})$, справа – переключения оценок $c_{k_t, t}^w$.

проистекающие из “неточных” коэффициентов тестируемой номинальной модели. Возмущение v_t в номинальной модели с параметрами θ моделируется в виде

$$(35) \quad v_t = c^w + \delta^w w_t + k_t^y \delta^y |y_{t-\mu}^{t-1}| + k_t^u \delta^u |u_{t-\mu}^{t-1}|, \quad |k_t^y| \leq 1, \quad |k_t^u| \leq 1, \quad \mu = 20.$$

Пусть сигнал слежения имеет вид $r_t = 10 \sin t$ для всех t .

Пример 1. Случайные возмущения. Пусть неизвестные параметры в описании (35) имеют значения

$$(36) \quad c^w = 5, \quad \delta^w = 1, \quad \delta^y = \delta^u = 0,1,$$

и w_t, k_t^y, k_t^u – независимые равномерно распределенные на $[-1, 1]$ псевдослучайные величины. Моделирование проводилось при следующих параметрах адаптивного управления: параметр мертвой зоны $\varepsilon_1 = 10^{-6}$, $c_{\min}^w = -10$, $c_{\max}^w = 10$, шаг сетки $\varepsilon_2 = 0,5$.

На рис. 1 представлены графики ошибки слежения $y - r$ (слева) и текущих оптимальных оценок качества $J(c_{k_t}^w, \delta_{k_t, t})$ (справа).

На рис. 2 слева представлены согласованные с измерениями на промежутке $[1, 1000]$ финальные значения $J(c_k^w, \delta_{k,1000})$ для всех k . Переключения оценок $c_{k,t}^w$ неизвестного смещения $c^w = 5$ представлены на рис. 2 справа. Заметим, что несмотря на симметричность распределений случайных величин w_t, k_t^y, k_t^u относительно нуля, установившаяся оценка смещения $c_{1000}^w = 4,5$ отлична от $c^w = 5$.

Во всех численных экспериментах со случайными возмущениями и детерминированными “колебательными” возмущениями установившиеся согласованные с измерениями верхние границы ошибки слежения $J(c_{k_\infty}^w, \delta_{k_\infty})$ значительно (в разы) меньше неизвестной оптимальной верхней границы $J(c^w, \delta)$. При этом при квантификации возмущений никак не проявляют себя операторные возмущения, поскольку текущие оценки δ_t неизвестного вектора δ , как правило, имеют вид $\delta_t = (c_t^w, 0, 0)$.

Замечание 2. Сторонниками стохастических моделей возмущений в теории идентификации систем множественное оценивание неизвестных параметров (set-membership approach) постоянно критикуется за кажущийся неизбежным консерватизм в связи с необходимой априорной информацией о верхних границах детерминированных возмущений (при этом подразумевается лишь консерватизм множественных оценок неизвестных параметров). Пример 1 иллюстрирует, что использование множественного оценивания и показателя качества задачи управления как идентификационного критерия гарантирует не просто неконсервативность согласованных с измерениями гарантий качества, но даже улучшенные гарантии качества по сравнению с оптимальным показателем качества (5), поскольку конкретные реализации возмущений в основной массе далеки от возмущений, максимизирующих ошибку слежения. Это аналогично тому, как в стохастическом случае усредненные показатели качества лучше, чем возможные наихудшие значения на конкретных “плохих” реализациях. При этом, однако, в задачах со стохастическими возмущениями задачи верификации модели возмущений, как правило, не обсуждаются. Оптимальность слежения в рамках детерминированной ℓ_1 -теории базируется на верификации модели возмущений и использовании достаточно полной получаемой в процессе управления информации о неизвестных параметрах, и платой за оптимальность служит соответствующий рост объема необходимых вычислений.

Пример 2. “Плохие” детерминированные возмущения. Цель этого примера – продемонстрировать “плохое” суммарное возмущение, при котором наглядно проявляется наличие операторных возмущений в суммарном возмущении v .

Рассмотрим управляемый объект (33) с суммарным возмущением вида (35) и параметрами (36) с уменьшенным $\delta^u = 0,05$ и детерминированными последовательностями

$$(37) \quad w_t = \cos(50t), \quad k_t^y = \sin(70t), \quad k_t^u = \cos(\ln(0,5t)).$$

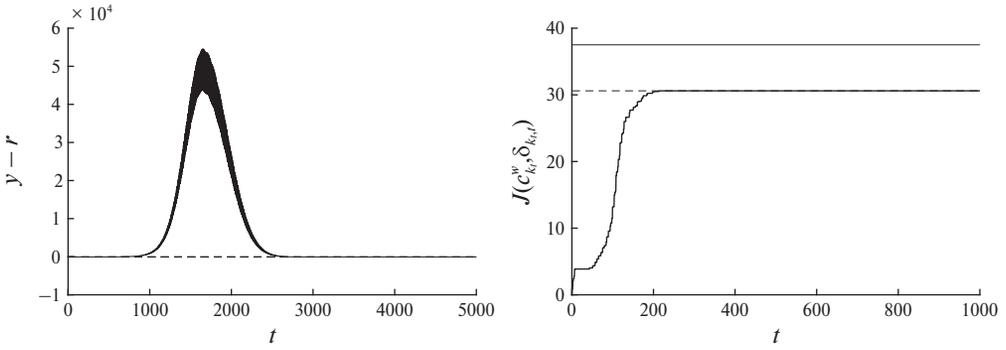


Рис. 3. Графики $y - r$ (слева) и $J(c_{k_t}^w, \delta_{k_t, t})$ (справа); $J(c^w, \delta)$ – сплошная прямая, $J(c_{k_\infty}^w, \delta_{k_\infty})$ – пунктирная прямая.

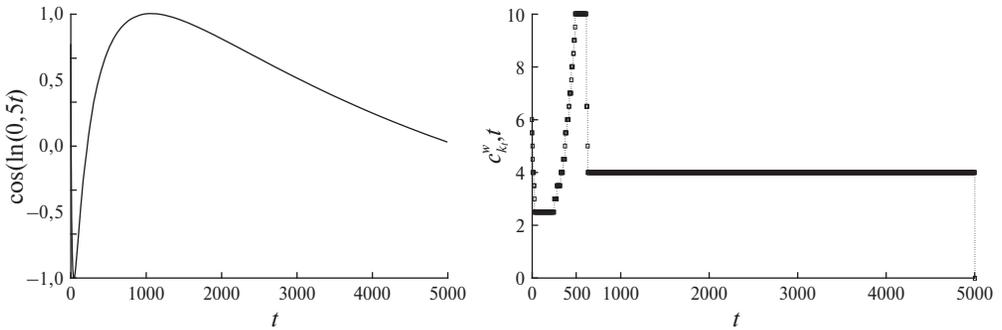


Рис. 4. Графики k_t^y и переключений оценок неизвестного смещения $c^w = 5$.

На левом графике рис. 3 представлена ошибка слежения $y - r$. При этом возмущении для всех $t \geq 498$ оценки $\delta_{k_t, t}^u > 0$, и $\delta_{k_t, t}^y = 0$, и последние оценки таковы: $\delta_{k_{5000}, 5000} = (1,6993; 0; 0,0581)$, $c_{5000}^w = 4$. Таким образом, с момента $t = 498$ в оценках $\delta_{k_t, t}$ норм возмущений проявляется операторное возмущение по управлению.

Известно, что в устойчивых линейных стационарных системах возможны большие отклонения от нуля, вызванные ненулевыми начальными данными или ограниченным возмущением [16, 17]. В примере 2 большое отклонение ошибки слежения (с $\max_t |y_t - r_t| = 5,4573 \cdot 10^4$) в нелинейной замкнутой системе (1)–(6) может быть вызвано как переключением соответствующих разным оценкам смещения регуляторов, так и “несимметричностью” относительно нуля последовательности $k_t^y = \cos(\ln(0,5t))$, представленной на левом графике рис. 4. График текущих оптимальных оценок $c_{k_t, t}^w$ неизвестного смещения c^w представлен на правом графике рис. 4, где установившаяся оценка смещения $c_{k_\infty}^w = 4 \neq c^w = 5$.

Замечание 3. Следует заметить, что несмотря на неприемлемые с точки зрения практических приложений значения ошибки слежения порядка 10^4 в переходном режиме, ее асимптотическое поведение при данном возмущении

характеризуется числами

$$\max_{t \in [4001, 5000]} |y_t - r_t| = 4,4163, \quad \max_{t \in [4901, 5000]} |y_t - r_t| = 2,7566.$$

Таким образом, фактическая установившаяся ошибка слежения на порядок меньше, чем финальная согласованная с данными измерений гарантируемая оценка ошибки слежения $J(c_{k_\infty}^w, \delta_{k_\infty}) = 30,4421$. Эта оценка, в свою очередь, лучше гарантируемого теоремой 1 оптимального (но неизвестного!) значения $J(c^w, \delta) = 37,2971$, несмотря на невысокую выбранную “точность” (размер сетки $\varepsilon_2 = 0,5$) оценки смещения. Наконец, само оптимальное значение $J(c^w, \delta)$ меньше наихудшей возможной асимптотической ошибки слежения, поскольку в $J(c^w, \delta)$ не учтено, что тестируемый объект управления имеет номинальные параметры θ^* , а не θ . В итоге алгоритм адаптивной компенсации неизвестного смещения c^w выполняет свое предназначение, несмотря на возможные большие отклонения ошибки слежения от нуля, и даже “подстраивается” к конкретным реализациям возмущения v , уменьшая излишний консерватизм в гарантируемых оценках качества при не “максимизирующих” возмущениях.

Замечание 4. Приведенные графики переключений оценок $c_{k_t}^w$ неизвестного смещения c^w иллюстрируют неидентифицируемость смещения c^w в описании (4) даже при отсутствии операторных возмущений, поскольку управление всегда происходит при конкретных реализациях возмущения v_t , для которых смещения (при любом их разумном определении) могут быть (правильнее сказать, будут) разными. То есть термин «смещение» по отношению к постоянной c^w в описании (4) относится именно к понятию смещения для класса всех возмущений, удовлетворяющих (4). В то же время текущие оценки $c_{k_t}^w$ можно (или, скорее, следует) считать корректным определением (в контексте решаемой задачи управления) текущих оценок смещения для конкретной реализации суммарного возмущения v .

Замечание 5. Объем и скорость вычислений в описанных примерах можно охарактеризовать следующими показателями. Время вычислений на ноутбуке с 15,2 ГиБ ОЗУ и процессором 18 Intel Core Ultra 5 125Н в примере 1 равно 2,99 с, в примере 2 – 15,1169 с. Число неравенств в полиэдральных оценках $P_{k,t}$ находится в интервале: 12–15 в примере 1 и в интервале 64–81 – в примере 2. Отношение этих показателей примерно соответствует отношению длин интервалов, равному 5. При этом показатели примера 2 на интервале времени $[1, 10\,000]$ остаются теми же, что означает, что переходные процессы для конкретного рассмотренного возмущения v к моменту времени $t = 5000$ уже завершились. Заметим, что время вычислений определяется в основном временем вычисления полиэдральных оценок $P_{k,t}$ и оптимальных векторных оценок $\delta_{k,t}$ в \mathbb{R}^3 и практически не зависит от размерности вектора номинальных параметров θ .

Замечание 6. Число неравенств в описании полиэдральных оценок $P_{k,t}$ и, как следствие, время вычисления оптимальных оценок в (21) можно сокра-

титель исключением возможных лишних неравенств после добавления новых неравенств (17) [18]. В приведенных результатах моделирования это не делалось, чтобы продемонстрировать число возможных обновлений оценок даже при очень маленьком значении (почти нулевом с точки зрения оценки качества модели) параметра мертвой зоны ε_1 .

6. Заключение

В статье рассмотрена задача оптимального, с заданной точностью, отслеживания заданного ограниченного сигнала для дискретного минимально-фазового объекта с известной или заданной для тестирования номинальной моделью, смещенным ограниченным внешним возмущением и операторными возмущениями с неизвестными нормами и смещением. Сложность задачи заключается в необходимости компенсации смещения на основе обоснованной оптимальной оценки качества управления в условиях неидентифицируемости всех неизвестных параметров. Решение задачи базируется на квантификации возмущений, множественном оценивании неизвестных параметров и использовании показателя качества управления как идеального идентификационного критерия. Такой подход позволяет глубже понять, продемонстрировать и реализовать максимальные возможности обратной связи, важность исследований которых подчеркивалась одним из ведущих экспертов по адаптивному управлению и идентификации систем L. Guo в аннотации к [19]:

“Основной целью адаптивной обратной связи является работа с динамическими системами с внутренними и/или внешними неопределенностями, используя доступную онлайн информацию. Таким образом, фундаментальной проблемой адаптивного управления является понимание максимальных возможностей (и ограничений) адаптивной обратной связи.”

и в аннотации к [20]:

“Наконец, мы рассмотрим более фундаментальные проблемы, связанные с максимальными возможностями и ограничениями механизма обратной связи при работе с неопределенными нелинейными системами, где механизм обратной связи определяется как класс всех возможных законов обратной связи.”

Полученное в статье решение не только заведомо гарантирует, с заданной точностью, ту же оценку качества слежения, что и при известных параметрах номинального объекта и возмущений, но дополнительно дает существенно лучшие гарантированные оценки качества в зависимости от конкретных реализаций детерминированных возмущений. Тем самым неявным образом учитывается то, что конкретные реализации возмущений, как правило, далеки от специальных реализаций возмущений, максимизирующих показатель качества управления, поскольку для максимизации оценки ошибки слежения суммарное возмущение v_t должно не только принимать максимальные зна-

чения на протяжении длительного интервала времени, но и иметь при этом определенные знаки на этом интервале.

СПИСОК ЛИТЕРАТУРЫ

1. *Khammash M., Pearson J.B.* Performance robustness of discrete-time systems with structured uncertainty // *IEEE Trans. Automat. Control.* 1991. V. AC-36. No. 4. P. 398–412.
2. *Khammash M., Pearson J.B.* Analysis and design for robust performance with structured uncertainty // *Syst. Control Lett.* 1993. V. 20. No. 3. P. 179–187.
3. *Kosut R., Goodwin G., Polis M.* (Eds.) Special issue on system identification for robust control design // *IEEE Transactions Automatic Control.* 1992. V. 37. No. 7. P. 900–1008.
4. *Soderstrom T., Astrom K.* (Eds.) Trends in System Identification (Special issue on system identification // *Automatica.* 1995. V. 31. No. 12. P. 1689–1907.
5. *Gevers M.* A personal View of the Development of System Identification // *IEEE Control Syst. Magazine.* 2006. V. 26. No. 6. P. 93–105.
6. *FLamnabhi-Lagarigue F., Annaswamy A., Engell S., et. al.* Systems & Control for the future of humanity, research agenda: Current and future roles, impact and grand challenges // *Annual Reviews in Control.* 2017. V. 43. P. 1–64.
7. *Ljung L.* Revisiting Total Model Errors and Model Validation // *J. Syst. Sci. Complex.* 2021. V. 34. P. 1598–1603.
8. *Khammash M.H.* Robust steady-state tracking // *IEEE Trans. Automat. Control.* 1995. V. 40. No. 11. P. 1872–1880.
9. *Khammash M.H.* Robust Performance: Unknown Disturbances and Known Fixed Inputs // *IEEE Trans. Automat. Control.* 1997. V. 42. No. 12. P. 1730–1734.
10. *Соколов В.Ф.* Асимптотическое робастное качество дискретной системы слежения в ℓ_1 -метрике // *АиТ.* 1999. № 1. С. 101–112.
11. *Соколов В.Ф.* Робастное управление при ограниченных возмущениях. Сыктывкар: Коми научный центр УрО РАН. 2011. 218 с.
12. *Sokolov V.F.* Control-Oriented Model Validation and Errors Quantification in the ℓ_1 Setup // *IEEE Trans. Autom. Control.* 2005. V. 50. No. 10. P. 1501–1509.
13. *Соколов В.Ф.* Задачи адаптивного оптимального управления дискретными системами с ограниченным возмущением и линейными показателями качества // *АиТ.* 2018. № 6. С. 155–171.
14. *Sokolov V.F.* Optimal Error Quantification and Robust Tracking under Unknown Upper Bounds on Uncertainties and Biased External Disturbance // *Mathematics.* 2024. V. 12. № 2. P. 197.
15. *Boyd S., Vandenberghe L.* Convex Optimization. Cambridge University Press. 2003.
16. *Polyak D.T., Shcherbakov P.S., Smirnov G.* Peak effects in stable linear difference equations // *J. Difference Equat. Appl.* 2018. V. 24. No. 9. P. 1488–1502.
17. *Поляк Б.Т., Тремба А.А., Хлебников М.В., Щербakov П.С., Смирнов Г.В.* Большие отклонения в линейных системах при ненулевых начальных условиях // *АиТ.* 2015. № 6. С. 18–41.

18. *Walter E., Piet-Lahanier H.* Exact recursive polyhedral description of the feasible parameter set for bounded error // IEEE Transactions on Automatic Control. 1989. V. 34. P. 911–915.
19. *Guo L.* Exploring the maximum capability of adaptive feedback // Int. J. Adapt. Control Signal Process. 2002. V. 16. N. 5. P. 341–354.
20. *Guo L.* Feedback and uncertainty: Some basic problems and results // Annual Reviews in Control. 2020. V. 49 P. 27–36.

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

Поступила в редакцию 03.03.2025

После доработки 20.05.2025

Принята к публикации 27.06.2025

© 2025 г. Р. ХИЛЬДЕБРАНД, д-р физ.-мат. наук (khildebrand.r@mipt.ru),
Т. ШИКАКЕ МАПУНГВАНА (tendaichikake@phystech.edu)
(Московский физико-технический институт, Долгопрудный)

О ПРОБЛЕМЕ ОПТИМАЛЬНОГО УПРАВЛЕНИЯ С УПРАВЛЕНИЕМ В ДИСКЕ¹

Рассматривается задача быстрогодействия с симметрией типа Фуллера и управлением в двухмерном диске единичного радиуса. Задача обладает аналитическим решением, с неявным представлением функции Беллмана. Оптимальное значение этой задачи служит в качестве верхней границы на оптимальное значение другой задачи оптимального управления с симметрией типа Фуллера, содержащей в оптимальном синтезе особый режим второго порядка и решаемой аналитически.

Ключевые слова: оптимальное управление, симметрия Фуллера, особый режим, задача быстрогодействия.

DOI: 10.31857/S0005231025080052, EDN: UTDQEM

1. Введение

Задачи оптимального управления обычно решаются с помощью принципа максимума Понтрягина (ПМП), который приводит к гамильтоновой системе с разрывной правой частью [1]. Так как вследствие этого правая часть не удовлетворяет условию Липшица, теорема Пикара о существовании и единственности решения системы обыкновенных дифференциальных уравнений (ОДУ) не применима, и фазовый портрет системы может содержать разного рода сингулярности. Эти сингулярности проявляют себя в оптимальном синтезе в виде особых режимов. Чтобы уметь строить оптимальный синтез, нужно владеть информацией о структуре особых режимов и порождающих их сингулярностей. Одним из приемов исследования этих объектов является решение *модельной системы*, которая, с одной стороны, достаточно простая, в лучшем случае решается аналитически, а с другой стороны, структурно устойчива и содержит в себе сингулярность исследуемого типа.

Феномен особых режимов в оптимальном управлении был представлен в [2], где приведен пример системы, в которой оптимальное управление претерпевает бесконечное число переключений из одного экстремального значения в другое за конечный отрезок времени. Это явление называют *четтерингом*. Особые режимы были систематически исследованы, в том числе в [3–6]. Феномен четтеринга детально исследован в [7, 8]. Более сложная система, в которой четтеринг присутствует в комбинации с фрактальной структурой

¹ Исследование выполнено при поддержке Министерства науки и высшего образования Российской Федерации (госзадание), номер проекта FSMG-2024-0011.

поведения оптимального управления, переключающегося хаотичным образом между тремя экстремальными значениями, была исследована в [9, 10].

В данной работе исследуются особые режимы второго порядка в продолжение программы, инициированной в [7]. Впервые задача оптимального управления с особым режимом второго порядка была аналитически решена в [2]. Она имеет вид

$$(1) \quad \min \int_0^{\infty} \frac{x^2}{2} dt : \quad \dot{x} = y, \quad \dot{y} = u, \quad u \in [-1, 1].$$

Особая траектория второго порядка, присутствующая в фазовом портрете гамильтоновой системы, есть траектория $x(t) = y(t) \equiv 0$. Стыковка неособой траектории оптимального синтеза задачи (1) с особой траекторией происходит за конечное время с бесконечным числом переключений управления u между экстремальными значениями ± 1 . Функция Беллмана задачи (1) была посчитана явно в [9] и задается по формуле

$$(2) \quad \omega_{1D}(x, y) = \begin{cases} -\frac{1}{2}x^2y - \frac{1}{3}xy^3 - \frac{1}{15}y^5 - \gamma(y^2 + 2x)^{\frac{5}{2}}, & x \geq -\beta y|y|; \\ \frac{1}{2}x^2y - \frac{1}{3}xy^3 + \frac{1}{15}y^5 - \gamma(y^2 - 2x)^{\frac{5}{2}}, & x \leq -\beta y|y|, \end{cases}$$

где $\beta \approx 0,4446$ есть корень алгебраического уравнения

$$36\beta^4 + 3\beta^2 - 2 = 0,$$

а $\gamma = \frac{-\beta^2 + 2\beta - \frac{2}{3}}{10(1-2\beta)^{\frac{3}{2}}} \approx 0,06753$. Напомним, что выражение $-\omega_{1D}(x_0, y_0)$ является оптимальным значением функционала в задаче с начальными условиями $x(0) = x_0, y(0) = y_0$.

В оптимальном синтезе задачи присутствует непрерывная группа симметрий. А именно, если $(x(t), y(t), u(t))$ является оптимальной траекторией в задаче (1), то для любого $\lambda > 0$ траектория

$$(3) \quad (x_\lambda(t), y_\lambda(t), u_\lambda(t)) = (\lambda^2 x(\lambda^{-1}t), \lambda y(\lambda^{-1}t), u(\lambda^{-1}t))$$

также является оптимальной [7]. Также для любого $\lambda > 0$ функция Беллмана удовлетворяет условию

$$(4) \quad \omega_{1D}(\lambda^2 x, \lambda y) = \lambda^5 \omega_{1D}(x, y).$$

В монографии [7] рассмотрен аналог задачи (1) с двухмерным управлением. Эта задача имеет вид

$$(5) \quad \min \int_0^{\infty} \frac{\|x\|^2}{2} dt : \quad \dot{x} = y, \quad \dot{y} = u, \quad u \in U = \mathbb{D},$$

где $\mathbb{D} = \{u \in \mathbb{R}^2 \mid \|u\| \leq 1\}$ – диск единичного радиуса, а $x(t), y(t)$ – векторно-значные функции. Оптимальный синтез этой задачи также содержит особую траекторию второго порядка $(x, y) = (0, 0)$ в пространстве $\mathbb{R}^2 \times \mathbb{R}^2$. Задача обладает той же симметрией (3), но вдобавок еще вращательной симметрией. А именно, для любой оптимальной траектории $(x(t), y(t), u(t))$ задачи (5) и любой ортогональной матрицы $O \in O(2)$ траектория

$$(6) \quad (Ox(t), Oy(t), Ou(t))$$

также является оптимальной. Функция Беллмана ω_{2D} в задаче (5) обладает симметриями

$$\omega_{2D}(\lambda^2 x, \lambda y) = \lambda^5 \omega_{2D}(x, y), \quad \omega_{2D}(Ox, Oy) = \omega_{2D}(x, y)$$

для любого $\lambda > 0$ и любой матрицы $O \in O(2)$. Вращательная симметрия позволяет рассматривать функцию Беллмана как функцию от грамиана $\begin{pmatrix} \langle x, x \rangle & \langle x, y \rangle \\ \langle x, y \rangle & \langle y, y \rangle \end{pmatrix}$ векторов x, y , что эффективно сокращает размерность на 1. Значение в момент t этой 2×2 матрицы на оптимальной траектории зависит только от начального значения грамиана.

В [7, Предложение 7.8] было показано, что в случае линейно зависимых начальных значений $x(0), y(0)$ задача (5) сводится к одномерной задаче (1), и векторы $x(t), y(t)$ все время остаются в одномерном подпространстве, натянутом на начальные значения. В частности, для линейно зависимых векторов $x = r_x(\cos \varphi, \sin \varphi)^T$, $y = r_y(\cos \varphi, \sin \varphi)^T$ функция Беллмана удовлетворяет условию

$$\omega_{2D}(x, y) = \omega_{1D}(r_x, r_y),$$

где числа r_x, r_y принимают любые вещественные значения.

Кроме оптимальных траекторий, исходящих из линейно зависимых начальных векторов, в задаче (5) существует семейство *автомодельных* оптимальных траекторий, полученное в [7, Предложение 7.9, Следствие 7.3]. На этих траекториях тангенс угла между векторами x и y не зависит от времени и равен $-\sqrt{5}/2$ (значения $\sqrt{5}$ и $\sqrt{5}/2$ в [7, стр. 233] оба ошибочны), а тангенс угла между векторами y и u также постоянен и равен $-\sqrt{5}$. Более того, справедливо равенство $2\langle y, y \rangle = \sqrt{6\langle x, x \rangle}$. Из этого следует, что грамиан начальных условий $x(0), y(0)$ на этих автомодельных траекториях задан выражением

$$(7) \quad \begin{pmatrix} \langle x(0), x(0) \rangle & \langle x(0), y(0) \rangle \\ \langle x(0), y(0) \rangle & \langle y(0), y(0) \rangle \end{pmatrix} = \begin{pmatrix} \frac{\lambda_0^4}{54} & -\frac{\lambda_0^3}{27} \\ -\frac{\lambda_0^3}{27} & \frac{\lambda_0^2}{6} \end{pmatrix}$$

для некоторого $\lambda_0 > 0$. Грамиан на траектории эволюционирует согласно закону

$$(8) \quad \begin{pmatrix} \langle x(t), x(t) \rangle & \langle x(t), y(t) \rangle \\ \langle x(t), y(t) \rangle & \langle y(t), y(t) \rangle \end{pmatrix} = \begin{pmatrix} \frac{\lambda(t)^4}{54} & -\frac{\lambda(t)^3}{27} \\ -\frac{\lambda(t)^3}{27} & \frac{\lambda(t)^2}{6} \end{pmatrix}, \quad \lambda(t) = \lambda_0 - t.$$

Из этого следует интерпретация параметра λ_0 как времени прибытия траектории в особый режим в начале координат $(x, y) = (0, 0)$.

Если углы между векторами x, y, u постоянны, то сами векторы все быстрее оборачиваются вокруг начала координат, делая бесконечное число оборотов за конечное время. Более детально, направление каждого из этих векторов задано зависящим от времени фазовым углом [7, Предложение 7.9, Следствие 7.3]

$$(9) \quad \varphi(t) = \sigma\sqrt{5} \log(\lambda_0 - t) + \text{const},$$

где $\sigma \in \{-1, +1\}$ определяет направление вращения вокруг начала координат, а аддитивная константа зависит от фаз начальных условий.

Полная картина оптимального синтеза в задаче (5) на данный момент неизвестна. В настоящей работе вычисляется верхняя оценка на оптимальное значение задачи, с предъявлением явного субоптимального решения. Эта оценка должна удовлетворять нескольким критериям, потенциально противоречащим друг другу:

- оценка должна быть достаточно близка к истинному значению;
- оценку нужно уметь численно эффективно использовать, например, если она задается единым аналитическим выражением.

Как стало понятным из предшествующего анализа, оптимальные решения для разных начальных значений могут сильно отличаться друг от друга. Построение оценки, которая, с одной стороны, всюду близка к истинному значению, а с другой стороны, не задается мозаикой разных выражений в разных областях фазового пространства, является непростой задачей.

Предлагаемый подход состоит в решении *задачи быстрого действия*

$$(10) \quad \begin{aligned} \min (T - t_0): \quad & \dot{x} = y, \quad \dot{y} = u, \quad u \in U = \mathbb{D}, \\ & x(T) = y(T) = 0, \quad x(t_0) = x_0, \quad y(t_0) = y_0, \end{aligned}$$

которая с точностью до смещения временной переменной t имеет то же множество допустимых траекторий, что и задача (5), но отличается от последней функционалом цены. Задача быстрого действия решается в разделе 2. В разделе 3 полученное решение подставляется в функционал цены исходной задачи (5), и получается искомую верхнюю оценку. В разделе 4 верхняя оценка сравнивается с истинным оптимальным значением в задаче (5) на тех траекториях, для которых последнее известно. Этот анализ показывает, что, с

одной стороны, качество оценки достаточно хорошее на всех таких начальных значениях, с другой стороны, оценка задается единым аналитическим выражением.

2. Задача быстродействия

В этом разделе явно решается задача быстродействия (10).

Применим ПМП. Введем сопряженные переменные ϕ, ψ и составим функцию Понтрягина

$$(11) \quad \mathcal{H} = -1 + \langle \phi, y \rangle + \langle \psi, u \rangle.$$

Оптимальное управление в случае $\psi \neq 0$ задается выражением $\hat{u} = \arg \max_{u \in U} \mathcal{H} = \frac{\psi}{\|\psi\|}$. Динамика сопряженных переменных следует закону

$$\dot{\psi} = -\frac{\partial \mathcal{H}}{\partial y} = -\phi, \quad \dot{\phi} = -\frac{\partial \mathcal{H}}{\partial x} = 0.$$

Так как конечное время T не фиксировано, имеем также условие трансверсальности

$$(12) \quad \mathcal{H}(T) = \langle \psi, \hat{u} \rangle - 1 = \|\psi\| - 1 = 0.$$

Отсюда следует, что функция $\psi(t) = \phi t + \psi(0)$ аффинная и в конечный момент времени $t = T$ приходит на окружность единичного радиуса.

В силу вращательной симметрии можно без ограничения общности полагать, что вектор $\dot{\psi} = -\phi = (\alpha, 0)^\top$ параллелен единичному базисному вектору $e_1 = (1, 0)^\top$ и $\alpha \geq 0$. В случае $\alpha > 0$ выберем начальный момент времени t_0 таким образом, чтобы значение $\psi(0) = (0, \beta)^\top$ было коллинеарно базисному вектору e_2 и $\beta \geq 0$.

Случай $\alpha\beta = 0$: В этом случае сопряженная переменная ψ пребывает в одномерном линейном подпространстве пространства \mathbb{R}^2 . Тогда и векторы u, y, x должны эволюционировать в этом подпространстве, и задача сводится к известной одномерной задаче быстродействия с ускорением, ограниченным по модулю единицей. Эта задача имеет также вид (10), но все переменные рассматриваются как скаляры.

В этой задаче сопряженная переменная ϕ является постоянным скаляром. Сдвинем переменную времени таким образом, что конечный момент времени T равен нулю. Тогда переменная ψ в конечный момент времени принимает значение $\psi(T) = \psi(0) = \sigma \in \{-1, +1\}$ и как функция от времени задается выражением $\psi(t) = \sigma - \phi t$. Оптимальное управление u кусочно-постоянно и имеет вид

$$\hat{u}(t) = \begin{cases} +1, & \text{если } \sigma > \phi t, \\ -1, & \text{если } \sigma < \phi t. \end{cases}$$

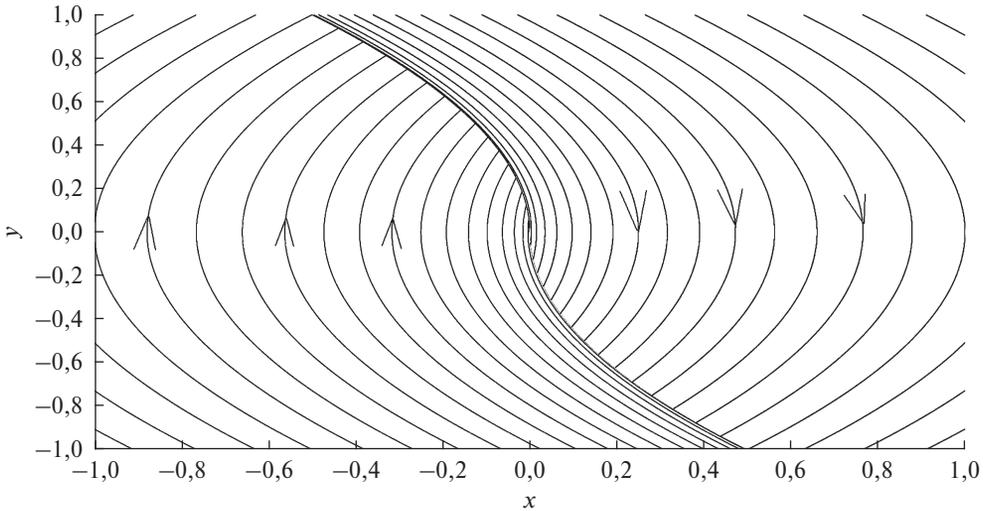


Рис. 1. Оптимальный синтез в одномерной задаче быстродействия.

Эквивалентно, полагая $\tilde{\phi} = \sigma\phi$, получаем

$$\hat{u}(t) = \begin{cases} +\sigma, & \text{если } 1 > \tilde{\phi}t, \\ -\sigma, & \text{если } 1 < \tilde{\phi}t. \end{cases}$$

Отсюда вытекает

$$y(t) = \int_0^t \hat{u}(s) ds = \begin{cases} \sigma t, & \text{если } 1 > \tilde{\phi}t, \\ -\sigma(t - \tilde{\phi}^{-1}) + \sigma\tilde{\phi}^{-1}, & \text{если } 1 < \tilde{\phi}t \end{cases}$$

и далее

$$x(t) = \int_0^t y(s) ds = \begin{cases} \sigma \frac{t^2}{2}, & \text{если } 1 > \tilde{\phi}t, \\ -\sigma \frac{t^2 - \tilde{\phi}^{-2}}{2} + 2\sigma\tilde{\phi}^{-1}(t - \tilde{\phi}^{-1}) + \sigma \frac{\tilde{\phi}^{-2}}{2}, & \text{если } 1 < \tilde{\phi}t. \end{cases}$$

Случай $1 > \tilde{\phi}t$ возможен, только если $x = \sigma \frac{y^2}{2} = -\frac{y|y|}{2}$, и тогда $t_0 = \sigma y_0$. Во всех других точках на плоскости переменных x, y имеем $1 < \tilde{\phi}t$. Отсюда следует, что множество решений уравнения $x + \frac{y|y|}{2} = 0$ разделяет плоскость переменных x, y на две области, в которых оптимальны два крайних значения управления. Разделяющая области кривая состоит из траекторий, которые приходят непосредственно в начало координат (см. рис. 1). Слева от разделяющей кривой имеем $\hat{u} = +1$ и $\sigma = -1$, справа $\hat{u} = -1$ и $\sigma = 1$. Траектории системы переключаются на противоположное значение управления, когда они достигают сепарирующую кривую.

В частности, в случае $1 < \tilde{\phi}t$ имеем

$$\hat{u} = -\sigma = -\text{sign} \left(x_0 + \frac{y_0|y_0|}{2} \right).$$

Далее, имеем $\sigma y_0 = -t_0 + 2\tilde{\phi}^{-1}$, $\sigma x_0 = -\frac{t_0^2}{2} + 2\tilde{\phi}^{-1}t_0 - \tilde{\phi}^{-2}$ и вследствие этого

$$\tilde{\phi}^{-1} = -\sqrt{\sigma x_0 + \frac{y_0^2}{2}}, \quad t_0 = -\sigma y_0 - 2\sqrt{\sigma x_0 + \frac{y_0^2}{2}}.$$

При $t_0 < t < -\sqrt{\sigma x_0 + \frac{y_0^2}{2}}$ окончательно получаем

$$y(t) = -\sigma t - 2\sigma\sqrt{\sigma x_0 + \frac{y_0^2}{2}}, \quad x(t) = -\frac{\sigma t^2}{2} - 2\sigma\sqrt{\sigma x_0 + \frac{y_0^2}{2}}t - x_0 - \frac{\sigma y_0^2}{2},$$

где $\sigma = \text{sign} \left(x_0 + \frac{y_0|y_0|}{2} \right)$.

Случай $\alpha > 0$, $\beta > 0$: В этом случае имеем $\psi(t) = (\alpha t, \beta)^\top$. В силу $\|\psi(T)\| = 1$ имеем $\beta \leq 1$ и $\alpha^2 T^2 + \beta^2 = 1$.

Введем нормализованную временную переменную $\tau = \alpha t$ и нормализованные начальный момент времени $\tau_0 = \alpha t_0$ и конечный момент времени $\bar{\tau} = \alpha T$. Тогда получим $\beta = \sqrt{1 - \bar{\tau}^2}$. Отсюда следует $\|\psi(t)\| = \sqrt{\tau^2 + 1 - \bar{\tau}^2}$. Таким образом, оптимальное управление имеет вид

$$(13) \quad \hat{u} = \frac{\psi}{\|\psi\|} = \frac{(\tau, \beta)^\top}{\sqrt{\tau^2 + \beta^2}}.$$

Так как динамика системы автономна, функция Понтрягина (11) постоянна вдоль траектории, и условие трансверсальности (12) дает интеграл энергии $\mathcal{H} = -1 - \alpha y_1 + \sqrt{\tau^2 + \beta^2} \equiv 0$. Отсюда имеем

$$(14) \quad y_1 = \frac{\sqrt{\tau^2 + \beta^2} - 1}{\alpha}.$$

Вторая компонента вектора $y(t)$ получается по формуле

$$(15) \quad y_2 = \int_T^t \frac{\beta}{\sqrt{\alpha^2 s^2 + \beta^2}} ds = \int_{\bar{\tau}}^{\tau} \frac{\beta}{\sqrt{s^2 + \beta^2}} \frac{ds}{\alpha} = \frac{\beta}{\alpha} \left(\text{arsinh} \frac{\tau}{\beta} - \text{artanh} \bar{\tau} \right).$$

Здесь воспользовались соотношением $\text{arsinh} \frac{\bar{\tau}}{\beta} = \text{arsinh} \frac{\bar{\tau}}{\sqrt{1 - \bar{\tau}^2}} = \text{artanh} \bar{\tau}$.

Интегрируя далее, получим для компонент вектора $x(t) = \int_T^t y(s) ds = \int_{\bar{\tau}}^{\tau} y(s/\alpha) \frac{ds}{\alpha}$

$$x_1 = \frac{1}{2\alpha^2} \tau \sqrt{\tau^2 + \beta^2} + \frac{\beta^2}{2\alpha^2} \text{arsinh} \frac{\tau}{\beta} - \frac{\tau}{\alpha^2} - \frac{1}{2\alpha^2} (\beta^2 \text{artanh} \bar{\tau} - \bar{\tau}),$$

$$x_2 = \frac{\beta\tau}{\alpha^2} \text{arsinh} \frac{\tau}{\beta} - \frac{\beta}{\alpha^2} \sqrt{\tau^2 + \beta^2} - \frac{\beta\tau}{\alpha^2} \text{artanh} \bar{\tau} + \frac{\beta}{\alpha^2}.$$

Вычислим элементы матрицы Грама векторов x, y . Подставляя полученные выше выражения для элементов x, y , получим после упрощения, что

$$\begin{aligned}
 \alpha^4 \|x\|^2 &= \frac{1}{4}\tau^4 + \left(\beta^2 \operatorname{artanh}^2 \bar{\tau} + \frac{5\beta^2}{4} + 1 \right) \tau^2 - (\beta^2 \operatorname{artanh} \bar{\tau} + \bar{\tau}) \tau + \\
 &+ \frac{1}{4} (4\beta^4 + 3\beta^2 + 1 + \beta^4 \operatorname{artanh}^2 \bar{\tau} - 2\beta^2 \bar{\tau} \operatorname{artanh} \bar{\tau}) - \\
 &- \tau^2 \sqrt{\tau^2 + \beta^2} + \frac{1}{2} (3\beta^2 \operatorname{artanh} \bar{\tau} + \bar{\tau}) \tau \sqrt{\tau^2 + \beta^2} - 2\beta^2 \sqrt{\tau^2 + \beta^2} - \\
 &- 2\beta^2 \tau^2 \operatorname{artanh} \bar{\tau} \operatorname{arsinh} \frac{\tau}{\beta} + \beta^2 \tau \operatorname{arsinh} \frac{\tau}{\beta} - \frac{\beta^2}{2} (\beta^2 \operatorname{artanh} \bar{\tau} - \bar{\tau}) \operatorname{arsinh} \frac{\tau}{\beta} - \\
 (16) \quad &- \frac{3\beta^2}{2} \tau \sqrt{\tau^2 + \beta^2} \operatorname{arsinh} \frac{\tau}{\beta} + \beta^2 \tau^2 \operatorname{arsinh}^2 \frac{\tau}{\beta} + \frac{\beta^4}{4} \operatorname{arsinh}^2 \frac{\tau}{\beta},
 \end{aligned}$$

$$\begin{aligned}
 \alpha^3 \langle x, y \rangle &= \frac{1}{2}\tau^3 + \left(\beta^2 \operatorname{artanh}^2 \bar{\tau} + \frac{\beta^2}{2} + 1 \right) \tau - \frac{\beta^2}{2} \sqrt{\tau^2 + \beta^2} \operatorname{arsinh} \frac{\tau}{\beta} + \\
 (17) \quad &+ \frac{1}{2} (\beta^2 \operatorname{artanh} \bar{\tau} + \bar{\tau}) \sqrt{\tau^2 + \beta^2} - \frac{3}{2} \tau \sqrt{\tau^2 + \beta^2} + \frac{\beta^2}{2} \operatorname{arsinh} \frac{\tau}{\beta} - \\
 &- \frac{1}{2} (\beta^2 \operatorname{artanh} \bar{\tau} + \bar{\tau}) + \beta^2 \tau \operatorname{arsinh}^2 \frac{\tau}{\beta} - 2\beta^2 \tau \operatorname{artanh} \bar{\tau} \operatorname{arsinh} \frac{\tau}{\beta},
 \end{aligned}$$

$$\begin{aligned}
 \alpha^2 \|y\|^2 &= \tau^2 - 2\sqrt{\tau^2 + \beta^2} + \beta^2 \operatorname{arsinh}^2 \frac{\tau}{\beta} - 2\beta^2 \operatorname{artanh} \bar{\tau} \operatorname{arsinh} \frac{\tau}{\beta} + \\
 (18) \quad &+ (\beta^2 \operatorname{artanh}^2 \bar{\tau} + \beta^2 + 1).
 \end{aligned}$$

2.1. Вычисление параметров $\alpha, \bar{\tau}, \tau_0$

Чтобы численно вычислить оптимальную траекторию в задаче быстрогодействия для заданной матрицы Грама $\begin{pmatrix} \|x_0\|^2 & \langle x_0, y_0 \rangle \\ \langle x_0, y_0 \rangle & \|y_0\|^2 \end{pmatrix}$, надо обратить зависимость начальных данных от параметров $\alpha, \bar{\tau}, \tau_0$, т.е. вычислить значения этих параметров.

Зависимость начальных данных от параметра α алгебраическое. Умножение параметра α на положительную константу λ приводит к умножению грамиана справа и слева на диагональную матрицу $\operatorname{diag}(\lambda^{-2}, \lambda^{-1})$. В конусе \mathcal{S}_+^2 неотрицательно определенных матриц размера 2×2 это действие определяет орбиты радиального типа. Каждая орбита пересекает аффинное компактное сечение конуса \mathcal{S}_+^2 в ровно одной точке. Сама орбита зависит только от параметров $\tau_0, \bar{\tau}$. Она может быть представлена, например, соотношениями $\frac{\|x_0\|^2}{\|y_0\|^4}, \frac{\langle x_0, y_0 \rangle}{\|y_0\|^3}$.

Зависимость орбиты от этих параметров не может быть обращена явно. Чтобы прояснить эту зависимость, вычислим пределы вышеупомянутых со-

отношений, когда параметры $\tau_0, \bar{\tau}$ стремятся к границе своей области определения. Напомним, что $-1 < \bar{\tau} < 1$, $\tau_0 < \bar{\tau}$.

В пределе $\tau_0 \rightarrow \bar{\tau}$ получим $\operatorname{arsinh} \frac{\tau_0}{\beta} \rightarrow \operatorname{arsinh} \frac{\bar{\tau}}{\beta} = \operatorname{artanh} \bar{\tau}$, $\sqrt{\tau_0^2 + \beta^2} \rightarrow 1$. Подставляя вместе с $\tau = \tau_0$ в (16), (17), (18), получим, что грамиан векторов x_0, y_0 стремится к 0. Однако если в то же время устремить $\alpha \rightarrow 0$ таким образом, что соотношение $\frac{\bar{\tau} - \tau_0}{\alpha} = T - t_0$ равно 1, то управление стремится к $\hat{u} \equiv (\bar{\tau}, \beta)^\top$ и траектория стремится к отрезку параболы, заданной по формуле $x(t) = \frac{1}{2}(t - T)^2 \hat{u}$, $y(t) = (t - T) \hat{u}$. Отсюда получаем $x_0 \rightarrow \frac{1}{2}(\bar{\tau}, \beta)^\top$, $y_0 \rightarrow -(\bar{\tau}, \beta)^\top$, и грамиан стремится к матрице $\begin{pmatrix} \frac{1}{4} & -\frac{1}{2} \\ -\frac{1}{2} & 1 \end{pmatrix}$. В частности,

$$\frac{\|x_0\|^2}{\|y_0\|^4} \rightarrow \frac{1}{4}, \quad \frac{\langle x_0, y_0 \rangle}{\|y_0\|^3} \rightarrow -\frac{1}{2}.$$

В пределе $\bar{\tau} \rightarrow \pm 1$ получаем $\beta \rightarrow 0$, $\beta \operatorname{arsinh} \frac{\tau}{\beta} \rightarrow 0$, $\beta \operatorname{artanh} \bar{\tau} \rightarrow 0$, $\sqrt{\tau^2 + \beta^2} \rightarrow |\tau|$. Подставляя в (16), (17), (18), имеем

$$\begin{aligned} \alpha^4 \|x\|^2 &\rightarrow \frac{1}{4} \tau^4 + \tau^2 + \frac{1}{4} - \tau^2 |\tau| + \frac{1}{2} \bar{\tau} \tau |\tau| - \bar{\tau} \tau, \\ \alpha^3 \langle x, y \rangle &\rightarrow \frac{1}{2} (|\tau| - 1) (\tau (|\tau| - 2) + \bar{\tau}), \\ \alpha^2 \|y\|^2 &\rightarrow (|\tau| - 1)^2. \end{aligned}$$

При $\bar{\tau} \rightarrow -1$ получаем $\tau \leq -1$ и $|\tau| = -\tau$. Положив $\alpha = -1 - \tau_0$, снова имеем $T - t_0 \rightarrow 1$ и получаем предел $\begin{pmatrix} \|x_0\|^2 & \langle x_0, y_0 \rangle \\ \langle x_0, y_0 \rangle & \|y_0\|^2 \end{pmatrix} \rightarrow \begin{pmatrix} \frac{1}{4} & -\frac{1}{2} \\ -\frac{1}{2} & 1 \end{pmatrix}$. В итоге соотношения $\frac{\|x_0\|^2}{\|y_0\|^4}$, $\frac{\langle x_0, y_0 \rangle}{\|y_0\|^3}$ снова стремятся к тем же предельным значениям $\frac{1}{4}, -\frac{1}{2}$.

При $\bar{\tau} \rightarrow 1$ получаем

$$\begin{pmatrix} \|x_0\|^2 & \langle x_0, y_0 \rangle \\ \langle x_0, y_0 \rangle & \|y_0\|^2 \end{pmatrix} \rightarrow \begin{pmatrix} \frac{1}{4} \frac{(\tau_0 (|\tau_0| - 2) + 1)^2}{\alpha^4} & -\frac{1}{2} \frac{(1 - |\tau_0|)(\tau_0 (|\tau_0| - 2) + 1)}{\alpha^3} \\ -\frac{1}{2} \frac{(1 - |\tau_0|)(\tau_0 (|\tau_0| - 2) + 1)}{\alpha^3} & \frac{(1 - |\tau_0|)^2}{\alpha^2} \end{pmatrix}$$

и $\frac{\|x_0\|^2}{\|y_0\|^4} \rightarrow \frac{1}{4} \frac{(\tau_0 (|\tau_0| - 2) + 1)^2}{(1 - |\tau_0|)^4}$, $\frac{\langle x_0, y_0 \rangle}{\|y_0\|^3} \rightarrow -\frac{1}{2} \frac{\tau_0 (|\tau_0| - 2) + 1}{|1 - |\tau_0|| (1 - |\tau_0|)}$. Если $\tau_0 \geq 0$, то эти пределы снова равны $\frac{1}{4}, -\frac{1}{2}$. При $\tau_0 \leq 0$ они равны $\frac{((1 + \tau_0)^2 - 2)^2}{4(1 + \tau_0)^4}$, $\frac{(1 + \tau_0)^2 - 2}{2|1 + \tau_0|(1 + \tau_0)}$, и $\lim_{\bar{\tau} \rightarrow 1} \frac{\|x_0\|^2}{\|y_0\|^4} = \left(\lim_{\bar{\tau} \rightarrow 1} \frac{\langle x_0, y_0 \rangle}{\|y_0\|^3} \right)^2$.

При $\bar{\tau} = 1$, $\tau_0 \in (-\infty, -1)$ соотношение $\frac{\langle x_0, y_0 \rangle}{\|y_0\|^3}$ монотонно растет от $-\frac{1}{2}$ до $+\infty$. При $\tau_0 \in (-1, 0]$ оно монотонно растет от $-\infty$ до $-\frac{1}{2}$. Предельные значения соотношений $\frac{\|x_0\|^2}{\|y_0\|^4}$, $\frac{\langle x_0, y_0 \rangle}{\|y_0\|^3}$ на границе представлены на рис. 2. Предел отличен от $(\frac{1}{4}, -\frac{1}{2})$ только при $\bar{\tau} \rightarrow 1$, $\tau_0 < 0$, в этом случае его геометрическое место – парабола.

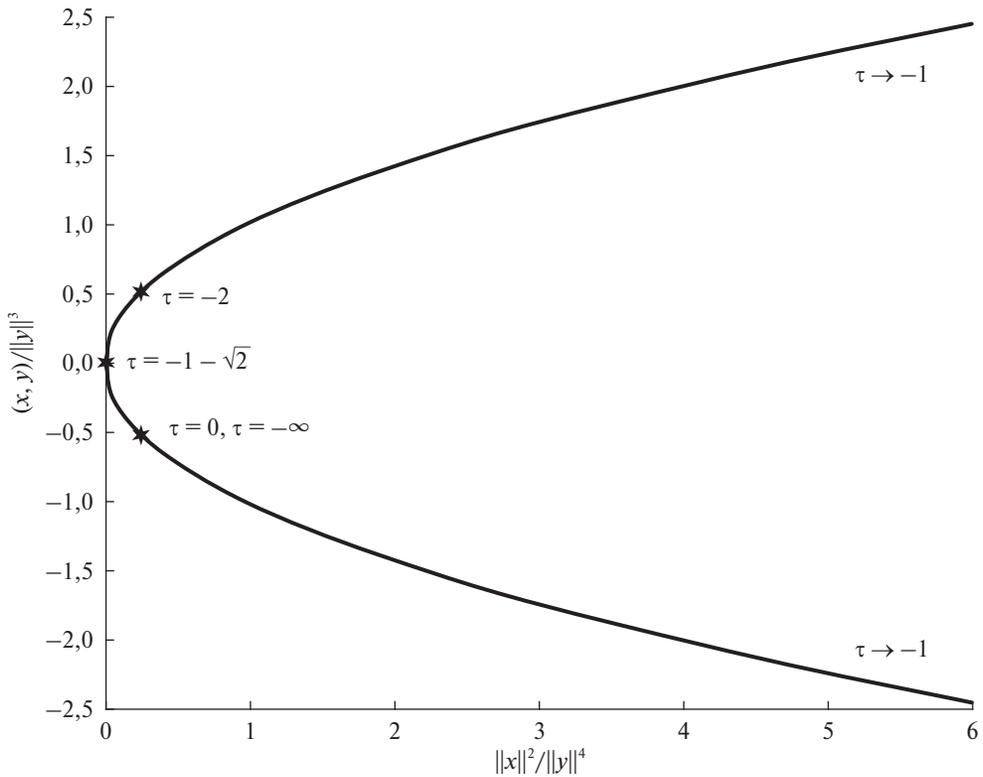


Рис. 2. Пределы соотношений $\frac{\|x_0\|^2}{\|y_0\|^4}, \frac{\langle x_0, y_0 \rangle}{\|y_0\|^3}$ при стремлении параметров $\bar{\tau}, \tau_0$ к границе их области определения.

Наконец, при $\tau_0 \rightarrow -\infty$ грамиан стремится к бесконечности. Однако если одновременно устремить $\alpha \rightarrow +\infty$ таким образом, что $\frac{-\tau_0}{\alpha} \rightarrow 1$, то ведущие члены в выражении для τ_0 доминируют, и снова получим $\begin{pmatrix} \|x_0\|^2 & \langle x_0, y_0 \rangle \\ \langle x_0, y_0 \rangle & \|y_0\|^2 \end{pmatrix} \rightarrow \begin{pmatrix} \frac{1}{4} & -\frac{1}{2} \\ -\frac{1}{2} & 1 \end{pmatrix}$.

В итоге, если параметры $\bar{\tau}, \tau_0$ обходят границу своей области определения, то пара соотношений $\left(\frac{\|x_0\|^2}{\|y_0\|^4}, \frac{\langle x_0, y_0 \rangle}{\|y_0\|^3}\right)$ двигается по параболе, представленной на рис. 2, включая бесконечно удаленную точку. За исключением интервала $(\bar{\tau}, \tau_0) \in \{1\} \times \mathbb{R}_-$ пара соотношений стремится к значениям $(\frac{1}{4}, -\frac{1}{2})$. При $\bar{\tau} \in (-1, +1), \tau_0 < \bar{\tau}$ пара принимает значения справа от параболы.

Значения параметров $\alpha, \bar{\tau}, \tau_0$, приводящие к заданному грамиану векторов x_0, y_0 , можно вычислить следующим образом. Сначала вычисляем соотношения $\frac{\|x_0\|^2}{\|y_0\|^4}, \frac{\langle x_0, y_0 \rangle}{\|y_0\|^3}$. Затем определяем значения параметров $\tau_0, \bar{\tau}$, дающие эти соотношения. Это можно осуществить, например, отследив кривые уровня соотношений на плоскости параметров $\bar{\tau}, \tau_0$, найдя графически их пересечение и уточнив значения методом Ньютона. Наконец, находим значение параметра α , например, из (18).

3. Верхняя граница на оптимальное значение

В этом разделе вычисляется значение функционала цены задачи (5) на оптимальном решении задачи быстрогодействия 2. Рассмотрим снова два случая из предыдущего раздела.

Случай $\alpha > 0, \beta > 0$: Значение функционала цены на оптимальной для задачи быстрогодействия траектории равно

$$\frac{1}{2} \int_{t_0}^T \|x(s)\|^2 ds = \frac{1}{2} \int_{\tau_0}^{\bar{\tau}} \|x(\tau/\alpha)\|^2 \frac{d\tau}{\alpha} = \frac{1}{2\alpha^5} \int_{\tau_0}^{\bar{\tau}} \alpha^4 \|x\|^2 d\tau.$$

Интегрируя выражение (16) по переменной τ , получим

$$\begin{aligned} & \frac{\tau^5}{20} + \frac{1}{3} \left(\frac{71}{36} \beta^2 + \beta^2 \operatorname{artanh}^2 \bar{\tau} + 1 \right) \tau^3 - \frac{1}{2} (\beta^2 \operatorname{artanh} \bar{\tau} + \bar{\tau}) \tau^2 + \\ & + \frac{1}{36} (9\beta^4 \operatorname{artanh}^2 \bar{\tau} - 18\beta^2 \bar{\tau} \operatorname{artanh} \bar{\tau} + 56\beta^4 + 27\beta^2 + 9) \tau - \\ & - \frac{1}{9} (3\beta^2 \bar{\tau} - 5\beta^4 \operatorname{artanh} \bar{\tau}) \sqrt{\tau^2 + \beta^2} - \frac{11\beta^2}{8} \tau \sqrt{\tau^2 + \beta^2} + \\ & + \frac{1}{18} (13\beta^2 \operatorname{artanh} \bar{\tau} + 3\bar{\tau}) \tau^2 \sqrt{\tau^2 + \beta^2} - \frac{1}{4} \tau^3 \sqrt{\tau^2 + \beta^2} - \frac{5\beta^4}{8} \operatorname{arsinh} \frac{\tau}{\beta} - \\ & - \frac{\beta^2}{2} (\beta^2 \operatorname{artanh} \bar{\tau} - \bar{\tau}) \tau \operatorname{arsinh} \frac{\tau}{\beta} + \frac{\beta^2}{2} \tau^2 \operatorname{arsinh} \frac{\tau}{\beta} - \frac{2\beta^2}{3} \tau^3 \operatorname{artanh} \bar{\tau} \operatorname{arsinh} \frac{\tau}{\beta} - \\ & - \frac{5\beta^4}{9} \sqrt{\tau^2 + \beta^2} \operatorname{arsinh} \frac{\tau}{\beta} - \frac{13\beta^2}{18} \tau^2 \sqrt{\tau^2 + \beta^2} \operatorname{arsinh} \frac{\tau}{\beta} + \\ & + \frac{\beta^4}{4} \tau \operatorname{arsinh}^2 \frac{\tau}{\beta} + \frac{\beta^2}{3} \tau^3 \operatorname{arsinh}^2 \frac{\tau}{\beta}. \end{aligned}$$

При $\tau = \bar{\tau}$ это выражение принимает вид

$$\frac{1}{1080} ((1024\beta^4 - 163\beta^2 + 54)\bar{\tau} - 675\beta^4 \operatorname{artanh} \bar{\tau}).$$

Поэтому значение $-\omega^{TO}$ функционала на оптимальной траектории для задачи быстрогодействия удовлетворяет условию

$$\begin{aligned} -\alpha^5 \omega^{TO} &= -\frac{\tau_0^5}{40} - \frac{1}{6} \left(\frac{71}{36} \beta^2 + \beta^2 \operatorname{artanh}^2 \bar{\tau} + 1 \right) \tau_0^3 + \frac{1}{4} (\beta^2 \operatorname{artanh} \bar{\tau} + \bar{\tau}) \tau_0^2 - \\ & - \frac{1}{72} (9\beta^4 \operatorname{artanh}^2 \bar{\tau} + 27\beta^2 + 56\beta^4 - 18\beta^2 \bar{\tau} \operatorname{artanh} \bar{\tau} + 9) \tau_0 + \\ & + \frac{1}{2160} ((1024\beta^4 - 163\beta^2 + 54)\bar{\tau} - 675\beta^4 \operatorname{artanh} \bar{\tau}) + \end{aligned}$$

$$\begin{aligned}
& + \frac{1}{18}(3\beta^2\bar{\tau} - 5\beta^4 \operatorname{artanh}\bar{\tau})\sqrt{\tau_0^2 + \beta^2} + \frac{11\beta^2}{16}\tau_0\sqrt{\tau_0^2 + \beta^2} - \\
& - \frac{1}{36}(13\beta^2 \operatorname{artanh}\bar{\tau} + 3\bar{\tau})\tau_0^2\sqrt{\tau_0^2 + \beta^2} + \frac{1}{8}\tau_0^3\sqrt{\tau_0^2 + \beta^2} + \\
& + \frac{5\beta^4}{16}\operatorname{arsinh}\frac{\tau_0}{\beta} + \frac{\beta^2}{4}(\beta^2 \operatorname{artanh}\bar{\tau} - \bar{\tau})\tau_0\operatorname{arsinh}\frac{\tau_0}{\beta} - \frac{\beta^2}{4}\tau_0^2\operatorname{arsinh}\frac{\tau_0}{\beta} + \\
& + \frac{\beta^2}{3}\tau_0^3\operatorname{artanh}\bar{\tau}\operatorname{arsinh}\frac{\tau_0}{\beta} + \frac{5\beta^4}{18}\sqrt{\tau_0^2 + \beta^2}\operatorname{arsinh}\frac{\tau_0}{\beta} + \\
& + \frac{13\beta^2}{36}\tau_0^2\sqrt{\tau_0^2 + \beta^2}\operatorname{arsinh}\frac{\tau_0}{\beta} - \frac{\beta^4}{8}\tau_0\operatorname{arsinh}^2\frac{\tau_0}{\beta} - \frac{\beta^2}{6}\tau_0^3\operatorname{arsinh}^2\frac{\tau_0}{\beta}.
\end{aligned}$$

Случай $\alpha\beta = 0$: Для одномерной задачи оптимального управления имеем

$$\frac{x(t)^2}{2} = \begin{cases} \frac{t^4}{8}, & \text{если } t > -\sqrt{\sigma x_0 + \frac{y_0^2}{2}}, \\ \frac{1}{2} \left(\frac{t^2}{2} + 2\sqrt{\sigma x_0 + \frac{y_0^2}{2}}t + \sigma x_0 + \frac{y_0^2}{2} \right)^2, & t < -\sqrt{\sigma x_0 + \frac{y_0^2}{2}}. \end{cases}$$

Отсюда в силу $\tilde{\phi}^{-1} = -\sqrt{\sigma x_0 + \frac{y_0^2}{2}}$, $\sigma = \operatorname{sign}\left(x_0 + \frac{y_0|y_0|}{2}\right)$ и $t_0 = -\sigma y_0 + 2\tilde{\phi}^{-1}$ получаем, что значение функционала задачи (1) на оптимальной для задачи быстродействия траектории равно

$$\begin{aligned}
(19) \quad \frac{1}{2} \int_{t_0}^0 x(t)^2 dt &= \frac{1}{2} \int_{t_0}^{\tilde{\phi}^{-1}} \left(\frac{t^2}{2} - 2\tilde{\phi}^{-1}t + \tilde{\phi}^{-2} \right)^2 dt + \int_{\tilde{\phi}^{-1}}^0 \frac{t^4}{8} dt \\
&= -\frac{t_0^5}{40} + \frac{t_0^4\tilde{\phi}^{-1}}{4} - \frac{5}{6}t_0^3\tilde{\phi}^{-2} + t_0^2\tilde{\phi}^{-3} - \frac{t_0\tilde{\phi}^{-4}}{2} + \frac{\tilde{\phi}^{-5}}{12} \\
&= -\frac{23}{60}\tilde{\phi}^{-5} + \frac{\tilde{\phi}^{-4}\sigma y_0}{2} - \frac{\tilde{\phi}^{-2}\sigma y_0^3}{6} + \frac{\sigma y_0^5}{40} \\
&= \frac{\sigma y_0 x_0^2}{2} + \frac{x_0 y_0^3}{3} + \frac{\sigma y_0^5}{15} + \frac{23}{60} \left(\sigma x_0 + \frac{y_0^2}{2} \right)^{5/2}.
\end{aligned}$$

4. Качество аппроксимации

В этом разделе сравнивается значение функционала на субоптимальной траектории с оптимальными значениями для тех начальных данных, для которых оптимальное значение известно.

Рассмотрим сначала одномерную задачу (1). Так как оптимальное значение (2) (помноженное на -1) и значение (19) на траектории задачи быстродействия удовлетворяют соотношению (4), относительная разница этих значений

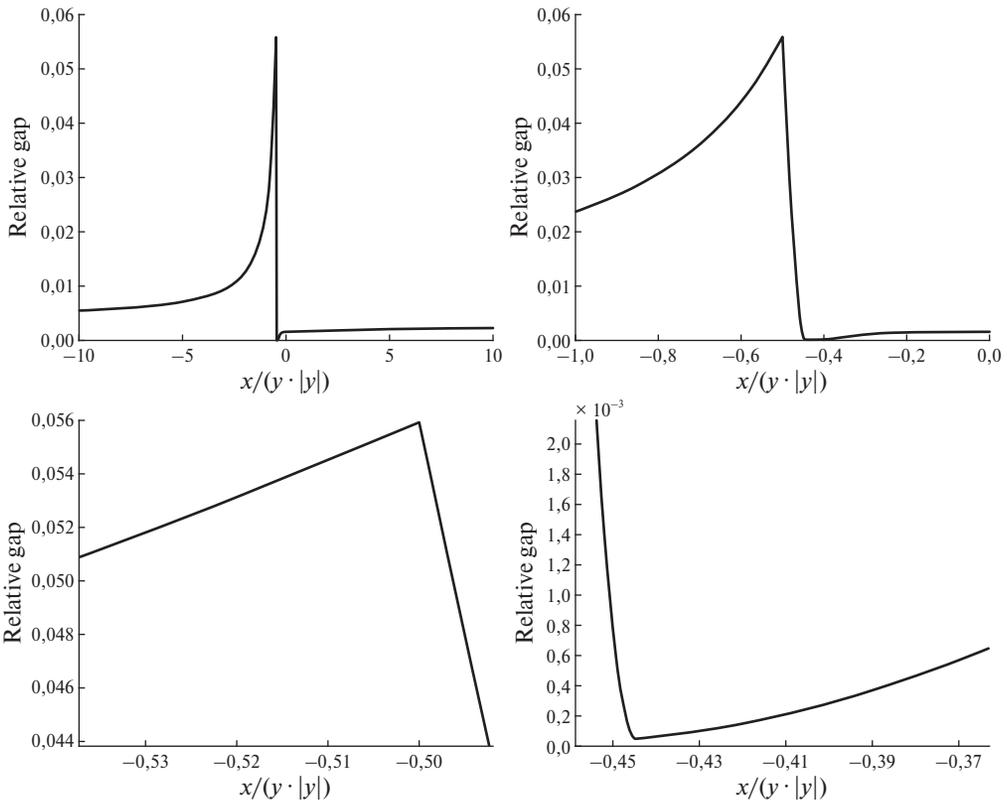


Рис. 3. Относительная разница между значением функционала задачи (1) на оптимальной для задачи быстрогодействия траектории и оптимальным значением.

зависит только от соотношения $\frac{x}{|y|y}$. Эта разница представлена на рис. 3. В задаче быстрогодействия управление на оптимальных траекториях переключается на кривой $x = -\frac{y|y|}{2}$, в то время как на оптимальных для исходной задачи (1) траекториях управление переключается на кривой $x = -\beta y|y|$, где $\beta \approx 0,4446$. Графики справа и снизу на рис. 3 являются увеличениями деталей графика в левом верхнем углу. Разница колеблется между $5,4 \cdot 10^{-5}$ и $5,6 \cdot 10^{-2}$.

Так как и задача быстрогодействия (10), и задача (5) сводятся к их одномерным аналогам рис. 1 и (1), в случае, если начальные значения векторов x, y коллинеарны, указанный разрыв имеет место также для значений в двухмерных задачах.

Рассмотрим теперь автомодельные траектории, найденные в [7]. Сначала вычислим оптимальные значения функционала задачи (5) на этих траекториях. В силу (4) функция Белмана на траекториях, удовлетворяющих условию (8), имеет вид

$$\omega_{2D}(x(t), y(t)) = \frac{\lambda(t)^5}{\lambda_0^5} \omega_{2D}(x(0), y(0)).$$

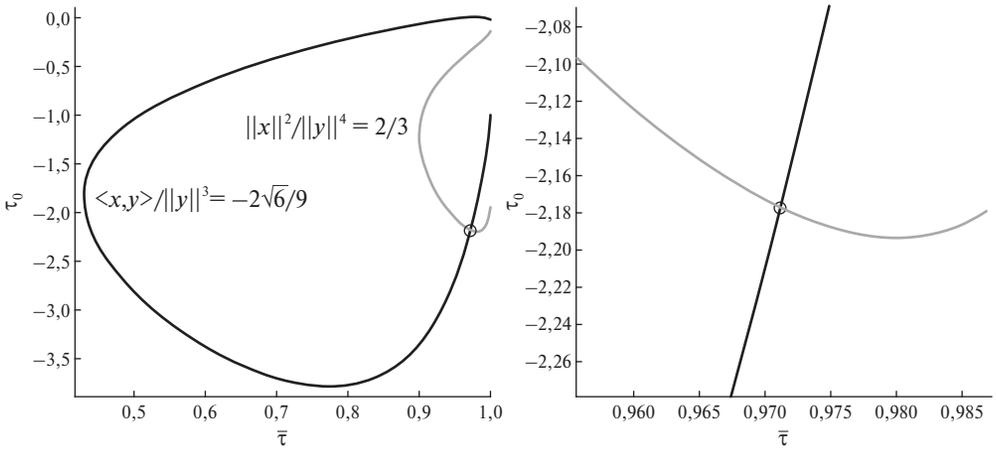


Рис. 4. Линии уровня соотношений $\frac{\|x_0\|^2}{\|y_0\|^4}$, $\frac{\langle x_0, y_0 \rangle}{\|y_0\|^3}$ в плоскости параметров $(\bar{\tau}, \tau_0)$ для значений, отвечающих автомодельной траектории.

Дифференцируя по t и пользуясь соотношением $\frac{d\omega(x(t), y(t))}{dt} = \frac{1}{2}\|x(t)\|^2$, получим

$$(20) \quad \omega_{2D}(x(0), y(0)) = -\frac{\lambda_0^5}{540}.$$

Вычислим значение функционала задачи (5) на оптимальных для задачи быстродействия траекториях с теми же начальными условиями (7). Для этого нужно обратить зависимость (16), (17), (18), т.е. вычислить значения параметров $\alpha, \tau, \bar{\tau}$, приводящих к нужным начальным векторам.

Следуя схеме, предложенной в конце раздела 2, сначала вычислим линии уровня соотношений $\frac{\|x_0\|^2}{\|y_0\|^4}$, $\frac{\langle x_0, y_0 \rangle}{\|y_0\|^3}$ в плоскости параметров $(\bar{\tau}, \tau_0)$. В силу (7) получаем

$$\frac{\|x_0\|^2}{\|y_0\|^4} = \frac{2}{3}, \quad \frac{\langle x_0, y_0 \rangle}{\|y_0\|^3} = \frac{2\sqrt{6}}{9}.$$

Соответствующие линии уровня представлены на рис. 4. Значения параметров $\bar{\tau}, \tau_0$, которые соответствуют автомодельной траектории, на рисунке даны единственной точкой пересечения линий уровня (кружочек). Уточнив полученные графически приближения методом Ньютона, получим значения параметров

$$\bar{\tau} \approx 0,97116420999, \quad \tau_0 \approx -2,17695799429.$$

Подставляя их в (18) и полагая $\|y_0\|^2 = \frac{\lambda_0^2}{6}$ в силу (7), получим значение параметра $\alpha \approx 4,13415835032 \lambda_0^{-1}$.

Подставляя полученные значения параметров в выражение для функционала цены, получим субоптимальное значение цены $\approx 0,0019779902706 \lambda_0^5$. Сравнивая с оптимальным значением (20), вычисленным выше, получаем относительную ошибку $\approx 6,8 \cdot 10^{-2}$.

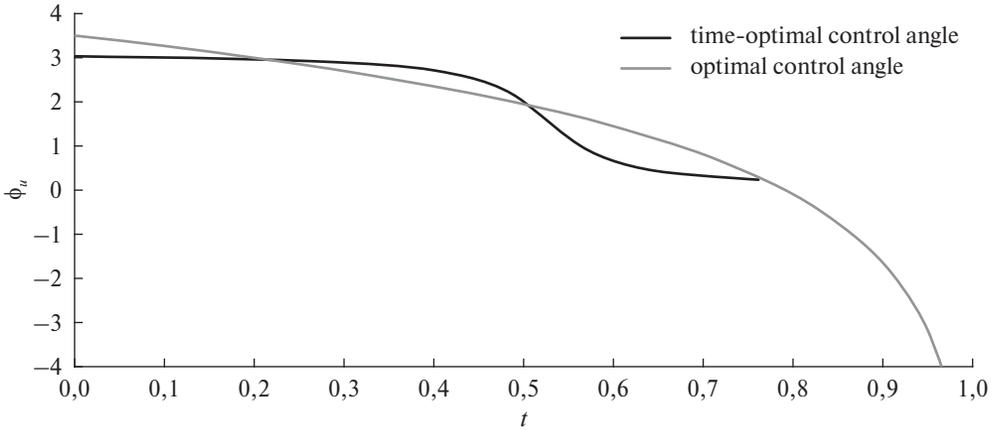


Рис. 5. Эволюция фазового угла оптимального управления для задачи (5) и для задачи быстродействия (10) при одном и том же начальном значении векторов x, y , лежащих на автомодельной траектории с временем перехода в начало координат $T = 1$.

Сравним теперь оптимальное управление на автомодельной траектории с управлением на траектории задачи быстродействия с теми же начальными значениями векторов x, y , соответствующими $\lambda_0 = 1$. Тогда оптимальной для задачи (5) траектории нужно единичное время, чтобы прийти в начало координат.

Оптимальное для задачи быстродействия управление задано (13), где τ пробегает интервал от τ_0 до $\bar{\tau}$, а исходная переменная времени t – интервал от 0 до $T^{TO} = \alpha^{-1}(\bar{\tau} - \tau_0) \approx 0,7614904746$. Отметим, что управление вращается по часовой стрелке вокруг начала координат. Заметим также, что время прихода в начало координат меньше, чем время прихода $T = 1$ автомодельной траектории, поскольку в задаче быстродействия минимизируется именно это время прихода. На рис. 5 фазовый угол управления в задаче быстродействия представлен как функция от времени.

В силу (9) оптимальное управление на автомодельной траектории задано по формуле

$$\hat{u} = (\cos \varphi(t), \sin \varphi(t))^T, \quad \varphi(t) = \sqrt{5} \log(T - t) + \text{const},$$

где $T = \lambda_0 = 1$ есть время прихода траектории в начало координат, а константа есть угол управления в начальный момент времени $t = 0$. Для вычисления этой константы сначала вычислим начальное значение вектора y по формулам (14), (15). Оно задается выражением

$$y(0) \approx \begin{pmatrix} 0,2878394861 \\ -0,2895083711 \end{pmatrix}.$$

Как было сказано в разделе 1, оптимальное управление направлено с разницей в угол $\pi - \arctg \sqrt{5}$ по отношению к y , что дает начальное значение

угла управления $\approx 3,5035658841$. При $t \rightarrow T = 1$ угол управления стремится как логарифм к $-\infty$, и управление совершает бесконечное число оборотов вокруг начала координат за конечное время. Зависимость угла управления от времени представлена на рис. 5. При $t \rightarrow 1$ фаза оптимального управления стремится к $-\infty$.

5. Заключение и перспективы

В данной работе рассмотрены две задачи оптимального управления. Эти задачи обладают одним и тем же множеством допустимых траекторий, но имеют разные функционалы цены. Задача быстродействия (10) решается аналитически, но в задаче (5), в которой возникает особый режим второго порядка, явно известен только ограниченный набор оптимальных траекторий.

Описано решение задачи быстродействия, которое использовано для построения верхней границы на оптимальное значение в задаче (5). Сравнение субоптимальных траекторий (оптимальных для задачи быстродействия) с известными оптимальными в задаче (5) траекториями показывает, что относительная ошибка по значению функционала колеблется от тысячных долей процента до нескольких процентов. Разница в фазе оптимального управления может быть существенной и достигать 45 градусов.

Верхняя оценка на оптимальное значение функционала может быть использована для ограничения области в расширенном фазовом пространстве, в которой могут лежать оптимальные траектории задачи (5). Это, в свою очередь, упростит изучение оптимального синтеза в этой задаче. Более конкретно в [9] было показано, что симметрия Фуллера (3) влечет за собой соотношение $\omega(x, y) = \frac{1}{5}(\langle \psi, y \rangle + 2\langle \phi, x \rangle)$ на функцию Беллмана задачи, где ϕ, ψ – сопряженные с x, y переменные. Поэтому верхняя оценка на значение функционала в некоторой точке (x, y) влечет линейное неравенство на оптимальные значения сопряженных переменных в этой точке.

Численные эксперименты показывают, что автомодельная траектория является отталкивающей в фактор-пространстве расширенного фазового пространства по действию групп симметрии, в то время как траектории, соответствующие одномерному аналогу (1) задачи, являются притягивающими. Формальное доказательство этого свойства и качественное описание оптимального синтеза в задаче (5) остаются открытыми и будут исследоваться в последующих работах.

СПИСОК ЛИТЕРАТУРЫ

1. *Pontryagin L.S., Boltyanskii V.G., Gamkrelidze R.V., Mischchenko E.F.* The Mathematical Theory of Optimal Processes / New York, London: Wiley, 1962.
2. *Fuller A.T.* Relay control systems optimized for various performance criteria / Proceedings of the First World Congress IFAC. Butterworth, 1960. P. 510–519.
3. *Kelley H.J., Kopp R.E., Moyer M.G.* Singular extremals / Topics in Optimization. New York: Acad. Press, 1967. P. 63–101.

4. *Marchal C.* Chattering arcs and chattering controls // J. Optimiz. Theory App. 1973. V. 11. No. 5. P. 441–468.
5. *Lewis R.M.* Definitions of order and junction condition in singular control problems // SIAM J. Contr. Optim. 1980. V. 18. No. 1. P. 21–32.
6. *Kupka I.* Generic properties of extremals in optimal control problems / Differential geometric control theory, Boston: Birkhäuser, 1983. Progr. Math. V. 27. P. 310–315.
7. *Zelikin M.I., Borisov V.F.* Theory of chattering control with applications to astronautics, robotics, economics, and engineering / Boston: Birkhäuser, 1994.
8. *Lokutsievskiy L.V.* Generic structure of the lagrangian manifold in chattering problems. // Sbornik Math. 2014. V. 205. No. 3. P. 432–458.
9. *Zelikin M.I., Melnikov N.B., Hildebrand R.* Topological structure of a typical fibre of optimal synthesis for chattering problems // P. Steklov Inst. Math. 2001. V. 233. P. 116–142.
10. *Zelikin M.I., Lokutsievskiy L.V., Hildebrand R.* Typicality of chaotic fractal behaviour of integral vortices in hamiltonian systems with discontinuous right hand side // Journal of Mathematical Sciences. 2017. V. 221. No. 1. P. 1–136.

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

Поступила в редакцию 03.03.2025

После доработки 20.05.2025

Принята к публикации 27.06.2025

© 2025 г. Ф. ЧЖАН (zfzhao@stanford.edu),
С. БОЙД, д-р философии (boyd@stanford.edu)
(Стэнфордский университет, Стэнфорд, США)

РЕШЕНИЕ ЗАДАЧ О МНОГОТОВАРНЫХ СЕТЕВЫХ ПОТОКАХ БОЛЬШОЙ РАЗМЕРНОСТИ НА ГРАФИЧЕСКИХ ПРОЦЕССОРАХ

Рассматривается задача о многотоварных сетевых потоках из всех пар узлов в сети с ребрами, имеющими заданные пропускные способности. При обычном подходе для каждой пары узлов “источник–назначение” на каждом ребре отслеживается отдельный поток. В статье используется более эффективная формулировка, в которой потоки с одним и тем же узлом-назначением объединяются, что позволяет уменьшить количество переменных в k раз, где k – размер сети. Задачи с сотнями узлов, с общим числом переменных порядка миллиона, могут быть решены стандартными общими методами внутренней точки на центральных процессорах (CPU); ниже используются совместимые с графическими процессорами (GPU) алгоритмы, которые могут решать такие задачи гораздо быстрее и, кроме того, масштабируются на гораздо большие задачи, с миллиардом переменных. Представленный метод основан на прямо-двойственном гибридном градиентном алгоритме и использует несколько особенностей задачи для эффективных вычислений на GPU. С помощью численных экспериментов показано, что прямо-двойственный метод многотоварных сетевых потоков ускоряет современные коммерческие решатели от 100 до 1000 раз и масштабируется на задачи гораздо большего размера. Приведена реализация данного метода с открытым исходным кодом.

Ключевые слова: многотоварные потоки, прямо-двойственный метод, GPU-совместимая оптимизация.

DOI: 10.31857/S0005231025080066, **EDN:** UTHPNJ

1. Введение

1.1. Оптимизация многотоварных сетевых потоков

Формулировка задачи о многотоварных сетевых потоках (multicommodity network flow, MCF), приведенная ниже, следует [1].

Сеть. Рассмотрим направленную сеть с n узлами и m ребрами, которая является полносвязной, т.е. между каждой парой узлов существует направленный путь. Обозначим через $A \in \mathbf{R}^{n \times m}$ ее матрицу инцидентности:

$$A_{i\ell} = \begin{cases} +1, & \text{если ребро } \ell \text{ входит в узел } i, \\ -1, & \text{если ребро } \ell \text{ выходит из узла } i, \\ 0 & \text{в остальных случаях.} \end{cases}$$

Ребро ℓ имеет положительную пропускную способность c_ℓ . Суммарный поток на ребре ℓ (определен ниже) не может превышать c_ℓ .

Матрица трафика. Рассмотрим постановку задачи с многотоварными потоками из всех пар, т.е. существует трафик, который исходит из каждого узла и предназначен для каждого другого узла. Будем характеризовать трафик между всеми парами “источник–назначение” с помощью матрицы трафика $T \in \mathbf{R}^{n \times n}$. Для любой пары различных узлов i, j обозначим через $T_{ij} \geq 0$ трафик из узла-источника j в узел-назначение i . Трафик из узла в самого себя отсутствует; для математического удобства определим диагональные элементы матрицы трафика как $T_{ii} = -\sum_{j \neq i} T_{ij}$ (отрицательное значение суммарного трафика с узлом-назначением i). При таком определении диагональных элементов имеем $T\mathbf{1} = 0$, где $\mathbf{1}$ – вектор, состоящий из единиц.

Полезность сети. Пусть u_{ij} – строго вогнутая возрастающая функция полезности для трафика из узла j в узел i , где $j \neq i$. Будем считать, что функции полезности дифференцируемы, с областью значений \mathbf{R}_{++} – множеством положительных чисел. (Описанные ниже методы легко распространяются на недифференцируемые функции полезности, если вместо градиентов использовать субградиенты.) Суммарная полезность, которую необходимо максимизировать, равна $\sum_{i \neq j} u_{ij}(T_{ij})$. Для простоты примем $u_{ii} = 0$, и тогда суммарную полезность можно записать как

$$U(T) = \sum_{i,j} u_{ij}(T_{ij}).$$

Функция U имеет область значений $\mathcal{T} = \{T \mid T_{ij} > 0 \text{ для } i \neq j\}$, т.е. матрица трафика должна содержать положительные внедиагональные элементы.

Примерами функций полезности являются взвешенная логарифмическая полезность $u(s) = w \log s$ и взвешенная степенная полезность $u(s) = ws^\gamma$, причем $\gamma \in (0, 1)$, а $w > 0$ – вес.

Матрица потоков, определяемая через узлы-назначения. Следуя [1], агрегируем все потоки с одним и тем же узлом-назначением, рассматривая их как один товар, который сохраняется во всех узлах, кроме источника и назначения, но может быть разделен и объединен. Товарные потоки задаются матрицей потоков $F \in \mathbf{R}^{n \times m}$ через узлы-назначения, где $F_{i\ell} \geq 0$ – поток на ребре ℓ , предназначенный для узла i . Ограничение пропускной способности ребра можно выразить как $F^T \mathbf{1} \leq c$, где неравенство является поэлементным. Похожая формулировка агрегирования потоков, хотя и основанная на источнике, рассмотрена в [2].

Сохранение потока. Поток, предназначенный для узла i , сохраняется во всех узлах $j \neq i$, включая дополнительное вливание трафика T_{ij} , который исходит из узла j и предназначен для узла i . Это значит, что

$$T_{ij} + \sum_{\ell} A_{j\ell} F_{i\ell} = 0, \quad i, j = 1, \dots, n, \quad j \neq i.$$

В узле-назначении весь трафик выходит, и по определению T_{ii} имеем:

$$T_{ii} + \sum_{\ell} A_{i\ell} F_{i\ell} = 0, \quad i = 1, \dots, n.$$

Объединяя эти два условия и используя определение T_{ii} , сохранение потока можно компактно записать в матричной форме как

$$T + FA^T = 0.$$

Задача о многотоварных потоках. В задаче MCF будем искать матрицу потоков, максимизирующую суммарную полезность сети при соблюдении ограничений на пропускную способность ребер и сохранение потока. Это можно выразить в виде следующей задачи:

$$(1) \quad \begin{array}{l} \text{максимизировать } U(T) \\ \text{при ограничениях } F \geq 0, \quad F^T \mathbf{1} \leq c, \quad T + FA^T = 0, \end{array}$$

с переменными F и T и неявным ограничением $T \in \mathcal{T}$. Здесь заданы топология сети A , пропускные способности ребер c и функции полезности трафика u_{ij} .

Можно исключить матрицу трафика T с помощью $T = -FA^T$ и сформулировать задачу MCF только в терминах переменной F :

$$(2) \quad \begin{array}{l} \text{максимизировать } U(-FA^T) \\ \text{при ограничениях } F \geq 0, \quad F^T \mathbf{1} \leq c, \end{array}$$

с переменной F и неявным ограничением $-FA^T \in \mathcal{T}$. Число скалярных переменных в этой задаче равно nm . Для дальнейшего использования определим множество допустимых потоков как

$$\mathcal{F} = \{F \mid F \geq 0, F^T \mathbf{1} \leq c\}.$$

Существование и единственность решения. Сначала покажем, что задача MCF (1) всегда разрешима. Рассмотрим единичный поток из каждого источника в каждое назначение по кратчайшему пути (с наименьшим числом ребер); такой путь существует, поскольку граф полносвязен. Обозначим эту матрицу потоков через F^{sp} . Теперь примем $F = \alpha F^{\text{sp}}$, где $\alpha = 1 / \max_{\ell} ((F^{\text{sp}T} \mathbf{1})_{\ell} / c_{\ell}) > 0$; в результате, $F^T \mathbf{1} \leq c$. Очевидно, что матрица потоков F допустима; имеем $T_{ij} = \alpha > 0$ для $i \neq j$ и $T = -FA^T \in \mathcal{T}$. Это показывает, что задача всегда разрешима. Пусть U^{sp} – соответствующая целевая функция.

Можно добавить к задаче ограничение $U(T) \geq U^{\text{sp}}$, не меняя при этом множества решений. При таком добавлении допустимое множество становится компактным. Отсюда следует, что задача MCF (1) всегда разрешима. Решение не обязательно должно быть единственным. Однако оптимальное решение T является единственным. Отметим также, что неявное ограничение $T = -FA^T \in \mathcal{T}$ является избыточным (см. выше).

Решение задачи MCF. Задача о многотоварных потоках (2) является выпуклой [3] и поэтому может быть эффективно решена. В [1] использованы

стандартные общие решатели метода внутренней точки, такие как коммерческий решатель MOSEK [4], совместно с CVXPY [5]; экземпляры задачи с десятками узлов и тысячами переменных решены за несколько секунд на одном CPU. В настоящей статье представлен алгоритм решения задачи MCF, который полностью использует возможности GPU. Для задач малой и средней размерности данный метод дает существенное ускорение по сравнению с общими методами; кроме того, он масштабируется на гораздо большие задачи, которые не могут быть решены общими методами.

1.2. Условие оптимальности и невязка

Условие оптимальности. Обозначим через $\tilde{\mathcal{F}}$ замыкание допустимого множества, включая неявное ограничение $T = -FA^T \in \mathcal{T}$:

$$\tilde{\mathcal{F}} = \mathcal{F} \cap \{F \mid -FA^T \in \text{cl}(\mathcal{T})\},$$

где $\text{cl}(\mathcal{T})$ – замыкание \mathcal{T} .

Матрица потоков F оптимальна для (2) тогда и только тогда, когда $F \in \mathcal{F}$, $-FA^T \in \mathcal{T}$, и неравенство

$$\text{Tr}(Z - F)^T G \geq 0$$

справедливо для всех $Z \in \tilde{\mathcal{F}}$, где $G = \nabla_F(-U)(-FA^T)$ (например, см. [3, раздел 4.2.3]). Имеем $G = U'A$, где $U'_{ij} = u'_{ij}((-FA^T)_{ij})$.

Условие оптимальности через проекцию на \mathcal{F} . Для дальнейшего использования выразим вышеприведенное условие оптимальности в терминах проекции матрицы Q на \mathcal{F} . Обозначим через Π евклидову проекцию на \mathcal{F} . Предположим, что $Q \in \mathbf{R}^{n \times m}$, и зададим $F = \Pi(Q)$; в результате, $F \in \mathcal{F}$. Предположим также, что $-FA^T \in \mathcal{T}$, так что $G = \nabla_F((-U)(-FA^T))$ существует. Тогда F также является евклидовой проекцией Q на $\tilde{\mathcal{F}}$. Отсюда следует, что $\text{Tr}(Z - F)^T G \geq 0$ для всех $Z \in \tilde{\mathcal{F}}$, поэтому условие оптимальности выше выполняется, и матрица потоков F оптимальна. Очевидно, оно будет выполняться и при более слабом условии:

$$G = \gamma(F - Q) \text{ для некоторого } \gamma \geq 0.$$

Подведем итог: F оптимальна, если $F = \Pi(Q)$ для некоторого Q , $-FA^T \in \mathcal{T}$ и $G = \gamma(F - Q)$ для некоторого $\gamma \geq 0$. Обратное утверждение также верно: если F оптимальна, то $F = \Pi(Q)$ для некоторого Q с $-FA^T \in \mathcal{T}$ и $G = \gamma(F - Q)$ для некоторого $\gamma \geq 0$. (Действительно, это справедливо при $\gamma = 1$ и $Q = F - G$.) Это условие оптимальности легко интерпретируется: F является неподвижной точкой операции проекции градиента с длиной шага γ .

Невязка оптимальности. Для любого $Q \in \mathbf{R}^{n \times m}$ с $F = \Pi(Q)$ определим невязку (оптимальности) как

$$r(Q) = \begin{cases} \min_{\gamma \geq 0} \|G - \gamma(F - Q)\|_F^2, & \text{если } -FA^T \in \mathcal{T}, \\ \infty & \text{в противном случае,} \end{cases}$$

где $\|\cdot\|_F^2$ обозначает квадрат фробениусовой нормы матрицы (сумму квадратов ее элементов). При $-FA^T \in \mathcal{T}$ правая часть является квадратичной функцией γ , поэтому минимум легко выражается в явном виде как

$$(3) \quad r(Q) = \begin{cases} \|G\|_F^2 - \frac{\text{Tr}^2 G^T (F - Q)}{\|F - Q\|_F^2}, & \text{если } -FA^T \in \mathcal{T}, F \neq Q, \\ & \text{Tr} G^T (F - Q) \geq 0, \\ \|G\|_F^2, & \text{если } -FA^T \in \mathcal{T}, F = Q \text{ или} \\ & \text{Tr} G^T (F - Q) < 0, \\ \infty & \text{в противном случае.} \end{cases}$$

Очевидно, что $F = \Pi(Q)$ оптимальна тогда и только тогда, когда $r(Q) = 0$.

1.3. Обзор имеющихся результатов

Многоотоварный сетевой поток. Различные постановки задач MCF были сформулированы и изучены ранее. Начало положено в публикациях [6, 7], посвященных версии с линейными функциями полезности, которая может быть сформулирована как задача линейного программирования (ЛП). В более поздних работах предложены нелинейные выпуклые постановки [8, 9] и (невыпуклые) смешанные целочисленные формулировки [10–12] задач MCF для различных приложений. Они широко используются в управлении перевозками [13–15], энергетике и экономике [8, 10, 16] и сетевых коммуникациях [11, 17, 18]. В обзорной статье [19] описано более двухсот исследований по задачам MCF за 2000–2019 гг. В настоящей статье рассматривается нелинейная выпуклая постановка задач MCF и разрабатываются GPU-совместимые алгоритмы для решения экземпляров таких задач большой размерности. Обзор нелинейных выпуклых задач MCF можно найти в [20]. Совсем недавно модели MCF стали использоваться при составлении расписаний коммуникации между несколькими GPU для задач глубокого обучения [21, 22], но в соответствующих задачах MCF применяются решатели на основе CPU.

Методы первого порядка для выпуклой оптимизации. Методы первого порядка, такие как алгоритм градиентного спуска, алгоритм проксимальной точки, прямо-двойственный гибридный градиентный алгоритм и их ускоренные версии, применяются для решения различных задач выпуклой оптимизации. По сравнению с методами второго порядка, использующими информацию о гессиане, методы первого порядка известны своей низкой вычислительной сложностью и поэтому привлекательны для решения задач оптимизации большой размерности. Недавно прямо-двойственный гибридный градиентный алгоритм был исследован для решения больших задач ЛП [23–25] и задач оптимальной транспортировки [26] на GPU. Другие методы первого порядка, такие как ADMM (Alternating Direction Method of Multipliers), использованы для разработки методов оптимизации с GPU-ускорением для задач об оптимальном потоке мощности (optimal power flow) [27, 28].

Оптимизация сетевого потока с GPU-ускорением. В статье [29], посвященной методам на базе GPU для оптимизации сетевых потоков, рассмотрена реализация параллельного алгоритма маршрутизации на GPU для программно-определяемых сетей (software-defined networking, SDN), который решает лагранжеву релаксацию задачи смешанного целочисленного ЛП. Авторы [30] реализовали генетический метод на GPU для решения задачи маршрутизации, сформулированной в виде целочисленной задачи ЛП. В [31] рассмотрена ЛП-формулировка задачи о многотоварных сетевых потоках и построена модель глубокого обучения для генерации новых столбцов в методе отложенной генерации столбцов. Авторы [32] реализовали асинхронный алгоритм push-relabel для задачи о максимальном сетевом потоке с одним товаром, который является гибридом CPU-GPU. В соответствии с [1] в статье [33] использована точно такая же формулировка агрегирования потоков MCF, что и в настоящей работе; обучена нейросетевая модель для минимизации лагранжевой релаксации без ограничений задачи, а результат передан в качестве теплого старта в решатель Gurobi [34] для получения окончательного ответа. Авторы [35] встроили формулировку задачи агрегирования многотоварных потоков на основе узлов-источников в решение комбинированной транспортной модели и использовали ускоренный вариант проксимального альтернирующего алгоритма “предиктор-корректор” (proximal alternating predictor-corrector algorithm). По их утверждению, предложенный алгоритм ориентирован на GPU, но численные эксперименты приведены на CPU на примерах сетей небольшого размера. В [36] использован прямо-двойственный градиентный метод для решения комбинированных моделей трафика, который, однако, не ориентирован на GPU.

1.4. Новизна

В настоящей работе, мотивированной последними достижениями в области методов оптимизации на основе GPU, решение нелинейных выпуклых задач MCF большой размерности ускоряется с помощью GPU. В частности, используется формулировка задачи MCF из [1] (также описанная выше), которая компактно представлена в матричном виде и требует меньшего количества переменных за счет агрегирования потоков. Показывается, что данная конкретная постановка задачи может быть эффективно решена с помощью прямо-двойственного гибридного градиентного метода первого порядка при работе на GPU.

Насколько известно авторам, настоящая статья является первой, посвященной точному решению выпуклых задач MCF на GPU. В классических работах по таким задачам MCF большой размерности обычно используется лагранжева релаксация для связанных ограничений, а полученные подзадачи меньшей размерности решаются параллельно (например, см. [20]). В настоящей работе не используется какая-либо явная стратегия декомпозиции задачи, а алгоритмическое ускорение является в основном эмпирическим и зависит от высокооптимизированных ядер CUDA для матричных операций.

Более того, размерность задачи уменьшается за счет агрегирования потоков. Поэтому представленный здесь метод имеет более простую форму, не требующую массового решения подзадач и синхронизации, к тому же он является точным.

1.5. План дальнейшего изложения

Предлагаемый алгоритм описывается в разделе 2. Экспериментальные результаты, использующие предложенную реализацию PyTorch, приводятся и обсуждаются в разделе 3; очень похожие результаты, полученные с помощью предложенной реализации JAX, даны в Приложении 7. Итоги настоящей работы подводятся в разделе 4. Код и все данные, необходимые для воспроизведения представленных здесь результатов, можно найти по адресу

<https://github.com/cvxgrp/pdmcf>

2. Прямо-двойственный гибридный градиентный метод

2.1. Прямо-двойственная формулировка с седловой точкой

Выведем прямо-двойственную формулировку с седловой точкой для задачи MCF (1). Обозначим через \mathcal{I} индикаторную функцию множества \mathcal{F} , т.е. $\mathcal{I}(F) = 0$ для $F \in \mathcal{F}$ и $\mathcal{I}(F) = \infty$ в противном случае. Перейдем к минимизации $-U$ в (1), чтобы получить эквивалентную задачу

$$(4) \quad \begin{aligned} & \text{минимизировать} \quad -U(T) + \mathcal{I}(F) \\ & \text{при ограничении} \quad T = -FA^T \end{aligned}$$

с переменными T и F . Введем двойственную переменную $Y \in \mathbf{R}^{n \times n}$, связанную с (матричным) ограничением-равенством. Тогда лагранжиан имеет вид

$$\mathcal{L}(T, F; Y) = -U(T) + \mathcal{I}(F) - \text{Tr} Y^T (T + FA^T)$$

(см. [3, глава 5]). Лагранжиан \mathcal{L} является выпуклым по прямым переменным (T, F) и аффинным (следовательно, вогнутым) по двойственной переменной Y . Если $(T, F; Y)$ – седловая точка \mathcal{L} , то (T, F) является решением задачи (4) (а F – решением задачи MCF (2)); обратное утверждение также справедливо.

Можно аналитически минимизировать \mathcal{L} на T , чтобы получить редуцированный лагранжиан

$$(5) \quad \hat{\mathcal{L}}(F; Y) = \inf_T \mathcal{L}(T, F; Y) = -(-U)^*(Y) + \mathcal{I}(F) - \text{Tr} Y^T F A^T,$$

где U^* – сопряженная функция к U [3, раздел 3.3]. Этот редуцированный лагранжиан является выпуклым по прямой переменной F и вогнутым по двойственной переменной Y . Если $(F; Y)$ – седловая точка $\hat{\mathcal{L}}$, то F – решение задачи MCF (2) (см. [37, раздел 1]). Заметим, что $\hat{\mathcal{L}}$ является выпукло-вогнутым, с билинейной связью.

2.2. Основной метод PDHG

Прямо-двойственный гибридный градиентный алгоритм (PDHG), впервые представленный в [38] и позже изученный в [39, 40], является методом первого порядка для нахождения седловой точки выпукло-вогнутой функции с билинейной связью. В [40, раздел 4.1] алгоритм расширен за счет включения избыточной релаксации (перерелаксации), которая, как было замечено, улучшает сходимость на практике. Для (5) алгоритм PDHG имеет вид

$$(6) \quad \begin{aligned} \hat{F}^{k+1/2} &= \mathbf{prox}_{\alpha\mathcal{T}}(F^{k-1/2} + \alpha Y^k A), \\ F^{k+1} &= 2\hat{F}^{k+1/2} - F^{k-1/2}, \\ \hat{Y}^{k+1} &= \mathbf{prox}_{\beta(-U)^*}(Y^k - \beta F^{k+1} A^T), \\ F^{k+1/2} &= \rho \hat{F}^{k+1/2} + (1 - \rho)F^{k-1/2}, \\ Y^{k+1} &= \rho \hat{Y}^{k+1} + (1 - \rho)Y^k, \end{aligned}$$

где $\mathbf{prox}_f(v) = \operatorname{argmin}_x (f(x) + (1/2)\|x - v\|_2^2)$ обозначает проксимальный оператор для f [41], $\alpha, \beta > 0$ – положительные длины шага, удовлетворяющие $\alpha\beta \leq 1/\|A\|_2^2$, а $\rho \in (0, 2)$ представляет собой параметр перерелаксации.

Разумными вариантами выбора параметров являются

$$\alpha = \beta = 1/\|A\|_2, \quad \rho = 1,9.$$

(Вместо $\|A\|_2$ можно использовать верхнюю границу для $\|A\|_2$.)

Сходимость. В статье [40] показано, что при существовании седловой точки функции $\hat{\mathcal{L}}$ последовательность $(F^k; Y^k)$ сходится к седловой точке функции $\hat{\mathcal{L}}$ при $k \rightarrow \infty$. Как известно, для MCF существуют оптимальная матрица потока и двойственная переменная, поэтому F^k сходится к оптимальной матрице потока. Отсюда следует, что $r(F^{k-1/2} + \alpha Y^k A) \rightarrow 0$ при $k \rightarrow \infty$. Заметим, что $-FA^T \in \mathcal{T}$ выполняется только в пределе.

2.3. Проксимальные операторы

Здесь более подробно рассмотрим два проксимальных оператора, присутствующих в PDHG.

Первый проксимальный оператор. Отметим, что $\mathbf{prox}_{\alpha\mathcal{T}}$, применяемый на шаге обновления $\hat{F}^{k+1/2}$ в (6), является проекцией на \mathcal{F} :

$$\mathbf{prox}_{\alpha\mathcal{T}}(F) = \Pi(F).$$

Поскольку ограничения, определяющие \mathcal{F} , разделяются по столбцам матрицы F , можно вычислить $\Pi(F)$, спроецировав каждый столбец f_ℓ из F на масштабированный симплекс $\mathcal{S}_\ell = \{f \mid f \geq 0, \mathbf{1}^T f \leq c_\ell\}$. Эта проекция имеет вид

$$\Pi_{\mathcal{S}_\ell}(f_\ell) = (f_\ell - \mu_\ell \mathbf{1})_+,$$

где μ_ℓ – оптимальный множитель Лагранжа, а оператор $(a)_+ = \max\{a, 0\}$ применяется к вектору поэлементно. Оптимальное значение μ_ℓ – наименьшее

неотрицательное значение, для которого $(f_\ell - \mu_\ell \mathbf{1})_+^T \mathbf{1} \leq c_\ell$. Его легко найти с помощью алгоритма биссекции; см. раздел 2.6.

Второй проксимальный оператор. Проксимальный оператор, применяемый на шаге обновления \hat{Y}^{k+1} в (6), может быть разложен покомпонентно, поскольку $\beta(-U)^*$ является суммой функций различных переменных. (Диагональные элементы $-u_{ii}$ равны нулю, поэтому $(-\beta u_{ii})^*$ является индикаторной функцией множества $\{0\}$, а ее проксимальный оператор – нулевой функцией.) Для каждого внедиагонального элемента $i \neq j$ необходимо оценить

$$\mathbf{prox}_{\beta(-u_{ij})^*}(y).$$

Эти одномерные проксимальные операторы легко вычисляются в общем случае. Для взвешенной логарифмической полезности $u(s) = w \log s$ имеем

$$\mathbf{prox}_{\beta(-u)^*}(y) = \frac{y - \sqrt{y^2 + 4\beta w}}{2}.$$

Для взвешенной степенной полезности $u(s) = ws^\gamma$ оператор $\mathbf{prox}_{\beta(-u)^*}(y)$ – единственное отрицательное число z , для которого

$$(-z)^{c_1+2} + y(-z)^{c_1+1} - c_1 c_2 = 0,$$

где

$$c_1 = \frac{\gamma}{1-\gamma} > 0, \quad c_2 = \beta \left(\frac{1}{\gamma} - 1 \right) (w\gamma)^{\frac{1}{1-\gamma}} > 0.$$

2.4. Адаптивный выбор длины шага

В базовом алгоритме PDHG (6) шаги α и β фиксированы. Было замечено [23], что на практике адаптивное изменение их длины в процессе работы алгоритма может существенно улучшить сходимость. Опишем реализацию адаптивного выбора длины шага.

Выразим длину шага как

$$\alpha^k = \eta/\omega^k, \quad \beta^k = \eta\omega^k,$$

где $\eta \leq 1/\|A\|_2$ и $\omega^k > 0$ задает вес прямых переменных. При $\omega^k = 1$ получаем базовый PDHG (6).

Вес ω^k инициализируется значением $\omega^0 = 1$ и затем адаптируется следующим образом [23, раздел 3.3]:

$$(7) \quad \omega^{k+1} = \left(\frac{\Delta_Y^{k+1}}{\Delta_F^{k+1}} \right)^\theta (\omega^k)^{1-\theta},$$

где $\Delta_F^{k+1} = \|F^{k+1/2} - F^{k-1/2}\|_F$, $\Delta_Y^{k+1} = \|Y^{k+1} - Y^k\|_F$ и θ – параметр, равный 0,5 в данной реализации. Интуиция, лежащая в основе обновления веса (7), состоит в балансе невязок прямых и двойственных переменных; подробности см. в [23, раздел 3.3]. В [23] ω обновляется при каждом перезапуске. В настоящей статье без использования перезапусков обнаружено, что

обновление ω^k каждые k^{adapt} итераций при одновременном выполнении условий $\Delta_F^k > 10^{-5}$ и $\Delta_Y^k > 10^{-5}$ хорошо работает на практике для задачи MCF. В экспериментах использовано $k^{\text{adapt}} = 100$. Также можно прекратить адаптацию ω^k после некоторого количества итераций, сохраняя его значение постоянным в последующих итерациях. По крайней мере, технически это означает, что доказательство сходимости при постоянном ω справедливо и для адаптивного алгоритма.

Простая граница для $\|A\|_2$. Можно легко вычислить простую верхнюю границу для

$$\|A\|_2 = \sqrt{\lambda_{\max}(AA^T)},$$

где λ_{\max} обозначает максимальное собственное значение. Заметим, что AA^T – матрица Лапласа, связанная с сетью; как известно, для нее имеет место верхняя граница

$$\lambda_{\max}(AA^T) \leq 2d_{\max},$$

где d_{\max} – наибольшая степень узла в графе. (Для полноты этот результат строго доказан в Приложении 6.) Таким образом, можно взять

$$(8) \quad \eta = 1/\sqrt{2d_{\max}}.$$

2.5. Алгоритм

Представим в окончательной форме алгоритм, который будем называть PDMCF. Задаем $r^0 = +\infty$, $\alpha^0 = \eta/\omega^0$ и $\beta^0 = \eta\omega^0$, где η определена в (8), а $\omega^0 = 1$.

Алгоритм 2.1. PDMCF

При заданных $F^{-1/2}$, Y^0 и параметре $\epsilon > 0$

для $k = 0, 1, \dots$:

1. Проверить критерий остановки. Выйти и вернуть $\hat{F}^{k-1/2}$ при выполнении условия $r^k < n\tau\epsilon$;
2. Базовые обновления PDHG (6):

$$\hat{F}^{k+1/2} = \Pi(F^{k-1/2} + \alpha^k Y^k A),$$

$$F^{k+1} = 2\hat{F}^{k+1/2} - F^{k-1/2},$$

$$\hat{Y}_{ij}^{k+1} = \begin{cases} \mathbf{prox}_{\beta^k(-u_{ij})^*}(Y_{ij}^k - \beta^k(F^{k+1}A^T)_{ij}), & j \neq i, \\ 0, & j = i, \end{cases}$$

$$F^{k+1/2} = \rho\hat{F}^{k+1/2} + (1 - \rho)F^{k-1/2},$$

$$Y^{k+1} = \rho\hat{Y}^{k+1} + (1 - \rho)Y^k;$$
3. Адаптивные обновления длины шага (7) (если k кратно k^{adapt} и $\Delta_F^{k+1}, \Delta_Y^{k+1} > \tau$):

$$\omega^{k+1} = \left(\Delta_Y^{k+1}/\Delta_F^{k+1}\right)^\theta (\omega^k)^{1-\theta},$$

$$\alpha^{k+1} = \eta/\omega^{k+1}, \quad \beta^{k+1} = \eta\omega^{k+1}.$$

Инициализация. Всегда выбираем $F^{-1/2} = 0$ и $Y^0 = I - \mathbf{1}\mathbf{1}^T$. В качестве альтернативы можно использовать более подходящие значения $F^{-1/2}$ и Y , например, при теплом старте, когда задача с похожими данными уже решена. Подробнее об этом расскажем в разделе 3.1.

Критерий остановки. Поскольку $\hat{F}^{k+1/2}$ является результатом проекции на \mathcal{F} , невязка оптимальности (3) имеет вид

$$r^{k+1} = r(F^{k-1/2} + \alpha^k Y^k A).$$

Используем критерий остановки $r^k < n\epsilon$, т.е. поэлементная нормализованная невязка $r^k/n\epsilon$ должна быть меньше заданного пользователем порога ϵ .

2.6. Детали реализации

Индексация матрицы инцидентности. Храним только индексы ненулевых элементов матрицы A . Умножение матриц на A и A^T можно эффективно вычислить с помощью функций `scatter` и `gather`, которые являются высокооптимизированными ядрами CUDA и доступны в большинстве основных вычислительных языков GPU.

Проекция на масштабированный симплекс. Для вычисления μ_ℓ при $(f_\ell)_+^T \mathbf{1} > c_\ell$, следуя [42], сначала сортируем f_ℓ от наибольшего элемента к наименьшему, чтобы сформировать f'_ℓ . Затем находим наибольший индекс t , при котором $f'_{t+1} - ((\sum_{i=1}^t f'_{li} - c_\ell)/t) > 0$. Наконец, берем $\mu_\ell = (\sum_{i=1}^t f'_{li} - c_\ell)/t$.

В недавних работах предложен более эффективный метод вычисления проекции на симплексное множество (например, см. [43, 44]). Для простоты реализации в настоящей статье используется более простой алгоритм, описанный выше.

3. Эксперименты

Все эксперименты были выполнены на одном графическом процессоре N100 с 80 Гб памяти, поддерживаемом 26 виртуальными ядрами CPU и 241 Гб оперативной памяти. Ниже представлены результаты для предложенной реализации PyTorch; аналогичные результаты, описанные в Приложении 7, получены с помощью предложенной реализацией JAX.

3.1. Примеры

Данные и параметры. Рассматривались взвешенные логарифмические полезности вида $u_{ij}(T_{ij}) = w_{ij} \log T_{ij}$. Величина $\log w_{ij}$ принималась равномерной на $[\log 0,3, \log 3]$. Для топологии сети сначала создавались n двумерных точек данных $\xi_i \in \mathbf{R}^2$, каждая из которых обозначалась как (ξ_{ix}, ξ_{iy}) для $i = 1, \dots, n$. Величины ξ_{ix} и ξ_{iy} принимались равномерными на $[0, 1]$. Затем добавлялись оба ребра (ξ_i, ξ_j) и (ξ_j, ξ_i) , когда либо ξ_i входит в число q -ближайших соседей ξ_j , либо наоборот. Для каждого ребра ℓ задавалась его пропускная способность c_ℓ , где величина $\log c_\ell$ принималась равномерной на $[\log 0,5, \log 5]$.

Использовался порог критерия остановки $\epsilon = 0,01/(n(n - 1))$ для задач малой и средней размерности и $\epsilon = 0,03/(n(n - 1))$ для задач большой размерности. Результаты сравнивались с коммерческим решателем MOSEK, работающим на CPU, с настройками по умолчанию. MOSEK способен решать задачи с высокой точностью; анализ показал, что для всех экземпляров задач нормированная разность полезностей между результатами PDMCF и MOSEK не превышает примерно 0,01. Парные нормализованные (оптимальные) полезности лежали в диапазоне от 1 до 10. То есть PDMCF находит потоки, которые являются от 0,1% до 1% субоптимальными по сравнению с потоками, найденными MOSEK.

Задачи малой и средней размерности. В табл. 1 показано время работы MOSEK и PDMCF, необходимое для решения задач различных размерностей. В столбце nm указано количество скалярных переменных оптимизации в экземпляре задачи. Видно, что предложенная реализация PDMCF на GPU дает ускорение по сравнению с MOSEK от 10 до 1000 раз, причем более значительное ускорение наблюдается для экземпляров задач большей размерности. Также приведено время работы PDMCF на CPU, которое все еще ниже, чем у MOSEK, но со значительно меньшим ускорением. Аналогичная производительность наблюдается и для предложенной реализации JAX (см. Приложение 7).

Таблица 1. Время работы для задач малой и средней размерности

Размерность задачи				время (с)			
n	q	t	nm	MOSEK	PDMCF (CPU)	PDMCF (GPU)	итерации
100	10	1178	1×10^5	5	1	0,5	490
200	10	2316	5×10^5	23	2	0,7	690
300	10	3472	1×10^6	95	6	0,8	840
500	10	5738	3×10^6	340	18	1,1	950
500	20	11 176	6×10^6	1977	34	1,4	790
1000	10	11 424	1×10^7	2889	1382	19,5	7220
1000	20	22 286	2×10^7	16 765	349	5,1	1040

Задачи большой размерности. В табл. 2 показано время работы для нескольких экземпляров задач большой размерности. MOSEK не может решить все эти задачи из-за ограничений памяти. PDMCF справляется со всеми этими задачами, причем самая большая из них содержит 10^9 переменных.

Таблица 2. Время работы для задач большой размерности

Размерность задачи				время (с)			
n	q	t	nm	MOSEK	PDMCF (CPU)	PDMCF (GPU)	итерации
3000	10	34 424	1×10^8	OOM	7056	96	4140
5000	10	57 338	3×10^8	OOM	19 152	395	3970
10000	10	114 054	1×10^9	OOM	87 490	1908	4380

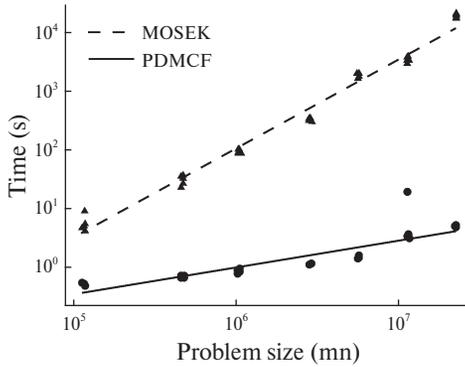


Рис. 1. Время работы для задач малой и средней размерности.

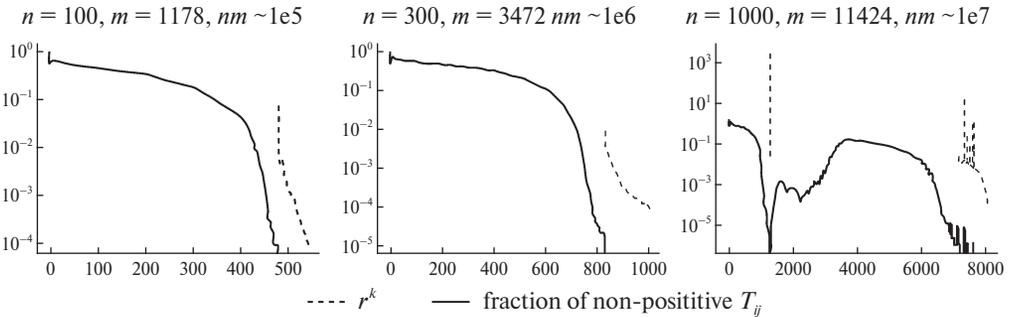


Рис. 2. График сходимости для задач малой и средней размерности.

Масштабирование. На рис. 1 представлена диаграмма времени работы для экземпляров задач малой и средней размерности. Здесь взяты 5 экземпляров задач, сгенерированных на случайной выборке из $\{0, 1, 2, 3, 4\}$ для различных значений n, q , перечисленных в табл. 1. Ось x соответствует размерности переменной оптимизации nm , а ось y – времени работы в секундах (в логарифмической шкале). Линиями показаны аффинные функции для этих данных, с наклоном около 1,5 для MOSEK и около 0,5 для PDMCF.

График сходимости. На рис. 2 показана сходимость для трех экземпляров задачи с размерностями $10^5, 10^6$ и 10^7 с PDMCF, где ось x соответствует числу итераций. Особенно на начальных итерациях наблюдается бесконечная невязка r^k , так как $-FA^T \notin \mathcal{T}$. Для этих итераций сплошной линией показана доля неположительных внедиагональных элементов в T . Для допустимых итераций прерывистой линией показана (конечная) невязка.

Теплый старт. В разделе 2.5 начинаем с некоторых простых начальных $F^{-1/2}$ и Y^0 . Также тестируем работу PDMCF с теплым стартом. На рис. 3 показано, как меняется время работы при различных теплых стартах. Чтобы сформировать эти старты, для некоторого коэффициента возмущения ν случайным образом возмущаем элементы весовой матрицы полезности, чтобы получить $\tilde{w}_{ij} = (1 \pm \nu)w_{ij}$, каждую с вероятностью 1/2. Решаем задачу о мно-

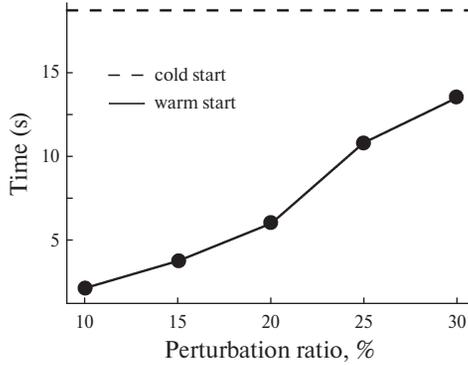


Рис. 3. График теплового старта для задачи средней размерности.

готоварных сетевых потоках с возмущенным весом полезности \tilde{w} с помощью PDMCF до тех пор, пока не попадем в допустимую точку $(F^{\text{feas}}, Y^{\text{feas}})$, удовлетворяющую $(-F^{\text{feas}} A^T)_{ij} > 0$ для всех различных i, j . Записываем прямой вес в этой точке как ω^{feas} . Затем решаем желаемую задачу о многотоварных сетевых потоках с исходным весом полезности w с $F^{-1/2} = F^{\text{feas}}, Y^0 = Y^{\text{feas}}$ и $\omega^0 = \omega^{\text{feas}}$.

Отметим, что настройка $\omega^0 = \omega^{\text{feas}}$ важна для ускорения сходимости, иначе обычно требуется такое же количество итераций для сходимости, как и при холодном старте, если просто выбрать $\omega^0 = 1$. На рис. 3 взят экземпляр задачи с $n = 1000, q = 10$. Ось x соответствует коэффициенту возмущения ν , а ось y – времени работы в секундах. Как видно, при коэффициенте возмущения $\nu = 10\%$ экономия времени работы составляет более 80%. При $\nu = 30\%$ эта экономия снижается примерно до 30%, что вполне логично, учитывая, что при большем возмущении веса полезности исходной задачи и возмущенной задачи отличаются. Поэтому ожидается, что теплый старт будет находиться дальше от оптимального решения исходного экземпляра задачи.

4. Заключение

В работе представлен алгоритм PDMCF, который ускоряет решение задач о многотоварных сетевых потоках на графических процессорах (GPU). Данный метод исходит из формулировки задачи о многотоварных сетевых потоках на основе узлов-назначений, что позволяет уменьшить количество переменных оптимизации по сравнению с классической формулировкой задачи. Затем для решения так сформулированной задачи применяется алгоритм PDHG. Эмпирические результаты подтверждают, что предложенный алгоритм дружелюбен к GPU и ускоряет вычисления (сокращает время работы) на три порядка по сравнению с классическими коммерческими решателями на базе CPU. Кроме того, этот алгоритм способен решать в десять раз большие задачи, чем те, которые могут быть решены коммерческими решателями на базе CPU.

Благодарности

Авторы благодарят Энтони Деглериса и Партха Нобеля за ценные обсуждения деталей реализации, а также Демьяна Ярмошика за очень полезные отзывы, которые помогли переработать первоначальную рукопись.

ПРИЛОЖЕНИЕ 1. Верхняя граница для $\lambda_{\max}(AA^T)$

Для направленного графа с матрицей инцидентности A обозначим через $d_i = (AA^T)_{ii}$ степень вершины i , а для $i \neq j$ величина $-(AA^T)_{ij}$ есть количество ребер, соединяющих вершины i и j , т.е. два, если оба ребра (i, j) и (j, i) существуют. Заметим, что $\lambda_{\max}(AA^T) = \max_{\|x\|_2=1} x^T(AA^T)x = \max_{x \neq 0} \frac{x^T(AA^T)x}{x^T x}$. Имеем

$$\begin{aligned} x^T(AA^T)x &= \sum_i (AA^T)_{ii}x_i^2 + \sum_{i \neq j} (AA^T)_{ij}x_i x_j = \\ &= \sum_i d_i x_i^2 + \sum_{i \neq j} (AA^T)_{ij}x_i x_j \leq \\ &\leq \sum_i d_i x_i^2 + \sum_{i \neq j} |(AA^T)_{ij}|(x_i^2/2 + x_j^2/2) = \\ &= \sum_i x_i^2(d_i + \sum_{j \neq i} |(AA^T)_{ij}|) = \\ &= \sum_i 2d_i x_i^2 \leq \\ &\leq 2d_{\max} x^T x. \end{aligned}$$

Следовательно, $\lambda_{\max}(AA^T) = \max_{x \neq 0} \frac{x^T(AA^T)x}{x^T x} \leq 2d_{\max}$.

ПРИЛОЖЕНИЕ 2. Результаты JAX

Результаты, представленные в разделе 3.1, относятся к предложенной реализации PyTorch. Здесь приводятся те же результаты для предложенной реализации JAX. В табл. 3 и 4 показано время работы на тех же экземплярах задачи, что и в табл. 1 и 2. Отметим, что JIT-компиляция (just-in-time) JAX добавляет накладные расходы на компиляцию функций в первый раз, поэтому на задачах малой размерности она работает хуже, чем ее аналог PyTorch. Время работы этих двух версий близко для задач средней и большой размерности, причем JAX работает несколько медленнее.

Таблица 3. Время работы для задач малой и средней размерности (JAX)

Размерность задачи				время (с)			итерации
n	q	m	nm	MOSEK	PDMCF (CPU)	PDMCF (GPU)	
100	10	1178	1×10^5	5	12	5	490
200	10	2316	5×10^5	23	57	6	690
300	10	3472	1×10^6	95	164	6	840
500	10	5738	3×10^6	340	548	7	950
500	20	11 176	6×10^6	1977	890	8	790
1000	10	11 424	1×10^7	2889	18 554	26	7150
1000	20	22 286	2×10^7	16 765	5143	15	1040

Таблица 4. Время работы для задач большой размерности (JAX)

Размерность задачи				время (с)			итерации
n	q	m	nm	MOSEK	PDMCF (CPU)	PDMCF (GPU)	
3000	10	34 424	1×10^8	OOM	10 6274	139	4140
5000	10	57 338	3×10^8	OOM	382 400	421	3970
10 000	10	114 054	1×10^9	OOM	1 809 517	2078	4380

СПИСОК ЛИТЕРАТУРЫ

1. *Yin P., Diamond S., Lin B., Boyd S.* Network optimization for unified packet and circuit switched networks // ArXiv Preprint ArXiv:1808.00586. 2019. <https://arxiv.org/abs/1808.00586>
2. *Bar-Gera H., Boyce D.* Origin-based algorithms for combined travel forecasting models // Transportation Research Part B: Methodological. 2003. V. 37. No. 5. P. 405–422.
3. *Boyd S., Vandenberghe L.* Convex Optimization. Cambridge: Cambridge University Press, 2004.
4. ApS MOSEK. The MOSEK optimization toolbox for MATLAB manual. Version 9.0. 2019. <http://docs.mosek.com/9.0/toolbox/index.html>
5. *Diamond S., Boyd S.* CVXPY: A Python-embedded modeling language for convex optimization // J. Machin. Learn. Res. 2016. V. 17. No. 83. P. 1–5.
6. *Ford Jr.L., Fulkerson D.* A suggested computation for maximal multi-commodity network flows // Management Science. 1958. V. 5. No. 1. P. 97–101.
7. *Hu T.* Multi-commodity network flows // Operations Research. 1963. V. 11. No. 3. P. 344–360.
8. *Gautier A., Granot F.* Forest management: A multicommodity flow formulation and sensitivity analysis // Management Science. 1995. V. 41. No. 10. P. 1654–1668.
9. *Ouorou A., Mahey P.* A minimum mean cycle cancelling method for nonlinear multicommodity flow problems // Eur. J. Oper. Res. 2000. V. 121. No. 3. P. 532–548.
10. *Manfren M.* Multi-commodity network flow models for dynamic energy management—mathematical formulation // Energy Procedia. 2012. V. 14. P. 1380–1385.
11. *Kabadurmus O., Smith A.* Multi-commodity k-splittable survivable network design problems with relays // Telecommunication Systems. 2016. V. 62. P. 123–133.

12. *Zantuti A.* The capacity and non-simultaneously multicommodity flow problem in wide area network and data flow management // Proceedings of the 18th International Conference on Systems Engineering. 2005. P. 76–80. <https://doi.org/10.1109/ICSENG.2005.81>
13. *Ereera A., Morales J., Savelsbergh M.* Global intermodal tank container management for the chemical industry // Transportation Research Part E: Logistics and Transportation Review. 2005. V. 41. P. 551–566.
14. *Mesquita M., Moz M., Paias A., Pato M.* A decompose-and-fix heuristic based on multi-commodity flow models for driver rostering with days-off pattern // European Journal of Operational Research. 2015. V. 245. No. 2. P. 423–437.
15. *Rudi A., Frohling M., Zimmer K., Schultmann F.* Freight transportation planning considering carbon emissions and in-transit holding costs: a capacitated multi-commodity network flow model // EURO J. Transport. Logist. 2016. V. 5. P. 123–160. <https://api.semanticscholar.org/CorpusID:53229695>
16. *Singh I.* A dynamic multi-commodity model of the agricultural sector: A regional application in Brazil // European Economic Review. 1978. V. 11. P. 155–179. <https://api.semanticscholar.org/CorpusID:152973282>
17. *Wagner D., Raidl G., Pferschy U., Mutzel P., Bachhiesl P.* A multi-commodity flow approach for the design of the last mile in real-world fiber optic networks // Operation Research Proceedings. 2006. <https://api.semanticscholar.org/CorpusID:17037020>
18. *Layeb S., Heni R., Balma A.* Compact MILP models for the discrete cost multicommodity network design problem // 2017 International Conference on Engineering & MIS. IEEE, 2017. P. 1–7.
19. *Salimifard K., Bigharaz S.* The multicommodity network flow problem: state of the art classification, applications, and solution methods // Operational Research. 2022. V. 22. No. 1. P. 1–47.
20. *Ouorou A., Mahey P., Vial J.* A survey of algorithms for convex multicommodity flow problems // Management Science. 2000. V. 46. No. 1. P. 126–147.
21. *Liu X., Arzani B., Kakarla S., Zhao L., Liu V., Castro M., Kandula S., Marshall L.* Rethinking machine learning collective communication as a multi-commodity flow problem // Proceedings of the ACM SIGCOMM 2024 Conference. 2024. P. 16–37.
22. *Basu P., Zhao L., Fantl J., Pal S., Krishnamurthy A., Khoury J.* Efficient all-to-all collective communication schedules for direct-connect topologies // Proceedings of the 33rd International Symposium on High-Performance Parallel and Distributed Computing. 2024. P. 28–41. <https://doi.org/10.1145/3625549.3658656>
23. *Applegate D., Díaz M., Hinder O., Lu H., Lubin M., O’Donoghue B., Schudy W.* Practical large-scale linear programming using primal-dual hybrid gradient // Neural Information Processing Systems. 2021. <https://api.semanticscholar.org/CorpusID:235376806>
24. *Lu H., Yang J.* cuPDLp.jl: A GPU implementation of restarted primal-dual hybrid gradient for linear programming in Julia // ArXiv Preprint ArXiv:2311.12180. 2024. <https://arxiv.org/abs/2311.12180>
25. *Lu H., Peng Z., Yang J.* MPAX: mathematical programming in JAX // ArXiv Preprint ArXiv:2412.09734. 2024. <https://arxiv.org/abs/2412.09734>

26. *Ryu E., Chen Y., Li W., Osher S.* Vector and matrix optimal mass transport: theory, algorithm, and applications // *SIAM J. Scientif. Comput.* 2018. V. 40. No. 5. P. A3675–A3698.
27. *Degleris A., Gamal A., Rajagopal R.* GPU accelerated security constrained optimal power flow // *ArXiv Preprint ArXiv:2410.17203.* 2024. <https://arxiv.org/abs/2410.17203>
28. *Ryu M., Byeon G., Kim K.* A GPU-accelerated distributed algorithm for optimal power flow in distribution systems // *ArXiv Preprint ArXiv:2501.08293.* 2025.
29. *Wang X., Zhang Q., Ren J., Xu S., Wang S., Yu S.* Toward efficient parallel routing optimization for large-scale SDN networks using GPGPU // *Journal Network Comput. Appl.* 2018. V. 113. P. 1–13.
30. *Kikuta K., Oki E., Yamanaka N., Togawa N., Nakazato H.* Effective parallel algorithm for GPGPU-accelerated explicit routing optimization // *2015 IEEE Global Communications Conference. IEEE,* 2015. P. 1–6.
31. *Zhang S., Ajayi O., Cheng Y.* A self-supervised learning approach for accelerating wireless network optimization // *IEEE Transactions on Vehicular Technology.* 2023. V. 72. No. 6. P. 8074–8087.
32. *Wu J., He Z., Hong B.* Chapter 5 – Efficient CUDA algorithms for the maximum network flow problem // *GPU Computing Gems Jade Edition.* Boston: Morgan Kaufmann, 2012. P. 55–66. <https://www.sciencedirect.com/science/article/pii/B9780123859631000058>
33. *Liu H., Huang S., Qin S., Yang T., Yang T., Xiang Q., Liu X.* Keep your paths free: Toward scalable learning-based traffic engineering // *Proceedings of the 8th Asia–Pacific Workshop on Networking.* 2024. P. 189–191. <https://doi.org/10.1145/3663408.3665813>
34. Gurobi Optimization. LLC Gurobi Optimizer Reference Manual. 2024. <https://www.gurobi.com>
35. *Yarmoshik D., Pershianov M.* On the application of saddle-point methods for combined equilibrium transportation models // *Proceedings of the 23rd International Conference on Mathematical Optimization Theory and Operations Research (MOTOR).* 2024. P. 432–448. https://doi.org/10.1007/978-3-031-62792-7_29
36. *Kubentayeva M., Yarmoshik D., Pershianov M., Kroshnin A., Kotliarova E., Tupitsa N., Pasechnyuk D., Gasnikov A., Shvetsov V., Baryshev L., Shurupov A.* Primal-dual gradient methods for searching network equilibria in combined models with nested choice structure and capacity constraints // *ArXiv Preprint ArXiv:2307.00427.* 2023. <https://arxiv.org/abs/2307.00427>
37. *Malitsky Y., Pock T.* A first-order primal-dual algorithm with linesearch // *SIAM J. Optim.* 2018. V. 28. No. 1. P. 411–432.
38. *Zhu M., Chan T.* An efficient primal-dual hybrid gradient algorithm for total variation image restoration // *UCLA CAM Report.* 2008. V. 34. No. 2.
39. *Chambolle A., Pock T.* A first-order primal-dual algorithm for convex problems with applications to imaging // *J. Math. Imaging Vision.* 2011. V. 40. P. 120–145.
40. *Chambolle A., Pock T.* On the ergodic convergence rates of a first-order primal-dual algorithm // *Mathematical Programming.* 2015. V. 159. P. 253–287.
41. *Parikh N., Boyd S.* Proximal algorithms // *Foundations and Trends in Optimization.* 2014. V. 1. No. 3. P. 127–239.

42. *Held M., Wolfe P., Crowder H.* Validation of subgradient optimization // *Mathematical Programming.* 1974. V. 6. P. 62–88.
43. *Duchi J., Shalev-Shwartz S., Singer Y., Chandra T.* Efficient projections onto the l_1 -ball for learning in high dimensions // *Proceedings of the 25th International Conference on Machine Learning.* 2008. P. 272–279.
<https://doi.org/10.1145/1390156.1390191>
44. *Condat L.* Fast projection onto the simplex and the l_1 ball // *Mathematical Programming.* 2016. V. 158. No. 1-2. P. 575–585.
<https://doi.org/10.1007/s10107-015-0946-6>

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

Поступила в редакцию 03.03.2025

После доработки 20.05.2025

Принята к публикации 27.06.2025

СОДЕРЖАНИЕ

Тематический выпуск

Щербаков П.С. К 90-летию Б.Т. Поляка	3
Акинфиев И.А., Граничин О.Н., Тарасова Е.Ю. Поисковый метод стохастической нестационарной оптимизации функции с гильдеровским градиентом	6
Бекри Я., Немировский А., Юдицкий А. О надежном восстановлении сигналов по непрямым наблюдениям	32
Мирасьерра В., Маммарелла М., Даббене Ф., Аламо Т. Точные аппроксимации множеств с вероятностными ограничениями с помощью пакетного вероятностного масштабирования	60
Соколов В.Ф. Оптимальное робастное слежение для дискретного минимально-фазового объекта при неизвестных смещении и норме внешнего возмущения и нормах неопределенностей	82
Хильдебранд Р., Шикаке Мапунгвана Т. О проблеме оптимального управления с управлением в диске	99
Чжан Ф., Бойд С. Решение задач о многотоварных сетевых потоках большой размерности на графических процессорах	116

C O N T E N T S

Topical issue

Shcherbakov P.S. To the 90th Anniversary of B.T. Polyak	3
Akinfiyev I.A., Granichin O.N., Tarasova E.Y. Search Method for Stochastic Non-Stationary Optimization of a Function with a Hölder Gradient	6
Bekri Ya., Nemirovski A., Juditsky A. On Robust Recovery of Signals from Indirect Observations	32
Mirasierra V., Mammarella M., Dabbene F., Alamo T. Tight Approximations of Chance Constrained Sets through Pack-Based Probabilistic Scaling	60
Sokolov V.F. Optimal Robust Tracking for Discrete-Time Minimum-Phase Plant under Unknown Bias and Norms of External Disturbance And Uncertainties	82
Hildebrand R., Chikake Mapungwana T. On Optimal Control Problems with Control in a Disc	99
Zhang F., Boyd S. Solving Large Multicommodity Network Flow Problems on GPUS	116