РОССИЙСКИЙ УНИВЕРСИТЕТ ДРУЖБЫ НАРОДОВ
ИМЕНИ ПАТРИСА ЛУМУМБЫ

Факультет физико-математических и естественных наук
Кафедра математического моделирования и искусственного интеллекта

Утверждаю
Заведующий кафедрой
математического моделирования
и искусственного интеллекта

__________________ **Малых М.Д.**

«_____»_________ **2024 г.**

НАУЧНО-ИССЛЕДОВАТЕЛЬСКАЯ РАБОТА

на тему

Обучение нейронных сетей для аппроксимации решений краевых задач с применением неклассических вариационных формулировок

Выполнил
Студент группы НКНбд0120
Студенческий билет №: 1032204249

_____________________________ (ФИО подпись)

«__________» 20_____ г.

Руководитель

_____________________________ (ФИО подпись)

Москва 2024

Содержание

1 Введение

В современном мире численные методы играют ключевую роль в решении широкого спектра задач математической физики. Особенно важными являются методы, основанные на вариационных принципах, которые позволяют эффективно моделировать разнообразные физические явления. Данная работа рассматривается проблема аппроксимации решений краевых задач для уравнений математической физики с применением неклассических вариационных формулировок.

Одним из ключевых направлений в данной области является использование нейронных сетей для приближенного решения дифференциальных уравнений. В частности, это связано с возможностью адаптации нейронных сетей к сложным структурам данных и уравнениям, которые трудно или невозможно решить с помощью классических аналитических методов.

Основные методы и исследования

Для проведения исследования в данной области необходимо рассмотреть как теоретические основы построения неклассических вариационных формулировок краевых задач, так и современные методы обучения нейронных сетей. Основными источниками, на которых основана эта работа, являются работы Филиппова, Савчина и Шорохова "Вариационные принципы для непотенциальных операторов а также работы Ядава, и Кумара "Введение в нейросетевые методы для дифференциальных уравнений"и Сириньяно и его коллег, которые представили глубокое обучение алгоритм для нейронных сетей для решения дифференциальных уравнений в частных производных (DGM).

Ядав, и Кумар подробно описали в своей книге методы нейросетей для решения дифференциальных уравнений. Они сосредоточились на важности более глубокого понимания методов нейронных сетей, которые были представлены в наиболее эвристической и интуитивно понятной манере.

Работа Филиппова, Савчина и Шорохова была связана с вариационным принципом для класса непотенциальных операторов. С. Савчин провел исследование, которое оказало значительное влияние на вариационные принципы, не потенциальные операторы, симметрии и обратные задачи вариационного исчисления.

Основываясь на этой идее, Сириньяно и Спилиопулос разработали метод глубокого обучения Galerkin (DGM), платформу глубокого обучения для решения PDE. Они аппроксимируют решение с помощью глубокой нейронной сети, обучая ее соблюдению дифференциального оператора, начальных и граничных условий.

Цели и перспективы

Целью данной работы является анализ и сравнение различных методов построения неклассических вариационных постановок краевых задач и методов обучения нейронных сетей для их аппроксимации, а также изучение их применимости в различных физических задачах. Ожидается, что в результате исследования будут получены новые знания о возможностях и ограничениях этих методов, которые могут оказаться важными для развития численных методов в математической физике и смежных областях науки и техники.

Прочие области применения нейросетей и не вариационных принципов для решения дифференциальных уравнений используются с большим обещанием и значительными вкладами. На данный момент такие методы, вероятно, предлагают большие преимущества по сравнению с традиционными численными методами, особенно в отношении задач высокой размерности. Однако дальнейшие исследования должны быть проведены для реализации этого потенциала и разработки решений современных проблем.

Стуктуры работы.

Работа состоит из трех глав.

. Во введении отражена обзор научных статей и монографий по теме научно-исследовательской работы.

. В первой главе представилы методы построения неклассических вариационных формулировок краевых задач для уравнений математической физики

. В второй главе посвящена выбор архитектуры нейронной сети для аппроксимации решений краевой задачи.

. В третьей главе приведены основные подходы к обучению нейронной сети для аппроксимации решений краевой задачи

2 Глава 1

Задачи краевого значения — это тип математической задачи, включающий дифференциальное уравнение и набор дополнительных ограничений, известных как граничные условия, которым должно удовлетворять решение в определенных точках или поверхностях в области интереса. Эти задачи часто встречаются в физике и инженерии, особенно в контекстах теплопроводности, динамики жидкостей и электромагнетизма.

2.1 Основные элементы

2.1.1 Дифференциальное уравнение

Центральным элементом задачи краевого значения является дифференциальное уравнение, которое может быть обычным (ОДУ) или частичным (ЧДУ). Это уравнение связывает некоторую функцию (часто представляющую физические величины, такие как температура, перемещение или электрический потенциал) с её производными.

2.1.2 Граничные условия

Это условия, которым должно удовлетворять решение на границах области, где определено дифференциальное уравнение. Область может быть пространственной (как края металлической пластины) или временной (начало и конец времени), в зависимости от задачи.

2.1.3 Область

Это относится к диапазону, в котором рассматривается дифференциальное уравнение, такому как длина стержня, поверхность сферы или объем цилиндра.

2.2 Типы граничных условий

- Условия Дирихле: Значение функции указывается на границе.
- Условия Неймана: Указывается производная функции на границе.
- Смешанные или условия Робена: Указывается комбинация значений функций и производных.

Задачи краевого значения отличаются от задач начального значения, которые указывают условия только в одной точке (обычно в начале временного домена). Решение задач краевого значения часто включает нахождение функции, которая не только удовлетворяет дифференциальному уравнению, но и соответствует граничным условиям, что делает эти задачи по своей сути более сложными и иногда требующими численных методов для решения.

2.3 Неклассические вариационные формулировки для краевых задач в математической физике

Определение математического представления неклассических вариационных формулировок для краевых задач в математической физике включает адаптацию классических вариационных принципов для учета более сложных и реальных сценариев, где классические предположения не выполняются. Ниже приведены подробные математические представления нескольких неклассических вариационных формулировок:

1. Слабые формулировки

Слабая формулировка краевой задачи заключается в выражении проблемы таким образом, чтобы её можно было применить к функциям в более широком функциональном пространстве, обычно в пространстве Соболева.

Математическое представление: Для дифференциального уравнения, обычно представленного как $Lu = f$ на области Ω с границей $\partial\Omega$ и граничным условием $u = g$ на $\partial\Omega$, слабая форма выводится путём умножения на тестовую функцию v (с $v = 0$ на $\partial\Omega$) и интегрирования по Ω:

$$\int_\Omega v \cdot Lu \, dx = \int_\Omega v \cdot f \, dx$$

После применения интегрирования по частям и предположения подходящих условий регулярности, мы получаем:

$$\int_\Omega a(u, v) \, dx = \int_\Omega v \cdot f \, dx,$$

где $a(u, v)$ — билинейная форма, охватывающая дифференциальные операции над u.

2. Обобщённые функциональные пространства

Здесь решения ищутся в пространствах, которые могут включать распределения или другие обобщённые функциональные пространства, особенно при работе с разрывами или сингулярностями.

Математическое представление: Формулировка в обобщённых функциональных пространствах часто не изменяет форму уравнений, но изменяет пространства, из которых берутся решения и тестовые функции, например, из $H^1(\Omega)$ в $H^s(\Omega)$ для некоторого s, или даже в распределения, если ожидается низкая регулярность решений.

3. Смешанные и гибридные методы

Смешанные методы включают формулировку дополнительных переменных для переформулировки краевой задачи в систему уравнений, которая может быть более естественно или легко обработана.

Математическое представление:

$$\begin{aligned} \sigma - \nabla u &= 0, \\ \nabla \cdot \sigma &= f. \end{aligned}$$

Это может привести к системе уравнений, каждая часть которой может быть обработана более непосредственно, часто приводя к лучшей числовой стабильности или упрощению реализации.

4. Методы штрафов

Методы штрафов добавляют к функционалу условие, которое сильно штрафует отклонения от граничных условий или внутренних ограничений.

Математическое представление: Для краевой задачи, где $u = g$ на $\partial\Omega$, метод штрафов может изменить функционал $J(u)$ до:

$$J(u) = \int_\Omega F(u, \nabla u)\, dx + \frac{1}{\epsilon} \int_{\partial\Omega} (u - g)^2\, ds,$$

где ϵ — малый параметр штрафа.

5. Множители Лагранжа

Этот метод используется для принуждения соблюдения ограничений, которые не обрабатываются естественно функциональным пространством или граничными условиями, путем введения множителей Лагранжа.

Математическое представление: Для ограничения $\int_\Omega u\, dx = C$, функционал с множителем Лагранжа λ является:

$$J(u, \lambda) = \int_\Omega F(u, \nabla u)\, dx + \lambda \left(\int_\Omega u\, dx - C \right).$$

Вариационные неравенства используются для решения проблем, где решение ограничено неравенствами.

Математическое представление: Для краевой задачи с ограничением типа $u \geq g$ на Ω, вариационное неравенство может быть:

$$\int_\Omega a(u, v - u)\, dx \geq \int_\Omega f \cdot (v - u)\, dx \quad \forall v \geq g.$$

Эти формулировки позволяют обрабатывать более сложные и разнообразные задачи, чем классические методы, что особенно полезно в физике, инженерии и прикладной математике. Каждый подход корректирует классический вариационный принцип для учета конкретных ограничений, областей или характеристик решения.

3 Глава 2

3.1 Определение и влияние архитектуры нейронной сети

Термин "архитектура нейронной сети"относится к структуре и конфигурации искусственной нейронной сети, включая расположение слоев, количество узлов в каждом слое и связи между этими узлами. Эта архитектура создана для имитации обработки информации человеческим мозгом, хотя и в упрощенной и структурированной форме.

Архитектура нейронной сети существенно влияет на ее способность точно моделировать и решать различные типы задач, от простых задач, таких как линейная регрессия, до сложных, таких как распознавание изображений и перевод текста. Решения о количестве слоев, типах слоев (например, сверточные, рекуррентные), функциях активации и другие играют ключевую роль в производительности сети и ее пригодности для конкретных приложений.

3.1.1 Основные компоненты архитектуры нейронной сети

- **Слои:**
  - **Входной слой:** первый слой, который получает входные данные. Он выступает в роли первичного приемника данных, передающего данные последующим слоям.
  - **Скрытые слои:** один или несколько слоев, где выполняется вычисление через взвешенные соединения. Эти слои трансформируют входные данные в формат, пригодный для использования выходным слоем.
  - **Выходной слой:** последний слой, производящий выходные данные сети, соответствующие выполняемой задаче (например, классификация, регрессия).

- **Узлы (или нейроны):** каждый узел в слое представляет собой нейрон человеческого мозга и содержит функцию активации, которая определяет, должен ли он активироваться и насколько сильно, передавая данные следующему слою.
- **Веса и смещения:** параметры сети, которые корректируются в процессе обучения. Веса контролируют силу влияния между двумя нейронами. Смещения добавляются к входным данным для корректировки выходных данных вместе с взвешенной суммой от входных данных.
- **Функции активации:** функции, применяемые к каждому узлу для внесения нелинейности в выход нейрона. Это важно, поскольку помогает сети учиться сложным паттернам.
- **Связность:** относится к тому, как нейроны в разных слоях соединены между собой. Это может быть плотное соединение (каждый нейрон соединен с каждым нейроном следующего слоя), сверточное соединение (соединения на основе фильтров), рекуррентное соединение (соединения между узлами через временные шаги) и т.д.

3.2 Обзор различных типов архитектур нейронных сетей

- **Прямые нейронные сети (FNN):** Самый простой тип искусственных нейронных сетей. В этой архитектуре информация передается только в одном направлении — вперед — от входных узлов, через скрытые узлы (если они есть), к выходным узлам. В сети нет циклов или петель.
- **Сверточные нейронные сети (CNN):** В основном используются в области компьютерного зрения, но также все чаще применяются в других областях. CNN характеризуются своими сверточными слоями, которые могут захватывать иерархические паттерны в пространственных данных, таких как изображения.
- **Рекуррентные нейронные сети (RNN):** Разработаны для распознавания последовательностей, таких как

текст или временные ряды. В отличие от FNN, RNN имеют связи, образующие петли назад, что позволяет им сохранять некоторый вид памяти о предыдущих входных данных. Это делает их идеальными для задач, где важен контекст из предыдущей части последовательности.

- **Долговременная кратковременная память (LSTM):** Специальный вид RNN, способный учиться зависимостям длительного действия. LSTM особенно полезны, когда требуется захватить информацию от входных данных, между которыми проходят длительные временные промежутки.

- **Автоэнкодеры:** Используются для обучения без учителя эффективных кодировок. Сеть обучается сжимать входные данные в код меньшей размерности, а затем восстанавливать выходные данные из этого представления. Автоэнкодеры широко используются для обнаружения аномалий, шумоподавления и сокращения размерности.

- **Генеративно-состязательные сети (GANs):** Состоят из двух нейронных сетей, генератора и дискриминатора, которые обучаются одновременно. Задача генератора — производить данные, максимально приближенные к реальным, а задача дискриминатора — определить, сгенерированы данные или реальны.

- **Трансформеры:** Основаны на механизме самовнимания, трансформеры оказались чрезвычайно эффективными в обработке последовательных данных, превосходя RNN и LSTM в задачах, таких как обработка естественного языка (NLP).

- **Физически-информированные нейронные сети (PINNs):** Используются для приближенного решения краевых задач (BVPs) в контексте распространения сейсмических волн. Сейсмические волны — это сложные явления, включающие распространение энергетических волн через Землю, которые могут быть смоделированы с помощью дифференциальных уравнений (PDEs). Точное решение этих уравнений критично для приложений в геофизике, таких как моделирование землетрясений, разведка нефти и понимание внутренней структуры Земли. PINNs включают управляющие физические законы (выраженные в виде дифференциальных уравнений) непосредственно в функцию потерь сети.

4 Глава 3

Обучение нейронной сети для приближенного решения краевых задач (BVP) в математике и физике включает несколько конкретных подходов, которые адаптируют процесс обучения для эффективного управления ограничениями и сложностями уравнений, управляющих проблемой.

4.1 Основные стратегии обучения

1. **Обучение с учителем**
   - Генерация данных: Создание набора данных из входных данных (например, координаты, временные точки) и соответствующих выходных данных (например, решения в этих точках) с использованием традиционных численных методов или аналитических решений, где это возможно.
   - Процесс обучения: Обучение нейронной сети отображению входных параметров в решения с использованием стандартных функций потерь, таких как среднеквадратичная ошибка между прогнозами сети и истинными решениями.
2. **Физически информированные нейронные сети (PINNs)**
   - Формулировка функции потерь: Вместо полагания исключительно на маркированные данные, PINNs включают управляющие дифференциальные уравнения в функцию потерь. Это достигается за счет оценки дифференциального уравнения в точках внутри домена и штрафов за отклонения от нуля.
   - Граничные и начальные условия: Граничные условия также включены в функцию потерь, обеспечивая их

соблюдение за счет штрафов за отклонения на границах.

3. **Обучение без учителя**
   - Энергетические модели: Использование подхода на основе энергии, где сеть обучается минимизировать физический энергетический функционал, соответствующий изучаемой BVP.
   - Архитектура сети: Разработка сети, которая выводит параметры, определяющие состояние системы, и затем рассчитывает энергию на основе этих параметров.
4. **Метод Дип Галеркина (DGM)**
   - Метод: Обучение глубокой нейронной сети для удовлетворения дифференциального уравнения по всему домену и соблюдение граничных условий без необходимости данных обучения решений.
   - Функция потерь: Состав функции потерь включает термины, обеспечивающие выполнение дифференциального уравнения, и термины для выполнения граничных и начальных условий.
5. **Вариационные методы**
   - Принцип: Обучение нейронных сетей для приближения минимизатора некоторого функционала в задачах, где существует вариационная формулировка.
   - Реализация: Обучение сети для вывода значений функции, минимизирующих интеграл плотности энергии.
6. **Трансферное обучение**
   - Применение: Использование предварительно обученных моделей на аналогичных типах дифференциальных уравнений или доменах и их дополнительная настройка для конкретной BVP.
   - Преимущество: Сокращение времени обучения и улучшение производительности модели.
7. **Состязательное обучение (для GANs)**

- Настройка: Использование GAN, где одна сеть генерирует решения, а другая различает между сгенерированными и реальными решениями.
- Цель: Улучшение способности генеративной модели производить реалистичные решения.

4.2 Рассмотрения стратегии обучения

- Выбор точек: Влияние выбора мест для оценки дифференциального уравнения и граничных условий на эффективность обучения и точность решения.
- Балансировка терминов: Важность правильной балансировки терминов в функции потерь для обеспечения адекватного учета всех аспектов задачи.
  Каждый из этих подходов может быть адаптирован к конкретным типам (BVP) и характеристикам области, в которой ставится задача, таким как ее размерность, сложность и характер граничных условий.

5 Заключение и перспективы

В заключение, изучение научных статей и монографий, включая работы Филиппова, Савчина, Шорохова, Шалова, а также Ядава и Кумара, позволяет нам увидеть, как теоретические принципы и практические методы взаимодействуют при решении задач с граничными условиями, особенно при анализе волновых уравнений. Фундаментальные идеи, заложенные Филипповым и другими, а также Шаловым о вариационных принципах и минимизации квадратичных функционалов, создают прочную математическую базу. Эту базу можно эффективно сочетать с новейшими вычислительными методами. Такое сочетание обещает принести революцию в нашем подходе к решению волновых уравнений, в том числе с помощью технологий нейронных сетей, о которых рассказали Ядав и его коллеги.

Двигаясь вперед на стыке математики и машинного обучения, наши исследования открывают двери в будущее, где мы сможем более точно предсказывать и управлять сложностями природной и искусственной среды. По мере развития этих методов они могут радикально изменить сферы, зависящие от понимания динамических систем, такие как метеорология, сейсмология и инженерное дело.

Список литературы

[1] В. М. Филиппов, В. М. Савчин, С. Г. Шорохов, *Вариационные принципы для непотенциальных операторов*, Итоги науки и техн. Сер. Соврем. пробл. мат. Нов. достиж., 1992, том 40, стр. 3–176.

[2] В. М. Шалов, *Принцип минимума квадратичного функционала для гиперболического уравнения*, Дифференц. уравнения, 1965, том 1, номер 10, стр. 1338–1365.

[3] N. Yadav, A. Yadav, M. Kumar, *An Introduction to Neural Network Methods for Differential Equations*, Springer, 2015.