

УДК 519.242.33

Ю. Н. Орлов, А. Д. Босов

Московский физико-технический институт (государственный университет)

## Кинетико-гидродинамический подход к прогнозированию нестационарных временных рядов на основе уравнения Фоккера–Планка

Выводится цепочка кинетических уравнений эволюции многомерных выборочных функций распределения нестационарного временного ряда. Строится система уравнений гидродинамики, замыкание которой на каком-либо моменте распределения приводит к модели временного ряда в виде порождающей его нелинейной динамической системы.

**Ключевые слова:** выборочная плотность функции распределения, уравнение Лиувилля, уравнение Фоккера–Планка, нестационарный временной ряд.

### 1. Введение

При анализе нестационарных временных рядов одной из центральных задач является оптимизация объема выборки для вычисления текущих статистик ряда. Необходимость оптимизации связана с тем, что малый объем статистически мало репрезентативен, а большой объем данных не дает нужной точности в силу того, что текущее распределение изменилось. Для практического решения этой задачи в работах [1, 2] была введена индикативная статистка, названная авторами «горизонтным рядом» и показывающая, каким объемом лучше всего, т.е. с минимальной ошибкой в среднем квадратичном, определяется текущая выборочная плотность функции распределения (далее ВПФР).

Пусть  $x(t)$  есть значение случайной величины  $\xi$  в момент времени  $t$ , и  $f_T(x, t)$  есть ВПФР, построенная по выборке объема  $T$  в момент времени  $t$ . Говоря о случайной величине  $\xi$ , мы предполагаем, что, по крайней мере, теоретически в каждый момент времени существует некоторая генеральная совокупность, определяющая вероятность появления того или иного значения  $x$ . Для простоты полагаем, что время  $t$  дискретно, ряд  $x(t)$  эквидистантный с единичным шагом, причем все значения равномерно ограничены по времени. В этих условиях без ограничения общности можно считать, что  $\forall t x(t) \in [0; 1]$ . Основное утверждение, показывающее возможность равномерного ограничения по времени для оценки неточности в определении ВПФР, состоит в том, что для функционала  $V(T, \tau; t)$  близости двух ВПФР

$$V(T, \tau; t) = \int_0^1 |f_T(x, t + \tau) - f_T(x, t)| dx \quad (1)$$

имеет место оценка [1]

$$0 \leq V(T, \tau; t) \leq \min(2\tau/T; 2). \quad (2)$$

Промежуток интегрирования далее везде считается отрезком  $[0; 1]$ , причем значения ВПФР в граничных точках полагаются равными нулю. Пределы интегрирования далее для краткости не указываются.

**Определение 1.** Горизонтным рядом  $h(t)$  для ряда  $x(t)$  при сдвиге на промежуток  $\tau$  называется такой минимальный объем выборки  $h(t, \tau; \varepsilon)$ , что при всех  $T \geq h(t, \tau; \varepsilon)$  выполнено условие  $V(T, \tau; t) \leq \varepsilon$ .

Таким образом, горизонтный ряд – это двухпараметрическая статистика, которая характеризует взаимозависимость двух выборок. Оптимальный объем выборки находится из распределения горизонтного ряда, если последнее квазистационарно на горизонте прогнозирования  $\tau$ , т.е. если различия в распределении горизонтной статистики составляют величину  $o(\varepsilon)$ . Примеры нестационарных рядов с таким свойством приведены в [3].

Оптимизация объема выборки приводит к тому, что распределение  $f_T(x, t)$  на горизонте  $\tau$  в первом приближении (с точностью  $\varepsilon$ ) можно считать стационарным. В следующем приближении можно построить уравнение, описывающее эволюцию ВПФР на этом горизонте. В силу нормировки ВПФР на единицу это уравнение должно иметь вид закона сохранения, т.е. быть аналогом уравнения Лиувилля в классической статистической механике. В [3] показано, что можно придать корректный смысл эмпирическому уравнению Лиувилля:

$$\frac{\partial f_T(x, t)}{\partial t} + \frac{\partial}{\partial x} u_T(x, t) f_T(x, t) = 0, \quad (3)$$

где локальная скорость  $u_T(x, t)$ , называемая в [3] эмпирической лиувиллевой скоростью (ЭЛС), определяется через двумерную ВПФР  $F(x, \dot{x}, t)$  значений ряда и его приращений (далее указание на объем выборки  $T$  для краткости опускаем) по формуле

$$u(x, t) f(x, t) = \int \dot{x} F(x, \dot{x}, t) d\dot{x}, \quad f(x, t) = \int F(x, \dot{x}, t) d\dot{x}. \quad (4)$$

Производные по времени в (3) и (4) трактуются как разностные производные с шагом 1, а интегралы представляют собой сумму по конечному множеству разбиения отрезка  $[0; 1]$ . Эта трактовка оправдана, если  $T \gg 1$ . Если удастся каким-либо образом определить ЭЛС, то из (3) можно будет найти и ВПФР. Прогноз самого ряда строится как среднее значение по найденной ВПФР.

В настоящей работе предложен метод определения ЭЛС с помощью системы уравнений для эволюции моментов ВПФР  $f(x, t)$ , которые замыкаются на третьем уравнении. В гидродинамической аналогии это уравнение энергии. Замыкание происходит в духе «наивного прогноза», когда момент третьего порядка считается постоянным. Корректность такой гипотезы проверяется постфактум, т.е. модель признается приемлемой для данного временного ряда, если изменение «постоянной» величины на горизонте  $\tau$  можно считать величиной порядка  $o(\varepsilon)$ . На практике обычно полагают, что  $\varepsilon = 0,05$ , а  $o(\varepsilon) \leq 0,01$ .

Прогнозная модель на основе уравнения Лиувилля (3) была предложена в [3]. Здесь будет построена модель, основанная на эволюционном уравнении Фоккера—Планка, которое моделирует не только снос плотности вероятности со скоростью  $u(x, t)$ , но и эффект случайного блуждания, оцениваемого по эмпирическим данным с предыдущего шага по времени. Эта модель оказалась более точной в плане прогнозирования временного ряда, чем модель, основанная на уравнении Лиувилля.

## 2. Цепочка кинетических уравнений для ВПФР

Опишем идеологию вывода кинетических уравнений для эволюции ВПФР, следуя подходу А.А. Власова, изложенному им в [4]. Предположим, что в нашем распоряжении имеется вся совокупность величин  $x(t)$ , так что можно построить совместные «генеральные» плотности функции распределения (ПФР)  $f_n(x, \dot{x}, \dots, x^{(n)}; t)$  любой размерности. Мера

$$d\Gamma_n = f_n(x, \dot{x}, \ddot{x}, \dots, x^{(n)}; t) \prod_{k=0}^n dx^{(k)} \quad (5)$$

представляет число кинематических состояний частиц в границах указанных промежутков. Согласно формулам теории вероятностей для плотностей совместных многомерных распределений, имеет место зацепление ПФР в виде

$$f_n(x, \dot{x}, \dots, x^{(n)}; t) = \int f_{n+1}(x, \dot{x}, \dots, x^{(n)}, x^{(n+1)}, t) dx^{(n+1)}. \quad (6)$$

Нетривиальным является вопрос о том, что представляет собой математический объект, получающийся из (5) предельным переходом  $n \rightarrow \infty$ . В [4] этот вопрос не обсуждался; видимо, предполагалось, что смысл объекта

$$d\Gamma = F(x, \dot{x}, \ddot{x}, \dots, x^{(k)}, \dots; t) \prod_{k=0}^{\infty} dx^{(k)} \quad (7)$$

интуитивно понятен. Однако далее для плотности меры (7) будет записано некое эволюционное уравнение, поэтому требуется определить ее более корректно. Такое определение дано в недавней работе В.В. Козлова и О.Г. Смолянова [5], где построена также и общая схема перехода от уравнения Лиувилля относительно мер на бесконечномерных пространствах к равносильной этому уравнению бесконечной системе уравнений относительно функций распределения на конечномерных пространствах. Мы не будем углубляться в теорию, развитую в [5], поскольку наша задача формулируется в терминах конечномерных пространств, но упомянуть об этой проблеме было необходимо для целостности концепции. Таким образом, можно считать, что уравнение Лиувилля для  $F$ , выражающее закон сохранения нормировки, корректно, как и представления проекций на конечномерные подпространства:

$$\begin{aligned} \frac{\partial F}{\partial t} + \sum_{k=0}^{\infty} \operatorname{div}_{x^{(k)}} (x^{(k+1)} F) &= 0; \\ f_n(x, \dot{x}, \dots, x^{(n)}; t) &= \int F(x, \dot{x}, \ddot{x}, \dots, x^{(k)}, \dots; t) \prod_{k=n+1}^{\infty} dx^{(k)}. \end{aligned} \quad (8)$$

После этого, считая, что при стремлении скоростей и ускорений всех порядков к бесконечности ПФР стремится к нулю, из (8) и из соотношений (6) получаем уравнения эволюции для  $n$ -частичных ПФР:

$$\begin{aligned} \frac{\partial f_0(x, t)}{\partial t} + \operatorname{div}_x \int \dot{x} f_1(x, \dot{x}; t) d\dot{x} &= 0; \\ \frac{\partial f_1(x, \dot{x}; t)}{\partial t} + \operatorname{div}_x \dot{x} f_1(x, \dot{x}; t) + \operatorname{div}_{\dot{x}} \int \ddot{x} f_2(x, \dot{x}, \ddot{x}; t) d\ddot{x} &= 0; \quad \dots; \\ \frac{\partial f_n(x, \dot{x}, \dots, x^{(n)}; t)}{\partial t} + \sum_{k=0}^{n-1} \operatorname{div}_{x^{(k)}} x^{(k+1)} f_n(x, \dot{x}, \dots, x^{(n)}; t) + \\ + \operatorname{div}_{x^{(n)}} \int x^{(n+1)} f_{n+1}(x, \dot{x}, \dots, x^{(n+1)}; t) dx^{(n+1)} &= 0; \quad \dots \end{aligned} \quad (9)$$

Таким образом, уравнение эволюции для каждой частичной ПФР не является замкнутым. Оно связывает между собой концентрации и потоки, причем последние выражаются через ПФР от большего числа переменных. Если на каком-либо этапе оборвать зацепление, т.е. задать последний интеграл в (9) независимо, то в пределе нулевых центральных моментов всех порядков относительно координат, скоростей и ускорений всех порядков будет получаться та или иная динамическая система [3].

Если замкнуть цепочку (9) на первом уравнении, считая интеграл в нем равным нулю, получится уже упомянутое выше в п. 1 стационарное решение. В следующем приближении интеграл по  $d\dot{x}$  учитывается, а обрыв происходит на втором уравнении, где интеграл по  $d\ddot{x}$  считается равным нулю и т.д. Так, при замыкании на втором уравнении система уравнений для  $f_0(x, t)$  и  $f_1(x, \dot{x}, t)$  имеет вид

$$\begin{aligned} \frac{\partial f_0(x, t)}{\partial t} + \frac{\partial}{\partial x} u(x, t) f_0(x, t) &= 0, \quad \frac{\partial f_1(x, \dot{x}, t)}{\partial t} + \dot{x} \frac{\partial f_1(x, \dot{x}, t)}{\partial x} = 0; \\ u(x, t) f_0(x, t) &= \int \dot{x} f_1(x, \dot{x}, t) d\dot{x}, \quad f_0(x, t) = \int f_1(x, \dot{x}, t) d\dot{x}. \end{aligned} \quad (10)$$

Заметим теперь, что в системе (10) скорость  $u(x, t)$  может быть любой, лишь бы при решении первого уравнения не появились бы отрицательные вероятности. Статистический смысл ЭЛС состоит в том, что ее среднее значение равно производной среднего выборочного значения ряда [2]:

$$\frac{d}{dt} \int x f_0(x, t) dx = \int x \frac{\partial f_0(x, t)}{\partial t} dx = - \int x \frac{\partial}{\partial x} u(x, t) f_0(x, t) dx = \int u(x, t) f_0(x, t) dx = U(t). \quad (11)$$

Действительно, эта же величина равна изменению эмпирического первого момента за один шаг по времени:

$$m(t+1) - m(t) = \frac{1}{T} \sum_{k=1}^T x(t-T+k+1) - \frac{1}{T} \sum_{k=1}^T x(t-T+k) = \frac{1}{T} \sum_{k=1}^T \dot{x}(t-T+k) = \int \dot{x} f_1(x, \dot{x}, t) dx,$$

а последнее выражение равно (11) в силу определения ЭЛС (4).

Очевидно, наблюдаемое изменение выборочного среднего значения временного ряда может быть получено при различных ЭЛС  $u(x, t)$ , которые в этом смысле определены с точностью до выражения вида  $\lambda(t) \partial f_0(x, t) / \partial x$ , где  $\lambda(t)$  — произвольная функция времени. Это наблюдение позволяет расширить класс кинетических моделей для описания эволюции ВПФР.

### 3. Эмпирическое уравнение Фоккера—Планка

Цепочка равенств (11) показывает, что эволюция первого момента ВПФР в силу уравнения Лиувилля (3) совпадает с производной по времени среднего выборочного значения, вычисленного непосредственно по временному ряду. Однако для высших моментов такой инвариантности нет. Тем не менее можно так «подправить» ЭЛС без нарушения нормировки ВПФР, чтобы первые два эмпирических момента ВПФР изменялись бы в соответствии с кинетическим уравнением эволюции самой ВПФР.

**Теорема 1.** Пусть  $m(t)$  и  $q(t)$  — первый и второй выборочные моменты нестационарного временного ряда,  $f_0(x, t)$  — его ВПФР,  $u(x, t)$  — ЭЛС, а  $U(t)$  — среднее значение ЭЛС. Тогда эволюция величин  $m(t)$  и  $q(t)$  в силу уравнения Фоккера—Планка:

$$\frac{\partial f_0}{\partial t} + \frac{\partial(u f_0)}{\partial x} - \lambda(t) \frac{\partial^2 f_0}{\partial x^2} = 0, \quad (12)$$

где

$$\lambda(t) = \frac{1}{2} \dot{q}(t) - \text{cov}(x, u) - m(t)U(t), \quad (13)$$

совпадает с их изменениями, вычисленными непосредственно по выборке.

**Доказательство.** Прямой проверкой убеждаемся в том, что (11) выполнено для  $m(t) = \int x f_0(x, t) dx$ . Рассмотрим далее производную  $dq/dt$  в силу уравнения (12):

$$\frac{dq}{dt} = \int x^2 \frac{\partial f_0}{\partial t} dx = \int x^2 \left( -\frac{\partial(u f_0)}{\partial x} + \lambda \frac{\partial^2 f_0}{\partial x^2} \right) dx.$$

После двукратного интегрирования по частям с учетом того, что на границах промежутка ВПФР равна нулю, получаем

$$\frac{dq}{dt} = 2 \int x \left( u f_0 - \lambda \frac{\partial f_0}{\partial x} \right) dx = 2 \text{cov}(x, u) + 2mU + 2\lambda,$$

где  $\text{cov}(x, u)$  обозначает ковариацию величин  $x$  и  $u(x, t)$ . Полученное выражение совпадает с выражением (13). Теорема 1 доказана.

Таким образом, величина  $\lambda(t)$ , представляющая собой коэффициент диффузии, равна половинной невязке между производной по времени второго выборочного момента ВПФР и его эволюции в силу уравнения Лиувилля (3). Уравнение (13) будем называть *эмпирическим уравнением Фоккера—Планка*.

Может показаться, что замена уравнения Лиувилля (3) на уравнение Фокера—Планка (13) ничем не лучше, поскольку в момент времени  $t$  производная  $\dot{q}(t)$  не известна, так как для ее независимого определения нужно знать значение ряда в момент  $t + 1$ . Тем не менее модель (13) можно использовать на практике после небольшой коррекции, состоящей в

замене  $\lambda(t)$  на  $\lambda(t-1)$ . На следующем шаге по времени этот коэффициент, естественно, пересчитывается с учетом нового значения  $x(t+1)$ . Ошибка, которая вносится при такой коррекции, часто оказывается величиной порядка  $o(\varepsilon)$ . Это связано с тем, что для  $\varepsilon$ -стационарных ВПФР, т.е. таких, для которых выполнено условие  $V(T, 1; t) = \|f(x, t) - f(x, t+1)\| \leq \varepsilon$ , изменение второго момента на  $[0; 1]$  сравнительно мало. Действительно, пусть  $N$  есть число промежутков разбиения отрезка  $[0; 1]$  при практическом вычислении интеграла нормы  $V(T, \tau; t)$ , и  $f(i, t)$  есть значение ВПФР в  $i$ -м промежутке. Тогда из того, что

$$\forall i \quad |f(i, t) - f(i, t+1)| \leq \sum_{i=1}^N |f(i, t) - f(i, t+1)| \leq \varepsilon,$$

следует оценка

$$\begin{aligned} |q(t+1) - q(t)| &= \left| \int x^2 f_0(x, t+1) dx - \int x^2 f_0(x, t) dx \right| \leq \\ &\leq \int x^2 |f_0(x, t+1) - f_0(x, t)| dx \leq \varepsilon \int_0^1 x^2 dx = \frac{\varepsilon}{3}. \end{aligned} \quad (14)$$

В то же время, как нетрудно убедиться, невязка производной второго момента ВПФР с вычисленной производной по уравнению Лиувилля составляет величину порядка  $\varepsilon$ . Таким образом, даже скорректированное уравнение Фоккера—Планка дает более точное приближение к эволюции моментов ВПФР, чем уравнение Лиувилля.

#### 4. Система уравнений гидродинамического типа и ее замыкание

Кинетическая модель эволюции ВПФР  $f_0(x, t)$ , основанная на уравнении Фоккера—Планка (12) и замкнутая на уравнении для двумерной ВПФР  $f_1(x, \dot{x}, t)$  так, как это сделано в (10), имеет вид

$$\begin{aligned} \frac{\partial f_0(x, t)}{\partial t} + \frac{\partial}{\partial x} u(x, t) f_0(x, t) - \lambda(t) \frac{\partial^2 f_0(x, t)}{\partial x^2} &= 0, \quad \frac{\partial f_1(x, \dot{x}, t)}{\partial t} + \dot{x} \frac{\partial f_1(x, \dot{x}, t)}{\partial x} = 0; \\ u(x, t) f_0(x, t) &= \int \dot{x} f_1(x, \dot{x}, t) d\dot{x}, \quad f_0(x, t) = \int f_1(x, \dot{x}, t) d\dot{x}. \end{aligned} \quad (15)$$

Заметим, что сама ВПФР  $f_0(x, t)$  построена по соответствующему оптимальному объему выборки при прогнозировании на 1 шаг вперед, как описано в п. 1. Вычисление же совместной ВПФР  $f_1(x, \dot{x}, t)$  требует гораздо большего объема машинной памяти, не говоря уже о работе со следующим приближением (обрыве на третьем уравнении цепочки (2.5) и т.д.). Поэтому практически важной задачей является создание эволюционной модели непосредственно в терминах одномерной ВПФР. Из (15) следует, что для этого надо определить ЭЛС  $u(x, t)$ . Это можно сделать, либо найдя  $f_1(x, \dot{x}, t)$  из оборванного уравнения Лиувилля для парной ВПФР, что требует больших вычислительных затрат, либо получить уравнение для самой  $u(x, t)$ .

Используя второе и третье уравнения системы (15), получаем

$$\frac{\partial (u f_0)}{\partial t} = \int \dot{x} \frac{\partial f_1(x, \dot{x}, t)}{\partial t} d\dot{x} = - \int \dot{x}^2 \frac{\partial f_1(x, \dot{x}, t)}{\partial x} d\dot{x} \equiv - \frac{\partial (e(x, t) f_0(x, t))}{\partial x}, \quad (16)$$

где  $e(x, t)$  представляет, как это видно из (16), плотность второго момента по скорости для ВПФР  $f_1(x, \dot{x}, t)$ . Таким образом, эволюция  $u(x, t)$  (т.е. первого момента  $f_1(x, \dot{x}, t)$ ) определяется вторым моментом. Возникает эволюционная цепочка для моментов, решение которой эквивалентно построению эволюции самой ВПФР в предположении, что проблема моментов однозначно разрешима. Но как следует из оценки (14), при нахождении приближенной эволюции ВПФР нет необходимости строить уравнения эволюции для моментов всех порядков, поскольку, чем выше порядок момента, тем меньше его вклад в эволюцию ВПФР. Следовательно, моментную цепочку также можно оборвать на каком-либо порядке,

считая интеграл в правой части известным. Тогда получается замкнутая система уравнений гидродинамического типа. Например, при замыкании на третьем уравнении (т.е. на эволюции момента  $e(x, t)$ ) получаем

$$\begin{aligned} \frac{\partial f_0(x, t)}{\partial t} + \frac{\partial}{\partial x} u(x, t) f_0(x, t) - \lambda(t) \frac{\partial^2 f_0(x, t)}{\partial x^2} &= 0, \\ \frac{\partial}{\partial t} u(x, t) f_0(x, t) &= -\frac{\partial}{\partial x} e(x, t) f_0(x, t), \quad \frac{\partial}{\partial t} e(x, t) f_0(x, t) = -\frac{\partial}{\partial x} g(x, t) f_0(x, t), \end{aligned} \quad (17)$$

где  $g(x, t) = \int \dot{x}^3 f_1(x, \dot{x}, t) d\dot{x}$  считается известной функцией. Например, эту функцию, как и величину  $\lambda(t)$ , можно считать известной с предыдущего шага по времени.

## 5. Пример модели прогнозирования нестационарного временного ряда

Опишем кратко результаты применения модели (17) к задаче прогнозирования конкретного временного ряда, образованного суточными данными  $c(t)$  курса ЦБ рубль/доллар за период с 2005 по 2010 гг. по данным [6]. Обозначим

$$M = \max_t (c(t+1) - c(t)), \quad m = \min_t (c(t+1) - c(t)) \quad (18)$$

и рассмотрим нормированные приросты курса:

$$x(t) = \frac{c(t+1) - c(t) - m}{M - m}. \quad (19)$$

Ряд  $x(t)$  существенно нестационарный, что может быть проверено стандартными методами, например по критерию Колмогорова—Смирнова [7]. Так если две выборки объемом в 500 данных, расположенные встык, взяты из одной генеральной совокупности, то они отличаются в норме непрерывных функций не более чем на 0,1 с вероятностью 0,9. Однако оказалось, что из 600 таких выборок, построенных скользящим образом по рассматриваемым данным, лишь 28 (менее 0,05) отличаются менее чем на 0,1, тогда как в среднем отличие между ними составило 0,23, а 0,9-квантиль распределения расстояний между выборками лежит от 0,11 до 0,35.

Исходный ряд  $c(t)$  является настолько нерегулярно меняющимся, что наилучшим его стандартным прогнозом (т.е. адаптивным, регрессионным или авторегрессионным, скользящим средним и т.п.) оказывается наивный (завтра будет как сегодня), ошибка которого равна среднеквадратичному относительному приросту и составляет для изучаемого фрагмента примерно 0,008. Ошибка вышеперечисленных стандартных прогнозов по оптимальному для них объему данных заключена в диапазоне от 0,009 до 0,011, т.е. все они работают хуже наивного прогноза.

Рекуррентно разрешимая модель для построения ВПФР  $f_0(x, t)$ , имеющая вид нелинейной динамической системы, получается из (17) путем дискретизации области значений  $x(t)$ . Величина ячейки при этом должна иметь порядок  $o(\varepsilon)$ . Например, при  $\varepsilon = 0,05$  можно взять 100 ячеек равномерного разбиения. Тогда, обозначая тильдой рассчитываемые величины, получаем в праворазностном координатном шаблоне следующую систему:

$$\begin{aligned} \mathbf{f}: \quad & \tilde{f}(x, t+1) = f(x, t) + f(x, t) \tilde{u}(x, t) - f(x+1, t) \tilde{u}(x+1, t) + \\ & + \frac{\lambda(t)}{2} (f(r+1, t) - 2f(r, t) + f(r-1, t)); \\ \mathbf{u}: \quad & (\tilde{u}(x, t) - u(x, t-1)) f(x, t-1) = \\ & = u(x, t-1) (u(x+1, t-1) f(x+1, t-1) - u(x, t-1) f(x, t-1)) + \\ & + \tilde{e}(x, t) f(x, t-1) - \tilde{e}(x+1, t) f(x+1, t-1); \\ \mathbf{e}: \quad & (\tilde{e}(x, t) - e(x, t-1)) f(x, t-1) = \\ & = e(x, t-1) (u(x+1, t-1) f(x+1, t-1) - u(x, t-1) f(x, t-1)) + \\ & + g(x, t-1) f(x, t-1) - g(x+1, t-1) f(x+1, t-1). \end{aligned} \quad (20)$$

Анализ горизонтной статистики для ряда  $x(t)$  величин, определенных в (19), дает оптимальный объем для прогнозирования на 1 шаг вперед, равный в среднем  $T_{\text{opt}} = 290$  данных. Эволюция такой ВПФР по системе уравнений (20) отличается от фактической на 0,003, что более чем в 2 раза меньше, чем верхняя грань ошибки, равная, согласно (2),  $2/T_{\text{opt}}$ . Прогноз самого ряда как среднего значения прогнозной ВПФР имеет относительную ошибку отклонения от факта, равную в среднем квадратичном по рассматриваемому фрагменту данных 0,004. Это в 2 раза лучше, чем ошибка наивного прогноза.

Таким образом, существуют примеры рядов, когда предложенная модель имеет заметные преимущества в точности перед стандартными методами прогнозирования, корректно применяемыми лишь для стационарных рядов. Дальнейшее уточнение модели возможно на пути более полного учета высших моментов ВПФР совместных приращений высших порядков для случайных величин. Ограничением здесь является лишь производительность вычислительных комплексов, поскольку с увеличением размерности объем анализируемых данных стремительно возрастает. Например, анализ одной только горизонтной статистики для трехмерной задачи (с учетом разностных ускорений) на том же объеме данных (порядка 1000) требует использования массива  $10^{16}$  точек. Заметим, что объем выборки в тысячу данных сравнительно мал, на практике более востребованы задачи по анализу выборок объемом несколько миллионов данных. Разработка соответствующих численных алгоритмов представляется актуальной и практически важной задачей.

## Литература

1. Орлов Ю.Н., Осминин К.П. Методика определения оптимального объема выборки для прогнозирования нестационарного временного ряда // Информационные технологии и вычислительные системы. — 2008. — № 3. — С. 3–13.
2. Орлов Ю.Н., Осминин К.П. Построение выборочной функции распределения для нестационарного временного ряда // Мат. мод. — 2008. — №9. — С. 23–33.
3. Орлов Ю.Н., Осминин К.П. Нестационарные временные ряды. — М.: Эдиториал УРСС, Книжный дом ЛИБРОКОМ, 2011.
4. Власов А.А. Статистические функции распределения. — М.: Наука, 1966.
5. Козлов В.В., Смолянов О.Г. Бесконечномерные уравнения Лиувилля относительно мер // Доклады РАН. — 2010. — Т. 432. — № 1. — С. 28–32.
6. Данные обменных курсов валют // [www.finam.ru](http://www.finam.ru)
7. Гнеденко Б.В. Курс теории вероятностей. — М.: Физматлит, 1961.

Поступила в редакцию 22.02.2012.