

На правах рукописи

Аботалей Мостафа

Аботалей Мостафа Салахелдин Абделсалам

**АЛГОРИТМЫ ПРОГНОЗИРОВАНИЯ ВРЕМЕННЫХ РЯДОВ
НА ОСНОВЕ ВЗВЕШЕННОГО МЕТОДА НАИМЕНЬШИХ
МОДУЛЕЙ**

1.2.3 — Теоретическая информатика, кибернетика

Автореферат

диссертации на соискание ученой степени
кандидата физико-математических наук

Челябинск — 2025

Работа выполнена на кафедре системного программирования
ФГАОУ ВО «Южно-Уральский государственный университет
(национальный исследовательский университет)»

**Научный
руководитель:**

МАКАРОВСКИХ Татьяна Анатольевна
доктор физ.-мат. наук, доцент,
профессор кафедры системного программирования
ФГАОУ ВО «Южно-Уральский государственный
университет (национальный исследовательский уни-
верситет)» (г. Челябинск)

**Официальные
оппоненты:**

ЕРОХИН Владимир Иванович
доктор физ.-мат. наук, профессор,
старший научный сотрудник отдела 51
ФГБВОУ ВО «Военно-космическая академия им.
А.Ф. Можайского» Министерства обороны Россий-
ской Федерации (г. Санкт-Петербург)

ЛАКМАН Ирина Александровна
кандидат техн. наук, доцент,
заведующий лабораторией исследования социально-
экономических проблем регионов
ФГБОУ ВО «Уфимский университет науки и техно-
логий» (г. Уфа)

**Ведущая организа-
ция:**

ФГБУН «Институт вычислительной математики и
математической геофизики Сибирского отделения
Российской академии наук» (г. Новосибирск)

Защита состоится 21 мая 2025 г. в 12:00 часов на заседании диссертационного совета
24.2.437.10 при Южно-Уральском государственном университете по адресу: 454080, г. Че-
лябинск, пр. Ленина, 76, ауд. 1001.

С диссертацией можно ознакомиться в библиотеке Южно-Уральского государственного
университета и на сайте [https://www.susu.ru/ru/dissertation/24243710-d-21229818/
abotaleb-mostafa-salaheldin-abdelsalam](https://www.susu.ru/ru/dissertation/24243710-d-21229818/abotaleb-mostafa-salaheldin-abdelsalam).

Автореферат разослан «___» _____ 2025 г.

Ученый секретарь дис. совета



Е.В. Иванова

Общая характеристика работы

Актуальность темы. На сегодняшний день накоплен значительный опыт в измерении вибрационных сигналов, разработке методов вибродиагностики и прогнозирования состояния и ресурса механических систем, а также в анализе больших объемов данных, описывающих различные промышленные, социальные и экономические процессы. Важным направлением исследований является повышение точности и скорости определения диагностических признаков, что особенно актуально для уникальных высоконагруженных механических систем. Решение этой задачи часто базируется на динамических характеристиках систем. Пусть состояние системы в момент времени t представлено вектором $\mathbf{x}(t) \in \mathbb{R}^n$, а диагностические признаки — вектором $\mathbf{y}(t) \in \mathbb{R}^m$. Ключевым аспектом является выбор диагностической математической модели $f : \mathcal{X} \rightarrow \mathcal{Y}$, устанавливающей связь между пространством состояний $\mathcal{X}$ и пространством диагностических признаков $\mathcal{Y}$. К таким моделям относятся динамические модели, регрессионные модели и модели машинного обучения, способные выявлять сложные зависимости в данных. Выбор модели зависит от характеристик анализируемого процесса и целей диагностики. Например, для нелинейных систем с высоким уровнем шума предпочтительны модели машинного обучения, тогда как для линейных систем подходят динамические и регрессионные модели. Таким образом, оптимальный выбор диагностической модели играет ключевую роль в повышении точности вибродиагностики механических систем, способствуя продлению их ресурса и снижению эксплуатационных рисков. Идентификация с использованием различных статистических методов, нейронных сетей (NN) или математических моделей (ММ) давно применяется в различных сферах. В настоящее время эти методы используются не только в промышленности, но и для прогнозирования развития социально значимых процессов, таких как пандемия COVID-19. Разработанное программное обеспечение для этих методов позволяет работать с временными рядами данных $(\{Y_t\})$, что делает такие системы применимыми для моделирования любых процессов, а не только эпидемиологических. Прогнозирование данных часто выполняется с использованием нейронных сетей, которые обеспечивают высокую точность для краткосрочных задач, таких как рыночные прогнозы. Однако их применение ограничено для процессов, описываемых дифференциальными уравнениями. Когнитивные модели показывают сопоставимую точность с ARIMA и успешно применяются для прогнозирования экономических параметров, таких как производство и логистика. Следует отметить, что все вышеуказанные модели в основном предназначены для краткосрочного прогнозирования. Следовательно, **актуальной**

задачей является разработка математического подхода, который позволит явно получать качественные квазилинейные разностные уравнения, адекватно описывающие рассматриваемый процесс, для среднесрочного и долгосрочного прогнозирования. Некоторые исследования в этой области включают модели, которые используют очистку данных, сглаживание и регрессионные модели для прогнозирования промышленного потребления электроэнергии. Однако, как и во многих других исследованиях, эти подходы остаются основанными на статистических методах.

Степень разработанности темы исследования

Первые исследования в области статистического анализа данных датированы 1927 г., а развитие методов прогнозирования имеет еще более длинную историю. В середине 1970-х гг. была популяризирована модель авторегрессионного интегрированного скользящего среднего (ARIMA) с использованием процедуры итеративного моделирования, состоящей из идентификации, оценки и проверки модели. Если модель ARIMA построена и признана адекватной, прогнозы будущих значений являются просто условными ожиданиями модели, если в качестве критерия используется минимальная среднеквадратическая ошибка. Тем не менее, большинство современных публикаций посвящено не столько развитию и улучшению прогностических свойств модели в общем случае, сколько ее использованию в ряде практических ситуаций (прогнозирование различных социальных явлений, динамики цен, природных явлений и прочих зависимостей).

Для анализа сложных закономерностей рассматриваются различные модификации моделей пространства состояний для экспоненциального сглаживания для обработки более широкого спектра сезонных закономерностей (модель BATS, описанная в 2011 г. Де Ливера А.М. и др.). Подбор параметров современных моделей осуществляется полностью автоматически. Это позволяет модели улавливать постепенно изменяющуюся сезонность путем введения в модель комбинаций из рядов Фурье. Одним из недостатков таких моделей является низкая скорость вычислений.

Важной движущей силой будущих исследований в области анализа временных рядов является прогресс в сборе больших объемов данных в различных сферах деятельности. Открытыми являются задачи сбора и хранения выборок большого объема, тяжелые хвосты, неравномерно распределенные наблюдения и комбинации многомерных дискретных и непрерывных переменных, возникновение т.н. «черного лебедя» (события с нулевыми дисперсией и математическим ожиданием) могут легко сделать существующие методы неадекватными. Таким образом, актуальной задачей является разработка новых моделей, использующих другие уравнения для управления эволюци-

ей условных моментов более низкого порядка. Решению данной задачи для практических целей (эпидемиология, финансы) посвящены, например, работы Elizar I., Gonzales M.R., Valeras A.S. и др.

В современных исследованиях и при моделировании сложных закономерностей зачастую использование одной модели не является эффективным, по этой причине разрабатываются различные гибридные модели, объединяющие в себе два и более методов. Наиболее популярны гибридные модели, в которых совмещаются только регрессионные подходы (один из которых – auto-ARIMA с автоматически определяемыми параметрами), либо регрессионные модели и нейросетевые методы. Наибольшую популярность здесь получили гибриды регрессионных моделей с моделями машинного обучения.

Отдельно хотелось бы отметить подход к анализу временных рядов, основанный на использовании квазилинейных рекуррентных соотношений (например, работы Панюкова А.В. и Мезаала Я.), развитию которого посвящена данная диссертационная работа. Существуют и другие методы, связанные с использованием детерминированных алгоритмов для подбора параметров модели (Тырсин А.Н., Петин Г.А., Зыкина А.В.). В отличие от нейронных сетей этот подход позволяет в явном виде получать качественные квазилинейные разностные уравнения (адекватно описывающие рассматриваемый процесс).

Цели и задачи исследования

Целью диссертационного исследования является разработка методов определения коэффициентов моделей для прогнозирования процессов, представленных в виде дискретных временных рядов. Для достижения указанной цели необходимо решить следующие задачи.

1. Разработать метод квазилинейного детерминированного анализа моделей нелинейной динамики для решения задачи прогнозирования развивающихся процессов. В качестве целевой функции для минимизации невязок использовать функцию, повышающую устойчивость модели к выбросам, что позволит получить более точные прогнозы.
2. Разработать систему прогнозирования временных рядов с возможностью подключения дополнительных модулей. Система должна давать возможность вывода значений погрешностей прогнозирования для дальнейшего выбора экспертом лучшей модели для использования.
3. Провести вычислительные эксперименты для рассмотренных в работе математических моделей с применением разработанной системы прогнозирования.

Методология и методы исследования. В работе развиваются методы регрессионного анализа динамики наблюдаемых процессов с приме-

нием статистических методов, методов построения нейронных сетей, глубокого обучения, линейного программирования, метода наименьших модулей, компьютерного моделирования.

Научная новизна

1. Разработан новый метод определения коэффициентов разностного квазилинейного уравнения по заданному временному ряду (по наблюдаемым отсчетам) для прогнозирования развивающихся процессов, основанный на методе наименьших модулей. Метод позволяет эффективно минимизировать значение функции потерь.
2. Спроектирован и реализован подход к выявлению закономерностей в данных и прогнозирования одномерного временного ряда, позволяющая получить результаты построения прогноза с использованием статистических, нейросетевых и квазилинейных моделей, проанализировать погрешности прогнозирования и выбрать лучшую модель для работы с конкретными исходными данными.

Теоретическая значимость результатов обусловлена решением актуальных задач обнаружения закономерностей развивающихся процессов с применением современного математического аппарата. При этом возможность учета нелинейности позволяет использовать полученные модели в различных областях.

Практическая ценность заключается в том, что разработанные Алгоритмы реализованы в виде программного обеспечения для прогнозирования временных рядов различными методами и проведения сравнительного анализа результатов работы реализованных алгоритмов и выбора лучших алгоритмов для обнаружения закономерностей в заданном наборе данных. Результаты, изложенные в диссертации, могут быть использованы в прогнозировании социальных, экономических, природных явлений, в медицине, промышленности и пр. Разработанный метод квазилинейного анализа одномерного временного ряда применен для исследования ряда прикладных задач (прогнозирование урожайности, распространение эпидемии, анализ температуры воздуха, прогнозирование скорости ветра).

Положения, выносимые на защиту

1. Предложен метод обобщенных наименьших отклонений (GLDM) для анализа моделей нелинейной динамики при прогнозировании развивающихся процессов. В данном методе функция арктангенса $\arctan(x)$ используется в качестве функции потерь, что повышает устойчивость модели к выбросам и обеспечивает более точные прогнозы.

2. Разработана система прогнозирования временных рядов с возможностью подключения дополнительных модулей. Система дает возможность вывода значений погрешностей прогнозирования для дальнейшего выбора экспертом лучшей модели для использования.
3. Проведены вычислительные эксперименты на реальных данных, подтвердившие эффективность разработанных методов и подходов.

Степень достоверности результатов. Достоверность научных результатов и выводов исследования определяется корректным использованием современных математических методов, подтверждены доказательствами в соответствии с современным уровнем математической строгости, объемом апробации и представления этапов работы на внутривузовских, всероссийских и международных научных конференциях и семинарах. Результаты и выводы не противоречат ранее полученным результатам других авторов.

Апробация результатов. Основные положения диссертационной работы, разработанные методы, алгоритмы и результаты вычислительных экспериментов докладывались на следующих международных и всероссийских научных конференциях.

1. The 7th International conference on Time-series and Forecasting (19-21 июля 2021 г., Испания, г. Гран-Канария).
2. Computer Science Protecting Human Society Against Epidemics. ANTICOVID 2021. IFIP Advances in Information and Communication Technology (28-29 июня 2021 г., Польша, г. Варшава).
3. International Conference on Information Technology and Nanotechnology (ITNT) (20-24 сентября 2021 г., Самара).
4. International Conference on Information Technology and Nanotechnology (ITNT) (23-27 мая 2022 г., Самара).
5. Mathematical Optimization Theory and Operations Research: Recent Trends. MOTOR (05-10 июля 2021 г., Иркутск).
6. Mathematical Optimization Theory and Operations Research: Recent Trends. MOTOR (02-06 июля 2022 г., Петрозаводск).
7. Mathematical Optimization Theory and Operations Research: Recent Trends. MOTOR (02-08 июля 2023 г., Екатеринбург).
8. Mathematical Optimization Theory and Operations Research: Recent Trends. MOTOR (30 июня-06 июля 2024 г., Омск).

9. Параллельные вычислительные технологии (ПаВТ) (29-31 марта 2022 г., Дубна).
10. Параллельные вычислительные технологии (ПаВТ) (28-30 марта 2023 г., Санкт-Петербург).
11. XIV Международная конференция «Оптимизация и приложения» (ОРТИМА), (26-30 сентября 2022 г., Черногория, г. Петровац).
12. XV Международная конференция «Оптимизация и приложения» (ОРТИМА), (18-22 сентября 2023, Черногория, г. Петровац).
13. 4th International Scientific Conference of Alkafeel University (ISCKU 2022), (20-21 декабря 2022 г., Ирак, г. Эн-Наджаф).
14. Цифровая индустрия: состояние и перспективы развития 2023 (ЦИ-СП'2023), (21-23 ноября 2023 г., Челябинск).
15. 8th International Conference on Emerging Applications of Information Technology (EAIT 2024) (12-13 января, 2024 г., Индия, г. Калькутта).
16. The 3rd IEEE International Conference on Electronic Engineering (ICEEM2023) (07-08 октября 2023 г., Египет, г. Минуф).
17. International Conference on Trends in Quantum Computing and Emerging Business Technologies (TQCEBT 2024) (22-23 марта 2024 г., Индия, г. Пуна).

Публикации. Основные результаты диссертации опубликованы в 7 научных работах, в том числе 3 статьи в журналах, входящих в категорию K2 перечня ВАК [1–3], 2 статьи в журналах, входящих в квартиль Q1 наукометрической базы Scopus [4, 7], и 1 статья в журнале, входящем в квартиль Q2 наукометрической базы Scopus [5]. Получено 5 свидетельств Роспатента о государственной регистрации программ для ЭВМ [8–12].

Личный вклад. Содержание диссертации и основные положения, выносимые на защиту, отражают личный вклад автора в опубликованные работы. Работы [1, 4, 5, 6] выполнены автором самостоятельно. В работах [2, 7–12] постановка задачи выполнена научным руководителем Макаровских Т.А.; все полученные результаты принадлежат М.С.А. Аботалёб. В работе [3] Панюкову А.В. принадлежит обзор полученных ранее результатов, Макаровских Т.А. – постановка задачи, все остальные результаты принадлежат М.С.А. Аботалёб.

Соответствие диссертации паспорту научной специальности.

Тематика, содержание и полученные результаты диссертационного исследования М.С.А. Аботалеба соответствуют следующему направлению исследований паспорта специальности 1.2.3. «Теоретическая информатика, кибернетика»: п. 9. Математическая теория исследования операций.

Структура и объем работы. Диссертация состоит из введения, 4 глав, заключения и библиографии. Общий объем диссертации 182 страниц, из них 160 страниц основного текста, 22 страницы библиография.

Содержание работы

Во введении приводится обоснование актуальности темы и степень ее разработанности; формулируются цели и задачи; раскрывается новизна, теоретическая и практическая значимость полученных результатов; формулируется методологическая основа диссертационного исследования; дается обзор содержания диссертации.

В первой главе приводится анализ существующих методов прогнозирования, которые можно подразделить на статистические, нейросетевые, а также методы, использующие дифференциальные и разностные уравнения.

Показано, что большинство из известных в настоящее время методов применимы для краткосрочного прогнозирования, поэтому актуальна задача разработки математического подхода, позволяющего в явном виде получать качественные квазилинейные разностные уравнения (адекватно описывающие рассматриваемый процесс). Кроме того, в настоящее время активно исследуются нейросетевые модели и модели глубокого обучения, обладающие одним существенным недостатком, связанным с невозможностью точно описать параметры процесса.

Несмотря на обилие методов прогнозирования временных рядов, среди них не существует универсального наилучшего метода прогнозирования. В течение долгого времени остается актуальным вопрос выбора наиболее подходящего метода прогнозирования имеющегося временного ряда. Данная задача не является тривиальной, поскольку выбор метода зачастую зависит от структуры временного ряда, а также самого процесса, частоты получения новых данных, объема выборки, горизонта прогнозирования и пр.

Поэтому создание единой гибкой системы прогнозирования, осуществляющей автоматический подбор метода прогнозирования является важной задачей. Существующее программное обеспечение позволяет использовать последовательно каждый из методов и не позволяет автоматически получить сводный анализ примененных методов. Также в систему прогнозирования представляется разумным добавить метод определения параметров квазили-

нейного рекуррентного уравнения. Поскольку в настоящее время не разработано универсального средства для построения прогнозов, предлагается разработать расширяемую систему прогнозирования, которая способна обрабатывать любые временные ряды произвольной длины и не зависит от характера прикладных задач.

В качестве прототипа взята модель, рассмотренная в работах Панюкова А.В. и Межаала Я.А., в которой рассматривается алгоритм определения коэффициентов $a_1, a_2, a_3, \dots, a_m \in \mathbb{R}$ квазилинейной авторегрессионной модели m -го порядка:

$$y_t = \sum_{j=1}^{n(m)} a_j g_j(\{y_{t-k}\}_{k=1}^m) + \varepsilon_t, \quad t = 1, 2, \dots, T, \quad (1)$$

$$\{a_j^*\}_{j=1}^{n(m)} = \arg \min_{\{a_j\}_{j=1}^{n(m)} \in \mathbb{R}} \left[\sum_{t=1}^T \arctan \left| \sum_{j=1}^{n(m)} a_j g_j(\{y_{t-k}\}_{k=1}^m) - y_t \right| \right]. \quad (2)$$

с актуальной информацией о значениях переменных состояния $\{y_t \in \mathbb{R}\}_{t=1-m}^T$ в моменты времени t , где $g_j : (\{y_{t-k}\}_{k=1}^m) \rightarrow \mathbb{R}$, $j=1, 2, \dots, n$, $n = 2m + C_m^2 = m(m+3)/2$, функции n заданы, и $\{\varepsilon_t \in \mathbb{R}\}_{t=1}^T$ — неизвестные ошибки.

Во второй главе рассмотрен алгоритм определения и оценки параметров квазилинейного рекуррентного уравнения. Оригинальность данного подхода заключается в том, что он вместо метода наименьших квадратов использует обобщенный метод наименьших модулей с подбором весов. Задача (2), т.е. задача GLDM-оценки, является задачей многоэкстремальной оптимизации. GLDM-оценки устойчивы к наличию корреляции значений в $\{y_t \in \mathbb{R}\}_{t=1-m}^T$ и (при соответствующих настройках) лучше всего подходят для вероятностных распределений ошибок с более тяжелыми (чем у нормального распределения) хвостами. Блок-схема рассматриваемого алгоритма приведена на рис. 1.

Оригинальность рассмотренного алгоритма определения параметров квазилинейного рекуррентного уравнения заключается в том, что он вместо метода наименьших квадратов использует обобщенный метод наименьших модулей с подбором весов. Показано, что задача GLDM-оценки, является задачей многоэкстремальной оптимизации, и GLDM-оценки устойчивы к наличию корреляции значений в $\{y_t \in \mathbb{R}\}_{t=1-m}^T$ и (при соответствующих настройках) лучше всего подходят для вероятностных распределений ошибок с более тяжелыми (чем у нормального распределения) хвостами.

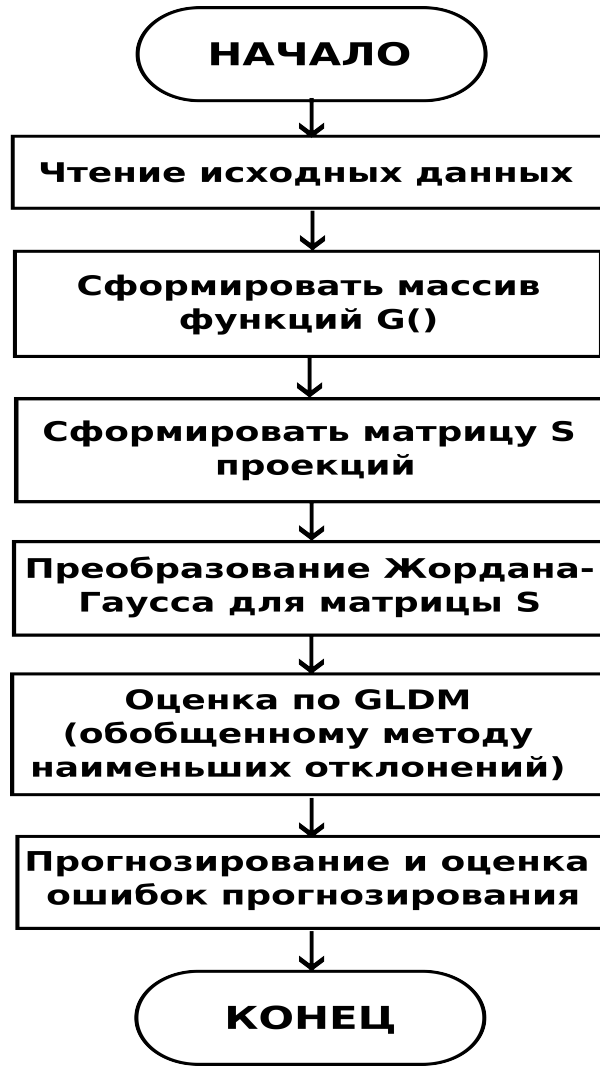


Рис. 1: Схема реализации модели

Алгоритм 1 **DualWLDMSolver** используется для решения задачи двойственной оптимизации. Этот алгоритм применяет метод проекционного градиента для нахождения оптимального решения на пересечении линейного подпространства и параллелепипеда, что делает его подходящим для выпуклых задач оптимизации.

Шаги работы алгоритма:

Шаг 1 (Инициализация): Алгоритм начинается с нулевой начальной точки $w_0 = 0$. - Вычисляется проекция градиента g на подпространство $\mathcal{L}$, что определяется через матрицу проекции $S = E - S^T(SS^T)^{-1}S$, где S — это матрица, задающая линейные ограничения. Проекция градиента вычисляется как $g = S \cdot \nabla$, где $\nabla = \{y_t\}_{t=1}^T$ — градиент целевой функции.

Шаг 2 (Итеративное обновление): На каждой итерации алгоритм вычисляет максимальный шаг α_* , при котором переменные w_t остаются в пределах допустимого диапазона $[-p_t, p_t]$ для всех t . - Если какая-либо коор-

дината переменной w_t достигает границы параллелепипеда, она фиксируется, и дальнейшее движение по этому направлению становится невозможным. Это означает, что градиент в этом направлении становится равным нулю.

Шаг 3 (Остановка): Алгоритм останавливается, когда градиент больше не может улучшить решение, то есть когда все направления движения исчерпаны и проекция градиента на допустимую область становится нулевой.

Таким образом, алгоритм **DualWLDMSolver** находит оптимальное решение задачи двойственной оптимизации с учетом всех линейных ограничений и ограничений параллелепипеда. Благодаря применению проекций и проверке ограничений на каждом шаге, алгоритм гарантирует нахождение решения с гарантированной сходимостью при выполнении условий выпуклости задачи.

Algorithm 1 . DualWLDMSolver

Require: :

∇_L

▷ Проекция градиента

$\{p_t \in \mathbb{R}^+\}_{t=1}^T$

▷ Весовые коэффициенты

Ensure: :

$w^* = \arg \max_{w \in \mathbb{R}^T} \sum_{i=1}^T w_i \cdot y_i$

▷ Оптимальное решение двойственной задачи

$R^* = \{t \in T : |w_t^*| = p_t\}$

▷ Несколько активных ограничений

1: $w \leftarrow \{w_i = 0 : i = 1, 2, \dots, T\}; R \leftarrow \emptyset; g = \nabla_L$

▷ Начальная точка

2: **while** $(\alpha_* \neq 0)$ **do**

3: $\{(\alpha_*, t_*) \leftarrow \arg \max_{\substack{\alpha \in \mathbb{R}^+ \\ t \notin R}} \{\alpha \geq 0 : -p_t \leq w_t + \alpha g_t \leq p_t\}\}$

▷ Параметры шага

4: $w \leftarrow w + \alpha_* g; g_{t_*} \leftarrow 0; R := R \cup \{t_*\};$

▷ Переход к следующей точке

5: **end while**

6: $w^* = w, R^* = R$

return (w^*, R^*)

Алгоритм 2 **WLDM-estimator** используется для нахождения оптимального решения прямой задачи по методу взвешенного линейного девиационного метода (WLDM). На вход алгоритма подаются такие данные, как проекция градиента, матрица линейного подпространства, веса и значения временных рядов. На выходе алгоритм возвращает оптимальные коэффициенты и значения невязок. Принцип работы алгоритма можно описать в виде следующих шагов:

Шаг 1 (Инициализация): Алгоритм получает вектор значений временного ряда $y = \{y_t \in \mathbb{R}^+\}_{t=1-m}^T$ и весовые коэффициенты $\{p_t \in \mathbb{R}^+\}_{t=1}^T$, которые отражают значимость каждой точки данных во временном ряду. - Матрица $S = \{S_t \in \mathbb{R}^N\}_{t \in T}$ представляет собой линейное подпространство $\mathcal{L}$, на которое проецируется градиент $\nabla_{\mathcal{L}}$.

Шаг 2 (Использование двойственного решения): Для нахождения двойственного решения применяется вспомогательный алгоритм *DualWLDMSolver*, который использует проекцию градиента ∇ и веса p_t . Результатом работы этого алгоритма являются оптимальные значения двойственного решения w^* и набор активных ограничений R^* . Эти активные ограничения определяют, какие временные точки соответствуют границам допустимого множества.

Шаг 3 (Формирование новой системы): После получения двойственного решения, пересчитываются матрица S^* и вектор y^* для временных шагов, которые не попали под активные ограничения R^* . Это позволяет создать новую систему уравнений, которая уже не зависит от влияния ограничений и лучше поддается дальнейшей оптимизации.

Шаг 4 (Решение системы прямой задачи): Решение прямой задачи осуществляется с помощью решения системы линейных уравнений, где пересчитанная матрица S^* инвертируется. Оптимальные коэффициенты A^* для прямой задачи вычисляются следующим образом:

$$(A^*)^T = y^* \cdot (S^*)^{-1}$$

Эти коэффициенты минимизируют разницу между прогнозируемыми и фактическими значениями временного ряда.

Шаг 5 (Вычисление невязок): Невязки z^* определяются как разница между прогнозируемыми значениями, вычисленными с использованием коэффициентов A^* , и фактическими значениями временного ряда:

$$z^* = (A^*)^T S - y$$

Эти невязки отражают степень отклонения прогнозируемых значений от реальных данных, что позволяет оценить точность модели.

Шаг 6 (Вывод результатов): На выходе алгоритм возвращает оптимальные коэффициенты A^* и невязки z^* . Эти значения могут быть использованы для анализа качества модели и ее прогностических возможностей.

Таким образом, алгоритм **WLDM-estimator** предоставляет точные и оптимальные решения для прогнозирования временных рядов, основываясь на решении двойственной задачи и минимизации невязок в системе линейных уравнений. Этот подход гарантирует высокую точность результатов и эффективность при работе с данными.

Algorithm 2 . WLDМ-estimator

Require: :

$$S = \{S_t \in \mathbb{R}^N\}_{t \in T}$$

▷ Матрица линейного подпространства $\mathcal{L}$

$$\nabla_{\mathcal{L}}$$

▷ Проекция градиента на $\mathcal{L}$

$$\{p_t \in \mathbb{R}^+\}_{t=1}^T$$

▷ Веса

$$\{y_t \in \mathbb{R}^+\}_{t=1-m}^T$$

▷ Значения переменных

Ensure: :

$$A^* \in \mathbb{R}^n$$

▷ Оптимальное решение прямой задачи

$$z^* \in \mathbb{R}^T$$

▷ Ограничения

$$1: (w^*, R^*) \leftarrow \text{DualWLDMSolver} \left(\nabla, \{p_t \in \mathbb{R}^+\}_{t=1}^T \right)$$

$$2: S^* \leftarrow \{S_t : t \notin R^*\}; y^* \leftarrow \{y_t : t \notin R^*\}$$

▷ Матрица системы

$$3: (A^*)^\top \leftarrow y^* \cdot (S^*)^{-1}$$

▷ Решение системы

$$4: z^* \leftarrow (A^*)^\top S - y$$

▷ Невязки

$$5: \text{Output: } (A^*, z^*)$$

Установленные результаты позволяют свести задачу определения оценки GLDM к итерационной процедуре с оценками WLDМ (см. алгоритм 3).

Алгоритм 3 **GLDM-estimator** используется для вычисления оптимального решения обобщенного метода наименьших отклонений (GLDM). Этот алгоритм работает итеративно, с пересчетом весов и остатков на каждом шаге до тех пор, пока решение не стабилизируется.

Шаги работы алгоритма:

Шаг 1 (Инициализация весов): Алгоритм инициализирует все весовые коэффициенты как равные 1 для всех t :

$$p \leftarrow \{p_t = 1 : t = 1, 2, \dots, T\}$$

Затем вычисляется начальное приближение решения с помощью вспомогательной функции **WLDMSolver**, которая принимает в качестве входных данных матрицу S , проекцию градиента ∇ , веса p_t и значения переменных состояния y_t . Это возвращает начальные коэффициенты $A^{(1)}$ и остатки $z^{(1)}$.

Шаг 2(Обновление весов): После получения начальных остатков $z^{(1)}$, весовые коэффициенты пересчитываются для каждого временного шага по следующей формуле:

$$p_t \leftarrow \left(1 / \left(1 + (z_t^{(1)})^2 \right) \right)$$

для всех $t = 1, 2, \dots, T$. Это обновление позволяет учесть величину ошибок и адаптировать веса в зависимости от величины остатка.

Шаг 3 (Повторное решение с обновленными весами): После пересчета весов алгоритм снова вызывает **WLDMSolver**, чтобы вычислить новые коэффициенты $A^{(2)}$ и остатки $z^{(2)}$. - Счетчик итераций k устанавливается равным 2.

Шаг 4 (Итеративный процесс): Алгоритм выполняет итерации до тех пор, пока решения на двух последовательных шагах не станут одинаковыми, т.е. $A^{(k)} = A^{(k-1)}$. На каждой итерации весовые коэффициенты обновляются на основе текущих остатков:

$$p_t^{(k)} \leftarrow \left(1 / \left(1 + (z_t^{(k)})^2 \right) \right)$$

Алгоритм снова вызывает **WLDMSolver**, чтобы вычислить новые значения $A^{(k+1)}$ и $z^{(k+1)}$, и увеличивает счетчик k на единицу.

Шаг 5 (Завершение): Алгоритм завершает работу, когда на двух последовательных шагах изменения в решении становятся незначительными (т.е. $A^{(k)} = A^{(k-1)}$). Окончательные коэффициенты A^* и остатки z^* принимаются равными значениям, вычисленным на последней итерации: $A^* = A^{(k)}$, $z^* = z^{(k)}$.

Шаг 6 (Вывод результатов): На последнем шаге алгоритм возвращает оптимальные коэффициенты A^* и остатки z^* , которые могут быть использованы для анализа и оценки точности модели.

Таким образом, алгоритм **GLDM-estimator** проводит итеративную процедуру оптимизации, где на каждом шаге пересчитываются веса и остатки для достижения более точного решения. Итерации продолжаются до тех пор, пока изменения в решении не стабилизируются.

Теорема 1. *Последовательность $\{(A^{(k)}, z^{(k)})\}_{k=1}^{\infty}$, построенная по алгоритму GLDM-оценки сходится к глобальному минимуму (a^*, z^*) задачи (2).*

Изложенное в данной главе показывает возможность решения задачи определения параметров уравнения с использованием алгоритма **GLDM-оценка**. Алгоритм показывает, что его вычислительная сложность пропорциональна вычислительной сложности алгоритма решения прямой и/или двойственной задачи WLDМ.

Наличие нулевых коэффициентов в моделях более высокого порядка свидетельствует об эффективности метода исключения факторов, не влияющих на прогнозирование, что упрощает модели и повышает их интерпретируемость. Это исследование подтверждает ключевую роль выбора оптимального порядка построения модели и интерпретации коэффициентов модели в контексте прогнозирования временных рядов.

Algorithm 3 . GLDM-estimator

Require:

$$S = \{S_t \in \mathbb{R}^N\}_{t \in T}$$

 $\triangleright$ Матрица линейного подпространства $\mathcal{L}$

$$\nabla_L$$

 $\triangleright$ Проекция градиента на $\mathcal{L}$

$$\{p_t \in \mathbb{R}^+\}_{t=1}^T$$

 $\triangleright$ Весовые коэффициенты

$$\{y_t \in \mathbb{R}^+\}_{t=1-m}^T$$

 $\triangleright$ Значения заданных переменных состояния**Ensure:**

$$A^* \in \mathbb{R}^n$$

 $\triangleright$ Оптимальное решение GLDM

$$z^* \in \mathbb{R}^T$$

 $\triangleright$ Остатки

```
1:  $p \leftarrow \{p_t = 1 : t = 1, 2, \dots, T\}$ 
2:  $(A^{(1)}, z^{(1)}) \leftarrow \mathbf{WLDMSolver}(S, \nabla_{\mathcal{L}}, \{p_t\}_{t=1}^T, \{y_t\}_{t=1-m}^T);$ 
3: for all  $t = 1, 2, \dots, T$  do
4:    $p_t \leftarrow \left(1 / \left(1 + (z_t^{(1)})^2\right)\right)$ 
5: end for
6:  $(A^{(2)}, z^{(2)}) \leftarrow \mathbf{WLDMSolver}(S, \nabla_{\mathcal{L}}, \{p_t\}_{t=1}^T, \{y_t\}_{t=1-m}^T);$ 
7:  $k \leftarrow 2$ 
8: while  $(A^{(k)} \neq A^{(k-1)})$  do
9:   for all  $t = 1, 2, \dots, T$  do
10:     $p_t^{(k)} \leftarrow \left(1 / \left(1 + (z_t^{(k)})^2\right)\right)$ 
11:   end for
12:    $(A^{(k+1)}, z^{(k+1)}) \leftarrow \mathbf{WLDMSolver}(S, \nabla_{\mathcal{L}}, \{p_t^{(k)}\}_{t=1}^T, \{y_t\}_{t=1-m}^T);$ 
13:    $k \leftarrow (k + 1)$ 
14: end while
15:  $z^* \leftarrow z^{(k)}, (A^*) \leftarrow A^{(k)}$ 
return  $(A^*, z^*)$ 
```

$\triangleright$ Поиск ограничений

Проведенные вычислительные эксперименты показывают, что среднее число итераций алгоритма **GLDM-оценка** равно количеству коэффициентов в идентифицированном уравнении. Вычислительная сложность решения практических задач не превышает $O((n(m))^3 T + n(m) \cdot T^2)$ в зависимости от m . таким образом, алгоритм позволяет получить решение задачи за сравнительно короткий срок.

Показано, что в контексте прогнозирования временных рядов с использованием моделей более высокого порядка часто наблюдается, что ряд коэффициентов может быть равен нулю. Данную ситуацию можно объяснить несколькими математическими и статистическими причинами, которые имеют решающее значение при оптимизации точности прогнозирования модели и эффективности вычислений. Обнуление коэффициента можно обосновать математически путем анализа градиента целевой функции по отношению к каждому коэффициенту в контексте обобщенного метода наименьших отклонений (GLDM).

Выбор оптимального порядка построения модели при прогнозировании временных рядов с использованием обобщенного метода наименьших отклонений (GLDM) имеет решающее значение для обеспечения баланса между

сложностью модели и ее способностью обобщаться для прогнозирования новых значений. Приводятся рассуждения, которые описывают методологию определения оптимального порядка m модели путем минимизации функции потерь. Описан алгоритм выявления значимых коэффициентов модели. Он может быть особенно полезен в тех случаях, когда решающее значение имеет баланс между сложностью и функциональной совместимостью модели и точностью, что позволяет создать оптимизированную модель, которая эффективно отражает существенную динамику данных.

Теорема 2 [Выбор оптимальной модели]. Пусть $\{y_t\}_{t=1}^T$ – временной ряд, и функция потерь $F(\mathbf{a})$ определена как:

$$F(\mathbf{a}) = \sum_{t=1}^T \arctan |y_t - f(\mathbf{a}, \mathbf{x}_t)|,$$

где функция f зависит от параметров $\mathbf{a}$. Тогда оптимальный порядок модели m^* равен:

$$m^* = \arg \min_m \left(\min_{\mathbf{a}} F(\mathbf{a}, m) \right).$$

Условия, при которых коэффициенты a_j модели могут быть обнулены без потери точности прогнозирования, приведены в следующей теореме. Если производная $F(\mathbf{a})$ по коэффициенту a_j равна нулю на всем интервале, это указывает на то, что изменения a_j не влияют на значение функции потерь, и коэффициент a_j может быть исключен из модели.

Теорема 3 [Условия наличия нулевых коэффициентов модели]. Если для коэффициента a_j :

$$\frac{\partial F(\mathbf{a})}{\partial a_j} = \sum_{t=1}^T \frac{\partial}{\partial a_j} \arctan |y_t - f(\mathbf{a}, \mathbf{x}_t)| = 0,$$

то a_j не влияет на функцию потерь $F(\mathbf{a})$ и может быть обнулен.

Приведенные теоремы описывают основные аспекты обобщенного метода наименьшего отклонения (GLDM) для анализа данных временных рядов. Они определяют метод определения эффективной структуры модели и определения избыточных коэффициентов модели.

В третьей главе рассмотрено созданное программное обеспечение, позволяющее построить прогноз для выбранного пользователем периода с использованием любой из рассмотренных выше моделей. Для достижения этой цели было разработано пять программных модуля.

1. Система прогнозирования **Epidemic.TA**, позволяющая построить прогнозы для одномерного временного ряда с использованием регрессионных моделей, и модели **SIR**. В состав первой версии системы входила функция, использующая модель **SIR**, но в более поздних версиях данная функция была исключена, поскольку она имеет низкую прогностическую способность. Вместо **SIR** используется модель на основе нейронной сети **NNAR**.
2. Система прогнозирования **DL-system**, использующий методы глубокого обучения для построения прогнозов для одномерного временного ряда.
3. Модуль **GLDM**, использующий методы определения параметров квазиди-нейного рекуррентного соотношения
4. Для работы с длинными временными рядами, требующими матриц большого порядка, в программном обеспечении реализована возможность использования двух видов параллелизма: функционального и параллелизма по данным. Реализация алгоритма прогнозирования поддерживает многопоточные вычисления и была протестирована на различных вычислительных платформах для повышения производительности.
5. Программное обеспечение для одномерных временных рядов Версия 1, предназначенное для анализа и прогнозирования временных рядов с использованием различных моделей, таких как классические статистические методы (**Auto ARIMA**, экспоненциальное сглаживание, **Prophet**), гибридные модели (**Auto ARIMA** + экспоненциальное сглаживание и полиномиальная регрессия), методы машинного обучения (**SVM**, многослойный перцептрон) и обобщенный метод наименьшего отклонения (**GLDM**) с различными порядками зависимостей.

Все модули реализованы на языках **R** (модуль **updated Epidemic.TA**), **Python** (модуль **DL-system**). Модуль **GLDM** – на языках **Python** и **C++**, кроме того, на языке **C++** имеется параллельная версия алгоритма, реализованная с использованием библиотеки **OpenMP**. Программное обеспечение для одномерных временных рядов Версия 1 – на языках **Python** и **R**.

Разработанное программное обеспечение предназначено для анализа и прогнозирования одномерных временных рядов с использованием различных моделей: от классических статистических методов до современных методов машинного обучения и обобщенного метода наименьшего отклонения (**GLDM**). Оно обеспечивает полный цикл анализа временных рядов, от загрузки данных до получения прогнозов. На рисунке 2 представлена схема процесса обработки данных, выбора моделей и оценки точности прогнозов.

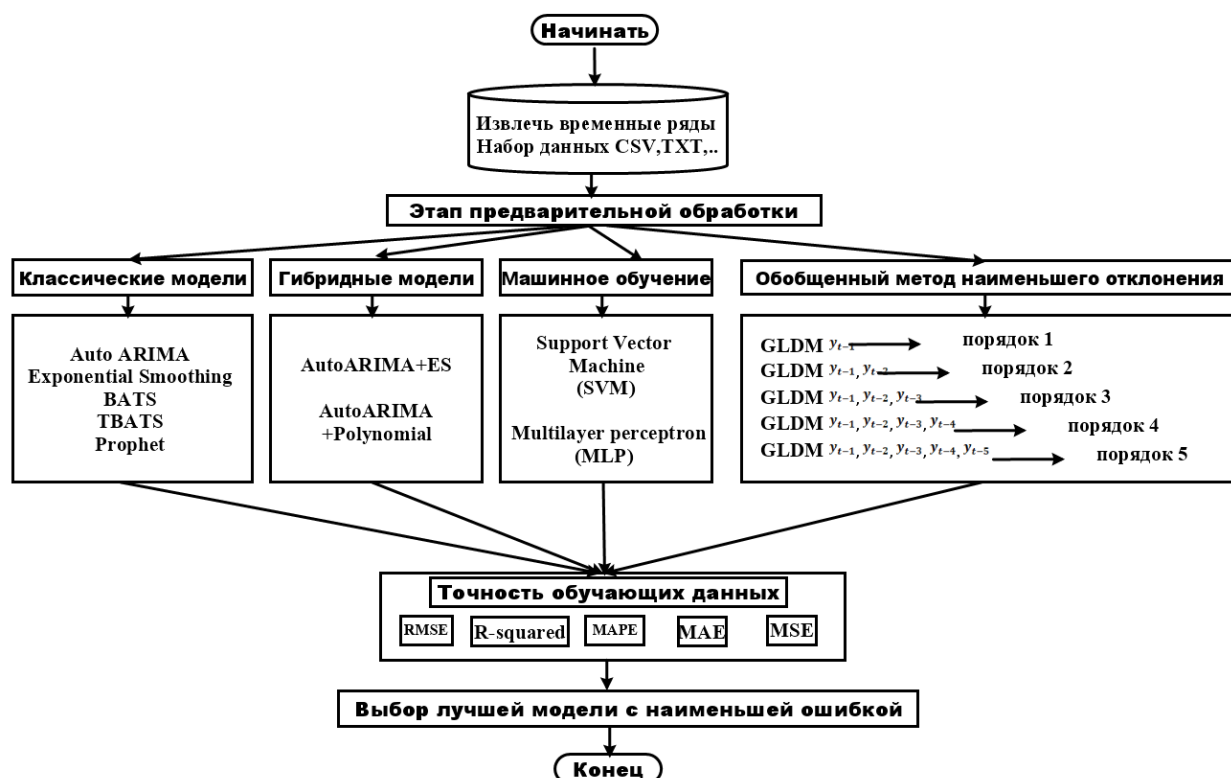


Рис. 2: Схема системы прогнозирования для одномерных временных рядов.

Показано, что последовательная версия алгоритма определения параметров квазилинейного рекуррентного уравнения с использованием персонального компьютера работает для временных рядов порядка 10^3 элементов и требует несколько минут времени. При попытке реализации функционального параллелизма выяснилось, что данный вид параллелизма для данного алгоритма не является эффективным. При задействовании более мощных вычислителей, ускорение отсутствует. С другой стороны, возможен достаточно эффективный параллелизм по данным, что позволяет работать с большим количеством временных рядов даже на персональном компьютере.

Поскольку для запуска разработанного алгоритма необходимо использовать матрицы высокого порядка, полезно распараллелить алгоритм с целью ускорения процесса получения результата. Проведенные вычислительные эксперименты показывают, что при простейших способах распараллеливания рассмотренных алгоритмов имеем ускорение алгоритма в три раза на современных персональных компьютерах при использовании половины из возможных потоков. Тестирование алгоритма на более мощных вычислителях показало не только отсутствие эффекта от распараллеливания для временных рядов длиной до 100 000 значений, но и значительное замедление вычислений за счет накладных расходов на организацию параллельных областей в каждой из функций. Таким образом, для реализации алгоритма

определения параметров одного квазилинейного рекуррентного соотношения нецелесообразно использовать возможности параллельных вычислений.

В четвертой главе на основе анализа временных рядов с данными из различных областей с помощью обобщенного метода наименьших отклонений (GLDM) был определен оптимальный порядок построения моделей для прогнозирования различных развивающихся процессов (таблица 1).

Таблица 1: Перечень используемых наборов данных и их длина n

№	Набор данных	Длина ряда n
1	NDVI (значение вегетационного индекса для полей с пшеницей)	15
2	Температура (климатические данные рядом с университетом Монаш, Клейтон, Виктория, Австралия)	9 999
3	Скорость ветра для SCADA (рабочие данные высокой точности, полученные от ветряной турбины, активно вырабатывающей электроэнергию в Турции)	50 530
4	Случаи смерти COVID-19 в РФ (ежедневные случаи смерти от Covid-19 в Российской Федерации с 25 марта 2020 г. по 23 августа 2022 г.)	882
5	Случаи заражения COVID-19 в Самарской области (Ежедневные случаи заражения COVID-19 в Самарской области с 16 апреля 2020 г по 13 января 2023 г)	1 003

Полученные результаты вычислительных экспериментов показывают, что порядок прогнозирующей модели, в значительной степени зависит от характеристик набора данных (характер данных, лежащая в их основе динамика и наличие нелинейных закономерностей, количество доступных данных).

Результаты проведенных вычислительных экспериментов показали, что в большинстве случаев для воссоздания уравнения, описывающего динамику временного ряда, достаточно использовать модель второго порядка. Тем не менее, для рядов, характеризующихся значительными сезонными колебаниями и автокорреляциями (например, данные о десятилетних колебаниях температуры) модели более высоких порядков дадут более точные результаты.

В заключении в краткой форме излагаются итоги выполненного диссертационного исследования; представляются отличия данной работы от ранее выполненных родственных работ других авторов; формулируются основные результаты диссертационной работы, выносимые на защиту; рассматриваются направления дальнейших исследований в данной области.

Заключение

В диссертации разработаны и исследованы новые методы и алгоритмы определения коэффициентов квазилинейного рекуррентного уравнения. В настоящее время эффективные детерминированные методы для анализа ди-

намических процессов, представленных временными рядами, является актуальной проблемой, решение которой востребовано при проведении прикладных исследований в различных областях: от промышленности, до медицины и социально-экономических приложений.

Оригинальность алгоритма определения параметров квазилинейного рекуррентного уравнения заключается в том, что он вместо метода наименьших квадратов использует обобщенный метод наименьших модулей с подбором весов.

В результате исследования разработана система прогнозирования, которая обладает следующими свойствами: (1) способна обрабатывать любые временные ряды произвольной длины и не зависит от характера прикладных задач; (2) автоматически настраивает параметры используемых моделей; (3) расширяема, т.е. допускает подключение дополнительных модулей для прогнозирования; (4) позволяет выбрать критерий отбора «лучшей» модели. В данную систему включен модуль, реализующий метод определения параметров квазилинейного рекуррентного уравнения.

Говоря о качестве самой модели на основе метода наименьших отклонений, можно отметить, что она работает не хуже нейросетевых моделей или классических статистических моделей, что подтверждается и проведенными экспериментами, в которых были сравнены результаты работы рассмотренной модели и результаты, полученные с помощью известных моделей. У модели, использующей обобщенный метод наименьших отклонений (GLDM) есть одно существенное преимущество по сравнению с этими моделями, заключающееся в возможности интерпретировать коэффициенты модели с точки зрения задачи исследования.

Результаты исследования могут использоваться для моделирования динамически развивающихся процессов в широком спектре областей, поскольку разработанный алгоритм определения коэффициентов квазилинейного уравнения и система прогнозирования, его включающая, не зависят от специфики предметной области.

Основные результаты, полученные в ходе выполнения диссертационного исследования, являются новыми и не покрываются работами других авторов, обзор которых был приведен в главе 1. Основные отличия заключаются в следующем.

1. Алгоритм **GLDM-оценки**, идеи которого были рассмотрены, позволяет получить коэффициенты квазилинейного рекуррентного соотношения, с помощью которых можно не только воссоздать заданные значениями временного ряда динамический процесс, но и построить качественный прогноз. В работе доказано, что коэффициенты, полученные при

помощи указанного алгоритма, дают минимальную невязку, а последовательность $\{(A^{(k)}, z^{(k)})\}_{k=1}^{\infty}$, построенная по алгоритму GLDM-оценки сходится к глобальному минимуму (a^*, z^*) задачи (2). Приведены теоретические обоснования выбора порядков модели для наборов данных различного объема и структуры, обсуждается роль нулевых коэффициентов в повышении эффективности модели. Приведены условия наличия нулевых коэффициентов модели, сформулированы условия, при которых ряд коэффициентов модели может быть обнулен.

2. Разработанная система прогнозирования получает в качестве входных данных только временной ряд, пользователь не должен обладать дополнительной информацией о структуре временного ряда, специфике предметной области, настраивать параметры моделей. Система является расширяемой, она допускает подключение дополнительных модулей прогнозирования и включение их в анализ точности работы моделей. Система позволяет анализировать временные ряды для различных прикладных задач. Исследования возможностей использования параллелизма для алгоритма идентификации коэффициентов модели с помощью обобщенного метода наименьших модулей для достаточно большого числа коротких временных рядов показали, что можно разработать параллельное персональное приложение для ноутбуков и/или мобильных телефонов, максимально используя ресурсы этих устройств и позволяя получить результат в 5–6 раз быстрее, поскольку разница во времени использования персонального компьютера для расчета больших наборов данных будет иметь важное значение.

Направления будущих исследований включают разработку двунаправленных GLDM моделей высокого порядка. Оптимизация модели позволяет найти оптимальные коэффициенты, минимизирующие отклонение модели от реальных данных. Это расширение позволяет значительно улучшить качество прогнозирования и адаптацию модели к реальным данным. Дальнейшие исследования будут включать адаптацию GLDM для многомерных временных рядов, исследование интеграции лаговых членов порядка выше пятого, а также исследование возможности включения дополнительных объясняющих переменных для дальнейшего улучшения предсказательных возможностей модели и адаптивности к сложным реальным данным.

Публикации автора по теме диссертации

Публикации в журналах из списка ВАК

1. Аботалёб М. С. А. Установление значимости коэффициентов квазилинейного уравнения N-факторной авторегрессии // Проблемы информатики. – 2024. – № 3. – С. 5–28. DOI: 10.24412/2073-0667-2024-3-5-28 (Перечень ВАК К2).
2. Макаровских Т. А., Аботалёб М. С. А. Автоматический подбор параметров модели ARIMA для прогноза количества случаев заражения и смерти от Covid-19 // Вестник ЮУрГУ. Серия: Вычислительная математика и информатика. – 2021. – Т. 10, № 2. – С. 20–37. DOI: 10.14529/cmse210202 (Перечень ВАК К2, RSCI).
3. Аботалёб М. С. А., Макаровских Т. А., Панюков А. В. Возможности параллелизма при идентификации квазилинейного рекуррентного уравнения // Вестник ЮУрГУ. Серия: Вычислительная математика и информатика. – 2023. – Т. 12, № 4. – С. 94–109. DOI: 10.14529/cmse230404 (Перечень ВАК К2, RSCI).

Публикации, индексируемые в Web of Science и Scopus

4. Abotaleb M. Solving The Optimizing Parameters Problem For Non-linear Datasets Using The High-Order General Least Deviations Method (GLDM) Algorithm // Computational Methods for Differential Equations. – 2024. DOI: 10.22034/cmde.2024.62441.2751 (Web of Science Q2, Scopus Q1).
5. Abotaleb M. Soft Computing-Based Generalized Least Deviation Method Algorithm for Modeling and Forecasting COVID-19 using Quasilinear Recurrence Equations // Iraqi Journal For Computer Science and Mathematics. – 2024. – Vol. 5, No. 3. – P. 441–472. DOI: 10.52866/ijcsm.2024.05.03.028 (Scopus Q2).
6. Abotaleb M. Proving Optimal Model Selection and Zero Coefficient Cases in Time Series Forecasting with the Generalized Least Deviation Algorithm // Advanced Mathematical Models & Applications. – 2024. – Vol. 9, No. 3. – P. 475–524. DOI: 10.62476/amma93475 (Scopus Q3).
7. Abotaleb M., Makarovskikh T. Advanced milk production modelling using high-order generalized least deviation method // Modeling Earth Systems and Environment. – 2024. – Vol. 10. – P. 7019–7047. DOI: 10.1007/s40808-024-02142-9 (Web of Science Q3, Scopus Q1).

Свидетельства о регистрации программ для ЭВМ

8. Аботалев М. С. А., Макаровских Т. А. Свидетельство Роспатента о государственной регистрации программы для ЭВМ «Параллельная программа для идентификации параметров квазилинейных рекуррентных соотношений для массива временных рядов» № 2023661570 от 01.06.2023.
9. Аботалев М. С. А., Макаровских Т. А. Свидетельство Роспатента о государственной регистрации программы для ЭВМ «Библиотека для квазилинейного анализа временных рядов» № 2022669812 от 25.10.2022.
10. Макаровских Т. А., Аботалев М. С. А. Свидетельство Роспатента о государственной регистрации программы для ЭВМ «Система прогнозирования для одномерных временных рядов» № 2022660664 от 08.06.2022.
11. Макаровских Т. А., Аботалев М. С. А. Свидетельство Роспатента о государственной регистрации программы для ЭВМ «Библиотека epidemic.ta для прогнозирования случаев заражения Covid-19» № 2021616747 от 26.04.2021.
12. Аботалев М. С. А., Макаровских Т. А. Свидетельство Роспатента о государственной регистрации программы для ЭВМ «Univariate time series software Version 1 (Программное обеспечение для одномерных временных рядов Версия 1)» № 2024682644 от 25.09.2024.

Исходные тексты программ, реализующих алгоритмы и методы, разработанные в рамках настоящего диссертационного исследования, а также наборы данных, использованные для проведения вычислительных экспериментов, подтверждающих эффективность указанных подходов, размещены в свободном доступе в следующих репозиториях:

<https://gitverse.ru/mostafa/GLDM>

<https://github.com/abotalebmostafa11/GLDMHO>

Код пакета `abotaleb1`, предназначенного для моделирования одномерных временных рядов с использованием метода обобщенных наименьших отклонений (GLDM), также опубликован в открытом доступе на <https://gitverse.ru/mostafa/GLDM>. Подробная информация о пакете и инструкции по его установке представлены на странице PyPI: <https://pypi.org/project/abotaleb1/>.