Implicit weak adversarial networks to solve stochastic partial differential equations

Кафедра интеллектуального анализа данных

Доклад подготовил: Папай Иван

Постановка задачи: Эллиптические PDE

Общая форма краевой задачи

Найти $u : \Omega \subset \mathbb{R}^d \to \mathbb{R}$ такую, что:

$$\begin{cases} -\sum_{i=1}^d \partial_i \left(\sum_{j=1}^d a_{ij} \partial_j u \right) + \sum_{i=1}^d b_i \partial_i u + cu = f, & \text{в } \Omega \\ u = g \text{ (Дирихле)} \quad \text{или} \quad \frac{\partial u}{\partial \vec{n}} = g \text{ (Нейман)}, & \text{на } \partial\Omega \end{cases}$$

Условие сильной эллиптичности

Существует $\theta > 0$ такое, что:

$$\xi^\top A(x) \xi \geq \theta |\xi|^2 \quad \forall \xi \in \mathbb{R}^d, \; x \in \Omega \text{ п.в.}$$

где $A(x) = [a_{ij}(x)] \in \mathbb{R}^{d \times d}$ симметрична.

Параболические PDE

Начально-краевая задача

Найти $u : \Omega \times [0, T] \to \mathbb{R}$ такую, что:

$$\begin{cases} u_t - \sum_{i=1}^d \partial_i \left(\sum_{j=1}^d a_{ij} \partial_j u \right) + \sum_{i=1}^d b_i \partial_i u + cu = f, & \text{в } \Omega \times [0, T] \\ \text{Граничные условия на } \partial\Omega \times [0, T] \\ u(x, 0) = h(x), & \text{в } \Omega \end{cases}$$

Операторная форма

$$u_t + L(x, t; u) = f$$

где

$$L(x, t; u) = -\nabla \cdot (A \nabla u) + \mathbf{b} \cdot \nabla u + cu$$

Слабая формулировка PDE

Общий вид эллиптического PDE

$$-\nabla \cdot (A \nabla u) + \mathbf{b} \cdot \nabla u + cu = f \quad \text{в } \Omega$$

Слабая форма

Найти $u \in H^1(\Omega)$ такую, что:

$$\langle A[u], \varphi \rangle = 0 \quad \forall \varphi \in H_0^1(\Omega)$$

где

$$\langle A[u], \varphi \rangle = \int_\Omega (A \nabla u \cdot \nabla \varphi + \mathbf{b} \varphi \cdot \nabla u + cu\varphi - f\varphi)\, dx$$

Ключевая идея WAN

Минимаксная переформулировка

$$\|A[u]\|_{op} = \sup_{\varphi \neq 0} \frac{|\langle A[u], \varphi \rangle|}{\|\varphi\|_2}$$

$$u \text{ – решение} \Leftrightarrow \|A[u]\|_{op} = 0$$

Задача оптимизации

$$\min_u \max_\varphi \frac{|\langle A[u], \varphi \rangle|^2}{\|\varphi\|_2^2}$$

- $\min_u$: найти решение PDE
- $\max_\varphi$: найти "наихудшую"тестовую функцию

Функция потерь и обучение

Общая функция потерь

$$L(\theta, \eta) = \underbrace{\log |\langle A[u_\theta], \varphi_\eta \rangle|^2 - \log \|\varphi_\eta\|_2^2}_{\text{Внутренняя область}} + \alpha \underbrace{L_{\text{bdry}}(\theta)}_{\text{Граница}}$$

Adversarial обучение

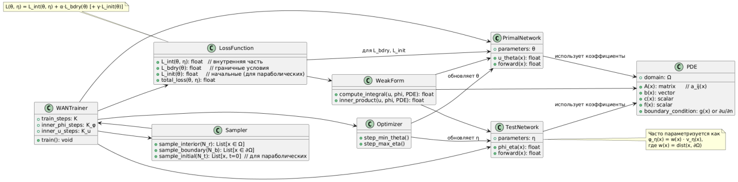
1. Фиксируем u_θ, обновляем φ_η:

$$\eta \leftarrow \eta + \tau_\eta \nabla_\eta L(\theta, \eta)$$

2. Фиксируем φ_η, обновляем u_θ:

$$\theta \leftarrow \theta - \tau_\theta \nabla_\theta L(\theta, \eta)$$

Архитектура WAN: Две соревнующиеся сети



Primal Network u_θ

- Аппроксимирует **решение** PDE
- Минимизирует остаток
- Выход: $u_\theta(x)$

Test Network φ_η

- Аппроксимирует **тестовую функцию**
- Максимизирует остаток
- Выход: $\varphi_\eta(x)$

Параболические PDE: Уравнение теплопроводности

Постановка задачи

$$\begin{cases} \dfrac{\partial u}{\partial t} = \nabla \cdot (D \nabla u) + f(x,t), & (x,t) \in \Omega \times [0, T] \\ u(x,0) = u_0(x), & x \in \Omega \\ u = g \text{ на } \partial\Omega \end{cases}$$

Слабая форма с интегрированием по времени

$$\langle A[u], \varphi \rangle = \int_\Omega (u_T \varphi_T - u_0 \varphi_0) dx - \int_0^T \int_\Omega u \partial_t \varphi dx dt +$$

$+\int_0^T \int_\Omega D \nabla u \cdot \nabla \varphi dx dt$

Сравнение архитектур: WAN, IWAN, XWAN

Параметр	**WAN**	**IWAN**	**XWAN**
Активации	ReLU/tanh	**SIREN (sin)**	Neural ODE
Производные	Кусочно-гладкие	**Аналитические**	Точные (ODE)
Скорость	Быстрая	Средняя	**Оч. медл.**
Сходимость	Медл. (гладк.)	**Быстр. (гладк.)**	Теор. идеал.
Память	Малая	Малая	**Огромная**
Практичность	**Высокая**	**Высокая**	Низкая

Подходы по моделированию primal network

- **IWAN** (SIREN): SIREN-Block
- **XWAN**: XNode-Block
- **WAN**: DNN

Применение к диффузионным моделям

Связь с уравнением Фоккера-Планка

Диффузионный процесс:

$$dx_t = f(x_t, t)dt + g(t)dw_t$$

Его плотность $p(x, t)$ удовлетворяет:

$$\frac{\partial p}{\partial t} = -\nabla \cdot (fp) + \frac{1}{2}g^2 \Delta p$$

WAN для обучения диффузионных моделей

- Параметризация: $p_\theta(x, t)$ или $s_\theta(x, t) = \nabla \log p_\theta$
- Обучение через слабую форму Фоккера-Планка
- **Преимущество**: Не нужны сэмплы из процесса

Заключение

Основные свойства WAN

- **Высокие размерности**: $d \sim 100$ и более
- **Физическая интерпретация**: Слабая форма

Что предлагается

- **IWAN**: Теория SIREN для PDE + доказательство сходимости
- **Связь**: Weak PDE solvers $\leftrightarrow$ Diffusion models
- **Приложения**: Генеративное моделирование через PDE