На правах рукописи

Крутиков Владимир Николаевич

ЧИСЛЕННЫЕ МЕТОДЫ БЕЗУСЛОВНОЙ ОПТИМИЗАЦИИ С ИТЕРАТИВНЫМ ОБУЧЕНИЕМ И ИХ ПРИМЕНЕНИЕ

05.13.18 – *Математическое моделирование, численные методы и комплексы программ*

Автореферат

диссертации на соискание ученой степени

доктора технических наук

Томск 2005

Работа выполнена в ГОУ ВПО «Кемеровский государственный университет»

Официальные оппоненты: доктор технических наук, профессор
Смагин Валерий Иванович

доктор технических наук, профессор
Антамошкин Александр Николаевич;

доктор физико-математических наук,
профессор Либерзон Марк Романович.

Ведущая организация: Институт вычислительного моделирования
СО РАН (г. Красноярск).

Защита состоится 14 апреля 2005 г. в14 час. 30 мин. на заседании диссертационного совета Д 212.267.08 при Томском государственном университете (634050, г. Томск, пр. Ленина, 36)

С диссертацией можно ознакомиться в Научной библиотеке Томского государственного университета

Ваш отзыв на автореферат (2 экз.), заверенный печатью, просьба направлять по адресу: 634050, г. Томск, пр. Ленина, 36, Томский государственный университет, ученому секретарю ТГУ.

Автореферат разослан «9» марта 2005 г.

Ученый секретарь диссертационного
совета, доктор технических наук А. В. Скворцов

Общая характеристика работы

Актуальность работы. Релаксационные методы безусловной оптимизации (РМО) относятся к средствам численного моделирования и используются при изучении явлений действительности. На их основе конструируются численные методы нелинейного программирования (НЛП). Распространенным приложением РМО являются оптимизационные задачи оценивания параметров математических моделей в схемах структурно-параметрического моделирования, т. е. при конструировании модели, когда требуется определить такие ее структуру и параметры, которые обеспечивали бы наилучшее согласование с реальностью, к числу которых относятся рассматриваемые в работе актуальные задачи: прогнозирование свойств терморегулирующих покрытий на основе результатов ускоренных испытаний; определение нестационарных законов горения пороха на основе манометрических испытаний; выбор модели минимальной сложности при аппроксимации нейросетями; экономико-математический анализ деятельности предприятий; оценивание параметров и структуры поверхности при аппроксимации горно-геологических объектов. Задачи структурно-параметрического моделирования сопряжены с многократным решением задач оптимизации с различными типами и степенями сложности, что предполагает наличие спектра надежных, быстросходящихся и универсальных РМО.

Эффективные РМО основываются на некоторой конечной многошаговой стратегии (*глобальной стратегии*) минимизации определенного класса функций (например квадратичных) и содержат параметры аппроксимативной модели функции и средства ее обучения. Несоответствие классов моделей метода и функции, невозможность выполнения точно операций алгоритма (например одномерного спуска), разнесённость в пространстве обучающей информации и изменчивость оцениваемых характеристик функции приводят к разрушению *глобальных свойств сходимости* конечных многошаговых стратегий обучения, что определяет актуальность целенаправленного создания *устойчивых алгоритмов обучения*, отдельные шаги которых одновременно реализуют и некоторую *итеративную стратегию (локальную)* улучшения качества аппроксимативной модели метода, т. е. обладают *локальными свойствами сходимости*.

В работах Я. З. Цыпкина применительно к различным областям знания развиты концепция и единообразный подход к решению проблем построения обучающихся систем и принципы построения алгоритмов обучения как алгоритмов оптимизации заданного показателя качества. Как выясняется, этот подход эффективен и при создании алгоритмов обучения в РМО, что определяет актуальность его использования при конструировании обучающихся методов оптимизации. В диссертации отмеченный подход используется как унифицированное средство построения и анализа устойчивых алгоритмов обучения – *итеративных алгоритмов изменения метрики пространства* в схемах РМО, обладающих локальными и глобальными свойствами сходимости.

Развитие области численных методов безусловной оптимизации связано с именами отечественных ученых Л. В. Канторовича, Н. Н. Моисеева, В. Ф. Демьянова, Е. Г. Гольштейна, Ю. Г. Евтушенко, Ю. М. Ермольева, В. С. Михалевича, В.

Г. Карманова, Б. Т. Поляка, Б. Н. Пшеничного, Ю. М. Данилина, А. С. Немировского, Д. Б. Юдина, Ф. П. Васильева, Н. З. Шора, Ю. Е. Нестерова и многих других авторов. РМО, в зависимости от используемой информации о функции, можно подразделить на градиентные, субградиентные и прямого поиска.

Прямые методы, использующие только вычисления функции, развивались в работах В. Г. Карманова, Б. Н. Пшеничного, О. П. Бурдакова, М. М. Потапова, Л. А. Растригина, Ф. Л. Черноусько, Пауэлла, Брента, Бродли и др. Одно из слабых мест релаксационных методов прямого поиска – это отсутствие эффективных методов минимизации сложных негладких функций. Например, метод деформируемого многогранника (МД), относимый к числу методов негладкой оптимизации, согласно исследованию С. И. Нестеровой и В. А. Скокова, среди множества негладких задач в состоянии решить лишь часть размерности 2. Более предпочтительным для этих целей является метод случайного поиска (СП), поскольку он сходится линейно (В. Г. Карманов), а адаптация шага Л. А. Расстригина, Г. С. Тарасенко обеспечивает его сходимость на негладких функциях, но он имеет медленную скорость сходимости на овражных функциях. Отмеченные обстоятельства определяют актуальность исследований диссертации по созданию алгоритмов обучения метрики в методах СП для ускорения их сходимости и расширения области решаемых ими гладких и негладких задач.

Прямые методы, основанные на минимизации вдоль векторов линейно-независимой системы, с одновременным преобразованием их в сопряженные, изучались Б. Н. Пшеничным, О. П. Бурдаковым, Пауэллом, Брентом, Бродли и др. В методах Пшеничного и Пшеничного-Редковского по значениям функции вычисляются необходимые вторые производные для построения системы сопряженных векторов и осуществляется спуск вдоль них. В этих алгоритмах не используется возможность совмещения затрат вычислений функции на аппроксимацию вторых производных и одномерный спуск и велики затраты вычислений функции на итерации ($\sim n^2$). К недостаткам методов Пауэлла и Брента относится необходимость точного одномерного спуска и постоянно возникающая линейная зависимость векторов спуска, что снижает их эффективность с ростом размерности. Метод Бродли – один из методов, в котором используется устойчивая схема обучения и не требуется точный одномерный спуск. Но его схема обучения не позволяет построить за конечное число шагов систему сопряженных векторов, что снижает эффективность метода с ростом размерности. В этой связи в работе для построения эффективного алгоритма минимизации прямого поиска решается задача создания алгоритмов обучения метрики, обладающих локальными и глобальными свойствами сходимости и дается исследование скорости сходимости на сильновыпуклых функциях класса методов минимизации вдоль векторов линейно-независимой системы в зависимости от ее свойств.

Методы сопряженных градиентов и квазиньютоновские методы (КНМ) изучались в работах Ю. Г. Евтушенко, В. Г. Карманова, Б. Т. Поляка, Б. Н. Пшеничного, Ю. М. Данилина, О. П. Бурдакова, Хестенса, Штифеля, Девидона, Флетчера, Ривза, Пауэлла, Бройдена, Денниса, Шнабеля, Морэ и многих других. В результате исследований выявлено значительное число формул пересчета матриц, проведен экспериментальный и теоретический анализ скорости сходимости существующих методов. Экспериментально установлено, но не объяснено, что в КНМ наилучшие результаты имеет формула пересчета BFGS. Оказывается, что КНМ без точного одномерного спуска более эффективны, нежели с одномерным спуском. Попытки повысить эффективность КНМ при неточном одномерном спуске за счет использования явных или неявных схем последовательного построения системы сопряженных векторов, не увенчались успехом в силу высокой степени линейной зависимости последовательностей векторов спуска и изменчивости свойств функции. В этой связи является актуальным объяснение установленных экспериментально фактов и создание устойчивых к линейной зависимости векторов спуска алгоритмов обучения метрики, обладающих локальными и глобальными свойствами сходимости при неточном одномерном спуске. Исследования диссертации направлены на изучение и разрешение этих проблем с привлечением формализма теории обучения.

Развитие численных методов негладкой оптимизации связано с именами ученых И. И. Астафьева, В. П. Булатова, Ф. П. Васильева, Е. Г. Гольштейна, А. М. Гупала, В. Ф. Демьянова, Ю. М. Ермольева, И. И. Еремина, Ю. А. Левина, В. Н. Малоземова, А. С. Немировского, Е. А. Нурминского, Ю. Е. Нестерова, Б. Т. Поляка, Н. З. Шора, Д. Б. Юдина, Лемарешаля, Вульфа, Келли и многих других. Результаты численного исследования С. И. Нестеровой и В. А. Скокова известных методов негладкой оптимизации: метода эллипсоидов (А. С. Немировский, Д. Б. Юдин.), r-алгоритма (Н. З. Шор), метода ортогонального спуска (В. А. Скоков, М. Б. Щепакин) и метода уровней (МУ – Е. Г. Гольштейн, А. С. Немировский, Ю. Е. Нестеров), с одной стороны, демонстрируют высокую эффективность метода МУ по критерию числа вычислений функции и субградиента, а с другой – выявляют необходимость создания эффективных по критерию времени счета, надежных методов негладкой оптимизации. Сегодня существует проблема создания универсальных, надежных методов негладкой оптимизации, эффективных при решении гладких и негладких в том числе и невыпуклых задач, конкурирующих по свойствам скорости сходимости с методами, основанными на квадратичной либо кусочно-линейной моделях функции, на классах функций, промежуточных отмеченным базовым классам.

Одна из возможностей создания подобных методов заключается в совершенствовании класса релаксационных методов ε-субградиентного типа (РСМ), к числу которых относятся методы Лемарешаля и Вульфа. Методы этого класса пригодны для решения невыпуклых гладких и негладких задач оптимизации, генерируют на квадратичных функциях сопряженные направления и находят минимум за конечное число итераций, но существенно уступают в скорости сходимости r-алгоритму. Происхождение r-алгоритма и его определяющие соотношения обучения не определены, а его существующие реализации на сложных задачах имеют либо низкую скорость сходимости, либо не позволяют их решить. В этой связи, для разработки надежных и эффективных РСМ, необходимо создание теории, назначение которой: определить удобные определяющие соотношения обучения, разработать эффективные алгоритмы обучения и методы оптимизации на их основе, объяснить происхождение r-алгоритма, выявить его обучающие соотношения, ускоряющие свойства при конечных параметрах растяжения пространства и создать эффективные способы реализации РСМ. Исследования диссертации направлены на решение поставленных актуальных задач с привлечением формализма теории обучения.

Одним из назначений создаваемых алгоритмов является их применение для реализации схем структурно-параметрического моделирования при решении отмеченных выше актуальных прикладных задач. Созданию конкретной системы построения модели предшествуют предварительные этапы формирования множества моделей, оценки качества моделей, подбора алгоритмов оценивания их параметров и т. д. Для их выполнения необходим комплекс надежных, быстросходящихся алгоритмов оптимизации, что определяет актуальность создания программного комплекса РМО, численного исследования и сравнения с известными методами его алгоритмов и определения областей их приложений.

Таким образом, разработка эффективных релаксационных методов безусловной оптимизации, основанных на итеративных алгоритмах обучения, создание комплекса их программ, определение областей их приложений и применение в вычислительных схемах структурно-параметрического моделирования является актуальной научной проблемой.

Диссертационная работа выполнена в соответствии: с планами НИР КемГУ по кафедре математической кибернетики 1991-2004 гг.; с грантом Российского фонда фундаментальных исследований (код проекта 04-03-33121); с программой «Университеты России» (код проекта УР.04.01.044); договорной тематикой с угольными компаниями Кузбасса, НТЦ «Кузбассуглетехнология» и с корпорацией «Кузбассинвестуголь»; с Комплексной региональной программой научных исследований и внедрения совместных работ СО АН СССР и Министерства угольной промышленности СССР, «Программа “Сибирь”», раздел «Уголь Кузбасса», и приказами Министерства угольной промышленности СССР от 16.11.79, и №315 от 24.06.80.

Цель работы состоит в разработке и исследовании эффективных релаксационных методов безусловной оптимизации, основанных на итеративных алгоритмах обучения, создании комплекса их программ и применении его алгоритмов в вычислительных схемах структурно-параметрического моделирования.

Идея работы заключается в использовании теории адаптации и обучения для создания устойчивых к возмущениям и смене ситуаций итеративных алгоритмов обучения в релаксационных методах безусловной оптимизации.

Задачи исследований:

1. Дать анализ условий функционирования алгоритмов обучения в методах безусловной оптимизации и определить для подобных условий алгоритмы обучения.
2. Разработать алгоритмы обучения метрики в методах прямого поиска и эффективные методы прямого поиска с изменением метрики пространства на их основе.

3. Разработать устойчивые к линейной зависимости векторов спуска алгоритмы обучения метрики в квазиньютоновских методах.
4. Разработать эффективные релаксационные методы сопряженных субградиентов ε-субградиентного типа с растяжением пространства.
5. Разработать программный комплекс релаксационных методов безусловной оптимизации, основанных на итеративных алгоритмах обучения, для решения прикладных задач структурно-параметрического моделирования.

Методы исследований: методы теории адаптации и обучения для создания алгоритмов обучения; методы линейной алгебры, статистики, нелинейного программирования, функционального анализа для разработки и теоретического анализа алгоритмов оптимизации; тестирование и сравнение алгоритмов для численного анализа эффективности новых методов минимизации; методы анализа данных и закономерности изучаемых явлений для оценивания параметров и структуры математических моделей в разрабатываемых системах математического моделирования.

Научная и практическая новизна работы заключается в том, что в ней:

1. Впервые в методах случайного поиска (СП) использовано обучение метрики и созданы эффективные методы СП с варьированием метрики для решения сложных гладких и негладких задач безусловной оптимизации.
2. Впервые для класса методов минимизации вдоль векторов линейно независимой системы получены оценки скорости сходимости на сильновыпуклых функциях и разработано семейство конечно-сходящихся итеративных алгоритмов обучения с ограниченной релаксацией для построения сопряженных направлений, сочетающих локальные и глобальные свойства сходимости процесса обучения. На этой основе создан эффективный метод минимизации без вычисления производных (МСН).
3. Впервые формализм теории адаптации и её алгоритмы минимизации показателя качества обучения (одношаговый и двухшаговый) применены в субградиентных и квазиньютоновских методах для разработки новых алгоритмов обучения, обладающих локальными и глобальными свойствами сходимости. Одношаговый алгоритм обучения в квазиньютоновских методах приводит к известным формулам преобразования матриц, а двухшаговый – обеспечивает конечную сходимость метода при неточном одномерном спуске. Одношаговый алгоритм обучения позволяет разработать новый метод решения неравенств и новый субградиентный метод (РСМК) на его основе.
4. Разработаны новые алгоритмы обучения с растяжением пространства: алгоритм для решения множества равенств; алгоритм для решения неравенств на отделимых множествах. На основе алгоритма решения неравенств разработан новый релаксационный субградиентный метод с растяжением пространства в направлении субградиента (МРП) для решения сложных задач оптимизации.
5. Разработан новый матричный алгоритм обучения для решения неравенств и построено одноранговое семейство релаксационных субградиентных методов на его основе, в которое, как частный случай, входит r-алгоритм и обобщение

метода Вульфа.

6. Впервые для долгосрочного прогнозирования изменений коэффициента поглощения солнечной радиации терморегулирующих покрытий космических летательных аппаратов по данным лабораторных испытаний разработаны: комплекс математических моделей оптической деградации терморегулирующих покрытий и схемы их построения; схема анализа построенного множества моделей и выбора модели для прогноза.
7. Впервые разработан комплекс программ релаксационных методов (ПКРМО), значительная часть алгоритмов которого предназначена для решения сложных негладких задач оптимизации, в том числе и невыпуклых, что существенно расширяет возможности решения задач структурно-параметрического моделирования.

Основные положения, выносимые на защиту.

1. Итеративные алгоритмы обучения метрики и численные методы случайного поиска с изменением метрики на их основе для решения сложных гладких и негладких задач безусловной оптимизации.
2. Скорость сходимости класса методов минимизации вдоль векторов линейно независимой системы, конечно-сходящиеся итеративные алгоритмы обучения с ограниченной релаксацией для построения сопряженных направлений и численный метод прямого поиска с изменением метрики на этой основе.
3. Формализм и алгоритмы минимизации показателя качества теории обучения позволяют разработать эффективные алгоритмы обучения в методах сопряженных субградиентов и квазиньютоновских методах: одношаговый алгоритм обучения в квазиньютоновских методах, приводящий к известным способам преобразования матриц, двухшаговый алгоритм обучения матриц, обеспечивающий конечную сходимость метода; одношаговый алгоритм обучения для решения неравенств и субградиентный метод РСМК на его основе для решения негладких задач оптимизации высокой размерности.
4. Алгоритм обучения с растяжением пространства для решения равенств и неравенств, релаксационный субградиентный метод с растяжением пространства в направлении субградиента МРП на его основе для решения сложных гладких и негладких задач оптимизации, его свойства и реализация.
5. Матричный алгоритм обучения для решения неравенств и одноранговое семейство релаксационных субградиентных методов на его основе, в которое, как частный случай, входит r-алгоритм и обобщение метода Вульфа.
6. Комплекс математических моделей оптической деградации терморегулирующих покрытий, схемы их построения, анализа и выбора модели по данным лабораторных испытаний для долгосрочного прогнозирования изменений коэффициента поглощения солнечной радиации терморегулирующих покрытий космических летательных аппаратов в заданных условиях эксплуатации.
7. Комплекс программ разработанных методов прямого поиска, квазиньютоновских и субградиентных методов ПКРМО является надежным и эффективным средством решения сложных задач безусловной оптимизации и реализации

оптимизационного инструментария в задачах структурно-параметрического моделирования.

Достоверность научных положений и выводов подтверждается: строгими теоретическими оценками скорости сходимости созданных алгоритмов обучения; корректным теоретическим анализом сходимости, скорости сходимости и ускоряющих свойств, созданных численных методов оптимизации; результатами тестирования на общепринятых тестах с известными решениями; сравнением результатов тестирования новых алгоритмов с результатами для известных методов, полученных другими авторами; корректным использованием статистических методов оценивания качества получаемых прикладных моделей в задачах оценивания параметров и структуры математических моделей сложных явлений.

Личный вклад автора состоит: в разработке и обосновании методов СП с изменением метрики, метода минимизации без вычисления производных, единообразного подхода создания, обоснования и реализации новых релаксационных методов минимизации негладких функций и квазиньютоновских методов, алгоритмов решения множества неравенств и субградиентных методов минимизации; в разработке программного комплекса методов оптимизации с обучением и программных средств определения параметров и структуры математических моделей в задачах математического моделирования; в создании системы математического моделирования и долгосрочного прогнозирования изменений коэффициента поглощения солнечной радиации терморегулирующих покрытий космических летательных аппаратов.

Практическая значимость результатов работы. Разработанные релаксационные методы безусловной оптимизации можно использовать при решении сложных задач оптимизации, для реализации программных средств оптимизации в прикладных задачах математического моделирования, и, в частности, при долгосрочном прогнозирования изменений коэффициента поглощения солнечной радиации терморегулирующих покрытий, при определении нестационарных законов горения пороха на основе манометрических испытаний, при обучении нейронных сетей, при построении нейронной сети минимальной сложности методом негладкой регуляризации, при выявлении структуры поверхности по данным наблюдения.

Апробация работы. Основные положения диссертационной работы докладывались и были одобрены на VIII и XII Всероссийских семинарах «Нейроинформатика и ее приложения» (Красноярск, 2000, 2004 гг.); Четвертой Международной школе-семинаре «Внутрикамерные процессы, горение и газовая динамика дисперсных систем» (Санкт-Петербург, Россия, 2004г); Всероссийской конференции ”Наука и практика: диалоги нового века” (Анжеро-Судженск, 2003, 2002 гг.); Всероссийской научно-практической конференции ”Наука и образование” (Белово, 2001, 2002 гг.); Международной школе-семинаре по методам оптимизации и их приложениям (Иркутск: СЭИ, 1989, 1995, 1998 гг.); Международной конференции по математическому моделированию (Якутск, 1994г.); Десятом Всесоюзном симпозиуме “Системы программного обеспечения решения задач оптимального планирования” (Москва: АН СССР, ЦЭМИ, 1988г.); на I и II Всесоюзных семинарах «Информатика недр» (г. Кемерово, 1987 г., 1989 г.); VII Международном конгрессе по маркшейдерскому делу (г. Ленинград, 1988г.); VII региональной конференции по математике и механике (Томск, 1981 г.), Всесоюзном научно-техническом семинаре “Численные методы нелинейного программирования” (Харьков, 1979 г.); Всесоюзном совещании “Применение случайного поиска при решении прикладных задач” (Кемерово, 1979 г.); симпозиуме “Вероятностные вычислительные методы и средства” (Москва, 1978 г.); Всесоюзном семинаре “Случайный поиск и системы автоматизированного проектирования в электронике” (Цахкадзор, 1978 г.); IV Всесоюзном совещании по статистическим методам в теории управления (Фрунзе, 1978 г.); Всесоюзной школе-семинаре по оптимизации динамических систем (Минск, 1977 г.).

Публикации. По теме диссертации автором опубликовано 72 научных работы, список которых приведен в конце автореферата.

Структура работы. Диссертация состоит из введения, шести глав, заключения и приложения. Объем диссертации составляет 290 страниц машинописного текста, в том числе содержит 31 таблицу и 6 рисунков. Список литературы включает 231 наименование. Приложения изложены на 8 страницах.

Основное содержание работы

В первой главе рассматриваются типы сложности задач безусловной оптимизации, условия обучения и свойства схем обучения в известных релаксационных методах безусловной оптимизации (РМО) и алгоритмы теории обучения, соответствующие выявленным условиям обучения.

Сложность задачи безусловной оптимизации определяется размерностью, степенью вырожденности, доступной информацией о функции, степенью гладкости. Отмеченные характеристики задачи обусловливают тип метода для ее решения. Проведен качественный анализ схем обучения известных РМО. В результате выявлено, что обучение в известных РМО осуществляется по принципу удовлетворения последнего или нескольких последних обучающих соотношений.

В целях единообразного рассмотрения и анализа алгоритмов обучения, которые становятся неузнаваемыми при решении различных задач, приведены показатели качества и используемые в работе алгоритмы обучения, соответствующие отмеченному принципу. Рассмотрены итеративные алгоритмы оценивания параметров $c^* \in R^n$ линейной модели $y=(z,c)$, $z,c \in R^n$, $y \in R^1$ по данным $y_k \in R^1$, $z_k \in R^n$, $k = 0,1,2,...$, где $y_k = (c^*, z_k) + \xi_k$, ξ_k – ошибка. В работе используются алгоритмы обучения, на итерации которых решается задача минимизации показателя качества обучения на серии наблюдений:

$$J_N(c_k) = \frac{1}{N} \sum_{i=k-N+1}^{k} Q(z_i, c_k), \quad Q(z,c) = \frac{1}{2}(y-(c,z))^2. \tag{1.1}$$

Алгоритмы минимизации при N=1 и N=2 в (1.1) будем называть соответственно ***одношаговым*** (алгоритм Качмажа) и ***двухшаговым*** алгоритмами обучения. Для реализации обучения в субградиентных методах, в результате модификации итерационного метода наименьших квадратов, разработан одношаговый ***алгоритм***

обучения с растяжением пространства:

$$c_{k+1}=c_k+\frac{H_k z_{k+1}(y_{k+1}-(c_k,z_{k+1}))}{(H_k z_{k+1},z_{k+1})},\quad H_{k+1}=H_k-(1-\frac{1}{\alpha^2})\frac{H_k z_{k+1}z_{k+1}{}^T H_k{}^T}{(H_k z_{k+1}z_{k+1})}\quad (\alpha>1). \tag{1.2}$$

Отмеченные алгоритмы обеспечивают выполнение последнего обучающего соотношения $y_k=(z_k,c_{k+1})$. На их основе в диссертации строятся эффективные методы минимизации.

Во второй главе дается обоснование скорости ходимости класса методов минимизации вдоль векторов линейно независимой системы, разрабатываются и обосновываются сходящиеся локальные стратегии обучения метрики, реализующие глобальные стратегии, и эффективные релаксационные методы прямого поиска на их основе.

Обоснование скорости сходимости класса методов минимизации вдоль векторов линейно независимой системы $p_i\in R^n$, $i=1,2\ldots,n$:

$$x_i=x_{i-1}+\beta_i p_i,\ \beta_i=\arg\min_{\beta} f(x_{i-1}+\beta p_i),\quad i=1,2\ldots,n. \tag{2.1}$$

Условие 2.1. *Пусть $f(x)$, $x\in R^n$, сильновыпуклая функция с константой ρ, а ее градиент $f'(x)$ удовлетворяет условию Липшица с константой L.*

Обозначим P – матрицу, столбцы которой p_i, $i=1,\ldots,n$, μ и M – минимальное и максимальное собственные значения матрицы P^TP, x^* – точку минимума функции. В теоремах 2.1, 2.2 приведены оценки скорости сходимости процесса (2.1) в зависимости от свойств системы векторов.

Теорема 2.1. *Пусть $f(x)$ удовлетворяет условию 2.1, для каждого цикла процесса (2.1) $\|p_i\|=1$, $i=1,\ldots,n$, матрица P удовлетворяет ограничению $\mu\ge\mu_0>0$, а точка x^m получена применением m циклов процесса (2.1). Тогда для оценки скорости сходимости имеет место неравенство:*

$$f(x^m)-f(x^*)\le[f(x_0)-f(x^*)]\exp\left(-\frac{m\rho^2\mu_0^2}{2L^2n^3}\right). \tag{2.2}$$

Теорема 2.2. *Пусть $f(x)$ удовлетворяет условию 2.1, для каждого цикла процесса (2.1) $\|p_i\|=1$, $i=1,\ldots,n$, матрица P удовлетворяет ограничению $|\det(P)|\ge\delta$, а точка x^m получена применением m циклов процесса (2.1). Тогда для оценки скорости сходимости имеет место неравенство:*

$$f(x^m)-f(x^*)\le[f(x_0)-f(x^*)]\exp\left(-\frac{m\rho^2\delta^4}{2L^2n^{2n-1}}\right). \tag{2.3}$$

Алгоритмы случайного поиска. Изучается скорость сходимости методов случайного поиска $x_{k+1}=x_k+h_k\xi_k$ $(k=0,1,...)$ с фиксированным шагом h_k в точке x_k (без одномерного спуска). Доказано, что при условии 2.1 и специальном выборе шага в точке такие методы имеют скорость сходимости, эквивалентную скорости сходимости метода скорейшего спуска, что определяет возможность создания эффективных алгоритмов случайного поиска при наличии эффективных процедур адаптации шага. Показано, что известные алгоритмы адаптации шага являются следствием адаптивного подхода и разработана схема оптимизации их параметров. На этой основе разработан алгоритм СПО. Развитием процедуры адаптации шага является разработанный ***алгоритм с пространственной адаптации шага*** (СПМ) для минимизации «овражных» функций, где используются операторы растяжения пространства в заданном направлении.

Минимум квадратичной функция $f(x)=(x,Ax)/2 + (b,x) + c,$ где матрица A>0, т. е. симметричная и строго положительно определенная, $x,\ b\in R^n$, c – скаляр, может быть найден процессом (2.1) за цикл минимизации вдоль сопряженных относительно матрицы A векторов (А-ортогональных), т. е. при $P^TAP=I$. Для создания метода СП с варьированием метрики разработан процесс построения матрицы P, удовлетворяющей соотношению $P^T AP = I$:

$$P_{k+1} = P_k\left[I-(1-a_k^{-0,5})\xi_k\xi_k^T\right],\ \ k=0,1,2,..,\ \ a_k=\xi_k^T P_k^T AP_k\xi_k\,,\ \|\ \xi_k\ \|=1. \quad (2.4)$$

Здесь $\xi\in R^n$ – вектор, равномерно распределенный на единичной сфере (РРЕС); A, P, I – $n\times n$ матрицы, A>0. Обозначая через W_k k-ю реализацию случайной ортогональной системой векторов (СОСВ) и объединяя n шагов процесса (2.4) для ортогональных векторов ξ_{ik} из системы W_k, получаем новый процесс:

$$P_0 = I,\quad P_{k+1}=P_k\left[I-\sum_{i=1}^{n}(1-a_i^{-0,5})\xi_{ik}\xi_{ik}^T\right],\ a_{ik}=\xi_{ik}^T P_k^T AP_k\xi_{ik},\quad k=0,1,2,\ldots \quad (2.5)$$

Модифицируем процесс (2.5) следующим образом.

М1. Пусть в (2.5) вместо А используется A_{ik}, причем:

$$\|\ A-A_{ik}\ \|\le\Delta_k\ (k=0,1,2...,i=\overline{1,n}),\ ||\,Y\,||^2\ m\le y^T Ay\le||\,Y\,||^2\ M\ ,\ \sum_{k=0}^{\infty}\Delta_k<\infty.$$

М2. Положим $P_k=I$, если после (2.5): $|\det(P_k)|\le 1/(z\alpha M)^{n/2}=\delta_1$, или $\|P_k\|\ge(nz\alpha/m)^{0.5}=\delta_2\ (\alpha>1)$, где $z=2nM/m$.

М3. В (2.5) ограничим a_{ik} : $a_{ik}=\begin{cases}\varepsilon_1=\beta nM\delta_2^2, & при\quad a_{ik}>\varepsilon_1,\\ \varepsilon_2=m\delta_1^2/\beta\delta_2^{2(n-1)}, & при\ \ a_{ik}\le\varepsilon_2,\end{cases}\quad(\beta>1).$

В следующей теореме обосновывается сходимость алгоритма обучения метрики в методах случайного поиска.

Теорема 2.3. *Процесс (2.5), модифицированный в соответствии с М1-М3, задает последовательность* $\{P_k\}$ *такую, что последовательность матриц* $\{B_k=P_k^T AP_k\}$ *сходится к I п.н.*

Алгоритм случайного поиска с варьированием метрики (СПВМ0). Задаем $h_0>0,\ x_0\in R^n,\ 0\le\varepsilon<0.5,\ \delta_1,\delta_2,\varepsilon_1,\varepsilon_2>0$. Пусть $P_0=I,\ x_{k-1},h_{k-1},\ P_{k-1}$ уже получены, *k*-е приближение имеет следующий вид.

1. $h_{k1}:=h_{k-1},\ P_{k1}:=P_{k-1},\ x_{k1}:=x_{k-1}.$

Для $i=\overline{1,n}$ выполняем п.2-4.

2. $\alpha_{ki}=\arg\min\limits_{\alpha} f\left(x_{ki}-\alpha s_{ki}\right)$, где $s_{ki}=P_{ki}\xi_{ki}$, ξ_{ki} – вектор из W_k. $x_{ki+1}=x_{ki}+\alpha_{ki}s_{ki}$.

3. $a_{ki} = \left(\dfrac{f_{1i} - f_{2i}}{\gamma_{1i} - \gamma_{2i}} - \dfrac{f_{2i} - f_{3i}}{\gamma_{2i} - \gamma_{3i}} \right) / (\gamma_{2i} - \gamma_{3i})$, где $f_{ji} = f(x_{ki} + \gamma_{ji} s_{ki})$, $|\gamma_{ji}| < ch_{ki}$, $j = 1,2,3$,

ограничим a_{ki} по формуле из М3.

4. $P_{k,i+1} = P_{ki}\left[I - (1 - a_{ki}^{-0.5})\xi_{ki}\xi_{ki}^T\right]$, $h_{k,i+1} = h_{ki}\theta + (1-\theta)|\alpha_{ki}|$, $\theta \in [\varepsilon, 1-\varepsilon]$.

5. Положить $P_k = P_{k,n+1}$, $x_k = x_{k,n+1}$, $h_k = h_{k,n+1}$.

6. Если $|\det(P_k)| < \delta_1$ или $\| P_k \| > \delta_2$, то $P_k = I$.

7. Положить $k = k+1$, перейти к выполнению п.1.

Свойства сходимости алгоритма отражены в теоремах 2.4 и 2.5.

Теорема 2.4. *Пусть f (x) удовлетворяет условию 2.1. Тогда последовательность* x_k*, построенная алгоритмом СПВМ0, сходится к точке минимума x*, для оценки сходимости имеют место неравенства:*

$$f(x_k) - f(x^*) \le [f(x_0) - f(x^*)] \exp(-c_1 k),\ 1 > c_1 > 0.$$

Теорема 2.5. *Пусть f (x) удовлетворяет условию 2.1 и трижды непрерывно дифференцируема, собственные значения матрицы* $f''(x^*)$ *принадлежат отрезку [m, M]. Пусть значения* $\delta_1, \delta_2, \varepsilon_1, \varepsilon_2$ *выбраны, как в выражениях из М2-М3. Тогда при достаточно больших k* $x_k \to x_*$ *п.н. сверхлинейно.*

На основе СПВМ0 получены реализации СПВМ1, СПВМ2, СПВМ3. Вычисления в алгоритмах организованы так, что они работоспособны при минимизации кусочно-линейных функций. Результаты численного исследования свидетельствуют, что алгоритмы СПВМ1 и СПВМ2 превосходят по скорости методы Розенброка, Хука-Дживса, деформируемого многогранника и обычные методы СП, а методы СПВМ2, СПВМ3 эффективны при минимизации сложных негладких задач. Основное назначение разработанных методов СП – это решение задач минимизации, где нет возможности вычислить градиент или субградиент и нельзя применить для решения по условиям гладкости некоторый метод сопряженных направлений.

Метод минимизации без вычисления производных на основе рассредоточенной схемы А-ортогонализации (МСН). Матрицу *P* назовем А-нормированной, если $(p_i, Ap_i) = 1$, $i = 1,...,n$. Для А-нормированной матрицы *P* определим преобразование *V=W(i,j,P)* в виде:

$$\begin{aligned} &\tilde{v}_i = p_i - \omega_{ij} a_{ij}^p p_j,\ v_i = \tilde{v}_i / (\tilde{v}_i, A\tilde{v}_i)^{0.5},\ 0 < \omega_{ij} < 1,\ i > j, \\ &v_q = p_q, \quad q = 1,...,i-1, i+1,...,n, \end{aligned} \tag{2.6}$$

где $a_{ij}(P) = a_{ij}^p = (p_i, Ap_j)/[(p_j, Ap_j)(p_i, Ap_i)]^{0.5}$, а ω_{ij} выбирается по правилу:

$$\omega_{ij} = \begin{cases} 1,\ если\ (a_{ij}^p)^2 \le 1 - \delta, \\ 1 - [(a_{ij}^p)^2 - 1 + \delta]^{0.5} / |a_{ij}^p|,\ если\ (a_{ij}^p)^2 > 1 - \delta, \end{cases} \qquad 0 < \delta < 1. \tag{2.7}$$

Матрица *V=W(i,j,P)* также А-нормированная. Формула (2.7) задает шаг неполной релаксации алгоритма обучения (2.6) в случае подозрительно больших значениях $(a_{ij}^p)^2$. Лемма 2.1 определяет свойство сходимости процесса (2.6) при ограниченной релаксации (2.7), что следует из ограниченности модуля определителя матриц $W(i,j,P)$ при его росте (2.8).

Лемма 2.1. *Преобразование (2.6) увеличивает значение определителя:*

$$|W(i,j,P)| = \begin{cases} |P|/[1-(a_{ij}^p)^2], \text{ если } (a_{ij}^p)^2 \le 1-\delta, \\ |P|/\delta, \qquad \text{если } (a_{ij}^p)^2 > 1-\delta. \end{cases} \tag{2.8}$$

Множество различных пар $\{(i_k,j_k),\ i_k > j_k\}$, расположенных в некотором порядке, общее число которых равно M=n(n-1)/2), будем называть *полной последовательностью различных пар*. Множество полных последовательностей различных пар, таких, что $\{(j_k,1),,(j_k,j_k-1)\} \in \{(i_1,j_1),...,(i_k,j_k)\}$, $k=1,...,n(n-1)/2$, будем обозначать L_P. Если пары берутся из $Q \subset L_P$, то применительно к процессу

$$P^{k+1} = W(i_k,j_k,P^k), \quad k=1,2,...M, \tag{2.9}$$

это означает, что прежде чем вектор с номером j_k будет выступать в роли вектора для A-ортогонализации других векторов, то прежде он должен быть сам A-ортогонализован ко всем векторам с меньшими номерами.

Глобальные свойства сходимости процесса (2.6) определяются ниже.

Теорема 2.6. *Пусть матрица P^1 A-нормирована и $|P^1| \neq 0$. Тогда за конечное число итераций N процесса (2.9) для пар (i_k,j_k) из последовательности пар $\{Q_1 \subset L_P,\ Q_2 \subset L_P,...\}$ будет получена матрица P^{N+1}, столбцы которой сопряженные векторы.*

Рассмотрим последовательность пар индексов, удобную для получения алгоритма минимизации. Обозначим $E_1=(1,...n)$, $E_l=(l,l+1,...,n,\ 1,...,l-1)$, $l=2,...,n$. Элементы T_{li} последовательности $T_l=(l,\ l-1,\ l+1,\ l-2,\ l+2,...)$ получим поочередным выбором первого и последнего элементов из E_l. Последовательность пар $\tilde{\Pi}_l = ([l,l-1],[l-1,l+1],[l+1,l-2],....)$ получим расстановкой скобок в T_l. Перестановкой индексов в парах $\tilde{\Pi}_l$ получим последовательность Π_l, где первый индекс в паре больше второго.

Теорема 2.7. *Пусть матрица P^1 A-нормирована и $|P^1| \neq 0$. Тогда за конечное число итераций N процесса (2.9) для пар (i_k,j_k) из $\{\Pi_1, \Pi_2,...,\Pi_n, \Pi_1, \Pi_2,...,\Pi_n,...\}$ будет получена матрица P^{N+1}, столбцы которой сопряженные векторы.*

На основании процесса (2.9) разработан алгоритм минимизации вдоль векторов МСН, с последовательностью номеров из пар П, с одновременным их преобразованием (2.9), где коэффициенты вычисляются как коэффициенты одномерных и двумерных квадратичных моделей. В следующей теореме определяются свойства алгоритма при минимизации квадратичных функций.

Теорема 2.8. *Пусть в алгоритме МСН производится точный одномерный спуск. Тогда, если минимизируемая функция квадратичная с невырожденной матрицей Гессе, то ее минимум будет найден за конечное число итераций.*

В следующих теоремах установлены свойства сходимости алгоритма при минимизации общих функций.

Теорема 2.9. *Пусть $f(x)$ удовлетворяет условию 2.1. Тогда для последовательности x_k, $k=1, 2,...$, генерируемой алгоритмом МСН при точном одномерном спуске, имеет место линейная скорость сходимости.*

Теорема 2.10. *Пусть $f(x)$ удовлетворяет условию 2.1, дважды непрерывно* *дифференцируема, а ее матрица вторых производных $f''(x)$ удовлетворяет условиям: $m||z||^2 \le (z, f''(x)z) \le M||z||^2 \quad \forall z \in R^n$, $||f''(x)-f''(y)|| \le L||x-y|| \quad \forall x,y \in R^n$ $(L>0)$. Тогда последовательность $\{x_k\}$, генерируемая алгоритмом МСН при точном одномерном спуске, сходится точке к x^* сверхлинейно.*

Таким образом, в этой главе изложены методы случайного поиска с варьированием метрики для решения сложных гладких и негладких задач безусловной оптимизации; для класса методов минимизации вдоль векторов линейно-независимой системы получены оценки скорости сходимости на сильновыпуклых функциях; разработано семейство алгоритмов обучения метрики, сочетающих локальные и глобальные свойства сходимости процесса обучения, и, на этой основе, получен эффективный метод минимизации без вычисления производных. Результаты тестирования и сравнения алгоритма МСН с известными методами этого класса устанавливают его высокую эффективность.

В третьей главе: на основе формализма теории обучения дается анализ схем обучения матриц в квазиньютоновских методах (КНМ); обосновываются их ускоряющие свойства; разрабатывается КНМ с двухшаговым алгоритмом обучения матриц.

Вывод и анализ на основе формализма теории обучения симметричных формул пересчета матриц. В схеме КНМ метода:

$$x_{k+1} = x_k - \gamma_k H_k f'_k, \tag{3.1}$$

$$H_{k+1} = H\left(H_k, \Delta x_k, y_k\right),\ \Delta x_k = x_{k+1} - x_k,\ y_k = \nabla f(x_{k+1}) - \nabla f(x_k), \tag{3.2}$$

рассчитываются матрицы B_k или H_k – приближения матриц A или $H^* = A^{-1}$. Формулы перерасчета обозначим $H\left(H, \Delta x, y\right)$ или $B\left(B, \Delta x, y\right)$. Например,

формула BFGS: $$H(H, \Delta x, y) = H - \frac{Hy\Delta x^T + \Delta x y^T H}{(y, \Delta x)} + \left(1 + \frac{(y, Hy)}{(y, \Delta x)}\right)\frac{\Delta x \Delta x^T}{(y, \Delta x)}. \tag{3.3}$$

На квадратичных функциях выполняются соотношения:

$$y = A\Delta x \ \text{ или } \ \Delta x = H^* y,\ \text{ где } y = \nabla f(x + \Delta x) - \nabla f(x),\ H^* = A^{-1}. \tag{3.4}$$

Векторные равенства (3.4) образуют n обучающих соотношений для строк матриц (в силу симметрии – и для столбцов), для каждой из которых можно применить алгоритмы обучения, основанные на минимизации показателя качества (1.1). Доказано, что одношаговый алгоритм обучения для минимизации показателя (1.1) при N=1 в форме невязок $r_{k+1} = W(z_k)r_k$, где $r_k = c_k - c^*$, $W(z) = I - zz^T / (z,z)$, применяемый к строкам, либо последовательно к строкам и столбцам матриц, приводит к известным формулам пересчета матриц, записанным в форме невязок:

1) $R_+ = R\,W(y)$ $(R = H - H^*)$ – второй метод Бройдена;
2) $Z_+ = Z\ W(\Delta x)$ $(Z = B - A)$ – первый метод Бройдена;
3) $R_+ = W(y)\,R\,W(y)$ $(R = H - H^*)$ – формула Дж. Гринстадта ;
4) $Z_+ = W(\Delta x) Z W(\Delta x)$ $(Z = B - A)$ – симметричная формула Пауэлла.

Формулы BFGS и DFP выводятся применением преобразований 3) и 4) в системе координат с единичной матрицей Гессе квадратичной функции. Анализ процессов обучения в методах BFGS и DFP позволяет сделать вывод о более высокой степени ортогональности векторов спуска и обучения матриц в методе BFGS, что объясняет его преимущества.

Ускоряющие свойства квазиньютоновских методов. Обозначим x^* – точку минимума функции, f^* значение функции в ней. В системе координат $\hat{x} = Px$ процесс (3.1)-(3.2) эквивалентен минимизации функции $\varphi(\hat{x}) = f\left(P^{-1}\hat{x}\right) = f(x)$, которая удовлетворяет условию 2.1 с константами ρ_p и L_p. Определим преобразование $\hat{x} = Vx$, где V – невырожденная матрица, такое, что $\rho_V / L_V \geq \rho_p / L_p$ для всякой невырожденной $(n \times n)$ – матрицы P.

Теорема 3.1. *Пусть функция $f(x)$ удовлетворяет условию 2.1, H_0 – симметричная и строго положительно определенная ($n \times n$) – матрица. Тогда для последовательности $f(x_k)$, $k=0,1,2,...$, заданной процессом (3.1)-(3.3), при точном одномерном спуске, имеет место оценка:*

$$f_k - f^* \leq (f_0 - f^*) exp\left\{-\frac{\rho_V^3}{54L_V^3}\left[k - 3 - \frac{nM}{L_V} - \frac{n\rho_V}{m}\right]\right\}, \tag{3.5}$$

где m и M – соответственно минимальное и максимальное собственные значения матрицы $P^{-T}A_0^{-1}P^{-1}$.

Оценка (3.5) получена без предположения существования вторых производных функции и свидетельствует о наличии ускоряющих свойств КНМ в области минимизации в случае, когда существует преобразование координат, обеспечивающее неравенство $\rho_V^3 / L_V^3 >> \rho / L$, где отношение ρ / L определяет скорость сходимости метода скорейшего спуска.

Квазиньютоновский метод на основе двухшагового алгоритма обучения (КНМР). Определим условия на выбор шага γ_k в (3.1):

Условие 3.1. $f(x_{k+1}) \leq f(x_k) + \alpha(f_k', x_{k+1} - x_k)$, $\alpha \in (0,1)$, (3.6)

$$(f_{k+1}', x_{k+1} - x_k) \geq \beta(f_k', x_{k+1} - x_k), \quad \beta \in (\alpha, 1), \tag{3.7}$$

причем выбирается $\gamma_k = 1$, если для него выполнены условия (3.8), (3.9).

Алгоритм КНМР основан на преобразованиях матриц в (3.2) для пары векторов Δx_k, предварительно преобразованных в сопряженные:

$$\Delta\tilde{x}_{k-1} = \Delta x_{k-1} - u_k \Delta x_k, \; \tilde{y}_{k-1} = y_{k-1} - u_k y_k \quad (u_k = (\Delta x_k, y_{k-1})/(\Delta x_k, y_k)), \tag{3.8}$$

$$\tilde{H}_k = H(H_k, \Delta\tilde{x}_{k-1}, \tilde{y}_{k-1}), \qquad H_{k+1} = H(\tilde{H}_k, \Delta x_k, y_k). \tag{3.9}$$

Преобразования (3.8), (3.9) определяют итерацию двухшагового алгоритма обучения в квазиньютоновских методах.

Алгоритм КНМР.

1. *Задать $x_0 \in R^n$, матрицу $H_0 > 0$, $\varepsilon \in (0, 1/6)$. Положить $k = 0$, $q_0 = 0$. Вычислить $f_0' = f'(x_0)$.*
2. *Найти новое приближение минимума $x_{k+1} = x_k - \gamma_k H_k f_k'$, где шаг γ_k удовлетворяет условию 3.1. Положить $q_{k+1} = q_k + 1$.*
3. *Вычислить $f_{k+1}' = f'(x_k), \Delta x_k = x_{k+1} - x_k$, $y_k = f_{k+1}' - f_k'$. Если $\| f_{k+1}' \| = 0$, то останов.*

4. Если $q_{k+1} \geq 2$, *и* $|(\Delta x_{k-1}, y_k)|/[(\Delta x_{k-1}, y_{k-1})(\Delta x_k, y_k)]^{1/2} < 1 - \varepsilon^2$, (3. 10)
то перейти на п. 5, иначе положить $H_{k+1} = H(H_k, \Delta x_k, y_k)$ *и перейти на п.6.*

5. Положить $q_{k+1} = 0$ *и провести преобразования (3.8), (3.9).*

6. Положить $k = k + 1$. *Если выполнен критерий останова, то закончить вычисления, иначе перейти на п. 2.*

В работе изучаются способы наращивания размерности подпространства выполнимости квазиньютоновского соотношения при условии применения двухшагового алгоритма обучения (3.8), (3.9). Обоснованы следующие свойства КНМ на квадратичных функциях:

1. Размерность подпространства минимизации n_M сокращается по мере роста размерности подпространства выполнения квазиньютоновского соотношения n_{KC}, причем $n_M \leq n + 1 - n_{KC}$.
2. Размерность подпространства выполнения квазиньютоновского соотношения в процессе работы КНМ не убывает.
3. Отдельные итерации с точным одномерным спуском повышают на единицу размерность подпространства квазиньютоновского соотношения.
4. Отдельные включения итераций с преобразованием матриц для пар сопряженных векторов, полученных из линейно-независимых пар векторов спуска, повышают на единицу размерность подпространства квазиньютоновского соотношения.

Следствия отмеченных свойств изложены в теоремах 3.2, 3.3.

Теорема 3.2. *Для решения задачи минимизации квадратичной функции за конечное число итераций процесса (3.1)–(3.3), при условии 3.1 на одномерный спуск, достаточно не более* $(n-1)-$*го включения пар последовательных итераций {(3.1)–(3.3)} и {(3.1), (3.8), (3.9)}, причем решение, если оно не получено ранее, обязательно будет получено на следующей итерации после* $(n-1)-$*го включения.*

Теорема 3.3. *Для решения задачи минимизации квадратичной функции за конечное число итераций процесса(3.1) – (3.3), при условии 3.1 на одномерный спуск, достаточно не более* $(n-1)-$*го включения итераций {(3.1) – (3.3)} с точным одномерным спуском, причем решение, если оно не получено ранее, обязательно будет получено на второй итерации после* $(n-1)-$*го включения.*

Результаты теорем 3.2, 3.3 свидетельствуют о возможности применения приемов увеличения размерности подпространства выполнения КНС в произвольные моменты, что позволяет разработать устойчивые к ошибкам методы.

В § 3.6 дается обоснование квазиньютоновского метода на основе двухшаговых схем обучения матриц.

Теорема 3.4. *Алгоритм КНМР с параметром* $\varepsilon = 0$ *позволяет решить задачу минимизации квадратичной функции с невырожденной матрицей вторых производных за число итераций, не превосходящее* [2(n-1)+1].

Теорема 3.5. *Пусть алгоритмом КНМР решается задача минимизации квадратичной функции с невырожденной матрицей вторых производных. Тогда решение будет найдено за конечное число итераций.*

Таким образом, применение конечного числа шагов двухшагового алгоритма обучения матриц приводит к конечному окончанию процесса минимизации. Отметим, что в обычной задаче обучения линейной модели, в отличие от квазиньютоновских методов, двухшаговые алгоритмы минимизации показателя качества обучения не обеспечивают конечности процесса обучения.

Численное исследование на многомерных тестах алгоритма КНМР и квазиньютоновского метода с формулой BFGS свидетельствует о преимуществе алгоритма КНМР, которое существенно на функциях с высокой степенью вырожденности при больших размерностях.

Таким образом, в этой главе формализм теории обучения применен в квазиньютоновских методах для построения алгоритмов обучения, обладающих локальными и глобальными свойствами сходимости. Одношаговый алгоритм обучения в квазиньютоновских методах приводит к известным способам преобразования матриц, а двухшаговый – обеспечивают конечную сходимость метода при неточном одномерном спуске.

В четвертой главе разрабатываются и обосновываются эффективные алгоритмы обучения, реализующие сходящиеся локальные стратегии обучения в релаксационных субградиентных методах (РСМ) типа «сопряженных субградиентов», которые, в случае минимизации квадратичных функций, реализуют глобальные стратегии обучения. РСМ – это пример класса методов, где применение теории обучения позволило построить ряд методов оптимизации, эффективных на гладких и негладких задачах минимизации, в том числе и невыпуклых.

Подход построения эффективных алгоритмов обучения в релаксационных субградиентных методах типа «сопряженных субградиентов». Для решения задачи минимизации функции $f(x)$ в процессах ε-субградиентного типа

$$x_{i+1} = x_i - \gamma_i s_{i+1}, \quad \gamma_i = \arg\min_{\gamma} f(x_i - \gamma s_{i+1}),$$

направление s_{i+1} выбирают из множества $S(G)$ решений неравенств:

$$(s,g) > 0, \ \forall g \in G, \qquad \text{где } G = \partial_\varepsilon f(x_i). \tag{4.1}$$

В диссертации направления спуска в РСМ выбирается как одно из решений неравенств (4.1). Если множество $G \subset R^n$ принадлежит некоторой гиперплоскости, тогда решение системы неравенств можно получить, решая равенства $(s,g) = 1 \ (\forall g \in G)$ методами теории обучения. В работе для этих целей используются одношаговый и двухшаговый алгоритмы минимизации показателя (1.1) и одношаговый алгоритм обучения с растяжением пространства. Для построения новых методов негладкой минимизации на их основе создано несколько методов решения неравенств, обладающих конечной сходимостью на отделимых множествах, изложение которых приводится ниже. К числу подобных методов принадлежит и одноранговое семейство РСМ, частным случаем которого является известный эвристический r – алгоритм Н.З.Шора.

Одношаговый алгоритм обучения с растяжением пространства (ОАОР) – это алгоритм обучения (1.2), записанный в виде:

$$s_{i+1} = s_i + H_i g_i \frac{[y_i - (s_i, g_i)]}{(g_i, H_i g_i)}, \quad H_{i+1} = H_i - (1 - \frac{1}{\alpha^2}) \frac{H_i g_i g_i^T H_i^T}{(g_i, H_i g_i)} \quad (\alpha \geq 1). \quad (4.2)$$

Задача 4.1. *По данным* $g_i \in R^n$, $y_i \in R^1$, $i = 0,1,2,\ldots$, *процесса* $y_i = (s^*, g_i) + \xi_i$, *где* ξ_i *– ошибка, найти вектор неизвестных параметров* $s^* \in R^n$.

Предположение 4.1. *Пусть для* $s_i \in R^n$ *векторы* g_i *и значения* $y_i = (s^*, g_i) + \xi_i$ *задачи 4.1 удовлетворяют соотношениям*:

$$q_i^2 = (\Delta_i, g_i)^2 \big/ (\Delta_i \Delta_i)(g_i g_i) \geq q^2 > 0, \text{ где } \Delta_i = s_i - s^*, \quad (4.3)$$

$$\lambda_i^2 \leq \lambda^2 < 1, \text{ где } \lambda_i = [(g_i, s_i) - y_i] / [(g_i, s_i) - (g_i s^*)] - 1. \quad (4.4)$$

Сформулируем свойства сходимости процесса (4.2).

Теорема 4.1. *Пусть последовательность* $\{s_i\}$ *– результат процесса (4.2),* $\|g_i\| \neq 0$ *для* $i = 0,1,2,\cdots$, $s_0 \in R^n, H_0 = I$. *Тогда, если выполнены условия (4.3), (4.4):* $q_i^2 \geq q^2 > 0$, $\lambda_i^2 \leq \lambda^2 < 1$ *и* $1 \leq \alpha^2 \leq 1/\lambda^2$, *то имеет место оценка:*

$$\min_{0 \leq j \leq i-1} (\Delta_j, \Delta_j) \leq \frac{i(\Delta_0, \Delta_0)(\alpha^2 - 1)}{q^2 n(\alpha^{2i/n} - 1)}, \quad i \geq 1. \quad (4.5)$$

Получим оценку скорости сходимости одношагового алгоритма обучения (ОАО):

$$s_{i+1} = s_i + g_i [y_i - (s_i, g_i)] / (g_i, H_i g_i), \quad (4.6)$$

на основе которого будет получен метод решения неравенств (ОАОН) и метод минимизации (РСМК).

Теорема 4.2. *Пусть последовательность* $\{s_i\}$ *– результат процесса (4.6) решения задачи 4.1,* $\|g_i\| \neq 0$ *для* $i = 0,1,2,\cdots$, $s_o \in R^n$. *Тогда, если выполнены условия (4.3), (4.4), имеет место оценка:* $(\Delta_i, \Delta_i) \leq (\Delta_0, \Delta_0)(1 - q^2(1 - \lambda^2))^i$.

Результаты теорем 4.1, 4.2 устанавливают преимущество в скорости сходимости алгоритма ОАОР сравнительно с алгоритмом ОАО.

Одношаговый алгоритм обучения с растяжением пространства для решения множества неравенств (ОАОРН) является распространением алгоритма ОАОР на задачу решения неравенств (4.1). Обозначим: $\eta(G)$ – вектор минимальной длины из *G*, $\rho_G \equiv \rho(G) = \| \eta(G) \|$, $R_G \equiv R(G) = \max_{g \in G} \| g \|$, $\mu_G = \eta(G) / \rho_G$, $s^* = \mu_G / \rho_G$, $R_S \equiv R_S(G) = \max_{g \in G} (\mu_G, g)$, $r(G) = \rho_G / R_S$, $v(G) = \rho_G / R_G$, $S_\delta(x) = \{ z \in R^n \mid \| z - x \| \leq \delta \}$.

Предположение 4.2. *Множество G выпуклое, замкнутое, ограниченное* $(R_G < \infty)$ *и удовлетворяет условию отделимости, то есть* $\rho_G > 0$.

Следующая теорема позволяет свести задачу нахождения $s \in S(G)$ к задаче 4.1 и обосновать применимость алгоритмов ОАО и ОАОР для ее решения.

Теорема 4.3. *Пусть для множества G выполнено предположение 4.2 и известно некоторое приближение* $s_i \in S_{\delta_0}(s^*)$ *искомого вектора* $s^* = \mu_G / \rho_G$. *Тогда* *для текущего приближения s_i вектор $g_i \in G$ и величина $y_i = 1$ в задаче 4.1 с искомым вектором s^* удовлетворяют условиям (4.3), (4.4):*

1) *если* $(s_i, g_i) \le 0$, *причем справедливы оценки:* $q^2 = 1/\delta_0^2 R_G^2, \quad \lambda^2 = (1-\rho_G / R_S)^2$;

2) *если* $(s_i, g_i) < 1$, *причем справедливы оценки:* $q^2=0, \quad \lambda^2=1$.

Алгоритм ОАОРН для решения неравенств.

1. *Положить* i=0. *Задать* $s_0 \in R^n$, $\alpha > 1$, $H_0 = I$.
2. *Найти* $g_i \in G$, *удовлетворяющий условию* $(s_i, g_i) \le 0$, *если такого вектора не существует, то* $s_i \in S(G)$, *закончить работу алгоритма.*
3. *Получить новое приближение* $s_{i+1} = V(s_i, g_i, H_i)$ *и* $H_{i+1} = H(g_i, H_i)$, *где*
$$V(s,g,H) = s + Hg[1-(s,g)]/(g,Hg), \quad H(g,H) = H - (1-1/\alpha^2)Hgg^T H^T /(g,Hg).$$
4. *Положить i=i+1. Перейти на пункт 2.*

Приведем оценку скорости сходимости алгоритма ОАОРН.

Теорема 4.4. *Пусть множество G удовлетворяет предположению 4.2, а последовательность* $\{s_i\}$ *генерируется алгоритмом* ОАОРН *при условии:* $1 < \alpha^2 \le 1/(1-\rho_G / R_s)^2 = 1/\lambda^2$. *Тогда вектор* $s_i \in S(G)$ *будет получен за конечное число итераций, причем пока* $s_i \notin S(G)$: *а) для величин* q_i *и* λ_i *выполнены условия (4.3), (4.4) с параметрами* q^2 *и* λ^2 *равными:* $q^2 = 1/\|\Delta_0\|^2 R_G^2, \quad \lambda^2 = (1-\rho_G / R_S)^2$; *б) для сходимости* $s_i \to s^*$ *имеет место оценка (4.5).*

РСМ с растяжением пространства в направлении субградиента (МРП) основан на методе решения неравенств ОАОРН.

Алгоритм МРП **(*r_g*).**

1. Задать начальное приближение $H_0 = I$, $x_0 \in R^n$, $s_0 = 0$, *целые* $k = i = q_0 = 0$, *параметр* $\alpha > 1$. *Вычислить* $g_0 \in \partial f(x_0)$. *Задать последовательности* $\varepsilon_j \ge 0$, $\sigma_j \ge 0$, $j=0, 1, \ldots$.

2. Если $g_i=0$, *то закончить вычисления.*

3. Если $(g_i, H_i g_i) \le \varepsilon_k^2$ *или* $(g_i, H_i g_i)/(g_i, g_i) \le \sigma_k^2$, *то произвести обновление* $k = k+1, q_k = i, s_i = 0, H_i = I$, $z_k = x_i$.

4. Получить новое приближение: $s_{i+1} = V(s_i, g_i, H_i)$, $H_{i+1} = H(g_i, H_i)$.

5. Вычислить новое приближение: $x_{i+1} = x_i - \gamma_i s_{i+1}$, $\gamma_i = \arg\min_{\gamma} f(x_i - \gamma s_{i+1})$.

6. Вычислить субградиент $g_{i+1} \in \partial f(x_{i+1})$, *исходя из условия* $(g_{i+1}, s_{i+1}) \le 0$.

7. Положить i=i+1, перейти на пункт 2.

Свойства сходимости алгоритма определены в теоремах 4.5-4.7. Обозначим: $d(x) = \rho(\partial f(x))$, x_* – точку минимума функции, x^* – предельные точки последовательности $\{z_k = x_{q_k}\}_{k=1}^{\infty}$, генерируемой алгоритмом r_g. Определим свойства сходимости алгоритма в окрестность минимума.

Теорема 4.5. *Пусть функция* $f(x)$ *строго выпукла на* R^n, *множество*

$D(x_0)$ ограничено, на $D(x_0)$ при $x \neq x_*$, $r(\partial f(x)) \geq r_0 > 0$, параметр α алгоритма r_g удовлетворяет соотношению: $1 < \alpha = 1/(1-r_\alpha) < 1/(1-r_0)$. Тогда, если $\varepsilon_k = E_0$, $\sigma_k = \Sigma_0$, то $d(x^*) \leq max\{E_0, 2\Sigma_0 Q_0\} \equiv d_0$, где $Q_0 = \max_{x \in D(x_0)} R(x)$.

На основе теоремы 4.5 формулируются условия сходимости алгоритма.

Теорема 4.6. *Пусть выполнены предположения теоремы 4.5, и функция* $f(x)$ *на множестве* $D(x_0)$, $x \neq x_*$ *удовлетворяет условию* $v(\partial f(x)) \geq v_0 > 0$. *Тогда, если* $\varepsilon_k = 0$, $\sigma_k = \Sigma_0 \leq v_0 / 4$, *то любая предельная точка последовательности* $\{z_k\}$ *является точкой минимума на* R^n.

Теорема 4.7. *Пусть выполнены предположения теоремы 4.5. Тогда, если* $\varepsilon_k \to 0$, $\sigma_k \to 0$, *то любая предельная точка последовательности* $\{z_k\}$ *является точкой минимума на* R^n.

Связь алгоритма МРП с методом сопряженных градиентов (МСГ).

Теорема 4.8. *Пусть функция f(x) квадратичная и в МРП* $H_0 = I, \alpha \geq 1$. *Тогда, если начальные точки алгоритмов МРП и МСГ совпадают, то, при* $n \geq k \geq 0$, *до тех пор, пока* $g_k \neq 0$, *оба процесса генерируют совпадающие последовательные приближения.*

***Метод решения неравенств на основе одношагового алгоритма обучения* (ОАОН)**:

1. Положить $i = 0$. *Задать* $s_0 \in R^n$.

9. Найти $g_i \in G$ *при условии* $(s_i, g_i) \leq 0$, *если такого вектора не существует, то* $s_i \in S(G)$, *закончить работу алгоритма.*

10. Получить s_{i+1} *по формуле (1.1) при* $y_i = 1$: $s_{i+1} = s_i + g_i \dfrac{[1-(s_i, g_i)]}{(g_i, g_i)}$.

1. Положить $i = i + 1$. *Перейти на пункт 2.*

Дадим оценку скорости сходимости $s_i \to s^*$.

Теорема 4.9. *Пусть множество* G *удовлетворяет предположению 4.2. Тогда для оценки скорости сходимости к точке* s^* *последовательности* $\{s_i\}$, *генерируемой алгоритмом A1 до момента останова, справедливо соотношение:* $\| s_i - s^* \|^2 \leq (\| s_0 \| + \rho_G^{-1})^2 - i / R_G^2$, *а при некотором значении* i, *удовлетворяющем неравенству* $i \leq i^* \equiv R_G^2 (\| s_0 \| + \rho_G^{-1})^2 + 1$, *будет получен вектор* $s_i \in S(G)$.

Алгоритм минимизации (РСМК) на основе алгоритма **ОАОН**:

Задать начальное приближение $x_0 \in R^n$, *целые* $k = i = 0$.

1. Положить $k = k + 1, q_k = i, s_i = 0, \Sigma_i = 0$.

2. Задать ε_k, m_k.

1. Вычислить $g_i \in \partial f(x_i)$ *такой, что* $(g_i, s_i) \leq 0$. *Если* $g_i = 0$, *то останов.*

2. Найти новое приближение направления спуска: $s_{i+1} = s_i + g_i [1-(s_i, g_i)] / (g_i, g_i)$.

3. Вычислить новое приближение критерия: $\Sigma_{i+1} = \Sigma_i + (g_i, g_i)^{-1}$.

4. Вычислить: $x_{i+1} = x_i - \gamma_i s_{i+1}$, $\quad \gamma_i = \arg\min_{\gamma} f(x_i - \gamma s_{i-1})$.

7. Положить $i = i+1$. *Если* $1/\sqrt{\Sigma_i} < \varepsilon_k$ *или* $i - q_k > m_k$, *то перейти на пункт 1.*

8. Перейти на пункт 3.

Определим свойства сходимости алгоритма в окрестность минимума.

Теорема 4.10. *Пусть функция* $f(x)$ *строго выпукла на* R^n, *множество* $D(x_0)$ *ограничено и параметры* ε_k, m_k, *задаваемые в пункте 2 алгоритма РСМК, фиксированы:* $\varepsilon_k = E_0 > 0$, $m_k = M_0$. *Тогда, если* x^* *– предельная точка последовательности* $\{x_{q_k}\}_{k=1}^{\infty}$, *генерируемой алгоритмом РСМК, то* $d(x^*) \le \max\{E_0, R(x_0)/\sqrt{M_0}\} \equiv d_0$, *где* $R(x_0) = \max_{x \in D(x_0)} \max_{v \in \partial f(x)} \|v\|$. *В частности, если* $M_0 \ge R^2(x_0) E_0^{-2}$, *то* $d(x^*) \le E_0$.

На основе теоремы 4.10 формулируются условия сходимости алгоритма.

Теорема 4.11. *Пусть функция* $f(x)$ *строго выпукла, множество* $D(x_0)$ *ограничено и* $\varepsilon_k \to 0, m_k \to \infty$. *Тогда любая предельная точка последовательности* $\{x_{q_k}\}$, *генерируемая алгоритмом РСМК, является точкой минимума функции* $f(x)$ *на* R^n.

Связь алгоритмов РСМК и сопряженных градиентов.

Теорема 4.13. *Пусть множество* $D(x_0) = \{x \in R^n \mid f(x) \le f(x_0)\}$ *ограничено и функция f(x) строго выпукла и непрерывно дифференцируема в* R^n. *Тогда, если начальные точки алгоритмов РСМК и МСГ совпадают, то при* $n \ge k \ge 0$ *оба процесса генерируют совпадающие последовательные приближения.*

Разработана эффективная технология реализации РСМ, которая позволяет создавать надежные реализации РСМ методов. Разработаны и численно исследованы реализации алгоритмов МРП и РСМК. Результаты тестирования свидетельствуют о высокой эффективности и надежности (получено решение всех тестовых задач) предложенных методов при решении сложных гладких и негладких задач минимизации, в том числе и невыпуклых.

Таким образом, обучающие соотношения в субградиентных методах сформулированы в виде задачи решения неравенств, которая, в свою очередь, сведена к задаче решения равенств с ошибками. Это позволило использовать одношаговый и двухшаговый алгоритмы обучения для построения РСМ. На этой основе построен метод сопряженных субградиентов РСМК, сходящийся на негладких функциях и конечный на квадратичных функциях.

Применение многошаговых алгоритмов теории обучения позволило получить алгоритмы обучения с растяжением пространства: алгоритм для решения множества равенств; алгоритм для решения неравенств на отделимых множествах. На основе алгоритма решения неравенств разработан новый релаксационный субградиентный метод с растяжением пространства в направлении субградиента МРП для решения сложных задач оптимизации, сходящийся на негладких функциях и конечный на квадратичных функциях.

В пятой главе получено одноранговое семейство релаксационных субградиентных методов, полюсами которого являются r-алгоритм и модификация алгоритма Вульфа – метод сопряженных субградиентов с растяжением пространства. Здесь теория обучения является одной из компонент, приведших к построению семейства методов минимизации.

Метод сопряженных субградиентов с растяжением пространства (AWM(α)). Для множества G и вектора g введем подмножество $U(G,g)=\{u\in G\,|\,(g,u)\le 0\}$. Для векторов g и u определим вектор:

$$g_0^W(u,g)=g+\beta y,\quad \text{где}\quad y=u-g,\quad \beta=-(y,g)/(y,y). \tag{5.1}$$

Если $u\in U(G,g)$, то $g_0^W(u,g)$ принадлежит оболочке векторов g и u.

Из алгоритма минимизации Ф. Вульфа выделим ***метод решения неравенств (AW)***: $g_0\in G$. Найти $g_i=g_0^W(g_{i-1},u_i)$, пока $\exists u_i\in U(G,g_{i-1})$, $i=0,1,\ldots$.

Теорема 5.1 *Пусть множество G удовлетворяет предположению 4.2. Тогда алгоритм AW находит решение* $g_i\in S(G)$ *неравенств (4.1) за конечное число итераций, которое не превосходит минимального целого i, удовлетворяющего неравенству:* $i>ln\left(\rho^2/R^2\right)/ln\left(1-\rho^2/(\rho^2+R^2)\right)$.

Алгоритм обучения (4.2) применим к модифицированной системе равенств $\{(s,y_i)=0,\, i=0,1,\ldots,k-1,\ (s,g_k)=1\quad (y_i=g_{i+1}-g_i)\}$, полученной из равенств $\{(s,g_i)=1,\ i=0,1,\ldots k\}$, и распространим на решение неравенств. В результате придем к ***алгоритму решения неравенств (AR(α))***, содержащемуся в схеме r-алгоритма:

$$s_i=H_ig_i,\ y_i=g_{i+1}-g_i,\ g_{i+1}\in U(G,s_i), \tag{5.2}$$

$$H_{i+1}=H_i-(1-1/\alpha^2)\frac{H_iy_iy_i^TH_i^T}{(y_i,H_iy_i)}, \tag{5.3}$$

Теорема 5.2. *Пусть множество G удовлетворяет предположению 4.2 и* $R_S=\rho_G$. *Тогда при* $\alpha^2>1$ *алгоритм AR(α) находит решение* s_i *неравенств (4.1) за конечное число итераций.*

Алгоритм решения неравенств на основе выбора ближайшего к началу координат вектора в текущей метрике (AW(α)) получим, объединяя алгоритмы AW и AR(α):

3. *Положить i=0. Задать* $\alpha>1$, $H_0=I$. *Выбрать произвольно* $g_0\in G$.
4. *Найти* $u_i\in G$, *такой, что* $(H_iu_i,g_i)\le 0$. *Если такого вектора не существует, то* $H_ig_i\in S(G)$, *закончить работу алгоритма.*
5. *Вычислить* $y_i=u_i-g_i$ *и получить приближение матрицы по формуле (5.3):*
6. *Выбрать* $g_{i+1}=g^W(g_i,u_i)$, *где* $\{g^W(g_i,y_i)=g_i+\beta_iy_i,\ \beta_i=-\dfrac{(H_iy_i,g_i)}{(H_iy_i,y_i)}\}$.
7. *Положить i=i+1. Перейти на пункт 2.*

Теорема 5.3. *Пусть множество G удовлетворяет предположению 4.2 и* $R_S=\rho_G$. *Тогда при* $\alpha^2>1$ *алгоритм AW(α) находит решение неравенств (4.1) за* *конечное число итераций.*

Метод сопряженных субградиентов с растяжением пространства AWM(α) получим, используя алгоритм решения неравенств AW(α),:

1. Задать начальное приближение $H_0 = I$, $x_0 \in R^n$, параметр α >1. Вычислить $g_0 \in \partial f(x_0)$. Если g_0=0, то x_0 точка минимума, закончить вычисления.

2. Для i =0,1,2, ..., выполнить действия:

2.1. $x_{i+1} = x_i - \gamma_i s_i$, $\quad \gamma_i = \arg\min_{\gamma} f(x_i - \gamma s_i)$, $s_i = H_i g_i$.

2.2. Найти субградиент $u_{i+1} \in \partial f(x_{i+1})$ *такой, что* $(u_{i+1}, s_i) \le 0$.

2.3. Если i не кратно m, то вычислить: $y_i = u_{i+1} - g_i$, $g_{i+1} = g^W(g_i, y_i)$, *и выполнить (5.3), в противном случае:* $H_0 = I$, $g_{i+1} = u_{i+1}$.

Теорема 5.4. *На квадратичной функции алгоритм AWM(α),* $\alpha \ge 1$, *и метод сопряженных градиентов, при равенстве их начальных точек и циклов обновления m≤n, генерируют одинаковые приближения минимума для i ≤m.*

Одноранговое семейство релаксационных субградиентных методов с растяжением пространства (ARWM(α, Δ)). Обозначим ***A(α) – общий алгоритм решения неравенств***, получаемый из AW(α) замененой пункта 4:

{*4. Выбрать произвольно* $g_{i+1} \in G$ }.

Обозначим: $\theta = \theta(r) = [(1 - r_G)/(1 + r_G)]^2$ и $Q = Q(\alpha, \theta) = 1 + (\alpha^2 - 1)\theta$. В следующей теореме дается обоснование сходимости алгоритма A(α).

Теорема 5.5. *Пусть множество G удовлетворяет предположению 4.2 и θ <1/n. Тогда при* $1 < \alpha^2 \le 1/n\theta$ *алгоритм A(α) находит решение неравенств (4.1) за конечное число итераций, которое не превосходит минимального целого i, удовлетворяющего неравенству* $1 > \dfrac{4iR_G^2(\alpha^2 - 1)Q^i}{n\rho_G^2(\alpha^{2i/n} - 1)}$.

Введем зависимости: $\theta(\omega) = (1 - \omega)^2/(1 + \omega)^2$, $\omega(\theta) = (1 - \theta^{1/2})/(1 + \theta^{1/2})$. Установим зависимость параметров алгоритма *A(α)* от характеристик множества G, обеспечивающую его сходимость за конечное число итераций.

Теорема 5.6. *Пусть множество G удовлетворяет предположению 4.2, $\theta(r_G)$<1/n и задано некоторое θ*, удовлетворяющее неравенству:* $\theta(r_G) < \theta^* < 1/n$. *Тогда при* $1 < \alpha^2 \le 1/n\theta^*$, $\varepsilon \le \varepsilon^* = \rho_G\Delta/2$, *где Δ= r_G –r*, r*= $\omega(\theta^*)$, алгоритм A(α) находит решение неравенств (4.1) на множестве* $S_\varepsilon(G)$ *за конечное число итераций, которое не превосходит минимального целого i, удовлетворяющего неравенству:*

$$1 > \frac{36iR_G^2(\alpha^2 - 1)[Q(\alpha,\theta^*)]^i}{n\rho_G^2(\alpha^{2i/n} - 1)} = \frac{36i(\alpha^2 - 1)[Q(\alpha,\theta^*)]^i}{n\nu_G^2(\alpha^{2i/n} - 1)}.$$

Алгоритмы решения неравенств AR(α), AW(α) отличаются от A(α) конкретизацией пункта 4. В AR(α) он имеет следующий вид: {*4. Задать* $g_{i+1} = u_i$}. ***Семейство алгоритмов решения неравенств ARW(α,Δ)*** получим из алгоритма A(α), заменив в нем пункт 4:

$\{4.\ Задать\ g_{i+1} = \Delta \cdot g^W(g_i, u_i) + (1-\Delta)\cdot u_i\ \ (0 \le \Delta \le 1)\}$.

Одноранговое семейство релаксационных субградиентных методов с растяжением пространства ARWM(α,Δ) на основе алгоритм ARW(α,Δ):

1. Задать $H_0 = I$, $x_0 \in R^n$, *целые i=0,* $m_0=0$, *N – период обновления,* $\alpha > 1$, $0 \le \Delta \le 1$. *Вычислить* $g_0 \in \partial f(x_0)$. *Если* $g_0=0$, *то* x_0 *точка минимума, останов.*

2. $s_{i+1} = H_i g_i$, $\quad x_{i+1} = x_i - \gamma_i s_{i+1}$, $\gamma_i = \arg\min_{\gamma} f(x_i - \gamma s_{i+1})$.

3. Вычислить $u_i \in \partial f(x_{i+1})$, $(u_i, s_{i+1}) \le 0$. *Если* $u_i=0$, *то* x_{i+1} *– точка минимума, останов.*

4. Если $i - m_i \ge N$, *то:* $m_{i+1} = i$, $g_{i+1} = u_i$, $H_{i+1} = I$*; перейти на пункт 7.*

5. Вычислить: $y_i = u_i - g_i$, $\beta_i = -\dfrac{(H_i y_i, g_i)}{(H_i y_i, y_i)}$, $g^W_{i+1} = g_i + \beta_i y_i$. *Если* $g^W_{i+1} = 0$, *то: произвести обновление {* $m_{i+1} = i$, $g_{i+1} = u_i$, $H_{i+1} = I$ *} и перейти на пункт 7.*

6. Положить $m_{i+1} = m_i$*, вычислить:*

$$g_{i+1} = \Delta \cdot g^W{}_{i+1} + (1-\Delta)\cdot u_i, \quad H_{i+1} = H_i - (1 - 1/\alpha^2)\frac{H_i y_i y_i^T H_i^T}{(y_i, H_i y_i)}.$$

7. Положить i=i+1, перейти на пункт 2.

Обозначим i_k, $k=0,1,\dots,$ – индексы i, при которых происходит обновление в пунктах 4 или 5. Обозначим z_k – точки x_{i_k}, x_* – точку минимума функции, x^* – дельные точки последовательности $\{z_k\}_{k=1}^{\infty}$.

Теорема 5.7. *Пусть множество* $D(x_0)$ *ограничено, функция f(x) строго выпуклая на* R^n, *и на* $D(x_0)$ *при* $x \ne x_*$, *ее характеристики удовлетворяют соотношениям:* $\theta(r(\partial f(x))) \le \theta_0 < 1/n$, $v(\partial f(x)) \ge v_0 > 0$, *где* θ_0 *и* v_0 *некоторые константы. Пусть задано некоторое* θ^**, такое, что* $\theta_0 < \theta^* < 1/n$. *Тогда при параметрах алгоритма минимизации ARWM(α, Δ):* $1 < \alpha^2 \le 1/n\theta^*$, $N > 2\, N(v_0, \theta^*)$, *где* $N(v_0, \theta^*)$ *– минимальное целое i, удовлетворяющее неравенству* $1 > \dfrac{36i(\alpha^2-1)[Q(\alpha,\theta^*)]^i}{nv_0^2(\alpha^{2i/n}-1)}$, *любая предельная точка последовательности* $\{z_k\}$ *является точкой минимума.*

Алгоритмы однорангового семейства РСМ с растяжением пространства и, в частности алгоритм r_{OM}, были реализованы и исследованы. Результаты тестирования свидетельствуют о высокой эффективности и надежности (получено решение всех тестовых задач) предложенных методов при решении сложных гладких и негладких задач минимизации, в том числе и невыпуклых.

Оценки скорости сходимости r-алгоритма при конечных значениях параметра растяжения пространства. Последовательные приближения r-алгоритма строятся по формулам:

$$x_{k+1} = x_k + \gamma_k r_k, \quad r_k = -H_k f'(x_k), \quad \gamma_k = \arg\min_{\gamma>0} f(x_k + \gamma r_k), \qquad (5.4)$$

$$H_{k+1} = H_k - \beta \frac{H_k y_k y_k^T H_k^T}{(H_k y_k, y_k)}, \quad y_k = f'(x_{k+1}) - f'(x_k), \quad \beta \in (0,1), \quad (5.5)$$

где x_0 – заданная начальная точка, $H_0 > 0$. Обозначим через x^* – точку минимума функции $f(x)$, $f_* = f(x^*)$, $f_k = f(x_k)$, $\mu_k = f_k - f_*$, $f_k' = f'(x_k)$.

В системе координат $\hat{x} = Px$ процесс (5.4)-(5.5) эквивалентен минимизации функции $\varphi(\hat{x}) = f\left(P^{-1}\hat{x}\right) = f(x)$, которая удовлетворяет условию 2.1 с константами ρ_p, L_p. Определим преобразование $\hat{x} = Vx$, где V – невырожденная матрица, такое, что $\rho_V / L_V \geq \rho_p / L_p$ для всякой невырожденной $(n \times n)$ – матрицы P. В следующей теореме обосновываются ускоряющие свойства r-алгоритма.

Теорема 5.8. *Пусть функция f(x) удовлетворяет условию A. Тогда для последовательности* $\{f_k\}$, $k = 0,1,2,\ldots$, *заданной процессом (5.4) – (5.5) с матрицей* H_0, *удовлетворяющей условию* $m_0 \leq (H_0 z, z)/(z,z) \leq M_0 \quad \forall z \in R^n$, *имеет место оценка:*

$$\mu_{k+1} \leq \mu_0 \exp\left\{ -\frac{\rho_V^2}{L_V^2} \left[\frac{k \ln(1+\alpha)}{n\alpha} + \frac{\ln(m/M)}{\alpha} \right] \right\}, \qquad (5.6)$$

где m и M – соответственно минимальное и максимальное собственные значения матрицы $\overline{H}_0 = V H_0 V^T$.

Таким образом, применение алгоритма обучения с растяжением пространства позволило получить новый матричный алгоритм обучения для решения неравенств и тем самым предсказать наличие метода решения неравенств в r-алгоритме. На этой основе построено одноранговое семейство релаксационных субградиентных методов, полюсами которого являются эвристический r-алгоритм Н. З. Шора и новый метод Ф. Вульфа в метрике, относимые ранее к различным классам. Обоснованы ускоряющие свойства r-алгоритма при конечных значениях параметра растяжения пространства. Разработана эффективная реализация алгоритмов семейства, проведено их численное исследование и сравнение с известными методами.

В шестой главе разрабатывается программный комплекс релаксационных методов безусловной оптимизациии (ПКРМО) и приводятся примеры их приложений в задачах математического моделирования.

Программный комплекс ПКРМО включает прямые методы (СПО (§ 2.4) – алгоритм случайного поиска с оптимизацией параметров адаптации шага; СПМ (§ 2.4) – алгоритм СП с пространственной адаптацией шага; СПВМ1, СПВМ2, СПВМ3 (§ 2.5) – методы СП с варьированием метрики; МСН (§ 2.6) – метод минимизации без вычисления производных на основе рассредоточенной схемы А-ортогонализации), градиентные методы (КНМ –квазиньютоновский метод минимизации с формулой BFGS; КНМР (§ 3.6) – квазиньютоновский метод минимизации на основе двухшагового алгоритма обучения; МСГ – метод сопряженных градиентов; МСГП – модификация метода сопряженных градиентов, предложенная Б. Т. Поляком), субградиентные методы (РСМК (гл. 4) – метод сопряженных субградиентов для минимизации негладких функций высокой размерности; МРП (§4.8) – метод с растяжением пространства в направлении субградиента;

$ARWM_{OM}(\alpha, \Delta)$ (гл. 5) – одноранговое семейство релаксационных субградиентных методов; $r_{OM}(\alpha)$ (гл. 5) – r-алгоритм, который является частным случаем семейства алгоритмов $ARWM(\alpha, \Delta)$ при Δ=0).

Среди представленных методов для решения задач минимизации большой размерности применимы методы сопряженных градиентов (МСГ, МСГП) и метод сопряженных субградиентов РСМК. Отмеченные методы не всегда справляются с задачами в случае высокой степени вырожденности функции. С целью испытания возможностей методов с изменением метрики пространства на сложных задачах высокой размерности были решены известные тестовые задачи при размерности n=1000 квазиньютоновскими (КНМ, КНМР) и субградиентными (МРП, $r_{OM}(\alpha)$) методами. Во всех методах в каждой точке функция и субградиент вычислялись одновременно. В табл. 6.1 приведено количество затраченных методом вычислений функции и субградиента, которое соответствует моменту выполнения условия $f_k - f^* \le \varepsilon$.

Таблица 6.1.

Функции (n=1000)	Точность ε	$r_{OM}(\alpha)$	МРП	КНМ	КНМР
$f_1(x)$	10^{-10}	921	1688	3059	1792
$f_2(x)$	10^{-5}	2179	541	2116	1234
$f_3(x)$	10^{-4}	40738	18088	–	–

Субградиентными методами получено решение задачи минимизации негладкой функции $f_3(x) = \max\limits_{1 \le i \le n} (\,| x_i \,| i^3)$, степень вырожденности которой соответствует квадратичной форме с числом обусловленности 10^{18}, что, вместе с негладкостью, создает значительные сложности для метода минимизации. На гладких функциях $f_1(x)$, $f_2(x)$ они конкурентоспособны с квазиньютоновскими методами. Методы с изменением метрики пространства комплекса ПКРМО эффективны при решении сложных гладких и негладких задач высокой размерности.

Анализ методов ПКРМО показывает, что они обеспечивают возможность решения задач безусловной оптимизации с различными типами и степенями сложности (высокая степень вырожденности, негладкость, невыпуклость, высокая размерность).

Алгоритмы программного комплекса ПКРМО, в силу возможности эффективно решать негладкие невыпуклые задачи, являются надежным средством реализации методов оценивания параметров математических моделей, например метода наименьших квадратов и метода наименьших модулей. Методы ПКРМО обеспечивают широкие возможности реализации схем поиска математической модели методами оптимизации. Дадим краткое описание прикладных задач построения математических моделей по экспериментальным данным, которые решены в диссертации.

Комплекс математических моделей коэффициента поглощения солнечной радиации (КПСР) терморегулирующих покрытий (ТРП). Оценка работоспособности ТРП в условиях длительной эксплуатации представляет собой важную практическую задачу, которая не поддается решению аналитическими методами. Поэтому возникает необходимость проведения ускоренных испытаний в условиях повышенной интенсивности облучения с последующей их интерпретацией с помощью методов математического моделирования. В этой связи возникает необходимость разработки комплекса математических моделей (ММ) оптической деградации ТРП, учитывающих возможные закономерности деградации широкого класса ТРП. Изменение коэффициента поглощения a_s необходимо представить в виде модели $\Delta a_S(x,t,J,T)$, где x – вектор неизвестных параметров, t,J,T – соответственно время, интенсивность и температура облучения.

В результате анализа существующих моделей прогнозирования оптической деградации ТРП КЛА, экспериментальных наблюдений за образованием радиационных дефектов и центрами оптического поглощения и многолетнего опыта разработки математических моделей для конкретных покрытий сформировано множество альтернативных моделей для описания изменений КПСР ТРП, каждая из моделей которого включает в себя определенное число составляющих двух видов $F_i(x,t,J,T),\ FF_j(x,t,J,T),\ i=1,...6,\ j=1,...4$.

Предварительные расчеты на имеющемся множестве экспериментальных данных показали, что для описания экспериментальных данных достаточно присутствия в модели только двух составляющих, причем наилучшие сочетания образуются по формулам:

$$\Delta a_S(x,t,J,T) = x_4 F_i + x_5 FF_j,\ i = 1,...6,\ j = 1,...4\,. \tag{6.1}$$

Модели вида (6.1) представляют множество для поиска эффективной модели для прогнозирования.

Схема долгосрочного прогнозирования изменений КПСР ТРП космических летательных аппаратов (КЛА) по данным лабораторных испытаний. Задача долгосрочного прогнозирования состоит в предсказании значения a_s при заданном постоянном режиме облучения для заданного времени облучения. Комплекс ММ может быть использован для прогнозирования оптической деградации ТРП на реальных орбитах. Для этого необходимо проведение наземных испытаний и определение изменений коэффициента поглощения a_S в условиях, имитирующих условия орбиты, расчет коэффициентов комплекса математических моделей, выбор «наилучшей» модели. Коэффициенты моделей x определяем по методу наименьших квадратов, посредством минимизации суммы квадратов:

$$\Phi(x) = \sum_{i=1}^{m} [\,\Delta a_s(x,t_i,J_i,T_i) - \Delta a_{si}\,]^2\ , \tag{6.2}$$

где m – количество экспериментальных данных по измерению Δa_s, x – вектор неизвестных параметров. При этом должны выполняться ограничения:

$$1)\ x_j > 0,\quad j = 1,..,n\,,\quad 2)\ x_4 + x_5 = \Delta a_{s\infty}\,,\quad 3)\ x_4 + x_5 \le 1 - a_{s0}\,, \tag{6.3}$$

где n – количество неизвестных параметров модели; $\Delta a_{s\infty}$ – предельное значение Δa_s при $t \to \infty$; a_{s0} – исходное значение интегральной отражательной способности; t – время облучения, $\Delta a_s(x,t,J,T)$ – изменение значения КПСР ТРП при облучении за время t. Значение $a_{s\infty} = a_{s0} + \Delta a_{s\infty} = 1$ соответствует КПСР абсолютно черного тела.

На основе алгоритмов из ПКРМО были разработаны методы оценки параметров моделей (6.1) посредством минимизации (6.2) при ограничениях (6.3), по данным наземных испытаний. Построение моделей осуществляется в два этапа. На первом этапе строится полное множество моделей (6.1). После этого вычисляется осредненное значение верхнего предела как некоторое среднее пределов полученных моделей. На втором этапе верхний предел всех моделей фиксируется в виде ограничения 2) из (6.3), с использованием осредненного предела, и заново строится полное множество моделей. Полученное множество моделей анализируется.

Критериев оценки качества моделей может быть несколько. От их количества и от того, насколько полно они отражают процессы деградации ТРП, будет зависеть точность выбора модели и точность долгосрочного прогноза. В качестве критериев использовались: критерий минимума остаточной дисперсии; критерий сравнения моделей для отдельных режимов, где худшим условиям функционирования ТРП должны соответствовать и большие изменения КПСР ТРП; критерий правдоподобности предельного значения Δa_s при $t \to \infty$. Для прогнозирования оптической деградации ТРП из множества моделей по критериям дискриминации выбираем наиболее оптимальную для данного ТРП и условий эксплуатации модель.

Изложенная методика математического моделирования и долгосрочного прогнозирования неоднократно проверялась и использовалась в практике прогнозирования изменений КПСР в условиях космоса по данным ускоренных лабораторных испытаний.

Прикладные задачи построения математических моделей по экспериментальным данным. Методы ПКРМО применены для определения нестационарных законов горения пороха на основе манометрических испытаний. Задача обработки экспериментальных данных зависимости давления пороховых газов от времени для определения силы пороха, коволюма пороховых газов и установления временной зависимости скорости горения сведена к ряду задач оптимизации алгоритмически заданных функций, которые эффективно решаются методом МСН из ПКРМО.

В диссертации обосновывается эффективность применения для оценивания параметров нейросетевых моделей релаксационных субградиентных методов. В § 6.7 при решении задачи построения искусственной нейронной сети минимальной сложности используются идеи негладкой регуляризации для целей построения рангового упорядочения переменных по их значимости и последующего удаления малозначимых. Здесь возникают серии невыпуклых негладких задач оптимизации, которые в состоянии решать развитые в работе субградиентные методы. В результате исследований §§ 6.6, 6.7 выявлено, что субградиентные методы МРП, $r_{OM}(\alpha)$ и РСМК являются эффективным средством решения задач оценивания параметров нейросетевых приближений и построения нейронной сети минимальной сложности.

Изучение систем многопродуктового производства на основании математических моделей приводит к необходимости решения задач нелинейного программирования высокой размерности. В § 6.8 сформулированы условия оптимальности плана в задаче оптимизации прибыли многопродуктового производства, дана их интерпретация и разработан метод декомпозиции по ограничениям получения оптимального решения, который заключается в решении двойственной задачи негладкой оптимизации малой размерности с последующим получением оптимального решения прямой. Задачи негладкой оптимизации схем декомпозиции могут быть решены субградиентными методами программного комплекса ПКРМО.

Важной задачей является задача мониторинга и анализа системы однотипных предприятий. В § 6.9 разработана схема и программный комплекс анализа системы шахт на основе микроэкономических характеристик предприятий, определяемых средствами аппарата производственных функций, параметры которых рассчитываются методами программного комплекса ПКРМО по данным «затраты-выпуск». В результате выделены качественно однородные группы предприятий, а в группах произведено их ранговое упорядочение по критерию эффективности, что позволяет сформулировать выводы о состоянии системы и мерах по ее реструктуризации.

В § 6.10 излагается комплекс оптимизационных методов аппроксимации, сглаживания и анализа поверхности для решения актуальной задачи разработки методов выявления структуры поверхности в виде ее разрывов и перегибов по данным наблюдения, которая находит приложения в геологии при изучении характера залегания полезных ископаемых.

Основные результаты диссертации:

1. Для класса методов случайного поиска разработаны итеративные алгоритмы обучения метрики и построены эффективные численные методы случайного поиска с изменением метрики на их основе для решения сложных гладких и негладких задач безусловной оптимизации.
2. Для класса методов минимизации вдоль векторов линейно независимой системы получены оценки скорости сходимости на сильновыпуклых функциях и разработано семейство конечно-сходящихся итеративных алгоритмов обучения с ограниченной релаксацией для построения сопряженных направлений, сочетающих локальные и глобальные свойства сходимости процесса обучения. На этой основе создан эффективный метод минимизации без вычисления производных (МСН).
3. Формализм теории адаптации и одношаговый и двухшаговый алгоритмы минимизации показателя качества обучения позволяют разработать эффективные алгоритмы обучения в субградиентных и квазиньютоновских методах, обладающие локальными и глобальными свойствами сходимости. Одношаговый алгоритм обучения в квазиньютоновских методах приводит к известным способам преобразования матриц, а двухшаговый – обеспечивает конечную сходимость метода при неточном одномерном спуске. Одношаговый алгоритм обучения позволяет построить метод решения неравенств на отделимых мно-

жествах и субградиентный метод (РСМК) на его основе.

4. Разработаны новые алгоритмы обучения с растяжением пространства: алгоритм для решения множества равенств; алгоритм для решения неравенств на отделимых множествах. На основе алгоритма решения неравенств разработан новый релаксационный субградиентный метод с растяжением пространства в направлении субградиента (МРП) для решения сложных задач оптимизации.
5. Разработан новый матричный алгоритм обучения для решения неравенств и построено одноранговое семейство релаксационных субградиентных методов на его основе, в которое, как частный случай, входит r-алгоритм и разработанное обобщение метода Вульфа.
6. Для долгосрочного прогнозирования изменений коэффициента поглощения солнечной радиации терморегулирующих покрытий космических летательных аппаратов по данным лабораторных испытаний разработаны: комплекс математических моделей оптической деградации терморегулирующих покрытий и схемы их построения; схема анализа множества моделей и выбора из него модели, на основе которой осуществляется прогноз при заданных условиях эксплуатации.
7. Разработан комплекс программ ПКРМО созданных методов: прямого поиска, квазиньютоновских и субградиентных. Алгоритмы комплекса ПКРМО являются надежным и эффективным средством решения сложных задач безусловной оптимизации и реализации оптимизационного инструментария в задачах структурно-параметрического моделирования.

Основные публикации автора по теме диссертации

1. Крутиков, В. Н. Релаксационный метод минимизации с растяжением пространства в направлении субградиента / В. Н. Крутиков, Т. В. Петрова // Экономика и мат. методы. – 2003. – Т. 39. – Вып. 1. – С 106-119.
2. Крутиков, В. Н. О скорости сходимости методов минимизации вдоль векторов линейно-независимой системы / В. Н. Крутиков // Журнал вычислительной математики и метематической физики. – 1983. – Т.23, №1. – С.218-220.
3. Крутиков, В. Н. Статистический метод вычисления сопряженных направлений / В. В. Захаров, В. Н. Крутиков // Кибернетика (СССР: Киев). – 1984. – №6. – С.95-100.
4. Крутиков, В. Н. Разработка комплекса математических моделей оптической деградации терморегулирующих покрытий космических летательных аппаратов / М. М. Михайлов, В. Н. Крутиков // Перспективные материалы. – 1997. – № 1. – С.21-27.
5. Крутиков, В. Н. Прогнозирование оптической деградации терморегулирующих покрытий космических летательных аппаратов по результатам наземных испытаний / М. М. Михайлов, В. Н. Крутиков // Перспективные материалы. – 1997. – № 2. – С.18-25.
6. Крутиков, В. Н. «Спецтема 1» / М. М. Михайлов, М. И.Дворецкий, Л. Г. Косицин, В. Н. Крутиков, Г. Г. Савельев, Б. И Кузнецов // Вопросы оборонной техники. – Серия 10, выпуск 151. Ленинград: ГОИ им. Вавилова. – 1980.

7. Крутиков, В. Н. «Спецтема 2» / М. М. Михайлов, М. И.Дворецкий, Л. Г. Косицин, В. Н. Крутиков // Вопросы оборонной техники. – Серия 10, выпуск 151. Ленинград: ГОИ им. Вавилова. – 1980.
8. Крутиков, В. Н. Алгоритм случайного поиска с адаптивной метрикой («SPM») / В. Н. Крутиков // Свидетельство об официальной регистрации программ № 2003612566. – М: РОСПАТЕНТ. Зарегистрировано в реестре программ для ЭВМ. – Москва, 25 ноября 2003г. – 2003.
9. Крутиков, В. Н. Релаксационный субградиентный метод с растяжением пространства в направлении субградиента («RSM») / В. Н. Крутиков // Свидетельство об официальной регистрации программ № 2003612567. – М: РОСПАТЕНТ. Зарегистрировано в реестре программ для ЭВМ. – Москва, 25 ноября 2003г. – 2003.
10. Крутиков, В. Н. Релаксационные методы безусловной оптимизации, основанные на принципах обучения: Учебное пособие / В. Н. Крутиков. – ГОУ ВПО «Кемеровский государственный университет». – Кемерово: Кузбассвузиздат. –2004. – 171с.
11. Крутиков, В. Н. Одноранговое семейство релаксационных субградиентных методов с растяжением пространства / В. Н. Крутиков // Электронный журнал "Исследовано в России". – 209. –2003. –С 2450-2459. http: // zhurnal.ape.relarn.ru / articles / 2003 / 209. pdf.
12. Крутиков, В. Н. Метод сопряженных субградиентов с растяжением пространства / В. Н. Крутиков, Д. В. Арышев // Электронный журнал "Исследовано в России" – 208. –2003. –С 2439-2449. –http: / zhurnal.ape.relarn.ru / articles / 2003 / 208.pdf.
13. Крутиков, В. Н. Реализация алгоритмов одноранговоrо семейства субградиентных методов / В. Н. Крутиков, Д. В. Арышев //Электронный журнал "Исследовано в России" – 43. –2004. –С. 464-473. http: // zhurnal.ape.relarn.ru / articles /2004 / 043.pdf.
14. Krutikov, V.N. Predicting the optical degradation of thermoregulating coatings of flying space systems on the basis of the results of tests carried out on earth / M. M. Mikhailov and V.N. Krutikov // Journal of Advanced Materials (Cambridge interscience publishing). –1996. –v.3. – №2. –pp. 106-113.
15. Krutikov, V.N. Method of examining the thermal adjusting coatings of space apparatus / M. M. Mikhailov and V.N. Krutikov // Journal of Advanced Materials (Cambridge interscience publishing). –1996. –v.3.– №1. – pp. 21-28.
16. Крутиков, В. Н. Метод определения коэффициента поглощения терморегулирующих покрытий в зависимости от времени, интенсивности излучения и температуры // М. М. Михайлов, М. И.Дворецкий, Л. Г. Косицин, В. Н. Крутиков // Космическая технология и материаловедение. – М.: Наука, 1982. – С. 95-100 .
17. Крутиков, В. Н. Методы минимизации на основе частной модели субградиентных множеств / В. Н. Крутиков // Методы оптимизации и их приложения / Труды 11-й Международной Байкальской школы-семинара. – Иркутск.1998. – Том 1. –С.101-104.
18. Крутиков, В. Н. Исследование субградиентных методов обучения нейронных сетей / В. Н. Крутиков, Д. В. Арышев // Вестник КемГУ. – Кемерово. –2004. –№1 (17). –С. 119-124 .
19. Крутиков, В. Н. Алгоритм последовательного отсева неинформативных переменных линейной модели / В. Н. Крутиков, Д. В. Арышев // Вестник КемГУ. – Ке-

мерово. –2004. –№1 (17). –С. 124-129.
20. Крутиков, В. Н. Новый релаксационный субградиентный метод с изменением метрики / В. Н. Крутиков // Вестник КемГУ. – Кемерово. –2001. – Вып. 4. – С. 16-22.
21. Крутиков, В. Н. Новый метод решения задач минимизации большой размерности / В. Н. Крутиков, Т. В. Петрова // Вестник КемГУ. – Кемерово. –2001. – Вып. 4. – С.65-71.
22. Крутиков, В. Н. Новый метод решения неравенств с растяжением пространства и субградиентные методы на его основе / В. Н. Крутиков, Н. Ю. Комаров // Вестник КемГУ. – Кемерово. – 2001. – №3(7). – С. 73-78.
23. Крутиков, В. Н. Субградиентный метод с неточным одномерным спуском / В. Н. Крутиков, Т. В. Петрова // Вестник КемГУ. – Кемерово. – 2001. – №3(7). –С. 85-91.
24. Крутиков, В. Н. Алгоритмы распознавания складок поверхности / В. Н. Крутиков, Е. Н. Глушко // Вестник КемГУ. – Кемерово. – 2001. –Вып. 4. – С. 60-65.
25. Крутиков, В. Н. Методы построения линейной функции полезности сложных объектов / В. Н. Крутиков, Я. С Ворошилов // Вестник КемГУ. – Кемерово. – 2001. Вып. 4. – С. 71-76.
26. Крутиков, В. Н. Оптимизационные задачи моделирования межотраслевого комплекса и методы их решения / В. Н. Крутиков // Вестник КемГУ. –Кемерово. – 2001. – №3(7). – С. 20-25.
27. Крутиков, В. Н. Метод восстановления метрики по упорядоченным расстояниям / В. Н. Крутиков, Н. Б. Пушкина // Вестник КемГУ. – Кемерово. – 2001. –Вып. 4. – С. 35-39.
28. Крутиков, В. Н. Абсолютные оценки скорости сходимости r-алгоритма и метода Ньютона / В. Н. Крутиков // Якутск: Матем. заметки ЯГУ. – 1997. – Т.4. – № 1. – С. 38-50.
29. Крутиков, В. Н. Новый релаксационный метод недифференцируемой минимизации / В. Н. Крутиков, Т. В. Петрова // Мат. заметки ЯГУ. – 2001. – Т.8. Вып. 1. – С. 50-60.
30. Крутиков, В. Н. Анализ динамики сходимости квазиньютоновских методов / В. Н. Крутиков // Матем. заметки ЯГУ. – Якутск. – 1994. – Т.1. Вып. 2. – С. 40 – 48.
31. Крутиков, В. Н. Квазиньютоновские методы на основе рассредоточенных способов восстановления Гессиана / В. Н. Крутиков // Мат. заметки ЯГУ. – 2000. –Т.7. – Вып. 2. – С. 62-81.
32. Крутиков, В. Н. Быстросходящийся метод решения задач безусловной минимизации, не требующий вычисления производных / В. Н. Крутиков // Газовая динамика. – Томск: ТГУ. – 1987. – С.85-99.
33. Крутиков, В. Н. Методы обнаружения тектонических нарушений / В. Н. Крутиков, Е. Н. Глушко // ТЭК и ресурсы Кузбасса: Вестник ТЭК Кузбасса. – 2001. – №3. – С. 22-23.
34. Крутиков, В. Н. Использование математического моделирования в задачах оптимизации межотраслевых связей в структуре регионального ТЭК / В. Н. Крутиков, В. Н. Вылегжанин // ТЭК и ресурсы Кузбасса. – 2001. – №4. – С.62-64.
35. Крутиков, В. Н. Алгоритмы аппроксимации поверхности, заданной значениями в узлах нерегулярной сетки / В. Н. Крутиков, С. Л. Злобина, Н. Ф. Бувальцев // Ма-

тем. заметки. ЯГУ. –1995. –Т.2, № 2. –С. 110-120.
36. Крутиков, В. Н. Повышение эффективности алгоритмов случайного поиска посредством включения в схему экстраполяции / В. Н. Крутиков, В. В. Захаров // Проблемы случайного поиска. – Рига: Зинатне. – 1978. – Вып.7. – С. 207-213.
37. Крутиков, В. Н. Теоретическое и экспериментальное исследование скорости сходимости двух алгоритмов случайного поиска / В. Н. Крутиков, В. В. Захаров // Вопросы разработки территориальных автоматизированных систем управления. – Кемерово: КемГУ. – 1984. – С. 65-70.
38. Крутиков, В. Н. Сравнение оценок скорости сходимости случайного покоординатного и циклического покоординатного спусков / В. Н. Крутиков // Применения случайного поиска при решении прикладных задач. – Кемерово: КемГУ. – 1981. – С.71-75.
39. Крутиков, В. Н. Новые адаптивные алгоритмы случайного поиска и численные эксперименты с ними / В. Н. Крутиков, В. В. Захаров // Оптимизация динамических систем. – Минск: изд.-во БГУ. – 1978. – С. 51-55 .
40. Крутиков, В. Н. Релаксационный субградиентный метод недифференцируемой минимизации с растяжением пространства / В. Н. Крутиков, Т. В. Петрова // Деп. в ВИНИТИ. – М. – 2000. – №3222-В00. – 28с.
41. Крутиков, В. Н. Методы распознавания разрывных форм пластовых месторождений / В. Н. Крутиков, Е. Н. Глушко // Деп. в ВИНИТИ. – 2000. – 191-В00. –31с.
42. Крутиков, В. Н. Методы распознавания складчатых форм пластовых месторождений / В. Н. Крутиков, Е. Н. Глушко // Деп. в ВИНИТИ. – 2000. – 190-В00. – 38с.
43. Крутиков, В. Н. Модифицированный алгоритм Качмажа для решения задачи построения разделяющей поверхности / В. Н. Крутиков, Т. В. Петрова // Деп. в ВИНИТИ. – М. – 1999. – 3940-В99. – 11с.
44. Крутиков, В. Н. Алгоритм Качмажа с изменением метрики / В. Н. Крутиков, Т. В. Петрова // Деп. в ВИНИТИ. – М. – 1999. – 3941-В99. – 9с.
45. Крутиков, В. Н. Экспериментальная оценка множества методов аппроксимации поверхности / В. Н. Крутиков, В. П. Потапов, Е. Н. Глушко // Деп. в ВИНИТИ. – М. – 1999. – 1401-В99. – 13с.
46. Крутиков, В. Н. Новый метод аппроксимации поверхности / В. Н. Крутиков, Е. Н. Глушко // Деп. в ВИНИТИ. – М. – 1999. – 1400-В99. – 11с.
47. Крутиков, В. Н. Метод негладкой регуляризации в задачах контрастирования / В. Н. Крутиков, Д. В. Арышев // Нейроинформатика и ее приложения: материалы XII Всероссийского семинара 1-3 октября 2004г. – Красноярск: ИВМ СО РАН, 2004. – С.80-81.
48. Крутиков, В. Н. Об определении нестационарных законов горения пороха на основе манометрических испытаний / Ю. П. Хоменко, В. М. Широков, В. Н. Крутиков // Материалы четвертой Международной школы-семинара: Внутрикамерные процессы, горение и газовая динамика дисперсных систем: сборник материалов. Том 1. – Санкт-Петербург, Россия, 2004г. – С 86-90.
49. Крутиков, В. Н. Релаксационные субградиентные методы на основе двухшаговых алгоритмов обучения / В. Н. Крутиков, Д. В. Арышев // Материалы Всероссийской научно-практической конференции «Наука и практика: диалоги нового века». – Часть 3. Информационные технологии и математическое моделирование. –

Томск: Твердыня. – 2003. – С. 125-127.
50. Крутиков, В. Н. Алгоритм построения нейронной сети с минимальным числом нейронов / В. Н. Крутиков, Д. В. Арышев // Наука и образование: Материалы Всероссийской научной конференции (20-21 февраля 2003 г.): Белово: БИ(Ф) КемГУ, 2003. – С 440-445.
51. Крутиков, В. Н. Об одной схеме анализа системы угольных предприятий / В. Н. Крутиков. // Наука и образование: Материалы Всероссийской научной конференции (12-13 апреля 2002 г.). – Ч2. – Белово: БИ(Ф) КемГУ, 2002. – С 332-336.
52. Крутиков, В. Н. Об одном методе выделения информативной системы признаков / В. Н. Крутиков, Д. В. Арышев // Материалы Всероссийской научно-практической конференции «Информационные технологии и математическое моделирование» – Томск: Твердыня, 2002. – С 194-196.
53. Крутиков, В. Н. Минимизация негладких функций методом случайного поиска / В. Н. Крутиков, Я. С. Ворошилов // Тез. докл. Второй научно-практической конференции «Наука и образование». –Белово: БФ КемГУ, 2001. – С 290-292.
54. Крутиков, В. Н. Методы регуляризации решения задачи обучения нейронной сети / В. Н. Крутиков, Д. И. Филинов // Нейроинформатика и ее приложения: Материалы VIII Всероссийского семинара, 6-8 октября 2000 г. – Красноярск: ИПЦ КГТУ, 2000. – С.91.
55. Крутиков, В. Н. Основные подходы к проектированию информационной модели углегазового месторождения / В. Т. Преслер, В. Н. Крутиков, А. В. Гарнага // Труды научно-технической конференции: Опыт и перспективы наукоемких технологий в угольной промышленности Кузбасса. – Кемерово: ИУ СО РАН, 1998. – С. 173-178.
56. Крутиков, В. Н. Построение математической модели метановыделения из пластов спутников / В. Н. Крутиков, С. К. Тризно, Г. Я. Полевшиков // Труды научно-технической конференции: Опыт и перспективы наукоемких технологий в угольной промышленности Кузбасса. – Кемерово: ИУ СО РАН. – 1998. – С.173-178.
57. Крутиков, В. Н. Квазиньютоновские методы с попарной А-ортоганализацией / В. Н. Крутиков // Тез. докл. 10-й Байкальской школы-семинара: Методы оптимизации и их приложения. – Иркутск: СЭИ, 1995. –С.88-89.
58. Крутиков, В. Н. Методы приближения поверхности по данным на хаотической сетке на основе алгоритма сдвига штрафов / В. Н. Крутиков, С. Л. Злобина // Тез. докл. 10-й Байкальской школы-семинара: Методы оптимизации и их приложения. – Иркутск: СЭИ, 1995. – С.196-197.
59. Крутиков, В. Н. Методы математического моделирования пластовых месторождений / В. Н. Крутиков, С. Л. Злобина // Тез. докл. Международной конференции по математическому моделированию. – Якутск. – 1994. – С.144.
60. Крутиков, В. Н. Прогноз особенностей пластовых месторождений / В. Н. Крутиков, С. Л. Злобина // Тез. докл. Международной конференции по математическому моделированию. – Якутск. – 1994. – С.143.
61. Крутиков, В. Н. Прогноз тектонической нарушенности угольных месторождений / В. Н. Вылегжанин, Э. И. Витковский, В. Н. Крутиков, В. П. Потапов // Горное давление в очистных и подготовительных выработках. – Новосибирск: ИГД, 1990. – С.42-45.
62. Крутиков, В. Н. О скорости сходимости метода BFS и его модификаций / В. Н.

Крутиков // Тез. докл. Международной школы-семинара по методам оптимизации и их приложениям. – Иркутск: СЭИ, 1989. – С.114-115.

63. Крутиков, В. Н. О свойствах модификаций алгоритма с растяжением пространства вдоль разности градиентов / В. Н. Крутиков, Г. Б. Кацэба // Тез. докл. Международной школы-семинара по методам оптимизации и их приложениям. – Иркутск: СЭИ, 1989. – С. 116-117.
64. Крутиков, В. Н. Сравнение оценок скорости сходимости r-алгоритма , метода Ньютона и DFP / В. Н. Крутиков // Тез. докл. десятого всесоюзного симпозиума: Системы программного обеспечения решения задач оптимального планирования. – М: АН СССР, ЦЭМИ, 1988. – С.45-46.
65. Крутиков, В. Н. Методы пространственного моделирования геомеханической обстановки шахтных полей средствами горной информатики / В. Н. Вылегжанин, Э. И. Витковский, В. Н. Крутиков, В. П. Потапов // Доклады \/II международного конгресса по маркшейдерскому делу. – Ленинград. – 1988. – Т6. – С.83-90.
66. Крутиков, В. Н. Управление метрикой окрестностей в алгоритмах дискретной оптимизации / В. Н. Крутиков, Корсакова О. Н. //Тез. докл.: Дискретная оптимизация и компьютеры. –М: ЦЭМИ–КемГУ, 1987. – С.127-128.
67. Крутиков, В. Н. Алгоритмы случайного поиска с переменной метрикой / В. Н. Крутиков, В. В. Захаров //Тезисы докладов Всесоюзного семинара: Случайный поиск и системы автоматизированного проектирования в электротехнике. – Ереван. – 1979. – С.7-8 .
68. Крутиков, В. Н. Об одной методике построения методов сопряженных направлений / В. Н. Крутиков //Тезисы докладов Всесоюзного научно-технического семинара (г. Харьков): Численные методы нелинейного программирования. – Москва. – 1979. – С.43-45 .
69. Крутиков, В. Н. Управление распределением испытаний в алгоритмах случайного поиска / В. Н. Крутиков. // Тезисы докладов 4 Всесоюзного совещания: Статистические методы теории управления. – М.: Наука. – 1978. – С.27-28.
70. Крутиков, В. Н. Алгоритмы случайного поиска с изменением метрики пространства испытаний / В. Н. Крутиков, В. В. Захаров //Системы управления. – Томск: изд.-во ТГУ, 1978. – С.131-135.
71. Крутиков, В. Н. Новые идеи и методы моделирования поверхностей пластов угля / В. Н. Крутиков, А. В. Паначев // Тез. докл. II Всесоюзного семинара: Информатика недр. – Кемерово: ИУ, 1989. – С.41.
72. Крутиков, В. Н. Оценка и сравнение двух методов аппроксимации поверхностей пластовых месторождений / В. Н. Крутиков, А. В. Паначев, В. П. Потапов // Тез. докл. II Всесоюзного семинара: Информатика недр. – Кемерово: ИУ, 1989. – С.47.

Подписано к печати 2.03.2005. Печать офсетная.
Печ.л.2,25. Тираж 120 экз. Заказ №

ГОУ ВПО «Кемеровский государственный университет». 650043,
г. Кемерово, ул. Красная, 6,
Отпечатано в издательстве «Кузбассвузиздат.» 650043, г. Кемерово, ул. Ермака, 7