А. Г. Предтеченский
Новосибирский государственный университет
Ул. Пирогова, 2, Новосибирск, 630090
e-mail:

ОЦЕНИВАНИЕ ОДНОМЕРНОЙ ПАРАМЕТРИЧЕСКОЙ ARCH МОДЕЛИ ПРИ НАРУШЕНИИ ГИПОТЕЗЫ О НОРМАЛЬНОСТИ. ОБОБЩЕННЫЙ МЕТОД МОМЕНТОВ

Спецификация модели.

Концепция модели авторегрессионной условной гетероскедастичности (AutoRegressive Conditional Heteroskedasticity, ARCH) была впервые предложена в работе Engle R.F. (1982). Основная идея привела к многочисленным модификациям, среди которых наиболее известны: обобщенная ARCH (Generalized ARCH, GARCH; Bollerslev T. 1986), ARCH–M (Engle R.F., Lilien D.M. and Robins R.P., 1987), экспоненциальная GARCH (Exponential GARCH, EGARCH; Nelson D.B., 1991) модели. Обзоры литературы по ARCH моделям представлены Bollerslev T., Chou R.Y., and Kroner K.F. (1992), Bollerslev T., Engle R.F., and Nelson D. B. (1993).

Ниже мы формулируем модель в достаточно общей форме, допускающей любые способы совместной параметризации функций условного среднего и дисперсии.

Пусть $\{y_t\}_{t=1}^{t=+\infty}$ — последовательность наблюдаемых скалярных случайных величин. $\Omega_t = (y_{t-1}, y_{t-2}, \cdots y_1)$ — набор предопределенных к моменту t переменных (набор переменных, которым обусловлены математические ожидания). Функции условного математического ожидания и условной дисперсии y_t совместно параметризованы вектором $\theta \in \Theta \subset R^m$: $\mu(t;\Omega_t;\theta)$, $\sigma^2(t;\Omega_t;\theta)$. μ, σ^2 — известные функции, которые далее будем обозначать, опуская аргумент Ω_t и используя нижний индекс t. Существует единственный $\theta_0 \in \Theta$ такой, что

$$E(y_t|\Omega_t) = \mu_t(\theta_0) \qquad (M.1)$$

$$V(y_t|\Omega_t) = \sigma_t^2(\theta_0) \qquad (M.2),$$

θ_0 называется вектором истинных параметров. Определим также функции остатков и стандартизованных остатков

$$\varepsilon_t(\theta) = y_t - \mu_t(\theta)$$
$$z_t(\theta) = \varepsilon_t(\theta)\sigma_t^{-1}(\theta).$$

Условия (M.1)–(M.2) позволяют вычислить математическое ожидание и дисперсию $\varepsilon_t(\cdot)$ в произвольной точке параметрического пространства. Например, в точке θ_0

$$E(\varepsilon_t(\theta_0)|\Omega_t) = 0 \qquad (M.3)$$

$$E(\varepsilon_t^2(\theta_0)|\Omega_t) = \sigma_t^2(\theta_0). \qquad (M.4)$$

Процедура, используемая наиболее часто для оценки θ_0, состоит в максимизации функции правдоподобия, построенной в предположении о том, что распределение y_t при условии Ω_t нормально со средним и дисперсией, определенными (M.1)-(M.2). Гипотеза об условной нормальности, однако, часто не выдерживает тестирования; в этой связи возникает проблема выбора метода, устойчивого к различным ее нарушениям.

Обоснован метод квази-максимального правдоподобия, который предполагает максимизацию нормальной функции правдоподобия при том, что распределение $y_t|\Omega_t$ в действительности не является нормальным (Bollerslev T. and Wooldridge J.M., 1992). При этом оценки сохраняют свойства состоятельности и асимптотической нормальности, однако утрачивают свойство асимптотической эффективности. В данной работе предпринята попытка оценить модель обобщенным методом моментов. Оказалось возможным построить оптимальные инструменты, приводящие к оценкам, асимптотически более эффективным, чем оценки метода квази-максимального правдоподобия.

Оценивание модели в предположении о нормальности условного распределения: метод максимального правдоподобия.

Применение метода максимального правдоподобия требует явного задания функции плотности распределения случайных величин y_t. Предположение о нормальном характере распределения $y_t|\Omega_t$ позволяет воспользоваться простой и детально разработанной процедурой оценивания неизвестных параметров (Цыплаков А., 1997; Bollerslev T., et al 1993). Гипотеза об условной нормальности процесса формально записывается как

$$y_t|\Omega_t \sim N\left(\mu_t(\theta_0), \sigma_t^2(\theta_0)\right). \tag{N}$$

Определения

Оценки метода максимального правдоподобия (ММП) доставляют максимум критериальной функции, составленной из вкладов отдельных наблюдений:

$$l(\theta) = \sum\nolimits_{t=1}^{n} l_t(\theta), \tag{1}$$

где вклад t-го наблюдения определяется как

$$l_t(\theta) = -\frac{1}{2}\ln\sigma_t^2(\theta) - \frac{\varepsilon_t^2(\theta)}{2\sigma_t^2(\theta)}. \tag{2}$$

$l(\theta_0)$ совпадает с логарифмом совместной плотности распределения вектора $y_n,\ldots,y_1$. Градиент и гессиан критериальной функции также составлены из вкладов отдельных наблюдений:

$$g(\theta) = \sum\nolimits_{t=1}^{n} g_t(\theta) \qquad H(\theta) = \sum\nolimits_{t=1}^{n} h_t(\theta),$$

вклады наблюдения t записываются как

$$g_t(\theta) \equiv \nabla_\theta l_t(\theta) = \frac{1}{2}\left[\frac{\varepsilon_t^2(\theta)}{\sigma_t^2(\theta)} - 1\right]\nabla_\theta \ln\sigma_t^2(\theta) + \left[\frac{\varepsilon_t(\theta)}{\sigma_t^2(\theta)}\right]\nabla_\theta \mu_t(\theta) \tag{3}$$

$$\begin{aligned} h_t(\theta) \equiv \nabla_{\theta\theta}^2 l_t(\theta) = {} & \frac{1}{2}\left[\frac{\varepsilon_t^2(\theta)}{\sigma_t^2(\theta)} - 1\right]\nabla_{\theta\theta}^2 \ln\sigma_t^2(\theta) + \left[\frac{\varepsilon_t(\theta)}{\sigma_t^2(\theta)}\right]\nabla_{\theta\theta}^2 \mu_t(\theta) - \\ & -\frac{1}{2}\frac{\varepsilon_t^2(\theta)}{\sigma_t^2(\theta)}\nabla_\theta \ln\sigma_t^2(\theta)\nabla_\theta \ln\sigma_t^2(\theta)^{\mathrm{T}} \\ & -\left[\frac{\varepsilon_t(\theta)}{\sigma_t^2(\theta)}\right]\left(\nabla_\theta\mu_t(\theta)\nabla_\theta \ln\sigma_t^2(\theta)^{\mathrm{T}} + \nabla_\theta\mu_t(\theta)^{\mathrm{T}}\nabla_\theta \ln\sigma_t^2(\theta)\right) \\ & -\frac{1}{\sigma_t^2(\theta)}\nabla_\theta\mu_t(\theta)\nabla_\theta\mu_t(\theta)^{\mathrm{T}}. \end{aligned} \tag{4}$$

Вычислим условные ожидания (3) и (4) в точке истинных параметров. Значения функций μ_t и σ_t^2, производные этих функций по θ предопределены к моменту t. Выражения, заключенные в квадратные скобки, обращаются в нуль. Таким образом,

$$E\left(g_t(\theta_0)\middle|\Omega_t\right) = 0 \tag{5}$$

$$\begin{aligned} H_t \equiv E\left(h_t(\theta_0)\middle|\Omega_t\right) = \\ = -0.5\nabla_\theta \ln\sigma_t^2(\theta_0)\,\nabla_\theta \ln\sigma_t^2(\theta_0)^{\mathrm{T}} - \sigma_t^{-2}(\theta_0)\nabla_\theta\mu_t(\theta_0)\nabla_\theta\mu_t(\theta_0)^{\mathrm{T}} \end{aligned} \tag{6}$$

Обозначим I_t матрицу условной ковариации вклада t-го наблюдения градиент:

$$I_t \equiv V\left(g(\theta_0)\middle|\Omega_t\right) = E\left(g_t(\theta_0)g_t(\theta_0)^{\mathrm{T}}\middle|\Omega_t\right). \qquad (7)$$

Вследствие (5) безусловное ожидание градиента критериальной функции равно нулю:

$$E\left(g(\theta_0)\right) = 0. \qquad (8)$$

Определим информационную матрицу как безусловную ковариацию градиента в точке θ_0:

$$I \equiv V\left(g(\theta_0)\right) = E\left(g(\theta_0)g(\theta_0)^{\mathrm{T}}\right). \qquad (9)$$

Поскольку последовательность вкладов наблюдений в градиент критериальной функции серийно не коррелирует (равенство 5), информационная матрица может быть также вычислена по формуле

$$I = \sum\nolimits_{t=1}^{n} E\left(g_t(\theta_0)g_t(\theta_0)^{\mathrm{T}}\right). \qquad (10)$$

Для дальнейшего изложения существенно, что соотношения (3)-(10) были выведены вне связи с гипотезой (*N*).

Информационная матрица

В теории метода максимального правдоподобия известно равенство

$$I_t = -H_t \qquad (11),$$

которое нетрудно установить и в данном случае. Внешнее произведение вклада t-го наблюдения в градиент $g_t g_t^{\mathrm{T}}$ содержит ε_t в степени от первой до четвертой. Условные ожидание и дисперсия истинных остатков определены равенствами (M.3)-(M.4), тогда как для вычисления третьего и четвертого моментов необходимо прибегнуть к дополнительному предположению (N). (11) влечет равенство

$$I = -E\left(H(\theta_0)\right) \qquad (12),$$

однако вычислить полные ожидания не представляется возможным.

В некоторых случаях вектор θ можно разделить на компоненты b и γ, первая из которых параметризует условное среднее, вторая — условную дисперсию. Тогда $\nabla_\gamma \mu_t(\theta) = \mathbf{0}$, однако даже в этом случае $\nabla_b \sigma_t^2(\theta) \neq \mathbf{0}$. Если, кроме того, распределение $y_t|\Omega_t$ симметрично, выполнены некоторые ограничения на функциональную форму σ^2, то информационная матрица является блочно-диагональной:

$$I_{b\gamma} = \mathbf{0}.$$

Engle (1982) приводит набор достаточных условий и формальное доказательство для ARCH модели. Блочная диагональность информационной матрицы между параметрами b и γ означает, что оценки параметров среднего состоятельны даже при неверной спецификации функции условной дисперсии. В частности, оценки методом наименьших квадратов являются состоятельными, однако, выигрыш в эффективности от использования ММП по сравнению с МНК может оказаться сколь угодно великим. Более того, оценки γ, полученные на основе состоятельных, но не эффективных оценок b (например, на основе МНК), сохраняют свойство асимптотической эффективности.

Не обладают свойством блочной диагональности информационные матрицы ARCH-M моделей: для них разбиения $\theta=(b,\gamma)$ не существует. Иное исключение составляют EGARCH модели и другие, в которых σ_t^2 является асимметричной функцией остатков. Присутствие ошибок в спецификации функции условной дисперсии приводит к несостоятельности оценок параметров среднего, и наоборот. Состоятельное оценивание требует верной спецификации полной модели.

При определенных условиях регулярности[1] оценки максимального правдоподобия состоятельны, асимптотически нормальны и эффективны с асимптотической матрицей ковариации[2]

$$\lim_{n} V\left(\sqrt{n}\left(\hat{\theta}^{\mathrm{ML}}-\theta_0\right)\right)=\left(\operatorname*{plim}_{\mathrm{n}}\ n^{-1}\sum\nolimits_{t=1}^{n} I_t\right)^{-1}=-\left(\operatorname*{plim}_{\mathrm{n}}\ n^{-1}\sum\nolimits_{t=1}^{n} H_t\right)^{-1} \quad (13).$$

Существуют два базовых способа состоятельно оценить информационную матрицу[3]. Первый способ основан на связи информационной матрицы и гессиана (равенства 11–12). В качестве оценки приемлема матрица $-\hat{H}$, где

$$\hat{H}=-\sum\nolimits_{t=1}^{n}\left(0.5\nabla_{\theta}\ln\sigma_t^2\left(\hat{\theta}\right)\nabla_{\theta}\ln\sigma_t^2\left(\hat{\theta}\right)^{\mathrm{T}}+\sigma_t^{-2}\left(\hat{\theta}\right)\nabla_{\theta}\mu_t\left(\hat{\theta}\right)\nabla_{\theta}\mu_t\left(\hat{\theta}\right)^{\mathrm{T}}\right). \quad (14)$$

Данная оценка построена как сумма выражений вида $-H_t$, причем участвующие в H_t функции исчислены в точке $\hat{\theta}$.

Второй способ вытекает из равенства (10) для информационной матрицы. Опустив знаки мат. ожиданий и воспользовавшись оценками неизвестных параметров, приходим к

$$\hat{B}=\sum\nolimits_{t=1}^{n} g_t\left(\hat{\theta}\right)g_t\left(\hat{\theta}\right)^{\mathrm{T}}. \quad (15)$$

Второй способ восходит к статье Berndt, Hall, Hall, and Hausman и потому называется BHHH. Другое название — метод внешнего произведения градиента (outer product of the gradient, OPG).

Как $-\hat{H}$, так и $\hat{B}$ состоятельны для *I*. Обращенная информационная матрица служит оценкой матрицы ковариации оценок максимального правдоподобия $V\left(\hat{\theta}\right)$. Возможны, следовательно, два выражения:

$$\hat{V}^{\mathrm{ML}}=-\hat{H}^{-1} \quad (16.\text{а})$$

$$\hat{V}^{\mathrm{ML}}=\hat{B}^{-1}. \quad (16.\text{б})$$

Нарушения гипотезы об условной нормальности: метод квази-максимального правдоподобия.

Гипотезу (*N*) позволяет протестировать верное в нуле свойство независимости и нормальной распределенности стандартизованных остатков. Как правило, гипотеза отклоняется из-за того, что оцененные $\hat{z}_t$ демонстрируют положительный куртозис. Реже причиной отклонения нулевой гипотезы становится асимметрия (необходимые ссылки представлены Bollerslev T., et al 1992).

Рядом авторов реализован ММП в предположении о том, что плотность распределения $z_t\left(\theta_0\right)$ принадлежит некоторому параметризованному семейству $f(z,\upsilon)$. Так, Bollerslev T. (1987), Nelson D.B. (1991), Bollerslev T., Engle R.F., Nelson D.B. (1993) применяют, соответственно, *t*-Стьюдента, GED (Generalized Error Distribution), и обобщенное *t*-Стьюдента распределения. Плотности *t* и GED имеют единственный параметр, регулирующий величину куртозиса, плотность обобщенного *t*-распределения имеет два параметра и включает *t* и GED как частные случаи. Параметры υ и θ оцениваются одновременно максимизацией логарифмической функции правдоподобия

$$l(\theta,\upsilon)=\sum\nolimits_{t=1}^{n}\left[\ln f\left(z_t(\theta);\upsilon\right)-0.5\ln\sigma_t^2(\theta)\right].$$

---

[1] Проверка условий регулярности, обеспечивающих свойства оптимальности оценок максимального правдоподобия, оказалась экстремально сложной для класса ARCH моделей.

[2] Формально (12) есть асимптотическая матрица ковариации $\sqrt{n}\left(\hat{\theta}-\theta_0\right)$, а не $\hat{\theta}$, но для простоты изложения мы будем называть ее асимптотической матрицей ковариации оценок.

[3] Здесь и далее под состоятельной понимаем такую оценку $\hat{I}$, что $n\quad I\quad p\quad n\quad I$.

С точки зрения асимптотической эффективности ММП с корректно определенной функцией плотности $f(\cdot,\cdot)$ является наилучшим решением. Реализация его, однако, технически крайне трудна: поскольку производные соответствующих функций правдоподобия не могут быть представлены аналитически, для максимизации их прибегают к методам численного дифференцирования.

Качество подбора функции плотности $f(\cdot,\cdot)$ можно установить, сравнивая фактическое и ожидаемое количества таких значений $|z_t|$, которые превосходят некоторое заданное Z. В этом смысле GED не вполне адекватно отражает частоту «хвостовых событий»: фактическое число выбросов гораздо больше, чем если бы z_t были реализациями GED-распределенной случайной величины со значением параметра υ, равным оцененному. Кроме того, *t* и GED симметричны, тогда как асимметрия — одна из важнейших особенностей российских финансовых временных рядов. Среди других распределений, применяемых при оценивании ARCH модели, — смесь нормального и логнормального, нормального и Пуассона распределений.

Установлено, что максимизация критериальной функции (1)-(2) приводит к состоятельным и асимптотически нормальным оценкам независимо от того, как именно распределены случайные величины y_t. В тех случаях, когда истинное распределение y_t неизвестно, эту процедуру принято называть методом квази- (псевдо-) максимального правдоподобия (МКМП). Отличие ее от традиционного ММП состоит в матрице ковариации оценок:

$$\lim_n V\left(\sqrt{n}\left(\hat\theta^{\,\mathrm{QML}} - \theta_0\right)\right) =$$

$$= \left(\operatorname*{plim}_n\, n^{-1}\sum\nolimits_{t=1}^{n} H_t\right)^{-1}\left(\operatorname*{plim}_n\, n^{-1}\sum\nolimits_{t=1}^{n} I_t\right)\left(\operatorname*{plim}_n\, n^{-1}\sum\nolimits_{t=1}^{n} H_t\right)^{-1}. \tag{17}$$

Равенство $I_t = -H_t$ неверно в общем случае без предположения об условной нормальности, поэтому (17) не эквивалентно (13). МКМП неизбежно приводит к потере асимптотической эффективности. Потери эффективности, возникающие, в частности, при *t*-распределенных ошибках невелики, однако могут быть весьма существенными, если распределение ошибок асимметрично.

Для вывода равенств (5), (6), (8), и (10) предположение (*N*) не привлекалось, все они являются следствием верной спецификации функций условного среднего и дисперсии, т.е. (*M*.1)–(*M*.2). Поэтому матрица $\hat H$ остается состоятельной для гессиана, $\hat B$ — состоятельной для информационной матрицы. Однако $-\hat H$ и $\hat B$ не являются асимптотически эквивалентными, как не являются асимптотически эквивалентными минус гессиан и информационная матрица. Оценкой ковариацонной матрицы КМП-оценок служит

$$\hat V^{\,QML} = \hat H^{-1}\hat B\hat H^{-1}. \tag{18}$$

Оценка (18) устойчива к нарушению гипотезы об условной нормальности в том смысле, что остается состоятельной для ковариации оценок, полученных максимизацией (1)-(2). Оценки $-\hat H^{-1}$ и $\hat B^{-1}$ при указанном нарушении свойства состоятельности не сохраняют.

Тестирование

Асимптотическая нормальность оценок КМП позволяет воспользоваться стандартными процедурами. Пусть нулевая гипотеза формулируется как

$$r\left(\theta_0\right) = \mathbf{0}, \tag{19}$$

где $r:\Theta\to R^l$ дифференцируема на $\operatorname{int}(\Theta)$ и $l<m$. Если $\theta_0 \in \operatorname{int}(\Theta)$ и матрица $\nabla_\theta r\left(\theta_0\right)$ имеет ранг *l*, то применима статистика Вальда

$$W = nr\left(\hat\theta\right)^{\mathrm T}\left(\nabla_\theta r\left(\hat\theta\right)\hat{\mathrm V}^{-1}\nabla_\theta r\left(\hat\theta\right)^{\mathrm T}\right)^{-1} r\left(\hat\theta\right), \tag{20}$$

где $\hat{\theta}$ — оценки параметров при альтернативной гипотезе (оценки полной модели, без ограничений (19)), $\hat{V}$ — состоятельная оценка ковариации $\hat{\theta}$. При верной гипотезе (*N*) следует использовать $\hat{V}^{ML}$, в противном случае — $\hat{V}^{QML}$. Верно предположение (*N*) или нет, в нуле статистика Вальда имеет асимптотическое хи-квадрат распределение с *m-l* степенями свободы. Тест Вальда

$$\begin{aligned} W < \tau_{0.05} &\Rightarrow \text{H}_0 \textit{ принимается} \\ W > \tau_{0.05} &\Rightarrow \text{H}_0 \textit{ отвергается}, \end{aligned} \tag{21}$$

где $\tau_{0.05}$ - 5%-й квантиль χ^2_{m-l} распределения, характеризуется асимптотической ошибкой первого рода 5 %: вероятность отвергнуть H_0 тогда как она верна

$$P\left(W > \tau_{0.05} | \text{H}_0\right)$$

при увеличении числа наблюдений сходится к

$$P\left(\chi^2_{m-l} > \tau_{0.05} | \text{H}_0\right) = 0.05.$$

Асимптотические результаты могут оказаться неприемлемыми для малых выборок и при некорректном выборе матрицы $\hat{V}$. Bollerslev T. и Wooldridge J.M. (1992) сообщают результаты имитационных экспериментов, проливающих свет на характер искажений, связанных с использованием в тестировании несостоятельных оценок $\hat{V}^{\text{ML}}$ при нарушении гипотезы (N). Общий вывод исследования состоит в следующем: ковариационные матрицы $\hat{V}^{\text{ML}}$ систематически недооценивают истинные размеры стандартных ошибок.

Схема исследования такова. Построены 1000 реализаций AR(1)-GARCH(1,1) процесса, имеющего условное t_5 распределение. Для каждой реализации вычислены

- оценки параметров истинной модели;
- ковариационные матрицы оценок трех типов: $\hat{H}^{-1}\hat{B}\hat{H}^{-1}$, $-\hat{H}^{-1}$, $\hat{B}$. Эти типы будем вслед за авторами называть соответственно RB (от robust — устойчивый), HE (от hessian — гессиан), OPG (от outer product of the gradient — внешнее произведение градиента).
- статистики Вальда для верной нулевой гипотезы H_0 (19) трех типов по общей формуле (20). Тип статистики определяется типом оценки вариационной матрицы, применяемой в (20) — RB, HE, или OPG.

Имитационные эксперименты позволяют построить эмпирические распределения трех вариантов статистики Вальда при верной нулевой гипотезе, которые затем сопоставляются с хи-квадрат распределением. Полученные распределения имеют более толстые хвосты, чем χ^2_{m-l}. Так, например, доля реализаций статистики Вальда типов HE и OPG, лежащих правее 5%-го квантиля χ^2_{m-l}, больше 0.05, скажем, 0.1. Это означает, что тест (18) имеет ошибку первого рода 10%, а не 5%. Иными словами, вероятность отвергнуть нулевую гипотезу H_0 в то время как она верна составляет 0.1:

$$P\left(W > \tau_{0.05} | \text{H}_0\right) = 0.1.$$

Распределение RB-статистики близко к χ^2_{m-l}. Использование устойчивой формы статистики Вальда, как и следовало ожидать, предпочтительнее двух других, причем OPG-статистика наименее точна. Таким образом, как $-\hat{H}^{-1}$, так и $\hat{B}^{-1}$ систематически преуменьшают вариацию оценок и вводят в заблуждение относительно того уровня значимости, с которым нуль может быть отвергнут.

Точность всех форм статистик снижается при переходе к несимметричному χ^2_1 распределению $y_t | \Omega_t$. Аналогичные результаты были получены и для LM статистики.

Обобщенный метод моментов.

Обобщенный метод моментов (ОММ) обладает следующими достоинствами:

(i) не требует явных предположений относительно плотности условного распределения и допускает присутствие ненулевых куртозиса и асимметрии;

(ii) использует лишь производные первого порядка функций μ_t и σ_t^2 и позволяет избежать тем самым применения методов численного дифференцирования;

(iii) асимптотически более эффективен, чем МКМП. Проверка метода на выборочных данных, демонстрирующих экстремально высокие коэффициенты асимметрии и куртозиса, свидетельствует о значительном выигрыше в эффективности.

Дополним модель (M.1)-(M.4) предположениями относительно третьего и четвертого моментов распределения $y_t|\Omega_t$:

$$E\left(\varepsilon_t^3(\theta_0)\middle|\Omega_t\right) = a_0\sigma_t^3(\theta_0) \qquad (M.5)$$

$$E\left(\varepsilon_t^4(\theta_0)\middle|\Omega_t\right) = (k_0+3)\sigma_t^4(\theta_0), \qquad (M.6)$$

где a_0, k_0 — постоянные коэффициенты асимметрии и куртозиса. Стандартизованные остатки в точке θ_0 тогда имеют первые четыре момента, равные соответственно 0, 1, a_0, k_0+3. Гипотеза нормальности формулируется как $a_0 = 0,\ k_0 = 0$.

Спецификация (M) обеспечивает две группы уравнений, идентифицирующих истинные значения параметров. Пусть

$$f_t(y_t, \Omega_t; \theta) = \left(\varepsilon_t(\theta) \quad \varepsilon_t^2(\theta) - \sigma_t^2(\theta)\right)$$

– строка из двух элементов, которую далее будем обозначать, опуская аргументы Ω_t и y_t. Мат. ожидания $E\left(f_t(\theta)\middle|\Omega_t\right)$ существуют для всех $\theta \in \Theta$ и обращаются в ноль единственным $\theta_0 \in \Theta$. В этом смысле система уравнений

$$E\left(f_t(\theta)\middle|\Omega_t\right) = 0 \qquad (19)$$

идентифицирует истинный вектор параметров[4]. Определим условные матрицу ковариации и якобиан f_t в точке θ_0 как

$$\Lambda_t \equiv V\left(f_t(\theta_0)\middle|\Omega_t\right) = \begin{bmatrix} \sigma_t^2(\theta_0) & a_0\sigma_t^3(\theta_0) \\ a_0\sigma_t^3(\theta_0) & (k_0+2)\sigma_t^4(\theta_0) \end{bmatrix} \qquad (20)$$

$$J_t \equiv E\left(\nabla_\theta f_t(\theta_0)\middle|\Omega_t\right) = \left(-\nabla_\theta \mu_t(\theta_0) \quad -\nabla_\theta \sigma_t^2(\theta_0)\right). \qquad (21)$$

Класс оценок ОММ порождается различными наборами инструментальных переменных, выбор которых ограничен последовательностью Ω_t. Асимптотическая ковариационная матрица ОММ оценок ограничена снизу, причем существует набор оптимальных инструментов, приводящий к эффективным оценкам.

Пусть l инструментов могут быть организованы в матрицу $\left[W_1 \vdots \dots \vdots W_n\right]$ размерности $l \times 2n$, где W_t - часть матрицы размерности $l \times 2$, относящаяся к наблюдению t (вклад данного наблюдения в матрицу инструментов). Требуется, чтобы число инструментов было не меньше числа оцениваемых параметров, т.е. $l \geq m$, и чтобы к моменту t значения W_t были известны: $W_t \subset \Omega_t$; можно указать бесконечное число инструментальных переменных. Эмпирические моменты, соответствующие данному набору инструментов могут быть выражены как

[4] Систематическое изложение метода моментов приведено в Davidson, McKinnon, откуда заимствованы многие обозначения. Авторы учебника ограничиваются теми моделями, для которых имеется единственный набор уравнений, определяющий параметры (это модели условного среднего). Рассмотрение двух или более уравнений (3.19) не приводит к концептуальным трудностям, но требует несколько более сложной системы обозначений.

$$\sum_{t=1}^{n} W_t f_t(\theta)^T . \tag{21}$$

Матрица условной ковариации эмпирических моментов в точке θ_0 равна

$$\sum_{t=1}^{n} V\left(W_t f_t(\theta_0)^T \middle| \Omega_t\right) = \sum_{t=1}^{n} W_t \Lambda_t W_t^T . \tag{22}$$

Если l=m, то оценки $\hat{\theta}$ находятся решением системы m уравнений

$$\sum_{t=1}^{n} W_t f_t(\theta)^T = \mathbf{0} , \tag{23}$$

если l>m, то минимизацией критериальной функции – квадратичной формы, построенной из (21) и (22):

$$Q(\theta) = n^{-2}\left(\sum_{t=1}^{n} f_t(\theta) W_t^T\right)\left(\sum_{t=1}^{n} W_t \Lambda_t W_t^T\right)^{-1}\left(\sum_{t=1}^{n} W_t f_t(\theta)^T\right). \tag{24}$$

При любом выборе инструментов оценки, определяемые (23) или (24), состоятельны и асимптотически нормальны с асимптотической матрицей ковариации

$$\lim V\left(\sqrt{n}\left(\hat{\theta} - \theta_0\right)\right) =$$

$$= \left[\left(p\lim n^{-1}\sum_{t=1}^{n} J_t W_t^T\right)\left(p\lim n^{-1}\sum_{t=1}^{n} W_t \Lambda_t W_t^T\right)^{-1}\left(p\lim n^{-1}\sum_{t=1}^{n} W_t J_t^T\right)\right]^{-1} . \tag{25}$$

Инструменты W, такие что $W_t = J_t \Lambda_t^{-1}$, приводят к оценкам, эффективным в классе ОММ. Существует ровно m оптимальных инструментов, поэтому эффективные оценки находятся решением системы

$$\sum_{t=1}^{n} J_t \Lambda_t^{-1} f_t(\theta)^T = \mathbf{0} . \tag{26}$$

Асимптотическая матрица ковариации таких оценок меньше, чем при любом ином выборе инструментов:

$$\lim V\left(\sqrt{n}\left(\hat{\theta} - \theta_0\right)\right) = \left(p\lim\, n^{-1}\sum_{t=1}^{n} J_t \Lambda_t^{-1} J_t^T\right)^{-1} . \tag{27}$$

Воспользуемся матрицами $W = \left[W_1 \vdots \ \ldots \ \vdots W_n\right]$ размерности l×2n, $J = \left[J_1 \vdots \ \ldots \ \vdots J_n\right]$ размерности m×2n, блочно-диагональной матрицей Λ размерности 2n×2n с диагональными блоками Λ_t. Тогда участвующие в (25) и (27) суммы записываются как WJ^T, $W\Lambda W^T$, $J\Lambda^{-1}J^T$. Опустим знаки $p$lim**,** множители n^{-1} и рассмотрим разность

$$J\Lambda^{-1}J^T - JW^T\left(W\Lambda W^T\right)^{-1}WJ^T$$

между обращенными матрицами ковариации, относящимися к оптимальному и произвольному наборам инструментов, соответственно. Если симметричная 2n×2n матрица Ψ такова, что $\Psi^2 = \Lambda$, то разность данная равна

$$J\Psi^{-1}\left(I - \Psi W^T\left(W\Psi^2 W^T\right)^{-1}W\Psi\right)\Psi^{-1}J^T .$$

Эта матрица положительно полуопределена, поскольку матрица в больших скобках идемпотентна. Отсюда немедленно следует положительная полуопределенность

$$\left[JW^T\left(W\Lambda W^T\right)^{-1}WJ^T\right]^{-1} - \left[J\Lambda^{-1}J^T\right]^{-1} .$$

Оптимальные в классе ОММ оценки асимптотически более эффективны, чем оценки МКМП. Достаточно показать, что асимптотическая матрица ковариации последних приводима к виду (25) с помощью какого-либо набора неоптимальных инструментов. Вклад наблюдения t в этот набор инструментов представляет собой матрицу $J_t\Lambda_t^{-1}$, вычисленную при $a_0 = 0,\ k_0 = 0$:

$$W_t = J_t\Lambda_t^{-1}\Big|_{\substack{a_o=0\\k_o=0}} = \left(-\sigma_t^{-2}(\theta_0)\nabla_\theta\mu_t(\theta_0) \quad -0.5\sigma_t^{-2}(\theta_0)\nabla_\theta \ln\sigma_t^2(\theta_0)\right). \quad (28)$$

Если коэффициенты асимметрии и куртозиса действительно равны нулю, то такой набор инструментов является оптимальным. Следовательно, при верной гипотезе (N) методы моментов и максимального правдоподобия асимптотически эквивалентны.

Выбор инструментов (28) приводит к следующим совпадениям:

- эмпирических моментов и градиента критериальной функции (3.1)-(3.2) в точке истинных параметров с точностью до знака «минус»: $W_t f_t(\theta_0) = -g_t(\theta_0)$;
- их условных ковариационных матриц: $W_t\Lambda_t W_t^T = I_t$, и
- условных вторых производных: $J_t W_t^T = H_t$.

Следовательно, совпадают и вероятностные пределы, участвующие в (15) и (25). Итак, (25) при соответствующем выборе инструментальных переменных характеризует асимптотическую ковариационную матрицу МКМП -оценок.

Davidson и MacKinnon предлагают двухшаговую или итеративную процедуры вычисления оценок ОММ. Требуется построить состоятельные, но, возможно, неэффективные оценки, используя их, определить приблизительно оптимальные инструменты; с помощью найденных инструментов вычислить оценки параметров. Если исходные оценки не очень точны, желательно повторить процедуру несколько раз.

Пусть имеются $\hat{\theta}$, с помощью которых можно состоятельно оценить коэффициенты асимметрии и куртозиса выборочными моментами стандартизованных остатков:

$$\hat{a} = \frac{\hat{\eta}_3}{\hat{\eta}_2^{1.5}}, \quad \hat{k} = \frac{\hat{\eta}_4}{\hat{\eta}_2^2} - 3. \quad (29)$$

Для всех $\theta\in\Theta$ определим выражения $\Lambda_t(\theta)$ и $J_t(\theta)$, причем для вычисления первого из них вместо истинных значений $a_0,\ k_0$ будем использовать оценки (29). Новые оценки параметров θ предлагается вычислять как решение системы

$$\sum_{t=1}^{n} J_t(\theta)\Lambda_t^{-1}(\theta) f_t(\theta)^T = \mathbf{0},$$

где t-е слагаемое равно

$$\left(\hat{k} - \hat{a}^2 + 2\right)^{-1}\nabla_\theta\mu_t(\theta)\left[\hat{a}\frac{\varepsilon_t^2(\theta)}{\sigma_t^3(\theta)} - \left(\hat{k}+2\right)\frac{\varepsilon_t(\theta)}{\sigma_t^2(\theta)} - \hat{a}\frac{1}{\sigma_t(\theta)}\right] +$$
$$+\left(\hat{k} - \hat{a}^2 + 2\right)^{-1}\nabla_\theta \ln\sigma_t^2(\theta)\left[1 + \hat{a}\frac{\varepsilon_t(\theta)}{\sigma_t(\theta)} - \frac{\varepsilon_t^2(\theta)}{\sigma_t^2(\theta)}\right]. \quad (31)$$

При этом оценки оптимальных инструментов и оценки параметров определяются одновременно. Такое вычисление эквивалентно реализации итеративной процедуры, в которой оценки на i-м шаге θ^i определяются с помощью инструментов, оцененных на (i-1)-м шаге как решение

$$\sum_{t=1}^{n} J_t\left(\theta^{i-1}\right)\Lambda_t^{-1}\left(\theta^{i-1}\right) f_t(\theta)^T = \mathbf{0}.$$

Полученные оценки приводят к новым значениям коэффициентов асимметрии и куртозиса, и процедура может быть повторена. На практике уже после третьей итерации достигается вполне удовлетворительная сходимость.

В качестве исходных оценок приемлемы оценки максимального правдоподобия. При $a = 0,\ k = 0$ вклад t-го наблюдения в эмпирический момент (31) совпадает со вкладом в градиент логарифмической функции правдоподобия. Следовательно, оценки максимального правдоподобия могут быть получены на первом шаге решением (30) при $a = 0,\ k = 0$.

Оценкой ковариационной матрицы для $\hat{\theta}$ служит

$$\hat{V}^{GMM} = \left(\sum_{t=1}^{n} J_t(\hat{\theta}) \Lambda_t^{-1}(\hat{\theta}) J_t(\hat{\theta})^T \right)^{-1},$$

где $$J_t(\theta)\Lambda_t^{-1}(\theta)J_t(\theta)^T =$$

$$= \left(\hat{k} - \hat{a}^2 + 2\right)^{-1}\left(\hat{k} + 2\right)\sigma_t^{-2}(\theta)\nabla_\theta \mu_t(\theta)\nabla_\theta \mu_t(\theta)^T -$$

$$-\left(\hat{k} - \hat{a}^2 + 2\right)^{-1}\hat{a}\sigma_t^{-1}(\theta)\left(\nabla_\theta \mu_t(\theta)\nabla_\theta \ln\sigma_t^2(\theta)^T + \nabla_\theta \ln\sigma_t^2(\theta)\nabla_\theta \mu_t(\theta)^T\right) +$$

$$+\left(\hat{k} - \hat{a}^2 + 2\right)^{-1}\nabla_\theta \ln\sigma_t^2(\theta)\nabla_\theta \ln\sigma_t^2(\theta)^T.$$

Примеры.

Мы попытались сравнить эффективность методов квази–максимального правдоподобия и моментов на примере нескольких временных рядов, сгенерированных российским финансовым рынком. Отличительной особенностью этих временных рядов являются экстремально высокие значения коэффициентов асимметрии и куртозиса.

Применяется EGARCH параметризация, логарифм условной дисперсии представлен как процесс AR(1). Функция условного среднего включает авторегрессионную и ARCH in mean компоненты:

$$\begin{aligned} &\ln\sigma_t^2 = \omega + \theta z_{t-1} + \gamma\left(\left|z_{t-1}\right| + E\left|z_{t-1}\right|\right) + \phi\ln\sigma_{t-1}^2 \\ &\mu_t = b + \beta y_{t-1} + \delta\sigma_t. \end{aligned} \qquad \text{(M)}$$

Стандартизованные остатки z_t предполагаются независимыми и одинаково распределенными со средним 0 дисперсией 1.

Данные.

- Индекс РТС. Привлечены ежедневные (на момент закрытия) данные по индексу Российской Торговой Системы за период 04.05.95–05.11.98; временной ряд содержит 793 наблюдения. Переменная *РТС* была получена на основе этих данных как
$$y_t = 100\left[P_t / P_{t-1} - 1\right].$$
- Доходность ГКО. Переменная представляет собой взвешенное среднее доходности по выпускам ГКО со сроками погашения от 1 до 360 дней; в качестве весов приняты текущие доли выпусков в суммарном обороте торгов. Временной ряд содержит 403 наблюдения за период 5.01.97-14.08.98.
- Цены ГКО. Переменная представляет собой взвешенное среднее цен выпусков ГКО со сроками погашения от 1 до 360 дней; в качестве весов приняты текущие доли выпусков в суммарном обороте торгов. Временной ряд содержит 403 наблюдения за период 5.01.97-14.08.98.

Характер условных распределений.

- Индекс РТС. Коэффициенты асимметрии и куртозиса стандартизованных остатков $\hat{z}_t$ равны соответственно 0,1231 и 1,7533. Гипотеза (*N*) о нормальной распределенности отклонена в пользу симметричного лептокуртического распределения. Данный вывод согласуется с многочисленными наблюдениями, относящимися к отдаче вложений в фондовые индексы США: плотность условного распределения подобных процессов симметрична, но по сравнению с нормальной имеет более толстые хвосты (См., например, Bollerslev T., Chou R.Y., and Kroner K.F., 1992; Bollerslev T., Engle R.F., and Nelson D. B., 1993).

- Доходность ГКО. Стандартизованные остатки демонстрируют экстремальные коэффициенты асимметрии и куртозиса: 1,9308 и 9,2601.
- Цены ГКО. Коэффициенты асимметрии и куртозиса стандартизованных остатков равны, соответственно, 2,5162 и -0,8475. Столь яркая асимметрия является, по–видимому, особенностью российских финансовых временных рядов: по данным зарубежных отчетов асимметричные распределения достаточно редки.

Методы квази-максимального правдоподобия и моментов: оценки модели (м).

Провал гипотезы (*N*) в каждом из рассмотренных случаев мотивирует применение метода квази–максимального правдоподобия (МКМП) и обобщенного метода моментов (ОММ). Результаты оценивания модели данными методами представлены в таблицах 1–3. Приведены стандартные отклонения оценок параметров, соответствующие трем различным ковариационным матрицам:

- $-\hat{H}^{-1}$. Данная форма (HE) ковариационной матрицы основана на гессиане логарифмической функции правдоподобия, используется в методе максимального правдоподобия и теряет свойство состоятельности при нарушении гипотезы (*N*) (столбец ML St.Dev.).
- $\hat{H}^{-1}\hat{B}\hat{H}^{-1}$. Такая форма (RB) устойчива к нарушению гипотезы (N) и является ковариационной матрицей МКМП (столбец QML St.Dev.).
- Ковариационной матрице ОММ (столбец GMM St.Dev.).

В последнем столбце показаны отношения стандартных отклонений оценок ОММ к стандартным отклонениям оценок МКМП (столбец GMM/QML).

HE–форма ковариационной матрицы систематически преуменьшает стандартные отклонения оценок метода квази–максимального правдоподобия. Смещение особенно велико, если распределение инноваций асимметрично. Данное сравнение иллюстрирует результаты Bollerslev T. и Wooldridge J.M. (1992).

Для большинства параметров стандартные отклонения оценок ОММ меньше стандартных отклонений оценок МКМП. Отношения их лежат в интервале: от 0,8077 до 1,0194 (РТС), от 0,5458 до 1,0647 (доходность ГКО), от 0,4824 до 1,1572 (цены ГКО). Общую закономерность нарушают: константа в модели для РТС (1,0194), θ для доходности ГКО (1,0647), δ для цен ГКО (1,1572).

Нам представляется возможным на основании теоретических и практических результатов выдвинуть следующую **гипотезу**. Обобщенный метод моментов более эффективен, чем метод квази–максимального правдоподобия не только асимптотически, но также и **для конечных выборок**. Выигрыш в эффективности тем больше, чем сильнее отклонение условного распределения процесса от нормального. Проведение экспериментов Монте Карло с целью верификации данной гипотезы, однако, не входит в план настоящей работы и остается перспективной задачей.

Тесты на спецификацию.

С целью обнаружения ошибок в спецификации модели мы проверяем набор из 12 условий ортогональности, сформулированных относительно z_t. Тестируются предположения о нулевом среднем, единичной дисперсии, отсутствии автокорреляции в стандартизованных ошибках и их квадратах до пятого порядка. Результаты тестирования (t–статистики) для трех временных рядов представлены в таблице 4.

- Индекс РТС. При 10 процентном уровне значимости ни одно из условий не может быть отвергнуто.
- Доходность ГКО. Достоверно могут быть отвергнуты два условия ортогональности:

$$E\left(z_t z_{t-2}\right) = 0$$

$$E\left[\left(z_t^2 - 1\right)\left(z_{t-5}^2 - 1\right)\right] = 0$$

Нарушение первого из них сигнализирует о пропущенной переменной y_{t-2} в уравнении условного среднего. Нарушение второго условия свидетельствует об ограничительности AR(1) репрезентации логарифма условной дисперсии. В более общей постановке $\ln\sigma_t^2$ моделируется как процесс ARMA(p,q), причем степени авторегрессии и скользящего среднего определяются на основе критериев Schwartz или Akaike. Выбор p=1, q=0 в настоящей работе обусловлен лишь техническими ограничениями: формального тестирования порядков p и q не проводилось.

Цены ГКО. Нарушены два условия ортогональности:

$$E\left(z_t z_{t-2}\right)=0$$

$$E\left[\left(z_t^2-1\right)\left(z_{t-1}^2-1\right)\right]=0$$

Автокорреляция квадратов стандартизованных остатков первого порядка свидетельствует, по-видимому, о неудовлетворительности EGARCH параметризации условной дисперсии; в частности, простая GARCH(1,1) модель приводит к меньшей t статистике для данного теста.

Таблица 1.

РТС. Методы квази-максимального правдоподобия и моментов.

	QMLE Estimate	*ML St.Dev.*	*QMLE St.Dev.*	*GMM Estimate*	*GMM St.Dev.*	*GMM / QML **
ω	0,211110	0,044967	0,070720	0,175348	0,057123	0,8077
θ	-0,024720	0,035895	0,060395	-0,030853	0,048906	0,8098
γ	0,498193	0,055513	0,086180	0,485103	0,074509	0,8646
ϕ	0,913217	0,019045	0,030457	0,919309	0,025042	0,8222
β	0,210746	0,038514	0,041496	0,272529	0,039843	0,9602
c	0,104878	0,083227	0,084562	0,109164	0,086201	1,0194

Таблица 2.

Доходность ГКО. Методы квази-максимального правдоподобия и моментов.

	QMLE Estimate	*ML St.Dev.*	*QMLE St.Dev.*	*GMM Estimate*	*GMM St.Dev.*	*GMM / QML **
ω	0,017306	0,016070	0,025163	0,009005	0,017849	0,7093
θ	0,241774	0,036985	0,061292	0,189461	0,065258	1,0647
γ	0,149522	0,034811	0,076191	0,052284	0,054374	0,7137
ϕ	0,991597	0,007343	0,011836	0,996323	0,007308	0,6175
β	0,724297	0,066200	0,118685	0,776934	0,067986	0,5728
c	3,980118	0,939140	1,484244	3,250043	0,961844	0,6480
δ	1,202641	0,328083	0,619242	1,029808	0,337989	0,5458

Таблица 3.

Цены ГКО. Методы квази-максимального правдоподобия и моментов.

	QML Estimate	*ML St.Dev.*	*QML St.Dev.*	*GMM Estimate*	*GMM St. Dev.*	*GMM / QML **
ω	0,013897	0,006430	0,009478	0,005152	0,006651	0,7017
θ	-0,057667	0,018708	0,041778	-0,053641	0,020155	0,4824
γ	0,086991	0,021409	0,031387	0,058078	0,021354	0,6803
ϕ	0,991719	0,004039	0,005536	0,996885	0,004042	0,7301
β	0,296550	0,058415	0,075724	0,289070	0,054709	0,7225
c	67,684135	5,663298	7,186756	68,669918	5,313474	0,7393
δ	-2,321332	0,340737	0,322848	-2,499164	0,373585	1,1572

QMLE Estimate — оценки МКМП.
ML St.Dev. — стандартные отклонения оценок МКМП по HE–форме ковариационной матрицы.
QML St.Dev. — стандартные отклонения оценок МКМП по RB–форме ковариационной матрицы.
GMM Estimate — оценки ОММ.
GMM St.Dev. — стандартные отклонения оценок ОММ.
GMM/QML — отношения стандартных отклонений оценок ОММ к стандартным отклонениям оценок МКМП.

Таблица 4

Тестирование спецификации моделей: значения t-статистик.

Условия	*Индекс РТС*	*Доходность ГКО*	*Цены ГКО*
$E(z_t) = 0$	-0,4297	0,8719	-0,1271
$E(z_t z_{t-1}) = 0$	1,5339	0,3303	0,6958
$E(z_t z_{t-2}) = 0$	0,2958	-1,9756	-1,6717
$E(z_t z_{t-3}) = 0$	1,6391	0,2427	-1,3434
$E(z_t z_{t-4}) = 0$	-0,1460	0,8085	0,7397
$E(z_t z_{t-5}) = 0$	-0,0537	0,9534	1,3670
$E(z_t^2) - 1 = 0$	0,1412	0,5830	0,1793
$E\left[(z_t^2 - 1)(z_{t-1}^2 - 1)\right] = 0$	0,8727	1,3365	2,1175
$E\left[(z_t^2 - 1)(z_{t-2}^2 - 1)\right] = 0$	-0,1589	1,5769	-0,9125
$E\left[(z_t^2 - 1)(z_{t-3}^2 - 1)\right] = 0$	-1,2705	0,5244	-0,6135
$E\left[(z_t^2 - 1)(z_{t-4}^2 - 1)\right] = 0$	0,5290	0,7398	1,3396
$E\left[(z_t^2 - 1)(z_{t-5}^2 - 1)\right] = 0$	-0,8055	-1,8189	1,3239

Список литературы

1. Цыплаков А. Некоторые эконометрические методы. Метод максимального правдоподобия в эконометрии. — Новосибирск: ЭФ НГУ, 1997.
2. Bollerslev Tim (1986) «Generalized Autoregressive Conditional Heteroskedasticity», Journal of Econometrics, 31, 307-327.
3. Bollerslev Tim (1987) «A Conditional Heteroskedastic Time Series Model for Speculative Prices and Rates of Return», Review of Economics and Statistics, 69, 542-547.
4. Bollerslev Tim and Jeffrey M. Wooldridge (1992), «Quasi Maximum Likelihood Estimation and Inference in Dynamic Models with Time Varying Covariances», Econometric Reviews, 11, 143-172.
5. Bollerslev Tim, Engle Robert F. and Nelson Daniel B. (1993). ARCH Models / University of California, San Diego.
6. Bollerslev Tim, Ray Y. Chou and Kenneth F. Kroner (1992) «ARCH modelling in finance: A Review of the Theory and Empirical Evidence», Journal of Econometrics, 52, 5-59.
7. Davidson Russell, MacKinnon James G. Estimation and Inference in Econometrics./New York Oxford - Oxford University Press, 1993.
8. Domowitz Ian, Hakkio Craige S. (1985) «Conditional Variance and the Risk Premium in the Foreign Exchange Market», Journal of International Economics, 19, 47-66.
9. Engle Robert F. (1982) «Autoregressive Conditional Heteroskedasticity with the Estimates of the Variance of U.K. Inflation», Econometrica, 50, 987-1008.
10. Engle Robert F., David M. Lilien and Russel P. Robins (1987) «Estimating Time Varying Risk Premia in the Term Structure: The ARCH-M model», Econometrica, 55, 391-407.
11. Nelson Daniel B. (1990), «Stationarity and Persistance in the GARCH(1,1) Model», Econometric Theory, 6, 318-334.
12. Nelson Daniel B. (1991), «Conditional Heteroskedasticity in Asset Returns: A New Approach», Econometrica, 59, 347-370.