INDIAN INSTITUTE OF TECHNOLOGY MADRAS

Department of Chemical Engineering

CH5350 : Applied Time Series Analysis
Solutions to Assignment #2

1 Given $x[k] = \phi_1 x[k-1] + e[k]$ where $e[k] \sim \mathcal{N}(0, \sigma_e^2)$ and $|\phi_1| < 1$

1.1

$$\text{corr}(x[k], x[k-l]) = \frac{\sigma_{xx}[l]}{\sqrt{\text{var}(x[k])\text{var}(x[k-l])}}$$
$$\sigma_{xx}[l] = \text{E}((x[k] - \mu_{x[k]})(x[k-l] - \mu_{x[k-l]}))$$

Rewriting $x[k]$ in terms of $x(k-2)$ will give

$$x[k] = \phi_1^2 x[k-2] + \phi_1 e[k-1] + e[k]$$

Extending the series to $x[k-l]$ gives

$$\begin{aligned} x[k] &= \phi_1^l x[k-l] + \phi_1^{l-1} e[k-l-1] + ... + e[k] \\ &= \phi_1^l x[k-l] + \sum_{r=0}^{l-1} \phi^r e[k-r] \end{aligned}$$

The mean of the sample $x[k]$ in terms of $x[k-l]$ is $\mu_{x[k]} = \phi_1^l \mu_{x[k-l]}$

$$\begin{aligned} \sigma_{xx}[l] &= \text{E}((x[k] - \phi_1^l \mu_{x[k-l]})(x[k-l] - \mu_{x[k-l]})) \\ &= \text{E}((\phi_1^l x[k-l] + \phi_1^{l-1} e[k-l+1] + ... + e[k] - \phi_1^l \mu_{x[k-l]})(x[k-l] - \mu_{x[k-l]}) \\ &= \text{E}((\phi_1^l (x[k-l] - \phi_1^l \mu_{x[k-l]})(x[k-l] - \mu_{x[k-l]})) \\ &= \phi_1^l \text{var}(x(k-l)) \end{aligned}$$

Now,

$$\operatorname{corr}(x[k], x[k-l]) = \frac{\phi_1^l \operatorname{var}(x(k-l)}{\sqrt{\operatorname{var}(x[k])\operatorname{var}(x[k-l])}}$$

$$\boxed{\operatorname{corr}(x[k], x[k-l]) = \frac{\phi_1^l \sqrt{\operatorname{var}(x(k-l)}}{\sqrt{\operatorname{var}(x[k])}}}$$

1.2

As given in solution of problem 1.1, $x[k]$ can be written in terms of $x[0]$ as

$$x[k] = \phi_1^k x[0] + e[k] + \phi_1 e[k-1] + \phi_1^2 e[k-2] + ...$$

Given $|\phi_1| < 1$, Hence $x[0]$ term in the above series will vanish as $k \to \infty$. Hence

$$\boxed{x[k] = \sum_{r=0}^{\infty} (\phi_1^r) e[k-r]}$$

1.3

It has been proved in the solution of problem 1.2 that

$$x[k] = \sum_{r=0}^{\infty} (\phi_1^r) e[k-r]$$

For the process to be stationary it should satisfy the following two conditions

1. The mean should be finite.

2. The autocorrelation function should only depends on lag.

The mean of the process is

$$\mathrm{E}(x[k]) = \mathrm{E}\left(\sum_{r=0}^{\infty} (\phi_1^r) e[k-r]\right) = \sum_{r=0}^{\infty} \phi_1^r \mathrm{E}\left(e[k-r]\right) = 0$$

The auto covariance of the process is given by

$$\operatorname{cov}\left(x[k], x[k-l]\right) = \mathrm{E}\left(\left(\sum_{r=0}^{\infty} (\phi_1^r) e[k-r]\right)\left(\sum_{m=0}^{\infty} (\phi_1^m) e[k-m-l]\right)\right)$$

which can be written as

$$\begin{aligned}\text{cov}\left(x[k], x[k-l]\right) &= \mathrm{E}\left(\left(e[k] + \phi_1 e[k-1] + \phi_1^2 e[k-2] + ...\right)\left(e[k-l] + \phi_1 e[k-l-1] + \phi_1^2 e[k-l-2] + ...\right.\right.\\ &= (\phi_1^l \sigma_e^2 + \phi_1^{l+2}\sigma_e^2 + \phi_1^{l+4}\sigma_e^2 + ...)\\ &= \frac{\phi_1^l \sigma_e^2}{1-\phi_1^2}\end{aligned}$$

Hence, regardless of initial conditions always an AR process will be asymptotically stable.

1.4

$$x[k] = \phi_1^k x[0] + \sum_{r=0}^{\infty}(\phi_1^r)e[k-r]$$

The mean of the process is

$$\mu_{x[k]} = \phi_1^k \mu_{x[0]}$$

. The ACVF of the process is

$$\text{cov}\left(x[k], x[k-l]\right) = \phi_1^{2k-l}\sigma_{x[0]}^2 + \frac{\phi_1^l \sigma_e^2}{1-\phi_1^2}$$

Both mean and ACVF of $x[k]$ depends on mean and variance of $x[0]$. For the system to be stationary independent of initial conditions, then $x[0]$ should come from normal distribution with mean 0 and variance $\frac{\sigma_e^2}{1-\phi_1^2}$.

2

2.1

Given MA(2) process $H(q^-1) = 1 + c_1 q^{-1} + C_2 q^{-2}$. The ACVF generating function is given by

$$g_\sigma(z) = \sum_{l=-\infty}^{\infty} \sigma_{vv}[l] z^{-l}$$

In class, it is proved that

$$g_\sigma(z) = \sigma_e^2 H(z^{-1})H(z)$$

Hence,

$$
\begin{aligned}
g_\sigma(z) &= \sigma_e^2(1 + c_1 z^{-1} + c_2 z^{-2})(1 + c_1 z + c_2 z^2) \\
&= \sigma_e^2(c_2 z^{-2} + (c_1 + c_1 c_2) z^{-1} + (1 + c_1^2 + c_2^2) + (c_1 + c_1 c_2) z + c_2 z^2)
\end{aligned}
$$

The ACVF of the series at various lags is given by respective coefficients as

$$
\sigma_{vv}[l] = \begin{cases} (1 + c_1^2 + c_2^2)\sigma_e^2 & l = 0 \\ (c_1 + c_1 c_2)\sigma_e^2 & l = \pm 1 \\ c_2 \sigma_e^2 & l = \pm 2 \\ 0 & |l| > 2 \end{cases}
$$

2.2

The ACF of the series is given by

$$
\rho_{vv}[l] = \begin{cases} 1 & l = 0 \\ \frac{(c_1 + c_1 c_2)}{(1 + c_1^2 + c_2^2)} & l = \pm 1 \\ \frac{c_2}{(1 + c_1^2 + c_2^2)} & l = \pm 2 \\ 0 & |l| > 2 \end{cases}
$$

It is to be admissible if and only if the values at all lags ($l \neq 0$) should be less than unity. *i.e*

$$
\left| \frac{(c_1 + c_1 c_2)}{(1 + c_1^2 + c_2^2)} \right| < 1 \quad \text{and} \quad \left| \frac{c_2}{(1 + c_1^2 + c_2^2)} \right| < 1
$$

Imposing condition of Invertibility the constraint will be

$$
\left| \frac{-c_1 \pm \sqrt{(c_1^2 - 4c_2)}}{2c_2} \right| > 1
$$

Solving the above equation, the constraints are

$$
\begin{gathered}
c_1 + c_2 > -1 \\
c_1 - c_2 < 1 \\
-1 < c_2 < 1
\end{gathered}
$$

The admissible region for coefficients c_1 and c_2 is shown in Figure 1.

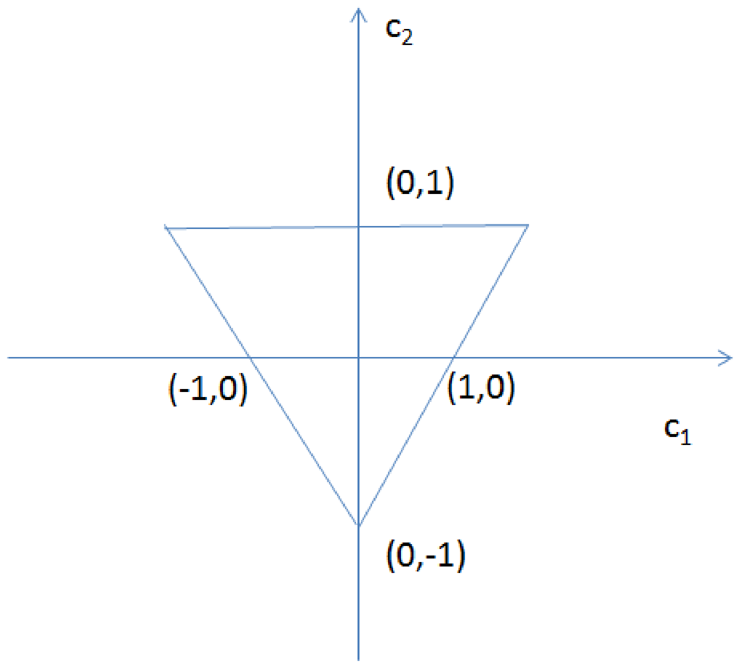

Figure 1: Admissible region for coefficients c_1 and c_2

2.3

Given

$$H(q^{-1}) = \frac{1}{1 - 1.3q^{-1} + 0.4q^{-2}}$$

. Using Yule-Walker equations, the auto-correlation of the series is calculated as

$$\rho_{yy}[l] = \begin{cases} 1 & l = 0 \\ 0.9286 & l = \pm 1 \\ 0.8071 & l = \pm 2 \\ 0.6778 & l = \pm 3 \end{cases}$$

The following table compares the all theoretical values as well as values given by both acf and ARMAacf routine in R.

lag	Theoretical Value	using ACF	using ARMAacf
1	0.926	0.9304	0.9285
2	0.8071	0.8166	0.80714

Table 1: Values of ACF for the given AR(2) process

3 Given $y(k) = A\sin(2\pi f_0 k) + e[k]$ and $e[k] \sim \mathcal{N}(0,1)$

3.1 Stationary

For the process to be stationary it should satisfy the following two conditions

1. The mean should be finite.

2. The autocorrelation function should only depends on lag.

The mean of the process is given by

$$\begin{aligned} E(y[k]) &= E(A\sin(2\pi f_0 k) + e[k]) \\ &= A\sin(2\pi f_0 k) \end{aligned}$$

which is a function of time k. Hence, the process is not stationary.

3.2 Sample autocorrelation

$$\rho_{yy}[l] = \lim_{N\to\infty} \frac{1}{N} \sum_{k=l+1}^{N} (y[k] - \bar{y})(y[k-l] - \bar{y})$$

Here $\bar{y} = \sum_{k=1}^{\infty} y[k]$. Since, it is periodic wave the value of sample mean will be zero. Then

$$\rho_{yy}[l] = \lim_{N\to\infty} \frac{1}{N} \sum_{k=l+1}^{N} (A\sin(2\pi f_0 k) + e[k]))(A\sin(2\pi f_0 k - l) + e[k-l])$$

Since, $\sin(2\pi f_0 k)$ is uncorrelated with $e[k]$,

$$\begin{aligned} \rho_{yy}[l] &= \lim_{N\to\infty} \frac{1}{N} \sum_{k=l+1}^{N} (A\sin(2\pi f_0 k))(A\sin(2\pi f_0 (k-l))) + e[k]e[k-l] \\ &= \lim_{N\to\infty} \frac{1}{N} \sum_{k=l+1}^{N} (A^2 \sin^2(2\pi f_0 k) cos(2\pi f_0 l)) - (A^2 \sin(2\pi f_0 k)\cos(2\pi f_0 k)\sin(2\pi f_0 l)) + \sigma_{ee}[l] \\ &= \lim_{N\to\infty} \frac{1}{N} \sum_{k=l+1}^{N} (0.5A^2(1 - cos(4\pi f_0 k)))\cos(2\pi f_0 l) - 0.5A^2 \sin(2\pi f_0 l)(\sin(4\pi f_0 k)) + \sigma_{ee}[l] \\ &= 0.5A^2\cos(2\pi f_0 l) \lim_{N\to\infty} \frac{1}{N} \sum_{k=l+1}^{N} (1 - cos(4\pi f_0 k)) - 0.5A^2 \sin(2\pi f_0 l) \lim_{N\to\infty} \frac{1}{N} \sum_{k=l+1}^{N} (\sin(4\pi f_0 k)) + \sigma_{ee}[l \\ &= \frac{A^2}{2}\cos(2\pi f_0 l) + \sigma_{ee}[l] \end{aligned}$$

$$\implies \boxed{\rho_{yy}[l] = \frac{A^2}{2}\cos(2\pi f_0 l) \quad \forall \quad l > 0}$$

Hence, ACF of sinusoidal signal is periodic with the same frequency as that of the signal.

3.3

It is very advantageous to detect periodicities from ACF of periodic signal. By visual inspection, we can't tell whether the signal is periodic. But, from ACF plot we can say whether the signal is periodic and about the period of the signal. For example the above sine series added with noise is shown in Figure 2.

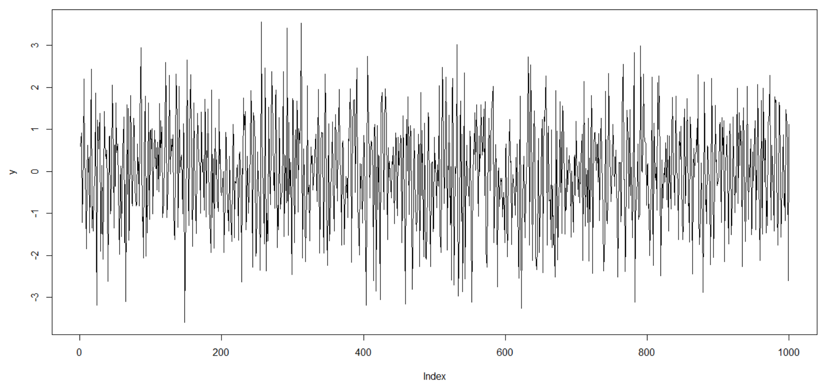

Figure 2: Sinusoidal signal added with noise

3.4

Figure 3 shows the ACF of the series shown in Figure 2

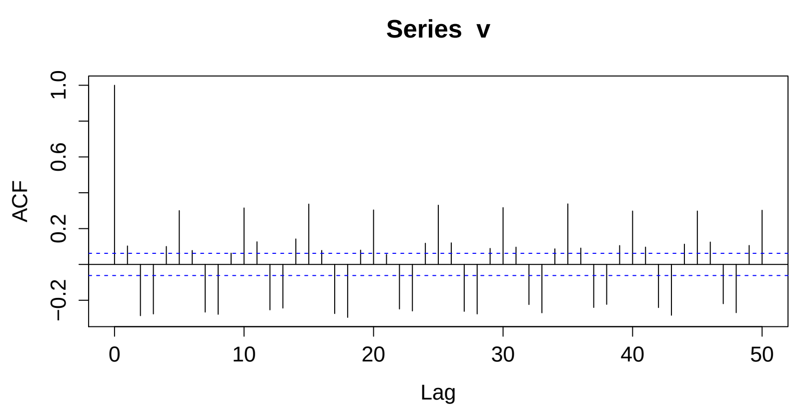

Figure 3: Acf of sine signal added with noise

From Figure 3, it is evident that ACF is periodic with the same frequency as that of the signal.

4

4.1

In class we have derived PACF of a process for two lags as

$$\phi_{11} = \rho_1$$
$$\phi_{22} = \frac{\rho_2 - \rho_1^2}{1 - \rho_1^2}$$

4.1.1 $x_1[k] = 0.7x_1[k-1] - 0.12x_1[k-2] + e[k]$

For the above process, $\rho_1 = 0.6250$ and $\rho_2 = 0.3175$.
So, PACF coefficients will be $\phi_{11} = 0.6250$ and $\phi_{22} = -0.12$
PACF coefficients calculated using ARMAacf in R are given as $\phi_{11} = 0.6250$ and $\phi_{22} = -0.120$. In class we have learnt that the value of PACF coefficient at lag p will be last coefficient of AR(p) process. The value of $\phi_{22} = -0.12$ which is the last coefficient of given process.

4.1.2 $x_2[k] = e[k] + 0.4e[k-1]$

For the above process $\rho_1 = \frac{c_1}{1+c_1^2} = 0.34482$ and $\rho_2 = 0$.
So, PACF coefficients will be $\phi_{11} = 0.34482$ and $\phi_{22} = -0.1349$
PACF coefficients calculated using ARMAacf in R are given as $\phi_{11} = 0.34482$ and $\phi_{22} = -0.1349$. Although there is no direct relation between $x[k]$ and $x[k-1]$,$x[k-2]$, we have got some values because of noise.

4.2

4.2.1 EuStockMarkets data

The data is shown in the following Figure 4.

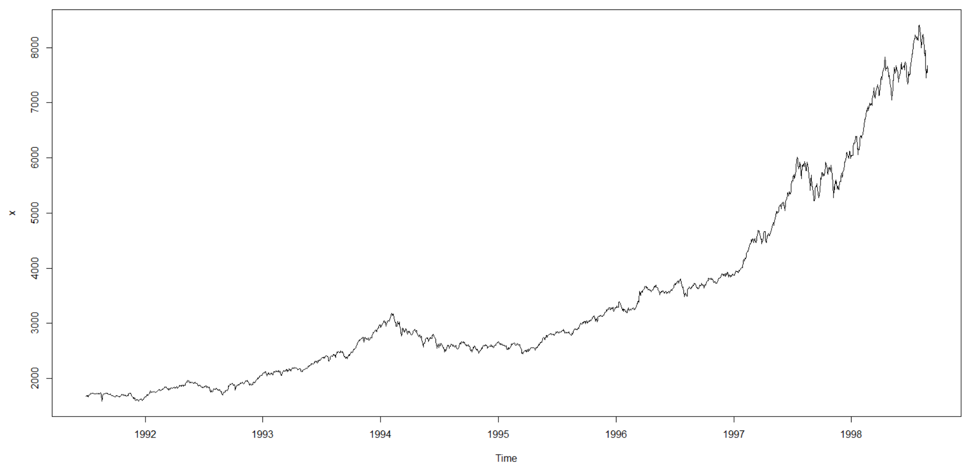

Figure 4: SMI data

In Figure 4, the data is **not stationary** (as mean is varying with time) and **not periodic** . The estimated ACF of the data is shown in Figure 5.

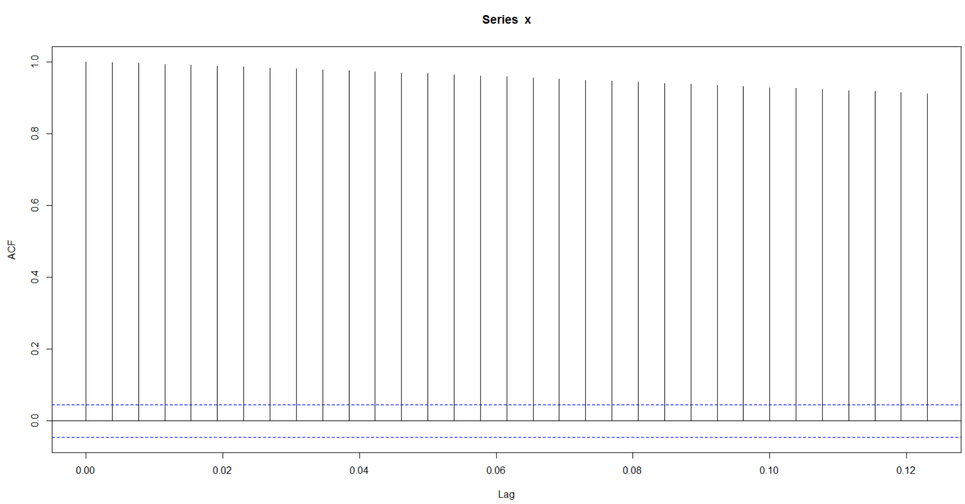

Figure 5: ACF of SMI data

ACF is not decaying with time, so the process is not stationary which can be decided by visual inspection. ACF plot suggest there is no periodicity in the series. The estimated PACF of the data is shown in the following Figure 6.

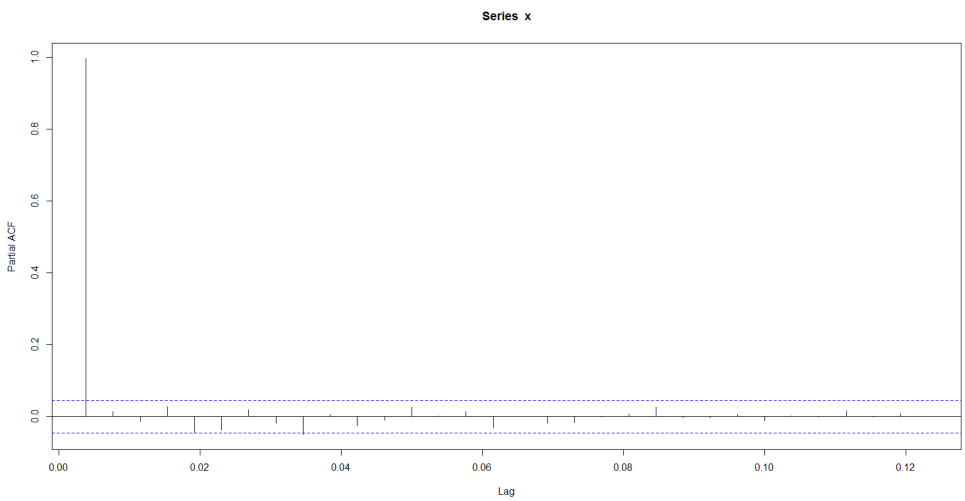

Figure 6: PACF of SMI data

PACF plot suggests that the there is only direct relation between $y[k]$ and $y[k-1]$ sample means we can model it as AR(1) process.

4.2.2 quakes data

The data is shown in Figure 7

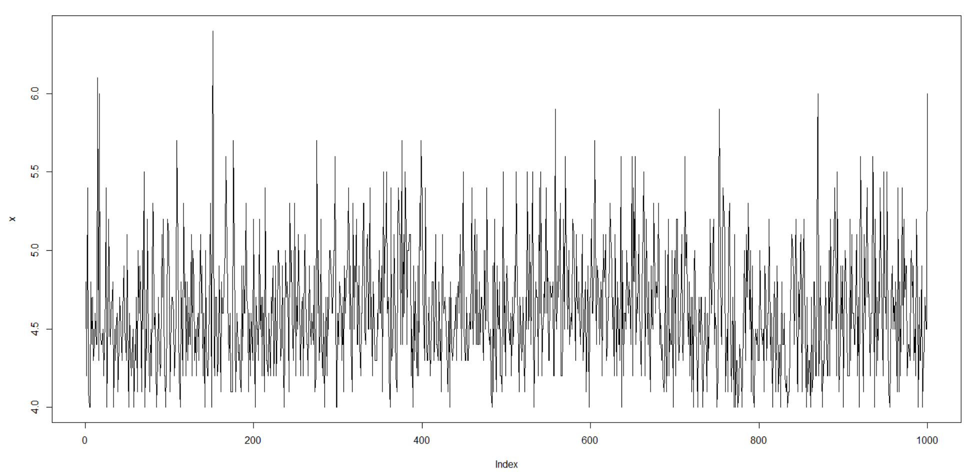

Figure 7: magnitude data

In Figure 7, the data is **stationary**. Although the data looks like periodic, we cannot decide it truly by visual inspection.
The estimated ACF of the data is shown in Figure 8

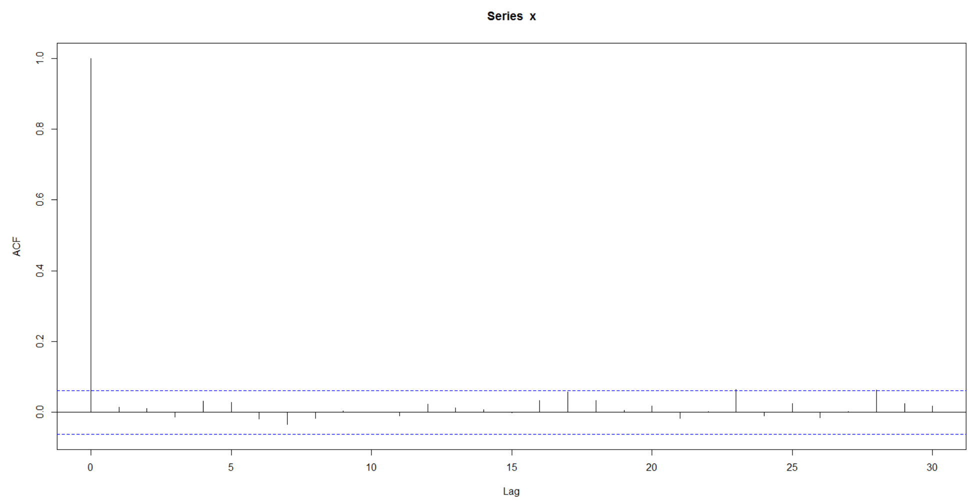

Figure 8: ACF of magnitude data

Estimated ACF plot resembles like the ACF plot of white noise process. This suggests the process cannot be modelled as a linear process. The estimated PACF of the data is shown in Figure 9

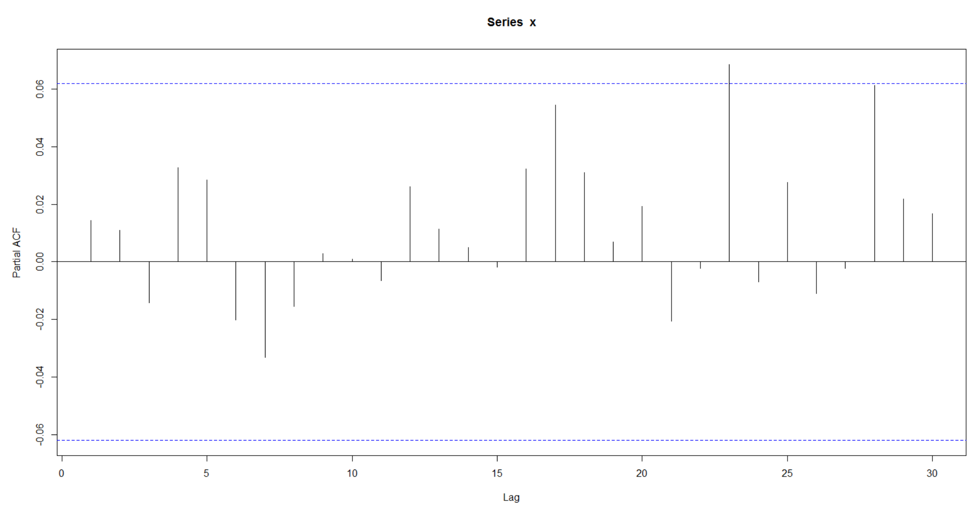

Figure 9: PACF of magnitude data

Estimated PACF plot confirms same thing as that of ACF plot.

4.2.3 nottem data

The data is shown in Figure 10

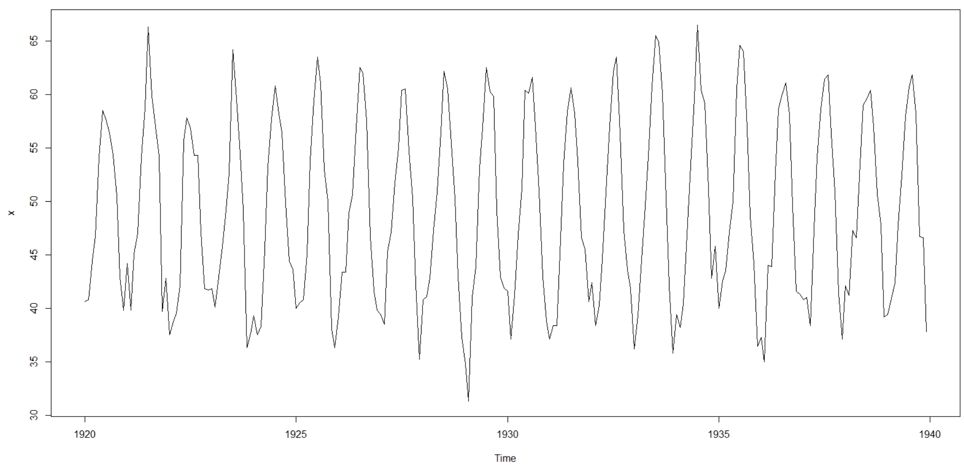

Figure 10: nottem data

In Figure 10, the data is **stationary** and is **periodic** by visual inspection.
The estimated ACF of the data is shown in Figure 11

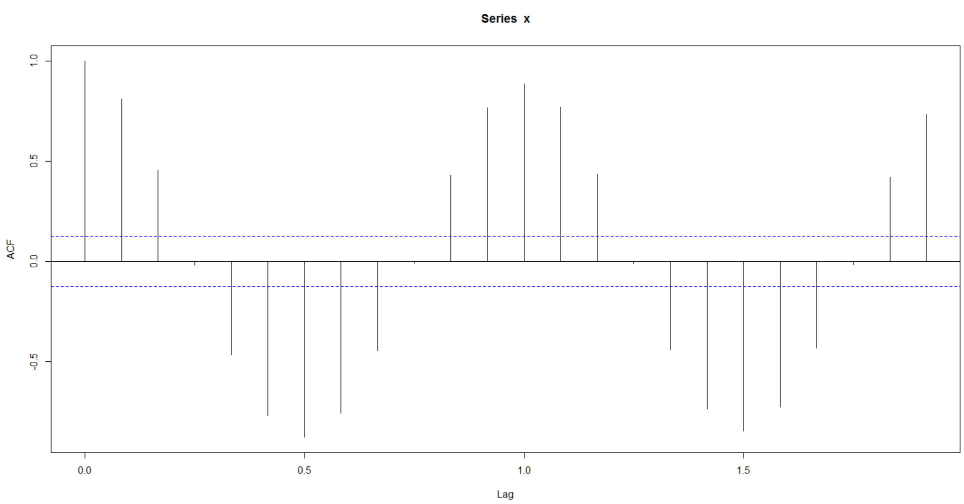

Figure 11: ACF of nottem data

The estimated ACF plot in Figure 11 confirms that the series is periodic. The estimated PACF of the data is given in Figure 12

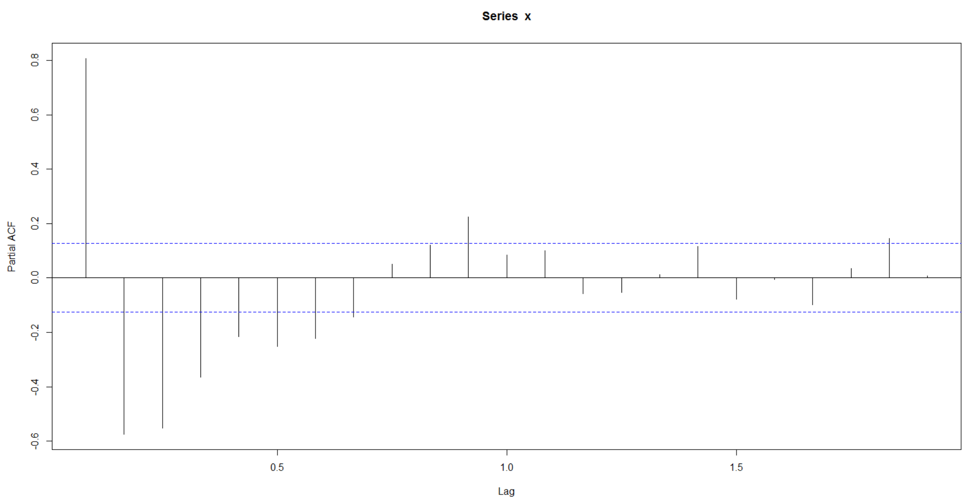

Figure 12: PACF of nottem data

The estimated PACF plot in Figure 12 shows that the series can be sampled as AR(4) process.

4.2.4 generated data

The data is shown in Figure 13

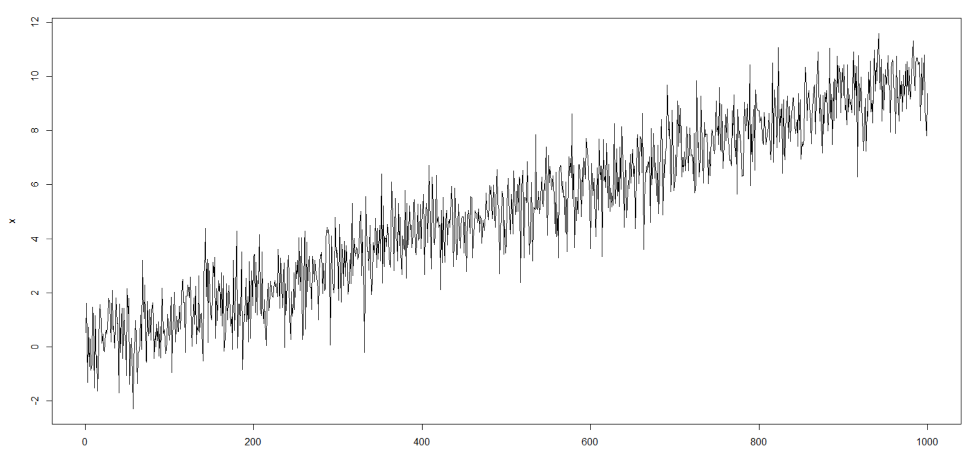

Figure 13: $x[k] = 0.01k + e[k]$ data

In Figure 13, the data is **not stationary** (as mean is varying with time) and is **not periodic**. The estimated ACF of the generated data is shown in Figure 14

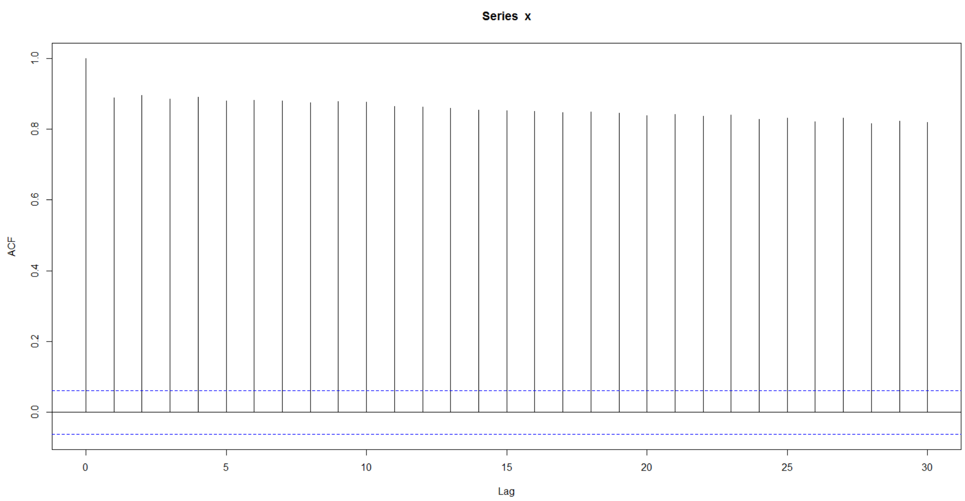

Figure 14: ACF of generated data

Figure 14 confirms the data is non-stationary. The estimated PACF of the data is given in Figure 15

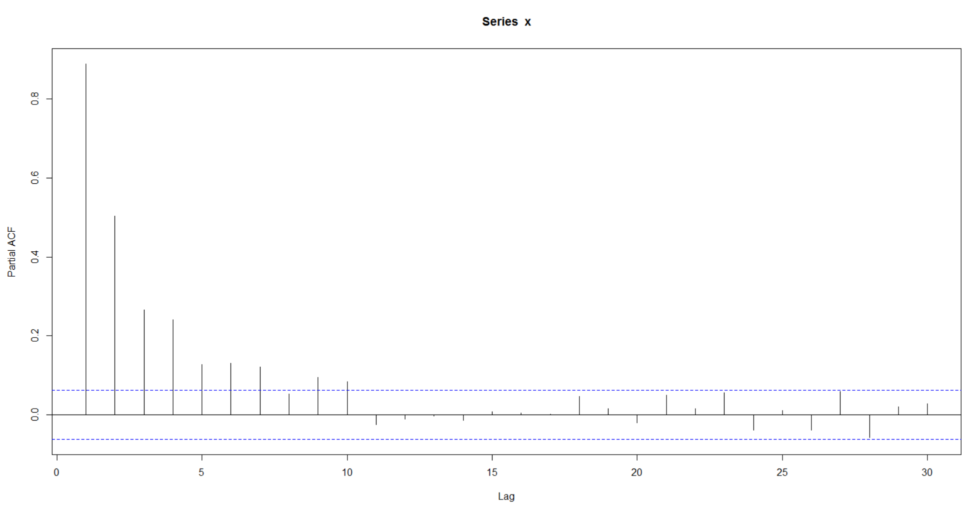

Figure 15: PACF of generated data

Estimated PACF plot shows that the series can be modelled as AR(6) process.

5

The code to generate distribution of ACF values at three different lags for white noise process is shown below

```
dev.new()
par(font.main=2,font.axis=2,font.lab=2,cex.axis=1.1,cex.lab=1.1) layout(matrix
    (1:3,3,1))
n=200
for (l in c(2,8,18))
{
acfests<-replicate(n,acf(rnorm(1000),plot=F)$acf[l]) hcolor<-floor(runif(1,min
    =1,max=657))
hist(acfests, main=paste(?Distribution of ACF estimate?), ylab=?Density?, prob
    =T, col=colors()[hcolor])
}
```

Three distributions shown in Figures 16,17 and 18 are of estimated ACF values of white noise process at three different lags ($l = 1, 7, 17$).

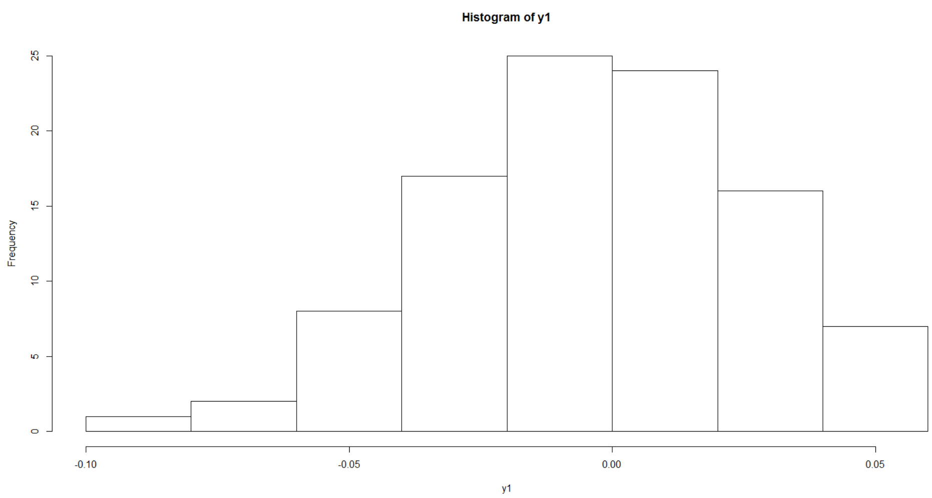

Figure 16: Distribution of ACF at lag 1

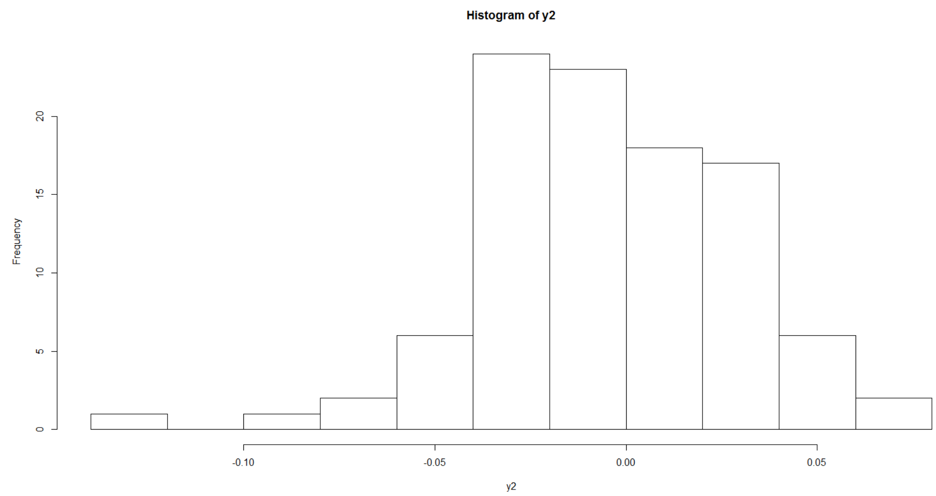

Figure 17: Distribution of ACF at lag 7

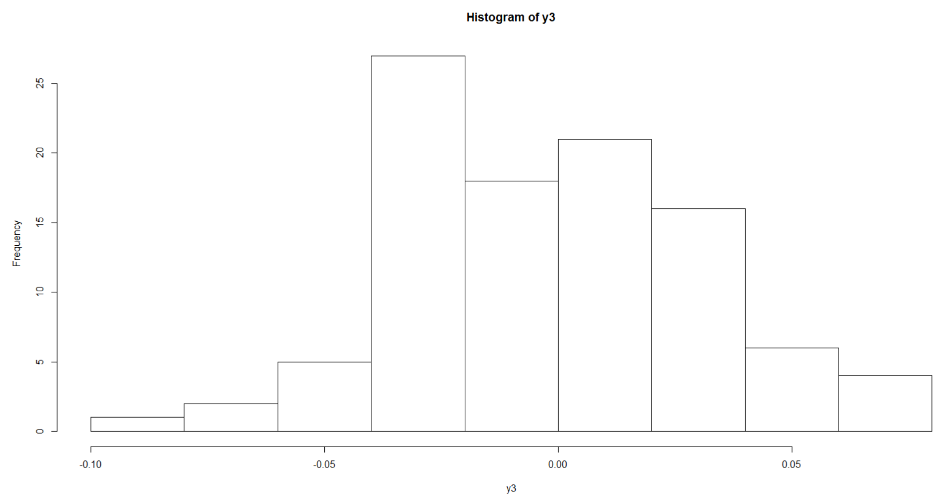

Figure 18: Distribution of ACF at lag 17

The true value of ACF of white noise process other than $(l \neq 0)$ should be zero. But the distribution shows the there are many non-zero values although they are negligible. This is because the distribution we have plotted is for only estimated values.