IL2233

Lab 1: ARIMA Model and Prediction

Shawn Nagar, Leon Moll

May 4, 2024

1 Stationarity of AR models

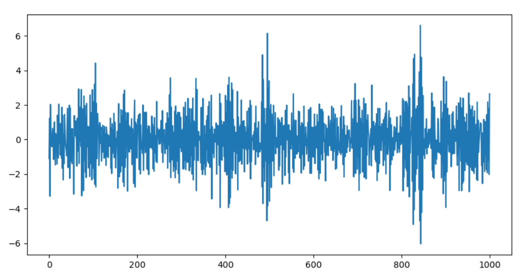

Figure 1: $AR(1)_1$

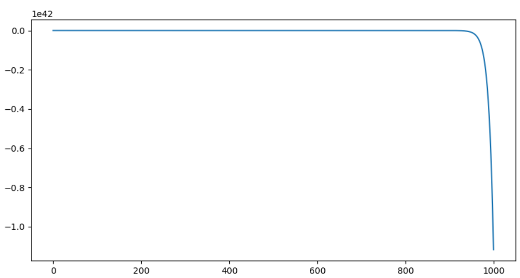

Figure 2: $AR(1)_2$

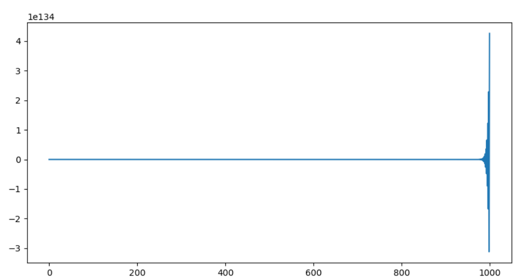

Figure 3: $AR(2)_3$

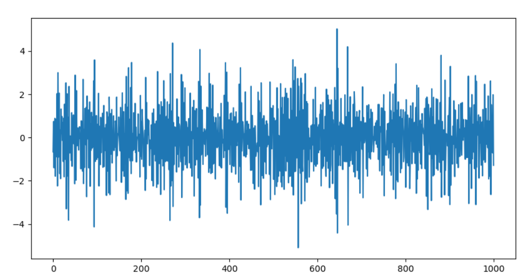

Figure 4: $AR4(2)_4$

Model	Visual Inspection	Coefficient Values	isstationary	ADF
$AR(1)_1$	stationary	$\|0.8\| < 1 \rightarrow$ stationary	true	0.0
$AR(1)_2$	not stationary	$\|-1.1\| \nless 1 \rightarrow$ not stationary	false	1.0
$AR(2)_3$	not stationary	$-0.5 + 1 < 1$ $-0.5 - 1 < 1$ $\|-0.5\| < 1 \rightarrow$ stationary	false	0.0
$AR(2)_4$	stationary	$0.5 + 1 \nless 1$ $0.5 - 1 < 1$ $\|0.5\| < 1 \rightarrow$ not stationary	true	0.0

For $AR(2)_3$ the results of the ADF do not match with my visual guess for stationarity. For the other it matches. This means that the ADF test is not always perfect. When interpreting the ADF results it is important to consider that the ADF test results are given in confidence intervalls and that its not an absolute value.

1.1 Questions

With visual inspection, how do you identify if a time series is stationary or not?
I check if I can see a trend, or if the difference between peaks and highs changes.

How do you judge the stationarity of time series using the unit-root method? Does it always give correct results?
If the p value is below 0.05 the series is stationary. For $AR(2)_3$ it gave a false result, therefore its not always correct.

What is the role of component ϵ_t in the model? Why is it important?
The component ϵ_t stands for the random error term or the residuals at time t. It is the component that captures the part of the time series data that cannot be explained by the autoregressive term.

To have an AR(p) model be stationary, is there any requirement on the auto-regressive coefficients? List the constraints for AR(1) and AR(2) models.

$$\text{For } p = 1 : -1 < \phi_1 < 1.$$
$$\text{For } p = 2 : -1 < \phi_2 < 1, \quad \phi_2 + \phi_1 < 1, \quad \phi_2 - \phi_1 < 1.$$

2 ACF, PACF of AR models

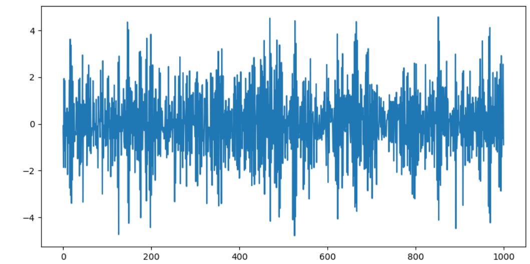

Figure 5: Line Plot for 1

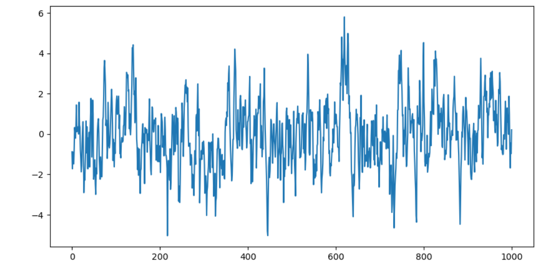

Figure 6: Line Plot for 2

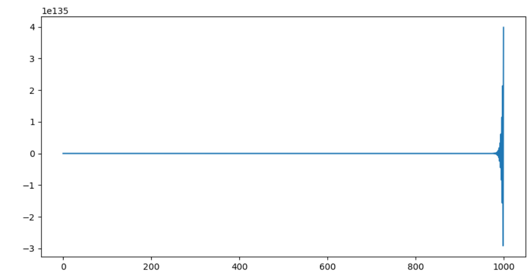

Figure 7: Line Plot for 3

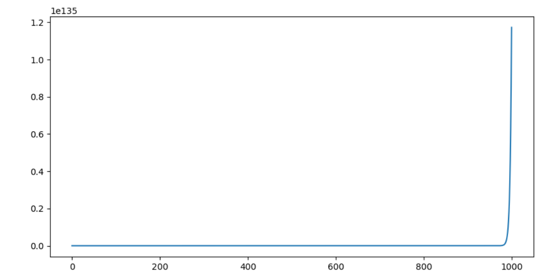

Figure 8: Line Plot for 4

1 and 2 are stationary, 3 and 4 are not stationary.

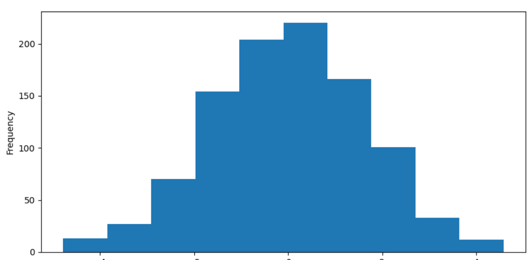

Figure 9: Histogram Plot for 1

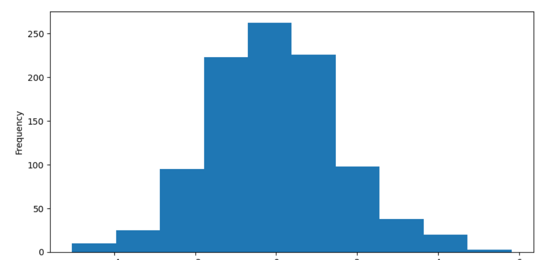

Figure 10: Histogram Plot for 2

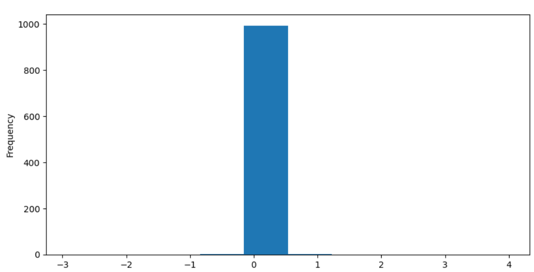

Figure 11: Histogram Plot for 3

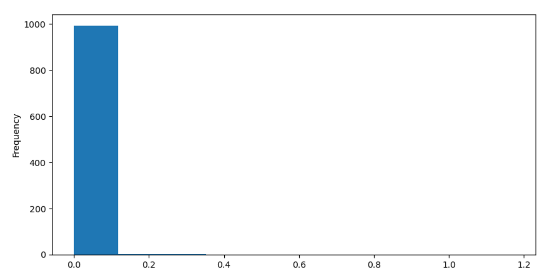

Figure 12: Histogram Plot for 4

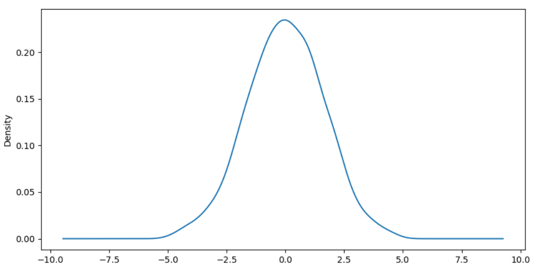

Figure 13: Density Plot for 1

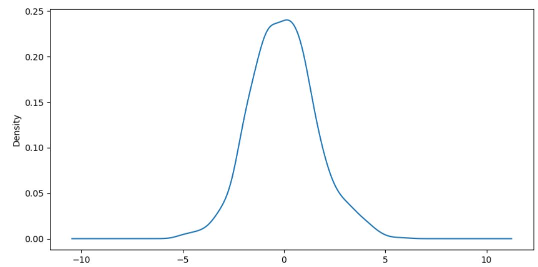

Figure 14: Density Plot for 2

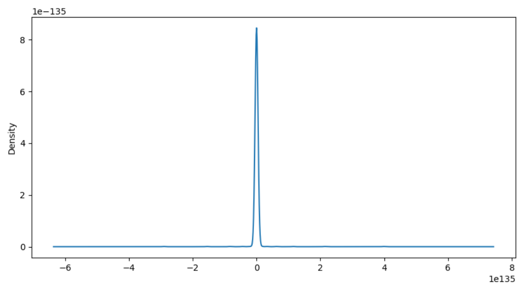

Figure 15: Density Plot for 3

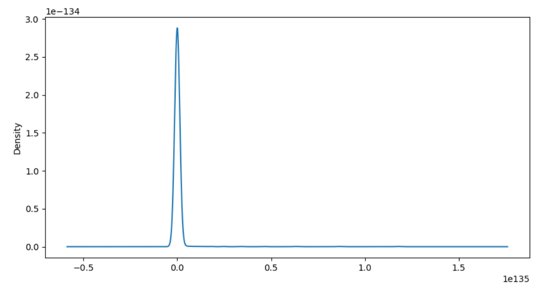

Figure 16: Density Plot for 4

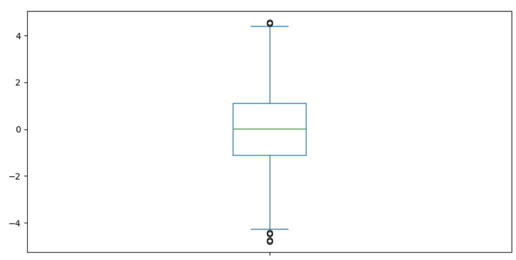

Figure 17: Box Plot for 1

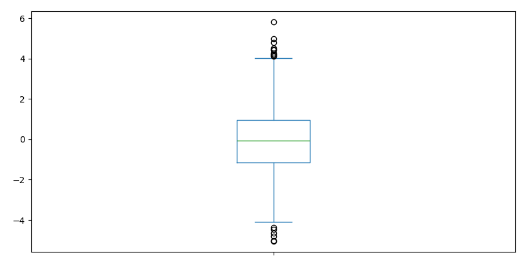

Figure 18: Box Plot for 2

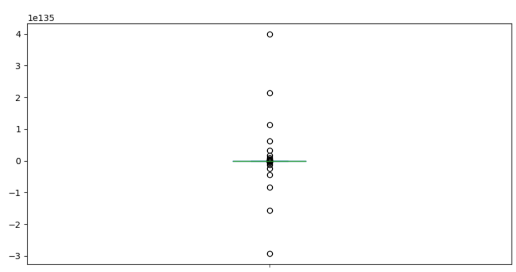

Figure 19: Box Plot for 3

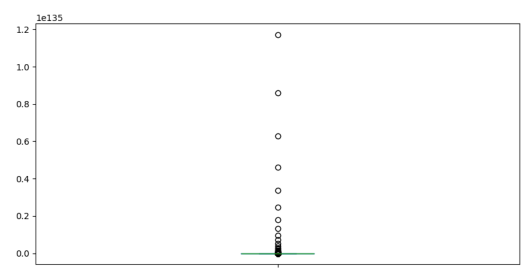

Figure 20: Box Plot for 4

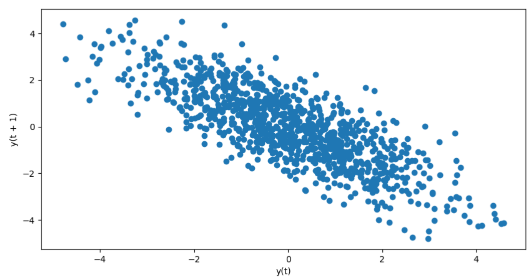

Figure 21: lag-1 plot for 1

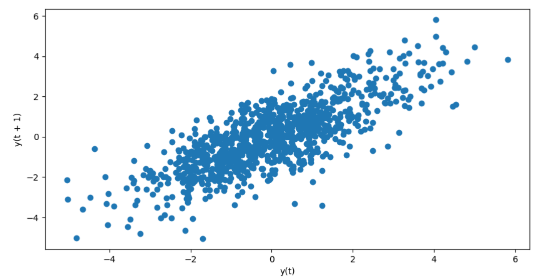

Figure 22: lag-1 plot for 2

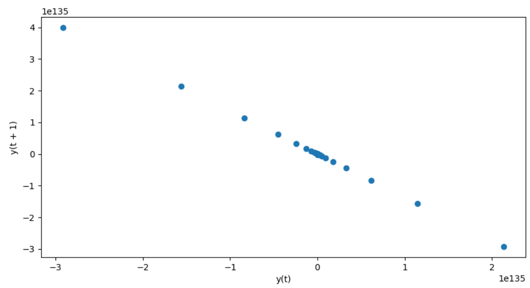

Figure 23: lag-1 plot for 3

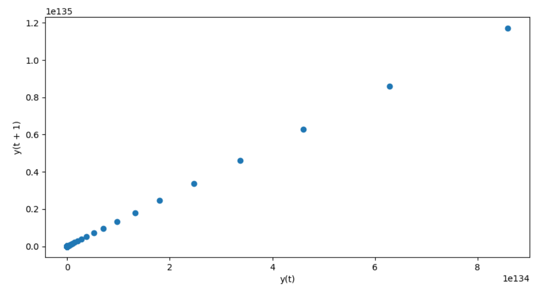

Figure 24: lag-1 plot for 4

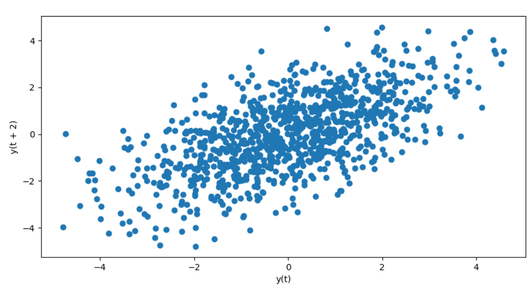

Figure 25: lag-2 plot for 1

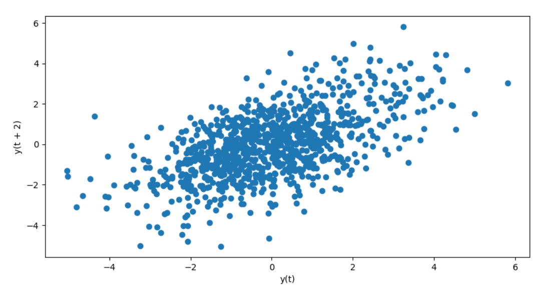

Figure 26: lag-2 plot for 2

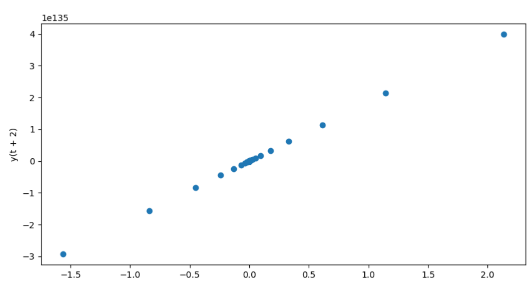

Figure 27: lag-2 plot for 3

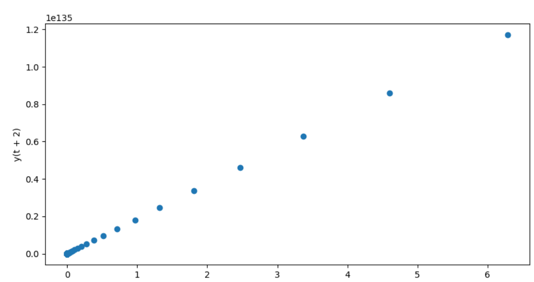

Figure 28: lag-2 plot for 4

In the lag-1 plot for 1 and 3 there is a negative autocorrelation. In the lag-2 plot for 1 and 3 there is a positive autocorrelation. For both lag plots for 2 and 4 there is a positive autocorrelation.

The lag-1 plot of 1 and 2 are more spread then the lag-1 plot of 3 and 4. This is due to the smaller absolute coefficient of the AR models. 1 and 2 have a smaller absolute first coefficient with 0.8 compared to 1. The lag-2 plot of 3 and 4 are also narrower than the plots for 1 and 2. The lag-2 plots for 1 and 2 are not completly randomly spread though. This is because they are still slightly dependent on the last values, although they only have a coefficient for t-1.

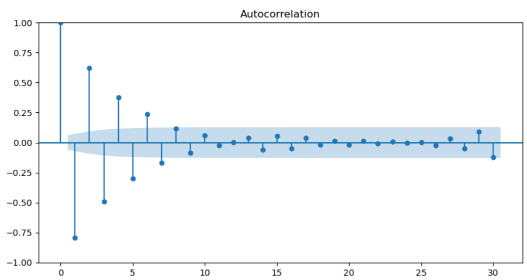

Figure 29: ACF plot for 1

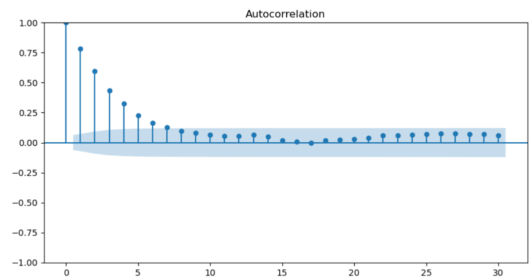

Figure 30: ACF plot for 2

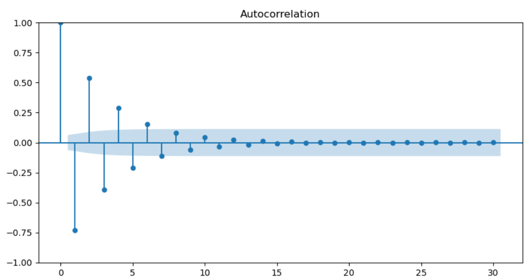

Figure 31: ACF plot for 3

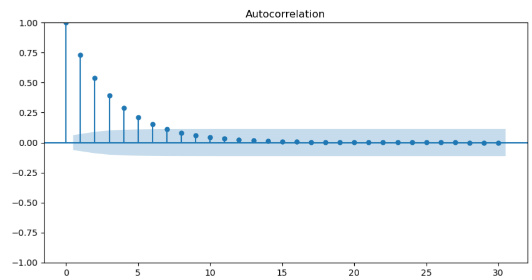

Figure 32: ACF plot for 4

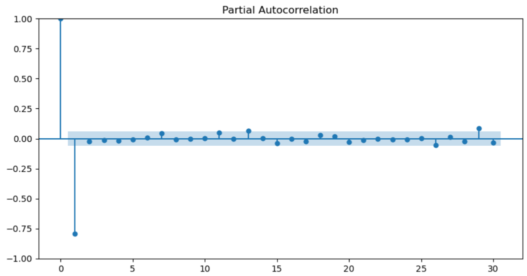

Figure 33: PACF plot for 1

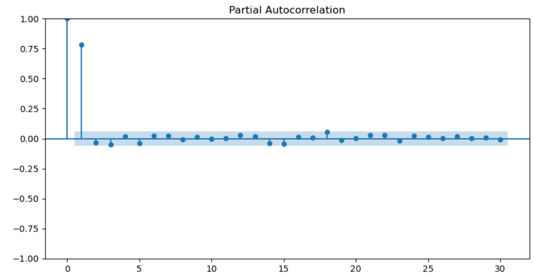

Figure 34: PACF plot for 2

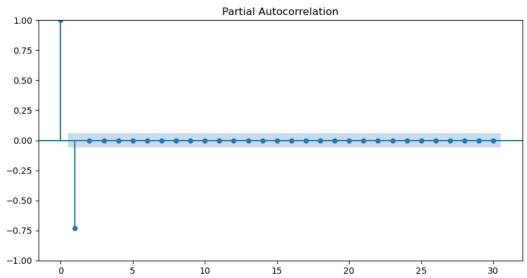

Figure 35: PACF plot for 3

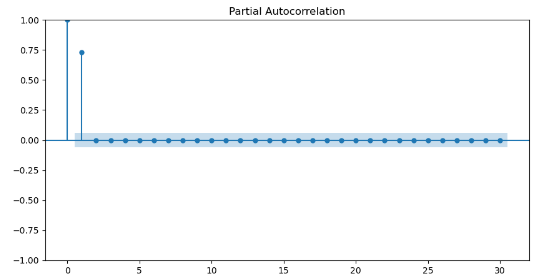

Figure 36: PACF plot for 4

2.1 Questions

What characteristics can you observe from the ACF graphs of the AR(p) models?
All ACF graphs show that the autocorrelation declines with the number of lags. All ACF graphs follow the observations from the lag-1 and log-2 plots.

What characteristics can you observe from the PACF graphs of the AR(p) models?
All PACF graphs show that there is only a partial autocorrelation upto a lag of 1. This would mean that also 3 and 4 are an AR(1) model which is not the case. From a PACF graph you can see the order of the series, which helps to pick the right order of the AR model.

3 Invertibility, ACF, PACF of MA models

Model	Visual Inspection	Coefficient Values	isinvertible
1	?	$\lvert -2\rvert \nless 1 \rightarrow$ not invertible	false
2	?	$\lvert -0.5\rvert < 1 \rightarrow$ invertible	true
3	?	$\lvert 16/25\rvert < 1$ $16/25 + (-4/5) > -1$ $-4/5 - 16/25 < 1$ $\rightarrow$ invertible	true
4	?	$\lvert 25/16\rvert \nless 1 \rightarrow$ not invertible	false

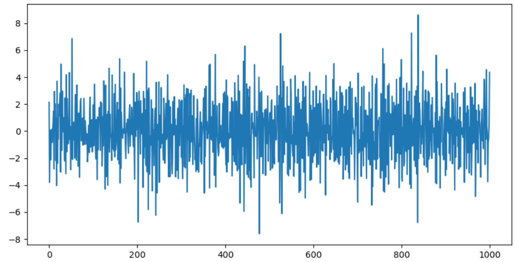

Figure 37: Line Plot of 1

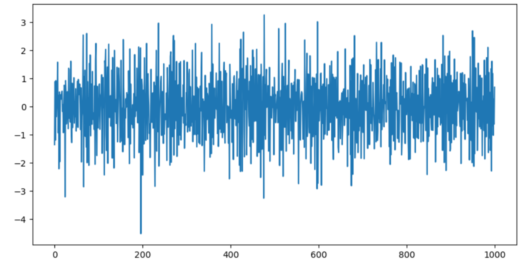

Figure 38: Line Plot of 1

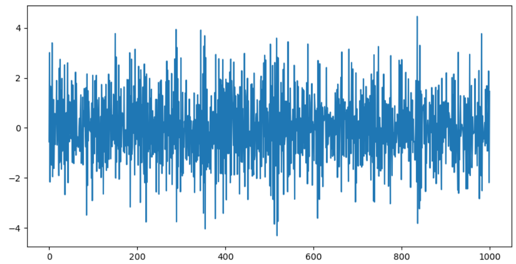

Figure 39: Line Plot of 1

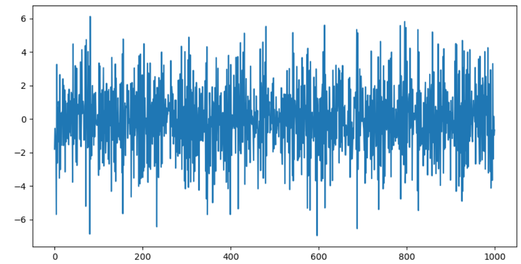

Figure 40: Line Plot of 1

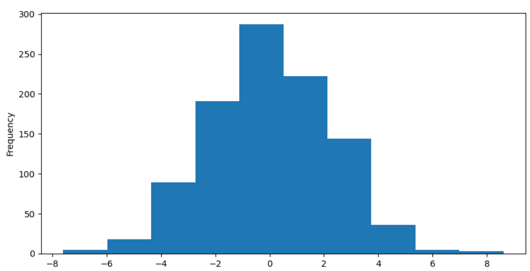

Figure 41: Histogram Plot for 1

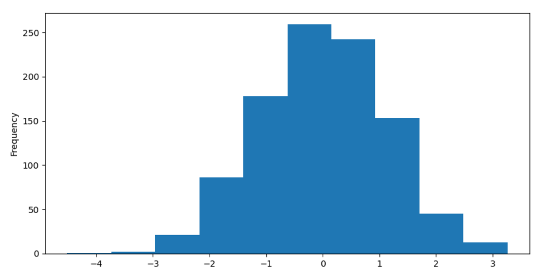

Figure 42: Histogram Plot for 2

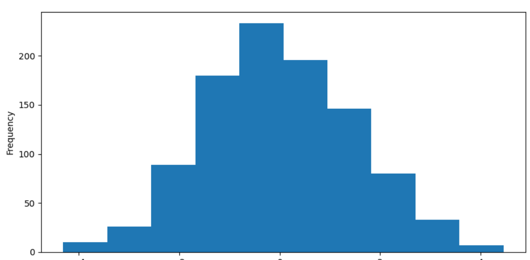

Figure 43: Histogram Plot for 3

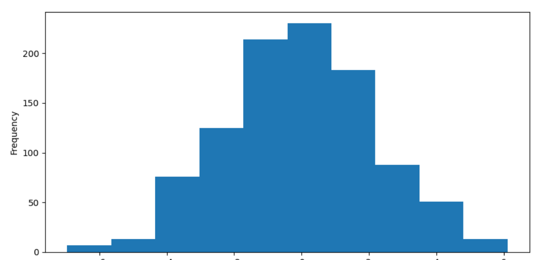

Figure 44: Histogram Plot for 4

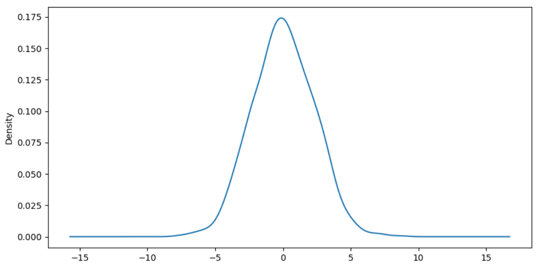

Figure 45: Density Plot for 1

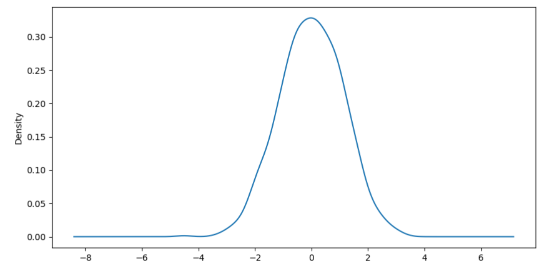

Figure 46: Density Plot for 2

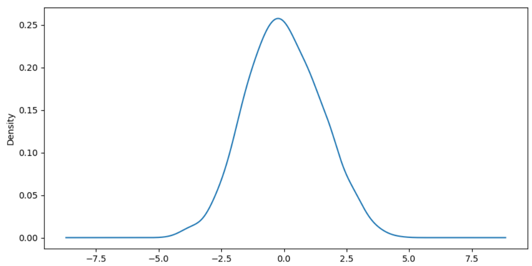

Figure 47: Density Plot for 3

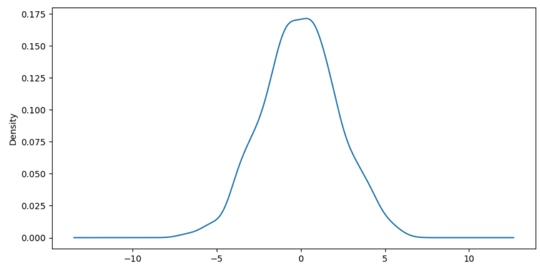

Figure 48: Density Plot for 4

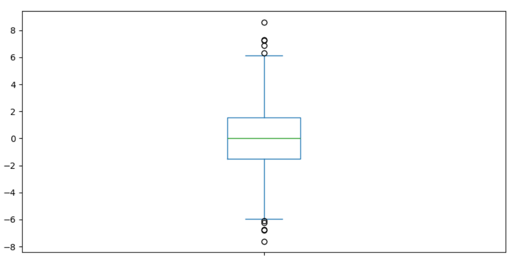

Figure 49: Box Plot for 1

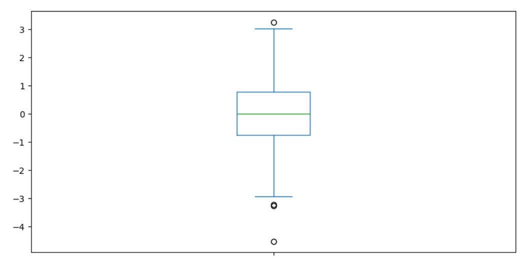

Figure 50: Box Plot for 2

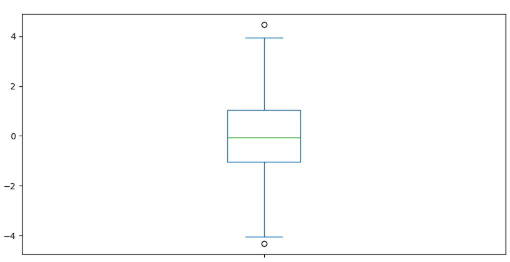

Figure 51: Box Plot for 3

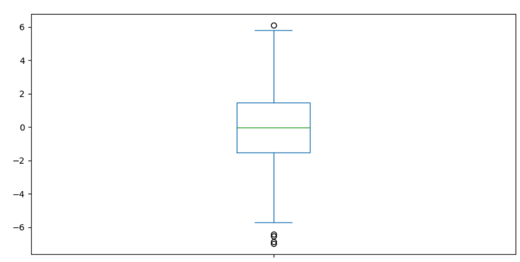

Figure 52: Box Plot for 4

There are no outliers, the only visible difference in the plots is the variance.

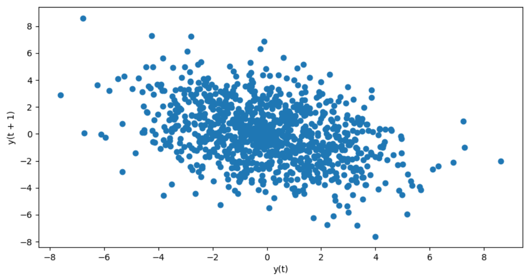

Figure 53: lag-1 plot for 1

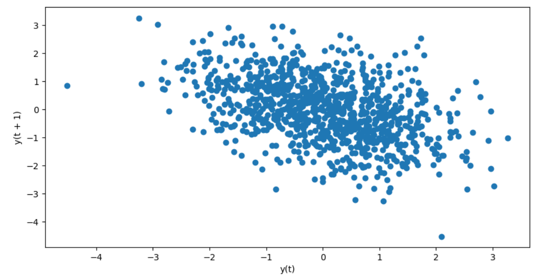

Figure 54: lag-1 plot for 2

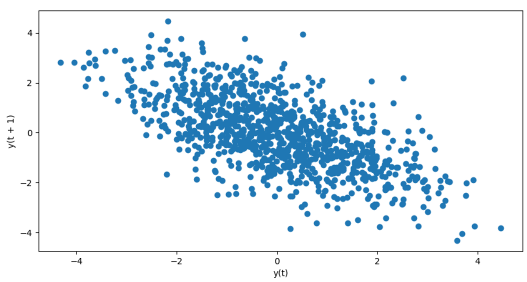

Figure 55: lag-1 plot for 3

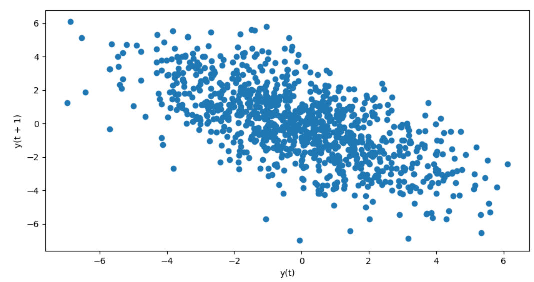

Figure 56: lag-1 plot for 4

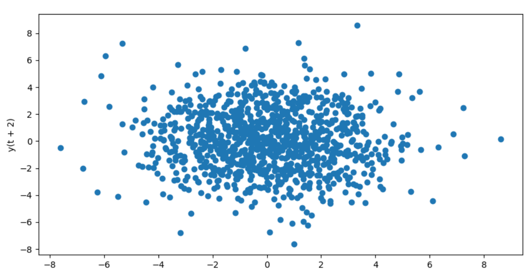

Figure 57: lag-2 plot for 1

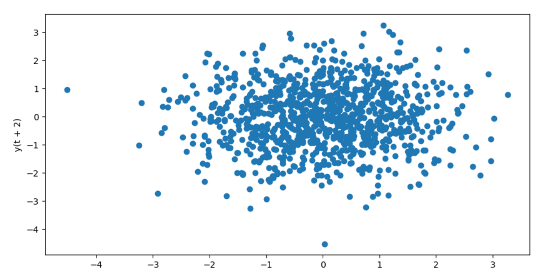

Figure 58: lag-2 plot for 2

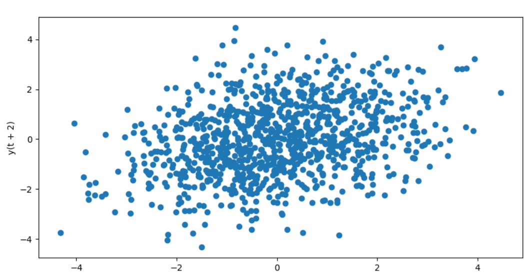

Figure 59: lag-2 plot for 3

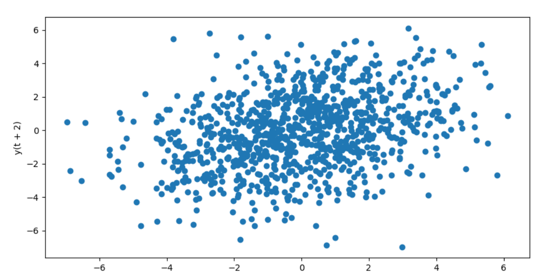

Figure 60: lag-2 plot for 4

The lag-1 plot of 3 and 4 seem to show an autocorrelation.

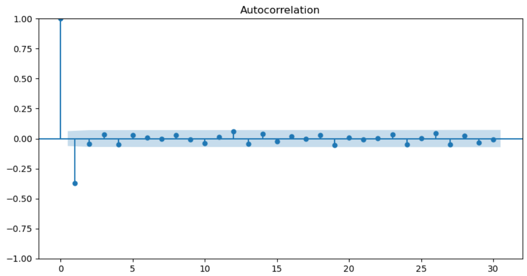

Figure 61: ACF plot for 1

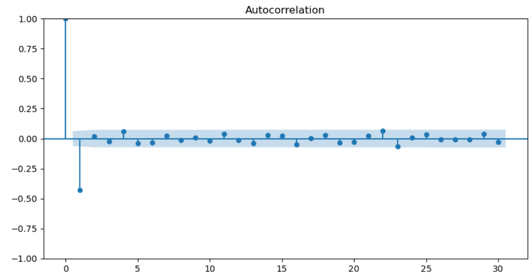

Figure 62: ACF plot for 2

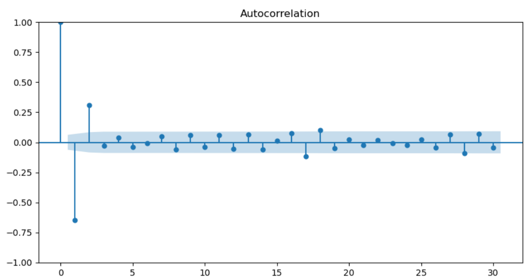

Figure 63: ACF plot for 3

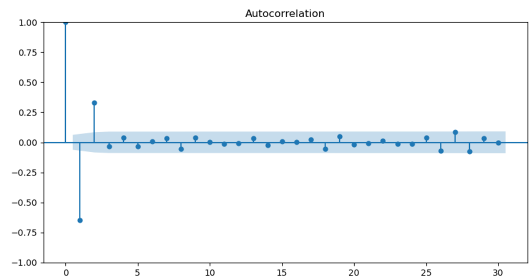

Figure 64: ACF plot for 4

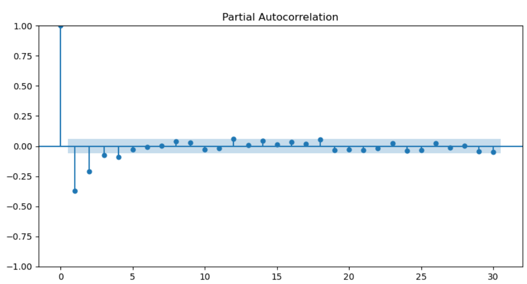

Figure 65: PACF plot for 1

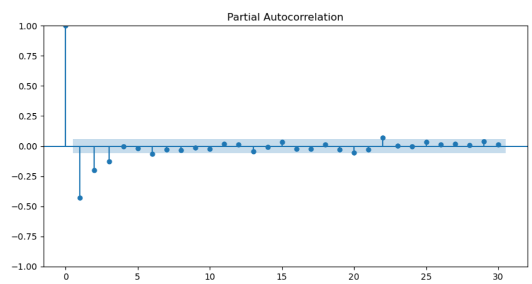

Figure 66: PACF plot for 2

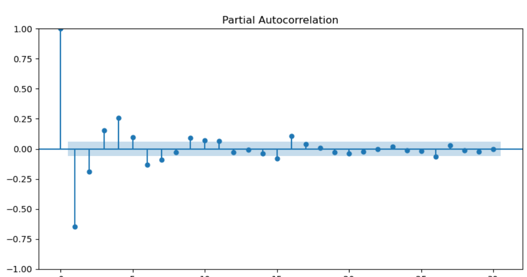

Figure 67: PACF plot for 3

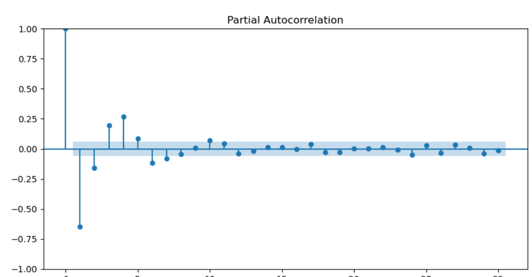

Figure 68: PACF plot for 4

3.1 Questions

Are all the MA models invertible? If not, which ones are invertible and which ones are not invertible? 2 and 3 are invertible, 1 and 4 are not.

What characteristics can you observe from the ACF graphs of the MA(q) models?
The ACF plots for 1 and 2 show a autocorrelation for a lag of 1, while the ACF plots for 3 and 4 show autocorrelation upto a lag of 2. This corresponds to the order of the MA(q) models.

What characteristics can you observe from the PACF graphs of the MA(q) models?
The The PACF plots for 1 and 2 only show a slight autocorrelation for lag -1 and a higher autocorellation for the lag-1 for 3 and 4. But they are not as clear as the ACF.

To have an MA(q) model be invertible, is there any requirement on the auto-regressive coefficients? List the constraints for MA(1) and MA(2) models.

$$\text{Für } q = 1 : -1 < \theta_1 < 1.$$
$$\text{Für } q = 2 : -1 < \theta_2 < 1, \quad \theta_2 + \theta_1 > -1, \quad \theta_1 - \theta_2 < 1.$$

4 Stationarity, ACF and PACF of ARMA models

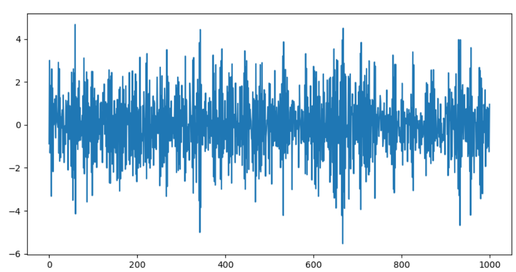

Figure 69: Line Plot of 1

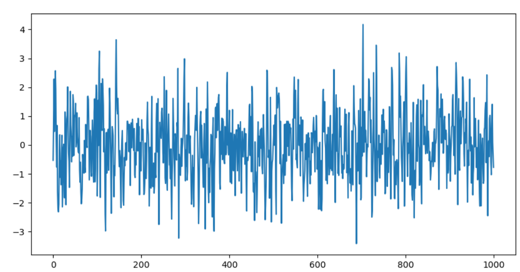

Figure 70: Line Plot of 1

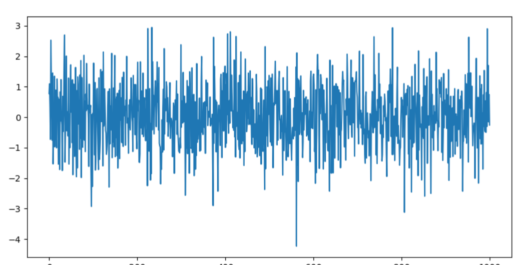

Figure 71: Line Plot of 1

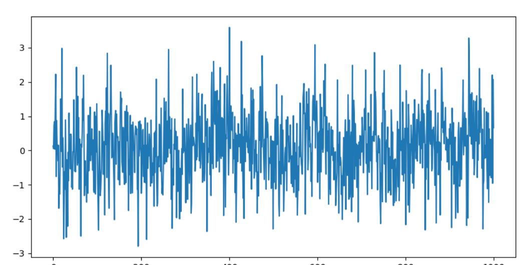

Figure 72: Line Plot of 1

Model	Visual Inspection	isstationary	isinvertible	ADF
1	Stationary	true	true	0.0
2	Stationary	true	true	0.0
3	Stationary	true	true	0.0
4	Stationary	true	true	0.0

The ADF results match the visual inspection for every model.

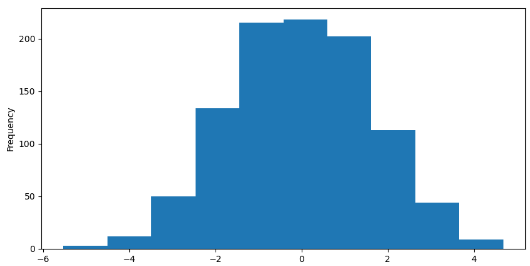

Figure 73: Histogram Plot for 1

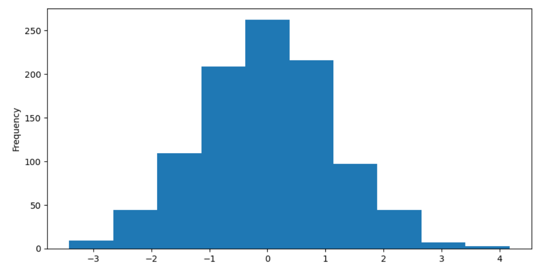

Figure 74: Histogram Plot for 2

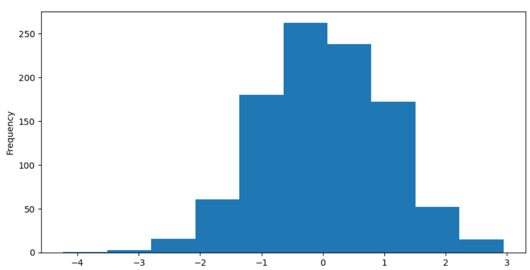

Figure 75: Histogram Plot for 3

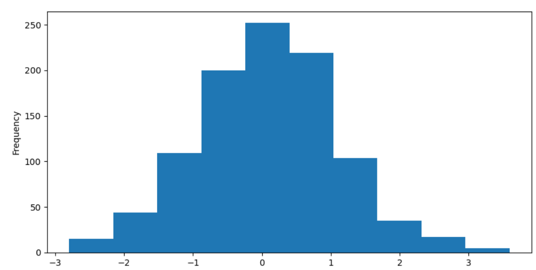

Figure 76: Histogram Plot for 4

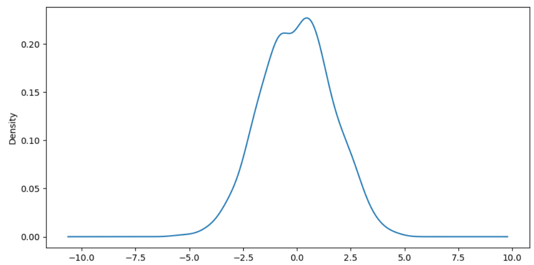

Figure 77: Density Plot for 1

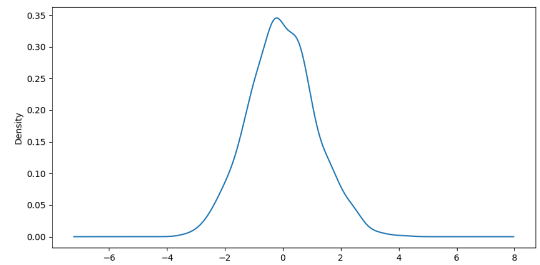

Figure 78: Density Plot for 2

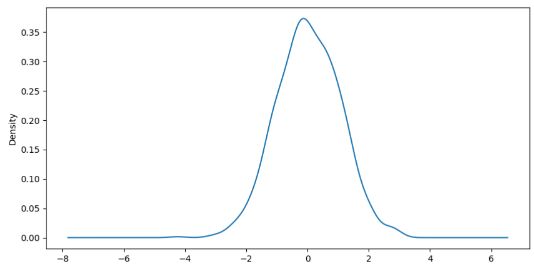

Figure 79: Density Plot for 3

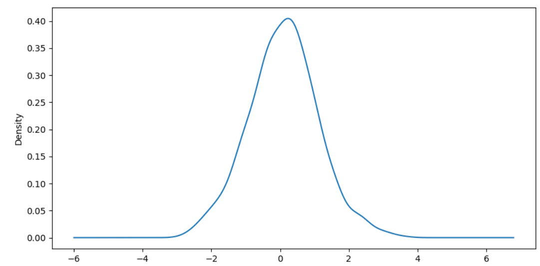

Figure 80: Density Plot for 4

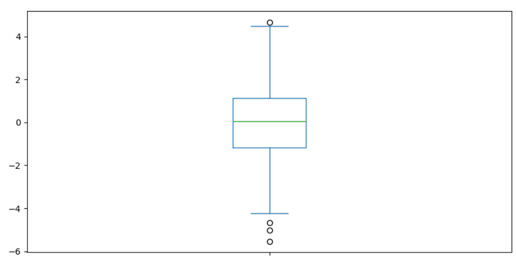

Figure 81: Box Plot for 1

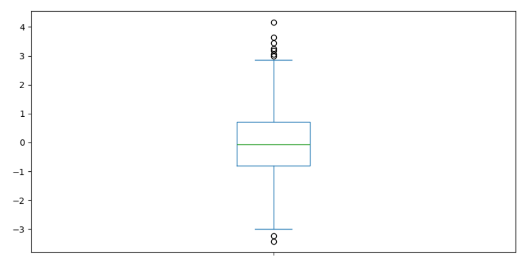

Figure 82: Box Plot for 2

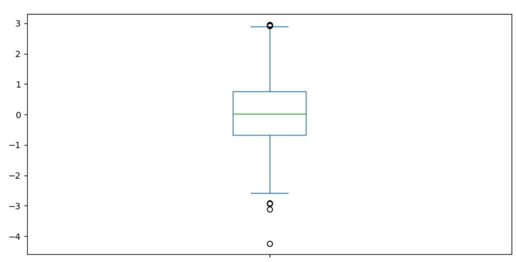

Figure 83: Box Plot for 3

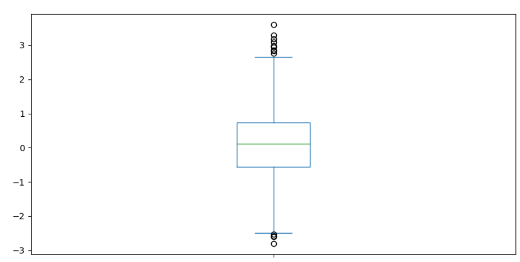

Figure 84: Box Plot for 4

There are no outliers from the plots, the only clear difference in the plots is the variance.

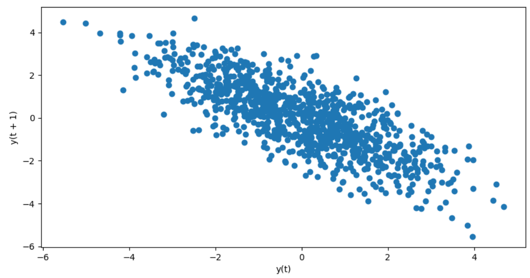

Figure 85: lag-1 plot for 1

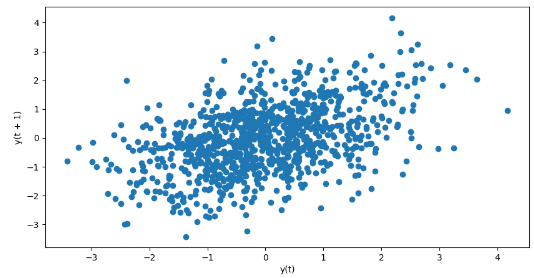

Figure 86: lag-1 plot for 2

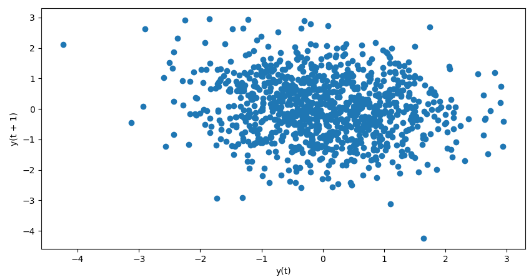

Figure 87: lag-1 plot for 3

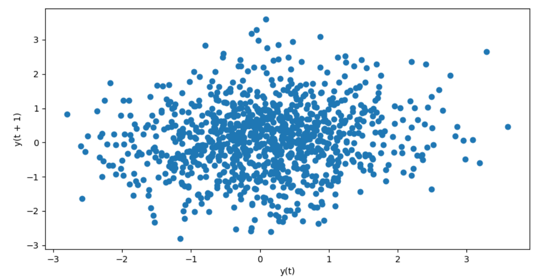

Figure 88: lag-1 plot for 4

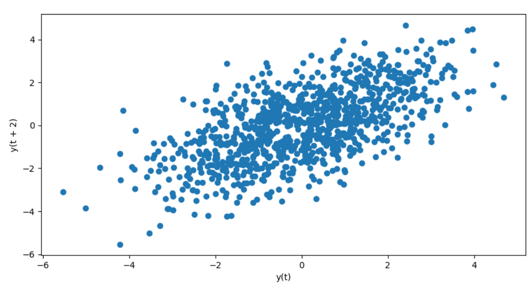

Figure 89: lag-2 plot for 1

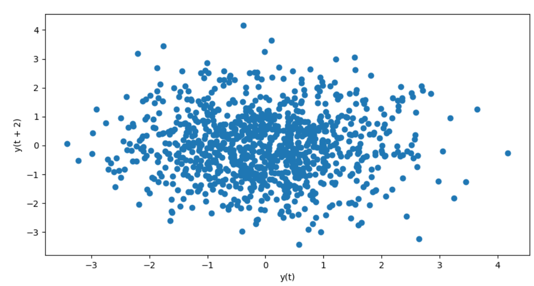

Figure 90: lag-2 plot for 2

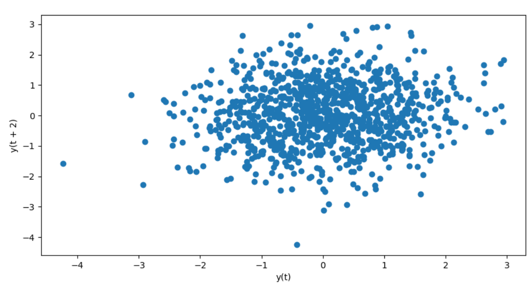

Figure 91: lag-2 plot for 3

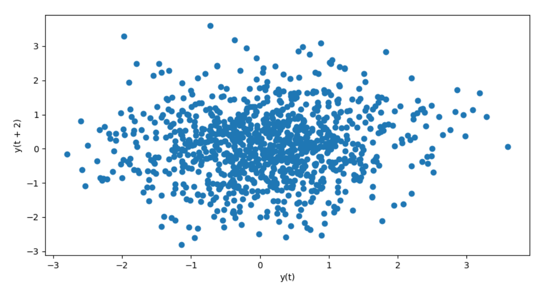

Figure 92: lag-2 plot for 4

The lag-1 and lag-2 plot of 1 show auto-correlation, the rest appears to be spread randomly.

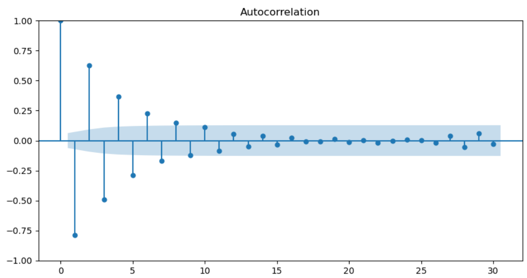

Figure 93: ACF plot for 1

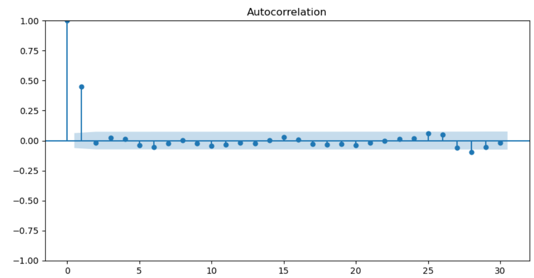

Figure 94: ACF plot for 2

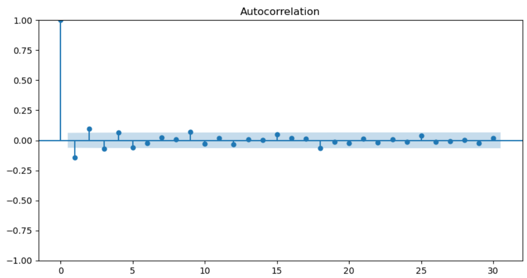

Figure 95: ACF plot for 3

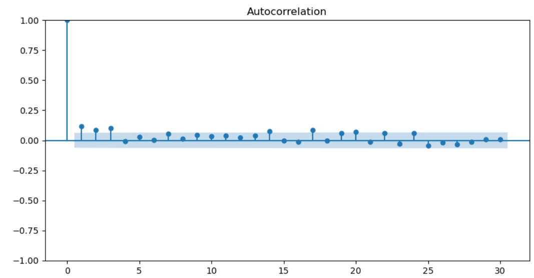

Figure 96: ACF plot for 4

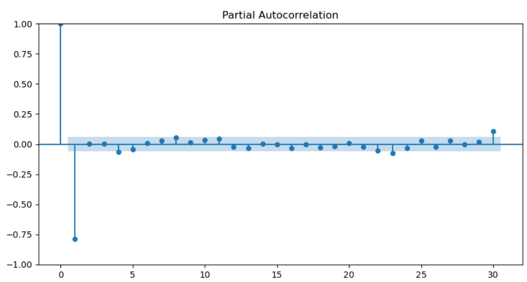

Figure 97: PACF plot for 1

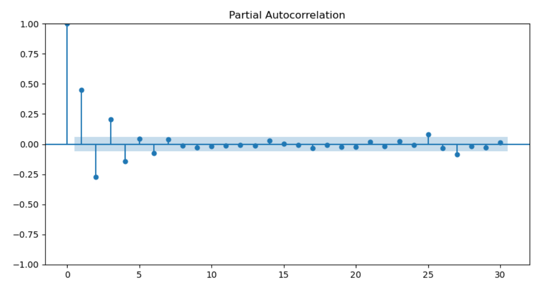

Figure 98: PACF plot for 2

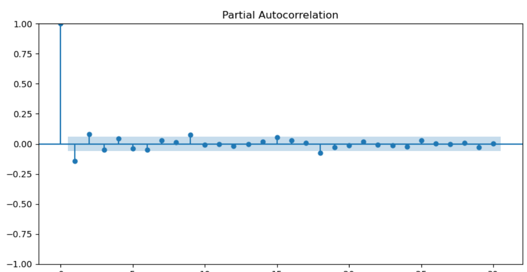

Figure 99: PACF plot for 3

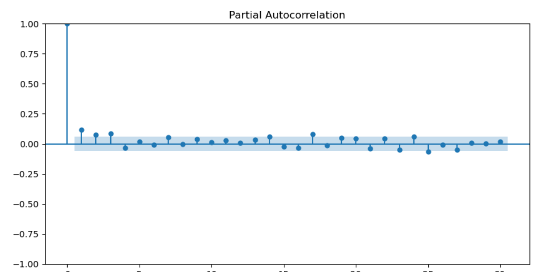

Figure 100: PACF plot for 4

4.1 Questions

What characteristics can you observe from the ACF, PACF graphs of the AR(p) model?
From the PACF graph I can observe that the model is of first order.

What characteristics can you observe from the ACF, PACF graphs of the MA(q) model? From the ACF graph I can observe that the model is of first order.

What characteristics can you observe from the ACF, PACF graphs of the ARMA(p, q) model? From the ACF graph I can observe that the MA part of the model is of first order. For the PACF/AR it is not that clear.

Model	ACF	PACF
AR(p)	subsides	cut of at p
MA(q)	cut of at q	subsides
ARMA(p,q)	subsides	subsides

5 ARIMA modeling and prediction

5.1 Step 1: Randomness test

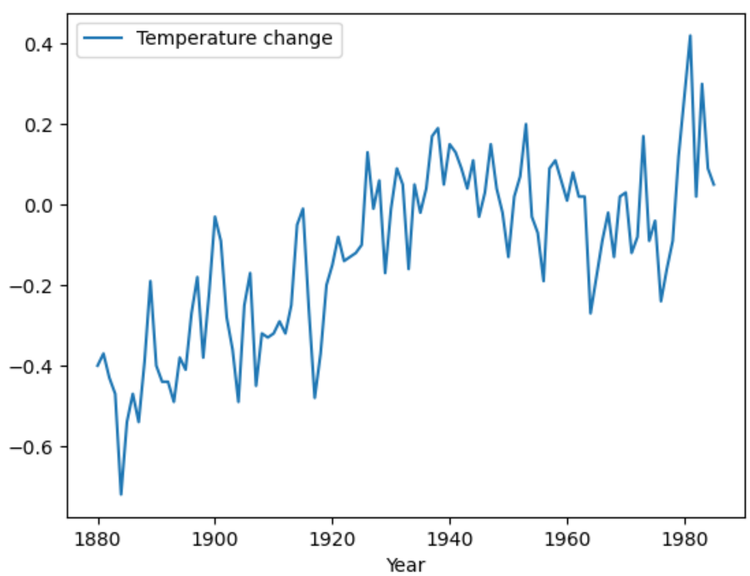

Figure 101: Line plot of Temperature Series Data

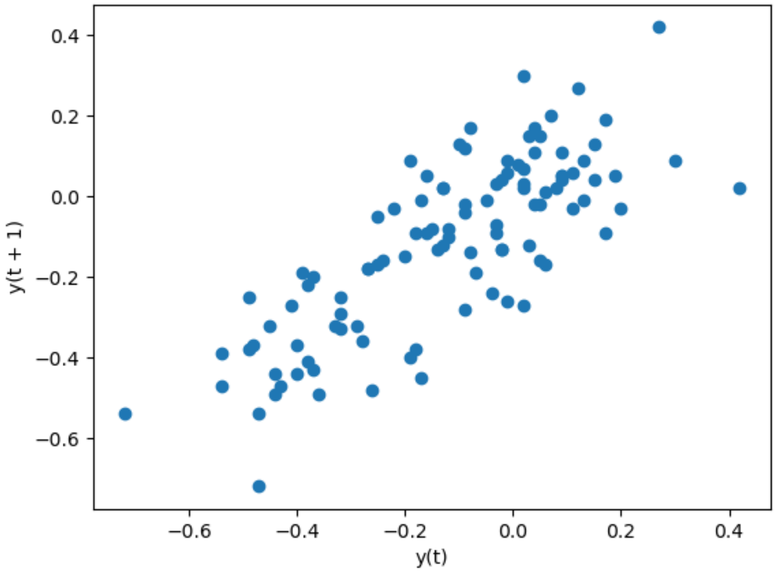

Figure 102: Lag-1 plot of series

Ljung-Box test returns a p value $= 4.7 \times 10^{-16}$
Since this value is less than 0.05, we reject the null hypothesis. There is correlation.

5.2 Step 2: Stationarity test and differencing

Through visual inspection of the line plot, we can tell the data is not stationary as it has an upward trend.

This can be confirmed by the Augmented Dickey-Fuller (ADF) test performed which return a p value = 0.327778. Since this is grater than 0.05, the data is not stationary.

We employ first-order differencing (d = 1) to remove this trend and make the data stationary. The result is plotted below.

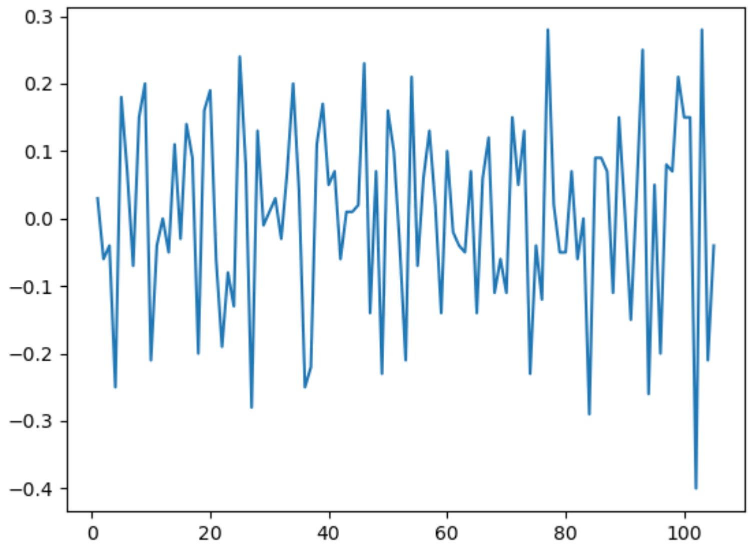

Figure 103: Line plot of differenced series data

ADF test performed on this differenced data returns a p value = 0.0. Thus differenced series is confirmed to be stationary.

5.3 Step 3: Model identification

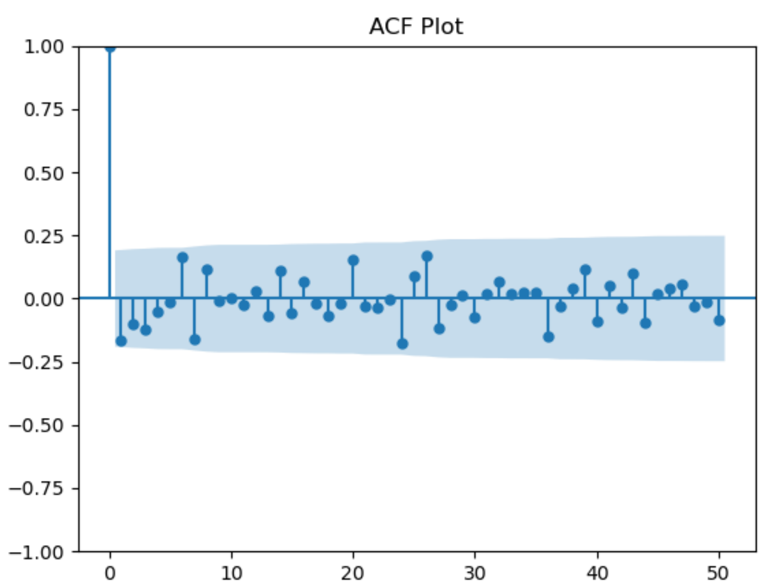

Figure 104: ACF of differenced series

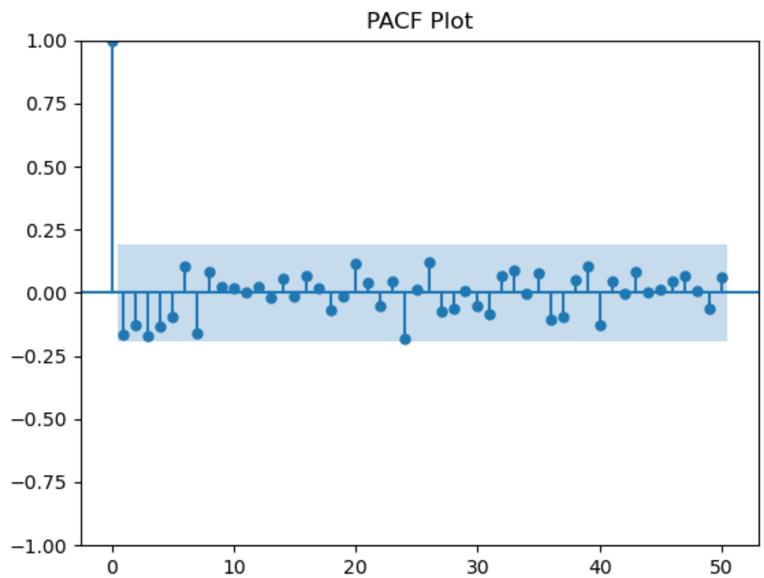

Figure 105: PACF of differenced series

A viable p value can be identified by the PACF plot as the lag points outside of the confidence region

A viable q value can be identified by the ACF plot as the lag points outside of the confidence region

Thus a [p,d,q] value of [1,1,1] was chosen.

5.4 Step 4: Parameter estimation and model optimization

The auto-arima function was used to validate our selection. It also chose [p,d,q] = [1,1,1] with an AIC = -130.7.

Auto-arima conducts a grid search, scanning a range of (p, q) value combinations. The best model with the least AIC will be chosen as the output.

5.5 Step 5: Model validation

We used in-sampling prediction to calculate the residual series.

Ljung-Box test performed returned a p value = 0.885733 . Since this value is greater than 0.05, the data is random.

This was further confirmed by the visual inspection of the lag-1 plot

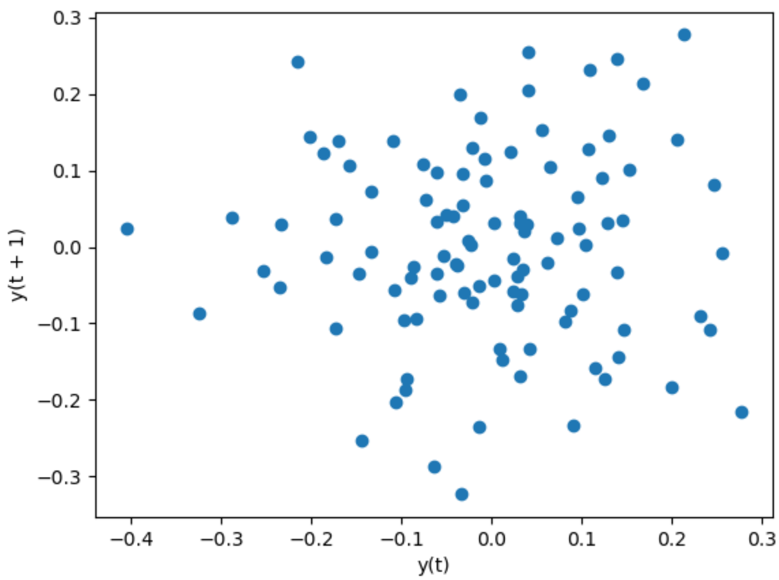

Figure 106: Lag-1 plot of residual series

5.6 Step 6: Model forecasting

We employ out-of-sample predictions for forecasting 10 values.

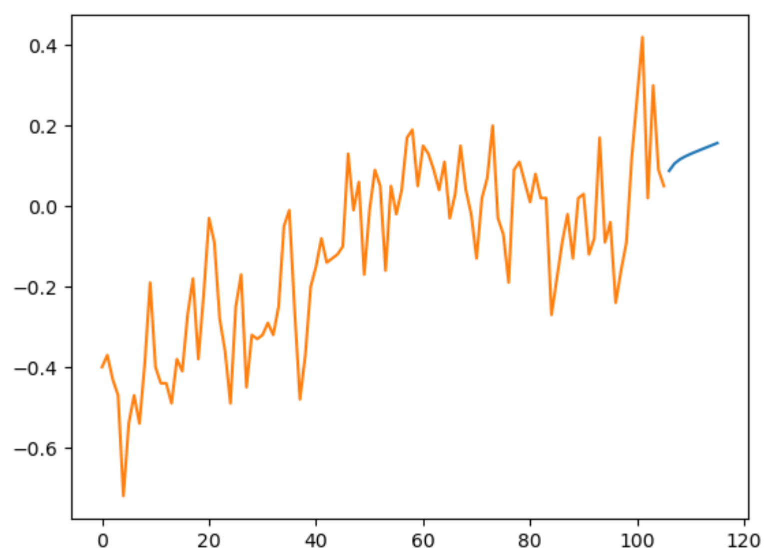

Figure 107: Forecast line plot of temperature series

5.7 Questions:

- Draw a line plot for the time series data. Do you observe any trend, season in the data?

  Line plot indicates that there is an upward trend. Data is not stationary.

- Draw histogram, density plot, heat map, and box plot for the time series data. Are there any outliers? Why?

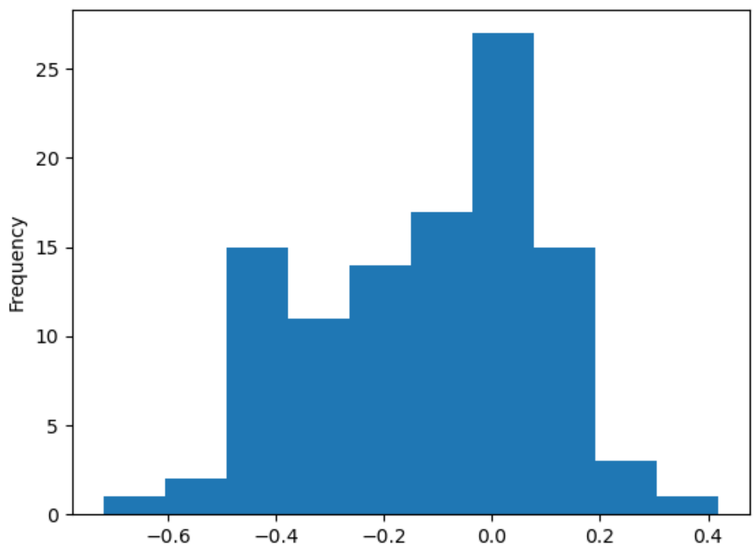

Figure 108: Histogram of time series data

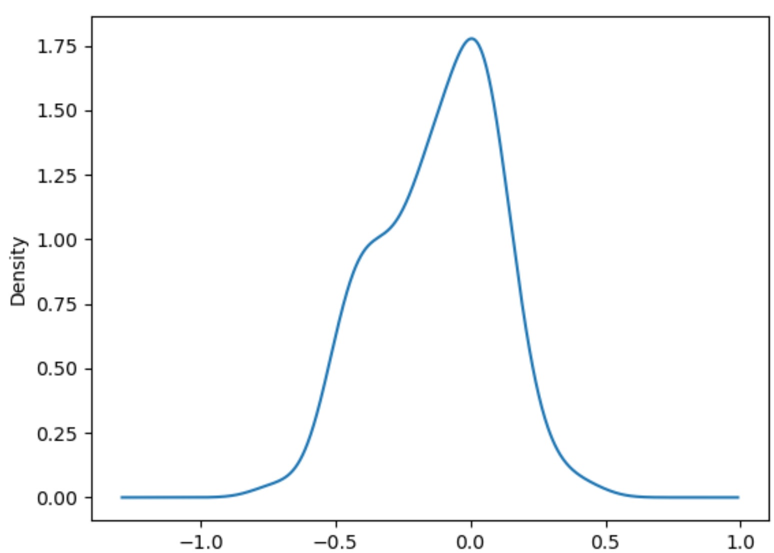

Figure 109: Density of time series data

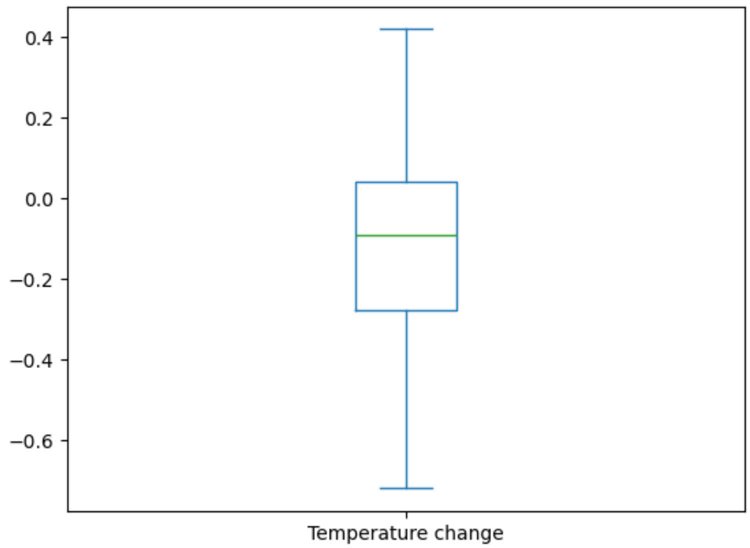

Figure 110: Box plot of time series data

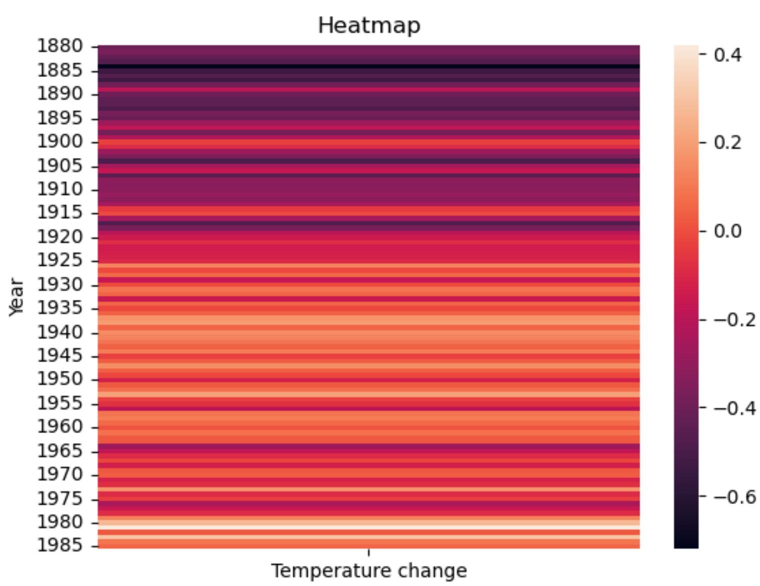

Figure 111: Heat map of time series data

There are no Outliers

- Draw lag-1 and lag-2 plots for the time series data. Do you observe any auto-correlation from the lag plots?

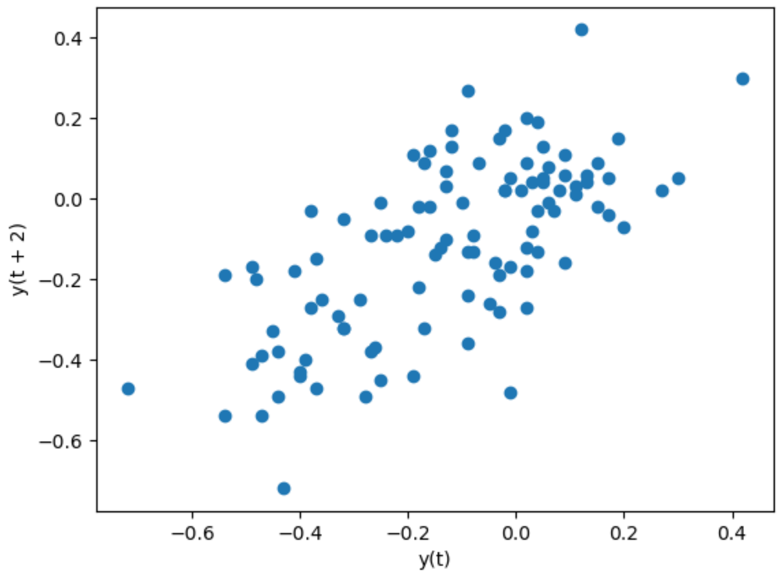

Figure 112: Lag-2 plot of time series data

Refer to Figure 102 for Lag-1 plot

Yes, there is auto-correlation though mild.

- Is the series random? How do you check it? Do the three methods (line plot, lag-1 plot, and Ljung-Box test) give the same results?

  The series is not random. There is clearly a trend in the line plot. The lag plots indicate an auto-correlation though it is not very apparent in the lag plot. The Ljung-Box test returns a p-value ¡ 0.05. Thus, we can reject the null hypthesis and confirm there is autocorrelation

- Is the series stationary? Try with visual inspection and ADF test. Do they give the same results?

  The series is not stationary. The ADF test returns a p value ¿ 0.05. Visual inspection clearly shows an upward trend.

- If the series is not stationary, how do you make it stationary? If you use the differencing operation, how do you decide a proper order of differencing without underdifferencing/over-differencing?

  We use differencing to make the series stationary. We employed different d values and used visual inspection of plots to determine if the data is stationary. Our selection was confirmed through mean, median and standard-deviation statistics.

- What (p, d, q) values do you use? How do you determine them?

  We chose p,d,q values of 1,1,1. d was chosen based on visual insection of plots after differencing. p was chosen based on visual inspection of PACF plots and q was chosen based on visual inspection of ACF plots. The selection was confirmed by the auto-arima function which chose similar values.

- After model fitting, is the remainder series (in-sample prediction) considered to be white noise?

  Yes. For the series to be white noise, it must be stationary and random.

  The remainder series is random. This was confirmed Ljung-Box test gave a p value of greater than 0.05 indicating independent data points. Visual inspection of the lag-1 plot also confirmed the same.

  The remainder series is stationary. This was confirmed by the ADF test which returned a p value of less than 0.05.

- What is the MSE of the fitted model for the data?

  MSE of the fitted model: 0.01699102907355269

- For out-of-sample prediction, do the predicted values (10 steps) reflect the trend and fluctuation of the series?

  The predictions follow the trend without the fluctuations

6 Series transformation

- Q1: What are the common transformation techniques applicable to turn a non-stationary series into a stationary series?

  Common methods are:

  1. Differencing: 1st order differencing where y(t) is subtracted from y(t-1)
  2. Decomposition: Decomposing a time series into its trend, seasonal and cyclical patterns
     (a) Additive Decomposition
     (b) Multiplicative Decomposition
     (c) STL: "Seasonal and Trend decomposition using Loess"
  3. Logarithmic Transformation: Removing logarithmic aspects of the data set
  4. Box-Cox transformation: Used to stabilise non-uniform inconsistent variance

- Q2: What is the Box-Cox transform? Give its definition and explain its generality

  The Box-Cox transformation is a mathematical transformation used to stabilize the variance of a dataset and make it approximately a Gaussian distribution.

  $$z = \begin{cases} \frac{y^\lambda - 1}{\lambda} & \text{if } \lambda \neq 0 \\ \ln(y) & \text{if } \lambda = 0 \end{cases}$$

  y is the original data. λ is the transformation parameter. It can take any real value, including zero. z is the transformed data.

λ is typically selected to maximize the log-likelihood of the transformed data. This can be done by trying different values and selecting the one that results in the most approximately normal distribution of the transformed data.

The Box-Cox transformation is applicable only to strictly positive data. If the data contains zero or negative values, one may need to shift the data or use alternative transformations, such as the Yeo-Johnson transformation, which extends the Box-Cox transformation to handle non-positive data.

- Q3: Can a differencing operation remove a linear trend? Give an example by generating a synthetic series, and draws the series before differencing and after differencing.

  Yes. first order differencing was used in task5 to make the time series stationary.

  Given below we generated a linear trend data set with a random noise component. It is clear that with first order differencing, the linear trend is removed.

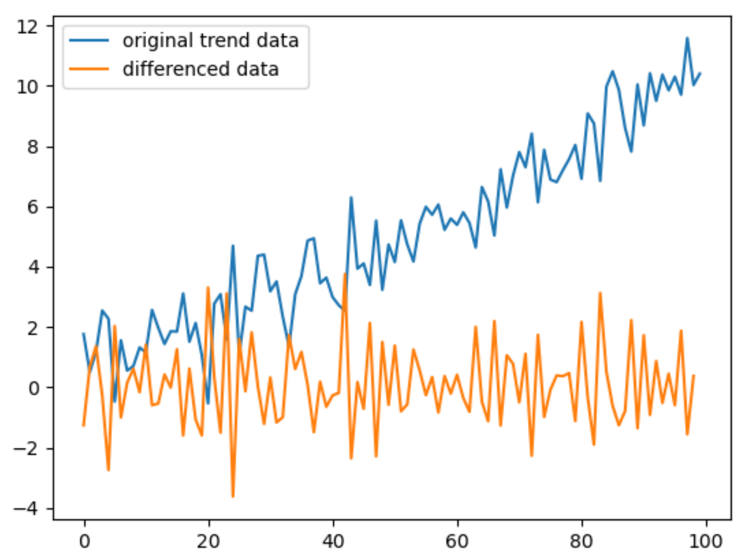

Figure 113: Line plots of linear series data

- Q4: Can a differencing operation remove an exponential trend? If not, which additional transformation needs to be taken? Give an example by generating a synthetic series, and plots the series before transformation and after transformation, before differencing and after differencing

  Differencing cannot remove an exponential trend. However, a logarithmic transformation followed by differencing may work.

  Below, an exponential data set was generated. first order differencing was applied on this data. Additionally, logarithmic transformation was also applied on this data.

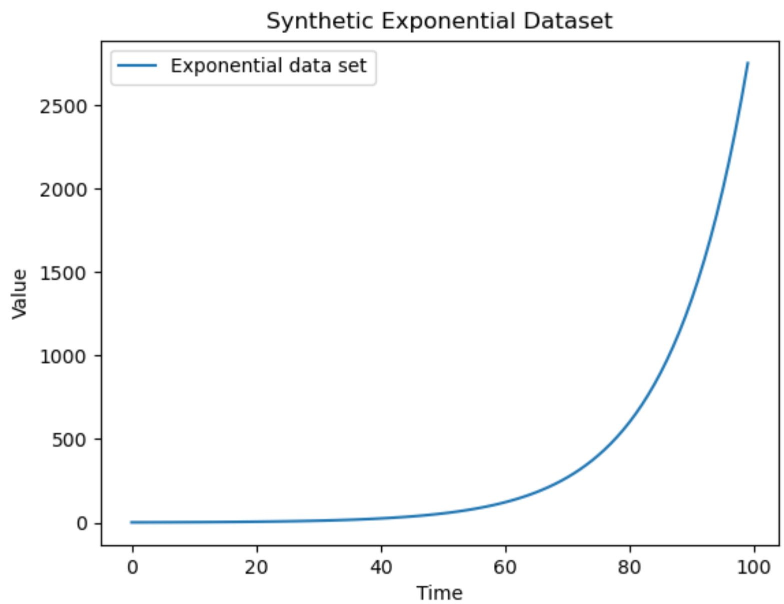

Figure 114: Line plots of exponential series data

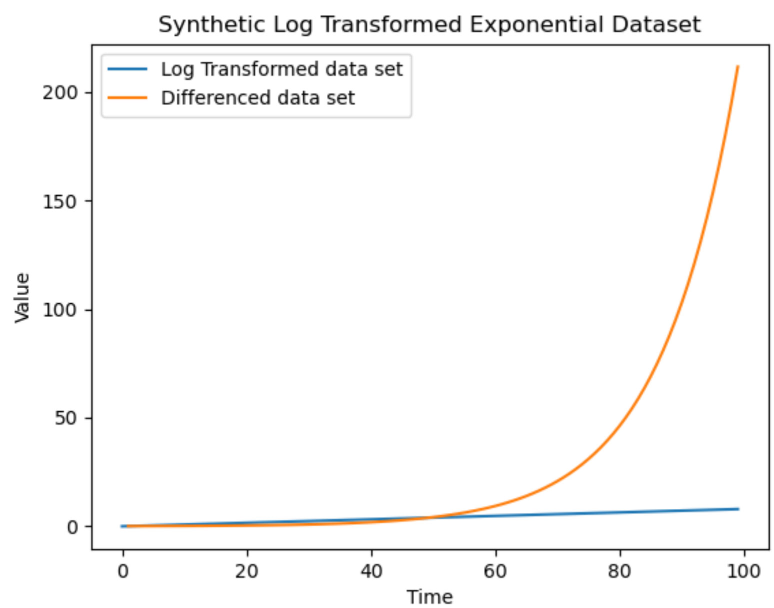

Figure 115: Line plots of transformed / differenced exponential series data

- Q5: Can a differencing operation remove a seasonal (periodic) trend? If yes, under what condition? Give an example by generating a synthetic series, draw its ACF, and draws the series before differencing and after differencing with different step length.

  Yes, possibly. If we know the periodicity of the seasonal effect, we can difference with a lag equivalent to that periodicity.

Given below, we generated a seasonal dataset (n = 12) with a linear upward trend line. first order differencing removes the linear trend. Lag-12 differencing removes the seasonal characteristic of the data.

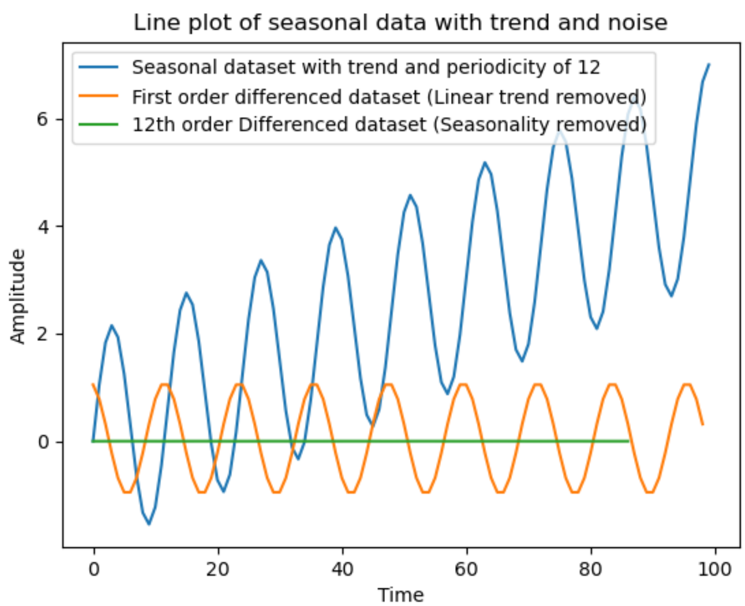

Figure 116: Line plots of seasonal and differenced series data

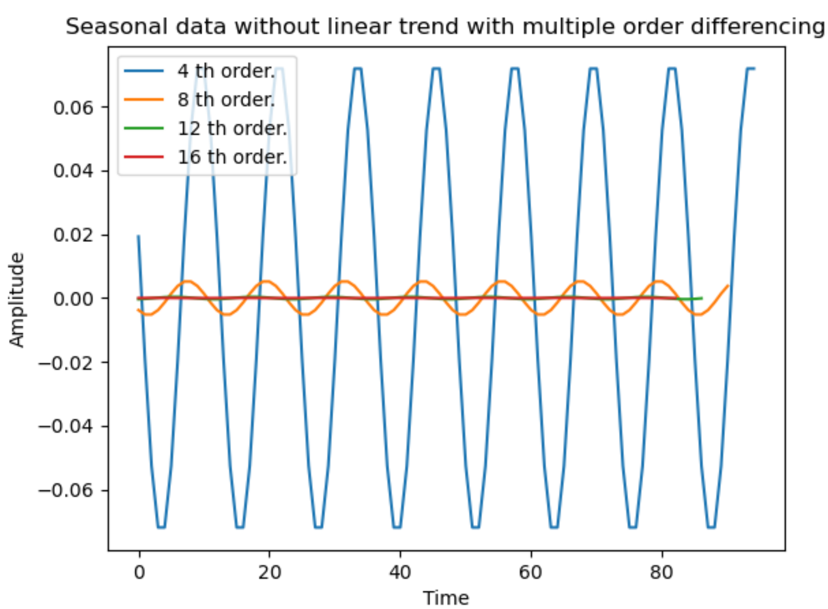

Figure 117: Line plots of differenced series data for multiple orders

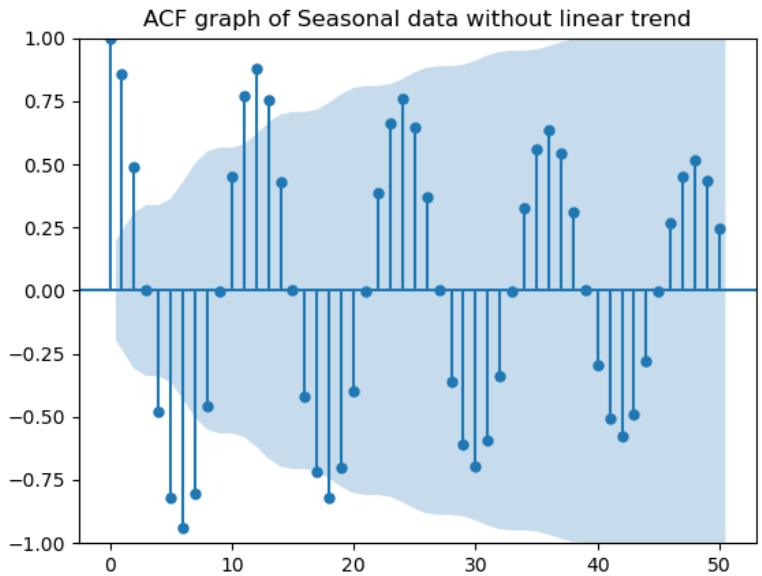

Figure 118: ACF of seasonal series data