UNIVERSITY of LIMERICK

OLLSCOIL LUIMNIGH

Time Series Analysis

MS 4218

joseph.lynch@ul.ie

Outline

Model for non-stationary time series

- Stationarity through differencing
- ARIMA models
  - IMA(1,1) model
  - IMA(2,2) model
  - ARI(1,1) model
- Constant terms in ARIMA model
- Other transformations

Revisit AR(1) model

$$Y_t = \phi Y_{t-1} + e_t.$$

This is stationary if $|\phi| < 1$.

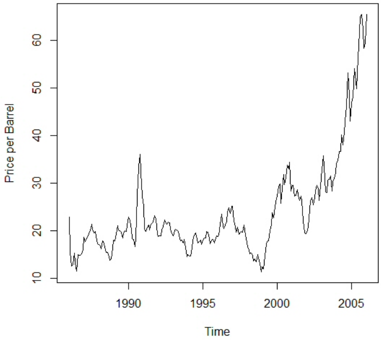
It is not stationary if $|\phi| \geq 1$.

The random walk model is a sum of white noise terms, and can be recursively defined by $Y_t = Y_{t-1} + e_t$.

It has $\phi = 1$ and is $\therefore$ not stationary.

Although it has zero mean, its variance is $t\sigma_e^2$ and thus increases with time.

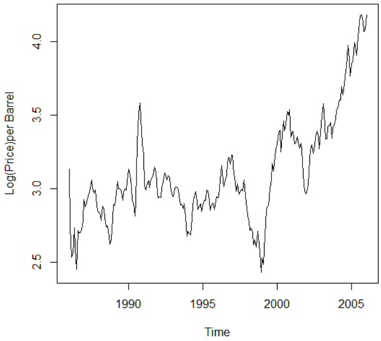
Plot of monthly oil prices: January 1986 to January2006




```
data(oil.price)
plot(oil.price,ylab="Price per barrel",type="l")
```

Unstable variance.

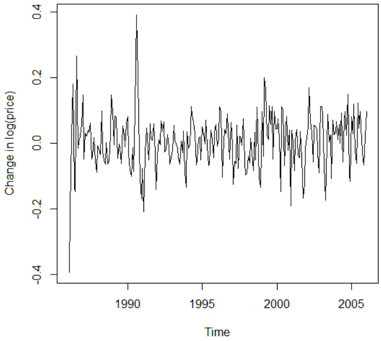
Use of logarithms to stabilise variance




```
data(oil.price); plot(log(oil.price),
ylab="Price per barrel",type="l")
```

Still not stationary.

Plot of differenced series of log(oil price)




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

Outlier still present though.

ARIMA: Auto-regressive Integrated Moving Average

If the d^{th} difference of a time series $\{Y_t\}$ is stationary, then

$$W_t = \nabla^d Y_t$$

is an ARIMA(p, d, q) process.

$d = 1$ or 2 is usually sufficient.

$$\begin{aligned} W_t &= \phi_1 W_{t-1} + \phi_2 W_{t-2} + \cdots + \phi_p W_{t-p} \\ &\quad + e_t - \theta_1 e_{t-1} - \cdots - \theta_q e_{t-q} \end{aligned}$$

$$\begin{aligned} Y_t - Y_{t-1} &= \phi_1(Y_{t-1} - Y_{t-2}) + \phi_2(Y_{t-2} - Y_{t-3}) + \ldots \\ &\quad + \phi_p(Y_{t-p} - Y_{t-p-1}) + e_t - \theta_1 e_{t-1} - \cdots - \theta_q e_{t-q}. \end{aligned}$$

$$\begin{aligned} Y_t &= (1 + \phi_1) Y_{t-1} + (\phi_2 - \phi_1) Y_{t-2} + \cdots + (\phi_p - \phi_{p-1}) Y_{t-p} \\ &\quad - \phi_p Y_{t-p-1} + e_t - \theta_1 e_{t-1} - \cdots - \theta_q e_{t-q}. \end{aligned}$$

Characteristic equation

This differenced form looks like it could be ARMA($p+1, q$) model.

The characteristic polynomial satisfies:

$$\begin{aligned} 1 &- (1+\phi_1)x - (\phi_2 - \phi_1)x^2 - \cdots - (\phi_p - \phi_{p-1})x^p + \phi_p x^{p+1} \\ &= (1 - \phi_1 x - \phi_2 x^2 - \cdots - \phi_p x^p)(1-x). \end{aligned}$$

As one root is $x = 1$, the process is not stationary and so it cannot be an ARMA($p+1, q$) process.

The remaining roots are the characteristic roots of the stationary ∇Y_t.

IMA(1,1)= ARIMA(0,1,1)

Differencing once confers stationarity as an MA(1) process.

$$
\begin{aligned}
\nabla Y_t &= Y_t - Y_{t-1} = e_t - \theta e_{t-1} \\
\Rightarrow Y_t &= Y_{t-1} + e_t - \theta e_{t-1} \\
&= (Y_{t-2} + e_{t-1} - \theta e_{t-2}) + e_t - \theta e_{t-1} \\
&= e_t + (1-\theta)e_{t-1} - \theta e_{t-2} + Y_{t-2} \\
&= e_t + (1-\theta)e_{t-1} - \theta e_{t-2} + (Y_{t-3} + e_{t-2} - \theta e_{t-3}) \\
&= e_t + (1-\theta)e_{t-1} + (1-\theta)e_{t-2} - \theta e_{t-3} + Y_{t-3}.
\end{aligned}
$$

IMA(1,1) cont.

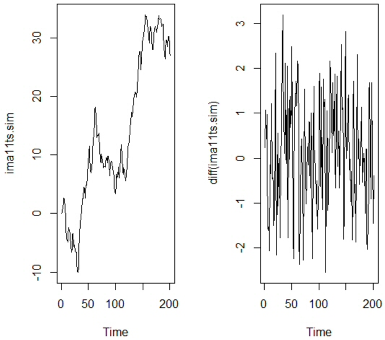
If we continue iterating, for $-m < 1$ and $t > 0$, we have

$$Y_t = e_t + (1-\theta)e_{t-1} + (1-\theta)e_{t-2} + \cdots + (1-\theta)e_{-m} - \theta e_{-m-1}.$$

The white noise does not die out as we go into the past.

Y_t is mostly an equally-weighted accumulation of a large number of white noise values.

IMA(1,1) simulation




```
set.seed(987); par(mfrow=c(1,2))
ima11ts.sim <- arima.sim(list(order = c(0,1,1),
ma = 0.7), n = 200)
plot(ima11ts.sim); plot(diff(ima11ts.sim))
```

Variance of IMA(1,1)

$$
\begin{aligned}
Var(Y_t) &= Var\{e_t + (1-\theta)e_{t-1} + (1-\theta)e_{t-2} + \\
&\quad \cdots + (1-\theta)e_{-m} - \theta e_{-m-1}\} \\
&= Var(e_t) + (1-\theta)^2 Var(e_{t-1}) + (1-\theta)^2 Var(e_{t-2}) + \\
&\quad \cdots + (1-\theta)^2 Var(e_{-m}) + \theta^2 Var(e_{-m-1}) \\
&= \sigma_e^2 + (1-\theta)^2\sigma_e^2 + (1-\theta)^2\sigma_e^2 + \\
&\quad \cdots + (1-\theta)^2\sigma_e^2 + \theta^2\sigma_e^2 \\
&= \sigma_e^2\{1 + (1-\theta)^2(t+m) + \theta^2\} \;=\; \gamma_0.
\end{aligned}
$$

As $t\uparrow$, $Var(Y_t)\uparrow$.

Auto-covariance for IMA(1,1)

$$
\begin{aligned}
Cov(Y_t, Y_{t-k}) &= Cov\{e_t + (1-\theta)e_{t-1} + (1-\theta)e_{t-2} + \dots \\
&\quad +(1-\theta)e_{-m} - \theta e_{-m-1}, \\
&\quad e_{t-k} + (1-\theta)e_{t-k-1} + (1-\theta)e_{t-k-2} + \dots \\
&\quad +(1-\theta)e_{-m-k} - \theta e_{-m-k-1}\} \\
&= \sigma_e^2\{1-\theta+\theta^2+(1-\theta)^2(t+m-k)\} \\
&= \gamma_k.
\end{aligned}
$$

Auto-correlation for IMA(1,1)

$$
\begin{aligned}
\rho_k &= \frac{\gamma_k}{\sqrt{Var(Y_t)Var(Y_{t-k})}} \\
&\approx \frac{t+m-k}{\sqrt{(t+m)(t+m-k)}} = \sqrt{\frac{t+m-k}{t+m}} \\
&\approx 1 \quad \text{for large } m \text{ and moderate } k \\
&> 0 \quad \text{for many lags.}
\end{aligned}
$$

```
cor(ima11ts.sim[-1],
ima11ts.sim[-length(ima11ts.sim)])
r_1=0.99554515
```

IMA(2,2)=ARIMA(0,2,2)

Differencing twice confers stationarity as an MA(2) process.

$$\begin{aligned}\nabla^2 Y_t &= \nabla(Y_t - Y_{t-1}) = (Y_t - Y_{t-1}) - (Y_{t-1} - Y_{t-2}) \\ Y_t - 2Y_{t-1} + Y_{t-2} &= e_t - \theta_1 e_{t-1} - \theta_2 e_{t-2}.\end{aligned}$$

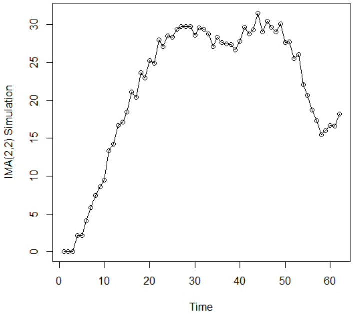
Like with IMA(1,1),

$Var(Y_t)$ increases rapidly with t and

$Corr(Y_t, Y_{t-k}) \to 1$ for all moderate k.

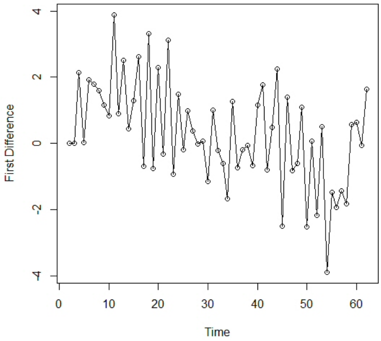
Simulation of IMA(2,2)= ima22.s series

$\theta_1 = 1$ and $\theta_2 = -0.6$. Smooth with $\uparrow$ variance as $t \uparrow$.




```
data(ima22.s); plot(ima22.s, ylab="IMA(2,2)
Simulation",type="o") cor(ima22.s[-1],
ima22.s[-length(ima22.s)]) 0.9866268
```

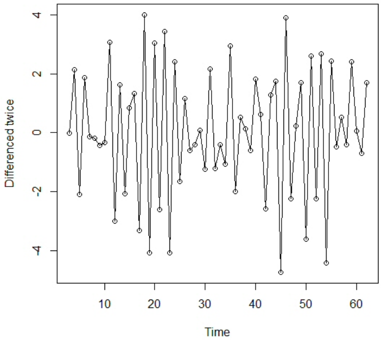
First difference of simulated IMA(2,2) series




```
plot(diff(ima22.s),
ylab="First Difference",type="o")
```

Non-stationary IMA(1,2).

Second difference of simulated IMA(2,2) series: Stationary




```
plot(diff(ima22.s, difference =2),
ylab="Differenced twice",type="o")
```

Stationary MA(2) with $\rho_1 = -0.678$ and $\rho_2 = 0.254$.

ARI(1,1) = ARIMA(1,1,0)

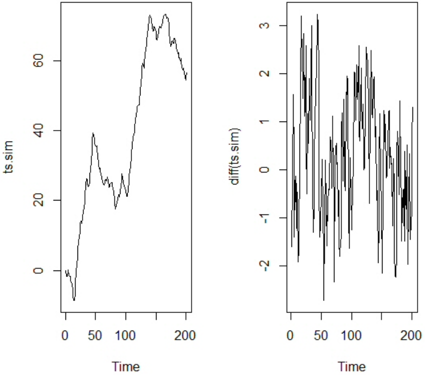
A stationary AR(1) model after differencing once.

$$Y_t - Y_{t-1} \quad = \quad \phi(Y_{t-1} - Y_{t-2}) + e_t.$$

We can simulate an ARI(1,1) series with arima.sim function.

```
ts.sim <- arima.sim(list(order = c(1,1,0),
ar = 0.7), n = 200)
par(mfrow=c(1,2));plot(ts.sim)
plot(diff(ts.sim))
```

Simulated ARI(1,1) and its differenced series



Differenced ARI(1,1) is a stationary AR(1) series.

Constant terms in ARIMA models, i.e., mean $\mu \neq 0$.

For an ARIMA (p, d, q),$\nabla^d Y_t = W_t$ is a stationary ARMA(p, q) process.

Non-zero mean can be accounted for 1. by subtraction:

$$\begin{aligned} W_t - \mu &= \phi_1(W_{t-1} - \mu) + \cdots + \phi_p(W_{t-p} - \mu) \\ &\quad + e_t - \theta_1 e_{t-1} - \cdots - \theta_q e_{t-q}, \end{aligned}$$

or 2. by inclusion as a constant:

$$\begin{aligned} W_t &= \theta_0 + \phi_1 W_{t-1} + \cdots + \phi_p W_{t-p} \\ &\quad + e_t - \theta_1 e_{t-1} - \cdots - \theta_q e_{t-q}. \end{aligned}$$

Expectation for non-zero mean, μ

$$
\begin{aligned}
E(W_t) &= E(\theta_0 + \phi_1 W_{t-1} + \cdots + \phi_p W_{t-p} \\
&\quad + e_t - \theta_1 e_{t-1} - \cdots - \theta_q e_{t-q}) \\
\mu &= \theta_0 + (\phi_1 + \cdots + \phi_p)\mu \\
\Rightarrow \theta_0 &= \mu(1 - \phi_1 - \cdots - \phi_p) \\
\text{and } \mu &= \frac{\theta_0}{1 - \phi_1 - \cdots - \phi_p}.
\end{aligned}
$$

IMA(1,1) with θ_0

$$
\begin{aligned}
W_t &= \theta_0 + e_t - \theta e_{t-1} \\
\equiv Y_t &= Y_{t-1} + \theta_0 + e_t - \theta e_{t-1} \\
&= Y_{t-2} + \theta_0 + e_{t-1} - \theta e_{t-2} + \theta_0 + e_t - \theta e_{t-1} \\
&= e_t + (1-\theta)e_{t-1} - \theta e_{t-2} + Y_{t-2} + 2\theta_0 \\
Y_t &= e_t + (1-\theta)e_{t-1} + (1-\theta)e_{t-2} + \dots \\
&\quad +(1-\theta)e_{-m} - \theta e_{-m-1} + (t+m+1)\theta_0.
\end{aligned}
$$

$\theta_0 \rightarrow$ Linear trend in IMA(1,1)

$$(t+m+1)\theta_0 = (m+1)\theta_0 + \theta_0 t$$

is an added linear deterministic time trend with slope of θ_0.

For IMA(2,1), i.e., a stationary MA(1) series after second differencing, in the presence of θ_0, a quadratic trend is implied.

Other transformations

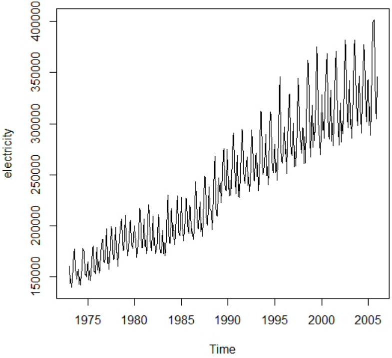
Differencing can lead to stationarity, as can percentage change (See Lab 3 Q6).

Logarithms can stabilise the variance when the standard deviation increases proportionately with time.

If the standard deviation increases exponentially with time, taking logarithms will convert the exponential series to a linear series.

Monthly U.S. electricity generation

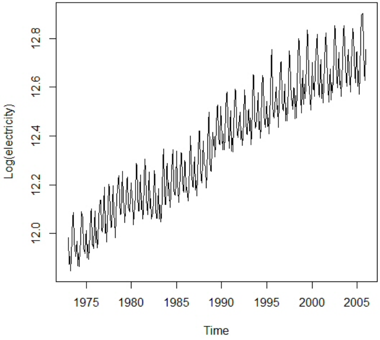
```
data(electricity);plot(electricity)
```



Larger variation with later times.

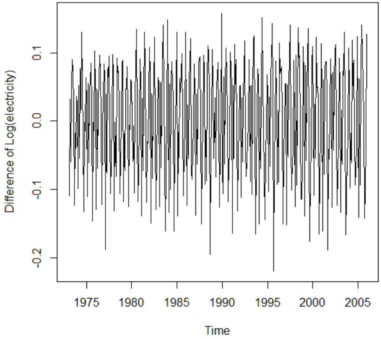
Log (Monthly U.S. electricity generation)

```
plot(log(electricity),ylab="Log(electricity)")
```



Trend still present, but variance more constant.

Differenced series of Log (Monthly U.S. electricity generation): Stationary




```
plot(diff(log(electricity)),
ylab="Difference of log(electricity)")
# log of data firstly, then diff
```

Power Transformations for positive data values only

Box-Cox transformation can be used to make data stationary.

For a given λ value, it is defined by:

$$g(x) = \begin{cases} \frac{x^{\lambda}-1}{\lambda} & \text{if } \lambda \neq 0. \\ \log(x) & \text{if } \lambda = 0. \end{cases}$$

When $\lambda = \frac{1}{2}$ the transformation is$\approx \sqrt{x}$.

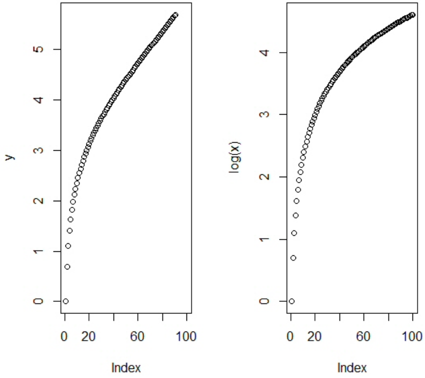
When λ is small

Using the following code, $\frac{x^{\lambda}-1}{\lambda}$ is shown overleaf to be a smooth function as $\lambda \rightarrow 0$.

```
x<-seq(1,100,by=1)
lambda<-seq(0.01,0.1,by=0.001)
y<-rep(NA,length(x))
for(i in 1:length(x))
y[i]<-(x[i]^(lambda[i])-1)/lambda[i]
par(mfrow=c(1,2))
plot(y); plot(log(x))
```

Plot of Y v $\log(x)$ when λ is small

$$Y = \frac{x^\lambda - 1}{\lambda}.$$



Log-likelihood theory

The best-fitting model to a given dataset of x_i values is the model with the highest likelihood function. For discrete data, this function is the product of the joint probabilities.

$$
\begin{aligned}
L(\theta) &= \prod_{i=1}^{n} Pr(X_i = x_i). \\
\log\{L(\theta)\} &= l(\theta) = \sum_{i=1}^{n} \log\{Pr(X_i = x_i)\}.
\end{aligned}
$$

$L(\theta)$ and $l(\theta)$ have a maximum at the same value of θ.

Parameter mles

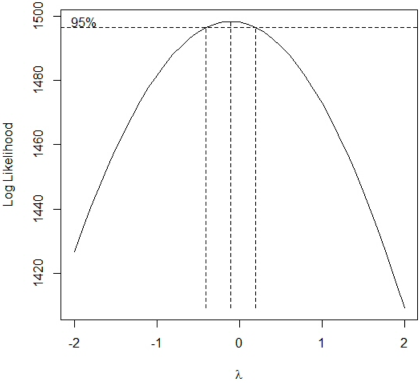
Differentiating $l(\theta)$ with respect to each of a given model's parameters yields the maximum likelihood estimates (mles) of each of the parameter values.

Thus for an AR(1) process, the mles of ϕ and σ_e^2 are required.

If, in addition, a Box-Cox transformation were required, the parameter λ has also to be estimated, but only approximately.

The BoxCox.ar function in R outputs the appropriate value of λ.

Plot of Loglikelihood v λ



R code

The R code for the preceding graph, and the mle confidence interval for λ is given by:

```
elect.transform<-BoxCox.ar(electricity,
lambda=seq(-2,2,0.05))
elect.transform
```

The mle is $\lambda = -0.1$.

The 95% confidence interval includes $\lambda = 0 \Rightarrow$ use $\log(x)$ to transform the data to stationarity.

Model Specification

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