

Chapter 9

ARCH and GARCH Processes

9.1 Introduction

When modelling time series, there are broadly speaking two approaches, the *fundamentalist* and the *data analyst*. The first of these aims to construct a model based on the fundamental principles of the situation and then use data to estimate the parameters so that the model may be used for forecasting and prediction. The second of these makes no attempt to search for underlying principles and simply tries to find a model that fits the available data. The ARCH and GARCH processes, motivated by observed volatility of financial series, fall firmly into the second category.

The ARCH and GARCH processes are designed for situations where, after fitting a stationary time series model with residuals that satisfy $\{X_t\} \sim \text{WN}(0, \sigma^2)$, the residuals do not correspond to $\text{IID}(0, \sigma^2)$, because the *volatility* seems variable, with bursts of high volatility. More information on the nature of the residuals can lead to more accurate prediction. The ARCH / GARCH processes are one attempt to study the situation of variable volatility.

Let $\{X_t : t \in \mathbb{Z}\}$ be a stationary, mean zero time series. That is, the mean and possible trend have already been removed. Suppose that the $\{X_t\}$ appear uncorrelated, so that the $\text{WN}(0, \sigma^2)$ seems justified, but that $\text{IID}(0, \sigma^2)$ does not seem justified, because the ‘volatility’ seems variable. One approach is to model this by processes of the form:

$$\begin{cases} X_t = \sigma_t Z_t & \{Z_t\} \sim \text{IID}N(0, 1) \\ \sigma_t \in \mathcal{F}_{t-1}^{(Z)} \end{cases} \quad (9.1)$$

where $\mathcal{F}_{t-1}^{(Z)}$ denotes the collection $\{Z_s : s \leq t-1\}$ and $\sigma_t \in \mathcal{F}_{t-1}^{(Z)}$ denotes that σ_t is a function of these random variables.

9.2 The ARCH Process

A prominent model of this class is the so-called ARCH process, introduced by R.F. Engle (1982). The abbreviation ARCH stands for *autoregressive conditional heteroscedasticity*.

Definition 9.1. *The process $\{X_t : t \in \mathbb{Z}\}$ is said to be an ARCH(p) process if it is stationary and*

$$\begin{cases} X_t = \sigma_t Z_t & \{Z_t\} \sim IIDN(0, 1) \\ \sigma_t^2 = \alpha_0 + \alpha_1 X_{t-1}^2 + \dots + \alpha_p X_{t-p}^2 & \alpha_0 > 0, \quad \alpha_j \geq 0 \quad j = 1, \dots, p \end{cases} \quad (9.2)$$

The requirement that $\alpha_0 > 0$ and $\alpha_j \geq 0$ for $j = 1, \dots, p$ ensure that $\sigma_t > 0$. The coefficients $\alpha_0, \alpha_1, \dots, \alpha_p$ have to be chosen such that the process $\{X_t : t \in \mathbb{Z}\}$ defined in this way is stationary. The following lemma, which holds for all processes of the form given by Equation (9.1), helps to establish when $\{X_t^2\}$ may be represented as a linear stationary process with $WN(0, \sigma^2)$ innovations, for some well defined $\sigma^2 < +\infty$.

Lemma 9.2. *Let $\{X_t\}$ be a stationary process satisfying Equation (9.1). Let*

$$\tilde{Z}_t = X_t^2 - \sigma_t^2 = \sigma_t^2(Z_t^2 - 1). \quad (9.3)$$

Assume that $\{\sigma_t : t \in \mathbb{Z}\}$ is a stationary process satisfying $\mathbb{E}[\sigma_t^4] < +\infty$, then $\{\tilde{Z}_t : t \in \mathbb{Z}\}$ is $WN(0, \sigma^2)$ where $\sigma^2 = 2\mathbb{E}[\sigma_t^4]$.

Proof By construction,

$$\mathbb{E}[\tilde{Z}_t] = \mathbb{E}[\sigma_t^2(Z_t^2 - 1)] = \mathbb{E}[\sigma_t^2](\mathbb{E}[Z_t^2] - 1) = 0$$

and

$$\mathbb{E}[\tilde{Z}_t^2] = \mathbb{E}[\sigma_t^4]\mathbb{E}[Z_t^4 - 2Z_t^2 + 1] = 2\mathbb{E}[\sigma_t^4] = \sigma^2.$$

and the proof is complete. □

Lemma 9.3. *Let $\{X_t : t \in \mathbb{Z}\}$ be an ARCH(p) process defined by Equation (9.2) satisfying $\mathbb{E}[X_t^4] < +\infty$. Let \tilde{Z} be defined by Equation (9.3). Then*

$$X_t^2 - \sum_{j=1}^p \alpha_j X_{t-j}^2 = \alpha_0 + \tilde{Z}_t.$$

If $\{X_t\}$ is defined by Equation (9.2), it follows that $\{X_t^2\}$ is an AR(p) process with mean

$$\mathbb{E}[X_t^2] = \mu = \frac{\alpha_0}{1 - \sum_{j=1}^p \alpha_j}$$

if and only if

1. $\sum_{j=1}^p \alpha_j < 1$ and
2. $\mathbb{E}[X_t^4] < +\infty$.

Proof

$$X_t^2 - \sigma_t^2 = \tilde{Z}_t = X_t^2 - \alpha_0 - \sum_{j=1}^p \alpha_j X_{t-j}^2$$

giving

$$X_t^2 - \sum_{j=1}^p \alpha_j X_{t-j}^2 = \alpha_0 + \tilde{Z}_t.$$

This has the form of an AR(p) process. Taking expectations gives

$$\mu \left(1 - \sum_{j=1}^p \alpha_j \right) = \alpha_0$$

from which it follows that

$$\mu = \frac{\alpha_0}{1 - \sum_{j=1}^p \alpha_j}.$$

It remains to show that $\mathbb{E}[\tilde{Z}_t^2] < +\infty$ if and only if $\mathbb{E}[X_t^4] < +\infty$. This follows directly from:

$$\begin{aligned} \mathbb{E} \left[\left(\tilde{Z}_t + \alpha_0 \right)^2 \right] &= \left(1 - \sum_{j=1}^p \alpha_j \right)^2 \mathbb{E}[X_t^4] \\ &+ 2 \sum_{j=1}^p \alpha_j \left(1 - \sum_{k=1}^p \alpha_k \right) \mathbb{E}[X_t^2 (X_t^2 - X_{t-j}^2)] + \sum_{j=1}^p \sum_{k=1}^p \alpha_j \alpha_k \mathbb{E}[(X_t^2 - X_{t-j}^2)(X_t^2 - X_{t-k}^2)] \\ &= I + II + III. \end{aligned}$$

Firstly, $\mathbb{E}[\tilde{Z}_t^2] < +\infty \Leftrightarrow \mathbb{E}[(\tilde{Z} + \alpha_0)^2] < +\infty$. Clearly each of these terms is non-negative;

$$I = c^2 \mathbb{E}[X_t^4] \quad c = 1 - \sum_{j=1}^p \alpha_j > 0.$$

Since $\mathbb{E}[X_t^2 X_{t-j}^2] \leq \mathbb{E}[X_t^4]$ by Hölder together with stationarity, it follows that

$$0 \leq II \leq 2 \left(\sum_{j=1}^p \alpha_j \right) \left(1 - \sum_{j=1}^p \alpha_j \right) \mathbb{E}[X_t^4] = 2c(1-c) \mathbb{E}[X_t^4]$$

Finally, let $\gamma(j, k) = \mathbb{E}[(X_t^2 - X_{t-j}^2)(X_t^2 - X_{t-k}^2)]$ then γ is a covariance matrix and hence non negative definite. It follows that $\underline{\alpha}^t \gamma \underline{\alpha} > 0$, so that $III > 0$. It follows that

$$0 \leq III \leq 4(1-c)^2 \mathbb{E}[X_t^4].$$

The if and only if condition now follows directly. □

Although it is the process $\{X_t\}$ that is of interest, the process $\{X_t^2\}$ and its representation as an AR(p) process, is very useful for parameter estimation. Most statistical tests rely on a central limit theorem effect, which requires the innovations to have a well defined finite variance.

For general models, there may be difficulties establishing conditions on the coefficients so that the innovations for the AR(p) process corresponding to $\{X_t^2\}$ are $\text{WN}(0, \sigma^2)$ for a well defined $\sigma^2 < +\infty$. The following example shows the parameter range for the ARCH(1) process.

Example 9.1 (ARCH(1) process).

Let $\{X_t\}$ satisfy

$$\begin{cases} X_t = \sigma_t Z_t & \{Z_t\} \sim \text{IID}N(0, 1) \\ \sigma_t^2 = \alpha_0 + \alpha_1 X_{t-1}^2. \end{cases}$$

Suppose that $\mathbb{E}[X_t^2] = \mu < +\infty$. By stationarity, it follows from the first equation that

$$\mathbb{E}[X_t^2] = \mathbb{E}[\sigma_t^2] = \mu > 0.$$

From the second, it therefore follows that

$$\mu = \alpha_0 + \alpha_1 \mu \Rightarrow \mu = \mathbb{E}[X_t^2] = \frac{\alpha_0}{1 - \alpha_1},$$

in line with the general result.

Now consider the time series $\{X_t^2 : t \in \mathbb{Z}\}$. With $\tilde{Z}_t = Z_t^2 - \sigma_t^2$,

$$X_t^2 - \sigma_t^2 = \tilde{Z}_t = X_t^2 - \alpha_0 - \alpha_1 X_{t-1}^2$$

so that

$$X_t^2 - \alpha_1 X_{t-1}^2 = \alpha_0 + \tilde{Z}_t.$$

It follows that, provided $\mathbb{E}[\tilde{Z}_t^2] < +\infty$ and $\alpha_1 < 1$, the process $\{X_t^2\}$ is a causal AR(1) process with mean $\mu = \frac{\alpha_0}{1 - \alpha_1}$. For the AR(1) process, precise conditions on the coefficients can be established to ensure that $\{X_t^2\}$ is an AR(1) process; the following computation shows that this is the case if and only if $\alpha_1^2 < \frac{1}{3}$.

Using

$$X_t = \sigma_t Z_t \Rightarrow X_t^4 = \sigma_t^4 Z_t^4 = (\alpha_0 + \alpha_1 X_{t-1}^2)^2 Z_t^4,$$

it follows, using $\mathbb{E}[Z_t^4] = 3$, that

$$\begin{aligned} \mathbb{E}[X_t^4] &= 3\mathbb{E}[\alpha_0^2 + 2\alpha_0\alpha_1 X_{t-1}^2 + \alpha_1^2 X_{t-1}^4] \\ &= 3\left(\alpha_0^2 + \frac{2\alpha_0^2\alpha_1}{1 - \alpha_1}\right) + 3\alpha_1^2\mathbb{E}[X_{t-1}^4] = \frac{3\alpha_0^2(1 + \alpha_1)}{1 - \alpha_1} + 3\alpha_1^2\mathbb{E}[X_t^4] \end{aligned}$$

Clearly, $\mathbb{E}[X_t^4] = +\infty$ for $3\alpha_1^2 \geq 1$. For $3\alpha_1^2 < 1$,

$$\mathbb{E}[X_t^4] = \frac{3\alpha_0^2(1 + \alpha_1)}{(1 - \alpha_1)(1 - 3\alpha_1^2)}.$$

It follows that $\{X_t^2\}$ is a causal AR(1) process if and only if $\alpha_1 < \frac{1}{\sqrt{3}}$. □

Since $\mathbb{E}[X_t] = 0$, it follows that the *kurtosis* of X_t is:

$$\frac{\mathbb{E}[X_t^4]}{\text{Var}(X_t)^2} - 3 = 3 \times \frac{\alpha_0^2(1 + \alpha_1)}{(1 - \alpha_1)(1 - 3\alpha_1^2)} \times \frac{(1 - \alpha_1)^2}{\alpha_0^2} - 3 = 3 \frac{1 - \alpha_1^2 - 1 + 3\alpha_1^2}{1 - 3\alpha_1^2} = \frac{6\alpha_1^2}{1 - 3\alpha_1^2} > 0.$$

The kurtosis of an ARCH(p) process is always greater than 0, indicating thicker tails than a normal distribution. □

Generating Polynomials for the ARCH(p) Process

Let

$$\alpha(z) = \alpha_1 z + \dots + \alpha_p z^p.$$

Note that the α_0 has not been included in the polynomial. Then the equation for σ_t^2 may be written in the form:

$$\sigma_t^2 = \alpha_0 + \alpha(B)X_t^2.$$

In this notation, $\sum_{j=1}^p \alpha_j = \alpha(1)$ and hence the formula for $\mathbb{E}[X_t^2]$ may be written as:

$$\mathbb{E}[X_t^2] = \alpha_0 + \alpha(1)\mathbb{E}[X_t^2] \Rightarrow \mathbb{E}[X_t^2] = \frac{\alpha_0}{1 - \alpha(1)}.$$

9.2.1 ARCH - t Process

Sometimes the underlying noise $\{Z_t\}$ is not IID $N(0, 1)$; it is heavier tailed. Instead, IID t may be used; the degrees of freedom of the t distribution may be altered to fit the empirical distribution of the residuals.

9.3 Testing for the ARCH Effect

Let $(z_t)_{t=1}^T$ denote the estimates of the innovations of the time series. There are two prominent tests available to determine whether a data set exhibits an ARCH effect. The first is simply the *Ljung - Box* test. The Ljung-Box test determines whether or not the acf is significantly different from the acf of white noise. This is applied to both the series $(z_t)_{t=1}^T$ and $(z_t^2)_{t=1}^T$. If the first series is white noise, but the second is not, then this is an indicator of ARCH effects.

9.3.1 Testing whether a series is white noise

In general, in Time Series, it is important to test whether or not a ‘white noise’ model fits. To indicate whether or not there are ARCH effects, we also consider the series $(Z_t)_{t \geq 1}$.

As ever, $\rho(0) = 1$. If the series is white noise, then $\rho(1) = \rho(2) = \dots = \rho(h) = 0$ for all $h \geq 1$. In this situation, the Bartlett formula (5.1) (recall estimating the ACVF and ACF) reduces to (for $1 \leq i \leq j < +\infty$):

$$w_{ij} = \begin{cases} 0 & i \neq j \\ 1 & i = j \end{cases}$$

so that $\hat{\rho} \sim AN(0, \frac{1}{n}I)$. Hence, consider testing:

$$\begin{cases} H_0 : & \text{The data are an observed random sample from } WN(0, \sigma^2) \\ H_1 & \text{The data are not an observed random sample from } WN(0, \sigma^2). \end{cases}$$

The test statistic

$$Q := n \sum_{j=1}^h \hat{\rho}(j)^2 \xrightarrow[n \rightarrow +\infty]{(d)} \chi_h^2;$$

the null hypothesis is rejected for $Q > \chi_{h;\alpha}^2$ where $\chi_{h;\alpha}^2$ is the value such that $\mathbb{P}(X > \chi_{h;\alpha}^2) = \alpha$ for $X \sim \chi_h^2$.

The Ljung - Box test The Ljung–Box test (named after Greta M. Ljung and George E. P. Box) is a statistical test of whether a of a group of autocorrelations of a time series are different from zero and the test statistic is defined as:

$$Q = n(n+2) \sum_{k=1}^h \frac{\hat{\rho}^2(k)}{n-k}$$

where n is the sample size, $\hat{\rho}(k)$ is the sample autocorrelation at lag k , and h is the number of lags being tested. Clearly, Q has asymptotic χ^2 distribution, since $\lim_{n \rightarrow +\infty} \frac{n+2}{n-k} = 1$.

The Ljung-Box statistic takes into account the fact that $\hat{\gamma}(h) = \frac{1}{n} \sum_{j=1}^{n-h} (X_j - \bar{X})(X_{j+h} - \bar{X})$ where a $\frac{1}{n}$ is used instead of $\frac{1}{n-h-1}$.

Degrees of Freedom for Ljung-Box Note that the test is applied to the *residuals* of the fitted model, not the original series. This means, for example, that if a causal invertible ARMA(p,q) is fitted to the original series $\{X_t\}_{t \in \mathbb{Z}}$:

$$\phi(B)X_t = \theta(B)Z_t$$

then $\{Z_t\}_{t \in \mathbb{Z}}$ is recovered by:

$$Z_t = \theta(B)^{-1}\phi(B)X_t$$

and the Ljung-Box test applied to $\{Z_t\}_{t \in \mathbb{Z}}$.

In such applications the hypothesis being tested is that the residuals have no autocorrelation. When testing the residuals of an estimated model, the degrees of freedom need to be adjusted to reflect the number of parameters estimated. In a model with $p+q$ estimated parameters (for example ARMA(p,q)), the degrees of freedom should be set to $h-p-q$.

The distribution of the test statistic is derived under the assumption that the residuals are *normally* distributed; when the data is not normal, it is approximate and relies on a central limit theorem effect.

9.3.2 Engle's test

The second test available is the *Lagrange multiplier test* of Engle (1982). This test is equivalent to the usual F statistic for testing $\alpha_j = 0$ for all $j = 1, \dots, p$ in the linear regression

$$z_t^2 = \alpha_0 + \alpha_1 z_{t-1}^2 + \dots + \alpha_p z_{t-p}^2 + \epsilon_t \quad t = p + 1, \dots, T. \quad (9.4)$$

where ϵ_t denotes the error term, p is a pre-specified positive integer and T is the sample size. The null hypothesis is:

$$H_0 : \alpha_1 = \dots = \alpha_m = 0.$$

Let $Q_0 = \sum_{t=p+1}^T (z_t^2 - \tilde{\omega})^2$, where $\tilde{\omega} = \frac{1}{T} \sum_{t=1}^T z_t^2$ is the sample mean of $(z_t^2)_{t=1}^T$ and let $Q_1 = \sum_{t=p+1}^T \hat{\epsilon}_t^2$ where $\hat{\epsilon}_t$ is the least squares residual estimate for the regression problem of Equation (9.4). Then

$$F = \frac{(SSR_0 - SSR_1)/p}{SSR_1/(T - 2p - 1)} \sim_{T \rightarrow +\infty} \chi_p^2.$$

The decision rule is to reject H_0 if $F > \chi_{p,\alpha}^2$. Again, note that the statistic is based on a central limit theorem effect.

9.3.3 Example

Consider the monthly log stock returns of Intel Corporation from 1973 - 2008, found in the file `m-intc7308.txt`. Figures 9.1, 9.2, 9.3, 9.4 and 9.5 indicate that there is conditional heteroscedasticity. Firstly, the time series plot indicates clusters of high volatility. Secondly, while the acf of the log stock returns shows no significant serial correlations with only minor exceptions at 7 and 14, the acf and pacf of the squared log returns indicates show serial correlations, indicating that the monthly returns are not independent.

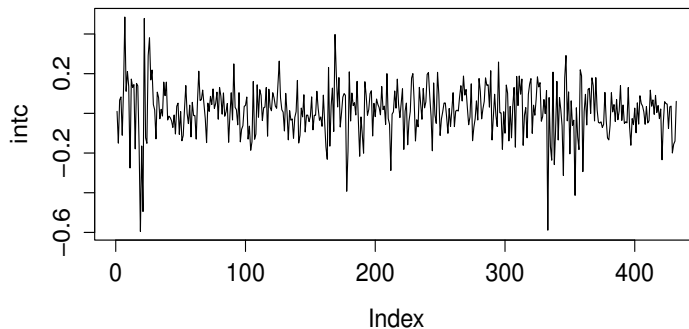


Figure 9.1: Intel log stock returns time series plot

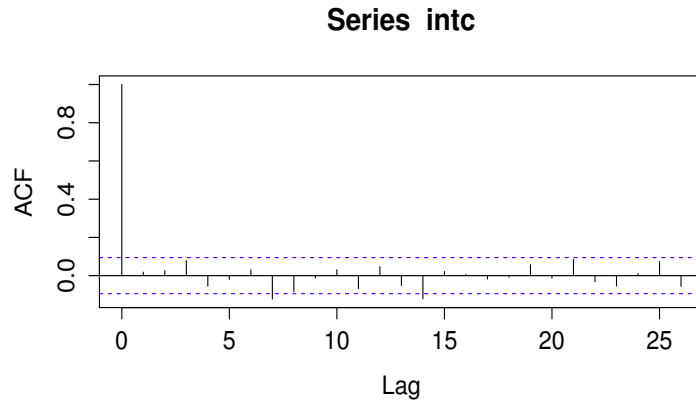


Figure 9.2: Intel log stock returns acf

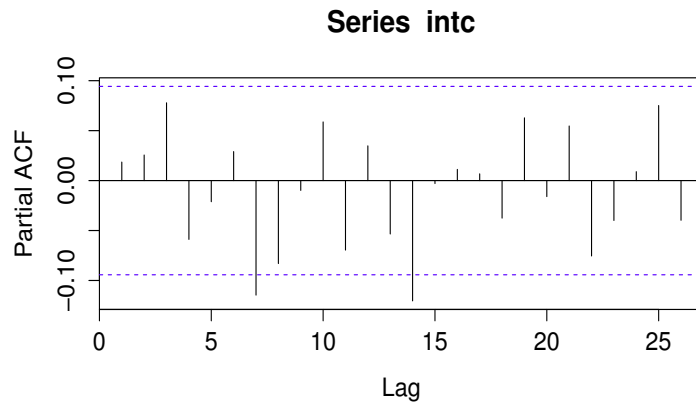


Figure 9.3: Intel log stock returns pacf

These observations are confirmed by the Ljung - Box test: the $Q(m)$ statistics of the log return series give $Q(12) = 18.26$ with a p value of 0.11 confirming no serial correlations in the data (significance level 5%). On the other hand, the Ljung - Box test on the squared log return series gives a value of $Q(12) = 89.85$ with a p value close to 0, thus indicating that strong ARCH effects are present.

The R script is as follows:

```
www =
"https://www.mimuw.edu.pl/~noble/courses/TimeSeries/data/m-intc7308.
txt"
mintc7308 = read.table(www,header=T)
da = mintc7308
intc = log(da[,2]+1)
plot(intc, type = "l")
```

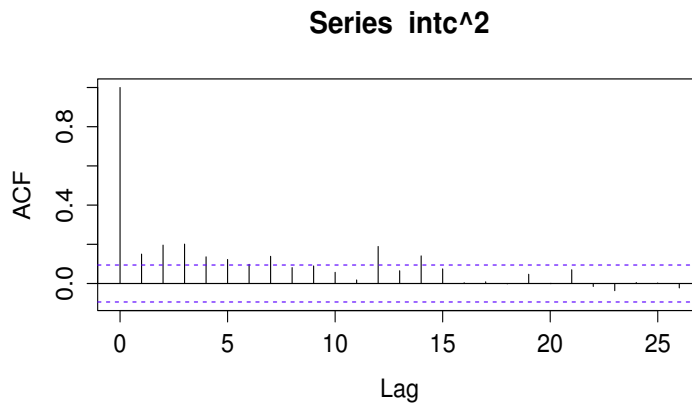



Figure 9.4: Intel log stock returns squared acf

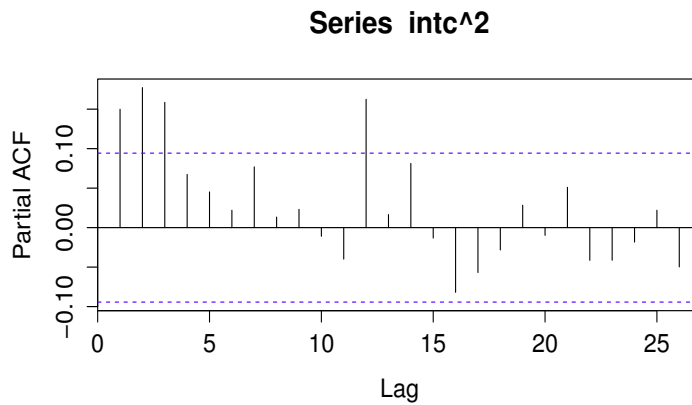


Figure 9.5: Intel log stock returns squared pacf

```
acf(intc)
pacf(intc)
acf(intc^2)
pacf(intc^2)
```

to get the acf and pacf of the series and the squared values. The command for the Ljung Box test is:

```
Box.test(intc,lag=12,type='Ljung')
```

and the output of the Box-Ljung test is:

```
Box-Ljung test
```

```
data: intc
```

X-squared = 18.2635, df = 12, p-value = 0.1079

Now remove the mean and take the square:

```
at=intc-mean(intc)
Box.test(at^2,lag=12,type='Ljung')
```

and the result of the test is:

Box-Ljung test

```
data: at^2
X-squared = 89.8509, df = 12, p-value = 5.274e-14
```

9.4 Building an ARCH model

Using the fact that X_t^2 is an AR(p) process, it follows that the pacf of $\{X_t^2\}$ is 0 for lags greater than p . The sample pacfs may therefore be used to select an ARCH model. Having selected the order, there are several methods for estimating the parameters. One way is to consider $\{X_t^2 : t \in \mathbb{Z}\}$, subtract the mean, and apply the techniques of fitting an AR(p) process.

Likelihood Method Under the distributional assumption that $\{Z_t\} \sim IIDN(0,1)$, the likelihood function of an ARCH(p) model is:

$$\begin{aligned} f(z_1, \dots, z_T | \underline{\alpha}) &= f(z_T | \mathcal{F}_{T-1}) f(z_{T-1} | \mathcal{F}_{T-2}) \dots f(z_{p+1} | \mathcal{F}_p) f(z_1, \dots, z_p | \underline{\alpha}) \\ &= \left(\prod_{t=p+1}^T \frac{1}{\sqrt{2\pi\sigma_t^2}} \exp \left\{ \frac{-z_t^2}{2\sigma_t^2} \right\} \right) f(z_1, \dots, z_p | \underline{\alpha}), \end{aligned}$$

where $\underline{\alpha} = (\alpha_0, \dots, \alpha_p)^t$. It follows that the log likelihood function \mathcal{L} is:

$$\mathcal{L}(z_{p+1}, \dots, z_T | \underline{\alpha}) = \text{const} - \sum_{t=p+1}^T \left\{ \frac{1}{2} \ln(\sigma_t^2) + \frac{1}{2} \frac{z_t^2}{\sigma_t^2} \right\},$$

where

$$\sigma_t^2 = \alpha_0 + \alpha_1 z_{t-1}^2 + \dots + \alpha_p z_{t-p}^2.$$

The conditional likelihood is usually used since $f(z_1, \dots, z_p | \underline{\alpha})$ is usually rather complicated.

In heavy tailed applications, the normal density may be replaced by a standardised Student-t distribution with degrees of freedom chosen to reflect the desired tail. Other distributions may be chosen to reflect the fact that the innovations are often skewed.

9.5 Forecasting

Forecasts from the ARCH model can be obtained recursively in exactly the same way as those for an AR model. Consider an ARCH(p) model; at the forecast origin h , the 1-step ahead forecast of σ_{h+1}^2 is:

$$\sigma_h^2(1) = \alpha_0 + \alpha_1 X_h^2 + \dots + \alpha_p X_{h+1-p}^2.$$

Since σ_{h+1}^2 is \mathcal{F}_h measurable, it follows that $\sigma_h^2(1) = \sigma_{h+1}^2$.

The two step ahead forecast is:

$$\sigma_h^2(2) = \alpha_0 + \alpha_1 \sigma_h^2(1) + \alpha_2 X_h^2 + \dots + \alpha_p X_{h+2-p}^2$$

which follows because $\mathbb{E}[X_{h+1}^2 | \mathcal{F}_h] = \sigma_{h+1}^2 = \sigma_h^2(1)$.

Similarly, the l step ahead forecast for σ_{h+l}^2 is:

$$\begin{aligned} \sigma_h^2(l) &= \alpha_0 + \sum_{i=1}^p \alpha_i \mathbb{E}[X_{h+l-1}^2 | \mathcal{F}_h] \\ &= \alpha_0 + \sum_{i=1}^p \alpha_i \sigma_h^2(l-i) \end{aligned}$$

where

$$\sigma_h^2(l-i) = X_{h+l-i}^2 \quad l-i \leq 0.$$

Example For fitting a GARCH model, the package `fGarch` package may be used; the command is `garchFit`. I found it easier (using Kubuntu 21.04) to install `fGarch` using ‘Synaptic Package Manager’; the package is `r-cran-fgarch`.

Again, the data considered is the log stock returns of Intel Corporation from 1973 - 2008. The ARCH(1) model returned is:

$$\begin{cases} R_t = 0.0126 + X_t \\ \sigma_t^2 = 0.0112 + 0.3795 X_{t-1}^2 \end{cases}$$

All the parameter estimates are highly significant.

The acf of the residuals $\{x_t\}$ gives $Q(10) = 12.54$ with a p value of 0.25 and those of $\{x_t^2\}$ give $Q(10) = 16.02$ with a p value of 0.10, neither significant at the 5% level, indicating that the model fits.

Using the same data as before (and the series `intc`), the command for fitting is:

```
library(fGarch)
m1 = garchFit(intc~garch(1,0),data=intc,trace=F)
```

A summary of the fitted model is found by:

```
> summary(m1)
```

Title:

GARCH Modelling

Call:

```
garchFit(formula = intc ~ garch(1, 0), data = intc, trace = F)
```

Mean and Variance Equation:

```
data ~ garch(1, 0)
```

```
<environment: 0x55b62d2ca050>
```

```
[data = intc]
```

Conditional Distribution:

norm

Coefficient(s):

```
      mu      omega      alpha1
0.012637 0.011195 0.379492
```

Std. Errors:

based on Hessian

Error Analysis:

	Estimate	Std. Error	t value	Pr(> t)
mu	0.012637	0.005428	2.328	0.01990 *
omega	0.011195	0.001239	9.034	< 2e-16 ***
alpha1	0.379492	0.115534	3.285	0.00102 **

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Log Likelihood:

```
288.0589      normalized: 0.6668031
```

Description:

```
Thu Apr 22 10:02:36 2021 by user:
```

Standardised Residuals Tests:

Statistic p-Value

Jarque-Bera Test	R	Chi ²	137.919	0
Shapiro-Wilk Test	R	W	0.9679248	4.024057e-08
Ljung-Box Test	R	Q(10)	12.54002	0.2505382
Ljung-Box Test	R	Q(15)	21.33508	0.1264607
Ljung-Box Test	R	Q(20)	23.19679	0.2792353
Ljung-Box Test	R ²	Q(10)	16.0159	0.09917814
Ljung-Box Test	R ²	Q(15)	36.08022	0.001721296
Ljung-Box Test	R ²	Q(20)	37.43683	0.01036728
LM Arch Test	R	TR ²	26.57744	0.008884586

Information Criterion Statistics:

AIC	BIC	SIC	HQIC
-1.319717	-1.291464	-1.319813	-1.308563

Let us describe the contents of the output. We'll deal with GARCH slightly later in the lecture. For now, an ARCH(p) process is a GARCH($p,0$) process. Here $p = 1$ and $q = 0$.

The μ , ω and α_1 refer to estimates for the model:

$$\begin{cases} X_t = \mu + a_t \\ a_t = \sigma_t \epsilon_t \end{cases}$$

where $\{\epsilon_t\} \sim IIDN(0, 1)$ and

$$\sigma_t^2 = \omega + \alpha_1 a_{t-1}^2.$$

Here the μ is insignificant, while the ω is very significant and, since α_1 is significant, there is a significant ARCH effect.

Next follow the results of a series of tests on the autocorrelations. We have considered the Ljung-Box test. Computing the *power* of the Ljung-Box test, and deciding on the optimal number of lags to give the most powerful test available is not feasible. Therefore, several lags are chosen. We only need *one* of these to reject the null hypothesis to conclude that the process is *not* white noise.

We see that, for the original series (denoted R), the null hypothesis is *not* rejected for a Ljung-Box test at 10, 15 or 20 lags. The white-noise hypothesis looks reasonable.

For the series R^2 , the null hypothesis of 'white noise' is clearly rejected; the test at 15 lags is the most powerful of the Ljung-Box tests offered.

An ARCH(1) model with Student t distribution may be fitted as follows:

```
m3 = garchFit(intc~garch(1,0),data=intc,trace=F,cond.dist='std')
```

The *beginning* of the summary is:

```
> summary(m3)
```

```

Title:
  GARCH Modelling

Call:
  garchFit(formula = intc ~ garch(1, 0), data = intc, cond.dist =
"std",
  trace = F)

Mean and Variance Equation:
  data ~ garch(1, 0)
<environment: 0x55b62fed17f0>
  [data = intc]

Conditional Distribution:
  std

Coefficient(s):
      mu      omega    alpha1    shape
0.016731  0.011939  0.285321  6.015194

Std. Errors:
  based on Hessian

Error Analysis:
      Estimate Std. Error t value Pr(>|t|)
mu      0.016731   0.005302   3.155 0.001603 **
omega   0.011939   0.001603   7.449 9.39e-14 ***
alpha1  0.285321   0.110608   2.580 0.009892 **
shape   6.015194   1.562619   3.849 0.000118 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

```

The additional piece here is the *shape*. Here $\{\epsilon_t\} \sim IIDt_\nu$ where ν is the number of degrees of freedom, estimated by *shape*.

The introduction of the additional parameter has led to a *slight* improvement in terms of the AIC. It turns out that this data set needs a GARCH process; we now turn to GARCH.

9.6 The GARCH Process

The GARCH process is defined as:

$$\begin{cases} X_t = \sigma_t Z_t & \{Z_t\} \sim \text{IIDN}(0, 1) \\ \sigma_t^2 = \alpha_0 + \alpha_1 X_{t-1}^2 + \dots + \alpha_p X_{t-p}^2 + \beta_1 \sigma_{t-1}^2 + \dots + \beta_q \sigma_{t-q}^2 \\ \alpha_0 > 0, \quad \alpha_j \geq 0 \quad j = 1, \dots, p \quad \beta_k \geq 0 \quad k = 1, \dots, q \end{cases}$$

The equation for σ^2 may be written, in terms of generating polynomials, as

$$\sigma_t^2 = \alpha_0 + \alpha(B)X_t^2 + \beta(B)\sigma_t^2,$$

where

$$\begin{cases} \alpha(z) = \alpha_1 z + \dots + \alpha_p z^p \\ \beta(z) = \beta_1 z + \dots + \beta_q z^q \end{cases}$$

Using $\mathbb{E}[X_t^2] = \mathbb{E}[\sigma_t^2]$ together with stationarity gives:

$$\mathbb{E}[X_t^2] = \alpha_0 + (\alpha(1) + \beta(1))\mathbb{E}[X_t^2] \Rightarrow \mathbb{E}[X_t^2] = \frac{\alpha_0}{1 - \alpha(1) - \beta(1)}.$$

Again using $\tilde{Z}_t = X_t^2 - \sigma_t^2 = \sigma_t^2(Z_t^2 - 1)$, so that $\tilde{Z}_t \sim \text{WN}(0, \sigma^2)$ provided that $\mathbb{E}[\sigma_t^4] < +\infty$, it follows that

$$X_t^2 - \tilde{Z}_t = \alpha_0 + \alpha(B)X_t^2 + \beta(B)(X_t^2 - \tilde{Z}_t).$$

This may be written as

$$X_t^2 - (\alpha(B) + \beta(B))X_t^2 = \alpha_0 + (1 - \beta(B))\tilde{Z}_t.$$

Provided $\mathbb{E}[\tilde{Z}_t^2] < +\infty$, this is an ARMA($\max(p, q), q$) process with mean $\mu = \mathbb{E}[X_t^2]$ given by:

$$\mu - (\alpha(1) + \beta(1))\mu = \alpha_0 \Rightarrow \mu = \frac{\alpha_0}{1 - (\alpha(1) + \beta(1))}.$$

9.7 An Illustrative Example

Specifying the order of a GARCH process is not so easy. Only lower order GARCH processes are used in most applications. Consider the monthly excess returns of S&P 500 index starting from 1926 for 792 observations. The ACF is shown in Figure 9.7 while the partial autocorrelations for the squared excess returns is shown in Figure 9.8. The series of returns has some serial correlations at lags 1 and 3, but the key feature of the pacf of x_t^2 is that it shows strong linear dependence and hence an MA(3) model may seem appropriate.

An AR(3) model has the same number of parameters; it is considered here. The fitted model is:

$$\begin{cases} x_t - 0.088x_{t-1} + 0.023x_{t-2} + 0.123x_{t-3} - 0.0066 = a_t \\ \hat{\sigma}_a^2 = 0.00333 \end{cases}$$

A joint estimation of the AR(3) - GARCH(1,1) model gives:

$$\begin{cases} x_t = 0.0078 + 0.032x_{t-1} - 0.029x_{t-2} - 0.008x_{t-3} + a_t \\ \sigma_t^2 = 0.000084 + 0.1213a_{t-1}^2 + 0.8523\sigma_{t-1}^2 \end{cases}$$

From the volatility equation, the implied (unconditional) variance of a_t is:

$$\frac{0.000084}{1 - 0.8523 - 0.1213} = 0.00317.$$

which is in line with the variance estimate for the AR(3) model.

In the AR(3) - GARCH(1,1) model, however, the parameters for the three AR coefficients are insignificant at the 5% level. It therefore seems appropriate to redefine the model by dropping all AR parameters. The redefined model is:

$$\begin{cases} x_t = 0.0076 + a_t \\ \sigma_t^2 = 0.000086 + 0.1216a_{t-1}^2 + 0.8511\sigma_{t-1}^2 \end{cases}$$

The Ljung-Box statistics suggest that the model is adequate. The R code is as follows:

```
> ww2 =
"https://www.mimuw.edu.pl/~noble/courses/TimeSeries/data/sp500.dat"
> sp500 = read.table(ww2)
> sp500fit = garchFit(~arma(3,0)+garch(1,1),data=sp500,trace=F)
> summary(sp500fit)
```

A student t distribution can be added; this does not improve the model significantly.

```
sptfit = garchFit(~garch(1,1),data=sp500,trace=F,cond.dist="std")
#obtain standardised residuals
stresi = residuals(sptfit,standardize=T)
plot(stresi,type="l")
Box.test(stresi,10,type='Ljung')
```

Box-Ljung test

```
data:  stres_i
X-squared = 11.38, df = 10, p-value = 0.3287
```

The result is not significant; the null hypothesis is not rejected.

Let us now predict 5 time steps ahead for this model.


```

predict(sptfit,5)
  meanForecast meanError standardDeviation
1  0.008455033 0.05330091      0.05330091
2  0.008455033 0.05327888      0.05327888
3  0.008455033 0.05325782      0.05325782
4  0.008455033 0.05323770      0.05323770
5  0.008455033 0.05321847      0.05321847

```

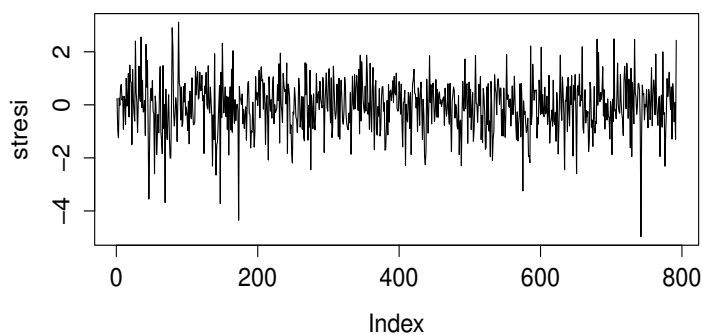


Figure 9.6: Standard and Poor standardised residuals

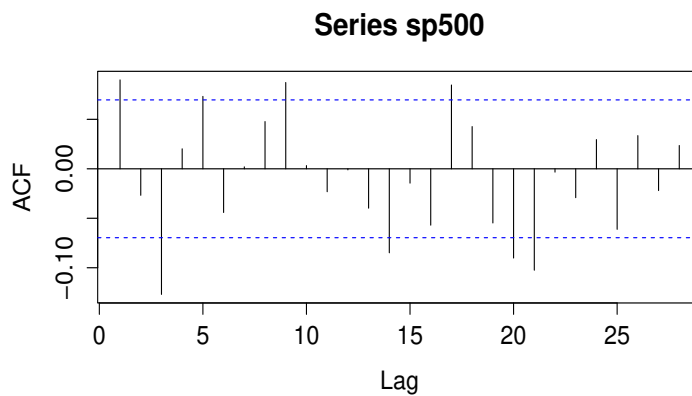


Figure 9.7: Standard and Poor acf for (excess) return series

9.8 Further Extensions of ARCH / GARCH

There are numerous associated models:

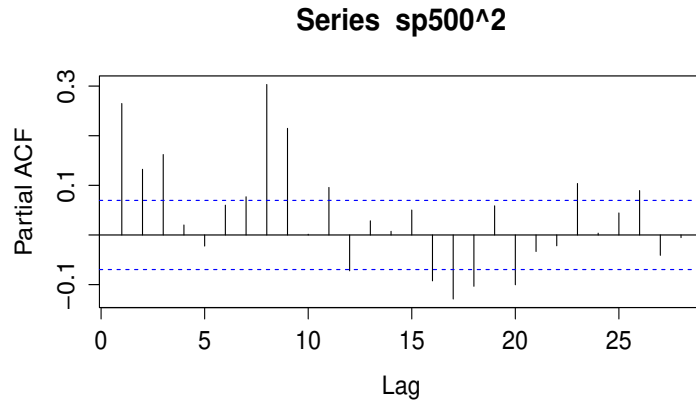


Figure 9.8: Standard and Poor partial autocorrelation for squared excess return series

Non-linear ARCH (NARCH)

$$\sigma_t^\gamma = \alpha_0 + \alpha(B)|X_t|^\gamma + \beta(B)\sigma_t^\gamma.$$

The choice $\gamma = 2$ yields the GARCH model; $\gamma = 1$ is sometimes used.

The standard approach for constructing variations on ARCH / GARCH is as follows: some financial data sets are considered. For each set, the hypothesis $H_0 : \gamma = 2$ (GARCH) is tested against the alternative $H_1 : \gamma \neq 2$ (or rather: the null hypothesis is that the data may be modelled by one of the current models; the alternative is that the wider class of models is more appropriate). The null hypothesis is generally rejected. The creator of the generalisation is then happy and some papers are written. The typical situation is that we have some further parameters to play with in $H_0 \cup H_1$. Since most (in fact all) models are simplifications of reality, a null hypothesis H_0 which is much smaller than $H_0 \cup H_1$ is (almost) always rejected if the data set is large enough.

Asymmetric ARCH Processes One drawback with ARCH, GARCH and NARCH is that positive and negative values have a symmetric effect on the volatility. Many financial time series are strongly asymmetric; negative returns are followed by larger increase in the volatility than positive returns. Typical examples of this may be prices of gasoline or the interest on a loan. Let

$$X_t^+ = \max(X_t, 0) \quad X_t^- = \min(X_t, 0)$$

There are several examples which address this problem; for example, the exponential GARCH (EGARCH) and quadratic GARCH (QGARCH).

The EGARCH is:

$$\ln \sigma_t^2 = \alpha_0 + \beta \ln \sigma_{t-1}^2 + \lambda Z_{t-1} + \phi (|Z_{t-1}| - \mathbb{E}[|Z_{t-1}|]).$$

Other examples may be found in the literature.