КАЗАХСКИЙ НАЦИОНАЛЬНЫЙ УНИВЕРСИТЕТ ИМЕНИ АЛЬ-ФАРАБИ · КАФЕДРА КОМПЬЮТЕРНЫХ НАУК

Научное машинное обучение

Лекция 6. PINN с жёсткими ограничениями и вариационные PINN: Deep Ritz и Deep Galerkin

Точное выполнение граничных условий и решение уравнений через энергию и слабую форму

Кенжебек Ержан Ғалымжанұлы

PhD, старший преподаватель кафедры компьютерных наук

Цель и план лекции

Цель

Изучить способы точного учёта граничных условий и вариационные подходы к решению ДУЧП нейросетями: Deep Ritz, Deep Galerkin и вариационные PINN.

План

1. Мягкие и жёсткие ограничения
2. Конструирование ансатца с жёсткими условиями
3. Сильная и слабая формы уравнения
4. Метод Deep Ritz
5. Метод Deep Galerkin и вариационные PINN
6. Сравнение подходов и выбор метода

Мягкие и жёсткие ограничения

Мягкие

Условия входят штрафом в функцию потерь: $w_b \cdot |u - g|^2$. Просто, но выполняются приближённо и требуют подбора веса.

Жёсткие

Архитектура гарантирует условия точно, поэтому член граничных потерь не нужен.

Жёсткие ограничения уменьшают число конкурирующих членов потерь — это снимает часть проблем из лекции 5.

Конструирование анзатца

$$u_\theta(x) = g(x) + \varphi(x) \cdot N_\theta(x)$$

g — любая функция, удовлетворяющая условию на границе; φ равна нулю на границе и положительна внутри; N_θ — нейросеть.

Отрезок [0, 1]

u = x(1 − x)·N_θ(x) даёт u(0) = u(1) = 0.

Периодичность

Вход заменяется парой sin(2πx/L), cos(2πx/L).

Жёсткие условия: плюсы и минусы

Плюсы

Условия выполняются точно; меньше членов потерь; обычно лучше обусловленность.

Минусы

Для сложной геометрии нужны функции расстояния до границы или R-функции; условия Неймана и Робина строятся сложнее; ансатц сужает класс решений.

Подходы к сложной геометрии: Sukumar и Srivastava (2022), Lu и соавторы (2021).

Сильная и слабая формы

Для задачи Пуассона $-\Delta u = f$ с нулевыми граничными условиями:

Сильная: $-\Delta u = f$ в Ω

Слабая: $\int \nabla u \cdot \nabla v \, dx = \int f \cdot v \, dx$ для всех v

Вариационная: u минимизирует $E[u] = \int \left(\tfrac{1}{2}|\nabla u|^2 - f \cdot u \right) dx$

Слабая форма требует лишь первых производных — это снижает порядок дифференцирования.

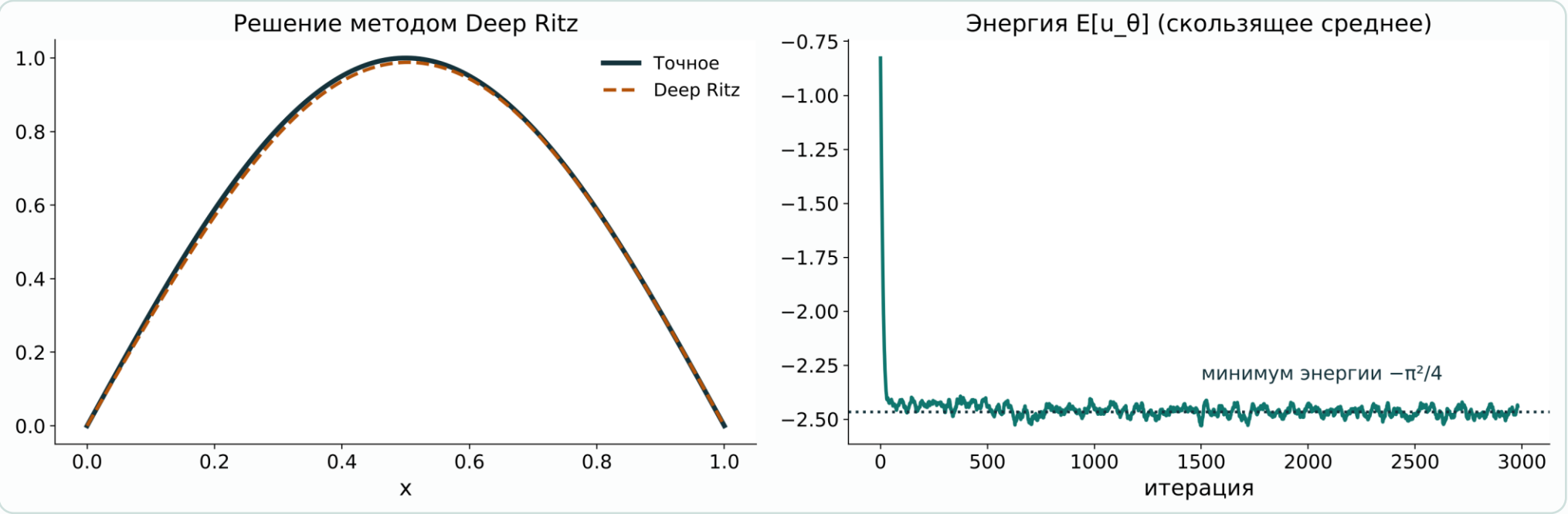
Метод Deep Ritz

Подставляем нейросеть в функционал энергии и минимизируем его по параметрам. Интегралы оцениваются методом Монте-Карло (E и Yu, 2018).

$$E[u_\theta] \approx (|\Omega|/N)\cdot\Sigma_j \left(\tfrac{1}{2}|\nabla u_\theta(x_j)|^2 - f(x_j)\cdot u_\theta(x_j) \right)$$

- Требуются только первые производные сети.
- Подходит для задач, имеющих вариационную формулировку.
- Граничные условия — штраф или жёсткий ансатц.

Пример: Deep Ritz для уравнения Пуассона



Задача $-u'' = \pi^2 \sin(\pi x)$, ансатц $u = x(1-x)\cdot N_\theta(x)$, 4096 случайных точек на шаге. Максимальная ошибка около 0,02; энергия стремится к минимуму $-\pi^2/4$.

Метод Deep Galerkin

Минимизируется среднеквадратичная невязка уравнения в случайно выбираемых точках области; применим к нелинейным и высокоразмерным задачам (Sirignano и Spiliopoulos, 2018).

$$L(\theta) = E \| \partial u_\theta/\partial t + N[u_\theta] \|^2 + \text{граничные и начальные члены}$$

По сути это PINN со стохастической выборкой точек и особой архитектурой типа LSTM; ориентирован на высокие размерности.

Вариационные PINN

Невязка проверяется не в точках, а интегрально — против набора тестовых функций v_k (Kharazmi, Zhang, Karniadakis, 2019–2021).

$$\int R(u_\theta)\cdot v_k \, dx = 0, \; k = 1, \ldots, K$$

- После интегрирования по частям порядок производных у сети понижается.
- Тестовые функции — полиномы или кусочные функции (hp-VPINN).
- Квадратура вместо случайных точек.

Сравнение подходов

Метод	Что минимизируется	Порядок производных	Ограничение
PINN	Невязка в точках	Порядок уравнения	Нужна гладкая сеть, баланс потерь
Deep Ritz	Функционал энергии	Первый	Нужна вариационная постановка
Deep Galerkin	Невязка, случайная выборка	Порядок уравнения	Высокая дисперсия оценок
VPINN	Слабая невязка	Понижен интегрированием по частям	Выбор тестовых функций, квадратур

Высокие размерности

Классические сетки требуют порядка N^d узлов. Методы Монте-Карло с нейросетями — Deep Ritz, Deep Galerkin — не строят сетку и масштабируются на десятки и сотни переменных.

Преимущество

Отсутствие сетки, оценки интегралов по случайным точкам.

Цена

Шум стохастических оценок; точность обычно ниже, чем у сеток в малых размерностях.

Как выбрать метод

- Есть вариационная форма и нужны первые производные — Deep Ritz.
- Сильно нелинейная или высокоразмерная задача — Deep Galerkin.
- Нужна точность на гладких решениях и контроль порядка — VPINN.
- Простая геометрия — жёсткие условия через ансатц; сложная — штраф или функции расстояния.

Выводы

- Жёсткие ограничения встраивают граничные условия в архитектуру и убирают штрафные члены потерь.
- Слабая форма и энергия требуют производных меньшего порядка, чем сильная форма.
- Deep Ritz минимизирует функционал энергии методом Монте-Карло.
- Deep Galerkin и вариационные PINN дополняют метод невязки случайной выборкой и тестовыми функциями.
- Выбор метода зависит от постановки, геометрии и размерности.

Контрольные вопросы

1. Чем жёсткие ограничения отличаются от мягких?
2. Как построить ансатц для условий Дирихле на отрезке?
3. Запишите слабую и вариационную формы задачи Пуассона.
4. В чём суть метода Deep Ritz?
5. Чем Deep Galerkin отличается от обычного PINN?
6. Какие преимущества у вариационных PINN?

Литература

1. E W., Yu B. The Deep Ritz method: a deep learning-based numerical algorithm for solving variational problems // Commun. Math. Stat. 2018. Vol. 6. P. 1–12.
2. Sirignano J., Spiliopoulos K. DGM: a deep learning algorithm for solving partial differential equations // J. Comput. Phys. 2018. Vol. 375. P. 1339–1364.
3. Kharazmi E., Zhang Z., Karniadakis G.E. hp-VPINNs: variational physics-informed neural networks with domain decomposition // Comput. Methods Appl. Mech. Eng. 2021. Vol. 374.
4. Lagaris I.E., Likas A., Fotiadis D.I. Artificial neural networks for solving ordinary and partial differential equations // IEEE Trans. Neural Networks. 1998. Vol. 9. P. 987–1000.
5. Lu L. et al. Physics-informed neural networks with hard constraints for inverse design // SIAM J. Sci. Comput. 2021. Vol. 43. P. B1105–B1132.
6. Sukumar N., Srivastava A. Exact imposition of boundary conditions with distance functions in physics-informed deep neural networks // Comput. Methods Appl. Mech. Eng. 2022. Vol. 389.