ПРИМЕНЕНИЕ ТЕОРИИ ВЕЙВЛЕТ-ПРЕОБРАЗОВАНИЙ В ИССЛЕДОВАНИИ ФИНАНСОВЫХ ВРЕМЕННЫХ РЯДОВ

Огородов А.П.

ГОУВПО «Мордовский государственный университет им. Н. П. Огарева», г. Саранск
Тел.: +7 (917) 692-35-77, e-mail: nkogorodov@mail.ru

Аннотация. В статье автор говорит о важности эффективного кодирования информации в прогнозировании финансовых временных рядов. Для минимизации шумовой составляющей входного сигнала предлагается использовать вейвлет-анализ. На примере исследования данных о ценах и объемах продаж акций ОАО «Сургутнефтегаз» за 2006 г. показываются положительные результаты использования данного подхода. Также проанализированы фрактальные свойства финансовых временных рядов.
Ключевые слова: вейвлет-анализ, фондовый рынок, показатель Херста, прогнозирование.

В каждой реальной задаче присутствует шум, и необходимо уметь справляться с ним. В особенности это относится к задачам обработки временных рядов, в которых переменные получены в результате измерений в некоторой физической системе, причем в самой системе и/или в механизме измерений шум присутствует естественным образом. В финансовых приложениях данные зашумлены особенно сильно. Например, совершение сделок может регистрироваться в базе данных с запозданием, причем в различных случаях – с разным. Пропуск значений или неполную информацию также иногда рассматривают как шум: в таких случаях берется среднее или наилучшее значение, и это приводит к зашумлению базы данных.

Поэтому для успешного прогнозирования необходима эффективная обработка входных данных, в частности, минимизация случайных флуктуаций и шума. Понизить шумовую составляющую можно, прибегнув к вейвлет-анализу.

При многоуровневом одномерном вейвлет-анализе сигнал раскладывается на аппроксимирующие коэффициенты cA_N и детализирующие коэффициенты $cD_1....cD_N$. Эти векторы получаются сверткой исходного сигнала s с фильтром нижних частот Lo_D для аппроксимации и с фильтром высоких частот Hi_D для детализации, а затем сопровождаются двоичной децимацией. Вейвлет-разложение сигнала s, проведенное до уровня N, является вектором, который имеет следующую структуру:

$[cA_N, cD_N, ..., cD_1]$

Результаты разложения удобно изображать графически в виде дерева, изображенного на рис. 1

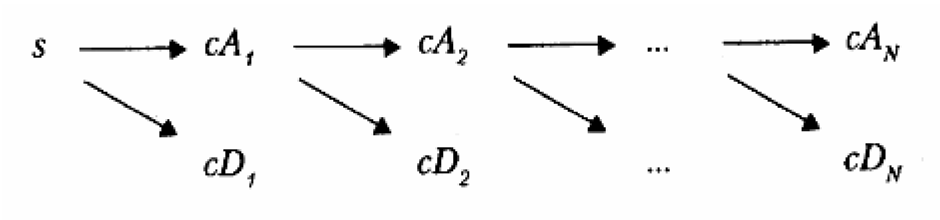

Рис. 1. Дерево разложения.

Полученные при разложении аппроксимирующие коэффициенты представляют сглаженный сигнал, а детализирующие коэффициенты описывают колебания. Следовательно, шумовая компонента больше отражается в детализирующих коэффициентах cD_j. Поэтому при удалении шума обрабатывают обычно детализирующие коэффициенты. Удаление шума реализуется с помощью метода пороговой обработки коэффициентов (трешолдинг) и заключается в обнулении значений коэффициентов, меньших некоторого порогового значения [1].

Для изучения возьмем 15-минутные данные о ценах и объемах продаж акций ОАО «Сургутнефтегаз» на ММВБ за 2006 г. Перед использованием эти массивы данных были предварительно обработаны. Недостающие данные в 18:45 заполнены данными в 10:45 следующего дня, а данные в 10:30 удалены. В этом случае рабочий день ММВБ будет состоять из восьми часов работы.

Полученный сигнал имеет шумовую компоненту, поэтому сначала вычислим порог для удаления шума, удалим этот шум и оценим фрактальность сигнала. Для удаления шума используем вейвлет db4, уровень разложения 3 и минимаксный порог 2.771. Исходный и очищенный сигнал изображены на рис. 2.

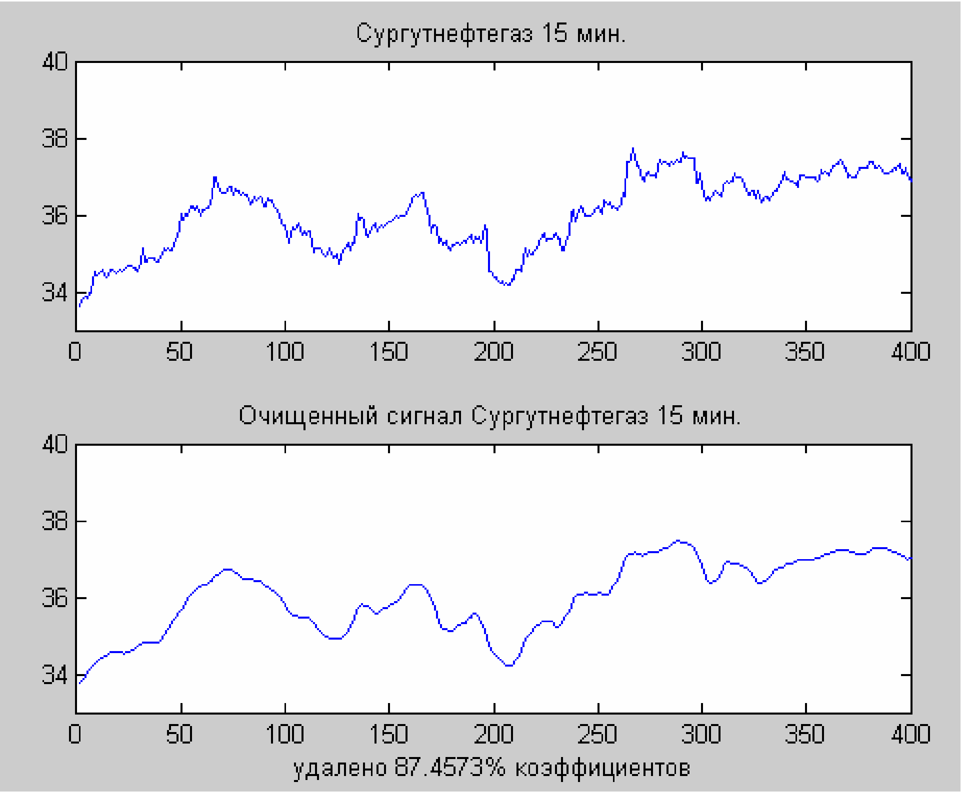

Рис. 2. Графики сигнала и его очищенной от шума версии.

В результирующем сигнале удалено 87.4573% коэффициентов. Это позволяет эффективно сжимать такие сигналы для их хранения. При этом в очищенном сигнале колебания цены видны более явно, хорошо отражена их форма волн, что позволяет четче идентифицировать фигуры технического анализа на графике. Очищенные от шума паттерны фигур удобно использовать для обучения нейронных сетей с целью распознавания подобных паттернов в будущем и соответственно прогнозирования тренда. Можно также получить всю удаленную шумовую компоненту в виде разности исходного сигнала и очищенного. Шумовая компонента может быть использована для получения статистических и стохастических характеристик высокочастотной части сигнала. Например, из шумовой составляющей сигнала можно выразить такую характеристику акции, как пределы колебания ее цены.

Проведем оценку показателя Херста для исходного S и очищенного от шума So сигналов.

HERST_S = 0.4492
HERST_So = 1.5924

Для сигнала S показатель Херста H около 1/2, что означает большую хаотичность движения цены, а для So H > 1, что означает большую регулярность движения.

Отметим, что применительно к финансовым данным показатель Херста H измеряет влияние информации на временной ряд, поэтому:

− значение H = 0,5 подразумевает случайное блуждание, что является подтверждением гипотезы эффективного рынка. В этом случае события некоррелированы, все новости уже учтены и обесценены рынком.

− значение H>0,5 подразумевает, что сегодняшние события будут иметь значение завтра, и полученная информация продолжает учитываться рынком некоторое время спустя. И это не просто последовательная корреляция, это функция долговременной памяти, которая обусловливает информационное влияние в течение больших периодов времени.

− значение H<0,5 сигнализирует антиперсистентный ряд. Такой ряд волатилен, т. е. более изменчив, чем ряд случайный. Он состоит из частых реверсов «спад-подъем» [2].

Кроме подавления шума, вейвлет-анализ позволяет из непомерно большого числа входных переменных извлекать наиболее значимые для предсказания признаки, что является плюсом для предобработки больших массивов значений финансовых временных рядов (тот же дискретный сигнал). Такое представление динамики ряда удобно использовать в качестве обучающей выборки для нейронных сетей. Связано это с тем, что чем дальше в прошлое уходит история ряда, тем меньше деталей его поведения влияет на результат предсказаний. Это обосновано психологией субъективного восприятия прошлого участниками торгов, которые, собственно, и формируют будущее. Поэтому, используя вейвлет-преобразование финансовых временных рядов, можно удалить лишние детали поведения рынка, сохранив общий вид кривой [3]. Выделить признаки временного ряда можно, используя аппроксимирующие коэффициенты cA_N (представляют низкочастотный сглаженный сигнал). Уровень сжатия сигнала зависит от уровня (глубины) N вейвлет-разложения. На рис 3 изображен исходный финансовый временной ряд и его реконструкция по 50 аппроксимирующим коэффициентам. Обратим внимание на то, что для этого потребовалось в четыре раза меньше данных, и временной ряд восстановлен достаточно точно, максимумы и минимумы отражены верно.

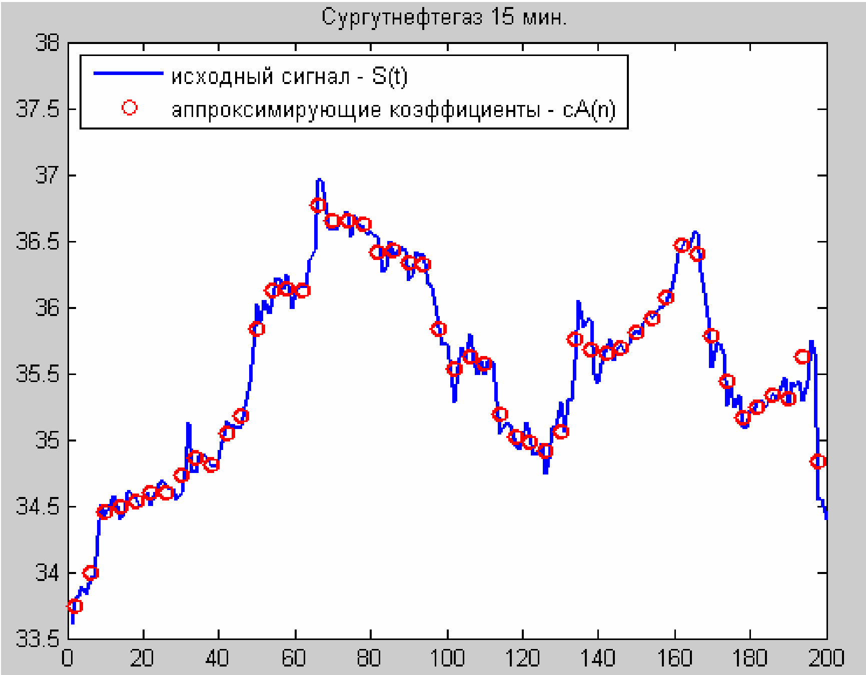

Рис. 3. Пример исходного сигнала и его реконструкции по 50 аппроксимирующим коэффициентам.

Как правило, в качестве финансовых временных рядов выступают массивы данных об объемах или ценах продаж, товарного или фондового рынка. Для их анализа более естественно использовать вейвлеты с параметром масштабирования N, отличным от 2. Дело в том, что такие временные ряды имеют периодичность по времени. Например, при анализе 15-минутных данных о цене акции естественно использовать вейвлеты с параметром масштабирования N = 4 (поскольку в часе таких периодов - 4). Далее, при анализе часовых данных о цене ценной бумаги фондового рынка естественно использовать вейвлеты с параметром масштабирования N = 8 (поскольку день работы рынка содержит 8 таких отсчетов). При анализе дневных данных естественно использовать вейвлеты с параметром масштабирования N = 5 (поскольку неделя работы рынка содержит 5 рабочих дней). При анализе дневных показателей продаж товара естественно использовать вейвлеты с параметром масштабирования N = 7 (недельный цикл), с отдельной выборкой данных за выходные дни.

В представленной работе был проведен анализ финансовых временных рядов на основе вейвлет-преобразования. В частности рассмотрена возможность подавления шумовой составляющей и флуктуаций в рядах, извлечение наиболее значимых для предсказания признаков. Также проанализированы их фрактальные свойства. На основе данной работы можно сделать вывод об эффективности применения вейвлет-анализа в качестве математического аппарата для обработки финансовых данных. Вейвлет-анализ позволяет подготовить информацию к прогнозированию на базе нейронных сетей.

Список использованной литературы

1. Смоленцев Н.К. Основы теории вейвлетов. Вейвлеты в Matlab. // М.: ДМК Пресс, 2008. – 448с.: ил.

2. Ширяев В. И. Финансовые рынки и нейронные сети // Учебное пособие. – М.: Издательство ЛКИ, 2007. – 224 с.

3. Нейрокомпьютинг и его применения в экономике и бизнесе [Электронный ресурс] – Режим доступа: http://www.intuit.ru/department/expert/neurocomputing/8/1.html – Загл. с экрана.