

# ARСН-модель

Определение 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  $\sigma_t^2 = D(\varepsilon_t | \varepsilon_{t-1}, \varepsilon_{t-2}, \dots)$ .

Определение 2: **ARСН-модель** первого порядка имеет вид:

$$\sigma_t^2 = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2, \quad (1)$$

где  $\varepsilon_t = y_t - \hat{y}_t$  - остатки, полученные после предварительной оценки какой-либо модели.

Здесь дисперсия в момент времени  $t$  зависит от квадрата ошибок в момент  $(t-1)$ , то есть условная дисперсия  $\sigma_t^2$  является AR-процессом квадрата ошибок модели.

**ARСН-модель ( $q$ )** (порядок авторегрессии дисперсии –  $q$ ) имеет вид:

$$\sigma_t^2 = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \alpha_2 \varepsilon_{t-2}^2 + \dots + \alpha_q \varepsilon_{t-q}^2 \quad (2)$$

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

Моделировать дисперсию можно лишь после того, как из данных убрали среднюю величину, поэтому полная модель ARCH( $q$ ) имеет вид:

$$y_t = \alpha_0 + \sum_{i=1}^p \alpha_i y_{t-i}; \quad (3)$$

$$\sigma_t^2 = \beta_0 + \sum_{j=1}^q \beta_j \varepsilon_{t-j}^2.$$

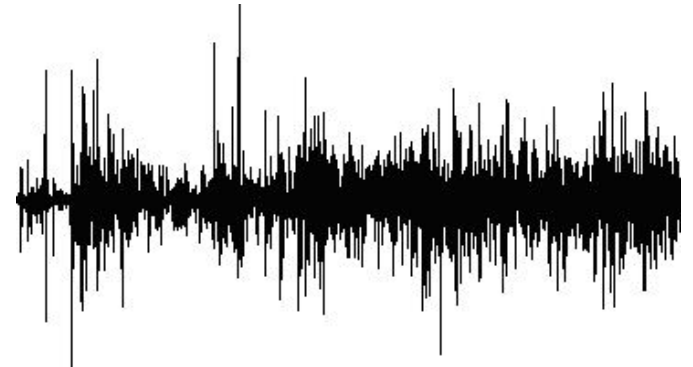
Здесь первое уравнение является AR-процессом  $p$ -ого порядка, а во втором уравнении моделируется условная дисперсия как линейная функция от  $q$  квадратов прошлых значений ошибок ( $\varepsilon_{t-q}$ ), полученных после оценки первого уравнения.

Условие «положительности» дисперсии:  $\beta_0 > 0, \beta_1 \geq 0, \beta_2 \geq 0, \dots, \beta_q \geq 0$ .

**Смысл ARCH-модели:** если абсолютная величина остатков модели  $\varepsilon_t$  первого уравнения (3) является большой, то это приведет к повышению условной дисперсии в последующие периоды (второе уравнение (3)), наоборот если остатки близки к нулю, то это приведет к снижению условной дисперсии.

Это проявляется свойство, называемое **кластеризацией волатильности**, которую как раз и позволяет измерить ARCH-модель.

Эффект кластеризации волатильности впервые было отмечено Манделбротом (1963 г.)



# Алгоритм определения наличия ARCH-эффектов.

1. необходимо построить AR-модель ряда  $x_t$  с ошибкой  $\varepsilon_t$  согласно первому уравнению из (3);
2. определить остатки  $\hat{\varepsilon}_t$  как оценки  $\varepsilon_t$ ;
3. построить линейную регрессию квадрата ошибок в момент времени  $t$  на квадрат остатков модели после моделирования AR:  $\hat{\varepsilon}_t^2 = \lambda_1 \hat{\varepsilon}_{t-1}^2$  ;
4. протестировать коэффициент  $\lambda$  на отсутствие значимости с помощью теста Стьюдента, теста Фишера,  $\chi^2$  теста, принимая за нулевую гипотезу:  
 $H_0: \lambda_1 = 0$ .  
Соответственно за альтернативную гипотезу  $H_1: \lambda_1 \neq 0$ .
5. Если  $\lambda_1$  значимо отличается от 0, то модель можно специфицировать как ARCH-модель первого порядка (ARCH (1)).

# Общая схема тестирования модели на ARCH-эффекты:

1. Проводится оценка модели (например, AR-модели, СС-модели, ARСС-модели или простой регрессии по времени);

2. на основе знания ошибок модели  $\varepsilon_t = y_t - \hat{y}_t$  ( $\hat{y}_t$  – расчётное значение модели, построенной в пункте 1)), оценивается модель:

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

Здесь модель тестируется на ARCH-эффекты  $p$ -го порядка.

3. для оценённой модели рассчитывается коэффициент детерминации  $R^2$ , отвечающий за качество подгонки модели;

4. формируются гипотезы (нулевая и альтернативная):

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

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

5. определяется значение статистики  $\chi^2_{расч} = TR^2$ , где  $T$  – объём выборки ряда,  $R^2$  – коэффициент детерминации;

6. сравнивается  $\chi^2_{расч}$  с  $\chi^2_{табл}$ , определённым для степеней свободы  $p$  ( $p$  – число временных лагов в модели ARCH( $p$ ))

7. 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  $p$ .

# GARCH-модель

Определение 3: **GARCH-модель** – модель с обобщённой авторегрессией условной гетероскедастичности. GARCH ( $p, q$ ) имеет в отличие от ARCH-модели два порядка и в общем виде

записывается:

$$\sigma_t^2 = \alpha_0 + \sum_{i=1}^p \alpha_i \varepsilon_{t-i}^2 + \sum_{j=1}^q \beta_j \sigma_{t-j}^2 \quad (4)$$

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

GARCH- модель показывает, что текущее значение условной дисперсии является функцией от константы –  $p$ -го значения квадратов остатков из уравнения условной средней (или любого другого уравнения) и  $q$ -го значения 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):

$$\sigma_t^2 = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \beta_1 \sigma_{t-1}^2 \quad (5)$$

# Волатильность GARCH

**Волатильность** (изменчивость) не является постоянным процессом, и может изменяться во времени. Если известна точная модель описания процесса, изменяющегося во времени, то для нахождения годовой волатильности этого процесса нужно определить квадратный корень из условной дисперсии и умножить модель на  $\sqrt{N}$ , где  $N$  – количество наблюдений в год. Полученная мера волатильности будет изменяться во времени, т.е. текущая волатильность будет определяться как функция от прошлой волатильности. Для предсказания волатильности при помощи GARCH-модели, можно использовать следующую рекурсивную модель:

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

Здесь  $\varepsilon_t^2$  - величина, неизвестная в будущем, которая при выполнении прогноза заменяется условной оценкой дисперсии  $\sigma_t$ . Таким образом, формула (7) позволяет предсказывать  $\sigma_{t+1}^2$  в момент времени  $(t+1)$ , затем  $\sigma_{t+2}^2$  в момент времени  $(t+2)$  и т.д. При этом, например,  $\sigma_{t+2}$  рассчитывается как условная дисперсия при условии известности значений  $y_1, y_2, \dots, y_t$  и прогноза  $y_{t+1}$ . Результат каждого расчёта является предсказанием условной дисперсии на  $j$  периодов вперёд.

# Оценка моделей ARCH и GARCH

ARCH и GARCH процессы имеют, как правило, островершинное безусловное распределение. Так эксцессы (момент четвертого порядка) для модели ARCH (1), представленной уравнением (1), и GARCH (1;1), представленной уравнением (5), соответственно равны

$$\frac{6\alpha_1^2}{1 - 3\alpha_1^2} \quad \text{и} \quad \frac{6\beta_1^2}{1 - \alpha_1^2 - 2\alpha_1\beta_1 - 3\beta_1^2} .$$

Коэффициенты асимметрии (моменты третьего порядка) для моделей волатильности равны нулю. Несмотря на это стандартным методом оценки моделей является **метод максимального правдоподобия**, основывающийся на нормальном распределении. При этом оценки моделей будут состоятельными, но асимптотически неэффективными (неэффективные в пределе при увеличении числа степеней свободы). Отметим, что наличие высоких куртозисов ARCH-процессов хорошо согласуется с поведением многих финансовых показателей, имеющих толстые хвосты в распределении.

# Оценка моделей ARCH и GARCH

Рассмотрим условное математическое ожидание в момент времени  $t$ :

$$m_t = \alpha_0 + \sum_{i=1}^m \alpha_i y_{t-i} + \varepsilon_t, \quad \varepsilon_t = y_t - \left( \alpha_0 + \sum_{i=1}^m \alpha_i y_{t-i} \right).$$

Отсюда ошибка определяется как:

Предположим, что  $\varepsilon_t = \sigma_t^2 z$ , где  $\sigma_t^2$  - условная дисперсия, а  $z$  - распределена по стандартизированному нормальному закону,

то есть  $z \sim N(0, 1)$ . Тогда  $\varepsilon_t \sim N(0, \sigma_t^2)$ , где для ARCH-модели:

$$\sigma_t^2 = \beta_0 + \sum_{i=1}^p \beta_i \varepsilon_{t-i}^2 ;$$

а для GARCH-модели:

$$\sigma_t^2 = \beta_0 + \sum_{i=1}^p \beta_i \varepsilon_{t-i}^2 + \sum_{j=1}^q r_j \sigma_{t-j}^2$$

В итоге для ARCH-модели мы имеем  $m+1+p+1$  неизвестных параметров, а для GARCH-модели  $m+1+p+1+q$  параметров, подлежащих оцениванию.

Оценка параметров производится методом максимального правдоподобия.



# Проверка адекватности GARCH/ARCH

## моделей.

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

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

здесь  $n$  – общее число наблюдений временного ряда,  $k$  – число степеней свободы модели (для GARCH  $k=p+q$ , для ARCH  $k=p$ ),  $D_{ост}^2$  – остаточная или объясненная моделью дисперсия,  $D_{общ}^2$  – общая дисперсия.

**Для проверки достоверности** оценок модели необходимо проанализировать стандартизованные остатки  $\hat{\epsilon}/\sigma$ , где  $\sigma$  – условное среднеквадратическое отклонение, рассчитанное по модели GARCH/ARCH, а  $\hat{\epsilon}$  – 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 Лjungа-Бокса (LB).** Рассчитывается  $LB$  статистика на основе предварительного расчета  $k$  коэффициентов автокорреляции по  $T$  наблюдениям ( $\rho_k$ ), с последующим возведением их в квадрат  $\rho_k^2$ :

$$LB_{\text{расчетное}} = \frac{T(T+2)}{8} \sum_{k=1}^m \frac{\rho_k^2}{(T-k)} \quad (8)$$

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  $\chi^2$ , определенным для степени свободы  $\nu=m$ . Если  $LB_{\text{расчетное}} > \chi^2$ , то гипотезу о незначимости  $m$  лагов в исходной модели авторегрессии отклоняют на заданном уровне значимости  $\alpha$ .

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

Рассчитывают коэффициенты автокорреляции по стандартизованным остаткам  $\gamma'_k$  с последующим возведением их в квадрат  $\gamma_k'^2$ :

$$LB_{\text{расчетное}} = \frac{T(T+2)}{9} \sum_{k=1}^m \frac{\gamma_k'^2}{(T-k)} \quad (9)$$

где  $m$  – максимальный лаг автокорреляции стандартизованных остатков.

Выдвигается гипотеза о несущественности порядков  $p$  и  $q$  GARCH модели.  $LB_{\text{расчетное}}$  сравнивают с  $\chi^2_{\text{табл}}$ , определенным для степени свободы  $\nu_1=m-p-q$ , где  $m$  – общее число наблюдений,  $p$  и  $q$  – порядки GARCH модели.

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

# Идентификация модели GARCH на основе анализа коррелограмм

1. После оценивания математического ожидания ряда данных (на основе АРСС моделей, выделения компонент временного ряда или обычной регрессии) получают остаточную компоненту.
2. Стандартизируют полученные остатки.
3. Строят коррелограммы АКФ и ЧАКФ на стандартизированных остатках.
4. Определяют количество лагов для коэффициентов АКФ и ЧАКФ, выходящих за границы белого шума. Полученное число является порядком ARCH-модели.

**Селекцию моделей ARCH и GARCH проводят исходя из минимума информационных критериев Акайке, Шварца и Ханена-Квина.**