На правах рукописи

КУЛИКОВСКИХ Илона Марковна

# НЕЯВНАЯ РЕГУЛЯРИЗАЦИЯ ФУНКЦИЙ ЭКСПОНЕНЦИАЛЬНОГО ТИПА В ЗАДАЧАХ КЛАССИФИКАЦИИ

Специальность 05.13.17 – теоретические основы информатики

Автореферат
диссертации на соискание ученой степени
доктора технических наук

ПЕНЗА – 2020

Работа выполнена в ФГАОУ ВО «Самарский национальный исследовательский университет имени академика С.П. Королева»на кафедре «Информационные системы и технологии».

| | |
|---|---|
| Научный консультант – | **Прохоров Сергей Антонович,**<br>Заслуженный работник высшей школы РФ,<br>доктор технических наук, профессор,<br>заведующий кафедрой «Информационные системы и технологии» |
| Официальные оппоненты: | **Амосов Олег Семенович,**<br>доктор технических наук, профессор,<br>главный научный сотрудник лаборатории № 40 «Интеллектуальные системы управления и моделирования»ФГБУН «Институт проблем управления имени В.А. Трапезникова» Российской академии наук |
| | **Горбаченко Владимир Иванович,**<br>доктор технических наук, профессор,<br>заведующий кафедрой «Компьютерные технологии»ФГБОУ ВО «Пензенский государственный университет» |
| | **Кулагин Владимир Петрович,**<br>доктор технических наук, профессор,<br>заведующий кафедрой «Аппаратное, программное и математическое обеспечение вычислительных систем» ФГБОУ ВО «МИРЭА – Российский технологический университет» |
| Ведущая организация – | ФГБУН «Санкт-Петербургский институт информатики и автоматизации» Российской академии наук, г. Санкт-Петербург |

Защита состоится « 06 » июля 2020 г., в 14 часов на заседании диссертационного совета Д 212.337.01 на базе ФГБОУ ВО «Пензенский государственный технологический университет»по адресу: 440039, г. Пенза, пр. Байдукова/ул. Гагарина, д. 1а/11, корпус 1, конференц-зал.

С диссертацией можно ознакомиться в библиотеке ФГБОУ ВО «Пензенский государственный технологический университет»и на сайте www.penzgtu.ru.

Автореферат разослан «____» ________________ 2020 г.

Ученый секретарь
диссертационного совета Д 212.337.01,
доктор технических наук, доцент В.А. Чулков

# Общая характеристика работы

**Актуальность темы диссертации**

Теория обучаемых систем предлагает эффективные алгоритмы анализа информации, которые задают направление развития во многих отраслях. Широкое применение данного класса алгоритмов объясняется их способностью выявлять сложные структуры в исходных наборах данных через построение взаимосвязи между последовательностью обучающих моделей и множеством уровней представления информации. Повышение качества алгоритмов обучения является фундаментальной задачей данной теории.

Основу теории обучаемых систем заложили В.Н. Вапник и А.Я. Червоненкис в конце 60-х - начале 70-х годов в предложенной статистической теории восстановления зависимостей по эмпирическим данным. Согласно их исследованиям, построение модели восстанавливаемой зависимости в виде оптимального параметрического семейства алгоритмов, допускающих наименьшее число ошибок на обучающей выборке, может также учитывать погрешность измерений при формировании данной выборки. Это приводит к понижению качества алгоритма вне обучающей выборки и низкой обобщающей способности. Таким образом, был сделан вывод об оптимальной сложности модели с наилучшей обобщающей способностью, которая определяется величиной ошибки на неизвестной контрольной выборке.

Предложенные В.Н. Вапником и А.Я. Червоненкисом в рамках теории оценки оказались завышенными, что требовало чрезмерно длинных выборок и сложных моделей. Такую особенность связывают с тем, что данные оценки не учитывают процесс поиска решения или метод обучения, который потенциально снижает сложность модели, формируя эффективную сложность подмножества алгоритмов. Как следствие, возникло большое множество исследований, направленных на уточнение оценок. В работах академика РАН В.Л. Матросова показано, что при выборе метода обучения обеспечивается корректное распознавание заданной обучающей выборки при использовании подмножества алгоритмов ограниченной емкости, лежащего в алгебраическом расширении семейства алгоритмов вычисления оценок. Еще одно направление связано с построением композиций алгоритмов, плохо восстанавливающих искомую зависимость по отдельности, с помощью корректирующей операции. Теория алгоритмических композиций была разработана в алгебраическом подходе к построению корректных алгоритмов академиком РАН Ю.И. Журавлевым и нашла свое развитие в работах его учеников. P. Bartlett показал, что эффективная сложность композиции алгоритмов равна средней взвешенной сложности отдельных алгоритмов, т.е. сглаживая прогнозы, построенная композиция не увеличивает результирующую сложность. Полученные таким образом оценки были завышены, хотя оказались точнее оценок Вапника-Червоненкиса, что натолкнуло Y. Freund и R. Schapire на создание улучшенного композиционного метода - бустинга, в котором базовые алгоритмы строятся последовательно, минимизируя ошибки, допущенные на предыдущих алгоритмах. Так как такой подход позволяет улучшать качество обучения посредством увеличения количества алгоритмов в композиции, исследования P. Bartlett, Y. Freund, R. Schapire существенным образом повлияли на представление о соотношении обобщающей способности и сложности алгоритмов, акцентируя внимание на первостепенной важности метода обучения, на основе которого по обучающей выборке строится алгоритм. В комбинаторном подходе к оценке качества обучаемых алгоритмов, предложенном К.В. Воронцовым для уточнения статистической теории Вапника-Червоненкиса, явным образом вводится понятие метода обучения. Данное направление обращает внимание на необходимость привлечения априорной информации для построения более точных оценок. Исследованию соотношения обучающей выборки (локальной информации) и априорных ограничений (универсальной информации) посвящена теория универсальных и локальных ограничений академика РАН К.В. Рудакова.

В своих недавних исследованиях по оценке качества моделей обучения с избыточным количеством параметров, B. Neyshabur и др. вводят понятие неявного смещения, задаваемого градиентным методом обучения [1]. Неявная регуляризация, вводимая таким смещением, рассматривается как разновидность метода обучения, где обобщающая способность, в отсутствии явной регуляризации и ограничений на сложность модели, продолжает повышаться. В результате исследования различных функций экспоненциального типа было выявлено, что метод градиентного спуска как в детерминированной, так и в стохастической постановке задачи в условиях идеальной разделимости классов сходится в направлении оптимальной гиперплоскости, максимизирующей зазор между классами с нормой $L_2$ [2]. Однако скорость сходимости к оптимальному решению с фиксированным шагом снижается до логарифмической, тогда как скорость сходимости самого метода является линейной.

Наиболее распространенным способом повышения скорости сходимости при неявной регуляризации является использование адаптивного градиентного шага. Как правило, такой подход предполагает модификацию градиентного метода, не гарантирующую оценку границы погрешности, эквивалентную оценке стандартного градиентного метода. Кроме того, применение адаптивных градиентных методов значительно повышает смещенность результирующих оценок, что является одной из наиболее существенных проблем, препятствующих широкому внедрению обучаемых систем в отраслях, связанных с жизнью и деятельностью человека, и требующих постоянного контроля со стороны экспертов.

[1] Neyshabur, B., Tomioka, R., and Srebro, N. In search of the real inductive bias: On the role of implicit regularization in deep learning. In ICLR, 2014.

[2] Soudry, D., Hoffer, E., Nacson, M., Gunasekar, S., and Srebro, N. The implicit bias of gradient descent on separable data. JMLR, 19:1–57, 2018.

В работах T. Poggio[3] и его учеников рассмотрены различные виды устойчивости в контексте моделей с избыточным количеством параметров. В частности, отмечается, что устойчивость глобального минимума по отношению к локальным в пространстве параметров модели, т.е. сокращение количества локальных пиков на поверхности функции потерь, повышает устойчивость модели к выбросам в пространстве исходной выборки наблюдений, наличие которых оказывает влияние на смещенность оценок.

В данной работе предлагаются функции экспоненциального типа с переменными параметрами, которые учитывают их скорость роста и величину асимптот. Неявная регуляризация таких функций гарантирует границу погрешности, эквивалентную границе метода градиентного спуска, более высокую скорость сходимости и менее смещенные оценки, так как регуляризация выполняется в пространстве параметров функции, и не зависит от параметров метода оптимизации, которые вносят в результирующие оценки численные погрешности. Преимуществом предлагаемого подхода также является тот факт, что предлагаемые функции имеют физический смысл - они построены на основе обобщения логистического уравнения [4], впервые предложенного для описания динамики роста популяций и нашедшего широкое применение в различных технических и социальных средах. В рамках данного исследования рассматривается применимость предложенных моделей обучения в более широком контексте обучения машин и людей [5], а именно, при задании частных параметров функций экспоненциального типа в теории фильтрации при построении рекурсивных фильтров Лагерра и Мейкснера и в теории тестирования при построении модели Бартона (4PL IRT, 4-Parameter Logistic Item Response Theory)[6,7].

Актуальность работы подтверждается соответствием Перечню критических технологий РФ от 7 июля 2011 г. № 899: «Нано-, био-, информационные, когнитивные технологии»(п. 8) и направлению из Стратегии научно-технологического развития Российской Федерации от 01.12.2016 № 642: «Переход к передовым цифровым, интеллектуальным производственным технологиям, роботизированным системам, к новым материалам и способам конструирования, создание систем обработки больших объемов данных, машин-

[3]Poggio, T., Kawaguchi, K., Liao, Q., Miranda, B., Rosasco, L. et al. Theory of Deep Learning III: explaining the non-overfitting puzzle. Center for Brains, Minds and Machines (CBMM), MIT, Memo No. 073, 2018.

[4]Verhulst, P. Notice sur la loi que la population poursuit dans son accroissement. Correspondance mathamatique et physique, 10:113–121, 1838.

[5]Wilson, R.C., Shenhav, A., Straccia, M. et al. The Eighty Five Percent Rule for optimal learning. Nature Communications, 10: 4646, 2019.

[6]Birnbaum, A. Some latent trait models and their use in inferring an examinee's ability. In F.M. Lord and M.R. Novick (Eds.), Statistical theories of mental test scores (pp. 397-479). Reading, MA: Addison-Wesley, 1968.

[7]Barton, M.A. and Lord, F.M. An upper asymptote for the three-parameter logistic item-response model. Princeton, NJ: Educational Testing Service, 1981.

ного обучения и искусственного интеллекта» (п. 1).

Работа выполнена при государственной поддержке гранта Президента РФ (№ МК-6218.2018.9), гранта РФФИ (№ 18-37-00219), гранта Минобрнауки РФ (№ 074-U01), проекта программы СТАРТ-1-18 (№ С1-51885), проекта DATACROSS Центра Превосходства, финансируемого Правительством Хорватии и Европейским Союзом через Европейский фонд регионального развития - Операционная программа конкурентоспособности и сплочения (KK.01.1.1.01.0009), а также при проведении научно-исследовательских работ в рамках стратегической академической единицы «Нанофотоника, перспективные технологии зондирования Земли и интеллектуальные геоинформационные системы»(№ 05в-Р008-035, № 05в-Р001-403).

В диссертационной работе разработаны теоретические положения, совокупность которых можно квалифицировать как новые научно обоснованные технические решения, внедрение которых вносит значительный вклад в развитие страны.

**Объектом исследования** в диссертационной работе являются функции экспоненциального типа.

**Предметом исследования** в диссертационной работе является теория обучаемых систем.

**Целью диссертационной работы** является повышение эффективности неявной регуляризации функций экспоненциального типа в задачах классификации.

**Задачи диссертационной работы:**

1. Провести анализ основных подходов к повышению эффективности градиентных методов для функций экспоненциального типа.
2. Определить функции экспоненциального типа с переменными параметрами.
3. Повысить скорость сходимости и устойчивость градиентных методов для функций экспоненциального типа с переменными параметрами.
4. Понизить сложность моделей в теории фильтрации и теории тестирования за счет применения неявной регуляризации функций экспоненциального типа с переменными параметрами.
5. Провести экспериментальные исследования для анализа функций экспоненциального типа с переменными параметрами в задачах классификации.
6. Предложить программные реализации моделей, построенных на основе функций экспоненциального типа с переменными параметрами.
7. Провести апробацию предложенных моделей и их программных реализаций при решении прикладных задач.

**Методы**, используемые в диссертации, включают градиентные методы построения регрессионных моделей, методы регуляризации, методы теории рекурсивной фильтрации, методы теории современного тестирования, а также методы, используемые для описания динамики роста популяций.

**Научная новизна работы:**

1. Введено определение функции экспоненциального типа с переменными параметрами, которое, в отличие от известных, учитывает ее скорость роста и величину асимптот.
2. Повышена скорость сходимости и устойчивость градиентных методов за счет неявной регуляризации функций экспоненциального типа с переменными параметрами.
3. Разработан метод неявной регуляризации моделей рекурсивных фильтров Лагерра и Мейкснера, сводящий соответствующий коэффициент фильтра к нулю.
4. Разработан метод неявной регуляризации модели Бартона при индивидуальном и групповом тестировании, замещающий явные вероятностные параметры модели на их нечеткие определения.

**Практическая значимость работы:**

1. Снижены вычислительные затраты на построение моделей на основе функций экспоненциального типа с переменными параметрами за счет гарантированной более высокой скорости сходимости градиентных методов.
2. Снижена смещенность оценок моделей на основе функций экспоненциального типа с переменными параметрами за счет более высокой устойчивости градиентных методов.
3. Снижена сложность рекурсивных фильтров Лагерра и Мейкснера, построенных на основе функций экспоненциального типа с переменными параметрами, за счет сведения соответствующего коэффициента фильтра к нулю.
4. Снижена сложность модели Бартона при индивидуальном и групповом тестировании, построенной на основе функций экспоненциального типа с переменными параметрами, за счет замещения явных вероятностных параметров модели на их нечеткие определения.
5. Получены результаты экспериментальных исследований по анализу неявной регуляризации на основе функций экспоненциального типа с переменными параметрами в задачах классификации.
6. Предложены программные реализации моделей, построенных на основе функций экспоненциального типа с переменными параметрами, включающие их функциональное и алгоритмическое описание.
7. Проведена апробация предложенных моделей и их программных реализаций при решении задач обработки реальных данных.

Методы неявной регуляризации функций экспоненциального типа с переменными параметрами и их программные реализации внедрены в АО «Самарский электромеханический завод»(г. Самара) при создании платформы для сбора информации и формирования контингента и АО «РКЦ «Прогресс»(г. Самара) при обработке результатов испытаний. Результаты применения неявной регуляризации функций экспоненциального типа с переменными параметрами для понижения сложности индивидуального и группового тестирования внедрены в ООО «ТимКемистри»(г. Самара) при анализе реальных проектов для выявления характеристик, указывающих на особенности членов команды и оптимально сконструированное взаимодействие между ними. Разработанные модели, алгоритмы и их программные реализации для решения прикладных задач используются в учебном процессе ФГАОУ ВО «Самарский национальный исследовательский университет имени академика С.П. Королева»(г. Самара) при подготовке специалистов по направлению «Информатика и вычислительная техника»(коды 09.06.01 и 09.04.01). Предложенные научно обоснованные технические решения приведены в актах внедрения. Акты внедрения прилагаются.

**Результаты, выносимые на защиту:**

1. Введено определение функции экспоненциального типа с переменными параметрами, которое, в отличие от известных, учитывают ее скорость роста и величину асимптот.
2. Повышена эффективность неявной регуляризации функций экспоненциального типа в задачах классификации, которая в отличие от ранее предложенных, обеспечивает более высокую скорость сходимости и устойчивость градиентных методов.
3. Снижены вычислительные затраты на неявную регуляризацию функций экспоненциального типа с переменными параметрами на порядок.
4. Снижена смещенность оценок моделей при неявной регуляризации функций экспоненциального типа с переменными параметрами в 5 раз.
5. Разработан метод неявной регуляризации моделей рекурсивных фильтров Лагерра и Мейкснера на основе функций экспоненциального типа с переменными параметрами.
6. Разработан метод неявной регуляризации модели Бартона при индивидуальном и групповом тестировании на основе функций экспоненциального типа с переменными параметрами.
7. Предложены программные реализации моделей, построенных на основе функций экспоненциального типа с переменными параметрами, включающие их функциональное и алгоритмическое описание.

**Область исследования** согласно паспорту специальности 05.13.17 – Теоретические основы информатики:

- исследование и разработка средств представления знаний. Разработка интегрированных средств представления знаний, средств представления знаний, отражающих динамику процессов, концептуальных и семиотических моделей предметных областей (п. 4);
- разработка и исследование моделей и алгоритмов анализа данных, обнаружения закономерностей в данных и их извлечениях (п. 6);

- разработка методов распознавания образов, фильтрации, распознавания и синтеза решающих правил. Моделирование формирования эмпирического знания (п. 7);
- исследование и когнитивное моделирование интеллекта, включая моделирование поведения, моделирование рассуждений различных типов, моделирование образного мышления (п. 8);
- применение бионических принципов, методов и моделей в информационных технологиях (п. 13).

Согласно формуле специальности «Теоретические основы информатики», к ней относятся, в числе прочего, «. . . исследования методов преобразования информации в данные и знания; создание и исследование. . . моделей данных и знаний, методов работы со знаниями, методов машинного обучения и обнаружения новых знаний. . . ; исследования принципов создания и функционирования аппаратных и программных средств автоматизации указанных процессов».

**Личный вклад.** В выполненных единолично работах опубликованы результаты, выносимые на защиту в данной диссертации. В выполненных с соавторами работах опубликованы частные решения, отражающие примеры применения результатов диссертации.

**Степень достоверности и апробация результатов.**

Достоверность результатов обеспечивается широкой экспериментальной базой и строгим теоретическим обоснованием справедливости предложенных результатов. Проведен сравнительный анализ разработанных методов с наиболее эффективными известными подходами на основе объективных критериев качества.

Результаты, полученные в диссертации, представлялись на Международном конгрессе студентов, аспирантов и молодых ученых «Перспектива-2007», Нальчик, Кабардино-Балкария (2007); Всероссийской научной конференции «Инновационные технологии в управлении, образовании, промышленности» «АСТИНТЕХ-2007», Астрахань (2007); Всероссийской научно-технической конференции студентов, молодых ученых и специалистов «Биомедсистемы - 2007»(Биотехнические, медицинские и экологические системы и комплексы), Рязань (2007); Международных научно-технических конференциях «Радиотехника и связь», Саратов (2007, 2009); Международной открытой конференции «Современные проблемы информатизации в анализе и синтезе технологических и программно-телекоммуникационных систем (СПИ-2008)», Воронеж (2008); Международной молодежной научной конференции «XXXIV ГАГАРИНСКИЕ ЧТЕНИЯ», Москва (2008); Всероссийских научных конференциях с международным участием «Математическое моделирование и краевые задачи», Самара (2008, 2011); Международной научно-технической конференции «Методы, средства и технологии получения и обработки измерительной информации (ИЗМЕРЕНИЯ-2008)», Пенза (2008); Международной конференции «Идентификация, измерение характеристик и имитация случайных сигналов (ИИИ-2009)», Новосибирск (2009); Материалы научно - практической конференции «Образование – инвестиции в успех», Алматы, Казахстан (2009); Российской школе-семинаре аспирантов, студентов и молодых ученых «Информатика, моделирование, автоматизация проектирования», Ульяновск (2009); Международной научно-практической конференции «Аналитические и численные методы моделирования естественнонаучных и социальных проблем», Пенза (2009, 2017); Международной научно-технической конференции «Проблемы автоматизации и управления в технических системах», Пенза (2011); Всероссийских научно-практических конференциях «Компьютерные технологии в науке, практике и образовании», Самара (2005, 2006, 2007, 2008, 2009, 2012); Международных научно-технических конференциях «Перспективные информационные технологии в научных исследованиях, проектировании и обучении», Самара (2006, 2010, 2012, 2013, 2014, 2016, 2017, 2018, 2019); 8th International Conference on Applied Mathematics and Scientific Computing, Šibenik, Croatia (2013); 7th Computer Science and Electronic Engineering Conference, Colchester, UK (2015); Международной конференции и молодежной школе «Информационные технологии и нанотехнологии», Самара (2017, 2018, 2019), 4th International Workshop on Data Science, Zagreb, Croatia (2019), 22nd International Conference on Discovery Science, Split, Croatia (2019).

**Публикации.** Результаты диссертации опубликованы в 106 работах, из них: 29 публикаций в журналах, рекомендованных ВАК, среди них 12 публикаций входят в базы WoS/Scopus, а также 2 монографии, 3 учебных пособия и 14 свидетельств о государственной регистрации программ для ЭВМ.

**Объем и структура диссертации.** Диссертация состоит из введения, пяти глав, заключения и четырех приложений. Общий объем диссертации 242 страницы. Диссертация содержит 28 таблиц, 113 рисунков и список литературы из 459 наименования.

Автор выражает благодарность за предоставление условий для проведения исследований и обсуждение результатов диссертационной работы в рамках проекта DATACROSS Центра Превосходства Prof., Dr. Sc. Sven Lončarić, Prof., Dr. Sc. Tomislav Šmuc, Prof., Dr. Sc. Tarzan Legović и Dr. Sc. Tomislav Lipić во время нахождения на постдокторском курсе в Laboratory for Machine Learning and Knowledge Representation, Ruđer Bošković Institute и Faculty of Electrical Engineering and Computing, University of Zagreb (2018-2019).

## Содержание диссертации

**Во введении** показана актуальность темы диссертации, определены цель и задачи работы, методы исследования, изложена научная новизна и практическая значимость полученных результатов, сформулированы основные положения, выносимые на защиту.

**В первой главе** рассмотрены основные подходы к повышению сходимости функций экспоненциального типа.

Пусть имеется выборка наблюдений $\{\mathrm{x}_i, y_i\}_{i=1}^m$, где $\mathrm{x}_i \in \mathrm{R}^n$ и $y_i \in \{-1, 1\}$. Поставим задачу минимизации эмпирической функции потерь

$$\mathscr{L}(\theta) = \sum_{i=1}^{m} \ell(y_i \theta^{\mathrm{T}} \mathrm{x}_i) \tag{1}$$

с вектором параметров $\theta \in \mathrm{R}^n$. По аналогии с постановкой задачи, предложенной D. Soundry и др. [2], рассмотрим случай, когда выборка наблюдений идеально разделима (Предположение 1), $\ell$ удовлетворяет условиям Предположения 2, а $-\ell'$ имеет экспоненциальный хвост (Предположение 3).

**Предположение 1.** *Выборка наблюдений линейно разделима:* $\exists\ \theta^*$ *такое, что* $\forall i : y_i \theta^{*\mathrm{T}} \mathrm{x}_i > 0$.
Для простоты преобразований, предположим также, что $\forall i \in \{1, ..., m\} : y_i = 1, \|x_i\| < 1$.

**Предположение 2.** $\forall u \in \mathrm{R}$*:* $\ell(u)$ *является гладкой строго убывающей неотрицательной функцией, ограниченной снизу:* $\ell(u) > 0$*,* $\ell'(u) < 0$*,* $\lim_{u \to -\infty} \ell'(u) \neq 0$*,* $\lim_{u \to \infty} \ell(u) = \lim_{u \to \infty} \ell'(u) = 0$*, имеющей непрерывный по Липшицу градиент с константой* $\beta$*:* $\ell(u') \leq \ell(u) + \langle \nabla \ell(u), u' - u \rangle + \frac{\beta}{2} \|u' - u\|^2, \quad \beta > 0$.

**Предположение 3.** *Пусть* $-\ell'(u) = \exp(-f(u))$*, где* $f(u) = \omega(\ln(u))$ [8]*.*

Под функцией **экспоненциального типа** будем понимать функцию **с экспоненциальным хвостом**, заданную согласно Предположению 3. Понятие *экспоненциального хвоста* было предложено в работе [8] для описания асимптотического поведения функции $\ell$, при котором $f(u)$ доминирует над $g(u)$, где $g(u) = \ln(u)$: $\forall C > 0, \exists U : \forall u \in U \;\; C|g(u)| < |f(u)|$. Если аргумент $u$ является случайной величиной, функционал $\ell$ может задавать функцию распределения. Если функционал $\ell$ является логистической функцией, то $\ell(u)$ определяет логистическую функцию распределения, используемую для определения логистической регрессии, сигмоидальной функции активации при построении модели нейронных сетей и т.д. В данной работе сделан акцент на асимптотическом анализе функционала $\ell$, имеющего экспоненциальный хвост согласно Предположению 3. Распределения вероятностей с различными хвостами, построенными на основе $\ell$, в рамках данной работы не рассматриваются.

Воспользуемся методом градиентного спуска для минимизации (1) с фиксированным шагом градиента $\eta$:

$$\theta(t+1) = \theta(t) - \eta \nabla \mathscr{L}(\theta(t)) = \theta(t) - \eta \sum_{i=1}^{m} \ell'(\theta(t)^{\mathrm{T}} \mathrm{x}_i) \mathrm{x}_i. \tag{2}$$

[8] Nacson, M., Lee, J.D., Gunasekar, S., Savarese, P.H.P., Srebro, N., and Soundry, D. Convergence of gradient descent on separable data. In AISTATS, 2019

Обобщая постановку задачи на случай многослойной нейронной сети, соотношение (2) примет вид:

$$\mathbf{\Theta}_l(t+1) = \mathbf{\Theta}_l(t) - \eta \sum_{i=1}^{m} \nabla_{\theta_l} \mathcal{L}(\Theta), \tag{3}$$

$$\mathcal{L}(\Theta) = \sum_{i=1}^{m} \ell(\langle \Pi(\Theta), \mathrm{x}_i \rangle), \tag{4}$$

где $\Pi(\Theta) = \mathbf{\Theta}_1 \times \mathbf{\Theta}_2 \times \cdots \times \mathbf{\Theta}_L$, $\Theta = \{\mathbf{\Theta}_l \in \mathrm{R}^{d_{l-1} \times d_l} : l = 1, 2, \ldots, L\}$, $L$ -количество слоев, $d_l$ - количество нейронов в слое $l$.

В случае стохастической постановки задачи, метод градиентного спуска обновляет (3) с учетом (4) для каждого пакета $B(t) \subseteq \{1, \ldots, m\}$ как:

$$\mathcal{L}(\Theta) = \sum_{i \in B(t)} \ell(\langle \Pi(\Theta), \mathrm{x}_i \rangle). \tag{5}$$

С учетом Предположений 1 и 2, метод градиентного спуска находит глобальный минимум, даже если функция потерь $\mathscr{L}(\theta)$ не является выпуклой [2,8]. Так как $\forall i : \theta^{*\mathrm{T}} \mathrm{x}_i > 0$ и $-\ell'(t) > 0$ для любого конечного $t$, согласно неравенству Коши-Шварца, $\|\theta(t)\| \geq \frac{\|\theta^{*\mathrm{T}} \theta(t)\|}{\|\theta^*\|}$, $\lim_{t \to \infty} \|\theta(t)\| = \infty$, $\forall i : \lim_{t \to \infty} \theta(t)^{\mathrm{T}} \mathrm{x}_i = \infty$.
С учетом сделанных допущений нормированный вектор параметров сходится в направлении гиперплоскости, максимизирующей зазор с нормой $L_2$ [2]:

$$\left\| \frac{\theta(t)}{\|\theta(t)\|} - \frac{\hat{\theta}}{\|\hat{\theta}\|} \right\| = \mathcal{O}\left(\frac{1}{g(t)}\right), \tag{6}$$

где $\min_i \frac{\theta_t^{\mathrm{T}} \mathrm{x}_i}{\|\theta_t\|} = d - \mathcal{O}\left(\frac{1}{g(t)}\right)$, $d = \max_\theta \min_i \frac{\theta^{\mathrm{T}} \mathrm{x}_i}{\|\theta\|} = \frac{1}{\hat{\theta}}$. Скорость сходимости (6) справедлива для большинства функций $\ell$, включая логистическую и экспоненциальную. Если $f'(t) = o(1)$ и $f$ является строго вогнутой, то $g(t) = \ln(t)$[8].

Скорость сходимости $\mathcal{O}\left(\frac{1}{\ln(t)}\right)$ значительно ниже скорости сходимости самой функции $\mathcal{O}\left(\frac{1}{t}\right)$. Наиболее широко используемые подходы к повышению скорости сходимости не гарантируют границу погрешности, эквивалентную границе метода градиентного спуска, и дают более смещенные оценки.

В данной работе предлагаются функции **экспоненциального типа с переменными параметрами**, иначе, функции **с переменными экспоненциальными хвостами**, которые позволяют выполнять неявную регуляризацию непосредственно в пространстве параметров хвоста, что, в отличие от пространства параметров метода оптимизации, не вносит в результирующие оценки численные погрешности. Такой подход гарантирует границу погрешности, эквивалентную границе метода градиентного спуска, более высокую скорость сходимости и менее смещенные оценки.

**Во второй главе** предложено определение функций экспоненциального типа с переменными параметрами. Обоснована их эффективность для повышения скорости сходимости и устойчивости градиентных методов.

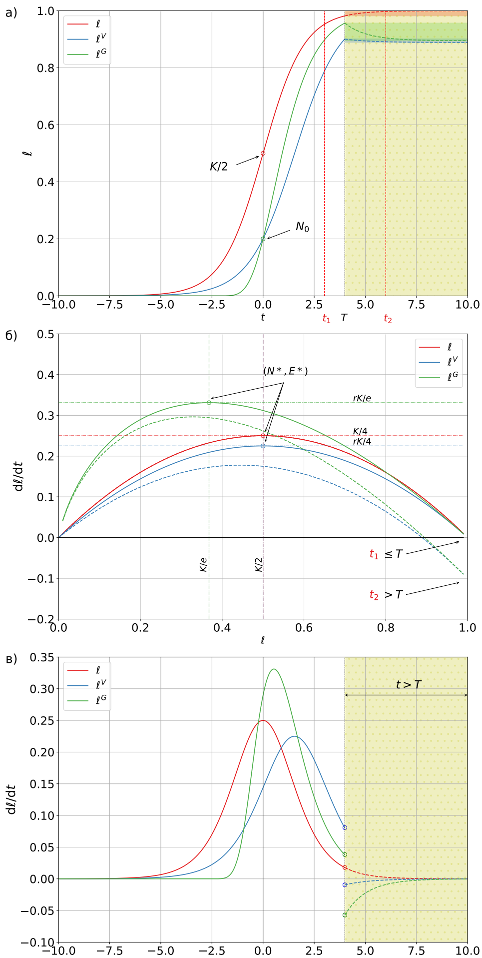



Рис. 1 – Функции $\ell$ с переменными экспоненциальными хвостами: а) $\ell(t)$; б) $\ell'(\ell)$; с) $\ell'(t)$

Зададим некоторую функцию, удовлетворяющую Предположениям 1, 2 и 3, в виде $\ell(t; T, K, r, N_0, E)$, которая имеет переменный экспоненциальный хвост и имитирует модель логистического роста популяции $N(t)$ с учетом пропорциональной «редукции»[9], преобразованную к диапазону $t \in \mathrm{R}$:

$\lim_{t\to-\infty} \ell(t; T, K, r, N_0, E) = 0$,

$\lim_{t\to\infty} \ell(t; T, K, r, N_0, E) = K$,

где «редукция»популяции $E$ начинается в момент времени $T$ со скоростью $H(t) = EN(t)$, $r$ задает скорость роста популяции, $K \leq 1$ - емкость среды, которая определяет максимально допустимый размер популяции, $N_0$ - начальный размер популяции.

Предложим функции с переменными экспоненциальными хвостами: построенные на уравнениях Ферхюльста $\ell^{\mathrm{V}}(t; T, K, r, N_0, E)$ и Гомперца $\ell^{\mathrm{G}}(t; T, K, r, N_0, E)$.

На рис. 1 приведены функция с постоянным экспоненциальным хвостом $T = 0$, $K = 1$, $r = 1$, $N_0 = 0.5$, $E = 0$ (красная кривая) и функции с переменными экспоненциальными хвостами $T = 4$, $K = 1$, $r = 0.9$, $N_0 = 0.2$, $E = 0.1$ на основе уравнения Ферхюльста (синяя кривая) и Гомперца (зеленая кривая).

Значение $(N^*, E^*)$ на рис. 1 б) задает точку равновесия, которая определяет максимально допусти-

[9] Legović, T. Dynamic population models. In Jorgensen, S. (ed.), Ecological model types, pp. 39–63. Elsevier, 2016.

мый уровень «редукции», не приводящий к вымиранию популяции. Моменты времени $t_1 = 3$ и $t_2 = 6$ определяют состояние системы до и после $T$ соответственно. Введенные параметры $T, K, r, N_0, E$ позволяют учитывать скорость роста функции $r$ и верхнюю асимптоту экспоненциального хвоста $K$, на изменение величины которой оказывают влияние параметры $T, N_0, E$. Обобщая определения $\ell^{\mathrm{V}}$ и $\ell^{\mathrm{G}}$, введем функцию $\ell(t; T, K, r, N_0, E, q)$ с переменным экспоненциальным хвостом и обобщающим параметром $q$.

**Лемма 1.** *Пусть* $\ln_q(x)$ *задает функцию обобщенного логарифма, которая определяет область под кривой, заданной* $u_q = \frac{1}{t^{1-q}}$, $t \in [1, x]$: $\ln_q(x) = \int_1^x \frac{\mathrm{d}t}{t^{1-q}} = \frac{x^q - 1}{q}, q \neq 0$. *Тогда функция с переменным экспоненциальным хвостом имеет вид:*

$$\ell(t; T, K, r, N_0, E, q) = \begin{cases} K\mathrm{e}^{\ln_q\left(\frac{N_0}{K}\right)\mathrm{e}^{-rt}}, & \textit{если } t \leq T, \\ \varepsilon K \mathrm{e}^{\left(\ln_q\left(\frac{N_T}{K}\right) + \frac{E}{r}\right)\mathrm{e}^{-r(t-T)}}, & \textit{если } t > T, \end{cases} \tag{7}$$

*где* $\varepsilon$ *определяет степень влияния редукции* $E$ *на скорость роста* $r$.

**Следствие 1.** *Из Леммы 1 следует, что:*

1. $\ell(t; T, K, r, N_0, E, q) \approx \ell^{\mathrm{V}}(t; T, K, r, N_0, E)$ *при* $q = 1$;

2. $\ell(t; T, K, r, N_0, E, q) = \ell^{\mathrm{G}}(t; T, K, r, N_0, E)$ *при* $q \to 0$, *принимая во внимание* $\lim_{q \to 0} \ln_q(x) = \ln(x)$;

3. $\ell(t; T, K, r, N_0, E, q) \approx \ell(t)$ *при* $q = 1$, $T = 0$, $K = 1$, $r = 1$, $N_0 = 0.5$, $E = 0$.

С учетом Леммы 1 перепишем (2) в виде:

$$\boldsymbol{\theta}(t+1) = \boldsymbol{\theta}(t) - \eta \sum_{i=1}^{m} \ell'(\boldsymbol{\theta}(t)^{\mathrm{T}} \mathrm{x}_i; T, K, r, N_0, E, q)\mathrm{x}_i, \tag{8}$$

$$\ell'(t; T, K, r, N_0, E, q) = \begin{cases} -Kr \ln_q\left(\frac{N_0}{K}\right) \mathrm{e}^{-rt + \ln_q\left(\frac{N_0}{K}\right)\mathrm{e}^{-rt}}, & \text{если } t \leq T, \\ -\varepsilon K r \left(\ln_q\left(\frac{N_0}{K}\right) + \frac{E}{r}\right) \times \\ \mathrm{e}^{-r(t-T) + \left(\frac{E}{r} + \ln_q\left(\frac{N_T}{K}\right)\right)\mathrm{e}^{-rt} - \frac{E}{r}}, & \text{если } t > T. \end{cases} \tag{9}$$

Проанализируем скорость сходимости нормированного вектора параметров (8) в направлении гиперплоскости, максимизирующей зазор между классами с нормой $L_2$.

**Теорема 1.** *Для почти любой выборки наблюдений, удовлетворяющей Предположению 1, любой функции* $\ell$, *удовлетворяющей Предположениям 2 и 3 итерации (8) в любой начальной точке* $\boldsymbol{\theta}_0$ *с единственным слоем* $L = 1$

*сходятся в направлении гиперплоскости, максимизирующей зазор между классами, с фиксированным шагом градиента* $\eta < \frac{2}{\beta}$ *и временем начала редукции* $T$ *как* $\min_i \frac{\theta_t^{\mathrm{T}} \mathrm{x}_i}{\|\theta_t\|} = d - \mathcal{O}\left(\frac{1}{g(t)}\right)$, $d = \max_{\theta} \min_i \frac{\theta^{\mathrm{T}} \mathrm{x}_i}{\|\theta\|} = \frac{1}{\hat{\theta}}$,

$$g(t) = \begin{cases} r\ln(t) + W_0\left(\frac{\ln_q\left(\frac{N_0}{K}\right)}{t}\right), & \text{если } t \leq T, \\ rT - \frac{E}{r} + \left(\ln(t) + W_0\left(\frac{\left(\ln_q\left(\frac{N_0}{K}\right) + \frac{E}{r}\right)\mathrm{e}^{-Tr + \frac{E}{r}}}{t}\right)\right), & \text{если } t > T, \end{cases} \tag{10}$$

*где* $W_0$ *задает основную ветвь функции Ламберта* $W$.

Из Теоремы 1 следует, что оценка $g(t)$ для моделей с переменными экспоненциальными хвостами с учетом асимптотического поведения основной ветви функции Ламберта дает более высокую скорость сходимости до $\mathcal{O}\left(\frac{1}{\sqrt{t}}\right)$ и $\mathcal{O}\left(\frac{1}{t}\right)$ в сравнении с ранее выявленной скоростью $\mathcal{O}\left(\frac{1}{\ln t}\right)$ для моделей с постоянными экспоненциальными хвостами при одних и тех же допущениях. Теорема 1 включает анализ метода градиентного спуска, если число слоев $L = 1$. Однако полученные результаты могут быть расширены на случай стохастической постановки задачи, если $L > 1$ (5).

На рис. 2-4 приведены результаты экспериментальных исследований скорости сходимости на различных наборах данных из *UCI Machine Learning Repository*: *Breast Cancer Wisconsin* (breast), *Pima Indians* (pima), *MNIST* (mnist), – при построении нейронной сети, включающей один скрытый слой с $d_l = 100$. Для обучения сети использовался метод стохастического градиента со стандартным значением шага $\eta = 0.01$, количеством пакетов $|B(t)| = 75$ и количестве итераций $n = 750$. Функции активации построены на основе определения (7). Параметры экспоненциального хвоста оптимизировались в диапазонах $T \in [0,3], K \in [0.85,1], r \in [0.5,5.1], N_0 \in [0.2,0.8], E \in [0,1]$ методом случайного поиска с выбором 7.5% возможных комбинаций значений поиска по полной сетке.
Исходная выборка наблюдений была разбита случайным выбором на обучающую и тестовую, которая составляла 20 % от исходной. Для оптимизации параметров хвоста использовалась контрольная выборка, которая составляла 20 % от обучающей.

Из приведенных рисунков видно, что использование функций с переменными экспоненциальными хвостами повышает скорость сходимости градиентного метода, в большей степени для хвоста, построенного на уравнении Ферхюльста в случае наборов breast, pima, и с небольшим перевесом в пользу уравнения Гомперца в случае набора mnist.

Переопределим функцию $\ell'(t; T, K, r, N_0, E, q)$ в отсутствии редукции $E = 0$, $T = 0$ при значении обобщающего параметра $q = 1$, но дополненную с учетом нижней асимптоты - критического размера популяции $0 \leq A < K$, ниже

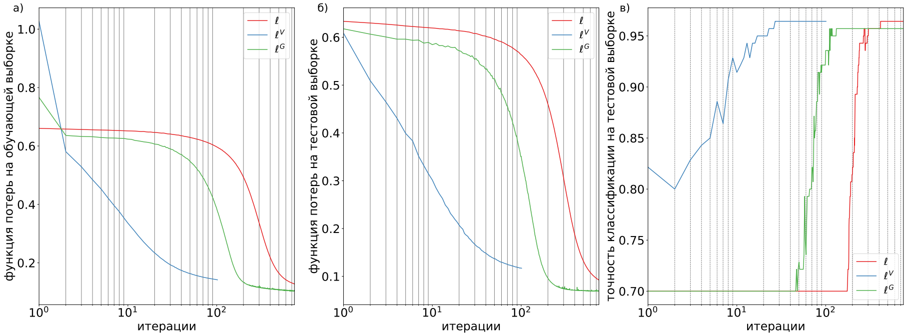



Рис. 2 – Результаты экспериментальных исследований на наборе breast: а) потери на обучающей выборке; б) потери на тестовой выборке; в) точность на тестовой выборке

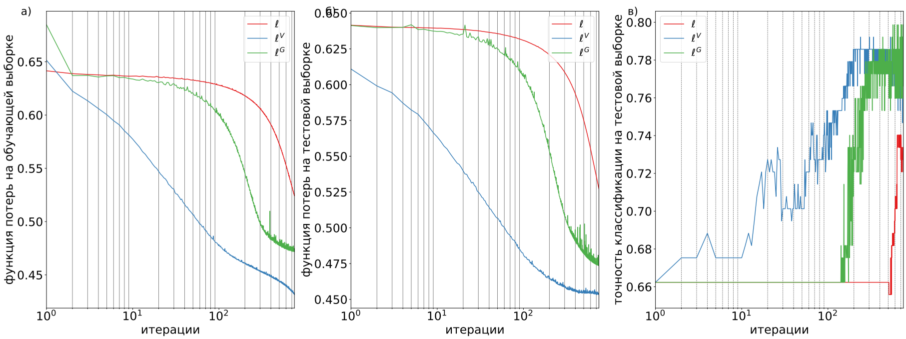



Рис. 3 – Результаты экспериментальных исследований на наборе pima: а) потери на обучающей выборке; б) потери на тестовой выборке; в) точность на тестовой выборке

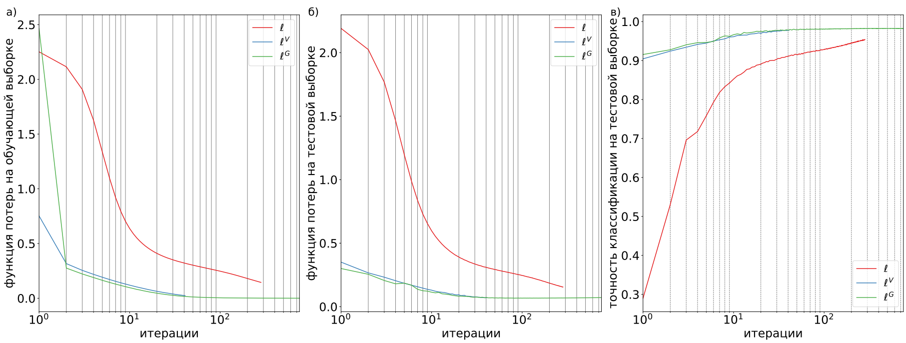



Рис. 4 – Результаты экспериментальных исследований на наборе mnist: а) потери на обучающей выборке; б) потери на тестовой выборке; в) точность на тестовой выборке которого происходит ее вымирание [10]:

$$\ell(t; K, A, r, N_0) = A + (K - A)\mathrm{e}^{\ln_1\left(\frac{N_0+A}{K-A}\right)\mathrm{e}^{-rt}}. \tag{11}$$

На рис. 5 приведено графическое представление влияния верхней и нижней асимптот при заданных $r = 1$ и $N_0 = 0.3$ переменного экспоненциального хвоста на устойчивость и смещенность оценок. Более высокие значения верхней асимптоты $K$ указывают на замедление или ускорения сходимости, присутствие преждевременной сходимости (зеленая зона), тогда как повышение значений нижней асимптоты $A$ сигнализирует об отсутствии сходимости (красная зона), в частности, ввиду наличия выбросов или намеренно искаженных наблюдений.

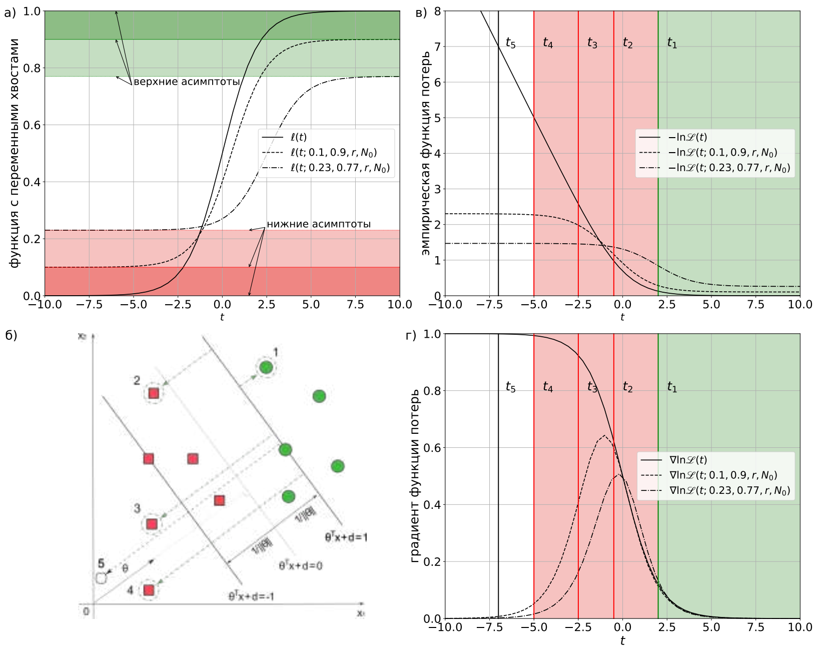



Рис. 5 – Влияние величины асимптот экспоненциального хвоста на устойчивость и смещенность оценок: а) функции с переменными хвостами; б) эмпирическая функция потерь; в) разделяющая гиперплоскость с верными (1), шумовыми (2,3,4) и намеренно искаженными (5) наблюдениями; г) градиент эмпирической функции потерь

На рис. 6 приведены результаты экспериментальных исследований на искусственно сгенерированной линейно разделимой выборке наблюдений, где количество признаков $n = 10$, количество наблюдений изменялось приращениями $m_k \in [20, 100]$, а метки классов для случайно выбранной доли наблю-

[10] Kulikovskikh, I., Prokhorov, S., Lipić, T., Legović, T., Šmuc, T. BioGD: Bio-inspired robust gradient descent. PLoS ONE 14(7): e0219004, 2019.

дений $p = \{0, 0.15, 0.3\}$ были изменены на противоположные. На рис. 7 приведены результаты для расширенного набора $p = \{0, 0.05, 0.1, 0.15, 0.2, 0.25, 0.3\}$ с усредненными значениями и разбросом значений эмпирической функции потерь при использовании переменных и постоянных экспоненциальных хвостов, просуммированные на всем диапазоне $m_k \in [20, 100]$. Для построения модели регрессии использовался метод градиентного спуска со значением шага по умолчанию $\eta = 1.5e-8$. Для оптимизации параметров переменного хвоста использовался поиск по сетке в диапазонах $A \in [0, 0.15]$ и $K \in [0.85, 1]$. Выборка наблюдений была разбита случайным выбором на обучающую и контрольную, которая составляла 20 % от исходной.

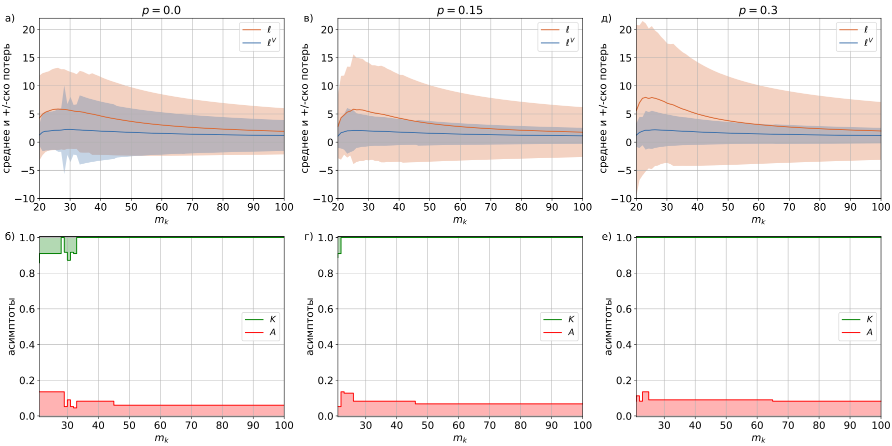



Рис. 6 – Результаты экспериментальных исследований на модельных данных с различных уровнем шумовых меток $p = \{0, 0.15, 0.3\}$: а) потери на обучении; б) потери на контроле; в) точность на контроле

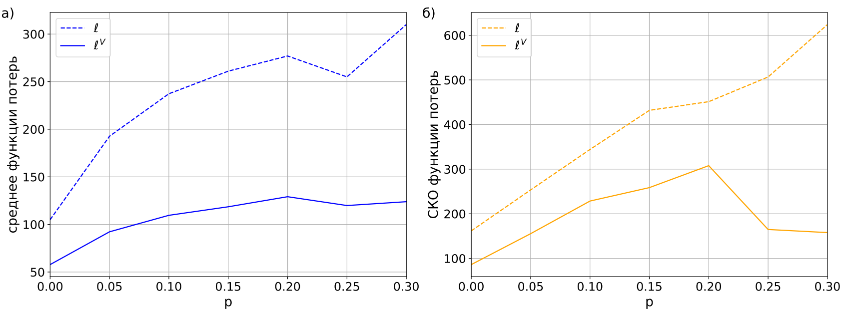



Рис. 7 – Результаты экспериментальных исследований для расширенного набора $p = \{0, 0.05, 0.1, 0.15, 0.2, 0.25, 0.3\}$: а) среднее эмпирической функции потерь; б) СКО эмпирической функции потерь

На рис. 8 приведены результаты экспериментальных исследований на различных наборах данных, которые показывают усредненные значения и разброс значений эмпирической функции потерь при использовании переменных и постоянных экспоненциальных хвостов, просуммированные на всем диапазоне $m_k \in [20, 100]$. Наилучший результат достигнут на наборах данных pima и *Heart Statlog* (heart) (см. рис. 9 ), где изменения значений асимптот более заметны и указывают на исключение проблем, связанных с отсутствием сходимости, ее замедлением или ускорением, а также наличием преждевременной сходимости.

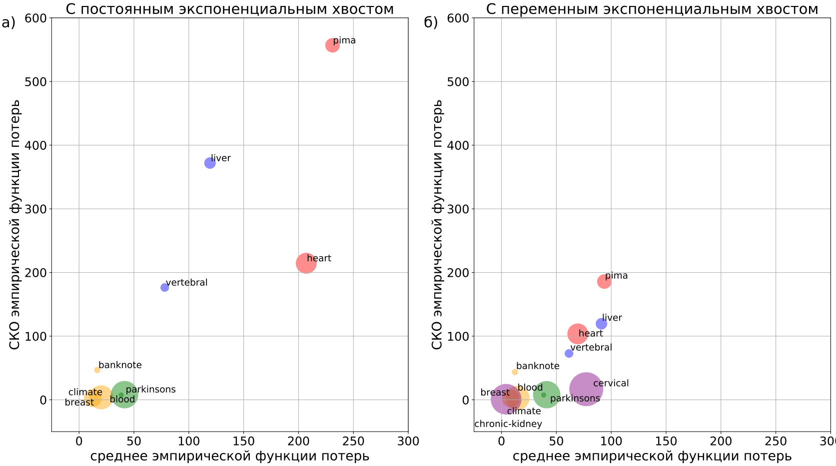



Рис. 8 – Результаты экспериментальных исследований на различных наборах данных: а) для функции с постоянным хвостом; б) для функции с переменным хвостом

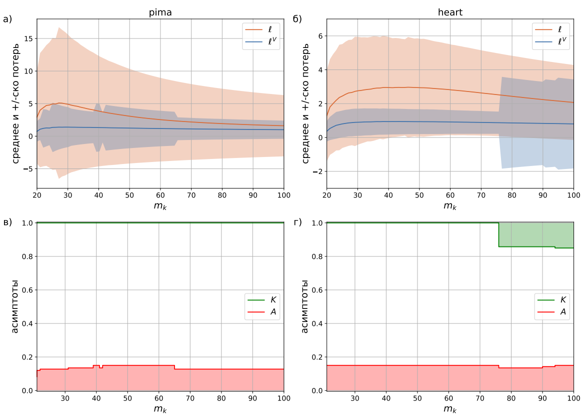



Рис. 9 – Результаты экспериментальных исследований на наборах pima и heart: а) потери на обучении; б) потери на контроле; в) точность на контроле

В главах 3 и 4 рассматриваются применения моделей с переменными экспоненциальными хвостами в теории фильтрации и теории тестирования. В целях компактности представления введем

$$c = f(T, K, r, N_0, E, q).$$

**В третьей главе** показано применение неявной регуляризации функций экспоненциального типа с переменными параметрами для понижения сложности фильтров Лагерра и Мейкснера.

Зададим переменный экспоненциальный хвост на основе пары параметров фильтра, описывающих его полюс и учитывающих скорость роста импульсной характеристики фильтра и величину ее асимптот. Интерпретируем параметры фильтра на основе (7) при $T = 0$, $E = 0$ и $q \to 0$. Вводя замену $\exp(-rt) = x, \quad -rt = \ln(x), \quad t = -\frac{\ln(x)}{r}$, получим $\exp(-cx)x \sim t^c \exp(-rt)$. Принимая во внимание последнее соотношение, зададим переменный экспоненциальный хвост полюсом фильтра $\gamma/\xi$, характеризующим скорость роста функции $r$, и параметром весовой функции $\alpha$, задающим величину $c$.

Применим модель с переменным экспоненциальным хвостом при формировании фильтра в $s$- и $z$-областях. Рассмотрим непрерывные обобщенные фильтры Лагерра с параметрами, описывающими полюс фильтра и весовую функцию.

**Определение 1.** *Предположим, что $k \in \mathbb{N}_0$, $\tau \in \mathbb{R}^+$, $A \in \mathbb{R}$ и при каждом $\alpha \in A$ с условием $\alpha > -1$ в русле работы* [11] *обобщенные полиномы Лагерра $L_k(\tau, \alpha)$ являются ортогональными с весовой функцией $\omega(\tau, \alpha) = \tau^\alpha \exp(-\tau)$ в гильбертовом пространстве $L_2(\mathbb{R}^+)$. Тогда для каждого $\gamma \in \Gamma \subset \mathbb{R}$ с условием $\gamma > 0$ систему $\left\langle \exp(-\gamma\tau/2) L_k(\gamma\tau, \alpha) \right\rangle_{k=0}^{\infty}$ будем называть системой обобщенных функций Лагерра и обозначать через $L_k(\tau, \gamma, \alpha)$.*

При этом система $\left\langle L_k(\tau, \gamma, \alpha) \right\rangle_{k=0}^{\infty}$ является ортогональной с неотрицательной весовой функцией $\omega(\tau, \alpha) = \tau^\alpha$ на интервале $\tau \in \mathbb{R}^+$ и нормой $\|L_k(\gamma, \alpha)\|^2 = \Gamma(k + \alpha + 1)/(k!\gamma^{\alpha+1})$.

**Лемма 2.** *Пусть для каждого $\alpha \in A$, где $A = \{\alpha > -1 : \alpha \in \mathbb{R}\}$, $\gamma \in \Gamma$, где $\Gamma = \{\gamma > 0 : \gamma \in \mathbb{R}\}$ и $k \in \mathbb{N}_0$, заданы обобщенные функции Лагерра $\left\langle L_k(\tau, \gamma, \alpha) \right\rangle_{k=0}^{\infty}$. Тогда обобщенные фильтры Лагерра $\left\langle \Lambda_k(s, \gamma, \alpha) \right\rangle_{k=0}^{\infty}$ в $s$-области с $s \in \mathbb{R}^+$ могут быть заданы в виде*

$$\Lambda_k(s, \gamma, \alpha) = \left( \frac{\gamma}{s + \gamma/2} \right)^{\alpha+1} \left( \frac{s - \gamma/2}{s + \gamma/2} \right)^k. \qquad (12)$$

Рассмотрим представления обобщенных фильтров Лагерра с параметрами полюса и весовой функции в дискретной области. Отличительной чертой структуры предлагаемых фильтров Мейкснера[12,13], построенных на основе функции с переменным экспоненциальным хвостом, от дискретных фильтров Лагерра является наличие ячеек бесконечной полосы, количество которых зависит от заданного параметра $\alpha$. При любых целочисленных значениях параметра $\alpha$, а не только для четных, как предлагалось ранее, фильтры имеют рациональную форму представления, и, таким образом, являются легко реализуемыми.

---

[11] Джексон, Д. Ряды Фурье и ортогональные полиномы: пер. с англ. М.,Л.: Главное издательство иностранной литературы, 1948.

[12] Prokhorov, S.A., Kulikovskikh, I.M. Pole position problem for Meixner filters. Signal Processing. 120: 8-12. 2016.

[13] Kulikovskikh, I.M. Meixner nonorthogonal filters//Automation and Remote Control. 79(8): 1458-1473, 2018.

**Определение 2.** *Пусть для каждого* $k \in \mathbb{N}_0$, $\xi \in |z| < 1$, $\alpha \subset \mathbb{R}$ *с условием* $\alpha > -1$ *задано* $z$*-преобразование обобщенных фильтров Лагерра* $Z\{\Lambda_k(s,\gamma,\alpha)\}$. *Тогда систему* $\left\langle Z\{\Lambda_k(s,\gamma,\alpha)\}\right\rangle_{k=0}^{\infty}$ *будем называть системой обобщенных фильтров Лагерра в* $z$*-области или фильтрами Мейкснера и обозначать через* $\left\langle G_k(z,\xi,\alpha)\right\rangle_{k=0}^{\infty}$.

**Лемма 3.** *Пусть задана система* $\left\langle \Lambda_k(s,\gamma,\alpha)\right\rangle_{k=0}^{\infty}$. *Тогда для каждого* $k \in \mathbb{N}_0$, $\xi \in |z| < 1$, $\alpha \subset \mathbb{R}$ *с условием* $\alpha > -1$ *и с учетом пары билинейных преобразований*

$$G_k(z,\xi,\alpha) \mapsto \frac{2a}{a+s} G_k\left(\frac{a+s}{a-s},\xi,\alpha\right) = \frac{(-1)^k}{1+\xi}\Lambda_k\left(s, 2a\frac{1-\xi}{1+\xi},\alpha\right), \qquad (13)$$

$$\Lambda_k(s,\gamma,\alpha) \mapsto \frac{z}{z+1}\Lambda_k\left(a\frac{z-1}{z+1},\gamma,\alpha\right) = \frac{(-1)^k 2a}{a+\gamma/2} G_k\left(z, \frac{a-\gamma/2}{a+\gamma/2},\alpha\right) \qquad (14)$$

*система фильтров Мейкснера* $\left\langle G_k(z,\xi,\alpha)\right\rangle_{k=0}^{\infty}$ *может быть представлена в виде*

$$G_k(z,\xi,\alpha) = \frac{(1-\xi^2)z}{z-\xi}\left(\frac{(1-\xi)(z+1)}{z-\xi}\right)^{\alpha}\left(\frac{1-\xi z}{z-\xi}\right)^{k}. \qquad (15)$$

Новое представление фильтров Мейкснера (15) приведено на рис. 10.

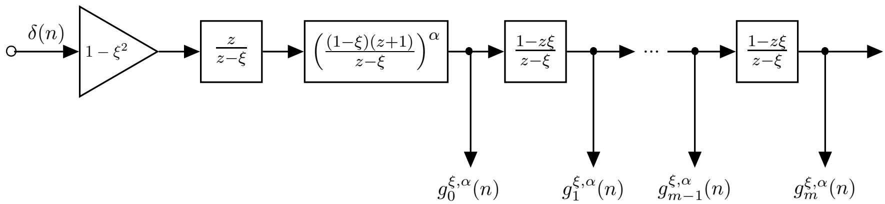



Рис. 10 – Новое представление БИХ фильтров Мейкснера

Пусть для каждого $\tau \in \mathbb{R}^+$, $m \in \mathbb{N}_0$, $\alpha \in A \subset \mathbb{R}$ с условием $\alpha > -1$ и $\gamma \in \Gamma \subset \mathbb{R}$ с условием $\gamma > 0$ функция $f(\tau)$ с учетом $\int_0^\infty \big(f(\tau)\big)^2 \mathrm{d}\tau < \infty$ может быть выражена как ограниченный ряд Лагерра с погрешностью $\Delta_m^\alpha(\gamma) = \sum_{k=m+1}^{\infty}\big(\beta_k^\alpha(\gamma)\big)^2 \big\|L_k^\alpha(\gamma)\big\|^2$, где коэффициенты $\beta_k^\alpha(\gamma) = \frac{\langle f(\tau), L_k^\alpha(\tau,\gamma)\rangle_{\omega^\alpha(\tau)}}{\|L_k^\alpha(\gamma)\|^2}$.

Применением неявную регуляризацию моделей с переменными экспоненциальными хвостами для понижения сложности непрерывного фильтра: $\gamma^* = \arg\min_{\gamma\in\Gamma} \Delta_m^\alpha(\gamma)$. В данной работе расширяются результаты предыдущих исследований[14,15], и предлагается строгое теоретическое обоснование единствен-

[14]Clowes, G. Choice of the time-scaling factor for linear system approximations using orthonormal Laguerre functions. IEEE Trans. Automatic Control. 10: 487-489, 1965.

[15]Волков, И.И., С.А. Прохоров Способ повышения точности аппроксимации корреляционных функций ортогональными функциями Лагерра. Приборостроение. 17: 66-72, 1974.

ности коэффициента, ведущего к оптимальному решению, для обобщенных функций Лагерра в непрерывной и дискретной областях [12,16]. Приведем основные положения.

**Лемма 4.** *Пусть $k \in \mathbb{N}_0$, $\alpha \in A$, где $A = \{\alpha > -1 : \alpha \in \mathbb{R}\}$, $\gamma \in \Gamma$, где $\Gamma = \{\gamma > 0 : \gamma \in \mathbb{R}\}$. Тогда справедливо рекуррентное соотношение $\beta_k^{\alpha+1}(\gamma) = \beta_k^{\alpha}(\gamma) - \beta_{k+1}^{\alpha}(\gamma)$.*

**Следствие 2.** *Если в условии Леммы 4 справедливо $\beta_{m+1}^{\alpha}(\gamma) = 0$, то $\beta_m^{\alpha}(\gamma) = \beta_m^{\alpha+1}(\gamma)$ и $\beta_{m+2}^{\alpha}(\gamma) = -\beta_{m+1}^{\alpha+1}(\gamma)$.*

**Лемма 5.** *Для каждого $m \in \mathbb{N}_0$, $\alpha \in A$, где $A = \{\alpha > -1 : \alpha \in \mathbb{R}\}$, $\gamma \in \Gamma$, где $\Gamma = \{\gamma > 0 : \gamma \in \mathbb{R}\}$, справедливо $|\beta_m^{\alpha+1}(\gamma)| > |\beta_m^{\alpha+2}(\gamma)|$ при $\beta_{m+1}^{\alpha}(\gamma) = 0$.*

**Лемма 6.** *Для каждого $m \in \mathbb{N}_0$, $\alpha \in A$, где $A = \{\alpha > -1 : \alpha \in \mathbb{R}\}$, $\gamma \in \Gamma$, где $\Gamma = \{\gamma > 0 : \gamma \in \mathbb{R}\}$, справедливо $\beta_m^{\alpha}(\gamma)\beta_{m+2}^{\alpha}(\gamma) < 0$ при $\beta_{m+1}^{\alpha}(\gamma) = 0$.*

Сформулируем основную теорему.

**Теорема 2.** *При любом $m \in \mathbb{N}_0$, $\alpha \in A$, где $A = \{\alpha > -1 : \alpha \in \mathbb{R}\}$, $\gamma \in \Gamma$, где $\Gamma = \{\gamma > 0 : \gamma \in \mathbb{R}\}$, каждое решение $\gamma^* = \gamma^*(\alpha) \in \Gamma$ может быть представлено в виде $\gamma^* = \arg\min\limits_{\gamma\in\Gamma} \Delta_m^{\alpha}(\gamma)$ или $\gamma^* = \arg\max\limits_{\gamma\in\Gamma} \Delta_m^{\alpha}(\gamma)$. При этом в первом случае $\beta_{m+1}^{\alpha}(\gamma^*) = 0$, а во втором случае $\beta_m^{\alpha}(\gamma^*) = 0$.*

В системе $\left\langle G_k(z,\xi,\alpha)\right\rangle_{k=0}^{\infty}$ для каждого $k \in \mathbb{N}_0$, $\xi \in |z| < 1$, $\alpha \subset \mathbb{R}$ с условием $\alpha > -1$ зададим передаточную функцию $H(z)$, которая может быть представлена ортогональным рядом с погрешностью приближения $\Delta_m(\xi,\alpha) = \left\| H - \sum_{k=0}^{m} \beta_k(\xi,\alpha) G_k(\xi,\alpha) \right\|^2$, где коэффициенты $\tilde{\beta}_k(\xi,\alpha) = \langle H, G_k(\xi,\alpha)\rangle$ в силу неортогональности системы определяются с помощью следующего преобразования $\langle H, G_n(\xi,\alpha)\rangle = \sum_{k=0}^{m} \langle G_k(\xi,\alpha), G_n(\xi,\alpha)\rangle \beta_k(\xi,\alpha)$. Отсюда, $\beta_k(\xi,\alpha) = \sum_{k=0}^{m} M_{k,n}^{-1}(\xi,\alpha)\tilde{\beta}_k(\xi,\alpha), \quad M_{k,n}(\xi,\alpha) = \langle G_k(\xi,\alpha), G_n(\xi,\alpha)\rangle$.

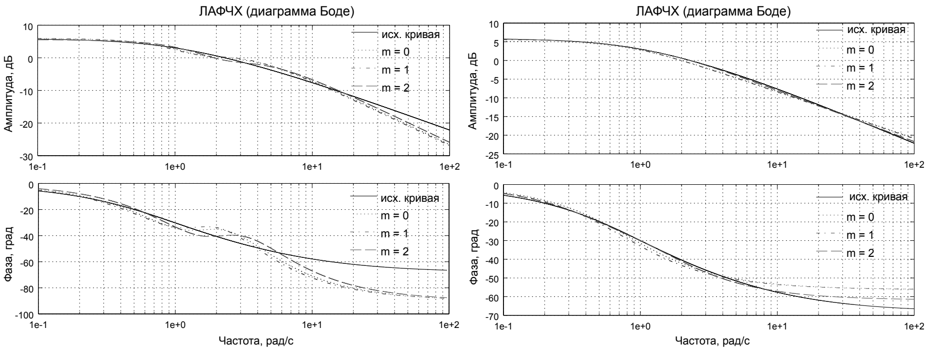
Применением неявную регуляризацию моделей с переменными экспоненциальными хвостами для понижения сложности дискретного фильтра: $(\xi,\alpha)^* = \arg\min\limits_{\substack{|\xi|<1 \\ \alpha>-1}} \Delta_m(\xi,\alpha)$. Обобщим Теорему 2 на случай фильтров Мейкснера. Введем $\Xi = \{\xi : \xi \in |z| < 1\}$. Тогда, имея $\Gamma = \{\gamma > 0 : \gamma \in \mathbb{R}\}$, зададим композицию областей определения полюсов в виде $\Phi \in \{\Gamma, \Xi\}$.

**Теорема 3.** *При любом $m \in \mathbb{N}_0$, $\alpha \in A$, где $A = \{\alpha > -1 : \alpha \in \mathbb{R}\}$, $\phi \in \Phi$, каждое решение $\phi^* = \phi^*(\alpha) \in \Phi$ при условии $\beta_k^{\alpha+1}(\phi) = \beta_k^{\alpha}(\phi) \mp \beta_{k+1}^{\alpha}(\phi)$ может быть представлено в виде $\phi^* = \arg\min\limits_{\phi\in\Phi} \Delta_m^{\alpha}(\phi)$ или $\phi^* = \arg\max\limits_{\phi\in\Phi} \Delta_m^{\alpha}(\phi)$. При этом в первом случае $\beta_{m+1}^{\alpha}(\phi^*) = 0$, а во втором случае $\beta_m^{\alpha}(\phi^*) = 0$.*

[16] Prokhorov, S.A., Kulikovskikh, I.M. Unique condition for generalized Laguerre functions to solve pole position problem. Signal Processing. 108: 25-29, 2015.

Из Теоремы (3) следует, что применение функций с переменным экспоненциальным хвостом позволяет понизить сложность рекурсивных фильтров, сводя соответствующий коэффициент фильтра к нулю в непрерывной и дискретной областях.

На рис. 11 приведены диаграммы Боде (ЛАФЧХ), построенные при синтезе фильтра дробного порядка, заданного передаточной функцией вида $H(s) = 1/(s^{0.7}+2)+1/(s^{0.8}+2)+1/(s^{0.9}+1)$ с помощью обобщенного фильтра Лагерра.




а) б)

Рис. 11 – Диаграммы Боде: а) $\gamma^* = 1.66$ и $\alpha = 0$; б) $\gamma^* = 1.66$ и $\alpha^* = -0.375$

Из рисунка видно, что при выборе целочисленного значения $\alpha = 0$, ближайшего к оптимальному $\alpha = -0.375$, увеличение порядка фильтра $m$ не приводит к улучшению результата. Неявная регуляризация модели фильтра со значением $\alpha^*$ позволяет получить лучший результат при идентичном порядке или эквивалентный результат при меньшем порядке, что, таким образом, свидетельствует о снижении сложности фильтра.

**В четвертой главе** показано применение неявной регуляризации функций экспоненциального типа с переменными параметрами для понижения сложности модели Бартона при индивидуальном и групповом тестировании.

Зададим переменный экспоненциальный хвост на основе пары параметров тестирования, описывающих поведение респондентов и учитывающих скорость роста уровня знаний и величину асимптот, а именно минимально и максимально возможных уровней.

Интерпретируем параметры модели тестирования на основе (11) при $q = 1$ и $N_0 = \frac{K-3A}{2}$, где переменный экспоненциальный хвост определяется скоростью роста уровня знаний $\alpha$, задающим скорость роста функции $r$, и параметрами угадывания $a$ и забывания $b$, задающими величину $c$ [17].

Применим модель с переменными экспоненциальными хвостами при формировании модели тестирования Бартона, определяющей вероятность $h$ пра-

[17] Kulikovskikh, I. Cognitive validation map for early occupancy detection in environmental sensing. Engineering Applications of Artificial Intelligence. 65: 330-335, 2017.

вильного ответа на задание $i$ [7]:

$$h_i(\theta) = a_i + \frac{b_i - a_i}{1 + \exp(-\alpha_i(\theta - \beta_i))}, \tag{16}$$

где $\theta$ задает уровень знаний респондента, $\beta_i$ - сложность задания $i$, $\alpha_i$ - дискриминационная мощность задания $i$, задающая скорость роста уровня знаний, $b_i$ - вероятность забывания правильного ответа на задание $i$, а $a_i$ - вероятность чистого угадывания правильного ответа на задание $i$.

Для выполнения неявной регуляризации модели тестирования, определим параметр переменного экспоненциального хвоста $c^r$, который учитывает значения и взаимное влияние асимптот $c$ при скорости роста уровня знаний $r$ при индивидуальном тестировании в терминах нечеткой логики.

**Определение 3.** *Пусть $U$ - нечеткое множество ответов респондента в условиях индивидуального тестирования. Тогда $c^r = \{(u, \mu_{c^r_L(u)})|u \in U\}$, где $\mu_{c^r_L(u)} : U \to [0,1]$ - функция принадлежности $c^r_L \subset U$, будем называть степенью угадывания.*

При этом величина $c^r_L$ связана с уровнем сложности заданий $L$, разработанных в соответствии с таксономией Блума, и является обобщением $c^r \in [0,1]$.

Для анализа влияния нечеткой задания параметров на оценки модели было проведено реальное тестирование $J = 9$ учащихся университета, которым было предложено 30 минут для выполнения набора тестов $I = 9$ с множественным выбором на предлоги английского языка. Тесты даны от более простых (LOTS) к более сложным (HOTS). Уровни сложности заданий для тестирования были заданы в виде $L = \{0.1, 0.26, 0.42, 0.58, 0.74, 0.9\}$, где каждое из значений соответствует центру треугольной функции принадлежности согласно Определению 3. Для каждого задания использовано одинаковое значение штрафа за попытку угадывания $p = 0.1$. Уровень способностей оценен косвенно и установлен равным $\theta = 0.6$ для каждого участника. При этом респондентам не было разрешено обсуждать задания и ответы с другими участниками тестирования.

На рис. 12 приведены результирующие оценки: истинные $\tilde{u}$, вероятностные $\tilde{u}^c$ и нечеткие $\tilde{u}^{c^r}$. Из рисунка следуют следующие выводы. Все приведенные оценки указывают на убывающий тренд в распределении ответов согласно требуемым уровням знаний. Видна взаимосвязь между уровнем знаний и степенью угадывания: если $\theta_j < l_i$, истинные оценки $\tilde{u}$ и предложенные нечеткие оценки $\tilde{u}^{c^r}$ демонстрируют более близкие результаты в сравнении с вероятностными $\tilde{u}^c$ ввиду возрастающей степени чистого угадывания. Таким образом, определение переменного хвоста в терминах нечеткой логики не требует явного задания параметров модели, строгого формирования инструкций согласно уровней сложности, что позволяет понизить сложность модели и получить менее смещенные оценки.

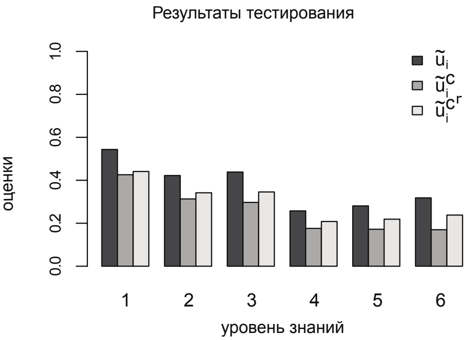



Рис. 12 – Результирующие оценки тестирования: истинные $\tilde{u}$, вероятностные $\tilde{u}^c$ и нечеткие $\tilde{u}^{c^r}$

Расширим Определение 3 на случай группового тестирования [18]. Для оценивания параметров модели (16) было проведено реальное тестирование, выполнен статистический анализ результатов тестирования $J = 84$ учащихся университета, которым было предложено 30 минут для выполнения набора тестов $I = 9$ с множественным выбором на предлоги английского языка. Тесты были даны от более простых (LOTS) к более сложным (HOTS). Респонденты были дополнительно смотивированы принимать активное участие в обсуждении и демонстрировать наилучшие результаты. Все участники имели (получали) инженерное образование и их возраст не превышал 30 лет. Вся выборка респондентов была разбита на 2 группы на основе уровня знаний английского языка согласно CEFR (The Common European Framework of Reference for Languages): тех, чей уровень выше $B2$ "hl-" и тех, чей языковой уровень ниже "ll-". При этом каждая из групп была случайно разбита на 2 подгруппы: одним был аннонсирован штраф на попытку угадывания $p = 0.25$ "-p", другим штраф не аннонсировался $p = 0$ "-np".

На рис. 13-15 приведены диаграммы с результатами тестирования по всем группам респондентов, в зависимости от их уровня знаний и наличия штрафа. На основании проведенного статистического анализа полученных оценок, были сформулированы следующие утверждения:

**H1** Чем ниже уровень знаний, тем больше чистых угадываний в результатах тестирования.

**H2** Чем выше уровень знаний, тем больше частичных угадываний в результатах тестирования.

**H3** Тестирование заданий, требующих низкого уровня знаний, стимулирует групповое тестирование в меньшей степени как в случае отсутствия штрафа за попытку угадывания, так и при наличии.

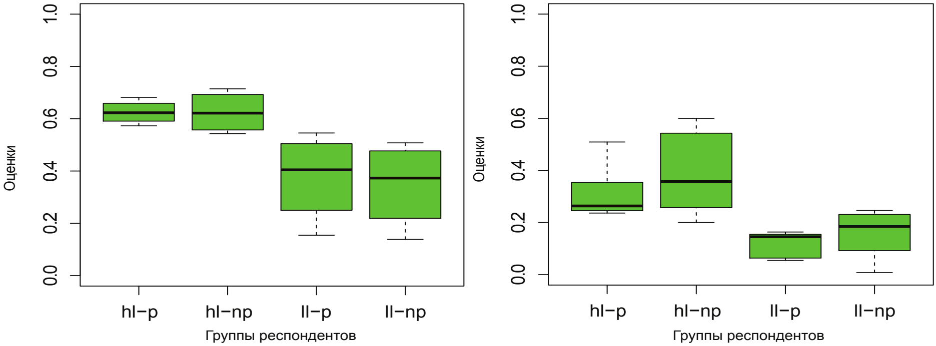
**H4** Тестирование заданий, требующих высокого уровня знаний, стимулирует групповое тестирование в большей степени в случае отсутствия штрафа за попытку угадывания.

**H5** Респонденты с низким уровнем знаний демонстрируют больший уровень чистых угадываний и учатся от группового тестирования в меньшей степени.

**H6** Респонденты с высоким уровнем знаний демонстрируют больший уро-

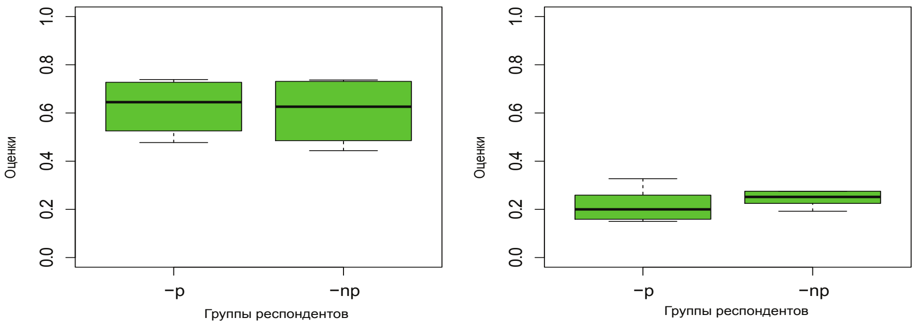
[18] Kulikovskikh, I., Prokhorov, S., Suchkova, S. Promoting collaborative learning through regulation of guessing in clickers. Computers in Human Behavior. 75: 81-91, 2017.

вень частичных угадываний и учатся от группового тестирования в большей степени.




а) б)

Рис. 13 – Результаты тестирования: а) LOTS; б) HOTS




а) б)

Рис. 14 – Результаты тестирования в зависимости от наличия штрафа: а) LOTS; б) HOTS

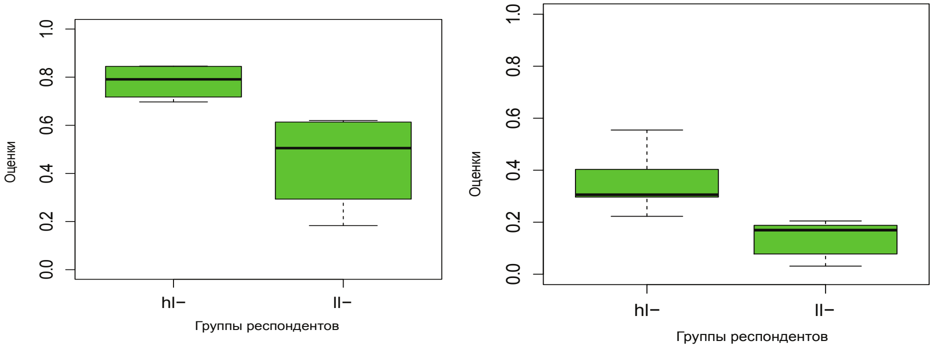
С учетом сформулированных утверждений, дополним Определение 3.

**Определение 4.** *Пусть* $u \in U$, $A([0,1])$ *- совокупность всех замкнутых подинтервалов в диапазоне* $[0,1]$. *В русле работы*[19], $A([0,1]) = \{u = [\underline{u}, \overline{u}]\} | (\underline{u}, \overline{u}) \in [0,1]^2 \cup \underline{u} \leq \overline{u}\}$. *Тогда отображение* $c_L^r : U \rightarrow A([0,1])$ *будем называть степенью угадывания в условиях группового тестирования.*

При этом $c_L^r(u) = [\underline{c_L^r}(u), \overline{c_L^r}(u)] \in A([0,1])$, где $\underline{c_L^r} : U \rightarrow [0,1]$ и $\overline{c_L^r} : U \rightarrow [0,1]$ - отображения, определяющие нижнюю и верхнюю границы интервала $c_L^r(u)$.

**Определение 5.** *Пусть* $c_L^{rup}(u) \in [0,1]$, $c_L^{rlow}(u) \in [0,1]$ *и* $c_L^{rlow}(u) \leq c_L^{rlow}(u)$ *для* $u \in U$. *Тогда* $FOU(c_L^r) = \bigcup_{u \in U} J_u$, *где* $c_L^{rup}(u) = \overline{FOU(c_L^r)}$, $c_L^{rlow}(u) = \underline{FOU(c_L^r)}$, *будем называть мерой влияния группового тестирования.*

[19] Bustince, H., Fernandez, J., Hagras, H., Pagola, M., Barrenechea, E. Interval type-2 fuzzy sets are generalization of interval-valued fuzzy sets: Towards a wider view on their relationship. IEEE Tran. On Fuzzy Sets. 23(5): 1876–1882, 2014.




а) б)

Рис. 15 – Результаты тестирования в зависимости от уровня знаний респондентов: а) LOTS; б) HOTS

Применение функции с переменным экспоненциальным хвостом позволяет понизить сложность модели Бартона для индивидуального и группового тестирования за счет замещения явных вероятностных параметров, описывающих степени забывания и угадывания обучаемых для каждого уровня сложности, на их нечеткие определения.

**В пятой главе** приведено описание комплекса программ для построения функций экспоненциального типа с переменными параметрами, включающее функциональное и алгоритмическое описание. Перечислены прикладные задачи, решенные с использованием комплекса программ. Приведены результаты апробации комплекса программ.

На рис. 16 приведено функциональное описание моделей на основе функций экспоненциального типа с переменными параметрами, положенное в основу алгоритмического описания. Для компактности представления алгоритмическая реализация представлена в терминах пространственной схемы взаимодействия объектов и состоит из 12 пронумерованных функциональных блоков (ФБ), соединенных ключом •— , показывающим переход от одного функционального пространства к другому (точкой обозначается исходное пространство):

- ФБ $\{1\}$ выявляет параметры переменного экспоненциального хвоста, требуемые для повышения сходимости градиентных методов при неявной регуляризации, со входным ключом $\{X, y\}$;
- ФБ $\{2,4,5,6\}$ определяет параметры переменного экспоненциального хвоста в задачах фильтрации со входным ключом $\{D, f\}$;
- ФБ $\{3,7,8,9\}$ определяет параметры переменного экспоненциального хвоста в задачах тестирования со входным ключом $\{X, u\}$;
- ФБ $\{10,11\}$ определяет параметры переменного экспоненциального хвоста, позволяющие повысить скорость сходимости градиентных методов при неявной регуляризации;

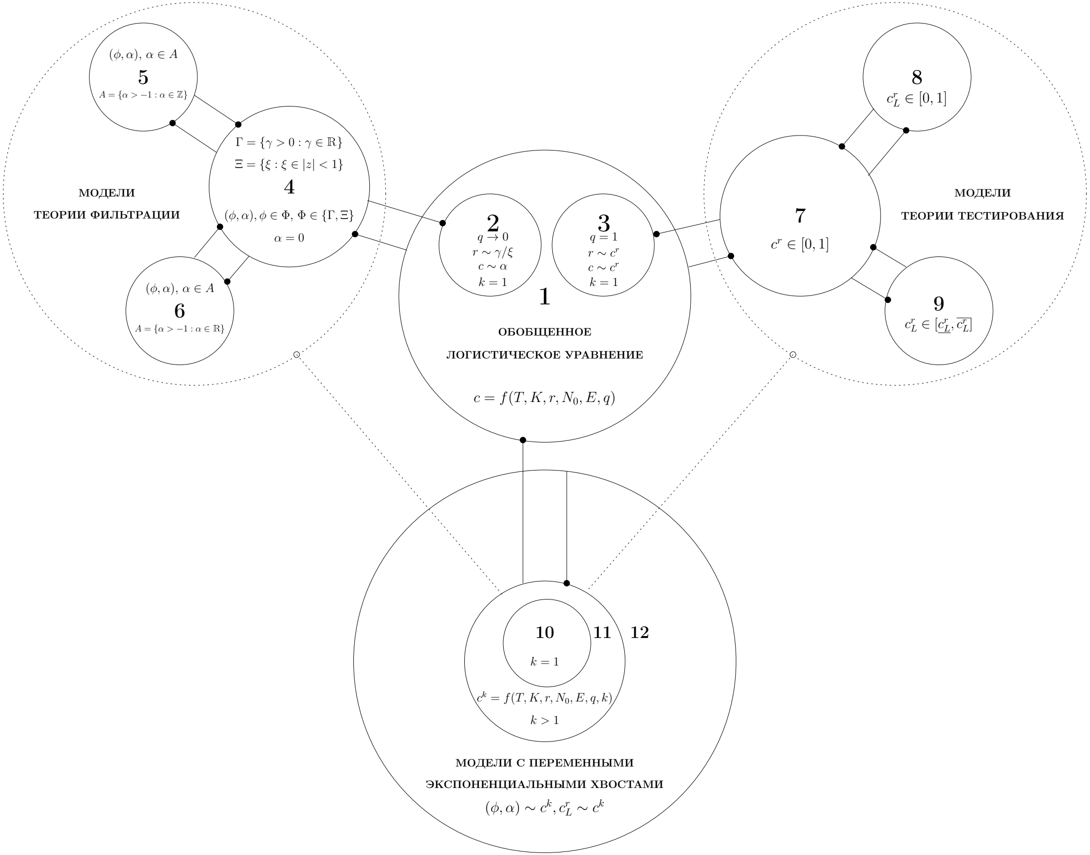



Рис. 16 – Функциональное описание моделей на основе функций экспоненциального типа с переменными параметрами

- ФБ {12} определяет параметры переменного экспоненциального хвоста при построении регрессионных моделей, на основе ее представлений в задачах фильтрации и тестирования.

Алгоритмическое описание переменных экспоненциальных хвостов для построения моделей на основе рассмотренных интерпретаций ее параметров в задачах фильтрации и тестирования представлено на рис. 17. Представленные алгоритмы были апробированы при решении ряда прикладных задач, в частности: проведении корреляционно-спектрального анализа в условиях ограниченных вычислительных ресурсов, интеллектуального анализа фотоплетизмографических сигналов, распознавании струнных музыкальных инструментов по спектру звуковых сигналов, нечеткого оценивания знаний студентов в условиях внутренней неопределенности, выявлении аномалий при решении задачи обнаружения живых объектов на основе неявных экологических факторов, реализации модели обучения машин на основе эффекта забывания, вызванного извлечением информации и т.д.

Комплекс программ, разработанный на основе алгоритмического описания, был также апробирован при формировании эффективных команд и оптимизации взаимодействия между их членами в организациях различного профиля. Для измерения параметров членов команды был использован набор тестовых заданий на основе проверенного инструментария, включающего опросник по социально-демографическим параметрам, тесты для выявления распределения ролей в команде, измерения выраженности социального интеллекта, определения стратегий поведения в конфликте, субъективной оценки сложившейся модели команды и т.д. Валидность пула проверялась дублированием, где схожесть по нескольким шкалам позволяла сделать вывод о достоверности полученных оценок. Апробация комплекса программ включала как сбор данных, включающий программные и аппаратные механизмы отслеживания поведения респондентов во время тестирования, так и дополнительный контроль уровня достоверности ответов на тестовые задания через выявление и маркирование смещенных оценок, построенных на некорректных, хаотичных и сфальсифицированных ответах.

На основе алгоритмического описания моделей на основе функций с переменными экспоненциальными хвостами была построена модель коллаборативной фильтрации ( Collaborative Filtering) со встраиванием слоев (Embeddings). На рис. 18 приведены результаты анализа ответов 1000 респондентов на 471 задание при проведении реального тестирования с помощью метода главных компонент (PCA). Ответы на тестовые задания были предварительно обработаны с использованием соответствующих психометрических шкал, которые сформировали карту из 62 параметров для каждого респондента. Красным маркером в пространствах тестовых заданий и респондентов помечены идентификаторы параметров и респондентов соответственно, которые внесли смещение в результирующие оценки, но были скорректированны с помощью разработанного комплекса программ.

Построение более достоверных оценок в результате проведенной апробации позволило снизить временные затраты на формирование эффективных команд, сформировать оптимальный пул методик, дающий достоверный результат и занимающий адекватное количество времени респондентов на заполнение, а также создать инструментарий для создания эффективных команд с возможностью присоединения нового участника команды к существующей с целью повышения производительности труда.

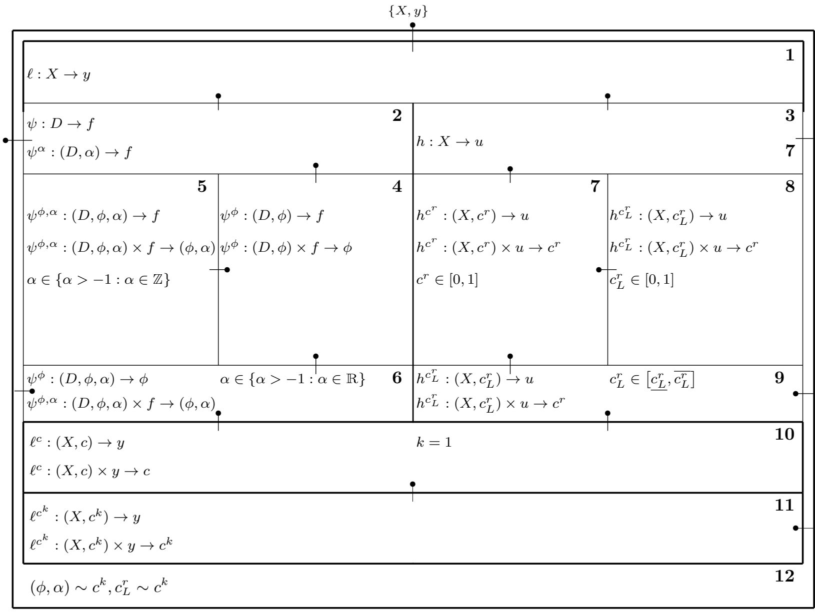



Рис. 17 – Алгоритмическое описание моделей на основе функций экспоненциального типа с переменными параметрами

**В заключении** сформулированы основные выводы, перечислены полученные в работе результаты.

В результате выполнения диссертационной работы были решены поставленные задачи и получены следующие результаты:

1. Проведен анализ основных подходов к повышению эффективности градиентных методов для функций экспоненциального типа. Выявлен недостаток адаптивных градиентных методов, связанный с низкой скоростью и устойчивостью. Предложено решение, связанное с построением функций экспоненциального типа с переменными параметрами.
2. Введено определение функций экспоненциального типа с переменными параметрами на основе обобщенного логистического уравнения, которое учитывает ее скорость роста и величину асимптот.

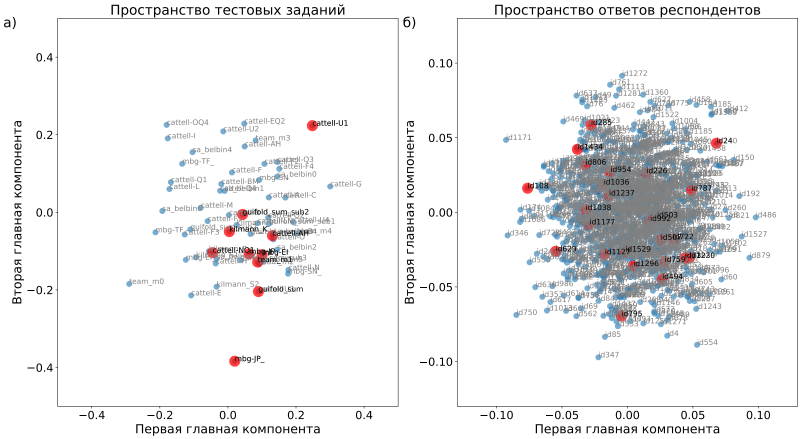



Рис. 18 – Выявление смещения при реальном психометрическом тестировании: а) в пространстве тестовых заданий; б) в пространстве ответов респондентов

3. Повышена скорость сходимости и устойчивость градиентных методов за счет неявной регуляризации функций экспоненциального типа с переменными параметрами.
4. Предложено определение фильтров Лагерра и Мейкснера на основе функций экспоненциального типа с переменными параметрами, заданных полюсами фильтров, характеризующими скорость роста, и параметром весовой функции, задающим величину асимптот. Снижена сложность рекурсивных фильтров Лагерра и Мейкснера за счет сведения соответствующего коэффициента фильтра к нулю. Предложено определение модели Бартона при индивидуальном и групповом тестировании на основе функций экспоненциального типа с переменными параметрами, заданных скоростью роста уровня знаний, и параметрами забывания и угадывания, определяющими величину асимптот. Снижена сложность модели Бартона при индивидуальном и групповом тестировании за счет замещения явных вероятностных параметров модели на их нечеткие определения.
5. Проведены экспериментальные исследования для анализа функций экспоненциального типа с переменными параметрами. Результаты вычислительных экспериментов подтвердили эффективность предлагаемых моделей. Снижены вычислительные затраты на получение результирующих оценок за счет повышения скорости сходимости. Снижена смещенность результирующих оценок за счет повышения устойчивости к выбросам, пропущенным и намеренно искаженным наблюдениям. Наилучший результат достигнут на наборах данных при заметных изменениях значений асимптот функций, указывающих на исключений проблем, связанных с отсутствием сходимости, ее замедлением или ускорением, а также наличием преждевременной сходимости.

6. Предложены программные реализации моделей, построенных на основе функций экспоненциального типа с переменными параметрами, включающие их компактное функциональное и алгоритмическое описание на основе пространственной схемы взаимодействия объектов.
7. Проведена апробация предложенных моделей и их программных реализаций при решении задач, связанных с формированием эффективных команд и оптимизацией взаимодействия между их членами в организациях различного профиля, разработке систем обработки данных реального времени, а также внедрение. Научно обоснованные технические решения приведены в актах внедрения, которые прилагаются.

Основные результаты диссертации опубликованы в работах:

**Статьи в изданиях, рекомендованных ВАК:**


1. **Куликовских, И.М.** Понижение сложности модели индивидуального и группового адаптивного тестирования с множественным выбором на основе нечеткой когнитивной карты/И.М. Куликовских, С.А. Прохоров//Программные системы и вычислительные методы. – 2018. – 14. – С. 15-26.
2. **Куликовских, И.М.** Неортогональные фильтры Мейкснера/И.М. Куликовских//Автоматика и Телемеханика. – 2018. – 8. – С. 111-118.
3. Прохоров, С.А. Регуляризованные ортогональные модели вероятностных характеристик с условием выполнения их основных свойств/С.А. Прохоров, **И.М. Куликовских**//Программные продукты и системы. – 2018. – Т. 31, №1. – С. 99-101.
4. **Куликовских, И.М.** Выявление аномалий в пространстве экологических признаков для повышения точности обнаружения живых объектов в здании/И.М. Куликовских//Компьютерная оптика. – 2017. – Т. 41, №1. – С. 126-133.
5. **Куликовских, И.М.** Метод повышения интерпретируемости регрессионных моделей на основе трехступенчатой модели развития мышления/И.М. Куликовских//Программные продукты и системы. – 2017. – Т. 30, №4. – С. 601-608.
6. **Куликовских, И.М.** Формирование пространства признаков для обнаружения живых объектов в здании на основе экологических факторов/И.М. Куликовских//Известия СНЦ РАН. – 2016. – Т. 18, №4(4). – С. 754-759.
7. **Куликовских, И.М.** Комплексная система коллаборативного обучения на основе нечетких моделей для описания поведения систем с частичным знанием/И.М. Куликовских, С.А. Прохоров, С.А. Сучкова, Е.В. Матыцин// Известия СНЦ РАН. – 2016. – Т. 18, №4(4). – С. 760-765.
8. **Куликовских, И.М.** Вычисление коэффициентов неортогональных фильтров Мейкснера в системе GNU Octave/И.М. Куликовских//ЖУРНАЛ РАДИОЭЛЕКТРОНИКИ. – 2016. – №6. url: `http://jre.cplire.ru/jre/jun16/3/text.pdf`.
9. Прохоров, С.А. Условие оптимальности фильтров Мейкснера/С.А. Прохоров, **И.М. Куликовских**//ЖУРНАЛ РАДИОЭЛЕКТРОНИКИ. – 2015. – №4. url: `http://jre.cplire.ru/mac/apr15/9/text.pdf`.
10. Прохоров, С.А. Система адаптивного обучения на основе иерархических ко-

нечных автоматов/С.А. Прохоров, **И.М. Куликовских**//Известия СНЦ РАН. – 2015. – Т. 17, №2(5). – С. 1087-1091.

11. Прохоров, С.А. Кластеризация диагностических тестов при изучении предлогов английского языка в соответствии с таксономией Блума/С.А. Прохоров, С.А. Сучкова, **И.М. Куликовских**//Известия СНЦ РАН. – 2015. – Т. 17, №2(5). – С. 1097-1103.
12. Прохоров, С.А. Программная реализация оценивания коэффициентов Фурье при ограниченных вычислительных ресурсах/С.А. Прохоров, **И.М. Куликовских**//Программные продукты и системы. – 2015. – №3(111). – С. 113-118.
13. Прохоров, С.А. Создание комплекса программ на основе пространственной схемы взаимодействия объектов/С.А. Прохоров, **И.М. Куликовских**// Программные продукты и системы. – 2012. – №3. – С. 5-8.
14. Прохоров, С.А. Численно-аналитический подход к вычислению интегралов при построении ортогональных моделей/С.А. Прохоров, **И.М. Куликовских**// Вестн. Сам. гос. техн. ун-та. Серия физ.-мат. науки. – 2009. – №2(19). – С. 140-146.
15. Прохоров, С.А. Аппроксимация корреляционных функций и спектральных плотностей мощности ортогональными функциями Сонина-Лагерра/С.А. Прохоров, **И.М. Куликовских**//Вестн. Сам. гос. техн. ун-та. Серия физ.-мат. науки. – 2008. – №2(17). – С. 185-191.
16. Прохоров, С.А. Частотные характеристики ортогональных функций Сонина-Лагерра / С.А. Прохоров, **И.М. Куликовских**//Вестн. Сам. гос. техн. ун-та. Серия физ.-мат. науки. – 2007. – №2(15). – С. 123-127.

**Статьи, индексируемые в международных базах WoS/Scopus:**

17. **Kulikovskikh, I.** BioGD: Bio-inspired robust gradient descent/I. Kulikovskikh, S. Prokhorov, T. Lipić, T. Legović , T. Šmuc//PLoS One. – 2019. – 14(7). – pp. e0219004.
18. **Kulikovskikh, I.** An SGD-based meta-learner with “growing” descent/I. Kulikovskikh, S. Prokhorov, T. Legović, T. Šmuc//Journal of Physics: Conference Series. – 2019. – 1368. – pp. 052008.
19. **Kulikovskikh, I.M.** Psychological perspectives on implicit regularization: A model of retrieval-induced forgetting/I.M. Kulikovskikh, S.A. Prokhorov//Journal of Physics: Conference Series. – 2018. – 1096. – pp. 012079.
20. **Kulikovskikh, I.M.** Meixner nonorthogonal filters/I.M. Kulikovskikh//Automation and Remote Control. – 2018. – 79(8). – pp. 1458-1473.
21. **Kulikovskikh, I.M.** Cognitive validation map for early occupancy detection in environmental sensing/I.M. Kulikovskikh//Engineering Applications of Artificial Intelligence. – 2017. – 65. – pp. 330-335.
22. **Kulikovskikh, I.M.** Promoting collaborative learning through regulation of guessing in clickers/I.M. Kulikovskikh, S.A. Prokhorov, S.A. Suchkova//Computers in Human Behavior. – 2017. – 75. – pp. 81-91.
23. **Kulikovskikh, I.M.** Minimizing the effects of floor and ceiling to improve the convergence of log-likelihood / I.M. Kulikovskikh, S.A. Prokhorov//Procedia Engineering. – 2017. – 201. – pp. 779-788.
24. **Kulikovskikh, I.M.** Anomaly detection in an ecological feature space to improve

the accuracy of human activity identification in buildings / I.M. Kulikovskikh// Computer Optics. – 2017. – 41(1). – pp. 126-133.

25. Prokhorov, S.A. Pole position problem for Meixner filters / S.A. Prokhorov, **I.M. Kulikovskikh**//Signal Processing. – 2016. – 120. – pp. 8-12.
26. Prokhorov, S.A. Unique condition for generalized Laguerre functions to solve pole position problem / S.A. Prokhorov, **I.M. Kulikovskikh** // Signal Processing. – 2015. – 108. – pp. 25-29.
27. Prokhorov, S.A. Fuzzy learning performance assessment based on decision making under internal uncertainty/S.A. Prokhorov, **I.M. Kulikovskikh**//7$^{th}$ Computer Science and Electronic Engineering Conference (CEEC 2015). – Colchester, UK, 2015. – pp. 65-70.

**Монографии:**

28. Прохоров, С.А. Основные ортогональные функции и их приложения. Часть 1. Ортогональные функции экспоненциального типа/С.А. Прохоров, **И.М. Куликовских**. – Самара: Изд-во «Инсома-пресс», 2019. – 200 с.
29. **Куликовских, И.М.** Автоматизированная система корреляционно-спектрального анализа в ортогональных базисах Якоби / И.М. Куликовских // Прикладной анализ случайных процессов / под ред. С.А. Прохорова. – Самара: СНЦ РАН, 2007. - С. 347-359.

**Учебные пособия:**

30. Прохоров, С.А. Численные методы и алгоритмы аппроксимативного анализа корреляционно-спектральных характеристик в ортогональных базисах: учебное пособие/С.А. Прохоров, **И.М. Куликовских**. – Самара: Изд-во «Инсома-пресс», 2019. – 254 с.
31. Прохоров, С.А. Численные методы, алгоритмы и комплексы программ для проведения вычислительного и натурного экспериментов: учебное пособие/С.А. Прохоров, **И.М. Куликовских**. – Самара: Изд-во «Инсома-пресс», 2019. – 208 с.
32. Прохоров, С. А. Ортогональные модели корреляционно-спектральных характеристик случайных процессов: лабораторный практикум/С.А. Прохоров, **И.М. Куликовских**. – Самара: СНЦ РАН, 2008. – 301 с.

**Свидетельства о государственной регистрации программ для ЭВМ:**

33. **Куликовских, И.М.** Программа анализа влияния эффекта RIF на сходимость градиентных методов обучения/И.М. Куликовских/Свидетельство о государственной регистрации программ для ЭВМ. Рег. №2019667794 от 27.12.2019.
34. **Куликовских, И.М.** Программная реализация устойчивой биологически обоснованной модели обучения/И.М. Куликовских/Свидетельство о государственной регистрации программ для ЭВМ. Рег. №2019667690 от 26.12.2019.
35. **Куликовских, И.М.** Программа обучения машин на основе эффекта забывания, вызванного извлечением информации/И.М. Куликовских, Е.А. Пономарев/Свидетельство о государственной регистрации программ для ЭВМ. Рег. №2018665162 от 03.12.2018.
36. **Куликовских, И.М.** Программа нечеткого обучения машин/И.М. Куликовских, Е.А. Назарова/Свидетельство о государственной регистрации программ для ЭВМ. Рег. №2018664363 от 15.11.2018.

37. **Куликовских, И.М.** Программная реализация модели долговременной памяти неокогнитрона/И.М. Куликовских, Д.В. Безруков/Свидетельство о государственной регистрации программ для ЭВМ. Рег. №2018664362 от 15.11.2018.
38. **Куликовских, И.М.** Программа нечеткого оценивания знаний студентов в условиях внутренней неопределенности/И.М. Куликовских, С.А. Прохоров, С.А. Сучкова/Свидетельство о государственной регистрации программ для ЭВМ. Рег. №2015660296 от 16.12.2015.
39. **Куликовских, И.М.** Машинное обучение использованию предлогов в английском языке/И.М. Куликовских, С.А. Прохоров, С.А. Сучкова / Свидетельство о государственной регистрации программ для ЭВМ. Рег.№2015618129 от 31.07.2015.
40. **Куликовских, И.М.** Мобильное приложение для анализа коэффициентов разложения в условиях ограниченных вычислительных ресурсов/ И.М. Куликовских, С.А. Прохоров, Д.В. Целищев / Свидетельство о государственной регистрации программ для ЭВМ. Рег. №2014616733 от 08.09.2014.
41. **Куликовских, И.М.** Справочник по ортогональным функциям на базе ОС Android: ортогональные функции экспоненциального типа/И.М. Куликовских, С.А. Прохоров, А.П. Майоров/Свидетельство о государственной регистрации программ для ЭВМ. Рег. №2013660124 от 24.10.2013.
42. **Куликовских, И.М.** Мобильное приложение для спектрального анализа данных на базе iOS/И.М. Куликовских, С.А. Прохоров, Я.Ю. Богданова / Свидетельство о государственной регистрации программ для ЭВМ. Рег. №2013611627 от 30.01.2013.
43. **Куликовских, И.М.** Программный комплекс интеллектуального анализа данных: модуль обработки фотоплетизмографических сигналов / И.М. Куликовских, С.А. Прохоров, Н.С. Филиппова / Свидетельство о государственной регистрации программ для ЭВМ. Рег. №2013611625 от 30.01.2013.
44. Гребнев, В.В. Программа обработки базы данных нейтрализаторов/ В.В. Гребнев, Г.Д. Мальчиков, В.И. Заражевский, И.Е. Кравченко, С.А. Прохоров, **И.М. Куликовских**/Свидетельство о государственной регистрации программ для ЭВМ. Рег. №2010612790 от 06.10.2010.
45. **Куликовских, И.М.** Автоматизированная информационная система исследования обобщенных ортогональных многочленов Якоби/И.М. Куликовских, С.А. Прохоров/Свидетельство о государственной регистрации программ для ЭВМ. Рег. №2009614285 от 14.08.2009.
46. **Куликовских, И.М.** Автоматизированная система спектрально-корреляционного анализа методом ортогональных разложений «СКАН»/И.М. Куликовских, С.А. Прохоров/Свидетельство о государственной регистрации программ для ЭВМ. Рег. №2009613943 от 24.07.2009.

**Статьи в других изданиях:**

47. **Kulikovskikh, I.** Machines in classroom: Towards human-like active learning/I. Kulikovskikh, T. Šmuc//22nd International Conference on Discovery Science (DS 2019). – Split, Croatia, 2019. – pp. 43.
48. **Kulikovskikh, I.** Bio-inspired robust machine learning/I. Kulikovskikh, T. Šmuc //4th International Workshop on Data Science (IWDS 2019). – Zagreb, Croatia, 2019. – pp. 31-33.

49. **Kulikovskikh, I.** Growing descent of stochastic gradient with the generalized logistic map/I. Kulikovskikh, S. Prokhorov, T. Legović, T. Šmuc//4th Int. Conference on Information Technology and Nanotechnology (ITNT-2019). – Samara, Russia, 2019. – pp. 338-344.
50. **Kulikovskikh, I.M.** A method of implicit regularization based on the phenomena of retrieval-induced forgetting/I.M. Kulikovskikh, S.A. Prokhorov//4th Int. Conference on Information Technology and Nanotechnology (ITNT-2018). – Samara, Russia, 2018. – pp. 2132-2137.
51. **Куликовских, И.М.** Метод обучения машин на основе энактивистского подхода к извлечению информации/И.М. Куликовских, С.А. Прохоров // Аналитические и численные методы моделирования естественно-научных и социальных проблем: сборник статей XII Международной научно-технической конференции – Пенза: ПГУ, 2017. – С. 175-179.
52. **Kulikovskikh, I.M.** Modifications of log-likelihood to measure floor and ceiling effects/I.M. Kulikovskikh, S.A. Prokhorov//3rd Int. Conference on Information Technology and Nanotechnology (ITNT-2017). – Samara, Russia, 2017. – pp. 1849-1853.
53. **Kulikovskikh, I.M.** Some lightweight algorithms for scientific computing in mobile technologies/I.M. Kulikovskikh, S.A. Prokhorov//8th International Conference on Applied Mathematics and Scientific Computing. – Šibenik, Croatia, 2013. – pp. 40-41.