UNIVERSITY of LIMERICK

OLLSCOIL LUIMNIGH

Time Series Analysis

MS 4218

joseph.lynch@ul.ie

Outline

Model Specification

- Acf
- Pacf
- Eacf
- Dicky-Fuller unit root test
- Application to data series from Lecture 1

Sample auto-correlation: r_k, the estimator of ρ_k

In lecture 3, we defined the sample auto-correlation function as

$$r_k = \frac{\sum_{t=1}^{n-k}(Y_t - \bar{Y})(Y_{t-k} - \bar{Y})}{\sum_{t=1}^{n}(Y_t - \bar{Y})^2}, \text{ for } k = 1, 2 \ldots$$

At this stage, we had not met any AR or MA models, only white noise.

For white noise, this sample statistic varies from sample to sample, and via the Central Limit Theorem, follows a Normal distribution with ≈ 0 mean and variance inversely related to sample size, i.e.,

$$r_k \sim N\left(0, \frac{1}{n}\right).$$

Properties r_k, the estimator of ρ_k of AR models

Unlike white noise, AR and MA models do not have zero auto-correlation.

$$r_k \sim N\left(\rho_k, Var(r_k) = \frac{c_{kk}}{n}\right)$$

$Var(r_k)$ depends on the model in question but is inversely related to the sample size n through c_{kk} which is a complicated function of ρ_k $\forall$ k values.

For an AR(1) process, with theoretical $\rho_k = \phi^k$,

$$Var(r_1) = \frac{1-\phi^2}{n}.$$

If ϕ is close to 1, $Var(r_1) \downarrow$, and thus the estimate of ρ_1 is more accurate.

$Var(r_k)$ for an AR(1) process cont.

For large lags, i.e., large k,

$$Var(r_k) = \frac{1}{n}\frac{1+\phi^2}{1-\phi^2}.$$

If ϕ is close to 1, $(1-\phi^2)\downarrow$, and hence, $Var(r_k)\uparrow$, and thus the estimate of ρ_k is less accurate.

$$Corr(r_1, r_2) = 2\phi\sqrt{\frac{1-\phi^2}{1+2\phi^2-3\phi^4}}.$$

There can be significant auto-correlation between successive terms. Correlation is independent of n.

$Var(r_k)$ for an MA(1) process

$$
\begin{aligned}
Var(r_1) &= \frac{1}{n}(1 - 3\rho_1^2 + 4\rho_1^4). \\
Var(r_k) &= \frac{1}{n}(1 + 2\rho_1^2), \qquad \text{for } k > 1. \\
Cov(r_1, r_2) &= \frac{1}{n}2\rho_1(1 - \rho_1^2).
\end{aligned}
$$

We have $Var(r_k) > Var(r_1)$ for $k > 1$.

Like for the AR(1) process, the sample auto-correlations can be highly correlated.

$Var(r_k)$ for an MA(q) process

$$Var(r_k) = \frac{1}{n}\left(1 + 2\sum_{j=1}^{q}\rho_j^2\right), \qquad \text{for } k > q.$$

For an observed time series, ρ_j can be replaced by r_j.

Taking the square root yields the standard errors of r_k.

In general, if r_k is more than 2 standard errors from ρ_k, then it is not a good estimator of ρ_k.

Partial auto-correlation function (pacf): ϕ_{kk}

For an MA(q) process, $\rho_k = 0$ for $k > q$ lags. (TSLecture4a)

For an AR(p) process, $\rho_k \neq 0$ for $k > p$ lags. It $\downarrow$ but does not cut off. (TSLecture4b)

The pacf measures the correlation between Y_t and Y_{t-k} after removing the intervening variables.

$$\phi_{kk} = Corr(Y_t, Y_{t-k} | Y_{t-1}, Y_{t-2}, \ldots, Y_{t-k+1}).$$

By convention, the lag 1 partial auto-correlation $\phi_{11} = 1$.

Pacf cont.

The lag 2 partial auto-correlation

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

If stationary, for an AR(1), $\rho_k = \phi^k$, and hence

$$\phi_{22} = \frac{\phi^2 - \phi^2}{1 - \phi^2} = 0.$$

For an AR(1) process, $\phi_{kk} = 0 \; \forall \; k > 1$.

For an AR(p) process, $\phi_{kk} = 0 \; \forall \; k > p$.

Pacf and MA processes

The lag 2 partial auto-correlation

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

If stationary, for an MA(1), $\rho_1 = -\frac{\theta}{1+\theta^2}$, and $\rho_2 = 0$,

$$\phi_{22} = -\frac{\theta^2}{1 + \theta^2 + \theta^4}.$$

For an MA(1) process, $\phi_{kk} \downarrow$ exponentially fast as $k \uparrow$. Pacf of MA(q) behaves just like acf of AR(p).

Acf and pacf behaviour for ARMA models

	AR(p)	MA(q)	ARMA(p, q), $p > 0, q > 0$
ACF	Tails off	Cuts off after lag q	Tails off
PACF	Cuts off after lag p	Tails off	Tails off

AR(p) and MA(q) models can, in theory, be identified easily enough.

To overcome the difficulties with identification of ARMA models, the EACF has been developed.

EACF$=$ Extended auto-correlation function.

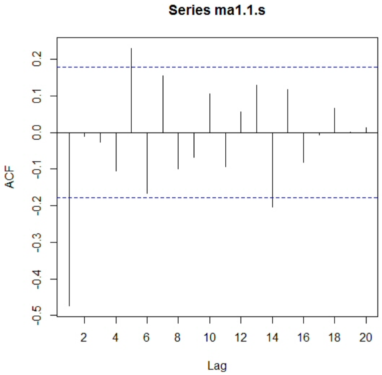
EACF

If the AR part of a mixed ARMA model is known, “filtering out" of the AR part from the observed time series results in a pure MA process that enjoys the cutoff property in the acf.

The eacf() in R outputs a table of 0’s and X’s.

We look for a theoretical pattern of a triangle of 0’s with the upper left-hand vertex corresponding to ARMA(p, q) order.

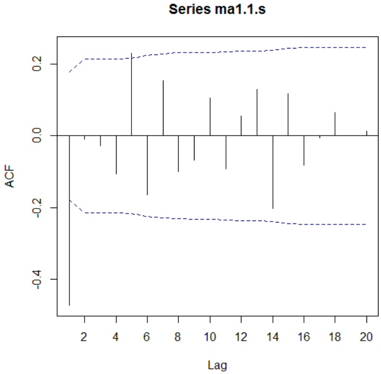
Sample acf of MA(1) process with $\theta = 0.9$ and limits of $\pm \frac{2}{\sqrt{120}}$



$\rho_1 = -0.497$. For $k > 1, \rho_k = 0$.

```
data(ma1.1.s)
macorr<-acf(ma1.1.s,xaxp=c(0,20,10))
macorr # Lags 1,5,14 non zero!
```

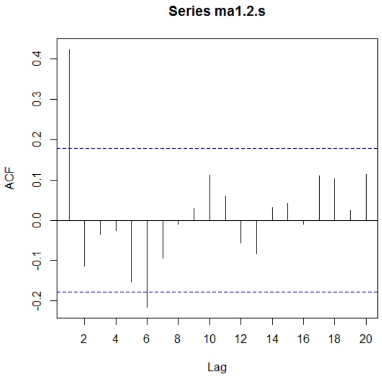
MA(1) sample acf with limits based on $Var(r_k)$ for MA process




```
acf(ma1.1.s,ci.type="ma",xaxp=c(0,20,10))
```

See page 7.

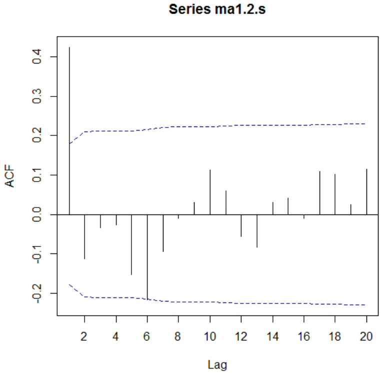
MA(1)sample acf with $\theta = -0.9$ and limits of $\pm\frac{2}{\sqrt{120}}$



$\rho_1 = 0.497$

```
data(ma1.2.s); acf(ma1.2.s,xaxp=c(10,20,10))
```

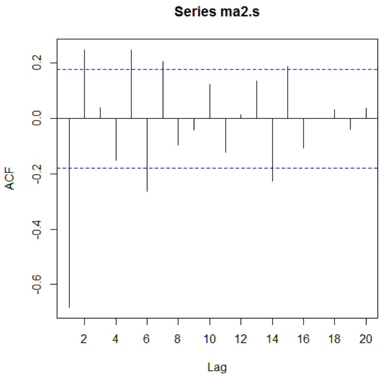
MA(1) sample acf with limits based on $Var(r_k)$ for MA process



$\rho_1 = 0.497$

```
data(ma1.2.s); acf(ma1.2.s,ci.type="ma",
xaxp=c(10,20,10))
```

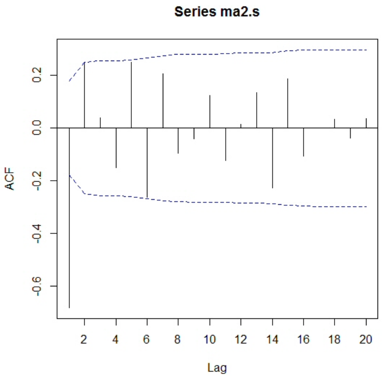
Sample acf of simulated MA(2) series with $\theta_1 = 1$ and $\theta_2 = -0.6$



$\rho_1 = -0.68$; $\rho_2 = 0.25$. $n = 120$; $\frac{2}{\sqrt{120}} = 0.1826$.

```
data(ma2.s)
ma2acf<-acf(ma2.s,xaxp=c(0,20,10))
ma2acf
```

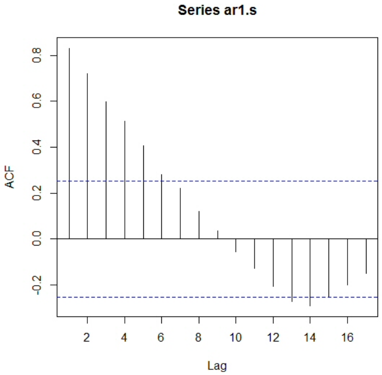
Sample acf of MA(2) series with alternative bounds




```
acf(ma2.s,ci.type="ma",xaxp=c(0,20,10))
```

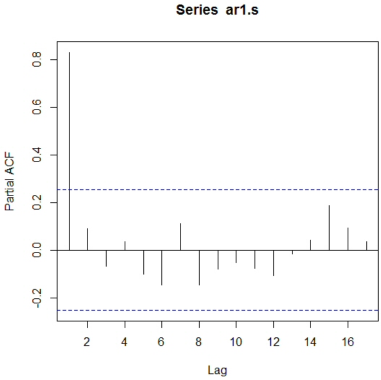
Much better result!

Sample acf of simulated AR(1) series with $\phi = 0.9$




```
data(ar1.s)
acf(ar1.s,xaxp=c(0,20,10))
```

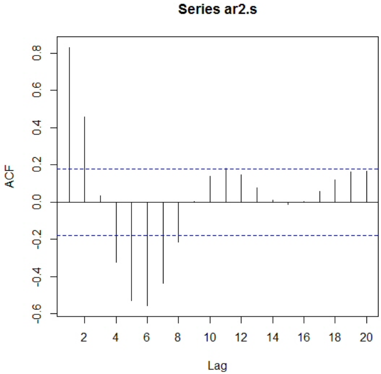
Sample pacf of simulated AR(1) series with $\phi = 0.9$




```
pacf(ar1.s,xaxp=c(0,20,10))
```

Cut-off at lag 1, typical of an AR(1) model.

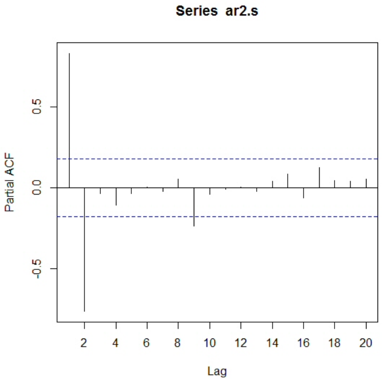
Sample acf of simulated AR(2) series with $\phi_1 = 1.5$ and $\phi_2 = -0.75$




```
acf(ar2.s,xaxp=c(0,20,10))
```

Damped cosine wave, theoretically not unexpected with $\phi_2 = -0.75$.

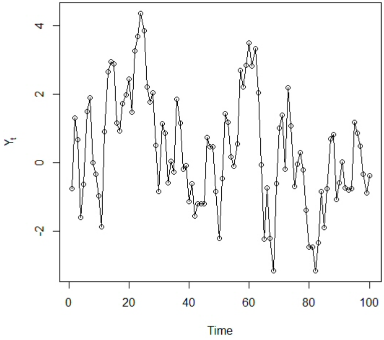
Sample pacf of simulated AR(2) series with $\phi_1 = 1.5$ and $\phi_2 = -0.75$




```
pacf(ar2.s,xaxp=c(0,20,10))
```

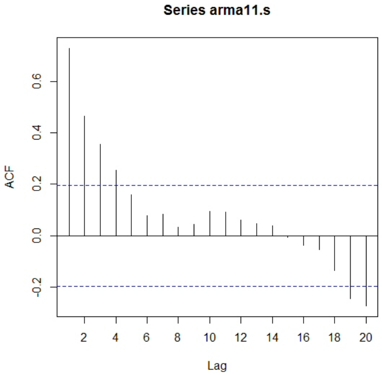
Typical of AR(2) process.

Simulated ARMA(1,1) series with $\phi = 0.6$ and $\theta = -0.3$



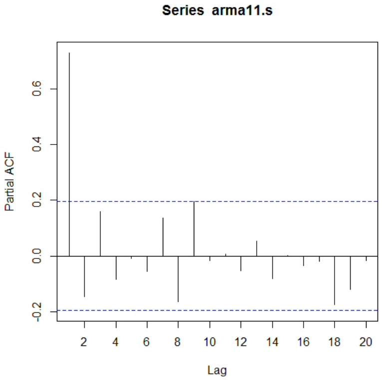

```
data(arma1.s)
plot(arma1.s, type='o',ylab=expression(Y[t]))
```

Sample acf of simulated ARMA(1,1) series




```
length(arma11.s)  #100
acf(arma11.s,xaxp=c(0,20,10))
```

Sample pacf of simulated ARMA(1,1) series




```
pacf(arma11.s,xaxp=c(0,20,10))
```

?AR(1) model.

Eacf of simulated ARMA(1,1) series

```
eacf(arma1.s)
```

```
AR/MA
  0 1 2 3 4 5 6 7 8 9 10 11 12 13
0 x x x x o o o o o o o  o  o  o
1 x o o o o o o o o o o  o  o  o
2 x o o o o o o o o o o  o  o  o
3 x x o o o o o o o o o  o  o  o
4 x o x o o o o o o o o  o  o  o
5 x o o o o o o o o o o  o  o  o
6 x o o o x o o o o o o  o  o  o
7 x o o o x o o o o o o  o  o  o
```

ARMA($p = 1$ or 2, $q = 1$) is more appropriate.

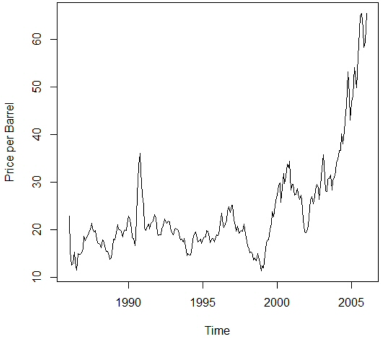
Sample acf and Non-stationary(NS) data

As sample acf implicitly assumes stationarity, it is difficult to know what it is measuring in non-statonary data.

NS data tends to slowly drift up or down with apparent trends.

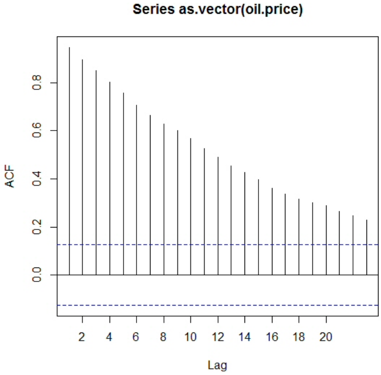
In any event, if data NS, sample acf will fail to die out rapidly as lags $\uparrow$.

Plot of non-stationary oil price series




```
data(oil.price)
plot(oil.price, ylab='Price per Barrel',type='l')
```

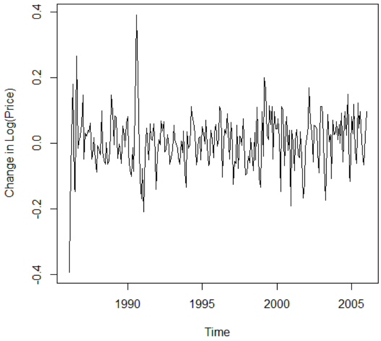
Sample acf of oil price series




```
acf(as.vector(oil.price),xaxp=c(0,20,10))
```

All values significantly far from zero.

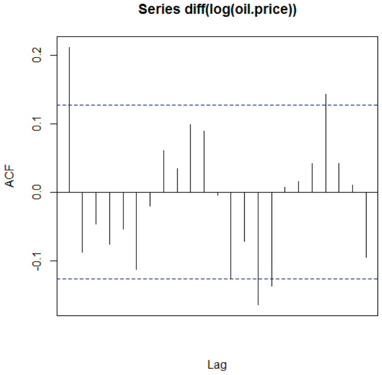
Plot of difference in log(oil price) series




```
plot(diff(log(oil.price)),
ylab='Change in Log(Price)',type='l')
```

More stationary in appearance with a couple of outliers.

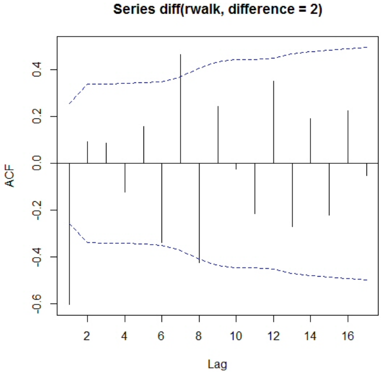
Sample acf of difference in log(oil price) series




```
acf(diff(log(oil.price)),xaxp=c(0,20,10))
```

After differencing, MA(1) seems likely, thus consider non-stationary IMA(1,1) model for log(original) series.

Over-differenced random walk




```
data(rwalk)
acf(diff(rwalk,difference=2),
ci.type='ma', xaxp=c(0,18,9))
```

Over-differencing introduces unnecessary correlations.

Over-differenced random walk cont.

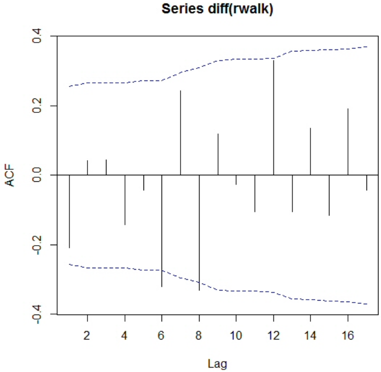
Over-differencing creates non-invertible models which have the effect of making parameter estimation difficult.

$$
\begin{aligned}
Y_t &= Y_{t-1} + e_t. \\
\nabla Y_t &= e_t. \\
\nabla^2 Y_t &= e_t - e_{t-1},
\end{aligned}
$$

ie., an IMA(2,1) model with one θ parameter to be estimated.

The correct model is IMA(1,1) with $\theta = 0$, or more succinctly an IMA(1,0) model.

Correctly differenced random walk




```
acf(diff(rwalk),ci.type='ma',
xaxp=c(0,18,9))
```

First difference looks like white noise.

Dickey-Fuller Unit Root test

Quantifies the degree of non-stationarity in data.

H_0: Data are non-stationary.
H_a: Data are stationary.

More extreme value of test statistic unlikely to have occurred by chance.

Thus the $p-$value is low and H_0, the data are non-stationary, is rejected.

```
adf.test(data,k)
```

The value of k, the lag order to calculate the test statistic, is outputted via

```
ar(diff(series))
```

Akaike's and Bayesian Information Criteria: AIC and BIC

AIC and BIC are model selection criteria.

For competing models, the model with the **lowest AIC or BIC** fits the data best.

$AIC = -2\log(\text{Maximum Likelihood}) + 2k$, where k is the number of parameters to be estimated.

$BIC = -2\log(\text{Maximum Likelihood}) + k\log(n)$, where n is the sample size.

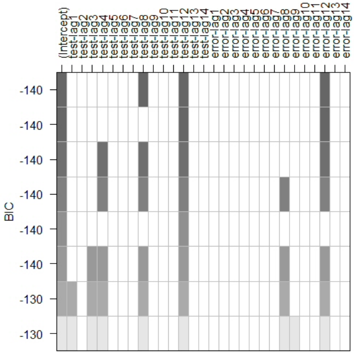
$k = p + q + 1$ for an ARMA(p, q) model with an intercept θ_0 term.

Corrected AIC: AIC_c

$$AIC_c = AIC + \frac{2(k+1)(k+2)}{n-k-2}.$$

The extra non-stochastic penalty term makes it the most efficient estimator of the correct model choice.

Best subset ARMA selection based on BIC



Each row corresponds to a model where the cells for the variables selected are shaded.

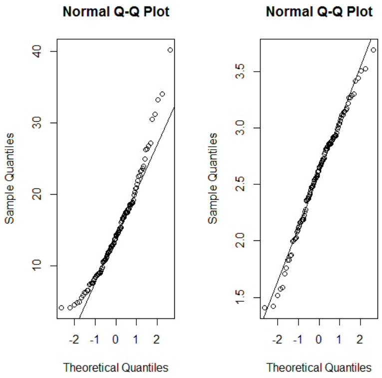
Top row best. Shading ↓ down the columns.

R code for best subset ARMA selection

ARMA(12,12) model simulation.

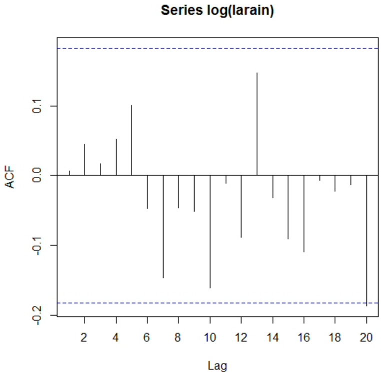
```
set.seed(92397)
test=arima.sim(model=list(ar=c(rep(0,11),0.8),
                          ma=c(rep(0,11),0.7)
                         )
                          ,n=120)
res=armasubsets(y=test,nar=14,nma=14,y.name='test',
ar.method='ols')
plot(res)
```

Log of LA rainfall Normal.




```
qqnorm(larain);qqline(larain) # slide 43 Lecture 3
shapiro.test(larain) # p = 0.0001614 < 0.05
#Reject H0
qqnorm(log(larain));qqline(log(larain))
```

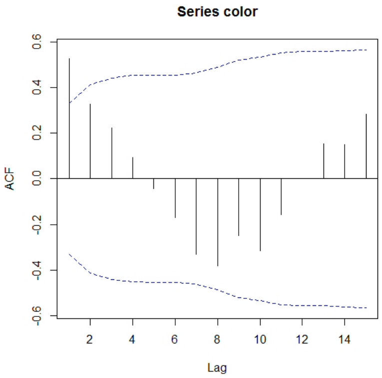
Sample acf of log (LA rainfall)



Log $\equiv$ to small λ in Box-Cox transformation.

No discernible dependence pattern in the log of LA rainfall series.

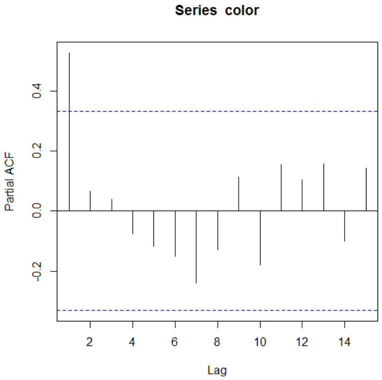
Chemical Process Colour Property series




```
data(color)
acf(color,ci.type="ma")
```

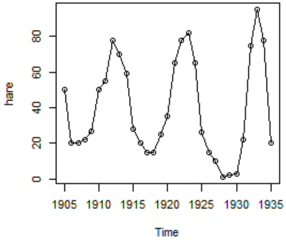
?MA(1), but damped cosine wave means we need to check for AR model.

Using pacf(), consider AR(1) model



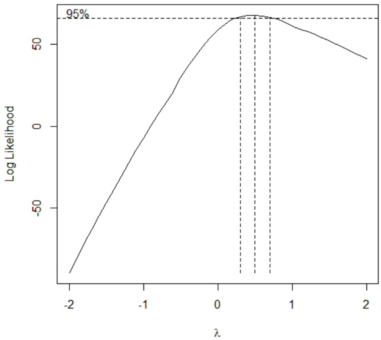

```
pacf(color)
```

Annual abundance of Canadian Hare series



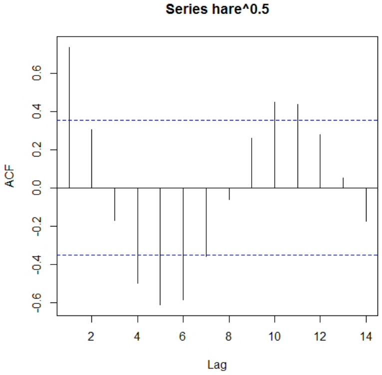

```
data(hare)
plot(hare)
ar(diff(hare))
adf.test(hare,k=8) # not stationary
```

Hare Analysis cont.




```
haretrans=BoxCox.ar(hare,lambda=seq(-2,2,0.05))
haretrans #mle of lambda =0.45
adf.test(hare^(0.45)) #p<0.01, i.e. stationary
```

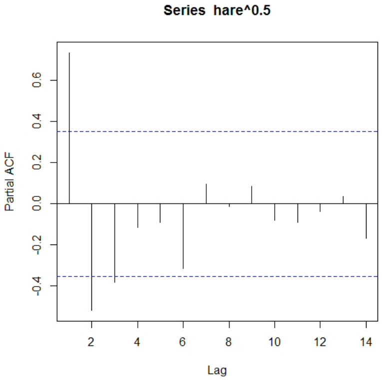
$\lambda = 0.45$. CI includes 0.5, use square root transform.




```
acf(hare^0.5)
```

Strong value at lag 1, but damped cosine wave warrants further investigation.

Sample pacf for square root of Hare abundance




```
pacf(hare^0.5)
```

Consider AR(2) or AR(3) model.

Test of stationarity of log(oil.price)

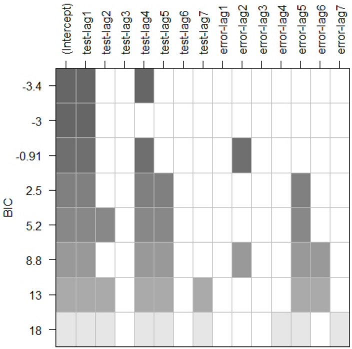
```
adf.test(log(oil.price))
Dickey-Fuller = -1.1119, p-value = 0.9189
alternative hypothesis: stationary

#p>0.05 thus implying do not reject H_0,
#i.e., data non-stationary,

#thus need for differencing.


adf.test(diff(log(oil.price)))
data:  diff(log(oil.price))
Dickey-Fuller = -6.6505, p-value = 0.01
alternative hypothesis: stationary
#p<0.05 thus reject H_0.
```

Eacf of diff(log(Oil price)) series

```
eacf(diff(log(oil.price)))
AR/MA
  0 1 2 3 4 5 6 7 8 9 10 11 12 13
0 x o o o o o o o o o o  o  o  o
1 x x o o o o o o o o x  o  o  o
2 o x o o o o o o o o o  o  o  o
3 o x o o o o o o o o o  o  o  o
4 o x x o o o o o o o o  o  o  o
5 o x o x o o o o o o o  o  o  o
6 o x o x o o o o o o o  o  o  o
7 x x o x o o o o o o o  o  o  o
ARMA(p=0,q=1)
```

An IMA(1,1) model on the log of the oil price. See page 31.

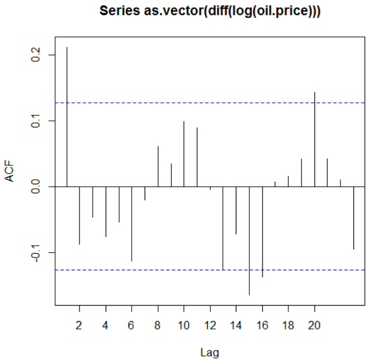
Best subset ARMA model for diff(log(Oil)) series




```
res=armasubsets(y=diff(log(oil.price)),
nar=7,nma=7,y.name='test',
ar.method='ols');plot(res)
```

$\nabla \log(oil)$ should be modelled in Y_{t-1} and Y_{t-4}.

Sample acf of diff(log(oil))

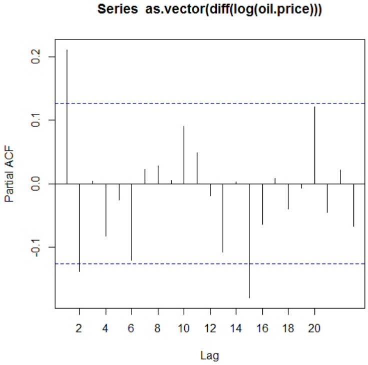



```
acf(as.vector(diff(log(oil.price))),
xaxp=c(0,20,10))
```

Consider MA(1) for diff(log(oil.price)), i.e.,

IMA(1,1) on log oil price.

Sample pacf of diff(log(oil))




```
pacf(as.vector(diff(log(oil.price))),
xaxp=c(0,20,10))
```

Consider AR(1) for diff(log(oil.price)), i.e.,

ARI(1,1) on log oil price.

Estimating parameters

- Method of moments
- Least squares
- Maximum likelihood