Topics in Time Series (3)

Yisen Du (Eason)

1 Question1

1. Consider the dataset Iynx that is available in base R. This gives the annual numbers of lynx trappings for 1821-1934 in Canada. Type help(lynx) to learn more about the dataset.

a) Fit the $AR(2)$ model to the first 90 observations of this dataset. Report the estimates of ϕ_0, ϕ_1, ϕ_2 and σ along with uncertainty quantification. (3 points)

b) Write down an explicit formula for the predictions generated by your fitted $AR(2)$ model for Y_t for $t \geq 91$. (4 points)

c) Use your $AR(2)$ to predict the data from time points $t = 91, \ldots, 114$. Also compute the standard deviations corresponding to the accuracy of prediction. (4 points).

d) Compare your predictions with the actual values from the dataset. Comment on the accuracy of the predictions. (2 points)

1.1 Answer of Q1(a)

By definition of $AR(2)$, our model is

$$Y_t = \phi_0 + \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + Z_t,\ Z_t \overset{\text{iid}}{\sim} N(0, \sigma^2)$$

Since we use the first 90 data as observations, $n = 90$. Note that the model fitting of AR model is very similar to linear model, which the only difference is the design matrix. We can write the $AR(2)$ model in compact form

$$\boldsymbol{y} = X\boldsymbol{\beta} + \mathbf{z},\ where\ \boldsymbol{y} = \begin{bmatrix} y_3 \\ y_4 \\ \vdots \\ y_{90} \end{bmatrix}_{88\times 1},\ X = \begin{bmatrix} 1 & y_2 & y_1 \\ 1 & y_3 & y_2 \\ \vdots & \vdots & \vdots \\ 1 & y_{89} & y_{88} \end{bmatrix}_{88\times 3},\ \boldsymbol{\beta} = \begin{bmatrix} \phi_0 \\ \phi_1 \\ \phi_2 \end{bmatrix}_{3\times 1},\ \boldsymbol{z} = \begin{bmatrix} z_3 \\ z_4 \\ \vdots \\ z_{90} \end{bmatrix}_{88\times 1}$$

We assume the priors

$$\phi_0, \phi 1, \phi 2, log\sigma \overset{\text{iid}}{\sim} N(0, \sigma^2)$$

Using the same method in Bayesian linear regression, we could get the point estimate of $\boldsymbol{\beta}$ and σ are

$$\hat{\boldsymbol{\beta}} = (X^T X) X^{-1} \boldsymbol{y},\ \hat{\sigma} = \sqrt{\frac{||\boldsymbol{y} - X\hat{\boldsymbol{\beta}}||^2}{n - 2p - 1}}$$

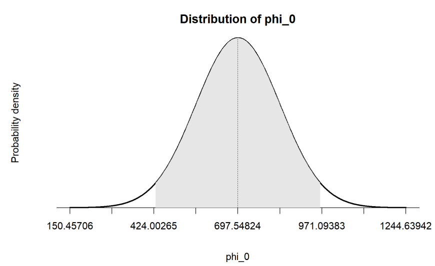
Also, the posterior distributions of $\boldsymbol{\beta}$ and $\frac{S(\hat{\boldsymbol{\beta}})}{\sigma^2}$ are

$$t_{n-2p-1}(\hat{\boldsymbol{\beta}}, \hat{\sigma^2}(X^TX)^{-1}),\ \chi^2_{n-2p-1},\ \textit{respectively}$$

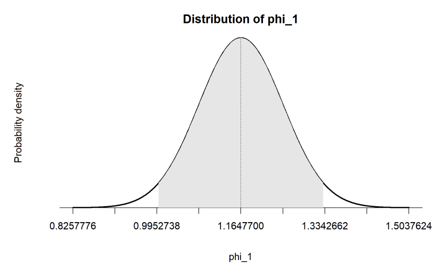
Since here $n - 2p - 1 = 85$, t distribution is approximately equal to normal distribution. We could use normal distribution to get 95% credential intervals of ϕ_0, ϕ_1, ϕ_2. To get uncertainty quantification of σ, we randomly draw samples from χ^2_{n-2p-1} and compute the realistic values of σ. Then we use the histogram of these realistic values to approximately find 95% confidential interval of σ. The results are shown in the following table and figures.

	point estimate	95% uncertainty interval
ϕ_0	697.55	[429.48, 965.62]
ϕ_1	1.16	[0.9986, 1.33]
ϕ_2	-0.62	[-0.79, -0.46]
σ	9.96	$[8.66, 11.71]$

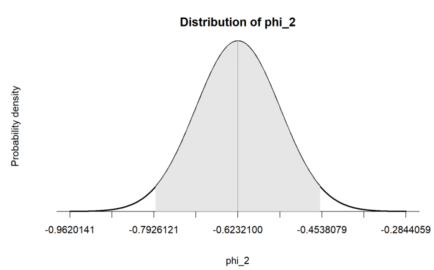
Table 1: Point estimate and uncertainty interval



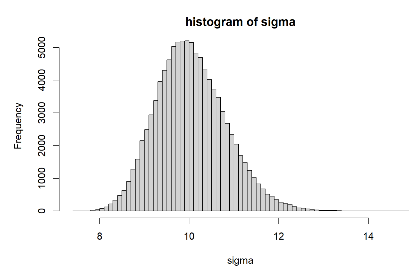
(a) Distribution of ϕ_0



(b) Distribution of ϕ_1



(a) Distribution of ϕ_2



(b) Distribution of σ

1.2 Answer of Q1(b)

To write the explicit formula for the predictions, firstly we write down the recurrent relation and then solve it to get analytical solution. The recurrent relation is given by

$$\hat{y_{90}} = \hat{\phi}_0 + \hat{\phi}_1 y_{90} + \hat{\phi}_2 y_{89}$$

$$\hat{y_{91}} = \hat{\phi}_0 + \hat{\phi}_1 \hat{y_{91}} + \hat{\phi}_2 y_{90}$$

$$......$$

let's define $v_t := y_t - \frac{\hat{\phi}_0}{1-\hat{\phi}_1-\hat{\phi}_2}$, then the recurrent relation for v_k is

$$v_k = \hat{\phi}_1 v_{k-1} + \hat{\phi}_2 v_{k-2}$$

The characteristic equation is given by

$$\phi(z) = 1 - \hat{\phi}_1 z - \hat{\phi}_2 z^2$$

From (a), $\hat{\phi}_1 \approx 1.16$ and $\hat{\phi}_2 \approx 0.62$. After solving the characteristic equation, the complex roots are $z \approx 1.27e^{\pm 0.743i}$. Therefore, the general solution for v_t is given by

$$v_t = c_1 1.27^{-t} \times cos(0.743t + c_2)$$

Note that c_1 and c_2 are two constants determined by initial conditions $v_{89} = 382 - \frac{\hat{\phi}_0}{1-\hat{\phi}_1-\hat{\phi}_2}$ and $v_{90} = 808 - \frac{\hat{\phi}_0}{1-\hat{\phi}_1-\hat{\phi}_2}$. The system is given by

$$\begin{cases} c_1(1.27)^{-89} cos(66.127 + c_2) = -1134.413 \\ c_1(1.27)^{-90} cos(66.87 + c_2) = -708.413 \end{cases}$$

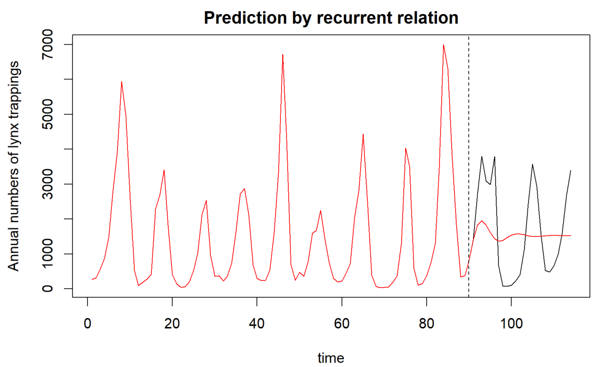
We could divide two equations to solve c_2 and then solve c_1. After solving c_1 and c_2, the prediction formula for Y_t is given by

$$y_t = (1.97 \times 10^{12}) 1.27^{-t} \times cos(0.743t - 0.237) + 1516.413$$

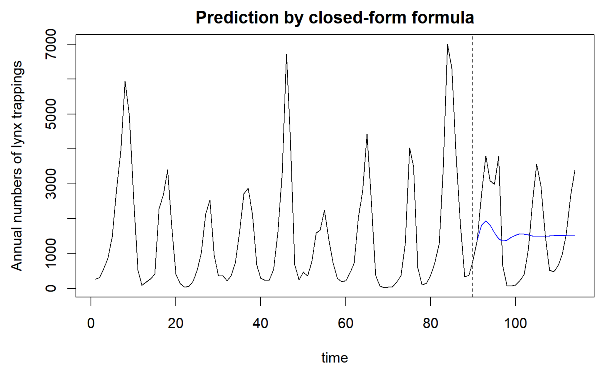
$$t = 91, 92, ..., 114$$

1.3 Answer of Q1(c)

We could use the formula in (b) to plot the prediction values (since there's some approximations, the results are not accurate). Also, we could directly use the recurrent relation to compute prediction values. The following figures show the prediction results are almost the same.



(a) Prediction by closed-form formula



(b) Prediction by recurrent relation

To compute prediction uncertainty, we use the method and algorithm proposed in lecture 15, the standard deviations are shown below.

```
 [1]  914.2592 1403.5246 1555.4971 1559.9234 1585.0549 1634.9047 1660.6794
 [8] 1663.1049 1664.8888 1671.2522 1675.6768 1676.4509 1676.5321 1677.3068
[15] 1678.0248 1678.2171 1678.2174 1678.3034 1678.4130 1678.4540 1678.4549
[22] 1678.4633 1678.4790 1678.4869
```

Figure 4: Prediction standard deviations

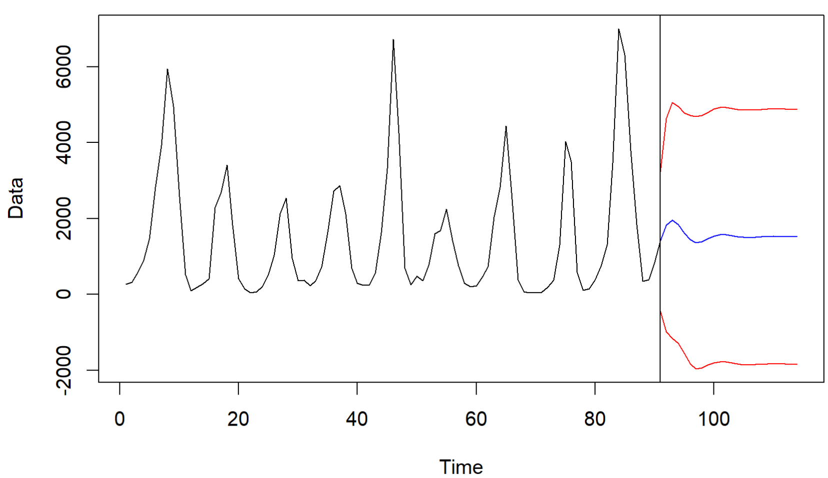
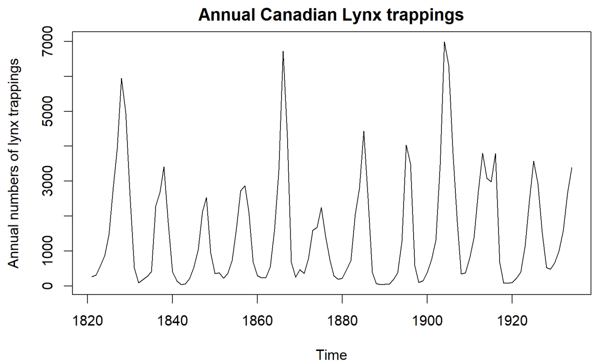
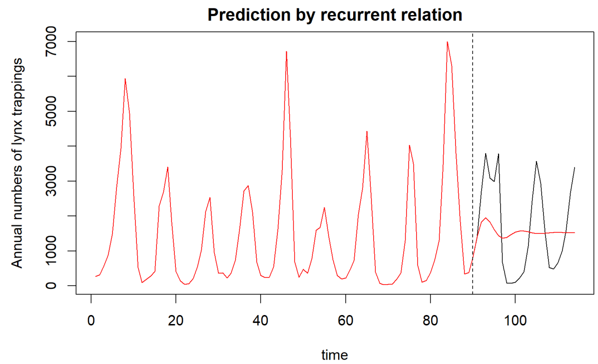


Figure 5: Prediction with uncertainty bands

1.4 Answer of Q1(d)



(a) Actual values



(b) Prediction

Comment: For the first few points, the predictions are relative accurate and the accuracy decreases then. Looking at the data, it seems like there's a periodicity. However, $AR(2)$ model is not sufficient to capture this pattern.

2 Question2

2. Consider the US population dataset from https://fred.stlouisfed.org/series/ POPTHM that we have worked with in class.

a) Fit an $AR(2)$ model to this dataset. Report the estimates of ϕ_1, ϕ_2 and σ along with uncertainty quantification (3 points).

b) Write down an explicit formula for the predictions generated by your fitted $AR(2)$ model for Y_t for the future months. (4 points)

c) Use your $AR(2)$ model to predict the data for 36 months immediately succeeding the last month in the dataset. Plot these predictions and uncertainty indicators along with the original data. Do these predictions make intuitive sense? (6) points)

d) Suppose that we want to predict the US population for the months preceding January 1959. For this purpose, fit the model:

$$Y_t = \alpha_0 + \alpha_1 Y_{t+1} + \alpha_2 Y_{t+2} + Z_t$$

Compare your fitted model with the forward model fitted earlier. Are there any similarities between the two models? (4 points)

e) Using your model from the previous part, predict the US population for the 36 months immediately preceding January 1959. Plot these predictions and uncer- tainty indicators along with the original data. Do these predictions make intuitive sense? (6 points).

2.1 Answer of Q2(a)

Using the same method in Q1, the point estimates and uncertainty quantification are shown below.

	point estimate	95% uncertainty interval
ϕ_0	12.99	[4.44, 21.54]
ϕ_1	1.95	[1.93, 1.974]
ϕ_2	-0.95	[-0.97, -0.93]
σ	0.78	[0.74, 0.82]

Table 2: Point estimate and uncertainty interval

2.2 Answer of Q2(b)

Using the same method in Q1, the difference here is that the roots of characteristic equation are real and distinct. By theorem, the general solution is given by

$$v_t = c_1 z_1^{-t} + c_2 z_2^{-t}$$

Using the last two data (populations in 202205 and 202206) as the initial values, we could solve c_1 and c_2. Finally, the explicit formula is given by

$$y_t = 334914.8 \times 1^{-t} - 2.32 \times 10^{20} \times 1.053^{-t} + 1298925$$

2.3 Answer of Q2(c)

The prediction standard deviation and prediction plots are shown below.

```
 [1]   16.75288   36.74020   60.34795   86.75067  115.37799  145.80942  177.71969
 [8]  210.84865  244.98322  279.94581  315.58641  351.77704  388.40761  425.38284
[15]  462.61985  500.04626  537.59870  575.22150  612.86572  650.48828  688.05123
[22]  725.52112  762.86851  800.06749  837.09530  873.93196  910.56001  946.96423
[29]  983.13136 1019.04998 1054.71026 1090.10382 1125.22356 1160.06358 1194.61902
[36] 1228.88596
```

Figure 7: Prediction standard deviation

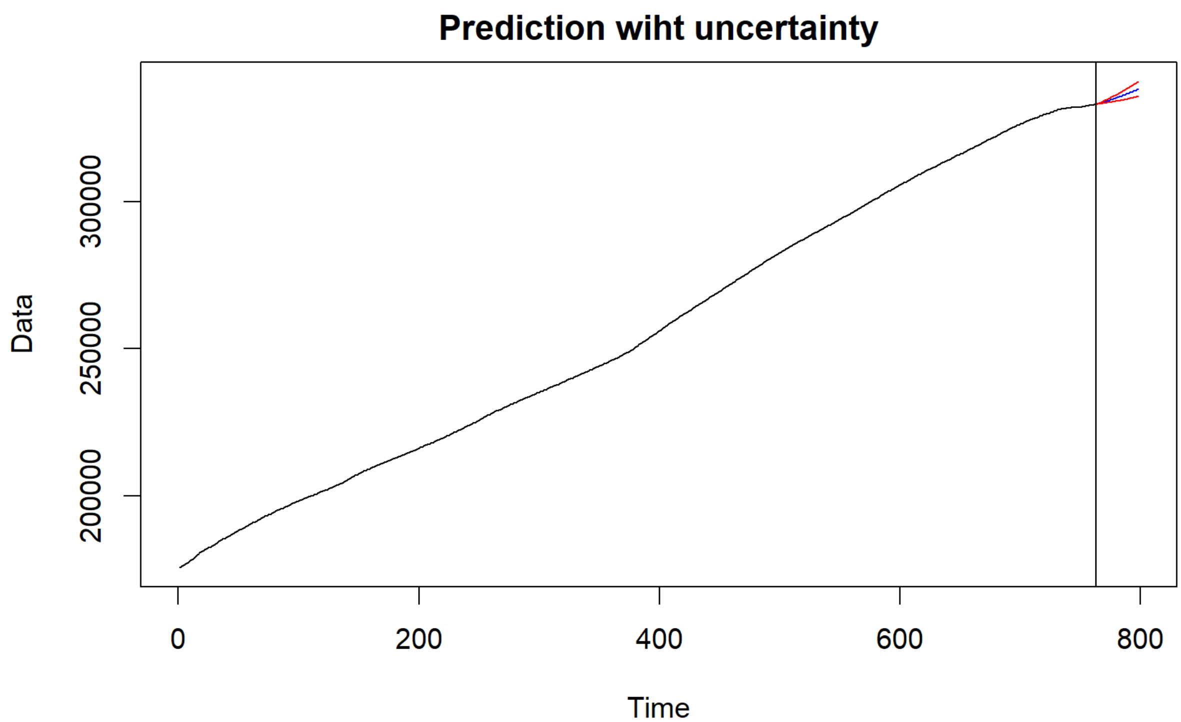


Figure 8: Prediction with uncertainty

Comment: It seems make sense since the prediction shows a increasing trend similar to the previous data.

2.4 Answer of Q2(d)

We could fit this model using linear regression with different design matrix. Compare the fitted parameters of backward model and forward model

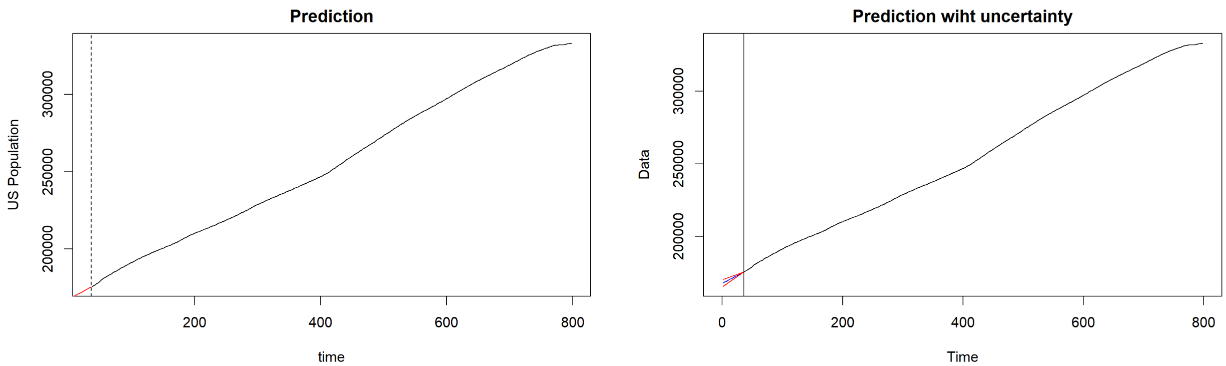
	fitted parameters for forward model	fitted parameters for backward model
ϕ_0	12.98925	-11.58669
ϕ_1	1.95181	1.94820
ϕ_2	-0.95182	-0.94820
σ	0.7803163	0.7795733

Table 3: Comparison between the two models

Comment: We could find the absolute values of parameters are very similar. The sign of ϕ_0 is different because one is forward model while the other is backward.

2.5 Answer of Q2(e)

Using the similar method in (b), the predictions are shown below.



(a) Prediction

(b) Prediction with uncertainty

Comment: It seems make sense since the prediction shows a increasing trend similar to the following data.

3 Question3

3. Consider again the US population dataset from https://fred.stlouisfed.org/series/ POPTHM that we have worked with in class.

a) Fit the $AR(3)$ model to this dataset and obtain point estimates of ϕ_1, ϕ_2, ϕ_3 and σ.(3 points).

b) For better interpretability, I want to fit a model to the twice differenced series:

$$D_t = (Y_t - Y_{t-1}) - (Y_{t-1} - Y_{t-2}) = Y_t - 2Y_{t-1} + Y_{t-2}.$$

Fit the $AR(1)$ model to D_t. Rewrite your fitted model:

$$D_t = \hat{\alpha}_0 + \hat{\alpha}_1 D_{t-1} + Z_t.$$

back in terms of Y_t by substituting $D_t = Y_t - 2Y_{t-1} + Y_{t-2}$ in the above equation. Compare this model to the $AR(3)$ model fitted in the previous part. Are they similar? (4 points)

c) Compare the predictions of the two models for the next 60 time points. Are the predictions similar? (4 points).

3.1 Answer of Q3(a)

Using the same method as before, the point estimates are shown below.

	Point estimate
ϕ_0	17.90
ϕ_1	2.36
ϕ_2	-1.78
ϕ_3	0.43
σ	0.74

Table 4: Point estimates

3.2 Answer of Q3(b)

Fitting the $AR(1)$ to D_t, the fitted parameters are $\hat{\alpha_0} \approx -0.0886$ and $\hat{\alpha_1} \approx 0.39$. Back in terms of Y_t, we could compute this model with $AR(3)$ model in (a).

	this model	$AR(3)$
ϕ_0	-0.0886	17.90
ϕ_1	2.39	2.36
ϕ_2	-1.78	-1.78
ϕ_3	0.39	0.43

Table 5: Comparison between this model and $AR(3)$

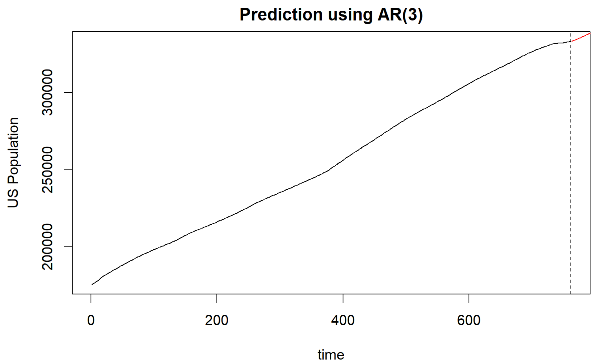
Comment: Only the constant term is quite different, the other parameters are quite similar for these two models. It implies the close relation between 'twice difference model' and $AR(3)$ model. Note that if we substitute D_t by Y_t, and compare it with $AR(3)$ model

$$Y_t = \alpha_0 + (\alpha_1 + 2)Y_{t-1} + (-2\alpha_1 - 1)Y_{t-2} + \alpha_1 Y_{t-3} + Z_t$$

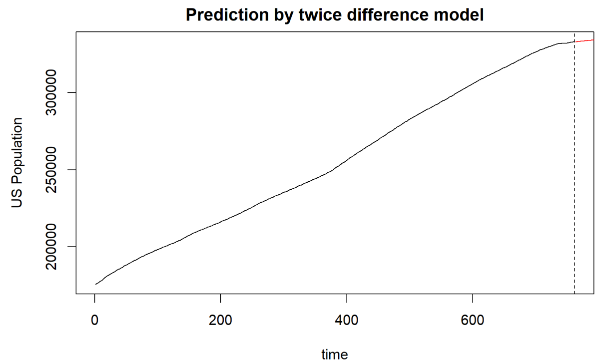
$$Y_t = \phi_0 + \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \phi_3 Y_{t-3} + Z_t$$

This model looks like a constrained version of $AR(3)$.

3.3 Answer of Q3(c)



(a) Prediction by $AR(3)$



(b) Prediction by twice difference model

Comment: Although they both show an increasing trend, prediction of $AR(3)$ is steeper than twice difference model.

4 Question4

4. 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. Separate the last 48 observations from this dataset and keep them as a test dataset. Fit the $AR(p)$ model for each $p = 1, 2, \ldots, 24$ for the training dataset and use it to predict the observations in the test dataset. Evaluate the 24 models based on the accuracy of prediction and report the model with the best prediction accuracy. (10 points).

4.1 Answer of Q4

The following figures shows the predictions of these 24 models. Using l-2 and l-1 lost to evaluate the accuracy of prediction, both methods shows $AR(15)$ is the best among all the models.

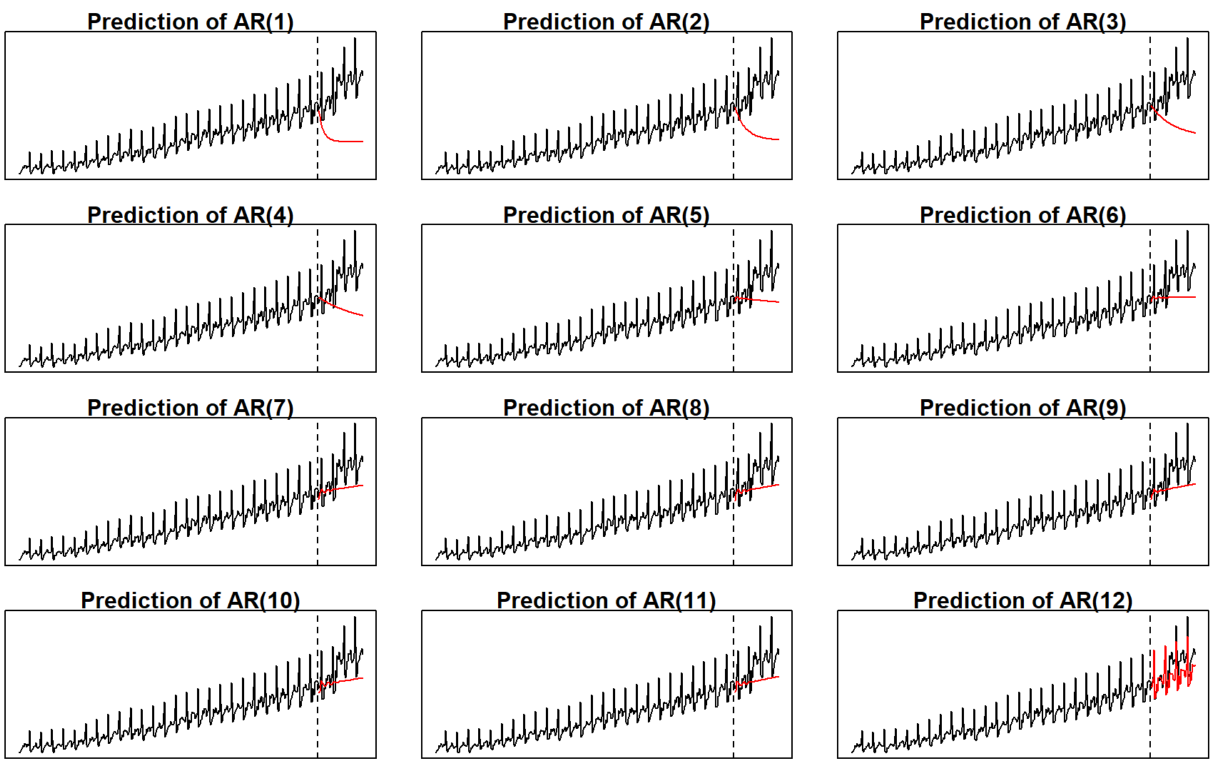


Figure 11: $AR(1)$ t0 $AR(12)$

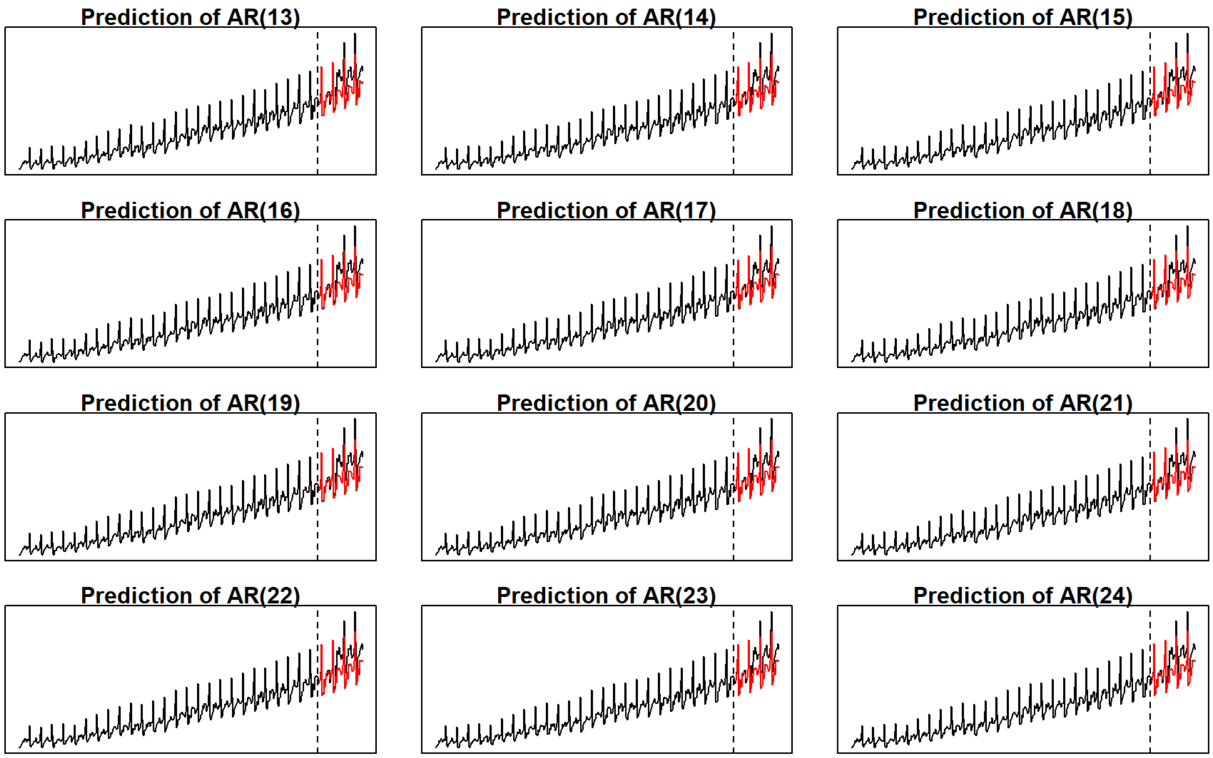
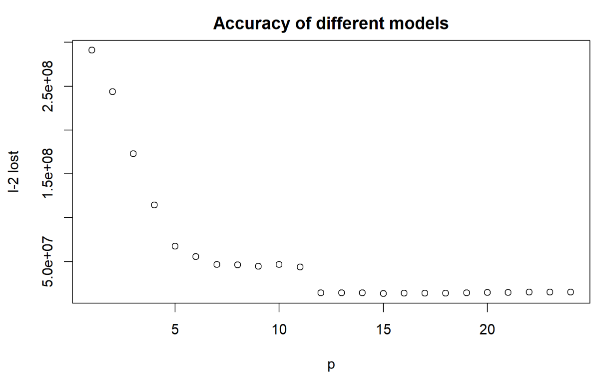
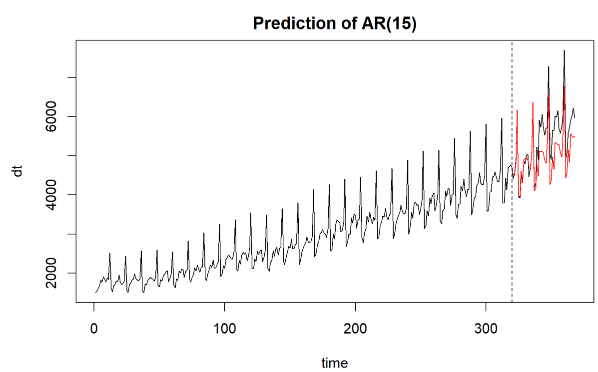


Figure 12: $AR(13)$ t0 $AR(24)$



(a) figure1



(b) figure2

5 Question5

5. We have seen in class that the $AR(1)$ model is said to belong to the causal stationary regime when $|\phi_1| < 1$. A similar characterization exists for the $AR(2)$ model. a) Show that the for the $AR(2)$ model, the roots of the characteristic equation are given by (2 points)

$$\frac{\phi_1 - \sqrt{\phi_1^2 + 4\phi_2}}{-2\phi_2} \text{ and } \frac{\phi_1 + \sqrt{\phi_1^2 + 4\phi_2}}{-2\phi_2}.$$

b) Show that both the roots have modulus strictly greater than 1 if and only if the pair (ϕ_1, ϕ_2) satisfy all the following three inequalities: (4 points)

$$\phi_2 + \phi_1 < 1 \quad \phi_2 - \phi_1 < 1 \quad |\phi_2| < 1.$$

5.1 Answer of Q5(a)

By definition of $AR(2)$ model,

$$Y_t = \phi_0 + \phi_1 Y_{k-1} + \phi_2 Y_{k-2}$$

Let $v_k = Y_k - \frac{\phi_0}{1-\phi_1-\phi_2}$, then $v_k = \phi_1 v_{k-1} + \phi_2 v_{k-2}$
The roots of characteristic equation is solved by

$$\phi(z) = 1 - \phi_1 z - \phi_2 z^2 = 0$$

The solutions are

$$z_1, z_2 = \frac{\phi_1 \pm \sqrt{\phi_1^2 + 4\phi_2}}{-2\phi_2}$$

Thus, proved.

5.2 Answer of Q5(b)

It's equivalent to show

$$|z| > 1 \iff \begin{cases} \phi_2 + \phi_1 < 1 \\ \phi_2 - \phi_1 < 1 \\ |\phi_2| < 1 \end{cases}$$

Note that z is the root of

$$1 - \phi_1 z - \phi_2 z^2 = 0 \tag{1}$$

Its reciprocal polynomial is given by

$$-\phi_2 - \phi_1 x + x^2 = 0 \tag{2}$$

The property of reciprocal polynomial guaranties that z is the root of (1) if and only if $\frac{1}{z}$ is the root of (2). Therefore, $|z| > 1 \iff |x| < 1$.

Then, we aim to show

$$|x| < 1 \iff \begin{cases} \phi_2 + \phi_1 < 1 \\ \phi_2 - \phi_1 < 1 \\ |\phi_2| < 1 \end{cases}$$

The roots of (2) is given by

$$x = \frac{\phi_1 \pm \sqrt{\phi_1^2 + 4\phi_2}}{2}$$

Consider the following three cases:

- (i) $\phi_1^2 + 4\phi_2 = 0 \Longrightarrow x = \frac{\phi_1}{2}$

  Then, we have

$$|x| < 1 \iff |\frac{\phi_1}{2}| < 2 \iff \phi_1^2 < 4 \iff \phi_2 > -1$$

- (ii) $\phi_1^2 + 4\phi_2 > 0 \Longrightarrow x_1, x_2 = \frac{\phi_1 \pm \sqrt{\phi_1^2 + 4\phi_2}}{2}$

  Then, we have

$$|x| < 1 \iff -1 < \frac{\phi_1 \pm \sqrt{\phi_1^2 + 4\phi_2}}{2} < 1 \iff \begin{cases} \phi_1^2 + 4\phi_2 < (2 - \phi_1)^2 \\ \phi_1^2 + 4\phi_2 < (2 + \phi_1)^2 \end{cases}$$

  After the simplification, we have,

$$\begin{cases} \phi_2 + \phi_1 < 1 \\ \phi_2 - \phi_1 < 1 \end{cases}$$

- (iii) $\phi_1^2 + 4\phi_2 < 0 \Longrightarrow x_1, x_2 = \frac{1}{2}\phi_1 \pm \frac{1}{2}\sqrt{-(\phi_1^2 + 4\phi_2)}i$

In this case,

$$|x| = \frac{1}{4}\phi_1^2 - \frac{1}{4}(\phi_1^2 + 4\phi_2) = -\phi_2$$

Then, we have

$$|x| < 1 \Longleftrightarrow \phi_2 > -1$$

Combining the result of these three cases, we have

$$|z| > 1 \Longleftrightarrow |x| < 1 \Longleftrightarrow \begin{cases} \phi_2 + \phi_1 < 1 \\ \phi_2 - \phi_1 < 1 \\ \phi_2 > -1 \end{cases}$$

Note that

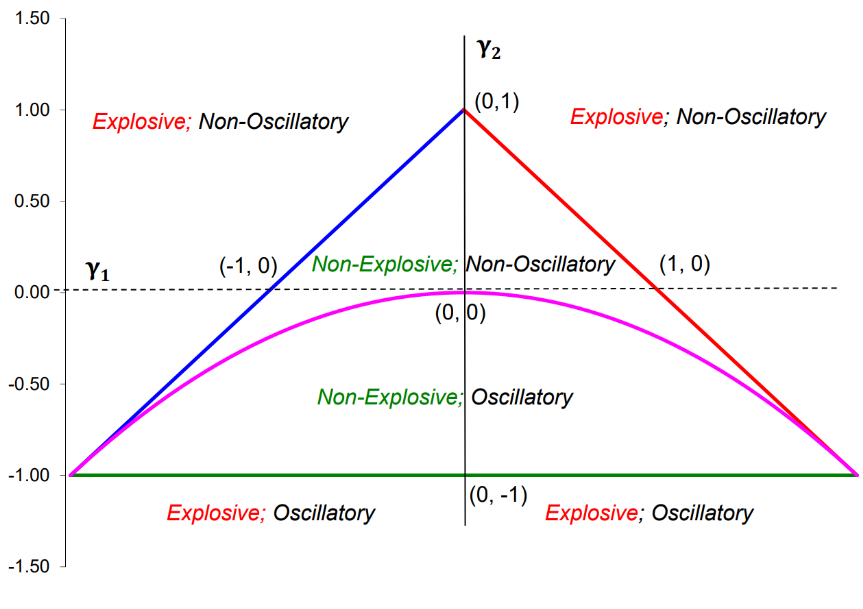
$$\begin{cases} \phi_2 + \phi_1 < 1 \\ \phi_2 - \phi_1 < 1 \end{cases} \Longleftrightarrow \phi_2 < 1$$

Therefore, we have

$$|z| > 1 \Longleftrightarrow \begin{cases} \phi_2 + \phi_1 < 1 \\ \phi_2 - \phi_1 < 1 \\ |\phi_2| < 1 \end{cases}$$

Proved.

Remark: In fact, the stationary region of $AR(2)$ is given by



Source: This diagram is based on Figure 7.1, on p.196 of A. Zellner, *An Introduction to Bayesian Inference in Econometrics*, Wiley, New York, 1971.

Figure 14: Stationarity requires that the roots have to be inside the red, blue and green triangle.