Тархов Д.А., Лазовская Т.В. Об использовании методов нейронных сетей для одного жесткого уравнения первого порядка. // Проблемы информатики в образовании, управлении, экономике и технике: Сб. статей XIV Междунар. научно-техн. конф. – Пенза: ПДЗ, 2014. – С. 171-175.

УДК 004.021

ОБ ИСПОЛЬЗОВАНИИ МЕТОДОВ НЕЙРОННЫХ СЕТЕЙ ДЛЯ ОДНОГО ЖЕСТКОГО УРАВНЕНИЯ ПЕРВОГО ПОРЯДКА

Д.А. Тархов, Т.В. Лазовская

USE OF THE METHODS OF NEURAL NETWORKS FOR SOLVING OF SOME STIFF DIFFERENTIAL FIRST-ORDER EQUATION

D.A. Tarkhov, T.V. Lazovskaya

Аннотация. Рассматривается проблема решения жесткого дифференциального уравнения первого порядка. Для обобщенной параметризованной задачи применяется аппарат нейронных сетей. Анализируются результаты численных экспериментов для базисных функций двух типов и разного числа нейронов в сети. Исследуется возможность и эффективность использования дополнительных данных.

Ключевые слова: нейронные сети, дифференциальные уравнения, жесткие задачи, гибридные методы, параметризованные задачи.

Abstract. The problem of solving the stiff differential first-order equation is considered. Neural Networks are used to model a solution of a parameterized equation. The results of numerical experiments for two types of the basic functions and different counts of neurons are compared and analyzed. The possibility and the productivity of use of extra data are investigated.

Keywords: neural network, differential equations, stiff problems, hybrid methods, parameterized problems.

Рассмотрим класс обыкновенных дифференциальных уравнений, для решения которых не применимы явные вычислительные методы, известный как жесткие дифференциальные уравнения. Классическим примером такой задачи является следующее уравнение первого порядка:

$$y' = -50(y - \cos(x)),$$

приводимое, в частности, в книге [1]. При решении этой задачи явным методом Эйлера с начальным условием $y(0) = 0$ возникает критическое значение для шага сетки, равное 2/50, при превышении которого приближенное решение становится неустойчивым с сильными колебаниями. При этом, для меньшего шага ошибка оказывается слишком велика.

Для решения жестких задач обычно используются неявные и полуявные вычислительные методы, результатом применения которых является точечная аппроксимация решения. Мы рассмотрим данное уравнение в обобщенном виде

$$y' = -\alpha(y - \cos(x)), \qquad y(0) = 0.$$

Применим нейросетевые методы для построения его приближенного параметризованного решения в виде некоторой функции (а не в отдельных точках как при решении методом Эйлера и иными классическими численными методами). Обобщение задачи позволяет искать ее решение сразу для всех значений параметра α, который будем считать принадлежащим интервалу $[\alpha_{min}; \alpha_{max}]$, где $\alpha_{max} = 50$. Поиск решения по x производится на интервале $[0; 1]$.

Ищем приближенное решение этой задачи в виде выхода искусственной нейронной сети заданной архитектуры $y(x) = \sum_{i=1}^{n} c_i v(x, \alpha, \boldsymbol{a}_i)$, веса которой определяются в процессе минимизации функционала ошибки

$$\sum_{j=1}^{m} \left(y'(\xi_j) - F\left(\xi_j, y(\xi_j), \alpha_j\right)\right)^2 + \delta\, y^2(0),$$

где $F(x, y, \alpha) = -\alpha(y - \cos(x))$.

При выборе базисных функций v рассматривались два варианта:

- сигмоиды $th[a(x - d)]th[a_1(x - d_1)]$ и

- асимметричные гауссианы $x\, exp[-a(x - d)^2]\, exp[-a_1(x - d_1)^2]$, заведомо удовлетворяющие начальному условию $y(0) = 0$.

Случайным образом на $[0; 1] \times [\alpha_{min}; \alpha_{max}]$ выбираются m первых тестовых точек (ξ_j, α_j), которые будут перегенерироваться в процессе обучения нейронной сети. Минимизация функционала ошибки происходит с помощью алгоритма, сочетающего в себе методы облака из трех частиц и RProb.

Для исследования применения методов нейронных сетей при решении поставленной выше задачи была проведена серия численных экспериментов для каждого варианта базисных функций.

Кроме простого нейросетевого подхода, рассматривался гибридный, когда построенное нейросетевое приближение $u(x, \alpha)$ использовалось в неявном методе Адамса второго порядка решения ДУ $y' = f(x, y)$. На каждом шаге решение уравнения $y_{i+1} = y_i + \frac{h}{2}(f(x_i, y_i) + f(x_{i+1}, y_{i+1}))$ заменялось формулой $y_{i+1} = y_i + \frac{h}{2}\left(f(x_i, y_i) + f(x_{i+1}, u(x_{i+1}, \alpha))\right)$.

Оба метода были применены при различном числе n нейронов в сетях ($n = 5, 20, 50$).

Также был исследован вопрос появления дополнительных данных. Чтобы не выходить за рамки задачи, в качестве таких данных использовались приближенные решения, полученные классическим методом Эйлера. С целью отразить влияние количества дополнительных данных на результат в одних экспериментах учитывалось только приближенное решение при $\alpha = 50$, а в других – при обоих значениях $\alpha = 5, \alpha = 50$.

	Без доп. данных		Данные, $\alpha = 50$		Данные, $\alpha = 5, 50$	
	сигмоиды	гауссианы	сигмоиды	гауссианы	сигмоиды	гауссианы
$n = 5$	4,078	1,503	2,176	3,746	1,561	3,376
$n = 20$	4,312	0,932	2,781	1,226	1,673	2,074
$n = 50$	8,811	1,787	4,482	1,587	1,260	1,556

Таблица содержит результаты экспериментов нейросетевого моделирования решения задачи для функционала ошибки. Все эксперименты проводились при числе итераций $m = 200$.

Очевидно, что при отсутствии дополнительных данных с лучшей стороны себя показал вариант с базисными функциями, удовлетворяющими начальному условию (гауссианы). Привлечение же дополнительной информации повышает точность только при использовании универсальных базисных функций (сигмоиды), тогда как в случае сети с подобранными под начальное условие функциями в основном происходит, наоборот, увеличение ошибки. Особенно заметен эффект от наличия данных для сетей с высоким количеством нейронов ($n = 50$).

Что касается большой ошибки в графе «без данных» в случае сигмоидов, ее можно объяснить особенностями соответствующего функционала, чувствительного к резкому росту решения. Увеличение числа нейронов при этом не меняет ситуации, так же, как и использование описанного выше гибридного метода. Стоит отметить, что есть возможность добиться хороших результатов при увеличении числа итераций, но это требует значительного роста времени вычислений.

В случае универсальной нейронной сети привлечение данных для $\alpha = 50$ не имеет особенного эффекта, кроме некоторого логичного улучшения при α, близких к 50, независимо от числа нейронов. Но при использовании данных для $\alpha = 5, \alpha = 50$ одновременно отмечается заметное повышение точности и существенное улучшение результатов при применении гибридного метода.

Как видно из таблицы, хороший эффект дает использование специальных базисных функций в сочетании с правильным числом нейронов в сети. Причем именно в этом случае гибридный метод позволяет получить значительное увеличение точности.

Уже отмеченное увеличение ошибки с привлечением дополнительных данных при малом числе нейронов в случае гауссиан, но ее уменьшение при $n = 50$ позволяют сделать вывод о недостатке числа нейронов для обработки всей информации при $n = 5, 20$.

Таким образом, нейронные сети позволяют решать жёсткие задачи без принципиального изменения алгоритма обучения. При этом важным преимуществом по сравнению с классическими методами является функциональный, а не дискретный характер решения. Широкие возможности открывает использование параметризованных сетей, позволяющих получить приближённое решение на всём рассматриваемом интервале изменения параметра.

Уточнение нейросетевого решения возможно при использовании в качестве дополнительной информации приближенных численных решений, полученных классическими методами, даже такими слабыми, как метод Эйлера. Использование гибридного алгоритма, основанного на неявном методе Адамса, эффективно для нейронных сетей, дающих достаточно точные приближения.

Описанный выше подход к решению жесткого дифференциального уравнения можно легко распространить на уравнения высшего порядка. То же самое касается и перехода от начальной задачи к краевым или решения уравнений, заданных в частных производных, и пр.

Библиографический список

1. Хайрер Э., Ваннер Г. Решение обыкновенных дифференциальных уравнений. Жёсткие и дифференциально-алгебраические задачи. – М.: Мир, 1999.– 685 с.

2. Тархов Д.А. Нейросетевые модели и алгоритмы. – М.: Радиотехника, 2014. – 348 с.

3. Горбаченко В.И., Артюхина Е.В. Обучение радиально-базисных нейронных сетей при решении дифференциальных уравнений в частных производных // Нейрокомпьютеры: разработка, применение. – 2007. – № 9. – С. 150–159.

4. Васильев А.Н., Тархов Д.А. Построение приближенных нейросетевых моделей по разнородным данным // Математическое моделирование. – 2007. – Т. 19. – №12. – С. 43–51.

Тархов Дмитрий Альбертович
Санкт-Петербургский государственный политехнический университет,
г. Санкт-Петербург, Россия
E-mail: dtarkhov@gmail.com

Tarkhov Dmitry Albertovich
Saint-Petersburg Politechnical University, Saint-Petersburg, Russia

Лазовская Татьяна Валерьевна
Санкт-Петербургский государственный политехнический университет,
г. Санкт-Петербург, Россия
E-mail: tatianala@list.ru

Lazovskaia Tatiana Valerievna
Saint-Petersburg Politechnical University, Saint-Petersburg, Russia