Time Series - Homework 3

Matt Isaac

October 16, 2018

Exercise 1: (Brockwell and Davis 5.3)

Consider the AR(2) process X_t satisfying

$$X_t - \phi X_{t-1} - \phi^2 X_{t-2} = Z_t$$

a. For what values of ϕ is this a causal process?

We define

$$\phi(z) = 1 - \phi z - \phi^2 z^2.$$

The AR(2) process given is causal for all values of ϕ for which the roots of $\phi(z) \leq |1|$.
Using the quadratic formula to find the roots of $\phi(z) = 1 - \phi z - \phi^2 z^2$ and imposing the needed constraint, we obtain the following:

$$\frac{-(-\phi) \pm \sqrt{(-\phi)^2 - 4(-\phi^2)(1)}}{2(-\phi^2)} \leq |1|$$

$$\implies \frac{\phi \pm \sqrt{5\phi^2}}{-2\phi^2} \leq |1|$$

$$\implies \frac{\phi \pm \sqrt{5}\phi}{-2\phi^2} \leq |1|$$

$$\implies \frac{1 \pm \sqrt{5}}{-2\phi} \leq |1|$$

This yields the following cases:

Case 1 :

$$\frac{1 \pm \sqrt{5}}{-2\phi} \geq 1$$

$$\implies \phi \geq -1.618$$

or

$$\implies \phi \leq 0.618$$

Case 2:

$$\frac{1 \pm \sqrt{5}}{-2\phi} \leq -1$$

$$\implies \phi \leq 1.618$$

or

$$\implies \phi \geq -0.618$$

Thus, $\mathbf{-0.618 \leq \phi \leq 0.618}$.

b. The following sample moments were computed after observing $X_1, ..., X_{200}$:

$$\hat{\gamma}(0) = 6.06,\ \hat{\rho}(1) = 0.687$$

Find estimates of ϕ and σ^2 by solving the Yule-Walker equations.

Since

$$\hat{\rho}(1) = \frac{\hat{\gamma}(1)}{\hat{\gamma}(0)},$$

we can see that

$$\hat{\gamma}(1) = \hat{\rho}(1)\hat{\gamma}(0).$$

Thus,

$$\hat{\gamma}(1) = (0.687)(6.06) = 4.163.$$

Recalling the system of linear equations,

$$\hat{\gamma}(0) - \phi_1\hat{\gamma}(0) - \phi_2\hat{\gamma}(2) = \hat{\sigma}^2$$
$$\hat{\gamma}(1) - \phi_1\hat{\gamma}(0) - \phi_2\hat{\gamma}(1) = 0$$
$$\hat{\gamma}(2) - \phi_1\hat{\gamma}(1) - \phi_2\hat{\gamma}(0) = 0,$$

Recalling that $\phi_2 = \phi_1^2$ we can solve for $\hat{\phi}$ using the second equation above:

$$\hat{\gamma}(1) - \phi_1\hat{\gamma}(0) - \phi_2\hat{\gamma}(1) = 0$$

$$\implies 4.163 - \phi(6.06) - \phi^2(4.163) = 0.$$

Rearranging into a more recognizable form,

$-4.163\phi^2 - 6.06\phi + 4.163 = 0$

```
quadratic <- polyroot(c(4.163, -6.06, -4.163))
Mod(quadratic)
```

```
## [1] 0.5089909 1.9646719
```

Obtaining the roots of the equation above, we see that $\phi = 0.509, 1.965$. In order to determine which value of ϕ results in a causal process, we will check to see which value of ϕ results in the following equation having roots outside of the unit circle:

$\phi(z) = 1 - \phi z - \phi^2 z^2$

```
phiz.1 <- polyroot(c(1, -0.509, -(0.509^2)))
Mod(phiz.1)
```

```
## [1] 1.214212 3.178849
```

```
phiz.2 <- polyroot(c(1, -1.965, -(1.965^2)))
Mod(phiz.2)
```

```
## [1] 0.3145211 0.8234270
```

So, $\phi = 0.509$ gives the equation $\phi(z) = 1 - 0.509z - 0.509^2 z^2$, which has roots of 1.214 and 3.179. No roots exist inside the unit circle, so this is the value of ϕ which results in a causal process. On the other hand, if we let $\phi = 1.965$, we do get roots of $\phi(z)$ which are inside the unit circle, which makes the process not causal.

Next, we will solve for $\hat{\gamma}(2)$, by solving the last equation in the system of linear equations:

$$\hat{\gamma}(2) - \phi_1\hat{\gamma}(1) - \phi_2\hat{\gamma}(0) = 0,$$

$$\implies \hat{\gamma}(2) - (0.509 * 4.163) - (0.509^2 * 6.06) = 0$$

$$\implies \hat{\gamma}(2) - 2.119 - 1.570 = 0$$

$$\implies \hat{\gamma}(2) = 3.689.$$

We can then solve for σ^2 by solving the first equation in the system of linear equations:

$$\hat{\gamma}(0) - \phi_1\hat{\gamma}(0) - \phi_2\hat{\gamma}(2) = \hat{\sigma}^2$$

$$\implies 6.06 - (0.509 * 4.163) - (0.509^2 * 3.689) = \hat{\sigma}^2$$

$$\implies \hat{\sigma}^2 = 2.985$$

In summary, $\hat{\phi} = \mathbf{0.509}$ and $\hat{\sigma}^{\mathbf{2}} = \mathbf{2.985}$.

Exercise 2

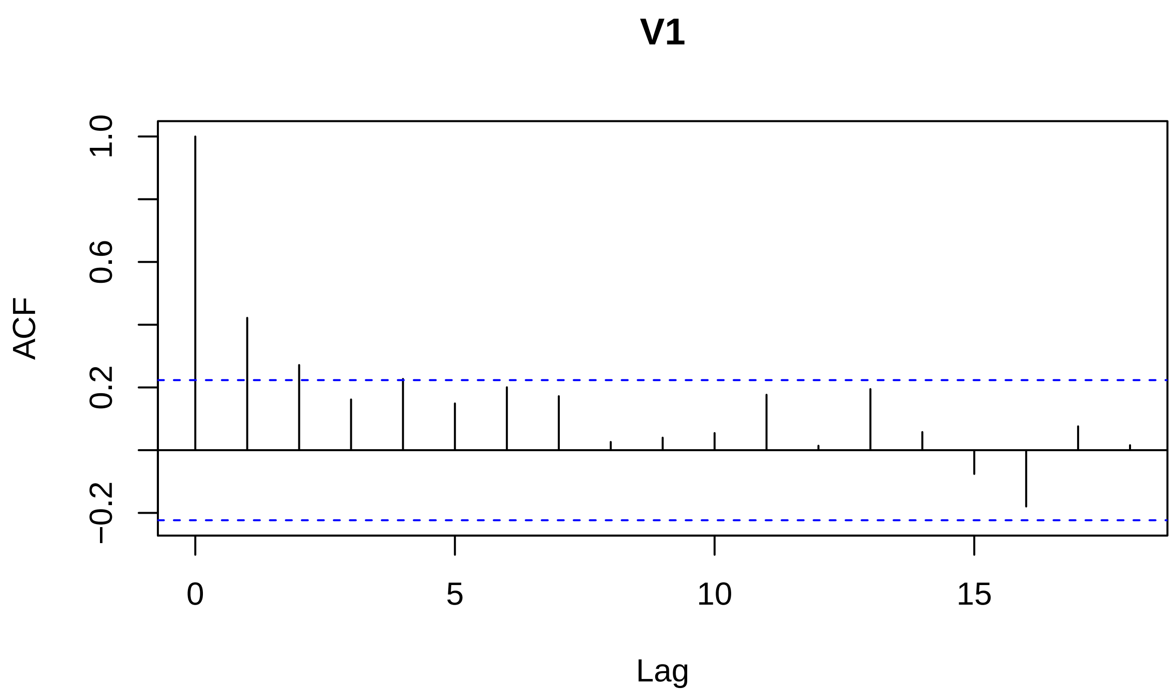
Consider fitting an AR(p) model to the Dow Jones Utility Index data.

a. Decide on the value of p. Justify your choice.

First, I read in the raw data and created a time series object. I next subtracted off the mean, and plotted the ACF to check stationarity. Because we see that the ACF decreases rapidly and lacks any complicated structure, we can safely assume that the process is stationary.

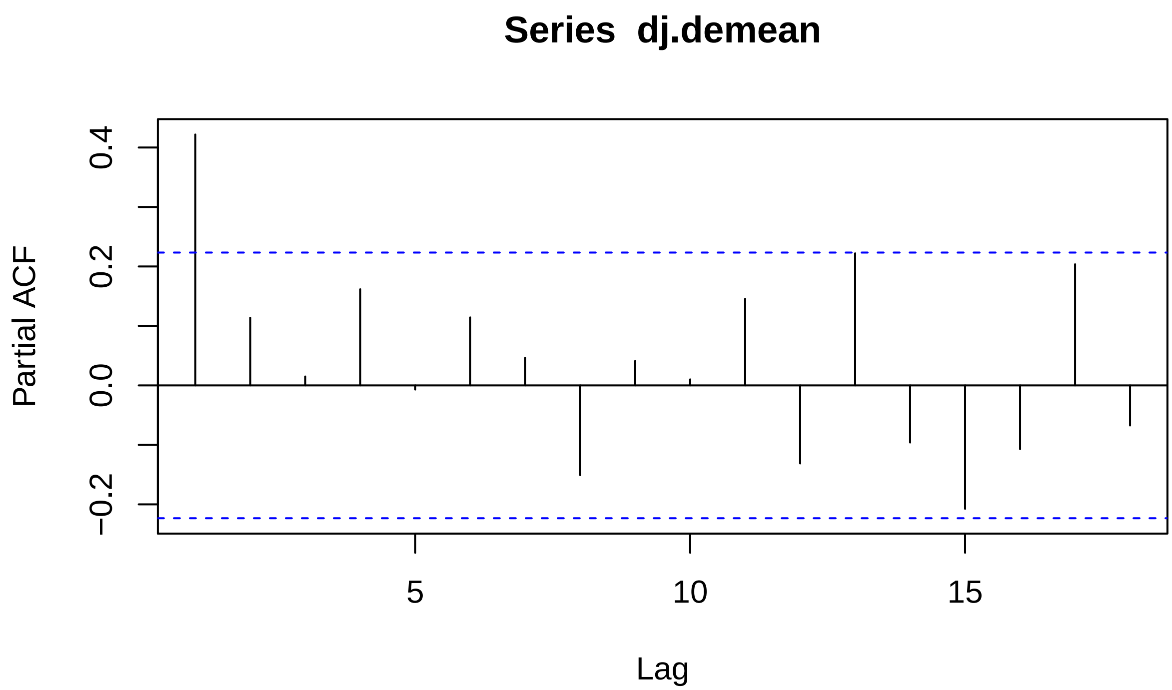
```
dj.raw <- read.csv("dj_index.csv", header = FALSE) # read in raw data
dj <- ts(dj.raw, start = 1, frequency = 1) # create time series object

dj.demean <- dj - mean(dj) # subtract off mean
acf(dj.demean) # check stationarity: stationarity acceptable.
```



Next, I plotted the PACF to determine what value to use for p.

```
pacf(dj.demean) # use acf to choose value of p
```



The only significant lag is at $p = 1$, so we will fit an AR(1) model.

b. Estimate the parameters (using one of the three methods we learned). Perform appropriate tests on whether the estimated AR coefficients are significant.

The R code below fits an AR(1) model, esitmating ϕ_1 with the 'CSS-ML' method.

```
# estimate parameters
dj.ar <- arima(dj.demean, order = c(1, 0, 0), method = "CSS-ML", include.mean = FALSE)
dj.ar$coef
```

```
##       ar1
## 0.4471124
```

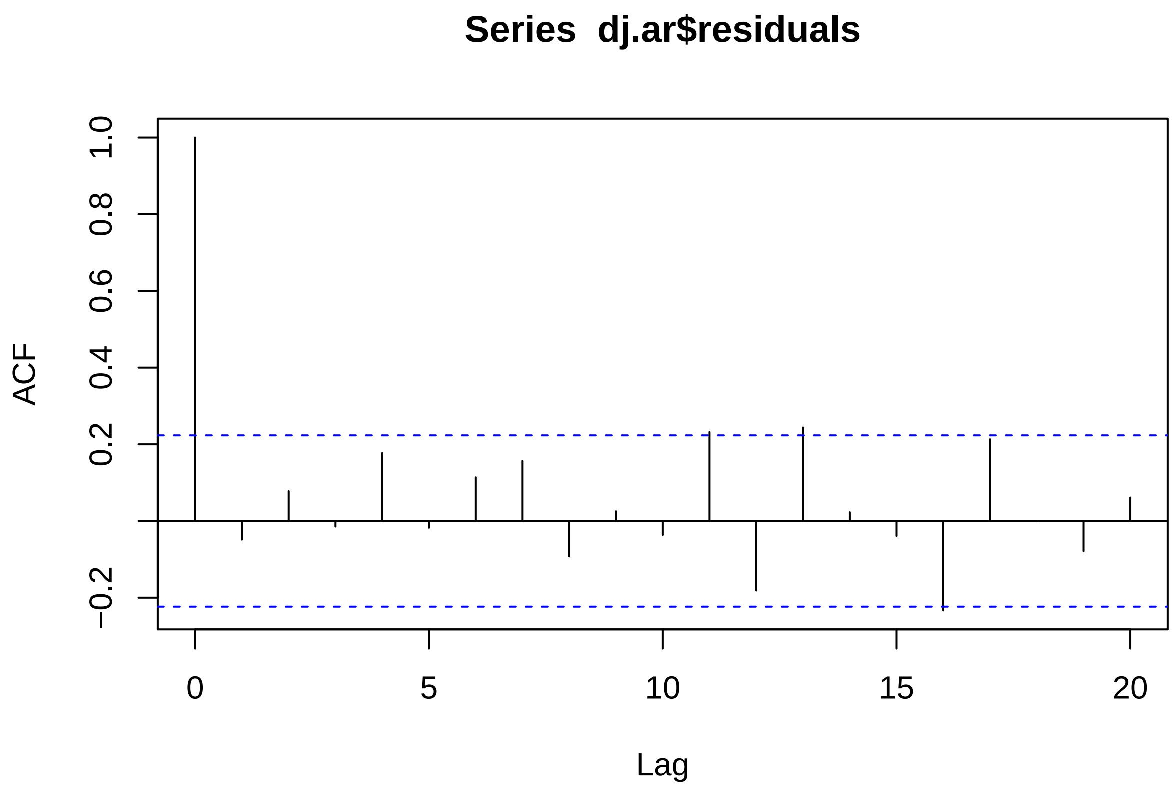
```
# appropriate tests for sig. of coeffs.
se <- sqrt(diag(dj.ar$var.coef)) #

CI_low <-  dj.ar$coef-1.96*se #
CI_upper = dj.ar$coef+1.96*se
```

The value obtained by the estimation method for ϕ_1 is 0.447. The confidence interval is (0.241, 0.653), which does not include 0, indicating that the parameter ϕ_1 is significant.

c. What do you learn from the residuals?

```
# check residuals to see if model fits data
acf(dj.ar$residuals, lag.max = 20)
```




```
lb.test <- Box.test(dj.ar$residuals, lag = 20, type = "Ljung-Box")
```

Plotting the ACF of the residuals, we see that there is not much indication of autocorrelation in the residuals before lag 11. It appears that lag 11, 13, and 16 may be slightly significant. Performing the Ljung-Box test can give us a second 'opinion' as to how well our model fits the data. The pvalue obtained from the test is 0.038. So both the ACF of the residuals and the Ljung-Box test indicate that the residuals are not completely white noise. However, recalling the PACF plot obtained from the data, we remember that we would have to extend our model to AR(13) or AR(15) in order to capture the significant lags. There is a tradeoff between model simplicity and how well it fits the data. Because we only had marginal significance in the Ljung-Box test, we will continue with the AR(1) model.

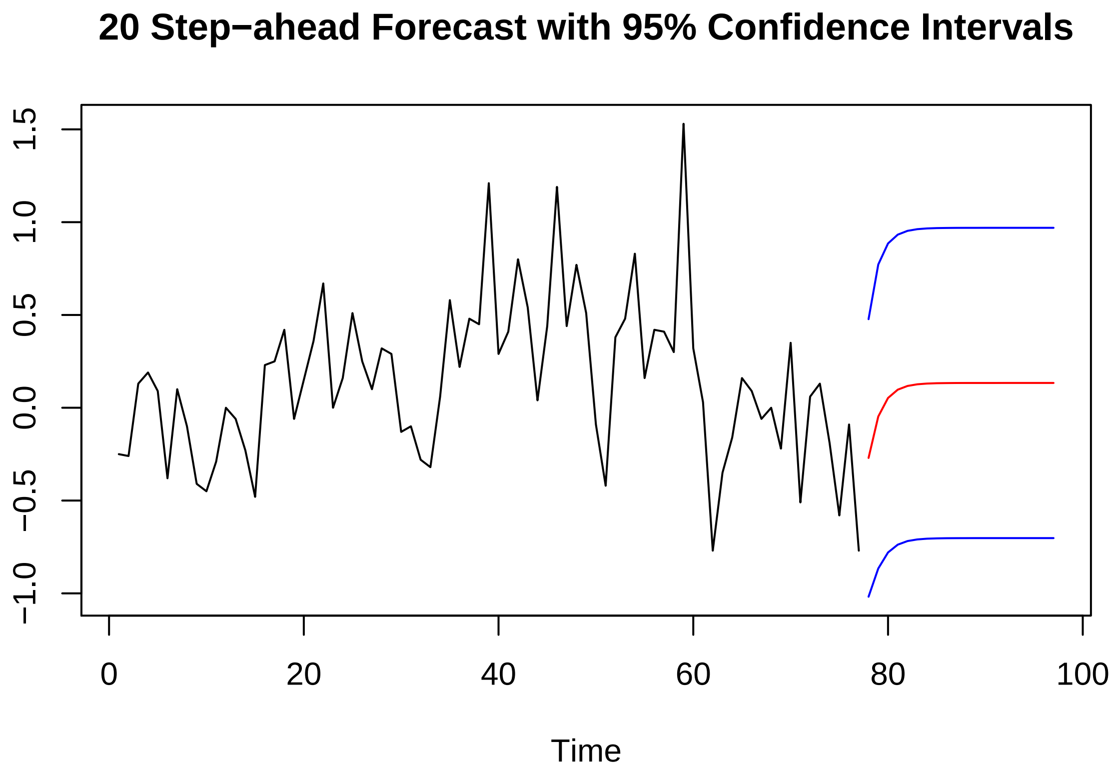
d. Make forecasts up to 20-step-ahead and calculate the 95% confidence intervals.
   This is performed in the R code below (see comments).

```
dj.predict = predict(dj.ar, n.ahead = 20, se.fit =TRUE) # 20 step ahead forecast

dj.predict.value <- dj.predict$pred + mean(dj) # add the mean back on to the predictions
cl.predict.upper <- dj.predict.value + 1.96 * dj.predict$se # calculate CI upper bounds
cl.predict.lower <- dj.predict.value - 1.96 * dj.predict$se # calculate CI lower bounds
```

e. Plot the forecasts and their 95% confidence intervals and the observed data in the same figure.
   Performed in the R code below:

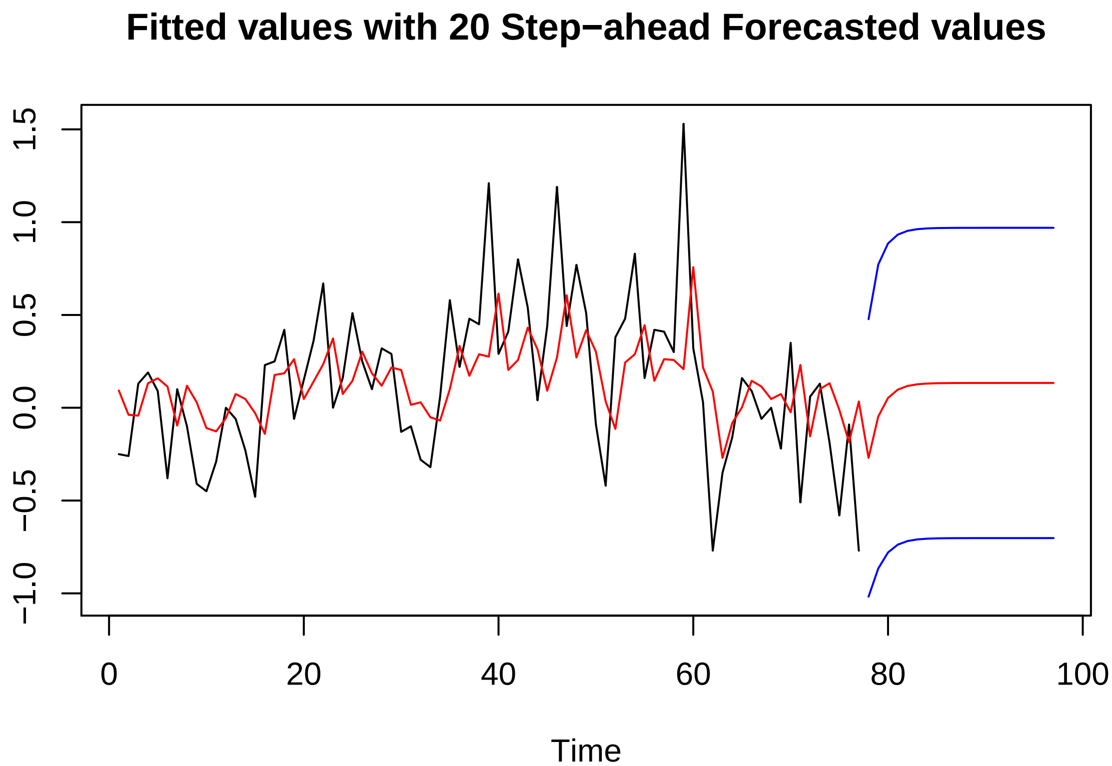
```
ts.plot(dj, dj.predict.value, cl.predict.upper, cl.predict.lower,
        gpars = list(col = c("black", "red", "blue", "blue")),
        main = "20 Step-ahead Forecast with 95% Confidence Intervals")
```



And lastly, just for fun, here is a plot of the fitted values (including forecasted values and confidence intervals) along with the observed time series values.

```
#plot fitted and predicted values together with the observed data
dj.fitted.pred=ts(rep(0,97),start=1,frequency=1)
dj.fitted.pred[1:77] <- dj-dj.ar$residuals
dj.fitted.pred[78:97]=dj.predict.value[1:20]

ts.plot(dj, dj.fitted.pred, cl.predict.upper, cl.predict.lower,
        gpars = list(col = c("black", "red", "blue", "blue")),
        main = "Fitted values with 20 Step-ahead Forecasted values")
```


Fitted values with 20 Step−ahead Forecasted values
1.5
1.0
0.5
0.0
−0.5
−1.0
0
20
40
60
80
100
Time