Time Series Practical 2 Hilary Term 2010 Prof. Reinert

In this practical we shall have a go at the Johnson & Johnson quarterly earnings per share time series from lectures. The series is compiled of 84 observations (21 years) measured from the first quarter of 1960 to the last quarter of 1980; a quarter is 3 months. The data are also discusses in the book by Shumway and Stoffer, Time Series Analysis and Its Applications.

Download the data, called `jj.dat`, either from the course website or at `http://www.stat.pitt.edu/stoffer/tsa2/`.

If you store the data in a directory called P:/, then use

```
jj<-scan("P:/jj.dat")
```

to read the data into `R`. To plot the data, you could use

```
jj<-ts(jj, start=1960, frequency=4)
ts.plot(jj)
```

Does this series look stationary? You could try a transformation such as taking logs;

```
jj.log<-ts(log(jj), start=1960, frequency=4)
ts.plot(jj.log)
```

Is there a trend? Try for example differencing; `jj.diff<-diff(jj.log)`. Is there seasonality? Look at

```
par(mfrow=c(2,1))
acf(jj.diff)
pacf(jj.diff)
```

or at

```
spectrum(jj.diff)
spectrum(jj.diff, spans=c(3,5))
```

If you think that there may be annual periodicity, try `jj.diff2<- diff(jj.diff, lag=4)`. To use the Box-Jenkins approach for fitting a time series, look at the acf and the pacf of the series. Then, for example if you think that an MA model would be a good idea, then you could use for example

```
aicres<-c(1:14)
for(i in 1:14){ aicres[i]<- arima(jj.diff2, order=c(0,0,i),n.cond=15)
$aic }
aicres
```

Which model would you select? Check its diagnostic plots. For example, if you choose an MA(13) model, then you could use

```
fit<- arima(jj.diff2, order=c(0,0,13))
tsdiag(fit)
```

Note that an MA(13) model may not be the best model; which criteria do you use to choose a good model? Remember also that `cpgram` is the command to give you a cumulative periodogram.

Once you have selected a model, you can check how well it forecasts future values. To this purpose we shorten the time series by one year, and see how well we can predict the last year. Put

```
years<-seq(1960, 1980+3/4, by=1/4)
jj.short<-jj.log[1:80] years.short<-years[1:80]
```

and fit a model for the series `jj.short` using the above procedure (but for the shorter time series). Write your model as a SARIMA model. For example, if you difference once and then you take the 4-lag difference and you find that the best fitting model is MA(13) (which may indeed not be the model you find to fit best!) then you could use

```
orig.fit.short<-arima(jj.short, order=c(0,1,13), seasonal=list(order=c(0,1,0
), period=4 ), n.cond=4)
tsdiag(orig.fit.short)
```

to fit a SARIMA model to the data. Check the residuals; are you satisfied with the model diagnostics? Note that this fit does not include an intercept. Is the intercept significant in the model you have fitted, or could you fit a model without the intercept also?

Using again an MA(13) fit, to plot the forecasts, you could use

```
fyears<-years[81:84]
pred.1<-predict(arima(jj.short, order=c(0,1,13), seasonal=list(order=c(0,1,0
), period=4 ), n.cond=4) , n.ahead=4)
pred.1
forecast.1<-pred.1$pred
flimits.1<-1.96* round(pred.1$se, 3)
```

and then

```
plot(years.short, jj.short, type="l", xlab="year", ylab="log earnings",
xlim=range(years))
lines(fyears, forecast.1, lwd=2)
lines(fyears, forecast.1 + flimits.1, lty=2)
lines(fyears, forecast.1 - flimits.1, lty=2)
```

and compare it to the original time series,

```
plot(jj.log, type="l", xlab="year", ylab="log earnings", xlim=range(years))
lines(fyears, forecast.1, col="red")
lines(fyears, forecast.1 + flimits.1, lty=2, col="blue") lines(fyears,
forecast.1 - flimits.1, lty=2, col="blue")
```

Finally, to compare different models, you could calculate the mean squared prediction error

`mse<- mean(`$(\exp(forecast.1) - \exp(jj.log[81:84]))$^2)