ESTIMATION AND AUTOMATIC SELECTION OF ARCH MODELS

Suppose our data was generated by an ARIMA model with ARCH innovations. Two questions arise here: (1) What are the dimensions of the ARIMA and ARCH models? (2) What are the values of the parameters of the ARIMA and ARCH models? We will return to (1) later, but for now let us suppose that we have decided (perhaps incorrectly) on the dimensions.

For the ARIMA \((k\,d\,l)\) model with ARCH \((q)\) innovations, it would be nice to estimate all \(k + l + q + 1\) parameters simultaneously from the data. This is not typically done in practice, however, due to the lack of good computer algorithms for this problem. Instead, we will perform the ARIMA and ARCH modeling and estimation separately, ignoring any possible interactions between the two. So the first step is to estimate the ARIMA model (and to select \(k\) and \(l\)), using traditional ARIMA estimation software. Thus, the ARIMA modeling proceeds exactly as described earlier in this course.

Next, we estimate the ARCH parameters (and select \(q\)) based on the residuals from the fitted ARIMA model. (It is always a good idea to examine the ACF and PACF of the residuals from any ARIMA model under consideration. If they show any strong structure, then the ARIMA model may be an inadequate description of the data.) We will treat these residuals as if they were the actual innovations \(\{\varepsilon_t\}\) of the process.

To estimate an ARCH \((q)\) process, we will use the R software (see separate handout on using R for more details. Always remember that R is case sensitive, so for example if you have a variable \(x\) and you type \(X\) the variable will not be found). The method of estimation used by R is maximum likelihood.

For an example of the use of R, consider the simulated ARCH(2) dataset, which is in the file ARCH2.SIM on the course website. To analyze this data set, you would first copy this file to an appropriate directory (see separate sheet).

To read in this dataset which has no missing values, stored in H:/ARCH2.SIM, the R command is:

```r
> x=scan("H:/ARCH2.SIM")
```

This will create the variable \(x\) in R.
If you want to read in a data set that starts with a single missing value (for example, residuals from an ARIMA model with \( d = 1 \), stored in H:/RES.DAT), the commands would be

\[
> x = \text{scan("H:/RES.DAT",what="list")[-1]}
\]

\[
> x = \text{as.numeric(x)}
\]

Either way, the data set you are reading in must have just one number (or asterisk) per line. Before estimating ARCH models, you must download and load the tseries package (see the document on using R). To then estimate an ARCH(2) model from \( x \), you can use the R command:

\[
> \text{model=garch(x,c(0,2))}
\]

This produces some output, which is not directly useful and should be ignored. To then see the parameter estimates and corresponding \( t \)-statistics and \( p \)-values, use

\[
> \text{summary(model)}
\]

and to see the log likelihood function for this model use

\[
> \text{logLik(model)}
\]

For any particular value of \( q = 1, 2, 3, \ldots \), the command

\[
> \text{model=garch(x,c(0,q))}
\]

would estimate an ARCH(\( q \)) model.

The software provides parameter estimates along with the corresponding standard errors and \( t \)-statistics. Note that \( a_0 \) is the same as our \( \omega \), and \( a_1, \ldots, a_q \) are the ARCH parameter estimates. For the simulated ARCH(2) data set, the estimates were \( \hat{\omega} = .261, \ \hat{\alpha}_1 = .608, \ \text{and} \ \hat{\alpha}_2 = .232 \). These estimates are reasonably close to the true values \( \omega = .25, \ \alpha_1 = .6, \ \alpha_2 = .35 \).

The R garch command assumes that the mean of the series is zero. This is reasonable for our purposes since we will be working with residuals from a fitted ARIMA model. As long as a correct decision was made as to whether to include a constant in the ARIMA model, it is safe to assume that the mean of the residuals is zero. So we will have no need in our ARCH models for any constant that controls the mean. By contrast, the constant \( \omega \) controls the variance, not the mean, and \( \omega \) must always be included in any ARCH model, since otherwise the unconditional variance would be zero.
A few eccentricities of the R ARCH estimation routine should be noted. First, R is in some situations unable to calculate the standard errors, $t$-statistics and $p$-values, and instead returns missing values (NA) for these entries. This can happen for various (technical) reasons, but even if it does happen the log likelihood will still be calculated and can be relied upon. Another problem is that the estimates provided by R do not necessarily correspond to a stationary model. Specifically, R sometimes gives \[ \sum_{i=1}^{q} \hat{\alpha}_i \geq 1. \] If the estimated model is not stationary, it not possible to directly calculate the estimated (unconditional) variance of the series. This should be \( \hat{\omega}(1 - \sum_{j=1}^{q} \hat{\alpha}_j) \), but this quantity will be either infinite or negative if \( \sum_{j=1}^{q} \hat{\alpha}_j \geq 1. \) The best thing to do in this case is to interpret the estimated unconditional variance as being infinite. It is still possible to compute the conditional variance \( h_t \) in this case, and therefore construct one-step forecast intervals. Note that \( h_t \) will be finite, even if the estimated model is nonstationary.

Sometimes, R estimates one of the ARCH parameters to be exactly 1.0. It can also estimate one or more of the ARCH parameters to be zero. These events can be attributed to the constraints placed by R. When they occur, the corresponding standard errors are typically meaningless (e.g., a standard error of zero). We should interpret an ARCH parameter estimate of zero to mean that the given parameter was not significant. The ARCH(1) model with \( \alpha_1 = 1 \) is of considerable practical interest, even though it is not stationary. Under this model, the best forecast of \( \varepsilon_{n+1}^2 \) is just \( \varepsilon_n^2 \) plus the constant \( \omega \), so \( \{\varepsilon_t^2\} \) is a random walk with drift. The model is called an Integrated ARCH, or \( I-ARCH \). Under the \( I-ARCH \) model, changes in the squared series are not forecastable. Still, the model is useful in practice since the width of the one-step forecast intervals will depend on the most recent observation.
Automatic ARCH Model Selection With $AIC_C$

We can compute $AIC_C$ for an $ARCH(q)$ model using the log likelihood output from R. Let $N = n - d$, the effective sample size after differencing. The formula is

$$AIC_C = -2 \log \text{likelihood} + 2(q+1) \frac{N}{N-q-2}.$$ 

There are a few differences between $AIC_C$ for ARCH, given above, and the $AIC_C$ formulas for ARIMA models. Here, we work directly with the log likelihood function instead of the residual sum of squares. Note that you will need to multiply the log likelihood output from R by $-2$. As usual, the selected model is the one which minimizes $AIC_C$.

You should also include $q = 0$ as one of the candidates. This corresponds to the "null hypothesis" that there are no ARCH effects, i.e., that the observations are independent. The log likelihood for this model can be generated by hand for this model by the command

```r
> -0.5*N*(1+log(2*pi*mean(x^2)))
```

where $x$ is the data set under consideration and $N$ is calculated as above. If $AIC_C$ selects $q = 0$, then an ARCH model is not needed, while if it selects a positive value of $q$ then the data shows signs of significant ARCH effects. The garch command will not work for $q=0$. 

Examples

Simulated ARCH(1)

For the simulated ARCH(1) data set (ARCH1.SIM), using candidate values of \( q \) from 0 to 5, \( AIC_C \) selected the correct value, \( q = 1 \).

<table>
<thead>
<tr>
<th>( q )</th>
<th>log likelihood</th>
<th>( AIC_C )</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>-154.648</td>
<td>311.337</td>
</tr>
<tr>
<td>1</td>
<td>-127.606</td>
<td>259.336</td>
</tr>
<tr>
<td>2</td>
<td>-127.698</td>
<td>261.646</td>
</tr>
<tr>
<td>3</td>
<td>-133.900</td>
<td>276.221</td>
</tr>
<tr>
<td>4</td>
<td>-135.667</td>
<td>281.972</td>
</tr>
<tr>
<td>5</td>
<td>-134.410</td>
<td>281.723</td>
</tr>
</tbody>
</table>

This can be compared with the selection of \( q = 1 \) or \( q = 2 \) based on the PACF of the squared series. The ARCH(1) estimates are \( \hat{\omega} = .323 \), \( \hat{\alpha}_1 = .790 \). (The true values were \( \omega = .25 \), \( \alpha_1 = .5 \).)

Simulated ARCH(2)

For the simulated ARCH(2) data set (arch2.sim), \( AIC_C \) again selected the correct value, \( q = 2 \).

<table>
<thead>
<tr>
<th>( q )</th>
<th>log likelihood</th>
<th>( AIC_C )</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>-136.538</td>
<td>275.117</td>
</tr>
<tr>
<td>1</td>
<td>-123.162</td>
<td>250.448</td>
</tr>
<tr>
<td>2</td>
<td>-118.025</td>
<td>242.300</td>
</tr>
<tr>
<td>3</td>
<td>-124.089</td>
<td>256.600</td>
</tr>
<tr>
<td>4</td>
<td>-122.437</td>
<td>255.512</td>
</tr>
<tr>
<td>5</td>
<td>-120.911</td>
<td>254.725</td>
</tr>
</tbody>
</table>

\( AIC_C \) selected the true \( q = 2 \), even though the PACF of the squares was not significant at lag 2.
Bear Stearns Returns

For the Bear Stearns daily returns, 2 Jan to 28 March 2008, the ACF and PACF are insignificant except at lag 4. The $AIC_C$ selects an ARIMA(2,0,4) model, and the ACF and PACF of the residuals indicate that the residuals are white noise. The ACF and PACF of the squared residuals, however, are both significant at lag one. If we fit low-order ARCH models, $AIC_C$ selects an $ARCH(3)$ model.

<table>
<thead>
<tr>
<th>$q$</th>
<th>log likelihood</th>
<th>$AIC_C$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>35.683</td>
<td>-69.297</td>
</tr>
<tr>
<td>1</td>
<td>49.048</td>
<td>-93.885</td>
</tr>
<tr>
<td>2</td>
<td>56.398</td>
<td>-106.368</td>
</tr>
<tr>
<td>3</td>
<td>57.717</td>
<td>-106.706</td>
</tr>
<tr>
<td>4</td>
<td>58.421</td>
<td>-105.731</td>
</tr>
<tr>
<td>5</td>
<td>58.146</td>
<td>-102.708</td>
</tr>
</tbody>
</table>

The $ARCH(3)$ estimates from R are $\hat{\omega} = .00268$, $\hat{\alpha}_1 = .2687$, $\hat{\alpha}_2 = .7420$, $\hat{\alpha}_3 = .0638$. Since the sum of the $ARCH$ parameter estimates exceeds 1, the estimated model may be thought of as a type of I-$ARCH$ model. Using the Minitab ARIMA command, we constructed the best one-step forecast of the returns, $f_{t,1}$. Using the $ARCH(3)$ model for the residuals (so that the residuals are not strict white noise), we have $h_{t+1} = \text{var}[\varepsilon_{t+1} | \psi_t] = .00268 + .2687\varepsilon_t^2 + .7420\varepsilon_{t-1}^2 + .0638\varepsilon_{t-2}^2$, and a 95% one-step forecast interval from the information available at time $t$ is given by $f_{t,1} \pm 1.96 \sqrt{h_{t+1}}$.

For illustrative purposes, we will use the above forecast interval to evaluate the performance of the model during the time just before and after the crash in Bear Stearns on Friday, March 14, 2008, due to the announcement of a 28-day emergency loan to the company from JP Morgan Chase and the Federal Reserve Bank of New York. (On Sunday, March 16’th, 2008, a merger agreement was signed between Bear Stearns and JP Morgan Chase, resulting in even higher volatility on March 17’th.) The use of parameter estimates based on the full data set is not entirely realistic, of course, since the estimated model coefficients would be different if based solely on pre-crash data.
On Thursday, March 13, 2008, (the last trading day before the crash), the return was −0.07437, and the 95% forecast interval based on our fitted ARCH(3) model would be (−.069,.194). The actual return on March 14’th was −.47368, so the interval failed to contain the true value. The ARCH(3) forecast interval constructed on March 14’th would be (−.619,.397), drastically widened by the crash, but still failing to contain the actual March 17’th return, −.83967. However, the forecast interval constructed on March 17’th would be (−1.297,1.052), further drastically widened by the additional volatility, and indeed containing the actual March 18’th return of .22869.

On the other hand, consider the forecast intervals based on the assumption that the residuals are independent, i.e., strict white noise. Using the residual MS for the entire series, .01980 (again, a quantity which would not be available solely from pre-crash data), the one-step forecast intervals would all be $f_{n,1}±1.96\sqrt{.01980}$, of constant width .5516. These intervals miss the true return by a wide margin, for all three dates, March 14’th, 17’th and 18’th. Before the crash, on March 13’th, $h_{t+1}$ is only 23 percent of the residual mean square, but after the crash $h_{t+1}$ increases sharply, and by March 17’th, it becomes 18 times larger than the residual mean square. This illustrates the ability of the ARCH model to adapt to changing volatility conditions.

<table>
<thead>
<tr>
<th>PI Construction Date</th>
<th>ARIMA PI</th>
<th>ARCH PI</th>
<th>Next Day’s Return</th>
<th>$h_{t+1}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>March 13’th</td>
<td>(-.2133,.3383)</td>
<td>(-.069,.194)</td>
<td>-.47368</td>
<td>.00453</td>
</tr>
<tr>
<td>March 14’th</td>
<td>(-.3868,.1648)</td>
<td>(-.619,.397)</td>
<td>-.83967</td>
<td>.0671</td>
</tr>
<tr>
<td>March 17’th</td>
<td>(-.3983,.1533)</td>
<td>(-1.297,1.052)</td>
<td>.22869</td>
<td>.359</td>
</tr>
</tbody>
</table>