TSA 1

Zhenning Zhang
Winston Bartle
Jamie Mo

April 19, 2022

Part 1

Question a

$y_t = \phi' x_t + \varepsilon_t$. The OLS goal is to find the best ϕ such that it minimizes $\sum_{t=1}^{T} \varepsilon_t^2 = \sum_{t=1}^{T} \left(y_t - \phi' x_t\right)^2$. taking the derivatives w.r.t ϕ':

$$
\begin{aligned}
0 &= -2 \sum_{t=1}^{T} \left(y_t - \phi' x_t\right) \left(-x_t'\right) \\
0 &= \sum_{t=1}^{T} y_t x_t' - \sum_{t=1}^{T} \phi' x_t x_t' \\
\hat{\phi_{OLS}}' &= \left(\sum_{t=1}^{T} x_t x_t'\right)^{-1} \left(\sum_{t=1}^{T} y_t x_t'\right) \\
\hat{\phi_{OLS}} &= \left(\sum_{t=1}^{T} x_t x_t'\right)^{-1} \left(\sum_{t=1}^{T} x_t y_t\right)
\end{aligned}
$$

Question b

$\varepsilon = y - X\phi$. The goal is to find ϕ such that it minimizes $\varepsilon' \varepsilon = y' y - y' X\phi - \phi' X\phi - \phi' X' y + \phi' X' X\phi$:

$$
\begin{aligned}
0 &= \frac{\partial y^{'} y}{\partial \phi} - 2\frac{\partial y^{'} X\phi}{\partial \phi} + \frac{\partial \phi^{'} X^{'} X\phi}{\partial \phi 1} \\
&= 0 - 2\left(y^{'} X\right)^{'} + 2\frac{\partial \phi^{'} X^{'}}{\partial \phi} X\phi \\
&= -2\left(y^{'} X\right)^{'} + 2X^{'} X\phi \\
\hat{\phi_{OLS}} &= \left(X^{'} X\right)^{-1} X^{'} y \\
&= \phi + \left(X^{'} X\right)^{-1} X^{'} \varepsilon \\
E\left[\hat{\phi_{OLS}}|X\right] &= E\left[\phi + \left(X^{'} X\right)^{-1} X^{'} \varepsilon|X\right] \\
&= \phi + \left(X^{'} X\right)^{-1} X^{'} E\left[\varepsilon|X\right] \\
&= \phi
\end{aligned}
$$

Hence unbiased. Here the assumption is that $E[\varepsilon|X] = 0$

Question c

The likelihood is a joint probability.

$$L(\theta) = P(y_1, \cdots, y_T|\theta)$$

The recursive definition of the process gives us the transition distribution directly , therefore We are simply going to input previous observations in order to decompose the joint distribution.

$$
\begin{aligned}
L(\theta) &= P(y_1, \cdots, y_T|\theta) \\
&= P(y_T|y_1, \cdots, y_{T-1}; \theta) \times P(y_1, \cdots, y_T|\theta) \\
&= P(y_T|y_1, \cdots, y_{T-1}; \theta) \times P(y_{T-1}|y_1, \cdots, y_{T-2}; \theta) \times P(y_1, \cdots, y_{T-2}|\theta) \\
&= \cdots \\
&= P(y_T|y_{T_1}, \cdots, y_1) \cdots P(y_2|y_1) \times P(y_1)
\end{aligned}
$$

Question d

From the previous sub question:

$$f(y) = P(y_T|y_{T_1}, \cdots, y_1) \cdots P(y_2|y_1) \times P(y_1)$$

MLE estimator is the solution of the equation $\frac{dL(\theta)}{d\theta} = 0$ with $\varepsilon_t \sim i.i.d.\ N(0, \sigma^2)$
The density of the p observations has mean vector having each element value as

$$\mu = \frac{c}{1 - \phi_1 - \phi_2 - \cdots - \phi_p}$$

The variance-covariance matrix will be

$$\sigma^2 \times V_p = \begin{pmatrix} \gamma_0 & \gamma_1 & \cdots & \gamma_{p-1} \\ \gamma_1 & \gamma_0 & \cdots & \gamma_{p-2} \\ \gamma_2 & \gamma_1 & \cdots & \gamma_{p-2} \\ \cdots & & & \\ \gamma_{p-1} & \gamma_{p-2} & \cdots & \gamma_1 \end{pmatrix}$$

Then we get the distribution of p observations as

$$f_{y_t|y_{t_1},\cdots,y_1}(y_t|y_{t_1},\cdots,y_1|\theta) = \frac{1}{\sqrt{2\pi\sigma^2}} \times \exp\left[\frac{[-(y_t - c - \phi_1 \times y_{t-1} - \cdots - \phi_p \times y_{t-p})^2]}{2\sigma^2}\right]$$

$$\begin{aligned} L(\theta) = \frac{-p}{2} \times log(2\pi) - \frac{p}{2} \times log(\sigma^2) - \frac{1}{2\sigma^2}(y_p - \mu_p)' \times V_P^-1 \times (y_p - \mu_P) - \frac{T-p}{2} \times log(2\pi) \\ -\frac{T-p}{2} \times log(\sigma^2) + 1/2log(|V_p^-1| - \sum_{t=p+1}^{T} (\frac{[-(y_t - c - \phi_1 \times y_{t-1} - \cdots - \phi_p \times y_{t-p})^2]}{2\sigma^2}) \end{aligned}$$

Question e

In FGLS we assume that:

$$\begin{aligned} Var(y_t) &= \sigma^2\Omega \\ &= \frac{\sigma^2}{1 - \phi_1^2 - \phi_2^2 - \cdots - \phi_p^2} \end{aligned}$$

As a result, we can write Ω^{-1} as:

$$\hat{\Omega}^{-1} = (1 - \hat{\phi}_1^2 - \hat{\phi}_2^2 - \cdots - \hat{\phi}_p^2)$$

And the corresponding FGLS is:

$$\hat{\beta}_{FGLS} = (z_t\hat{\Omega}z_t')^{-1}z_t\hat{\Omega}^{-1}r_t$$

Question f

y_t stationarity means $|z| > 1 \Rightarrow |\alpha| < 1$ and $\alpha = 0$

$$\begin{aligned} &y_t = \phi_0 - 1.3y_{t-2} - 0.4y_{t-4} + \epsilon_t \\ &y_t + 1.3y_{t-2} + 0.4y_{t-4} = \epsilon_t \\ &(1 + 1.3 \times L^2 + 0.4 \times L^4) \times y_t = \epsilon_t \end{aligned}$$

The characteristic equation

$$\theta(z) = 1 + 1.3 \times Z^2 + 0.4 \times Z^4 = 0$$

has characteristic roots z_1, z_2, z_3, z_4.

We use substitution method to get the roots.

Assume $z^2 = k$

$$\begin{aligned}\theta(k) &= 1 + 1.3 \times k + 0.4 \times k^2 = 0 \\ k &= \frac{-1.3 \pm \sqrt{(1.69 - 1.6)}}{0.8} \\ k = z^2 &= \frac{-1}{0.8}, -\frac{-1.6}{0.8}\end{aligned}$$

The values we got here are $\pm\sqrt{\frac{-5}{4}}, \pm\sqrt{-2}$
The values lies outside the unit circle. Then the process y_t is stationary and there is no unit root .

Question g

$AR(1) : y_t = \phi y_{t-1} + \varepsilon_t$. Assuming stationarity ρ_k is: $\frac{\gamma_k}{\gamma_0} = \frac{\phi\gamma_{k-1}}{\gamma_0} = \frac{\phi^k\gamma_0}{\gamma_0} = \phi^k$ therefore is follows that if the ACF decays exponentially, it means that $|\phi^k|$ is very small implying $|\phi| < 1$ which is the condition for stationarity of an AR(1) model.

Question h

We can compute the general expression:

$$\begin{aligned}E(y_t y_{t-k}) &= 0.6E(y_{t-1}y_{t-k}) + 0.25E(y_{t-2}y_{t-k}) \\ \rho_k &= 0.6\rho_{k-1} + 0.25\rho_{k-2}\end{aligned}$$

Set $k = 1$, can easily write:

$$\gamma_1 = 0.6\gamma_0 - 0.25\gamma_1$$

Knowing that $\rho_0 = 1$:

$$\begin{aligned}\rho_1 = \frac{\gamma_1}{\gamma_0} &= 0.6 - 0.25\rho_1 \\ &= 0.48\end{aligned}$$

Write the evolution dynamics in matrix form:

$$\begin{pmatrix}\rho_k \\ \rho_{k-1}\end{pmatrix} = \begin{pmatrix}0.6 & -0.25 \\ 1 & 0\end{pmatrix}\begin{pmatrix}\rho_{k-1} \\ \rho_{k-2}\end{pmatrix}$$

To write a deterministic form, we need to utilize the eigendecomposition. Thus we need to find the eigenvalues of this matrix.

$$\det\begin{pmatrix}0.6 - \lambda & -0.25 \\ 1 & -\lambda\end{pmatrix} = \lambda^2 - 0.6\lambda + 0.25 = 0$$

Then we find two complex eigenvalues:

$$\lambda_{1,2} = 0.3 \pm 0.4i$$

Due to the decomposition form, as well as $\rho_k \in \mathbb{R}$, we know the deterministic form can must be written as:

$$\rho_k = (a + bi)(0.3 + 0.4i)^k + (a - bi)(0.3 - 0.4i)^k$$

We know the value of ρ_0 and ρ_1 so:

$$\begin{aligned}\rho_0 &= 2a = 1\\ \rho_1 &= 0.6a - 0.8b = 0.48\end{aligned}$$

This is fortunately easy to calculate:

$$a = \frac{1}{2}, b = -\frac{9}{40}$$

For simplicity just keep a and b and start to rewrite the function more:

$$\begin{aligned}\rho_k = &\ (a + bi)(0.3 + 0.4i)^k + (a - bi)(0.3 - 0.4i)^k\\ = &\ (a + bi) \cdot \frac{1}{2^k} \cdot (\cos(k \cdot \arctan\frac{4}{3}) + i \sin(k \cdot \arctan\frac{4}{3}))\\ &\quad + (a - bi) \cdot \frac{1}{2^k} \cdot (\cos(k \cdot \arctan\frac{4}{3}) - i \sin(k \cdot \arctan\frac{4}{3}))\\ = &\ \frac{1}{2^k}(2a \cdot \cos(k \cdot \arctan\frac{4}{3}) - 2b \cdot \sin(k \cdot \arctan\frac{4}{3}))\end{aligned}$$

So finally put a and b into the equation:

$$\rho_k = \frac{1}{2^k}(\cos(k \cdot \arctan\frac{4}{3}) + \frac{9}{20}\sin(k \cdot \arctan\frac{4}{3}))$$

yields the deterministic form.

Part 2

Question a

ARCH(m) can be written as AR(m) for $\{\varepsilon_t^2\}$. ARCH(m): $\sigma_t^2 = \alpha_0 + \sum_{i=1}^m \alpha_i \varepsilon_{t-i}^2$

Definition of Garch is:

$$\begin{aligned}\varepsilon_t^2 &= \sigma_t^2 + w_t \\ &= \alpha_0 + \sum_{i=1}^m \alpha_i \varepsilon_{t-i}^2 + w_t\end{aligned}$$

Where w_t is White noise, such that under the equation, the AR(m) process for $\{\varepsilon_t^2\}$ is obtained.

$$Var(\varepsilon_t|Y_{t-1}) = E(\varepsilon_t^2|Y_{t-1}) = \sigma_t^2$$

At the same time, following the model assumption:

$$Var(\varepsilon_t|Y_{t-1}) = \sigma_t^2 \cdot E(v_t^2|Y_{t-1}) = \sigma_t^2$$

Question b

Likelihood of GARCH(m,s)

From the normal distribution we can derive the conditional probability:

$$f(y_t|Y_{t-1}) = \frac{1}{\sqrt{2\pi\sigma_t^2}} e^{-\varepsilon_t^2/2\sigma_t^2}$$

Where, according to the model, $\sigma_t^2 = \alpha_0 + \sum_{i=1}^m \alpha_i \varepsilon_{t-i}^2 + \sum_{j=1}^s \beta_j \sigma_{t-j}^2$

Since we cannot assume y_0 is given,the unconditional probability of y_1 is:

$$f(y_1) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\varepsilon_1^2/2\sigma^2}$$

Where σ^2 is the unconditional variance.

As the error decomposition of Q1c, we can write the likelihood:

$$L = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\varepsilon_1^2/2\sigma^2} \times \prod_{t=2}^T \frac{1}{\sqrt{2\pi\sigma_t^2}} e^{-\varepsilon_t^2/2\sigma_t^2}$$

Take the logarithm obtains:

$$l = \log L \propto -\frac{1}{2} \sum_{t=k}^T \left(\log \alpha_0 + \sum_{i=1}^m \alpha_i \varepsilon_{t-i}^2 + \sum_{j=1}^s \beta_j \sigma_{t-j}^2 \right) - \sum_{t=k}^T \frac{\varepsilon_t^2}{2\sigma_t^2}$$

Whereby $k = \max(m, s) + 1$

Question c

Positivity Restriction: $\implies \alpha_0, \alpha_1, \alpha_2, \ldots, \alpha_m \geq 0$
Stationarity conditions: $\implies \sum_{i=1}^{m} \alpha_i + \sum_{j=1}^{s} \beta_j < 1$

Question d

The h-step-ahead forecast for the condition variance

$$\begin{aligned}
\hat{\sigma}^2_{n+h} &= E(\sigma^2_{n+h}|F_n) \\
&= \alpha_0 + \sum_{i=1}^{m} E(\alpha_i \varepsilon^2_{n+h-i}|F_n) + \sum_{j=1}^{s} E(E(\beta_j \varepsilon^2_{n+h-j}|F_{n+h-j-1})|F_n) \\
&= \alpha_0 + \sum_{i=1}^{m} \alpha_i \hat{\sigma}^2_{n+h-i} + \sum_{j=1}^{s} \beta_j \hat{\sigma}^2_{n+h-j}
\end{aligned}$$

There is no closed formula for the expression. But by writing recursively we will be able to write to $\hat{\sigma}^2_{n+1}$ which all variables are contained in F_n.

Question e

From 2b We know that the conditional probability is:

$$f(y_t|\mathcal{F}_{t-1}) = \frac{1}{\sqrt{2\pi\sigma^2_t}} e^{-\varepsilon^2_t/2\sigma^2_t}$$

Taking the logarithm to the conditional probability yields:

$$\log f(y_t|\mathcal{F}_{t-1}) = -\frac{1}{2}\log 2\pi - \frac{1}{2}\log \sigma^2_t - \frac{\varepsilon^2_t}{2\sigma^2_t}$$

Followed by taking the derivative:

$$\frac{\partial \log f(y_t|\mathcal{F}_{t-1})}{\partial \sigma^2_t} = -\frac{1}{2\sigma^2_t} + \frac{\varepsilon^2_t}{2\sigma^4_t} = \frac{\varepsilon^2_t - \sigma^2_t}{2\sigma^4_t}$$

Then we can obtain:

$$\begin{aligned}
E\left[\left(\frac{\partial \log f(y_t|\mathcal{F}_{t-1})}{\partial \sigma^2_t}\right)^2 \middle| \mathcal{F}_{n-1}\right] &= \frac{1}{4\sigma^8_t} E\left[(\varepsilon^2_t - \sigma^2_t)^2|\mathcal{F}_{n-1}\right] \\
&= \frac{1}{4\sigma^8_t} E\left[\sigma^4_t(v^2_t - 1)^2|\mathcal{F}_{n-1}\right] \\
&= \frac{1}{4\sigma^4_t} E\left[(v^2_t - 1)^2|\mathcal{F}_{n-1}\right] = \frac{1}{2\sigma^4_t}
\end{aligned}$$

As a result, $\mathcal{I}_t^{-1} = 2\sigma^4_t$

Question f

Using the result in 2e, we can write the equation given by:

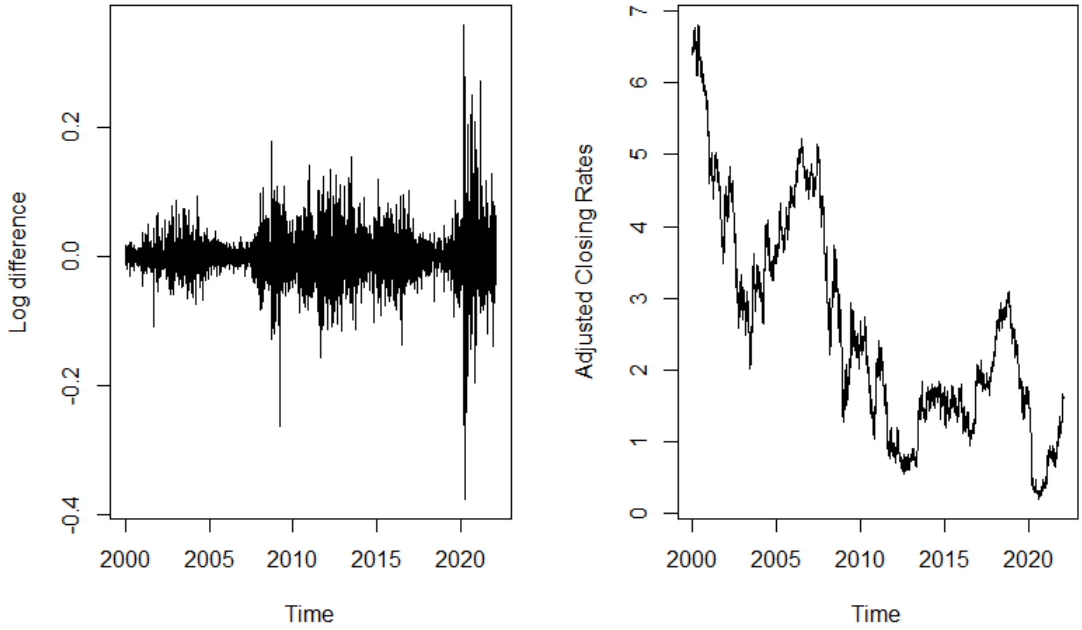
$$\begin{aligned}\sigma_{t+1}^2 &= \tilde{\alpha}_0 + \tilde{\alpha}_1 \; 2\sigma_t^4 \; \frac{(\varepsilon_t^2 - \sigma_t^2)}{2\sigma_t^4} + \tilde{\beta}_1 \sigma_t^2 \\ &= \tilde{\alpha}_0 + \tilde{\alpha}_1 \varepsilon_t^2 + (\tilde{\alpha}_1 + \tilde{\beta}_1)\sigma_t^2\end{aligned}$$

Which is indeed equivalent to GARCH(1,1) updating equation for σ_t^2

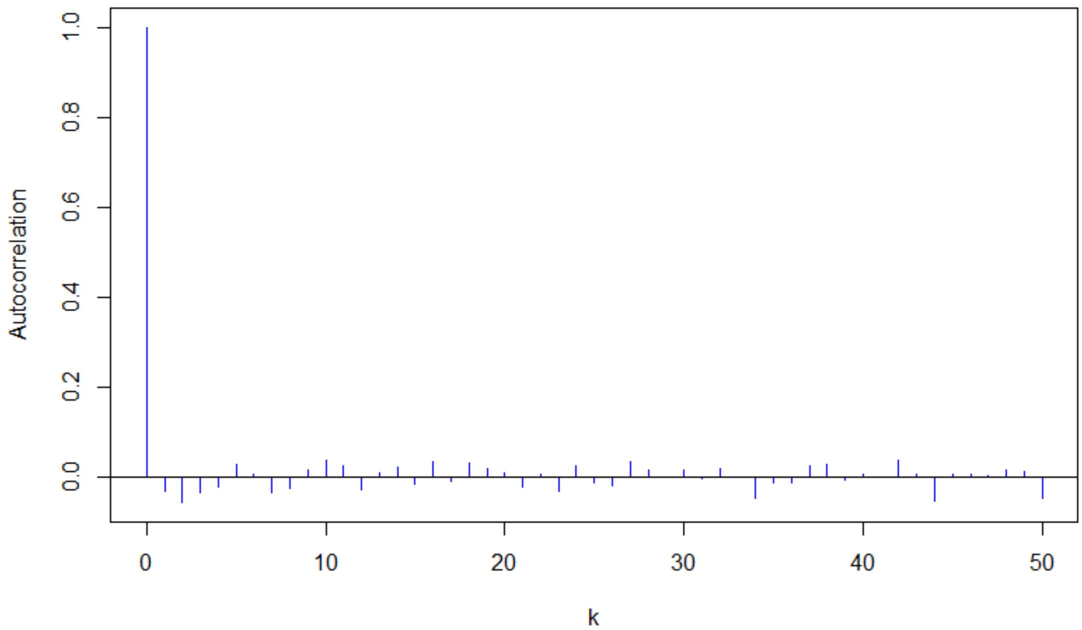
Part 3

Question b

Log returns on the left, Rate's on the right:



ACF plot:



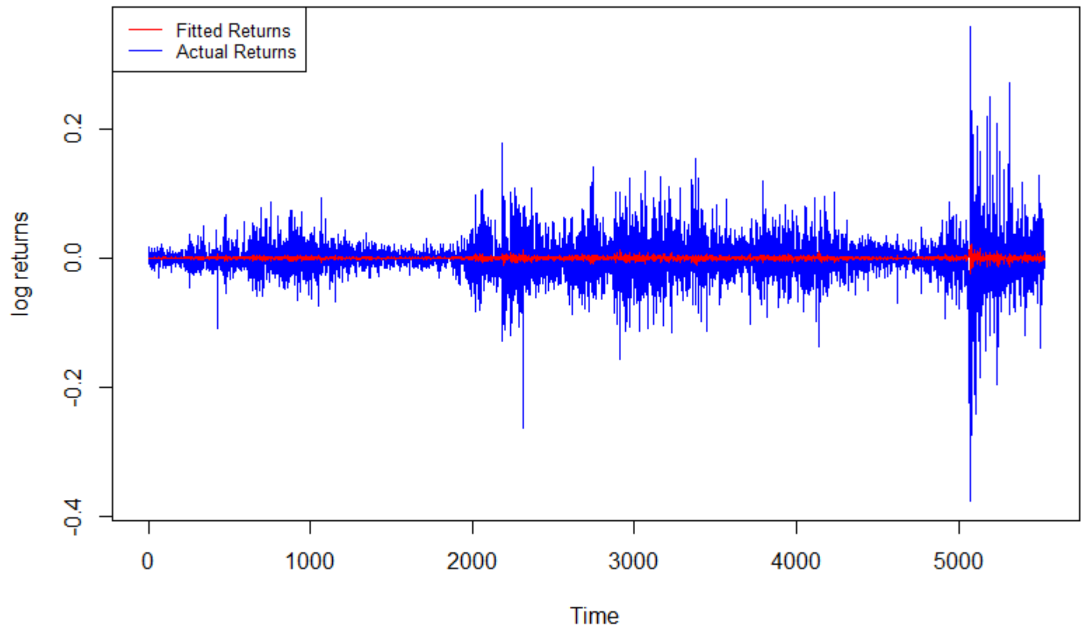
ACF shows an exponential decay, furthermore the graph for the log difference looks stable with constant mean around 0 and doesn't appear to have any correlation with the lags therefore we believe the log returns time series is stationary.

Question c

	b_ols	b_mle
intercept	-0.0003227341	-0.0003092789
ar1	-0.0344720061	-0.0344957241
ar2	-0.0567299534	-0.0564862934

Question d

Fitted returns and action returns:



	b_ols	b_mle	b_tseries
intercept	-0.0003227341	-0.0003092789	-0.0003209261
ar1	-0.0344720061	-0.0344957241	-0.0344415023
ar2	-0.0567299534	-0.0564862934	-0.0566947113

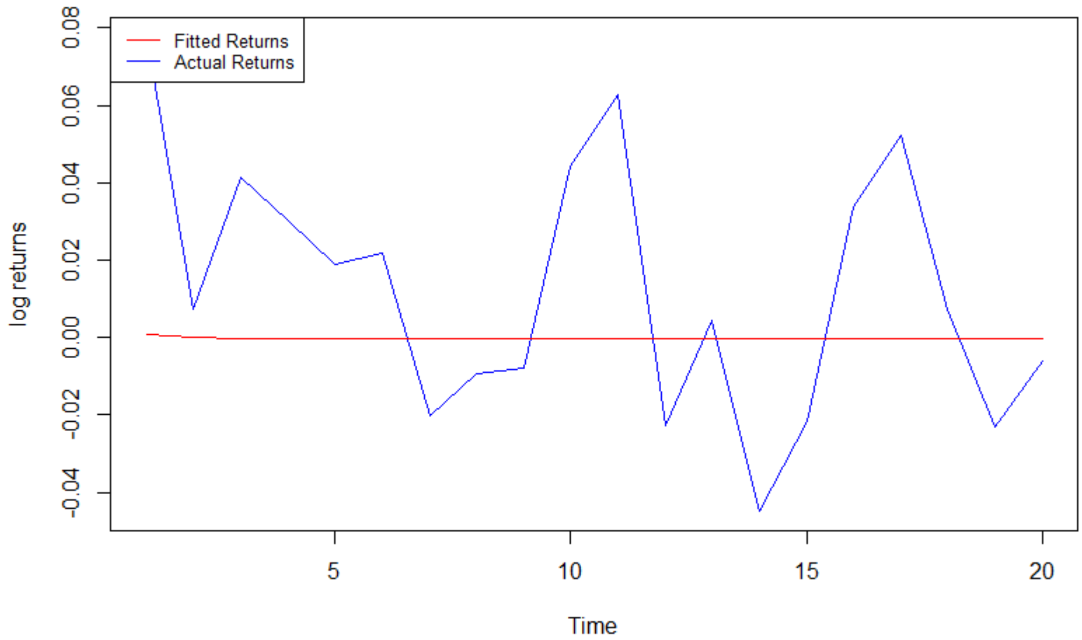
Quality of fit is very poor as seen in the red plot.

e

Normality assumption is rejected at 95% by jacque-Bera Test.

f

Action returns and forecast:

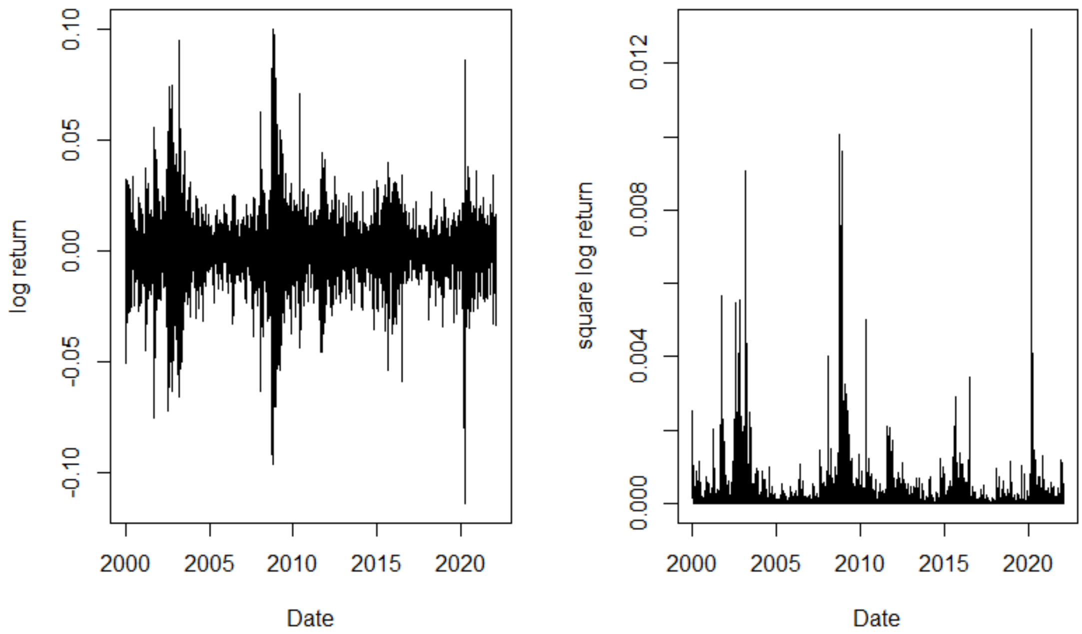


Quality of the forecast is rather poor as the difference between the blue curve to the red is quite large.

Part 4

Question b

Log-returns on the left and squared log-returns on the right:



Volatility is not constant over time as can be seen with the multiple huge spikes.

Question c

The estimation of $\sigma_t^2 = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \alpha_2 \sigma_{t-1}^2$

	α_0	α_1	α_2
value	2.272072e-06	1.100749e-01	8.772182e-01

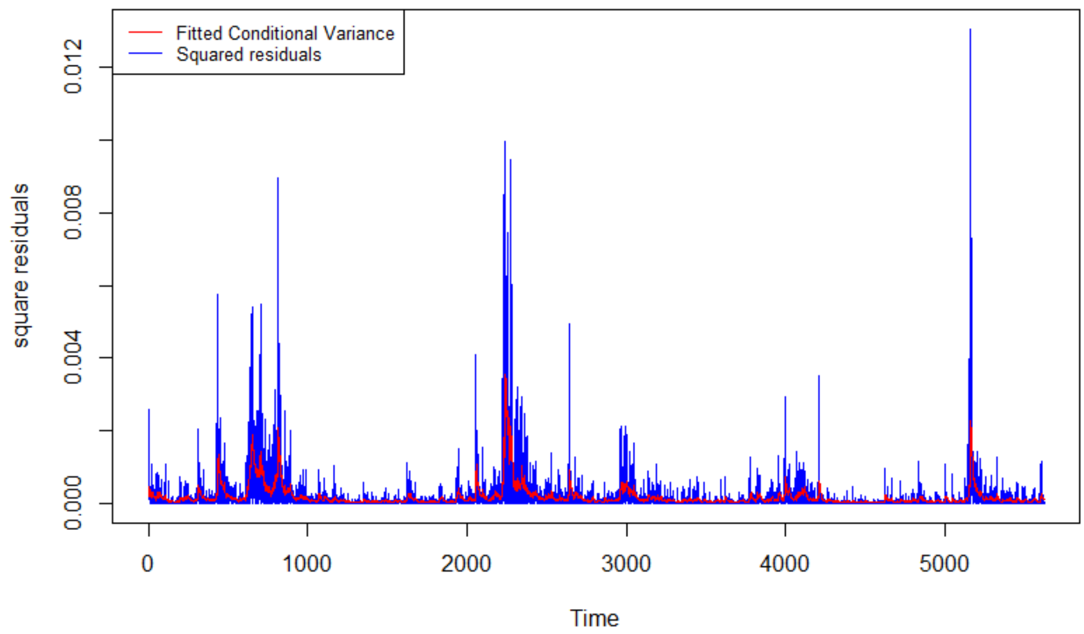
0.11 of the volatility shock today feeds through to tomorrow and we see that the coefficient of ε_{t-1} and σ_{t-1} sums nearly to 1, it might suggest that volatility is persistant.

Question d

Normality assumption rejected at 95% confidence with the Jacque-Bera Test, other assumption we can think of for the error term could be a student-t distribution.

Question e

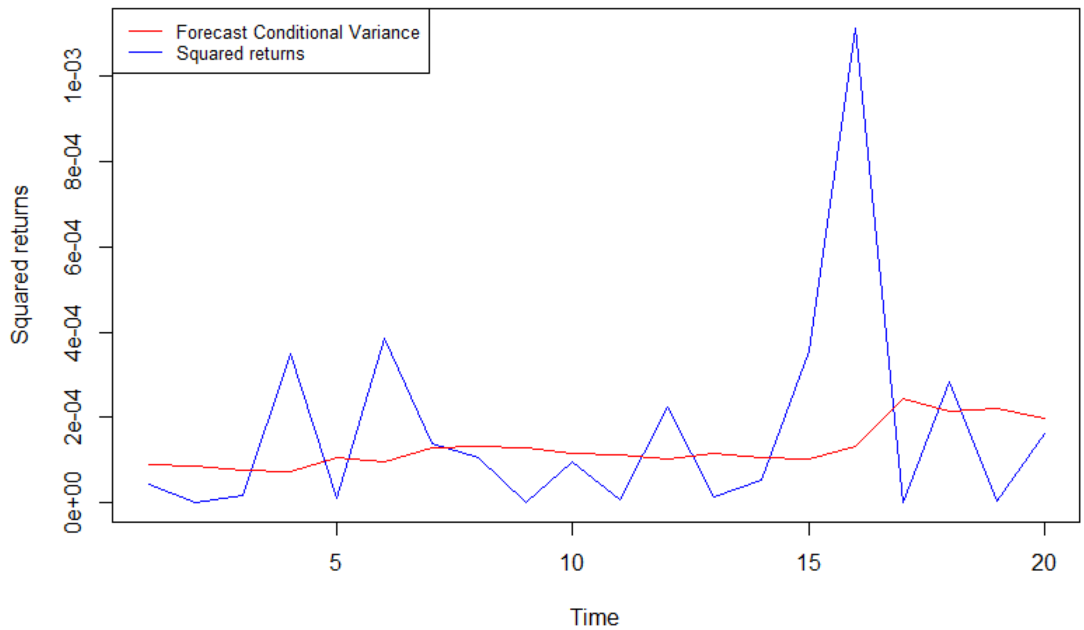
	b_ug	b_mle
omega	2.360946e-06	2.272072e-06
alpha1	1.136631e-01	1.100749e-01
beta1	8.731156e-01	8.772182e-01



Quality of fit is quite good except for values whereby the square residuals are large.

Question f

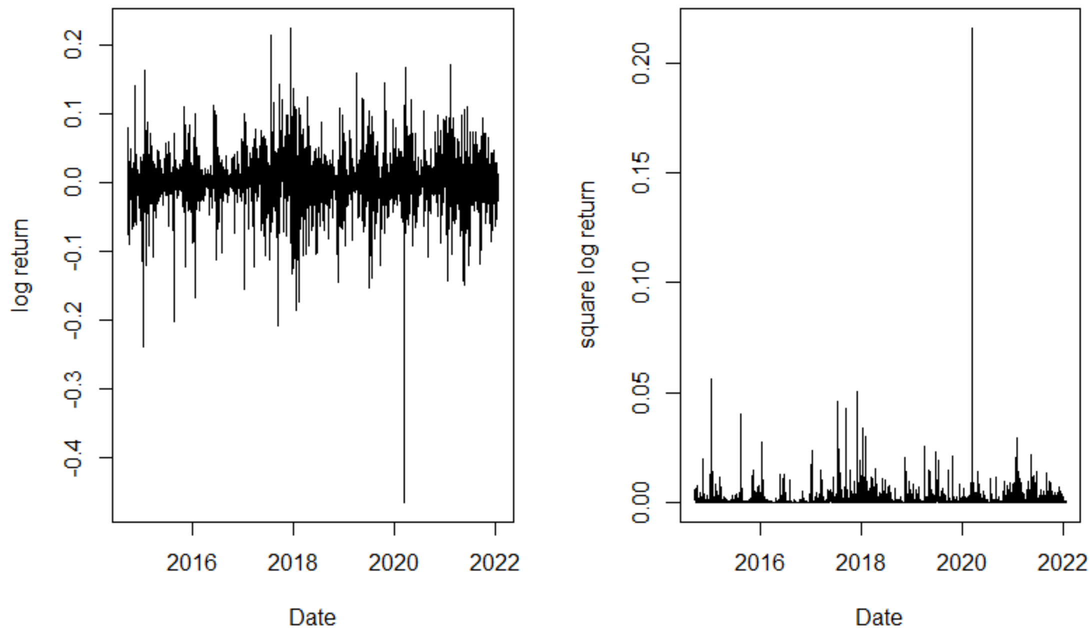
Forecast plot:



Quality of fit better than in question 3 however observation 16 one can see that the forecast is extremely poor.

Question g

Bitcoin log-returns and squared log-returns:



Volatility doesn't seem to be constant over time.

The estimation of $\sigma_t^2 = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \alpha_2 \sigma_{t-1}^2$

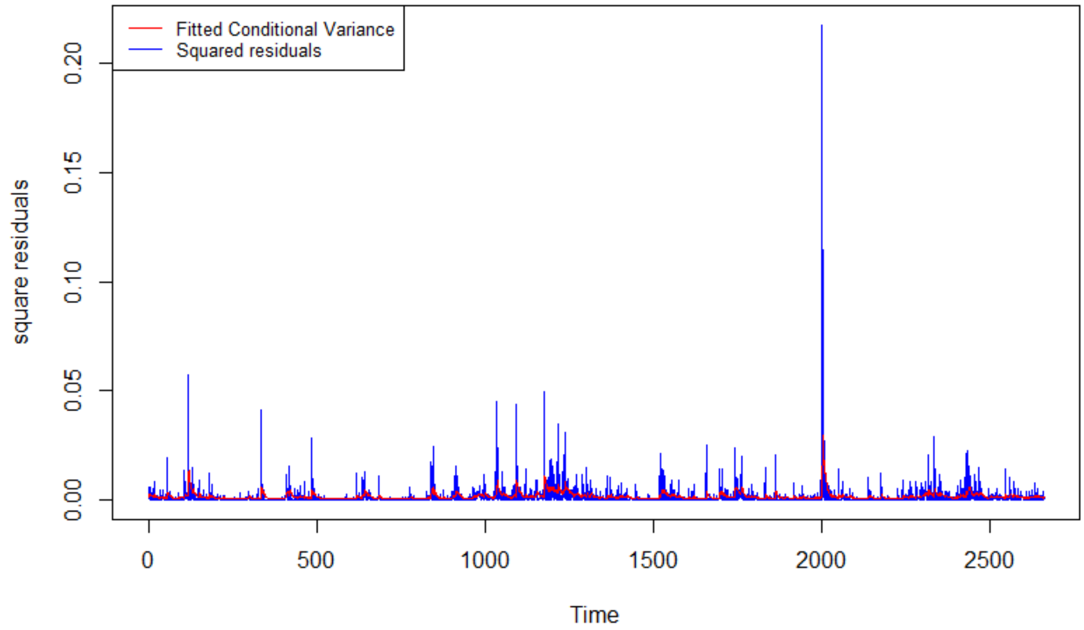
	α_0	α_1	α_2
value	6.840742e-05	1.295358e-01	8.399760e-01

0.13 of the volatility shock today feeds through to tomorrow and we see that the coefficient of ε_{t-1} and σ_{t-1} sums nearly to 1, it might suggest that volatility is persistant.

Normality assumption rejected by jacque-bera test at 95%

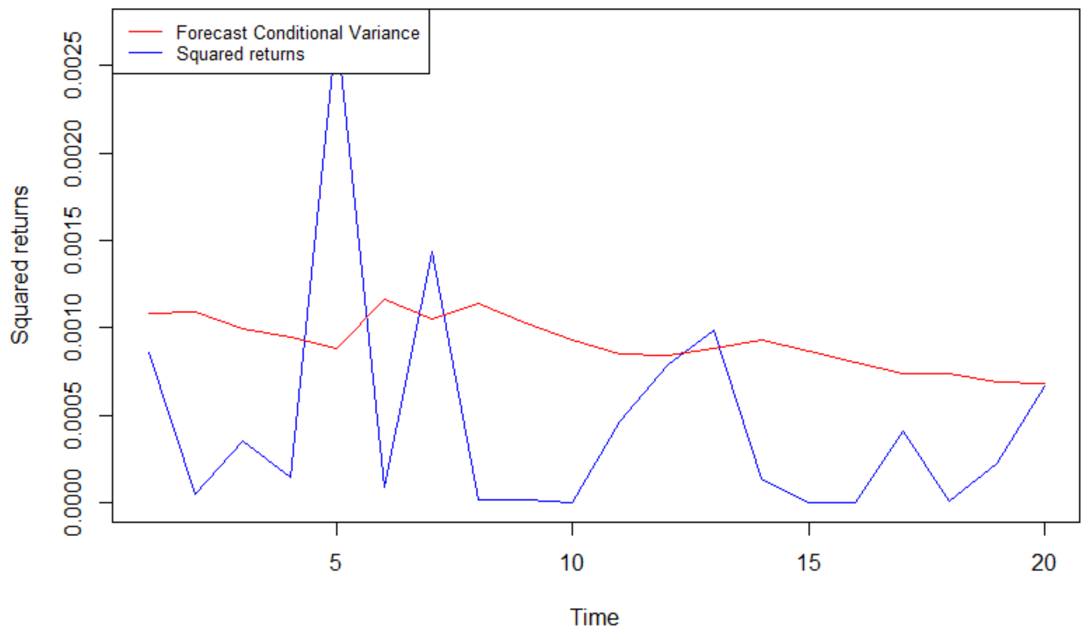
	b_ug	b_mle
omega	6.824091e-05	6.840742e-05
alpha1	1.300869e-01	1.295358e-01
beta1	8.395176e-01	8.399760e-01

Fitted condition variances and squared residuals for Bitcoin:



Quality of fit good except for large values of square residuals.

Conditional variance forecasts and actual square returns Bitcoin:



The forecast is not as good as for AEX data, the forecast is very poor.