УДК 62-501.12

А.А. Мусаев

МОДЕЛИРОВАНИЕ КОТИРОВОК ТОРГОВЫХ АКТИВОВ

Мусаев А.А. **Моделирование котировок торговых активов.**
Аннотация. Рассмотрена задача моделирования изменения цен торговых активов, котируемых на срочных рынках (валютном, фондовом и т.п.). Приведен исторический обзор эволюции моделей процесса котировок и осуществлен критический анализ основных подходов к решению данной задачи. Представлены материалы численных исследований, уточняющих структуру процесса ценообразования.
Ключевые слова: котировки активов, моделирование, нестационарный процесс, хаотический процесс.

Musaev A.A. **Modeling of trading assets quotations.**
Abstract. The problem of the modeling of trading assets quoted at the currency markets, shares etc. is considered. The historical review of quotations' process models evolution is presented. Critical analysis of the basic approaches in the decision of the given problem is carried out. Numerical researches specifying structure of price formation process are presented.
Keywords: quotations of actives, modeling, non-stationary process, chaotic process.

1. Введение. Задачи моделирования динамических систем в большинстве прикладных задач ориентированы на формирование алгоритма или генератора временных последовательностей, свойства которых в каком-то заранее оговоренном смысле близки к свойствам наблюдений за изменением состояния моделируемой системы. При этом, как правило, в качестве меры близости, характеризующей степень адекватности модели, выступает субъективно выбранная метрика, определяющая интегральное рассогласование между наблюдениями и моделью. Применение модели, обладающей свойствами, подобными реальной динамической системе, позволяет целенаправленно изучать свойства системы, ее реакцию на имитацию различных воздействий и получать необходимые знания в контексте стоящих перед исследователем прикладных или теоретических задач.

В отношении процессов, связанных с изменением котировок торговых активов, задача моделирования имеет существенные отличия от традиционной схемы. Дело в том, что проведение вычислительных экспериментов по формированию и применению торговых стратегий не нуждается в создании моделей реализаций динамики котировок. Существуют богатые архивы многомерных рядов котировок, полностью соответствующие реальным процессам ценообразования и позволяющие проводить любые вычислительные эксперименты. Тем не менее, задача моделирования динамики котировок является важным элементом проведения исследовательских работ, ориентированных на построение эффективных торговых стратегий. В настоящей статье осуществлен критический обзор существующих подходов к задаче моделирования динамики котировок, а также представлены некоторые результаты численного анализа, ориентированного на построение таких моделей.

2. Моделирование динамики котировок: от броуновского движения до модели Самуэльсона. По-видимому, первым значимым подходом к формализации процессов ценообразования, протекающих на фондовых и иных биржах, была диссертационная работа французского математика Луи Башелье «Theorie de la speculation», представленная к защите в 1900г. [21]. В основу его концепции была положена гипотеза о полной статистической независимости рядов котировок биржевых активов. Иными словами, корреляция между двумя следующими друг за другом наблюдениями равна нулю, а автокорреляционная функция вырождается в $\delta-$*функцию* Дирака. Природным аналогом такого процесса является *броуновское движение*, открытое шотландским ботаником Робертом Брауном в 1827г.

С математической точки зрения броуновское движение представляет собой *случайный процесс с независимыми приращениями* $\{X_t\}_{t\in T\subset[0,\infty)}$, у которого для любой возрастающей временной последовательности $t_0=0, t_{1,} \dots, t_n$ случайные величины $X_{t_0}, X_{t_1}-X_{t_0}, \dots,$ $X_{t_n}-X_{t_{n-2}}$ являются независимыми. В дальнейшем будем рассматривать простейшую равномерную дискретную временную последовательность $T=\{0,1,\dots,N\}$.

Одним из наиболее известных вариантов процесса с независимыми приращениями является процесс одномерных *случайных блужданий*

$$X_t = X_0 + \sum_{k=1}^{t}\delta X_k$$

где $\delta X_t = X_t - X_{t-1} = \begin{cases} 1, & с\ вероятностью\ p_{t,} \quad 0<p_t<1, \\ -1, & с\ вероятностью\ q_t = p_t - 1, \end{cases}$

последовательность независимых случайных величин. Приложение этого процесса к моделированию ценообразования приведено, например, в [3].

Естественным обобщением случайного блуждания для моделирования динамики котировок явился винеровский процесс $W_t = \{X_t\}_{t \in T \subset [0,\infty)}$ [13], у которого независимые приращения ΔX_t представляют собой гауссовский процесс, т.е. для $\forall t, s \in [0,\infty)$, $s < t, \quad \sigma = const$

$$\Delta X_{ts} \in N\{0, \sigma^2 \cdot |t - s|\}.$$

Обозначим через v – случайную величину, подчиненную нормальному закону с параметрами $N\{0,1\}$. Тогда для винеровского процесса $\Delta X_{\Delta t} = v\sqrt{\Delta t}$ или, переходя в пределе к непрерывному случаю, $dX = v\sqrt{dt}$.

Винеровский процесс и до сих пор остается базовой моделью динамики котировок биржевых активов, несмотря на многочисленную и, в целом, справедливую критику. Типовыми недостатками модели, отражающими ее прикладную неадекватность процессам ценообразования торговых активов, считаются:

1. Наличие отрицательных значений котировок. Заметим, что данный недостаток не является существенным, т.к. для трейдеров важны не сами котировки, а их приращения. Тем не менее, еще Башелье для устранения этого недостатка предложил использовать логарифмическое преобразование, что существенно снизило наглядность и интерпретируемость процессов.

2. Существенные отклонения формы эмпирического распределения приращений ΔX_{ts} от гауссовой кривой $N\{0, \sigma^2 \cdot |t - s|\}$. В частности, применение критериев нормальности типа Харке-Бера (*JB*) [1, 5], основанного на равенстве нулю коэффициентов асимметрии и эксцесса, или непараметрического критерия Колмогорова подтверждает негауссовскую природу данных. Особенно остро данное несоответствие проявляется при обработке данных с утяжеленными хвостами распределений, что соответствует более высокому уровню (по сравнению с нормальным распределением) больших уклонений.

3. Существенная нестационарность рядов наблюдений.

4. Наличие хаотической составляющей в рядах наблюдений [8, 12].

Перечисленные несоответствия модели привели к различным обобщениям, усложняющим базовую модель, но обладающую, по мнению авторов этих дополнений, большей степенью экономического или математического подобия. Проблема подобия является совершенно неоднозначной и связана с содержательной составляющей задачи, ради которой осуществляется моделирование. В случае, когда речь идет о моделировании с целью создания полигона данных, используемого для отладки торговых стратегий или торговых роботов, требуется получить соответствие статистических свойств модели и реальных данных. В то же время ряды наблюдений за изменением котировок образуют нестационарные процессы, у которых вероятностные параметры непрерывно изменяются во времени, причем характер этих изменений определяется хаотическим процессом.

Одним из направлений обобщения винеровской модели является добавление составляющей, имитирующей системный тренд αdt. Для непрерывного случая получаем соотношение, представляющее собой стохастическое дифференциальное уравнение Ито [28]:

$$dX = \alpha dt + \sigma v\sqrt{dt}\ .$$

В качестве σ обычно используется оценка параметра рассеянья (например, *среднеквадратическое отклонение* (ско)). Оценка ско осуществляется по ряду котировок на предшествующем участке наблюдения. Для дискретного времени данная модель (известная как модель Башелье) имеет вид

$$\Delta X = \alpha\Delta t + \sigma v\sqrt{\Delta t} \in N\{\alpha\Delta t, \sigma\sqrt{\Delta t}\}\,.$$

Обобщением модели Башелье явилось стохастическое дифференциальное уравнение, в котором параметры тренда α и волатильности σ являются функциями от цены актива:

$$dX = \alpha(X, t)dt + \sigma(X, t)\varepsilon\sqrt{dt}\ .$$

Используя понятие среднего приращения цены μ и абсолютного дохода μX, получим дифференциальное соотношение $dX = \mu X dt$. С учетом винеровской компоненты можно записать уравнение в терминах ставки доходности dX / X [6, 17]

$$\frac{dX}{X} = \mu dt + \sigma\varepsilon\sqrt{dt} \in N\{\mu dt, \sigma\sqrt{dt}\}\ .$$

Приведенное уравнение изменения цены актива называют моделью Самуэльсона [30]. Решение этого уравнения имеет вид

$$X_t = X_0 \exp\left(\mu t + \sigma W_t - \frac{\sigma^2}{2}t\right).$$

где W_t – стандартный винеровский процесс.

3. Моделирование динамики котировок: временные ряды. Второе направление моделирования траекторий биржевых активов связано с применением моделей временных рядов. Классификация основных типов моделей временных рядов, используемых для этой цели, представлена на рис. 1. При этом подходе базовыми математическими конструкциями являются модели *авторегрессии, скользящего среднего* и их производные.

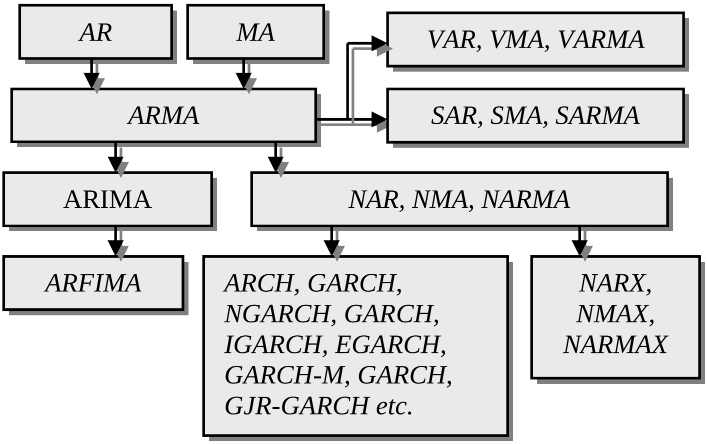

Рис. 1. Классификация основных типов моделей временных рядов.

*Модель авторегрессии (*Autoregressive *process,* $AR(p)$*)* задается соотношением

$$X_t = \sum_{j=1}^{p} a_j X_{t-j} + v_t ,$$

где X_t, X_{t-i}, $i = 1,\ldots,p,$ – значения переменной x в соответствующие моменты времени; p – порядок модели; $\alpha_1,\ldots,\alpha_p$ – параметры модели; v_t – случайная составляющая с нулевым математическим ожиданием, конечной дисперсией и единичной автокорреляционной матрицей, свидетельствующей об отсутствии автокорреляционной связи в рядах наблюдений, т. е. $v_t \in N\{0, \sigma_v^2\}$, $\operatorname{cov}(v) = \sigma_v^2 E$, E – единичная матрица.

В некоторых случаях используется нотация на основе оператора лага $L^d X_t = X_{t-d}$. В этом случае процесс $AR(p)$ может быть записан в виде

$$(1 - \sum_{j=1}^{p} a_j L^j) X_t = v_t .$$

Статистическая подгонка модели $AR(p)$, как и любой другой модели временного ряда, предполагает использование двухэтапной процедуры. На первом этапе уточняется структура модели, т.е. определяется рациональный порядок модели (величина p), на втором осуществляется оценка значений ее коэффициентов.

Модель скользящего среднего (Moving Average, $MA(q)$) определяется в виде линейной комбинации текущего и прошедших значений случайной составляющей $v_t, v_{t-1},\ldots,v_{t-q}$, соответствующей «белому шуму»:

$$X_t = \sum_{j=1}^{q} \beta_j v_{t-j} + v_t$$

где $\beta_1,\ldots,\beta_q$ – параметры модели. При использовании оператора лага эту модель можно записать в виде

$$X_t = (1 - \sum_{j=1}^{q} \beta_j L^j) v_t = \beta(L) v_t .$$

Модель авторегрессии-скользящего среднего (Autoregressive-Moving average, $ARMA(p, q)$) определяется как аддитивная комбинация выше рассмотренных моделей $AR(p)$ и $MA(q)$:

$$X_t = \sum_{j=1}^{p} a_j X_{t-j} + v_t - \sum_{j=1}^{q} \beta_j v_{t-j} + v_t ,$$

где $\alpha_1,\ldots,\alpha_p$, $\beta_1,\ldots,\beta_q$ – параметры модели; p – порядок авторегрессии; q – порядок скользящего среднего. Другая запись этой модели, использующая оператор лага, имеет вид $\alpha(L) X_t = \beta(L) v_t$.

Процесс, отвечающий модели $ARMA(p, q)$, является стационарным, если корни уравнения $\alpha(L) = 0$ лежат вне единичного круга. В противном случае это процесс нестационарен.

Считается, что данная модель наиболее адекватна наблюдениям процесса, в котором «собственная» стохастика усугублена рядами внешних неконтролируемых возмущений, формируемых новостной информацией и групповой психологией участников торгов [18].

Модель $ARMA$ оказалась весьма конструктивной и получила целый ряд обобщений, авторы которых сочли их более адекватными реальным рядам наблюдений котировок исходя из различных критериев экономического подобия.

В частности, очевидным обобщением описанных выше временных рядов являются модели с нелинейной зависимостью от предшествующих наблюдений и шумовой компоненты. Соответствующие нелинейные аналоги получили наименование нелинейной авторегрессии (*Nonlinear Autoregressive*, NAR), нелинейного скользящего среднего (*Nonlinear Moving Average*, NMA) и нелинейной авторегрессии-скользящего среднего ($NARMA$).

Важным обобщением $ARMA(p, q)$ процесса является *модель Бокса-Дженкинса*, известная также как модель авторегрессии-интегрированного скользящего среднего *(Autoregressive Integrated Moving Average*, $ARIMA(p, d, q)$) [2]. Данная модель используется для описания нестационарных процессов, обладающих некоторой однородностью. Однородность проявляется в том, что, если не учитывать локальных трендов, различные участки этого процесса до определенной степени подобны. Указанное свойство характерно для процессов, у которых конечная разность некоторого порядка d является стационарным процессом. Модель Бокса-Дженкинса обычно задается уравнением вида

$$\alpha(L)\nabla^d X_t = \beta(L)v_t, \tag{1}$$

где ∇^d – оператор взятия конечной разности порядка d, L – оператор лага.

Уравнение (1) может быть записано в виде

$$(1-\sum_{i=1}^{p}\alpha_i L^i)(1-L)^d X_t = (1+\sum_{j=1}^{q}\beta_i L^i)v_t \tag{2}$$

Более лаконичную запись (2) можно получить, используя матричную нотацию:

$$\mathrm{A}(L)(1-L)^d X_t = \mathrm{B}(L)v_t,$$

где $L^d X_t = X_{t-d}$, $A(L) = 1-\alpha_1 L - \ldots - \alpha_p L^p$, $\mathrm{B}(L) = 1-\beta_1 L - \ldots - \beta_q L^q$.

В качестве наглядного примера можно привести модель $ARIMA(0, d, 0)$, для которой соотношение (2) будет иметь вид $(1-L)^d X_t = v_t$ или, в развернутом виде

$$X_t - dX_t + \frac{d(d-1)}{2} X_{t-2} - \ldots = v_t$$

Как указывается в [2], модель (1, 2) соответствует процессу на выходе неустойчивого (при $d \neq 0$) линейного фильтра, на входе которого – белый шум v_t. На возможность применения данной модели в задачах моделирования динамики котировок указано, например, в [5].

Следующим обобщением данной линии моделей временных рядов является модель *авторегрессии-дробно интегрированного скользящего среднего* (*Autoregressive Fractionally Integrated Moving Average*, $ARFIMA(p, d, q)$), допускающая использование дробных значений дифференциального параметра d [18, 19, 26, 27]. При этом в дробной модели для оператора лага используется биномиальное представление вида

$$(1-L)^d = \sum_{k=1}^{\infty} C_d^k (-L)^k = 1 - dL + \frac{d(d-1)}{2} L^2 - \ldots,$$

что рассматривается как математическая интерпретация долговременной памяти модели. Модель $ARFIMA(p, d, q)$ описывается тем же соотношением (2), что и модель $ARIMA(p, d, q))$, однако величина d, как уже указывалось, является дробной. При этом сторонники применения этой модели утверждают, что она позволяет воспроизводить многие устойчивые статистические закономерности, присущие реальному рынку, такие как длинная память, кластеризация волатильности, негауссовость распределения доходностей, информационные послешоки и пр. [10, 11, 17].

В случае если модели временного ряда типа $ARMA$ учитывают сезонные или периодические эффекты, то соответствующие обобщения образуют класс сезонных (*seasonal*) моделей типа SAR, $SARMA$ или $SARIMA$ [23]. Поиск сезонных процессов в хаотических рядах наблюдений, содержащих явно выраженную непериодическую компоненту, представляется мало перспективным занятием. Тем не менее, экономисты и трейдеры продолжают поиск различного рода «волн» на основании каких-то общих представлений о сезонности в экономике, политике, астрономических и астрологических закономерностях etc.

Расширение этого же класса моделей на случай многомерных (*multivariate*) временных рядов [15] приводит к классу векторных моделей типа *VAR (Vector AR), VMA (Vector MA)* или *VARMA (Vector ARMA).* Многомерное направление моделирования позволяет осуществлять совместное исследование динамики различных инструментов с учетом их взаимосвязи, однако алгоритмы идентификации и прогноза становятся существенно более сложными.

В ситуациях, связанных с *нелинейными* (nonlinear) моделями временных рядов, используется спецификация вида NAR, NMA или $NARMA$. Для проверки исходного ряда на нелинейность можно использовать тест *BDS* (*Brock-Dechert-Scheinkman*), разработанный для эконометрических моделей [22].

Говорить о повышенной эффективности нелинейных моделей можно только в контексте конкретной задачи с определенными критериями качества ее решения и совокупностью ограничений. В условиях таких ограничений иногда удается сформулировать условие сходимости и состоятельности ожидаемых решений, основанных на нелинейных моделях. В большинстве случаев незначительные положительные эффекты от применения нелинейных схем не окупают резкое усложнение вычислительных схем и возникновение дополнительной неопределенности относительно их состоятельности в изменяющихся условиях функционирования рынков.

Тем не менее, нелинейность многолика, и в ряде случаев ее введение представляется вполне оправданным. В частности, большой класс моделей временных рядов связан с попыткой имитации гетероскедастичности, т.е. изменения во времени параметра волатильности котировки. В частности модель *ARCH* (*Autoregressive Conditional Heteroskedasticity*), рассмотренная в [25], использует стандартное представление случайной составляющей наблюдений в виде $v_t = \sigma_t \cdot n_t(0,1)$, где $n_t(0,1)$ – ряд наблюдений, моделируемый независимой нормально распределенной случайной величиной с параметрами (0, 1), а σ_t задается соотношением

$$\sigma_t^2 = \gamma_0 + \sum_{i=1}^{m} \gamma_i v_{t-i}^2, \quad \gamma_0 > 0, \quad \gamma_i \geq 0, \quad \forall i > 0.$$

Развитие этого направления привело к возникновению таких моделей, как *GARCH (Generalized Autoregressive Conditional Heteroskedasticity), NGARCH (Nonlinear GARCH), IGARCH (Integrated Generalized Autoregressive Conditional Heteroskedasticity), EGARCH (Exponential General Autoregressive Conditional Heteroskedastic), GARCH-M*

(GARCH-in-mean), QGARCH (Quadratic GARCH), GJR-GARCH (Glosten-Jagannathan-Runkle GARCH), с которыми можно познакомиться, например, в [7] или на сайте [20].

Другой пример нелинейных моделей связан с учетом группы детерминированных экзогенных факторов $U_{<t,\, t-s>} = (u_t, \ldots, u_{t-s})$ [29]. Общий вид нелинейной модели авторегрессии с экзогенными факторами (*nonlinear autoregressive exogenous model, NARX*) можно записать в виде $X_t = f(X_{<t,\, t-p>},\ U_{<t,\, t-s>}) + v_t$, где $X_{<t,\, t-p>} = (x_t, \ldots, x_{t-p})$, f – нелинейная функция времени, в качестве которой часто используют полиномы. Аналогично и другие базовые модели временных рядов могут быть расширены с введением экзогенных воздействий. При этом в аббревиатуре этих моделей в конце добавляется символ *«X»*.

В заключение обзора нелинейных модификаций следует упомянуть процессы с двойной стохастичностью, когда один или несколько параметров в исходной модели $X_t = f(X_{<t,\, t-p>}) + v_t$ сами являются случайными величинами. Для одномерной задачи данный подход известен как концепция смешанных распределений (*compounded distributions*). Иногда его интерпретируют как наличие базового параметрического временного ряда X_t и возмущающего случайного процесса V_t, воздействующего на отдельные параметры базового ряда X_t. Данная технология моделирования аналогична моделям со скрытыми (латентными) переменными [24].

Приведенный обзор представляет собой неполный перечень основных направления в моделировании динамики котировок биржевых активов. Возникает вопрос об эффективности использования тех или иных моделей в задачах, ориентированных на управление торговыми операциями.

4. Критический анализ. Математические модели сами по себе не нужны. Разумеется, существует «чистая» математика, допускающая *l'art pour l'art* . Однако при решении прикладных задач разработка математической модели полностью определяется практическими целями и задачами. Это – банальное утверждение. Тем не менее, разработка сложных математических моделей, допускающих имитацию траекторий котировок, в ряде случаев превращается в моделирование ради самого моделирования.

Зачем нужно моделировать котировки?

Прежде всего, для создания полигона данных, необходимого для тестирования торговых стратегий, советников, торговых роботов и других программно-алгоритмических средств, ориентированных на управление торговыми операциями. Однако на сайтах брокерских компаний и трейдерских центров на открытом доступе лежат многолетние архивы рядов котировок основных финансовых инструментов. По-видимому, трудно усомниться в том, что ни одна модель не будет отражать динамику котировок лучше, чем сами котировки.

Вторая задача моделирования котировок состоит в использовании модели в качестве базового элемента торговой стратегии. При этом в качестве модели выступает некоторая системная составляющая, представляющей собой сглаженную модель наблюдений и определяющая основные тенденции развития ситуаций. В этом случае формируемая модель является базовой структурой, согласованной с выбранной торговой стратегией, а эффективность управляющих торговых решений будет определяться качеством восстановления системной составляющей наблюдений. При такой постановке задачи вопрос о степени адекватности математической модели и наблюдаемого процесса, в традиционной постановке, основанной на мерах близости, теряет смысл. Модель, формируемая как системная составляющая динамики котировок, может быть далека от подобия рядам наблюдений. Ее структура должна определяться лишь терминальными показателями эффективности торговой стратегии и соответствующим ей торговым операциям [8].

Заметим, что простейшее применение модели в качестве экстраполяционной прогностической схемы, как справедливо указывают некоторые эксперты [16], не позволяет построить эффективную торговую стратегию. Это связано с существенной нестационарностью случайной составляющей динамики котировок, обусловленной, как будет показано ниже, наличием хаотической компоненты рядов наблюдений.

В качестве третьей задачи моделирования можно выделить задачу идентификации статистических свойств «квазистационарных» участков случайной составляющей динамики котировок. Наличие и знание систематических свойств на определенном участке траектории позволяет выбрать из предварительно сформированной базы знаний вариант торговой стратегии, согласованный с текущей динамикой котировок и, как следствие, обеспечивающий эффективную торговлю. По существу речь идет о структурной (выбор торговой стратегии) и/или параметрической (определение оптимальных параметров торговой стратегии) адаптации системы управления торговыми операциями. Очевидно, что такой подход имеет смысл в ситуациях, когда изменение статистических свойств в целом нестационарного процесса происходит достаточно инерционно. Иначе говоря, должна выполняться гипотеза о том, что частота значимых изменений в динамике котировок позволяет осуществлять статистический анализ динамических свойств, формировать модель, отвечающую результатам анализа одной из торговых стратегий, и замыкать контур управления, т.е. реализовывать адаптивный вариант торговых операций. Однако возможность применения такой гипотезы на практике, даже с учетом современного уровня быстродействия компьютеров, не является очевидной. По существу, речь идет об установлении принципиальной возможности, а в случае положительного решения – об определении границ допустимости адаптивного (или оптимизационного) подхода в условиях нестационарной динамики.

Таким образом, традиционный перечень прикладных задач, реализуемых на основе математического моделирования, в отношении рассматриваемой задачи управления торговыми операциями можно разделить на два класса: задачи, для которых моделирование не нужно, и задачи, для которых моделирование не позволяет получить эффективные торговые решения в силу нестационарности и хаотичности наблюдаемых процессов.

Тем не менее, гипотеза о принципиальной возможности построения эффективной стратегии имеет право на существование, о чем свидетельствует успешная работа. Возможным вариантом реализации такой стратегии являются робастифицированные алгоритмы, обладающие пониженной чувствительностью к вариациям статистической структуры случайной составляющей динамики котировок [9, 16]. При этом степень эффективности таких алгоритмов существенно повышается при наличии (и соответствующем использовании) каких-либо ограничений, сужающих спектр вариаций динамических и статистических характеристик наблюдаемых процессов. Исследования задачи выявления таких ограничений связаны с применением апостериорных моделей, базирующихся на исследовании больших (и очень больших) массивов ретроспективных данных. Основная цель этих моделей в этом случае состоит не в генерации полигона тестовых данных и не в попытке использовать их в контуре управления торговыми операциями, а в формировании общих представлений о структуре и особенностях динамических процессов, протекающих на торговых площадках. Данный вопрос рассмотрен в следующем пункте настоящей статьи.

5. Моделирование: декомпозиция и анализ. В традиционной теории динамических систем [14] простейшая модель прямых наблюдений обычно задается соотношением

$$z_k = x_k + v_k, \quad k = 1,\ldots,N \tag{3}$$

где $x_k, \quad k = 1,\ldots,N$ – детерминированный, но неизвестный процесс, отражающий истинную динамику объекта анализа, $v_k, \quad k = 1,\ldots,N$ — некоторая шумовая компонента, определяемая погрешностями наблюдений. В более общем случае, основанном на байесово-калмановской парадигме [4], траектория движения описывается случайным процессом $x_k = \Phi_{k/k-1}x_k + w_k, \quad k = 1,\ldots,N$, где $\Phi_{k/k-1}$ — переходная матрица, а $w_k, k = 1,\ldots,N$ – случайная составляющая, образующая так называемые шумы системы. Как правило, шумы системы моделируются стационарным гауссовским процессом. Задача фильтрации в этом случае состоит в выделении системного процесса $\hat{x}_k, k = 1,\ldots,N$, представляющего собой условное среднее наблюдаемого случайного процесса.

Очевидно, что такая трактовка компонент наблюдений совершенно непригодна для представления рядов котировок. В данной задаче наблюдаемый процесс полностью (с точностью до выбранного разряда округления в процессе оцифровки) соответствует истинной динамике котировок $z_k = x_k, \quad k = 1,\ldots,N$, а шумов наблюдений $v_k, \quad k = 1,\ldots,N$ практически нет. «Шумовая» составляющая процесса $w_k, \quad k = 1,\ldots,N$ также не может быть механически отброшена путем сглаживания, т.к. ее значения содержательно неотделимы от ряда котировок. Поэтому модифицированное двухкомпонентное аддитивное представление изучаемого процесса $x_k = y_k + v_k, \quad k = 1,\ldots,N$ в данном классе прикладных задач требует качественно иной интерпретации. Здесь под $y_k, \quad k = 1,\ldots,N$ понимается сглаженная квазирегулярная системная составляющая, используемая для определения тренда, прогноза или непосредственно в процессе формирования решений, а $v_k, \quad k = 1,\ldots,N$ — случайная составляющая, образованная невязками между наблюдениями $z_k = x_k, k = 1,\ldots,N$ и процессом $y_k = \hat{x}_k, k = 1,\ldots,N$, полученным в результатами апостериорной обработки.

Заметим, что при таком определении возникает принципиальная неопределенность, связанная с самим понятием системной компоненты. По существу, разделение наблюдаемого ряда на системный процесс и помеховую составляющую является неоднозначным и требует дополнительного определения, связанного с такими субъективными факторами, как выбор стратегии игры и критерия ее эффективности. В этом случае на процесс идентификации системной составляющей определяющим образом влияют такие экзогенные факторы, как стратегические предпочтения участников торговых операций [8].

По существу, успех или проигрыш в торговых операциях, так или иначе, связан с выбором торговой стратегии, формирующей управляющие решения на основе представлений о последующих изменениях значений котировок торговых активов. Принято различать краткосрочные стратегии, ориентированные на торговые операции в течение дневной сессии, среднесрочные операции (от нескольких дней до нескольких месяцев) и долгосрочные (годовые или многолетние) операции, близкие, по сути, к инвестиционным процессам. При этом очевидно, что динамические тренды, определяющие управляющие решения, будут существенно разными. Отсюда непосредственно следует, что и определение системной составляющей будет разным. В частности, для долгосрочных торговых операций практически не существенны внутрисуточные колебания котировок, полностью определяющие результативность краткосрочных суточных спекуляций.

Таким образом, разделение общей динамики котировок на системную и случайную носит условный характер, определяется внешними факторами и требует уточнения при решении каждой конкретной задачи. Это, в свою очередь, означает, что «случайная» компонента динамики, являющаяся шумовой по отношению к выбранной «системной компоненте», может быть совершенно неслучайной с точки зрения традиционных статистических критериев случайности и независимости рядов наблюдений.

Иными словами, шумовая компонента, которую необходимо отфильтровать для выявления принципиально важной для торговли системной составляющей динамики котировок, по своей статистической природе не должна являться и не является «шумовой» в традиционном (для вероятностно-статистической парадигмы) смысле этого слова. В частности, случайная составляющая не является стационарным и нормальным процессом и, как правило, содержит свою системную составляющую, определяемую условным средним.

Важно отметить, что формализованную селекцию рядов наблюдений на системную и случайную можно осуществить, проводя последовательную идентификацию квазисистемных компонент до тех пор, пока оставшиеся невязки не превратятся в стационарный шумовой процесс. При таком подходе оказывается целесообразным представить исходный процесс в виде аддитивной трехкомпонентной модели наблюдения, включающей в себя системную компоненту, используемую в процессе формирования торговых решений, квазисистемную компоненту с динамикой более высокого порядка и чисто случайную составляющую, образующую стационарный временной ряд. Однако реализация подобного представления на практике связана со специфическими трудностями, характерными для процессов торговой динамики. В частности, когда квазисистемная компонента представляет собой колебательный непериодический процесс (в синергетике такие процессы называют хаотическими), традиционную идентификацию путем полиномиальной подгонки (или в каком-то другом ортогональном базисе) реализовать невозможно. В этом случае можно использовать последовательные вычислительные схемы динамической фильтрации.

Рассмотрим вопрос о фильтрации случайной погрешности в процессе последовательной подгонки рядов наблюдений на суточном интервале наблюдения. С этой целью будем последовательно повышать степень аппроксимирующего полинома, имитирующего системную составляющую, до такого уровня, пока статистические критерии согласия не подтвердят случайный характер соответствующих невязок между моделью и реальными наблюдениями.

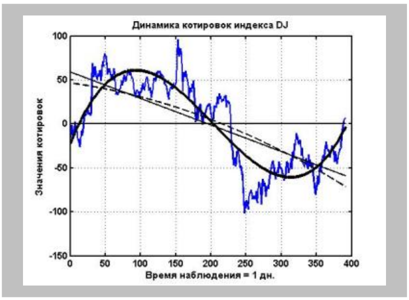

Рис. 2. Пример 1 динамики изменения котировок индекса DJ и его полиномиальных аппроксимаций в течение одной торговой сессии.

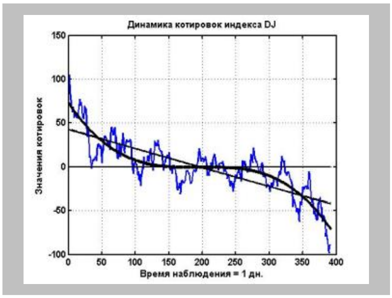

Рис. 3. Пример 2 динамики изменения котировок индекса DJ и его полиномиальных аппроксимаций в течение одной торговой сессии.

На рис. 2 и 3 приведены графики изменения котировок индекса DJ и его полиномиальных аппроксимаций (1-й, 2-й и 3-й степени) в течение двух однодневных торговых сессий, а на рис. 4 и 5 – соответствующие им невязки аппроксимаций.

Вопрос о соответствии рядов невязок аппроксимаций условию независимости и стационарности решается на основе применения известных статистических критериев [1]. Для проверки гипотезы H_0 о стохастической независимости ряда наблюдений используем медианный критерий серий. Для выбранного примера при N=392 (число минутных отсчетов за время работы нью-йоркской биржи), критические значения, соответственно, равны $\vartheta^* = 177.1$ $\tau^*_{\max} = 8{,}6$.

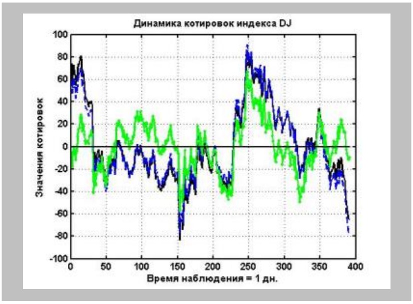

Рис. 4. Невязки полиномиальных аппроксимаций для примера на рис. 2.

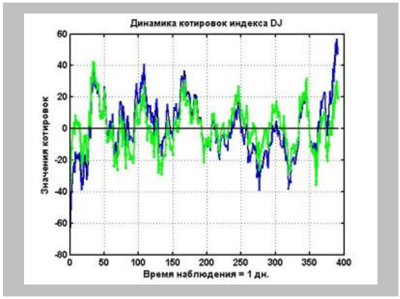

Рис. 5. Невязки полиномиальных аппроксимаций для примера на рис. 3.

Соответствующие значения решающих статистик и решение о принятии гипотезы H_0 о независимости рядов ошибок регрессии для трех торговых сессий на нью-йоркской бирже приведены в таблице 1.

Значение «1» в столбце H_0 означает, что гипотеза не противоречит данным, «0» – гипотеза H_0 отвергается.

Таблица 1. Анализ независимости на основе медианного критерия серий.

	1 сессия			2 сессия			3 сессия		
m	ϑ	$\tau_{\max}$	H_0	ϑ	$\tau_{\max}$	H_0	ϑ	$\tau_{\max}$	H_0
1	194	7	1	193	10	0	198	8	1
2	224	7	1	197	7	1	202	6	1
3	200	6	1	214	5	1	201	5	1
4	202	6	1	202	8	1	198	7	1
5	227	5	1	188	9	0	199	7	1

Из приведенных данных видно, что гипотеза о независимости исследуемых рядов наблюдений в большинстве случаев не отвергается.

При этом вполне достаточно использовать аппроксимирующие полиномы 2-3 порядка. Некоторое снижение качества оценивания при m>=4 объясняется плохой обусловленностью матрицы нормальных уравнений при высоких порядках аппроксимирующего полинома.

Следует заметить, что стационарность исследуемых процессов этот критерий проверить не может. Так, например, если искусственно увеличить амплитуду первой половины временного ряда, представленного на рис. 2, а вторую оставить без изменения, то параметры ϑ и $\tau_{\max}$ не изменятся, и гипотеза H_0 о независимости не будет отвергнута, а стационарность процесса очевидным образом выполняться не будет.

Проверка независимости и стационарности невязок оценивания не является самоцелью. Главной задачей является выявление системной составляющей, значимой для процесса принятия торговых решений. При этом нестационарность невязок, обусловленная вариациями параметров рассеяния, не является существенной. Главную роль в этом смысле играют вариации среднего, образующие тренды в виде смесей апериодических и колебательных процессов.

Для анализа независимости таких рядов наблюдений можно использовать критерий восходящих и нисходящих серий [1]. Для численной проверки данного теста используем те же данные для трех торговых сессий на нью-йоркской бирже, что и при составлении таблицы 1. Однако аппроксимации полиномами степени $p > 3$ использовать не будем, т.к. они приводят к вырожденности матрицы нормальных уравнений. Результат исследований приведен в таблице 2.

Таблица 2. Анализ независимости на основе критерия восходящих и нисходящих серий.

	1 сессия			2 сессия			3 сессия		
m	ϑ	$\tau_{\max}$	H_0	ϑ	$\tau_{\max}$	H_0	ϑ	$\tau_{\max}$	H_0
1	246	4	1	271	3	0	264	3	1
2	254	3	1	249	5	1	267	4	1
3	266	4	1	258	3	1	251	5	1

При этом критические значения равны, соответственно $\vartheta(n)^* = 245$ и $\tau(n)^* = 7$.

Таким образом, при выполнении традиционной технологии оценки параметров регрессионной модели по полной выборке ранее сделанное предположение о наличии третьей, квазисистемной компоненты в динамике исследуемого процесса оказалось несостоятельным. Оставшаяся после оценивания случайная составляющая представляет собой последовательность независимых отсчетов. Разумеется, это еще не означает, что данная составляющая образует стационарный процесс (в узком или широком смысле этого определения). В частности, из приведенных выше графиков на рис. 4 и 5 можно видеть наличие системного колебательного, но непериодического процесса.

Именно этот процесс, не улавливаемый традиционными критериями проверки независимости, вносит в исходный процесс элемент хаотической динамики, который не позволяет эффективно прогнозировать развитие ситуации на торговых площадках валютного, фондового и других рынков. Однако прежде, чем изучать данный вопрос аналитическими или числовыми методами, необходимо приблизить процесс оценивания состояния фондового рынка к практике.

Рассмотренный выше подход носит лишь теоретический характер. Это связано с тем, что процесс оценивания осуществлялся на основе совместной обработки данных, наблюдаемых в процессе всей торговой сессии. На практике решение формируется последовательно, по мере поступления данных с использованием выборки нарастающего объема, либо на основе скользящего окна наблюдения. В связи с этим проведем повторный анализ невязок регрессионных оценок с учетом последнего замечания. Для этого будем использовать схему последовательного регрессионного восстановления данных по выборке нарастающего объема и для различных размеров скользящего окна наблюдений, используемого для расчета коэффициента передачи *МНК (метод наименьших квадратов*) фильтра.

На рис. 6-7 представлены графики последовательной МНК аппроксимации для котировок индекса DJ в течение трех игровых дней (здесь и далее «вырезаны» интервалы времени, соответствующие времени, когда торговая площадка не работала) и одного игрового дня.

Нетрудно заметить, что попытка использования всей накопленной информации для оценки текущих значений существенно повышает инерционность процесса оценивания и приводит к значительным задержкам. Для скользящего окна наблюдения качество восстановления данных, как это видно из графиков на рис. 8 (скользящее окно с $\tau = 10$ одноминутных отсчетов) и 9 ($\tau = 20$ отсчетов), значительно выше.

Однако в обоих случаях основной проблемой качественного восстановления остается запаздывание аппроксимации относительно исходного процесса, что и приводит в конечном итоге к принципиальному снижению эффективности технического анализа, основанного на анализе трендов. При этом, естественно, рост окна повышает степень сглаживания и одновременно ведет к росту величины задержки.

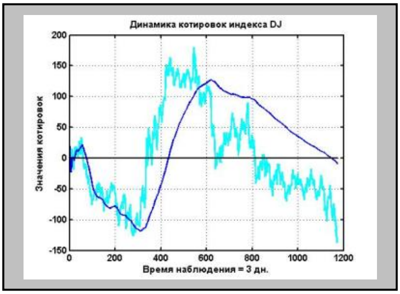

Рис. 6. Последовательная МНК аппроксимация для выборки нарастающего объема для котировок индекса DJ в течение трех игровых дней.

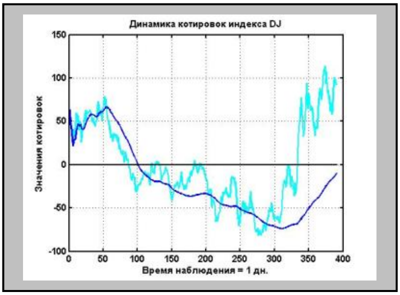

Рис. 7. Последовательная МНК аппроксимация для выборки нарастающего объема для котировок индекса DJ в течение одного игрового дня.

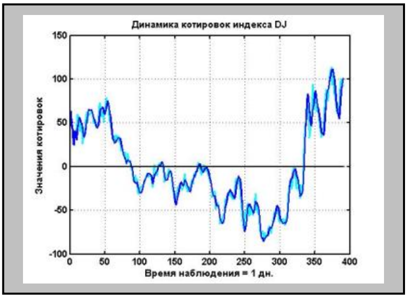

Рис. 8. Последовательная МНК аппроксимация для скользящей выборки объема $\tau = 10$ для котировок индекса DJ в течение одного игрового дня.

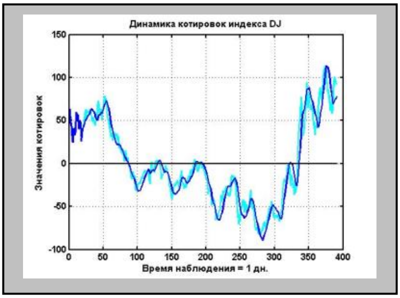

Рис. 9. Последовательная МНК аппроксимация для скользящей выборки объема $\tau = 20$ для котировок индекса DJ в течение одного игрового дня.

Сформируем для случая оценивания по скользящей выборке объема $\tau = 10:10:30$ однодневные последовательности погрешностей регрессионных оценок и проверим их на независимость с помощью приведенных выше критериев. Графики для этих погрешностей со скользящими окнами наблюдения $\tau = 10$ и $\tau = 30$ представлены на рис. 10. Можно видеть, что в приведенных обстоятельствах, имитирующих реальную последовательную обработку, а не идеалистическую схему совместной обработки данных, описанную в начале настоящего подраздела, имеется явно выраженная квазисистемная составляющая, о существовании которой выдвигалось предположение. При этом данная составляющая становится все более явно выраженной по мере роста скользящего окна наблюдения и обусловленной им задержки при формировании оценки текущего состояния индикатора фондового рынка.

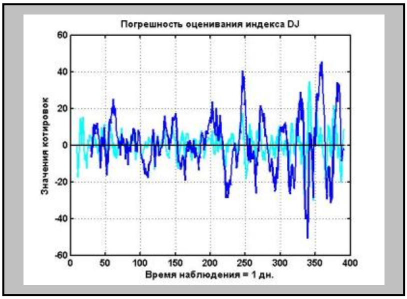

Рис. 10. Последовательности погрешностей регрессионных оценок со скользящими окнами наблюдения $\tau = 10$ и $\tau = 30$.

Для строгого подтверждения этого наблюдения проверим статистическую гипотезу о независимости выбранных рядов наблюдений с помощью критерия восходящих и нисходящих серий. Критические значения этого критерия для данного ряда равны $v^* = 244$; $\tau^*_{\max} = 7$. Численные значения решающих статистик равны для окна наблюдения $\tau = 10$, соответственно, $v = 191$ и $\tau_{\max} = 7$, для окна наблюдения $\tau = 30$ — $v = 180$ и $\tau_{\max} = 6$. Таким образом, в соответствии с выбранным критерием гипотеза о независимости ряда наблюдения отвергается, что и подтверждает наличие системной компоненты в ряду невязок.

Для повторного подтверждения данного вывода используем медианный критерий серий. Критические значения в этом случае равны, соответственно, $v^* = 109$ и $\tau^*_{\max} = 8.5$. Численные значения решающих статистик равны для окна наблюдения $\tau = 10$, соответственно, $v = 109$ и $\tau_{\max} = 8$, для окна наблюдения $\tau = 30$ – $v = 56$ и $\tau_{\max} = 25$.

Таким образом, в соответствии с медианным критерием данная гипотеза о независимости также отвергается в обоих случаях.

Принципиальным моментом для полученного вывода является то, что квазисистемная компонента практически не содержит явно выраженного линейного тренда, она носит колебательный непериодический характер. Динамику такого рода невозможно подогнать традиционными вычислительными схемами совместной обработки данных. Концепция адаптации также не допускает эффективной реализации в силу того, что непериодические колебания сами обусловлены задержкой оценивания. Если к ним добавить время, необходимое для идентификации изменений в динамике контролируемого процесса, то построенный контур адаптации будет подстраиваться к уже устаревшей динамике и, по существу, только ухудшать точность оценивания. Единственный конструктивный выход из создавшегося положения связан с попыткой применения робастифицированных схем оценивания, устойчивых к влиянию выявленной квазисистемной компоненты.

В заключение данной работы рассмотрим вопрос о возможности снижения уровня системной компоненты в ряду невязок за счет повышения степени аппроксимирующего полинома. В качестве примера численного анализа использовались те же параметры, что и в предыдущем случае. Разница состояла лишь в том, что была выбрана другая торговая сессия. Порядок аппроксимирующего полинома был выбран равным, соответственно, 2 (рис. 11) и 3 (рис. 12).

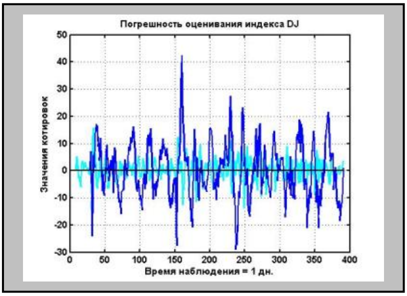

Рис. 11. Последовательности погрешностей регрессионных оценок со скользящими окнами наблюдения $\tau = 10$ и $\tau = 30$ и аппроксимирующим полиномом порядка 2.

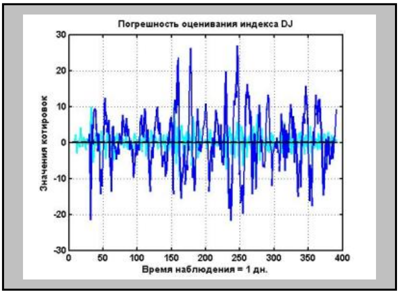

Рис. 12. Последовательности погрешностей регрессионных оценок со скользящими окнами наблюдения $\tau = 10$ и $\tau = 30$ и аппроксимирующим полиномом порядка 3.

Проверка независимости показала, что значения решающих статистик критерия восходящих и нисходящих серий в первом случае равны $\nu = 145$ и $\tau_{\max} = 8$, а во втором случае, $\nu = 88$ и tau_max=8. Численные значения решающих статистик при использовании аппроксимирующего полинома порядка 2 и медианного критерия серий равны для окна наблюдения $\tau = 10$, соответственно, $\nu = 145$ и $\tau_{\max} = 8$, для окна наблюдения $\tau = 30$ v=66 и tau_max=18. При пороговых значениях с уровнем доверия $p=(1-\alpha)=0.90-0.95$ $v^*=176.6$; $\tau^*_{\max}=8.56$ и, следовательно, гипотеза о стохастической независимости отвергается. Для окна наблюдения $\tau=30$ имеем значения решающих статистик $v=66$ и $\tau_{\max}=18$, при тех критических значениях гипотеза о стохастической независимости также отвергается.

Аналогичные результаты получаются при использовании аппроксимирующего полинома порядка 3.

Заметим, что помимо указанных критериев для проверки независимости можно использовать критерий Аббе [1]. Однако его использование предполагает предварительную проверку соответствия погрешностей наблюдений гауссовскому закону распределения. В данных условиях эта проверка приводит к отрицательному результату, что нарушает условие применимости данного критерия. Заметим, что при анализе нормальности использовался критерий, основанный на оценивании значений асимметрии и эксцесса, причем условие нормальности не выполнялось в силу большого значения оценки эксцесса.

Таким образом, из приведенных исследований можно сделать вывод, что исходный процесс изменения котировок может быть представлен в виде *трехкомпонентной (ТК) модели*

$$x_k = y_k + v_k + \xi_k, \quad k=1,\dots,N, \tag{4}$$

которая включает в себя:

- системную составляющую y_k, $k=1,\dots,N$, используемую при формировании торговых решений и образованную сложным сглаженным нелинейным процессом с явно выраженными трендами и колебаниями,

- квазисистемную помеховую компоненту v_k, $k=1,\dots,N$, представляющую собой несмещенный колебательный непериодический (хаотический) процесс,

- стационарный случайный процесс ξ_k, $k=1,\dots,N$ с распределением, как это будет показано ниже, стремящимся к гауссовскому закону.

Графически описанное представление можно проиллюстрировать кривыми, приведенными на рис. 13 … 15.

На рис. 13 показан график изменения системной компоненты, сформированной на основе МНК аппроксимации с вычислением коэффициента передачи на скользящем окне размером L=20, на фоне последовательности наблюдений за изменением котировок индекса DJ в течении трех торговых сессий.

Можно видеть, что системная компонента представляет собой непериодический колебательный процесс с явно выраженными трендовыми участками. При этом тренд выражен косвенно и, как правило, содержит так называемые участки «коррекции» обратной направленности.

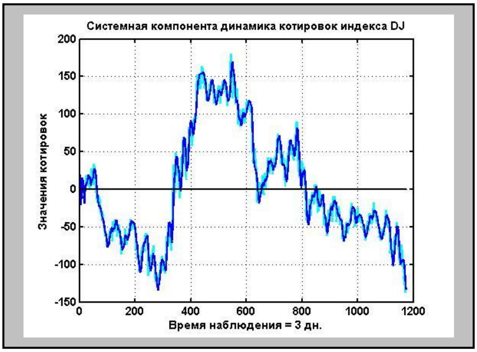

Рис. 13. Системная компонента на фоне последовательности наблюдений за изменением котировок индекса DJ.

Важно еще раз указать на субъективность процесса выделения системной компоненты, формируемой из общих соображений трейдера о степени сглаженности исходного ряда в интересах обеспечения процесса выработки эффективных торговых решений.

Очевидно, что данный параметр будет во многом определяться субъективно выбранной стратегией проведения торговых операций.

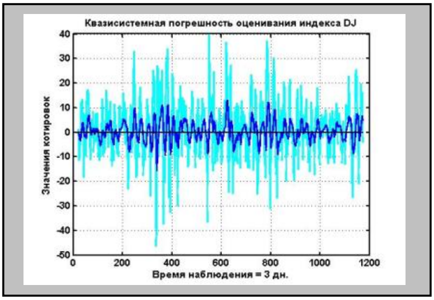

Рис. 14. Квазисистемная компонента модели на фоне невязок между процессом изменения котировок DJ и системной компонентой

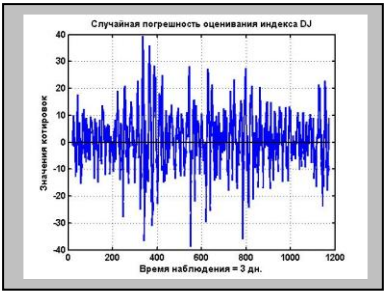

Рис. 15. Случайная компонента последовательности наблюдений изменения котировок DJ

Далее, на рис. 14. приведена квазисистемная компонента $v(t)$ модели (4) на фоне невязок между процессом изменения котировок DJ и выбранной системной компонентой. Для ее определения можно повторно использовать МНК аппроксимацию или какой-то более простой и оперативный способ оценки условного среднего. В частности, в данном случае, использовалась технология экспоненциального сглаживания с коэффициентом передачи $\alpha = 0.3$. Данный процесс также представляет собой вариант хаотической динамики, т.е. колебательный непериодический процесс, однако он уже не содержит явно выраженных трендов и достаточно хорошо центрирован.

На последнем в данной группе рисунке 15 представлена третья чисто случайная составляющая модели изменения котировок DJ (4), образованная разностью между невязками системной компоненты и значениями квазисистемной компоненты. Нетрудно увидеть, что приведенный процесс достаточно близок к стационарному гауссовскому шуму. Для проверки этого факта можно использовать указанные выше критерии.

6. Заключение. Следует заметить, что на практике критерий нормальности может не выполняться, так как сглаживающая технология, использованная для идентификации квазирегулярной модели, приводит, как уже отмечалось выше, к возникновению заметного эксцесса. Тем не менее, приближенное описание вида распределения случайной составляющей гауссовской кривой для многих прикладных задач может оказаться вполне допустимым.

Главным выводом из приведенных исследований является заключение о целесообразности применения трехкомпонентной аддитивной модели динамики процессов изменения котировок (4), принципиально отличающейся от традиционной модели наблюдений (3) с аддитивными шумами:

- спецификой данных, полученных в процессе мониторинга за торговой ситуацией и практически не содержащих погрешностей наблюдений;

- наличием субъективной составляющей, формируемой в процессе выбора торговой стратегии, и вычислительной схемы формирования системной составляющей;

- наличием хаотической компоненты в форме колебательного непериодического процесса.

Литература

1. *Айвазян С. А., Мхитарян В. С.* Прикладная статистика и основы эконометрии: Учебник для вузов. М.: ЮНИТИ, 1998. 1022 с.
2. *Бокс Дж., Дженкинс Г.* Анализ временных рядов. Прогноз и управление / Пер. с англ. под ред. В.Ф. Писаренко. М.: Мир, 1994. 407 с.

3. *Ильинский А*. Случайное блуждание и ценообразование на биржевых рынках [электронный ресурс *http://www.spekulant.ru/archive/Sluchajnoebluzhdanie_i_cenoobrazovanie_na_birzhevyh_rynkah.html*].
4. *Калман Р., Фалб П., Арбиб М.* Очерки по математической теории систем / Пер. с англ. под ред. Я.З. Цыпкина. М.: Мир, 1971. 400 с.
5. *Канторович Г.Г.* Анализ временных рядов // Экономический журнал ВШЭ. 2002. №2. С. 251–273.
6. *Кляшторная О.В.* Модели активов. Лекция по курсу «Рынок ценных бумаг». М: МИЭМ, 2007. [электронный ресурс *http://www.* fem.miem.edu.ru /u4ebmateriali/ klashtornaa /modeli_aktivov.doc].
7. *Льюнг Л.* Идентификация систем. Теория для пользователя. / Пер. с англ. под ред. Я.З. Цыпкина. М.: Наука, 1991. 432с.
8. *Мусаев А.А*. Quod est veritas. Трансформация взглядов на системную составляющую наблюдаемого процесса // Труды СПИИРАН. 2010. Вып. 15. С. 53–74.
9. *Мусаев А.А.* Методы построения робастифицированных систем анализа торговых ситуаций // Труды СПИИРАН. 2010. Вып. 14. С. 187–215.
10. *Остапенко Е.С., Дунаева Т.А.* Прогнозирование временных рядов с долговременной памятью с помощью моделей ARFIMA // Еконімичний вісник НТТУ КПІ. 2010. С. 270–273.
11. *Перцовский О.Е.* Моделирование валютных рынков на основе процессов с длинной памятью: Препринт WP2/2004/03. М.: ГУ ВШЭ, 2003. 52 с.
12. *Петерс Э.* Хаос и порядок на рынках капитала. / Пер. с англ. под ред. А. Н. Романова. М.: Мир, 2000. 336 с.
13. *Прохоров Ю.В., Розанов Ю.А.* Теория вероятностей. Основные понятия. Предельные теоремы. Случайные процессы. М.: Наука, 1967. 496 с.
14. *Сейдж Э.П., Уайт Ч.С.* Оптимальное управление системами / Пер. с англ. под ред. Б.Р. Левина. М.: Радио и связь, 1982. 392 с.
15. *Хеннан Э.* Многомерные временные ряды / Пер. с англ. М.: Мир, 1974. 576 с.
16. *Чеботарев Ю.* Торговые роботы на российском фондовом рынке. М.: Омега-Л, 2006. 144 с.
17. *Яковлев А., Франгуриди Г.* Стохастический подход к решению задач алгоритмической торговли // Журнал D` . 2010. №16. С. 46–49.
18. Autoregressive-moving average model [электронный ресурс http://en.wikipedia.org/wiki/Autoregressive_moving_average_model].
19. Autoregressive fractionally integrated moving average [электронный ресурс http://en.wikipedia.org/wiki/Autoregressive_fractionally_integrated_moving_average].
20. Autoregressive conditional heteroskedasticity [электронный ресурс http://en. wikipedia.org/wiki/Autoregressive_conditional_heteroskedasticity].
21. *Bachelier L.* Theorie de la speculation // Annales scientifiques de l'E.N.S. 1900. Ser. 3. Tome 17. P. 21–86.
22. *Brock W.A., Scheinkman J.A., Dechert W.D., LeBaron B.* A Test for Independence based on the Correlation Dimension // *Econometric Reviews.* 1996. V.15. P. 197–235.
23. *Cox D. R. Some statistical methods connected with series of events //* J. R. Statist. Soc. Ser. B. 1955. V.17. P.129–164.
24. Doubly stochastic model [электронный ресурс http://en.wikipedia.org/wiki/Doubly_stochastic_model#cite_note-0#cite_note-0].
25. *Engle R.F., Ng. V.K.* Measuring and testing the impact of news on volatility // Journal of Finance. 1982. v.48 (5). P. 1749–1778.

26. *Granger C. W. J., Joyeux R.* An introduction to long-memory time series and fractional differencing // Journal of Time Series Analysis, 1980. V.1. P. 15–29.
27. *Hosking J. R. M.* Fractional differencing. // Biometrika. 1981. V. 68(1). P. 165–176.
28. *Ito K.* Essentials of stochastic processess. Providence: AMS, 2006. 170 p.
29. *Nelles O.* Nonlinear System Identification: From Classical Approaches to Neural Networks and Fuzzy Models Book Description / Berlin: Springer Berlin, 2000. 785 p.
30. *Samuelson P.* Foundations of Economic Analysis / Cambridge, MA: Harvard University Press, 1947. 850 p.
31. *Time series.* [электронный ресурс http://en.wikipedia.org/wiki/ Time _series].
32. *Time Series With Long Memory.* Advanced Texts in Econometrics *// Ed. by* P. M. Robinson. Oxford: Oxford University Press, 2003. 392 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 — 188. 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 — 188. 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 24.03.2011.

РЕФЕРАТ

Мусаев А.А. **Моделирование котировок торговых активов.**

Основной целью задачи моделирования является попытка создания прогностической модели, позволяющей предсказать значения котировок рыночных активов на определенный период времени. Однако все попытки пролонгации накопленного и оперативного опыта не привели к успеху. Тем не менее, попытки найти философский камень биржевой торговли продолжались и вели к построению все более и более сложных моделей.

В настоящей статье приведен краткий обзор эволюции моделей, имитирующих динамику котировок рыночных активов. В частности, выделено два основных направления. Первое направление связано с построением моделей, в которых котировки ведут себя подобно частицам в броуновском движении. Наиболее корректной моделью в этом случае являлся винеровский процесс. Альтернативное направление моделирование связано с использованием моделей временных рядов. Однако и в этом случае многообразные попытки усложнения базовых моделей типа AR или ARMA не увеличили адекватность модели и, тем более, не способствовали повышению качества прогноза и эффективности торговых решений.

В настоящей статье выдвигается предположение, что основные неудачи при решении данной задачи связаны с двумя системными причинами.

Первая причина состоит в том, что задача декомпозиции рядов наблюдений, необходимая для построения эффективных моделей, должна разделять системную и помеховую компоненты, причем системная компонента не должна отвечать каким-либо критериям математического или эконометрического подобия реальным процессам. Критерием качества формирования системной составляющей является терминальный критерий эффективности торговых операций. Вторая, не менее важная причина состоит в том, что помеховая составляющая, образованная разностью между рядами наблюдений и восстановленной системной составляющей, является крайне сложным нестационарным процессом, содержащим хаотическую компоненту.

В качестве иллюстрационной модели в статье предложена трехкомпонентная модель котировок, состоящая из системной компоненты, согласованной с выбранной торговой стратегией и ее параметрами, чисто случайной составляющей, близкой к белому шуму, и нестационарной компоненты с элементами хаотической динамики, борьба с которой и определяет успешность (или неуспешность) технического анализа.

SUMMARY

Musaev A.A. **Modeling of trading assets quotations.**

The main purpose of modeling is an attempt to create a forecasting model to predict the market assets quotes values. However all attempts for prolongation of accumulated and operational experience have not lead to success. Nevertheless, the attempts to find the stock trading Philosopher's Stone are continued and lead to formation more and more complex models.

This article gives a brief overview of the market quotations dynamics models evolution. In particular, it singled out two main approaches. The first approach is associated with the construction of models in which the quotes behave like particles in Brownian motion. The most correct model in this case is the Wiener process. Attempts to switch to the longer time periods led to the model based on the Ito process. Approach with addition artificial trend components produced Samuelson model.

An alternative direction of modeling involves the use of time series models. However, even in this case multiple attempts of complication of basic AR or ARMA models did not contribute to improving the precision of forecasting or trading decisions efficiency.

In this paper, an assumption that the major failure in solving this problem is connected with two systemic causes is made. The first reason is that the observations series decomposition task, which is very important for effective model formation, should be shared for the system and noise components. At that system component has not to be a response to any criteria of mathematical or econometric similarity to real quotation process. Quality criterion of forming of a system component must be the terminal criterion of trade operations efficiency. This implies the subjectivity of the modeling process (recovery system components), its dependence on exogenous factors, and relativity to hierarchically superior system (trade strategy and its parameters).

The second reason lies in the nature of noise component, formed as a difference between observations series and restored systematic component. The noise component is represented by extremely complicated time-dependent process that contain chaotic composite. Under the chaotic component we mean here a continuous non-periodic oscillatory process, essentially limiting the possibility of sequential optimization or adaptations.

As an illustrative example the three-component quotations of dynamic model are proposed in the article. This model consists of system component, concerted with the chosen trading strategy and its parameters, a purely random component, close to white noise, and transitional component, including chaotic dynamics. Control of third component determines the technical analysis of success (or failure).