Problemset 8

Nikolai German (12712506)

8.1 - ARMA processes

a)

We call a time series white noise if and only if $U_t \sim \mathcal{N}(0, \sigma^2)$ and $\gamma_U(h) = 0$ for$h \neq 0$

We know X_t is white noise.

Case 1: $Y_t = X_t^2$

Since $X_t \sim \mathcal{N}(0, \sigma^2)$ we know $\mathbb{E}(X_t) = 0$ and $Var(X_t) = \mathbb{E}(X_t^2) - \mathbb{E}(X_t)^2 = \sigma^2$. Thus $\mathbb{E}(Y_t) = \mathbb{E}(X_t^2) = \sigma^2 \neq 0$.

So Y_t is not white noise.

An other way to show this, would be using Jensen's inequality and the fact, that $f : \mathbb{R} \to \mathbb{R}_0^+$ with $x \mapsto x^2$ is a strictly convex function.

Case 2: $Z_t = |X_t|$

We want to show that $\mathbb{E}(Z_t) > 0$, thus making Z_t no white noise.

Let's have a closer look at $\mathbb{E}(Z_t)$:

$$
\begin{aligned}
\mathbb{E}(Z_t) &= \int_{-\infty}^{\infty} |x| \, f_{X_t}(x)dx \\
&= \int_{-\infty}^{0} -x f_{X_t}(x)dx + \int_{0}^{\infty} x f_{X_t}(x)dx
\end{aligned}
$$

Since $\mathbb{E}(X_t) = 0$ and X_t is symetric around $x = 0$ we know $\int_{-\infty}^{0} x f_{X_t}(x)dx + \int_{0}^{\infty} x f_{X_t}(x)dx = 0$, so $\int_{-\infty}^{0} x f_{X_t}(x)dx = -\int_{0}^{\infty} x f_{X_t}(x)dx$.

$$\begin{aligned}
\mathbb{E}(Z_t) &= -\int_{-\infty}^{0} x f_{X_t}(x)dx + \int_{0}^{\infty} x f_{X_t}(x)dx \\
&= 2 \cdot \int_{0}^{\infty} x f_{X_t}(x)dx
\end{aligned}$$

For $\sigma^2 > 0$, $\int_{0}^{\infty} x f_{X_t}(x)dx > 0$, thus $\mathbb{E}(Z_t) > 0$, and Z_t is no white noise.

b)

Let $Y_t = m + \epsilon_t + \theta\epsilon_{t-1}$ with $m \in \mathbb{R}$, $\theta \in (-1, 1)$, $\epsilon_t \sim \mathcal{N}(0, \sigma^2) \quad \text{with} \quad \sigma > 0$

Expectation

$$\begin{aligned}
\mathbb{E}(Y_t) &= \mathbb{E}(m + \epsilon_t + \theta\epsilon_{t-1}) \\
&= m + \mathbb{E}(\epsilon_t) + \theta \cdot \mathbb{E}(\epsilon_{t_1}) = m
\end{aligned}$$

Variance

$$\begin{aligned}
Var(Y_t) &= Var(m + \epsilon_t + \theta\epsilon_{t-1}) \\
&= Var(\epsilon_t) + \theta^2 \cdot Var(\epsilon_{t-1}) + 2Cov(\epsilon_t, \theta \cdot \epsilon_{t-1})
\end{aligned}$$

Since ϵ_t and ϵ_{t-1} are independent from another, $Cov(\epsilon_t, \theta \cdot \epsilon_{t-1}) = 0$.

$$\begin{aligned}
Var(Y_t) &= Var(\epsilon_t) + \theta^2 \cdot Var(\epsilon_{t-1}) \\
&= \sigma^2 + \theta^2 \cdot \sigma^2 = (1 + \theta^2) \cdot \sigma^2
\end{aligned}$$

Correlation

$$\begin{aligned}\rho_1 &= Cor(Y_t, Y_{t-1}) \\ &= \frac{Cov(Y_t, Y_{t-1})}{Var(Y_t)^{1/2} \cdot Var(Y_{t-1})^{1/2}} \\ &= \frac{1}{(1+\theta^2)\cdot\sigma^2} Cov(Y_t, Y_{t-1})\end{aligned}$$

Now we look at $Cov(Y_t, Y_{t-1})$:

$$\begin{aligned}Cov(Y_t, Y_{t-1}) &= \mathbb{E}\left[(Y_t - \mathbb{E}(Y_t))(Y_{t-1} - \mathbb{E}(Y_{t-1}))\right] \\ &= \mathbb{E}\left[(Y_t - m)(Y_{t-1} - m)\right] \\ &= \mathbb{E}(m^2) - \mathbb{E}(mY_t) - \mathbb{E}(mY_{t-1}) + \mathbb{E}(Y_t \cdot Y_{t-1}) \\ &= -m^2 + \mathbb{E}(m^2 + m\epsilon_{t-1} + m\theta\epsilon_{t-2} + m\epsilon_t + \epsilon_t\epsilon_{t-1} + \theta\epsilon_t\epsilon_{t-2} + m\theta\epsilon_{t-1} + \theta\epsilon_{t-1}^2 + \theta^2\epsilon_{t-1}\epsilon_{t-2}) \\ &= \theta\mathbb{E}(\epsilon_{t-1}^2) = \theta\sigma^2\end{aligned}$$

Thus we obtain:

$$\rho_1 = \frac{\theta\sigma^2}{(1+\theta^2)\cdot\sigma^2} = \frac{\theta}{1+\theta^2}$$

c)

Let $Y_t - m = \phi(Y_{t-1} - m) + \epsilon_t$ with $m \in \mathbb{R}$, $\phi \in (-1, 1)$, $\epsilon_t \sim \mathcal{N}(0, \sigma^2) \quad \text{with} \quad \sigma > 0$

Expectation

$$\begin{aligned}\mathbb{E}(Y_t) &= \mathbb{E}\left(m - \phi m + \phi Y_{t-1} + \epsilon_t\right) \\ &= (1-\phi)m + \phi\mathbb{E}(Y_{t-1}) \\ &= (1-\phi)m + \phi((1-\phi)m + \phi\mathbb{E}(Y_{t-2})) \\ &= (1-\phi)m \cdot \sum_{i=0}^{1} \phi^i + \phi^2\mathbb{E}(Y_{t-2}) \\ &= (1-\phi)m \cdot \sum_{i=0}^{t} \phi^i + \phi^t\mathbb{E}(Y_0)\end{aligned}$$

With $t \to \infty$ we get $\phi^t = 0$ and $\sum_{i=0}^{t} \phi^i = \frac{1}{1-\phi}$, and subsequently:

$$\mathbb{E}(Y_t) = (1-\phi)m \cdot \frac{1}{1-\phi} + 0 \cdot \mathbb{E}(Y_0) = m$$

Variance

$$\begin{aligned}
Var(Y_t) &= Var\left((1-\phi)m + \phi Y_{t-1} + \epsilon_t\right) \\
&= Var\left(\phi Y_{t-1} + \epsilon_t\right) \\
&= \phi^2 Var(Y_{t-1}) + \sigma^2 \quad \text{since } Y_{t-1} \text{ and } \epsilon_t \text{ are independent} \\
&= \sigma^2 + \phi^2 \left[\phi^2 Var(Y_{t-2}) + \sigma^2\right] \\
&= \sigma^2 \sum_{i=0}^{1} (\phi^2)^i + \phi^4 Var(Y_{t-2}) \\
&= \sigma^2 \sum_{i=0}^{t} (\phi^2)^i + \phi^{2t} Var(Y_0)
\end{aligned}$$

Again we look at $t \to \infty$ and obtain $\phi^{2t} = 0$ and $\sum_{i=0}^{t} (\phi^2)^i = \frac{1}{1-\phi^2}$. Thus we obtain for the variance:

$$Var(Y_t) = \frac{\sigma^2}{1-\phi^2}$$

Correlation

$$\begin{aligned}
\rho_j &= Cor(Y_t, Y_{t-j}) \\
&= \frac{Cov(Y_t, Y_{t-j})}{Var(Y_t)^{1/2} \cdot Var(Y_{t-j}^{1/2})} \\
&= Cov(Y_t, Y_{t-j}) \frac{1-\phi^2}{\sigma^2}
\end{aligned}$$

We look at $Cov(Y_t, Y_{t-j})$

$$\begin{aligned}
Cov(Y_t, Y_{t-j}) &= \mathbb{E}\left[(Y_t - \mathbb{E}(Y_t)) \cdot (Y_{t-j} - \mathbb{E}(Y_{t-j}))\right] \\
&= \mathbb{E}\left[(Y_t - m) \cdot (Y_{t-j} - m)\right] \\
&= \mathbb{E}\left[Y_t \cdot Y_{t-j} - m \cdot Y_t - m \cdot Y_{t-j} + m^2\right] \\
&= m^2 - m\mathbb{E}[Y_t] - m\mathbb{E}[Y_{t-j}] + \mathbb{E}\left[Y_t \cdot Y_{t-j}\right] \\
&= \mathbb{E}\left[Y_t \cdot Y_{t-j}\right] - m^2
\end{aligned}$$

Further expanding on $\mathbb{E}\left[Y_t \cdot Y_{t-j}\right]$ leaves us with:

$$
\begin{aligned}
&\mathbb{E}\left[Y_t \cdot Y_{t-j}\right] \\
&\qquad = \mathbb{E}\left[m(1-\phi) + \phi Y_{t-1} + \epsilon_t) \cdot (Y_{t-j})\right] \\
&\qquad = \mathbb{E}\left[m(1-\phi) \cdot Y_{t-j}\right] + \mathbb{E}\left[\phi \cdot Y_{t-1} \cdot Y_{t-j}\right] + \mathbb{E}\left[\epsilon_t \cdot Y_{t-j}\right] \\
&\qquad = \phi \mathbb{E}\left[Y_{t-1} \cdot Y_{t-j}\right] + m^2(1-\phi)
\end{aligned}
$$

We use this result and obtain:

$$
Cov(Y_t, Y_{t-j}) = \phi \mathbb{E}\left[Y_{t-1} \cdot Y_{t-j}\right] + m^2(1-\phi) - m^2
$$

Expanding again with $\phi \mathbb{E}\left[Y_{t-1} \cdot Y_{t-j}\right] = \phi\left[\phi \mathbb{E}\left[Y_{t-2} \cdot Y_{t-j}\right] + m^2 \cdot (1-\phi)\right]$ results in:

$$
Cov(Y_t, Y_{t-j}) = \phi^2 \mathbb{E}\left[Y_{t-2} \cdot Y_{t-j}\right] + \phi^1 m^2(1-\phi) + \phi^0 m^2(1-\phi) - m^2
$$

Using this result and expanding unti j gives us:

$$
\begin{aligned}
Cov(Y_t, Y_{t-j}) &= \phi^j \cdot \mathbb{E}[Y_{t-j}^2] + m^2 \cdot (1-\phi) \cdot \sum_{i=0}^{j-1} \phi^i - m^2 \\
&= \phi^j \cdot \left(Var[Y_{t-j}] + \mathbb{E}[Y_{t-j}]^2\right) + m^2 \cdot (1-\phi) \frac{1-\phi^j}{1-\phi} - m^2 \\
&= \phi^j \frac{\sigma^2}{1-\phi^2} + \phi^j m^2 + (1-\phi^j) m^2 - m^2 \\
&= \phi^j \frac{\sigma^2}{1-\phi^2}
\end{aligned}
$$

And finally we obtain:

$$
\rho_j = Cov(Y_t, Y_{t-j}) \frac{1-\phi^2}{\sigma^2} = \phi^j \frac{\sigma^2}{1-\phi^2} \frac{1-\phi^2}{\sigma^2} = \phi^j
$$

d)

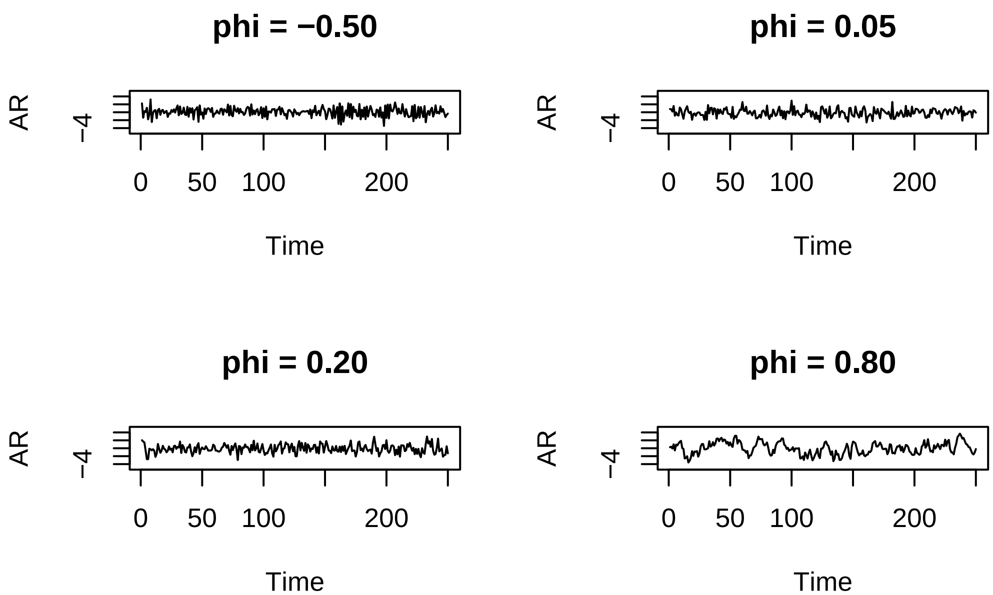
AR(1)

We simulate 250 observations of an AR(1) modell:

$$Y_t = \phi Y_{t-1} + Z_t$$

with $Z_t \stackrel{i.i.d}{\sim} \mathcal{N}(0, \sigma^2)$, and $\phi \in \{-0.05, 0.05, 0.2, 0.8\}$

```
par(mfrow = c(2,2))
for (ar in c(-.5, .05, .2, .8)) {
  AR <- arima.sim(model = list(order = c(1, 0, 0), ar = ar), n = 250)
  plot(AR, main = sprintf("phi = %.2f", ar), ylim = c(-5, 5))
}
```



We see, that for negative values of ϕ, the timeseries starts oscilating. Bigger absolute values of ϕ results in higher amplitudes. Finally the closer ϕ gets to 1, the smoother the time series gets.

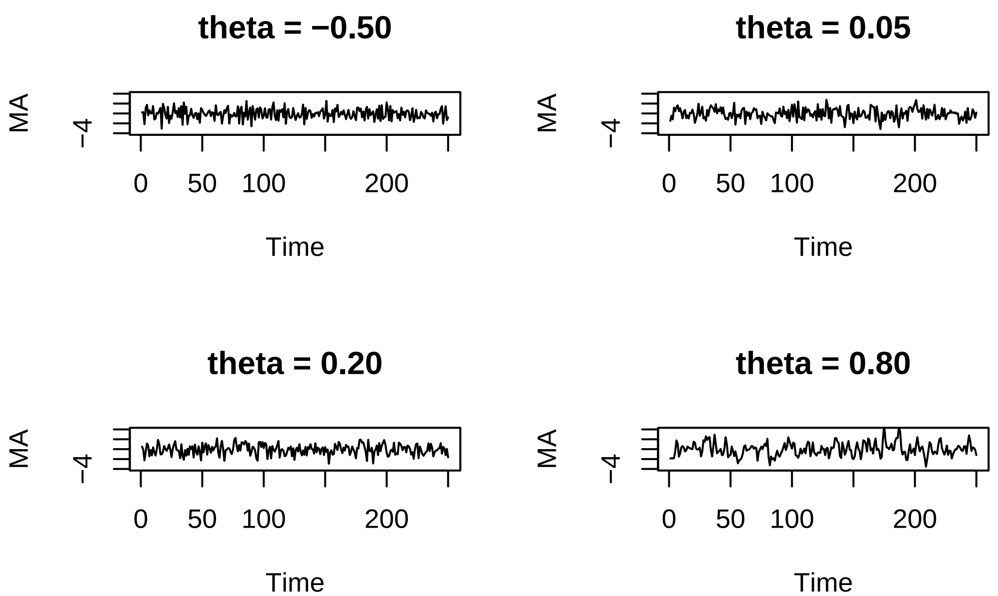
MA(1)

We simulate 250 observations of an MA(1) modell:

$$Y_t = Z_t + \theta Z_{t-1}$$

with $Z_t \overset{i.i.d}{\sim} \mathcal{N}(0, \sigma^2)$, and $\theta \in \{-0.05, 0.05, 0.2, 0.8\}$

```
par(mfrow = c(2,2))
for (ma in c(-.5, .05, .2, .8)) {
  MA <- arima.sim(model = list(order = c(0, 0, 1), ma = ma), n = 250)
  plot(MA, main = sprintf("theta = %.2f", ma), ylim = c(-4, 4))
}
```



Again we observe, that for negative values of θ, the time series seems to start oscilating. The closer the absolute value of θ, the higher and less frequent the oscilation becomes.

8.2 - Stationary Time Series

a)

X_t, Y_t are uncorrelated weakly stationary time series, thus $Cov(X_t, Y_s) = 0$ for each t, s.

Weakly stationary means $Cov(X_t, X_{t'}) = \gamma_X(t, t') = \gamma_X(|t - t'|) = \gamma_X(|t - (t + h)|) = \gamma_X(h)$, (analogue for Y) i.e. the covariance function only depends on the lag, rather than the cooncrete value of t.

We look at the covariance of $X + Y$:

$$
\begin{aligned}
\gamma_{X+Y}(t,t+h) &= Cov((X+Y)_t,(X+Y)_{t+h}) \\
&= Cov(X_t+Y_t, X_{t+h}+Y_{t+h}) \\
&= Cov(X_t, X_{t+h}+Y_{t+h}) + Cov(Y_t, X_{t+h}+Y_{t+h}) \\
&= Cov(X_{t+h}+Y_{t+h}, X_t) + Cov(X_{t+h}+Y_{t+h}, Y_t) \\
&= Cov(X_{t+h}, X_t) + Cov(Y_{t+h}, Y_t) + Cov(Y_{t+h}, X_t) + Cov(X_{t+h}, Y_t) \\
&= Cov(X_{t+h}, X_t) + Cov(Y_{t+h}, Y_t) \\
&= \gamma_X(h) + \gamma_Y(h)
\end{aligned}
$$

Thus making the covariance function of $X+Y$ only a function of the lag h.

b)

$\epsilon_t \sim \mathcal{N}(0,1)$ for all $t = 1, 2, ...$

1 $X_t = \epsilon_t - \epsilon_{t+2}$

X_t is stationary. To proof, we look at $\gamma_X(t,t+h)$, $h \in \mathbb{N}$

$$
\begin{aligned}
\gamma_X(t,t+h) &= Cov(X_t, X_{t+h}) \\
&= Cov(\epsilon_t - \epsilon_{t+2}, \epsilon_{t+h} - \epsilon_{t+h+2}) \\
&= Cov(\epsilon_t, \epsilon_{t+h} - \epsilon_{t+h+2}) - Cov(\epsilon_{t+2}, \epsilon_{t+h} - \epsilon_{t+h+2}) \\
&= Cov(\epsilon_t, \epsilon_{t+h}) - Cov(\epsilon_t, \epsilon_{t+h+2}) - Cov(\epsilon_{t+2}, \epsilon_{t+h}) + Cov(\epsilon_{t+2}, \epsilon_{t+h+2}) \\
&= \begin{cases} 0 & h \neq 2 \\ -\sigma^2 & h = 2 \end{cases}
\end{aligned}
$$

We observe, that $\gamma_X(t,t+h)$ is only a function of the lag h, thus making the time series statioary.

2 $X_t = \epsilon_t + t\epsilon_{t+2}$

Again, we will look at the covariance function $\gamma_X(t,t+h)$, $h \in \mathbb{N}$:

$$
\begin{aligned}
\gamma_X(t,t+h) &= Cov(X_t, X_{t+h}) \\
&= Cov(\epsilon_t + t\epsilon_{t+2}, \epsilon_{t+h} + t\epsilon_{t+h+2}) \\
&= Cov(\epsilon_t, \epsilon_{t+h}) + Cov(t\epsilon_{t+2}, \epsilon_{t+h}) + Cov(\epsilon_t, t\epsilon_{t+h+2}) + Cov(t\epsilon_{t+2}, t\epsilon_{t+h+2}) \\
&= Cov(\epsilon_t, \epsilon_{t+h}) + tCov(\epsilon_{t+2}, \epsilon_{t+h}) + tCov(\epsilon_t, \epsilon_{t+h+2}) + t^2Cov(\epsilon_{t+2}, \epsilon_{t+h+2}) \\
&= \begin{cases} 0 & h \neq 2 \\ t\sigma^2 & h = 2 \end{cases}
\end{aligned}
$$

Thus making the covariance function $\gamma_X(t,t+h)$ a function of both t, h, which implies it is not stationary.

3 $X_t = X_{t-1} + \epsilon_t$

$$
\begin{aligned}
\gamma_X(t,t+h) &= Cov(X_t, X_{t+h}) \\
&= Cov(X_{t-1} + \epsilon_t, X_{t+h-1} + \epsilon_{t+h}) \\
&= Cov(X_{t-1}, X_{t+h-1}) + Cov(X_{t-1}, \epsilon_{t+h}) + Cov(\epsilon_t, X_{t+h-1}) + Cov(\epsilon_t, \epsilon_{t+h}) \\
&= Cov(X_{t-1}, X_{t+h-1}) + Cov(\epsilon_t, X_{t+h-1}) \\
&= Cov(X_t, X_{t+h-1})
\end{aligned}
$$

We observe, that $\gamma_X(t,t+h) = \gamma_X(t,t+h-1) = ... = \gamma_X(t,t+h-(h-1)) = \gamma_X(t,t) = Var(X_t)$

$$
\begin{aligned}
\gamma_X(t,t) &= Var(X_t) \\
&= Var(X_{t-1} + \epsilon_t) \\
&= Var(X_{t-1}) + \sigma^2 \\
&= Var(X_0) + t\sigma^2 \\
&= Var(X_0) + t
\end{aligned}
$$

Thus making $\gamma_X(t,t+h)$ a function of t and not stationary.

4 $X_t = \phi X_{t-1} + 1 + \epsilon_t$ with $|\phi| < 1$

$$
\begin{aligned}
\gamma_X(t, t+h) &= Cov(X_t, X_{t+h}) \\
&= Cov(\phi X_{t-1} + 1 + \epsilon_t, \phi X_{t+h-1} + 1 + \epsilon_{t+h}) \\
&= Cov(\phi X_{t-1}, \phi X_{t+h-1} + 1 + \epsilon_{t+h}) \\
&+ Cov(1, \phi X_{t+h-1} + 1 + \epsilon_{t+h}) \\
&+ Cov(\epsilon_t, \phi X_{t+h-1} + 1 + \epsilon_{t+h}) \\
&= Cov(\phi X_{t-1}, \phi X_{t+h-1}) + Cov(\phi X_{t-1}, 1) + Cov(\phi X_{t-1}, \epsilon_{t+h}) \\
&+ Cov(1, \phi X_{t+h-1}) + Cov(1, 1) + Cov(1, \epsilon_{t+h}) \\
&+ Cov(\epsilon_t, \phi X_{t+h-1}) + Cov(\epsilon_t, 1) + Cov(\epsilon_t, \epsilon_{t+h}) \\
&= Cov(\phi X_{t-1}, \phi X_{t+h-1}) + Cov(\epsilon_t, \phi X_{t+h-1}) \\
&= Cov(\phi X_{t-1} + \epsilon_t, \phi X_{t+h-1}) \\
&= \phi Cov(X_t, X_{t+h-1}) \\
&= \phi^h Var(X_t)
\end{aligned}
$$

We follow up by looking closely at $Var(X_t)$:

$$
\begin{aligned}
Var(X_t) &= Var(\phi X_{t-1} + 1 + \epsilon_t) \\
&= \phi^2 Var(X_{t-1}) + 1 \\
&= (\phi^2)^t Var(X_0) + \sum_{i=0}^{t} (\phi^2)^i
\end{aligned}
$$

For $t \to \infty$ we get $(\phi^2)^t = 0$ and $\sum_{i=0}^{t} (\phi^2)^i = \frac{1}{1-\phi^2}$

Thus giving us for the variance $Var(X_t) = \frac{1}{1-\phi^2}$

We use this result in the covariance function and obtain

$$
\gamma_X(t, t+h) = \phi^h Var(X_t) = \frac{\phi^h}{1-\phi^2}
$$

We observe once again, that the covariance function only depends on h and not on t, thus making X_t stationary.

c)

Since the definitness of potentially infinite sized matrices can be hard to tackle analytically, even for band-matrices, we will instead computationally check if all eigenvalues of the band matrices Γ_i, $i = 1, 2$ are positive for sizes $n = 4, \ldots, 100$:

```
# function for checking if all eigenvalues of a band matrix are positive:
band.pos.def <- function(corr, n) {
  diags <- lapply(corr, function(gamma) rep(gamma, n))
  mat <- Matrix::bandSparse(n = n,
                            k = -seq_len(length(corr)),
                            diag = diags,
                            symm = TRUE)
  all(eigen(mat)$values > 0)
}

# check if all eigenvalues are positive for first set and n = 4,...,100:
all(
  vapply(
    4:100,
    function(n) band.pos.def(c(1, .8, .2), n),
    logical(1)
    )
  )
```

```
[1] FALSE
```

```
# check if all eigenvalues are positive for second set and n = 4,...,100:
all(
  vapply(
    4:100,
    function(n) band.pos.def(c(1, .8, .5), n),
    logical(1)
    )
  )
```

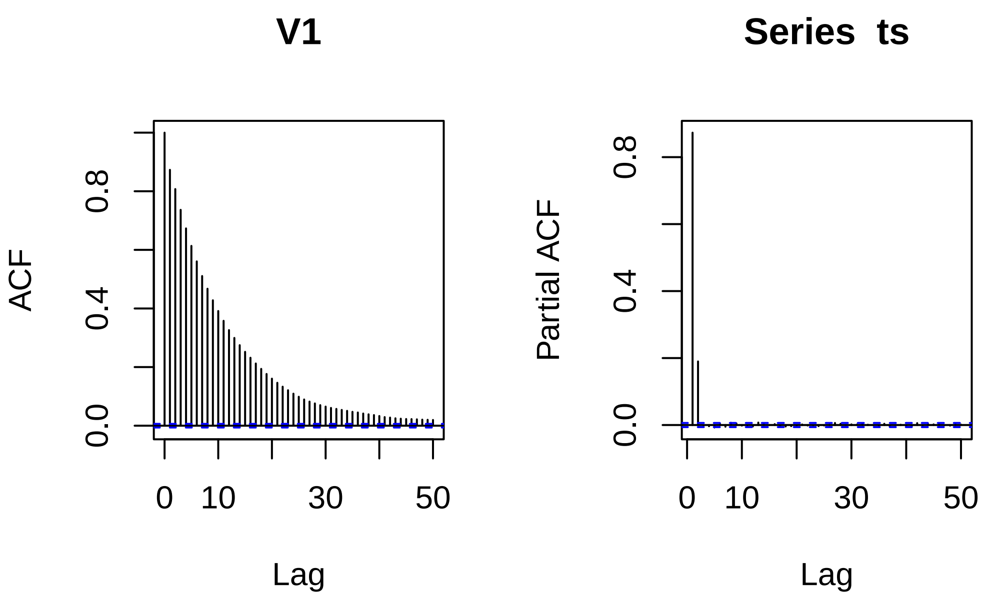
```
[1] FALSE
```

8.3 - Fitting Models to Time Series Data

```
ts <- read.delim("../data/Ex_8_3.txt", header = FALSE)
```

a)

```
par(mfrow = c(1,2))
acf(ts)
pacf(ts)
```



Since the acf decays of in geometric fashion, and the pacf cuts off after a lag of 2, we assume to see an AR(2) process.

The AR(2) process is specified by:

$$Y_t = \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + Z_t$$

with $Z_t \overset{i.i.d.}{\sim} \mathcal{N}(0, \sigma^2)$.

b)

We fit the linear model and compute the 95% confidence intervals:

```r
df <- data.frame("t0" = ts[3:100000, 1], "t1" = ts[2:99999, 1], "t2" = ts[1: 99998, 1])

model <- lm(t0 ~ ., df)
coefs <- coef(summary(model))[,1:2]

(CI = cbind(lower =  coefs[,1]  - 1.96*coefs[,2],
            estimate = coefs[,1],
            upper = coefs[,1]  + 1.96*coefs[,2]))
```

```
                lower  estimate     upper
(Intercept) 0.9976130 1.0289191 1.0602252
t1          0.7012150 0.7073004 0.7133858
t2          0.1838265 0.1899118 0.1959972
```

c)

An ARMA(p,q) process combines an AR(p) with a MA(q) process. In detail:

$$Y_t = \underbrace{\phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \cdots + \phi_p y_{t-p}}_{\text{AR(p) part}} + \underbrace{Z_t + \theta_1 Z_{t-1} + \cdots + \theta_q Z_{t-q}}_{\text{MA(q) part}}$$

We fet the ARIMA(2,1) model and compute the 95% CI:

```r
mod2 <- arima(ts, order = c(2, 0, 1))

est <- mod2$coef
se <- sqrt(diag(mod2$var.coef))

(CI2 <- cbind(lower = est - 1.96 * se,
              estimate = est,
              upper = est + 1.96 * se))
```

```
                lower    estimate       upper
ar1        0.66888744  0.70033202  0.73177660
ar2        0.16837714  0.19598549  0.22359383
ma1       -0.02483867  0.00721764  0.03927395
intercept  9.94982227 10.01001354 10.07020481
```

d)

The linear model results in similiar point estimates, compared to the ARMA(2,1) model. The obvious difference is, the linear model underestimates the variance in the parameters. A reason for that can be, that the linear model does interpret Y_{t-1} and Y_{t-2} as independent. In other words, the fact of present autocorrelation violates the assumption of independent covariates.