PSTAT 174 HW #4

Chris Meade

2/14/2017

Problem 2

Let $\{X_t\}$ be a $MA(1)$ process such that $X_t = Z_t + \theta Z_{t-2}$. Assume that $\{X_t\}$ has mean μ, $\theta = -0.6$, $\sigma^2 = 1$, and $\bar{x}_{100} = 0.157$. Since $\{X_t\}$ is a $MA(1)$ process, we have

$$\begin{gathered}\gamma_X(h) = \sigma^2(1+\theta^2) \text{ if } h = 0 \\ \gamma_X(h) = \sigma^2\theta \text{ if } h = 1, -1 \\ \gamma_X(h) = 0 \text{ if } |h| > 1\end{gathered}$$

Since $n = 100$ is sufficiently large, we can estimate the variance of $\{X_t\}$ with $n^{-1}v$ where

$$\begin{aligned}v &= \gamma_X(0) + 2\sum_{1\le h\le n}((1-\frac{h}{n})\gamma_X(h)) \\ &= \sigma^2(1+\theta^2) + 2((1-\frac{1}{100})\theta) \\ &= \sigma^2 + \sigma^2\theta^2 + 2\theta - 2\theta\frac{1}{100}\end{aligned}$$

Then

$$n^{-1}v = 100^{-1}(\sigma^2 + \sigma^2\theta^2 + 2\theta - 2\theta\frac{1}{100})$$

Plugging in estimate for σ^2 and θ yields

$$n^{-1}v = \frac{0.172}{100} = 0.00172$$

Since X_t is approximately normal distributed, we may we the Z test statistic in the construction of our confidence interval.

We now construct the confidence interval

$$\begin{aligned}I &= (\bar{x}_{100} - Z_{.025} * 0.00172, \bar{x}_{100} + Z_{.025} * 0.00172) \\ &= (0.157 - 1.96 * 0.00172, 0.157 + 1.90 * 0.00172) \\ &= (0.153629, 0.160268)\end{aligned}$$

Since $0 \notin I$, we would reject the hypothesis that $\mu = 0$.

Problem 3

Using Bartlet's formula, we cacluate the diagonal entries of the matrix W, $w_{1,1}$ and $w_{2,2}$.

$$w_{1,1} = 1 - 3\rho(1)^2 + 4\rho(1)^4$$

and

$$w_{2,2} = 1 + 2\rho(1)^2$$

To contruct the 95% confidence interval for $\rho(1)$, we plug our estimates of $\rho(1)$ and $\rho(2)$ into the equation for $w_{1,1}$.

$$\begin{aligned}\hat{w}_{1,1} &= 1 - 3\rho(1)^2 + 4\rho(1)^4 \\ &= 1 - 3(0.438)^2 + 4(0.145)^4 \\ &= 0.426\end{aligned}$$

The confidence interval of $\rho(1)$ is of the form $(\hat{\rho}(1) - Z_{0.025}\sqrt{w_{1,1}/n}, \hat{\rho}(1) + Z_{0.025}\sqrt{w_{1,1}/n})$. Substitution our estimates yields $(0.310, 0.566)$.

We now repeat the process to construct a 95% confidence interval for $\rho(2)$.

$$\begin{aligned}\hat{w}_{2,2} &= 1 + 2\rho(1)^2 \\ &= 1 + 2(0.438)^2 \\ &= 1.387\end{aligned}$$

Then the confidence interval for

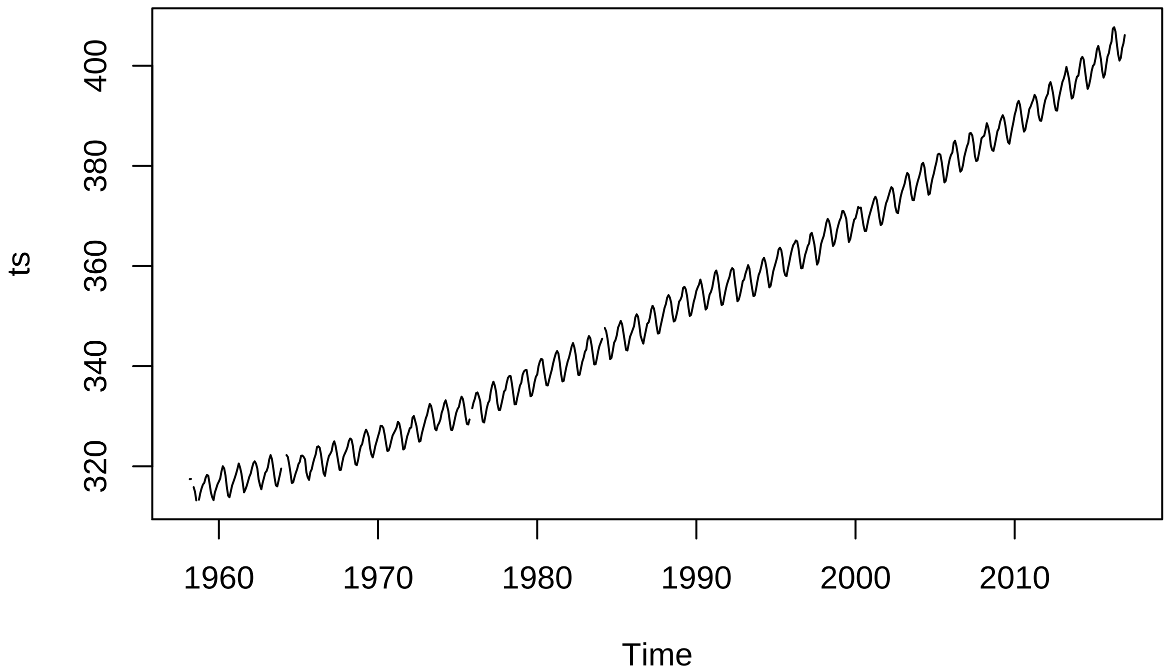
$$\rho(2) = (\hat{\rho}(2) - Z_{0.025}\sqrt{(}w_{2,2}/100), \hat{\rho}(2) + Z_{0.025}\sqrt{(}w_{2,2}/100)) = (-0.085, 0.376)$$

If $\theta = 0.6$, then $\rho(1) = \dfrac{0.6}{1+0.6^2} = 0.441$. Since this value in an element of the confidence interval for $\rho(1)$, and because the confidence interval for $\rho(2)$ contains 0, we fail to reject the hypothesis that $\theta = 0.6$.

Problem 4

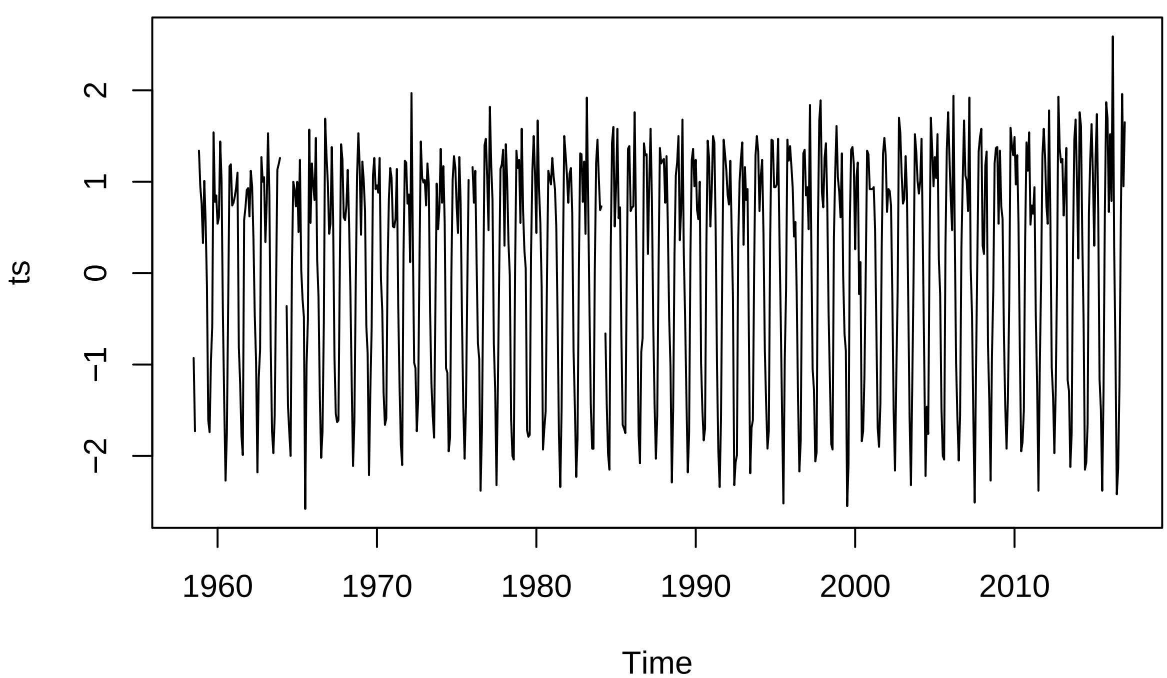
This dataset contains the average monthly atomospheric $C0_2$ concentration. The data is available from https://www.esrl.noaa.gov/gmd/ccgg/trends/data.html. This data is important due to the relationship between atmospheric CO_2 concentration and global warming. In forecasting this time series, we may be able to predict what effects will we face from global warming.

```
# Import Data
ts <- read.table("c02.txt", skip = 3, header = F, sep = "")

# Remove exraneous featurs
ts <- ts$V4

# Encode missing values
ts <- ifelse(ts == -99.99, NA, ts)

# Convert vector to time series object
ts <- ts(ts, start=c(1958,3), frequency=12)

# Plot the time series
plot.ts(ts)
```

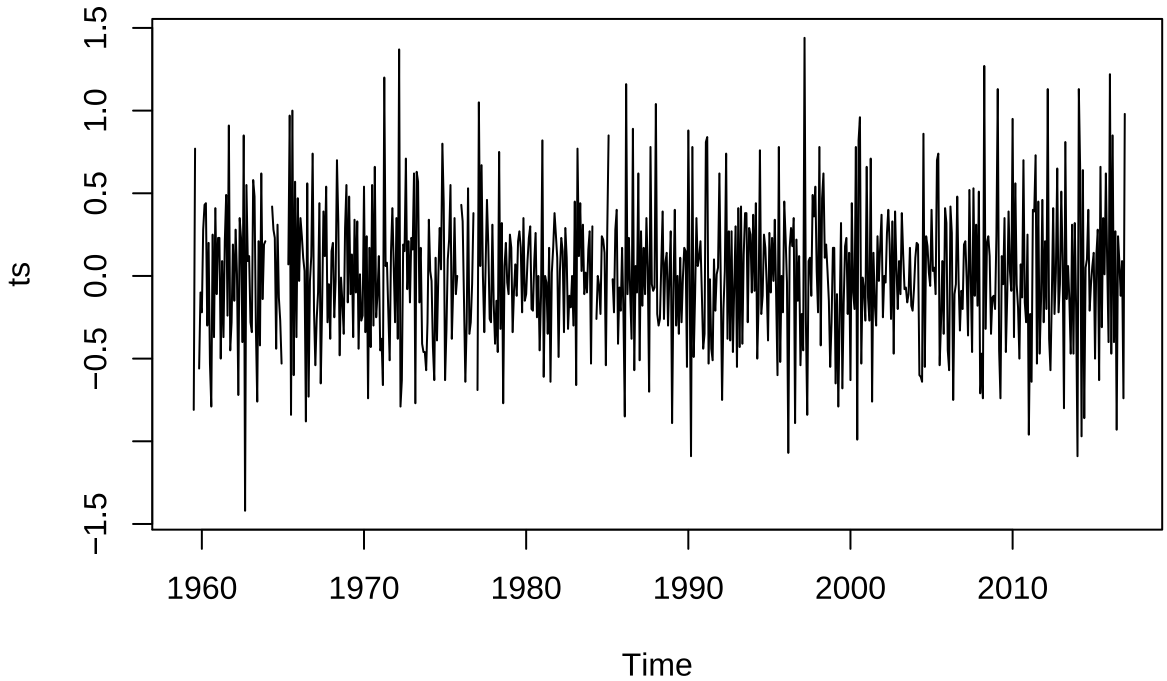


This plot clearly indicates an upward trend with seasonality. This implies that we will need to perform differencing at lag 1 to remove the trend and differencing at lag 12 to remove seasonality. Since variance appears to stable, we need not perform a data transformation.

```r
# Remove trend
ts <- diff(ts, 1)
plot.ts(ts)
```

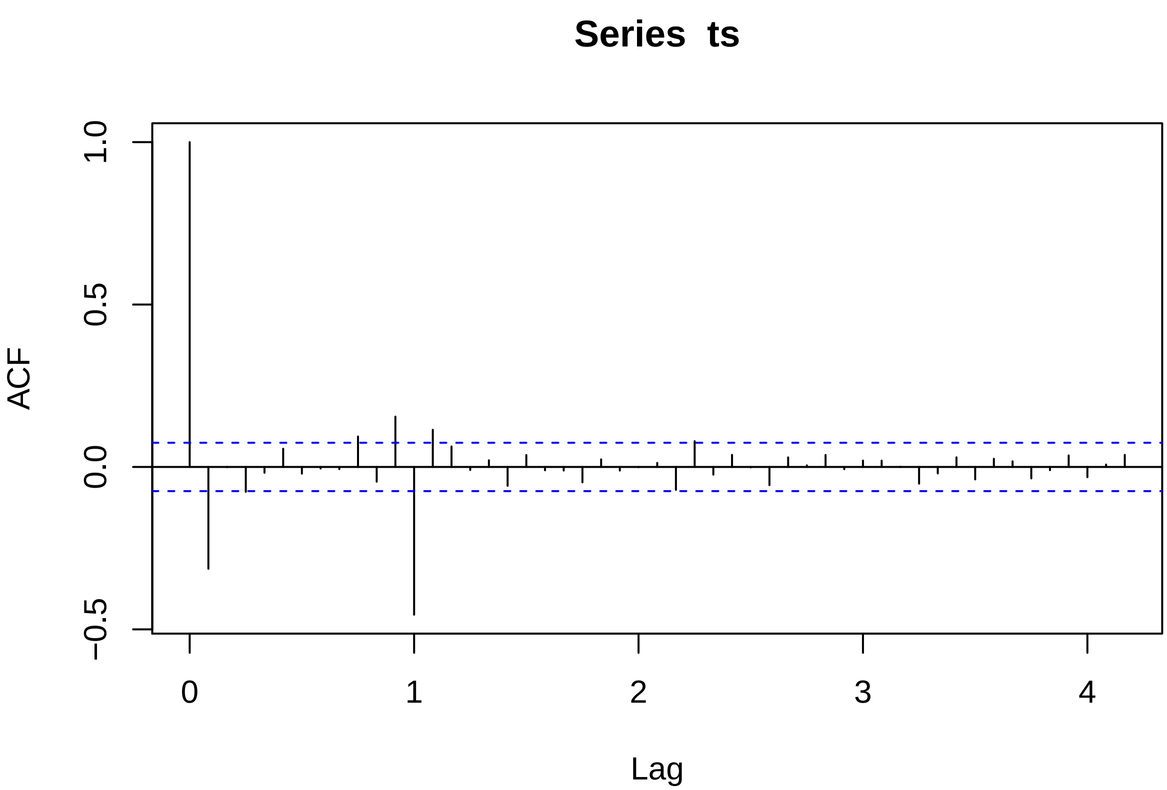



```r
# Remove Seasonality
ts <- diff(ts, 12)
plot.ts(ts)
```

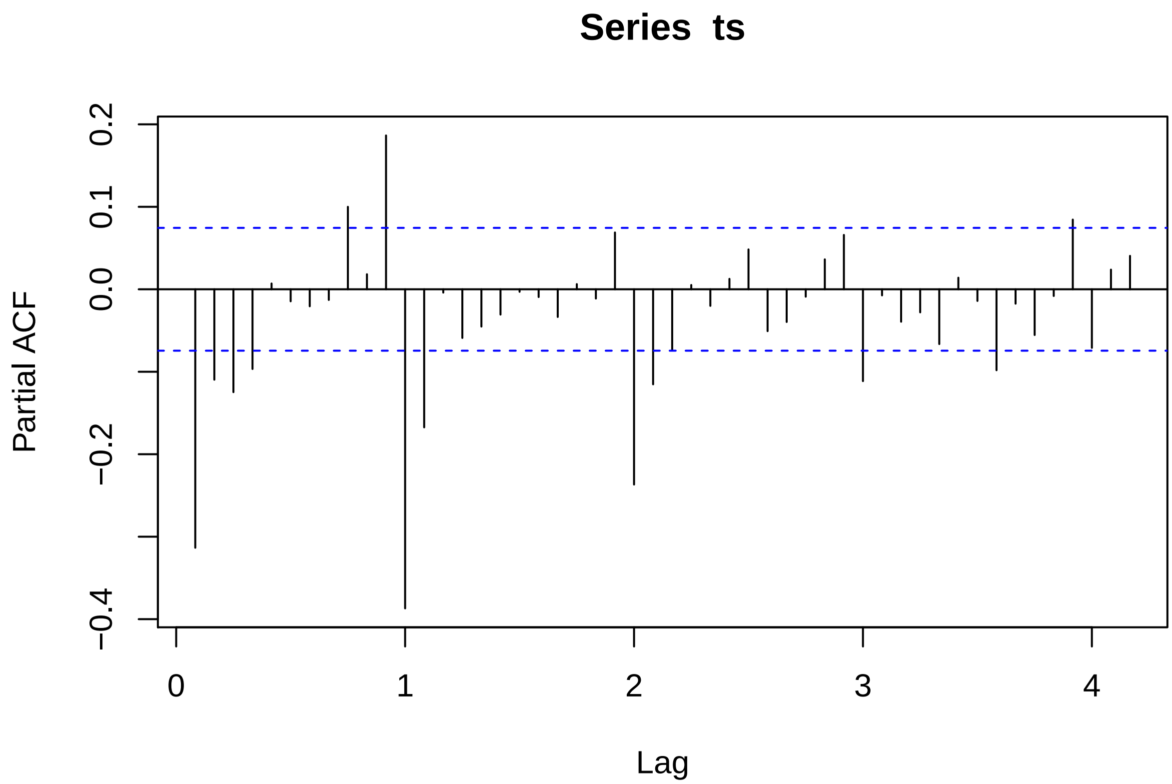


After differencing, the model appears stationary. We now examine ACF and PACF to preliminarily identify our model.

```
acf <- acf(ts, lag = 50, type = "correlation", plot = TRUE, na.action = na.pass)
```



```r
pacf <- pacf(ts, lag = 50, plot = TRUE, na.action = na.pass)
```




```r
arima(ts)
```

```
##
## Call:
## arima(x = ts)
##
## Coefficients:
##       intercept
##          0.0070
## s.e.     0.0167
##
## sigma^2 estimated as 0.1878:  log likelihood = -392.11,  aic = 788.22
```

From these plots, we should consider $AR(1)$, $AR(2)$, $MA(1)$, and $ARMA(p, q)$ $p \leq 2,\ q \leq 1$ models.