

ИССЛЕДОВАНИЕ АППРОКСИМАТИВНЫХ ВОЗМОЖНОСТЕЙ РАДИАЛЬНО-БАЗИСНОЙ СЕТИ С РАЗЛИЧНЫМИ ФУНКЦИЯМИ АКТИВАЦИИ

© 2013 С.А.Прохоров, И.А.Лёзин, И.В. Лезина, С.К. Латыш, С.А. Саиян

Самарский государственный аэрокосмический университет имени академика С.П. Королёва
(национальный исследовательский университет)

Поступила в редакцию 07.06.2013

В статье описана разработанная авторами автоматизированная система для аппроксимации плотности вероятности случайных процессов. В системе используется радиально базисная сеть (RBF), в которой в качестве функции активации нейронов помимо функции Гаусса используются ядерные функции. В статье анализируются результаты аппроксимации, полученные сетями с различными функциями активации нейронов, и делается вывод о целесообразности замены функции Гаусса ядерными функциями.

Ключевые слова: автоматизированная система, нейронная сеть, радиально базисная сеть, аппроксимация, случайный процесс, функция Гаусса, ядерные функции.

Задача аппроксимации [1] является актуальной математической задачей. Для моделирования различных процессов, в частности экономических или физических, необходимо иметь функциональную зависимость, описывающую исследуемый процесс. На практике, для получения такой зависимости необходимо решить задачу аппроксимации, позволяющую на основании имеющихся данных построить функцию, описывающую исследуемый процесс. Для решения задачи аппроксимации существует множество математических методов, однако их использование в каждом случае сугубо индивидуально и не всегда возможно. В данной статье рассматривается использование нейронных RBF-сетей для аппроксимации плотности распределения вероятности случайного процесса, которую можно представить в виде [2] (1):

$$\hat{f}(x) = \sum_{k=0}^m w_k \cdot \varphi_k(x, \alpha_1, \dots, \alpha_n), \quad (1)$$

где w_k – веса, настраиваемые в процессе обучения сети, φ_k – функция активации нейронов, $\alpha_1, \dots, \alpha_n$ – параметры функции активации.

RBF-сети используют в качестве функции

Прохоров Сергей Антонович, доктор технических наук, профессор, заведующий кафедрой информационных систем и технологий. E-mail: sp@smr.ru

Лёзин Илья Александрович, кандидат технических наук, доцент кафедры информационных систем и технологий. E-mail: ilyozin@yandex.ru

Лёзина Ирина Викторовна, кандидат технических наук, доцент кафедры информационных систем и технологий. E-mail: chuchuck@yandex.ru

Латыш Сергей Константинович, студент четвёртого курса факультета информатики. E-mail: avatar135@yandex.ru
Саиян Сергей Андраникович, студент четвёртого курса факультета информатики. E-mail: saiyana@haulmont.com

активации нейронов радиально-базисные функции. Общий вид радиально-базисной функции можно описать уравнением:

$$\varphi_k(x) = e^{-(x_k - c_k)^2 / 2\sigma_k^2}, \quad (2)$$

Чаще всего в качестве функции активации нейронов радиально базисной сети используется функция Гаусса (2). В данной статье в качестве функции активации помимо функции Гаусса выбраны две ядерные функции [3]: квадратичная ядерная функция (3) и ядро Епанечникова (4).

$$\frac{15}{16} \cdot \left(1 - \frac{x - c}{2\sigma^2}\right)^2, \quad (3)$$

$$0,75 \cdot \left(1 - \frac{x - c}{2\sigma^2}\right). \quad (4)$$

В исследовании в качестве случайного процесса использовался набор выборок случайных величин с заданным законом распределения. Выборки генерировались при помощи метода обратной функции. В качестве входных данных для нейронной сети использовались значения высот столбцов гистограммы.

В качестве алгоритмов обучения был использован метод обратного распространения ошибки [4], а также алгоритм K-means [5]. Данный алгоритм относится к классу градиентных алгоритмов, т.е. изменения весов связей производятся в направлении минимизации градиента ошибки. Ошибка при обучении равна разности сигнала на выходе сети и эталонного значения выхода, соответствующего входным данным.

Обучение по алгоритму K-means заключается в размещении центров радиально базисных функций только в тех точках пространства, где имеются информативные данные, и нахождении

таких значений коэффициента сигма, чтобы области охвата всех радиальных функций покрывали все пространство входных данных, причем любые две зоны могут перекрываться лишь в незначительной степени.

Критерием остановки обучения является достижение заданной погрешности. Результаты аппроксимации можно оценить по формуле:

$$\sigma = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - y_{\phi_i})^2}, \quad (5)$$

где y_i – значение функции, полученное сетью, y_{ϕ_i} – эталонное значение функции.

Для решения поставленных задач было разработано программное обеспечение, моделирующее работу нейронной сети. Внешний вид программы представлен на рис. 1.

В исследовании использовались различные законы распределения случайной величины, ниже рассмотрены результаты для нормального, равномерного закона и закона Симпсона. На рисунках представлено сравнение выходного сигнала сети с эталонным значением.

Закон распределения: закон Симпсона. Параметры: $a=1$, $b=4$. Количество нейронов скрытого слоя: 10. Метод обучения: метод обратного распро-

странения ошибки. Коэффициент обучения: 0,2.

В табл.1 представлены значения среднеквадратического отклонения выходного сигнала сети от эталонного значения для каждого исследования. Значение слева от черты соответствует методу обратного распространения ошибки, значение справа – алгоритму K-means. Число нейронов скрытого слоя в эксперименте равно десяти. Анализируя таблицу, можно отметить, что выбор функции Гаусса в качестве функции активации сети является целесообразным для нормального закона распределения и закона Симпсона. Для равномерного закона распределения лучшим выбором будут ядерные функции.

В таблице 2 представлена зависимость значения среднеквадратического отклонения от числа нейронов скрытого слоя для случайной величины, распределенной по закону Симпсона. Значение слева от черты соответствует методу обратного распространения ошибки, значение справа – алгоритму K-means.

Как видно из таблицы, вариация числа нейронов в скрытом слое не оказывает серьезное влияние на аппроксимативные возможности сети. Из таблицы следует, что для определения функциональной зависимости оптимально выбрать пять-десять нейронов в скрытом слое.

В табл. 3 представлена зависимость значения

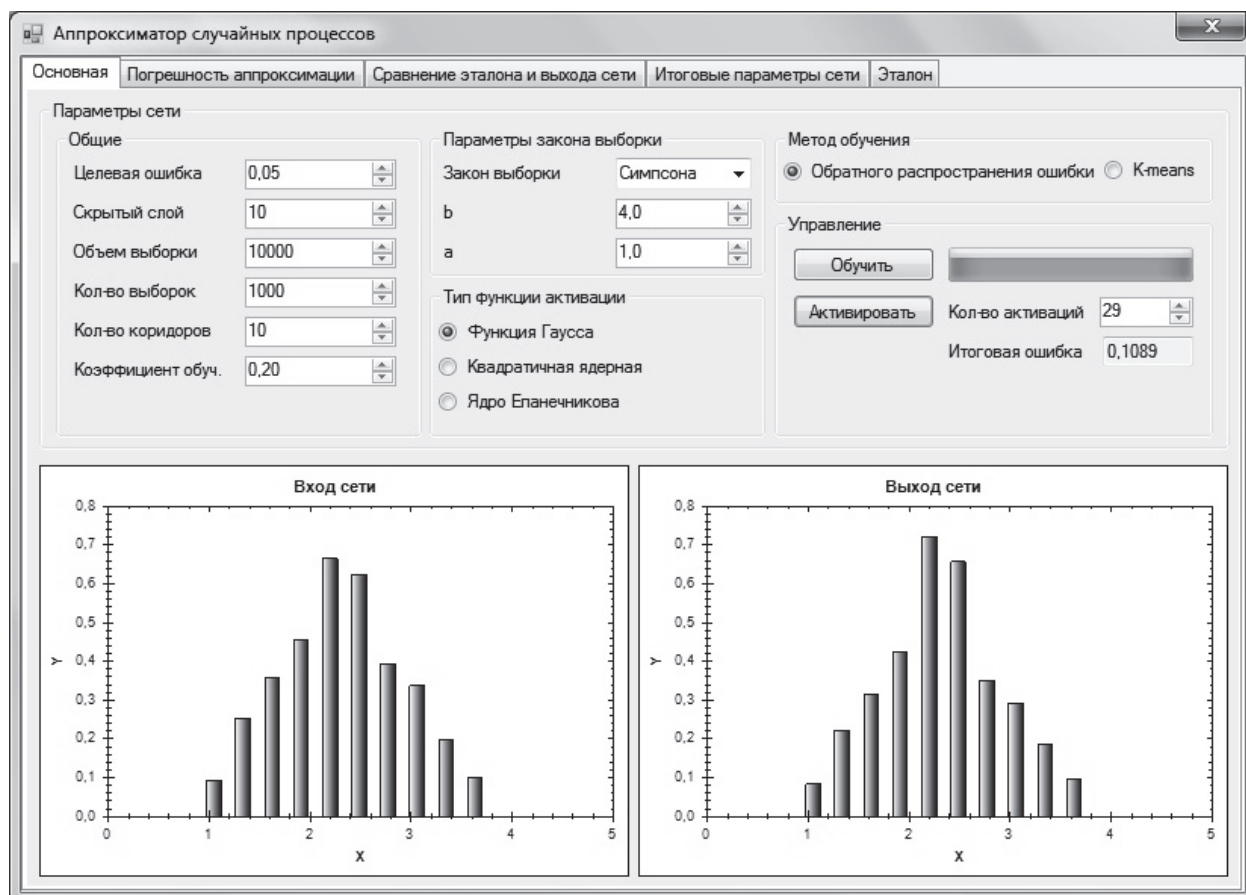


Рис. 1. Интерфейс программы

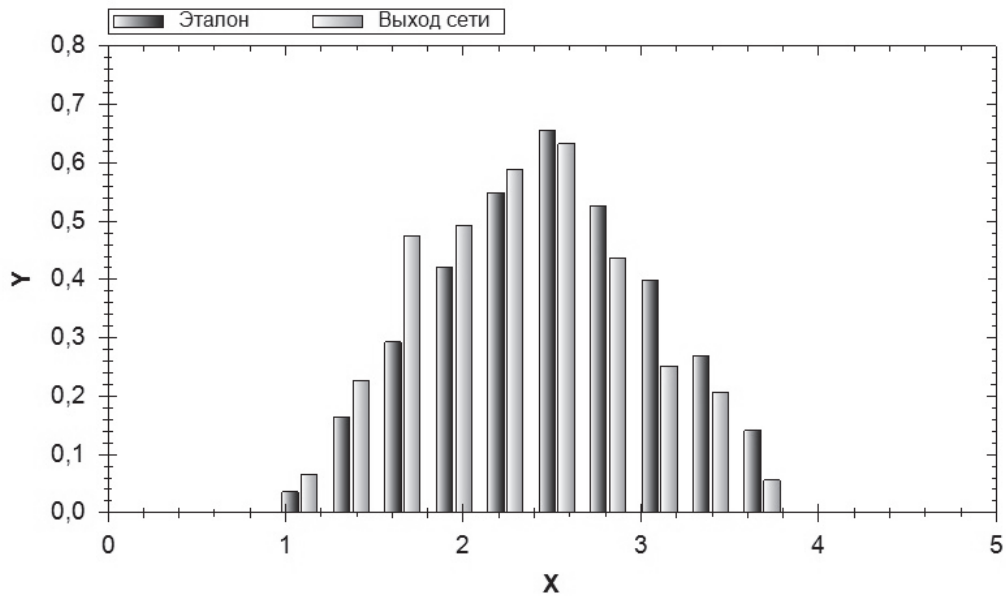


Рис. 2. Использование функции Гаусса

Таблица 1. Значения СКО для различных законов распределения.

	Нормальный	Равномерный	Симпсона
Функция Гаусса	0,024 / 0,013	0,005 / 0,003	0,033 / 0,005
Квадратичная ядерная	0,039 / 0,015	0,001 / 0,001	0,017 / 0,005
Ядро Епанечникова	0,027 / 0,015	0,001 / 0,001	0,018 / 0,007

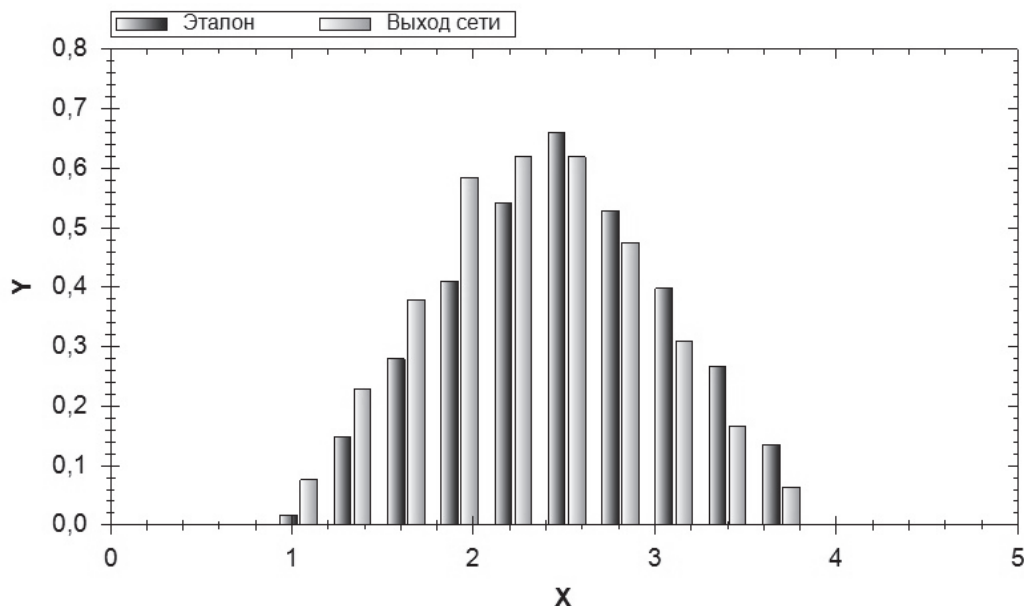


Рис. 3. Использование ядра Епанечникова

среднеквадратического отклонения от коэффициента обучения. На практике от коэффициента обучения зависит сходимость алгоритма, слишком большой коэффициент может привести к неустойчивости и расхождению. Анализ таблицы показывает, что при маленьком коэффициенте

обучения алгоритм может не успеть сойтись при заданном числе шагов. Оптимальным значением коэффициента обучения является 0,2.

Исследования показали, что использование ядерных функций активации дает лучшие результаты для равномерного закона распределения и

Таблица 2. Значения СКО для различного числа нейронов скрытого слоя.

	5	10	20	30
Функция Гаусса	0,019 / 0,008	0,033 / 0,005	0,012 / 0,007	0,012 / 0,006
Квадратичная ядерная	0,039 / 0,007	0,017 / 0,005	0,011 / 0,004	0,016 / 0,004
Ядро Епанечникова	0,014 / 0,007	0,018 / 0,007	0,098 / 0,004	0,012 / 0,003

Таблица 3. Значения СКО для различного коэффициента обучения

	0,05	0,1	0,2	0,3
Функция Гаусса	0,024 / 0,007	0,015 / 0,007	0,033 / 0,005	0,054 / 0,006
Квадратичная ядерная	0,014 / 0,006	0,009 / 0,005	0,017 / 0,005	0,034 / 0,004
Ядро Епанечникова	0,022 / 0,009	0,012 / 0,006	0,018 / 0,007	0,013 / 0,006

закона Симпсона при обучении методом обратного распространения ошибки. Алгоритм K-means показывает идентичные результаты в обучении сети для различных функций активации, и в большинстве случаев дает лучшее значение СКО, чем метод обратного распространения ошибки. В общем случае, для получения наилучших результатов под конкретную задачу целесообразно подбирать оптимальную функцию активации.

СПИСОК ЛИТЕРАТУРЫ

1. Прохоров С.А. Аппроксимативный анализ случай-

ных процессов . Самара: СГАУ, 2001. 329 с.
 2. Исследование аппроксимативных возможностей радиально-базисной нейронной сети в зависимости от алгоритма обучения / И.А. Лёзин, И.В. Лёзина, Д.Д. Телегина. Цифровые модели в проектировании и производстве РЭС: межвузовский сборник научных трудов [под редакцией проф. Н.К.Юркова]. Пенза: изд-во ПГУ, 2012. Вып. 17.С. 279-289.
 3. Деврой Л., Дверфи Л. Непараметрическое оценивание плотности. L1 – подход [пер. с англ]. М.: Мир, 1988. 408 с.
 4. Хайкин С. Нейронные сети: полный курс .2-е издание.: [пер. с англ]. М.: Издательский дом «Вильямс», 2006. 1104 с.
 5. Осовский С. Нейронные сети для обработки информации [пер. с польского И.Д. Рудинского]. М.: Финансы и статистика, 2002. 344 с.

THE RESEARCH OF APPROXIMATIVE POSSIBILITIES OF RADIAL-BASIS NETWORKS WITH DIFFERENT ACTIVATION FUNCTIONS

© 2013 S.A. Prokhorov, I.A. Lyozin, I.V. Lyozina, S.K. Latysh, S.A. Saiyan

Samara State Aerospace University named after Academician S.P. Korolev
 (National Research University)

The article describes an automated system developed by the authors to approximate the probability density function of random processes. This system uses radial-basis network (RBF), which uses kernel functions as the activation of neurons in addition to the Gaussian functions. The article analyzes the results of approximation produced by networks with different activation functions of neurons and concludes an advisability of replacing the Gaussian functions by kernel functions.

Keywords: automated system, neural network, radial-basis network, approximation, random process, Gaussian function, kernel function.

Sergey Prokhorov, Doctor of Technics, Professor, Head at the Information Systems and Technology Department. E-mail: sp@smr.ru
 Ilya Lyozin, Candidate of Technics, Associate Professor at the Information Systems and Technology Department. E-mail: ilyozin@yandex.ru
 Irina Lyozina, Candidate of Technics, Associate Professor at the Information Systems and Technology Department. E-mail: chuchyck@yandex.ru
 Sergey Latysh, Fourth-Year Student of the Faculty of Informatics. E-mail: avatar135@yandex.ru
 Sergey Saiyan, Fourth-Year Student of the Faculty of Informatics. E-mail: saiyan@haulmont.com