Экспоненциальная скользящая средняя с переменным фактором сглаживания.

Станислав Булашев

Динамика движения цен активов характеризуется тем, что в ней в каждый момент времени присутствуют более или менее ярко выраженные трендовая и шумовая составляющие. Целью сглаживания ценовых рядов является фильтрация шумовых колебаний с целью выявления устойчивых тенденций движения. В идеальной скользящей средней (MA) максимально возможное сглаживание шума должно сочетаться с минимальным искажением трендов. Однако очевидно, что эти два условия противоречат друг другу. К тому же, дело осложняется тем, что характеристики и трендов и шума нерегулярны во времени. Поэтому при сглаживании в каждый момент времени необходимо поддерживать разумный баланс между мерой гладкости MA и мерой близости этой MA к исходному ценовому ряду, причем этот баланс должен соответствовать текущей ценовой динамике.

Одним из наиболее популярных методов сглаживания является экспоненциальная скользящая средняя (EMA). Напомним методику ее вычисления. Введем обозначения:

- $\{x_t\}$ - ряд цены (или логарифма цены) актива,
- $\{y_t\}$ - скользящая средняя ряда $\{x_t\}$.

В качестве меры близости рядов $\{x_t\}$ и $\{y_t\}$ в момент времени t выберем функцию $(x_t - y_t)^2$, а в качестве меры гладкости ряда $\{y_t\}$ функцию $(y_t - y_{t-1})^2$. Если взять взвешенную сумму этих функций, то есть

$$\Delta S_t = w \cdot (x_t - y_t)^2 + (1 - w) \cdot (y_t - y_{t-1})^2$$

то для расчета текущей величины y_t, которая минимизирует функцию ΔS_t, надо решить следующее уравнение:

$$\frac{\partial \Delta S_t}{\partial y_t} = 0$$

Решая его, получим, что $y_t = y_{t-1} + w \cdot (x_t - y_{t-1})$. Это и есть формула для классической EMA с постоянным фактором сглаживания w. Обычно в торговых алгоритмах на основе EMA, сигналами к покупкам/продажам являются моменты изменения знака ее первой производной, или в барной нотации - первой разности, то есть величины $\Delta y_t = y_t - y_{t-1}$.

Очевидным недостатком EMA является именно постоянство фактора сглаживания. Формула для ее расчета никак не учитывает, что, как говорилось выше, характеристики и трендов и шума нерегулярны во времени. Поэтому при сильном сглаживании (маленьком w), EMA дает мало ложных сигналов к покупкам/продажам на флэте, но зато сигналы о смене тренда сильно запаздывают во времени. Напротив, при слабом сглаживании (большом w), EMA с небольшим опозданием реагирует на смену тренда, но генерирует массу ложных сигналов на флэте.

*   *   *

В этой статье рассматривается один из возможных методов расчета EMA *с переменным во времени фактором сглаживания*, который адаптируется в текущей динамике цены. Адаптация состоит в том, что чем сильнее в движении цены выражена трендовая составляющая, тем большие значения должен принимать фактор сглаживания, что уменьшает запаздывание скользящей средней относительно цены. Напротив, на флэтах фактор сглаживания должен уменьшаться, при этом скользящая средняя вырождается в горизонтальную (или почти горизонтальную) линию.

До того, как рассказать о разработанном мной алгоритме, вкратце остановлюсь на аналогичных работах. Если не вдаваться в подробности, то все авторы, которых мне удалось прочесть на эту тему, эксплуатируют одну и ту же идею, которая состоит в следующем. Для вычисления на текущем баре значения переменного во времени фактора сглаживания, рассчитывается индикатор тренда. У разных авторов он может называться Trend Indicator, Chande Momentum Oscillator, Efficiency Ratio и т.д. Но формула для его вычисления всегда одна и та же:

$$TI_t = \frac{\left|\sum_{k=t-T+1}^{t}(x_k - x_{k-1})\right|}{\sum_{k=t-T+1}^{t}\left|x_k - x_{k-1}\right|}$$

То есть индикатор тренда на текущем баре t представляет собой отношение модуля суммы изменений цен на последних T барах, к сумме модулей изменений цен на последних T барах. Причем количество баров, на котором вычисляется индикатор тренда, постоянно. Очевидно, что $0 \le TI_t \le 1$, и тем больше, чем сильнее на периоде T в движении цены выражена трендовая составляющая. Далее вычисляется переменный фактор сглаживания по формуле $w_t = w_{\max} \cdot TI_t$. Может быть использована и более сложная монотонно возрастающая функция от TI_t. И наконец определяется собственно значение скользящей средней $y_t = y_{t-1} + w_t \cdot (x_t - y_{t-1})$.

Описанный алгоритм дает преимущество по сравнению с классической EMA (с постоянным w), потому что он решает поставленную выше задачу, а именно увеличивает переменный фактор w_t в случае наличия направленного движения на последних T барах, и уменьшает его на флэте. Однако на мой взгляд в методике его расчета содержится принципиально неустранимое противоречие, которое состоит в том, что *переменный* фактор w_t вычисляется через величину (индикатор тренда), рассчитываемую с *постоянным* периодом T.

Еще одним, но уже менее принципиальным, недостатком является то, что в прочитанных мной работах зависимость фактора сглаживания от индикатора тренда была жестко заданной, то есть отсутствовала возможность ее варьирования без переписывания соответствующих индикаторов в программе теханализа.

*   *   *

Итак, нужно было решить следующую задачу - найти формулу для расчета переменного во времени фактора сглаживания w_t, который не должен являться функцией от каких либо индикаторов, вычисляемых с постоянным периодом суммирования или усреднения.

Вспомним то, о чем говорилось выше - при сглаживании в каждый момент времени необходимо поддерживать разумный баланс между мерой гладкости МА и мерой близости этой МА к исходному ценовому ряду, причем этот баланс должен соответствовать текущей ценовой динамике.

Остановимся подробнее на мере близости $(x_t - y_t)^2$, которая является квадратом ошибки аппроксимации скользящей средней исходного ценового ряда $e_t = x_t - y_t$. Для ЕМА с постоянным фактором сглаживания наблюдаются следующие эффекты:

- на флэте ошибка аппроксимации изменяется как правило в достаточно узком диапазоне, при этом y_t колеблется с той же частотой, что и x_t, но запаздывая по времени и с меньшей амплитудой.
- на выходе из флэта в тренд, ошибка аппроксимации резко возрастает по абсолютной величине, цена x_t как бы убегает от своей средней y_t.

Итак для того, чтобы попытаться устранить одновременно обе присущих классической ЕМА особенности (то есть колебания y_t на флэте и существенное отставание y_t от x_t на тренде), попробуем для расчета переменного во времени фактора сглаживания w_t использовать ошибку аппроксимации. Однако, в качестве этой ошибки величину $e_t = x_t - y_t$ использовать нельзя. Дело в том, что до того, как вычислить $y_t = y_{t-1} + w_t \cdot (x_t - y_{t-1})$, уже должен быть известен w_t. Поэтому для целей расчета w_t будем использовать другую величину $\delta_t = x_t - y_{t-1}$.

Сформулируем в общем виде требования, которым должен удовлетворять переменный во времени фактор сглаживания w_t:

- w_t может меняться в пределах от $w_{\min}$ до $w_{\max}$,
- w_t должен быть монотонно возрастающей функцией от $|\delta_t|$, причем форма этой зависимости должна быть варьируемой.

Так как исходный ряд может иметь размерность, введем в рассмотрение нормирующий коэффициент E, имеющий такую же размерность, как у величин x_t и y_t. Назовем этот коэффициент квантом ошибки аппроксимации. Разделив $|\delta_t|$ на E, получим безразмерную переменную, показывающую количество квантов, на которое в текущий момент y_t отклоняется от x_t:

$$z_t = \frac{|\delta_t|}{E}$$

Итак, можно записать, что

$w_t = w_{\min} + (w_{\max} - w_{\min}) \cdot q(z_t)$, где $q(z_t)$ представляет собой монотонно возрастающую функцию, лежащую в пределах от 0 до 1, при том что z_t меняется вообще говоря в пределах от 0 до ∞. Выше говорилось, что вид функции w_t, а значит и $q(z_t)$, должен быть варьируемым. Следовательно, она должна содержать некоторое количество параметров. На мой взгляд, наиболее приемлемо использовать в этом случае экспоненциальную зависимость $q(z_t)$, содержащую единственный параметр, который однако позволяет изменять форму $q(z_t)$ в очень широких пределах:

$q(z_t) = 1 - \exp\left(-z_t^a\right)$, где положительный параметр a отвечает за форму зависимости.

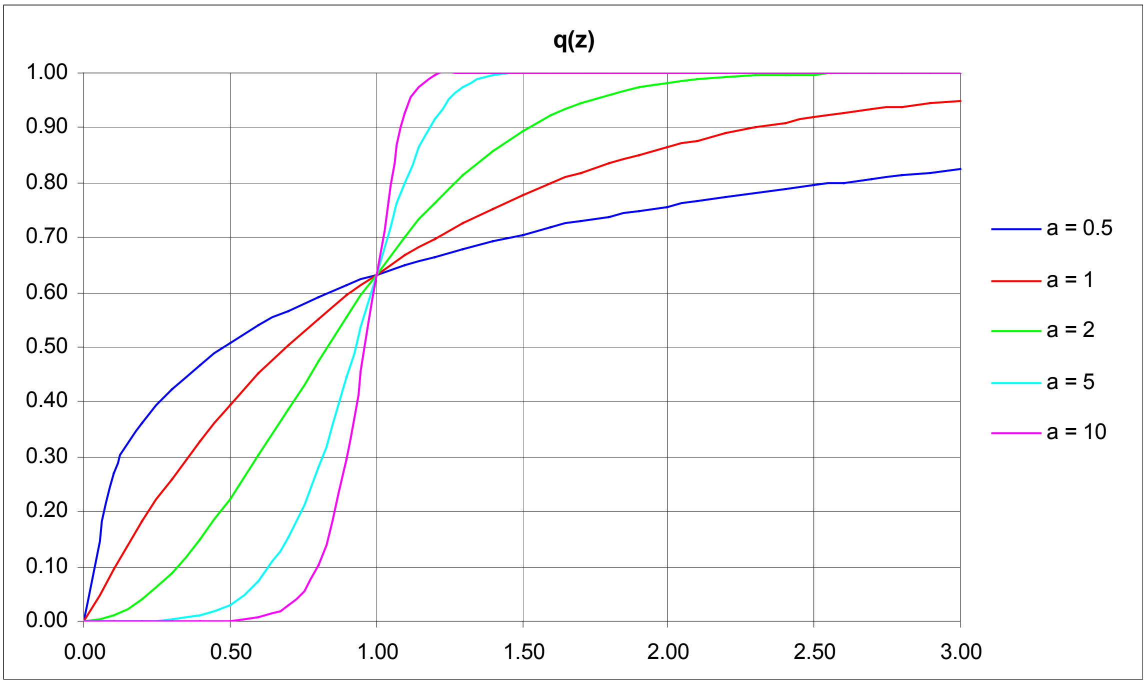

Рис.1. Зависимость $q(z)$ при различных значениях параметра формы a

Резюмируем сказанное. Мы определили алгоритм сглаживания временных рядов, который является экспоненциальной скользящей средней с переменным фактором сглаживания в виде:

$$y_t = y_{t-1} + w_t \cdot (x_t - y_{t-1})$$

где

$$w_t = w_{\min} + (w_{\max} - w_{\min}) \cdot \left(1 - \exp\left(-\left|\frac{x_t - y_{t-1}}{E}\right|^a\right)\right)$$

Назовем этот алгоритм EMAVFS (Exponential Moving Average with Variable Factor of Smoothing). Формально он имеет 4 параметра:

- $w_{\min}$ и $w_{\max}$ ограничивают пределы изменения w_t,
- a отвечает за форму зависимости w_t от ошибки аппроксимации $\delta_t = x_t - y_{t-1}$,
- E определяет чувствительность w_t к изменению ошибки $\delta_t = x_t - y_{t-1}$.

Несмотря на то, что формально у алгоритма достаточно много параметров - 4, некоторые из них не обязательно использовать как оптимизируемые переменные. Например $w_{\min}$ логично априорно во всех случаях принять равным 0. Параметр формы a после некоторого исследования также можно зафиксировать и не менять в дальнейшем. То есть, на мой взгляд, лишь верхний предел $w_{\max}$ и параметр чувствительности E обязательно должны являться оптимизируемыми переменными.

На рисунке 2 приведен пример построения EMAVFS.

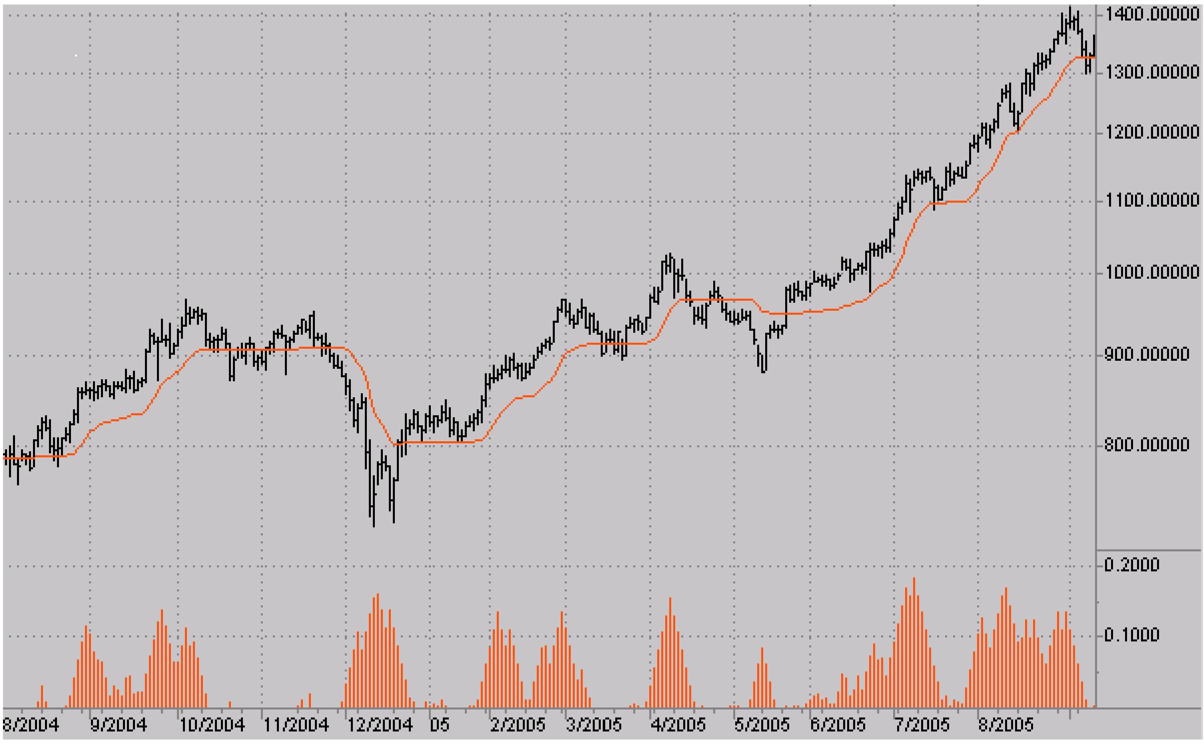

Рис.2. Вверху – EMAVFS от цены закрытия по акциям Лукойла (дневные бары).
Внизу – переменный фактор сглаживания w_t для EMAVFS.

В заключение хотелось бы подчеркнуть одно важное обстоятельство - в отличие от классической EMA, EMAVFS является *нелинейным* методом сглаживания, то есть если $x_t^{(1)}, x_t^{(2)}, \ldots$ - набор временных рядов, $c^{(1)}, c^{(2)}, \ldots$ - набор констант, то

$$EMAVFS\left(\sum_k c^{(k)} \cdot x_t^{(k)}\right) \neq \sum_k c^{(k)} \cdot EMAVFS\left(x_t^{(k)}\right)$$