For this assignment, you may work in groups of two or three students. Part 1 is theoretical and part 2 is an application of your results from part 1. For each sub-question, you can earn three points. It is crucial to motivate your answers: No points will be awarded for answers without motivation. Furthermore, reports that are not typeset as a single PDF file in LaTeX will not be considered. Please also make sure to submit your R (and/or Python) codes for question 2, separately, together with a CSV file of the data. For question 2, your code should reproduce all reported results without returning errors. Results that are not returned by your code cannot be graded.

The deadline for this assignment is Monday, February 26, 15:00.

Part 1. Reconsider the (cases) of the GARCH(m, s) model from Assignment 2, that is,

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

where $\nu_t \sim$ i.i.d. $\mathcal{N}(0,1)$. The h-step-ahead forecast for the conditional variance is defined by: $\hat{\sigma}_{n+h}^2 = \mathbb{E}[\sigma_{n+h}^2 | \mathcal{F}_n]$. Here $\mathcal{F}_n$ contains all information (conditional variances and residuals) up to time n.

(a) Give a general expression for the h-step-ahead forecast for the conditional variance following a GARCH(m, s) model, in terms of known values at time n.

(b) Denote the conditional information matrix by

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

and show that $\mathcal{I}_t^{-1} = 2\sigma_t^4$. Hint: use the information matrix equality.

(c) Under information matrix scaling and Gaussian errors, the GAS(1,1) model boils down to the following updating equation for the time-varying parameter $f_t = \sigma_t^2$

$$
\sigma_{t+1}^2 = \tilde{\alpha}_0 + \tilde{\alpha}_1 \mathcal{I}_t^{-1} \frac{\partial \log f(y_t|\mathcal{F}_{t-1})}{\partial \sigma_t^2} + \tilde{\beta}_1 \sigma_t^2.
$$

Show that this updating equation for GAS(1,1) is equivalent to the GARCH(1,1) updating equation for σ_t^2.

In the following sub-questions, you are taken through some steps to show that model averaging can lead to reduced variance and MSE. Model averaging means combining forecasts from different estimated models. Suppose we have two different one-step ahead forecasts, $\hat{y}_{1,t+1}$ and $\hat{y}_{2,t+1}$. We can combine these forecasts, to obtain a new one:

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

Now, assume that the two individual forecasts have forecast errors $e_{1,t+1}$ and $e_{2,t+1}$ and both forecasts are unbiased, so $\mathbb{E}[e_{1,t+1}] = \mathbb{E}[e_{2,t+1}] = 0$. Furthermore, the prediction errors have variances of σ_1^2 and σ_2^2, respectively, and covariance σ_{12}.

(d) Show that the combined forecast $\hat{y}_{t+1}^{MA}$ is unbiased for any value of ω.

(e) Derive the optimal weights ω and $(1-\omega)$ that minimize the variance of the model averaging forecast error (and hence the forecast MSE):

$$\sigma_{MA}^2 = \mathbb{E}[(y_{t+1} - \hat{y}_{t+1}^{MA})^2]$$

and show that the optimal ω can be written as:

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

(f) Show that the individual forecast with lower prediction error variance gets a larger weight in the optimal model averaging forecast.

(g) Show by explicit calculation that the variance of the model averaging forecast error is at most as high as either of the individual forecast error variances, under the optimal weights:

$$\sigma_{MA}^2(\omega^*) \leq \min(\sigma_1^2, \sigma_2^2)$$

[Hint: first show that the variance of the model averaging forecast error can be written as $\frac{\sigma_1^2\sigma_2^2 - \sigma_{12}^2}{\sigma_1^2+\sigma_2^2-2\sigma_{12}} = \frac{\sigma_1^2\sigma_2^2(1-\rho_{12}^2)}{\sigma_1^2+\sigma_2^2-2\sigma_1\sigma_2\rho_{12}}$].

This concept can also be extended to a combination of more forecasts. Define a model averaging forecast as: $\hat{y}_{t+1}^{MA} = \sum_{i=1}^{N} \omega_i \hat{y}_{i,t+1}$, under the restriction $\sum_{i=1}^{N} \omega_i = \omega' \iota = 1$. Furthermore, assume a covariance matrix $\mathbf{\Sigma_e}$ of all forecast errors $e_{i,t+1} = y_{t+1} - \hat{y}_{i,t+1}$.

(h) Derive an expression for the optimal weights of the model averaging forecast $\omega^* = (\omega_1^*, ..., \omega_N^*)'$, that minimizes the variance of the combined forecast error $e_{t+1}^{MA} : \omega' \mathbf{\Sigma_e} \omega$ under the restriction $\omega' \iota = 1$.

Part 2. In this part, you are asked to download data yourself again and make estimates and forecasts, assuming a GARCH(1,1) model for the conditional variance. Unless mentioned otherwise, it is not allowed to use special packages for estimating and forecasting GARCH models, so you should write your own functions and code.

(a) Download daily data from the website of Yahoo Finance for the Amsterdam Exchange Index (ticker: ˆAEX) from February 1, 2000, up to and including January 31, 2024. Save as a CSV file and enable your script to read the data. Use the adjusted closing prices as our time series. You are allowed to use a package for downloading the data.

(b) Calculate the daily log-returns and plot the log-returns and the squared log-returns over time. What can you say about the series' volatility based on this plot (is volatility constant over time or is there heteroskedasticy of some sort)?

The last 20 data points are not used for estimating, but for comparing to our forecasts later. Split your data points into a set containing these last 20 values and an estimation set, containing all other data points.

(c) Estimate the model using the Maximum Likelihood. Just as in Q2b, you may assume that $\mu = 0$. As starting values, you may use the sample variance of the log-returns over the whole estimation sample as the first conditional variance and you may assume $\varepsilon_t = 0$ as starting value for the error term. Interpret the results. What do the estimated coefficients imply for the conditional variance based on information of earlier returns?

(d) Test the model assumption of normally distributed errors. In case this assumption is rejected, what other distribution for the error terms can you think of?

(e) Now, estimate the model using the "rugarch" package in R (so now you are allowed to use functions defined in R to estimate the GARCH model). Compare these results to the results from MLE obtained in (c).
Also, plot the fitted conditional variances and the squared residuals over time. Comment on the quality of the fit.

(f) Use the estimated parameters from (c) to make forecasts for the conditional variance up to 20 days forward. Note: for constructing the forecasts, you must write your own code again! Plot the forecasts and use them in your report. Also plot the actual squared returns and compare them to the conditional variance forecasts.

(g) Repeat all steps of part 4 with daily data from September 17, 2014, up to and including January 20, 2023 for Bitcoin (ticker: BTC-USD). You can re-use a large part of your code and functions, only the data and interpretations are different. Compare the results for the GARCH model between the AEX and Bitcoin. What are the most important differences for the conditional variances?

