

ОТЗЫВ ОФИЦИАЛЬНОГО ОППОНЕНТА
на диссертационную работу Мотренко Анастасии Петровны
«Выбор моделей прогнозирования мультикоррелирующих рядов»,
представленную на соискание ученой степени
кандидата физико-математических наук по специальности
05.13.17 – «Теоретические основы информатики».

Временные ряды являются одним из общепризнанных инструментов анализа данных в области восстановления эмпирических закономерностей, характеризующих изменяющееся во времени состояние изучаемого динамического процесса. В последние годы, в эпоху больших данных (Big Data) и четвертой промышленной революции (Industry 4.0), наряду с традиционными приложениями временных рядов в эконометрике, социологии и медицине на передний план выходят прикладные задачи в области моделирования и управления автоматизированными производствами, сложными роботизированными системами, мониторинга и предсказания глобальных климатических изменений, анализа мультиспектральных данных дистанционного зондирования и т. п. Во всех перечисленных приложениях в качестве математической модели изучаемого процесса выступают временные ряды, фазовые пространства которых описываются большим количеством разнотипных и не всегда независимых параметров. Наряду с классической для теории временных рядов задачей выявления искомой эмпирической закономерности, эти новые практические постановки подтверждают актуальность изучения и ряда близких математических задач, среди которых:

- задача выбора наиболее адекватного признакового описания анализируемых процессов, в том числе связанного с погружением их фазовых множеств в пространства минимальной допустимой размерности;
- задача эффективной сегментации рядов;
- задача оценивания их длины, достаточной для классификации с заданным уровнем качества;
- а также проблема оценивания вычислительной сложности соответствующих комбинаторных задач, проектирования, теоретического и численного обоснования высокопроизводительных алгоритмов их решения.

Поэтому, **актуальность** тематики диссертационной работы А.П. Мотренко, связанной с исследованием перечисленных выше задач, **не вызывает сомнения**.

Результаты диссертации получены в рамках активно развиваемого алгебраического подхода к проблеме распознавания, опирающегося на фундаментальные работы акад. Ю.И. Журавлева и акад. К.В. Рудакова. В частности, А.П. Мотренко предложена авторская модификация квадратичного метода отбора признаков для рядов с тензорным описанием состояния; разработан новый метод вложения фазового пространства в пространство меньшей размерности, сохраняющий вероятностную структуру соседства описываемых объектов; разработаны методы автоматической сегментации квазипериодического временного ряда и оценивания достаточной длины обучающей выборки для классификации рядов. Предложенные автором работы алгоритмы прошли численное тестирование при решении ряда актуальных прикладных задач.

Предлагаемые подходы являются новыми, а форма представления материала работы и полученные выводы – достоверными. Научные положения, выводы и рекомендации, сформулированные в диссертации, являются в достаточной степени обоснованными и подтверждены математическими теоремами. Предложенные автором решения аргументированы и оценены по сравнению с другими известными решениями.

Диссертационная работа состоит из введения, пяти глав и заключения.

Во введении обоснована актуальность работы, определены цели и задачи исследования.

В первой главе введены основные понятия и обозначения. Сформулированы решаемые задачи, введено определение задачи декодирования временного ряда. Предложен итерационный алгоритм фильтрации признаков для многоиндексных матриц (тензоров), основанный на приближенном решении вещественной релаксации соответствующей задачи булева квадратичного программирования. Основным результатом главы является Теорема 3, из которой следует сходимость предложенного метода к точке локального оптимума релаксированной задачи со скоростью геометрической прогрессии.

Во второй главе автором предложено обобщение известного метода SNE, применяемого для снижения размерности фазового пространства при визуализации многомерных данных и основанного на максимально точно сохраняемом вероятностном отношении соседства между описываемыми объектами.

В третьей главе рассмотрена задача сегментирования квазипериодического временного ряда. Автором предложен алгоритм автоматического выделения «периодов», основанный на анализе фазовой траектории временного ряда. Основным результатом главы, по-видимому, является Теорема 4, в которой обосновывается аналитическое представление двух главных компонент сингулярного разложения траекторной матрицы исследуемого ряда.

В четвертой главе поставлена задача оценки достаточного объема выборки. Доказаны теоремы об асимптотическом распределении значений функции сходства между выборками фиксированного объема. На основании полученных результатов предложен эмпирический алгоритм оценки длины выборки, достаточной для классификации рядов с наперед заданным уровнем качества.

Пятая глава целиком посвящена описанию результатов экспериментальных исследований и сравнительного анализа методов, разработанных в предыдущих главах в приложении к задачам декодирования электрокортикограмм, классификации текстовых временных рядов и анализа биомедицинских временных рядов.

В диссертационной работе А.П. Мотренко получены следующие **основные результаты**.

1. Предложен итерационный алгоритм отбора признаков, учитывающий многоиндексную структуру корреляций между признаками.
2. Предложен алгоритм снижения размерности признакового пространства, сохраняющий распределение отношений ближайшего соседства между объектами.
3. Предложен алгоритм сегментирования квазипериодических временных рядов для идентификации интерпретируемых объектов классификации.
4. Предложен алгоритм оценки достаточного объема выборки, основанный на сравнении распределений параметров модели при заданном объеме выборки.
5. Проведены вычислительные эксперименты, подтверждающие адекватность предложенных в работе методов и успешную применимость их при решении актуальных прикладных задач.

В целом диссертация А.П. Мотренко является завершенной и целостной исследовательской работой, в которой получен ряд новых научных результатов.

Замечания

1. В главе 1 автором предложен новый приближенный алгоритм решения задачи выбора информативной подсистемы признаков, реализующий схему покоординатного спуска для вещественной релаксации соответствующей

тензорной квадратичной задачи булевой оптимизации, и обосновываются условия его сходимости к точке локального оптимума данной релаксации. Данный результат представляет несомненный научный интерес с точки зрения развития методов оптимизации. Тем не менее, вопросы теоретического или численного обоснования качества получаемого приближения в терминах исходной дискретной задачи выбора оптимальной подсистемы признаков в настоящий момент остались открытыми.

2. Использованная автором в методе сокращения размерности признакового пространства (в описанном в главе 2) методика кодирования меток классов, предполагающая их линейную упорядоченность и вполне уместная в задаче восстановления регрессии, представляется не вполне удачной в контексте рассматриваемой в главе задачи многоклассовой классификации.
3. Задача эффективной сегментации квазипериодических временных рядов — одна из наиболее хорошо известных и активно изучаемых задач анализа динамических данных. Представляет интерес проведение теоретического или численного сравнения предложенного в главе 3 метода с известными подходами в этой области, например, с последними результатами проф. А.В. Кельманова и коллег.

Высказанные замечания не влияют на высокую квалификационную оценку работы и могут рассматриваться как пожелания к продолжению исследований.

Результаты диссертации А.П. Мотренко своевременно и полно опубликованы в рецензируемых научных изданиях, в том числе в пятнадцати статьях, из которых девять - в изданиях, входящих в перечень ВАК РФ.

Результаты работы прошли апробацию на нескольких международных и всероссийских конференциях и семинарах и рекомендуются к использованию в МФТИ, ФИЦ ИУ РАН, ИМ СО РАН, ИММ УрО РАН при проектировании и обосновании алгоритмов анализа многомерных временных рядов, а также к включению в программы специальных курсов университетов по специальности прикладная математика и статистика.

Автореферат адекватно отражает содержание диссертационной работы.

Диссертация соответствует критериям, установленным действующим Положением о присуждении ученых степеней, обладает внутренним единством, содержит новые научные результаты и рекомендации по их практическому использованию.

Работа полностью соответствует требованиям ВАК РФ, предъявляемым к диссертациям на соискание ученой степени кандидата физико-математических наук по специальности 05.13.17 – Теоретические основы информатики – а ее автор, **А.П. Мотренко, заслуживает присуждения ей ученой степени кандидата наук по данной специальности.**

Официальный оппонент,
доктор физико-математических наук, профессор РАН,
заведующий отделом математического программирования
Института математики и механики им. Н.Н. Красовского Уральского отделения РАН
Тел. +7 (343) 375-35-05
e-mail: mkhachay@imm.uran.ru

Подпись заверяю
Ученый секретарь
ИММ УрО РАН

Хачай Михаил Юрьевич
20 августа 2019 г.
(Фигас Борис Вагнорович)