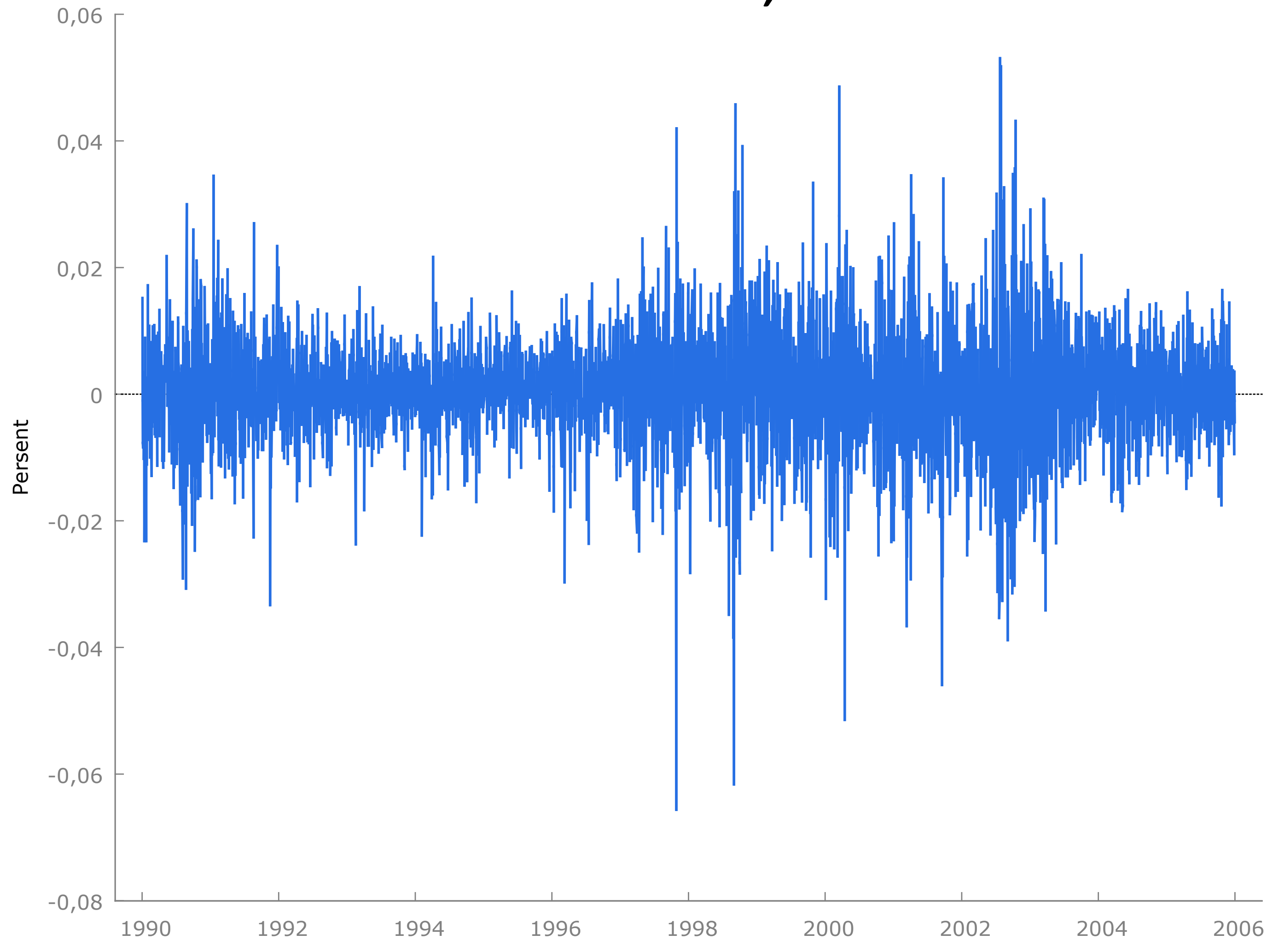
ARCH и GARCH модели

Ежедневные данные о темпах прироста индекса Нью-Йоркской фондовой биржи в 1990-2005 гг., %



Для некоторых временных рядов, характеризующих макроэкономические процессы, такие модели как *ARIMA* не в полной мере отражают закономерности, которым следует временной ряд и его характеристики.

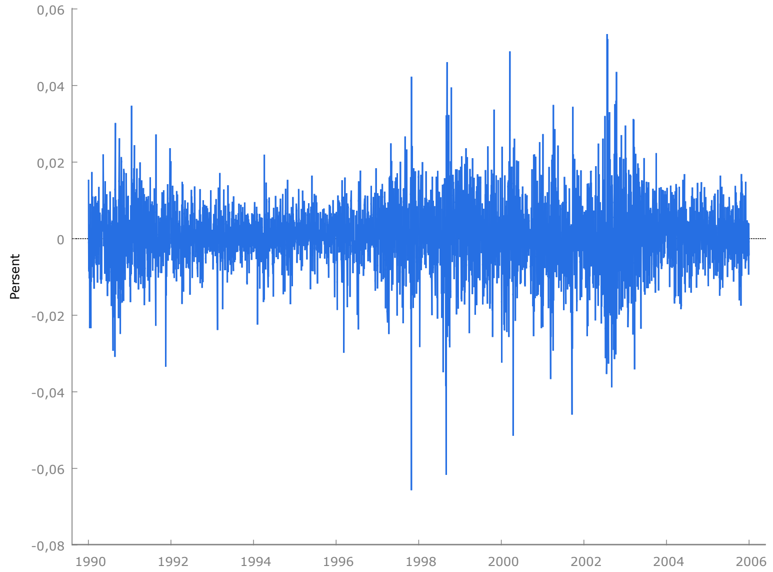
Так для финансовых временных рядов была выявлена закономерность в поведении уровней ряда, заключающаяся в том, что малые и большие значения уровней ряда группировались целыми сериями или кластерами. Несмотря на это, гипотеза о гомоскедастичности дисперсии для больших временных интервалов не противоречила данным.

Так получен класс моделей временных рядов, в которых учитываются изменения дисперсии и ковариаций – модели *волатильности*.

Под **волатильностью** (*volatility* – изменчивость) понимают статистический показатель, характеризующий степень вариабельности цены актива во времени.

Как правило, для оценки волатильности рассчитывается среднеквадратическое отклонение стоимости финансового актива за определенный период времени.

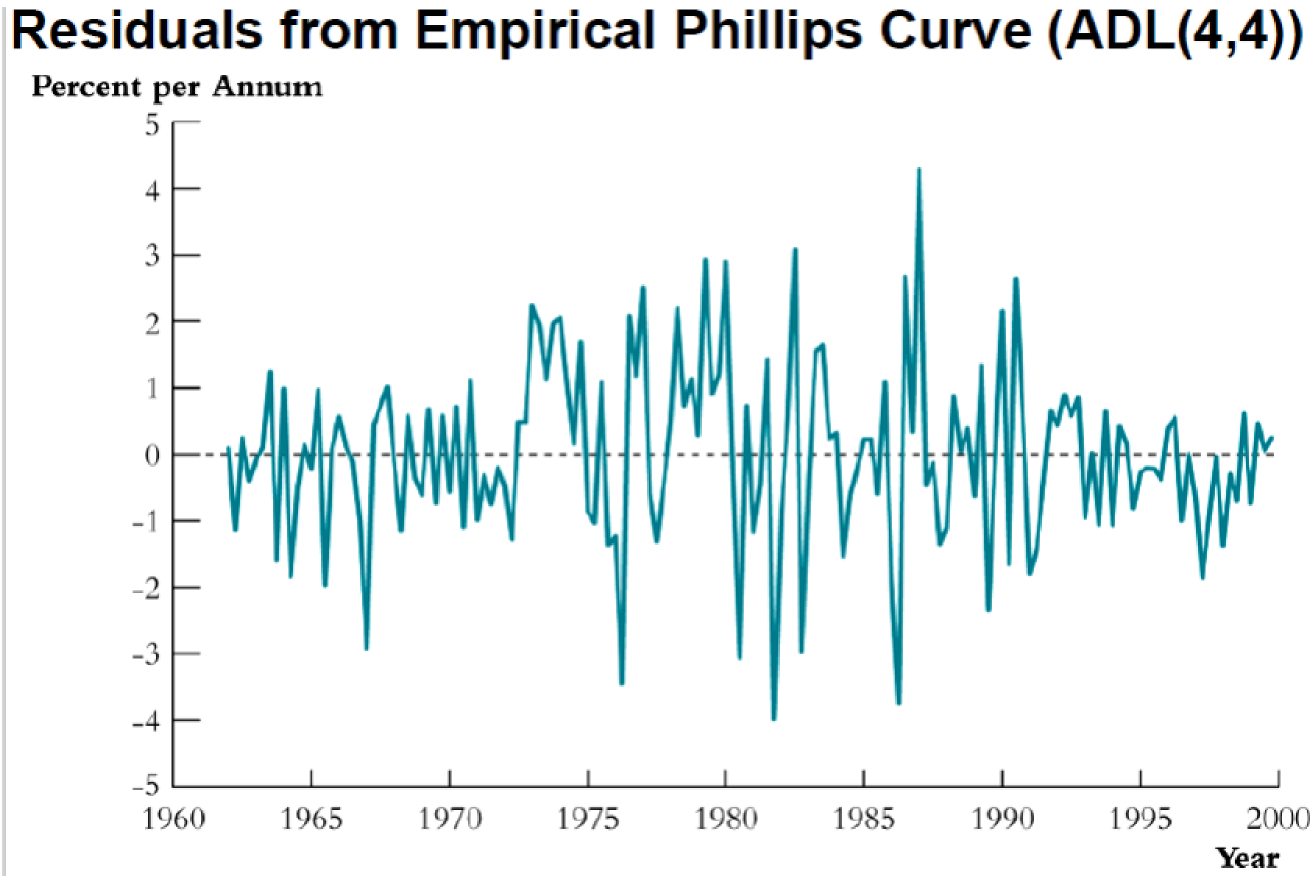
Ежедневные данные о темпах прироста индекса Нью-Йоркской фондовой биржи в 1990-2005 гг., %



Кластеризация волатильности часто наблюдается на финансовых рынках: периоды высокой и низкой дисперсии временного ряда последовательно сменяют друг друга.

Тут наблюдается так называемая **кластеризация волатильности**, которую как раз и позволяют измерить ARCH и GARCH модели.

Эффект кластеризации волатильности впервые был отмечен Манделбротом (1963 г.).



Иногда кластеризация волатильности встречается и в макроданных: последние десятилетия дисперсия остатков в модели инфляции гораздо меньше, чем в 70-х.

На слайде представлены остатки модели эмпирической кривой Филлипса.

Условная гетероскедастичность

Описанные примеры представляют собой специфический вид гетероскедастичности, характерный для временных рядов:

условную гетероскедастичность - conditional heteroskedasticity (HC).

Один из подходов к ее моделированию: ARCH и GARCH модели, а также их обобщения.

Условная гетероскедастичность

Зачем учитывать условную гетероскедастичность при моделировании?

1. **Эффективность.** Мы знаем, что корректный учет гетероскедастичности позволяет получить более точные оценки коэффициентов.
2. **Предсказание волатильности.** Иногда при анализе динамики финансовых переменных важно уметь предсказывать их волатильность (дисперсию). Например, для оценки степени рискованности актива.

Ключевой особенностью моделей *ARCH* (AutoRegressive Conditional Heteroscedasticity) и *GARCH* (General ARCH) является разграничение условных и безусловных дисперсий.

В то время как безусловная матрица ковариаций для представляющих интерес переменных может быть неизменной во времени, условные дисперсии и ковариации часто зависят нетривиальным образом от состояний изучаемых явлений или процессов в прошлом.

Понимание точного характера этой временной зависимости крайне важно для многих проблем в макроэкономике и финансах, таких как необратимые инвестиции, цены на опционы, структура процентных ставок по срокам и общие динамические соотношения для цен активов.

Кроме того, с точки зрения получения эконометрических выводов потеря в асимптотической эффективности из-за неучета гетероскедастичности может быть сколь угодно большой, и при составлении социально-экономических прогнозов, как известно, можно использовать намного более точную оценку неопределенности ошибки прогноза, если получать ее как условную по текущему информационному множеству.

ARCH(1)

Условной дисперсией называется дисперсия случайной переменной, обусловленная информацией о других случайных переменных, то есть дисперсия, найденная при условии наличия знаний о дисперсии в предыдущие моменты времени

$$\sigma_t^2 = D(u_t \,|\, u_{t-1}, u_{t-2 \ldots}).$$

ARCH(1) - модель первого порядка имеет вид:

$$\sigma_t^2 = \alpha_0 + \alpha_1 u_{t-1}^2$$

где $u_t = y_t - \hat{y}_t$

- остатки, полученные после предварительной оценки какой-либо модели временного ряда: AR, MA, ARMA, ARIMA, ADL, ...

Р.Энгл предложил модель в 1982 году.

ARCH(p)

Модель ARCH(p) (порядок авторегрессии дисперсии – p) имеет вид:

$$\sigma_t^2 = \alpha_0 + \alpha_1 u_{t-1}^2 + \alpha_2 u_{t-2}^2 + \ldots + \alpha_p u_{t-p}^2$$

Тут условная дисперсия представлена как линейная функция квадратов прошлых ошибок в моменты времени *t-1, t-2, ..., t-p* AR-процесс временных рядов – процесс авторегрессии, когда текущие значения ряда линейно зависят от предыдущих значений.

ARCH(p)

В модели *ARCH(p)* величину σ_t^2 называют волатильностью процесса.

Смысл модели в том, что если абсолютная величина u_t оказывается большой, то это приводит к повышению условной дисперсии в последующие периоды. При высокой условной дисперсии более вероятно появление больших по абсолютной величине значений u_t. Так, что модель характеризуется инерционностью условной дисперсии (кластеризацией волатильности).

ARCH(p)

Модель ARCH(p) - это модель AR(p) применительно к дисперсии остатков временного ряда, которая создает эту дисперсию в момент времени t в зависимости от сдвинутых наблюдений дисперсии.

Члены ошибки (u_t) являются остатками линейной модели, такой, например, как ARIMA, на исходном временном ряде.

Таким образом, строятся 2 модели:

- модель временного ряда;
- модель условной гетероскедастичности.

ARCH(p)

Процесс *ARCH* не автокоррелирован:

$$\mathbf{E}(u_t u_{t-j}) = \mathbf{E}(\mathbf{E}(u_t u_{t-j} | u_{t-1})) = \mathbf{E}(u_{t-j}\mathbf{E}(u_t | u_{t-1})) = 0$$

Поскольку процесс имеет постоянное (нулевое) математическое ожидание и не автокоррелирован, то он является слабо стационарным, если у него есть дисперсия. Еще одно из свойств *ARCH* процессов состоит в том, что безусловное распределение u_t имеет более толстые хвосты и остроконечную вершину, чем нормальное распределение.

ARCH(p)

Оценки модели *ARCH*(p) могут быть получены помощью обычного МНК. Эти оценки являются наилучшими линейными несмещенными оценками, но оказываются неэффективными.
Поэтому для нахождения оценок модели *ARCH(p)* используется метод максимального правдоподобия (квазиправдоподобия), считая, что u_t имеют нормальную асимптотику.

ARCH(p)

Пример:

$$y_t = \beta_0 + \delta_0 x_t + \delta_1 x_{t-1} + \delta_2 x_{t-2} + \ldots + u_t,$$

где случайные ошибки описываются моделью ARCH(p):

$$u_t = \varepsilon_t * \sqrt{\alpha_0 + \alpha_1 u_{t-1}^2 + \alpha_2 u_{t-2}^2 + \ldots + \alpha_p u_{t-p}^2},$$

ε_t — белый шум с единичной дисперсией

ARCH(p)

Условная дисперсия случайной ошибки:

$$\sigma_t^2 = \alpha_0 + \alpha_1 u_{t-1}^2 + \alpha_2 u_{t-2}^2 + \ldots + \alpha_p u_{t-p}^2$$

Безусловная дисперсия:

$$\mathrm{E}(\sigma_t^2) = \alpha_0 + \alpha_1 E(u_{t-1}^2) + \alpha_2 E(u_{t-2}^2) + \ldots + \alpha_p E(u_{t-p}^2)$$

$$\sigma^2 = \alpha_0 + \alpha_1 \sigma^2 + \cdots + \alpha_p \sigma^2$$

$$\sigma^2 = \frac{\alpha_0}{1 - \alpha_1 - \cdots - \alpha_p}$$

$$(\alpha_1 + \cdots + \alpha_p) < 1$$

ARCH(p)

Условная гетероскедастичность **не противоречит** условию стационарности ряда, так как безусловная дисперсия постоянна.

$$\sigma^2 = \frac{\alpha_0}{1 - \alpha_1 - \cdots - \alpha_p}$$

Модель GARCH(p,q)

Базовая модель ARCH(p) может быть обобщена в различных направлениях.

Наиболее важным является обобщенная ARCH, называемая GARCH, в которую включены слагаемые, соответствующие скользящим средним. Автором модели считается Тим Питер Боллерслев.

Модель GARCH(p,q) – модель с обобщённой авторегрессионной условной гетероскедастичностью -

Generalized autoregressive conditional heteroskedasticity (GARCH).

GARCH (*p,q*) имеет следующий вид:

$$\sigma_t^2 = \alpha_0 + \alpha_1 u_{t-1}^2 + \alpha_2 u_{t-2}^2 + \ldots + \alpha_p u_{t-p}^2 + \\ + \varphi_1 \sigma_{t-1}^2 + \varphi_2 \sigma_{t-2}^2 + \ldots + \varphi_p \sigma_{t-q}^2$$

Модель GARCH(p,q)

Модель GARCH(p,q) может быть записана в общем виде:

$$\sigma_t^2 = \alpha_0 + \sum_{i=1}^{p} \alpha_i u_{t-i}^2 + \sum_{j=1}^{q} \varphi_j \sigma_{t-j}^2$$

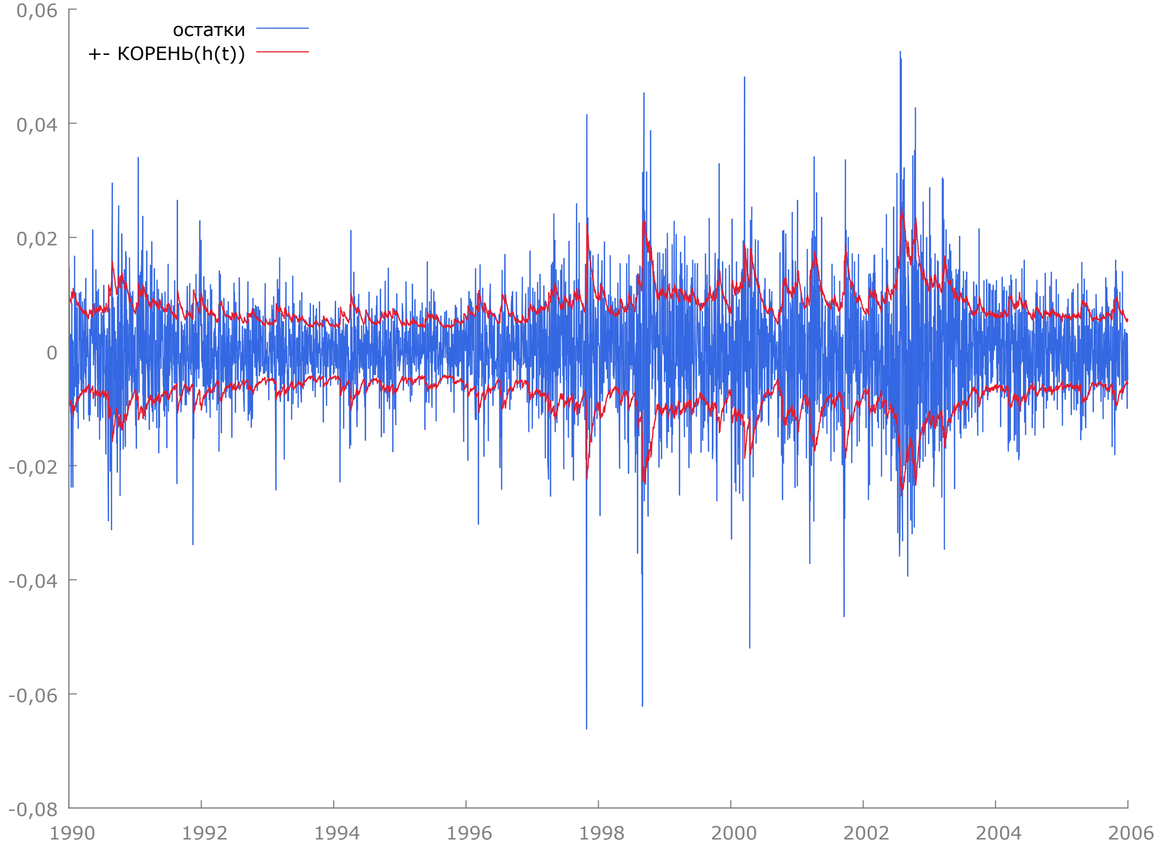
где α_i и $\varphi_j > 0$ ($i=1, 2,\ldots, p$; $j=1, 2,\ldots, q$) иначе дисперсия была бы меньше нуля.

GARCH- модель показывает, что текущее значение условной дисперсии является функцией от константы, *p* квадратов остатков из уравнения условной средней (или любого другого уравнения, описывающего временной ряд) и *q* значений предыдущей условной дисперсии (то есть AR-процесс *q*-го порядка от условной дисперсии).

Наиболее популярной для предсказания изменчивости доходности финансовых активов является модель GARCH(1, 1):

$$\sigma_t^2 = \alpha_0 + \alpha_1 u_{t-1}^2 + \varphi_1 \sigma_{t-1}^2$$

Оценивание GARCH(1,1) для доходности индекса Нью-Йоркской Фондовой биржи (1990–2005)



$$\sigma_t^2 = 0{,}0000008 + 0{,}0728u_{t-1}^2 + 0{,}9185\sigma_{t-1}^2$$

Оценивание GARCH(1,1) для доходности индекса Нью-Йоркской Фондовой биржи (1990–2005) в Gretl

gretl: модель 4

Файл Правка Тесты Сохранить Графики Анализ LaTeX

```
Оценок функции: 87
Оценок градиента: 18

Модель 4: GARCH, использованы наблюдения 1990-01-02:2005-12-30 (T = 4036)
Зависимая переменная: Persent
Стандартные ошибки рассчитаны на основе Гессиана

             коэффициент   ст. ошибка      z      p-значение
  -----------------------------------------------------------
  const      0,000523299   0,000110285    4,745   2,09e-06    ***

  alpha(0)   7,87499e-07   2,01951e-07    3,899   9,64e-05    ***
  alpha(1)   0,0728166     0,00889480     8,186   2,69e-016   ***
  beta(1)    0,918493      0,00992245    92,57    0,0000      ***

Среднее завис. перемен  0,000368   Ст. откл. завис. перем  0,008945
Лог. правдоподобие      13799,90   Крит. Акаике           -27589,81
Крит. Шварца           -27558,29   Крит. Хеннана-Куинна   -27578,64
обратите внимание на сокращенные обозначения статистики

Безусловная дисперсия ошибок = 9,06214e-005
Отношение правдоподобия для (G)ARCH компонент:
  Хи-квадрат(2) = 979,886 [1,66135e-213]
```

$$\sigma_t^2 = \underset{(0{,}0000002)}{0{,}0000008} + \underset{(0{,}0089)}{0{,}0728}u_{t-1}^2 + \underset{(0{,}0099)}{0{,}9185}\sigma_{t-1}^2$$

Оценивание GARCH(1,1) для доходности индекса Нью-Йоркской Фондовой биржи (1990–2005) в EViews

Equation: UNTITLED   Workfile: КОПИЯ NYSE::Untitled\

View | Proc | Object | Print | Name | Freeze | Estimate | Forecast | Stats | Resids

Dependent Variable: PERSENT
Method: ML ARCH - Normal distribution (BFGS / Marquardt steps)
Date: 05/22/22   Time: 22:03
Sample: 1/02/1990 12/30/2005
Included observations: 4036
Convergence achieved after 30 iterations
Coefficient covariance computed using outer product of gradients
Presample variance: backcast (parameter = 0.7)
GARCH = C(2) + C(3)*RESID(-1)^2 + C(4)*GARCH(-1)

Variable	Coefficient	Std. Error	z-Statistic	Prob.
C	0.000524	0.000115	4.543618	0.0000
Variance Equation				
C	7.93E-07	1.36E-07	5.814830	0.0000
RESID(-1)^2	0.072990	0.005191	14.06173	0.0000
GARCH(-1)	0.918212	0.006030	152.2711	0.0000

R-squared	-0.000306	Mean dependent var	0.000368
Adjusted R-squared	-0.000306	S.D. dependent var	0.008945
S.E. of regression	0.008947	Akaike info criterion	-6.836554
Sum squared resid	0.322974	Schwarz criterion	-6.830307
Log likelihood	13800.17	Hannan-Quinn criter.	-6.834341
Durbin-Watson stat	1.918555		

$$\underset{(0{,}0000013)}{\sigma_t^2 = 0{,}0000008} + \underset{(0{,}0052)}{0{,}0730u_{t-1}^2} + \underset{(0{,}0060)}{0{,}9182\sigma_{t-1}^2}$$

Волатильность GARCH

Волатильность (изменчивость) не является постоянным процессом, и может изменяться во времени.

Для предсказания волатильности при помощи GARCH-модели, можно использовать следующую рекурсивную модель:

$$\sigma_{t+1}^2 = \alpha_1 + \beta_1 u_t^2 + \gamma_1 \sigma_t^2$$

$$\sigma_{t+j}^2 = \alpha_1 + (\beta_1 + \gamma_1)\sigma_{t+j-1}^2$$

Здесь u_t^2 - величина, неизвестная в будущем, которая при выполнении прогноза заменяется условной оценкой дисперсии σ_t. Таким образом, рекурсивная модель позволяет предсказывать σ_t^2 в момент времени $(t+1)$, затем σ_t^2 в момент времени $(t+2)$ и т.д. При этом, например, σ_{t+2} рассчитывается как условная дисперсия при условии, что известны значения $y_1, y_2, \ldots, y_t$ и прогноза y_{t+1}. Результат каждого расчёта является предсказанием условной дисперсии на j периодов вперёд.

Достоинства и недостатки ARCH(p) и GARCH(p,q)

Одной из **сильных сторон** моделей ARCH(p) и GARCH(p,q) является то, что они оценивают положительный избыточный эксцесс, т.е. «толстые» хвосты относительно нормального распределения, что согласуется с эмпирическими данными финансовых временных рядов.

Слабые стороны данных моделей заключаются в том, что они исходят из одинакового эффекта для положительных и отрицательных шоков волатильности, поскольку она зависит от квадрата предидущих шоков, в то время как цены активов реагируют на положительные и отрицательные шоки по-разному.

Модели ARCH(p) и GARCH(p,q) также не предлагают нового понимания источника вариаций финансового временного ряда, поскольку они просто механически описывает условную дисперсию.

Модели ARCH(p) и GARCH(p,q), скорее всего, предсказывают волатильность излишне, поскольку они медленно откликаются на крупные изолированные шоки.

Алгоритм идентификации ARCH и GARCH-моделей

1. Построить модель стационарного временного ряда y_t с ошибкой u_t;
2. Определить остатки u_t;
3. Построить график остатков u_t;
4. Протестировать остатки на ARCH-эффекты;
5. Построить графики АКФ и ЧАКФ;
6. Определить порядок ARCH или GARCH модели
7. Оценить несколько ARCH и/или GARCH моделей
8. Проверить значимость моделей и их оценок.
9. Сравнить модели, используя информационные критерии Акаике, Шварца и Хеннана-Куина (рекомендуется использовать критерий Шварца).
10. Выбрать наилучшую модель.
11. Тестировать остатки модели на автокорреляцию.

Общая схема тестирования модели на ARCH-эффекты

1. Проводится оценка модели стационарного временного ряда (AR, MA, ARMA, ARIMA, тренда, …);
2. Определяются остатки $u_t = y_t - \hat{y}_t$
3. Оценивается модель ARCH(p):

$$\hat{\sigma}_t^2 = \alpha_0 + \alpha_1 \hat{u}_{t-1}^2 + \alpha_2 \hat{u}_{t-2}^2 + \ldots + \alpha_q \hat{u}_{t-p}^2$$

Модель тестируется на ARCH-эффекты p-го порядка.

4. для оценённой модели рассчитывается коэффициент детерминации R^2
5. формируются гипотезы (нулевая и альтернативная):

$$H_0 : \alpha_1 = 0, \alpha_2 = 0, \ldots, \alpha_p = 0 \quad ,$$

$$H_1 : \alpha_1 \neq 0, \alpha_2 \equiv 0, \ldots, \alpha_p \neq 0 \quad ;$$

6. определяется значение статистики $\chi^2_{расч} = (T-p)R^2$, где T – число наблюдений, R^2 – коэффициент детерминации;
7. сравнивается $\chi^2_{расч}$ с $\chi^2_{табл}$, определённым для степеней свободы p (p – число временных лагов в модели ARCH(p))
8. если $\chi^2_{расч} > \chi^2_{табл}$, то H_0 отклоняется, и считается, что ARCH-модель существенна на заданном уровне значимости и её порядок равен p.

Проверка адекватности GARCH/ARCH моделей.

Качество подгонки GARCH/ARCH модели под исходные данные можно контролировать на основе близости к единице коэффициента детерминации (R^2) или скорректированного коэффициета детерминации ($R^2_{Adjusted}$).

$$R^2 = 1 - \frac{D^2_{ост}}{D^2_{общ}} \qquad \text{или} \qquad R^2_{\text{Adjusted}} = R^2 \cdot \frac{(n-k)}{(n-1)} \quad ,$$

где n – общее число наблюдений,

k – число степеней свободы модели (для GARCH $k=p+q$, для ARCH $k=p$),

$D^2_{ост}$ – дисперсия остатков,

$D^2_{общ}$ – общая дисперсия.

Проверка адекватности GARCH/ARCH моделей.

Необходимо также проанализировать стандартизованные остатки u/σ,

где σ – условное среднеквадратическое отклонение, рассчитанное по модели GARCH/ARCH,

а u - остатки в уравнении условного математического ожидания (первоначального уравнения).

Если модель GARCH/ARCH достаточно хорошо описана, то стандартизованные остатки являются независимыми одинаково распределенными случайными величинами с нулевым математическим ожиданием и единичным средним квадратичным отклонением.

Идентификация модели GARCH

I этап: **Расчет критерия Люнга-Бокса (*LB*).** Рассчитывается *LB* статистика на основе предварительного расчета k коэффициентов автокорреляции по T наблюдениям (ρ_k), с последующим возведением их в квадрат ρ_k^2 :

$$LB_{расчетное} = T(T+2)\sum_{k=1}^{m}\frac{\rho_k^2}{(T-k)}$$

где m – максимальный лаг коэффициентов автокорреляции, T – длина временного ряда.

Выдвигается гипотеза о незначимости m лагов в исходной модели авторегрессии. Сравнивают расчетное значение *LB* с критическим значением χ^2, определенным для степени свободы $v=m$. Если $LB_{расчетное} > \chi^2$, то гипотезу о незначимости m лагов в исходной модели авторегрессии отклоняют на заданном уровне значимости α.

Идентификация модели GARCH

II этап: **Расчет критерия Люнга-Бокса по стандартизованным остаткам.**

Рассчитывают коэффициенты автокорреляции по стандартизованным остаткам γ'_k с последующим возведением их в квадрат γ'^2_k :

$$LB_{расчетное} = T(T+2)\sum_{k=1}^{m}\frac{\gamma'^2_k}{(T-k)}$$

где m – максимальный лаг автокорреляции стандартизованных остатков.

Выдвигается гипотеза о несущественности порядков p и q GARCH модели. $LB_{расчетное}$ сравнивают с $\chi^2_{табл}$, определенным для степени свободы $v_1 = m-p-q$, где m – общее число наблюдений, p и q – порядки GARCH модели.

Если $LB_{расчетное} < \chi\ (m-p-q)$, то нулевую гипотезу отклоняют и считают, что GARCH модель хорошо определена, и ее порядки соответственно равны p и q.

На практике модель *GARCH* дополняют какой-либо моделью, описывающей поведение условного или безусловного математического ожидания наблюдаемого ряда. С точки зрения прогнозирования перспективной является модель, сочетающая *ARIMA* с *GARCH*, первая используется для моделирования условного математического ожидания ряда, а вторая − условной дисперсии.

Модели авторегрессионно условно гетероскедастичных процессов могут различаться тем, какой именно функцией задается зависимость условной дисперсии от своих лагов и лагов u_t. Например, в **логарифмической *GARCH* модели** условная дисперсия задается уравнением

$$\ln \sigma_t^2 = \mu_0 + \sum_{j=1}^{p} \theta_j \ln \sigma_{t-j}^2 + \sum_{j=1}^{q} \mu_j \ln u_{t-j}^2$$

Следующая **нелинейная *GARCH* модель** включает в себя как частный случай обычную *GARCH* модель

$$\sigma_t^\lambda = \mu_0 + \sum_{j=1}^{p} \theta_j \sigma_{t-j}^\lambda + \sum_{j=1}^{q} \mu_j \left| u_{t-j} \right|^\lambda$$

Логарифмическая *GARCH* модель является предельным частным случаем этой модели при $\lambda \to 0$.

В приведенных моделях условная дисперсия зависит от абсолютной величины лагов u_t.

В финансовых данных часто наблюдается *эффект левериджа*: снижение рыночной стоимости акционерного капитала увеличивает соотношение заемных средств к собственным и, следовательно, повышает рискованность вложений в фирму.

Последнее проявляется в увеличении волатильности и будущие значения волатильности отрицательно коррелируют с текущей доходностью, т.е. модель асимметрична относительно u_t. Например, Д.Нельсон предложил **экспоненциальную *GARCH* модель (*EGARCH*)**:

$$u_t = \varepsilon_t \exp\left[\mu_0 + \theta\sigma_{t-1}^2 + \mu_1\left|\frac{u_{t-1}}{\sigma_{t-1}}\right| + \mu_2\frac{u_{t-1}}{\sigma_{t-1}}\right]$$

Модель *EGARCH* несимметрично реагирует на всплески. Асимметрия информации позволяет дисперсии быстрее реагировать на понижение активности рынка, чем на взлеты.

Некоторые другие виды моделей *ARCH*:

Абсолютные остатки $\sigma_t^2 = \mu_0 + \mu_1 |u_{t-1}|$.

Пороговые (*TGARCH*) модели вида

$$\sigma_t^2 = \mu_0 + \mu_1 u_{t-1}^2 + \mu_2 u_{t-1}^2 d_{\mathrm{t}-1} + \theta \sigma_{t-1}^2$$

где d_t=1, если u_t<0 и 0 в противном случае.

В отличие от рассмотренных выше моделей *ARCH*, которые являются моделями, управляемыми наблюдениями (или данными), в другом классе моделей волатильности – моделях управляемых параметрами – предполагается, что волатильность зависит не от прошлых наблюдений, а от некоторых ненаблюдаемых компонент. Наиболее широко известна модель **стохастической волатильности (*SV* – stochastic volatility) С.Тэйлора:**

$$u_t=\varepsilon_t \exp[h_t/2],\ h_{t+1}=\mu_0+\mu_1 h_t+\eta_t,$$

где h_t – латентная компонента, которую можно интерпретировать как случайный и неустойчивый поток новой информации, который очень трудно моделировать непосредственно, но можно оценить по данным наблюдений; η_t – гауссовский белый шум.

Модель Тэйлора называется логнормальной *SV* моделью. Оценки модели получают обобщенным методом наименьших квадратов.

Более подробно о моделях волатильности можно прочитать в: Модели теории временных рядов в финансах и эконометрике//Обозрение прикладной и промышленной математики. Т. 3. Вып. 6. – М.: ТВП, 1996.