

На правах рукописи



АЛКЕЗУИНИ Мухи Муртада

**ОБУЧЕНИЕ СЕТЕЙ РАДИАЛЬНЫХ БАЗИСНЫХ ФУНКЦИЙ
ДЛЯ РЕШЕНИЯ КРАЕВЫХ ЗАДАЧ
МОДЕЛИРОВАНИЯ ОБЪЕКТОВ С РАСПРЕДЕЛЕННЫМИ
ПАРАМЕТРАМИ**

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

**А в т о р е ф е р а т
диссертации на соискание ученой степени
кандидата технических наук**

ПЕНЗА 2020

Работа выполнена на кафедре «Компьютерные технологии» Федерального государственного бюджетного образовательного учреждения высшего образования «Пензенский государственный университет».

Научный руководитель – доктор технических наук, профессор
Горбаченко Владимир Иванович

Официальные оппоненты: **Тархов Дмитрий Альбертович**,
доктор технических наук, доцент, ФГАОУ ВО
«Санкт-Петербургский политехнический
университет Петра Великого», профессор
кафедры «Высшая математика»;

Солдатова Ольга Петровна,
кандидат технических наук, доцент,
ФГАОУ ВО «Самарский национальный
исследовательский университет
имени академика С. П. Королева»,
доцент кафедры «Информационные системы
и технологии»

Ведущая организация – АО «Научно-производственное предприятие
"Рубин"», г. Пенза

Защита диссертации состоится 18 декабря 2020 г., в _____ часов на заседании диссертационного совета Д 212.186.04 в Федеральном государственном бюджетном образовательном учреждении высшего образования «Пензенский государственный университет» по адресу: 440026, г. Пенза, ул. Красная, д. 40.

С диссертацией и авторефератом можно ознакомиться в библиотеке ФГБОУ ВО «Пензенский государственный университет» и на сайте: https://dissov.pnzggu.ru/ecspertiza/Tehnicheskie_nauki/alkezuini

Автореферат разослан «___» _____ 2020 г.

Ученый секретарь
диссертационного совета



Бождай Александр Сергеевич

ОБЩАЯ ХАРАКТЕРИСТИКА РАБОТЫ

Актуальность темы исследования. Многие объекты, рассматриваемые в теплофизике, расчетах на прочность, аэродинамике, ядерной физике, моделировании нефтяных месторождений, гидрогеологии, управлении объектами с распределенными параметрами и во многих других областях, являются объектами с распределенными параметрами. Широко используемыми математическими моделями объектов с распределенными параметрами являются краевые задачи для дифференциальных уравнений в частных производных (ДУЧП).

Аналитическое решение существует для ограниченного круга краевых задач, описывающих реальные объекты с распределенными параметрами, поэтому решение таких задач производится численными методами. Популярными численными методами решения краевых задач являются сеточные методы конечных разностей и конечных элементов. Но построение расчетных сеток для этих методов в случае реальных задач является сложным и трудоемким. Методы конечных разностей и конечных элементов требуют решения плохо обусловленных систем алгебраических уравнений высокой размерности. Другим подходом, получающим распространение в последнее время, являются бессеточные методы, которые не требуют построения сетки точек, соединенных между собой для аппроксимации уравнений. Бессеточные методы дают приближенное решение в виде суммы базисных функций, умноженных на веса функций. В качестве базисных популярны радиальные базисные функции (РБ-функции). Использование РБ-функций позволяет получать решение в произвольных точках расчетной области. Веса вычисляются из условия минимизации невязок в выбранных пробных точках, а параметры базисных функций назначаются до решения. Неформальный подбор параметров РБ-функций является существенным недостатком их использования.

Перспективной является реализация бессеточных методов на нейронных сетях. Решение краевых задач для ДУЧП возможно на многослойных персептронах. Но наиболее перспективным является применение сетей радиальных базисных функций (РБФ-сетей), так как РБФ-сети содержат всего два слоя, один из которых является линейным, а формирование решения носит локальный характер, что упрощает обучение таких сетей. Применение РБФ-сетей позволяет в процессе обучения сетей настраивать как веса, так и параметры РБ-функций.

Большой вклад в теорию и практику применения РБФ-сетей внесли L. Jianyu, L. Siwei, Q. Yingjian, H. Yaping, N. Mai-Duy, T. Tran-Cong, S. Sarra, H. Chen, L. Kong, W. Leng, M. Kumar, N. Yadav, A. H. Васильев, Д. А. Тархов, В. И. Горбаченко.

Решение задачи формируется в процессе обучения РБФ-сети. Поэтому важно сокращение времени обучения сетей. Но в настоящее время для обучения РБФ-сетей при решении краевых задач используются в основном простейшие градиентные методы первого порядка на основе градиентного спуска.

Быстрые методы второго порядка практически не используются при решении краевых задач на РБФ-сетях. Исключение представляет предложенный В. И. Горбаченко и М. В. Жуковым метод доверительных областей обучения РБФ-сетей. Но метод достаточно сложен, так как требует на каждой итерации обучения сети решать задачу условной минимизации. Поэтому тема диссертационного исследования, посвященная совершенствованию алгоритмов обучения РБФ-сетей при решении краевых задач, является актуальной.

Целью диссертационной работы является разработка алгоритмов обучения сетей радиальных базисных функций для численного решения краевых задач моделирования объектов с распределенными параметрами, позволяющих сократить время решения задачи.

Для достижения поставленной цели необходимо решить следующие **задачи**:

1. Проведение анализа применения РБ-функций и РБФ-сетей для решения краевых задач. Проведение анализа алгоритмов обучения сетей радиальных базисных функций для решения краевых задач. Рассмотрение и выделение недостатков в существующих алгоритмах, определение путей их устранения.

2. Адаптация для обучения РБФ-сетей современных быстрых градиентных алгоритмов первого порядка.

3. Разработка алгоритма обучения РБФ-сетей на основе градиентного метода второго порядка.

4. Разработка комплекса программ, реализующего разработанные алгоритмы, проведение вычислительных экспериментов с использованием разработанного комплекса программ.

Объектом исследования диссертационной работы являются численные методы решения краевых задач на РБФ-сетях.

Предмет исследования – алгоритмы обучения РБФ-сетей повышенной производительности для решения краевых задач.

Методы исследования. Для решения поставленных задач использовались теория искусственных нейронных сетей, теория численной оптимизации, численные методы.

Соответствие паспорту специальности. Результаты исследования соответствуют пункту 3 «Разработка, обоснование и тестирование эффективных вычислительных методов с применением современных компьютерных технологий» и 4 «Реализация эффективных численных методов и алгоритмов в виде комплексов проблемно-ориентированных программ для проведения вычислительного эксперимента» паспорта научной специальности 05.13.18 «Математическое моделирование, численные методы и комплексы программ».

Научная новизна диссертационной работы состоит в следующем:

1. Для обучения РБФ-сетей при решении задач аппроксимации функций и краевых задач математического моделирования объектов с распределенными параметрами разработаны численные алгоритмы на основе современных быстрых градиентных алгоритмов первого порядка: алгоритм градиентного спуска с импульсом, алгоритм метода Нестерова, алгоритм с адаптивной скоростью обуче-

ния RMSProp. Разработанные алгоритмы отличаются от известных алгоритмов учетом архитектуры РБФ-сетей, использованием аналитического вычисления градиента функционала ошибки сети, управляемыми значениями скорости настройки и коэффициентами коррекции различных компонентов вектора градиента. Алгоритмы позволяют существенно сократить время обучения РБФ-сетей при решении задач аппроксимации функций и краевых задач.

2. Для обучения РБФ-сетей, предназначенных для решения задач аппроксимации функций и краевых задач математического моделирования объектов с распределенными параметрами, разработан численный алгоритм, основанный на градиентном методе второго порядка – методе Левенберга – Марквардта, отличающийся от известных реализаций учетом специфики архитектуры сети и аналитическим вычислением параметров. Метод кардинально сокращает время обучения по сравнению с известными методами первого порядка и не требует решения на каждой итерации обучения задачи условной минимизации, как метод доверительных областей.

3. Для формирования вектора поправки параметров РБФ-сети, формируемого в алгоритме метода Левенберга – Марквардта, предложено использование итерационных методов подпространств Крылова, позволившее повысить численную устойчивость вычислений и снизить требования к выбору параметра регуляризации.

Практическая ценность работы. Разработанные алгоритмы обучения РБФ-сетей, а также программный комплекс, созданный с применением разработанных алгоритмов, могут быть использованы специалистами в области математического моделирования для численного моделирования объектов с распределенными параметрами.

Достоверность и обоснованность результатов, полученных в диссертационной работе, обеспечена корректным использованием математических методов, компьютерным моделированием разработанных алгоритмов и обсуждением полученных результатов в научных изданиях и выступлениями на международных и всероссийских научных конференциях.

Основные положения, выносимые на защиту:

1. Численные градиентные алгоритмы первого порядка для обучения РБФ-сетей при решении задач аппроксимации функций и краевых задач математического моделирования объектов с распределенными параметрами: алгоритм градиентного спуска с импульсом, алгоритм метода Нестерова, алгоритм с адаптивной скоростью обучения RMSProp, отличающиеся от известных алгоритмов учетом архитектуры РБФ-сетей, использованием аналитического вычисления градиента функционала ошибки сети, управляемыми значениями скорости настройки и коэффициентами коррекции различных компонентов вектора градиента.

2. Численные алгоритмы на основе метода Левенберга – Марквардта для обучения РБФ-сетей при решении задач аппроксимации функций и решения краевых задач моделирования объектов с распределенными параметрами, описываемых дифференциальными уравнениями в частных производных, отли-

чающиеся от известных реализаций учетом специфики архитектуры сети и аналитическим вычислением параметров.

3. Комплекс программ решения краевых задач на сетях радиальных базисных функций.

Реализация и внедрение. Предложенные в работе алгоритмы и разработанные на их основе программы внедрены на кафедре «Компьютерные технологии» Пензенского государственного университета при выполнении гранта РФФИ 16-08-00906 «Обучение сетей радиальных базисных функций при построении моделей процессов в сложных технических системах», что подтверждено актом о внедрении. Результаты работы используются в учебном процессе при преподавании курса «Нейронные сети» на кафедре «Компьютерные технологии» ФГБОУ ВО «Пензенский государственный университет», что подтверждено справкой о внедрении.

Апробация работы. Основные результаты работы докладывались и обсуждались на 12 международных и всероссийских конференциях: International conference on mathematical modelling in applied sciences (St. Petersburg, 2017); 2017 Annual Conference on New Trends in Information and Communications Technology Applications, NTICT 2017 (Baghdad, Iraq, 2017); 2019 International Conference on Applied and Engineering Mathematics, ICAEM (Taxila, Pakistan, 2019); 2019 International Russian Automation Conference – RusAutoCon, (Sochi, Russia, 2019); на XV, XVI и XVII Международных научно-технических конференциях «Проблемы информатики в образовании, управлении, экономике и технике» (Пенза, 2015, 2016, 2017); IX и X Международных научно-технических конференциях «Математическое и компьютерное моделирование естественно-научных и социальных проблем» (Пенза, 2015, 2016); XIII, XIV, XV и XVI Международных научно-технических конференциях «Новые информационные технологии и системы» (Пенза, 2016, 2017, 2018, 2019); XI, XII и XIII Международных научно-технических конференциях «Аналитические и численные методы моделирования естественно-научных и социальных проблем» (Пенза, 2016, 2017, 2018); XXV, XXVI и XXVII Всероссийских семинарах «Нейроинформатика, ее приложения и анализ данных» (Красноярск, 2017, 2018, 2019); XXI Международной научно-технической конференции «Нейроинформатика 2019» (Москва, 2019); Международной конференции «Telecommunications, Computing and Control» (Санкт-Петербург, 2019); Всероссийской конференции «Нейронные сети послезавтра: проблемы и перспективы» (Нижний Новгород, 2019).

Публикации. По результатам исследования опубликовано 25 работ, в том числе 2 статьи в изданиях, рекомендованных ВАК, 4 статьи в изданиях, индексируемых в международной базе Scopus. Опубликовано 18 материалов докладов на международных и всероссийских конференциях, индексируемых в базе данных РИНЦ, 1 статья опубликована в зарубежном рецензируемом журнале. Получено свидетельство о государственной регистрации комплекса программ для ЭВМ.

Личный вклад автора. Все основные научные результаты, приведенные в диссертации, получены автором самостоятельно. Из работ, выполнен-

ных в соавторстве, в диссертации использованы материалы, подготовленные лично автором. Автором разработаны новые градиентные алгоритмы первого порядка с аналитическим вычислением параметров сети, разработан алгоритм метода Левенберга – Марквардта с аналитическим расчетом параметров сети.

Структура и объем работы. Диссертация состоит из введения, четырех глав, заключения, библиографического списка из 140 наименований и одного приложения. Общий объем работы составляет 122 страницы, из которых 117 страниц занимает основной текст диссертации, включающий 23 рисунка и 4 таблицы.

ОСНОВНОЕ СОДЕРЖАНИЕ ДИССЕРТАЦИИ

Во введении обоснована актуальность темы диссертации; сформулированы цели и задачи исследования; отражены научная новизна работы и практическая значимость; перечислены методы исследования; приведены положения, выносимые на защиту.

В первой главе проведен анализ бессеточных методов численного решения краевых задач математической физики. ГОСТ Р 57188–2016 дает следующее определение бессеточных методов: «Численные методы, которые не требуют сетки точек, соединенных между собой для аппроксимации уравнений». Эти методы являются проекционными методами и дают приближенное дифференцируемое решение в виде разложения по системе базисных функций. Бессеточные методы производят аппроксимацию решения. Для этого в области решения выбирается множество фиксированных точек, в общем виде не располагающихся на сетке. С каждой точкой связывается своя базисная функция, равная единице в этой точке. Обычно используются РБ-функции, центры которых располагаются в этих точках. Выбирается множество пробных точек, в которых проверяются ошибки аппроксимации. Это множество в общем случае не совпадает со множеством центров РБ-функций. Для аппроксимации решения используются два метода: интерполяционный метод и среднеквадратический метод аппроксимации. В первом случае параметры разложения по базису вычисляются из условия равенства нулю ошибок (невязок) решения. Во втором случае параметры разложения, возможно, и параметры базисных функций определяются посредством минимизации функции ошибки, представляющей собой сумму квадратов невязок. Для вычисления невязок при решении краевых задач, описываемых уравнениями в частных производных, используется дифференцируемость базисных функций.

Радиальные базисные функции (РБ-функции) – это функции расстояния точки пространства от параметра функции, называемого центром функции: $\varphi(\|\mathbf{x} - \mathbf{c}\|, \mathbf{p})$, где $\mathbf{x}$ – точка пространства; $\mathbf{p}$ – вектор параметров функции; $\mathbf{c}$ – центр радиальной базисной функции; $\|\mathbf{x} - \mathbf{c}\|$ – евклидова норма (расстояние) между точкой и центром. Применяются различные РБ-функции. В работе используется функция Гаусса (Гауссиан)

$$\varphi(\|\mathbf{x} - \mathbf{c}\|, a) = \exp\left(-\frac{\|\mathbf{x} - \mathbf{c}\|^2}{2a^2}\right), \quad (1)$$

где $\mathbf{c}$ – положение центра функции; a – параметр формы, часто называемый шириной. Функция Гаусса выбрана, так как ее значения ограничены, а область определения совпадает с областью определения ее производных. Это удобно при решении уравнений в частных производных.

Краевая задача в операторной форме имеет вид

$$Lu(\mathbf{x}) = f(\mathbf{x}), \mathbf{x} \in \Omega, \quad Bu(\mathbf{x}) = p(\mathbf{x}), \mathbf{x} \in \partial\Omega, \quad (2)$$

где u – решение задачи; L – дифференциальный оператор; B – оператор граничных условий; Ω – область решения; $\partial\Omega$ – граница области; f и p – известные функции.

Внутри области решения и на границе задается множество пробных точек

$$\{\mathbf{x}_i \mid i=1,2,\dots,N \subset \Omega\} \cup \{\mathbf{x}_i \mid i=N_1+1,N,\dots,N+K \subset \partial\Omega\}, \quad (3)$$

где N – количество пробных точек во внутренней области Ω ; K – количество пробных точек на границе $\partial\Omega$.

При использовании РБ-функций решение задачи находится в виде взвешенной суммы базисных функций

$$u_{\text{RBF}}(\mathbf{x}) = \sum_{j=1}^{n_{\text{RBF}}} w_j \varphi_j(\mathbf{x}), \quad \mathbf{x} \in \bar{\Omega} = \Omega \cup \partial\Omega, \quad (4)$$

где φ_j – РБ-функция; w_j – вес, n_{RBF} – количество РБ-функций.

В (4) количество РБ-функций берется равным количеству пробных точек: $M = N + K$. Параметры РБ-функций задаются. Неизвестные коэффициенты в (4) находятся как решение системы линейных алгебраических уравнений с квадратной матрицей, которая получается из невязок задачи (2) в пробных точках (3) после подстановки (4) в (2). Для этого РБ-функция должна быть дифференцируема необходимое число раз. Решением сформированной системы является вектор весов $\mathbf{w}$. Матрица системы является несимметричной и плохо обусловленной, что затрудняет решение системы при большом числе пробных точек. Наиболее серьезным недостатком является неформализуемый подбор параметров РБ-функций. Этот недостаток преодолевается использованием сетей радиальных базисных функций.

РБФ-сеть включает два слоя. Первый слой состоит из РБ-функций, производящих нелинейное преобразование входного вектора $\mathbf{x} = [x_1, x_2, \dots, x_d]$ – координат точки, в которой вычисляется приближение к решению (d – размерность пространства). Второй слой РБФ-сети представляет собой линейный взвешенный сумматор

$$u(\mathbf{x}) = \sum_{m=1}^{n_{\text{RBF}}} w_m \varphi_m(\mathbf{x}; \mathbf{p}_m), \quad (5)$$

где n_{RBF} – количество РБФ; w_m – вес РБФ φ_m ; $\mathbf{p}_m$ – вектор параметров.

Решение краевой задачи (2) на РБФ-сети является аппроксимацией неизвестного решения на множестве пробных точек (3). Поскольку решение в пробных точках неизвестно, то возможна только минимизация невязок на

множестве пробных точек. Для минимизации используется метод наименьших квадратов. Функционал ошибки для поиска весов $\mathbf{w}$ и параметров $\mathbf{p}$ РБ-функций, минимизирующих невязки в пробных точках, имеет вид

$$I(\mathbf{w}, \mathbf{p}) = \sum_{i=1}^N [Lu_{\text{RBF}}(\mathbf{x}_i; \mathbf{w}, \mathbf{p}) - f(\mathbf{x}_i)]^2 + \lambda \sum_{i=N+1}^{N+K} [Bu_{\text{RBF}}(\mathbf{x}_i; \mathbf{w}, \mathbf{p}) - p(\mathbf{x}_i)]^2, \quad (6)$$

где $\mathbf{x}_i$ – пробные точки; λ – подбираемый штрафной множитель; u_{RBF} – приближенное решение (5), получаемое на РБФ-сети. Штрафной множитель λ обеспечивает выполнение граничных условий, так как в бессеточных методах условия на границе не фиксируются. Как видно из (6), использование РБФ-сетей позволяет вычислить не только веса, но и параметры РБ-функций (в случае функции Гаусса – координаты центров и ширину).

Абсолютное большинство алгоритмов обучения РБФ-сетей основано на градиентных методах оптимизации, среди которых выделяют методы нулевого, первого и второго порядков. Методы нулевого порядка используют только значения оптимизируемой функции. Методы первого порядка основаны на применении первых производных оптимизируемой функции (градиента функции). Методы второго порядка основаны на применении вторых производных (матрицы Гессе). В известных работах используется простейший метод первого порядка – метод градиентного спуска. Реализация метода градиентного спуска на примере двумерной задачи (2) и использования Гауссиана в качестве РБ-функций производится следующим образом. Рассмотрим вектор параметров РБФ-сети

$$\boldsymbol{\theta} = [w_1, w_2, \dots, w_{n_{\text{RBF}}}, c_{11}, c_{21}, \dots, c_{n_{\text{RBF}}1}, c_{12}, c_{22}, \dots, c_{n_{\text{RBF}}2}, a_1, a_2, \dots, a_{n_{\text{RBF}}}]^T, \quad (7)$$

где w_j , c_{j1} и c_{j2} , a_j – веса, координаты центров, ширина РБ-функции, $j = 1, 2, 3, \dots, n_{\text{RBF}}$, n_{RBF} – количество РБ-функций.

Коррекция вектора (7) на итерации k в методе градиентного спуска производится по формуле

$$\boldsymbol{\theta}^{(k+1)} = \boldsymbol{\theta}^{(k)} + \Delta \boldsymbol{\theta}^{(k+1)}, \quad (8)$$

где $\Delta \boldsymbol{\theta}^{(k+1)} = -\eta \nabla J(\boldsymbol{\theta}^{(k)})$ – вектор поправки параметров, η – подбираемая скорость обучения, $\nabla J(\boldsymbol{\theta}^{(k)})$ – вектор градиента функционала (6) по компонентам вектора $\boldsymbol{\theta}^{(k)}$ (7) на итерации k .

Вычисления по (8) заканчиваются при малом значении функционала ошибки (6). Метод градиентного спуска обладает низкой скоростью сходимости, что не позволяет решать задачи с приемлемой точностью и требует больших затрат времени.

Методы второго порядка пока не нашли практического применения в обучении РБФ-сетей при решении краевых задач. На кафедре «Компьютерные технологии» Пензенского государственного университета предложен

быстрый алгоритм обучения РБФ-сетей, основанный на эффективном методе оптимизации – методе доверительных областей. Но этот алгоритм сложен, так как минимум функционала ошибки отыскивается в ограниченных областях, что требует на каждом шаге процесса оптимизации решения условной задачи оптимизации.

Таким образом, РБФ-сети являются перспективным средством реализации бессеточных методов решения краевых задач. Важной проблемой использования РБФ-сетей является отсутствие быстрых и относительно несложных алгоритмов обучения.

Во второй главе разработаны алгоритмы обучения РБФ-сетей для решения задач аппроксимации функций. В процессе решения краевой задачи на РБФ-сети находится аппроксимированное решение задачи. Поэтому целесообразно начать исследования алгоритмов обучения РБФ-сетей с решения более простых задач аппроксимации функций, для которых известны целевые значения в пробных точках. Кроме того, применение РБФ-сетей для аппроксимации функций имеет самостоятельное значение.

Обучение РБФ-сети при решении задач аппроксимации представляет собой минимизацию функционала ошибки

$$I = \frac{1}{2} \sum_{j=1}^n e_j^2 = \frac{1}{2} \sum_{j=1}^n (u(\mathbf{x}_j) - T_j)^2, \quad (9)$$

где n – количество пробных точек; e_j – невязка (ошибка) аппроксимации в j -й пробной точке; $\mathbf{x}_j$ – вектор координат j -й пробной точки (для аппроксимации функции двух переменных $\mathbf{x}_j = [x_{j1}, x_{j2}]^T$); $u(\mathbf{x}_j)$ – приближенное значение функции (5) в j -й пробной точке, полученное на РБФ-сети; T_j – известное (целевое) значение функции в j -й пробной точке, множитель $1/2$ введен для упрощения вычислений.

Так как при обучении РБФ-сети настраиваются параметры различной природы, например при использовании функции Гаусса настраиваются веса, центры и ширина РБ-функций, то предложено применять различные значения параметров обучения для различных параметров РБ-функций. Для этого вектор градиента функционала (9) предложено представлять в блочной форме (номер итерации опущен):

$$\mathbf{g} = \nabla I = [\mathbf{g}_w \mid \mathbf{g}_{c_1} \mid \mathbf{g}_{c_2} \mid \mathbf{g}_a]^T, \quad (10)$$

где $\mathbf{g}_w$, $\mathbf{g}_{c_1}$, $\mathbf{g}_{c_2}$, $\mathbf{g}_a$ – компоненты вектора градиента по весам, координатам центра и ширине. Вектор поправки параметров в методе градиентного спуска также представляется в блочном виде:

$$\Delta \Theta^{(k+1)} = [\Delta \Theta_w^{(k+1)} \mid \Delta \Theta_{c_1}^{(k+1)} \mid \Delta \Theta_{c_2}^{(k+1)} \mid \Delta \Theta_a^{(k+1)}]^T, \quad (11)$$

где $\Delta\theta_w^{(k+1)} = -\eta_w \mathbf{g}_w(\theta^{(k)})$, $\Delta\theta_{c_1}^{(k+1)} = -\eta_{c_1} \mathbf{g}_{c_1}(\theta^{(k)})$, $\Delta\theta_{c_2}^{(k+1)} = -\eta_{c_2} \mathbf{g}_{c_2}(\theta^{(k)})$,
 $\Delta\theta_a^{(k+1)} = -\eta_a \mathbf{g}_a(\theta^{(k)})$.

В работе для обучения РБФ-сетей разработаны алгоритмы быстрых современных методов первого порядка. В алгоритме градиентного спуска с импульсом (моментом) поправка к вектору параметров формируется следующим образом:

$$\Delta\theta^{(k+1)} = \alpha\Delta\theta^{(k)} - \eta \mathbf{g}_\theta(\theta^{(k)}), \quad (12)$$

где η и α – подбираемые коэффициенты, принимающие значения в интервале $[0, 1]$, $\mathbf{g}_\theta(\theta^{(k)}) = \nabla I(\theta^{(k)})$ – вектор градиента функционала ошибки.

Вектор поправки параметров (12) ускоряет обновление параметров, если направление градиента совпадает с направлением поправки в предыдущей итерации, и замедляет обновление параметров в противном случае. В результате скорость сходимости возрастает и появляется возможность выхода из областей локального минимума и плоских участков функционала ошибки, в которых градиент близок к нулю. Для задания различных параметров обучения для различных параметров РБ-функций использованы блочные представления (10) и (11). Тогда компоненты вектора поправки (12) строятся по формулам

$$\Delta\theta_w^{(k+1)} = \alpha_w \Delta\theta_w^{(k)} - \eta_w \mathbf{g}_w(\theta^{(k)}), \quad \Delta\theta_{c_1}^{(k+1)} = \alpha_{c_1} \Delta\theta_{c_1}^{(k)} - \eta_{c_1} \mathbf{g}_{c_1}(\theta^{(k)}),$$

$$\Delta\theta_{c_2}^{(k+1)} = \alpha_{c_2} \Delta\theta_{c_2}^{(k)} - \eta_{c_2} \mathbf{g}_{c_2}(\theta^{(k)}), \quad \Delta\theta_a^{(k+1)} = \alpha_a \Delta\theta_a^{(k)} - \eta_a \mathbf{g}_a(\theta^{(k)}).$$

Алгоритм метода Нестерова можно представить в виде алгоритма метода с импульсом с модифицированным вычислением градиента

$$\theta^{(k+1)} = \theta^{(k)} + \Delta\theta^{(k+1)}, \quad \Delta\theta^{(k+1)} = \alpha\Delta\theta^{(k)} - \eta \mathbf{g}_\theta(\theta^{(k)} + \alpha\Delta\theta^{(k)}),$$

где α и η – подбираемые коэффициенты, $\mathbf{g}_\theta(\theta^{(k)} + \alpha\Delta\theta^{(k)}) = \nabla I(\theta^{(k)} + \alpha\Delta\theta^{(k)})$ – вектор градиента функционала ошибки в промежуточной точке. Для задания различных параметров обучения для различных параметров РБ-функций и различных коэффициентов коррекции вектора градиента использованы блочные представления (10) и (11).

В алгоритмах с адаптивной скоростью обучения скорость обучения пересчитывается в каждой итерации. В диссертации выбран хорошо зарекомендовавший себя алгоритм RMSProp (Root Mean Square Propagation):

$$\mathbf{s}^{(k+1)} = \beta \mathbf{s}^{(k)} + (1 - \beta) \mathbf{g}(\Theta^{(k)}) \otimes \mathbf{g}(\Theta^{(k)}), \quad \Theta^{(k+1)} = \Theta^{(k)} - \eta \mathbf{g}(\Theta^{(k)}) / \sqrt{\mathbf{s}^{(k+1)} + \varepsilon},$$

где $\mathbf{g}(\Theta^{(k)})$ – вектор градиента; Θ – вектор параметров; $\otimes$ – поэлементное умножение; η – подбираемый коэффициент; ε – сглаживающий член, для из-

бегания деления на ноль, обычно $\varepsilon \approx 10^{-10}$; ./ – поэлементное деление, $\beta \approx 0,9$ – коэффициент ослабления.

Структура РБФ-сети и дифференцируемость РБ-функций позволяют вычислять аналитически компоненты градиента функционала ошибки, что позволяет ускорить вычисления. В диссертации получены формулы аналитического вычисления компонентов градиента функционала. Для аппроксимации функции двух переменных компоненты градиента равны

$$\frac{\partial I}{\partial w_i} = \sum_{j=1}^n (u(\mathbf{x}_j) - T_j) \phi_i(\mathbf{x}_j), \quad \frac{\partial I}{\partial c_{i1}} = w_i \sum_{j=1}^n (u(\mathbf{x}_j) - T_j) \phi_i(\mathbf{x}_j) \frac{x_{j1} - c_{i1}}{a_i^2},$$

$$\frac{\partial I}{\partial c_{i2}} = w_i \sum_{j=1}^n (u(\mathbf{x}_j) - T_j) \phi_i(\mathbf{x}_j) \frac{x_{j2} - c_{i2}}{a_i^2}, \quad \frac{\partial I}{\partial a_i} = w_i \sum_{j=1}^n (u(\mathbf{x}_j) - T_j) \phi_i(\mathbf{x}_j) \frac{\|\mathbf{x}_j - \mathbf{c}_i\|}{a_i^3},$$

где c_{i1} и c_{i2} – координаты центра i -й РБ-функции; x_{j1} и x_{j2} – координаты пробной точки $\mathbf{x}_j$; $\|\mathbf{x}_j - \mathbf{c}_i\|$ – евклидова норма.

В работе для обучения РБФ-сетей разработан градиентный алгоритм второго порядка – алгоритм метода Левенберга – Марквардта. Рассмотрим применение алгоритма при аппроксимации функций двух переменных. Коррекция вектора параметров Θ (7) в k -м цикле (итерации) обучения описывается формулой (8), в которой вектор поправки $\Delta\Theta^{(k)}$ является решением системы линейных алгебраических уравнений

$$(\mathbf{J}_{k-1}^T \mathbf{J}_{k-1} + \mu_k \mathbf{E}) \Delta\Theta^{(k)} = -\mathbf{g}_{k-1}, \quad (13)$$

где $\mathbf{E}$ – единичная матрица; μ_k – параметр регуляризации; $\mathbf{g} = \mathbf{I}^T \mathbf{e}$ – градиент функционала ошибки (9) по вектору параметров θ ; $\mathbf{e} = [e_1 \ e_2 \ \dots \ e_n]^T$ – вектор ошибок, используемый при вычислении функционала ошибки (9), $\mathbf{J}_{k-1}$ – матрица Якоби, которая вычисляется по значениям параметров сети в $k-1$ итерации.

Матрицу Якоби предложено представить в блочном виде $\mathbf{J} = [\mathbf{J}_w \mid \mathbf{J}_{c_1} \mid \mathbf{J}_{c_2} \mid \mathbf{J}_a]$, где элементы матриц $\mathbf{J}_w$, $\mathbf{J}_{c_1}$, $\mathbf{J}_{c_2}$ и $\mathbf{J}_a$ представляют собой частные производные ошибок по весам и параметрам РБ-функций. С учетом структуры РБФ-сети и дифференцируемости РБ-функций элементы матрицы $\mathbf{J}$ вычисляются аналитически:

$$\frac{\partial e_i}{\partial w_j} = \frac{\partial u(\mathbf{x}_i)}{\partial w_j} = \phi_j(\mathbf{x}_i), \quad \frac{\partial e_i}{\partial c_{j1}} = w_j \phi_j(\mathbf{x}_i) \frac{x_{i1} - c_{j1}}{a_j^2}, \quad \frac{\partial e_i}{\partial c_{j2}} = w_j \phi_j(\mathbf{x}_i) \frac{x_{i2} - c_{j2}}{a_j^2},$$

$$\frac{\partial e_i}{\partial a_j} = w_j \phi_j(\mathbf{x}_i) \frac{\|\mathbf{x}_i - \mathbf{c}_j\|^2}{a_j^3},$$

где $\phi_j(\mathbf{x}_i)$ – значение j -й радиальной базисной функции в точке $\mathbf{x}_i$.

Разработанный алгоритм Левенберга – Марквардта обучения РБФ-сетей для решения задач аппроксимации в форме псевдокода имеет вид

```

while критерий останова не выполнен do
    вычислить вектор ошибок  $\mathbf{e}$ 
    вычислить матрицы  $\mathbf{J}_{c_1}$ ,  $\mathbf{J}_{c_2}$  и  $\mathbf{J}_a$ , сформировать матрицу  $\mathbf{J}$ 
    решить систему  $(\mathbf{J}_{k-1}^T \mathbf{J}_{k-1} + \mu_k \mathbf{E}) \Delta \theta^{(k)} = -\mathbf{J}^T \mathbf{e}_{k-1}$ 
    обновить вектор параметров  $\Theta$ 
    скорректировать параметр  $\mu$ 
end while

```

Недостатком метода Левенберга – Марквардта является плохая обусловленность системы (13). Число обусловленности системы зависит от начальных значений ширины РБ-функций и точности вычислений. Для улучшения обусловленности используется параметр регуляризации μ . Но по мере уменьшения погрешности обучения значение параметра уменьшается, что ухудшает обусловленность. В диссертации для формирования вектора параметров РБФ-сети исследовано использование итерационных методов подпространств Крылова. Эти методы за счет численной устойчивости позволяют ослабить негативное влияние плохой обусловленности системы. Ввиду положительной определенности матрицы системы (13) можно использовать метод сопряженных градиентов. Уменьшение параметра регуляризации в процессе обучения РБФ-сети ухудшает обусловленность системы (13), поэтому требуется подбор достаточно большого начального значения этого параметра. Применение итерационных методов снижает требования к точности подбора параметра регуляризации. Еще больший эффект, как показали эксперименты, дает применение устойчивого метода бисопряженных градиентов. Необходимость выполнения итерационного процесса решения системы (13) на каждой итерации метода Левенберга – Марквардта существенно увеличивает время решения (на 15–20 %), так как предлагаемые алгоритмы являются быстрыми, а решение системы может производиться с относительной нормой невязки $10^{-2} - 10^{-3}$.

Третья глава посвящена разработке алгоритмов обучения РБФ-сетей для решения краевых задач. Так как решение ДУЧП является, в сущности, аппроксимацией неизвестного решения, то разработанные для аппроксимации функций алгоритмы могут быть применены для решения краевых задач. Решение краевой задачи на РБФ-сети рассмотрим на примере модельной задачи, описываемой уравнением Пуассона с граничными условием Дирихле:

$$\frac{\partial^2 u}{\partial x_1^2} + \frac{\partial^2 u}{\partial x_2^2} = f(x_1, x_2), \quad (x_1, x_2) \in \Omega, \quad u = p(x_1, x_2), \quad (x_1, x_2) \in \partial\Omega, \quad (14)$$

где $\partial\Omega$ – граница области; f и p – известные функции координат (x, y) .

Функционал ошибки (6) для модельной задачи имеет вид

$$I = \left[\sum_{i=1}^N (\Delta u_i - f_i)^2 + \lambda \sum_{j=1}^K (u_j - p_j)^2 \right], \quad (15)$$

где Δu_i – Лапласиан в точке i .

Так как Гауссиан является дифференцируемой функцией, то Лапласиан и компоненты вектора градиента функционала ошибки предложено вычислять в аналитическом виде. Поэтому реализация на РБФ-сетях градиентных алгоритмов первого порядка при решении краевых задач принципиально не отличается от реализации этих алгоритмов для решения задачи аппроксимации функций, что позволяет реализовать не применявшиеся ранее для РБФ-сетей алгоритм моментов и алгоритм Нестерова. Выражение для Лапласиана имеет вид

$$\Delta u_i = \frac{\partial^2 u_i}{\partial x_{i1}^2} + \frac{\partial^2 u_i}{\partial x_{i2}^2} = \sum_{k=1}^{n_{\text{RBF}}} \left[w_k e^{-\frac{\|\mathbf{x} - \mathbf{c}_k\|^2}{2a_k^2}} \left(\frac{\|\mathbf{x}_i - \mathbf{c}_k\|^2 - 2a_k^2}{a_k^4} \right) \right].$$

Компоненты градиента по весам сети вычисляются по формуле

$$\frac{\partial I}{\partial w_p} = \sum_{i=1}^N (\Delta v_i - f_i) \varphi_p(\mathbf{x}_i) \frac{\|\mathbf{x}_i - \mathbf{c}_p\|^2 - 2a_p^2}{a_p^4} + \lambda \sum_{j=1}^K (v_j - p_j) w_p \varphi_p(\mathbf{x}_j).$$

Компоненты градиента по координате c_{p1} центров РБФ-функций вычисляются по формуле

$$\frac{\partial I}{\partial c_{p1}} = w_p \sum_{i=1}^N (\Delta v_i - f_i) (x_{i1} - c_{p1}) \frac{\|\mathbf{x}_i - \mathbf{c}_p\|^2 - 4a_p^2}{a_p^6} \varphi_p(\mathbf{x}_i) + \lambda w_p \sum_{j=1}^K (v_j - p_j) \frac{(x_{j1} - c_{p1})}{a_p^2} \varphi_p(\mathbf{x}_j).$$

Аналогичный вид имеют компоненты градиента по координате c_{p2} . Компонент градиента по ширине равен

$$\begin{aligned} \frac{\partial I}{\partial a_p} = w_p \sum_{i=1}^N (\Delta v_i - f_i) \frac{\|\mathbf{x}_i - \mathbf{c}_p\|^4 - 6a_p \|\mathbf{x}_i - \mathbf{c}_p\|^2 + 4a_p^4}{a_p^7} \varphi_p(\mathbf{x}_i) + \\ + \lambda w_p \sum_{j=1}^K (v_j - p_j) \frac{\|\mathbf{x}_j - \mathbf{c}_p\|^2}{a_p^3} \varphi_p(\mathbf{x}_j). \end{aligned}$$

Процесс обучения сети заканчивается при малом значении функционала ошибки (15) или средней квадратической погрешности.

Разработанный алгоритм метода Левенберга – Марквардта рассмотрим на примере задачи (14). Фигурирующими в матрице Якоби ошибками являются невязки r_i в пробных точках $\mathbf{x}_i$. Матрица Якоби представляется в блочном виде

$$\mathbf{J} = \begin{bmatrix} \mathbf{J}_w & | & \mathbf{J}_{c_1} & | & \mathbf{J}_{c_2} & | & \mathbf{J}_a \end{bmatrix}. \text{ Поэлементное представление матрицы имеет вид}$$

$$\mathbf{J}_w = \begin{bmatrix} \frac{\partial r_1}{\partial w_1} & \dots & \frac{\partial r_1}{\partial w_{n_{\text{RBF}}}} \\ \frac{\partial r_2}{\partial w_1} & \dots & \frac{\partial r_2}{\partial w_{n_{\text{RBF}}}} \\ \dots & \dots & \dots \\ \frac{\partial r_n}{\partial w_1} & \dots & \frac{\partial r_n}{\partial w_{n_{\text{RBF}}}} \end{bmatrix}, \quad \mathbf{J}_{c_1} = \begin{bmatrix} \frac{\partial r_1}{\partial c_{11}} & \dots & \frac{\partial r_1}{\partial c_{n_{\text{RBF}}1}} \\ \frac{\partial r_2}{\partial c_{11}} & \dots & \frac{\partial r_2}{\partial c_{n_{\text{RBF}}1}} \\ \dots & \dots & \dots \\ \frac{\partial r_n}{\partial c_{11}} & \dots & \frac{\partial r_n}{\partial c_{n_{\text{RBF}}1}} \end{bmatrix},$$

$$\mathbf{J}_{c_2} = \begin{bmatrix} \frac{\partial r_1}{\partial c_{12}} & \dots & \frac{\partial r_1}{\partial c_{n_{\text{RBF}}2}} \\ \frac{\partial r_2}{\partial c_{12}} & \dots & \frac{\partial r_2}{\partial c_{n_{\text{RBF}}2}} \\ \dots & \dots & \dots \\ \frac{\partial r_n}{\partial c_{12}} & \dots & \frac{\partial r_n}{\partial c_{n_{\text{RBF}}2}} \end{bmatrix}, \quad \mathbf{J}_a = \begin{bmatrix} \frac{\partial e_1}{\partial a_1} & \dots & \frac{\partial e_1}{\partial a_{n_{\text{RBF}}}} \\ \frac{\partial e_2}{\partial a_1} & \dots & \frac{\partial e_2}{\partial a_{n_{\text{RBF}}}} \\ \dots & \dots & \dots \\ \frac{\partial e_n}{\partial a_1} & \dots & \frac{\partial e_n}{\partial a_{n_{\text{RBF}}}} \end{bmatrix},$$

где r_i – невязка в i -й пробной точке; $n = N + K$ – суммарное количество пробных точек внутри и на границе области решения.

Для элементов матрицы Якоби получены аналитические выражения. Элементы матрицы $\mathbf{J}_w$ для внутренних пробных точек имеют вид

$$\frac{\partial r_i}{\partial w_p} = \frac{\partial(\Delta v_i - f_i)}{\partial w_p} = \frac{\|\mathbf{x}_i - \mathbf{c}_p\|^2 - 2a_p^2}{a_p^4} \varphi_p(\mathbf{x}_i).$$

Для граничных пробных точек $\frac{\partial r_i}{\partial w_p} = \varphi_p(\mathbf{x}_i)$.

Элементы матрицы $\mathbf{J}_{c_1}$ для внутренних пробных точек имеют вид

$$\frac{\partial r_i}{\partial c_{p1}} = \frac{w_p}{a_j^4} (x_{i1} - c_{p1}) \frac{\|\mathbf{x}_i - \mathbf{c}_p\|^2 - 4a_p^2}{a_p^2} \varphi_p(\mathbf{x}_i).$$

Для граничных точек элементы матрицы записываются в виде

$$\frac{\partial r_i}{\partial c_{p1}} = w_p \frac{(x_{i1} - c_{p1})}{a_p^2} \varphi_p(\mathbf{x}_i).$$

Аналогично вычисляются элементы матрицы $\mathbf{J}_{c_2}$. Элементы матрицы $\mathbf{J}_a$ для внутренних пробных точек вычисляются по формуле

$$\frac{\partial r_i}{\partial a_p} = \frac{w_p}{a_p^5} \varphi_p(\mathbf{x}_i) \left[\frac{\|\mathbf{x}_i - \mathbf{c}_p\|^2}{a_p^2} \left(\|\mathbf{x}_i - \mathbf{c}_p\|^2 - 2a_p^2 \right) - 4 \left(\|\mathbf{x}_i - \mathbf{c}_p\|^2 - a_p^2 \right) \right].$$

Для граничных точек элементы матрицы записываются в виде

$$\frac{\partial r_i}{\partial a_p} = w_p \frac{\|\mathbf{x}_i - \mathbf{c}_p\|^2}{a_p^3} \varphi_p(\mathbf{x}_i).$$

В приведенных формулах r_i – невязка в точке i . Условия завершения процесса обучения методом Левенберга – Марквардта те же, что и в методах первого порядка.

В диссертации рассмотрено применение разработанных алгоритмов в унифицированном подходе к численному решению различных классов краевых задач моделирования объектов с распределенными параметрами: нелинейных задач, нестационарных задач, обратных задач.

В четвертой главе представлены результаты экспериментального исследования разработанных алгоритмов обучения РБФ-сетей для решения краевых задач. Для экспериментального исследования разработанных алгоритмов создан комплекс программ в системе MATLAB для решения двумерных задач аппроксимации функций и решения краевых задач. Комплекс программ включает главные программы, позволяющие реализовать задание параметров сети, организовать обучение сети, вывод кривых изменения функционалов ошибки в процессе обучения и вывод результатов решения задач, и программы, вычисляющие параметры РБФ-сетей и алгоритмов обучения.

Эксперименты проводились на компьютере со следующими характеристиками: процессор Intel Core i5-8500, частота – 3,0 GHz, оперативная память – 16,0 GB.

Эксперименты показали, что для несложной задачи аппроксимации функции одной переменной $\sin$ с успехом можно использовать для обучения РБФ-сети как метод градиентного спуска, так и метод Нестерова. Метод Нестерова обладает некоторым преимуществом, которое для этой задачи не является решающим.

Вторая задача аппроксимации является аппроксимацией функции $z = x^2 + y^2$. 100 узлов интерполяции располагались случайным образом в области $(x = -3 \dots + 3, y = -3 \dots + 3)$. Количество РБ-функций равно 16. При инициализации сети функции располагались в узлах сетки. На рисунке 1 показано расположение центров РБ-функций, в виде радиусов кругов показана ширина, значения весов условно показаны заливкой кругов. Рисунок 1 показывает важность настройки не только весов, но и центров и ширины РБ-функций, что можно реализовать с использованием РБФ-сетей. Сеть обучалась с использованием алгоритмов первого и второго порядка до достижения значения средней квадратической ошибки, равной 0,01. Ускоренные методы первого порядка показали большое преимущество перед алгоритмом градиентного спуска. Импульсный метод сократил число итераций обучения в 3 раза, но оказался очень чувствителен к случайным изменениям начальных параметров. Лучшие и наиболее стабильные результаты показал метод Нестерова. За счет более

точного вычисления вектора градиента и несущественного усложнения алгоритма метод позволил уменьшить число итераций обучения в 175 раз, а время решения задачи – в 170 раз.



а) при инициализации сети б) после обучения сети методом Левенберга – Марквардта

Рисунок 1 – Расположение центров, значения весов и ширины РБ-функций при решении второй задачи аппроксимации

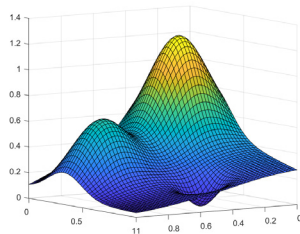
Алгоритм RMSProp работает устойчиво, но показал существенно меньшую скорость сходимости, чем метод Нестерова, сократив число итераций в 6 раз по сравнению с градиентным спуском. Метод Левенберга – Марквардта на порядок сокращает число итераций по сравнению с методом Нестерова, уменьшив число итераций в 35 раз по сравнению с методом Нестерова, а время решения задачи – в 17 раз. Методы первого порядка не позволили достичь средней квадратической ошибки, меньшей 0,01. Метод Левенберга – Марквардта позволил за 20 итераций получить среднюю квадратическую погрешность, равную 10^{-6} .

Третьей задачей аппроксимации была аппроксимация популярной в теории аппроксимации мультимодальной функции Франке:

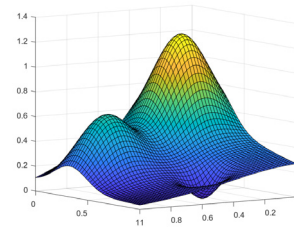
$$f(x, y) = 0,75 \exp \left(-\frac{(9x-2)^2}{4} - \frac{(9y-2)^2}{4} \right) + 0,75 \exp \left(-\frac{(9x+1)^2}{49} - \frac{9y+1}{10} \right) + \\ + 0,5 \exp \left(-\frac{(9x-7)^2}{4} - \frac{(9y-3)^2}{4} \right) - 0,2 \exp \left(-(9x-4)^2 - (9x-7)^2 \right).$$

Узлы интерполяции в количестве 100 штук располагались случайным образом в области $(x = 0 \dots 1, y = 0 \dots 1)$. Аппроксимация методами первого порядка не позволила получить среднюю квадратическую погрешность менее 10^{-1} . При этом результат даже визуально не соответствовал аппроксимируемой функции. Метод Левенберга – Марквардта позволил достичь средней квадратической погрешности, равной 10^{-6} , за 15 итераций. Количество РБ-функций равно 16. Результаты аппроксимации показаны на рисунке 2 и демонстрируют полное визуальное совпадение аппроксимированной функции с аналитическим представлением. На рисунке 3 показаны расположения центров, условные обозначения ширины РБ-функций и весов при инициализации и после обучения сети. Рисунок 3 свидетельствует о важности настройки не только весов сети, но и параметров РБ-функций. При аппроксимации функции Франке исследовано применение для решения системы (13) итерационных методов подпро-

пространств Крылова. Для решения системы уравнений применены метод сопряженных градиентов и устойчивый метод бисопряженных градиентов. Как показали эксперименты, достаточно решать систему до значения относительной нормы невязки, равного 10^{-2} , что несущественно увеличивает время решения. Высокая численная устойчивость методов подпространств Крылова, особенно устойчивого метода бисопряженных градиентов, ослабляет негативное влияние плохой обусловленности системы и упрощает подбор параметра регуляризации.

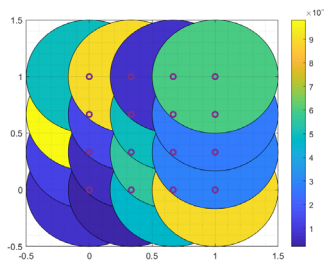


а) аналитически вычисленная функция

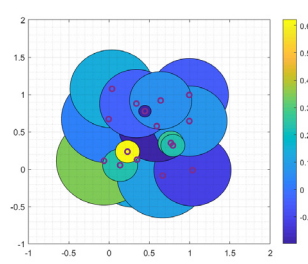


б) результат аппроксимации

Рисунок 2 – Аппроксимация функции Франке



а) при инициализации сети



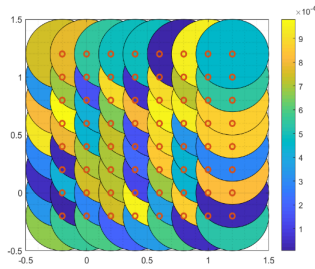
б) после обучения сети методом Левенберга – Марквардта

Рисунок 3 – Расположение центров, значения весов и ширины РБ-функций при аппроксимации функции Франке

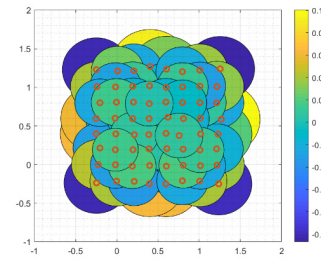
Эксперименты по решению краевых задач проводились на двух задачах. Первой являлась задача (14) при $f(x_1, x_2) = \sin(\pi x_1) \sin(\pi x_2)$ и $p(x_1, x_2) = 0$. Она решалась в единичном квадрате. Количество внутренних пробных точек $N = 100$. Количество граничных пробных точек $K = 40$. Пробные точки располагались случайным образом в области решения и на границе области. При инициализации центры РБ-функций располагались на квадратной сетке 8×8 . Расположение центров, условное представление ширины и величины весов перед обучением сети (рисунок 4, а) и после обучения (рисунок 4, б) показывают важность настройки параметров РБФ-функций. Зависимость среднеквадратической ошибки различных алгоритмов от номера итерации при решении первой краевой задачи показана на рисунке 5.

Как видно из рисунка 5, метод градиентного спуска позволил решить первую модельную задачу с небольшой точностью. Для решения с большой точностью метод практически не применим. Несколько бóльшую точность обеспечивает метод Нестерова. Только метод Левенберга – Марквардта позволил решить задачу с высокой точностью за приемлемое время. Метод Левен-

берга – Марквардта позволил решить задачу за то же число итераций, что и метод доверительных областей, но реализация метода Левенберга – Марквардта проще, что позволило в 1,5 раза сократить время решения.



а) при инициализации сети



б) после обучения сети методом Левенберга – Марквардта

Рисунок 4 – Центры, веса и ширина РБ-функций при решении первой краевой задачи

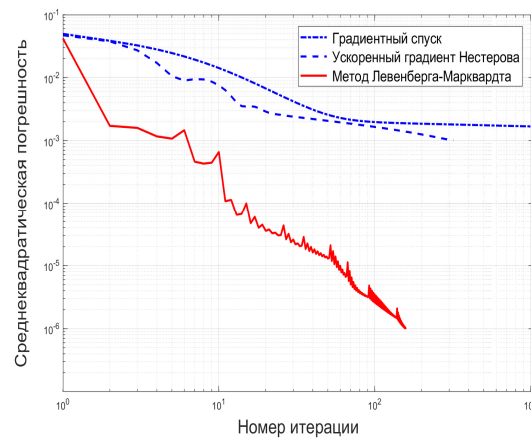


Рисунок 5 – Зависимости среднеквадратической ошибки различных алгоритмов от номера итерации

Вторая задача также описывается (14), но правая часть и граничные условия рассчитаны так, чтобы задача имела аналитическое мультимодальное решение:

$$u(x, y) = 3(x-1)^2 e^{-x^2-(y+1)^2} - 10 \left(\frac{x}{5} - x^3 - y^5 \right) e^{-x^2-y^2} - \frac{1}{3} e^{-(x+1)^2-y^2}.$$

Задача решалась в квадрате $[-3, 3]^2$. Задачу не удалось решить методами первого порядка. Разработанный алгоритм метода Левенберга – Марквардта позволил за 20 итераций и 11 с решить задачу до средней квадратической ошибки 10^{-6} (рисунок 6).

Вторая задача решалась при следующих условиях: количество внутренних пробных точек $N = 200$, количество граничных пробных точек $K = 80$, штрафной коэффициент функционала ошибки по границе $\lambda = 100$. Пробные точки располагались случайным образом в области решения и на границе области. Веса инициализировались как и при решении первой краевой задачи. Начальная ширина всех РБФ была постоянной, равной 0,6 для всех методов. При инициализации центры РБ-функций располагались на квадратной сетке 16×16 .

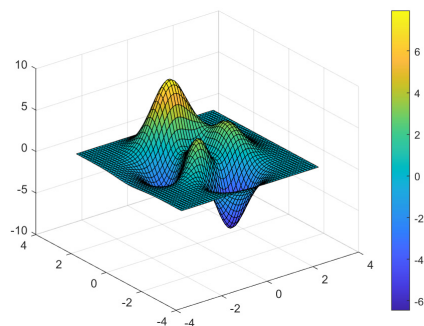


Рисунок 6 – Результаты численного решения второй краевой задачи

Таким образом, разработанные алгоритмы обучения РБФ-сетей, особенно алгоритм метода Левенберга – Марквардта, показали преимущество перед известными алгоритмами.

В заключении сформулированы основные результаты диссертационной работы.

Приложения содержат копию свидетельства о регистрации программного комплекса и копию акта о внедрении результатов диссертационной работы.

ОСНОВНЫЕ РЕЗУЛЬТАТЫ РАБОТЫ

В диссертации решена научно-техническая задача, заключающаяся в разработке новых алгоритмов повышенной производительности обучения сетей радиальных базисных функций при численном решении краевых задач, являющихся математическими моделями объектов с распределенными параметрами.

Основные результаты работы состоят в следующем.

1. Проведен анализ существующих численных методов решения краевых задач, описываемых дифференциальными уравнениями в частных производных, с использованием радиальных базисных функций и сетей радиальных базисных функций. Выявлены преимущества использования сетей радиальных базисных функций. Показано, что известные методы обучения сетей радиальных базисных функций не обеспечивают быстрого обучения сетей радиальных базисных функций. В качестве пути устранения этого недостатка предложено совершенствовать алгоритмы обучения сетей.

2. Для обучения РБФ-сетей при решении задач аппроксимации функций и краевых задач математического моделирования объектов с распределенными параметрами разработаны численные алгоритмы на основе современных быстрых градиентных методов первого порядка: алгоритм градиентного спуска с импульсом, алгоритм метода Нестерова, алгоритм с адаптивной скоростью обучения RMSProp. Разработанные алгоритмы учитывают архитектуру сетей радиальных базисных функций и используют управляемые параметры обучения для различных параметров сети. Получены аналитические выражения для вычисления градиента функционала ошибки. При решении модельных задач до средней квадратической ошибки, равной 0,01, ускоренный алгоритм Нестерова обеспечил сокращение числа итераций на два порядка по сравнению с используемым в настоящее время алгоритмом градиентного спуска.

3. Для обучения РБФ-сетей, предназначенных для решения задач аппроксимации функций и краевых задач математического моделирования объектов с распределенными параметрами, разработан численный алгоритм метода второго порядка – метода Левенберга – Марквардта, отличающийся учетом специфики архитектуры сети и аналитическим вычислением параметров. Метод позволил на модельных задачах достичь средней квадратической погрешности 10^{-6} , не достижимой известными алгоритмами первого порядка. Предложенный алгоритм позволяет получить такую погрешность за число итераций, равное числу итераций наиболее быстрого метода доверительных областей, но за счет более простого алгоритма сокращает время решения задачи в 1,5 раза.

4. В алгоритме метода Левенберга – Марквардта предложено применение методов подпространств Крылова для вычисления вектора коррекции компонентов РБФ-сети, позволившее при увеличении времени решения на 15–20 % снизить требования к выбору начального значения параметра регуляризации.

5. Проведен анализ возможностей решения на РБФ-сетях различных краевых задач математического моделирования объектов с распределенными параметрами, показавший возможность унифицированного подхода к решению таких задач с применением разработанных алгоритмов обучения РБФ-сетей.

6. На основе предложенных алгоритмов разработан комплекс программ решения двумерных задач аппроксимации функций и краевых задач для дифференциальных уравнений в частных производных второго порядка. Комплекс программ использован для экспериментального исследования разработанных алгоритмов и позволил сравнить различные алгоритмы обучения РБФ-сетей.

СПИСОК ОПУБЛИКОВАННЫХ РАБОТ ПО ТЕМЕ ДИССЕРТАЦИИ

Публикации в изданиях, рекомендованных ВАК РФ

1. Алкезуини, М. М. Совершенствование алгоритмов обучения сетей радиальных базисных функций для решения задач аппроксимации / М. М. Алкезуини, В. И. Горбаченко // Модели, системы, сети в экономике, технике, природе и обществе. – 2017. – № 3 (23). – С. 123–138.

2. Алкезуини, М. М. Моделирование объектов с распределенными параметрами на нейронных сетях / В. И. Горбаченко, М. М. Алкезуини // Модели, системы, сети в экономике, технике, природе и обществе. – 2019. – № 4 (32). – С. 50–64.

Публикации в изданиях, индексируемых в международных базах данных Scopus и Web of Science

3. Alqezweeni, M. M. Application of parametric identification method and radial basis function networks for solution of inverse boundary value problems / V. I. Gorbachenko, M. M. Alqezweeni, M. S. Jaafar // 2017 Annual Conference on New Trends in Information and Communications Technology Applications (Baghdad, Iraq, 7–9 March 2017). – Baghdad, Iraq, 2017. – P. 18–21.

4. Alqezweeni, M. M. Efficient Solving of Boundary Value Problems Using Radial Basis Function Networks Learned by Trust Region Method / M. M. Alqezweeni, V. I. Gorbachenko, M. V. Zhukov, M. S. Jaafar // Hindawi. International Journal of Mathematics and Mathematical Sciences. – 2018. – Vol. 2018. – Article ID 9457578.

5. Alqezweeni, M. Improving algorithms for learning of radial basis functions networks for approximation problems and solving partial differential equations / V. Gorbachenko, M. Alqezweeni // 2019 International Conference on Applied and Engineering Mathematics (Taxila, Pakistan, 27–29 August 2019). – Taxila, Pakistan, 2019. – P. 264–268.

6. Alqezweeni, M. M. Learning Radial Basis Functions Networks in Solving Boundary Value Problems / V. I. Gorbachenko, M. M. Alqezweeni // 2019 International Russian Automation Conference (Sochi, Russia, 8–14 September, 2019). – Sochi, 2019. – P. 1–6.

Публикации в других изданиях

7. Alkazweeny, M. Fundamental solutions method for coupled mathematical models of heat and mass transfers / V. Gorbachenko, O. Yaremko, N. Yaremko, M. Alkazweeny // International Journal of Pure and Applied Mathematics. – 2018. – Vol. 118, № 3. – P. 637–649.

8. Alqezweeni, M. M. Solution partial differential equation on radial basis function networks // Математическое и компьютерное моделирование естественно-научных и социальных проблем : сб. ст. IX Междунар. науч.-техн. конф. молодых специалистов, аспирантов и студентов (г. Пенза, 20–22 мая 2015 г.) / под ред. И. В. Бойкова. – Пенза : Изд-во ПГУ, 2015. – С. 276–281.

9. Алкезуини, М. М. Решение уравнений в частных производных на сетях радиальных базисных функций / М. М. Алкезуини, В. И. Горбаченко, М. В. Жуков // Проблемы информатики в образовании, управлении, экономике и технике : сб. ст. XV Междунар. науч.-техн. конф. (г. Пенза, 12–13 ноября, 2015 г.). – Пенза : Приволжский дом знаний, 2015. – С. 58–66.

10. Алкезуини, М. М. Обучение сетей радиальных базисных функций для решения краевых задач математической физики / М. М. Алкезуини, В. И. Горбаченко, М. В. Жуков // Проблемы информатики в образовании, управлении, экономике и технике : сб. ст. XVI Междунар. науч.-техн. конф. (г. Пенза, 17–18 ноября 2016 г.). – Пенза : Приволжский дом знаний, 2016. – С. 103–109.

11. Алкезуини, М. М. Обучение сетей радиальных базисных функций при решении задач аппроксимации / М. М. Алкезуини, В. И. Горбаченко // Новые информационные технологии и системы : тр. XIII Междунар. науч.-техн. конф. (г. Пенза, 23–25 ноября 2016 г.). – Пенза : Изд-во ПГУ, 2016. – С. 34–36.

12. Алкезуини, М. М. Алгоритм обучения методом доверительных областей сетей радиальных базисных функций для решения краевых задач математической физики / М. М. Алкезуини, В. И. Горбаченко, М. В. Жуков // Аналитические и численные методы моделирования естественнонаучных и социальных проблем : материалы XI Междунар. науч.-техн. конф. (г. Пенза, 6–9 декабря 2016 г.) / под ред. проф. И. В. Бойкова. – Пенза : Изд-во ПГУ, 2016. – С. 168–173.

13. Alqezweeni, M. M. Approximation of Functions Through Radial Basis Function Networks / V. I. Gorbachenko, M. M. Alqezweeni, M. S. Jaafar // Математическое и компьютерное моделирование естественно-научных и социальных проблем : сб. ст. X Междунар. науч.-техн. конф. молодых специалистов, аспирантов и студентов (г. Пенза, 23–27 мая 2016 г.) / под ред. И. В. Бойкова. – Пенза : Изд-во ПГУ, 2016. – С. 281–286.

14. Alqezweeni, M. M. Solving boundary value problems using radial basis functions networks learned by trust region method / V. I. Gorbachenko, M. V. Zhukov, M. M. Alqezweeni // International conference on mathematical modelling in applied sciences (Saint Petersburg, 24–28 July, 2017 г.). – St. Petersburg : Peter the Great St. Petersburg Polytechnic University, 2017. – P. 97–98.

15. Alqezweeni, M. M. Kernel basis functions method for coupled mathematical models of heat and mass transfers / V. I. Gorbachenko, O. E. Iaremko, M. M. Alqezweeni // International conference on mathematical modelling in applied sciences (Saint Petersburg, 24–28 July, 2017 г.). – St. Petersburg : Peter the Great St. Petersburg Polytechnic University, 2017. – P. 79–80.

16. Алкезуини, М. М. Ускоренные методы первого порядка обучения сетей радиальных базисных функций / М. М. Алкезуини, В. И. Горбаченко // Нейроинформатика,

ее приложения и анализ данных : материалы XXV Всерос. семинара (г. Красноярск, 29 сентября – 1 октября 2017 г.). – Красноярск : Ин-т вычислительного моделирования СО РАН, 2017. – С. 3–7.

17. Алкезуини, М. М. Обучение сетей радиальных базисных функций методом Левенберга – Марквардта / М. М. Алкезуини, В. И. Горбаченко // Проблемы информатики в образовании, управлении, экономике и технике : сб. ст. XVII Междунар. науч.-техн. конф. (г. Пенза, 26–27 октября 2017 г.). – Пенза : Приволжский дом знаний, 2017. – С. 74–79.

18. Алкезуини, М. М. Адаптация метода Левенберга – Марквардта для обучения сетей радиальных базисных функций / М. М. Алкезуини, В. И. Горбаченко // Новые информационные технологии и системы : тр. XIV Междунар. науч.-техн. конф. (г. Пенза, 22–24 ноября 2017 г.). – Пенза : Изд-во ПГУ, 2017. – С. 190–194.

19. Алкезуини, М. М. Обучение сетей радиальных базисных функций методом Нестерова при решении краевых задач математической физики / М. М. Алкезуини, В. И. Горбаченко // Аналитические и численные методы моделирования естественнонаучных и социальных проблем : материалы XII Междунар. науч.-техн. конф. (г. Пенза, 4–6 декабря 2017 г.) / под ред. И. В. Бойкова. – Пенза : Изд-во ПГУ, 2017. – С. 171–175.

20. Алкезуини, М. М. Решение краевых задач на сетях радиальных базисных функций / М. М. Алкезуини, В. И. Горбаченко // Нейроинформатика, ее приложения и анализ данных : материалы XXVI Всерос. семинара (г. Красноярск, 28–30 сентября 2018 г.). – Красноярск : Ин-т вычислительного моделирования СО РАН, 2018. – С. 9–15.

21. Алкезуини, М. М. Решение краевых задач на сетях радиальных базисных функций / М. М. Алкезуини, В. И. Горбаченко // Новые информационные технологии и системы : материалы XV Междунар. науч.-техн. конф., посвященной 75-летию ПГУ (г. Пенза, 20–22 ноября 2018 г.). – Пенза : Изд-во ПГУ, 2018. – С. 72–78.

22. Алкезуини, М. М. Обучение методом Левенберга – Марквардта сетей радиальных базисных функций при решении краевых задач / М. М. Алкезуини, В. И. Горбаченко // Аналитические и численные методы моделирования естественнонаучных и социальных проблем : материалы XIII Междунар. науч.-техн. конф. (г. Пенза, 4–6 декабря 2018 г.) / под ред. И. В. Бойкова. – Пенза : Изд-во ПГУ, 2018. – С. 181–187.

23. Алкезуини, М. М. Обучение сетей радиальных базисных функций при решении задач аппроксимации и уравнений в частных производных / М. М. Алкезуини, В. И. Горбаченко // Нейроинформатика 2019 : сб. науч. тр. XXI Междунар. науч.-техн. конф. : в 2 ч. – Москва : МФТИ, 2019. – Ч. 1. – С. 85–93.

24. Алкезуини, М. М. Решение задач аппроксимации на сетях радиальных базисных функций, обучаемых методом Левенберга – Марквардта / М. М. Алкезуини, В. И. Горбаченко // Нейроинформатика, ее приложения и анализ данных : материалы XXVII Всерос. семинара (г. Красноярск, 27–29 сентября 2019 г.). – Красноярск : Ин-т вычислительного моделирования СО РАН, 2018. – С. 3–8.

25. Алкезуини, М. М. Решение прямых и обратных краевых задач на сетях радиальных базисных функций / М. М. Алкезуини, В. И. Горбаченко, Д. А. Стенькин // Новые информационные технологии и системы : сб. науч. ст. XVI Междунар. науч.-техн. конф. (г. Пенза, 27–29 ноября 2019 г.). – Пенза : Изд-во ПГУ, 2019. – С. 242–245.

26. Alqezweeni, M. M. The Solution of Approximation Problems on RBF Networks Learned by The Levenberg Marquardt Method / M. M. Alqezweeni, V. I. Gorbachenko // 2020 International Russian Automation Conference (Sochi, 6–12 September, 2020). – Sochi, 2020. – P. 637–641.

Свидетельства о государственной регистрации программ для ЭВМ

1. Свидетельство об официальной регистрации программы для ЭВМ № 2019665780. Программный комплекс для решения краевых задач на сетях радиальных базисных функций / Алкезуини М. М., Горбаченко В. И. ; дата государственной регистрации в Реестре программ для ЭВМ 28 ноября 2019 г.

Научное издание

АЛКЕЗУИНИ Мухи Муртада

**ОБУЧЕНИЕ СЕТЕЙ РАДИАЛЬНЫХ БАЗИСНЫХ
ФУНКЦИЙ ДЛЯ РЕШЕНИЯ КРАЕВЫХ ЗАДАЧ
МОДЕЛИРОВАНИЯ ОБЪЕКТОВ
С РАСПРЕДЕЛЕННЫМИ ПАРАМЕТРАМИ**

Специальность 05.13.18 – Математическое моделирование, численные
методы и комплексы программ

Редактор *М. М. Кузнецова*
Технический редактор *М. Б. Жучкова*
Компьютерная верстка *М. Б. Жучковой*

Распоряжение № 54/191-2020 от 14.10.2020

Подписано в печать 14.10.2020.
Формат 60×84¹/₁₆. Усл. печ. л. 1,16.
Тираж 100. Заказ № 331.

Издательство ПГУ.
440026, Пенза, Красная, 40.
Тел./факс: (8412) 56-47-33; e-mail: iic@pnzgu.ru