Time Series Analysis: Assignment 3

Bowen Ma (12960780), Mianyun He (13605275), Shiyi Yang (13627295)

Part 1

a.

Consider the following GARCH(m, s) model,

$$\begin{aligned} y_t &= \mu + \varepsilon_t \\ \varepsilon_t &= \sigma_t \nu_t \\ \sigma_t^2 &= \alpha_0 + \sum_{i=1}^{m} \alpha_i \varepsilon_{t-i}^2 + \sum_{j=1}^{s} \beta_j \sigma_{t-j}^2 \end{aligned}$$

where $\nu_t \overset{iid}{\sim} \mathcal{N}(0,1)$. By definition, the h-step-ahead forecast for the conditional variance is given as: $\hat{\sigma}_{n+h}^2 = E\left[\sigma_{n+h}^2 | \mathcal{F}_n\right]$. We first calculate the one-step ahead forecast and then generalize the recursive formulation for the h-step ahead forecast.

One-step ahead:

$$\hat{\sigma}_{n+1}^2 = \sigma_{n+1}^2 = \alpha_0 + \sum_{i=1}^{m} \alpha_i \varepsilon_{n+1-i}^2 + \sum_{j=1}^{s} \beta_j \sigma_{n+1-j}^2 \quad (\sigma_{n+1}^2 \text{ is actually in } \mathcal{F}_n)$$

h-step ahead:

$$\begin{aligned} \hat{\sigma}_{n+h}^2 &= \alpha_0 + \sum_{i=1}^{m} E\left[\alpha_i \varepsilon_{n+h-i}^2 | \mathcal{F}_n\right] + \sum_{j=1}^{s} E\left[\beta_j \sigma_{n+h-j}^2 | \mathcal{F}_n\right] \\ \Rightarrow \hat{\sigma}_{n+h}^2 &= \alpha_0 + \sum_{i=1}^{m} \alpha_i E\left[\varepsilon_{n+h-i}^2 | \mathcal{F}_n\right] + \sum_{j=1}^{s} \beta_j E\left[\sigma_{n+h-j}^2 | \mathcal{F}_n\right] \\ \Rightarrow \hat{\sigma}_{n+h}^2 &\overset{(*)}{=} \alpha_0 + \sum_{i=1}^{m} \alpha_i \hat{\sigma}_{n+h-i}^2 + \sum_{j=1}^{s} \beta_j \hat{\sigma}_{n+h-j}^2 \end{aligned}$$

Since $\mathcal{F}_n \subset \mathcal{F}_{n+h-i-1}$:

$$\begin{aligned} (*) : \sigma_{n+h-i}^2 = E\left[\varepsilon_{n+h-i}^2 | \mathcal{F}_{n+h-i-1}\right] &\Rightarrow E\left[\sigma_{n+h-i}^2 | \mathcal{F}_n\right] = E\left[E\left[\varepsilon_{n+h-i}^2 | \mathcal{F}_{n+h-i-1}\right] | \mathcal{F}_n\right] \\ &\Rightarrow E\left[\sigma_{n+h-i}^2 | \mathcal{F}_n\right] = E\left[\varepsilon_{n+h-i}^2 | \mathcal{F}_n\right] \\ &\Rightarrow \hat{\sigma}_{n+h-i}^2 = E\left[\varepsilon_{n+h-i}^2 | \mathcal{F}_n\right] \end{aligned}$$

Hence, the recursive formulation of the h-step-ahead forecast is:

$$\hat{\sigma}_{n+h}^2 = \alpha_0 + \sum_{i=1}^{m} \alpha_i \hat{\sigma}_{n+h-i}^2 + \sum_{j=1}^{s} \beta_j \hat{\sigma}_{n+h-j}^2$$

b.

From our assumptions: $\varepsilon_t | \mathcal{F}_{t-1} \overset{iid}{\sim} \mathcal{N}(0, \sigma_t^2)$, since $y_t = \mu + \varepsilon_t$, $y_t | \mathcal{F}_{t-1} \overset{iid}{\sim} \mathcal{N}\left(\mu, \sigma_t^2\right)$

$$f\left(y_t | \mathcal{F}_{t-1}\right) = \frac{1}{\sigma_t \sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{y_t - \mu}{\sigma_t}\right)^2}$$

$$\log f\left(y_t | \mathcal{F}_{t-1}\right) \propto -\frac{1}{2} \log \sigma_t^2 - \frac{1}{2\sigma_t^2} \left(y_t - \mu\right)^2$$

The time varying parameter is the conditional variance σ_t^2, we can compute the conditional information matrix $\mathcal{I}_t$ by:

$$\mathcal{I}_t = E\left[\left(\frac{\partial \log f\left(y_t|\mathcal{F}_{t-1}\right)}{\partial \sigma_t^2}\right)^2 \middle| \mathcal{F}_{t-1}\right] \overset{\text{I.M.E.}}{=} -E\left[\frac{\partial^2 \log f\left(y_t|\mathcal{F}_{t-1}\right)}{\partial \left(\sigma_t^2\right)^2} \middle| \mathcal{F}_{t-1}\right]$$

The first and the second order derivative of the likelihood:

$$\frac{\partial \log f\left(y_t|\mathcal{F}_{t-1}\right)}{\partial \sigma_t^2} = -\frac{1}{2\sigma_t^2} + \frac{1}{2\left(\sigma_t^2\right)^2}\left(y_t - \mu\right)^2$$

$$\frac{\partial^2 \log f\left(y_t|\mathcal{F}_{t-1}\right)}{\partial (\sigma_t^2)^2} = \frac{1}{2\left(\sigma_t^2\right)^2} + \frac{\left(y_t-\mu\right)^2}{2} \cdot (-2) \cdot \left(\sigma_t^2\right)^{-3} = \frac{1}{2\sigma_t^4} - \frac{\left(y_t-\mu\right)^2}{\sigma_t^6}$$

Since we know from a) that $\sigma_t^2 \in \mathcal{F}_{t-1}$:

$$\begin{aligned}
\mathcal{I}_t &= -E\left[\frac{1}{2\sigma_t^4} - \frac{\left(y_t-\mu\right)^2}{\sigma_t^6} \middle| \mathcal{F}_{t-1}\right] \\
&= -\left(\frac{1}{2\sigma_t^4} - \frac{E\left[\left(y_t-\mu\right)^2|\mathcal{F}_{t-1}\right]}{\sigma_t^6}\right) \\
&= -\left(\frac{1}{2\sigma_t^4} - \frac{\sigma_t^2}{\sigma_t^6}\right) \\
&= \frac{1}{2\sigma_t^4}
\end{aligned}$$

Hence, the conditional information matrix is: $\mathcal{I}_t = \frac{1}{2\sigma_t^4}$, which results in $\mathcal{I}_t^{-1} = 2\sigma_t^4$.

c.

We know from part b) and the formulation of GAS(1,1) model:

$$\begin{aligned}
\sigma_{t+1}^2 &= \tilde{\alpha}_0 + \tilde{\alpha}_1\left(2\sigma_t^4\right)\left(-\frac{1}{2\sigma_t^2} + \frac{\left(y_t-\mu\right)^2}{2\sigma_t^4}\right) + \tilde{\beta}_1\sigma_t^2 \\
\Rightarrow \sigma_{t+1}^2 &= \tilde{\alpha}_0 + \tilde{\alpha}_1\left(\left(y_t-\mu\right)^2 - \sigma_t^2\right) + \tilde{\beta}_1\sigma_t^2 \\
\Rightarrow \sigma_{t+1}^2 &= \tilde{\alpha}_0 + \tilde{\alpha}_1\varepsilon_t^2 + \left(\tilde{\beta}_1 - \tilde{\alpha}_1\right)\sigma_t^2 \\
\Rightarrow \sigma_t^2 &= \tilde{\alpha}_0 + \tilde{\alpha}_1\varepsilon_{t-1}^2 + \left(\tilde{\beta}_1 - \tilde{\alpha}_1\right)\sigma_{t-1}^2
\end{aligned}$$

Define $\alpha_0 := \tilde{\alpha}_0,\ \alpha_1 := \tilde{\alpha}_1,\ \alpha_2 := (\tilde{\beta}_1 - \tilde{\alpha}_1):\ \sigma_t^2 = \alpha_0 + \alpha_1\varepsilon_{t-1}^2 + \alpha_2\sigma_{t-1}^2$.

Hence, under the information matrix scaling and Gaussian erros, the updating equation for GAS(1,1) is equivalent to the GARCH(1,1) updating equation for σ_t^2

d.

From $\hat{y}_{t+1}^{MA} = \omega\hat{y}_{1,t+1} + (1-\omega)\hat{y}_{2,t+1}$:

$$\begin{aligned}
E\left[\hat{y}_{t+1}^{MA}\right] &= E\left[\omega\hat{y}_{1,t+1} + (1-\omega)\hat{y}_{2,t+1}\right] \\
&= \omega E\left[\hat{y}_{1,t+1}\right] + (1-\omega)E\left[\hat{y}_{2,t+1}\right] \\
&= \omega E\left[y_{t+1} - e_{1,t+1}\right] + (1-\omega)E\left[y_{t+1} - e_{2,t+1}\right] \\
&= \omega\left(E\left[y_{t+1}\right] - E\left[e_{1,t+1}\right]\right) + (1-\omega)\left(E\left[y_{t+1}\right] - E\left[e_{2,t+1}\right]\right)
\end{aligned}$$

Since each individual forecast is unbiased: $E\left[e_{1,t+1}\right]=E\left[e_{2,t+1}\right]=0$:

$$E\left[\hat{y}_{t+1}^{MA}\right]=\omega E\left[y_{t+1}\right]+(1-\omega)E\left[y_{t+1}\right]=E\left[y_{t+1}\right]\Rightarrow E\left[y_{t+1}-\hat{y}_{t+1}^{MA}\right]=E\left[e_{t+1}^{MA}\right]=0$$

Hence, the combined forecast $\hat{y}_{t+1}^{MA}$ is unbiased for any value of ω.

e.

To minimize $\sigma_{MA}^2=E\left[\left(y_{t+1}-\hat{y}_{t+1}^{MA}\right)^2\right]$ for ω and $1-\omega$, we first simplify σ_{MA}^2:

$$\begin{aligned}
\sigma_{MA}^2 &= E\left[\left(y_{t+1}-\omega\hat{y}_{1,t+1}-(1-\omega)\hat{y}_{2,t+1}\right)^2\right]\\
&= E\left[\left(\omega\left(y_{t+1}-\hat{y}_{1,t+1}\right)+(1-\omega)\left(y_{t+1}-\hat{y}_{2,t+1}\right)\right)^2\right]\\
&= E\left[\left(\omega e_{1,t+1}+(1-\omega)e_{2,t+1}\right)^2\right]\\
&= E\left[\omega^2 e_{1,t+1}^2+2\omega(1-\omega)e_{1,t+1}e_{2,t+1}+(1-\omega)^2 e_{2,t+1}^2\right]\\
&= \omega^2 E\left[e_{1,t+1}^2\right]+2\omega(1-\omega)E\left[e_{1,t+1}e_{2,t+1}\right]+(1-\omega)^2 E\left[e_{2,t+1}^2\right]
\end{aligned}$$

Since $E\left[e_{1,t+1}\right]=E\left[e_{2,t+1}\right]=0$: $E\left[e_{1,t+1}^2\right]=\operatorname{Var}\left(e_{1,t+1}\right)=\sigma_1^2$, $E\left[e_{2,t+1}^2\right]=\operatorname{Var}\left(e_{2,t+1}\right)=\sigma_2^2$, $E\left[e_{1,t+1}e_{2,t+1}\right]=\operatorname{cov}\left(e_{1,t+1},e_{2,t+1}\right)=\sigma_{12}$

Hence, $\sigma_{MA}^2=\omega^2\sigma_1^2+2\omega(1-\omega)\sigma_{12}+(1-\omega)^2\sigma_2^2$.

F.O.C:

$$2\omega\sigma_1^2+(2-4\omega)\sigma_{12}-2(1-\omega)\sigma_2^2=0\Rightarrow\omega\left(\sigma_1^2+\sigma_2^2-2\sigma_{12}\right)=\sigma_2^2-\sigma_{12}\Rightarrow\omega^*=\frac{\sigma_2^2-\sigma_{12}}{\sigma_1^2+\sigma_2^2-2\sigma_{12}}$$

S.O.C

$$2\sigma_1^2-4\sigma_{12}+2\sigma_2^2>0$$

Note that the S.O.C must hold since $\sigma_{MA}^2=\left(\sigma_1^2-2\sigma_{12}+\sigma_2^2\right)\omega^2+\left(2\sigma_{12}-2\sigma_2^2\right)\omega+\sigma_2^2$. To ensure σ_{MA}^2 is always positive for $\forall\omega$, the parabola should be convex $\Rightarrow\sigma_1^2-2\sigma_{12}+\sigma_2^2>0$

Hence, the optimal $\omega^*=\frac{\sigma_2^2-\sigma_{12}}{\sigma_1^2+\sigma_2^2-2\sigma_{12}}$.

f.

We know $\omega^*=\frac{\sigma_2^2-\sigma_{12}}{\sigma_1^2+\sigma_2^2-2\sigma_{12}}$ is the weight for $\hat{y}_{1,t+1}$, and $(1-\omega^*)=\frac{\sigma_1^2-\sigma_{12}}{\sigma_1^2+\sigma_2^2-2\sigma_{12}}$ is the weight for $\hat{y}_{2,t+1}$:

If $\omega^*>1-\omega^*$: $\sigma_2^2-\sigma_{12}>\sigma_1^2-\sigma_{12}\Rightarrow\sigma_2^2>\sigma_1^2$. If the prediction error of model 2 is larger than model 1, model 1 has larger weight.

If $1-\omega^*>\omega^*$: $\sigma_1^2-\sigma_{12}>\sigma_2^2-\sigma_{12}\Rightarrow\sigma_1^2>\sigma_2^2$. In this case model 2 has larger weight as model 1 has a larger prediction error.

g.

From part e) we know: $\sigma_{MA}^2=\omega^2\sigma_1^2+2\omega(1-\omega)\sigma_{12}+(1-\omega)^2\sigma_2^2$, plug in the optimal weight ω^*:

$$\sigma_{MA}^2(\omega^*) = (\omega^*)\,\sigma_1^2 + 2\omega^*(1-\omega^*)\,\sigma_{12} + (1-\omega^*)^2\,\sigma_2^2$$

$$= \left(\frac{\sigma_2^2-\sigma_{12}}{\sigma_1^2+\sigma_2^2-2\sigma_{12}}\right)^2 \sigma_1^2 + 2\cdot\left(\frac{\sigma_2^2-\sigma_{12}}{\sigma_1^2+\sigma_2^2-2\sigma_{12}}\cdot\frac{\sigma_1^2-\sigma_{12}}{\sigma_1^2+\sigma_2^2-2\sigma_{12}}\right)\cdot\sigma_{12} + \left(\frac{\sigma_1^2-\sigma_{12}}{\sigma_1^2+\sigma_2^2-2\sigma_{12}}\right)^2 \sigma_2^2$$

$$= \frac{\sigma_2^4\sigma_1^2+\sigma_1^4\sigma_2^2-4\sigma_1^2\sigma_2^2\sigma_{12}+2\sigma_1^2\sigma_2^2\sigma_{12}+\sigma_1^2\sigma_{12}^2+\sigma_2^2\sigma_{12}^2-2\sigma_1^2\sigma_{12}^2-2\sigma_2^2\sigma_{12}^2+2\sigma_{12}^3}{\left(\sigma_1^2+\sigma_2^2-2\sigma_{12}\right)^2}$$

$$= \frac{\sigma_1^2\sigma_2^2\left(\sigma_1^2+\sigma_2^2-2\sigma_{12}\right)-\sigma_{12}^2\left(\sigma_1^2+\sigma_2^2-2\sigma_{12}\right)}{\left(\sigma_1^2+\sigma_2^2-2\sigma_{12}\right)^2}$$

$$= \frac{\sigma_1^2\sigma_2^2-\sigma_{12}^2}{\sigma_1^2+\sigma_2^2-2\sigma_{12}}$$

Since the correlation coefficient $\rho_{12} = \frac{\sigma_{12}}{\sigma_1\sigma_2}$, we have $\sigma_{MA}^2(\omega^*) = \frac{\sigma_1^2\sigma_2^2\left(1-\rho_{12}^2\right)}{\sigma_1^2+\sigma_2^2-2\sigma_1\sigma_2\rho_{12}}$.
Since $\sigma_1^2>0, \sigma_2^2>0, \left(1-\rho_{12}^2\right)\geq 0 \Rightarrow \sigma_1^2\sigma_2^2\left(1-\rho_{12}^2\right)\geq 0 \Rightarrow \sigma_1^2+\sigma_2^2-2\sigma_1\sigma_2\rho_{12}>0$

We can apply a trick to show $\sigma_{MA}^2(w^*) \leq \min\left(\sigma_1^2,\sigma_2^2\right)$:

Consider $\left(\sigma_1-\rho_{12}\sigma_2\right)^2 \geq 0$, which always holds:

$$\begin{aligned}
&\Rightarrow \sigma_1^2-2\sigma_1\sigma_2\rho_{12}+\rho_{12}^2\sigma_2^2\geq 0\\
&\Rightarrow \sigma_1^2+\sigma_2^2-2\sigma_1\sigma_2\rho_{12}\geq\sigma_2^2-\rho_{12}^2\sigma_2^2\\
&\Rightarrow \sigma_1^2+\sigma_2^2-2\sigma_1\sigma_2\rho_{12}\geq\sigma_2^2\left(1-\rho_{12}^2\right)\\
&\Rightarrow 1\geq\frac{\sigma_2^2\left(1-\rho_{12}^2\right)}{\sigma_1^2+\sigma_2^2-2\sigma_1\sigma_2\rho_{12}} \quad \text{as} \quad \sigma_1^2+\sigma_2^2-2\sigma_1\sigma_2\rho_{12}>0\\
&\Rightarrow \sigma_1^2\geq\frac{\sigma_1^2\sigma_2^2\left(1-p_{12}^2\right)}{\sigma_1^2+\sigma_2^2-2\sigma_1\sigma_2\rho_{12}}\\
&\Rightarrow \sigma_1^2\geq\sigma_{MA}^2(\omega^*)
\end{aligned}$$

Consider $\left(\sigma_2-\rho_{12}\sigma_1\right)^2 \geq 0$, which always holds:

$$\begin{aligned}
&\Rightarrow \sigma_2^2-2\sigma_1\sigma_2\rho_{12}+p_{12}^2\sigma_1^2\geq 0\\
&\Rightarrow \sigma_1^2+\sigma_2^2-2\sigma_1\sigma_2\rho_{12}\geq\sigma_1^2-\rho_{12}^2\sigma_1^2\\
&\Rightarrow 1\geq\frac{\sigma_1^2\left(1-\rho_{12}^2\right)}{\sigma_1^2+\sigma_2^2-2\sigma_1\sigma_2\rho_{12}} \quad \text{as} \quad \sigma_1^2+\sigma_2^2-2\sigma_1\sigma_2\rho_{12}>0\\
&\Rightarrow \sigma_2^2\geqslant\frac{\sigma_2^2\sigma_1^2\left(1-\rho_{12}^2\right)}{\sigma_1^2+\sigma_2^2-2\sigma_1\sigma_2\rho_2}\\
&\Rightarrow \sigma_2^2\geqslant\sigma_{MA}^2(\omega^*)
\end{aligned}$$

Hence $\sigma_{MA}^2(\omega^*) \leq \min\left\{\sigma_1^2,\sigma_2^2\right\}$, since it's smaller than both of them.

h.

By definition, $\hat{y}_{t+1}^{MA} = \sum_{i=1}^N w_i\hat{y}_{i,t+1}$, where $\sum_{i=1}^N \omega_i = \omega'\iota = 1$. Σ_e is the variance-covariance matrix of N forecast errors. We need to find $\omega^* = (\omega_1^*,\ldots,\omega_N^*)'$ that minimizes $\operatorname{Var}\left(y_{t+1}-\hat{y}_{t+1}^{MA}\right) = \omega'\Sigma_e\omega$, under the restriction $w'\iota = 1$

Lagrangian:

$$\mathcal{L}(\omega, \lambda) = \omega' \Sigma_e \omega - \lambda \left(\omega' \iota - 1 \right)$$

F.O.C

$$\begin{aligned}
\frac{\partial \mathcal{L}}{\partial \omega} = 0 &\Rightarrow \frac{\partial \omega'}{\partial \omega} \Sigma_e \omega + \frac{\partial \omega' \Sigma_e'}{\partial \omega} \omega - \lambda \iota = 0 \\
&\Rightarrow 2 \Sigma_e \omega - \lambda \iota = 0 \\
&\Rightarrow \omega = \frac{\lambda}{2} \left(\Sigma_e \right)^{-1} \iota \\
&\Rightarrow \omega' = \frac{\lambda}{2} \iota' \left(\Sigma_e \right)^{-1}
\end{aligned}$$

$$\begin{aligned}
\frac{\partial \mathcal{L}}{\partial \lambda} = 0 &\Rightarrow \omega' \iota - 1 = 0 \\
&\Rightarrow \frac{\lambda}{2} \iota' \left(\Sigma_e \right)^{-1} \iota - 1 = 0 \\
&\Rightarrow \lambda = 2 \left[\iota' \left(\Sigma_e \right)^{-1} \iota \right]^{-1} \\
&\Rightarrow \omega^* = \frac{\left(\Sigma_e \right)^{-1} \iota}{\iota' \left(\Sigma_e \right)^{-1} \iota}
\end{aligned}$$

Hence, the optimal weights of the model averaging forecast $\omega^* = \frac{(\Sigma_e)^{-1} \iota}{\iota' (\Sigma_e)^{-1} \iota}$

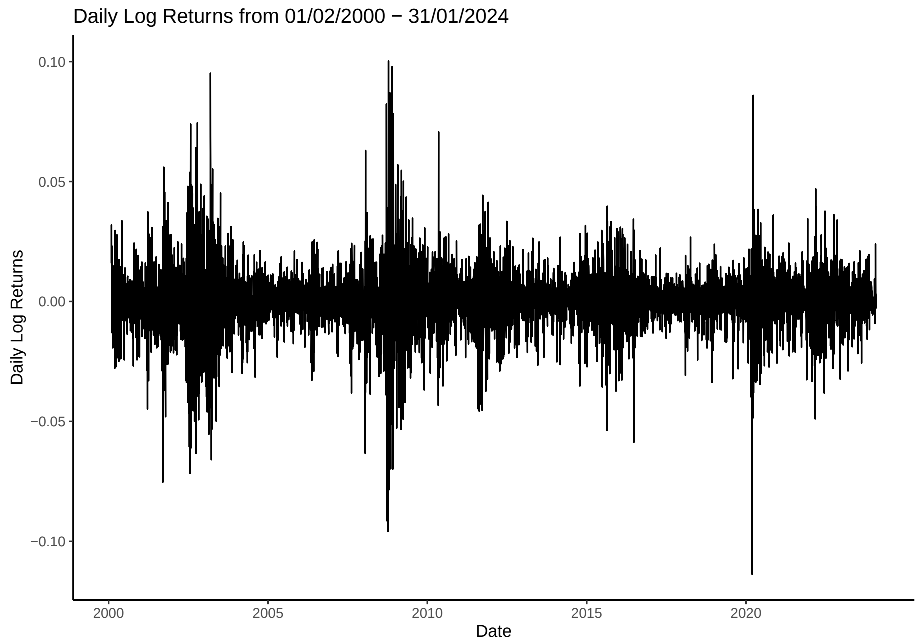
Part 2

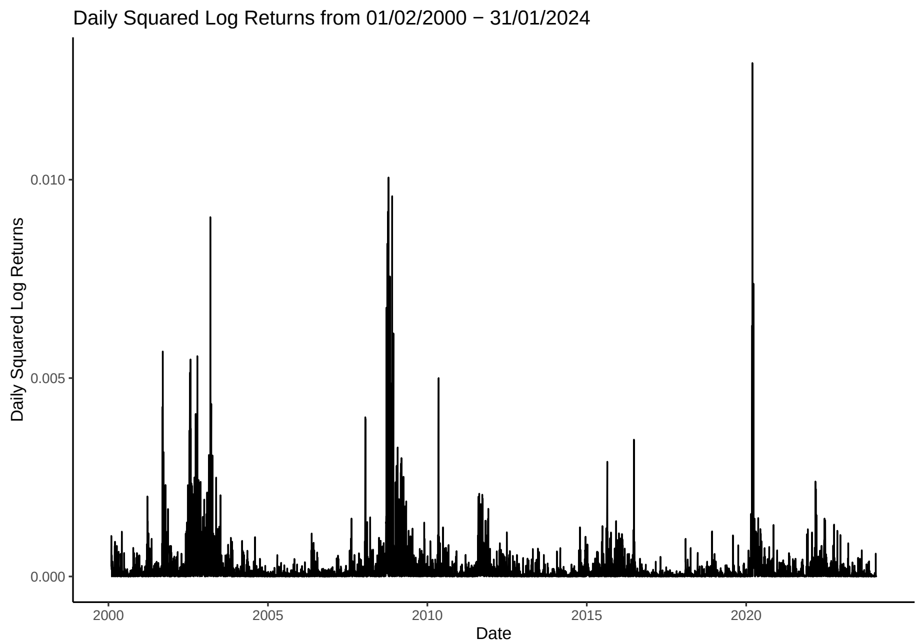
a.

The R script imports the data from the website of Yahoo Finance for the for the Amsterdam Exchange Index (ticker: ˆAEX) from February 1, 2000, up to and including January 31, 2024, and extracts the adjusted closing prices to form a time series $\{y_t\}$. Any null values in the dataset are replaced with NA and then removed. The script further processes the data to extract prices and dates, ensuring they are converted to the appropriate data types.

b.

As in the previous assignment, we compute the daily log-return at time t as $\log(\frac{y_t}{y_{t-1}})$, the daily squared daily log-return at time t as $\left(\log(\frac{y_t}{y_{t-1}})\right)^2$, and plot them as follows:





Analysis of the depicted plots reveals discernible periods characterized by elevated volatility, notably in the years 2003, 2008, and 2022, interspersed with intervals featuring diminished volatility. This observation underscores the non-constancy of volatility across time, implying the presence of heteroskedasticity in the data. Furthermore, an observed pattern indicates that elevated positive or negative returns tend to be succeeded by similarly pronounced positive or negative returns, implying a clustering of volatilities in the data.

c.

In part c, we employ Maximum Likelihood Estimation (MLE) to estimate the GARCH(1,1) model, assuming $\mu = 0$. For initializing σ_t^2, we use the sample variance of log-returns across the entire estimation sample as the initial conditional variance, and take ε_t as the starting value for the error term. It's crucial to ensure that the estimated α coefficients remain strictly positive to prevent negative fitted conditional variances, as emphasized in the second assignment. To address this, we employ the exponential function as our link function, ensuring the positivity of our parameters in the GARCH(1,1) model, as advised in the first tutorial question. After this, we use the *optim* function to determine the maximum likelihood estimates for the model parameters. Finally, our parameter estimates for $(\alpha_0, \alpha_1, \alpha_2)$ are $(0.000002305149, 0.1104659, 0.8763329)$, with a maximum log-likelihood of 19030.81. The fitted GARCH(1,1) is as follows:

$$\sigma_t^2 = 0.000002305149 + 0.1104659\varepsilon_{t-1}^2 + 0.8763329\sigma_{t-1}^2$$

The estimated coefficient for the lagged log returns, α_1, is 0.1104659. This suggests that, given the coefficients $(\alpha_0, \alpha_1, \alpha_2) < 1$, the estimated conditional variance is stationary. Additionally, the estimated coefficient for the lagged log returns implies that its effect on the conditional variance is positive and less than 1.

d.

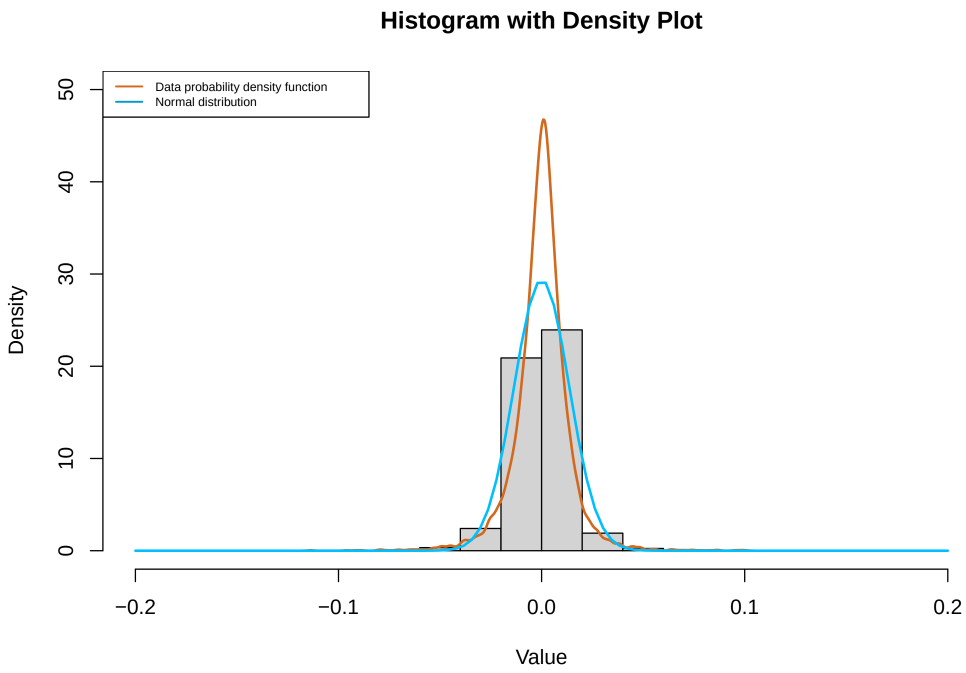
To assess the normality of the distribution, the Jarque–Bera test was employed. Initially, we computed the skewness (S) and kurtosis (K) of the distribution, utilizing the following formulas:

$$S = \frac{\frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^3}{\left(\frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2\right)^{\frac{3}{2}}}, \quad \text{and} \quad K = \frac{\frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2}{\left(\frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2\right)^2}$$

Subsequently, the Jarque-Bera test statistic (JB) was calculated using the formula:

$$JB = \frac{n}{6}(S^2 + \frac{1}{4}(K-3)^2)$$

The calculated Jarque-Bera statistic yielded a value of 13965.44, surpassing the critical value of $\chi_{2,0.05}$, thus leading to the rejection of the null hypothesis of normality at 5% significance level. Consequently, we infer that the errors deviate from a normal distribution.

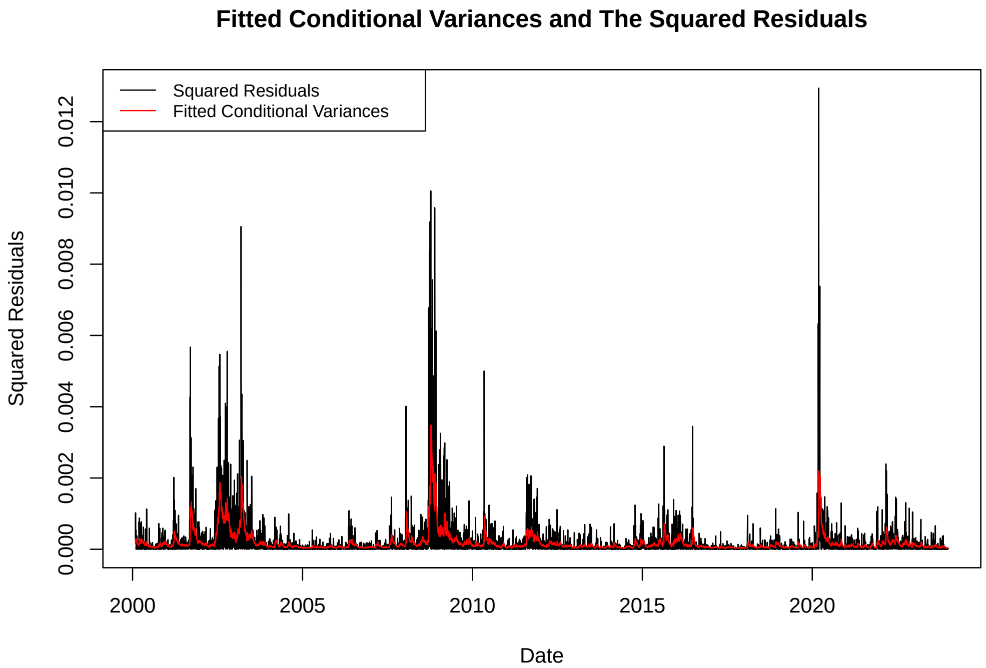


To assess the distribution of the error terms, a histogram was constructed, accompanied by a density function overlay depicting a normal distribution. The observed characteristics of the error distribution suggest a plausible conformity to a t-distribution. This conjecture is supported by the higher kurtosis and fatter tails exhibited in comparison to the normal distribution.

e.

In this part, we use the "rugarch" package in R to estimate the GARCH(1,1) model. Assuming normality in the distribution specification of the "ugarchspec" function, we obtain parameter estimates of $(\alpha_0, \alpha_1, \alpha_2) = (0.000002315583, 0.1108334, 0.8759291)$, with a maximum log-likelihood of 19027.62. Upon comparison with the results obtained using Maximum Likelihood Estimation (MLE) in part (c), we observe that these estimates are very close to each other.

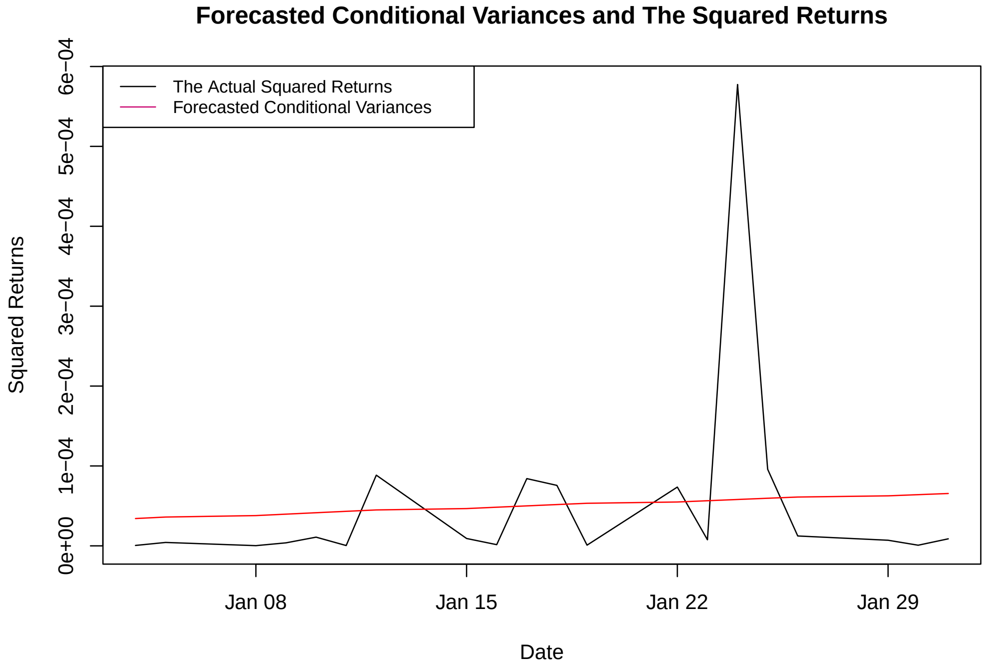
Subsequently, we proceeded to visualize the fitted conditional variances and the squared residuals over time. Upon inspection of the plot, it becomes evident that the fit adequately captures the general trend of the squared residuals. However, it is notable that the fitted conditional variances failed to accurately capture the occurrence of extreme high values in the squared residuals. This discrepancy can be attributed to the normality assumption in the model specification, leading to an underestimation of the squared residuals at extreme values.



f.

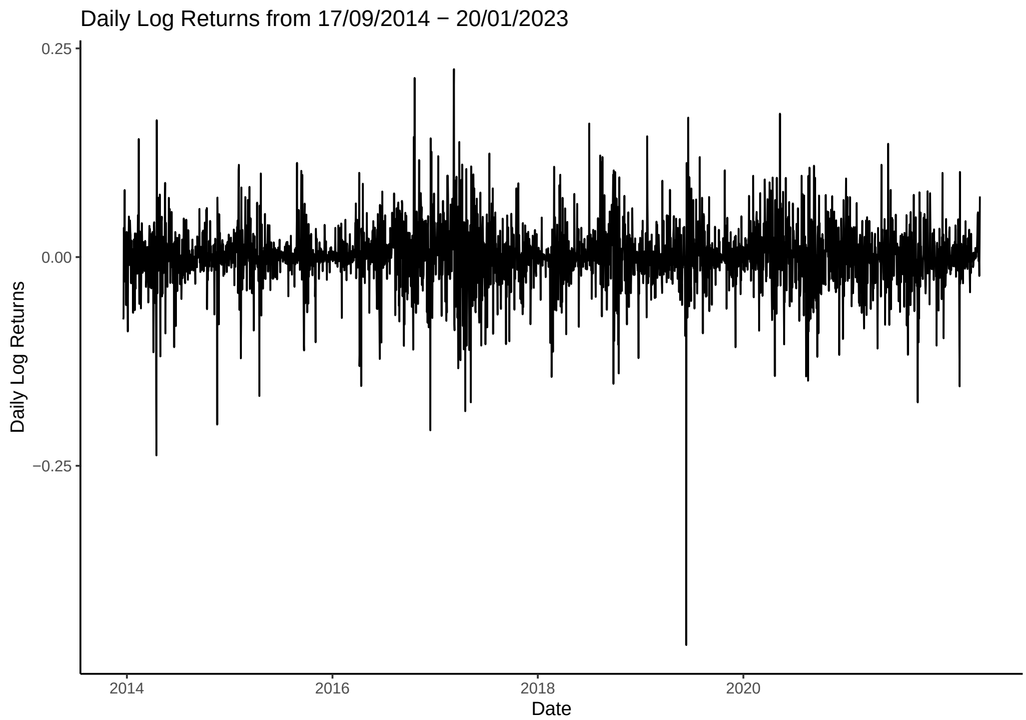
Next, we utilize the estimated parameters obtained in part c to generate forecasts for the conditional variance for the next 20 days. Employing the recursive formulation outlined in 1a), our forecast follows the pattern: $\hat{\sigma}^2_{n+h} = \alpha_0 + (\alpha_1 + \alpha_2)\hat{\sigma}^2_{n+h-1}$. Similar to previous assignments, we initialize our forecast with the last in-sample $\hat{\sigma}^2_t$ and proceed recursively.

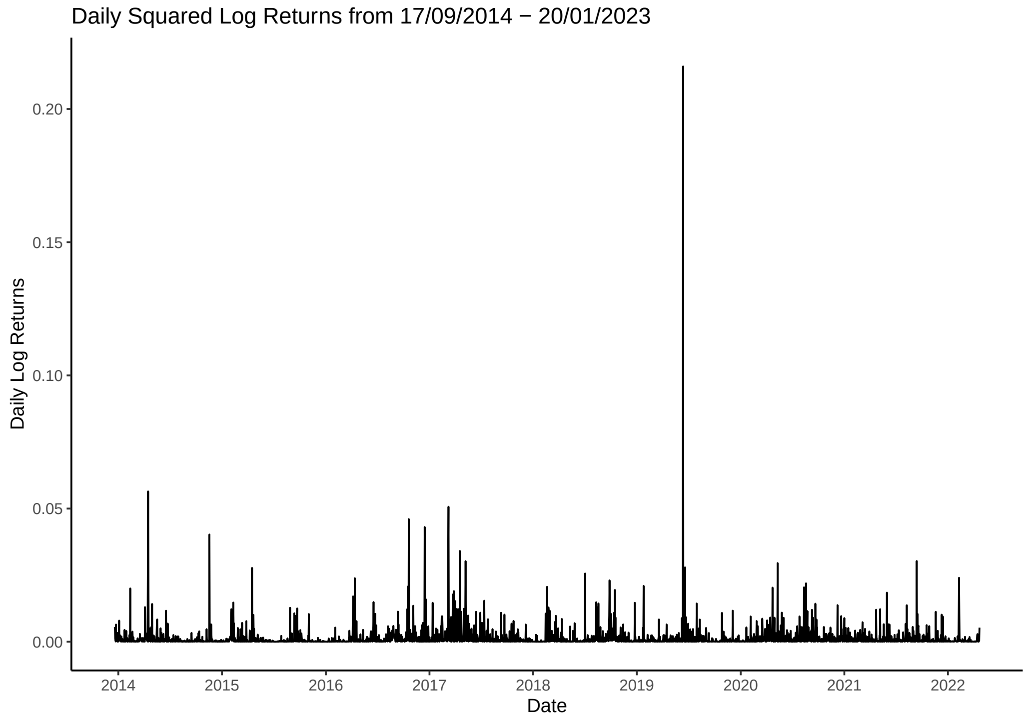
Subsequently, we visualize the forecasted conditional variance alongside the actual squared returns for these 20 days within a single plot. The forecasted conditional variance exhibits a generally flat trajectory with a slight upward trend. While the forecast values generally align with the average of the actual squared returns, they fail to capture the sharp spike and subsequent drastic drop in the squared returns observed between January 23rd and January 26th, 2023.



g.

In this section, we replicate the computational procedure outlined in 2a) to compute the log-returns and squared log-returns for Bitcoin daily data spanning from September 17, 2014, to January 20, 2023. Upon examination, it becomes apparent that the log-returns of Bitcoin typically are larger than those of the AEX, as evidenced by the values on the vertical axis. Additionally, the log-returns of Bitcoin display less significant volatility clustering when compared to those of the AEX, and we observe a prominent surge in volatility for Bitcoin, notably occurring between 2019 and 2020. This noteworthy fluctuation could be attributed to the outbreak of global pandemic.





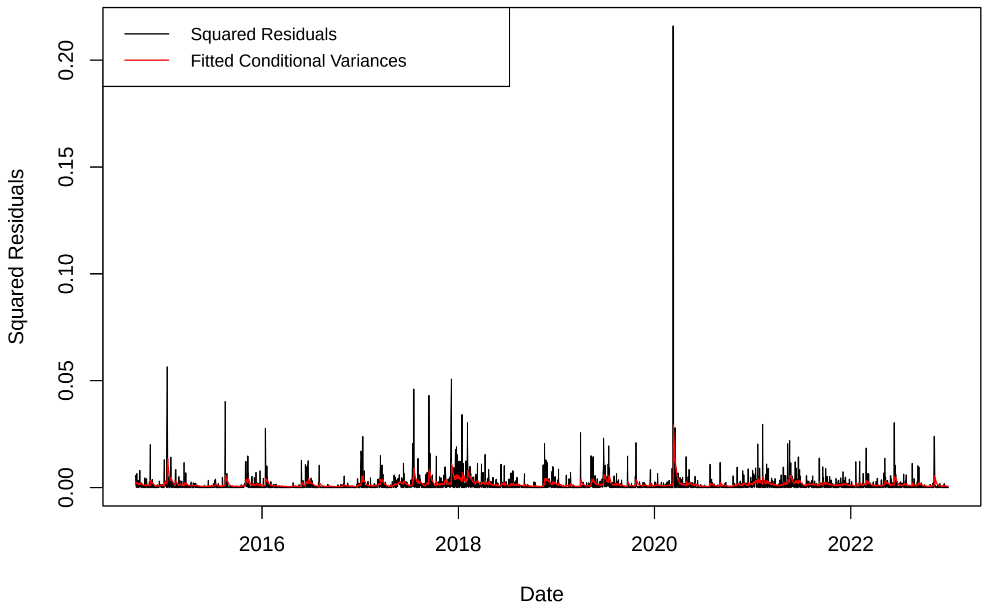
Once more, we utilize Maximum Likelihood Estimation (MLE) along with the "rugarch" package to estimate the GARCH(1,1) model for the Bitcoin dataset. Through MLE, we obtain the estimated coefficient $(\alpha_0, \alpha_1, \alpha_2) = (0.000075, 0.130113, 0.832645)$, with a log-likelihood of 5784.861. Utilizing the package and assuming normality, we obtain $(\alpha_0, \alpha_1, \alpha_2) = (0.000076, 0.130426, 0.832051)$, yielding a log-likelihood of 5782.752. Once again, the estimated coefficients from MLE and "rugarch" are both positive and very close to each other. The impact of previous log-returns on the conditional variance remains positive and is slightly below 0.13. Comparing these results to those obtained from the AEX data, we observe similar estimated coefficients for previous log-returns and previous conditional variance. However, the GARCH(1,1) model for the Bitcoin data exhibits a higher intercept, suggesting a larger average conditional variance for Bitcoin data.

Subsequently, we proceeded to visualize the fitted conditional variances and the squared residuals over time. Again, it becomes evident that the fit adequately captures the general trend of the squared residuals, but the fitted conditional variances failed to accurately capture the occurrence of extreme high values in the squared residuals. This discrepancy can be again attributed to the normality assumption in the model specification, leading to an underestimation of the squared residuals at extreme values.

Ultimately, we present a visualization showcasing the forecasted conditional variance alongside the actual squared returns for a span of 20 days, all within a single plot. The forecasted conditional variance displays an upward trend, mirroring the overall increasing trend observed in the actual squared returns. However, the forecasts are unable to accurately capture the sharp spike and subsequent drastic drop observed in the actual squared returns.

In summary, the plots of daily squared log returns reveal distinct characteristics for both the AEX and Bitcoin. The AEX exhibits lower values but displays more significant heteroskedasticity and volatility clustering compared to Bitcoin, which demonstrates relatively higher values for squared log returns. The results from the GARCH(1,1) model confirm that Bitcoin indeed has a larger intercept, indicating higher average conditional variance. Additionally, the forecast of the conditional variance for Bitcoin shows an increasing trend, suggesting higher volatility compared to the forecast for the AEX, and indicating the returns of Bitcoin has a higher volatility.

Fitted Conditional Variances and The Squared Residuals



Forecasted Conditional Variances and The Squared Returns