WAN-ы для решения высокоразмерных дифференциальных уравнений в частных производных

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: Математические основы

Пространства Соболева

- $H^1(\Omega) = \{u \in L^2(\Omega) : \nabla u \in L^2(\Omega)^d\}$
- $H_0^1(\Omega) = \{u \in H^1(\Omega) : u|_{\partial\Omega} = 0\}$
- Норма: $\|u\|_{H^1} = \left(\int_\Omega |u|^2 + |\nabla u|^2 dx\right)^{1/2}$

Слабая форма эллиптического PDE

Найти $u \in H^1(\Omega)$ такую, что:

$$\langle A[u], \varphi \rangle = \int_\Omega \left(\sum_{i,j=1}^{d} a_{ij} \partial_j u \partial_i \varphi + \sum_{i=1}^{d} b_i \varphi \partial_i u + cu\varphi - f\varphi \right) dx = 0$$

для всех $\varphi \in H_0^1(\Omega)$.

Операторная норма и минимаксная формулировка

Оператор $A[u]$ как линейный функционал

$$A[u] : H_0^1(\Omega) \to \mathbb{R}, \quad A[u](\varphi) = \langle A[u], \varphi \rangle$$

Индуцированная операторная норма

$$||A[u]||_{op} = \sup \left\{ \frac{|\langle A[u], \varphi \rangle|}{||\varphi||_2} : \varphi \in H_0^1(\Omega), \varphi \neq 0 \right\}$$

где $||\varphi||_2 = \left(\int_\Omega |\varphi(x)|^2 dx \right)^{1/2}$.

Ключевое наблюдение

u - слабое решение $\Leftrightarrow$ $||A[u]||_{op} = 0$ и $B[u] = 0$ на $\partial\Omega$

Минимаксная переформулировка

Theorem (Эквивалентность)

Следующие задачи эквивалентны:

$$\text{(1)}\quad \min_{u \in H^1} ||A[u]||_{op}^2$$

$$\text{(2)}\quad \min_{u \in H^1} \max_{\varphi \in H_0^1} \frac{|\langle A[u], \varphi \rangle|^2}{||\varphi||_2^2}$$

Доказательство.

$$\begin{aligned} ||A[u]||_{op}^2 &= \sup_{\varphi \neq 0} \frac{|\langle A[u], \varphi \rangle|^2}{||\varphi||_2^2} \\ &= \max_{\varphi \in Y} |\langle A[u], \varphi \rangle|^2, \quad Y = \{\varphi \in H_0^1 : ||\varphi||_2 = 1\} \end{aligned}$$

$\square$

Weak Adversarial Network: Архитектура

Параметризация сетями

- **Primal network**: $u_\theta : \mathbb{R}^d \to \mathbb{R}$ - слабое решение
- **Adversarial network**: $\varphi_\eta : \mathbb{R}^d \to \mathbb{R}$ - тестовая функция

Функция потерь для внутренней области

$$L_{\text{int}}(\theta, \eta) = \log |\langle A[u_\theta], \varphi_\eta \rangle|^2 - \log ||\varphi_\eta||_2^2$$

Граничные условия

- **Дирихле**: $L_{\text{bdry}}(\theta) = \frac{1}{N_b} \sum_{j=1}^{N_b} |u_\theta(x_b^{(j)}) - g(x_b^{(j)})|^2$
- **Нейман**: $L_{\text{bdry}}(\theta) = \frac{1}{N_b} \sum_{j=1}^{N_b} \left| \frac{\partial u_\theta}{\partial \vec{n}}(x_b^{(j)}) - g(x_b^{(j)}) \right|^2$

Общая задача оптимизации

Минимаксная задача

$$\min_{\theta} \max_{\eta} L(\theta, \eta) = L_{\text{int}}(\theta, \eta) + \alpha L_{\text{bdry}}(\theta)$$

где $\alpha > 0$ - параметр баланса.

Седловая точка

Оптимальные параметры (θ^*, η^*) удовлетворяют:

$$L(\theta^*, \eta) \leq L(\theta^*, \eta^*) \leq L(\theta, \eta^*)$$

$$\Updownarrow$$

u_{θ^*} - слабое решение PDE

Вычисление градиентов

Интегральное ядро

Для эллиптического PDE:

$$I(x;\theta,\eta) = \sum_{i,j=1}^{d} a_{ij}(x)\partial_j u_\theta(x)\partial_i \varphi_\eta(x) + \sum_{i=1}^{d} b_i(x)\varphi_\eta(x)\partial_i u_\theta(x) +$$

$$+ c(x)u_\theta(x)\varphi_\eta(x) - f(x)\varphi_\eta(x)$$

Градиенты

$$\nabla_\theta I = \sum_{i,j} a_{ij}\partial_j \nabla_\theta u_\theta \partial_i \varphi_\eta + \sum_i b_i \varphi_\eta \partial_i \nabla_\theta u_\theta + c\nabla_\theta u_\theta \varphi_\eta$$

$$\nabla_\eta I = \sum_{i,j} a_{ij}\partial_j u_\theta \partial_i \nabla_\eta \varphi_\eta + \sum_i b_i \nabla_\eta \varphi_\eta \partial_i u_\theta + c u_\theta \nabla_\eta \varphi_\eta - f\nabla_\eta \varphi_\eta$$

Дискретизация и приближение

Монте-Карло интегрирование

$$\langle A[u_\theta], \varphi_\eta \rangle \approx \frac{1}{N_r} \sum_{j=1}^{N_r} I(x_r^{(j)}; \theta, \eta)$$

$$||\varphi_\eta||_2^2 \approx \frac{1}{N_r} \sum_{j=1}^{N_r} |\varphi_\eta(x_r^{(j)})|^2$$

Градиент по θ

$$\nabla_\theta L_{\text{int}} \approx 2 \left(\sum_{j=1}^{N_r} I(x_r^{(j)}) \right)^{-1} \left(\sum_{j=1}^{N_r} \nabla_\theta I(x_r^{(j)}) \right)$$

Функция веса для граничных условий

Факторизация тестовой функции

$$\varphi_\eta = w \cdot v_\eta$$

где $w : \Omega \to \mathbb{R}$ удовлетворяет:

- $w(x) > 0$ для $x \in \Omega$
- $w(x) = 0$ для $x \in \partial\Omega$

Функция расстояния

$$w(x) = \mathrm{dist}(x, \partial\Omega) = \inf\{|x - y| : y \in \partial\Omega\}$$

Преимущества

- Автоматическое выполнение $\varphi_\eta|_{\partial\Omega} = 0$
- Упрощение тренировки
- Повышение устойчивости

WAN для стационарных PDE

1. Инициализировать параметры θ, η
2. Для $k = 1, 2, \ldots, K$:
   2.1 Сэмплировать $\{x_r^{(j)}\}_{j=1}^{N_r} \subset \Omega$, $\{x_b^{(j)}\}_{j=1}^{N_b} \subset \partial\Omega$
   2.2 Для $m = 1, \ldots, K_\varphi$:

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

   2.3 Для $n = 1, \ldots, K_u$:

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

Параболические PDE: Два подхода

Метод 1: Полудискретизация по времени

- Crank-Nicolson схема:

$$u^{n+1} - u^n = \frac{h}{2}\left(L(t_{n+1}, u^{n+1}) + f^{n+1} + L(t_n, u^n) + f^n\right)$$

- На каждом шаге решаем эллиптическую задачу

Метод 2: Совместная обработка (x, t)

Рассматриваем $\Omega \times [0, T]$ как $(d+1)$-мерную область

Алгоритм 1: Полудискретизация по времени

1. Задать $u^0(x) = h(x)$
2. Для $n = 0, 1, \ldots, N-1$:
   2.1 Решить эллиптическую задачу для u^{n+1}:

   $$u^{n+1} - \frac{h}{2} L(t_{n+1}, u^{n+1}) = u^n + \frac{h}{2}\left(L(t_n, u^n) + f^{n+1} + f^n\right)$$

   2.2 Использовать алгоритм стац.PDE для нахождения $u_{\theta_{n+1}} \approx u^{n+1}$

Слабая формулировка параболических PDE

Интегрирование по времени и пространству

$$\langle A[u], \varphi \rangle = \int_\Omega \left(u(x, T)\varphi(x, T) - h(x)\varphi(x, 0) \right) dx \\ - \int_0^T \int_\Omega u \partial_t \varphi dx dt + \int_0^T \int_\Omega \\ \left(\sum_{i,j} a_{ij} \partial_j u \partial_i \varphi + \sum_i b_i \varphi \partial_i u + cu\varphi - f\varphi \right) dx dt = 0$$

Обобщенная функция потерь

$$L(\theta, \eta) = L_{\text{int}}(\theta, \eta) + \gamma L_{\text{init}}(\theta) + \alpha L_{\text{bdry}}(\theta)$$

где $\gamma > 0$ - вес начального условия.

Алгоритм 2: Совместная обработка (x, t)

1. Параметризовать $u_\theta(x, t)$, $\varphi_\eta(x, t)$
2. Сэмплировать точки в $\Omega \times [0, T]$
3. Минимизировать расширенную функцию потерь:

$$L(\theta, \eta) = L_{\text{int}}(\theta, \eta) + \gamma L_{\text{init}}(\theta) + \alpha L_{\text{bdry}}(\theta)$$

4. Использовать факторизацию $\varphi_\eta(x, t) = w(x) v_\eta(x, t)$

Оценка ошибки и сходимости

Относительная L^2 ошибка

$$\text{Error} = \frac{||u_\theta - u^*||_2}{||u^*||_2}$$

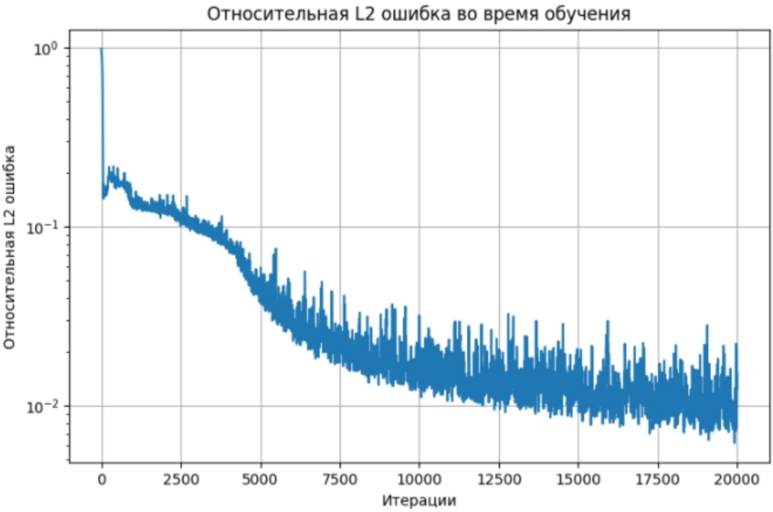
где

$$||u||_2^2 = \int_\Omega |u(x)|^2 dx$$

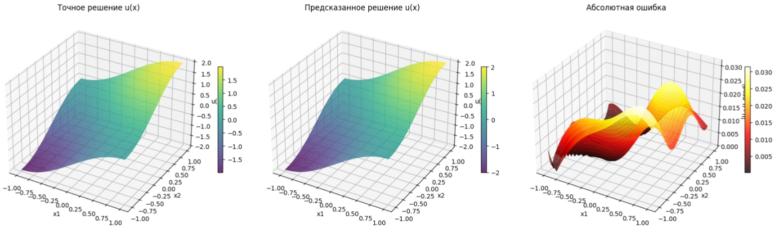
Дискретная аппроксимация

$$||u_\theta - u^*||_2^2 \approx \frac{1}{M} \sum_{j=1}^{M} |u_\theta(x^{(j)}) - u^*(x^{(j)})|^2$$

Пример обучения



Пример обучения



Заключение

- Новая минимаксная формулировка для слабых решений PDE
- Adversarial подход с разделением на primal и test networks
- Эффективное решение высокоразмерных задач
- Работа на произвольных областях, не только в физике