

Разметка финансовых временных рядов

Волобой В.С.

*Математико-механический факультет, Санкт-Петербургский
государственный университет*

Цель: данная работа посвящена изучению возможного варианта разметки финансового временного ряда.

В современном мире задача прогнозирования является весьма популярной и востребованной. Для того, чтобы провести прогноз, данные сначала необходимо подготовить. В данной работе предлагается новый вариант разметки финансовых временных рядов, который в последующем можно использовать как этап подготовки данных к прогнозированию

- **Определение 1:** Временной ряд f на интервале $[a, b]$ с разбиением N , длины T - это множество пар (t_i, x_i) , где t_i - момент времени из промежутка $[a, b]$, $T = b - a$, $t_i = a + i * \frac{T}{N}$ ($i \in \{0..N\}$), x_i - значение временного ряда в момент времени t_i
- **Определение 2:** Временной ряд f на интервале $[a, b]$ с разбиением N будем называть размеченным, если каждому его элементу (t_i, x_i) сопоставляется символ z из алфавита $\{d, s, u\}$

Согласно определению 2, необходимо исследовать принцип разметки временного ряда. Суть этого принципа заключается в том, насколько сильно изменяется угол наклона между двумя соседними элементами временного ряда.

Исходя из вышесказанного, введены следующие правила:

- Элементу (t_i, x_i) временного ряда сопоставим “ s ”, если $\leq \text{tg}(5)$
 - Элементу (t_i, x_i) временного ряда сопоставим “ u ”, если $> \text{tg}(5)$ и $x_{i+1} > x_i$
 - Элементу (t_i, x_i) временного ряда сопоставим “ d ”, если $> \text{tg}(5)$ и $x_{i+1} < x_i$
- **Определение 3:** Тренд — это основная тенденция развития процесса во времени

Необходимо придумать алгоритм, выделяющий начала и концы трендов во временном ряде и определяющий характер этих трендов согласно следующей классификации:

- Восходящий тренд – временной ряд в промежутке $[a,b]$, где всем элементам ряда внутри этого промежутка сопоставлен символ “u”
- Нисходящий тренд – временной ряд в промежутке $[a,b]$, где всем элементам ряда внутри этого промежутка сопоставлен символ “d”
- Боковой тренд (флет) – временной ряд в промежутке $[a,b]$, где всем элементам ряда внутри этого промежутка сопоставлен символ “s”

После того, как во временном ряде будут выделены начала и концы трендов, к этим точкам будет применён принцип разметки временного ряда. Всем точкам внутри одного тренда сопоставится тот же символ из алфавита $\{d,s,u\}$, что и началу тренда, к которому принадлежат эти точки, вне зависимости от принципа разметки.

Выделение тренда внутри временного ряда, было сделано с помощью экспоненциальных скользящих средних. Для начала необходимо познакомиться с некоторыми определениями:

- **Определение 4:** Скользящая средняя - функция, значения которой в каждой точке определения равно среднему значению исходной функции за предыдущий период
- **Определение 5:** Взвешенная скользящая средняя – скользящее среднее, у которой каждый член исходной функции имеет свой вес : последние значения исходной функции более значимы чем предыдущие
- **Определение 6:** Экспоненциальная скользящая средняя - взвешенная скользящая средняя, веса которой убывают экспоненциально

В данной работе для выделения трендов во временном ряде была использована экспоненциальная скользящая средняя, которая вычисляется по формуле:

- $EMA(i) = EMA(i-1) + k * (Price(i) - EMA(i-1))$, где i – текущее время, $Price(i)$ – значение ряда в момент времени i , $k = 2/(n+1)$, где n – период средней.

Суть работы скользящей средней заключается в сглаживании колебаний цен временного ряда, путём его усреднения, на рассматриваемом историческом интервале.

Однако экспоненциальная скользящая средняя не следует за графиком моментально, она несколько запаздывает за его тенденцией, но этот факт можно обратить в плюс: так как известно значение цен ряда на всём временном промежутке, то можно рассматривать экспоненциальную скользящую среднюю во временном интервале не только слева направо, но и справа налево.

Для рассмотрения выбран индекс RTS:

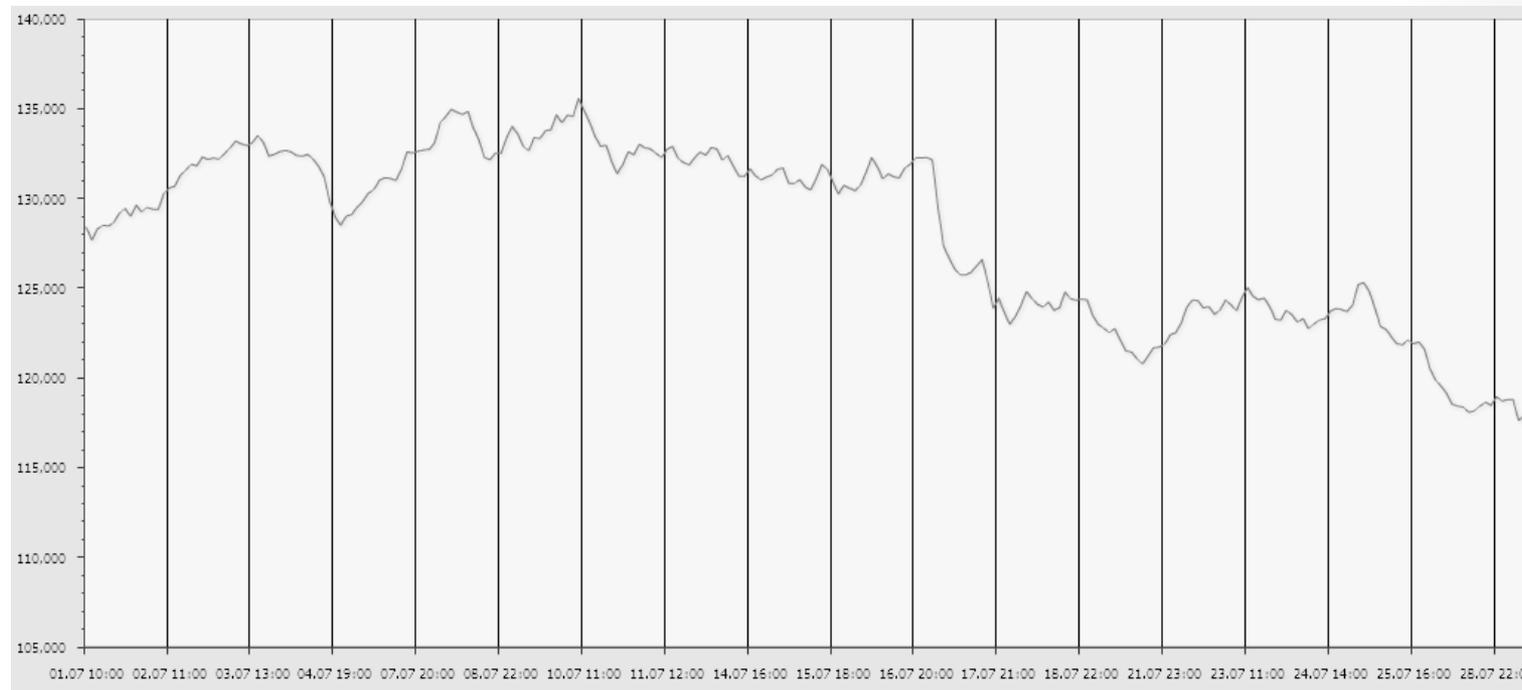


Рис.1 График индекса RTS за период с 1.07.2014 по 28.07.2014

Для выделения трендов во временном ряде рассмотрим следующий алгоритм:

- Построение экспоненциальной скользящей средней временного ряда “слева направо”
- Построение экспоненциальной скользящей средней временного ряда “справа налево”
- Абсциссы точек пересечения двух скользящих средних берутся за начала и концы трендов
- Определяется характер трендов согласно вышеприведённой классификации

Результаты работы алгоритма:

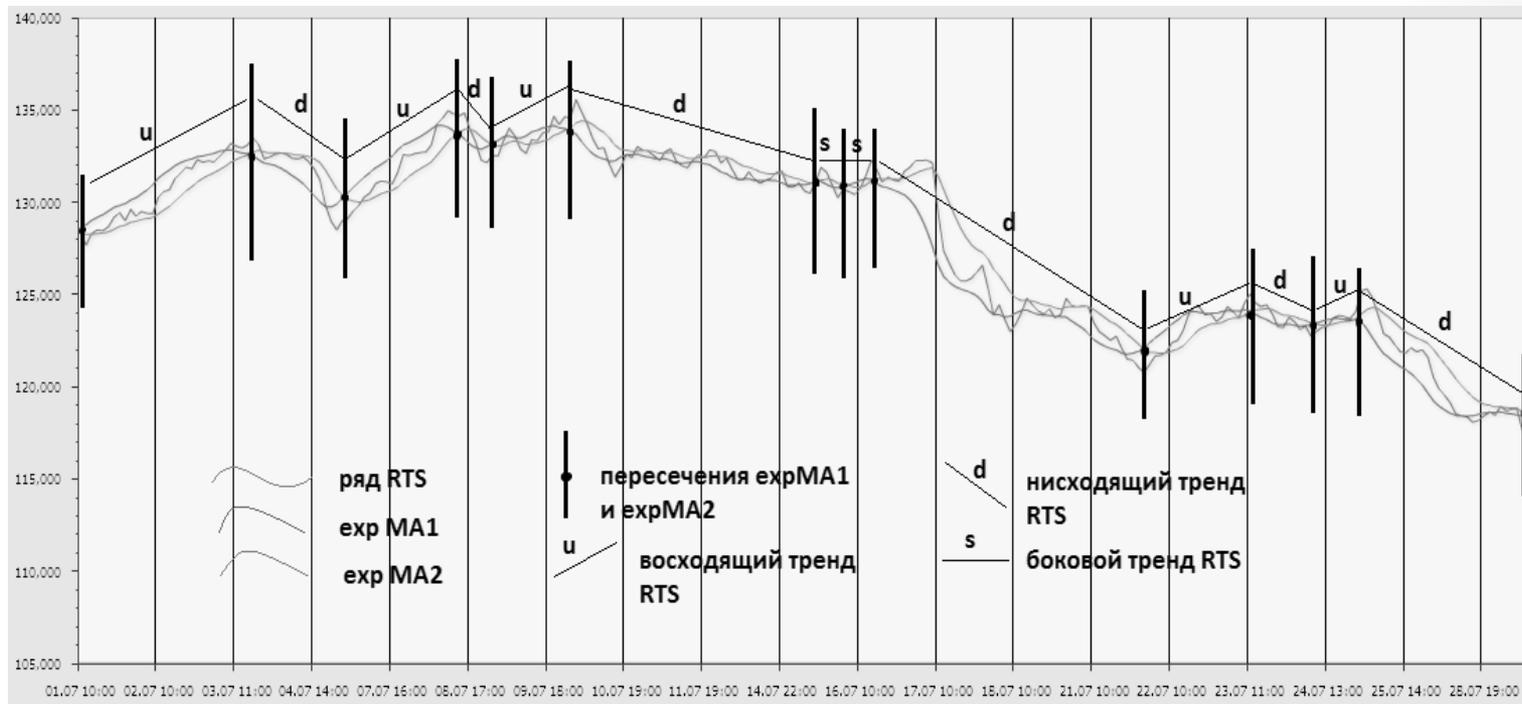


Рис.2 График индекса RTS вместе с двумя экспоненциальными средними. Выделены точки пересечения экспоненциальных средних, а так же показаны тренды индекса RTS.

Заключение: В данной работе предложен новый метод разметки финансовых рядов, который может быть в дальнейшем использован для прогнозирования.