Exercises and Solutions: week 2

Financial Econometrics

2024-2025

Paolo Gorgi

CHAPTER 4: Parameter estimation

1. We have a sample of observed data $\{y_1, \ldots, y_5\}$ generated from a GARCH(1,1) model. We want to estimate the model and therefore we need to recover the conditional variance σ_t^2 using the recursion

$$\sigma_t^2 = \omega + \beta_1 \sigma_{t-1}^2 + \alpha_1 y_{t-1}^2, \quad \text{for } t = 2, \ldots, 5.$$

Set the initial condition σ_1^2 equal to the sample variance.

The table below reports the observed sample. Fill in the empty cells for the conditional variance when the above recursion is evaluated at $\theta = (\omega, \beta_1, \alpha_1) = (0.1, 0.8, 0.1)$.

t	1	2	3	4	5
y_t	1.1	-0.6	-1.3	0.1	0.9
σ_t^2					

Do the same for the table below but evaluate the recursion at $\theta = (\omega, \beta_1, \alpha_1) = (0.2, 0.3, 0.1)$.

t	1	2	3	4	5
y_t	1.1	-0.6	-1.3	0.1	0.9
σ_t^2					

Solution:

We set as initialization the sample variance. Since the GARCH model has unconditional mean zero, we can obtain the variance without estimating the mean and setting it to zero.

$$\sigma_1^2 = \frac{1}{5} \sum_{i=1}^{5} y_t^2 = 0.82.$$

Once we have the initial value σ_1^2 we can use the updating equation to recover σ_t^2:

t=2) $\sigma_2^2 = \omega + \beta_1 \sigma_1^2 + \alpha_1 y_1^2 = 0.1 + 0.8 \times 0.82 + 0.1 \times 0.1 \times 1.1^2 = 0.89.$

t=3) $\sigma_3^2 = \omega + \beta_1 \sigma_2^2 + \alpha_1 y_2^2 = \ldots.$

t=4) $\sigma_4^2 = \ldots.$ *and so on.*

The same can be done for the second table.

2. Obtain the updating equation for the derivative sequence $\frac{\partial \sigma_t^2}{\partial \theta}$ for the following models

   (a) $y_t = \sigma_t \epsilon_t, \ \ \sigma_t^2 = \omega + \alpha_1 y_{t-1}^2;$

   (b) $y_t = \sigma_t \epsilon_t, \ \ \sigma_t^2 = \omega + \alpha_1 y_{t-1}^2 + \alpha_2 y_{t-2}^2;$

   (c) $y_t = \sigma_t \epsilon_t, \ \ \sigma_t^2 = \omega + \alpha_1 y_{t-1}^2 + \beta_1 \sigma_{t-1}^2;$

   (d) $y_t = \sigma_t \epsilon_t, \ \ \sigma_t^2 = \omega + \sum_{i=1}^{2} \alpha_i y_{t-i}^2 + \sum_{j=1}^{2} \beta_j \sigma_{t-j}^2;$

Solution:

(a) The parameter vector is given by $\theta = (\omega, \alpha_1)^\top$. Therefore the derivative process $\frac{\partial \sigma_t^2}{\partial \theta}$ is given by

$$\frac{\partial \sigma_t^2}{\partial \theta} = \begin{bmatrix} \frac{\partial \sigma_t^2}{\partial \omega} \\ \frac{\partial \sigma_t^2}{\partial \alpha_1} \end{bmatrix}.$$

We can now calculate the derivatives $\frac{\partial \sigma_t^2}{\partial \omega}$ and $\frac{\partial \sigma_t^2}{\partial \alpha_1}$ as

$$\frac{\partial \sigma_t^2}{\partial \omega} = \frac{\partial(\omega + \alpha_1 y_{t-1}^2)}{\partial \omega} = 1,$$
$$\frac{\partial \sigma_t^2}{\partial \alpha_1} = \frac{\partial(\omega + \alpha_1 y_{t-1}^2)}{\partial \alpha_1} = y_{t-1}^2.$$

Therefore we conclude that

$$\frac{\partial \sigma_t^2}{\partial \theta} = \begin{bmatrix} 1 \\ y_{t-1}^2 \end{bmatrix}.$$

(b) The parameter vector is given by $\theta = (\omega, \alpha_1, \alpha_2)^\top$. Therefore the derivative process $\frac{\partial \sigma_t^2}{\partial \theta}$ is given by

$$\frac{\partial \sigma_t^2}{\partial \theta} = \begin{bmatrix} \frac{\partial \sigma_t^2}{\partial \omega} \\ \frac{\partial \sigma_t^2}{\partial \alpha_1} \\ \frac{\partial \sigma_t^2}{\partial \alpha_2}. \end{bmatrix}$$

We can now calculate the derivatives $\frac{\partial \sigma_t^2}{\partial \omega}$, $\frac{\partial \sigma_t^2}{\partial \alpha_1}$ and $\frac{\partial \sigma_t^2}{\partial \alpha_2}$ as

$$\begin{aligned}
\frac{\partial \sigma_t^2}{\partial \omega} &= \frac{\partial(\omega + \alpha_1 y_{t-1}^2 + \alpha_2 y_{t-2}^2)}{\partial \omega} = 1, \\
\frac{\partial \sigma_t^2}{\partial \alpha_1} &= \frac{\partial(\omega + \alpha_1 y_{t-1}^2 + \alpha_2 y_{t-2}^2)}{\partial \alpha_1} = y_{t-1}^2, \\
\frac{\partial \sigma_t^2}{\partial \alpha_2} &= \frac{\partial(\omega + \alpha_1 y_{t-1}^2 + \alpha_2 y_{t-2}^2)}{\partial \alpha_2} = y_{t-2}^2.
\end{aligned}$$

Therefore we conclude that

$$\frac{\partial \sigma_t^2}{\partial \theta} = \begin{bmatrix} 1 \\ y_{t-1}^2 \\ y_{t-2}^2 \end{bmatrix}.$$

(c) The parameter vector is given by $\theta = (\omega, \alpha_1, \beta_1)^\top$. Therefore the derivative process $\frac{\partial \sigma_t^2}{\partial \theta}$ is given by

$$\frac{\partial \sigma_t^2}{\partial \theta} = \begin{bmatrix} \frac{\partial \sigma_t^2}{\partial \omega} \\ \frac{\partial \sigma_t^2}{\partial \alpha_1} \\ \frac{\partial \sigma_t^2}{\partial \beta_1} \end{bmatrix}.$$

We can now calculate the derivatives $\frac{\partial \sigma_t^2}{\partial \omega}$, $\frac{\partial \sigma_t^2}{\partial \alpha_1}$ and $\frac{\partial \sigma_t^2}{\partial \beta_1}$ as

$$\begin{aligned}
\frac{\partial \sigma_t^2}{\partial \omega} &= \frac{\partial(\omega + \alpha_1 y_{t-1}^2 + \beta_1 \sigma_{t-1}^2)}{\partial \omega} = 1 + \beta_1 \frac{\partial \sigma_{t-1}^2}{\partial \omega}, \\
\frac{\partial \sigma_t^2}{\partial \alpha_1} &= \frac{\partial(\omega + \alpha_1 y_{t-1}^2 + \beta_1 \sigma_{t-1}^2)}{\partial \alpha_1} = y_{t-1}^2 + \beta_1 \frac{\partial \sigma_{t-1}^2}{\partial \alpha_1}, \\
\frac{\partial \sigma_t^2}{\partial \beta_1} &= \frac{\partial(\omega + \alpha_1 y_{t-1}^2 + \beta_1 \sigma_{t-1}^2)}{\partial \beta_1} = \sigma_{t-1}^2 + \beta_1 \frac{\partial \sigma_{t-1}^2}{\partial \beta_1},
\end{aligned}$$

Therefore, we obtain that

$$\frac{\partial \sigma_t^2}{\partial \theta} = \begin{bmatrix} 1 \\ y_{t-1}^2 \\ \sigma_{t-1}^2 \end{bmatrix} + \beta_1 \frac{\partial \sigma_{t-1}^2}{\partial \theta}.$$

(d) The parameter vector is given by $\theta = (\omega, \alpha_1, \alpha_2, \beta_1, \beta_2)^\top$. Therefore the derivative process $\frac{\partial \sigma_t^2}{\partial \theta}$ is given by

$$\frac{\partial \sigma_t^2}{\partial \theta} = \begin{bmatrix} \frac{\partial \sigma_t^2}{\partial \omega} \\ \frac{\partial \sigma_t^2}{\partial \alpha_1} \\ \frac{\partial \sigma_t^2}{\partial \alpha_2} \\ \frac{\partial \sigma_t^2}{\partial \beta_1} \\ \frac{\partial \sigma_t^2}{\partial \beta_2} \end{bmatrix}.$$

We can now calculate the derivatives $\frac{\partial \sigma_t^2}{\partial \omega}$, $\frac{\partial \sigma_t^2}{\partial \alpha_1}$, $\frac{\partial \sigma_t^2}{\partial \alpha_2}$, $\frac{\partial \sigma_t^2}{\partial \beta_1}$ and $\frac{\partial \sigma_t^2}{\partial \beta_2}$ as

$$\begin{aligned}
\frac{\partial \sigma_t^2}{\partial \omega} &= \frac{\partial(\omega + \sum_{i=1}^2 \alpha_i y_{t-i}^2 + \sum_{j=1}^2 \beta_j \sigma_{t-j}^2)}{\partial \omega} = 1 + \beta_1 \frac{\partial \sigma_{t-1}^2}{\partial \omega} + \beta_2 \frac{\partial \sigma_{t-2}^2}{\partial \omega}, \\
\frac{\partial \sigma_t^2}{\partial \alpha_1} &= \frac{\partial(\omega + \sum_{i=1}^2 \alpha_i y_{t-i}^2 + \sum_{j=1}^2 \beta_j \sigma_{t-j}^2)}{\partial \alpha_1} = y_{t-1}^2 + \beta_1 \frac{\partial \sigma_{t-1}^2}{\partial \alpha_1} + \beta_2 \frac{\partial \sigma_{t-2}^2}{\partial \alpha_1}, \\
\frac{\partial \sigma_t^2}{\partial \alpha_2} &= \frac{\partial(\omega + \sum_{i=1}^2 \alpha_i y_{t-i}^2 + \sum_{j=1}^2 \beta_j \sigma_{t-j}^2)}{\partial \alpha_2} = y_{t-2}^2 + \beta_1 \frac{\partial \sigma_{t-1}^2}{\partial \alpha_2} + \beta_2 \frac{\partial \sigma_{t-2}^2}{\partial \alpha_2}. \\
\frac{\partial \sigma_t^2}{\partial \beta_1} &= \frac{\partial(\omega + \sum_{i=1}^2 \alpha_i y_{t-i}^2 + \sum_{j=1}^2 \beta_j \sigma_{t-j}^2)}{\partial \beta_1} = \sigma_{t-1}^2 + \beta_1 \frac{\partial \sigma_{t-1}^2}{\partial \beta_1} + \beta_2 \frac{\partial \sigma_{t-2}^2}{\partial \beta_1}, \\
\frac{\partial \sigma_t^2}{\partial \beta_2} &= \frac{\partial(\omega + \sum_{i=1}^2 \alpha_i y_{t-i}^2 + \sum_{j=1}^2 \beta_j \sigma_{t-j}^2)}{\partial \beta_2} = \sigma_{t-2}^2 + \beta_1 \frac{\partial \sigma_{t-1}^2}{\partial \beta_2} + \beta_2 \frac{\partial \sigma_{t-2}^2}{\partial \beta_2}.
\end{aligned}$$

Therefore we conclude that

$$\frac{\partial \sigma_t^2}{\partial \theta} = \begin{bmatrix} 1 \\ y_{t-1}^2 \\ y_{t-2}^2 \\ \sigma_{t-1}^2 \\ \sigma_{t-2}^2 \end{bmatrix} + \beta_1 \frac{\partial \sigma_{t-1}^2}{\partial \theta} + \beta_2 \frac{\partial \sigma_{t-2}^2}{\partial \theta}.$$

3. Write a Newton-Raphson algorithm that finds the maximum of the function $-(x-1)^4/4$ starting from the point $x^{(1)} = 3$. Find $x^{(3)}$.

   Solution:
   First, we notice that $f'(x) = -(x-1)^3$ and $-3(x-1)^2$. Therefore,

   $$x^{(k+1)} = x^{(k)} - (x^{(k)} - 1)/3.$$

   Given that $x^{(1)} = 3$, we have

   $$x^{(2)} = x^{(1)} - (x^{(1)} - 1)/3 = 2.33,$$

   and

   $$x^{(3)} = x^{(2)} - (x^{(2)} - 1)/3 = 1.89.$$

4. Write the log-likelihood function for the following ARCH model with non-zero constant conditional mean μ

   $$y_t = \mu + \sigma_t \varepsilon_t, \quad \sigma_t^2 = \omega + \alpha_2 (y_{t-2} - \mu)^2.$$

   The sample size is 500. Define all quantities including the parameter vector θ and the initial conditions for σ_t^2 (if needed).

   Solution:

   First we note that the parameter vector is $\theta = (\mu, \omega, \alpha_2)^\top$. The conditional distribution of y_t given the past $Y^{t-1} = \{y_{t-1}, y_{t-2}, \dots\}$ is $y_t | Y^{t-1} \sim N(\mu, \sigma_t^2)$. Therefore the conditional density function $p(y_t|Y^{t-1})$ is given by

   $$p(y_t|Y^{t-1}) = \frac{1}{\sqrt{2\pi\sigma_t^2}} \exp\left(-\frac{(y_t - \mu)^2}{2\sigma_t^2}\right),$$

   and thus the log-density is

   $$\log p(y_t|Y^{t-1}) = -\frac{1}{2}\log(2\pi) - \frac{1}{2}\log(\sigma_t^2) - \frac{1}{2}\frac{(y_t - \mu)^2}{\sigma_t^2}.$$

   The term $\frac{1}{2}\log(2\pi)$ is just a constant (it does not depend on parameters) and thus it can be dropped. The log-likelihood function $L(y_1, \dots, y_T; \theta)$ is given by the sum of the conditional log-densities

   $$L(y_1, \dots, y_{500}; \theta) = \sum_{t=3}^{500} \left(-\frac{1}{2}\log(\sigma_t^2) - \frac{1}{2}\frac{(y_t - \mu)^2}{\sigma_t^2}\right),$$

   where $\sigma_t^2 = \omega + \alpha_2 (y_{t-2} - \mu)^2$. Note that the sum in the above expression starts from 3 because the first observation available is y_1 and σ_t^2 depends on y_{t-2}.

5. Write the log-likelihood function for the following GARCH model with non-zero dynamic conditional mean $\mathbb{E}(y_t|Y^{t-1}) = \phi y_{t-1}$

$$y_t = \phi y_{t-1} + \sigma_t \varepsilon_t, \quad \sigma_t^2 = \omega + \alpha_1 (y_{t-1} - \phi y_{t-2})^2 + \beta_1 \sigma_{t-1}^2.$$

The sample size is 500. Define all quantities including the parameter vector θ and the initial conditions for σ_t^2 (if needed).

Solution:

First we note that the parameter vector is $\theta = (\phi, \omega, \alpha_1, \beta_1)^\top$. The conditional distribution of y_t given the past $Y^{t-1} = \{y_{t-1}, y_{t-2}, \dots\}$ is $y_t|Y^{t-1} \sim N(\phi y_{t-1}, \sigma_t^2)$. Therefore the conditional density function $p(y_t|Y^{t-1})$ is given by

$$p(y_t|Y^{t-1}) = \frac{1}{\sqrt{2\pi\sigma_t^2}} \exp\left(-\frac{(y_t - \phi y_{t-1})^2}{2\sigma_t^2}\right),$$

and thus the log-density is

$$\log p(y_t|Y^{t-1}) = -\frac{1}{2}\log(2\pi) - \frac{1}{2}\log(\sigma_t^2) - \frac{1}{2}\frac{(y_t - \phi y_{t-1})^2}{\sigma_t^2}.$$

The term $\frac{1}{2}\log(2\pi)$ is just a constant (it does not depend on parameters) and thus it can be dropped. The log-likelihood function $L(y_1, \dots, y_T; \theta)$ is given by the sum of the conditional log-densities

$$L(y_1, \dots, y_{500}; \theta) = \sum_{t=3}^{500} \left(-\frac{1}{2}\log(\sigma_t^2) - \frac{1}{2}\frac{(y_t - \phi y_{t-1})^2}{\sigma_t^2}\right),$$

where $\sigma_t^2 = \omega + \alpha_1(y_{t-1} - \phi y_{t-2})^2 + \beta_1 \sigma_{t-1}^2$. The recursion for σ_t^2 can be initialized setting σ_1^2 and σ_2^2 equal to the sample variance. Note that the sum in the above expression starts from 3 because the first observation available is y_1 and σ_t^2 depends on both y_{t-1} and y_{t-2}.

6. Consider following output for the estimation of a GARCH(1,1) model from a sample of data of size $T = 1000$.

$$\hat{\theta}_T = \begin{bmatrix} \hat{\omega} \\ \hat{\alpha}_1 \\ \hat{\beta}_1 \end{bmatrix} = \begin{bmatrix} 0.25 \\ 0.32 \\ 0.64 \end{bmatrix}, \quad \hat{\Omega} = \begin{bmatrix} 1.32 & 0.81 & 0.22 \\ 0.81 & 1.75 & 0.30 \\ 0.30 & 0.22 & 0.98 \end{bmatrix},$$

where $\hat{\Omega}$ is an estimate of the asymptotic covariance matrix of $\sqrt{T}(\hat{\theta}_T - \theta_0)$. Find 95% confidence intervals for the true ω, α_1 and β_1.

Solution:
From, the asymptotic distribution of the ML estimator, we know that for large T,

$$\hat{\theta}_T \overset{app}{\sim} N\big(\theta_0, \hat{\Omega}/T\big).$$

Therefore, $\hat{\omega} \overset{app}{\sim} N\big(\omega, \hat{\Omega}_{11}/T\big)$, $\hat{\alpha} \overset{app}{\sim} N\big(\alpha_1, \hat{\Omega}_{22}/T\big)$ and $\hat{\beta} \overset{app}{\sim} N\big(\beta_1, \hat{\Omega}_{33}/T\big)$. Therefore, we can construct 95% level confidence intervals as follows

$$CI_{95\%}(\omega) = \hat{\omega} \pm 1.96\sqrt{\hat{\Omega}_{11}/T} = 0.25 \pm 1.96\sqrt{1.32/1000} = (0.18, 0.32).$$

$$CI_{95\%}(\alpha_1) = \hat{\alpha}_1 \pm 1.96\sqrt{\hat{\Omega}_{22}/T} = 0.32 \pm 1.96\sqrt{1.75/1000} = (0.24, 0.40).$$

$$CI_{95\%}(\beta_1) = \hat{\beta}_1 \pm 1.96\sqrt{\hat{\Omega}_{33}/T} = 0.64 \pm 1.96\sqrt{0.98/1000} = (0.58, 0.70).$$

7. Use the estimation results of the previous question to calculate the p-value for the test $H_0 : \alpha_1 \geq 0.45$ vs $H_1 : \alpha_1 < 0.45$.

Solution:
From the estimation results of the previous exercise we have the following approximate distribution $\hat{\alpha}_1 \overset{app}{\sim} N\left(\alpha_1, \hat{\Omega}_{22}/T\right)$, where $\hat{\Omega}_{22}$ is the element in position $(2, 2)$ of the covariance matrix $\hat{\Omega}$ in the previous exercise. Therefore, under the null hypothesis H_0 we have that

$$z = \sqrt{T}\frac{\hat{\alpha}_1 - 0.45}{\sqrt{\hat{\Omega}_{2,2}}} \overset{app}{\sim} N(0, 1).$$

Using the estimation results of the previous exercise, we have that the observed value of z is given by

$$z_{obs} = \sqrt{1000}\frac{0.32 - 0.45}{\sqrt{1.75}} = -3.11.$$

As a result, since we have left-tailed test, the p-value is given by

$$\text{p-value} = P(z \leq z_{calc}) = \Phi(z_{calc}) = \Phi(-3.11) = 0.001,$$

where $\Phi(\cdot)$ is the cdf of a standard normal.

CHAPTER 5: Econometric analysis with GARCH models

1. Plot the news impact curve for the following ARCH models

$$\text{(i)} \quad y_t = \sigma_t \varepsilon_t, \quad \sigma_t^2 = 0.1 + 0.1 y_{t-1}^2$$
$$\text{(ii)} \quad y_t = \sigma_t \varepsilon_t, \quad \sigma_t^2 = 0.1 + 0.3 y_{t-1}^2.$$

Comment on the following statements:

(a) The impact of the log-return y_{t-1} on σ_t^2 is bigger for model (ii) than for model (i).

(b) For both models (i) and (ii), negative log-returns have smaller impact on the conditional variance than positive log-returns.

Solution:

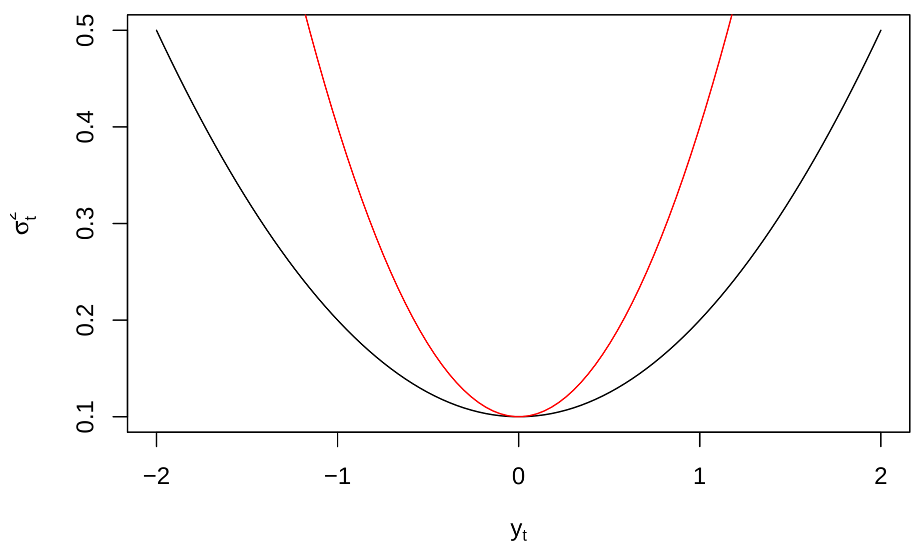

Figure 1: The black line shows the impact curve for model (i) whereas the red line shows the impact curve for model (ii).

(a) The statement is correct because the coefficient α_1 of the ARCH(1) model in (ii) is bigger than the coefficient α_1 of the ARCH(1) model in (i). Therefore, the observed value y_{t-1} has a bigger impact on σ_t^2 for model (ii) than for model (i).

(b) The statement is incorrect. This because the impact of y_{t-1} on σ_t^2 depends only on the absolute value of y_{t-1} (in the expression of σ_t^2 the last observation y_{t-1} is squared). In fact observing for instance $y_{t-1} = 1$ or $y_{t-1} = -1$ leads exactly to the same conditional variance σ_t^2.

2. Which model would you select according to the following table? Explain why.

Model	AIC	BIC	log-lik
ARCH(1)	3668.2	3678.0	-1832.1
ARCH(2)	3651.2	3665.9	-1822.6
ARCH(3)	3636.0	3655.6	-1814.0
GARCH(1,1)	3596.2	3610.9	-1795.1
GARCH(2,1)	3597.0	3616.6	-1794.5
GARCH(1,2)	3595.6	3615.2	-1793.8

Solution:
As we can see the model with the lowest AIC is the GARCH(1,2) and, instead, the model with the lowest BIC is the GARCH(1,1). Therefore, we can select either the GARCH(1,1) or the GARCH(1,2). If we want a more parsimonious model we could opt for the GARCH(1,1) as it has less parameters. The log-likelihood should not be used as a selection criterion.

3. The table below reports AIC, BIC and log-likelihood values obtained from the estimation of some ARCH/-GARCH models. Each case refers to a different dataset. Find the mistake in each of the 3 cases and explain why.

	Model	AIC	BIC	log-lik
Case 1	ARCH(1)	2475.2	2485.0	-1235.6
	GARCH(1,1)	2486.4	2501.1	-1240.2
Case 2	ARCH(2)	2484.4	2494.2	-1240.2
	GARCH(1,1)	2486.4	2501.1	-1240.2
Case 3	GARCH(1,1)	2679.1	2664.4	-1329.2
	GARCH(1,2)	2682.6	2663.0	-1327.5

Solution:
Case 1) The ARCH(1) is nested in the GARCH(1,1) (GARCH(1,1) with $\beta_1 = 0$ is ARCH(1)). Therefore, the GARCH(1,1) cannot have a lower likelihood than ARCH(1,1).
Case 2) ARCH(2) and GARCH(1,1) have the same number of parameters, which is 3. Therefore, if they have the same value for the log-likelihood they also need to have the same values of AIC and BIC.
Case 3) The BIC suggests that the better model is the GARCH(1,2) and the AIC suggests that the better model is the $GARCH(1,1)$. This cannot be true because the BIC penalizes additional parameters more than the AIC ($\log(T)$ vs 2 as penalty). Therefore, since GARCH(1,1) has less parameters, if AIC is smaller also the BIC should be smaller than those of the GARCH(1,2) (or the other way around).

4. Consider the GARCH model with normal error distribution

$$y_t = \sigma_t \varepsilon_t, \qquad \sigma_t^2 = 0.1 + 0.3 y_{t-1}^2 + 0.7 \sigma_{t-1}^2 \ .$$

The table below reports the observed y_T and σ_T^2. Fill in the table by calculating the α-VaR at time $T+1$.

Solution:
The α-VaR$_{T+1}$ is defined as the value such that

$$P\left(y_{T+1} \leq \alpha\text{-VaR}_{T+1} | Y^T\right) = \alpha.$$

level α	y_T	σ_T^2	VaR_{T+1}
0.05	-0.8	1.0	
0.01	-0.8	1.0	
0.05	1.3	1.0	
0.01	1.3	1.0	

Next, we can write

$$P\left(\frac{y_{T+1}}{\sigma_{T+1}} \leq \frac{\alpha\text{-VaR}_{T+1}}{\sigma_{T+1}}|Y^T\right) = \alpha.$$

Therefore, since $\frac{y_{T+1}}{\sigma_{T+1}}|Y^T \sim N(0,1)$, we have that

$$\Phi\left(\frac{\alpha\text{-VaR}_{T+1}}{\sigma_{T+1}}\right) = \alpha.$$

Inverting, the cdf of the standard normal we get

$$\frac{\alpha\text{-VaR}_{T+1}}{\sigma_{T+1}} = \Phi^{-1}(\alpha) = z_\alpha,$$

which implies

$$\alpha\text{-VaR}_{T+1} = \sigma_{T+1} z_\alpha.$$

We can now use this expression to obtain the desired VaR.
Case $\alpha = 0.05$, $y_T = -0.8$, $\sigma_T^2 = 1.0$) We first obtain σ_{T+1}^2,

$$\sigma_{T+1}^2 = 0.1 + 0.3y_T^2 + 0.7\sigma_T^2 = 0.1 + 0.3 \times (-0.8)^2 + 0.7 \times 1.0 = 0.99.$$

Therefore,

$$\alpha\text{-VaR}_{T+1} = \sigma_{T+1} z_{0.05} = \sqrt{0.99} \times (-1.64) = -1.62.$$

The rest of the table can be easily filled-in following the same steps.

5. Suppose that the daily percentage return y_t on the stock of Microsoft is well described by the following ARCH(1) model

$$y_t = \sigma_t \varepsilon_t, \qquad \sigma_t^2 = 0.1 + 0.52y_{t-1}^2\ .$$

Suppose that the stock price of Microsoft just went down from \$64.35 to \$63.1. What is the probability the stock price will fall below \$61 during the next trading day? What is the probability that it will rise above \$65?

Solution:
The last observed percentage return is given by

$$Y_T = \frac{p_T - p_{T-1}}{p_{T-1}} \times 100 = \frac{63.1 - 64.35}{64.35} \times 100 = -2.$$

Therefore, the conditional variance is

$$\sigma_{T+1}^2 = 0.1 + 0.5Y_T^2 = 0.1 + 0.5 \times (-2)^2 = 2.18.$$

a) We compute the conditional probability that the price will fall below 61 as follows

$$\begin{aligned}
P(p_{T+1} \leq 61|Y^T) &= P\left(\frac{p_{T+1} - p_T}{p_T} \times 100 \leq \frac{61 - p_T}{p_T} \times 100|Y^T\right) \\
&= P\left(y_{T+1} \leq \frac{61 - 63.1}{63.1} \times 100|Y^T\right) \\
&= P\left(y_{T+1} \leq -3.33|Y^T\right) \\
&= P\left(\frac{y_{T+1}}{\sigma_{T+1}} \leq \frac{-3.33}{\sigma_{T+1}}|Y^T\right) \\
&= \Phi\left(\frac{-3.33}{\sigma_{T+1}}\right) = \Phi\left(\frac{-3.33}{\sqrt{2.18}}\right) = \Phi(-2.26).
\end{aligned}$$

b) Similarly, the probability that the price will rise above 65 is obtained as follows

$$\begin{aligned}
P(p_{T+1} > 65|Y^T) &= P\left(y_{T+1} > \frac{65-63.1}{63.1} \times 100|Y^T\right) \\
&= P\left(y_{T+1} > 3.01|Y^T\right) \\
&= P\left(\frac{y_{T+1}}{\sigma_{T+1}} > \frac{3.01}{\sigma_{T+1}}|Y^T\right) \\
&= 1 - \Phi\left(\frac{3.01}{\sigma_{T+1}}\right) = 1 - \Phi(2.04).
\end{aligned}$$

6. Assume that $\{y_t\}_{t\in\mathbb{Z}}$ follows the following ARCH model

$$y_t = \sigma_t \varepsilon_t, \qquad \sigma_t^2 = 0.5 + 0.45 y_{t-1}^2 + 0.2 y_{t-3}^2 \ .$$

We are at time T and we have observed $y_T = -1.2$, $y_{T-1} = 1.95$ and $y_{T-2} = 0.9$. Find the forecast for the variance 1 and 2 steps ahead, namely $\sigma_T^2(h) = \mathbb{V}ar(y_{T+h}|Y^T)$, $h = 1, 2$. Find a good approximation for the forecast of the conditional variance 25 steps ahead $\sigma_T^2(25)$.

Solution:

First we obtain that σ_{T+1}^2 is given by

$$\sigma_{T+1}^2 = 0.5 + 0.45 y_T^2 + 0.2 y_{T-2}^2 = 0.5 + 0.45 \times (-1.2)^2 + 0.2 \times (0.9)^2 = 1.31.$$

We now find that $\sigma_T^2(1)$ is given by

$$\sigma_T^2(1) = \mathbb{V}ar(y_{T+1}|Y^T) = \mathbb{E}(y_{T+1}^2|Y^T) = \mathbb{E}(\sigma_{T+1}^2\varepsilon_{T+1}^2|Y^T) = \mathbb{E}(\sigma_{T+1}^2|Y^T)\mathbb{E}(\varepsilon_{T+1}^2|Y^T) = \sigma_{T+1}^2 = 1.31.$$

Instead we obtain that $\sigma_T^2(2)$ is given by

$$\begin{aligned}
\sigma_T^2(2) &= \mathbb{V}ar(y_{T+2}|Y^T) = \mathbb{E}(y_{T+2}^2|Y^T) = \mathbb{E}(\sigma_{T+2}^2\varepsilon_{T+2}^2|Y^T) = \mathbb{E}(\sigma_{T+2}^2|Y^T)\mathbb{E}(\varepsilon_{T+2}^2|Y^T) = \mathbb{E}(\sigma_{T+2}^2|Y^T) \\
&= \mathbb{E}(0.5 + 0.45 y_{T+1}^2 + 0.2 y_{T-1}^2|Y^T) = 0.5 + 0.45\mathbb{E}(y_{T+1}^2|Y^T) + 0.2\mathbb{E}(y_{T-1}^2|Y^T) \\
&= 0.5 + 0.45\sigma_{T+1}^2 + 0.2 y_{T-1}^2 = 0.5 + 0.45 \times 1.31 + 0.2 \times (1.95)^2 = 1.85.
\end{aligned}$$

Finally, we need to find an approximation to $\sigma_T^2(25)$. We know that $\sigma_T^2(h) \to \mathbb{V}ar(y_t)$ at an exponential rate as $h \to \infty$. Therefore, we can use the approximation $\sigma_T^2(25) \approx \mathbb{V}ar(y_t)$. Indeed the model in question is an ARCH(3) with parameters $\omega = 0.5$, $\alpha_1 = 0.45$, $\alpha_2 = 0$ and $\alpha_3 = 0.2$. As we know the unconditional variance of an ARCH(3) is given by $\mathbb{V}ar(y_t) = \omega/(1 - \alpha_1 - \alpha_2 - \alpha_3)$. Therefore, we conclude that

$$\sigma_T^2(25) \approx \frac{0.5}{(1 - 0.45 - 0.2)} = 1.43.$$

7. Assume that $\{y_t\}_{t\in\mathbb{Z}}$ follows the following GARCH(1,1) model

$$y_t = \sigma_t \varepsilon_t, \qquad \sigma_t^2 = 0.5 + 0.7\sigma_{t-1}^2 + 0.2 y_{t-1}^2 \ .$$

We are at time T and we have observed $y_T = -1.2$, $\sigma_T^2 = 0.8$. Find the forecast for the variance 1 and 2 steps ahead, namely $\sigma_T^2(h) = \mathbb{V}ar(y_{T+h}|Y^T)$, $h = 1, 2$. Find a good approximation for the forecast of the conditional variance 25 steps ahead $\sigma_T^2(25)$.

Solution:

First we obtain that σ_{T+1}^2 is given by

$$\sigma_{T+1}^2 = 0.5 + 0.7\sigma_T^2 + 0.2 y_T^2 = 0.5 + 0.7 \times 0.8 + 0.2 \times (-1.2)^2 = 1.35$$

We now find that $\sigma_T^2(1)$ is given by

$$\sigma_T^2(1) = \mathbb{V}ar(y_{T+1}|Y^T) = \mathbb{E}(y_{T+1}^2|Y^T) = \mathbb{E}(\sigma_{T+1}^2\varepsilon_{T+1}^2|Y^T) = \mathbb{E}(\sigma_{T+1}^2|Y^T)\mathbb{E}(\varepsilon_{T+1}^2|Y^T) = \sigma_{T+1}^2 = 1.35.$$

Instead we obtain that $\sigma_T^2(2)$ is given by

$$\begin{aligned}\sigma_T^2(2) &= \mathbb{V}ar(y_{T+2}|Y^T) = \mathbb{E}(y_{T+2}^2|Y^T) = \mathbb{E}(\sigma_{T+2}^2\varepsilon_{T+2}^2|Y^T) = \mathbb{E}(\sigma_{T+2}^2|Y^T)\mathbb{E}(\varepsilon_{T+2}^2|Y^T) = \mathbb{E}(\sigma_{T+2}^2|Y^T)\\ &= \mathbb{E}(0.5 + 0.7\sigma_{T+1}^2 + 0.2y_{T+1}^2|Y^T) = 0.5 + 0.7\mathbb{E}(\sigma_{T+1}^2|Y^T) + 0.2\mathbb{E}(y_{T+1}^2|Y^T)\\ &= 0.5 + 0.7\sigma_{T+1}^2 + 0.2y_{T+1}^2 = 0.5 + 0.9\sigma_{T+1}^2 = 0.5 + 0.9 \times 1.35 = 1.71.\end{aligned}$$

Finally, we need to find an approximation to $\sigma_T^2(25)$. We know that $\sigma_T^2(h) \to \mathbb{V}ar(y_t)$ at an exponential rate as $h \to \infty$. Therefore, we can use the approximation $\sigma_T^2(25) \approx \mathbb{V}ar(y_t)$. Indeed the model in question is an GARCH(1,1) with parameters $\omega = 0.5$, $\beta_1 = 0.7$ and $\alpha_1 = 0.2$. As we know the unconditional variance of an GARCH(1,1) is given by $\mathbb{V}ar(y_t) = \omega/(1 - \beta_1 - \alpha_1)$. Therefore, we conclude that

$$\sigma_T^2(25) \approx \frac{0.5}{(1 - 0.7 - 0.2)} = 5.$$

8. Consider the following GARCH models

$$\begin{aligned}&\text{(i)} \quad y_t = \sigma_t\varepsilon_t, \quad \sigma_t^2 = 0.5 + 0.8\sigma_{t-1}^2 + 0.1y_{t-1}^2\\ &\text{(ii)} \quad y_t = \sigma_t\varepsilon_t, \quad \sigma_t^2 = 0.5 + 0.1\sigma_{t-1}^2 + 0.1y_{t-1}^2\end{aligned}$$

Comment on the following statement: the approximation to the conditional variance $\sigma_T^2(25)$ as obtained in the previous should be more accurate for model (ii) than for model (i).

Solution:

The statement is correct. For a GARCH(1,1) model we have that

$$\sigma_T^2(h) = \omega + (\alpha_1 + \beta_1)\sigma_T^2(h-1),$$

and thus

$$\sigma_T^2(h) = \omega\sum_{i=0}^{h-1}(\beta_1 + \alpha_1)^i + (\beta_1 + \alpha_1)^h\sigma_{T+1}^2.$$

As a result, when $\beta_1 + \alpha_1$ is small the term $(\beta_1 + \alpha_1)^{25}$ is closer to zero than when $\beta_1 + \alpha_1$ is large. Similarly, when $\beta_1 + \alpha_1$ is small the term $\omega\sum_{i=0}^{24}(\alpha_1 + \beta_1)^i$ is closer to $\omega/(1 - \beta_1 - \alpha_1)$ than when $\beta_1 + \alpha_1$ is large. In our case, we have that $\beta_1 + \alpha_1$ is smaller for model (ii) than for model (i). Therefore, for model (ii), the unconditional variance $\omega/(1 - \beta_1 - \alpha_1)$ can be considered a better approximation to $\sigma_T^2(25)$.

An intuitive explanation why the statement is correct is that model (ii) has a shorter memory (dependence on the past) than model (i). Therefore, the past has less predictive power for model (ii) and thus the prediction converges more quickly to the unconditional variance.