



Юрий Балагула

Юлия Абакумова

Длинная память
на рынке нефти:
спектральный подход

Препринт Ес-01/11

Факультет экономики

Санкт-Петербург
2011

УДК 330.43
ББК 65.01
Б20

Балагула Ю.М., Абакумова Ю.А.

Б20 **Длинная память на рынке нефти: спектральный подход** / Юрий Балагула, Юлия Абакумова : Препринт Ес-01/11; Факультет экономики. — СПб.: ЕУСПб, 2011. — 40 с.

В данной работе предлагается спектральный подход к оценке эффекта длинной памяти во временных рядах и его практическое применение для анализа цен на нефть.

Информация об авторах:

Балагула Юрий Моисеевич — слушатель третьего курса факультета экономики Европейского университета в Санкт-Петербурге, ubalagula@eu.spb.ru;

Абакумова Юлия Александровна — слушатель третьего курса факультета экономики Европейского университета в Санкт-Петербурге, yabakumova@eu.spb.ru.

Издание осуществлено за счет средств проекта создания специализации по природным ресурсам и экономике энергетики «ЭксонМобил».

© Ю.М. Балагула, Ю.А. Абакумова, 2011

ДЛИННАЯ ПАМЯТЬ НА РЫНКЕ НЕФТИ: СПЕКТРАЛЬНЫЙ ПОДХОД

Ю.М. Балагула, Ю.А. Абакумова

ВВЕДЕНИЕ

Важной задачей финансовой теории является моделирование динамики доходности и волатильности цен на сырьевых рынках. Результаты этих исследований могут быть применены для характеристики свойств рынков – в какой степени рынки являются эффективными, насколько складывающиеся цены адекватны фундаментальным показателям (нет ли пузырей) – то есть в какой степени рынки выполняют свою экономическую роль правильной оценки активов. С другой стороны, адекватное моделирование соответствующих временных рядов дает возможности для их прогноза, который рыночные агенты могут использовать в своих интересах для извлечения прибыли.

Долгое время было принято считать, что цены активов на финансовых рынках следуют процессу «случайного блуждания» и, следовательно, совершенно непредсказуемы. Эта точка зрения полностью соответствует гипотезе эффективного рынка, которая лежит в основании классической теории финансовых рынков, в частности, наиболее популярной модели ценообразования активов CAPM. Однако многочисленные исследования показывают, что зачастую временные ряды не удовлетворяют этой предпосылке из-за наличия эффекта «длинной памяти», который проявляется в том, что текущие значения временного ряда в значительной степени зависят от прошлых значений (Перцовский, 2003).

Для описания эффекта длинной памяти в последнее время были предложены разнообразные подходы – введена модель ARFIMA, разработано большое количество тестов, базирующихся на разных методологиях. В данной работе проводится исследование временных рядов цен на нефть с помощью спектрального подхода и их прогнозирование с использованием модели ARFIMA.

В первой главе вводятся понятия длинной памяти, персистентности, рассматривается их связь с теорией фракталов, концепциями эффективного и фрактального рынков.

Во второй главе раскрывается математическое содержание понятий длинной памяти и показателя Хёрста, излагаются основы преобразования Фурье и спектрального анализа временных рядов, описывается сущность предлагаемого авторами спектрального подхода к моделированию эффекта длинной памяти.

Третья глава содержит описание эконометрических инструментов анализа длинной памяти – модели авторегрессии и скользящего среднего с дробным порядком интегрирования (ARFIMA) и теста Geweke-Porter-Hudak (GPH), позволяющего определить дробный порядок интегрирования временного ряда.

Четвёртая глава включает в себя оценку моделей ARFIMA и моделей, не учитывающих наличие длинной памяти, для искусственного ряда со значительным дробным порядком интегрирования, сравнение полученных прогнозов согласно выбранным критериям сопоставления. Также проверяется гипотеза о наличии длинной памяти у ряда цен на нефть, строятся модели ARFIMA, ARIMA и GARCH, и проводится сопоставление их прогнозных свойств.

Для проведения необходимых тестов и построения прогнозов использовались пакеты R, Eviews и OxMetrics. Следует отметить, что

каждый из пакетов отличается набором удобных функций, не доступных другому пакету. EViews удобен в построении моделей ARIMA, проведении необходимых тестов на единичные корни. В R легко доступна функция генерирования ряда с дробным порядком интегрирования, но в то же время, еще не встроена команда для прогнозирования модели ARFIMA, которой обладает пакет OхMetrics с очень дружественным интерфейсом и легко осваиваемыми командами.

1. ДЛИННАЯ ПАМЯТЬ И ЭФФЕКТИВНОСТЬ РЫНКА

Концепция эффективности рынка играет весьма важную роль в теории финансовых рынков. Понятие эффективности означает, что совокупность информации, относящейся к оценке актива, торгуемого на рынке, мгновенно отражается в стоимости этого актива. Таким образом, цена, наблюдаемая на эффективном рынке, представляет собой корректную оценку фундаментальной стоимости актива, поскольку эта цена уже отражает всю информацию, связанную с прошлыми событиями и ожиданиями относительно будущих событий. В этих условиях изменения цен обусловлены главным образом какими-либо непредвиденными событиями. Отсюда можно сделать вывод, что изменения цен являются случайными, а сами цены следуют процессу «случайного блуждания», тогда как доходности представляют собой «белый шум» (Перцовский, 2003).

Однако эмпирические исследования показывают наличие существенных отклонений от этой модели. В частности, в рамках модели случайного блуждания доходности рассматриваются, как правило, одинаково независимо (и обычно нормально) распределенными, с нулевым средним и постоянной во времени дисперсией. На практике это не так –

соответствующие распределения имеют так называемые «толстые хвосты». Существуют и другие отклонения от идеальной модели – такие, как кластеризация волатильности, эффекты понедельника, января и т.п.

Весьма существенным эмпирическим наблюдением, ставящим под сомнение гипотезу эффективного рынка, является наличие длинной памяти во временных рядах доходностей или волатильностей цен активов. Наличие длинной памяти означает, что существует зависимость между изменениями цен, далеко отстоящими друг от друга во времени. Такая корреляционная структура открывает возможность для прогнозирования, что невозможно в случае эффективного рынка.

С понятием длинной памяти тесно связана концепция персистентности, введённая Хёрстом. Персистентность характеризует свойство временного ряда сохранять тенденцию изменения. Если временной ряд на протяжении больших промежутков времени в среднем возрастает или убывает - говорят, что он является *персистентным*. Если наоборот, он постоянно меняет знак приращений – он является *антиперсистентным* (Мандельброт, 2002). Чтобы количественно характеризовать это свойство, Хёрст придумал величину, ныне носящую его имя – показатель (экспонента) Хёрста, и соответствующий метод анализа – метод нормированного размаха (R/S-statistics, rescaled range analysis), получивший широкое распространение при анализе структуры памяти временных рядов.

Необходимо отметить, что понятия персистентности и памяти временных рядов часто анализируются в контексте теории фракталов. В работах основоположника этой теории (Мандельброт, 2002) показано, что наличие памяти и персистентности во временных рядах тесно связаны с наличием в их структуре самоподобия (точнее, самоаффинности) – главного характерного признака фрактального объекта. Поэтому

показатель Хёрста и другие характеристики памяти могут рассматриваться как некоторые фрактальные размерности или связанные с ними величины; и наоборот, аппарат фрактального анализа временных рядов и теория фракталов могут привлекаться для анализа финансовых временных рядов и моделирования соответствующих рынков в задачах экономической теории и эконометрики.

Следуя этому подходу, взамен традиционной гипотезы эффективного рынка, Э.Петерс разработал альтернативу – гипотезу фрактального рынка (Петерс, 2004). Она основывается на следующих предположениях:

- разные инвесторы имеют разные горизонты инвестирования;
- информация по разному влияет на инвесторов с разными горизонтами инвестирования;
- долгосрочные инвесторы больше учитывают фундаментальные факторы, краткосрочные базируются на техническом анализе, принимая решения об инвестировании.

Рынок, удовлетворяющий этим предположениям, имеет фрактальную (самоподобную) структуру и является стабильным. Стабильность рынка теряется при разрушении этой структуры, когда горизонт инвестирования становится одинаковым для всех инвесторов, например, если все инвесторы бегут с рынка вследствие каких-то критических новостей. Из этой теории следует, что в периоды стабильности рынка (в краткосрочном периоде) эволюция цен описывается моделью случайного блуждания, в долгосрочном периоде необходимы другие модели, типа обобщенного броуновского движения. Адекватной характеристикой временных рядов цен активов на таком рынке является показатель Хёрста, который одновременно может рассматриваться и как фрактальная размерность, и как характеристика памяти временного ряда в смысле наличия

персистентности или антиперсистентности. Например, в работе (Tabak, 2007) показатель Хёрста используется, чтобы определить, стал ли рынок сырой нефти (Brent and West Texas Intermediate) более эффективным после введения определённых новаций.

В качестве характеристики памяти временного ряда могут использоваться и другие величины, например, спектральный параметр β и параметр дробного интегрирования d (порядок интегрирования временного ряда). Ниже будет показано, что они связаны с показателем Хёрста и даже более удобны для практического применения.

2. ДЛИННАЯ ПАМЯТЬ ВО ВРЕМЕННЫХ РЯДАХ

Процессы Хёрста и персистентность

Когда говорят о наличии памяти во временных рядах, имеют в виду следующее. Традиционно рассматриваемые эконометрикой временные ряды, описываемые моделями типа ARMA, имеют *короткую память (short memory)*. Это означает, что их автокорреляционная функция убывает экспоненциально, то есть очень быстро. В противоположность этому, для рядов, обладающих свойством длинной памяти (*long memory* or *long-range-dependent processes*), автокорреляционная функция убывает крайне медленно, гиперболически. Основываясь на этом свойстве автокорреляционной функции, дают следующее строгое математическое определение длинной памяти:

Стационарный процесс X_t является процессом с длинной памятью, если существует действительное число α , $0 < \alpha < 1$ и константа $c > 0$, для которых выполняется условие

$$\lim_{k \rightarrow \infty} \frac{\rho_k}{ck^{-\alpha}} = 1,$$

где ρ – функция автокорреляции, k – номер лага. То есть, приближаясь к бесконечности, автокорреляционная функция убывает как степенная функция с коэффициентом α .

По-видимому, первым такие процессы ввёл в рассмотрение Хёрст, изучая статистику годового стока реки Нил. Чтобы описать отклонение процесса от привычного марковского, он ввёл понятие *персистентности*, характеризующее свойство временного ряда сохранять тенденцию изменения. Если временной ряд на протяжении больших промежутков времени в среднем возрастает или убывает - говорят, что он является *персистентным*. Если наоборот, он имеет тенденцию постоянно менять знак приращений – он является *антиперсистентным*. Чтобы количественно характеризовать это свойство, Хёрст предложил новый метод анализа временных рядов – метод нормированного размаха или R/S-анализ.

Рассмотрим выборку $\{y_1, \dots, y_n\}$ из стационарного процесса с длинной памятью. Пусть x_t - частичные суммы от $\{y_i\}$, то есть $x_t = \sum_{j=1}^t y_j$ для $t = 1 \dots n$ и s_n^2 – выборочная дисперсия $\{y_i\}$. R/S-статистика определяется следующим образом:

$$R_n = \frac{1}{s_n} \left[\max_{1 \leq t \leq n} \left(x_t - \frac{t}{n} x_n \right) - \min_{1 \leq t \leq n} \left(x_t - \frac{t}{n} x_n \right) \right]$$

То есть сначала исследуемая выборка численно интегрируется. Затем подсчитывается R_n – размах получившегося ряда (максимальное значение минус минимальное), нормированный среднеквадратичным отклонением исходной выборки. Процедура метода состоит в том, что из анализируемого временного ряда извлекается выборка точек, отстоящих друг от друга на k лагов $\{y_i, \dots, y_{i+k-1}\}$, $1 \leq i \leq n-k+1$ и для данной выборки рассчитывается R/S-статистика R_k .

Показатель Хёрста H определяется методом наименьших квадратов как коэффициент регрессии $\log R_k$ на $\log k$.

При $1 > H > 0.5$ случайный процесс является персистентным; при $0.5 > H > 0$ – антиперсистентным. Случай $H = 0.5$ означает классический «белый шум». Примеры соответствующих временных рядов приведены на рис. 1.

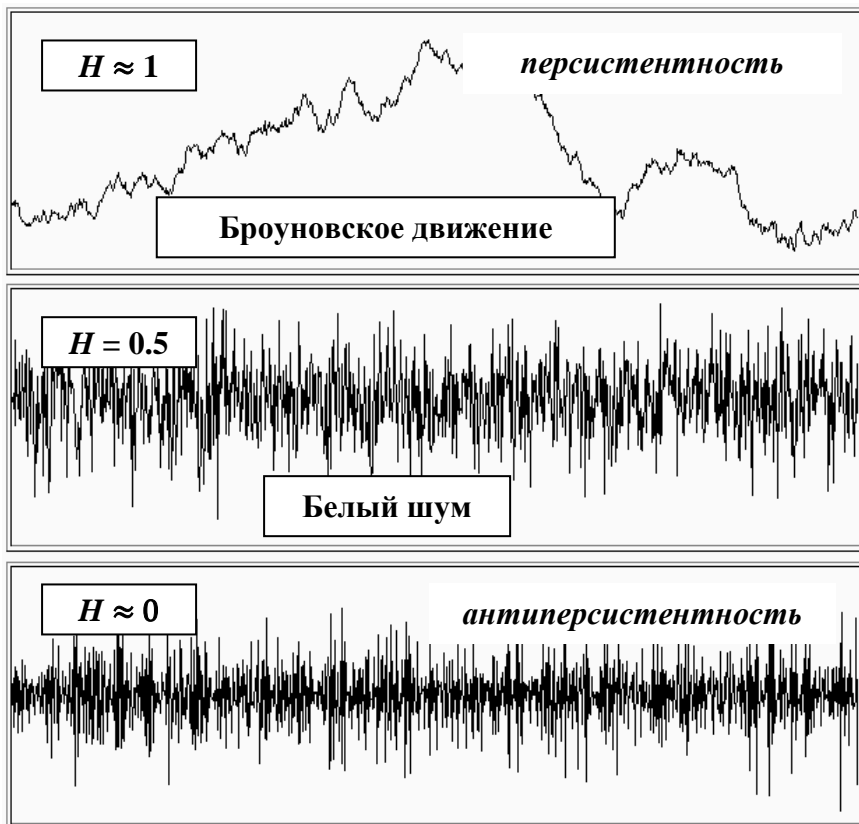


Рис. 1. Процессы Хёрста.

Спектральный анализ временных рядов

Альтернативный подход к анализу временных рядов с памятью использует аппарат спектрального (Фурье) анализа. Основная идея спектрального анализа состоит в том, что любую периодическую функцию $X(t)$ можно представить в виде суммы синусоид, т.е. собственно ряда Фурье:

$$X(t) = \sum_f A(f) \sin(2\pi f t + \varphi)$$

Обычно t интерпретируется как время. Тогда f – частота данной синусоиды, φ – её начальная фаза, то есть угол сдвига относительно начала координат, $A(f)$ – амплитуда синусоиды, то есть коэффициент ряда Фурье при члене, соответствующем частоте f . Он определяется выражением вида

$$A(f) = \int_{-\infty}^{+\infty} X(t) \sin(2\pi f t + \varphi) dt$$

В общем случае, когда функция $X(t)$ не периодическая, а любая, ряд Фурье обобщается интегралом Фурье, $A(f)$ становится некоторой функцией от непрерывной частоты f . В этом случае формулу для $A(f)$ можно рассматривать (упрощённо) как некоторое интегральное преобразование (преобразование Фурье) из плоскости t — X (время vs значение функции), на которой изображается исходная функция, в плоскость f — A (частота vs амплитуда синусоиды), которая называется *спектром*. В этом случае можно определить и обратное преобразование Фурье, согласно которому можно однозначно восстановить исходный сигнал по его спектру:

$$X(t) = \int_{-\infty}^{+\infty} A(f) \sin(2\pi ft + \varphi) df$$

На практике часто используют понятие спектральной плотности мощности (Power Spectral Density, PSD) функции $X(t)$, которая показывает, какую часть энергии сигнала несёт синусоида частоты f :

$$S(f) \sim |A(f)|^2$$

Зависимость $S(f)$, вычисленную с помощью дискретного преобразования Фурье, называют *периодограммой*.

Анализ Фурье (он же спектральный или гармонический анализ) широко используется в физических и технических приложениях с целью выявления периодичностей в сигналах, описания поведения периодических функций и систем уравнений, порождающих их. В случае непериодических функций, например, случайных процессов (временных рядов), преобразование Фурье может использоваться для описания того, какие изменения преобладают в исследуемом процессе – быстрые (высокочастотные) или медленные (низкочастотные). Удобство применения спектрального анализа связано с тем, что вследствие своей широкой распространенности в физике и технике, он реализован во всех стандартных общематематических пакетах (Matlab, MathCad и т.д.). Многие программные продукты, особенно ориентированные на обработку сигналов (такие, как LabVIEW), содержат готовые программные компоненты для дискретного быстрого преобразования Фурье (FFT), получения периодограммы и других инструментов. Наряду с ними всегда программно реализуется и обратное быстрое дискретное преобразование Фурье (IFFT, Inverse Fast Fourier Transform), которое позволяет восстановить исходный временной ряд по его спектру.

Примеры, иллюстрирующие преобразование Фурье, приведены на рис. 2.

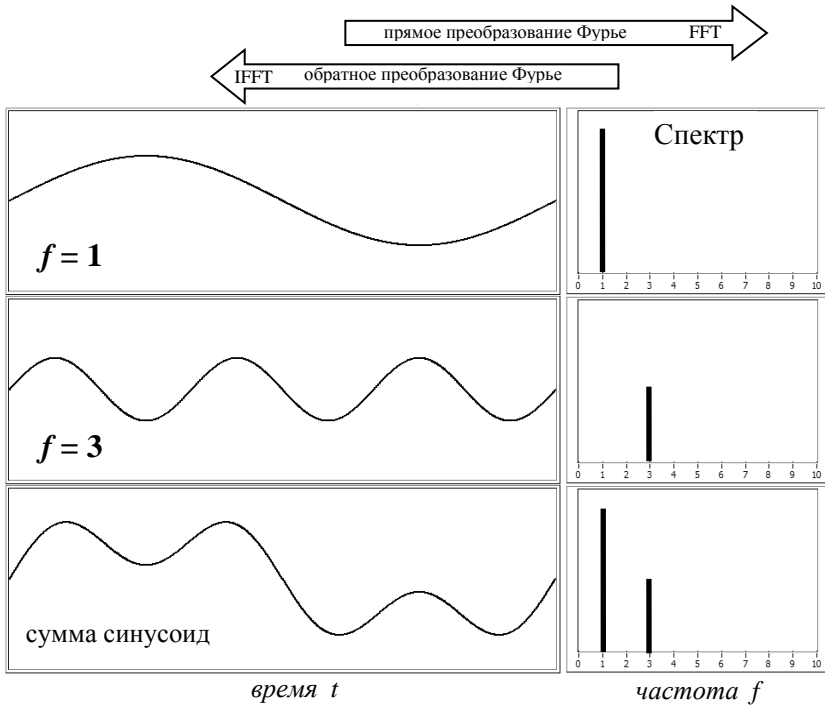


Рис. 2. Преобразование Фурье. На верхнем рисунке показана синусоида частотой $f = 1$ и её преобразование Фурье (спектр). Частота – число колебаний (циклов) в единицу времени. Спектр этой синусоиды (в упрощенном рассмотрении, не учитывающем специфические эффекты дискретного преобразования Фурье) есть дельта-функция в точке $f = 1$, амплитуда которой равна амплитуде исходной синусоиды с некоторым нормировочным коэффициентом.

Как было указано выше, прямое дискретное преобразование Фурье можно применить к любому временному ряду. Например, для временного ряда, показанного на рис. 3, спектр можно приблизительно аппроксимировать гиперболически убывающей функцией от частоты. Если этот спектр представить в дважды логарифмическом масштабе (логарифм амплитуды или спектральной плотности мощности от

логарифма частоты), то его удобно аппроксимировать линейной функцией, коэффициенты которой можно находить как результат соответствующей регрессии.

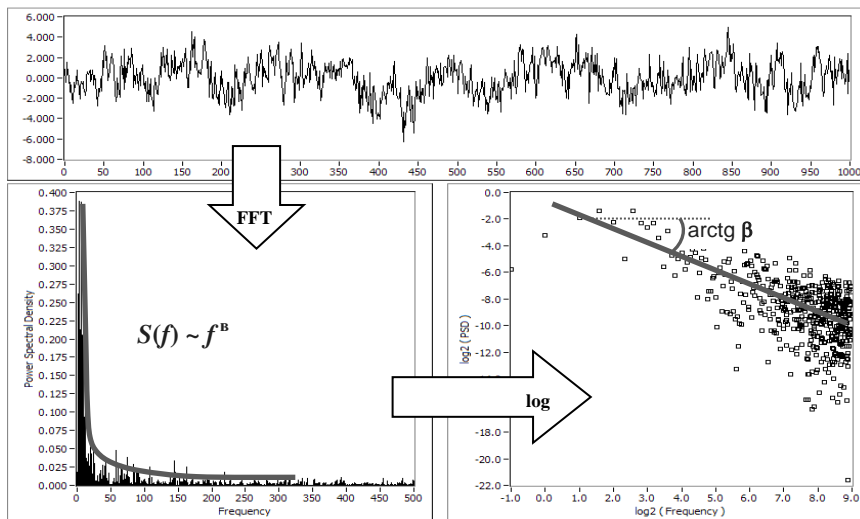


Рис. 3. Спектр мощности (периодограмма) временного ряда в обычном (слева) и дважды логарифмическом (справа) масштабах.

Преобразование Фурье (прямое) является хорошо разработанным и широко применяемым методом анализа временных рядов (см. например, Дженкинс, 1971). Однако лишь сравнительно недавно выяснилось, что спектральное представление временного ряда является весьма удобным инструментом при изучении вопросов, связанных с памятью, персистентностью и т.п. (Malamud, 1999). Обратное преобразование Фурье удобно использовать как метод генерации временных рядов с заданными свойствами памяти.

Будем рассматривать процессы, спектральная плотность мощности которых является степенной функцией частоты:

$$S(f) \sim f^\beta$$

На практике это означает, что периодограмма такого временного ряда будет хорошо аппроксимироваться степенной функцией с отрицательным или положительным показателем степени β . В дважды логарифмическом масштабе эта периодограмма выглядит как некоторая наклонная полоса или «комета», то есть хорошо аппроксимируется прямой, угловой коэффициент которой равен β (рис. 3). В этом случае β удобно интерпретировать как управляющий параметр, задающий соотношение между высокочастотной и низкочастотной частями спектра, а наклон соответствующей прямой – как наглядное представление этого соотношения. В случае $\beta=0$ вклад высокочастотной и низкочастотной составляющих одинаков, спектр в дважды логарифмическом масштабе аппроксимируется горизонтальной линией, а соответствующий временной ряд представляет собой хорошо известный *белый шум* (рис. 4).

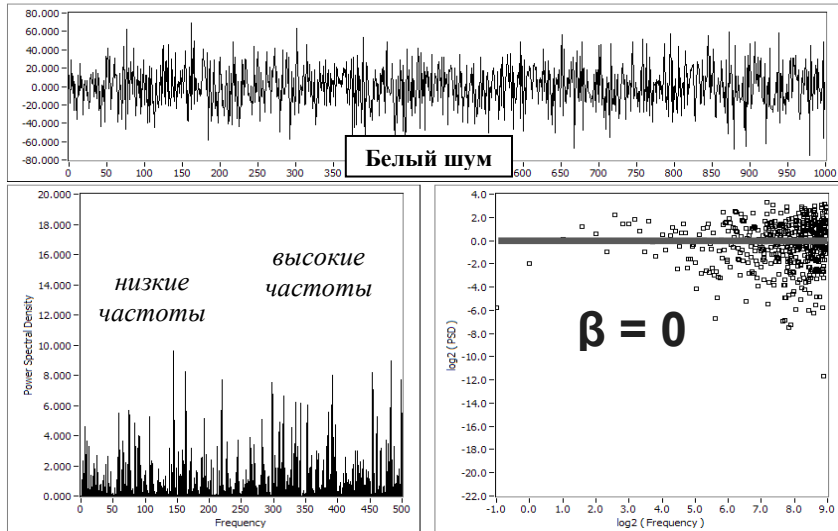


Рис. 4. Спектральный анализ белого шума.

Если $\beta < 0$ (рис. 3), то низкочастотная часть спектра преобладает над высокочастотной, медленные изменения во временном ряде превалируют над быстрыми. Это соответствует свойству персистентности. Если $\beta > 0$, преобладает высокочастотная составляющая спектра, соответствующие временные ряды выглядят как антиперсистентные.

Примеры временных рядов с разными β , сгенерированных с помощью алгоритма обратного преобразования Фурье, и их периодограммы приведены на рис. 5. Видно, что варьируя β , мы изменяем «наклон» периодограммы временного ряда в дважды логарифмическом масштабе. Тем самым мы меняем соотношение высокочастотной и низкочастотной частей спектра, соответственно этому изменяется вид сгенерированного по данному спектру временного ряда. Таким образом, β может рассматриваться как параметр (один!), позволяющий описывать довольно разные по внешнему виду ряды и «схватывающий» свойство персистентности, а значит и наличие долгой памяти. За некоторую «точку отсчёта» естественно принять значение $\beta=0$, соответствующее белому шуму. Очевидно, β может принимать и нецелые значения, и может быть расширено за границы диапазона, приведённого на рис. 5, в сторону как положительных, так и отрицательных значений. При изменении β в сторону $-\infty$ ряд становится визуально всё более гладким (не переставая быть случайным процессом); при изменении β в сторону $+\infty$, напротив, ряд становится более «изрезанным», «изломанным».

Важно отметить, что при $\beta=-2$ временной ряд представляет собой классическое броуновское движение, то есть процесс, приращения которого образуют белый шум. Это наблюдение есть проявление очень важного правила: если мы образуем временной ряд приращений некоторого процесса, то есть дифференцируем (численно) данный процесс, β получившегося временного ряда приращений равна β

исходного процесса плюс 2. И наоборот, интегрирование временного ряда соответствует уменьшению его β на 2.

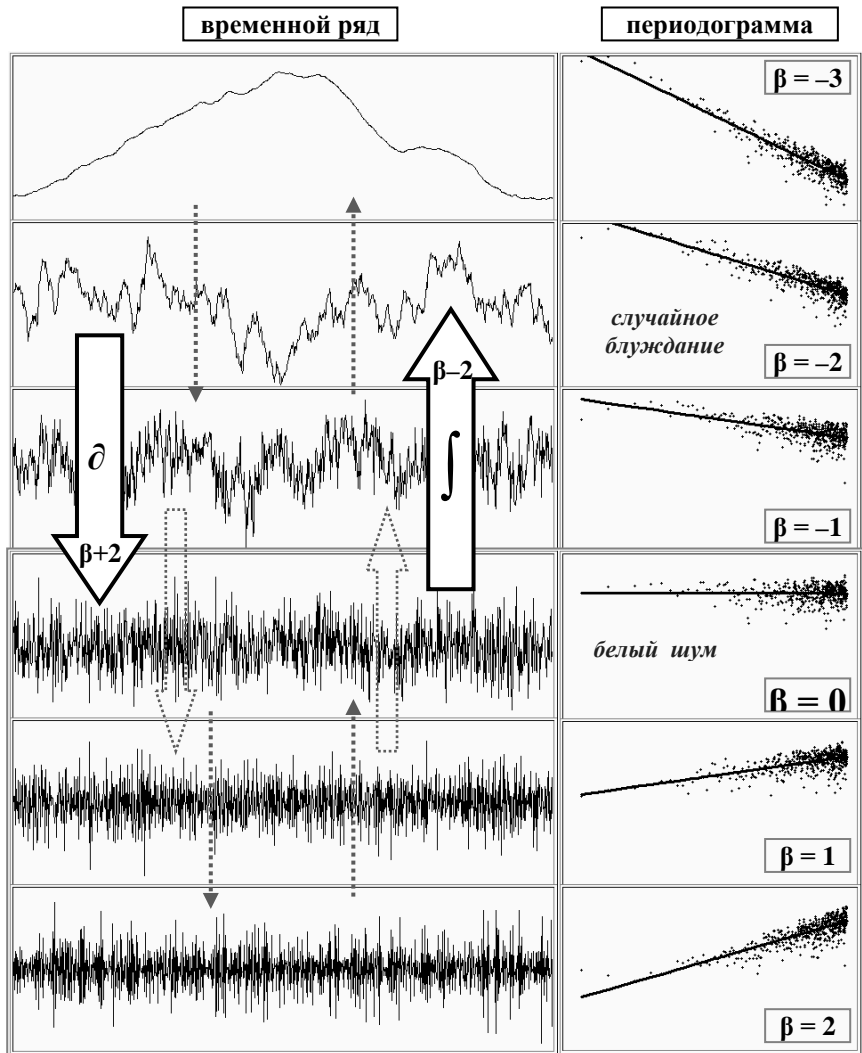


Рис. 5. Временные ряды, сгенерированные методом обратного преобразования Фурье, и их периодограммы в дважды логарифмическом масштабе для различных значений спектрального параметра β . Для разных временных рядов масштаб по вертикальной оси сильно отличается.

3. ДЛИННАЯ ПАМЯТЬ В ЭКОНОМЕТРИЧЕСКИХ МОДЕЛЯХ

Традиционным и хорошо разработанным инструментом эконометрики финансовых временных рядов является модель авторегрессии и скользящего среднего ARMA (p,q). В рамках этой модели автокорреляционная функция убывает экспоненциально, то есть сравнительно быстро. Про такие временные ряды говорят, что они обладают короткой памятью. Расширением данной модели являются процессы ARIMA (p,d,q) – интегрированные процессы авторегрессии и скользящего среднего, описываемые следующим уравнением

$$\Phi(L)(1-L)^d X_t = \mu + \Theta(L)\varepsilon_t,$$

где X_t – исследуемый процесс (временной ряд);

L - оператор сдвига;

$\Phi(L)$ - полином степени p от L ;

$\Theta(L)$ - полином степени q от L ;

d – порядок интегрирования процесса X_t .

При $d = 0$ X_t является процессом, удовлетворяющим модели ARMA (p,q) (ARIMA ($p,0,q$)). Этот процесс стационарен и обладает короткой памятью, то есть действие шока сохраняется в течение считанного количества лагов. При $d = 1$ можно говорить о бесконечной памяти ряда, имея в виду, что каждый шок оказывает влияние на поведение процесса бесконечно долго. Также можно рассматривать $d = 2,3$ и т.д. Таким образом, промежуточная ситуация, когда последствия шока оказываются временными, но длительными, остается вне поля зрения. Для решения этой проблемы была предложена модель дробно-интегрированного авторегрессионного процесса скользящего среднего ARFIMA (p,d,q), расширяющая предыдущую на случай нецелых значений d . Уравнение для этой модели остается неизменным, а оператор дробного интегрирования аппроксимируется следующим рядом:

$$(1 - L)^d = \sum_{k=0}^{\infty} \frac{\Gamma(k-d)}{\Gamma(-d)\Gamma(k+1)} L^k,$$

где Γ – гамма-функция.

При $0 < d < 1$ автокорреляционная функция такого процесса убывает гиперболически, то есть крайне медленно. Параметр d может служить мерой длинной памяти – чем больше d , тем память «длиннее».

Общий подход, развиваемый эконометрикой временных рядов, состоит в построении моделей, связывающих последующие значения временного ряда с предыдущими, таких, как ARMA, ARIMA, ARFIMA, GARCH, FIGARCH и многих других. Очевидно, с точки зрения практики, такие модели «заточены» на решение задачи прогнозирования. При этом, следуя методологии эконометрики, должны быть сформулированы основная и альтернативная гипотезы, построена статистика метода и доказано её предельное распределение. Изложенные выше спектральный и «фрактальный» подходы (метод Хёрста) пришли в эконометрику из физических приложений, и их перевод на язык эконометрики временных рядов продолжает оставаться актуальной задачей. В частности, вопрос о предельном распределении и робастности показателя Хёрста является дискуссионным (см. работы Andrew Lo, например Lo, 1991). Спектральный параметр β представляется более удобным для описания эффекта длинной памяти, однако для него (пока) не получено предельное распределение.

В рамках спектрального подхода разработан метод GPH (по именам создателей Geweke и Porter-Hudak), позволяющий получать дробный порядок интегрирования временного ряда d , чтобы затем подставить его в модель ARFIMA (Geweke, 1983).

Если спектральная плотность ряда представима в виде

$$S(f) = S_0(f) [2\sin(f/2)]^{-2d},$$

где $S_0(f)$ – медленно меняющаяся функция (соответствует компонентам короткой памяти в ряде), то порядок интегрирования d может быть получен как коэффициент регрессии $\log I(f_j)$ на $\log[2\sin(f_j/2)]^{-2}$, где $I(f_j)$ – периодограмма.

Тогда, при соблюдении ряда дополнительных условий, предельное распределение будет нормальным гауссовским:

$$\sqrt{m}(\hat{d} - d) \rightarrow N(0, \frac{\pi^2}{24}),$$

при $m \rightarrow \infty$, где m – число частот f_j , используемых в регрессии (Palma, 2007).

Существуют теоретические соотношения, связывающие характеристики длинной памяти, полученные разными методами. Так, спектральный параметр β связан с порядком интегрирования d из модели ARFIMA соотношением

$$\beta = -2d$$

При этом белый шум ($\beta=0$) соответствует процессу ARFIMA($p,0,q$), а случайное блуждание ($\beta=-2$) – процессу ARFIMA($p,1,q$).

4. ИССЛЕДОВАНИЕ ПРОГНОЗНЫХ СВОЙСТВ МОДЕЛЕЙ

Сопоставление прогнозов модели с длинной памятью и авторегрессионной модели скользящего среднего для сгенерированного ряда

Для анализа свойств моделей ARFIMA и ARIMA был сгенерирован ряд X длиной 3000 наблюдений с порядком интегрирования 0,45. Для этой цели был использован пакет R, который поддерживает широкий спектр статистических и численных методов для специальных областей применения.

Ряд X был сгенерирован следующим образом: сначала был получен ряд Y с заданными коэффициентами AR и MA, а затем к нему был применен оператор взятия разности порядка 0,45, и полученные значения были смещены на 100 единиц вверх. Динамика реализовавшихся значений показана на рис. 6.

$$y_t = 0.2y_{t-1} + \varepsilon_t + 0.4\varepsilon_{t-1}, \quad \varepsilon_t \sim N(0,1),$$

$$x_t = 100 + (1 - L)^{0.45} y_t.$$

Для этой цели в R был включен пакет `fracdiff` и выполнена следующая программа:

```
k <- fracdiff.sim(3000, ar = .2, ma = .4, d = .45)
t <- k$series
x <- t+100
```

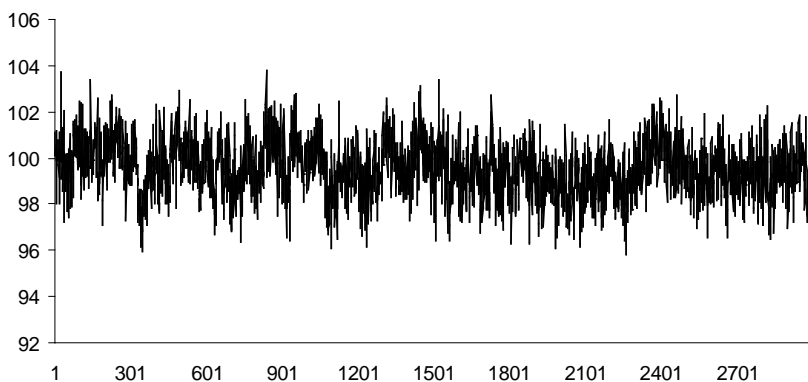


Рис. 6. Сгенерированный ряд X .

Для проверки переменной X на стационарность был использован расширенный тест Дикки-Фуллера (ADF test). В нем нулевая гипотеза означает наличие единичного корня. Отрицание нулевой гипотезы

означает, что все корни по абсолютной величине меньше 1, что является условием стационарности. В данном случае тест не отрицает наличия единичных корней, статистика лежит левее критических точек, что говорит о стационарности сгенерированного ряда (Таблица 1).

Таблица 1. Расширенный тест Дикки-Фуллера для сгенерированного ряда X

Нулевая гипотеза: X содержит единичный корень			
Экзогенная переменная: Константа			
Количество включенных лагов: 8			
Статистика ADF		t-статистика	Вероятность*
		-9,77604	0,0000
Критические точки:	1% уровень	-3,43235	
	5% уровень	-2,86231	
	10% уровень	-2,56722	
*Односторонние р-значения MacKinnon			

Для ряда из первых 2000 наблюдений переменной X была оценена модель ARMA(1,1) (Таблица 2). Все коэффициенты согласно показателям p-value оказались значимыми. Для устойчивости процесса необходимо, чтобы обратные корни полиномов AR и MA лежали внутри единичного круга, что выполняется в построенной модели (обратные корни равны 0,93 и 0,70, соответственно).

Таблица 2. Модель ARMA(1,1) для ряда X

Метод Наименьших Квадратов				
Выборка: 2 2000				
Переменная	Коэфф.	Стандартная ошибка	t-статистика	Вер-сть
C	99,633	0,097	1027,231	0,000
AR(1)	0,930	0,013	73,725	0,000
MA(1)	-0,705	0,024	-28,977	0,000
R^2	0,273	Инф. критерий Akaike		2,900
Скорректированный	0,272	Инф. критерий Schwarz		2,908
Ст. ошибка регрессии	1,031	Статистика Durbin-Watson		1,893
Сумма квадратов остатков	2120,303			
F-статистика	374,432			

Тест White, в котором за нулевую гипотезу принимается гомоскедастичность дисперсий ошибок, показал отсутствие гетероскедастичности в этой модели (Таблица 3).

Таблица 3. Тест White на гетероскедастичность в модели ARMA(1,1)

F-статистика	0,4684	Вер-ть. F(9,1989)	0,8965
Набл.* R^2	4,2276	Вер-ть. $\chi^2(9)$	0,8958

После проведенных тестов был построен прогноз для последних 1000 наблюдений. На рис. 7 изображены прогнозные значения переменной X, реальные значения и доверительный интервал, рассчитанный как прогнозные значение плюс удвоенная стандартная ошибка.

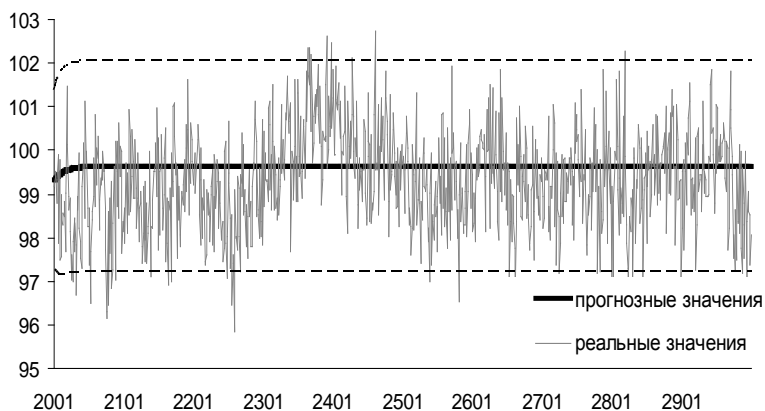


Рис. 7. Прогноз модели ARMA(1,1) для ряда X .

Для этого же ряда была оценена модель ARFIMA(1, d ,1). В различных эконометрических пакетах встроены разные способы оценивания параметра d , в частности метод GPH, который применялся при оценке этой модели в пакете OxMetrics (Таблица 4). OxMetrics – модульное программное обеспечение, обеспечивающее интегрированное решение для эконометрического анализа временных рядов, прогнозирования и финансового эконометрического моделирования.

Таблица 4. Модель ARFIMA(1,d,1) для ряда X

Выборка: 1 – 2000

	Коэффициент	Стандартная ошибка	t-значение	t-вероятность
d	0,455	0,065	7,000	0,000
AR-1	0,315	0,122	2,600	0,010
MA-1	-0,488	0,154	-3,160	0,002
Константа	99,620	1,066	93,500	0,000

Параметр d оказался значим и принимает значение $0,455 \pm 0,13$ (значение коэффициента \pm две стандартные ошибки), что очень близко к заданной величине (0,45). Другие коэффициенты модели значимы на 1% уровне. Значения вероятности основаны на $t(T-s)$ - распределении, где s - число оцененных параметров.

Затем был построен прогноз на тот же интервал, что и в модели ARMA, результаты продемонстрированы на рис. 8.

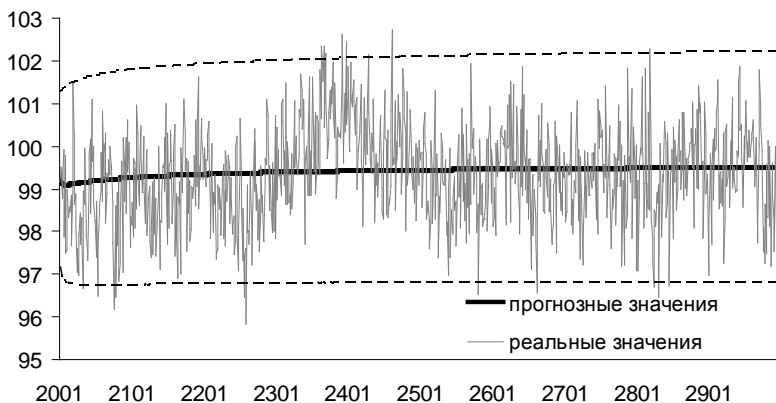


Рис. 8. Прогнозные значения ряда X по модели ARFIMA(1,d,1).

Для сравнения прогнозных значений моделей широко применяются два следующих показателя (Shittu, 2009) – корень среднеквадратической ошибки прогноза (RMSFE) и средняя абсолютная процентная ошибка прогноза (MAPFE), которые вычисляются, соответственно, по формулам:

$$RMSFE = \sqrt{\frac{1}{h+1} \sum_{t=s}^{h+s} (X_t^F - X_t)^2},$$

$$MAPFE = \frac{100}{h+1} \sum_{t=s}^{h+s} \left| \frac{X_t^F - X_t}{X_t} \right|,$$

где X_t^F - прогнозные значения ряда X ,

X_t - реальные значения ряда X ,

$h+1$ – период прогнозирования.

Очевидно, что чем меньше значения этих показателей, тем меньше отклонения ожидаемых значений от реализовавшихся и, следовательно, точнее построенный прогноз.

Для сгенерированного ряда X корень среднеквадратической ошибки прогноза составляет 1,16 для модели ARMA(1,1) и 1,11 для модели ARFIMA(1, d ,1), а средняя абсолютная процентная ошибка прогноза равна 0,93% и 0,89%, соответственно.

Наименьшие значения этих показателей получены для прогнозов модели ARFIMA, причем, в среднем прогнозные значения этой модели на 4 базисных пункта ближе к реализовавшимся, чем прогнозные значения, полученные с помощью модели ARMA. Отсюда можно сделать вывод о более реалистичных оценках ожидаемых значений данного ряда в модели, учитывающей длинную память.

На рис. 9 приведены стандартные ошибки моделей ARMA(1,1) и ARFIMA(1, d ,1). Видно, что в начале периода прогнозирования ошибки

моделей незначительно отличаются, а затем ошибки модели ARFIMA начинают резко увеличиваться.

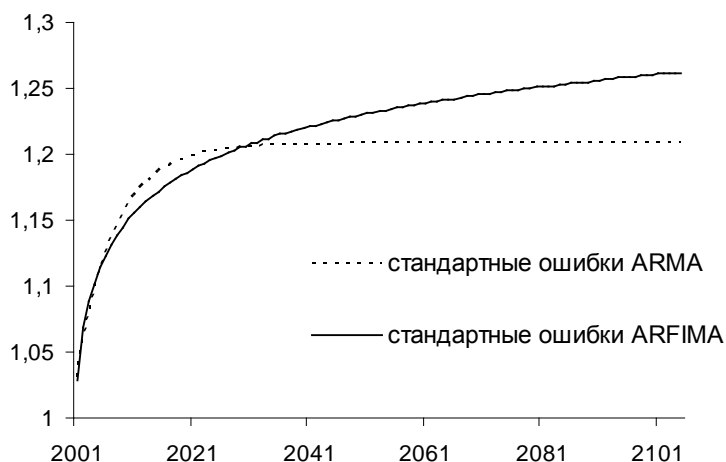


Рис. 9. Стандартные ошибки моделей ARMA(1,1) и ARFIMA(1,d,1) для ряда X.

Различия в полученных графиках обуславливаются разницей в порядке интегрирования, а также тем, что ошибка прогноза в случайном блуждании имеет свойство накапливаться. Это заставляет ошибки модели ARFIMA возрастать намного быстрее, чем ошибки авторегрессионной модели скользящего среднего.

Исследование прогнозных свойств моделей для ряда цен на нефть

Полученные результаты были проверены на реальных экономических данных, а именно - на ряде цен на нефть марки Brent (рис. 10). Ежедневные цены закрытия в долларах США за баррель с 6 мая 2005

года по 23 апреля 2010 года ($T=1814$) были взяты с сайта компании Финам.

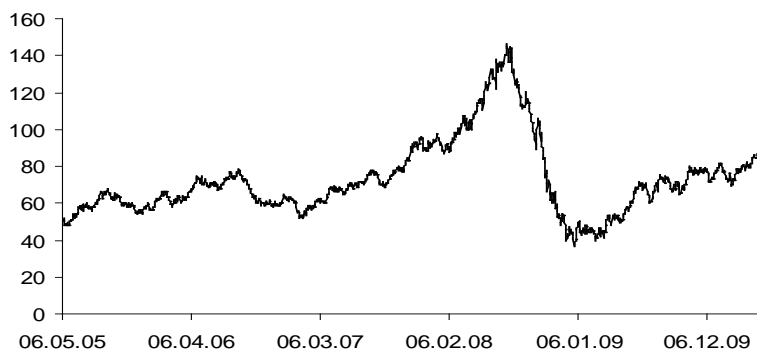


Рис. 10. Цены на нефть марки Brent, долл США/баррель.

При проверке гипотезы о стационарности ряда расширенный тест Дикки-Фуллера выявил, что цены на нефть обладают первым порядком интегрирования (Таблицы 5 и 6).

Таблица 5. Тест Дикки-Фуллера для ряда цен на нефть

Нулевая гипотеза: Ряд имеет единичный корень			
Экзогенная переменная: Константа			
Количество лагов: 1			
		t-статистика	Вероятность*
Статистика ADF		-0,708	0,839
Критические точки:	1% уровень	-3,497	
	5% уровень	-2,891	
	10% уровень	-2,582	
* Односторонние р-значения MacKinnon			

Таблица 6. Тест Дикки-Фуллера для первой разности ряда цен на нефть

Нулевая гипотеза: первая разность ряда содержит единичный корень			
Экзогенная переменная: Константа			
Количество лагов: 0			
		t-статистика	Вероятность*
Статистика ADF		-7,799	0,000
Критические точки:	1% уровень	-3,497	
	5% уровень	-2,891	
	10% уровень	-2,582	
* Односторонние р-значения MacKinnon			

В пакет R встроена функция `auto.arima`, возвращающая лучшую модель ARIMA согласно информационным критериям Akaike (AIC) и/или Schwarz (BIC):

$$AIC = 2 \frac{p+q}{n} + \ln \left(\frac{\sum_{t=1}^n e_t^2}{n} \right),$$

$$BIC = \frac{(p+q) \ln n}{n} + \ln \left(\frac{\sum_{t=1}^n e_t^2}{n} \right),$$

где p и q - параметры модели ARMA,

n - объем выборки.

Критерий Akaike является попыткой свести в один показатель два требования: уменьшение числа параметров модели и качество подгонки

модели. Главное отличие критерия Schwarz от AIC состоит в большем штрафе за количество параметров. Согласно этим критериям, из двух моделей следует выбирать модель с меньшим значением AIC и/или BIC (Магнус, 2004).

Согласно критериям AIC и Schwarz лучшей моделью для первых 1700 наблюдений оказалась модель ARIMA(1,1,1) (Таблица 7). Обратные корни полиномов AR и MA лежат внутри единичного круга, а коэффициенты модели значимы на 1% уровне.

Таблица 7. Модель ARIMA(1,1,1) для цен на нефть

Зависимая переменная: Первая разность ряда				
Метод Наименьших Квадратов				
Выборка: 3 1714				
Переменная	Коэфф.	Ст. ошибка	t-стат.	Вер-сть
AR(1)	0,986719	0,01215	81,2169	0,000
MA(1)	-0,974671	0,01671	-58,32	0,000
R^2	0,005498	Akaike info criterion		3,41111
Скорректированный R^2	0,004916	Schwarz criterion		3,41747
Ст. ошибка регрессии	1,331138			
Сумма квадратов остатков	3029,997			
Статистика Durbin-Watson	2,018945			

По данной модели был построен прогноз на последние 100 наблюдений (рис. 11). Показатели RMSFE и MAPFE для полученных прогнозных значений составляют 4,43 и 4,58%, соответственно.

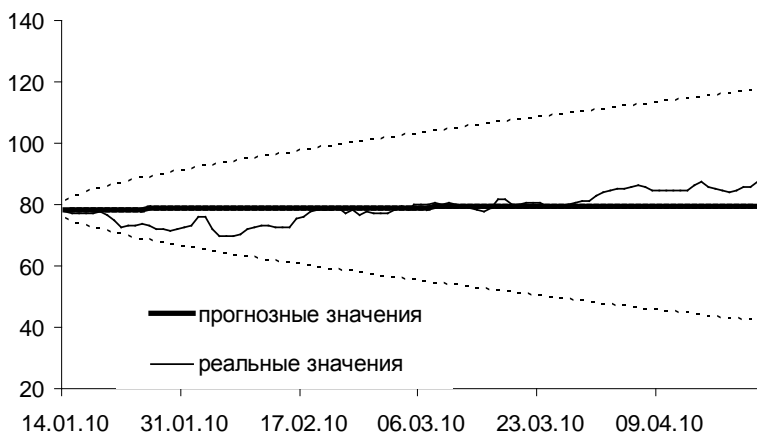


Рис. 11. Прогноз модели ARIMA(1,1,1) для цен на нефть.

Тест White на гетероскедастичность выявил наличие ARCH-эффекта в модели ARIMA (Таблица 8). Это означает, что реальные значения оценок имеют условно гетероскедастичные остатки, т.е. остатки с переменными дисперсиями. При этом оценки параметров модели, получаемые обычным методом наименьших квадратов, который использовался при оценке параметров модели ARIMA, не являются наилучшими и не обеспечивают минимума дисперсии и ковариации. В этом случае предлагается использовать обобщенную модель авторегрессионной условной гетероскедастичности.

Таблица 8. Тест White на гетероскедастичность в модели ARIMA(1,1,1)

F-статистика	40,82122	Вер-сть F(3,1708)	0,000
Набл.* R^2	114,5381	Вер-сть. $\chi^2(3)$	0,000

Для этого ряда была построена модель GARCH(1,1) (Таблица 9), в которой все коэффициенты оказались значимы на 1% уровне. По этой

модели был построен прогноз на последние 100 наблюдений (рис. 12). RMSFE и MAPFE для прогнозных значений этой модели составляют 4,55 и 4,71%, соответственно.

Таблица 9. Модель GARCH(1,1) для первой разности ряда цен на нефть

Зависимая переменная: Первая разность ряда				
Метод: ML - ARCH (Marquardt)				
Выборка: 3 1714				
Переменная	Коэфф.	Ст. ошибка	z-статистика	Вер-сть
AR(1)	0,983585	0,02003	49,1008	0,000
MA(1)	-0,977517	0,02356	-41,484	0,000
Уравнение дисперсии				
C	0,004958	0,00199	2,49146	0,013
RESID(-1)^2	0,030083	0,00396	7,60477	0,000
GARCH(-1)	0,966753	0,00472	204,996	0,000
R^2	0,004341	Инф. критерий Akaike		3,1088
Скорректированный R^2	0,002008	Критерий Schwarz		3,1247 1
Ст. ошибка регрессии	1,333082			
Сумма квадратов остатков	3033,523			
Статистика Durbin-Watson	2,004579			

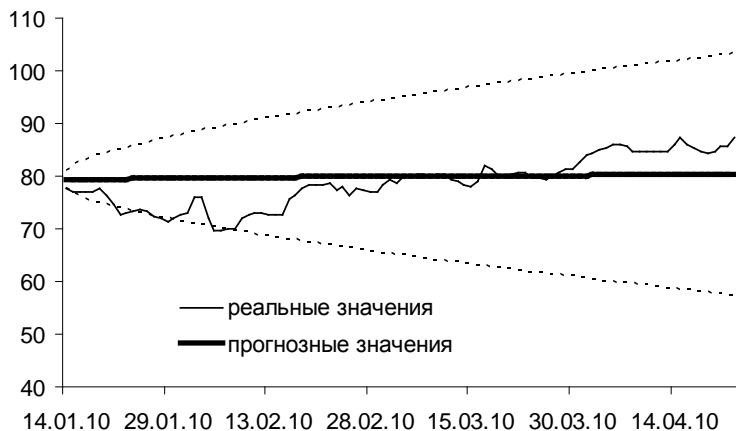


Рис. 12. Прогнозные значения ряда цен на нефть по модели GARCH(1,1).

Для оценки дробного порядка интегрирования была построена модель ARFIMA(1,d,1) (Таблица 10), в которой параметр d значим на 5% уровне и составляет $0,16 \pm 0,139$ (значение коэффициента \pm удвоенная стандартная ошибка). При оценке этого параметра по-прежнему использовался метод GPH, а вероятность рассчитывалась на основе t(T-s)-распределения. Остальные параметры значимы на 1% уровне.

Таблица 10. Модель ARFIMA(1,d,1) для первой разности ряда цен на нефть

Выборка: 2-1714				
	Коэффициент	Стандартная ошибка	t-значение	t-вероятность
d	0,162	0,070	2,33	0,020
AR-1	0,550	0,091	6,01	0,000
MA-1	-0,701	0,094	-7,45	0,000
log-lik	-2919,693	AIC	3,414	

На последние 100 наблюдений, не включенных в модель, был построен прогноз (рис. 13). Показатели ошибки прогноза составляют 4,39 и 4,55% и по сравнению со значениями моделей условной гетероскедастичности и авторегрессионной модели скользящего среднего оказываются наименьшими (в частности, прогнозные значения модели ARFIMA в среднем ближе к реальным на 3 процентных пункта по сравнению со значениями модели ARIMA и на 16 процентных пунктов по сравнению со значениями модели GARCH).

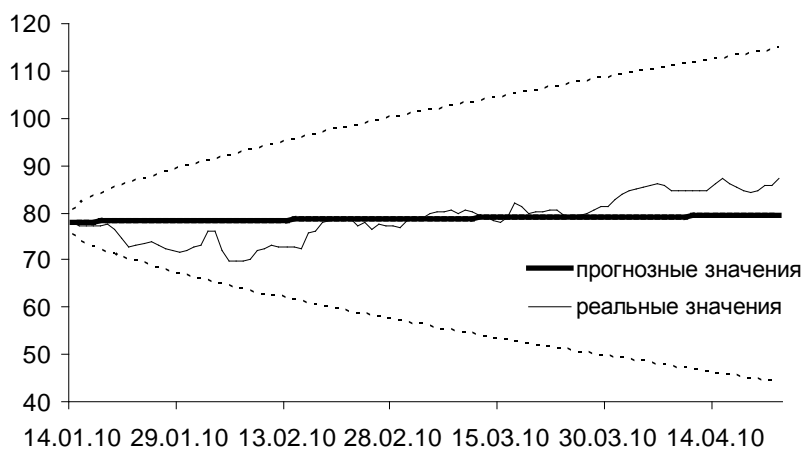


Рис. 13. Прогнозные значения ряда цен на нефть по модели ARFIMA(1,d,1).

На рис. 14 показаны стандартные ошибки всех оцененных моделей. Ошибки модели GARCH оказались наименьшими, а ошибки модели ARFIMA меньше ошибок ARIMA вследствие разницы порядков интегрирования.

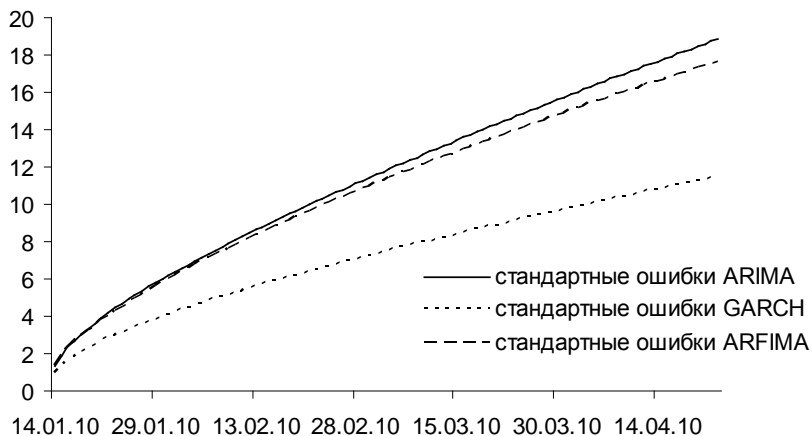


Рис. 14. Стандартные ошибки моделей $ARIMA(1,1,1)$, $ARFIMA(1,d,1)$ и $GARCH(1,1)$ для ряда цен на нефть.

В таблице 11 приведены показатели RMSFE и MAPFE для моделей, построенных для сгенерированного ряда X и ряда цен на нефть.

Таблица 11. Показатели RMSFE и MAPFE

	ряд X		ряд цен на нефть	
	RMSFE	MAPFE	RMSFE	MAPFE
ARIMA	1,162	0,931	4,429	4,582
GARCH	-	-	4,551	4,712
ARFIMA	1,113	0,893	4,391	4,551

Наименьшие значения обоих критериев достигаются у прогнозов модели ARFIMA. Это означает, что модель, учитывающая наличие длинной памяти, оказалась более точной в плане построения прогнозов для рядов, обладающих этим свойством.

ЗАКЛЮЧЕНИЕ

В данной работе предлагается спектральный подход к оценке эффекта длинной памяти во временных рядах и его практическое применение для анализа цен на нефть.

В теоретической части подчеркнута связь явления длинной памяти с концепцией эффективного рынка, сделан вывод о том, что задача выявления длинной памяти важна как с теоретической точки зрения – для подтверждения или опровержения теории эффективного рынка, так и с практической – для прогнозирования финансовых временных рядов. Также рассмотрена связь понятий памяти временного ряда с концепциями персистентности и фрактальности.

В плане математического обоснования даны основы теории преобразования Фурье и рассмотрен спектральный подход к описанию эффекта длинной памяти. Показано, что спектральный параметр β является удобной характеристикой персистентности, а следовательно, и длинной памяти. Подчеркнута его связь с операциями дифференцирования и интегрирования временного ряда.

Для сгенерированного ряда с порядком интегрирования 0,45 были оценены модели ARFIMA и ARMA, прогнозные свойства которых согласно значениям корня среднеквадратической ошибки и средней абсолютной процентной ошибки прогноза отличаются, причем отклонения прогнозных значений модели с дробным порядком интегрирования от реальных значений оказываются меньше.

Для определения практической значимости полученного результата был исследован ряд цен на нефть, в котором была обнаружена длинная память. Для этого ряда были оценены модели ARFIMA, ARIMA и GARCH, анализ которых показал, что наиболее точные прогнозы для ряда с

дробным порядком интегрирования достигаются с помощью модели, учитывающей наличие длинной памяти.

СПИСОК ЛИТЕРАТУРЫ

- Магнус Я.Р., Катышев П.К., Персецкий А.А.** *Эконометрика. Начальный курс: Учеб.* – 6-е изд., М.: Дело, 2004. – 307 с.
- Дженкинс Г., Ваттс Д.** *Спектральный анализ и его приложения* М.: Мир, 1971.
- Мандельброт Б.** *Фрактальная геометрия природы.* М.: Институт компьютерных исследований, 2002.
- Перцовский О.Е.** *Моделирование валютных рынков на основе процессов с длинной памятью.* Препринт WP2/2004/03 – М.: ГУ ВШЭ, 2003. – 52 с.
- Петерс Э.Э.** *Фрактальный анализ финансовых рынков.* М.: Интернет-трейдинг, 2004.
- Geweke J., Porter-Hudak S.** *The estimation and application of long memory time series models* // Journal of Time Series Analysis, 4, 1983, 221-38.
- Malamud B.D., Turcotte D.L.** *Self-affine time series: measures of weak and strong persistence* // Journal of Statistical Planning and Inference, 80, pp. 173-196, 1999.
- Lo A.W.** *Long term memory in stock market prices*// Econometrica 59(5), 1991, 1279-1313.
- Palma W.** *Long-memory time series. Theory and methods.* Wiley, 2007.
- Shittu O. I., Yaya O. S.** *Measuring forecast performance of ARMA and ARFIMA models: an application to US dollar/ UK pound foreign exchange rate*// European journal of scientific research, Vol.32, No.2, 2009, 170.

Tabak В.М., Сажеиро D.O. *Are the crude oil markets becoming weakly efficient over time? A test for time-varying long-range dependence in prices and volatility* // Energy Economics 29 (2007) 28 – 36.

Сайт холдинга «Финам» <http://www.finam.ru>,
временной ряд цен на нефть марки Brent импортирован с адреса
<http://www.finam.ru/analysis/export/default.asp>.

ОГЛАВЛЕНИЕ

ВВЕДЕНИЕ	2
1. ДЛИННАЯ ПАМЯТЬ И ЭФФЕКТИВНОСТЬ РЫНКА	4
2. ДЛИННАЯ ПАМЯТЬ ВО ВРЕМЕННЫХ РЯДАХ	7
Процессы Хёрста и персистентность	7
Спектральный анализ временных рядов.....	10
3. ДЛИННАЯ ПАМЯТЬ В ЭКОНОМЕТРИЧЕСКИХ МОДЕЛЯХ	17
4. ИССЛЕДОВАНИЕ ПРОГНОЗНЫХ СВОЙСТВ МОДЕЛЕЙ	19
Сопоставление прогнозов модели с длинной памятью и авторегрессионной модели скользящего среднего для сгенерированного ряда	19
Исследование прогнозных свойств моделей для ряда цен на нефть.....	26
ЗАКЛЮЧЕНИЕ.....	35
СПИСОК ЛИТЕРАТУРЫ.....	36

Балагула Ю.М., Абакумова Ю.А. Длинная память на рынке нефти: спектральный подход. — СПб. : Европейский университет в Санкт-Петербурге, 2011. — 40 с. — (Серия препринтов; факультет экономики, Ес-01/11).

Отпечатано с оригинал-макета, предоставленного авторами

Подписано в печать 13.01.11

Формат 60x88 1/16. Тираж 50 экз

Европейский университет в Санкт-Петербурге
191187, Санкт-Петербург, ул. Гагаринская, д. 3
econ@eu.spb.ru