

УДК 004.032.26

Физически-информированные нейронные сети при решении коэффициентных обратных задач

Кузнецов Е.Б.¹, Леонов С.С.^{1,2}

ФГАОУ ВО «Московский авиационный институт (национальный исследовательский университет)»¹,

ФГАОУ ВО «Российский университет дружбы народов имени Патриса Лумумбы»²

Аннотация: В работе рассмотрен процесс создания полуэмпирических рекуррентных нейронных сетей, относящихся к классу физически-информированных нейронных сетей. Предложен подход к построению архитектуры полуэмпирической сети, использующий неявные схемы семейства Рунге-Кутты, который позволяет получить аналитическое функциональное выражение эквивалентное предобученной рекуррентной нейронной сети. Апробация предложенного подхода проведена на тестовой линейной слабо жесткой начальной задаче с одним скалярным параметром.

Ключевые слова: полуэмпирические нейронные сети, физически-информированные нейронные сети, идентификация, коэффициентная обратная задача.

1. Введение

При моделировании физических и технологических процессов значимую роль играют методы определения параметров модели. Подобные задачи относятся к коэффициентным обратным задачам и активно исследуются с 50-х годов прошлого века. По своей природе обратные задачи существенно отличаются от задач анализа и имеют как свою внутреннюю специфику, например, неединственность решения, так и разнообразие методов исследования, что обуславливает сложность решения этих задач в общем виде. В данной работе для решения коэффициентных обратных задач для обыкновенных дифференциальных уравнений, описывающих поведение динамических систем, предложено использовать один из вариантов физически-информированных нейронных сетей, называемых полуэмпирическими рекуррентными нейронными сетями. Подобные нейросетевые архитектуры тесно связаны с традиционными конечно-разностными схемами дискретизации систем дифференциальных уравнений и с начала 90-х годов прошлого века активно исследуются в работах G. Dreyfus'a [1], Д. А. Тархова и А. Н. Васильева [2], Ю. В. Тюменцева [3]. Предложен новый подход к построению полуэмпирических рекуррентных нейронных сетей на основе неявных методов семейства Рунге-Кутты.

2. Полуэмпирические рекуррентные нейронные сетей и диагонально-неявные методы семейства Рунге-Кутты

Рассмотрим задачу Коши для системы обыкновенных дифференциальных уравнений, записанных относительно неизвестной вектор-функции $\mathbf{y}(x) = (y_1(x), \dots, y_n(x))^T$ на отрезке $x \in [x_0, X]$:

$$\frac{d\mathbf{y}(x)}{dx} = \mathbf{F}(x, \mathbf{y}(x), \bar{\alpha}) \quad (1)$$

с начальными условиями

$$\mathbf{y}(x_0) = \mathbf{y}_0. \quad (2)$$

Здесь $\bar{\alpha} = (\alpha_1, \dots, \alpha_p)^\top$ – вектор параметров задачи, $\mathbf{F}(x, \mathbf{y}(x), \bar{\alpha})$ – заданная вектор-функция, зависящая от $n + p + 1$ аргумента, $\mathbf{y}_0$ – вектор значений функции $\mathbf{y}(x)$ в начальной точке $x = x_0$.

При заданном векторе параметров $\bar{\alpha}$ для численного решения задачи (1)-(2) проведем разбиение отрезка $[x_0, X]$ узловыми точками x_i , $i = \overline{1, N-1}$:

$$x_0 < x_1 < \dots < x_{N-1} < x_N = X. \quad (3)$$

Приближенное значение вектор функции $\mathbf{y}(x)$ в узловой точке x_k может быть найдено по следующей схеме [4]:

$$\begin{aligned} \mathbf{y}_k &= \mathbf{y}_{k-1} + h \sum_{i=1}^s b_i \mathbf{k}_i, \quad k = \overline{1, N}, \\ \mathbf{k}_i &= \mathbf{F} \left(x_{k-1} + c_i h, \mathbf{y}_{k-1} + h \sum_{j=1}^s a_{ij} \mathbf{k}_j \right). \end{aligned} \quad (4)$$

Класс методов вида (4) называются s -стадийными методами Рунге-Кутты. Векторы $\mathbf{b}$, $\mathbf{c}$ и матрица коэффициентов A задаются матрицей Бутчера

$$\begin{array}{c|cccc} c_1 & a_{11} & a_{12} & \dots & a_{1s} \\ c_2 & a_{21} & a_{22} & \dots & a_{2s} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ c_s & a_{s1} & a_{s2} & \dots & a_{ss} \\ \hline & b_1 & b_2 & \dots & b_s \end{array} \quad (5)$$

В случае $a_{ij} = 0$ при $i \leq j$ методы (4) являются явными. Если $a_{ij} = 0$ при $i > j$ и хотя бы один диагональный элемент $a_{ii} \neq 0$, имеем диагонально неявный метод Рунге-Кутты. Во всех остальных случаях методы Рунге-Кутты являются неявными.

Остановимся подробнее на скалярном варианте задачи (1)-(2), записанной относительно скалярной функции $y(x)$ на отрезке $x \in [x_0, X]$:

$$\frac{dy}{dx} = f(x, y, \bar{\alpha}), \quad y(x_0) = y_0. \quad (6)$$

Рассмотрим более общую постановку для задачи (6). Будем полагать, что вектор параметров $\bar{\alpha}$ либо не задан, либо определен со значительной погрешностью. Поэтому для задачи (6) требуется найти не только решение, но и вектор $\bar{\alpha}$ (определить модель рассматриваемого процесса) таким образом, чтобы в точках $x = \xi_i$, $i = \overline{1, M}$, искомая функция принимала значения $y(\xi_i) = y_i^e$.

Для решения этой задачи воспользуемся вариантом полуэмпирической рекуррентной архитектуры, предложенной А. Н. Васильевым и Д. А. Тарховым [2]. Согласно ему заменим конечный рассматриваемый интервал $x \in [x_0, X]$ на интервал

переменной длины $[x_0, x]$ с переменной правой границей. Тогда при использовании какой-либо численной схемы

$$y_k = \mathcal{A}(y_k, y_{k-1}, f, h, N) \quad (7)$$

с выбранной функцией $\mathcal{A}(\cdot)$, каждое узловое значение y_k будет являться функцией аргумента x , т.е. $y_k = y_k(x)$, $k = \overline{1, N}$. Приближенное решение задачи (6) будет равно значению в последнем узле:

$$\hat{y}(x) = y_N(x). \quad (8)$$

Стоит отметить, что $y_N(x)$ является выходом рекуррентной нейронной сети эквивалентной используемой численной схеме (7). Само решение $\hat{y}(x)$ содержит в себе не только информацию о всех пройденных состояниях, но и информацию о заданной функции $f(x, y, \bar{\alpha})$, начальных условиях и значениях параметров $\bar{\alpha}$.

Первым шагом к построению приближенного решения является создание нейросетевой архитектуры, которая отражала бы специфику текущей задачи. Для примера воспользуемся одностадийными методами семейства Рунге-Кутты, объединенными названием θ -метод [4]:

$$y_k = y_{k-1} + hf(x_{k-1} + \theta h, y_{k-1} + \theta(y_k - y_{k-1}), \bar{\alpha}), \quad k = \overline{1, N}. \quad (9)$$

При фиксированном значении θ рекуррентная формула (9) задает один из численных методов решения задачи Коши, например, явный ($\theta = 0$) и неявный ($\theta = 1$) методы Эйлера, неявный метод трапеций ($\theta = 1/2$).

При $k = N$ (с учетом равенства (8)) можно отметить, что выражение

$$\hat{y} = y_{N-1} + hf(x_{N-1} + \theta h, y_{N-1} + \theta(\hat{y} - y_{N-1}), \bar{\alpha})$$

невозможно в произвольном случае задания функции f разрешить относительно $\hat{y}$ при $\theta \neq 1$. При $\theta = 1$ имеем

$$\begin{aligned} \hat{y} - hf(x, \hat{y}, \bar{\alpha}) &= L_1(x, \hat{y}), \quad y_0 - L_{N-1}(x, \hat{y}) + hf(x - (N-1)h, L_{N-1}(x, \hat{y}), \bar{\alpha}) = 0, \\ L_{N-k}(x, \hat{y}) - hf(x - (N-k)h, L_{N-k}(x, \hat{y}), \bar{\alpha}) &= L_{N-k+1}(x, \hat{y}). \end{aligned} \quad (10)$$

Подставляя y_0 из начального условия и задавая нейросетевое разложение функции $\hat{y}(x)$:

$$\hat{y}(x, W_1, \dots, W_Q) = \varphi_Q(\varphi_{Q-1}(\dots \varphi_1(x, W_1), W_{Q-1}), W_Q), \quad (11)$$

поставим задачу минимизации функционала ошибки вида

$$\begin{aligned} \mathcal{E}(W_1, \dots, W_Q, \bar{\alpha}) &= \delta_1 \left(\hat{y} \Big|_{x=x_0} - y_0 \right)^2 + \delta_2 \sum_{i=1}^M \left(\hat{y} \Big|_{x=\xi_i} - y_i^e \right)^2 + \\ &+ \delta_3 \sum_{j=1}^P \left(y_0 - L_{N-1}(x, \hat{y}) + hf(x - (N-1)h, L_{N-1}(x, \hat{y}), \bar{\alpha}) \right)^2 \Big|_{x=\eta_j} \rightarrow \min_{W_1, \dots, W_Q, \bar{\alpha}} \end{aligned} \quad (12)$$

для поиска нейросетевых коэффициентов $W_1, \dots, W_Q$ и вектора параметров $\bar{\alpha}$. Здесь η_j – пробные точки, случайным образом распределенные на отрезке $[x_0, X]$, Q – количество слоев в нейросетевом разложении (11), $W_1, \dots, W_Q$ – матрицы нейросетевых коэффициентов каждого слоя разложения (11).

3. Численный пример

В качестве примера рассмотрим задачу идентификации параметра α для линейной начальной задачи

$$\frac{dy}{dx} = -\alpha (y - \cos x), \quad y(0) = 0, \quad x \in [0, 1], \quad (13)$$

которая при $\alpha = 50$ дается в качестве тестовой в монографии [5].

Используя формулы (10), получим предобученную полуэмпирическую рекуррентную сеть

$$\begin{aligned} \hat{y} + \alpha h (\hat{y} - \cos x) &= L_1(x, \hat{y}), \quad L_{N-1}(x, \hat{y}) + \alpha h [L_{N-1}(x, \hat{y}) - \cos(x - (N-1)h)] = 0, \\ L_{N-k}(x, \hat{y}) + \alpha h [L_{N-k}(x, \hat{y}) - \cos(x - (N-k)h)] &= L_{N-k+1}(x, \hat{y}). \end{aligned} \quad (14)$$

Нейросетевое разложение (11) выберем двухслойным ($Q = 2$) с 9 нейронами на каждом слое, задаваемыми матрицами $W^{(1)}$ и $W^{(2)}$ размерности 9×3 :

$$\begin{aligned} \hat{y}(x, W^{(1)}, W^{(2)}) &= \sum_{i=1}^9 W_{i1}^{(1)} \tanh(W_{i2}^{(1)} \cdot u(x, W^{(2)}) + W_{i3}^{(1)}), \\ u(x, W^{(2)}) &= \sum_{i=1}^9 W_{i1}^{(2)} \tanh(W_{i2}^{(2)} \cdot x + W_{i3}^{(2)}). \end{aligned} \quad (15)$$

Минимизируем функционал вида (12) при $N = 7$, $M = P = 200$, $\delta_1 = 5025$, $\delta_2 = 100500$ и $\delta_3 = 100000$ с η_j , $j = \overline{1, P}$ – величинами равномерно распределенными на отрезке $[0, 1]$, и ξ_i , $i = \overline{1, M}$ – значениями аналитического решения задачи (13) при $\alpha = 30$ и наложенным стандартным гауссовским шумом со среднеквадратичным отклонением $\sigma = 0.01$. В результате определили константу $\alpha_* = 30.51628$ и весовые матрицы нейросетевого разложения $W_*^{(1)}$ и $W_*^{(2)}$. Графики аналитического решения при $\alpha = 30$, α_* и нейросетевого решения даны на рис. 1.

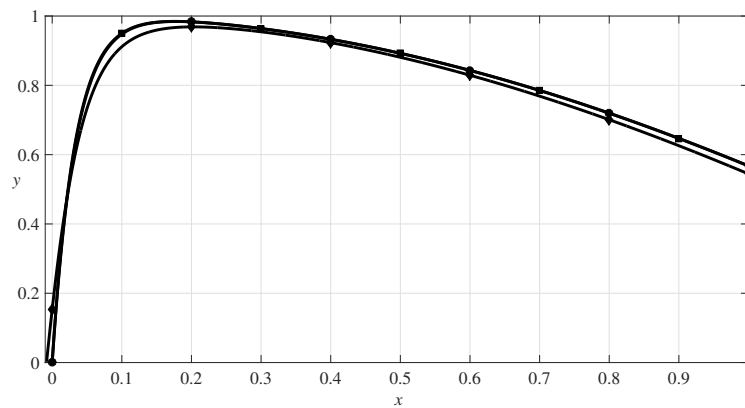


Рис. 1. Решение задачи (13): кружки – аналитическое решение при $\alpha = 30$, квадраты – аналитическое решение при $\alpha_* = 30.51628$, ромбы – нейросетевое решение (15)

Литература

1. Dreyfus G. Neural Networks: Methodology and Applications. Berlin: Springer, 2005. 515 p.

2. Vasilyev A.N., Tarkhov D.A. Semi-Empirical Neural Network Modeling and Digital Twins Development. US: Elsevier, Academic Press, 2019. 288 p.
3. Tiumentsev Yu.V., Egorchev M.V. Neural Network Modeling and Identification of Dynamical Systems. US: Elsevier, Academic Press, 2019. 319 p.
4. Хайпер Э., Нёрсетт С., Ваннер Г. Решение обыкновенных дифференциальных уравнений. Нежесткие задачи. Москва: Мир, 1990. 512 с.
5. Хайпер Э., Ваннер Г. Решение обыкновенных дифференциальных уравнений. Жесткие и дифференциально-алгебраические задачи. Москва: Мир, 1999. 685 с.

MSC 68T07

Physics Informed Neural Networks in Solving Coefficient Inverse Problems

E.B. Kuznetsov,¹ S.S. Leonov^{1,2}

Moscow Aviation Institute (National Research University)¹
Peoples' Friendship University of Russia named after Patrice Lumumba (RUDN
University)²

Abstract: The paper considers the process of constructing a semi-empirical recurrent neural network belonging to the class of physics informed neural networks. The authors propose an approach to building a semi-empirical network architecture using implicit Runge-Kutta family schemes, which allows to obtain an analytical functional expression equivalent to a pre-trained recurrent neural network. The proposed approach was tested on a linear, weakly rigid initial problem with one scalar parameter.

Keywords: semi-empirical recurrent neural networks, physics informed neural networks, identification, coefficient inverse problem, ordinary differential equations, initial problem.