ИПМ им.М.В.Келдыша РАН • Электронная библиотека

Препринты ИПМ • Препринт № 47 за 2008 г.

Орлов Ю.Н., Осминин К.П.

# Кинетические уравнения для прогнозирования нестационарных временных рядов

***Рекомендуемая форма библиографической ссылки:*** Орлов Ю.Н., Осминин К.П. Кинетические уравнения для прогнозирования нестационарных временных рядов // Препринты ИПМ им. М.В.Келдыша. 2008. № 47. 30 с. URL: http://library.keldysh.ru/preprint.asp?id=2008-47

РОССИЙСКАЯ АКАДЕМИЯ НАУК

Ордена Ленина

ИНСТИТУТ ПРИКЛАДНОЙ МАТЕМАТИКИ

им. М.В. Келдыша

Ю.Н. Орлов, К.П. Осминин

# КИНЕТИЧЕСКИЕ УРАВНЕНИЯ ДЛЯ ПРОГНОЗИРОВАНИЯ НЕСТАЦИОНАРНЫХ ВРЕМЕННЫХ РЯДОВ

Москва, 2008

Ю.Н. Орлов, К.П. Осминин

Кинетические уравнения для прогнозирования нестационарных временных рядов

В работе излагается методика прогнозирования нестационарных временных рядов с помощью кинетических уравнений, которые записываются для эмпирической выборочной функции распределения. Объем выборки зависит от требуемой точности прогноза на заданном промежутке времени и определяется соответствующим функционалом. В зависимости от статистики т.н. горизонтного ряда для прогноза выбирается либо уравнение Лиувилля, либо Фоккера-Планка. Рассмотрены примеры прогнозирования некоторых временных рядов, возникающих на рынках ценных бумаг.

Yu.N. Orlov, K.P. Osminin

Kinetic equations for forecasting of non-stationary time-series

In this work the method of forecasting of non-stationary time-series with the use of kinetic equations for empiric distribution function is proposed. The series volume depends on the accuracy for any given horizon of forecasting. In dependence of statistics properties of so-called horizon series the evolution Liouville equation or Fokker-Planck equation are used. As an example the time-series of stock-exchange is considered.

# Введение

В настоящей работе исследуется возможность применения кинетических уравнений к задаче прогнозирования нестационарных временных рядов. Такие ряды встречаются во многих практических задачах, связанных с обработкой данных и прогнозированием в медицине, геологии, биологии, экономике и других областях. Случайные временные ряды (но, как правило, стационарные) возникают также и при численном решении некоторых систем обыкновенных дифференциальных уравнений, когда при определенных значениях параметров дискретизации получается разбегание траекторий для сколь угодно близких начальных условий. Использование кинетических уравнений для описания прогнозируемого изменения выборочной функции распределения временного ряда является новым подходом, который для нестационарных случайных процессов имеет определенные преимущества по сравнению с традиционными адаптивными методами анализа данных.

Основная проблема при исследовании нестационарных рядов состоит в том, что любая конечная выборка не принадлежит генеральной совокупности, т.е. не существует стационарного распределения вероятности, реализацией которого является данный случайный процесс. В силу этого обстоятельства нет и теорем об асимптотической нормальности отклонений выборочных моментов от моментов генеральной совокупности. Но тогда с увеличением объема выборки не происходит уточнения статистических свойств временного ряда. Отсутствует также и такое традиционное (для стационарных процессов) свойство как асимптотическая несмещенность оценок выборочных моментов.

Напомним некоторые определения из теории анализа выборочных статистик, важные для понимания отличий подхода, предлагаемого в настоящей работе, от традиционных методов, развитых для стационарных временных рядов [1, 2].

Оценкой $\widetilde{\theta}_n$ величины $\theta$, получаемой в результате $n$ наблюдений над некоторой случайной величиной $\omega$, называется статистика, т.е. некоторая функция от результатов наблюдений, по значениям которой судят о значениях величины $\theta$. Такой величиной, например, являются моменты гипотетического стационарного распределения, которым, как предполагается, определяются вероятности осуществления изучаемых событий.

Оценка $\widetilde{\theta}_n$ величины $\theta$ называется несмещенной, если ее математическое ожидание равно оцениваемой величине: $M[\widetilde{\theta}_n] = \theta$. Это равенство означает, что если бы нам было известно распределение вероятностей, то, усреднив по этому распределению наблюдаемую эмпирическую статистику, мы получили бы среднее значение параметра $\theta$. Оценка называется асимптотически несмещенной, если ее математическое ожидание стремится к оцениваемой величине при увеличении объема выборки.

Оценка называется состоятельной, если она сходится по вероятности к оцениваемой величине: $\forall \varepsilon > 0 \quad P\left(\left|\widetilde{\theta}_n - \theta\right| \le \varepsilon\right) \to 1, n \to \infty$. Если оценка состоятельна, то увеличение объема выборки приводит к уточнению эмпирического значения изучаемой величины. Оценка называется суперсостоятельной, если указанная сходимость по вероятности имеет асимптотику, главный член которой убывает быстрее, чем $1/\sqrt{n}$, т.е. такая сходимость более быстрая, чем сходимость к нормальному распределению в рамках центральной предельной теоремы.

Несмещенная оценка $\widetilde{\theta}_n$ величины $\theta$ называется эффективной, если она имеет наименьшую дисперсию среди всех возможных несмещенных оценок величины $\theta$, вычисленных по выборкам одного и того же объема $n$.

Доверительным интервалом с доверительной вероятностью $p$ для, например, математического ожидания $\mu$, оценкой которого по выборке объема $n$ является выборочное среднее $\mu_n$, называется интервал с границами

$$\mu_n \pm q_{(p+1)/2} \frac{\sigma_n}{\sqrt{n}},$$

где

$$\mu_n = \frac{1}{n}\sum_{k=1}^{n} x(k), \ \sigma_n^2 = \frac{1}{n-1}\sum_{k=1}^{n} (x(k) - \mu_n)^2,$$

а величина $q_{(p+1)/2}$ является $\frac{p+1}{2}$-квантилем стандартного нормального распределения, т.е. такой величиной $q_\alpha$, что

$$\int_{-\infty}^{q_\alpha} f_{norm}(\xi) d\xi = \alpha = \frac{\Phi(q_\alpha) + 1}{2}\Big|_{\alpha=(p+1)/2},$$

причем

$$\Phi(q) = 2\int_0^q f_{norm}(\xi)d\xi$$

есть вероятность того, что случайная величина $\xi$ не превосходит по модулю значение $q$, а $f_{norm}(\xi) = \frac{1}{\sqrt{2\pi}}\exp\left(-\frac{\xi^2}{2}\right)$ есть плотность нормального распределения с нулевым средним и единичной дисперсией. Такой подход к оценке средних величин основан на том, что для стационарных процессов отклонения выборочных средних от среднего по генеральной совокупности распределены асимптотически нормально.

Отличие рассматриваемого нами случая от традиционного стационарного в том, что мы имеем дело с нестационарной выборкой, поэтому невозможно взять среднее по генеральной совокупности ввиду отсутствия таковой. Следовательно, нет понятий состоятельности, эффективности и несмещенности оценок. Поскольку мы не можем говорить об отклонениях выборочных оценок от средних по стационарному распределению, то для них отсутствуют также понятия доверительных интервалов в смысле определений, данных выше. Следовательно, чтобы формализовать задачу прогнозирования с заданной точностью для нестационарных процессов, необходимо ввести соответствующие определения, которые были бы практически полезны и имели бы корректное статистическое содержание.

В математическом плане задача прогноза вероятности того, что значение временного ряда будет лежать в заданных пределах, сводится к построению модели временной эволюции выборочной функции распределения (ВФР), построенной по выборке конечного объема, один или оба конца которой, возможно, скользящие.

В работах авторов [3, 4] предложена методика анализа нестационарных временных рядов, основанная на том факте, что для любого ряда можно подобрать выборку такого объема, что изменение построенной ВФР при сдвиге на любое конечное число шагов будет меньше любого наперед заданного числа. Это означает, что, задавая желательную точность анализа данных, можно определить промежуток времени (горизонт прогноза), на котором с указанной точностью справедливы методы, разработанные для стационарного случая. При этом важной частью методики является установление связи между объемом выборки, точностью прогноза и горизонтом прогноза. В настоящей работе показано, что этими величинами определяется также и тип моделей, которые предлагаются для описания эволюции ВФР.

## 1. Определение промежутка квазистационарности ВФР

Пусть $x(t)$ представляет значение случайной величины $x$ в дискретный момент времени $t$. Обозначим $f_T(x,t)$ ВФР, полученную по выборке $\{x(n)\}$ объема $T$, где натуральное число $n$ принадлежит скользящему окну $\Delta_T(t)=[t-T,t]$. Символ геометрического отрезка $\Delta_T(t)$ здесь означает набор соответствующих натуральных чисел. Среднее значение функции $a(x)$ от случайной величины $x$ по ВФР $f_T(x,t)$ будем обозначать угловыми скобками:

$$\langle a(x)\rangle_{T,t}=\int a(x)f_T(x,t)dx\,.$$

(1)

В тексте для краткости изложения мы иногда будем опускать указание зависимости от объема выборки $T$ и/или момента времени $t$, если это не будет препятствовать пониманию смысла предложения.

В каждый момент времени ВФР $f_T(x,t)$ представляет собой функцию одного аргумента $x$. Введем норму в пространстве функций распределения, являющихся суммируемыми по $x$ функциями, т.е. в пространстве $L_1$:

$$\|f(x,t)\|=\int|f(x,t)|dx\,.$$

(2)

*Определение 1. Расстоянием между двумя ВФР f и h называется величина*

$$\rho(f,h)=\|f(x,t_1)-h(x,t_2)\|=\int|f(x,t_1)-h(x,t_2)|dx\,.$$

(3)

В этом определении сравниваемые ВФР могут быть построены по выборкам разных объемов и/или быть отнесены к различным моментам времени.

*Определение 2. ВФР $f_T(x,t)$ временного ряда $x(t)$ будем называть $\theta$-$\varepsilon$-стационарной на временном промежутке $[0;\theta]$, если*

$$\forall\,\tau\in[0;\theta],\ \forall\,t\ge 0\quad\int|f_T(x,t+\tau)-f_T(x,t)|dx\le\varepsilon\,.$$

(4)

Это определение, как и расстояние (3), символически записано в виде интегрального критерия, который для ВФР фактически представляет собой конечную сумму на ограниченном промежутке изменения значений ряда.

Если неравенство (4) выполняется для всех $\theta$, то ВФР (и сам ряд) будем называть просто $\varepsilon$-стационарным. Если неравенство (4) выполнено лишь для конкретного значения $\tau$ (или для некоторого набора этих значений), то ряд и соответствующую ВФР будем называть *ограниченно $\tau$-$\varepsilon$-стационарными.*

Если исследуемый ряд таков, что промежуток изменения величины $x$ равномерно ограничен по времени, область интегрирования в (4) без ограничения общности можно считать отрезком $[-1;1]$. В этом случае из $\varepsilon$-стационарности распределения следует также и $\varepsilon$-стационарность его

моментов. Функционал

$$V_T(t,\tau) \equiv \rho\left(f_T(x,t+\tau), f_T(x,t)\right) = \int \left| f_T(x,t+\tau) - f_T(x,t) \right| dx$$

(5)

совместно с определением 2 играет роль критерия $\varepsilon$-стационарности распределения случайной величины $x$. Он показывает, насколько ВФР с одинаковым объемом статистической базы отличаются одна от другой в моменты времени, разделенные интервалом $\tau$.

В [3] показано, что из неотрицательности ВФР и ее нормированности на единицу в любой момент времени следует оценка

$$0 \le V_T(t,\tau) \le \min\left(2\tau / T; 2\right).$$

(6)

Неравенство (6) дает возможность сделать важный вывод о том, что при фиксированном $\tau$ функционал (5) равномерно ограничен по $t$. Поэтому $\forall\, \varepsilon > 0$ всегда можно подобрать такой объем выборки $T > 2\tau / \varepsilon$, что ВФР будет $\varepsilon$-стационарной. Таким образом, если нельзя сравнить выборочную функцию распределения со стационарным распределением, то можно добиться близости двух нестационарных выборочных распределений. Это позволит прогнозировать временной ряд с некоторой заданной точностью возможного изменения функции распределения его значений.

При увеличении точности в определении $\varepsilon$-стационарности, т.е. при уменьшении $\varepsilon$, объем выборки, при которой достигается условие $V_T(t,\tau) \le \varepsilon$, растет. В силу равномерной ограниченности для каждого момента времени $t$ существует такое минимальное значение $h\left(t,\tau;\varepsilon\right)$, что при всех $T \ge h(t,\tau,\varepsilon)$ и при всех $\tau' \le \tau$ значения функционала $V_T(t,\tau)$ не превосходят $\varepsilon$. Рассмотрим максимальное по периоду наблюдений значение статистики $h\left(t,\tau;\varepsilon\right)$

$$H(\tau;\varepsilon) = \max_t h\left(t,\tau;\varepsilon\right).$$

(7)

В силу оценки (6) имеем $H(\tau,\varepsilon) \le 2\tau / \varepsilon$.

Статистику $h\left(t,\tau;\varepsilon\right)$ будем называть горизонтным рядом для прогнозирования ряда $x(t)$. Ее анализ позволяет получить эмпирическую оценку максимального горизонта прогноза, внутри которого распределение остается $\varepsilon$-стационарным, а также оценку минимального объема выборки, необходимого для прогнозирования ВФР внутри заданного горизонта.

## 2. Точность прогнозирования ВФР и временного ряда

Вычисление текущих оценок статистических характеристик нестационарного процесса $x(t)$ приводит к появлению принципиально неустранимых погрешностей: погрешности, возникающей за счет конечности времени усреднения, т.е. вследствие недостаточной репрезентативности объема выборки, а также погрешности, возникающей за счет изменения

статистики на интервале усреднения. Будем называть погрешность (в смысле среднего квадратичного) оценки статистики за счет конечности промежутка $\Delta$ погрешностью первого типа и обозначать $\Sigma_1$, а погрешность за счет нестационарности статистики – погрешностью второго типа и обозначать $\Sigma_2$. Задача оптимизации прогнозного алгоритма состоит в минимизации функционала полной ошибки $\Sigma$ в оценке статистики. Предполагая, что процессы, приводящие к ошибкам первого и второго типов, независимы, в качестве функционала полной ошибки выбираем

$$\Sigma = \sqrt{\Sigma_1^2 + \Sigma_2^2}\,.$$

(8)

При прогнозировании возникает задача оптимизации объема $T$ выборки в терминах функционала ошибки (8). Эта задача состоит в определении такого объема $T_{opt}$, при котором $\Sigma(T_{opt}) = \min$, и построении соответствующего численного алгоритма нахождения $T_{opt}$ для заданного нестационарного временного ряда. Оптимизационная задача возникает в силу того, что ошибки $\Sigma_1(T)$ и $\Sigma_2(T)$ как функции объема выборки имеют различное поведение. Именно, чтобы уменьшить ошибку вследствие неполной статистической репрезентативности, следует увеличивать объем $T$, а для уменьшения влияния нестационарности на статистические характеристики временного ряда следует уменьшать объем выборки. Алгоритмические методы нахождения $T_{opt}$ описаны одним из авторов в [5].

Таким образом, формулировка условия близости двух ВФР объединяет две статистики данного временного ряда $x(t)$: статистику объема выборки $T_{opt}(t)$ и статистику горизонта прогноза $\tau(t)$. Именно, требуется определить $T_{opt}(t)$ так, чтобы ВФР, построенные по выборкам этих оптимальных объемов на промежутке $[t, t+\tau(t)]$, различались бы меньше, чем на заданное число, характеризующее ошибку прогноза.

*Определение 3. Ошибкой $\delta$ прогноза для временного ряда $x(i)$ будем называть среднеквадратичное отклонение прогнозных значений $\tilde{x}(i)$ от фактических $x(i)$ на промежутке горизонта прогноза:*

$$\delta = \sqrt{\frac{1}{\tau}\sum_{i=t+1}^{t+\tau}\delta_i^2}\,, \quad \delta_i = \tilde{x}(i) - x(i)\,.$$

(9)

*Определение 4. Ошибкой $\varepsilon(t)$ прогноза ВФР в момент времени t будем называть интегральное абсолютное отличие прогнозной ВФР $\tilde{f}(x,t)$ от фактической $f(x,t)$, построенных по выборкам равных объемов:*

$$\varepsilon(t) = \int_{-1}^{1} \left|\tilde{f}(x,t) - f(x,t)\right| dx\,.$$

(10)

Предположим, что задана допустимая ошибка прогноза $\delta$. Естественно, она не может быть меньше, чем корень из дисперсии прогнозной ВФР $\tilde{f}$. Эту дисперсию обозначим через $\tilde{\sigma}^2$:

$$\tilde{\sigma}^2(t) = \int_{-1}^{1} (x - \tilde{\mu}(t))^2 \tilde{f}(x,t)dx, \quad \tilde{\mu}(t) = \int_{-1}^{1} x\tilde{f}(x,t)dx.$$

(11)

Различие между прогнозным и фактическим средними выборочными величинами можно оценить из неравенства

$$|\tilde{\mu} - \mu| \le \int_{-1}^{1} |x| \left| \tilde{f} - f \right| dx \le \varepsilon.$$

(12)

Рассмотрим ошибку прогнозирования значения ряда в некоторый момент времени. Прогнозным значением ряда в момент времени $t$ естественно считать среднее значение $\tilde{\mu}(t)$. Формальной ошибкой такого прогноза является $\tilde{\sigma}(t)$. Однако при сравнении с фактом следует учесть, что фактическая ВФР изменилась по сравнению с прогнозной, так что оценка квадрата фактической ошибки прогноза составит

$$\delta^2 = \int_{-1}^{1} (x - \tilde{\bar{x}})^2 f(x,t)dx = \int_{-1}^{1} (x - \tilde{\bar{x}} + \bar{x} - \bar{x})^2 f(x,t)dx =$$

$$= \sigma^2 + (\bar{x} - \tilde{\bar{x}})^2 \le \sigma^2 + \varepsilon^2.$$

(13)

Минимизация верхней оценки ошибки прогноза, представляемой формулой (13), и есть задача минимизации ошибки (8). Идея такой минимизации состоит в следующем.

Введем плотность распределения $\psi_{\tau,\varepsilon}(T)$ значений горизонтного ряда $h(t,\tau;\varepsilon)$, т.е. определим вероятность того, что расстояние между двумя ВФР, построенных по выборке объема $T$ и сдвинутых по времени на фиксированный промежуток $\tau$, не превосходит $\varepsilon$ для всех $T' \ge T$. Для краткости зависимость от момента времени $t$ в аргументах $\psi_{\tau,\varepsilon}(T)$ опущена.

Эта плотность $\psi_{\tau,\varepsilon}(T)$ строится по имеющимся данным к текущему моменту времени $t$ следующим образом. Для каждого момента времени $t'$, $1 \le t' < t - \tau$, строятся ВФР по выборкам объемов $T = 1,2,...,t'$, после чего для каждой из этих ВФР вычисляется функционал $V_T(t',\tau)$. Промежуток значений $[0;2]$, принимаемых функционалом $V_T(t',\tau)$, разбивается при этом на некоторое количество $N$ отрезков, например, равномерно, так что правый конец $k$-го отрезка есть $v_k = 2k/N, k = 1,2,...,N$. Номер отрезка, фиксирующий заданную в (5) величину близости $\varepsilon$ между ВФР, есть $k_\varepsilon = [N\varepsilon/2]$. Промежуток $[1;t-\tau]$ возможных значений объемов выборок содержит $t - \tau$ целочисленных точек, так что плоскость $\{T \times V\}$ покрыта

$1 \times \frac{2}{N}$-сетью. Отдельную ячейку сети нумеруем двумя индексами *i,k*. Если значение $V_T(t',\tau)$ соответствует номерам *i,k* некоторой ячейки, помещаем в нее индикатор «1». Ячейки, номерам которых не отвечают значения функционала $V_T(t',\tau)$, заполняем нулями. Совокупность ячеек с фиксированными значениями *i* или *k* образует полосу. В каждый момент времени $t'$ *k*-ая полоса состоит, таким образом, из нулей и единиц. Значением $h_k(t')$ горизонтного ряда в *k*-ой полосе называется индекс *i* ячейки, следующей за ячейкой с наибольшим индексом, содержащей ноль. Если нулевая ячейка – последняя, то значение горизонтного ряда в этой полосе отсутствует. Плотность $\psi_{\tau,\varepsilon}(T)$ представляет собой выборочную функцию распределения временного ряда $h_{k_\varepsilon}(t')$, построенную по выборке объема $t-\tau$, доставляющей исходному ряду условие ограниченной τ-ε-стационарности (4) в определении 2.

Если требуется, чтобы было выполнено условие полной τ-ε-стационарности, то процедуру формирования членов ряда $h_{k_\varepsilon}(t')$ следует дополнить рассмотрением всех значений $\tau'$ от 1 до $\tau$. Именно, при каждом значении $\tau'$ рассматриваются величины $h_k(t',\tau')$, которые определяются по вышеописанному алгоритму. Затем находятся величины $h_{k_\varepsilon}(t') = \max\limits_{\tau' \le \tau} h_{k_\varepsilon}(t',\tau')$.

Определим интегральную функцию распределения горизонтного ряда

$$\Psi_{\tau,\varepsilon}(T) = \sum_{n=1}^{T} \psi_{\tau,\varepsilon}(k), \quad \Psi_{\tau,\varepsilon}(2\tau/\varepsilon) = 1. \tag{14}$$

Смысл функции (14) в том, что если вместо величины $H(\tau,\varepsilon)$ из (7), во многих случаях оказывающейся равной своему теоретическому максимуму $2\tau/\varepsilon$, взять некоторое меньшее значение $T^*$, то с вероятностью $\Psi_{\tau,\varepsilon}(T^*)$ будут выполнены условия соответственно ограниченной или полной τ-ε-стационарности.

Введенные понятия позволяют окончательно сформулировать задачу об оптимизации объема выборки для прогнозирования временного ряда на заданный горизонт $\tau$ с заданной точностью $\delta$ в смысле определения (9). Поскольку для τ-ε-стационарных рядов справедлива оценка (13) $\delta^2 \le \sigma^2 + \varepsilon^2$, то оптимальный объем для прогнозирования на горизонт $\tau$ при соблюдении условия τ-ε-стационарности ВФР определим следующим образом.

*Определение 5. Оптимальным объемом $T_{opt}$ для прогнозирования временного ряда $x(t)$ на горизонт $\tau$ называется такой объем выборки, при котором сумма относительной дисперсии ВФР, построенной по выборке этого*

*объема на текущий момент начала прогноза, и квадрата расстояния между двумя ВФР, сдвинутыми на интервал $\tau$, минимальна.*

Чтобы провести оптимизацию объема выборки согласно определению 5, положим сначала $T_{opt} = 2\tau/\varepsilon$. Рассмотрим зависимость эмпирической выборочной дисперсии от объема выборки, усредненной по промежутку времени, на котором распределение горизонтного ряда τ-ε-стационарно. Уменьшение оптимального объема выборки играет принципиальную роль при прогнозировании, поскольку выборочная дисперсия часто ведет себя немонотонно как функция объема выборки. Простая оценка оптимального объема дается тогда формулой

$$T_{1opt} = \max\{2\tau/\varepsilon, \arg\min\sigma(T)\}.$$

(15)

Более тонкая оптимизация основана на анализе суммы выборочной дисперсии и квадрата ожидаемого расстояния между двумя ВФР, сдвинутыми на промежуток $\tau$. Эта оптимизация использует локальные свойства плотности распределения горизонтного ряда. В частности, может оказаться, что максимальных значений объемов выборки, равных $2\tau/\varepsilon$, среди значений $h(t,\tau,\varepsilon)$ весьма мало, так что соответствующая вероятность $\psi_{\tau,\varepsilon}(2\tau/\varepsilon)$ близка к нулю. Выберем тогда некоторое значение $T^* < 2\tau/\varepsilon$. Для него функционал $V_{T*}(t,\tau)$ не превосходит $\varepsilon$ с вероятностью $\Psi_{\tau,\varepsilon}(T^*)$, определяемой формулой (14). Однако для выборок объемов от $T^*+1$ до $2\tau/\varepsilon$ функционал нормы может быть больше $\varepsilon$. Этим увеличением неточности в близости двух ВФР можно пренебречь, если $T^*$ есть $1-\varepsilon$-квантиль распределения $\psi_{\tau,\varepsilon}(T)$, т.е. такая величина, для которой

$$\Psi_{\tau,\varepsilon}(T^*) = 1-\varepsilon.$$

(16)

В качестве объема выборки $T^*$ берется ближайшее целое число к решению уравнения (16). Тогда ожидаемая неточность функционала (5) будет не больше, чем

$$\varepsilon' = \varepsilon\Psi_{\tau,\varepsilon}(T^*) + \frac{2\tau}{T^*}\left(1-\Psi_{\tau,\varepsilon}(T^*)\right) = \varepsilon\left(1+\frac{2\tau}{T^*}-\varepsilon\right) = \varepsilon + o(\varepsilon).$$

(17)

Последнее равенство в (17) следует из того, что при $\varepsilon \to 0$ $T^* \to 2\tau/\varepsilon$.

Итак, в результате получается уточненная оценка оптимального объема выборки при сохранении требования τ-ε-стационарности ВФР:

$$T_{2opt} = \max\{T^*, \arg\min\sigma(T)\}.$$

(18)

Наконец, при минимизации ошибки прогноза исходного временного ряда можно еще расширить допустимую неточность в прогнозе ВФР с тем, чтобы

уменьшить совокупную ошибку прогноза за счет возможного уменьшения выборочной дисперсии. Описанная методика определения оптимального объема выборки является пошаговой с шагом, равным горизонту прогноза τ.

*Определение 6. Временной ряд $x(t)$, для которого оптимальный объем выборки при прогнозировании на заданный промежуток времени $\tau$, является квазистационарным (т.е. τ-ε-стационарным) временным рядом, будем называть устойчиво прогнозируемым рядом.*

Из (6) следует, что в качестве максимальной оценки минимально допустимого объема выборки $H(\tau,\varepsilon)$ можно взять $2\tau/\varepsilon$, т.е. при желании каждый временной ряд можно считать устойчиво прогнозируемым в смысле определения 6. Оптимизация же прогноза ряда заключается в том, что в некоторых случаях суммарная ошибка (13) может быть меньше для меньших допустимых объемов выборки. Поэтому желательно, чтобы именно эти оптимальные выборки представляли собой квазистационарную статистику.

Развиваемая далее теория может применяться к различным временным рядам, но корректно обоснованной она является именно для устойчиво прогнозируемых рядов, поскольку при прогнозировании распределения на любой промежуток, как превышающий границы квазистационарности, так и находящийся внутри них, необходимо быть уверенным в том, что объем выборки сохраняет свойство своей оптимальности. Это означает, что оптимальный объем выборки, определяемый до начала прогнозирования, в последующем не должен иметь статистически значимого тренда в течение промежутка времени τ.

В заключение этого параграфа отметим, что моменты статистики горизонтного ряда $h(t,\tau,\varepsilon)$ являются локальными индикаторами статистических свойств ряда $x(t)$. Их изменение с течением времени будет свидетельствовать об изменении условий функционирования той физической системы, которая и порождает данный случайный процесс.

## 3. Выборочные моменты и их эволюция

В связи с тем, что для выборочных моментов нестационарных рядов отсутствуют утверждения об их состоятельности как оценок моментов соответствующих распределений, то нет необходимости определять их так, чтобы они давали несмещенную оценку, как в случае стационарных рядов. Мы будем работать с выборочными средними, не предполагая их асимптотической несмещенности, поскольку, как отмечалось во введении, асимптотическая сходимость не определена для наших процессов ни в слабом смысле, ни по вероятности.

Выборочный момент первого порядка по выборке на промежутке $\Delta_T(t)$ определяем как

$$\mu_T(t) = \frac{1}{T} \sum_{k=t-T+1}^{t} x(k).$$

(19)

Выборочные центральные моменты порядка $r \geq 2$ определяем по формуле

$$m_{r,T}(t) = \frac{1}{T} \sum_{k=t-T+1}^{t} \left(x(k) - \mu_T(t)\right)^r .$$

(20)

Рассмотрим, например, временные ряды, образованные дневной ценой закрытия акций компаний General Motors (GM), General Electric (GE) и Microsoft (MS) за 2004-2005 гг. по данным [6]. Результаты усреднения (19-20) по объему выборки в 900 значений представлены на Рис. 1-2.

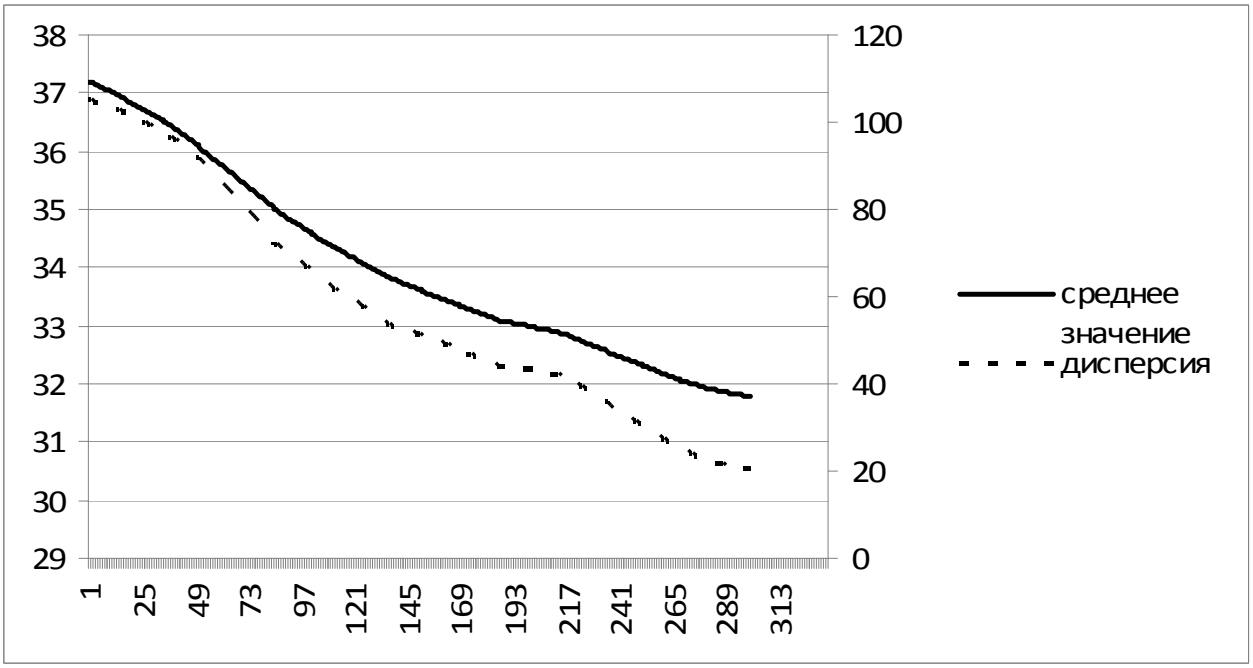


Рис. 1. Среднее скользящее значение (левая шкала) и дисперсия (правая шкала) цен на акции GE.

Здесь показан только фрагмент графика эволюции моментов ряда GE. Локальное уменьшение среднего значения и дисперсии (абсолютной, а не относительной) обусловлено снижением цен, а не стабилизацией случайного процесса. График приведен именно с целью демонстрации существенного непостоянства во времени моментов распределения. Похожим образом выглядит зависимость от времени средних выборочных значений ВФР для цен на акции компании MS (Рис. 2):

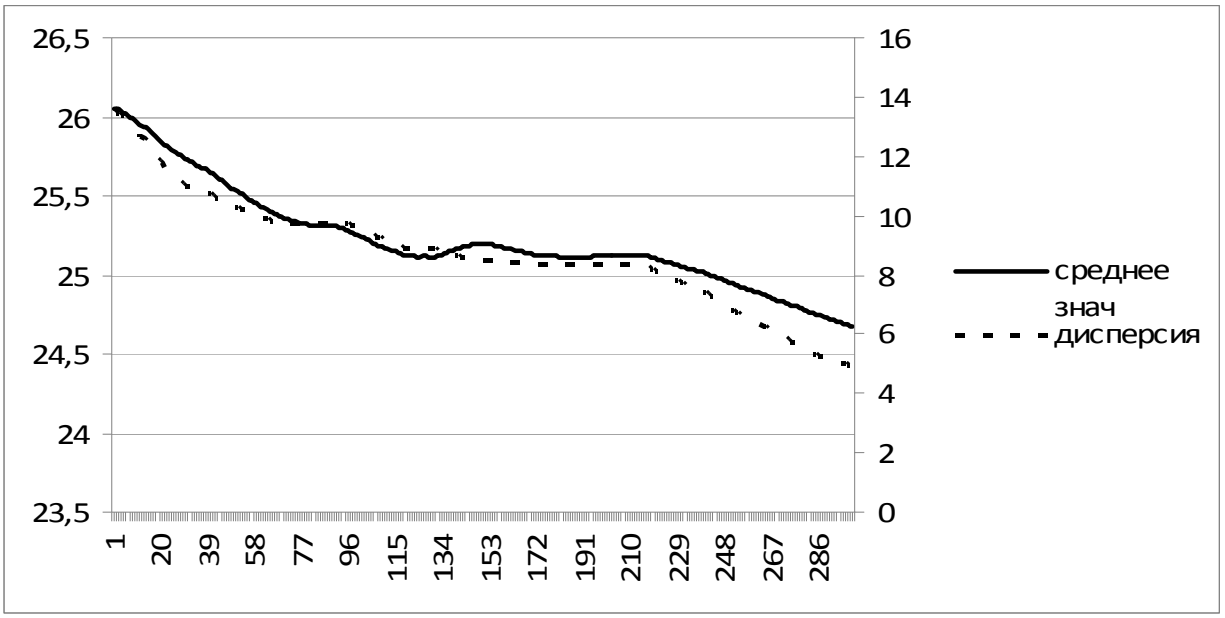


Рис. 2. Среднее скользящее значение (левая шкала) и дисперсия (правая шкала) цен на акции MS.

На Рис. 3 приведены для сравнения скользящие средние временного ряда цен на электроэнергию на НОРЭМ, исследованного в нашей предыдущей работе [3].

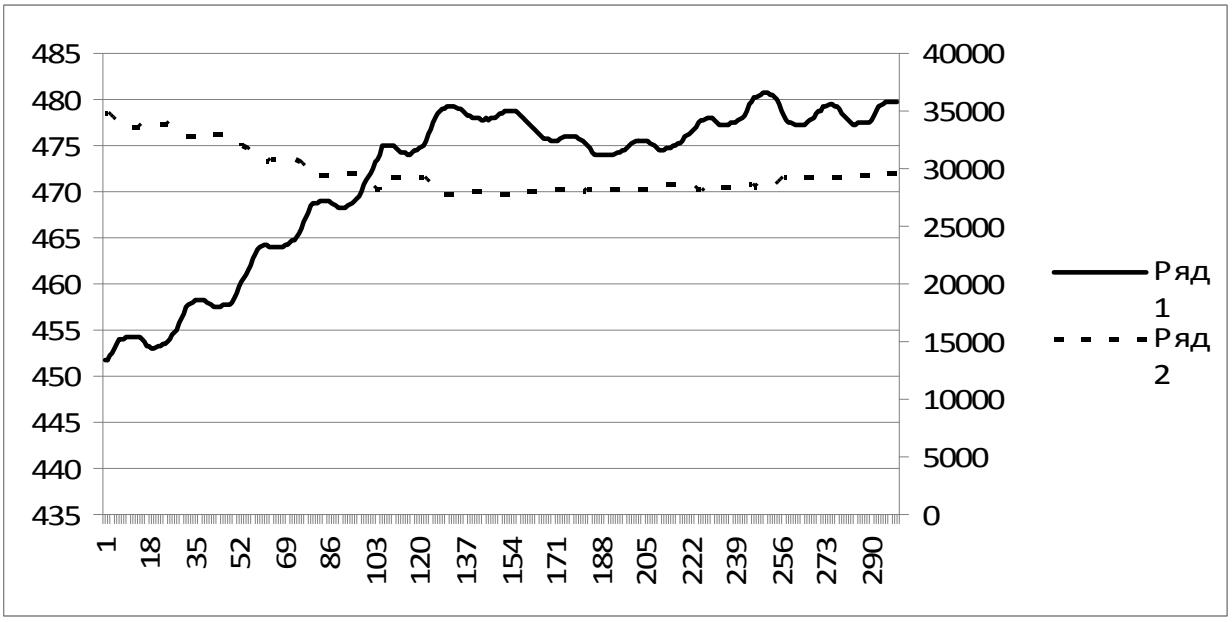


Рис. 3. Среднее скользящее значение (левая шкала, Ряд 1) и дисперсия (правая шкала, Ряд 2) цен на электроэнергию на НОРЭМ.

Поскольку приведенные эмпирические величины моментов имеют заметные временные тренды, то учет их в уравнении эволюции ВФР может быть полезным для повышения точности прогнозирования. В следующем параграфе мы рассмотрим с этой точки зрения различные модели эволюции ВФР.

## 4. Кинетические уравнения для ВФР

Исследуем вопрос о том, какими уравнениями имеет смысл описывать эволюцию эмпирической ВФР. С одной стороны, поскольку ВФР строится по конечной выборке данных, то такое уравнение по своему существу должно быть дискретным. С другой стороны, качественные черты модели более заметны в непрерывном описании, дискретная форма которого будет представлять собой численную схему расчетов в каждом конкретном случае. Поэтому сначала рассмотрим модели эволюции ВФР в дифференциальной форме.

Введем двумерную ВФР $F_T(x,\dot{x},t)$, случайных величин $x$ и $\dot{x}$ исходного ряда и ряда его производной, полученного взятием разностей $x(t)-x(t-1)$ в соседние моменты времени. Можно рассмотреть ВФР, зависящую и от большего числа производных, т.е. функцию вида $F(x,\dot{x},\ddot{x},...,t)$, но надо помнить об ограниченности объема выборки заданной величиной $T$. В силу конечности выборки невозможно эмпирически определить производную более высокого порядка, чем $T$-1. Более того, чтобы оперировать одинаковыми массивами данных по каждой из $r$ производных, необходимо иметь исходный ряд в количестве $T+r$ элементов.

Следовательно, чтобы можно было содержательно обсуждать задачу об эволюции ВФР $F_T(x,\dot{x},t)$, число доступных для анализа элементов ряда должно быть на 1 больше, чем в выборке. Такое включение естественным образом содержится в отрезке $[t-T,t]$: крайний левый элемент $x(t-T)$ окна $\Delta(t)$ будем считать виртуальным, т.е. не принадлежащим анализируемой выборке, в соответствии с определениями (19-20), а сами выборочные средние будем определять по данным $x(t-T+1)$, $x(t-T+2)$, …, $x(t)$. Тем самым мы будем считать, что в окне $\Delta(t)$ содержится не $T$+1, а ровно $T$ элементов.

Формально из сохранения во времени нормировки функции распределения следует уравнение Лиувилля

$$\frac{\partial F(\xi,t)}{\partial t}+div_{\xi}\left(\dot{\xi}F(\xi,t)\right)=0,\quad \xi=(x,\dot{x},...).$$

(21)

Обычно предполагается, что дивергенция в (21) берется в пространстве фазовых переменных, в котором введена динамически-инвариантная мера. Это означает, что в системе нет источников вероятности, т.е. вероятностное пространство является полным. Тогда

$$div_{\xi}\dot{\xi}=0,$$

после чего уравнение (21) может быть записано в виде

$$\frac{\partial F(\xi,t)}{\partial t}+\dot{\xi}\frac{\partial F(\xi,t)}{\partial \xi}=0\ .$$

Такая запись согласуется с представлениями статистической механики, в которой компоненты фазовой координаты $\xi$ считаются независимыми. Однако следует заметить, что в случае выборочного распределения конечного объема такой независимости может и не быть, т.к. корреляция между выборочными значениями $x$ и $\dot{x}$ для рядов, встречающихся, например, в эконометрике, как правило, отлична от нуля. В то же время «интеграл» от совместного эмпирического распределения двух величин по одной из них, т.е. сумма по всем соответствующим ее значениям, даст эмпирическое распределение по оставшейся переменной в силу самого правила построения совместной ВФР. Таким образом, на стадии вывода уравнения эволюции эмпирической ВФР методически правильно считать компоненты $\xi$ независимыми. То, что при одной реализации случайного процесса (т.е. на некоторой фазовой траектории) фазовые координаты коррелируют, не препятствует выбору удобного базиса в фазовом пространстве.

Имея целью создание прогнозной модели для «однофазной» ВФР $f(x,t)$, мы ограничимся рассмотрением кинетических уравнений для распределения одной переменной, хотя на стадии вывода соответствующего уравнения понадобится совместное распределение $F(x,\dot{x},t)$.

Поскольку

$$f(x,t)=\int F(x,\dot{x},t)d\dot{x},$$

(22)

то из (21) следует уравнение Лиувилля

$$\frac{\partial f}{\partial t}+\frac{\partial (uf)}{\partial x}=0,$$

(23)

где введена функция $u(x,t)$ по формуле

$$u(x,t)f(x,t)=\int \dot{x}F(x,\dot{x},t)d\dot{x}.$$

(24)

Введем средние по *x* величины в момент времени *t*:

$$\mu(t)=\langle x\rangle=\int xf(x,t)dx,\quad U(t)=\langle u(x,t)\rangle=\int u(x,t)f(x,t)dx.$$

(25)

Поскольку уравнение (23) записано относительно эмпирической ВФР, то будем его называть «эмпирическим» уравнением Лиувилля. Это означает, что скорость изменения ВФР $u(x,t)$ является параметром, определяемым по данной конкретной выборке в момент *t*, а не выводится из каких-либо уравнений, как в традиционной статистической механике.

Найдем скорость изменения скользящего среднего $\mu(t)$ в силу уравнения (23). Поскольку, как отмечалось выше, фактическая область изменения наблюдаемых величин *x* конечна, то мы не имеем возможности строить эмпирическую ВФР с бесконечной областью определения, т.к. все

интегралы, вычисляемые численно, – принципиально собственные. Тогда при реализации численной процедуры следует ввести ряд фиктивных промежутков изменения случайной величины $x$, количество которых на единицу больше порядка разностных операторов, фигурирующих в модели. В этих фиктивных промежутках значения ВФР и ее производных равны нулю. Удобно считать эти граничные точки без ограничения общности точками $\pm 1$. Тогда из (24) и (25) получаем интегрированием по частям эмпирическое уравнение эволюции первого момента ВФР в виде

$$\frac{d\mu}{dt} = \int x \frac{\partial f}{\partial t} = -\int x \frac{\partial uf}{\partial x} = \int uf dx = U .$$

(26)

Это уравнение означает, что изменение со временем среднего значения ряда равно средней скорости изменения значений ряда по скользящей выборке, что совпадает с представлениями математической статистики в дискретном случае. Таким образом, уравнение Лиувилля (23) является адекватной моделью для описания эволюции первого момента ВФР.

Рассмотрим эволюцию центральных выборочных моментов, определяемых в соответствии с (20) как

$$m_k(t) = \int \left(x - \mu(t)\right)^k f(x,t) dx, \ k \geq 2 .$$

(27)

Скорость изменения этих величин также может быть найдена по уравнению (23). Имеем

$$\frac{dm_k}{dt} = -km_{k-1}\frac{d\mu}{dt} + \int (x-\mu)^k \frac{\partial f}{\partial t} dx = -km_{k-1}\frac{d\mu}{dt} - \int (x-\mu)^k \frac{\partial uf}{\partial x} dx =$$

$$= -km_{k-1}\frac{d\mu}{dt} + km_{k-1}U + k\int (x-\mu)^{k-1}(u-U) f dx.$$

Преобразуем полученное выражение. Первые два слагаемых в нем, содержащие $m_{k-1}$, в силу (26) взаимно уничтожаются, а оставшийся интеграл может быть выражен через ковариации скорости и степеней исходной случайной величины $x$:

$$\int (x-\mu)^{k-1}(u-U) f dx = \sum_{j=0}^{k-1} (-1)^j \binom{k-1}{j} \mu^j \operatorname{cov}\left(x^{k-1-j}, u\right).$$

Поскольку при $j = k-1$ одна из величин под знаком ковариации является константой, то соответствующая ковариация равна нулю. Таким образом, окончательно получаем уравнение

$$\frac{dm_k}{dt} = k\sum_{j=0}^{k-2} (-1)^j \binom{k-1}{j} \mu^j \operatorname{cov}\left(x^{k-1-j}, u\right)$$

(28)

В частности, для уравнения эволюции выборочной дисперсии получаем

$$\dot{m}_2 = 2\,\mathrm{cov}(x,u)\,.$$

(29)

С другой стороны, изменение тех же величин (27) может быть вычислено в момент времени *t* непосредственно по имеющимся выборкам (см. выше Рис. 1-3). Получающиеся при этом значения будут, вообще говоря, отличаться от (28), поскольку эмпирическое уравнение Лиувилля (23) без включения в рассмотрение высших разностей (ускорений и т.д.) представляет лишь приближенную эволюционную модель ВФР. Мы сознательно пошли на такие упрощения, желая сократить объем вычислений, т.к. величина массива данных растет как степень объема одномерной выборки. Возникает вопрос: можно ли, не увеличивая существенно объем вычислений, т.е. оставаясь в том же пространстве, уточнить модель (23), чтобы некоторые, по крайней мере первые центральные моменты ВФР, действительно эволюционировали бы в соответствии с уточненным уравнением?

Для ответа на этот вопрос используем эвристические соображения о том, что при надлежащих условиях на моменты некоторого стационарного распределения оно (распределение) определяется ими однозначно (теорема Гамбургера [7]). Именно, теорема Гамбургера утверждает, что если некоторая последовательность чисел $\{\mu_{2n}\}$ такова, что ряд по обратным степеням этих величин расходится

$$\sum_{n=1}^{\infty}(\mu_{2n})^{-1/(2n)} = \infty\,,$$

то существует единственная функция распределения $f(x)$, моменты которой равны этим числам $\mu_{2n}$.

Как было показано в п.1, с некоторой точностью можно говорить об $\varepsilon$-стационарных выборочных распределениях, которые хотя и не имеют пределов (по вероятности или в слабом смысле и т.п.), могут доставлять оценки тех или иных статистик, удовлетворяющие практические нужды.

Поэтому, хотя требование состоятельности оценки, например, асимптотически несмещенной дисперсии не выполняется, желательно, чтобы с точностью, контролируемой исследователем, соответствующим выборочным средним по нестационарным распределениям можно было придать практический смысл. Полезно, например, ввести понятие $\varepsilon$-несмещенной оценки по выборке объема *T*. Такой оценкой мы будем называть статистику, отклонение среднего по которой от среднего по предположительно стационарному распределению не превосходит $\varepsilon$. В этом случае отличие модулей разности смещенных и несмещенных оценок, построенных по двум близким (т.е. $\varepsilon$-стационарным) ВФР будет иметь порядок $o\left(\varepsilon^2\right)$.

*Определение 7. Оценка A величины a по $\varepsilon$-стационарному распределению называется $\varepsilon$-несмещенной, если отличие модулей разности этой оценки и*

*оценки $B$ той же величины, но являющейся несмещенной, если распределение стационарно, имеет порядок $o(\varepsilon^2)$.*

Например, пусть $A$ есть оценка дисперсии $m_2$ согласно определениям (19-20). Если бы распределение было стационарным, то величина $B = \frac{1}{T-1}\sum_{k=t-T+1}^{t}(x(k)-\mu)^2$ являлась бы несмещенной оценкой дисперсии этого распределения. Поскольку моменты $\varepsilon$-стационарного распределения тоже $\varepsilon$-стационарны, то $\left|A-\widetilde{A}\right| \le \varepsilon$. Так как оценки $B$ и $A$ связаны равенством $B = \frac{T}{T-1}A$ и, по оценке (6), $\frac{1}{T} = O(\varepsilon)$, то для $\left|B-\widetilde{B}\right|$ получаем оценку $\left|B-\widetilde{B}\right| = \left|A-\widetilde{A}\right| + o(\varepsilon^2)$. Таким образом, выборочная дисперсия (21) и аналогично высшие центральные моменты являются $\varepsilon$-несмещенными оценками $\varepsilon$-стационарного распределения.

Поскольку выборочный момент порядка $k$ имеет $k$-1 степень свободы, то для выборки фиксированного объема $T$ осмысленно можно ставить вопрос об исследовании только первых $T$-1 моментов. Более того, чтобы получающиеся оценки были $\varepsilon$-несмещенными, требуется выполнение условия $k/T \le \varepsilon$. Это означает, что корректными оценками моментов нестационарного распределения конечной выборки являются только первые несколько выборочных моментов. Именно, если при прогнозировании на $\tau$ шагов при заданном уровне $\varepsilon$-стационарности ВФР определен оптимальный объем выборки $T \le 2\tau/\varepsilon$, то при построении модели эволюции ВФР мы можем аргументированно использовать оценки только для моментов порядка до $2\tau$.

Приведенные соображения позволяют уточнить эволюционную модель для ВФР. Например, пусть у нас есть эмпирическое значение производной по времени дисперсии данной ВФР, равное $\dot{m}_2^{(e)}$. Обозначим невязку этого значения с уравнением (29) через $\lambda$:

$$\lambda(t) = \dot{m}_2^{(e)} - 2\,\mathrm{cov}(x,u).$$

(30)

Рассмотрим для ВФР вместо уравнения Лиувилля (23) уравнение Фоккера-Планка:

$$\frac{\partial f}{\partial t} + \frac{\partial (uf)}{\partial x} - \frac{\lambda(t)}{2}\frac{\partial^2 f}{\partial x^2} = 0.$$

(31)

Для этого уравнения в силу его дивергентной формы уравнение эволюции первого момента остается тем же, что и выше, т.е. имеет вид (26), а уравнение эволюции выборочной дисперсии, как легко проверить, совпадает с соответствующей эмпирической производной $\dot{m}_2^{(e)}$.

Уравнение Фоккера-Планка применяется, как правило, для описания процесса с независимыми нормально распределенными приращениями. При независящих от времени величинах $u$ и $\lambda$ оно описывает перенос и диффузию вероятности для гауссовского белого шума. Для нестационарного процесса оно играет роль модельного и является определенным уточнением уравнения Лиувилля, если есть основания считать, что в исследуемом процессе присутствуют случайные блуждания. Некоторым «оправданием» такой не вполне корректно обоснованной модели служит то, что уравнение Фоккера-Планка активно (и также не вполне обоснованно) применяется во многих задачах статистической физики. Необоснованность же уравнения связана с тем, что гауссов белый шум является удобной, но абстрактной математической моделью, на практике не реализующейся, поскольку не может реально наблюдаться процесс, имеющий автокорреляционную функцию в виде дельта-функции.

С целью обобщения уравнения (31) при построении модели эволюции ВФР, которое должно приводить к верным эмпирическим значениям производной по времени для выборочных моментов вплоть до порядка $r$, вводим эмпирически определяемые величины

$$\lambda_k(t) = \frac{dm_k^{(e)}}{dt} - k\sum_{j=0}^{k-2}(-1)^j \binom{k-1}{j} \mu^j \operatorname{cov}\left(x^{k-1-j}, u\right)$$

(32)

и составляем уравнение

$$\frac{\partial f}{\partial t} + \frac{\partial (uf)}{\partial x} + \sum_{k=2}^{r}(-1)^{k-1}\frac{\lambda_k(t)}{k!}\frac{\partial^k f}{\partial x^k} = 0\,.$$

(33)

Легко проверяется, что в силу этого уравнения $\dot{\mu} = U, \dot{m}_k = \dot{m}_k^{(e)}$.

Важно подчеркнуть, что количество учитываемых эмпирических выборочных моментов в (33) зависит от горизонта и желаемой точности прогноза. Если оказывается, что при этом близость между ВФР должна быть не больше заданного числа $\varepsilon$, то, как указывалось выше, учитываемое количество моментов не должно превосходить $\varepsilon T$, которое, в свою очередь, не превосходит 2τ. Например, при ε~(1/$T$; 2/$T$) корректные модели эволюции ВФР могут быть основаны на уравнениях Лиувилля (23) или Фоккера-Планка (31). Если ε~(2/$T$; 4/$T$), то можно использовать уравнение (31) или включить в рассмотрение третью пространственную производную, и т.д.

## 4. Сравнительный анализ точности прогнозных моделей

При численном решении уравнений (23) и (31) производные по аргументу $x$ аппроксимировались схемами с центральными разностями. Для модели, основанной на уравнении Лиувилля, численно решалось уравнение

$$\widetilde{f}(x,t+1) = f(x,t) + \frac{1}{2}u(x-1,t)f(x-1,t) - \frac{1}{2}u(x+1,t)f(x+1,t).$$

(34)

Для уравнения Фоккера-Планка дискретная схема определения ВФР на следующем временном слое имеет вид

$$\widetilde{f}(x,t+1) = f(x,t) + \frac{1}{2}u(x-1,t)f(x-1,t) - \frac{1}{2}u(x+1,t)f(x+1,t) +$$

$$+ \frac{\lambda(t)}{2}\big(f(x+1,t) - 2f(x,t) + f(x-1,t)\big).$$

(35)

При решении конечно-разностного уравнения Лиувилля (34) на каждом шаге по времени проводилась коррекция найденной функции распределения. Отрицательные значения «насильственно» зануялись, а затем распределение вновь нормировалось на единицу.

После нахождения прогнозной ВФР для следующего момента времени строился прогноз собственно значения временного ряда $x(t+1)$. Этот прогноз может быть выполнен в рамках различных вероятностных моделей: ожидаемое значение $x(t+1)$ можно определить как среднее значение по прогнозной ВФР, как наиболее вероятное значение, или иными способами.

В данной работе прогноз временного ряда строился исходя из того, что точка наибольшего (не абсолютного, а алгебраического) изменения ВФР может быть интерпретирована как аргумент наиболее вероятного значения временного ряда. Тогда аргумент, при котором максимальна дивергенция $\partial(uf(x,t))/\partial x$, и будет прогнозным значением ряда.

Для полученного таким путем прогноза вычислялись текущие математическое ожидание, дисперсия и ковариация, определялись их разностные производные по времени, и по этим данным строился прогноз ВФР на следующий шаг.

Аналогичные действия выполнялись и при численном решении уравнения Фоккера-Планка.

Прогноз ВФР на основе численного расчета по уравнениям (34) и (35) показал, что модель на основе уравнения Фоккера-Планка является более точной. Хотя в абсолютных величинах повышение точности незначительное – в среднем ошибка прогноза по уравнению (35) меньше ошибки прогноза по уравнению (34) на слабо меняющуюся от объема выборки величину, равную 0,21%, – в относительных величинах (т.е. по отношению к ошибке прогноза) это может быть заметным эффектом. Например, при прогнозировании ВФР с точностью 5% относительный эффект использования более точной модели составил приблизительно 4%, тогда как при прогнозировании с точностью 1,5% – уже 14%.

Фактическая ошибка прогноза ряда оказывается несколько больше, чем декларируемая точность прогноза ВФР. Это связано с тем, что кроме чисто статистической ошибки прогноза погрешность содержится и в самом методе. Так, при построении ВФР ошибка в определении принадлежности элементов

ряда некоторому интервалу $\Delta x$ имеет значение порядка $1/N$, где $N$ есть количество ячеек, на которое разбит интервал изменения величины $x$. В наших расчетах использовалось разбиение на 100 ячеек, т.е. ВФР строилась с точностью 1%. Кроме того, ошибка возникает и при процедуре перенормировки ВФР после каждого шага расчета по времени. Оказалось, что из-за появления отрицательных значений нормировка не сохраняется примерно на 0,5%. Наконец, неточность вносится и тем, что скорость $u(x,t)$, дающая прогноз изменения ВФР, берется в один – предшествующий – момент времени. На самом деле эта скорость также является случайной величиной, которая должна прогнозироваться на основе анализа ускорений и т.д., однако увеличение размерности фазового пространства привело бы к существенному замедлению счета.

Отметим, что точность конструирования ВФР по данным наблюдения, которые считаются «абсолютно точными», может быть увеличена единственным образом – путем использования более мелкого разбиения. Естественно, эта мелкость является ограничивающим фактором для самой постановки вопроса о прогнозе ВФР с заданной точностью, поскольку точность прогноза не может превышать точность конструирования самой ВФР. Однако и сама мелкость разбиения не может быть выбрана произвольной. Дело в том, что эмпирическая вероятность попадания значения случайной величины в некоторый интервал является по своему существу величиной стационарной, т.к. она определяется по одной конкретной выборке. Для нестационарных процессов такая выборка единственна, т.к. у нас нет возможности построить ансамбль вместо данной реализации случайного процесса. Поэтому, чтобы можно было корректно говорить о частотах попадания значений элементов выборки в заданный интервал как о квазистационарных вероятностях, надо, чтобы число элементов в выборке было во всяком случае не меньше, чем число ячеек разбиения. Для выборок малого объема это условие заведомо нарушается, т.е. ВФР не всегда статистически репрезентативна. Поскольку закон распределения вероятностей не известен априори, то под него невозможно заранее подогнать разбиение области значений случайной величины, чтобы эти значения заполняли бы ячейки равномерно. Приемлемость такого подхода связана с тем, что для нестационарных процессов частоты попадания значений в некоторый интервал все равно нельзя трактовать как оценку соответствующей вероятности, а возможность использования их в этом качестве обеспечивается функционалом (5) как мерой близости двух распределений.

Подчеркнем, что кинетическое уравнение относительно ВФР как функции только одного фазового аргумента $x$ является прогнозной моделью первого порядка, когда по «известной» скорости строится прогноз самого ряда. Поэтому корректно сравнивать этот прогноз с другими – стандартными – методами также первого порядка: прогнозами методом скользящих средних для регрессионной и авторегрессионной моделей. Для получения модели второго порядка в уравнении Лиувилля (21) следует рассмотреть фазовое

пространство координат, скоростей и ускорений. Строго говоря, уравнение Фоккера-Планка неявно учитывает зависимость от высших производных, поскольку в нем используется некоторая «поправленная» локальная средняя скорость, однако эта модель является все же менее точной, чем уравнение Лиувилля в фазовом пространстве большей размерности.

В качестве тестовых примеров оценки точности метода были взяты ценовые ряды акций компаний GM, GE, MSFT, FORD по данным [6].

По результатам анализа можно сделать два основных вывода. Во-первых, среди стандартных прогнозных методов первого порядка наша модель имеет более высокую точность (хотя и не очень значительно). Во-вторых, ошибка прогнозов по стандартным моделям скользящих средних минимальна, если в качестве объема усреднения берется оптимальный объем, вычисленный по горизонтной статистике для данного ряда. Например, для ряда GE (Рис. 4) ошибка по оптимальному и неоптимальному объемам выборки в среднем составила для стандартных методов соответственно 10-11% (показаны на Рис. 4) и 17-19% (не показаны), а для прогноза по кинетическому уравнению – 9%. Корреляционные методы имеют более высокую точность, чем регрессионные. Аналогичное поведение демонстрируют и ряды для GM, MSFT, FORD.

Разумеется, регрессионные методы первого порядка не претендуют на высокую точность в описании даже стационарных процессов, что и демонстрируется графиками на Рис. 4. Однако следует отметить, что кинетический прогноз, представленный гладкой кривой, имеет более высокую адаптивность, хотя и построен по тем же самым скользящим выборкам. Кроме того, важным его преимуществом является то, что в регрессионных методах ошибка прогноза накапливается с ростом горизонта прогнозирования, тогда как точность кинетического прогноза остается в фиксированных пределах.

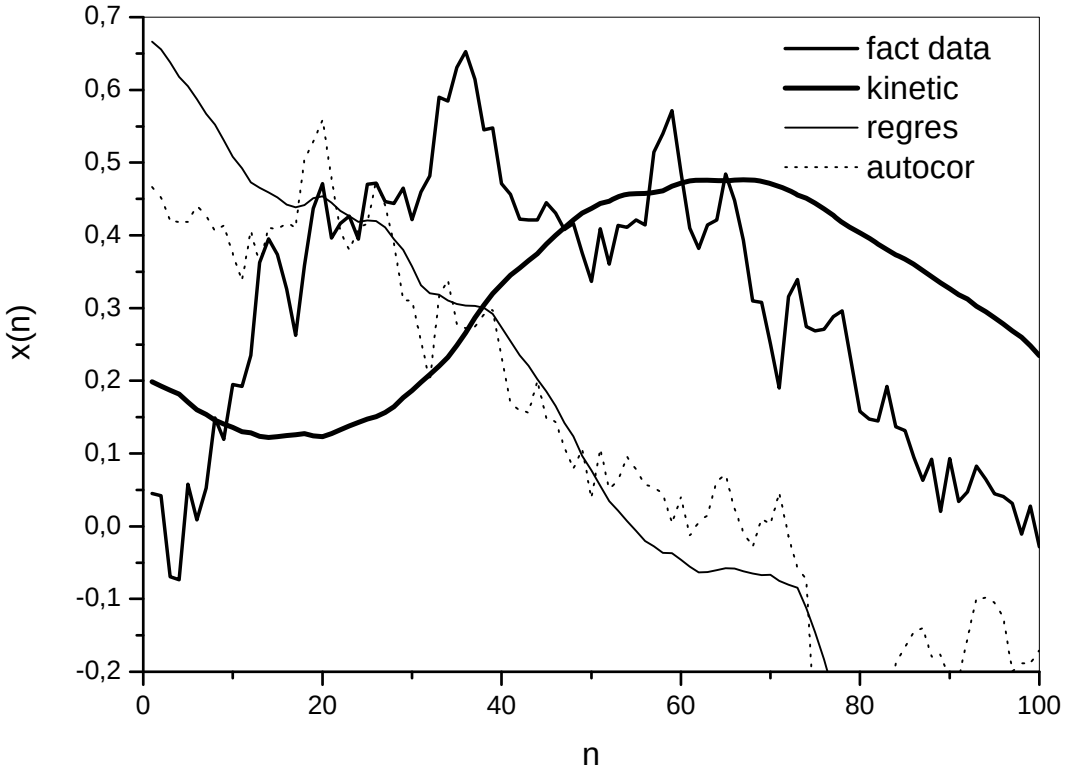


Рис. 4. Сравнение прогнозов ряда GE на 1 шаг вперед

Проведенный анализ показывает, что метод кинетических уравнений можно успешно использовать для более точного прогнозирования нестационарных временных рядов (по крайней мере тех, которые рассмотрены в наших примерах). Поскольку, однако, точность метода первого порядка в принципе не очень высока, а кинетический подход второго порядка (с учетом ускорений) требует больших вычислительных затрат, то для повышения точности прогноза можно рекомендовать использование стандартных методов второго или более высокого порядков, дополненных алгоритмом построения горизонтной статистики и определения оптимального объема выборки.

## 5. Уравнение Лиувилля для динамических систем с особенностями

Используемое в настоящей работе эволюционное уравнение для ВФР является не вполне законным с той точки зрения, что формально у нас нет динамической системы, в силу которой эволюционирует распределение начальных условий, т.е. распределение ансамбля таких динамических систем. Кроме того, нет и самого ансамбля. В качестве псевдо-скоростей, псевдо-ускорений и т.п. мы использовали выборочные эмпирические зависимости по текущему распределению, которые носят лишь формальную аналогию с величинами, входящими в уравнение Лиувилля статистической механики. В данной работе мы не обсуждаем метод выделения т.н. главных компонент

ряда [8], который с некоторой точностью позволяет подобрать динамическую систему, дающую похожую псевдо-случайную траекторию. Термин «уравнение Лиувилля» используется нами здесь в том смысле, что это уравнение эволюционного типа, которое по виду ВФР в предыдущий момент времени позволяет найти ее в следующий момент. В силу вероятностного характера динамики, порождаемой случайным временным рядом, эта эволюция должна пониматься в некотором усредненном аспекте. Такой моделью и является предлагаемое в работе уравнение (25) и его дальнейшие модификации (33).

В связи с использованием уравнения Лиувилля для прогнозирования ВФР необходимо отметить еще один теоретический аспект, возникающий при использовании этого уравнения «по прямому назначению», т.е. для динамических систем – но с хаотическим поведением. При этом хаос в динамических системах может быть не только детерминированным, т.е. порождаемым специальной схемой дискретизации некоторой непрерывной регулярной динамической системы (как в случае логистического отображения), но и появляться в результате вырожденности траектории в фазовом пространстве, т.е. в силу неединственности решений уравнений движения. Если нет указания на поведение системы в особой точке, то она может с некоторой вероятностью продолжить движение по любой из доступных ей траекторий. Статистические свойства таких систем изучались в [9, 10].

Поскольку мы строим эмпирическое уравнение Лиувилля, которому по смыслу отвечает некоторая динамическая система с хаотическим поведением, следует предусмотреть возможность того, что, получаемая численно, такая система может быть дискретной реализацией динамической системы с особыми точками в фазовом пространстве. Для такой системы уравнение Лиувилля должно быть модифицировано введением дополнительных, не содержащихся изначально в системе условий, позволяющих «пройти» особое решение.

В качестве простейшей иллюстрации проблемы построения уравнения Лиувилля для динамических систем с особенностями рассмотрим систему, порождаемую дифференциальным уравнением, не разрешенным относительно производной по времени:

$$F(x,p,t)=0, \quad p=\dot{x}. \tag{36}$$

Достаточным условием разрешимости уравнения (36) относительно $p$ является $\partial F/\partial p \neq 0$. В этом случае через каждую такую точку $\{x,t\}$ в расширенном фазовом пространстве, называемую регулярной, проходит ровно одна интегральная кривая. В точке $\{x,t\}$ определено единственное векторное поле $\{\dot{x},1\}\equiv\{p(x,t),1\}$, где $p(x,t)$ является локальным решением уравнения (36) относительно $p$. Исследуем эволюцию ансамбля таких систем, различающихся начальными условиями, и введем плотность распределения координаты $x$, учитывая, что эта координата является

решением уравнения (36). Тогда эволюция распределения $\rho(x,t)$ в силу системы (36) описывается уравнением Лиувилля

$$\frac{\partial\rho}{\partial t}+p(x,t)\frac{\partial\rho}{\partial x}=0.$$

(37)

Т.к. полная вероятность для системы пребывать в фазовом пространстве $\gamma$ равна единице и сохраняется во времени, то должна существовать плотность динамически-инвариантной меры, такой, что $\int\rho d\gamma=\int\rho(x,t)\sqrt{g(x,t)}dx=1$, где $\sqrt{g}$ и есть эта плотность. Легко убедиться, что $\sqrt{g}=1/p$ есть требуемая плотность, т.к. тогда

$$div\rho\dot{x}=\frac{1}{\sqrt{g}}\frac{\partial}{\partial x}\left(\rho\sqrt{g}\dot{x}\right)=\frac{1}{\sqrt{g}}\frac{\partial}{\partial x}\left(\rho p\sqrt{g}\right)=p\frac{\partial\rho}{\partial x},$$

и уравнение (39) можно записать в дивергентной форме закона сохранения.

В случае же, если $\partial F/\partial p=0$, такой меры в общем случае может и не быть, а уравнение вида (37) записать не удастся, т.к. скорость *p* может быть не определена. Если $\partial^2F/\partial p^2\neq0$, то уравнение $\partial F/\partial p=0$ можно локально разрешить относительно *p*, после чего исключить *p* и из исходного уравнения (36). Получившаяся в результате дискриминантная кривая $S(x,t)=F(x,p(x,t),t)=0$ в расширенном фазовом пространстве $\{x,t\}$ является особым решением, имеющим общую точку с каждой регулярной интегральной кривой. Движение ансамбля систем в этой точке не определено, т.к. в ней нет единственности решения уравнения (36). Следовательно, на дискриминантной кривой следует поставить граничные условия для плотности вероятности $\rho(x,t)$, которые могут быть сформулированы в виде некоторого правила рассеяния вероятности на особой траектории. Эти правила можно записать в том же виде, что и для движения по регулярной траектории:

$$\frac{\partial\rho(x,t)}{\partial t}=\int P(x,x';t)\rho(x',t)dx',$$

(38)

где $P(x,x';t)$ есть скорость перехода (вероятность перейти из точки $x'$ в точку *x* в единицу времени). Для регулярной точки траектории скорость перехода $P(x,x';t)$ определяется через фазовый поток $G_{t,t_0}$, который представляет собой символическую запись интегральных кривых векторного поля $\{p(x,t),1\}$ для уравнения (36): $x(t)=G_{t,t_0}x(t_0)$. Именно, в этом случае $P(x,x';t)=\frac{\partial}{\partial t}\delta\left(G_{t,t_0}(x')-x\right)$. Для сингулярной точки выражение $P(x,x';t)$ содержит фактор $\delta(S(x',t))$ и правила продолжения траектории через эту точку. Важно подчеркнуть, что эти правила являются внешними по отношению к задаче (36), они не выводимы из этого уравнения. Например,

можно потребовать, чтобы фазовый поток «не замечал» особую траекторию. Тогда, если обозначить $t_S(x)$ неявную функцию, определяемую дискриминантной кривой, получим правило

$$\lim_{t\to t_s+0}\rho(x,t)=\lim_{t\to t_s-0}\rho(x,t).$$

После этого $\rho(x,t)$ доопределяется по непрерывности во всех точках.

Характерным примером, обобщающим случай (36), является движение лагранжевой системы с вырождающимся гессианом [10]. Рассмотрим систему, динамика которой задается лагранжианом $L(q,v)\equiv L(x),\ x=\{q,v\},\ v=\dot{q}$. Если действие

$$S[L]=\int_{t_0}^{t_1}L(q,v)dt$$

имеет стационарную точку, то эта точка называется экстремалью и является решением уравнений Эйлера-Лагранжа

$$\frac{d}{dt}\frac{\partial L}{\partial v_i}=\frac{\partial L}{\partial q_i},$$

которые представляют собой систему дифференциальных уравнений второго порядка по времени $t$ относительно координат:

$$\frac{\partial^2 L}{\partial v_i\partial v_j}\dot{v}_j=f_i{}^{\alpha}=-\frac{\partial}{\partial q_j}\left(v_j\frac{\partial L}{\partial v_i}-L\delta_{ij}\right).$$

(39)

Совокупность величин $M_{ij}=\dfrac{\partial^2 L}{\partial v_i\partial v_j}$ образуют тензор массы динамической системы. Обозначим определитель соответствующей матрицы Гесса через $J=\det(M_{ij})$. Согласно определениям [11], лагранжиан $L$ называется невырожденным, если этот определитель не обращается в ноль ни в одной точке фазового пространства. Лагранжиан называется сильно невырожденным, если он невырожденный, и уравнения $p_i=\dfrac{\partial L}{\partial v_i}$ имеют единственные непрерывно дифференцируемые и обратимые решения $v_i=v_i(q,p)$. Отображение $\Im:\{q,v\}\to\{q,p\}$ называется преобразованием Лежандра, если $L$ сильно невырожден. Если в некоторой области фазового пространства $\gamma$ гессиан вырожден, то там ускорения могут не быть представлены как однозначные функции фазовых переменных. В работе [9] эта ситуация трактуется как хаотизация классической динамики.

Классическим примером является случай, когда лагранжиан зависит от модулей координаты и скорости: $L=L(q,v),\quad q=|\mathbf{q}|,\ v=|\mathbf{v}|$. Тогда

$$M_{ij}=\frac{1}{v}\frac{\partial L}{\partial v}\Delta_{ij}^{\perp}+\frac{1}{v}\frac{\partial E}{\partial v}\Delta_{ij},$$

где $E = v\dfrac{\partial L}{\partial v} - L$ – энергия системы, а $\Delta_{ij}$ и $\Delta_{ij}^{\perp}$ – взаимно-ортогональные проекторы на направление вектора скорости в *n*-мерном пространстве:

$$\Delta_{ij} = e_i e_j, \quad \Delta_{ij}^{\perp} = \delta_{ij} - \Delta_{ij}, \quad \mathbf{e} = \mathbf{v}/v.$$

Определитель матрицы $M_{ij}$ легко вычисляется, поскольку $\dfrac{1}{v}\dfrac{\partial E}{\partial v}$ – ее собственное значение кратности 1, а $\dfrac{1}{v}\dfrac{\partial L}{\partial v}$ – собственное значение кратности $n-1$, так что

$$J = \left(\frac{1}{v}\frac{\partial L}{\partial v}\right)^{n-1}\frac{1}{v}\frac{\partial E}{\partial v}, \quad \left(M^{-1}\right)_{ij} = \left(\frac{1}{v}\frac{\partial L}{\partial v}\right)^{-1}\Delta_{ij}^{\perp} + \left(\frac{1}{v}\frac{\partial E}{\partial v}\right)^{-1}\Delta_{ij}.$$

(40)

Отсюда следует, что особыми точками являются нетривиальные (с ненулевым значением скорости) стационарные точки энергии и лагранжиана, причем последние существуют только в пространстве размерности больше 1. Стационарные точки энергии отвечают разрыву компоненты ускорения, направленной вдоль скорости и потому касательной к пространственной траектории тела, а стационарные точки лагранжиана – разрыву составляющей ускорения, ортогональной скорости. Согласно [10], сингулярным множеством $S_D$ такой динамической системы называется множество точек, в которых $J = 0$. В данном примере $S_D$ разбивается в прямую сумму двух множеств: сингулярного множества $S_D^1$ первого типа, отвечающего разрыву компоненты ускорения, направленной вдоль движения, и сингулярного множества $S_D^2$ второго типа, отвечающего разрыву нормальной компоненты ускорения.

Пусть система имеет набор независимых первых интегралов $I(x)$. В [10] показано, что траекторию в некоторых случаях можно продолжить через множество $S_D^1$ с сохранением по непрерывности первых интегралов и зеркальным отражением направления движения, если это точка локального максимума, либо прохождением без изменения направления движения, если это точка локального минимума. В частности, при отражении

$$\lim_{t\to t_s+0}\rho(x,t) = \lim_{t\to t_s-0}\int\rho(x',t)\delta\big(S(x')\big)\delta\big(I(x)-I(x')\big)\delta(\mathbf{q}-\mathbf{q}')\delta\big(\mathbf{x}-2\mathbf{n}(\mathbf{x}'\mathbf{n})\big)dx'$$

точка сохраняет координаты в физическом пространстве, но меняет на противоположные нормальную составляющую скорости. При прохождении без отражения и «конденсации» на особом множестве получаем

$$\lim_{t\to t_s+0}\rho(x,t) = \lim_{t\to t_s-0}\int\rho(x',t)\delta\big(S(x')\big)\delta\big(I(x)-I(x')\big)\delta(x-x')dx'.$$

В случае с сингулярным множеством $S_D^2$ второго типа $n-1$ компонент ускорения являются неопределенными. Это можно трактовать как рассеяние на произвольный угол с сохранением интеграла энергии.

В литературе весьма подробно исследованы вопросы возникновения хаоса в динамических системах (см., напр., монографии [12, 13]), а также изучены и классифицированы особенности динамических систем в виде дифференцируемых отображений [14-16]. В первых системах хаос возникает при определенных параметрах дискретизации дифференциальных уравнений движения, в которых не было каких-либо особенностей. Во вторых системах особенности присущи именно дифференциальному уравнению, а при дискретизации они проявляются лишь приближенно и при достаточно мелком шаге разбиения. В последнем случае уравнение Лиувилля имеет некоторые черты «эмпирического», т.к. оно дополняется некоторыми внешними условиями, обеспечивающими проход через особое решение, возможно, в виде вероятностного выбора каждой из траекторий. Тем самым статистическая механика динамических систем с особенностями образует связь с теорией случайных процессов и, в частности, временных рядов.

## Заключение

Итак, в работе предложен метод анализа нестационарных временных рядов, который направлен на решение следующих задач:

- определение оптимального объема выборки для формирования квазистационарной ВФР и ее прогнозирования с заданной точностью;
- минимизация совокупного функционала ошибки аппроксимации временного ряда на заданном интервале времени, а также определение интервала времени, на котором такая ошибка минимальна;
- прогнозирование ВФР и временного ряда внутри границ квазистационарности ВФР.

Поскольку в данном случае отсутствуют классические понятия доверительных интервалов, то прогноз в нашем подходе означает совокупность вероятного значения ряда в наперед заданный момент времени и распределения его значений, выборочные квантили которого и дадут оценку доверительных интервалов. В работе показано, что для целей прогнозирования нестационарных временных рядов с использованием понятия эволюции ВФР могут быть применены кинетические уравнения классической статистической механики – уравнение Лиувилля и Фоккера-Планка.

Показано также, что по крайней мере для рассмотренных временных рядов наш подход дает более высокую точность прогноза, чем скользящие адаптивные методы.

## Благодарности

Авторы выражают благодарность своим коллегам профессорам Г.Г. Амосову и Г.Г. Малинецкому, а также доцентам Н.А. Митину и В.Ж. Сакбаеву за плодотворные обсуждения и ценные указания.

## Цитированные источники

1. Королюк В.С., Портенко Н.И., Скороход А.В., Турбин А.Ф. Справочник по теории вероятностей и математической статистике. – М.: Наука, 1985.
2. Уилкс С. Математическая статистика. – М.: Наука, 1967.
3. Орлов Ю.Н., Осминин К.П. Анализ нестационарных временных рядов. / Препринт ИПМ им. М.В. Келдыша РАН, № 36, 2007. – 24 с.
4. Орлов Ю.Н., Осминин К.П. Построение выборочной функции распределения для прогнозирования нестационарного временного ряда. // Мат. Мод., 2008, № 7 (в печати).
5. Осминин К.П. Алгоритмы построения статистик для анализа и прогнозирования нестационарных временных рядов. // ИТВС, № 3, 2008 (в печати).
6.http://moneycentral.msn.com/detail/stock_quote?Symbol=GE&pkw=PI&vendor=Paid+Inclusion&OCID=iSEMPI
7. Ахиезер Н.И. Классическая проблема моментов. – М.: Физматлит, 1961.
8. Главные компоненты временных рядов. Сб. статей / Ред. Д.Л. Данилов и А.А. Жиглявский. СПбГУ, 1997.
9. Pavlotsky I.P., Strianese M., Toscano R. Prolongation of the integral curve on the singular set via the first integral. // J. of Interdisciplinary Math. 1999. V.2. Nos.2-3. P.101-119.
10. Орлов Ю.Н. Основы квантования вырожденных динамических систем. – М.: МФТИ, 2004. – 236 с.
11. Дубровин Б.А., Новиков С.П., Фоменко А.Т. Современная геометрия. – М.: Наука, 1986. – 759 с.
12. Кузнецов С.П. Динамический хаос. – М.: Физматлит, 2001. – 296 с.
13. Малинецкий Г.Г., Потапов А.Б. Современные проблемы нелинейной динамики. – Москва-Ижевск: Регулярная и хаотическая динамика, 2000.
14. Каток А.Б., Хасселблат Б. Введение в современную теорию динамических систем. – М.: Факториал, 1999. – 767 с.
15. Арнольд В.И., Ильяшенко Ю.С. Обыкновенные дифференциальные уравнения. / Итоги науки и техники ВИНИТИ. – М., 1985.
16. Арнольд В.И., Варченко А.Н., Гусейн-Заде С.М. Особенности дифференцируемых отображений. Т.1. – М.: Наука, 1982.