

ОТ РЕДАКЦИИ

За последние 10 лет в России заметно увеличилось число молодых специалистов, которые активно работают в области численных методов оптимизации и их различных приложениях к задачам анализа данных и моделирования. Особенно много молодежи, работающей в данных направлениях, сконцентрировалось в школе Прикладной математики и информатики (ПМИ) МФТИ. Это заслуга многих людей, среди которых особо хочется выделить двух: проф. Б. Т. Поляка и проф. Ю. Е. Нестерова. С 2011 по 2015 г. Борис Теодорович и Юрий Евгеньевич активно работали в лаборатории ПреМоЛаб МФТИ (рук. — проф. В. Г. Спокойный), увлекая молодежь своими идеями и своим примером. Плоды той работы привели к тому, что каждый год появляется несколько десятков статей по оптимизации от молодых сотрудников школы ПМИ МФТИ (среди которых стабильно есть статьи на ICML, NeurIPS, AISTATS и в различные журналы уровня Q1), при поддержке школы ПМИ МФТИ в России проводятся крупные международные конференции по численным методам оптимизации, организован общероссийский семинар по оптимизации, выигран большой грант РНФ по данному направлению (рук. — проф. А. М. Райгородский, среди основных участников — проф. Б. Т. Поляк, акад. Ю. Г. Евтушенко), в работу которого вовлечена активная молодежь. Все это дает возможность периодически делать специальные выпуски журнала, всецело посвященные вопросам, связанным с современными численными методами оптимизации. Настоящий специальный выпуск является первым таким сборником, в который были приглашены в основном молодые, активно работающие сотрудники школы ПМИ МФТИ, как правило аспиранты или недавно защитившие диссертации кандидаты наук. Отобранные в этот сборник статьи в целом неплохо отражают современные тенденции развития области численных методов выпуклой оптимизации. Практически во всех работах речь идет именно о выпуклых задачах. Основная причина рассмотрения выпуклых задач — возможность достаточно глубокого теоретического анализа. Отобранные статьи в основном сконцентрированы именно на теоретических аспектах анализа предлагаемых методов. Перечислим вкратце работы, вошедшие в этот сборник.

В статье Артёма Агафонова «Нижние оценки для методов типа условного градиента для задач минимизации гладких сильно выпуклых функций» исследуются методы типа условного градиента, т. е. использующие линейный минимизационный оракул (без проектирования на допустимое множество). Такие методы хорошо себя зарекомендовали при решении различных задач выпуклой минимизации на множествах типа единичного симплекса. В своей статье Артём исследует вопрос об эффективности таких методов для задач сильно выпуклой оптимизации. В статье показано, что без дополнительных предположений не стоит ожидать, что методы типа условного градиента будут эффективны для сильно выпуклых задач. Причина связана с наличием полученной в статье нижней оценки на число вызовов линейного минимизационного оракула, в которую входит корень из размерности пространства.

В статье Мохаммада Алкусы с соавторами «Подход к решению невыпуклой равномерно вогнутой седловой задачи со структурой» исследуются седловые задачи, равномерно вогнутые по max-группе переменных и невыпуклые по min-группе переменных. Такие задачи интенсивно исследовались в предыдущие пару лет. Многие важные результаты для таких задач уже были получены в случае, когда равномерная выпуклость заменялась обычной выпуклостью. В данной работе при разработке численного метода использовалось сочетание тензорного метода для равномерно вогнутой max-задачи с невыпуклым градиентным спуском для min-задачи. Отметим, что первый автор данной статьи Мохаммад Алкуса учился в аспирантуре школы ПМИ, успешно защитил кандидатскую диссертацию, затем вернулся в Сирию и продолжает активно работать в области численных методов оптимизации.

В статье Александры Базаровой, выполненной под руководством одного из самых активных сейчас молодых оптимизаторов школы ПМИ МФТИ, Александра Безносикова (об активности может сказать, например, такой факт, что на конференцию ICML 2022 Александр направил 7 статей, в большинстве из которых он первый автор), рассматривается задача невыпуклой оптимизации, которая имеет специальную структуру — целевой функционал можно представить как выпуклую функцию «с небольшой рябью». Хотя формально рассматриваемый класс задач невыпуклый, в работе Александры с соавторами «Линейно сходящиеся безградиентные методы минимизации параболической аппроксимации» показывается, как можно использовать специфику постановки задачи, чтобы получить более богатую теорию для данного класса задач, по сравнению с общим классом задач не выпуклой оптимизации. Отметим здесь, что Александр Безносиков за свои прорывные результаты на днях был номинирован на наивысшую стипендию школы ПМИ по оптимизации. Эта стипендия вручается наиболее активным студентам школы ПМИ, которые пишут статьи в области численных методов оптимизации. В этом семестре 16 студентов школы были номинированы на стипендии разного уровня.

В статье Егора Гладина и Екатерины Бородич «Редукция дисперсии для минимаксных задач с небольшой размерностью одной из переменных» исследуются седловые выпукло-вогнутые задачи с композитной структурой. Предполагается, что седловая функция имеет структуру суммы (с большим числом слагаемых). Другой важной особенностью постановки задачи является предположение о небольшой размерности \min -переменных. Отмеченные особенности приводят к необходимости сочетания приема редукции дисперсии с методами типа центров тяжести (в работе используется метод Вайды), пригодных для эффективного решения задач в пространствах небольшой размерности. В статье Егора и Екатерины приводится возможный способ такого сочетания. Работа существенно базируется на анализе скорости сходимости метода Вайды, проделанном ранее Е. Л. Гладиным.

В статье Марины Даниловой (аспирантка Б. Т. Поляка) и Григория Малиновского «Метод тяжелого шарика с усреднением» исследуется усредненный (по итерациям) метод тяжелого шарика Поляка. Показывается положительный эффект от усреднения. Данная работа продолжает большой цикл (в основном) западных работ по изучению данного метода. Статья содержит ряд интересных новых наблюдений, которые, с одной стороны, хорошо согласуются с полученными ранее результатами, с другой — удачным образом их дополняют.

В работе Дарины Двинских с соавторами «О связях задач стохастической выпуклой минимизации с задачами минимизации эмпирического риска на шарах в p -нормах» обобщаются недавние результаты ее диссертации (рук. — проф. В. Г. Спокойный), посвященной в том числе вопросам связи онлайн-подходов к решению задач выпуклой стохастической оптимизации (к таким подходам, например, относится метод стохастического градиентного спуска) и офлайн-подходов (методов Монте-Карло) на случай, когда константы липшицевости и сильной выпуклости определяются в произвольных нормах (не обязательно евклидовых). С точки зрения офлайн-подхода изучен класс задач стохастической выпуклой оптимизации с острым минимумом (с произвольным выбором нормы). Впервые исследован вопрос о связи онлайн- и офлайн-подходов для сильно выпукло-вогнутых стохастических седловых задач.

В статье Павла Двуреченского (первого в России доктора наук по направлению Computer Science) «Градиентный метод с неточным оракулом для задач композитной невыпуклой оптимизации» изучается скорость сходимости проксимального градиентного метода для задач невыпуклой композитной оптимизации в условиях, когда градиент целевой функции удовлетворяет условию Гельдера и при этом доступен только неточный градиент. Несмотря на кажущуюся специфичность задачи, в такую постановку (ввиду достаточно большой общности) погружается много важных классов прикладных задач. В частности, в работе Павла предложен эффективный

метод поиска экстремальных точек для невыпуклых функционалов с гильбертовым градиентом в условиях шума и неевклидова проектирования.

В работах Екатерины Котляровой с соавторами изучаются модели равновесного распределения транспортных потоков по путям. В первой работе «Обоснование связи модели Бэкмана с вырождающимися функциями затрат с моделью стабильной динамики» исследуется связь двух моделей (задач выпуклой оптимизации): Бэкмана и Нестерова – деПальмы. Исследуется вопрос о возможности получения одной модели (задачи оптимизации) как предела по параметру другой модели (параметризованной задачи оптимизации). Во второй статье «Ускорение работы двухстадийной модели равновесного распределения потоков по сети» изучаются различные алгоритмы поиска равновесия в упомянутых ранее двух моделях. Изучается вопрос о том, в какой степени стабилизация потоков по ребрам по ходу работы алгоритмов может сокращать время работы алгоритма Дейкстры или его более специализированных аналогов. На каждой новой итерации метода необходимо заново искать всевозможные кратчайшие пути из различных источников в различные стоки. Но на старших итерациях за счет стабилизации потоков по ребрам, веса ребер также стабилизируются. В какой степени это наблюдение может сократить вычисления на перезапуск алгоритма Дейкстры или его аналогов? Во второй статье предпринята попытка дать ответ на этот вопрос. Результаты этих статей развивают (пополняют) результаты недавней монографии одного из соавторов этих работ — Е. В. Гасниковой.

В работах Петра Остроухова с соавторами исследуются тензорные методы. После работы Ю. Е. Нестерова 2018 года интерес к использованию методов высокого порядка (старшие производные целевой функции) для решения достаточно гладких задач выпуклой оптимизации резко возрос. В первой статье П. А. Остроухова с соавторами «Тензорные методы для сильно выпуклых сильно вогнутых седловых задач и сильно монотонных вариационных неравенств» исследуется вопрос о распространении тензорных методов на седловые (сильно) выпукло-вогнутые задачи. Полученные здесь результаты говорят о том, что такого же большого эффекта, который имел место для задач выпуклой оптимизации для седловых задач, вряд ли стоит ожидать. То есть результаты скорее отрицательные. Однако результаты представляются достаточно ценными, поскольку это направление рассматривалось как естественное и перспективное продолжение исследований, начатое Ю. Е. Нестеровым. Во второй статье «Тензорные методы внутри смешанного оракула для решения задач типа min-min» Пётр исследует задачу выпуклой минимизации со специальной структурой min-min. По первой группе min-переменных функционал (сильно) выпуклый, и размерность этой группы предполагается большой. По второй группе min-размерность переменных небольшая, функционал предполагается (сильно) выпуклым и достаточно гладким. Пётр показывает, как можно сочетать быстрый градиентный метод для первой группы переменных с тензорными методами для второй группы переменных. В случае если внутренняя задача определена на всем пространстве, используется супербыстрый тензорный метод третьего порядка (предложенный Ю. Е. Нестеровым и активно развиваемый в школе ПМИ МФТИ). Если же внутренняя задача определена на компакте, используется ускоренный тензорный метод для задачи с композитом.

В статье Дмитрия Пасечнюка с соавторами (А. В. Алпатов и др., «Стохастическая оптимизация в задаче цифрового предсказания сигнала») исследуется задача обучения модели цифрового преобразователя усилителя сигнала на базовой станции. Проводится сопоставительный анализ различных численных подходов к данной задаче. Изучаются как вопросы непосредственно оптимизационного характера, так и вопросы, связанные с переобучением. Работа довольно техническая, поскольку связана с вполне конкретной реальной задачей, поставленной отделом Андрея Воробьева и Евгения Яницкого компании «Хуавей». Дмитрий был активным участником проекта. Интересно отметить, что соавторами Дмитрия являются школьники, которые делали летом 2021 года под руководством Дмитрия этот проект в рамках проектной смены «Большие

вызовы» (рук. направления — проф. А. М. Райгородский). Также интересно отметить, что сам Дмитрий является студентом первого курса школы ПМИ МФТИ, что не помешало иметь ему уже 20 статей по численным методам оптимизации, среди которых есть статья в журнале уровня Q1.

В статье Никиты Плетнева с соавторами «Применение градиентных методов оптимизации для решения задачи Коши для уравнения Гельмгольца» изучается некорректная обратная начально-краевая задача для линейного эллиптического уравнения. Никита исследует вопрос о том, какая концепция неточного градиента более адекватна при анализе ускоренных градиентных методов решения таких задач. Точнее говоря, исходная задача сводится к задаче квадратичной оптимизации в гильбертовом пространстве. Однако для вычисления действия линейного оператора (и его сопряженного), входящего в функционал, требуется решить начально-краевую задачу для эллиптического уравнения, что в общем случае может быть сделано только численно, то есть заведомо неточно. Существуют как минимум две конкурирующие модели неточности в градиенте (по Б. Т. Поляку): аддитивная неточность, относительная аддитивная неточность. Последняя неточность дружественнее (меньше портит градиентного типа методы). Численные исследования, проведенные Никитой, показывают, что для рассматриваемой задачи неточность в градиенте скорее аддитивная, нежели относительная аддитивная. Также показано, что ускоренные методы Нестерова работают на таких задачах заметно лучше неускоренных градиентных методов, несмотря на наличие шума в градиенте.

В первой статье Фёдора Стонякина с соавторами (О. С. Савчук и др., «Адаптивные методы первого порядка для относительно сильно выпуклых задач оптимизации») впервые предложены адаптивные градиентные методы для задач минимизации относительно сильно выпуклых функций с гарантиями оптимальной скорости сходимости на классе относительно липшицевых (непрерывных) функций. Такой класс задач был предложен несколько лет назад Ю. Е. Нестеровым и оказался довольно важным в целом ряде приложений. Часть предложенных методов приводит к оптимальным оценкам сложности как на классе относительно гладких, так и на классе относительно липшицевых задач. Во второй работе Фёдора Стонякина с соавторами (С. С. Аблаев и др., «Субградиентные методы для задач негладкой оптимизации с некоторой релаксацией условия острого максимума») исследуется класс задач выпуклой оптимизации с некоторой релаксацией условия «острого минимума». Этот класс задач был введен еще в 60-е годы XX века Б. Т. Поляком. Отметим интересные связи предложенного обобщения с возникшей несколько лет назад в оптимизации концепцией неточного оракула Деволдера – Глинера – Нестерова. В статье исследован комплекс вопросов разработки эффективных методов для соответствующих типов негладких задач: сочетание возможности использования в оценке скорости сходимости адаптивно подбираемых параметров, «неточности» рассматриваемого обобщения острого минимума, применимость к некоторым типам невыпуклых задач, а также возможность использования на итерациях неточного субградиента.

В статье Назария Тупицы «Об адаптивных ускоренных методах и их модификациях для альтернативной минимизации» развивается направление исследований, которое легло в основу диссертации Назария (рук. — проф. В. Г. Спокойный): ускоренные методы альтернированных направлений. Методы альтернированных направлений и их ускоренные варианты, как правило, заметно лучше (быстрее) сходятся на практике, чем обычные градиентные методы и их ускоренные варианты, хотя в теоретическом плане никакой разницы обнаружить не удастся. Исследование Назария — это попытка объяснить описанное наблюдение за счет лучшей адаптивной настройки методов альтернированных направлений на локальные константы сильной выпуклости.

А. В. Гасников, А. И. Лобанов