STAT211 Mandatory Homework 8

Yapi Donatien Achou

April 2, 2019

Contents

1 Problem 8.1

1.1 Part a

```
yearlySunsplts <- na.omit(read.table(file="data/yearly_sunspots.txt",header=TRUE))
dat <- yearlySunsplts$sunspots
png("yearly.png")
plot(dat,col="blue",type="b")
```

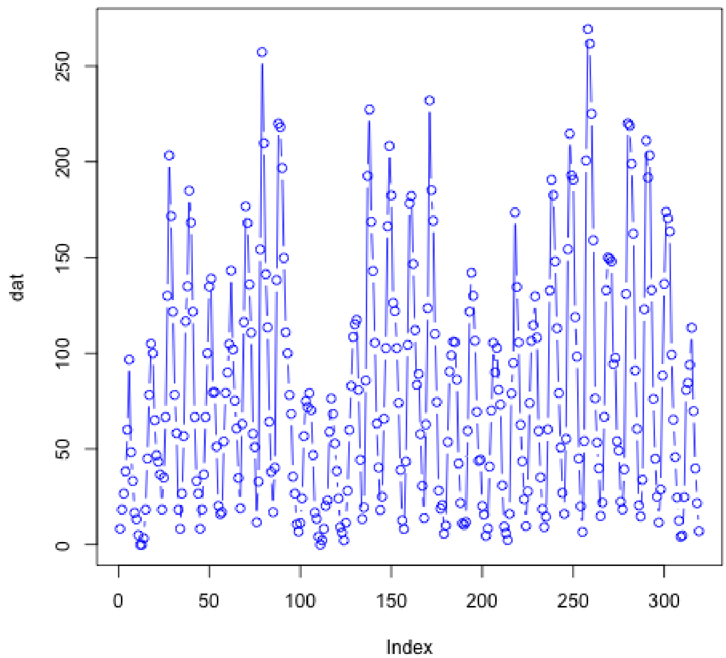

Figure 1: Yearly sunspot

1.2 Part b

Let

$$Y_t = X_t - \bar{X}_t \tag{1}$$

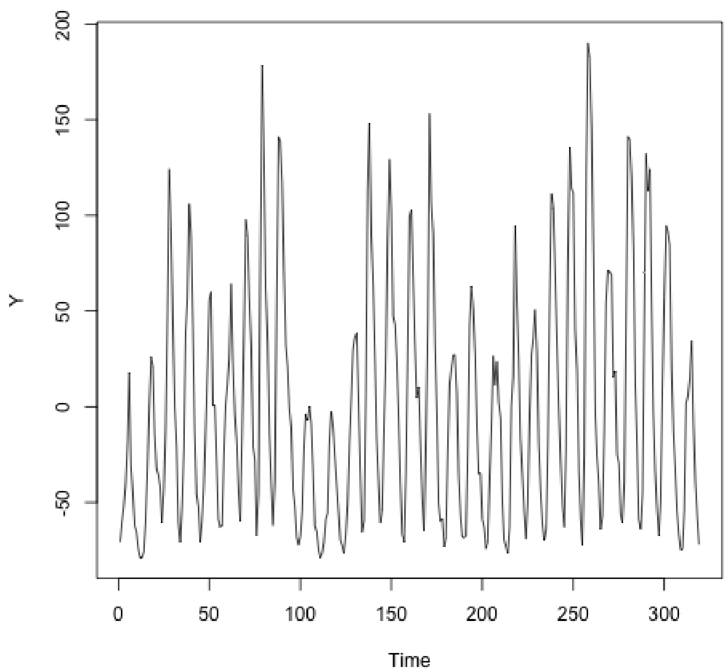

Figure 2: Plot of $Y_t = X_t - \bar{X}_t$

Figure 2 shows the plot of $Y_t = X_t - \bar{X}_t$.

```
yearlySunsplts <- na.omit(read.table(file="data/yearly_sunspots.txt",header=TRUE))
X <- yearlySunsplts$sunspots
Y <- X-mean(X)
png("note/acf.png")
acf(Y,col="blue")
```

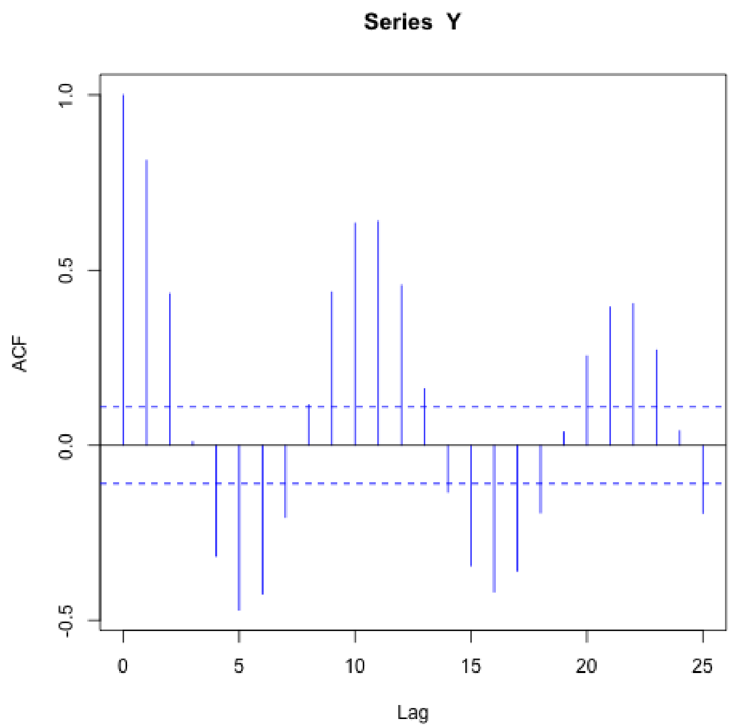

Figure 3: ACF

```
yearlySunsplts <- na.omit(read.table(file="data/yearly_sunspots.txt",header=TRUE))
X <- yearlySunsplts$sunspots
Y <- X-mean(X)
png("note/pacf.png")
pacf(Y,col="blue")
```

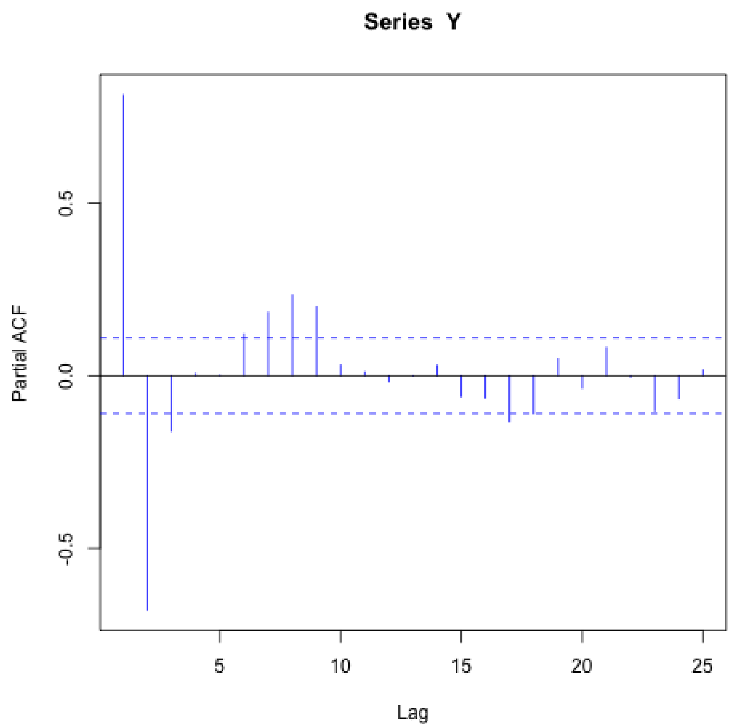

Figure 4: PACF

1.3 Part c: Fit AR(2) model

```
yearlySunsplts <- na.omit(read.table(file="data/yearly_sunspots.txt",header=TRUE))
X <- yearlySunsplts$sunspots
Y <- X-mean(X)
ar2 <- ar(Y,order.max=2)
print(ar2)

Call:
ar(x = Y, order.max = 2)

Coefficients:
      1        2
 1.3666  -0.6792

Order selected 2  sigma^2 estimated as  704.1
```

From the simulation result $\hat{\phi}_1 = 1.3666, \hat{\phi}_1 = 0.6792, \sigma_2 = 704.1$

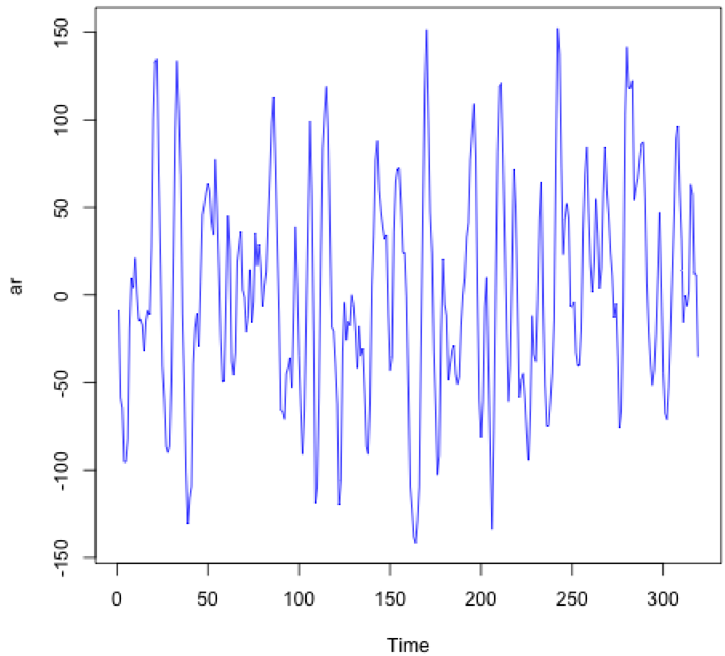

Figure 5: The simulated AR(2) process

1.4 Part d: Periodogram

```
library(TSA)
set.seed(10)
yearlySunsplts <- na.omit(read.table(file="data/yearly_sunspots.txt",header=TRUE))
X <- yearlySunsplts$sunspots
Y <- X-mean(X)
m <- length(Y)
ar.sim <- arima.sim(model=list(ar=c(1.3666, -0.6792)),n=m)
per <- periodogram(ar.sim)
spectrum <- per$spec
png("periodogram.png")
plot(spectrum,col="blue",type="b")
```

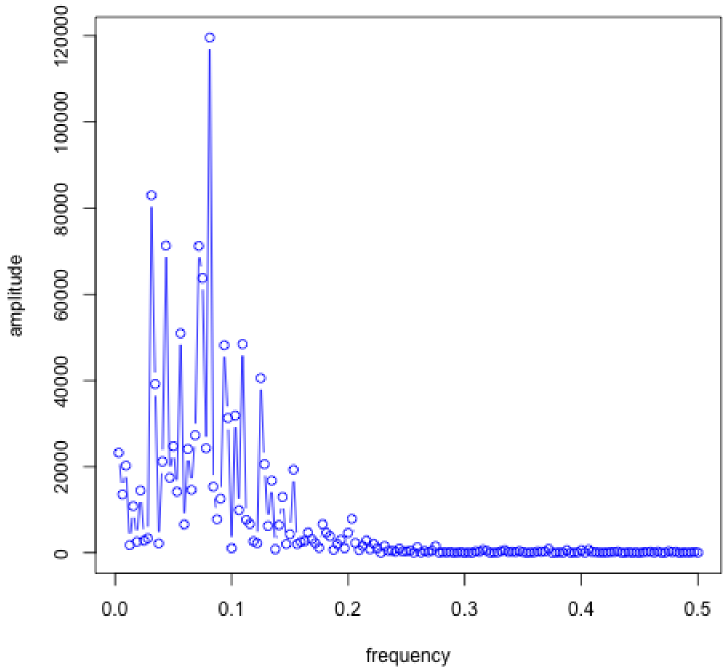

Figure 6: Periodogram

1.5 Part d: Smooth periodogram

```
library(TSA)
set.seed(10)
yearlySunsplts <- na.omit(read.table(file="data/yearly_sunspots.txt",header=TRUE))
X <- yearlySunsplts$sunspots
Y <- X-mean(X)
m <- length(Y)
ar.sim <- arima.sim(model=list(ar=c(1.3666, -0.6792)),n=m)
per <- periodogram(ar.sim)
amplitude <- per$spec
smoothPeriodogram = filter(amplitude, filter = c(1/3,3), sides=2)
png("smoothPeriodogram.png")
plot(smoothPeriodogram,col="blue",type="b")
```

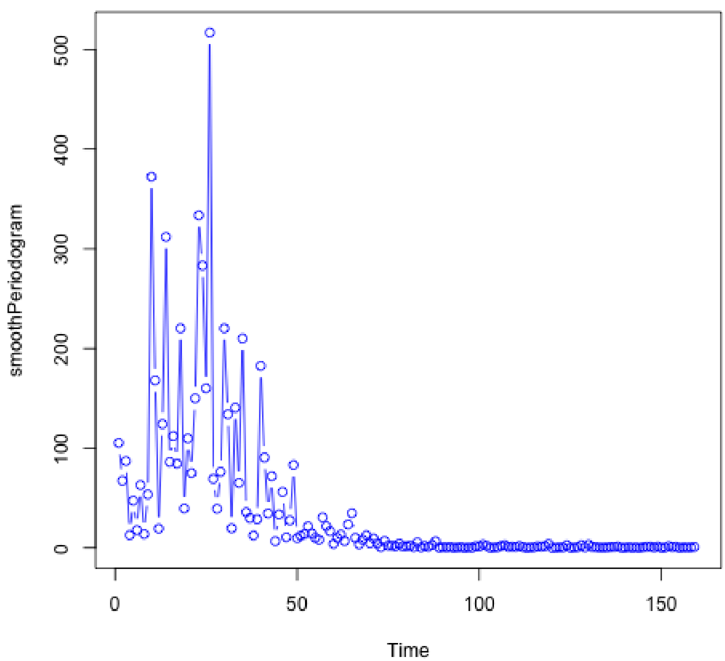

Figure 7: Smoother periodogram

1.6 Part f: Spectral density

```
library(TSA)
set.seed(10)
yearlySunsplts <- na.omit(read.table(file="data/yearly_sunspots.txt",header=TRUE))
X <- yearlySunsplts$sunspots
Y <- X-mean(X)
m <- length(Y)
ar.sim <- arima.sim(model=list(ar=c(1.3666, -0.6792)),n=m)
spectralDensity <- spectrum(ar.sim,method="ar")
png("spectralDensity.png")
plot(spectralDensity,col="blue")
```

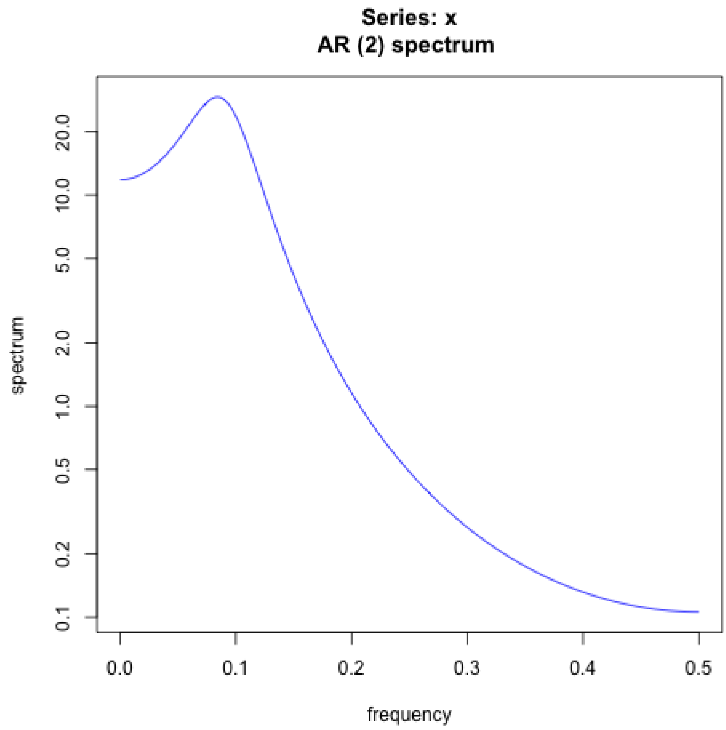

Figure 8: Spectral density

1.7 Part g

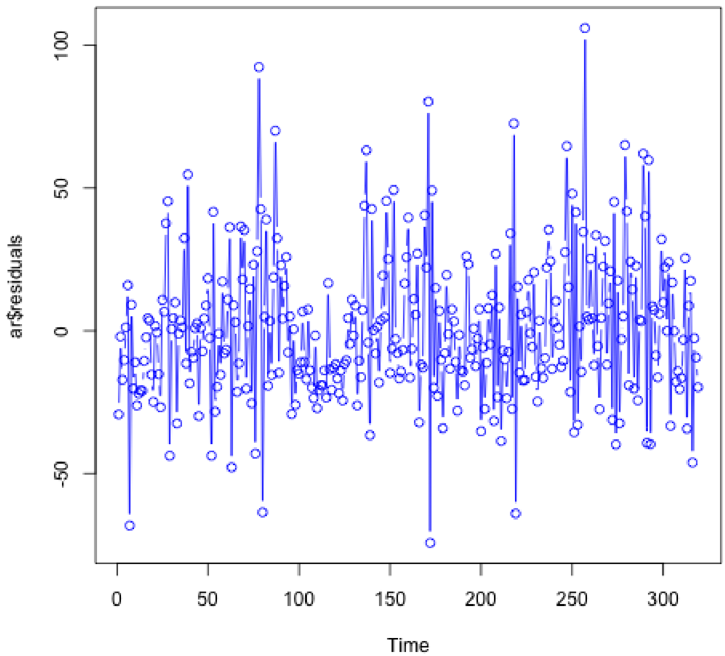

Figure 9: Residual

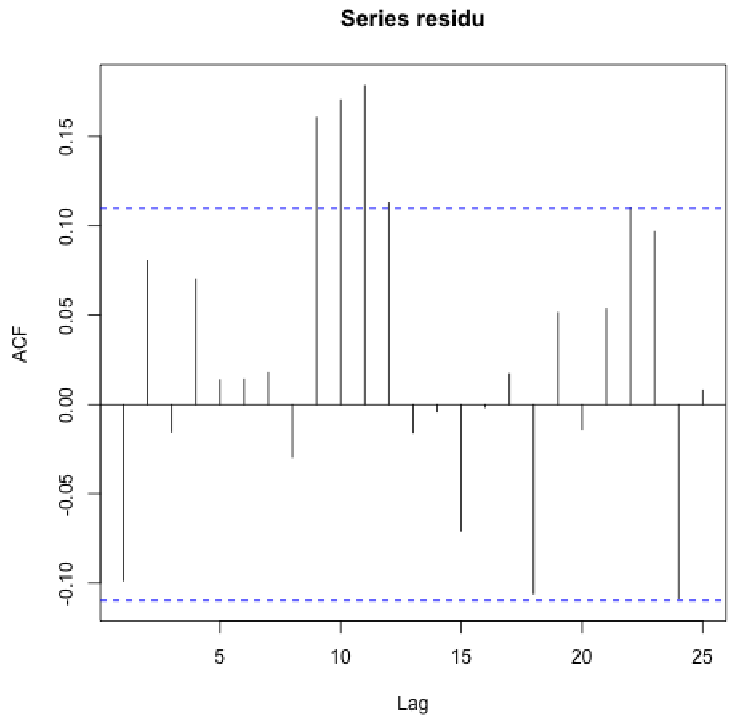

Figure 10: ACF of Residual

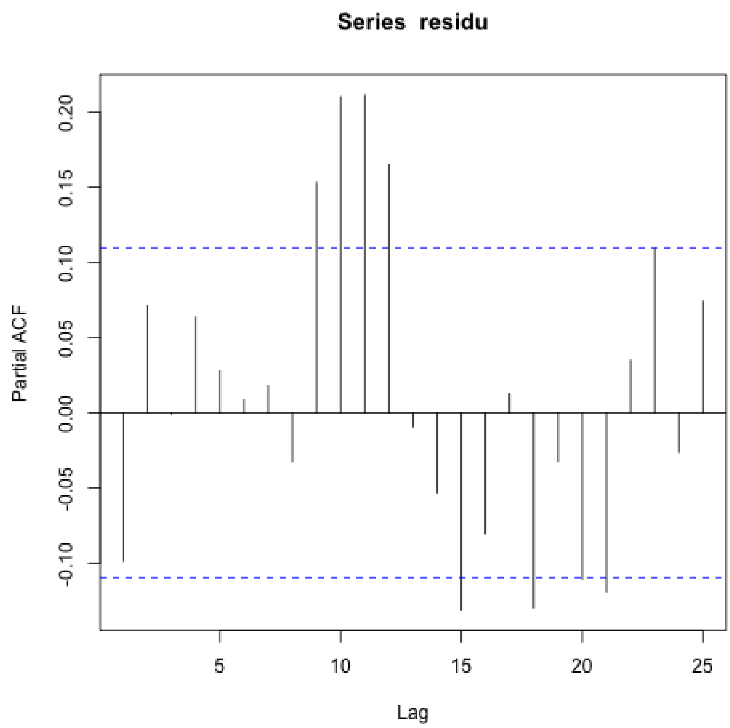

Figure 11: PACF of Residual

1.8 Part h

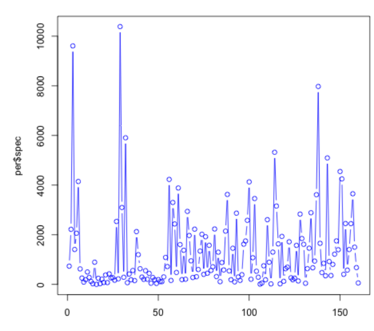

Figure 12: Periodogram of residual

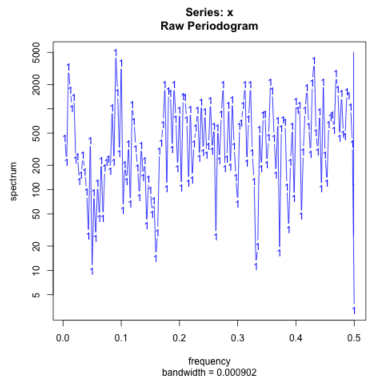

Figure 13: Spectral density of residual

2 Problem 8.2

In this exercise we study the spectral representation of a time series. The spectral representation of a time series is used to decompose a time series into sums of sinusoidal components [1]. By doing so, we can identify the dominant periods or frequencies of a time series by using its periodogram [2]. The periodogram is a sample based function that gives an estimate of the spectral density [1]. If $f(\cdot)$ is the spectral density and $I_n(\cdot)$ is the periodogram of n observations, then $I_n(\cdot)$ can be view as an estimation of $2\pi f(\cdot)$ [1].

The spectral density for an AR(2) model is given by

$$\begin{aligned} f(w) &= \frac{\sigma^2}{2\pi} \frac{1}{|\phi(\exp(-iw))|^2}, \quad w \in (-\pi, \pi], \quad \phi(z) = 1 - \phi_1 z - \phi_2 z^2. \\ &= \frac{\sigma^2}{2\pi} \frac{1}{|1 - \phi_1 e^{-iw} - \phi_2 e^{-2iw}|^2} \end{aligned} \tag{2}$$

The spectral density of a stationary process $\{X_t\}$ specifies the frequency decomposition of the autocovariance function (ACF) [1].

2.1 Part a: Spectral density

```
f <- function(w){
  phi1 <- 1.4
  phi2 <- -0.9
  phi <- 1-phi1*w-phi2*w**2
  sigma <- 1
  denominator <- abs(1-phi1*exp(-(0+1i)*w) - phi2*exp(-(0+2i)*w))**2
```

```
  factor <- (sigma*sigma)/(2*pi)
  value <- factor*(1./denominator)
  return(value)

}
set.seed(10)
w <- seq(-pi, pi)
png("theoreticalDensity.png")
plot(f(w), col="blue",type="b",xlab="frequency")
```

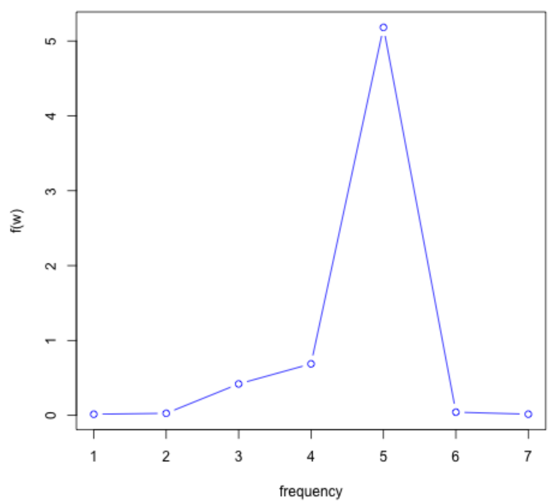

Figure 14: Spectral density of the AR(2) process

2.2 Part b: Comparing spectral density

For spectral for $\phi_2 = 0.95$

```
f <- function(w){
  phi1 <- 1.4
  phi2 <- 0.95
  phi <- 1-phi1*w-phi2*w**2
  sigma <- 1
  denominator <- abs(1-phi1*exp(-(0+1i)*w) - phi2*exp(-(0+2i)*w))**2
  factor <- (sigma*sigma)/(2*pi)
  value <- factor*(1./denominator)
  return(value)
```

```
}
set.seed(10)
w <- seq(-pi, pi)
png("theoreticalDensity.png")
plot(f(w), col="blue",type="b",xlab="frequency")
```

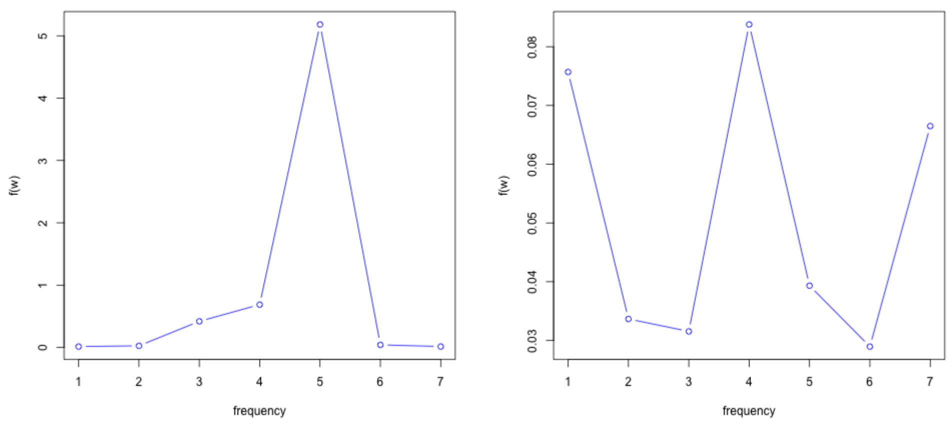

Figure 15: Spectral density of the AR(2) process. Left $\phi_2 = -0.90$, right $\phi_2 = 0.95$

We can observe from Figure (14) that the spectral density for $\phi_2 = 0.95$ oscillate more compared to the one with $\phi_2 = -0.9$.

2.3 Part c: Simulation of the AR(2) process

Simlating the AR(2) process with 100 obervations

```
# Plot of AR(2) model
ar.sim <- arima.sim(model=list(ar=c(1.4, -0.9)), n=100)
ts.plot(ar.sim)
```

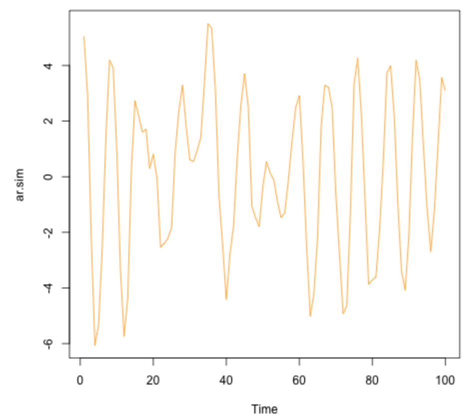

Figure 16: AR(2) simulation for $\phi_1 = 1.4, \phi_2 = -0.9$

```
#Periodogram  of AR(2) model
ar.sim <- arima.sim(model=list(ar=c(1.4, -0.9)), n=100)
per <- periodogram(ar.sim)
plot(per)
```

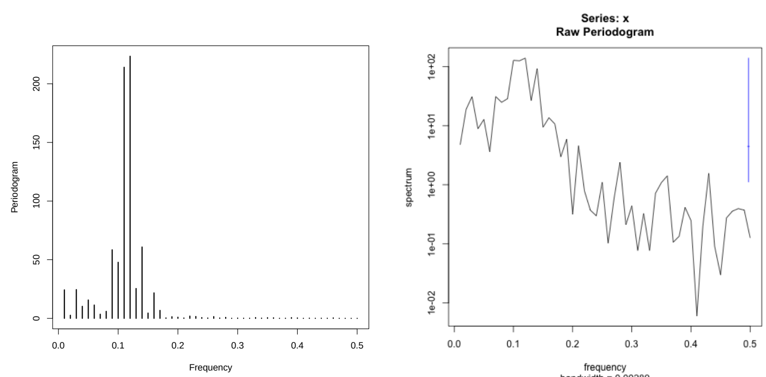

Figure 17: Periodogram of the AR(2) process

```
#Moving average smoothening of periodogram
set.seed(10)
ar.sim <- arima.sim(model=list(ar=c(1.4, -0.9)), n=100)
periodogram <- periodogram(ar.sim)
spectrum <- periodogram$spec
trendpattern = filter(spectrum, filter = c(1/3,3), sides=2)
```

```
plot(trendpattern)
```

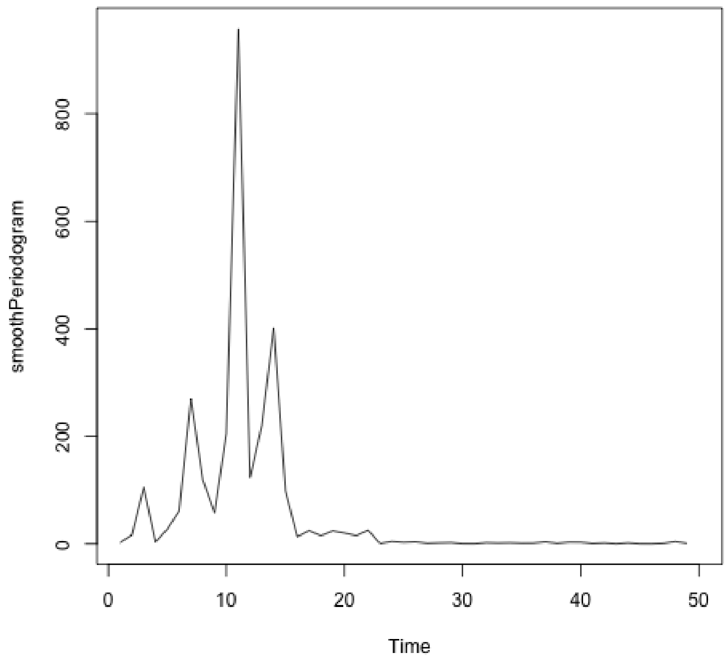

Figure 18: Smooth periodogram

```
#Estimated spectral density
set.seed(10)
ar.sim <- arima.sim(model=list(ar=c(1.4, -0.9)), n=100)
estimatedSpecralDensity <- spectrum(ar.sim,method="ar")
png("EstimatedSpectralDensity.png")
plot(estimatedSpecralDensity)
```

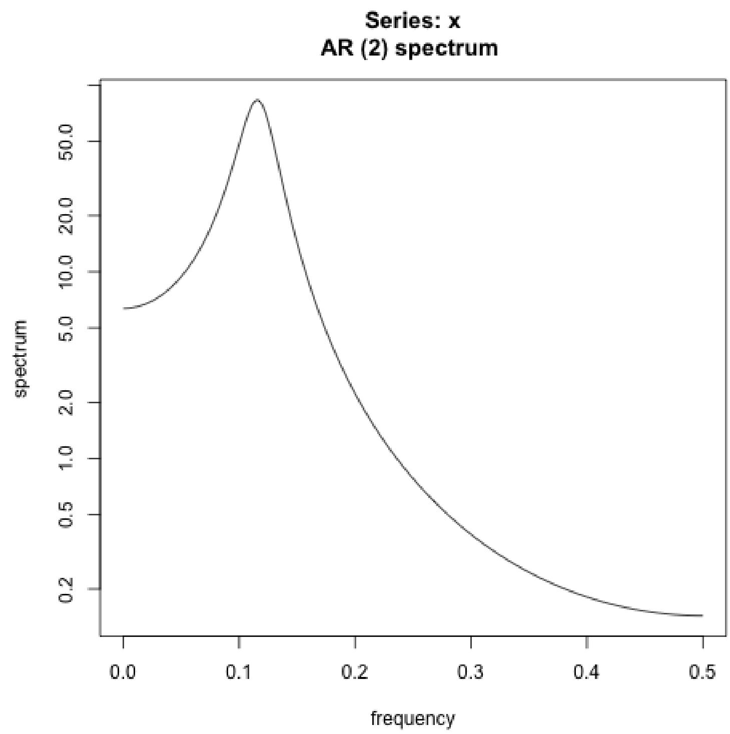

Figure 19: Estimated spectral density from R function spectrum

Figure 15 shows the simulated AR(2) process, Figure 16 shows the periodogram of the AR(2) process, Figure 17 shows the smoothen periodogram and Figure 18 shows the estimated spectral density from the AR(2) model.

2.4 Part d

Recall the plot of the AR(2) model

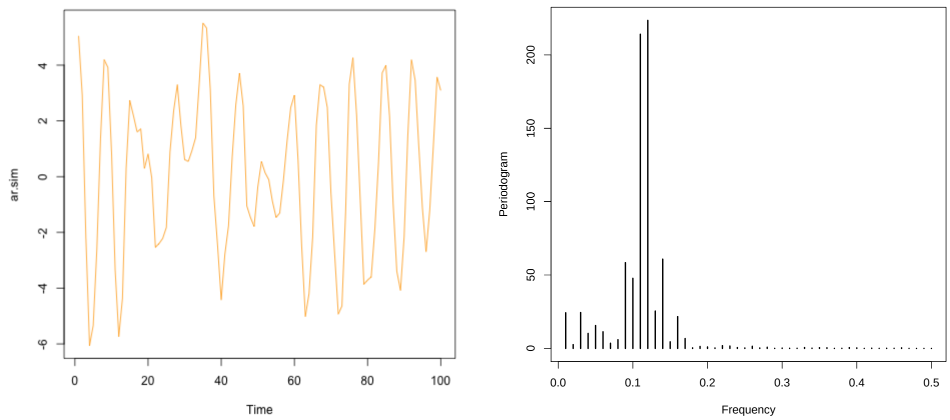

Figure 20: AR(2) simulation for $\phi_1 = 1.4, \phi_2 = -0.9$ left and periodogram right

From Figure 19 we can clearly see a periodic structure of the time series modelled by the AR(2) process. From the periodogram, we can see a pick. We get the peak and corresponding frequency as follow

```
ar.sim <- arima.sim(model=list(ar=c(1.4, -0.9)), n=100)
periodogram <- periodogram(ar.sim)
spectrum <- periodogram$spec
frequency <- periodogram$freq
pick <- max(spectrum)
pickIndex = match(pick, spectrum)
freq <- frequency[pickIndex]
print(pick)
print(freq)

>>[1] 315.5908
>>[1] 0.11
```

We observe a peak at a frequency of 0.11 which corresponds to period

$$T = \frac{1}{0.11} = 9 \tag{3}$$

which means that a cycle is completed in 9 time periods.

2.5 Part e

Simulation with 1000 observation

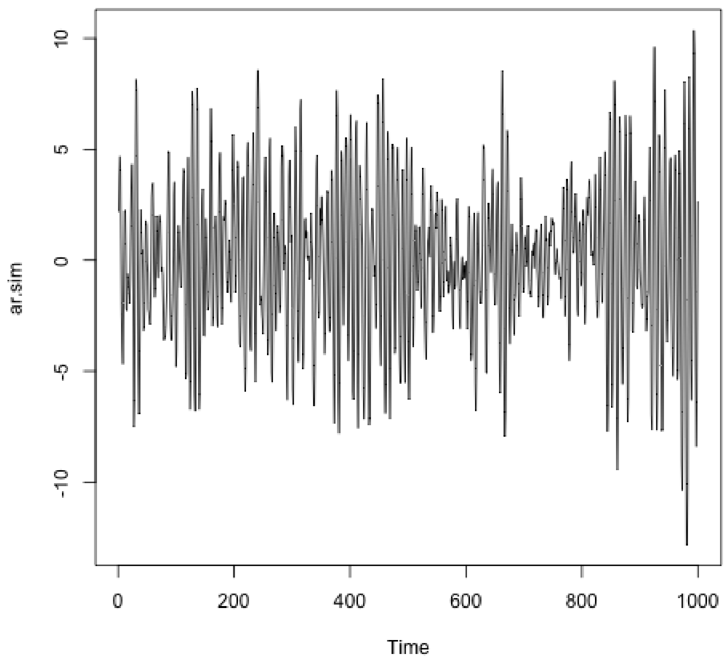

Figure 21: Simulation with 1000 obervartion

Periodogram for 1000 observations

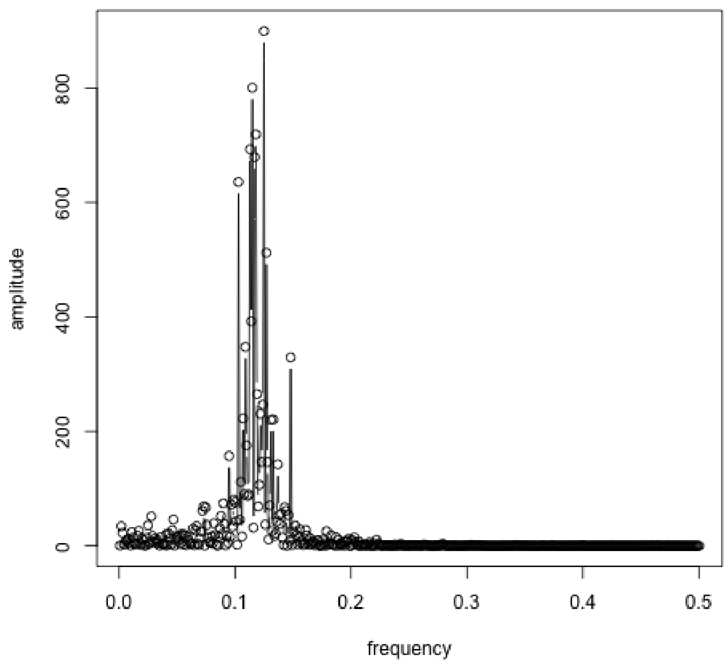

Figure 22: periodogram for 1000 obervation

Smooth peridogram

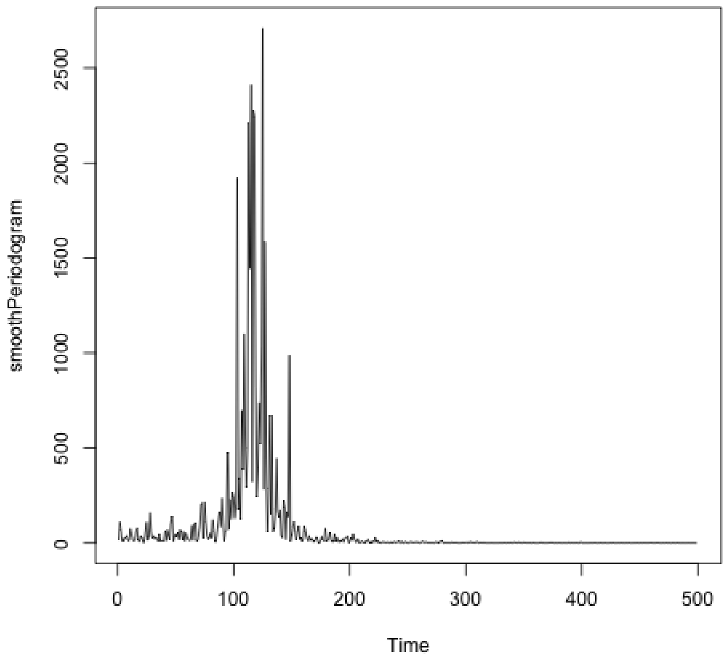

Figure 23: Smoother periodogram

Estimated spectral density

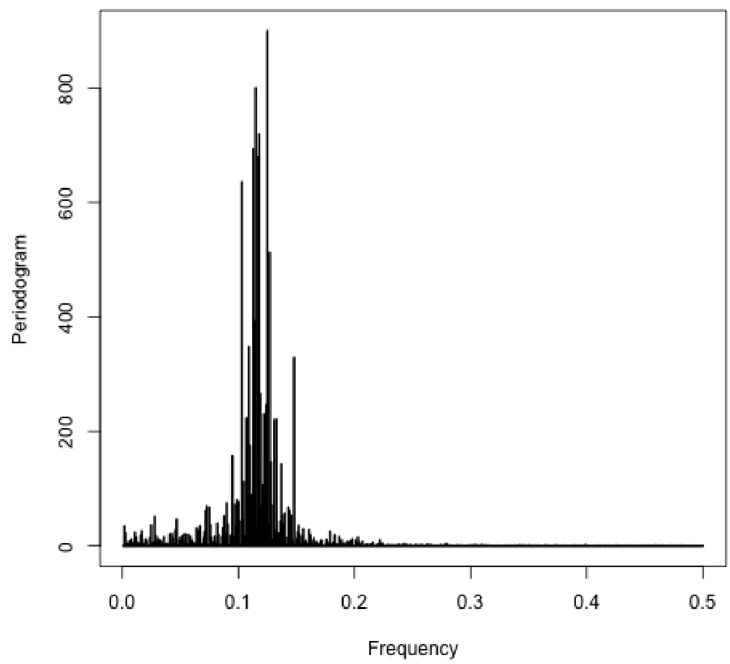

Figure 24: Estimated spectral density

For 1000 observations, the the peak is at at frequency of 0.125 corresponding to 8 time periods

3 Problem 8.3

If the process $\{X_t\}$ is a causal ARMA(p,q), its spectral density is given by [1]

$$f_X(\lambda) = \frac{\sigma^2}{2\pi} \frac{|\theta(e^{-i\lambda})|^2}{|\phi(e^{-i\lambda})|^2}, \quad -\pi \leq \lambda \leq \pi. \tag{4}$$

Furthermore,

$$\begin{aligned} \theta(e^{-i\lambda}) &= 1 + \theta_1 e^{-i\lambda} + \theta_2 e^{-i2\lambda} + \cdots + \theta_n e^{-in\lambda} \\ &= 1 + \sum_{k=0}^{n} \theta_n e^{-ik\lambda}, \end{aligned} \tag{5}$$

and

$$\begin{aligned}\phi(e^{-i\lambda}) &= 1 - \phi_1 e^{-i\lambda} - \phi_2 e^{-i2\lambda} - \cdots - \phi_n e^{-in\lambda} \\ &= 1 - \sum_{k=0}^{n} \phi_n e^{-ik\lambda}\end{aligned} \tag{6}$$

where

$$e^{-ik\lambda} = \cos(k\lambda) - i\sin(k\lambda). \tag{7}$$

Now let

$$\begin{aligned}f(\lambda) &= e^{-ik\lambda} \\ &= \cos(k\lambda) - i\sin(k\lambda),\end{aligned} \tag{8}$$

and let λ_0 be arbitrary taken from the interval $[-\pi, \pi]$. Then

$$\begin{aligned}\lim_{\lambda\to\lambda_0} f(\lambda) &= \lim_{\lambda\to\lambda_0} \cos(k\lambda) - i\sin(k\lambda) \\ &= cos(k\lambda_0) - i\sin(k\lambda_0) \\ &= f(\lambda_0)\end{aligned} \tag{9}$$

and this prove that $e^{-ik\lambda}$ is continuous and so is its linear combination

$$\theta(e^{-ik\lambda}) = 1 + \sum_{k=0}^{n} \theta_n e^{-ik\lambda}$$

and

$$\phi(e^{-ik\lambda}) = 1 - \sum_{k=0}^{n} \phi_n e^{-ik\lambda}.$$

Since we have a causal process, the roots of $\phi(e^{-ik\lambda})$ lies outside the unit circle, and $\phi(e^{-ik\lambda}) \neq 0$, and since the quotient of two continuous functions is continuous we have that

$$\frac{\theta(e^{-i\lambda})}{\phi(e^{-i\lambda})}$$

is continuous and

$$\frac{|\theta(e^{-i\lambda})|^2}{|\phi(e^{-i\lambda})|^2}$$

is also continuous. Therefore

$$f_X(\lambda) = \frac{\sigma^2}{2\pi} \frac{|\theta(e^{-i\lambda})|^2}{|\phi(e^{-i\lambda})|^2}, \quad -\pi \leq \lambda \leq \pi.$$

is continuous.

Assume now that the minimum of the spectral density is attained at some value, say λ_0. Then the minimum of the spectral density is

$$
\begin{aligned}
f_X(\lambda_0) &= \frac{\sigma^2}{2\pi}\frac{|\theta(e^{-i\lambda_0})|^2}{|\phi(e^{-i\lambda_0})|^2}, \quad -\pi \leq \lambda \leq \pi. \\
&= \frac{\sigma^2}{2\pi}\left|\frac{\theta(e^{-i\lambda_0})}{\phi(e^{-i\lambda_0})}\right|^2 \\
&= \frac{\sigma^2}{2\pi}\left|\frac{1+\sum_{k=0}^{n}\theta_k e^{-ik\lambda}}{1-\sum_{k=0}^{n}\phi_k e^{-ik\lambda}}\right|^2 \geq 0.
\end{aligned}
$$

For $f_X(\lambda_0)$ to be equal to zero, we need

$$
1+\sum_{k=0}^{n}\theta_k e^{-ik\lambda} = 1+\sum_{k=0}^{n}\theta_k\cos(k\lambda) + i\sum_{k=0}^{n}\theta_k\sin(k\lambda) = 0.
$$

This mean that we need to have simultaneously

$$
\sum_{k=0}^{n}\theta_k\sin(k\lambda) = 0.
$$

and

$$
\sum_{k=0}^{n}\theta_k\cos(k\lambda) = -1.
$$

Furthermore,

$$
\sum_{k=0}^{n}\theta_k\sin(k\lambda) = 0,
$$

means that either $\theta_k = 0$ or $\sin(k\lambda) = 0$, $\forall k = 1,\cdots,n$. Since all θ_k can not be equal to zero, the latter must hold, that is

$$
\begin{aligned}
\sin(\lambda) &= 0 \\
\sin(2\lambda) &= 0 \\
&\vdots \\
\sin(n\lambda) &= 0
\end{aligned}
$$

which means that

$$
n\lambda = \pm\frac{\pi}{2} \Rightarrow \lambda = \pm\frac{\pi}{2n}. \tag{10}
$$

Similarely,

$$\sum_{k=0}^{n} \theta_k \cos(k\lambda) = -1.$$

means that at $\lambda = \pm\frac{\pi}{2n}$ we must have

$$\theta_1 \cos\left(\frac{\pi}{2}\right) + \theta_2 \cos\left(2\frac{\pi}{4}\right) + \cdots + \theta_n \cos\left(n\frac{\pi}{2n}\right) = -1,$$

but

$$\theta_1 \cos\left(\frac{\pi}{2}\right) + \theta_2 \cos\left(2\frac{\pi}{4}\right) + \cdots + \theta_n \cos\left(n\frac{\pi}{2n}\right) = 0$$

so

$$\theta_1 \cos\left(\frac{\pi}{2}\right) + \theta_2 \cos\left(2\frac{\pi}{4}\right) + \cdots + \theta_n \cos\left(n\frac{\pi}{2n}\right) \neq -1$$

Therefore

$$f_X(\lambda_0) = \frac{\sigma^2}{2\pi} \left| \frac{1 + \sum_{k=0}^{n} \theta_k e^{-ik\lambda}}{1 - \sum_{k=0}^{n} \phi_k e^{-ik\lambda}} \right|^2 > 0.$$

4 Problem 8.4

Let $\{X_t, t = 1, \cdots, n\}$ be a data from a time series The likelihood estimator of the autocovariance function Γ_{LM} is given by [1]

$$\Gamma_{LM} = (2\pi)^{-n/2}(det(\Gamma_n))^{-1/2} \exp\left(-\frac{1}{2}(X'\Gamma_n^{-1}(X_n))\right) \tag{11}$$

and the sample autocovariance is given by

$$\hat{\Gamma}_n = \begin{bmatrix} \hat{\gamma}(0) & \hat{\gamma}(1) & \dots & \hat{\gamma}(k-1) \\ \hat{\gamma}(1) & \hat{\gamma}(0) & \dots & \hat{\gamma}(k-2) \\ \vdots & \vdots & \vdots & \vdots \\ \hat{\gamma}(k-1) & \hat{\gamma}(k-2) & \dots & \hat{\gamma}(0) \end{bmatrix} \tag{12}$$

where

$$\hat{\gamma}(h) = \frac{1}{n} \sum_{t=1}^{n-|h|} (X_{t+|h|} - \bar{X})(X_t - \bar{X}) \tag{13}$$

5 Problem 8.5

Let $\{X_t, t = 1, \cdots, n\}$ be a data from a time series. Supposed that we fit the AR(p) process to the data using the Yule Walker estimation of ϕ and σ^2. Let explain $\hat{\Gamma_{YW}}(h) = \hat{\gamma}(h)$ for $|h| \leq p$.

From [1], the Yule-Walker estimators $\hat{\phi}$ and $\hat{\sigma}^2$ of ϕ and σ^2 are solutions of

$$\hat{\Gamma}_p \hat{\phi} = \hat{\gamma}_p \tag{14}$$

and

$$\hat{\sigma}^2 = \hat{\gamma}(0) - \hat{\phi}' \hat{\gamma}_p \tag{15}$$

respectively. Where

$$\hat{\Gamma}_p = [\hat{\gamma}(i-j)]_{i,j=1}^{p}, \tag{16}$$

and

$$\hat{\gamma}_p = (\hat{\gamma}(0), \cdots, \hat{\gamma}(p). \tag{17}$$

For every nonsingular covariance matrix of the form (16), there is an AR(p) process whose autocovariance at lages $0, \cdots, p$ are $\gamma(0), \cdots, \gamma(p)$ where the coefficients and white noise are computed from the sample Yule-Walker equations [1]. Consequently, from [1] the Solutions of (14) and (15) is

$$\gamma_F(h) = \hat{\gamma}(h), \quad , h = 0, \cdots, p \tag{18}$$

Which are found from the Durbin-Levinson algorithm [1]. Therefore the autocovariance of the fitted model at lags $0, \cdots, p$ coincide with that of the sample autocovariance [1]

6 Problem 8.6

Let $\{X_t,\}$ be a MA(2) process

$$X_t = \theta(B) Z_t = (1 - B\xi_1^{-1})(1 - B\xi_2^{-1}) Z_t, \quad \{Z_t,\} \sim WN(0, \sigma^2) \tag{19}$$

6.1 Part a

$$X_t = \theta_1 Z_{t-1} + \theta_2 Z_{t-2} + Z_t \tag{20}$$

For an ARMA(p,q) process we have

$$\hat{\gamma}(k) - \phi_1 \hat{\gamma}(k-1) - \cdots - \phi_p \hat{\gamma}(k-p) = \sigma^2 \sum_{j=k}^{q} \theta_j \psi_{j-k}, \quad 0 \le k \le p+q. \tag{21}$$

For an MA(2) process equation (21) reduces to

$$\hat{\gamma}(k) = \sigma^2 \sum_{j=k}^{2} \theta_j \psi_{j-k}, \quad 0 \le k \le 2. \tag{22}$$

or

$$\hat{\gamma}(0) = \sigma^2 (1 + \theta_1 \psi_1 + \theta_2 \psi_2) \tag{23}$$

$$\hat{\gamma}(1) = \sigma^2(\theta_1 + \theta_2\psi_1) \quad (24)$$

$$\hat{\gamma}(2) = \sigma^2\theta_2 \quad (25)$$

from which θ_1, θ_2 can be computed from

6.2 Part b

The coefficient $\{\tilde{\theta}_j, j = 1, 2\}$ Applying the Durbin-Levinson algorithm to the sample auto-covariance we can fit the MA(2) process [1]. The model then becomes

$$X_t = \hat{\theta}_{21}Z_{t-1} + \hat{\theta}_{22}Z_{t-2} + Z_t \quad (26)$$

6.3 Partc

Finding the filter that fits

$$X_t = \tilde{\theta}(B)\tilde{Z}_t \quad (27)$$

References

[1] Petter J. Brockwell. Richard A. Davis *Introduction to Time Series and Forecasting.* Springer. Second edition. 2001

[2] PennState Eberly College of science, *STAT 510. Applied Time Series Analysis.* https://newonlinecourses.science.psu.edu/stat510/node/71/