КАЗАХСКИЙ НАЦИОНАЛЬНЫЙ УНИВЕРСИТЕТ ИМЕНИ АЛЬ-ФАРАБИ · КАФЕДРА КОМПЬЮТЕРНЫХ НАУК

Научное машинное обучение

Лекция 5. Проблемы обучения PINN: балансировка потерь, спектральное смещение, адаптивная выборка

Почему PINN плохо обучаются и как это исправить

Кенжебек Ержан Ғалымжанұлы

PhD, старший преподаватель кафедры компьютерных наук

Цель и план лекции

Цель

Разобрать основные причины трудностей обучения PINN и методы их преодоления: адаптивное взвешивание потерь, борьба со спектральным смещением, адаптивный выбор коллокационных точек.

План

1. Почему обучение PINN трудно
2. Дисбаланс членов функции потерь
3. Методы балансировки весов
4. Спектральное смещение
5. Адаптивная выборка точек
6. Причинность и другие приёмы

Почему PINN трудно обучать

Конкурирующие цели

Невязка, граница и данные тянут параметры в разные стороны.

Плохая обусловленность

Дифференциальные операторы усиливают плохую обусловленность ландшафта потерь.

Низкая частота

Сеть первой учит гладкие компоненты и медленно — резкие.

Малая невязка в точках не гарантирует малую ошибку решения (Krishnapriyan et al., 2021).

Дисбаланс потерь

Градиенты членов L_r и L_b могут отличаться на порядки; оптимизатор подстраивается под доминирующий член, а остальные недообучаются (Wang, Teng, Perdikaris, 2021).

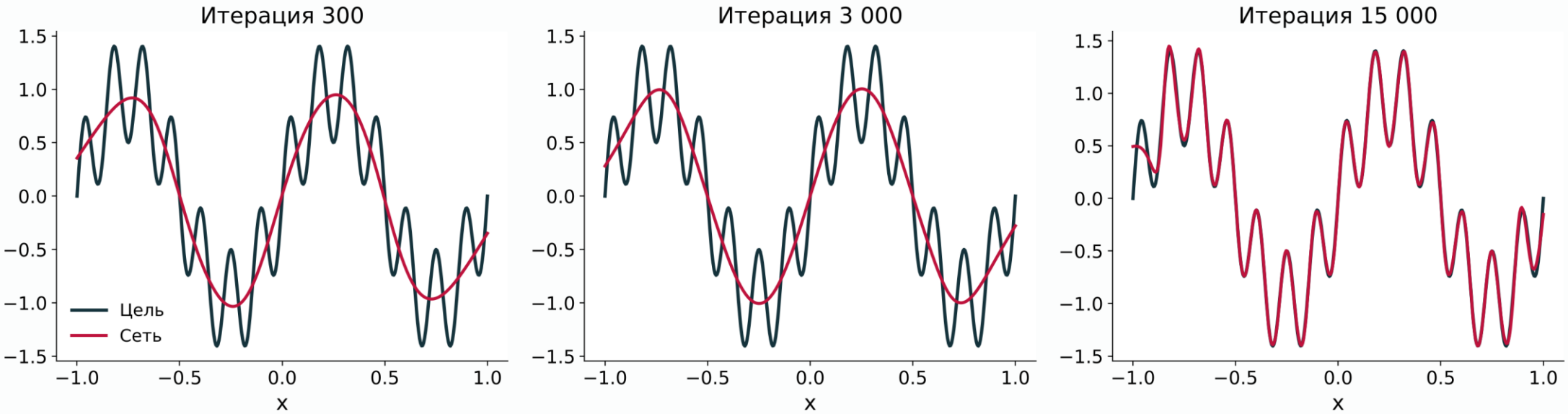
$$L(\theta) = \lambda_r \cdot L_r + \lambda_b \cdot L_b + \lambda_o \cdot L_o$$

Симптом: невязка мала, но граничные условия нарушены — или наоборот.

Методы балансировки весов

Подход	Идея
Ручной подбор	Перебор λ по сетке; дорого и нестабильно
По статистике градиентов	Вес подстраивается по отношению норм градиентов членов (Wang et al., 2021)
По NTK	Веса из собственных значений ядра нейрокасательной (Wang, Yu, Perdikaris, 2022)
Самоадаптивные веса	Вес каждой точки — обучаемый параметр, растущий там, где ошибка велика (McClenny, Braga-Neto)

Спектральное смещение



Цель: sin(2πx) + 0,5·sin(14πx). За 3000 итераций сеть выучила лишь низкую частоту; высокая частота появляется только после 15 000 итераций.

Как смягчить спектральное смещение

- Фурье-признаки на входе: $\gamma(x) = [\sin(2\pi Bx), \cos(2\pi Bx)]$ (Tancik et al., 2020).
- Многомасштабные сети (Wang, Wang, Perdikaris, 2021).
- Периодические активации (SIREN, Sitzmann et al., 2020).
- Разбиение области и обучение отдельных сетей.

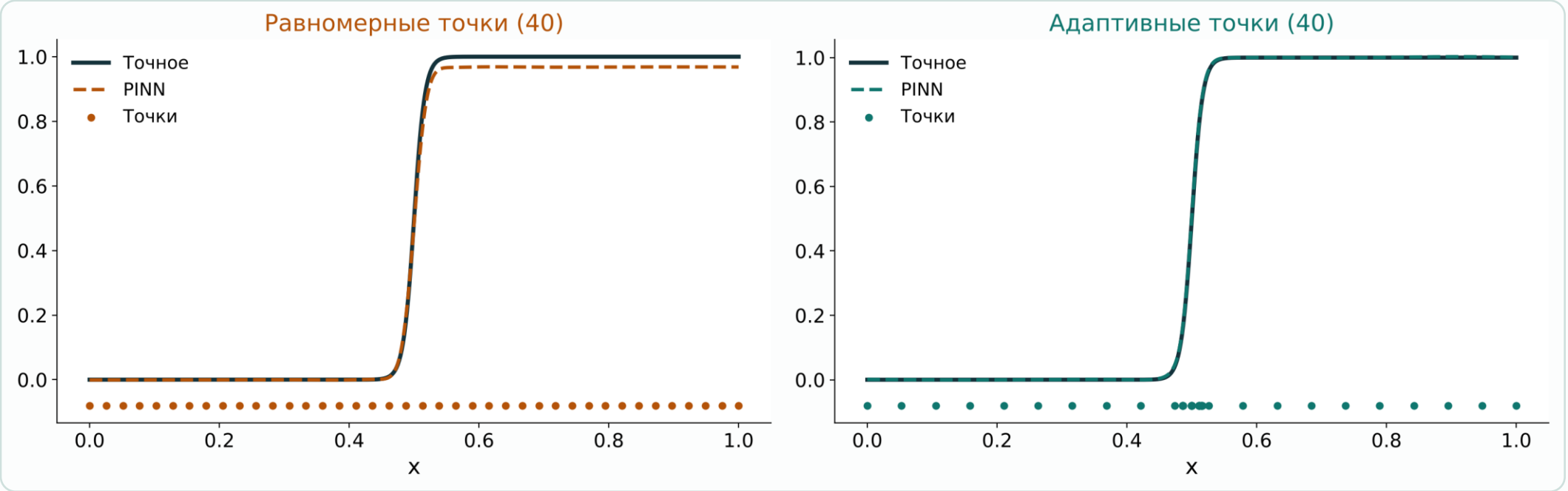
Эффект связан с принципом частоты (Rahaman et al., 2019): сети быстрее подгоняют низкие частоты.

Адаптивная выборка точек

Точки лучше ставить там, где невязка велика: у фронтов, пограничных слоёв и ударных волн.

- RAR: периодически добавлять точки с наибольшей невязкой (Lu et al., 2021).
- RAD и RAR-D: вероятность выбора пропорциональна невязке (Wu et al., 2023).
- Периодическая перевыборка всех точек.

Эксперимент: равномерные и адаптивные точки



Задача $u' = g$ с резким ступенчатым профилем $\tanh(60(x-0{,}5))$. При 40 точках среднеквадратичная ошибка 0,023 для равномерной сетки и 0,002 для адаптивной; точки сгустились у фронта.

Причинность во времени

Для эволюционных уравнений PINN могут минимизировать невязку в поздние моменты раньше, чем выучат начальное условие, и получать неверное решение.

Каузальное обучение (Wang, Sankaran, Perdikaris, 2024) взвешивает вклад момента времени t по накопленной невязке на более ранних моментах: сначала учим начало, затем дальше.

Чек-лист диагностики

- Следить за каждым членом потерь отдельно.
- Сравнивать с эталоном на простой задаче с известным решением.
- Проверять карту невязки: где она сосредоточена.
- Менять веса, активацию, число и положение точек по одному.
- Завершать L-BFGS и проверять устойчивость к инициализации.

Другие приёмы

Декомпозиция области

XPINN и cPINN: отдельные сети на подобластях.

Обучение по этапам

Постепенное усложнение: от малых времён и вязкости к целевым.

Жёсткие условия

Ансатц, автоматически выполняющий границу — лекция 6.

Выводы

- Обучение PINN осложняют конкурирующие члены потерь, плохая обусловленность и спектральное смещение.
- Дисбаланс градиентов лечат адаптивным взвешиванием потерь.
- Сеть учит низкие частоты раньше высоких; помогают Фурье-признаки и многомасштабные сети.
- Адаптивная выборка сгущает точки у резких особенностей и повышает точность при том же числе точек.
- В эволюционных задачах важно учитывать причинность.

Контрольные вопросы

1. Какие основные трудности возникают при обучении PINN?
2. Что такое дисбаланс градиентов и как он проявляется?
3. Опишите два способа адаптивного взвешивания потерь.
4. Что такое спектральное смещение и как с ним бороться?
5. Как работает метод RAR?
6. Зачем в эволюционных задачах нужно каузальное обучение?

Литература

1. Wang S., Teng Y., Perdikaris P. Understanding and mitigating gradient flow pathologies in physics-informed neural networks // SIAM J. Sci. Comput. 2021. Vol. 43. P. A3055–A3081.
2. Wang S., Yu X., Perdikaris P. When and why PINNs fail to train: a neural tangent kernel perspective // J. Comput. Phys. 2022. Vol. 449.
3. Krishnapriyan A. et al. Characterizing possible failure modes in physics-informed neural networks // NeurIPS. 2021.
4. Rahaman N. et al. On the spectral bias of neural networks // ICML. 2019.
5. Tancik M. et al. Fourier features let networks learn high frequency functions in low dimensional domains // NeurIPS. 2020.
6. Lu L. et al. DeepXDE: a deep learning library for solving differential equations // SIAM Review. 2021. Vol. 63. P. 208–228.
7. Wang S., Sankaran S., Perdikaris P. Respecting causality for training physics-informed neural networks // Comput. Methods Appl. Mech. Eng. 2024. Vol. 421.