УДК 681.3 doi: 10.26102/2310-6018/2018.23.4.008

И.Н. Крючкова, Е.Е. Красновский, Е.В. Болнокина, О.Я. Кравец

АЛГОРИТМЫ ИССЛЕДОВАНИЯ МНОГОМЕРНЫХ ВРЕМЕННЫХ РЯДОВ С УЧЕТОМ ОТСРОЧЕННОГО ВЛИЯНИЯ ФАКТОРОВ НА ОСНОВЕ МАТЕМАТИЧЕСКОГО МОДЕЛИРОВАНИЯ

Воронежский государственный технический университет
Московский государственный технический университет имени Н.Э.Баумана

Исследованы модели и методы нейросетевого моделирования динамики на основе анализа многомерных временных рядов с учетом отсроченного влияния значимых факторов. В связи с невозможностью одновременного определения оптимального временного лага и обучения сети необходимо рассматривать нахождение многомерного лага как отдельную оптимизационную задачу. Изложена математическая постановка задачи построения нейросети для ненулевого запаздывания, приведено описание особенностей оптимизации величины запаздывания для одной независимой переменной (входа), конкретизирована информационная база моделирования и прогнозирования и нейросетевые алгоритмы обработки данных, проведена алгоритмизация множественного регрессионного анализа с оптимизацией вектора запаздываний для значимых факторов. Принципиальная возможность применения анализа чувствительности для нахождения оптимального многомерного временного лага была подтверждена в ходе вычислительного эксперимента. Анализ чувствительности проводился на тестовых данных, полученных расчетом значений наборов функций нескольких переменных с известным запаздыванием по некоторым переменным. Анализ ошибок обучения, обобщения и прогнозирования на исходных и смещенных рядах позволил сделать вывод о существенном снижении ошибки обучения и ошибки прогнозирования на смещенных рядах при практически неизменной ошибке обобщения, что свидетельствует об эффективности предложенного алгоритма и отсутствии структурных эффектов в изменении качества прогноза.

Ключевые слова: математическое моделирование, нейронные сети, запаздывание, прогноз.

Введение

Существует большой класс объектов, которые описываются многомерными временными рядами. Особенностью объектов является априорное существование зависимости результирующего временного ряда от факторных. Однако зависимость не прямая, а «смещенная» - «последействие» возникает не немедленно, а с некоторым лагом. Задачи с запаздыванием для немногомерных временных рядов широко известны, наш случай – обобщенный. Как следствие, нет прямых способов решения такой задачи (в т.ч. выбора вектора оптимальных лагов), и оправданным является применение нейросетевых технологий как логический базис разработки нейросетевых алгоритмов решения задач [8].

Задача прогнозирования временного ряда в нейросетевом логическом базисе [8] рассматривается в следующей формулировке [1-4]. Имеются M временных рядов X_1, X_2,…X_M по N точек наблюдений в каждом. Имеется исследуемая переменная Y - такой же по параметрам временной ряд, зависящий от рядов наблюдений. Необходимо построить и обучить нейросеть, реализующую отображение $Y=\varphi(X_1,X_2,\dots X_M)$ с заданной степенью точности, на массиве исходных данных, содержащих значения наблюдений в моменты времени t_i, (i=0,1…..N-1). Обученная сеть используется для предсказания значения $Y(t+\Delta_Y)$ на основе значений рядов $Xj(tN-\Delta_X)$, где Δ_Y – период предсказания, Δ_X – временной лаг.

Материалы и методы

Согласно следствию из теоремы Колмогорова - Арнольда - Хехт-Нильсена [5], для любого множества пар входных-выходных векторов произвольной размерности существует однородная двухслойная нейронная сеть с последовательными связями, с сигмоидальными функциями активации [6] и с конечным числом нейронов, которая для каждого входного вектора формирует соответствующий ему выходной вектор. Поэтому в качестве функции активации нейронов используется сигмоида:

$$f(x)=\frac{1}{1+e^{-x}}. \qquad (1)$$

Обучение двухслойной нейронной сети при наличии временного лага представляет собой процесс поиска решения задачи минимизации целевой функции, называемой функцией ошибки, в виде:

$$\min_{\forall(w_i,\Delta_j)} E=\frac{1}{2}\sum_{p=1}^{N}(y_p-f_p(\sum_{i=1}^{K}w_{ki}^{(2)}\cdot(f(\sum_{j=0}^{M}w_{ij}^{(1)}\cdot(x_j(t-\Delta_j))))))^2, \qquad (2)$$

где N - количество обучающих примеров, M - количество входов, $w_{ij}^{(1)}$ - вес i-го нейрона скрытого слоя, $w_{ki}^{(2)}$ - вес i-го нейрона выходного слоя, K - количество нейронов в скрытом слое, $x_0=1$, $w_0^{(1)}=1$, $w_0^{(2)}=1$, Δ_j - временной лаг для j-го входа, $x_j(t-\Delta_j)$ - смещенное на величину временного лага значение независимой переменной.

Сформулируем ограничения на Δ_j.

1. Временной лаг Δ_j должен быть неотрицательным, что определяется физическим смыслом задачи. Действительно, изменение независимой переменной в момент времени t+1 не может влиять на значение зависимой переменной в момент времени t, отсюда $\Delta_j\geq 0$.

2. Смещение точки наблюдений в сторону более ранних времен не должно приводить к выходу за нижнюю границу исходного временного ряда, так как явление, описываемое данным временным рядом, может

отсутствовать в этот момент как таковое, и экстраполированные значения не будут иметь физического смысла. Следовательно, $x_j(t-\Delta_j)\in X_j$.

Добавление в выражение целевой функции переменной Δ_j не позволяет использовать для обучения ни один из стандартных алгоритмов, основанных на подборе весовых коэффициентов, так как наличие временного лага приводит к неопределенности выбора обучающих векторов. Следовательно, в связи с невозможностью одновременного определения оптимального временного лага и обучения сети необходимо рассматривать нахождение многомерного лага как отдельную оптимизационную задачу.

Результаты

Вектор оптимальных временных лагов

Задача построения прогнозирующей нейросети с учетом запаздывания сведена к задаче построения вектора оптимальных временных лагов и задания с его помощью смещенных значений независимых переменных $x_j(t-\Delta_j)$, входящих в выражение (1). Тогда задача в постановке (1) может быть решена как стандартная задача обучения нейросети одним из известных методов, например, методом обратного распространения ошибки [7], реализованным практически во всех программных средствах нейросетевого моделирования.

Пусть учитываемые «прошлые» значения независимой переменной X_i описываются подмножеством входов $S_i=\{x_i(t), x_i(t-1),\dots x_i(t-\Delta_s),\dots x_i(t-h_i)\}$ двухслойной нейросети типа «многослойный персептрон» структуры M^*-К-1. Здесь К - количество нейронов в скрытом слое, Δ_s - временной лаг, h_i - глубина выборки по i-му фактору, M^* - количество входов нейросети, определяемое по формуле:

$$M^* = \sum_{i=1}^{M}(1+h_i), \qquad (3)$$

где M - количество учитываемых факторов. Данная формула позволяет определить состав входов при произвольной глубине погружения для каждого фактора. При h=0 значения независимых переменных в предыдущие моменты времени не включаются в состав входного вектора [2].

Тогда чувствительность нейросети к удалению входа, соответствующего значению независимой переменной $x_i(t-\Delta_s)$, может быть описана модулем абсолютной ошибки прогнозирования в виде:

$$E_{\Delta is} = \left| y_p - f(\sum_{r=1}^{K} w_r^{(2)}(f(\sum_{j=0}^{M^*} w_{rj}^{(1)} \cdot x_j^{is} \cdot \lambda_{sj}^{i}))) \right|, \qquad (4)$$

где $E_{\Delta is}$ - ошибка прогнозирования при удалении входа, соответствующего значению i-го фактора в момент времени t - Δ_s, $w_r^{(2)}, w_{rj}^{(1)}$ - веса нейронов обученной сети; x_j - значение j-го входа нейросети, y_p - наблюдаемое значение прогнозируемого параметра в момент времени t.

В формуле (4) λ^i_{sj} - элемент матрицы исключений, определяемый как

$$\lambda^{i}_{sj} = \begin{cases} 0,\ j = s + \sum\limits_{k=1}^{i}(1+h_i); \\ 1,\ j \neq s + \sum\limits_{k=1}^{i}(1+h_i). \end{cases} \tag{5}$$

Матрица исключений λ^i_{sj} введена для обеспечения возможности обнуления весов связей входа, соответствующего лагу Δ_s.

Для набора входов x_j^{is}, являющегося подмножеством входного вектора $(x_0,\ldots x_j)$ и описывающего историю изменения i-ого фактора, выполняется следующее соотношение:

$$x^{is} = x^{i}(t-\Delta_s)\,, \tag{6}$$

где Δ_s - временной лаг, s=0,…h_i.

Оптимальным будем считать такое значение временного лага Δ_s, при котором удаление соответствующего входа приведет к получению максимальной ошибки прогнозирования [3]:

$$\begin{cases} E_{\Delta is} \xrightarrow[x^{is}\in X_i]{} \max, \\ x^{is} = x^{i}(t-\Delta_s). \end{cases} \tag{7}$$

Максимальная глубина исторической выборки по каждому фактору не должна превышать половину длины исходного временного ряда. Временной лаг Δ_s должен быть неотрицательным, поэтому

$$0 \le \Delta_s \le N/2. \tag{8}$$

Окончательно задача оптимизации многомерного лага с учетом (2), (5), (6) и (8) выглядит так:

$$\begin{cases} E_i = \left| y_p - f\left(\sum\limits_{r=1}^{K} w_r^{(2)} \left(f\left(\sum\limits_{j=0}^{M^*} w_{rj}^{(1)} \cdot x_j^{is} \cdot \lambda^i_{sj}\right)\right)\right)\right| \xrightarrow[x_j^{is}\in X_i]{} \max, \\ M^* = \sum\limits_{i=1}^{M}(1+h_i), h_i \le \dfrac{N}{2}, i = \{1,\ldots M\}, y_p \in Y. \end{cases} \tag{9}$$

Нелинейный характер функции (2), входящей в выражение (9), неопределенность выбора глубины выборки по каждому фактору, возрастающая при увеличении количества учитываемых факторов и глубин выборок размерность задачи, требуют разработки специального нейросетевого алгоритма, учитывающего описанные ограничения.

Анализ чувствительности

Принципиальная возможность применения анализа чувствительности для нахождения оптимального многомерного временного лага была подтверждена в ходе вычислительного эксперимента.

Анализ чувствительности проводился на тестовых данных, полученных расчетом значений наборов функций нескольких переменных с известным запаздыванием по некоторым переменным и значениям лагов 3, 5, 7. В рассматриваемом ниже примере функция имеет вид (10):

$$Y(t) = 10 + \frac{(X_1(t) + \dfrac{X_2(t)}{3} - \ln(X_4(t)) \cdot X_3(t_3))}{X_6(t) + 2{,}5X_5(t_5)} - 0{,}44X_7(t_7) + X_8, \quad (10)$$

где t_3=t-1, t_5=t-3, t_7=t-2.

Взаимное расположение рядов исходных данных $X_1,\ldots X_8$ и зависящей от них функции Y, использованных в одной из серий эксперимента, приведены на Рисунке 1. Вариации параметров уравнения (9) породили набор нейросетей, обобщенная структура которых представлена на Рисунке 2.

В табл. 1 представлены значения ошибок обучения, обобщения и прогнозирования, полученные в ходе вычислительного эксперимента.

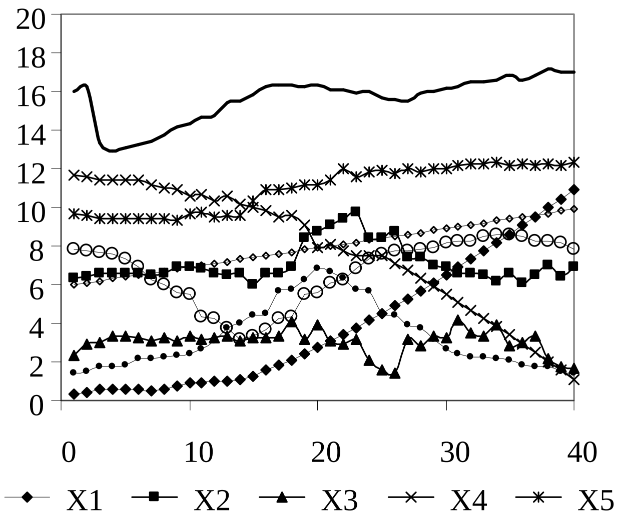

Рисунок 1 - Взаимное расположение восьми рядов значений независимых переменных $X_1(t),\ldots X_8(t)$ и зависящей от них переменной *Y*

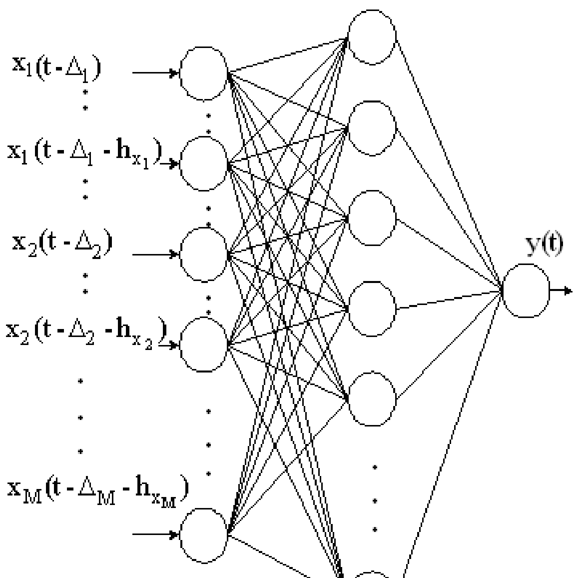

Рисунок 2 - Обобщенная структура нейросети после оптимизации вектора временных лагов

Выборочные результаты анализа чувствительности набора нейросетей приведены на Рисунке 3. В рядах, изначально не обладавших временным лагом, например, $X_4(t)$, наибольшая чувствительность отмечена для входов $X_i(t)$ (Рисунок 3, а). Для рядов $X_3(t)$, $X_5(t)$, $X_7(t)$ наиболее чувствительными оказались входы $X_i(t\text{-}\Delta)$ (Δ_3=1, Δ_5=3, Δ_7=2) (Рисунок 3, б).

Таблица 1 - Качество обучения нейросети и прогноза Y

Набор входных переменных	Средне квадратическое отклонение на обучающем множестве	Средне квадратическое отклонение на контрольном множестве	Средняя абсолютная ошибка, %
$X_1(t)$,… $X_8(t)$	0.240844	0.184005	2.78
$X_1(t)$, $X_2(t)$, $X_3(t-1)$, $X_4(t)$, $X_5(t-3)$, $X_6(t)$, $X_7(t-2)$, $X_8(t)$	0.09965	0.199094	1.37

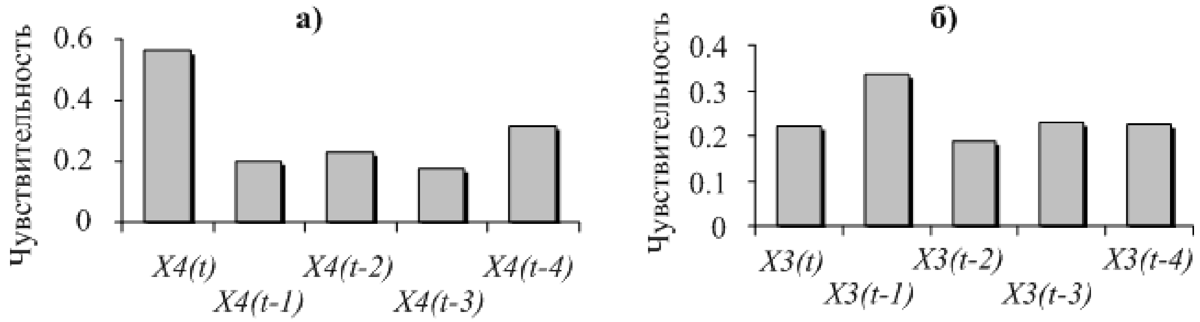

Рисунок 3 - Выборочные результаты анализа чувствительности входов, отражающих историю изменения независимых переменных

Алгоритмизация процедур неоднородного нейросетевого анализа

Далее осуществлена алгоритмизация процедур неоднородного нейросетевого анализа, позволяющих построить вектор оптимальных временных лагов значимых независимых переменных. В ходе решения основной задачи прогнозирования решены следующие подзадачи:

- определение вектора оптимальных лагов;
- сдвиг исходных временных рядов на величину найденных лагов.

Для определения оптимального временного лага по каждому временному фактору предложена процедура оценки чувствительности нейросети к значениям независимой переменной в различных временных точках с увеличением глубины погружения по анализируемому фактору. Максимальная глубина погружения определяется как N_i/2, где N_i - длина временного ряда i-й независимой переменной. Алгоритм представлен на Рисунке 4. Анализ чувствительности проводился по следующей схеме.

1 этап. Определение максимальной глубины погружения.

2 этап. Построение и обучение набора нейросетей с различным составом входов: глубина погружения исследуемого ряда принимается максимальной, остальные ряды представлены одним входом на ряд.

3 этап. Определение временного лага для каждого ряда с помощью анализа чувствительности нейросети к удалению входа.

После определения оптимальных временных лагов была построена новая совокупность данных, использованная для обучения сети методом обратного распространения ошибки и получения прогнозного значения зависимой переменной.

Найденные значения оптимальных временных лагов использовались для сдвига временных рядов независимых переменных и построения скорректированного множества исходных данных по следующей схеме:

1) определение максимального из найденных оптимальных лагов;

2) сдвиг каждого ряда на величину оптимального лага данной независимой переменной;

3) выбор тактики продолжения процесса: редукция рядов исходных данных на величину максимального лага или дополнение недостающих значений.

Сдвиг рядов исходных данных осуществляется с помощью линии задержки, принцип действия которой представлен на Рисунке 5. Здесь $x_i(t_k)$ - значение *i*-й независимой переменной в *k*-й временной точке, $\Delta(\Delta_1 \ldots \Delta_M)$ - вектор оптимальных временных лагов.

После определения оптимальных временных лагов была построена новая совокупность данных, использованная для обучения сети с целью получения уточненного прогноза значения зависимой переменной.

Для получения уточненной выборки для обучения и прогноза предлагается следующий алгоритм.

Шаг 1. Выбирается первый параметр X_1.

Шаг 2. Выбирается соответствующий временной лаг Δ_1.

Шаг 3. Если значение временного лага Δ_i равно нулю, то значения *i*-ой независимой переменной в выбранные моменты времени переносятся в результирующую выборку без изменений. Иначе ряд сдвигается. Если значение X_i в момент t-Δ отсутствует, то выбирается тактика редукции или дополнения временного ряда, и в соответствии с выбранным вариантом получают укороченные (Рисунок 6) или дополненные ряды наблюдений.

Шаг 4. Присвоение нового значения X_i в каждой временной точке.

Дополнение недостающих данных велось с использованием одной из трех тактик получения «виртуальных» значений.

Тактика I - производится дополнение временного ряда данными непосредственных наблюдений в данной точке.

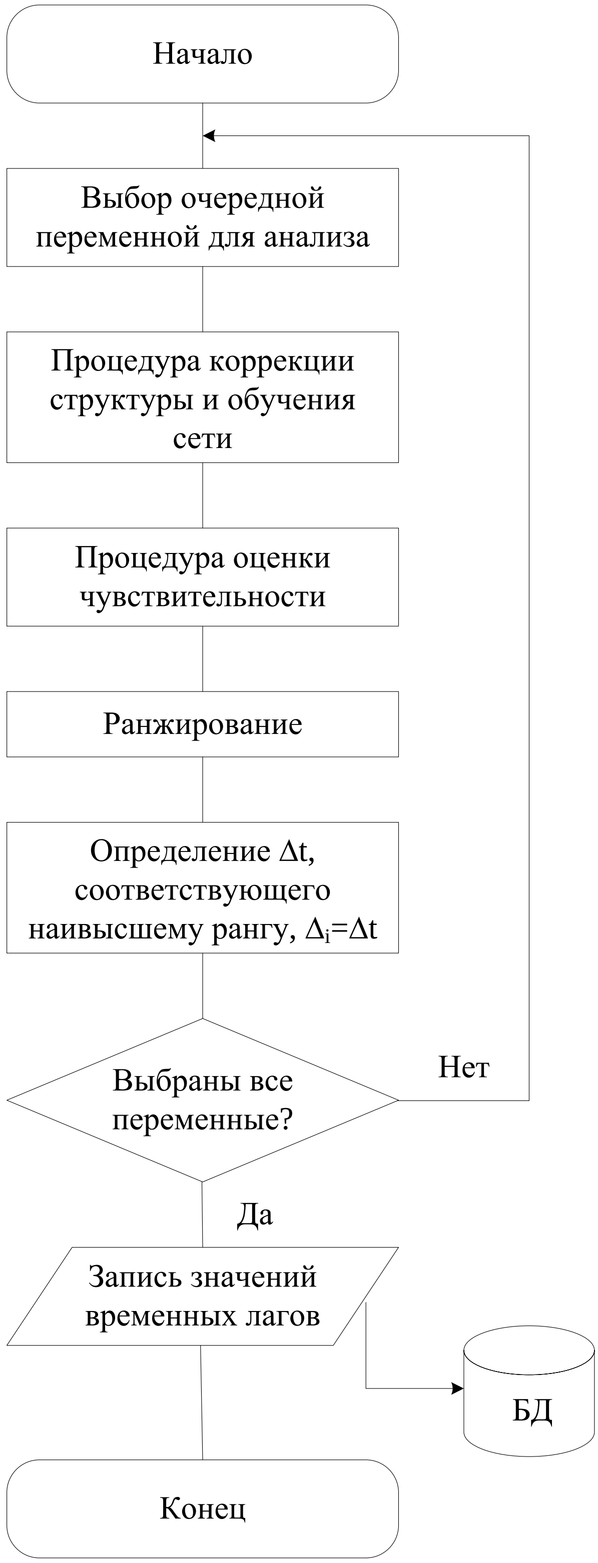

Рисунок 4 - Алгоритм определения вектора оптимальных лагов

Тактика II - дополнение временного ряда расчетными данными по выбранному алгоритму (средним по ряду, средним по соседним значениям, заполнение нулями).

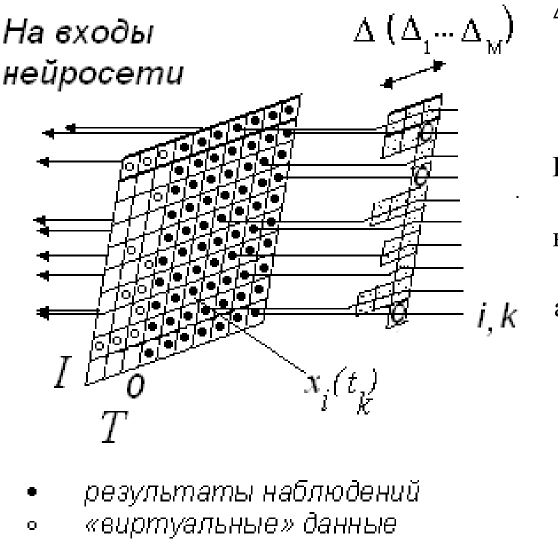

Рисунок 5 - Принцип действия линии задержки

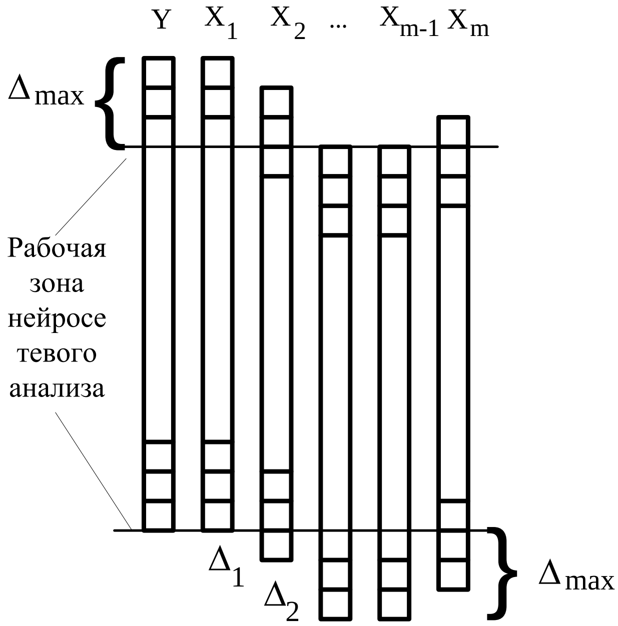

Рисунок 6 - Формирование редуцированного множества исходных данных

Рисунок 6 иллюстрирует построение редуцированной матрицы исходных данных. Ограничением этого способа продолжения процесса служит необходимость поддержания количества обучающих примеров на уровне, достаточном для обучения нейронной сети.

Укрупненный алгоритм построения скорректированной выборки с дополнением пропущенных значений представлен на Рисунке 7.

Тактика сохранения подразумевает наличие в рядах исходных данных пропущенных значений, которые должны быть заполнены до начала процедур обучения и прогнозирования. При выборе тактики дополнения временного ряда необходимо учитывать особенности предметной области рассматриваемой задачи.

Обсуждение

Анализ ошибок обучения, обобщения и прогнозирования на исходных и смещенных рядах, приведенных в Таблице 1, позволил сделать вывод о снижении ошибки обучения и ошибки прогнозирования на смещенных рядах при практически неизменной ошибке обобщения, что свидетельствует об эффективности предложенного алгоритма и отсутствии структурных эффектов в изменении качества прогноза.

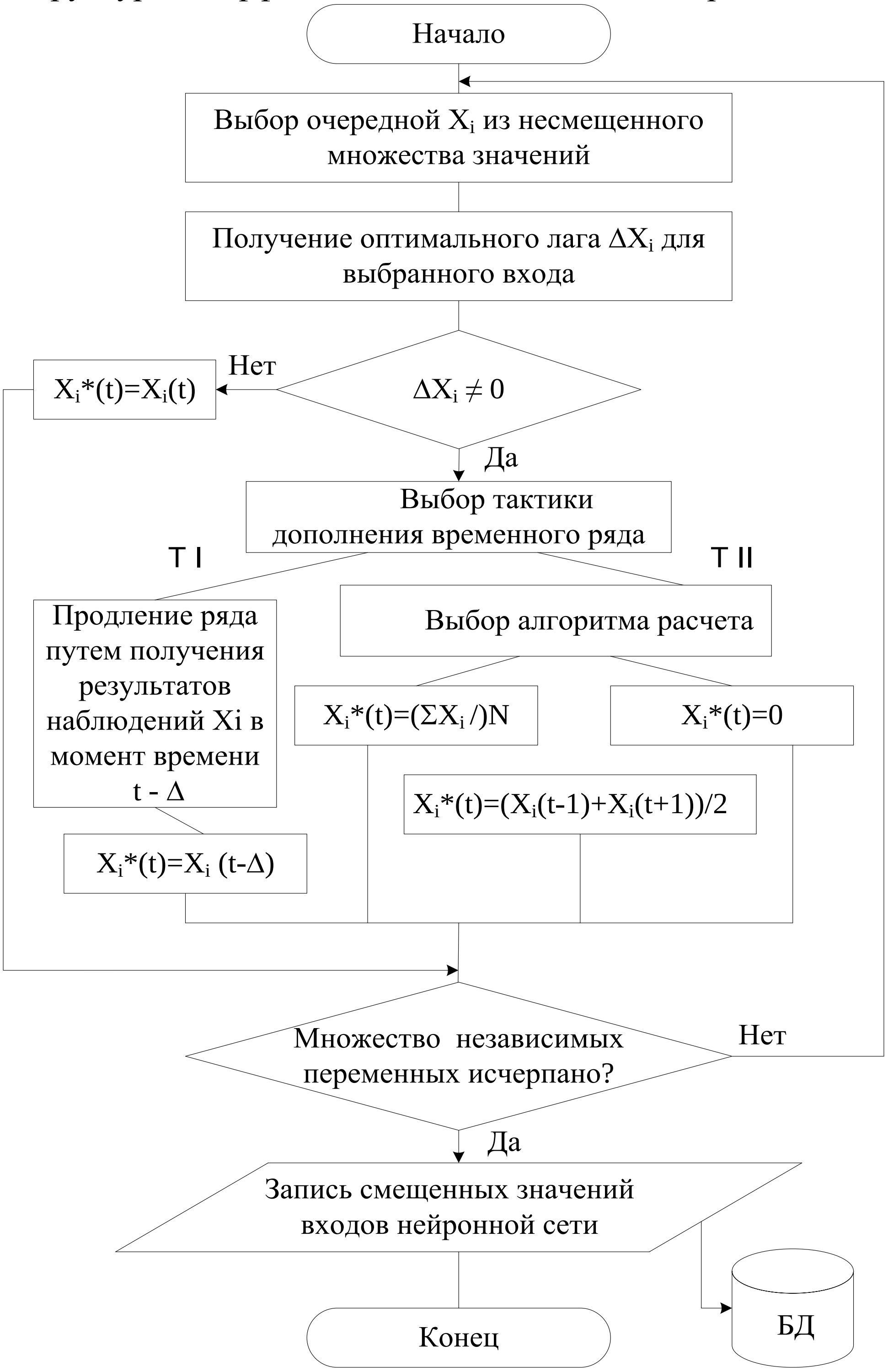

Рисунок 7 - Укрупненный алгоритм расчета смещенных значений входов нейронной сети

Разработанный алгоритм оптимизации многомерного лага на основе анализа чувствительности с применением линии задержки на входе нейросети позволяет учесть отсроченное влияние значимых факторов на прогнозируемую величину и устранить неопределенность в выборе вектора входных данных при обучении, тестировании нейросети и получении прогноза.

Разработанный алгоритм оптимизации многомерного лага на основе анализа чувствительности с применением линии задержки на входе нейросети позволяет учесть отсроченное влияние значимых факторов на прогнозируемую величину и устранить неопределенность в выборе вектора входных данных при обучении, тестировании нейросети и получении прогноза.

Заключение

Исследованы модели и методы нейросетевого моделирования динамики на основе анализа многомерных временных рядов с учетом отсроченного влияния значимых факторов. В связи с невозможностью одновременного определения оптимального временного лага и обучения сети нахождение многомерного лага рассмотрено как отдельная оптимизационная задача. Изложена математическая постановка задачи построения нейросети для ненулевого запаздывания, приведено описание особенностей оптимизации величины запаздывания для одной независимой переменной (входа), конкретизирована информационная база моделирования и прогнозирования и нейросетевые алгоритмы обработки данных, проведена алгоритмизация множественного регрессионного анализа с оптимизацией вектора запаздываний для значимых факторов. Принципиальная возможность применения анализа чувствительности для нахождения оптимального многомерного временного лага подтверждена в ходе вычислительного эксперимента.

Результаты

1. Обоснована необходимость учета запаздывания влияния независимых переменных на прогнозируемую величину. Показано, что наличие неучтенного временного лага не позволяет получить корректные обучающие примеры и ухудшает качество прогнозирования.
2. Получена математическая формулировка постановки задачи построения прогнозирующей нейронной сети в случае ненулевого запаздывания для одной и нескольких независимых переменных.

3. Предложен способ определения оптимального временного лага, основанный на поиске максимума абсолютной ошибке предсказания при исключении входа, соответствующего определенному значению лага.
4. Проведена верификация предложенного алгоритма путем исследования чувствительности нейросети, построенной для аппроксимации функции с заранее определенными временными лагами. Результаты свидетельствуют, что ошибка обучения на смещенных временных рядах значительно уменьшается вследствие установления однозначного соответствия примеров ($x_i(t)$, $y(t)$) и построения корректного обучающего множества, наиболее полно характеризующего реальные зависимости между переменными, а ошибка прогнозирования уменьшается практически вдвое.
5. Предложен дополнительный этап предварительной обработки исходных данных, заключающийся в сдвиге рядов исходных данных на величины оптимальных временных лагов.
6. Разработан алгоритм оптимизации вектора временных лагов с применением метода анализа чувствительности нейросети к удалению входа, соответствующего оптимальному значению временного лага для каждой независимой переменной.
7. Разработана процедура коррекции структуры нейросети с уменьшением глубины исторической выборки, определяемой графиками автокорреляционных функций, что особенно актуально при нехватке обучающих примеров при исследовании систем, имеющих короткую историю наблюдений. Предложена структура линии задержки, позволяющей учесть найденные значения временных лагов для каждой независимой переменной и сформировать корректные множества для обучения, тестирования нейросети и получения уточненного прогноза.

ЛИТЕРАТУРА

1. Крючкова И.Н. Специальное математическое обеспечение краткосрочного прогнозирования на основе нейросетевого моделирования и анализа многомерного лага: Автореф. дис. … канд. техн. наук. Воронеж, 2007.
2. Авдеева В.М., Кравец О.Я. Теоретические основы прогнозирования налоговых поступлений на основе кросскорреляционного анализа многомерных временных рядов// Системы управления и информационные технологии, 2006, №1.2(23). - С. 212-216.

3. Авдеева В.М., Кравец О.Я., Крючкова И.Н. Территориальное прогнозирование налоговых поступлений с применением многомерных кросскорреляционных технологий// Инновационный Вестник Регион, 2007, №3(9). - С. 31-36.
4. Авдеева В.М., Крючкова И.Н. Исследование технологии нейросетевого прогнозирования налоговых поступлений территории с применением техники многомерного кросскорреляционного анализа// Территория науки, 2007, №4(5). - С. 428-436.
5. Hecht-Nielsen R. Theory of the Backpropagation Neural Network// Neural Networks, 1989, 1(1):593 - 605.
6. Gibbs M.N. Variational Gaussian process classifiers// IEEE Transactions on Neural Networks, 2000, 11 (6): 1458-1464.
7. Deep Learning/ I. Goodfellow, Y. Bengio, A. Courville. MIT Press, 2016, 196 p.
8. Галушкин А.И. О методике решения задач в нейросетевом логическом базисе// Нейроинформатика – 2006. Часть 1. https://refdb.ru/download/1480079.html.

I.N. Kryuchkova, E.E. Krasnovskiy, E.V. Bolnokina, O.Ja. Kravets

ALGORITHMS FOR RESEARCH OF MULTIDIMENSIONAL TIME SERIES TAKING INTO ACCOUNT THE EXTENDED INFLUENCE OF FACTORS ON THE BASIS OF MATHEMATICAL MODELING

Voronezh State Technical University
Moscow State Technical University named after NE Bauman

The models and methods of neural network modeling of dynamics based on the analysis of multidimensional time series, taking into account the delayed influence of significant factors, are investigated. In connection with the impossibility of simultaneous determination of the optimal time lag and network training, it is necessary to consider finding a multidimensional lag as a separate optimization problem. The mathematical formulation of the problem of building a neural network for a non-zero delay is described, a description of the optimization characteristics of the latency for one independent variable (input) is given, the information base for modeling and forecasting and neural network data processing algorithms are specified, and the latency vector for significant factors is optimized. The fundamental possibility of using sensitivity analysis to find the optimal multidimensional time lag was confirmed during the computational experiment. The sensitivity analysis was carried out on test data obtained by calculating the values of the sets of functions of several variables with a known delay for some variables. Analysis of learning errors, generalization and forecasting on the original and offset series allowed to conclude that there was a significant decrease in the training error and prediction error on the shifted series with a practically *unchanged generalization error, which indicates the effectiveness of the proposed algorithm and the absence of structural effects in changing the quality of the forecast.*

Keywords: mathematical modeling, neural networks, lag, forecast.

REFERENCES

1. Kryuchkova I.N. Spetsial'noe matematicheskoe obespechenie kratkosrochnogo prognozirovaniya na osnove neyrosetevogo modelirovaniya i analiza mnogomernogo laga: Avtoref. dis. … kand. tekhn. nauk. Voronezh, 2007.
2. Avdeeva V.M., Kravets O.Ya. Teoreticheskie osnovy prognozirovaniya nalogovykh postupleniy na osnove krosskorrelyatsionnogo analiza mnogomernykh vremennykh ryadov// Sistemy upravleniya i informatsionnye tekhnologii, 2006, No.1.2(23). - pp. 212-216.
3. Avdeeva V.M., Kravets O.Ya., Kryuchkova I.N. Territorial'noe prognozirovanie nalogovykh postupleniy s primeneniem mnogomernykh krosskorrelyatsionnykh tekhnologiy// Innovatsionnyy Vestnik Region, 2007, No.3(9). - pp. 31-36.
4. Avdeeva V.M., Kryuchkova I.N. Issledovanie tekhnologii neyrosetevogo prognozirovaniya nalogovykh postupleniy territorii s primeneniem tekhniki mnogomernogo krosskorrelyatsionnogo analiza// Territoriya nauki, 2007, No.4(5). - pp. 428-436.
5. Hecht-Nielsen R. Theory of the Backpropagation Neural Network// Neural Networks, 1989, 1(1):593 - 605.
6. Gibbs M.N. Variational Gaussian process classifiers// IEEE Transactions on Neural Networks, 2000, 11 (6): 1458-1464.
7. Deep Learning/ I. Goodfellow, Y. Bengio, A. Courville. MIT Press, 2016, 196 p.
8. Galushkin A.I. O metodike resheniya zadach v neyrosetevom logicheskom bazise// Neyroinformatika – 2006. Chast' 1. https://refdb.ru/download/1480079.html.