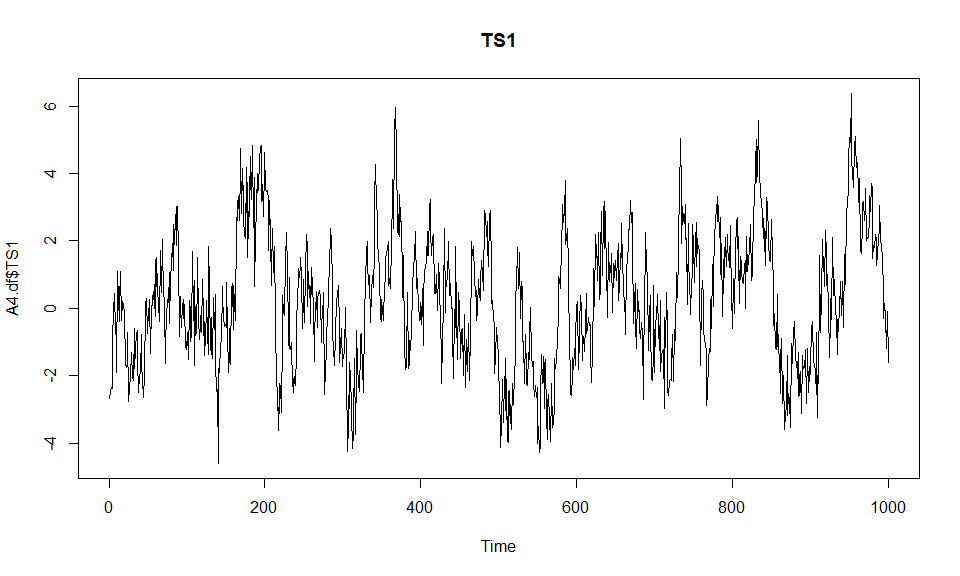
Stats 326: Assignment 4

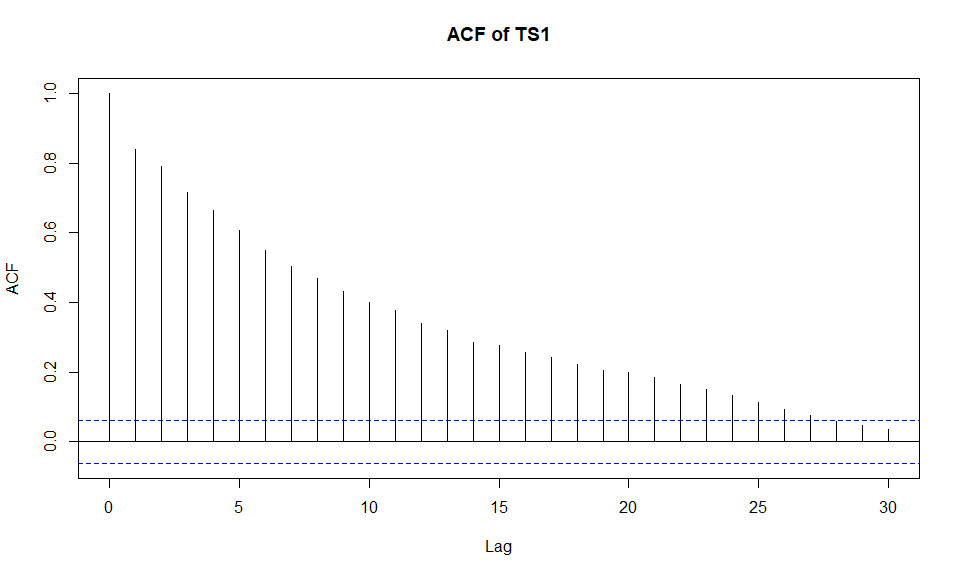
Hasnain Cheena

16/04/2020

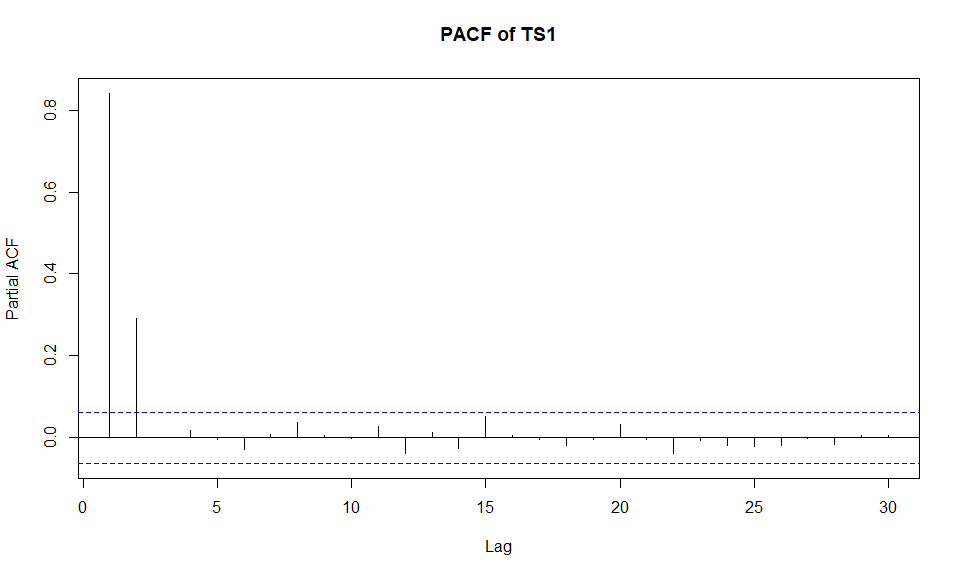
## Question 1

plot.ts(A4.df$TS1, main="TS1")

acf(A4.df$TS1, main="ACF of TS1")



pacf(A4.df$TS1, main="PACF of TS1")



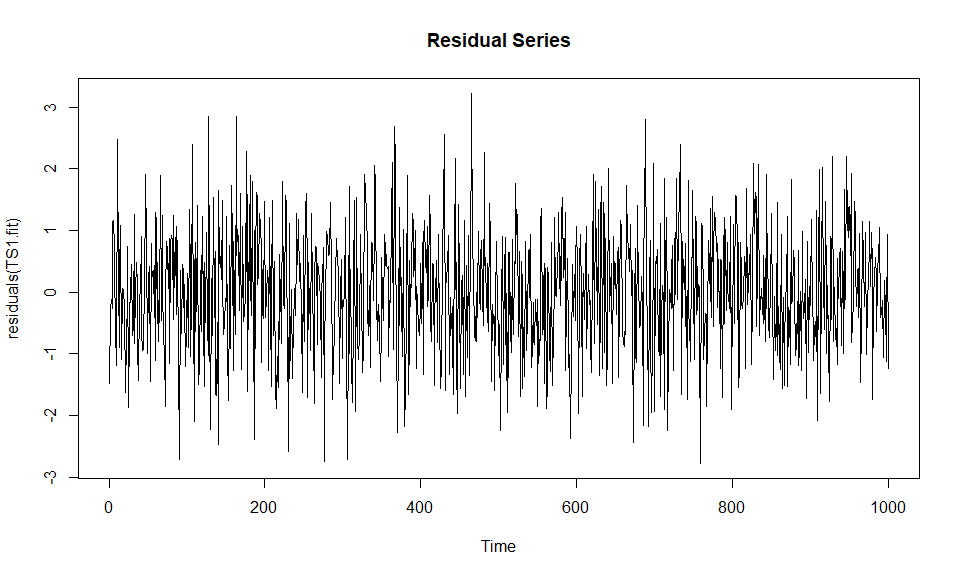
The plot of the series shows clustering indicating positive autocorrelation. The acf shows decay while the pacf shows cut-off at lag 2. This suggests AR(2) is the most suitable model. A general equation of the model is shown below:

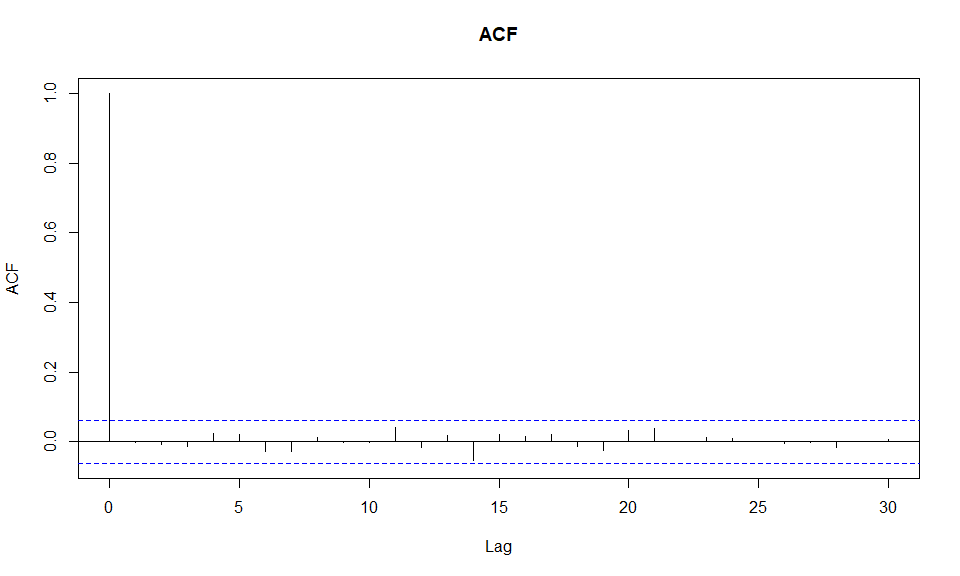
TS1.fit = arima(A4.df$TS1, order=c(2,0,0))  
TS1.fit

##   
## Call:  
## arima(x = A4.df$TS1, order = c(2, 0, 0))  
##   
## Coefficients:  
## ar1 ar2 intercept  
## 0.5958 0.2928 0.2106  
## s.e. 0.0302 0.0303 0.2821  
##   
## sigma^2 estimated as 1.008: log likelihood = -1423.72, aic = 2855.44

Estimated equation of model:

plot(residuals(TS1.fit), main="Residual Series")



acf(residuals(TS1.fit), main="ACF")

The Residual Series appear to be random scatter about 0. The plot of the autocorrelation function of the Residual Series shows no significant lags. Therefore, AR(2) is appropriate.

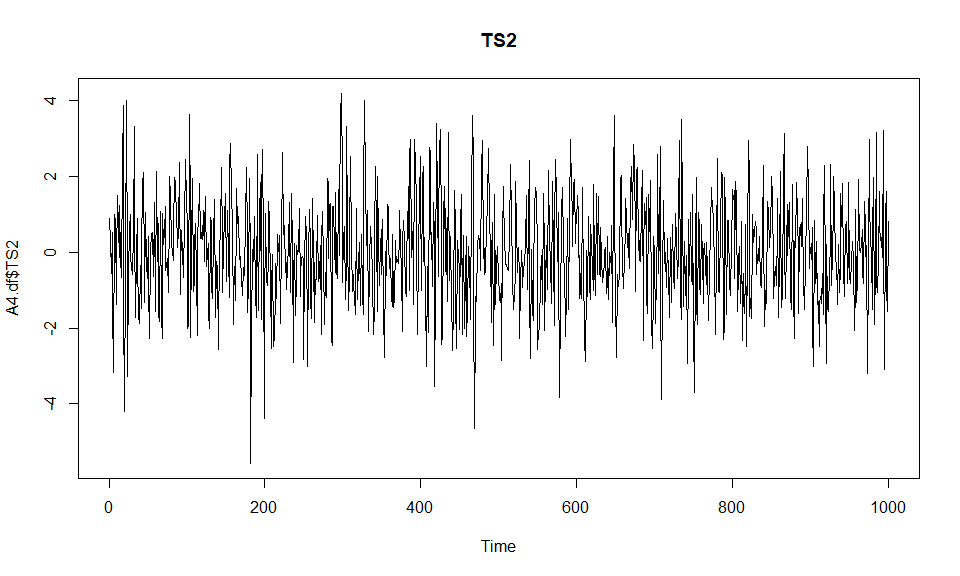
Other models tried:

AR(3) AIC = 2857.86  
ARMA(2,1) AIC = 2889.29

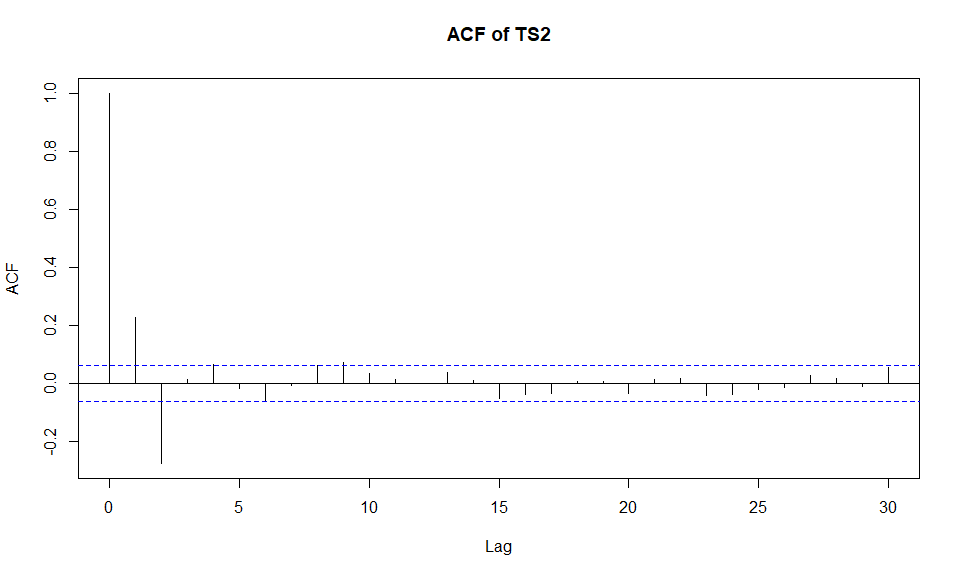
AR(2) is the best model. This is because AR(2) had the smallest AIC score and all terms were significant.

## Question 2

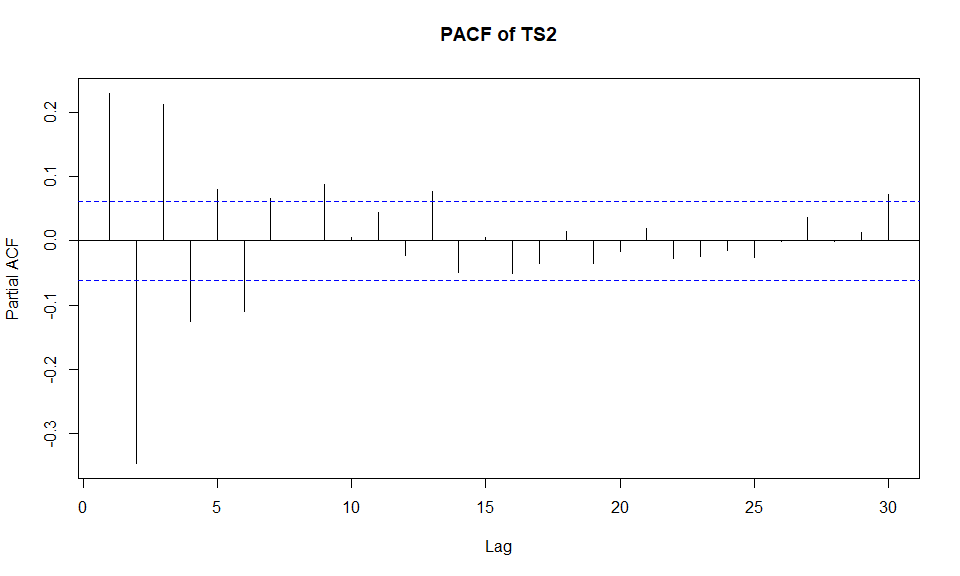
plot.ts(A4.df$TS2, main="TS2")



acf(A4.df$TS2, main="ACF of TS2")



pacf(A4.df$TS2, main="PACF of TS2")

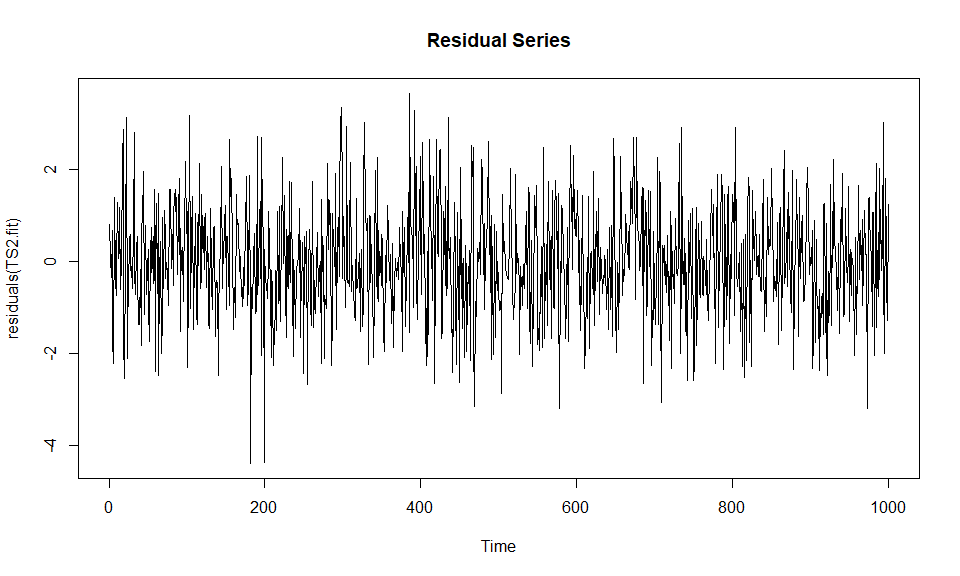


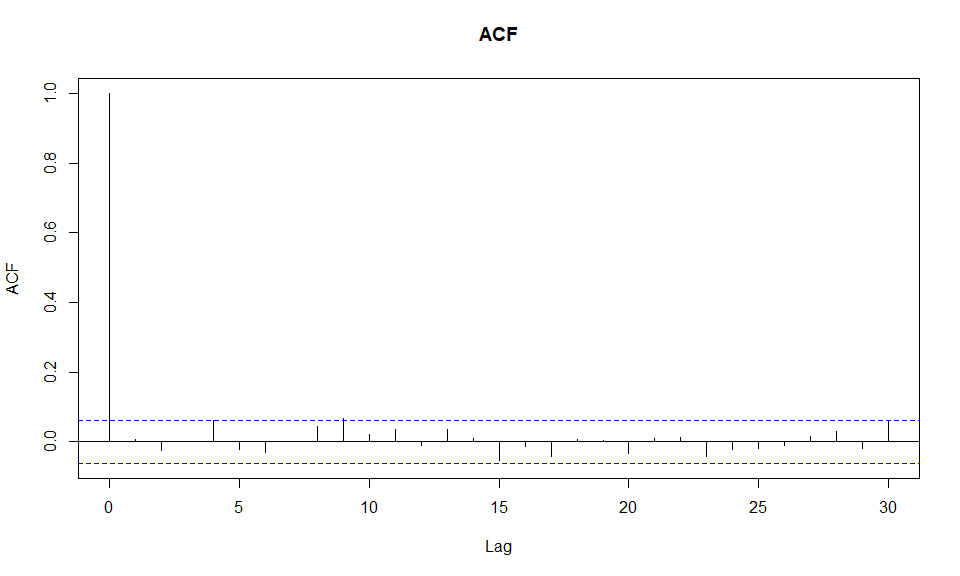
The plot of the series shows no discrenable pattern. The acf shows cut-off at lag 2 and the pacf shows decay (or persistence). This suggests MA(2) is the most suitable model.   
A general equation of the model is shown below:

TS2.fit = arima(A4.df$TS2, order=c(0,0,2))  
TS2.fit

##   
## Call:  
## arima(x = A4.df$TS2, order = c(0, 0, 2))  
##   
## Coefficients:  
## ma1 ma2 intercept  
## 0.4377 -0.311 -0.0086  
## s.e. 0.0302 0.030 0.0433  
##   
## sigma^2 estimated as 1.475: log likelihood = -1613.7, aic = 3235.4

Estimated equation of model:

plot(residuals(TS2.fit), main="Residual Series")

acf(residuals(TS2.fit), main="ACF")

The Residual Series appears to be random scatter about 0. The plot of the autocorrelation function of the Residual Series shows no significant lags.

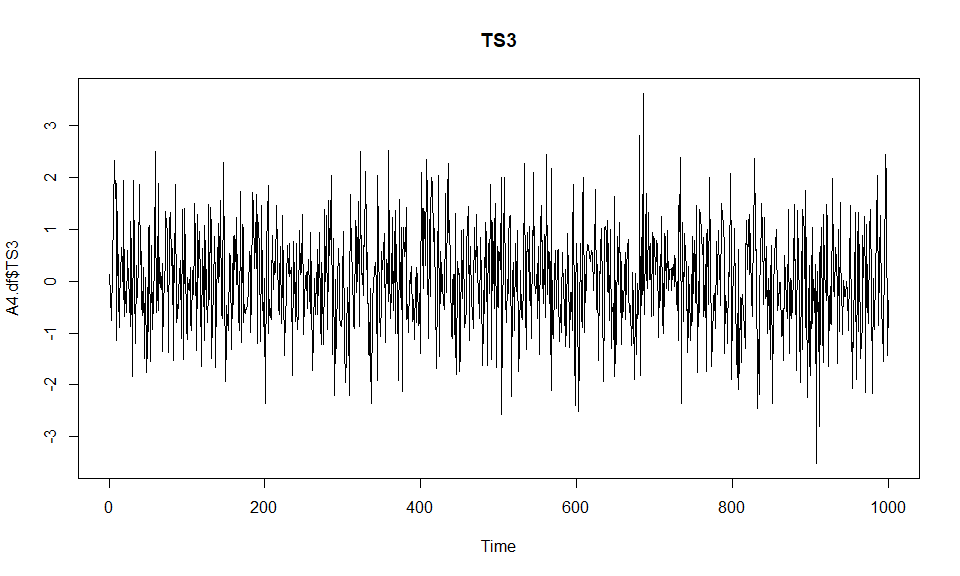
Other models tried:

MA(3) AIC = 3236.86  
ARMA(1,2) AIC = 3236.98

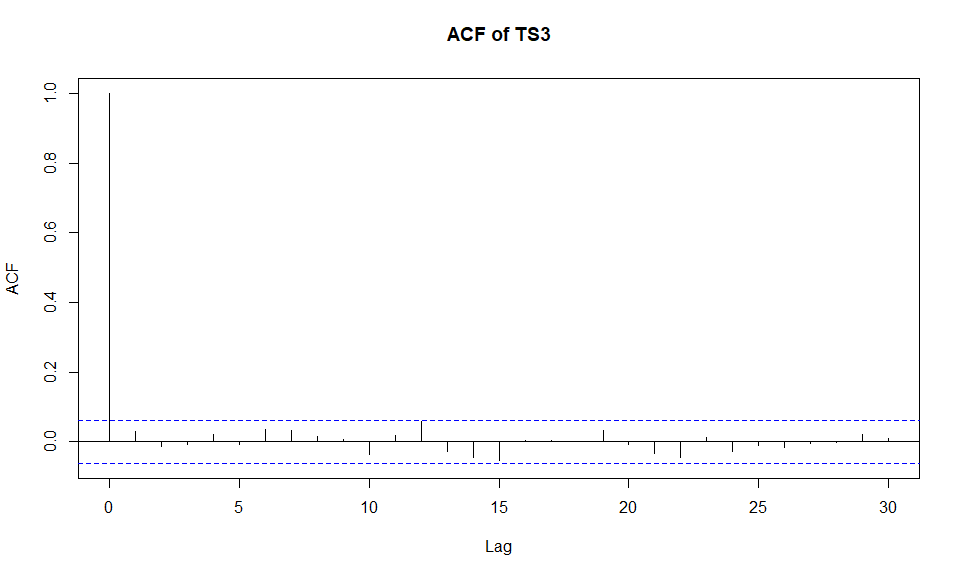
MA(2) is the best model. This is because MA(2) had the smallest AIC score and all terms were significant.

## Question 3

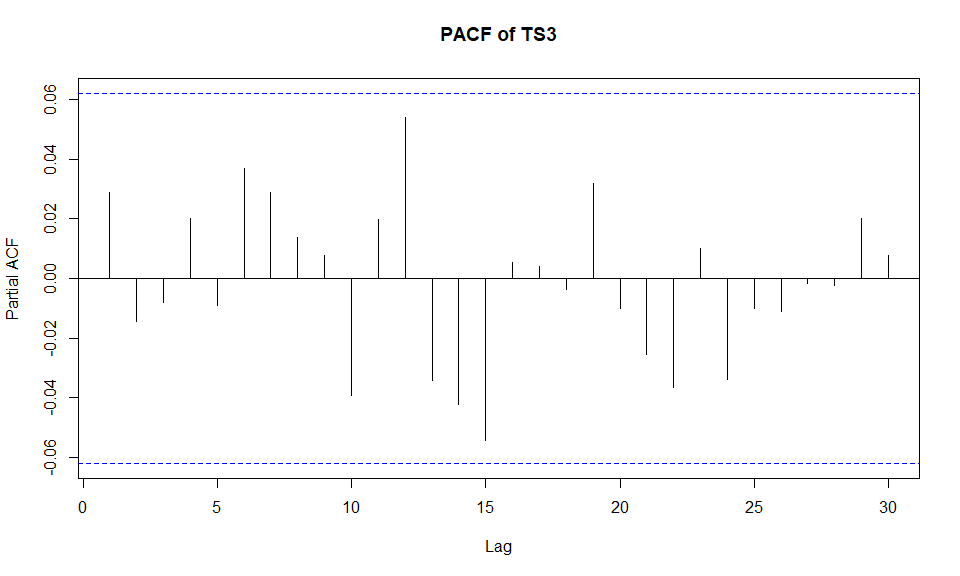
plot.ts(A4.df$TS3, main="TS3")



acf(A4.df$TS3, main="ACF of TS3")



pacf(A4.df$TS3, main="PACF of TS3")

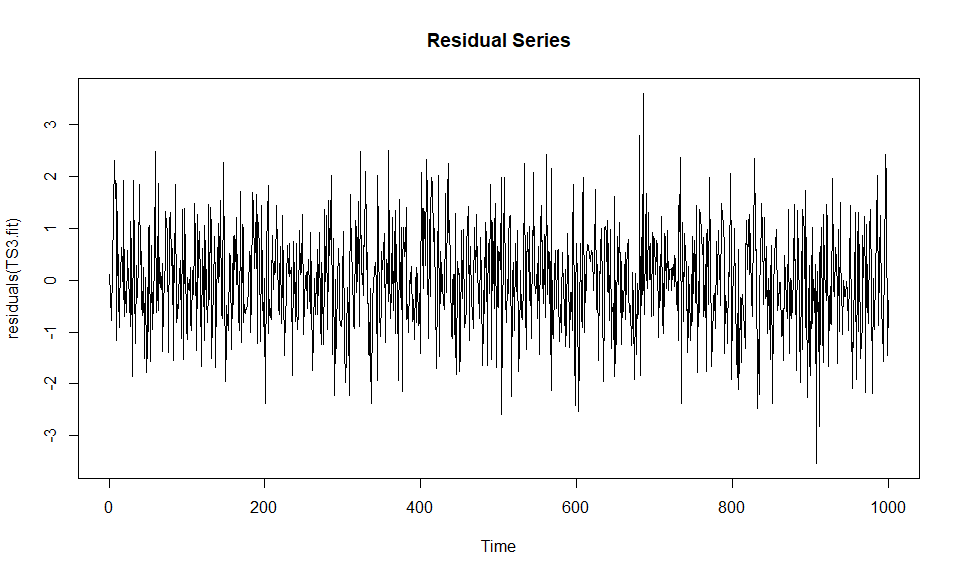


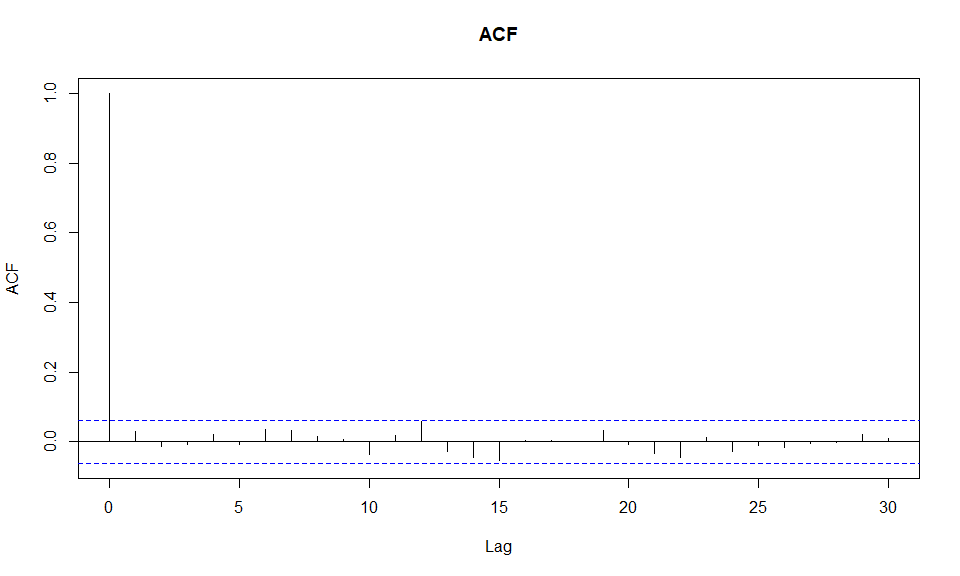
The plot of the series shows no discernible pattern. The acf and pacf show no significant lags. This suggests the series is White Noise. A general form of the model is shown below:

TS3.fit = arima(A4.df$TS3, order=c(0,0,0))  
TS3.fit

##   
## Call:  
## arima(x = A4.df$TS3, order = c(0, 0, 0))  
##   
## Coefficients:  
## intercept  
## 0.0211  
## s.e. 0.0316  
##   
## sigma^2 estimated as 0.9961: log likelihood = -1417, aic = 2838

Estimated model:

plot(residuals(TS3.fit), main="Residual Series")

acf(residuals(TS3.fit), main="ACF")

The Residual Series appears to be random scatter about 0. The plot of the autocorrelation function of the Residual Series shows no significant lags.

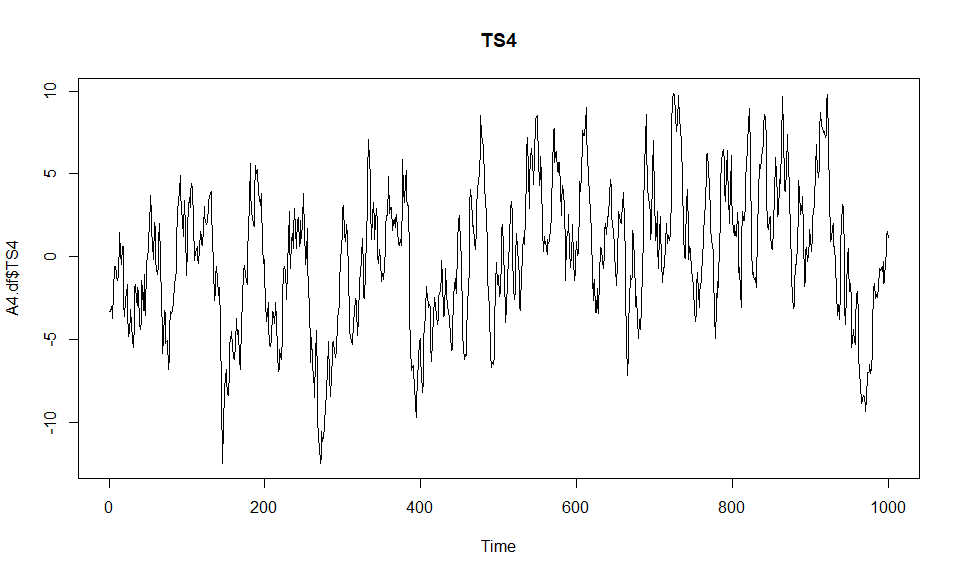
Other models tried:

AR(1) AIC: 2839.17   
MA(1) AIC: 2839.14

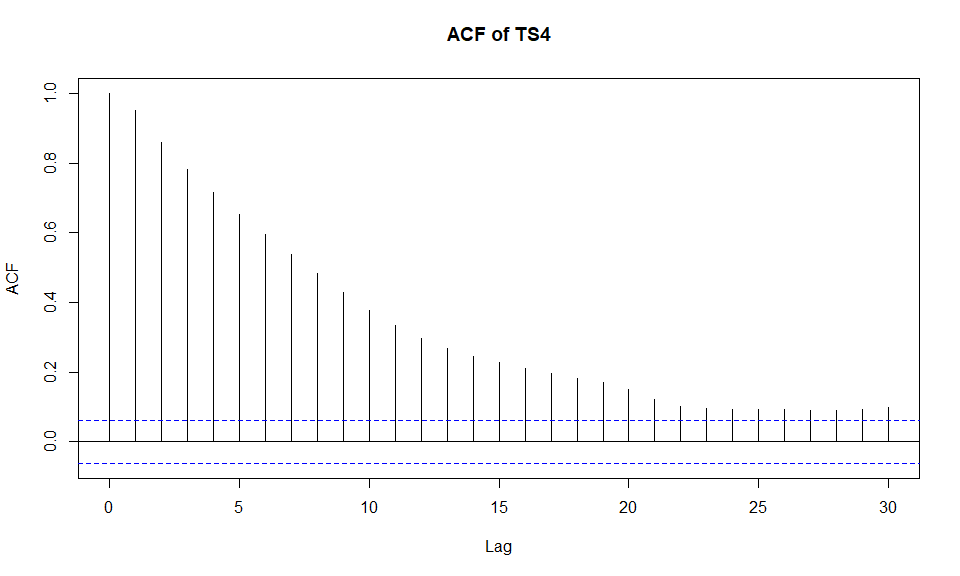
The white noise model is the best model as it met all assumptions and has the lowest AIC score relative to the other models tried.

## Question 4

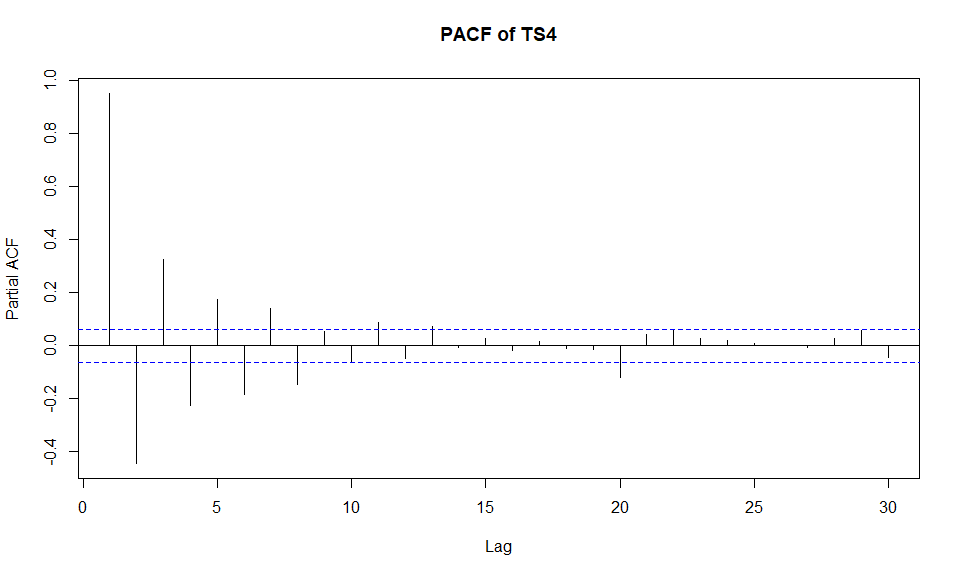
plot.ts(A4.df$TS4, main="TS4")



acf(A4.df$TS4, main="ACF of TS4")



pacf(A4.df$TS4, main="PACF of TS4")

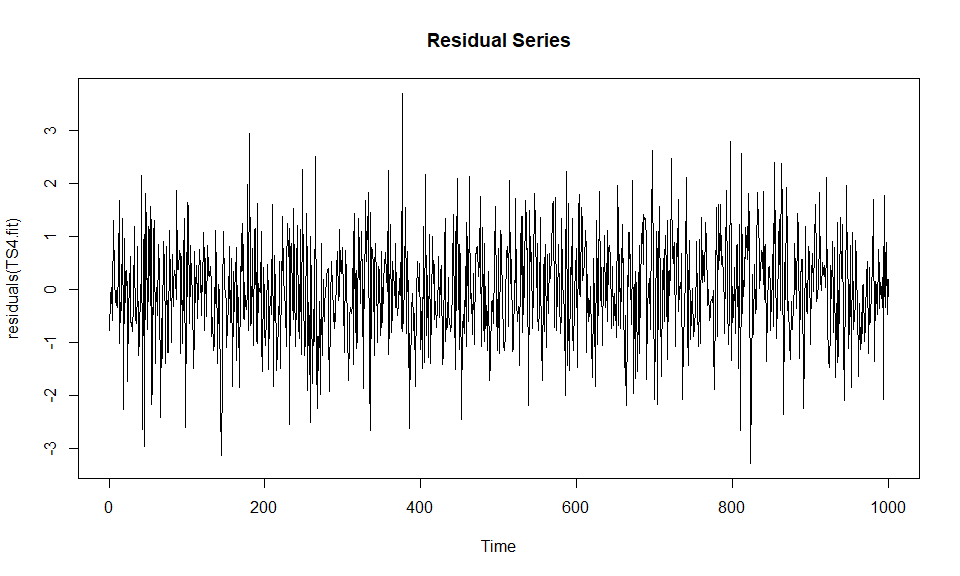


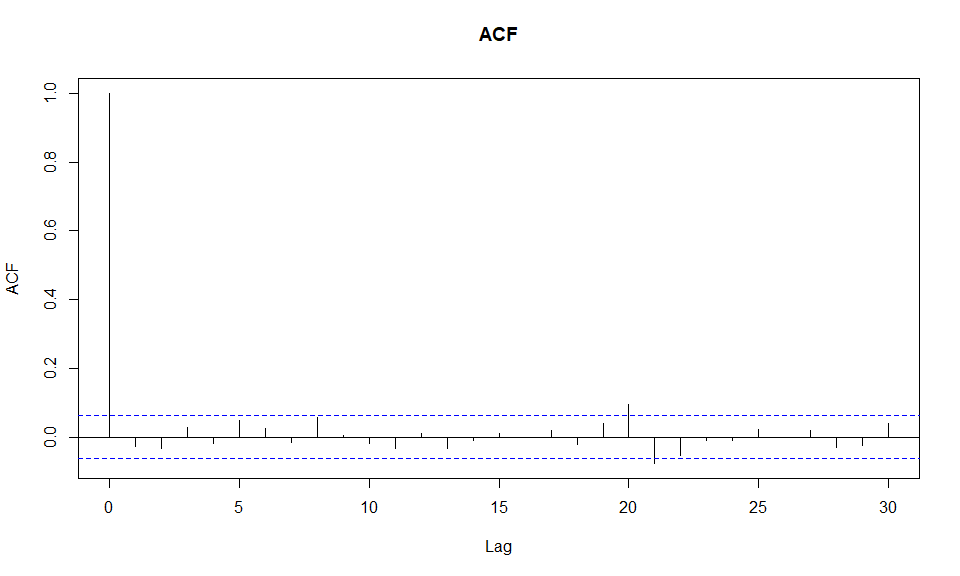
The plot of the series shows clustering indicating positive autocorrelation. Both the acf and pacf show decay. This suggests an ARMA(p,q) is an appropriate model. However, from the plots we have no indication of what order of the ARMA model must be used. Therefore, I began with ARMA(1,1). General form of ARMA(1,1) shown below:

TS4.fit = arima(A4.df$TS4, order=c(1,0,1))  
TS4.fit

##   
## Call:  
## arima(x = A4.df$TS4, order = c(1, 0, 1))  
##   
## Coefficients:  
## ar1 ma1 intercept  
## 0.8974 0.9121 -0.0147  
## s.e. 0.0139 0.0128 0.5786  
##   
## sigma^2 estimated as 0.9828: log likelihood = -1412.55, aic = 2833.11

Estimate equation of the model:

plot(residuals(TS4.fit), main="Residual Series")

acf(residuals(TS4.fit), main="ACF")

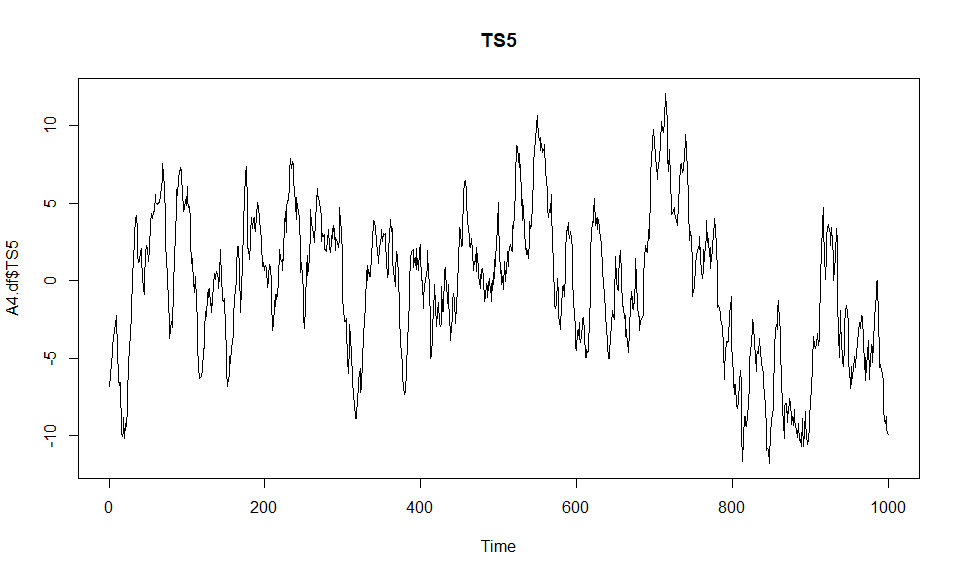
The Residual Series appears to be random scatter about 0. The plot of the autocorrelation function of the Residual Series shows 2 weakly signficant lags at lags 20 and 21. As they are weakly significant, they are not a concern.

Other models tried:   
ARMA(2,1) AIC: 2833.86   
ARMA(1,2) AIC: 2833.73

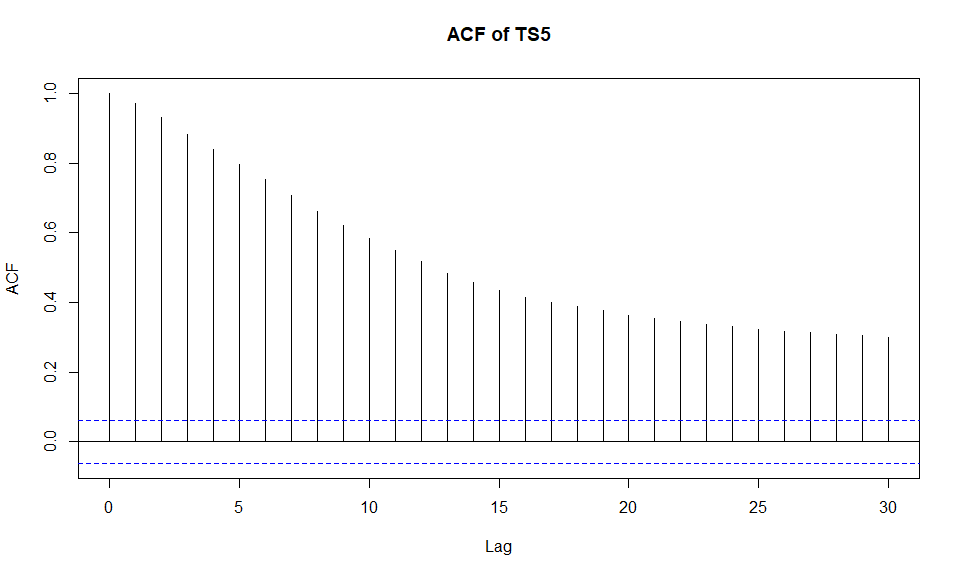
ARMA(1,1) is the best model because all terms are significant and it has the lowest AIC score relative to the other models tried.

## Question 5

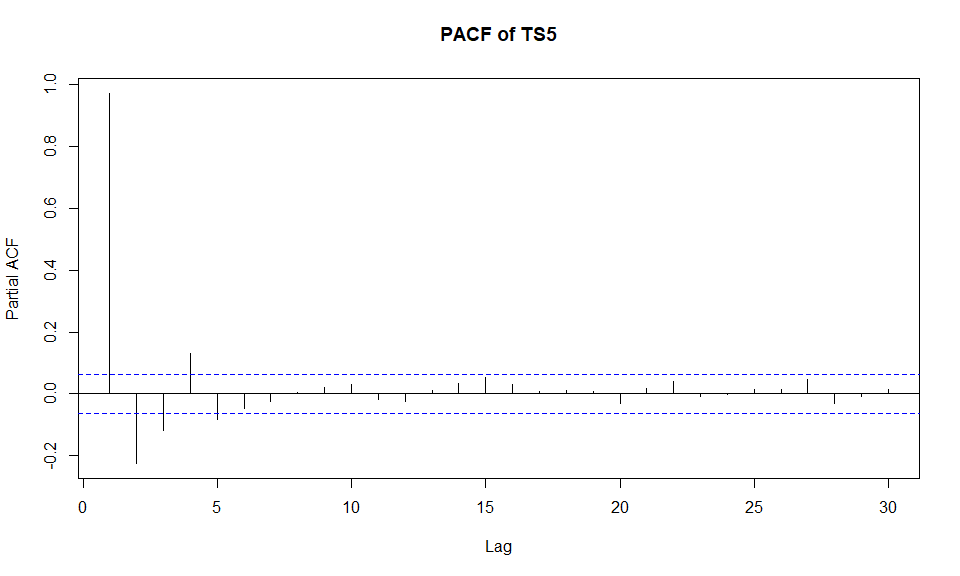
plot.ts(A4.df$TS5, main="TS5")



acf(A4.df$TS5, main="ACF of TS5")



pacf(A4.df$TS5, main="PACF of TS5")



The plot of the series shows clustering indicating positive autocorrelation. Both the acf and pacf show decay/persistence. This suggests an ARMA(p,q) is an appropriate model. However, from the plots we have no indication of what order of the ARMA model must be used. Therefore, I began with ARMA(1,1). General form of ARMA(1,1) shown below:

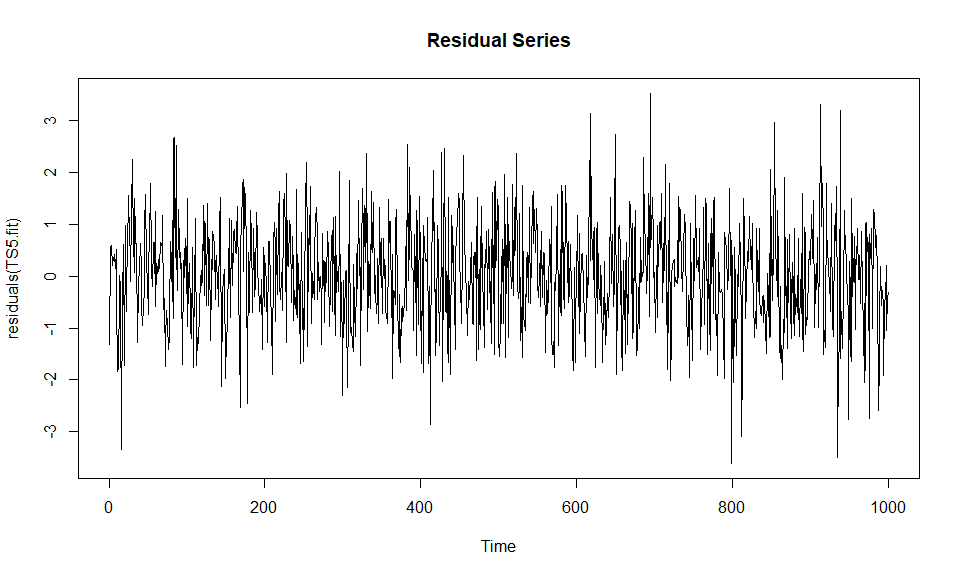
TS5.fit = arima(A4.df$TS5, order=c(1,0,1))

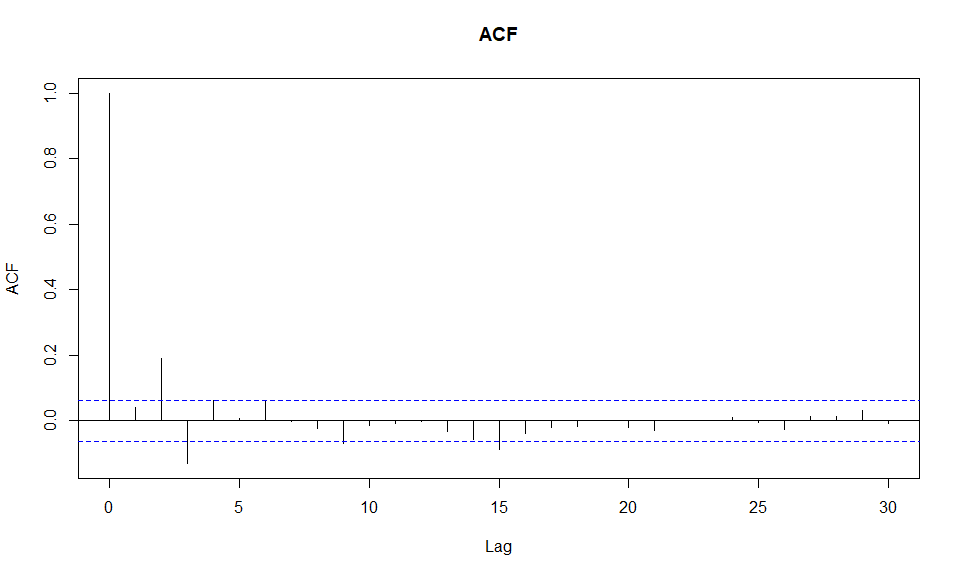
## Warning in arima(A4.df$TS5, order = c(1, 0, 1)): possible convergence problem:  
## optim gave code = 1

TS5.fit

##   
## Call:  
## arima(x = A4.df$TS5, order = c(1, 0, 1))  
##   
## Coefficients:  
## ar1 ma1 intercept  
## 0.9674 0.1876 -0.6895  
## s.e. 0.0082 0.0260 1.1571  
##   
## sigma^2 estimated as 1.063: log likelihood = -1450.83, aic = 2909.67

Estimated equation:

plot(residuals(TS5.fit), main="Residual Series")

acf(residuals(TS5.fit), main="ACF")

The Residual Series appears to be random scatter about 0. The plot of the autocorrelation function of the Residual Series shows 4 significant lags at lags 2, 3, 9 and 15. A better model is outlined on the next page.

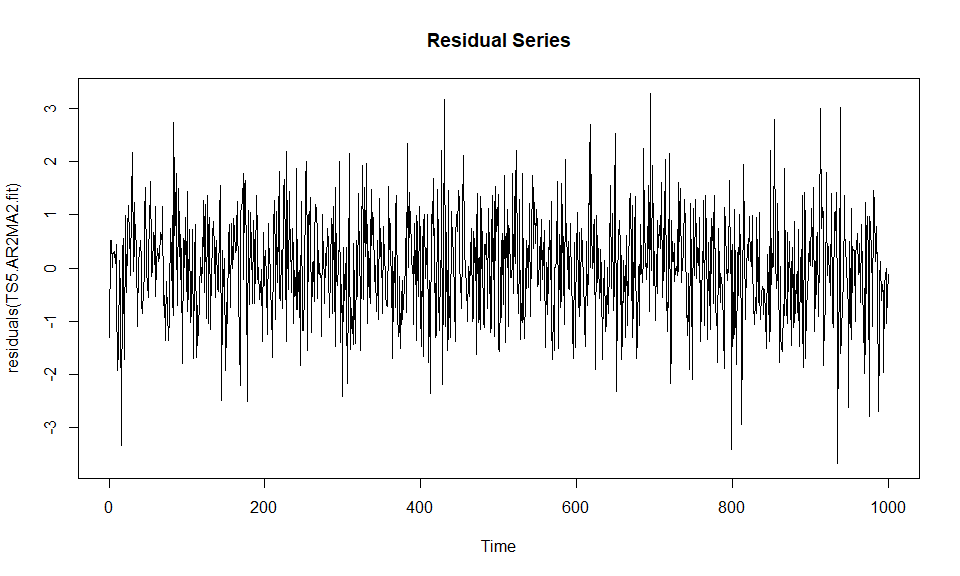
**Better Model:**

TS5.AR2MA2.fit = arima(A4.df$TS5, order=c(2,0,2))  
TS5.AR2MA2.fit

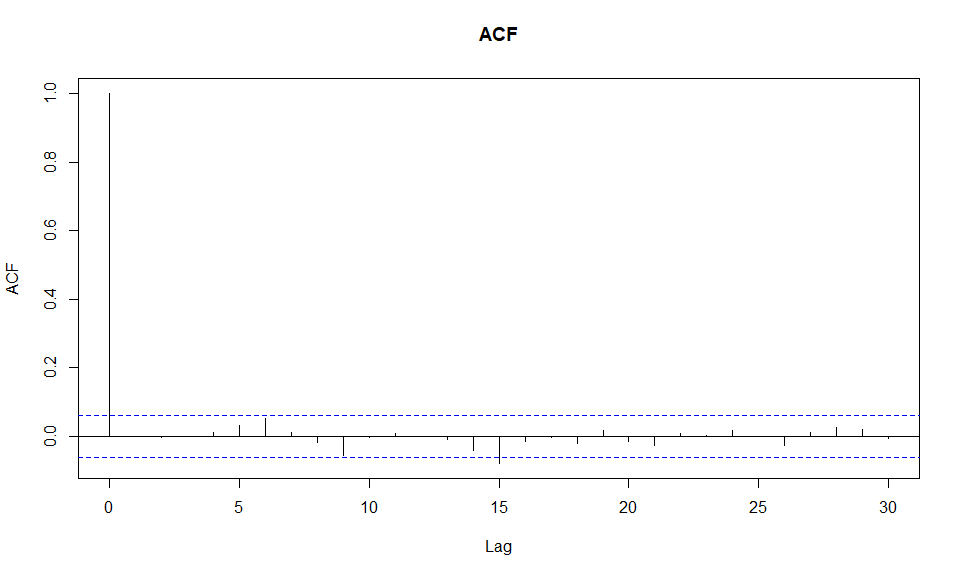
##   
## Call:  
## arima(x = A4.df$TS5, order = c(2, 0, 2))  
##   
## Coefficients:  
## ar1 ar2 ma1 ma2 intercept  
## 0.5840 0.3552 0.6361 0.3235 -0.5989  
## s.e. 0.1063 0.1040 0.1014 0.0325 0.9989  
##   
## sigma^2 estimated as 0.9954: log likelihood = -1418.31, aic = 2848.63

Estimated equation:

plot(residuals(TS5.AR2MA2.fit), main="Residual Series")



acf(residuals(TS5.AR2MA2.fit), main="ACF")



The Residual Series appears to be random scatter about 0. The plot of the autocorrelation function of the residual series shows a significant lag at lag 15. However, as this lag is weakly significant is not of concern.

Therefore, ARMA (2,2) was decided as the best model for this series as all estimates are significant and it has the lowest AIC score relative to the other models tried.