УДК 519.2 URL: http://ptsj.ru/catalog/iemim/sta/992.html

ПРИМЕНЕНИЕ КИНЕТИЧЕСКОГО ПОДХОДА К ПРОГНОЗИРОВАНИЮ ЦЕН ЗАКРЫТИЯ АКЦИЙ

Е.Р. Сайкова saykovaer@student.bmstu.ru

МГТУ им. Н.Э. Баумана, Москва, Российская Федерация

Рассмотрена задача прогнозирования значений временного ряда на основе кинетического подхода. Проблема незамкнутости системы уравнений для моделирования эволюции выборочной плотности функции распределения решена с помощью введения прогнозного предположения. Предложен альтернативный способ получения прогноза на более чем один шаг по времени вперед с использованием базовой модели на основе уравнения Лиувилля. Реализован программный код для построения прогноза с использованием кинетического подхода для цен закрытия акций различных компаний. Представлены результаты его работы (в виде оценки ошибки прогнозирования как средней абсолютной ошибки в процентах — MAPE) для десяти нестационарных временных рядов.

Ключевые слова: нестационарные временные ряды, выборочная плотность функции распределения, уравнение Лиувилля, прогнозирование значений временного ряда, кинетический подход к прогнозированию временных рядов, цена закрытия акции

Введение. Задача прогнозирования значений временного ряда возникает в самых разных областях человеческой деятельности, таких как планирование бизнеса, распределение ресурсов и др. Цель данной работы — применение кинетического подхода к прогнозированию цен закрытия акций.

Рассматриваемые ряды являются нестационарными. И если стационарные ряды хорошо изучены, то общая теория анализа нестационарных временных рядов все еще находится в разработке. Трудности, возникающие при построении математической модели нестационарного временного ряда, связаны с тем, что статистические свойства такого ряда меняются со временем.

В монографии [1] предложен кинетический подход, в рамках которого задачу прогнозирования значений временного ряда можно рассматривать как задачу построения модели эволюции выборочной плотности функции распределения (ВПФР).

Предобработка ряда. Считаем, что значения временного ряда принадлежат отрезку $[0;1]$. Для произвольных данных этого можно добиться следующим преобразованием:

$$\tilde{x} = \frac{x - \min x}{\max x - \min x},$$

 ISSN 2541-8009

где $\tilde{x}$ — ряд, преобразованный на отрезок [0;1]; x — преобразуемый ряд.

Также из рассматриваемых данных стандартными методами удалим полиномиальный и периодический тренды.

Кинетический подход. Для моделирования эволюции ВПФР примéним кинетическое уравнение Лиувилля. Для ВПФР $f_T(x,t)$, построенной по выборке объема T, оно записывается следующим образом:

$$\frac{\partial f_T(x,t)}{\partial t}+\frac{\partial}{\partial x}\big(f_T(x,t)u_T(x,t)\big)=0, \tag{1}$$

где $u_T(x,t)$ — эмпирическая лиувиллева скорость, или ЭЛС [1].

Поскольку ВПФР $f_T(x,t)$ строится по набору данных, причем конечного объема T, описывать ее эволюцию следует дискретным уравнением. Однако для удобства восприятия будем использовать непрерывное описание. Так как объем выборки полагается фиксированным, индекс T в дальнейшем опустим.

Из (1) видно, что ЭЛС в момент времени t определяется значениями ВПФР в моменты времени t и $t+1$. То есть в текущий момент t известно только значение ЭЛС в момент времени $t-1$. Таким образом, для построения прогноза ВПФР в момент времени $t+1$ нужно предварительно спрогнозировать значение ЭЛС в момент времени t.

Введем двумерную выборочную плотность совместного распределения ряда и его приращений $\dot{x}(t)\equiv x(t+1)-x(t)$ ($\dot{x}\in[-1;1]$), которую обозначим как $F(x,\dot{x},t)$. Тогда, согласно [1], для ЭЛС имеет место следующее соотношение:

$$f(x,t)u(x,t)=\int_{-1}^{1}\dot{x}\,F(x,\dot{x},t)d\dot{x}.$$

Однако двумерная ВПФР $F(x,\dot{x},t)$ пока никак не определена и ее нельзя получить из каких-либо соотношений (поскольку при ее построении используется значение ряда в момент времени $t+1$). Для $F(x,\dot{x},t)$ можно также записать уравнение Лиувилля, введя приращение $\ddot{x}$ и расширив тем самым фазовое пространство, однако это приведет к аналогичной проблеме [1]. Таким образом получается цепочка бесконечных расширений фазового пространства, которую непонятно, в какой момент оборвать. При этом хорошо бы вообще замкнуть уравнение Лиувилля для одномерной ВПФР $f(x,t)$. Но тогда необходимо независимо определить величину $u(x,t)$.

Прогнозное предположение. Если записать уравнение (1) в дискретном виде, выбрав единичный шаг дискретизации по обеим переменным (x и t), и оставить слева только $f(x,t+1)$, получится следующее выражение:

$$f(x,t+1)=f(x,t)+f(x,t)u(x,t)-f(x+1,t)u(x+1,t). \tag{2}$$

Введем следующее прогнозное предположение: пусть скорость в момент времени t равна скорости в момент времени $t-1$ [1]. Таким образом можно избавиться от возникшей при записи уравнения Лиувилля проблемы необходимости прогнозировать текущее значение ЭЛС. Получим для выражения (2) следующее:

$$\tilde{f}(x,t+1) = f(x,t) + f(x,t)u(x,t-1) - f(x+1,t)u(x+1,t-1), \quad (3)$$

где $\tilde{f}(x,t+1)$ — прогнозируемое значение ВПФР.

Поскольку требуется построить прогноз не на один, а на τ шагов вперед, предлагаем добавлять полученное на предыдущем этапе (спрогнозированное) значение временного ряда в выборку (убирая при этом из нее самое «старое» значение) и повторять процедуру получения прогноза на один шаг.

Определение объема выборки. Подробно методика определения оптимального объема выборки описана в [1] и [2]. Однако это представляет собой отдельную задачу, поэтому будем использовать минимально допустимый объем выборки, который оценивается сверху величиной $[2\tau/\varepsilon]$, где ε — допустимый уровень нестационарности ВПФР, τ — горизонт прогноза, а квадратные скобки обозначают целую часть числа [2]. В дальнейшем будем считать, что $\varepsilon = 0{,}05$, а объем выборки равен $[2\tau/\varepsilon]$.

Вычисление ЭЛС. Для определения ЭЛС нужно предварительно построить:

– массив из данных объема T обработанного временного ряда x и массив из ряда первых разностей $p \equiv \dot{x}$;

– одномерную ВПФР $f(x,t)$;

– двумерную ВПФР $F(x,p,t)$.

Строить одномерную и двумерную ВПФР $f(x,t)$ и $F(x,p,t)$ будем с помощью ядерной оценки плотности распределения Парзена [3]. В качестве ядра Парзена будем использовать гауссово ядро, которое выглядит как плотность стандартного нормально распределения [4]:

$$\mathcal{K}(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}.$$

Теперь, чтобы найти ЭЛС, вычислим интеграл $\int_{-1}^{1} pF(x,p,t)dp$ по формуле Симпсона [5].

Затем массив полученных значений поделим на массив значений $f(x,t)$. По краям, где значение ВПФР близко к нулю, ЭЛС полагается также равной нулю. Будем считать, что там, где $f(x,t) < 10^{-7}$, $u(x,t) = 0$.

 ISSN 2541-8009

Получение прогнозных значений. После применения прогнозного предположения и схемы (3) получаем спрогнозированную ВПФР. С ее помощью ищем прогноз самого временного ряда:

$$\hat{x} = \int_0^1 x f(x, t+1) dx,$$

где $\hat{x}$ — прогноз для предварительно обработанного временного ряда; x — значения, в которых вычислена спрогнозированная ВФПР; $f(x, t+1)$ — спрогнозированная ВПФР.

Затем с $\hat{x}$ надо выполнить ряд преобразований: обратное линейное преобразование из отрезка $[0;1]$, прибавление последнего известного элемента (обратное преобразование из ряда первых разностей). Таким образом получается прогноз следующего (после последнего известного) значения временного ряда. Как уже упоминалось, для получения дальнейших прогнозных значений будем использовать модификацию выборки и повторение процедуры получения прогноза на один шаг по времени вперед.

Применение к прогнозированию цен закрытия акций. Рассмотрим процесс построения прогноза значений временного ряда на $\tau = 7$ дней на примере данных о ценах закрытия акций компании Microsoft (Microsoft Corporation Common Stock) в период с 28.02.2013 по 24.02.2023 (всего 2516 наблюдений) [6]. Исходный временной ряд изображен на рис. 1.

Автокорреляционная функция (АКФ) ряда представлена на рис. 2.

График АКФ первых разностей показан на рис. 3.

Так, путем перехода к первым разностям из рассматриваемого временного ряда был удален линейный тренд. Далее полученный ряд отобразим на отрезок $[0;1]$.

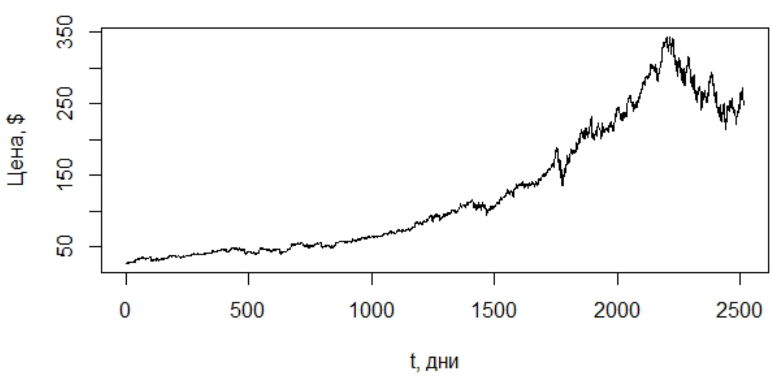

Рис. 1. Временной ряд из цен закрытия акций Microsoft

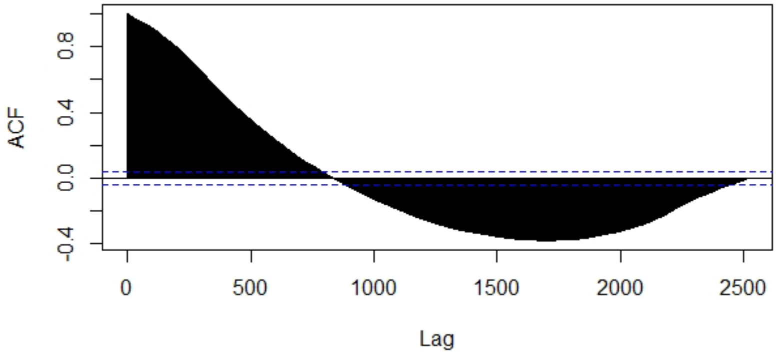

Рис. 2. График АКФ

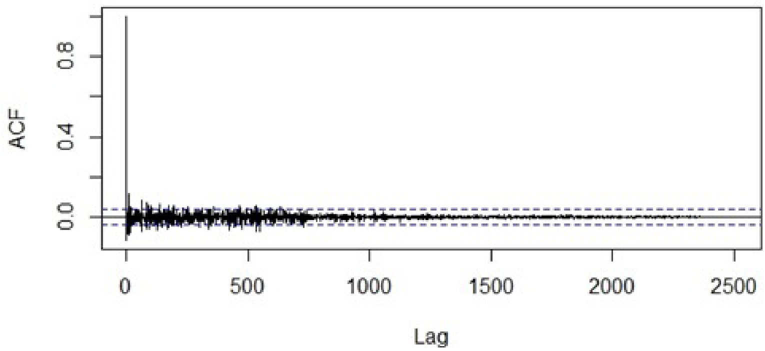

Рис. 3. График АКФ первых разностей

Объем выборки, который будет использоваться для построения прогноза, равен 280 элементам. Чтобы составить массив x, необходимо взять не T последних элементов, а T элементов, исключая последний, поскольку можно найти лишь ЭЛС $u(x,t-1)$.

Графическое изображение ядерной оценки двумерной плотности функции распределения в качестве иллюстрации представлено на рис. 4.

Приступим к прогнозированию ВПФР. Будем использовать уже упомянутое прогнозное предположение, согласно которому за один шаг по времени ЭЛС не изменилась $\left(u(x,t)=u(x,t-1)\right)$, и разностную схему (5). Для этого предварительно построим другую ВПФР того же объема — $f(x,t)$. Для нее нужна уже выборка, в которую включены последние T элементов ряда.

Визуальное сравнение исходной и спрогнозированной ВПФР представлено на рис. 5.

Полученный прогноз для цен закрытия акций компании Microsoft продемонстрирован на рис. 6.

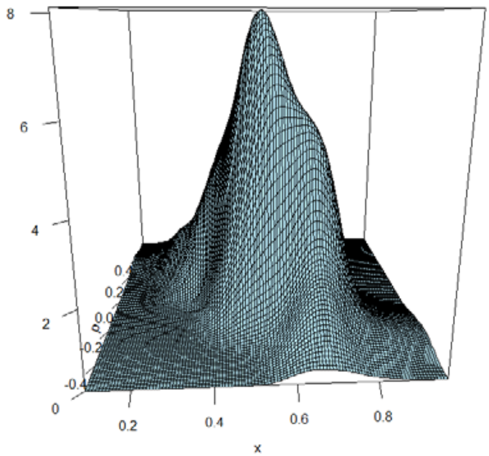

Рис. 4. Двумерная ВПФР

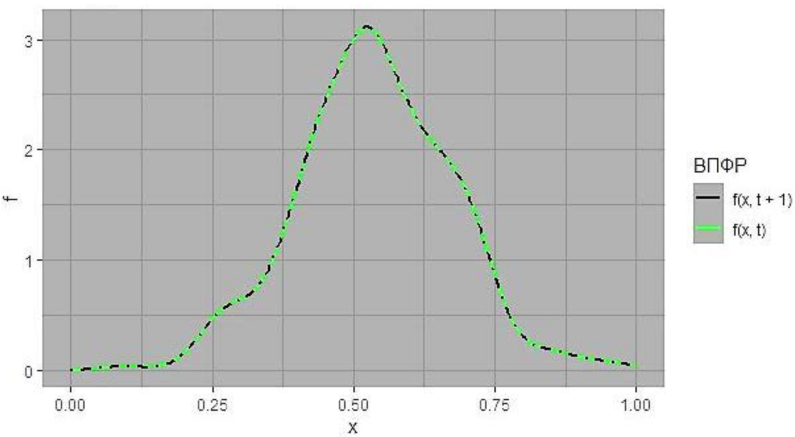

Рис. 5. Исходная и спрогнозированная ВПФР

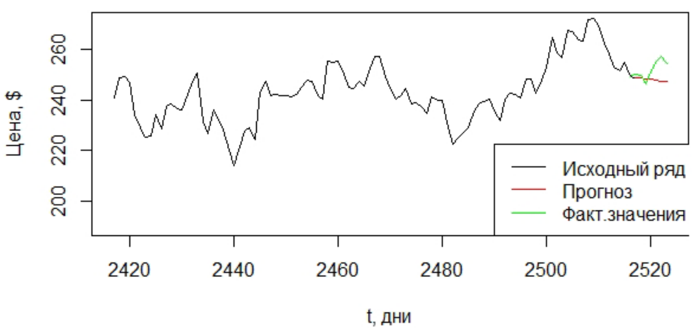

Рис. 6. Прогноз цен закрытия акций компании Microsoft (последние 100 значений и прогноз)

Аналогичные действия были применены ко всем рассматриваемым временным рядам (данные взяты из [6]), после чего по фактическим значениям и полученным прогнозам были оценены ошибки прогнозирования. Для оценки ошибок применяли среднюю абсолютную ошибку в процентах (MAPE) [7]. Результаты прогнозирования (MAPE) с помощью модели на основе кинетического подхода представлены ниже:

Amazon.com (Amazon.com, Inc. Common Stock) 1,61
Apple (Apple Inc. Common Stock) 2,16
Disney (Walt Disney Company (The) Common Stock) 0,86
Ford (Ford Motor Company Common Stock) 6,79
General Motors (General Motors Company Common Stock) 2,31
Lenovo (Lenovo Group Ltd) 5,13
Microsoft (Microsoft Corporation Common Stock) 1,77
Panasonic (Panasonic Holdings Corporation) 3,17
Visa (Visa Inc.) 1,26
Volkswagen (Volkswagen AG) 7,36

Выводы. Таким образом, применение кинетического подхода к временным рядам цен закрытия акций дает приемлемые результаты — лишь в 3 случаях из 10 значение ошибки превысило 5 %. Ни в одном из случаев ошибка не превысила 8 %.

Отметим, что в перспективе точность прогноза по этой модели может быть еще повышена, например, с помощью определения оптимального объема выборки [1].

Литература

[1] Орлов Ю.Н., Осминин К.П. *Нестационарные временные ряды: Методы прогнозирования с примерами анализа финансовых и сырьевых рынков*. Москва, ЛЕНАНД, 2023, 384 с.

[2] Орлов Ю.Н., Осминин К.П. Методика определения оптимального объема выборки для прогнозирования нестационарного временного ряда. *Информационные технологии и вычислительные системы*, 2008, № 3, с. 3–13.

[3] Lerasle M., Magalhães N., Reynaud-Bouret P. Optimal kernel selection for density estimation. High Dimensional *Probability VII. The Cargese. Prog. Probab. Birkhauser*, 2016, vol. 71, pp. 425–460. https://doi.org/10.1007/978-3-319-40519-3_19

[4] Мэрфи К.П. *Вероятностное машинное обучение: введение*. Москва, ДМК Пресс, 2022, 990 с.

[5] Горлач Б.А. *Математический анализ*. Санкт-Петербург, Лань, 2022, 608 с.

[6] *Фондовая биржа Nasdaq*. URL: https://www.nasdaq.com/ (дата обращения 25.02.2023).

[7] Хапке Х., Нельсон К. *Разработка конвейеров машинного обучения. Автоматизация жизненных циклов модели с помощью TensorFlow*. Москва, ДМК Пресс, 2021, 346 с.

Поступила в редакцию 07.05.2024

Сайкова Евгения Романовна — студентка кафедры «Вычислительная математика и математическая физика», МГТУ им. Н.Э. Баумана, Москва, Россия.

Научный руководитель — Облакова Татьяна Васильевна, канд. физ.-мат. наук, доцент кафедры «Вычислительная математика и математическая физика», МГТУ им. Н.Э. Баумана, Москва, Россия.

Ссылку на эту статью просим оформлять следующим образом:
Сайкова Е.Р. Применение кинетического подхода к прогнозированию цен закрытия акций. *Политехнический молодежный журнал*, 2024, № 04 (93). URL: https://ptsj.bmstu.ru/catalog/math/compmath/992.html

KINETIC APPROACH APPLICATION IN FORECASTING THE STOCK CLOSING PRICES

E.R. Saykova saykovaer@student.bmstu.ru

Bauman Moscow State Technical University, Moscow, Russian Federation

The problem of forecasting time series values based on the kinetic approach is considered. The problem of non-closure of the system of equations for modeling the evolution of the sample density of the distribution function is solved by introducing a forecast assumption. An alternative method for obtaining a forecast for more than one time step ahead using a basic model based on the Liouville equation is proposed. A program code for constructing a forecast using the kinetic approach for closing prices of shares of various companies is implemented. The results of its work (in the form of an estimate of the forecast error as the mean absolute percentage error — MAPE) for ten non-stationary time series are presented.

Keywords: non-stationary time series, distribution function sample density, Liouville equation, time series values forecasting, kinetic approach to stock closing price forecasting

Received 07.05.2024

Saykova E.R. — Student of the Department of Computational Mathematics and Mathematical Physics, Bauman Moscow State Technical University, Moscow, Russia.

Scientific advisor — Oblakova T.V., Ph. D. (Phys. & Math.), Associate Professor of the Department of Computational Mathematics and Mathematical Physics, Bauman Moscow State Technical University, Moscow, Russia.

Please cite this article in English as:
Saykova E.R. Kinetic approach application in forecasting the stock closing prices. *Politekhnicheskiy molodezhnyy zhurnal*, 2024, no. 04 (93). (In Russ.). URL: https://ptsj.bmstu.ru/catalog/math/compmath/992.html

 ISSN 2541-8009