

УДК 519.258

ОЦЕНКА И ПРИМЕНЕНИЕ МОДЕЛЕЙ ВРЕМЕННЫХ РЯДОВ С ДОЛГОЙ ПАМЯТЬЮ В ЭКОНОМИЧЕСКИХ ЗАДАЧАХ

Л. А. Осипов,

доктор техн. наук, профессор

А. М. Кричевский,

аспирант

Санкт-Петербургский государственный университет аэрокосмического приборостроения

Рассмотрены модели временных рядов, характеризующихся наличием долговременной зависимости (долгой памяти). Для идентификации таких рядов предложено использовать модели класса ARIMA (p, d, q) с дробным показателем d. Показаны пути оценки параметра «памяти» временного ряда, решения задачи прогнозирования в таких рядах.

Time series models with long memory are considered. For the identification of such time series the ARIMA (p, d, q) model with fractional parameter d is proposed. Ways of estimation of the «memory» parameter and methods of forecasting in such time series are also studied.

Изучение временной структуры рядов различной природы, например экономика, телекоммуникация, астрономия и т. п., играет ключевую роль в моделировании и получении прогнозных оценок в различных областях науки и техники. В последнее время значительный интерес проявился к временным рядам (ВР), которые можно охарактеризовать термином «временные ряды с долгой памятью» — *time series with long memory* [1]. Существующие синонимы для этих рядов: долговременная зависимость (*long-range dependence*), сильная зависимость (*strong dependence*) или персистентность (*persistence*). Под рядами с долгой памятью понимаются не только стационарные ряды, но также и нестационарные, в которых зависимость от времени спадает очень медленно.

Естественно предположить, что долгая память может быть обнаружена в данных, занимающих достаточно большой промежуток времени. Но, как и в других областях статистики, теория конечной выборки является обычно математически трудной даже в простых моделях и требует строгих допущений. В теории большой выборки необходимо обеспечить правила вывода, которые становятся более надежными при увеличении объема выборки. Однако эта теория может поставить значительные более трудные математические проблемы в долговременной памяти временных рядов, чем в кратковременной памяти.

Долгая память обычно описывается в виде автоковариаций или спектральной плотности. Положим, что X_t , $t = 0, \pm 1, \dots$ является временным рядом. Если ряд стационарный, то $E(X_t) = \mu$, $\text{cov}(X_t, X_{t+j}) = \rho(j)$ не зависят от t . В случае, если X_t имеет непрерывную функцию распределения, то его спектральная плотность выражается следующим образом:

$$S(f) = \frac{1}{2\pi} \sum_{j=-\infty}^{\infty} \rho(j) e^{-ijf}, \quad -\pi < f < \pi. \quad (1)$$

Здесь $S(f)$ — неотрицательная четная функция с периодом 2π при ее продолжении за диапазон Найквиста $[-\pi, +\pi]$.

Принято считать, что ряд X_t имеет долговую память, если

$$S(0) = \frac{1}{2\pi} \sum_{j=-\infty}^{\infty} \rho(j) = \infty, \quad (2)$$

т. е. $S(f)$ имеет полюс на нулевой частоте.

Противоположная ситуация с нулевым значением спектральной плотности на нулевой частоте:

$$S(0) = \frac{1}{2\pi} \sum_{j=-\infty}^{\infty} \rho(j) = 0, \quad (3)$$

определяет отрицательную зависимость или антиперсистентность.

Учитывая выражения (2), (3), можно сказать, что ряд X_t имеет короткую память, если

$$0 < S(0) < \infty.$$

Кратко поясним возникновение новой модели, отталкиваясь от методологии Бокса—Дженкинса [2]. Наиболее распространенными моделями для стационарных ВР являются модели авторегрессии и скользящего среднего. Авторегрессионную модель порядка p , которая сокращенно обозначается $AR(p)$ (*autoregressive process*), можно записать в виде

$$X_t = \Phi_1 X_{t-1} + \Phi_2 X_{t-2} + \dots + \Phi_p X_{t-p} + a_t, \quad (4)$$

где $\Phi_1, \Phi_2, \dots, \Phi_p$ — весовые коэффициенты.

Выражение (4) определяет *процесс авторегрессии порядка p* , в котором текущее значение ряда в момент t выражается через конечное число прошлых значений и величину возмущения a_t , не зависящую от прошлого. С помощью оператора сдвига $B = X_{t-1}/X_t$ модель (4) можно записать в эквивалентной форме

$$(1 - \Phi_1 B - \Phi_2 B^2 - \dots - \Phi_p B^p) y_t = a_t,$$

которая после введения оператора авторегрессии $\Phi(B)$ принимает вид

$$\Phi(B) X_t = a_t. \quad (5)$$

Модель скользящего среднего (*moving average*) предполагает, что в ошибках модели в предшествующие периоды сосредоточена информация по всей предыстории ряда. Эта модель порядка q запишется в виде

$$X_t = a_t - \theta_1 a_{t-1} - \dots - \theta_q a_{t-q}, \quad (6)$$

где символы $\theta_1, \dots, \theta_q$ используются для обозначения конечного набора весовых параметров.

Соотношение (6) определяет *процесс скользящего среднего порядка q* , или сокращенно $MA(q)$, который представляет собой линейную комбинацию текущего и прошлых значений шума a_t . Используя оператор сдвига, можно записать для процесса (6) эквивалентное выражение

$$X_t = (1 - \theta_1 B - \dots - \theta_q B^q) a_t \equiv \theta(B) a_t. \quad (7)$$

Моделями $AR(p)$ и $MA(q)$ за счет выбора их порядков p и q можно удовлетворительно описывать многие реальные процессы. Однако на практике для достижения большей гибкости в подгонке моделей к наблюдаемым ВР иногда целесообразно объединить в одной модели и авторегрессию, и скользящее среднее; при этом цель должна состоять в построении наиболее экономных моделей, дающих хорошую аппроксимацию с помощью небольшого числа параметров. Достижению этого помогает рассмотрение *смешанных моделей авторегрессии — скользящего среднего*, т. е. моделей $ARMA(p, q)$, которые имеют вид

$$X_t = \Phi_1 X_{t-1} + \dots + \Phi_p X_{t-p} + a_t - \theta_1 a_{t-1} - \dots - \theta_q a_{t-q}, \quad (8)$$

или в другой эквивалентной форме через операторы $\Phi(B), \theta(B)$

$$\Phi(B) y_t = \theta(B) a_t. \quad (9)$$

Модель (8) может интерпретироваться как линейная модель множественной регрессии, в которой в качестве объясняющих переменных выступают прошлые значения самой зависимой переменной, а в качестве регрессионного остатка — скользящие средние из элементов белого шума.

Существует неограниченное число различных проявлений нестационарности. Однако можно выделить обширный класс встречающихся в приложениях ВР со специфической однородной нестационарностью, которая удовлетворительно описывается стохастической моделью, являющейся модифицированной формой модели $ARMA$. Условие стационарности модели (9) означает, что корни полинома $\Phi(B)$ лежат вне единичного круга [2]. Естественный путь получения нестационарного процесса, описываемого таким же уравнением, заключается в ослаблении этого ограничения. В частности, оказывается, что во многих случаях наблюдаемые в реальности процессы хорошо описываются моделями типа (9), у которых один или несколько корней $\Phi(B)$ равны единице. Такой класс моделей называется *процессами авторегрессии — проинтегрированного скользящего среднего*. В английской аббревиатуре такой процесс запишется как $ARIMA$ (с добавлением к уже известному сокращению $ARMA$ слова *integrated*).

Рассмотрим модель

$$\varphi(B) X_t = \theta(B) a_t, \quad (10)$$

где в отличие от равенства (9) $\varphi(B)$ — нестационарный оператор авторегрессии порядка $p + d$, такой, что d корней уравнения $\varphi(B) = 0$ равны единице, а остальные p корней лежат вне единичного круга; оператор же скользящего среднего $\theta(B)$ по-прежнему обладает порядком q и является обратимым (все его корни лежат вне единичного круга).

Тогда можно записать, что

$$\varphi(B) = \Phi(B)(1 - B)^d,$$

где $\Phi(B)$ — уже стационарный порядка p оператор авторегрессии (т. е. с корнями вне единичного круга). Если ввести оператор разности $\Delta = 1 - B, \Delta X_t = X_t - X_{t-1}$, то $\varphi(B)$ запишется как $\varphi(B) \Delta^d X_t = \theta(B) a_t$ и модель (10) можно представить в виде

$$\Phi(B) \Delta^d X_t = \theta(B) a_t. \quad (11)$$

Здесь d -я разность ряда X_t вычисляется по формуле

$$w_t \equiv \Delta^d X_t = (1 - B)^d X_t$$

и, следовательно, удовлетворяет уравнению

$$\Phi(B) w_t = \theta(B) a_t, \quad (12)$$

т. е. является уже стационарным обратимым процессом $ARMA(p, q)$.

Таким образом, процесс X_t , описываемый уравнением (11), можно получить d -кратным суммированием (или интегрированием) процесса $\{w\}$, являющегося в соответствии с (12) процессом $ARMA$. Вследствие этого процесс, задаваемый моделью (11), называют процессом $ARIMA$. Если в формуле (11) оператор авторегрессии $\Phi(B)$ имеет порядок p , а оператор скользящего среднего $\theta(B)$ — порядок q , то кратко модель (11) записывается как $ARIMA(p, d, q)$. В частности, при $d = 0$ эта более общая модель сводится к смешанной модели $ARMA(p, q)$. Тем самым модель $ARIMA(p, d, q)$ охватывает широкий класс как стационарных (при $d = 0$), так и нестационарных (при $d \geq 1$) процессов.

В работах [3, 4] было впервые предложено рассмотреть дробные значения d из интервала $d \in (-1/2, 1/2)$, что привело к дробной (*fractional*) авторегрессионной модели скользящего среднего порядков p, d, q ($ARFIMA(p, d, q)$ или $FARIMA(p, d, q)$).

Примем, что X_t удовлетворяет следующему дифференциальному уравнению:

$$(1 - B)^d X_t = a_t, \quad a_t \sim N(0, \sigma_a^2). \quad (13)$$

При расширении показателя d в (13) до нецелых степеней результатом является ряд с преобразованием, которое включает в себя разложение члена $(1 - B)^d$ по биномиальной теореме для нецелых показателей:

$$(1 - B)^d = \sum_{k=0}^{\infty} (-1)^k \binom{d}{k} B^k,$$

где

$$\binom{d}{k} = \frac{d(d-1)(d-2)\dots(d-k-1)}{k!}.$$

Применяя это разложение к X_t , получаем

$$(1 - B)^d X_t = \sum_{k=0}^{\infty} (-1)^k \binom{d}{k} B^k X_t = \sum_{k=0}^{\infty} A_k X_{t-k} = a_t,$$

где коэффициенты авторегрессии A_k выражаются через гамма-функцию:

$$A_k = (-1)^k \binom{d}{k} = \frac{\Gamma(k+d)}{\Gamma(-d)\Gamma(k+1)}.$$

Грейнджер и Хоскинг показали [3, 4], что характеристики таких временных рядов обладают важными свойствами: например, X_t является стационарным и обратимым для $d \in (-1/2, 1/2)$. Кроме того, оказывается, что положительная или отрицательная зависимость определяются знаком при параметре d , т. е. автокорреляционные коэффициенты процесса X_t имеют тот же знак, что и d .

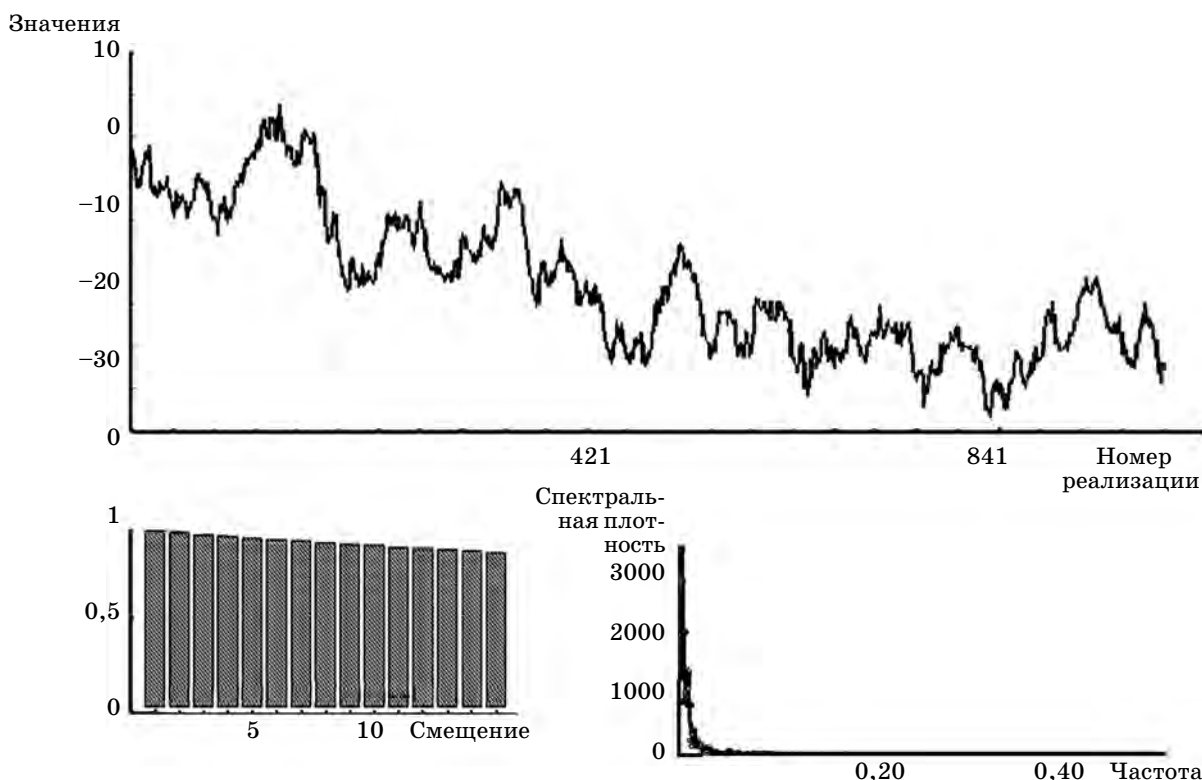
Медленный спад автокорреляций объясняется тем, что при положительном d сумма последних сходится к бесконечности, а при отрицательном d — к нулю.

Простейшей реалистической моделью для стационарного ряда является параметрическая модель, которая выражает $\rho(j)$ для всех j и $S(f)$ для всех f как параметрическую функцию только двух параметров: d и неизвестного масштабного фактора. Возможно, что самой ранней такой моделью явился фрактальный шум, концепция которого возникла из рассмотрения самоподобия.

Определим самоподобный процесс следующим образом [1]: непрерывный стохастический процесс $\{y(t); -\infty < t < \infty\}$ является самоподобным с параметром самоподобия $H \in (0; 1)$, если для любого $a > 0$ процесс $\{y(at); -\infty < t < \infty\}$ имеет то же распределение, что и процесс $\{a^H y(t); -\infty < t < \infty\}$.

У каждого реального самоподобного процесса должен быть наибольший и наименьший масштаб: нельзя бесконечно увеличивать или уменьшать масштаб. Например, в случае броуновского движения (БД), представляющего образец самоподобного процесса, диапазон масштабов, в пределах которого сохраняется самоподобие, охватывает много порядков величины: от размеров сосуда с жидкостью (допустим, 0,1 м) до длины свободного пробега молекул между столкновениями, которая для малых частиц может достигнуть 10^{-9} м. Во многих случаях объект считается самоподобным, если его можно масштабировать с коэффициентом подобия 10 и меньше (до трех дискретных шагов) [5]. Оценим спектральную плотность броуновской функции $S(f)$, которую определим как проекцию БД на одно пространственное направление в зависимости от времени. БД порождается независимыми приращениями и имеет плоский спектр. Следовательно, сумма (интеграл) приращений обладает спектральной плотностью, пропорциональной f^{-2} . Отметим, что в общем случае зависимость спектральной плотности от частоты характеризуется степенным законом вида $f^{-\beta}$. Среди шумов большой известностью пользуется белый шум со спектральным показателем $\beta = 0$. Иначе говоря, спектр белого шума не зависит от частоты. Проинтегрировав белый шум один раз по времени, получаем коричневый шум (проекцию БД на одно пространственное измерение), который имеет спектральную плотность, пропорциональную f^{-2} . Но белый и коричневый шумы далеко не исчерпывают все спектральные возможности: между ними располагается розовый шум со спектром f^{-1} , а за коричневым — черный, пропорциональный $f^{-\beta}$, где $\beta > 2$.

В качестве примера самоподобного процесса приведем результаты моделирования коричневого и розового шумов. Генерирование коричневого шума сводится к суммированию независимых случайных чисел и реализуется сравнительно легко посредством табличного расчета в Excel. На рис. 1 приведена реал-



■ Рис. 1. Реализация процесса коричневого шума вместе с автокорреляционной функцией и спектральной плотностью

лизация процесса, определяющего коричневый шум вместе с автокорреляционной функцией и спектральной плотностью: и автокорреляционная функция, и спектральная плотность свидетельствуют о наличии долгой памяти в коричневом шуме.

Сравнительно простой метод генерирования розового шума состоит в том, чтобы сложить несколько релаксационных процессов со значениями времен релаксации τ , образующими самоподобную прогрессию с коэффициентом подобия 10 (или еще меньше — для лучшей сходимости).

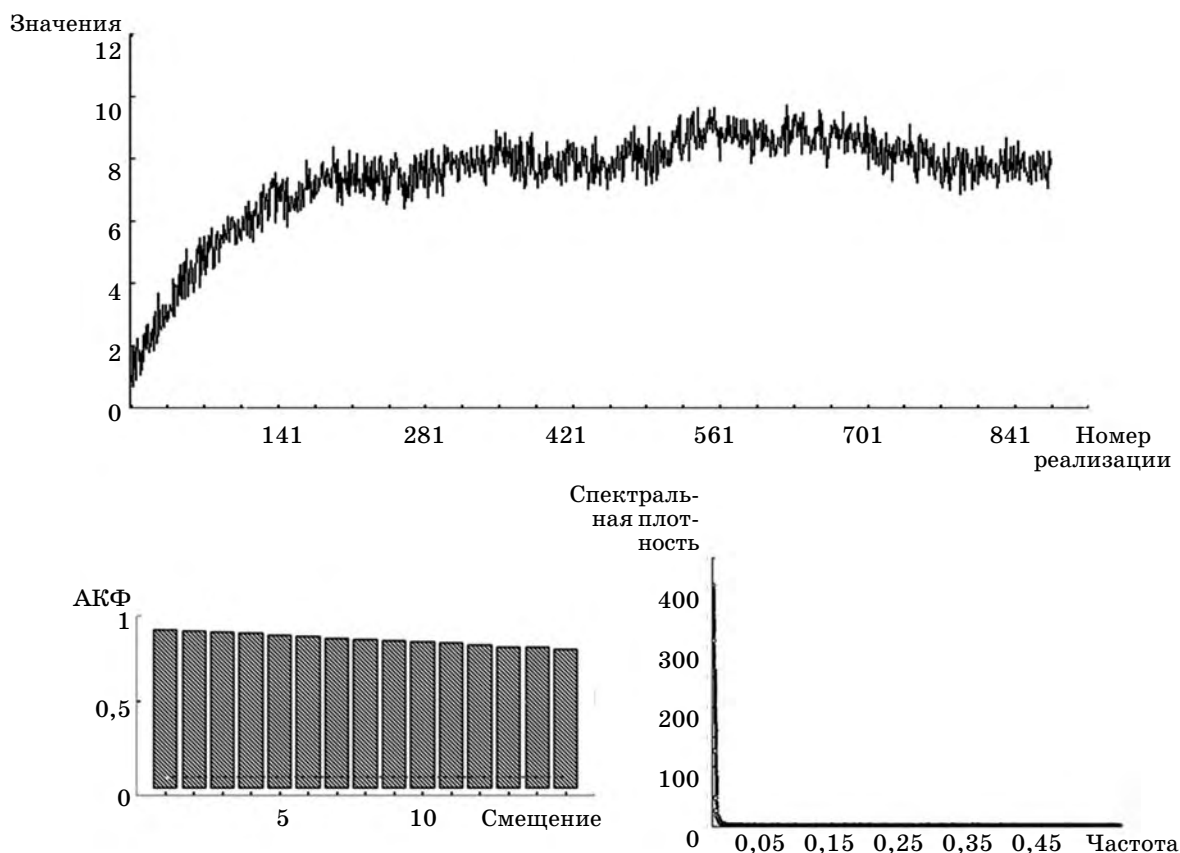
Релаксационный процесс с дискретными значениями времени x_n можно задавать с помощью генератора случайных чисел, который позволяет получать независимые случайные числа r_n , подставляемые затем в рекуррентное соотношение [5]

$$x_{n+1} = \rho x_n + \sqrt{1 - \rho^2} r_n, \quad x_0 = 0, \quad (14)$$

где ρ — требуемый коэффициент корреляции между соседними случайными значениями. Со временем релаксации τ этот коэффициент связан соотношением $\rho = \exp(-1/\tau)$. Таким образом, для набора значений времени релаксации, каждое из которых в 10 раз превосходит предыдущее ($\tau = 1, 10, 100, \dots$), коэффициенты корреляции получаются вычислением последовательных корней десятой степени (т. е. $\rho = 0,37; 0,90; 0,99, \dots$). Результаты

расчета, выполненные по (14), приведены на рис. 2 вместе с автокорреляционной функцией и спектральной плотностью.

При использовании модели класса *ARFIMA* (p, d, q) важно правильно определить параметры этой модели. Неверное определение параметров p и q приводит к несогласованной оценке коэффициентов *AR*- и *MA*-моделей, но ошибка в оценке d дает неверную интерпретацию обеих моделей из-за потери идентификации. Асимптотическое поведение спектра указывает на то, что при моделировании рядов с короткой памятью несущественно влияние очень низких частот и очень длинных лагов, т. е. в ситуациях доминирования d . Вследствие этого становится понятно, что оценки d должны основываться на информации о низких частотах или длинных лагах. С точки зрения требований устойчивости, оценки должны быть основаны на очень малой части данных при увеличении объема выборки данных, поэтому естественно ожидать более медленной скорости сходимости, чем для оценок, полученных по модели с целыми значениями параметров. Однако в очень длинных временных рядах (в экономике, метеонаблюдениях) доступное число степеней свободы может быть достаточным для обеспечения адекватной точности. Такие оценки обычно называются *полупараметрическими* (*semiparametric*) [1].



■ Рис. 2. Реализация процесса розового шума вместе с автокорреляционной функцией и спектральной плотностью

Существует несколько методов для оценки параметра d , являющихся в то же время и тестами для обнаружения долгой памяти во временных рядах. Параметр d можно оценить во временной или частотной областях [6].

Во временной области для получения оценки d используется асимптотическое представление для ковариаций [1]

$$\rho(j) \sim c_1 j^{2d-1} \text{ при } j \rightarrow \infty,$$

где $c_1 > 0$ — постоянная; j — временной лаг.

В этом случае для формирования оценки d могут быть применены несколько подходов:

- нелинейная регрессия выборочных автоковариаций;
- обычный метод наименьших квадратов для построения регрессии задержанных выборочных автоковариаций;
- метод максимального правдоподобия.

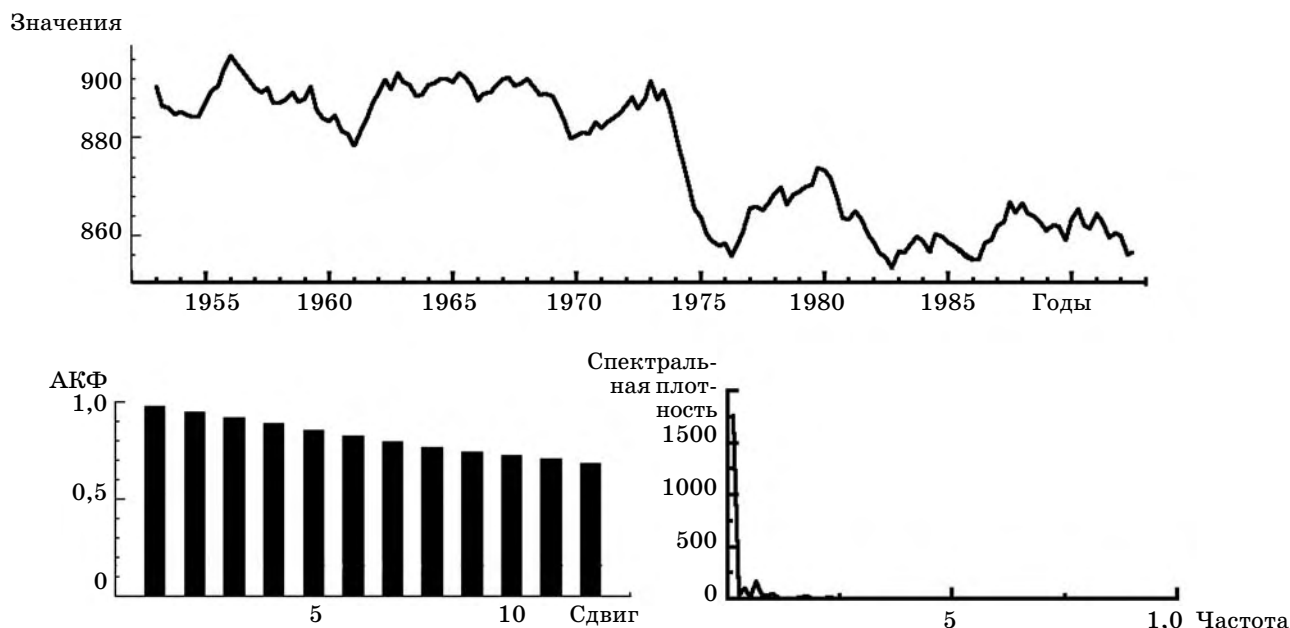
Однако распределения этих оценок достаточно сложны, поэтому использование предложенных оценок носит ограниченный характер.

В частотной области одна из первых оценок d была предложена в работе [7]. Сущность метода заключа-

ется в построении уравнения регрессии логарифма периодограммы на низких частотах как функции частоты: ожидаемый наклон зависит от параметра d . Использовались только несколько первых ординат периодограммы, и авторы работы пришли к выводу, что результирующая оценка регрессии для d может описать характеристику долгой памяти ВР без искажения ее свойствами краткой памяти процесса.

Для экспериментальной проверки изложенного в отношении ВР с долгой памятью здесь воспользуемся данными из экономики за 1953–1992 гг. Изменение выбранного параметра во времени показано на рис. 3. Характер автокорреляционной функции и периодограммы, являющейся оценкой спектральной плотности, указывает на то, что рассматриваемый временной ряд, скорее всего, обладает долгой памятью и характер его изменения можно описать моделью *ARFIMA* (p, d, q).

Для анализа ВР, показанного на рис. 3, воспользуемся модулем динамического моделирования PcGive из программного пакета GiveWin2.20. Этот модуль позволяет идентифицировать и тестировать модель и оценивать ее параметры.

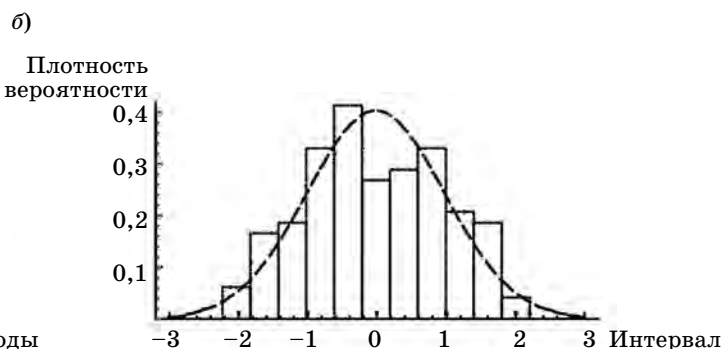
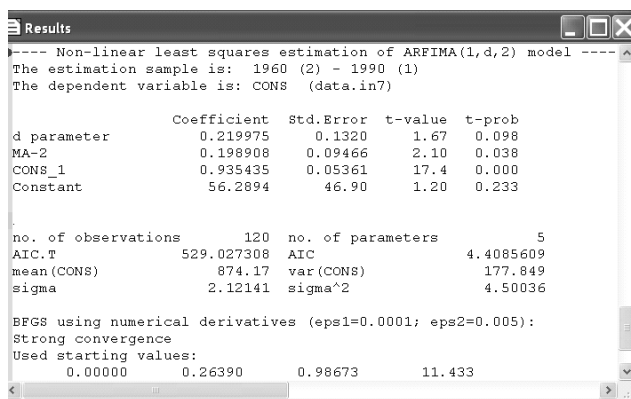


■ Рис. 3. Временной ряд, автокорреляционная функция, периодограмма

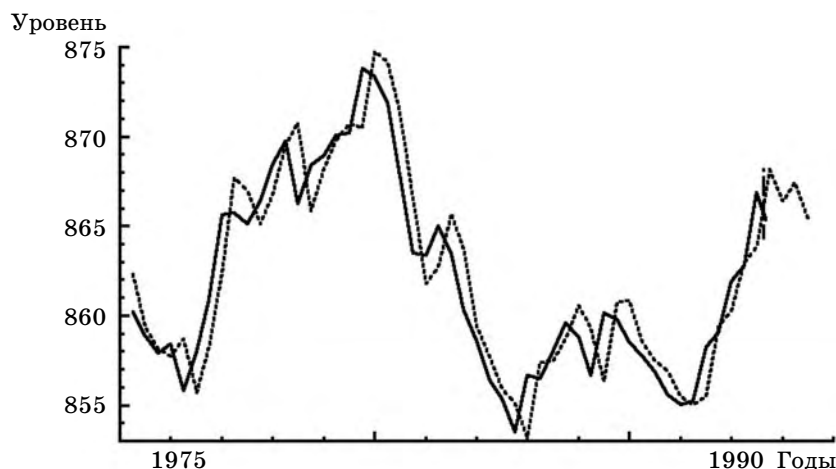
Для оценки параметров модели в данном модуле используются точный метод максимального правдоподобия (*Exact Maximum Likelihood — EML*) и нелинейный метод наименьших квадратов (*Nonlinear Least Squares — NLS*). Решение задачи выполняется в форме последовательного диалога за несколько шагов, в частности: выбор данных для анализа, установка параметров модели, оценка параметров модели. Для данного ряда была выбрана модель *ARFIMA(1, d, 2)*. Параметр *d* оценивался с помощью нелинейного метода наименьших квадратов. Результаты расчетов параметров модели приведены в таблице.

Значения *AR-1* и *MA-1* не приводятся, так как при редактировании модели были выбраны фик-

■ Результаты расчетов



■ Рис. 4. Исходный (—) и подобранный (---) ряды (а) и гистограмма остатков и кривая нормального распределения (б)



■ Рис. 5. Прогноз ряда на 4 интервала времени вперед: — — исходный ряд; - - - - — подобранный ряд

сированные лаги для $AR-1$ и $MA-1$, равные единице.

На рис. 4, а приведены исходный ВР и подобранный, который описывается моделью $ARFIMA(1, d, 2)$ при $d = 0,22$. На рис. 4, б приведена гистограмма остатков для проверки адекватности модели.

Подобранную модель можно использовать и для прогноза. Например, на рис. 5 видно, что предсказанные значения сохраняют тенденцию ряда.

Таким образом, в работе показана возможность использования моделей ВР с долгой памятью для анализа и прогнозирования.

Литература

1. Time Series with Long Memory / Ed. P. M. Robinson. Oxford University Press, 2003.
2. Бокс Дж., Дженкинс Г. Анализ временных рядов. Прогноз и управление. Вып. 1. М.: Мир, 1974.
3. Granger C. W., Joyeux R. E. An introduction to long-memory series models and fractional differencing // Journal of Time Series Analysis. 1980. Vol. 1. P. 15–29.
4. Hosking J. R. M. Fractional differencing // Biometrika. 1981. Vol. 68. P. 165–176.

5. Шредер М. Фракталы, хаос, степенные законы / НИЦ «Регулярная и хаотическая динамика». Ижевск, 2005.
6. Breidt F. J., Crato N., Lima P. The detection and estimation of long memory in stochastic Volatility // Journal of Econometrics. 1998. Vol. 73. P. 325–348.
7. Geweke J., Porter-Hudak S. The estimation of long memory time series models // Journal of Time Series Analysis. 1983. Vol. 4. P. 221–238.