

ОБ ОДНОМ МЕТОДЕ ПРЕДВАРИТЕЛЬНОГО ПРОГНОЗА ВРЕМЕННЫХ РЯДОВ ФИНАНСОВЫХ ДАННЫХ

Аннотация. Рассмотрена проблема прогнозирования временных рядов цен акций ведущих мировых компаний, которым свойственна долговременная память. Делается предположение, что игнорирование наличия подобной корреляционной структуры временных рядов с применением традиционных методов анализа приводит к появлению значительно большей погрешности, чем учет долговременной памяти при фактическом ее отсутствии. Предполагается, что колебания цен на инструменты финансового рынка описываются процессом Херста, которым моделируют процессы с долговременной памятью. Такой временной ряд не может быть эффективно проанализирован с помощью традиционных стационарных моделей, которые полностью игнорируют этот факт. Ставится задача: с использованием рассматриваемого метода установить наличие долговременной памяти у исходного временного ряда и определить его тип.

Ключевые слова: временные ряды, фрактал, нейронные сети.

Разработанный в настоящей статье алгоритм позволяет идентифицировать и численно оценить такие фундаментальные характеристики временных рядов, как наличие и глубина долговременной памяти, устойчивость и т.д.

Используем фрактальный анализ — проверенный метод описания эволюционных процессов и прогнозирования экономических временных рядов. Основным инструментом фрактального анализа временных рядов является алгоритм R/S-анализа. Его методология была разработана в середине XX века гидрологом Херстом в процессе исследования временных рядов объемов речного стока. При проверке предположения, что ряды данных подчиняются нормальному закону, Херст определил новую статистику — индекс Херста (H). В ходе исследования Херст измерял флуктуации воды в пласте, сравнивая со средним по времени, и вводил безразмерное соотношение делением амплитуды R на стандартное отклонение S (R/S-анализ). Этот метод анализа был вызван изменением диапазона. Херст обнаружил, что большинство природных явлений, включая речные стоки, могут быть описаны регрессионной моделью с шумом (тенденция + шум). Силу тренда и уровень шума можно измерить по нормализованной амплитуде или, другими словами, до значения H не более 0.5 ($H < 0.5$).

Опишем алгоритм R/S-анализа в том виде, в котором он реализован в современных методах фрактального анализа [1, 2]. С учетом временных рядов

$$X = \{x_i\}, \quad i = 1, 2, \dots, n, \quad (1)$$

которые последовательно выделяют начальные сегменты

$$X_\tau = x_1, x_2, \dots, x_\tau, \quad \tau = 3, 4, \dots, n,$$

вычисляем текущее среднее значение

$$\bar{X}_\tau = \frac{1}{\tau} \sum_{i=1}^{\tau} x_i.$$

Тогда для каждого фиксированного X_τ , $\tau = 3, 4, \dots, n$, рассчитаем накопленное отклонение для его сегментов длины t :

$$Y_{\tau,t} = \sum_{i=1}^t (x_i - \bar{X}_\tau), \quad t = \bar{1}, \tau.$$

Далее вычисляем разность между максимальным и минимальным накопленными отклонениями: $R = R(\tau) = \max_{1 \leq t \leq \tau} (Y_{\tau,t}) - \min_{1 \leq t \leq \tau} (Y_{\tau,t})$, которая называется диапазоном R . Этот масштаб нормирован и представлен как доля от R/S , где

$$S = S(\tau) = \sqrt{\frac{1}{\tau} \sum_{j=1}^{\tau} (x_j - \bar{X}_{\tau})^2},$$

S — стандартное отклонение для интервальных временных рядов X_{τ} , $\tau = 3, 4, \dots, n$.

Показатель Херста $H = H(\tau)$, который характеризует фрактальную размерность рассматриваемых временных рядов и соответствующий цветовой шум, получаем в виде соотношения [1]

$$\frac{R}{S} = (\alpha\tau)^H.$$

Логарифмируя обе части этого уравнения и предполагая, что $\alpha = 0.5$ (см. [3]), получаем в декартовых координатах (x_{τ}, y_{τ}) точки H — ординату и абсциссу траектории соответственно:

$$y_{\tau} = H(\tau) = \frac{\log(R(\tau)/S(\tau))}{\log(\tau/2)}, \quad x_{\tau} = \tau.$$

Для фрактального анализа (1) R/S -траектории в декартовых логарифмических координатах последовательности точек, где абсцисса $x_{\tau} = \log(\tau/2)$ и ордината $y_{\tau} = \log(R(\tau)/S(\tau))$, сегмента, соединяющего соседние точки (x_{τ}, y_{τ}) и $(x_{\tau+1}, y_{\tau+1})$, $\tau = 3, 4, \dots, n-1$, требуется получить графическое представление R/S -пути (H -пути) в логарифмических координатах (в обычных декартовых координатах).

Одной из основных характеристик фрактальных временных рядов является цветовой шум, который соответствует этой серии в определенный момент времени. Значения $H \geq 0.6$ определяют черный шум. Чем больше значение, тем выше устойчивость тренда к соответствующему сегменту временных рядов. Значения в окрестности $\sim 0.5 \pm 0.1$ определяет область белого шума, что соответствует «хаотическому поведению временного ряда», и, следовательно, имеем менее надежный прогноз. Как показано ниже, присущий черным рядам серый шум соответствует области нечеткого различия между областями черного и белого шума.

Нельзя утверждать или отрицать наличие долговременной памяти во временном ряду (1), если его H -траектория недолгое время находится в области черного шума, а поведение R/S -траектории хаотично с начальной точки [4].

Основанием для утверждения того, что временной ряд (1) имеет долговременную память, является выполнение следующих условий.

1. H -путь, проходящий через некоторые его начальные точки, находится в черном шуме, а путь, определяющий точки, входящие в черный шум, является трендом. H -путь, при котором глубина состояния памяти при H -траектории уменьшается, а R/S -траектория в этой точке показывает резкое изменение тренда.

2. Если в этом временном ряду случайно перетасовать его элементы, то максимальное значение алгоритма Херста и R/S -траектории будет намного ниже, чем значения H для исходного временного ряда; в этом случае временной ряд имеет долговременную память.

В настоящей статье R/S -анализ проводится на основании следующих временных рядов. Анализируемый временной ряд stock quotation — MSFT (X_1) имеет размер образца 1050, а временной ряд stock quotation — IBM (X_2) имеет размер образца 1100.

На рис. 1 показаны данные, полученные на выходе R/S-анализа, H- и R/S-пути для соответствующих временных рядов: для графиков H-траектории на отрезках длины абсцисс τ , для графиков R/S-траектории по значениям абсцисс $\ln(\tau/2)$.

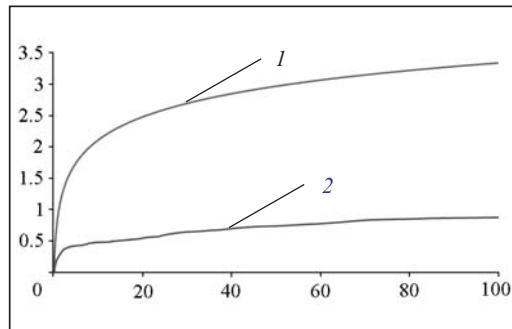


Рис. 1. Графики R/S-пути (1) и H-траектории (2) для временных рядов (временные ряды X_1)

На основании R/S-анализа были получены результаты, общие для всех рассмотренных наблюдений.

- H-траектория каждого временного ряда после получения первого значения переходит в области черного шума, что позволяет свидетельствовать о сохраняемости трендового времени и временных характеристиках долговременной памяти.

- Для каждого временного ряда можно оценить количество первых элементов (m), после чего показатель H переходит в зону черного шума. Эта величина характеризует минимально допустимый образец временного ряда, который позволяет прогнозировать поведение временных рядов. Число первых элементов временного ряда составляет от пяти до девяти.

- Результаты R/S-анализа позволяют также утверждать, что некоторые рассматриваемые временные ряды по своей природе являются циклическими (скорее, квазициклическими). Более того, из анализа R/S-пути вытекает, что точки изменения тренда часто соответствуют окончанию квазицикла. Дальнейшее расширение объема квазициклов позволяет оценить глубину долговременной памяти.

- Определение свойств квазициклических временных рядов и глубины долговременной памяти временного ряда с помощью только R/S-анализа может оказаться недостаточным. Необходимы дополнительные методы и алгоритмы.

Рассмотрим некоторые проблемы, связанные с погрешностью стандартного алгоритма обратного распространения, который используется в нейронных сетях. В стандартном алгоритме обратного распространения проблема выбора скорости обучения нейронной сети имеет большое значение для повышения эффективности алгоритма. При нахождении адаптивной полосы пропускания выбирается наиболее эффективный алгоритм спуска. Согласно этому алгоритму нейронная сеть для каждой скорости обучения выбиралась таким образом, чтобы средняя квадратическая ошибка была минимальной:

$$\alpha(t) = \min E(y_j(t+1)).$$

Решение зависит от нейронных элементов функции активации и в общем случае определяется следующим образом:

$$w_{ij}(t+1) = w_{ij}(t) - \alpha(t) \frac{\partial E}{\partial w_{ij}(t)}, \quad (2)$$

$$T_j(t+1) = T_j(t) - \alpha(t) \frac{\partial E}{\partial T_j(t)}. \quad (3)$$

Средняя квадратическая ошибка нейронной сети определяется как

$$E = \frac{1}{2} \sum_j (y_j - t_j)^2.$$

Необходимо решить уравнение

$$\frac{\partial E}{\partial \alpha(t)} = \frac{\partial E}{\partial y_j(t+1)} \cdot \frac{\partial y_j(t+1)}{\partial \alpha(t)},$$

которое является сложным для общего случая. Поэтому в [3] для поиска адаптивной скорости обучения $\alpha(t)$ предлагается использовать метод наискорейшего спуска, что связано с большим объемом вычислений. Предложим приближенный метод нахождения $\alpha(t)$, который основан на разложении в ряд Тейлора функции активации нейрона. Рассмотрим этот метод более подробно.

Пусть выходное значение j -го нейрона последнего слоя нейронной сети вычисляется следующим образом:

$$\begin{aligned} y_j(t) &= F(S_j(t)), \\ S_j(t) &= \sum_i y_i(t)w_{ij}(t) - T_j(t), \end{aligned} \quad (4)$$

где $y_i(t)$ — выходное значение i -го нейрона скрытого слоя.

Для вычисления выходного значения j -го нейрона за время $t+1$ (см. (2) и (3))

$$\begin{aligned} S_j(t+1) &= \sum_i y_i \left(w_{ij} - \alpha \frac{\partial E}{\partial w_{ij}} \right) - T_j + \alpha \frac{\partial E}{\partial T_j} = \\ &= \sum_i y_i w_{ij} - T_j + \alpha \left(\sum_j y_j \frac{\partial E}{\partial w_{ij}} - \frac{\partial E}{\partial T_j} \right) \end{aligned} \quad (5)$$

введем обозначения:

$$a_j = \sum_j y_j \frac{\partial E}{\partial w_{ij}} - \frac{\partial E}{\partial T_j}. \quad (6)$$

Тогда уравнение (5) можно представить следующим образом:

$$S_j(t+1) = S_j(t) - \alpha \cdot a_j. \quad (7)$$

Значение j -го нейрона за время $(t+1)$:

$$y_j(t+1) = F(S_j(t+1)).$$

Разложим последнее выражение в ряд Тейлора:

$$y_j(t+1) = F(0) + F'(0) \cdot F(S_j(t+1)), \quad (8)$$

где

$$F'(0) = \frac{\partial F}{\partial S_j},$$

если $S_j(0) = 0$.

Из выражений (8), (7) получим

$$y_j(t+1) = F(0) + F'(0) \cdot S_j(t) - \alpha F'(0) a_j. \quad (9)$$

Поскольку

$$y_j(t) = F(0) + F'(0) \cdot S_j(t),$$

то выражение (9) можно представить следующим образом:

$$y_j(t+1) = y_j(t)F(0) - \alpha F'(0) a_j. \quad (10)$$

Для поиска адаптивной скорости обучения следует найти

$$E = \frac{1}{2} \sum_j (y_j(t+1) - t_j)^2 \rightarrow \min;$$

тогда

$$\frac{\partial E}{\partial \alpha} = \sum_j (y_j(t) - t_j - \alpha F'(0) a_j) \cdot (-F'(0) a_j) = 0.$$

Отсюда получаем

$$\alpha(t) = \frac{\sum_j (y_j(t) - t_j) a_j}{F'(0) \sum_j a_j^2}. \quad (11)$$

Так как $\frac{\partial^2 E}{\partial \alpha^2} > 0$, то $\alpha(t)$ является минимальным стандартным отклонением.

Находим выражение для a_j . Имеем

$$\frac{\partial E}{\partial w_{ij}} = \frac{\partial E}{\partial y_j} \cdot \frac{\partial y_j}{\partial S_j} \cdot \frac{\partial S_j}{\partial w_{ij}} = (y_j - t_j) F'(S_j) y_j, \quad (12)$$

$$\frac{\partial E}{\partial T_j} = \frac{\partial E}{\partial y_j} \cdot \frac{\partial y_j}{\partial S_j} \cdot \frac{\partial S_j}{\partial T_j} = -(y_j - t_j) F'(S_j); \quad (13)$$

если (12) и (13) подставить в уравнение (6), то получим

$$a_j = \left(1 + \sum_i y_i^2 \right) \cdot (y_j - t_j) \cdot F'(S_j). \quad (14)$$

Пусть

$$\gamma_j = y_j - t_j. \quad (15)$$

Тогда с учетом (11)–(15) получим формулу для расчета приблизительной скорости обучения нейронной сети:

$$\alpha(t) = \frac{\sum_j \gamma_j^2 F'(S_j)}{F'(0) \cdot \left(1 + \sum_i y_i^2 \right) \left(\sum_j \gamma_j^2 (F'(S_j))^2 \right)}, \quad (16)$$

где γ_j является j -м элементом ошибки нейронной сети.

Поскольку производная сигмоидной функции имеет вид

$$y'_j = F'(S_j) = y_j(1 - y_j),$$

$$y'_j(0) = F'(0) = \frac{1}{4},$$

то скорость адаптивного обучения нейронной сети представим следующим образом:

$$\alpha(t) = \frac{4 \sum_j \gamma_j^2 y_j (1 - y_j)}{\left(1 + \sum_i y_i^2 \right) \sum_j \gamma_j^2 y_j^2 (1 - y_j)^2},$$

где γ_j — ошибка скрытого слоя j -го нейронного элемента,

$$\gamma_j = \sum_i \gamma_i y_i (1 - y_i) w_{ij},$$

значение i характеризуется нейронами следующего слоя j .

Если для функции активации использовать гиперболический тангенс, то получим

$$y'_j = F'(S_j) = \frac{1}{\operatorname{ch}^2(S_j)} = (1 - y_j^2),$$

$$y'_j(0) = F'(0) = 1.$$

Следовательно,

$$\alpha(t) = \frac{\sum_j \gamma_j^2 (1 - y_j^2)}{\left(1 + \sum_i y_i^2\right) \sum_j \gamma_j^2 (1 - y_j^2)^2},$$

где γ_j — ошибка скрытого слоя j -го нейронного элемента,

$$\gamma_j = \sum_i \gamma_i (1 - y_i^2) w_{ij}.$$

Отметим, что для каждого слоя нейронной сети значение $\alpha(t)$ рассчитывается отдельно. Как показывают численные эксперименты, с использованием повышенной скорости адаптивного обучения абсолютные значения могут быть получены через $\alpha(t)$. Это может привести к появлению эффекта десинхронизации учебного процесса, когда синаптические связи весовых коэффициентов во время обучения в широком диапазоне резко изменяют свои значения, в результате возникает функция колебательной составляющей стандартной ошибки отклонения. Поэтому рекомендуется ограничивать абсолютное значение $\alpha(t)$. Выражения, полученные выше для обучения адаптивной скорости нейронной сети, могут ускорить процесс сетевого обучения и исключить проблему произвольного выбора скорости обучения. Это ощутимое преимущество перед стандартным обратным алгоритмом.

Как и в случае линейной нейронной сети, адаптивная скорость обучающей сети рассчитывается следующим образом:

$$\alpha(t) = \frac{\sum_k \sum_j (y_j^k)^2}{F'(0) \sum_k \sum_j (a_j^k)^2},$$

$$a_j^k = \sum_p \gamma_j^p F'(S_j^p) \left(1 + \sum_i y_i^p y_i^k\right),$$

где $k = \overline{1, L}$; L — количество изображений, после подачи которых вход нейронной сети регулируется синаптическими коэффициентами и порогами нейронов. Следует отметить, что в этом случае прослеживается изменение синаптических связей:

$$w_{ij}(L) = w_{ij}(0) - \alpha(t) \cdot \sum_k y_j^k F'(S_j^k) y_i^k,$$

$$T_j(L) = T_j(0) + \alpha(t) \cdot \sum_k y_j^k F'(S_j^k).$$

Последние выражения можно использовать в пакетном режиме путем изменения весовых коэффициентов нейронной сети.

Адаптивная скорость обучения приближительная. Точность определяется числом членов ряда Тейлора и независимостью слоев. Этот принцип обобщает результаты всех уровней нейронной сети, полученные для последнего слоя. Естественно, для самого общего случая это не всегда правильно, и тогда можно использовать разные скорости для разных уровней нейронной сети.

Теорема 1. Адаптивная скорость обучения последнего скрытого слоя нейронной сети определяется по формуле

$$\alpha_2(t) = \frac{\sum_i C_i \sum_j (y_j - t_j) w_{ij}}{(F'(0))^2 \sum_j \left(\sum_i C_i w_{ij} \right)^2},$$

где

$$C_i = \gamma_i F'(S_i) \left(1 + \sum_k y_k^2 \right),$$

j — выходной слой нейронной сети, представляющий последний скрытый слой; k — скрытый предпоследний слой.

Доказательство. Для i -го нейрона последней взвешенной суммы скрытого слоя на основе (4) запишем следующее соотношение:

$$S_i(t+1) = S_i(t) - \alpha_2 C_i,$$

где

$$C_i = \sum_k y_k \frac{\partial E}{\partial w_{ij}} - \frac{\partial E}{\partial T_i} = \gamma_i F'(S_i) \left(1 + \sum_k y_k^2 \right).$$

Тогда в результате разложения в ряд Тейлора $S_i(t+1)$ выходного значения i -го нейрона имеем

$$y_i(t+1) = F(0) + F'(0)S_i(t+1).$$

После преобразования этого выражения получим

$$y_i(t+1) = y_i(t) - \alpha_2 F'(0)C_i. \quad (17)$$

Взвешенная сумма j -го элемента может быть представлена следующим образом:

$$S_j(t+1) = \sum_i w_{ij} y_i(t+1) - T_j. \quad (18)$$

После подстановки (17) в (18) получаем

$$S_j(t+1) = S_j(t) - \alpha_2 F'(0) \sum_i w_{ij} C_i.$$

Выходное значение j -го элемента последнего слоя определяется как

$$y_j(t+1) = y_j(t) + F'(0)S_j(t+1) = y_j(t) - (F'(0))^2 \alpha_2 \sum_j w_{ij} C_j.$$

Тогда

$$\frac{\partial E}{\partial \alpha_2} = \sum_j \left(y_j(t) - t_j - \alpha_2 (F'(0))^2 \sum_i w_{ij} C_i \right) \left(- (F'(0))^2 \sum_i w_{ij} C_i \right).$$

Приравняв это выражения к нулю и упростив его, получим

$$\alpha_2(t) = \frac{\sum_i C_i \sum_j (y_j - t_j) w_{ij}}{(F'(0))^2 \sum_j \left(\sum_i C_i w_{ij} \right)^2},$$

что и требовалось доказать.

Приведенные выше выражения позволяют выбрать скорость адаптивного обучения для нейронных сетей, имеющих один или два скрытых слоя.

Таким образом, имеется возможность предварительного прогнозирования финансовых временных рядов с использованием алгоритма Херста (с использованием R/S-анализа). Из анализа следует, что во временном ряду можно влиять на долговременную память для оценивания ее глубины и при этом обнаруживать наличие циклов (квазициклов). Тем не менее, R/S-анализ не является исчерпывающим предварительным исследованием, поскольку он не всегда дает полную информацию о поведении временных рядов без применения дополнительных методов и алгоритмов.

СПИСОК ЛІТЕРАТУРИ

1. Peters E.E. Chaos and order in the capital markets: A new view of cycles, prices, and market violability. Wiley Finance, 1996. 288 p.
2. Feder J. Fractals. Premium Press, 1990. 283 p.
3. Mandelbrot B. The fractal geometry of nature. New York: W.H. Freeman, 1983. 270 p.
4. Дидманідзе І., Кахіані Г. Искусственная нейронная сеть и управление инвестициями фондового рынка. *Теоретичні та прикладні аспекти побудови програмних систем. ТААПСД'2013: Праці конференції*. Ялта, 2013. С. 52–53.

Надійшла до редакції 21.02.2019

А.Д. Шаташвілі, І.Ш. Дідманідзе, Г.О. Кахіані, Т.О. Фоміна **ПРО ОДИН МЕТОД ПОПЕРЕДНЬОГО ПРОГНОЗУВАННЯ ЧАСОВИХ РЯДІВ** **ФІНАНСОВИХ ДАНИХ**

Анотація. Розглянуто проблему прогнозування часових рядів цін акцій провідних світових компаній, яким властива довготермінова пам'ять. Зроблено припущення, що у разі застосування традиційних методів аналізу ігнорування наявності подібної кореляційної структури часових рядів призводить до появи значно більшої похибки, ніж врахування довготермінової пам'яті за фактичної її відсутності. Передбачається, що коливання цін на інструменти фінансового ринку описуються процесом Герста, який моделює процеси з довготерміновою пам'яттю. Такий часовий ряд не можна ефективно аналізувати за допомогою традиційних стаціонарних моделей, які повністю ігнорують цей факт. Ставиться задача — з використанням розглянутого методу встановити наявність довготермінової пам'яті у вихідного часового ряду і визначити його тип.

Ключові слова: часові ряди, фрактал, нейронні мережі.

A.D. Shatashvili, I.Sh. Didmanidze, G.A. Kakhiani, T.A. Fomina **A METHOD OF PRELIMINARY FORECASTING OF TIME SERIES OF FINANCIAL DATA**

Abstract. The problem of forecasting the time series of stock prices of leading global companies that are characterized by long-term memory is considered. It is assumed that ignoring the presence of such a correlation structure in time series using traditional methods of analysis leads to a much greater error than taking into account long-term memory in its actual absence. It is assumed that the daily fluctuations in prices for financial market instruments are the Hurst process, that is, they have long-term memory, which means such a time series cannot be effectively analyzed using traditional stationary models that completely ignore this fact. Thus, the task is set, using the R/S analysis method, to determine the presence of long-term memory in the initial time series, to determine its type.

Keywords: data series, fractal, artificial neural networks.

Шаташвіли Альберт Даниелович,

доктор физ.-мат. наук, профессор Батумского государственного университета имени Шота Руставели, Грузия, e-mail: shatal@bk.ru.

Дидманідзе Ібраїм Шотаєвич,

кандидат физ.-мат. наук, доктор философии, профессор Батумского государственного университета имени Шота Руставели, Грузия, e-mail: ibraim.didmanidze@bsu.edu.ge; ibraimd@mail.ru.

Кахіані Григорій Александрович,

академический доктор информатики, ассоциированный профессор Батумского государственного университета имени Шота Руставели, Грузия, e-mail: gregory.kakhiani@bsu.edu.ge.

Фоміна Тамара Александровна,

кандидат физ.-мат. наук, доцент Батумского государственного университета имени Шота Руставели, Грузия, e-mail: shatal@bk.ru.