STAT211 Homework 1
Solutions

Teaching Assistant: Sondre Hølleland
10:15-12:00, 23.01.2019
Auditorium 5, Realfagsbygget

1.4 (Brockwell et al., 2016, p. 35)
Let $\{Z_t\}$ be a sequence of independent normal random variables, each with mean 0 and variance σ^2, and let a, b, and c be constants. Which, if any, of the following processes are stationary? For each stationary process specify the mean and autocovariance function.

a) $X_t = a + bZ_t + cZ_{t-2}$

Solution: $\mu_X(t) = \mathbb{E}\, X_t = a + b\,\mathbb{E}\, Z_t + c\,\mathbb{E}\, Z_{t-2} = a$ and

$$\begin{aligned}
\gamma_X(t+h,t) &= \text{Cov}\,(X_{t+h}, X_t) = \text{Cov}\,(a + bZ_{t+h} + cZ_{t+h-2}, a + bZ_t + cZ_{t-2}) \\
&= \text{Cov}\,(bZ_{t+h} + cZ_{t+h-2}, bZ_t + cZ_{t-2}) \\
&= b^2\,\text{Cov}\,(Z_{t+h}, Z_t) + bc\,\text{Cov}\,(Z_{t+h-2}, Z_t) + bc\,\text{Cov}\,(Z_{t+h}, Z_{t-2}) + c^2\,\text{Cov}\,(Z_{t+h-2}, Z_{t-2}) \\
&= (b^2\delta_{0,h} + bc\delta_{h,2} + bc\delta_{h,-2} + c^2\delta_{h,0})\sigma^2 \\
&= \begin{cases} (b^2+c^2)\sigma^2, & \text{if } h = 0 \\ b\,c\,\sigma^2, & \text{if } |h| = 2 \\ 0, & \text{ohterwise.} \end{cases}
\end{aligned}$$

Here we have used the Kronecker delta: $\delta_{i,j} = 1$ if $i = j$ and 0 otherwise. Since $\mu_X(t)$ and $\gamma_X(t+h,t)$ do not depend on t, the process $\{X_t\}$ is stationary.

b) $X_t = a + bZ_1\cos(ct) + cZ_2\sin(ct)$

Solution: $\mu_X(t) = \mathbb{E}\, X_t = a + b\cos(ct)\,\mathbb{E}\, Z_1 + c\sin(ct)\,\mathbb{E}\, Z_2 = a$ and

$$\begin{aligned}
\gamma_X(t+h,t) &= \text{Cov}\,(X_{t+h}, X_t) = \text{Cov}\,(a + bZ_1\cos(ct+ch) + cZ_2\sin(ct+ch), a + bZ_1\cos(ct) + cZ_2\sin(ct)) \\
&= b^2\,\sigma^2\,\cos(ct)\cos(ct+ch) + c^2\,\sigma^2\,\sin(ct)\sin(ct+ch),
\end{aligned}$$

Since $\gamma(t+h,t)$ depends on t the process $\{X_t\}$ is not stationary.

c) $X_t = a + bZ_t\cos(ct) + cZ_{t-1}\sin(ct)$

Solution: $\mu_X(t) = \mathbb{E}\, X_t = a + b\cos(ct)\,\mathbb{E}\, Z_1 + c\sin(ct)\,\mathbb{E}\, Z_2 = a$ and

$$\begin{aligned}
\gamma_X(t+h,t) &= \text{Cov}\,(X_{t+h}, X_t) = \text{Cov}\,(bZ_{t+h}\cos(ct+ch) + cZ_{t+h-1}\sin(ct+ch), bZ_t\cos(ct) + cZ_{t-1}\sin(ct)) \\
&= \{b^2\cos(ct)\cos(c(t+h))\delta_{h,0} + bc\sin(ct)\cos(c(t+h))\delta_{h,-1} \\
&\quad + bc\cos(ct)\sin(c(t+h))\delta_{h,1} + c^2\sin(ct)\sin(c(t+h))\delta_{h,0}\}\sigma^2 \\
&= \begin{cases} \{b^2\cos(ct)\cos(ct+ch) + c^2\sin(ct)\sin(ct+ch)\}\sigma^2, & \text{if } h = 0, \\ bc\sin(ct)\cos(ct+ch)\sigma^2, & \text{if } h = -1, \\ bc\cos(ct)\sin(ct+ch)\sigma^2, & \text{if } h = 1, \\ 0, & \text{otherwise.} \end{cases}
\end{aligned}$$

Since $\gamma_X(t+h,t)$ depends on t the process $\{X_t\}$ is not stationary.

d) $X_t = a + bZ_0$

Solution: $\mu_X(t) = \mathbb{E}\,X_t = a + b\,\mathbb{E}\,Z_0 = a$ and

$$\gamma_X(t+h,t) = \mathrm{Cov}\,(X_{t+h}, X_t) = \mathrm{Cov}\,(a+bZ_0, a+bZ_0) = b^2\mathrm{Var}(Z_0) = b^2\sigma^2$$

Since $\mu_X(t)$ and $\gamma_X(t+h,t)$ do not depend on t, the process $\{X_t\}$ is stationary.

e) $X_t = Z_0\cos(ct)$

Solution: $\mu_X(t) = \mathbb{E}\,X_t = \cos(ct)\,\mathbb{E}\,Z_0 = 0$ and

$$\begin{aligned}\gamma_X(t+h,t) &= \mathrm{Cov}\,(X_{t+h}, X_t) = \mathrm{Cov}\,(Z_0\cos(c(t+h)), Z_0\cos(ct))\\ &= \cos(c(t+h))\cos(ct)\mathrm{Var}(Z_0) = \cos(c(t+h))\cos(ct)\sigma^2.\end{aligned}$$

Since $\gamma_X(t+h,t)$ depends on t the process $\{X_t\}$ is not stationary.

f) $X_t = Z_tZ_{t-1}$

Solution: $\mu_X(t) = \mathbb{E}\,X_t = \mathbb{E}\,(Z_tZ_{t-1}) = \mathbb{E}\,(Z_t)\,\mathbb{E}\,(Z_{t-1}) = 0$ and

$$\begin{aligned}\gamma_X(t+h,t) &= \mathrm{Cov}\,(X_{t+h}, X_t) = \mathrm{Cov}\,(Z_{t+h}Z_{t+h-1}, Z_tZ_{t-1}) = \mathbb{E}\,(Z_{t+h}Z_{t+h-1}Z_tZ_{t-1})\\ &= \mathbb{E}\,(Z_t^2Z_{t-1}^2)\delta_{h,0} = \mathbb{E}\,(Z_t^2)\mathbb{E}\,(Z_{t-1}^2)\delta_{h,0} = \sigma^2\sigma^2\delta_{h,0}\\ &= \begin{cases}\sigma^4, & \text{if } h=0\\ 0, & \text{otherwise.}\end{cases}\end{aligned}$$

Since $\mu_X(t)$ and $\gamma_X(t+h,t)$ do not depend on t, the process $\{X_t\}$ is stationary.

1.5 (Brockwell et al., 2016, p. 35)
Let $\{X_t\}$ be the moving-average process of order 2 given by $X_t = Z_t + \theta Z_{t-2}$, where $\{Z_t\}$ is WN(0,1).

a) Find the autocovariance and autocorrelation functions for this process when $\theta = 0.8$.

Solution: We found the autocovariance function in 1.4 a) inserting $a=0$, $b=1$ and $c=\theta$, we get that $\gamma_X(h) = (1+\theta^2)\delta_{h,0}\sigma^2 + \theta\sigma^2\delta_{|h|,2} = 1.64\cdot\delta_{h,0} + 0.8\cdot\delta_{|h|,2}$ For the autocorrelation function, we get

$$\rho_X(h) = \frac{\gamma_X(h)}{\gamma_X(0)} = \frac{(1+\theta^2)\delta_{h,0}\sigma^2 + \theta\sigma^2\delta_{|h|,2}}{(1+\theta^2)\sigma^2} = \begin{cases}1, & h=0\\ \frac{\theta}{1+\theta^2} = 0.488, & |h|=2\end{cases}$$

b) Compute the variance of the sample mean $(X_1+X_2+X_3+X_4)/4$ when $\theta = 0.8$.

Solution: Let $\overline{X}_4 = (X_1+X_2+X_3+X_4)/4$. Then

$$\begin{aligned}4\overline{X}_4 &= Z_1 + \theta Z_{-1} + Z_2 + \theta Z_0 + Z_3 + \theta Z_1 + Z_4 + \theta Z_2\\ &= Z_1+Z_2+Z_3+Z_4 + \theta(Z_{-1}+Z_0+Z_1+Z_2)\\ &= (1+\theta)(Z_1+Z_2) + Z_3 + Z_4 + \theta(Z_{-1}+Z_0)\\ \mathrm{Var}\overline{X}_4 &= \frac{1}{4^2}\mathrm{Var}((1+\theta)(Z_1+Z_2) + Z_3 + Z_4 + \theta(Z_{-1}+Z_0))\\ &= \frac{1}{16}\mathbb{E}\,((1+\theta)(Z_1+Z_2) + Z_3 + Z_4 + \theta(Z_{-1}+Z_0))^2\\ &= \frac{1}{16}\{(1+\theta)^2(\mathbb{E}\,Z_1^2 + \mathbb{E}\,Z_2^2) + \mathbb{E}\,Z_3^2 + \mathbb{E}\,Z_4^2 + \theta^2(\mathbb{E}\,Z_{-1}^2 + \mathbb{E}\,Z_0^2)\}\\ &= 2\frac{(1+\theta)^2 + 1 + \theta^2}{16} = \frac{1+\theta+\theta^2}{4}.\end{aligned}$$

With $\theta = 0.8$, the variance is 0.61.
You may control the answers by simulation:

```
xt <- arima.sim(list(order=c(0,0,2), ma=c(0,.8)), n=1e8) # you may want to choose a lower n
var(filter(xt, rep(1/4,4),sides=1)[-(1:4)])

## [1] 0.6100163

xt <- arima.sim(list(order=c(0,0,2), ma=c(0,-.8)), n=1e8) # you may want to choose a lower n
var(filter(xt, rep(1/4,4),sides=1)[-(1:4)])

## [1] 0.2100262
```

c) Repeat (b) when $\theta = 0.8$ and compare your answer with the result obtained in (b).

Solution: With $\theta = -0.8$ the variance is 0.21, i.e. lower than when $\theta = 0.8$.

1.6 (Brockwell et al., 2016, p. 35)
Let $\{X_t\}$ be the AR(1) process defined in Example 1.4.5.

a) Compute the variance of the sample mean $(X_1 + X_2 + X_3 + X_4)/4$ when $\phi = 0.9$ and $\sigma^2 = 1$.

Solution: $X_t = \phi X_{t-1} + Z_t$, with $Z_t \sim \text{WN}(0, 1)$. Let $\overline{X}_4 = (X_1 + X_2 + X_3 + X_4)/4$.

$$\begin{aligned}
X_4 &= \phi^3 X_1 + \phi^2 Z_2 + \phi Z_3 + Z_4\\
X_3 &= \phi^2 X_1 + \phi Z_2 + Z_3\\
X_2 &= \phi X_1 + Z_2\\
4\overline{X}_4 &= (1+\phi+\phi^2+\phi^3)X_1 + (1+\phi+\phi^2)Z_2 + (1+\phi)Z_3 + Z_4\\
\text{Var}(\overline{X}_4) &= \frac{1}{4^2}\text{Var}((1+\phi+\phi^2+\phi^3)X_1 + (1+\phi+\phi^2)Z_2 + (1+\phi)Z_3 + Z_4)\\
&= \frac{1}{16}\big\{(1+\phi+\phi^2+\phi^3)^2\text{Var}\, X_1 + (1+\phi+\phi^2)^2\text{Var}\, Z_2 + (1+\phi)^2\text{Var}\, Z_3 + \text{Var}\, Z_4\big\}\\
&= \frac{1}{16}\big\{(1+\phi+\phi^2+\phi^3)^2\gamma_X(0) + (1+\phi+\phi^2)^2 + (1+\phi)^2 + 1\big\}
\end{aligned}$$

We have from example 1.4.5 that $\gamma_X(0) = \sigma^2/(1-\phi^2) = (1-\phi^2)^{-1}$. Hence,

$$\text{Var}(\overline{X}_4) = \frac{1}{16}\big\{(1+\phi+\phi^2+\phi^3)^2(1-\phi^2)^{-1} + (1+\phi+\phi^2)^2 + (1+\phi)^2 + 1\big\} = 4.638.$$

b) Repeat (a) when $\phi = -0.9$ and compare your answer with the result obtained in (a).

Solution: $\text{Var}(\overline{X}_4) = 0.126$.
You may control the answers by simulation:

```
xt <- arima.sim(list(order=c(1,0,0), ar=.9), n=1e8) # you may want to choose a lower n
var(filter(xt, rep(1/4,4),sides=1)[-(1:4)])

## [1] 4.640985

xt <- arima.sim(list(order=c(1,0,0), ar=-.9), n=1e8) # you may want to choose a lower n
var(filter(xt, rep(1/4,4),sides=1)[-(1:4)])

## [1] 0.1256778
```

1.7 (Brockwell et al., 2016, p. 35)
If $\{X_t\}$ and $\{Y_t\}$ are uncorrelated stationary sequences, i.e., if X_r and Y_s are uncorrelated for every r and

s, show that $\{X_t + Y_t\}$ is stationary with autocovariance function equal to the sum of the autocovariance functions of $\{X_t\}$ and $\{Y_t\}$.

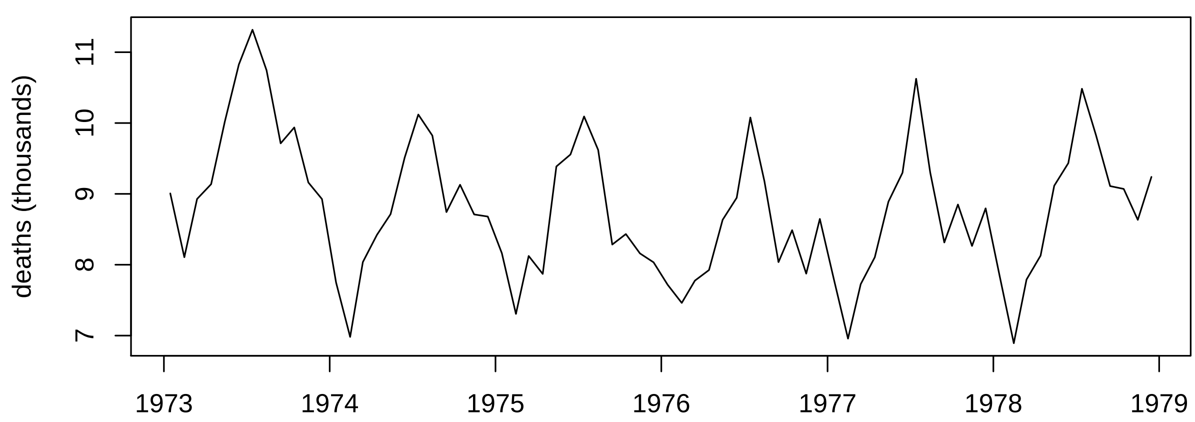
Solution: Let $U_t = X_t + Y_t$. Then $\mu_U(t) = \mathbb{E}(X_t + Y_t) = \mu_X(t) + \mu_Y(t)$, which we know is independent of t since $\{X_t\}$ and $\{Y_t\}$ are stationary. Further,

$$\begin{aligned}\gamma_U(t+h,t) &= \mathrm{Cov}\left(X_{t+h} + Y_{t+h}, X_t + Y_t\right)\\ &= \mathrm{Cov}\left(X_{t+h}, X_t\right) + \underbrace{\mathrm{Cov}\left(X_{t+h}, Y_t\right)}_{=0} + \underbrace{\mathrm{Cov}\left(Y_{t+h}, X_t\right)}_{=0} + \mathrm{Cov}\left(Y_{t+h}, Y_t\right)\\ &= \gamma_X(t+h,t) + \gamma_Y(t+h,t) = \gamma_X(h) + \gamma_Y(h),\end{aligned}$$

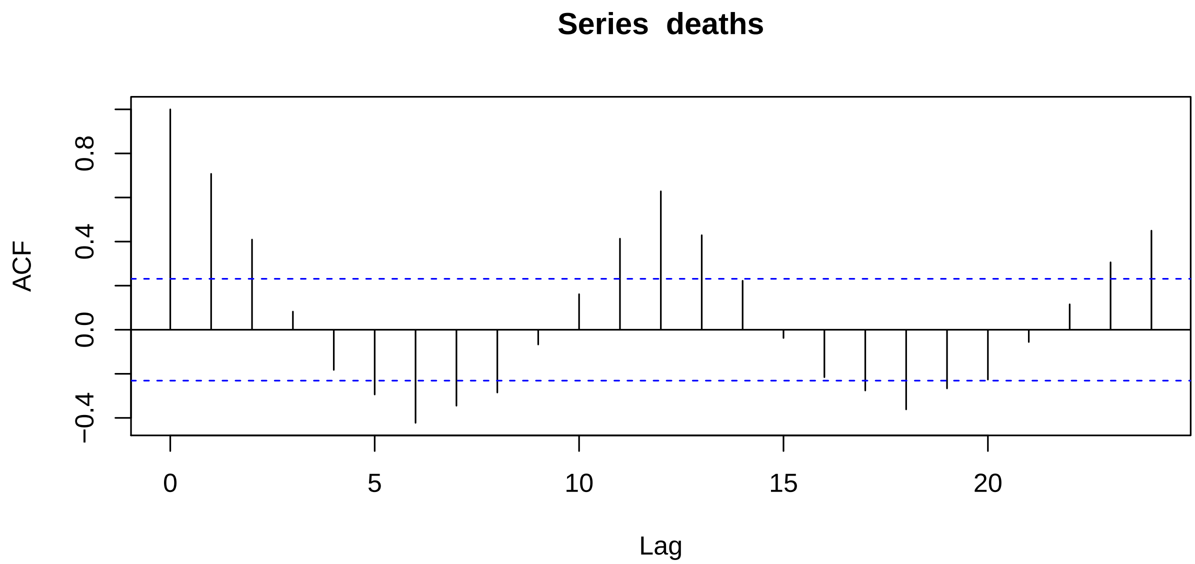
where the last equality holds since $\{X_t\}$ and $\{Y_t\}$ are stationary. Hence, $\{X_t + Y_t\}$ is stationary.

1.17 (Brockwell et al., 2016, p. 36)
Load the dataset `deaths` in R using the `read.table` function. Plot the data. Also create a histogram of the data using the R function `hist`. Plot the sample autocorrelation function using the `acf` function. The presence of a strong seasonal component with period 12 is evident in the graph of the data and in the sample autocorrelation function.

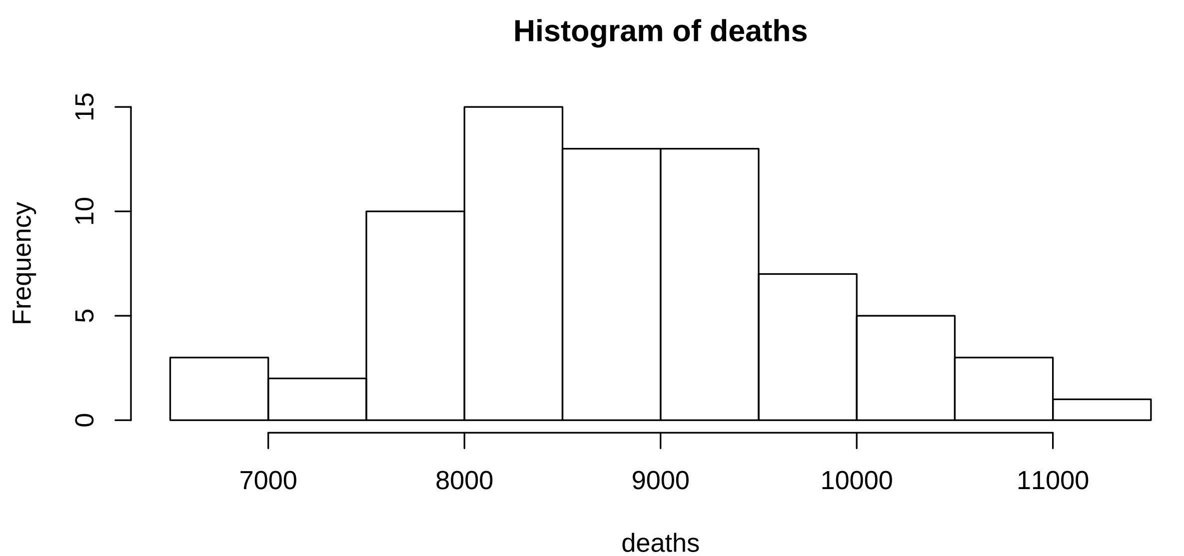
```
data<-read.table("../../Data/deaths.txt", skip = 9, header=FALSE)
names(data)<-c("month","year","deaths")
data$time <- as.Date(paste(data$year,data$month,"15",sep="-"))
attach(data)
plot(time,deaths/1000, type="l", xlab="", ylab = "deaths (thousands)")
```




```
acf(deaths ,lag.max = 24)
```




```r
hist(deaths, breaks = 15)
```




```r
detach(data)
```

1.18 (Brockwell et al., 2016, p. 37)
We are still studying the dataset `deaths`. In this exercise, you are supposed to reproduce the figures 1-24 and 1-25 in (Brockwell et al., 2016, pp. 27-28). In 1.17, we found a period of length 12. Fit a seasonal component using the procedure described in section 1.5.2.1 on page 26. You may use the following functions or write your own:

```
# Function for calculating a moving average when d is even
ma <- function(x,n=12){filter(x,c(.5,rep(1,n-1),.5)/n, sides=2)}
# Function for finding the seasonal component
seasonal.component <- function(x){
  # First step: detrending
  detrended <- x - ma(x)
  # Second step: Calculating sesonal component from detrended data
  wt<-rowMeans(matrix(detrended[!is.na(detrended)], nrow=12,byrow=FALSE))
  st<-(wt-mean(wt))[c(7:12,1:6)] #seasonal component
  return(st)
}
```

Plot the deseasonalized data (as in figure 1-24). Fit a quadratic trend (polynomial of order two) to the deseasonalized data and add the curve to the plot you just created. The trend should be $\widehat{m}_t = 9952 - 71.82t + 0.8260t^2$ for $1 \leq t \leq 72$. This can be done using the following code:

```
# Let dtr be the deseasonalized observations
M <- poly(1:72, degree=2, raw=TRUE)
trend<-lm(dtr ~ M) # Re-estimating trend of the deseasonalized data
```

Calculate the detrended and deseasonalized data, i.e.

$$\widehat{Y}_t = x_t - \widehat{m}_t - \widehat{s}_t, \quad t = 1, \dots, 72.$$

Plot the sample autocorrelation function of $\{\widehat{Y}_t\}$.
Forecast the data for the next 24 months without allowing for this dependence, based on the assumption that the estimated seasonal and trend components are true values and that $\{Y_t\}$ is a white noise sequence with zero mean. Calculate $\widehat{m}_{72+k}$ for $k = 1, \dots, 24$ and do the forecasting by

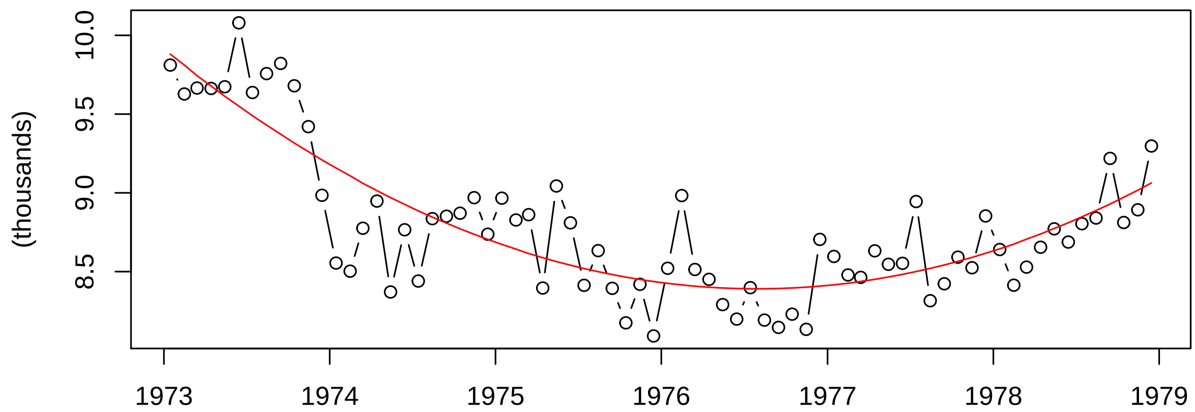
$$\widehat{X}_{72+k} = \widehat{m}_{72+k} + \widehat{s}_{72+k}, \quad k = 1, \dots 24.$$

Plot the original data with the forecasts appended. Later we shall see how to improve on these forecasts by taking into account the dependence in the series $\{Y_t\}$.
Tips: To calculate $\widehat{m}_{72+k}$ the following code may be useful:

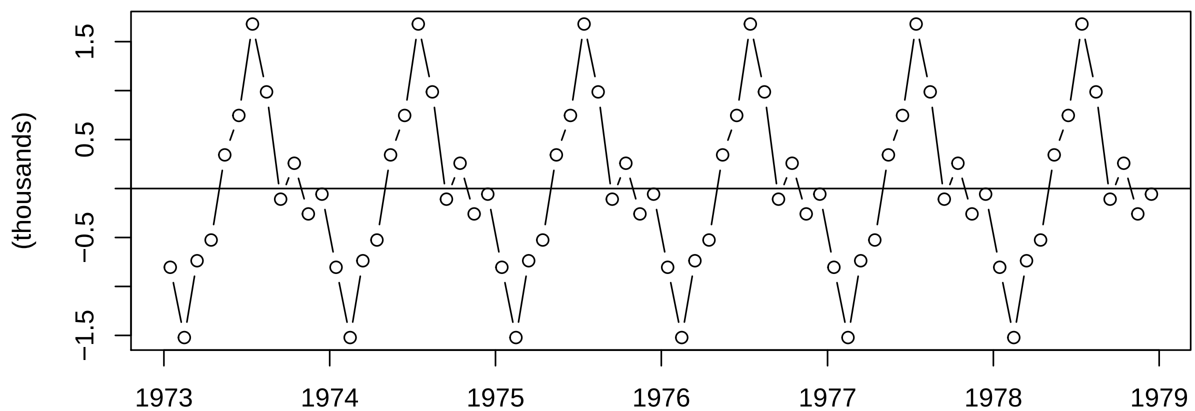
```
M <- poly(72 + 1:24, 2, raw=TRUE)
m.hat <- predict(trend, newdata= M)
```

Solution:
This problem can be solved using the R-package itsmr, which is an R-version of the ITSM software used in Brockwell et al. (2016):

```
library(itsmr)
dtr<-deaths-season(deaths,12) #removing season
plot(data$time,dtr/1000,type="b",   #plotting deseasonalized data
     xlab="",ylab="(thousands)")
lines(data$time,trend(dtr,2)/1000,col=2)   # adding trend line
```



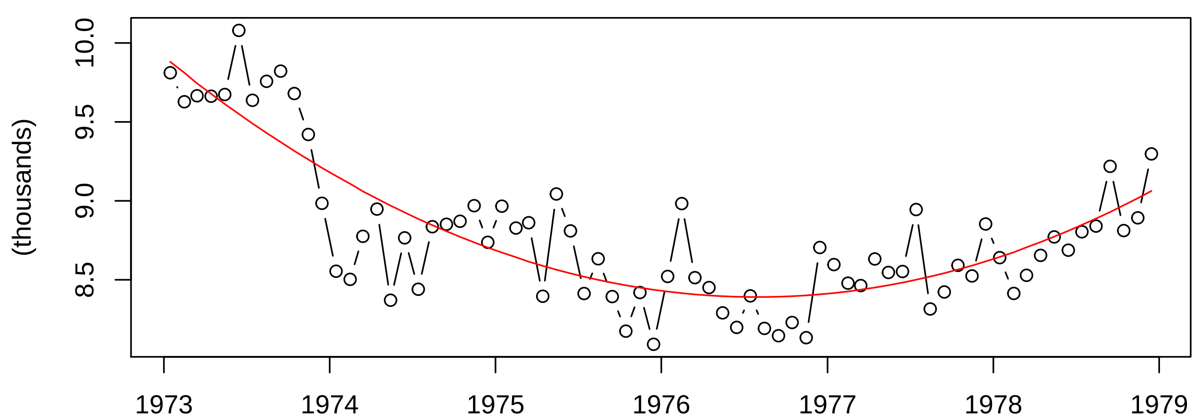

```
plot(data$time,season(deaths,12)/1000, type="b",  # Plotting seasonal component
     xlab="",ylab="(thousands)")
abline(h=0) # adding horizontal line at zero
```



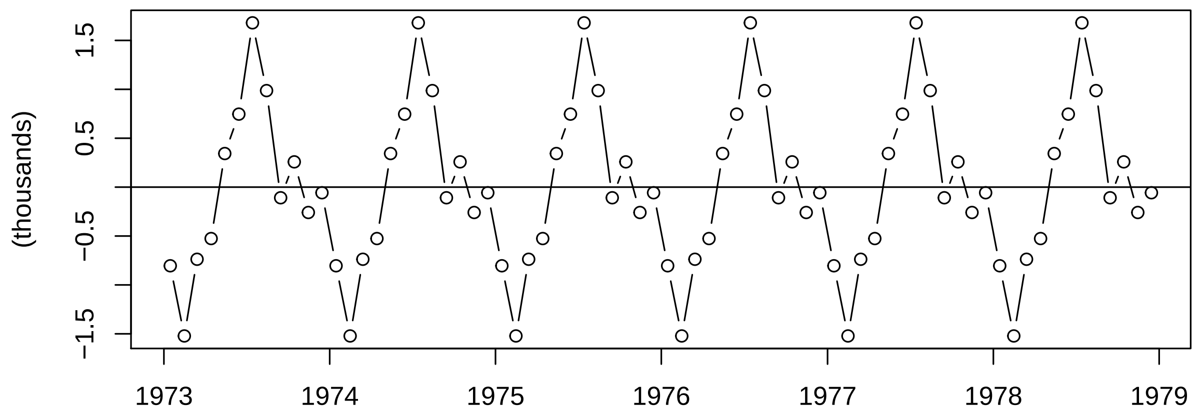

```
detach("package:itsmr", unload=TRUE)
```

Or you can write your own code:

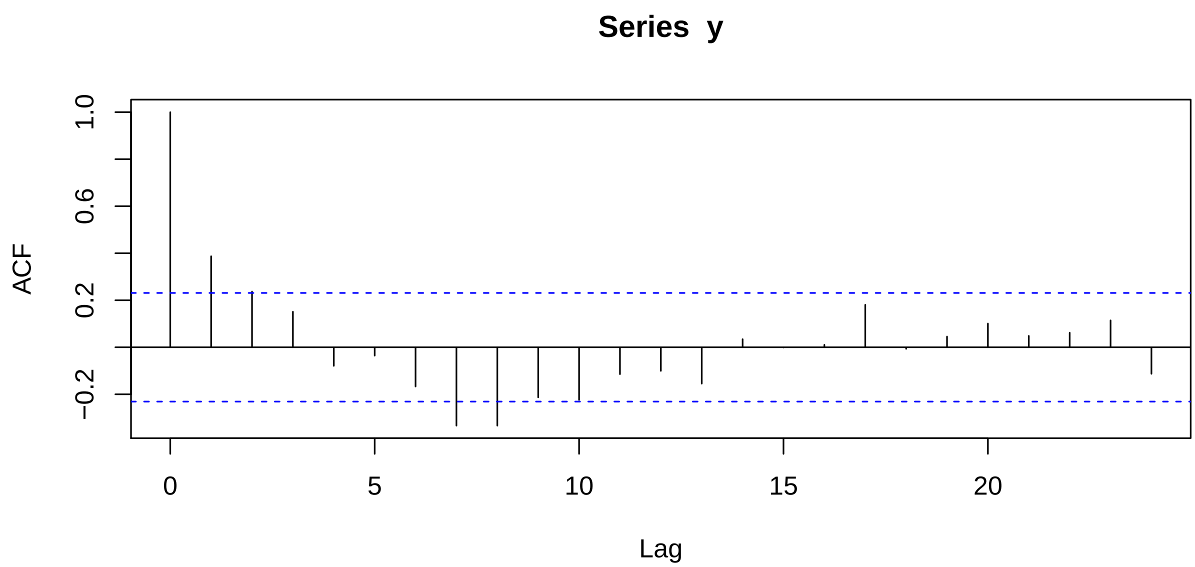
```
attach(data)

st <- seasonal.component(x = deaths)
dtr<-deaths-st # removing seasonal component from data

# Reproducing Fig 1.24
plot(time,dtr/1000, type="b", xlab="",ylab="(thousands)") # plotting de-seasonalized data
M <- poly(1:72,degree=2, raw=TRUE)
trend<-lm(dtr ~ M) # Re-estimating trend
trend$coefficients

##   (Intercept)            M1            M2
## 9951.8220098  -71.8171689    0.8260222

lines(time,trend$fitted/1000,"l",col=2) # adding trend to plot
```




```
# Reproducing Fig 1.25
plot(time, rep(st/1000,length.out=72),type="b",xlab="",ylab="(thousands)")
abline(h=0)
```

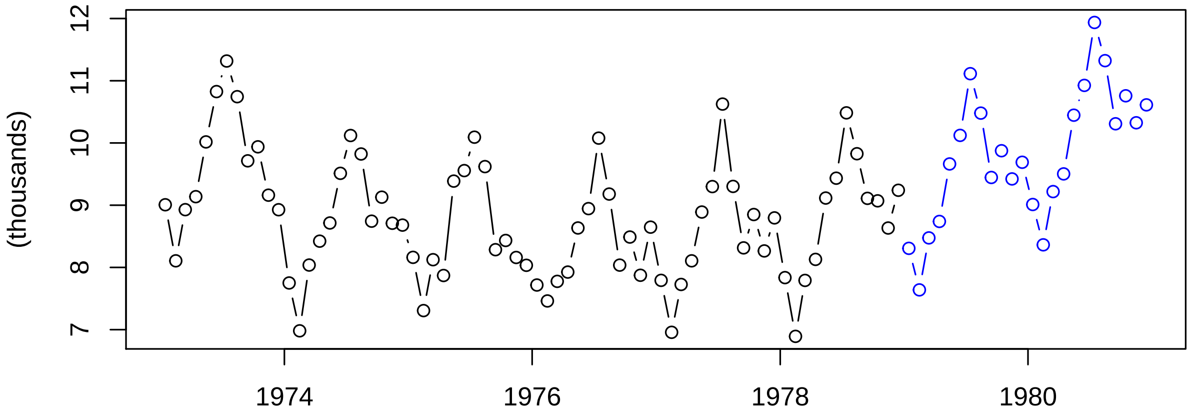



```r
y <- deaths - trend$fitted - st
acf(y,lag.max = 24)
```




```r
p.time <- seq(as.Date("1979-01-15"),by="month",length.out=24)

M <- poly(72+1:24,2,raw=TRUE)
m.hat <- predict(trend, newdata= M)
arima.mod<-arima(y)
yhat <- predict(arima.mod, n.ahead = 24)$pred + m.hat+rep(st,2)
plot(time,deaths/1000,xlim=range(time,p.time),type="b",
```

```
     ylim = range(deaths, yhat)/1000,
     ylab = "(thousands)", xlab ="")
lines(p.time,yhat/1000, type="b", col = "blue")
```




```
detach(data)
```

References

Brockwell Peter J, Davis Richard A, Calder Matthew V. Introduction to time series and forecasting. 3. 2016.