Б

УДК 519.6 **DOI** 10.24412/2413-7383-2024-3-4-19

В. Н. Беловодский
Федеральное государственное бюджетное образовательное учреждение
высшего образования «Донецкий национальный технический университет»
283001, Донецкая Народная Республика, г. Донецк, ул. Артёма, 58

ОБ ИНТЕРПОЛЯЦИИ И АППРОКСИМАЦИИ ФУНКЦИЙ С ИСПОЛЬЗОВАНИЕМ НЕЙРОННЫХ СЕТЕЙ

V. N. Belovodskiy
Federal State Budgetary Educational Institution of Higher Education
"Donetsk National Technical University"
283001, Donetsk People's Republic, Donetsk, st. Artyoma, 58

ON INTERPOLATION AND APPROXIMATION OF FUNCTIONS USING NEURAL NETWORKS

Постоянно расширяющаяся область эффективного использования нейронных сетей влечет необходимость ознакомления с ними уже на ранних стадиях обучения студентов. И представляется естественным начинать этот процесс при изучении курса методов вычислений с рассмотрения типовых задач теории приближения функций. В данной статье на примере двухслойной нейронной сети изучаются возможности нейронных сетей по интерполяции и аппроксимации функций одной переменной. Полученные результаты сравниваются с классическими алгебраическими подходами. Расчеты проводятся на основе специально разработанной программы, в которой минимизация ошибки сети осуществляется с использованием подпрограммы, реализующей методы Левенберга – Марквардта. Для нахождения глобального минимума ошибки при идентификации параметров сети их начальные значения варьируются и в качестве таковых назначаются различные пробные точки равномерно распределенной в единичном кубе K^n последовательности Соболя. Отмечены достоинства и недостатки нейронного подхода, сделаны обобщения.
Ключевые слова: интерполяция, аппроксимация, нейронная сеть, минимизация, последовательность Соболя

The evolution of neural networks and the ever-expanding field of their application dictate the need to study its theory basics at early stages of student learning. And it is natural to begin this process with consideration of typical problems of the approximation theory of functions in course of computational methods. Such tasks are discussed in this article. Using the example of a two-layer neural network, the possibilities of neural networks for interpolation and approximation of functions of one variable are studied. Computational experiments are being performed with the help of a specially developed program in which the error minimization of the neural network is carried out using a subroutine implementing Levenberg – Marquardt methods. Minimizing errors in the identification of network parameters is carried out by varying their initial values using the Sobol sequence. The results obtained are compared with classical algebraic approaches. The advantages and disadvantages of the neural approach are noted, the generalizations are made.
Key words: interpolation, approximation, neural network, minimization, Sobol sequence

Введение

Расширяющаяся область использования нейронных сетей [1-3] вызывает необходимость начального ознакомления с ними и в разделах теории приближения функций учебных курсов Методов вычислений. На первых порах, представляется достаточным ограничиться обсуждением теоретических основ и базовых принципов формирования нейронных сетей, иллюстрацией интерполяционных возможностей и сравнением их с классическими алгебраическими подходами. Не вызывает сомнения, что такое предварительное ознакомление может оказаться полезным, как для студентов, предметно изучающих нейронные сети на старших курсах обучения, так и специалистов различного профиля, впервые сталкивающихся с нейронными сетями. В данной статье и предпринимается такая попытка.

Следуя классической терминологии [4], под интерполяцией ниже будем понимать формирование функции, принимающей заданные значения в заданных узловых точках, а под экстраполяцией – определение приближенного значения рассматриваемой функции посредством интерполирующей вне заданного интервала.

1 Теоретические основы

Для определенности рассмотрим нейронную сеть, представленную на рис. 1.

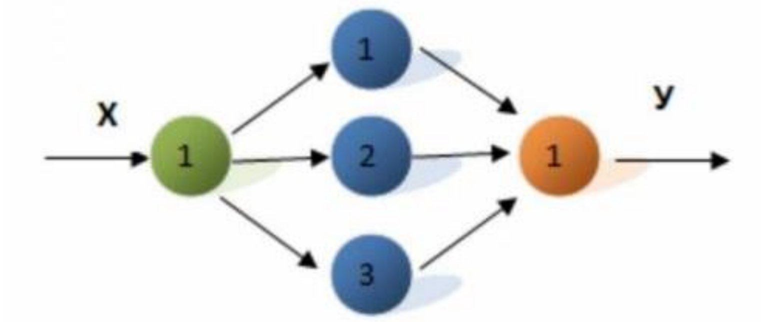


Рисунок 1 – Рассматриваемая нейронная сеть

Она содержит три слоя: первый – входной, состоящий из одного нейрона, второй, внутренний – содержащий три нейрона, и третий – выходной, также состоящий из одного. Такие сети, обычно, называются двухслойными, входной слой в терминологии не фигурирует. Функционирует такая сеть следующим образом. Входной числовой сигнал x после преобразований $П_1$, $П_2$, $П_3$ поступает, соответственно, на нейроны внутреннего слоя $1, 2, 3$, внутренние нейроны выполняют над ними преобразования $П_4$, $П_5$, $П_6$, после чего их результирующая комбинация $П_7$ передается на выходной нейрон, что и является результатом работы сети. Таким образом, значение

$$y=П_7(П_4(П_1), П_5(П_2), П_6(П_3)).$$

В нейронных сетях преобразования $П_1$, $П_2$, $П_3$ являются линейными, с добавлением к ним постоянной составляющей, т.е., по существу, – аффинными. Преобразования, реализуемые на внутренних нейронах, обычно, принимаются одинаковыми, т.е. $П_4 = П_5 = П_6$, они реализуются посредством, так называемой функции активации, а результирующее преобразование $П_7$ также является аффинным.

Математически функционирование сети описывается так:

$$x \to \left\{ \begin{matrix} z_1 = c_1 x + c_0,\ u_1 = f(z_1) \\ z_2 = c_2 x + c_0, u_2 = f(z_2) \\ \dots \qquad \dots \\ z_n = c_n x + c_0, u_n = f(z_n) \end{matrix} \right\} \to y = c_{n+1}u_1 + c_{n+2}u_2 + \dots + c_{2n}u_n + c_{2n+1}$$

- в общем случае, или, в конкретном, для рис. 1 -

$$x \to \begin{cases} z_1 = c_1 x + c_0, \quad u_1 = f(z_1) \\ z_2 = c_2 x + c_0, \quad u_2 = f(z_2) \to y = c_4 u_1 + c_5 u_2 + c_6 u_n + c_7 \\ z_3 = c_3 x + c_0, \quad u_3 = f(z_3) \end{cases} \qquad (1)$$

Здесь, для упрощения модели, в преобразованиях $П_1$, $П_2$, $П_3$ постоянные смещения приняты одинаковыми и функции активации равны, т.е. $f = П_4 = П_5 = П_6$. Таким образом, в свернутом виде, результат работы сети описывается выражением

$$y = c_4 f(c_1 x + c_0) + c_5 f(c_2 x + c_0) + c_6 f(c_3 x + c_0) + c_7. \qquad (1.1)$$

Теперь задача состоит в назначении функции активации, задании исходных данных и последующем определении параметров модели. Прежде, чем перейти к этой части сделаем историческое отступление и коснемся теоретических основ нейронных сетей.

Принято считать, что истоки математического обоснования нейронных сетей начинаются с, так называемой, аппроксимационной теоремы Вейерштрасса (1885 г.) [5], в которой доказана принципиальная возможность представления непрерывной функции одной переменной алгебраическим многочленом со сколь угодно высокой степенью точности, а в работе [6] предложен порядок формирования такого многочлена.

В 1957 г. в результате научной дискуссии с В. Арнольдом при решении 13-й проблемы Гильберта А. Колмогоров опубликовал теорему о возможности представления непрерывной функции нескольких переменных с помощью суперпозиции непрерывных функций одной переменной и операции сложения [7], т.е. в виде

$$f(x_1, x_2, \dots, x_n) = \sum_{q=1}^{2n+1} h_q \left[\sum_{p=1}^{n} \varphi_q^p (x_p) \right]. \qquad (2)$$

Развивая этот результат Дж. Лоренц в 1962 г. показал [8], что число внешних функций может быть уменьшено до одной и разложение (2) представимо в виде

$$f(x_1, x_2, \dots, x_n) = \sum_{q=1}^{2n+1} h \left[\sum_{p=1}^{n} \varphi_q^p (x_p) \right].$$

В конце 80-х была опубликована работа Хехт-Нильсена [9], на которую ссылаются некоторые авторы в качестве обоснования возможностей нейронных сетей. В ней, в частности, показано, что непрерывная функция $y = f(x_1, \dots, x_n)$ может быть представлена с помощью двухслойной сети, содержащей на внутреннем слое *(2n+1)* нейрон, выходные значения которых определяются соотношениями

$$z_k = \sum_{j=1}^{n} \lambda^k \varphi(x_j + \varepsilon k) + k, k = \overline{1, 2n+1},$$

а результирующее значение сети - выражением

$$y = \sum_{k=1}^{2n+1} g(z_k),$$

где λ – некоторая действительная постоянная, $\varphi(x)$ – произвольная ограниченная возрастающая функция, $0<\varepsilon\leq\delta$ – произвольное положительное число, $g(x)$ – непрерывная функции, вид которой определяется значением ε и f .

Заключения, предполагающие аффинный характер преобразования сигналов, поступающих на внутренние нейроны, содержатся в статье [10]. В ней показано, что множество таких образом сформированных «нейронных» функций является плотным во множестве непрерывных и поэтому любая непрерывная функция со сколь угодно высокой степенью точности может быть аппроксимирована нейронной сетью с произвольной нелинейной функцией активации.

Несколько ранее к близким результатам пришел и автор работы [11]. Он ввел понятие сигмоидальной функции

$$\sigma(t) \to \begin{cases} 1, & t \to +\infty \\ 0, & t \to -\infty \end{cases}$$

и показал, что множество функций

$$G(x) = \sum_{j=1}^{N} \alpha_j \sigma(y_j^T x + \theta_j),$$

где $y_j \in R^n, \alpha_j, \theta_j$ – постоянные, $x \in R^n$ – переменная, а $y_j^T x$ – скалярное произведение, плотно во множестве непрерывных функций.

Заметим, что число нейронов, необходимое для получения аппроксимации нужной точности, в перечисленных выше работах не указано.

Такие рекомендации содержатся в работах [12], [13]. В них, без ссылок на источники, формулируется теорема, которая «доказывает представимость функции многих переменных достаточно общего вида с помощью двухслойной нейронной сети с прямыми полными связями с n нейронами входного слоя, ($2n\ +1$) нейронами скрытого слоя с заранее известными ограниченными функциями активации (например, сигмоидальными) и m нейронами выходного слоя с неизвестными функциями активации». Там же приводятся и формулы для установления числа нейронов в скрытых слоях на основе оценки необходимого числа весов L_w в сети с сигмоидальными передаточными функциями:

$$\frac{mN}{1+\log_2 N} \leq L_w \leq m\left(\frac{N}{m}+1\right)(n+m+1)+m,$$

где n – размерность входного сигнала, m – размерность выходного, а N – число элементов обучающей выборки. Откуда, рекомендуемое число нейронов для двухслойной сети, как отмечают авторы, должно составлять

$$L = \frac{L_w}{n+m} .$$

Там же приводятся и другие оценочные выражения, в частности,

$$2(n+L+m) < N < 10(n+L+m)$$

или

$$\frac{N}{10} - n - m < L < \frac{N}{2} - n - m.$$

Заметим, что данные соотношения ставят в зависимость число нейронов в сети от объема исходных данных, что не согласовывается с приведенными выше теоремами, в которых их количество определяется числом переменных в рассматриваемой функции. Если же допустить, что они гарантируют достижение заданной точности аппроксимации, то величина погрешности в этих соотношениях не фигурирует.

Сведения о количественных соотношениях отсутствуют, также, и в монографиях по глубокому обучению [14-17]. В них рассматриваются нейронные сети, как уже сложившиеся конструкции, обсуждаются их архитектурные особенности, алгоритмы обучения и многочисленные приложения.

Вопросы аппроксимации немонотонных функций на примере двухслойной сети (рис. 1) рассматриваются в работе [18]. В качестве алгоритма обучения сети используется градиентный спуск, а в качестве функций активации – гиперболический тангенс или *ReLU*. На основе результатов экспериментов авторы отмечают проблемы, связанные с инициализацией параметров сети случайным образом, иллюстрируют влияние многократного повторения процесса обучения и вида функции активации на качество аппроксимации.

Перейдем к предметному проектированию рассматриваемой сети.

2 Описание сети

Архитектура сети выбрана (рис. 1), ее функционирование описано (1.1), теперь, необходимо выбрать числовой ряд, определиться с алгоритмом обучения, реализовать все это программно и провести эксперименты.

В сети (1) восемь параметров. Временной ряд такого же объема сформируем на основе синусоиды на отрезке *[0 , 2pi]*, разделив его на 7 равных частей. Таким образом,

$$x_i = (2pi/7)k\ ,\ y_i = sin(x_i)\ ,\ k=0,\ \dots,\ 7. \qquad (3)$$

Функцию активации будем варьировать.

Результат решения задачи минимизации, вообще говоря, зависит от выбора начальной точки. Поэтому инициализацию параметров сети выполним на основе последовательности Соболя, которая, как отмечают ее авторы, обладает наилучшими, из известных распределений, равномерными свойствами [19], беря в качестве начальных заданное число ее первых точек. Обучение сети проведем в среде Matlab с использованием функции *lsqnonln* [20], которая осуществляет минимизацию суммарной ошибки нейронной сети

$$mist(c_0,\dots,c_7) = \sum_{k=0}^{7}\left(\sin(x_k) - \left(\sum_{i=1}^{3} c_{i+3} f(c_i x_k + c_0) + c_7\right)\right)^2 \to min \Rightarrow c_0,\dots,c_7 \qquad (4)$$

на заданном числовом наборе методами Гаусса-Ньютона и Левенберга - Марквардта. Эти методы, как отмечается в [21] являются модификациями метода Ньютона и комбинирует его с методом градиентного пуска, учитывают информацию о кривизне поверхности и позволяют обеспечивать увеличение шага вдоль пологих участков поверхности отклика и уменьшение вдоль крутых ее спусков.

3 Интерполяция

Эксперимент № 1. Функция активации - сигмоид, т.е. в (1), (3) $f(x)=1/(1+e^{-x})$, сумма квадратов его отклонений в узловых точках (3) равна *10.5129.*
Число пробных начальных точек при инициализации параметров сети возьмем равным 32. В качестве иллюстрации распределения точек последовательности Соболя в единичном кубе K_8 на рис.2 приведены, для примера, расположение проекций этих точек на плоскости координат x_1 - x_2 .

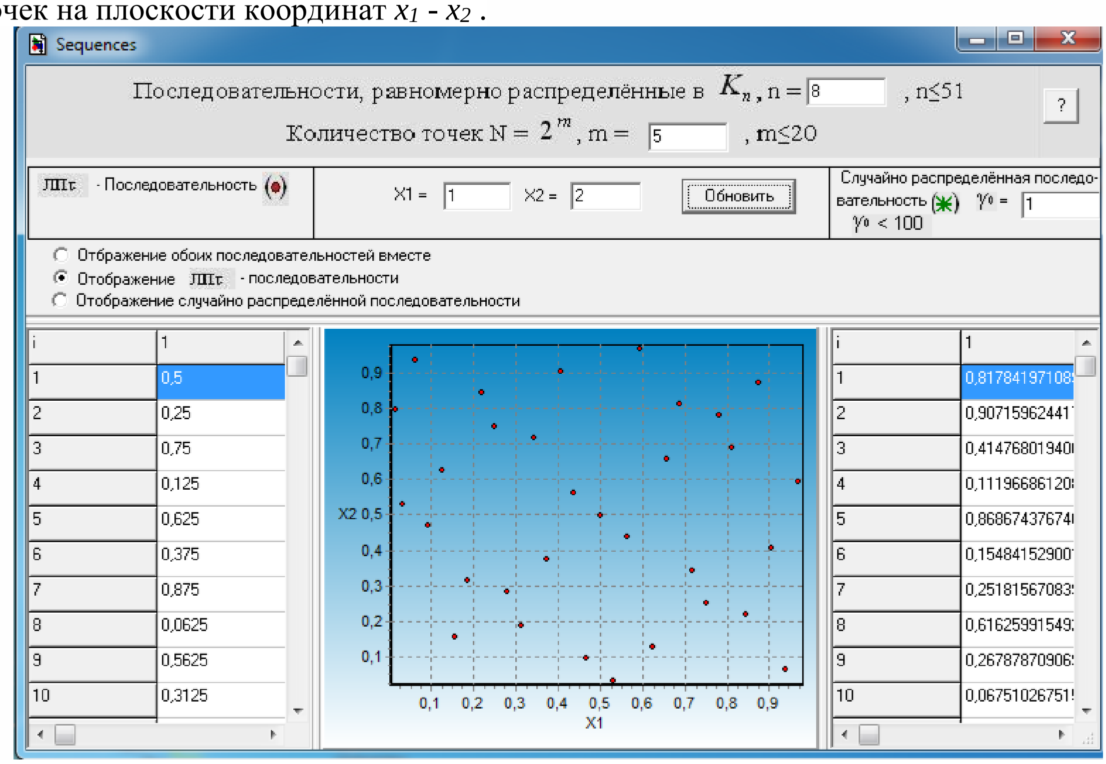


Рисунок 2 – Расположение проекций первых 32-х пробных точек последовательности Соболя в пространстве K_8 на плоскость x_1 - x_2

Дальнейший порядок действий состоял в следующем. Для каждой пробной точки проводился запуск режима обучения сети, т.е. нахождение минимума (4). Оказалось, что при настройках программы по умолчанию, а это 800 итераций вычисления функции (4), желаемый «нулевой» минимум задачи оптимизации не достигался, поэтому набор $(c_0,\ldots,c_7)$, полученный после первого запуска программы, принимался в качестве начального и программа *lsqnonln* запускалась повторно. В результате такой двукратной минимизации были достигнуты следующие значения ошибки сети (4), с точностью до четырех десятичных знаков после запятой:

mist =

Columns 1 through 8

1.8557 0.6682 0.5383 0.7487 1.0704 0.0343 1.2080 0.5555

Columns 9 through 16

0.8144 0.6671 1.8640 2.3487 2.0511 1.6084 1.4635 0.8872

Columns 17 through 24

0.0000 0.8494 1.0082 0.5985 2.1903 0.8537 0.5594 0.5557

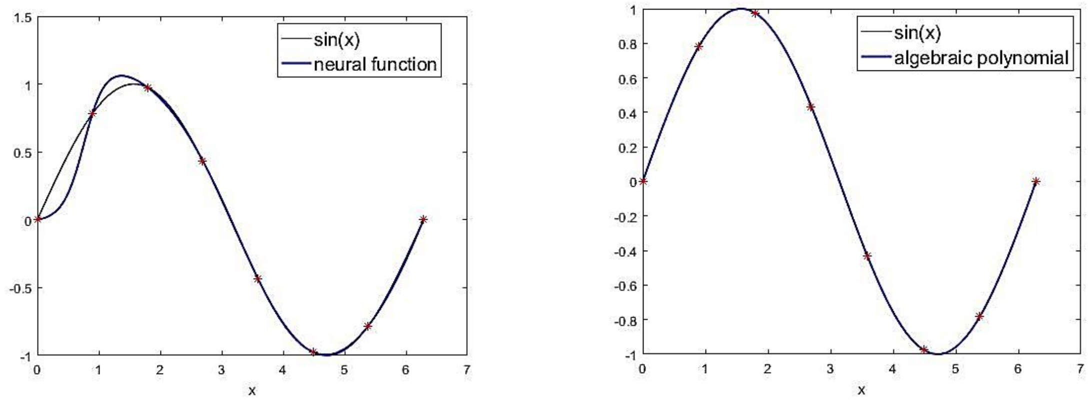
Columns 25 through 32

0.9718 2.4378 2.2375 0.7043 2.5136 0.9860 0.6793 0.8831

Наилучший результат получен для 17-й пробной точки, соответствующее ей значение минимума (4) составляет *mistake = 1.5859e-18,* что, наверное, можно считать идеальным. Соответствующий оптимальный набор параметров сети был следующим:

$$copt = [c_0 \ c_1 \ c_2 \ c_3 \ c_4 \ c_5 \ c_6 \ c_7] =$$
$$= [\,-4.4531 \quad 0.5432 \quad 1.3161 \quad 5.8293 \quad 11.7780 \quad -4.2617 \quad 1.1943 \quad -0.1002].$$

Графики синусоиды на рассматриваемом участке *[0, 2pi]* и результаты ее интерполяции приведены на рис. 3: на рис. 3а – интерполяция с помощью нейронной сети, на рис. 3б - с помощью алгебраического многочлена 7-й степени. Визуально - качество интерполяции при алгебраической интерполяции выше, да, и трудоемкость ее построения существенно ниже, чем при нейронной.

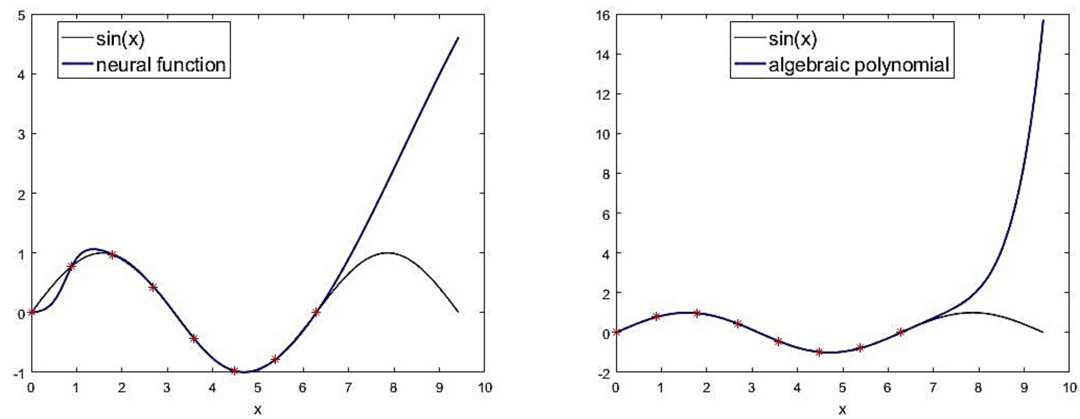


а) б)

Рисунок 3 – Интерполяция синусоиды с помощью нейронной сети а) и алгебраическая интерполяция б)

На рис. 4 приведены графики на промежутке *[0, 3pi]*, иллюстрирующие экстраполяционные возможности полученных конструкций, откуда следует, что прогностические возможности обоих типов интерполяции явно оставляют желать лучшего.

В качестве замечания, отметим, что оптимальная, в данном случае, 17-я пробная точка, при иной нумерации точек, могла бы оказаться и другой, но, тем не менее, понятно, что единичный случайный выбор пробной точки не является надежным для нахождения глобального экстремума.



а) б)

Рисунок 4 – Нейронная а) и алгебраическая б) экстраполяция синусоиды

Эксперимент № 2. Функция активации - гиперболический тангенс, т.е. в (1) и (3) $f(x)=(e^x-e^{-x})/(e^x+e^{-x})$, сумма квадратов его отклонений в узловых точках (2) равно *10.4413* . Алгоритм обучения остается прежний.

В этом случае набор ошибок

mist =

Columns 1 through 8

0.8420 0.6392 0.8770 0.5635 1.0704 0.7631 0.3099 0.5408

Columns 9 through 16

0.7354 0.7621 0.5383 0.6860 0.5383 0.7940 2.2869 1.4934

Columns 17 through 24

0.5383 1.1230 1.0486 0.5383 1.7664 1.6062 0.5330 2.2137

Columns 25 through 32

0.5506 0.7646 0.5384 0.5383 0.5519 0.9423 0.6146 0.5537

Наилучшим является результат *0.3099,* что заметно хуже, чем в случае сигмоида.

Эксперимент № 3. Функция активации – функция *ReLU*, т.е. в (1) и (3) $f(x)=max\{0 , x\})$, сумма квадратов ее отклонений в узловых точках равно *129.3426.* В этом случае набор ошибок

mist =

Columns 1 through 8

2.2424 2.2424 2.2424 2.2424 1.9349 2.2424 1.6135 2.2424

Columns 9 through 16

2.1366 2.1774 2.2424 2.2424 2.2424 0.5508 2.0019 3.5000

Columns 17 through 24

1.9363 1.9374 2.2424 1.9372 2.0348 0.5908 2.2424 2.2424

Columns 25 through 32

2.2424 2.2424 1.9774 3.5000 0.5508 1.9401 2.2424 2.1562

Наилучшим является результат *0.5508,* что значительно хуже, чем в первых двух случаях.

Эксперимент № 4. Функция активации *f(x)=sin(x),* сумма квадратов ее отклонений в узловых точках = 0.

В этом случае набор ошибок:

mist =

Columns 1 through 8

0.0000 0.0000 0.0000 0.0000 0.0000 0.0000 0.3146 0.0000

Columns 10 through 16

0.0000 0.0000 0.2604 0.0000 0.0000 0.3762 0.0000 3.4904

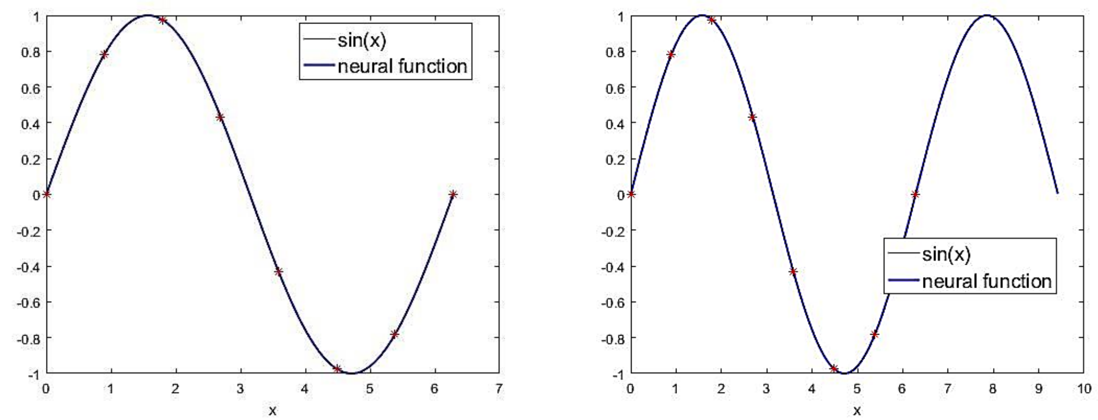
Columns 17 through 24

0.0000 0.0000 0.3766 0.0000 0.0000 0.0000 0.2775 0.0000

Columns 25 through 32

0.0157 0.0000 0.8371 0.0000 0.1583 0.0000 0.1005 0.0788

Наилучший результат достигается на 22-й пробной точке и соответствующая ей погрешность равна *6.7765e-26.* Графические материалы, иллюстрирующие этот результат представлены на рис. 5.



а) б)

Рисунок 5 – Графики интерполяции (слева) и экстраполяции (справа) синусоиды нейронной сетью с функцией активации *f(x)=sin(x)*

Высокий уровень совпадения, как по интерполяции, так и экстраполяции вполне естественен, так как функция активации адекватна исходной информации. В этом случае оптимальный набор весовых коэффициентов имеет вид

copt = -0.0000 1.0000 0.6177 -0.6177 1.0000 1.3080 1.3080 0.0000

и, после подстановки их в (1_1) имеем

$$y = f(x) + 1.3080 \cdot f(0.6177x) + 1.3080 \cdot f(-0.6177x)$$

или

$$y = \sin(x) + 1.3080 \cdot \sin(0.6177x) + 1.3080 \cdot \sin(-0.6177x)$$

или

$$y = \sin(x)\,.$$

Это наводит на мысль, что подбором функции активации можно улучшить качество нейронной сети и наличие адекватности *f(x)* исходной информации, по-видимому, можно принять в качестве критерия при ее подборе в процессе проектировании нейронной сети. Безусловно, эта гипотеза требует экспериментального подтверждения.

Можно ли улучшить качество аппроксимации за счет увеличения объема исходных данных? Ответ на этот вопрос иллюстрируется ниже.

4 Аппроксимация

В качестве функции активации используем сигмоид, как наиболее надежный из традиционных в проведенных экспериментах, и варьируем число узловых точек на промежутке *[0, 2pi]*. Обучение сети проводим по прежнему алгоритму. Число точек последовательно удваиваем и принимаем равным *n=16, 32, 64, 128.* Полученные результаты представлены на рис. 6.

Следует отметить, что визуально, судя по графикам, качество аппроксимации, с увеличением объема данных, улучшается и достигает максимума при числе точек *n=64*. Правда, обращает на себя внимание его последующее ухудшение при последующем удвоении узловых точек, т.е. при *n=128.* Так, при *n=64* достигнутое минимальное значение функции цели составляет *minmist = 0.0017,* а при *n=128* уже *minmist = 4.1742*. Сравнивая соответствующие графики, можно предположить, что это явление следует отнести на счет неудачного завершения выполнения процедуры *lsqnonln* программой *matlab.* Действительно, этот факт подтверждается, так как значение функции цели *mist* (4), рассчитанное для числа точек $n=128$ для нейронной сети, построенной по результатам $n=64$, равно $mist = 0.0032$. И к этому значению можно приблизится, взяв оптимальный набор коэффициентов сети, определенный для $n=64$, в качестве начального при обучении сети для $n=128$. Однако это достаточно трудоемко. Поэтому отмеченное обстоятельство еще раз подчеркивает капризность процедуры идентификации параметров нейронной сети.

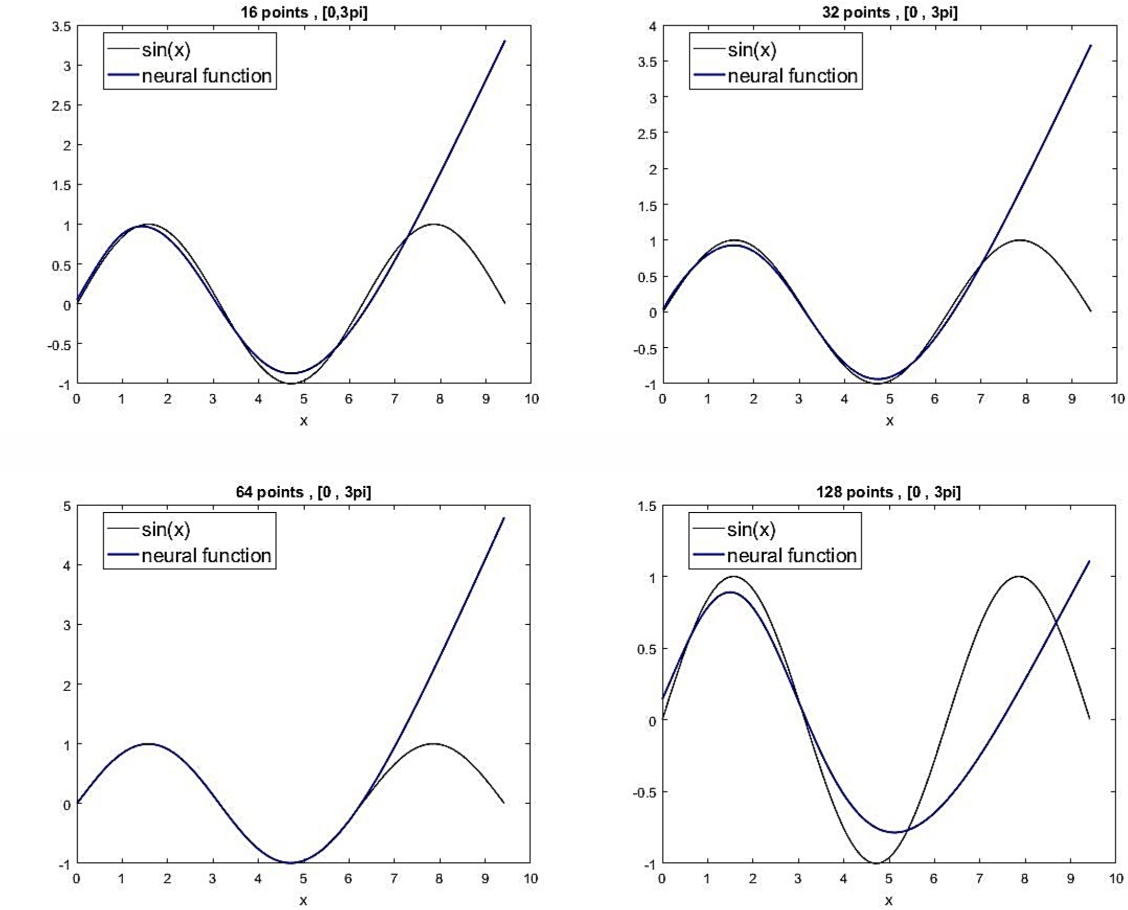


Рисунок 6 – Результаты аппроксимации синуса для числа узловых точек $n=16$, 32 , 64 , 128.

Признаков улучшения экстраполяции функции с увеличением числа интерполяционных узловых точек не наблюдается. К числу резервов ее улучшения, возможно, следует отнести усложнение архитектуры нейронной сети и (или) надлежащий подбор функции активации (см. результаты Эксперимента № 4).

Проведение алгебраической аппроксимации не проводилось.

Заключение

Проведенный анализ позволяет отметить некоторые недостатки и определенные достоинства нейронных сетей.

Так, эксперименты с функцией одной переменной продемонстрировали невысокие интерполяционные возможности нейронной сети и высокую трудоемкость ее обучения, по сравнению с алгебраическими подходами. Для получения в пакете *Matlab* удовлетворительного результата при интерполяции синуса по восьми равномерно расположенным узловым точкам на периоде его колебаний пришлось провести двукратную оптимизацию с использованием процедуры *lsqnonln*, реализующую минимизацию евклидовой нормы векторной функции. Причем качество аппроксимации синусоиды вне узловых точек с помощью нейронной интерполирующей сети заметно уступает многочленной интерполяции, построенной по тем же узловым точкам. Увеличение узловых точек позволило существенно повысить качество нейронной аппроксимации и приблизится к алгебраической, достигнутой при меньшем количестве узлов. Однако экстраполяция синусоиды с помощью оптимальной интерполирующей нейронной сети, даже, качественно не отражает ее поведения. Поэтому целесообразность использования нейронных сетей к задачам интерполяции и экстраполяции функции одной переменной вызывает большие сомнения.

Следует отметить важность удачного выбора функции активации. Наиболее надежно в проведенных экспериментах проявил себя сигмоид. Обратим внимание, также, на возможную связь между степенью адекватности функции активации исходной информации и качеством проектируемой нейронной сети (см. Эксперименты № 1 и 3).

Если же перейти к рассмотрению функции нескольких переменных, то мнение о достоинствах нейронных сетей начинает усиливаться. В случае нейронной сети увеличение числа переменных в исходной функции повлечет естественные изменения в ее архитектуре, в частности, увеличение числа нейронов на входном слое и, вполне возможные, внутренние изменения в числе скрытых слоев и количестве нейронов на каждом из них. Но, некоторые принципиальные моменты можно отметить исходя из общих соображений.

Так, при произвольном расположении узлов, задача алгебраической интерполяция функции нескольких переменных, вообще говоря, не имеет решения [4]. Известные ее подходы для функции двух переменных состоят в формировании прямоугольной сетки в плоскости независимых переменных или выполнении триангуляции исходных данных и построении для каждой грани соответствующего многогранника билинейной или плоской интерполирующей функции [20]. В результате этого исходная функция описывается набором ее интерполянтов на каждой грани этого многогранника. Однако использование для этих целей нейронной сети позволяет надеяться на получение аналитической зависимости, обеспечивающей интерполяцию на всем объеме данных.

При большем числе переменных аппроксимация значения функции в отдельных математических пакетах [20] осуществляется посредством покоординатной линейной интерполяции, однако в этом случае число слагаемых при числе переменных равным 10 становится больше тысячи и ее реализация становится технически затруднительной. В этой связи, из числа классических вспоминаются методы группового учета аргументов [22], позволяющие проводить пошаговое усложнение модели при неизменном объеме данных. Они отличаются универсальностью, однако, прозрачностью своих алгоритмов уступают нейронным сетям.

Список литературы

1. Ермоленко, Т. В. Исследование эффективности прогностических моделей для системы анализа и мониторинга энергопотребления на предприятиях угольной промышленности/ Т.В. Ермоленко, В.Н. Котенко, В.В. Винник. *Проблемы искусственного интеллекта*. 2022. №4 (27) . С. 25-34.

2. Анцыферов, С. С. Методология развития интеллектуальных систем/ С.С. Анцыферов, А.С. Сигов, К.Н. Фазилова. *Проблемы искусственного интеллекта*. 2022. №2(25).С.42-47.
3. Ермоленко, Т. В. Классификация аномалий сердцебиения с помощью глубокого обучения /Т.В. Ермоленко, Д.В. Ролик. *Проблемы искусственного интеллекта*. 2022. № 1(24). С. 40-53.
4 Березин, И.С. *Методы вычислений,* т. 1/ И.С. Березин, Жидков Н.П. Москва: изд-во физ.-мат. литературы, 1962. 464 с.
5. Фихтенгольц, Г.М. *Курс дифференциального и интегрального исчисления*. Т.III / Г.М. Фихтенгольц. Москва: Наука, 1966.- 656 с.
6. Бернштейн, С.Н. Доказательство теоремы Вейерштрасса, основанное на теории вероятностей/ С.Н. Бернштейн. *Собрание сочинений,* т.1, изд-во АН СССР, 1952. - с. 105-106.
7. Колмогоров, А.Н. *Избранные труды. Математика и механика*. Москва: Наука, 1985. с. 179-182.
8. Lorentz, George. Metric entropy, widths, and superpositions of functions (англ.*). American Mathematical Monthly : journal*. — 1962. — Vol. 69. — P. 469—485.
9. Hecht-Nielsen, R. Kolmogorov’s Mapping Neural Network Exsistence Theorem / R. Hecht-Nielsen. *Материалы Первой международной конференции IEEE по нейронным сетям*. 1987, III: pp. 11-13.- Режим доступа https://cs.uwaterloo.ca/~y328yu/classics/Hecht-Nielsen.pdf
10. Горбань, А.Н. Обобщенная аппроксимационная теорема и вычислительные возможности нейронных сетей/ А.Н. Горбань. *Сибирский журнал вычислительной математики*, 1998. Т.1, № 1. С. 12-24.
11. Cybenko, G. Approximation by Superpositions of a Sigmoidal Function/ G. Cybenko. *Mathematics of Control, Signals, and Systems*, 1989, 2. p.p.303 - 314.
12. Круглов, В.В. *Искусственные нейронные сети. Теория и практика* / В.В. Круглов, В.В. Борисов. Москва: Горячая линия - Телеком, 2002 - 382 с.
13. Ясницкий, Л. Искусственный интеллект: популярное введение для учителей и школьников / Л. Ясницкий. *Журнал «Информатика»*, № 23 (600), 1-15.12.2009 Режим доступа: https://inf.1sept.ru/view_article.php?ID=200902304
14. Mitchell, T.M. Machine Learning / T.M.Mitchel. – McGraw. Hill Science/ Engineering/ Math, 1997. - 432 p.
15. Хайкин, С. *Нейронные сети: полный курс/* С. Хайкин. Москва: Издательский дом «Вильямс», 2006. - 1104 с.
16. Галушкин, А.И. *Нейронные сети: основы теории* / А.И. Галушкин. Москва: Горячая линия - Телеком. 2012. - 496 с.
17. Николенко, С. *Глубокое обучение* / С.Николенко, А.Кадурин, Е. Архангельская. Санкт- Петербург: Питер, 2018. - 480 с.
18. Галкин, В.А. Некоторые аспекты аппроксимации и интерполяции функций искусственными нейронными сетями/ В.А. Галкин, Т.В. Гавриленко, А. Д. Смородинов. *Вестник КРАУНЦ. Физ.-мат. науки.* 2022. Т. 38. № 1. С. 54-73.DOI: 10.26117/2079-6641-2022-38-1-54-73.
19. Соболь, И.М. *Выбор оптимальных параметров в задачах со многими критериями*/ И.М. Соболь, Р.Б. Статников. - Москва: Наука, 1981. - 110 с.
20. Кетков, Ю.Л. *Матлаб 7: программирование, численные методы* / Ю.Л., Кетков, А.Ю. Кетков, М.М. Шульц. Санкт - Петербург: БХВ-Петербург, 2005.- 752 с.
21. *Алгоритм Левенберга — Марквардта* Режим доступа: https://ru.wikipedia.org/wiki/ Алгоритм_ Левенберга _—_Марквардта
22. Ивахненко, А.Г. *Индуктивный метод самоорганизации моделей сложных систем* / А.Г. Ивахненко. Киев : Наукова думка , 1981 296 с.

References

1 Ermolenko, T. V. *Issledovanie effektivnosti prognosticheskih modelej dlya sistemy analiza i monitoringa energopotrebleniya na predpriyatiyah ugol'noj promyshlennosti*/ T.V. Ermolenko, V.N. Kotenko, V.V. Vinnik// Problemy iskusstvennogo intellekta. 2022. №4 (27) . S. 25-34.
2. Ancyferov, S. S. *Metodologiya razvitiya intellektual'nyh sistem*/ S.S. Ancyferov, A.S. Sigov, K.N. Fazilova // Problemy iskusstvennogo intellekta. 2022. №2(25).S.42-47.
3. Ermolenko, T. V. *Klassifikaciya anomalij serdcebieniya s pomoshch'yu glubokogo obucheniya* /T.V. Ermolenko, D.V. Rolik // Problemy iskusstvennogo intellekta. 2022. № 1(24). S. 40-53.
4 Berezin, I.S. *Metody vychislenij*, t. 1/ I.S. Berezin, ZHidkov N.P. - Moskva: izd-vo fiz.-mat. literatury, 1962. - 464 s.

5. Fihtengol'c, G.M. *Kurs differencial'nogo i integral'nogo ischisleniya*. T.III/ G.M. Fihtengol'c. - Moskva: Nauka, 1966.- 656 s.
6. Bernshtejn, S.N. *Dokazatel'stvo teoremy Vejershtrassa, osnovannoe na teorii veroyatnostej*/ S.N. Bernshtejn, Sobranie sochinenij, t.1, izd-vo AN SSSR, 1952. - s. 105-106.
7. Kolmogorov, A.N. *Izbrannye trudy. Matematika i mekhanika*. - Moskva: Nauka, 1985. s. 179-182
8. Lorentz, George. Metric entropy, widths, and superpositions of functions (angl.) // American Mathematical Monthly : journal. — 1962. — Vol. 69. — P. 469—485.
9. Hecht-Nielsen, R. Kolmogorov's Mapping Neural Network Exsistence Theorem// R. Hecht-Nielsen, Materialy Pervoj mezhdunarodnoj konferencii IEEE po nejronnym setyam.1987, III: pp. 11-13.- https://cs.uwaterloo.ca/~y328yu/classics/Hecht-Nielsen.pdf Rezhim dostupa - Zaglavie s ekrana. 09.03.2024
10. Gorban', A.N. *Obobshchennaya approksimacionnaya teorema i vychislitel'nye vozmozhnosti nejronnyh setej*/ A.N. Gorban' //Sibirskij zhurnal vychislitel'noj matematiki, 1998. T.1, № 1. S. 12-24.
11. Cybenko, G. Approximation by Superpositions of a Sigmoidal Function/ G. Cybenko// Mathematics of Control, Signals, and Systems, 1989, 2. - p.p.303 - 314.
12. Kruglov, V.V. *Iskusstvennye nejronnye seti. Teoriya i praktika* /V.V. Kruglov, V.V. Borisov.- Moskva: Goryachaya liniya - Telekom, 2002 - 382 s.
13. YAsnickij, L. *Iskusstvennyj intellekt: populyarnoe vvedenie dlya uchitelej i shkol'nikov*./ L. YAsnickij // ZHurnal «Informatika», № 23 (600), 1-15.12.2009 .- https://inf.1sept.ru/view_article.php?ID=200902304 - Rezhim dostupa - Zaglavie s ekrana. 09.03.2024.
14. Mitchell, T.M. *Machine Learning*/T.M.Mitchel. - McGraw - Hill Science/ Engineering/ Math, 1997. - 432 p.
15. Hajkin, S. *Nejronnye seti: polnyj kurs*/ S. Hajkin. - Moskva: Izdatel'skij dom «Vil'yams», 2006. - 1104 s.
16. Galushkin, A.I. *Nejronnye seti: osnovy teorii*/ A.I. Galushkin. - Moskva: Goryachaya liniya - Telekom. 2012. - 496 s.
17. Nikolenko, S. *Glubokoe obuchenie* /S.Nikolenko, A.Kadurin, E. Arhangel'skaya. - Sankt- Peterburg: Piter, 2018. - 480 s.
18. Galkin, V.A. *Nekotorye aspekty approksimacii i interpolyacii funkcij iskusstvennymi nejronnymi setyami*/ V.A. Galkin, T.V. Gavrilenko, A. D. Smorodinov // Vestnik KRAUNC. Fiz.-mat. nauki. 2022. T. 38. № 1. S. 54-73.DOI: 10.26117/2079-6641-2022-38-1-54-73.
19. Sobol', I.M. *Vybor optimal'nyh parametrov v zadachah so mnogimi kriteriyami*/ I.M. Sobol', R.B. Statnikov. - Moskva: Nauka, 1981. - 110 s.
20. Ketkov, YU.L. *Matlab 7: programmirovanie, chislennye metody*/ YU.L., Ketkov, A.YU. Ketkov, M.M. SHul'c - Sankt - Peterburg: BHV-Peterburg, 2005.- 752 s.
21. Algoritm Levenberga — Markvardta - https://ru.wikipedia.org/wiki/ Algoritm_ Levenberga _—_Markvardta - Rezhim dostupa - Zaglavie s ekrana. 09.03.2024
22. Ivahnenko, A.G. *Induktivnyj metod samoorganizacii modelej slozhnyh sistem* / A.G. Ivahnenko — Kiev : Naukova dumka , 1981 — 296 s.

RESUME

V. N. Belovodskiy

On Interpolation and Approximation of Functions Using Neural Networks

In paper the typical problems of the theory of approximation of functions are considered. A two-layer neural network was implemented in the matlab environment and interpolation and approximation of one of the elementary functions were performed on its basis, the results obtained were compared with algebraic approaches. The advantages and disadvantages of the neural technique are noted, generalizations are made.

The conducted computational experiments have demonstrated the low interpolation capabilities of the neural network and the high complexity of its training. The quality of the function approximation outside the nodal points is noticeably inferior to the polynomial interpolation. Increasing the nodal points allows you to significantly improve the quality of neural approximation and get closer to the algebraic one. However, extrapolation of a

sinusoid using a neural network does not even qualitatively reflect its behavior. Therefore, the use of neural networks for the tasks of interpolation and extrapolation of a function of one variable is highly questionable.

Slightly different opinions may be expressed regarding the functions of several variables. The problem of algebraic interpolation of a function of several variables with an arbitrary disposition of nodes, generally speaking, is unsolvable. However, the network usage for these purposes allows us to hope for obtaining an analytical dependence that provides interpolation on the entire volume of data. And due to the transparency of the algorithms for constructing neural networks, it seems more promising to use them for approximating such functions.

РЕЗЮМЕ

В. Н. Беловодский

Об интерполяции и аппроксимации функций с использованием нейронных сетей

В статье рассматриваются типовые задачи теории приближения функций, на примере двухслойной нейронной сети проводится интерполяция и аппроксимация одной из элементарных функций и полученные результаты сравниваются с алгебраическими подходами. Расчеты проводятся на основе программы, разработанной в среде Matlab. Отмечены достоинства и недостатки нейронного подхода, сделаны обобщения.

В статье отмечается, что вычислительные эксперименты продемонстрировали невысокие интерполяционные возможности нейронной сети и высокую трудоемкость ее обучения. Для получения удовлетворительного результата при интерполяции синуса по восьми равномерно расположенным узловым точкам на периоде его колебаний пришлось проводить двукратную минимизацию с использованием процедуры *lsqnonln* программы Matlab. Качество аппроксимации функции вне узловых точек заметно уступает многочленной интерполяции. Увеличение узловых точек позволяет существенно повысить качество нейронной аппроксимации и приблизится к алгебраической. Однако экстраполяция синусоиды с помощью нейронной сети, даже, качественно не отражает ее поведения. Поэтому целесообразность использования нейронных сетей к задачам интерполяции и экстраполяции функции одной переменной вызывает большие сомнения.

Несколько иные соображения можно высказать относительно функций нескольких переменных. Задача алгебраической интерполяции функции нескольких переменных при произвольном расположении узлов, вообще говоря, не имеет решения. Однако использование для этих целей нейронной сети позволяет надеяться на получение аналитической зависимости, обеспечивающей интерполяцию на всем объеме данных. В силу прозрачности своих алгоритмов использование нейронных сетей представляется перспективным и для аппроксимации таких функций.

Беловодский В. Н. – к.т.н., доцент, ФГБОУ ВО ДонНТУ, кафедра компьютерного моделирования и дизайна, 283001, Донецк, ул. Артема, 58 ,тел +7(949) 334-9150, v.belovodskiy@gmail.com. *Область научных интересов*: моделирование технических систем, нелинейная динамика, фракталы и математический дизайн, нейронные сети

Статья поступила в редакцию 17.04.2024.