Цель работы

Рассмотреть задачу распределения тепла в диске и задачу электрокинетики, решить их с помощью PINN и сравнить полученные данные с решениями, полученными другими способами, оценить целесообразность применения PINN к задаче.

Описание задачи

Создать нейросеть, которой на вход подаются пространственные координаты. На выходе хотим значение физической величины в данной точке. Обучить данную нейросеть используя методику PINN. Оценить полученные результаты.

Слайд 2 Дифференциальные уравнения повсеместно встречаются в математической физике, однако редко когда их удаётся решить аналитически. Одним из подходов к решению дифференциальных уравнений могут служить Physics-informed neural networks (PINN, физически информированные нейронные сети) – разновидность нейронных сетей, способные решать задачи математической физики. В отличии от классических нейронных сетей, использующих большую выборку данных для обучения, PINN используют уравнения, описывающие физическую систему, что позволяет им обучаться на сравнительно небольших объёмах обучающих данных.

Слайд 3 PINN основаны на глубоких нейронных сетях, рассмотрим их устройство: Нейроном сети будем называть такую функцию

$$f(z, \theta) = h(wz + b),$$

Здесь z – вектор-столбец входов нейрона, θ – параметры нейрона, состоящие из вектор-строки весов w_j и смещения b. h – функция активации нейрона, обычно это tanh, сигмоида или ReLu.

Глубокая нейронная есть, состоящая из L скрытых слоёв по N нейронов в каждом может быть записана следующим образом:

$$\begin{aligned} q^{(0,n)} &= h(w^{(0,n)}z + b^{(0,n)}), n = 1, ..., N \\ q^{(l,n)} &= h(w^{(l,n)q^{(l-1,i)}+b^{(l,n)}}), n = 1, ..., N,\ l = 1, ..., L-1 \\ q^{(L)} &= w^{(L)}q^{(L-1)} + b^{(L)}. \end{aligned}$$

Кратко записать её можно так $\bar{u}(z) = q^{(L)}$. За θ обозначим все параметры всех нейронов.

Обучением сети $\bar{u}$ на выборке $T = \{z_i, u_i\}_{i=1}^{N_f}$ называется поиск таких

$$\theta = \min_{\theta} \|\bar{u}(z_i, \theta) - u_i\|, i = 1, ..., N_f.$$

Для поиска параметров θ используется градиентный спуск и его модификации, я буду использовать Adam.

Слайд 4

Рассмотрим теперь устройство самой PINN. Пусть у нас имеется система уравнений,

$$F_j(z, u, \lambda_j) = F_j(z, u, u'_{z^1}, u''_{z^1}, ..., \lambda_j) = 0, z \in \Omega, j = \overline{1, N},$$

определённая в области Ω, z – независимые переменные, u – искомая функция, λ_j – параметры системы. Граничные условия для неё

$$B_k(z_0) = 0, z_0 \in \partial\Omega, k = \overline{1, K},$$

z_0 – лежат на границе области Ω.

Обозначим за $\bar{u}$ – нейронную сеть, приближающую решение u.

Составим следующую функцию потерь:

$$\begin{aligned} MSE &= MSE_f + MSE_b \\ &= \sum_{j=1}^{N} \frac{1}{N_f} \sum_{i=1}^{N_f} F_j^2(z_i, \bar{u}(z_i), \lambda_j) + \sum_{k=1}^{K} \frac{1}{N_b} \sum_{b=1}^{N_b} (\bar{u}(z_b) - u_b)^2. \end{aligned}$$

Здесь MSE_f отвечает за выполнение уравнений системы F_j, $\{z_i\}$ – точки коллокации, они равномерно распределены по области Ω. MSE_b отвечает за выполнение граничных условий, z_b – равномерно распределены по границе Ω, пара (z_b, u_b) получено из граничных условий.

Заметим, что вычисления MSE_f нужно уметь считать производные $\bar{u}$ в точках коллокации z_i. Для этого мы воспользуемся автоматическим дифференцированием. Суть автоматического дифференцирования заключается в том, что мы можем выразить производную сложной функции, составленную из элементарных функций через значения этих функций в точке и значения их производных и так как производные всех элементарных функций известны, то и посчитать производную сложной функции не составляет труда. Наша нейронная сеть $\bar{u}$ как раз и является такой сложной функцией.

Слайд 5

Здесь изображена принципиальная схема работы PINN, сеть принимает на вход переменные, вычисляет функцию, затем с помощью автоматического дифференцирования вычисляются её частные производные, из них составляются дифференциальные уравнения, они входят в функцию потерь, и, если функция потерь достаточно маленькая то мы считаем что сеть обучилась и завершаем обучение, иначе обновляем параметры сети и запускаем процесс заново.

Слайд 6

Рассмотрим в начале относительно простую задачу распределения тепла в кольце, без внутренних источников тепла, с такими граничными условиями.

Сеть возьмём с 4 слоями по 20 нейронов в каждом, функция активации $tanh$, оптимизатор Adam, Эпох 10000. Поэкспериментируем с различными соотношениями количества точек коллокации и граничных точек. А именно N_f возьмём 1000, 5000, 10000. N_b возьмём 100, 500, 1000.

Слайд 7

Посмотрим на графики обучения. По ним видно, что при $N_f = 1000$ сеть обучается не очень хорошо, однако при 5000 и 10000 обучение проходит хорошо.

Слайд 8

Посмотрим на графики решений. По ним так же видно что сети с $N_f = 1000$ имеют разрыв при $\phi = 0$. Так же полученные решения хорошо согласуются с аналитическим решением.

Слайд 8

Рассмотрим теперь более сложную задачу электрокинетики для щелевой поры, состоящей из двух одноименно заряженных параллельных пластин бесконечной протяженности, удерживающих раствор воды и противоионов пластин. Система описывается следующим набором уравнений.

$$\begin{aligned}
\vec{j} &= -D\nabla c - \xi zec\nabla\Phi + cv,\\
\frac{\partial c}{\partial t} &= -\nabla\cdot\vec{j},\\
\nabla^2\Phi &= -4\pi l_{\mathrm{B}} k_{\mathrm{B}} Tzc,\\
\rho\big(\frac{\partial v}{\partial t} + (v\cdot\nabla)v\big) &= -\nabla p_H + \eta\nabla^2 v - (k_{\mathrm{B}}T\nabla c + zec\nabla\Phi),\\
\nabla\cdot v &= 0.
\end{aligned}$$

И граничными условиями

$$\begin{aligned}
&c(t, X_l) = 0.01, c(t, X_r) = 0.01, c(0, X) = 0.002\\
&v(t, X_l) = 0, v(t, X_r) = 0, v(0, X) = 0\\
&\Phi(t, X_l) = -0.05, \Phi(t, X_r) = -0.05, \Phi(0, X) = -0.009x^2 + 2.
\end{aligned}$$

здесь t – время, $t > 0$, X – пространственные координаты. c – концентрация противоионов, v – скорость жидкости, Φ – потенциал.

Слайд

Рассмотрим двухмерный случай, тогда сеть $\bar{u}$ имеет 4 выхода, 1 для c, 2 для v, 1 для Φ. В качестве архитектуры так же возьмём 4 слоя по 20 нейронов, с функцией активации tanh, оптимизатор Adam, эпох 10000. График всего обучения не слишком информативен, так как в начале функция потерь уменьшается очень быстро. Рассмотрим участок с 100-ой по 1000-ную эпохи. Из этого участка явно видно, что обучение происходило примерно до 500 эпохи после чего остановилось. Если мы взглянем на график концентрации полученный PINN и сравним его с эталонным решением системы то увидим, что сеть не справилась с нахождением решения, вероятнее всего она нашла нулевое решение и застряла в нём.

Слайд

Рассмотрим трёхмерный случай сеть тут имеет 5 выходов 1 для концентрации c, 3 для скорости v, 1 для потенциала Φ. В целом ситуация аналогична двухмерному случаю, обучение остановилось в районе 400-ой эпохи. Найденное решение похоже на нулевое.

Слайд

Результаты работы:

- Для тестовой системы применен метод решения с помощью PINN и получено хорошее согласование с аналитическим решением.

- Для системы описывающей задачу электрокинетики в системе щелевой поры применён метод решения с помощью PINN и получено неудовлетворительное согласование и поиск причин - предмет дальнейших исследований.

Дополнительно

Электрокинетические явления – физические явления переноса (движения) дисперсной фазы либо дисперсионной среды коллоидной системы относительно друг друга, которые происходят под действием приложенного электрического поля.

Аналитическое решение для тепла

$$\begin{aligned} u(r,\phi) = 1 - \frac{\ln r}{\ln 2} + \left(\frac{-r}{3} + \frac{4}{3r}\right)\sin(\phi) \\ + \left(\frac{-r}{3} + \frac{4}{3r}\right)\cos(\phi) + \left(\frac{r^2}{5} + \frac{4}{5r^2}\right)\sin(2\phi) \\ + \left(\frac{3r^3}{63} + \frac{312}{64r^3}\right)\sin(3\phi) + \left(\frac{16r^4}{255} - \frac{16}{255r^4}\right)\cos(4\phi). \end{aligned}$$