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 $j = 1, \ldots, m$)

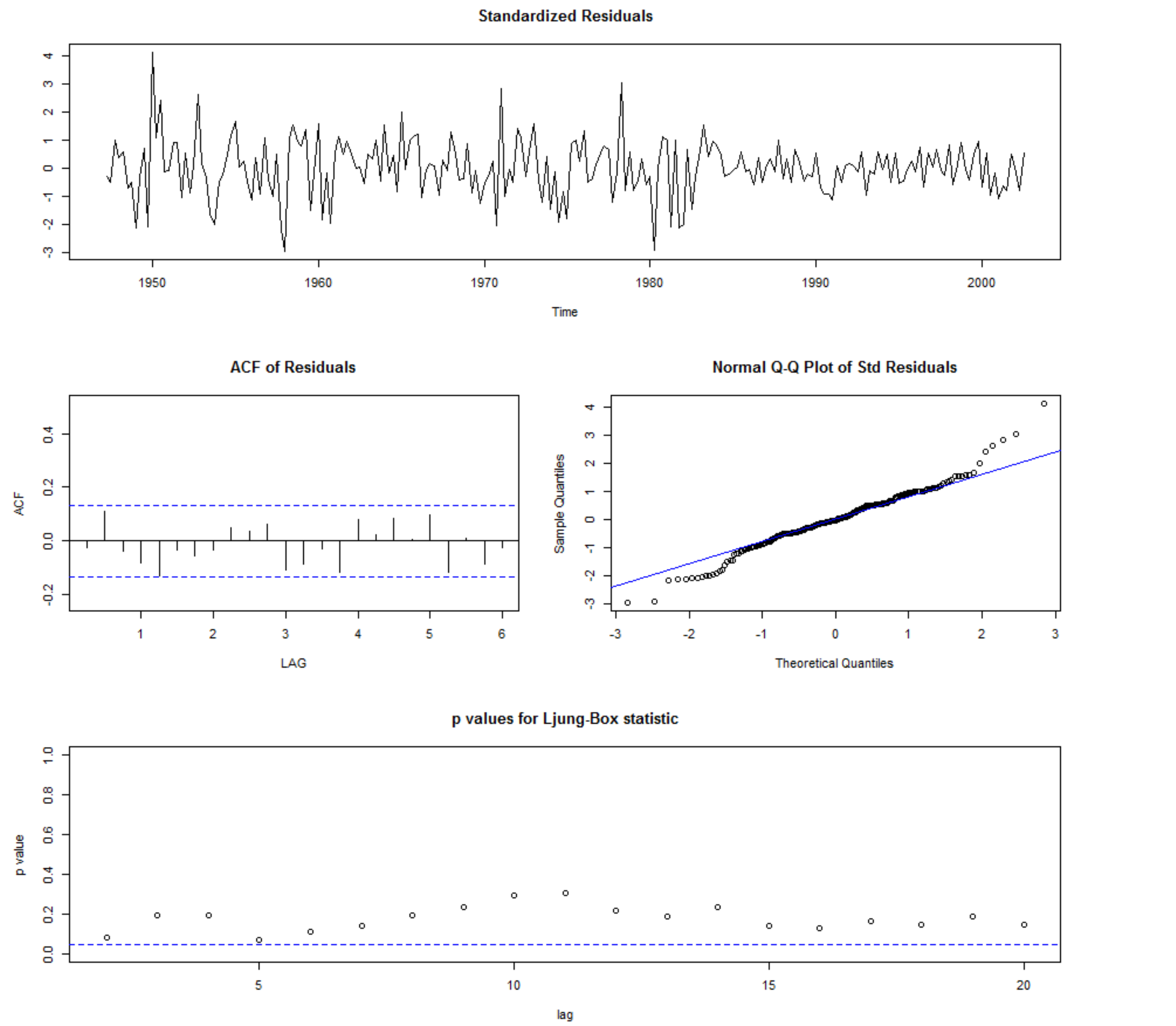
$$\Rightarrow \sum_{j=1}^{m} \left(x_{n+j}^{n} - x_{n+j-1}^{n} \right) = \sum_{j=1}^{m} y_{n+j}^{n} = \sum_{j=1}^{m} \left(\delta \frac{1-\varphi^{j}}{1-\varphi} + \varphi^{j} y_{n} \right) \Rightarrow$$

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

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

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

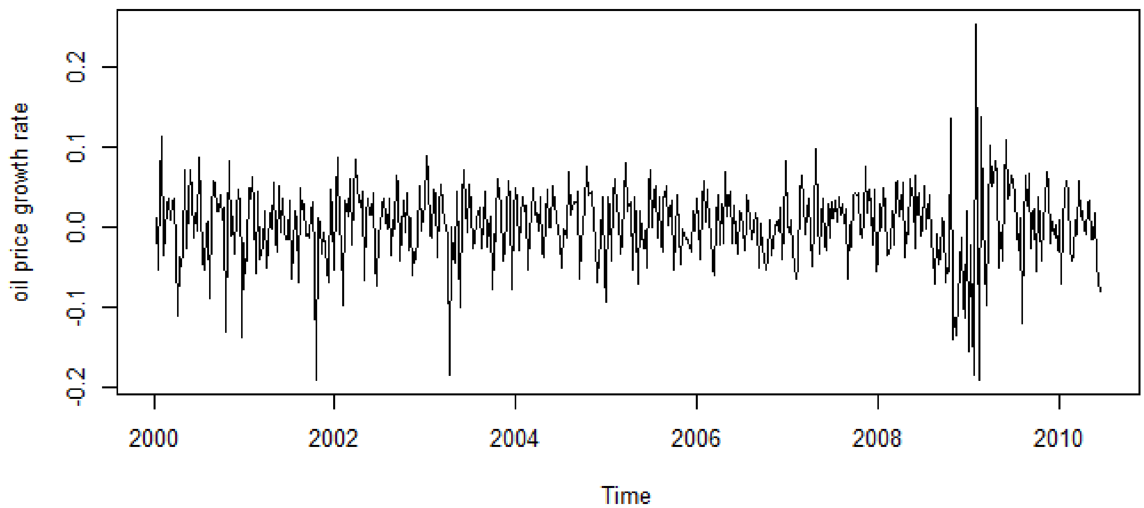
3. Exercise 3.31 from the textbook.

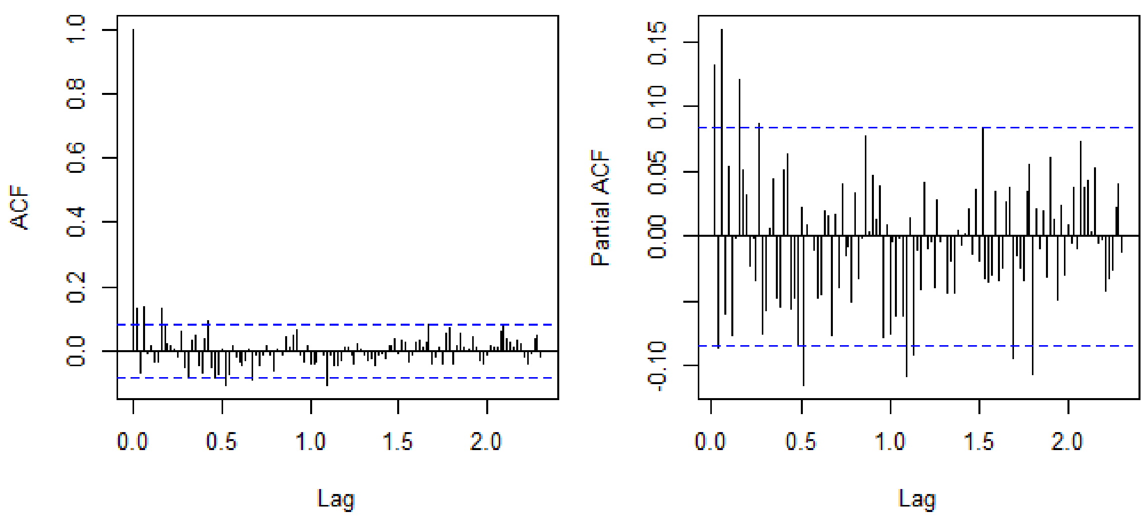


Overall, the model residuals do not show significant deviations from the assumptions. Their mean and variance seem relatively constant. The Q-Q plot shows heavier than Normal tails (which is usual in practice). There seems to be no significant residual auto-correlation in the residuals.

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)





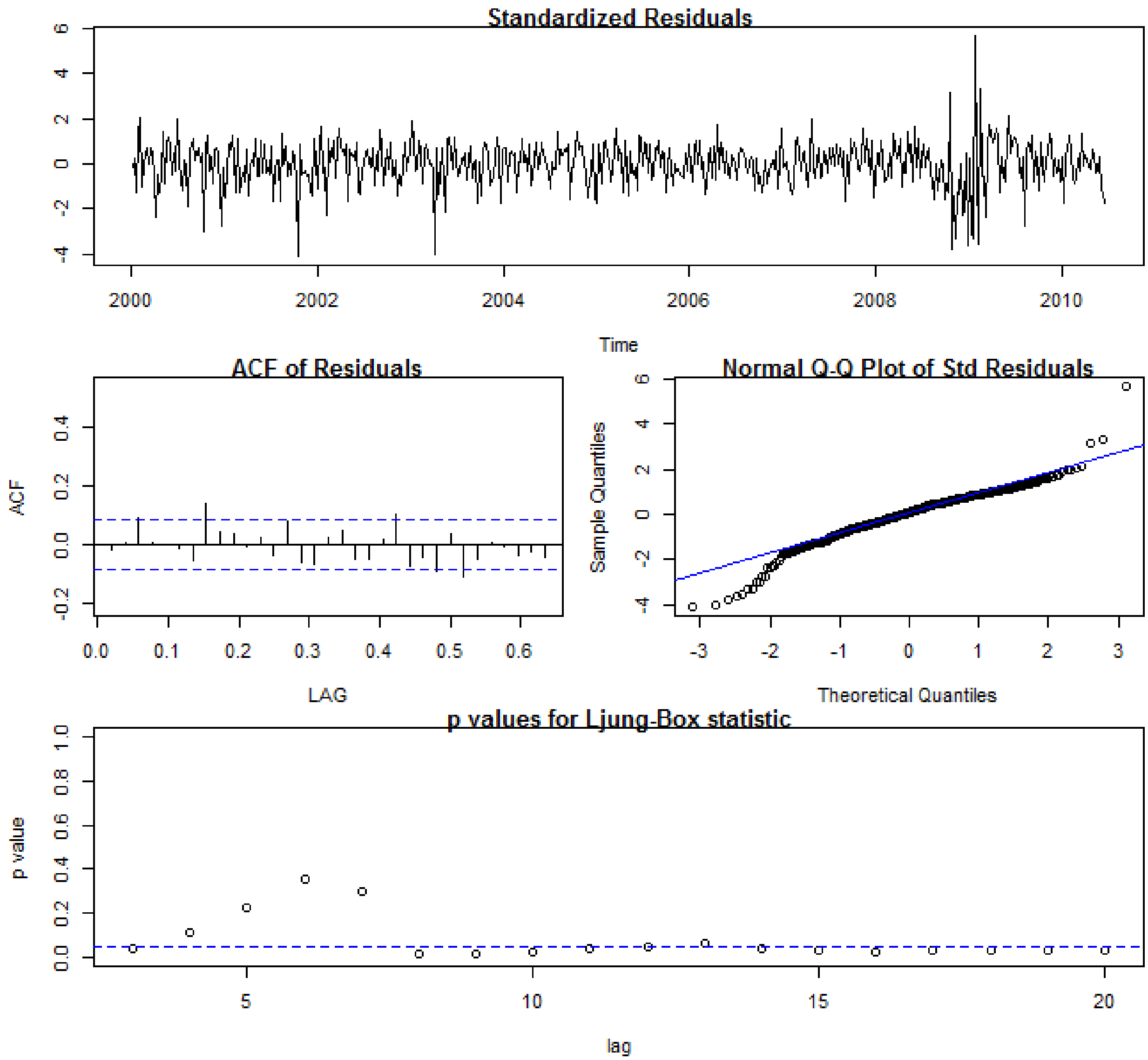
It is not clear from the ACF/PACF what the best model specification is, so we can try different models and compare some criterion, like the AIC. The `auto.arima()` function in the `forecast` package (with option `seasonal=FALSE`) gives an ARIMA(1,0,1) specification for the growth rate, with the following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



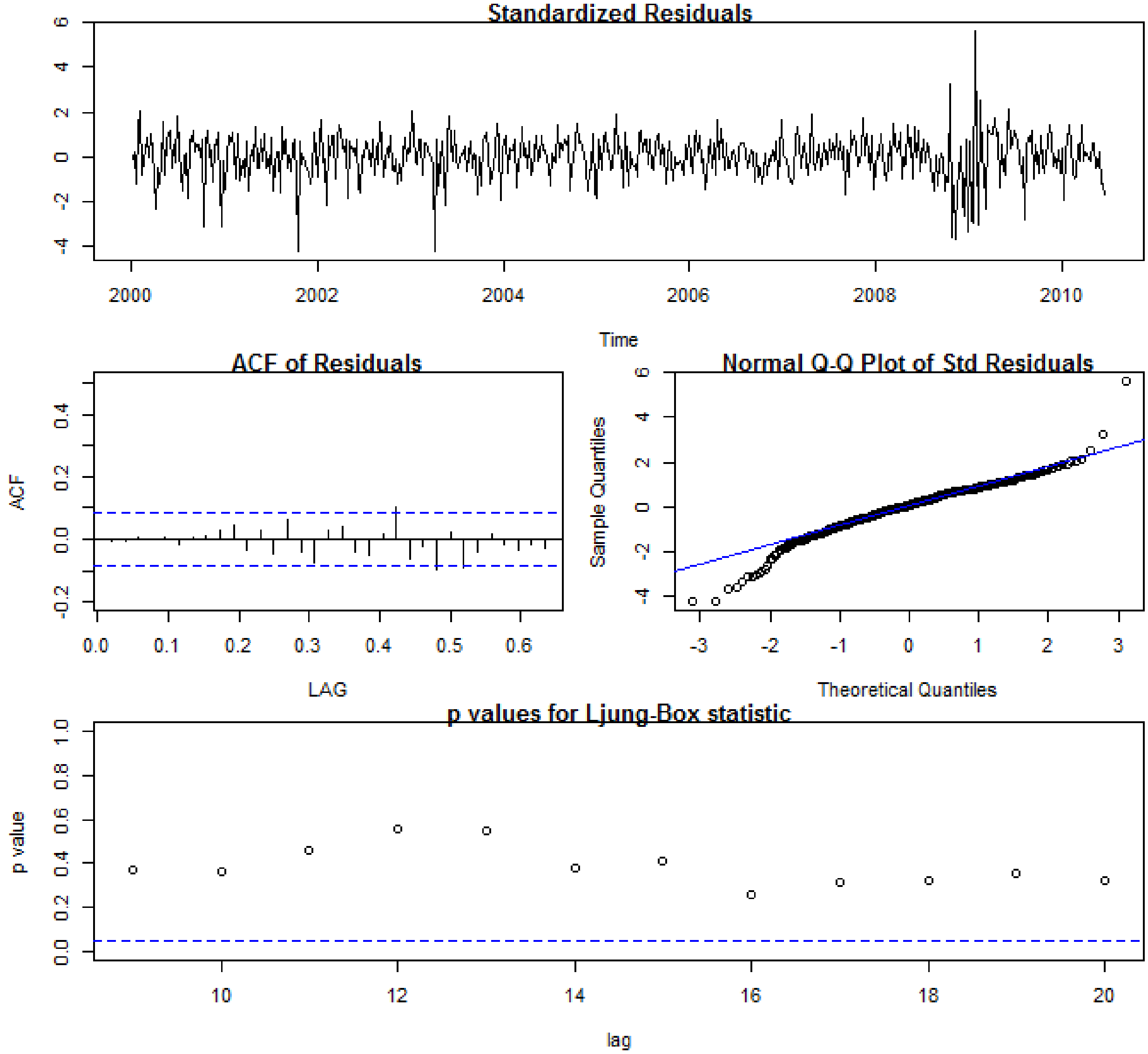
Therefore, we can try a bigger model (higher lags). In particular, Yule-Walker estimation for purely autoregressive models gives an AR(8) model specification with lower AIC:

```
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:



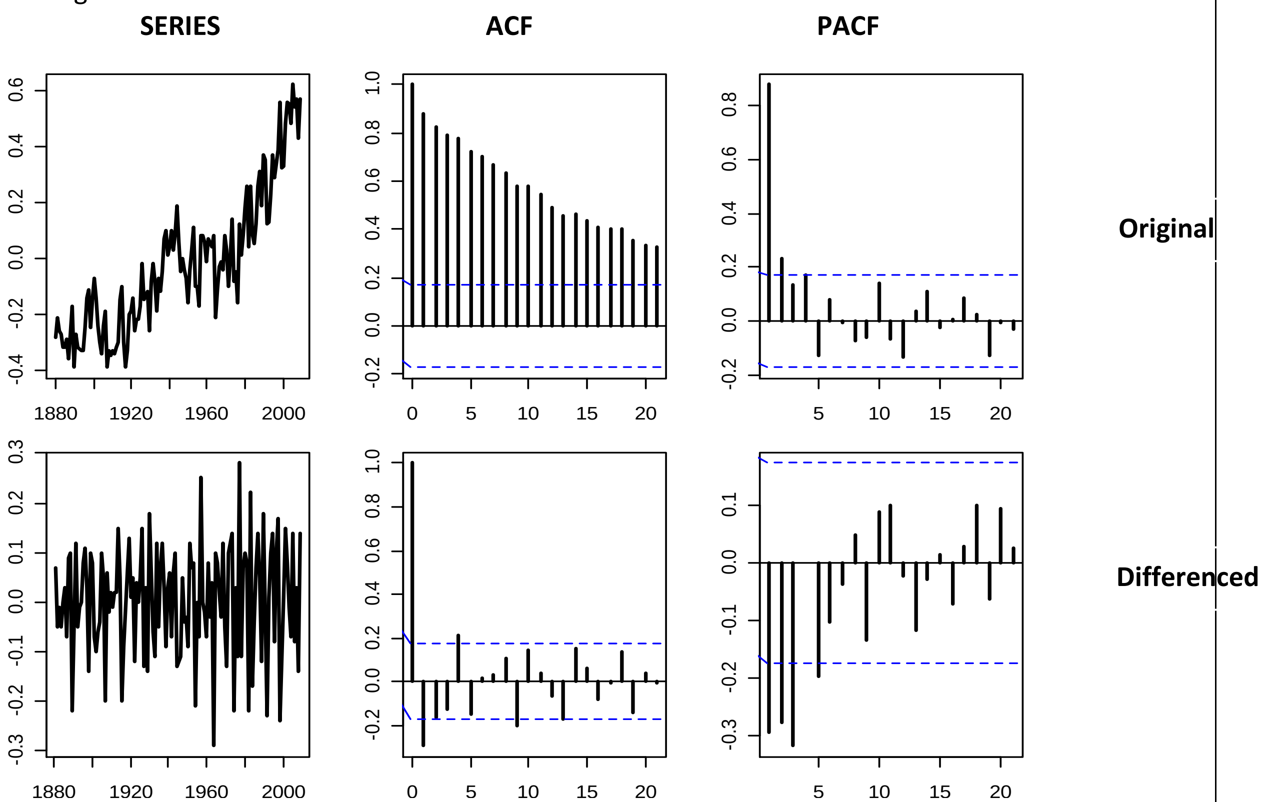
Adding an MA component, i.e. trying an ARMA(8,1) model, does not offer substantial improvements, so we can stick with the AR(8) model for the growth rate, (or, equivalently, an ARIMA(8,1,0) model for the log-price, if we want to use the model to forecast the actual price).

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:

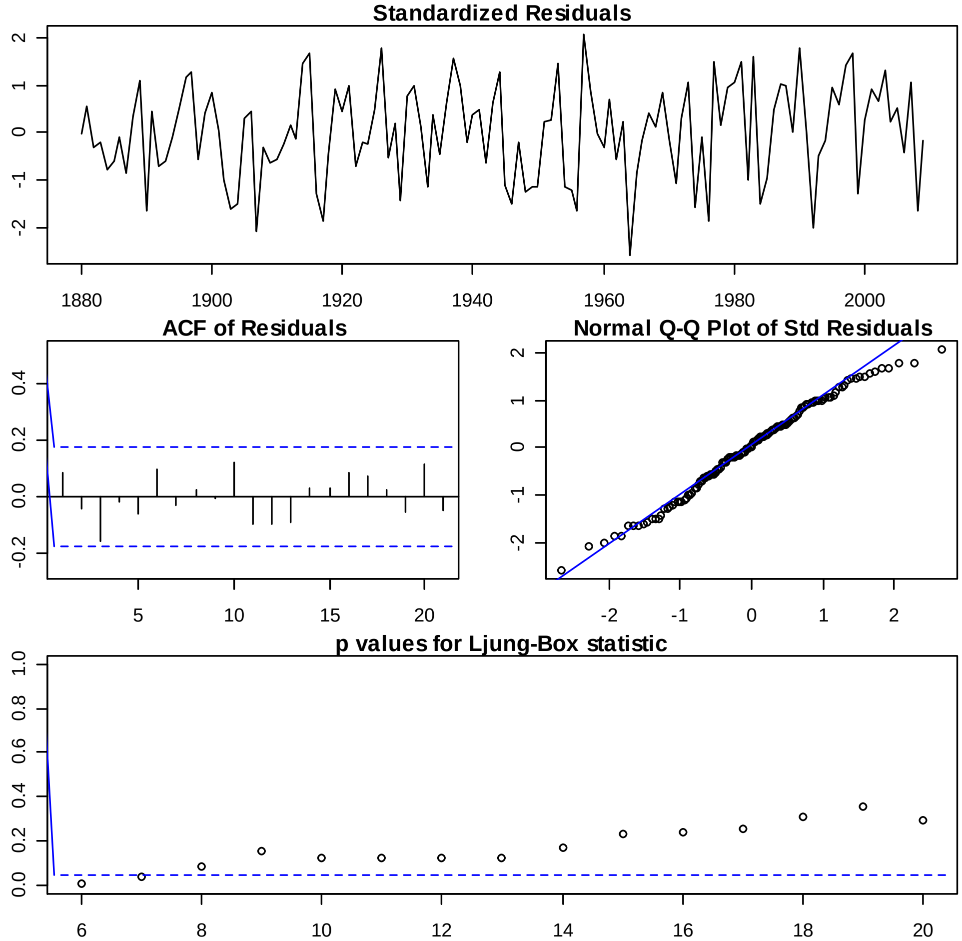


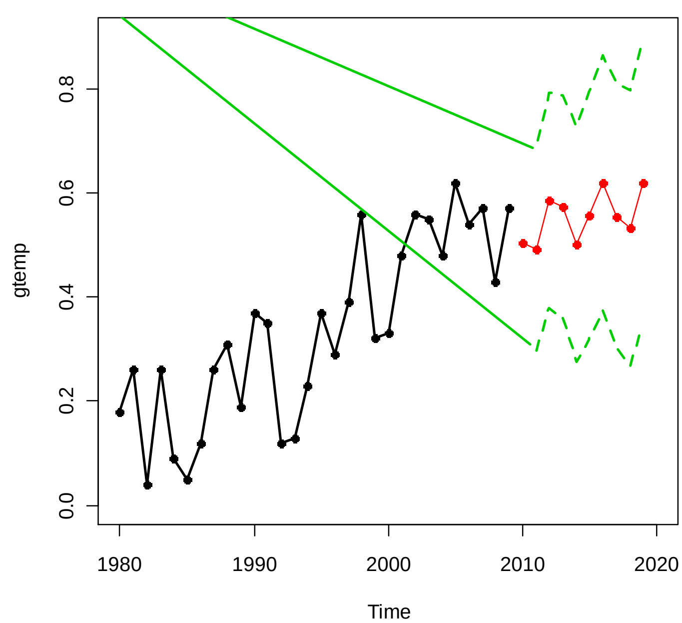
The best model selected using AIC over all possible ARIMA(p,1,q) with p,q≤7 is ARIMA(2,1,3) with parameter estimates:

```
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.

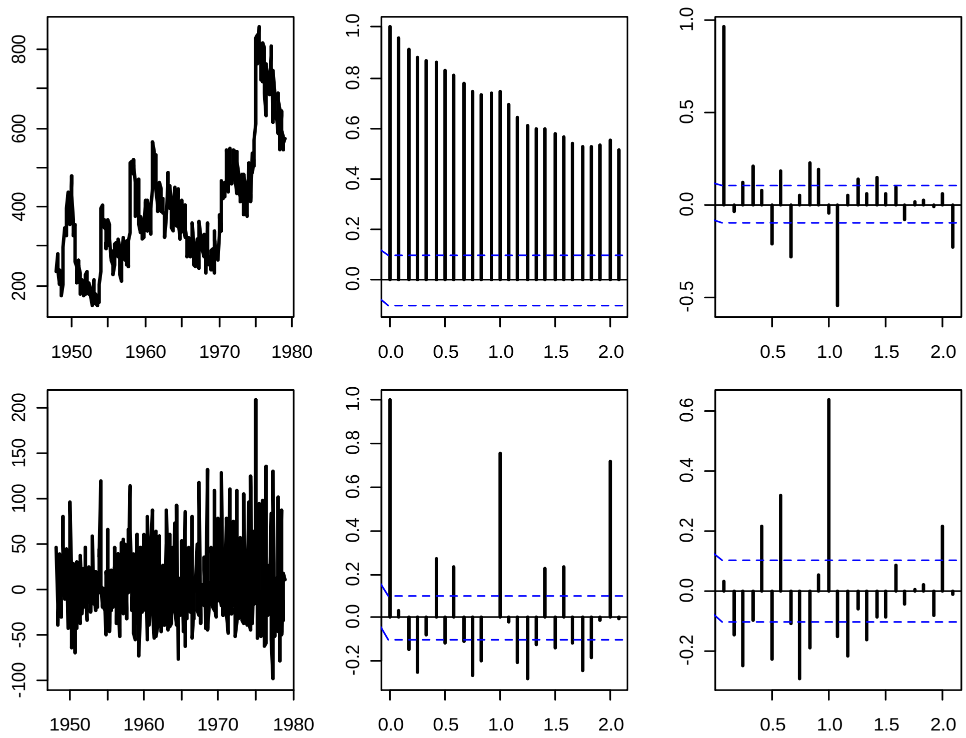
(b) We have $X_t = (1+\Theta B^2)W_t = \Theta(B^2)W_t$ where $\Theta(z) = 1+\Theta z \Rightarrow \Theta^{-1}(z) = \sum_{j=0}^{\infty}(-\Theta z)^j$, so the invertible representation of the series is $W_t = \Theta^{-1}(B^2)X_t = \sum_{j=0}^{\infty}(-\Theta B^2)^j X_t = \sum_{j=0}^{\infty}(-\Theta)^j X_{t-2j}$.

(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**

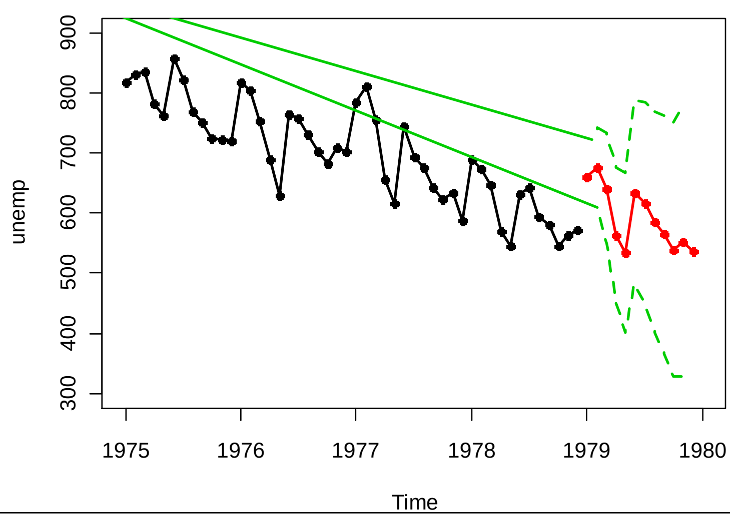


The auto.arima() function picked the SARIMA(2,1,0)×$(2,0,0)_{12}$ model

```
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



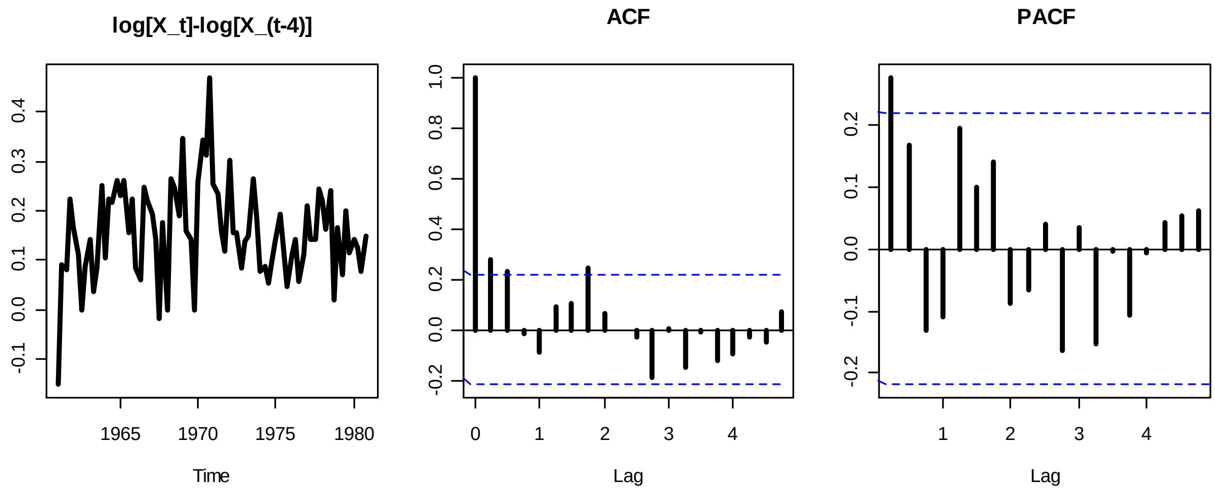
8. Exercise 3.44 from the textbook.

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

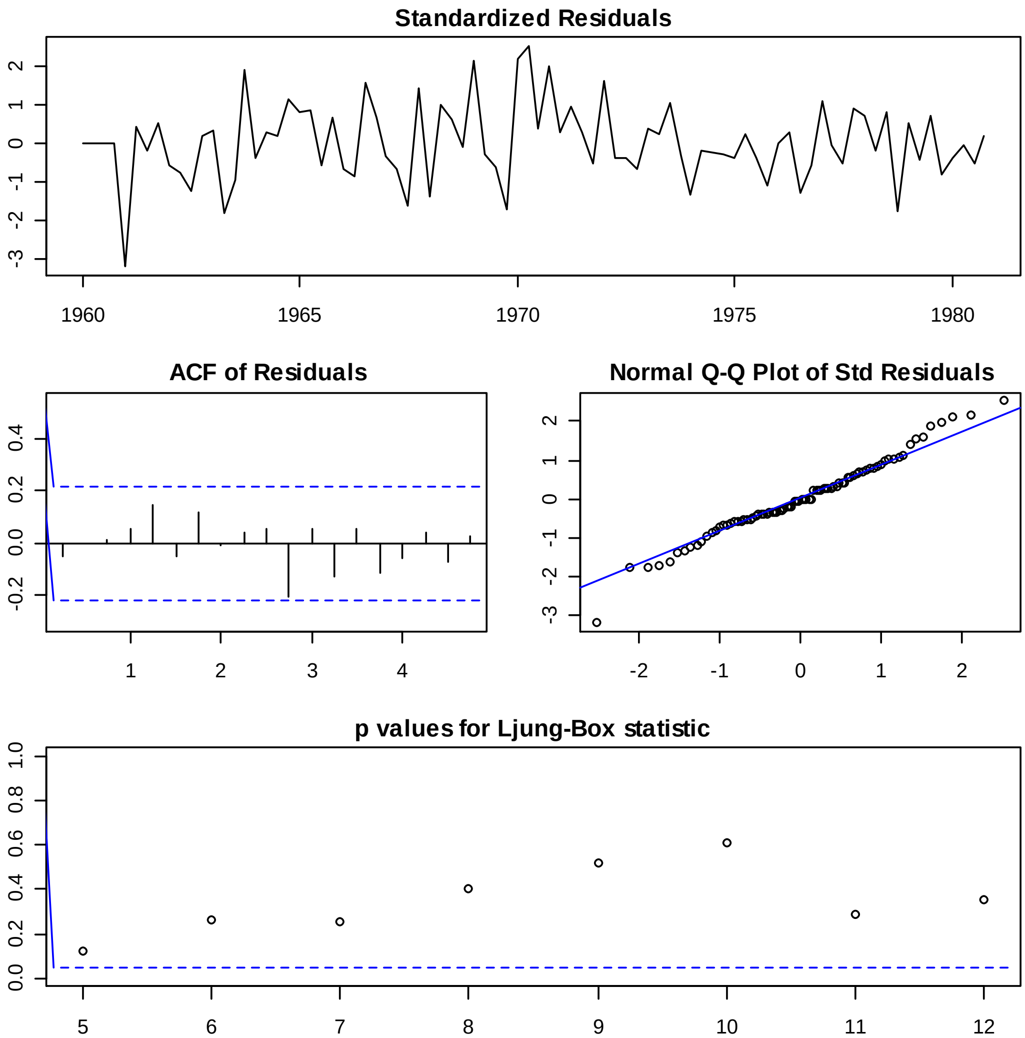
```
Coefficients:
         ar1      ar2      ma1     ma2  constant
      0.6798  -0.6133  -0.4031  0.7998    0.0386
s.e.  0.1650   0.1693   0.1228  0.1270    0.0035

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

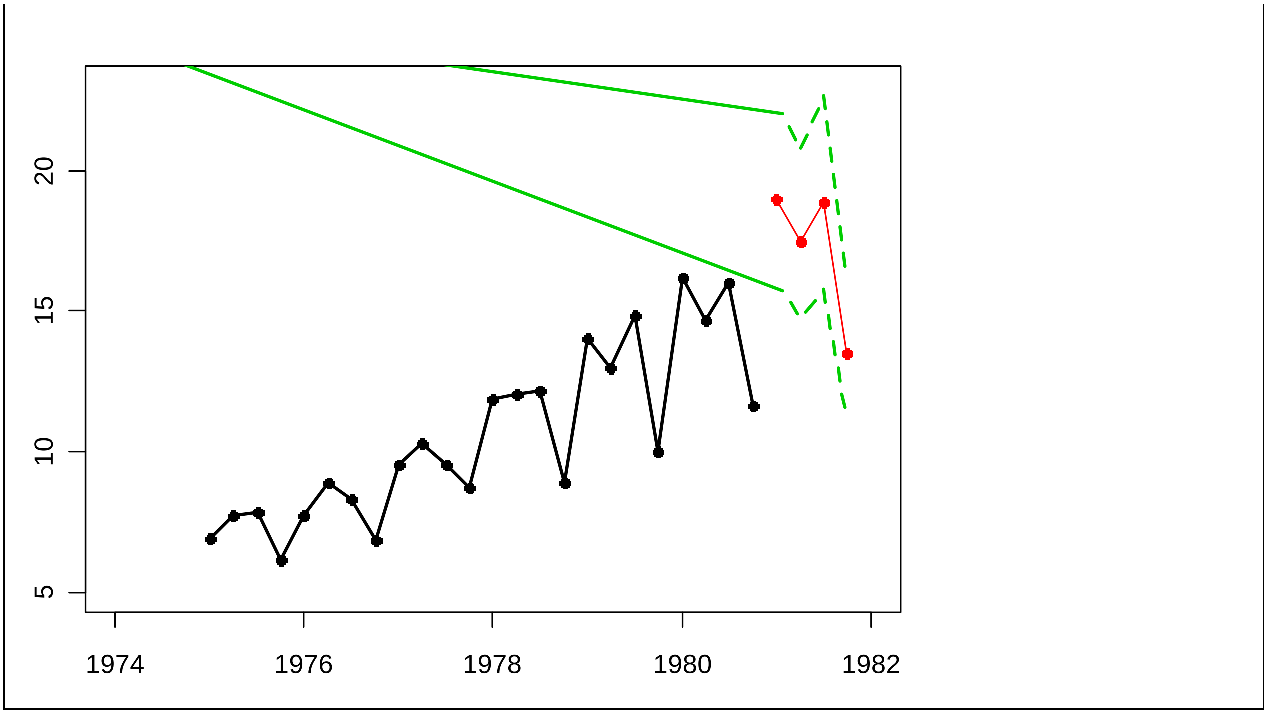
Note that the model has an integrated multiplicative seasonal component with period 4 (i.e. annual period for quarterly data). The seasonally differenced log-series at lag 4 (i.e. the stationary ARMA part of the model) is shown below, together with its ACF & PACF:



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 $|\Theta|<1$, so the model is invertible if $|\Theta|<1$. 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 $j \geq m$. 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 $(1-\varphi B)^{-1} = (1+\varphi B + \varphi^2 B^2 + \cdots) \Rightarrow$

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

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

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

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

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

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