

Эконометрический ликбез: волатильность

Одномерные GARCH-модели: обзор*

Эдуардо Росси[†]

Университет Павии, Павия, Италия

В эссе представлен обзор одномерных GARCH-моделей. Рассматриваются модели ARCH, GARCH, EGARCH и другие возможные нелинейные расширения. Приведены условия стационарности (слабой и сильной). Инференция и тестирование гипотез рассматриваются в рамках оценивания методом максимального квазиправдоподобия. Обсуждаются непрерывные приближения к GARCH.

1 Введение

Изучение волатильности доходности активов внесло важный вклад в понимание современных финансовых рынков. Волатильность считается мерой риска, а рисковость любого финансового актива – решающая характеристика, определяющая его равновесную цену. Модели авторегрессионной условной гетероскедастичности (ARCH) были разработаны для учета эмпирических закономерностей в финансовых данных. Многим финансовым временным рядам свойственны следующие стилизованные факты:

- Цены активов, вообще говоря, нестационарны. Доходности обычно стационарны. Некоторые финансовые временные ряды являются дробно интегрированными.
- Автокорреляция в рядах доходностей обычно слаба или отсутствует.
- Серийная независимость между квадратами значений ряда часто отвергается в пользу наличия нелинейной зависимости между последовательными наблюдениями.
- Волатильность рядов доходностей кластеризована.
- Нормальность отвергается в пользу некоторого распределения с толстыми хвостами.
- В некоторых рядах присутствует так называемый *эффект рычага*, то есть изменения цен акций отрицательно коррелируют с изменениями волатильности. Когда рыночная стоимость фирмы падает, соотношение заемного и собственного капитала обычно растет. Это повышает волатильность акционерного капитала, если доходность постоянна. Однако Блэк утверждает, что реакция волатильности на направление изменения доходности слишком велика, чтобы полностью объясняться эффектом рычага.
- Волатильности различных ценных бумаг очень часто движутся вместе.

Модель ARCH и ее обобщения применяются для моделирования волатильности доходностей большого числа финансовых активов, таких как ценные бумаги с фиксированным доходом, обменные курсы, индивидуальные акции и фондовые индексы. В классе ARCH-моделей с дискретным временем ожидания формулируются в терминах напрямую наблюдаемых величин,

*Перевод Б. Гершмана. Цитировать как: Росси, Эдуардо (2010). «Одномерные GARCH-модели: обзор», Квантиль, №8, стр. 1–67. Citation: Rossi, Eduardo (2010). “Univariate GARCH Models: A Survey,” *Quantile*, No.8, pp. 1–67.

[†]Адрес: Dipartimento economia politica e metodi quantitativi, via San Felice 5, 27100 Pavia, Italy. Электронная почта: erossi@eco.unipv.it

в то время как модели стохастической волатильности в дискретном или непрерывном времени включают латентные переменные состояния. Предложено, оценено и изучено огромное число моделей. Продолжают появляться новые модификации. Однако Энгл (Engle, 2002) выделяет как наиболее влиятельные модель GARCH, предложенную Боллерслевом (Bollerslev, 1986), и модель EGARCH Нельсона (Nelson, 1991). Имеется множество обзоров разрастающейся литературы об ARCH-моделях, например Andersen & Bollerslev (1998), Andersen et al. (2006), Bauwens et al. (2006), Bera & Higgins (1993), Bollerslev et al. (1992), Bollerslev et al. (1994), Degiannakis & Xekalaki (2004), Diebold (2004), Diebold & Lopez (1995), Engle (2001, 2004), Engle & Patton (2001), Pagan (1996), Palm (1996), Shephard (1996) и Teräsvirta (2009).

Эти обзоры и главы в учебниках содержат все увеличивающийся список акронимов и аббревиатур, используемых для обозначения всего изобилия предложенных моделей и методов. Например, Bollerslev (2009) в качестве приложения к традиционным обзорам приводит справочник с длинным списком связанных с ARCH-моделями сокращений.

Первые исследования в области эконометрического моделирования волатильности были исключительно параметрическими, но в последние годы происходит сдвиг в сторону менее параметрических и даже полностью непараметрических методов. Непараметрические подходы к моделированию волатильности, которые, как правило, не делают предположений о функциональных формах, позволяют получать гибкие и в то же время состоятельные (по мере увеличения частоты выборки для ряда доходностей) оценки реализованной волатильности, см. обзор в Andersen et al. (2009).

2 Для чего нужны ARCH-модели?

Теорема Вольда о разложении устанавливает, что любой ковариационно стационарный ряд $\{y_t\}$ можно записать в виде суммы линейно детерминированной компоненты и линейно стохастической, представимой в виде квадратично суммируемого одностороннего скользящего среднего. Таким образом,

$$y_t = d_t + u_t,$$

где d_t – линейно детерминированная компонента, а u_t – линейно регулярный ковариационно стационарный случайный процесс:

$$u_t = B(L) \varepsilon_t,$$

$$B(L) = \sum_{i=0}^{\infty} b_i L^i, \quad \sum_{i=0}^{\infty} b_i^2 < \infty, \quad b_0 = 1,$$

где L – лаговый оператор, то есть $x_{t-1} = Lx_t$, и

$$\begin{aligned} \mathbb{E}[\varepsilon_t] &= 0, \\ \mathbb{E}[\varepsilon_t \varepsilon_\tau] &= \begin{cases} \sigma_\varepsilon^2 < \infty, & t = \tau, \\ 0, & \text{иначе.} \end{cases} \end{aligned}$$

Некоррелированная последовательность инноваций необязательно гауссова, а следовательно, необязательно независимая. Зависимые инновации свойственны нелинейным временным рядам вообще и условно гетероскедастичным временным рядам в частности.

Предположим, что y_t – линейный ковариационно стационарный процесс с IID инновациями, а не просто белый шум. Безусловные среднее и дисперсия равны

$$\begin{aligned} \mathbb{E}[y_t] &= 0, \\ \mathbb{E}[y_t^2] &= \sigma_\varepsilon^2 \sum_{i=0}^{\infty} b_i^2, \end{aligned}$$

и оба неизменны во времени. Условное среднее меняется во времени и имеет вид

$$\mathbb{E}[y_t | \Phi_{t-1}] = \sum_{i=1}^{\infty} b_i \varepsilon_{t-i},$$

где $\Phi_{t-1} = \{\varepsilon_{t-1}, \varepsilon_{t-2}, \dots\}$ – информационное множество. Эта модель неспособна уловить динамику условной дисперсии. Действительно, условная дисперсия y_t постоянна:

$$\mathbb{E}[(y_t - \mathbb{E}[y_t | \Phi_{t-1}])^2 | \Phi_{t-1}] = \sigma_\varepsilon^2.$$

Это ограничение проявляется в свойствах условной дисперсии ошибки прогноза на k шагов вперед. Условный прогноз на k шагов вперед имеет вид

$$\mathbb{E}[y_{t+k} | \Phi_t] = \sum_{i=0}^{\infty} b_{k+i} \varepsilon_{t-i},$$

а соответствующая ошибка прогноза равна

$$y_{t+k} - \mathbb{E}[y_{t+k} | \Phi_t] = \sum_{i=0}^{k-1} b_i \varepsilon_{t+k-i}$$

с условной дисперсией

$$\mathbb{E}[(y_{t+k} - \mathbb{E}[y_{t+k} | \Phi_t])^2 | \Phi_t] = \sigma_\varepsilon^2 \sum_{i=0}^{k-1} b_i^2.$$

При $k \rightarrow \infty$ условная дисперсия ошибки прогноза сходится к безусловной дисперсии $\sigma_\varepsilon^2 \sum_{i=0}^{\infty} b_i^2$. Для любого k условная дисперсия ошибки прогноза зависит только от k , но не от Φ_t . Таким образом, простая модель с *IID* инновациями неспособна учесть важную информацию, доступную в момент времени t .

3 Модель ARCH: общее определение

Обозначим за $\{\varepsilon_t(\theta)\}$ случайный процесс в дискретном времени с условным средним и дисперсией, параметризованными конечномерным вектором $\theta \in \Theta \subseteq \mathfrak{R}^m$, где θ_0 – истинное значение вектора параметров. Предположим для начала, что $\varepsilon_t(\theta_0)$ является скалярным.

Пусть $\mathbb{E}_{t-1}[\cdot]$ обозначает условное матожидание, когда информационное множество (обозначенное как Φ_{t-1}) состоит из прошлых значений процесса и другой информации, доступной в момент времени $t-1$:

$$\mathbb{E}_{t-1}[\cdot] \equiv \mathbb{E}[\cdot | \Phi_{t-1}].$$

Аналогично определяется условная дисперсия:

$$\mathbb{V}_{t-1}[\cdot] \equiv \mathbb{V}[\cdot | \Phi_{t-1}].$$

Определение 1 (Bollerslev et al., 1994). Процесс $\{\varepsilon_t(\theta_0)\}$ следует модели авторегрессионной условной гетероскедастичности (ARCH), если

$$\mathbb{E}_{t-1}[\varepsilon_t(\theta_0)] = 0, \quad t = 1, 2, \dots, \quad (1)$$

и условная дисперсия

$$\sigma_t^2(\theta_0) \equiv \mathbb{V}_{t-1}[\varepsilon_t(\theta_0)] = \mathbb{E}_{t-1}[\varepsilon_t^2(\theta_0)], \quad t = 1, 2, \dots, \quad (2)$$

нетривиально зависит от σ -поля, порожденного прошлыми наблюдениями $\{\varepsilon_{t-1}(\theta_0), \varepsilon_{t-2}(\theta_0), \dots\}$.

Пусть $\{y_t(\theta_0)\}$ – исследуемый случайный процесс с условным средним

$$\mu_t(\theta_0) \equiv \mathbb{E}_{t-1}(y_t), \quad t = 1, 2, \dots \quad (3)$$

Как $\mu_t(\theta_0)$, так и $\sigma_t^2(\theta_0)$ измеримы относительно информационного множества на момент времени $t - 1$.¹ Обозначим за $\{\varepsilon_t(\theta_0)\}$ процесс

$$\varepsilon_t(\theta_0) \equiv y_t - \mu_t(\theta_0). \quad (4)$$

Из уравнений (1) и (2) следует, что стандартизированный процесс

$$z_t(\theta_0) \equiv \varepsilon_t(\theta_0) \sigma_t^2(\theta_0)^{-1/2}, \quad t = 1, 2, \dots, \quad (5)$$

имеет нулевое условное среднее ($\mathbb{E}_{t-1}[z_t(\theta_0)] = 0$) и неизменную во времени единичную условную дисперсию. Можно считать, что $\varepsilon_t(\theta_0)$ порождается процессом

$$\varepsilon_t(\theta_0) = z_t(\theta_0) \sigma_t^2(\theta_0)^{1/2},$$

где $\varepsilon_t^2(\theta_0)$ – несмещенная оценка $\sigma_t^2(\theta_0)$. Предположим, что $z_t(\theta_0) \sim NID(0, 1)$ и не зависит от $\sigma_t^2(\theta_0)$. Тогда

$$\mathbb{E}_{t-1}[\varepsilon_t^2] = \mathbb{E}_{t-1}[\sigma_t^2] \mathbb{E}_{t-1}[z_t^2] = \mathbb{E}_{t-1}[\sigma_t^2],$$

поскольку $z_t^2 | \Phi_{t-1} \sim \chi_{(1)}^2$. Медиана $\chi_{(1)}^2$ равна 0.455, так что $\mathbb{P}[\varepsilon_t^2 < \frac{1}{2} \sigma_t^2] > \frac{1}{2}$. Прокси ε_t^2 создает потенциально значительную ошибку при анализе малых выборок σ_t^2 , $t = 1, 2, \dots, T$, хотя ошибка уменьшается по мере роста T .

Если условное распределение z_t неизменно во времени и имеет конечный четвертый момент, то

$$\mathbb{E}[\varepsilon_t^4] = \mathbb{E}[z_t^4] \mathbb{E}[\sigma_t^4] \geq \mathbb{E}[z_t^4] (\mathbb{E}[\sigma_t^2])^2 = \mathbb{E}[z_t^4] (\mathbb{E}[\varepsilon_t^2])^2,$$

то есть $\mathbb{E}[\varepsilon_t^4] \geq \mathbb{E}[z_t^4] (\mathbb{E}[\varepsilon_t^2])^2$, согласно неравенству Йенсена². Строгое равенство имеет место только в случае постоянной условной дисперсии. Если $z_t \sim NID(0, 1)$, то $\mathbb{E}[z_t^4] = 3$, а следовательно, безусловное распределение ε_t имеет положительный эксцесс:

$$\mathbb{E}[\varepsilon_t^4] \geq 3 (\mathbb{E}[\varepsilon_t^2])^2, \quad \mathbb{E}[\varepsilon_t^4] / (\mathbb{E}[\varepsilon_t^2])^2 \geq 3.$$

Более того, куртозис можно выразить как функцию от изменчивости условной дисперсии. Если $\varepsilon_t | \Phi_{t-1} \sim N(0, \sigma_t^2)$, то $\mathbb{E}_{t-1}[\varepsilon_t^4] = 3(\mathbb{E}_{t-1}[\varepsilon_t^2])^2$ и

$$\mathbb{E}[\varepsilon_t^4] = 3\mathbb{E}[(\mathbb{E}_{t-1}[\varepsilon_t^2])^2] \geq 3(\mathbb{E}[\mathbb{E}_{t-1}[\varepsilon_t^2]])^2 = 3(\mathbb{E}[\varepsilon_t^2])^2.$$

Далее

$$\mathbb{E}[\varepsilon_t^4] = 3(\mathbb{E}[\varepsilon_t^2])^2 + 3\mathbb{E}[(\mathbb{E}_{t-1}[\varepsilon_t^2])^2] - 3(\mathbb{E}[\mathbb{E}_{t-1}[\varepsilon_t^2]])^2$$

¹Andersen (1996) различает детерминированный, условно гетероскедастичный, условно стохастический и современно стохастический процессы для волатильности. Процесс волатильности является детерминированным, если информационное множество (σ -поле), в данном эссе обозначаемое как Φ , совпадает с σ -полем всех случайных векторов в системе до момента $t = 0$ включительно, условно гетероскедастичным, если Φ содержит информацию, доступную и наблюдаемую в момент времени $t - 1$, условно стохастическим, если Φ содержит все случайные вектора до момента $t - 1$, и современно стохастическим, если информационное множество Φ содержит случайные векторы вплоть до момента времени t .

²Неравенство Йенсена. Пусть X и $g(X)$ – интегрируемые случайные величины. Тогда $\mathbb{E}[g(X)] \leq g(\mathbb{E}[X])$, если $g(\cdot)$ – вогнутая функция, и $\mathbb{E}[g(X)] \geq g(\mathbb{E}[X])$, если $g(\cdot)$ выпуклая.

и

$$k = \frac{\mathbb{E}[\varepsilon_t^4]}{(\mathbb{E}[\varepsilon_t^2])^2} = 3 + 3 \frac{\mathbb{E}\left[\left(\mathbb{E}_{t-1}[\varepsilon_t^2]\right)^2\right] - \left(\mathbb{E}\left[\mathbb{E}_{t-1}[\varepsilon_t^2]\right]\right)^2}{(\mathbb{E}[\varepsilon_t^2])^2} = 3 + 3 \frac{\mathbb{V}[\sigma_t^2]}{(\mathbb{E}[\varepsilon_t^2])^2}.$$

Другое важное свойств ARCH-процесса состоит в отсутствии условной серийной корреляции. Учитывая, что $\mathbb{E}_{t-1}[\varepsilon_t] = 0$, по закону повторных матожиданий, имеем

$$\mathbb{E}_{t-h}[\varepsilon_t] = \mathbb{E}_{t-h}[\mathbb{E}_{t-1}(\varepsilon_t)] = \mathbb{E}_{t-h}[0] = 0.$$

Из этого свойства ортогональности следует, что процесс $\{\varepsilon_t\}$ условно некоррелирован:

$$\begin{aligned} C_{t-h}[\varepsilon_t, \varepsilon_{t+k}] &= \mathbb{E}_{t-h}[\varepsilon_t \varepsilon_{t+k}] - \mathbb{E}_{t-h}[\varepsilon_t] \mathbb{E}_{t-h}[\varepsilon_{t+k}] \\ &= \mathbb{E}_{t-h}[\varepsilon_t \varepsilon_{t+k}] = \mathbb{E}_{t-h}[\mathbb{E}_{t+k-1}[\varepsilon_t \varepsilon_{t+k}]] \\ &= \mathbb{E}[\varepsilon_t \mathbb{E}_{t+k-1}[\varepsilon_{t+k}]] = 0. \end{aligned}$$

3.1 Модель ARCH(q)

Модель ARCH(q), предложенная в Engle (1982), представляет условную дисперсию как линейную функцию квадратов прошлых возмущений:

$$\sigma_t^2 = \omega + \sum_{i=1}^q \alpha_i \varepsilon_{t-i}^2. \quad (6)$$

В этой модели для обеспечения положительности условной дисперсии параметры должны удовлетворять следующим условиям: $\omega > 0$, $\alpha_1 \geq 0$, $\alpha_2 \geq 0$, \dots , $\alpha_q \geq 0$. Определив

$$\sigma_t^2 \equiv \varepsilon_t^2 - v_t,$$

где $\mathbb{E}_{t-1}[v_t] = 0$, можно записать модель (6) в виде AR(q)-процесса для ε_t^2 :

$$\varepsilon_t^2 = \omega + \alpha(L) \varepsilon_t^2 + v_t,$$

где $\alpha(L) = \alpha_1 L + \alpha_2 L^2 + \dots + \alpha_q L^q$. Этот процесс является слабо стационарным тогда и только тогда, когда $\sum_{i=1}^q \alpha_i < 1$; в этом случае безусловная дисперсия равна

$$\mathbb{E}[\varepsilon_t^2] = \omega / (1 - \alpha_1 - \dots - \alpha_q). \quad (7)$$

Процесс характеризуется положительным эксцессом. Например, в случае модели ARCH(1) с $\varepsilon_t | \Phi_{t-1} \sim N(0, \sigma_t^2)$, куртозис равен

$$\mathbb{E}[\varepsilon_t^4] / (\mathbb{E}[\varepsilon_t^2])^2 = 3(1 - \alpha_1^2) / (1 - 3\alpha_1^2), \quad (8)$$

если $3\alpha_1^2 < 1$, а при $3\alpha_1^2 = 1$

$$\mathbb{E}[\varepsilon_t^4] / (\mathbb{E}[\varepsilon_t^2])^2 = \infty.$$

В обоих случаях куртозис превышает значение 3, характерное для нормального распределения.

3.2 ARCH как нелинейная модель

Важнейшее свойство ARCH-модели состоит в том, что $\mathbb{C}[\varepsilon_t^2, \varepsilon_{t-j}^2] \neq 0$, хотя $\mathbb{C}[\varepsilon_t, \varepsilon_{t-j}] = 0$ для $j \neq 0$. Рассмотрим связь ARCH-модели с билинейной моделью. Временной ряд $\{\varepsilon_t\}$ следует билинейной модели, если

$$\varepsilon_t = \sum_{i=1}^p \phi_i \varepsilon_{t-i} + \sum_{j=1}^r \sum_{k=1}^s b_{jk} \varepsilon_{t-j} u_{t-k} + u_t,$$

где u_t – последовательность $IID(0, \sigma_u^2)$ величин. Первые два условных момента имеют вид

$$\begin{aligned} \mathbb{E}_{t-1}[\varepsilon_t] &= \sum_{i=1}^p \phi_i \varepsilon_{t-i} + \sum_{j=1}^r \sum_{k=1}^s b_{jk} \varepsilon_{t-j} u_{t-k} + u_t, \\ \mathbb{V}_{t-1}[\varepsilon_t] &= \sigma_u^2. \end{aligned}$$

В отличие от ARCH-модели, в которой условная дисперсия меняется во времени, в билинейной модели условная дисперсия постоянна. Тем не менее, безусловные моменты в обеих моделях могут быть похожи. Например, билинейной модели

$$\varepsilon_t = b_{21} \varepsilon_{t-2} u_{t-1} + u_t, \quad \mathbb{E}[\varepsilon_t] = 0, \quad \mathbb{C}[\varepsilon_t^2, \varepsilon_{t-1}^2] = b_{21}^2 \sigma_u^2,$$

автокоррелированной для квадратов наблюдений, свойственна кластеризация больших и малых отклонений, как и ARCH-процессу.

4 Модель GARCH(p, q)

Для лаконичного моделирования условной гетероскедастичности Bollerslev (1986) и независимо от него Taylor (1986) предложили обобщенную ARCH-модель – GARCH(p, q):

$$\sigma_t^2 = \omega + \alpha(L) \varepsilon_t^2 + \beta(L) \sigma_t^2, \quad (9)$$

где $\alpha(L) = \alpha_1 L + \dots + \alpha_q L^q$, $\beta(L) = \beta_1 L + \dots + \beta_p L^p$. Модель GARCH(1, 1) наиболее популярна в прикладных исследованиях³:

$$\sigma_t^2 = \omega + \alpha_1 \varepsilon_{t-1}^2 + \beta_1 \sigma_{t-1}^2. \quad (10)$$

Для корректного определения условной дисперсии в модели GARCH(p, q) все коэффициенты в соответствующей линейной ARCH-модели бесконечного порядка должны быть положительными. Записывая GARCH(p, q)-модель как ARCH(∞), получаем:

$$\sigma_t^2 = \left(1 - \sum_{i=1}^p \beta_i L^i\right)^{-1} \left(\omega + \sum_{j=1}^q \alpha_j \varepsilon_{t-j}^2\right) = \omega^* + \sum_{k=0}^{\infty} \phi_k \varepsilon_{t-k-1}^2. \quad (11)$$

Таким образом, $\sigma_t^2 \geq 0$, если $\omega^* \geq 0$ и все $\phi_k \geq 0$. Неотрицательность ω^* и ϕ_k является также и необходимым условием неотрицательности σ_t^2 . Чтобы ω^* и $\{\phi_k\}_{k=0}^{\infty}$ были корректно определены, предположим, что:

- i. Корни уравнения $\beta(x) = 1$ лежат вне единичного круга, и $\omega \geq 0$. Это условие гарантирует, что величина ω^* конечна и положительна.

³Модель GARCH принадлежит к классу моделей детерминированной условной гетероскедастичности, в которых условная дисперсия является функцией от переменных, находящихся в информационном множестве на момент времени t .

ii. Полиномы $\alpha(x)$ и $1 - \beta(x)$ не имеют общих корней.

Эти условия не гарантируют ни то, что $\sigma_t^2 \leq \infty$, ни строгую стационарность ряда $\{\sigma_t^2\}_{t=-\infty}^{\infty}$. В простой GARCH(1, 1)-модели для положительности σ_t^2 почти наверное требуется (Nelson & Cao, 1992), наряду с условиями (i) и (ii), чтобы

$$\begin{aligned}\omega &\geq 0, \\ \beta_1 &\geq 0, \\ \alpha_1 &\geq 0.\end{aligned}\tag{12}$$

В моделях GARCH(1, q) и GARCH(2, q) эти условия можно ослабить. Например, в модели GARCH(1, 2) необходимыми и достаточными условиями являются:

$$\begin{aligned}\omega &\geq 0, \\ 1 > \beta_1 &\geq 0, \\ \beta_1\alpha_1 + \alpha_2 &\geq 0, \\ \alpha_1 &\geq 0.\end{aligned}\tag{13}$$

Для модели GARCH(2, 1) условия имеют вид:

$$\begin{aligned}\omega &\geq 0, \\ \alpha_1 &\geq 0, \\ \beta_1 &\geq 0, \\ \beta_1 + \beta_2 &< 1, \\ \beta_1^2 + 4\beta_2 &\geq 0.\end{aligned}\tag{14}$$

Эти условия не такие строгие, как предложенные в Bollerslev (1986):

$$\begin{aligned}\omega &\geq 0, \\ \beta_i &\geq 0, \quad i = 1, \dots, p, \\ \alpha_j &\geq 0, \quad j = 1, \dots, q.\end{aligned}\tag{15}$$

Модель GARCH(2, 2) изучена в He & Teräsvirta (1999). Эти результаты нельзя перенести на многомерный случай, в котором условие положительности $\{\sigma_t^2\}$ означает положительную определенность условной ковариационной матрицы. Для идентификации GARCH-модели, в которой по крайней мере один из коэффициентов $\beta_j > 0$, необходимо также потребовать, чтобы как минимум один из коэффициентов $\alpha_j > 0$. Если $\alpha_1 = \dots = \alpha_q = 0$, условная и безусловная дисперсии ε_t равны, и β_1, \dots, β_p – неидентифицируемые мешающие параметры. С точки зрения оценивания модели GARCH(p, q) методом максимального правдоподобия необходимо рекурсивно подсчитать $\{\sigma_t^2\}_{t=0}^{\infty}$, начиная с момента 0 и применяя (9), предполагая произвольные значения $\{\sigma_{-1}^2, \dots, \sigma_{-p}^2, \varepsilon_{-1}^2, \dots, \varepsilon_{-q}^2\}$ для периодов, предшествующих выборке. Условия (15) гарантируют, что ряд $\{\sigma_t^2\}_{t=0}^{\infty}$ неотрицательный при любых неотрицательных значениях $\{\sigma_{-1}^2, \dots, \sigma_{-p}^2, \varepsilon_{-1}^2, \dots, \varepsilon_{-q}^2\}$. С другой стороны, условия, гарантирующие, что $\omega^* \geq 0$ и $\phi_k \geq 0$ (см. (13) для модели GARCH(1, 2) и (14) для модели GARCH(2, 1)), этого обеспечить не могут. Проблему можно решить, выбрав начальные значения, которые гарантируют неотрицательность $\{\sigma_t^2\}_{t=0}^{\infty}$ с вероятностью 1 при неотрицательных ω^* и $\{\phi_k\}_{k=0}^{\infty}$. Nelson & Cao (1992) предлагают случайно выбрать $\varepsilon_t^2 \geq 0$ и положить $\varepsilon_t^2 = \varepsilon^2$ для t от -1 до $-\infty$ и $\sigma_t^2 = \sigma^2$ для $1 - p \leq t \leq 0$, где

$$\sigma^2 = \left(1 - \sum_{i=1}^p \beta_i\right)^{-1} \left[\omega + \varepsilon^2 \sum_{j=1}^q \alpha_j\right] = \omega^* + \varepsilon^2 \sum_{k=0}^{\infty} \phi_k.$$

Таким образом получаем последовательность $\{\sigma_t^2\} \geq 0$ для всех $t \geq 0$ с вероятностью 1, поскольку

$$\sigma_t^2 = \omega^* + \sum_{k=0}^{t-1} \phi_k \varepsilon_{t-k-1}^2 + \sum_{k=t}^{\infty} \phi_k \varepsilon^2.$$

Предполагая, что $\sum_{i=1}^p \beta_i + \sum_{j=1}^q \alpha_j < 1$, можно положить σ^2 и ε^2 равными их общему безусловному среднему:

$$\sigma^2 \equiv \varepsilon^2 \equiv \left(1 - \sum_{i=1}^p \beta_i - \sum_{j=1}^q \alpha_j \right)^{-1} \omega.$$

В стационарной GARCH-модели Engle & Mezrich (1996) предлагают *таргетирование дисперсии*, а именно, заменяют константу ω в (9) на $(1 - \sum_j \alpha_j - \sum_i \beta_i) \sigma^2$. Оценка σ^2 , то есть $\hat{\sigma}^2 = \sum_{t=1}^T \varepsilon_t^2$, подставляется вместо σ^2 перед оцениванием других параметров. Модель содержит на один параметр меньше, чем стандартная модель GARCH(p, q).

4.1 Уравнения Юла-Уолкера для процесса квадратов

Процесс $\{\varepsilon_t^2\}$ можно представить в виде ARMA(m, p):

$$\varepsilon_t^2 = \omega + \sum_{j=1}^m (\alpha_j + \beta_j) \varepsilon_{t-j}^2 + \left(v_t - \sum_{i=1}^p \beta_i v_{t-i} \right),$$

где $m = \max(p, q)$, $\mathbb{E}_{t-1}[v_t] = 0$, $v_t \in [-\sigma_t^2, \infty)$. Таким образом, применимы классические результаты для ARMA-моделей. В частности, рассмотрим автоковариационную функцию (см. Bollerslev, 1988), то есть

$$\begin{aligned} \gamma^2(k) &= \mathbb{C}[\varepsilon_t^2, \varepsilon_{t-k}^2], \\ \gamma^2(k) &= \mathbb{C} \left[\omega + \sum_{j=1}^m (\alpha_j + \beta_j) \varepsilon_{t-j}^2 + \left(v_t - \sum_{i=1}^p \beta_i v_{t-i} \right), \varepsilon_{t-k}^2 \right], \\ \gamma^2(k) &= \sum_{j=1}^m (\alpha_j + \beta_j) \mathbb{C}[\varepsilon_{t-j}^2, \varepsilon_{t-k}^2] + \mathbb{C} \left[v_t - \sum_{i=1}^p \beta_i v_{t-i}, \varepsilon_{t-k}^2 \right]. \end{aligned} \quad (16)$$

Когда k достаточно велико, последний член в правой части выражения (16) равен нулю. Последовательность автоковариаций удовлетворяет линейному разностному уравнению порядка $\max(p, q)$ для $k \geq p + 1$:

$$\gamma^2(k) = \sum_{j=1}^m (\alpha_j + \beta_j) \gamma^2(k - j).$$

Эту систему можно использовать для идентификации лаговых порядков m и p , то есть p и q , если $q \geq p$, и p при $q < p$. Автокорреляционная функция ε_t^2 , если таковая существует, убывает медленно, хотя и экспоненциально. В ARCH-модели скорость убывания слишком высокая по сравнению с той, которая обычно наблюдается в финансовых временных рядах, если только максимальный лаг q не велик.

4.2 Стационарность

Процесс $\{\varepsilon_t\}$, следующий модели GARCH(p, q), является последовательностью мартингаловых приращений. Для доказательства ковариационной стационарности достаточно показать, что дисперсия

$$\mathbb{V}[\varepsilon_t] = \mathbb{V}[\mathbb{E}_{t-1}[\varepsilon_t]] + \mathbb{E}[\mathbb{V}_{t-1}[\varepsilon_t]] = \mathbb{E}[\sigma_t^2]$$

асимптотически постоянна во времени.

Утверждение. Процесс $\{\varepsilon_t\}$, следующий модели GARCH(p, q) с положительными коэффициентами $\omega \geq 0$, $\alpha_i \geq 0$, $i = 1, \dots, q$, $\beta_i \geq 0$, $i = 1, \dots, p$, является ковариационно стационарным тогда и только тогда, когда

$$\alpha(1) + \beta(1) < 1. \quad (17)$$

Это условие является достаточным, но не необходимым для строгой стационарности. Каждый слабо стационарный GARCH-процесс является также строго стационарным. Поскольку ARCH-процессы имеют толстые хвосты, условия ковариационной стационарности часто более жесткие, чем условия строгой стационарности. Если $\alpha(1) + \beta(1) < 1$, слабо стационарное решение единственно и совпадает с единственным стационарным решением.

Для поиска строго стационарного решения рассмотрим $\varepsilon_t^2 = z_t^2 \sigma_t^2$. Модель GARCH(1, 1) можно записать в виде

$$\sigma_t^2 = \omega + \sigma_{t-1}^2 (\alpha_1 z_{t-1}^2 + \beta_1). \quad (18)$$

Рекурсивная подстановка приводит к записи GARCH(1, 1)-модели в виде

$$\sigma_t^2 = \omega \left(1 + \sum_{k=1}^{\infty} \prod_{i=1}^k (\beta_1 + \alpha_1 z_{t-i}^2) \right). \quad (19)$$

Обозначая

$$\begin{aligned} A_t &= \beta_1 + \alpha_1 z_t^2, \\ B_t &= \omega, \\ Y_t &= \sigma_{t+1}^2, \end{aligned}$$

получаем, что $Y_t = \sigma_{t+1}^2$ является решением стохастического рекуррентного уравнения $Y_t = A_t Y_{t-1} + B_t$, где $\{A_t, B_t\} - IID$. Каждое строго стационарное решение $\{\sigma_t^2\}$ уравнения (18) можно выразить в виде функции процесса $\{z_t\}$, так что из стационарности $\{\sigma_t^2\}$ следует стационарность $\{\sigma_t^2, z_t\}$, а значит, и $\{\varepsilon_t, \sigma_t\}$. Таким образом, существование строго стационарных решений для GARCH(1, 1)-процесса сводится к изучению строго стационарных решений уравнения (18) (см. Lindner, 2009b). Решение $\{Y_t\}$ стохастического рекуррентного уравнения (с IID коэффициентами) – это последовательность случайных величин. Каждое решение удовлетворяет

$$\begin{aligned} Y_t &= A_t Y_{t-1} + B_t \\ &= A_t A_{t-1} Y_{t-2} + A_t B_{t-1} + B_t \\ &= \dots \\ &= \left(\prod_{j=0}^k A_{t-j} \right) Y_{t-k-1} + \sum_{i=0}^k \left(\prod_{j=0}^{i-1} A_{t-j} \right) B_{t-i} \end{aligned}$$

для всех $k \in \mathbb{N} \cup \{0\}$, где $\prod_{j=0}^{-1} A_{t-j} = 1$ для произведения по пустому множеству индексов. Для существования стационарного решения необходимо, чтобы

$$\lim_{k \rightarrow \infty} \left(\prod_{j=0}^k A_{t-j} \right) Y_{t-k-1} = 0 \quad \text{п.н.}$$

и сумма $\sum_{i=0}^k \left(\prod_{j=0}^{i-1} A_{t-j} \right) B_{t-i}$ сходилась почти наверное при $k \rightarrow \infty$. В случаях GARCH(1, 1) и ARCH(1)

$$\sigma_{t+1}^2 = Y_t = \left(\prod_{i=0}^k A_{t-i} \right) \sigma_{t-k}^2 + \omega \sum_{i=0}^k \left(\prod_{j=0}^{i-1} A_{t-j} \right). \quad (20)$$

Поскольку это сумма неотрицательных компонент, ряд $\sum_{i=0}^{\infty} \prod_{j=0}^{i-1} A_{t-j}$ сходится почти наверное для каждого t , и следовательно, $\prod_{i=0}^k A_{t-i}$ сходится почти наверное к 0 при $k \rightarrow \infty$. Если процесс $\{\sigma_{t+1}^2\}$ строго стационарный, то $\left(\prod_{i=0}^k A_{t-i} \right) \sigma_{t-k}^2$ сходится к 0 по распределению и по вероятности при $k \rightarrow \infty$. Следовательно, существует не более одного строго стационарного решения $\sigma_t^2 = Y_{t-1}$:

$$Y_t = \omega \sum_{i=0}^{\infty} \left(\prod_{j=0}^{i-1} A_{t-j} \right). \quad (21)$$

Таким образом, из существования строго стационарного процесса GARCH(1, 1) или ARCH(1) следует сходимости к 0 почти наверное произведения $\prod_{i=0}^k A_{t-i}$ при $k \rightarrow \infty$. Но верно также и обратное утверждение, а значит, строго стационарное решение GARCH(1, 1)/ARCH(1)-процесса существует тогда и только тогда, когда $\prod_{i=0}^k A_{t-i}$ сходится п.н. к 0 при $k \rightarrow \infty$.

Nelson (1990b) показал, что, при $\omega > 0$ $\sigma_t^2 < \infty$ п.н. и $\{\varepsilon_t, \sigma_t^2\}$ строго стационарны тогда и только тогда, когда $\mathbb{E} [\ln (\beta_1 + \alpha_1 z_t^2)] < 0$. Учитывая, что

$$\mathbb{E} [\ln (\beta_1 + \alpha_1 z_t^2)] \leq \ln (\mathbb{E} [\beta_1 + \alpha_1 z_t^2]) = \ln (\alpha_1 + \beta_1),$$

при $\alpha_1 + \beta_1 = 1$ модель строго стационарна. Поэтому $\mathbb{E} [\ln (\beta_1 + \alpha_1 z_t^2)] < 0$ является более мягким требованием, чем $\alpha_1 + \beta_1 < 1$. Для модели ARCH(1) при $\alpha_1 = 1$, $\beta_1 = 0$ и $z_t \sim IID(0, 1)$

$$\mathbb{E} [\ln (z_t^2)] \leq \ln (\mathbb{E} [z_t^2]) = \ln (1).$$

Таким образом, интегрированная ARCH(1)-модель является строго стационарной, но не ковариационно стационарной.

Для анализа условий существования строго стационарных решений модели GARCH(p, q) необходимо расширить стохастическое рекуррентное уравнение на многомерный случай (см. Bougerol & Picard, 1992a). Строгая стационарность многомерных стохастических рекуррентных уравнений исследуется в терминах максимальной экспоненты Ляпунова. Bougerol & Picard (1992b) показали, что неприводимые рекуррентные уравнения с IID коэффициентами $\{A_t, B_t\}$, такими, что $\mathbb{E}[\ln^+ \|A_0\|] < \infty$ и $\mathbb{E}[\ln^+ \|B_0\|] < \infty$, допускают неупреждающее строго стационарное решение тогда и только тогда, когда максимальная экспонента Ляпунова, связанная с $\{A_t\}$, строго отрицательна. Bougerol & Picard (1992a) также показали, что GARCH(p, q)-процесс допускает строго стационарное решение тогда и только тогда, когда максимальная экспонента Ляпунова для последовательности $\{A_t\}$ строго отрицательна. Это решение единственно.

4.3 Существование моментов GARCH-процессов

Для GARCH(1, 1)-процессов с нормальными инновациями необходимое и достаточное условие существования моментов дано в Bollerslev (1986), а в He & Teräsvirta (1999) оно расширено на случай общих шумовых последовательностей. Ling (1999) и Ling & McAleer (2002) приводят необходимое и достаточное условие существования моментов более высокого порядка. Для ARCH(p)-процесса необходимое и достаточное условие существования моментов высокого порядка было получено ранее в Milhøj (1985). Для строго стационарных процессов GARCH(1, 1) и ARCH(1) момент $\mathbb{E}[\sigma_t^{2m}]$ величины σ_t^2 , $m > 0$, существует тогда и только тогда, когда

$$\mathbb{E}[(\beta_1 + \alpha_1 z_0^2)^m] < 1. \quad (22)$$

Если m – положительное целое число, выполняется условие (22), и $\mu_j = \mathbb{E}[\sigma_t^{2j}]$ обозначает j -й момент σ_t^2 , то μ_m можно рекурсивно вычислить по формуле

$$\mu_m = (1 - \mathbb{E}[(\beta_1 + \alpha_1 z_0^2)^m])^{-1} \sum_{j=0}^{m-1} \binom{m}{j} \omega^{m-j} \mathbb{E}[(\beta_1 + \alpha_1 z_0^2)^j] \mu_j. \quad (23)$$

Отсюда следует, что момент порядка $2m$ величины ε_t равен

$$\mathbb{E}[\varepsilon_t^{2m}] = \mu_m \mathbb{E}[z_0^{2m}].$$

Чтобы понять, почему условие (22) необходимо и достаточно для конечности $\mathbb{E}[\sigma_t^{2m}]$, $m \in (0, \infty)$, перепишем модель GARCH(1, 1) в виде (19). Если $\mathbb{E}[(\beta_1 + \alpha_1 z_0^2)^m] < 1$ и $m \in [1, m)$, то, согласно неравенству Минковского,

$$(\mathbb{E}[\sigma_t^{2m}])^{1/m} \leq \omega \sum_{i=0}^{\infty} (\mathbb{E}[\beta_1 + \alpha_1 z_0^{2i}])^{i/m} < \infty, \quad (24)$$

а для $m < 1$ можно похожим образом применить неравенство $\mathbb{E}[(U + V)^m] \leq \mathbb{E}[U^m] + \mathbb{E}[V^m]$, выполняющееся для положительных случайных величин U, V . В обратную сторону, если $\mathbb{E}[\sigma_t^{2m}] < \infty$, то величина $\mathbb{E} \prod_{j=0}^i (\beta_1 + \alpha_1 z_{t-j}^2)^m$ должна сходиться к 0 при $i \rightarrow \infty$, что может выполняться только при условии (22). Если m целое и выполняется условие (22), то (23) следует после возведения (10) в степень m и взятия математического ожидания.

Момент $\mathbb{E}[\sigma_t^{2m}]$ является конечным для целого m тогда и только тогда, когда

$$\sum_{j=0}^m \binom{m}{j} \beta_1^{m-j} \alpha_1^j \mathbb{E}[z_t^{2j}] < 1.$$

Если $z_t \sim IIDN(0, 1)$, получаем

$$\sum_{j=0}^m \binom{m}{j} \beta_1^{m-j} \alpha_1^j \prod_{i=1}^j (2i - 1) < 1.$$

Так, например, четвертый момент σ_t существует тогда и только тогда, когда $\beta_1^2 + 2\alpha_1\beta_1 + 3\alpha_1^2 < 1$.

4.4 Прогнозирование волатильности

Модель GARCH(p, q) можно представить в виде ARMA-процесса, учитывая, что $\varepsilon_t^2 = \sigma_t^2 + v_t$, где $\mathbb{E}_{t-1}[v_t] = 0$, $v_t \in [-\sigma_t^2, \infty)$:

$$\varepsilon_t^2 = \omega + \sum_{j=1}^{\max(p,q)} (\alpha_j + \beta_j) \varepsilon_{t-j}^2 + \left(v_t - \sum_{i=1}^p \beta_i v_{t-i} \right).$$

Таким образом, $\varepsilon_t^2 \sim \text{ARMA}(m, p)$, где $m = \max(p, q)$. Прогнозирование для GARCH(p, q)-модели рассмотрено в Engle & Bollerslev (1986). Запишем

$$\sigma_{t+k}^2 = \omega + \sum_{i=1}^n (\alpha_i \varepsilon_{t+k-i}^2 + \beta_i \sigma_{t+k-i}^2) + \sum_{i=k}^m (\alpha_i \varepsilon_{t+k-i}^2 + \beta_i \sigma_{t+k-i}^2),$$

где $n = \min\{m, k-1\}$ и, по определению, суммы от 1 до 0 и от $k > m$ до m обе равны нулю. Тогда

$$\mathbb{E}_t [\sigma_{t+k}^2] = \omega + \sum_{i=1}^n ((\alpha_i + \beta_i) \mathbb{E}_t [\sigma_{t+k-i}^2]) + \sum_{i=k}^m (\alpha_i \varepsilon_{t+k-i}^2 + \beta_i \sigma_{t+k-i}^2).$$

В частности, для GARCH(1, 1)-модели при $k > 2$ имеем

$$\begin{aligned} \mathbb{E}_t [\sigma_{t+k}^2] &= \sum_{i=0}^{k-2} (\alpha_1 + \beta_1)^i \omega + (\alpha_1 + \beta_1)^{k-1} \sigma_{t+1}^2 \\ &= \omega \frac{(1 - (\alpha_1 + \beta_1)^{k-1})}{(1 - (\alpha_1 + \beta_1))} + (\alpha_1 + \beta_1)^{k-1} \sigma_{t+1}^2 \\ &= \sigma^2 (1 - (\alpha_1 + \beta_1)^{k-1}) + (\alpha_1 + \beta_1)^{k-1} \sigma_{t+1}^2 \\ &= \sigma^2 + (\alpha_1 + \beta_1)^{k-1} (\sigma_{t+1}^2 - \sigma^2). \end{aligned}$$

Когда процесс является ковариационно стационарным, $\mathbb{E}_t [\sigma_{t+k}^2]$ сходится к σ^2 при $k \rightarrow \infty$.

4.5 Модель IGARCH(p, q)

Определение 2. GARCH(p, q)-процесс, имеющий первые два условных момента вида

$$\mathbb{E}_{t-1} [\varepsilon_t] = 0,$$

$$\sigma_t^2 \equiv \mathbb{E}_{t-1} [\varepsilon_t^2] = \omega + \sum_{i=1}^q \alpha_i \varepsilon_{t-i}^2 + \sum_{i=1}^p \beta_i \sigma_{t-i}^2,$$

где $\omega \geq 0$, $\alpha_i \geq 0$ и $\beta_i \geq 0$ для всех i , и для которого уравнение

$$1 - \alpha(x) - \beta(x) = 0$$

имеет $d > 0$ единичных корней, а $\max\{p, q\} - d$ корней лежат вне единичного круга, называют:

- i) интегрированным относительно дисперсии порядка d , если $\omega = 0$.
- ii) интегрированным относительно дисперсии порядка d с трендом, если $\omega > 0$.

Интегрированные GARCH(p, q)-модели (IGARCH), как с трендом, так и без него, являются, таким образом, частью более широкого класса моделей со свойством «устойчивой дисперсии», когда текущая информация остается важной для прогнозирования условных дисперсий при любом горизонте планирования.

Интегрированная GARCH(p, q)-модель имеет место, если выполняется необходимое условие

$$\alpha(1) + \beta(1) = 1.$$

Для примера рассмотрим модель IGARCH(1, 1), в которой $\alpha_1 + \beta_1 = 1$. Процесс можно записать в виде

$$\begin{aligned} \sigma_t^2 &= \omega + \alpha_1 \varepsilon_{t-1}^2 + (1 - \alpha_1) \sigma_{t-1}^2, \\ \sigma_t^2 &= \omega + \sigma_{t-1}^2 + \alpha_1 (\varepsilon_{t-1}^2 - \sigma_{t-1}^2), \quad 0 < \alpha_1 \leq 1. \end{aligned} \tag{25}$$

Для этой конкретной модели прогноз условной дисперсии на k шагов вперед имеет вид

$$\mathbb{E}_t [\sigma_{t+k}^2] = (k-1)\omega + \sigma_{t+1}^2.$$

Nelson (1991) показал, что при инициализации процесса IGARCH в некоторый конечный момент времени его поведение зависит от параметра ω . Если $\omega > 0$, то безусловная дисперсия ε_t линейно растет со временем. Если $\omega = 0$, реализации процесса почти наверное сходятся к нулю. Параметр β_1 влияет на скорость этой сходимости. Важно отметить, что безусловная дисперсия ε_t не существует, когда условная дисперсия следует модели IGARCH. В случае GARCH(1, 1)-модели возможное объяснение того факта, что оценка $\alpha_1 + \beta_1$ близка к единице, состоит в том, что во время периода оценивания GARCH-модели происходит смена константы (Diebold, 1986, и Lamoureux & Lastrapes, 1990). Это означает, что порождающий данные GARCH-процесс нестационарный.

Модель IGARCH(1, 1) в (25) также можно записать в форме ARIMA(0, 1, 1):

$$(1-L)\varepsilon_t^2 = \omega + v_t - \beta_1 v_{t-1}, \quad (26)$$

где $v_t = \varepsilon_t^2 - \sigma_t^2$ является мартингаловым приращением относительно σ_t^2 .

4.6 Устойчивость

При изучении поведения временных рядов волатильности активов часто возникает вопрос о том, как долго сохраняются шоки условной дисперсии. Если шоки волатильности сохраняются навсегда, они могут сдвинуть всю временную структуру рискованных премий. В теории вероятностей существует много типов сходимости (почти наверное, по вероятности, в L^p), так что является ли шок временным или устойчивым может зависеть от определения понятия сходимости. В линейных моделях обычно не имеет значения, какое из стандартных определений использовать, поскольку результаты в основном согласуются. В GARCH-моделях ситуация более запутанная. В случае IGARCH(1, 1)

$$\sigma_t^2 = \omega + \alpha_1 \varepsilon_{t-1}^2 + \beta_1 \sigma_{t-1}^2,$$

где $\alpha_1 + \beta_1 = 1$. Учитывая, что $\varepsilon_t^2 = z_t^2 \sigma_t^2$, перепишем IGARCH(1, 1)-процесс в виде

$$\sigma_t^2 = \omega + \sigma_{t-1}^2 ((1 - \alpha_1) + \alpha_1 z_{t-1}^2), \quad 0 < \alpha_1 \leq 1.$$

Когда $\omega = 0$, σ_t^2 представляет собой мартингал. На основании природы устойчивости в линейных моделях может показаться, что модель IGARCH(1, 1) с $\omega > 0$ и $\omega = 0$ аналогична случайному блужданию со смещением и без него, соответственно, а следовательно, является естественной моделью с устойчивыми шоками. Оказывается, это рассуждение обманчиво: в модели IGARCH(1, 1) с $\omega = 0$ последовательность σ_t^2 сходится к нулю почти наверное, а в модели IGARCH(1, 1) с $\omega > 0$ σ_t^2 строго стационарна и эргодична, а значит, не ведет себя как случайное блуждание, поскольку случайное блуждание почти наверное расходится. Имеется два определения устойчивости.

1. Первое определение устойчивости основано на сходимости распределения прогноза σ_t^2 . Предположим, что σ_t^2 строго стационарна и эргодична. Пусть $F(\sigma_t^2)$ – безусловная КФР для σ_t^2 , а $F_s(\sigma_t^2)$ – КФР для σ_t^2 , условная на информации на момент времени $s < t$. Для любого s $F(\sigma_t^2) - F_s(\sigma_t^2) \rightarrow 0$ во всех точках непрерывности при $t \rightarrow \infty$. Это означает, что устойчивость отсутствует, если процесс $\{\sigma_t^2\}$ стационарный и эргодичный.
2. Устойчивость также можно определить в терминах моментов прогноза. Для некоторой $\eta > 0$ шоки у σ_t^2 неустойчивы тогда и только тогда, когда для каждого s , $\mathbb{E}_s[\sigma_t^{2\eta}]$ сходится при $t \rightarrow \infty$ к конечному пределу, не зависящему от информационного множества на момент времени s .

Являются ли шоки у $\{\sigma_t^2\}$ устойчивыми или нет, напрямую зависит от используемого определения. Условный момент может расходиться для одного значения η , но сходиться к конкретному пределу, не зависящему от начальных условий, для другого значения η , даже если $\{\sigma_t^2\}$ стационарна и эргодична.

Для GARCH(1, 1)-модели

$$\mathbb{E}_s [\sigma_t^2] = \omega \sum_{k=0}^{t-s-1} (\alpha_1 + \beta_1)^k + \sigma_{t-s}^2 (\alpha_1 + \beta_1)^{t-s}.$$

$\mathbb{E}_s [\sigma_t^2]$ сходится к безусловной дисперсии $\omega / (1 - \alpha_1 - \beta_1)$ при $t \rightarrow \infty$ тогда и только тогда, когда $\alpha_1 + \beta_1 < 1$. В модели IGARCH(1, 1) с $\omega > 0$ и $\alpha_1 + \beta_1 = 1$ величина $\mathbb{E}_s [\sigma_t^2] \rightarrow \infty$ п.н. при $t \rightarrow \infty$. Тем не менее IGARCH-модели строго стационарны и эргодичны. Nelson (1990) показал, что в IGARCH(1, 1)-модели $\mathbb{E}_s [\sigma_t^{2\eta}]$ сходится к конечному пределу, не зависящему от информации на момент времени s при $t \rightarrow \infty$, если $\eta < 1$. Если носитель z_t неограничен, из этого следует, что в каждой стационарной и эргодичной GARCH(1, 1)-модели $\mathbb{E}_s [\sigma_t^{2\eta}]$ расходится для всех достаточно больших значений η и сходится для всех достаточно малых значений η .

4.7 Модель волатильности с постоянной и временной компонентами

Engle & Lee (1999) предположили, что высокая устойчивость и долгая память в рядах волатильности может быть связана с меняющимся во времени долгосрочным уровнем волатильности. Наличие единичного корня в процессе волатильности указывает на присутствие стохастического тренда наряду с временной компонентой волатильности доходности капитала. Декомпозиция условной дисперсии доходности активов на постоянную и временную компоненты – это способ исследовать долгосрочную и краткосрочную динамику волатильности фондового рынка (Engle & Lee, 1999). Используя представление GARCH(1, 1)-модели в форме с таргетированием дисперсии, запишем

$$\sigma_t^2 = (1 - \alpha_1 - \beta_1) \sigma^2 + \alpha_1 \varepsilon_{t-1}^2 + \beta_1 \sigma_{t-1}^2 = \sigma^2 + \alpha_1 (\varepsilon_{t-1}^2 - \sigma^2) + \beta_1 (\sigma_{t-1}^2 - \sigma^2). \quad (27)$$

Последние два члена имеют нулевое среднее. Эта модель расширяется на случай, допускающий непостоянство волатильности в долгосрочном периоде. Пусть q_t – *постоянная компонента* условной дисперсии. *Компонентная модель* условной дисперсии определяется как

$$\begin{aligned} \sigma_t^2 &= q_t + \alpha_1 (\varepsilon_{t-1}^2 - q_{t-1}) + \beta_1 (\sigma_{t-1}^2 - q_{t-1}) \\ &= q_t - (\alpha_1 + \beta_1) q_{t-1} + \alpha_1 \varepsilon_{t-1}^2 + \beta_1 \sigma_{t-1}^2. \end{aligned} \quad (28)$$

Используя лаговый оператор,

$$(1 - \beta_1 L) \sigma_t^2 = [1 - (\alpha_1 + \beta_1) L] q_t + \alpha_1 \varepsilon_{t-1}^2,$$

где

$$q_t = \omega + q_{t-1} + \phi (\varepsilon_{t-1}^2 - \sigma_{t-1}^2).$$

Постоянная волатильность σ^2 заменена на меняющийся во времени тренд, q_t , и его предыдущее значение. Ошибка прогноза, $\varepsilon_{t-1}^2 - \sigma_{t-1}^2$, является движущей силой зависящей от времени динамики тренда. Разность между условной дисперсией и ее трендом, $\sigma_{t-1}^2 - q_{t-1}$, составляет *временную компоненту* условной дисперсии.

Многошаговый прогноз тренда – это просто сумма текущего тренда и постоянного смещения:

$$\begin{aligned} q_{t+k} &= \omega + q_{t+k-1} + \phi (\varepsilon_{t+k-1}^2 - \sigma_{t+k-1}^2), \\ \mathbb{E}_{t-1} [q_{t+k}] &= \omega + \mathbb{E}_{t-1} [q_{t+k-1}] + \phi \mathbb{E}_{t-1} [\varepsilon_{t+k-1}^2 - \sigma_{t+k-1}^2], \end{aligned}$$

но $\mathbb{E}_{t-1} [\varepsilon_{t+k-1}^2] = \mathbb{E}_{t-1} [\sigma_{t+k-1}^2]$, так что $\mathbb{E}_{t-1} [\varepsilon_{t+k-1}^2 - \sigma_{t+k-1}^2] = 0$. Следовательно,

$$\begin{aligned} \mathbb{E}_{t-1} [q_{t+k}] &= \omega + \omega + \mathbb{E}_{t-1} [q_{t+k-2}] + \phi \mathbb{E}_{t-1} [\varepsilon_{t+k-2}^2 - \sigma_{t+k-2}^2] \\ &= \dots \\ &= k\omega + q_t. \end{aligned} \quad (29)$$

Из (28) следует, что

$$\sigma_{t+1}^2 - q_{t+1} = \alpha_1 (\varepsilon_t^2 - q_t) + \beta_1 (\sigma_t^2 - q_t).$$

Прогноз в момент времени $t - 1$ равен

$$\begin{aligned} \mathbb{E}_{t-1} [\sigma_{t+k}^2] - \mathbb{E}_{t-1} [q_{t+k}] &= (\alpha_1 + \beta_1) (\mathbb{E}_{t-1} [\sigma_{t+k-1}^2] - \mathbb{E}_{t-1} [q_{t+k-1}]) \\ &= (\alpha_1 + \beta_1)^k (\sigma_t^2 - q_t). \end{aligned}$$

Прогноз $\mathbb{E}_{t-1} [\sigma_{t+k}^2] - \mathbb{E}_{t-1} [q_{t+k}]$ при $\alpha_1 + \beta_1 < 1$ сходится к нулю по мере расширения горизонта прогнозирования на далекое будущее:

$$\mathbb{E}_{t-1} [\sigma_{t+k}^2] - \mathbb{E}_{t-1} [q_{t+k}] = 0 \text{ при } k \rightarrow \infty. \quad (30)$$

Таким образом, в долгосрочном периоде нет разницы между условной дисперсией и трендом. Поэтому q_t называют постоянной компонентой условной дисперсии. Объединяя (30) и (29), долгосрочный прогноз условной дисперсии – это просто сумма текущего ожидания тренда и постоянного смещения

$$\mathbb{E}_{t-1} [\sigma_{t+k}^2] = k\omega + q_t \text{ при } k \rightarrow \infty.$$

Компонентную модель можно расширить на случай процессов без единичных корней. Общая модель компонент имеет вид

$$\begin{aligned} \sigma_t^2 &= q_t + \alpha_1 (\varepsilon_{t-1}^2 - q_{t-1}) + \beta_1 (\sigma_{t-1}^2 - q_{t-1}), \\ q_t &= \omega + \rho q_{t-1} + \phi (\varepsilon_{t-1}^2 - \sigma_{t-1}^2), \end{aligned} \quad (31)$$

где q_t по-прежнему является компонентой условной дисперсии с более долгой памятью при $\rho > (\alpha_1 + \beta_1)$. Многошаговые прогнозы условной дисперсии и тренда имеют вид

$$\mathbb{E}_{t-1} [\sigma_{t+k}^2] - \mathbb{E}_{t-1} [q_{t+k}] = (\alpha_1 + \beta_1)^k (\sigma_t^2 - q_t), \quad (32)$$

$$\mathbb{E}_{t-1} [q_{t+k}] = \frac{(1 - \rho^k)}{(1 - \rho)} \omega + \rho^k q_t \quad (33)$$

для $\rho < 1$ и $(\alpha_1 + \beta_1) < 1$. Если $\rho > (\alpha_1 + \beta_1)$, временная компонента в (32) убывает быстрее тренда в (33), так что тренд доминирует в прогнозе условной дисперсии по мере расширения горизонта прогнозирования. По мере расширения горизонта планирования условная дисперсия сходится к константе, поскольку сам тренд является стационарным:

$$\mathbb{E}_{t-1} [\sigma_{t+k}^2] = \mathbb{E}_{t-1} [q_{t+k}] = \frac{\omega}{1 - \rho} \text{ при } k \rightarrow \infty.$$

Перепиывая (31) в виде

$$\sigma_t^2 = (1 - \alpha_1 L - \beta_1 L) q_t + \alpha_1 \varepsilon_{t-1}^2 + \beta_1 \sigma_{t-1}^2,$$

а (31) как

$$(1 - \rho L) q_t = \omega + \phi (\varepsilon_{t-1}^2 - \sigma_{t-1}^2), \quad (34)$$

и домножая на $(1 - \rho L)$, получаем общую модель компонент:

$$(1 - \rho L) \sigma_t^2 = (1 - \rho L) [(1 - \alpha_1 L - \beta_1 L) q_t + \alpha_1 \varepsilon_{t-1}^2 + \beta_1 \sigma_{t-1}^2]. \quad (35)$$

Подставляя (34) в (35), имеем

$$(1 - \rho L) \sigma_t^2 = (1 - \alpha_1 - \beta_1) \omega + (\phi + \alpha_1) \varepsilon_{t-1}^2 + (-\rho \alpha_1 - (\alpha_1 + \beta_1) \phi) \varepsilon_{t-2}^2 + (\rho - \phi + \beta_1) \sigma_{t-1}^2 + (\phi (\alpha_1 + \beta_1) - \beta_1 \rho) \sigma_{t-2}^2.$$

Процесс GARCH(2, 2) является процессом, порождающим данные, для условной дисперсии, определенной в модели компонент. При $\rho = \phi = 0$ модель упрощается до GARCH(1, 1). Таким образом, модель GARCH(1, 1) описывает лишь одну динамическую компоненту условной дисперсии.

5 Асимметричные модели

Для повышения гибкости исходная GARCH-модель была обобщена и расширена в разных направлениях. Первоначальная спецификация GARCH-модели предполагает, что реакция на шок не зависит от знака шока, а является функцией только от его размера. Тем не менее, один из стилизованных фактов о финансовой волатильности состоит в том, что плохие новости (отрицательные шоки) обычно оказывают большее влияние на волатильность, чем хорошие новости (положительные шоки). То есть волатильность стремится быть выше на падающем рынке, чем на растущем. Black (1976) объясняет этот эффект тем, что плохие новости обычно снижают цену акций, увеличивая коэффициент рычага (то есть отношение заемного и собственного капитала) и приводя к большей волатильности капитала. На основании этого предположения асимметричное воздействие новостей на волатильность обычно называют эффектом рычага. Альтернативные параметризации пытаются учесть асимметрию поведения волатильности. В этой главе представлены наиболее популярные в литературе асимметричные модели.

5.1 Модель EGARCH(p, q)

Простая структура (9) накладывает важные ограничения на GARCH-модели (Nelson, 1991).

- Существует отрицательная корреляция между доходностью капитала и изменениями волатильности доходов, то есть волатильность обычно растет в ответ на «плохие новости» (избыточная доходность меньше ожидаемой) и падает в ответ на «хорошие новости» (избыточная доходность выше ожидаемой). GARCH-модели, тем не менее, предполагают, что только величина, но не положительность или отрицательность неожиданной избыточной доходности определяет σ_t^2 . Если распределение z_t симметрично, изменение дисперсии завтра условно не коррелирует с избыточной доходностью сегодня (Nelson, 1991). Если записать σ_t^2 как функцию лагов σ_t^2 и z_t^2 , где $\varepsilon_t^2 = z_t^2 \sigma_t^2$, то есть

$$\sigma_t^2 = \omega + \sum_{j=1}^q \alpha_j z_{t-j}^2 \sigma_{t-j}^2 + \sum_{i=1}^p \beta_i \sigma_{t-i}^2, \quad (36)$$

становится очевидным, что условная дисперсия не зависит от изменений знака z_t^2 . Более того, инновации $z_{t-j}^2 \sigma_{t-j}^2$ не являются IID.

- Другое ограничение GARCH-моделей возникает из-за условий неотрицательности на ω^* и ϕ_k в (11), которые накладываются для обеспечения неотрицательности σ_t^2 для всех t с вероятностью 1. Из этих ограничений следует, что рост z_t^2 в любом периоде увеличивает σ_{t+m}^2 для всех $m \geq 1$, исключая случайное осциллирование процесса σ_t^2 .

- GARCH-модели неспособны объяснить наблюдаемую ковариацию между ε_t^2 и ε_{t-j} . Это возможно, только если условная дисперсия является асимметричной функцией от ε_{t-j} .
- В модели GARCH(1, 1) шоки могут быть устойчивыми в одной норме и затухающими в другой, так что условные моменты GARCH(1, 1)-модели могут взрываться даже когда процесс строго стационарный и эргодичный.
- GARCH-модели по сути специфицируют поведение квадратов данных. Поэтому несколько больших наблюдений могут доминировать в выборке.

Асимметричные модели дают объяснение так называемому *эффекту рычага*, когда неожиданное падение цены увеличивает волатильность больше, чем аналогичное неожиданное повышение цены. В экспоненциальной GARCH(p, q)-модели (EGARCH(p, q)), предложенной в Nelson (1991), σ_t^2 впервые зависит как от размера, так и от знака лагированных шоков. Модель определяется следующим образом:

$$\ln(\sigma_t^2) = \omega + \sum_{i=1}^q \alpha_i (\phi z_{t-i} + \psi (|z_{t-i}| - \mathbb{E}[|z_{t-i}|])) + \sum_{i=1}^p \beta_i \ln(\sigma_{t-i}^2), \quad (37)$$

где $\alpha_1 \equiv 1$, $\mathbb{E}[|z_t|] = (2/\pi)^{1/2}$, если $z_t \sim IIDN(0, 1)$, и где на параметры ω , β_i , α_i не накладывается ограничение неотрицательности. Положим

$$g(z_t) \equiv \phi z_t + \psi (|z_t| - \mathbb{E}[|z_t|]).$$

По построению, $\{g(z_t)\}_{t=-\infty}^{\infty}$ – IID случайная последовательность с нулевым средним. Компонентами $g(z_t)$ являются ϕz_t и $\psi (|z_t| - \mathbb{E}[|z_t|])$, и каждая имеет нулевое среднее. Если распределение z_t симметрично, компоненты ортогональны, хотя не являются независимыми. На множестве $0 < z_t < \infty$ функция $g(z_t)$ линейна по z_t с углом наклона $\phi + \psi$, а на множестве $-\infty < z_t \leq 0$ $g(z_t)$ линейна с углом наклона $\phi - \psi$. Таким образом, $g(z_t)$ позволяет процессу для условной дисперсии $\{\sigma_t^2\}$ асимметрично реагировать на увеличение и падение цены акций. Член $\psi (|z_t| - \mathbb{E}[|z_t|])$ представляет собой эффект величины. Если $\psi > 0$ и $\phi = 0$, инновация в $\ln(\sigma_{t+1}^2)$ положительна (отрицательна), когда величина z_t больше (меньше), чем ее матожидание. Если $\psi = 0$ и $\phi < 0$, инновация в $\ln(\sigma_{t+1}^2)$ положительна (отрицательна), когда инновации доходностей отрицательны (положительны). Отрицательный шок доходностей, увеличивающий отношение заемного и собственного капитала и, следовательно, неопределенность будущих доходов, можно учесть при $\alpha_i > 0$ и $\phi < 0$.

Как и в случае стандартной GARCH-модели, модель первого порядка на практике является самой популярной в семействе EGARCH.

Многие авторы, например Nelson (1991), обнаружили, что распределение стандартизированных остатков оцененных GARCH-моделей имеет положительный эксцесс, см. также Engle & Gonzalez-Rivera (1991). Nelson (1991) предполагает, что z_t имеет GED-распределение (также называемое семейством экспоненциально-степенных распределений). Плотность нормированной (с нулевым средним и единичной дисперсией) случайной величины, имеющей GED-распределение, имеет вид

$$f(z; v) = \frac{v \exp\left[-\left(\frac{1}{2}\right)|z/\lambda|^v\right]}{\lambda 2^{(1+1/v)} \Gamma(1/v)}, \quad -\infty < z < \infty, \quad 0 < v \leq \infty,$$

где $\Gamma(\cdot)$ – гамма-функция, а

$$\lambda \equiv \left(2^{(-2/v)} \Gamma(1/v) / \Gamma(3/v)\right)^{1/2}.$$

Параметр v отвечает за толщину хвостов распределения. Если $v = 2$, z имеет стандартное нормальное распределение. При $v < 2$ распределение z имеет более толстые хвосты, чем

нормальное (например, при $v = 1$ z имеет двойное экспоненциальное распределение), а при $v > 2$ распределение z имеет более тонкие хвосты, чем нормальное (например, при $v = \infty$ z равномерно распределена на отрезке $[-3^{1/2}, 3^{1/2}]$). Для такой функции плотности $\mathbb{E}[|z_t|] = \lambda 2^{1/v} \Gamma(2/v) / \Gamma(1/v)$ (Hamilton, 1994).

Для анализа стационарности перепишем модель EGARCH(p, q) в виде

$$\left(1 - \sum_{i=1}^p \beta_i L^i\right) \ln(\sigma_t^2) = \omega + \sum_{i=1}^q \alpha_i L^i (\phi z_t + \psi (|z_t| - \mathbb{E}[|z_t|])),$$

$$\ln(\sigma_t^2) = \left(1 - \sum_{i=1}^p \beta_i\right)^{-1} \omega + \left(1 - \sum_{i=1}^p \beta_i L^i\right)^{-1} \left(\sum_{i=1}^q \alpha_i L^i\right) g(z_t),$$

$$\ln(\sigma_t^2) = \omega^* + \sum_{i=1}^{\infty} \varphi_i g(z_{t-i}).$$

В EGARCH(p, q)-модели $\ln(\sigma_t^2)$ – линейный процесс, поэтому его стационарность (ковариационную или строгую) и эргодичность легко установить. В этом важная разница по сравнению с GARCH-моделями, для которых сложно оценить устойчивость шоков дисперсии. Если $\phi \neq 0$ или $\psi \neq 0$, процессы $\{\exp(-\omega^*) \sigma_t^2\}$ и $\{\exp(-\omega^*/2) \varepsilon_t\}$ строго стационарны и эргодичны, а процесс $\{\ln(\sigma_t^2) - \omega^*\}$ ковариационно стационарен тогда и только тогда, когда $\sum_{i=1}^{\infty} \varphi_i^2 < \infty$ (см. Nelson, 1991). Поскольку процесс $\ln(\sigma_t^2)$ записан в форме ARMA(p, q), если полиномы $1 - \sum_{i=1}^p \beta_i x^i$ и $\sum_{i=1}^q \alpha_i x^i$ не имеют общих корней, то условия строгой стационарности $\ln(\sigma_t^2)$ эквивалентны требованию, чтобы все корни $1 - \sum_{i=1}^p \beta_i x^i$ лежали вне единичного круга.

Из строгой стационарности $\{\exp(-\omega^*) \sigma_t^2\}$ и $\{\exp(-\omega^*/2) \varepsilon_t\}$ необязательно следует ковариационная стационарность σ_t^2 , так как $\{\exp(-\omega^*) \sigma_t^2\}$ и $\{\exp(-\omega^*/2) \varepsilon_t\}$ могут не иметь конечные безусловные среднее и дисперсию. Для некоторых распределений $\{z_t\}$ (например, для t -распределения Стьюдента с конечным числом степеней свободы), конечные безусловные моменты $\{\exp(-\omega^*) \sigma_t^2\}$ и $\{\exp(-\omega^*/2) \varepsilon_t\}$ не существуют. Если z_t имеет GED-распределение с более тонкими хвостами, чем у двойного экспоненциального распределения, и если $\sum_{i=1}^{\infty} \varphi_i^2 < \infty$, то $\{\sigma_t^2\}$ и $\{\varepsilon_t\}$ не только строго стационарны и эргодичны, но и имеют произвольные конечные моменты, из чего следует их ковариационная стационарность.

He et al. (2002) заметили, что убывание автокорреляций квадратов наблюдений в модели EGARCH первого порядка быстрее экспоненциального в начале, а затем замедляется, достигая экспоненциального падения. Более того, Malmsten & Teräsvirta (2004) показали, что симметричная EGARCH(1, 1)-модель с нормальными ошибками недостаточно гибка для характеристики рядов как с большим куртозисом, так и с медленно убывающими автокорреляциями. В этом случае выбор распределения стандартизированных ошибок особенно важен. Предположение о нормальных ошибках означает, что автокорреляция квадратов наблюдений первого порядка растет достаточно быстро как функция от куртозиса для любого фиксированного значения β_1 , прежде чем увеличение замедляется. Nelson (1991) предложил использовать для ошибок GED-распределение, в то время как выбор t -распределения может означать бесконечную безусловную дисперсию $\{\varepsilon_t\}$. Как и в случае модели GARCH(1, 1), выбор распределения ошибок с более толстыми, чем у нормального распределения, хвостами увеличивает куртозис и в то же время уменьшает автокорреляцию квадратов или абсолютных значений наблюдений.

5.2 Другие асимметричные модели

Нелинейная GARCH(p, q)-модель (Engle & Bollerslev, 1986):

$$\sigma_t^2 = \omega + \sum_{i=1}^q \alpha_i |\varepsilon_{t-i} - k|^\delta + \sum_{i=1}^p \beta_i \sigma_{t-i}^2, \quad (38)$$

где $k \neq 0$, и инновации в σ_t^2 зависят как от размера, так и от знака лагированных остатков, улавливая таким образом эффект рычага в волатильности доходности капитала.

Модель GJR-GARCH, предложенная в Glosten et al. (1993):

$$\sigma_t^2 = \omega + \sum_{i=1}^p \beta_i \sigma_{t-i}^2 + \sum_{i=1}^q (\alpha_i \varepsilon_{t-i}^2 + \gamma_i S_{t-i}^- \varepsilon_{t-i}^2), \quad (39)$$

где

$$S_t^- = \begin{cases} 1, & \varepsilon_t < 0, \\ 0, & \varepsilon_t \geq 0. \end{cases} \quad (40)$$

Асимметричная GARCH(p, q)-модель (AGARCH), предложенная в Engle (1990):

$$\sigma_t^2 = \omega + \sum_{i=1}^q \alpha_i (\varepsilon_{t-i} + \gamma)^2 + \sum_{i=1}^p \beta_i \sigma_{t-i}^2. \quad (41)$$

Отрицательные значения γ означают, что положительные шоки ведут к меньшим увеличениям будущей волатильности, чем отрицательные шоки той же абсолютной величины.

Нелинейная AGARCH(1, 1)-модель (NAGARCH) из Engle & Ng (1993):

$$\sigma_t^2 = \omega + \alpha_1 \left(\frac{\varepsilon_{t-1}}{\sigma_{t-1}} + \gamma \right)^2 + \beta_1 \sigma_{t-1}^2. \quad (42)$$

Модель QGARCH(p, q) из Sentana (1995):

$$\sigma_t^2 = \sigma^2 + \Psi' \mathbf{x}_{t-q} + \mathbf{x}'_{t-q} \mathbf{A} \mathbf{x}_{t-q} + \sum_{i=1}^p \beta_i \sigma_{t-i}^2, \quad (43)$$

где $\mathbf{x}_{t-q} = (\varepsilon_{t-1}, \dots, \varepsilon_{t-q})'$. Линейный член $(\Psi' \mathbf{x}_{t-q})$ допускает асимметрию. Внедиагональные элементы симметричной матрицы параметров \mathbf{A} учитывают эффекты взаимодействия лагированных значений x_t на условную дисперсию. Условная дисперсия σ_t^2 положительна тогда и только тогда, когда \mathbf{A} в квадратичной форме положительно определена. Модель QGARCH включает как частные случаи несколько асимметричных моделей. Обобщенная расширенная ARCH-модель (AARCH), являющаяся обобщением AARCH-модели из Bera et al. (1992), предполагает, что $\Psi = \mathbf{0}$. Модель ARCH(q) соответствует случаю, в котором $\Psi = \mathbf{0}$, $\beta_i = 0$, $i = 1, \dots, q$, и \mathbf{A} диагональна. Асимметричная GARCH-модель предполагает диагональную \mathbf{A} . Модель линейного стандартного отклонения соответствует значениям $\beta_i = 0$, $\sigma^2 = \rho^2$, $\Psi = 2\rho\phi$ и $\mathbf{A} = \phi\phi'$:

$$\sigma_t^2 = (\rho + \phi' \mathbf{x}_{t-q})^2.$$

5.3 Кривая воздействия новостей

Новости оказывают асимметричное воздействие на волатильность. В модели с асимметричной волатильностью хорошие и плохие новости имеют разную предсказательную силу для будущей волатильности. Кривая воздействия новостей (КВН), предложенная в Pagan & Schwert (1990) и получившая свое название в Engle & Ng (1993), характеризует воздействие прошлых шоков доходности на волатильность доходности, неявное в модели волатильности. Фиксируя информацию на момент $t-2$ и раньше, можно исследовать подразумеваемую взаимосвязь между ε_{t-1} и σ_t^2 . Все лагированные условные дисперсии оцениваются на уровне безусловной дисперсии доходности акций. КВН связывает прошлые шоки доходности (новости) с текущей волатильностью. Эта кривая указывает на то, как новая информация включается в оценки волатильности.

В GARCH-модели КВН центрирована в $\varepsilon_{t-1} = 0$. В случае EGARCH-модели кривая достигает наименьшего значения в $\varepsilon_{t-1} = 0$ и экспоненциально возрастает в обоих направлениях, но с разными параметрами.

GARCH(1, 1):

$$\sigma_t^2 = \omega + \alpha \varepsilon_{t-1}^2 + \beta \sigma_{t-1}^2.$$

Если $\sigma_{t-1}^2 = \sigma^2$, КВН имеет следующий вид:

$$\sigma_t^2 = A + \alpha \varepsilon_{t-1}^2, \quad A \equiv \omega + \beta \sigma^2.$$

EGARCH(1, 1):

$$\ln(\sigma_t^2) = \omega + \beta \ln(\sigma_{t-1}^2) + \phi z_{t-1} + \psi (|z_{t-1}| - \mathbb{E}[|z_{t-1}|]),$$

где $z_t = \varepsilon_t / \sigma_t$. Кривая воздействия новостей имеет вид

$$\sigma_t^2 = \begin{cases} A \exp\left(\frac{\phi + \psi}{\sigma} \varepsilon_{t-1}\right), & \text{при } \varepsilon_{t-1} > 0, \\ A \exp\left(\frac{\phi - \psi}{\sigma} \varepsilon_{t-1}\right), & \text{при } \varepsilon_{t-1} < 0, \end{cases}$$

где $A \equiv \sigma^{2\beta} \exp\left(\omega - \psi \sqrt{2/\pi}\right)$, $\phi < 0$, $\psi + \phi > 0$.

- Модель EGARCH позволяет хорошим и плохим новостям иметь различный эффект на волатильность, в отличие от стандартной GARCH-модели.
- Модель EGARCH позволяет важным новостям иметь большее воздействие на волатильность по сравнению с обычной GARCH-моделью. В EGARCH-модели дисперсия в обоих направлениях выше, поскольку экспонента в конечном счете доминирует квадратичную функцию.

Для асимметричной GARCH(1, 1)-модели (Engle, 1990)

$$\sigma_t^2 = \omega + \alpha (\varepsilon_{t-1} + \gamma)^2 + \beta \sigma_{t-1}^2,$$

КВН, имеющая вид

$$\sigma_t^2 = A + \alpha (\varepsilon_{t-1} + \gamma)^2, \tag{44}$$

где $A \equiv \omega + \beta \sigma^2$, $\omega > 0$, $0 \leq \beta < 1$, $\sigma > 0$, $0 \leq \alpha < 1$, асимметрична и центрирована в $\varepsilon_{t-1} = -\gamma$.

Для GJR-GARCH-модели

$$\sigma_t^2 = \omega + \alpha \varepsilon_t^2 + \beta \sigma_{t-1}^2 + \gamma S_{t-1}^- \varepsilon_{t-1}^2,$$

КВН имеет вид

$$\sigma_t^2 = \begin{cases} A + \alpha \varepsilon_{t-1}^2, & \varepsilon_{t-1} \geq 0, \\ A + (\alpha + \gamma) \varepsilon_{t-1}^2, & \varepsilon_{t-1} < 0, \end{cases}$$

где $A \equiv \omega + \beta \sigma^2$, $\omega > 0$, $0 \leq \beta < 1$, $\sigma > 0$, $0 \leq \alpha < 1$, $\alpha + \beta < 1$, и центрирована в $\varepsilon_{t-1} = -\gamma$.

Эти различия между кривыми воздействия новостей имеют большие последствия для выбора портфеля и ценообразования активов. Поскольку предсказуемая волатильность рынка связана с рыночной премией, две модели ведут к очень разным рыночным рисковым премиям, а следовательно, разным рисковым премиям для индивидуальных акций в условной версии модели CAPM. Различия в предсказуемой волатильности после выхода некоторой важной новости ведут к существенным различиям текущей цены опционов и разным динамическим стратегиям хеджирования.

6 Нелинейная GARCH

6.1 GARCH с плавным переходом

Реакция условной дисперсии на шоки может быть нелинейной, как, например, в модели GJR-GARCH. Нелинейная версия модели GJR-GARCH получается, если сделать переход между режимами плавным. Модель GARCH с плавным переходом STGARCH (Gonzalez-Rivera, 1998, и Anderson et al., 1999) позволяет эффекту квадратов прошлых инноваций зависеть как от знака, так и от величины ε_{t-i} через функцию плавного перехода:

$$\sigma_t^2 = \omega_1 + \sum_{i=1}^q \alpha_{1i} \varepsilon_{t-i}^2 + \left(\omega_2 + \sum_{i=1}^q \alpha_{2i} \varepsilon_{t-i}^2 \right) G(\gamma, c; \varepsilon_{t-i}) + \sum_{i=1}^p \beta_i \sigma_{t-i}^2, \quad (45)$$

где функция перехода имеет вид

$$G(\gamma, c; \varepsilon_{t-i}) = \left(1 + \exp \left\{ -\gamma \prod_{k=1}^K (\varepsilon_{t-i} - c_k) \right\} \right)^{-1}. \quad (46)$$

- При $K = 1$ G упрощается до простой логистической функции, которая контролирует переход коэффициента ε_{t-i}^2 от α_{1i} к $\alpha_{1i} + \alpha_{2i}$ как функцию от ε_{t-i} , и аналогично для константы. При $\gamma \rightarrow \infty$ функция плавного перехода становится ступенчатой функцией и соответствует резкому переключению из одного режима в другой. Если дополнительно положить $c_1 = 0$, получится GJR-GARCH-модель, поскольку ε_t и z_t имеют одинаковый знак. Модель STGARCH при $K = 1$ предназначена для моделирования асимметричных откликов на шоки.
- При $K = 2$ и $c_1 = -c_2$ в (46) модель остается симметричной относительно нуля, но реакция условной дисперсии на шоки является нелинейной функцией от лагов ε_t^2 . Модель STGARCH при $K = 2$ полезна, когда предположение о двух различных режимах не является адекватным приближением асимметричного поведения условной дисперсии.

Lanne & Saikkonen (2005) предложили модель GARCH с плавным переходом, чей первый порядок имеет вид

$$\sigma_t^2 = \omega + \alpha_1 \varepsilon_{t-1}^2 + \delta_1 G_1(\vartheta; \sigma_{t-1}) + \beta_1 \sigma_{t-1}^2. \quad (47)$$

$G_1(\vartheta; \sigma_{t-1})$ – непрерывная ограниченная функция, как, например, в (45). Lanne & Saikkonen (2005) используют КФР гамма-распределения. В этой модели переменная перехода – лагированная условная дисперсия. Такая параметризация смягчает нежелательное свойство стандартной GARCH-модели, которая часто преувеличивает устойчивость волатильности.

6.2 Пороговая GARCH

В пороговой GARCH-модели (TGARCH), предложенной в Zakoian (1994), условное стандартное отклонение определяется как

$$\sigma_t = \omega + \sum_{i=1}^q (\alpha_i^+ \varepsilon_{t-i}^+ - \alpha_i^- \varepsilon_{t-i}^-) + \sum_{i=1}^p \beta_i \sigma_{t-i}, \quad (48)$$

где $\varepsilon_{t-i}^+ = \max(\varepsilon_{t-i}, 0)$, $\varepsilon_{t-i}^- = \min(\varepsilon_{t-i}, 0)$, и α_i^+ , α_i^- , $i = 1, \dots, q$, – параметры. TGARCH-модель линейна по параметрам, поскольку пороговый параметр полагается равным нулю. ARCH-модель с двойным порогом (DARCH), предложенная в Li & Li (1996), позволяет параметрам условного среднего и дисперсии меняться в зависимости от режима. Условное среднее определяется следующим образом:

$$y_t = \sum_{k=1}^K \left(\phi_{0k} + \sum_{i=1}^{p_k} \phi_{ik} y_{t-i} \right) \mathbb{I}\{c_{k-1}^{(m)} < y_{t-b} \leq c_k^{(m)}\} + \varepsilon_t, \quad (49)$$

а условная дисперсия имеет вид

$$\sigma_t = \sum_{l=1}^L \left(\alpha_{0l} + \sum_{i=1}^{p_l} \alpha_{il} \varepsilon_{t-i}^2 \right) \mathbb{I}\{c_{l-1}^{(v)} < y_{t-d} \leq c_l^{(v)}\}, \quad (50)$$

где b и d – параметры запаздывания, $b, d \geq 1$. Число режимов в (49) и (50), K и L , соответственно, не должно совпадать, как и две пороговые переменные. Можно использовать другие пороговые переменные, помимо лагов y_t .

6.3 GARCH с марковским переключениями

Модель с марковскими переключениями (MS) имеет нелинейную спецификацию, в которой различные состояния мира воздействуют на динамику временного ряда. Динамические свойства зависят от текущего режима, а сами режимы являются реализациями скрытой марковской цепи с конечным пространством состояний. Модели с марковским переключениями были введены в мейнстрим эконометрики в Hamilton (1989, 1990). Hamilton & Susmel (1994) заметили, что последствия очень больших шоков, таких как биржевой крах в октябре 1987 года, для последующей волатильности могут настолько отличаться от последствий малых шоков, что стандартная ARCH или GARCH-модель неспособны это уловить надлежащим образом. Более того, Hamilton & Susmel (1994) различают режимы с низкой, средней и высокой волатильностью в данных о недельных доходностях акций, причем режимы высокой волатильности связаны с рецессиями, в то время как Maheu & McCurdy (2000) выделяют рынки медведей и быков и обнаруживают, что волатильность намного выше на медвежьих рынках.

Модель ARCH с переключением режимов (SWARCH), независимо предложенная в Cai (1994) и Hamilton & Susmel (1994), расширяет стандартную линейную ARCH(q)-модель, позволяя константе, $\omega_{s(t)}$, или величине квадратов инноваций, входящих в уравнение условной дисперсии, зависеть от некоторой латентной переменной $s(t)$, причем переход между разными состояниями происходит в соответствии с марковской цепью. Модель из Hamilton & Susmel (1994) выглядит следующим образом:

$$\sigma_t^2 = \omega(s_t) + \sum_{i=1}^q \alpha_i(s_t) \varepsilon_{t-i}^2, \quad (51)$$

где s_t – дискретная ненаблюдаемая эргодичная марковская цепь, заданная на множестве $\mathbb{S} = \{1, \dots, S\}$ индикаторов режимов с переходными вероятностями

$$p_{ij} = \mathbb{P}\{s_t = j | s_{t-1} = i\}, \quad i, j = 1, \dots, S,$$

где, по определению, $\sum_{j \in \mathbb{S}} p_{ij} = 1, \forall i \in \mathbb{S}$. В моделях с марковскими переключениями переменная переключения s_t экзогенна в том смысле, что она не подвержена обратному воздействию со стороны наблюдаемого процесса. При $q = 1$ $\mathbb{V}[\varepsilon_t | s_t = i, \varepsilon_{t-1}] = \omega(i) + \alpha_1(i)\varepsilon_{t-1}^2$. Эта модель позволяет σ_t^2 меняться между режимами высокой и низкой волатильности, что измеряется значениями параметров ARCH. Cai (1994) рассматривает частный случай (51), в котором переключается только константа $\omega(s_t)$ и $S = 2$. Обобщение моделей с марковскими переключениями на случай GARCH-моделей было предложено в Gray (1996) и Dueker (1997) и позже модифицировано в Klaassen (2002). Модель GARCH(1, 1) с марковскими переключениями имеет вид

$$\sigma_t^2 = \omega(s_t) + \alpha_1(s_t)\varepsilon_{t-1}^2 + \beta_1(s_t)\sigma_{t-1}^2. \quad (52)$$

Условия стационарности, существование моментов, автокорреляционные функции и эргodicность были изучены в Francq et al. (2001) и Francq & Zakoian (2005).

В этой модели есть одна существенная проблема. Если переписать (52), используя рекурсивную подстановку, при данном начальном значении σ_0^2 ,

$$\sigma_t^2 = \sum_{i=0}^{t-1} (\omega(s_{t-i}) + \alpha_1(s_{t-i})\varepsilon_{t-i-1}^2) \prod_{j=0}^{i-1} \beta_1(s_{t-j}) + \sigma_0^2 \prod_{i=0}^{t-1} \beta_1(s_{t-i}),$$

то ясно, что σ_t^2 зависит от всей истории смены режимов. Нахождение логарифмической функции правдоподобия для выборки длины T требует интегрирования по всем \mathbb{S}^T возможным (ненаблюдаемым) траекториям режимов, что делает оценивание (52) недоступным на практике. В этой функции следует избавиться от ненаблюдаемых величин путем интегрирования. При $z_t \sim IIDN(0, 1)$ ε_t является смесью нормальных распределений с меняющимися во времени весами смешивания. Веса смешивания можно подсчитать как $\mathbb{P}\{s_t = j | (\varepsilon_n, s_n)_{n \leq t-1}\}$, где $(\varepsilon_n, s_n)_{n \leq t-1}$ обозначает информацию до момента $t - 1$ включительно. Условную дисперсию ε_{t-1} при наличии информации в момент $t - 2$ можно найти по формуле

$$h_{t-1} = \sum_{j=1}^S \mathbb{P}\{s_{t-1} = j | (\varepsilon_n, s_n)_{n \leq t-2}\} \sigma_{jt-1}^2, \quad (53)$$

где $\sigma_{jt-1}^2 = \mathbb{V}[\varepsilon_{t-1} | s_{t-1} = j]$. Таким способом можно избавиться от ненаблюдаемых величин $s_{t-i}, i = 0, \dots, t$. Gray (1996) использует h_{t-1} вместо σ_{t-1}^2 в (52) для условной дисперсии в каждом режиме:

$$\sigma_t^2(s_t) = \omega(s_t) + \alpha_1(s_t)\varepsilon_{t-1}^2 + \beta_1(s_t)h_{t-1}^2, \quad s_t = 1, \dots, S. \quad (54)$$

Функцию правдоподобия можно найти, используя рекурсивную схему первого порядка, как в базовой модели из Hamilton (1989).

Klaassen (2002) предлагает использовать $\mathbb{P}\{s_{t-1} = j | (\varepsilon_n, s_n)_{n \leq t-1}\}$ в (53), то есть информацию вплоть до момента $t - 1$, а не $t - 2$. Вдобавок он использует тот факт, что режим в момент t принадлежит информационному множеству, на котором обусловлена σ_t^2 , и использует вероятности $\mathbb{P}\{s_{t-1} = j | s_t = i, (\varepsilon_n, s_n)_{n \leq t-1}\}$ для обновления дисперсии в i -ом режиме.

В модели MS-GARCH из Gray (1996) интерпретация параметра неоднозначна. Действительно, $\beta_1(s_t)$ нельзя рассматривать как инерцию, поскольку h_{t-1} состоит из дисперсий всех компонент, а также меняющихся во времени условных вероятностей режимов. Таким образом, дисперсии режимов (а не только общая дисперсия, как в базовой модели с марковским переключениями) будут меняться просто потому, что условные вероятности режимов непостоянны, что противоречит определению шоков, движущих процессы внутри режима.

Haas et al. (2004) предлагают многорежимную GARCH-модель, которая сохраняет соответствие между параметрами GARCH-модели в j -ом режиме и процессом для дисперсии. S -мерный вектор $\sigma_t^{(2)} = (\sigma_{1t}^2, \dots, \sigma_{St}^2)'$ дисперсий режимов следует модели GARCH(1, 1):

$$\sigma_t^{(2)} = \omega + \alpha \varepsilon_{t-1}^2 + \beta \sigma_{t-1}^{(2)}, \quad (55)$$

где $\omega = (\omega_1, \dots, \omega_S)'$, $\alpha = (\alpha_1, \dots, \alpha_S)'$, $\beta = \text{diag}(\beta_1, \dots, \beta_S)'$. Предполагается, что неравенства $\omega > 0$, $\alpha, \beta \geq 0$ выполняются поэлементно, чтобы обеспечить положительность процесса для дисперсии. Если $\max\{\beta_1, \dots, \beta_S\} < 1$, уравнение (55) можно обратить и получить $\sigma_t^{(2)} = (I_S - \beta)^{-1}\omega + (I_S - \beta L)^{-1}\alpha \varepsilon_{t-1}^2$. В этой модели динамика дисперсии меняется, если происходит смена режима. Условная дисперсия в данном режиме является функцией только лагированной условной дисперсии в том же режиме, в отличие от других моделей. Если истинная модель имеет меньше режимов, чем специфицированная, последняя содержит неидентифицируемые мешающие параметры. Haas et al. (2004) устанавливают условия стационарности второго порядка.

7 Обобщающие модели

Распространение GARCH-моделей вдохновило некоторых авторов на определение семейств GARCH-моделей, которые включали бы в себя как можно больше частных моделей. Вложенность показывает связь между моделями и позволяет применять стандартные тесты на вложенность для определения относительного качества подгонки для каждой модели. В целом обобщающие модели основаны на трансформации условного стандартного отклонения. Taylor (1986) и Schwert (1989) первыми предложили ARCH-модели для условного стандартного отклонения:

$$\sigma_t = \omega + \sum_{i=1}^q \alpha_i |\varepsilon_{t-i}| + \sum_{i=1}^p \beta_i \sigma_{t-i}. \quad (56)$$

Их мотивация опирается на Davidian & Carroll (1987), где показано в рамках модели регрессии, что оценки дисперсии, основанные на абсолютной величине остатков, устойчивы к выбросам в данных. Вследствие неравенства Йенсена большие шоки имеют меньший эффект на условную дисперсию, чем в стандартной GARCH-модели. Nelson & Foster (1994) показывают, что GARCH-расширение модели Тейлора-Шверта дает состоятельную оценку процессов, близким к диффузионным; более того, они показывают, что для распределений с положительным эксцессом GARCH для абсолютных значений является более эффективным фильтром для условной дисперсии, чем стандартная GARCH-модель.

Нелинейную GARCH-модель в Higgins & Bera (1992), имеющую вид

$$\sigma_t^\delta = \omega + \alpha_1 |\varepsilon_{t-1}|^\delta + \beta_1 \sigma_{t-1}^\delta, \quad (57)$$

можно трактовать как расширение модели условного стандартного отклонения, поскольку она параметризует условное стандартное отклонение в степени δ как функцию от лагированных условных стандартных отклонений и лагированных абсолютных значений инноваций, возведенных в ту же степень. Эта формулировка, очевидно, сводится к стандартной GARCH(p, q)-модели при $\delta = 2$.

Первая в литературе обобщающая модель – это асимметричная степенная GARCH(p, q)-модель (APGARCH) из Ding et al. (1993), имеющая вид

$$\sigma_t^\delta = \omega + \sum_{i=1}^q \alpha_i (|\varepsilon_{t-i}| - \gamma_i \varepsilon_{t-i})^\delta + \sum_{j=1}^p \beta_j \sigma_{t-j}^\delta, \quad (58)$$

где

$$\begin{aligned}\omega &> 0, \quad \delta \geq 0, \\ \alpha_i &\geq 0, \quad 1 > \gamma_i > -1, \quad i = 1, \dots, q, \\ \beta_j &\geq 0, \quad j = 1, \dots, p.\end{aligned}$$

Эта модель применяет преобразование Бокса-Кокса к процессу для условного стандартного отклонения и асимметричным абсолютным значениям остатков. Асимметричная реакция волатильности на положительные и отрицательные шоки – это хорошо известный эффект рычага. Если предположить, что распределение ε_t условно нормальное, то условие существования $\mathbb{E}[\sigma_t^\delta]$ и $\mathbb{E}[|\varepsilon_t|^\delta]$ имеет вид

$$\frac{1}{\sqrt{2\pi}} \sum_{i=1}^q \alpha_i \left((1 + \gamma_i)^\delta + (1 - \gamma_i)^\delta \right) 2^{\frac{\delta-1}{2}} \Gamma\left(\frac{\delta+1}{2}\right) + \sum_{j=1}^p \beta_j < 1.$$

Если это условие удовлетворено, то при $\delta \geq 2$ процесс ε_t ковариационно стационарный. Но $\delta \geq 2$ является достаточным условием ковариационной стационарности процесса ε_t . Karanasos & Kim (2006) изучают автокорреляционную структуру общей APGARCH(p, q)-модели. Эта обобщенная версия GARCH-модели включает в себя семь других моделей как частные случаи.

1. ARCH(q)-модель, при $\delta = 2$ и $\gamma_i = 0, i = 1, \dots, q, \beta_j = 0, j = 1, \dots, p$.
 2. GARCH(p, q)-модель, при $\delta = 2$ и $\gamma_i = 0, i = 1, \dots, q$.
 3. Модель условного стандартного отклонения из Taylor (1986) и Schwert (1989), при $\delta = 1$ и $\gamma_i = 0, i = 1, \dots, q$.
 4. GJR-GARCH-модель, при $\delta = 2$.
- Если $\delta = 2$ и $0 \leq \gamma_i < 1$,

$$\begin{aligned}\sigma_t^2 &= \omega + \sum_{i=1}^q \alpha_i (|\varepsilon_{t-i}| - \gamma_i \varepsilon_{t-i})^2 + \sum_{j=1}^p \beta_j \sigma_{t-j}^2 \\ &= \omega + \sum_{i=1}^q \alpha_i \left(|\varepsilon_{t-i}|^2 + \gamma_i^2 \varepsilon_{t-i}^2 - 2\gamma_i |\varepsilon_{t-i}| \varepsilon_{t-i} \right) + \sum_{j=1}^p \beta_j \sigma_{t-j}^2\end{aligned}$$

Раскрывая модуль, получаем

$$\sigma_t^2 = \begin{cases} \omega + \sum_{i=1}^q \alpha_i^2 (1 + \gamma_i)^2 \varepsilon_{t-i}^2 + \sum_{j=1}^p \beta_j \sigma_{t-j}^2, & \varepsilon_{t-i} < 0, \\ \omega + \sum_{i=1}^q \alpha_i (1 - \gamma_i)^2 \varepsilon_{t-i}^2 + \sum_{j=1}^p \beta_j \sigma_{t-j}^2, & \varepsilon_{t-i} > 0.\end{cases}$$

Следовательно,

$$\begin{aligned}\sigma_t^2 &= \omega + \sum_{i=1}^q \alpha_i (1 - \gamma_i)^2 \varepsilon_{t-i}^2 + \sum_{i=1}^q \alpha_i \left((1 + \gamma_i)^2 - (1 - \gamma_i)^2 \right) S_i^- \varepsilon_{t-i}^2 + \sum_{j=1}^p \beta_j \sigma_{t-j}^2 \\ &= \omega + \sum_{i=1}^q \alpha_i (1 - \gamma_i)^2 \varepsilon_{t-i}^2 + \sum_{j=1}^p \beta_j \sigma_{t-j}^2 + \sum_{i=1}^q 4\alpha_i \gamma_i S_i^- \varepsilon_{t-i}^2.\end{aligned}$$

Полагая $\alpha_i^* = \alpha_i(1 - \gamma_i)^2$, $\gamma_i^* = 4\alpha_i\gamma_i$, получим

$$\sigma_t^2 = \omega + \sum_{i=1}^p \alpha_i(1 - \gamma_i)^2 \varepsilon_{t-i}^2 + \sum_{j=1}^q \beta_j \sigma_{t-j}^2 + \sum_{i=1}^p \gamma_i^* S_i^- \varepsilon_{t-i}^2,$$

то есть модель GJR-GARCH.

При $-1 \leq \gamma_i < 0$ имеем

$$\begin{aligned} \sigma_t^2 &= \omega + \sum_{i=1}^q \alpha_i (|\varepsilon_{t-i}| - \gamma_i \varepsilon_{t-i})^2 + \sum_{j=1}^p \beta_j \sigma_{t-j}^2 \\ &= \begin{cases} \omega + \sum_{i=1}^q \alpha_i (1 - \gamma_i)^2 \varepsilon_{t-i}^2 + \sum_{j=1}^p \beta_j \sigma_{t-j}^2, & \varepsilon_{t-i} > 0, \\ \omega + \sum_{i=1}^q \alpha_i (1 + \gamma_i)^2 \varepsilon_{t-i}^2 + \sum_{j=1}^p \beta_j \sigma_{t-j}^2, & \varepsilon_{t-i} < 0. \end{cases} \end{aligned}$$

Перепишывая, получаем

$$\begin{aligned} \sigma_t^2 &= \omega + \sum_{i=1}^q \alpha_i (1 + \gamma_i)^2 \varepsilon_{t-i}^2 + \sum_{j=1}^p \beta_j \sigma_{t-j}^2 + \sum_{i=1}^q \alpha_i \left((1 - \gamma_i)^2 - (1 + \gamma_i)^2 \right) S_i^+ \varepsilon_{t-i}^2 \\ &= \omega + \sum_{i=1}^q \alpha_i (1 + \gamma_i)^2 \varepsilon_{t-i}^2 + \sum_{j=1}^p \beta_j \sigma_{t-j}^2 - \sum_{i=1}^q 4\alpha_i \gamma_i S_i^+ \varepsilon_{t-i}^2, \end{aligned}$$

где

$$S_i^+ = \begin{cases} 1, & \varepsilon_{t-i} \geq 0, \\ 0, & \text{иначе.} \end{cases}$$

Полагая $\alpha_i^* = \alpha_i(1 + \gamma_i)^2$, $\gamma_i^* = -4\alpha_i\gamma_i$, получаем модель

$$\sigma_t^2 = \omega + \sum_{i=1}^q \alpha_i^* \varepsilon_{t-i}^2 + \sum_{j=1}^p \beta_j \sigma_{t-j}^2 + \sum_{i=1}^q \gamma_i^* S_i^+ \varepsilon_{t-i}^2,$$

позволяющую положительным шокам иметь больший эффект на волатильность.

5. TARСН-модель из Zakoian (1994), при $\delta = 1$ и $\beta_j = 0$, $j = 1, \dots, p$. Получаем

$$\begin{aligned} \sigma_t &= \omega + \sum_{i=1}^q \alpha_i (|\varepsilon_{t-i}| - \gamma_i \varepsilon_{t-i}) = \\ &= \omega + \sum_{i=1}^q \alpha_i (1 - \gamma_i) \varepsilon_{t-i}^+ - \sum_{i=1}^q \alpha_i (1 + \gamma_i) \varepsilon_{t-i}^-, \end{aligned}$$

где $\varepsilon_{t-i}^+ = \max\{\varepsilon_{t-i}, 0\}$ и $\varepsilon_{t-i}^- = \varepsilon_{t-i} - \varepsilon_{t-i}^+$. Полагая $\alpha_i^+ = \alpha_i(1 - \gamma_i)$ и $\alpha_i^- = \alpha_i(1 + \gamma_i)$, получаем

$$\sigma_t = \omega + \sum_{i=1}^q \alpha_i^+ (1 - \gamma_i) \varepsilon_{t-i}^+ - \sum_{i=1}^q \alpha_i^- (1 + \gamma_i) \varepsilon_{t-i}^-.$$

При $\beta_j \neq 0$, $j = 1, \dots, q$, можно получить более общий класс TARСН-моделей.

6. NGARCH-модель из Higgins & Bera (1992), при $\gamma_i = 0, i = 1, \dots, p$.

Хентшель (Hentschel, 1995) вводит следующую модель:

$$\frac{\sigma_t^\lambda - 1}{\lambda} = \omega + \alpha \sigma_{t-1}^\lambda f^\nu(z_{t-1}) + \beta \frac{\sigma_{t-1}^\lambda - 1}{\lambda}, \quad (59)$$

где $\lambda > 0$ и $f^\nu(z_t) = |z_t - b| - c(z_t - b)$. Параметр λ определяет форму преобразования. При $\lambda > 1$ преобразование σ_t выпукло, а при $\lambda < 1$ вогнуто. Параметр ν служит для преобразования функции абсолютного значения $f(\cdot)$. Частные случаи получаются путем соответствующего выбора параметров λ, ν, b и c . Преобразование $f^\nu(\cdot)$ контролирует воздействие шоков, ε_t , на трансформированное условное стандартное отклонение σ_t . Преобразование типа Бокса-Кокса для σ_t делает возможным вложение большого числа GARCH-моделей, позволяя $\lambda \rightarrow 0$. В частности, при $\lambda = \nu = 1$ модель принимает вид

$$\sigma_t = \tilde{\omega} + \alpha \sigma_{t-1} [|z_{t-1} - b| - c(z_{t-1} - b)] + \beta \sigma_{t-1}, \quad (60)$$

который Хентшель назвал моделью GARCH для абсолютных значений. Положительное значение b вызывает сдвиг КВН вправо и позволяет получить асимметрию. Но асимметрию можно также получить поворотом КВН. Разрешая наклоны различной величины с каждой стороны от начала координат, КВН такого типа также допускает асимметричную реакцию дисперсии. Поворот задается параметром c : $c > 0$ соответствует повороту по часовой стрелке, который означает, что отрицательные шоки увеличивают волатильность больше, чем положительные. Модели EGARCH, GJR-GARCH и пороговая модель из Zakoïan (1994) имеют повернутую КВН. Когда и сдвиг и поворот присутствуют в одной КВН, они могут либо усиливать, либо компенсировать друг друга. Среди моделей (первого порядка), вложенных в модель Хентшеля, имеются следующие:

1. EGARCH-модель при $\lambda = 0, \nu = 1$, произвольных b и c , используя тот факт, что $\lim_{\lambda \rightarrow 0} (\sigma_t^\lambda - 1)/\lambda = \log \sigma_t$.
2. GJR-GARCH-модель в (39) при $\lambda = \nu = 2, b = 0$ и свободном c .
3. TGARCH-модель в (48) при $\lambda = \nu = 1$ и $b = 0$.
4. GARCH-модель в (9) при $\lambda = \nu = 2$ и $b = c = 0$.
5. NAGARCH-модель в (42) при $\lambda = \nu = 2$ и $c = 0$. В этой модели КВН сдвинута вправо на расстояние b .
6. APGARCH-модель в (58) при $\lambda = \nu, b = 0$
7. Нелинейная GARCH в (57) при $\lambda = \nu, b = c = 0$.

Модели GARCH, NAGARCH и GJR-GARCH различаются только ограничениями на b и c .

Расширенная GARCH-модель из Duan (1997) основана на вспомогательном процессе, который можно вольно интерпретировать как преобразование Бокса-Кокса условной дисперсии. Вспомогательный авторегрессионный процесс определяется в терминах некоторого преобразования ошибок. Расширенная GARCH(1, 1)-модель имеет вид

$$\varepsilon_t = z_t \sigma_t, \quad z_t \sim D(0, 1), \quad (61)$$

$$\phi_t = \omega + \zeta_{1,t} \phi_{t-1} + \zeta_{2,t-1}, \quad (62)$$

$$\sigma_t^2 = \begin{cases} |\lambda \phi_t - \lambda - 1|, & \text{при } \lambda \neq 0, \\ \exp\{\phi_t - 1\}, & \text{при } \lambda = 0, \end{cases} \quad (63)$$

где $(\zeta_{1,t}, \zeta_{2,t})$ – строго стационарная последовательность случайных векторов с непрерывным распределением, измеримым относительно информации, доступной на момент времени t . Предположение о непрерывности распределения важно, поскольку оно исключает возможность присвоения положительной вероятностной массы точке взрыва, то есть $\lambda\phi_t - \lambda + 1 = 0$ при $\lambda < 0$. Процесс ε_t является расширенным GARCH(1, 1)-процессом, если он удовлетворяет уравнениям (61), (62) и (63), и

$$\zeta_{1,t} = \alpha_1 + \alpha_2|z_t - c|^\delta + \alpha_3 \max(0, c - z_t)^\delta, \quad (64)$$

$$\zeta_{2,t} = \alpha_4 \frac{|z_t - c|^\delta - 1}{\delta} + \alpha_5 \frac{\max(0, c - z_t)^\delta - 1}{\delta}. \quad (65)$$

Расширенная GARCH(1, 1)-модель включает в себя как частные случаи несколько видов GARCH-моделей, таких как GARCH(1, 1), EGARCH(1, 1) и GJR-GARCH(1, 1). Этот процесс строго стационарен при простом достаточном условии.

He & Teräsvirta (1999) определили еще одно семейство GARCH-моделей:

$$\sigma_t^{2\delta} = \sum_{i=1}^q g(z_{t-i}) + \sum_{i=1}^p c_i(z_{t-i})\sigma_{t-i}^{2\delta}, \quad \delta > 0, \quad (66)$$

где $\{g(z_{t-i})\}$ и $\{c_i(z_{t-i})\}$ – последовательности IID случайных величин. Стандартная GARCH (p, q) -модель получается при $g(z_t) = \omega/q$ и $c(z_{t-i}) = \alpha_i z_{t-i}^2 + \beta_i, i = 1, \dots, q$, в (66). Модель GJR-GARCH, GARCH для абсолютных значений, квадратичная GARCH и степенная GARCH принадлежат этому классу.

8 GARCH в среднем

GARCH-модели часто используют для прогнозирования риска портфеля в определенный момент времени. Engle et al. (1987) рассматривают экономику, в которой не склонные к риску агенты выбирают из двух видов финансовых вложений, чтобы максимизировать свою ожидаемую полезность. Первая возможность представлена рискован активом с нормально распределенной доходностью, для которого риск измеряется дисперсией доходности, а компенсация – ростом ожидаемой доходности. Второй тип инвестиций представлен безрисковым активом. Максимизация агентами своей функции полезности при условиях баланса на рынках приводит к стандартной связи между матожиданием и дисперсией доходности рискованного актива. Engle et al. (1987) исследуют эту взаимосвязь в случае, когда дисперсия доходности рискованного актива меняется во времени, а следовательно, меняется и цена рискованного актива. Приведенные предположения определяют взаимосвязь между средним и дисперсией доходности актива, которая по-прежнему положительна, но не постоянна. Из этого следует, что модель условной дисперсии наподобие GARCH может быть полезным представлением меняющейся во времени рискованной премии при объяснении избыточной доходности. Избыточная доходность в таком случае является комбинацией непредсказуемой разности ε_t между ожидаемой и реализованной нормами доходности и функции условной дисперсии портфеля (Teräsvirta, 2009). Таким образом, если y_t – избыточная доходность на момент t , модель GARCH в среднем (GARCH-M), предложенная в Engle et al. (1987), представляет собой систему

$$\begin{aligned} y_t &= \gamma + g(\sigma_t^2) - \mathbb{E}[g(\sigma_t^2)] + \varepsilon_t, \\ \sigma_t^2 &= \omega + \sum_{i=1}^q \alpha_i \varepsilon_{t-i}^2 + \sum_{i=1}^p \beta_i \sigma_{t-i}^2. \end{aligned} \quad (67)$$

Эта модель одновременно характеризует динамику среднего и дисперсии временного ряда.

Engle et al. (1987) использовали функцию $g(\sigma_t^2) = \delta\sqrt{\sigma_t^2}$, но альтернативные варианты $g(\sigma_t^2) = \ln(\sigma_t^2)$ и $g(\sigma_t^2) = \sigma_t^2$ также предлагались в литературе. Процесс GARCH-M обладает интересной структурой моментов. Предположим, что $\mathbb{E}[z_t^3] = 0$ и $\mathbb{E}[\varepsilon_t^4] < \infty$. Из уравнения (67) следует, что

$$\mathbb{E}[(y_t - \mathbb{E}[y_t])(y_{t-k} - \mathbb{E}[y_t])] = \mathbb{E}[\varepsilon_{t-k}g(\sigma_t^2)] + \mathbb{C}[g(\sigma_t^2), g(\sigma_{t-k}^2)].$$

Это означает, что в y_t присутствует прогнозируемая компонента. Более того,

$$\mathbb{E}[(y_t - \mathbb{E}[y_t])^3] = 3\mathbb{E}[\sigma_t^2(g(\sigma_t^2) - \mathbb{E}[g(\sigma_t^2)])] + \mathbb{E}[(g(\sigma_t^2) - \mathbb{E}[g(\sigma_t^2)])^3] \neq 0.$$

Отсюда следует, что модель GARCH-M подразумевает скошенное частное распределение y_t , если только $g(\sigma_t^2)$ не константа. При $g(\sigma_t^2) = \delta\sigma_t$ и $\delta < 0$ это частное распределение отрицательно скошено.

9 GARCH-модели с долгой памятью

На практике часто случается, что оценка суммы параметров GARCH(1,1)-модели, $\hat{\alpha}_1 + \hat{\beta}_1$, близка к единице. Другими словами, оказывается, что модель IGARCH(1,1) является разумным приближением процесса, порождающего данные. Однако применение IGARCH(1,1) означает предположение о том, что безусловная дисперсия моделируемого процесса не существует. Более того, в модели IGARCH воздействие шока на оптимальный прогноз будущей условной дисперсии приводит к тому, что соответствующие накопленные веса импульсного отклика сходятся к ненулевой константе, так что прогнозы линейно возрастают по мере расширения горизонта прогнозирования. Это означает, что ценообразование рискованных ценных бумаг, включая долгосрочные опционы и фьючерсы, может крайне сильно зависеть от начальных условий или текущего состояния экономики. Но такая сильная зависимость не соответствует наблюдаемому поведению цен. Taylor (1986), Dacorogna et al. (1993), Ding et al. (1993) указывают на явные признаки долгой памяти в рядах эмпирических автокорреляций для абсолютных значений и квадратов доходностей различных финансовых активов. Эти исследования четко свидетельствуют в пользу моделей с автоковариациями, медленно убывающими по лагу как функция $k^{-\gamma}$, для некоторого $0 < \gamma < 1$.

Хотя эмпирический анализ свидетельствует о том, что долгая память – это свойство квадратов или абсолютных значений доходностей, нет согласия по поводу возможных причин этого стилизованного факта. Granger & Ding (1996) показали, что текущее агрегирование стабильных GARCH(1,1)-процессов дает агрегированный процесс с гиперболически убывающими автокорреляциями. Хотя это свойство вроде бы согласуется с долгой памятью, Zaffaroni (2007) показал, что автокорреляционная функция этого процесса суммируема, что не позволяет классифицировать его как процесс с долгой памятью. Andersen & Bollerslev (1997) показывают, что текущее агрегирование потоковых процессов со слабо зависимой информацией может привести к долгой памяти в рядах волатильности. Еще одно обоснование приводят Müller et al. (1997), предполагая, что долгая память в рядах волатильности может возникать из-за реакции краткосрочных дилеров на динамику прокси-переменной для тренда ожидаемой волатильности (грубой волатильности), что ведет к устойчивости процессов волатильности для более высоких частот (очищенной волатильности).

Если приведенные выше статьи исследовали причины долгой памяти в рядах волатильности, то другие в сущности ставили под сомнение само наличие долгой памяти. В частности, один из аргументов состоит в том, что структурные сдвиги различных типов способны объяснить крайнюю устойчивость волатильности, а также могут порождать ряд, который выглядит как ряд с долгой памятью. Mikosch & Starica (1998) и Granger & Hyung (2004) приводят теорию и симуляции в подтверждение того, что ложные свидетельства долгой памяти можно обнаружить во временном ряде со сдвигами. Более того, хотя Granger & Hyung

(2004) показали для ряда абсолютных значений доходности индекса S&P500, что модель с нерегулярными сдвигами обладает худшей предсказательной силой, чем модель с долгой памятью; для того же ряда Starica & Granger (2005) получили, что нестационарная модель, позволяющая сдвиги в безусловной дисперсии, превосходит модель с долгой памятью при прогнозировании, но не на коротком горизонте. Diebold & Inoue (2001) также показали, как процессы с марковскими переключениями могут порождать долгую память в условном среднем, а Granger & Teräsvirta (1999) показали, что процесс с переключениями знака обладает признаками процесса с долгой памятью. Возможное наличие структурных сдвигов в условной дисперсии как причину крайней устойчивости в форме IGARCH изначально предложили Lamoureaux & Lastrapes (1990) и Diebold (1986).

Долгосрочная зависимость, или долгая память (иногда также называемые строгой зависимостью или устойчивостью), означает устойчивое поведение временных рядов. Концепция долгой памяти впервые была разработана в Hurst (1951). Долгую память можно определить в терминах скорости убывания автокорреляций для больших лагов. В частности, стационарный процесс обладает долгой памятью (или долгосрочной зависимостью), если существует вещественное число d и константа $c_\rho > 0$, такие, что

$$\lim_{k \rightarrow \infty} \frac{\rho(k)}{c_\rho k^{2d-1}} = 1,$$

где $\rho(k)$ – автокорреляция для лага k , а d – параметр долгой памяти. Автокорреляции процесса с долгой памятью несуммируемы. Альтернативное, хотя и не эквивалентное, определение долгосрочной зависимости можно дать, используя спектральную плотность $f(\lambda)$ процесса:

$$\lim_{\lambda \rightarrow 0^+} \frac{f(\lambda)}{c_f |\lambda|^{-2d}} = 1, \quad 0 < c_f < \infty.$$

Спектральная плотность $f(\lambda)$ имеет полюс и в начале координат ведет себя как константа c_f , помноженная на λ^{-2d} . Автокорреляции стационарного и обратимого ARMA-процесса геометрически ограничены, то есть $|\rho(k)| \leq cm^{-k}$, где $0 < m < 1$, и следовательно, это процесс с короткой памятью. Популярный подход к моделированию долгой памяти состоит в использовании класса моделей ARFIMA, предложенного в Granger & Joyeux (1980) и Hosking (1981). Он обобщает ARIMA-модели, позволяя дробную степень интегрированности. В классе $I(d)$ дробно интегрированных дискретных процессов с долгой памятью в дискретном времени распространение шоков среднего происходит с медленной гиперболической скоростью убывания, по сравнению с крайними случаями: классом стационарных и обратимых $I(0)$ процессов ARMA с экспоненциальным убыванием и классом $I(1)$ -процессов с бесконечной устойчивостью.

Класс моделей ARFIMA(k, d, l) для дискретного процесса $\{y_t\}$ с вещественными значениями имеет вид (см. Granger & Joyeux, 1980, и Hosking, 1981)

$$a(L)(1-L)^d y_t = b(L)\varepsilon_t, \quad (68)$$

где $a(L)$ и $b(L)$ – полиномы от L порядков k и l , соответственно, а $\{\varepsilon_t\}$ – серийно некоррелированный процесс с нулевым средним. Все корни $a(L)$ и $b(L)$ лежат вне единичного круга. Процесс определяется как $I(d)$. Оператор дробной разности, $(1-L)^d$, имеет биномиальное разложение. Используя разложение Тейлора-Маклорена в точке $z = 0$, получаем

$$(1-z)^d = 1 - dL + \frac{d(d-1)}{2!}L^2 + \dots = \sum_{j=0}^{\infty} \binom{d}{j} (-1)^j z^j, \quad \binom{d}{j} = \frac{d!}{(d-j)!j!}.$$

Коэффициенты при L^j убывают медленно и асимптотически пропорциональны $j^{-(1+d)}$. Учитывая, что $\Gamma(j+1) = j! = j\Gamma(j)$, где $\Gamma(j) = \int_0^\infty x^{j-1} \exp(-x) dx$, $j > 0$, – гамма-функция, можно записать

$$\binom{d}{j} = \frac{d!}{(d-j)!j!} = \frac{\Gamma(j-d)}{\Gamma(-d)\Gamma(j+1)}.$$

Тогда получаем

$$(1-L)^d = \sum_{j=0}^{\infty} \binom{d}{j} (-1)^j z^j = \sum_{j=0}^{\infty} \frac{\Gamma(j-d)}{\Gamma(-d)\Gamma(j+1)} L^j = \sum_{j=0}^{\infty} \phi_j L^j = F(-d, 1, 1; L),$$

где $F(-d, 1, 1; L)$ – гипергеометрическая функция, определяемая как

$$F(m, n, s; x) = \Gamma(s)\Gamma(m)^{-1}\Gamma(n)^{-1} \sum_{j=0}^{\infty} \Gamma(m+j)\Gamma(n+j)\Gamma(s+j)^{-1}\Gamma(j+1)^{-1} x^j.$$

Если $\mathbb{V}[\varepsilon_t] < \infty$ и $-1/2 < d < 1/2$, процесс $\{y_t\}$ является слабо стационарным и обратимым и единственным образом представим в виде скользящего среднего и авторегрессии бесконечных порядков. При $d < 1$ процесс возвращается к среднему. При $1/2 < d < 1$, y_t не является ковариационно стационарным, но возвращается к среднему. Модель ARFIMA разделяет краткосрочную и долгосрочную динамику, улавливая краткосрочное поведение с помощью обыкновенных лаговых полиномов ARMA, $a(L)$ и $b(L)$, а долгосрочное – параметром дробной интегрированности d .

9.1 Модель FIGARCH

Процесс FIGARCH, предложенный в Baillie et al. (1996), объединяет многие свойства дробно интегрированных процессов для среднего с обыкновенным GARCH-процессом для условной дисперсии. Модели FIGARCH свойственна медленная гиперболическая скорость убывания лагированных квадратов инноваций в условной дисперсии, хотя накопленные веса импульсных откликов, связанных с воздействием шока волатильности на оптимальные прогнозы будущей условной дисперсии, сходятся к нулю. Это свойство модель разделяет со слабо стационарными GARCH-процессами. Модель FIGARCH предлагает альтернативу подходу, рассматривающему изменения параметров GARCH-модели как главную причину медленного убывания автокорреляций. Тем не менее, свойство долгой памяти (и даже существование стационарного режима) для модели FIGARCH не доказано (см. Giraitis et al., 2000). GARCH(p, q)-процесс как ARMA(m, p) для ε_t^2 , где $m = \max(p, q)$,

$$(1 - \alpha(L) - \beta(L))\varepsilon_t^2 = \omega + (1 - \beta(L))v_t, \quad v_t = \varepsilon_t - \sigma_t^2, \quad (69)$$

является последовательностью мартингальных приращений. В ковариационно стационарной GARCH-модели эффект квадратов прошлых инноваций на текущую условную дисперсию экспоненциально убывает по величине лага. Если авторегрессионный полином $1 - \alpha(x) - \beta(x)$ содержит единичный корень, получаем модель IGARCH(p, q), определяемую как

$$\phi(L)(1-L)\varepsilon_t^2 = \omega + (1 - \beta(L))v_t, \quad (70)$$

где $\phi(L) \equiv (1 - \alpha(L) - \beta(L))(1-L)^{-1}$ имеет порядок $m-1$, $m = \max(p, q)$. Модель FIGARCH получается простой заменой оператора $(1-L)$ оператором дробной разности. Аналогично ARFIMA(k, d, l)-процессу для среднего, FIGARCH(p, d, q) для $\{\varepsilon_t\}$ естественно определяется как

$$\phi(L)(1-L)^d \varepsilon_t^2 = \omega + (1 - \beta(L))v_t, \quad (71)$$

где $0 < d < 1$, и все корни $\phi(L)$ и $(1 - \beta(L))$ лежат вне единичного круга. Это означает, что процесс для $\{\varepsilon_t^2\}$ является ARFIMA($m - 1, d, p$). Учитывая, что $v_t \equiv \varepsilon_t^2 - \sigma_t^2$, альтернативное представление модели FIGARCH(p, d, q) имеет вид

$$(1 - \beta(L))\sigma_t^2 = \omega + (1 - \beta(L) - \phi(L)(1 - L)^d)\varepsilon_t^2, \quad (72)$$

$$\sigma_t^2 = (1 - \beta(1))^{-1}\omega + \lambda(L)\varepsilon_t^2, \quad (73)$$

где

$$\lambda(L) \equiv 1 - (1 - \beta(L))^{-1}\phi(L)(1 - L)^d \quad (74)$$

и $\lambda(L) = \lambda_1 L + \lambda_2 L^2 + \lambda_3 L^3 + \dots$ FIGARCH(p, d, q)-модель корректно определена, и $\sigma_t^2 > 0$ п.н. для любого t , если все коэффициенты бесконечного ARCH-представления неотрицательны, то есть $\lambda_k > 0$ для $k = 0, 1, 2, \dots$. При $0 < d \leq 1$ $F(-d, 1, 1; L = 1) = 0$, так что $\lambda(1) = 1$. Следовательно, второй момент безусловного распределения ε_t бесконечен, и FIGARCH-процесс не является слабо стационарным. Это свойство он делит с классом IGARCH-процессов. Baillie et al. (1996) утверждают, что, поскольку лаги высоких порядков в бесконечном ARCH-представлении любой FIGARCH-модели могут доминироваться в смысле абсолютных значений соответствующими IGARCH-коэффициентами, класс FIGARCH(p, d, q) строго стационарен и эргодичен для $0 \leq d \leq 1$.

Однако доказательство существования стационарного решения (73), приведенное в Baillie et al. (1996), по всей видимости, неверно (см. Giraitis et al., 2007). Вопрос существования стационарного решения уравнения FIGARCH («проблема FIGARCH») остается открытым и весьма сложным. Giraitis et al. (2000), Mikosch & Starica (2000, 2003) обсуждают споры вокруг FIGARCH-модели.

Модель FIGARCH(p, d, q) включает ковариационно стационарную GARCH(p, q)-модель при $d = 0$ и IGARCH(p, q)-модель при $d = 1$. Если позволить параметру d принимать значения в интервале от нуля до единицы, модель получает дополнительную гибкость, что важно при моделировании долговременной зависимости в условной дисперсии.

Следует проявлять осторожность при интерпретации устойчивости в нелинейных моделях. В случае, когда условная дисперсия является линейной функцией прошлых ε_t^2 , устойчивость просто характеризовать в терминах коэффициентов импульсного отклика для оптимального прогноза будущей условной дисперсии как функции от инновации в момент времени t , v_t :

$$\gamma_k \equiv \frac{\partial \mathbb{E}_t[\varepsilon_{t+k}^2]}{\partial v_t} - \frac{\partial \mathbb{E}_t[\varepsilon_{t+k-1}^2]}{\partial v_t}.$$

В более общих моделях условной дисперсии γ_i зависят от информационного множества на момент t . В моделях FIGARCH коэффициенты импульсного отклика не зависят от t . Их можно определить по коэффициентам лагового полинома $\gamma(L)$:

$$(1 - L)\varepsilon_t^2 = (1 - L)^{1-d}\phi(L)^{-1}\omega + (1 - L)^{1-d}\phi(L)^{-1}(1 - \beta(L))v_t \equiv \zeta + \gamma(L)v_t \quad (75)$$

Долгосрочное воздействие прошлых шоков на процесс волатильности теперь можно оценить в терминах предела накопленных весов импульсного отклика. Действительно,

$$\sum_{i=0}^k \gamma_i = \sum_{i=0}^k \left[\frac{\partial \mathbb{E}_t[\varepsilon_{t+i}^2]}{\partial v_t} - \frac{\mathbb{E}_t[\varepsilon_{t+i-1}^2]}{\partial v_t} \right] = \frac{\partial \mathbb{E}_t[\varepsilon_{t+k}^2]}{\partial v_t}, \quad (76)$$

а

$$\lambda_k = \frac{\partial \mathbb{E}_t[\varepsilon_{t+k}^2]}{\partial v_t}.$$

Тогда

$$\gamma(1) = \lim_{k \rightarrow \infty} \sum_{i=0}^k \gamma_i = \lim_{k \rightarrow \infty} \lambda_k = F(d-1, 1, 1; 1) \phi(1)^{-1} (1 - \beta(1)). \quad (77)$$

В ковариационно стационарной GARCH(p, q)-модели и FIGARCH(p, d, q)-модели с $0 < d < 1$ шоки в условной дисперсии в конечном счете исчезают, с точки зрения прогнозирования. В то время как шоки в GARCH-процессах ($d = 0$) убывают с быстрой экспоненциальной скоростью, в FIGARCH-модели λ_k доминируется гиперболической скоростью убывания. Таким образом, несмотря на то, что кумулятивная функция импульсного отклика сходится к нулю при $0 \leq d < 1$, параметр дробной разности несет важную информацию относительно того, как и с какой скоростью распространяются шоки волатильности. С другой стороны, при $d = 1$ $F(d-1, 1, 1; 1) = 1$, и накопленные веса импульсных откликов сходятся к не равной нулю константе $\gamma(1) = \phi(1)^{-1} (1 - \beta(1))$. Поэтому с точки зрения прогнозирования, шоки σ_t^2 в модели IGARCH сохраняются навсегда. При $d > 1$ $F(d-1, 1, 1; 1) = \infty$, что ведет к не наблюдаемому в данных взрывному процессу для условной дисперсии, и $\gamma(1)$ становится не определена.

Рассмотрим в качестве примера модель GARCH(1, 1)

$$\sigma_t^2 = \omega + \alpha_1 \varepsilon_{t-1}^2 + \beta \sigma_{t-1}^2$$

в форме ARMA:

$$(1 - \phi_1 L) \varepsilon_t^2 = \omega + (1 - \beta_1 L) v_t,$$

где $\phi_1 \equiv \alpha_1 + \beta_1$. Веса импульсных откликов для этой модели являются коэффициентами полинома

$$\gamma(L) = (1 - L)(1 - \phi_1 L)^{-1} (1 - \beta_1 L),$$

то есть

$$\begin{aligned} \gamma_0 &= 1, \\ \gamma_1 &= \phi_1 - \beta_1 - 1, \\ \gamma_k &= (\phi_1 - \beta_1)(\phi_1 - 1) \phi_1^{k-2}, \quad k > 2. \end{aligned}$$

Накопленные веса импульсных откликов равны

$$\lambda_k = (\phi_1 - \beta_1) \phi_1^{k-1}, \quad k > 1,$$

и в пределе $\gamma(1) = 0$, если $0 < \phi_1 < 1$. Следовательно, эффект шока на прогноз будущей условной дисперсии сходится к нулю с быстрой экспоненциальной скоростью. В модели IGARCH(1, 1), то есть при $\phi_1 = 1$,

$$(1 - L) \varepsilon_t^2 = \omega + (1 - \beta_1 L) v_t,$$

коэффициенты $\lambda(L)$ имеют вид

$$\lambda_k = (1 - \beta_1), \quad \forall k > 1,$$

и накопленные веса импульсных откликов равны ненулевой константе $\gamma(1) = 1 - \beta_1$. Модель FIGARCH(1, d , 0) имеет вид

$$(1 - \phi_1 L)^d \varepsilon_t^2 = \omega + (1 - \beta_1 L) v_t.$$

Можно показать, что накопленные коэффициенты отклика в ARCH-представлении бесконечного порядка для FIGARCH(1, d , 0)-модели

$$\lambda(L) \equiv 1 - (1 - \beta_1 L)^{-1} (1 - L)^d$$

равны

$$\lambda_k = (1 - \beta_1 - (1 - d)k^{-1})\Gamma(k + d - 1)\Gamma(k)^{-1}\Gamma(d)^{-1}$$

при $k > 1$ и $\lambda_0 = 1$. Поэтому, если $\omega > 0$, условие $0 \leq \beta_1 < d \leq 1$ является необходимым и достаточным для гарантии, что условная дисперсия в FIGARCH(1, d , 0)-модели положительна п.н. для всех t . Более того, из формулы Стирлинга следует, что для больших лагов k

$$\lambda_k = ((1 - \beta_1)\Gamma(d)^{-1})k^{d-1}.$$

В отличие от ковариационно стационарной GARCH(1, 1)-модели или модели IGARCH(1, 1), в которых воздействие шоков условной дисперсии либо экспоненциально убывает, либо сохраняется бесконечно, в модели FIGARCH(1, d , 0) реакция условной дисперсии на прошлые шоки убывает медленно, гиперболически.

9.2 FIEGARCH

Поскольку оценки стандартной GARCH(p, q)-модели часто указывают на близкий к единичному корень в авторегрессионном полиноме, при оценивании модели EGARCH(p, q), имеющей вид (37), наибольший корень оцененного полинома $1 - \hat{\beta}(x)$ очень близок к единице. Тем не менее, как заметил Nelson (1991), модель EGARCH(p, q) также можно расширить, позволив дробный порядок интегрированности. Bollerslev & Mikkelsen (1996) разлагают на множители авторегрессионный полином $(1 - \beta(L)) = \phi(L)(1 - L)^d$, где все корни $\phi(x) = 0$ лежат вне единичного круга, чтобы записать модель в виде

$$\ln(\sigma_t^2) = \omega + \phi(L)^{-1}(1 - L)^{-d}(1 + \alpha(L))g(z_{t-1}), \quad (78)$$

где $\alpha(L) = \alpha_1 L + \alpha_2 L^2 + \dots + \alpha_q L^q$. Очевидно, модель FIEGARCH(p, d, q) включает в себя обычную EGARCH-модель при $d = 0$ и интегрированную EGARCH-модель при $d = 1$. По аналогии с классом ARFIMA моделей условного среднего, $\{\ln(\sigma_t^2)\}$ является ковариационно стационарным и обратимым при d в интервале от $-1/2$ до $1/2$. Шоки на оптимальные прогнозы будущих значений $\ln(\sigma_t^2)$ исчезают при всех $d < 1$. Более того, в отличие от формулировки FIGARCH, параметры FIEGARCH необязательно должны удовлетворять каким-либо ограничениям неотрицательности, чтобы модель была корректно поставлена.

10 Процедуры оценивания

По сравнению с другими моделями волатильности (например, стандартной моделью стохастической волатильности), GARCH-модели легче оценивать, что сильно способствовало их популярности. Поскольку волатильность является функцией прошлых наблюдений, функция правдоподобия имеет явный вид и с ней просто работать (см. Francq & Zakoïan, 2009). Оценивание ARCH-моделей методом наименьших квадратов и максимального квазиправдоподобия рассмотрено в пионерской статье Энгла (Engle, 1982). Асимптотические свойства оценки методом максимального квазиправдоподобия (ММКП) получили широкий интерес за последние 20 лет. Первые статьи ограничивались моделями ARCH (см. Weiss, 1986) или GARCH(1, 1) (Lee & Hansen, 1994, Lumsdaine, 1996). ММКП-оценивание общей GARCH(p, q)-модели изучено в Berkes & Horváth (2003, 2004), Berkes, Horváth & Kokosza (2003), Francq & Zakoïan (2004) и Hall & Yao (2003). Straumann (2005) является обширной монографией об оценивании GARCH-моделей.

10.1 Подсчет функции правдоподобия

Процедура, наиболее часто используемая для оценивания $\theta_0 \in \Theta$ в ARCH-моделях заключается в максимизации функции правдоподобия при дополнительном предположении о распределении стандартизированных инноваций $z_t(\theta)$. Пусть $f(z_t(\theta); \eta)$ обозначает функцию плотности $z_t(\theta) \equiv \varepsilon_t(\theta) / \sigma_t(\theta)$, с нулевым средним и единичной дисперсией, где η – мешающий параметр, $\eta \in H \subseteq R^k$. Пусть $(y_T, y_{T-1}, \dots, y_1)$ – реализация ARCH-модели, определенной уравнениями с (1) по (5), а $\psi' \equiv (\theta', \eta')$ – объединенный $(m+k) \times 1$ -вектор оцениваемых параметров условного среднего, дисперсии и функций плотности. Логарифмическая функция правдоподобия для t -го наблюдения имеет вид

$$l_t(y_t; \psi) = \ln(f(z_t(\theta); \eta)) - \frac{1}{2} \ln(\sigma_t^2(\theta)), \quad t = 1, 2, \dots \quad (79)$$

Член $-\frac{1}{2} \ln(\sigma_t^2(\theta))$ в правой части – это якобиан преобразования стандартизированных инноваций, $z_t(\theta)$, в наблюдаемые y_t : $f(y_t; \psi) = f(z_t(\theta); \eta) |J|$, где $J = \partial z_t / \partial y_t = 1 / \sigma_t(\theta)$. Условная логарифмическая функция правдоподобия для всей выборки равна сумме условных логарифмических правдоподобий в (79):

$$L_T(y_T, y_{T-1}, \dots, y_1; \psi) = \sum_{t=1}^T l_t(y_t; \psi). \quad (80)$$

ММП-оценка истинных параметров $\psi'_0 \equiv (\theta'_0, \eta'_0)$, скажем, $\widehat{\psi}_T$, получается из максимизации функции (80). Предполагая дифференцируемость функции условной плотности, $\mu_t(\theta)$, и $\sigma_t^2(\theta)$ при всех $\psi \in \Theta \times H \equiv \Psi$, ММП-оценка является решением уравнения

$$S_T(y_T, y_{T-1}, \dots, y_1; \widehat{\psi}) \equiv \sum_{t=1}^T s_t(y_t; \widehat{\psi}) = 0, \quad (81)$$

где $s_t \equiv \frac{\partial l_t(y_t, \psi)}{\partial \psi}$ – скор-вектор для t -го наблюдения. В частности, для параметров условного среднего и дисперсии

$$\frac{\partial l_t(y_t, \psi)}{\partial \theta} = f(z_t(\theta); \eta)^{-1} f'(z_t(\theta); \eta) \frac{\partial z_t(\theta)}{\partial \theta} - \frac{1}{2} (\sigma_t^2(\theta))^{-1} \frac{\partial \sigma_t^2}{\partial \theta}, \quad (82)$$

где $f'(z_t(\theta); \eta) \equiv \frac{\partial f(z_t(\theta); \eta)}{\partial z_t}$ и

$$\frac{\partial z_t(\theta)}{\partial \theta} = \frac{\partial}{\partial \theta} \left(\frac{\varepsilon_t(\theta)}{\sqrt{\sigma_t^2}} \right) = -\frac{\partial \mu_t}{\partial \theta} (\sigma_t^2(\theta))^{-1/2} - \frac{1}{2} (\sigma_t^2(\theta))^{-3/2} \frac{\partial \sigma_t^2}{\partial \theta} \varepsilon_t(\theta),$$

где $\varepsilon_t(\theta) \equiv y_t - \mu_t(\theta)$. На практике решение системы $m+k$ нелинейных уравнений в (81) ищется с помощью численных методов оптимизации.

Для осуществления процедуры ММП-оценивания необходимо сделать явное предположение об условной плотности (79). Как правило, в литературе используют нормальное распределение:

$$f(z_t(\theta); \eta) = (2\pi)^{-1/2} \exp \left\{ -\frac{z_t(\theta)^2}{2} \right\}.$$

Поскольку нормальное распределение однозначно задается своими двумя моментами, только параметры условного среднего и дисперсии входят в логарифмическую функцию правдоподобия в уравнении (80), то есть $\psi = \theta$. Логарифмическое правдоподобие для t -го наблюдения имеет вид

$$l_t(\theta) = -\frac{1}{2} \ln(2\pi) - \frac{1}{2} z_t(\theta)^2 - \frac{1}{2} \ln(\sigma_t^2),$$

а следовательно, скор-вектор из уравнения (82) равен

$$\begin{aligned} s_t &= -z_t \frac{\partial z_t}{\partial \theta} - \frac{1}{2} (\sigma_t^2(\theta))^{-1} \frac{\partial (\sigma_t^2(\theta))}{\partial \theta} \\ &= \frac{\partial \mu_t(\theta)}{\partial \theta} \frac{\varepsilon_t(\theta)}{\sigma_t^2(\theta)} + \frac{1}{2} (\sigma_t^2(\theta))^{-1} \frac{\partial \sigma_t^2(\theta)}{\partial \theta} \left(\frac{\varepsilon_t^2(\theta)}{\sigma_t^2(\theta)} - 1 \right). \end{aligned} \quad (83)$$

Как обсуждается в McCullough & Renfro (1999) и Brooks, Burke & Pesand (2001), на практике при максимизации гауссовской логарифмической функции правдоподобия

$$L_T(y_T, y_{T-1}, \dots, y_1; \theta) = -\frac{T}{2} \ln(2\pi) - \frac{1}{2} \sum_{t=1}^T \frac{\varepsilon_t^2(\theta)}{\sigma_t^2(\theta)} - \frac{1}{2} \sum_{t=1}^T \ln(\sigma_t^2(\theta)) \quad (84)$$

возникает несколько практических вопросов. На практике используется условная функция правдоподобия (84), поскольку аналитически безусловное распределение начальных значений неизвестно (Diebold & Schuermann (1993) предложили вычислительно сложную численную процедуру для аппроксимации точной логарифмической функции правдоподобия). Когда $\sigma_t^2(\theta)$ моделируется как GARCH(p, q), необходимо выбрать начальные значения параметров модели ω , α_i , $i = 1, \dots, q$ и β_j , $j = 1, \dots, p$ и инициализировать ε_t^2 и σ_t^2 . Если $\mathbb{E}_{t-1}[y_t] = \mu$, в качестве начального значения для μ обычно используют выборочное среднее y_t , параметры условной дисперсии, помимо ω и α_1 , часто полагают равными нулю, а ω присваивается значение безусловной дисперсии y_t . В качестве начального значения для σ_t^2 возможен вариант

$$\sigma_t^2 = \varepsilon_t^2 = \frac{1}{T} \sum_{s=1}^T \varepsilon_s^2, \quad t \leq 0,$$

где начальные значения для ε_s находятся как остатки регрессии y_t на константу. После инициализации логарифмической функции правдоподобия ее можно максимизировать, используя численные методы оптимизации.

10.2 Оценивание методом максимального квазиправдоподобия

Общие свойства ММКП-оценки, определяемой как $\hat{\theta}_T = \arg \max_{\theta \in \Theta} L_T(\theta)$, где $L_T(\theta)$ задается уравнением (84), зависят как от истинного распределения, так и от используемого при построении функции правдоподобия. При стандартных условиях регулярности (White, 1981, Gourieroux et al., 1984, Gallant, 1987, и Gourieroux, 1997) эта оценка состоятельна, даже если истинное распределение не является условно нормальным:

$$\sqrt{T} (\hat{\theta}_T - \theta_0) \xrightarrow{d} N(0, A_0^{-1} B_0 A_0^{-1}).$$

Матрицы A_0 и B_0 равны, соответственно,

$$\begin{aligned} A_0 &= -\frac{1}{T} \mathbb{E}_0 \left[\frac{\partial^2 \log L(\theta_0)}{\partial \theta \partial \theta'} \right], \\ B_0 &= \frac{1}{T} \mathbb{E}_0 \left[\frac{\partial \log L(\theta_0)}{\partial \theta} \frac{\partial \log L(\theta_0)}{\partial \theta'} \right], \end{aligned}$$

где \mathbb{E}_0 обозначает матожидание относительно истинного распределения.

Можно показать, что $A_0 = B_0$ только в особом случае. Матрица вторых производных логарифмической функции правдоподобия для t -го наблюдения имеет вид

$$\begin{aligned} \frac{\partial^2 l_t}{\partial \theta \partial \theta'} &= \frac{1}{2} (\sigma_t^2(\theta))^{-2} \frac{\partial \sigma_t^2(\theta)}{\partial \theta} \frac{\partial \sigma_t^2(\theta)}{\partial \theta'} - \frac{1}{2} (\sigma_t^2(\theta))^{-1} \frac{\partial^2 \sigma_t^2(\theta)}{\partial \theta \partial \theta'} - \frac{\varepsilon_t^2(\theta)}{(\sigma_t^2(\theta))^3} \frac{\partial \sigma_t^2(\theta)}{\partial \theta} \frac{\partial \sigma_t^2(\theta)}{\partial \theta'} \\ &+ \frac{1}{2} \frac{\varepsilon_t^2(\theta)}{(\sigma_t^2(\theta))^2} \frac{\partial^2 \sigma_t^2(\theta)}{\partial \theta \partial \theta'} - \frac{\partial \mu_t(\theta)}{\partial \theta} \frac{\partial \sigma_t^2(\theta)}{\partial \theta'} \frac{\varepsilon_t(\theta)}{(\sigma_t^2(\theta))^2} + \frac{\varepsilon_t(\theta)}{\sigma_t^2(\theta)} \frac{\partial^2 \mu_t(\theta)}{\partial \theta \partial \theta'} \\ &- (\sigma_t^2(\theta))^{-1} \frac{\partial \mu_t(\theta)}{\partial \theta} \frac{\partial \mu_t(\theta)}{\partial \theta'} - \frac{\varepsilon_t(\theta)}{(\sigma_t^2(\theta))^2} \frac{\partial \sigma_t^2(\theta)}{\partial \theta} \frac{\partial \mu_t(\theta)}{\partial \theta'}. \end{aligned}$$

Поскольку $\mathbb{E}_{t-1} \left[\frac{\varepsilon_t(\theta)}{(\sigma_t^2(\theta))^{1/2}} \right] = 0$ и $\mathbb{E}_{t-1} \left[\frac{\varepsilon_t^2(\theta)}{(\sigma_t^2(\theta))} \right] = 1$, получаем, что

$$A_t = -\mathbb{E}_0 \left[\frac{\partial^2 l_t}{\partial \theta \partial \theta'} \right] = \mathbb{E}_0 \left[\frac{1}{2} (\sigma_t^2(\theta))^{-2} \frac{\partial \sigma_t^2(\theta)}{\partial \theta} \frac{\partial \sigma_t^2(\theta)}{\partial \theta'} + (\sigma_t^2(\theta))^{-1} \frac{\partial \mu_t(\theta)}{\partial \theta} \frac{\partial \mu_t(\theta)}{\partial \theta'} \right].$$

Информационная матрица имеет вид

$$\begin{aligned} B_t = \mathbb{E}_0 \left[\frac{\partial l_t}{\partial \theta} \frac{\partial l_t}{\partial \theta'} \right] &= \mathbb{E}_0 \left[\left(\frac{\partial \mu_t(\theta)}{\partial \theta} \frac{\varepsilon_t(\theta)}{\sigma_t^2(\theta)} + \frac{1}{2} \frac{\partial \sigma_t^2(\theta)}{\partial \theta} \frac{\varepsilon_t^2(\theta)}{(\sigma_t^2(\theta))^2} - \frac{1}{2} \frac{\partial \sigma_t^2(\theta)}{\partial \theta} (\sigma_t^2(\theta))^{-1} \right) \right. \\ &\times \left. \left(\frac{\partial \mu_t(\theta)}{\partial \theta} \frac{\varepsilon_t(\theta)}{\sigma_t^2(\theta)} + \frac{1}{2} \frac{\partial \sigma_t^2(\theta)}{\partial \theta} \frac{\varepsilon_t^2(\theta)}{(\sigma_t^2(\theta))^2} - \frac{1}{2} \frac{\partial \sigma_t^2(\theta)}{\partial \theta} (\sigma_t^2(\theta))^{-1} \right)' \right], \end{aligned}$$

что можно записать в виде

$$\begin{aligned} B_t &= \mathbb{E}_0 \left[\frac{1}{4} \frac{1}{(\sigma_t^2(\theta))^2} \frac{\partial \sigma_t^2(\theta)}{\partial \theta} \frac{\partial \sigma_t^2(\theta)}{\partial \theta'} (K_t(\theta) - 1) + (\sigma_t^2(\theta))^{-1} \frac{\partial \mu_t(\theta)}{\partial \theta} \frac{\partial \mu_t(\theta)}{\partial \theta'} \right] \\ &+ \mathbb{E}_0 \left[\frac{1}{2} \frac{1}{(\sigma_t^2(\theta))^3} \left(\frac{\partial \sigma_t^2(\theta)}{\partial \theta} \frac{\partial \mu_t(\theta)}{\partial \theta'} + \frac{\partial \mu_t(\theta)}{\partial \theta} \frac{\partial \sigma_t^2(\theta)}{\partial \theta'} \right) M_{3t}(\theta) \right] \end{aligned}$$

где $M_{3t}(\theta) = \mathbb{E}_{t-1} [\varepsilon_t^3(\theta)]$ и $K_t(\theta) = (\sigma_t^2(\theta))^{-2} \mathbb{E}_{t-1} [\varepsilon_t^4(\theta)]$. При $\theta = (a', b)'$, где a – параметры условного среднего, а b – параметры условной дисперсии, скор-функция принимает вид (Gourieroux, 1997):

$$s_t = \begin{pmatrix} \partial l_t / \partial a \\ \partial l_t / \partial b \end{pmatrix},$$

где

$$\begin{aligned} \frac{\partial l_t}{\partial a} &= \frac{\partial \mu_t(a)}{\partial a} \frac{\varepsilon_t(a)}{\sigma_t^2(b)}, \\ \frac{\partial l_t}{\partial b} &= \frac{1}{2} (\sigma_t^2(b))^{-1} \frac{\partial \sigma_t^2(b)}{\partial b} \left(\frac{\varepsilon_t^2(a)}{\sigma_t^2(b)} - 1 \right). \end{aligned}$$

В этом случае матрица Гессе для t -го наблюдения равен

$$A_t = \begin{bmatrix} \mathbb{E} \left[(\sigma_t^2(b))^{-1} \frac{\partial \mu_t(a)}{\partial a} \frac{\partial \mu_t(a)}{\partial a'} \right] & 0 \\ 0 & \mathbb{E} \left[\frac{1}{2} (\sigma_t^2(b))^{-2} \frac{\partial \sigma_t^2(b)}{\partial b} \frac{\partial \sigma_t^2(b)}{\partial b'} \right] \end{bmatrix},$$

а информационная матрица имеет вид

$$B_t = \begin{bmatrix} \mathbb{E} \left[(\sigma_t^2(b))^{-1} \frac{\partial \mu_t(a)}{\partial a} \frac{\partial \mu_t(a)}{\partial a'} \right] & \mathbb{E} \left[\frac{1}{2(\sigma_t^2(b))^3} \frac{\partial \mu_t(a)}{\partial a} \frac{\partial \sigma_t^2(b)}{\partial b'} M_{3t}(\theta) \right] \\ \mathbb{E} \left[\frac{1}{2(\sigma_t^2(b))^3} \frac{\partial \sigma_t^2(b)}{\partial b} \frac{\partial \mu_t(a)}{\partial a'} M_{3t}(\theta) \right] & \mathbb{E} \left[\frac{1}{4(\sigma_t^2(b))^2} \frac{\partial \sigma_t^2(b)}{\partial b} \frac{\partial \sigma_t^2(b)}{\partial b'} (K_t(\theta) - 1) \right] \end{bmatrix}.$$

Асимптотические ковариационные матрицы ММКП-оценок \hat{a}_T и \hat{b}_T имеют вид

$$\begin{aligned} \mathbb{V}^{asy} \left[\sqrt{T} (\hat{a}_T - a) \right] &= \left(\frac{1}{T} \sum_t \mathbb{E} \left[(\sigma_t^2(b))^{-1} \frac{\partial \mu_t(a)}{\partial a} \frac{\partial \mu_t(a)}{\partial a'} \right] \right)^{-1}. \\ \mathbb{V}^{asy} \left[\sqrt{T} (\hat{b}_T - b) \right] &= \left(\frac{1}{T} \sum_t \mathbb{E} \left[\frac{1}{2} (\sigma_t^2(b))^{-2} \frac{\partial \sigma_t^2(b)}{\partial b} \frac{\partial \sigma_t^2(b)}{\partial b'} \right] \right)^{-1} \\ &\quad \times \left(\frac{1}{T} \sum_t \mathbb{E} \left[\frac{1}{4(\sigma_t^2(b))^2} \frac{\partial \sigma_t^2(b)}{\partial b} \frac{\partial \sigma_t^2(b)}{\partial b'} (K_t(\theta) - 1) \right] \right) \\ &\quad \times \left(\frac{1}{T} \sum_t \mathbb{E} \left[\frac{1}{2} (\sigma_t^2(b))^{-2} \frac{\partial \sigma_t^2(b)}{\partial b} \frac{\partial \sigma_t^2(b)}{\partial b'} \right] \right)^{-1}. \end{aligned}$$

Если истинное распределение нормально, $M_{3t}(\theta) = 0$ и $K_t(\theta) = 3$, то выражения для A_t и B_t совпадают. Более того, асимптотическая ковариационная матрица ММКП-оценки \hat{b}_T упрощается до

$$\mathbb{V}^{asy} \left[\sqrt{T} (\hat{b}_T - b) \right] = \left(\frac{1}{T} \sum_t \mathbb{E} \left[\frac{1}{2} (\sigma_t^2(b))^{-2} \frac{\partial \sigma_t^2(b)}{\partial b} \frac{\partial \sigma_t^2(b)}{\partial b'} \right] \right)^{-1}.$$

В GARCH-моделях, в которых распределение z_t симметрично, и параметры уравнений для условного среднего и дисперсии неизменны, информационная матрица логарифмического правдоподобия блочно-диагональна. Из этого следует, что параметры уравнения условного среднего можно оценить отдельно от параметров уравнения условной дисперсии без потери асимптотической эффективности. Это может существенно облегчить оценивание. Примером модели, в которой информационная матрица не является блочно-диагональной, является GARCH-M.

10.2.1 Состоятельность

Weiss (1986) впервые изучил асимптотические свойства ММП-оценок в ARCH-модели. Он показал, что ММП-оценки состоятельны и асимптотически нормальны, при условии, что нормализованные данные имеют конечные четвертые моменты. Это требование отсекает IGARCH-модели. Bollerslev & Wooldridge (1992) получили асимптотическое распределение ММКП-оценки при более строгих предположениях об асимптотической нормальности скор-вектора и равномерной слабой сходимости функции правдоподобия и ее вторых производных. Они не проверяют условия и не показывают, как их можно проверить для GARCH-моделей.

Рассмотрим стандартную модель GARCH(p, q), где $y_t = \varepsilon_t$, то есть $\mu_t = 0$, и

$$\begin{aligned} \varepsilon_t &= z_t \sqrt{\sigma_{0t}^2}, \\ \sigma_{0t}^2 &= \omega_0 + \sum_{i=1}^q \alpha_{0i} \varepsilon_{t-i}^2 + \sum_{j=1}^p \beta_{0j} \sigma_{0t-j}^2, \end{aligned} \quad (85)$$

где $y_t = \varepsilon_t$, $\omega_0 > 0$, $\alpha_{0i} \geq 0$ ($i = 1, \dots, q$), $\beta_{0j} \geq 0$ ($j = 1, \dots, p$), $\{z_t\}$ – *IID* случайные величины, такие что $\mathbb{E}[z_1^2] = 1$, и порядки p и q известны (см. Francq & Zakoïan, 2009). Francq & Zakoïan (2004) и Straumann (2005) показывают, что гауссовская ММКП-оценка (85) строго состоятельна при следующих условиях (см. Francq & Zakoïan, 2009):

- A1. $\theta_0 \equiv (\omega_0, \alpha_{01}, \dots, \alpha_{0q}, \beta_{01}, \dots, \beta_{0p})' \in \Theta \subset (0, +\infty) \times [0, \infty)^{p+q}$, где Θ – компакт,
- A2. Процесс в θ_0 имеет строго стационарное, эргодичное и неупреждающее решение,
- A3. $\forall \theta \in \Theta \sum_{i=1}^p \beta_i < 1$,
- A4. z_t^2 имеет невырожденное распределение с $\mathbb{E}[z_1^2] = 1$,
- A5. При $p > 0$ ($\sum_{i=1}^q \alpha_{0i} z^i$) и $(1 - \sum_{i=1}^p \beta_{0i} z^i)$ не имеют общих корней, $\sum_{i=1}^q \alpha_{0i} \neq 0$ и $\alpha_{0q} + \beta_{0p} \neq 0$,

то есть

$$\hat{\theta}_T \xrightarrow{\text{п.н.}} \theta_0 \quad \text{при} \quad T \rightarrow \infty.$$

Условие $\mathbb{E}[z_1] = 0$ не требуется. Предположение $\mathbb{E}[z_1^2] = 1$ делается для идентифицируемости и не является ограничивающим, достаточно $\mathbb{E}[z_1^2] < \infty$ (см. Berkes & Horváth, 2003).

10.2.2 Асимптотическая нормальность

Lumsdaine (1996) делает предположения о масштабированной переменной, ε_t/σ_t , а не о наблюдаемых данных. В качестве вспомогательных предположений требуется, чтобы масштабированная переменная была *IID* из симметричного одномодального распределения с конечным 32-м моментом. Lee & Hansen (1994) внесли свой вклад, рассмотрев гораздо более широкий класс GARCH-процессов, сосредотачиваясь на свойствах ММКП-оценок. Предполагается, что условное среднее и дисперсия верно специфицированы, и для оценивания параметров используется принцип максимального правдоподобия. Авторы подчеркивают, что нет причины полагать, что вся условная зависимость содержится в условном среднем и дисперсии, так что ε_t/σ_t , масштабированная переменная, не обязана быть независимой по времени. Они допускают некоторую зависимость от времени, полагая, что масштабированная переменная строго стационарна и эргодична. Для случая IGARCH им удается доказать лишь существование состоятельного корня функции правдоподобия. Этот результат требует, чтобы условный момент порядка $2 + \delta$ масштабированной переменной был равномерно ограничен. Доказательство асимптотической нормальности (включая IGARCH-модель) дополнительно требует предположения о том, что условный четвертый момент масштабированной переменной равномерно ограничен.

Что касается модели (85), как следует из Francq & Zakoïan (2004), при изложенных предпосылках и

$$A6. \theta_0 \in \text{int}\{\Theta\},$$

$$A7. \mathbb{E}[z_1^4] < \infty,$$

ММКП-оценка асимптотически нормальна (см. Francq & Zakoïan, 2004):

$$\sqrt{T} (\hat{\theta}_T - \theta_0) \xrightarrow{d} N \left(0, \mathbb{E} \left[(9z_1^4 - 1) (\sigma_t^2(\theta))^{-2} \frac{\partial \sigma_1^2(\theta_0)}{\partial \theta} \frac{\partial \sigma_1^2(\theta_0)}{\partial \theta'} \right] \right). \quad (86)$$

Следует отметить, что для наблюдаемого процесса не накладываются никакие условия на моменты. Это особенно интересно для финансовых временных рядов, для которых существование моментов четвертого или даже второго порядка находится под вопросом. Условие на

момент в A7 для *IID* процесса z_t , очевидно, необходимо для существования дисперсии гауссовского распределения в (86). Вместо предположения об *IID* можно потребовать, чтобы z_t была последовательностью мартингалов, ценой дополнительных предположений о моментах.

Когда $\mu_t \neq 0$ и μ_t можно представить в виде ARMA-модели, GARCH-процесс не наблюдается напрямую. Если допустить ARMA-часть, существенно расширяется спектр приложений, но это приводит к серьезным техническим трудностям при доказательстве асимптотических результатов. В Francq & Zakoian (2004) доказано, что состоятельность MMKP-оценок имеет место при предположениях, схожих со случаем чистой GARCH-модели. В частности, наблюдаемый процесс не обязан иметь конечную дисперсию для состоятельности MMKP-оценок, но требуется предположение $\mathbb{E}[z_1] = 0$. Для асимптотической нормальности MMKP-оценок инновации должны иметь вторые моменты; это неудивительно, поскольку в случае ARMA-моделей с *IID* инновациями MMKP-оценки асимптотически нормальны только при существовании вторых моментов (см. Brockwell & Davis, 1991). Можно показать, что дисперсия скор-вектора для MMKP-оценок может не существовать, если $\mathbb{E}[y_t^2] = +\infty$, и асимптотическое распределение MMKP-оценок неизвестно при $\mathbb{E}[y_t^4] = +\infty$.

Для получения асимптотического распределения MMKP-оценок в GARCH-моделях критическим предположением является строгая положительность компонент истинного вектора параметров. Если некоторые компоненты равны нулю, параметр, на который наложено условие неотрицательности, лежит на границе пространства параметров и предположение A5 не выполняется. Это предположение является серьезным ограничением теории оценивания GARCH-моделей. Асимптотическое распределение MMKP-оценок, когда один или несколько коэффициентов GARCH-модели равны нулю, не может быть гауссовым. Важное следствие негауссова поведения MMKP-оценок состоит в том, что тесты Вальда и отношения правдоподобия не подчиняются стандартному асимптотическому χ^2 -распределению (см. Francq & Zakoian, 2009).

10.2.3 Негауссовы функции правдоподобия

Меняющейся во времени волатильности недостаточно для полного объяснения толстых хвостов: распределение очищенных от воздействия волатильности остатков имеет более толстые хвосты, чем хвосты нормального распределения. Некоторые исследователи ARCH-моделей утверждают, что гауссово правдоподобие часто не подходит для анализа из-за большого числа выбросов в переменной. Отсюда следует, что стандартизированные инновации z_t не являются нормальными и необходима более общая спецификация условного распределения.

В литературе применялись альтернативные условные распределения для учета толстых хвостов и асимметрии спекулятивных цен. Есть несколько способов работы с асимметрией или толстыми хвостами распределения. Во-первых, некоторые распределения допускают асимметрию или толстые хвосты. Например, скошенное t -распределение или распределение IV Пирсона обладают такими характеристиками. Во-вторых, асимметрию можно получить, используя разложение вокруг симметричного распределения. Для улавливания асимметрии и толстых хвостов было предложено несколько альтернативных распределений. Например, Bollerslev (1987) и Baillie & Bollerslev (1989) использовали t -распределение Стьюдента для моделирования обменных курсов, то есть предположили, что z_t следует стандартизированному t -распределению Стьюдента с нулевым средним и единичной дисперсией, имеющему плотность

$$f(z_t; \nu) = \frac{\Gamma(\frac{\nu+1}{2})}{\Gamma(\frac{1}{2}\nu)\sqrt{\pi(\nu-2)}} \left[1 + \frac{z_t^2}{\nu-2} \right]^{-\frac{(\nu+1)}{2}}, \quad (87)$$

где ν – число степеней свободы, $\nu > 2$. Куртозис условного распределения z_t равен

$$k_c = 3 \frac{\nu - 2}{\nu - 4}$$

при $\nu > 4$. Безусловный куртозис ошибки ε_t равен

$$k_u = 3 \frac{\nu - 2}{\nu - 4} \frac{\mathbb{E}[\sigma_t^4]}{(\mathbb{E}[\sigma_t^2])^2}.$$

Множитель, включающий число степеней свободы, улавливает чрезмерный эксцесс, поскольку он по построению больше единицы. Hsieh (1989) и Nelson (1991) моделируют обменные курсы и цены акций, соответственно, используя негауссовское правдоподобие на основе обобщенного распределения ошибок (GED) (см. раздел 5). Куртозис условного распределения z_t равен

$$k_c = \frac{\Gamma(1/v)\Gamma(5/v)}{(\Gamma(3/v))^2},$$

то есть он больше 3 при $v > 2$. Безусловный куртозис ε_t равен

$$k_u = \frac{\Gamma(1/v)\Gamma(5/v)}{(\Gamma(3/v))^2} \frac{\mathbb{E}[\sigma_t^4]}{(\mathbb{E}[\sigma_t^2])^2}.$$

Распределение Стьюдента симметрично, так что оно не может уловить асимметрию. Обобщение, которое исправляет этот недостаток, было предложено в Hansen (1994) и называется скошенным t -распределением Стьюдента:

$$f(z_t; \nu, \lambda) = \begin{cases} bc \left(1 + \frac{1}{\nu-2} \left(\frac{bz_t+a}{1-\lambda}\right)^2\right)^{-(\nu+1)/2} & \text{при } z_t < -\frac{a}{b}, \\ bc \left(1 + \frac{1}{\nu-2} \left(\frac{bz_t+a}{1+\lambda}\right)^2\right)^{-(\nu+1)/2} & \text{при } z_t \geq -\frac{a}{b}, \end{cases},$$

где

$$a = 4\lambda c \left(\frac{\nu-2}{\nu-1}\right), \quad b = \sqrt{1 + 3\lambda^2 - a^2}, \quad c = \frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\Gamma\left(\frac{\nu}{2}\right) \sqrt{\pi(\nu-2)}}.$$

Эта плотность определена при $2 < \nu < \infty$ и $-1 < \lambda < +1$. Она включает в себя множество частных случаев, что позволяет использовать стандартные тесты в рамках метода максимального правдоподобия:

1. При $\lambda=0$ это обыкновенное t -распределение Стьюдента.
2. При $\lambda=0$ и $\nu=\infty$ это нормальное распределение.

Дальнейшие расширения распределения даны в Theodossiou (1998) и Jondeau & Rockinger (2003). Vai et al. (2003) приводят точное выражение для безусловного куртозиса в GARCH-модели и модели стохастической волатильности, когда инновации не являются условно нормальными. Они показывают, что безусловный куртозис для обеих моделей определяется совместно условным распределением z_t и кластеризацией волатильности ε_t .

Были предложены некоторые другие негауссовские распределения ошибок. Fernández & Steel (1998) предложили использовать асимметричное t -распределение Стьюдента для учета как скошенности, так и эксцесса стандартизированных остатков. Venter & Jongh (2002) предложили нормальное обратное гауссовское распределение. Gallant & Tauchen (2001) предложили очень гибкое полупараметрическое распределение инноваций на основе эрмитова

разложения гауссовской плотности. Это разложение позволяет улавливать общие отклонения формы распределения от нормальной в стандартизованных остатках GARCH-модели.

Panorska et al. (1995) предполагают, что распределение z_t принадлежит семейству негауссовских стабильных распределений. Индекс стабильности α полагается большим единицы, чтобы гарантировать конечность среднего. Индекс стабильности также определяет толщину хвостов распределения. Хвосты становятся тоньше по мере приближения α к 2; при $\alpha = 2$ стандартное распределение Парето совпадает с нормальным распределением $N(0, 2)$; при $\alpha < 2$, у z_t не существует моментов порядка α и выше. Помимо допущения бесконечных шоков дисперсии, стабильный GARCH-процесс можно обобщить, рассматривая асимметричные инновации и процессы, в которых GARCH присутствует не только в условных моментах второго порядка, но и в абсолютных моментах порядка $\delta > 0$. Для этого предполагается, что z_t имеет стандартное асимметричное стабильное распределение Парето с индексом стабильности α , параметром скошенности $\beta \in [-1, 1]$ (стабильное распределение симметрично для $\beta = 0$, скошено вправо для $\beta > 0$ и влево для $\beta < 0$), нулевым параметром положения и единичным параметром масштаба. ММКП-оценивание является приблизительным в том смысле, что плотность стабильного распределения Парето необходимо аппроксимировать. Для этого можно применить алгоритм из Mittnik et al. (1999), приближающий плотность стабильного распределения Парето с помощью быстрого преобразования Фурье характеристической функции.

В важной статье Newey & Steigerwald (1997) обсуждают дилемму прекращения использования гауссовского ММКП-оценивания. Когда используется неверное негауссовское правдоподобие, оценки могут терять состоятельность. Авторы показывают, что состоятельность достигается, если (i) условное среднее есть тождественный ноль, или (ii) предполагаемая (теоретическая) и истинная (эмпирическая) плотности распределения ошибок симметричны вокруг нуля. Если условие симметрии не выполняется, правильной спецификации условного среднего и дисперсии недостаточно для обеспечения состоятельности ММКП-оценок, так как среднее и дисперсия не обязаны соответствовать естественным параметрам положения и масштаба. Для плотности $f(x)$ естественный параметр положения, μ , и естественный параметр масштаба, σ , минимизируют выражение

$$\mathbb{E} \left[-\frac{1}{2} \log(\sigma^2) + \log \left(f \left(\frac{x - \mu}{\sigma} \right) \right) \right].$$

В этом случае требуется дополнительный параметр положения распределения, чтобы выполнить условие идентификации, необходимое для состоятельности. Параметр положения учитывает асимметрию z_t (то есть разницу между условным средним и естественным параметром положения), и его можно использовать в спецификации условного среднего либо функции распределения. Newey & Steigerwald (1997) предлагают оценивать следующую модель:

$$y_t = \mu_t(\theta) + \sigma_t(\theta)(\alpha + z_t(\theta)),$$

так что

$$l_t(\theta, \alpha, \eta) = -\frac{1}{2} \log(\sigma_t^2(\theta)) + \log \left(f \left(\frac{y_t - \mu_t - \alpha \sigma_t(\theta)}{\sigma_t}; \eta \right) \right).$$

Таким образом, при использовании негауссовской плотности в построении функции правдоподобия важно, подтверждают ли тесты на адекватность, что предполагаемое распределение соответствует данным. Li & Turtle (2000) рассматривают подход с оценивающей функцией для оценки параметров ARCH-модели на основе некоторых предположений об условных моментах как попытку улучшить свойства ММКП-оценок для негауссовских ошибок.

Известно, что, если распределение ошибок имеет конечный четвертый момент, ММКП-оценки асимптотически нормальны в случае ARCH-модели (Weiss, 1986) и GARCH(1, 1)-модели

(Lee & Hansen, 1994, Lumsdaine, 1996). Тем не менее, для других моделей сложно добиться много большего, чем состоятельность, особенно в случае распределений с достаточно толстыми хвостами, представляющих особый интерес в финансовых приложениях. Hall & Yao (2003) разработали теорию оценок для ARCH и GARCH-моделей, уделяя особое внимание случаю распределений с толстыми хвостами. Возникающие предельные распределения являются многомерными стабильными законами, и их особенно трудно оценивать напрямую.

10.2.4 Эффективность ММКП-оценок

Engle et al. (1991) и Bollerslev & Wooldridge (1992) оценили потерю эффективности, связанную с использованием метода максимального квазиправдоподобия при ненормальном распределении для модели GARCH(1,1). Они обнаружили, что если распределение z_t симметрично, то ММКП-оценки часто близки к ММП-оценкам. Но если z_t имеет скошенное распределение, ММКП-оценки могут значительно отличаться от ММП-оценок. Engle et al. (1991) заключают, что «имеет смысл искать лучшие оценки, чем ММКП-оценки».

Berkes & Horváth (2004) рассматривают весьма общую модель, в которой функция $f(z_t(\theta); \eta)$, используемая для построения логарифмической функции правдоподобия, необязательно является истинной функцией плотности для z_t . В случае чистой GARCH-модели (то есть при $\mu_t = 0$) при некотором условии регулярности Berkes & Horváth (2004) и Straumann (2005) показали, что эта (негауссовская) ММКП-оценка почти наверное сходится к

$$\theta_0^* = (d\omega, d\alpha_1, \dots, d\alpha_q, \beta_1, \dots, \beta_p)', \quad d > 0,$$

если плотность неверно специфицирована и негауссовская, d обычно не равна 1, и ММКП-оценки несостоятельны. Интересно, что гауссовские ММКП-оценки не всегда неэффективны, если распределение ошибок ненормально. Francq & Zakoïan (2009) показывают, что при некоторых предположениях о плотности $f(z_t(\theta))$ ММКП-оценки имеют такую же асимптотическую дисперсию, что и ММП-оценки, когда плотность z_t имеет вид

$$f(z_t) = \frac{a^a}{\Gamma(a)} \exp(-az_t^2) |z_t|^{2a-1}, \quad a > 0.$$

Если плотность $f(z_t)$ не принадлежит этому семейству распределений, то ММКП-оценки асимптотически неэффективны. Относительная асимптотическая эффективность ММП по сравнению с ММКП в случае нормализованной студентовской плотности с ν степенями свободы равна

$$\frac{\mathbb{V}^{asy} \left[\sqrt{T}(\hat{\theta}_T^{QML} - \theta_0) \right]}{\mathbb{V}^{asy} \left[\sqrt{T}(\hat{\theta}_T^{ML} - \theta_0) \right]} = \frac{\nu(\nu - 1)}{\nu(\nu - 1) - 12} \quad \text{при } \nu > 3.$$

В общем случае f неизвестна и ее можно заменить оценкой, полученной из стандартизованных остатков $\hat{z}_t = \varepsilon_t / \sigma_t(\hat{\theta}_T^{QML})$, $t = 1, \dots, T$. Можно использовать непараметрическую ядерную оценку плотности f . Адаптивное оценивание параметров GARCH-модели возможно вплоть до параметра масштаба, с той же асимптотической точностью, как если бы распределение ошибок было известно; см. Drost & Klaassen (1997). Тем не менее, адаптивное оценивание всех параметров GARCH-модели невозможно. Адаптивное оценивание в GARCH-моделях исследовано несколькими авторами, см. Drost & Klaassen (1997), Engle et al. (1991), Ling & McAleer (2003), Linton (1993).

11 Процедуры инференции

11.1 Тестирование на ARCH-эффекты

Тесты на наличие ARCH-эффектов подробно рассмотрены в литературе. Простой и часто применяемый тест нулевой гипотезы

$$H_0 : \alpha_1 = \alpha_2 = \dots = \alpha_q = 0$$

против альтернативы

$$H_1 : \alpha_1 \geq 0, \alpha_2 \geq 0, \dots, \alpha_q \geq 0$$

хотя бы с одним строгим неравенством, – это тест множителей Лагранжа (LM-тест), предложенный в Engle (1982). Тест основан на скор-функции и информационной матрице при нулевой гипотезе, в предположении, что стандартизированные инновации z_t нормально распределены. Рассмотрим ARCH-модель в (6): $\sigma_t^2 = \zeta_t \alpha$. Пусть $\zeta_t = (1, \hat{\varepsilon}_{t-1}^2, \dots, \hat{\varepsilon}_{t-q}^2)$, $\alpha = (\alpha_0, \alpha_1, \dots, \alpha_q)'$, где $\hat{\varepsilon}_t = y_t - \mu_t(\hat{\theta})$ – остатки. При верной нулевой гипотезе σ_t^2 постоянна, то есть $\sigma_t^2 = \sigma_0^2$. Скор-функция гауссовского логарифмического правдоподобия в (84) при верной нулевой гипотезе имеет вид

$$\left. \frac{\partial L_T}{\partial \alpha} \right|_0 = \frac{1}{2\sigma_0^2} \sum_t \zeta_t' \left(\frac{\hat{\varepsilon}_t^2}{\sigma_0^2} - 1 \right) = \frac{1}{2\sigma_0^2} Z' f^0, \quad (88)$$

где $f^0 = \left[\left(\frac{\hat{\varepsilon}_1^2}{\sigma_0^2} - 1 \right), \dots, \left(\frac{\hat{\varepsilon}_T^2}{\sigma_0^2} - 1 \right) \right]'$, и $Z' = (\zeta_1', \dots, \zeta_T')$ – $(q+1) \times T$ матрица. Матрица вторых производных имеет вид

$$\frac{\partial^2 l_t}{\partial \alpha \partial \alpha'} = - \left(\frac{1}{\sigma_t^2} \right)^2 \frac{\hat{\varepsilon}_t^2}{\sigma_t^2} \zeta_t' \zeta_t + \frac{1}{2} \left(\frac{1}{\sigma_t^2} \right)^2 \zeta_t' \zeta_t.$$

Отсюда находим информационную матрицу при нулевой гипотезе, равную среднему по всем t ожидаемых значений условного матожидания:

$$\begin{aligned} A_{\alpha\alpha,0} &= -\frac{1}{T} \mathbb{E} \left[\mathbb{E} \left[\sum_t \frac{\partial^2 l_t}{\partial \alpha \partial \alpha'} \mid \Phi_{t-1} \right] \right] \Big|_0 = \\ &= -\frac{1}{T} \mathbb{E} \left[\sum_t \mathbb{E} \left[- \left(\frac{1}{\sigma_t^2} \right)^2 \frac{\hat{\varepsilon}_t^2}{\sigma_t^2} \zeta_t' \zeta_t + \frac{1}{2} \left(\frac{1}{\sigma_t^2} \right)^2 \zeta_t' \zeta_t \mid \Phi_{t-1} \right] \right] \Big|_0 = \\ &= \frac{1}{T} \sum_{t=1}^T \left\{ -\frac{1}{2} \left(\frac{1}{\sigma_0^2} \right)^2 \mathbb{E} [\zeta_t' \zeta_t] + \left(\frac{1}{\sigma_0^2} \right)^2 \mathbb{E} [\zeta_t' \zeta_t] \right\} = \\ &= \frac{1}{2} \left(\frac{1}{\sigma_0^2} \right)^2 \frac{1}{T} \sum_{t=1}^T \mathbb{E} [\zeta_t' \zeta_t]. \end{aligned}$$

LM-статистика равна

$$\xi_{LM} = \frac{1}{T} \left(\left. \frac{\partial L_T}{\partial \alpha} \right|_0 \right)' A_{\alpha\alpha,0}^{-1} \left(\left. \frac{\partial L_T}{\partial \alpha} \right|_0 \right). \quad (89)$$

LM-статистику можно состоятельно оценить:

$$\xi_{LM} = f^{0'} Z \frac{1}{2\sigma_0^2} \left[\frac{1}{2} \left(\frac{1}{\sigma_0^2} \right)^2 \sum_{t=1}^T \mathbb{E} [\zeta_t' \zeta_t] \right]^{-1} \frac{1}{2\sigma_0^2} Z' f^0 = \frac{1}{2} f^{0'} Z (Z' Z)^{-1} Z' f^0 \quad (90)$$

Если предположить нормальность, то $p \lim (f^{0'} f^0 / T) = 2$. Таким образом, асимптотически эквивалентная статистика имеет вид

$$\xi^* = T f^{0'} Z (Z' Z)^{-1} Z' f^0 / (f^{0'} f^0) = TR^2, \quad (91)$$

где R^2 – квадрат коэффициента множественной корреляции между f^0 и Z . Поскольку добавление константы и умножение на число не меняют R^2 регрессии, это также R^2 регрессии $\hat{\varepsilon}_t^2$ на константу и q лагированных значений $\hat{\varepsilon}_t^2$. Статистика имеет асимптотическое хи-квадрат распределение с q степенями свободы, если нулевая гипотеза верна. Тест состоит в регрессировании квадратов остатков на константу и q лагов и сравнением TR^2 с χ_q^2 . Это асимптотически локально наиболее мощный тест.

Однако Lumsdaine & Ng (1999) показали, что LM-тест (91) может отвергать нулевую гипотезу при неверной спецификации уравнения для условного среднего (10). Они показали, что ошибки спецификации приводят к серийной корреляции остатков $\hat{\varepsilon}_t$, а следовательно, и их квадратов, $\hat{\varepsilon}_t^2$. Следовательно, нужно быть аккуратным при спецификации уравнения условного среднего (10) прежде, чем тестировать на ARCH-эффекты.

Lee & King (1993) предложили локально наиболее мощный в среднем скор-тест на наличие ARCH и GARCH-эффектов. Тест основан на сумме скор-функций при нулевой гипотезе, когда мешающие параметры заменяются ММП-оценками. В отсутствие мешающих параметров этот тест локально наиболее мощный в среднем. Затем сумма скор-функций стандартизируется путем деления на асимптотическую стандартную ошибку. Полученный тест имеет асимптотическое $N(0, 1)$ -распределение. Тест-статистики, применяемые для тестирования против ARCH(q)-процесса, можно также использовать для тестирования против GARCH(p, q)-процесса. В малых выборках этот тест более мощный, чем LM-тест, и его асимптотические критические значения по меньшей мере так же точны.

Критерии Вальда и отношения правдоподобия (LR) также можно применять для тестирования гипотезы об условной гомоскедастичности, например, против GARCH(1, 1) в качестве альтернативы.

Статистика для тестирования $H_0 : \alpha_1 = \beta_1 = 0$ против $H_1 : \alpha_1 \geq 0$ или $\beta_1 \geq 0$ с хотя бы одним строгим неравенством не следует χ^2 -распределению с двумя степенями свободы, поскольку стандартное предположение о том, что истинный параметр при H_0 не лежит на границе пространства параметров, нарушается.

11.2 IGARCH(1,1) против GARCH(1,1)

Если необходимо тестировать наличие единичного корня в GARCH(1, 1), можно использовать тест отношения правдоподобия (LR), скор-тест (LM) или тест Вальда. Хорошо известно, что при нулевой гипотезе все три статистики асимптотически распределены как χ_1^2 . Однако их свойства в конечных выборках различаются. Предположим, что истинная модель имеет вид

$$y_t = \mu + \varepsilon_{0t}, \quad \varepsilon_{0t} | \Phi_{t-1} \sim N(0, \sigma_{0t}^2), \\ \sigma_{0t}^2 = \omega_0 + \alpha_0 \varepsilon_{0t-1}^2 + \beta_0 \sigma_{0t-1}^2, \quad \theta_0 = (\mu_0, \omega_0, \alpha_0, \beta_0)'$$

Определим нулевую гипотезу:

$$H_0 : \alpha_0 + \beta_0 = 1,$$

или

$$H_0 : g(\theta_0) = \alpha_0 + \beta_0 - 1 = 0,$$

против альтернативы

$$H_1 : \alpha_0 + \beta_0 < 1.$$

Теперь предположим, что оцениваемая модель верно специфицирована для условного среднего и дисперсии:

$$\begin{aligned} y_t &= \mu + \varepsilon_t, \\ \sigma_t^2 &= \omega + \alpha \varepsilon_{t-1}^2 + \beta \sigma_{t-1}^2, \quad \theta = (\mu, \omega, \alpha, \beta)', \\ L_T(\theta) &= \sum_{t=1}^T l_t(\theta), \quad A_0 = -\mathbb{E} \left[\frac{\partial^2 l_t(\theta_0)}{\partial \theta \partial \theta'} \right] \end{aligned}$$

Оцененная матрица Гессе имеет вид

$$A_T(\theta) = -\frac{1}{T} \sum_t \frac{\partial^2 l_t(\theta)}{\partial \theta \partial \theta'},$$

а в случае GARCH-моделей

$$A_T(\theta) = \frac{1}{T} \left[\frac{1}{2} \sum_{t=1}^T (\sigma_t^2)^{-2} \frac{\partial \sigma_t^2}{\partial \theta} \frac{\partial \sigma_t^2}{\partial \theta'} + \sum_{t=1}^T (\sigma_t^2)^{-1} \frac{\partial \varepsilon_t}{\partial \theta} \frac{\partial \varepsilon_t}{\partial \theta'} \right].$$

Информационная матрица

$$B_0 = \frac{1}{T} \mathbb{E} \left[\frac{\partial L_T(\theta_0)}{\partial \theta} \frac{\partial L_T(\theta_0)}{\partial \theta'} \right]$$

оценивается с помощью

$$B_T(\theta) = \frac{1}{T} \sum_{t=1}^T \frac{\partial l_t}{\partial \theta} \frac{\partial l_t}{\partial \theta'}.$$

Здесь $A_T(\hat{\theta})$ и $B_T(\hat{\theta})$ – состоятельные оценки A и B , где $\hat{\theta}$ – ММП-оценка θ_0 . Если условное распределение ε_0 нормально, то $A = B$. Более того,

$$D\sqrt{T}(\hat{\theta} - \theta_0) \xrightarrow{d} N(0, I_k)$$

D равна $A^{1/2}$ в случае условного нормального распределения и $B^{-1/2}A$ в общем случае.

В Lumsdaine (1995) исследованы свойства в конечных выборках в предположении, что истинное распределение нормально. Пусть $\hat{\theta}_R$ и $\hat{\theta}_U$ – ММП-оценки моделей с ограничением и без ограничения, соответственно. При нулевой гипотезе следующие статистики имеют χ_1^2 -распределение:

$$\xi_{LR} = -2 \left(L_T(\hat{\theta}_R) - L_T(\hat{\theta}_U) \right), \quad (92)$$

$$\xi_{LM}^{NR} = \frac{1}{T} \left(\frac{\partial L_T(\hat{\theta}_R)}{\partial \theta} \right)' A_T^{-1}(\hat{\theta}_R) \left(\frac{\partial L_T(\hat{\theta}_R)}{\partial \theta} \right), \quad (93)$$

$$\xi_W^{NR} = Tg(\hat{\theta}_U)' \left[\left(\frac{\partial g(\hat{\theta}_U)}{\partial \theta'} \right) A_T^{-1}(\hat{\theta}_U) \left(\frac{\partial g(\hat{\theta}_R)}{\partial \theta} \right) \right]^{-1} g(\hat{\theta}_U). \quad (94)$$

Робастные версии (93) и (94) имеют вид

$$\begin{aligned}\xi_{LM}^R &= \frac{1}{T} \left[\left(\frac{\partial L_T(\hat{\theta}_R)}{\partial \theta} \right)' A_T^{-1}(\hat{\theta}_R) \left(\frac{\partial g(\hat{\theta}_R)}{\partial \theta} \right) \right] \\ &\quad \times \left[\left(\frac{\partial g(\hat{\theta}_R)}{\partial \theta'} \right) A_T^{-1}(\hat{\theta}_R) B_T(\hat{\theta}_R) A_T^{-1}(\hat{\theta}_R) \right]^{-1} \\ &\quad \times \left[\left(\frac{\partial g(\hat{\theta}_R)}{\partial \theta'} \right) A_T^{-1}(\hat{\theta}_R) \left(\frac{\partial L_T(\hat{\theta}_R)}{\partial \theta} \right) \right], \\ \xi_W^R &= T g(\hat{\theta}_U)' \left[\left(\frac{\partial g(\hat{\theta}_U)}{\partial \theta'} \right) A_T^{-1}(\hat{\theta}_U) B_T(\hat{\theta}_U) A_T^{-1}(\hat{\theta}_U) \left(\frac{\partial g(\hat{\theta}_R)}{\partial \theta} \right) \right]^{-1} g(\hat{\theta}_U).\end{aligned}$$

Статистики Вальда (ξ_W^R и ξ_W^{NR}) равны квадратам (робастной и неробастной, соответственно) t -статистик для $g(\theta) = 0$. Lumsdaine (1995) сравнивает робастную и неробастную версии LM и вальдовской статистик методом Монте-Карло и обнаруживает, что робастные версии обладают более желательными свойствами размера теста и частоты отвержения нулевой гипотезы. LM-тесты склонны к значительному превышению размера при нулевой гипотезе.

11.3 Тест с искусственной регрессией в модели GARCH в среднем

Стандартной практикой в прикладной эконометрике является проверка адекватности оцененной модели. В случае регрессионных GARCH-моделей одним из ключевых вопросов при выборе модели является тест на отсутствие обратных эффектов условной дисперсии на условное среднее. Если последнее зависит от условной дисперсии, получаем модель GARCH в среднем. Рассмотрим процесс GARCH(1, 1) в среднем:

$$y_t = \pi x_t + \delta \sigma_t^2 + \varepsilon_t = \mu_t + \varepsilon_t, \quad (95)$$

где

$$\begin{cases} \varepsilon_t | \Phi_{t-1} \sim N(0, \sigma_t^2), \\ \sigma_t^2 = \omega + \alpha_1 \varepsilon_{t-1}^2 + \beta_1 \sigma_{t-1}^2. \end{cases} \quad (96)$$

Процесс предполагается слабо стационарным. Важной задачей в этом контексте является тест на GARCH-модель против модели GARCH в среднем, то есть $H_0 : \delta = 0$, для данной спецификации условной дисперсии. В общем случае, анализ моделей GARCH в среднем гораздо сложнее анализа простых GARCH-моделей (см. Pagan, 1996). В последних остатки, полученные из регрессии условного среднего можно использовать для диагностики и проверки спецификации. Однако в моделях GARCH в среднем невозможно оценить случайный член, заранее не специфицировав верную модель для условной дисперсии, так что анализ до оценивания затруднителен.

Стандартная практика полагается на анализ после оценивания. Возможная стратегия решения этой проблемы – использовать LM-тест. Есть несколько асимптотически эквивалентных способа подсчета тестов типа LM; некоторые из них требуют явного подсчета матрицы Гессе, а другие – нет. Поскольку последние часто более удобны вычислительно, Lucchetti & Rossi (2005) рассматривают два метода подсчета тестов типа LM для $H_0 : \delta = 0$ в модели (95), не требующих вычисления матрицы Гессе, известные как методы OPG (регрессия внешнего произведения градиентов) и DLR (регрессия двойной длины) (см. Davidson & McKinnon, 1984; 1988; 2001).

OPG-регрессии впервые были использованы для сокращения вычислений в алгоритме ВННН при максимизации правдоподобия, но впоследствии также применялись для подсчета

статистик типа LM (см. Godfrey & Wickens, 1981). Регрессия двойной длины менее известна. DLR – это вычислительный метод, который можно использовать для нахождения величин, связанных с ММП-оцениванием и тестированием, таких как информационная матрица.

Концепция *искусственной регрессии* была предложена в Davidson & McKinnon (1984). Рассмотрим параметрическую модель с неизвестным вектором параметров θ , принадлежащим пространству параметров $(\Theta \subseteq \mathbb{R}^m)$, который можно оценить методом максимального правдоподобия. Стандартную искусственную регрессию можно записать в виде

$$r(\theta) = R(\theta)b + \text{остатки}, \quad (97)$$

где b – m -вектор коэффициентов. Термин «остатки» употребляется в нейтральном смысле, во избежании восприятия (97) как статистической модели. Чтобы (97) была искусственной регрессией, вектор $r(\theta)$ и матрица $R(\theta)$ должны удовлетворять определенным свойствам. Два главных условия имеют вид:

$$\begin{aligned} R(\theta)'r(\theta) &= \mathbf{g}(\theta), \\ \frac{1}{T}R(\bar{\theta})'R(\bar{\theta}) &\xrightarrow{P} \mathcal{I}(\theta), \end{aligned}$$

где $\bar{\theta}$ – произвольная состоятельная оценка θ , $\mathbf{g}(\cdot)$ – скор-вектор, а $\mathcal{I}(\theta)$ – информационная матрица; см. Davidson & McKinnon (2001). OPG-регрессия основана на том факте, что информационная матрица, по определению, является ковариационной матрицей скор-вектора. Для выборки из T наблюдений определим $\ell_t(\theta)$ как t -ый вклад в выборочную логарифмическую функцию правдоподобия, где θ – вектор параметров с m элементами, а \mathbf{G} – матрица вкладов градиентов, то есть $T \times m$ матрица, такая что

$$\mathbf{G}_{ti}(\theta) = \frac{\partial \ell_t(\theta)}{\partial \theta_i}.$$

Тест множителей Лагранжа (LM) можно записать в так называемой скор-форме:

$$LM = T^{-1}\mathbf{g}(\theta)'\mathcal{I}(\theta)^{-1}\mathbf{g}(\theta), \quad (98)$$

где

$$\mathbf{g}(\theta) = \mathbf{G}(\theta)'\boldsymbol{\nu}$$

– скор-вектор для логарифмического правдоподобия, суммированный по всей выборке, а $\mathcal{I}(\theta)$ – информационная матрица ($\boldsymbol{\nu}$ обозначает вектор единиц). OPG-регрессию можно записать в виде

$$\boldsymbol{\nu} = \mathbf{G}b_0 + \text{остатки}.$$

Пусть $\tilde{\theta}$ обозначает ММП-оценку при ограничении, полученную при r ограничениях в максимизации логарифмического правдоподобия. Тогда объясненная сумма квадратов (ESS) в OPG-регрессии

$$\boldsymbol{\nu} = \tilde{\mathbf{G}}b_0 + \text{остатки},$$

где $\tilde{\mathbf{G}} = \mathbf{G}(\tilde{\theta})$ – это OPG-форма LM-статистики, равная T , помноженному на нецентрированный R^2 . Другими словами, LM-тест на основе OPG просто заменяет информационную матрицу на ее выборочный аналог. OPG-регрессия применима к широкому ряду моделей и требует только первые производные. Но в общем случае как оцененные ковариации, так и тест-статистики не очень надежны в конечных выборках. Альтернативой является регрессия двойной длины (DLR), для которой доказаны лучшие свойства в конечных выборках.

Класс моделей, к которым применима DLR, можно записать в виде

$$f_t(w_t, \theta) = u_t, \quad t = 1, \dots, T, \quad (99)$$

где $f_t(\cdot)$ – гладкая функция, зависящая от случайного вектора w_t и вектора параметров θ ; w_t содержит зависимую переменную y_t и некоторые экзогенные и/или предопределенные переменные и/или лагированные зависимые переменные x_t . При данной последовательности информационных множеств Φ_t (обычно включающей x_t) несущественно, чтобы $y_t|\Phi_t$ имело нормальное распределение, хотя критично, чтобы модель можно было преобразовать так, чтобы выполнялось (99), и $f_t|\Phi_t \stackrel{\text{iid}}{\sim} N(0, 1)$ при нулевой гипотезе. t -ый вклад в логарифмическое правдоподобие можно записать в виде

$$\ell_t(\theta) = \text{const} - \frac{1}{2} f_t^2(\theta) + k_t(\theta), \quad (100)$$

где $k_t(\theta)$ – логарифм якобиана $\log \left| \frac{\partial f_t}{\partial y_t} \right|$. Полезно взглянуть на производные f_t и k_t относительно вектора параметров θ , то есть на два якобиана

$$\begin{bmatrix} df_t \\ dk_t \end{bmatrix} = \begin{bmatrix} F_t(\theta) \\ K_t(\theta) \end{bmatrix} d\theta. \quad (101)$$

Из определений F_t и K_t следует, что скор-вектор для наблюдения t можно записать в виде

$$g_t(\theta) = -f_t(\theta)F_t(\theta) + K_t(\theta), \quad (102)$$

так что матрицы \mathbf{F} , \mathbf{K} и вектор \mathbf{f} можно тривиально определить, и градиент равен

$$\mathbf{g} = -\mathbf{F}'\mathbf{f} + \mathbf{K}'\boldsymbol{\iota}.$$

В терминах уравнения (97), для DLR имеем

$$R(\theta) = \begin{bmatrix} -\mathbf{F}(\theta) \\ \mathbf{K}(\theta) \end{bmatrix}, \quad r(\theta) = \begin{bmatrix} \mathbf{f}(\theta) \\ \boldsymbol{\iota} \end{bmatrix}.$$

Регрессия двойной длины имеет вид

$$\begin{bmatrix} \mathbf{f} \\ \boldsymbol{\iota} \end{bmatrix} = \begin{bmatrix} -\mathbf{F} \\ \mathbf{K} \end{bmatrix} b_D + \text{остатки}.$$

Фундаментальный результат, который позволяет DLR работать, состоит в том, что в этом классе моделей (см. Davidson & McKinnon, 1993)

$$\frac{1}{T} (\mathbf{F}'\mathbf{F} + \mathbf{K}'\mathbf{K}) \xrightarrow{p} \mathcal{I}(\theta),$$

если матрицы \mathbf{F} и \mathbf{K} удовлетворяют надлежащим условиям регулярности.

Если оценить DLR, в которой величины $\mathbf{F}(\theta)$, $\mathbf{f}(\theta)$ и $\mathbf{K}(\theta)$ оценены в $\tilde{\theta}$, LM-тест на основе DLR имеет вид

$$\frac{1}{T} \tilde{\mathbf{g}}' \left(\tilde{\mathcal{I}}^{-1} \right) \tilde{\mathbf{g}} = \left(\boldsymbol{\iota}' \tilde{\mathbf{K}} - \tilde{\mathbf{f}}' \tilde{\mathbf{F}} \right) \left(\tilde{\mathbf{F}}' \tilde{\mathbf{F}} + \tilde{\mathbf{K}}' \tilde{\mathbf{K}} \right)^{-1} \left(\tilde{\mathbf{K}}' \boldsymbol{\iota} - \tilde{\mathbf{F}}' \mathbf{f} \right),$$

где правая сторона – объясненная сумма квадратов из DLR, и

$$\tilde{\mathcal{I}} = \frac{1}{T} \left(\tilde{\mathbf{F}}' \tilde{\mathbf{F}} + \tilde{\mathbf{K}}' \tilde{\mathbf{K}} \right)$$

является состоятельной оценкой \mathcal{I} при верной нулевой гипотезе. Заметим, что b_O и b_D можно записать, с точностью до параметра масштаба, как произведение состоятельной оценки \mathcal{I}^{-1}

на \mathbf{g} . На практике OPG и DLR-регрессии ведут к асимптотически эквивалентным формулировкам LM-теста. Тем не менее, они различаются, поскольку используют разные оценки информационной матрицы. Lucchetti & Rossi (2005) рассматривают поведение этих двух статистик при тестировании гипотезы $\delta = 0$ в модели GARCH в среднем (95).

Если собрать все параметры в вектор $\theta = (\pi, \delta, \omega, \alpha_1, \beta_1)'$, уравнение (95) означает, что процесс f_t можно определить как

$$f_t = \frac{\varepsilon_t}{\sqrt{\sigma_t^2}},$$

то есть гауссовский белый шум с единичной дисперсией. При такой формулировке применимо уравнение (100), где k_t равно $-\frac{1}{2} \log \sigma_t^2$ в этом случае. Векторы F_t и K_t , определенные в уравнении (101), необходимо оценивать рекурсивно, из-за рекурсивных членов в (96). Определим «вектор состояния» z_t :

$$\zeta_t = \begin{bmatrix} \varepsilon_t \\ \sigma_t^2 \end{bmatrix}.$$

Для F_t имеем

$$F_t = J_t^f J_t^z = \frac{\partial f_t(\theta)}{\partial \zeta_t(\theta)} \frac{\partial \zeta_t(\theta)}{\partial \theta}$$

и

$$K_t = J_t^k J_t^z = \frac{\partial k_t(\theta)}{\partial \zeta_t(\theta)} \frac{\partial \zeta_t(\theta)}{\partial \theta},$$

так что $df_t = J_t^f d\zeta_t$ и $dk_t = J_t^k d\zeta_t$. Из определений f_t и k_t немедленно следует, что

$$J_t^f = \sigma_t^{-1} \begin{bmatrix} 1 & -\frac{\varepsilon_t}{2\sigma_t^2} \end{bmatrix} \quad (103)$$

$$J_t^k = \begin{bmatrix} 0 & -\frac{1}{2\sigma_t^2} \end{bmatrix} \quad (104)$$

Выражение для J_t^z содержит рекурсивный член, так как из (96) имеем:

$$\begin{aligned} d\zeta_t &= \begin{bmatrix} d\varepsilon_t \\ d\sigma_t^2 \end{bmatrix} \\ &= \begin{bmatrix} 1 & -\delta \\ 0 & 1 \end{bmatrix} \left(\begin{bmatrix} 0 & 0 \\ 2\alpha_1\varepsilon_{t-1} & \beta_1 \end{bmatrix} \begin{bmatrix} d\varepsilon_{t-1} \\ d\sigma_{t-1}^2 \end{bmatrix} + \begin{bmatrix} -x_t & -\sigma_t^2 & 0 & 0 & 0 \\ 0 & 0 & 1 & \varepsilon_{t-1}^2 & \sigma_{t-1}^2 \end{bmatrix} d\theta \right) \end{aligned}$$

или в более компактной форме,

$$d\zeta_t = M_t d\zeta_{t-1} + Q_t d\theta,$$

где определения M_t и Q_t ясны. По индукции, J_t^z должно подчиняться уравнению

$$J_t^z = M_t J_{t-1}^z + Q_t,$$

а следовательно, имея начальную точку⁴ для J_0^z , все необходимые величины можно подсчитать рекурсивно, что позволяет найти \mathbf{F} , \mathbf{K} и \mathbf{G} для любого θ .

Симуляции Монте-Карло для множества одномерных AR-моделей временных рядов с ошибками GARCH(1, 1) показывают, что в конечных выборках размер DLR-LM-теста гораздо ближе к его номинальному значению, чем размер OPG-LM-теста: OPG-LM-тест слишком часто отвергает нулевую гипотезу, что может привести к неверной инференции о параметре при рисковомой премии.

⁴Удобно брать в качестве начальной точки $\sigma_0 = \omega / (1 - \alpha_1 - \beta_1)$, то есть безусловную дисперсию ε_t (см. Nelson & Cao, 1992). Ее производная равна

$$\frac{\partial \sigma_0}{\partial \theta} = \begin{bmatrix} 0 & 0 & \frac{1}{1-\alpha_1-\beta_1} & \frac{\omega}{(1-\alpha_1-\beta_1)^2} & \frac{\omega}{(1-\alpha_1-\beta_1)^2} \end{bmatrix}$$

11.4 Тест на асимметричные эффекты

В каждой модели волатильности неявно присутствует определенная кривая воздействия новостей. В стандартной GARCH-модели кривая воздействия новостей симметрична и центрирована в $\varepsilon_{t-1} = 0$. То есть положительные и отрицательные шоки доходности одного размера приводят к одинаковой волатильности. Кроме того, большие шоки доходности предсказывают большую волатильность, пропорционально квадрату размера шока доходности. Если отрицательный шок доходности ведет к большей волатильности, чем положительный шок того же размера, прогноз GARCH-модели занижает объем волатильности вследствие плохих новостей и завышает его для хороших новостей. Далее, если большие шоки доходности ведут к большей волатильности, чем допускает квадратичная функция, прогноз стандартной GARCH-модели занижает волатильность вследствие большого шока доходности и завышает ее для малого шока доходности.

Engle & Ng (1993) предложили три диагностических теста для моделей волатильности: *тест на смещение для шоков разного знака, тест на смещение для размера отрицательных шоков* и *тест на смещение для размера положительных шоков*. Эти тесты проверяют, можно ли прогнозировать квадрат нормализованных остатков некоторыми переменными, наблюдаемыми в прошлом, которые не включены в используемую модель волатильности. Если эти переменные могут прогнозировать квадрат нормализованного остатка, то модель для дисперсии неверно специфицирована. Тест на смещение для шоков разного знака рассматривает воздействие положительных и отрицательных шоков доходности на волатильность, не прогнозируемое рассматриваемой моделью. Тест на смещение для размера отрицательных шоков рассматривает разные эффекты, оказываемые большими и малыми отрицательными шоками доходности на волатильность, которые не предсказываются моделью для волатильности. Тест на смещение для размера положительных шоков рассматривает разные эффекты, которые большие и малые положительные шоки доходности оказывают на волатильность, не предсказываемые моделью для волатильности.

Для получения оптимальной формы этих тестов Engle & Ng (1993) предполагают, что модель волатильности при верной нулевой гипотезе является частным случаем более общей модели следующего вида:

$$\ln(\sigma_t^2) = \ln(\sigma_{0t}^2 (\delta_0' z_{0t})) + \delta_a' z_{at}, \quad (105)$$

где $\sigma_{0t}^2 (\delta_0' z_{0t})$ – модель волатильности при верной нулевой гипотезе, δ_0 – $(k \times 1)$ вектор параметров при верной нулевой гипотезе, z_{0t} – $(k \times 1)$ вектор объясняющих переменных при верной нулевой гипотезе, δ_a – $(m \times 1)$ вектор дополнительных параметров, z_{at} – $(m \times 1)$ вектор пропущенных объясняющих переменных. Эта форма включает как GARCH, так и EGARCH модели. Например, GARCH(1, 1)-модель соответствует случаю

$$\sigma_{0t}^2 (\delta_0' z_{0t}) = \delta_0' z_{0t}, \quad z_{0t} \equiv [1, \sigma_{t-1}^2, \varepsilon_{t-1}^2]',$$

$$\delta_0 \equiv [\omega, \beta, \alpha]', \quad \delta_a = [\beta^*, \phi^*, \psi^*]', \quad z_{at} = \left[\log(\sigma_{t-1}^2), \frac{\varepsilon_{t-1}}{\sigma_{t-1}}, \left(\frac{|\varepsilon_{t-1}|}{\sigma_{t-1}} - \sqrt{2/\pi} \right) \right]'$$

Обобщающая модель имеет вид

$$\log(\sigma_t^2) = \log(\omega + \beta \sigma_{t-1}^2 + \alpha \varepsilon_{t-1}^2) + \beta^* \log(\sigma_{t-1}^2) + \phi^* \frac{\varepsilon_{t-1}}{\sigma_{t-1}} + \psi^* \left(\frac{|\varepsilon_{t-1}|}{\sigma_{t-1}} - \sqrt{2/\pi} \right),$$

и при $\alpha = \beta = 0$ это модель EGARCH(1, 1), а при $\beta^* = \phi^* = \psi^* = 0$ это GARCH(1, 1)-модель.

Нулевая гипотеза имеет вид $\delta_a = 0$. Пусть v_t – нормализованный остаток, соответствующий наблюдению t при модели волатильности, соответствующей нулевой гипотезе, т.е. $v_t \equiv \varepsilon_t / \sigma_t$. LM-тест статистика для $H_0 : \delta_a = 0$ в (105) – это тест на $\delta_a = 0$ во вспомогательной регрессии

$$v_t^2 = z_{0t}' \delta_0 + z_{at}' \delta_a + u_t, \quad (106)$$

где $z_{0t}^* \equiv \sigma_{0t}^{-2} \left(\frac{\partial \sigma_t^2}{\partial \delta_0} \right)$, $z_{at}^* \equiv \sigma_{0t}^{-2} \left(\frac{\partial \sigma_t^2}{\partial \delta_a} \right)$. Как $\frac{\partial \sigma_t^2}{\partial \delta_0}$, так и $\frac{\partial \sigma_t^2}{\partial \delta_a}$ оцениваются в $\delta_a = 0$ и δ_0 (ММП-оценка δ_0 при верной H_0). Если ограничения на параметры соблюдены, переменные в правой части (106) не должны иметь объясняющей силы. Поэтому тест часто подсчитывается как

$$\xi_{LM} = TR^2,$$

где R^2 – квадрат коэффициента множественной корреляции в (106), а T – число наблюдений в выборке⁵. LM-статистика имеет асимптотическое распределение хи-квадрат с m степенями свободы при верной нулевой гипотезе, где m – число ограничений на параметры. Для обобщающей модели (105) $\frac{\partial \sigma_t^2}{\partial \delta_a}$ равна⁶ $\sigma_{0t}^2 z_{at}^*$ при верной нулевой гипотезе, а следовательно, $z_{at}^* = z_{at}$. v_t^2 регрессируется на константу z_{0t}^* и z_{at} . Переменные в z_{at} – это S_{t-1} , $S_{t-1}^- \varepsilon_{t-1}$ и $S_{t-1}^+ \varepsilon_{t-1}$. Оптимальная форма теста на смещение для шоков разного знака следующая:

$$v_t^2 = a + b_1 S_{t-1}^- + \gamma' z_{0t}^* + e_t,$$

где

$$S_{t-1}^- = \begin{cases} 1, & \varepsilon_{t-1} < 0, \\ 0, & \text{иначе.} \end{cases}$$

Регрессия для теста на смещение для размера отрицательных шоков имеет вид

$$v_t^2 = a + b_2 S_{t-1}^- \varepsilon_{t-1} + \gamma' z_{0t}^* + e_t,$$

а для теста на смещение для размера отрицательных шоков

$$v_t^2 = a + b_3 S_{t-1}^+ \varepsilon_{t-1} + \gamma' z_{0t}^* + e_t, \quad S_{t-1}^+ = \begin{cases} 1, & \varepsilon_{t-1} > 0, \\ 0, & \text{иначе.} \end{cases}$$

t -статистики для b_1 , b_2 и b_3 являются статистиками для тестов на смещение знака, отрицательное смещение знака и положительное смещение знака, соответственно. Совместный тест – это LM-тест на добавление трех переменных в уравнение дисперсии (105) при предполагаемой спецификации:

$$v_t^2 = a + b_1 S_{t-1}^- + b_2 S_{t-1}^- \varepsilon_{t-1} + b_3 S_{t-1}^+ \varepsilon_{t-1} + \gamma' z_{0t}^* + e_t.$$

Тестовая статистика – это TR^2 . Если модель волатильности верна, то $b_1 = b_2 = b_3 = 0$, $\gamma = 0$ и e_t является IID. Если z_{0t}^* не включать, то тест будет консервативен; размер теста будет меньше либо равен номинальному, а мощность, вероятно, снижена.

12 Приближения GARCH-моделей в непрерывном времени

12.1 Слабая GARCH(1,1)-модель в непрерывном времени

Хотя финансовые данные обычно наблюдаются в дискретном времени, количественные финансовые модели формулируются в непрерывном времени, что оказывается удобнее при моделировании ценообразования опционов. Однако модели с непрерывным временем также могут быть хорошим приближением при работе с дискретными наблюдениями. Типичные примеры – высокочастотные или неравномерно наблюдаемые данные. Если в дискретном

⁵Для сильно нелинейных моделей алгоритмы численной оптимизации обычно не гарантируют точную ортогональность v_t^2 и z_{0t}^* . Engle & Ng (1993) предлагают регрессировать y_t^2 только на z_{0t} , а затем использовать полученные остатки (которые ортогональны z_{0t}) вместо v_t^2 в (106).

⁶Действительно, $\sigma_t^2 = \sigma_{0t}^2 (\delta_0' z_{0t}) \exp(\delta_a' z_{at})$, $\partial \sigma_t^2 / \partial \delta_a = \sigma_{0t}^2 z_{at} \exp(\delta_a' z_{at})$ при нулевой гипотезе, $\delta_a = 0$, $\partial \sigma_t^2 / \partial \delta_a = \sigma_{0t}^2 z_{at}$.

времени необходимо специфицировать закон движения для приращения логарифма цены, то в непрерывном времени моделируется поведение собственно логарифма цены. Обычно ненаблюдаемый процесс для волатильности $\{\sigma_t\}_{t \geq 0}$ моделируется как полумартингал, а логарифм цены (p_t) задается выражением

$$p_t = \int_0^t (\mu + b\sigma_s^2) ds + \int_0^t \sigma_s - dM_s,$$

где $\{M_t\}_{t \geq 0}$ – процесс Леви, а μ, b – вещественные константы. В моделях стохастической волатильности с непрерывным временем процесс $\{M_t\}_{t \geq 0}$ обычно независим от $\{\sigma_t\}_{t \geq 0}$, и полагается стандартным броуновским движением. В последнем случае квадратичная вариация p_t равна $\int_0^t \sigma_s^2 ds$. Если модели с непрерывным временем служат для приближения GARCH-процесса, то процесс, из которого делается выборка в дискретном времени, должен быть близок к GARCH-процессу. Таким образом, в оптимальном случае сам процесс является GARCH-процессом, из которого осуществляется выборка через равные промежутки времени $\{nh\}_{n \in \mathbb{N}}$ для каждого $h > 0$. Тем не менее, Drost & Nijman (1993) показали, что GARCH-процессы не замкнуты при агрегировании по времени, то есть, если $\{\varepsilon_t\}_{t \in \mathbb{N}}$ – GARCH-процесс, движимый некоторым шумом $\{z_t\}_{t \in \mathbb{N}}$ с процессом волатильности $\{\sigma_t\}_{t \in \mathbb{N}}$, то – за исключением случая вырожденного шума – не существует IID последовательности шума $\{\tilde{z}_{2t}\}_{t \in \mathbb{N}}$ (где $2t$ означает, что время между двумя последовательными наблюдениями увеличивается вдвое) и процесса волатильности $\{\tilde{\sigma}_{2t}\}_{t \in \mathbb{N}}$, таких, что $\{\tilde{\varepsilon}_{2t}\}_{t \in \mathbb{N}}$ – GARCH-процесс, движимый $\{\tilde{z}_{2t}\}_{t \in \mathbb{N}}$ с процессом волатильности $\{\tilde{\sigma}_{2t}\}_{t \in \mathbb{N}}$. В частности, непрерывный процесс $\{Y_t\}_{t \geq 0}$, который является GARCH(1, 1)-процессом при выборке в моменты времени $\{0, h, 2h, \dots\}$ для некоторой частоты h , не будет GARCH-процессом при выборке в моменты времени $\{0, 2h, 4h, \dots\}$. Аналогично, если для некоторого процесса для логарифма цены $\{p_t\}_{t \geq 0}$ приращения $\{p_{nh} - p_{(n-1)h}\}_{n \in \mathbb{N}}$ длины h являются GARCH(1, 1)-процессом с невырожденным шумом, то приращения $\{p_{n2h} - p_{(n-1)2h}\}_{n \in \mathbb{N}}$ длины $2h$, вообще говоря, не будут GARCH-процессом.

Решение этой проблемы с помощью слабых GARCH-процессов предложили Drost & Nijman (1993). Процесс GARCH(1, 1) с параметрами $(\omega, \alpha_1, \beta_1)$, где $\omega > 0, \alpha_1 \geq 0, \beta_1 \geq 0$, называется *слабым*, если $\{\varepsilon_t\}_{t \in \mathbb{N}}$ имеет конечный четвертый момент и существует слабо стационарный процесс волатильности $\{\sigma_t\}_{t \in \mathbb{N}}$, для которого $LP[\varepsilon_t | \sigma_0^2, z_0, \dots, z_{t-1}, z_0^2, \dots, z_{t-1}^2] = 0$ и $LP[\varepsilon_t^2 | \sigma_0^2, z_0, \dots, z_{t-1}, z_0^2, \dots, z_{t-1}^2] = \sigma_t^2$, где $LP[X | \sigma_0^2, z_0, \dots, z_{t-1}, z_0^2, \dots, z_{t-1}^2]$ обозначает наилучший линейный прогноз квадратично интегрируемой случайной переменной X . В отличие от GARCH-процессов, класс слабых GARCH-процессов замкнут относительно агрегирования по времени, то есть, если $\{\varepsilon_t\}_{t \in \mathbb{N}}$ – симметричный слабый GARCH(1, 1)-процесс, то таковым же является и $\{\varepsilon_{mn}\}_{n \in \mathbb{N}} \forall m \in \mathbb{N}$, см. пример 1 в Drost & Nijman (1993). Drost & Werker (1996) определяют непрерывный слабый GARCH(1, 1)-процесс как одномерный процесс $\{p_t\}_{t \geq 0}$, такой, что $\{p_{t_0+nh} - p_{t_0+(n-1)h}\}_{n \in \mathbb{N}}$ – слабый GARCH(1, 1)-процесс для каждого $h > 0$ и $t_0 \geq 0$. Drost & Werker (1996) также показывают, что параметры дискретизированного слабого GARCH-процесса соответствуют определенным параметрам непрерывного слабого GARCH-процесса, так что методы оценивания для дискретных слабых GARCH-процессов переносятся на определенные параметры непрерывных слабых GARCH-процессов.

Один значительный недостаток слабых GARCH-процессов состоит в том, что линейные проекции сложно интерпретировать как концепции волатильности, и, как подчеркивают Meddahi & Renault (1996, 2004), модели ценообразования активов основаны на условных матожиданиях, а не линейных проекциях. Таким образом, хотя разница между линейной проекцией и истинным условным матожиданием может быть численно небольшой в реалистичных приложениях, этот факт ограничивает применимость и формальную интерпретацию класса слабых GARCH-процессов.

12.2 Диффузионный предел GARCH-модели

Анализ ARCH-моделей можно упростить путем приближения стохастических разностных уравнений более удобными в обращении стохастическими дифференциальными уравнениями. Nelson (1990b) изучил предел ARCH-моделей в непрерывном времени. В частности, он предложил условия, при которых система стохастических разностных уравнений для ARCH-модели сходится по распределению к процессу Ито по мере приближения длины дискретного интервала к нулю. Он показал, что разные классы GARCH-процессов, например, GARCH(1, 1) и экспоненциальный ARCH-процесс, после надлежащей перепараметризации сходятся по распределению к двумерной невырожденной диффузии по мере сокращения интервала. Обычный метод построения непрерывных процессов из дискретных состоит в диффузионном приближении. Берется ряд в дискретном времени, определенный на сетке (например, $h\mathbb{N}$) с $h \downarrow 0$, и расширяется на пространство между точками сетки подходящим способом (например, с помощью интерполяции или кусочно-постоянной аппроксимации) в надежде, что эта последовательность процессов, определенная на $[0, \infty)$, слабо сходится к некоторому предельному процессу (см. Lindner, 2009b). Nelson (1990a) получил диффузионные пределы для GARCH(1, 1) и EGARCH-процессов. Он показал, что GARCH(1, 1) и EGARCH после надлежащей перепараметризации сходятся по распределению к двумерной невырожденной диффузии, по мере уменьшения длины временного интервала h . Предел в непрерывном времени может дать удобные приближения для прогноза и других моментов, когда дискретная модель приводит к трудным для анализа распределениям.

Дискретный GARCH(1, 1)-процесс имеет вид

$$p_k - p_{k-1} = \sigma_k z_k, \quad (107)$$

$$\sigma_k^2 = \omega + (\alpha_1 z_{k-1}^2 + \beta_1) \sigma_{k-1}^2, \quad k \in \mathbb{N}. \quad (108)$$

Для демонстрации результатов положим, что для каждого дискретного временного интервала $h > 0$ $\{z_{kh,h}\}_{k \in \mathbb{N}} \sim IIDN(0, 1)$, $\omega_h, \alpha_{1,h} > 0$ и $\beta_{1,h} \geq 0$, и пусть $\{p_{0,h}, \sigma_{0,h}^2\}$ – начальные случайные переменные, не зависящие от $\{z_{kh,h}\}_{k \in \mathbb{N}}$. Логарифм доходности и волатильность $\{p_{kh,h} - p_{(k-1)h,h}, \sigma_{kh,h}^2\}_{k \in \mathbb{N}}$ задаются рекурсивно уравнениями

$$p_{kh,h} - p_{(k-1)h,h} = h^{1/2} \sigma_{kh,h} z_{kh,h}, \quad k \in \mathbb{N}, \quad (109)$$

$$\sigma_{kh,h}^2 = \omega_h + (\alpha_{1,h} z_{(k-1)h,h}^2 + \beta_{1,h}) \sigma_{(k-1)h,h}^2, \quad k \in \mathbb{N}, \quad (110)$$

являющимися приближением GARCH(1, 1)-процесса в (107). Диффузионные модели, аппроксимирующие ARCH-модели, должны удовлетворять условию стационарности. Последовательность многомерных процессов в дискретном времени, индексированную временем h между двумя последовательными наблюдениями, можно расширить на случай непрерывного времени, предполагая, что они остаются постоянными между двумя последовательными моментами наблюдения. Вектор $\{p_{kh,h}, \sigma_{kh,h}^2\}_{k \in \mathbb{N}}$ включается в непрерывный процесс $\{p_{t,h}, \sigma_{t,h}^2\}_{t \geq 0}$ путем определения

$$p_{t,h} := p_{kh,h}, \quad \sigma_{t,h}^2 := \sigma_{kh,h}^2, \quad kh \leq t < (k+1)h, \quad k \in \mathbb{N}.$$

Последний процесс имеет выборочную траекторию в пространстве Скорохода $D([0, \infty), \mathbb{R}^2)$ \mathbb{R}^2 -значных кадлаг-функций, определенных на $[0, \infty)$. Распределение $\{p_{t,h}, \sigma_{t,h}^2\}_{t \geq 0}$ характеризуется распределением дискретных временных интервалов $\{p_{kh,h}, \sigma_{kh,h}^2\}_{k \in \mathbb{N}}$ и может быть задано последовательностью условных распределений. Поскольку интерес представляет предельное дифференциальное уравнение первого порядка, естественно наложить ограничения на распределение дискретного процесса. В частности, оно должно быть марковским первого порядка. Тогда распределение $\{p_{kh,h}, \sigma_{kh,h}^2\}_{k \in \mathbb{N}}$ характеризуется начальным распределением

и условным распределением при данном последнем наблюдаемом значении. Далее вычислим первые два условных момента, а затем при некоторых технических условиях применим теоремы о слабой сходимости марковских цепей к диффузионным процессам из Strook & Varadhan (1979). Идея заключается в том, чтобы обеспечить сходимость смещения и волатильности по единице времени при $h \downarrow 0$.

Пусть $\mathcal{F}_{(k-1)h} = \sigma(p_{(k-1)h,h}, \sigma_{(k-1)h,h}^2)$ – сигма-алгебра, порожденная $(p_{(k-1)h,h}, \sigma_{(k-1)h,h}^2)$. Заметим, что $\forall k$ и любого заданного h , $\{p_{kh,h} - p_{(k-1)h,h}\}, \{\sigma_{kh,h}^2 - \sigma_{(k-1)h,h}^2\}$ – марковские цепи первого порядка.

$$h^{-1}\mathbb{E} [p_{kh,h} - p_{(k-1)h,h} | \mathcal{F}_{(k-1)h}] = 0, \quad (111)$$

$$h^{-1}\mathbb{E} \left[(p_{kh,h} - p_{(k-1)h,h})^2 | \mathcal{F}_{(k-1)h} \right] = \sigma_{kh,h}^2, \quad (112)$$

$$h^{-1}\mathbb{E} \left[\sigma_{kh,h}^2 - \sigma_{(k-1)h,h}^2 | \mathcal{F}_{(k-1)h} \right] = h^{-1} \left[\omega_h + (\alpha_{1,h} + \beta_{1,h} - 1) \sigma_{(k-1)h,h}^2 \right], \quad (113)$$

$$\begin{aligned} h^{-1}\mathbb{E} \left[\left(\sigma_{kh,h}^2 - \sigma_{(k-1)h,h}^2 \right)^2 | \mathcal{F}_{(k-1)h} \right] &= h^{-1} \left[\omega_h^2 + (\alpha_{1,h} + \beta_{1,h} - 1)^2 \sigma_{(k-1)h,h}^4 \right] + \\ &+ 2h^{-1}\omega_h (\alpha_{1,h} + \beta_{1,h} - 1) \sigma_{(k-1)h,h}^2 + \\ &+ 2h^{-1}\alpha_{1,h}^2 \sigma_{(k-1)h,h}^4. \end{aligned} \quad (114)$$

Nelson (1990a) приводит условия слабой сходимости (обозначаемой \Rightarrow) $\{p_{t,h}, \sigma_{t,h}^2\}$ к некоторому процессу $\{p_t, \sigma_t^2\}$. Предположим, что имеются константы $\omega, \theta \in \mathbb{R}$ и $\alpha_1 > 0$, и обозначим за \mathcal{P}_0^h и \mathcal{P}_0 вероятностные меры для $\{p_{0,h}, \sigma_{0,h}^2\}$ и $\{p_0, \sigma_0^2\}$, соответственно. Если при $h \downarrow 0$ $\{p_{0,h}, \sigma_{0,h}^2\} \rightarrow \{p_0, \sigma_0^2\}$ или $\mathcal{P}_0^h \rightarrow \mathcal{P}_0$, то, при параметризации

$$\lim_{h \rightarrow 0} h^{-1}\omega_h = \omega, \quad (115)$$

$$\lim_{h \rightarrow 0} h^{-1}(1 - \alpha_{1,h} - \beta_{1,h}) = \theta, \quad (116)$$

$$\lim_{h \rightarrow 0} 2h^{-1}\alpha_{1,h}^2 = \alpha_1^2, \quad (117)$$

при $h \rightarrow 0$, $\{p_{t,h}, \sigma_{t,h}^2\} \Rightarrow \{p_t, \sigma_t^2\}$, где $\{p_t, \sigma_t^2\}$ – диффузионный процесс, являющийся решением

$$dp_t = \sigma_t dB_t, \quad (118)$$

$$d\sigma_t^2 = (\omega - \theta\sigma_t^2)dt + \alpha_1\sigma_t^2 dW_t, \quad (119)$$

где $\{B_t, W_t\}$ – два стандартные независимые броуновские движения, не зависящие от $\{p_0, \sigma_0^2\}$. Предельный процесс волатильности $\{\sigma_t^2\}$ – обобщенный процесс Орнштейна-Уленбека. Nelson (1990a) также показал, что (119) имеет строго стационарное решение $\{\sigma_t^2\}_{t \geq 0}$, если $2\theta/\alpha_1^2 > -1$ и $\omega > 0$, и тогда частное стационарное распределение σ_0^2 является обратным гамма-распределением с параметрами $1 + 2\theta/\alpha_1^2$ и $2\omega/\alpha_1^2$. Примером возможного выбора параметров, удовлетворяющего (115)–(117), является

$$\omega_h = \omega h, \quad (120)$$

$$\beta_{1,h} = 1 - \alpha_1\sqrt{h/2} - \theta h, \quad (121)$$

$$\alpha_{1,h} = \alpha_1\sqrt{h/2}. \quad (122)$$

Если волатильность не является торгуемым активом, невырожденность (119) ведет к неполноте рынков, поэтому модели ценообразования опционов требуют дополнительных предположений о предпочтениях и поведении инвесторов. Важное различие между GARCH-процессами в дискретном времени и их диффузионными пределами состоит в том, что квадрат процесса волатильности $\{\sigma_t^2\}_{t \geq 0}$ в (119) не зависит от броуновского движения $\{B_t\}_{t \geq 0}$, описывающего

процесс для логарифма цены. Процесс волатильности (118)–(119) имеет два независимых источника случайности, а именно $\{B_t\}_{t \geq 0}$ и $\{W_t\}_{t \geq 0}$. С другой стороны, дискретные GARCH-процессы определяются в терминах единственной последовательности шума $\{\varepsilon_t\}_{t \in \mathbb{N}}$, а не двух. Диффузионный предел Нельсона – пример непрерывного слабого GARCH-процесса, при условии существования конечного четвертого момента. Аппроксимацию Нельсона можно изменить для получения предельного процесса, задаваемого лишь одним броуновским движением (см. Corradi, 2000), но в этом случае предельный процесс волатильности детерминированный, что является нежелательным свойством процесса для цены.

Можно иначе взглянуть на проблему, предположив, что модель специфицирована в непрерывном времени как диффузионная модель. Можно задать ARCH-процесс, порождающий данные, похожий на истинный процесс, в том смысле, что распределение траекторий, генерируемое ARCH-моделью и диффузионной моделью сближаются при возрастающей мелкости дискретизации. Рассмотрим диффузионную модель, задаваемую уравнениями

$$dp_t = \left(\mu - \frac{\sigma_t^2}{2} \right) dt + \sigma_t dW_{1,t}, \quad (123)$$

$$d \ln \sigma_t^2 = \beta(\alpha - \ln \sigma_t^2) dt + \psi dW_{2,t}, \quad (124)$$

где $W_{1,t}, W_{2,t}$ – броуновские движения без смещения, независимые от (Y_0, σ_0^2) , такие, что $\text{Corr}(dW_{1,t}, dW_{2,t}) = \rho dt$. ARCH-приближение диффузии (131)–(132) имеет вид

$$y_{t+h} = y_t + \left(\mu - \frac{\sigma_t^2}{2} \right) h + h^{1/2} \sigma_t z_{t+h}, \quad (125)$$

$$\ln \sigma_{t+h}^2 = \ln \sigma_t^2 + \beta(\alpha - \ln \sigma_t^2) h + h^{1/2} g(z_{t+h}), \quad (126)$$

где $g(\cdot)$ – измеримая, $\mathbb{E}[|g(z_{t+h})|^{2+\delta}] < \infty$ для некоторого $\delta > 0$, и

$$\mathbb{V} \begin{bmatrix} z_t \\ g(z_t) \end{bmatrix} = \begin{bmatrix} 1 & \rho\psi \\ \rho\psi & \psi^2 \end{bmatrix}, \quad (127)$$

$$g(z_t) = \rho\psi z_t + \psi \left(\frac{1 - \rho^2}{1 - 2/\pi} \right)^{1/2} \left(|z_t| - \left(\frac{2}{\pi} \right)^{1/2} \right),$$

что точно соответствует EGARCH-модели. Альтернативная функция

$$g(z_t) = \rho\psi z_t + \psi \left(\frac{1 - \rho^2}{2} \right)^{1/2} (z_t^2 - 1)$$

также удовлетворяет уравнению (127). Эта спецификация является асимптотически оптимальным фильтром при $h \downarrow 0$, как обсуждается в Nelson & Foster (1994; 1995).

12.3 ARCH-модели как фильтры

Асимптотику для непрерывных моделей, предложенную в Nelson (1990a), можно применить к ряду других задач, возникающих при использовании ARCH-моделей. Nelson (1992) применяет эти результаты к GARCH(1, 1)- и EGARCH-моделей. Он исследует свойства оценок условной ковариационной матрицы, генерируемых неверно специфицированной ARCH-моделью. Nelson (1992) показывает, что при достаточно слабых условиях регулярности, когда диффузия наблюдается через дискретные интервалы длины h , разность между оценкой условной мгновенной дисперсии на основе GARCH(1, 1)-модели или EGARCH-модели и истинным значением сходится по вероятности к нулю при $h \downarrow 0$. Термин «оценка» соответствует его использованию в литературе о фильтрах, а не в статистике.

Если процесс хорошо приближается диффузией, широкий класс ARCH-моделей на удивление робастен к определенным типам ошибок спецификации. ARCH-модели могут допускать серьезные ошибки в спецификации условного среднего процесса и динамического поведения условной дисперсии. Ошибки могут быть настолько серьезными, что ARCH-модели лишены смысла как процессы, порождающие данные, и дают ужасные средне- и краткосрочные прогнозы, не влияя на состоятельность оценок условной дисперсии на один шаг вперед (Nelson, 1992). Условия регулярности требуют, чтобы условное распределение наблюдаемого ряда обладало не слишком толстыми хвостами, а условная дисперсия гладко менялась во времени. Предположим, что дискретный процесс, порождающий данные, – это AR(1)-EGARCH(1, 0):

$$p_{h,kh} = p_{(k-1)h,h} + h \cdot \mu_h(p_{(k-1)h,h}, \sigma_{kh,h}^2, kh) + h^{1/2} \sigma_{kh,h} z_{kh,h}, \quad k \in \mathbb{N}, \quad (128)$$

$$\ln(\sigma_{(k+1)h,h}^2) = \ln(\sigma_{kh,h}^2) - \beta_h(\ln(\sigma_{kh,h}^2) - \alpha_h) + \theta_h z_{kh,h} + \gamma_h(|z_{kh,h}| - m), \quad (129)$$

где $m = \sqrt{2/\pi}$. Nelson (1990a) показал, что при $\mu_h \rightarrow \mu$ равномерно на компакте

$$\beta_h = \beta + o(1), \quad \alpha_h = \alpha + o(1), \quad \theta_h = \theta + o(1), \quad \gamma_h = \gamma + o(1). \quad (130)$$

Если начальные значения $p_{0,h}$ и $\ln(\sigma_0^2)$ сходятся по распределению, и μ удовлетворяет достаточно слабым условиям регулярности, то при $h \downarrow 0$ последовательность процессов ступенчатых функций $\{p_{t,h}, \sigma_{t,h}^2\}_{h \downarrow 0}$, порождаемая (128)–(130), слабо сходится к диффузии

$$dp_t = \mu(p_t, \sigma_t^2, t) dt + \sigma_t dW_{1,t}, \quad (131)$$

$$d \ln \sigma_t^2 = \beta(\alpha - \ln \sigma_t^2) dt + \psi dW_{2,t}, \quad (132)$$

где $W_{1,t}$ и $W_{2,t}$ – одномерные броуновские движения без смещения и

$$\begin{bmatrix} dW_{1,t} \\ dW_{2,t} \end{bmatrix} \begin{bmatrix} dW_{1,t} & dW_{2,t} \end{bmatrix} = \begin{bmatrix} 1 & \theta \\ \theta & \theta^2 + (1 - m^2)\gamma^2 \end{bmatrix}. \quad (133)$$

Предположим теперь, что данные порождаются процессом

$$p_t = p_0 + \int_0^t \mu(p_s) ds + \int_0^t \sigma(p_s) ds,$$

и для получения оценок истинного мгновенного процесса условной дисперсии $\{\sigma_t^2(p_t)\}$ применяется модель EGARCH. Значения $\{\hat{\sigma}_{kh,h}^2\}$ находятся рекурсивно с использованием (129) с некоторым начальным $\hat{\sigma}_{0,h}^2$, а $\hat{z}_{kh,h}$ находятся по формуле

$$\hat{z}_{kh,h} \equiv (p_{kh,h} - p_{(k-1)h,h} - h \cdot \mu_h(p_{(k-1)h,h}, \hat{\sigma}_{kh,h}^2, kh)) / \hat{\sigma}_{kh,h}.$$

Если модель неверно специфицирована, $\hat{z}_{kh,h}$ не является ни IID, ни $N(0, 1)$. Требования на $\hat{\mu}_h, \beta_h, \alpha_h, \theta_h$ и γ_h для получения состоятельного фильтрования гораздо слабее, чем в (130), см. Nelson (1992). Определяя ошибку измерения как

$$Y_{t,h} \equiv \ln(\hat{\sigma}_{t,h}^2) - \ln(\sigma(p_t)^2),$$

Nelson (1992) дает условия, при которых $Y_{t,h} \xrightarrow{p} 0$. Эффект неверной спецификации условного среднего исчезает из условной дисперсии при $h \downarrow 0$.

В принципе ARCH-модель можно считать лаконичной моделью ожидаемой волатильности как явной функции от дискретно наблюдаемой доходности, то есть параметризованного условного матожидания. Тем не менее, как отмечает Nelson (1992), этим же самым дискретным параметрическим моделям можно дать непараметрическую интерпретацию как фильтрам для получения информации о (латентной) непрерывной волатильности. Предполагая, что выборочная траектория цены и соответствующий процесс непрерывной волатильности

непрерывны, – хотя формально неверно специфицированные при дискретной частоте выборки, – определенные надлежащим образом последовательности ARCH-моделей могут тем не менее служить состоятельными оценками волатильности истинной диффузии, в том смысле, что разность между истинной непрерывной волатильностью и оценками ARCH-фильтра сходится по вероятности к нулю по мере того как интервал частоты выборки уменьшается. Для иллюстрации рассмотрим GARCH(1, 1)-фильтр для доходности за период h , определенный в Nelson (1992). Обозначая за $r_{kh,h} \equiv p_{kh,h} - p_{(k-1)h,h}$ доходность за период kh , имеем

$$\begin{aligned}\hat{\sigma}_{kh,h}^2 &= \omega_h + \alpha_{1,h} r_{(k-1)h,h}^2 + \beta_{1,h} \hat{\sigma}_{(k-1)h,h}^2 \\ &= \omega_h (1 - \beta_{1,h})^{-1} + \sum_{i=1}^{\infty} \alpha_{1,h} \beta_{1,h}^i r_{(k-i)h,h}^2, \quad k \in \mathbb{N},\end{aligned}\quad (134)$$

где ω_h , $\alpha_{1,h}$ и $\beta_{1,h}$ удовлетворяют (120)–(121). Этот фильтр достигает состоятельности для σ_t^2 , чей процесс имеет непрерывные траектории. Однако скачки и разрывы в процессах для цены или диффузии исключены.

Мотивация ARCH-фильтров и сглаживателей, разработанных в Nelson (1992; 1996; 1996b), Nelson & Foster (1994; 1995) и Nelson & Schwartz (1992), становится ясной (см. Andersen et al., 2009), если сравнить их с выборочной оценкой σ_t^2 , найденной в предположении, что наблюдения цены доступны в $N + 1$ равномерно распределенных точках временного интервала $[0, Nh]$, то есть в $0, h, 2h, \dots, Nh$:

$$\hat{\sigma}_t^2 = \frac{1}{Nh} \sum_{k=1}^N r_{kh,h}^2. \quad (135)$$

Это состоятельная оценка непрерывной волатильности, если $N \rightarrow \infty$ и $Nh \rightarrow 0$ с соответствующими скоростями. Компромисс между длиной интервала выборки, $Nh \rightarrow 0$, и количеством наблюдений, $N \rightarrow 0$, аналогичен стандартному компромиссу между смещением и дисперсией, возникающему при непараметрическом ядерном оценивании. GARCH(1, 1)-фильтр работает так же, как и выборочная оценка дисперсии в (135), используя бесконечную взвешенную среднюю квадратов доходности, получаемых с растущей частотой вплоть до момента Nh . В случае GARCH-фильтров асимптотика все увеличивающегося числа наблюдений в интервале, $N \rightarrow \infty$, при уменьшающейся длине временных интервалов, $Nh \rightarrow 0$, достигается одним асимптотическим правилом, диктующим частоту выборки доходности и одновременно занижающим вес более давних наблюдений квадратов доходности, $r_{(k-i)h,h}^2$, для больших значений i .

Поскольку многие ARCH-модели могут служить состоятельными фильтрами для непрерывной волатильности, естественным образом возникает вопрос об эффективности. Nelson & Foster (1994) исследуют вопрос выбора ARCH-процесса для состоятельного и эффективного оценивания условной дисперсии диффузионного процесса, порождающего данные. Они аппроксимируют распределение ошибки измерения, возникающей при использовании приближенного ARCH-фильтра. Их результат позволяет сравнивать эффективность различных ARCH-фильтров и характеризовать асимптотически оптимальные ARCH-оценки условной дисперсии. Авторы выводят оптимальные ARCH-фильтры для трех диффузионных моделей и анализируют фильтрационные свойства нескольких GARCH-моделей. Nelson & Foster (1994) рассматривают случай, когда данные генерируются в непрерывном времени диффузией

$$\begin{aligned}d \begin{bmatrix} x_t \\ y_t \end{bmatrix} &= \begin{bmatrix} \mu(x_t, y_t, t) \\ \kappa(x_t, y_t, t) \end{bmatrix} dt + \\ &+ \begin{bmatrix} \sigma(y_t) & \rho(x_t, y_t, t) \Lambda(x_t, y_t, t) \sigma(y_t) \\ \Lambda(x_t, y_t, t) \sigma(y_t) & \Lambda(x_t, y_t, t)^2 \end{bmatrix}^{1/2} \begin{bmatrix} dW_{1,t} \\ dW_{2,t} \end{bmatrix},\end{aligned}\quad (136)$$

где $\{W_{1,t}, W_{2,t}\}$ – двумерное броуновское движение, x_0 и y_0 – случайные переменные, не зависящие от $\{W_{1,t}, W_{2,t}\}$, а $[\cdot]^{1/2}$ – квадратный корень матрицы. Функции $\mu, \kappa, \sigma, \rho$ и Λ предполагаются непрерывными и принимающими вещественные значения, $\Lambda(\cdot)$ и $\sigma(\cdot)$ положительны, $|\rho| \leq 1$, и $\sigma(y)$ дифференцируемая и строго возрастающая по y . Предполагается, что x_t наблюдается через дискретные интервалы времени длины h , а y_t – ненаблюдаемый процесс, контролирующий мгновенную условную дисперсию x_t , то есть $\sigma(y_t)^2$. Для получения оценки \hat{y}_t для y_t при данных дискретных наблюдениях $(x_0, x_h, x_{2h}, \dots, x_{[t/h]h})$ применяется ARCH-модель. Нормализованные инновации x_t и y_t определяются как

$$\xi_{x,t+h} \equiv h^{-1/2}(x_{t+h} - x_t - \mathbb{E}_t[x_{t+h} - x_t]), \quad (137)$$

$$\xi_{y,t+h} \equiv h^{-1/2}(y_{t+h} - y_t - \mathbb{E}_t[y_{t+h} - y_t]). \quad (138)$$

Nelson & Foster (1994) рассматривают класс ARCH-моделей, генерирующих \hat{y}_t (когда t – целое число, кратное h) рекурсивно:

$$\hat{y}_{t+h} = \hat{y}_t + h \cdot \hat{\kappa}(x_t, \hat{y}_t, t, h) + h^{1/2} \cdot g(\hat{\xi}_{x,t+h}, x_t, \hat{y}_t, t, h), \quad (139)$$

$$\hat{\xi}_{x,t+h} \equiv h^{-1/2}(x_{t+h} - x_t - h \cdot \hat{\mu}(x_t, \hat{y}_t, t, h)), \quad (140)$$

где $\hat{\kappa}(x_t, \hat{y}_t, t, h)$, $\hat{\mu}(x_t, \hat{y}_t, t, h)$ и $g(\hat{\xi}_{x,t+h}, x_t, \hat{y}_t, t, h)$ – выбираемые эконометристом функции, $\hat{\kappa}$ и $\hat{\mu}$ непрерывны по всем аргументам, а g дифференцируема по \hat{y}_t , $\hat{\xi}_{x,t+h}$ и h почти везде и всюду имеет односторонние производные. $\mathbb{E}_t[\cdot]$ – матожидание, условное на информации в момент времени t (то есть σ -алгебре, порожденной $\{x_\tau, y_\tau\}_{0 \leq \tau \leq t}$). Процесс $\{x_t, y_t\}$ марковский. Точно так же, как μ и κ – истинные смещения в x_t и y_t , $\hat{\mu}$ и $\hat{\kappa}$ – спецификации эконометриста (предположительно, неверные) для этих смещений, а $\hat{\xi}_{x,t+h}$ – остаток, полученный при использовании $\hat{\mu}$ вместо μ . ARCH-модель (139) воспринимает остаток $\hat{\xi}_x$ и оценки \hat{y} как если бы они были истинными. В ARCH-моделях шум в процессе $\{y_t\}$ является функцией оценки остатка наблюдаемого процесса. $g(\hat{\xi}_x, x, y, t, h)$ – спецификация этой функции эконометристом. Хотя процесс y_t ненаблюдаем, ARCH-модель воспринимает $h^{1/2}g(\hat{\xi}_x, x, \hat{y}, t, h)$ как истинную инновацию процесса y_t .

Выражение (139) включает частные случаи:

- GARCH(1, 1) при $y = \sigma^2$, $\hat{\kappa} \equiv \omega - \theta\sigma^2$, $g \equiv \alpha(\hat{\xi}_x^2 - \hat{\sigma}^2)$;
- EGARCH(1, 0) при $y = \ln(\sigma^2)$, $\hat{\kappa} \equiv -\beta(\ln(\hat{\sigma}^2) - \alpha)$, $g \equiv \theta\hat{\xi}_x/\hat{\sigma} + \gamma \left(|\hat{\xi}_x/\hat{\sigma}| - E_t|\xi_x/\sigma| \right)$;
- Модель Тейлора-Шверта при $y \equiv \sigma$, $\hat{\kappa} \equiv \omega - \theta\hat{\sigma}$, $g \equiv \alpha\hat{\sigma} - \left(|\hat{\xi}_x/\hat{\sigma}| - E_t|\xi_x/\sigma| \right)$.

Для целого t , кратного h , например, $t = [t/h] \cdot h$, определим нормализованную ошибку измерения в y :

$$q_t \equiv h^{-1/4}(\hat{y}_t - y_t),$$

которая является случайной ступенчатой функцией со скачками через временные интервалы длины h . Nelson & Foster (1994), меняя масштаб времени, получают диффузионный предел и асимптотическое распределение процесса для ошибки измерения, $\{q_t\}$, который зависит от $\rho(\cdot)$, $\sigma(\cdot)$, $\Lambda(\cdot)$ и $g(\cdot)$, но не зависит от $\hat{\kappa}$, κ , μ или $\hat{\mu}$. В то время как ошибки в компонентах дрейфа $\hat{\kappa}$, κ , μ и $\hat{\mu}$ влияют на $\{q_t\}$ при фиксированных $h > 0$, это влияние асимптотически исчезает при $h \downarrow 0$.

Nelson & Foster (1994) находят условия, при которых асимптотическая дисперсия $\{q_t\}$ минимальна (Теорема 4.1). Они также показывают, что GARCH(1, 1) является оптимальным фильтром, если данные генерируются моделью

$$dx_t = \mu dt + \sigma_t dW_{1,t}, \quad (141)$$

$$d\sigma_t^2 = (\omega - \theta\sigma_t^2)dt + \sqrt{2}\sigma_t^2 dW_{2,t}. \quad (142)$$

Асимптотически оптимальный фильтр полагает $\hat{\mu} = \mu$,

$$\hat{\sigma}_{t+h}^2 = \hat{\sigma}_t^2 + (\omega - \theta\hat{\sigma}_t^2) \cdot h + h^{1/2}(\xi_{x,t+h}^2 - \hat{\sigma}_t^2), \quad (143)$$

что узнаваемо как GARCH(1, 1), если переписать в виде

$$\hat{\sigma}_{t+h}^2 = \omega \cdot h + (1 - \theta h - \alpha h^{1/2})\hat{\sigma}_t^2 + h^{1/2}\xi_{x,t+h}^2. \quad (144)$$

Если $Corr(dW_{1,t}, dW_{2,t}) = \rho dt$, где $\rho \neq 0$, GARCH(1, 1) не является оптимальным фильтром. Но можно показать, что модификация GARCH(1, 1), нелинейная асимметричная GARCH в уравнении (42), оптимальна. Более того, авторы рассматривают две модели из литературы по ценообразованию опционов. В каждой модели P_t – цена акции, а σ_t – мгновенная волатильность доходности. Цена наблюдается в дискретные интервалы длины h . В каждой модели уравнение цены имеет вид

$$dP_t = P_t \mu dt + P_t \sigma_t dW_{1,t}.$$

Модель 1:

$$d \ln(\sigma_t^2) = -\beta(d \ln(\sigma_t^2) - \alpha)dt + \psi dW_{1,t}. \quad (145)$$

Модель 2:

$$d\sigma_t^2 = -\beta(\sigma_t^2 - \alpha)dt + \psi \sigma_t dW_{2,t}. \quad (146)$$

Пусть $x_t \equiv \ln(S_t)$ и $y_t \equiv \ln(\sigma_t^2)$. Тогда Лемма Ито позволяет записать Модель 1 в виде

$$dx_t = (\mu - \exp(y_t)/2)dt + \exp(y_t/2)dW_{1,t}, \quad (147)$$

$$dy_t = -\beta(y_t - \alpha)dt + \psi dW_{2,t}. \quad (148)$$

Асимптотически оптимальная ARCH-модель для модели (147)–(148) имеет вид (см. Теорему 4.4 в Nelson & Foster, 1994):

$$\hat{\mu}(x, \hat{y}) \equiv \mu - \exp(\hat{y})/2, \quad (149)$$

$$\begin{aligned} \hat{y}_{t+h} &= \hat{y}_t - \beta(\hat{y}_t - \alpha) \cdot h + h^{1/2}\psi \left(\rho \hat{\xi}_{x,t+h} \cdot \exp(-\hat{y}_t/2) + \right. \\ &\quad \left. + ((1 - \rho^2)/2)^{1/2} \left(\hat{\xi}_{x,t+h}^2 \exp(-\hat{y}_t) - 1 \right) \right). \end{aligned} \quad (150)$$

Получаемая минимальная дисперсия $h^{-1/4}(\sigma(\hat{y})^2 - \sigma(y)^2)$ равна $(2(1 - \rho^2))^{1/2}\psi\sigma^4$. В диффузионном пределе первые два условных момента полностью характеризуют процесс, и оптимальный ARCH-фильтр подгоняет эти моменты. Модель 2 состоит из уравнений

$$dx_t = (\mu - \exp(y_t)/2)dt + \exp(y_t/2)dW_{1,t}, \quad (151)$$

$$dy_t = (-\beta + \exp(-y_t)(\beta\alpha - \psi^2/2))dt + \psi \exp(-y_t/2)dW_{2,t}. \quad (152)$$

Пусть $2\beta\alpha > \psi^2$. Асимптотически оптимальный фильтр задается уравнением (149) и

$$\begin{aligned} \hat{y}_{t+h} &= \hat{y}_t + (-\beta + \exp(-\hat{y}_t)(\beta\alpha - \psi^2/2)) \cdot h + h^{1/2}\psi \exp(-\hat{y}_t/2) \\ &\quad \times \left(\rho \hat{\xi}_{x,t+h} \cdot \exp(-\hat{y}_t/2) + ((1 - \rho^2)/2)^{1/2} \cdot (\hat{x}_{x,t+h}^2 \exp(-\hat{y}_t) - 1) \right). \end{aligned} \quad (153)$$

Минимальная асимптотическая дисперсия $h^{-1/4}(\sigma(\hat{y})^2 - \sigma(y)^2)$ равна $(2(1 - \rho^2))^{1/2}\psi\sigma^3$. В (145) и соответствующей оптимальной ARCH-модели условная дисперсия σ_t^2 линейно растет по σ_t^4 , в то время как в (146) условная дисперсия σ_t^2 линейно растет по σ_t^2 . Большинство применяемых ARCH-моделей в сущности предполагают, что «дисперсия дисперсии» линейно растет по σ_t^4 . Если (146) порождает данные, то GARCH, EGARCH и другие подобные модели являются неэффективными фильтрами, если σ^2 очень мала или очень велика, поскольку

функции $g(\cdot)$ в этих ARCH-моделях неспособны подогнать истинную «дисперсию дисперсии» всюду в пространстве состояний. Одно важное свойство оптимального ARCH-фильтра касается способности отражать истинный процесс, порождающий данные. В частности, если данные были сгенерированы асимптотически оптимальным ARCH-фильтром, функциональная форма для второго условного момента переменных состояния будет такой же, как в истинном процессе, порождающем данные. Если условные первые моменты также совпадают, смещение второго порядка аналогично исчезает. Nelson & Foster (1995) доказывают желаемое свойство, что прогнозы, получаемые на основе возможно неверно специфицированной ARCH-модели, приближаются к прогнозам в истинной модели при $h \downarrow 0$. Таким образом, даже если ARCH-модели неверно специфицированы, они могут давать состоятельные оценки условной дисперсии. Однако поведение ARCH-фильтров с оцененными, а не известными параметрами, и свойства самих оценок параметров еще недостаточно изучены.

12.4 Модель GARCH в непрерывном времени

Непрерывная GARCH(1, 1)-модель (COGARCH), предложенная в Klüppelberg et al. (2004), основана на предельном переходе в явном представлении дискретного GARCH(1, 1)-процесса для получения непрерывного аналога. Начиная с (18), рекурсивная подстановка дает

$$\begin{aligned}\sigma_t^2 &= \omega \sum_{i=0}^{t-1} \prod_{j=i+1}^{t-1} (\beta_1 + \alpha_1 z_j^2) + \sigma_0^2 \prod_{j=0}^{t-1} (\beta_1 + \alpha_1 z_j^2) \\ &= \left(\omega \int_0^t \exp \left\{ - \sum_{j=0}^{\lfloor s \rfloor} \log (\beta_1 + \alpha_1 z_j^2) \right\} ds + \sigma_0^2 \right) \exp \left\{ \sum_{j=0}^{t-1} \log (\beta_1 + \alpha_1 z_j^2) \right\}, \\ p_t &= \sigma_t z_t = \sum_{j=0}^T z_j - \sum_{j=0}^{T-1} z_j,\end{aligned}$$

где $\lfloor s \rfloor$ означает наибольшее целое число, не превосходящее z , а $\sum_{j=0}^{t-1} \log (\beta_1 + \alpha_1 z_j^2) = t \log \beta_1 + \sum_{j=0}^{t-1} \log (1 + \alpha_1 z_j^2 / \beta_1)$ и $\sum_{j=0}^t \varepsilon_j$ – случайные блуждания, связанные таким образом, что первое можно восстановить из второго. Идея в Klüppelberg et al. (2004) состоит в том, чтобы далее заменить случайные блуждания на процессы Леви, являющиеся непрерывными аналогами случайных блужданий, а z_j заменить на скачки процесса Леви L . Авторы определяют вспомогательный процесс Леви ξ_t :

$$\xi_t = -t \log \beta_1 - \sum_{0 < s \leq t} \log \left(1 + \frac{\alpha_1}{\beta_1} (\Delta L_s)^2 \right), \quad t \geq 0.$$

При $\omega, \alpha_1, \beta_1 > 0$ процесс Леви имеет ненулевую меру Леви ν_L и соответствует случайному блужданию $-t(\log \delta + \sum_{j=0}^{t-1} \log (1 + \alpha_1 \varepsilon_j^2 / \alpha_1))$ в дискретном времени. При начальной случайной величине σ_0^2 , независимой от $\{L_t\}_{t \geq 0}$, определяются (непрерывный справа) процесс волатильности $\{\sigma_t\}_{t \geq 0}$ и модель COGARCH(1, 1):

$$\sigma_t^2 = \left(\omega \int_0^t \exp(\xi_s) ds + \sigma_0^2 \right) \exp(-\xi_t), \quad t \geq 0, \quad (154)$$

$$p_t = \int_0^t \sigma_{s-} dL_s, \quad t \geq 0. \quad (155)$$

Применение COGARCH позволяет проводить анализ неравномерно расположенных данных, не прибегая к приближениям, включающим оценивание отсутствующих значений, а при анализе высокочастотных данных неравномерное расположение – из-за уик-эндов, праздников

и перерывов между закрытием и открытием торгов – доминирующая черта данных. Процесс $\{\xi_t\}_{t \geq 0}$ действительно является процессом Леви, который противоположен по знаку субординатору с дрейфом $-\log \beta_1$. Следовательно, квадрат процесса волатильности снова является обобщенным процессом Орнштейна-Уленбека, движимым $\{\xi_t, \omega t\}$. Применяя Лемму Ито к (154), получаем, что $\{\sigma_t^2\}_{t \geq 0}$ удовлетворяет стохастическому дифференциальному уравнению

$$d\sigma_t^2 = (\omega + \log \beta_1 \sigma_{t-}^2)dt + \frac{\alpha_1}{\beta_1} \sigma_{t-}^2 d[L, L]_t^d,$$

где $[L, L]_t^d = \sum_{0 < s \leq t} (\Delta L_s)^2$ обозначает дискретную часть квадратичной вариации L . Заметим, что G имеет только один источник вариации, L , который движет как σ_t^2 , так и p_t . Скачок в L означает скачок в p_t , где размер скачка равен $\Delta p_t = \sigma_{t-} \Delta L_t$. Условия стационарности и существования моментов для $\{\sigma_t^2\}_{t \geq 0}$ даны в Klüppelberg et al. (2004), и из них следует, что $\{\sigma_t^2\}_{t \geq 0}$ имеет стационарное представление тогда и только тогда, когда

$$\int_{\mathbb{R}} \log(1 + \alpha_1 x^2 / \beta_1) \nu_L(dx) < -\log \beta_1,$$

что, в частности, требует $\delta < 1$. Доказано, что процесс COGARCH(1, 1) со стационарной волатильностью обладает многими свойствами дискретного GARCH(1, 1)-процесса. Как показано в Klüppelberg et al. (2004) и Klüppelberg et al. (2006), COGARCH-процесс имеет некоррелированные приращения, а автокорреляционные функции волатильности σ^2 и квадратов приращений p_t убывают экспоненциально. Что касается дискретных GARCH-процессов, стационарная COGARCH волатильность имеет Парето-хвосты при слабых предположениях, см. Klüppelberg et al. (2006). Brockwell et al. (2006) расширяют модель COGARCH(1, 1) до COGARCH(p, q).

Литература

- Andersen, T.G. (1996). Return volatility and trading volume: an information flow interpretation of stochastic volatility. *Journal of Finance* 51, 169–204.
- Andersen, T.G. & T. Bollerslev (1998). ARCH and GARCH models. Глава в *Encyclopedia of Statistical Sciences*, (под редакцией S. Kotz, C. Read & D. Banks), том 2. John Wiley and Sons.
- Andersen, T.G., T. Bollerslev, P. Christoffersen & F.X. Diebold (2006). Volatility and correlation forecasting. Глава в *Handbook of Economic Forecasting* (под редакцией C.W.J. Granger, G.Elliott & A. Timmermann), 777–878. North-Holland.
- Andersen, T.G. & T. Bollerslev (1997). Heterogeneous information arrivals and return volatility dynamics: uncovering the long-run in high frequency returns. *Journal of Finance* 52, 975–1005.
- Andersen, T.G., T. Bollerslev & F.X. Diebold (2009). Parametric and nonparametric volatility measurement. Глава в *Handbook of Financial Econometrics* (под редакцией Y. Ait-Sahalia & L. Hansen). North-Holland.
- Anderson, H., K. Nam & F. Vahid (1999). Asymmetric nonlinear smooth transition GARCH models. Глава в *Nonlinear Time Series Analysis of Economic and Financial Data* (под редакцией P. Rothman), 191–207. Kluwer.
- Bai, X., J.R. Russell & G.C. Tiao (2003). Kurtosis of GARCH and stochastic volatility models with non-normal innovations. *Journal of Econometrics* 114, 349–360.
- Baillie, R.T. & T. Bollerslev (1989). The message in daily exchange rates; a conditional-variance tale. *Journal of Business and Economic Statistics* 7, 297–305.
- Baillie, R.T., T. Bollerslev & H. Mikkelsen (1996). Fractionally integrated generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics* 74, 3–30.
- Bauwens, L., S. Laurent & J. Rombouts (2006). Multivariate GARCH models: a survey. *Journal of Applied Econometrics* 21, 79–110.
- Bera, A. & M. Higgins (1993). ARCH models: properties, estimation and testing. *Journal of Economic Surveys* 7, 305–362.

- Bera, A., M. Higgins & S. Lee (1992). Interaction between autocorrelation and conditional heteroskedasticity: a random-coefficient approach. *Journal of Business & Economic Statistics* 10, 133–142.
- Berkes, I. & L. Horváth (2003). The rate of consistency of the quasi-maximum likelihood estimator. *Statistics and probability letters* 61, 133–143.
- Berkes, I. & L. Horváth (2004). The efficiency of the estimators of the parameters in GARCH processes. *Annals of Statistics* 32, 633–655.
- Berkes, I., L. Horváth & P. Kokoszka (2003). GARCH processes: structure and estimation. *Bernoulli* 9, 201–227.
- Black, F. (1976). Studies in stock price volatility changes. Глава в *Proceedings of the 1976 Business Meeting of the Business and Economics Statistics Section, American Statistical Association*, 177–181.
- Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics* 31, 307–327.
- Bollerslev, T. (1987). A conditionally heteroscedastic time series model for speculative prices and rates of return. *Review of Economics and Statistics* 69, 542–547.
- Bollerslev, T. (1988). On the correlation structure for the generalized autoregressive conditional heteroskedastic process. *Journal of Time Series Analysis* 9, 121–131.
- Bollerslev, T. (2009). Glossary to ARCH (GARCH). Глава в *Volatility and Time Series Econometrics: Essays in Honour of Robert F. Engle* (под редакцией Т. Bollerslev, J. R. Russell & M. Watson). Oxford University Press.
- Bollerslev, T., R. Chou & K. Kroner. (1992). ARCH modeling in finance: a selective review of the theory and empirical evidence. *Journal of Econometrics* 52, 5–59.
- Bollerslev, T., R.F. Engle & D.B. Nelson (1994). ARCH models. Глава в *Handbook of Econometrics* (под редакцией R. Engle & D. McFadden), том 4. Elsevier Science.
- Bollerslev, T. & H.O. Mikkelsen (1996). Modeling and pricing long memory in stock market volatility. *Journal of Econometrics* 73, 151–184.
- Bollerslev, T. & J. Wooldridge (1992). Quasi-maximum likelihood estimation and inference in dynamic models with time-varying covariances. *Econometric Reviews* 11, 143–172.
- Bougerol, P. & N. Picard (1992a). Stationarity of GARCH processes and of some nonnegative time series. *Journal of Econometrics* 52, 115–127.
- Bougerol, P. & N. Picard (1992b). Strict stationarity of generalized autoregressive process. *The Annals of Probability* 20, 1714–1730.
- Brockwell, P. & R.A. Davis (1991). *Time Series: Theory and Methods*. Springer.
- Brockwell, P., C. Erdenebaatar & A.M. Lindner (2006). Continuous time GARCH processes. *The Annals of Applied Probability* 16, 790–826.
- Brooks, C., S. Burke & G. Pesand (2001). Benchmarks and the accuracy of GARCH model estimation. *International Journal of Forecasting* 17, 45–56.
- Cai, J. (1994). A Markov model of switching-regime ARCH. *Journal of Business & Economic Statistics* 12, 309–316.
- Corradi, V. (2000). Reconsidering the continuous time limit of the GARCH(1,1) process. *Journal of Econometrics* 96, 145–153.
- Dacorogna, M. M., U.A. Müller, R.J. Nagler, R.B. Olsen & O.V. Pictet (1993). Geographical model for the daily and weekly seasonal volatility in the foreign exchange market. *Journal of International Money and Finance* 12, 413–438.
- Davidian, M. & R. Carroll (1987). Variance function estimation. *Journal of American Statistical Association* 82, 1079–1091.
- DeGiannakis, S. & E. Kekalaki (2004). Autoregressive conditional heteroscedasticity (ARCH) models: a review. *Quality Technology and Quantitative Management* 1, 271–324.
- Diebold, F. (1986). Modeling the persistence of conditional variances: a comment. *Econometric Reviews* 5, 51–56.
- Diebold, F. (2004). The Nobel memorial prize for Robert F. Engle. *Scandinavian Journal of Economics* 106, 165–185.
- Diebold, F. & A. Inoue (2001). Long memory and regime switching. *Journal of Econometrics* 105, 131–159.

- Diebold, F. & J. Lopez (1995). Modeling volatility dynamics. Глава в *Macroeconometrics: Developments, Tensions and Prospects* (под редакцией К. Hoover), 427–472. Kluwer Academic Press.
- Diebold, F. and T. Schuermann (1993). Exact maximum likelihood Estimation of ARCH models. Технический доклад, Department of Economics, University of Pennsylvania.
- Ding, Z., C. Granger & R. Engle (1993). A long memory property of stock market returns and a new model. *Journal of Empirical Finance* 1, 83–106.
- Drost, F.C. & T. E. Nijman (1993). Temporal aggregation of GARCH processes. *Econometrica* 61, 909–927.
- Drost, F.C. & C.A.J. Klaassen (1997). Efficient estimation in semiparametric GARCH models. *Journal of Econometrics* 81, 193–221.
- Drost, F.C. & B.J.M. Werker (1996). Closing the GARCH gap: Continuous time GARCH modeling. *Journal of Econometrics* 74, 31–57.
- Duan, J.-C. (1997). Augmented GARCH(p, q) process and its diffusion limit. *Journal of Econometrics* 79, 97–127.
- Dueker, J. (1997). Markov switching in GARCH processes and mean-reverting stock market volatility. *Journal of Business & Economic Statistics* 15, 26–34.
- Engle, R.F. (1982). Autoregressive conditional heteroskedasticity with estimates of the variance of U.K. inflation. *Econometrica* 50, 987–1008.
- Engle, R.F. (1990). Discussion: stock market volatility and the crash of 87. *Review of Financial Studies* 3, 103–106.
- Engle, R.F. (2001). GARCH 101: The use of ARCH/GARCH models in applied econometrics. *Journal of Economic Perspectives* 15, 157–168.
- Engle, R.F. (2002). New frontiers for ARCH models. *Journal of Applied Econometrics* 17, 425–446.
- Engle, R.F. (2004). Nobel lecture. Risk and volatility: econometric models and financial practice. *American Economic Review* 94, 405–420.
- Engle, R.F. & T. Bollerslev (1986). Modeling the persistence of conditional variances. *Econometric Reviews* 5, 1–50.
- Engle, R.F. & G. Gonzalez-Rivera (1991). Semiparametric ARCH models. *Journal of Business & Economic Statistics* 19, 3–29.
- Engle, R.F. & R.G. Lee (1999). A permanent and transitory component model of stock return volatility. Глава в *Cointegration, Causality, and Forecasting: A Festschrift in Honor of Clive W.J. Granger* (под редакцией R. Engle & H. White), 475–497. Oxford University Press.
- Engle, R.F., D. Lilien & R. Robins (1987). Estimating time varying risk premia in the term structure: the ARCH-M model. *Econometrica* 55, 391–407.
- Engle, R.F. & J. Mezrich (1996). GARCH for groups. *Risk* 9, 36–40.
- Engle, R.F. & V.K. Ng (1993). Measuring and testing the impact of news on volatility. *Journal of Finance* 48, 1749–1778.
- Engle, R.F. & A.J. Patton (2001). What good is a volatility model? *Quantitative Finance* 1, 237–245.
- Fernández, C. & M. Steel (1998). On Bayesian modeling of fat tails and skewness. *Journal of the American Statistical Association* 93, 359–371.
- Francq, C., M. Roussignol & J. Zakoïan (2001). Conditional heteroskedasticity driven by hidden Markov chains. *Journal of Time Series Analysis* 22, 197–220.
- Francq, C. & J. Zakoïan (2005). The L2-structures of standard and switching-regime GARCH models. *Stochastic Processes and their Applications* 115, 1557–1582.
- Francq, C. & J. Zakoïan (2009). A tour in the asymptotic theory of GARCH estimation. Глава в *Handbook of Financial Time Series* (под редакцией T. Andersen, R.A. Davis, J.-P. Kreiss & T. Mikosch). Springer.
- Giraitis, L., P. Kokoszka & R. Leipus (2000). Stationary ARCH models: dependence structure and Central Limit Theorem. *Econometric Theory* 16, 3–22.
- Giraitis, L., R. Leipus & D. Surgailis (2007). Recent advances in ARCH modelling. Глава в *Long Memory in Economics* (под редакцией G. Teyssière & A.P. Kirman). Springer.
- Glosten, L., R. Jagannathan & D. Runkle (1993). On the relationship between the expected value and the volatility of the nominal excess return on stocks. *Journal of Finance* 48, 1779–1801.

- Gonzalez-Rivera, G. (1998). Smooth transition GARCH models. *Studies in Nonlinear Dynamics and Econometrics* 3, 61–78.
- Gourieroux, C. (1997). *ARCH Models and Financial Applications*. Springer.
- Gourieroux, C., A. Monfort & A. Trognon (1984). Pseudo maximum likelihood methods: theory. *Econometrica* 52, 681–700.
- Granger, C. & Z. Ding (1996). Modeling volatility persistence of speculative returns. *Journal of Econometrics* 73, 185–215.
- Granger, C. & N. Hyung (2004). Occasional structural breaks and long memory with an application to the S&P500 absolute returns. *Journal of Empirical Finance* 11, 399–421.
- Granger, C. & T. Teräsvirta (1999). A simple nonlinear time series model with misleading linear properties. *Economics Letters* 62, 161–165.
- Granger, C. & R. Joyeux (1980). An introduction to long-memory time series models and fractional differencing. *Journal of Time Series Analysis* 4, 221–238.
- Gray, S. (1996). Modeling the conditional distribution of interest rates as a regime-switching process. *Journal of Financial Economics* 42, 27–62.
- Haas, M., S. Mittnik & M. Paolella (2004). A new approach to Markov-switching GARCH models. *Journal of Financial Econometrics* 2, 493–530.
- Hall, P. & Q. Yao (2003). Inference in ARCH and GARCH models with heavy-tailed errors. *Econometrica* 71, 285–317.
- Hamilton, J. (1989). A new approach to the economic analysis of nonstationary time series and business cycle. *Econometrica* 57, 357–384.
- Hamilton, J. (1990). Analysis of time series subject to changes in regime. *Journal of Econometrics* 45, 39–70.
- Hamilton, J. (1994). *Time Series Analysis*. Princeton University Press.
- Hamilton, J. & R. Susmel (1994). Autoregressive conditional heteroskedasticity and changes in regime. *Journal of Econometrics* 64, 307–333.
- Hansen, B. (1994). Autoregressive conditional density estimation. *International Economic Review* 35, 705–730.
- He, C. & T. Teräsvirta (1999). Properties of moments of a family of GARCH processes. *Journal of Econometrics* 92, 173–192.
- He, C., T. Teräsvirta & H. Malmsten (2002). Moment structure of a family of first-order exponential GARCH models. *Econometric Theory* 18, 868–885.
- Hentschel, L. (1995). All in the family. Nesting symmetric and asymmetric GARCH models. *Journal of Financial Economics* 39, 71–104.
- Higgins, M. & A. Bera (1992). A class of nonlinear ARCH models. *International Economic Review* 33, 137–158.
- Hosking, J. (1981). Fractional differencing. *Biometrika* 68, 165–176.
- Hsieh, D. (1989). Modelling heteroscedasticity in daily foreign exchange rates. *Journal of Business and Economic Statistics* 7, 307–317.
- Hurst, H. (1951). Long term storage capacity of reservoirs. *Transactions of the American Society of Civil Engineers* 116, 770–799.
- Jondeau, E. & M. Rockinger (2003). Conditional volatility, skewness and kurtosis: existence, persistence, and comovements. *Journal of Economic Dynamics and Control* 27, 1699–1737.
- Karanasos, M. & J. Kim (2006). A re-examination of the asymmetric power ARCH model. *Journal of Empirical Finance* 13, 113–128.
- Klaasen, F. (2002). Improving GARCH volatility forecasts with regime-switching GARCH. *Empirical Economics* 27, 363–394.
- Klüppelberg, C., A. Lindner & R. Maller (2004). A continuous-time GARCH process driven by a Lévy process: Stationarity and second order behaviour. *Journal of Applied Probability* 41, 601–622.
- Klüppelberg, C., A. Lindner & R. Maller (2006). Continuous time volatility modelling: COGARCH versus Ornstein models. Глава в *From Stochastic Calculus to Mathematical Finance. The Shiryayev Festschrift* (под редакцией Y. Kabanov, R. Liptser, & J. Stoyanov), 393–419. Springer.

- Lamoureux, C. & W. Lastrapes (1990). Persistence in variance, structural change and the GARCH model. *Journal of Business & Economic Statistics* 8, 225–234.
- Lanne, M. & P. Saikkonen (2005). Nonlinear GARCH models for highly persistent volatility. *Econometrics Journal* 8, 251–276.
- Lee, S. W. & B.E. Hansen (1994). Asymptotic theory for the GARCH(1,1) quasimaximum likelihood estimator. *Econometric Theory* 10, 29–52.
- Li, C. & W. Li (1996). On a double threshold autoregressive heteroskedasticity time series model. *Journal of Applied Econometrics* 11, 253–274.
- Lindner, A. (2009a). Continuous time approximations to GARCH and stochastic volatility models. Глава в *Handbook of Financial Time Series* (под редакцией Т. Andersen, R. Davis, J.-P. Kreiss & Т. Mikosch). Springer.
- Lindner, A. (2009b). Stationarity, mixing, distributional properties and moments of GARCH(p, q)-processes. Глава в *Handbook of Financial Time Series* (под редакцией Т. Andersen, R. Davis, J.-P. Kreiss & Т. Mikosch). Springer.
- Ling, S. (1999). On the probabilistic properties of a double threshold ARMA conditional heteroskedastic model. *Journal of Applied Probability* 36, 688–705.
- Ling, S. & M. McAleer (2002). Necessary and sufficient moment conditions for the GARCH(r, s) and asymmetric power GARCH(r, s) models. *Econometric Theory* 18, 722–729.
- Ling, S. & M. McAleer (2003). Adaptive estimation in nonstationary ARMA models with GARCH noises. *Annals of Statistics* 31, 642–674.
- Linton, O. (1993). Adaptive estimation in ARCH models. *Econometric Theory* 9, 539–564.
- Lumsdaine, R. (1996). Consistency and asymptotic normality of the quasi-maximum likelihood estimator in IGARCH(1,1) and covariance stationary GARCH(1,1) models. *Econometrica* 64, 575–596.
- Maheu, J. & T. McCurdy (2000). Identifying bear and bull markets in stock markets. *Journal of Business & Economic Statistics* 18, 100–112.
- Malmsten, H. & T. Teräsvirta (2004). Stylized facts on financial time series and three popular models of volatility. *SSE/EFI Working Paper Series in Economics and Finance* 563. Stockholm School of Economics.
- McCullough, B. & C. Renfro (1999). Benchmarks and software standards: a case study of GARCH procedures. *Journal of Economic and Social Measurement* 25, 59–71.
- Meddahi, N. & E. Renault (1996). Aggregation and Marginalization of GARCH and Stochastic Volatility Models. Технический доклад, Department of Economics, University of Montreal and CIRANO.
- Meddahi, N. & E. Renault (2004). Temporal aggregation of volatility models. *Journal of Econometrics* 119, 355–379.
- Mikosch, T. & C. Starica (1998). Change of structure in financial time series, long range dependence and the GARCH model. Technical report, University of Groningen.
- Mikosch, T. & C. Starica (2000). Is it really long memory we see in financial returns? Глава в *Extremes and Integrated Risk Management* (под редакцией Р. Embrechts). Risk Books.
- Mikosch, T. & C. Starica (2003). Long-range dependence effects and ARCH modeling. Глава в *Theory and Applications of Long-Range Dependence* (под редакцией Р. Doukhan, G. Oppenheim & M.S. Taqqu). Birkhäuser.
- Milhøj, A. (1985). The moment structure of ARCH processes. *Scandinavian Journal of Statistics* 12, 281–292.
- Mittnik, S., S.T. Rachev, T. Doganoglu & D. Chenyao (1999). Maximum likelihood estimation of stable paretian models. *Mathematical and Computer Modelling* 29, 275–293.
- Müller, U.A., M.M. Dacorogna, R. D. Davé, R.B. Olsen, O.V. Pictet & J.E. von Weizsäcker (1997). Volatilities of different time resolutions – analyzing the dynamics of market components. *Journal of Empirical Finance* 4, 213–239.
- Nelson, D. (1990a). ARCH models as diffusion approximations. *Journal of Econometrics* 45, 7–38.
- Nelson, D. (1990b). Stationarity and persistence in the GARCH(1,1) model. *Econometric Theory* 6, 318–334.
- Nelson, D. (1991). Conditional heteroskedasticity in asset returns: a new approach. *Econometrica* 59, 347–370.
- Nelson, D. (1992). Filtering and forecasting with misspecified ARCH models I: Getting the right variance with the wrong model. *Journal of Econometrics* 52, 61–90.
- Nelson, D. (1996a). Asymptotic filtering theory for multivariate ARCH models. *Journal of Econometrics* 71, 1–47.

- Nelson, D. (1996b). Asymptotically optimal smoothing with ARCH models. *Econometrica* 64, 561–573.
- Nelson, D. & C. Cao (1992). Inequality constraints in the univariate GARCH model. *Journal of Business & Economic Statistics* 10, 229–235.
- Nelson, D. & D. Foster (1994). Asymptotic filtering theory for univariate ARCH models. *Econometrica* 62, 561–573.
- Nelson, D. & D. Foster (1995). Filtering and forecasting with misspecified ARCH models II: Making the right forecast with the wrong model. *Journal of Econometrics* 67, 303–335.
- Nelson, D. & B. Schwartz (1992). Filtering with ARCH models: A Monte Carlo investigation. *Proceedings of the American Statistical Association, Business and Economic Statistics Section*, 1–6.
- Newey, W. & D. Steigerwald (1997). Asymptotic bias for quasi-maximum-likelihood estimators in conditional heteroscedasticity models. *Econometrica* 65, 587–599.
- Pagan, A. (1996). The econometrics of financial markets. *Journal of Empirical Finance* 3, 15–102.
- Pagan, A. & G. Schwert (1990). Alternative models for conditional volatility. *Journal of Econometrics* 45, 267–290.
- Palm, F. (1996). GARCH Models of volatility. Глава в *Handbook of Statistics* (под редакцией С. Rao & G. Maddala), том 14. North-Holland.
- Panorska, A.K., S. Mittnik & S.T. Rachev (1995). Stable GARCH models for financial time series. *Applied Mathematics Letters* 5, 33–37.
- Schwert, W. (1989). Why does stock market volatility change over time? *Journal of Finance* 44, 1115–1153.
- Sentana, E. (1995). Quadratic ARCH models. *Review of Economic Studies* 62, 639–661.
- Shephard, N. (1996). Statistical aspects of ARCH and stochastic volatility models. Глава в *Time Series Models in Econometrics, Finance and Other Fields* (под редакцией D. Cox, D. Hinkley & O. Barndorff-Nielsen), 1–67. Chapman & Hall.
- Starica, C. & C. Granger (2005). Non-stationarities in stock returns. *Review of Economics and Statistics* 87, 503–522.
- Straumann, D. (2005). *Estimation in conditionally heteroscedastic time series models*. Lecture notes in statistics. Springer.
- Strook, D. & S.R.S. Varadhan (1979). *Multidimensional Diffusion Processes*. Springer.
- Taylor, S. (1986). *Modeling Financial Time Series*. John Wiley and Sons.
- Teräsvirta, T. (2009). An introduction to univariate GARCH models. Глава в *Handbook of Financial Time Series* (под редакцией T. Andersen, R. Davis, J.-P. Kreiss & T. Mikosch). Springer.
- Theodossiou, P. (1998). Financial data and skewed generalized t distribution. *Management Science* 44, 1650–1661.
- Venter, J. & P.J. de Jongh (2002). Risk estimation using the normal inverse Gaussian distribution. *Journal of Risk* 4, 1–23.
- Weiss, A.A. (1986). Asymptotic theory for ARCH models: estimation and testing. *Econometric Theory* 2, 107–131.
- White, H. (1981). Maximum likelihood estimation of misspecified models. *Econometrica* 50, 1–25.
- Zaffaroni, P. (2007). Contemporaneous aggregation of GARCH processes. *Journal of Time Series Analysis* 28, 521–544.
- Zakoïan, J.-M. (1994). Threshold heteroskedastic models. *Journal of Economic Dynamics and Control* 18, 931–955.

Univariate GARCH Models: A Survey

Eduardo Rossi

University of Pavia, Pavia, Italy

This article presents a survey of the developments of univariate GARCH models. ARCH, GARCH, EGARCH and other possible nonlinear extensions are examined. Conditions for stationarity (weak and strong) are presented. Inference and testing is presented in the quasi-maximum likelihood framework. Continuous GARCH approximations are discussed.

