На правах рукописи

Азарян Алексан Артурович

БЫСТРЫЕ АЛГОРИТМЫ МОДЕЛИРОВАНИЯ МНОГОМЕРНЫХ ЛИНЕЙНЫХ РЕГРЕССИОННЫХ ЗАВИСИМОСТЕЙ НА ОСНОВЕ МЕТОДА НАИМЕНЬШИХ МОДУЛЕЙ

05.13.18 – Математическое моделирование, численные методы и комплексы программ

Автореферат
диссертации на соискание ученой степени кандидата
физико-математических наук

Екатеринбург – 2019

Работа выполнена в ФГАОУ ВО «Уральский федеральный университет имени первого Президента России Б.Н. Ельцина» на кафедре прикладной математики

Научный руководитель:	доктор технических наук, доцент, Тырсин Александр Николаевич
Официальные оппоненты:	Гитман Михаил Борисович, доктор физико-математических наук, профессор, ФГБОУ ВО «Пермский национальный исследовательский политехнический университет», профессор кафедры «Вычислительная математика, механика и биомеханика» Кризский Владимир Николаевич, доктор физико-математических наук, профессор, Стерлитамакский филиал ФГБОУ ВО «Башкирский государственный университет», заместитель директора по научной работе
Ведущая организация:	ФГАОУ ВО «Российский университет дружбы народов», г. Москва

Защита состоится «20» марта 2019 г. в 13:00 на заседании диссертационного совета Д 212.285.25 на базе ФГАОУ ВО «Уральский федеральный университет имени первого Президента России Б.Н. Ельцина» по адресу: 620000, г. Екатеринбург, пр. Ленина, 51, зал заседаний диссертационных советов, комн. 248.

С диссертацией можно ознакомиться в библиотеке и н сайте ФГАОУ ВО «Уральский федеральный университет имени первого Президента России Б.Н. Ельцина», по адресу: http://lib.urfu.ru/mod/data/view.php?d=51&rid=287876

Автореферат разослан «____» ____________ 2019 г.

Ученый секретарь диссертационного совета,
доктор физико-математических наук, профессор В.Г. Пименов

ОБЩАЯ ХАРАКТЕРИСТИКА РАБОТЫ

Актуальность темы исследования. Математическое моделирование регрессионных зависимостей по экспериментальным данным выполняется с помощью статистических методов оценки параметров моделей. Во многих случаях имеется достаточная информация об изучаемых объектах, процессах и о свойствах, действующих на них возмущений. Это позволяет воспользоваться эффективными методами оценивания неизвестных параметров с использованием классических методов максимального правдоподобия. Для задачи оценивания линейных регрессионных моделей в предположении нормального распределения случайных погрешностей измерений, методом максимального правдоподобия является метод наименьших квадратов (МНК). На основе МНК создана целостная система статистической обработки. С учетом простоты реализации он является наиболее распространенным статистическим методом построения зависимостей.

Однако при построении математических моделей, например, в задачах мониторинга и диагностики, приходится сталкиваться со стохастической неоднородностью. Строгого определения неоднородности нет. Приведем определение[1]: «Будем считать однородной такую совокупность, элементы которой формируются под воздействием общих основных причин и условий, а их законы распределения имеют простую структуру, и неоднородной – если разные ее элементы формируются под влиянием разных причин и условий либо если она может быть представлена в виде объединения некоторого числа однородных совокупностей с более простой структурой законов распределения элементов». Применительно к регрессионным моделям к основным признакам стохастической неоднородности следует отнести: не полное соответствие модели части наблюдений; возможное наличие в выборке резко выделяющихся наблюдений, не обязательно обусловленных ошибками измерений; зачастую не экспериментальный, не однородный характер данных; использование различных группировок и округлений; возможная зависимость результатов наблюдений.

Данные особенности при использовании классических процедур оценивания могут привести к грубым ошибкам. В этой ситуации используют устойчивые (робастные и непараметрические) методы оценивания. Однако эти методы проигрывают МНК в быстродействии. Поэтому актуальным направлением исследований является повышение вычислительной эффективности регрессионного моделирования по экспериментальным данным в условиях стохастической неоднородности.

Степень разработанности темы исследования. В основе устойчивого регрессионного моделирования лежит метод наименьших модулей (МНМ)[2,3], также называемый l_1-аппроксимацией. Его важная особенность – детерминированный характер, т.к. не требуется привлечения гипотез о вероятностных свойствах изучаемых явлений[4].

---

[1] Смоляк С.А., Титаренко Б.П. Устойчивые методы оценивания. М.: Статистика, 1980.

[2] Мудров В.И., Кушко В.Л. Методы обработки измерений. Квазиправдоподобные оценки. М.: Радио и связь, 1983.

[3] Bloomfield P., Steiger W.L. Least Absolute Deviations: Theory, Applications, and Algorithms. Brikhauser, 1983.

[4] Акимов, П.А. Метод l_1–аппроксимации в навигационных задачах оценивания: дис. ... канд. физ.-мат. наук: 01.02.01 / Акимов Павел Александрович. – Москва, 2011.

Первые упоминания о МНМ связываются с работами Р. Босковича (R.J. Boscovich) и П.С. Лапласа (P.-S. de Laplace) второй половины XVIII века. Из современных исследований в области l_1-аппроксимации отметим работы М.В. Болдина, Г.И. Симоновой и Ю.Н. Тюрина, В.И. Мудрова и В.Л. Кушко, А.В. Панюкова, И.Б. Челпанова, П. Блумфилда (P. Bloomfield) и У. Стейгера (W.L. Steiger), Я. Додже (Ya. Dodge), Г. Бассета (G. Basset) и Р. Коенкера (R. Koenker), Д. Полларда (D. Pollard), С. Бойда (S. Boyd) и др. Вопросы алгоритмической реализации МНМ в линейном регрессионном моделировании рассматривались в работах А.И. Матасова и П.А. Акимова, С.И. Зуховицкого и Л.И. Авдеевой, В.И. Мудрова и В.Л. Кушко, А.Н. Тырсина, Р. Армстронга (R.D. Armstrong) и Д. Кунга (D.S. Kung), А. Барродейла (I. Barrodale) и Ф. Робертса (F. D. K. Roberts), Е. Вейсфельда (E. Weiszfeld), Г. Весоловски (G. O.Wesolowsky), С. Нарула (S.C. Narula) и Дж. Веллингтона (J.F. Wellington), В. Фишера (W.D. Fisher) и др.

Используемая при оценивании параметров регрессии функция потерь, либо является модулем (МНМ-оценки), либо функцией от модуля. При оценивании пространственных линейных регрессионных моделей обычно удается ограничиться МНМ-оценками или оценками, использующими выпуклые функции потерь. Однако одностороннее засорение экспериментальных данных или включение в состав объясняющих переменных временных лагов от выходной переменной приводит к смещению и неустойчивости МНМ-оценок[5]. С целью устранения этих недостатков используют оценки с функциями потерь, имеющими горизонтальную асимптоту[6]. Непосредственное использование выпукло-вогнутых функций потерь приводит к появлению множества неизвестных локальных минимумов у целевой функции, что затрудняет поиск глобального минимума. Этот недостаток можно устранить за счет использования в качестве начального приближения вектора параметров модели его МНМ-оценки или оценки, полученной с помощью обобщенного метода наименьших модулей (ОМНМ)[7]. Отметим, что ОМНМ можно непосредственно использовать для устойчивого оценивания моделей.

Известные точные алгоритмы реализации МНМ и ОМНМ при оценивании линейных регрессионных зависимостей являются достаточно эффективными лишь для малых размерностей моделей и ограниченного объема выборок, а приближенные алгоритмы имеют ограниченную точность, поскольку требование увеличения точности приводит к резкому росту вычислительных затрат. Это затрудняет использование данных методов в динамических задачах мониторинга и диагностики. Поэтому в этих приложениях алгоритмы численного оценивания моделей за ограниченное время и с приемлемой точностью представляют значительный интерес. Отметим, что вопрос о сходимости приближенных алгоритмов остается открытым.

Таким образом, **актуальны** разработка единого подхода к вычислительно эффективному моделированию многомерных линейных регрессионных зависимостей

[5] Болдин, М.В. Знаковый статистический анализ линейных моделей / М.В. Болдин, Г.И. Симонова, Ю.Н. Тюрин. – М.: Наука. Физматлит, 1997. – 288 с.

[6] Айвазян, С.А. Прикладная статистика: Исследование зависимостей / С.А. Айвазян, И.С. Енюков, Л.Д. Мешалкин. – М.: Финансы и статистика, 1985. – 488 с.

[7] Тырсин, А.Н. Робастная параметрическая идентификация моделей диагностики на основе обобщенного метода наименьших модулей: дис. … докт. тех. наук: 05.13.18 / Тырсин Александр Николаевич. – Челябинск, 2007. – 327 с.

в условиях стохастической неоднородности на основе МНМ и ОМНМ, не имеющих ограничений на порядок моделей и объем экспериментальных данных и проведение исследований для его теоретического обоснования.

Цели и задачи исследования. Целью работы является разработка и теоретическое обоснование нового подхода к вычислительно эффективному моделированию многомерных линейных регрессионных зависимостей в условиях стохастической неоднородности, а также создание на его основе комплекса алгоритмов и программ реализации метода наименьших модулей и обобщенного метода наименьших модулей.

Для достижения поставленной цели необходимо решить следующие задачи:

1. Предложить и исследовать новый подход для вычислительно эффективного моделирования многомерных линейных регрессионных зависимостей на основе методов наименьших модулей и обобщенных наименьших модулей.

2. Разработать и исследовать алгоритмы вычислительно эффективного моделирования многомерных линейных регрессионных зависимостей на основе методов наименьших модулей и обобщенных наименьших модулей.

3. Выполнить анализ вычислительной трудоемкости предложенных алгоритмов и провести их сравнение с известными результатами.

4. Разработать комплекс проблемно-ориентированных программ для проведения вычислительных экспериментов с целью исследования эффективности предложенных алгоритмов моделирования многомерных линейных регрессионных зависимостей.

Научная новизна.

В области математического моделирования:

1. Впервые установлена закономерность, позволяющая осуществлять моделирование многомерных линейных регрессионных зависимостей методом наименьших модулей локально – посредством спуска по узловым прямым. На основе этого предложен новый подход к вычислительно эффективному математическому моделированию многомерных линейных регрессионных зависимостей в условиях стохастической неоднородности, основанный на спуске по узловым прямым.

2. Предложенный подход к моделированию многомерных линейных регрессионных зависимостей, основанный на спуске по узловым прямым, реализован для обобщенного метода наименьших модулей. Установлена закономерность сокращения числа рассматриваемых возможных решений с увеличением размерности данных и числа наблюдений, позволившая обеспечить вычислительную эффективность моделирования линейных регрессионных зависимостей обобщенным методом наименьших модулей.

3. Установлено, что обобщенный метод наименьших модулей при некоторых ограничениях можно распространить и на случай моделирования многомерных нелинейных регрессионных зависимостей.

В области численных методов:

1. Разработаны вычислительно эффективные алгоритмы для моделирования многомерных линейных регрессионных зависимостей методом наименьших модулей.

2. Доказана сходимость предложенных алгоритмов оценивания параметров многомерных линейных регрессионных моделей методом наименьших модулей к точному решению за конечное число шагов.

3. Разработан вычислительно эффективный алгоритм моделирования линейных зависимостей методом обобщенных наименьших модулей.

4. Выполнен анализ вычислительной трудоемкости разработанных алгоритмов моделирования многомерных линейных регрессионных зависимостей.

В области комплексов программ:

1. Разработан программный комплекс, позволяющий: проводить вычислительные эксперименты как на модельных, так и на реальных данных с целью исследования эффективности предложенных алгоритмов оценивания многомерных линейных регрессионных моделей; применять и строить линейные модели с помощью разработанных алгоритмов для проведения вычислительных экспериментов; в качестве платформы для реализации разработанных алгоритмов используется язык программирования R.

2. С помощью разработанного комплекса программ решено несколько задач моделирования в технике и экономике.

Теоретическая и практическая значимость работы. Значимость диссертационного исследования обусловлена решением актуальных задач моделирования многомерных линейных регрессионных зависимостей в условиях стохастической неоднородности с применением современного математического аппарата. Полученные результаты развивают теорию моделирования регрессионных зависимостей на основе метода наименьших модулей. Разработанные алгоритмы моделирования реализуют общую идею спуска по узловым прямым и превосходят по вычислительной эффективности все известные аналоги. Алгоритмическая реализация в рамках предложенной идеи спуска по узловым прямым обобщенного метода наименьших модулей позволяет моделировать авторегрессионные зависимости. Наряду с высоким быстродействием они обладают достаточно простой структурой.

Использование разработанных алгоритмов и программ позволит существенно снизить вычислительные затраты при практическом моделировании реальных систем и явлений в виде регрессионных зависимостей. Представленные результаты вычислительных экспериментов свидетельствуют об адекватности проведенного математического моделирования и эффективности подхода на основе спуска по узловым прямым для дальнейшего его развития в технике и экономике в задачах диагностики систем и объектов в режиме реального времени.

Методология и методы диссертационного исследования.

Объектом исследования являются многомерные линейные регрессионные модели, параметры которых оцениваются в условиях стохастической неоднородности экспериментальных данных.

Предметом исследования является единый подход к построению алгоритмов моделирования многомерных линейных регрессионных зависимостей в условиях стохастической неоднородности и его теоретическое обоснование.

Для решения поставленных задач в работе используются методы математического моделирования, математической статистики, теории случайных процессов, статистических испытаний Монте–Карло, матричной алгебры, численных методов решения экстремальных задач, линейного программирования. Для программной реализации предложенных методов и алгоритмов были применены современные информационные технологии и средства программирования. Был разработан программный комплекс в среде RStudio с применением языка программирования R.

Степень достоверности результатов. Достоверность и обоснованность полученных результатов обеспечены математической строгостью постановки задач, корректным использованием математического аппарата и методов математического моделирования, согласованием результатов вычислительных экспериментов с модельными примерами, решением большого количества тестовых задач и практическим применением разработанного комплекса программ. Адекватность математической модели подтверждалась примерами ее использования. Полученные в работе результаты и выводы согласуются с результатами других авторов. Все результаты, выносимые на защиту, опубликованы.

Апробация работы. Теоретические и практические результаты исследований докладывались на следующих конференциях: 2-я научно-техническая конференция молодых ученых Уральского энергетического института (Екатеринбург, 2017), XI Международная школа–симпозиум «Анализ, моделирование, управление, развитие социально-экономических систем» (Симферополь–Судак, 2017), XVIII Всероссийский Симпозиум по прикладной и промышленной математике (Сочи, 2017), XIX Всероссийский Симпозиум по прикладной и промышленной математике (Санкт–Петербург, 2018), 12–я международная конференция «Новые информационные технологии в исследовании сложных структур» (Алтайский край, пос. Катунь, 2018). Также результаты работы обсуждались на научном семинаре Уральского федерального университета имени первого Президента России Б.Н. Ельцина (УрФУ) «Модели и методы оптимизации, оценивания данных и управления в технических и экономических системах» под руководством д.ф.-м.н. А.Ф. Шорикова (Екатеринбург, 2018), научном семинаре отдела математического программирования Института математики и механики им. Н.Н. Красовского УрО РАН под руководством д.ф.-м.н. М.Ю. Хачая (Екатеринбург, 2018) и семинарах кафедры прикладной математики УрФУ (Екатеринбург, 2016–2018).

Работа выполнялась в соответствии с планами научно–исследовательских работ по грантам РФФИ № 16–06–00048 и № 17–01–00315.

Комплекс из трех программ, предназначенный для моделирования и исследования регрессионных зависимостей с помощью МНМ и ОМНМ, зарегистрирован в Федеральной службе по интеллектуальной собственности (РОСПАТЕНТ).

Публикации. По теме диссертации опубликовано 18 работ [1–18], в т.ч. 5 статей в рецензируемых научных изданиях и журналах, рекомендованных ВАК Минобрнауки РФ [1–5], из которых одна включена в наукометрические базы MathSciNet и Zentralblatt MATH [1], одна – в Zentralblatt MATH [2] и одна – в GeoRef и Chemical Abstracts [4], глава монографии [9] и 3 свидетельства о государственной регистрации программ для ЭВМ [6–8].

Личное участие автора. Содержание диссертации и основные положения, выносимые на защиту, отражают персональный вклад автора в опубликованные работы. Из работ, выполненных в соавторстве, в диссертацию вошли только результаты, полученные ее автором. Все результаты диссертации, в том числе разработка, исследование и обоснование математических моделей и методов их исследования, разработка комплекса компьютерных моделей и экспериментальных методик, доказательство всех утверждений, проведение численных расчетов и моделирования, получены лично автором диссертации. Научным руководителем предложены постановки задач.

Структура и объем работы. Диссертация состоит из введения, 4 глав, заключения, списка литературы и приложений. Полный объем составляет 148 страниц, включая 21 рисунок, 27 таблиц, список литературы из 135 наименований и 3 приложения.

КРАТКОЕ СОДЕРЖАНИЕ ДИССЕРТАЦИИ

Во **Введении** обосновывается актуальность и научная новизна работы, ставятся цели работы и формулируются основные задачи для достижения поставленных целей, приводятся методы исследования, содержится информация по апробации работы, описана структура диссертации и ее краткое содержание.

Первая глава посвящена обзору устойчивых методов моделирования линейных регрессионных зависимостей.

В **первом параграфе** рассматривается проблема устойчивости моделирования линейных зависимостей в условиях стохастической неоднородности. Одной из типовых задач при статистической обработке результатов экспериментальных исследований является оценивание коэффициентов линейных регрессионных моделей

$$\mathbf{y} = \mathbf{X}\mathbf{a} + \boldsymbol{\varepsilon}, \tag{1}$$

где $\mathbf{y} = \begin{pmatrix} y_1 \\ y_2 \\ \dots \\ y_n \end{pmatrix}$ — наблюдаемые значения зависимой переменной; $\mathbf{X} = \{x_{ij}\}_{n\times m} = \begin{pmatrix} 1 & x_{12} & \dots & x_{1m} \\ 1 & x_{22} & \dots & x_{2m} \\ \dots & \dots & \dots & \dots \\ 1 & x_{n2} & \dots & x_{nm} \end{pmatrix}$ — значения объясняющих переменных, $\mathbf{x}^{(j)} = \begin{pmatrix} x_{1j} \\ x_{2j} \\ \dots \\ x_{nj} \end{pmatrix}$; $\boldsymbol{\varepsilon} = \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \dots \\ \varepsilon_n \end{pmatrix}$ — случайные ошибки измерений; $\mathbf{a} = \begin{pmatrix} a_1 \\ a_2 \\ \dots \\ a_m \end{pmatrix}$ — неизвестные коэффициенты.

При оценивании параметров линейной модели, использование классических процедур на основе МНК, ориентированных на выполнение основных предпосылок математической статистики может привести к грубым ошибкам. Особенно это проявляется, если в экспериментальных данных имеются аномальные наблюдения в виде выбросов, а также при нестационарности случайных ошибок. В этой ситуации используют устойчивые (робастные и непараметрические) методы оценивания. В их основе лежит МНМ. Используемая при оценивании параметров регрессии функция потерь либо является модулем, либо функцией от модуля.

Наиболее актуальным является обеспечение устойчивости оценок параметров моделей временных рядов, поскольку они гораздо более чувствительны к стохастической неоднородности данных, ввиду зависимости наблюдений. В результате при расчете коэффициентов моделей временных рядов, неустойчивыми к выбросам оказываются и МНМ–оценки. Недостатком МНМ также является то, что, он рассчитан на симметричный характер выбросов. С целью устранения этих недостатков используют оценки с функциями потерь, имеющими горизонтальную асимптоту. Одним из таких методов является обобщенный метод наименьших модулей (ОМНМ).

Во **втором параграфе** описывается метод наименьших модулей, который для модели (1) представляет собой минимизацию суммы модулей невязок

$$Q(\mathbf{a}) = \sum_{i=1}^{n} \left| y_i - \sum_{j=1}^{m} a_j x_{ij} \right| = \sum_{i=1}^{n} |y_i - \langle \mathbf{x}_i, \mathbf{a} \rangle| \to \min_{\mathbf{a} \in \mathbf{R}^m}. \qquad (2)$$

Целевая функция $Q(\mathbf{a})$ является выпуклой, непрерывной и кусочно-линейной, но у нее «плохая» геометрия – имеется большое число изломов в виде отрезков прямых, причем они вытянуты, скрещены, многие из них почти параллельны и находятся на близком расстоянии друг от друга, образуя так называемый «пучок» (рис. 1). Минимум у $Q(\mathbf{a})$ находится внутри этого «пучка». Стенки изломов представляют собой выпуклые линейные гиперграни, которые при приближении к минимуму целевой функции постепенно становятся почти параллельными. Эта «плохая» геометрия целевой функции приводит ко многим затруднениям при реализации разных алгоритмов нахождения точного решения задачи (2) на основе идей линейного программирования. Это вызвано тем, что:

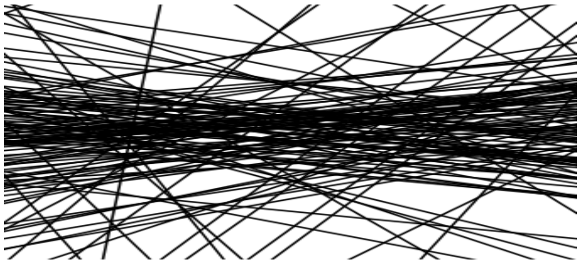

Рис. 1. График прямых, на которых расположены изломы целевой функции $Q(\mathbf{a})$, m = 2 и n = 128

- не учет особенностей целевой функции при переходе от одного к другому базисному решению приводит к отклонениям от «пучка» изломов, что замедляет сходимость к точному решению;
- на каждом шаге алгоритмов преобразуются все данные. В результате этого накапливаются вычислительные погрешности из-за ошибок округлений, что может привести к не точному решению задачи (2). Кроме того, преобразование всех данных для выпуклой целевой функции является вычислительно излишним и можно обойтись локальными данными для перехода от одного базисного решения к другому, при этом не происходит накопление вычислительных погрешностей.
- требуется хранить большой массив данных в виде симплекс-таблиц.

Другим методом для нахождения точного решения задачи (2) является алгоритм полного перебора, который осуществляет перебор всех особых точек, в которых не существует производная функции $Q(\mathbf{a})$ ни по одному из возможных направлений пространства $\mathbf{R}^m$. Данный алгоритм имеет $O\left(C_n^m \cdot (m^3 + m \cdot n)\right)$ вычислительную сложность. Следовательно, с ростом объема выборки и количества переменных наблюдается экспоненциальный рост вычислительных затрат.

Приближенным методом решения задачи (2) является итерационный метод вариационно-взвешенных квадратических приближений, на k-м шаге решается задача

$$F\left(\mathbf{a},\mathbf{a}^{(\boldsymbol{k}-\mathbf{1})}\right)=\sum_{i=1}^{n}\frac{1}{\left|y_i-\langle\mathbf{x}_i,\mathbf{a}^{(\boldsymbol{k}-\mathbf{1})}\rangle\right|}(y_i-\langle\mathbf{x}_i,\mathbf{a}\rangle)^2\rightarrow\min_{\mathbf{a}\in\mathbf{R}^m},$$

где k— номер итерации, $\mathbf{a}^{(\mathrm{k}-1)}$— вектор, полученный на предыдущей итерации. Но известно, что в точке решения задачи (2) m невязок равны нулю. Для решения этой проблемы используется прием регуляризации. Задается некоторое достаточно малое положительное число λ и i-й весовой коэффициент выбирается следующим образом:

$$\begin{cases}\left|y_i-\langle\mathbf{x}_i,\mathbf{a}^{(\boldsymbol{k})}\rangle\right|^{-1}, & \text{если } \left|y_i-\langle\mathbf{x}_i,\mathbf{a}^{(\boldsymbol{k})}\rangle\right|\geq\lambda,\\ \dfrac{1}{\lambda}, & \text{если } \left|y_i-\langle\mathbf{x}_i,\mathbf{a}^{(\boldsymbol{k})}\rangle\right|<\lambda.\end{cases}$$

В случае регуляризации, вопрос о сходимости алгоритма вариационно-взвешенных-квадратических приближений остается открытым.

Для решения задачи (2), также применяют методы поиска безусловного экстремума нулевого порядка. Но из-за «плохой» геометрии целевой функции при повышении точности наблюдается быстрый рост вычислительных затрат.

В **третьем параграфе** описывается ОМНМ, который для модели (1) имеет вид:

$$W(\mathbf{a})=\sum_{i=1}^{n}\rho(|y_i-\langle\mathbf{x}_{\mathrm{i}},\mathbf{a}\rangle|)\rightarrow\min_{\mathbf{a}\in\mathbf{R}^m}, \tag{3}$$

где $\rho(\cdot)$ − некоторая монотонно возрастающая, дважды непрерывно-дифференцируемая на положительной полуоси функция, причем $\rho(0)=0, \forall t>0 \quad 0<\rho'(t)<\infty, -\infty<\rho''(t)<0$.

Так же, как и в случае МНМ, для нахождения точного решения задачи (3) можно применить переборный алгоритм. Но и здесь, с ростом объема выборки и количества переменных наблюдается экспоненциальный рост вычислительных затрат.

Известный итерационный алгоритм для нахождения решения задачи (3) основанный на идеях линейного программирования, по тем же соображениям, которые были изложены для МНМ, не гарантирует сходимость к точному решению.

Оценки, полученные с помощью модифицированного ОМНМ на основе разбиения исходной выборки на подвыборки, могут значительно отличаться от точного решения. А повышение точности оценивания за счёт осреднения многократного повторного формирования подвыборок приводит к росту вычислительных затрат.

Во **второй главе** описан и исследован новый подход к вычислительно эффективному построению многомерных линейных регрессионных моделей в условиях стохастической неоднородности на основе МНМ и ОМНМ. Доказана его сходимость для МНМ. Описаны эффективные алгоритмы оценивания линейных моделей на основе предложенного подхода и проведен анализ их вычислительной трудоемкости.

В **первом параграфе**, для предложения нового подхода к эффективному построению линейных моделей на основе МНМ и доказательства его сходимости, введутся некоторые обозначения и определения.

Обозначим $\Omega = \{\Omega_1,\ \Omega_2,\ \dots\ ,\ \Omega_n\}$ – множество всех гиперплоскостей вида

$$\Omega_i = \Omega(\mathbf{a}, \mathbf{x}_i, y_i) = y_i - \langle \mathbf{x}_i, \mathbf{a} \rangle = 0, (i = 1, 2, \dots, n). \qquad (4)$$

Определение 1. Назовем узловой точкой, точку пересечения m различных гиперплоскостей вида (4):

$$\mathbf{u} = \bigcap_{s \in M} \Omega_s, M = \{k_1, \dots, k_m\}, k_1 < k_2 < \dots < k_m, k_l \in \{1, 2, \dots, n\}. \qquad (5)$$

Обозначим U – множество всех узловых точек (5).

Определение 2. Назовем узловой такую прямую, которая является пересечением $m - 1$ различных гиперплоскостей Ω_i:

$$l_{(k_1,\dots,k_{m-1})}: \bigcap_{i=k_1}^{k_{m-1}} \Omega_i\ , k_l \in \{1, 2, \dots\ , n\}.$$

Поскольку целевая функция выпуклая, то спуск у нее наиболее эффективен по ее «изломам». Но этими «изломами» являются отрезки узловых прямых. Следовательно, построение линейных моделей будет более эффективным, при спуске по узловым прямым. При этом мы устраняем все указанные выше проблемы известных методов решения задачи (2). В основе описанного подхода лежит теорема.

Теорема 1. Рассмотрим модель (1), для которой имеется выборка наблюдений $(\mathbf{x}_i, y_i) = (x_{i1}, x_{i2},\ \dots\ ,\ x_{im}, y_i)$, $(i\ =\ 1,\ \dots\ ,\ n)$, и пусть заданы функция $Q(\mathbf{a})$ задачи (2), гиперплоскости (4) и множество U всех узловых точек (5). Тогда функция $Q(\mathbf{a})$ всегда имеет точку глобального минимума, эта точка либо единственна и принадлежит U, либо состоит из выпуклого линейного многогранника, вершины которого являются точками из U.

У целевой функции $W(\mathbf{a})$ глобальный минимум также находится в узловой точке, и она имеет вытянутые «изломы», расположенные на узловых прямых. Поэтому поиск ее глобального минимума также осуществляем по узловым прямым.

Во **втором параграфе** описываются алгоритмы спуска по узловым прямым для метода наименьших модулей. Он выполняется следующим образом (рис. 2).

В качестве начального приближения берется узловая точка, являющаяся пересечением первых m гиперплоскостей $\Omega_1,\dots,\Omega_m$. Исключив одну из гиперплоскостей, получим узловую прямую l. В любой узловой точке можно построить m таких узловых прямых. Выберем ту, вдоль которой целевая функция достигает наименьшего значения, которое всегда будет достигаться в одной из узловых точек. Найдя эту точку, продолжим движение из нее по тому же принципу. В результате будет найдена узловая точка, спуск из которой невозможен. И эта узловая точка будет являться точным решением задачи (2).

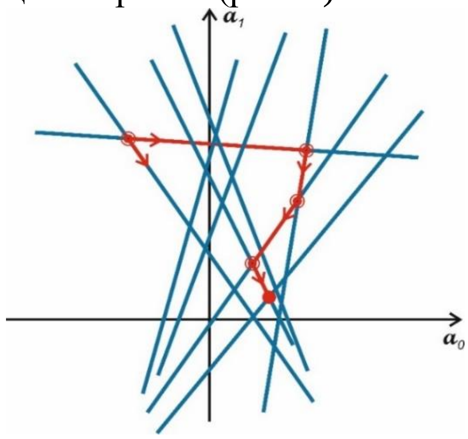

Рис 2. Спуск по узловым прямым

Доказана теорема о сходимости вышеописанного алгоритма.

Теорема 2. Алгоритм спуска вдоль узловых прямых сходится к точному решению задачи (2) за конечное число шагов.

Двигаясь вдоль прямой $l_{(k_1,\dots,k_{m-1})}$, для нахождения узловых точек, принадлежащих этой прямой, нужно для каждой точки решать СЛАУ порядка m:

$$\begin{cases} a_1 + a_2 x_{k_1,2} + a_3 x_{k_1,3} + \dots + a_m x_{k_1,m} = y_{k_1} \\ a_1 + a_2 x_{k_2,2} + a_3 x_{k_2,3} + \dots + a_m x_{k_2,m} = y_{k_2} \\ \dots \\ a_1 + a_2 x_{k_{m-1},2} + a_3 x_{k_{m-1},3} + \dots + a_m x_{k_{m-1},m} = y_{k_{m-1}} \\ a_1 + a_2 x_{i,2} + a_3 x_{i,3} + \dots + a_m x_{i,m} = y_i \end{cases}, \qquad (6)$$

где $k_1 < k_2 < \dots < k_{m-1}, i \in \{1, 2, \dots, n\}, i \notin \{k_1, k_2, \dots, k_{m-1}\}$.

Очевидно, что СЛАУ двух различных узловых точек, принадлежащих этой прямой, отличаются лишь одним уравнением. Следовательно, вычислительная эффективность алгоритма спуска существенно повысится, если для нахождения узловых точек, которые лежат на прямой $l_{(k_1,\dots,k_{m-1})}$, первые $(m-1)$ строк расширенной матрицы соответствующей (6), с помощью элементарных преобразований, предварительно преобразуем к ступенчатому виду:

$$\mathbf{A}'_{l_{(k_1,\dots,k_{m-1})}} = \begin{pmatrix} 1 & x_{k_1,2} & x_{k_1,3} & \dots & x_{k_1,m-1} & x_{k_1,m} & y_{k_1} \\ 0 & 1 & x'_{k_2,3} & \dots & x'_{k_2,m-1} & x'_{k_2,m} & y'_{k_2} \\ 0 & 0 & 1 & \dots & x'_{k_3,m-1} & x'_{k_3,m} & y'_{k_3} \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & 1 & x'_{k_{m-1},m} & y'_{k_{m-1}} \end{pmatrix}.$$

Использование ступенчатой матрицы $\mathbf{A}'_{l_{(k_1,\dots,k_{m-1})}}$ позволяет значительно сократить вычислительные затраты на нахождение всех узловых точек, лежащих на прямой $l_{(k_1,\dots,k_{m-1})}$: для каждой i-й узловой точки имеем расширенную матрицу

$$\mathbf{A}_{u_{(k_1,\dots,k_{m-1},i)}} = \begin{pmatrix} 1 & x_{k_1,2} & x_{k_1,3} & \dots & x_{k_1,m-1} & x_{k_1,m} & y_{k_1} \\ 0 & 1 & x'_{k_2,3} & \dots & x'_{k_2,m-1} & x'_{k_2,m} & y'_{k_2} \\ 0 & 0 & 1 & \dots & x'_{k_3,m-1} & x'_{k_3,m} & y'_{k_3} \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & 1 & x'_{k_{m-1},m} & y'_{k_{m-1}} \\ 1 & x_{i,2} & x_{i,3} & \dots & x_{i,m-1} & x_{i,m} & y_i \end{pmatrix}, \qquad (7)$$

где $k_1 < k_2 < \dots < k_{m-1}, i \in \{1, 2, \dots, n\}, i \notin \{k_1, k_2, \dots, k_{m-1}\}$. Для нахождения узловой точки остается выполнить только обратный ход метода Гаусса. Варьируя номер i в (7), найдем все узловые точки, лежащие на прямой $l_{(k_1,\dots,k_{m-1})}$.

Поскольку все остальные строки матрицы (7) остаются теми же, то наряду с экономией вычислительных затрат не накапливаются вычислительные погрешности.

Благодаря выпуклости целевой функции можно дополнительно повысить вычислительную эффективность алгоритма спуска за счет учета направления спуска – при этом для нахождения минимума на узловой прямой не требуется рассматривать все принадлежащие ей узловые точки.

Выполнен анализ вычислительной трудоемкости алгоритма (теорема 3).

Теорема 3. Алгоритм спуска по узловым прямым с использованием матриц (7) и с учетом направления спуска имеет среднюю вычислительную сложность

$$V = O(m^2 \cdot n^2 + m^4 \cdot n \cdot \ln n + m^2 \cdot n \cdot \ln^2 n).$$

В **третьем параграфе** описывается алгоритм спуска по узловым прямым для обобщенного метода наименьших модулей. Алгоритм решения задачи (3) также основан на использовании узловых прямых и состоит в следующем:

Шаг 1. Используем в качестве начального приближения $\mathbf{a}^{(1)}$ узловую точку решение задачи (2) $\mathbf{a}^{(0)}$, т.е. считаем $\mathbf{a}^{(1)} = \mathbf{a}^{(0)}$.

Шаг 2. Уточняем решение путем спуска из узловой точки $\mathbf{a}^{(1)}$, поочередно двигаясь вдоль каждой из m узловых прямых, проходящих через $\mathbf{a}^{(1)}$. Выберем ту узловую точку, в которой функция $W(\mathbf{a})$ достигает наименьшего значения. Найдя эту точку, продолжим движение из нее по тому же принципу. В результате будет найдена узловая точка $\mathbf{a}^{(2)}$, спуск из которой невозможен. При движении вдоль узловой прямой используем, как и в случае решения задачи (2) ступенчатые расширенные матрицы.

Шаг 3. Для найденной узловой точки $\mathbf{a}^{(2)}$ определим множество гиперплоскостей $\Omega(\mathbf{a}^{(2)}) \subset \Omega$, состоящее из $\alpha \cdot 100\%$ гиперплоскостей, наименее удаленных от $\mathbf{a}^{(2)}$.

Шаг 4. С помощью перебора среди всех узловых точек, получаемых в результате пересечения гиперплоскостей из множества $\Omega(\mathbf{a}^{(2)})$ находим ту точку $\mathbf{a}^{(3)}$, в которой целевая функция принимает наименьшее значение.

Шаг 5. Найденную узловую точку $\mathbf{a}^{(3)}$ сравниваем с $\mathbf{a}^{(2)}$. Если они совпадают, то работа алгоритма останавливается и в качестве решения задачи (3) берем полученную узловую точку. В противном случае в качестве начального приближения $\mathbf{a}^{(1)}$ берем узловую точку $\mathbf{a}^{(3)}$, переходим к *Шагу 2* и повторяем вышеописанные действия.

Выполнен анализ вычислительной трудоемкости предложенного алгоритма, который описывается теоремой 4.

Теорема 4. Алгоритм решения задачи оценивания параметров линейных регрессионных моделей обобщенным методом наименьших модулей на основе спуска по узловым прямым имеет среднюю вычислительную сложность

$$V = O\left(n^2 \cdot m^3 \cdot \ln n + C_{\alpha \cdot n}^{m} \cdot (m^3 + m \cdot n)\right).$$

В **четвертом параграфе** рассмотрены вопросы оценивания параметров нелинейных регрессионных моделей с помощью ОМНМ, сформулированы условия, при которых возможно получение решения.

В **третьей главе** представлено описание программного комплекса (рис. 3) для устойчивого построения линейных многомерных регрессионных моделей на основе МНМ и ОМНМ, а также для проведения вычислительных экспериментов с целью исследования эффективности предложенных алгоритмов оценивания многомерных линейных регрессионных моделей. В качестве платформы для реализации разработанных алгоритмов используется язык программирования R.

В **первом параграфе** представлено описание архитектуры программного комплекса, дано подробное описание всех модулей, входящих в его состав. Программный комплекс позволяет решать задачи (2) и (3), проанализировать вычислительную сложность решения этих задач, а также провести сравнительный анализ алгоритма спуска по узловым прямым для МНМ и ОМНМ с известными точными и приближенными методами их решения.

Во **втором параграфе**, на основе разработанного программного комплекса проведен сравнительный анализ по быстродействию решения задачи (2) алгоритмом спуска по узловым прямым с известными алгоритмами. Результаты сравнения с точными алгоритмами на основе полного перебора узловых точек и решения эквивалентной задачи линейного программирования отражены в табл. 1 и на рис. 4.

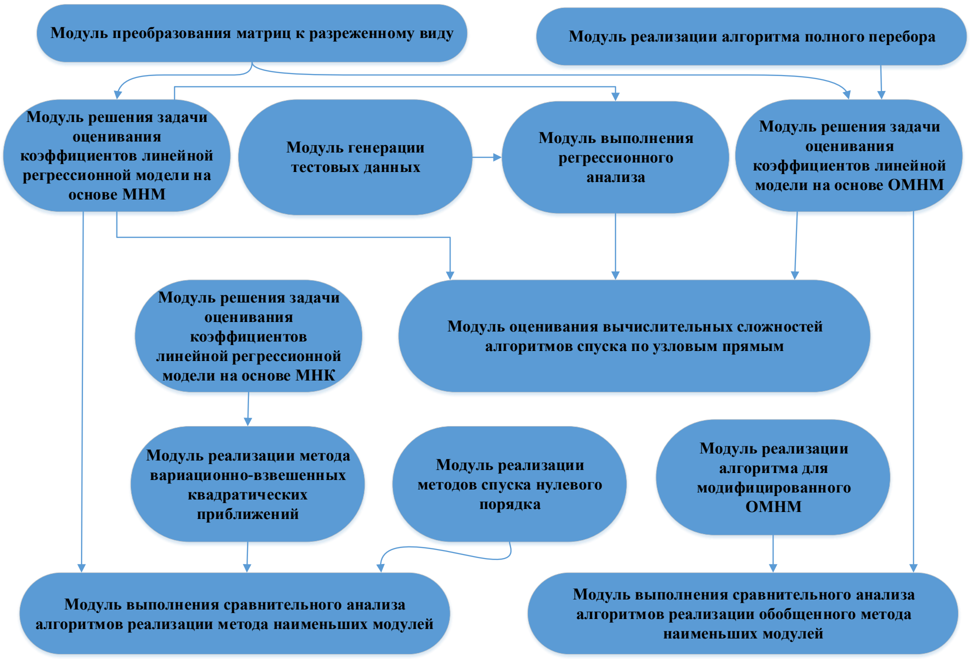

Рис. 3. Общая структура программного комплекса

Также методом статистических испытаний проведен ряд вычислительных экспериментов для сравнения алгоритма спуска по узловым прямым с алгоритмом на основе решения эквивалентной задачи линейного программирования. В табл. 2, 3 приведены результаты для решения задачи (2) для числа испытаний $N = 1000$ и количества параметров модели $m = 5$, где случайные ошибки ε имеют распределение

$$F(x) = (1-\gamma)N(0,\sigma^2) + \gamma\left(\frac{1}{\pi}\cdot \operatorname{arctg}\left(\frac{x-a_H}{\gamma_H}\right)+\frac{1}{2}\right), \qquad (8)$$

где $\sigma = 1;\ a_H = 0;\ \gamma_H = 1;\ \gamma = 0.1$.

Таблица 1. Вычислительная сложность алгоритмов нахождения точного решения задачи (2)

Алгоритм	Вычислительная сложность
Алгоритм полного перебора	$M_1 = O\left(C_n^m \cdot (m^3 + m\cdot n)\right)$
Симплекс-метод	$M_2 = O(n^3 \cdot \ln n)$
Алгоритм спуска по узловым прямым	$M_3 = O(m^2\cdot n^2 + m^4\cdot n\cdot \ln n + m^2\cdot n\cdot \ln^2 n)$

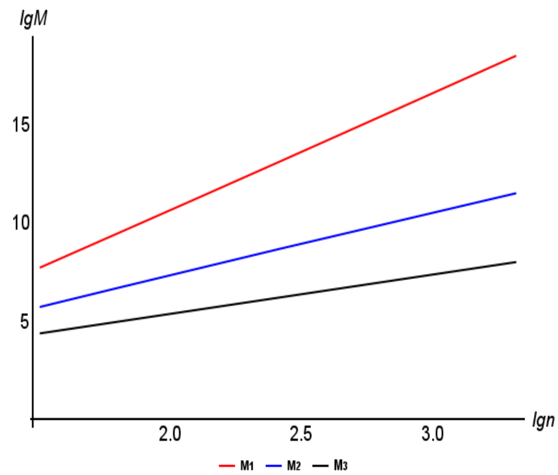

Рис. 4. Вычислительная сложность алгоритмов точного решения задачи (2)

В табл. 2, 3 M – число вычислительных операций, $\widetilde{M}$ – среднее число вычислительных операций. Все доверительные интервалы числа вычислительных операций для алгоритма спуска по узловым прямым лежат левее интервалов симплекс-алгоритма. Кроме того, симплекс-алгоритм не всегда находит точное решение задачи (2).

Проведен сравнительный анализ по точности и быстродействию алгоритма спуска по узловым прямым с приближенными алгоритмами на основе вариационно-взвешенных квадратических приближений и численных методов спуска нулевого порядка. Результаты статистических испытаний для модели (1) приведены в табл. 4, где случайные ошибки ε имеют распределение (8) и $\sigma = 1;\ a_H = 0;\ \gamma_H = 1; \gamma = 0.1;$ $N = 1000 -$ число испытаний; $m = 5$; $\Delta_1 = \Delta_2 = \cdots = \Delta_m = 1 -$ начальные величины шагов по координатным направлениям; $\mu = 10^{-6} -$ число для остановки алгоритма; $\lambda = 1.5 -$ ускоряющий множитель; $\alpha = 2 -$ коэффициент уменьшения шага; $\tau = 1.618 -$ коэффициент расширения; $\beta = 0.618 -$ коэффициент сжатия; $M = 4m$ – максимальное число неудачно выполненных испытаний на текущей итерации, $t_0 = 1\ -$ начальная величина шага, $T = 10^{-6}$ –минимальная величина шага.

Таблица 2. Результаты решений задачи (2) с помощью симплекс-алгоритма

n	Среднее число итераций	99%-й доверительный интервал числа итераций		$\lg\widetilde{M}$	99%-й доверительный интервал $\lg M$		Число расхождений от точного решения, %
		Левая граница	Правая граница		Левая граница	Правая граница	
32	20.72	20.25	21.18	**4.390**	**4.380**	**4.399**	***5.5***
64	40.26	39.45	41.05	**5.250**	**5.241**	**5.258**	***7.9***
128	75.67	74.18	77.15	**6.110**	**6.101**	**6.118**	***8.4***
256	143.34	140.65	146.02	**6.981**	**6.973**	**6.989**	***9.0***
512	343.14	326.40	359.00	**7.958**	**7.937**	**7.978**	***10.1***

Таблица 3. Результаты решений задачи (2) с помощью спуска по узловым прямым

n	Среднее число рассмотренных узловых точек	99%-й доверительный интервал рассмотренных узловых точек		$\lg\widetilde{M}$	99%-й доверительный интервал $\lg M$		Число расхождений от точного решения, %
		Левая граница	Правая граница		Левая граница	Правая граница	
32	134.74	131.70	137.78	**4.334**	**4.324**	**4.343**	***0***
64	241.00	235.70	246.26	**4.887**	**4.877**	**4.897**	***0***
128	431.86	422.00	441.73	**5.442**	**5.431**	**5.451**	***0***
256	784.07	765.48	802.65	**6.002**	**5.991**	**6.012**	***0***
512	1858.45	1766.14	1951.92	**6.677**	**6.655**	**6.699**	***0***

Таблица 4. Среднее квадратическое отклонение выборочных оценок параметров модели (1), найденной приближенным методом, относительно точного решения задачи (2), m = 5

Алгоритм вариационно-взвешенных квадратических приближений для n = 64			Методы поиска безусловного экстремума нулевого порядка		
			n	Метод конфигураций	Адаптивный метод случайного поиска
δ –вычислительная точность	$s_v(n,m)$	$t(n,m) = \dfrac{t_v(n,m)}{t_u(n,m)}$		$s_p(n,m)$	$s_r(n,m)$
10^{-1}	0.632	0.30	32	41.32	38.27
10^{-2}	0.352	0.72	64	44.00	41.21
10^{-3}	0.192	1.56	128	46.91	42.89
10^{-4}	0.127	3.0	256	47.96	44.57
10^{-5}	0.064	4.4	512	51.38	48.36
10^{-6}	0.041	5.6	1024	53.11	50.22

В табл. 4 обозначено: $s_v(n,m)$; $s_p(n,m)$ и $s_r(n,m)$ – среднее квадратическое отклонение вектора $\widehat{\mathbf{a}_v}$; $\widehat{\mathbf{a}_p}$ и $\widehat{\mathbf{a}_r}$ выборочных оценок параметров модели множественной линейной регрессии относительно вектора $\mathbf{a}^*$ точного решения задачи (2) соответственно для алгоритма Вейсфельда; метода конфигураций и адаптивного метода случайного поиска. $t_v(n,m)$ – среднее время вычислений для алгоритма Вейсфельда, $t_u(n,m)$– среднее время вычислений для алгоритма спуска по узловым прямым с использованием разреженных матриц и с учетом направления.

По результатам анализа сделан вывод, что предложенный алгоритм спуска по узловым прямым с использованием разреженных матриц и с учетом направления спуска, значительно выигрывает по сравнению с известными точными и приближенными методами и может эффективно использоваться на практике.

В **третьем параграфе** проводится сравнительный анализ алгоритма спуска по узловым прямым с известными точными (табл. 5, рис. 5) и приближенными методами реализации обобщенного метода наименьших модулей (табл. 6).

Таблица 5. Вычислительная сложность алгоритмов точного решения задачи (3)

Алгоритм полного перебора	Алгоритм спуска по узловым прямым для ОМНМ
$M_1 = O\left(C_n^m \cdot (m^3 + m \cdot n)\right)$	$M_2 = O\left(n^2 \cdot m^3 \cdot \ln n + C_{\alpha \cdot n}^m \cdot (m^3 + m \cdot n)\right)$

Результаты статистических испытаний показали, что с увеличением объема выборки, оптимальное значение α уменьшается обратно пропорционально. В случае, когда ошибки имеют распределение вида (8) $\alpha = 45/n$. Проведено сравнение алгоритма спуска по узловым прямым для ОМНМ и алгоритма модифицированного ОМНМ. Результаты статистических испытаний для задачи (2), где ошибки имеют распределение вида (8) и $\sigma = 1$; $a_H = 0$; $\gamma_H = 1$; $q = 2$ – количество подвыборок; $m = 4$ – размерность модели; $n = 100$ – объем выборки; $N = 1000$ – число испытаний, приведены в табл. 6.

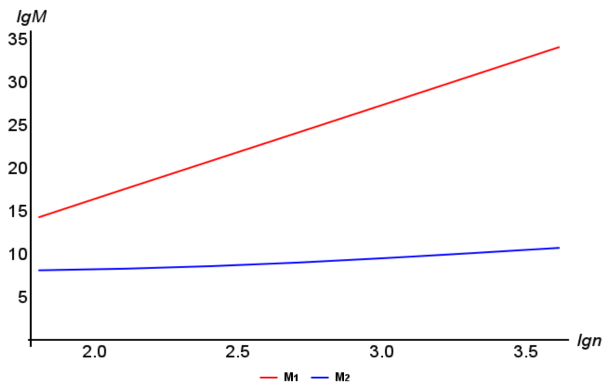

Рис. 5. Графики функции вычислительных сложностей алгоритмов точного решения задачи (3)

Таблица 6. Результаты сравнения алгоритмов модифицированного ОМНМ и спуска по узловым прямым для ОМНМ

γ	$p'(n,m)$	$t(n,m)/t'(n,m)$	$p''(n,m)$	$t(n,m)/t''(n,m)$
0.05	100	49.3	22.3	20.6
0.1	100	49.3	22	20.6
0.2	99.9	49.3	21.5	20.6

Здесь обозначено: $p'(n,m)$ – процент совпадения вектора $\widehat{\mathbf{a}'}$ выборочных оценок параметров модели множественной линейной регрессии относительно вектора $\mathbf{a}^*$ точного решения задачи (3) для алгоритма спуска по узловым прямым, $p''(n,m)$ – процент совпадения вектора $\widehat{\mathbf{a}''}$ выборочных оценок параметров модели множественной линейной регрессии относительно вектора $\mathbf{a}^*$ точного решения задачи (3) для алгоритма модифицированного ОМНМ, $t'(n,m)$ – среднее время вычислений для алгоритма спуска по узловым прямым, $t''(n,m)$ – среднее время вычислений для алгоритма модифицированного ОМНМ, $t(n,m)$ – время вычислений для алгоритма полного перебора.

По результатам анализа сделан вывод, что разработанный новый ОМНМ-алгоритм по быстродействию превосходит известные решения.

В **четвертой главе** рассмотрены примеры реализации и практического использования разработанного программного комплекса.

В **первом параграфе** на модельных примерах выполнено сравнение разработанных алгоритмов на основе спуска по узловым прямым с известными алгоритмами.

Во **втором параграфе** приведены четыре примера применения разработанного программного комплекса при решении практически важных задач в технике и экономике. В первых двух примерах был проведен сравнительный анализ результатов регрессионного моделирования, выполненный с помощью предложенного алгоритмов спуска по узловым прямым и известных алгоритмов МНМ-оценивания. Рассмотрены задачи моделирования условного среднего экономического ущерба муниципальных образований Свердловской области от пожаров и прогнозирования производительности центрального процессора. В третьем примере описаны результаты решения задачи оптимизации периода эксплуатации высоконагруженной техники на основе анализа средних удельных затрат. Для устойчивости к выбросам оценки средних ежемесячных затрат на эксплуатацию автомобиля вычислялись МНМ. В последнем примере рассмотрена задача оптимизация численности плательщиков страховых взносов в пенсионную систему за счет легализации неформальной занятости в регионах. При решении оптимизационной задачи был использован алгоритм ОМНМ-оценивания.

В **Заключении** приведены итоги проведенного исследования, рекомендации и перспективы дальнейшей разработки темы.

Положения, выносимые на защиту

В части *«Разработка новых математических методов моделирования объектов и явлений»:*

1. Разработан новый подход к математическому моделированию многомерных линейных регрессионных зависимостей в условиях стохастической неоднородности методом наименьших модулей, основанный на спуске по узловым прямым.

В части *«Развитие качественных и приближенных аналитических методов исследования математических моделей»:*

1. Аналитически и с помощью метода статистических испытаний исследованы методы моделирования многомерных линейных регрессионных зависимостей в условиях стохастической неоднородности экспериментальных данных; установлена закономерность, состоящая в вычислительно эффективном нахождении параметров математических моделей с помощью спуска по узловым прямым.

2. Повышение вычислительной эффективности моделирования линейных зависимостей на основе метода обобщенных наименьших модулей с помощью спуска по узловым прямым за счет установленной закономерности сокращения числа рассматриваемых возможных решений при использовании предложенного подхода.

3. Установлены классы многомерных нелинейных регрессионных зависимостей, для которых можно применять обобщенный метод наименьших модулей для вычислительно эффективного моделирования.

В части *«Разработка, обоснование и тестирование эффективных вычислительных методов с применением современных компьютерных технологий»:*

1. Разработаны вычислительно эффективные алгоритмы оценивания параметров линейных регрессионных моделей методом наименьших модулей на основе спуска по узловым прямым.

2. Установлена сходимость разработанных алгоритмов оценивания параметров многомерных линейных регрессионных моделей методом наименьших модулей к точному решению за конечное число шагов.

3. Разработан вычислительный алгоритм оценивания параметров линейных моделей методом обобщенных наименьших модулей на основе спуска по узловым прямым.

4. Выполнен анализ вычислительной трудоемкости предложенных алгоритмов, основанный на сочетании математических методов матричной алгебры и комбинаторики с современными технологиями математического моделирования, вычислительного эксперимента и статистических испытаний.

В части *«Реализация эффективных численных методов и алгоритмов в виде комплексов проблемно-ориентированных программ для проведения вычислительного эксперимента»:*

1. Разработан комплекс проблемно-ориентированных программ для проведения вычислительных экспериментов с целью исследования эффективности предложенных алгоритмов оценивания параметров многомерных линейных моделей, написанный на языке программирования R.

2. Поведены вычислительные эксперименты, подтверждающие эффективность предложенных алгоритмов по сравнению с известными решениями и адекватность проведенного моделирования.

ПУБЛИКАЦИИ АВТОРА ПО ТЕМЕ ДИССЕРТАЦИИ

Публикации в изданиях, рекомендованных ВАК

1. Azaryan A.A. Analysis of algorithms for stable estimation of coefficients of multiple linear regression models // Journal of Computational and Engineering Mathematics. 2018. Vol. 5. № 3. P. 17-23; 0,45 п.л. (MathSciNet, ZentralblattMATH).

2. Тырсин А.Н., Азарян А.А. Точное оценивание линейных регрессионных моделей методом наименьших модулей на основе спуска по узловым прямым // Вестник Южно-Уральского государственного университета. Серия «Математика. Механика. Физика». 2018. № 2. С. 47–56.; 0,65 п.л. / 0,4 п.л. (MathSciNet, ZentralblattMATH).

3. Тырсин А.Н., Азарян А.А. Методы устойчивого построения линейных моделей на основе спуска по узловым прямым // Модели, системы, сети в экономике, технике, природе и обществе. 2018. № 1(25). С. 188–202; 1,3 п.л. / 0,8 п.л.

4. Тырсин А.Н., Азарян А.А. Оптимизация периода эксплуатации высоконагруженной техники на основе анализа средних удельных затрат // Известия вузов. Горный журнал. 2017. № 5. С. 4–8; 0,3 п.л. / 0,2 п.л. (GeoRef, ChemicalAbstracts).

5. Тырсин А.Н., Азарян А.А. Точные алгоритмы реализации метода наименьших модулей на основе спуска по узловым прямым // Вестник Бурятского государственного университета. Математика, информатика. 2017. № 4. С. 21–32; 0,75 п.л. / 0,5 п.л.

Свидетельства о государственной регистрации программ для ЭВМ

6. Тырсин А.Н., Азарян А.А. Программа реализации метода наименьших модулей на основе спуска по узловым прямым: свидетельство № 2018610336; заявл. 13.11.2017; зарегистр. 10.01.2018; Реестр программ для ЭВМ.

7. Тырсин, А.Н., Чистова Е.В., Куклин А.А., Азарян А.А. Оптимизация численности плательщиков страховых взносов в пенсионную систему за счет легализации неформальной занятости в регионах: свидетельство № 2018610916; заявл. 22.11.2017; зарегистр. 19.01.2018; Реестр программ для ЭВМ.
8. Тырсин А.Н., Азарян А.А. Оценка линейных моделей методом обобщенных наименьших модулей на основе спуска по узловым прямым: свидетельство № 2018614491; заявл. 27.02.2018; зарегистр. 06.04.2018; Реестр программ для ЭВМ.

Монографии

9. «Лукавые» данные и реальная динамика социально-экономического развития субъектов РФ / Куклин А.А., Чичканов В.П., Никулина Н.Л., Чистова Е.В., Берсенев В.Л., Печеркина М.С., Васильева А.В., Наслунга К.С., Шипицына С.Е., Коробков И.В., Тырсин А.Н., Найденов А.С., Пыхов П.А., Яндыганов П.Я., Азарян А.А., Сурина А.А.; под ред. А.А. Куклина и В.П. Чичканова. Екатеринбург: Институт экономики УрО РАН, 2017. 364 с. (§§ 7.1, 7.2. Моделирование сбалансированности пенсионной системы России. С. 182–199); 1,15 п.л. / 0,4 п.л.

Другие публикации

10. Чистова Е.В., Тырсин А.Н., Азарян А.А. Легализация неформальной занятости как резерв повышения поступлений в пенсионную систему России // Пространственная экономика. 2017. № 4. С. 130–147; 1,15 п.л. / 0,4 п.л.
11. Азарян А.А., Тырсин А.Н. Эффективные алгоритмы оценивания линейных регрессионных моделей на основе метода наименьших модулей // Анализ, моделирование, управление, развитие социально-экономических систем: сборник научных трудов XI Международной школы-симпозиума, Симферополь-Судак, 14-27 сентября 2017 г. – Симферополь: ИП Корниенко А.А., 2017. С. 11–16; 0,4 п.л. / 0,25 п.л.
12. Азарян А.А., Тырсин А.Н. Повышение быстродействия точного алгоритма реализации метода наименьших модулей при оценивании параметров линейных регрессионных моделей // Труды второй научно-технической конференции молодых ученых Уральского энергетического института, Екатеринбург, 15-19 мая 2017. – Екатеринбург: УрФУ, 2017. С. 385–387. URL: http://elar.urfu.ru/handle/10995/55268; 0,2 п.л. / 0,15 п.л.
13. Тырсин А.Н., Азарян А.А., Корчёмкина Л.Н. Робастное оценивание стохастических моделей временных рядов в задачах диагностики // Современные тенденции развития науки и технологий. 2016. № 11-2. С.132–137; 0,4 п.л. / 0,2 п.л.
14. Тырсин А.Н., Азарян А.А. Об одном алгоритме реализации обобщенного метода наименьших модулей // Обозрение прикладной и промышленной математики. 2017. Т. 24, В. 4. С. 375–376; 0,15 п.л. / 0,1 п.л.
15. Азарян А.А. Комплекс проблемно-ориентированных программ для реализации, оценивания и исследования алгоритмов устойчивого построения линейных моделей // Проблемы внедрения результатов инновационных разработок: Сборник статей по итогам международной научно-практической конференции (Самара, 22 июня 2018 г.). Ч.2. – Стерлитамак: АМИ, 2018. С. 27–36; 0,65 п.л.
16. Тырсин А.Н., Азарян А.А. Оценивание нелинейных регрессионных зависимостей на основе обобщенного метода наименьших модулей // Материалы XIX Всероссийского Симпозиума по прикладной и промышленной математике (весенняя сессия), Санкт-Петербург, п. Репино, 21-27 апреля 2018 г. URL: http://www.tvp.ru/conferen/vsppmXIX/repso044.pdf; 0,15 п.л. / 0,1 п.л.
17. Азарян А.А., Тырсин А.Н. Быстрый алгоритм оценивания линейных регрессионных зависимостей на основе обобщенного метода наименьших модулей // Актуальные направления фундаментальных и прикладных исследований. Т. 2: Материалы XV международной научно-практической конференции, 9-10 апреля 2018 г. – North Charleston, USA, 2018. С. 68–73; 0,4 п.л. / 0,3 п.л.
18. Азарян А.А. Тырсин А.Н. Быстрые алгоритмы устойчивого оценивания линейных регрессионных моделей / А.А. Азарян, А.Н. Тырсин // Новые информационные технологии в исследовании сложных структур: Материалы 12-й международной конференции, Алтайский край, пос. Катунь, 4-8 июня 2018 г. – Томск: ТГУ, 2018. С. 103–104; 0,15 п.л. / 0,1 п.л.