[STAT 4540] HW-4

Michael Ghattas

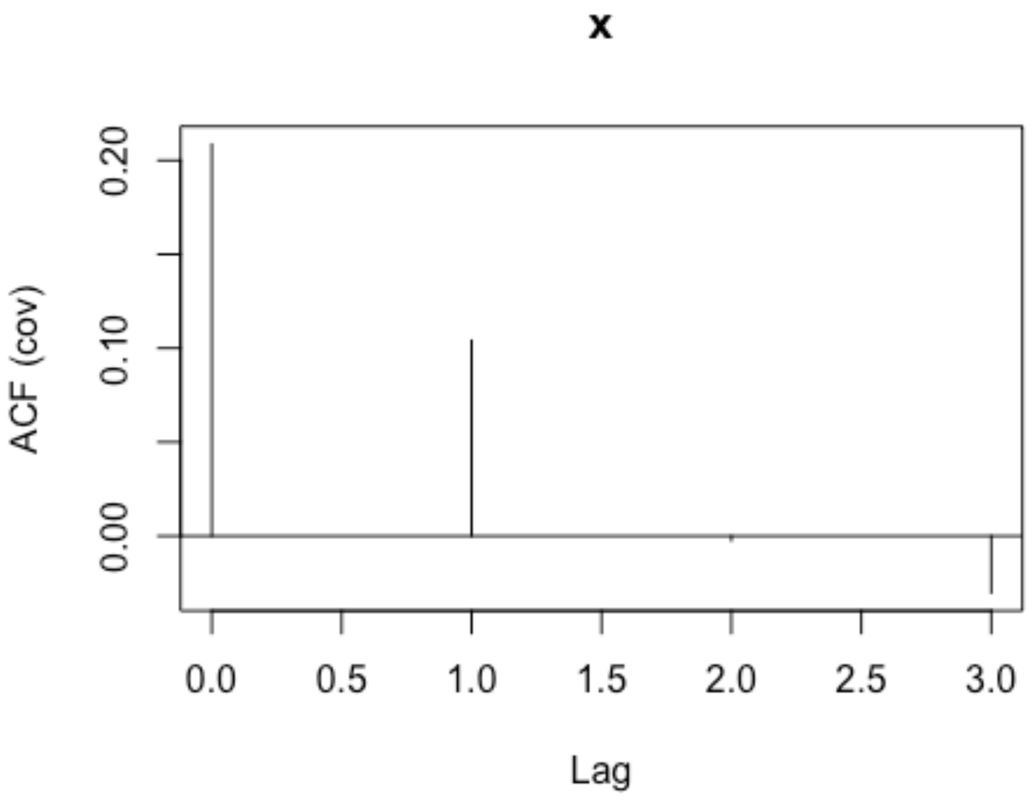
3/16/2022

Problem 1

(a)

```
library(lubridate)

##
## Attaching package: 'lubridate'

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

library(gridExtra)
library(ggplot2)
library(dplyr)

## Warning: package 'dplyr' was built under R version 4.1.2

##
## Attaching package: 'dplyr'

## The following object is masked from 'package:gridExtra':
##
##     combine

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

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

df <- load("/Users/Home/Documents/Michael_Ghattas/School/CU_Boulder/2022/
Spring 2022/STAT - 4540/HW/4/HW4_problem1_data.RData");
```

```
dat <- data.frame(x = x)
p = 3

phi = ar.yw(dat, aic = FALSE, order.max = p)

datACF = acf(dat, p, type = "covariance")
```




```
gam <- datACF[1:p]
g1 = as.numeric(datACF[1])[1]

## Warning: NAs introduced by coercion

## Warning: NAs introduced by coercion

## Warning: NAs introduced by coercion
```

```r
g2 = as.numeric(datACF[2])[1]

## Warning: NAs introduced by coercion

## Warning: NAs introduced by coercion

## Warning: NAs introduced by coercion

g3 = as.numeric(datACF[3])[1]

## Warning: NAs introduced by coercion

## Warning: NAs introduced by coercion

## Warning: NAs introduced by coercion

gamma = c(g1, g2, g3)
g0 = as.numeric(datACF[0])[1]

## Warning: NAs introduced by coercion

## Warning: NAs introduced by coercion

## Warning: NAs introduced by coercion

g1 = as.numeric(datACF[1])[1]

## Warning: NAs introduced by coercion

## Warning: NAs introduced by coercion

## Warning: NAs introduced by coercion

g2 = as.numeric(datACF[2])[1]

## Warning: NAs introduced by coercion

## Warning: NAs introduced by coercion

## Warning: NAs introduced by coercion
```

```
g = c(g0, g1, g2)
Gamma = toeplitz(g)

phi_hat = solve(Gamma, gamma)

phi

##
## Call:
## ar.yw.default(x = dat, aic = FALSE, order.max = p)
##
## Coefficients:
##       1        2        3
##  0.6908  -0.3834   0.0543
##
## Order selected 3  sigma^2 estimated as  0.1527

phi_hat

## [1]  0.69082236 -0.38339654  0.05429467
```

(b)

```
sigma2_hat = g0 - (gamma %*% phi_hat)
sigma2_hat

##            [,1]
## [1,] 0.1373881
```

(c)

```
AR <- arima.sim(list(ar = c(-1, -0.1)), 100)

mod1 = ar.yw(AR, order.max = 1, aic = FALSE)
mod2 = ar.yw(AR, order.max = 2, aic = FALSE)
mod99 = ar.yw(AR, order.max = 99, aic = FALSE)

predict(mod1)$pred[1]

## [1] -0.1570918

predict(mod2)$pred[1]
```

```
## [1] 0.09162709
```

```
predict(mod99)$pred[1]
```

```
## [1] 0.2076318
```

```
p1 = rep(0, 1000)
p2 = rep(0, 1000)
p99 = rep(0, 1000)

for (i in 1:1000)
{
  AR <- arima.sim(list(ar = c(-1, -0.1)), 101); AR

  mod1 = ar.yw(AR[1:100], order.max = 1, aic = FALSE)
  mod2 = ar.yw(AR[1:100], order.max = 2, aic = FALSE)
  mod99 = ar.yw(AR[1:100], order.max = 99, aic = FALSE)

  pred1 = predict(mod1)$pred[1]
  pred2 = predict(mod2)$pred[1]
  pred99 = predict(mod99)$pred[1]

  p1[i] = pred1 - AR[101]
  p2[i] = pred2 - AR[101]
  p99[i] = pred99 - AR[101]
}
```

```
mean(p1); mean(p2); mean(p99)
```

```
## [1] 0.02505485
```

```
## [1] 0.02826125
```

```
## [1] 0.03398891
```

```
sd(p1); sd(p2); sd(p99)
```

```
## [1] 0.9940399
```

```
## [1] 0.994255
```

```
## [1] 1.668998
```

With enough iterations we can start to see a trend of increasing values for the average prediction error and standard deviation.

Problem 2

(a)

$$X_t = Z_t + \theta \cdot Z_{t-1}$$
$$Z_t \sim WN(0, \sigma^2)$$

$$a = \Gamma^{-1} \cdot \gamma$$
$$\hat{x}_3 = a_1 \cdot x_1 + a_2 \cdot X_2 + a_3 \cdot X_4 + a_4 \cdot X_5$$
$$W = (X_1, X_2, X_4, X_5)$$

$$\gamma = [cov(X_3, X_1), cov(X_3, X_2), cov(X_3, X_4), cov(X_3, X_5)] = [\gamma(2), \gamma(1), \gamma(1), \gamma(2)] = [0, \theta \cdot \sigma^2, \theta \cdot \sigma^2, 0]$$

$$\Gamma =$$
$$|var(x_1), cov(x_1, x_2), cov(x_1, x_4), cov(x_1, x_5)|$$
$$|cov(x_2, x_1), var(x_2)), cov(x_2, x_4), cov(x_2, x_5)|$$
$$|cov(x_4, x_1), cov(x_4, x_2), var(x_4)), cov(x_4, x_5)|$$
$$|cov(x_5, x_1), cov(x_5, x_2), cov(x_5, x_4), var(x_5)|$$
$$=$$
$$|\gamma(0), \gamma(1), \gamma(3), \gamma(4)|$$
$$|\gamma(1), \gamma(0), \gamma(2), \gamma(3)|$$
$$|\gamma(3), \gamma(2), \gamma(0), \gamma(1)|$$
$$|\gamma(4), \gamma(3), \gamma(1), \gamma(0)|$$
$$=$$
$$|\sigma^2 \cdot (1 + \theta^2), \theta \cdot \sigma^2, 0, 0|$$
$$|\theta \cdot \sigma^2, \sigma^2 \cdot (1 + \theta^2), 0, 0|$$
$$|0, 0, \sigma^2 \cdot (1 + \theta^2), \theta \cdot \sigma^2|$$
$$|0, 0, \theta \cdot \sigma^2, \sigma^2 \cdot (1 + \theta^2)|$$

$\Gamma^{-1} =$

$$
\begin{array}{l}
|\frac{1}{\sigma^2 \cdot (1+\theta^2)}, \frac{1}{\theta \cdot \sigma^2}, 0, 0| \\
|\frac{1}{\theta \cdot \sigma^2}, \frac{1}{\sigma^2 \cdot (1+\theta^2)}, 0, 0| \\
|0, 0, \frac{1}{\sigma^2 \cdot (1+\theta^2)}, \frac{1}{\theta \cdot \sigma^2}| \\
|0, 0, \frac{1}{\theta \cdot \sigma^2}, \frac{1}{\sigma^2 \cdot (1+\theta^2)}|
\end{array}
$$

$$\hat{X}_3 = P(X_3 | W) = a^T \cdot W = \Gamma^{-1} \cdot \gamma \cdot [X_1, X_2, X_4, X_5]^T$$

(b)

```
sig2 = (1/3)
theta = sqrt(2)

gamma0 = sig2 * (1 + theta^2)
gamma1 = sig2 * theta
gam = c(0, gamma1, gamma1, 0)

gc1 = c(gamma0, gamma1, 0, 0)
gc2 = c(gamma1, gamma0, 0, 0)
gc3 = c(0, 0, gamma0, gamma1)
gc4 = c(0, 0, gamma1, gamma0)
Gam = as.matrix(cbind(gc1, gc2, gc3, gc4))

a = as.matrix(solve(Gam, gam))

MSE = gamma0 - (t(a) %*% gam); MSE

##            [,1]
## [1,] 0.4285714
```

Problem 3

(a)

If $X \sim N(0, \Sigma) \rightarrow X \sim N(\overrightarrow{0}, \sigma^2 I)$, this means x_i and x_j are uncorrelated and independent for MVN. Therefore $Y = \Sigma^{-\frac{1}{2}} X$ is $Y \sim MVN(\mu, \Sigma)$ and iid.

(b)

$f(X) = f(X_1, X_2, X_3, X_4) = f(\overrightarrow{X})$
Joint distribution for $X \sim N(\mu, \Sigma) = f(\mu, \theta, \sigma^2)$

$$f(\mu, \theta, \sigma^2) = \prod_{i=1}^{4} \frac{1}{2\pi \cdot |\Sigma^{-\frac{1}{2}}|} \exp - \frac{1}{2} \cdot (\overrightarrow{x} - \overrightarrow{\mu}) \cdot \Sigma^{-\frac{1}{2}} \cdot (\overrightarrow{x} - \overrightarrow{\mu})$$

Where:

$\overrightarrow{x} = [x_1, x_2, x_3, x_4]$
$\overrightarrow{\mu} = [0, 0, 0, 0]$
$\Sigma^{-\frac{1}{2}} =$

$$|\sqrt{var(x_1)}, \sqrt{cov(x_1, x_2)}, \sqrt{cov(x_1, x_3)}, \sqrt{cov(x_1, x_4)}|$$
$$|\sqrt{cov(x_2, x_1)}, \sqrt{var(x_2)}, \sqrt{cov(x_2, x_3)}, \sqrt{cov(x_2, x_4)}|$$
$$|\sqrt{cov(x_3, x_1)}, \sqrt{cov(x_3, x_2)}, \sqrt{var(x_3))}, \sqrt{cov(x_3, x_4)}|$$
$$|\sqrt{cov(x_4, x_1)}, \sqrt{cov(x_4, x_2)}, \sqrt{cov(x_4, x_3)}, \sqrt{var(x_4)}|$$

$=$

$$|\sqrt{\gamma(0)}, \sqrt{\gamma(1)}, \sqrt{\gamma(2)}, \sqrt{\gamma(3)}|$$
$$|\sqrt{\gamma(1)}, \sqrt{\gamma(0)}, \sqrt{\gamma(1)}, \sqrt{\gamma(2)}|$$
$$|\sqrt{\gamma(2)}, \sqrt{\gamma(1)}, \sqrt{\gamma(0)}, \sqrt{\gamma(1)}|$$
$$|\sqrt{\gamma(3)}, \sqrt{\gamma(2)}, \sqrt{\gamma(1)}, \sqrt{\gamma(0)}|$$

$=$

$$|\sigma\sqrt{1 + \theta_1^2 + \theta_2^2}, \sigma\sqrt{\theta_1 \cdot (1 + \theta_2)}, \sigma\sqrt{\theta_2}, 0|$$
$$|\sigma\sqrt{\theta_1 \cdot (1 + \theta_2)}, \sigma\sqrt{1 + \theta_1^2 + \theta_2^2}, \sigma\sqrt{\theta_1 \cdot (1 + \theta_2)}, \sigma\sqrt{\theta_2}|$$
$$|\sigma\sqrt{\theta_2}, \sigma\sqrt{\theta_1 \cdot (1 + \theta_2)}, \sigma\sqrt{1 + \theta_1^2 + \theta_2^2}, \sigma\sqrt{\theta_1 \cdot (1 + \theta_2)}|$$
$$|0, \sigma\sqrt{\theta_2}, \sigma\sqrt{\theta_1 \cdot (1 + \theta_2)}, \sigma\sqrt{1 + \theta_1^2 + \theta_2^2}|$$