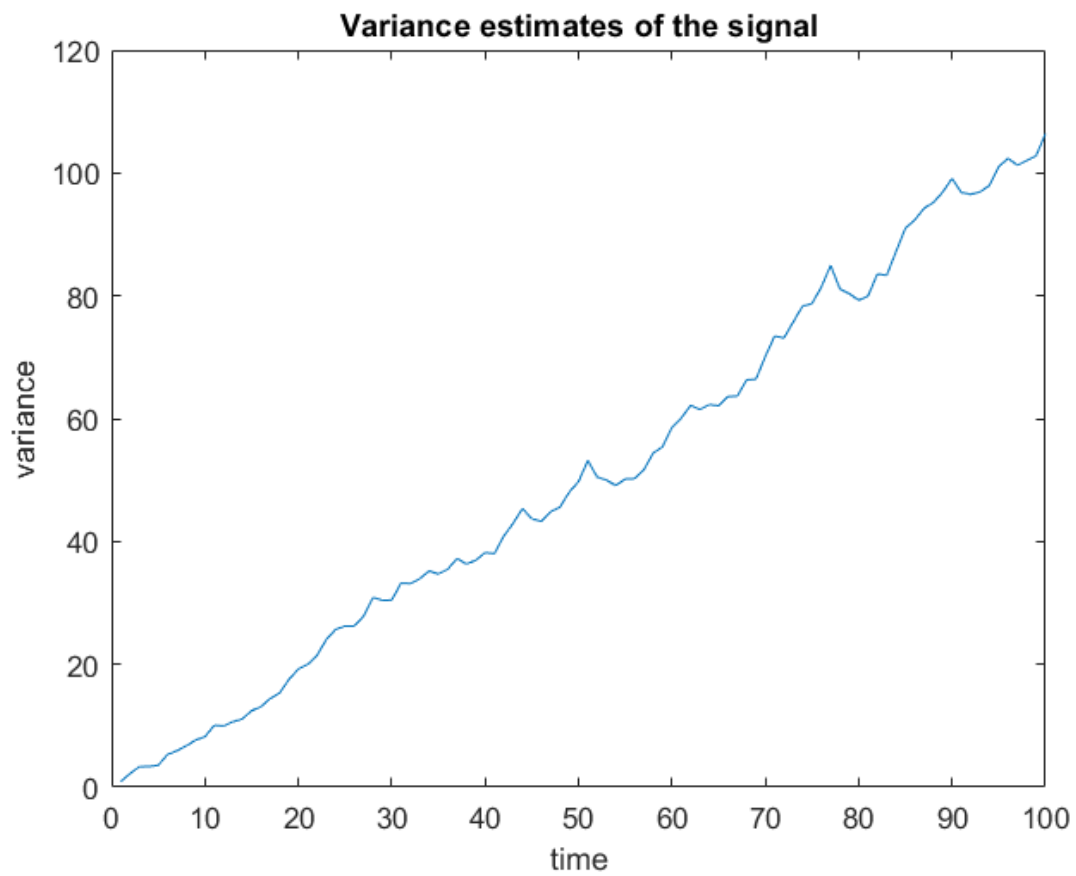


Question 1

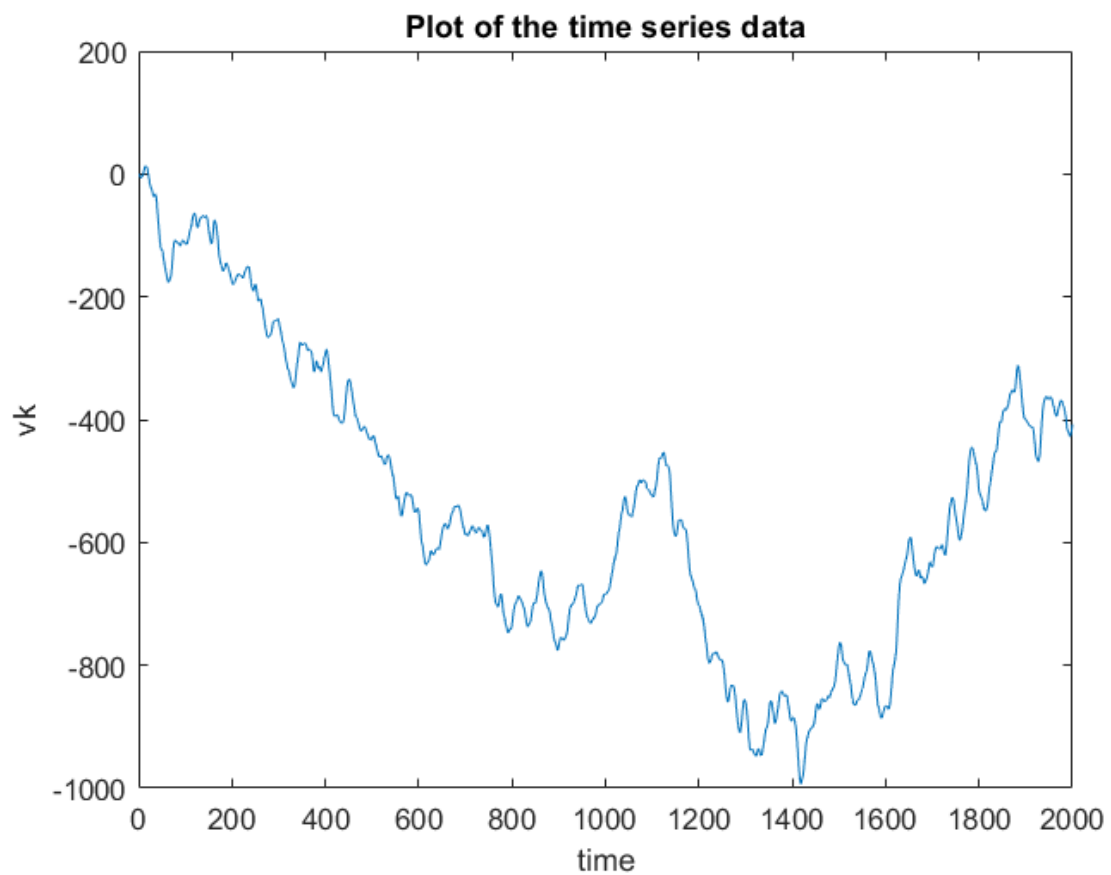


As expected, we see variance of the process keep increasing with time.

⇒ **The process is variance non-stationary.**

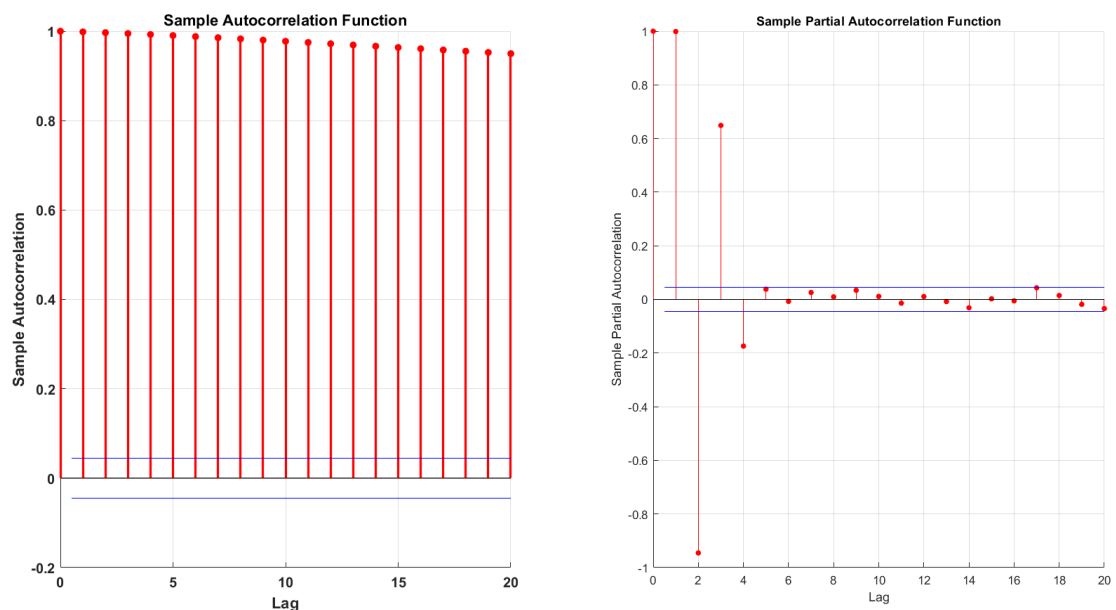
Question 3

Investigating the presence of random walk effects



Plot of the data indicates the presence of non-stationarity which is not of trend-type.

To determine the presence of integrating effects, let us examine the ACF and PACFs of the given data.



ACF decays **very very slowly**, while PACF[1] is 1. These 2 observations strongly indicate the presence of random walk effect. We shall now perform the **ADF test** to solidify our conclusion.

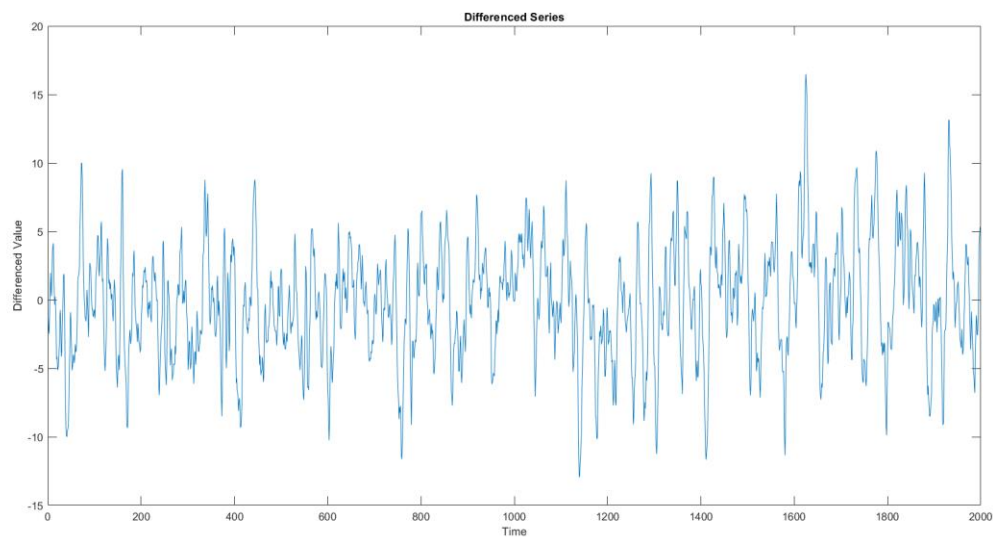
```
ADF Test
h =
    0

pval =
    0.8455
```

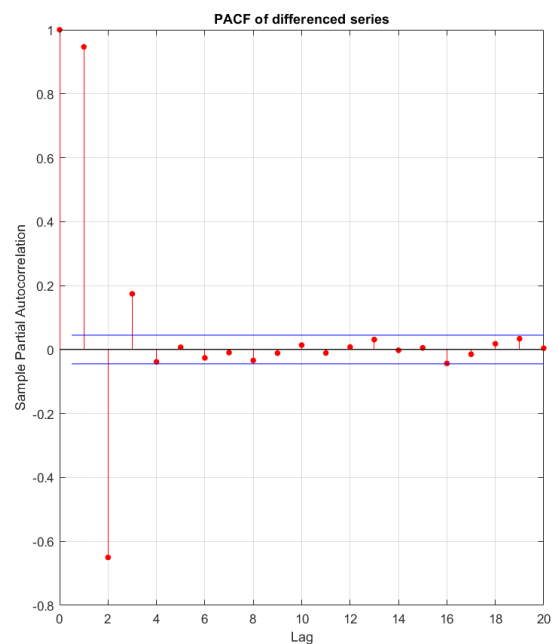
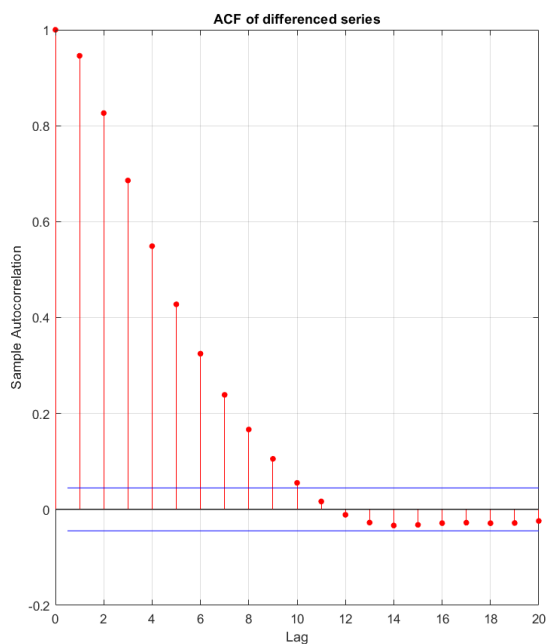
We obtain a high p-value of $0.845 > \alpha = 0.05$. This implies the null hypothesis of a unit root in the time series is rejected.

Therefore, the series contains random walk effects.

Differencing the series and finding the order of integration effects



Unlike the original series, the differenced series looks to have some sense of stationarity.



We can see an exponential decay of ACF indicating AR component and no abrupt change to 0 in PACF indicating MA component. This could mean that the integrating effects have been removed.

Conduct the **ADF test on the differenced series**:

```
ADF test on differenced series
h =
    1

pval =
    1.0000e-03
```

pvalue = 0.001 < alpha = 0.05

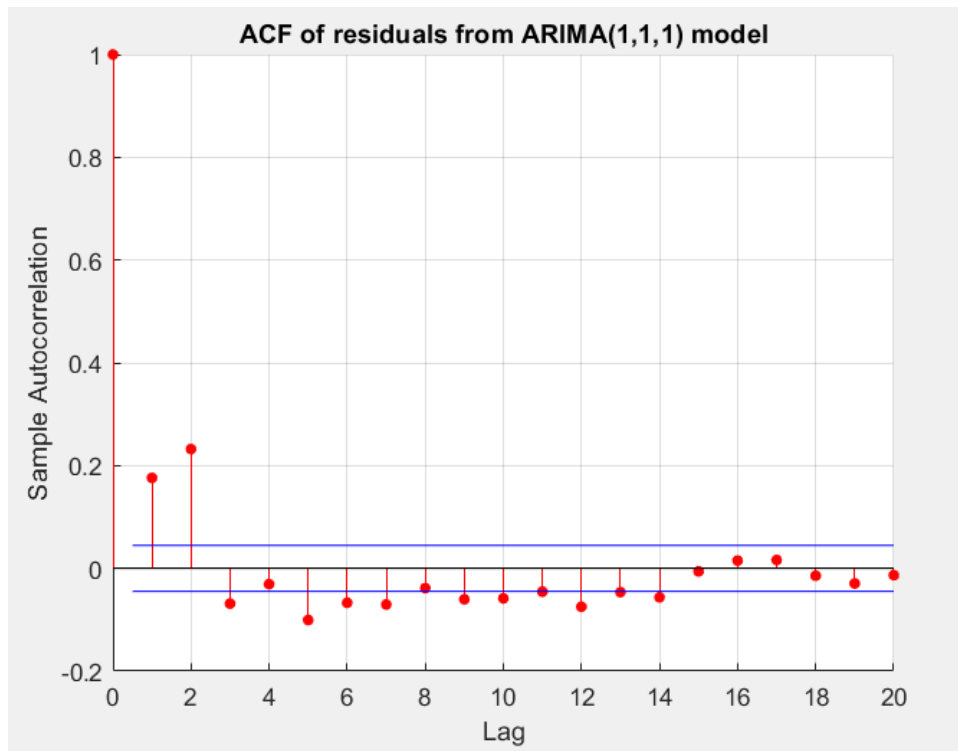
=> Null hypothesis is not rejected.

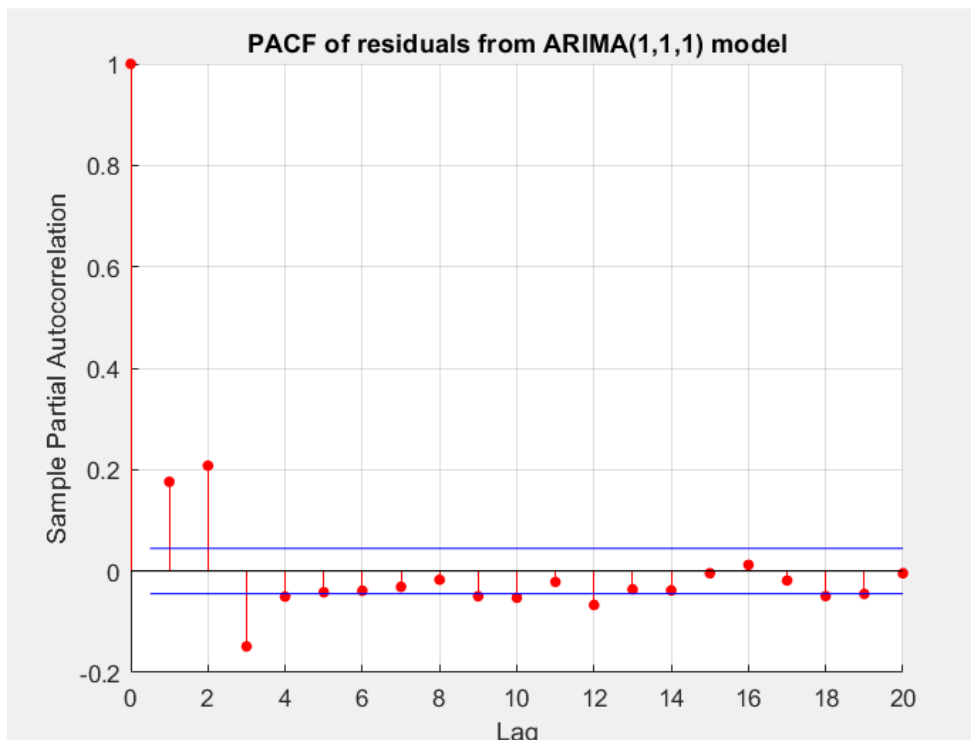
Therefore, we can say that the **differenced series doesn't contain random walk effects**. Since it seems to contain both AR and MA effects, we go for an ARIMA model. (Or we can model the differenced series as an ARMA model.)

ARIMA model-1: AR-1, I-1, MA-1

Since with a single differencing we are able to remove the random walk effects, we can say that, the **order of integrating process is 1**. That fixes the 'I' value of the ARIMA model.

Let us now examine the model residuals.





We can see that the ACF and PACF are still far away from zero at small lags. So, the residuals don't appear to be White Noise. We will confirm this using the lbq test.

Whiteness Test for Residuals results

1

0

With pvalue of 0, the lbqtest indicates **the residuals are not white**. This means that the model is underfitting. Errors in the parameter estimates:

ARIMA(1,1,1) Model (Gaussian Distribution)

Effective Sample Size: 2000

Number of Estimated Parameters: 3

LogLikelihood: -2960.14

AIC: 5926.28

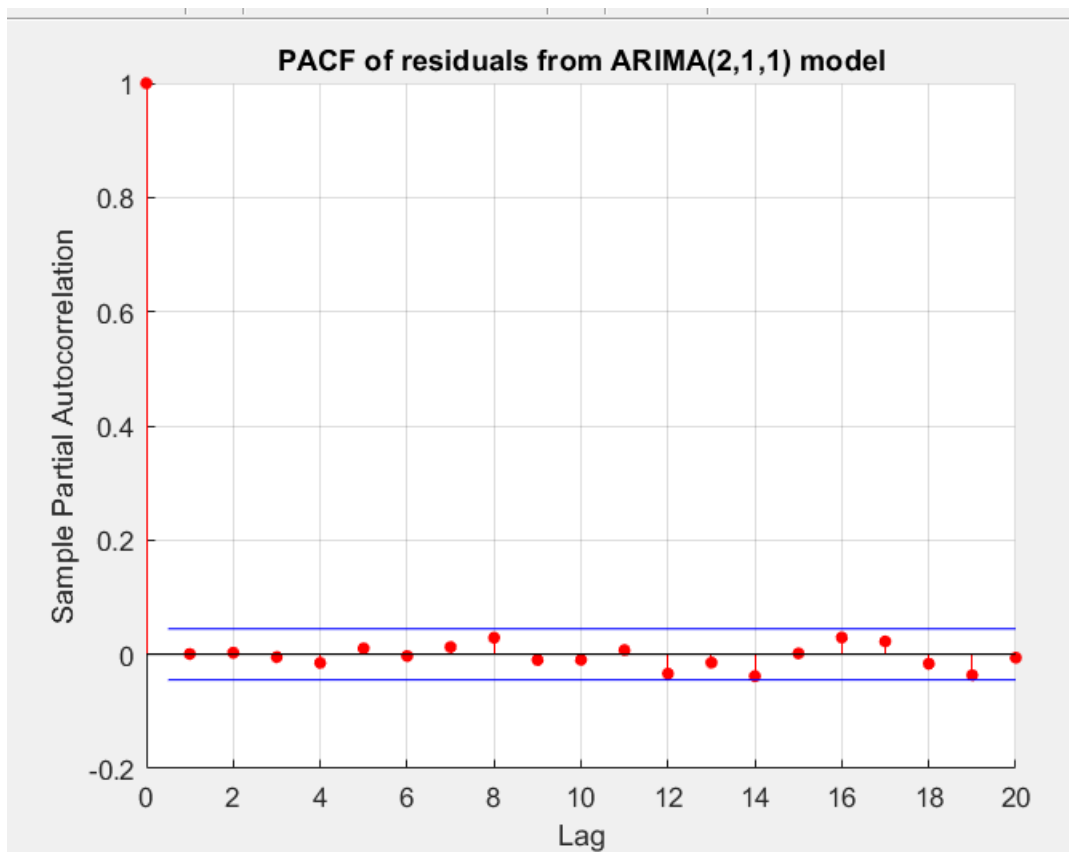
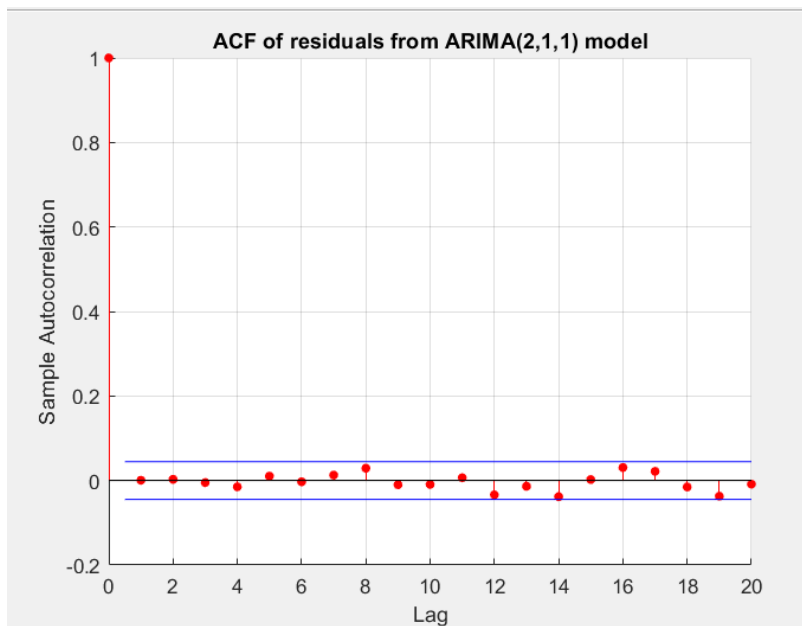
BIC: 5943.08

	Value	StandardError	TStatistic	PValue
Constant	0	0	NaN	NaN
AR{1}	0.91364	0.0090083	101.42	0
MA{1}	0.62323	0.018013	34.598	2.6672e-262
Variance	1.1301	0.033596	33.637	4.9076e-248

The ACF seems to decay and the PACF seems to cut off to zero after a set of initial lags. So let's increase the AR component and build a new ARIMA model.

ARIMA model-1: AR-2, I-1, MA-1

Examining the residuals:



PACF and ACF are extremely close to zero at all lags $\neq 0$. So, the residuals could be white noise!!

Performing the lbq test:

```
Whiteness Test for Residuals results
0
0.7668
```

Model passes the whiteness test! (Implies the model is not under-fitting)

Errors in parameter estimates:

ARIMA(2,1,1) Model (Gaussian Distribution)

```
Effective Sample Size: 2000
Number of Estimated Parameters: 4
LogLikelihood: -2843.6
AIC: 5695.2
BIC: 5717.61
```

	Value	StandardError	TStatistic	PValue
Constant	0	0	NaN	NaN
AR{1}	1.4103	0.02846	49.554	0
AR{2}	-0.5069	0.027996	-18.106	2.8568e-73
MA{1}	0.27114	0.031749	8.5402	1.3396e-17
Variance	1.0057	0.030717	32.742	3.9533e-235

We further observe that **all the coefficients are statistically significant**. (Implies the model is not over-fitting).

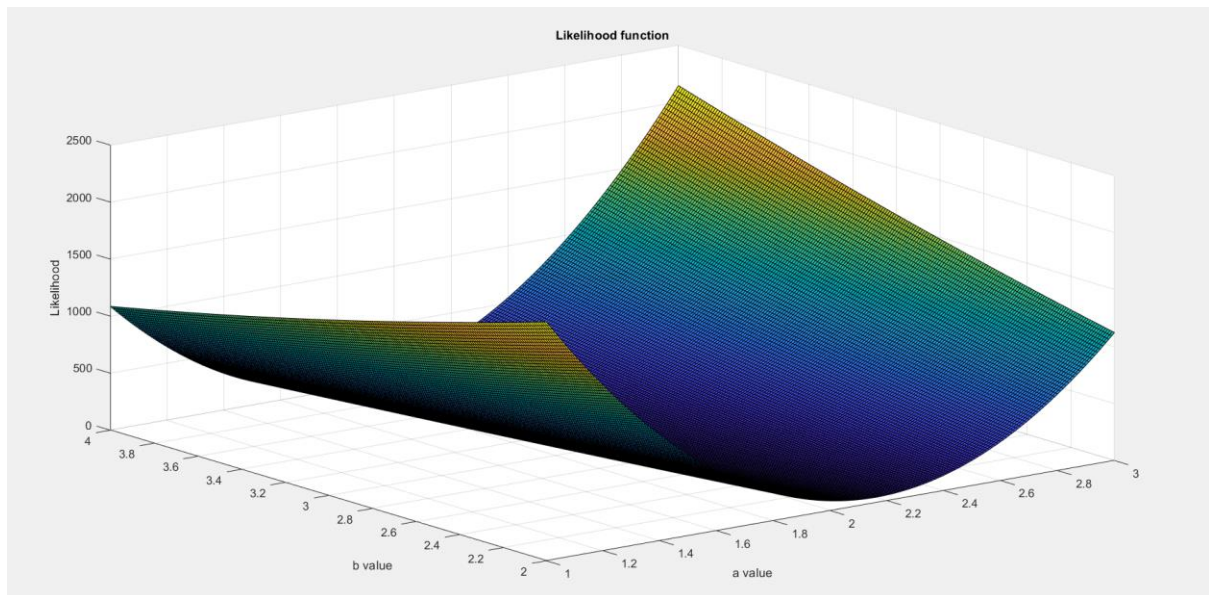
Also, ARIMA(2,1,1) has a **lower Akaike Information Criterion (AIC) value** compared to ARIMA(1,1,1). This backs our observation that the ARIMA(2,1,1) is the better model.

Conclusion: The built ARIMA (2,1,1) model satisfactorily represents the given data.

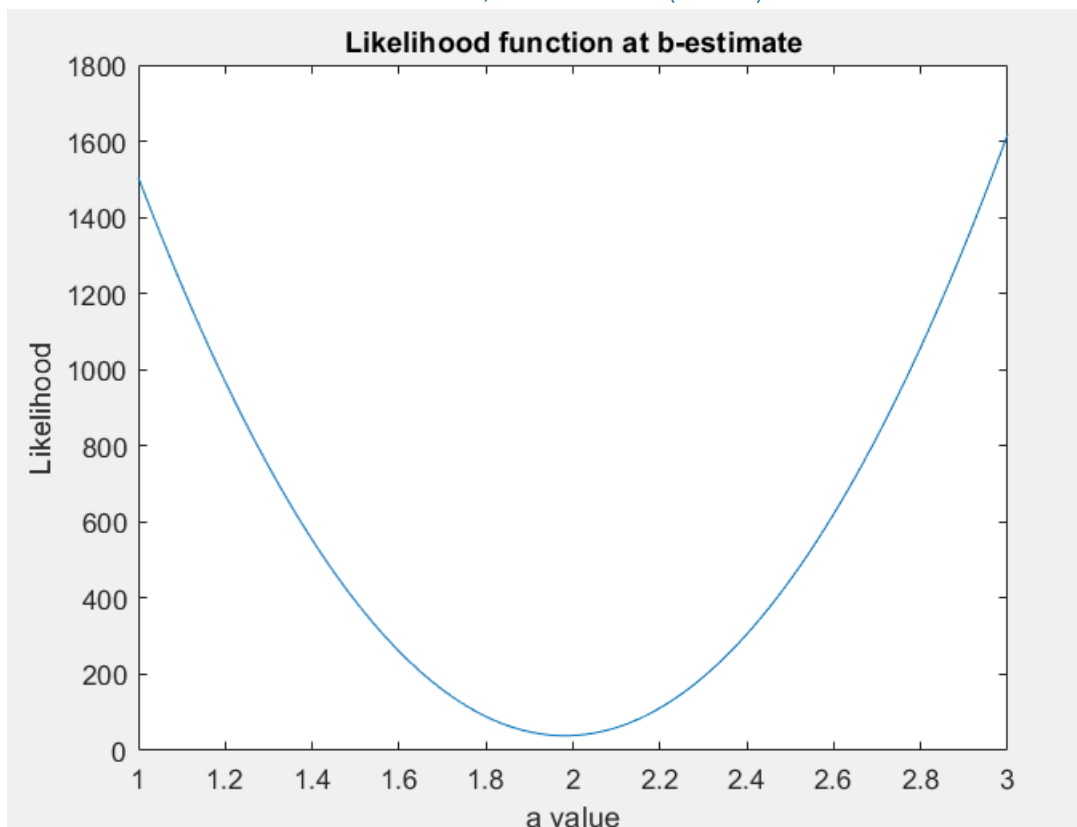
ARIMA (1,1,2) model was also tried but its residuals did not pass the whiteness test.

Question 4

From the graphs we can see that, the minimum of the likelihood function is close to the true values of a and b (2 and 3) and also to the analytical estimate.



Likelihood evaluated at different a , for a fixed b (=MLE).



Likelihood evaluated at different b , for a fixed a (=MLE).

