Topics in Time Series (4)

Yisen Du (Eason)

1 Question1

Consider the dataset in "norm.nao.monthly.b5001.current.ascii04Nov2022.txt' which gives (use the third column in the dataset) monthly data on the Northern Oscillation Index (this data has been taken from https://www. cpc.ncep.noaa.gov/products/ precip/CWlink/pna/nao.shtml; see this page for more details about the data).

a) I want to fit the $MA(q)$ model to this dataset. Look at the sample autocorrelation function of the data and figure out an appropriate value of q. (2 points)

b) Fit the $MA(q)$ (with your selected choice of q in the previous part) to the data. Use the conditional sum of squares method described in class (do not use any inbuild function in R for this part). Report point estimates and standard errors for $\mu, \theta_1, \ldots, \theta_q$ (5 points)

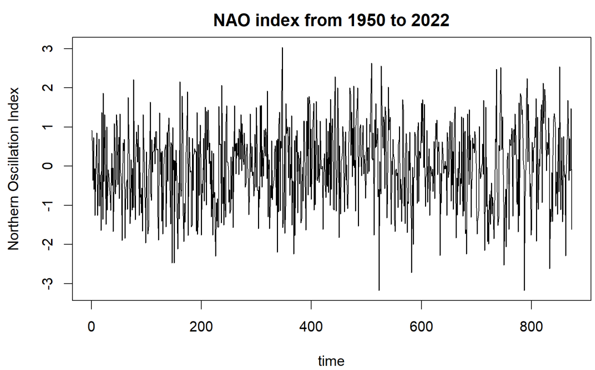
c) Compare your answers to that given by the arima function in R. (2 points)

d) Use your fitted model to obtain point predictions for the next 24 months. Comment on whether the predictions appear reasonable. (2 points)

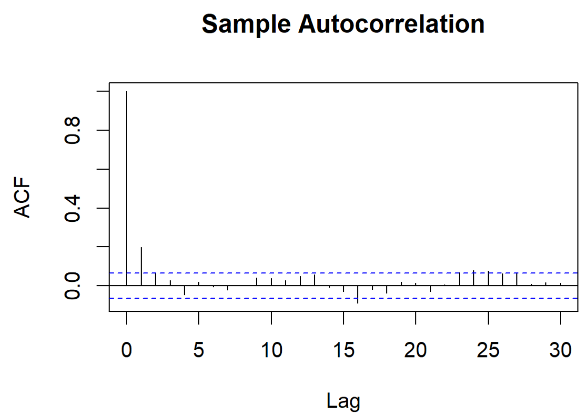
1.1 Answer of Q1(a)

Firstly, let's see the data. Since the level of the data is unchanged with time, we could directly apply MA model to it. In the lecture, we have shown how to use ACF to decide value of q. If the sample ACF exhibits the property that the values for lags larger than the particular lag q are relative small, then $MA(q)$ should be considered. The following ACF figure implies $q = 1$ is a good choice. The model is

$$Y_t = \mu + Z_t + \theta_1 Z_{t-1}, \text{ where } Z_t \overset{\text{iid}}{\sim} N(0, \sigma^2)$$



(a) Data overview



(b) ACF

1.2 Answer of Q1(b)

Using the conditional sum of squares method, the point estimates and standard errors are shown below.

	point estimate	standard error
μ	-0.0013	0.0403
θ_1	0.1852	0.032

Table 1: Inference result (conditional sum of squares)

1.3 Answer of Q1(c)

The two results are very similar.

	point estimate	standard error
μ	-0.0013	0.0402
θ_1	0.1851	0.0321

Table 2: Inference result (R function arima)

1.4 Answer of Q1(d)

Since the prediction result of $MA(1)$ is quite simple, it's reasonable because the value should around the mean. It's not sufficient because the future values could not all be constant.

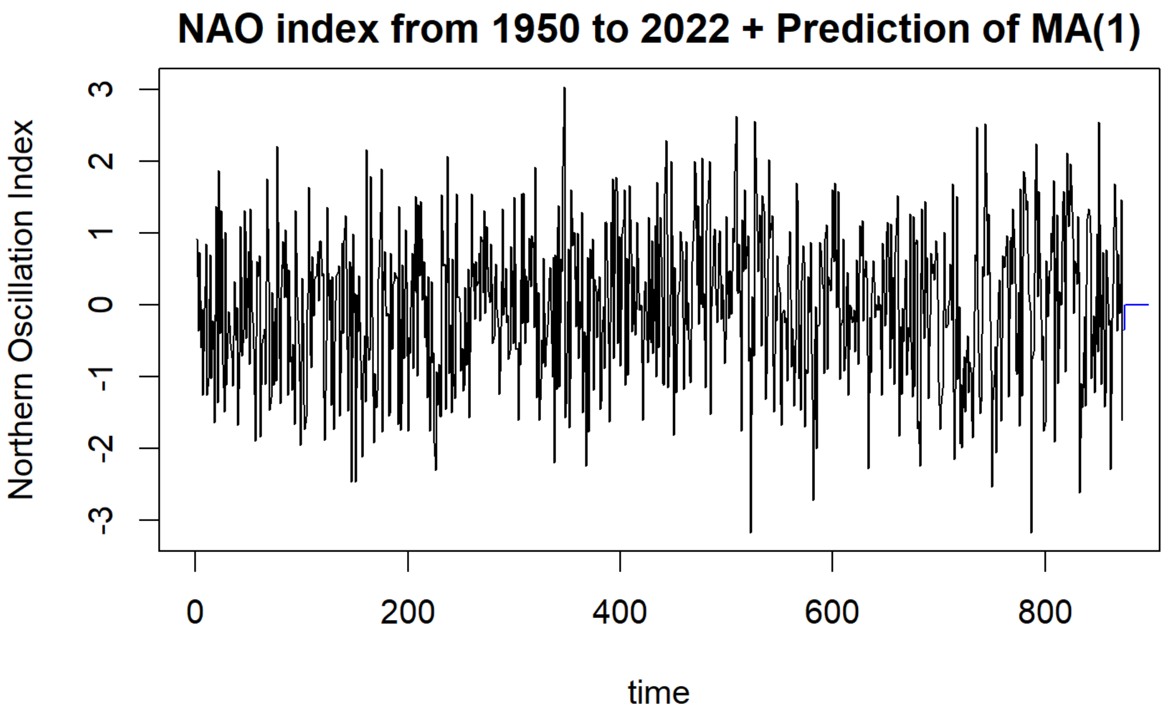


Figure 2: Prediction

2 Question2

Download the FRED dataset on "Long-Term Government Bond Yields: 10-year Main (including benchmark) for the United States" from https://fred.stlouisfed.org/ series/IRLTLT01USM156N. This is a monthly dataset (units are in percent) and it is not seasonally adjusted.

a) Fit an $AR(p)$ model to this datset with $p = 4$. Write the model as

$$Y_t = \hat{\phi}_0 + \hat{\phi}_1 Y_{t-1} + \cdots + \hat{\phi}_p Y_{t-p} + Z_t$$

and report parameter estimates and standard errors for $\hat{\phi}_j, j = 0, 1, \ldots, p$. Use the model to obtain predictions for the next 100 months. Do the predictions look reasonable? (5 points)

b) Let Y_t denote the original dataset. Construct a new dataset D_t by differencing $Y_t : D_t = Y_t - Y_{t-1}$. Plot the dataset D_t with time on the x-axis. Also plot the sample autocorrelation function of $\{D_t\}$. Would the $MA(1)$ model be reasonable for $\{D_t\}$?(3 points)

c) Fit the $MA(1)$ model to $\{D_t\}$ and obtain point estimates and standard errors of the parameters (you can use the R function arima). Denote this model by (**2** points)

$$D_t = \hat{\mu} + \epsilon_t + \hat{\theta}\epsilon_{t-1} \quad \text{where } \epsilon_t \overset{\text{i.i.d}}{\sim} N\left(0, \sigma^2\right).$$

d) Rewriting (2) as

$$\epsilon_t = (I + \hat{\theta}B)^{-1}(D_t - \hat{\mu}) = \left(I - \hat{\theta}B + \hat{\theta}^2 B^2 - \hat{\theta}^3 B^3 + \ldots\right)(D_t - \hat{\mu}),$$

approximate the model (2) by an autoregressive model of the form:

$$D_t = \hat{\psi}_0 + \hat{\psi}_1 D_{t-1} + \hat{\psi}_2 D_{t-2} + \hat{\psi}_3 D_{t-3} + \epsilon_t$$

and report the values of $\hat{\psi}_0, \hat{\psi}_1, \hat{\psi}_2, \hat{\psi}_3$. (2 points)

e) Replace $D_s = Y_s - Y_{s-1}$ on both sides of the equation in (3) to obtain an AR model for Y_t. Compare the coefficients of this AR model with those of (1). Are they similar? (3 points)

f) Use the AR model from the previous part to obtain predictions for the next 100 months. Compare these predictions with those obtained from part (a). Comment on the differences between these two predictions. (5 points).

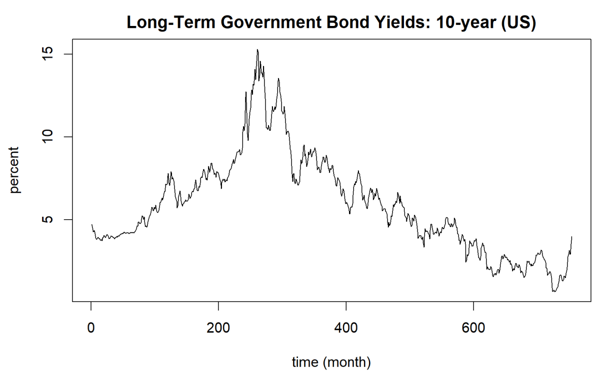
2.1 Answer of Q2(a)

The estimation of $AR(4)$ is shown below.

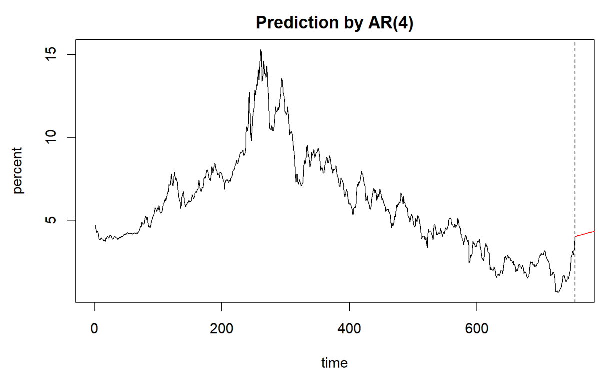
	point estimate	standard error
$\hat{\phi}_0$	0.0278	0.0004
$\hat{\phi}_1$	1.3814	0.0013
$\hat{\phi}_2$	-0.6132	0.0038
$\hat{\phi}_3$	0.3011	0.0038
$\hat{\phi}_4$	-0.0741	0.0013

Table 3: Estimations

The prediction looks not so reasonable because it's **linear** while previous data shows a **oscillation** pattern.



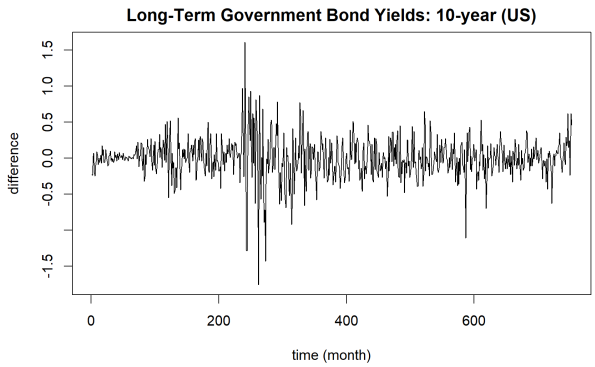
(a) Data overview



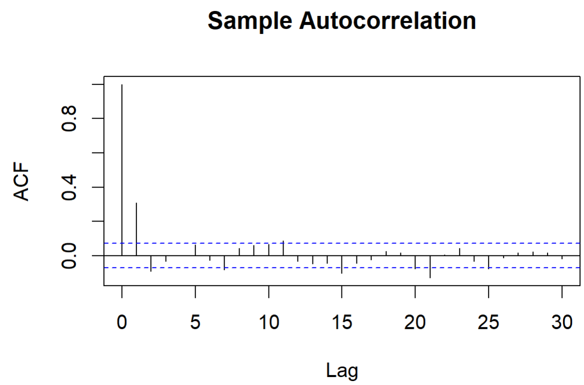
(b) Prediction

2.2 Answer of Q2(b)

The ACF shows $MA(1)$ is reasonable since all the lags larger than 1 are relative small.



(a) Data overview: difference



(b) ACF

2.3 Answer of Q2(c)

The estimation of $MA(1)$ is shown below.

	point estimate	standard error
$\hat{\mu}$	-0.0010	0.0134
$\hat{\theta}$	0.4338	0.0365

Table 4: Estimations

2.4 Answer of Q2(d)

The values are shown below.

	point estimate
$\hat{\phi}_0$	-0.0003
$\hat{\phi}_1$	0.3841
$\hat{\phi}_2$	-0.2296
$\hat{\phi}_3$	0.0709

Table 5: Estimations

2.5 Answer of Q2(e)

Replace $D_s = Y_s - Y_{s-1}$ to equation (3), we have

$$Y_t = \hat{\phi}_0 + (\hat{\phi}_1 + 1)Y_{t-1} + (\hat{\phi}_2 - \hat{\phi}_1)Y_{t-2} + (\hat{\phi}_3 - \hat{\phi}_2)Y_{t-3} + (-\hat{\phi}_3)Y_{t-4} + \epsilon_t$$

The parameters are given by

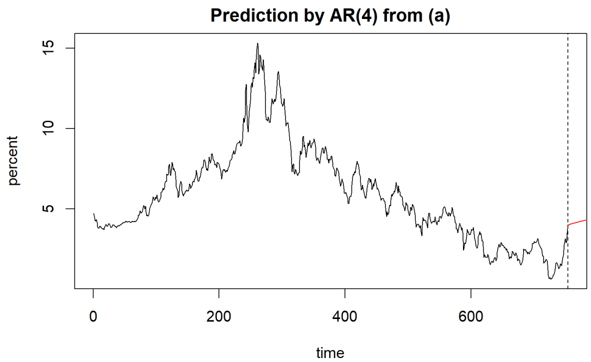
	point estimates from (d)	point estimates from (a)
$\hat{\phi}_0$	-0.0003	0.0278
$\hat{\phi}_1$	1.3841	1.3814
$\hat{\phi}_2$	-0.6138	-0.6132
$\hat{\phi}_3$	0.3005	0.3011
$\hat{\phi}_4$	-0.0709	-0.0741

Table 6: Estimations

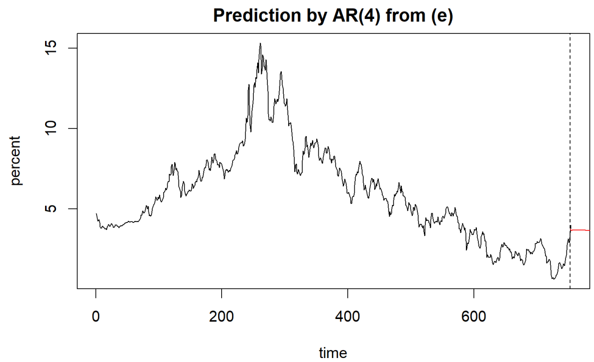
The results shows these two estimations are very similar except the intercept value.

2.6 Answer of Q2(f)

Due to the difference of intercept term, the two AR models are quite different. While the first one shows increasing pattern, the second one shows stationary pattern.



(a) using parameters from (a)



(b) using parameters from (e)

3 Question3

Download the FRED dataset on "Retail Sales: Beer, Wine, and Liquor Stores" from https://fred.stlouisfed.org/series/MrTSSM4453USN. This is a monthly dataset (the units are millions of dollars) and is not seasonally adjusted.

a) Fit an $AR(p)$ model to this dataset with $p = 16$. Write the model as

$$Y_t = \hat{\phi}_0 + \hat{\phi}_1 Y_{t-1} + \cdots + \hat{\phi}_p Y_{t-p} + Z_t$$

and report parameter estimates and standard errors for $\hat{\phi}_j, j = 0, 1, \ldots, p$. Use the model to obtain predictions for the next 36 months (3 years). Do the predictions look reasonable? (5 points)

b) Would any Moving Average model work directly on this dataset? Answer this question by trying out $MA(q)$ for a range of values of q. You can evaluate models by looking at their future predictions. Use the R function arima to fit models and the function predict to obtain future predictions (5 points).

c) Let Y_t denote the original dataset. Construct a new dataset D_t via:

$$D_t = Y_t - Y_{t-1} - Y_{t-12} + Y_{t-13}$$

This can be created in R by, for example, the command diff(diff(Yt, lag = 12)). Plot the dataset D_t with time on the x-axis. Also plot the sample autocorrelation function of $\{D_t\}$. Would the $MA(1)$ model be reasonable for $\{D_t\}$?(3 points)

d) Fit the $MA(1)$ model to $\{D_t\}$ and obtain point estimates and standard errors of the parameters (you can use the R function arima). Denote this model by (**2** points)

$$D_t = \hat{\mu} + \epsilon_t + \hat{\theta}\epsilon_{t-1} \quad \text{where } \epsilon_t \overset{\text{i.i.d}}{\sim} N\left(0, \sigma^2\right).$$

e) Rewriting (5) as

$$\epsilon_t = (I + \hat{\theta}B)^{-1}(D_t - \hat{\mu}) = \left(I - \hat{\theta}B + \hat{\theta}^2 B^2 - \hat{\theta}^3 B^3 + \ldots\right)(D_t - \hat{\mu}),$$

approximate the model (5) by an autoregressive model of the form:

$$D_t = \hat{\psi}_0 + \hat{\psi}_1 D_{t-1} + \hat{\psi}_2 D_{t-2} + \hat{\psi}_3 D_{t-3} + \epsilon_t$$

and report the values of $\hat{\psi}_0, \hat{\psi}_1, \hat{\psi}_2, \hat{\psi}_3$. (2 points)

f) Replace $D_s = Y_s - Y_{s-1} - Y_{s-12} + Y_{s-13}$ on both sides of the equation in (6) to obtain an AR model for Y_t. Compare the coefficients of this AR model with those of (4). Are they similar? (3 points)

g) Use the AR model from the previous part to obtain predictions for the next 36 months. Compare these predictions with those obtained from part (a). Comment on the differences between these two predictions. (5 points).

3.1 Answer of Q3(a)

The parameters and standard errors are shown below.

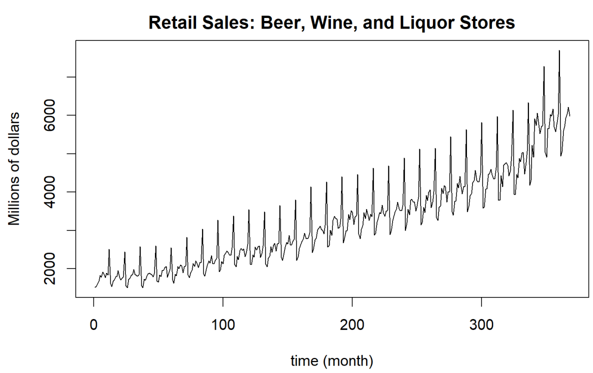
	point estimate	standard error
$\hat{\phi}_0$	-6.5251	3.190382e+02
$\hat{\phi}_1$	0.2932	2.945594e-03
$\hat{\phi}_2$	0.4054	2.922943e-03
$\hat{\phi}_3$	0.2977	2.946972e-03
$\hat{\phi}_4$	-0.1632	3.040883e-03
$\hat{\phi}_5$	0.0081	1.705457e-04
$\hat{\phi}_6$	0.0057	1.708012e-04
$\hat{\phi}_7$	-0.0011	1.709566e-04

Table 7: Estimations

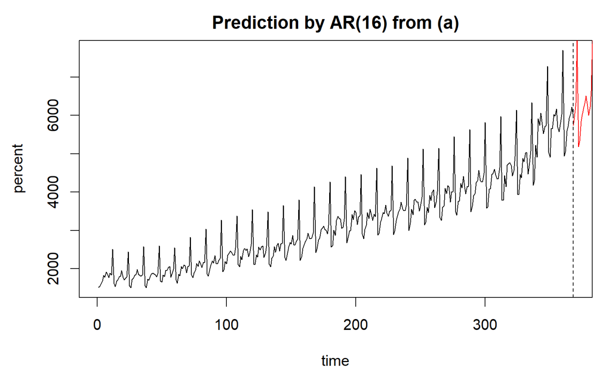
	point estimate	standard error
$\hat{\phi}_8$	-0.0010	1.738275e-04
$\hat{\phi}_9$	-0.0077	1.825807e-04
$\hat{\phi}_{10}$	-0.0186	1.827593e-04
$\hat{\phi}_{11}$	0.0083	1.842124e-04
$\hat{\phi}_{12}$	1.0137	1.842901e-04
$\hat{\phi}_{13}$	-0.3046	3.208488e-03
$\hat{\phi}_{14}$	-0.4107	3.175240e-03
$\hat{\phi}_{15}$	-0.3013	3.159911e-03
$\hat{\phi}_{16}$	0.1858	3.240494e-03

Table 8: Estimations (continues)

The predictions look reasonable because it both captures the trend and the periodicity.



(a) Data overview



(b) Prediction

3.2 Answer of Q3(b)

No, it wouldn't. The reason is there's obvious increasing trend of data, but MA models can only be applied to stationary data. The predictions of MA are very strange.

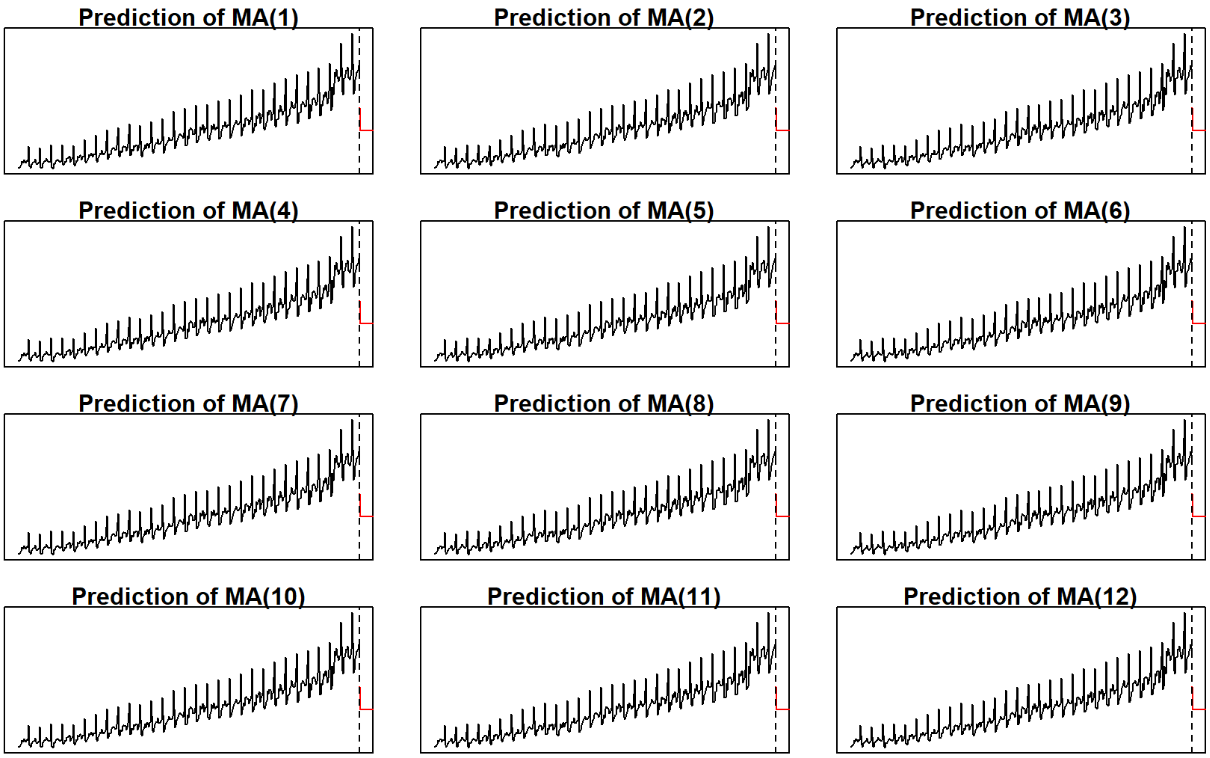
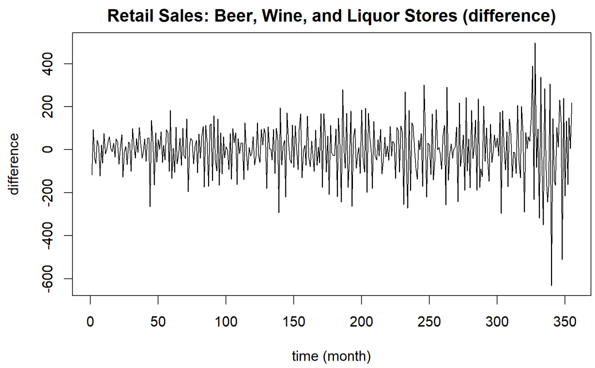


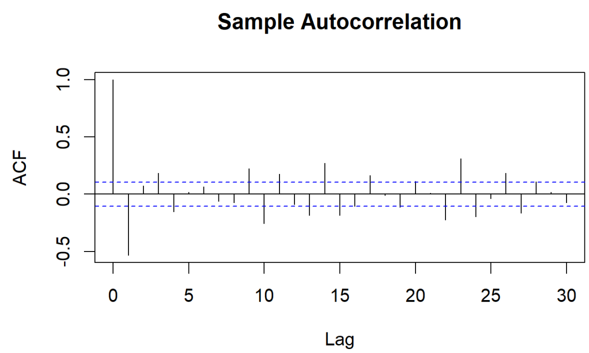
Figure 7: Predictions of MA models

3.3 Answer of Q3(c)

After taking the difference in this way, the data looks more stationary than before. The ACF shows $MA(1)$ is reasonable since all the lags larger than 1 are relative small.



(a) Difference



(b) ACF

3.4 Answer of Q3(d)

The estimation of $MA(1)$ is shown below.

	point estimate	standard error
$\hat{\mu}$	0.2200	2.7178
$\hat{\theta}$	-0.5423	0.0375

Table 9: Estimations

3.5 Answer of Q3(e)

The values are shown below.

	point estimate
$\hat{\phi}_0$	0.73596
$\hat{\phi}_1$	-0.66725
$\hat{\phi}_2$	-0.22924
$\hat{\phi}_3$	0.10759

Table 10: Estimations

3.6 Answer of Q3(f)

Replace $D_s = Y_s - Y_{s-1} - Y_{s-12} + Y_{s-13}$ to equation (6), we have

$$\begin{aligned} Y_t = &\hat{\phi}_0 + (\hat{\phi}_1 + 1)Y_{t-1} + (\hat{\phi}_2 - \hat{\phi}_1)Y_{t-2} + (\hat{\phi}_3 - \hat{\phi}_2)Y_{t-3} + (-\hat{\phi}_3)Y_{t-4} \\ &+ Y_{t-12} + (-1 - \hat{\phi}_1)Y_{t-13} + (\hat{\phi}_1 - \hat{\phi}_2)Y_{t-14} + (\hat{\phi}_2 - \hat{\phi}_3)Y_{t-15} + \hat{\phi}_3 Y_{t-16} + \epsilon_t \end{aligned}$$

The parameters are given by

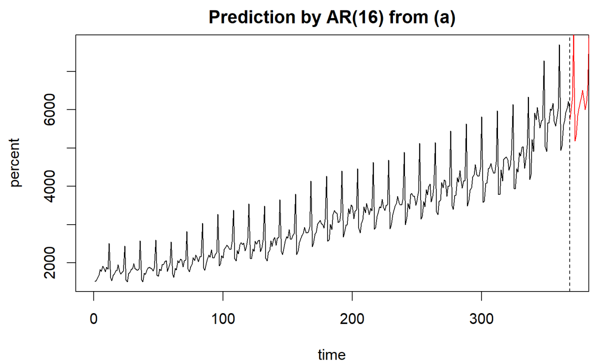
	point estimates from (e)	point estimates from (a)
constant term	0.7360	-6.5251
parameter before Y_{t-1}	0.3327	0.2932
parameter before Y_{t-2}	0.4380	0.4054
parameter before Y_{t-3}	0.3368	0.2977
parameter before Y_{t-4}	-0.1076	-0.1632
parameter before Y_{t-12}	1.0000	1.0137
parameter before Y_{t-13}	-0.3327	-0.3046
parameter before Y_{t-14}	-0.4380	-0.4107
parameter before Y_{t-15}	-0.3368	-0.3013
parameter before Y_{t-16}	0.1076	0.1858

Table 11: Estimations

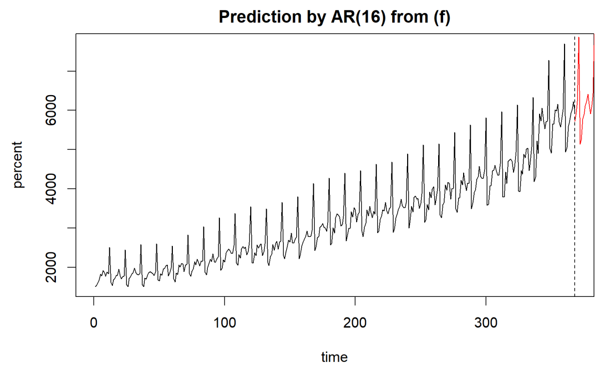
The results shows these two estimations are very similar except the intercept value.

3.7 Answer of Q3(g)

The predictions are quite similar for these two models



(a) Predictions from (a)



(b) Predictions from (f)

4 Question4

Consider the sunspots data that we looked at in class.
a) Plot the sample acf and pacf for this dataset. Based on these plots, argue that $AR(9)$ is an appropriate model for this dataset. (4 points)

b) Split this dataset by removing the last 40 datapoints and keeping them aside as a test dataset. The remaining observations will form the training dataset. Fit the $AR(p)$ model for $p = 1, 2, \ldots, 15$ as well as the $MA(q)$ model for $q = 1, 2, \ldots, 15$ to the training dataset. You can inbuilt R functions for fitting these models. Obtain predictions for each of these models for the future 40 datapoints and compare them to the actual observations in the test dataset. Which model performs best in terms of mean squared error of prediction? Compare the performance of the best model with the $AR(9)$ model (if they are different) obtained in the previous part. (8 points)

4.1 Answer of Q4(a)

We know that the PACF of an $AR(p)$ model satisfies that

$$\phi_{hh} = 0, \text{ for } h > p$$

According to the PACF below, we find that the PACF values greater than 9 are all very small. It implies $AR(9)$ is an appropriate model for this dataset.

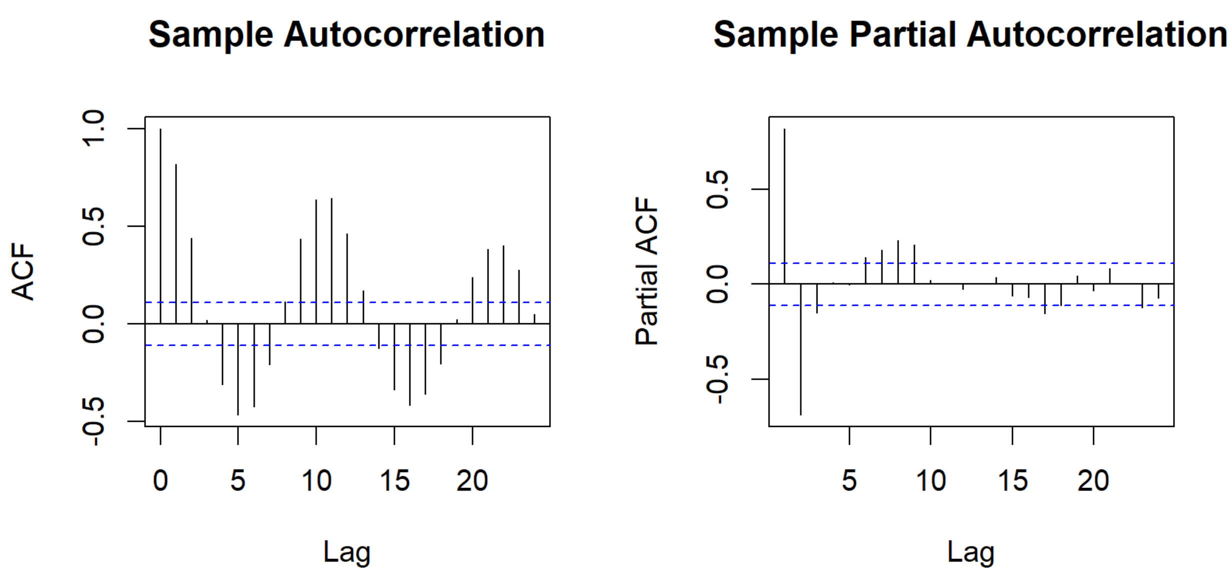


Figure 10: ACF and PACF

4.2 Answer of Q4(b)

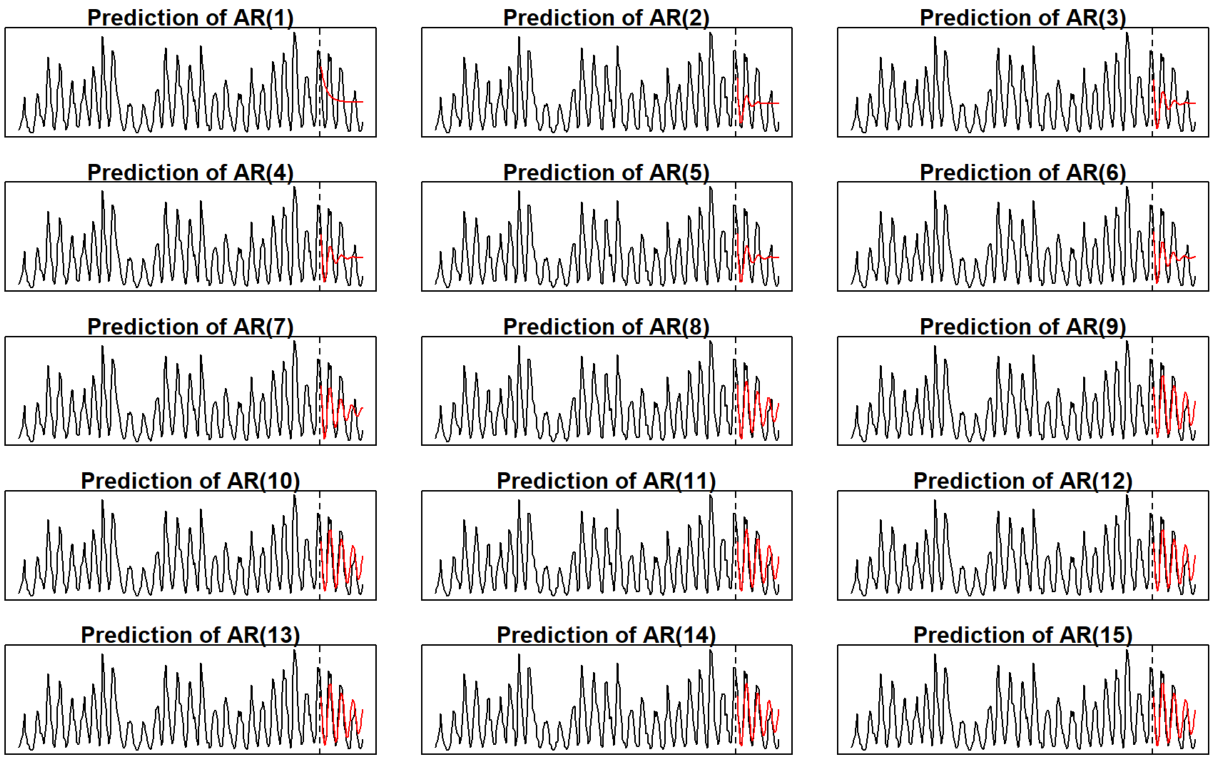


Figure 11: Predictions of AR models

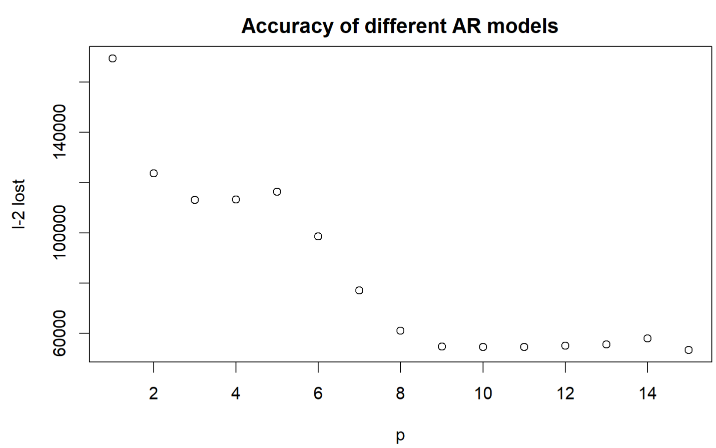


Figure 12: Accuracy of AR models

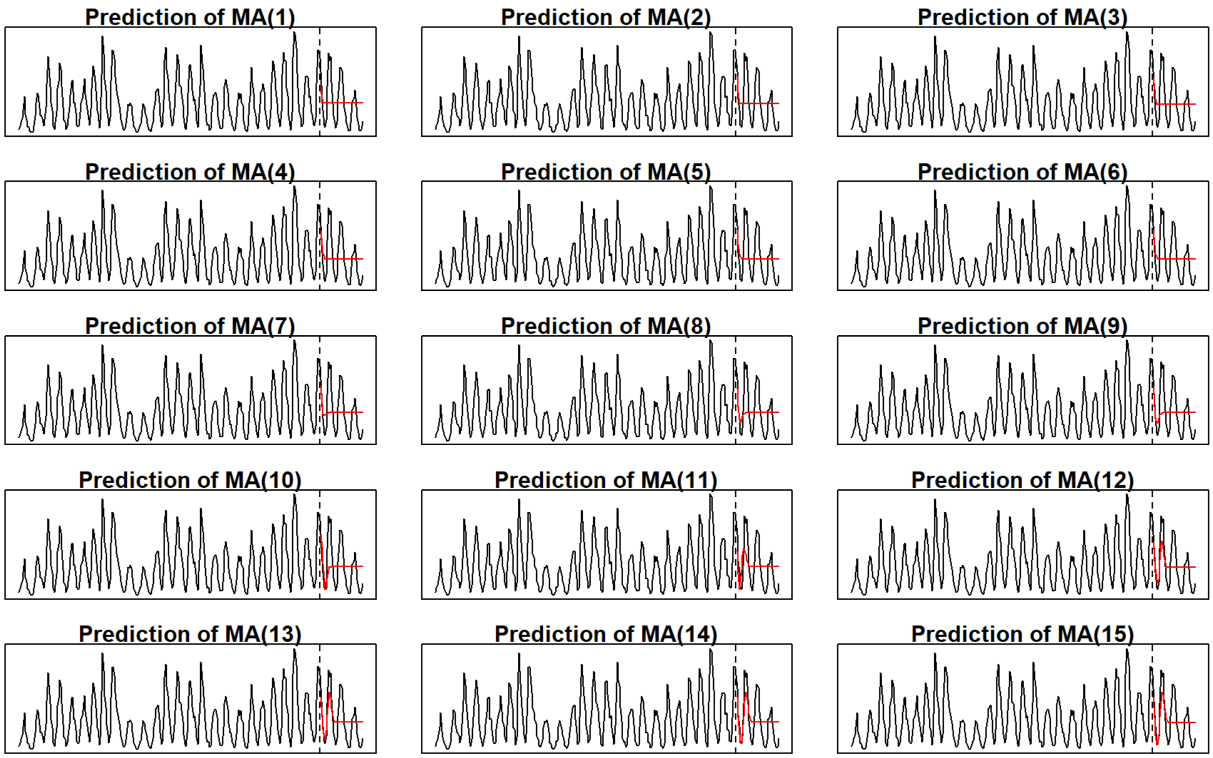


Figure 13: Predictions of MA models

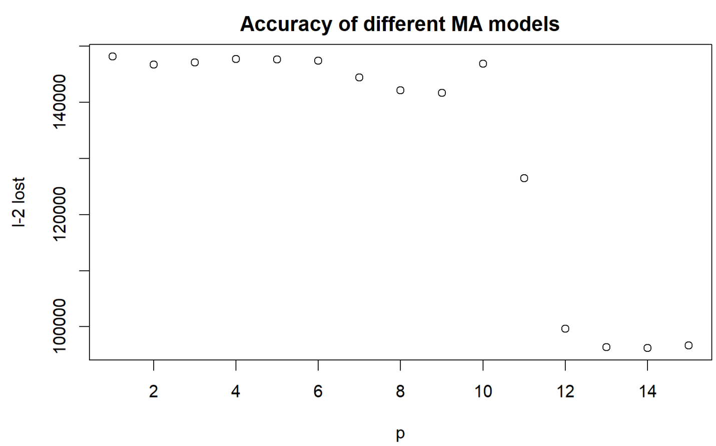


Figure 14: Accuracy of MA models

Comment: The $AR(15)$ model performs the best among all the models.The $AR(9)$ model is great as well. While the SSE of $AR(15)$ is 53251.34, the SSE of $AR(9)$ is 54699.88.

5 Question5

Let Y_1 and Y_2 be two uncorrelated random variables having mean zero and the same variance. Let $Y_3 := Y_1 + Y_2$.

a) What is the Best Linear Predictor (BLP) of Y_1 in terms of Y_3 ? (3 points)

b) What is the Best Linear Predictor (BLP) of Y_2 in terms of Y_3 ? (3 points)

c) What is the partial correlation $\rho_{Y_1,Y_2|Y_3}$?(3 points)

5.1 Answer of Q5(a)

We have these equations

$$\begin{cases} Cov(Y_1, Y_2) = 0 \\ EY_1 = EY_2 = 0 \\ Var(Y_1) = Var(Y_2) = \sigma^2 \\ Y_3 = Y_1 + Y_2 \end{cases}$$

From the class, we derived the formula of BLP when $p = 1$. The BLP of Y_1 in terms of Y_3 is given by

$$BLP = EY_1 + \frac{Cov(Y_1, Y_3)}{Var(Y_3)}(Y_3 - EY_3)$$

From the above equations, we have

$$\begin{aligned} Cov(Y_1, Y_3) &= Var(Y_1) + Cov(Y_1, Y_2) = \sigma^2 \\ Var(Y_3) &= Var(Y_1) + Var(Y_2) + 2Cov(Y_1, Y_2) = 2\sigma^2 \end{aligned}$$

Therefore, we have

$$BLP = \frac{\sigma^2}{2\sigma^2}Y_3 = \frac{1}{2}Y_3$$

5.2 Answer of Q5(b)

According to the symmetry of Y_1 and Y_2, we have

$$BLP = \frac{\sigma^2}{2\sigma^2} Y_3 = \frac{1}{2} Y_3$$

5.3 Answer of Q5(c)

By definition of partial correlation, we have

$$\rho_{Y_1, Y_2 | Y_3} = Corr(r_{Y_1|Y_3}, r_{Y_2|Y_3})$$

By definition of residual and results of (a) and (b), we have

$$r_{Y_1|Y_3} = Y_1 - \frac{1}{2} Y_3$$
$$r_{Y_2|Y_3} = Y_2 - \frac{1}{2} Y_3$$

Then, by definition of correlation, we have

$$\begin{aligned}
\rho_{Y_1, Y_2 | Y_3} &= \frac{Cov(Y_1 - \frac{1}{2}Y_3, Y_2 - \frac{1}{2}Y_3)}{\sqrt{Var(Y_1 - \frac{1}{2}Y_3)Var(Y_2 - \frac{1}{2}Y_3)}} \\
&= \frac{-\frac{1}{2}\sigma^2}{\frac{1}{2}\sigma^2} \\
&= -1
\end{aligned}$$

6 Question6

Let Y_1, Y_2 and ϵ be three uncorrelated random variables having mean zero and the same variance. Let $Y_3 := Y_1 + Y_2 + \epsilon$.

a) What is the Best Linear Predictor (BLP) of Y_1 in terms of Y_3 ? (3 points)

b) What is the Best Linear Predictor (BLP) of Y_2 in terms of Y_3 ? (3 points)

c) What is the partial correlation $\rho_{Y_1,Y_2|Y_3}$?(3 points)

6.1 Answer of Q6(a)

Similar to Q5(a), we have

$$BLP = EY_1 + \frac{Cov(Y_1, Y_1 + Y_2 + \epsilon)}{Var(Y_1 + Y_2 + \epsilon)}(Y_3 - EY_3) = \frac{1}{3}Y_3$$

6.2 Answer of Q6(b)

By symmetry, we have

$$BLP = \frac{1}{3}Y_3$$

6.3 Answer of Q6(c)

Similar to Q5(c), we have,

$$\begin{aligned} Cov(Y_1 - \frac{1}{2}Y_3, Y_2 - \frac{1}{2}Y_3) &= \frac{1}{4}\sigma^2 \\ Var(Y_1 - \frac{1}{2}Y_3) &= \frac{3}{4}\sigma^2 \end{aligned}$$

Therefore,

$$\rho_{Y_1,Y_2|Y_3} = \frac{\frac{1}{4}\sigma^2}{\frac{3}{4}\sigma^2} = \frac{1}{3}$$

7 Question7

Let Y be a 4×1 random vector with components Y_1, Y_2, Y_3 and Y_4. Suppose that each Y_i has mean zero. Suppose that the covariance matrix, Σ, of Y is given by

$$\Sigma = \mathrm{Cov}(Y) = \begin{pmatrix} 1 & 0.5 & 0 & 1 \\ 0.5 & 1.25 & 2 & -1.5 \\ 0 & 2 & 5 & -5 \\ 1 & -1.5 & -5 & 7 \end{pmatrix}$$

The inverse of Σ is given by

$$\Sigma^{-1} = \begin{pmatrix} 3.25 & -2.5 & 0 & -1 \\ -2.5 & 5 & -2 & 0 \\ 0 & -2 & 2 & 1 \\ -1 & 0 & 1 & 1 \end{pmatrix}$$

a) List all pairs (i, j) with $1 \leq i < j \leq 4$ such that the correlation between Y_i and Y_j is strictly positive. Give reasons for your answer. (2 points).

b) List all pairs (i, j) with $1 \leq i < j \leq 4$ such that the partial correlation between Y_i and Y_j given all the other Y_k 's equals zero. Give reasons for your answer. (2 points).

c) List all pairs (i, j) with $1 \leq i < j \leq 4$ such that the partial correlation between Y_i and Y_j given all the other Y_k 's is strictly positive. Give reasons for your answer. (3 points).

d) Let $\beta_0^* + \beta_1^* Y_1 + \beta_2^* Y_2 + \beta_3^* Y_3$ be the Best Linear Predictor of Y_4 in terms of Y_1, Y_2 and Y_3. For what values of $i \in \{0, 1, 2, 3\}$ is the coefficient β_i^* exactly zero? For what values of $i \in \{0, 1, 2, 3\}$ is the coefficient β_i^* strictly positive? Give reasons for your answers. ($2 + 2 = 4$ points).

e) What is the variance of the residual $r_{Y_4|Y_1,Y_2,Y_3}$? (2 points).

7.1 Answer of Q7(a)

By definition of covariance matrix, the positive entries imply positive correlation. Therefore, the following pairs have strictly positive correlation.

$$(1, 2), (1, 4), (2, 3)$$

7.2 Answer of Q7(b)

In the class, we have derived that

$$\rho_{Y_i, Y_j | Y_k, k \neq i, j} = \frac{-(\Sigma^{-1})(i, j)}{\sqrt{(\Sigma^{-1})(i, i) \cdot (\Sigma^{-1})(j, j)}}$$

Therefore, the following pairs have zero conditional partial correlation.

$$(1, 3), (2, 4)$$

7.3 Answer of Q7(c)

Similar to (b), the following pairs have positive conditional partial correlation.

$$(1,2),(1,4),(2,3)$$

7.4 Answer of Q7(d)

In the class, we have derived that

$$\beta_i^* = \rho_{Y,X_i|X_k,k\neq i}\sqrt{\frac{Var(r_{Y|X_k,k\neq i})}{Var(r_{X_i|X_k,k\neq i})}}$$

Therefore, we have

$$\beta_2^* \text{is exactly zero}$$
$$\beta_1^* \text{is strictly positive}$$

7.5 Answer of Q7(e)

Let's define

$$X = \begin{bmatrix} Y_1 \\ Y_2 \\ Y_3 \end{bmatrix}$$
$$Y = Y_4$$

Then, we have

$$Cov(X,Y) = \begin{bmatrix} Cov(X) & Cov(X,Y) \\ Cov(Y,X) & Var(Y) \end{bmatrix}$$

In the class, we have derived

$$\begin{aligned}
&Var(r_{Y_4|Y_1,Y_2,Y_3}) \\
&\qquad = Var(Y_4) - Cov(Y_4,X)(Cov(X))^{-1}Cov(X,Y_4) \\
&\qquad = 7 - \begin{bmatrix} 1 & -1.5 & -5 \end{bmatrix} \begin{bmatrix} 2.25 & -2.5 & 1 \\ -2.5 & 5 & -2 \\ 1 & -2 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ -1.5 \\ -5 \end{bmatrix} \\
&\qquad = 7 - 6 = 1
\end{aligned}$$

8 Question8

Suppose X_1, Z_2, Z_3, Z_4 are uncorrelated random variables having mean zero. Also suppose that X_1 has variance 1 while each of Z_2, Z_3, Z_4 has variance 3/4. Using these, we define new random variables X_2, X_3, X_4 via

$$X_2 = (-0.5)X_1 + Z_2, \quad X_3 = (-0.5)X_2 + Z_3 \quad \text{and} \quad X_4 = (-0.5)X_3 + Z_4.$$

a) What is the 3×3 covariance matrix of the 3×1 random vector with components X_1, X_2, X_3 ? (3 points)

b) What is the 3×3 covariance matrix of the 3×1 random vector with components X_2, X_3, X_4?(3 points)

c) What is the partial correlation between X_2 and X_4 given X_3 ? (3 points)

d) What is the partial correlation between X_1 and X_4 given X_2, X_3 ? (3 points)

e) What is the best linear predictor of X_4 in terms of X_1, X_2, X_3 ? (3 points)

8.1 Answer of Q8(a)

$$Cov(\begin{bmatrix} X_1 \\ X_2 \\ X_3 \end{bmatrix}) = \begin{bmatrix} Var(X_1) & Cov(X_1, X_2) & Cov(X_1, X_3) \\ Cov(X_1, X_2) & Var(X_2) & Cov(X_2, X_3) \\ Cov(X_1, X_3) & Cov(X_2, X_3) & Var(X_3) \end{bmatrix} = \begin{bmatrix} 1 & -\frac{1}{2} & \frac{1}{4} \\ -\frac{1}{2} & 1 & -\frac{1}{8} \\ \frac{1}{4} & -\frac{1}{8} & 1 \end{bmatrix}$$

8.2 Answer of Q8(b)

$$Cov(\begin{bmatrix} X_2 \\ X_3 \\ X_4 \end{bmatrix}) = \begin{bmatrix} Var(X_2) & Cov(X_2, X_3) & Cov(X_2, X_4) \\ Cov(X_2, X_3) & Var(X_3) & Cov(X_3, X_4) \\ Cov(X_2, X_4) & Cov(X_3, X_4) & Var(X_4) \end{bmatrix} = \begin{bmatrix} 1 & -\frac{1}{8} & \frac{1}{16} \\ -\frac{1}{8} & 1 & -\frac{1}{32} \\ \frac{1}{16} & -\frac{1}{32} & 1 \end{bmatrix}$$

8.3 Answer of Q8(c)

- Method 1

$$\rho_{X_2,X_4|X_3} = \frac{\rho_{X_2,X_4} - \rho_{X_2,X_3}\rho_{X_3,X_4}}{\sqrt{1-\rho^2_{X_2,X_3}}\sqrt{1-\rho^2_{X_3,X_4}}} \approx 0.059$$

- Method 2

  Using the Σ in 8(b)

$$\Sigma = \begin{bmatrix} 1 & -\frac{1}{8} & \frac{1}{16} \\ -\frac{1}{8} & 1 & -\frac{1}{32} \\ \frac{1}{16} & -\frac{1}{32} & 1 \end{bmatrix}$$

$$\rho_{X_2,X_4|X_3} = \frac{-(\Sigma)^{-1}(1,3)}{\sqrt{(\Sigma)^{-1}(1,1) \cdot (\Sigma)^{-1}(3,3)}} = \frac{0.0598}{\sqrt{1.0194 \cdot 1.0045}} \approx 0.059$$

8.4 Answer of Q8(d)

The covariance matrix of $X = \begin{bmatrix} X_1 \\ X_2 \\ X_3 \\ X_4 \end{bmatrix}$ is given by

$$\Sigma = \begin{bmatrix} 1 & -\frac{1}{2} & \frac{1}{4} & -\frac{1}{8} \\ -\frac{1}{2} & 1 & -\frac{1}{8} & \frac{1}{16} \\ \frac{1}{4} & -\frac{1}{8} & 1 & -\frac{1}{32} \\ -\frac{1}{8} & \frac{1}{16} & -\frac{1}{32} & 1 \end{bmatrix}$$

The inverse matrix is given by

$$\Sigma^{-1} = \begin{bmatrix} 1.4159 & 0.6667 & -0.2667 & 0.1270 \\ 0.6667 & 1.3333 & 0 & 0 \\ -0.2667 & 0 & 1.0667 & 0 \\ 0.1270 & 0 & 0 & 1.0159 \end{bmatrix}$$

Therefore, we have

$$\begin{aligned} \rho_{X_1,X_4|X_2,X_3} &= \frac{-(\Sigma)^{-1}(1,4)}{\sqrt{(\Sigma)^{-1}(1,1) \cdot (\Sigma)^{-1}(4,4)}} \\ &\approx -0.1059 \end{aligned}$$

8.5 Answer of Q8(e)

Let's define

$$X := \begin{bmatrix} X_1 \\ X_2 \\ X_3 \end{bmatrix}$$

The BLP is given by

$$\begin{aligned} BLP &= EX_4 + Cov(X_4, X)(Cov(X))^{-1}(X - EX) \\ &= \begin{bmatrix} -\frac{1}{8} & \frac{1}{16} & -\frac{1}{32} \end{bmatrix} \begin{bmatrix} 1.4 & 0.6667 & -0.2667 \\ 0.6667 & 1.3333 & 0 \\ -0.2667 & 0 & 1.0667 \end{bmatrix} \begin{bmatrix} X_1 \\ X_2 \\ X_3 \end{bmatrix} \\ &= -0.525X_1 + 16X_2 + 2.1X_3 \end{aligned}$$