

# Прогнозирование изменения параметров временных рядов и непрерывных функций

А.Ю. Проскураков<sup>а</sup>, Ю.А. Кропотов<sup>а</sup>

<sup>а</sup>Муромский институт (филиал) ФГБОУ ВО «Владимирский государственный университет имени Александра Григорьевича и Николая Григорьевича Столетовых», 602264, ул. Орловская, 23, Муром, Россия

## Аннотация

Рассмотрены различные классы алгоритмов прогнозирования изменений параметров непрерывных функций и временных рядов, имеющих место на интервале, называемом горизонт предсказания. Показаны ограниченные возможности методов декомпозиции процессов на эмпирические моды и методы параметрического прогнозирования, основанные на представлении временного ряда обобщённым многочленом по системе линейно независимых функций. Получены алгоритмы прогноза, основанные на моделях авторегрессии. Получены рекуррентные уравнения, определяющие явную зависимость оценок прогноза от коэффициентов модели. Рассмотрены вопросы нахождения оценок прогноза посредством минимизации функции потерь – квадрата нормы отклонения оценок от наблюдаемых значений временного ряда. Исследован метод обобщения алгоритма прогноза с помощью линейной модели, представленной функциональными рядами или ИНС. Представлена система прогнозирования с помощью ИНС.

*Ключевые слова:* прогнозирование; параметрическое предсказание; авторегрессия; функциональный ряд; нейронная сеть; временной ряд; трехслойный персептрон прямого распространения.

## 1. Введение

Задача прогнозирования заключается в нахождении будущих значений параметров временного ряда или непрерывного процесса на интервале, называемом горизонтом предсказания [1, 2], в пределах которого обеспечивается необходимая точность решения задачи. Для непрерывных процессов – это интервал  $(t, t + \tau]$ , который для временного ряда, полученного, в частности, посредством дискретизации непрерывного процесса, записывается как  $(n, n + N]$ . Здесь  $t$  и  $nT$  – текущие моменты времени, при этом  $T$  – период дискретизации. Прогнозирование обычно осуществляется по значениям временного ряда или процесса на конечном, предшествующем, интервале  $[t - T, t]$  времени.

Горизонт предсказания является не только одной из важнейших мер качества предсказуемости, но и используется в качестве критерия степени детерминированности и случайности наблюдаемых явлений, служит характеристикой динамического хаоса (характеристикой хаотических колебаний в динамических системах). В основе этого утверждения лежит зависимость горизонта предсказания не только от используемых алгоритмов, но и от свойств анализируемых временных рядов и процессов. В задачах прогнозирования выбор алгоритма осуществляется исходя из соображений максимизации горизонта предсказания. Один из принципов прогнозирования временных рядов или процессов основывается на их представлении непрерывными или дискретными моделями. В работах [1, 2] модель схематически описывается дифференциальным уравнением, зависящим от неизвестных параметров системы  $a$  и факторов  $f_k$ , отражающих неопределённость модели. При этом в рассмотрение вводятся три процесса: наблюдаемый процесс  $y$ , исследуемый (истинный) процесс  $x$  и модельный (прогностический) процесс  $z$ . Истинный процесс в силу неопределенности факторов  $f_k$  является (из множества возможных) неизвестным решением дифференциального уравнения

$$P(d/dt, x, a, f_k) = 0.$$

Пренебрежение указанными факторами позволяет получить дифференциальное уравнение, описывающее модельный процесс  $z$ , при соответствующих условиях близкий к истинному процессу  $x$ . Это уравнение можно записать в виде

$$G(d/dt, z, a) = 0.$$

Здесь критерием качества прогнозирования может является среднеквадратическое значение нормы отклонения модельного процесса по факторам  $f_k$  от истинного на интервале предсказания, то есть величина [1]

$$\langle \|x - z\|^2 \rangle = \frac{1}{M} \sum_{k=1}^M \int_t^{t+\tau} |x(v, f_k) - z(v)|^2 dv.$$

Аналогично определяется качество предсказуемости дискретных процессов или временных рядов, представленных дискретными моделями, которые, в частности, могут быть получены из непрерывных моделей путём замены производных конечными разностями. При этом операция интегрирования заменяется операцией суммирования по конечному на интервале прогнозирования множеству данных. Переход к дискретной модели эквивалентен численному решению дифференциального уравнения со свойственными этому решению проблемами чувствительности к возмущающим воздействиям.

На практике построение моделей истинных процессов основывается на данных о соответствующих наблюдаемых процессах: модели могут относиться к классам линейных дискретных [3] и регрессионных [4] систем.

Сложность прогнозирования нестационарных процессов обусловила появление метода, сочетающего в себе их декомпозицию на эмпирические моды (empirical mode decomposition – EMD) и преобразование Гильберта [5]. Этот

метод обеспечивает приближение наблюдаемого процесса суммой слагаемых вида  $A_k(t)\exp(j\varphi_k(t))$ , прогнозирование каждого из которых становится более надёжным и простым. В совокупности операция EMD и преобразование Гильберта называется преобразованием Гильберта – Хуанга [5]. При этом за эквивалент операции EMD принимаются и другие средства декомпозиции, в частности, банки фильтров [5].

Распространёнными методами прогнозирования являются параметрические методы регрессионной аппроксимации [4], динамические модели авторегрессии – скользящего среднего или скользящего среднего [3], методы импульсных функций [7] и искусственные нейронные сети [9, 10]. Решение задачи прогнозирования осуществляется не только в дискретном, но и в непрерывном времени [9].

## 2. Параметрическое регрессионное прогнозирование

Прогноз временного ряда этим методом предполагает нахождение функции регрессии или обобщённого многочлена по системе линейно независимых функций  $\{\varphi_k\}_1^M$ . Многочлен находится методом наименьших квадратов по известным на текущем интервале данным. Многочлен может быть записан в виде

$$P(t) = \sum_{k=1}^M a_k \varphi_k(t) = \Phi^T(t) \mathbf{a}, \quad (2.1)$$

где  $\Phi = (\varphi_1, \varphi_2, \dots, \varphi_M)^T$  и  $\mathbf{a} = (a_1, a_2, \dots, a_M)^T$ .

Если ввести матрицу, столбцами которой являются векторы значений базисных функций в узлах данных, то есть матрицу

$$\Phi = (\Phi(t_1), \Phi(t_2), \dots, \Phi(t_N)),$$

то вектор коэффициентов искомого многочлена

$$\mathbf{a} = (\Phi\Phi^T)^{-1} \Phi \mathbf{y}, \quad (2.2)$$

где  $\mathbf{y} = (y_1, y_2, \dots, y_N)^T$  - вектор наблюдаемых в текущий момент данных.

При этом за искомым прогнозом временного ряда в момент  $t_N + \tau$  берётся [4] величина

$$\hat{y}(t_N + \tau) = \Phi^T(t_N + \tau) \mathbf{a}. \quad (2.3)$$

Горизонт предсказания  $\tau$  полученной оценки таким образом невелик, поэтому прогнозирование по правилу (2.3) не получило широкого распространения. Кроме того, эта оценка весьма чувствительна к степени многочлена  $M$ , что выражается в значительном росте ошибки прогнозирования при увеличении степени  $n$ . Эта закономерность имеет место для многочленов по степенной системе базисных функций вида  $\Phi = (1, t, t^2, \dots, t^{M-1})$ .

Параметрический метод можно применить, если на интервале прогнозирования функциональный ряд по системе базисных функций  $\{\varphi_k\}_1^M$  является сходящимся, а его слагаемыми, начиная с номера  $M+1$ , можно пренебречь. Кроме того, функции системы  $\{\varphi_k\}_1^M$  должны, желательно, иметь на интервале их определения примерно одинаковую вариацию. В этом плане интересна система функций, ортонормированных на интервале  $[0, \infty]$ . Следует, однако, отметить, что приведённые предположения, требуют ещё своего доказательства. Данный метод является по существу методом статической локальной аппроксимации, в котором никак не отражается динамика временного ряда. Учесть эту динамику в принципе можно, установив зависимость коэффициентов  $a_k$  обобщённого многочлена (2.1) от времени, точнее, от номера окна локальной аппроксимации. Этим, фактически, осуществляется декомпозиция временного ряда на множество переменных коэффициентов многочлена (2.1), что позволяет отнести эту операцию к классу операций EMD.

Горизонт предсказуемости можно увеличить, если дополнить функцию потерь, минимизация которой даёт формулу (2.2), регуляризующим функционалом, стабилизирующим решение на интервале прогнозирования. При этом коэффициенты многочлена можно получить, решая задачу минимизации функции

$$\sum_{l=1}^N (\Phi^T(t_l) \mathbf{a} - y_l)^2 + \alpha \int_{t_N - v}^{t_N + \tau} (q \cdot \Phi^T(t) \mathbf{a})^2 dt,$$

где  $q$  - некоторый, в частности дифференциальный, оператор.

Как и выше, здесь не учитывается динамический характер временного ряда, а проблема выбора коэффициента регуляризации  $\alpha$  и интервала интегрирования  $[t_N - v, t_N + \tau]$  остается открытой.

## 3. Динамические модели прогнозирования

Задача прогнозирования при определённых условиях сводится к построению динамической модели, например, авторегрессии или авторегрессии–скользящего среднего, то есть к задаче представления временного ряда процессом на выходе некоторой динамической системы. Теоретические и чисто математические аспекты указанного представления, а точнее- представления временных рядов линейными системами, освещаются в [3]. В этой работе, в частности, освещаются проблемы точного и приближённого моделирования, а также рассматриваются модели многомерной авторегрессии и модели в переменных состояниях [2].

Прогноз в рамках подобного подхода осуществляется по известному состоянию модели в текущий момент времени и известному входному воздействию. Линейная дискретная модель, задаваемая импульсной функцией, описывается уравнением

$$x(n) = \sum_{l=-\infty}^n h(n-l)u(l),$$

где  $h(l)$  и  $u(l)$  – соответственно импульсная функция и входное воздействие модели.

Усечение правой части приведённого уравнения до конечного числа слагаемых даёт [4] модель скользящего среднего:

$$x(n) = \sum_{l=n-m}^n h(n-l)u(l) \equiv \sum_{l=0}^m b_l u(n-l).$$

Модель авторегрессии–скользящего среднего описывается разностным уравнением

$$x(n) = \sum_{l=1}^p \alpha_l x(n-l) + \sum_{l=0}^m b_l u(n-l). \tag{3.1}$$

Частной формой уравнения (3.1) является уравнение авторегрессии [3]

$$x(n) = \sum_{l=1}^p \alpha_l x(n-l) + u(n). \tag{3.2}$$

Если входное воздействие известно, то уравнение (3.1), включая и его частный случай – уравнение (3.2), это фактически уравнения прогноза на один шаг вперед. Коэффициенты уравнений (3.1) и (3.2) находятся посредством идентификации системы на текущем интервале данных.

Общее решение уравнения (3.1) или (3.2) с постоянными коэффициентами является, как известно [4], суммой его частного решения  $f(n)$  и общего решения однородного уравнения. Если линейно независимые частные решения однородного уравнения обозначить как  $\varphi_l(n), l = 1, 2, \dots, r$  то его общее решение можно записать в виде

$$\varphi(n) = \sum_{k=1}^p c_k \varphi_k(n).$$

Функции  $\varphi_k(n)$  выражаются через собственные значения характеристического многочлена

$$q^n - \sum_{l=1}^p \alpha_l q^{r-l} = 0.$$

При этом, если корень  $\lambda_k$  этого многочлена простой, то собственная функция имеет вид

$$\varphi_k(n) = \lambda_k^n \equiv p_k^n e^{s_k n}.$$

Корню же кратности  $r_k$  отвечают функции

$$\varphi_{k,1}(n) = \lambda_k^n \equiv p_k^n e^{s_k n},$$

$$\varphi_{k,2}(n) = n \lambda_k^n \equiv n p_k^n e^{s_k n}, \dots, \varphi_{k,r_k}(n) = n^{r_k-1} \lambda_k^n \equiv n^{r_k-1} p_k^n e^{s_k n}.$$

В качестве модели прогноза на один шаг вперед естественно выбрать автономную систему, описываемую уравнением

$$\hat{x}(n) = \sum_{l=1}^p \alpha_l x(n-l). \tag{3.3}$$

Здесь  $\hat{x}(n)$  - прогноз на один шаг. Оценки коэффициентов модели  $\alpha_k$  можно, например, получить, используя соответствующие (3.3) уравнения Юла – Уокера [4].

Если принять, что оценки  $\hat{\alpha}_k$  являются действительными значениями коэффициентов модели, то решение задачи прогноза на любое число шагов вперед может быть получено рекуррентным способом по формуле

$$\hat{x}(n+1) = \sum_{j=1}^l \hat{\alpha}_j \hat{x}(n-j) + \sum_{j=l+1}^p \hat{\alpha}_j x(n-j). \tag{3.4}$$

Следует учесть, что оценки прогноза (3.4) и коэффициенты модели  $\hat{\alpha}_k$ , полученные в ходе прогнозирования на один шаг, не являются оптимальными в случае прогноза на большее число шагов,  $k \geq 1$ . Более естественной мерой погрешности является величина нормы

$$e_n = \|\mathbf{x}_n - \hat{\mathbf{x}}_n\|,$$

где  $\mathbf{x}_n = (x_n, x_{n+1}, \dots, x_{n+k})^T, x_{n+l} \equiv x(n+l)$  и  $\hat{\mathbf{x}}_n = (\hat{x}_n, \hat{x}_{n+1}, \dots, \hat{x}_{n+k})^T, \hat{x}_{n+l} \equiv \hat{x}(n+l)$ .

Пусть задача прогноза заключается в нахождении  $k+1$  оценок  $\hat{x}(n+l)$  при значениях  $l = 0, 1, \dots, k$ .

В соответствии с уравнением (3.4), оценки  $\hat{x}(n+l)$  находятся последовательно, начиная со значения  $l = 0$  и кончая  $l = k$ .

Если ввести вектор коэффициентов модели  $\mathbf{a} = (a_0, a_1, \dots, a_r)^T$ , матрицу

$$\hat{\mathbf{X}}_n = \begin{pmatrix} x_{n-1} & x_{n-2} & x_{n-3} & \cdots & x_{n-p} \\ \hat{x}_n & x_{n-1} & x_{n-2} & \cdots & x_{n-p+1} \\ \hat{x}_{n+1} & \hat{x}_n & x_{n-1} & \cdots & x_{n-p+2} \\ \vdots & \vdots & \vdots & & \vdots \\ \hat{x}_{n+p-2} & \hat{x}_{n+p-3} & \hat{x}_{n+p-4} & \cdots & x_{n-1} \\ \hat{x}_{n+p-1} & \hat{x}_{n+k-2} & \hat{x}_{n+p-3} & \cdots & \hat{x}_n \\ \vdots & \vdots & \vdots & & \vdots \\ \hat{x}_{n+k-1} & \hat{x}_{n+k-2} & \hat{x}_{n+k-3} & \cdots & \hat{x}_{n+k-p} \end{pmatrix}$$

и вектор прогноза  $\hat{\mathbf{x}}_n = (\hat{x}_n, \hat{x}_{n+1}, \dots, \hat{x}_{n+k})^T$ , то уравнение (3.4) на интервале  $[n, n+k]$  может быть записано в виде

$$\hat{\mathbf{x}}_n = \hat{\mathbf{X}}_n \mathbf{a}. \quad (3.5)$$

Если ввести вектор данных  $\mathbf{y}_n = (y_n, y_{n+1}, \dots, y_{n+k})^T$ , наблюдаемых на интервале  $[n, n+k]$  прогнозирования, и матрицу  $\mathbf{Y}_n$  полученную из матрицы  $\hat{\mathbf{X}}_n$  посредством замены её компонентов  $\hat{x}_{n+l}$  и  $x_{n+l}$  на наблюдаемые данные  $y_{n+l}$ , то коэффициенты модели в соответствии с (3.5) можно найти из уравнения

$$\mathbf{y}_n = \mathbf{Y}_n \mathbf{a}. \quad (3.6)$$

Исходя из уравнения (3.5) коэффициенты модели можно также найти посредством минимизации функционала

$$J(\mathbf{a}) = \|\mathbf{y}_n - \hat{\mathbf{x}}_n\|^2 = \|\mathbf{y}_n - \mathbf{X}_n \mathbf{a}\|^2. \quad (3.7)$$

где компоненты матрицы  $\hat{\mathbf{X}}_n$ , отвечающие прогнозируемым значениям  $\hat{x}_{n+l}$  временного ряда, должны быть представлены в форме явных функций от коэффициентов модели.

Учитывая, что второй и последующие компоненты вектора, определяемого произведением  $\hat{\mathbf{X}}_n \mathbf{a}$ , являются нелинейными функциями коэффициентов модели, минимизация функции (3.7) требует применения численных методов.

Явный вид зависимости оценок прогноза от коэффициентов модели можно получить посредством рекуррентного решения уравнения (3.5).

Введём переменные

$$\sigma_l = \sum_{j=l+1}^p \alpha_j x_{n-j+l}, l = 0, 1, \dots, p-1.$$

Можно показать, что оценки прогноза являются линейными комбинациями этих переменных, то есть

$$\hat{x}_{n+l} = \sum_{j=0}^l d_{l-j+1} \sigma_j, l = 0, 1, \dots \quad (3.8)$$

Коэффициенты уравнения (3.8) находятся при этом рекуррентным способом:

$$d_l = \sum_{j=1}^l \alpha_j d_{l-j}, d_0 = 0, d_1 = 1.$$

Так, например, отсюда получаем

$$d_2 = \alpha_1,$$

$$d_3 = \alpha_1^2 + \alpha_2,$$

$$d_4 = \alpha_1^3 + 2\alpha_1\alpha_2 + \alpha_3,$$

$$d_5 = \alpha_1^4 + 3\alpha_1^2\alpha_2 + 2\alpha_1\alpha_3 + \alpha_2^2 + \alpha_4, \dots$$

Здесь следует отметить, что если величина  $l \geq p$ , то переменная  $\sigma_l = 0$  и оценка прогноза принимает вид

$$\hat{x}_{n+l} = \sum_{j=0}^p d_{l-j+1} \sigma_j.$$

Также могут быть получены аналогичные выражения оценки прогноза для модели авторегрессии – скользящего среднего. Оценивание коэффициентов модели временного ряда, как с помощью уравнения (3.6), так и посредством минимизации функционала (3.7), предполагает, что наблюдаемые данные порождаются детерминированным процессом. В стохастическом случае возможно применение методов корреляционного анализа.

#### 4. Нейросетевые методы прогнозирования

Горизонт предсказания любой модели прогноза зависит от того, насколько точно эта модель воспроизводит динамику временного ряда или системы, порождающей наблюдаемый процесс. В этой связи возникает вопрос о

горизонте предсказания моделей в форме искусственных нейронных сетей прямого распространения. Проблема здесь заключается в том, насколько точно динамика процесса может быть представлена весовыми коэффициентами сети. В этом случае отсутствие обратных связей компенсируется алгоритмами обучения типа обратного распространения ошибки. Алгоритм функционирования многослойной нейронной сети прямого распространения при прохождении сигналов по направлению от входа к выходу зададим уравнениями в матричной форме. Введем вектор выходов  $j$  – слоя сети, состоящего из  $p_j$  нейронов,

$$\mathbf{y}_j = (y_{j,1}, y_{j,2}, \dots, y_{j,p_j})^T, \quad (4.1)$$

и вектор весовых коэффициентов  $l$  – нейрона в  $j$  – слое сети

$$\mathbf{w}_{j,l} = (w_{j,l,1}, w_{j,l,2}, \dots, w_{j,l,p_j})^T. \quad (4.2)$$

Тогда синаптическая сумма  $l$  – нейрона в  $j$  – слое сети

$$s_{j,l} = \mathbf{w}_{j,l}^T \mathbf{y}_{j-1} + w_{j,l,0}, \quad (4.3)$$

где  $w_{j,l,0}$  – смещение нейрона.

При этом выходы  $j$  – слоя сети можно представить вектором

$$\mathbf{y}_j = \begin{pmatrix} \varphi(\mathbf{w}_{j,1}^T \mathbf{y}_{j-1} + w_{j,1,0}) \\ \varphi(\mathbf{w}_{j,2}^T \mathbf{y}_{j-1} + w_{j,2,0}) \\ \vdots \\ \varphi(\mathbf{w}_{j,p_j}^T \mathbf{y}_{j-1} + w_{j,p_j,0}) \end{pmatrix}, \quad (4.4)$$

где  $\varphi(\cdot)$  – функция активации нейронов сети.

Формула (4.4) является рекуррентным уравнением, позволяющим найти последовательно выходы всех слоёв сети, начиная с первого слоя,  $j=1$ , и кончая последним слоем, совпадающим с выходом сети. Вектор  $\mathbf{y}_0$  при этом – это известный выход сенсорного слоя сети.

При решении задач прогноза выходами сети в пределах одного цикла её функционирования являются результаты предсказания временного ряда или процесса на заданное число шагов вперёд, начиная с 1 и кончая  $P_{\text{вых.сл.}}$  (число нейронов в выходном слое сети).

Алгоритм обучения нейронной сети методом обратного распространения ошибки также можно представить уравнениями в матричной форме.

Введём матрицу весов  $j$  – слоя сети

$$\mathbf{W}_j = (\mathbf{w}_{j,1}, \mathbf{w}_{j,2}, \dots, \mathbf{w}_{j,p_j}), \quad (4.5)$$

вектор смещения нейронов  $j$  – слоя

$$\mathbf{w}_{j,0} = (w_{j,1,0}, w_{j,2,0}, \dots, w_{j,p_j,0})^T, \quad (4.6)$$

вектор синаптических сумм

$$\mathbf{s}_j = (s_{j,1}, s_{j,2}, \dots, s_{j,p_j})^T, \quad (4.7)$$

и векторную функцию активации  $j$  – слоя

$$\boldsymbol{\varphi}(\mathbf{s}_j) = (\varphi(s_{j,1}), \varphi(s_{j,2}), \dots, \varphi(s_{j,p_j}))^T. \quad (4.8)$$

Весовые коэффициенты  $j$  – слоя сети находятся при этом в ходе минимизации функционала

$$J_j(\mathbf{e}_j) = (\boldsymbol{\varphi}(\mathbf{s}_j) - \bar{\mathbf{y}}_j)^T (\boldsymbol{\varphi}(\mathbf{s}_j) - \bar{\mathbf{y}}_j), \quad (4.9)$$

где  $\bar{\mathbf{y}}_j = (\bar{y}_{j,1}, \bar{y}_{j,2}, \dots, \bar{y}_{j,p_j})^T$  – вектор требуемых выходов  $j$  – слоя сети и  $\boldsymbol{\varphi}(\mathbf{s}_j) - \bar{\mathbf{y}}_j = \mathbf{e}_j$  – вектор ошибок по выходам  $j$ –слоя.

Если ввести матрицы

$$\mathbf{W}_j = \frac{\partial \mathbf{s}_j^T}{\partial \mathbf{y}_{j-1}}, \quad (4.10)$$

$$\boldsymbol{\Phi}_j = \frac{\partial \boldsymbol{\varphi}(\mathbf{s}_j)}{\partial \mathbf{s}_j} = \text{diag} \left( \frac{\partial \varphi(s_{j,1})}{\partial s_{j,1}}, \frac{\partial \varphi(s_{j,2})}{\partial s_{j,2}}, \dots, \frac{\partial \varphi(s_{j,p_j})}{\partial s_{j,p_j}} \right),$$

то алгоритм обучения методом обратного распространения ошибки принимает вид разностного уравнения

$$\mathbf{e}_{j-1} = \mathbf{W}_j \boldsymbol{\Phi}_j \mathbf{e}_j, \quad (4.11)$$

с начальными условиями

$$\mathbf{e}_j = \boldsymbol{\varphi}(\mathbf{s}_j) - \bar{\mathbf{y}}_j, \text{ при } j = j_{\text{вых.сл.}} \text{ (номер выходного слоя сети), } \bar{\mathbf{y}}_j \text{ – вектор требуемых выходов } j\text{-го слоя сети.}$$

Коррекция весовых коэффициентов сети по ошибкам, полученным с помощью уравнения (4.11), осуществляется, если воспользоваться градиентным методом, по итерационным формулам

$$\begin{aligned}
\mathbf{w}_{j,l}(k+1) &= \mathbf{w}_{j,l}(k) - \alpha h_{j,l}(s_{j,l}) e_{j,l} \mathbf{y}_{j-1}, \\
\mathbf{w}_{j,0}(k+1) &= \mathbf{w}_{j,0}(k) - \alpha \Phi_j \mathbf{e}_j, \\
h_{j,l}(s_{j,l}) &= \frac{\partial \varphi(s_{j,l})}{\partial s_{j,l}},
\end{aligned}
\tag{4.12}$$

где  $l=1,2,\dots,p_j$ ,  $e_{j,l}$  является  $l$ -компонентом вектора  $\mathbf{e}_j$  ошибки, шаг настройки весовых коэффициентов  $\alpha$  выбирается в диапазоне  $0 < \alpha < 1$ .

Таким образом алгоритм (4.1 ÷ 4.12) обучения нейронной сети при прогнозировании изменения значений функции, реализуемый нейронной сетью прямого распространения, построенной по правилу многослойного персептрона, сводится в таблицу 1 [9].

| Этап обучения                         | Алгоритм   |
|---------------------------------------|--|
| Определение выходов (прямой проход)   | $ \mathbf{y}_j = \begin{pmatrix} \varphi(\mathbf{w}_{j1}^T \mathbf{y}_{j-1} + w_{j01}^j) \\ \varphi(\mathbf{w}_{j2}^T \mathbf{y}_{j-1} + w_{j02}^j) \\ \vdots \\ \varphi(\mathbf{w}_{jp_j}^T \mathbf{y}_{j-1} + w_{jp_j}^j) \end{pmatrix}, \quad j = 1, 2, \dots; \quad y_0 = C_m. $   |
| Определение ошибок (обратный проход)  | $ \mathbf{e}_{j-1} = \mathbf{W}_j \Phi_j \mathbf{e}_j, \quad e_j = \varphi(s_j) - \bar{y}_j, $ $ s_j = \mathbf{W}_j^T \mathbf{y}_{j-1} + w_{j0} = (s_{j1}, s_{j2}, \dots, s_{jp_j})^T, $ <p>где <math>\mathbf{W}_j = (w_{j1} \ w_{j2} \ \dots \ w_{jp_j})</math> - матрица весовых синаптических коэффициентов.</p>  |
| Коррекция синаптических коэффициентов | $ w_{j,l}(k+1) = w_{j,l}(k) - \alpha h_{j,l}(s_{j,l}) e_{j,l} y_{j-1}, $ $ w_{j0}(k+1) = w_{j0}(k) - \alpha \Phi_j e_j, $ $ \Phi_j \equiv \text{diag} \left( \frac{\partial \varphi(s_{j1})}{\partial s_{j1}}, \frac{\partial \varphi(s_{j2})}{\partial s_{j2}}, \dots, \frac{\partial \varphi(s_{jp_j})}{\partial s_{j,p_j}} \right) = \frac{\partial \varphi^T(s_j)}{\partial s_j}, \quad h_{j,l}(s_{j,l}) = \frac{\partial \varphi(s_{j,l})}{\partial s_{j,l}}, $ <p><math>\alpha</math> - шаг настройки, выбирается в диапазоне <math>0 &lt; \alpha &lt; 1</math>.</p> |

## 5. Структура прогнозирования на трехслойном персептроне.

Структурная схема, реализующая нейросетевое прогнозирование изменений значений параметров функции с её предварительной вейвлет-обработкой, представлена на рис. 1 [12].

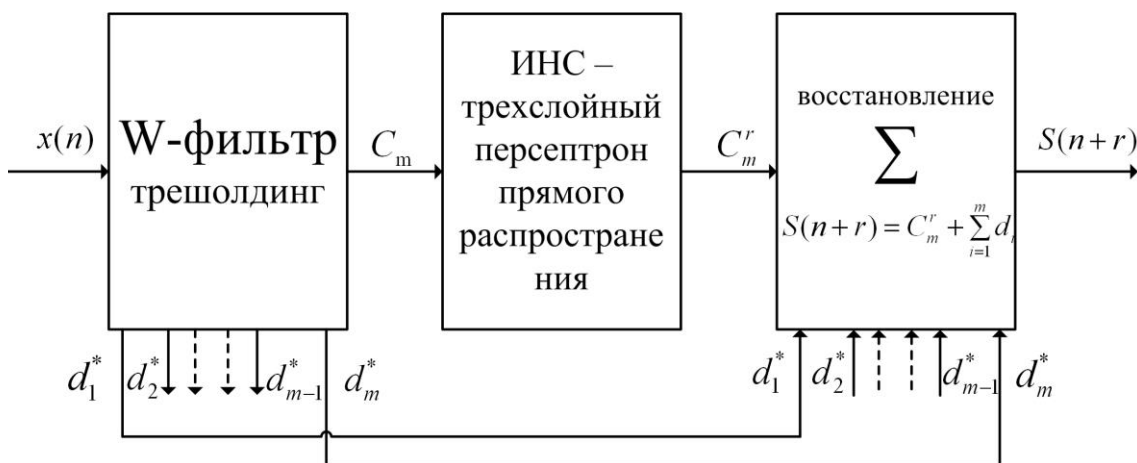


Рис. 1. Структурная схема нейросетевого прогнозирования изменений значений функции.

Как видно из рис.1, система прогнозирования, реализованная на трехслойном персептроне прямого распространения, используется применительно к задаче мониторинга информации в финансовых системах, которая формирует временной ряд  $x(n)$  отсчетов значений экономических параметров. Аналогично могут быть исследованы непрерывные функции, либо временные ряды данных, отображающие информацию изменения параметров различных процессов для решения задач прогнозирования в экономических системах.

Входные сигналы в виде непрерывной функции  $x(t)$  или временного ряда  $x(n)$  подаются на W-фильтр предварительной обработки вейвлет-преобразованием. В W-фильтре формируются аппроксимирующие коэффициенты  $C_i$ , вычисленные по формуле  $C_i = \frac{1}{p} C_{i-1} \varphi_i(2^i t - n)$ , где  $\varphi_i(2^i t - n)$  - скейлинг-функция, и детализирующие коэффициенты  $d_i$ ,  $i$ -ого уровня вейвлет-разложения, вычисленные по формулам  $d_1 = \frac{1}{p} x(n) \psi_1(2t - n)$ ,  $d_i = \frac{1}{p} \sum_{i=2}^m C_{i-1} \psi_i(2^i t - n)$ , где  $\psi_i(2^i t - n)$  - вейвлет-функция [8,14,15].

Выделенные аппроксимирующие коэффициенты  $C_i$  до уровня  $i=m$ , где  $1 \leq m \leq 6$  [7, 13], подаются на вход  $N$ -разрядного регистра сдвига, построенного на элементах задержки с передаточной функцией  $Z^{-1}$ , в котором формируется выборка входных сигналов нейронной сети в виде движущегося окна данных из  $N$  отсчетов, обработанных в W-фильтре. В нейронной сети, построенной по правилу многослойного персептрона прямого распространения вычисляются выходные сигналы в виде аппроксимирующих коэффициентов прогноза  $C_m^{*r}$ , где  $r$ - число периодов прогноза ( $r = 1, 2 \dots, 10$ ).

Многослойный персептрон состоит из входного слоя,  $j$ -внутренних слоев, где  $j \in \{0, 1\}$ , и выходного слоя. При  $j=1$ , структурная схема персептрона представляет собой трехслойный персептрон. Увеличение числа слоев, т.е.  $j > 1$ , в персептроне приводит к увеличению вычислительных затрат при незначительном уменьшении погрешности [10,13].

Число нейронов во входном слое персептрона влияет на число анализируемых входных аппроксимирующих коэффициентов временного ряда  $C_m(n)$ , соответственно на время анализа предыдущих значений отсчетов исследуемой функции, на длительность движущегося окна, на число разрядов во входном регистре сдвига, на общую погрешность прогнозирования. В соответствии с [10] погрешность прогноза на  $r = 10T$  достигает  $|\delta| \leq 5\%$ , при числе нейронов во входном слое  $p_{\text{вх}}=64$ , внутреннем слое  $p_{\text{внутр}}=10$ , выходном слое  $p_{\text{вых}}=10$ . При уменьшении числа нейронов во входном слое до  $p_{\text{вх}}=32$ , погрешность повышается до  $|\delta| \leq (7 \div 10)\%$  [13].

Число нейронов внутреннего слоя  $p_{\text{внутр}}$  и число нейронов выходного слоя  $p_{\text{вых}}$  определяется числом периодов прогноза  $r$  и, соответственно, подчиняется условию  $p_{\text{внутр}} = p_{\text{вых}} = r_{\text{max}}$ .

Полученные аппроксимирующие коэффициенты прогноза  $C_m^{*r}$ , с учетом необходимых для разных выходов нейронной сети задержек, сравниваются с входными значениями  $C_m$  на входе регистра, что обеспечивает формирование ошибок предсказаний  $e^{*r}$  на всех  $r$  выходах нейронов выходного слоя нейронной сети. При этом формализуем процедуру вычисления нейронной сетью аппроксимирующих коэффициентов прогноза, вычисление ошибок на выходах нейронов каждого из слоев, а также процедуру адаптивной настройки синаптических весовых коэффициентов, для каждого 1-го нейрона в рамках всех слоев нейронной сети.

Аппроксимирующий коэффициент прогноза на один период вычисляется на выходе 1-го нейрона выходного слоя в соответствии с выражением [13]:  $C_m^{*1} = y_{31} = \varphi_3(s_{10}) = \frac{1}{1 + \exp(-s_{10})}$ .

На выходе 1-го нейрона выходного (3-его) слоя вычисляется ошибка предсказания на один период -  $e_{31}^{*1}$ , по которой вычисляется ошибка на выходе первого нейрона предыдущего второго слоя нейронной сети -  $e_{21}$ . Ошибки на выходе 1-

ых нейронов 2-го и первого слоев вычисляются в соответствии с выражениями:  $e_{21} = \sum_{l=1}^{10} w_{3l} \cdot e_{31}$ ,  $e_{11} = \sum_{l=1}^{64} w_{2l} \cdot e_{21}$ .

Вычисленные таким образом ошибки обратного прохода, в свою очередь адаптивно перестраивают синаптические коэффициенты связей между нейронами различных слоев. Таким образом, веса 1-х нейронов каждого слоя адаптивно настраиваются в соответствии с выражениями:

$$w'_{31} = w_{31} + \alpha e_{31} \frac{d\varphi_3(s_{10})}{ds_{10}} C_m^{*1},$$

$$w'_{21} = w_{21} + \alpha e_{21} \frac{d\varphi_2(s_{10})}{ds_{10}} y_{21},$$

$$w'_{11} = w_{11} + \alpha e_{11} \frac{d\varphi_1(s_{64})}{ds_{64}} y_{11}.$$

Минимизация ошибки обеспечивается выше рассмотренной итерационной процедурой обучения нейронной сети, в ходе которой осуществляется настройка весовых или синаптических коэффициентов сети. Таким образом, задача предсказания заключается в оценивании будущих значений процесса по имеющимся данным  $C_m$  в текущий момент, в настройке значений коэффициентов нейронной сети по критерию минимальной величины ошибки предсказания, в оценивании управляемого объекта, основанных на сравнении выходных сигналов управляемого объекта  $C_m^{*r}$  и его модели, в качестве которой используется нейронная сеть.

Детализирующие коэффициенты  $d_i$ , вычисленные до уровня  $m$  ( $d_1, d_2, \dots, d_m$ ), после обработки алгоритмом сглаживания, поступают на блоки восстановления выходного временного ряда прогноза. Алгоритм сглаживания детализирующих коэффициентов по критерию адаптивного штрафного порога имеет вид [10, 11]:

$$d_i^* = \begin{cases} d_i, & d_i \geq \beta_0; \\ 0, & d_i < \beta_0, \end{cases}$$

где  $\beta_0 = b\sqrt{\sigma^2}$ ,  $\sigma^2$  - дисперсия шумовых составляющих во входном сигнале  $x(n)$ ,  $b$ -коэффициент пропорциональности,  $b = \sqrt{2 \ln N}$ ,  $N$  - число периодов анализа.

Выходные сигналы с выходов нейронной сети в виде выходных аппроксимирующих коэффициентов  $C_m^{*r}$  (где  $r$ -номер выхода нейронной сети, в соответствии с числом периодов прогноза  $r \in \{1 \div 10\}$ ) также поступает на  $r$ -тый блок восстановления выходного временного ряда прогноза. На выходе  $r$ -блоков восстановления формируется выходные сигналы прогноза в виде временного ряда  $s(n+rT)$  [10, 11],

$$s(n+r) = \frac{1}{p} \left[ \sum_n x(n) \psi_1(2t-n) + \sum_{i=2}^m C_{i-1} \cdot \psi_i(2^i t - n) + \sum_n C_m^{*r} \right], \quad (5.1)$$

где  $rT$ -время прогноза.

Также в структурной схеме представлен алгоритм вейвлет обработки временного ряда  $x(n)$  и алгоритм восстановления выходного сигнала  $s(k)$  с пониженной погрешностью, в соответствии с выражением восстановления [10, 11].

$$s(n) = \frac{1}{p} \left[ \sum_n x(n) \psi_1(2t-n) + \sum_{i=2}^m C_{i-1} \cdot \psi_i(2^i t - n) + \sum_n C_m \right], \quad (5.2)$$

В полученных выражениях (4.1) и (4.2), благодаря предварительной вейвлет-обработки в W-фильтре, существенно ослабляются флуктуации входного сигнала  $x(n)$  за счет формирования аппроксимирующих коэффициентов  $m$ -го уровня  $C_m$  и подавляются компоненты шума, имеющие место во входном сигнале, путем обработки детализирующих коэффициентов алгоритмом сглаживания, что заметно снижает погрешность представления информации. Таким образом, получены очищенный от помех обработанный выходной временной ряд в реальном времени и обработанный выходной временной ряд прогноза, которые представляют информацию об экономических показателях финансовой системы в устройствах отображения и в системах принятий решений с более низкой погрешностью.

Разработанные алгоритмы нейросетевого прогнозирования изменения значения функций могут быть реализованы в различных программных средах, таких как: MatLab Neural Network Toolbox, PyBrain, и NeurophStudio.

Ошибки прогноза зависят от того, насколько размеры входного слоя ИНС соответствуют характерной величине интервала временного ряда, по которому можно восстановить его динамику, что подтверждается графиками на рис. 2 [2].

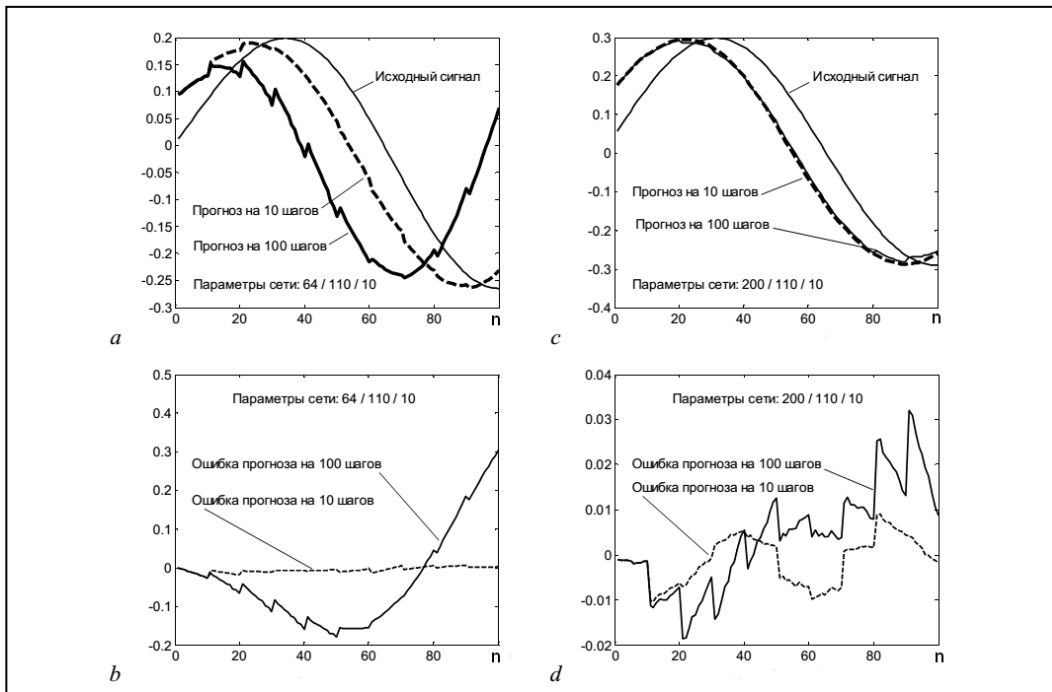


Рис. 2. Результаты и ошибки прогноза при различных размерах сенсорного слоя.

При моделировании, результаты которого приведены рис. 2, в качестве функции активации выбран гиперболический тангенс.

Из рис.2 видно, что увеличение числа сенсорных нейронов способствует увеличению горизонта и точности прогноза. Если для сети с 64-мя сенсорными нейронами горизонт предсказания может быть принят равным в пределах  $10 \div 20$  (рис. 2 a, b), то в случае 200 сенсорных нейронов (рис. 2 c, d), он увеличивается до 100. Можно также отметить, что



погрешность прогнозирования зависит не только от числа сенсорных нейронов, но и от длительности движущегося временного окна или почти периода исследования сигнала и числа выборок сигнала на этом периоде.

## Заключение

Существующие методы прогнозирования временных рядов и непрерывных процессов характеризуются большим разнообразием, что в свою очередь, обусловлено большим разнообразием задач. Так, велико число работ по методам прогноза, основанным на идеях декомпозиции процессов по эмпирическим модам (метод EMD). Однако данный подход применим, как правило, только при условии, что наблюдаемый процесс представлен небольшим числом составляющих.

Метод параметрической регрессии, используемый в ряде задач аппроксимации, обладает недостатком, заключающимся в возможности возникновения неконтролируемой ошибки на интервале прогнозирования, особенно при высоком порядке прогнозирующей функции.

Особый интерес в плане прогнозирования связан с представлением временных рядов моделями авторегрессии и нейронными сетями. В работе показано, что в рамках модели авторегрессии прогноз на несколько шагов вперед зависит от коэффициентов модели нелинейным образом. При этом задачу прогноза, состоящую в нахождении коэффициентов модели посредством минимизации целевой функции, можно решить итерационным методом. Поэтому интерес представляют рекуррентные и другие сети с обратными связями. В данной работе рассмотрены искусственные нейронные сети, в частности рассмотрены нейронные сети на персептроне прямого распространения. В работе получены уравнения функционирования и обучения ИНС в матричной форме, получен алгоритм обратной подстановки, с помощью которого можно увеличить глубину прогноза. Результаты моделирования показали эффективность разработанных алгоритмов. Результат применения разработанных алгоритмов заключается в повышении эффективности прогнозирования изменений значений функций по критериям длительности и минимизированной погрешности получения прогноза, а также быстрейшего действия, адаптивности системы при изменяющихся условиях. Техническим результатом является также возможность гибкого изменения архитектуры нейронной сети в случае изменения требований на длительность прогноза.

Предложенная структурная схема реализации нейросетевого прогнозирования изменений параметров временных рядов и непрерывных функций с предварительной вейвлет-обработкой обеспечивает возможность более эффективного мониторинга исследуемых процессов, например в экономических системах.

## Литература

- [1] Кравцов, Ю.Н. Случайность, детерминированность, предсказуемость // Успехи физических наук. – 1989. – Т. 158, Вып. 1. – С. 93 – 122.
- [2] Ермолаев, В.А. О методах прогнозирования временных рядов и непрерывных процессов // Радиотехнические и телекоммуникационные системы. – 2016. – №2. – С.52-63.
- [3] Бокс, Дж. Анализ временных рядов. Прогноз и управление. В 2-х выпусках./ Бокс Дж., Дженкинс Г – М.: Мир, 1974. Вып. 1 – 408 С., Вып. 2 – 199 С.
- [4] Маевский, В.В. Робастность регрессионного прогнозирования при наличии функциональных искажений модели / В.В. Маевский, Ю.С. Харин // АиТ, 2012. – №11, – С. 118 – 137.
- [5] Huang, N.E. The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis / Huang N.E., Shen Z., Long S.R., Wu M.C., Shin H.H., Zheng Q., Yen N.-C., Tung C.C., Liu H.H. // Proc. R. Soc. Lond. A – 1998, 454. – P. 903 – 995.
- [6] Ермолаев, В.А. Метод интерполяционной фильтрации в задачах обработки речевых сигналов во временной области / В.А. Ермолаев, О.Е. Карасев, Ю.А. Кропотов // Вестник компьютерных и информационных технологий. – 2008. – № 7. – С. 12 – 17.
- [7] Кропотов, Ю.А. Об адаптивном алгоритме наименьших квадратов в задачах компенсации акустического эха / Ю.А. Кропотов, В.А. Ермолаев // Радиотехнические и телекоммуникационные системы, 2012. – №2(6). – С. 40-44.
- [8] Дремин, И.М. Вейвлеты и их использование / И.М. Дремин, О.В. Иванов, В.А. Нечитайло // Успехи физических наук – 2001. – Т. 17, №5. – С. 465-501.
- [9] Ермолаев, В.А. Нейрокомпьютерные системы. Учебное пособие для студентов образовательной программы 230105.65 Программное обеспечение вычислительной техники и автоматизированных систем / В.А. Ермолаев, А.Ю. Проскуряков – Муром: Изд.полиграфический центр МИ ВлГУ, 2011. – 120 с.
- [10] Проскуряков, А.Ю. Модель прогнозирования временных рядов на трехслойном персептроне в информационной системе мониторинга загрязняющих выбросов // Радиотехнические и телекоммуникационные системы. – 2014. – №2. – С.57-63.
- [11] Кропотов, Ю.А. Информационно-управляющие телекоммуникационные системы аудиообмена и автоматизированного мониторинга / Ю.А. Кропотов, Д.В. Бейлекчи, А.А. Белов, В.А. Ермолаев, О.Е. Карасев, А.А. Колпаков, А.Н. Коноплев, А.Ю. Проскуряков // Радиотехнические и телекоммуникационные системы. – 2014. – № 1. – С. 99-113.
- [12] Патент на изобретение № 2600099. Способ нейросетевого прогнозирования изменения значений функции с ее предварительной вейвлет-обработкой и устройство его осуществления / А.А. Белов, В.А. Ермолаев, Ю.А. Кропотов, А.Ю. Проскуряков, Заявка № 2015110284/08 (016173) от 23.03.2015. Приоритет изобретения 23 марта 2015 г. Зарегистрировано в Государственном реестре изобретений Российской Федерации от 22 сентября 2016 г.
- [13] Проскуряков, А.Ю. Алгоритмы автоматизированных систем экологического мониторинга промышленных производств / А.Ю. Проскуряков, А.А. Белов, Ю.А. Кропотов // Монография. М.-Берлин: Директ-Медиа. – 2015. – 121 с.
- [14] Яковлев, А.Н. Введение в вейвлет-преобразование: учебное пособие / А.Н. Яковлев. – Новосибирск: Издательство Новосибирского технического университета. – 2003. – 102 с.
- [15] Kropotov, Yu.A. Identification of models for discrete linear systems with variable, slowly varying parameters / Yu.A. Kropotov, V.A. Ermolaev, V.T. Eremenko, O.E. Karasev // Journal of Communications Technology and Electronics. . – 2010. – Vol. 55, Issue 1. – P. 52-57.