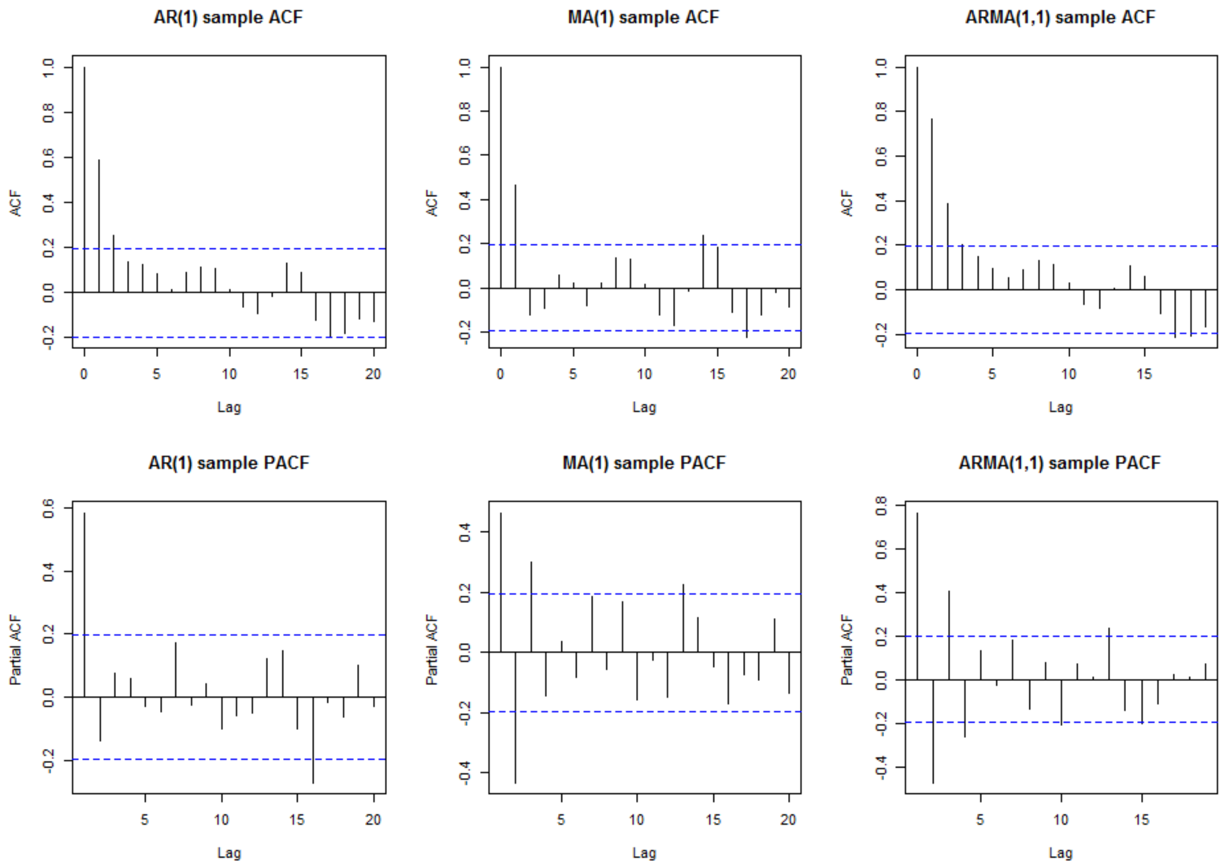
STAD57: Time Series Analysis
Problem Set 4

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

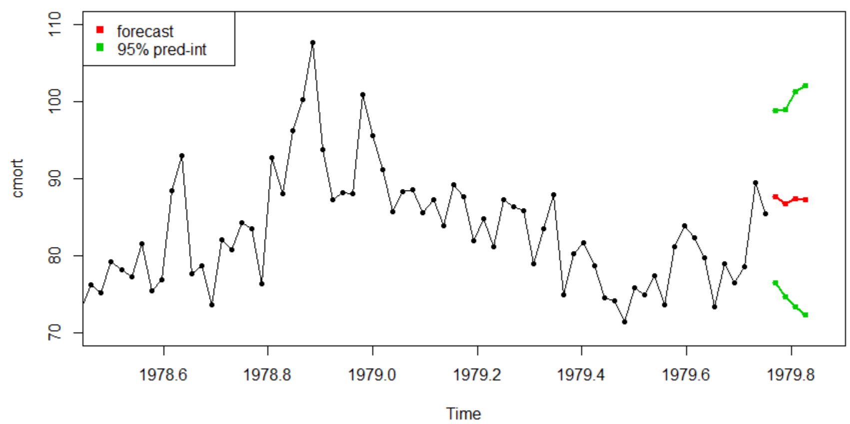
```
Call:
ar.ols(x = cmort, order.max = 2, demean = FALSE, intercept = TRUE)

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 $g(X)=\mathbb{E}[Y \mid X]$ (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:

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

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

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

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

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

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

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

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 $X_{n+m}^n=\sum_{j=1}^n\varphi_{nj}^{(m)}X_{n+1-j}=\varphi^mX_n$, with MSE

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

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 $\pi_j=-1.4\left(-.5\right)^{j-1},\ \forall j\ge 1$ (see Example .3.7 on p. 95). Thus, the truncated predictions from equation (3.91) are: $\tilde{x}_{n+m}^n=-\sum_{j=1}^{m-1}\pi_j\tilde{x}_{n+m-j}^n-\sum_{j=m}^{1+m-1}\pi_jx_{n+m-j}^n\Rightarrow$

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

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

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

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

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

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

$$\vdots$$

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

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

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

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

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

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

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

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

$$\vdots$$

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

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

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

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

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:

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

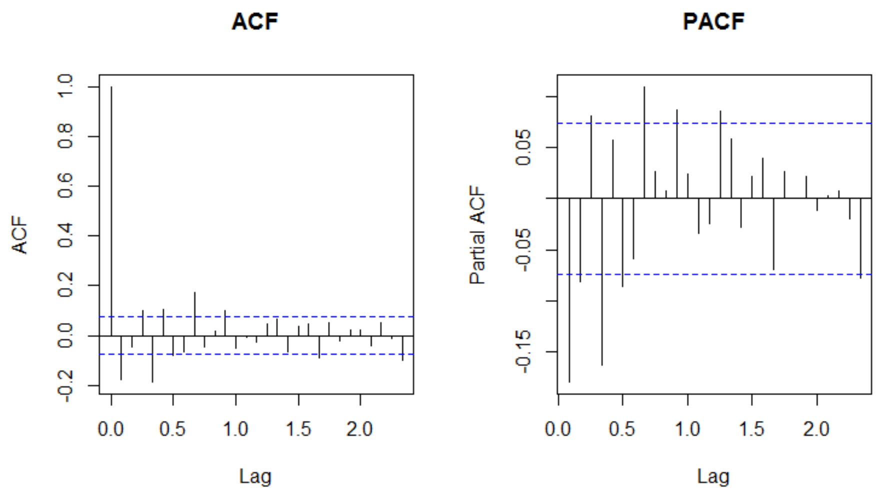
$$= E\left[\left(\sum_{j=0}^{m-1}\psi_j W_{n+m-j}\right)\left(\sum_{i=0}^{m-1}\psi_i W_{n+m+k-i}\right)\right] =$$

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

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.

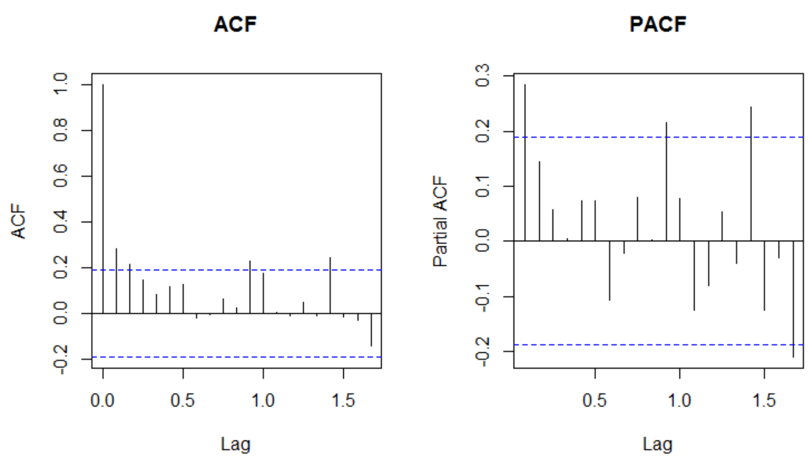
We have the following ACF/PACF plots:

a. Monthly Canadian reserves (in $)



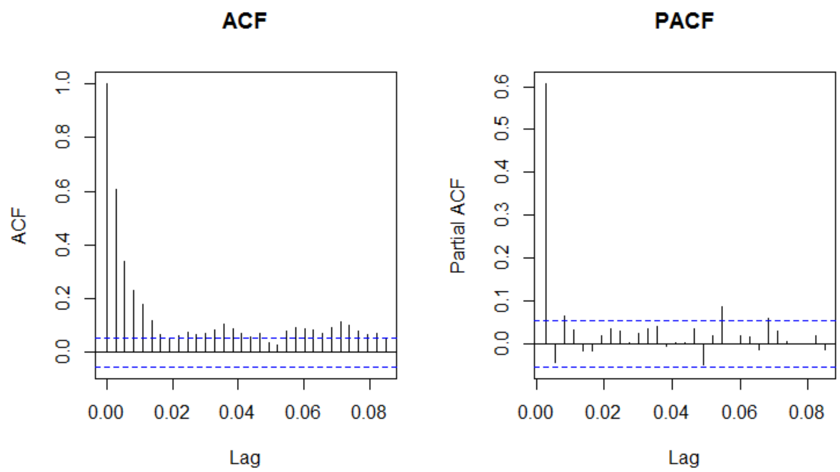
Since neither the ACF or the PACF seem to cut-off, we can go with a general ARMA(p,q) model. The exact order of the model is not obvious, but we should have p≥1 *AND* q≥1.

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	$\frac{.15}{.3}=\frac{1}{2}$

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

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

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

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

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

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

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

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

$$\mathrm{Cov}[Y,X]=E[X\cdot Y]-\cancel{E[X]\cdot E[Y]}=\sum_{x=-1}^{+1} y\cdot x\cdot P(Y=y,X=x)=$$

$$=(-1)(-1)\times.05+(0)(-1)\times.10+(1)(-1)\times.15+$$

$$+(-1)(0)\times.20+(0)(0)\times.10+(1)(0)\times.10+$$

$$+(-1)(1)\times.15+(0)(1)\times 0+(1)(1)\times.15=\underline{-0.1}$$

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

$$\mathbb{E}\left[\left(Y-g(X)\right)^2\right]=\sum_{y=-1}^{+1}\sum_{x=-1}^{+1}\left(y-g(x)\right)^2 P(Y=y,X=x)=$$

$$=(-1-1/6)^2\times.05+(0-1/6)^2\times.10+(1-1/6)^2\times.15+$$

$$+(-1-0)^2\times.20+(0-0)^2\times.10+(1-0)^2\times.10+$$

$$+(-1+1/6)^2\times.15+(0+1/6)^2\times 0+(1+1/6)^2\times.15=\underline{0.7833333}$$