

А.А. МУСАЕВ

Quod est veritas. Трансформация взглядов на системную составляющую наблюдаемого процесса

Musaev A.A. Quod est veritas. Трансформация взглядов на системную составляющую наблюдаемого процесса.

Аннотация. Рассмотрена задача формирования системной составляющей наблюдаемого случайного процесса, отражающего значимые изменения состояния динамической системы. Показано, что критерии качества формирования системной составляющей должны определяться требованиями иерархически вышестоящей метасистемы. При этом возникает новая постановка задачи оценивания, приводящая к необходимости построения вычислительных схем, существенно отличающихся от известных алгоритмов статистической фильтрации.

Ключевые слова: системная составляющая, динамическая система, метасистема, статистическое оценивание, фильтрация.

Musaev A.A. Quod est veritas. Views Transformation at a System Component of Observable Process.

Abstract. The design problem of observed stochastic process system component, reflecting dynamic system significant changes, is considered. It is shown that system component formation quality criteria should be defined by requirements of hierarchic higher metasystem. Thus is arise new estimation statement, leading to necessity of calculating schemes construction, essentially different from known statistical process estimation algorithms.

Keywords: system component, dynamic system, metasystem, statistical estimation, filtering.

1. Введение. Что такое системная составляющая наблюдаемого процесса? Очевидно, что любые дефиниции, генерируемые пресловутым «здравым смыслом», будут обладать расплывчатостью и неоднозначностью. Тем не менее, данный этап постановки задачи является необходимым и существенным с точки зрения понимания формируемой проблемы.

В качестве первого приближения под системной составляющей наблюдаемого процесса будем понимать некоторую оценку истинной траектории (сглаженную кривую), отражающую значимые изменения состояния и свойств исследуемой динамической системы.

Достоинством данного определения является содержащаяся в нем предпосылка для перехода к формализованной постановке на основе калмановской концепции пространства состояний [5]. Недостаток то же достаточно очевиден, термин «значимые изменения» является не определенным без дополнительных уточнений, вытекающих из содержательной постановки конкретной задачи. По-видимому, любое определение, связанное с реальными прикладными задачами, потребует, в

той или иной форме, «внешнего дополнения», что вполне согласуется с логическим постулатом, заложенным в теореме Геделя о неполноте [6, 10]. В этом случае сама постановка задачи определения системной компоненты теряет универсальность и оказывается обусловленной требованиями метасистемы, задающей критерии качества ее восстановления из наблюдаемого процесса. Последнее утверждение полностью согласуется с базовым положением системного анализа, согласно которому цель и качество функционирования любой системы определяется исходя из требований, предъявляемых к ней со стороны иерархически вышестоящей метасистемы, в которую она погружена и с которой она взаимодействует [8].

Приведенные рассуждения, несмотря на их кажущуюся очевидность, могут привести к существенным изменениям в отношении некоторых базовых понятий и даже к самому подходу решения задач определения движения. В частности, категория «адекватности» восстановленной системной составляющей и истинной траектории движения (изменения состояния) системы в фазовом пространстве, основанная на метриках близости, может оказаться просто ненужной в силу того, что алгоритмика метасистемы не нуждается в этом подобии. Дополнительную неопределенность создает само понятие «истинной траектории», относительно которого формируется оценка траектории движения (системная компонента). В некоторых задачах это понятие допускает неоднозначную формализацию. Соответствующие примеры будут приведены ниже.

2. Традиционная формализация. Первоначально рассмотрим традиционный подход восстановления системной составляющей, при котором со стороны метасистемы выдвигается требование ее максимального подобия истинной траектории движения системы. Для оценки качества восстановления системной составляющей обычно используются традиционные для математической статистики квадратичные метрики. Как правило, исходным информационным сырьем служат массивы наблюдений - «зашумленные» эмпирические данные $Y = \{y_k, k = 1, \dots, N\}$, полученные в результате мониторинга определяемого процесса $x(t)$ и представляющие собой реализации некоторого случайного процесса. В первом приближении естественно представить полученные наблюдения $y(t)$ в виде двухкомпонентной аддитивной модели

$$y(k) = F(x(k)) + v(k), \quad k = 1, \dots, N, \quad (1)$$

где $v(k)$, $k = 1, \dots, N$ образует шумовую компоненту наблюдений, обусловленную погрешностями системы измерений (системы мониторинга). Здесь F - некоторый функционал, отражающий зависимость между косвенными наблюдениями и изучаемым процессом.

Сразу заметим, что на практике модель (1) часто используется в упрощенном линеаризованном виде $y_k = A \cdot x_k + v_k$. В случае прямых наблюдений задача становится еще более наглядной $y_k = x_k + v_k$.

Традиционный статистический подход предполагает, что x_k , $k = 1, \dots, N$ представляет собой неизвестный, но детерминированный процесс, подлежащий определению, а погрешности наблюдений v_k , $k = 1, \dots, N$ - стационарный процесс с известной функцией распределения. Именно такой подход позволяет широко использовать традиционную вероятностно-статистическую парадигму, одним из важнейших результатов которой явился *метод наименьших квадратов* (МНК) и его многочисленные модификации [7].

В качестве первого примера рассмотрим задачу определения движений космических тел, в рамках которой впервые был сформулирован МНК. В частности, первое применение МНК было осуществлено К. Ф. Гауссом при определении движения малой планеты Цереры. При этом, по своей постановке, МНК решает задачу минимизации суммы квадратов невязок между траекторными измерениями и их расчетными значениями, соответствующими «истинной траектории». При этом в качестве истинной орбиты использовались расчетные значения траектории, определяемые законами Кеплера с грубым начальным приближением, предоставленным астрономами. Технически данные значения вычислялись как решение системы дифференциальных уравнений движения тела в центральном поле тяготения с правыми частями, учитывающими известные возмущающие факторы.

Таким образом, в классической линейной постановке истинная траектория динамической системы уже известна, задача построения системной составляющей сводится лишь к итерационному уточнению начальных условий краевой задачи. Итерационность решения вытекает из условий линеаризации модели наблюдения, при хорошем начальном приближении решение может быть достигнуто за одну итерацию.

Таким образом, ответ на вопрос «*Quod est veritas?*» во времена Гаусса был достаточно прост – системная компонента полностью соответствует истинной орбите и представляет собой расчетную траек-

торию небесного тела, определяемую законами Кеплера и уточненными по МНК начальными условиями.

3. Байесова парадигма. Повышение требований к точности навигационных определений в задачах космической баллистики сделали и без того очевидным факт того, что законы Кеплера позволяют описать движение космических тел лишь в первом приближении. Влияние множества неизвестных или не полностью определенных факторов (неравномерное распределение масс в телах притяжения, сопротивление верхних слоев атмосферы, давление солнечного ветра и связанная с ней динамика солнечной активности, влияние внешних гравитационных масс и т.п.) создало неопределенность, относящуюся к самой траектории движения, когда движение космического тела оказалось случайным процессом. При этом особую проблему для навигаторов создали нестационарные факторы влияния, такие, как плотность потока излучение Солнца или движение полюсов Земли.

Наличие неопределенности в динамике наблюдаемого процесса привело к возникновению байесовой парадигмы, в соответствии с которой системная компонента сама моделируется в виде случайного процесса. Простейшая форма представления изменения состояния наблюдаемой системы для дискретного времени в этом случае имеет вид $x_k = \Phi_{k/k-1}x_{k-1} + w_k$, $k = 1, \dots, N$, где $\Phi_{k/k-1}$ - переходная матрица (соответствующая фундаментальной матрице системы уравнений возмущенного движения тела в центральном поле тяготения [3]), w_k - шумы системы, обусловленные всей совокупностью неучтенных возмущающих факторов. В случае, когда указанные факторы оказываются достаточно многочисленными и соизмеримыми по уровню воздействия (т.е. когда выполняется условие Линденберга), оказывается удобным, со ссылкой на предельные теоремы теории вероятностей [11, 12], считать w_k , $k = 1, \dots, N$ стационарной гауссовской последовательностью. Данное допущение позволяет использовать практически весь арсенал статистических методов обработки рядов наблюдений.

Таким образом, в соответствии с байесовым подходом, решение задачи определения движения динамической системы сопряжено с использованием трех видов траекторий. Истинная траектория динамической системы представляет собой зашумленную реализацию случайного процесса, для которой, тем не менее, существует теоретическая модель, позволяющая сформировать опорную траекторию (определяющую переходную матрицу $\Phi_{k/k-1}$). Опорная траектория, по сути, представляет собой интегральную кривую, полученную в результа-

те численного интегрирования системы дифференциальных уравнений движения, и, следовательно, для ее определения необходимо уточнение граничных условий. С этой целью можно воспользоваться последовательными алгоритмами фильтрации, позволяющие сформировать системную составляющую. Требование подобия системной составляющей и истинной траектории приводит к тому, что для ее формирования используются алгоритмы последовательного оценивания условного математического ожидания процесса $E\{x(k)\}$, $k=1, \dots, N$. Классической схемой последовательной фильтрации в условиях байесовой постановки стал фильтр Калмана и его многочисленные версии, обладающие повышенной динамической устойчивостью [2, 5].

4. Немного истории: цена математического комфорта. Нетрудно заметить, что обе предыдущие постановки были осуществлены в рамках комфортных для обработчика предположений о стационарности шумовой компоненты рядов наблюдений. Данное предположение обусловлено фундаментальной аксиоматикой, заложенной в основе современной теории вероятностей [4, 13], в соответствии с которой вероятности события определяется путем предельного перехода частоты появления события, происходящего в процессе многократного повторения опыта в идентичных условиях. Если данный опыт разворачивается во времени, то результаты наблюдения могут быть представлены в виде стационарного процесса (т.е. процесса с неизменяющимися вероятностными характеристиками). Если условия опыта постоянно изменяются, как это имеет место на практике, то процесс становится нестационарным, в этом случае все основные понятия теории вероятностей (такие, как функция распределения) теряют смысл.

Известный выход из создавшегося положения состоит в условном разделении нестационарного процесса на системную компоненту, определяющую содержательное изменение контролируемого процесса, и случайную составляющую, представляющую собой оставшуюся часть и образующую стационарный процесс (обычно с известной гауссовой функцией распределения). В случае если системная компонента представляет собой некоторый инерционный процесс, допускающий хотя бы локальную аппроксимацию, и ее можно описать детерминированной математической моделью, например, полиномом. В условиях, когда подобное разделение исходного процесса оказывается возможным, борьба с неопределенностью сводится к снижению влияния на оценку случайной компоненты (сглаживанию, фильтрации) на основе того или иного метода статистической обработки данных. В частности, широкое применение нашли метод наименьших квадратов, ме-

тод максимального правдоподобия, метод наименьших модулей et setega. При этом качество восстановления чаще всего оценивается квадратическими метриками (дисперсией оценки или, в более общем случае, полным квадратом ошибки погрешности оценивания). Сами погрешности определяются, очевидным образом, в виде невязок между оценкой (значениями модели системной компоненты) и исходным рядом наблюдений.

Практически все технические, физические, химические, биологические и другие процессы, имеющие дело с преобразованиями материи и энергии, являются инерционными и допускают указанный подход. При этом инерционность наблюдаемых систем, при правильно выбранном масштабе времени, гарантирует относительную гладкость системной компоненты, допускающую, как минимум, ее локальную аппроксимацию с помощью той или иной математической модели.

Однако, для многих прикладных задач подобный подход, основанный на приведенных выше предположениях, оказывается некорректным и недееспособным. В частности, в процессе торговых операций на фондовом, валютном или ином рынках в неопределенный момент времени какой-то из факторов влияния может оказаться превалирующим в отношении того или иного финансового инструмента [9]. В этом случае трудно ожидать, что шумовая составляющая образует стационарный гауссовский процесс. Следствием некорректного использования традиционных алгоритмов статистического анализа в задачах такого типа является возникновение переменного смещения. Более того, появляется неопределенность относительно самого понятия системной составляющей.

Однако проблема применения условного среднего, формируемого одним из известных алгоритмов МНК или фильтром Калмана, в качестве оценки системной составляющей процесса, на этом не исчерпываются. Так, например, восстановление системной составляющей, используемой в техническом анализе при формировании торговых решений, связано с рядом субъективных факторов, важнейшим из которых является априори выбранная стратегия игры на биржевых площадках. В частности, при долгосрочной инвестиционной стратегии рабочий системный тренд будет существенно отличаться от тренда, используемого при внутрисуточных спекулятивных торговых операциях. Тем не менее, в обоих случаях трейдер или торговый робот будет использовать некоторую виртуальную кривую, описывающую значимые, с точки зрения эффективности торговых решений, изменения состояния используемого финансового инструмента. При этом если трейдер мо-

жет формировать эту системную кривую ментально, путем визуального анализа, интуиции и накопленного опыта, то для торгового робота необходимым условием эффективной работы является построение строгой математической модели, описывающей рабочую системную составляющую наблюдаемого процесса.

Построение подобной модели требует ответа на ряд вопросов, не имеющих однозначного и/или объективного ответа. В частности, очевидным вопросом являются определение понятия системной составляющей, зависящей от требований вышестоящей метасистемы, в роли которой выступает формируемая трейдером торговая стратегия.

5. Системная составляющая безынерционных процессов. Переход к задачам, связанным с игрой на торговых площадках, приводит к необходимости работы с качественно иными безынерционными процессами. В частности, к ним относятся процессы изменения котировок финансовых инструментов на фондовых и валютных рынках. Генезис этих процессов образован ментальными представлениями торгово-экономического сообщества о текущих состояниях этих инструментов и их изменениях под влиянием среды погружения, формирующей совокупность множества разнообразных финансовых, экономических, политических, социальных и других факторов. При этом представления социума о таком нематериальном параметре, как стоимость финансового инструмента, может меняться практически мгновенно и совершенно непредсказуемым образом. В этих условиях формализовано разделить системную и случайную (шумовую) компоненту оказывается крайне сложно, т.к. в состав регулярной составляющей входит колебательный неперіодический процесс, не допускающий эффективной аппроксимации детерминированным процессом [9].

В современной литературе нелинейные динамические системы, допускающие возможность параметрической неустойчивости, относят к категории детерминированного хаоса. Наиболее явной характеристикой таких систем является наличие точек бифуркации и участков с колебательными неперіодическими процессами. В данном случае неопределенность, обусловленная неперіодическими колебаниями, усилена чисто случайной компонентой, и соответствующие ей процессы следует отнести к категории стохастического хаоса. Традиционные методы вероятностно-статистического анализа, как уже отмечалось, не способны осуществлять эффективное оценивание таких сугубо нестационарных процессов. В результате этого многочисленные попытки получения дееспособных управлений торговыми операциями в техни-

ческом анализе на основе традиционных математических технологий не привели к успешным решениям.

В качестве модели описанных выше безынерционных процессов, описывающих, в частности, динамику котировок на биржевых площадках, удобно использовать трехкомпонентную аддитивную модель вида

$$x_k = y_k + v_k + \xi_k, \quad (2)$$

где y_k представляет собой системную составляющую, на основе которой формируются торговые решения, и образованную сложным нелинейным процессом с явно выраженными трендами и колебаниями, v_k - квазисистемная помеховая компонента, образованная несмещенным колебательным непериодическим процессом, и ξ_k - стационарный случайный процесс с распределением, которое во многих случаях можно отнести к классу засоренных нормальных распределений.

Рассмотрим отдельно составляющие предложенной трехкомпонентной модели (2). Основной вопрос состоит в том, что следует понимать под системной компонентой y_k . В случае использования традиционных алгоритмов обработки статистических данных, в качестве такой компоненты обычно выбирают гладкую детерминированную структуру, параметры которой определяются из условия минимизации априори заданной функции ошибок на обучающей выборке. При этом обучающая выборка должна быть получена из той же самой генеральной совокупности. Для динамических систем это означает, что обучающая выборка, представляющая собой ретроспективные данные, полученные, например, на предыдущем участке или образованные скользящим окном наблюдения, должна обладать той же вероятностной структурой, что и данные, используемые в процессе формирования оценки.

Как правило, традиционные статистические технологии опираются на совокупность достаточно строгих теоретических исследований, ограничивающих условия их применимости и оптимальности. На практике, данные ограничения (условия стационарности, независимости, нормальности, однородности et cetera) во многих случаях полностью или частично игнорируются, и, тем не менее, удается получать вполне осмысленные результаты, удовлетворяющие потребности прикладной науки. Но не всегда.

В частности, при решении задач выявления системной составляющей в динамике котировок в принципе не удастся построить сглаженной кривой, отвечающую требованиям трейдеров и других лиц,

принимающих решения (ЛПР). И одной из самых существенных причин является недоопределенность самого понятия системной компоненты.

6. Прокрустово ложе торговых решений. Практика работы с естественными динамическими системами в той или иной форме опирается на представление об их инерционности. Это означает, что для таких систем существуют значимые интервалы времени, в течение которых параметры модели либо остаются неизменными, либо меняются очень незначительно. В этом случае статистические оценки, представляющие собой условные средние, позволяют минимизировать влияние случайной «шумовой» компоненты и отразить наиболее значимые для практики изменения в состоянии системы. Применение традиционной статистической технологии к задачам безынерционной динамики приводит к тому, что в качестве системной компоненты, используемой для оценки трендов, выбирается условное среднее, обычно определяемое последовательно, по мере поступления отсчетов, по скользящей выборке наблюдений. Принципиальной особенностью такого подхода для любого статистического алгоритма оценивания (МНК, ММП, фильтр Калмана, экспоненциальные фильтры и т.п.), является наличие запаздывания, обусловленного влиянием ретроспективной информации, содержащейся в обучающей выборке. При этом увеличении размера обучающей выборки (увеличение коэффициента передачи фильтра, относящегося к накопленным данным) неизбежно приводит к увеличению запаздывания оценки системной компоненты по отношению к текущей динамике котировок, что приводит к регулярным ошибкам первого рода (пропуск решения об открытии или закрытии позиций). Попытки повысить чувствительность оценок условного среднего путем снижения размера обучающей выборки (увеличение коэффициента передачи фильтра, относящегося к текущим данным) неизбежно приводит к росту флуктуационной компоненты и росту ошибок второго рода (ложное открытие и закрытие позиций).

Для традиционной двухкомпонентной модели инерционных систем можно попытаться найти оптимальный размер обучающего скользящего окна, минимизирующего ошибки первого рода при заданном уровне ошибок второго рода (или наоборот) или построить адаптивную схему с контуром регулирования размера памяти. Для безынерционных динамических систем, содержащих хаотическую компоненту, такой подход реализовать не удастся, поскольку крайне узкий диапазон приемлемого размера окна наблюдения определяется не кри-

терием оптимальности, а критерием пригодности. Динамика торгового процесса может в любой момент времени существенно измениться, а несоответствие настройки фильтра системной компоненты, используемой для принятия торговых решений, свойствам текущих наблюдений, как правило, означает проигрыш при проведении торговой операции.

Очевидно, что описанное ограничение на применимость традиционных статистических методов в задачах технического анализа, проводимого при проведении торговых операций, является принципиальным, поскольку, при любом подходе, исходным «сырьем» для формирования оценки системной компоненты являются ретроспективные данные. Тем не менее, качество формируемых торговых решений, возможно, удастся повысить, если перейти к более корректному описанию аддитивных компонент модели (2).

7. Опорные траектории. Прежде всего, попытаемся сформировать представление об «идеальной системной составляющей» с точки зрения ЛППР. Очевидно, что такая модель должна содержать характеристические легко обнаруживаемые изменения в динамике в моменты времени, отвечающие успешным решениям, и не содержать подобным им ложных изменений в другое время наблюдения. Очевидно, что такая постановка должна включать в себя формализованное и параметризованное определение «успешного» решения. Иными словами, понятие идеальной системной компонент существенно зависит от параметров выбранной торговой стратегии. Разумеется, стратегии игры, которая, в свою очередь, может меняться в зависимости от динамических характеристик наблюдаемых процессов. Это означает, что будут изменяться и представления о системной компоненте, используемой для формирования соответствующих решений.

Главный вывод из приведенных рассуждений состоит в том, что системная компонента, выделяемая из наблюдаемого процесса, не обязана отвечать условию адекватности в традиционном смысле этого слова, то есть минимизировать значения какой-то априори выбранной метрики или представлять собой условное среднее. Основное требование к оценке системной компоненты является ее пригодность к обеспечению терминального результата. Разумеется, терминальный результат получается опосредовано, через систему принятия управляющих (торговых) решений, и, следовательно, требованию, выдвигаемые к виду и критериям качества системной компоненты, будут определяться выбранной управляющей (торговой) стратегией.

При всей очевидности приведенных рассуждений, полученный вывод приводит к радикально новому представлению, как о технологиях статистического оценивания, так и о математическом моделировании в целом.

Центральным утверждением терминального подхода состоит в том, что математическое моделирование в целом, и требования к формированию системной компоненты наблюдаемого случайного процесса, в частности, должно определяться исходя из терминальных критериев управления и отвечающих им стратегиям реализации управляющих решений.

Для пояснения данного утверждения ниже будут рассмотрены иллюстрирующие примеры.

8. Торговые стратегии. Критерии эффективности. В качестве иллюстрирующей платформы рассмотрим задачу проведения автоматизированных торговых операций на фондовом рынке. Вариант реализации торговых операции в этих задачах связан с использованием параметров TP (take profit) и SL (stop loss), закрывающими позиции при пересечении значений котировок рабочего инструмента уровнями $x_0 + TP$ или $x_0 - SL$, где x_0 - значение инструмента в момент открытия позиции. Указанные уровни описаны для игры на повышение, в противном случае знаки в приведенных соотношениях надо поменять на противоположные. Заметим, что возможны и другие способы закрытия позиции, например, при одновременном выполнении условий превышения уровня $x_0 + TP$ и развороте сглаженной кривой x_s или с переменными значениями в TP и SL . Подобные модификации не меняют сути предлагаемого подхода.

Выигрышная по вероятности стратегия открытия позиций состоит в выборе последовательности моментов времени открытий $TT_o = (t_1^*, t_2^*, \dots, t_m^*)$ такой что, вероятность того, что после открытия позиции в момент t_j^* , процесс изменения котировок $x(t)$, $t > t_j^*$ превысит уровень $x(t_j^*) + TP$ больше, чем вероятность его выхода за уровень $x(t_j^*) - SL$:

$$P_{win} = P\{x(t) > x(t_j^*) + TP\} > P_{loss} = P\{x(t) < x(t_j^*) - SL\}, \quad (3)$$

где $j = 1, \dots, m$, $t > t_j^*$.

Очевидно, что критерий (3) может быть записан в более «мягкой» форме, используя понятие среднего:

$$\hat{E}\{x(t) > x(t_j^*) + TP\} > \hat{E}\{x(t) < x(t_j^*) - SL\}, \quad (4)$$

где $j=1, \dots, m$, $t > t_j^*$, $\hat{E}\{\cdot\}$ - среднее число событий, описанных в фигурной скобке.

Заметим, что оба приведенных условия должны быть дополнены естественным ограничением типа $TP^*n_{win} \gg SL^*n_{loss}$, где n_{win} - число выигрышных открытий позиции, а n_{loss} - число проигрышных открытий. Иначе, даже при выполнении критериев (3) и (4) «выигрыш» может оказаться отрицательной величиной.

Критерии (3) и (4), в соответствии с терминологией общей теории эффективности систем [8], относятся к категории критериев пригодности. Реализация этих критериев позволяет получить непроигрышное управление торговыми операциями. В случае, когда необходимо осуществить сравнение пригодных стратегий или составляющих их схем моделирования, используют критерии превосходства или оптимальности. Так, например, критерий превосходства стратегии, опирающейся на модель 1 системной компоненты ряда наблюдений, над той же стратегией, использующую модель 2 этой же компоненты, можно представить в виде

$$\hat{E}_{M1}\{x(t) > x(t_j^*) + TP\} - \hat{E}_{M1}\{x(t) < x(t_j^*) - SL\} > \dots \\ \hat{E}_{M2}\{x(t) > x(t_j^*) + TP\} - \hat{E}_{M2}\{x(t) < x(t_j^*) - SL\}, \quad (5)$$

где $j=1, \dots, m$, $t > t_j^*$.

Соотношение (5) означает, что среднее значение разности между числом успешных открытий и ошибочных управлений (открытий позиции) при использовании модели 1 системной составляющей больше соответствующего числа при использовании модели 2. Данный критерий, опирающийся на соотношение (4), можно также записать, используя вместо разности, отношение правой и левой части этого же выражения. Заметим, что представленный критерий (5) должен быть дополнен критерием пригодности (4), в противном случае можно столкнуться с ситуацией, когда обе сравниваемые модели ведут к общему проигрышу.

Более жесткий критерий оптимальности предполагает, что выбранная стратегия или модель являются наилучшими в заданном классе с точки зрения выбранного показателя эффективности. Так, например, использую схему усредненного показателя пригодности (4), мож-

но определить модель M_{opt} , как наилучшую в заданном классе $\{M\}$, если она отвечает условию оптимальности вида

$$\hat{E}_{M_{opt}}\{x(t) > x(t_j^*) + TP\} - \hat{E}_{M_{opt}}\{x(t) < x(t_j^*) - SL\} = \max(M \in \{M\}),$$

$$j = 1, \dots, m, \quad t > t_j^*. \quad (6)$$

9. Примеры. В качестве наглядного примера, иллюстрирующего приведенную концепцию терминального моделирования, рассмотрим задачу выявления системной составляющей ряда наблюдений, полученных за изменением котировок индекса DJ в течение двух случайно выбранных однодневных торговых сессий. Соответствующий график динамики котировок представлен на рис. 1 и 2 и имеют вид хаотически изменяющегося колебательного нестационарного процесса с участками аперiodической динамики.

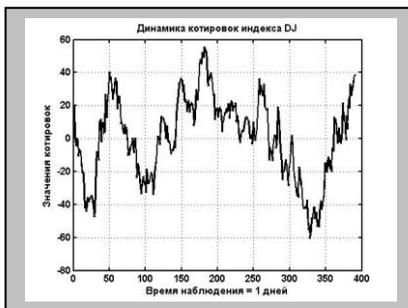


Рис. 1. Пример 1 изменения котировок индекса DJ

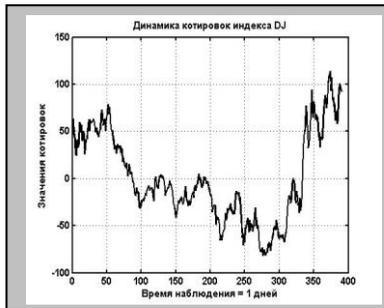


Рис. 2. Пример 2 изменения котировок индекса DJ

В качестве базовой торговой стратегии S_0 используем «классическую» схему формирования торговых решений, принятую в техническом анализе: позиция открывается при переломе системной составляющей в сторону второй ветви перелома. Предположим, что уровень $TP = SL = 50$ пунктов. Исходя из выбранной стратегии управления и значения параметра TP , сформируем апостериорно, «вручную», идеальную системную модель системной компоненты, позволяющую получить максимальный выигрыш. В дальнейшем такую модель будем называть опорной.

Первоначально рассмотрим схему формирования опорной модели в виде скорректированной линии апостериорного условного среднего.

Данная работа осуществляется в два этапа. На первом этапе формируется традиционная апостериорная модель на основе одного из известных алгоритмов статистического сглаживания. В частности, в качестве опорной траектории можно использовать полином заданного порядка с подгонкой параметров по МНК. На втором этапе осуществляется вывод массива данных, отвечающих полученной на первом этапе кривой, и их ручная коррекция, основанная на визуальном сравнении графика полученной опорной кривой и ее виртуальной идеальной формой.

Идеальная форма опорной траектории предполагает выполнение двух, достаточно противоречивых условий:

1. идеальная траектория имеет явно выраженные идентификаторы правильных открытий. В роли таких идентификаторов могут выступать переломы системной (сглаженной) траектории в моменты времени $TT_o = (t_1^*, t_2^*, \dots, t_m^*)$, после которых кривая котировок условно монотонно возрастает или убывает до уровня TP . При этом под условной монотонностью понимается возможность временного и многократного изменения направления динамики котировки рабочего инструмента на противоположный, однако величина просадки не должна превышать априори выбранный уровень SL . Иными словами, условная монотонность в этом случае обеспечивает достижения уровня TP раньше, чем уровня SL .

2. идеальная траектория не должна провоцировать «ложных тревог», в роли которых выступают локальные переломы, в результате которых реальная кривая, не достигнув уровня TP , изменит свое направление и сохранит условную монотонность вплоть до достижения уровня SL . В этом случае понятие условной монотонности охватывает динамику с возможными многократными переломами, однако в результате уровень SL будет достигнут раньше, чем уровень TP .

Рассмотрим реализацию первого этапа формирования опорных траекторий для выбранных выше примеров изменения котировок индекса DJ (рис. 1-2). На рис. 3-4 приведены графики апостериорной подгонки по МНК соответствующих зашумленным кривых.

В первом случае наилучшую, в смысле соответствия указанным выше требованиям, аппроксимацию обеспечивала подгонка полиномом 10 степени, во втором – 6 степени. При этом в обоих случаях при решении системы нормальных уравнений использовалось псевдообращение, т.к. применение аппроксимирующих полиномов степени 4 и выше приводило к вырожденности матрицы наблюдений.

Из приведенных графиков видно, что их применение в системе поддержки принятия решений (СППР) обеспечивает в первом случае 4 успешных открытия, а во втором 2 при полном отсутствии «ложных тревог» и отвечающих им проигрышных открытий позиций.

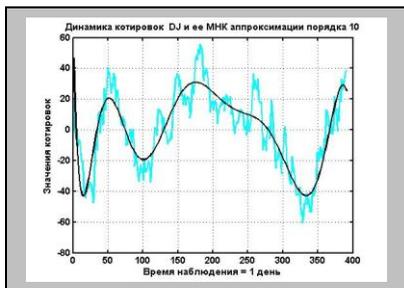


Рис. 3. График опорной траектории для примера 1 изменения котировок индекса DJ

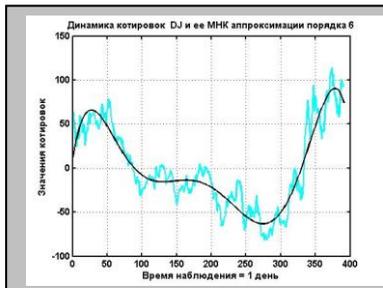


Рис. 4. График опорной траектории для примера 2 изменения котировок индекса DJ

Повышение порядка аппроксимации приводит к росту числа выигрышных открытий, но одновременно повышает вероятность ошибок второго рода, «ложных тревог», и соответствующему увеличению числа проигрышей. Соответствующие примеры приведены на рис. 5-6.

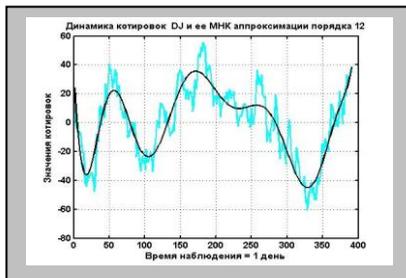


Рис. 5. График опорной траектории для примера 1 с увеличенным порядком аппроксимации

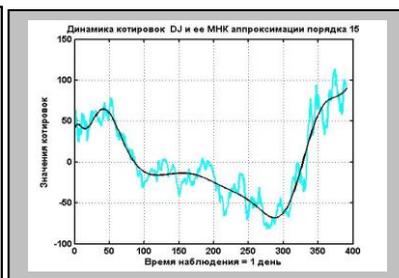


Рис. 6. График опорной траектории для примера 2 с увеличенным порядком аппроксимации

Заметим, что представленные аппроксимации достаточно близки к оптимальным для выбранных параметров торговой стратегии S_0 . В частности, в примере 1 получен наилучший результат для выбранного значения $TP = 50$. В примере 2 возможно увеличение общего выигрыша при дополнительном открытии позиций на 180, 355 и, возможно, на 248 шагах. Однако такой подход потребовал бы идентификации

крайне «острых» перегибов, и как следствие, к непредсказуемому росту «ложных тревог».

Таким образом, удачный выбор порядка модели позволил построить опорную кривую уже на первом этапе, без дополнительной ручной коррекции. Однако если в качестве параметра стратегии было бы выбрано значение $TP=20-25$, то это привело бы к большому числу пропусков открытий с выигрышным результатом. Непосредственное и существенное увеличение порядка аппроксимирующего полинома позволило бы обнаружить соответствующие перегибы, но в то же время, как уже отмечалось, неизбежно возросло бы и число ложных тревог. Данное противоречие является принципиальным и не позволяет сформировать идеальной стратегии технического анализа при наличии описанной выше квазисистемной хаотической составляющей. В этом случае для формирования идеальной опорной траектории по результатам апостериорных данных необходим либо второй, ручной этап коррекции, либо переход к другим методам восстановления системной составляющей, не требующим существенного роста порядка аппроксимирующей модели.

Один из таких подходов связан с применением технологии сплайн-аппроксимации [1]. При этом выбор узлов сплайн-аппроксимации осуществляется вручную, а сама аппроксимация кубическими сплайнами реализуется соответствующей программой. Пример построения опорной кривой для однодневной торговой сессии приведен на рис. 7.

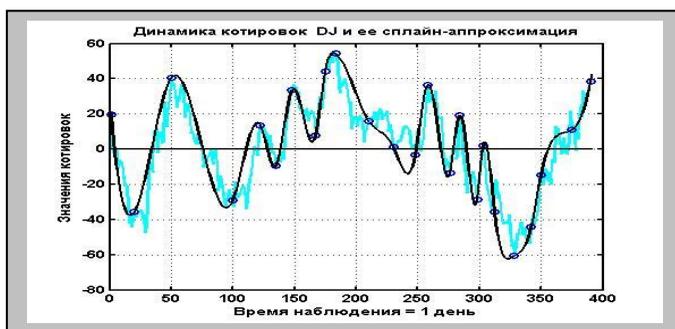


Рис. 7. Пример построения опорной траектории процесса изменения котировок индекса DJ на основе кубической сплайн-аппроксимации

Из приведенного примера видно, что построенная модель опорной траектории, в сочетании с описанной выше стратегией открытия позиций в моменты ее перегибов, позволяет получить оптимальный результат даже при незначительных величинах TP . Так например, в приведенном примере $TP = 20$, точками на графике обозначены узлы аппроксимации.

В качестве окончательного результата для приведенного примера укажем, что сформированная опорная траектория позволяет получить 15 выигрышных открытий и ни одной «ложной тревоги». Имеется два пропуска выигрышных открытий в районе 370 наблюдения, однако данное изменение носит флуктуационный характер. Для восстановления таких шумоподобных вариаций кубические сплайны слишком инерционны. При этом попытки увеличить порядок сплайна или существенно увеличить число узлов аппроксимации неизбежно привели бы к существенному росту «ложных тревог» за счет возникновения перегибов опорной траектории на флуктуационных колебаниях котировок.

Дальнейшее развитие приведенного подхода к формированию опорных траекторий еще дальше уводит от традиционных статистических схем выделения системной составляющей из его смеси с шумами.

В отличие от задач наблюдений за космическим и другими материальными объектами, в которых изменение состояния образует системную составляющую, а чисто случайная или «цветная» (то есть содержащая значимую автокорреляционную составляющую) помеховая компонента, образованная погрешностями системы наблюдения, в задачах торговых операций истинная динамика определяется самими наблюдениями. Отсюда непосредственно вытекает, что опорная траектория, используемая при формировании управляющих решений, носит чисто виртуальный характер и не обязана отражать системную динамику наблюдаемых процессов в традиционном смысле этого слова. Она может явно расходиться с системной тенденцией, как это видно на последнем примере, ее главная цель – отражать идентификационные характеристики, используемые в процессе выработки торговых решений.

В качестве примера на рис. 8-9 приведены два примера построения кусочно-линейной опорной траектории, описывающих изменения котировок индекса DJ с позиции выбранной торговой стратегии S_0 с параметром $TP = 50$.

Нетрудно видеть, что полученные опорные траектории позволяют получить явный выигрыш, в первом случае, в 300 пунктов, а во втором

– в 150 пунктов. Другой вариант выигрышной стратегии S_1 для данных опорных траекторий состоит в одновременном закрытии и открытии позиций в точках перегиба. В этом случае выигрыш в первом примере выигрыш составит приблизительно 520 пунктов, а во втором – 850 пунктов.

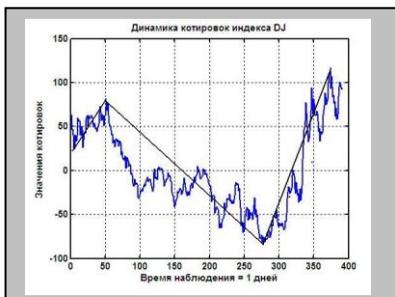


Рис. 8. Пример 1 кусочно-линейной опорной траектории изменения котировок индекса DJ

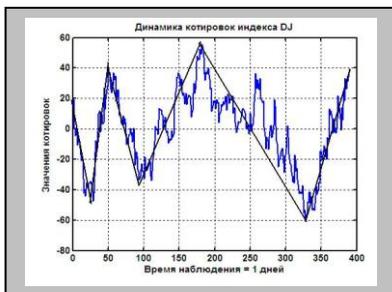


Рис. 9. Пример 2 кусочно-линейной опорной траектории изменения котировок индекса DJ

Заметим, что полученные результаты являются апостериорными и, следовательно, непригодными для реализации в процессе реальной работы на торговых площадках. Однако они позволяют получить опорные траектории, необходимые для формирования эффективной системной составляющей наблюдаемого процесса. Кроме того, они позволяют построить платформу для сравнительного анализа используемых стратегий и последовательной оптимизации структуры отвечающих им системных составляющих.

Заключение. Основной вывод, вытекающий из настоящего подраздела, состоит в том, что применение терминального подхода к задаче формирования системных составляющих наблюдаемого процесса качественно изменяет представление о методологии статистического восстановления искомым траекторий, включая вопросы оценки эффективности полученного результата.

Традиционные подходы, с использованием в качестве базовой динамики условного среднего или других кривых, построенных на основе минимизаций функций от различных вариантов метрик разброса, не всегда обеспечивают эффективной поддержки работы систем принятия решения или другого терминального пользователя получаемой информации.

Важным элементом анализа эффективности построения системных составляющих является процесс апостериорного формирования опорных траекторий. Технология формирования опорных траекторий должна быть строго согласована с требованиями терминального пользователя. В частности, для систем поддержки торговых решений опорные траектории должны формироваться с учетом структуры выбранной стратегии торговых операций.

Критериями качества системной составляющей рядов наблюдений, формируемой с использованием предложенной технологии опорных траекторий или иным способом, должны выступать не степень ее близости к условному среднему, отражающей сглаженную динамику изменения случайного процесса, а степень ее информационной эффективности по отношению к терминальному звену решаемой задачи.

Литература

1. *Алберг Дж., Нильсон Э., Уолш Дж.* Теория сплайнов и ее приложения // Пер. с англ. под ред. К.И. Бабенко. М.: Мир, 1972. 319с.
2. *Браммер К., Зиффлинг Г.* Фильтр Калмана - Бьюси // Пер. с нем. под ред. И.Е. Казакова. М.: Наука, 1982. 199с.
3. *Жданюк Б.Ф.* Основы статистической обработки траекторных измерений. - М.: Сов. радио. 1978. 384 с.
4. *Колмогоров А.Н.* Основные понятия теории вероятностей. М.: Мир, 1974. 120с.
5. *Калман Р., Фалб П., Арбиб М.* Очерки по математической теории систем // Пер. с англ. под ред. Я.З. Цыпкина. М.: Мир, 1971. 400с.
6. *Клини С. К.* Введение в метаматематику // Пер. с англ. М.: Либроком, 2008. 526с.
7. *Линник Ю.В.* Метод наименьших квадратов и основы теории обработки наблюдений М.: Физматгиз. 1958. 349 с.
8. *Морозов Л.М., Петухов Г.Б., Сидоров В.Н.* Методологические основы теории эффективности: Учебное пособие. - Л.: ВИКИ им. А. Ф. Можайского, 1982. - 236с.
9. *Мусаев А. А.* Структуризация хаоса на рынках капитала. Труды СПИИРАН, вып. 9. СПб.: Наука, 2009. 66-80 с.
10. *Нагель Э., Ньюмен Д. Р.* Теорема Гёделя, пер. с англ., М.: URSS (Красанд). 2010. 120с.
11. *Феллер В.* Введение в теорию вероятностей и её приложения. Т. 1 // Пер. с англ. М., Мир. 1963. 512 с.
12. *Феллер В.* Введение в теорию вероятностей и ее приложения. Т. 2 // Пер. с англ. Мир, 1967. 752 с.
13. *Mises Richard von.* *Mathematical Theory of Probability and Statistics*, New York, Academic Press, 1964. 577 p.

Мусаев Александр Азерович — д.т.н., профессор; ведущий научный сотрудник научно-исследовательской группы информационных технологий в образовании Учреждения Российской академии наук Санкт-Петербургский институт информатики и автоматизации РАН (СПИИРАН), научный консультант ОАО Специализированная инжиниринговая компания «Севзапмонтажавтоматика». Область научных интересов: анализ данных, управление и прогнозирование в сложных динамических системах, стохастические хаотические системы. Число научных публикаций — 172. amusaev@szma.com,

www.szma.com; СПИИРАН, 14-я линия В.О., д. 39, г. Санкт-Петербург, 199178, РФ; р.т. +7(812)350-5885, факс +7 (812)350-1113.

Musaev Alexander Azerovich — Dr. in Appl. Math., professor; leading researcher, Education Information Technology Group, St. Petersburg Institute for Informatics and Automation of the Russian Academy of Sciences (SPIIRAS), expert, public corporation Specialized engineering company "Sevzapmontageautomatica". Research interests: data analysis, complicated dynamic systems prognosis and control, stochastic chaos systems. The number of publications — 172. amusaev@szma.com, www.szma.com; SPIIRAS, 39, 14-th Line V.O., St. Petersburg, 199178, Russia; office phone +7(812)350-5885, fax +7(812)350-1113.

Рекомендовано ИГИТО СПИИРАН, рук. ктн, доц. А. В. Тишков.
Статья поступила в редакцию 23.12.2010.

РЕФЕРАТ

Мусаев А.А. **Quod est veritas. Трансформация взглядов на системную составляющую наблюдаемого процесса.**

Традиционная схема анализа данных, полученных в результате слежения за изменением состояния динамической системы, предполагает формирование некоторой системной составляющей, отражающей наиболее характерные свойства и основные тенденции в динамике исследуемого объекта. Общим положением в большинстве подходов к решению данной задачи является формирование системной компоненты, как условного среднего.

При решении ряда прикладных задач такой подход является оправданным. Действительно, в соответствии с общим положением системного анализа, требования к задаче формирования системной составляющей определяется со стороны вышестоящей метасистемы. Для большинства технических задач, связанных с инерционными процессами, приведенный подход означает подобие в среднем системной составляющей истинному движению наблюдаемой системы. Соответствие истинному движению позволяет формировать объективно обоснованные управленческие решения.

Однако в случае наблюдения ряда безынерционных виртуальных объектов, требование соответствия истинному процессу вообще не стоит. Так, например, при наблюдении состояния фондовых рынков, валютных рынков, динамики котировок различных активов и т.п. истинным процессом являются сами котировки. Шумы наблюдения в процессе мониторинга состояния полностью отсутствуют. Сам процесс изменения состояния рынка моделируется трехкомпонентной моделью включающую в себя участки с аperiodической динамикой неопределенной длительности, образующие локальные тренды, на которые накладываются колебательный неperiodический процесс (хаотическая составляющая) и случайная составляющая, образованная квазистационарным процессом с засоренным гауссовским распределением.

В роли метасистемы в этой ситуации выступает система выработки управляющих решений, для которой важнейшим условием успешной работы является выделение системной составляющей, чувствительной к значимым перегибам тренда, и не реагирующей на изменения, обусловленные хаотической и случайной составляющими процесса. В качестве апостериорного эталона можно сформировать опорную траекторию, идеально удовлетворяющую системе выработки управляющих решений с учетом выбранной торговой стратегии и ее параметров. В этом случае критерием качества формирования системной компоненты будет степень ее подобия опорной траектории, причем степень подобия определяется не метрикой близости, а синхронностью перегибов их трендов. При этом возникает качественно новая постановка задачи оценивания, приводящая к необходимости построения вычислительных схем, существенно отличающихся от известных алгоритмов статистической фильтрации.

SUMMARY

Musaev A.A. **Quod est veritas. Sights Transformations at a System Component of Observable Process.**

The analysis traditional scheme of the data, received as a result of dynamic system state changes tracking, assumes formation of some system component, reflecting the most typical properties and the basic tendencies in investigated object dynamics. In essence it is a question of system basic trends allocation from observations, containing noise component. The given problem has been solved on the basis of known data processing methods, such as least-squares method, maximum-likelihood method, Kalman filter and its modifications, etc. However in all examined approaches, the generality is formation system constituent as conditional average. In other words, for each moment of time required system components meaning is searched or as average value on realization set, or, at the consecutive scheme, as average value on sliding sample, taking into account an available on this time interval trend.

At the decision of some applied problems such approach is justified. Really, according to system analysis general provisions, requirement to system component formation is defined from higher metasytem. For the majority of technical tasks, connected with inertial processes, the resulted approach means similarity on the average a system component to true movement of observable system. Conformity to true movement allows to form objectively well-founded control decisions.

However in case of number instantaneous virtual objects observations, conformity to true process requirement at all don't cost. So, for example, at supervision of stock markets state, the currency markets state, dynamics of various actives quotations, etc., true process are observations themselves. Observation noises in state monitoring process are absent completely. Process of market state changes is modeled by ternary model, including sites with uncertain duration periodic dynamics, forming local trends, on which oscillatory acyclic process (a chaotic component) and the random component, formed by quasi-stationary process with contaminated normal distribution, are imposed.

In this situation metasytem is presented as decision support system, for which the major condition of successful work is allocation of the system component. For all that designed system component must be sensitive to significant trend excesses, and, in the same time, do not reacting to the changes, caused by chaotic and random process components. As a case in point of a posteriori standard it is possible to generate supporting trajectory, ideally satisfying to decision support system, taking into account chosen trade strategy and its parameters. In this case similarity degree between supporting trajectory and system component will be quality criterion of system component designed system. At that similarity degree is defined not by the affinity metrics, but their trends excesses synchronism. Thus there is qualitatively new statement of a process estimation problem, leading to necessity of computing schemes construction, essentially different from known statistical filtration algorithms.