

Предполагаемое название статьи

Применение методов искусственного интеллекта для восстановления коэффициентов в обратной задаче для уравнения типа реакция-диффузия-адвекция с дополнительными данными в моделях самоорганизации в биологических системах.

Цели и задачи работы

Биофизиками МГУ на основе работы [1] была построена следующая модель:

$$\begin{cases} \frac{\partial u}{\partial t} - \frac{\partial^2 u}{\partial x^2} = \lambda_1(u - v(x, t - t_0) - q(x))(u - \alpha(x) + \lambda_2 v(x, t - t_0)), & x \in (0, l), \quad t \in (0, T], \\ u(0, t) = 0, \quad u(1, t) = 0, & t \in (0, T], \\ u(x, 0) = q(x), & x \in [0, l]. \end{cases} \quad (1)$$

Здесь, $\lambda_1, \lambda_2, t_0$ – параметры, а функция $\alpha(x)$ является непрерывной при всех $x \in [0, l]$. Функция $v(x, t)$ определяется следующим образом: она равна нулю при $t < 0$, а при $t \geq 0$ для каждого фиксированного $x \in [0, l]$, выполняющего роль параметра, определяется как решение задачи Коши:

$$\begin{cases} \frac{\partial v}{\partial t} = \varepsilon(-v(x, t) + \alpha(x)), & t \in (0, T], \\ v(x, 0) = 0. \end{cases}$$

Задача заключается в определении функции $0 \leq q(x) \leq \alpha(x)$ (из физических соображений), $x \in [0, l]$, по известной дополнительной информации о значениях функции на известных кривых:

$$x_i(t) = f_1^i(t), \quad u(x_i(t), t) = f_2^i(t), \quad t \in [0, T], \quad i = \overline{1, I}.$$

Ранее мной и нашей научной группой уже решались подобные задачи [2], [3]. Основной идеей решения задачи была итерационная минимизация, основанная на градиентных методах для следующего функционала Тихонова:

$$J[q] = \sum_{i=1}^I \int_0^T \left(u(f_1^i(t), t; q) - f_2^i(t) \right)^2 dt + \alpha \Omega[q].$$

А в качестве начального приближения для минимизации функционала выбиралось хорошее начальное приближение, выбранное с помощью асимптотических методов. Для задачи в этой постановке возможность построения такого асимптотического приближения находится под вопросом, потому что дополнительной информацией ранее были функции "переходного" слоя $f_1(t)$. Однако в данной постановке такой информацией являются произвольные функции в области, на которой определена задача. В данной задаче может иметься несколько локальных минимумов, поэтому использование произвольного начального приближения для решения с помощью численных методов является неоптимальным. Нами было выдвинуто предложение построить хорошее начальное приближение с помощью нейронной сети, которое будет уточнено градиентными методами.

Описание данных

Для решения задачи вводится сетка $x_n = \frac{l}{N}n, t_m = \frac{T}{M}m, i = \overline{0, N}, j = \overline{0, M}$, где N, M - количество узлов. Входными данными для нейронной сети будет I пар функций $f_1^i(t), f_2^i(t)$, выходными данными будут значения заданной функции $q(x)$, в узлах сетки. Пример таких данных (1). Количество данных, для обучения может быть сгенерировано путем решения прямой задачи (1), для набора функций $q(x)$, которые будут задаваться следующим образом:

$$q(x_n) = c + a_0(x_n - b) + a_1(x_n - b)^2 + a_2(x_n - b)^3 + a_3(x_n - b)^4 + a_4(x_n - b)^5 + a_5(x_n - b)^6, \quad n = \overline{0, N},$$

и если $q(x_n)$, меньше нуля, то данное значения ставится равным нулю, а если это значение больше $\alpha(x_i)$, то данный пример отбрасывается. Коэффициенты подбираются из некоторых промежутков, в согласовании с биофизиками, чтобы удовлетворять физическим свойствам задачи. Основное требование – чтобы функция $q(x)$ выглядела, как одиночный пик, примерный вид зеленый график на (2). Для валидационной и тестовой выборки планируется взять функции $q(x)$ в виде суммы пар функций Гаусса, как функции другого типа.

$$q(x_n) = d_0 \exp\left(-\frac{(x_n - e_0)^2}{f_0}\right) + d_1 \exp\left(-\frac{(x_n - e_1)^2}{f_1}\right), \quad n = \overline{0, N},$$

Датасет будет состоять из 100000 функций для обучения, 20000 функций для валидации и 20000 для теста. Данные для обучения модели планируется зашумлять с некоторой относительной ошибкой.

Также остается вопрос выбора функций $f_1^i(t)$. Первый вариант – обучить нейросеть на произвольных парах функций. Второй – взять для задачи фиксированные экспериментальные функции $f_1^i(t)$, и затем обучать нейросеть на функциях $f_2^i(t)$. На данный момент выбран второй вариант, так как для его решения можно использовать более простую модель.

Метрики качества

В качестве метрики будет использоваться L_2 (MSELoss), которая чаще всего используется для таких задач. Так как такая задача еще не была решена, то чем меньше мы получим ошибку, тем лучше. Использование нейронной сети позволит быстрее работать градиентному методу решения задачи, так как он сходится к правильному за бесконечное число итераций. Использование хорошего начального приближения позволит ускорить сходимость на приемлимую точность гораздо быстрее.

Модели и алгоритмы

На данный момент уже есть несколько простых линейных и сверточных моделей, однако полученная с их помощью точность недостаточна. Линейная нейронная сеть была использована для похожей, но более простой обратной задачи [4], другая нейронная сеть применялась для определения коэффициентов при нахождении решения обратной задачи [5]. В представленной работе планируется использовать архитектуру ResNet [6], с заменой двумерных сверток на одномерные.

Результаты

На данный момент существует несколько простых линейных и сверточных нейронных сетей. Результат решения для сверточной нейронной сети представлен черным графиком на (2), и решение обратной задачи, полученное с их помощью и уточненное классическим алгоритмом, изображено на красном графике. В дальнейшем планируется улучшить результаты, используя более сложную структуру.

Список литературы

- [1] Uchimura A, Higuchi M, Minakuchi Y, Ohno M, Toyoda A, Fujiyama A, Miura I, Wakana S, Nishino J, Yagi T. Germline mutation rates and the long-term phenotypic effects of mutation accumulation in wild-type laboratory mice and mutator mice. *Genome Res.* 2015 Aug;25(8):1125-34. doi: 10.1101/gr.186148.114.
- [2] Argun, R.L., Gorbachev, A.V., Lukyanenko, D.V. et al. Features of Numerical Reconstruction of a Boundary Condition in an Inverse Problem for a Reaction–Diffusion–Advection Equation with Data on the Position of a Reaction Front. *Comput. Math. and Math. Phys.* 62, 441–451 (2022). <https://doi.org/10.1134/S0965542522030022>
- [3] Argun, R.; Gorbachev, A.; Levashova, N.; Lukyanenko, D. Inverse Problem for an Equation of the Reaction-Diffusion-Advection Type with Data on the Position of a Reaction Front: Features of the Solution in the Case of a Nonlinear Integral Equation in a Reduced Statement. *Mathematics* 2021, 9, 2342. <https://doi.org/10.3390/math9182342>

- [4] Lukyanenko, D.; Yeleskina, T.; Prigorniy, I.; Isaev, T.; Borzunov, A.; Shishlenin, M. *Inverse Problem of Recovering the Initial Condition for a Nonlinear Equation of the Reaction–Diffusion–Advection Type by Data Given on the Position of a Reaction Front with a Time Delay. Mathematics* 2021, 9, 342. <https://doi.org/10.3390/math9040342>
- [5] Gorbachenko, V.I., Lazovskaya, T.V., Tarkhov, D.A., Vasilyev, A.N., Zhukov, M.V. (2016). *Neural Network Technique in Some Inverse Problems of Mathematical Physics. In: Cheng, L., Liu, Q., Ronzhin, A. (eds) Advances in Neural Networks – ISSN 2016. ISSN 2016. Lecture Notes in Computer Science()*, vol 9719. Springer, Cham. https://doi.org/10.1007/978-3-319-40663-3_36
- [6] Sasha Targ, Diogo Almeida, Kevin Lyman, *Resnet in Resnet: Generalizing Residual Architectures, CoRR, 2016, <https://doi.org/10.48550/arXiv.1603.08029>*

Графики

Рис. 1: Примерный вид функций $f_1^i(t)$, и соответствующие им $f_2^i(t)$

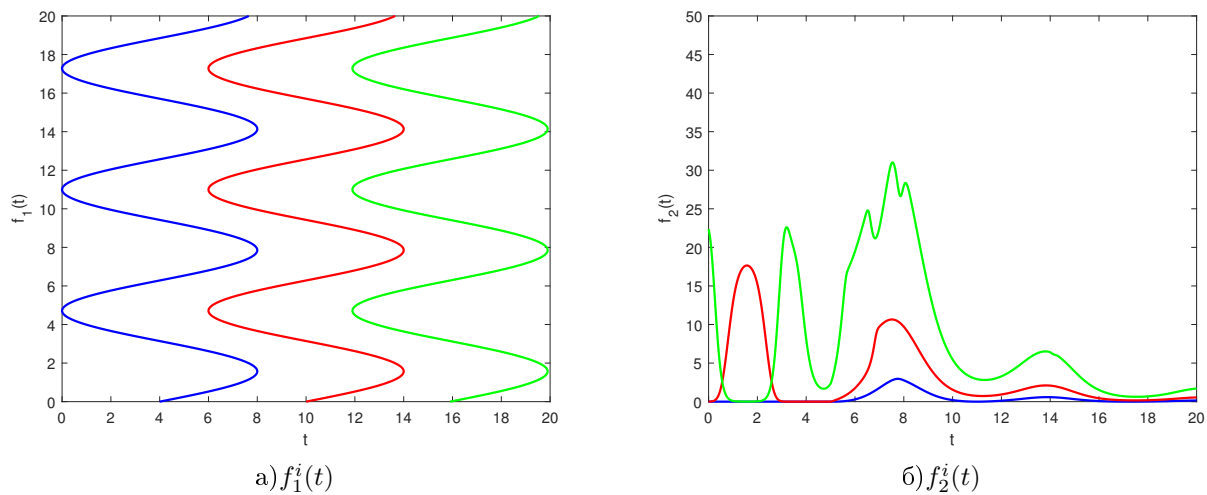


Рис. 2: Результаты решения для модельного примера

