

УДК 519.2

Вероятностно-статистический анализ сигналов классических алгоритмов скользящего среднего и нового оригинального алгоритма синтетического скользящего среднего

Тихонова О.А.

Донецкий национальный технический университет
oks_mail777@mail.ru

Abstract

Tikhonova O. A. "The likelihood-statistical analysis of signals of classical algorithms of sliding average and new original algorithm of a synthetic sliding average". The comparative analysis of likelihood sizes of generation of "useful" and "harmful" signals of classical algorithms of sliding average and new original algorithm of a synthetic sliding average is carried out.

Введение

Характерной особенностью временных дискретных рядов является то, что наблюдения за некоторым объектом производятся последовательно во времени. Например: температура воздуха в середине каждого часа суток, ежегодная урожайность зерновых, ежедневный объем продаж какого-нибудь товара, систематических контроль за качеством воды водоисточника и т.д.

При анализе временного ряда наблюдаемые значения интерпретируются как сумма систематической (тренд) и случайной циклической составляющих. Исключение случайных колебаний значений уровней ряда осуществляется с помощью нахождения «усредненных» значений. Метод скользящей средней применяется для характеристики тенденции развития исследуемой статистической совокупности и основан на расчете средних уровней ряда за определенный период.

С помощью графического метода исследователь может выделить следующие моменты: возрастающую тенденцию, убывающую тенденцию и неупорядоченное, хаотичное движение (флэт).

Качественной оценкой работы метода скользящей средней может быть показатель суммарной величины вероятности «полезных» сигналов на тенденции.

Целью данной работы является сравнительный анализ вероятностных величин генерации «полезных» и «вредных» сигналов классических алгоритмов скользящего среднего и нового оригинального алгоритма синтетического скользящего среднего [1] и индикаторов построенных на их основе.

В идеальном случае, применяя индикатор, пользователь должен был бы с его помощью получать надежный сигнал о наличие тенденции и пропускать участки хаотического движения.

Алгоритм синтетического скользящего среднего

Алгоритм синтетического скользящего среднего (ССС) был разработан с целью устране-

ния недостатков алгоритмов скользящего усреднения. Полный материал и анализ данного алгоритма были опубликованы в работе [1]. Остановимся на основных ее положениях.

Алгоритм ССС сводится к многократному усреднению исследуемого ряда «назад – вперед». В таблице 1 приведен алгоритм реализации ССС для величины временного окна усреднения $m=4$.

Таблица 1. Алгоритм формирования синтетической скользящей средней.

C_i Q_i	C_1	C_2	C_3	C_4
I	$Q_4=Q_3+\alpha(C_1 - Q_3)$	$Q_3=Q_2+\alpha(C_2 - Q_2)$	$Q_2=Q_1+\alpha(C_3 - Q_1)$	$Q_1=C_4$
II	$Q_5=Q_4$	$Q_6=Q_5+\alpha(Q_3 - Q_5)$	$Q_7=Q_6+\alpha(Q_2 - Q_6)$	$Q_8=Q_7+\alpha(Q_1 - Q_7)$

В таблице 1 C_i это уровни исследуемого временного ряда. Значения Q_i – отчеты выделяемого тренда с помощью экспоненциального скользящего среднего (ЭСС). При усреднении «назад» накапливается временная задержка усреднения (групповая задержка), которая компенсируется при обратном проходе «вперед». Комбинация проходов «назад – вперед» способствует получению значений ССС в нужное время, соответствующего моменту существования C_4 (это позволяет реализовать оперативное управление). Величина α (коэффициент экспоненциального усреднения) определяется как

$$\alpha = \frac{2}{\left(\frac{m}{2}\right) + 1}, \quad (1)$$

где: m – величина временного окна усреднения.

Отличие (1) от традиционного выражения заключается в том, что используется не m , а $m/2$. Поскольку при двойном проходе (усреднении уровней временного ряда C_i) эквивалентная величина m удваивается. Далее как и при реализации

любого алгоритма скользящего усреднения уровень C1 отбрасывается и добавляется новый C5. Для реализации, например, m=8 количество проходов «назад – вперед» равно соответственно четырем («назад – вперед – назад – вперед»). Главный недостаток ССС – значение m должно быть кратно четырем.

Построение вероятностной модели процесса

Рассматривая динамику изменения дискретного ряда, в обобщенном виде, можно классифицировать два ее состояния: наличие тенденции и хаос. Рассматривая сигнал, полученный с помощью индикатора можно классифицировать так же два его состояния: данный сигнал верен или данный сигнал не верен.

Обозначим вероятность события «сигнал верен» как $P_{\text{прав}}$, а вероятность события «сигнал не верен» как $P_{\text{ошиб}}$.

Событие «сигнал верен» является глобальным и объединяющим событием ряда более мелких событий образующих полную группу, как и событие «сигнал не верен» является глобальным и объединяющим событием ряда противоположных более мелких событий. Следовательно, вероятность $P_{\text{прав}}$ состоит из суммы вероятностей элементарных событий образующих полную группу события «сигнал верен».

Данное верно и для $P_{\text{ошиб}}$. В таблице 2 приведен список вероятностей, входящих в вероятностную модель события «генерация сигнала», и образующих вероятностное пространство.

Таблица 2.

Вероятности образующие $P_{\text{прав}}$	Вероятности образующие $P_{\text{ошиб}}$
Вероятность сигнала, указывающего на верную тенденцию P_s	Вероятность пропуска сигнала, указывающего на верную тенденцию P_{no_s}
Вероятность пропуска сигнала, указывающего на не верную тенденцию P_{no_b}	Вероятность сигнала, указывающего на не верную тенденцию P_b
Вероятность пропуска сигнала, попадания на флэт P_{no_f}	Вероятность сигнала, попадания на флэт P_f

Следовательно:

$$P_{\text{прав}} = P_s + P_{\text{no}_b} + P_{\text{no}_f} \quad (2)$$

$$P_{\text{ошиб}} = P_{\text{no}_s} + P_b + P_f$$

По теореме вероятностей сумма вероятностей *противоположных* событий равна 1:

$$P_{\text{прав}} + P_{\text{ошиб}} = 1 \quad (3)$$

В случае применения идеального индикатора $P_{\text{прав}} \rightarrow 1$, а $P_{\text{ошиб}} \rightarrow 0$. На практике, придерживаются следующего правила $P_{\text{прав}} \rightarrow \max$, а $P_{\text{ошиб}} \rightarrow \min$.

Расчет вероятностей:

$$P_s = N_s / N_{\text{общ}}$$

$$P_{\text{no}_b} = N_{\text{no}_b} / N_{\text{общ}}$$

$$P_{\text{no}_f} = N_{\text{no}_f} / N_{\text{общ}}$$

$$P_{\text{no}_s} = N_{\text{no}_s} / N_{\text{общ}}$$

$$P_b = N_b / N_{\text{общ}}$$

где

$N_{\text{общ}}$ – общее количество трендовых и не трендовых участков дискретного ряда;

$$N_{\text{общ}} = N_{\text{всх}} + N_{\text{нсх}} + N_{\text{флэт}}$$

$N_{\text{всх}}$ – общее количество возрастающих трендов;

$N_{\text{нсх}}$ – общее количество убывающих трендов;

$N_{\text{флэт}}$ – общее количество флэтов;

N_s – количество сигналов указывающих на верную тенденцию;

N_{no_s} – количество пропущенных сигналов указывающих на верную тенденцию;

N_{no_b} – количество пропущенных сигналов указывающих на не верную тенденцию;

N_b – количество сигналов указывающих на не верную тенденцию;

N_f – количество не пропущенных участков хаоса;

N_{no_f} – количество пропущенных участков хаоса;

Кроме вероятностных характеристик, в данной работе, выполнен расчет дисперсии, как показателя рассеивания [8] [9] данных дискретного ряда от результатов усреднения. Глубина возможного прогноза зависят от дисперсии отчетов выделяемых трендов относительно их истинных значений. Чем менее «колеблемый» выделенный тренд, тем точнее и дальше по времени в будущем осуществляется статистический прогноз. Оптимальным алгоритмом скользящих средних будет алгоритм с минимальной дисперсией.

Экспериментальная часть.

В качестве дискретного ряда в данной работе были взяты суточные цены закрытия валютной пары Евро-Доллар за 2,5 года (2003г. – 2005г.). Построение графика, применение индикаторов и анализ полученных сигналов были проведены в трейдерской программе TradeStation 2000i. Для создания программ новых индикаторов, базирующихся на алгоритме синтетического скользящего среднего, был использован внутренний редактор языка EasyLanguage PowerEditor программы TradeStation 2000i.

Исследования проводились при использовании следующих индикаторов: простое скользящее среднее (МА) 1 линия, экспоненциальное скользящее среднее (ЕМА) 1 линия, синтетическое скользящее среднее (SMA) 1 линия, МА 2 линии, ЕМА 2 линии, SMA 2 линии, MACD (moving average convergence divergence метод – схождение - расхождение скользящих средних) классический, MACD синтетический,

синтетический, CCI (Commodity Channel Index, Индекс товарного канала) классический, CCI синтетический.

Генерация сигналов с использованием скользящих средних 1 линии происходит по следующему, общему для них, правилу:

$$\left\{ \begin{array}{l} \text{раст_тренд, если } MovAver(m)_i = X_i, \\ MovAver(m)_{i+1} < X_{i+1} \\ \text{убыв_тренд, если } MovAver(m)_i = X_i, \\ MovAver(m)_{i+1} > X_{i+1} \end{array} \right. \quad (4)$$

где m – величина окна скользящего;

MovAver – один из предложенных алгоритмов скользящего среднего;

X – цифровой дискретный ряд;

На рисунке 1 представлен фрагмент дискретного ряда с графиком индикатора MA 1 линия и сгенерированные сигналы BUY (возрастающий тренд). Начало каждого сигнала на графике указано темной стрелкой с подписью R1.

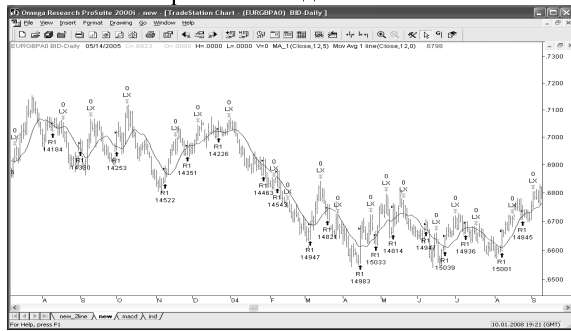


Рисунок 1 Индикатора MA 1 линия и сигналы возрастающего тренда.

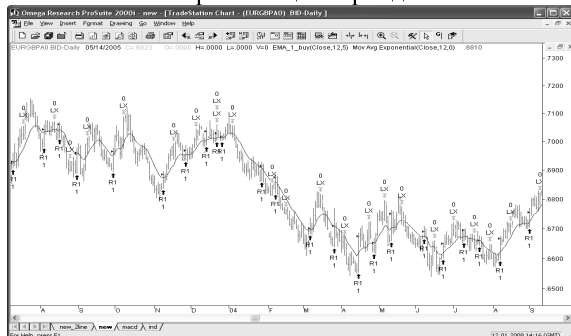


Рисунок 2 Индикатора EMA 1 линия и сигналы возрастающего тренда.

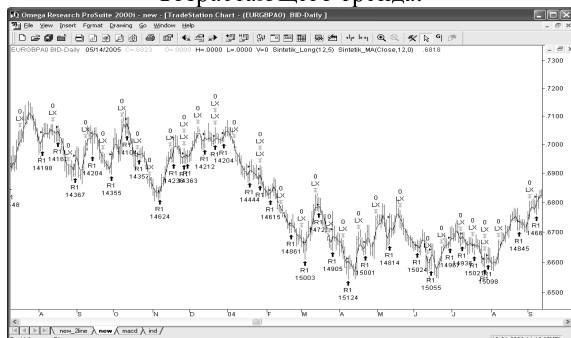


Рисунок 3 Индикатора SMA 1 линия и сигналы возрастающего тренда.

$$\left\{ \begin{array}{l} \text{раст_тренд,} \\ \text{если } MovAver(m1)_i = MovAver(m2)_i, \\ MovAver(m1)_{i+1} > MovAver(m2)_{i+1}, MovAver(m1)_i < X_i \\ \text{убыв_тренд,} \\ \text{если } MovAver(m1)_i = MovAver(m2)_i, \\ MovAver(m1)_{i+1} < MovAver(m2)_{i+1}, MovAver(m1)_i > X_i \end{array} \right. \quad (5)$$

где $m1, m2$ – величины окон скользящего быстрого и медленного рядов скользящего среднего, $m1 < m2$;

MovAver – один из предложенных алгоритмов скользящего среднего;

X – цифровой дискретный ряд;

На рисунках 4 и 5 отражены соответственно фрагменты графиков с использованием индикаторов MACD классический и MACD основанный на алгоритме синтетического скользящего среднего.

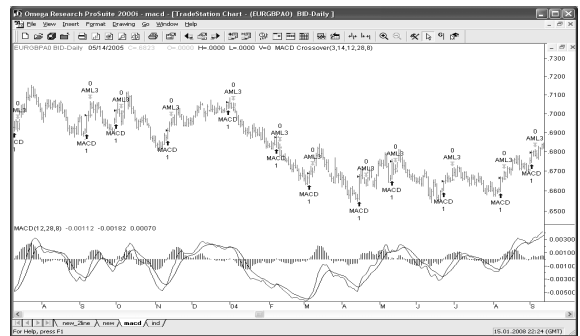


Рисунок 4 Индикатора MACD классический и сигналы возрастающего тренда.

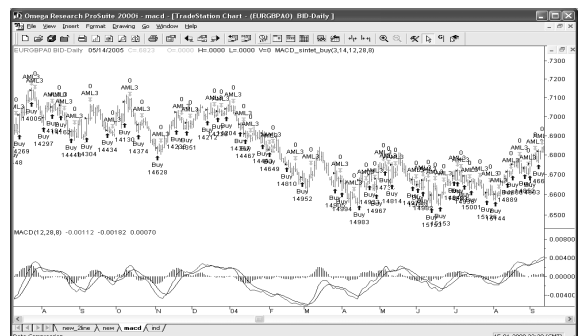


Рисунок 5 Индикатора MACD на базе алгоритма SMA и сигналы возрастающего тренда.

Правила генерации сигналов для классического индикатора MACD:

$$\left\{ \begin{array}{l} \text{раст_тренд,} \\ \text{если } EMA(m)_i = MACD_i > 0, \\ MACD_{i+1} < EMA(m)_{i+1} \\ \text{убыв_тренд,} \\ \text{если } EMA(m)_i = MACD_i < 0, \\ MACD_{i+1} > EMA(m)_{i+1} \end{array} \right. \quad (6)$$

Правила генерации сигналов для синтетического индикатора MACD:

Продолжение таблицы 3.

Вероятность	Составляющие вероятности	МА		ЕМА		SMA	
		2 линия		2 линия		2 линия	
		Восх. тренд	Убыв. тренд	Восх. тренд	Убыв. тренд	Восх. тренд	Убыв. тренд
P _{прав}	P _s	0,17	0,192	0,192	0,234	0,404	0,34
	P _{no_b}	0,298	0,383	0,362	0,404	0,234	0,34
	P _{no_f}	0,149	0,149	0,192	0,192	0,021	0,021
P _{ошиб}	P _{no_s}	0,255	0,192	0,234	0,149	0,021	0,043
	P _b	0,085	0,043	0,021	0,021	0,149	0,085
	P _f	0,043	0,043	0,000	0,000	0,17	0,17
Дисперсия		53,5	41,44	41,44	67,64	3,42	7,75

Продолжение таблицы 3.

Вероятность	Составляющие вероятности	MACD		MACD_S	
		Восх. тренд	Убыв. тренд	Восх. тренд	Убыв. тренд
		P _{прав}	P _s	0,362	0,255
P _{no_b}	0,319		0,255	0,213	0,192
P _{no_f}	0,149		0,149	0,021	0,000
P _{ошиб}	P _{no_s}	0,064	0,128	0,021	0,043
	P _b	0,064	0,17	0,17	0,234
	P _f	0,043	0,043	0,17	0,192
Дисперсия		12,88	40,92	2,84	5

Продолжение таблицы 3.

Вероятность	Составляющие вероятности	CCI		CCI_S	
		Восх. тренд	Убыв. тренд	Восх. тренд	Убыв. тренд
		P _{прав}	P _s	0,234	0,17
P _{no_b}	0,340		0,298	0,298	0,277
P _{no_f}	0,128		0,128	0,064	0,106
P _{ошиб}	P _{no_s}	0,192	0,213	0,149	0,085
	P _b	0,043	0,128	0,082	0,149
	P _f	0,064	0,064	0,128	0,085
Дисперсия		44,91	61,88	23,23	21,07

Таблица 4. Результаты расчетов неравенства теоремы Бернулли

	SMA 1 линия	ЕМА 1 линия	МА 1 линия
Восх. тренд			
m=	17	14	15
k=	47	47	47
p=	0,426	0,426	0,426
m/k=	0,3617	0,2979	0,3191
eps=	0,0643	0,1281	0,1069
прав. часть=	1,258	0,317	0,456
	ВЕРНО	ВЕРНО	ВЕРНО
Нисх. тренд			
m=	16	16	16
k=	47	47	47
p=	0,383	0,383	0,383
m/k=	0,3404	0,3404	0,3404
eps=	0,0426	0,0426	0,0426
прав. часть=	2,774	2,774	2,774
	ВЕРНО	ВЕРНО	ВЕРНО

Продолжение таблицы 4

	SMA 2 линии	ЕМА 2 линии	МА 2 линии
Восх. тренд			
m=	19	9	8
k=	47	47	47
p=	0,426	0,426	0,426
m/k=	0,4043	0,1915	0,1702
eps=	0,0217	0,2345	0,2558
прав. часть=	11,003	0,095	0,080
	ВЕРНО	НЕ верно	НЕ верно
Нисх. тренд			
m=	16	11	9
k=	47	47	47
p=	0,383	0,383	0,383
m/k=	0,3404	0,2340	0,1915
eps=	0,0426	0,1490	0,1915
прав. часть=	2,774	0,227	0,137
	ВЕРНО	ВЕРНО	НЕ верно

Продолжение таблицы 4

	MACD_S	MACD	CCI_S	CCI
Восх. тренд				
m=	19	17	13	9
k=	47	47	47	47
p=	0,426	0,426	0,426	0,426
m/k=	0,4043	0,3617	0,2766	0,1915
eps=	0,0217	0,0643	0,1494	0,2345
прав. часть=	11,003	1,258	0,233	0,095
	ВЕРНО	ВЕРНО	ВЕРНО	НЕ верно
Нисх. тренд				
m=	16	12	14	8
k=	47	47	47	47
p=	0,383	0,383	0,383	0,383
m/k=	0,3404	0,2553	0,2979	0,1702
eps=	0,0426	0,1277	0,0851	0,2128
прав. часть=	2,774	0,308	0,694	0,111
	ВЕРНО	ВЕРНО	ВЕРНО	НЕ верно

Анализ результатов.

МА 1 линия и ЕМА 1 линия: обладают средними вероятностями $R_{\text{прав}}$ и $R_{\text{ошиб}}$; высокий показатель дисперсии характеризует высокую величину запаздывания подачи правильного сигнала. Это явление объясняется временным запаздыванием и линеаризацией нелинейных трендов [1], и приводит к тому, что решение, исследователем, будет принято не в начале тенденции, а в конце.

SMA 1 линия. Данный алгоритм отличает высокая вероятность подачи правильных сигналов при минимальной величине дисперсии, что гарантирует своевременную подачу правильного сигнала. Декорреляция результатов усреднения, уменьшение эффекта линейных частотных искажений и эффекта линеаризации выделения не линейных трендов [1], [3], [5] приводят к минимальному эффекту смещения и повышают «чувствительность» данного индикатора к цифровому дискретному ряду. Эти особенности приводят к тому, что у данного индикатора высокая вероятность подачи ложных сигналов.

МА 2 линии и ЕМА 2 линии. Обладают низкой вероятностью подачи правильных сигналов при высокой вероятности неподачи неправильных сигналов и высокой вероятностью пропуска правильных сигналов. Использование пересечения «быстрой» и «медленной» линий приводит к «симбиозу» характеристик скользящих

средних: индикатор с меньшим окном скользяния обладает меньшими линейными частотными искажениями и меньшим эффектом линеаризации не линейных трендов; и на оборот, у второго индикатора, с большей величиной окна скользяния. Данный алгоритм улучшает качество генерируемых сигналов, при этом увеличивает временной запаздывание, которое определяется величиной окна скользяния «медленной» линии.

SMA 2 линии. Вероятностью подачи правильных сигналов – высокая (стремится к идеальной). Особенностью данного алгоритма является тот факт, что с ростом окна скользяния эффективная частота среза остается постоянной и происходят незначительные отставания результатов усреднения от реальных трендов. По этому, в данном случае, симбиоз «быстрой» и «медленной» приводит небольшому отсеиванию ложных сигналов. Данный алгоритм, также как и алгоритм SMA 1 линия, не менее чувствительно реагирует на малейшие колебания значений цифрового дискретного ряда.

Индикатор MACD. Имеет высокую вероятность подачи правильного сигнала и высокую вероятность неподачи не правильного сигнала. Низкая вероятность сгенерировать сигнал на участке флэта. Но, базируясь на алгоритме экспоненциального скользящего среднего, обладает всеми недостатками данного алгоритма и высоким показателем дисперсии.

Индикатор MACD_S. Аналогично индикатору SMA 2 линии, имеет самую высокую вероятность подачи правильного сигнала. Базируясь на алгоритме синтетического скользящего среднего, обладает минимальным показателем соотношений между двумя синтетическими скользящими линиями. Это приводит к генерации большого количества ложных сигналов.

Выводы.

1. Индикаторы технического анализа, основанные на классических алгоритмах скользящего среднего, имеют $R_{\text{прав}}$ выше, чем индикаторы, основанные на алгоритме синтетического скользящего среднего. Однако данное преимущество достигается не высоким показателем вероятности генерации правильного сигнала, а высоким показателем вероятности не сгенерировать не правильный сигнал.

2. При линейных частотных искажениях и линеаризации не линейных трендов, генерация правильного сигнала происходит при том условии, что длина выделяемого тренда превосходит величину окна скользяния. Но даже в этом случае генерация правильного сигнала может оказаться бесполезной из-за временного запаздывания. Данные индикаторы рекомендуются к применению только на трендовых участках дискретного ряда, при условии, что длина выделяемого тренда значительно выше величины окна скользяния.

3. Индикаторы, основанные на алгоритме синтетического скользящего среднего, имеют очень высокую вероятность генерации правильных сигналов. При этом остальные слагаемые $R_{\text{прав}}$ имеют низкие показатели.

4. Данный алгоритм формирует статистически независимые результаты усреднения, обладающие минимальными линейно-частотными искажениями в полосе прозрачности. Это приводит к минимальному эффекту смещения полученных результатов усреднения от данных цифрового дискретного ряда. Следствием этого, является, с одной стороны, минимальный показатель дисперсии, как оценки своевременности генерации правильного сигнала.

5. С другой стороны алгоритм SMA тонко реагирует на малейшие колебания дискретного ряда. Это позволяет использовать данные индикаторы как осцилляторы и исследовать с их помощью кратковременные тенденции.

Литература.

1. Смирнов А., Тихонова О. Секрет совершенства индикаторов Марка Джурика раскрыт?: Валютный спекулянт. -2006 №1, с.32-35
2. Смирнов А., Тихонова О. Выделение трендов с минимально возможной колеблемостью: Научные работы ДонНТУ, серия Информатика, кибернетика и вычислительная техника, выпуск

8. 2007, -с.256-261.

3. Тихонова О. А. Новый класс синтетических индикаторов: Материалы второй международной научно-технической конференции «Моделирование и компьютерная графика»-Донецк, ДонНТУ: -2007, с.271-277

4. Венцель Е. С. Теория вероятностей: -М., 1969, -576 с.

5. Тихонова О.А. Сравнительный анализ алгоритма синтетических скользящих средних с классическими алгоритмами скользящих средних с точки зрения цифровой фильтрации: Научные работы ДонНТУ, серия Информатика, кибернетика и вычислительная техника, выпуск 9. 2008, -с. 287-291.

6. Стивен Б. Акелис Технический анализ от А до Я, Москва, 1999г., -735 с.

7. Орлов А.И. Прикладная статистика: Москва, Экзамен. 2004, -656 с.

8. Орлов А.И. Теория принятия решений: Москва, Экзамен. 2004, -597 с.

9. Орлов А.И. Математика случая: Вероятность и статистика – основные факты: Учебное пособие, Москва.: МЗ-Пресс. 2004, – 110 с.

Поступила в редколлегию 11.02.2009