STA6856 **HW #4** **Instructor: Tharindu De Alwis**

Farooq Mahmud

Student Name (print): ______________________________

Question:	1	2	3	4	Total
Points:	10	10	20	20	60
Score:					

1. (10 points) Consider the following time series model.

$$Y_t = 0.25y_{t-1} - 0.25Y_{t-12} + 0.0625Y_{t-13} + e_t - 0.1e_{t-1} + 0.1e_{t-12} \text{ - } 0.01e_{t-13}$$

   (a) Recognize the model as $ARIMA(p, d, q) \times (P, D, Q)$ model. That is, what are the values for p, d, q, P, D, and Q.

   (b) Write down all the coefficient values using the standard notation. That is, ϕ's, θ's, Φ's, and Θ's.

2. (10 points) An AR model has AR characteristic polynomial

$$(1 - 1.6x + 0.7x^2)(1 - 0.8x^{12})$$

   (a) Is th model stationary?

   (b) Identify the model as a certain seasonal ARIMA model.

3. (20 points) Consider *electricity* dataset in the *TSA* R package that contains the monthly electricity generated in the United States.

   (a) Construct the sample ACF of the data, is the data stationary?

   (b) Calculate the sample ACF of the first difference of the logged transformed series. Is the seasonality visible in this display? If so, what is your seasonal component?

   (c) Plot the time series of seasonal difference, use $s = 12$ or the s value you have suggested in part (b), and first difference of the logged series. Does a stationary model seem appropriate now?

   (d) Display the sample ACF of the series after a seasonal difference and a first difference have been taken of the logged series in part (d). What model(s) might you consider for the electricity series?

4. (20 points) Consider the air passenger miles time series in *TSA* R package. The file is named *airmiles*.

   (a) Create the sample ACF and PACF for the *airmiles* data. Suggest the candidate values for p, d, q, P, D, Q, and s.

   (b) Fit $ARIMA(0, 1, 1) \times (0, 1, 0)_{12}$ and assess its adequacy.

   (c) Use auto.arima{forecast} to select the parameters p, d, q, P, D, Q, and s.

   (d) Refit the selected model from part(c) using arima{stats} function and assess its adequacy.

1 $Y_T = 0.25Y_{t-1} - 0.25Y_{t-12} + 0.0625Y_{t-13}$
$+e_t - 0.1e_{t-1} + 0.1e_{t-12} - 0.01e_{t-13}$

$Y_{t-1} \Rightarrow \phi_1 = 0.25 \rightarrow$ non-seasonal AR(1) $\Rightarrow$ $p=1$

$Y_{t-12} \Rightarrow \Phi_1 = -0.25 \rightarrow$ Seasonal AR(1) $\Rightarrow$ $P=1$

$e_{t-1} \Rightarrow \theta_1 = -0.1 \rightarrow$ nonseasonal MA(1) $\Rightarrow$ $q=1$

$e_{t-12} \Rightarrow \Theta_1 = 0.1 \rightarrow$ seasonal MA(1) $\Rightarrow$ $Q=1$

No differencing applied $\Rightarrow$ $d = D = 0$

2A $1-1.6x+0.7x^2 \Rightarrow$ get the roots $\Rightarrow$ $x = \frac{1.6\sqrt{(-1.6)^2-4(0.7)}}{2(0.7)} = \frac{1.6 \pm \sqrt{2.56-2.8}}{1.4}$

$= \frac{1.6 \pm \sqrt{-0.24}}{1.4} = \frac{1.6 \pm i\sqrt{0.24}}{1.4}$

$|x| = \frac{\sqrt{1.6^2+0.24}}{1.4} = \frac{\sqrt{2.56+0.24}}{1.4} \approx \frac{1.673}{1.4} \Rightarrow$ greater than 1 so stationary

$1-0.8x^{12} = 0 \Rightarrow \frac{-0.8x^{12}}{-0.8} = \frac{-1}{-0.8} \Rightarrow x^{12} = \frac{1}{0.8} \Rightarrow x = \left(\frac{1}{0.8}\right)^{\frac{1}{12}} \approx 1.019$

$\Rightarrow$ greater than 1 so stationary

Since all roots of polynomial lie outside of unit circle, model is stationary

2B $\phi_1 = 1.6 \quad \phi_2 = -0.7 \Rightarrow p=2$

$\Phi_1 = 0.8 \Rightarrow P=1$

no MA terms; no differencing $\Rightarrow d = D = 0$

$\Rightarrow$ ARIMA $(2,0,0) \times (1,0,0)_{12}$

HW4

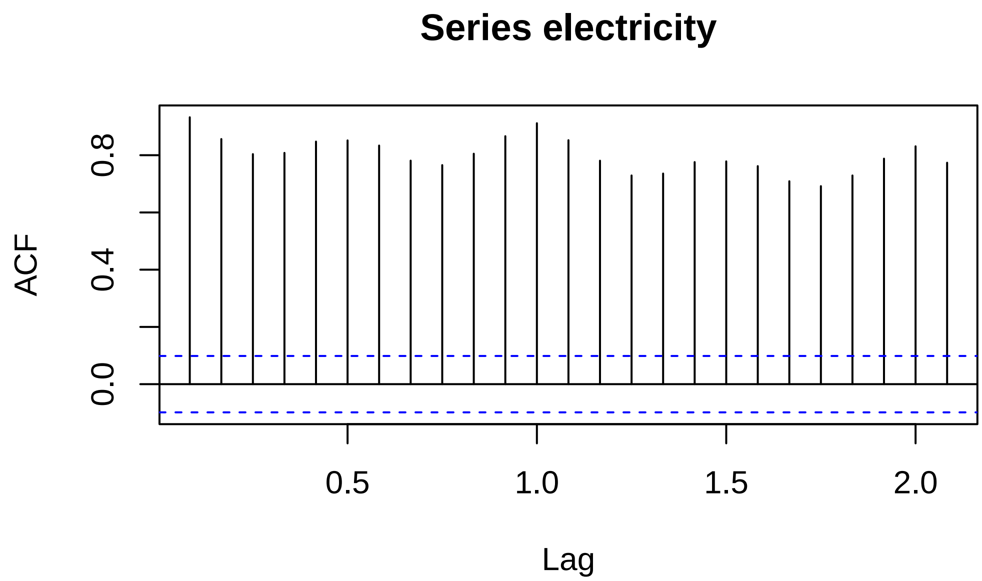
Farooq Mahmud

Problem 3a

```r
library(TSA)
```

```

Attaching package: 'TSA'

The following objects are masked from 'package:stats':

    acf, arima

The following object is masked from 'package:utils':

    tar
```

```r
data(electricity)
acf(electricity)
```

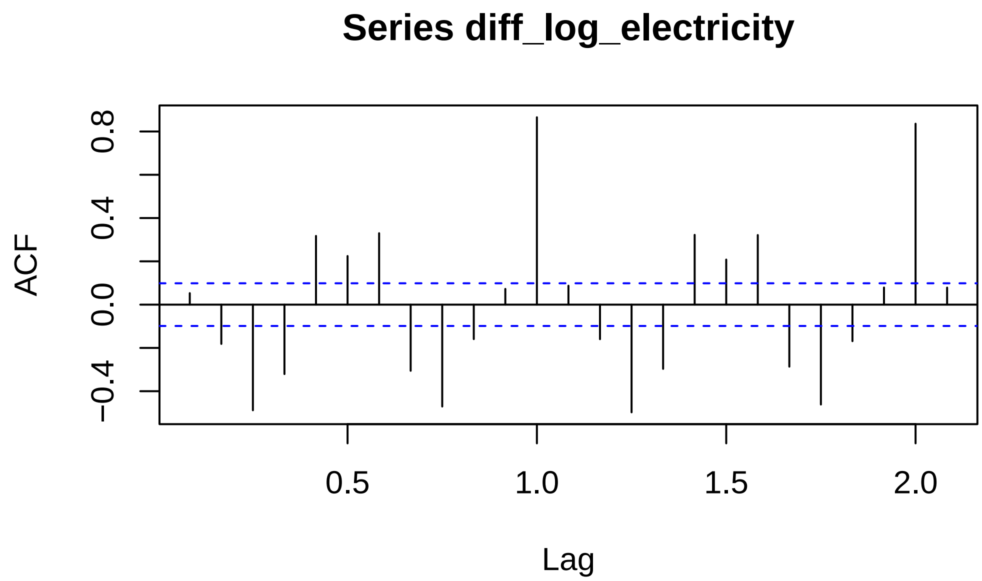


The data does not appear to be stationary based on the ACF plot because:

- Significant ACF at all lags.
- Seasonal pattern with peaks every 12 lags.
- Slow, if any decay in ACF.

Problem 3b

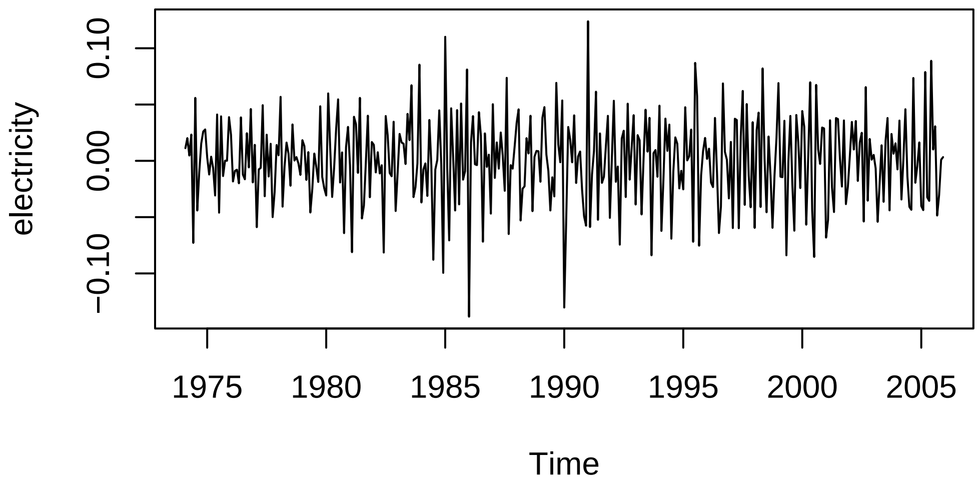
```
log_electricity <- log(electricity)
diff_log_electricity <- diff(log_electricity)
acf(diff_log_electricity)
```



Yes, seasonality is visible. Since the electricity data is monthly and there are significant ACF's at every 12 lags, the seasonal component is 12 months.

Problem 3c

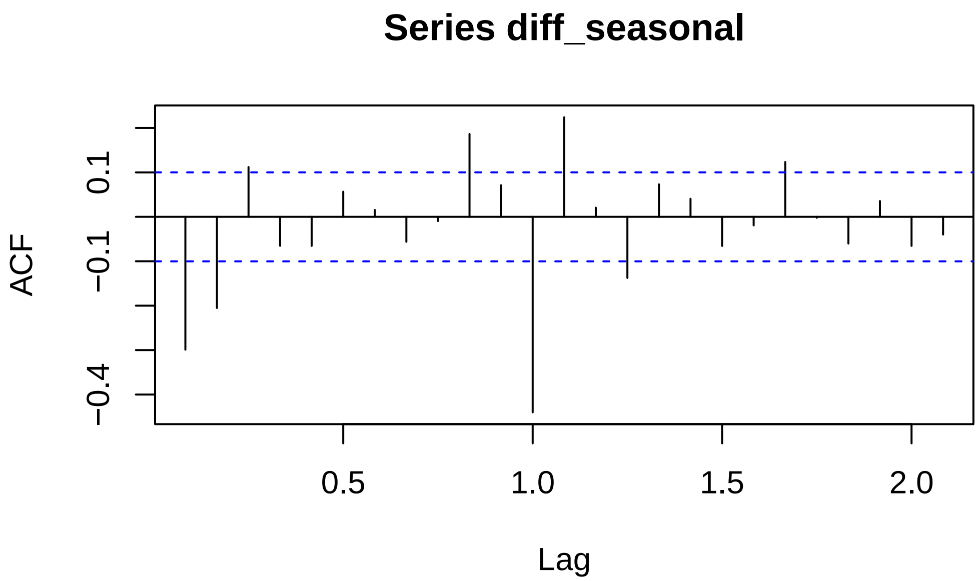
```
diff_seasonal <- diff(diff_log_electricity, lag = 12)
plot(diff_seasonal)
```



Yes, the plot looks stationary now.

Problem 3d

```
acf(diff_seasonal)
```



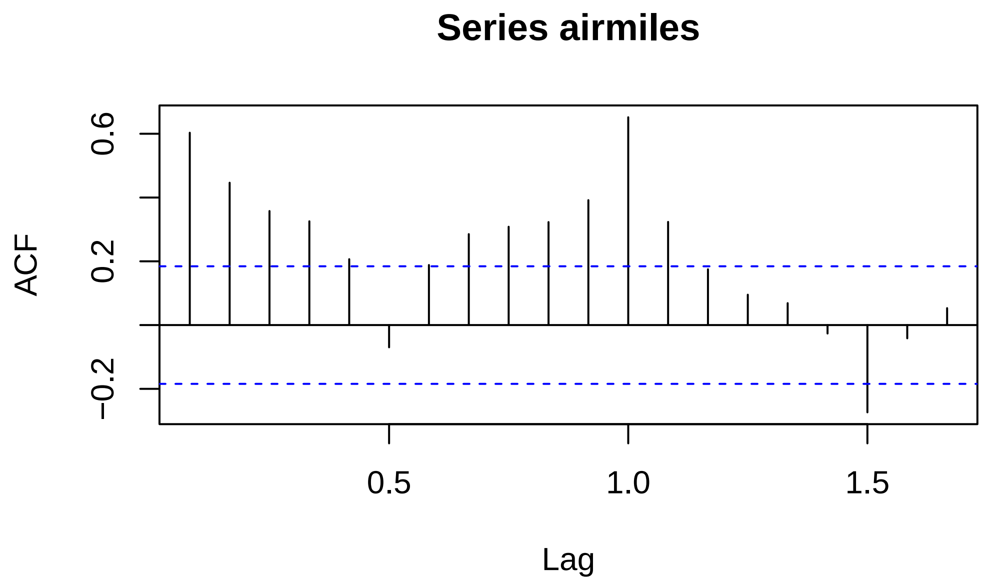
- $p = 0$ because no gradual decay appears in ACF.
- $d = 1$ because we needed first differencing to remove trend/non-stationarity.
- $q = 1$ because of a sharp drop in ACF after lag 1.
- $P = 0$ because no gradual decay or spike pattern.
- $D = 1$ because seasonal differencing of order 1 is needed.
- $Q = 1$ because the large spike at lag 12.
- Seasonal period of 12

Therefore a model to consider would be:

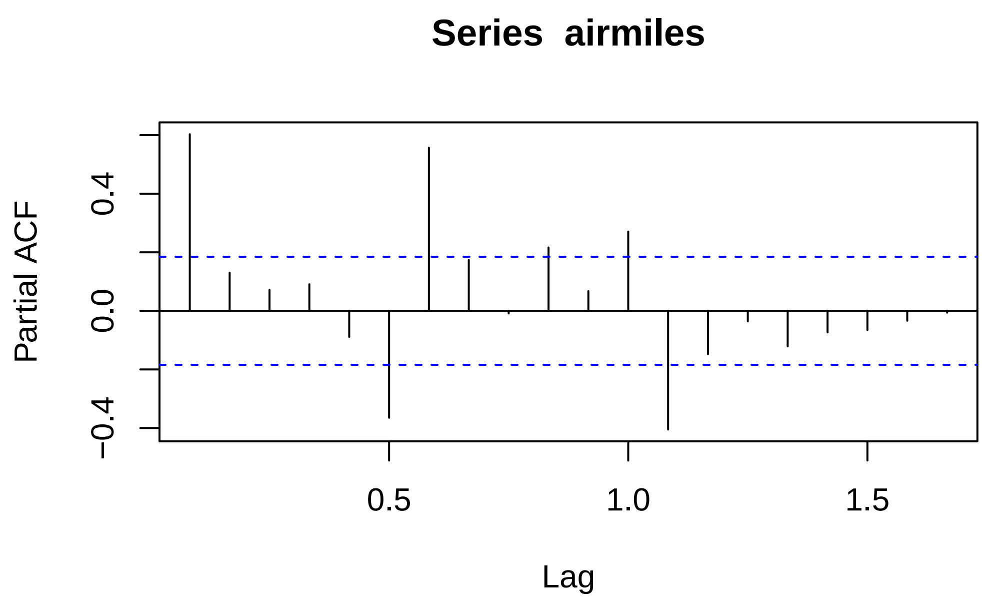
$ARIMA(0, 1, 1)X(0, 1, 1)_{12}$

Problem 4a

```
library(TSA)
data("airmiles")
acf(airmiles)
```




```r
pacf(airmiles)
```



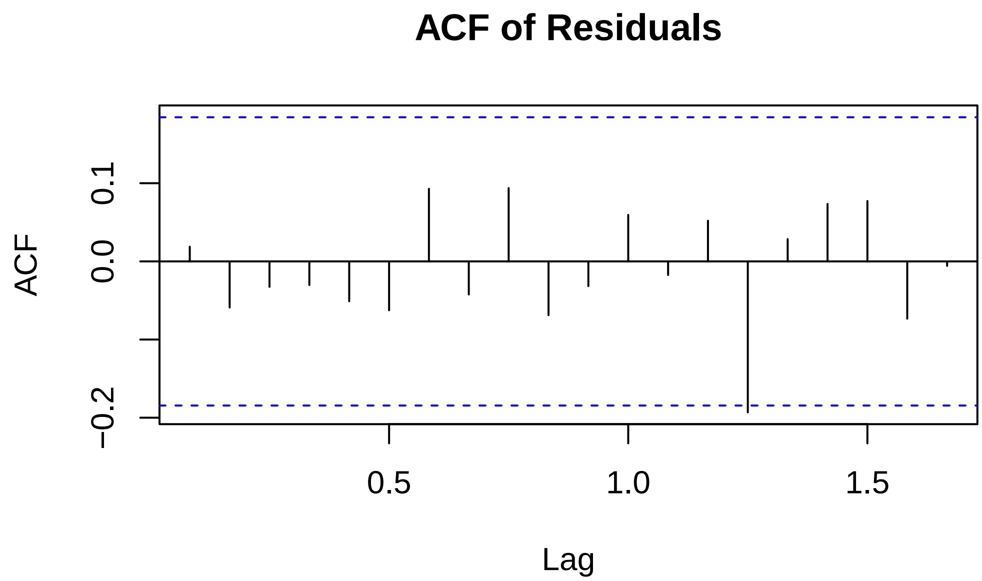
Just going by the ACF and PACF:

- $p = 0$ because there is no sharp cutoff in PACF.
- $d = 1$ because the series appears non-stationary and a first differencing is likely needed.
- $q = 1$ because ACF cuts of after lag 1 which suggests $MA(1)$.
- $P = 0$ because PACF does not show a seasonal spike at lag 12 suggesting no strong seasonal AR.
- $D = 1$ because ACF shows spikes every 12 lags suggesting one seasonal difference is needed.
- $Q = 1$ because of the significant spike at lag 12 suggesting seasonal $MA(1)$.
- $s = 12$ because this is monthly data.

Problem 4b

```
fit_b <- arima(airmiles,
               order = c(0, 1, 1),
               seasonal = list(order = c(0, 1, 1), period = 12))

ts.plot(residuals(fit_b), main = "Residual Plot", ylab = "Residuals")
```


```
acf(residuals(fit_b), main = "ACF of Residuals")
```

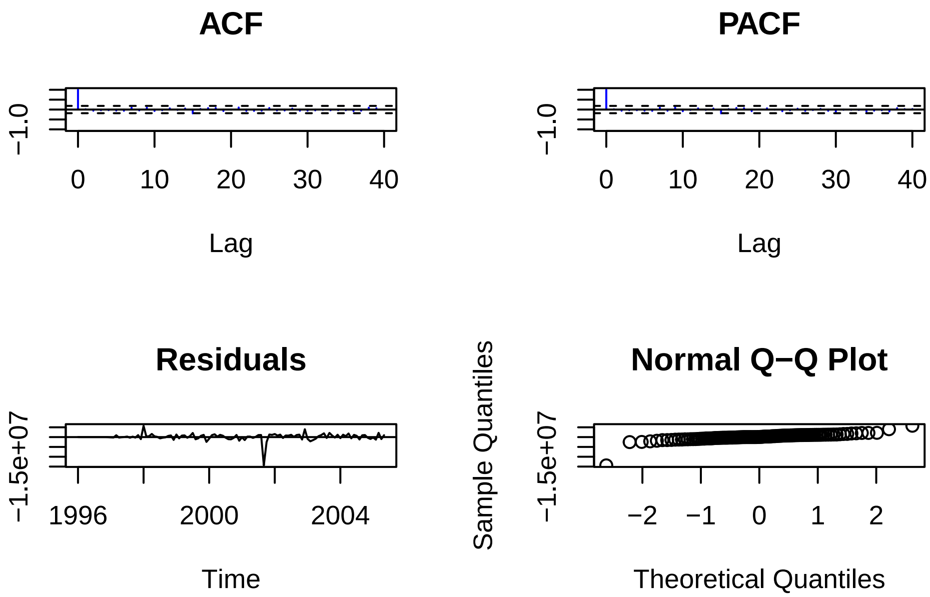



```r
library(itsmr)
```

```

Attaching package: 'itsmr'

The following objects are masked from 'package:TSA':

    periodogram, season
```

```r
test(fit_b$residuals)
```

```
Null hypothesis: Residuals are iid noise.
Test                        Distribution Statistic   p-value
Ljung-Box Q                Q ~ chisq(20)     12.78    0.8868
McLeod-Li Q                Q ~ chisq(20)      0.91         1
Turning points T      (T-74)/4.4 ~ N(0,1)        73     0.822
Diff signs S          (S-56)/3.1 ~ N(0,1)        61    0.1048
Rank P            (P-3164)/201.5 ~ N(0,1)      3340    0.3824
```



$ARIMA(0,1,1)X(0,1,0)_{12}$ appears to be a reasonable model because:

- The residual plot shows no obvious trend or seasonality.
- The ACF of residuals has all non-significant values.
- For the box test, $p > 0.05$ therefore the null hypothesis that the residuals resemble white-noise cannot be rejected.
- The Q-Q plot shows the data is normal but there is an outlier.

Problem 4c

```
library(forecast)
```

```
Registered S3 method overwritten by 'quantmod':
  method            from
  as.zoo.data.frame zoo
```

```
Registered S3 methods overwritten by 'forecast':
  method       from
  fitted.Arima TSA
  plot.Arima   TSA
```

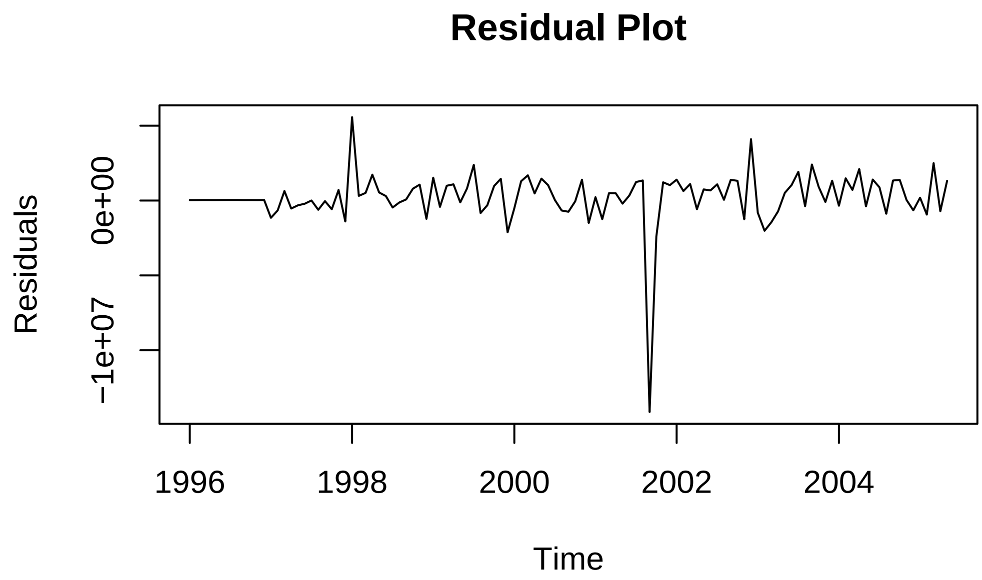
```
Attaching package: 'forecast'


The following object is masked from 'package:itsmr':

    forecast
```

```
fit_auto <- auto.arima(airmiles)
summary(fit_auto)
```

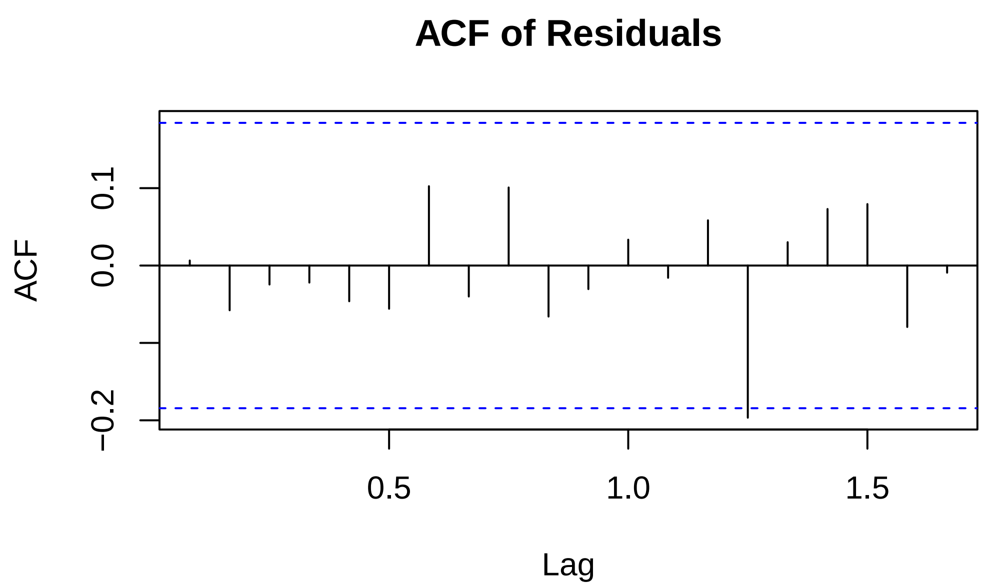
```
Series: airmiles
ARIMA(1,0,1)(0,1,1)[12] with drift

Coefficients:
         ar1      ma1     sma1      drift
      0.9021  -0.3742  -0.6902  108191.18
s.e.  0.0575   0.1226   0.1207   37540.33

sigma^2 = 3.588e+12:  log likelihood = -1605.31
AIC=3220.62   AICc=3221.25   BIC=3233.69

Training set error measures:
                    ME    RMSE      MAE        MPE     MAPE      MASE
Training set -10206.84 1754930 951540.7 -0.2950434 2.522186 0.3177697
                   ACF1
Training set 0.01256401
```

Problem 4d

```
model_d <- arima(airmiles,
                 order = c(1, 0, 1),
                 seasonal = list(order = c(0, 1, 1), period = 12))

ts.plot(residuals(model_d), main = "Residual Plot", ylab = "Residuals")
```

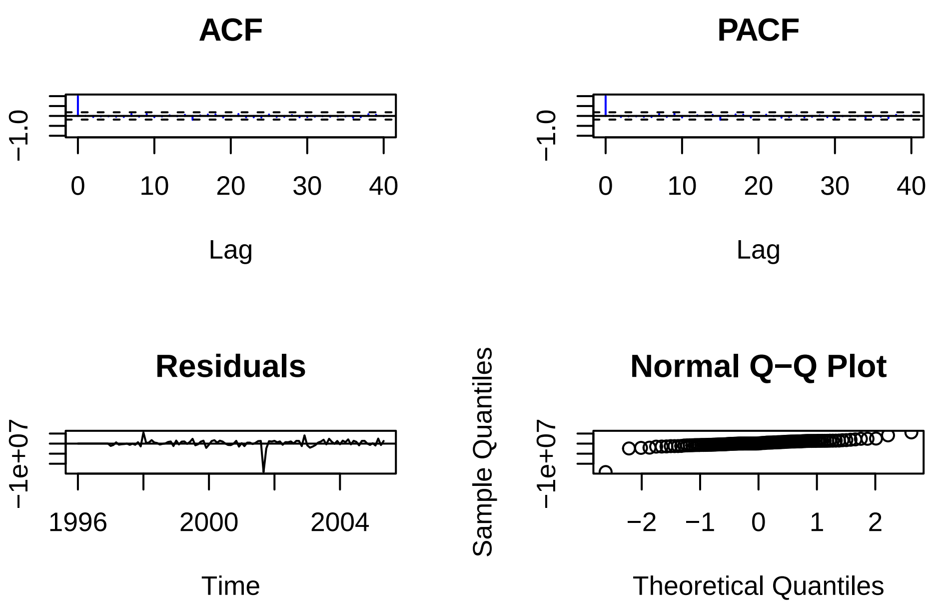



```r
acf(residuals(model_d), main = "ACF of Residuals")
```



```r
test(model_d$residuals)
```

```
Null hypothesis: Residuals are iid noise.
Test                          Distribution Statistic   p-value
Ljung-Box Q                  Q ~ chisq(20)     12.93    0.8803
McLeod-Li Q                  Q ~ chisq(20)      1.13         1
Turning points T      (T-74)/4.4 ~ N(0,1)        74         1
Diff signs S          (S-56)/3.1 ~ N(0,1)        64    0.0094 *
Rank P            (P-3164)/201.5 ~ N(0,1)      3541    0.0614
```



$ARIMA(1,0,1)X(0,1,1)_{12}$ suggested by `auto.arima` appears to be a reasonable model because:

- The residual plot shows no obvious trend or seasonality.
- The ACF of residuals has all non-significant values.
- For the box test, $p > 0.05$ therefore the null hypothesis that the residuals resemble white-noise cannot be rejected.
- The Q-Q plot shows the data is normal but there is an outlier.