

PS3Q3

Xiang Li

2024-02-08

```
source("create_monthly_price.R")
source("model_selection_function.R")
source("ols_function.R")
source("t_test_function.R")
source("generate_data_functions.R")
library("quantmod")
```

(a) Compile quarterly data for the U.S. GNP between 1947Q1 and 2023Q1.

```
getSymbols(Symbols = "GNP", src = "FRED")

## [1] "GNP"
```

(b) Compute quarterly GNP growth and filter the data to only include observations between 1960Q1 and 2019Q4. Provide summary statistics of the filtered quarterly GNP growth and plot the data.

```
head(GNP)
```

```
##              GNP
## 1947-01-01 244.142
## 1947-04-01 247.063
## 1947-07-01 250.716
## 1947-10-01 260.981
## 1948-01-01 267.133
## 1948-04-01 274.046
```

```
tail(GNP)
```

```
##              GNP
## 2022-04-01 25746.28
## 2022-07-01 26185.89
## 2022-10-01 26594.00
## 2023-01-01 26972.53
## 2023-04-01 27236.10
## 2023-07-01 27774.19
```

```
GNP = as.matrix(GNP[,1])
n_GNP = dim(GNP)[1]
GNP = as.matrix(GNP[5:n_GNP,1]/GNP[1:(n_GNP-4),1] - 1)
```

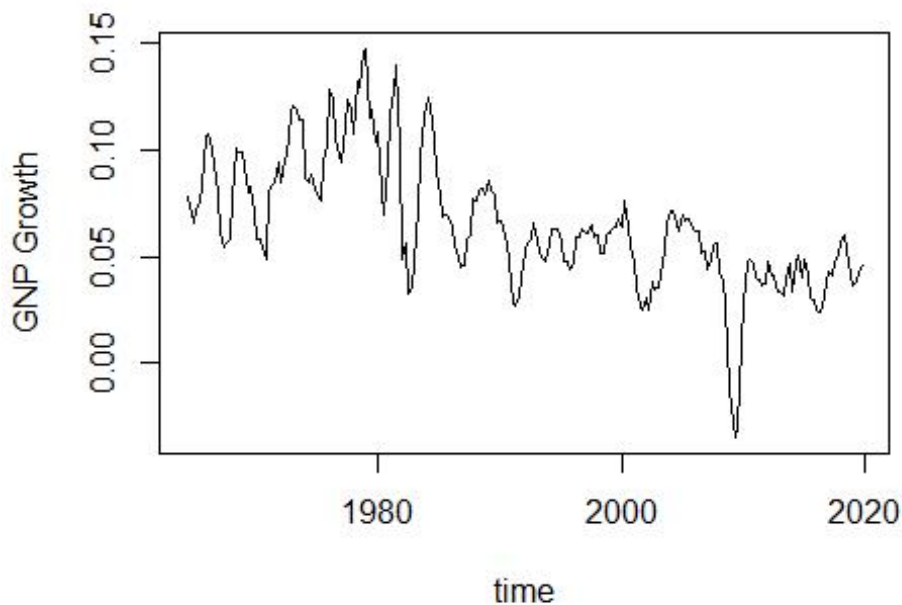
```

keep_data <- seq(from = as.Date("1964-04-01"), to = as.Date("2019-10-01"), by = "quarter")

GNP_G = as.matrix(GNP[as.Date(rownames(GNP)) %in% keep_data,])
n_obs = dim(GNP_G)[1]
date = as.Date(row.names(GNP_G))

plot(x = date, y = GNP_G, xlab='time', ylab='GNP Growth', type='l', col="black")

```



```
summary(GNP_G)
```

```

##          V1
##  Min.    :-0.03456
##  1st Qu.: 0.04668
##  Median : 0.06072
##  Mean    : 0.06535
##  3rd Qu.: 0.08193
##  Max.    : 0.14717

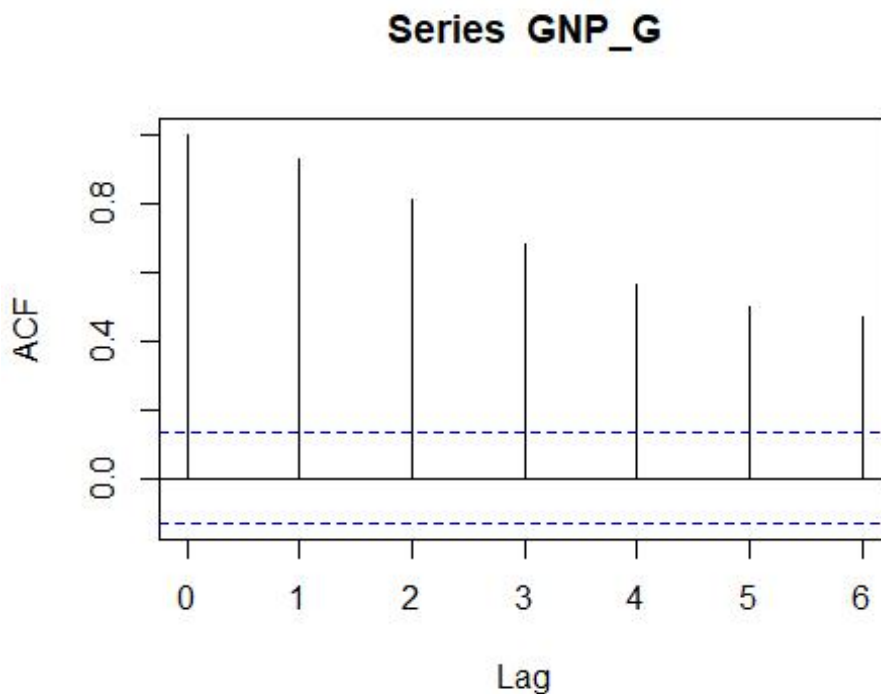
```

(c)

- i. Compute and plot empirical autocorrelation function (ACF). Given the plot, do you expect any time-series correlation among the observations? Explain why?
- ii. Set the maximum number of lags to the integer closest to the number of observations to the power one over third. Perform a test of joint

autocorrelation in GNP growth and report your result. Is your finding consistent with that of the ACF plot? Explain why?

```
acf(GNP_G, lag = round(n_obs^(1/3)))
```



Yes, because all the the bars extend beyond the dotted lines.

```
Box.test(GNP_G, lag = round(n_obs^(1/3)), type = "Ljung-Box")
```

```
##  
## Box-Ljung test  
##  
## data: GNP_G  
## X-squared = 630.72, df = 6, p-value < 2.2e-16
```

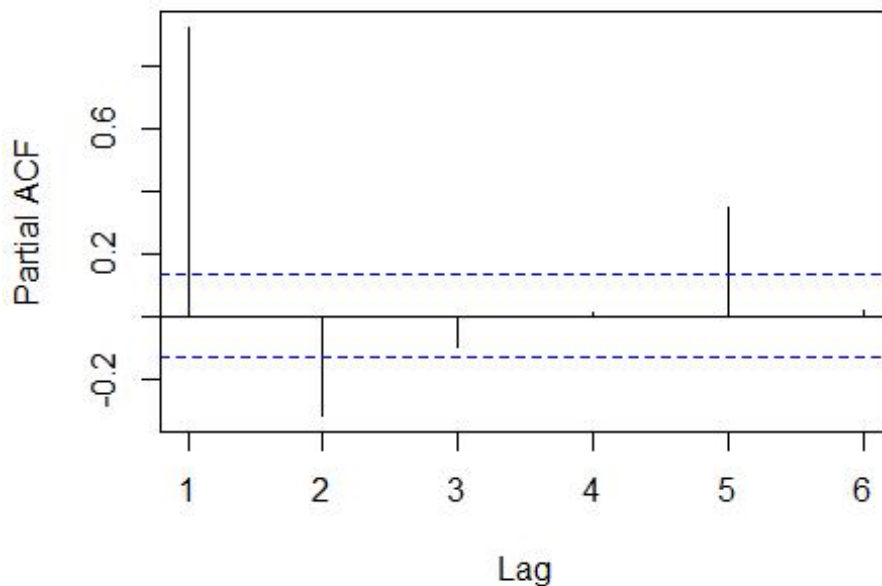
Reject the null hypothesis

(d)

- i. Compute and plot partial autocorrelation function (PACF). Given the plot, how many lags you would select for an AR process? Explain why?
- ii. Use AIC and BIC to determined the number of lags for an AR process. Compare your findings with that of the PACF plot.

```
pacf(GNP_G, lag = round(n_obs^(1/3)), main = "GNP Grpwth")
```

GNP Grpwth



```
results <- model_selection(n_obs^(1/3),GNP_G)
aic_values = results$AIC
bic_values = results$BIC
num_lags_aic = results$op_lag_AIC
num_lags_bic = results$op_lag_BIC
num_lags_aic

## [1] 5
```

(e)

- i. Consider an AR model with the number of lags selected by AIC. Estimate the coefficients of the AR model and compute the theoretical autocorrelation functions. Compare your findings with that of Part 3c.
- ii. Perform a joint test of autocorrelation among the residuals of the estimated model and report your result.

```
max_lags <- round(n_obs^(1/3))
sel_result_GNP = model_selection(max_lags,GNP_G)
num_lag_GNP = results$op_lag_AIC

lags_GNP <- matrix(NA, nrow = n_obs, ncol = num_lag_GNP)
for (j in 1:num_lag_GNP) {
  lags_GNP[(j+1):n_obs,j] <- GNP_G[1:(n_obs-j),1]
}
intercept = matrix(1,n_obs)
X = cbind(intercept,lags_GNP)
```

```

y = GNP_G
reg_result = ols(X[(num_lag_GNP+1):n_obs,], as.matrix(y[(num_lag_GNP+1):
n_obs,1]))
residuals = reg_result$u_hat # get the AR model residuals

beta_hat = reg_result$beta_hat
beta_hat

##           [,1]
## [1,]  0.004195339
## [2,]  1.189331251
## [3,] -0.148242170
## [4,] -0.058113331
## [5,] -0.405073904
## [6,]  0.355633283

Box.test(residuals, lag = round(n_obs^(1/3)), type = "Ljung-Box")

##
## Box-Ljung test
##
## data: residuals
## X-squared = 3.6032, df = 6, p-value = 0.7302

```

Cannot reject the null hypothesis.