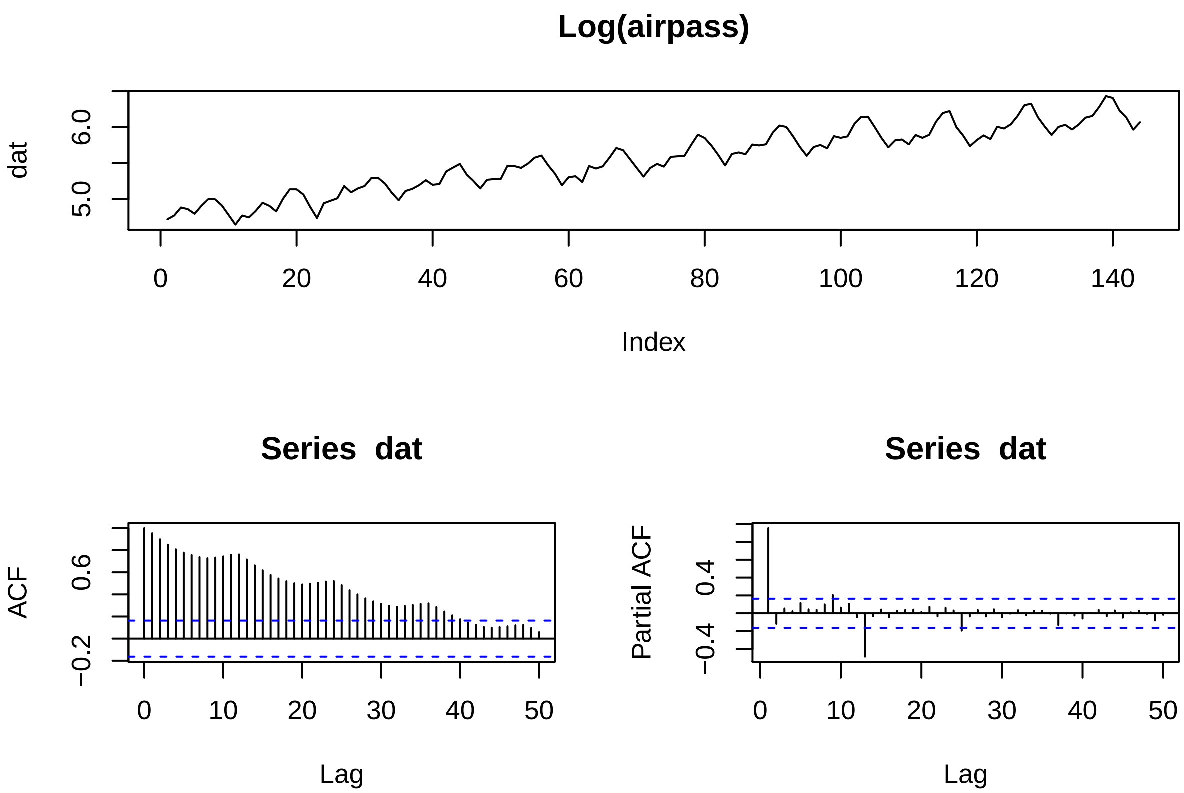
CH6-SARIMA models

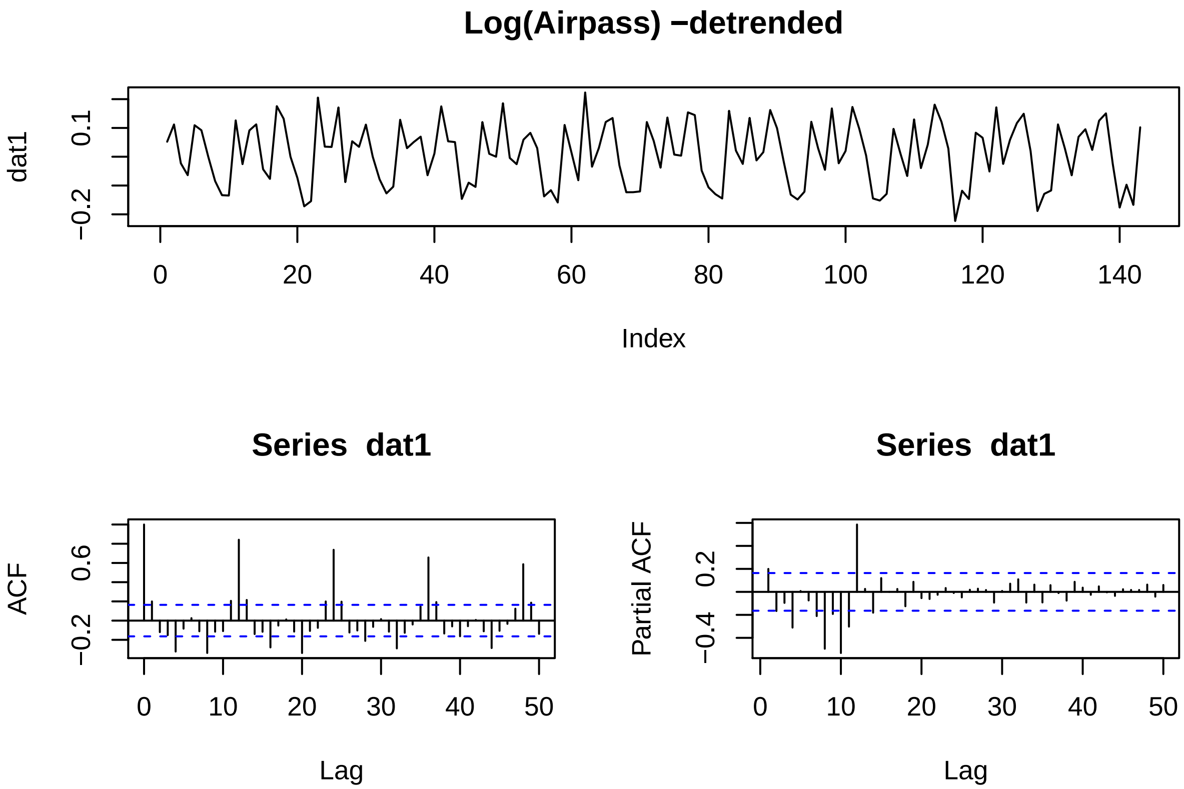
Here we present how to fit sARIMA model in R. The data is airpass data, and as we have seen in Box-Cox transformation we will take the log-transformation before fit SARIMA models.

```
library(itsmr)
data = airpass;
dat = log(data);
layout(matrix(c(1,1,2,3), 2, 2, byrow = TRUE))
plot(dat, type="l")
title("Log(airpass)")
acf(dat, lag=50);
pacf(dat, lag=50);
```



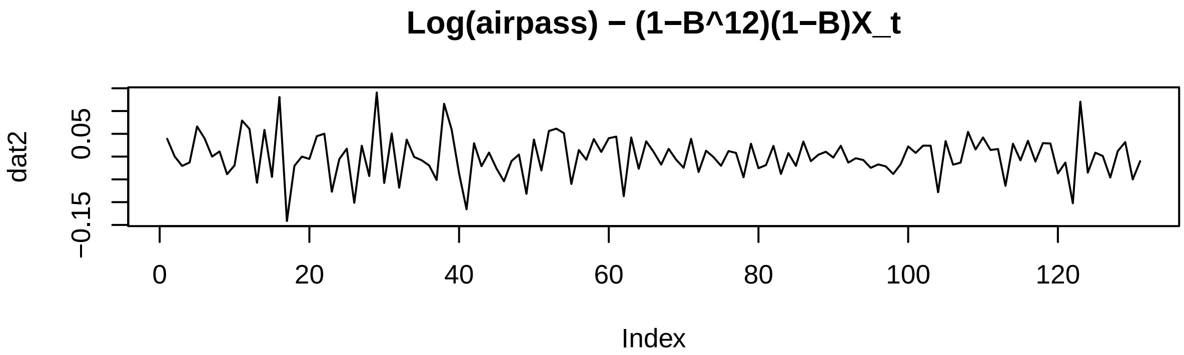
Since linear trend and sesaonality is clear, we will remove them by differencing.

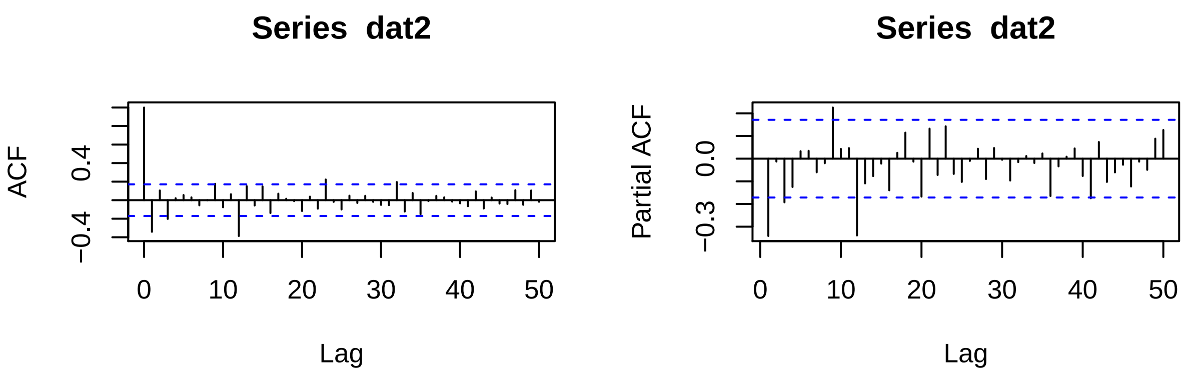
```
dat1 = diff(dat, 1);
layout(matrix(c(1,1,2,3), 2, 2, byrow = TRUE))
plot(dat1, type="l")
title("Log(Airpass) -detrended")
acf(dat1, lag=50);
pacf(dat1, lag=50);
```



From ACF plots, observe that ACF is decaying at evey 12th lag, and decreasing so P=1, ACF(1) and PACF(1) is also away from zero so take p=1. To confirm this take seasonal differencing

```
dat2 = diff(dat1, 12);
layout(matrix(c(1,1,2,3), 2, 2, byrow = TRUE))
plot(dat2, type="l")
title("Log(airpass) - (1-B^12)(1-B)X_t")
acf(dat2, lag=50);
pacf(dat2, lag=50);
```





Most of seasonality disappeared, but remained at lag 12. ACF(1) and PACF(1) is still nonzero, henct p=1 seems reasonable. Seasonality with period 12 is observed and consider SARIMA(1,1,0)X(1, 0, 0). Or we can ignore AR(1) coefficient and fit SARIMA(0,1,1)(1,0,0).

```
fit.1 = arima(dat, order = c(1,1,0), seasonal=list(order=c(1,0,0), period=12))
fit.1
```

```
##
## Call:
## arima(x = dat, order = c(1, 1, 0), seasonal = list(order = c(1, 0, 0), period = 12))
##
## Coefficients:
##           ar1    sar1
##       -0.2905  0.9287
## s.e.   0.0822  0.0229
##
## sigma^2 estimated as 0.001777:  log likelihood = 237.94,  aic = -469.89
```

```
fit.2 = arima(dat, order = c(0,1,0), seasonal=list(order=c(1,0,0), period=12))
fit.2
```

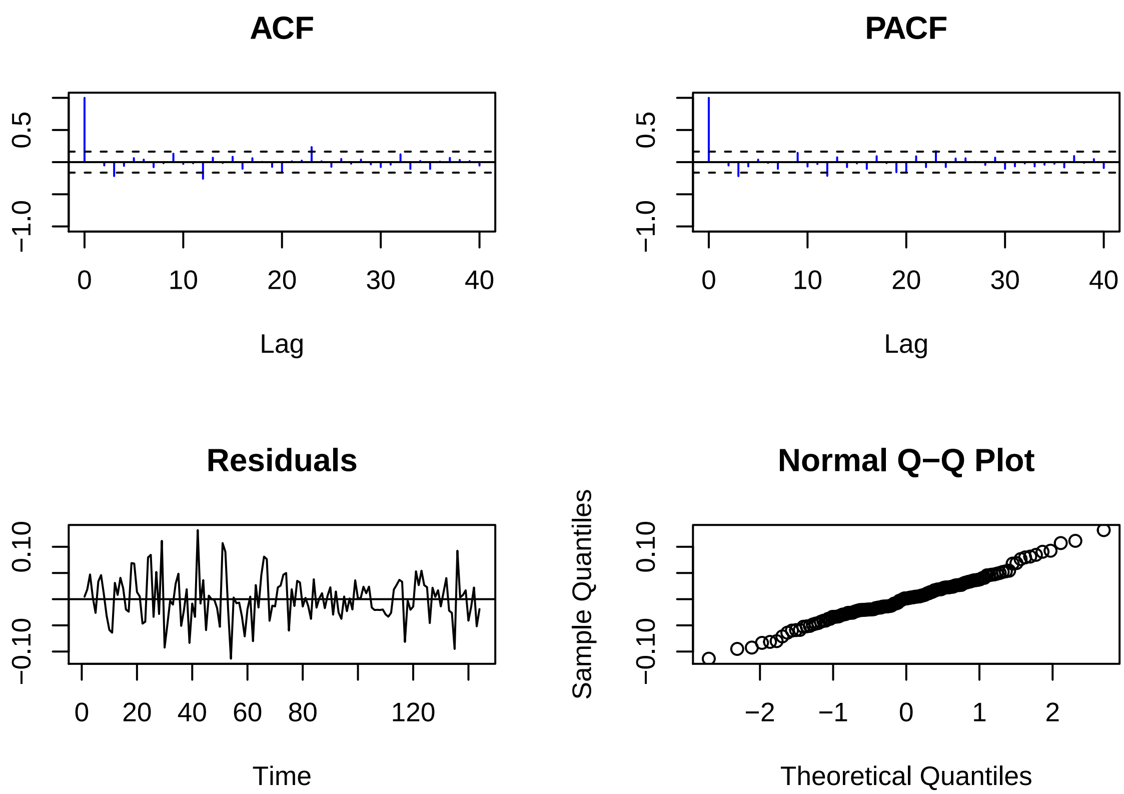
```
##
## Call:
## arima(x = dat, order = c(0, 1, 0), seasonal = list(order = c(1, 0, 0), period = 12))
##
## Coefficients:
##         sar1
##       0.9032
## s.e.  0.0278
##
## sigma^2 estimated as 0.001978:  log likelihood = 232.08,  aic = -460.17
```

AIC is smaller for SARIMA(1,1,0)(1,0,0), so we take it as the final model.

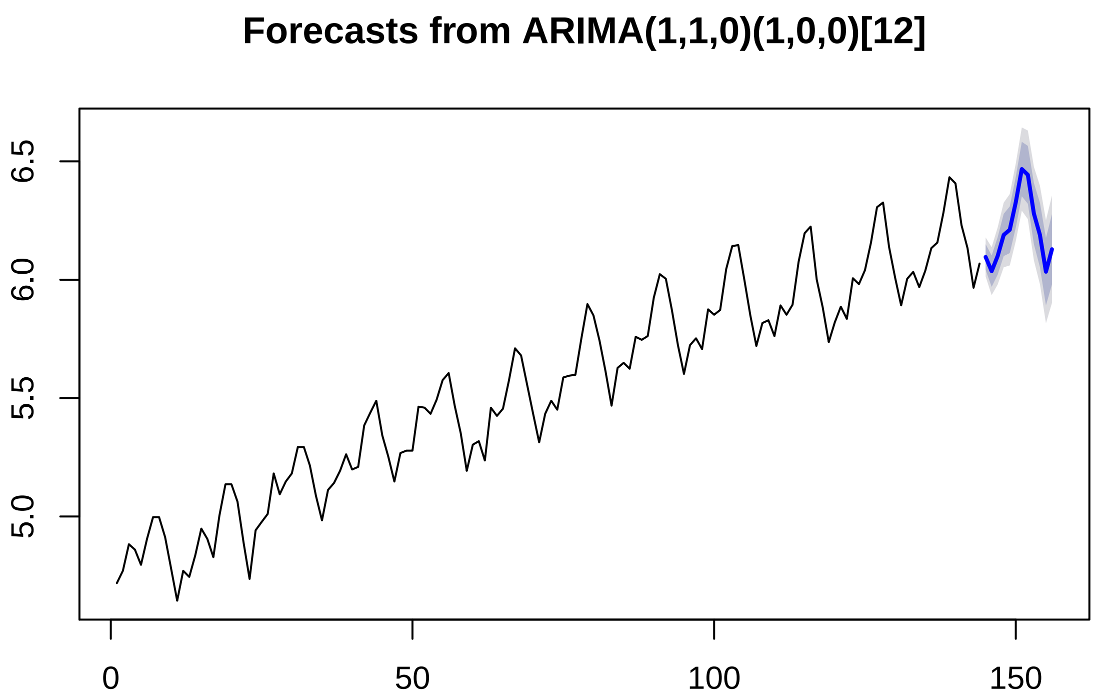
```
library(itsmr)
test(residuals(fit.1));
```

```
## Null hypothesis: Residuals are iid noise.
## Test                        Distribution Statistic   p-value
## Ljung-Box Q                Q ~ chisq(20)     31.83    0.0451 *
## McLeod-Li Q                Q ~ chisq(20)     40.23    0.0047 *
## Turning points T      (T-94.7)/5 ~ N(0,1)        94    0.8945
## Diff signs S        (S-71.5)/3.5 ~ N(0,1)        70    0.6661
## Rank P            (P-5148)/289.5 ~ N(0,1)      4996    0.5995
```

```
detach("package:itsmr")
library(forecast)
```




```
plot(forecast(fit.1, h=12))
```



Model selection by AICC/BIC finds

```
dat.ff = ts(dat, frequency=12);
auto.arima(dat.ff)
```

```
## Series: dat.ff
## ARIMA(0,1,1)(0,1,1)[12]
##
## Coefficients:
##           ma1     sma1
##       -0.4018  -0.5569
## s.e.   0.0896   0.0731
##
## sigma^2 estimated as 0.001371:  log likelihood=244.7
## AIC=-483.4   AICc=-483.21   BIC=-474.77
```

Automatic selection gives the best model as ARIMA(0, 1, 1)(0,1,1)[12].

Seasonal ARIMA model practice

We will consider the monthly time series observed from January 1992 to August 2003. Your analysis should include (but not limited to) the followings:

0. Open data set 'seasonal-example.csv'
1. Time plot, correlograms (ACF/PACF) and discuss key features of the data.
2. The best SARIMA$(p, d, q)(P, D, Q)$ model to describe the data and reasonings for model selection including model diagnostics/checking.
3. Forecasting of the next 4 months with 95% prediction interval. Please provide actual numbers also.