11 Chapter 11: Estimation and testing of ARCH and GARCH models

We have observations $X_1, \ldots, X_T$ and would like to estimate the parameters of ARCH/GARCH model.

11.1 Conditional MLE

For the ARCH(p) model

$$\begin{aligned} X_t &= \sigma_t \varepsilon_t, \\ \sigma_t^2 &= c_0 + b_1 X_{t-1}^2 + \ldots + b_p X_{t-p}^2, \end{aligned} \tag{11.1}$$

if we know the probability density function (pdf) $p(\cdot)$ of ε_t, the conditional log-likelihood function is

$$l(\{b_i\}, c_0, \sigma_t^2) = \sum_{t=p+1}^{T} (\ln(p(X_t/\sigma_t) - \ln(\sigma_t))). \tag{11.2}$$

The MLE for c_0, $\{b_i\}$ and σ_t^2 could be obtained by maximizing the above function.
If $\varepsilon_t \sim N(0, 1)$, then

$$l(\{b_i\}, c_0, \sigma_t^2) = -\frac{1}{2} \sum_{t=p+1}^{T} (\frac{X_t}{\sigma_t} + \ln(\sigma_t))). \tag{11.3}$$

Remarks.

- Under some mild regularity conditions, the MLE is consistent. The asymptotic normality requires more conditions such as $E\varepsilon_t^4 < \infty$, which is often questionable in financial applications.

- For a GARCH model

$$\sigma_t^2 = c_0 + \sum_{i=1}^{p} b_i X_{t-i}^2 + \sum_{j=1}^{q} a_j \sigma_{t-j}^2, \tag{11.4}$$

  it can be shown that

$$\begin{aligned} \sigma_t^2 = & \frac{c_0}{1 - \sum_{j=1}^{q} a_j} + \sum_{i=1}^{p} b_i X_{t-i}^2 + \\ & \sum_{i=1}^{p} b_i \sum_{k=1}^{\infty} \sum_{j_1=1}^{q} \ldots \sum_{j_k=1}^{q} a_{j_1} \ldots a_{j_k} X_{t-i-j_1-\ldots-j_k}^2. \end{aligned} \tag{11.5}$$

  If we apply truncation for $t > p$

$$\begin{aligned} \sigma_t^2 = & \frac{c_0}{1 - \sum_{j=1}^{q} a_j} + \sum_{i=1}^{p} b_i X_{t-i}^2 + \\ & \sum_{i=1}^{p} b_i \sum_{k=1}^{\infty} \sum_{j_1=1}^{q} \ldots \sum_{j_k=1}^{q} a_{j_1} \ldots a_{j_k} X_{t-i-j_1-\ldots-j_k}^2 \times \\ & I(t - i - j_1 - \ldots - j_k > 0), \end{aligned} \tag{11.6}$$

  the conditional MLE minimizes

$$\sum_{t=s}^{T} (\ln(\tilde{\sigma}_t) - \ln(f(X_t/\tilde{\sigma}_t))),$$

  where $s > p + 1$.

- The estimation of the conditional second moments is more difficult than the estimation of conditional mean; the likelihood function tend to be rather flat unless the sample size T is very large.
- To increase the fatness of tails, we may let ε_t to follow a distributions that has heavier tails than normal (all the distributions have been normalized to have a unit variance):

1. $t(k)$ with the pdf
$$f_k(x) = \frac{\Gamma((k+1)/2)}{(\pi k)^{0.5}\Gamma(k/2)(1+x^2/k)^{(k+1)/2}}.$$
2. Double exponential distribution
$$f(x) = 2^{-1/2}\exp -\sqrt{2}|x|.$$
3. Generalized Gaussian distribution
$$f(x) = \frac{k}{\lambda 2^{1+1/k}\Gamma(1/k)} exp - \frac{1}{2}|x/\lambda|^k,$$
where $\lambda = (2^{-2/k}\Gamma(1/k)/\Gamma(3/k))^{1/2}$.
4. We may use AIC and BIC to determine the order of the model.
5. In the case of unknown pdf $p(\cdot)$, we use the Gaussian likelihood to derive the quasi-MLE.

Theorem(Hall and Yao (2003)).

Let $\theta = (c_0, b_1, \ldots, b_p, a_1, \ldots, a_q)', U_t = \frac{d\sigma_t^2}{d\theta}$. (Then U_t/σ_t^2 has all its moments finite.) Suppose $M \equiv E(U_t U_t'/\sigma_t^4) > 0$.

Now if $E\varepsilon_t^4 < \infty$, then
$$\frac{T^{0.5}}{(E\varepsilon_t^4 - 1)^{0.5}}(\hat{\theta} - \theta) \to N(0, M^{-1})$$
in distribution.

11.2 Tests for ARCH effect

Test on GARCH models can be divided into two categories:

1. The likelihood ratio (LR) test can be constructed in the standard manner.
2. The standard test for ARMA models applied to the process X_t^2.

Suppose that $\{X_t\}$ is a strictly stationary process defined by
$$X_t = \sigma_t \varepsilon_t, \qquad \sigma_t^2 = c_0 + \sum_{i=1}^{p} b_i X_{t-i}^2, \tag{11.7}$$
where $c_0 > 0$ and $b_i > 0, \forall i$.

We test the null hypothesis
$$H_0 : b_1 = \ldots = b_p = 0$$

against the alternative hypothesis

$$H_1 : \exists j \in N, b_j \neq 0.$$

Suppose the density function $f(\cdot)$ of ε_t is known, then the conditional likelihood ratio test based on the statistic

$$S_{t,1} = \prod_{t=p+1}^{T} \frac{\sigma_t(\hat{c_0}, \hat{b})^{-1} f(X_t/\sigma_t(\hat{c_0}, \hat{b})}{\sigma_t(\hat{c_0}, 0)^{-1} f(X_t/\sigma_t(\hat{c_0}, 0)},$$

where $f(\hat{c_0}, \hat{b})$ is the conditional MLE for an ARCH(p) model, and $f(\hat{c_0}, 0)$ is the MLE under H_0.

Under H_0 in distribution

$$2lnS_{T,1} \to \chi_p^2$$

provided that the density function $f(\cdot)$ has sufficient degree of smoothness and the Fisher information matrix

$$I(c_0, b) \equiv \begin{pmatrix} I_{11}(c_0, b) & I_{12}(c_0, b) \\ I_{21}(c_0, b) & I_{22}(c_0, b) \end{pmatrix} = E(\dot{l}(X_t, c_0, b)\dot{l}(X_t, c_0, b)')$$

exists and positive-definite (see Serfling (1980)). Those conditions are fulfilled for GARCH models with $\varepsilon_t \sim N(0, 1)$. Here $\dot{l}(X_t, c_0, b)$ denotes the derivative of $\ln \sigma_t^{-1} f(X_t/\sigma_t)$ with respect to (c_0, b).

Engle (1982) proposed a test in terms of the statistic TR^2 where R^2 is the multiple correlation coefficient of e_t^2 and $(1, e_{t-1}^2, \ldots, e_{t-p}^2)$ where e_t are the sample residuals obtained from a curve fit (e.g., a regression model).

11.3 Tests for residuals from ARCH models

Our assumptions for residuals ε_t from ARCH/GARCH models are:

- uncorrelated (this assumption is assessed using the Ljung-Box test and acf plots); typically is satisfied;
- homoscedastic (this assumption is assessed using the scatter plot of the residuals and tests for variance homogeneity);
- normally distributed, with a particular focus on heavy tailness (the most popular test for heavy tailed alternatives to normality is the Jarque-Bera test); frequently is not satisfied.

The Jarque-Bera test Let $X_1, X_2, \ldots, X_n$ be a sample of independent and identically distributed random variables. Let μ, ν and σ be the population mean, median and standard deviation respectively. Let $\bar{X}$, M and s_n be the corresponding sample estimates of μ, ν and σ. For any positive integer k define the k-th population central moment $\mu_k = E(X-\mu)^k$ and its sample estimate $\hat{\mu}_k = \sum_{i=1}^{n}(X_i - \bar{X})^k$. (Note that for $k = 1$ we obtain the population mean $\mu_1 = \mu$ with the sample estimate $\hat{\mu}_1 = \bar{X}$; for $k = 2$ we obtain the population variance $\mu_2 = \sigma^2$ with the sample estimate $\hat{\mu}_2 = s^2$.)

The Jarque-Bera (1980) test statistic (JB) is given by

$$\mathrm{JB} = \frac{\mathrm{n}}{6}\left(\frac{\hat{\mu}_3}{\hat{\mu}_2^{3/2}}\right)^2 + \frac{\mathrm{n}}{24}\left(\frac{\hat{\mu}_4}{\hat{\mu}_2^2} - 3\right)^2, \tag{11.8}$$

where $\hat{\mu}_3/\hat{\mu}_2^{3/2}$ is the sample estimate of the skewness $\sqrt{b_1}$ and $\hat{\mu}_4/\hat{\mu}_2$ is the sample estimate of the kurtosis b_2.

Under the null hypothesis of normality, the sample skewness and kurtosis are asymptotically normally distributed with covariance matrix defined as

$$\sqrt{n}\begin{bmatrix} \hat{\mu}_3/\hat{\mu}_2^{3/2} \\ \hat{\mu}_4/\hat{\mu}_2 - 3 \end{bmatrix} \Rightarrow N\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 6 & 0 \\ 0 & 24 \end{bmatrix}\right) \tag{11.9}$$

Since the two components of the Jarque-Bera test are asymptotically independent and normally distributed, the JB test statistic has the asymptotic χ^2-distribution with 2 degrees of freedom.

The Robust Jarque-Bera test, or the barefaced impudent promotion of our research.:) Since the sample variance is known to be even more sensitive to extreme observations than the mean (Kendalls Advanced Theory of Statistics by Stuart, A. and Ord, K.), the sample skewness and kurtosis will be even more sensitive to atypical observations in the data. In the paper by Gel and Gastwirth (Economics Letters, 2007), we propose to utilize the robust estimate of standard deviation which is less affected by outlying values, namely the average absolute deviation from the sample median (MAAD), in the denominators of the sample estimates of skewness and kurtosis instead of the classical estimator of spread. MAAD is used to evaluate the fairness of tax assessments (Gastwirth, 1982) and defined by

$$J_n = \frac{C}{n}\sum_{i=1}^{n}|X_i - M|, \qquad C = \sqrt{\pi/2}. \tag{11.10}$$

We define the robust sample estimates of skewness and kurtosis by $\hat{\mu}_3/J_n^3$ and $\hat{\mu}_4/J_n^4$ respectively, which leads to the new robust Jarque-Bera (RJB) test statistic

$$\mathrm{RJB} = \frac{\mathrm{n}}{\mathrm{C_1}}\left(\frac{\hat{\mu}_3}{\mathrm{J_n^3}}\right)^2 + \frac{\mathrm{n}}{\mathrm{C_2}}\left(\frac{\hat{\mu}_4}{\mathrm{J_n^4}} - 3\right)^2, \tag{11.11}$$

where C_1 and C_2 are positive constants.

Theorem Let $X_1, X_2, \ldots, X_n \sim N(\mu, \sigma)$. Then

$$\sqrt{n}\begin{bmatrix} \hat{\mu}_3/J_n^3 \\ \hat{\mu}_4/J_n^4 - 3 \end{bmatrix} \Rightarrow N\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} C_1 & 0 \\ 0 & C_2 \end{bmatrix}\right), \tag{11.12}$$

where C_1 and C_2 are positive constants.

The simulation power study demonstrates that the new RJB test is more powerful than the Jarque-Bera and Shapiro-Wilk (SW) tests in detecting moderately heavy tailed departures from normality that are symmetric or moderately skewed.

Corollary The RJB test statistic asymptotically follows the χ^2-distribution with 2 degrees of freedom, i.e.

$$\mathrm{RJB} \sim \chi_2^2.$$

Therefore, the implied one-sided rejection region is

$$\text{reject } \mathrm{H_0} : \text{normality, if } \mathrm{RJB} \geq \chi^2_{1-\alpha,2}, \tag{11.13}$$

where $\chi^2_{1-\alpha,2}$ is the upper α-percentile of the χ_2-distribution with 2 degrees of freedom. The RJB is implemented in the R package *lawstat*.

Example 1 of fitting ARCH/GARCH in R. Simulated ARCH data.

```
> library(tseries)

> n <- 1100
>       a <- c(0.1, 0.5, 0.2)  # ARCH(2) coefficients
>       e <- rnorm(n)
>       x <- double(n)
>       x[1:2] <- rnorm(2, sd = sqrt(a[1]/(1.0-a[2]-a[3])))
>       for(i in 3:n)  # Generate ARCH(2) process
+ {x[i] <- e[i]*sqrt(a[1]+a[2]*x[i-1]^2+a[3]*x[i-2]^2)}
>
>  x.arch100 <- garch(x[1:100], order = c(0,2))  # Fit ARCH(2)

 ***** ESTIMATION WITH ANALYTICAL GRADIENT *****

> summary(x.arch100)   # Diagnostic tests

Call: garch(x = x[1:100], order = c(0, 2))

Model: GARCH(0,2)

Residuals:
      Min        1Q    Median        3Q       Max
-2.250054 -0.663296 -0.006635  0.795360  2.191260

Coefficient(s):
    Estimate  Std. Error  t value Pr(>|t|)
a0   0.06036     0.02138    2.823  0.00476 **
a1   0.78769     0.26860    2.933  0.00336 **
a2   0.01526     0.10977    0.139  0.88943
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Diagnostic Tests:
        Jarque Bera Test

data:  Residuals X-squared = 1.7561, df = 2, p-value = 0.4156


        Box-Ljung test

data:  Squared.Residuals X-squared = 0.0869, df = 1, p-value =
0.7682
```

```
>library(lawstat)
> rjb.test(na.omit(x.arch100$residuals), option = c("RJB"))
#### approximated critical values #####

        Robust Jarque Bera Test

data:  na.omit(x.arch100$residuals) X-squared = 1.4612, df = 2,
p-value = 0.4816

> rjb.test(na.omit(x.arch100$residuals), option = c("JB"))
#### approximated critical values #####
        Jarque Bera Test

data:  na.omit(x.arch100$residuals) X-squared = 1.7561, df = 2,
p-value = 0.4156

> plot(x.arch100)
```

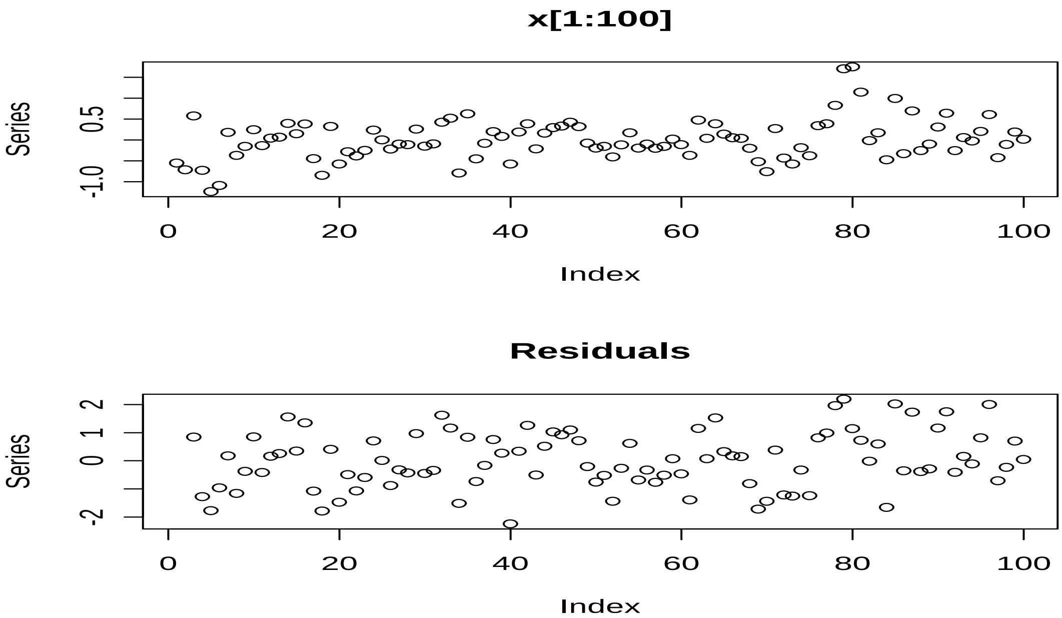

Figure 11.1: Time series plot of the data and ARCH(2) residuals.

Now compare the results if we take 1100 observations instead of 100 observations. We obtained MUCH more accurate estimates of ARCH parameters!

```
> summary(x.arch)
```

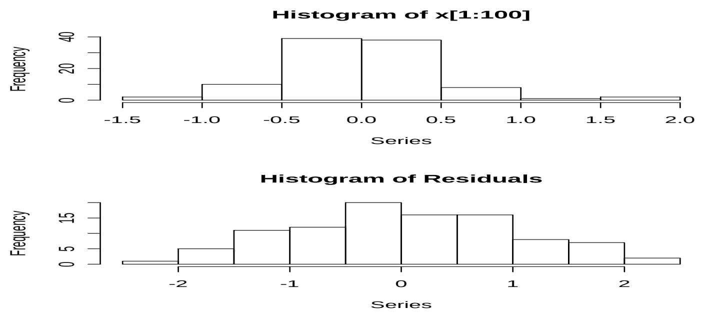

Figure 11.2: Histograms of the data and ARCH(2) residuals.

```
Call: garch(x = x, order = c(0, 2))

Model: GARCH(0,2)

Residuals:
     Min       1Q   Median       3Q      Max
-3.17715 -0.63414 -0.01502  0.68058  3.23393

Coefficient(s):
    Estimate  Std. Error  t value Pr(>|t|)
a0  0.101759    0.009489   10.724  < 2e-16 ***
a1  0.614048    0.067883    9.046  < 2e-16 ***
a2  0.159495    0.037664    4.235   2.29e-05 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Diagnostic Tests:
        Jarque Bera Test

data:  Residuals X-squared = 0.6398, df = 2, p-value = 0.7262


        Box-Ljung test

data:  Squared.Residuals X-squared = 0.675, df = 1, p-value = 0.4113
```

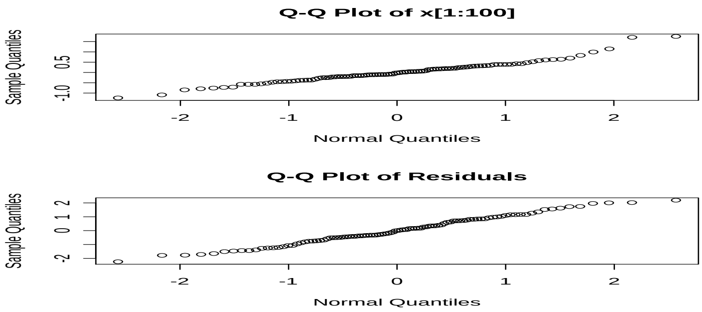

Figure 11.3: QQ plots of the data and ARCH(2) residuals.

```
###### to compare the models you can use "n.likeli"#######
###### the negative log-likelihood function evaluated######
######at the coefficient estimates (apart from some constant)######

> x.arch100$n.likeli
[1] -42.58789
> x.arch$n.likeli
[1] -223.7616
```

Example 2 of fitting ARCH/GARCH in R. Quarterly returns of the Canadian Real Estate Markets, from 1961 to 2005 (see Lecture 10).

```
>  m1 <- garch(Canada$Return, order = c(0,1))  # Fit ARCH(1)
 ***** ESTIMATION WITH ANALYTICAL GRADIENT *****
> m1$n.likeli
[1] -245.4842


>  m2 <- garch(Canada$Return, order = c(0,2))  # Fit ARCH(2)
 ***** ESTIMATION WITH ANALYTICAL GRADIENT *****
> m2$n.likeli [1] -246.1289

>  m11 <- garch(Canada$Return, order = c(1,1))  # Fit ARCH(1,1)
 ***** ESTIMATION WITH ANALYTICAL GRADIENT *****
> m11$n.likeli
[1] -251.2662
```

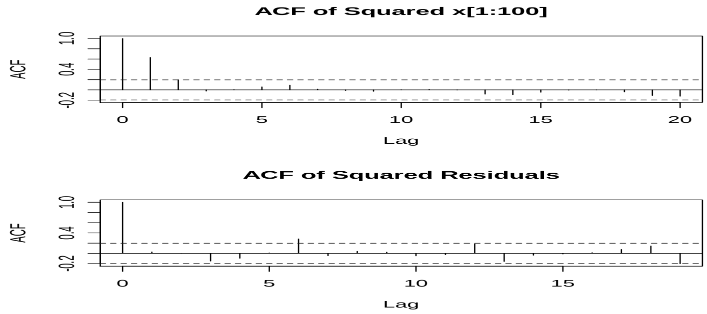

Figure 11.4: ACF plots of the data and ARCH(2) residuals.

```
>  m21 <- garch(Canada$Return, order = c(1,2))  # Fit ARCH(2,1)
 ***** ESTIMATION WITH ANALYTICAL GRADIENT *****
> m21$n.likeli
[1] -249.5579

>  m12 <- garch(Canada$Return, order = c(2,1))  # Fit ARCH(1,2)
 ***** ESTIMATION WITH ANALYTICAL GRADIENT *****
> m12$n.likeli
[1] -249.5985

>  m22 <- garch(Canada$Return, order = c(2,2))  # Fit ARCH(2,2)
 ***** ESTIMATION WITH ANALYTICAL GRADIENT *****
> m22$n.likeli
[1] -249.8246
```

So the most accurate model is GARCH(1,1) with the smallest negative loglikelihood of -251.27.
Now let us run the residual diagnostics.

```
> summary(m11)

Call: garch(x = Canada$Return, order = c(1, 1))

Model: GARCH(1,1)

Residuals:
```

```
     Min       1Q   Median       3Q      Max
-3.02966 -0.44995  0.02212  0.80259  3.19477

Coefficient(s):
    Estimate  Std. Error  t value Pr(>|t|)
a0  0.001864    0.001447    1.288   0.1977
a1  0.207490    0.085880    2.416   0.0157 *
b1  0.725111    0.100218    7.235 4.64e-13 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Diagnostic Tests:
        Jarque Bera Test

data:  Residuals X-squared = 2.7986, df = 2, p-value = 0.2468


        Box-Ljung test

data:  Squared.Residuals X-squared = 0.0456, df = 1, p-value =
0.8309
```

There are many other packages in R for modeling ARCH/GARCH processes.