

Заключение диссертационного совета ФПМИ.1.2.3.001 от 28 декабря 2023 года о присуждении Остроухову Петру Алексеевичу (гражданство РФ) ученой степени кандидата физико-математических наук

Диссертация «Применение методов высокого порядка к задачам оптимизации со структурными особенностями» принята к защите 17 октября 2023 года решением экспертного совета по физико-математическим наукам Аттестационной комиссии МФТИ, протокол №ФМ-2023/19.

Соискатель Остроухов Петр Алексеевич, 1995 года рождения.

В 2019 г. Остроухов П.А. окончил федеральное государственное автономное образовательное учреждение высшего образования «Московский физико-технический институт (государственный университет)» по специальности «Прикладная математика и физика». В 2023 г. соискатель окончил очную аспирантуру федерального государственного автономного образовательного учреждения высшего образования «Московский физико-технический институт (национальный исследовательский университет)» по направлению подготовки «09.06.01 Информатика и вычислительная техника». Справка о сдаче кандидатских экзаменов выдана в 2023 г. федеральным государственным автономным образовательным учреждением высшего образования «Московский физико-технический институт (национальный исследовательский университет)».

Диссертация «Применение методов высокого порядка к задачам оптимизации со структурными особенностями» прошла апробацию на кафедре интеллектуальных систем федерального государственного автономного образовательного учреждения высшего образования «Московский физико-технический институт (национальный исследовательский университет)» Министерства науки и высшего образования Российской Федерации.

Научный руководитель – доктор физико-математических наук, доцент, Гасников Александр Владимирович. Основное место работы – федеральное государственное автономное образовательное учреждение высшего образования «Московский физико-технический институт (национальный исследовательский университет)», кафедра математических основ управления.

Выбор организации Федеральное государственное автономное образовательное учреждение высшего образования «Омский государственный технический университет» в качестве ведущей организации обосновывается их согласием и компетентностью, подтвержденной публикациями, которые соответствуют тематике диссертации, а также способностью определить научную и практическую ценность работы.

Ведущая организация Федеральное государственное автономное образовательное учреждение высшего образования «Омский государственный технический университет» в своем положительном отзыве, подписанном Зыкиной Анной Владимировной, доктором физико-математических наук, профессором, заведующим кафедрой «Прикладная математика

и фундаментальная информатика», указала, что:

Актуальность темы

Диссертация затрагивает две основные темы: методы оптимизации высокого порядка для различных задач оптимизации с особенностями структуры и применение метода Франк-Вульфа к самосогласованным функциям. Обе эти темы актуальны. Методы, использующие производные порядка p , где p больше 2, стали широко развиваться после революционной работы Юрия Нестерова 2018 года. В этой работе он показал, как правильно регуляризовать аппроксимацию Тейлора целевой функции так, чтобы сделать ее выпуклой. Это сделало задачу минимизации данного приближения решаемой. Вторая тема также актуальна ввиду того, что многие современные задачи машинного обучения являются самосогласованными. Метод Франк-Вульфа впервые был предложен Маргаритой Франк и Филипом Вульфом в 1956 году. Привлекательность метода заключается в его применимости к различным задачам машинного обучения, ограниченным некоторым выпуклым множеством. Проблема применения метода Франк-Вульфа к самосогласованным функциям заключается в том, что этот метод опирается на наличие ограниченной константы кривизны, но самосогласованные функции часто не обладают таким свойством.

Научная новизна

Все результаты диссертации являются новыми. Прежде всего, данная работа расширяет область применения методов высокого порядка с учетом структуры, отличной от обычной выпуклой оптимизации без ограничений. В ней предложены новые почти-оптимальные методы для поиска стационарных точек выпуклых функций. Кроме того, в данной работе методы высокого порядка впервые внедряются в область задач сильно выпуклых сильно вогнутых седловых точек. Также впервые исследуется применимость методов высокого порядка в составе смешанных оракулов для решения задач минимизации типа $\min\text{-}\min$. Кроме того, в работе предлагаются новые модификации метода Франк-Вульфа, позволяющие его применять для решения самосогласованных задач.

Цели работы

Цели исследования охватывают широкий спектр задач оптимизации. Во-первых, внимание сосредотачивается на применении методов высокого порядка к выпуклым задачам с ограничениями типа равенств, исследуя два основных подхода: поиск стационарных точек двойственной задачи и прямо-двойственный метод. Цель состоит в том, чтобы подробно исследовать и численно сравнить оба этих подхода с точки зрения применения методов высокого порядка. Во-вторых, исследуется применимость методов высокого порядка для решения сильно выпуклых, сильно вогнутых седловых задач, а также сильно монотонных вариационных неравенств. Это включает в себя поиск приближенных стационарных точек для этих задач. Кроме того, целью является

внедрение методов высокого порядка в смешанные оракулы для задач типа min-min, что позволяет разделить задачу на две подзадачи, каждая из которых решается различными методами. Наконец, исследования направлены на внедрение новых модификаций метода Франк-Вульфа, адаптируя его для самосогласованных функций с особым акцентом на достижение сублинейной или линейной сходимости.

Теоретический и практический вклад работы

Достижение поставленных целей приводит к следующей теоретической и практической значимости. Во-первых, уменьшается разрыв между существующими методами и нижними оценками в задаче поиска стационарных точек в выпуклых функциях методами высокого порядка, что отражает значительное развитие методов оптимизации. Во-вторых, вводятся инновационные методы, адаптированные для конкретных классов задач с лучшей скоростью сходимости по сравнению с аналогами. Эти методы представляют собой многообещающий подход к решению задач со структурными особенностями. Кроме того, предлагаемые методы оптимизации высокого порядка обладают потенциалом для решения различных задач оптимизации с меньшим количеством итераций по сравнению с существующими подходами, что может потенциально сократить затраты времени и вычислительных ресурсов в таких областях, как машинное обучение, финансы, инженерия и научные эксперименты. Это может помочь исследователям и разработчикам выбрать подходящий метод оптимизации для своих конкретных задач. Наконец, интеграция методов Франк-Вульфа с самосогласованными функциями, часто встречающимися в реальных приложениях, предлагает практичный и эффективный подход к их оптимизации. Известно, что у самосогласованных функций часто отсутствует сильная выпуклость и гладкость, что делает обычные методы первого порядка менее практичными. Внутренняя природа метода Франк-Вульфа позволяет исследовать небольшое количество вершин, чья выпуклая комбинация образует множество, на котором определена задача.

Содержание диссертации

В главе 1 (введение) представлен обзор научной литературы по изучаемой проблеме, формулируются цели, научная новизна, теоретическая и практическая ценность представленной диссертации.

В главе 2 исследуются выпуклые задачи оптимизации с линейными ограничениями типа равенств. Вводится метод высокого порядка для минимизации нормы градиента и прямо-двойственный тензорный метод. Эти два метода основаны на двух различных подходах к решению рассматриваемой задачи: поиске стационарной точки двойственной функции с последующим восстановлением прямого решения и прямо-двойственному подходу, который решает одновременно и прямую, и двойственную задачи.

В главе 3 рассматриваются сильно выпуклые сильно вогнутые седловые задачи.

Предлагаются два новых метода для решения этой задачи. Один из них основан на методе рестартов, применяемом к другому существующему алгоритму. Другой – это модификация первого, но с переключением на другой алгоритм в его области квадратичной сходимости. Наконец, последний из предлагаемых методов также модифицирован для решения задачи минимизации нормы градиента для упомянутой функции.

В главе 4 исследуется применимость методов высокого порядка в смешанных оракулах для задач типа $\min\text{-}\min$, где вся задача может быть разделена на две подзадачи по разным группам переменных. Внешняя задача рассматривается на компакте, а внутренняя либо на компакте, либо на всем пространстве. Внешняя задача решается с использованием метода быстрого градиента, а внутренняя либо с помощью композитного тензорного метода (для компакта), либо с помощью супербыстрого тензорного метода (для всего пространства).

Глава 5 единственная, не посвященная методам высокого порядка. Здесь предлагаются две новые стратегии выбора шага в методе Франка-Вульфа. Эти модификации позволяют применять метод Франка-Вульфа к самосогласованным функциям, которые не обладают гладкостью по Липшицу.

Комментарии и предложения

1. Было бы интересно увидеть в главе 2 прямо-двойственный тензорный метод, ускоренный техникой Монтейро-Свайтера, которая дает лучшие результаты сходимости, чем ускорение Нестерова. Затем сравнить полученные алгоритмы в численных экспериментах.
2. В главах 3 и 4 отсутствуют численные эксперименты. Таким образом, несмотря на превосходство предложенных методов в некоторых сценариях, было бы интересно узнать, показывают ли они хорошие результаты на практике.

Эти замечания являются незначительными и не связаны с результатами диссертации, которые составляют основной вклад работы.

Автореферат полностью и точно отражает содержание диссертации.

Диссертация «Применение методов высокого порядка к задачам оптимизации со структурными особенностями» соответствует критериям, установленным п. 3 Положения о присуждении ученых степеней, предъявляемым к кандидатским диссертациям в МФТИ. Остроухов Петр Алексеевич заслуживает присуждения ученой степени кандидата физико-математических наук по специальности 1.2.3. Теоретическая информатика, кибернетика.

Основные результаты диссертации опубликованы в следующих 3 работах соискателя:

- [Индексируется базой данных Scopus] Ostroukhov, P.A., Kamalov, R.A., Dvurechenskii, P.E. and Gasnikov, A.V., 2022. Tensor methods for strongly convex strongly concave saddle point problems and strongly monotone variational inequalities. Компьютерные исследования и моделирование, 14(2), pp.357-376.
- [Индексируется базой данных Scopus] Остроухов, П.А., 2022. Тензорные методы внутри смешанного оракула для решения задач типа min-min. Компьютерные исследования и моделирование, 14(2), pp.377-398.
- [Индексируется базой данных Scopus] Dvurechensky, P., Ostroukhov, P., Safin, K., Shtern, S. and Staudigl, M., 2020, November. Self-concordant analysis of Frank-Wolfe algorithms. In International Conference on Machine Learning (pp. 2814-2824). PMLR.

В заключениях членов диссертационного совета по содержанию диссертации работа охарактеризована положительно, как выполненная на необходимом научном уровне, обладающая как теоретической, так и практической значимостью. Замечания, отраженные в заключениях:

- 1) Диссертация, в основном, фокусируется на получении теоретических результатов, и численные эксперименты приведены только во второй и пятой главах. Почему автор не предоставил результаты экспериментов для методов высокого порядка на седловых задачах и задачах типа minmin?
- 2) Может ли смешанный оракул, предложенный в этой работе для задач типа minmin, быть распространен на седловые задачи?
- 3) Существуют ли нижние оценки для седловых задач с гладкостью высокого порядка?
- 4) Если таковые существуют, совпадают ли предлагаемые в данной диссертации верхние оценки с нижними оценками?
- 5) В четвертой главе вы упоминаете, что для достижения линейной сходимости для метода Франк-Вульфа вам необходимо иметь некоторый локальный оракул линейной минимизации. В чем основное препятствие использования этого оракула на практике и почему вы используете его только для решения задачи оптимизации портфолио?
- 6) Вы используете почти оптимальный тензорный метод внутри предложенной вами "обертки" для поиска стационарной точки выпуклых функций. Возможно ли вместо этого использовать оптимальный тензорный метод для выпуклых функций, предложенный недавно в статье Ковалева и Гасникова?
- 7) Вы используете довольно много предположений в третьей главе, когда говорите о смешанном оракуле для задач типа min-min. Возможно ли как-то облегчить эти ограничения?
- 8) Возможно ли расширить данный подход на выпуклые и невыпуклые задачи?

Диссертационный совет отмечает, что в результате проведенных соискателем исследований получены следующие результаты:

1. Предложены прямо-двойственный метод высокого порядка и почти-оптимальный метод минимизации нормы градиента. Данные методы

разрабатывались применительно к задаче выпуклой оптимизации с линейными ограничениями типа равенств.

2. Предложен метод высокого порядка для решения сильно-выпуклых сильно-вогнутых седловых задач или поиска решений сильно-монотонных вариационных неравенств. Предлагаемый метод обладает локальной квадратичной сходимостью.
3. Предложен метод, основанный на смешанном оракуле, для задач типа min-min, где внутренняя задача решается методом высокого порядка.
4. Предлагаются модификации метода Франк-Вульфа для самосогласованных функций. Так как на практике многие задачи не обладают свойством гладкости, но являются самосогласованными, данный результат может быть очень полезен в индустриальных задачах.

Теоретическая и практическая значимость исследований обоснована тем, что

- Тема работы связана с применением методов высокого порядка, чья скорость сходимости, как известно, выше, чем у методов градиентного типа.
- В работе предлагаются методы высокого порядка применительно к различным задачам оптимизации, чьи структурные особенности позволяют получать оценки сложности лучше, чем у существующих аналогов.
- Также предлагаются модификации метода Франк-Вульфа применительно к самосогласованным функциям. Предлагаемые методы как закрывают теоретический вопрос о возможности применения данного метода к упомянутому классу задач, так и полезны на практике, так как многие индустриальные задачи обладают свойством самосогласованности.

Оценка достоверности результатов исследования выявила следующее

- Результаты являются теоретически обоснованными, а также они подкреплены численными экспериментами.
- Результаты диссертации опубликованы в рецензируемых научных изданиях.
- Результаты диссертации докладывались на международных научных конференциях.

Личный вклад соискателя состоит в том, что ключевые результаты получены автором работы лично.

На заседании 28 декабря 2023 г. диссертационный совет принял решение присудить Остроухову Петру Алексеевичу учёную степень кандидата физико-математических наук по специальности 1.2.3. Теоретическая информатика, кибернетика.

При проведении тайного голосования диссертационный совет в количестве 5 человек, из 5 человек, входящих в состав, проголосовали: за – 5, против – 0, недействительных бюллетеней – 0.

Председатель диссертационного совета:



И.Б. Петров

Учёный секретарь:
28.12.2023



К.Ю. Войтиков