Simulation of AR(p) and Seasonal processes

Assignment 2 – 02417 Time Series Analysis – Anders Hørsted (s082382)

This report consists of four parts. In the first part a few theoretical results for the AR(2) process are obtained and a prediction is made for a seasonal model. In part two a few different techniques for simulation of AR-processes are tested and the results from using different coefficients in the lag polynomial are compared and commented on. In part three the random walk process is simulated and plotted along with the autocorrelation function of the process. Finally in part four a couple of seasonal models are simulated and the autocorrelation function is plotted.

Part 1: Theory and prediction of seasonal model

The AR(2)-process

We consider the AR(2)-process

$$X_t + \phi_1 X_{t-1} + \phi_2 X_{t-2} = \varepsilon_t$$

where ε_t is white noise. First the parameters for which the AR(2)-process is stationary are determined.

From theorem 5.9 in [1] the roots of the equation $\phi(z^{-1}) = 0$, with respect to z, should all lie within the unit circle for the AR(2)-process to be stationary. Now assuming $z = 0$ we get

$$\phi(z^{-1}) = \phi_2 z^{-2} + \phi_1 z^{-1} + 1$$

so the roots are found from

$$\begin{aligned} \phi_2 z^{-2} + \phi_1 z^{-1} + 1 = 0 \quad &\Leftrightarrow \\ z^2 + \phi_1 z + \phi_2 = 0 \end{aligned}$$

With $D = \phi_1^2 - 4\phi_2$ we find the roots as

$$z_1 = \frac{-\phi_1 + \sqrt{D}}{2}, \quad z_2 = \frac{-\phi_1 - \sqrt{D}}{2}$$

We now look at the three cases $D = 0, D < 0, D > 0$ one at a time.

First $D = 0$. Then we get a real double root given by $z_1 = z_2 = -\frac{\phi_1}{2}$ and since this root should be within the unit circle, we get the inequality

$$\frac{|\phi_1|}{2} < 1 \quad \Leftrightarrow \quad \phi_1 \in (-2, 2)$$

Since $D = 0 \Leftrightarrow \phi_1^2 = 4\phi_2$ we get the first set

$$M_1 = \{(\phi_1, \phi_2) \,|\, \phi_1^2 = 4\phi_2, \phi_1 \in (-2, 2)\}$$

for which the AR(2)-process is stationary

For the second case $D < 0$ we get two complex roots given by

$$z_1 = \frac{-\phi_1}{2} + \frac{\sqrt{4\phi_2 - \phi_1^2}}{2} i, \quad z_2 \frac{-\phi_1}{2} - \frac{\sqrt{4\phi_2 - \phi_1^2}}{2} i$$

where both roots have the same modulus $|z_1| = |z_2| = \sqrt{\phi_2}$. Therefore to get stationarity $\sqrt{\phi_2} < 1 \Leftrightarrow \phi_2 \in [0, 1)$, which combined with $D < 0 \Leftrightarrow \phi_1^2 < 4\phi_2$ gives $\phi_1 \in (-2, 2)$ and the second parameter set becomes

$$M2 = \{(\phi_1, \phi_2) \,|\, \phi_1^2 < 4\phi_2, \phi_1 \in (-2, 2)\}$$

For the last case $D > 0$ the root of interest depends on the sign of ϕ_1. If $\phi_1 < 0$ we look at z_1 and for $\phi_1 > 0$ we look at z_2. Both cases can be handled by one inequality given by

$$\begin{aligned} \frac{|\phi_1| + \sqrt{\phi_1^2 - 4\phi_2}}{2} &< 1 \quad \Leftrightarrow \\ |\phi_1| + \sqrt{\phi_1^2 - 4\phi_2} &< 2 \end{aligned}$$

From this we get that $\phi_1 \in (-2, 2)$. We now look at all lines for which $\phi_2 = |\phi_1| - 1$, giving

$$\begin{aligned} |\phi_1| + \sqrt{\phi_1^2 - 4\phi_2} &= |\phi_1| + \sqrt{\phi_1^2 - 4(|\phi_1| - 1)} \\ &= |\phi_1| + \sqrt{(|\phi_1| - 2)^2} \\ &= |\phi_1| + \big|\, |\phi_1| - 2 \,\big| \\ &= 2 \end{aligned}$$

using that for $\phi_1 \in (-2, 2)$, $\big|\,|\phi_1| - 2\,\big| = 2 - |\phi_1|$. From the expression for D we conclude that $\phi_2 > |\phi_1| - 1$. Combined with $D > 0 \Leftrightarrow \frac{1}{4}\phi_1^2 > \phi_2$ the last set is

$$M_3 = \{(\phi_1, \phi_2) \,|\, |\phi_1| - 1 < \phi_2 < \frac{1}{4}\phi_1^2, \phi_1 \in (-2, 2)\}$$

The final result is that the AR(2)-process is stationary for parameters in the set

$$M = M_1 \cup M_2 \cup M_3 = \{(\phi_1, \phi_2) \,|\, |\phi_1| - 1 < \phi_2 < 1\}$$

The set M is sketched in figure **??**.

We now need to determined the set of parameters for which the autocorrelation function of the AR(2)-process shows damping harmonic oscillations. Based on equation 5.83

and the surrounding text in [1] the autocorrelation function shows damping harmonic oscillations exactly when the roots of the characteristic equation are complex conjugates. Using the result from the previous paragraphs the set where the autocorrelation function shows damping harmonic oscillations is given by

$$M2 = \{(\phi_1, \phi_2) \mid \phi_1^2 < 4\phi_2, \phi_1 \in (-2, 2)\}$$

The set $M2$ is sketched in figure **??**

Prediction in seasonal model

We now model the deviation Y_t between observed and calculated water levels as a linear model given by

$$(1 - 0.8B)(1 - 0.2B^6)(1 - B)Y_t = \varepsilon_t \tag{1}$$

where ε_t is white noise with variance σ_ε^2. An estimate based on around 1500 observations is calculated as $\hat{\sigma}_\varepsilon^2 = 0.31\,\mathrm{dm}^2$. Based on the given data $D = \{Y_1, Y_2, \ldots, Y_{10}\}$ shown in table 1 in the assignment paper, we will predict Y_{12}.

First the operator polynomial in model (1) is expanded giving

$$(1 - 1.8B + 0.8B^2 - 0.2B^6 + 0.36B^7 - 0.16B^8)Y_t = \varepsilon_t$$

that rewritting gives

$$Y_t = 1.8Y_{t-1} - 0.8Y_{t-2} + 0.2Y_{t-6} - 0.36Y_{t-7} + 0.16Y_{t-8} + \varepsilon_t$$

We can now predict Y_t at time $t = 11$ by

$$\begin{aligned}
\widehat{Y}_{11} &= \mathrm{E}[Y_{11} \mid D] \\
&= 1.8Y_{10} - 0.8Y_9 + 0.2Y_5 - 0.36Y_4 + 0.16Y_3 \\
&= -5.44
\end{aligned}$$

The prediction $\widehat{Y}_{11}$ is then used to find $\widehat{Y}_{12}$

$$\begin{aligned}
\widehat{Y}_{12} &= \mathrm{E}[Y_{12} \mid D] \\
&= 1.8\widehat{Y}_{11} - 0.8Y_{10} + 0.2Y_6 - 0.36Y_5 + 0.16Y_4 \\
&= -6.43
\end{aligned}$$

To get a 95% confidence interval for $\widehat{Y}_{12}$ we need to find the variance of the prediction error $Y_{12} - \widehat{Y}_{12}$.

$$\begin{aligned}
\mathrm{Var}[Y_{12} - \widehat{Y}_{12} \mid D] &= \mathrm{Var}[1.8Y_{11} + \varepsilon_{12} \mid D] \\
&= \mathrm{Var}[1.8(1.8Y_{10} - 0.8Y_9 + 0.2Y_5 - 0.36Y_4 + 0.16Y_3 + \varepsilon_{11}) + \varepsilon_{12} \mid D] \\
&= \mathrm{Var}[1.8\varepsilon_{11} + \varepsilon_{12}] \\
&= (1.8^2 + 1)\sigma_\varepsilon^2
\end{aligned}$$

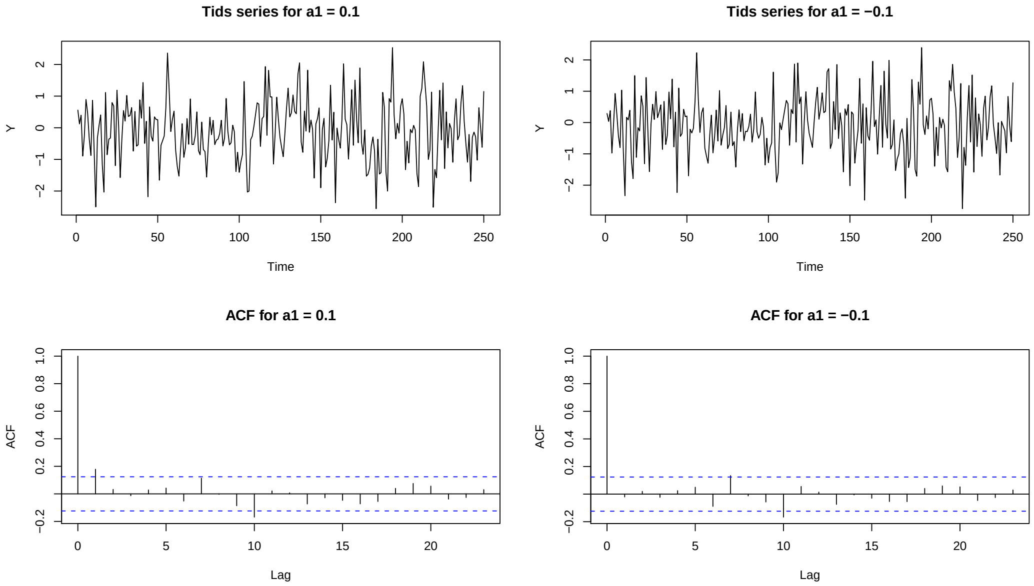

Figure 1: Picture time

Since σ_ε^2 is unknown we use the estimate $\hat{\sigma}_\varepsilon^2$ instead which gives

$$\begin{aligned}\mathrm{Var}[Y_{12} - \widehat{Y}_{12} \mid D] &= (1.8^2 + 1)\hat{\sigma}_\varepsilon^2 \\ &= 1.31\end{aligned}$$

And finally we get the 95% confidence interval for $\widehat{Y}_{12}$ as

$$-6.43 \pm 1.96 \cdot \sqrt{1.31} = [-8.68; -4.18]$$

Part 2: Simulating AR-processes

In this part we will simulate a few different AR-processes using the `filter` and `arima.sim` functions in R.

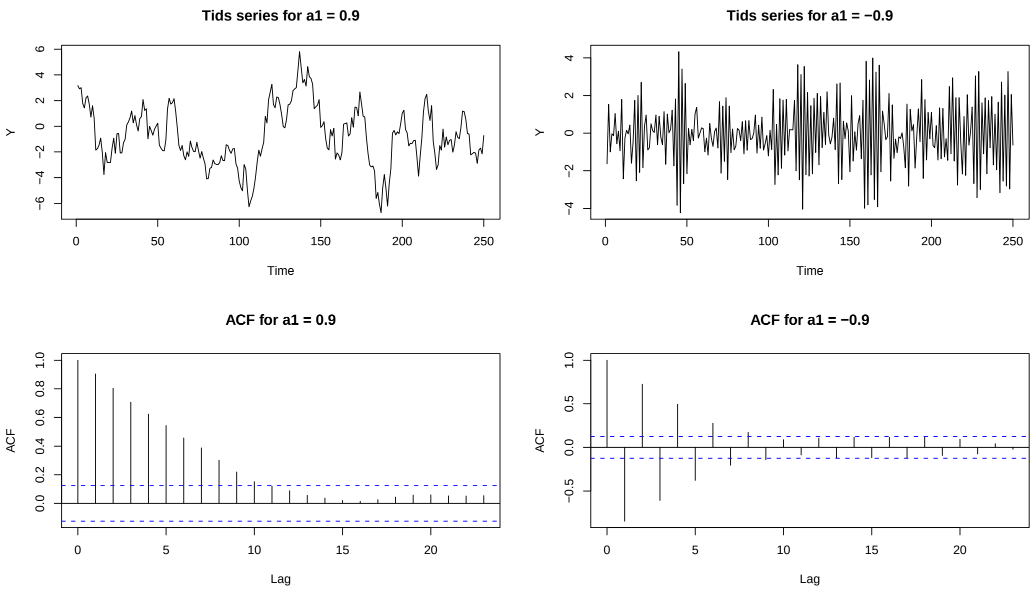

Figure 2: Picture time

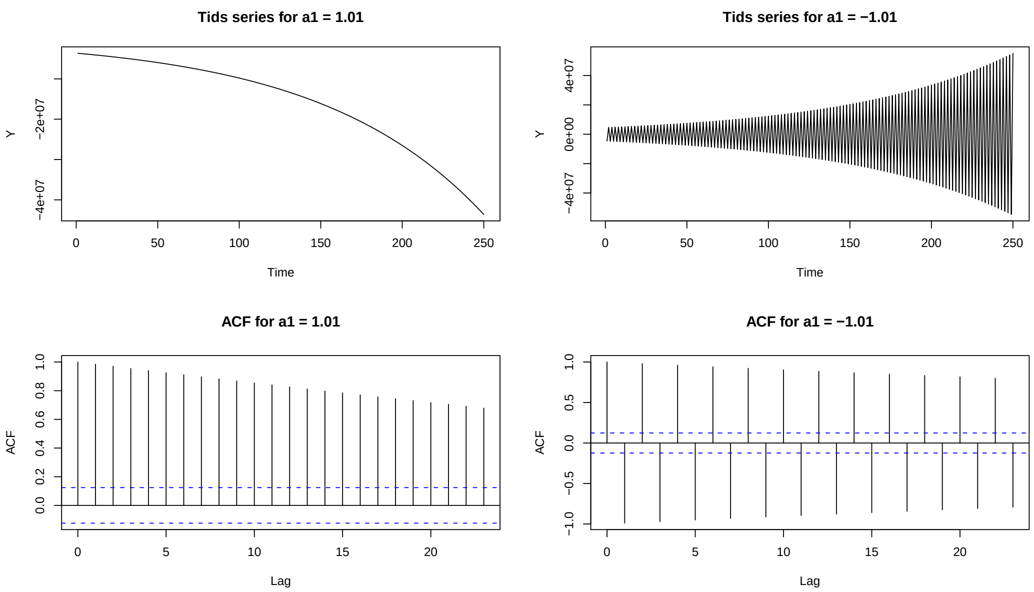

Figure 3: Picture time

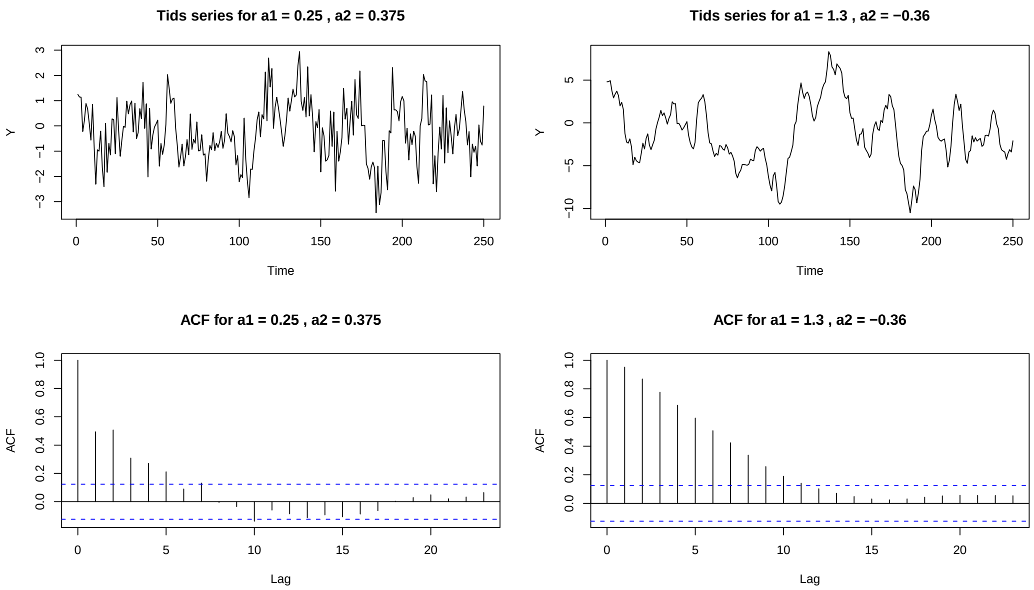

Figure 4: Picture time

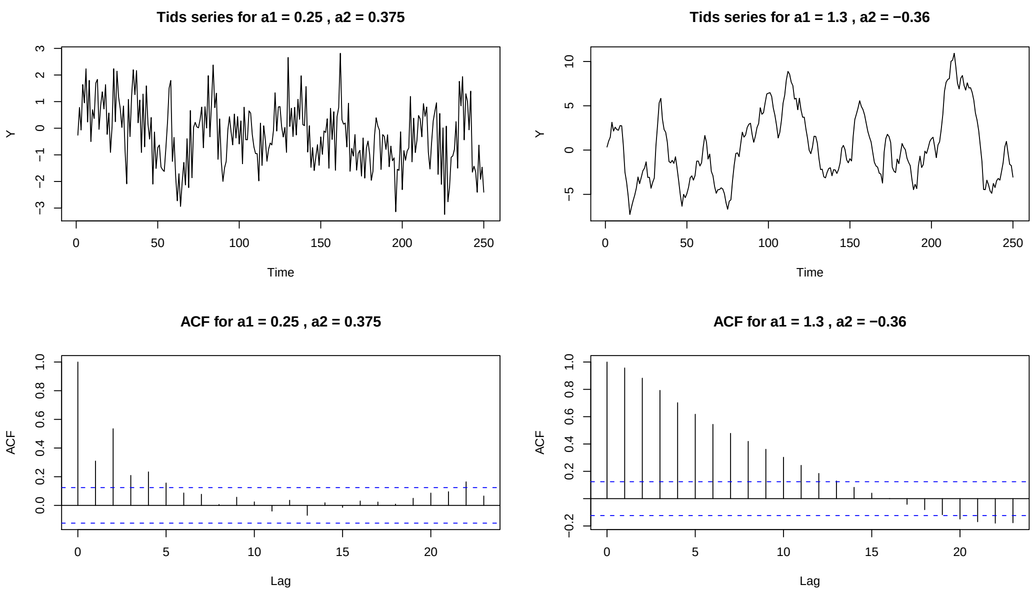

Figure 5: Picture time

Appendices

All R code used for this assignment is included here. All source code incl. latex code for this report can be found at `https://github.com/alphabits/dtu-fall-2011/tree/master/02417/assignment-2`

(part2.R)

```
numpoints = 1500
e = rnorm(numpoints)
ind.to.include = (numpoints-249):numpoints

my.filter = function (coeffs) {
    ts(filter(e, filter=coeffs, method="recursive")[ind.to.include])
}

my.arima.sim = function (coeffs) {
    arima.sim(model=list(ar=coeffs), n=250)
}

sim = function (simfns, coeffs) {
    ys = list()
    acfs = list()
    i = 1
    for (cs in coeffs) {
        ys[[i]] = simfns(cs)
        acfs[[i]] = acf(ys[[i]], plot=FALSE)
        i = i + 1
    }
    return(list(ys=ys, acfs=acfs, coeffs=coeffs))
}

###################
# AR(1) processes #
###################

coeffs = list(0.1, -0.1, 0.9, -0.9, 1.01, -1.01)
ar1.res = sim(my.filter, coeffs)


###################
# AR(2) processes #
###################

# Roots: (-1/2, 3/4) and (2/5, 9/10)
coeffs.ar2 = list(c(1/4, 3/8), c(13/10, -9/25))
ar2.res = sim(my.filter, coeffs.ar2)


########################
# AR(2) with arima.sim #
########################

# Roots: (-1/20, 1/10) and (5/6, 5/6)
coeffs.ar2.sim = list(c(1/20, 1/200), c(5/3, 25/36))
ar2.sim.res = sim(my.arima.sim, coeffs.ar2)
```

```
#################################################
# Helper functions for creating and saving plots #
#################################################

ar1.filename = "ar1-filter-%s.pdf"
ar2.filename = "ar2-filter-%s.pdf"
ar2.sim.filename = "ar2-sim-%s.pdf"

plot.timeseries = function (series, acfs, coeffs) {
    cols = length(acfs)
    par(mfcol=c(2,cols))
    for (ci in 1:cols) {
        cs = coeffs[[ci]]
        coeff.text = paste("for a1 =", cs[1])
        if (length(cs) == 2) {
            coeff.text = paste(coeff.text, ", a2 =", cs[2])
        }
        plot(series[[ci]], ylab="Y",
             main=paste("Tids series", coeff.text))
        plot(acfs[[ci]], main=paste("ACF", coeff.text))
    }
}

save.plots = function (sim.res, filename) {
    numplots = length(sim.res$ys)
    for (i in 1:ceiling(numplots/2)) {
        s = 2*i - 1
        e = if (s == numplots) s else s+1
        ind = s:e
        pdf(sprintf(filename, i), 12, 7)
        with(sim.res, {
            plot.timeseries(ys[ind], acfs[ind], coeffs[ind])
        })
        dev.off()
    }
}
```

References

[1] Henrik Madsen, *Time Series Analysis*. Chapman & Hall/CRC, 1st Edition, 2008.