УДК 004.942

Хайруллина Лилия Эмитовиа, Хакимов Зульфат Ниязович,
Хабибуллина Гузель Забировна

ПРИМЕНЕНИЕ ТЕХНОЛОГИЙ WOLFRAM MATHEMATICA В ВЕЙВЛЕТ-ОБРАБОТКЕ РЯДОВ ФИНАНСОВЫХ ДАННЫХ

В анализе финансовых временных рядов одним из ключевых моментов является предварительная обработка данных. Целью данной работы является демонстрация возможностей системы Wolfram Mathematica в препроцессинге финансовых данных. В качестве объекта исследования выбрана история ежедневных котировок акций Сбербанка за последние 3 года. Трешолдинг проводился на ортогональном вейвлете Добеши 6-го порядка. По соотношению сигнал/шум наиболее привлекательным оказался способ очистки на многоуровневом пороге на основе несмещенной оценки риска Штейна.

Финансовый ряд, трешолдинг, вейвлеты, Wolfram Mathematica.

Khairullina Lilia Emitovna, Khakimov Zulfat Niyazovich,
Khabibullina Guzel Zabirovna

APPLICATION OF WOLFRAM MATHEMATICA TECHNOLOGIES IN THE WAVELET PROCESSING OF FINANCIAL DATA SERIES

In the analysis of financial time series, one of the key points is the preprocessing of data. The purpose of this work is to demonstrate the capabilities of the Wolfram Mathematica system in financial data preprocessing. The history of Sberbank's daily stock quotes over the past 3 years has been selected as the object of the study. Tresholding was carried out on orthogonal Daubechies wavelets of the 6th order. In terms of signal-to-noise ratio, the most attractive method turned out to be cleaning at a multi-level threshold based on an unbiased Stein risk assessment.

Financial series, threading, wavelets, Wolfram Mathematica.

Введение. Анализ временных рядов финансовых данных является сложной задачей из-за их нелинейности и нестационарности [1-4]. Нестационарность выражается в изменчивости статистических характеристик ряда (математического ожидания, дисперсии) с течением времени и проявляется в виде тренда, цикличности, сезонности, случайном блуждании или их сочетании. Нестационарные сигналы чаще всего состоят из кратковременных высокочастотных и длительных низкочастотных компонентов, поэтому для анализа таких сигналов предпочтительнее использовать вейвлет-анализ. Чаще всего при проведении вейвлет-анализа исследователями используется система Matlab, языки программирования R или Python. В настоящее время актуальной научной теоретической и практической задачей является поиск новых программно-аппаратных средств, позволяющих быстро и качественно проводить сложные вычисления. Одним из лидеров среди существующих программных продуктов является Wolfram Mathematica. К преимуществам применения данной системы относится и то, что благодаря возможности облачных вычислений (WolframCloud) пользователь становится слабозависимым от аппаратного обеспечения. Данная статья имеет своей целью демонстрацию предварительной обработки временного ряда финансовых данных с помощью вейвлет-анализа в системе компьютерной математики Wolfram Mathematica.

Вейвлет-анализ и пороговая обработка сигнала. Вейвлет-преобразование сигнала $f(t)$ состоит в его разложении в ряд:

$$f(t) = c_0\psi(t) + \sum_{j=1}^{N}\sum_{k=0}^{2^j-1} c_{jk}\varphi_{jk}(t),$$

где $\psi(t)$ – называют отцовским вейвлетом, $\varphi_{jk}(t)$ – функции, полученные из материнского вейвлета $\varphi(t)$ путем сжатия и сдвига. Эти функции удовлетворяют условиям $\int_{-\infty}^{+\infty} \varphi(t)dt = 1,\ \int_{-\infty}^{+\infty} \psi(t)dt = 0.$

Согласно теории многоуровневой аппроксимации сигналов, построенной Mallat [5], любой сигнал можно представить в виде суммы аппроксимирующего и детализирующих сигналов. При этом коэффициенты аппроксимации содержат полезную информацию о поведении сигнала (грубое приближение сигнала). Коэффициенты детализации на каждом уровне разложения содержат уточняющую информацию о сигнале. В этих же коэффициентах могут содержаться шумы - случайные колебания данных. Они не несут полезной информации о главных тенденциях временного ряда данных и могут исказить интерпретацию сигнала. Поэтому анализ временных рядов целесообразно предварять очисткой от шумов. Одним из способов удаления шумов является трешолдинг - разложение сигнала на вейвлет-спектр и последующее обнулении тех вейвлет-коэффициентов разложения, значения которых меньше определенного порогового значения [7-9].

Существуют разные виды пороговой обработки сигнала, из которых выделяют мягкий и жесткий. Жесткий трешолдинг предполагает обнуление всех коэффициентов, меньших заданного порога, остальные коэффициенты сохраняются неизменными. Мягкий трешолдинг дополнительно предполагает уменьшение по модулю всех остальных коэффициентов на величину порога. Порог может быть единым для всех коэффициентов и уровней разложения или различным в зависимости от частотных особенностей сигнала. От выбора порогового уровня зависит качество сглаживания сигнала – слишком малый порог не устранит шумы, слишком большой может привести к потере коэффициентов, содержащих существенную информацию о сигнале [7].

Инструменты вейвлет-обработки в Wolfram Mathematica. Язык Wolfram Language обеспечивает полнофункциональную реализацию вейвлет-анализа [10]. Дискретное вейвлет-преобразование реализуется с помощью функции DiscreteWaveletTransform[]. По умолчанию применяется вейвлет Хаара. В качестве материнского вейвлета в зависимости от характерных особенностей сигнала можно использовать вейвлеты: BiorthogonalSplineWavelet[], CoifletWavelet[], DaubeshiesWavelet[] и др.

Восстановление сигнала осуществляется с помощью функции InverseWaveletTransform [].

Для пороговой очистки используется функция WaveletThreshold:

WaveletThreshold[*dwd*, *tspec*, *wind*],

где *dwd* -коэффициенты дискретного вейвлет-преобразования, *tspec* – характеристика порога, *wind* – индексы вейвлет- коэффициентов, к которым будет применяться пороговая обработка.

В качестве характеристики порога можно выбрать встроенные: Hard (жесткая обработка с порогом δ), Soft (мягкая обработка с порогом δ) и некоторые другие. Пороговое значение δ также можно выбрать из имеющихся, например:

- «SURE» - Stein’s unbased risk estimation, несмещенная оценка риска Штейна,

- «SURELevel» - порог «SURE», применяемый для каждого уровня разложения сигнала,

- «Universal» - универсальный порог Донохо-Джонстона, и другие.

3. Эмпирическое исследование. В качестве объекта исследования выберем историю котировок акций Сбербанка за период с 1.01.2020 по 30.04.2024, в качестве единицы исследования примем ежедневные торги. Фурье-спектр этого ряда наглядно демонстрирует наличие высокочастотных шумов (рис.1):

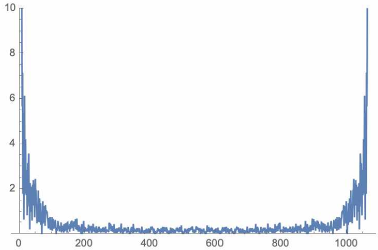

Рис.1 Фурье-спектр сигнала

Эффективность процедуры очистки сигнала в значительной степени зависит от вейвлет-базиса. Для выбора оптимального вейвлета будем использовать критерий минимума энтропий:

$$WE = -\sum_{j=1}^{N} p_j \ln(p_j),$$

где $p_j = \dfrac{E_j}{\sum_{j=1}^{N} E_j}$ - относительная вейвлет-энергия, $E_j = \sum_k d_{j,k}^2$, N – максимальный уровень разложения. В качестве наилучшей выберем ту вейвлет-функцию, которой будет соответствовать наименьшая энтропия.

Для расчета вейвлет-энтропии использовалась опция EnergyFraction функции DiscreteWaveletTransform системы Wolfram Mathematica:

```
dwd = DiscreteWaveletTransform[dat, DaubechiesWavelet[4], 4]
efrac = dwd["EnergyFraction"];
ED4 = -Sum[efrac[[i, 2]] Log[efrac[[i, 2]]], {i, 1, Length[efrac}]
```

Энтропия рассчитывалась для вейвлетов Хаара, Мейера, Добеши и Симлета разных порядков до 4 уровня разложения. Анализ вейвлет-энтропий показал, что наиболее предпочтительным является использование вейвлета Добеши 6-го порядка.

Очистка от шумов проводилась по детализирующим коэффициентам. В качестве порогового значения использовалась встроенная несмещенная оценка риска Штейна и его модификация, порог Донохо-Джонстона. В Wolfram Mathematica очистка сигнала реализуется командами:

```
dwt = DiscreteWaveletTransform[dat, DaubechiesWavelet[6], 4]
WT = WaveletThreshold[dwt, {"Hard", "SURE"}];
datthr = InverseWaveletTransform[WT];
```

Качество очистки оценивалось с помощью отношения сигнал/шум, определяемого формулой:

$$SNR = 10 lg \frac{\sigma^2_{восст}}{\sigma^2_{шум}},$$

где $\sigma^2_{\text{восст}}$ – дисперсия восстановленного сигнала, $\sigma^2_{\text{шум}}$ – дисперсия шума.

Результаты очистки временного финансового ряда сведем в таблицу.

Табл.1

Результаты пороговой очистки вейвлетом Добеши 6 порядка

Порог	Шумовой порог на каждом уровне разложения	SNR при пороговой обработке «Hard»	SNR при пороговой обработке «Soft»
"SURE"	1) 0,443732 2) 0,443732 3) 0,443732 4) 0,443732	56,822	44,286
"SURELevel"	1) 0,443732 2) 0,340379 3) 0,007522 4) 0,213248	57,492	45,849
"Universal"	1) 7,43896 2) 7,43896 3) 7,43896 4) 7,43896	28,274	25,509

Выводы. Система Wolfram Mathematica обладает мощным функционалом, позволяющим проводить качественную обработку временных рядов. Анализ полученных результатов показал, что на качество очистки сигнала влияет выбор базисного вейвлета – в нашем случае предпочтительным оказалось использование вейвлета Добеши 6-го порядка. Максимальное отношение сигнал/шум достигается при жесткой пороговой обработке с порогом «SURELevel». Проведенные исследования показали, что вейвлет-трешолдинг над детализирующими коэффициентами вейвлет-разложения является эффективным методом подавления выбросов и флуктуаций временного ряда. Очищенный сигнал повторяет форму исходного сигнала, все пики хорошо выражены

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. *Yasar M., Ray A.* Trend detection and data mining via wavelet and Hilbert-Huang transforms. Proceedings of the American Control Conference. 2008. p. 4292 - 4297. DOI = 10.1109/ACC.2008.4587168.

2. *Chaovalit P., Gangopadhyay A., Karabatis G. and Chen Z.* 2011. Discrete wavelet transform-based time series analysis and mining. ACM Comput. Surv. 43, 2, Article 6 (January 2011), 37 pages. URL://doi.acm.org/10.1145/1883612.1883613

3. *Lee H.Y., Beh W.L. & Lem K.H.* Wavelet as a Viable Alternative for Time Series Forecasting // Austrian Journal of Statistics. 2020. 49, p.38-47.

4.*Shaikh W., Syed F., Pandhiani S. & Solangi M.* Wavelet Decomposition Impacts on Traditional Forecasting Time Series Models // Computer Modeling in Engineering & Sciences. 2022. V.130. p.1517-1532. DOI=10.32604/cmes.2022.017822.

5. *Mallat S.G.* A Theory For Multiresolution Signal Decomposition: The Wavelet Representation / IEEE Transactions on Pattern Analysis and Machine Intelligence, 1989, №11, pp. 674–693.

6. *Манонина И.В.* Обработка детализирующих вейвлет-коэффициентов для повышения точности рефлектометрических измерений // Научный вестник МГТУ ГА. 2016. №5. URL: https://cyberleninka.ru/article/n/obrabotka-detaliziruyuschih-veyvlet-koeffitsientov-dlya-povysheniya-tochnosti-reflektometricheskih-izmereniy (дата обращения: 22.04.2024).

7. *Белов А.А.* Сглаживание временных рядов на основе вейвлет-преобразования в системах автоматизированного экологического мониторинга / А.А. Белов, А.Ю. Проскуряков // Методы и устройства передачи и обработки информации. - 2010. - № 1 (12). - С. 21-24.

8. *Кропотов Ю.А., Белов А.А., Проскуряков А.Ю.* Обработка временных рядов с применением вейвлет-преобразований для повышения точности представления информации // Транспортное машиностроение. 2018. №8 (69). URL: https://cyberleninka.ru/article/n/obrabotka-vremennyh-ryadov-s-primeneniem-veyvlet-preobrazovaniy-dlya-povysheniya-tochnosti-predstavleniya-informatsii (дата обращения: 22.04.2024).

9. *Московский С.Б., Сергеев А.Н., Лалина Н.А.* Очистка сигнала от шумов с использованием вейвлет-преобразования // Universum: Технические науки: электрон. научн. журн. 2015. № 2 (15) URL: http://7universum.com/ru/tech/archive/item/1958

10. Wavelet Analysis. URL: https://reference.wolfram.com/language/guide/Wavelets.html

Хайруллина Лилия Эмитовна, кандидат физико-математических наук, доцент кафедры информационных систем Института вычислительной математики и информационных технологий КФУ, Россия, город Казань, улица Кремлевская, 35, 420111, телефон: +7 (9272) 46-40-03, email: lxayrullina@yandex.ru.

Хакимов Зульфат Ниязович, аспирант, Институт вычислительной математики и информационных технологий КФУ, Россия, город Казань, улица Кремлевская, 35, 420111, email: znkhakimov@stud.kpfu.ru.

Хабибуллина Гузель Забировна, кандидат педагогических наук, доцент кафедры теории относительности и гравитации Института физики КФУ, Россия, город Казань, улица Кремлёвская, 16а, 420008, телефон +7(9274)05-70-21, email: hgz1980@rambler.ru

Khairullina Lilia Emitovna, Candidate of Physical and Mathematical Sciences, Associate Professor of the Department of Information Systems of the Institute of Computational Mathematics and Information Technologies of KFU, Kazan, 35 Kremlevskaya Street, 420111, Russia, phone: +7 (9272) 46-40-03, email: lxayrullina@yandex.ru.

Khakimov Zulfat Niyazovich, PhD student, Institute of Computational Mathematics and Information Technologies of KFU, Kazan, 35 Kremlevskaya Street, 420111, Russia, email: znkhakimov@stud.kpfu.ru

Khabibullina Guzel Zabirovna, Candidate of Pedagogical Sciences, Associate Professor of the Department of Relativity and Gravity at the Institute of Physics of KFU, Kazan, Russia, Kremlevskaya Street, 16a, 420008, phone +7(9274)05-70-21, email: hgz1980@rambler.ru