Homework7

Hsueh-Pin Liu

2023-04-18

Problem 1

Suppose an AR(1) is fitted to a time series data of length $n = 144$ and the estimated values of the parameters are $\mu = 99$, $\phi = 0{:}6$ and $\sigma = 1$. Assume the last three values of the time series are $X_{144} = 100$, $X_{143} = 100$, $X_{142} = 99$. Compute the forecasts and 95% forecast intervals for the next four values.

1. $\hat{X}_{144}(1) = 99 + 0.6(100-99) = 99.6$

$\hat{X}_{144}(2) = 99 + 0.6^2(100-99) = 99.36$

$\hat{X}_{144}(3) = 99 + 0.6^3(100-99) = 99.216.$

$\hat{X}_{144}(4) = 99 + 0.6^4(100-99) = 99.1296$ and the forecast intervals are:

$X_{145} = 99.6 \pm 1.96 \times 1 = 99.6 \pm 1.96$

$X_{146} = 99.36 \pm 1.96 \times 1\sqrt{\frac{1-0.6^4}{1-0.6^2}} = 99.36 \pm 2.29$

$X_{147} = 99.216 \pm 1.96 \times 1\sqrt{\frac{1-0.6^6}{1-0.6^2}} = 99.216 \pm 2.39$

$X_{148} = 99.1296 \pm 1.96 \times 1\sqrt{\frac{1-0.6^8}{1-0.6^2}} = 99.1296 \pm 2.43$

Figure 1: Problem 1

Problem 2

Suppose the annual sales of a company (in millions of \$) follow an AR(2) model given by $X_t = 5 + 1.1X_{t-1} - 0.5X_{t-2} + e_t$ with $\sigma^2 = 2$.

(a) If the sales of the company in 2011, 2012 and 2013 were \$9 million, \$11 million and \$10 million respectively, forecast sales for 2014 and 2015.

(b) Construct 95% forecast intervals for 2014 and 2015.

(c) If the sales in 2014 turns out to be \$12 million, update your forecast for 2015.

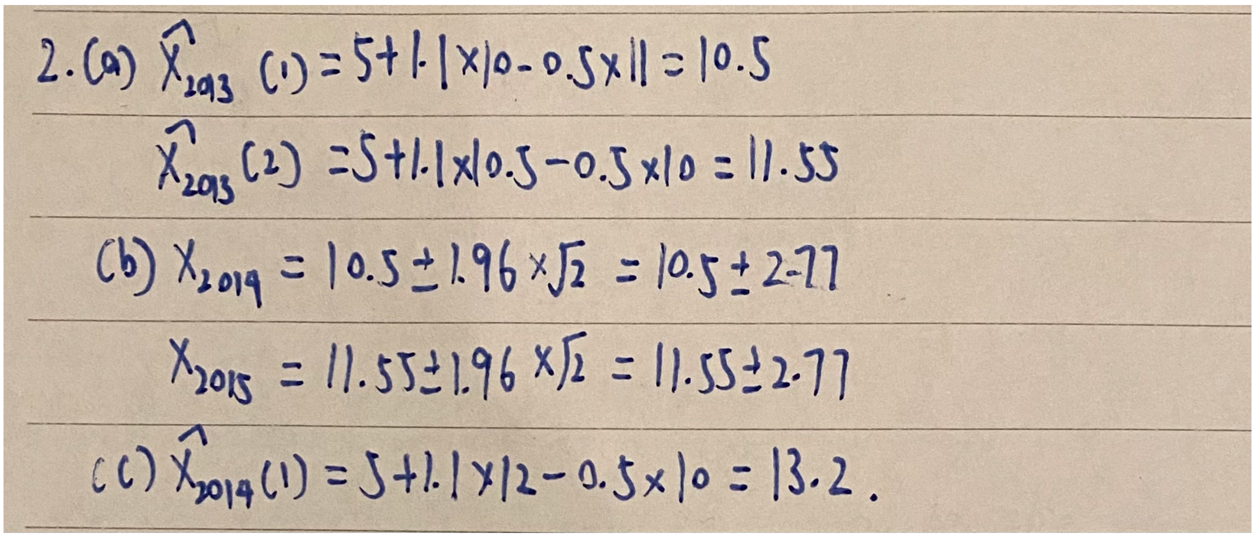


Figure 2: Problem 2

Problem 3

Simulate an AR(2) process with ϕ_1 = 1.5, ϕ_2 = -0.75, and μ = 100. Simulate 100 values, but set aside the last 10 values to compare forecasts to actual values.

(a) Using the first 90 observations in the series, find the MLE of the model parameters. Are the estimates comparable to the true values?

(b) Use the fitted model to forecast the 10 future values and obtain 95% forecast intervals.

(c) What percentage of the observed values are covered by the forecast intervals?

(d) Simulate a new sample data of the same size from the sample model and repeat steps (a),(b) and (c).

(a)

```
set.seed(1)
phi1 <- 1.5
phi2 <- -0.75
mu <- 100
n <- 100
pro3 <- arima.sim(model=list(ar=c(phi1, phi2)), n=n, mean=mu)
fit <- arima(pro3[1:90], order=c(2,0,0))
estimates <- coef(fit)
estimates
```

```
##         ar1         ar2   intercept
##   1.5256869  -0.7708933 400.3950711
```

According to the results, the estimates are comparable to the true values.

(b)

```
library(forecast)
```

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

```
pred <- predict(fit, n.ahead=10, interval="95%")
forecast_mean <- pred$pred
forecast_upper <- pred$pred + 1.96*pred$se
forecast_lower <- pred$pred - 1.96*pred$se
forecast(fit,level=c(0.95))
```

```
##     Point Forecast    Lo 95    Hi 95
##  91       398.4302 396.7207 400.1397
##  92       398.9846 395.8661 402.1031
##  93       399.7578 395.6581 403.8576
```

```
## 94        400.5102 395.9265 405.0938
## 95        401.0619 396.3537 405.7701
## 96        401.3237 396.6151 406.0323
## 97        401.2979 396.5321 406.0636
## 98        401.0565 396.1492 405.9639
## 99        400.7083 395.6521 405.7645
## 100       400.3631 395.2173 405.5088
```

(c)

```
actual <- pro3[91:100]
coverage <- sum(actual > forecast_lower & actual < forecast_upper)/length(actual)*100
coverage
```

```
## [1] 100
```

So it's 100% covered by the forecast intervals.

(d)

```
#a
set.seed(1)
phi1 <- 1.5
phi2 <- -0.6
mu <- 100
n <- 100
pro3 <- arima.sim(model=list(ar=c(phi1, phi2)), n=n, mean=mu)
fit <- arima(pro3[1:90], order=c(2,0,0))
estimates <- coef(fit)
estimates
```

```
##          ar1          ar2    intercept
##    1.5320819   -0.6860704 1001.0052997
```

```
#b
library(forecast)
pred <- predict(fit, n.ahead=10, interval="95%")
forecast_mean <- pred$pred
forecast_upper <- pred$pred + 1.96*pred$se
forecast_lower <- pred$pred - 1.96*pred$se
forecast(fit,level=c(0.95))
```

```
##    Point Forecast    Lo 95    Hi 95
## 91      1000.6047 998.9077 1002.302
## 92       999.9149 996.8102 1003.020
## 93       999.6095 995.4160 1003.803
## 94       999.6150 994.7147 1004.515
## 95       999.8328 994.5587 1005.107
```

```
##  96      1000.1628 994.7429 1005.583
##  97      1000.5189 995.0687 1005.969
##  98      1000.8382 995.3879 1006.288
##  99      1001.0829 995.6209 1006.545
## 100      1001.2389 995.7478 1006.730
```

```
#c
actual <- pro3[91:100]
coverage <- sum(actual > forecast_lower & actual < forecast_upper)/length(actual)*100
coverage
```

```
## [1] 100
```

I changed the two ϕs and still fit an AR(2) model, the result is showed as above.

Problem 4

Consider a MA(1) process given by $\mu = 5$, $\theta = 0.6$ and $\sigma = 0.1$. Suppose a sample realization of n = 5 is given by (starting from 1 ending at 5) 4.16, 5.76, 5.77, 4.02, 3.67. Find the forecast of the sixth and seventh observations and construct 95% forecast intervals.

```
mu <- 5
theta <- 0.6
sigma <- 0.1
n <- 5
x <- c(4.16,5.76,5.77,4.02,3.67)
fit <- arima(x, order=c(0,0,1))
forecast(fit,h=2,level=c(0.95))
```

```
##   Point Forecast    Lo 95    Hi 95
## 6       4.303336 2.977522 5.629150
## 7       4.533339 2.797441 6.269237
```

Problem 5

Consider the Johnson and Johnson Data from the last homework.

(a) Carry out appropriate Holt-Winters forecast for the next eight values. List and plot the forecasts along with the forecast intervals.

(b) Identify an appropriate ARIMA model and use the model to forecast for the next eight values. List and plot the forecasts along with the forecast intervals.

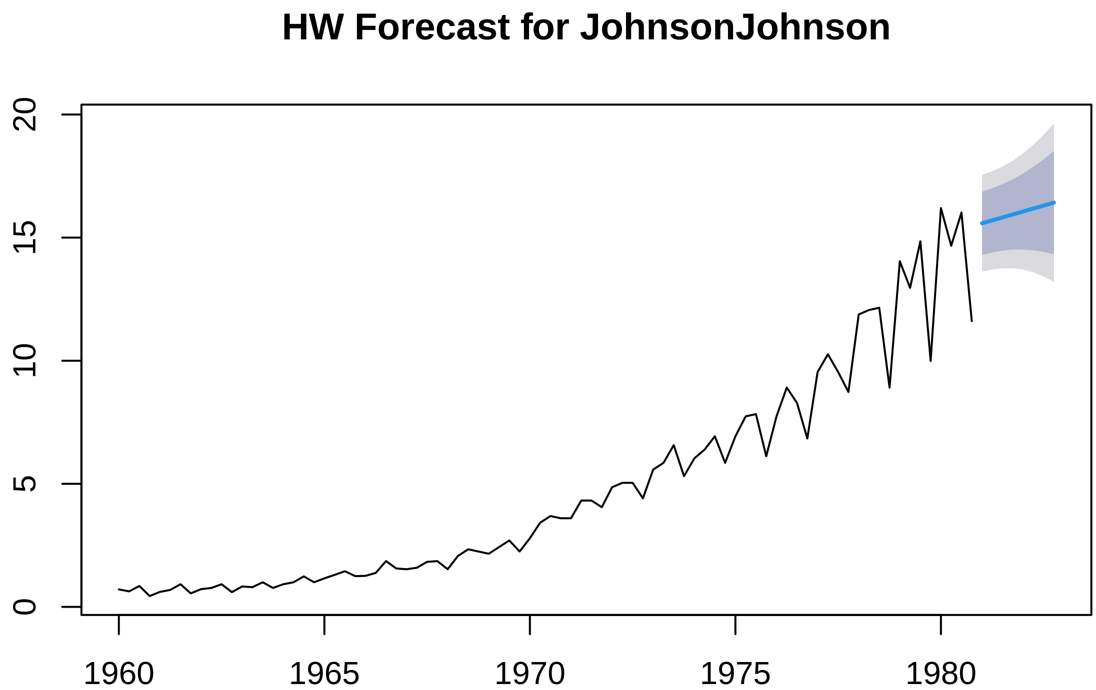
(c) Set aside the last eight observations in the data set as the validations sample and using the remaining data as the training sample, predict the eight observations. Compute RMSE, MAE and MAPE criteria of forecast comparison. What is your conclusion?

(a)

```
data("JohnsonJohnson")
library(forecast)
fit <- HoltWinters(JohnsonJohnson,gamma=F)
forecast.hw <- forecast(fit,h=8)
forecast.hw
```

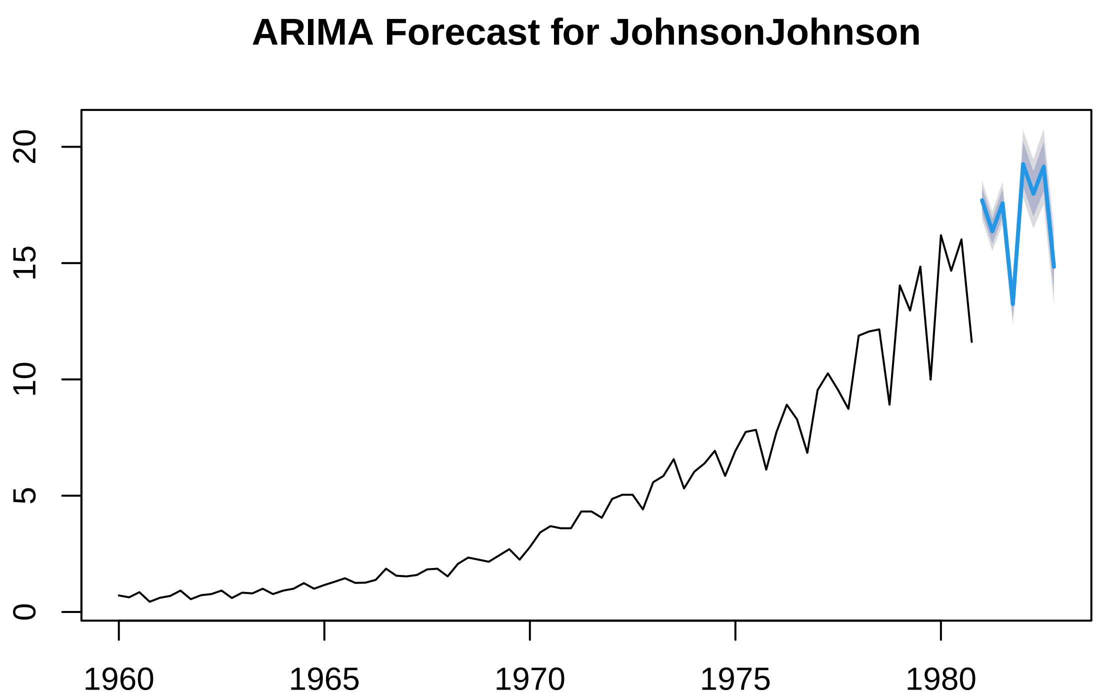
```
##         Point Forecast    Lo 80    Hi 80    Lo 95    Hi 95
## 1981 Q1       15.58283 14.29926 16.86641 13.61977 17.54590
## 1981 Q2       15.70254 14.39788 17.00720 13.70723 17.69785
## 1981 Q3       15.82224 14.47134 17.17315 13.75622 17.88827
## 1981 Q4       15.94195 14.51253 17.37137 13.75584 18.12805
## 1982 Q1       16.06165 14.51753 17.60578 13.70011 18.42319
## 1982 Q2       16.18136 14.48563 17.87709 13.58796 18.77475
## 1982 Q3       16.30106 14.41850 18.18362 13.42193 19.18019
## 1982 Q4       16.42077 14.31904 18.52249 13.20646 19.63507
```

```
plot(forecast.hw,main="HW Forecast for JohnsonJohnson")
```



(b)

```
fit.arima <- auto.arima(JohnsonJohnson)
forecast.arima <- forecast(fit.arima, h = 8)
plot(forecast.arima, main = "ARIMA Forecast for JohnsonJohnson")
```



(c)

```r
validation <- tail(JohnsonJohnson, 8)
training <- head(JohnsonJohnson, -8)
fit.hw.training <- hw(training)
forecast.hw.training <- forecast(fit.hw.training, h = 8)
fit.arima.training <- auto.arima(training)
forecast.arima.training <- forecast(fit.arima.training, h = 8)
rmse.hw <- sqrt(mean((forecast.hw.training$mean - validation)^2))
mae.hw <- mean(abs(forecast.hw.training$mean - validation))
mape.hw <- mean(abs((forecast.hw.training$mean - validation)/validation)) * 100
rmse.arima <- sqrt(mean((forecast.arima.training$mean - validation)^2))
mae.arima <- mean(abs(forecast.arima.training$mean - validation))
mape.arima <- mean(abs((forecast.arima.training$mean - validation)/validation)) * 100
cbind(rmse.hw,mae.hw,mape.hw,rmse.arima,mae.arima,mape.arima)
```

```
##       rmse.hw   mae.hw  mape.hw rmse.arima mae.arima mape.arima
## [1,] 1.346486 1.042347 6.952342  0.8401643      0.72   4.985207
```

As the result showed, the ARIMA model has lower values of RMSE,MAE and MAPE.