Министерство науки и высшего образования Российской Федерации

Санкт-Петербургский политехнический университет Петра Великого

Физико-механический институт

Высшая школа прикладной математики и вычислительной физики

Работа допущена к защите

Руководитель ОП

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ К.Н. Козлов

«\_\_\_» \_\_\_\_\_\_\_\_\_ 2025 г.

**ВЫПУСКНАЯ КВАЛИФИКАЦИОННАЯ РАБОТА   
РАБОТА БАКАЛАВРА**

**АПОСТЕРИОРНЫЕ ОЦЕНКИ ОШИБОК МЕТОДА DEEP RITZ РЕШЕНИЯ ДВУХТОЧЕЧНЫХ КРАЕВЫХ ЗАДАЧ**

по направлению подготовки 01.03.02 Прикладная математика и информатика

Направленность (профиль) 01.03.02\_02 Системное программирование

Выполнил

студент гр. 5030102/10201 Д.А. Басалаев

Руководитель

Доцент,

Кандидат физико-математических наук А.В. Музалевский

Консультант

По нормоконтролю Л.А. Арефьева

Санкт-Петербург

2025

**РЕФЕРАТ**

На 18 с., 2 рисунка, 10 таблицы, 0 приложений

КЛЮЧЕВЫЕ СЛОВА: ДВУХТОЧЕЧНАЯ КРАЕВАЯ ЗАДАЧА, МЕ-ТОД DEEP RITZ, НЕЙРОННЫЕ СЕТИ, ЧИСЛЕННОЕ РЕШЕНИЕ, АПОСТЕРИОРНАЯ ОЦЕНКА ОШИБКИ, ТОЧНОСТЬ РЕШЕНИЙ.

Тема выпускной квалификационной работы: «Апостериорные оценки ошибок метода Deep Ritz решения двухточечных краевых задач».

В данной работе объектами исследования являются апостериорные оценки ошибок решения обыкновенного дифференциального уравнения 2-ого порядка с краевыми условиями первого рода с помощью машинного обучения. Основная цель работы – выяснить практическую применимость данных апостериорных оценок ошибки. Решаемые задачи в ходе исследования:

* Описание метода Deep Ritz.
* Реализация и математическое объяснения апостериорных оценок ошибки решения дифференциальных уравнений с помощью метода Deep Ritz.
* Проведение серии экспериментов для сравнения реальной ошибки с апостериорной оценкой ошибки решения дифференциального уравнения с помощью метода Deep Ritz с различными параметрами нейронной сети.
* Проведения анализа точности полученных апостериорных оценок ошибки решения дифференциальных уравнений.

По проведённому анализу проведённых экспериментов сделан вывод для каких уравнений данные апостериорные оценки ошибок решения дифференциальных уравнений применимы, а также выяснена связь точности оценки в зависимости от параметров нейронной сети

**ABSTRACT**

On 18 pages, 2 figures, 10 tables, 0 appendices

KEYWORDS: TWO-POINT BOUNDARY VALUE PROBLEM, DEEP RITZ METHOD, NEURAL NETWORKS, NUMERICAL SOLUTION, A POSTERIORI ERROR ESTIMATION, SOLUTION ACCURACY.

Title of the Bachelor’s Thesis: “A Posteriori Error Estimates for the Deep Ritz Method for Solving Two-Point Boundary Value Problems.”

This work investigates a posteriori error estimates for the solution of a second-order ordinary differential equation with first-kind boundary conditions using machine learning. The main goal of the work is to determine the practical applicability of these a posteriori error estimates. The research addresses the following tasks:

* Description of the Deep Ritz method.
* Implementation and mathematical explanation of a posteriori error estimates for solving differential equations using the Deep Ritz method.
* Conducting a series of experiments to compare the actual error with the a posteriori error estimate for the solution of the differential equation using the Deep Ritz method with various neural network parameters.
* Analyzing the accuracy of the obtained a posteriori error estimates for solving differential equations.

Based on the analysis of the conducted experiments, a conclusion is drawn regarding which equations these a posteriori error estimates for solving differential equations are applicable, and the relationship between the accuracy of the estimate and the parameters of the neural network is clarified.

# Содержание

[Содержание 4](#_Toc197806118)

[Введение 5](#_Toc197806119)

[Глава 1. Постановка задачи 8](#_Toc197806120)

[1.1. Классическая постановка 8](#_Toc197806121)

[Глава 2. Методы решения на основе нейронных сетей 9](#_Toc197806122)

[2.1. Модель нейронной сети Deep Ritz 9](#_Toc197806123)

[Глава 3. Контроль точности нейросетевых решений 10](#_Toc197806124)

[3.1. Проблема контроля ошибок 10](#_Toc197806125)

[3.2. Гарантированные двусторонние оценки нормы ошибки 10](#_Toc197806126)

[3.2. Реализация вычисления апостериорной оценки ошибки 10](#_Toc197806127)

[Глава 4. Численные эксперименты 11](#_Toc197806128)

[4.1. Тест №1: Модельная задача с синусоидальным решением 11](#_Toc197806129)

[4.1.1. Результаты для аналитического решения 11](#_Toc197806130)

[4.1.2. Результаты для приближённого решения 13](#_Toc197806131)

[Заключение 16](#_Toc197806132)

[Список использованных источников 17](#_Toc197806133)

# Введение

Дифференциальные уравнения (ДУ) занимают важнейшее место в математическом моделировании широкого спектра физических явлений и инженерных задач. Они служат мощным инструментом для описания динамики систем, изменения во времени и пространстве, охватывая такие области, как физика, химия, биология, экономика и многие другие. Получение точных решений позволяет принимать обоснованные решения, оптимизировать процессы, проектировать новые устройства и материалы, а также углублять понимание фундаментальных законов природы. Сложность многих реальных задач приводит к тому, что аналитические решения ДУ, как правило, недоступны, что стимулирует разработку и совершенствование приближенных методов решения.

Традиционно для приближенного решения дифференциальных уравнений использовались методы конечных разностей, конечных элементов, Эйлера, Рунге-Кутты, Галеркина, Ритца и некоторые другие. В зависимости от свойств уравнения, требований к скорости и точности решения происходит выбор метода с учётом всех его преимуществ и ограничений. Однако, в последние годы наблюдается растущий интерес к применению машинного обучения для решения ДУ. Особенную эффективность и перспективность показали методы, основанные на обучении нейронных сетей (Deep Neural Networks (DNN)). Данные методы, предоставляет альтернативный способ построения приближенных решений, потенциально обладающий преимуществами в плане гибкости, адаптивности и возможности обработки сложных задач. Особенно перспективным направлением является использование вариационного принципа, например, методом Deep Ritz, позволяющим использовать архитектуру нейронной сети для минимизации функционала энергии, что обеспечивает эффективное решение краевых задач.

Вопрос о принципиальной возможности аппроксимации решений дифференциальных уравнений нейронными сетями активно исследуется с опорой на Универсальную теорему аппроксимации [3]. Однако, доказательство существования нейросети, способной с заданной точностью аппроксимировать решение, не является достаточным для практического применения метода. Для обеспечения надежности и контроля качества получаемых приближенных решений необходимы инструменты апостериорной оценки ошибок. Такие оценки позволяют, используя информацию о найденном приближении, получить верхнюю границу погрешности решения в некоторой норме. Большинство известных апостериорных методов основаны на использовании специальных свойств решений и аппроксимаций, таких как галеркинская ортогональность, достаточная регулярность точного решения или явление суперсходимости. Нейросетевые решения зачастую не удовлетворяют этим условиям, что делает невозможным прямое применение стандартных апостериорных оценок. Однако, в работе [4] предложен альтернативный подход к построению апостериорных оценок для широкого класса задач, не требующий выполнения вышеуказанных условий и адаптированный для нейросетевых решений.

Методы машинного обучения для решения дифференциальных уравнений, в частности, применение метода Deep Ritz, представляют собой активно развивающуюся область исследований. Несмотря на теоретические обоснования и подтвержденную возможность аппроксимации решений, практическое применение данных методов требует более глубокого изучения. В частности, вопрос об эффективности и точности апостериорных оценок, предоставляющих верхнюю границу погрешности решения, остается открытым. **Целью работы** является исследование применимости апостериорных оценок ошибок метода Deep Ritz для решения двухточечных краевых задач. Особое внимание уделяется изучению зависимости точности оценок от архитектуры и параметров нейронной сети, а также от свойств решаемого дифференциального уравнения. Для достижения поставленной цели были сформулированы следующие **задачи**:

* Изучение теоретических основ метода Deep Ritz и принципов построения апостериорных оценок ошибок для вариационных задач.
* Разработка программной реализации метода Deep Ritz для решения двухточечных краевых задач и алгоритма вычисления апостериорных оценок ошибок.
* Проведение серии численных экспериментов с различными дифференциальными уравнениями второго порядка для оценки эффективности и точности апостериорных оценок в зависимости от параметров нейронной сети и свойств решаемой задачи.
* Анализ полученных результатов и формулирование выводов о применимости и ограничениях использования апостериорных оценок ошибок метода Deep Ritz.

# Глава 1. Постановка задачи

1.1. Классическая постановка

Рассмотрим обыкновенное дифференциальное уравнение второго порядка, дополненное краевыми условиями первого рода:

где 𝑥 ∈ Ω = (𝑎, 𝑏) – открытый интервал на вещественной прямой ℝ; 𝑝(𝑥), 𝑞(𝑥) – заданные функции, 𝑝, 𝑞 : (𝑎, 𝑏) → ℝ; 𝑓(𝑥) – заданная функция, 𝑓 ∈ 𝐿2(Ω); 𝑢(𝑥) : ℝ → ℝ – искомая вещественнозначная функция вещественного аргумента. Классическая постановка задачи предполагает (1.1) и (1.2), что решение 𝑢(𝑥) является дважды непрерывно дифференцируемой функцией на замкнутом интервале, то есть 𝑢 ∈ 𝐶2().

# Глава 2. Методы решения на основе нейронных сетей

В последние годы наблюдается растущий интерес к применению искусственных нейронных сетей для решения дифференциальных уравнений. Несмотря на то, что классические вычислительные методы, такие как метод конечных элементов (МКЭ), во многих случаях остаются наиболее эффективными, существуют классы задач, для которых применение традиционных подходов сопряжено со значительными трудностями. К таким задачам относятся, например, экономические модели большой размерности, стохастические дифференциальные уравнения, а также задачи с неполностью определенными данными, характерные для медицины и биологии.

Кроме того, исследования в области применения нейронных сетей для решения дифференциальных уравнений представляют значительный интерес с точки зрения развития теории и методов машинного обучения. Дифференциальные уравнения являются хорошо изученным классом математических моделей, что позволяет использовать их для тестирования и сравнения эффективности различных методов оптимизации (обучения) нейронных сетей, поскольку успешность обучения может быть оценена количественно на основе точности полученного решения. Это обуславливает актуальность рассмотрения нейросетевых методов как альтернативного подхода к решению краевых задач, в том числе и в контексте данной работы.

2.1. Модель нейронной сети Deep Ritz

# Глава 3. Контроль точности нейросетевых решений

3.1. Проблема контроля ошибок

3.2. Гарантированные двусторонние оценки нормы ошибки

3.2. Реализация вычисления апостериорной оценки ошибки

# Глава 4. Численные эксперименты

4.1. Тест №1: Модельная задача с синусоидальным решением

В качестве первого тестового примера рассмотрим краевую задачу для уравнения Пуассона с однородными граничными условиями Дирихле:

Данная задача имеет аналитическое решение 𝑢(𝑥) = sin(𝜋𝑥) и выбрана в качестве модельной в силу простоты и наличия гладкого аналитического решения, что позволяет более наглядно оценить эффективность предложенных апостериорных оценок ошибок.

4.1.1. Результаты для аналитического решения

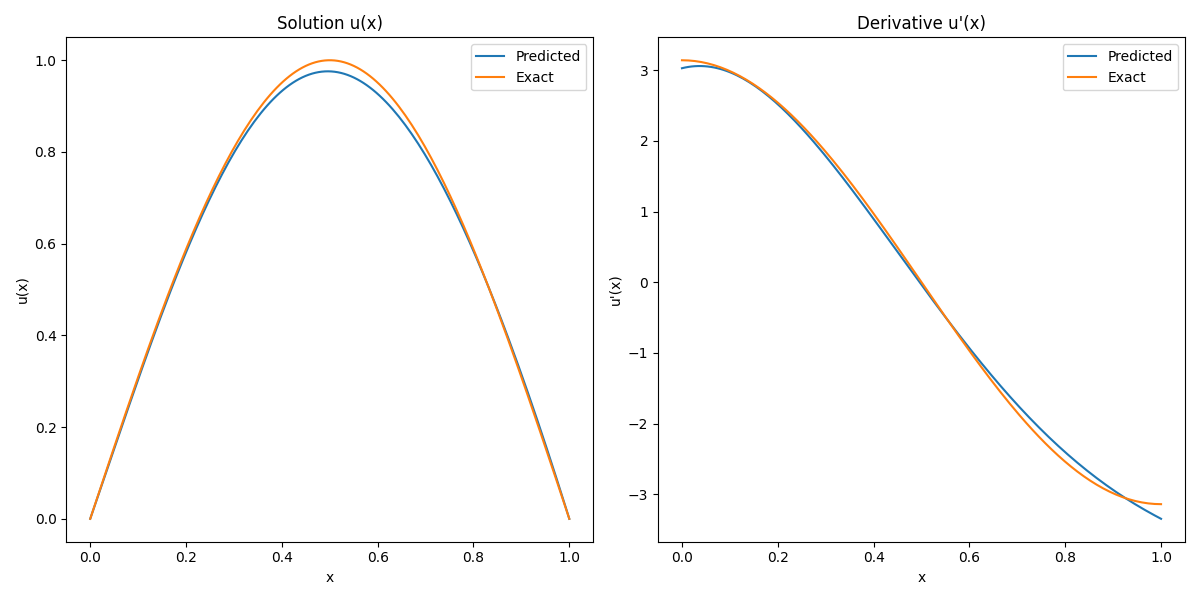


Рисунок 1: Графики аналитического и нейросетевого решений для тестовой задачи (4.1)-(4.2).

Далее представлены таблицы, содержащие значения индекса эффективности для различных конфигураций нейронной сети, используемой в методе Deep Ritz.

Таблица 1: Зависимость индекса эффективности от размера обучающей выборки.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Batch size |  |  |  | Training Time |
| 64 | 0.0576832 | 0.0288416 | 1.9999993 | 1.146 |
| 128 | 0.0126187 | 0.0063093 | 1.9999985 | 1.206 |
| 256 | 0.0157407 | 0.0078704 | 2.0000008 | 1.299 |
| 512 | 0.0161689 | 0.0080844 | 2.0000028 | 1.467 |
| 1024 | 0.0029355 | 0.0014678 | 1.9999942 | 1.854 |

Таблица 2: Зависимость индекса эффективности от количества итераций.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Iterations |  |  |  | Training Time |
| 1000 | 0.0259679 | 0.0129839 | 2.0000009 | 0.799 |
| 2000 | 0.0097438 | 0.0048719 | 1.9999995 | 1.199 |
| 3000 | 0.0359227 | 0.0179614 | 1.9999986 | 1.568 |
| 4000 | 0.0126987 | 0.0063494 | 1.9999984 | 1.998 |

Таблица 3: Зависимость индекса эффективности от количества слоёв нейронной сети.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Layers |  |  |  | Training Time |
| 2 | 0.0547187 | 0.0273593 | 2.0000016 | 0.992 |
| 3 | 0.0345096 | 0.0172548 | 2.0000020 | 1.195 |
| 4 | 0.0582991 | 0.0291490 | 1.9999986 | 1.411 |
| 5 | 0.0292883 | 0.0146442 | 1.9999983 | 1.705 |

Таблица 4: Зависимость индекса эффективности от количества слоёв нейронной сети.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Layers |  |  |  | Training Time |
| 2 | 0.0462305 | 0.0231152 | 1.9999999 | 1.334 |
| 3 | 0.0305819 | 0.0152910 | 2.0000018 | 1.150 |
| 4 | 0.0150327 | 0.0075163 | 1.9999996 | 1.199 |
| 5 | 0.0314196 | 0.0157098 | 1.9999991 | 1.343 |

Таблица 5: Зависимость индекса эффективности от оптимизатора нейронной сети.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Optimizer |  |  |  | Training Time |
| adam | 0.0165207 | 0.0082604 | 2.0000005 | 1.371 |
| sgd | 0.1634730 | 0.0817386 | 2.0000006 | 1.787 |
| msprop | 0.0154928 | 0.007464 | 1.9999985 | 1.049 |

4.1.2. Результаты для приближённого решения

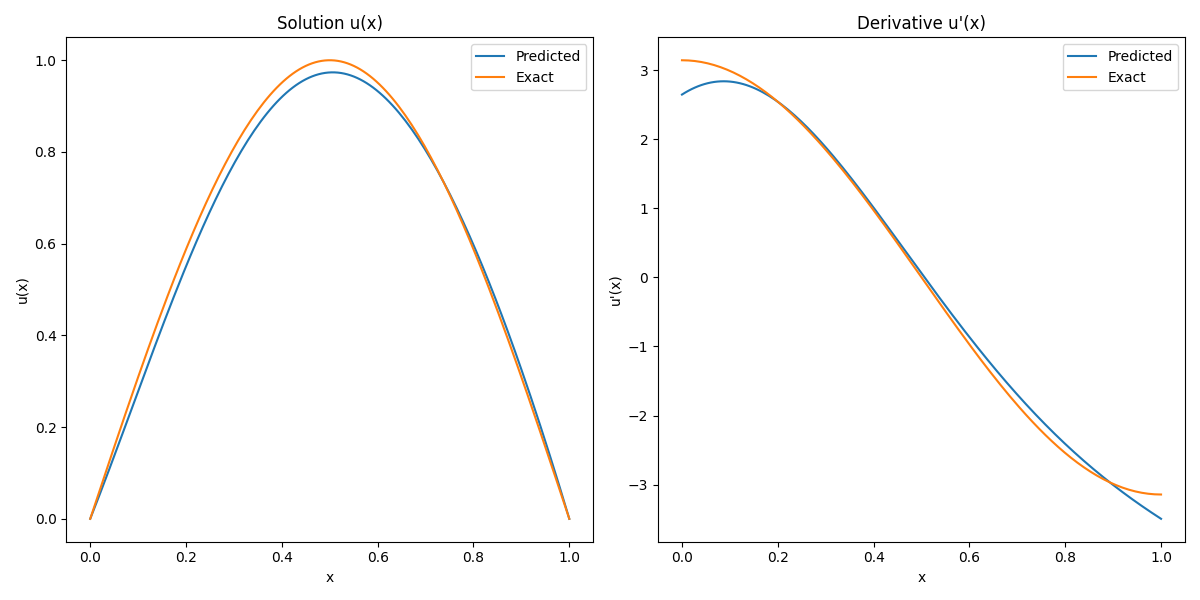


Рисунок 2: Графики приближённого и нейросетевого решений для тесто-вой задачи (4.1)-(4.2).

Таблица 6: Зависимость индекса эффективности от размера обучающей выборки.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Batch size |  |  |  | Training Time |
| 64 | 0.0768758 | 0.0384379 | 2.0000180 | 1.192 |
| 128 | 0.0768758 | 0.0101890 | 2.0000063 | 1.228 |
| 256 | 0.0189322 | 0.0094661 | 2.0000730 | 1.287 |
| 512 | 0.0056256 | 0.0028128 | 2.0000246 | 1.521 |
| 1024 | 0.0089707 | 0.0448530 | 2.0000154 | 1.827 |

Таблица 7: Зависимость индекса эффективности от количества итераций.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Iterations |  |  |  | Training Time |
| 1000 | 0.0259679 | 0.0129839 | 2.0000009 | 0.799 |
| 2000 | 0.0097438 | 0.0048719 | 1.9999995 | 1.199 |
| 3000 | 0.0359227 | 0.0179614 | 1.9999986 | 1.568 |
| 4000 | 0.0126987 | 0.0063494 | 1.9999984 | 1.998 |

Таблица 8: Зависимость индекса эффективности от количества слоёв нейронной сети.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Layers |  |  |  | Training Time |
| 2 | 0.0547187 | 0.0273593 | 2.0000016 | 0.992 |
| 3 | 0.0345096 | 0.0172548 | 2.0000020 | 1.195 |
| 4 | 0.0582991 | 0.0291490 | 1.9999986 | 1.411 |
| 5 | 0.0292883 | 0.0146442 | 1.9999983 | 1.705 |

Таблица 9: Зависимость индекса эффективности от количества слоёв нейронной сети.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Layers |  |  |  | Training Time |
| 2 | 0.0462305 | 0.0231152 | 1.9999999 | 1.334 |
| 3 | 0.0305819 | 0.0152910 | 2.0000018 | 1.150 |
| 4 | 0.0150327 | 0.0075163 | 1.9999996 | 1.199 |
| 5 | 0.0314196 | 0.0157098 | 1.9999991 | 1.343 |

Таблица 10: Зависимость индекса эффективности от оптимизатора нейронной сети.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Optimizer |  |  |  | Training Time |
| adam | 0.0165207 | 0.0082604 | 2.0000005 | 1.371 |
| sgd | 0.1634730 | 0.0817386 | 2.0000006 | 1.787 |
| msprop | 0.0154928 | 0.007464 | 1.9999985 | 1.049 |

# Заключение

В рамках данной дипломной работы были исследованы апостериорные оценки ошибок метода Deep Ritz для решения двухточечных краевых задач. Для различных дифференциальных уравнений второго порядка с различными свойствами были получены апостериорные оценки ошибок при варьировании параметров нейронной сети, используемой в методе Deep Ritz. Проведен анализ зависимости индекса эффективности от свойств дифференциального уравнения и параметров нейронной сети.

В ходе работы была разработана программная реализация метода Deep Ritz, которая использовалась для получения приближенных решений рассматриваемых краевых задач и вычисления апостериорных оценок ошибок.

Полученные результаты свидетельствуют о высокой устойчивости апостериорных оценок к изменениям свойств дифференциального уравнения и параметров нейронной сети. Значения индекса эффективности, близкие к 2, указывают на то, что апостериорные оценки дают достаточно точную оценку реальной ошибки, что подтверждает практическую применимость данного подхода для контроля качества решений, полученных методом Deep Ritz.

В качестве перспективных направлений дальнейших исследований можно выделить применение разработанных методов для решения задач большей размерности, а также исследование влияния различных архитектур нейронных сетей и методов оптимизации на точность и эффективность апостериорных оценок.

# Список использованных источников

1. Репин, С. И. (2008). A posteriori estimates for partial differential equations. Walter de Gruyter.
2. Музалевский, А. В., & Репин, С. И. (2024). О решении дифференциальных уравнений с помощью нейронных сетей: Целевые функционалы и верификация результатов. DOI: 10.20948/mm-2024-06-12
3. Чурилова, М. А. (2014). Вычислительные свойства функциональных апостериорных оценок для стационарной задачи реакции-диффузии. Вестник СПбГУ. Сер. 1. Математика. Механика. Астрономия, 1(59)(1), 161-173.
4. E W., Yu B. The Deep Ritz method: A deep learning-based numerical algorithm for solving variational problems. — 2017. — DOI 10.48550/ARXIV.1710.00211. — URL: https://arxiv.org/abs/1710.00211.
5. TensorFlow tutorials. — URL: https://www.tensorflow.org/tutorials (visited on 10.05.2025).