STAD57: Time Series Analysis
Problem Set 1

1. Do the following:

a. Download and install R (https://www.r-project.org/) & RStudio (https://www.rstudio.com/), if you don't have them already.

b. In R, there is a special function for creating time series called `ts()`. This function creates an object of type "ts". Try the following in R:

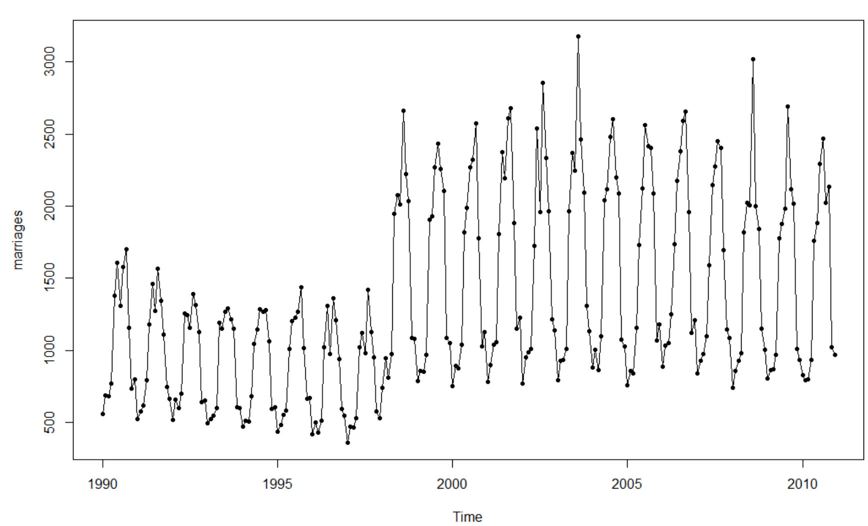
```
X=rnorm(120)
plot(X, pch=20)
X.ts=ts(data=X, start=c(2005,1), frequency=12 )
plot(X.ts)
class(X); class(X.ts)
```

c. The file `ontario_marriages.csv` contains counts of marriages in the province of Ontario by year, month, and city. Load these data in R and create & plot a time series of the total number of marriages in Toronto & Ottawa from Jan 1990 to Dec 2010. You can use the following starter code:

```
my_data=read.csv("ontario_marriages.csv")
X=aggregate(COUNT~MONTH+YEAR, data=my_data, function=sum,
  subset=(CITY %in% c('TORONTO','OTTAWA')) )
```

SOL:

```
idx=(X$YEAR>=1990) # Index for year >= 1990
marriages=ts(X$COUNT[idx],start=c(1990,1), frequency=12)
plot(marriages, type='o', pch=20)
```



2. Consider the series $\{X_1, X_2, X_3\}$, where X_1, X_2 are independent standard Normal and $X_3 = X_1 \times X_2$. Show that this is a White Noise series, i.e. that the three variables are uncorrelated. Is the series strictly stationary?

SOL:

Obviously, X_1, X_2 are uncorrelated since they are independent. We just need to show that X_3 is uncorrelated with X_1, X_2, i.e. show that $\mathrm{Cov}(X_1, X_3) = \mathrm{Cov}(X_2, X_3) = 0$.

We have $\mathrm{Cov}(X_1, X_3) = \mathrm{Cov}(X_1, X_1X_2) = \mathbb{E}\left(X_1(X_1X_2)\right) - \overbrace{\mathbb{E}(X_1)}^{=0}\mathbb{E}(X_1X_2) =$

$= \mathbb{E}(X_1^2X_2) \overset{\text{by indep.}}{=} \mathbb{E}(X_1^2)\underbrace{\mathbb{E}(X_2)}_{=0} = 0$, and similarly for $\mathrm{Cov}(X_2, X_3)$. Since the mean of all variables is (constant) 0: $\mathbb{E}(X_1) = \mathbb{E}(X_2) = 0$ & $\mathbb{E}(X_3) = \mathbb{E}(X_1X_2) \overset{\text{by indep.}}{=} \mathbb{E}(X_1)\mathbb{E}(X_2) = 0$ and the variance of all variables is (constant) 1: $\mathbb{V}(X_1) = \mathbb{V}(X_2) = 1$ &

$\mathbb{V}(X_3) = \mathbb{V}(X_1X_2) = \mathbb{E}\left((X_1X_2)^2\right) - \overbrace{\left(\mathbb{E}(X_1X_2)\right)^2}^{=0} = \mathbb{E}(X_1^2X_2^2) \overset{\text{by indep.}}{=} \mathbb{E}(X_1^2)\mathbb{E}(X_2^2) = 1 \cdot 1 = 1$,

the series is a weakly stationary & uncorrelated, i.e. WN(0,1). Nevertheless, the series is *not* strictly stationary, since the marginal distribution of $X_3 = X_1X_2$ is different from that of X_1, X_2.

3. Exercise 1.6 from textbook.

SOL:

a.

We can simply show that the mean is not constant

$E[X_t] = E[\beta_1 + \beta_2 t + W_t] = \beta_1 + \beta_2 t + \cancel{E[W_t]} = \beta_1 + \beta_2 t$, which is generally a function of *t*.

Moreover, we have: $Var[X_t] = Var[\beta_1 + \beta_2 t + W_t] = Var[W_t] = \sigma_W^2 = \gamma_X(0)$, and

$Cov[X_{t+h}, X_t] = Cov[\beta_1 + \beta_2(t+h) + W_{t+h}, \beta_1 + \beta_2 t + W_t] = Cov[W_{t+h}, W_t] = 0 = \gamma_X(h), \forall h \neq 0$

b.

We have $Y_t = X_t - X_{t-1} = \left(\cancel{\beta_1} + \cancel{\beta_2 t} + W_t\right) - \left(\cancel{\beta_1} + \beta_2(\cancel{t} - 1) + W_{t-1}\right) = \beta_2 + W_t - W_{t-1}$, so that:

$E[Y_t] = E\left[\beta_2 + W_t - W_{t-1}\right] = \beta_2 + \cancel{E[W_t]} - \cancel{E[W_{t-1}]} = \beta_2$, which is constant (indep. of *t*)

$Var[Y_t] = Var\left[\beta_2 + W_t - W_{t-1}\right] = Var[W_t] + Var[W_{t-1}] = 2\sigma_W^2 = \gamma_Y(0)$, which is also constant

$$Cov[Y_{t+h}, Y_t] = Cov\left[\left(\beta_2 + W_{t+h} - W_{t+h-1}\right), \left(\beta_2 + W_t - W_{t-1}\right)\right] =$$
$$= Cov\left[W_{t+h}, W_t\right] - Cov\left[W_{t+h}, W_{t-1}\right] - Cov\left[W_{t+h-1}, W_t\right] + Cov\left[W_{t+h-1}, W_{t-1}\right]$$

For $h=1$, $\gamma_Y(1) = \cancel{Cov[W_{t+1}, W_t]} - \cancel{Cov[W_{t+1}, W_{t-1}]} - Cov[W_t, W_t] + \cancel{Cov[W_t, W_{t-1}]} = -\sigma_W^2$

For $h \geq 2$, $\gamma_Y(h) = 0$ (there is no overlap in the $\{W_t\}$ time subscripts)

So $\{Y_t\}$ is stationary (it behaves like a MA(1) process)

c.

$$E[V_t] = E\left[\frac{1}{2q+1}\sum_{j=-q}^{q} X_{t-j}\right] = \frac{1}{2q+1}\sum_{j=-q}^{q} E\left[X_{t-j}\right] = \frac{1}{2q+1}\sum_{j=-q}^{q} \beta_1 + \beta_2(t-j) =$$

$$= \frac{1}{\cancel{(2q+1)}}\left\{\cancel{(2q+1)}(\beta_1 + \beta_2 t) - \cancel{\underbrace{\sum_{j=-q}^{q} j}_{=0}}\right\} = \beta_1 + \beta_2 t .$$

$$Cov[V_{t+h}, V_t] = Cov\left[\frac{1}{2q+1}\sum_{i=-q}^{q} X_{t+h-i}, \frac{1}{2q+1}\sum_{j=-q}^{q} X_{t-j}\right] =$$

$$= \frac{1}{(2q+1)^2}\sum_{i=-q}^{q}\sum_{j=-q}^{q} Cov\left[X_{t+h-i}, X_{t-j}\right] = \frac{1}{(2q+1)^2}\sum_{i=-q}^{q}\sum_{j=-q}^{q} \gamma_X\left((t+h-i)-(t-j)\right) =$$

$$= \frac{1}{(2q+1)^2}\sum_{i=-q}^{q}\sum_{j=-q}^{q} \underbrace{\gamma_X\left(h-i+j\right)}_{=\sigma_W^2 \text{ only if } h-i+j=0 \Leftrightarrow i=h+j} = \begin{cases} \dfrac{(2q+1)-h}{(2q+1)^2}\sigma_W^2, & \forall 0 \leq h \leq 2q \\ 0, & \forall h > 2q \end{cases}$$

Note that every non-zero term in the double sum will be equal to $\gamma_X(0) = \sigma_W^2$, which will happen only if $i = h + j$, for $i, j \in \{-q, \ldots, q\}$. The maximum number of non-zero terms occurs for $h = 0$, where we have $(2q-1) \times \gamma_X(0)$ terms. If $|h| \geq 2q-1$, then there are no non-zero terms, i.e. we can't have $i = h + j$ for any $i, j \in \{-q, \ldots, q\}$.

4. Exercise 1.7 from textbook.

SOL:

$$E[X_t] = E\left[W_{t-1} + 2W_t + W_{t+1}\right] = E\left[W_{t-1}\right] + 2E\left[W_t\right] + E\left[W_{t+1}\right] = 0$$

$$Cov[X_{t+h}, X_t] = Cov\left[W_{t+h-1} + 2W_{t+h} + W_{t+h+1}, W_{t-1} + 2W_t + W_{t+1}\right] =$$
$$= Cov\left[W_{t+h-1}, W_{t-1}\right] + 2Cov\left[W_{t+h-1}, W_t\right] + Cov\left[W_{t+h-1}, W_{t+1}\right] +$$
$$+2Cov\left[W_{t+h}, W_{t-1}\right] + 4Cov\left[W_{t+h}, W_t\right] + 2Cov\left[W_{t+h}, W_{t+1}\right] +$$
$$+Cov\left[W_{t+h+1}, W_{t-1}\right] + 2Cov\left[W_{t+h+1}, W_t\right] + Cov\left[W_{t+h+1}, W_{t+1}\right] =$$
$$\Rightarrow \gamma(h) = \begin{cases} (1+4+1)\sigma_W^2 = 6\sigma_W^2, & h=0 \\ (2+2)\sigma_W^2 = 4\sigma_W^2, & h=1 \\ \sigma_W^2 & h=2 \\ 0 & h \geq 2 \end{cases}$$
$$\Rightarrow \rho(h) = \frac{\gamma(h)}{\gamma(0)} = \begin{cases} 1, & h=0 \\ 4/6, & h=1 \\ 1/6 & h=2 \\ 0 & h \geq 2 \end{cases}$$

5. Exercise 1.14 from textbook.

SOL:

a.

$$E[Y_t] = E\left[\exp\left\{X_t\right\}\right] = \exp\left\{E[X_t] + Var[X_t]/2\right\} = \exp\left\{\mu_X + \gamma(0)/2\right\} = \mu_Y$$

b.

$$\gamma_Y(h) = Cov\left[Y_{t+h}, Y_t\right] = E[Y_{t+h}Y_t] - E[Y_{t+h}]E[Y_t] = E\left[\exp\left\{X_{t+h}\right\}\exp\left\{X_t\right\}\right] - \mu_Y^2 =$$
$$= E\left[\exp\left\{X_{t+h} + X_t\right\}\right] - \exp\left\{2\mu_X + \gamma(0)\right\}$$

But $\left(X_{t+h} + X_t\right) \sim N\left(E\left[X_{t+h} + X_t\right], Var\left[X_{t+h} + X_t\right]\right)$, where: $E\left[X_{t+h} + X_t\right] = 2\mu_X$ and $Var\left[X_{t+h} + X_t\right] = Var\left[X_{t+h}\right] + Var\left[X_t\right] + 2Cov\left[X_{t+h}, X_t\right] = 2\gamma(0) + 2\gamma(h),\ \Rightarrow$

$$\gamma_Y(h) = E\left[\exp\left\{X_{t+h} + X_t\right\}\right] - \exp\left\{2\mu_X + \gamma(0)\right\} =$$
$$= \exp\left\{2\mu_X + \gamma(0) + \gamma(h)\right\} - \exp\left\{2\mu_X + \gamma(0)\right\}$$
$$= \exp\left\{2\mu_X + \gamma(0)\right\}\left[\exp\left\{\gamma(h)\right\} - 1\right]$$

6. Exercise 1.15 from textbook.

SOL:

$$\mu_t = E[X_t] = E[W_t W_{t-1}] = E[W_t]E[W_{t-1}] = 0$$
$$\sigma_t^2 = Var[X_t] = Var[W_t W_{t-1}] = Var[W_t]Var[W_{t-1}] = \sigma_w^4$$
$$\gamma(1) = Cov\left[X_{t+1}, X_t\right] = Cov\left[W_{t+1}W_t, W_t W_{t-1}\right] =$$
$$= E[W_{t+1}W_t^2 W_{t-1}] = E[W_{t+1}]E[W_t^2]E[W_{t-1}] = 0$$
$$\gamma(h) = Cov\left[X_{t+h}, X_t\right] = Cov\left[W_{t+h}W_{t+h-1}, W_t W_{t-1}\right] =$$
$$= E\left[W_{t+h}W_{t+h-1}, W_t W_{t-1}\right] = E[W_{t+h}]E[W_{t+h-1}]E[W_t]E[W_{t-1}] = 0, \ \ \forall h \geq 1$$

So $\{X_t\}$ is stationary, and is actually a white noise with variance σ_w^4

7. Exercise 1.19 from textbook.

SOL:

a.

$$\mu_t = E[X_t] = E[\mu + W_t - .8W_{t-1}] = \mu + E[W_t] - .8E[W_{t-1}] = \mu$$

b.

For this MA(1) process: $\gamma(0) = \sigma_t^2 = \left[1 + (-.8)^2\right]\sigma_w^2 = 1.64\sigma_w^2$, $\gamma(1) = \gamma(-1) = -.8\sigma_w^2$, a $\gamma(h) = 0, \ \ \forall |h| \geq 2$. So, the sample mean variance becomes:

$$Var\left[\bar{X}\right] = \frac{1}{n}\sum_{h=-n}^{n}\left(1 - \frac{|h|}{n}\right)\gamma(h) = \frac{1}{n}\sum_{h=-1}^{1}\left(1 - \frac{|h|}{n}\right)\gamma(h) =$$
$$= \frac{1}{n}\left[\left(1 - \frac{1}{n}\right)\gamma(-1) + \gamma(0) + \left(1 - \frac{1}{n}\right)\gamma(1)\right] =$$
$$= \frac{1}{n}\left[\gamma(0) + 2\frac{n-1}{n}\gamma(1)\right] = \frac{1}{n}\left[1.64\sigma_w^2 - 2\frac{n-1}{n}.8\sigma_w^2\right] =$$
$$= \frac{\sigma_w^2}{n}\left[1.64 - 1.6\frac{n-1}{n}\right] \approx .04\frac{\sigma_w^2}{n}, \ \text{ for large } n$$

c.

If $\{X_t\}$ was just white noise, then the sample mean variance would just be the usual $\frac{\sigma_w^2}{n}$, which is always smaller than $\frac{\sigma_w^2}{n}\left[1.64 - 1.6\frac{n-1}{n}\right]$. The reason why this MA(1) model's variance is smaller is that the W's will partially cancel out because of the negative -.8 coefficient. To see this, compare what happens when n=2:

[MA(1)] $\bar{X} = \frac{1}{3}(X_1 + X_2 + X_3) = \frac{1}{3}\left[(\mu + W_1 - .8W_0) + (\mu + W_2 - .8W_1) + (\mu + W_3 - .8W_2)\right]$

$$= \frac{1}{3}\left[3\mu - .8W_0 + .2W_1 + .2W_2 + W_3\right] = \mu + \frac{1}{3}(-.8W_0 + .2W_1 + .2W_2 + W_3)$$

[WN] $\bar{X} = \frac{1}{3}(X_1 + X_2 + X_3) = \frac{1}{3}\left[(\mu + W_1) + (\mu + W_2) + (\mu + W_3)\right]$

$$= \frac{1}{3}\left[3\mu + W_1 + W_2 + W_3\right] = \mu + \frac{1}{3}(W_1 + W_2 + W_3)$$

Where $\frac{1}{3}(W_1 + W_2 + W_3)$ has higher variance than $\frac{1}{3}(-.8W_0 + .2W_1 + .2W_2 + W_3)$

8. Assume that the time series $\{Y_t\}$ is weakly stationary with mean μ and ACVF $\gamma_Y(h)$. Show that the differenced series $X_t = \nabla Y_t = Y_t - Y_{t-1}$ is also stationary and find its ACVF $\gamma_X(h)$.

SOL:

The mean of $\{X_t\}$ is constant equal to $\mathbb{E}(X_t) = \mathbb{E}(Y_t - Y_{t-1}) = \mathbb{E}(Y_t) - \mathbb{E}(Y_{t-1}) = \mu - \mu = 0$.

For the ACVF of $\{X_t\}$ let $t = s + h$, so that

$$\gamma_X(s,t) = \gamma_X(s,+h) = \mathrm{Cov}\left(X_s, X_{s+h}\right) = \mathrm{Cov}\left(Y_s - Y_{s-1}, Y_{s+h} - Y_{s+h-1}\right) =$$
$$= \mathrm{Cov}(Y_s, Y_{s+h}) - \mathrm{Cov}(Y_s, Y_{s+h-1}) - \mathrm{Cov}(Y_{s-1}, Y_{s+h}) + \mathrm{Cov}(Y_{s-1}, Y_{s+h-1}) =$$
$$= \gamma_Y(s, s+h) - \gamma_Y(s, s+h-1) - \gamma_Y(s-1, s+h) + -\gamma_Y(s-1, s+h-1) =$$
$$= \gamma_Y(h) - \gamma_Y(h-1) - \gamma_Y(h-1) + \gamma_Y(h) = 2\gamma_Y(h) - 2\gamma_Y(h-1)$$

The result is a function of the time-lag h, i.e. $\gamma_X(s,t) = \gamma_X(h) = 2\gamma_Y(h) - 2\gamma_Y(h-1)$, so $\{X_t\}$ is also weakly stationary.

STAD57: Time Series Analysis
Problem Set 2 Solutions

1. Use R to generate 400 observations from the following models:

a. $X_t = W_t + 2W_{t-1} - W_{t-2}$ (MA)

b. $X_t = -.8X_{t-1} + .4X_{t-2} + W_t$ (AR)

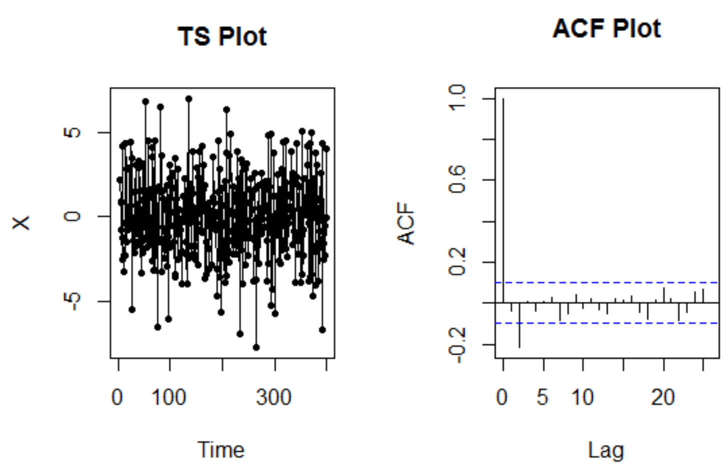
c. $X_t = .01 + X_{t-1} + W_t$ (Random walk)

(Hint: first generate a Normal white noise sequence $\{W_t\}$ with function `rnorm` and then use function `filter`). Create time series plots and ACF plots for all the series.

SOL:

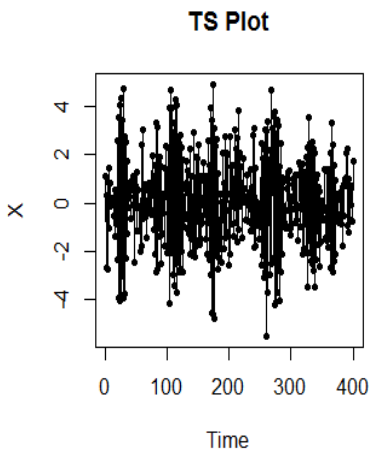
a.

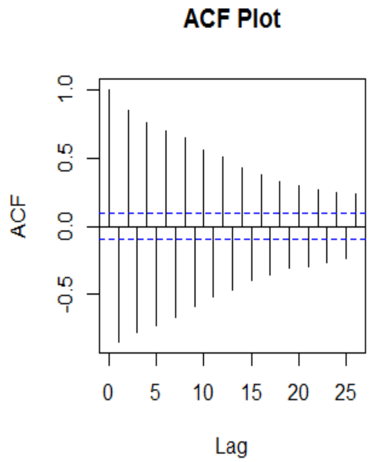
```
W=rnorm(n=400, mean=0, sd=1)
X=filter(W, c(1,2,-1), sides=1)
par(mfrow=c(1,2))
plot(X,type='o',pch=20, main="TS Plot")
acf(X, na.action = na.pass, main="ACF Plot")
```



b.

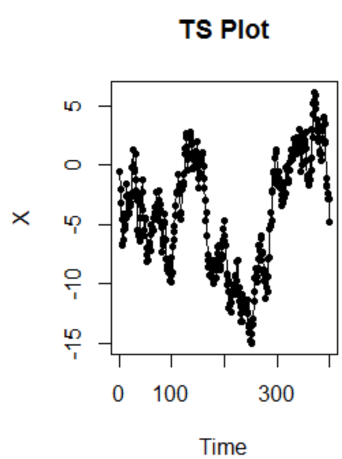
```
X=filter(W, c(-.5,.4), method = "recursive")
plot(X,type='o',pch=20, main="TS Plot")
acf(X, na.action = na.pass, main="ACF Plot")
```

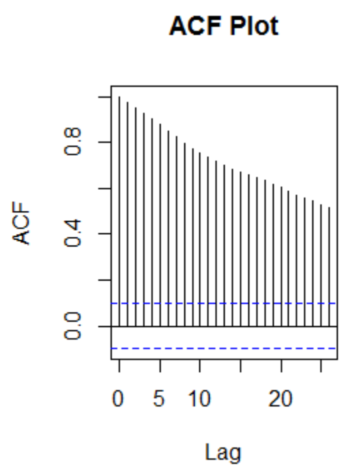




c.

```
X=filter(W+.01, c(1), method = "recursive")
plot(X,type='o',pch=20, main="TS Plot")
acf(X, na.action = na.pass, main="ACF Plot")
```





2. Load the following TS data sets in R and pre-process them (using de-trending, differencing, transformations etc) so that they become as close to stationary as possible.

a. Monthly Canadian reserves (in $)

b. Monthly car sales in Quebec (in # cars)

c. Daily average temperatures in Toronto (in °C)

Create time series plots and ACF plots for all of the original and processed series.

SOL:

There are typically more than one ways to model a time series, so the following answers are not strictly right (or wrong). R code is given at the end.

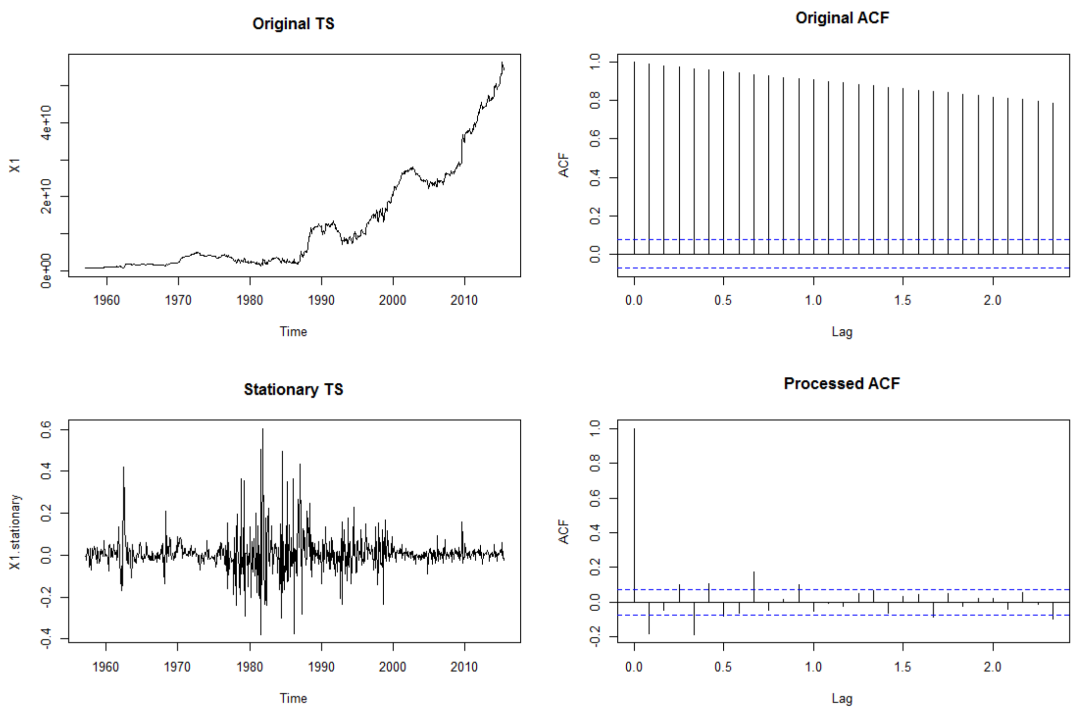
a.

I used first order differences of the logarithm of the series, i.e.:

$$Y_t = \nabla \log(X_t) = \log(X_t) - \log(X_{t-1}) = \log\left(X_t / X_{t-1}\right)$$

These are sometimes called log-returns, or continuously compounded returns (https://en.wikipedia.org/wiki/Rate_of_return#Logarithmic_or_continuously_compounded_return) and they are approximately equal to usual returns b/c

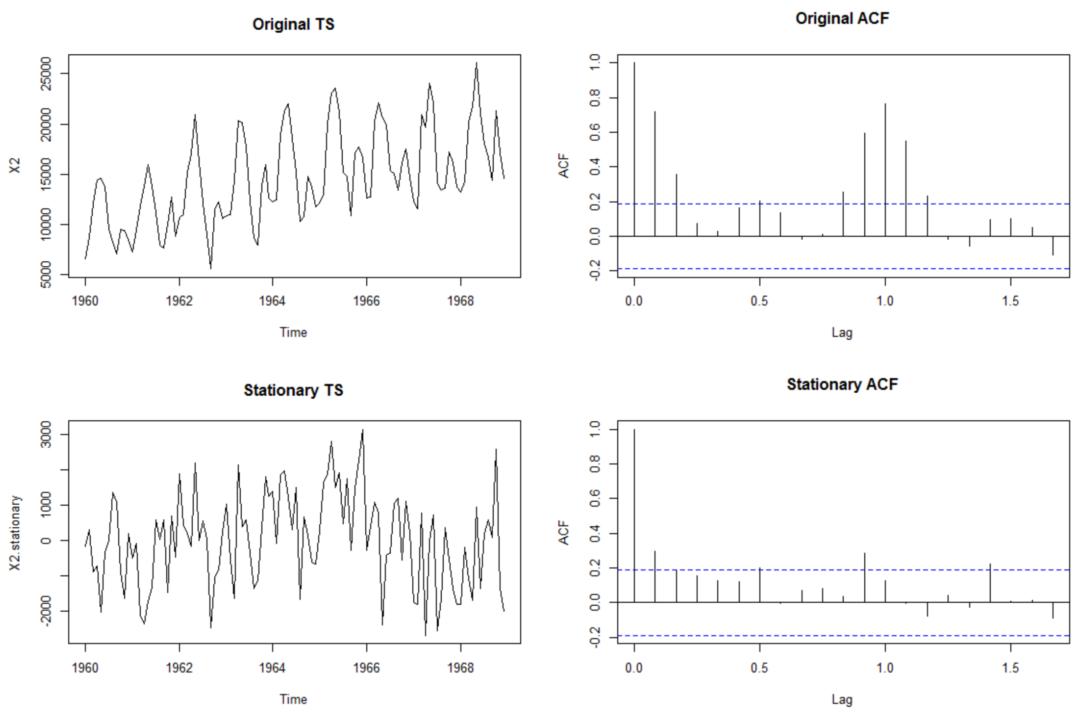
$\log(X_t) - \log(X_{t-1}) \approx \log'(X_{t-1})\nabla X_t = \frac{\nabla X_t}{X_{t-1}} = \frac{X_t - X_{t-1}}{X_{t-1}}$ (by using the 1st order Taylor expansion of the log function $\log(x+h) - \log(x) \underset{\text{for } h \to 0}{\approx} \log'(x) \times h$). The log-return (i.e. diff-log) transformation is very common for financial data (and has interesting modeling implications).



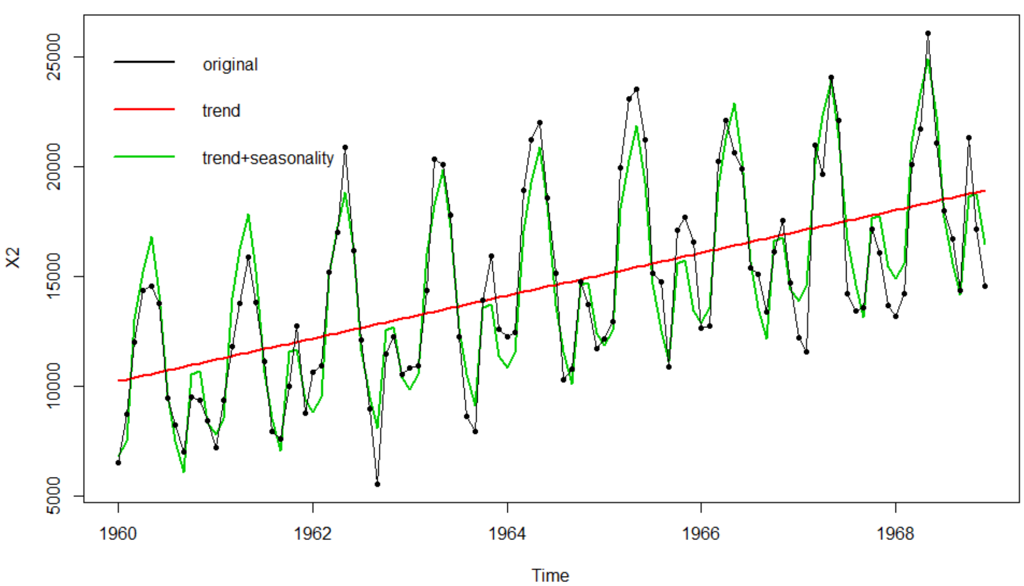
b.

I used detrending w/ linear trend and additive seasonality: $Y_t = X_t - T_t - S_t$,

where $T_t = \hat{\beta}_0 + \hat{\beta}_1 t$ from the regression model $X_t = \beta_0 + \beta_1 t + \varepsilon_t$, and $S_t = \hat{\mu}_{(t \bmod 12)}$ from the ANOVA model $X_t = \mu_{(t \bmod 12)} + \eta_t$

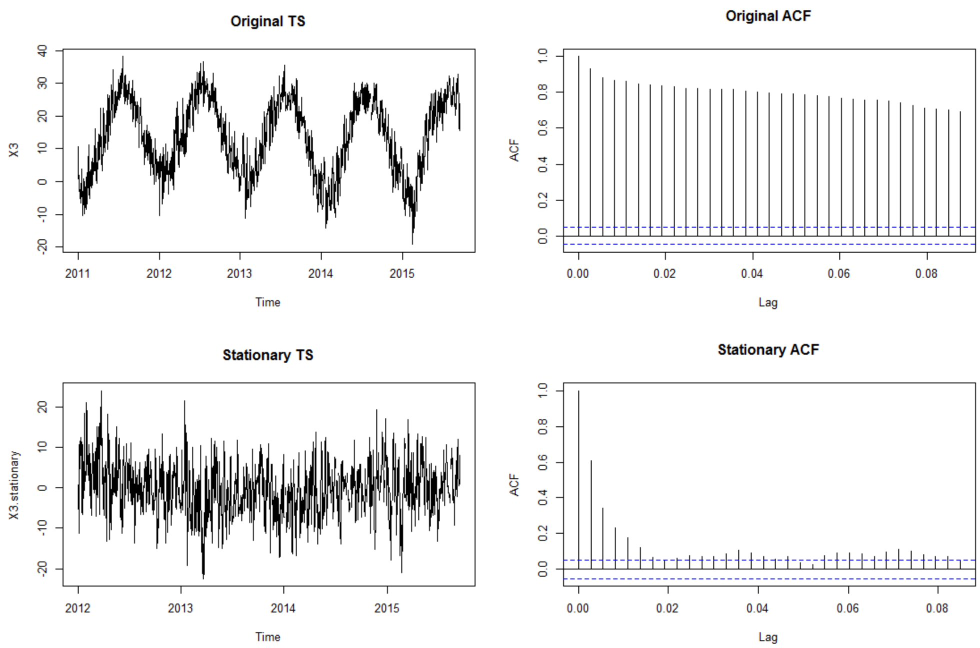


The plot of the fitted linear trend & seasonality is give below:



c.

Since these are daily temperature data, they should have an annual seasonal pattern. So, I just used differencing at lag 365 (i.e. 1 year): $Y_t = X_t - X_{t-365}$


Original TS
X3
Time
Original ACF
ACF
Lag
Stationary TS
X3.stationary
Time
Stationary ACF
ACF
Lag

STAD57: Time Series Analysis
Problem Set 3 Solutions

1. Exercise 3.1 from the textbook.

SOL:

We have $X_t = W_t + \theta W_{t-1}$, so that

$$\begin{aligned}\gamma_X(h) &= Cov\left(X_{t+h}, X_t\right) = Cov\left(W_{t+h} + \theta W_{t+h-1}, W_t + \theta W_{t-1}\right) = \\ &= Cov\left(W_{t+h}, W_t + \theta W_{t-1}\right) + \theta Cov\left(W_{t+h-1}, W_t + \theta W_{t-1}\right) = \\ &= Cov\left(W_{t+h}, W_t\right) + \theta Cov\left(W_{t+h}, W_{t-1}\right) + \theta Cov\left(W_{t+h-1}, W_t\right) + \theta^2 Cov\left(W_{t+h-1}, W_{t-1}\right) = \\ &= \begin{cases} \sigma_W^2(1+\theta^2), & h=0 \\ \sigma_W^2\theta, & h=1 \\ 0, & h\ge 2 \end{cases} \Rightarrow \rho_X(h) = \begin{cases} 1, & h=0 \\ \theta/(1+\theta^2), & h=1 \\ 0, & h\ge 2 \end{cases}\end{aligned}$$

To find the maximum/minimum of $\rho_X(1)$ we differentiate it w.r.t. to θ and set to 0:

$$\frac{d}{d\theta}\rho_X(h) = \frac{d}{d\theta}\left(\frac{\theta}{1+\theta^2}\right) = \frac{(1+\theta^2) - \theta(2\theta)}{(1+\theta^2)^2} = \frac{1-\theta^2}{(1+\theta^2)^2} = 0 \Rightarrow \theta = \pm 1$$

Substituting $\theta = \pm 1$ into $\rho_X(1)$ we get that the maximum/minimum values are:

$$\rho_X(1) = \frac{\theta}{1+\theta^2} = \frac{\pm 1}{1+(\pm 1)^2} = \pm\frac{1}{2}$$

(Note: to be technically correct we also have to check that the 2nd derivatives are negative/positive at the maximum/minimum, but you can check this is the case).

2. Exercise 3.2 from textbook.

SOL:

We have $\begin{cases} X_1 = W_1 \\ X_t = \varphi X_{t-1} + W_t, t \ge 2 \end{cases} \Rightarrow X_t = \varphi X_{t-1} + W_t = \varphi\left(\varphi X_{t-2} + W_{t-1}\right) + W_t =$

$$= \varphi^2\left(\varphi X_{t-3} + W_{t-2}\right) + \varphi W_{t-1} + W_t = \cdots = \varphi^{t-1}\underbrace{X_1}_{=W_1} + \varphi^{t-2}W_2 + \cdots + \varphi W_{t-1} + W_t \Rightarrow$$

$$\Rightarrow X_t = \varphi^{t-1}W_1 + \varphi^{t-2}W_2 + \cdots + \varphi W_{t-1} + W_t = \sum_{j=0}^{t-1}\varphi^j W_{t-j}$$

a) $\mathbb{E}\left[X_t\right] = \mathbb{E}\left[\sum_{j=0}^{t-1}\varphi^j W_{t-j}\right] = \sum_{j=0}^{t-1}\varphi^j \underbrace{\cancel{\mathbb{E}\left[W_{t-j}\right]}}_{=0} = 0$

$$\mathbb{V}\left[X_t\right] = \mathbb{V}\left[\sum_{j=0}^{t-1}\varphi^j W_{t-j}\right] = \sum_{j=0}^{t-1}\varphi^{2j}\underbrace{\mathbb{V}\left[W_{t-j}\right]}_{=\sigma_W^2} = \sigma_W^2\sum_{j=0}^{t-1}\varphi^{2j} = \sigma_W^2\frac{1-\varphi^{2t}}{1-\varphi^2}.$$

Since the variance depends on t, the series is not stationary.

b) $Cor\left(X_t, X_{t-h}\right) = Cor\left(\sum_{j=0}^{t-1}\varphi^j W_{t-j}, \sum_{k=0}^{t-1}\varphi^k W_{t-h-k}\right) = \sum_{j=0}^{t-1}\sum_{k=0}^{t-h-1}\varphi^j\varphi^k \underbrace{Cov\left(W_{t-j}, W_{t-h-k}\right)}_{=\begin{cases}\sigma_W^2, & j=h+k\\ 0, & j\neq h+k\end{cases}} =$

$$= \sigma_W^2 \sum_{k=0}^{t-h-1}\varphi^{k+h}\varphi^k = \varphi^h\left(\sigma_W^2\sum_{k=0}^{t-h-1}\varphi^{2k}\right) = \varphi^h\left(\sigma_W^2\,\frac{1-\varphi^{2(t-h)}}{1-\varphi^2}\right) = \varphi^h\mathbb{V}\left[X_{t-h}\right] \Rightarrow$$

$$\Rightarrow Cor\left(X_t, X_{t-h}\right) = \frac{Cov\left(X_t, X_{t-h}\right)}{\sqrt{\mathbb{V}\left[X_t\right]\mathbb{V}\left[X_{t-h}\right]}} = \frac{\varphi^h\mathbb{V}\left[X_{t-h}\right]}{\sqrt{\mathbb{V}\left[X_t\right]\mathbb{V}\left[X_{t-h}\right]}} = \varphi^h\sqrt{\frac{\mathbb{V}\left[X_{t-h}\right]}{\mathbb{V}\left[X_t\right]}}$$

c) Simply take limits of the above expressions as $t\to\infty$. We have:

$$\lim_{t\to\infty}\mathbb{V}\left[X_t\right] = \lim_{t\to\infty}\left(\sigma_W^2\,\frac{1-\varphi^{2t}}{1-\varphi^2}\right) = \sigma_W^2\,\frac{1-\lim_{t\to\infty}\left(\varphi^{2t}\right)}{1-\varphi^2} = \sigma_W^2\,\frac{1}{1-\varphi^2}\ \ (\text{since}\ |\varphi|<1)$$

$$\lim_{t\to\infty}Cor\left(X_t, X_{t-h}\right) = \lim_{t\to\infty}\left(\sigma_W^2\varphi^h\,\frac{1-\varphi^{2(t-h)}}{1-\varphi^2}\right) = \sigma_W^2\varphi^h\,\frac{1-\lim_{t\to\infty}\left(\varphi^{2(t-h)}\right)}{1-\varphi^2} = \sigma_W^2\,\frac{\varphi^h}{1-\varphi^2}.$$

d) To generate n values from the Gaussian AR(1) series, you can iteratively generate $N \gg n$ values starting from $X_1 = W_1 = \sigma_W Z_1 \sim N(0,\sigma_W^2)$ and using $X_t = \varphi X_{t-1} + W_t = \varphi X_{t-1} + \sigma_W Z_t,\ \forall t\geq 2$, where $\sigma_W Z_t = W_t \sim N(0,\sigma_W^2)$, but only keep the last n values.

e) If you start with $X_t = \dfrac{W_1}{\sqrt{1+\varphi^2}} \sim N\left(0, \dfrac{\sigma_W^2}{1-\varphi^2}\right)$, the resulting series will be stationary. We have $X_t = \varphi^{t-1}W_1 + \varphi^{t-2}W_2 + \cdots + \varphi W_{t-1} + W_t = \varphi^{t-1}W_1 + \sum_{j=0}^{t-2}\varphi^j W_{t-j}$. The mean will trivially be zero again, but the variance and covariances will be given by:

$$\mathbb{V}\left[X_t\right] = \mathbb{V}\left[\varphi^{t-1}X_1 + \sum_{j=0}^{t-2}\varphi^j W_{t-j}\right] = \varphi^{t-1}\mathbb{V}\left[X_1\right] + \sum_{j=0}^{t-2}\varphi^j\mathbb{V}\left[W_{t-j}\right] =$$

$$= \sigma_W^2\,\frac{\varphi^{t-1}}{1-\varphi^2} + \sigma_W^2\sum_{j=0}^{t-2}\varphi^j = \sigma_W^2\left(\frac{\varphi^{t-1}}{1-\varphi^2} + \frac{1-\varphi^{t-1}}{1-\varphi^2}\right) = \frac{\sigma_W^2}{1-\varphi^2}\ \ (\text{indep. of } t)$$

$$Cor\left(X_t, X_{t-h}\right) = Cor\left(\varphi^{t-1}X_1 + \sum_{j=0}^{t-2}\varphi^j W_{t-j}, \varphi^{t-h-1}X_1 + \sum_{k=0}^{t-h-2}\varphi^k W_{t-h-k}\right) =$$

$$= \varphi^{t-1}\varphi^{t-h-1}\underbrace{Cov\left(X_1, X_1\right)}_{=\mathbb{V}[X_1]} + \sum_{j=0}^{t-2}\sum_{k=0}^{t-h-2}\varphi^j\varphi^k\underbrace{Cov\left(W_{t-j}, W_{t-h-k}\right)}_{=\begin{cases}\sigma_W^2, & j=h+k\\ 0, & j\neq h+k\end{cases}} \Rightarrow$$

$$Cor\left(X_t, X_{t-h}\right) = \sigma_W^2\varphi^h\,\frac{\varphi^{2(t-h-1)}}{1-\varphi^2} + \sigma_W^2\sum_{k=0}^{t-h-2}\varphi^{k+h}\varphi^k = \sigma_W^2\varphi^h\left(\frac{\varphi^{2(t-h-1)}}{1-\varphi^2} + \sum_{k=0}^{t-h-2}\varphi^{2k}\right) =$$

$$= \sigma_W^2\varphi^h\left(\frac{\varphi^{2(t-h-1)}}{1-\varphi^2} + \frac{1-\varphi^{2(t-h-1)}}{1-\varphi^2}\right) = \sigma_W^2\,\frac{\varphi^h}{1-\varphi^2}\ (\text{indep. of } t)$$

$$\Rightarrow Cor\left(X_t, X_{t-h}\right) = \frac{Cov\left(X_t, X_{t-h}\right)}{\sqrt{\mathbb{V}\left[X_t\right]\mathbb{V}\left[X_{t-h}\right]}} = \frac{\sigma_W^2\,\dfrac{\varphi^h}{1-\varphi^2}}{\sigma_W^2\,\dfrac{1}{1-\varphi^2}} = \varphi^h,\ \ \forall h\geq 0$$

3. Exercise 3.4 from the textbook.

SOL:

a) $X_t = .8X_{t-1} - .15X_{t-2} + W_t - .3W_{t-1} \Leftrightarrow X_t - .8X_{t-1} + .15X_{t-2} = W_t - .3W_{t-1} \Leftrightarrow$

$\Leftrightarrow (1-.8B+.15B^2)X_t = (1-.3B)W_t \Leftrightarrow (1-.5B)\cancel{(1-.3B)}X_t = \cancel{(1-.3B)}W_t \Leftrightarrow$

$\Leftrightarrow (1-.5B)X_t = W_t \Leftrightarrow X_t = .5X_{t-1} + W_t$

which is an AR(1) model with $\varphi_1 = .5$. Since $|\varphi_1| < 1$, the model is causal (→stationary) and invertible (b/c all pure AR models are invertible).

b) $X_t = X_{t-1} - .5X_{t-2} + W_t - W_{t-1} \Leftrightarrow X_t - X_{t-1} + .5X_{t-2} = W_t - W_{t-1} \Leftrightarrow \underbrace{(1-B+.5B^2)}_{\varphi(B)}X_t = \underbrace{(1-B)}_{\theta(B)}W_t$

The roots of the AR polynomial are given by:

$$\varphi(z) = 0 \Rightarrow .5z^2 - z + 1 = 0 \Rightarrow z = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a} = \frac{-(-1) \pm \sqrt{(-1)^2 - 4(.5)(1)}}{2(.5)}$$

$$\Rightarrow z = \frac{1 \pm \sqrt{1-2}}{1} = 1 \pm \sqrt{-1} = 1 \pm i$$

These are conjugate complex roots with norm $|1 \pm i| = \sqrt{1^2 + 1^2} = \sqrt{2} > 1$, so the model is causal. However, the root of the MA polynomial $\theta(z) = 1 - z = 0$ is trivially $z = 1$, which lies *on* the unit circle → the model is *not* invertible.

4. Exercise 3.5 from the textbook.

SOL:

For the AR(2) model $\overbrace{(1-\varphi_1 B - \varphi_2 B^2)}^{=\varphi(B)}X_t = W_t$ to be stationary, we need the roots of the characteristic equation $1 - \varphi_1 z - \varphi_2 z^2 = 0$ to be outside the unit disk. Let z_1, z_2 be the possibly complex roots; we can equivalently write the equation as $1 - \varphi_1 z - \varphi_2 z^2 = \left(1 - z_1^{-1}z\right)\left(1 - z_2^{-1}z\right) =$
$= 1 - (z_1^{-1} + z_2^{-1})z + (z_1^{-1}z_2^{-1})z^2$, from which we get that $\varphi_1 = z_1^{-1} + z_2^{-1}$ & $\varphi_2 = -z_1^{-1}z_2^{-1}$. We want $|z_1|, |z_2| > 1 \Rightarrow |z_1^{-1}|, |z_2^{-1}| < 1$, which implies:

$$\varphi_1 = z_1^{-1} + z_2^{-1} \Rightarrow |\varphi_1| = |z_1^{-1} + z_2^{-1}| \le |z_1^{-1}| + |z_2^{-1}| < 1 + 1 = 2$$

$$\varphi_2 = -z_1^{-1}z_2^{-1} \Rightarrow |\varphi_2| = |-z_1^{-1}z_2^{-1}| \le |z_1^{-1}||z_2^{-1}| < 1$$

Moreover, the roots are given by $z = \dfrac{\varphi_1 \pm \sqrt{\varphi_1^2 + 4\varphi_2}}{-2\varphi_2}$, and they are real provided that

$\varphi_1^2 + 4\varphi_2 \ge 0 \Rightarrow \varphi_2 \ge -\dfrac{\varphi_1^2}{4}$. Assuming we have two real roots, let $z_1 \le z_2$ without loss of generality. For both roots to be outside the unit circle, we have 3 cases $\begin{cases} a: \quad 1 < z_1 \le z_2 \\ b: \quad z_1 \le z_2 < -1 \\ c: z_1 < -1 \,\&\, 1 < z_2 \end{cases}$. Let $\varphi_2 < 0$, so that $z_1 = \dfrac{\varphi_1 - \sqrt{\varphi_1^2 + 4\varphi_2}}{-2\varphi_2} \le \dfrac{\varphi_1 - \sqrt{\varphi_1^2 + 4\varphi_2}}{-2\varphi_2} = z_2$; the 3 case become:

$$a.\ 1 < z_1 \Rightarrow 1 < \frac{\varphi_1 - \sqrt{\varphi_1^2 + 4\varphi_2}}{-2\varphi_2} \Rightarrow \sqrt{\varphi_1^2 + 4\varphi_2} < \varphi_1 + 2\varphi_2 \Rightarrow \varphi_1^2 + 4\varphi_2 < (\varphi_1 + 2\varphi_2)^2 \Rightarrow$$

$$\Rightarrow \cancel{\varphi_1^2} + \cancel{4}\varphi_2 < \cancel{\varphi_1^2} + \cancel{4}\varphi_1\varphi_2 + \cancel{4}\varphi_2^2 \overset{(\text{div } \varphi_2)}{\Rightarrow} \boxed{1 > \varphi_1 + \varphi_2}$$

$$b.\ z_2 < -1 \Rightarrow \frac{\varphi_1 + \sqrt{\varphi_1^2 + 4\varphi_2}}{-2\varphi_2} < -1 \Rightarrow \sqrt{\varphi_1^2 + 4\varphi_2} < 2\varphi_2 - \varphi_1 \Rightarrow \varphi_1^2 + 4\varphi_2 < (2\varphi_2 - \varphi_1)^2 \Rightarrow$$

$$\Rightarrow \cancel{\varphi_1^2} + \cancel{4}\varphi_2 < \cancel{4}\varphi_2^2 - \cancel{4}\varphi_1\varphi_2 + \cancel{\varphi_1^2} \overset{(\text{div } \varphi_2)}{\Rightarrow} \boxed{1 > \varphi_2 - \varphi_1}$$

$$c.\ z_2 > 1 \Rightarrow \frac{\varphi_1 + \sqrt{\varphi_1^2 + 4\varphi_2}}{-2\varphi_2} > 1 \Rightarrow \sqrt{\varphi_1^2 + 4\varphi_2} > 2\varphi_2 + \varphi_1 \Rightarrow \varphi_1^2 + 4\varphi_2 > (2\varphi_2 + \varphi_1)^2 \Rightarrow$$

$$\Rightarrow \cancel{\varphi_1^2} + \cancel{4}\varphi_2 > \cancel{4}\varphi_2^2 + \cancel{4}\varphi_1\varphi_2 + \cancel{\varphi_1^2} \overset{(\text{div } \varphi_2)}{\Rightarrow} \boxed{1 < \varphi_2 - \varphi_1},\ AND$$

$$z_1 < -1 \Rightarrow \frac{\varphi_1 - \sqrt{\varphi_1^2 + 4\varphi_2}}{-2\varphi_2} < -1 \Rightarrow \sqrt{\varphi_1^2 + 4\varphi_2} > 2\varphi_2 + \varphi_1 \Rightarrow \varphi_1^2 + 4\varphi_2 > (2\varphi_2 + \varphi_1)^2 \Rightarrow$$

$$\Rightarrow \cancel{\varphi_1^2} + \cancel{4}\varphi_2 > \cancel{4}\varphi_2^2 + \cancel{4}\varphi_1\varphi_2 + \cancel{\varphi_1^2} \overset{(\text{div } \varphi_2)}{\Rightarrow} \boxed{1 < \varphi_2 + \varphi_1}$$

Note that for case c. we need $(1 < \varphi_2 + \varphi_1) \& (1 < \varphi_2 - \varphi_1) \Rightarrow \varphi_2 > 1$, which the assumption $\varphi_2 < 0$, so we reject it. Repeating the above analysis for $\varphi_2 > 0$, we get similar results. So, overall, we have the conditions $(|\varphi_2| < 1)\ \&\ (\varphi_2 + \varphi_1 < 1)\ \&\ (\varphi_2 - \varphi_1 < 1)$.

5. Exercise 3.6 from the textbook.

SOL:

We have $X_t = -.9X_{t-2} + W_t \Leftrightarrow X_t + .9X_{t-2} = W_t \Leftrightarrow \underbrace{(1 + 0B + .9B^2)}_{\varphi(B)} X_t = W_t$. The roots of the AR polynomial are: $\varphi(z) = 1 + .9z^2 = 0 \Rightarrow z^2 = -\dfrac{1}{.9} \Rightarrow z = \pm\dfrac{i}{\sqrt{.9}}$. The ACF is given by

$\rho(h) = -.9\rho(h-2), \forall h \geq 2$ with initial conditions $\rho(0) = 1$ $\rho(1) = -.9\rho(-1) = -.9\rho(1) \Rightarrow \rho(1) = 0$

$$\Rightarrow \begin{cases} \rho(2) = -.9\rho(0) = -.9 \\ \rho(3) = -.9\rho(1) = 0 \\ \rho(4) = -.9\rho(2) = (-.9)^2 \Rightarrow \rho(h) = \begin{cases} (-.9)^{h/2}, & h = 0, 2, 4, \dots \\ 0, & h = 1, 3, 5, \dots \end{cases} \\ \rho(5) = -.9\rho(3) = 0 \\ \vdots \end{cases}$$

6. Exercise 3.7 from the textbook.

SOL:

a) $X_t + 1.6X_{t-1} + .64X_{t-2} = W_t \Rightarrow X_t = -1.6X_{t-1} - .64X_{t-2} + W_t \Rightarrow \rho(h) = -1.6\rho(h-1) - .64\rho(h-2) \Rightarrow$

$$\rho(0) = 1$$

$$\rho(1) = -1.6\rho(0) - .64\rho(-1) = -1.6 - .64\rho(1) \Rightarrow \rho(1) = -\frac{1.6}{1.64} = -0.9756098$$

$$\rho(2) = -1.6\rho(1) - .64\rho(0) = 1.6\frac{1.6}{1.64} - .64 = 0.9209756$$

$$\rho(3) = -1.6\rho(2) - .64\rho(1) = \cdots = -0.8491707$$

$$\vdots$$

From R, we get:

```
> ARMAacf(ar=c(-1.6,-.64), ma=0, 10)
         0          1          2          3          4 
 1.0000000 -0.9756098  0.9209756 -0.8491707  0.7692488 
         5          6          7          8          9 
-0.6873288  0.6074068 -0.5319605  0.4623964 -0.3993796 
        10 
 0.3430736 
```

b) $X_t - .4X_{t-1} - .45X_{t-2} = W_t \Rightarrow X_t = .4X_{t-1} + .45X_{t-2} + W_t \Rightarrow \rho(h) = .4\rho(h-1) + .45\rho(h-2) \Rightarrow$

$$\rho(0) = 1$$

$$\rho(1) = .4\rho(0) + .45\rho(-1) = .4 + .45\rho(1) \Rightarrow \rho(1) = \frac{.4}{.55} = 0.7272727$$

$$\rho(2) = .4\rho(1) + .45\rho(0) = .4\frac{.4}{.55} - .45 = 0.7409091$$

$$\rho(3) = .4\rho(2) + .45\rho(1) = \cdots = 0.6236364$$

$$\vdots$$

```
> ARMAacf(ar=c(.4,.45), ma=0, 10)
        0         1         2         3         4         5 
1.0000000 0.7272727 0.7409091 0.6236364 0.5828636 0.5137818 
        6         7         8         9        10 
0.4678014 0.4183224 0.3778396 0.3393809 0.3057802 
```

c) $X_t - 1.2X_{t-1} + .85X_{t-2} = W_t \Rightarrow X_t = 1.2X_{t-1} - .85X_{t-2} + W_t \Rightarrow \rho(h) = 1.2\rho(h-1) - .85\rho(h-2) \Rightarrow$

$$\rho(0) = 1$$

$$\rho(1) = 1.2\rho(0) - .85\rho(-1) = 1.2 - .85\rho(1) \Rightarrow \rho(1) = \frac{1.2}{1.85} = 0.64864865$$

$$\rho(2) = 1.2\rho(1) - .85\rho(0) = 1.2\frac{1.2}{1.85} - .85 = -0.07162162$$

$$\rho(3) = 1.2\rho(2) - .85\rho(1) = \cdots = -0.63729730$$

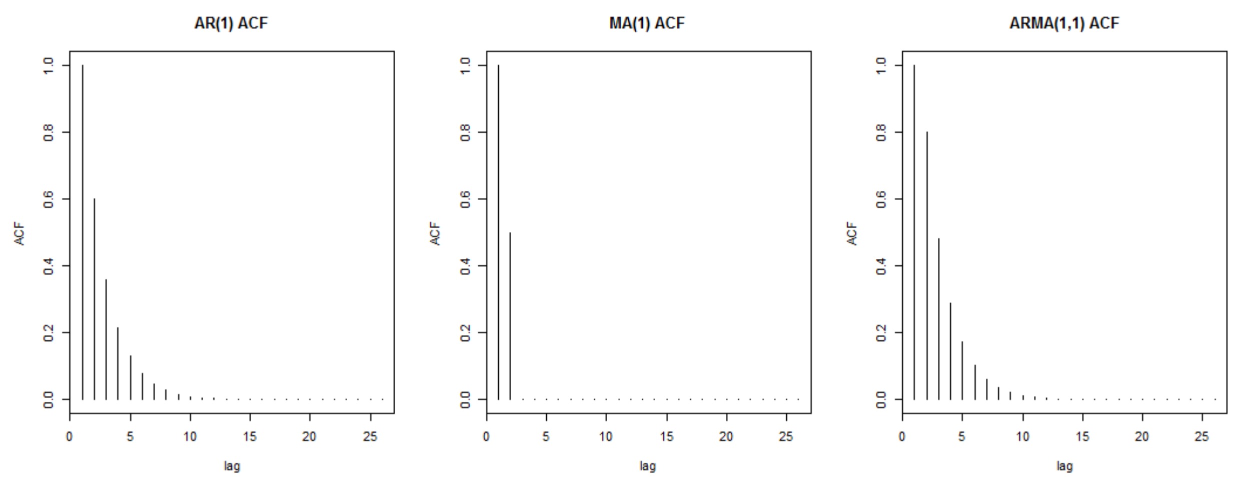
$$\vdots$$

```
> ARMAacf(ar=c(1.2,-.85), ma=0, 10)
          0           1           2           3           4 
 1.00000000  0.64864865 -0.07162162 -0.63729730 -0.70387838 
          5           6           7           8           9 
-0.30295135  0.23475500  0.53921465  0.44751583  0.07868654 
         10 
-0.28596460 
```

7. Exercise 3.8 from the textbook.

SOL:

The calculations of the general ARMA(1,1) ACF are verified in the example from Lecture 7, p. 14-15. The resulting ACF's for the various models using $\varphi = .6 \;\&\; \theta = .9$ are:



Code:

```
# Q3.8

phi=.6; theta=.9
AR_1.acf=ARMAacf(ar=phi, lag=25)
MA_1.acf=ARMAacf(ma=theta, lag=25)
ARMA_1.1.acf=ARMAacf(ar=phi, ma=theta, lag=25)

par(mfrow=c(1,3))
```

```
plot(AR_1.acf, type='h', main="AR(1) ACF", xla='lag', ylab='ACF' )
plot(MA_1.acf, type='h', main="MA(1) ACF", xla='lag', ylab='ACF' )
plot(ARMA_1.1.acf, type='h', main="ARMA(1,1) ACF", xla='lag',
ylab='ACF' )
```

8. Determine whether the following AR models are stationary, and calculate the first 4 coefficients of their causal representation.

   a. $X_t = -.8X_{t-1} + .4X_{t-2} + W_t$

   b. $X_t = -.5X_{t-1} + .4X_{t-2} + W_t$

SOL:

a.

The AR model is stationary if its characteristic polynomial has all roots outside the unit circle.

We have $X_t + .8X_{t-1} - .4X_{t-2} = W_t \Rightarrow \varphi(B)X_t = W_t$ where $\varphi(z) = 1 + .8z - .4z^2 \Rightarrow$

$$z = \frac{-.8 \pm \sqrt{(.8)^2 - 4\times(-.4)\times 1}}{2\times(-.4)} = \frac{.8 \pm \sqrt{.64 + 1.6}}{.8} = 1 \pm \sqrt{\frac{2.24}{.8}} = 1 \pm \sqrt{3.5} = \begin{cases} 2.870829 \\ -0.8708287 \end{cases}$$

Since only one root is outside the unit circle, the model is not stationary, and it does not have a causal representation (i.e. you can try to invert the AR polynomial operator, but the resulting ψ weights will be exploding to ∞. The first 4 terms are $\psi_1 = -0.8,\ \psi_2 = 1.04,\ \psi_3 = -1.152,$

$\psi_4 = 1.3376,\ \psi_5 = -1.53088$.

b.

We have $X_t + .5X_{t-1} - .4X_{t-2} = W_t \Rightarrow \varphi(B)X_t = W_t$ where $\varphi(z) = 1 + .5z - .4z^2 \Rightarrow$

$$z = \frac{-.5 \pm \sqrt{(.5)^2 - 4\times(-.4)\times 1}}{2\times(-.4)} = \frac{.5 \pm \sqrt{.25 + 1.6}}{.8} = \frac{.5 \pm \sqrt{1.85}}{.8} = \begin{cases} 2.325184 \\ -1.075184 \end{cases}$$

Since *both* roots are outside the unit disk (i.e. $|z| > 1$), the model is stationary and its causal representation is given by:

$$\psi(B)\varphi(B) = 1 \Leftrightarrow (1 + \psi_1 B + \psi_2 B^2 + \cdots)(1 + .5B - .4B^2) = 1$$

$$\Rightarrow 1 + (\psi_1 + .5)B + (\psi_2 + .5\psi_1 - .4)B^2 + \cdots = 1$$

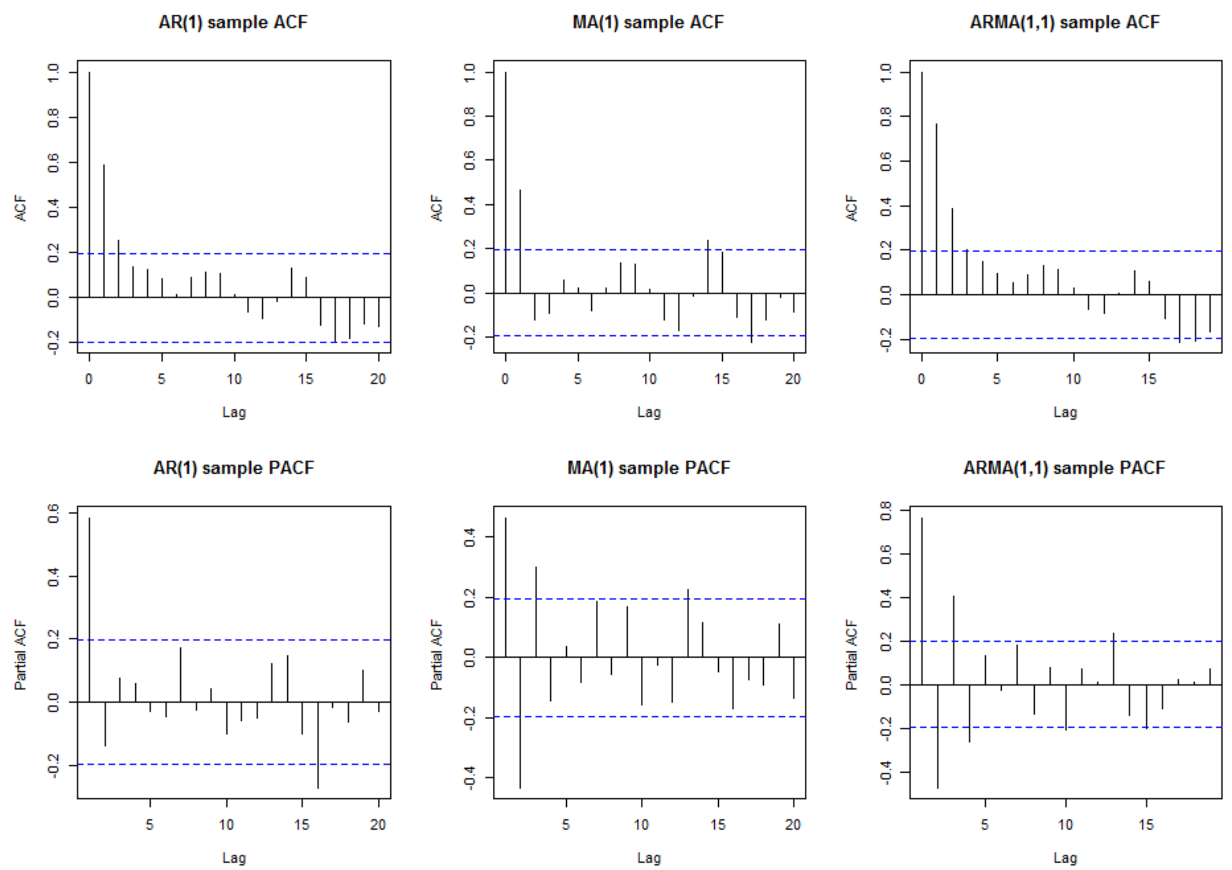
$$\Rightarrow \begin{cases} \psi_1 + .5 = 0 \Rightarrow \psi_1 = -.5 \\ \psi_2 + .5\psi_1 - .4 = 0 \Rightarrow \psi_2 = -.5\psi_1 + .4 = .5^2 + .4 = .65 \\ \psi_3 + .5\psi_2 - .4\psi_1 = 0 \Rightarrow \psi_3 = -.5\psi_2 + .4\psi_1 = -.5\times .65 - .4\times .5 = -.525 \\ \psi_4 + .5\psi_3 - .4\psi_2 = 0 \Rightarrow \psi_4 = -.5\psi_3 + .4\psi_2 = +.5\times .525 + .4\times .65 = .5225 \end{cases}$$

STAD57: Time Series Analysis
Problem Set 4

1. Exercise 3.9 from the textbook.

SOL:

The sample ACF/PACF's roughly match their theoretical behavior based on Table 3.1. Had we used a larger sample size, e.g. $n \geq 1000$, the match would be even closer.



2. Exercise 3.10 from the textbook.

a. The OLS-fitted model is: $X_t = 11.45 + .4286 X_{t-1} + .4418 X_{t-2} + W_t,\ \{W_t\} \sim \mathrm{WN}(0, \sigma_W^2 = 32.32)$.

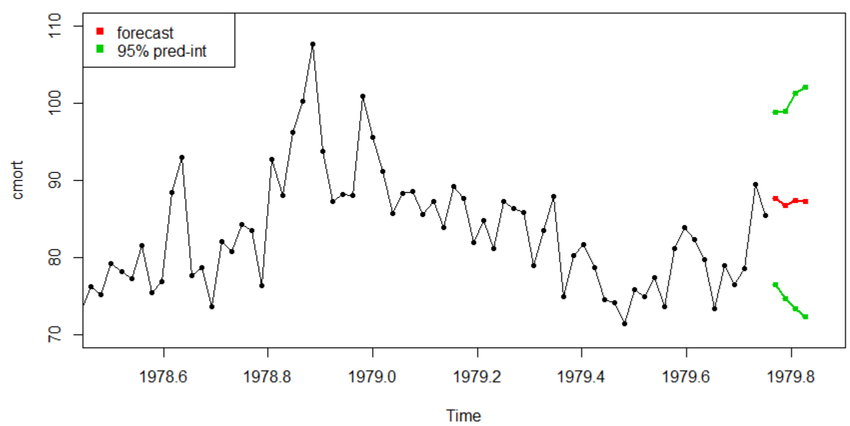
```
Call:
ar.ols(x = cmort, order.max = 2, demean = FALSE, intercept = TRUE)

Coefficients:
     1       2
0.4286  0.4418

Intercept: 11.45 (2.394)

Order selected 2  sigma^2 estimated as  32.32
```

b. The forecasts & 95% prediction interval are shown below



3. Exercise 3.11 from the textbook.

a. For the MA(1) model $X_t = W_t + \theta W_{t-1}$ the causal weights are $\psi_0 = 1, \psi_1 = \theta,$ and $\psi_j = 0, \forall j \ge 2$, and the invertible weights are $\pi_j = (-\theta)^j, \forall j \ge 0$. Thus, the 1-step-ahead BLP based on the infinite past of the series is $\tilde{X}_{n+1} = -\sum_{j=1}^{\infty} \pi_j X_{n+1-j} = -\sum_{j=1}^{\infty} (-\theta)^j X_{n+1-j}$, and its MSE is $P_{n+1}^n = \sigma_w^2 \sum_{j=0}^{1-1} \psi_j^2 = \sigma_w^2$.

b. For truncated prediction using eqn. (3.92) in the textbook, we have $\tilde{X}_{n+1}^n = \theta \tilde{W}_n^n$, where $\tilde{W}_0^n = 0$ & $\tilde{W}_t^n = X_t - \theta \tilde{W}_{t-1}^n$, so that:

$\tilde{W}_1^n = X_1$

$\tilde{W}_2^n = X_2 - \theta \tilde{W}_1^n = X_2 - \theta X_1$

$\tilde{W}_3^n = X_3 - \theta \tilde{W}_2^n = X_3 - \theta\left(X_2 - \theta X_1\right) = X_3 - \theta X_2 + (-\theta)^2 X_1$

$\vdots$

$\tilde{W}_n^n = X_n - \theta X_{n-1} + (-\theta)^2 X_{n-2} + \cdots + (-\theta)^{n-1} X_1 = \sum_{j=0}^{n-1} (-\theta)^j X_{n-j}$

The truncated 1-step-ahead predictor becomes

$$\tilde{X}_{n+1}^{n} = \theta \tilde{W}_{n}^{n} = \theta\left(\sum_{j=0}^{n-1}(-\theta)^{j} X_{n-j}\right) = -\sum_{j=1}^{n}(-\theta)^{j} X_{n+1-j}$$

Note that this is the same as the truncated 1-step-ahead formula given in eqn. (3.91), where

$$\tilde{X}_{n+1}^{n} = -\sum_{j=1}^{n}\pi_{j} X_{n+1-j} = -\sum_{j=1}^{n}(-\theta)^{j} X_{n+1-j}$$

The 1-step-ahead MSE is $\mathbb{E}\left[\left(X_{n+1}-\tilde{X}_{n+1}^{n}\right)^{2}\right] = \mathbb{E}\left[\left(X_{n+1}+\sum_{j=1}^{n}(-\theta)^{j} X_{n+1-j}\right)^{2}\right]$

$$= \mathbb{E}\left[\left(\sum_{j=0}^{n}(-\theta)^{j} X_{n+1-j}\right)^{2}\right] = \sum_{j=0}^{n}\sum_{k=0}^{n}(-\theta)^{j}(-\theta)^{j}\mathbb{E}\left[X_{n+1-j}X_{n+1-k}\right]$$

$$= \sum_{j=0}^{n}\sum_{k=0}^{n}(-\theta)^{j}(-\theta)^{j}\gamma(|j-k|)$$

The auto-covariance function of the MA(1) model is $\gamma(h) = \begin{cases}\sigma_w^2(1+\theta^2), & h=0\\ \sigma_w^2\theta, & h=1\\ 0, & h\geq 2\end{cases}$, so

$$\begin{aligned}
\mathbb{E}\left[\left(X_{n+1}-\tilde{X}_{n+1}^{n}\right)^{2}\right] &= \gamma(0)\sum_{j=0}^{n}(-\theta)^{2j} + 2\gamma(1)\sum_{j=0}^{n-1}(-\theta)^{2j+1}\\
&= \sigma_w^2(1+\theta^2)\sum_{j=0}^{n}\theta^{2j} + 2\sigma_w^2\theta\sum_{j=0}^{n-1}(-\theta)\theta^{2j}\\
&= \sigma_w^2(1+\theta^2)\frac{1-\theta^{2(n+1)}}{1-\theta^2} - 2\sigma_w^2\theta^2\frac{1-\theta^{2n}}{1-\theta^2}\\
&= \frac{\sigma_w^2}{1-\theta^2}\left(1+\cancel{\theta^2}-\cancel{\theta^{2n+2}}-\theta^{2n+4}-\cancel{2}\theta^2+\cancel{2}\theta^{2n+2}\right)\\
&= \frac{\sigma_w^2}{1-\theta^2}\left(1-\theta^{2n+4}-\theta^2+\theta^{2n+2}\right)\\
&= \frac{\sigma_w^2\left(1-\theta^2\right)\left(1+\theta^{2n+2}\right)}{1-\theta^2} = \sigma_w^2\left(1+\theta^{2n+2}\right)
\end{aligned}$$

Note that $\mathbb{E}\left[\left(X_{n+1}-\tilde{X}_{n+1}^{n}\right)^{2}\right] = \sigma_w^2\left(1+\theta^{2n+2}\right) \to P_{n+1}^{n} = \sigma_w^2$ as $n\to\infty$ (since $|\theta|<1$). So, the MSE of the truncated prediction converges exponentially fast (in the sample size *n*) to the MSE of the optimal predictor given the infinite past.

4. Exercise 3.14 from the textbook.

a) The MSE is $\mathbb{E}\left[\left(Y-g(X)\right)^{2}\right]$. Using the law of total expectation (a.k.a. tower law), we get $\mathbb{E}\left[\left(Y-g(X)\right)^{2}\right] = \mathbb{E}\left[\mathbb{E}\left[\left(Y-g(X)\right)^{2} \mid X\right]\right]$, so in order to minimize the MSE, we have to minimize the conditional expectation $\mathbb{E}\left[\left(Y-g(X)\right)^{2} \mid X\right]$, for any value of X. But we know that for any random variable Y, the value of c that minimizes $\mathbb{E}\left[\left(Y-c\right)^{2}\right]$ is the mean of Y, i.e. $c=\mathbb{E}[Y]$. Similarly for conditional expectations, the value of $g(X)$ that minimizes

$\mathbb{E}\left[\left(Y-g(X)\right)^2 \mid X\right]$ is $g(X)=\mathbb{E}\left[Y \mid X\right]$ (note that we can view $g(X)$ as a constant here, because, given X, any function of X behaves like a constant in the conditional expectation). For the model $Y = X^2 + Z$, where $X, Z \sim^{iid} N(0,1)$, we have:

$$g(X) = E\left[Y \mid X\right] = \mathbb{E}\left[X^2 + Z \mid X\right] = X^2 + \mathbb{E}\left[Z \mid X\right] = X^2$$

Using this predictor, the minimum MSE is:

$$\mathbb{E}\left[\left(Y-g(X)\right)^2\right] = \mathbb{E}\left[\left(X^2 + Z - X^2\right)^2\right] = \mathbb{E}\left[Z^2\right] = 1$$

b) If we restrict ourselves to linear functions $g(X) = a + bX$, the optimal (minimum MSE) parameter values are given by:

$$\begin{cases} \mathbb{E}\left[\left(Y-g(X)\right)1\right]=0 \\ \mathbb{E}\left[\left(Y-g(X)\right)X\right]=0 \end{cases} \Rightarrow \begin{cases} \mathbb{E}\left[X^2+Z-a-bX\right]=0 \\ \mathbb{E}\left[(X^2+Z-a-bX)X\right]=0 \end{cases}$$

$$\Rightarrow \begin{cases} \mathbb{E}[X^2]+\mathbb{E}[Z]-a-b\mathbb{E}[X]=0 \\ \mathbb{E}[X^3]+\mathbb{E}[ZX]-a\mathbb{E}[X]-b\mathbb{E}[X^2]=0 \end{cases}$$

$$\Rightarrow \begin{cases} 1-a=0 \\ b=0 \end{cases} \Rightarrow \begin{cases} a=1 \\ b=0 \end{cases}$$

Thus, the BLP is constant, $g(X)=1$, and its MSE is:

$$\mathbb{E}\left[\left(Y-g(X)\right)^2\right] = \mathbb{E}\left[\left(X^2+Z-1\right)^2\right] = \mathbb{E}\left[(X^2-1)^2+Z^2+2Z(X^2-1)^2\right] =$$

$$= \mathbb{E}\left[(X^2-1)^2\right] + \mathbb{E}[Z^2] + 2\mathbb{E}[Z]\mathbb{E}\left[(X^2-1)\right] =$$

$$= \mathbb{E}[X^4+1-2X^2]+1 = \mathbb{E}[X^4]+2-2\mathbb{E}[X^2] = 3+2-2 = 3$$

5. Exercise 3.15 from the textbook.

For the AR(1) model $X_t = \varphi X_{t-1} + W_t$ we have $\gamma(h) = \sigma_w^2 \dfrac{\varphi^h}{1-\varphi^2}$, $\forall h \geq 0$. Also, the m-step-ahead BLP is $X_{n+m}^n = \sum_{j=1}^n \varphi_{nj}^{(m)} X_{n+1-j}$, where the coefficients are given by

$$\sum_{j=1}^{n}\varphi_{nj}^{(m)}\gamma(k-j)=\gamma(m+k-1),\ \ \forall k=1,\ldots,n \Leftrightarrow$$

$$\Leftrightarrow\left\{\begin{array}{c}(k=1)\ \ \varphi_{n1}^{(m)}\gamma(1-1)+\varphi_{n2}^{(m)}\gamma(1-2)+\cdots+\varphi_{nn}^{(m)}\gamma(1-n)=\gamma(m+1-1)\\(k=2)\ \ \varphi_{n1}^{(m)}\gamma(2-1)+\varphi_{n2}^{(m)}\gamma(2-2)+\cdots+\varphi_{nn}^{(m)}\gamma(2-n)=\gamma(m+2-1)\\\vdots\\(k=n)\ \ \varphi_{n1}^{(m)}\gamma(n-1)+\varphi_{n2}^{(m)}\gamma(n-2)+\cdots+\varphi_{nn}^{(m)}\gamma(n-n)=\gamma(m+n-1)\end{array}\right\}$$

$$\Leftrightarrow\left\{\begin{array}{c}\varphi_{n1}^{(m)}\gamma(0)+\varphi_{n2}^{(m)}\gamma(1)+\cdots+\varphi_{nn}^{(m)}\gamma(n-1)=\gamma(m)\\\varphi_{n1}^{(m)}\gamma(1)+\varphi_{n2}^{(m)}\gamma(0)+\cdots+\varphi_{nn}^{(m)}\gamma(n-2)=\gamma(m+1)\\\vdots\\\varphi_{n1}^{(m)}\gamma(n-1)+\varphi_{n2}^{(m)}\gamma(n-2)+\cdots+\varphi_{nn}^{(m)}\gamma(0)=\gamma(m+n-1)\end{array}\right\}$$

$$\Leftrightarrow\underbrace{\begin{bmatrix}\gamma(0)&\gamma(1)&\cdots&\gamma(n-1)\\\gamma(1)&\gamma(0)&\ddots&\vdots\\\vdots&\ddots&\ddots&\gamma(1)\\\gamma(n-1)&\cdots&\gamma(1)&\gamma(0)\end{bmatrix}}_{=\boldsymbol{\Gamma}_n}\times\underbrace{\begin{bmatrix}\varphi_{n1}^{(m)}\\\varphi_{n2}^{(m)}\\\vdots\\\varphi_{nn}^{(m)}\end{bmatrix}}_{=\boldsymbol{\varphi}_n^{(m)}}=\underbrace{\begin{bmatrix}\gamma(m)\\\gamma(m+1)\\\vdots\\\gamma(m+n-1)\end{bmatrix}}_{=\boldsymbol{\gamma}_n^{(m)}}=\frac{\sigma_w^2}{1-\varphi^2}\begin{bmatrix}\varphi^m\\\varphi^{m+1}\\\vdots\\\varphi^{m+n-1}\end{bmatrix}=\varphi^{m-1}\underbrace{\frac{\sigma_w^2}{1-\varphi^2}\begin{bmatrix}\varphi^1\\\varphi^2\\\vdots\\\varphi^n\end{bmatrix}}_{=\boldsymbol{\gamma}_n}\Leftrightarrow$$

$$\Leftrightarrow\boldsymbol{\varphi}_n^{(m)}=\boldsymbol{\Gamma}_n^{-1}\boldsymbol{\gamma}_n^{(m)}=\varphi^{m-1}\left(\boldsymbol{\Gamma}_n^{-1}\boldsymbol{\gamma}_n\right)=\varphi^{m-1}\boldsymbol{\varphi}_n$$

where $\boldsymbol{\varphi}_n=\left[\varphi_{n1}\ \varphi_{n2}\cdots\varphi_{nn}\right]'$ are the coefficients of the 1-step-ahead BLP. But for the AR(1) model, we know that $\varphi_{n1}=\varphi$ and $\varphi_{nj}=0,\ \forall j=2,\ldots,n\Rightarrow\varphi_{n1}^{(m)}=\varphi^{m-1}\varphi_{n1}=\varphi^m$ and $\varphi_{nj}^{(m)}=0,\ \forall j=2,\ldots,n$. The m-step-ahead BLP becomes $X_{n+m}^n=\sum_{j=1}^{n}\varphi_{nj}^{(m)}X_{n+1-j}=\varphi^m X_n$, with MSE

$$\mathbb{E}\left[\left(X_{n+m}-X_{n+m}^n\right)^2\right]=\mathbb{E}\left[\left(X_{n+m}-\varphi^m X_n\right)^2\right]=$$
$$=\mathbb{E}[X_{n+m}^2]+\varphi^{2m}\mathbb{E}[X_n^2]-2\varphi^m\mathbb{E}[X_{n+m}X_n]=$$
$$=\gamma(0)+\varphi^{2m}\gamma(0)-2\varphi^m\gamma(m)=$$
$$=\frac{\sigma_w^2}{1-\varphi^2}\left[1+\varphi^{2m}-2\varphi^m\varphi^m\right]=\sigma_w^2\frac{1-\varphi^{2m}}{1-\varphi^2}$$

6. Exercise 3.16 from the textbook.

First, the invertible weights of the ARMA(1,1) model $X_t=.9X_{t-1}+W_t+.5W_{t-1}$ are given by $\pi_j=-1.4\left(-.5\right)^{j-1},\ \forall j\geq 1$ (see Example .3.7 on p. 95). Thus, the truncated predictions from equation (3.91) are: $\tilde{x}_{n+m}^n=-\sum_{j=1}^{m-1}\pi_j\tilde{x}_{n+m-j}^n-\sum_{j=m}^{1+m-1}\pi_j x_{n+m-j}^n\Rightarrow$

$$\Rightarrow\tilde{x}_{n+m}^n=1.4\sum_{j=1}^{m-1}(-.5)^{j-1}\tilde{x}_{n+m-j}^n+1.4\sum_{j=m}^{1+m-1}\left(-.5\right)^{j-1}x_{n+m-j}^n$$

Using equation (3.92), we have: $\tilde{x}_{n+m}^n=.9\tilde{x}_{n+m-1}^n+.5\tilde{w}_{n+m-1}^n,\ \forall m\geq 1$, where

$$\begin{cases} \tilde{x}_t^n = 0, \forall t \le 0, \ \& \ \tilde{x}_t^n = x_t, \forall 1 \le t \le n \\ \tilde{w}_t^n = 0, \forall \left[t \le 0, t > n \right] \\ \tilde{w}_t^n = \tilde{x}_t^n - .9\tilde{x}_{t-1}^n + .5\tilde{w}_{t-1}^n \end{cases} \Rightarrow$$

$$\tilde{w}_1^n = \tilde{x}_1^n - .9\tilde{x}_0^n - .5\tilde{w}_0^n = x_1$$

$$\tilde{w}_2^n = \tilde{x}_2^n - .9\tilde{x}_1^n - .5\tilde{w}_1^n = x_2 - 1.4x_1$$

$$\tilde{w}_3^n = \tilde{x}_3^n - .9\tilde{x}_2^n - .5\tilde{w}_2^n = x_3 - .9x_2 - .5\left(x_2 - 1.4x_1\right) = x_3 - 1.4\left(x_2 + (-.5)x_1\right)$$

$$\vdots$$

$$\tilde{w}_n^n = \tilde{x}_n^n - .9\tilde{x}_{n-1}^n - .5\tilde{w}_{n-1}^n = x_n - 1.4\left(x_{n-1} + (-.5)x_{n-2} + \ldots + (-.5)^{n-2}x_1\right) = x_n - 1.4\sum_{j=1}^{n-1}(-.5)^{j-1}x_{n-j}$$

Using the last result in the prediction equation, we get:

$$\tilde{x}_{n+1}^n = .9\tilde{x}_n^n + .5\tilde{w}_n^n = .9x_n + .5\left(x_n - 1.4\sum_{j=1}^{n-1}(-.5)^{j-1}x_{n-j}\right) =$$

$$= 1.4x_n + 1.4\sum_{j=1}^{n-1}(-.5)^{j}x_{n-j} = 1.4\sum_{j=1}^{n}(-.5)^{j-1}x_{n+1-j}$$

$$\tilde{x}_{n+2}^n = .9\tilde{x}_{n+1}^n + .5\,\underbrace{\cancel{\tilde{w}_{n+1}^n}}_{=0} = \left(1.4 - .5\right)\tilde{x}_{n+1}^n = 1.4\tilde{x}_{n+1}^n + (-.5)\left(1.4\sum_{j=1}^{n}(-.5)^{j-1}x_{n+1-j}\right) =$$

$$= 1.4\tilde{x}_{n+1}^n + 1.4\sum_{j=1}^{n}(-.5)^{j}x_{n+1-j} = 1.4\tilde{x}_{n+1}^n + 1.4\sum_{j=2}^{n+1}(-.5)^{j-1}x_{n+2-j}$$

$$\tilde{x}_{n+3}^n = .9\tilde{x}_{n+2}^n + .5\,\underbrace{\cancel{\tilde{w}_{n+2}^n}}_{=0} = \left(1.4 - .5\right)\tilde{x}_{n+2}^n = 1.4\tilde{x}_{n+2}^n + (-.5)\left(1.4\tilde{x}_{n+1}^n + 1.4\sum_{j=2}^{n+1}(-.5)^{j-1}x_{n+2-j}\right) =$$

$$= 1.4\tilde{x}_{n+2}^n + 1.4(-.5)\tilde{x}_{n+1}^n + 1.4\sum_{j=2}^{n+1}(-.5)^{j}x_{n+2-j} = 1.4\left(\tilde{x}_{n+2}^n + (-.5)\tilde{x}_{n+1}^n\right) + 1.4\sum_{j=3}^{n+2}(-.5)^{j-1}x_{n+3-j}$$

$$\vdots$$

$$\tilde{x}_{n+m}^n = .9\tilde{x}_{n+m-1}^n + .5\,\underbrace{\cancel{\tilde{w}_{n+m-1}^n}}_{=0} = \left(1.4 - .5\right)\tilde{x}_{n+m-1}^n =$$

$$= 1.4\tilde{x}_{n+m-1}^n + (-.5)\left(1.4\sum_{j=1}^{m-2}(-.5)^{j-1}\tilde{x}_{n+m-1-j}^n + 1.4\sum_{j=m-1}^{n+m-2}(-.5)^{j-1}x_{n+m-1-j}\right) =$$

$$= 1.4\tilde{x}_{n+m-1}^n + 1.4\sum_{j=1}^{m-2}(-.5)^{j}\tilde{x}_{n+m-1-j}^n + 1.4\sum_{j=m-1}^{n+m-2}(-.5)^{j}x_{n+m-1-j} =$$

$$= 1.4\sum_{j=1}^{m-1}(-.5)^{j-1}\tilde{x}_{n+m-j}^n + 1.4\sum_{j=m}^{n+m-1}(-.5)^{j-1}x_{n+m-j}$$

which is exactly the same as the truncated prediction formula from (3.91).

7. Exercise 3.17 from the textbook.

We have $X_{n+m} - \tilde{X}_{n+m} = \sum_{j=0}^{m-1}\psi_j W_{n+m-j}$, so that:

$$E\left[\left(X_{n+m} - \tilde{X}_{n+m}\right)\left(X_{n+m+k} - \tilde{X}_{n+m+k}\right)\right] =$$

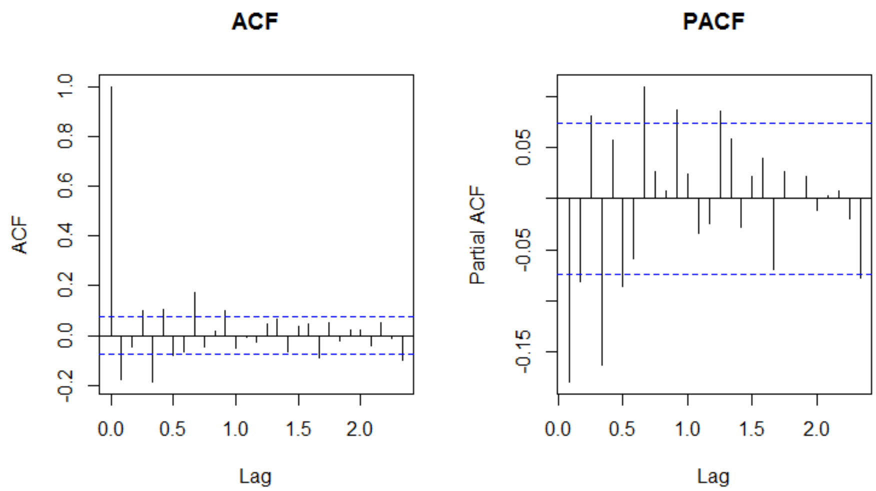
$$= E\left[\left(\sum_{j=0}^{m-1}\psi_j W_{n+m-j}\right)\left(\sum_{i=0}^{m-1}\psi_i W_{n+m+k-i}\right)\right] =$$

$$= \sum_{j=0}^{m-1}\sum_{i=0}^{m-1}\psi_j\psi_i \underbrace{E\left[W_{n+m-j}W_{n+m+k-i}\right]}_{=\sigma_w^2 \text{ only if } j=i-k, \text{ otherwise } =0} = \sigma_w^2\sum_{j=0}^{m-1}\psi_j\psi_{j+k}$$

8. Revisit the data in PS2, Q2: Plot the ACF & PACF for each of the 3 stationary series you produced (i.e. the series *after* any preprocessing). Based on these plots, try to identify an appropriate ARMA(p,q) model.

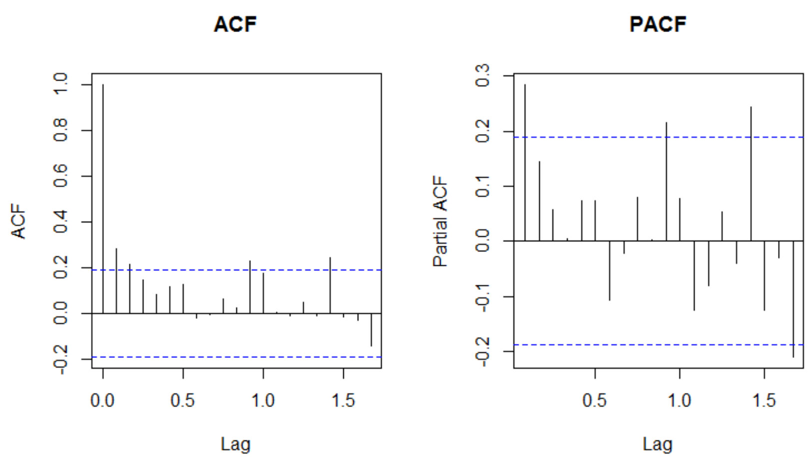
We have the following ACF/PACF plots:

a. Monthly Canadian reserves (in $)



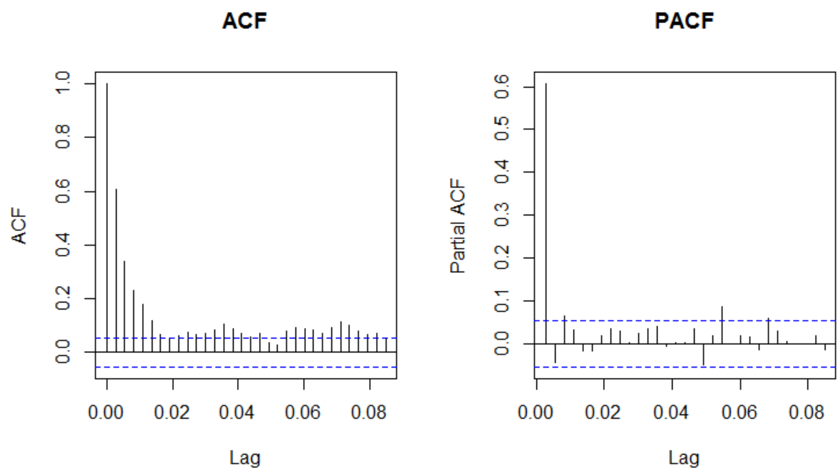
Since neither the ACF or the PACF seem to cut-off, we can go with a general ARMA(p,q) model. The exact order of the model is not obvious, but we should have p≥1 *AND* q≥1.

b. Monthly car sales in Quebec (in # cars)



The PACF looks like it cuts off after lag 1, while the ACF decreases more smoothly → we can go with an AR(1) model.

c. Daily average temperatures in Toronto (in °C)



The situation here is similar to part **b.**, but even clearer. The PACF cuts off after lag 1, and the ACF tails off exponentially → we can go with an AR(1) model.

9. Consider the discrete random variables $X, Y \in \{-1, 0, 1\}$ with joint bivariate probabilities given by the following contingency table:

	$Y=-1$	$Y=0$	$Y=1$
$X=-1$	.05	.10	.15
$X=0$	.20	.10	.10
$X=+1$	.15	0	.15

a. Find the minimum mean square error (MMSE) predictor of *Y* given *X* (i.e. the conditional expectation $g(X) = \mathbb{E}[Y \mid X]$) and the MSE it achieves (i.e. $\mathbb{E}\left[(Y - g(X))^2\right]$).

b. Find the best linear predictor (BLP) of *Y* given *X* (i.e. $g(X) = \alpha_0 + \alpha_1 X$ for the BLP coefficients α_0, α_1) and the MSE it achieves.

(Note: This is an example where the MMSE predictor and the BLP are *different*. The two would be equal only if the random variables were Gaussian, i.e. their joint distribution was Normal.)

a. The following table has the conditional distributions of Y given X=– 1,0,1

	$Y=-1$	$Y=0$	$Y=1$
$X=-1$	$\frac{.05}{.3} = \frac{1}{6}$	$\frac{.10}{.3} = \frac{1}{3}$	$\frac{.15}{.3} = \frac{1}{2}$
$X=0$	$\frac{.2}{.4} = \frac{1}{2}$	$\frac{.1}{.4} = \frac{1}{4}$	$\frac{.1}{.4} = \frac{1}{4}$
$X=+1$	$\frac{.15}{.3} = \frac{1}{2}$	0	$\frac{.15}{.3} = \frac{1}{2}$

$\left(\text{from } P(Y=y \mid X=x) = \frac{P(Y=y, X=x)}{P(X=x)}\right)$

$$\mathbb{E}[Y \mid X=-1]=\sum_{y=-1}^{+1} yP(Y=y \mid X=-1)=-1\times\frac{1}{6}+0\times\frac{1}{3}+1\times\frac{1}{2}=\frac{1}{3}$$

$$\mathbb{E}[Y \mid X=0]=\sum_{y=-1}^{+1} yP(Y=y \mid X=0)=-1\times\frac{1}{2}+0\times\frac{1}{4}+1\times\frac{1}{4}=-\frac{1}{4}$$

$$\mathbb{E}[Y \mid X=+1]=\sum_{y=-1}^{+1} yP(Y=y \mid X=+1)=-1\times\frac{1}{2}+0\times 0+1\times\frac{1}{2}=0$$

The MMSE is $g(X)=\mathbb{E}[Y \mid X]=\begin{cases} 1/3, & X=-1 \\ -1/4, & X=0 \\ 0, & X=+1 \end{cases}$, and the MSE it achieves is:

$$\mathbb{E}\left[\left(Y-g(X)\right)^2\right]=\sum_{y=-1}^{+1}\sum_{x=-1}^{+1}\left(y-g(x)\right)^2 P(Y=y,X=x)=$$
$$=(-1-1/3)^2\times.05+(0-1/3)^2\times.10+(1-1/3)^2\times.15+$$
$$+(-1+1/4)^2\times.20+(0+1/4)^2\times.10+(1+1/4)^2\times.10+$$
$$+(-1-0)^2\times.15+(0-0)^2\times 0+(1-0)^2\times.15=\underline{0.7416667}$$

b. Solving the prediction equations we get $\alpha_1=\dfrac{Cov(Y,X)}{\mathbb{V}(X)}$, $\alpha_0=\mathbb{E}[Y]-\alpha_1\mathbb{E}[X]$, which are exactly the same coefficient estimators as in the simple linear regression of Y on X.

$$\mathbb{E}[Y]=\sum_{y=-1}^{+1} yP(Y=y)=-1\times.4+0\times.2+1\times.4=0$$

$$\mathbb{E}[X]=\sum_{x=-1}^{+1} xP(X=x)=-1\times.3+0\times.4+1\times.3=0$$

$$\mathbb{V}[X]=\sum_{x=-1}^{+1}(x-E[X])^2 P(X=x)=(-1-0)^2\times.3+(0-0)^2\times.4+(1-0)^2\times.3=.6$$

$$\mathrm{Cov}[Y,X]=E[X\cdot Y]-\cancel{E[X]\cdot E[Y]}=\sum_{x=-1}^{+1} y\cdot x\cdot P(Y=y,X=x)=$$
$$=(-1)(-1)\times.05+(0)(-1)\times.10+(1)(-1)\times.15+$$
$$+(-1)(0)\times.20+(0)(0)\times.10+(1)(0)\times.10+$$
$$+(-1)(1)\times.15+(0)(1)\times 0+(1)(1)\times.15=\underline{-0.1}$$

$\Rightarrow \alpha_1=\dfrac{-.1}{.6}=-\dfrac{1}{6},\ \alpha_0=0$, so the BLP is $g(X)=-\dfrac{1}{6}X=\begin{cases} 1/6, & X=-1 \\ 0, & X=0 \\ -1/6, & X=+1 \end{cases}$ and its MSE is

$$\mathbb{E}\left[\left(Y-g(X)\right)^2\right]=\sum_{y=-1}^{+1}\sum_{x=-1}^{+1}\left(y-g(x)\right)^2 P(Y=y,X=x)=$$
$$=(-1-1/6)^2\times.05+(0-1/6)^2\times.10+(1-1/6)^2\times.15+$$
$$+(-1-0)^2\times.20+(0-0)^2\times.10+(1-0)^2\times.10+$$
$$+(-1+1/6)^2\times.15+(0+1/6)^2\times 0+(1+1/6)^2\times.15=\underline{0.7833333}$$

STAD57: Time Series Analysis
Problem Set 5

1. Exercise 3.18 from the textbook

a. $\varphi_1^{ols} = 0.4285906$ & $\varphi_2^{ols} = 0.4417874$

$\varphi_1^{yw} = 0.4339481$ & $\varphi_2^{yw} = 0.4375768$

b. $s.e.(\varphi_1^{ols}) = 0.03979433$ & $\text{s.e.}(\varphi_2^{ols}) = 0.03976163$

$s.e.(\varphi_1^{yw}) = 0.04001303$ & $\text{s.e.}(\varphi_2^{yw}) = 0.04001303$

The standard errors for Yule-Walker estimation are higher than those of regression (ordinary least squares). This is typical since for regression we treat the lagged TS values as fixed (i.e. non-random), which reduces the uncertainty in estimation.

2. Exercise 3.20 from the textbook.

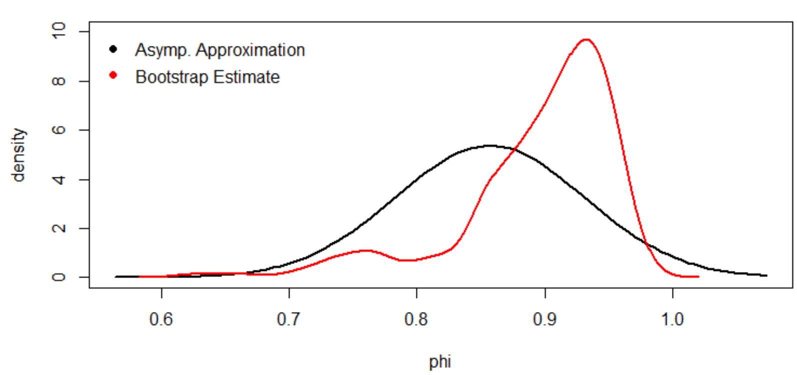
The ACF & PACF do not appear significantly different from 0 at all lags h≥1, which means that the simulated series resembles a White Noise. This is correct since the ARMA(1,1) model is redundant: $X_t = .9X_{t-1} + W_t - .9W_{t-1} \Leftrightarrow \cancel{(1-.9B)}X_t = \cancel{(1-.9B)}W_t \Leftrightarrow X_t = W_t$

3. Exercise 3.21 from the textbook.

	ϕ	θ	σ^2
True	.9	.5	1
Sim 1 Estimates	0.8935255	0.6206927	1.04339
Sim 2 Estimates	0.9305838	0.5301091	1.041926
Sim 3 Estimates	0.874694	0.513024	0.9972095

4. Exercise 3.22 from the textbook.

The asymptotic approximate distribution of $\hat{\varphi}$ and its Bootstrap estimate are shown below:



The asymptotic approximation is not so good when the number of data n is small, and the value of the parameter φ is close to the boundary of the stationary region (in this case, close to 1).

5. Exercise 3.24 from the textbook.

a. We have $\mathbb{E}[X_t]=\mathbb{E}[\alpha+\varphi X_{t-1}+W_t+\theta W_{t-1}]\Rightarrow \mu=\alpha+\varphi\mathbb{E}[X_{t-1}]+\mathbb{E}[W_t]+\theta\mathbb{E}[W_{t-1}]\Rightarrow$

$\Rightarrow \mu=\alpha+\varphi\mu\Rightarrow \mu=\dfrac{\alpha}{1-\varphi}$. We know that for the ARMA(1,1) model the autocovariance function is $\gamma(0)=\sigma_W^2\dfrac{1+2\varphi\theta+\theta^2}{1-\varphi^2},\quad \gamma(1)=\sigma_W^2\dfrac{(1+\varphi\theta)(\varphi+\theta)}{1-\varphi^2},\quad \gamma(h)=\varphi^{h-1}\gamma(1)\quad \forall h\geq 2$. The ACF is thus $\rho(0)=1,\quad \rho(1)=\dfrac{(1+\varphi\theta)(\varphi+\theta)}{1+2\varphi\theta+\theta^2},\quad \rho(h)=\varphi^{h-1}\rho(1)\quad \forall h\geq 2$. The series is weakly strationary, but not necessarily strictly stationary. It would be strictly stationary if the series were also Gaussian. (Note: the time series' mean does not affect its autocovariance or ACF)

b. From theorem A.5 in the Appendix (basically, a version of the Central Limit Theorem for stationary sequences of variables) we have that $\bar{X}_n \to AN\left(\mu,\frac{1}{n}V\right)$, where $\mu=\dfrac{\alpha}{1-\varphi}$, and

$$V=\sum_{h=-\infty}^{\infty}\gamma(h)=\gamma(0)+2\sum_{h=1}^{\infty}\gamma(h)=\gamma(0)+2\gamma(1)\sum_{h=1}^{\infty}\varphi^{h-1}\Rightarrow$$

$$\Rightarrow V=\gamma(0)+2\gamma(1)\sum_{h=0}^{\infty}\varphi^{h}=\gamma(0)+\frac{2\gamma(1)}{1-\varphi}=\sigma_W^2\,\frac{1+2\varphi\theta+\theta^2+2\dfrac{(1+\varphi\theta)(\varphi+\theta)}{1-\varphi}}{1-\varphi^2}$$

6. Consider a zero-mean, stationary *Gaussian* ARMA(p,q) model with autocovariance function $\gamma(h),\ \forall h\geq 0$.

a. Write down the likelihood $L=f(x_1,x_2)$ of the first two observations from the model using the fact that $\begin{bmatrix}X_1\\X_2\end{bmatrix}\sim N_2\left(\begin{bmatrix}0\\0\end{bmatrix},\begin{bmatrix}\gamma(0)&\gamma(1)\\\gamma(1)&\gamma(0)\end{bmatrix}\right)$.

Note: if $\mathbf{X}\sim N_d(\boldsymbol{\mu},\boldsymbol{\Sigma})$, the d-dimensional multivariate Normal density is given by

$$f(\mathbf{x})=(2\pi)^{-d/2}|\boldsymbol{\Sigma}|^{-1/2}\exp\left\{-\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu})'\boldsymbol{\Sigma}^{-1}(\mathbf{x}-\boldsymbol{\mu})\right\}$$

b. Write down the likelihood of the first two observations from the model using the fact that $f(x_1,x_2)=f(x_1)\times f(x_2\,|\,x_1)$, for $X_1\sim N(0,\gamma(0))$ and $X_2\,|\,X_1\sim N(X_2^1,P_2^1)$, where X_2^1 is the 1-step-ahead BLP of X_2 given X_1 & P_2^1 is its corresponding MSE. Show that this is the same as the expression in **a.**

(This problem justifies the likelihood breakdown using 1-step-ahead forecasts that we used in class for deriving the ML estimation)

a.

$$L = f(x_1, x_2) = (2\pi)^{-2/2} \left|\mathbf{\Sigma}\right|^{-1/2} \exp\left\{-\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu})'\mathbf{\Sigma}^{-1}(\mathbf{x}-\boldsymbol{\mu})\right\} =$$

$$= \frac{1}{2\pi}\sqrt{\begin{vmatrix}\gamma(0) & \gamma(1)\\ \gamma(1) & \gamma(0)\end{vmatrix}} \exp\left\{-\frac{1}{2}\begin{bmatrix}x_1 & x_2\end{bmatrix}\begin{bmatrix}\gamma(0) & \gamma(1)\\ \gamma(1) & \gamma(0)\end{bmatrix}^{-1}\begin{bmatrix}x_1\\ x_2\end{bmatrix}\right\} =$$

$$= \frac{1}{2\pi\sqrt{\gamma^2(0)-\gamma^2(1)}} \exp\left\{-\frac{1}{2\left(\gamma^2(0)-\gamma^2(1)\right)}\begin{bmatrix}x_1\gamma(0)-x_2\gamma(1) & -x_1\gamma(1)+x_2\gamma(0)\end{bmatrix}\begin{bmatrix}x_1\\ x_2\end{bmatrix}\right\} =$$

$$= \frac{1}{2\pi\sqrt{\gamma^2(0)-\gamma^2(1)}} \exp\left\{-\frac{x_1^2\gamma(0)-2x_1x_2\gamma(1)+x_2^2\gamma(0)}{2\left(\gamma^2(0)-\gamma^2(1)\right)}\right\}$$

b. The 1-step-ahead BLP is given by $X_2^1 = \varphi_{1,1}X_1 = \dfrac{\gamma(1)}{\gamma(0)}X_1$ (by solving the prediction eqn $\gamma(1) = \varphi_{1,1}\gamma(0)$), with MSE given by $P_2^1 = P_1^0 - \varphi_{1,1}^2\gamma(0) = \gamma(0) - \dfrac{\gamma^2(1)}{\gamma^2(0)}\gamma(0) = \gamma(0) - \dfrac{\gamma^2(1)}{\gamma(0)}$. The likelihood is:

$$L = f(x_1, x_2) = f(x_1)\times f(x_2 \mid x_1) =$$

$$= \left(\frac{1}{\sqrt{2\pi\gamma(0)}}\exp\left\{-\frac{x_1^2}{2}\right\}\right)\times\left(\frac{1}{\sqrt{2\pi\left(\gamma(0)-\gamma^2(1)/\gamma(0)\right)}}\exp\left\{-\frac{\left(x_2-x_1\gamma(1)/\gamma(0)\right)^2}{2\left(\gamma(0)-\gamma^2(1)/\gamma(0)\right)}\right\}\right) =$$

$$= \frac{1}{2\pi\sqrt{\gamma(0)\left(\gamma(0)-\gamma^2(1)/\gamma(0)\right)}} \exp\left\{-\frac{x_1^2}{2\gamma(0)} - \frac{x_2^2-2x_1x_2\,\gamma(1)/\gamma(0)+x_1^2\,\gamma^2(1)/\gamma^2(0)}{2\left(\gamma(0)-\gamma^2(1)/\gamma(0)\right)}\right\} =$$

$$= \frac{1}{2\pi\sqrt{\gamma^2(0)-\gamma^2(1)}} \exp\left\{-\frac{x_1^2\left(1-\cancel{\gamma^2(1)/\gamma^2(0)}\right)+x_2^2-2x_1x_2\,\gamma(1)/\gamma(0)+\cancel{x_1^2\,\gamma(1)/\gamma(0)}}{2\left(\gamma(0)-\gamma^2(1)/\gamma(0)\right)}\right\} =$$

$$= \frac{1}{2\pi\sqrt{\gamma^2(0)-\gamma^2(1)}} \exp\left\{-\frac{x_1^2-2x_1x_2\,\gamma(1)/\gamma(0)+x_2^2}{2\left(\gamma(0)-\gamma^2(1)/\gamma(0)\right)}\right\} =$$

$$= \frac{1}{2\pi\sqrt{\gamma^2(0)-\gamma^2(1)}} \exp\left\{-\frac{x_1^2\gamma(0)-2x_1x_2\gamma(1)+x_2^2\gamma(0)}{2\left(\gamma^2(0)-\gamma^2(1)\right)}\right\}$$

which is exactly the same as in part **a.**

STAD57: Time Series Analysis
Problem Set 7 Solutions

1. Exercise 3.27 from the textbook.

First, note that if X_t is stationary with covariance function $\gamma_X(h)$ then $\nabla^k X_t$ is stationary $\forall k \geq 1$. We only need to show this for ∇X_t, since the general result follows by induction, because $\nabla^k X_t = \nabla\left(\nabla^{k-1} X_t\right) = \nabla\left(\nabla \cdots \left(\nabla X_t\right)\right)$. We have:

$$\gamma_{\nabla X}(h) = Cov[\nabla X_{t+h}, \nabla X_t] = Cov[X_{t+h} - X_{t+h-1}, X_t - X_{t-1}] =$$
$$= Cov[X_{t+h}, X_t] - Cov[X_{t+h}, X_{t-1}] - Cov[X_{t+h-1}, X_t] + Cov[X_{t+h-1}, X_{t-1}] =$$
$$= \gamma_X(h) - \gamma_X(h+1) - \gamma_X(h-1) + \gamma_X(h) = 2\gamma_X(h) - \gamma_X(h+1) - \gamma_X(h-1)$$

which is independent of $t \Rightarrow \nabla X_t$ is stationary.

Next, we show that for any polynomial of order q, $P(t) = \beta_0 + \beta_1 t + \cdots + \beta_q t^q$, if you take k^th order differences you end up with a polynomial of order q−k, if k<q, or with a constant if k≥q.

Note that $(t-1)^k = t^k - \{\text{polynomial of order } k-1\}$, so that:

$$\nabla P(t) = P(t) - P(t-1) =$$
$$= \left(\beta_0 + \beta_1 t + \cdots + \beta_q t^q\right) - \left(\beta_0 + \beta_1 (t-1) + \cdots + \beta_q (t-1)^q\right) =$$
$$= \cancel{\left(\beta_0 + \beta_1 t + \cdots + \beta_q t^q\right)} - \cancel{\left(\beta_0 + \beta_1 t + \cdots + \beta_q t^q\right)} +$$
$$+ \left(\beta_1 + \beta_2 \times \{\text{polynomial of order } 1\} + \beta_q \times \{\text{polynomial of order } q-1\}\right)$$
$$= \{\text{polynomial of order } q-1\}$$

By induction, we get $\nabla^k P(t) = \nabla\left(\nabla^{k-1} P(t)\right) = \nabla\left(\nabla \cdots \left(\nabla P(t)\right)\right)$ is equal to a polynomial of order q-k, where if k≥q we end up with a constant. Combining this with the first result, we see that for $Y_t = \beta_0 + \beta_1 t + \cdots + \beta_q t^q + X_t,\ \beta_q \neq 0$ to be stationary (i.e. to have constant mean), we need to differentiate it at least q times.

2. Exercise 3.29 from the textbook.

(a) We have $y_t = \nabla x_t = x_t - x_{t-1} = \delta + \varphi y_{t-1} + w_t$, which means that y_t follows a non 0-mean AR(1) process. Let $z_t = \left(y_t - \mu\right) = \varphi\left(y_{t-1} - \mu\right) + w_t$ be the corresponding 0-mean AR(1) process, where the mean is such that $\left(y_t - \mu\right) = \varphi\left(y_{t-1} - \mu\right) + w_t \Leftrightarrow y_t = \delta + \varphi y_{t-1} + w_t \Rightarrow$ $\Rightarrow \mu - \varphi\mu = \delta \Rightarrow \mu = \delta/(1-\varphi)$. We also know that the m-step-ahead BLP estimator of the AR(1) model is $z_{n+m}^n = \varphi^m z_n$ (by recursive application of the 1-step-ahead predictor formula), so

$$z_{n+j}^n = \varphi^j z_n \Rightarrow y_{n+j}^n - \mu = \varphi^j\left(y_n - \mu\right) \Rightarrow y_{n+j}^n = (1-\varphi^j)\mu + \varphi^j y_n =$$
$$= \delta\frac{1-\varphi^j}{1-\varphi} + \varphi^j y_n = \delta\left[1 + \varphi + \cdots + \varphi^{j-1}\right] + \varphi^j y_n \quad \left(\text{since } \sum_{k=0}^{n-1} x^k = \frac{1-x^n}{1-x}\right)$$

(b) Note that $\nabla x_{n+j}^{n} = x_{n+j}^{n} - x_{n+j-1}^{n} = y_{n+j}^{n} \Rightarrow$ (summing both sides over $j = 1, \ldots, m$)

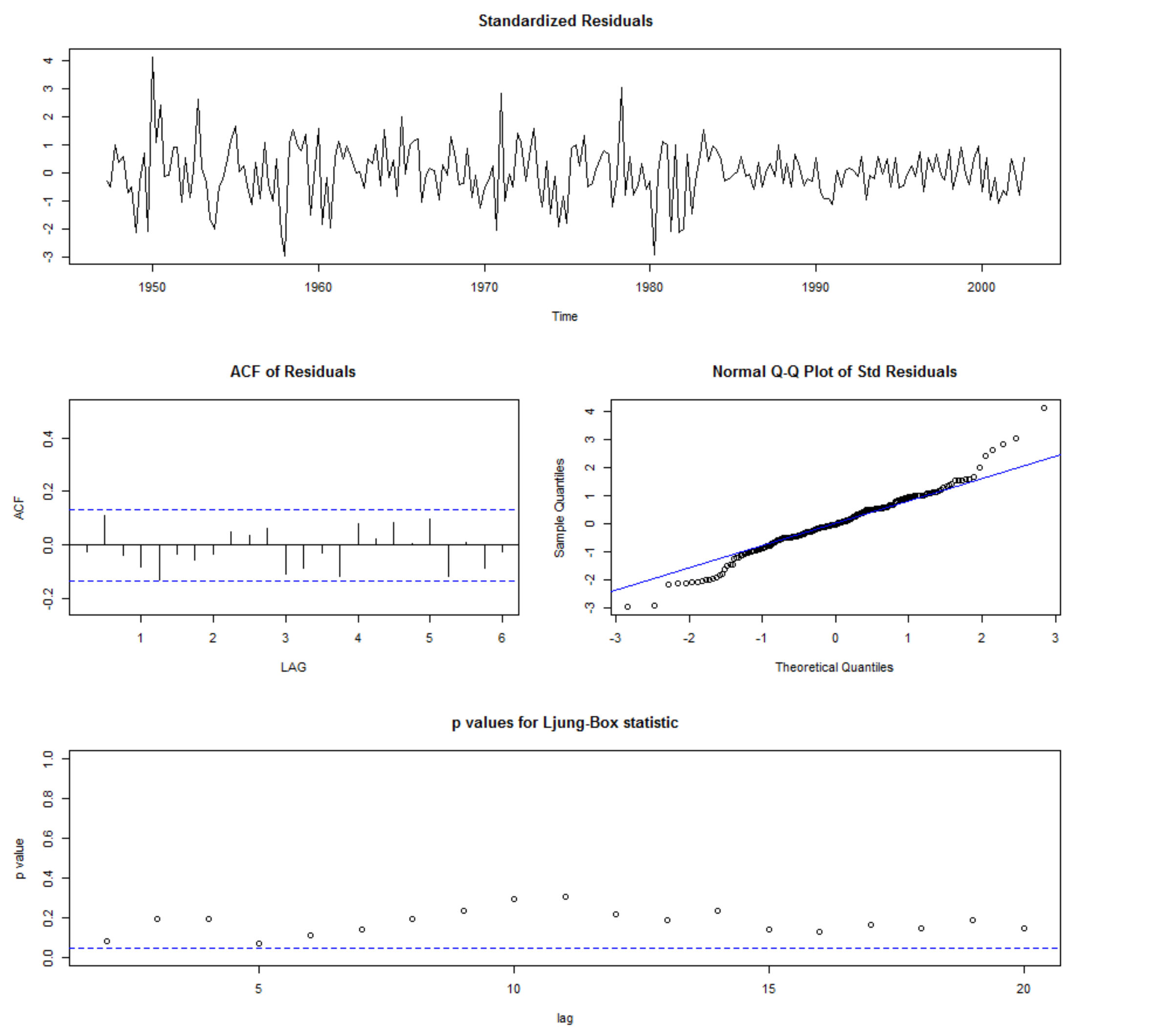
$$\Rightarrow \sum_{j=1}^{m} \left(x_{n+j}^{n} - x_{n+j-1}^{n} \right) = \sum_{j=1}^{m} y_{n+j}^{n} = \sum_{j=1}^{m} \left(\delta \frac{1-\varphi^{j}}{1-\varphi} + \varphi^{j} y_{n} \right) \Rightarrow$$

$$\Rightarrow \left(x_{n+m}^{n} - \cancel{x_{n+m-1}^{n}} \right) - \left(\cancel{x_{n+m-1}^{n}} - \cancel{x_{n+m-2}^{n}} \right) - \cdots - \left(\cancel{x_{n+1}^{n}} - x_{n}^{n} \right) = \frac{\delta}{1-\varphi} \sum_{j=1}^{m} \left(1-\varphi^{j} \right) + y_{n} \sum_{j=1}^{m} \varphi^{j} \Rightarrow$$

$$\Rightarrow x_{n+m}^{n} - x_{n} = \frac{\delta}{1-\varphi} \left[\sum_{j=1}^{m} 1 + \varphi \sum_{j=0}^{m-1} \varphi^{j} \right] + \left(x_{n} - x_{n-1} \right) \varphi \sum_{j=0}^{m-1} \varphi^{j} \Rightarrow \left(\text{since} \begin{cases} x_{n}^{n} = x_{n} \\ y_{n} = \nabla x_{n} = x_{n} - x_{n-1} \end{cases} \right)$$

$$\Rightarrow x_{n+m}^{n} = x_{n} + \frac{\delta}{1-\varphi} \left[m + \frac{\varphi\left(1-\varphi^{m}\right)}{1-\varphi} \right] + \left(x_{n} - x_{n-1} \right) \frac{\varphi\left(1-\varphi^{m}\right)}{1-\varphi}$$

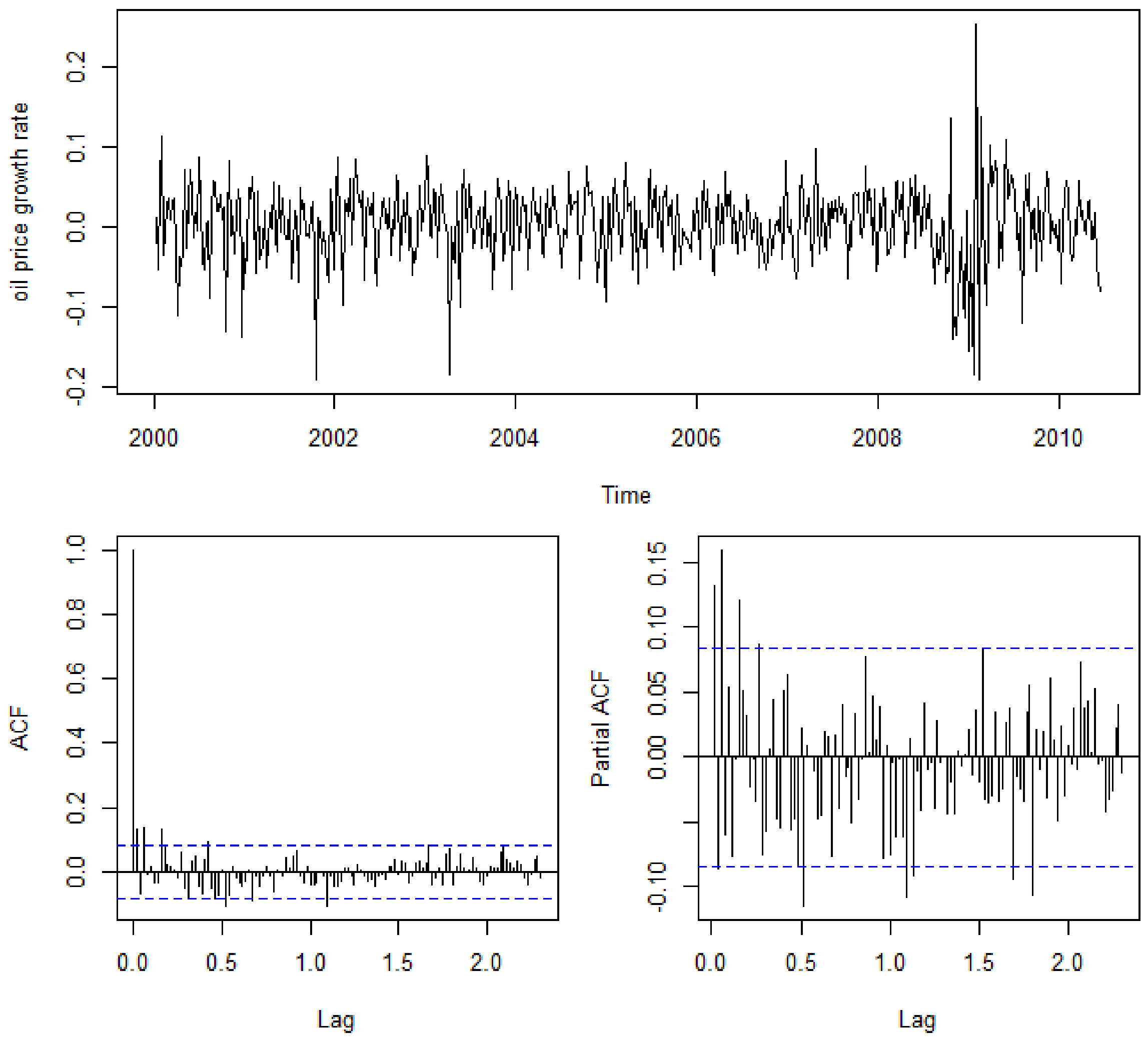
3. Exercise 3.31 from the textbook.



Overall, the model residuals do not show significant deviations from the assumptions. Their mean and variance seem relatively constant. The Q-Q plot shows heavier than Normal tails (which is usual in practice). There seems to be no significant residual auto-correlation in the residuals.

4. Exercise 3.32 from the textbook.

The plot of the growth rate (i.e. the log-difference of oil prices), with ACF & PACF is as follows: (notice the drop from the 2008 from the financial crisis)



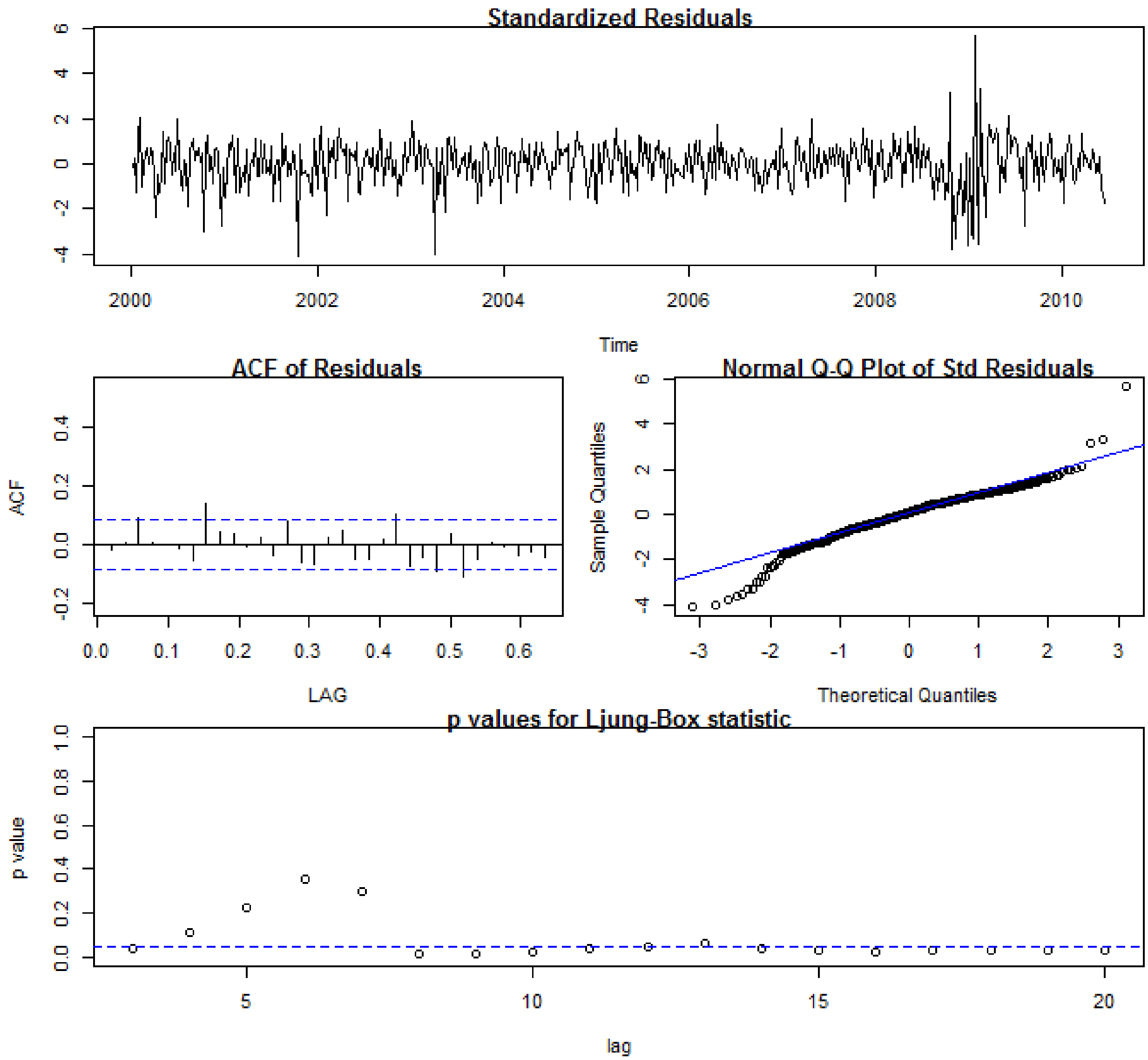
It is not clear from the ACF/PACF what the best model specification is, so we can try different models and compare some criterion, like the AIC. The `auto.arima()` function in the `forecast` package (with option `seasonal=FALSE`) gives an ARIMA(1,0,1) specification for the growth rate, with the following model fit:

```
Coefficients:
          ar1     ma1   xmean
      -0.5264  0.7146  0.0018
s.e.   0.0871  0.0683  0.0022
```

```
sigma^2 estimated as 0.002102:  log likelihood = 904.89,  aic = -1801.79

$AIC
[1] -5.153838
```

Nevertheless, the diagnostics plots show a significant residual autocorrelation at lag 8 (especially the Ljung-Box test).



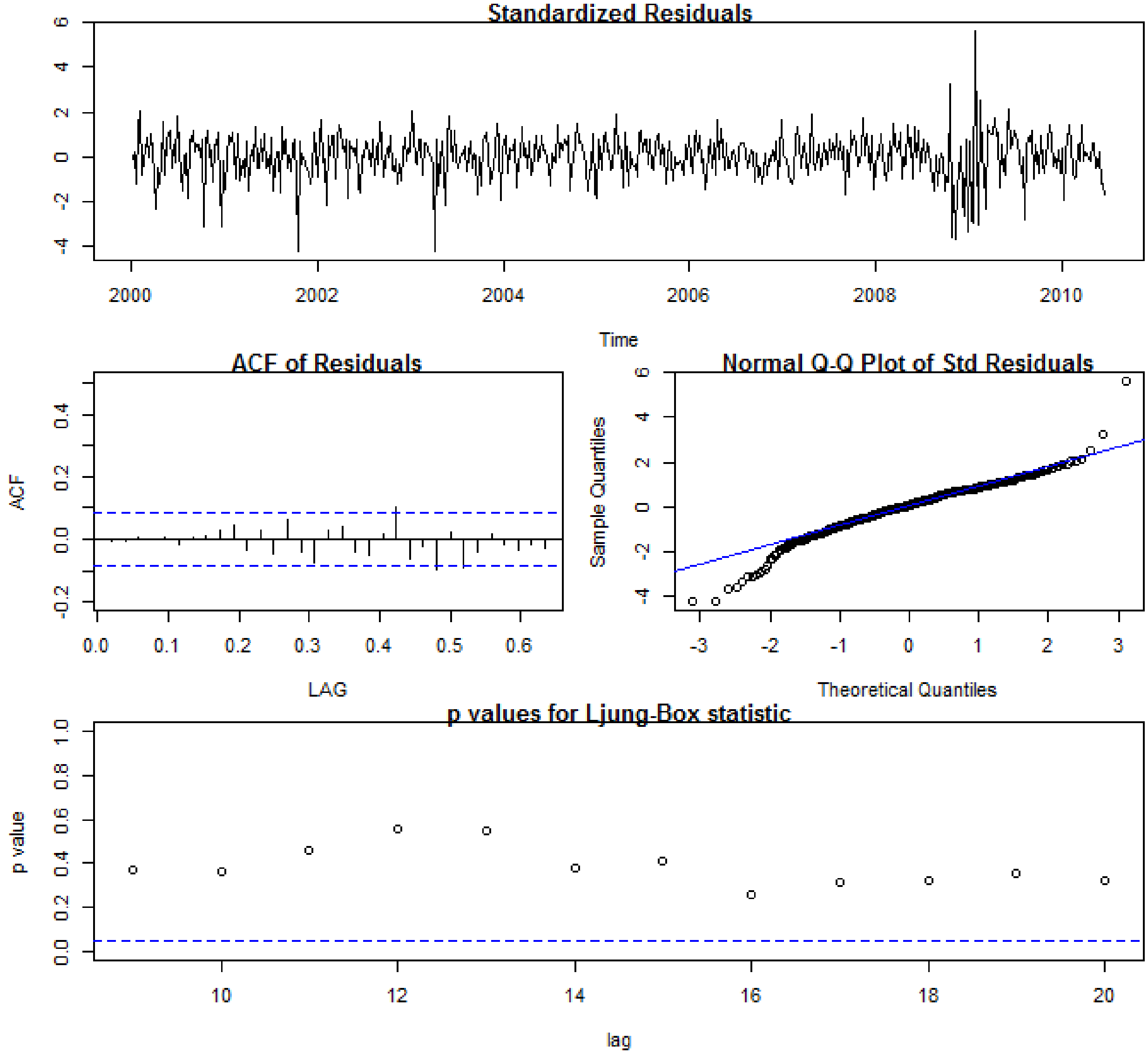
Therefore, we can try a bigger model (higher lags). In particular, Yule-Walker estimation for purely autoregressive models gives an AR(8) model specification with lower AIC:

```
Coefficients:
         ar1      ar2     ar3      ar4     ar5      ar6      ar7     ar8   xmean
      0.1742  -0.1200  0.1814  -0.0689  0.0448  -0.0621  -0.0218  0.1224  0.0017
s.e.  0.0426   0.0433  0.0436   0.0442  0.0443   0.0437   0.0435  0.0428  0.0026

sigma^2 estimated as 0.002038:  log likelihood = 913.19,  aic = -1806.39
```

```
$AIC
[1] -5.162664
```

Moreover, the residuals from this model look better:



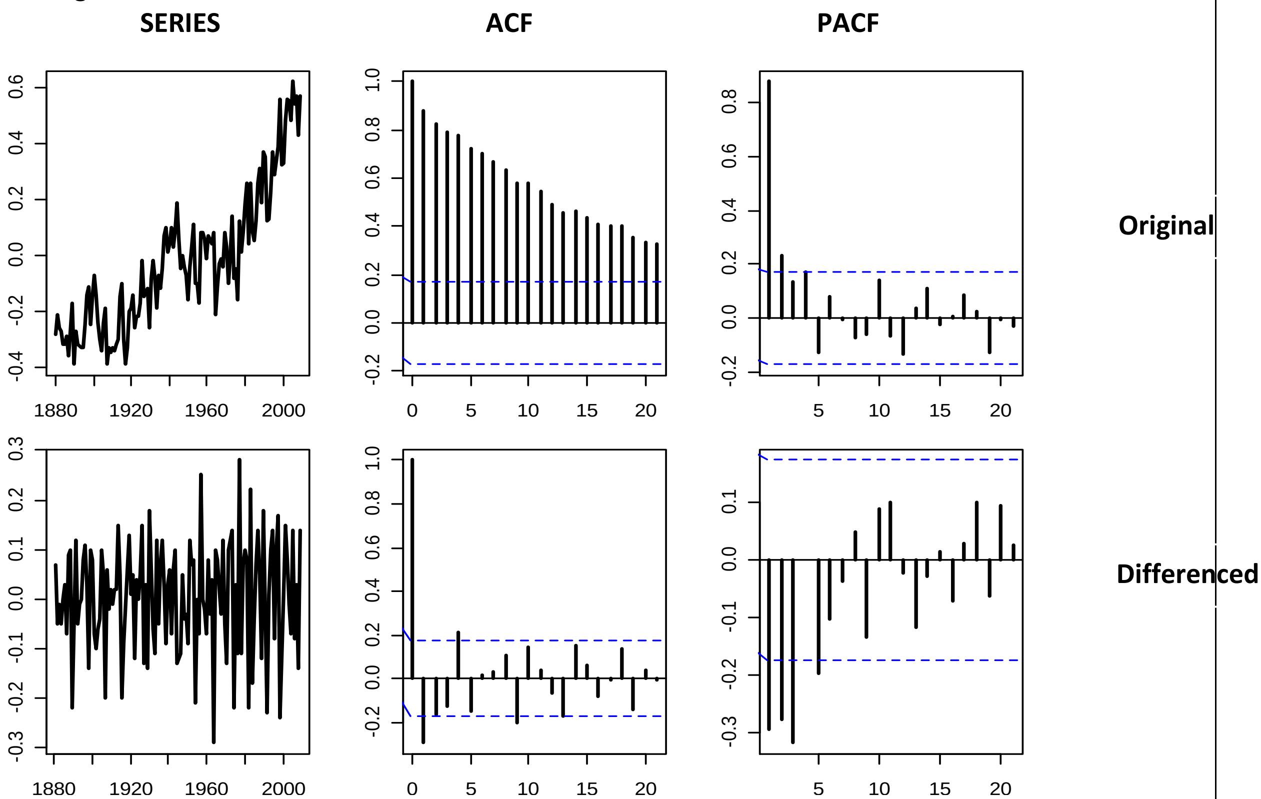
Adding an MA component, i.e. trying an ARMA(8,1) model, does not offer substantial improvements, so we can stick with the AR(8) model for the growth rate, (or, equivalently, an ARIMA(8,1,0) model for the log-price, if we want to use the model to forecast the actual price).

Note: The `auto.arima()` function is not always guaranteed to give you the best model specification for a given criterion (AIC/BIC). This can happen even for specifications that are considered by the functions; for an explanation why, see: http://stats.stackexchange.com/questions/122704/should-auto-arima-in-r-ever-report-a-model-with-higher-aic-aicc-and-bic-than-ot

So, it is important to always look at the diagnostics of the selected model and consider plausible alternatives. Finally, when comparing AIC values across specifications resulting from different R functions (like `auto.arima()` or `ar()`), you should use the same estimation function (i.e. `arima()` or `sarima()`) on the same data to calculate the criterion, because different functions can use different conventions for criterion formula.

5. Exercise 3.33 from the textbook.

The ACF of the original series looks integrated, so we take 1st order differences which give stationary looking ACF & PACF:

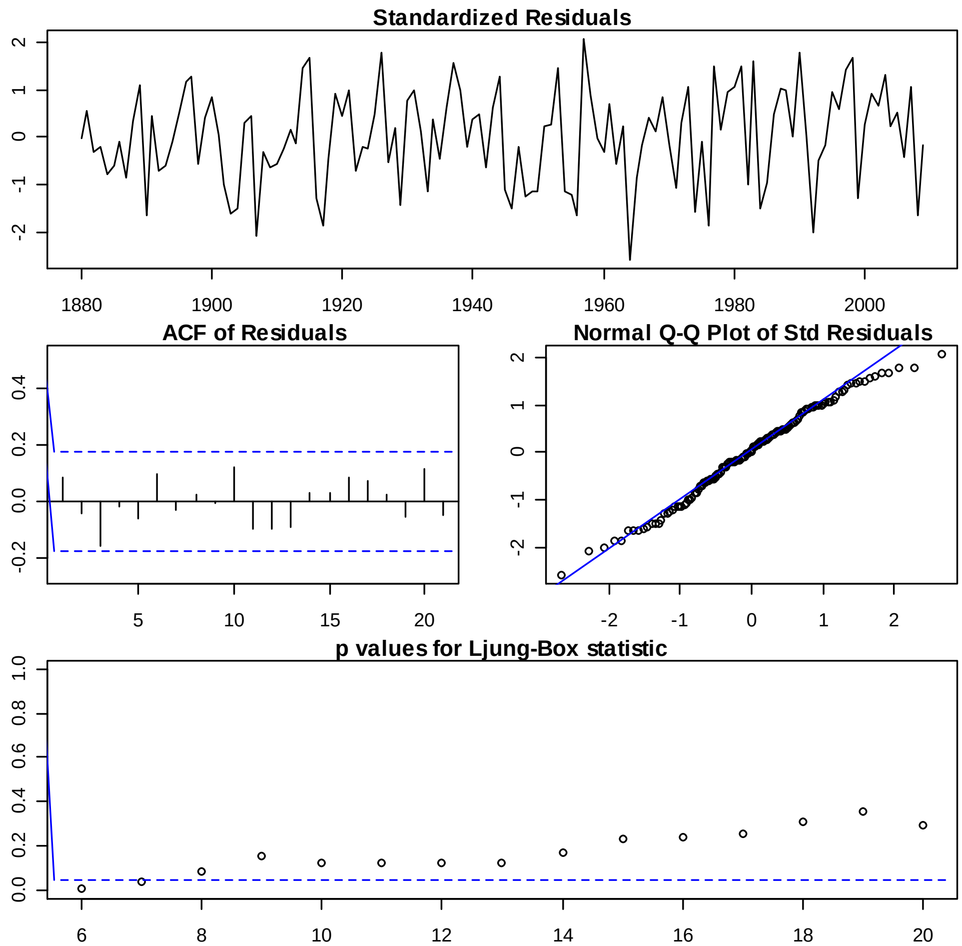


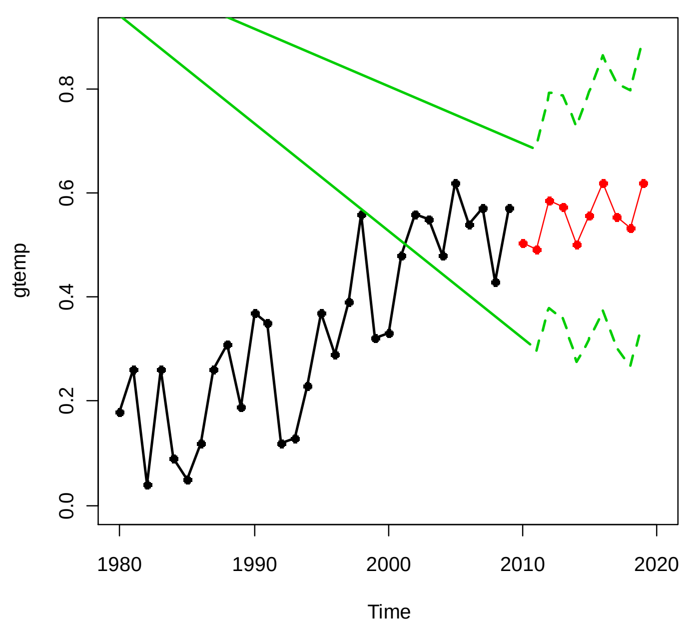
The best model selected using AIC over all possible ARIMA(p,1,q) with p,q≤7 is ARIMA(2,1,3) with parameter estimates:

```
Coefficients:
          ar1      ar2      ma1     ma2      ma3  constant
      -0.4033  -0.9807  -0.1563  0.6927  -0.6378    0.0060
s.e.   0.0212   0.0181   0.0841  0.0680   0.0915    0.0031

sigma^2 estimated as 0.008257:  log likelihood=123.69
AIC=-233.38   AICc=-232.46   BIC=-213.37
```

The diagnostics plots show that the model has a good fit, with 0-mean & constant variance residuals that seem uncorrelated & with somewhat fatter tails than normal. The 1- to 10-step-ahead forecasts are also given below:





6. Exercise 3.38 from the textbook.

(a) This is a SARIMA(0,0,0)×(0,0,1)$_2$ which is equivalent to a purely seasonal SMA(1)$_2$ model, with period 2.

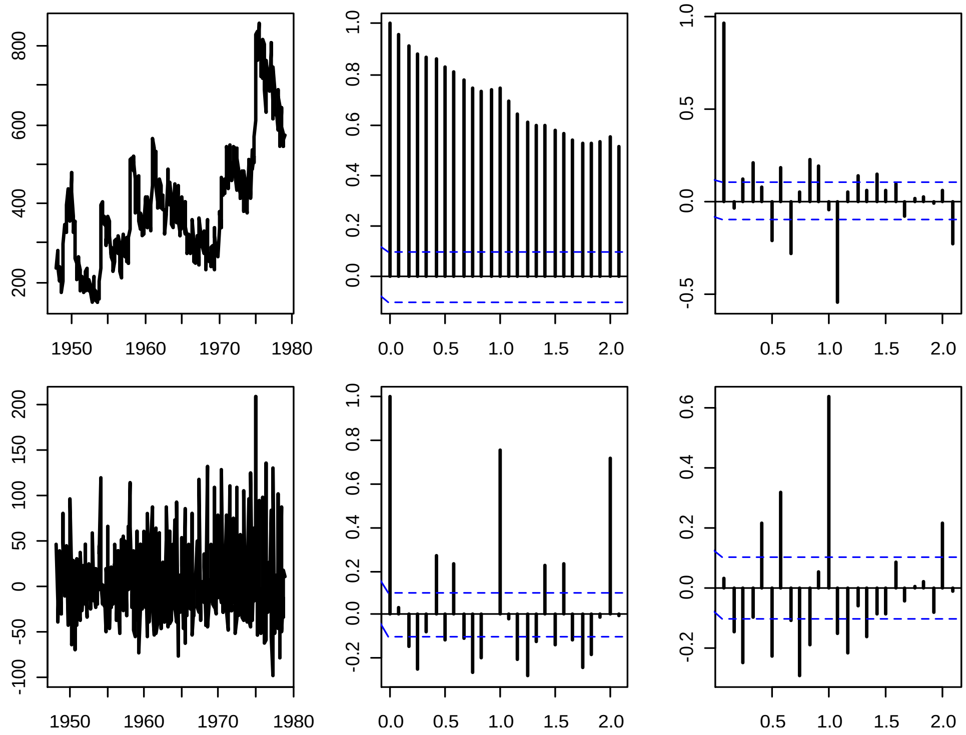
(b) We have $X_t = (1+\Theta B^2)W_t = \Theta(B^2)W_t$ where $\Theta(z) = 1+\Theta z \Rightarrow \Theta^{-1}(z) = \sum_{j=0}^{\infty}(-\Theta z)^j$, so the invertible representation of the series is $W_t = \Theta^{-1}(B^2)X_t = \sum_{j=0}^{\infty}(-\Theta B^2)^j X_t = \sum_{j=0}^{\infty}(-\Theta)^j X_{t-2j}$.

(c) Form equation (3.85) in the textbook we have $\tilde{X}_{n+m} = -\sum_{j=1}^{\infty}\pi_j \tilde{X}_{n+m-j}$ where $\tilde{X}_j = X_j, \forall j \le n$. From the previous part we have $\pi_j = \begin{cases} 0, & \text{odd } j \\ (-\Theta)^{j/2}, & \text{even } j \end{cases}$, so that the m-step-ahead forecast is $\tilde{X}_{n+m} = -\sum_{j=1}^{\infty}(-\Theta)^j \tilde{X}_{n+m-2j}$. We also know from equation (3.86) that $P_{n+m}^n = \sigma_W^2 \sum_{j=0}^{m-1}\psi_j^2$, where the ψ 's are the causal weights. In our case, $X_t = W_t + \Theta W_{t-2}$ so $\psi_0 = 1, \psi_1 = 0, \psi_2 = \Theta, \psi_j = 0 \ \forall j > 2$. Thus, $P_{n+1}^n = P_{n+2}^n = \sigma_W^2$, $P_{n+3}^n = P_{n+4}^n = \cdots = \sigma_W^2(1+\Theta^2)$.

7. Exercise 3.41f from the textbook.

The ACF of the original series looks integrated, so we take 1st order differences which give stationary looking ACF & PACF with what looks like annual seasonality (s=12 months):

SERIES **ACF** **PACF**



Original

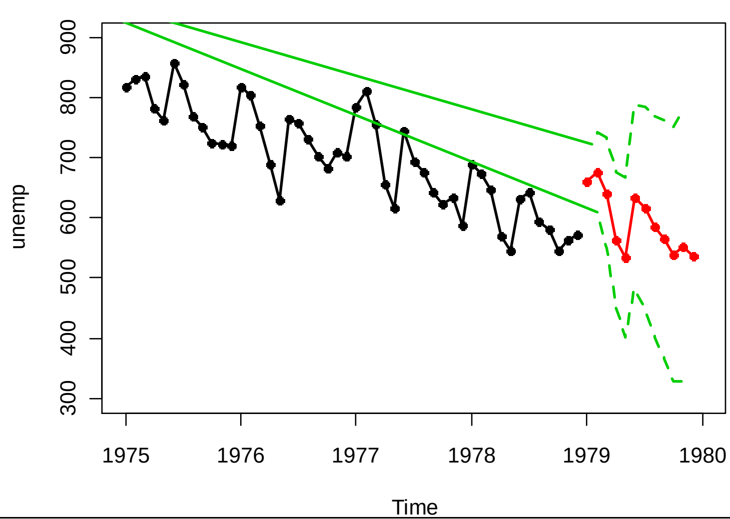
Differenced

The `auto.arima()` function picked the SARIMA(2,1,0)×(2,0,0)$_{12}$ model

```
Coefficients:
         ar1     ar2    sar1    sar2
      0.1348  0.2694  0.4600  0.4570
s.e.  0.0500  0.0508  0.0472  0.0489

sigma^2 estimated as 511.7:  log likelihood=-1694.01
AIC=3398.02   AICc=3398.19   BIC=3417.6
```

With the following forecasts



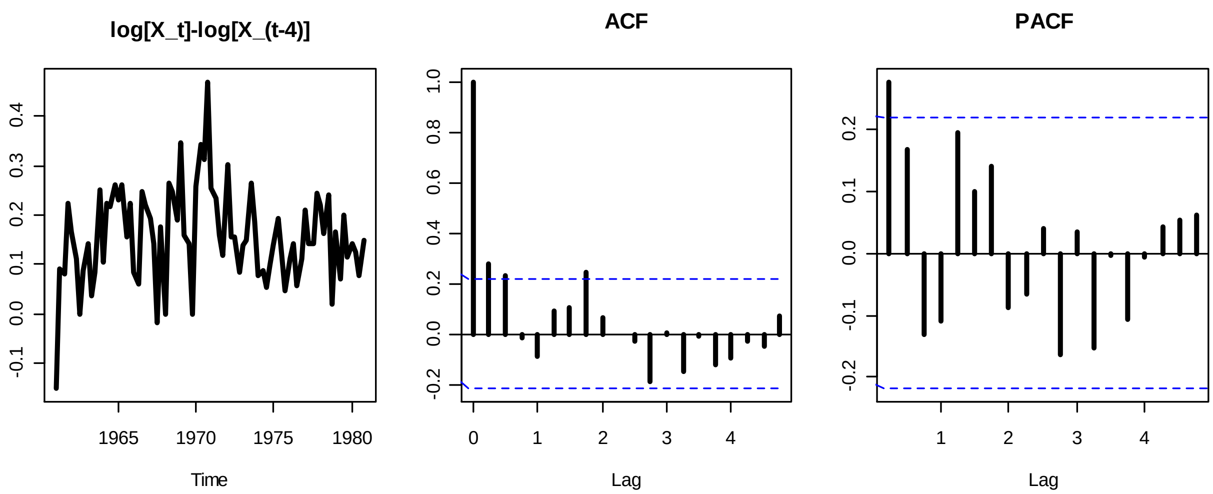
8. Exercise 3.44 from the textbook.

The selected model (minimum AIC) is a SARIMA(2,0,2)×(0,1,0)$_4$

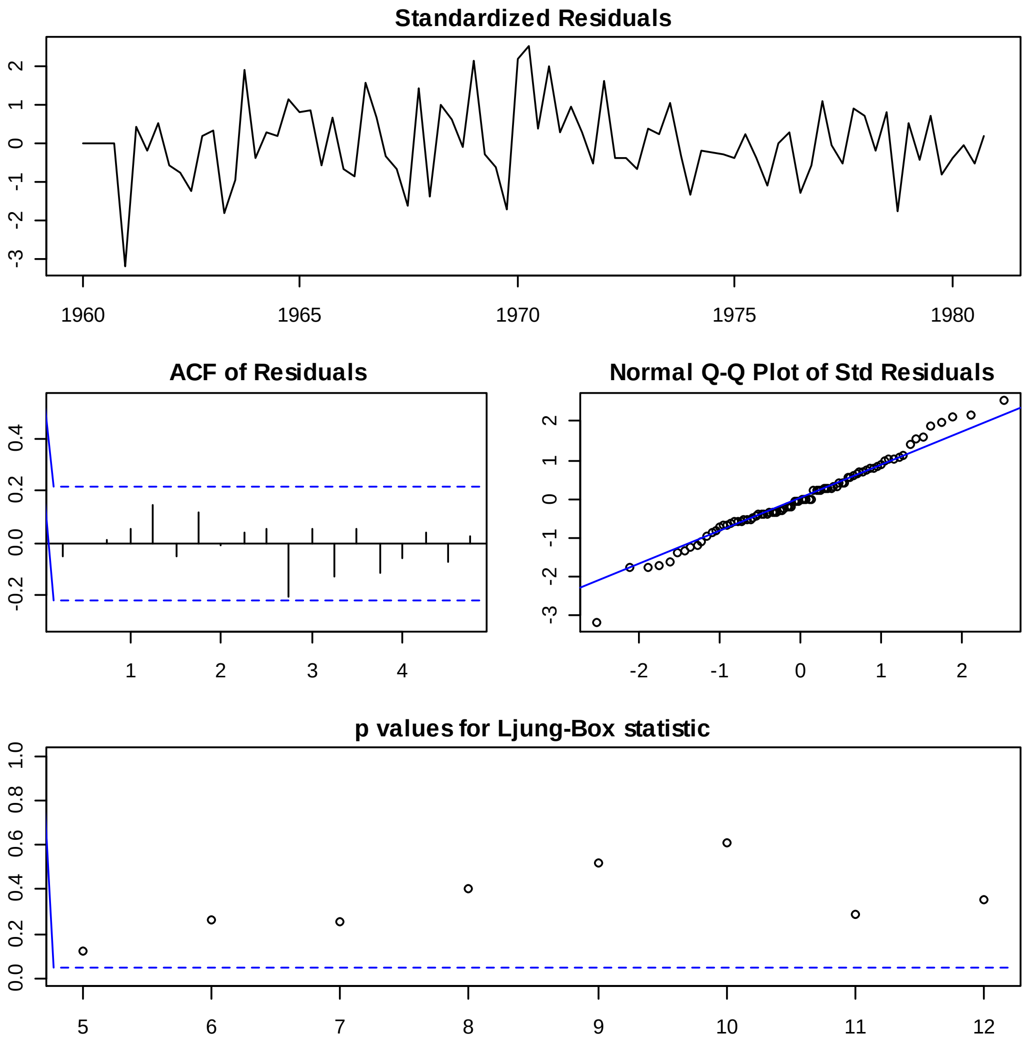
```
Coefficients:
         ar1      ar2      ma1     ma2  constant
      0.6798  -0.6133  -0.4031  0.7998    0.0386
s.e.  0.1650   0.1693   0.1228  0.1270    0.0035

sigma^2 estimated as 0.0071:  log likelihood=83.92
AIC=-155.83   AICc=-154.68   BIC=-141.54
```

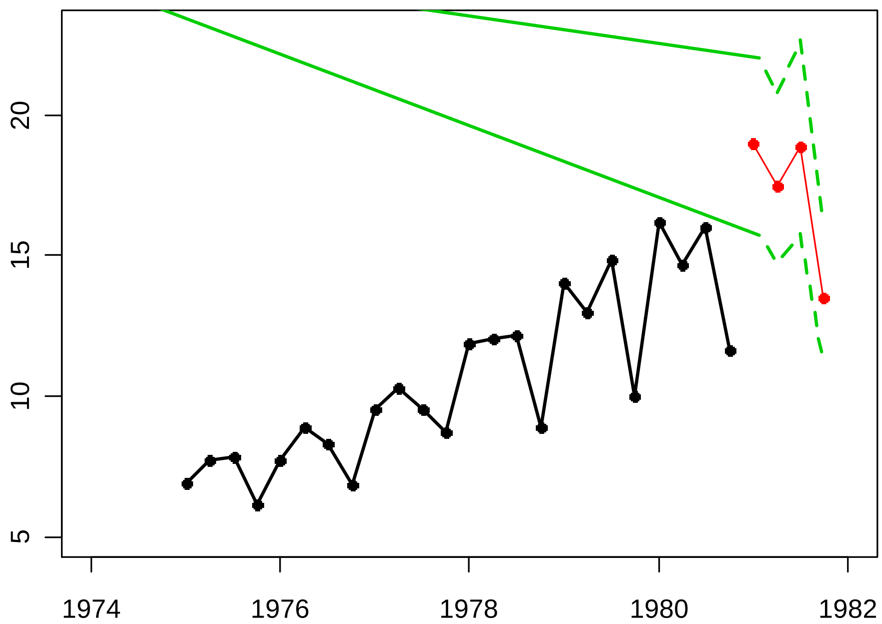
Note that the model has an integrated multiplicative seasonal component with period 4 (i.e. annual period for quarterly data). The seasonally differenced log-series at lag 4 (i.e. the stationary ARMA part of the model) is shown below, together with its ACF & PACF:



The diagnostics plots for the model are shown below; the residuals seem to be uncorrelated with 0-mean & constant variance, without significant departures from Normality.



Finally, the 1- to 4-step-ahead forecasts with 95% prediction intervals are given below:



9. Derive the ACF function of the general seasonal MA(2) model with period s, i.e. the SMA(2)$_s$:

$$X_t = W_t + \Theta_1 W_{t-s} + \Theta_2 W_{t-2s},\ \{W_t\} \sim \mathrm{WN}(0, \sigma_w^2)$$

We have:

$$\gamma(0) = Cov(X_t, X_t) = Cov(W_t + \Theta_1 W_{t-s} + \Theta_2 W_{t-2s}, W_t + \Theta_1 W_{t-s} + \Theta_2 W_{t-2s}) =$$

$$= Var(W_t) + \Theta_1^2 Var(W_{t-s}) + \Theta_2^2 Var(W_{t-2s}) = \sigma_w^2 (1 + \Theta_1^2 + \Theta_2^2)$$

$$\gamma(s) = Cov(X_{t+s}, X_t) = Cov(W_{t+s} + \Theta_1 W_t + \Theta_2 W_{t-s}, W_t + \Theta_1 W_{t-s} + \Theta_2 W_{t-2s}) =$$

$$= \Theta_1 Var(W_t) + \Theta_1 \Theta_2 Var(W_{t-s}) = \sigma_w^2 \Theta_1 (1 + \Theta_2)$$

$$\gamma(2s) = Cov(X_{t+2s}, X_t) = Cov(W_{t+2s} + \Theta_1 W_{t+s} + \Theta_2 W_t, W_t + \Theta_1 W_{t-s} + \Theta_2 W_{t-2s}) =$$

$$= \Theta_2 Var(W_t) = \sigma_w^2 \Theta_2$$

It is easy to see that $\gamma(h) = 0$ for any other $h \neq 0, s, 2s$. Thus, the ACF becomes:

$$\rho(h) = \frac{\gamma(h)}{\gamma(0)} = \begin{cases} \dfrac{1}{1+\Theta_1^2+\Theta_2^2}, & h = 0 \\ \dfrac{\Theta_1(1+\Theta_2)}{1+\Theta_1^2+\Theta_2^2}, & h = s \\ \dfrac{\Theta_2}{1+\Theta_1^2+\Theta_2^2}, & h = 2s \\ 0 & \text{otherwise} \end{cases}$$

10. Repeat problem 3.35 from the textbook for the model $X_t = \varphi X_{t-1} + W_t + \Theta W_{t-2}$

(a) The model is SARIMA(1,0,0)×(0,0,1)$_2$

(b) We have

$$X_t = \varphi X_{t-1} + W_t + \Theta W_{t-2} \Rightarrow (1-\varphi B)X_t = (1+\Theta B^2)W_t \Rightarrow$$

$$\Rightarrow W_t = (1+\Theta B^2)^{-1}(1-\varphi B)X_t = \sum_{j=0}^{\infty} \pi_j X_{t-j}$$

where the inverse polynomial $(1+\Theta B^2)^{-1} = \sum_{j=0}^{\infty} \lambda_j B^j$ is such that:

$$(1+\Theta B^2)^{-1}(1+\Theta B^2) = 1 \Leftrightarrow (\lambda_0 + \lambda_1 B + \lambda_2 B^2 + \cdots)(1+\Theta B^2) = 1 \Rightarrow$$

$$\Rightarrow \begin{cases} \lambda_0 = 1 \\ \lambda_1 B = 0 \Rightarrow \lambda_1 = 0 \\ \lambda_2 B^2 + \lambda_0 \Theta B^2 = 0 \Rightarrow \lambda_2 = -\Theta \\ \lambda_3 B^3 = 0 \Rightarrow \lambda_3 = 0 \\ \lambda_4 B^4 + \lambda_2 \Theta B^4 = 0 \Rightarrow \lambda_4 = \Theta^2 \\ \vdots \end{cases} \Rightarrow \lambda_j = \begin{cases} 0, \text{ for } j \text{ odd} \\ (-\Theta)^{j/2}, \text{ for } j \text{ even} \end{cases}$$

Obviously the coefficients λ_j are absolutely summable if |Θ|<1, so the model is invertible if |Θ|<1. Moreover:

$$W_t = \sum_{j=0}^{\infty} \pi_j X_{t-j} = (1+\Theta B^2)^{-1}(1-\varphi B)X_t = (1-\Theta B^2 + \Theta^2 B^4 - \Theta^3 B^6 + \cdots)(1-\varphi B)X_t$$

$$= (1 - \varphi B - \Theta B^2 + \varphi\Theta B^3 + \Theta^2 B^4 - \varphi\Theta^2 B^5 - \Theta^3 B^6 + \varphi\Theta^3 B^7 + \cdots)X_t$$

$$\Rightarrow \pi_j = \begin{cases} -\varphi(-\Theta)^{(j-1)/2}, \text{ for } j \text{ odd} \\ (-\Theta)^{j/2}, \text{ for } j \text{ even} \end{cases}$$

(c) The m-step ahead forecasts based on the infinite past ($\tilde{X}_{n+m}$) are given by $\tilde{X}_{n+m} = \sum_{j=1}^{\infty} \pi_j \tilde{X}_{n+m-j}$, using the invertible weight π_j defined in (b) and where $\tilde{X}_{n+m-j} = X_{n+m-j}$ for $j \ge m$. The forecast variance is given by $P_{n+m}^n = \sigma_w^2 \sum_{j=0}^{m-1} \psi_j^2$, where ψ_j are the model's causal weights, which are found as:

$$X_t = \varphi X_{t-1} + W_t + \Theta W_{t-2} \Rightarrow (1-\varphi B)X_t = (1+\Theta B^2)W_t \Rightarrow$$

$$\Rightarrow X_t = \sum_{j=0}^{\infty} \psi_j W_{t-j} = (1-\varphi B)^{-1}(1+\Theta B^2)W_t$$

But we already know that $(1-\varphi B)^{-1} = (1+\varphi B + \varphi^2 B^2 + \cdots) \Rightarrow$

$$\Rightarrow X_t = \sum_{j=0}^{\infty}\psi_j W_{t-j} = (1+\varphi B+\varphi^2 B^2+\cdots)(1+\Theta B^2)W_t =$$

$$= (1+\varphi B+\varphi^2 B^2+\varphi^3 B^3+\varphi^4 B^4+\cdots$$

$$+\Theta B^2+\varphi\Theta B^3+\varphi^2\Theta B^4+\varphi^3\Theta B^5+\cdots)W_t =$$

$$= W_t+\varphi W_{t-1}+(\varphi^2+\Theta)W_{t-2}+\varphi(\varphi^2+\Theta)W_{t-3}+\varphi^2(\varphi^2+\Theta)W_{t-4}+\cdots$$

$$= W_t+\varphi W_{t-1}+\sum_{j=2}^{\infty}\varphi^{j-2}(\varphi^2+\Theta)W_{t-j} \Rightarrow \psi_j = \begin{cases} 1, \text{ for } j=0 \\ \varphi, \text{ for } j=1 \\ \varphi^{j-2}(\varphi^2+\Theta), \text{ for } j\ge 2 \end{cases}$$

Thus, $P_{n+m}^{n} = \sigma_w^2\sum_{j=0}^{m-1}\psi_j^2 = \sigma_w^2\left(1+\varphi^2+\sum_{j=2}^{m-1}\varphi^{2(j-2)}(\varphi^2+\Theta)^2\right)$.