

DOI: 10.5862/JPM.218.14

УДК: 004.021

Т.В. Лазовская, Д.А. Тархов

Санкт-Петербургский политехнический университет Петра Великого

НОВЫЕ ПОДХОДЫ К ПОСТРОЕНИЮ ПАРАМЕТРИЗОВАННОГО НЕЙРОСЕТЕВОГО РЕШЕНИЯ ЖЕСТКОГО ДИФФЕРЕНЦИАЛЬНОГО УРАВНЕНИЯ

На примере одной модельной задачи исследуются новые алгоритмы нейросетевого моделирования, такие как специальная регенерация тестовых точек, использование дополнительной информации и гибридный метод. Приведенные в статье подходы можно естественно обобщить на системы обыкновенных дифференциальных уравнений, дифференциальные уравнения более высокого порядка и дифференциальные уравнения в частных производных.

НЕЙРОННАЯ СЕТЬ, ДИФФЕРЕНЦИАЛЬНОЕ УРАВНЕНИЕ, ПАРАМЕТРИЗОВАННАЯ МАТЕМАТИЧЕСКАЯ МОДЕЛЬ, ГИБРИДНЫЙ АЛГОРИТМ.

Введение

Методология построения нейросетевых моделей по дифференциальным уравнениям и иным данным (краевые условия, измерения и т. п.), разработанная профессорами Санкт-Петербургского политехнического университета Петра Великого А.Н. Васильевым и Д.А. Тарховым [3], обеспечивает решение сложных и некорректных задач математической физики [4 – 7]. Особенно перспективными представляются параметризованные нейросетевые модели, включающие один или несколько параметров задачи в качестве входных переменных [6 – 8] и позволяющие решать одновременно семейство задач, объединенных параметрами.

В данной работе на примере одной не-сложной модельной задачи поставлен и решен ряд новых принципиальных вопросов.

Во-первых, изучена возможность расширения интервала изменения параметра в рамках одной нейросетевой модели без потери точности, то есть увеличения множества одновременно решаемых задач.

Во-вторых, исследовано влияние на точность решения нового подхода к выбору тестовых точек, названного специальной регенерацией тестовых точек.

В-третьих, продолжено исследование [3], направленное на уточнение решения

с помощью гетерогенных дополнительных данных. Это точечные данные об искомой функции, в том числе неточные, что часто встречается в реальных моделях.

В данной статье новизна, по сравнению с прежними работами [3], заключается в том, что упомянутые точечные данные получены заведомо неточным численным методом. Кроме того, помимо точечных данных, используются дополнительные условия в виде уравнений, полученных с помощью асимптотического разложения.

Для поиска ответов на перечисленные выше вопросы было целесообразно в первую очередь взять простейшую модельную задачу, которая бы имела аналитическое решение, чтобы с ним сравнивать построенные приближенные решения и затем объективно оценить полученные результаты.

В качестве такой модельной задачи выбрано жесткое дифференциальное уравнение первого порядка [1]. Исследования [2 – 8] позволяют предположить, что выводы из сравнительного анализа изучаемых методов и алгоритмов останутся справедливыми и для более сложных задач, включая задачи математической физики; таким образом, рассмотрение столь простой задачи представляется оправданным.

Введение параметра в дополнительные данные задачи (они выражены с помощью

некоторого уравнения) позволяет, в частности, провести сравнение классической и неклассической постановок. В последнем случае условия накладываются на искомую функцию вне выбранной при решении области. В качестве отправной точки для такого условия используется естественная асимптотика изучаемой задачи. Роль неточных дополнительных точечных данных играет приближенное решение задачи, полученное одним из классических методов.

Нейронная сеть позволяет объединить информацию как в виде данных, так и уравнений с помощью функционала минимизации, характеризующего качество модели. Кроме того, в данной статье исследуется новая схема выбора пробных точек на разных этапах минимизации (процедура регенерация тестовых точек) при различных типах исходных условий.

Нейросетевые модели с дополнительными данными

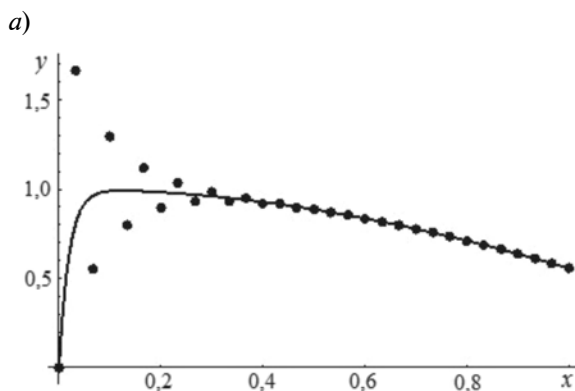
Особый интерес представляют задачи, которые обычно плохо решаются классическими явными методами или требуют большого количества итераций. Среди обыкновенных дифференциальных уравнений (ДУ) такими являются жесткие ДУ [1].

В книге [1] рассматривается классический пример жесткого уравнения

$$y' = -50(y - \cos x) \quad (1)$$

с начальным условием $y(0) = 0$.

При решении этой задачи явным мето-



дом Эйлера возникает критическое значение для шага сетки, равное $2/50$, при превышении которого приближенное решение становится неустойчивым с сильными колебаниями (рис. 1, а). При этом ошибка оказывается слишком большой для меньшего шага.

Мы будем рассматривать обобщенную параметризованную задачу

$$y' = -\alpha(y - \cos x), \quad y(0) = 0, \quad (2)$$

где $\alpha \in [5, 50]$ или $\alpha \in [0, 5; 50]$, $x \in [0, 1]$.

Жесткость задачи проявляется для переменной x в окрестности нуля, чем и обусловлен выбор соответствующего интервала. Пробные запуски показали, что качество нейросетевого решения сохраняется и для больших промежутков. Задача решается сразу для всех рассматриваемых значений параметра α . Заметим, что такие интервалы изменения параметра существенно больше рассмотренных в работах [6, 8].

Приближенное решение ищется в виде выхода искусственной нейронной сети заданной архитектуры

$$y(x) = \sum_{i=1}^n c_i v_i(x, \alpha, \mathbf{a}_i),$$

веса которой $\{c_i, \mathbf{a}_i\}_{i=1}^n$ определяются в процессе минимизации функционала ошибки

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

причем в нашем случае $F(x, y, \alpha) = -\alpha(y - \cos x)$.

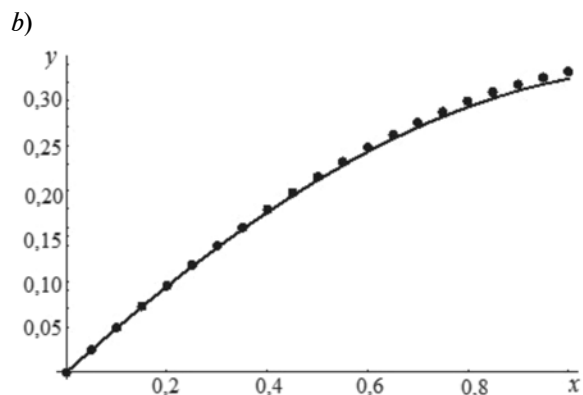


Рис. 1. Решения жесткого дифференциального уравнения (2) явным методом Эйлера при $\alpha = 50$ (а) и 5 (б). Символы соответствуют поточечным решениям методом Эйлера, линии – истинным решениям

Пробные точки (ξ_j, α_j) выбираются случайными, распределенными равномерно в рассматриваемой области изменения переменной x и параметра α , их выбор повторяется после нескольких (трех-пяти) итераций алгоритма оптимизации. Новый случайный выбор тестовых точек на некотором шаге будем называть регенерацией тестовых точек.

Оценка качества полученного решения производится по точному аналитическому решению уравнения (2) с начальным условием $y(0) = 0$, имеющему вид

$$y(x, \alpha) = \frac{\alpha^2 (\cos x - \exp(-\alpha x)) + \alpha \sin x}{\alpha^2 + 1}. \quad (3)$$

В настоящей работе исследованы два типа моделей, соответствующих различным базисным функциям, при разном количестве нейронов в сети. В первом случае выбирались универсальные сигмоиды вида

$$\text{th}[a(x - d)]\text{th}[a_1(\alpha - d_1)],$$

во втором – асимметричные гауссианы вида

$$x \exp[-a(x - d)^2] \exp[-a_1(\alpha - d_1)^2],$$

которые заведомо удовлетворяют начальному условию.

Оптимизация функционала ошибки проводилась по алгоритму, сочетающему RProp и метод облака [2]; точки случайным образом регенерировались каждые три шага, облако состояло из трех частиц.

Следует отметить, что сложность процесса оптимизации вызвана тем обстоятельством, что оптимизируемый функционал меняется после каждой регенерации пробных точек; таким образом мы избегаем возможности получения хорошей аппроксимации в фиксированном наборе точек и плохой в других точках рассматриваемой области, что может произойти при применении метода коллокаций.

Необходимо подчеркнуть, что применяемая вычислительная процедура содержит двойную стохастичность: помимо указанной случайной регенерации точек, случайными выбираются начальные веса нейронной сети.

Кроме того, были изучены алгоритмы

построения модели, использующие дополнительные данные об искомом решении, оценен эффект такого уточнения для разных типов базисных функций и числа нейронов в сети. В качестве таких данных рассматривалось соответствие искомым решениям уже найденным с помощью явного метода Эйлера для значений параметра α , равных 5 и 50. Отметим, что при $\alpha = 5$ уравнение перестает быть жестким и решается довольно точно (см. рис. 1, *b*). «Плохое» решение при $\alpha = 50$ позволяет изучить реакцию модели на неточные данные.

Введение в модель новой информации происходит с помощью добавления в функционал минимизации дополнительного слагаемого вида

$$\delta_1 \sum_{j=1}^m (f(x_j) - y(x_j))^2,$$

где $f(x_j)$ – поточечное решение Эйлера; вес δ_1 можно варьировать, при этом следует учитывать точность имеющихся данных или какие-либо специальные условия.

В первой модификации модели использовались описанные выше данные для $\alpha = 50$, во второй учитывались также данные для $\alpha = 5$. Некоторые результаты вычислительных экспериментов для функционала ошибки для двух типов нейронных сетей приведены в табл. 1. Рассматривались сети с разным числом нейронов и числом итераций, равным 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 применения данных для сетей с высоким количеством нейронов ($n = 50$).

Что касается большой ошибки в графе $\delta_1 = 0$ (означает отсутствие дополнительных данных) в случае сети персептрон, ее можно объяснить особенностями соответ-

Таблица 1

Значения функционала ошибки с различным набором данных для двух типов нейронных сетей

n	Базовая модель ($\delta_1 = 0$)		1-я модификация ($\alpha = 50$)		2-я модификация ($\alpha = 5; \alpha = 50$)	
	Сигмоиды	Гауссианы	Сигмоиды	Гауссианы	Сигмоиды	Гауссианы
5	4,078	1,503	2,176	3,746	1,561	3,376
20	4,312	0,932	2,781	1,226	1,673	2,074
50	8,811	1,787	4,482	1,587	1,260	1,556

Обозначения: n – количество нейронов, α – параметр; δ_1 – вес дополнительных данных; при δ_1 , отличном от нуля, использованы дополнительные данные при различных значениях α . Число итераций – 200.

ствующего функционала, чувствительно к резкому росту решения. Увеличение числа нейронов при этом не меняет ситуации. При отсутствии данных или их малом количестве большое число нейронов лишь замедляет обучение и ухудшает результат. Важно отметить, что результаты улучшаются при существенном увеличении числа итераций.

Таким образом, уточнение нейросетевого решения возможно при использовании в качестве дополнительной информации даже неточных данных, например приближенных численных решений, полученных классическими методами, в том числе такими слабыми, как явный метод Эйлера.

Модификации нейросетевой модели с помощью специальной регенерации тестовых точек

Продолжим уточнение модели с дополнительными данными, используя новую процедуру регенерации тестовых точек – выбор тестовых точек по определенному правилу на каждой итерации. Введем параметр d_t , принимающий значения 0; 0,3; 0,5; 0,7; 1,0 (вообще говоря, любые в промежутке от 0 до 1), отражающий долю точек, фиксирующихся от одной итерации к другой. Например, $d_t = 0$ означает полную регенерацию, т. е. все точки заново выбираются случайным образом (распределены равномерно по рассматриваемому интервалу) перед каждой итерацией, $d_t = 1$ означает, что точки фиксируются с первой итерации и больше не изменяются. Для промежуточных значений параметра

используется следующее правило: фиксируется $d_t \cdot m$ точек из всех m тестовых точек с наибольшими значениями функционала ошибки, а остальные регенерируются случайным образом. При первой итерации во всех случаях точки выбираются так, чтобы они были случайно и равномерно распределены на рассматриваемой области.

В экспериментах использовалась сеть типа персептрон, показавшая себя в предыдущем исследовании как наиболее восприимчивая к появлению дополнительной информации о модели. Число нейронов в сети выбиралось равным двадцати ($n = 20$), использовались данные о соответствии искомого решения приближенному, полученному явным методом Эйлера при $\alpha = 5$ и 50, количество итераций составляло 300.

Для объективной оценки полученных результатов введем следующую меру. Поскольку для нашего уравнения существует решение в явном виде (3), мы имеем возможность сравнить решение, построенное с помощью нейронных сетей, с истинным. В качестве такой меры используется среднеквадратичное отклонение s , найденное в 100 тыс. точек (a, x) , причем a и x равномерно распределены на соответствующих промежутках. Выбранное число точек позволяет говорить об устойчивости оценки относительно разных выборок.

Была проведена серия испытаний для различных значений параметра d_t . Качество построенных нейронной сетью решений определялось с помощью упомянутой выше среднеквадратичной оценки s . В табл. 2 представлены результаты экспериментов в

Таблица 2

Среднеквадратичная оценка качества построенной нейросетевой модели

Параметр регенерации	$d_i = 0$	$d_i = 0,3$	$d_i = 0,5$	$d_i = 0,7$	$d_i = 1,0$
$E_c, 10^{-2}$	6,39	6,45	6,80	6,62	6,85
$s^2, 10^{-2}$	0,96	0,59	0,48	0,92	1,25

Обозначения: E_c, s^2 – среднее и среднеквадратичное отклонения, d_i – введенный параметр, отражающий долю точек, фиксирующихся от одной итерации к другой. Число итераций – 300; число тестовых точек $m = 20$.

виде среднего (E_c) и среднеквадратичного (s^2) отклонений полученной выборки для c . Из данных таблицы видно, что среднее значение ошибки отличается лишь во втором знаке. Полученные результаты позволяют сделать вывод, что в данном случае регенерация по вышеописанному правилу при $d_i = 0,3$ и $d_i = 0,5$ обеспечивает более устойчивый результат нейросетевого моделирования.

Очевидно, что с увеличением размерности, т. е. при переходе к уравнениям в частных производных, данный эффект должен усилиться. Влияние регенерации точек на результат мы продолжим рассматривать и в следующем разделе.

Уточнение нейросетевой модели с помощью асимптотического условия

Следует отметить, что при достаточно малых значениях α справедливо соотношение $y \cong \alpha \sin x$. Полученное таким образом асимптотическое условие будем считать дополнительной информацией о модели. Кроме того, будем учитывать данные о соответствии искомого решения приближенному, полученному методом Эйлера при $\alpha = 50$; другими словами, изучим использование в модели гетерогенных данных. Помимо этого, будем продолжать исследование эффекта регенерации тестовых точек, описанной в предыдущем разделе.

Асимптотическое условие будем учитывать в модели добавлением в функционал минимизации слагаемого

$$\delta_2 \sum_{k=1}^m \left[y \left(x_k, \frac{\alpha_k}{M} \right) - \frac{\alpha_k}{M} \sin x_k \right]^2,$$

где M – достаточно большое фиксирован-

ное положительное число, а переменные x_k и α_k регенерируются так же, как и ранее на рассматриваемых промежутках $\alpha \in [0, 5; 50]$, $x \in [0, 1]$.

Описанный выше подход был исследован для сети с двадцатью базисными функциями ($n = 20$) и с двадцатью и пятьюдесятью тестовыми точками ($m = 20$ и 50). Нейронная сеть применялась с асимптотическим условием для $M = 50, 100$ и 200 . Отметим, что значение $1/M$ в этих случаях будет выходить за рамки интервала, в котором предполагается выполнение задачи (2), и можно говорить о решении неклассической задачи.

Для каждого набора параметров проводилась серия испытаний. Качество построенных нейронной сетью решений определялось с помощью упомянутой выше среднеквадратичной оценки. Результаты экспериментов приведены в табл. 3. Для модели с пятьюдесятью тестовыми точками представлен только случай использования асимптотического условия при $M = 50$, так как при других M значимых различий в результатах не выявлено. Очевидно, что при больших значениях параметра M задача является в большей степени неклассической, что выражается в ухудшении результатов при полной регенерации для значения $m = 20$. В этих случаях лучше работает метод с частичной регенерацией тестовых точек, а метод коллокаций дает наибольшую ошибку. Важно, что именно для модели с фиксированными точками для $m = 50$ мы получаем такой же хороший результат, как и при полной регенерации для $m = 20$. При этом сильно возрастает время обучения (оно линейно зависит от числа тестовых

Таблица 3

Среднеквадратичная оценка качества нейросетевых моделей, построенных с учетом асимптотического условия при различных значениях параметра M и параметра регенерации d_t

M	Регенерация	$d_t = 0$	$d_t = 0,3$	$d_t = 0,5$	$d_t = 0,7$	$d_t = 1,0$
$m = 20$						
50	$E_c, 10^{-2}$	5,51	6,83	8,38	7,36	7,54
	$s^2, 10^{-2}$	0,41	0,99	0,32	0,83	1,24
100	$E_c, 10^{-2}$	7,03	6,84	8,13	6,87	9,12
	$s^2, 10^{-2}$	0,68	1,25	1,70	1,10	1,35
200	$E_c, 10^{-2}$	7,71	7,09	6,96	6,68	7,63
	$s^2, 10^{-2}$	1,49	0,96	1,27	1,01	1,95
$m = 50$						
50	$E_c, 10^{-2}$	6,13	6,05	6,23	6,96	5,62
	$s^2, 10^{-2}$	0,95	0,46	1,18	0,71	0,27

точек). Таким образом, при прочих равных условиях полная регенерация позволяет сократить время работы алгоритма за счет уменьшения количества контрольных точек при сохранении точности результата.

Приведем явный вид полученного решения при $m = 20$ и $M = 50$:

$$\begin{aligned}
 u(x, \alpha) = & 0,095 + 0,103 \operatorname{th}[4,294 \cdot 10^{11} \times \\
 & \times (3,941 \cdot 10^{11} + x)] \operatorname{th}[0,086(-49,788 + \alpha)] + \\
 & + 0,147 \operatorname{th}[6,398(-0,37 + x)] \times \\
 & \times \operatorname{th}[0,112(-45,994 + \alpha)] + \\
 & + 0,235 \operatorname{th}[4,21(-0,76 + x)] \operatorname{th}[0,025 \times \\
 & \times (-45,33 + \alpha)] - 0,787 \operatorname{th}[1,88 \times \\
 & \times (-0,305 + x)] \operatorname{th}[0,046(-34,378 + \alpha)] + \\
 & + 0,083 \operatorname{th}[4,474(-0,643 + x)] \times \\
 & \times \operatorname{th}[0,149(-31,034 + \alpha)] + \\
 & + 0,579 \operatorname{th}[1,523(-0,503 + x)] \operatorname{th}[0,166 \times \\
 & \times (-26,922 + \alpha)] - 0,934 \operatorname{th}[0,083 \times \\
 & \times (-0,054 + x)] \operatorname{th}[0,565(-26,823 + \alpha)] - \\
 & - 0,163 \operatorname{th}[1,818(-0,533 + x)] \operatorname{th}[0,246 \times \\
 & \times (-25,466 + \alpha)] - 0,819 \operatorname{th}[2,017 \times \\
 & \times (-0,88 + x)] \operatorname{th}[0,14(-23,668 + \alpha)] + \\
 & + 0,577 \operatorname{th}[0,938(-0,831 + x)] \operatorname{th}[0,237 \times \\
 & \times (-23,39 + \alpha)] - 0,012 \operatorname{th}[27,835 \times \\
 & \times (-0,284 + x)] \operatorname{th}[0,407(-22,059 + \alpha)] - \\
 & - 0,276 \operatorname{th}[1,099 \cdot 10^{62}(2,145 \cdot 10^{65} + x)] \times \\
 & \times \operatorname{th}[0,091(-11,042 + \alpha)] - \\
 & - 0,317 \operatorname{th}[1,464(-0,569 + x)] \times \\
 & \times \operatorname{th}[0,244(-9,012 + \alpha)] - 0,105 \operatorname{th}[5,303 \times \\
 & \times (-0,489 + x)] \operatorname{th}[0,576(-3,596 + \alpha)] +
 \end{aligned}$$

$$\begin{aligned}
 & + 0,893 \operatorname{th}[1,588(-0,798 + x)] \operatorname{th}[0,105 \times \\
 & \times (-3,546 + \alpha)] + 0,23 \operatorname{th}[1,061 \times \\
 & \times (-0,581 + x)] \operatorname{th}[0,341(-2,658 + \alpha)] + \\
 & + 0,445 \operatorname{th}[62,244(-0,025 + x)] \times \\
 & \times \operatorname{th}[(0,084(-1,548 + \alpha)] - \\
 & - 1,009 \operatorname{th}[1,812(-0,976 + x)] \operatorname{th}[(0,121 \times \\
 & \times (-1,166 + \alpha)] - 0,026 \operatorname{th}[1,222 \times \\
 & \times (-0,375 + x)] \operatorname{th}[(16,58(-0,559 + \alpha)] + \\
 & + 0,849 \operatorname{th}[0,595(-0,056 + x)] \times \\
 & \times \operatorname{th}[2,416(-0,55 + \alpha)].
 \end{aligned}$$

Увеличение ошибки при полной регенерации 50 тестовых точек свидетельствует о некотором переобучении сети, то есть метод и число точек необходимо выбирать корректно относительно условий задачи. Частичную же регенерацию можно успешно применять для неклассической задачи, позволяя тем самым дорабатывать модель на «сложных» точках.

Особый интерес представляет сравнение нейросетевых решений, построенных с использованием дополнительных данных разных типов при конкретных значениях среднеквадратичной ошибки. На рис. 2 изображены нейросетевые приближения для моделей с параметрами $m = 20$, $d_t = 0$ без применения асимптотического условия и с его использованием при $\alpha = 0,5$ и одной и той же среднеквадратичной ошибке. Представленные на рис. 2 результаты

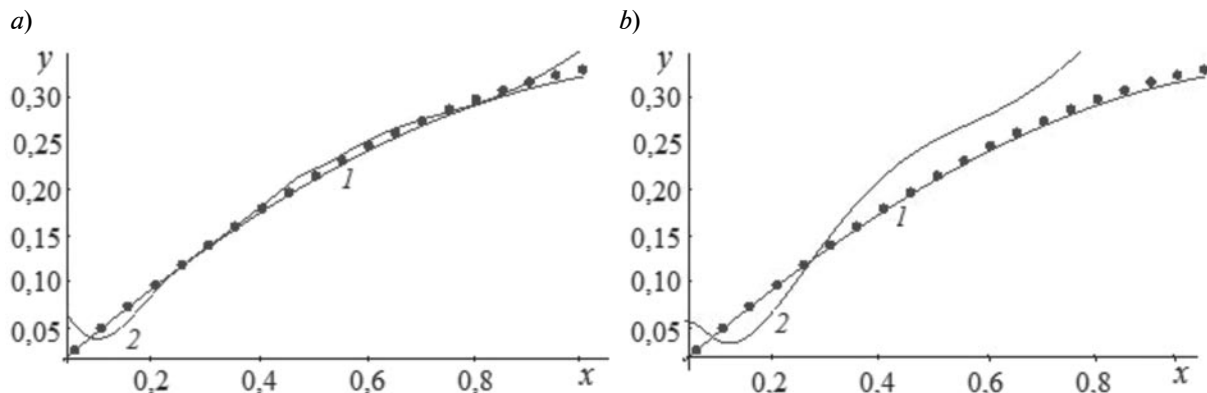


Рис. 2. Нейросетевые приближения решения, полученные при использовании точечных данных (а) и асимптотического условия (б) при $\alpha = 0,5$ и одной и той же среднеквадратичной ошибке 0,027. Параметры моделей $m = 20$, $d_i = 0$. Символы – точечные решения методом Эйлера, линии 1, 2 – истинное и нейросетевое решения соответственно

демонстрируют, что нейросетевая модель, построенная с использованием асимптотического условия, решает нежесткую задачу с меньшей точностью. При этом среднее значение оценки ошибки по всему промежутку изменения параметра у нее меньше (см. табл. 2 и 3), что говорит о более равномерном приближении к истинному решению на всем рассматриваемом интервале.

Отметим, что данный эффект сохраняется и для асимптотической модели при $M = 200$ (неклассическая задача), где приближение при малых значениях α еще грубее для равных значений среднеквадратичной оценки.

Таким образом, нейронная сеть успешно решает неклассические задачи, тогда как асимптотическая модель дает более равномерное приближение на всем интервале изменения параметра.

Способ сочетания нейросетевого и классического методов

Возможен еще один гибридный подход, который состоит в том, что нейросетевое приближение $u(x, \alpha)$ для решения задачи (2) используется при модификации неявного метода Адамса второго порядка точности для решения дифференциального уравнения $y' = f(x, y)$:

$$y_{i+1} = y_i + \frac{h}{2}(f(x_i, y_i) + f(x_{i+1}, y_{i+1})).$$

При классическом подходе переменная y_{i+1} задается неявно и требует на каждом шаге применения некоторого метода решения данного нелинейного уравнения. Чаще всего в этом случае используется метод «предиктор-корректор», состоящий в двухэтапных вычислениях на каждом шаге. В качестве первого этапа применяется метод Эйлера, вычисляется

$$\hat{y}_{i+1} = y_i + hf(x_i, y_i),$$

а в качестве второго – формула метода Адамса:

$$y_{i+1} = y_i + \frac{h}{2}(f(x_i, y_i) + f(x_{i+1}, \hat{y}_{i+1})).$$

Используем нейросетевое приближение $u(x, \alpha)$ для замены двух формул одной:

$$y_{i+1} = y_i + \frac{h}{2}(f(x_i, y_i) + f(x_{i+1}, u(x_{i+1}, \alpha))).$$

Данный подход применялся и анализировался для нейросетевых моделей с сетью перцептрон и базисными функциями-гауссианами для разного числа нейронов, а также при наличии и отсутствии дополнительных данных. Опишем наиболее интересные результаты.

Как и ожидалось, в условиях отсутствия дополнительных данных лучший результат и для рассматриваемого гибридного метода показывает применение нейронной сети с гауссианами. Число нейронов $n = 5$ оказы-

вается недостаточным, а при $n = 50$ возникает эффект переобучения. При $n = 20$ наблюдается значительное уменьшение ошибки при $\alpha = 50$, и таким образом, гибридный метод позволяет улучшить результат метода предиктор-корректор в жестком случае.

Что касается сети персептрон, эффект от применения гибридного метода проявляется в модели с использованием дополнительных данных для значений параметра $\alpha = 5$ и 50 (см. раздел «Нейросетевые модели с дополнительными данными»). Уже для сети с малым количеством нейронов число $n = 5$ дает улучшение результата для жесткого уравнения при $\alpha = 50$, по сравнению с нейросетевым и классическим методами, хотя для малых α этого не наблюдается. По-видимому, сказывается недостаток числа нейронов сети, так как при $n = 20$ и 50 сохраняется лучший результат для жесткого случая и достигается точность классических методов для уравнения при малых значениях параметра α .

Заключение

В результате проведенного исследования установлено, что нейросетевой подход при построении математической модели по разнородным данным, включающим дифференциальные уравнения, позволяет путем модификации функционала ошибки учитывать дополнительные условия различного типа без необходимости существенно

го изменения алгоритма.

Применение нейронных сетей дает возможность решать неклассические задачи, а также задачи построения моделей по неточным данным, что часто встречается в реальных приложениях. В рассмотренной задаче оказалось, что дополнительную информацию лучше воспринимает персептрон с сигмоидными базисными функциями, известными как универсальные аппроксиматоры.

Предложенный в статье алгоритм регенерации тестовых точек обеспечивает экономию времени работы сети в простых задачах и уточнение модели при наличии сложной дополнительной информации.

При использовании нейронных сетей, дающих достаточно точные приближения, эффективен гибридный алгоритм, использующий в классических неявных методах приближенные нейросетевые решения. В этом случае он дает возможность значительно повысить точность приближенного решения в дискретном наборе точек.

Методология, представленная в книге [3], позволяет рассмотреть естественные обобщения приведенного в настоящей статье подхода на системы обыкновенных дифференциальных уравнений, уравнения более высокого порядка и уравнения в частных производных.

Работа выполнена при финансовой поддержке Российского фонда фундаментальных исследований (грант №14-01-00660).

СПИСОК ЛИТЕРАТУРЫ

[1] Хайер Э., Ваннер Г. Решение обыкновенных дифференциальных уравнений. Жесткие и дифференциально-алгебраические задачи. М.: Мир, 1999. 685 с.
 [2] Тархов Д.А. Нейросетевые модели и алгоритмы. М.: Радиотехника, 2014. 348 с.
 [3] Васильев А.Н., Тархов Д.А. Нейросетевое моделирование. Принципы. Алгоритмы. Приложения. СПб.: Изд-во Политехнического университета, 2009. 528 с.
 [4] Васильев А.Н., Тархов Д.А. Построение приближенных нейросетевых моделей по разнородным данным // Математическое моделирование. 2007. Т. 19. № 12. С. 43–51.
 [5] Васильев А.Н., Осипов В.П., Тархов Д.А. Унифицированный процесс моделирования

физико-технических объектов с распределенными параметрами // Научно-технические ведомости СПбГПУ. Физико-математические науки. 2010. № 3(104). С. 39–52.

[6] Васильев А.Н., Тархов Д.А. Параметрические нейросетевые модели классических и неклассических задач для уравнения теплопроводности // Научно-технические ведомости СПбГПУ. Физико-математические науки. 2012. Т. 3. № 153. С. 136–144.

[7] Tarkhov D.A., Vasilyev A.N. Mathematical models of complex systems on the basis of artificial neural networks // Nonlinear Phenomena in Complex Systems. 2014. Vol. 17. No. 3. Pp. 327–335.

[8] Васильев А.Н., Тархов Д.А. Математические модели систем с интервально заданными

параметрами на основе гетерогенных нейронных сетей. Пористый катализатор // Нейрокомпьютеры: разработка, применение. 2014. № 5. С. 3–7.

СВЕДЕНИЯ ОБ АВТОРАХ

ЛАЗОВСКАЯ Татьяна Валерьевна – старший преподаватель кафедры высшей математики Санкт-Петербургского политехнического университета Петра Великого.
195251, Российская Федерация, г. Санкт-Петербург, Политехническая ул., 29
tatianala@list.ru

ТАРХОВ Дмитрий Альбертович – доктор технических наук, профессор кафедры высшей математики Санкт-Петербургского политехнического университета Петра Великого.
195251, Российская Федерация, г. Санкт-Петербург, Политехническая ул., 29
dtarkhov@gmail.com

Lazovskaya T.V., Tarkhov D.A. FRESH APPROACHES TO THE CONSTRUCTION OF PARAMETERIZED NEURAL NETWORK SOLUTIONS OF A STIFF DIFFERENTIAL EQUATION.

A number of new fundamental problems expanding Vasiliev's and Tarkhov's methodology worked out for neural network models constructed on the basis of differential equations and other data has been stated and solved in this paper. The possibility of extending the parameter range in the same neural network model without loss of accuracy was studied. The influence of the new approach to choosing test points and using the heterogeneous complementary data on the solution accuracy was analyzed.

The additional conditions in equation form derived from the asymptotic decomposition were used apart from the point data. The classical and non-classical definitions of the problem were compared by entering a parameter into the additional data. A new sampling scheme of test point choice at different stages of minimization (the procedure of test point regeneration) under various initial conditions was investigated. A way of combining two approaches (classical and neural-network) based on the Adams PECE method was considered.

NEURAL NETWORK, DIFFERENTIAL EQUATION, PARAMETRIZED MATHEMATICAL MODEL, HYBRID ALGORITHM.

REFERENCES

- [1] E. Hairer, G. Wanner, Reshenie obyknovennykh differentsial'nykh uravnenij. Zhestkie i differentsial'no-algebraicheskie zadachi [The solution of ordinary differential equations. Stiff and differential algebraic problems], Mir, Moscow, 1999.
- [2] D.A. Tarkhov, Nejrosetevye modeli i algoritmy [Neural network models and algorithms], Radiotekhnika, Moscow, 2014.
- [3] A.N. Vasiliev, D.A. Tarkhov, Nejrosetevoe modelirovanie. Printsipy. Algoritmy. Prilozheniya [Neural network modeling. Principles. Algorithms. Applications], Polytechnical University Publishing House, St. Petersburg., 2009.
- [4] A.N. Vasiliev, D.A. Tarkhov, Postroenie priblizhennykh nejrosetevykh modelej po raznorodnym dannym [Construction of approximate neural network models according to heterogeneous data], Math. Models & Comp. Simulations. 19 (12) (2007) 43–51.
- [5] A.N. Vasiliev, V.P. Osipov, D.A. Tarkhov, Unified process of modeling of physicotecnical objects with distributed parameters, St. Petersburg State Polytechnical University Journal: Physics and Mathematics. 3 (104) (2010) 39–52.
- [6] A.N. Vasiliev, D.A. Tarkhov, Parametrical neural network models of classical and nonclassical problems for heat conduction equation, St. Petersburg State Polytechnical University Journal: Physics and Mathematics. 3 (153) (2012) 136–144.
- [7] D.A. Tarkhov, A.N. Vasiliev, Mathematical models of complex systems on the basis of artificial neural networks, Nonlinear Phenomena in Complex Systems. 17 (3) (2014) 327–335.
- [8] A.N. Vasiliev, D.A. Tarkhov, Mathematical models of systems with interval representation of parameters on the basis of heterogeneous neural networks. Porous catalyst, Neurocomputers: design and applications. 5 (2014) 3–7.

THE AUTHORS

LAZOVSKAYA Tatiana V.

Peter the Great St. Petersburg Polytechnic University
29 Politekhnikheskaya St., St. Petersburg, 195251, Russian Federation
tatianala@list.ru

TARKHOV Dmitry A.

Peter the Great St. Petersburg Polytechnic University
29 Politechnikheskaya St., St. Petersburg, 195251, Russian Federation
dtarkhov@gmail.com