HW7

Siwei Tang

3/23/2022

Q1

(a)

```
amazon = read.csv("AMZN_March16_2020_March16_2022.csv")
google = read.csv("GOOGL_March16_2020_March16_2022.csv")
apple = read.csv("AAPL_March16_2020_March16_2022.csv")
nasdaq = read.csv("Nasdaq_March16_2000_March 16_2022.csv")
```

```
# calculate the closing price
n = dim(nasdaq)[1]
fix_return = 0.025/253
EX_R_nasdaq = nasdaq$Adj.Close[2:n]/ nasdaq$Adj.Close[1:n - 1] - 1 - fix_return
EX_R_apple = apple$Adj.Close[2:n]/apple$Adj.Close[1:n-1] - 1- fix_return
EX_R_amazon = amazon$Adj.Close[2:n]/amazon$Adj.Close[1:n-1] - 1 - fix_return
EX_R_google = google$Adj.Close[2:n]/google$Adj.Close[1:n-1] - 1 - fix_return
```

```
fit_apple = lm(EX_R_apple ~ EX_R_nasdaq)
fit_amazon = lm(EX_R_amazon ~ EX_R_nasdaq)
fit_google = lm(EX_R_google ~ EX_R_nasdaq)
model_names = c("fit_apple","fit_amazon","fit_google")
list_models = lapply(model_names, get)
alpha_list = rep(0,3)
beta_list = rep(0,3)
R2_list = rep(0,3)

# iterate over the fit results
for (i in 1:3){
  fit = list_models[[i]]
  alpha = fit$coefficients[1]
  beta = fit$coefficients[2]
  r = summary(fit)$r.squared
  # store results
  alpha_list[i] = alpha
  beta_list[i] = beta
  R2_list[i] = r
}
res_df = data.frame(alpha = alpha_list,
                    beta = beta_list,
                    precent_mkt_risk = R2_list,
                    row.names = c("apple","amazon","google")
                    )
```

```r
print(res_df)
```

```
##                alpha      beta precent_mkt_risk
## apple  6.296803e-04 1.0851737        0.6689344
## amazon 1.334689e-06 0.9620414        0.5133706
## google 5.987528e-04 0.9623872        0.6379769
```

##(2)

```r
var_F = var(EX_R_nasdaq)
epsilon_mat = diag(c(var(summary(fit_apple)$residuals),
                     var(summary(fit_amazon)$residuals),
                     var(summary(fit_google)$residuals)))
est_cov = as.matrix(beta_list,3,1)%*%var_F%*%as.vector(beta_list) + epsilon_mat
cat("Estimated cov among three stocks, \n")
```

```
## Estimated cov among three stocks,
```

```r
print(est_cov)
```

```
##              [,1]         [,2]         [,3]
## [1,] 0.0004286985 0.0002542319 0.0002543233
## [2,] 0.0002542319 0.0004390293 0.0002254658
## [3,] 0.0002543233 0.0002254658 0.0003535344
```

```r
emp_cov = matrix(rep(0,9),3,3)
a = EX_R_apple
b = EX_R_amazon
c = EX_R_google
emp_cov[1,] = c(var(a), cov(a,b),cov(a,c))
emp_cov[2,] = c(cov(a,b), cov(b,b),cov(b,c))
emp_cov[3,] = c(cov(a,c), cov(c,b),cov(c,c))
cat("Empirical covariance of return \n")
```

```
## Empirical covariance of return
```

```r
print(emp_cov)
```

```
##              [,1]         [,2]         [,3]
## [1,] 0.0004286985 0.0002659781 0.0002432843
## [2,] 0.0002659781 0.0004390293 0.0002323363
## [3,] 0.0002432843 0.0002323363 0.0003535344
```

The estimated result is very close to the empirical,as the estimates 2-3 sigificant figures. The diagnal is exactly the same, which means the estimated variance is the the true variance. # Q2 ## (a)

```r
library(ggplot2)
library(lubridate)
```

```
##
## Attaching package: 'lubridate'

## The following objects are masked from 'package:base':
##
##     date, intersect, setdiff, union
```

```r
library(dplyr)
```

```
##
## Attaching package: 'dplyr'
```

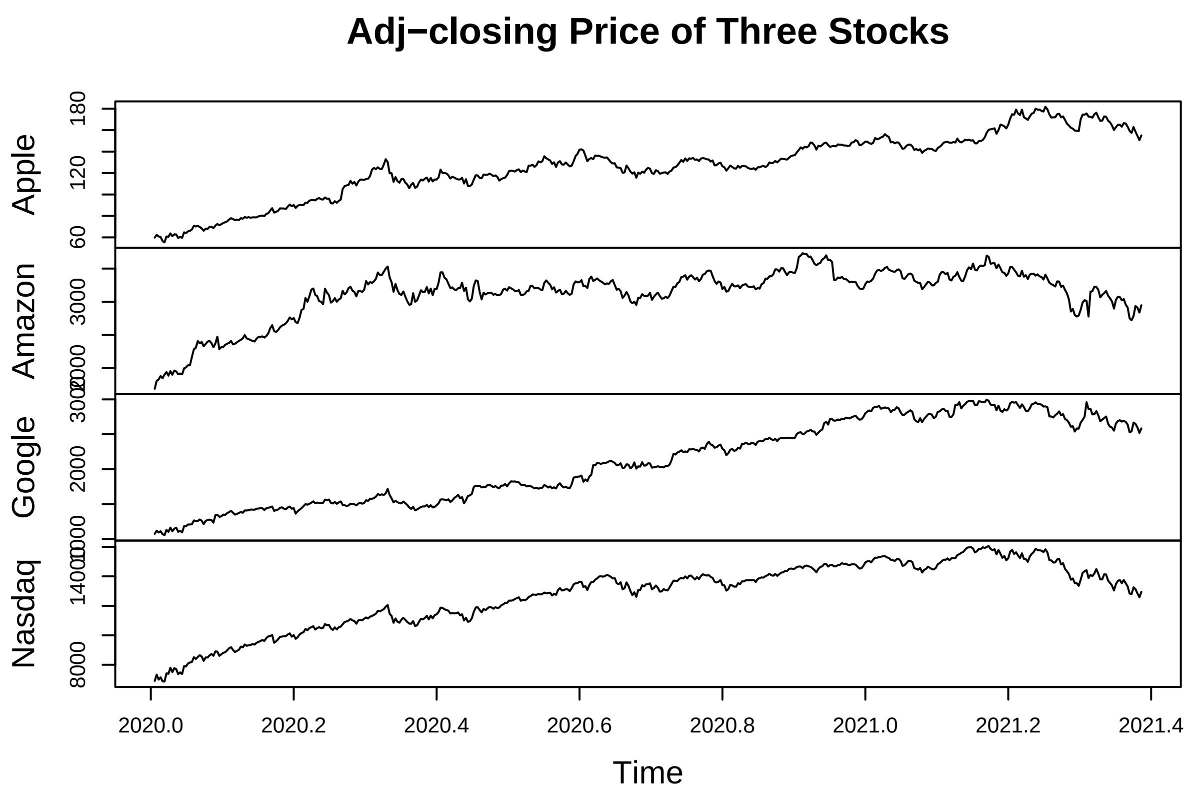
```
## The following objects are masked from 'package:stats':
##
##     filter, lag
```

```
## The following objects are masked from 'package:base':
##
##     intersect, setdiff, setequal, union
```

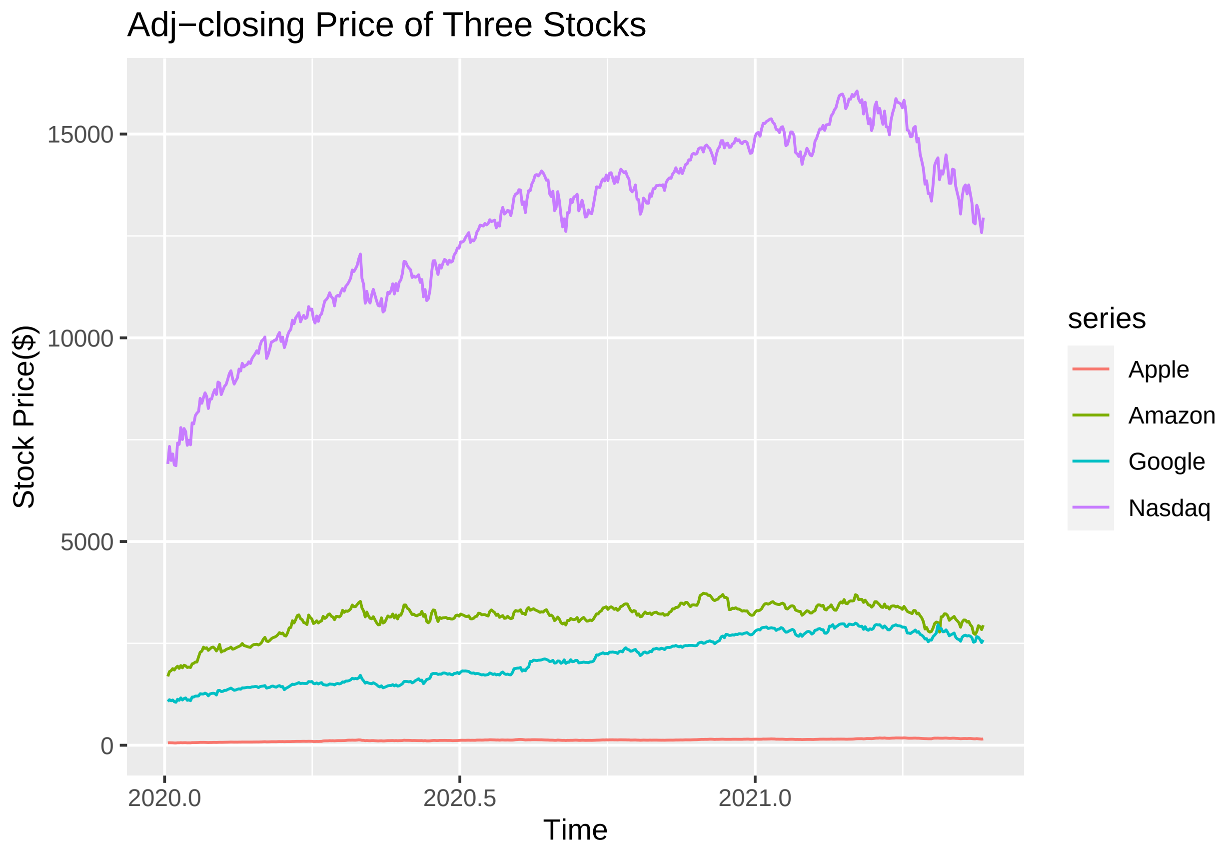
```
library(forecast)
```

```
## Registered S3 method overwritten by 'quantmod':
##   method            from
##   as.zoo.data.frame zoo
```

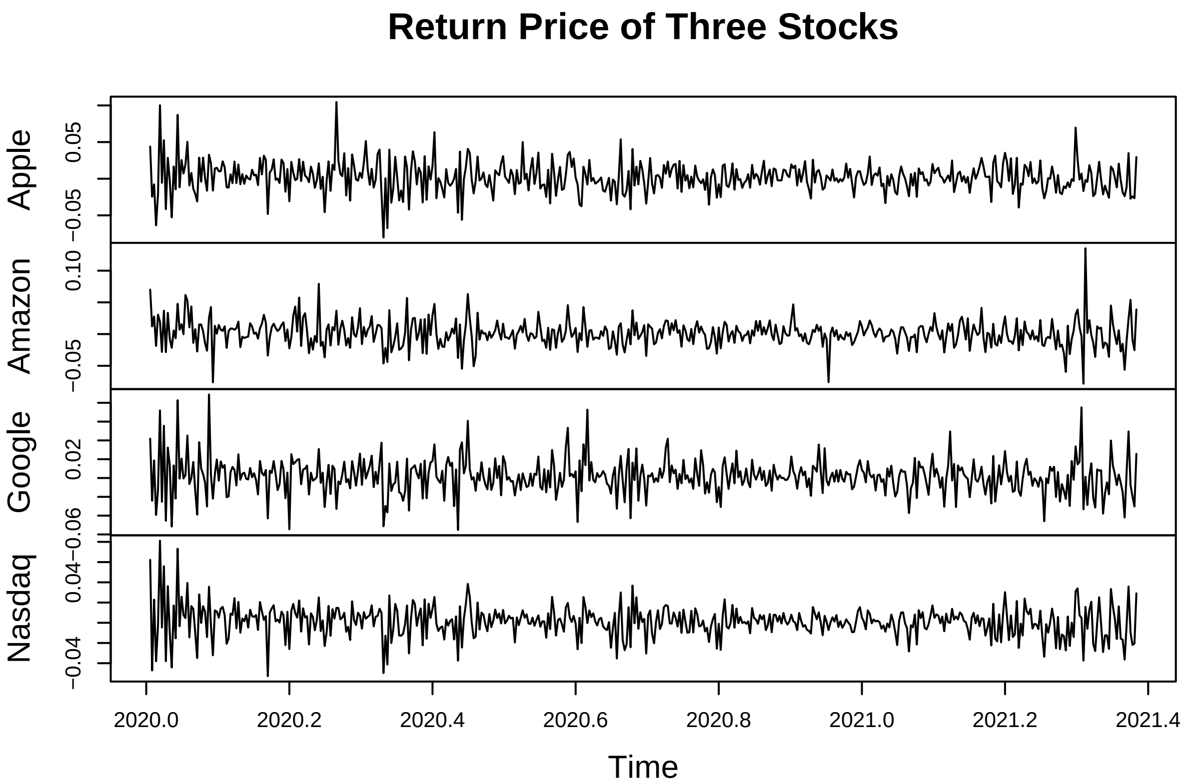
```
data_mat_close = data.frame(Apple = apple$Adj.Close,
                            Amazon = amazon$Adj.Close,
                            Google = google$Adj.Close,
                            Nasdaq = nasdaq$Adj.Close)
close_ts = ts(data = data_mat_close,
              start = c(2020,3,16),
              frequency = 365)
plot(close_ts,
     main = "Adj-closing Price of Three Stocks",  ylab ="Stock Price($)")
```




```
autoplot(close_ts,
     main = "Adj-closing Price of Three Stocks") +  ylab("Stock Price($)")
```




```
data_mat_return = data.frame(Apple = EX_R_apple,Amazon = EX_R_amazon,Google = EX_R_google,Nasdaq = EX_R_
close_ts = ts(data = data_mat_return,
              start = c(2020,3,16),
              frequency = 365)
plot(close_ts,
     main = "Return Price of Three Stocks"
     )
```

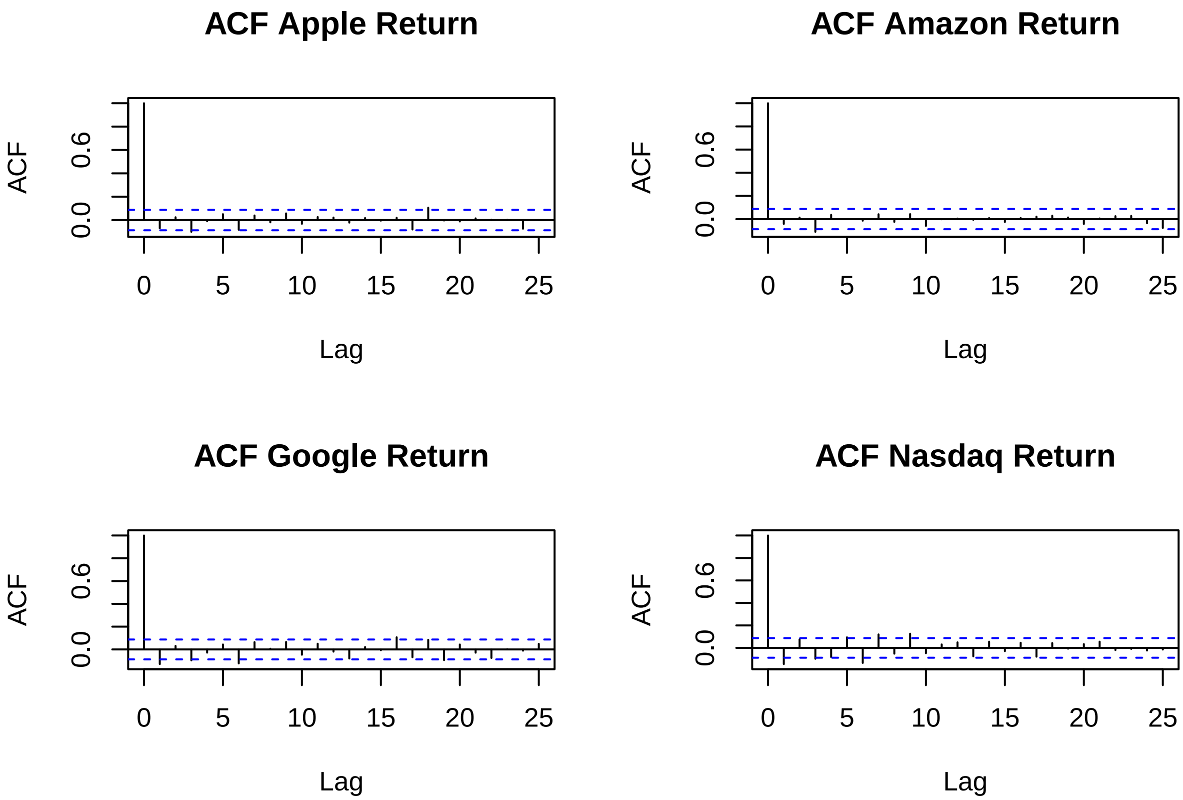


As for returns, they seems to have similar volatility, as their variance are similiar in the same time periods. As for the stock prices, they are quite similiar in the early periods, but differ when t approximates to infinity. ## (b)

```
par(mfrow = c(2,2))
acf(as.vector(data_mat_close$Apple), lag =25, main  = "ACF Apple Adj Closing Price")
acf(as.vector(data_mat_close$Amazon), lag = 25, main = "ACF Amazon Adj Closing Price")
acf(as.vector(data_mat_close$Google), lag = 25, main = "ACF Google Adj Closing Price")
acf(as.vector(data_mat_close$Nasdaq), lag = 25, main = "ACF Nasdaq Adj Closing Price")
```

There is no stationarity for any of the four adj-closing prices, because the auto-correlation doesn't decay to zero when lag<=25.

```
par(mfrow=c(2,2))
acf(as.vector(data_mat_return$Apple), lag =25, main  = "ACF Apple Return")
acf(as.vector(data_mat_return$Amazon), lag = 25, main = "ACF Amazon Return")
acf(as.vector(data_mat_return$Google), lag = 25, main = "ACF Google Return")
acf(as.vector(data_mat_return$Nasdaq), lag = 25, main = "ACF Nasdaq Return")
```



The Returns for individual stocks are generally stationary, though Google's return have some violations when lag = 1,6 and 16. For Nasdaq, the stationarity doesn't maifest itself until lag = 10.

Q3 $\quad Y_n, Z_n \sim AR(1)$ $\qquad$ Let $\mu_Y = \mu$, $\mu_Z = \mu'$

$\alpha_Y = \alpha$, $\alpha_Z = \alpha'$

(a) $Y_n = (1-\alpha)\mu + \alpha Y_{n-1} + \varepsilon_n$

$Z_n = (1-\alpha)\mu' + \alpha' Z_{n-1} + \varepsilon'_n$

$\mu_X = E(X_n) = E(Y_n - Z_n) = \alpha E Y_{n-1} - \alpha\mu + \mu - \alpha' E Z_{n-1} + \alpha'\mu' - \mu'$

$\because Y_n, Z_n \sim AR(1)$

$\therefore E Y_{n-1} = \mu, \quad E Z_{n-1} = \mu'$

$\mu_X = \cancel{\alpha\mu} - \cancel{\alpha\mu} + \mu - \cancel{\alpha'\mu'} + \cancel{\alpha'\mu'} - \mu'$

$= \mu_Y - \mu_Z$

$\Sigma_X(m,n) = \mathrm{cov}(Y_m - Z_m, Y_n - Z_n)$

$= \mathrm{cov}(Y_m, Y_n) - \mathrm{cov}(Y_n, Z_m) - \mathrm{cov}(Y_m, Z_n) + \mathrm{cov}(Z_m, Z_n)$

$\because Y_n, Z_n \sim AR(1)$

$\therefore \gamma_Y(t) = \mathrm{cov}(Y_m, Y_{m+t}) = \dfrac{\sigma_Y^2 \alpha_Y^{|t|}}{1-\alpha_Y^2}$

Similarly $\gamma_Z(t) = \mathrm{cov}(Z_m, Z_{m+t}) = \dfrac{\sigma_Z^2 \alpha_Z^{|t|}}{1-\alpha_Z^2}$

$\because Y_n, Z_n$ are independent

$\therefore \mathrm{cov}(Y_i, Z_j) = 0$

$$\Sigma_X(m,n) = \frac{\sigma_Y^{\alpha} \alpha_Y^{|m-n|}}{1-\alpha_Y^2} + \frac{\sigma_Z^2 \cdot \alpha_Z^{|m-n|}}{1-\alpha_Z^2}$$

This is only depending on $(m-n)$, thus X_n is stationary.

$$\gamma_X(0) = \frac{\sigma_Y^2}{1-\alpha_Y^2} + \frac{\sigma_Z^2}{1-\alpha_Z^2}$$

$$\rho_X(t) = \frac{\gamma_X(t)}{\gamma_X(0)} = \frac{\sigma_Y^2 \alpha_Y^{|t|}(1-\alpha_Z^2) + \sigma_Z^2 \cdot \alpha_Z^{|t|}(1-\alpha_Y^2)}{\sigma_Y^2(1-\alpha_Z^2) + \sigma_Z^2(1-\alpha_Y^2)}$$

(b) $X_n' = Y_n + Y_n \cdot Z_n$

$$\mu_{X_n'} = \mathbb{E}(Y_n + Y_n \cdot Z_n)$$
$$= \mathbb{E}Y_n + \mathbb{E}Y_n \cdot Z_n$$
$$= \mathbb{E}Y + \operatorname{cov}(Y_n, Z_n) + \mathbb{E}Y_n \cdot \mathbb{E}Z_n$$
$$= \mu_Y + 0 + \mu_Y \cdot \mu_Z$$

$$\mu_{X_n'} = \mu_Y(1+\mu_Z)$$

$$\Sigma_{X'}(m,n) = \operatorname{cov}(X_m', X_n')$$
$$= \operatorname{cov}(Y_m + Y_m \cdot Z_m,\ Y_n + Y_n \cdot Z_n)$$
$$= \operatorname{cov}(Y_m, Y_n) + \underbrace{\operatorname{cov}(Y_m Z_m, Y_n)}_{A} + \underbrace{\operatorname{cov}(Y_m, Y_n \cdot Z_n)}_{B} + \underbrace{\operatorname{cov}(Y_m Z_m, Y_n Z_n)}_{C}$$

A: $\mathbb{E}Y_m Z_m Y_n - \mathbb{E}Y_m Z_m \cdot \mathbb{E}Y_n = \mathbb{E}(Y_m \cdot Y_n) \cdot \mathbb{E}Z_m - \operatorname{cov}(Y_m, Z_m) - \mathbb{E}Y_m \mathbb{E}Z_m \cdot \mathbb{E}Y_n$

$= \gamma_Y(m-n) \cdot \mu_Z$

B: similarly $= \gamma_Y(m-n) \cdot \mu_Z$

C: $\mathbb{E}(Y_m Z_m Y_n Z_n) - \mathbb{E}(Y_m Z_m) \cdot \mathbb{E}(Y_n Z_n) = \mathbb{E}(Y_m Y_n) \cdot \mathbb{E}(Z_m Z_n) - \mathbb{E}Y_m \mathbb{E}Z_n \cdot \mathbb{E}Y_n \mathbb{E}Z_n$

$= (\operatorname{cov}(Y_m, Y_n) + \mu_Y^2)(\operatorname{cov}(Z_m, Z_n) + \mu_Z^2) - \mu_Y^2 \mu_Z^2$

$= (\gamma_Y(m-n) + \mu_Y^2)(\gamma_Z(m-n) + \mu_Z^2) - \mu_Y^2 \mu_Z^2$

$= \gamma_Y(m-n) \cdot \gamma_Z(m-n) + \mu_Y^2 \gamma_Z(m-n) + \mu_Z^2 \gamma_Y(m-n)$

$$\Sigma_X(m,n) = \gamma_Y(m-n) + 2\gamma_Y(m-n)\mu_Z + \gamma_Y(m-n)\gamma_Z(m-n) + \mu_Y^2 \gamma_Z(m-n) + \mu_Z^2 \gamma_Y(m-n)$$
$$= \gamma_Y(m-n)(1+\mu_Z)^2 + \gamma_Y(m-n)\gamma_Z(m-n) + \mu_Y^2 \gamma_Z(m-n)$$

Because μ_X is constant, $\Sigma_X(m,n)$ only depends on $(m-n)$

X_n is stationary

(c) $X_n = Y_n + Y_{n+1}$

$$\mu X_n = \mathbb{E}(Y_n + Y_{n+1})$$
$$= \mathbb{E}Y_n + \mathbb{E}Y_{n+1}$$

$$\mu X_n = 2\mu_Y$$

$$\Sigma_{X_n}(m,n) = \mathrm{cov}(Y_m + Y_{m+1}, Y_n + Y_{n+1})$$
$$= \mathrm{cov}(Y_m, Y_n) + \mathrm{cov}(Y_{m+1}, Y_n) + \mathrm{cov}(Y_m, Y_{n+1}) + \mathrm{cov}(Y_{m+1}, Y_{n+1})$$

let $m-n=t$

$$= \gamma_Y(t) + \gamma_Y(t+1) + \gamma_Y(t+1) + \gamma_Y(t)$$
$$= 2[\gamma_Y(t) + \gamma_Y(t+1)]$$
$$= 2\left[\frac{\sigma_Y^2 \alpha^{|t|} + \sigma_Y^2 \alpha^{|t+1|}}{1-\alpha_Y^2}\right]$$
$$= \frac{2\sigma_Y^2}{1-\alpha_Y^2}\left(\alpha^{|t|} + \alpha^{|t+1|}\right)$$

Because μX_n is constant, γX_n is only dependent on t

X_n is stationary.