

УДК 661.322
DOI: 10.26467/2079-0619-2018-21-2-40-50

О ПРИМЕНЕНИИ ИСКУССТВЕННОЙ НЕЙРОННОЙ СЕТИ ДЛЯ РЕШЕНИЯ ЗАДАЧИ АППРОКСИМАЦИИ НЕЛИНЕЙНЫХ ЗАВИСИМОСТЕЙ

В.Н. АГЕЕВ¹

¹Московский государственный технический университет гражданской авиации,
г. Москва, Россия

В работе рассматриваются вопросы, связанные с использованием искусственной нейронной сети (ИНС) для аппроксимации экспериментальных данных. Одной из проблем при разработке ИНС является выбор подходящей функции активации для нейронов скрытого слоя и регулировка параметров функции в процессе обучения сети. В статье рассматривается трехслойный перцептрон с одним скрытым слоем, каждый нейрон которого имеет функцию активации в виде гауссовой кривой. Выбор радиально-базисной функции активации позволяет применить в процессе обучения сети прямой метод определения весовых коэффициентов – метод наименьших квадратов. Качество аппроксимации при этом во многом зависит от правильности выбора значения параметра функции активации, которым в данном случае является ширина колокола гауссовой кривой. На практике этот параметр определяют путем проведения численных экспериментов. Это достаточно трудоемкий процесс. В данной работе предлагается определять значение этого параметра по обучающей выборке, представляющей собой координаты набора точек тестовой кривой с заданными свойствами. Эти свойства задаются исходя из априорных сведений об аппроксимируемой функции (линейная, квадратичная, логарифмическая, экспоненциальная зависимость). Поскольку тестовая кривая задается в явном виде, параметр функции активации определяется из условия достижения минимума интеграла от квадрата разности между значениями тестовой функции и выходным сигналом сети. Такой подход гарантирует получение аппроксимирующей кривой с хорошими свойствами, в частности, характеризуется отсутствием в ее графике так называемых «осцилляций» – многочисленных точек перегиба.

Ключевые слова: искусственная нейронная сеть, функция активации, обучающая выборка, аппроксимация функций.

ВВЕДЕНИЕ

Задача аппроксимации функции для искусственной нейронной сети (ИНС) формулируется как задача контролируемого обучения. Суть задачи состоит в следующем. Имеются значения функции в отдельных точках (обучающая выборка), система базисных функций и набор регулируемых весовых коэффициентов. Необходимо обучить сеть, т. е. подобрать весовые коэффициенты так, чтобы зависимость выходного сигнала от сигнала на входе наилучшим образом аппроксимировала обучающий набор точек.

Простейшая ИНС состоит всего из одного суммирующего нейрона с двумя входными элементами и одним выходом (рис. 1).

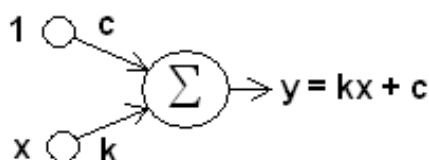


Рис. 1. Искусственная нейронная сеть с одним суммирующим нейроном
Fig. 1. Artificial neural network with one summing neuron

На один из входов подается сигнал с уровнем 1, на другой – измеряемая величина x . Параметрами сети являются весовые коэффициенты c и k , определяющие значения сигнала.

лов, поступающих на вход суммирующего нейрона. На его выходе появляется сигнал $y = k \cdot x + c$. Таким образом, данная сеть моделирует устройство преобразования входного сигнала x по линейному закону, параметры которого можно изменять, варьируя весовые коэффициенты c и k .

Эту сеть можно использовать для решения задачи линейной аппроксимации некоторой функции $y = f(x)$, заданной набором пар чисел (x_i, y_i) , $i = 1, 2, \dots, M$. Такая задача возникает при обработке экспериментальных данных, полученных в результате измерений выходных значений y на выходе некоторого устройства при подаче на его вход сигнала x . Требуется провести прямую линию через множество точек на плоскости, наилучшим образом отражающую характер зависимости y от x , как показано на рис. 2.

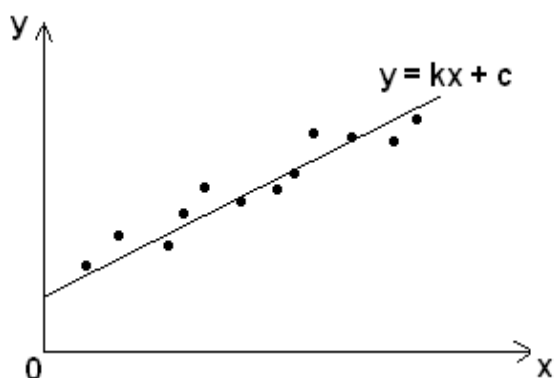


Рис. 2. Линейная аппроксимация экспериментальных данных
Fig. 2. Linear approximation of experimental data

Наилучшее приближение дает прямая линия $y = kx + c$, с минимальным значением ошибки аппроксимации (целевой функции)

$$S = \sum_{i=1}^M (kx_i + c - y_i)^2,$$

где M – количество экспериментальных точек. Значения коэффициентов k и c определяются в соответствии со стандартным методом наименьших квадратов (см., например, [1]). Этот метод сводится к решению системы линейных уравнений, получаемых из необходимого условия минимума функции $S(k, c)$:

$$\begin{aligned} \frac{\partial S}{\partial k} &= 2 \sum_{i=1}^M (kx_i + c - y_i) \cdot x_i = 0, \\ \frac{\partial S}{\partial c} &= 2 \sum_{i=1}^M (kx_i + c - y_i) = 0. \end{aligned} \tag{2}$$

С точки зрения теории искусственных нейронных сетей заданное множество экспериментальных точек (x_i, y_i) является обучающим набором, а найденные параметры k и c – весовыми коэффициентами связей.

После обучения (определения весов связей для заданного обучающего набора) нейронная сеть способна выполнять поставленную задачу: моделировать исследуемое устройство. Любой сигнал x на входе сети вызовет появление на выходе сигнала y , с минимальной погрешностью соответствующего реальному отклику на подаваемый сигнал в исследуемом устройстве.

Рассмотренная сеть из одного нейрона позволяет моделировать устройства, у которых выходной сигнал линейно зависит от входного, и отклонения от этой зависимости связаны лишь с ошибками измерения.

АПРОКСИМАЦИЯ НЕЛИНЕЙНОЙ ЗАВИСИМОСТИ

Более сложная задача – аппроксимация нелинейной зависимости $y = F(x)$.

Увеличив количество нейронов, можно построить сеть для отображения нелинейной зависимости. Показанная на рис. 3 схема является частным случаем так называемого трехслойного персептрона – ИНС с тремя слоями [2].

Первый слой – входной – состоит из одного элемента, сигнал с которого без изменения поступает на входы элементов второго – скрытого – слоя. Каждый элемент второго слоя N_i ($i = 1, 2, \dots, K$) характеризуется своей функцией активации $f_i(x)$, определяющей сигнал на выходе этого элемента. Сигналы с выходов элементов второго слоя передаются с весовыми коэффициентами $\omega_1, \omega_2, \dots, \omega_K$ на вход элемента третьего слоя, состоящего из одного нейрона, на выходе которого появляется сигнал вида

$$y = \sum_{i=1}^K \omega_i \cdot f_i(x).$$

В качестве функций активации можно использовать функцию Гаусса

$$f(x) = \exp\left[-\frac{(x-c)^2}{2 \cdot \sigma^2}\right], \quad (3)$$

где c – точка на оси абсцисс, в которой значение функции достигает максимума, σ – параметр, характеризующий ширину колокола кривой Гаусса. Точки $x = \pm\sigma$ являются точками перегиба графика функции (3).

На рис. 4 показано семейство из 11 функций активации на отрезке $[0, 1]$.

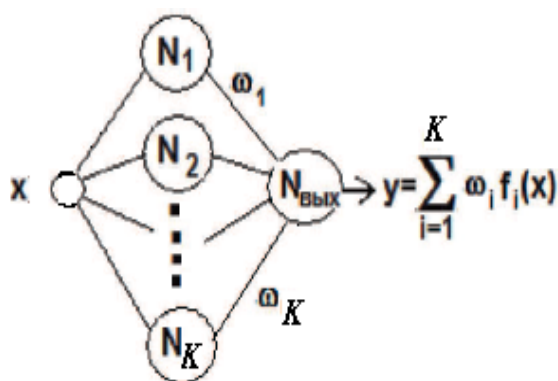


Рис. 3. Схема трехслойного перцептрона
Fig. 3. Scheme of three-layer perceptron

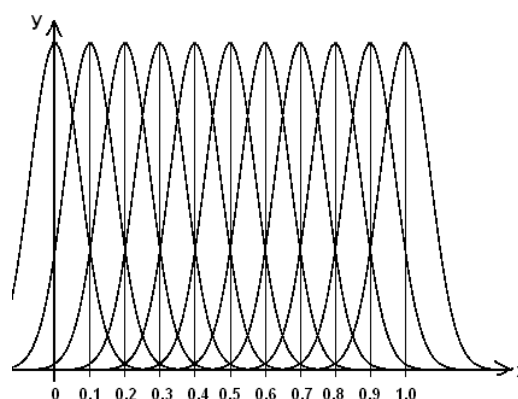


Рис. 4. Семейство гауссовых кривых при $\sigma = 0,05$
Fig. 4. A family of Gaussian curves with $\sigma = 0.05$

Функция (3) относится к радиально-базисным функциям, а ИНС с такой функцией активации скрытых нейронов – к классу радиально-базисных нейронных сетей (RBF-сети, см., например, [3, 4]).

Преимущество RBF-сетей по сравнению с традиционными сетями с сигмоидальными функциями активации (логистическими, линейными, квадратичными и др.) состоит в том, что их обучение (определение синаптических весов, минимизирующих ошибку аппроксимации) не требует сложных итерационных алгоритмов типа метода обратного распространения ошибки, а выполняется с помощью прямых методов (например, метода наименьших квадратов, градиентных методов и т. д.).

Пусть задан обучающий набор из M точек: $(x_i, y_i), i = 1, 2, \dots, M$.

Настройка сети состоит в определении значений весовых коэффициентов $\omega_1, \omega_2, \dots, \omega_M$ так, чтобы получить наилучшее соответствие между сигналами на выходе и значениями из обучающего набора. В качестве меры соответствия можно использовать тот же критерий, что и в приведенном выше примере – сумму квадратов разностей (ошибок аппроксимации).

$$S = \sum_{i=1}^M \left(\sum_{j=1}^K \omega_j \cdot f_j(x_i) - y_i \right)^2. \quad (4)$$

Условия минимума целевой функции S как функции весовых коэффициентов – равенство нулю ее частных производных:

$$\frac{\partial S}{\partial \omega_p} = 2 \cdot \sum_{i=1}^M \left(\sum_{j=1}^K \omega_j \cdot f_j(x_i) - y_i \right) \cdot f_p(x_i) = 0, \quad p = 1, 2, \dots, K. \quad (5)$$

Получившуюся систему линейных уравнений можно записать в матричной форме

$$CW = B, \quad (6)$$

где $W = (\omega_1, \omega_2, \dots, \omega_K)^T$ – искомый вектор весовых коэффициентов, C – числовая матрица размером $K \times K$, B – вектор-столбец правых частей:

$$C_{pq} = \sum_{i=1}^M f_p(x_i) \cdot f_q(x_i), \quad B_p = \sum_{i=1}^M y_i \cdot f_p(x_i). \quad (7)$$

$p, q = 1, 2, \dots, K.$

Основная проблема при практическом применении RBF-сети для нелинейной аппроксимации таблично заданных функций состоит в правильном выборе количества нейронов в скрытом слое и значения параметра σ в функции активации.

Очевидно, что увеличение числа нейронов скрытого слоя делает систему более гибкой в том смысле, что ИНС становится способной аппроксимировать функции с сильно выраженной нелинейностью, например, когда имеется несколько точек перегиба. Увеличение числа нейронов, однако, приводит к росту стоимости устройства в случае его аппаратной реализации, поэтому стараются ограничиться минимальным их количеством, но достаточным для достижения требуемой точности аппроксимации [3].

Относительно параметра σ можно предположить, что чем он меньше, тем лучше будут описываться локальные свойства аппроксимируемой зависимости. Однако при слишком малых значениях σ может появиться эффект, схожий с тем, который возникает при аппроксимации небольшого количества экспериментальных точек многочленом высокой степени: аппроксимирующая кривая проходит через все точки (ошибка аппроксимации равна нулю), но при этом сильно осциллирует, как показано на рис. 5. Количество нейронов $K = 11$. Кривая 1 получена при $\sigma = 0,04$, кривая 2 – при $\sigma = 0,1$.

Итак, при проектировании аппроксимирующей ИНС, разработчик сталкивается с проблемой выбора структуры сети (количество слоев и количество нейронов в каждом слое), а также параметров функций активации.

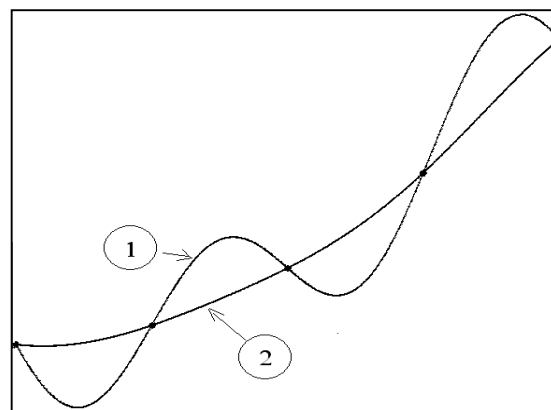


Рис. 5. Аппроксимация отрезка параболы по пяти точкам
Fig. 5. Approximation of a parabola cut at five points

Ниже рассматривается один из возможных подходов к решению этой проблемы.

Будем исходить из того, что имеется априорная информация о характере зависимости выходного сигнала от входного (экспоненциальная, логарифмическая, синусоидальная и т. д.).

Для того, чтобы сузить диапазон возможных допустимых значений искомых параметров сети (количество нейронов, параметры функций активации), предлагается использовать тестовые обучающие наборы, полученные для конкретных функциональных зависимостей.

Предположим, что есть основание полагать, что зависимость выходного сигнала некоторого устройства от сигнала на его входе имеет вид $y = F(x)$, $0 < x < 1$. Множество точек $(x_i, F(x_i))$, $i = 1, 2, \dots, M$, будем использовать как обучающий набор для ИНС, схема которой показана на рис. 3. Скрытые нейроны имеют функции активации типа (3), равномерно распределенные на отрезке $(0,1)$, как показано на рис. 4.

Целью численных экспериментов является выявление влияния на качество аппроксимации количества скрытых нейронов n в сети и параметра σ в функциях активации $f(x_i)$.

Качество аппроксимации $S(K, \sigma)$ будем оценивать интегрально по всему интервалу возможных значений аргумента:

$$S(K, \sigma) = \int_0^1 \left[F(x) - \sum_{i=1}^K \omega_i \cdot f_i(x) \right]^2 dx, \quad (8)$$

где ω_i – весовые коэффициенты функций активации, найденные в результате обучения сети на заданном обучающем наборе значений входных и выходных сигналов.

Численный эксперимент заключался в том, что при заданном обучающем наборе и фиксированном значении числа нейронов n определялось то значение σ , при котором величина интегрального отклонения минимально:

$$S(K, \sigma) \xrightarrow{\sigma} \min \quad (9)$$

Начальное значение σ выбирается так, чтобы погрешность аппроксимации $S(K, \sigma)$ оказалась достаточно большой. Величина σ затем уменьшается с некоторым шагом до тех пор, пока не будет достигнут минимум погрешности.

В качестве примера рассмотрим задачу аппроксимации затухающей синусоиды

$$y = \exp(-3x) \cdot \sin(2\pi \cdot x), \quad 0 \leq x, y \leq 1. \quad (10)$$

Для аппроксимации используется RBF-сеть со скрытым слоем из 11 нейронов. Обучающим набором являются координаты семи точек, лежащих на синусоиде, как показано на рис. 6. Оптимальное значение σ , обеспечивающее выполнение условия (9), равно 0,05. Кружками обозначены точки обучающей выборки.

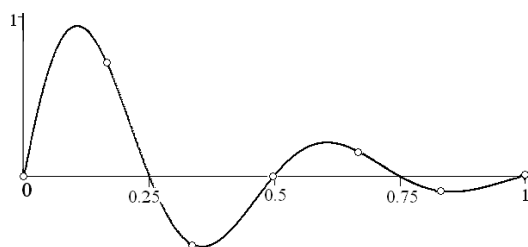


Рис. 6. График функции
Fig. 6. The Graph of function

Внесем теперь в координаты точек обучающей выборки случайные отклонения с помощью датчика случайных чисел:

$$y_i = y_i + \varepsilon_i, \quad \varepsilon_i = h \cdot (1 - \text{RND}), \quad i = 1, 2, \dots, 7,$$

где RND – случайное число из интервала $(0,1)$, полученное с помощью генератора случайных чисел, h – коэффициент, определяющий максимально возможное отклонение координат от их стационарных значений.

Такое изменение обучающей выборки имитирует погрешности измерений значений входных и выходных сигналов устройства, для которого разрабатывается ИНС.

Результаты численных экспериментов показаны на рис. 7.

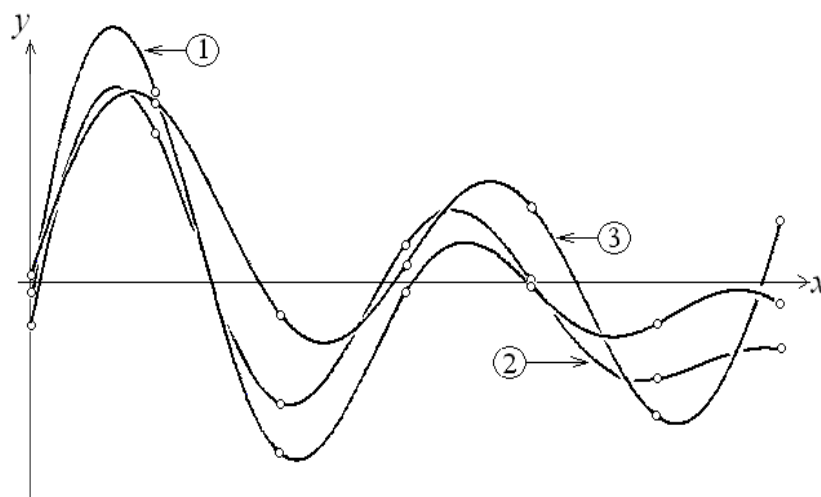


Рис. 7. Графики зависимости сигнала на выходе нейронной сети от входного сигнала для разных обучающих наборов со случайными отклонениями $\varepsilon \in (0, h)$: 1) $h = 0,2$, 2) $h = 0,4$, 3) $h = 0,6$

Fig. 7. The Graphs of the dependence of the signal y at the output of the neural network on the input signal for different training sets with random deviations $\varepsilon \in (0, h)$: 1) $h = 0.2$, 2) $h = 0.4$, 3) $h = 0.6$

Аппроксимирующие кривые получены при одном и том же значении $\sigma = 0,05$, полученном при обучении сети на тестовом примере (рис. 6).

Обращает на себя внимание тот факт, что все три кривые отражают характер зависимости выходного сигнала от входного, в соответствии с заданным эталоном (10).

МНОГОМЕРНАЯ АППРОКСИМАЦИЯ

Проблема аппроксимации функций многих переменных имеет давнюю историю. Основной вклад в решение этой проблемы внесли в середине прошлого века советские математики А.Н. Колмогоров и В.И. Арманд [6, 7]. В дальнейшем полученные ими результаты неоднократно уточнялись и развивались [8–11]. Одним из важных результатов в этой области является доказательство возможности аппроксимации произвольной непрерывной функции нейронной сетью с одним скрытым слоем (см., например, [12]).

Рассмотрим RBF-сеть, аналогичную представленной в предыдущем разделе (см. рис. 3). Отличия состоят в том, что теперь входным сигналом является вектор $X \in R^n$, а функция активации i -го скрытого нейрона имеет вид

$$f_i(X) = \exp\left[-\frac{\|X - C^{(i)}\|^2}{2 \cdot \sigma^2}\right],$$

где $C^{(i)} = (c_1^{(i)}, c_2^{(i)}, \dots, c_n^{(i)})$ – центр активации i -го нейрона, $\|\bullet\|$ – евклидова норма (длина вектора), σ – параметр функции активации, характеризующий размер области, в которой каждый нейрон скрытой области реагирует максимальным образом на входной сигнал X .

В двумерном случае $X = (x_1, x_2)$, $C^{(i)} = (c_1^{(i)}, c_2^{(i)})$.

Предположим, что аппроксимируемая зависимость $Z = Z(X)$ задана обучающим набором $(X^{(i)}, Z_i)$, $i = 1, 2, \dots, M$. Пусть K – количество скрытых нейронов, каждый из которых характеризуется своим центром активации $C^{(i)} = (c_1^{(i)}, c_2^{(i)})$, $i, j = 1, 2, \dots, K$. Синаптический вес связи скрытого нейрона с центром активации $C^{(i)}$ с выходным, обозначим ω_i . В результате обучения аппроксимирующая зависимость представляется функцией

$$F(X, \sigma) = \sum_{i=1}^K \omega_i \cdot f_i(X). \quad (11)$$

Как и в рассмотренном выше одномерном случае, ошибка аппроксимации S представляется в виде суммы квадратов разностей между заданными значениями Z_i и аппроксимирующей функции $F(X^{(i)}, \sigma)$ по всем точкам обучающего набора $X^{(i)}$, $i = 1, 2, \dots, M$:

$$S = \sum_{i=1}^M \left[F(X^{(i)}, \sigma) - Z_i \right]^2.$$

Вектор синаптических весов $W = (\omega_1, \omega_2, \dots, \omega_K)^T$ определяется в соответствии со стандартным методом наименьших квадратов путем решения системы линейных уравнений $GW = B$, где G – числовая матрица с коэффициентами G_{pq} , ($p, q = 1, 2, \dots, K$), B – столбец правых частей:

$$G_{pq} = \sum_{i=1}^M f_p(X^{(i)}) \cdot f_q(X^{(i)}), \quad B_p = \sum_{i=1}^M y_i \cdot f_p(X^{(i)}), \\ p, q = 1, 2, \dots, K.$$

Подбор оптимального значения параметра σ для заданного количества нейронов скрытого слоя выполняется так же, как в рассмотренном выше случае одномерной аппроксимации. Для этого используется обучающий набор значений некоторой эталонной функции, выбираемой исходя из априорных сведений о характере исследуемой зависимости $Z = Z(X)$.

В качестве примера рассмотрим задачу аппроксимации функции двух переменных, заданной таблично в узлах координатной сетки с шагом $h = 0,2$ (см. табл. 1 и рис. 8).

Таблица 1
Table 1

Значения функции $Z(x_i, y_j)$
The values of the function $Z(x_i, y_j)$

x\y	0,0	0,2	0,4	0,6	0,8	1,0
0,0	0,003	0,033	0,165	0,368	0,368	0,165
0,2	0,008	0,080	0,373	0,820	0,819	0,368
0,4	0,034	0,210	0,503	0,846	0,820	0,368
0,6	0,138	0,704	0,836	0,503	0,373	0,165
0,8	0,136	0,677	0,704	0,210	0,08	0,033
1,0	0,027	0,136	0,138	0,034	0,008	0,003

Количество нейронов в скрытом слое $K = 9$. Их центры активации расположены равномерно внутри квадрата $0 \leq x, y \leq 1$.

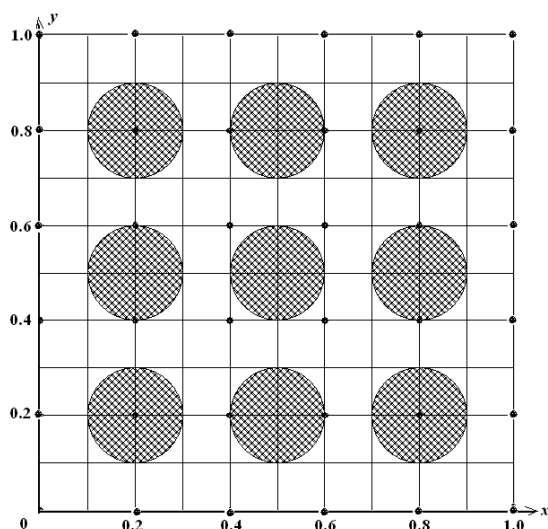


Рис. 8. Схема расположения центров активации скрытых нейронов
Fig. 8. The scheme of location of activation centers of hidden neurons

Результат обучения данной ИНС представлен на рис. 9. Функция аппроксимации (11) представлена линиями уровня.

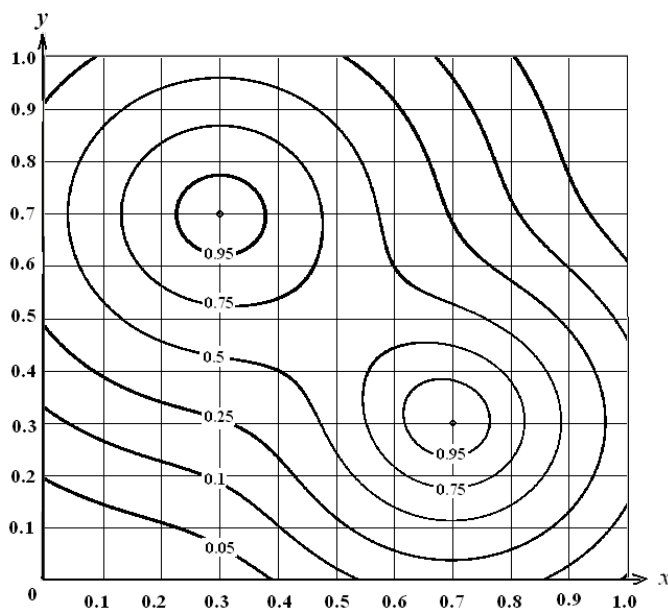


Рис. 9. Линии уровня аппроксимирующей функции для данных из таблицы 1
Fig. 9. Level lines of approximating function for data from table 1

ВЫВОДЫ

Для решения задачи аппроксимации, как отмечается во многих публикациях, лучше всего подходят радиально-базисные нейронные сети (RBF-сети) с одним скрытым слоем. В качестве функций активации скрытых нейронов используются радиально-базисные функции, к которым принадлежит и функция Гаусса.

В этом случае отдельный скрытый нейрон реагирует максимальным образом на входные сигналы из небольшой окрестности центра активации этого нейрона.

Такая организация сети позволяет проводить процесс обучения прямыми методами, без использования рекуррентных алгоритмов, таких как метод обратного распространения ошибки.

Этим RBF-сети в лучшую сторону отличаются от традиционных сетей с сигмоидальными функциями активации.

Одними из основных при построении RBF-сети являются задачи выбора точек центров активации для каждого нейрона и размеров их окрестностей.

В данной статье свойства сети определяются параметрами функции Гаусса – точкой максимума c_i и шириной кривой σ . Эти параметры характеризуют свойства сети и задаются заранее. Они не участвуют в процессе обучения, направленном на определение синоптических весов, минимизирующих ошибку аппроксимации.

Таким образом, процесс обучения RBF-сети может быть описан следующим образом.

1. Исходя из общих соображений и априорных сведений о характере зависимостей между входными и выходными сигналами, выбирается эталонная функция $f(X)$, $X \in R^n$.
2. Составляется обучающая выборка $\{X^{(i)}, Z_i = f(X^{(i)})\}$, $i = 1, \dots, M$.
3. Задается начальное значение параметра функции активации σ .
4. Определяются весовые коэффициенты для каждого нейрона скрытого слоя ω_i , $i = 1, \dots, K$.
5. Подсчитывается интегральная ошибка аппроксимации (8).
6. Параметр σ уменьшается и пункты 4, 5 повторяются до тех пор, пока ошибка аппроксимации станет минимальной.

Целью статьи было продемонстрировать один из возможных подходов к решению задачи аппроксимации с помощью нейронной сети. Конечно, вопрос о качестве аппроксимации при этом остается открытым. Как поведет себя сеть при аппроксимации других функций, заданных таблично, – это предмет дальнейших исследований.

СПИСОК ЛИТЕРАТУРЫ

1. **Калиткин Н.Н.** Численные методы. СПб., 2011. 592 с.
2. **Каллан Р.** Основные концепции нейронных сетей: пер. с англ. М., 2003. 288 с.
3. **Хайкин С.** Нейронные сети. Полный курс: пер. с англ. 2-е изд. М., 2006. 1104 с.
4. **Круглов В.В., Борисов В.В.** Искусственные нейронные сети. Теория и практика. М., 2002. 382 с.
5. **Осовский С.** Нейронные сети для обработки информации: пер. с пол. М., 2002. 344 с.
6. **Колмогоров А.Н.** О представлении непрерывных функций нескольких переменных в виде суперпозиции непрерывных функций одного переменного // ДАН СССР. 1957. Т. 114, № 5. С. 953–956.
7. **Арнольд В.И.** О представлении непрерывных функций нескольких переменных в виде суперпозиции непрерывных функций меньшего числа переменных // Математическое просвещение. 1958. № 3. С. 41–61.
8. **Круглов В.В., Дли М.И., Голунов Р.Ю.** Нечеткая логика и искусственные нейронные сети. М.: Физматлит, 2001. 224 с.
9. **Поспелов В.В.** О приближении функций нескольких переменных произведениями функций одного переменного: препринт № 32. М.: ИПМ АН СССР, 1978. 72 с.
10. Нейросетевые системы управления / В.А. Терехов, Д.В. Ефимов, И.Ю. Тюкин, В.И. Антонов. СПб.: Изд-во СПбГУ, 1999. 265 с.
11. **Шура-Бура М.Р.** Аппроксимация функций многих переменных функциями, каждая из которых зависит от одного переменного // Вычислительная математика. 1957. Вып. 27. С. 3–19.
12. **Бутырский Е.Ю., Кувалдин И.А., Чалкин В.П.** Аппроксимация многомерных функций // Научное приборостроение. 2010. Т. 20, № 2. С. 82–92.

СВЕДЕНИЯ ОБ АВТОРЕ

Агеев Владимир Николаевич, доктор технических наук, профессор, профессор кафедры прикладной математики МГТУ ГА, rv3bd@mail.ru.

THE SOLUTION OF THE APPROXIMATION PROBLEM OF NONLINEAR DEPENDANCES USING ARTIFICIAL NEURAL NETWORKS

Vladimir N. Ageyev¹

¹Moscow State Technical University of Civil Aviation, Moscow, Russia

ABSTRACT

The paper discusses issues connected with the use of an artificial neural network (ANN) to approximate the experimental data. One of the problems in the development of the ANN is the choice of an appropriate activation function for neurons of the hidden layer and adjusting the parameters of the function in the learning process of the network. The article discusses a three-layer perceptron with one hidden layer, each neuron of which has the activation function in the form of a Gaussian curve. The choice of radial basis activation function allows the use of the direct method of determining the weight coefficients – method of least squares in the process of network training. Thus the quality of the approximation depends on the correct choice of the value parameter of the activation function, which in this case is the width of the Gaussian bell curve. In practice, this parameter is determined by conducting numerical experiments. This is a rather time-taking process. In this paper we propose to define the value of this parameter by the training set, representing the coordinates of the test curve points set with the desired properties. These properties are based on the a priori data of the approximated functions (linear, quadratic, logarithmic, exponential relationship). Because the test curve is given in explicit form, the parameter of activation function is determined from the condition of reaching the minimum of the integral from the squared difference between the values of the test functions and the output of the network. This approach guarantees obtaining the approximating curve with good properties, in particular, it is characterized by the absence of so-called "oscillations" – many inflection points in its graph.

Key words: artificial neural network, activation function, training samples, function approximation.

REFERENCES

1. Kalitkin N.N. *Chislenniye metody* [Numerical methods]. Saint-Petersburg, 2011. 592 p. (in Russian)
2. Kallan R. *Osnovniye kontseptsii neyronnyh setey* [Basic concepts of neural networks]. Translated from English. M.: Williams, 2003, 288 pp. (in Russian)
3. Haykin S. *Neyronniye seti. Polniy kurs* [Neural networks. A complete course]. Translated from English. 2nd edition. 2006. 1104 p. (in Russian)
4. Kruglov V.V., Borisov V.V. *Iskusstvenniye neyronniye seti. Teoria i praktika* [Artificial neural network. Theory and practice]. M: 2002. 382 pp. (in Russian)
5. Osovskiy S. *Neyronniye seti dlya obrobтки informatsii* [Neural network for information processing]. Translated from polish. M.: Finance and Statistics, 2002, 344 p. (in Russian)
6. Kolmogorov A.N. *O predstavlenii nepreryvnykh funktsiy neskol'kih peremennykh v vide superpozitsii nepreryvnykh funktsiy odnogo peremennogo* [On the representation of continuous functions of several variables as a superposition of continuous functions of one variable]. Proceedings of the USSR Academy of Sciences, 1957. vol. 114, No 5, pp. 953–956. (in Russian)
7. Arnold V.I. *O predstavlenii nepreryvnykh funktsiy neskol'kih peremennykh v vide superpozitsii nepreryvnykh funktsiy men'shego chisla peremennykh* [On the representation of continuous functions of several variables as superposition of continuous functions of fewer variables] // *Matematicheskoye prosveshchenie* [Mathematical education], 1958, No. 3, pp. 41–61. (in Russian)

8. Kruglov V.V., Dlee M.I., Golunov R.Yu. *Nechetkaya logika i iskusstvenniy intellekt* [Fuzzy logic and artificial neural network]. M.: Fizmatlit, 2001, 224 p. (in Russian)

9. Pospelov V.V. *O priblizhenii funktsiy neskolkih peremennykh proizvedeniyami funktsiy odnogo peremennogo: preprint №32* [On the approximation of several variables functions by products of functions of one variable: preprint. No. 32]. M.: Keldysh Institute of Applied Mathematics, 1978, 72 p. (in Russian)

10. Terekhov V.A., Efimov D.V., Tyukin I.Yu., Antonov V.I. *Neyrosetevye sistemy upravleniya* [Neural network control systems]. Saint-Petersburg: SPb. State University, 1999, 265 p. (in Russian)

11. Shura-Bura M.R. *Aproksimatsiya funktsiy mnogih peremennykh funktsiyami. Kazhdaya is kotorykh zavisit ot odnogo peremennogo* [Approximation of functions of many variables functions, each of which depends on one variable]. *Vychislitel'naya matematika* [Computational mathematics] – 1957, issue 27, pp. 3–19. (in Russian)

12. Butirsky E.Yu., Kuvaldin I.A., Chalkin V.P. *Aproksimatsiya mnogomernykh funktsiy* [Approximation of multidimensional functions]. *Nauchnoye priborostroyeniye* [Scientific instrumentation]. 2010, vol. 20, № 2, 82–92 pp. (in Russian)

INFORMATION ABOUT THE AUTHOR

Vladimir N. Ageyev, Doctor of Technical Sciences, Professor of Applied Mathematics Chair, Moscow State Technical University of Civil Aviation, rv3bd@mail.ru.

Поступила в редакцию
Принята в печать

28.10.2017
14.03.2018

Received
Accepted for publication

28.10.2017
14.03.2018