

**ОТЗЫВ официального оппонента
о диссертации Гуртовой Ольги Владимировны
«Методы онлайн оптимизации квадратичной функции потерь, основанные
на использовании случайных признаков Фурье»,
представленной на соискание ученой степени
кандидата физико-математических наук по специальности
1.2.2. «Математическое моделирование, численные методы и комплексы
программ»**

Успехи современных систем искусственного интеллекта достигнуты в значительной степени благодаря тому, что обучение моделей происходит на данных очень большого объема. Данные оказываются распределенными по времени поступления и по месту хранения. С учетом этого методы построения моделей при последовательном поступлении данных становятся особенно значимыми. Большие размерности признаков пространств и другие объемные характеристики моделей заставляют с особым вниманием относиться к эффективности вычислительных алгоритмов. Во многих случаях объем вычислений, необходимых для построения канонически точных моделей оказывается неприемлемым. Использование стохастических методов (случайные выборки и т.п.) позволяет эффективно строить модели, обладающие необходимыми свойствами.

В этом русле актуальных современных исследований выполнена и диссертационная работа О.В.Гуртовой, посвященная разработке и анализу вычислительно эффективных методов онлайн оптимизации для задач регрессии. Данные для построения модели могут поступать последовательно и характеризуются сложными нелинейными зависимостями. Актуальность данного исследования обусловлена стремительным развитием областей, требующих обработки потоковых данных в реальном времени — от финансового прогнозирования и интернета вещей до адаптивных систем управления и рекомендательных сервисов.

Классические подходы, основанные на пакетной обработке, требуют полного переобучения на всех накопленных данных, что является вычислительно затратным и неприемлемым для многих практических приложений. В этом контексте онлайн оптимизация, исследуемая в работе, представляет собой одно из наиболее перспективных направлений. Особую сложность представляет работа с нелинейными моделями, где методы, основанные на ядрах (kernel methods), сталкиваются с проблемой высокой вычислительной сложности. Использование техники случайных признаков Фурье, являющейся центральным элементом диссертации, представляет собой

современный и эффективный подход к решению этой проблемы, позволяя свести нелинейную задачу к линейной регрессии в конечномерном пространстве признаков. Таким образом, разработка новых теоретически обоснованных и практически применимых онлайн алгоритмов на основе этой идеи является важной и актуальной научной задачей.

Теоретической основой работы служат фундаментальные труды в области онлайн выпуклой оптимизации, восходящие к работам Дж. Ханнана, Д. Блэкуэлла, Н. Литтлстоуна и В. Вовка, и систематизированные в классических монографиях Н. Чеза-Бьянки, Г. Лугози и Ш. Шалева-Шварца. Традиционно для работы с нелинейными зависимостями применялись ядерные методы, однако их масштабируемость была ограничена. Значительным прорывом, ставшим методологическим ядром данной диссертации, стала разработка метода случайных признаков Фурье А. Рахими и Б. Рехтом. Этот подход позволяет с высокой точностью аппроксимировать ядерные функции, заменяя работу в бесконечномерных пространствах на эффективную линейную регрессию в конечномерном пространстве случайных признаков.

Диссертация О.В. Гуртовой развивает эти идеи, интегрируя технику случайных признаков Фурье с современными алгоритмами онлайн агрегирования, такими как метод Вовка-Азури-Вармута (VAW), и применяя их к решению сложных задач, включая регрессию на данных с марковской зависимостью и многоядерное обучение. Таким образом, работа находится на пересечении нескольких актуальных направлений исследований и вносит вклад в создание теоретически обоснованных и практически применимых вычислительных методов.

Диссертационная работа состоит из введения, трёх глав, заключения, списка литературы и приложений.

Во введении обозначена проблематика, сформулированы цели и задачи, обоснована научная новизна и практическая значимость. Приведен краткий обзор современных направлений в онлайн оптимизации, многоядерном обучении и автоматизированном обучении (AutoML).

В главе 1 исследуется задача оценивания условного математического ожидания вида $h(x) = E[g(X_0, \dots, X_S) | X_0 = x]$ для стационарного марковского процесса по единственной наблюдаемой траектории. Предполагается что процесс X_t удовлетворяет условию -перемешивания. Указанная задача сводится к задаче регрессии, где в качестве признаков выступают значения X_t , а в качестве целевых переменных – величины $Y_t = g(X_t, \dots, X_{t+s})$. Для ее решения применяется алгоритм Вовка–Азури–Вармута (VAW), используемый в сочетании с нелинейными признаками.

В теореме 1.2 установлено, что при прямом применении метода оценка ожидаемых потерь полиномиально зависит от числа признаков m : $O(m^3/n)$, где n – количество данных. Этот результат выявляет ключевой недостаток такого подхода: его высокую вычислительную сложность и, как следствие, практическую неприменимость для моделей с большим числом признаков.

Для преодоления указанного недостатка в работе применяется техника случайных признаков Фурье. Основным результатом главы является теорема 1.3, в которой для случая, когда истинная регрессионная функция f принадлежит соответствующему гильбертову пространству с воспроизводящим ядром (RKHS), получена оценка сожаления порядка $O(\|f\|_p^2/\sqrt{n})$. Данная оценка не зависит полиномиально от размерности m , а определяется гладкостью целевой функции через ее норму $\|f\|_p$, что представляет собой существенное теоретическое улучшение и расширяет область практической применимости метода.

Далее приведены результаты апробации предложенного подхода на модели векторной авторегрессии (VAR). Для данной модели подтверждено выполнение необходимых условий: стационарность и свойство β -перемешивания. Проведенные вычислительные эксперименты демонстрируют конкурентоспособность метода даже в сравнении с нейросетевыми подходами.

В заключительной части главы теоретические выводы применяются к анализу модели векторной авторегрессии (VAR). Показано, что для этой модели выполняются ключевые предположения теории, а именно стационарность и свойство β -перемешивания. Проведенные численные эксперименты подтверждают эффективность предложенного подхода, точность которого сопоставима с нейросетевыми моделями.

В Главе 2 приведено описание предложенного автором двухуровневого алгоритма VAW^2 , и дано его обоснование. На первом уровне формируется ансамбль «экспертных» моделей, каждая из которых соответствует своему ядру и обучается с помощью алгоритма VAW . На втором уровне прогнозы агрегируются с помощью того же метода. Показано, что предложенная архитектура обладает значительно меньшей вычислительной сложностью по сравнению с прямым применением VAW к конкатенированным признакам, при этом сохраняя сопоставимые теоретические гарантии.

В теореме 2.2 установлена оценка сожаления порядка $O(T^{1/2}\ln T)$, где T – длина горизонта. В теореме 2.3 показано, что эту оценку можно незначительно улучшить (хотя указанный порядок при этом сохраняется) при наличии априорной информации о границах целевой переменной, используя усеченные прогнозы и мета-алгоритм EWA. Эффективность алгоритма VAW^2

подтверждена также экспериментально: сравнительный анализ на реальных и искусственных наборах данных демонстрирует, что предложенный метод превосходит по точности такие современные подходы к многоядерному обучению, как Raker и OMKL-GF.

Для решения практической проблемы предобработки данных предложен трехуровневый алгоритм S-VAW², который автоматически агрегирует прогнозы, полученные с использованием различных методов масштабирования признаков и целевой переменной. Это позволяет алгоритму адаптироваться к разнородности данных и снижает чувствительность к выбору конкретной предобработки. Сравнение с ведущим AutoML-фреймворком AutoGluon-Tabular на широком наборе табличных данных (Таблица 8) показало, что S-VAW² обеспечивает сопоставимую или лучшую точность при значительно меньших вычислительных затратах.

В главе 3 автор демонстрирует универсальность разработанного подхода, применяя методы онлайн оптимизации к задаче вариационного исчисления с неизвестным внешним воздействием. Ключевой шаг состоит в сведении исходной бесконечномерной задачи к конечномерной посредством тригонометрической аппроксимации. Далее проводится анализ сожаления для различных сценариев, включая статический, динамический и случай медленно меняющегося воздействия.

Теоретическую основу указанного сведения составляют леммы 3.1 и 3.2. В первой из них устанавливается связь между сожалением в исходной бесконечномерной задаче и сожалением в ее конечномерной аппроксимации, а во второй доказывается равномерная ограниченность ключевых параметров конечномерной задачи, таких как диаметр допустимого множества, константы Липшица и гладкости, а также параметр сильной выпуклости. Опираясь на эти теоретико-методологические основания, в теоремах 3.1–3.5 получены оценки сожаления для различных алгоритмов онлайн обучения. Теоретические выводы находят убедительное подтверждение в результатах вычислительных экспериментов, представленных в таблице 9, где продемонстрированы преимущества адаптивных стратегий в условиях нестационарной среды.

Полученные здесь результаты представляют особый интерес в связи с проблематикой распределенных вычислений. Недавно было показано (Farhadkhani et al., 2022), что при достаточно общих предположениях загрязнение данных и вмешательство в протокол вычислений (возможные действия противника) эквивалентны. Таким образом, полученные в диссертации результаты могут быть использованы для работы с неочищенными данными, содержащими, в том числе, злонамеренные искажения.

В заключении приведены основные результаты, полученные в диссертационной работе.

В приложениях приведены реализации предложенных численных методов и сертификат о регистрации программного комплекса.

Следует подчеркнуть, что все основные результаты, изложенные в диссертации, обладают научной новизной и подтверждены корректными математическими доказательствами. Методологический аппарат исследования интегрирует современные достижения функционального анализа, выпуклой оптимизации и теории онлайн обучения. Работа характеризуется высокой степенью технической завершенности, внутренней логической согласованностью и сбалансированным сочетанием теоретического и прикладного аспектов. Практическая значимость подтверждается реализацией всех предложенных алгоритмов в виде программного комплекса на языке Python, что гарантирует их воспроизводимость и применимость для практического использования.

По работе имеются следующие замечания и предложения.

1. Основные результаты работы получены с использованием алгоритма агрегирования прогнозов типа VAW, которые имеют сильные теоретические гарантии. В качестве дополнительной точки сравнения (помимо нейросетевых подходов) было бы ценно рассмотреть сравнительный анализ предложенных методов с современными адаптивными градиентными методами.

2. Алгоритм VAW² использует ансамбль из N ядер. В диссертации не обсуждается подробно стратегия выбора оптимального числа N ядер и соответствующего им набора гиперпараметров при отсутствии априорной информации о структуре данных. Представляется важным предложить эвристический или теоретически обоснованный подход к динамическому выбору параметра N в условиях онлайн обучения, что сделало бы метод более применимым в AutoML-контексте.

Высказанные замечания не умаляют научной и практической ценности диссертации и не влияют на ее общую положительную оценку.

Диссертационная работа О.В. Гуртовой представляет собой завершенное научное исследование, в котором решена актуальная научная задача разработки и теоретического обоснования вычислительно эффективных методов онлайн оптимизации для задач нелинейной регрессии. Основные научные положения и выводы работы прошли широкую апробацию, что подтверждается наличием 7 публикаций в рецензируемых научных изданиях, включая статьи в журналах, индексируемых в международных базах данных Scopus и Web of Science, а также докладами на 5 российских и международных конференциях. Практическая значимость разработанных алгоритмов дополнительно

подтверждается свидетельством о государственной регистрации программы для ЭВМ.

Содержание диссертации в полной мере соответствует паспорту научной специальности 1.2.2. – «Математическое моделирование, численные методы и комплексы программ».

Считаю, что диссертация О.В. Гуртовой на тему «Методы онлайн оптимизации квадратичной функции потерь, основанные на использовании случайных признаков Фурье» отвечает требованиям, предъявляемым к кандидатским диссертациям, а ее автор заслуживает присуждения искомой ученой степени кандидата физико-математических наук.

Официальный оппонент:

Профессор Кафедры математики

и анализа данных Факультета информационных технологий и анализа больших данных федерального государственного образовательного бюджетного учреждения высшего образования

«Финансовый университет

при Правительстве Российской Федерации»,

кандидат физико-математических наук,

профессор

Гисин Владимир Борисович

125167, Москва, пр-т Ленинградский, д. 49/2,
телефон: +7 (916) 253-27-26, E-mail: vgisin@fa.ru

Специальность, по которой оппонентом
защищена диссертация:

01.03.03 «Математическая логика,
алгебра и теория чисел», 1979 г.

