СПИСОК ЛИТЕРАТУРЫ

1. Стоянов А.К. Нейронная сеть, основанная на точечных отображениях // Известия Томского политехнического университета. – 2008. – Т. 313. – № 5. – С. 96–101.
2. Малинецкий Г.Г. Математические основы синергетики. – М.: Изд-во ЛКИ, 2007. – 312 с.
3. Лоскутов А.Ю., Михайлов А.С. Введение в синергетику. – М.: Наука, 1990. – 270 с.
4. Мандель И.Д. Кластерный анализ. – М.: Финансы и статистика, 1988. – 176 с.
5. Калан Р. Основные концепции нейронных сетей. – М.: Издательский дом «Вильямс», 2001. – 290 с.
6. Kohonen T. The self-organizing map // Proc. of the IEEE. – 1990. – V. 78. – № 9. – P. 1464–1480.
7. Хайкин С. Нейронные сети: полный курс. Пер. с англ. – М.: Издательский дом «Вильямс», 2006. – 1104 с.
8. Миронов С. Ирисы Фишера [Электронный ресурс]. – режим доступа: http://www.delphikingdom.com/asp/viewitem.asp?catalogid=400. – 13.10.2009.

Поступила 13.10.2009 г.

УДК 004.8.032.26

МОДИФИЦИРОВАННЫЙ ГРАДИЕНТНЫЙ АЛГОРИТМ ОБУЧЕНИЯ РАДИАЛЬНО-БАЗИСНЫХ НЕЙРОННЫХ СЕТЕЙ

В.Н. Вичугов

Томский политехнический университет
E-mail: vlad@aics.ru

Приведена структура радиально-базисных нейронных сетей. Определены недостатки классического градиентного алгоритма обучения нейронных сетей в задачах идентификации объектов управления. Предложен модифицированный градиентный алгоритм обучения, позволяющий устранить недостатки классического. Показан пример применения модифицированного алгоритма в задаче аппроксимации двумерной функции.

***Ключевые слова**:*
Искусственная нейронная сеть, радиально-базисная нейронная сеть, алгоритм обучения, идентификация.
Key words:
Artificial neural network, radial-basis neural network, learning algorithm, identification.

При использовании нейросетевых методов в задачах автоматического управления часто возникает необходимость построения нейросетевой модели объекта управления на основе полученных входных и выходных сигналов в реальном времени. Использование многослойных перцептронов для построения нейросетевой модели является затруднительным в связи с тем, что дополнительное обучение многослойного перцептрона в некотором участке рабочей области приводит к потере обученного состояния во всей рабочей области нейронной сети, что не позволяет использовать этот тип нейронных сетей в задачах реального времени. Указанный недостаток отсутствует в радиально-базисных нейронных сетях (РБНС), т. к. каждый их элемент влияет на значение выходного сигнала преимущественно только в ограниченном участке рабочей области, который характеризуется положением центра элемента и параметром σ, называемым шириной радиальной функции. Чем больше значение параметра σ, тем больше размер области, на которую оказывает влияние данный элемент.

Структура РБНС показана на рис. 1 [1, 2]. Сеть состоит из двух слоев. Входные сигналы поступают на элементы первого слоя без изменений.

На рисунке использованы обозначения: n – количество элементов в первом слое; $x_1, x_2, ..., x_n$ – входные сигналы; m – количество элементов во втором слое; $c_{i1}, c_{i2}, ..., c_{in}$ – координаты центра i-го элемента; σ_i – ширина радиальной функции i-го элемента; θ_i – выходной сигнал i-го элемента; w_i – весовой коэффициент выходной связи i-го элемента; y – выходной сигнал РБНС.

Выходной сигнал каждого элемента определяется функцией Гаусса [3]

$$\theta_i = \exp\left(-\frac{\sum_{j=1}^{n}(x_j - c_{ij})^2}{2\sigma_i^2}\right).$$

Выходной сигнал РБНС вычисляется как взвешенная сумма сигналов элементов:

$$y = \sum_{i=1}^{m} w_i \cdot \theta_i.$$

Для обучения РБНС используется градиентный алгоритм, основанный на минимизации целевой функции ошибки сети. В соответствии с этим алгоритмом для каждого элемента вычисляются величины изменений весового коэффициента Δw_i, ши-

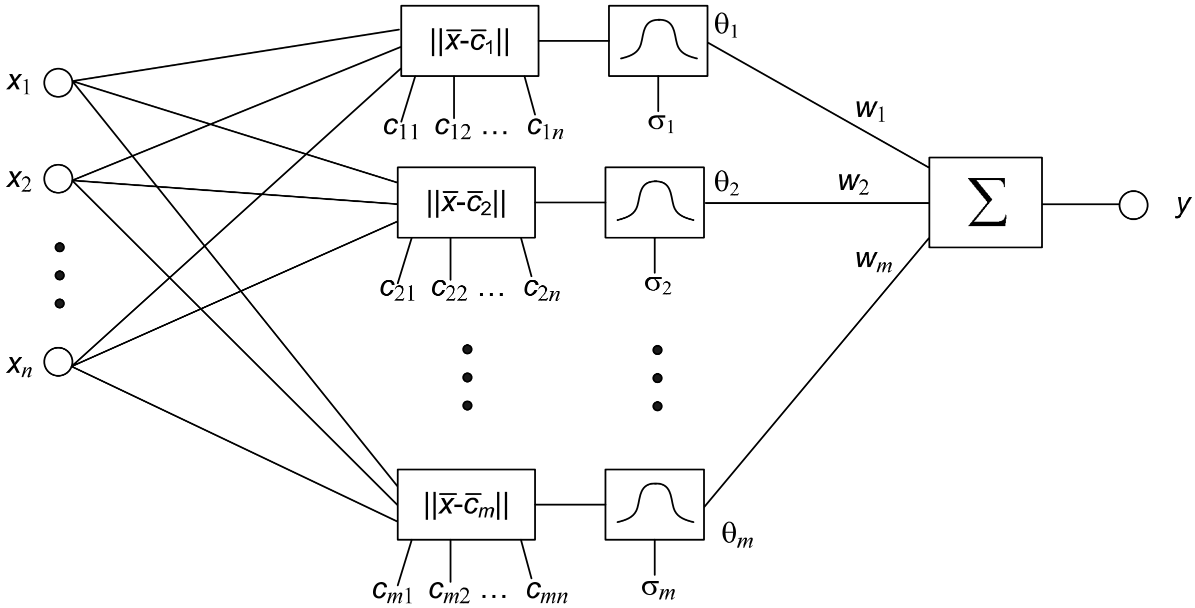

Рис. 1. *Структура радиально-базисной нейронной сети*

рины элемента $\Delta\sigma_i$ и координат центра элемента Δc_{ij}.

В результате проведенных экспериментов были выявлены некоторые недостатки классического градиентного алгоритма обучения РБНС:

1. В алгоритме обучения РБНС нет правил для первоначального задания количества элементов сети и их параметров, а также нет правил для изменения количества элементов в процессе обучения. Равномерное распределение элементов в рабочей области не всегда является оптимальным. Также может возникнуть ситуация, когда количество элементов, заданное первоначально, является недостаточным для достижения требуемого качества обучения.
2. В процессе обучения изменяются параметры всех элементов сети. В результате при увеличении количества элементов вычислительные затраты на обучение также увеличиваются.
3. РБНС не может достичь устойчивого состояния в процессе обучения в случаях, когда существуют элементы с близкими значениями параметров c_{ij} и σ_i. Появление подобных ситуаций во многом зависит от выбранного количества элементов и их начальных параметров. Причина ухудшения качества обучения заключается в том, что в градиентном алгоритме предполагается, что на выходное значение РБНС в каждой точке рабочей области в основном влияет только один элемент. При наличии нескольких элементов в одном участке рабочей области изменение их параметров в соответствии с градиентным алгоритмом не всегда приводит к уменьшению ошибки обучения.

С целью определения ситуаций, когда параметры некоторых элементов становятся близкими друг к другу, было введено понятие коэффициента взаимного пересечения элементов. Для вычисления этого коэффициента для некоторого элемента РБНС необходимо найти второй элемент, центр которого расположен ближе всего к центру рассматриваемого элемента. Значение коэффициента взаимного пересечения определяется как сумма выходной величины текущего элемента в центре второго элемента и выходной величины второго элемента в центре текущего элемента:

$$\rho_i = \exp\left(-\frac{\sum_{j=1}^{n}(c_{ij}-c_{dj})^2}{2\sigma_i^2}\right) + \exp\left(-\frac{\sum_{j=1}^{n}(c_{ij}-c_{dj})^2}{2\sigma_d^2}\right),$$

где i – номер элемента, для которого вычисляется значение коэффициента взаимного пересечения; d – номер элемента, центр которого расположен ближе всего к центру элемента с номером i. Номер элемента d определяется по формуле

$$d = \arg\min_k \sqrt{\sum_{j=1}^{n}(c_{ij}-c_{kj})^2}.$$

Значение коэффициента взаимного пересечения находится в интервале (0; 2]. Коэффициент принимает максимальное значение в том случае, когда центры рассматриваемых элементов совпадают. В ходе экспериментов по аппроксимации различных функций с помощью РБНС было определено, что ошибка РБНС начинает увеличиваться в случае, если значение коэффициента взаимного пересечения превышает 1,9. Поэтому для достижения максимального качества обучения РБНС необходимо ограничить максимальное значение коэффициента взаимного пересечения величиной 1,9.

С целью исключения недостатков классического градиентного алгоритма обучения РБНС был разработан модифицированный градиентный алго-

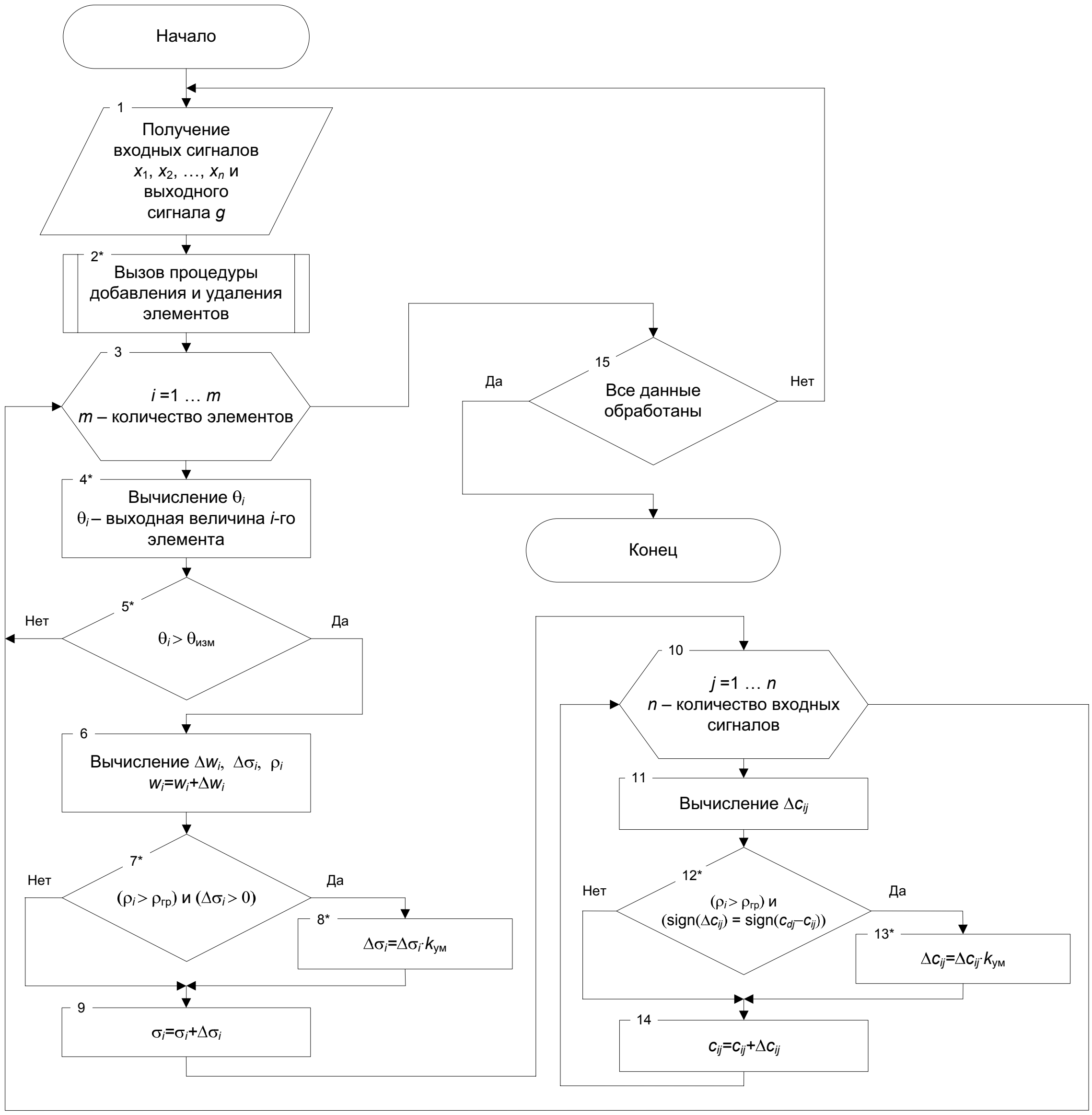

Рис. 2. *Блок-схема модифицированного градиентного алгоритма обучения РБНС*

ритм, блок-схема которого показана на рис. 2. Блоки, которые отсутствуют в классическом алгоритме, отмечены звездочками. Основные отличия от классического алгоритма заключаются в следующем:

1. Добавлены правила изменения структуры РБНС в процессе обучения (блок 2). В начале обучения РБНС не содержит элементов. По мере необходимости новые элементы добавляются, а неиспользуемые элементы удаляются.
2. Уменьшены вычислительные затраты, требуемые для каждого цикла обучения. Это достигается изменением параметров не всех элементов, как в классическом алгоритме, а только элементов, выходная величина которых в рассматриваемой точке больше величины $\theta_{изм}$ (блоки 4 и 5).
3. Исключена возможность возникновения ситуации, когда параметры некоторых элементов практически совпадают. Для этого вычисленные величины Δc_{ij} и $\Delta\sigma_i$ уменьшаются, если ко-

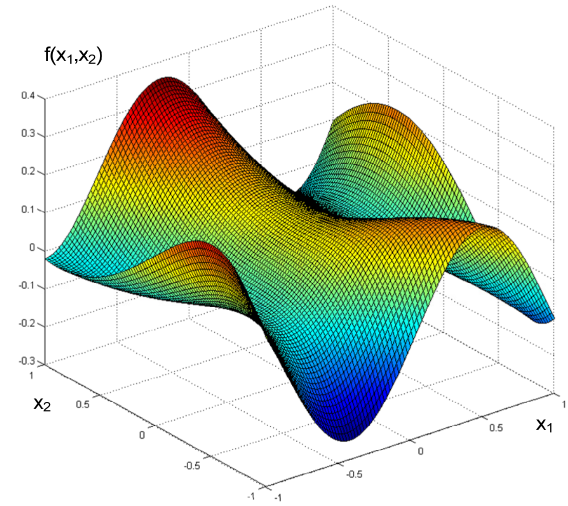

Рис. 3. *Поверхность функции $f(x_1,x_2)$*

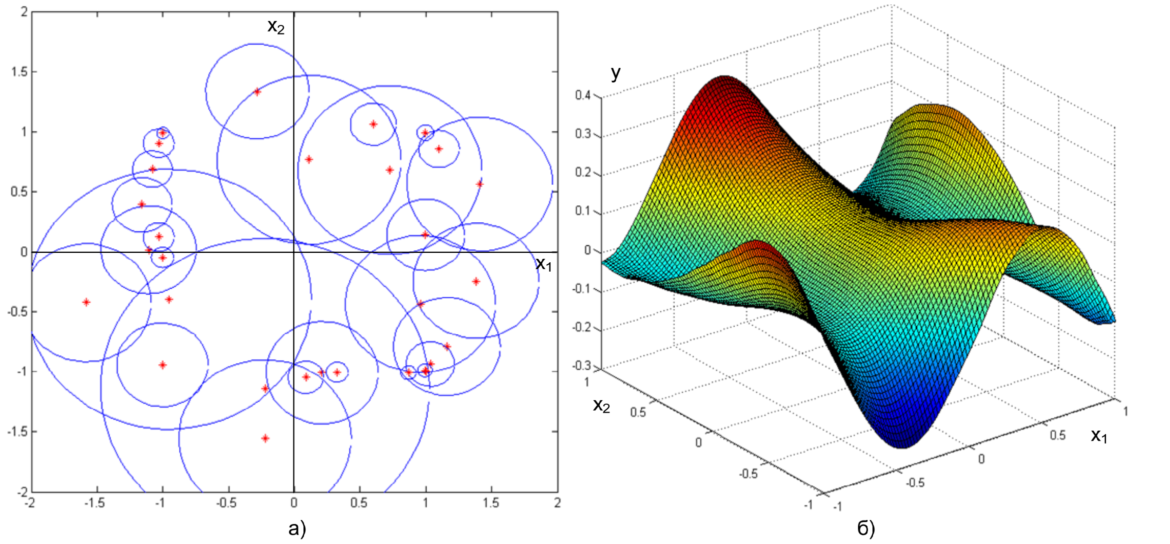

Рис. 4. *Результат аппроксимации функции $f(x_1,x_2)$: а) распределение элементов; б) поверхность, показывающая зависимость выхода РБНС от входных значений*

эффициент взаимного пересечения элементов превышает пороговую величину $\rho_{гр}$, равную 1,9 (блоки 7, 8, 12, 13).

Изменение структуры РБНС за счет добавления или удаления элементов приводит к изменению выходного значения РБНС только в окрестности центра добавляемого или удаляемого элемента, а не во всей рабочей области, как в случае с изменением структуры многослойного перцептрона. Поэтому добавление и удаление элементов РБНС возможно осуществлять в процессе обучения без необходимости запуска процесса обучения с самого начала.

Рассмотрим пример аппроксимации двумерной функции

$$f(x_1,x_2)=\sin\left(\frac{x_1^2}{2}-\frac{x_2^2}{4}+3\right)\cdot\cos(2x_1+1-\exp(-x_2))$$

на участке $x_1\in[-1;1]$, $x_2\in[-1;1]$ с помощью РБНС. Поверхность данной функции показана на рис. 3. При использовании классического градиентного алгоритма перед началом обучения была задана структура РБНС в виде 36 элементов с начальной шириной σ_0=0,2, равномерно распределенных в рабочей области. После приблизительно 10^6 циклов обучения среднеквадратическая ошибка обучения перестала уменьшаться и достигла значения $1{,}554\cdot10^{-3}$.

При использовании модифицированного градиентного алгоритма структура РБНС была определена автоматически в процессе обучения. После приблизительно трех миллионов циклов обучения количество элементов увеличилось до 30, а среднеквадратическая ошибка обучения составила $1{,}225\cdot10^{-3}$. Результаты обучения РБНС показаны на рис. 4.

Отсюда следует, что даже при меньшем количестве элементов модифицированный градиентный алгоритм позволяет достичь меньшей ошибки обучения по сравнению с классическим градиентным алгоритмом за счет динамического формирования структуры нейронной сети, но при этом требуется больше вычислительных ресурсов. Добавление новых элементов происходит только в те участки, которые характеризуются максимальной ошибкой аппроксимации, что приводит к уменьшению ошибки обучения при меньшем количестве элементов по сравнению с классическим алгоритмом обучения.

Выводы

1. На основе анализа классического градиентного алгоритма обучения радиально-базисных нейронных сетей разработан модифицированный алгоритм, позволяющий изменять структуру сети в процессе обучения.
2. Для исключения ситуаций, когда параметры элементов становятся близкими друг к другу, введен коэффициент взаимного пресечения элементов.
3. Экспериментально показано, что модифицированный алгоритм обучения сети позволяет автоматически формировать ее структуру в виде количества элементов второго слоя и их параметров.

СПИСОК ЛИТЕРАТУРЫ

1. Хайкин С. Нейронные сети: полный курс. 2-е изд. / под ред. Н.Н. Куссуль. – М.: Издат. дом «Вильямс», 2006. – 1104 с.
2. Осовский С. Нейронные сети для обработки информации: пер. с польск. – М.: Финансы и статистика, 2002. – 344 с.
3. Jianyu L., Luo Siwei, Qi Yingjiana, Huang Yapinga. Numerical solution of elliptic partial differential equation using radial basis function neural networks // Neural Networks. – 2003. – № 5/6. – P. 729–734.

Поступила 21.04.2009 г.