

РОБАСТНЫЙ АЛГОРИТМ ОБУЧЕНИЯ РАДИАЛЬНО-БАЗИСНОЙ АДАПТИВНОЙ ФАЗЗИ-ВЭЙВЛЕТ-НЕЙРОННОЙ СЕТИ

В статье предложен робастный алгоритм обучения радиально-базисной адаптивной фаззи-вейвлет-нейронной сети на основе критерия Р. Велша. Предложенный алгоритм обучения позволяет обрабатывать сигналы в условиях существенной зашумленности, а так же при наличии выбросов с большой амплитудой. Результаты эксперимента подтверждают эффективность предложенного робастного алгоритма обучения.

Введение

В настоящее время искусственные нейронные сети (ИНС) получили широкое распространение для решения широкого класса задач обработки информации и, прежде всего, для идентификации, эмуляции, интеллектуального управления, прогнозирования временных рядов произвольной природы в условиях сильной зашумленности, а также структурной и параметрической неопределенности.

Наиболее известными и популярными является многослойные сети с прямой передачей информации типа трехслойного персептрона, элементарными узлами которых есть, так называемые, P – нейроны с монотонными функциями активации. Эффективность многослойных сетей объясняется их универсальными аппроксимирующими свойствами в сочетании с относительно компактным представлением моделируемой нелинейной системы. Это означает, что они с успехом могут быть использованы в задачах моделирования (эмуляции) нелинейных систем, описываемых уравнениями вида

$$y(k) = F(x(k)) + \xi(k), \quad (1)$$

где $y(k)$ – выходной сигнал системы в k -й момент дискретного времени $k = 0, 1, 2, \dots$, $x(k) \in X - (n \times 1)$ – вектор входных сигналов, включающий в себя как экзогенные переменные, так и прошлые значения выходного сигнала, $F(\bullet)$ – произвольная функция в общем случае неизвестного вида, $\xi(k)$ – ненаблюдаемое возмущение с неизвестными характеристиками. Обычно предполагается также, что функция $F(\bullet)$ задана либо на единичном гиперкубе, либо на ортотопе

$$x_i(k) \in [x_i^{min}, x_i^{max}], i = 1, 2, \dots, n,$$

где x_i^{min}, x_i^{max} известные нижняя и верхняя границы варьирования i -го входного воздействия.

© Е.В. Бодянский, Е.А. Винокурова, 2007

Основным недостатком многослойных сетей является низкая скорость их обучения, основанного на обратном распространении ошибок, что делает невозможным их использование в задачах реального времени.

Альтернативой многослойным ИНС являются радиально–базисные–нейронные сети (РБНС), имеющие один скрытый слой, образованный, так называемыми, R – нейронами, при этом обучение этих сетей реализуется на уровне выходного слоя, представляющего собой адаптивный линейный ассоциатор [1–6]. В отличие от P –нейронов, R –нейроны имеют, как правило, колоколообразную функцию активации $f_j(x)$, аргументом которой является расстояние (обычно в евклидовой метрике) между текущим значением входного сигнала $x(k)$ и центром $c_j - j$ -го нейрона, т.е.

$$\varphi_j(x(k)) = \varphi_j \left(\sum_{i=1}^n (x_i(k) - c_{ji})^2 \right) = \varphi_j(\|x(k) - c_j\|^2). \quad (2)$$

Основным преимуществом РБНС является высокая скорость обучения в выходном слое, определяемая тем, что настраиваемые параметры входят в описание сети линейно. В то же время остается открытой проблема размещения центров R – нейронов, неудачное решение которой ведет к возникновению ”проклятия размерности”. Применение методов кластеризации, хотя и позволяет уменьшить размеры сети, исключает возможность работы в реальном времени. Здесь же следует отметить, что в [7] описана градиентная рекуррентная процедура покомпонентной настройки параметров c_{ji} , однако она характеризуется низкой скоростью сходимости.

Наряду с нейронными сетями для обработки сигналов различной природы последнее время достаточно часто используется теория вэйвлет–преобразования [8–10], обеспечивающая компактное локальное представление сигналов как в частотной, так и во временной области. На стыке теорий ИНС и вэйвлетов возникли вэйвлет–нейронные сети [11–14], подтвердившие свою эффективность в задачах обработки нестационарных нелинейных сигналов и процессов.

Элементарными узлами вэйвлет–нейронных сетей являются, так называемые, радиальные вэйвлоны [15], активационными функциями которых являются четные вэйвлеты с аргументом в виде евклидова расстояния между $x(k)$ и центром вэйвлета c_j , при этом каждая компонента расстояния $|x_i(k) - c_{ji}|$ нормируется на параметр ширины σ_{ji} так, что

$$\varphi_j(x(k)) = \varphi_j \left(\sum_{i=1}^n \left(\frac{x_i(k) - c_{ji}}{\sigma_{ji}} \right)^2 \right), \quad (3)$$

где $\varphi(\bullet)$ – активационная функция–вэйвлет. Рецепторными полями таких вэйвлонов являются гиперэллипсоиды с осями коллинеарными координатным осям пространства X .

Учитывая эквивалентность радиально–базисных ИНС и систем нечеткого вывода [16,17], а также возможность использования в качестве

функций принадлежности четных вэйвлетов [18,19], в рамках парадигмы унификации [15] можно говорить о такой гибридной системе как радиально–базисная фаззи–вэйвлет–нейронная сеть (РБФВНС), обладающая возможностью быстрого обучения радиально–базисных нейронных сетей, интерпретируемостью систем нечеткого вывода и локальными свойствами вэйвлетов.

В большинстве все алгоритмы обучения основываются на квадратичных критериях качества и в случае, если обрабатываемые данные зашумлены возмущениями с неизвестными законами распределения, то эти алгоритмы оказываются очень чувствительными к аномальным выбросам. Таким образом, актуальной задачей является синтез робастных алгоритмов обучения, которые позволяют обрабатывать зашумленные сигналы с аномальными возмущениями.

Настоящая статья посвящена синтезу робастного алгоритма обучения РБФВНС, обладающего регулируемым уровнем нечувствительности к различного рода выбросам, грубым ошибкам, негауссовским возмущениям, имеющего повышенную скорость сходимости и обеспечивающего улучшенные аппроксимирующие свойства по сравнению с традиционными системами вычислительного интеллекта.

Архитектура радиально–базисной–фаззи–вэйвлет–нейронной сети

Рассмотрим двуслойную архитектуру, приведенную на рис. 1 и совпадающую по сути с традиционной радиально–базисной нейронной сетью.

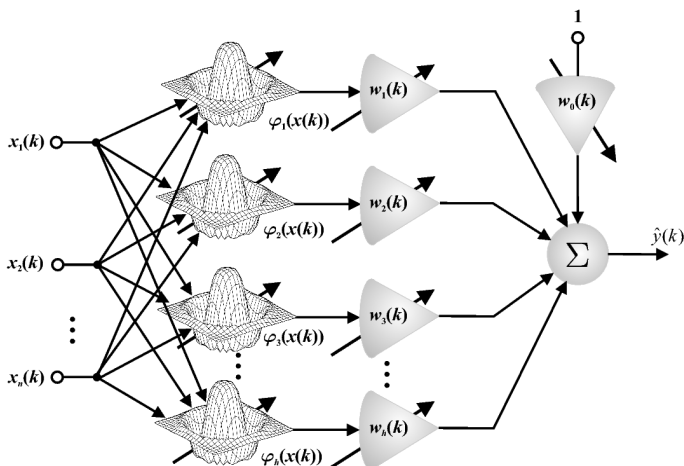


Рис. 1 – Радиально–базисная фаззи–вэйвлет–нейронная сеть

Нулевой слой архитектуры является рецепторным и в текущий момент времени k на него подается входной сигнал в форме вектора

$x(k) = (x_1(k), x_2(k), \dots, x_n(k))^T$. Скрытый слой в отличие от радиально-базисных сетей образован не R -нейронами, а вэйвлонами с активационными функциями-вэйвлетами вида

$$\varphi_j(x(k)) = \varphi_j((x_i(k) - c_{ji})^T Q_j^{-1}(x_i(k) - c_{ji})), \quad (4)$$

в которых вместо параметров смещения σ_{ji} в (1) используется положительно определенная матрица Q_j , т.е. используется не евклидова метрика, а расстояние Махаланобиса. Это приводит к тому, что рецепторные поля – гиперэллипсоиды вэйвлонов (2) могут иметь произвольную ориентацию относительно координатных осей пространства X , что расширяет функциональные возможности РБФВНС.

И, наконец, выходной слой представляет собой обычный адаптивный линейный ассоциатор с настраиваемыми синаптическими весами w_j

$$\hat{y}(k) = w_0 + \sum_{j=1}^h w_j \varphi_j((x_i(k) - c_{ji})^T Q_j^{-1}(x_i(k) - c_{ji})) = w^T \varphi(x(k)), \quad (5)$$

где $\varphi_0(x(k)) \equiv 1$, $w = (w_0, w_1, w_2, \dots, w_h)^T$, $\varphi(x(k)) = (1, \varphi_1(x(k)), \varphi_2(x(k)), \dots, \varphi_h(x(k)))$.

Настраиваемыми параметрами архитектуры, подлежащими определению в процессе обучения, будем полагать $(h+1)$ синаптических весов w_j , h ($n \times 1$)-векторов c_j и h ($n \times n$)-матриц Q_j^{-1} . Всего же такая сеть содержит $h(1+n+n^2) + 1$ уточняемых параметров.

Робастный алгоритм обучения радиально-базисной-фаззи-вэйвлет-нейронной сети

Опыт показывает, что методы идентификации, основанные на критерии наименьших квадратов, оказываются чрезвычайно чувствительными к отклонениям фактического закона распределения данных от нормального. В условиях различного рода выбросов, грубых ошибок, негауссовских возмущений с “тяжелыми хвостами” методы, связанные с критерием наименьших квадратов, теряют свою эффективность.

В этой ситуации на первый план выходят методы робастного оценивания [20], которые к настоящему времени получили известное распространение и для обучения искусственных нейронных сетей [21–23].

Вводя в рассмотрение ошибку обучения

$$e(k) = (y(k) - \hat{y}) = y(k) - w^T(k) \varphi(k) \quad (6)$$

и робастный критерий идентификации Р. Велша [24,25].

$$E(k) = f(k) = \beta^2 \ln \left(\cosh \frac{e(k)}{\beta} \right), \quad (7)$$

где β – положительный параметр, выбираемый из эмпирических соображений и определяющий размер зоны нечувствительности к выбросам.

Сравнение робастного критерия идентификации и критерия наименьших квадратов показано на рис. 2.

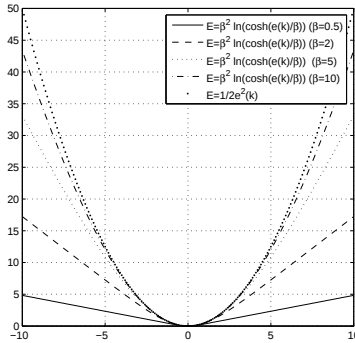


Рис. 2 – Сравнение критериев оптимизации

Необходимо отметить, что робастный критерий (7) удовлетворяет аксиомам метрического пространства:

1. **Невырожденность:** $f(x - y) = 0$ в том и только в том случае, когда x и y один и тот же элемент.

$$\text{Действительно, } y = x \implies f(x - x) = f(0) = \beta^2 \ln \left(\cosh \left(\frac{0}{\beta} \right) \right).$$

2. **Симметричность:** $f(x - y) = f(y - x)$.

$$\begin{aligned} \text{Действительно, } f(x - y) &= \beta^2 \ln \left(\cosh \left(\frac{x - y}{\beta} \right) \right) \text{ и } f(y - x) \\ &= \beta^2 \ln \left(\cosh \left(\frac{y - x}{\beta} \right) \right) = \beta^2 \ln \left(\cosh \left(-\frac{x - y}{\beta} \right) \right) - \\ &\beta^2 \ln \left(\cosh \left(\frac{x - y}{\beta} \right) \right) \implies f(x - y) = f(y - x) \end{aligned}$$

3. **Неотрицательность:** $f(x - y) \geq 0$.

$$\text{Действительно, } f(x - y) = \beta^2 \ln \left(\cosh \left(\frac{x - y}{\beta} \right) \right) \geq 0, \text{ поскольку функция } \cosh(\bullet) \text{ – положительная.}$$

4. **Неравенство треугольника:** $f(x - y) \leq f(x - z) + f(z - y)$.

Действительно,

$$\begin{aligned}\beta^2 \ln \left(\cosh \left(\frac{x-y}{\beta} \right) \right) &\leq \beta^2 \ln \left(\cosh \left(\frac{x-z}{\beta} \right) \right) + \beta^2 \ln \left(\cosh \left(\frac{x-y}{\beta} \right) \right); \\ \ln \left(\cosh \left(\frac{x-y}{\beta} \right) \right) &\leq \ln \left(\cosh \left(\frac{x-z}{\beta} \right) \right) + \ln \left(\cosh \left(\frac{x-y}{\beta} \right) \right); \\ 0 < \cosh \left(\frac{x-y}{\beta} \right) &\leq \cosh \left(\frac{x-z}{\beta} \right) \cosh \left(\frac{x-y}{\beta} \right).\end{aligned}$$

Проведя ряд несложных преобразований, получаем, что

$$2 \cosh \left(\frac{x-2z+y}{\beta} \right) \geq 0.$$

Данное неравенство справедливо при любых значениях x, y, z , поскольку функция $\cosh(\bullet)$ – положительная.

Далее рассмотрим процесс синтеза алгоритмов обучения. Для настройки синаптических весов и параметров вэйвлонов (векторов c_j и матриц Q_j^{-1}) будем использовать градиентную минимизацию критерия (7), при этом в отличие от покомпонентного обучения, рассмотренного в [7], будем производить уточнение в векторно-матричной форме, что, во-первых, проще с вычислительной точки зрения, а, во-вторых, позволит оптимизировать процесс обучения по быстродействию.

В общем случае алгоритм обучения может быть записан в виде

$$\begin{cases} w(k+1) = w(k) - \eta_w \nabla_w E(k), \\ c_j(k+1) = c_j(k) - \eta_{c_j} \nabla_{c_j} E(k), j = 1, \dots, h, \\ Q_j^{-1}(k+1) = Q_j^{-1}(k) - \eta_{Q_j^{-1}} \left\{ \frac{\partial E(k)}{\partial Q_j^{-1}} \right\}, j = 1, \dots, h, \end{cases} \quad (8)$$

где $\nabla_w E(k)$ – вектор-градиент критерия (7) по w , $\nabla_{c_j} E(k)$ – вектор-градиент критерия (7) по c_j ; $\left\{ \frac{\partial E(k)}{\partial Q_j^{-1}} \right\} - (n \times n)$ – матрица, образованная частными производными E по компонентам Q_j^{-1} ; $\eta_w, \eta_{c_j}, \eta_{Q_j^{-1}}$ – параметры шага алгоритма обучения.

Для произвольно взятого вэйвлета $\varphi((x(k) - c_j(k))^T Q_j^{-1} (x(k) - c_j(k)))$ можно записать

$$\left\{ \begin{array}{l} \nabla_w E(k) = -\beta \tanh\left(\frac{e(k)}{\beta}\right) \varphi_j((x(k) - c_j(k))^T Q_j^{-1}(x(k) - c_j(k))) = \\ \quad = -\tanh\left(\frac{e(k)}{\beta}\right) J_w(k), \\ \nabla_{c_j} E(k) = \beta \tanh\left(\frac{e(k)}{\beta_c}\right) w_j \varphi'_j((x(k) - c_j(k))^T Q_j^{-1}(x(k) - c_j(k))) \cdot \\ \quad \cdot Q_j^{-1}(x(k) - c_j(k)) = \tanh\left(\frac{e(k)}{\beta}\right) J_{c_j}(k), \\ \left\{ \frac{\partial E(k)}{\partial Q_j^{-1}} \right\} = -\beta \tanh\left(\frac{e(k)}{\beta}\right) w_j \varphi'_j((x(k) - c_j(k))^T Q_j^{-1}(x(k) - c_j(k))) \cdot \\ \quad \cdot (x(k) - c_j(k))(x(k) - c_j(k))^T = -\tanh\left(\frac{e(k)}{\beta}\right) J_{Q_j^{-1}}(k), \end{array} \right. \quad (9)$$

где $\varphi'_j(\bullet)$ – производная j -го вэйвлета по аргументу $((x(k) - c_j(k))^T Q_j^{-1}(x(k) - c_j(k)))$.

Тогда алгоритм обучения вэйвлонов скрытого слоя с учетом (9) принимает форму

$$\left\{ \begin{array}{l} w(k+1) = w(k) + \eta_w \beta \tanh\left(\frac{e(k)}{\beta_w}\right) \varphi_j((x(k) - c_j(k))^T Q_j^{-1}(x(k) - c_j(k))) = \\ \quad = w(k) + \eta_w \tanh\left(\frac{e(k)}{\beta_w}\right) J_w(k), \\ c_j(k+1) = c_j(k) - \eta_{c_j} \beta \tanh\left(\frac{e(k)}{\beta_c}\right) \cdot \\ \quad \cdot w_j(k) \varphi'_j((x(k) - c_j(k))^T Q_j^{-1}(x(k) - c_j(k))) \cdot \\ \quad \cdot Q_j^{-1}(x(k) - c_j(k)) = c_j(k) - \eta_{c_j} \tanh\left(\frac{e(k)}{\beta_c}\right) J_{c_j}(k), \\ Q_j^{-1}(k+1) = Q_j^{-1}(k) + \eta_{Q_j^{-1}} \beta \tanh\left(\frac{e(k)}{\beta_Q}\right) \cdot \\ \quad \cdot w_j \varphi'_j((x(k) - c_j(k))^T Q_j^{-1}(x(k) - c_j(k))) \cdot \\ \quad \cdot (x(k) - c_j(k))(x(k) - c_j(k))^T = -\tanh\left(\frac{e(k)}{\beta_Q}\right) J_{Q_j^{-1}}(k), \end{array} \right. \quad (10)$$

при этом скорость сходимости к оптимальным значениям w , c_j и Q_j^{-1} полностью определяется параметрами шага η_w , η_{c_j} и $\eta_{Q_j^{-1}}$.

Повышение скорости сходимости может быть достигнуто путем использования более сложных, нежели градиентные, процедур типа Хартли или Марквардта, которые для первого соотношения (10) могут быть записаны в обобщенной форме [26]

$$w(k+1) = w(k) - \lambda_w \left(J_w(k) J_w^T(k) + \eta_w I \right)^{-1} J_w(k) \tanh \left(\frac{e(k)}{\beta_w} \right), \quad (11)$$

где $I - (n \times n)$ -единичная матрица, λ_w – положительный демпфирующий параметр, η_w – параметр регуляризации.

Воспользовавшись леммой обращения матриц, после несложных преобразований можно получить простой и эффективный алгоритм обучения параметров в виде

$$\begin{cases} w(k+1) = w(k) + \lambda_w \frac{\tanh \left(\frac{e(k)}{\beta_w} \right) J_w(k)}{\eta_w + \|J_w(k)\|^2}, \\ c_j(k+1) = c_j(k) - \lambda_{c_j} \frac{\tanh \left(\frac{e(k)}{\beta_c} \right) J_{c_j}(k)}{\eta_{c_j} + \|J_{c_j}(k)\|^2}, \\ Q_j^{-1}(k+1) = Q_j^{-1}(k) + \lambda_{Q_j^{-1}} \frac{\tanh \left(\frac{e(k)}{\beta_Q} \right) J_{Q_j^{-1}}(k)}{\eta_{Q_j^{-1}} + \text{Tr}(J_{Q_j^{-1}}^T(k) J_{Q_j^{-1}}(k))}. \end{cases} \quad (12)$$

Для придания процессу обучения дополнительных сглаживающих свойств, используя подход, предложенный в [27], можно ввести следующую модифицированную процедуру обучения:

$$\begin{cases} w(k+1) = w(k) + \lambda_w \frac{\tanh \left(\frac{e(k)}{\beta_w} \right) J_w(k)}{\eta_w(k)}, \\ \eta_w(k+1) = \alpha_w \eta_w(k) + \|J_w(k+1)\|^2, \\ c_j(k+1) = c_j(k) - \lambda_{c_j} \frac{\tanh \left(\frac{e(k)}{\beta_c} \right) J_{c_j}(k)}{\eta_{c_j}(k)}, \\ \eta_{c_j}(k+1) = \alpha_{c_j} \eta_{c_j}(k) + \|J_{c_j}(k+1)\|^2, \\ Q_j^{-1}(k+1) = Q_j^{-1}(k) + \lambda_{Q_j^{-1}} \frac{\tanh \left(\frac{e(k)}{\beta_Q} \right) J_{Q_j^{-1}}(k)}{\eta_{Q_j^{-1}}(k)}, \\ \eta_{Q_j^{-1}}(k+1) = \alpha_{Q_j^{-1}} \eta_{Q_j^{-1}}(k) + \text{Tr}(J_{Q_j^{-1}}^T(k+1) J_{Q_j^{-1}}(k+1)) \end{cases} \quad (13)$$

(здесь $0 \leq \alpha_w \leq 1$, $0 \leq \alpha_{c_j} \leq 1$, $0 \leq \alpha_{Q_j^{-1}} \leq 1$ – параметры взвешивания устаревшей информации), являющуюся нелинейным гибридом алгоритмов Качмажа-Уидроу-Хоффа и Гудвина-Рэмеджа-Кэйнеса и обладающую как следящими, так и фильтрующими свойствами.

Результаты экспериментальных исследований

Экспериментальное исследование разработанного робастного алгоритма обучения проводилось на основе сигнала, зашумленного интенсив-

ными выбросами. Сигнал получен на основе ряда Мэки-Гласса, который был сгенерирован с помощью разностного уравнения [28]

$$\dot{x} = \frac{0.2x(t - \tau)}{1 + x^{10}(t - \tau)} - 0.1x(t), \quad (14)$$

(значения временного ряда для каждой целой точки были получены с помощью метода Рунге-Кутты четвертого порядка, используемый временной шаг для данного метода был выбран 0.1, а начальные условия $x(0) = 1.2$, задержка $\tau = 17$ и $x(t)$ были получены для $t = 0 \dots 51000$) и наложенного на него случайного шума, сгенерированного по распределению Коши с помощью метода обратного преобразования, которое описывается выражением

$$F_X^{-1}(x) = x_0 + \gamma \tan \left[\pi \left(x - \frac{1}{2} \right) \right], \quad (15)$$

где x_0 – коэффициент сдвига; γ – коэффициент масштаба ($\gamma > 0$); x – носитель ($x \in (-\infty, +\infty)$).

На рис. 3 представлен общий вид обрабатываемого сигнала.

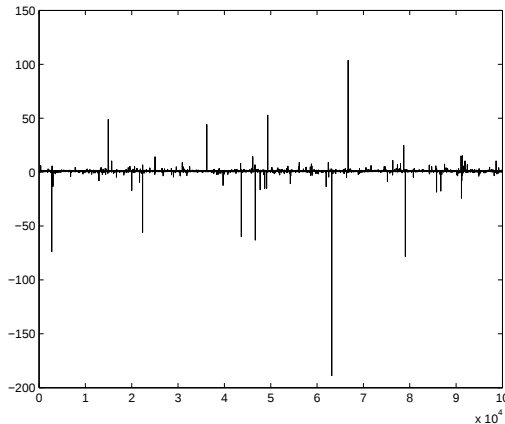


Рис. 3 – Временной ряд Мэки-Гласса с шумом, имеющим распределение Коши

Значения $x(t - 18)$, $x(t - 12)$, $x(t - 6)$, $x(t)$ были использованы для прогнозирования $x(t + 6)$. В режиме on-line обучения, РБФНС была обучена с помощью алгоритма обучения (13) для 50000 итераций (50000 точек для $t = 118, \dots, 50117$). Параметры обучающего алгоритма были взяты $\beta_w = 1$, $\beta_c = 0.5$, $\beta_Q = 0.5$, $\alpha_w = \alpha_c = \alpha_Q = 0.99$, $\lambda_w = \lambda_{c_j} = \lambda_{Q_j^{-1}} = 0.99$. Начальные значения были взяты $\eta_w(0) = \eta_{c_j}(0) = \eta_{Q_j^{-1}} = 10000$. После 50000 итераций обучение было остановлено и следующие 500 точек

для $t = 50118, \dots, 50617$ были использованы как тестовые данные для прогноза. В качестве активационной функции был взят вэйвлет “Mexican hat”. Начальные значения синаптических весов были сгенерированы случайным образом в диапазоне от -0.1 до 0.1 .

В качестве критерия качества прогноза была использована среднеквадратическая ошибка

$$RMSE = \frac{1}{N} \sum_{k=1}^N (x(k) - \hat{x})^2. \quad (16)$$

На рис. 4 а представлены результаты прогнозирования зашумленного временного ряда (реальные значения (пунктирная линия) и прогнозируемые значения (сплошная линия)). На рис. 4 б представлен сегмент процесса обучения; как видно, ряд выбросов с большой амплитудой, присутствующих в начале выборки, не оказали большого влияния на алгоритм обучения.

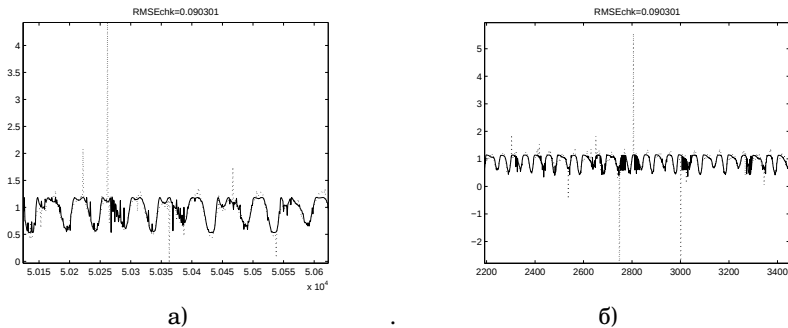


Рис. 4 – Результаты обработки зашумленного сигнала с помощью робастного алгоритма обучения

Сравнение результатов прогнозирования на основе робастного алгоритма обучения проводилось с результатами прогнозирования на основе градиентного алгоритма и алгоритма на основе рекуррентного метода наименьших квадратов, где структура сети, и количество параметров настройки было одинаковым. На рис.5 а представлены результаты прогнозирования сигнала на основе градиентного алгоритма обучения (реальные значения (пунктирная линия) и прогнозируемые значения (сплошная линия)). На рис.5 б представлен сегмент процесса обучения; как видно, первый же выброс в начале выборки сильно повлиял на процесс обучения и как результат – большая ошибка прогнозирования.

При обучении РБФВНС рекуррентным методом наименьших квадратов при первом выбросе происходит “взрыв параметров” ковариационной матрицы и как результат – невозможность прогнозирования сигналов,

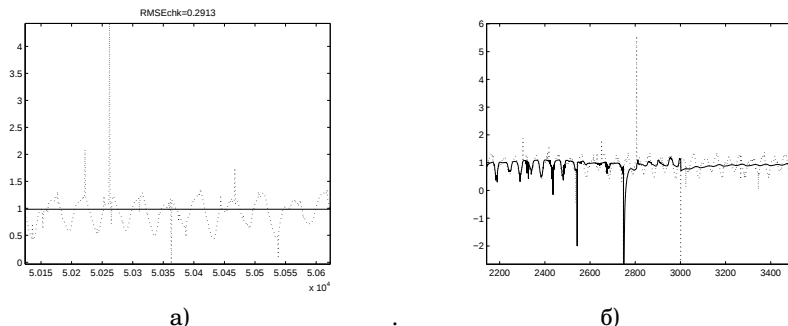


Рис. 5 – Результаты обработки зашумленного сигнала с помощью градиентного алгоритма обучения

зашумленных аномальными выбросами. В Таблице 1 приведены результаты сравнения.

Табл. 1 – Сравнение результатов прогнозирования зашумленного сигнала

Нейронная сеть / Алгоритм обучения	RMSE
Радиально-базисная адаптивная фаззи-вэйвлет-нейронная сеть / Предложенный робастный алгоритм обучения (14)	0.0903
Радиально-базисная адаптивная фаззи-вэйвлет-нейронная сеть / Градиентный алгоритм обучения	0.2913
Радиально-базисная адаптивная фаззи-вэйвлет-нейронная сеть / РМНК	∞

Таким образом видно, что предложенный робастный алгоритм обучения позволяет обрабатывать сигналы в условиях существенного загрязнения выбросами.

Выводы

В статье предложен численно простой и эффективный робастный алгоритм обучения всех параметров радиально-базисной фаззи-вэйвлет-нейронной сети, обладающий следящими и сглаживающими свойствами и позволяющий в реальном времени обрабатывать нестационарные нелинейные сигналы в условиях различного рода выбросов и возмущений с “тяжелыми хвостами”. Надстройка рецепторных полей вэйвлонов, включающая в себя их преобразования (сдвиг, растяжение, вращение), позволяет улучшить аппроксимирующие свойства сети, что подтверждается результатами экспериментального исследования.

Литература

1. Moody J., Darken C. J. Fast learning in networks of locally-tuned processing units // Neural Computation. – 1989. – №1. – P. 281–294.

2. Moody J., Darken C.J. Learning with localized receptive fields // Proc. of the 1988 Connectionist Models Summer School. – San Mateo: Morgan-Kaufmann, 1988. – P. 133–143.
3. Park J., Sandberg I. W. Universal approximation using radial-basis-function networks // Neural Computation. – 1991. – N3. – P. 246–257.
4. Leonard J. A., Kramer M. A., Ungar L. H. Using radial basis functions to approximate a function and its error bounds // IEEE Trans. on Neural Networks. – 1992. – N3. – P. 614–627.
5. Sunil E. V. T., Yung C. Sh. Radial basis function neural network for approximation and estimation of nonlinear stochastic dynamic systems // IEEE Trans. on Neural Networks. – 1994. – N5. – P. 594–603.
6. Poggio T., Girosi F. A Theory of Networks for Approximation and Learning. Massachusetts Institute of Technology: A. I. Memo – 1140, C.B.I.P. Paper – 31, 1994. – 63 p.
7. Bishop C. M. Neural Networks for Pattern Recognition. Oxford: Clarendon Press, 1995. – 482 p.
8. Chui C.K. An Introduction to Wavelets. New York: Academic, 1992. – 264 p.
9. Daubechies I. Ten Lectures on Wavelets. Philadelphia, PA: SIAM, 1992. – 228 p.
10. Meyer Y. Wavelets: Algorithms and Applications. Philadelphia, PA: SIAM, 1993. – 133 p.
11. Billings S.A., Wei H.-L. A new class of wavelet networks for nonlinear system identification // IEEE Trans. on Neural networks. – 2005. – 16. – N4. – P. 862–874.
12. Zhang Q. H., Benveniste A. Wavelet networks // IEEE Trans. on Neural Networks. – 1992. – 3. – N6. – P. 889–898.
13. Zhang J., Walter G.G., Miao Y., Lee W.N. Wavelet neural networks for function learning // IEEE Trans. on Signal Processing. – 1995. – 43. – N6. – P. 1485–1497.
14. Zhang Q.H. Using wavelet network in nonparametric estimation // IEEE Trans. on Neural Networks. – 1997. – 8. – N2. – P. 227–236.
15. Reyneri L. M. Unification of neural and wavelet networks and fuzzy systems // IEEE Trans. on Neural Networks. – 1999. – N10. – P. 801–814.
16. Jang J. S. R., Sun C.T. Functional equivalence between radial basis function networks and fuzzy inference systems // IEEE Trans. on Neural Networks. – 1993. – N4. – P. 156–159.
17. Hunt K. J., Haas R., Smith R. M. Extending the functional equivalence of radial basis function networks and fuzzy inference systems // IEEE Trans. on Neural Networks. – 1996. – N7. – P. 776–781.

18. Kosko B. Mitaim S. What is the best shape for a fuzzy set in function approximation? // Proc. 5th IEEE Int. Conf on Fuzzy Systems “Fuzz-96”. – 1996. – 2. – P 1237–1213.
19. Kosko B. Mitaim S. Adaptive joint fuzzy sets for function approximation // Proc. 1997 Int. Conf. on Neural Networks “ICNN-97”. – 1997. – P 537–542.
20. Rey W.J.J. Robust Statistical Methods // Lecture Note in Mathematics. – 2. Berlin-Heidelberg-New York: Springer-Verlag, 1978. – 128 p.
21. Cichocki A., Lobos T. Adaptive analogue network for real time estimation of basic waveforms of voltages and currents // IEEE Proc. – 1992. – 139, Part C. – P 343–350.
22. Unbehauen R. Cichocki A. Neural Networks for Optimization and Signal Processing. Stuttgart: Teubner, 1993. – 526 p.
23. Sheng-Tun Li, Shu-Ching Chen. Function approximation using robust wavelet neural networks // Proc. of the 14th IEEE International Conference on Tools with Artificial Intelligence (ICTAI-02). – 2002. – P 483–488.
24. Holland P.W., Welsh R.E. Robust regression using iteratively reweighted least squares // Commun. Statist. - Theory and Methods. – 1977. – N46. – P 813–828.
25. Welsh R.E. Nonlinear statistical data analysis // Proc. Comp. Sci. and Statist., Tenth Ann Symp. Interface. Held at Nat'l Bur. Stds. Gaithersburg, MD. – 1977. – P 77–86.
26. Бодянский Е. В. Адаптивные алгоритмы идентификации нелинейных объектов управления // АСУ и приборы автоматики. – Харьков: Выща шк., 1987. – N81. – С. 43–46.
27. Ye. Bodyanskiy, V. Kolodyazhniy, A. Stephan. An adaptive learning algorithm for a neuro-fuzzy network. // Computational Intelligence. Theory and Applications, Lecture Notes in Computer Science. – 2001. – 2206. – P 68–75.
28. Glass L. Mackey M. C. Oscillation and chaos in physiological control systems // Science. – 1977. – N197. – P 287–289.

Получено 15.11.2007