

УДК 681.5

А. М. ВУЛЬФИН, В. М. ГИНИЯТУЛЛИН**ГЛАДКАЯ КУСОЧНО-КВАДРАТИЧНАЯ АППРОКСИМАЦИЯ
В КОМПЛЕКСНОЗНАЧНОМ НЕЙРОСЕТЕВОМ БАЗИСЕ**

В работе обсуждается алгоритм гладкой кусочной-квадратичной аппроксимации функций одного переменного. Элементарные функции аппроксимации генерируются комплекснозначными нейронами (эллиптические звенья) и паракомплекснозначными нейронами (гиперболические звенья). *Аппроксимация ; гладкая аппроксимация ; нейрон ; комплексный вектор ; паракомплексный вектор*

В теории аппроксимаций рассматриваются различные постановки задачи обработки данных. В случае если известен некоторый набор функций, то задача заключается в том, чтобы подобрать для него соответствующие коэффициенты. С другой стороны, можно поставить задачу о поиске тех функций, которыми удобнее всего аппроксимировать имеющиеся в наличии опытные данные. Аппроксимация экспериментальных данных с минимизацией суммы невязок согласно первому подходу [1] требует задания формы зависимости – модели исследуемого процесса. Подобрать адекватную модель бывает достаточно трудно. На практике в таких случаях часто используется сплайн-аппроксимация, в частности кубический сплайн [2]. Использование сплайнов гарантирует непрерывность первой (второй) производной на всей области определения имеющегося набора экспериментальных данных.

1. СОСТОЯНИЕ ВОПРОСА

В работе [3] рассмотрен метод кусочно-эллиптической аппроксимации, его недостатком можно считать разрыв первой производной в точках сшивки соседних эллиптических звеньев.

2. ПОСТАНОВКА ЗАДАЧИ

В данном исследовании предлагается использовать алгоритмы нейросетевой аппроксимации, порождающие непрерывные кусочные модели функций одного аргумента.

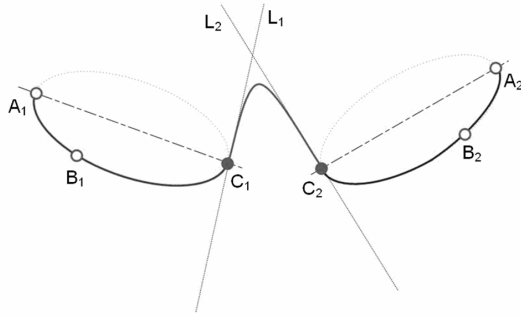
3. МЕТОДИКА ИССЛЕДОВАНИЯ**Гиперболические сшивки эллиптических звеньев аппроксимирующей модели**

Элементарным звеном кусочно-эллиптической аппроксимации является одна из половин эллипса. Известно, что в вершинах эллипса касательная перпендикулярна к его диаметру. Для того чтобы обеспечить непрерывность первой производной используется гиперболическая сшивка соседних эллиптических звеньев. Основанием для такого метода служит работа [4], в которой показано наличие бесконечно большого количества мнимых дополнительных гипербол вокруг любого эллипса.

В качестве гладкой сшивки можно использовать одну из ветвей гиперболы, построенной следующим образом: в вершинах соседних эллипсов строятся перпендикуляры к диаметрам, точка их пересечения полагается центром гиперболы, а перпендикуляры – ее асимптотами (рис. 1). Остается подобрать такие параметры гиперболы, чтобы в окрестности точки сшивки расстояние от ветви гиперболы до ее асимптоты не превышало заданной погрешности.

Однако построение гладкой сшивки одной гиперболой не всегда возможно. Как видно из рис. 2, для всех четырех возможных случаев взаимного расположения эллиптических участков гладкую сшивку можно построить с помощью двух гипербол. Для этого необходимо построить вспомогательную прямую, которая будет общей асимптотой соседних гипербол. В данной работе предполагается, что гиперболическая сшивка строится в промежутке между экспериментальными точками, т. е., фактически, решается задача интерполяции. Поскольку никаких дополнительных ограничений на качество интерполяции не накладывается, то месторас-

положение центров гипербол (точки O_1 и O_2 на рис. 2) выбираются следующим образом. В зависимости от расположения половин эллипсов (выше/ниже диаметра эллипса) выбираются точки O_1 и O_2 на перпендикулярах к диаметрам эллипсов (ниже/выше вершины эллипса). Расстояние от точек O_1 и O_2 до соответствующих вершин эллипсов равно $\alpha \cdot d$, где d – расстояние между вершинами соседних эллипсов.



$A_1B_1C_1$ – опорные точки первого эллиптического звена;
 $A_2B_2C_2$ – опорные точки второго эллиптического звена;
 L_1, L_2 – асимптоты гиперболы;
 C_1, C_2 – точки сшивки эллиптических и гиперболических звеньев

Рис. 1. Сшивка эллиптических звеньев одной гиперболой

Для совпадения веток соседних гипербол значения одной из полуосей, каждой гиперболы, устанавливаются одинаковыми и равными $\beta \cdot d$.

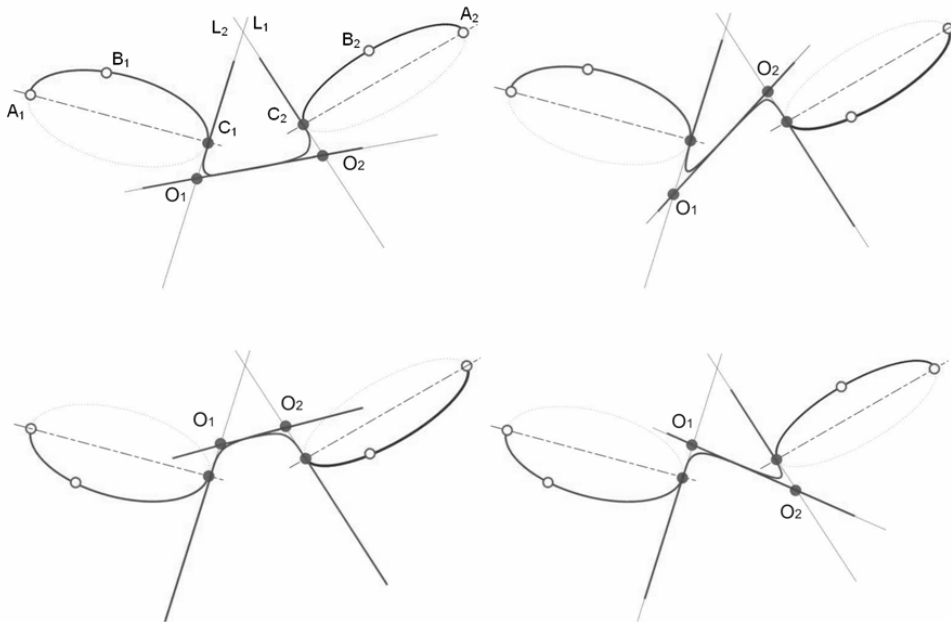
Таким образом, о каждой из гипербол известны: уравнения асимптот (и, как следствие, координаты центра), значения одной полуоси и квадрант, в котором находится нужная ветвь гиперболы. Остается определить неизвестные характеристики гипербол.

Паракомплекснозначный персептрон

Существуют три двухкомпонентные числовые системы (алгебры) [5]:

- комплексные числа (квадрат модуля есть сумма квадратов компонент);
- паракомплексные числа (квадрат модуля есть разность квадратов компонент);
- дуальные числа (квадрат модуля есть квадрат действительной части).

Последние две алгебры содержат делители нуля [5]. Тем не менее, при условии отсутствия в алгоритме операции деления, использование этих алгебр вполне оправдано (в данной работе используются паракомплексные числа).



$A_1B_1C_1$ – опорные точки первого эллиптического звена;
 $A_2B_2C_2$ – опорные точки второго эллиптического звена;
 L_1, L_2 – асимптоты гиперболы;
 C_1, C_2 – точки сшивки эллиптических и гиперболических звеньев;
 O_1, O_2 – вершины гиперболических звеньев;
 O_1O_2 – общая асимптота гиперболы

Рис. 2. Варианты расположения эллиптических звеньев и гиперболических сшивок

Рассмотрим нейрон паракомплекснозначного персептрона, имеющий три входа и паракомплекснозначный вектор весов, модель которого представлена на рис. 3.

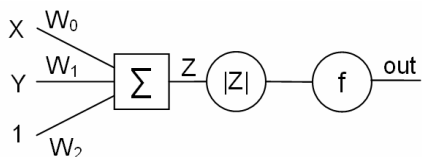


Рис. 3. Модель нейрона паракомплекснозначного персептрона

Результатом скалярного произведения вещественнозначного входного вектора $\{x, y, 1\}$ и паракомплексного вектора весов $\{c_0+id_0, c_1+id_1, c_2+id_2\}$ является паракомплексное число z :

$$\begin{pmatrix} x \\ y \\ 1 \end{pmatrix}^T (c_0 + id_0 \quad c_1 + id_1 \quad c_2 + id_2)^T = z = (xc_0 + yc_1 + c_2) + i(xd_0 + yd_1 + d_2). \quad (1)$$

Квадрат модуля этого числа является аргументом пороговой функции активации:

$$\begin{aligned} \text{abs}^2(z) &= \text{Re}^2(z) - \text{Im}^2(z), \\ \text{out} &= f(\text{abs}^2(z)). \end{aligned} \quad (2)$$

При моделировании двузначных булевых функции [6] с использованием однослойного паракомплекснозначного персептрона (линейная функция активации нейрона), обученного путем полного перебора, было установлено, что сеть порождает функцию двух переменных, линии уровня которой проецируются на плоскость XOY в виде семейства сопряженных гипербол (рис. 4).

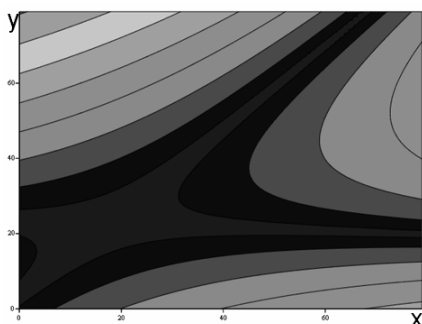


Рис. 4. Линии уровня, порождаемые нейроном с паракомплексным вектором весов

Изолинии на рис. 4 соответствуют разным значениям квадрата модуля, на асимптотах гипербол квадрат модуля равен нулю. В соседних

квадрантах знаки квадрата модуля противоположны, в противоположных квадратах – одинаковы.

Если в качестве значения порога функции активации выбрать единицу, то можно отказаться от операции извлечения корня. В процессе обучения паракомплекснозначный персептрон оперирует гиперболической кривой с требуемыми параметрами из семейства гипербол. Такая гипербола является проекцией линии уровня поверхности, значение функции на которой равно единице, на плоскость XOY.

Расчет параметров гиперболической кривой

Рассмотрим метод определения коэффициентов уравнения, описывающего гиперболическую кривую, и установим их связь с компонентами вектора весов. Приравняем квадрат модуля выхода нейрона к единице и приведем полученное выражение к виду:

$$\text{out} = \text{Re}^2(z) - \text{Im}^2(z) = 1, \quad (3)$$

$$\begin{aligned} \text{Re}^2(z) &= (c_0x + c_1y + c_2)^2 = \\ &= c_0^2x^2 + 2c_0c_1xy + 2c_0c_2x + c_1^2y^2 + 2c_1c_2y + c_2^2, \end{aligned}$$

$$\begin{aligned} \text{Im}^2(z) &= (d_0x + d_1y + d_2)^2 = \\ &= d_0^2x^2 + 2d_0d_1xy + 2d_0d_2x + d_1^2y^2 + 2d_1d_2y + d_2^2. \end{aligned}$$

Воспользуемся общим уравнением кривой второго порядка на плоскости XOY в виде [7]:

$$\begin{aligned} a_{11}x^2 + 2a_{12}xy + a_{22}y^2 + \\ + 2a_{13}x + 2a_{23}y + a_{33} = 0. \end{aligned} \quad (4)$$

Сопоставляя коэффициенты уравнений (3) и (4), можно установить соотношение между коэффициентами уравнения, описывающего семейство эллипсов и компонентами вектора весов нейрона:

$$\begin{aligned} a_{11} &= c_0^2 - d_0^2, \\ a_{12} &= c_0c_1 - d_0d_1, \\ a_{22} &= c_1^2 - d_1^2, \\ a_{13} &= c_0c_2 - d_0d_2, \\ a_{23} &= c_1c_2 - d_1d_2, \\ a_{33} &= c_2^2 - d_2^2 - 1. \end{aligned} \quad (5)$$

Необходимо определить коэффициенты выражения (5) для каждой гиперболы. Воспользуемся каноническим уравнением гиперболы [7]:

$$\frac{x^2}{m^2} - \frac{y^2}{n^2} = 1, \quad (6)$$

где m и n – величины полуосей гиперболы.

Такое уравнение описывает гиперболу с центром в начале координат. Чтобы перейти от общего уравнения гиперболы к ее каноническому виду, необходимо осуществить поворот и перенос текущей системы координат. Координаты каждой точки необходимо подвергнуть преобразованию согласно следующим формулам [8]:

$$\begin{aligned} x' &= x \cos(\alpha) + y \sin(\alpha) - x_0, \\ y' &= -x \sin(\alpha) + y \cos(\alpha) - y_0, \end{aligned} \quad (7)$$

где α – угол поворота осей координат – определяется как угол между положительным направлением оси OX базовой системы координат и положительным направлением оси OX смещенной и повернутой системы [8]; (x_0, y_0) – точку, в которую осуществляется параллельный перенос базовой системы координат.

Подставим в (6) формулы перехода к исходным координатам (7):

$$\begin{aligned} &\frac{(x \cos(\alpha) + y \sin(\alpha) - x_0)^2}{m^2} - \\ & - \frac{(-x \sin(\alpha) + y \cos(\alpha) - y_0)^2}{n^2} = 1. \end{aligned} \quad (8)$$

Сопоставляя результат преобразования и соотношение (4), найдем выражение, позволяющее определить коэффициенты общего уравнения кривой второго порядка посредством базовых характеристик гиперболы – длины полуосей, угла поворота относительно базовой системы координат и координат центра гиперболы:

$$\begin{aligned} a_{11} &= \frac{1}{m^2} \cos^2(\alpha) - \frac{1}{n^2} \sin^2(\alpha), \\ a_{12} &= -\frac{1}{m^2} \sin(\alpha) \cos(\alpha) - \frac{1}{n^2} \sin(\alpha) \cos(\alpha) \\ a_{22} &= \frac{1}{m^2} \sin^2(\alpha) - \frac{1}{n^2} \cos^2(\alpha), \\ a_{13} &= -\frac{x_0}{m^2} \cos^2(\alpha) + \frac{y_0}{n^2} \sin^2(\alpha), \\ a_{23} &= -\frac{x_0}{m^2} \sin^2(\alpha) + \frac{y_0}{n^2} \cos^2(\alpha), \\ a_{33} &= \frac{x_0^2}{m^2} - \frac{y_0^2}{n^2} - 1. \end{aligned} \quad (9)$$

Далее выразим компоненты вектора весов через характеристики эллипса, сопоставив системы (5) и (9):

$$\begin{aligned} c_0 &= \frac{1}{m} \cos(\alpha), & d_0 &= \frac{1}{n} \sin(\alpha), \\ c_1 &= -\frac{1}{m} \sin(\alpha) & d_1 &= \frac{1}{n} \cos(\alpha), \quad (10) \\ c_2 &= -\frac{x_0}{m}, & d_2 &= -\frac{y_0}{n}, \end{aligned}$$

где m и n – величины полуосей гиперболы; (x_0, y_0) – координаты центра исходной гиперболы; α – угол поворота системы координат.

Таким образом, имея геометрические характеристики произвольной гиперболы, можно получить соответствующие ей веса паракомплекснозначного нейрона.

Схема работы аппроксиматора

Схематическое представление работы обученного аппроксиматора приведено на рис. 5.

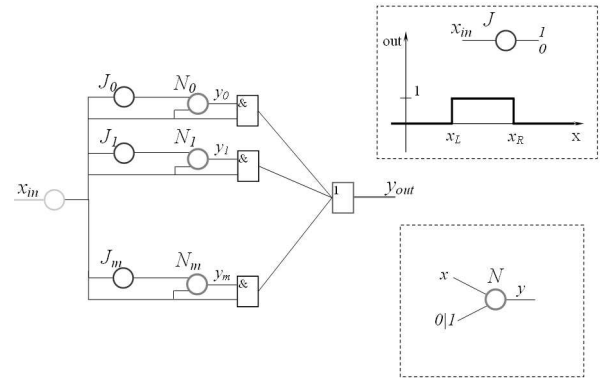


Рис. 5. Схема работы обученной сети

Входной элемент подает значение аргумента аппроксимируемой функции на каждый J элемент, который принимает решение о принадлежности данного сигнала интервалу активности соответствующего нейрона N . Используя параметрическое преобразование гиперкомплексного вектора весов [3], элемент N вычисляет значение функции (если значение второго параметра равно 1, в противном случае выход элемента равен нулю), а выходной элемент, действующий по принципу логической функции ИЛИ, формирует выход всей сети.

4. ОБСУЖДЕНИЕ РЕЗУЛЬТАТОВ

Гладкая кусочно-эллиптическая аппроксимация экспериментальных данных

В качестве примера была рассмотрена функция одного переменного, определяемая уравнением:

$$y(x) = -0.2x + 0.4 \sin(2x) + e^{0.25x}, \quad (11)$$

и табулированная на интервале $[-10; 10]$ с шагом 0,5. Аппроксимация выполнена с ошибкой равной 0,05. Ошибка аппроксимации в рассматриваемом случае определяется согласно следующей методике [3]. На этапе поиска максимально возможной покрывающей полосы (определение понятия покрывающей полосы дано в [3]) под ошибкой понимается величина Δ , определяющая максимальное евклидово расстояние от каждой точки, попавшей в полосу, до элементарного звена аппроксимации на этом участке. Таким образом, отклонение экспериментальных точек на каждом участке аппроксимирующей модели не превышает заданной величины. На рис. 6 приведен график функции (11) и график аппроксимирующей функции, а в табл. 1 дано сравнение с кусочно-эллиптическим методом и кубическим сплайном. Количество звеньев каждого типа определяет количество соответствующих нейронов N .

Таблица 1
Сравнение гладкой кусочно-эллиптической, кусочно-эллиптической аппроксимации и интерполяции кубическим сплайном

Аппроксимация	Кол-во элементарных звеньев		Кол-во точек	Ошибка
	Эллиптических	Гиперболических		
Гладкая кусочно-эллиптическая	14	13	41	0,05
	17			
Кусочно-эллиптическая	17		41	0,05
Кубический сплайн	40		41	–

Полученная аппроксимация функции (11), как и сама исходная функция, представлены на рис. 6. Графики функций практически совпадают. Построим аппроксимирующую модель предложенной таблично заданной функции (11) с помощью полносвязной нейронной сети прямого распространения – многослойного персептрона.

В результате серии экспериментов по подбору оптимальной архитектуры и метода обучения многослойного персептрона были выбраны параметры сети:

- один вход и один выход;
- один скрытый слой с 20 нейронами; функция активации гиперболический тангенс;
- выходной слой из одного нейрона с линейной функцией активации;
- обучение согласно алгоритму обратного распространения ошибки;
- ошибка обучения 0,01.

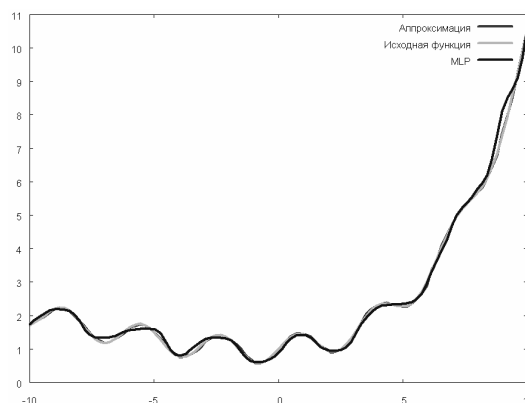


Рис. 6. Гладкая кусочно-эллиптическая аппроксимация

Результаты аппроксимации исходного набора данных приведены на рис. 6. Для того чтобы оценить качественные характеристики различных аппроксимирующих моделей, табулируем функцию (11) на интервале $[-10; 10]$ с шагом, равным половине исходного, т. е., 0,25. На рис. 7 приведены коэффициенты детерминации R^2 и графики линейной регрессии исходной функции и аппроксимирующей модели.

Среднеквадратичное отклонение на тестовом наборе данных для многослойного персептрона составило 0,019, а максимальное отклонение 0,552. Для гладкой кусочно-эллиптической аппроксимации 0,003 и 0,010 соответственно.

Таким образом, гладкая кусочно-эллиптическая аппроксимация может быть использована для построения моделей некоторого класса функций, в том случае, если требуется непрерывность первой производной модели. Построенная модель содержит дополнительную информацию, которая может быть из нее извлечена – для каждого элементарного звена можно получить его аналитическое выражение и выражение, определяющее первую производную. Многослойный персептрон может быть успешно применен для аппроксимации наборов данных, имеющих сравнительно большой объем (несколько десятков примеров в обучающем множестве), к тому же, обучение и тестирование такой сети представляет собой нетривиальную задачу (проблема переобучения и т. п.), как по сложности применяемых методов, так и по затраченному времени.

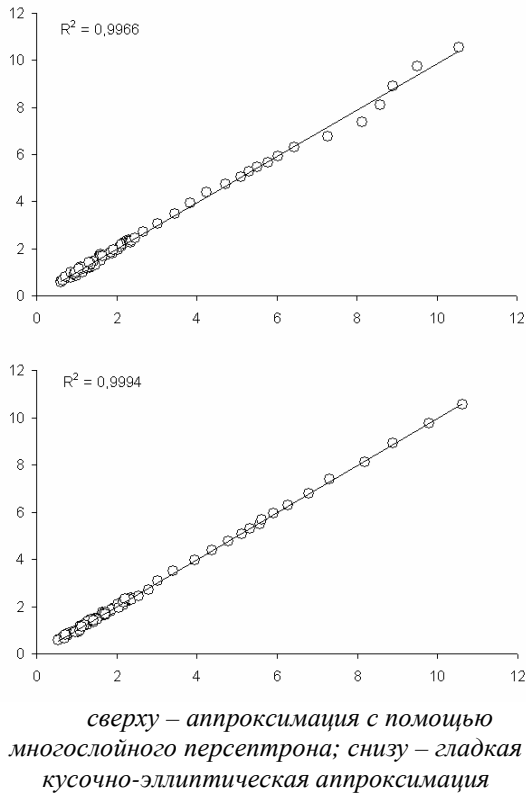


Рис. 7. Показатели качества аппроксимирующих моделей

Аппроксимация экспериментальных данных вискозиметрических измерений нефти месторождения Кюрсангя. Применим разработанные методы для аппроксимации экспериментальной зависимости напряжения сдвига τ от скорости γ сдвига при течении жидкости в капилляре. Исходные данные взяты из [9] и приведены в табл. 2.

Таблица 2
Данные вискозиметрических измерений месторождения Кюрсангя

γ, c^{-1}	3,0	5,4	9,0	26,2	27,0	48,6	81,0	145,8
$\tau, 10^{-1} \text{Па}$	12,2	16,7	22,8	31,5	43,2	56,3	85,4	121,5

Применим методы кусочно-эллиптической и гладкой кусочно-эллиптической аппроксимации – рис. 8.

Из рисунков видно, что применение кусочных методов для текущего набора экспериментальных данных предпочтительнее интерполяции кубическим сплайном в силу меньших осцилляций результирующей модели и сокращения числа элементарных участков. Использование многослойного перцептрона для построения этой аппроксимирующей модели невозможно ввиду малого количества обучающего множества.

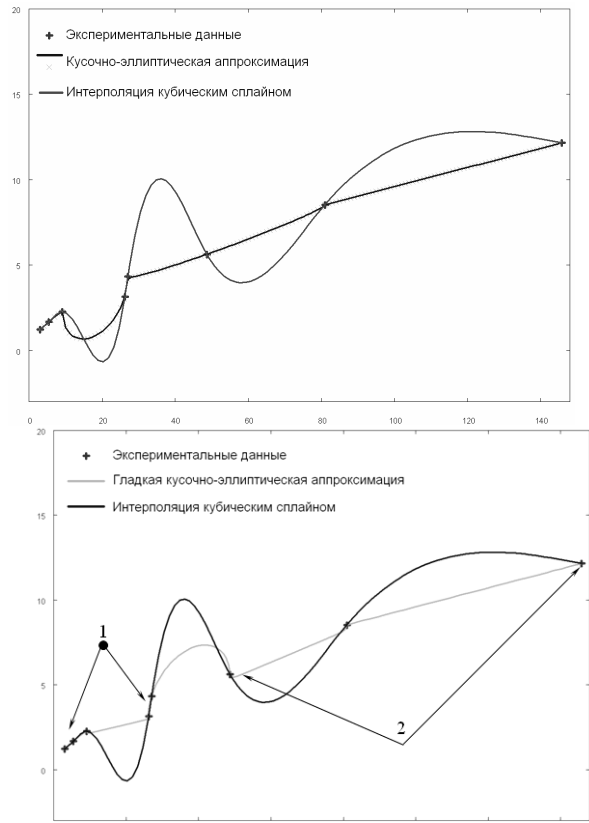


Рис. 8. Аппроксимация экспериментальных данных

В случае гладкой кусочно-эллиптической аппроксимации практически полностью сохраняется монотонность исходного набора экспериментальных точек. Например, на интервале точек 1–5 рис. 8, кубический сплайн как возрастает, так и убывает, в случае же гладкой кусочной аппроксимации сохраняется возрастающая тенденция экспериментального ряда. Аналогичное поведение аппроксимирующих кривых наблюдается и на интервале точек 6–8.

ВЫВОДЫ

1. Разработанный алгоритм формирует гладкую кусочно-квадратичную аппроксимацию функции одного переменного с помощью комплекснозначаной нейронной сети, и применим для моделирования небольших объемов данных по сравнению с многослойными сетями прямого распространения.

2. По сравнению с кубическим сплайном сокращается количество элементарных звеньев и достигается аналогичная точность.

СПИСОК ЛИТЕРАТУРЫ

1. **Кори, Г.** Справочник по математике для научных работников и инженеров / Г. Кори, Т. Кори. СПб. : Лань, 2003. 832 с.
2. **Завьялов, Ю. С.** Методы сплайн-функций / Ю. С. Завьялов, Б. И. Квасов. М. : Наука, 1980. 362 с.
3. **Вульфин, А. М.** Кусочно-эллиптическая аппроксимация в комплекснозначном нейросетевом базисе / А. М. Вульфин, В. М. Гиниятуллин // Научная сессия МИФИ-2009. XI Всероссийск. науч.-техн. конф. «Нейроинформатика-2009» : Сб. науч. тр. Ч. 2. М. : МИФИ, 2009. С. 75–84.
4. **Гирш, А. Г.** Теория и практика мнимых образов в геометрии. МАИ Прикладная геометрия / А. Г. Гирш. М., 2005. 206 с.
5. **Кантор, И. Л.** Гиперкомплексные числа / И. Л. Кантор, А. С. Солодовников. М. : Наука, 1973. 144 с.
6. **Гиниятуллин, В. М.** Моделирование логических функций в нейросетевом базисе / В. М. Гиниятуллин // Нефтегазовое дело. 2008. Т. 6, № 1. С. 35–43.
7. **Ефимов, Н. В.** Квадратичные формы и матрицы / Н. В. Ефимов. М. : Наука, 1975. 160 с.
8. **Письменный, Д. Т.** Конспект лекций по высшей математике. Ч. 1 / Д. Т. Письменный. М. : Айрис-пресс, 2003. 288 с.
9. **Хисамутдинов, Н. И.** Разработка нефтяных месторождений. Т.1: Разработка нефтяных месторо-

ждений на поздней стадии / Н. И. Хисамутдинов, М. М. Хасанов, А. Г. Телин. М. : ВНИИОЭНГ, 1994. 240 с.

ОБ АВТОРАХ



Вульфин Алексей Михайлович, асп. Дипл. инж. по спец. «Программное обеспечение вычислительной техники» (УГНТУ, 2008). Иссл. в обл. нейросетевых технологий.



Гиниятуллин Вахит Мансурович, доц. Дипл. инж. по спец. «Химическая технология топлив и углеродных материалов» (УГНТУ, 1992). Канд. техн. наук по спец. «Химическая технология топлива» (УГНТУ, 1998). Иссл. в обл. искусственных нейронных сетей и параллельного программирования.