Roll No-12

M.sc. 3rd semester

Date of Assignment-21/12/2020

Date of Submission-12/01/2021

Experiment No -12

Topic- Fitting of auto regressive (AR) series to time series data.

Problem – Fit an auto regressive series of the form-

$$u_{t+2} + au_{t+1} + bu_t = \varepsilon_{t+2}$$

to the following data and then find the period of the fitted AR(2). Also fit an AR(3) considering the same data.

t	u_t
1	5.5933
2	-2.4558
3	-7.4251
4	-4.0501
5	5.1414
6	5.614
7	-0.0622
8	0.7227
9	0.7997
10	2.6251
11	-2.1324
12	-0.854
13	-0.7279
14	2.0284
15	7.967
16	6.1003
17	-1.0224
18	-2.5796
19	-2.7597
20	-5.755
21	-0.5003
22	3.533
23	2.6143
24	0.7917

Theory and Calculation-

Part-A- Fitting of AR(2)

The auto-regressive series of order k is given by-

$$u_t = \phi_1 u_{t-1} + \phi_2 u_{t-2} + ... + \phi_k u_{t-k} + \varepsilon_t \quad \text{where } \varepsilon_t \sim N\left(0, \sigma^2\right)$$

Now, taking the deviation of the observations u_t from their mean $\bar{u}$ and replacing the $u_t's$ with these deviations, the series becomes-

$$u'_t = \phi_1 u'_{t-1} + \phi_2 u'_{t-2} + ... + \phi_k u'_{t-k} + \varepsilon_t$$

When k=2, the AR series of order k reduces to-

$$u'_t = \phi_1 u'_{t-1} + \phi_2 u'_{t-2} + \varepsilon_t \qquad \text{----(1)}$$

Multiplying (1) by u'_{t-k}, we get-

$$u'_t u'_{t-k} = \phi_1 u'_{t-1} u'_{t-k} + \phi_2 u'_{t-2} u'_{t-k} + \varepsilon_t u'_{t-k} \qquad \text{----(2)}$$

Taking expectation on both sides of equation (2), we get-

$$E(u'_t u'_{t-k}) = \phi_1 E(u'_{t-1} u'_{t-k}) + \phi_2 E(u'_{t-2} u'_{t-k}) + E(\varepsilon_t u'_{t-k})$$

$$\Rightarrow C_k = \phi_1 C_{k-1} + \phi_2 C_{k-2} \qquad \text{----(3)}$$

where C_k represents the auto-covariance function at lag k.

Dividing both sides by $C_0 = E(u_t{}^2)$, we get from (3)

$$\rho_k = \phi_1 \rho_{k-1} + \phi_2 \rho_{k-2} \qquad \text{----(4)}$$

where ρ_k represents the auto-correlation at lag k.

In practice, ρ_k is determined by the sample counter-part of ρ_k i.e. r_k. Putting k=1 and k=2 in equation (4), we get-

$$\rho_1 = \phi_1 + \phi_2 \rho_1 \qquad [\because \rho_0 = 1 \quad \& \quad \rho_{-1} = \rho_1]$$

$$\rho_2 = \phi_1 \rho_1 + \phi_2$$

Replacing ρ_1 & ρ_2 by r_1 & r_2 respectively, we get-

$r_1 = \phi_1 + \phi_2 r_1$

$r_2 = \phi_1 r_1 + \phi_2$

$$\Rightarrow \begin{pmatrix} r_1 \\ r_2 \end{pmatrix} = \begin{pmatrix} 1 & r_1 \\ r_1 & 1 \end{pmatrix} \begin{pmatrix} \phi_1 \\ \phi_2 \end{pmatrix}$$

$$\Rightarrow \begin{pmatrix} \phi_1 \\ \phi_2 \end{pmatrix} = \begin{pmatrix} 1 & r_1 \\ r_1 & 1 \end{pmatrix}^{-1} \begin{pmatrix} r_1 \\ r_2 \end{pmatrix} \qquad \text{----(5)}$$

Solving (5), we get the estimate of the coefficients $\phi_i 's$. Therefore, the fitted AR series of order 2 is-

$$\hat{u}'_t = \hat{\phi}_1 u'_{t-1} + \hat{\phi}_2 u'_{t-2}$$

The period of the fitted second order AR series is given by-

$$Period = \frac{2\pi}{\theta}, \quad \text{where } \theta = \cos^{-1}\left(\frac{\hat{\phi}_1}{2\sqrt{|\hat{\phi}_2|}} \right)$$

Finally, we have $\hat{u}'_t = \hat{\phi}_1 u'_{t-1} + \hat{\phi}_2 u'_{t-2}$

$\therefore \hat{u}_t = \hat{u}'_t + \bar{u}$

Calculation-

we construct the following table-

t	Est_ut
1	0
2	0
3	-3.99353
4	-2.50595
5	2.239708
6	5.838234
7	1.0494
8	-2.58308
9	1.013877
10	0.626706
11	1.668575
12	-2.16431
13	1.219092
14	0.588897
15	2.156666
16	4.164123
17	-0.22018
18	-3.42169
19	-0.41818
20	0.333706
21	-1.34625
22	3.427244
23	2.924907
24	0.154608

Programming in R

```
library(readxl)

df_1 = read_excel("Autocorrelation_2.xlsx")

View(df_1)

ut = c(5.5933,-2.4558,-7.4251,-4.0501,5.1414,5.614,-0.0622,0.7227,0.7997,2.6251,-2.1324,-
0.854,-0.7279,2.0284,7.967,6.1003,-1.0224,-2.5796,-2.7597,-5.755,-
0.5003,3.533,2.6143,0.7917)

ut

N = length(ut)

N

ut_bar = mean(ut)

ut_bar

c0 = sum((ut-ut_bar)^2)/N

c1 = (sum((ut[1:23]-ut_bar)*(ut[2:24]-ut_bar)))/N

c2 = (sum((ut[1:22]-ut_bar)*(ut[3:24]-ut_bar)))/N

r1 = c1/c0

r2 = c2/c0

A = array(c(r1,r2),dim=c(2,1))

B = array(c(1,r1,r1,1),dim=c(2,2))

coeff = solve(B)%*%A

coeff

theta = acos(coeff[1,1]/(2*sqrt(abs(coeff[2,1]))))

period = (2*pi)/theta

period

f_term = coeff[1,1]*ut[2:23]

s_term = coeff[2,1]*ut[1:22]
```

```
utp_est = f_term+s_term

ut_est = utp_est+ut_bar

ut_est

df_a = append(0,ut_est)

df_2 = append(0,df_a)

View(df_2)

Data = cbind(df_1,df_2)

Data

#using ggplot to plot the graph

library(ggplot2)

ggp = ggplot(NULL, mapping = aes(t)) +
        geom_point(data = Data, mapping = aes(y=ut), col = "black") + geom_line(data = Data,
mapping = aes(y=ut), col = "orange", size = 1) +
        geom_point(data = Data, mapping = aes(y=df_2), col = "blue") + geom_line(data = Data,
mapping = aes(y=df_2), col = "green", size = 1) +
 labs(
  title = paste("Plotting of observed ut and estimated ut against t"),
  subtitle = paste("orange_line=observed ut and green_line=estimated ut"),
  caption = "Data from Model",
  x = "t",
  y = "observed_ut_and_estimated_ut"
  )
 ggp
```

Graph plotting

We plot the graph of observed and estimated values for the second order series considering t along X-axis and u_t & $\hat{u}_t$ along Y-axis by Using ggplot2-

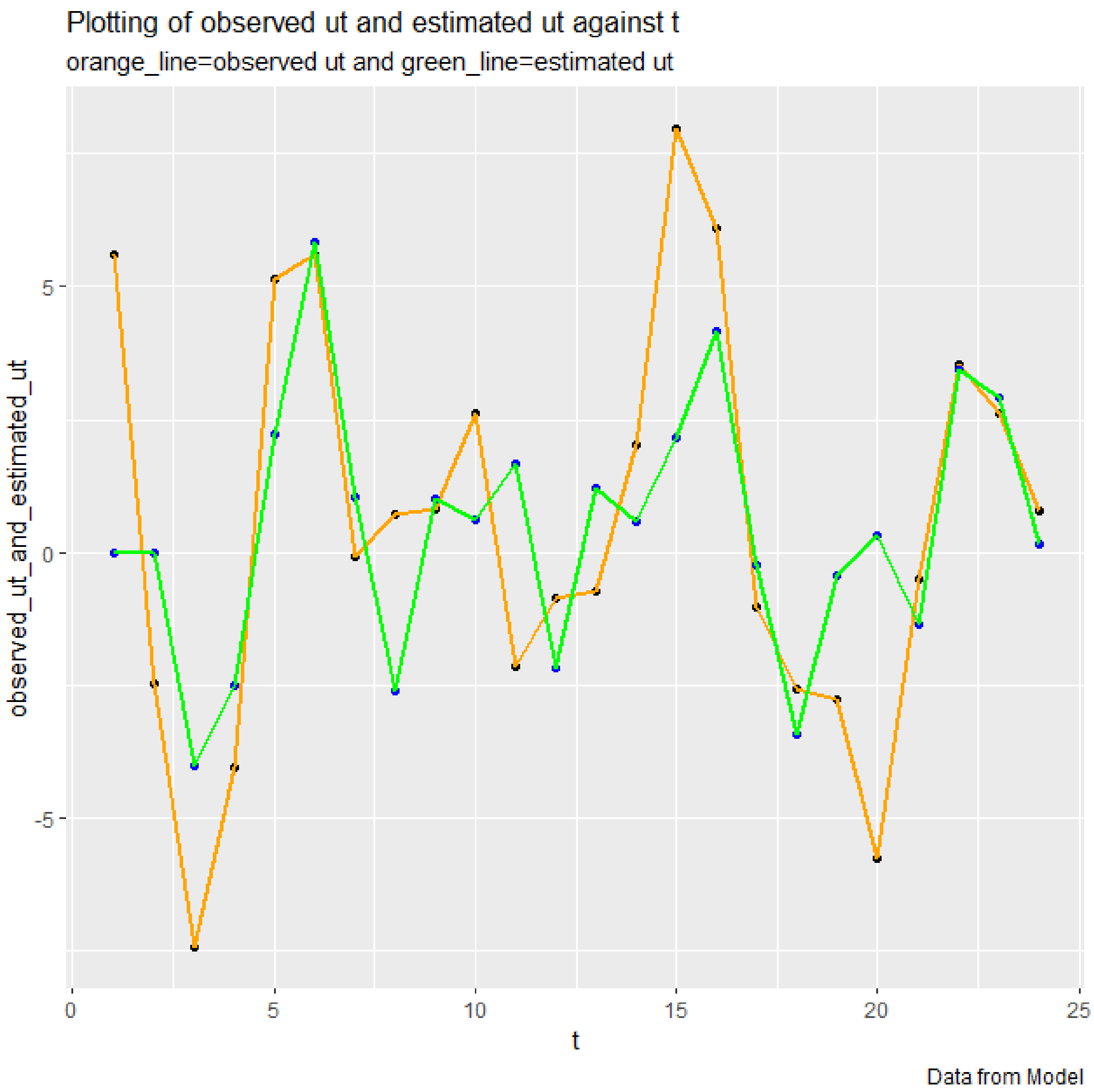
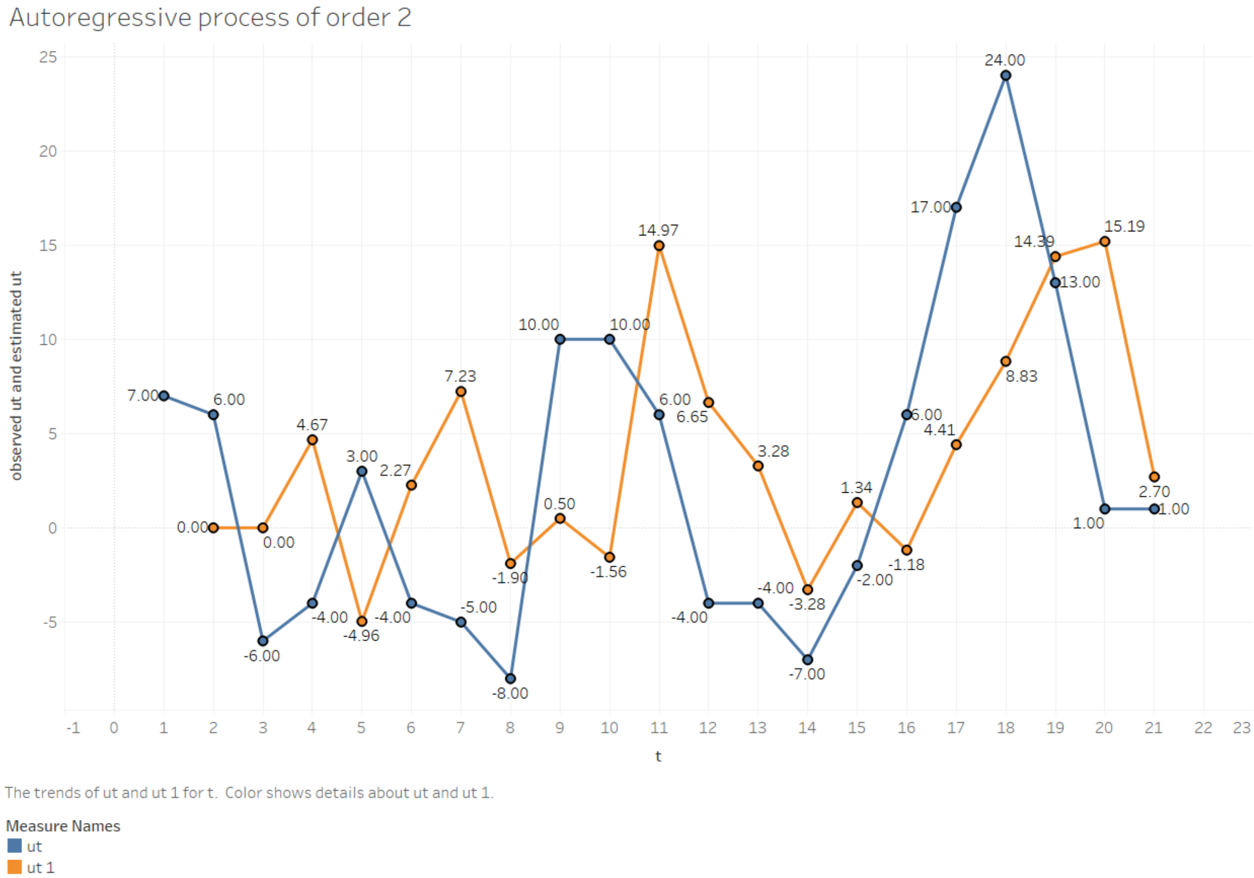


Tableau-



Conclusion-

The fitted AR(2) is-

$$\hat{u}_t = (0.5940278)u_{t-1} + (-0.5515499)u_{t-2}$$

The period of the fitted AR series of order 2 is <u>5.41955</u>.

Part-B-Fitting of AR(3)

The auto-regressive series of order k is given by-

$$u_t = \phi_1 u_{t-1} + \phi_2 u_{t-2} + ... + \phi_k u_{t-k} + \varepsilon_t \quad \text{where } \varepsilon_t \sim N(0, \sigma^2)$$

Now, taking the deviation of the observations u_t from their mean $\bar{u}$ and replacing the $u_t's$ with these deviations, the series becomes-

$$u'_t = \phi_1 u'_{t-1} + \phi_2 u'_{t-2} + ... + \phi_k u'_{t-k} + \varepsilon_t$$

when k=3, the AR series reduces to-

$$u'_t = \phi_1 u'_{t-1} + \phi_2 u'_{t-2} + \phi_3 u'_{t-3} + \varepsilon_t \qquad \text{----(1)}$$

Multiplying both sides of (1) by u'_{t-k} and taking expectation, we get-

$$E(u'_t u'_{t-k}) = \phi_1 E(u'_{t-1} u'_{t-k}) + \phi_2 E(u'_{t-2} u'_{t-k}) + \phi_3 E(u'_{t-3} u'_{t-k}) + E(\varepsilon_t u'_{t-k})$$

$$\Rightarrow C_k = \phi_1 C_{k-1} + \phi_2 C_{k-2} + \phi_3 C_{k-3} \qquad \text{----(2)}$$

Dividing both sides of equation (2) by C_0, we get-

$$\rho_k = \phi_1 \rho_{k-1} + \phi_2 \rho_{k-2} + \phi_3 \rho_{k-3} \qquad \text{----(3)}$$

In practice, ρ_k is determined by the sample counter-part of ρ_k i.e. r_k. Putting k=1,k=2 and k=3 in equation (3), we get-

$$\rho_1 = \phi_1 + \phi_2 \rho_1 + \phi_3 \rho_2 \qquad [\because \rho_0 = 1 \quad \& \quad \rho_{-1} = \rho_1]$$

$$\rho_2 = \phi_1 \rho_1 + \phi_2 + \phi_3 \rho_1$$

$$\rho_3 = \phi_1 \rho_2 + \phi_2 \rho_1 + \phi_3$$

Replacing ρ_1, ρ_2 & ρ_3 by r_1, r_2 & r_3 respectively, we get-

$$r_1 = \phi_1 + \phi_2 r_1 + \phi_3 r_2$$

$$r_2 = \phi_1 r_1 + \phi_2 + \phi_3 r_1$$

$$r_3 = \phi_1 r_2 + \phi_2 r_1 + \phi_3$$

$$\Rightarrow \begin{pmatrix} r_1 \\ r_2 \\ r_3 \end{pmatrix} = \begin{pmatrix} 1 & r_1 & r_2 \\ r_1 & 1 & r_1 \\ r_2 & r_1 & 1 \end{pmatrix} \begin{pmatrix} \phi_1 \\ \phi_2 \\ \phi_3 \end{pmatrix}$$

$$\Rightarrow \begin{pmatrix} \phi_1 \\ \phi_2 \\ \phi_3 \end{pmatrix} = \begin{pmatrix} 1 & r_1 & r_2 \\ r_1 & 1 & r_1 \\ r_2 & r_1 & 1 \end{pmatrix}^{-1} \begin{pmatrix} r_1 \\ r_2 \\ r_3 \end{pmatrix} \qquad \text{----(4)}$$

Solving (4), we get the estimate of the coefficients $\phi_i 's$. Therefore, the fitted AR series of order 3 is-

$$\hat{u}'_t = \hat{\phi}_1 u'_{t-1} + \hat{\phi}_2 u'_{t-2} + \hat{\phi}_3 u'_{t-3}$$

Finally, we have $\hat{u}'_t = \hat{\phi}_1 u'_{t-1} + \hat{\phi}_2 u'_{t-2} + \hat{\phi}_3 u'_{t-3}$

$$\therefore \hat{u}_t = \hat{u}'_t + \bar{u}$$

Calculation-

we construct the following table-

t	Est_ut
1	0
2	0
3	0
4	-2.68617
5	2.256708
6	5.971319
7	1.293682
8	-2.6911
9	0.645151
10	0.629779
11	1.56523
12	-2.04632
13	1.010592
14	0.712419
15	2.114174
16	4.014099
17	-0.26044
18	-3.65204
19	-0.7403
20	0.395398
21	-1.09547
22	3.403901
23	3.138796
24	0.225133

Programming in R

```
library(readxl)

df_1 = read_excel("Autocorrelation_2.xlsx")

View(df_1)

ut = c(5.5933,-2.4558,-7.4251,-4.0501,5.1414,5.614,-0.0622,0.7227,0.7997,2.6251,-
2.1324,-0.854,-0.7279,2.0284,7.967,6.1003,-1.0224,-2.5796,-2.7597,-5.755,-
0.5003,3.533,2.6143,0.7917)

ut

N = length(ut)

N

ut_bar = mean(ut)

ut_bar

c0=sum((ut-ut_bar)^2)/N

c1=(sum((ut[1:23]-ut_bar)*(ut[2:24]-ut_bar)))/N

c2=(sum((ut[1:22]-ut_bar)*(ut[3:24]-ut_bar)))/N

c3=(sum((ut[1:21]-ut_bar)*(ut[4:24]-ut_bar)))/N

r1=c1/c0

r2=c2/c0

r3=c3/c0

A=array(c(r1,r2,r3),dim=c(3,1))

B=array(c(1,r1,r2,r1,1,r1,r2,r1,1),dim=c(3,3))

coeff=solve(B)%*%A

coeff

f_term=coeff[1,1]*ut[3:23]

s_term=coeff[2,1]*ut[2:22]

t_term=coeff[3,1]*ut[1:21]
```

```
utp_est=f_term+s_term+t_term

ut_est=utp_est+ut_bar

ut_est

df_a = append(0,ut_est)

df_b = append(0,df_a)

df_2 = append(0,df_b)

View(df_2)

Data = cbind(df_1,df_2)

Data

#using ggplot to plot the graph

library(ggplot2)

ggp = ggplot(NULL, mapping = aes(t)) +

    geom_point(data = Data, mapping = aes(y=ut), col = "black") + geom_line(data = Data, mapping = aes(y=ut), col = "orange", size = 1) +

    geom_point(data = Data, mapping = aes(y=df_2), col = "blue") + geom_line(data = Data, mapping = aes(y=df_2), col = "green", size = 1) +

 labs(

  title = paste("Plotting of observed ut and estimated ut against t"),

  subtitle = paste("orange_line=observed ut and green_line=estimated ut"),

  caption = "Data from Model",

  x = "t",

  y = "observed_ut_and_estimated_ut"

  )

 ggp
```

Graph plotting

We plot the graph of observed and estimated values for the third order series considering t along X-axis and u_t & $\hat{u}_t$ along Y-axis by using ggplot 2-

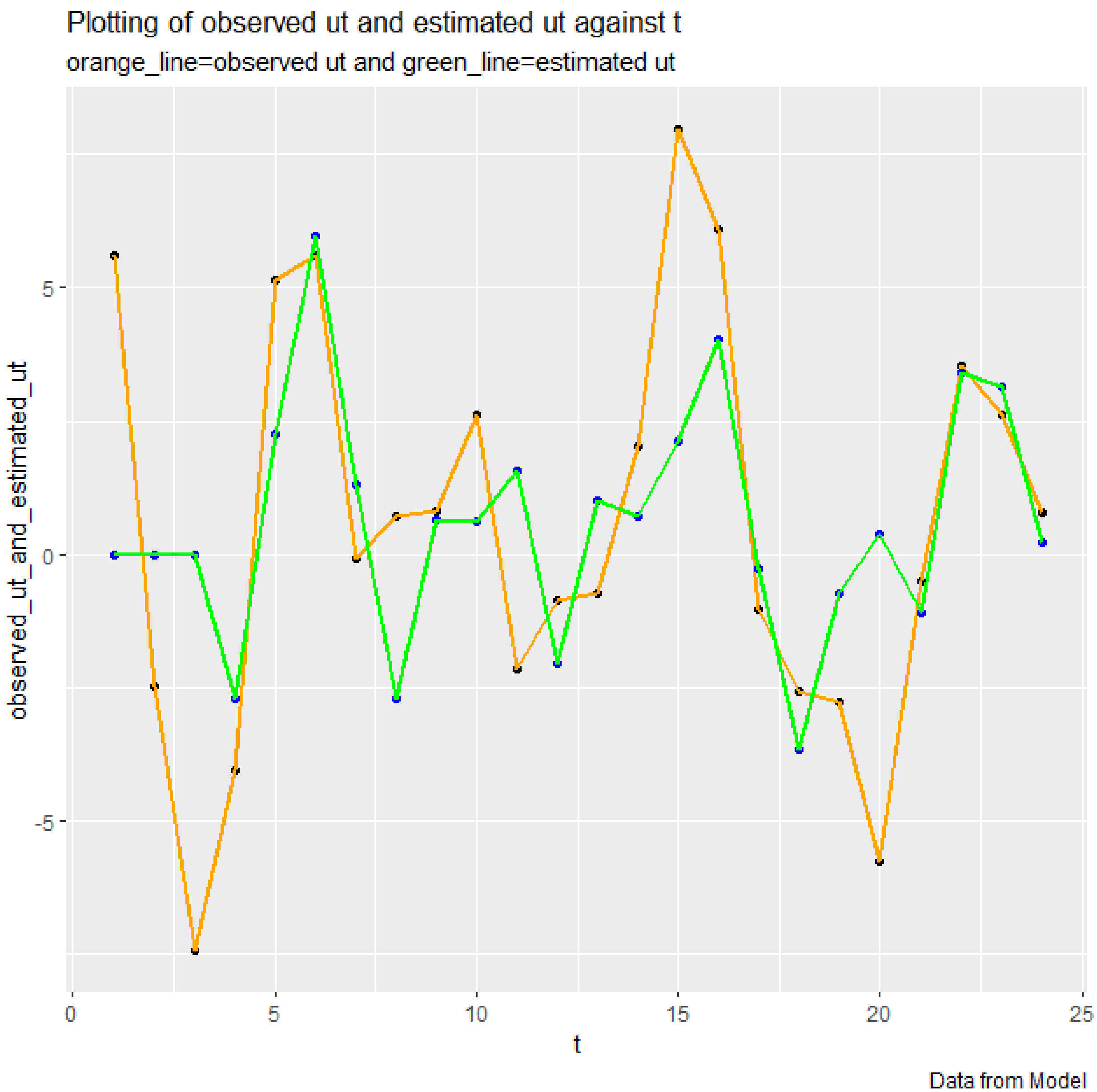
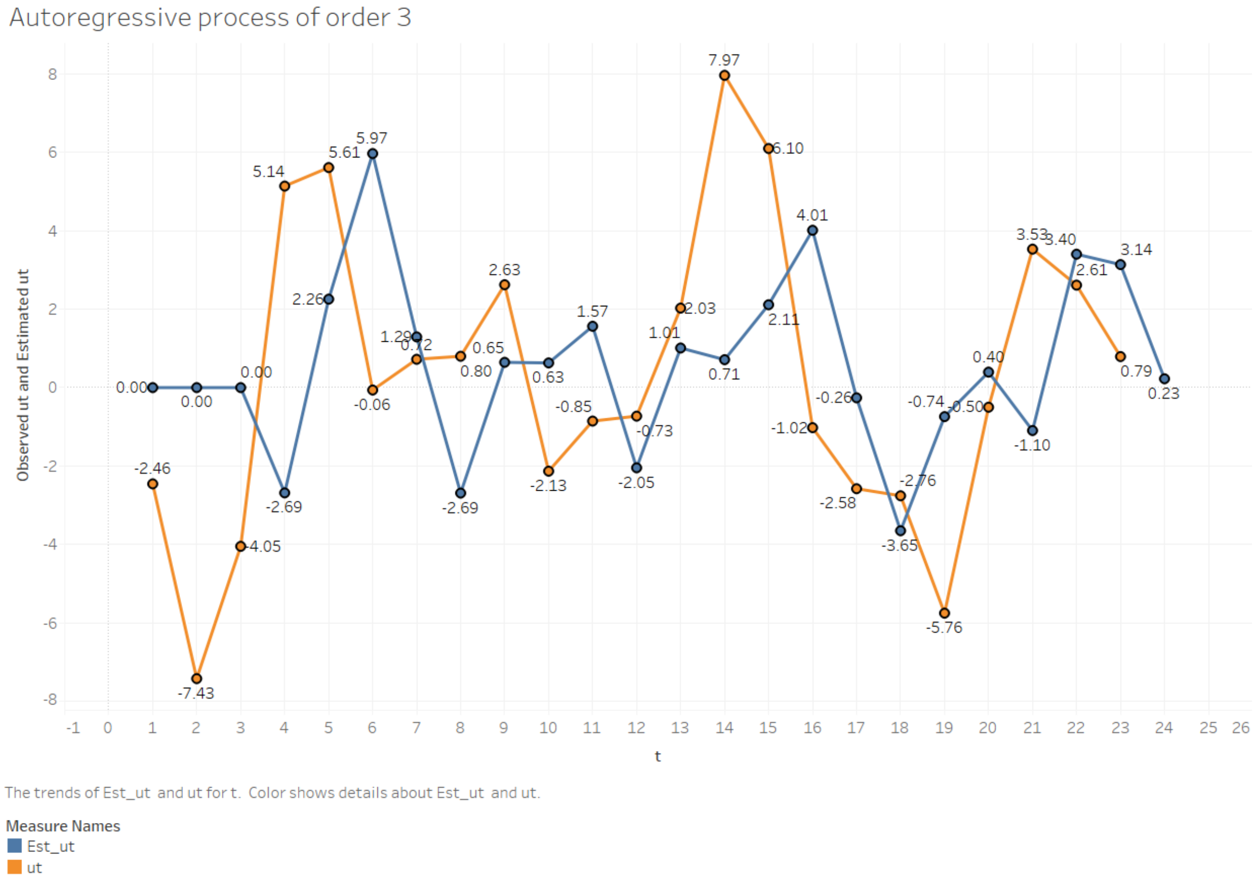


Tableau-



Conclusion-

The fitted AR(3) is-

$$\hat{u}_t = (0.56041030)u_{t-1} + (-0.51534329)u_{t-2} + (-0.06095102)u_{t-3}$$