

Time Series Analysis
Homework Assignment 3
Submitted by Sruthy Annie Santhosh,312213

Autoregressive Processes

Exercise 1:

Let $\{Z_t\}$ be a sequence of independent normal random variables, each with mean 0 and variance σ^2 . Which, if any, of the following processes are causal? For each causal process, calculate the first three coefficients ψ_1, ψ_2, ψ_3 in the representation of an AR(p) process as the infinite sum of white noise Z_{t-j} .

We know that from the slides,

Theorem 5.6

An $AR(p)$ process $(X_t)_{t \in \mathbb{Z}}$ is causal if

$$\Phi(x) := 1 - \phi_1 x - \phi_2 x^2 - \dots - \phi_p x^p$$

has no root inside the complex unit circle. It then has a representation

$$X_t = \sum_{j=0}^{\infty} \psi_j Z_{t-j},$$

where $\sum_{j=0}^{\infty} |\psi_j| < \infty$ and $\sum_{j=0}^{\infty} \psi_j^2 < \infty$ and the coefficients are defined recursively by $\psi_k = 0$ for $k < 0$, $\psi_0 = 1$ and

$$\psi_j = \sum_{i=1}^p \phi_i \psi_{j-i}, \quad j \geq 1.$$



A sufficient condition for an $AR(p)$ process to be causal is

$$\sum_{k=1}^p |\phi_k| < 1.$$

$$(1) X_t = 2.25X_{t-1} - 0.5X_{t-2} + Z_t$$

Writing in polynomial form,

$$\Phi(x) = 1 - 2.25x + 0.5x^2$$

The solutions of the polynomial $\Phi(x) = 0$ are $x_1 = 0.5-0i$ and $x_2 = 4.0+0i$

The roots of the function can be found using the `polyroot()` function in R.

Here the root x_1 is less than 1 i.e., is inside the complex unit circle. Hence the process $\{X_t\}$ is not causal.

```
> polyroot(c(1, -2.25, 0.5))
```

```
[1] 0.5-0i 4.0+0i
```

$$(2) X_t = X_{t-1} - 0.21X_{t-2} + Z_t .$$

Writing in polynomial form,

$$\Phi(x) = 1 - x + 0.21 x^2$$

The solutions of the polynomial $\Phi(x) = 0$ are $x_1 = 1.428571+0i$ and $x_2 = 3.333333-0i$ as found by the polyroot() function.

```

> polyroot(c(1, -1, 0.21))
[1] 1.428571+0i 3.333333-0i

```

Here the roots are greater than 1 i.e, both roots are outside the complex unit circle. Hence the process $\{X_t\}$ is causal.

We know that $p=2$, $\phi_1 = 1$ and $\phi_2 = -0.21$.

From the theorem given above, we can find the first three coefficients ψ_1, ψ_2, ψ_3 in the representation of an AR(p) process as the infinite sum of white noise Z_{t-j} .

$$\psi_1 = \sum_{i=1}^p \phi_i \psi_{1-i} = \sum_{i=1}^2 \phi_i \psi_{1-i} = \phi_1$$

i.e, $\psi_1 = 1$

$$\psi_2 = \sum_{i=1}^p \phi_i \psi_{2-i} = \sum_{i=1}^2 \phi_i \psi_{2-i} = \phi_1 \psi_1 + \phi_2 \psi_0$$

i.e, $\psi_2 = 1 + -0.21*1 = 0.79$

$$\psi_3 = \sum_{i=1}^p \phi_i \psi_{3-i} = \sum_{i=1}^2 \phi_i \psi_{3-i} = \phi_1 \psi_2 + \phi_2 \psi_1$$

i.e, $\psi_3 = 0.79 + -0.21 = 0.58$

Exercise 2:

1. Simulate a time series of length 1000 for the following model, giving appropriate R code and placing the simulated data in a vector `x`:

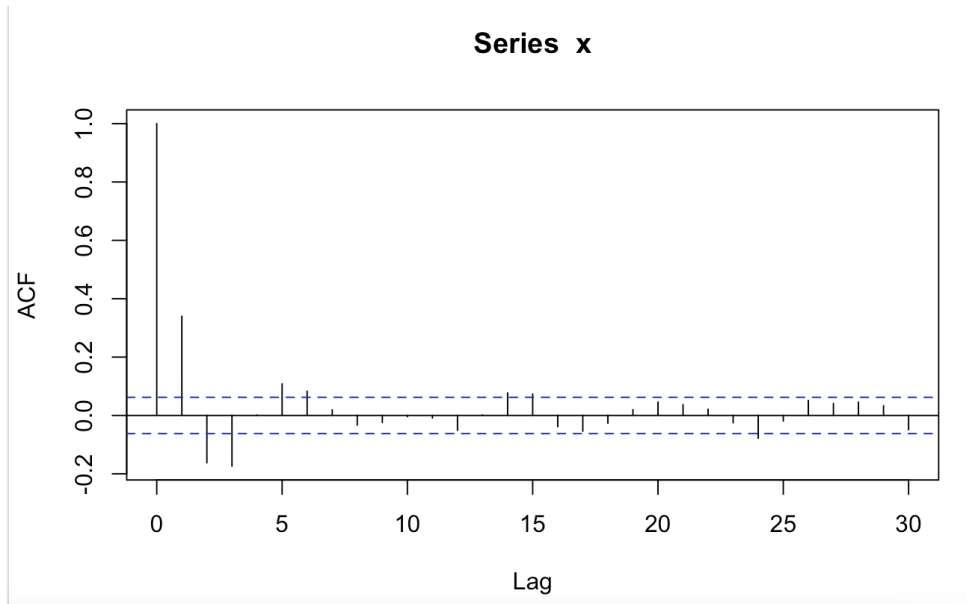
$X_t = 0.4X_{t-1} - 0.3X_{t-2} + Z_t$, where $\{Z_t\}$ be a sequence of independent normal random variables, each with mean 0 and variance 1.

A sample of the `x` values generated

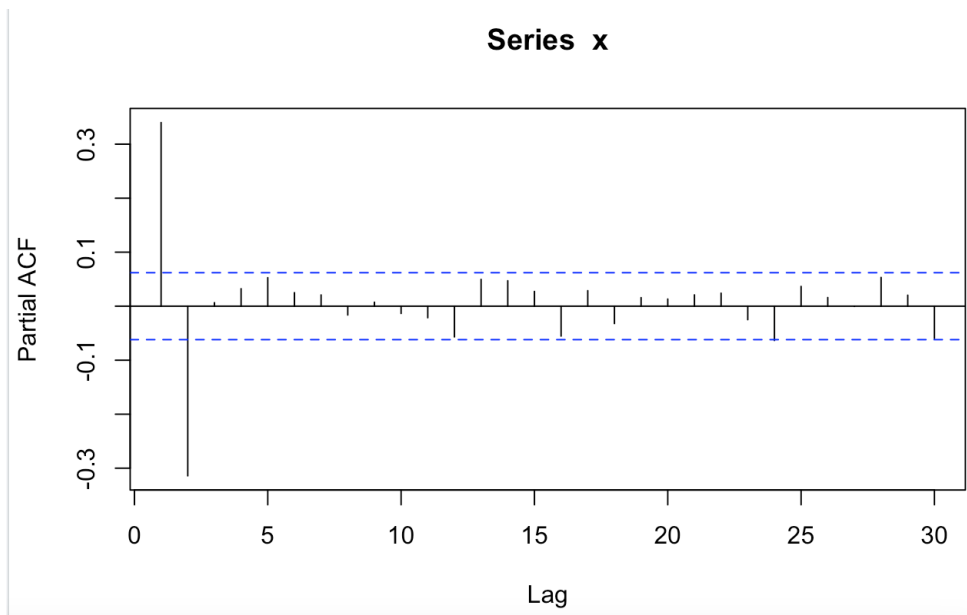
```
x
[1] 0.0187461709 -0.1842525421 -1.4506554180 -1.1241541204 0.2800801038 0.8390725783
[7] -0.9564711752 -0.9979862611 -1.7389258336 -0.6526528492 1.3623961135 1.4965358082
[13] -0.0483380668 0.5191487342 0.9635510421 0.3190230631 -1.1163999436 -0.7374172810
[19] 0.9654743328 1.0903934422 -0.4497955597 -2.6923230947 -1.6168565079 -1.9581068666
[25] -1.5633838158 -0.4115830215 -0.3831734942 -0.9019533180 -0.3475902851 -0.1222306488
[31] -1.7983556284 -0.7606191228 1.2038253799 0.8946418488 -1.3832344528 -2.2572006981
[37] -0.1258227148 -1.1322556302 -0.7396994472 -0.6077660784 1.0653548023 -0.1540731356
[43] -1.2098982299 0.3967365518 -0.4459878971 -0.3262314598 0.2358289378 -0.1090076684
[49] -0.7919663319 0.3711434040 -0.0145902859 -0.4517357006 1.1916367587 2.7499425173
[55] 1.2483052438 0.4606817266 -1.0924308267 -0.0422798563 -0.3350769481 0.1696406661
[61] -1.0692151180 -0.9347545222 -0.8834599282 0.2671580290 1.4382775858 1.7112894637
[67] 0.9887231674 -0.5991061895 0.0264853369 -1.0559937203 -0.0494208766 -1.1333974873
[73] -1.4869782368 -0.4732755991 -1.2331530053 0.8214277588 -0.7813100165 -0.9893401503
[79] -1.2129816972 1.3341957103 1.4904008479 -0.0267598830 0.2550700687 0.8266568261
[85] 0.6943835742 0.1885870032 0.5268838673 2.3746971089 -0.3921313907 -0.9432175235
[91] -0.6760022671 -0.1789179935 0.2007782641 1.2893350217 1.0504578764 -1.3862624643
[97] -2.4763195940 0.3182768013 1.0183745539 1.5388951713 -0.4517586368 -0.2229966002
[103] -0.9936144137 0.3810271806 -0.1827178186 0.3757793828 0.8661137845 -1.4243391584
[109] 0.1985981783 1.6346946328 -0.6858562038 0.3641173560 -0.1127307236 -0.4700877058
```

2. Plot the correlogram and partial correlogram for the simulated data. Comment on the results.

ACF plot



PACF plot



From the above two plots we can observe that while in ACF plot, many lags can be considered significant, in PACF plot only the first two lags can be considered as significant. This can mean that the underlying process is AR(2).

3. Fit an AR model to the data in x giving the parameter estimates and order of the fitted AR process.

```

Call:
ar(x = x, method = "mle")

Coefficients:
      1      2
0.4464 -0.3136

Order selected 2  sigma^2 estimated as  0.9804

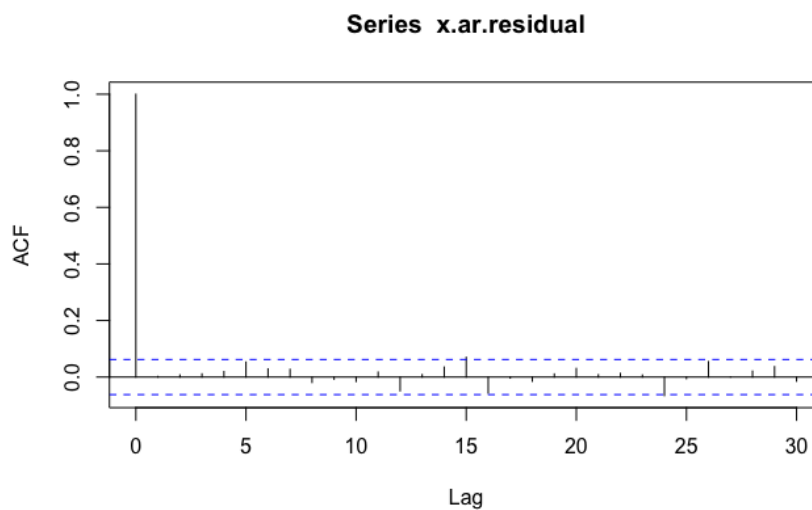
```

The order of the fitted process is 2 and the variance is 0.9804

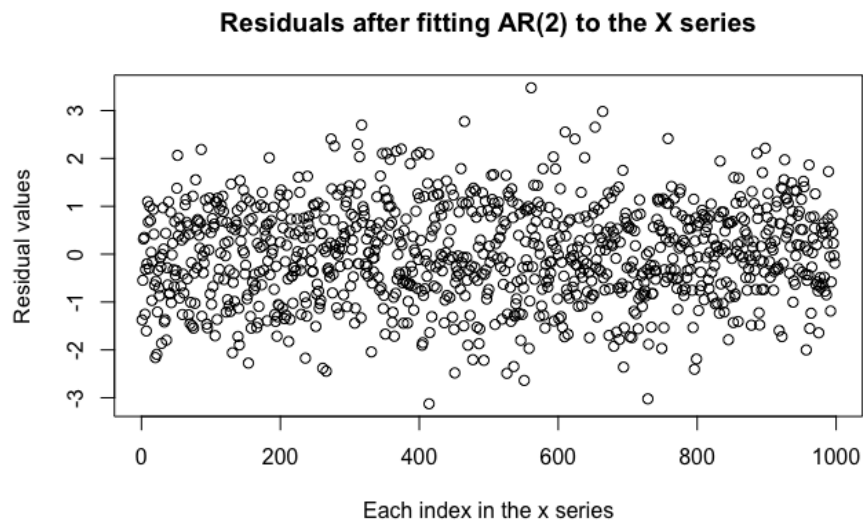
So we get a fitted AR(2) Model with $X_t = 0.4464 \cdot X_{t-1} - 0.3136 \cdot X_{t-2} + Z_t$ where $\{Z_t\}$ is a sequence of independent normal random variables each with mean 0 and variance 0.9804

4. Perform two tests for IID on the residuals of the fitted AR process, as well as one test for Gaussianity. Comment on the results.

The ACF of the residuals of the fitted AR process



Plotting the residuals after fitting AR(2) to the X series



Performing the Box-Pierce test for testing IID hypothesis. Here we are taking the threshold value as 0.05. A p-value of 0.05 or lower is generally considered statistically significant.

```
> #for lag = 1 to 5
> for(i in 1:5){ print(Box.test(x.ar.residual, lag = i, type = "Box-Pierce"))}
```

Box-Pierce test

```
data: x.ar.residual
X-squared = 0.0047606, df = 1, p-value = 0.945
```

Box-Pierce test

```
data: x.ar.residual
X-squared = 0.062943, df = 2, p-value = 0.969
```

Box-Pierce test

```
data: x.ar.residual
X-squared = 0.19177, df = 3, p-value = 0.9789
```

Box-Pierce test

```
data: x.ar.residual
X-squared = 0.56678, df = 4, p-value = 0.9667
```

Box-Pierce test

```
data: x.ar.residual
X-squared = 3.3742, df = 5, p-value = 0.6425
```

For the Box-Pierce test, the p value is always above the threshold value of 0.05. The Box-Pierce test does not detect a significant contradiction with the IID hypothesis. But this does not prove the IID hypothesis.

Performing the turning point test

```
> turning.point.test(x.ar.residual)
```

Turning Point Test

```
data: x.ar.residual  
statistic = 0.37572, n = 998, p-value = 0.7071  
alternative hypothesis: non randomness
```

We can see that the value for turning point test statistic is 0.37572 and the corresponding p-value is 0.7071. This p value is greater than a threshold $\alpha = 0.05$. Hence we cannot reject the i.i.d hypothesis at this level 0.05. But this also does not prove the hypothesis.

Performing the Shapiro-Wilk Test for Gaussianity

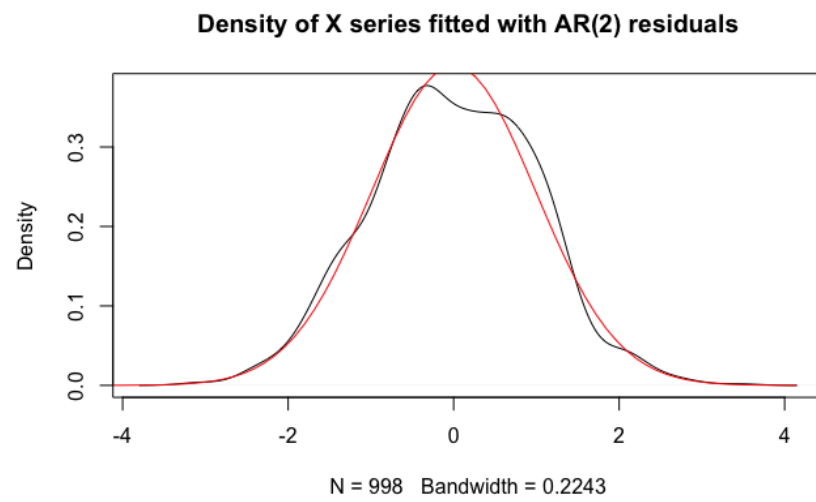
```
> #Shapiro test for Gaussianity  
> shapiro.test(x.ar.residual)
```

Shapiro-Wilk normality test

```
data: x.ar.residual  
W = 0.99803, p-value = 0.2949
```

From the output, the p-value > 0.05 shows that we fail to reject the null hypothesis, which means the distribution of our data is not significantly different from the normal distribution. In other words distribution of our data is normal.

We can also check the gaussianity using the dnorm.



This also shows a normal distribution.