Analysis of ARMA(2,1) process

TMA4285 Time series models - Exercise 3

Anders Christiansen Sørby, Edvard Hove

Abstract

Blabla

Introduction

The ARMA process

A stationary process $\{X_t\}$ is an $ARMA(p, q)$ process if the following relation exists

$$\phi(B)X_t = \theta(B)Z_t \tag{1}$$

where $Z_t \sim \mathbf{WN}(\mathbf{0}, \sigma^2)$, and $\phi(\cdot)$ and $\theta(\cdot)$ are polynomials (with no common factors) given by

$$\phi(z) = 1 - \phi_1 z - \phi_2 z^2 - \cdots - \phi_p z^p \tag{2}$$

and

$$\theta(z) = 1 + \theta_1 z + \theta_2 z^2 + \cdots + \theta_q z^q, \tag{3}$$

and B is the backward shift operator. That is $B^j X_t = X_{t-j}, j \in \mathbb{Z}$.

In particular an $ARMA(2, 1)$ process is given by

$$X_t - \phi_1 X_{t-1} - \phi_2 X_{t-2} = Z_t + \theta_1 Z_{t-1}$$

Theory

An expression for $\{X_t\}$ satisfying equation (1) on the form

$$X_t = \sum_{j=-\infty}^{\infty} \psi_j Z_{t-j} \tag{4}$$

is said to be a stationary solution. A stationary solution exists, and is also unique, if no roots of the autoregressive polynomial $\phi(z)$ lie on the unit circle. If X_t can be expressed in terms of $Z_s, s \leq t$ we say that the process is causal. Such an expression is guaranteed to exist if no roots of $\phi(z)$ lie on the unit disk. Include how the $\phi\theta$ is estimated

Model parameter estimation

The $ARMA(2,1)$ process (with zero mean) consists of 4 parameters that need to be estimated. That is ϕ_1, ϕ_2, θ_1 and σ^2. This can be done in a number of ways. A simple and consistent way is the method of moments. The idea is to take the expected value of the expression of X_t multiplied with X_{t-h}.

$$\mathbb{E}[X_t X_{t-h}] = \mathbb{E}[(Z_t + \theta_1 Z_{t-1} + \phi_1 X_{t-1} + \phi_2 X_{t-2})X_{t-h}] \tag{5}$$

This then gives a set of equations containing the autocovariance function $\gamma(h)$ and the parameters. By using the sample autocovariance function $\hat{\gamma}(h)$ in place of the autocovariance function we can solve the equations, resulting in estimates of the parameters of interest.

The estimators found using this method may be biased and they often have high variance compared to other estimators. They are however useful as first approximations in numerical methods of estimation.

Another way is to use the maximum likelihood estimators.

REPEATED ESTIMATION ON SIMULATED DATA COMES HERE: IN ORDER TO USE arima.sim WE NEED AN EXPRESSION FOR THE INNOVATIONS

We need to confirm that a given process is indeed an $ARMA(2,1)$

Data analysis

The language R has many built-in functions for analysing timeseries.

Our recieved realization of the process can be seen in the next plot.
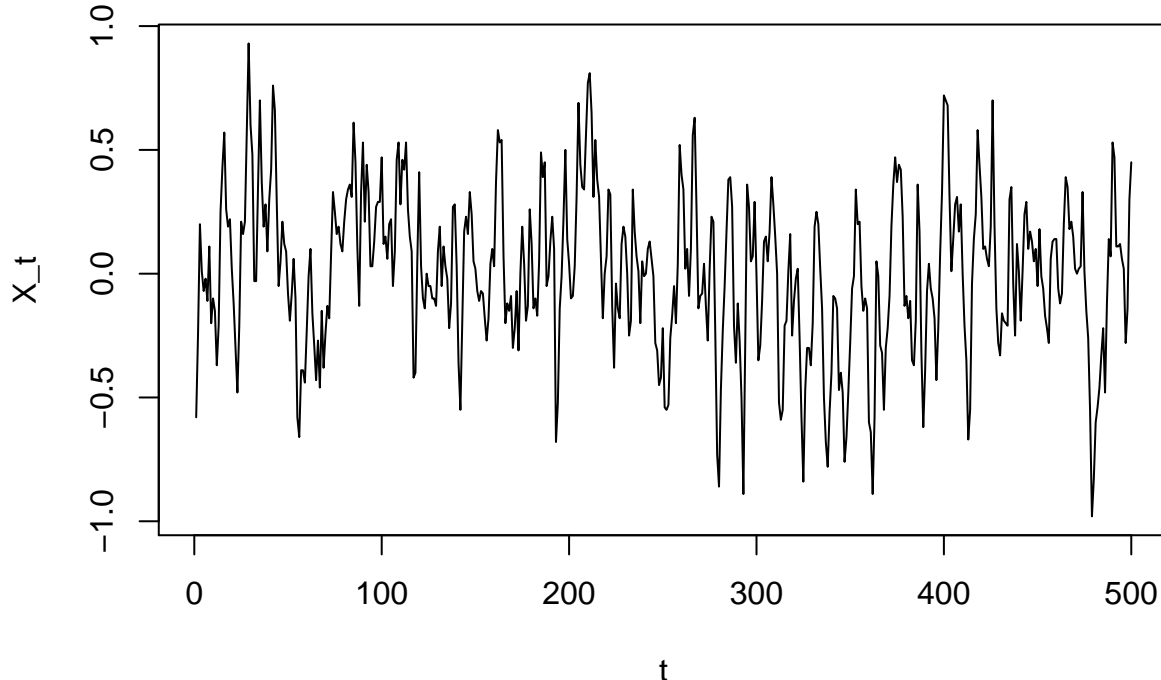

There is no obvious need for a transformation of the data to fit an ARMA model. Based on the plot it seems probable that both the mean and variance is constant with respect to time. (DETTE HADDE VÆRT FINT Å DEMONSTRERE)
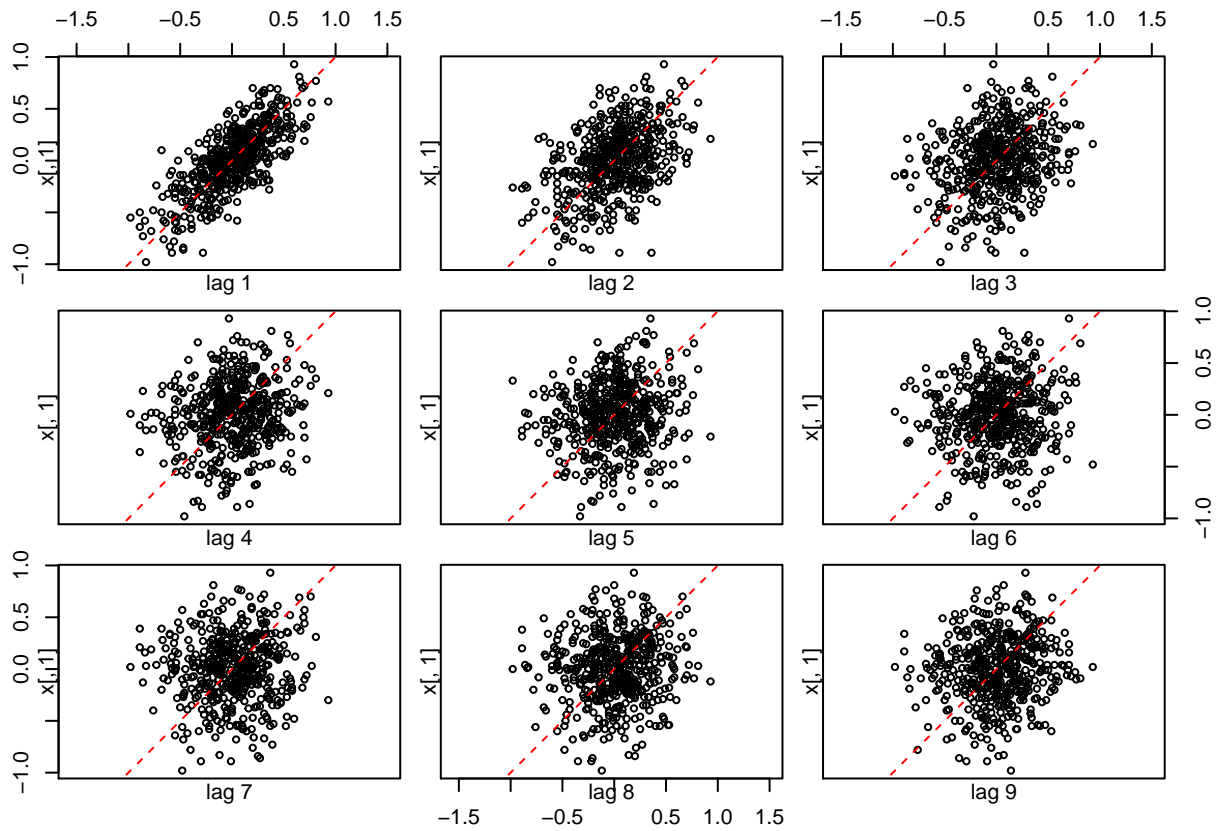
```
arima=arima(x,order=c(2,0,1))
```

```
lag.max=100
acf=acf(x,lag.max=lag.max)
lines(seq(0,lag.max),ARMAacf(ar=c(0.7411,-0.1282),ma=0.2789,lag.max=lag.max))
n=seq(0,25)
s=0.1*sin(2*pi*(n-6)/14)
lines(n,s)
```
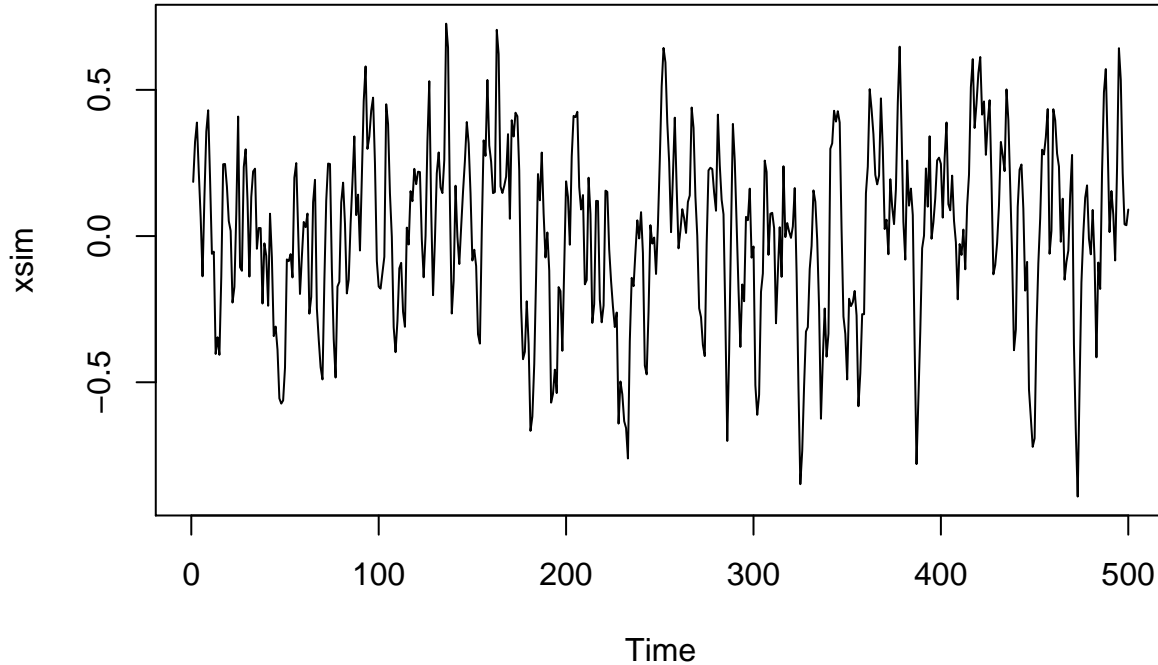
V1
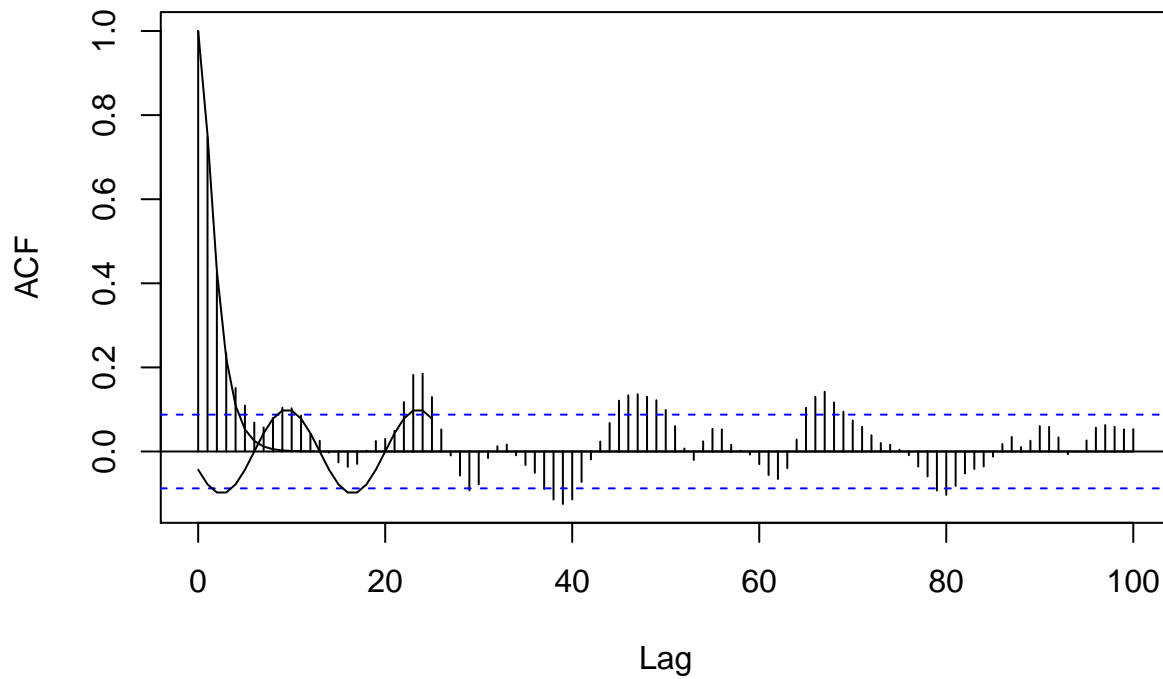

```
p=10
c=2*pi/p
y=asin(c*x[,1])
```

It is conceivable, based on the sample autocorrelation function, that there is some underlying seasonal lag. More importantly the sample ACF forces us to reject the hypothesis that the data is generated by an $MA(q)$ model (atleast for $q \leq 50$).

```
xsim <- arima.sim(n=500, list(ar=c(arima$coef[1],arima$coef[2]), ma=arima$coef[3]),innov=rnorm(1000, sd=
plot(xsim)
```

Using the function `arima` we get MLE parameter estimates for our the underlying distribution.

```
print(arima)
```

```
##
## Call:
## arima(x = x, order = c(2, 0, 1))
##
## Coefficients:
##          ar1      ar2     ma1  intercept
##       0.7411  -0.1282  0.2789     0.0045
## s.e.  0.1216   0.1000  0.1181     0.0295
##
## sigma^2 estimated as 0.04004:  log likelihood = 94.49,  aic = -178.98
```

We can then check the causality criterion by solving

$$\phi(z) = 1 - 0.741z + 0.128z^2 = 0, \tag{6}$$

which gives us the roots $z_1 \approx 2.145$ and $z_2 \approx 3.637$. Since both these roots are outside the unit disk there exists an expression for X_t on the form

$$X_t = \sum_{i=0}^{\infty} \psi_i Z_{t-i} \tag{7}$$

Model prediction

Discussion

Bla bla

Conclusion

This was ..

Appendix

Everything not in the main section This also includes R-code as a whole

References