УДК 004.032.26

Е.В. Бодянский, О.А. Романюк, О.С. Удовенко

Харьковский национальный университет радиоэлектроники, Харьков

АДАПТИВНЫЙ ФИЛЬТР-ПРЕДИКТОР МНОГОМЕРНЫХ СУЩЕСТВЕННО НЕСТАЦИОНАРНЫХ ВРЕМЕННЫХ РЯДОВ

Решена задача адаптивного прогнозирования многомерных временных рядов, характеристики которых могут резко изменяться непредсказуемым образом с течением времени. Введена нейросетевая архитектура с синапсами – динамическими фильтрами и предложен алгоритм обучения, сочетающий в себе как высокое быстродействие, так и фильтрующие свойства. Предлагаемый фильтр-предиктор предназначен для использования в задачах технического анализа, связанных с оперативным предсказанием биржевых курсов.

Ключевые слова: *адаптивное прогнозирование и фильтрация, многомерные нестационарные временные ряды, обучение.*

Введение

Задача прогнозирования многомерных временных рядов достаточно часто возникает во многих технических, медико-биологических, финансово-экономических приложениях, где качество принимаемых решений существенным образом зависит от точности формируемых прогнозов. Особое место этой задаче отводится в техническом анализе, где неверный прогноз ведет к неоправданным финансовым потерям [1]. Необходимо отметить, что временные ряды, являющиеся предметом рассмотрения технического анализа параметров компьютерных систем биржевой торговли, как правило, характеризуются высоким уровнем нестационарности прогнозируемых параметров, наличием существенных нерегулярных нелинейных трендов, (а также аномальных выбросов), стохастической и хаотической неопределенностью. Пример такого типичного ряда, описывающего курс акций компании Apple с 1998 г. по 2008 г., приведен на рис. 1.

Данные этого ряда свидетельствуют о наличии некоторой предсказуемой составляющей, хотя здесь нельзя также и говорить о случае полностью детерминированной хаотической динамики. Следовательно, попытка применения технического анализа для прогнозирования имеет здесь веские основания. Особое внимание следует обратить на преимущества адаптивных методов прогнозирования, процесс реализации которых состоит в вычислении последовательных значений прогнозируемых показателей с учетом степени влияния предшествующих уровней. При краткосрочном прогнозе наиболее эффективными являются адаптивные предикторы, учитывающие неравноценность уровней временного ряда и быстро приспосабливающие свою структуру и параметры к изменяющимся условиям. Здесь важно правильно выбрать глубину учитываемой предыстории.

Понятно, что традиционные методы анализа временных рядов, основанные на регрессионном, корелляционном, спектральном и других подобных подходах в данном случае неэффективны.

Альтернативой традиционным методам может служить аппарат вычислительного интеллекта и, прежде всего, искусственные нейронные сети [2], благодаря своим универсальным аппроксимирующим возможностям.

Вместе с тем, элементарный визуальный анализ последовательности, приведенной на рис. 1, показывает, что из высокого качества аппроксимации вовсе не следуют экстраполирующие свойства нейронной сети. Для обеспечения прогнозирующих свойств в данном случае следует отказаться от процедур обучения, основанных на обратном распространении ошибок (многослойные персептроны, рекуррентные сети Элмана, Джордана, Вильямса-Зипсера и т.п.) или стандартном методе наименьших квадратов (радиально-базисные нейронные сети, функционально связанные нейронные сети и т.п.) и воспользоваться алгоритмами, основанными на локальных критериях обучения и «короткой памяти» (в идеале - одношаговыми). При этом, конечно, алгоритмы обучения должны обладать не только высоким быстродействием, но и фильтрующими свойствами для подавления шумовой компоненты в анализируемой последовательности. Короткая память алгоритма обучения позволяет воспользоваться предположением о локальной линейности временного ряда на коротких временных интервалах [3], что, в свою очередь, дает возможность применения линейных алгоритмов адаптивной идентификации [4], соответствующим образом модифицированных с целью придания им фильтрующих свойств [5]. И, наконец, заметим, что прогнозирование одномерного изолированного временного ряда не имеет особого смысла, поскольку любое реальное явление или объект, как правило, характеризуется множеством взаимозависимых показателей, совместная обработка которых позволяет в итоге улучшить качество, как самих прогнозов, так и принимаемых на их основе решений.

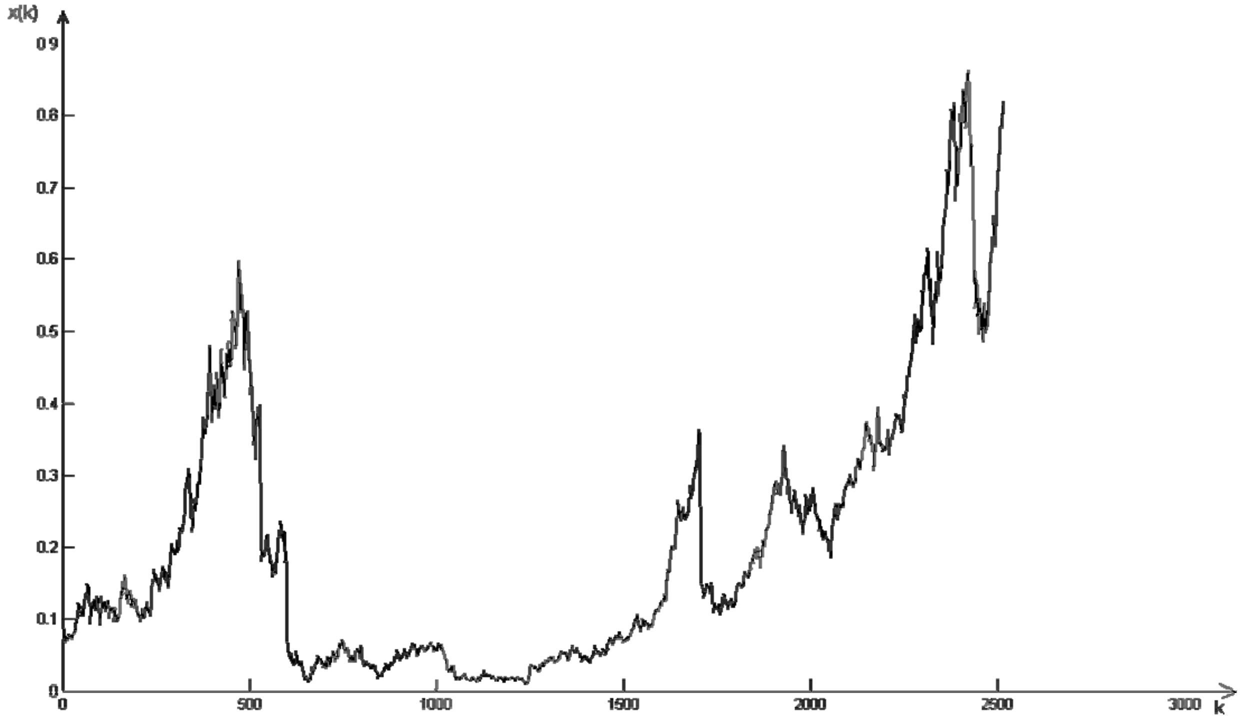

Рис. 1. Временной ряд, описывающий цены акций компании Apple на время открытия торговых периодов с 1998 г. по 2008 г.

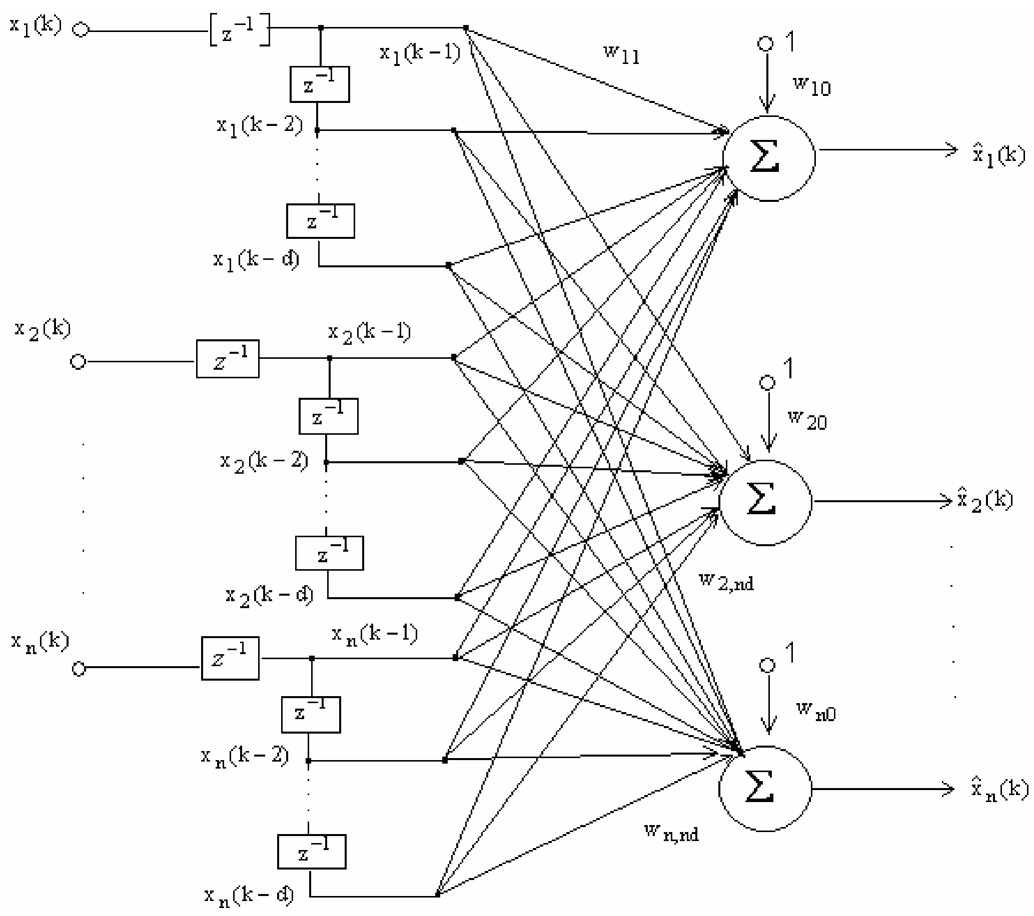

Рис. 2. Адаптивный фильтр-предиктор

Архитектура адаптивного фильтра-предиктора

Архитектура предлагаемого ниже адаптивного нейросетевого фильтра-предиктора многомерного существенно нестационарного временного ряда $x(k)=(x_1(k),x_2(k),...,x_n(k))^T$, где k = 1, 2.. – текущее дискретное время, приведена на рис. 2. Архитектура содержит n параллельно соединенных адаптивных линейных ассоциаторов с $nd+1$ входами каждый, где $d\geq 1$ – глубина используемой для прогнозирования временного ряда его же предыстории. Каждый из ассоциаторов содержит $nd+1$ настраиваемых синаптических весов w_{ij}, $j = 1, 2,\ldots, n$; $i = 0, 1,\ldots, nd$, а архитектура в целом – $n(nd + 1)$ весов. С каждым из ассоциаторов связана цепочка из d элементов чистого запаздывания z^{-1}, формирующая на его входах предысторию $x_i(k-1),x_i(k-1),...,x_i(k-d)$ так, что фактически можно говорить о синапских фильтрах [6 – 8] (фильтрах с конечной импульсной характеристикой [5]) на входах каждого нейрона. Синаптические веса w_{jo} определяют уровень смещения на выходах каждого нейрона - ассоциатора.

Таким образом, преобразование, реализуемое данной архитектурой, может быть записано в виде

$$\hat{x}(k)=WX(k)\,, \tag{1}$$

где $\hat{x}(k)$ – (nx1)-вектор прогнозов, полученных по предыстории доступной к моменту времени k; $X(k)=(1,x_1(k-1),...,x_1(k-d),x_2(k-1),...,x_n(k-d))^T$ – ((nd + 1)x1)-вектор этой предыстории; W – (nx(nd + 1))-матрица синаптических весов, подлежащих определению в процессе обучения фильтра-предиктора и обеспечивающая оптимальность в некотором смысле получаемых прогнозов.

Обучение фильтра-предиктора

Обучение параметров архитектуры, приведенной на рис. 2, будем проводить путем пошаговой минимизации локального критерия

$$\begin{aligned}E(k)&=\|x(k)-\hat{x}(k)\|^2=\|e(k)\|^2=\\&=\|x(k)-WX(k)\|^2\end{aligned} \tag{2}$$

с помощью градиентной процедуры

$$\begin{aligned}W(k)&=W(k-1)-\eta(k)\nabla_W E(k)=\\&=W(k-1)+2\eta(k)e(k)X^T(k),\end{aligned} \tag{3}$$

где $\eta(k)$ – параметр шага обучения, определяющий как «скоростные», так и фильтрующие свойства алгоритма (3) и выбираемый из тех или иных (часто эмпирических) соображений. Для оптимизации процесса обучения по быстродействию, используя технику, рассмотренную в [9], введем функцию Ляпунова

$$V(k)=Tr\tilde{W}^T(k-1)\tilde{W}(k-1)-Tr\tilde{W}^T(k)\tilde{W}(k)\,, \tag{4}$$

где Tr – символ следа матрицы; $Tr\tilde{W}^T(k)\tilde{W}(k)$ – квадрат сферической нормы матрицы $\tilde{W}$; $\tilde{W}(k)=W^*-W(k)$ – ненаблюдаемая матрица теоретических ошибок параметров; W^* – некая идеальная матрица весов, обеспечивающих оптимальное предсказание. Преобразуя (3) к форме

$$\tilde{W}(k)=\tilde{W}(k-1)-2\eta(k)e(k)X^T(k) \tag{5}$$

и подставляя ее в (4), приходим к зависимости

$$\begin{aligned}V(k)=&-Tr\eta(k)\tilde{W}^T(k-1)e(k)X^T(k)-\\&-Tr\eta(k)X(k)e^T(k)\tilde{W}(k-1)-\\&-2Tr\eta^2(k)X(k)e^T(k)e(k)X^T(k),\end{aligned} \tag{6}$$

оптимизация которой по η путем решения дифференциального уравнения

$$\frac{\partial V(k)}{\partial \eta}=0\,, \tag{7}$$

позволяет получить оптимальное значение

$$\eta(k)=\frac{Tr\tilde{W}^T(k-1)e(k)X^T(k)+TrX(k)e^T(k)\tilde{W}(k-1)}{4TrX(k)e^T(k)e(k)X^T(k)}, \tag{8}$$

которое, однако, не может быть использовано в расчетах из-за ненаблюдаемости матрицы ошибок $\tilde{W}(k-1)$.

Принимая во внимание очевидное соотношение

$$\tilde{W}(k-1)X(k)=e(k)\,, \tag{9}$$

можно переписать (8) в виде

$$\eta(k)=\frac{Tre(k)e^T(k)}{2TrX(k)e^T(k)e(k)X^T(k)}=\frac{1}{2\|X(k)\|^2}, \tag{10}$$

после чего процедура обучения может быть записана в виде оптимальной по быстродействию форме:

$$\begin{aligned}W(k)&=W(k-1)+\frac{e(k)X^T(k)}{\|X(k)\|^2}=\\&=W(k-1)+e(k)X^+(k),\end{aligned} \tag{11}$$

где $X^+(k)$ – вектор, псевдообратный по Муру–Пенроузу к X(k). Несложно видеть, что процедура (11) является обобщением на многомерный случай одношагового алгоритма Качмажа-Уидроу-Хоффа, достаточно широко распространенного в теории и практике обучения искусственных нейронных сетей.

Известно, что одношаговый алгоритм Качмажа, гарантируя максимальное быстродействие, совершенно не обладает фильтрующими свойствами, для обеспечения которых введем модификацию процедуры обучения с экспоненциальным взвешиванием информации вида

$$\begin{cases}W(k)=W(k-1)+r^{-1}(k)e(k)X^T(k);\\ r(k)=\alpha r(k-1)+\|x(k)\|^2,\end{cases} \tag{12}$$

где $0\leq\alpha\leq 1$ – параметр сглаживания, обеспечивающий компромисс между следящими и фильтрующими свойствами алгоритма. При $\alpha = 0$ алгоритм (12) приобретает вид (11), обеспечивая максимально возможное быстродействие при слежении за изменяющимися свойствами временного ряда x(k). При $\alpha = 1$ (12) превращается в процедуру стохастической аппроксимации [11], обладающую фильтрующими свойствами.

Заметим также, что в каждый текущий момент времени k с приходом нового вектора наблюдений x(k), сначала рассчитывается уточненная матрица синаптических весов W(k), а на ее основе – одношаговый прогноз

$$\hat{x}(k+1) = W(k)X(k+1), \qquad (13)$$

где вектор предыстории X(k + 1) содержит наблюдения, доступные к моменту k.

Результаты численного эксперимента

Рассматривалась тестовая задача текущего одношагового прогнозирования значений четырехмерного вектора $x(k) = (x_1(k), x_2(k), x_3(k), x_4(k))^T$, где $x_1(k)$ – цена акций компании Apple на время открытия торгового дня, $x_2(k)$ – минимальная цена акций компании Apple за весь торговый день, $x_3(k)$ – максимальная цена компании Apple за весь торговый день, $x_4(k)$ – цена акций компании Apple на время закрытия торгового дня.

Заметим, что на рис. 1 приведена компонента $x_1(k)$. Глубина предыстории была принята d = 5, а значение параметра сглаживания $\alpha = 0{,}8$. На рис. 3 приведен график прогнозов компоненты $\hat{x}_1(k)$.

Как видно, временная последовательность, образованная прогнозными значениями, очень близка к исходному прогнозируемому ряду $x_1(k)$.

Качество прогнозирования оценивалось с помощью оценок среднеквадратичной ошибки (RMSE)

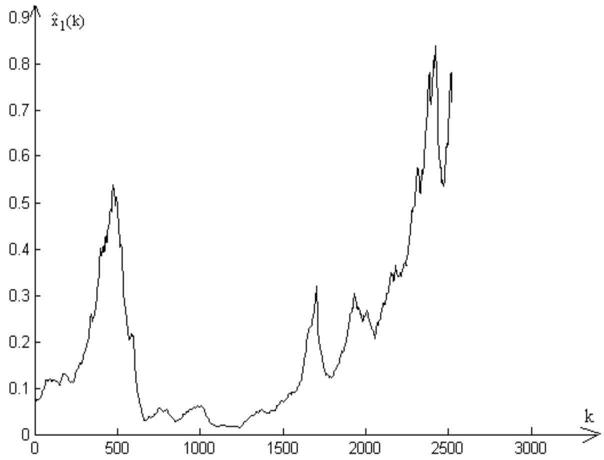

Рис. 3. Прогноз временного ряда

$$\bar{e}_j^2(k) = \frac{k-1}{k}\bar{e}_j^2(k-1) + \frac{1}{k}e_j^2(k) \qquad (14)$$

и среднего нормированного модуля ошибок (MAPE)

$$\left|\bar{e}_j(k)\right| = \frac{k-1}{k}\left|\bar{e}_j(k-1)\right| + \frac{1}{k}\frac{\left|e_j(k)\right|}{\left|x_j(k)\right|}. \qquad (15)$$

На рис. 4 приведен график изменения $\bar{e}_1^2(k)$, а на рис. 5 – $\left|\bar{e}_1(k)\right|$.

С помощью введенного адаптивного фильтра-предиктора удалось обеспечить достаточно высокое качество прогнозирования.

На рис. 6 приведена также автокорреляционная функция ошибок прогнозирования

$$R_{e_1e_1}(\tau) = \frac{1}{N-\tau}\sum_{k=\tau+1}^{N}(e_1(k) - \bar{e}_1(N))\times$$
$$\times(e_1(k+\tau) - \bar{e}_1(N)), \quad \tau = 0,1,2..., \qquad (16)$$

где N – объем обучающей выборки.

Как видно, временной ряд ошибок прогнозирования есть по сути «белый» шум, что свидетельствует о правильном выборе структуры прогноза, определяемой глубиной предыстории d.

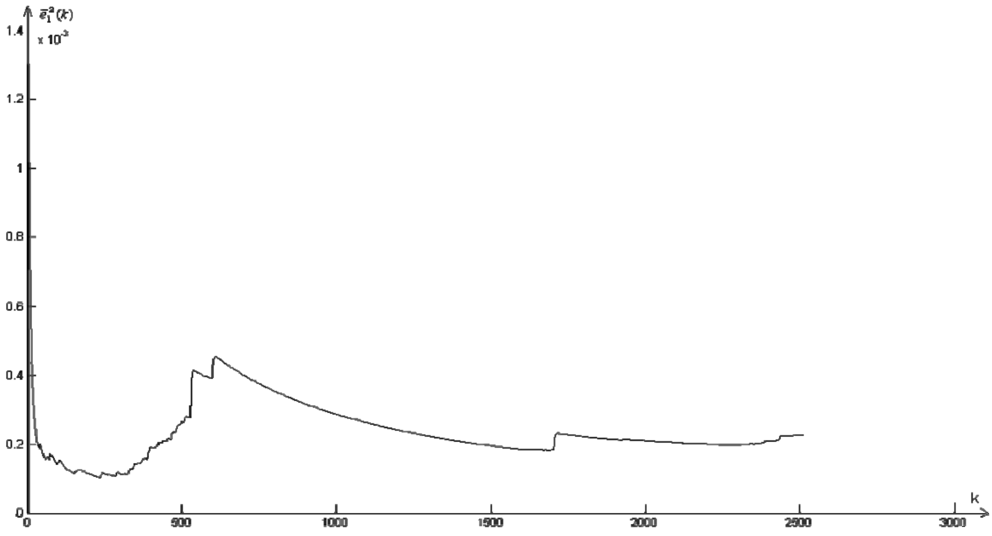

Рис. 4. Изменение среднеквадратичной ошибки $\bar{e}_1^2(k)$

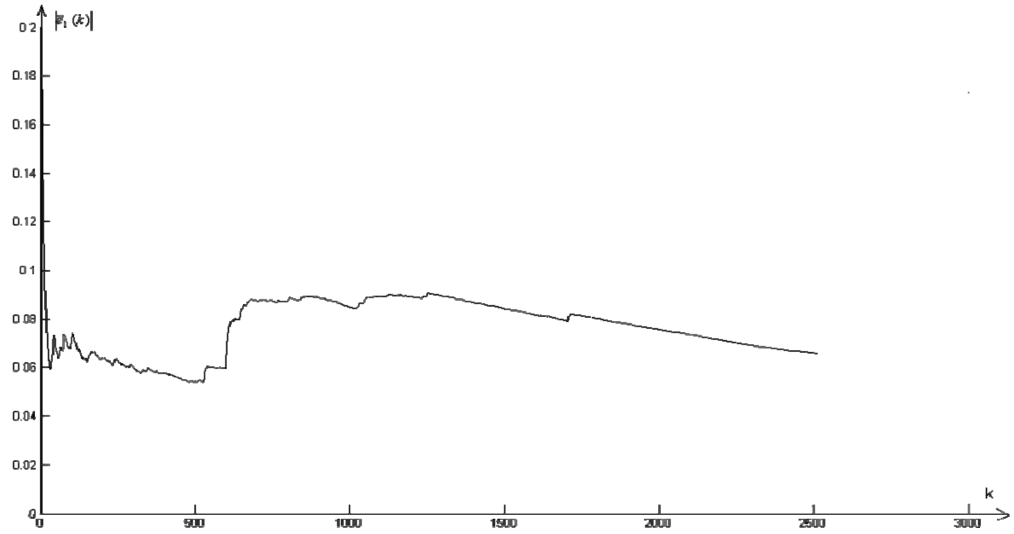

Рис. 5. Изменение среднего нормированного модуля ошибки $\left|\overline{e}_1\ (k)\right|$

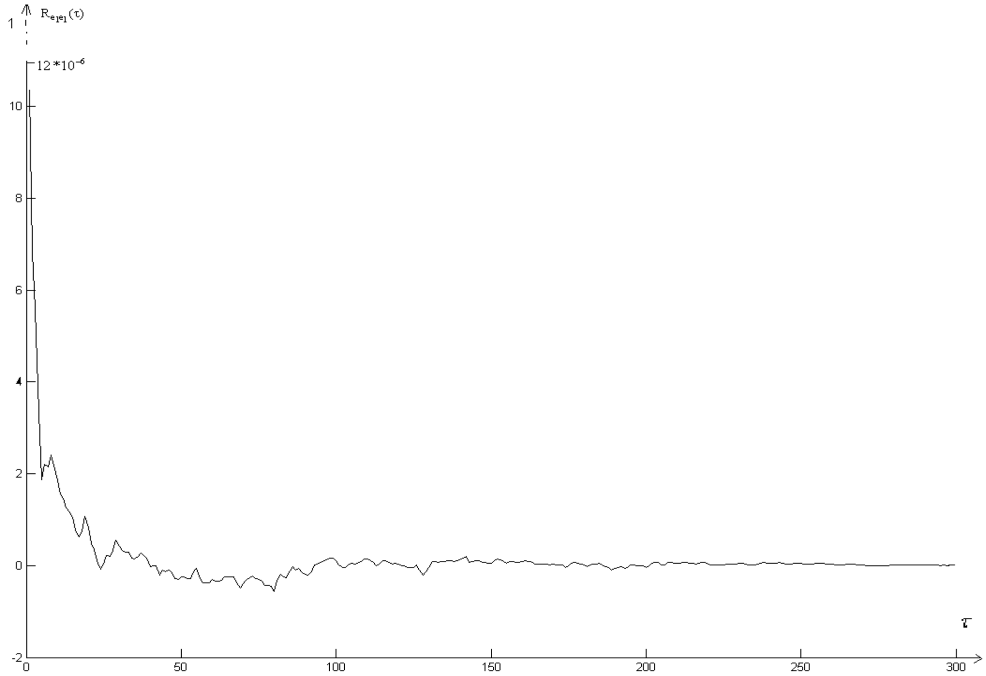

Рис. 6. Автокорреляционная функция ошибок прогнозирования

Выводы

Предложенные архитектура и алгоритм обучения адаптивного нейросетевого фильтра-предиктора многомерных временных рядов обеспечивают высокие быстродействие и качество прогнозирования в условиях существенной нестационарности и неопределенности и могут найти применение в задачах технического анализа и других приложениях, связанных с предсказанием многомерных стохастических и хаотических последовательностей.

Список литературы

1. Найман Э. Малая энциклопедия трейдера / Э. Найман. – М.: Альфа Капитал, 1999. – 236 с.

2. Zirilli J.S. Financial Prediction Using Neural Networks // J.S. Zirilli // Int. Thomson Computer Press. – London, 1997. – 135 p.

3. Райбман Н.С. Построение моделей процесс производства / Н.С. Райбман, В.М. Чадеев. – М.: Энергия, 1975. – 376 с.

4. Льюнг Л. Идентификация системы. Теория для пользователя / Л. Льюнг. – М.: Наука, 1991. – 432 с.

5. Уидроу Б. Адаптивная обработка сигналов / Б. Уидроу, С. Стирнс. – М.: Радио и связь, 1989 – 440 с.

6. Wan E. Temporal backpropagation for FIR neural networks // Int.Joint Conf. on Neural Networks.– San Diego, 1990. – V. 1. – P. 575-580.

7. Wan E.A. Time series prediction by using a connectionist network with integral delay lines / Eds. by A. Weigend, N. Gershenfeld "Times Series Prediction. Forecasting the Future and Understanding the Past". – SFI Studies in the Science of Complexity. – V. XVII. – Reading: Addison-Wesley, 1994. – P. 195-218.

8. Back A.D. A unifying view of some training algorithms for multilayer perceptrons with FIR filter synapses / A.D. Back, E.A. Wan, S. Lawrence, A.C. Tsoi; Eds. by J. Vlontzos, J. Hwang, E. Wilson // Neural Networks for Signal Processing 4. – N.Y.: IEEE Press, 1994. – P. 146-154.

9. Руденко О.Г. Адаптивный алгоритм прогнозирования случайных последовательностей / О.Г. Руденко, Е.В. Бодянский, И.П. Плисс // Автоматика. – 1979. – № 1. – С. 51-54.

10. Бодянский Е.В. Многошаговые оптимальные упредители многомерных нестационарных стохастических процессов / Е.В. Бодянский, И.П. Плисс, Т.В. Соловьева. // Докл. АН УССР. – 1986. – Сер. А. – № 12. – С. 47-49.

11. Caines P.E. Adaptive control with recursive identification for stochastic linear systems/ P.E. Caines, S. Lafortune // IEEE Trans. on Autom. Control. – 1984. – 29, #4. – P. 312-321.

Поступила в редколлегию 9.06.2009

Рецензент: д-р техн. наук, проф. В.А. Филатов, Харьковский национальный университет радиоэлектроники, Харьков.

АДАПТИВНИЙ ФІЛЬТР-ПРЕДИКТОР БАГАТОВИМІРНИХ СУТТЄВО НЕСТАЦІОНАРНИХ ЧАСОВИХ РЯДІВ

Є.В. Бодянський, О.А. Романюк, О.С. Удовенко

Вирішено задачу адаптивного прогнозування багатовимірних часових рядів, характеристики яких можуть різко змінюватися непередбаченим чином з плином часу. Введено нейромережеву архітектуру з синапсами – динамічними фільтрами та запропоновано алгоритм навчання, що поєднує в собі як високу швидкість, так і фільтруючі властивості. Фільтр-предиктор, що запропоновано, призначений для використання в задачах технічного аналізу, пов'язаних з оперативним передбаченням біржевих курсів.

***Ключові слова:** адаптивне прогнозування та фільтрація, багатовимірні нестаціонарні часові ряди, навчання.*

THE ADAPTIVE FILTER-PREDICTOR FOR MULTIVARIATE ESSENTIALLY NONSTATIONARY TIME SERIES

Ye.V. Bodyanskiy, O.A. Romaniuk, O.S. Udovenko

The adaptive forecasting task for multivariate time series whose characteristics can abruptly change in time by nonpredictable way is solved. The neural network architecture with synapses-dynamic filter is introduced and learning algorithm that combines both high speed and filtering properties is proposed. The proposed filter-predictor is intended for using in technical analysis tasks connected with on-line stock indexes prediction.

***Keywords:** adaptive prediction and filtering, adaptive multivariate nonstationary time series, learning.*