Time Series Analysis: Assignment 1

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

Part 1

a.

Consider the AR(p) model:

$$y_t = \phi_0 + \phi_1 y_{t-1} + \ldots + \phi_p y_{t-p} + \varepsilon_t, \quad t = 1, \ldots T$$

for some given initial values $y_0, \ldots y_{-p+1}$ and $\varepsilon_t \overset{iid}{\sim} \mathcal{N}\left(0, \sigma^2\right)$.

We can rewrite the model in terms of $x_t = (1, y_{t-1}, ..., y_{t-p})'$, and a vector of coefficients $\phi = (\phi_0, ...\phi_p)$, such that the model becomes:

$$y_t = x_t'\phi + \varepsilon_t$$

The OLS estimator can be found by minimizing the sum of squared residuals:

$$min \sum_{t=1}^{T} (y_t - x_t'\phi)^2 = min \sum_{t=1}^{T} (y_t^2 - 2y_t x_t'\phi + (x_t'\phi)^2)$$

Follow the first order condition:

$$\frac{\partial}{\partial \phi} \sum_{t=1}^{T} (y_t^2 - 2y_t x_t'\phi + (x_t'\phi)^2) = \sum_{t=1}^{T} (-2y_t x_t + 2x_t x_t'\phi) = 0$$

$$\Rightarrow \sum_{t=1}^{T} y_t x_t = \sum_{t=1}^{T} x_t x_t'\phi \ \Rightarrow \hat{\phi}_{\text{OLS}} = (\sum_{t=1}^{T} x_t x_t')^{-1} \sum_{t=1}^{T} x_t y_t$$

Check the second order condition for sufficiency:

$$\frac{\partial}{\partial \phi'} \sum_{t=1}^{T} (-2y_t x_t + 2x_t x_t'\phi) = 2 \sum_{t=1}^{T} x_t x_t'$$

We verify that the matrix $x_t x_t'$ is positive definite since the matrix $X := \begin{pmatrix} 1 & y_{T-1} & \cdots & y_{T-p} \\ 1 & y_{T-2} & \cdots & y_{T-p-1} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & y_0 & \cdots & y_{-p+1} \end{pmatrix}$ has full column rank. Hence, we indeed reached the minimum, with the OLS estimator:

$$\hat{\phi}_{\text{OLS}} = (\sum_{t=1}^{T} x_t x_t')^{-1} \sum_{t=1}^{T} x_t y_t$$

b.

The AR(p) model in terms of matrix $X := \begin{pmatrix} 1 & y_{T-1} & \cdots & y_{T-p} \\ 1 & y_{T-2} & \cdots & y_{T-p-1} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & y_0 & \cdots & y_{-p+1} \end{pmatrix}$, vector $y = (y_T, y_{T-1}, ..., y_1)'$, and vector $\varepsilon = (\varepsilon_T, \varepsilon_{T-1}, ..., \varepsilon_1)'$ can be rewritten as:

$$y = X\phi + \varepsilon$$

The OLS estimator can be found by minimizing the sum o squared residuals:

$$\min\ e'e = \min(y - X\phi)'(y - X\phi) = \min y'y - 2y'X\phi + (X\phi)'(X\phi)$$

Follow the first order condition:

$$\frac{\partial}{\partial \phi} y'y - 2y'X\phi + (X\phi)'X\phi = 0 \Rightarrow -2X'y + 2X'X\phi = 0 \leftarrow \text{Here we used } \frac{\partial f'g}{\partial x} = \frac{\partial f'}{\partial x} g + \frac{\partial g'}{\partial x} f$$

$$\Rightarrow X'y = X'X\phi \Rightarrow \hat{\phi}_{\text{OLS}} = (X'X)^{-1}X'y$$

Check the second order condition:

$$\frac{\partial(-2X'y + 2X'X\phi)}{\partial \phi'} = 2X'X$$

We again verify that the matrix $X\prime X$ is positive definite since the matrix X has full column rank. Hence, we reached the minimum, with the OLS estimator:

$$\hat{\phi}_{\text{OLS}} = (X'X)^{-1}X'y$$

We can rewrite the OLS estimator in terms of ε and compute the expectation:

$$\hat{\phi}_{\text{OLS}} = (X'X)^{-1} X'(X\phi + \varepsilon) = \phi + (X'X)^{-1} X'\varepsilon$$

We know $\varepsilon_t \overset{iid}{\sim} \mathcal{N}\left(0, \sigma^2\right)$, therefore $\varepsilon \sim N\left(0, \sigma^2 I_n\right)$:

$$E\left[\hat{\phi}_{\text{OLS}}\right] = \phi + E\left[(X'X)^{-1} X\varepsilon\right] = \phi + E_X\left[E\left[(X'X)^{-1} X'\varepsilon \mid X\right]\right] = \phi + E_X\left[(X'X)^{-1} X'E[\varepsilon \mid X]\right]$$

Since $E[\varepsilon \mid X] = 0$, we have $E\left[\hat{\phi}_{\text{OLS}}\right] = \phi$. Hence, the OLS estimator is unbiased.

c.

The definition of conditional density function gives: $f(x \mid y) = \frac{f(x,y)}{f(y)}$. Apply the formula to $f(y)$ recursively we have:

$$\begin{aligned} f(y) &= f(y_T, y_{T-1}, ..., y_1) = f(y_T \mid y_{T-1}, y_{T-2}, ..., y_1) f(y_{T-1}, y_{T-2}, ..., y_1) \\ &= f(y_T \mid y_{T-1}, ..., y_1) f(y_{T-1} \mid y_{T-2}, ..., y_1) f(y_{T-2}, y_{T-3} ... y_1) \\ &= ... \\ &= f(y_T \mid y_{T-1}, ..., y_1) f(y_{T-1} \mid y_{T-2}, ..., y_1) ... f(y_3 \mid y_2, y_1) f(y_2 \mid y_1) f(y_1) \end{aligned}$$

d.

By definition: $\mathcal{F}_t = \{y_s \mid s \leq t\}$; $\mathcal{F}_0 = \{y_{-p+1}, \ldots, y_0\}$

From part c we have: $f(y) = \prod_{t=2}^T f\left(y_t \mid y_{t-1}, \ldots, y_1\right) f\left(y_1\right) = \prod_{t=2}^T f\left(y_t \mid \mathcal{F}_{t-1}\right) f\left(y_1\right)$

Also apply the formula of conditional probability to $f(y_1)$: $f\left(y_1\right) = f\left(y_1 \mid \mathcal{F}_0\right) f\left(\mathcal{F}_0\right)$, hence:

$$f\left(y \mid \mathcal{F}_0\right) = \frac{f\left(y, \mathcal{F}_0\right)}{f\left(\mathcal{F}_0\right)} = \frac{\prod_{t=2}^T f\left(y_t \mid \mathcal{F}_{t-1}\right) f\left(y_1 \mid \mathcal{F}_0\right) f\left(\mathcal{F}_0\right)}{f\left(\mathcal{F}_0\right)} \Rightarrow f\left(y \mid \mathcal{F}_0\right) = \prod_{t=1}^T f\left(y_t \mid \mathcal{F}_{t-1}\right)$$

Based on this result, since $\varepsilon_t \overset{iid}{\sim} \mathcal{N}\left(0, \sigma^2\right)$, we have $y_t \mid F_{t-1} \sim \mathcal{N}\left(x_t'\phi, \sigma^2\right)$

The probability density function of $y_t \mid F_{t-1}$ is: $f\left(y_t \mid F_{t-1}\right)=\frac{1}{\sigma \sqrt{2 \pi}} e^{-\frac{1}{2}\left(\frac{y_t-x_t^{\prime} \phi}{\sigma}\right)^2}$

We can derive the log-likelihood function:

$$
\begin{aligned}
l\left(\phi, \sigma^2\right) & =\sum_{t=1}^T \log \left(f\left(y_t \mid \mathcal{F}_{t-1}\right)\right)=\sum_{t=1}^T \log \left(\frac{1}{\sigma \sqrt{2 \pi}} e^{-\frac{1}{2}\left(\frac{y_t-x_t^{\prime} \phi}{\sigma}\right)^2}\right)=\sum_{t=1}^T\left[\log \left(\frac{1}{\sigma \sqrt{2 \pi}}\right)-\frac{1}{2}\left(\frac{y_t-x_t^{\prime} \phi}{\sigma}\right)^2\right] \\
& =-\frac{T}{2} \log \left(2 \pi \sigma^2\right)-\frac{1}{2 \sigma^2} \sum_{t=1}^T\left(y_t-x_t^{\prime} \phi\right)^2
\end{aligned}
$$

To maximize $l\left(\phi, \sigma^2\right)$, by the first order condition:

$$
\frac{\partial l\left(\phi, \sigma^2\right)}{\partial \phi}=0 \Rightarrow-\frac{1}{2 \sigma^2} \sum_{t=1}^T 2\left(y_t-x_t^{\prime} \phi\right)\left(-x_t\right)=0
$$

$$
\Rightarrow \sum_{t=1}^T y_t x_t-\sum_{t=1}^T x_t x_t^{\prime} \phi=0 \Rightarrow \hat{\phi}_{\mathrm{MLE}}=\left(\sum_{t=1}^T x_t x_t^{\prime}\right)^{-1} \sum_{t=1}^T x_t
$$

Under normally distributed error terms, the OLS and MLE are equivalent:

$$
\hat{\phi}_{\mathrm{OLS}}=\hat{\phi}_{\mathrm{MLE}}=\left(\sum_{t=1}^T x_t x_t^{\prime}\right)^{-1} \sum_{t=1}^T x_t
$$

e.

We can rewrite the new model as follows to obtain the white noise, where $y_t \sim \mathrm{AR(p)}$:

$$
\gamma_t=z_t^{\prime} \beta+y_t \Rightarrow \gamma_t=z_t^{\prime} \beta+x_t^{\prime} \phi+\varepsilon_t \Rightarrow \gamma_t-x_t^{\prime} \phi=z_t^{\prime} \beta+\varepsilon_t
$$

Least-Square solves:

$$
\min \sum_{t=1}^T\left(\gamma_t-x_t^{\prime} \phi-z_t^{\prime} \beta\right)^2
$$

F.O.C:

$$
\frac{\partial}{\partial \beta} \sum_{t=1}^T\left(\gamma_t-x_t^{\prime} \phi-z_t^{\prime} \beta\right)^2=\sum_{t=1}^T 2\left(\gamma_t-x_t^{\prime} \phi-z_t^{\prime} \beta\right) \cdot\left(-z_t\right)=0
$$

$$
\Rightarrow \sum_{t=1}^T\left(\gamma_t-x_t^{\prime} \phi-z_t^{\prime} \beta\right) z_t=0 \Rightarrow \sum_{t=1}^T\left(\gamma_t z_t-z_t x_t^{\prime} \phi\right)=\sum_{t=1}^T z_t z_t^{\prime} \beta
$$

$$
\Rightarrow \hat{\beta}=\left(\sum_{t=1}^T z_t z_t^{\prime}\right)^{-1} \sum_{t=1}^T\left(z_t \gamma_t-z_t x_t^{\prime} \phi\right)
$$

Since we have derived the $\hat{\phi}_{\mathrm{OLS}}$ in previous questions, $\beta(\hat{\phi}_{\mathrm{OLS}})=\left(\sum_{t=1}^T z_t z_t^{\prime}\right)^{-1} \sum_{t=1}^T\left(z_t \gamma_t-z_t x_t^{\prime} \hat{\phi}_{\mathrm{OLS}}\right)$
In conclusion, we advocate for the adoption of the Feasible Generalized Least Squares (FGLS) estimator as opposed to the Ordinary Least Squares (OLS) estimator in the present context. The rationale behind this preference lies in the nonconformity of the model to the assumptions of the standard framework. Consequently, assessments such as tests and confidence intervals relying on OLS are rendered inconclusive due to the deviation from the underlying assumptions.

f.

$$y_t = \phi_0 - 1.3y_{t-2} - 0.4y_{t-4} + \varepsilon_t$$
$$\phi(L) = 1 + 1.3L^2 + 0.4L^4, \phi(L)y_t = \varepsilon_t$$

AR roots:

$$\phi(Z) = 1 + 1.3Z^2 + 0.4Z^4Z$$
$$Z_{1,2}^2 = \frac{-1.3 \pm \sqrt{1.3^2 - 1.6}}{0.8}$$
$$Z_1^2 = -1.25, \quad Z_2^2 = -2$$
$$Z_{1,1} = \frac{\sqrt{5}i}{2}, \quad Z_{1,2} = \frac{-\sqrt{5}i}{2}, \quad Z_{2,1} = \sqrt{2}i, \quad Z_{2,2} = -\sqrt{2}i$$

All roots are outside the unit circle in $\mathbb{C}$, hence $\{y_t\}$ is stationary.

g.

$$y_t = \alpha + \phi y_{t-1} + \varepsilon_t$$

For AR(1) process, the ACF is $\rho_k = \frac{\gamma_k}{\gamma_0}$,

$$\gamma_k = E[(y_t - \mu)(y_{t-k} - \mu)] = \frac{\phi^k \sigma^2}{1 - \phi^2}, \quad \gamma_0 = \frac{\sigma^2}{1 - \phi^2}, \quad \rho_k = \frac{\gamma_k}{\gamma_0} = \phi^k$$

We know that the correlation coefficient satisfies $|\rho| \leq 1$. Consequently, this condition must also hold true for the parameter $|\phi|$ in an AR(1) process: $|\phi| \leq 1$. Additionally, the requirement of a constant mean imposes the constraint $|\phi| \neq 1$. Thus, it follows that $|\phi| < 1$. As the result of this condition, ρ_k goes to 0, when k goes to infinity, which suggests a stationary process. Therefore, based on the behavior of the autocorrelation function (ACF), $|\phi| < 1$ is the condition for a stationary AR(1) process.

h.

$$\begin{aligned} y_t = 0.6y_{t-1} - 0.25y_{t-2} + \varepsilon_t, \quad \gamma_k = cov(y_t, y_{t-k}) &= cov((0.6y_{t-1} - 0.25y_{t-2} + \varepsilon_t), y_{t-k}) \\ &= 0.6cov(y_{t-1}, y_{t-k}) - 0.25cov(y_{t-2}, y_{t-k}) + cov(\varepsilon_t, y_{t-k}) \end{aligned}$$

Since

$$cov(\varepsilon_t, y_{t-k}) = E[\varepsilon_t \cdot y_{t-k}] = E_x[E[\varepsilon_t \cdot y_{t-k}|\mathcal{F}_{t-1}]] = E_x[y_{t-k} \underbrace{E[\varepsilon_t|\mathcal{F}_{t-1}]}_{=0}] = 0$$

Therefore,

$$\gamma_k = 0.6\gamma k - 1 - 0.25\gamma_{k-2} \to \rho_k = 0.6\rho_{k-1} - 0.25\rho_{k-2}, \text{ with } \rho_0 = 1$$

then

$$\rho_1 = 0.6\rho_0 - 0.25\rho{-1} \xrightarrow{\rho_1 = \rho_{-1}} \rho_1 = 0.48$$

In order to make an analytical solvable system,

$$\begin{pmatrix} \rho_k \\ \rho_{k-1} \end{pmatrix} = \begin{pmatrix} 0.6 & -0.25 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} \rho_k \\ \rho_{k-1} \end{pmatrix} \overset{\rho_1 = \rho_{-1}}{=} \underbrace{\begin{pmatrix} 0.6 & -0.25 \\ 1 & 0 \end{pmatrix}^k}_{:=A} \begin{pmatrix} \rho_0 \\ \rho_1 \end{pmatrix}$$

$$det(A - \lambda) = 0 \to det \begin{pmatrix} 0.6 - \lambda & -0.25 \\ 1 & -\lambda \end{pmatrix} = 0 \to \lambda^2 - 0.6\lambda + 0.25 = 0 \to \lambda = 0.3 \pm 0.4i$$
$$\rho_k = (a + bi)(0.3 + 0.4i)^k + (a - bi)(0.3 - 0.4i)^k = 2Re((a + bi)(0.3 + 0.4i)^k)$$

We already know that $\rho_0 = 1$ and $\rho_1 = 0.48$, hence a and b can be solved.

$$\begin{cases} \rho_0 = 1 \\ \rho_1 = 0.48 \end{cases} \to \begin{cases} 2a = 1 \\ 2 \cdot (0.3a - 0.4b) = 0.48 \end{cases} \to \rho_k = 2Re((0.5 - 0.225i)(0.3 + 0.4i)^k)$$

$$= 2Re(se^{i\varphi}(\gamma e^{i\omega})^k) = 2Re(\gamma^k se^{i(\varphi+\omega k)}) = 2r^k scos(\varphi + \omega k)$$

where

$$s = \sqrt{0.5^2 + 0.225^2} \approx 0.55, \gamma = \sqrt{0.3^2 + 0.4^2} = 0.5, \phi = \arctan(\frac{-9}{20}) =\approx -0.42, \omega = \arctan(\frac{4}{3}) \approx 0.93$$

We find that

$$\rho_k \approx 1.10 \cdot 0.5^k cos(-0.42 + 0.93k)$$

i.

Given ARMA$(1,1)$, $y_t = \alpha + \phi y_{t-1} + \varepsilon_t + \theta\varepsilon_{t-1}$, where $|\phi| < 1$ and $|\theta| < 1$, $\varepsilon_t \overset{iid}{\sim} \mathcal{N}\left(0, \sigma^2\right)$

We can rewritten the model using the lag operator:

$$(1 - \phi L)y_t = \alpha + (1 + \theta L)\varepsilon_t \Rightarrow \varepsilon_t = \frac{(1 - \phi L)y_t - \alpha}{1 + \theta L}$$

By the definition of the h-step-ahead forecast: $\hat{y}_{n+h} = E\left[y_{n+h} \mid Y_n\right]$, we can derive the first prediction:

$$\begin{aligned}
\hat{y}_{n+1} &= E\left(y_{n+1} \mid Y_n\right) = E\left(\alpha + \phi y_n + \varepsilon_{n+1} + \theta\varepsilon_n \mid Y_n\right) \\
&= \alpha + E\left(\phi y_n \mid Y_n\right) + E\left(\varepsilon_{n+1} \mid Y_n\right) + E\left(\theta\varepsilon_n \mid Y_n\right) \\
&= \alpha + \phi y_n + 0 + \theta\varepsilon_n = \alpha + \phi y_n + \theta\frac{(1 - \phi L)y_n - \alpha}{1 + \theta L} \\
&= \frac{1 + \theta L - \theta}{1 + \theta L}\alpha + \frac{\phi + \theta}{1 + \theta L}y_n
\end{aligned}$$

Similarly,

$$\hat{y}_{n+2} = E\left[y_{n+2} \mid Y_n\right] = E\left[\alpha + \phi y_{n+1} + \varepsilon_{n+2} + \theta\varepsilon_{n+1} \mid Y_n\right] = \alpha + \phi\hat{y}_{n+1}$$
$$\hat{y}_{n+3} = \alpha + \phi\hat{y}_{n+2} = \alpha + \phi\left(\alpha + \phi\hat{y}_{n+1}\right) = \alpha(1 + \phi) + \phi^2\hat{y}_{n+1}$$

Therefore in general,

$$\begin{aligned}
\hat{y}_{n+h} &= \alpha\sum_{j=0}^{h-2}\phi^j + \phi^{h-1}\hat{y}_{n+1} \\
&= \alpha\sum_{j=0}^{h-2}\phi^j + \phi^{h-1}\left[\frac{1 + \theta L - \theta}{1 + \theta L}\alpha + \frac{\phi + \theta}{1 + \theta L}y_n\right] \\
&= \alpha\left(\frac{(1 + \theta L - \theta)\phi^{h-1}}{1 + \theta L} + \sum_{j=0}^{h-2}\phi^j\right) + \phi^{h-1}\frac{\phi + \theta}{1 + \theta L}y_n
\end{aligned}$$

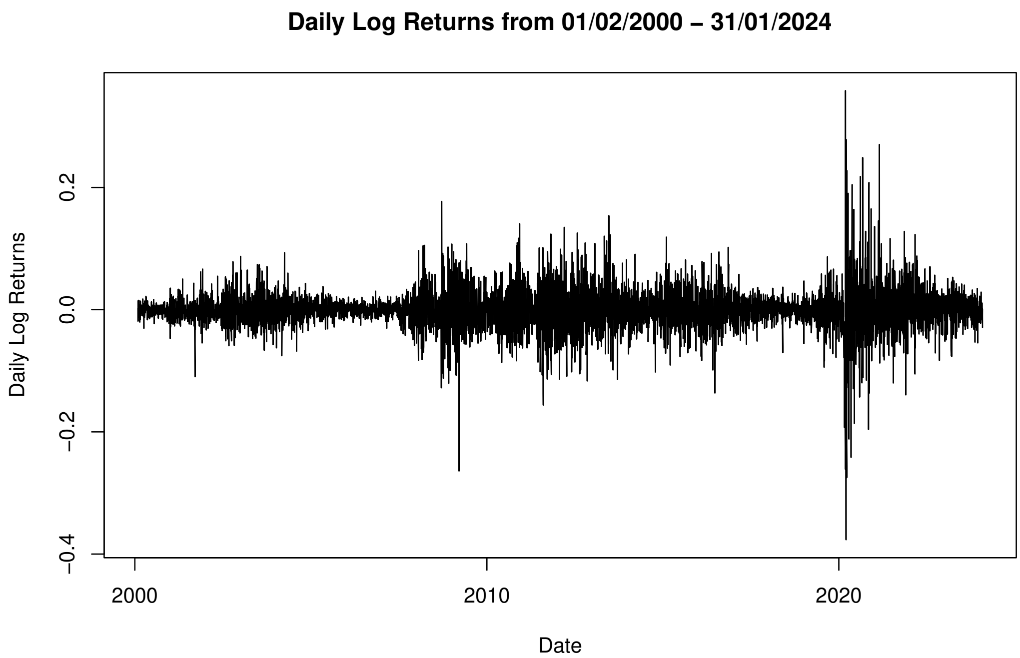
Part 2

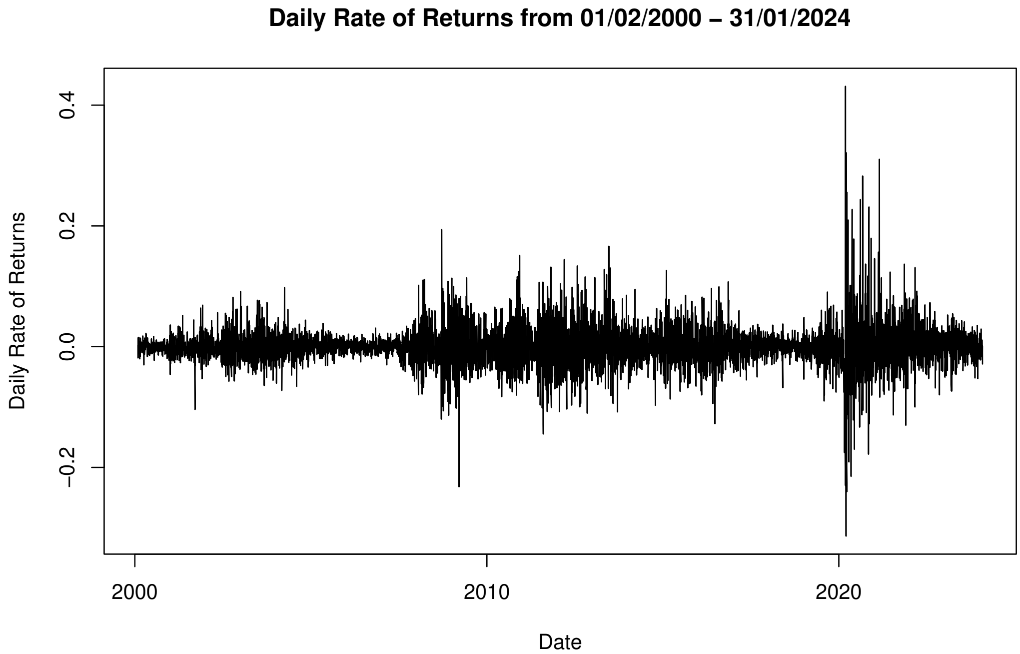
a.

The R script imports the data from the file, extracting 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.

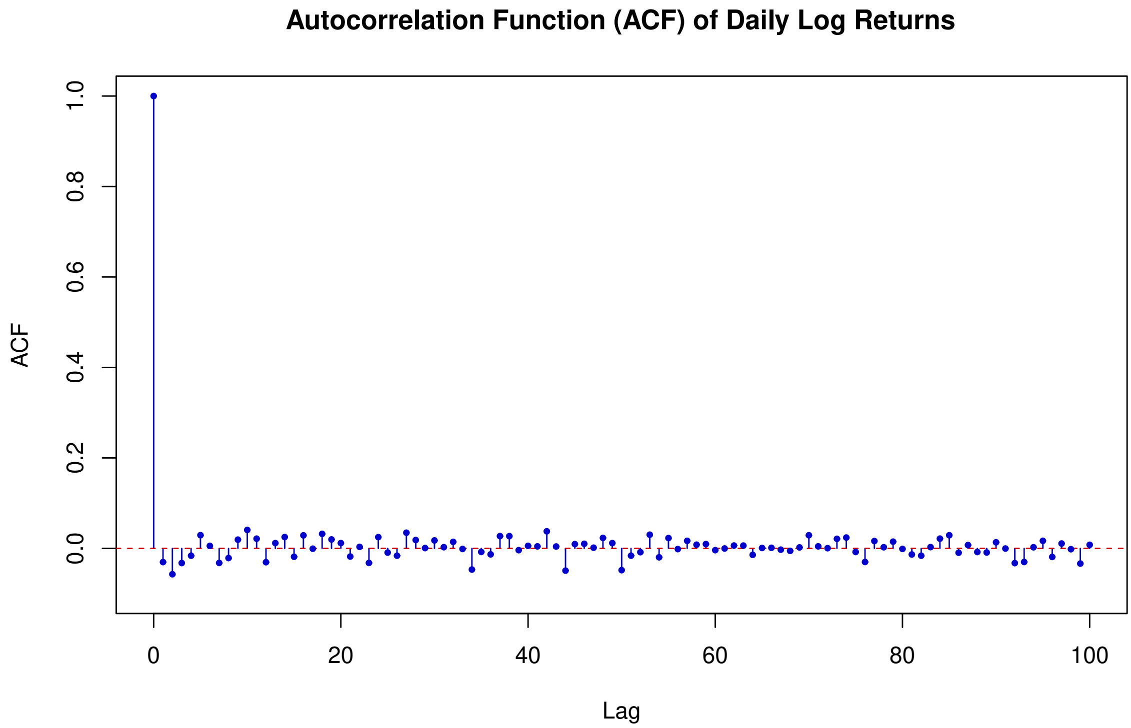
b.

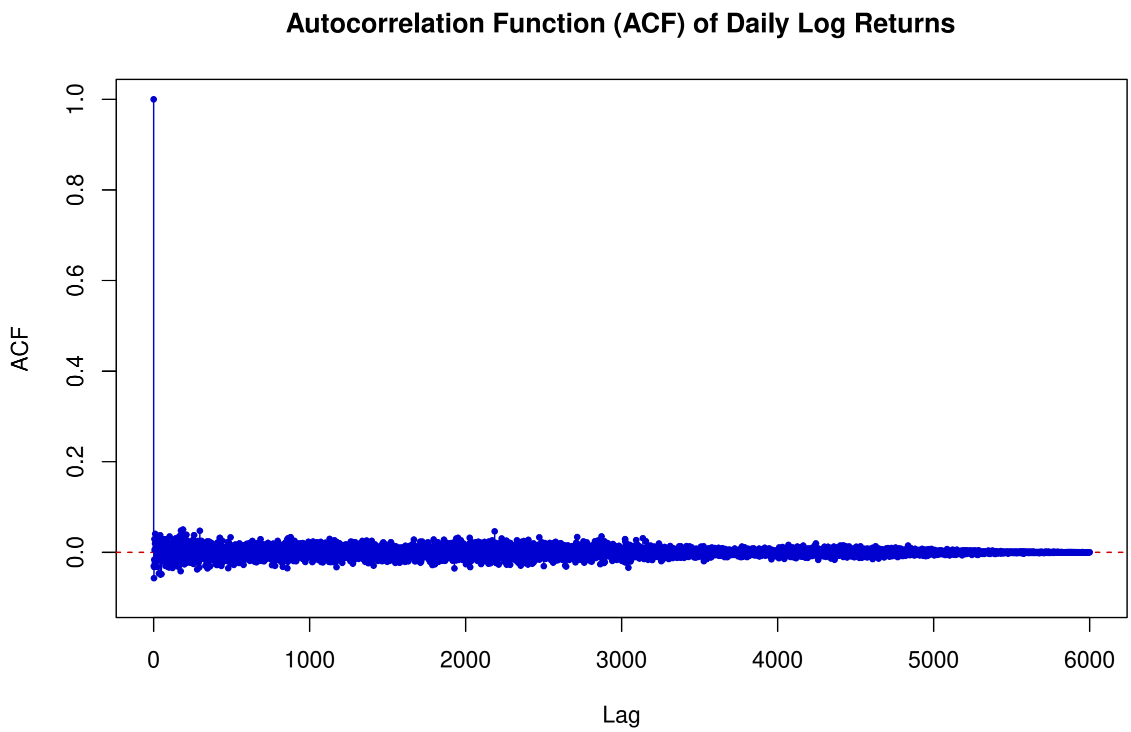
The daily log-return at time t is computed as $\log(\frac{y_t}{y_{t-1}})$, while the daily rate at time t is calculated as $\frac{y_t - y_{t-1}}{y_{t-1}} = \frac{y_t}{y_{t-1}} - 1$. The autocorrelation function (ACF) is determined by $\rho_k = \frac{\mathrm{Cov}(y_t, y_{t-k})}{\sqrt{\mathrm{Var}(y_t)\mathrm{Var}(y_{t-k})}}$.





When examining the daily log returns alongside the adjusted closing price of the 5-Year Treasury Yields, it's evident that the mean remains consistent for the log returns, but not for the closing prices, indicating they are not stationary. The plots depicting daily rate of returns closely resemble those of log returns, indicating a close relationship between the two, as evidenced by the approximation $\log(\frac{y_t}{y_{t-1}}) \approx \frac{y_t - y_{t-1}}{y_{t-1}} = \frac{y_t}{y_{t-1}} - 1$. While the mean of the daily log returns seems stable, the variance displays instability over time, contradicting a crucial requirement for stationarity. However, the range of the daily log returns, approximately (-0.38, 0.35), suggests relatively minor fluctuations. Consequently, we can not draw conclusions regarding the stationarity of the daily log returns process at this stage.





The initial value in the autocorrelation function (ACF) is 1, subsequently, as the number of lag increases , the ACF gradually decreases toward zero. This trend aligns with theoretical expectations of a stationary series and is evident in part 1, exercise h, particularly when observing the limit as the lag approaches infinity. Therefore, based on the behavior the ACF, we conclude that the three conditions of stationarity are loosely satisfied.

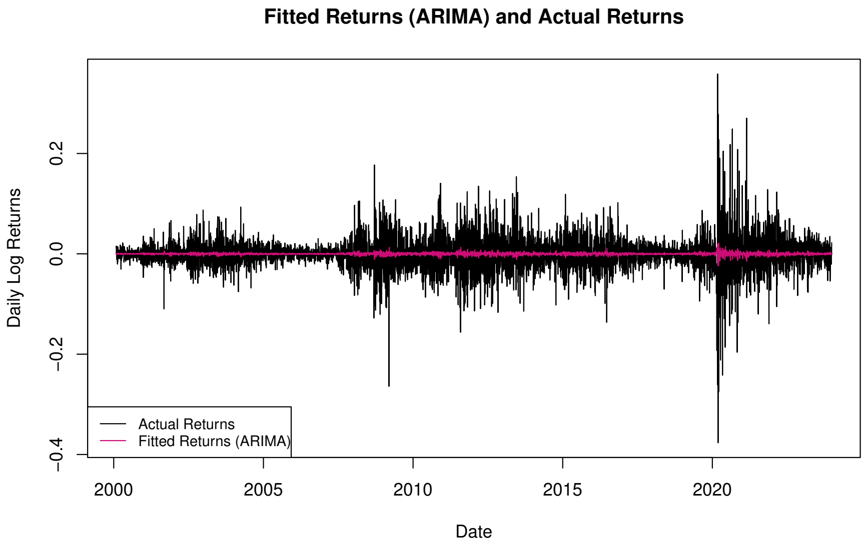
The autocorrelation function (ACF) begins with a value of 1, signifying perfect correlation with itself. As the number of lags increases, the ACF decreases geometrically toward zero. This pattern corresponds with theoretical expectations for a stationary series and is exemplified in question exercise h of part 1. Consequently, judging from the ACF's behavior, we conclude that the three conditions of stationarity are satisfied.

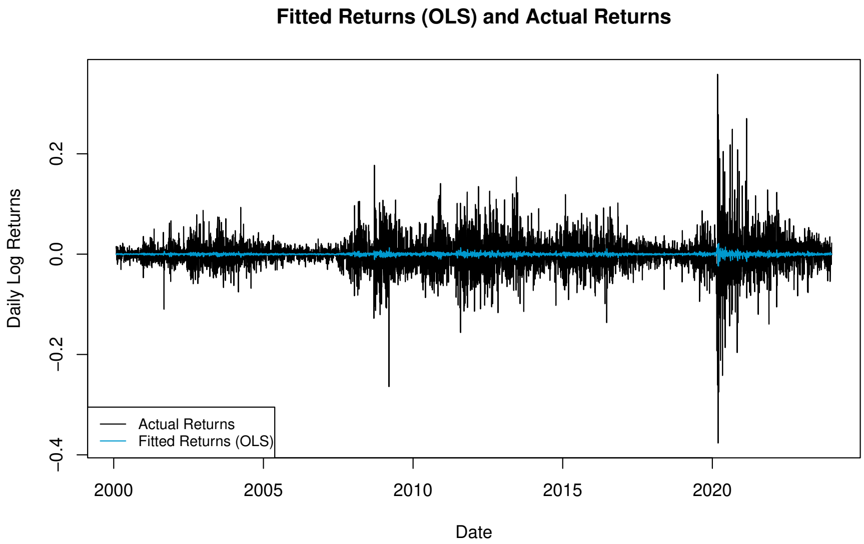
c.

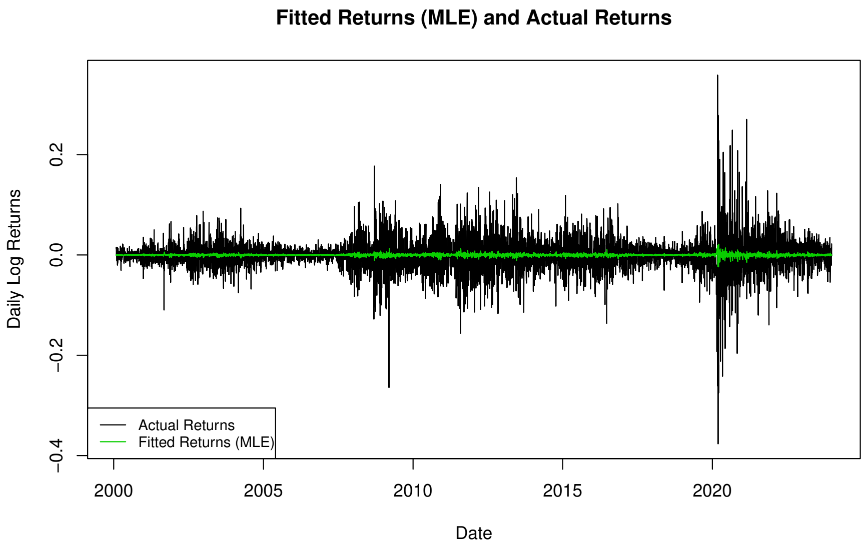
Estimating the AR(2) model using OLS yields $\hat{\phi}_{\text{OLS}} = (-0.0000935067, -0.03200305, -0.05817223)$, while employing MLE results in $\hat{\phi}_{\text{MLE}} = (-0.0000935065 - 0.03200305 - 0.05817223)$ when setting the initial values equal to the first three log-returns. Theoretically, MLE should provide the same estimate as OLS since we assume normally distributed error terms. However, the slight discrepancies observed are likely due to numerical computation errors.

d.

Estimating the AR(2) model using the "arima" function, we have $\hat{\phi}_{\text{ARIMA}} = (-0.0000878354, -0.03200474, -0.05815605)$. Here, the coefficients of the two lag variables closely resemble those obtained from OLS and MLE, while the intercept is slightly larger. Upon comparing the three associated fitted plots, we observe similarities among the results of MLE, OLS, and ARIMA, as the fitted log-returns all center around zero. Additionally the actual values exhibit greater volatility compared to the fitted model.

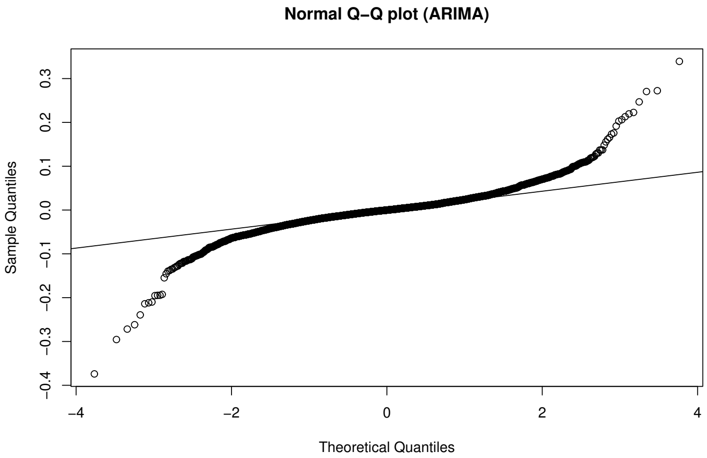

Fitted Returns (ARIMA) and Actual Returns
Daily Log Returns
0.2
0.0
−0.2
−0.4
Actual Returns
Fitted Returns (ARIMA)
2000
2005
2010
2015
2020
Date

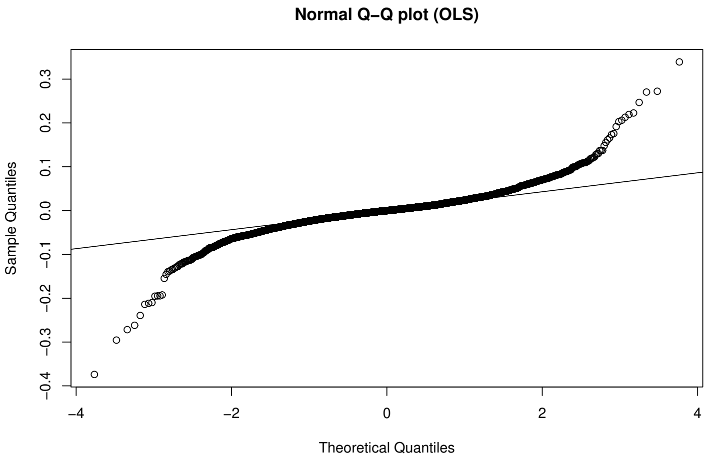

Fitted Returns (OLS) and Actual Returns
Daily Log Returns
0.2
0.0
−0.2
−0.4
Actual Returns
Fitted Returns (OLS)
2000
2005
2010
2015
2020
Date

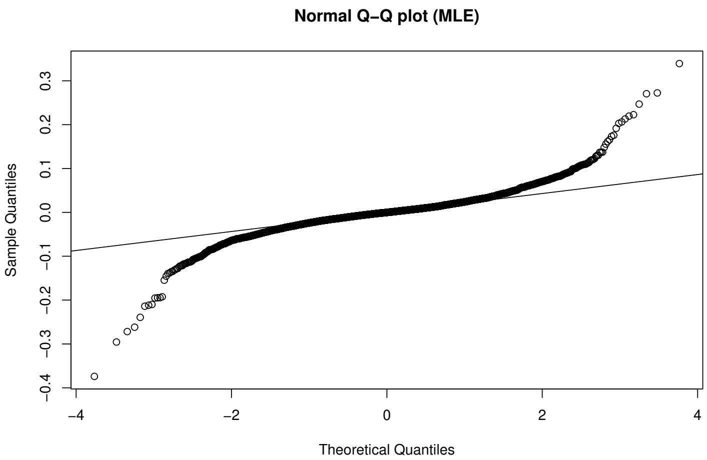

Fitted Returns (MLE) and Actual Returns
Daily Log Returns
0.2
0.0
−0.2
−0.4
Actual Returns
Fitted Returns (MLE)
2000
2005
2010
2015
2020
Date

e.

Based on the results of the Jarque-Bera test, it is evident that the residuals exhibit non-normality. The null hypothesis assumes normality, and with p-values for the residuals of OLS, MLE, and ARIMA all smaller than 2.2e-16, we reject the null hypothesis that the residuals from these methods follow a normal distribution. This conclusion is further supported by examining the associated QQ-plots, where it is apparent that not all residuals align with the straight line, suggesting deviations from normal distribution.







f.

The forecasted returns generated by the ARIMA model are centered around zero and demonstrate comparatively smaller fluctuations when compared to the actual values.