Practice 8: Simulating and fitting ARMA models

The goal of this practice is to identify a possible ARMA model to a given data, to fit it and to validate it as a good model.

The outline is the following:

1. Simulation of a model.
2. Heuristic identification of a model. Auto-correlation and partial auto-correlation functions.
3. Fitting a model.
4. Validation.

1. Simulation of a model.

Do

```
SIM1<-arima.sim(list(order=c(1,0,0), ar=0.7),n=1000)

SIM2<-arima.sim(list(order=c(0,0,2), ma=c(1.5,0.75)),n=1000)

SIM3<-arima.sim(list(order=c(2,0,1), ar=c(0.3,0.2), ma=0.3), n=1000)
```

We have simulated an AR(1) model with parameter 0.7, an MA(2) model with parameters –1.5 and –0.75, and an ARMA(2,1) model with parameters AR 0.3 and 0.2 and parameter MA –0.3.

Make the charts of SIM1, SIM2 and SIM3.

2. Heuristic identification of the model. Auto-correlation and partial auto-correlation functions.

Here we want to compute and represent graphically functions acf and pacf of our series. If x represents any of our series, we have to write

```
acf(x,k)
pacf(x,k)
```

where k represents the number of represented lags. Recall that k has to be at most one third of n, the number of simulated data. Some authors recommend at most k=ln(n).

Recall that an AR(p) model has an exponentially decreasing acf and a pacf that is null after p lags. An MA(q) model has an exponentially decreasing pacf and an acf that is null after q lags.

3. Fitting a model.

Observe the graphics of acf and pacf of the three series and deduce heuristically a good model. Afterwards fit the proposed model using the instruction

```
S1<-arima(SIM3, order=c(1,0,0), method=”ML”)
S1
```

And similarly, for the other two. Check if the estimations are good.

4. Validation.

We can apply the following instructions to S3:

```
x<-SIM3
h<-log(n)
Box.test (x, lag = h, type = c("Ljung-Box"))
```

Here we reject the null hypothesis; SIM3 is not an IID noise!

But we can compute

```
y<-resid (S3)
Box.test (y, lag =h, type = c("Ljung-Box"))
```

and now the p-value is greater than 0.05 and so we can accept the series of residuals is an IID noise. Thus, we validate our ARMA(2,0,1) model.

A faster way to do it is using the instruction

```
tsdiag(S3, gof.lag=20)
```

See the graphics. The third one shows clearly that data are well fitted; the p-values for all the lags are high. Option gof.lag indicates until what order we compute the statistics of Ljung-Box test.

5. A complete example:

Do:

```
n<-10000
h<-log(n)
s3<-arima.sim(list(order=c(2,0,1),ar=c(0.3,0.2),ma=0.3), n)
plot(s3, type="l")
acf(s3)
pacf(s3)
p3<-arima(s3,order=c(2,0,1),method="ML")
p3
Box.test(p3,lag=h,type=c("Ljung-Box"))
y<-resid(p3)
Box.test(y,lag=h,type=c("Ljung-Box"))
```

6. Roots of a polynomial:

Given a polynomial $1+a_1 x+a_2 x^2+\cdots\cdots+a_n x^n$ we can obtain its roots using the R instruction

Polyroot (c(1,a_1,….,a_n))