Forecasting Time Series Homework 4

Yoon Tae Park[yp2201], Yeong Koh[yk2678], Yoobin Cheong[yc5206]

1) Consider the $AR(2)$ process $x_t = x_{t-1} - 5x_{t-2} + \varepsilon_t$. Determine whether the process is stationary.

A)

1) $x_t = x_{t-1} - 5x_{t-2} + \varepsilon_t$

Need to find a root for $z^2 = z - 5$

$z^2 - z + 5 = 0$

$$z = \frac{1 \pm \sqrt{1-4(5)}}{2} = \frac{1 \pm \sqrt{19}\,i}{2}$$

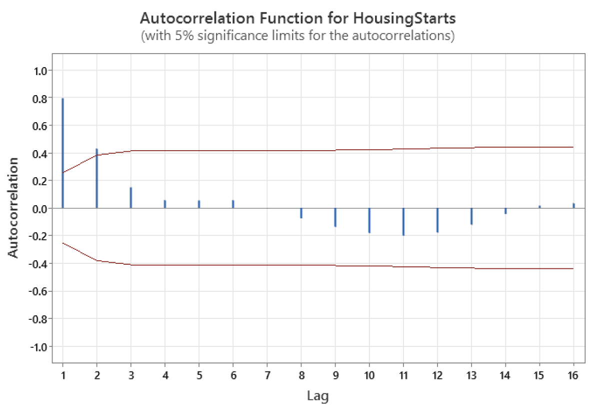
$$|z| = \sqrt{\left(\frac{1}{2}\right)^2 + \left(\frac{\sqrt{19}}{2}\right)^2} = \sqrt{\frac{1}{4} + \frac{19}{4}} = \sqrt{5}$$

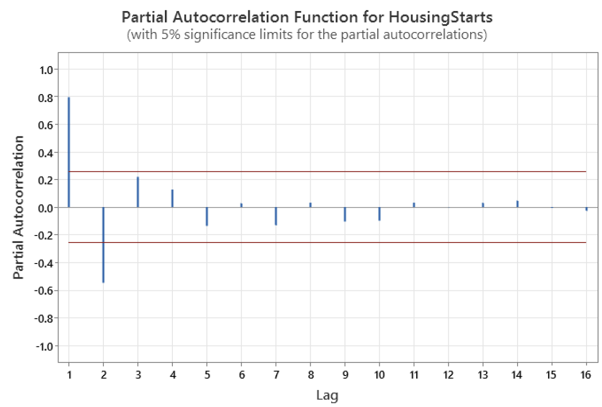
$\Rightarrow$ Since $|z| = \sqrt{5}$ which is greater than 1,

the process is <u>not</u> stationary.

2) Use the ACF and PACF to identify $ARIMA(p, d, q)$ models for the Housing Starts series, the log of the GDP series, the first differences of the log of the GDP series, and the first differences of the log of the CPI series (commonly known as "inflation"). Give reasons for your choices of p, d, q for each series. Do *not* try to estimate parameters. Just select p, d, q.

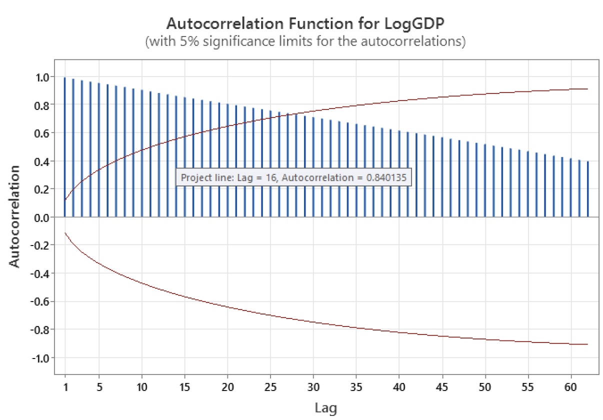
Housing Starts series

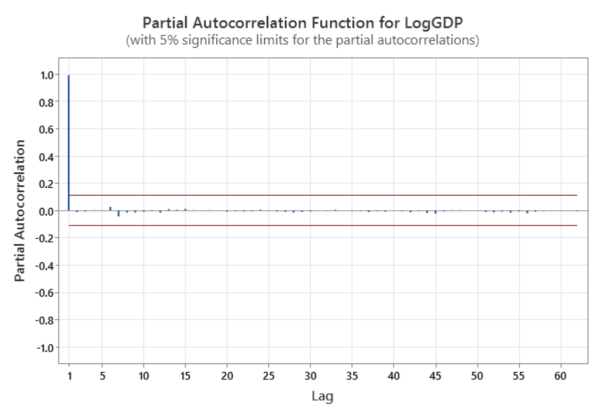




We would identify the series of Housing Starts as ARIMA(0, 0, 2). By looking at the ACF, we can see that ACF cuts off after k=2. For all k > 2, values are close to 0. By looking at the PACF, we can also see that PACF cuts off after k=2. However, ACF shows a stronger cutoff compared to PACF, and therefore we conclude that the model follows MA(2). Considering that we didn't implement any differencing, we can identify this model as ARIMA (0, 0, 2).

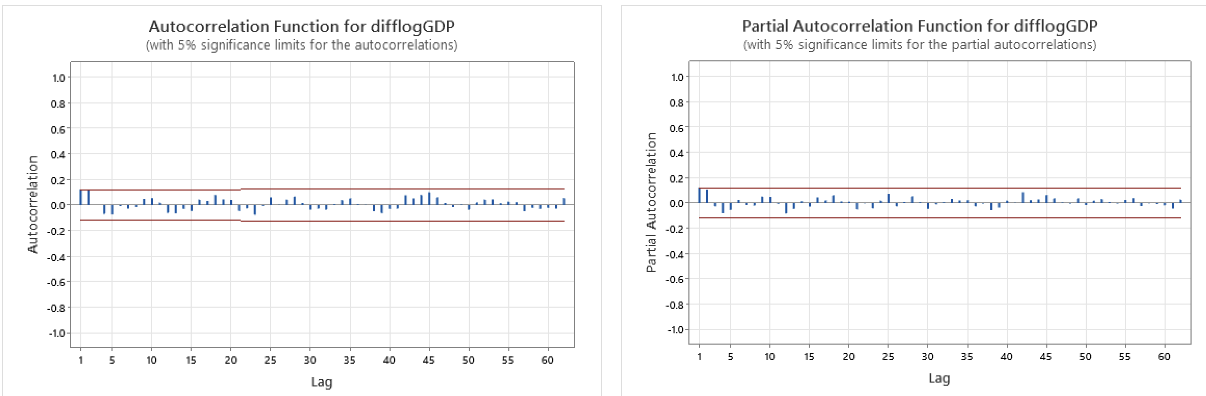
Log of GDP series





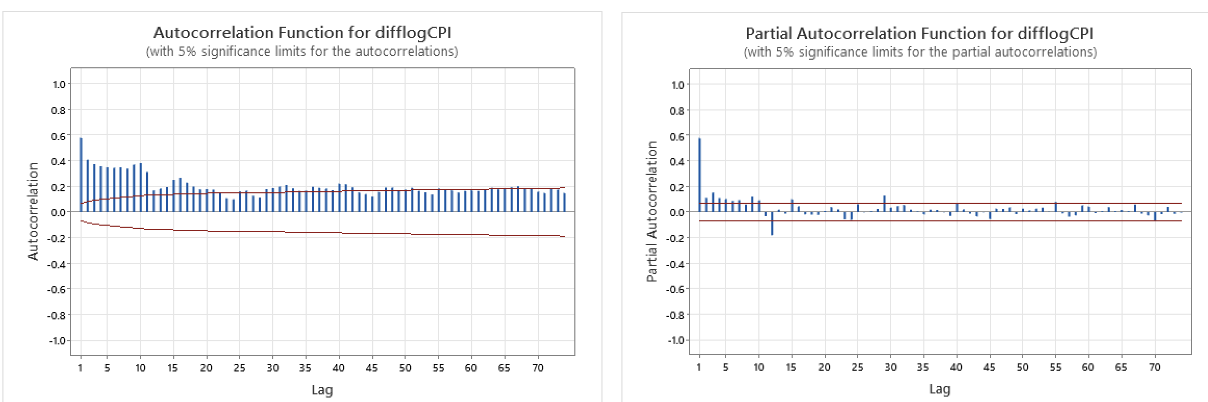
We would identify the series of Log of GDP as ARIMA(1, 0, 0). By looking at the ACF, we can see that ACF dies down, and for PACF, we can see that PACF cuts off after k=1. Therefore, we conclude that the model follows AR(1). Considering that we didn't implement any differencing, we can identify this model as ARIMA (1, 0, 0).

1st Difference of Log of GDP series



We would identify the series of 1st Difference of Log GDP as ARIMA(1, 0, 1). By looking at the ACF, we can see that ACF cuts off after k=1 because for all k > 1, values are getting close to 0. By looking at the PACF, we can see that PACF also cuts off after k=1.Therefore, the model follows MA(1) and AR(1). Considering that we didn't implement any differencing, we can identify this model as ARIMA(1, 0, 1).

1st Difference of Log of CPI series



We would identify the series of 1st Difference of Log CPI as ARIMA(1, 0, 0). By looking at the ACF, there is a noticeable decrease after k=1 but it is still ambiguous whether the series cuts off or dies down. By looking at the PACF, we can see that PACF cuts off after k=1 because for all k > 1, autocorrelation values are getting close to 0. Therefore, we select PACF as a stronger indicator, and conclude the model follows AR(1). Considering that we didn't implement any differencing, we can identify this model as ARIMA(1, 0, 0).

3) For the first difference of the log GDP series, use the method described in the handout for Chapter 3, Part IV, page 6 to estimate b in the invertible $MA(1)$ model $x_t = \varepsilon_t + b\varepsilon_{t-1}$.

Lag	ACF
1	0.122262

3. $MA(1)$ $\quad x_t = \varepsilon_t + b\,\varepsilon_{t-1}$

We can get estimates of b as :

$$\hat{b} = \frac{1 \pm \sqrt{1-4r_1^2}}{2r_1}$$

From the ACF table in question 2,

we know that $r_1 = 0.122262$

$$\therefore \quad \hat{b} = \frac{1 \pm \sqrt{1-4(0.122262)^2}}{2 \times (0.122262)}$$

$$b = 8.055 \quad \text{or} \quad 0.124146$$

as $|b| = 0.12416 < 1$, but $|b| = 8.055 > 1$

we select $b = 0.124146$

(For the invertible $MA(1)$ model)

4) For the first difference of the log GDP series, use the Yule-Walker equation $r_1 = \hat{a}_1 r_0$ to estimate a_1 in the $AR(1)$ model $x_t = a_1 x_{t-1} + \varepsilon_t$. Is your fitted model stationary?

4. $r_1 = \hat{a}_1 r_0$, $x_t = a_1 x_{t-1} + \varepsilon_t$
(AR(1))

Given ACF lag 1 as 0.122262,

we can plug in as $r_1 = 0.122262$ $r_0 = 1$

$\boxed{\therefore 0.122262 = \hat{a}_1}$

For stationary analysis,
Need to find a root for $z^1 = a_1 \cdot z^0$

$|z| = \sqrt{(0.122262)^2} = 0.122262$

Since $|z| = 0.122262 < 1$,

the process is stationary

5)

A) For the first difference of the log GDP series, use the two Yule-Walker equations

$$r_2 = \hat{a}_1 r_1 + \hat{a}_2 r_0$$

$$r_1 = \hat{a}_1 r_0 + \hat{a}_2 r_1$$

to estimate a_1 and a_2 in the $AR(2)$ model $x_t = a_1 x_{t-1} + a_2 x_{t-2} + \varepsilon_t$.

Lag	ACF	T	LBQ
1	$\boxed{0.122262}$	2.11	4.50
2	$\boxed{0.119354}$	2.03	8.80

5. (a) $AR(2): x_t = a_1 x_{t-1} + a_2 x_{t-2} + \varepsilon_t$

$$\left[\begin{array}{l} r_2 = \hat{a}_1 r_1 + \hat{a}_2 r_0 \\ r_1 = \hat{a}_1 r_0 + \hat{a}_2 r_1 \end{array}\right.$$

From the ACF of first difference of the log GDP series,

r_1 = lag1 ACF $= 0.122262$

r_2 = lag2 ACF $= 0.119354$

$r_0 = 1$

$$\rightarrow \left[\begin{array}{l} 0.119354 = \hat{a}_1 \cdot 0.122262 + \hat{a}_2 \\ 0.122262 = \hat{a}_1 + \hat{a}_2 \cdot 0.122262 \end{array}\right.$$

Solving two equations, we get $\hat{a}_1 \fallingdotseq 0.10930$

$\hat{a}_2 \fallingdotseq 0.10599$

B) Prove that your fitted $AR(2)$ model is stationary. (It must be stationary, since it can be proved in general that AR models estimated by solving the Yule-Walker equations are *always* stationary).

(b) to check if AR(2) model is stationary,

which is $x_t = 0.10930\, x_{t-1} + 0.10599\, x_{t-2} + \varepsilon_t$,

we need to find a root for

$$z^2 = 0.10930\, z + 0.10599$$

$$z^2 - 0.10930\, z - 0.10599 = 0$$

$$z = -0.274766 \quad \text{or} \quad 0.384766$$

as both $|z| < 1$, the model is stationary

C) Use your fitted model to forecast the log GDP (*not* just the first difference of the log GDP, but the log GDP itself) for the third quarter of 2021. (This is a one-step-ahead forecast for log GDP, based on an $ARIMA(2, 1, 0)$ model).

DATE	GDP	logGDP	lagGDP	diffGDP
2021-04-01	19368.310	9.87139	9.85512	0.0162743
2021-07-01	19478.893	9.87709	9.87139	0.0056932

5. (c) $x_t = a_1 x_{t-1} + a_2 x_{t-2} + \varepsilon_t$

We plug in values we have and get x_t.

$x_t = 0.10930 \times 0.0056932 + 0.1599 \times 0.0162743 + 0$

(Note: x_{t-1} - lag 1 of diff log GDP,
x_{t-2} - lag 2 of diff log GDP)

$= 0.00234717981$

As $x_t = Y_t - Y_{t-1}$,

$Y_t = Y_{t-1} + x_t$

$= 9.87709 + 0.00234717981$

$\doteqdot 9.87944$