UNIVERSITY of LIMERICK

OLLSCOIL LUIMNIGH

Time Series Analysis

MS 4218

joseph.lynch@ul.ie

Outline

Model Specification

Estimation of model parameters

- Method-of-Moments(MoM)
- Least squares
- Maximum likelihood

Arima model ?(p, d, q)

If model is an ARMA(p, q) stationary model, the values of $\phi_1, \ldots, \phi_p$ and $\theta_1, \ldots, \theta_q$ need to be found.

If model is non-stationary Arima(p, d, q), the d^{th} difference of the original time series is a stationary ARMA(p, q), and again, the $\phi_1, \ldots, \phi_p$ and $\theta_1, \ldots, \theta_q$ parameters need to be estimated.

If there is a non-zero mean, μ, then it also needs to be estimated.

Likewise $\sigma_e^2 = Var(e_t)$ needs to be estimated.

Method-of-Moments estimation of parameters

Uses sample moments to estimate corresponding theoretical moments, e.g.,

sample mean $\bar{x}$ to estimate the theoretical population mean μ.

In TSLecture3 Page 9, we saw that this estimator was heavily influenced by auto-correlation in the data.

$$Var(\bar{Y}) = \frac{\gamma_0}{n}\left\{1 + 2\sum_{k=1}^{n-1}\left(1 - \frac{|k|}{n}\right)\rho_k\right\}.$$

If $\rho_k < 0$, $Var(\bar{Y})$ decreases and if $\rho_k > 0$, $Var(\bar{Y})$ increases.

MoM estimation for AR models: See TSLecture 4b

In an AR(1) model, the theoretical auto-correlation $\rho_1 = \phi$.

The MoM estimate of ρ_1 is r_1, the sample auto-correlation coefficient, and r_1 is then used to estimate ϕ, i.e., $r_1 = \hat{\phi}$.

For an AR(2) process, ϕ_1 and ϕ_2 need to be estimated.

The Yule-Walker equations here are:

$$\rho_k = \phi_1\rho_{k-1} + \phi_2\rho_{k-2}.$$

$$\rho_1 = \phi_1\rho_0 + \phi_2\rho_{-1} = \phi_1 + \phi_2\rho_1.$$

$$\rho_2 = \phi_1\rho_1 + \phi_2\rho_0 = \phi_1\rho_1 + \phi_2.$$

MoM estimator of AR(2) models cont.

For MoM estimates of ϕ_1 and ϕ_2, replace ρ's with r's.

$$r_1 = \phi_1 + \phi_2 r_1.$$

$$r_2 = r_1\phi_1 + \phi_2.$$

Solving these equations for ϕ_1 and ϕ_2 yields

$$\hat{\phi}_1 = \frac{r_1(1-r_2)}{1-r_1^2}.$$

$$\hat{\phi}_2 = \frac{r_2 - r_1^2}{1-r_1^2}.$$

MoM estimator of AR(p) models cont.

Replace ρ's with r's in full Yule-Walker equations.

$$
\begin{aligned}
r_1 &= \phi_1 + \phi_2 r_1 + \phi_3 r_2 + \cdots + \phi_p r_{p-1}.\\
r_2 &= \phi_1 r_1 + \phi_2 + \phi_3 r_1 + \cdots + \phi_p r_{p-2}.\\
&\vdots\\
r_p &= \phi_1 r_{p-1} + \phi_2 r_{p-2} + \phi_3 r_{p-3} + \cdots + \phi_p.
\end{aligned}
$$

These linear equations are then solved for $\hat{\phi}_1, \ldots, \hat{\phi}_p$, which are called the Yule-Walker estimates.

MoM estimates for MA(1) models: See TSLecture4a

For an MA(1) model, the theoretical auto-correlation function is

$$\rho_1 = -\frac{\theta}{1+\theta^2}, \qquad (\rho_{max} = |0.5|).$$

$$r_1 = -\frac{\theta}{1+\theta^2}.$$

$$r_1\theta^2 + \theta + r_1 = 0.$$

$$\Rightarrow \hat{\theta} = -\frac{1}{2r_1} \pm \sqrt{\frac{1}{4r_1^2} - 1}.$$

$\exists$ a Real solution if $\frac{1}{4r_1^2} - 1 > 0$, i.e., $r < |0.5|$.

If $r > |0.5|$, $\nexists$ MoM estimate, so ? wrong model choice.

MoM estimates for MA(2) models: See TSLecture4a

For an MA(2) model, the theoretical auto-correlation function is

$$\rho_k = \begin{cases} 1 & \text{if } k = 0. \\ \frac{-\theta_1+\theta_1\theta_2}{1+\theta_1^2+\theta_2^2} & \text{if } k = 1. \\ \frac{-\theta_2}{1+\theta_1^2+\theta_2^2} & \text{if } k = 2. \\ 0 & \text{if } k > 2. \end{cases}$$

Replacing ρ_k with r_k and solving $\hat{\theta}$'s results in equations that are non-linear in θ's.

Thus solutions may possibly be numerical and only invertible solutions are retained.

MoM estimates for MA(q) models: See TSLecture4a

For an MA(q) model, the theoretical auto-correlation function is

$$\rho_k = \begin{cases} \frac{-\theta_k+\theta_1\theta_{k+1}+\cdots+\theta_{q-k}\theta_q}{1+\theta_1^2+\theta_2^2+\cdots+\theta_q^2} & \text{if } k = 1, 2, \ldots, q \\ 0 & \text{if } k > q. \end{cases}$$

Replacing ρ_k with r_k and solving $\hat{\theta}$'s results in equations that are highly non-linear in θ's and thus solutions, if any, are only numerical.

Only invertible solutions are retained.

MoM estimates for ARMA($p = 1, q = 1$) models: See TSLecture4c

For an ARMA$(1, 1)$ model, the theoretical auto-correlation function is

$$\rho_k = \frac{(\phi - \theta)(1 - \phi\theta)}{1 - 2\phi\theta + \theta^2}\phi^{k-1}, \text{ for } k \geq 1.$$

We have $\frac{\rho_2}{\rho_1} = \frac{\rho_3}{\rho_2} = \frac{\rho_k}{\rho_{k-1}} = \phi$, and so,

$$\hat{\phi} = \frac{r_2}{r_1}.$$

Then solve for $\hat{\theta}$ using

$$r_1 = \frac{(1 - \theta\hat{\phi})(\hat{\phi} - \theta)}{1 - 2\theta\hat{\phi} + \theta^2},$$

retaining only the invertible solution.

Estimates of the noise variance σ_e^2

After estimating ϕ's and $\theta' s$, only σ_e^2 still needs estimating.

The process variance γ_0 can be estimated by the sample variance s^2.

$$s^2 = \frac{\sum_{t=1}^{n}(Y_t - \bar{Y})^2}{n-1}.$$

Estimates of the noise variance σ_e^2 for AR models

Equations for γ_0 in terms of the AR(p)parameters in TSLecture4b Page 35.

$$
\begin{aligned}
\gamma_0 &= \frac{\sigma_e^2}{1 - \phi_1\rho_1 - \phi_2\rho_2 - \cdots - \phi_p\rho_p} \\
\Rightarrow \hat{\sigma}_e^2 &= s^2(1 - \phi_1 r_1 - \phi_2 r_2 - \cdots - \phi_p r_p).
\end{aligned}
$$

For an AR(1) process

$$
\begin{aligned}
\hat{\sigma}_e^2 &= s^2(1 - \phi_1 r_1) \\
&= s^2(1 - r_1^2),
\end{aligned}
$$

where $r_1 = \hat{\phi}_1$.

Estimates of the noise variance σ_e^2 for MA models

Equations for γ_0 for MA(q) parameters in TSLecture4a Page 30.

$$\begin{aligned}\gamma_0 &= (1+\theta_1^2+\theta_2^2+\cdots+\theta_q^2)\sigma_e^2 \\ \Rightarrow \hat{\sigma}_e^2 &= \frac{s^2}{1+\theta_1^2+\theta_2^2+\cdots+\theta_q^2}.\end{aligned}$$

Because MoM estimators for MA processes can be numerically difficult to evaluate, $\hat{\sigma}_e^2$ can likewise be difficult to estimate.

Estimates of the noise variance σ_e^2 for ARMA(1,1) model

In TSLecture4c Page 13 we derived an expression for γ_0 in terms of the parameters for this model.

$$\begin{aligned}\gamma_0 &= \frac{1 - 2\phi\theta + \theta^2}{1 - \phi^2}\sigma_e^2. \\ \Rightarrow \hat{\sigma}_e^2 &= \frac{s^2(1 - \hat{\phi}^2)}{1 - 2\hat{\phi}\hat{\theta} + \hat{\theta}^2}.\end{aligned}$$

As with purely MA processes, the computations can be numerically difficult to evaluate, and so $\hat{\sigma}_e^2$ can likewise be difficult to estimate.

Least Squares Estimation

Because with MA and ARMA models, MoM estimates may not be available, other methods are required.

For an AR(1) process with non-zero mean, we need to estimate ϕ, σ_e^2 and μ.

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

The residual for this model is given by:

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

The best fitting model has the minimum value of the total sum of squares of the residuals.

$$S_c(\phi, \mu) \quad = \quad \sum_{t=2}^{n} \{(Y_t - \mu) - \phi(Y_{t-1} - \mu)\}^2, \qquad \text{for } (Y_{t_1}, \ldots, Y_{t_n}).$$

This is the conditional sum-of-squares function.

LSE and AR(1) models

As S_c is a function of two parameters, ϕ and μ, we can get partial derivatives w.r.t. each parameter, set the results to 0 and solve the respective equation for ϕ and for μ.

$$
\begin{aligned}
\frac{\partial S_c}{\partial \mu} &= \sum_{t=2}^{n} 2\{(Y_t - \mu) - \phi(Y_{t-1} - \mu)\}(-1 + \phi) \\
&= 0, \\
\Rightarrow \hat{\mu} &= \frac{1}{(n-1)(1-\phi)} \left(\sum_{t=2}^{n} Y_t - \phi \sum_{t=2}^{n} Y_{t-1} \right).
\end{aligned}
$$

LSE and AR(1) models

If $n \uparrow\uparrow$,

$$\frac{1}{n-1}\sum_{t=2}^{n} Y_t \approx \frac{1}{n-1}\sum_{t=2}^{n} Y_{t-1}$$

$$\approx \bar{Y}.$$

$$\begin{aligned} \hat{\mu} &\approx \frac{1}{1-\phi}(\bar{Y} - \phi\bar{Y}) \\ &= \frac{\bar{Y}(1-\phi)}{1-\phi} \\ &= \bar{Y}. \end{aligned}$$

LSE and AR(1) models cont.

$$
\begin{aligned}
\frac{\partial S_c}{\partial \phi} &= \sum_{t=2}^{n} 2\{(Y_t - \bar{Y}) - \phi(Y_{t-1} - \bar{Y})\}(Y_{t-1} - \bar{Y}) = 0 \\
\Rightarrow \hat{\phi} &= \frac{\sum_{t=2}^{n}(Y_t - \bar{Y})(Y_{t-1} - \bar{Y})}{\sum_{t=2}^{n}(Y_{t-1} - \bar{Y})^2}.
\end{aligned}
$$

Apart from one missing denominator term, this is the same as r_1.

For large n, this missing term can be ignored, and thus LSE and MoM estimators are almost identical.

LSE and AR(p) models cont.

Using the same techniques, but for $\phi_1, \ldots, \phi_p$ and μ parameters, we get $\hat{\phi}_1, \ldots, \hat{\phi}_p$ and $\hat{\mu}$ results for LSE which are almost identical to their MoM counterparts, i.e., we get the sample Yule-Walker equations:

$$r_1 = \phi_1 + r_1\phi_2 \text{ and } r_2 = r_1\phi_1 + \phi_2.$$

$$\hat{\phi}_1 = \frac{r_1(1 - r_2)}{1 - r_1^2}.$$

$$\hat{\phi}_2 = \frac{r_2 - r_1^2}{1 - r_1^2}.$$

The estimate of the mean, $\hat{\mu}$ is likewise $\bar{Y}$.

LSE and MA models

An MA(1) process can be expressed as an AR(∞) process, - see TSLecture4c Page 3 and LabSolutions5 Q3.

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

LSE is obtained from minimising the sum of squared residuals, i.e.,

$$
\begin{aligned}
S_c(\theta) &= \sum (e_t)^2 \\
&= \sum (Y_t + \theta Y_{t-1} + \theta^2 Y_{t-2} + \theta^3 Y_{t-3} + \dots)^2
\end{aligned}
$$

LSE and MA models cont.

The S_c equation is non-linear in the parameters and requires numerical optimisation.

Conditional on $e_0 = 0$, $S_c(\theta)$ can be evaluated for a single value of θ.

$$\begin{aligned} e_1 &= Y_1. \\ e_2 &= Y_2 + \theta e_1 \\ e_n &= Y_n + \theta e_{n-1}, \end{aligned}$$

and so, $\sum(e_t)^2$ can be evaluated for that single θ value.

An MA(1) process is invertible for $-1 < \theta < 1$, and the minimisation process can be computed over a selection of θ values in this range to find the optimum LSE.

LSE and ARMA(1,1) and (p, q) model

The ARMA(1,1) model is defined by:

$$\begin{aligned} Y_t &= \phi Y_{t-1} + e_t - \theta e_{t-1}, \\ \Rightarrow e_t &= Y_t - \phi Y_{t-1} + \theta e_{t-1} \end{aligned}$$

The LSE is found by minimising the sum of residual squares

$$S_c(\phi, \theta) = \sum_{t=2}^{n} e_t^2,$$

For ARMA(p, q) minimise $S_c(\phi_1, \ldots, \phi_p, \theta_1, \ldots \theta_q)$ numerically to get LSE of the parameters.

Maximum Likelihood and Unconditional least squares

The likelihood function is the joint probability distribution of obtaining the data actually observed, $Y_1, \ldots, Y_n$, as a function of the unknown parameters $L(\phi, \theta, \mu, \sigma_e^2)$ in the model.

The maximum likelihood estimates (mles) of the parameters maximise the likelihood, or equivalently the log likelihood $l(\phi, \theta, \mu, \sigma_e^2)$ which is mathematically more tractable.

For σ_e^2, the joint density is the product of n independent Normal densities of $e_1, \ldots, e_n$ terms.

In an AR(1) process, for given values of ϕ and μ, $l(\phi, \mu, \sigma_e^2)$ can be maximised analytically w.r.t. σ_e^2.

In theory, is most comprehensive and R does all the work!

Properties of estimates

The large-sample properties of mle and lse are identical, and for large n, the estimators are $\approx$ Normally distributed.

For an AR(1) model

$$Var(\hat{\phi}) = \frac{1-\phi^2}{n} \qquad (Var \downarrow \text{ as } \phi \rightarrow 1).$$

For an AR(2) model

$$Var(\hat{\phi}_1) \approx Var(\hat{\phi}_2) = \frac{1-\phi_2^2}{n}.$$

$$Corr(\hat{\phi}_1, \hat{\phi}_2) \approx -\frac{\phi_1}{1-\phi_2} = -\rho_1.$$

If AR(2) model fitted instead of AR(1), $Var(\hat{\phi}_1) \uparrow$.

Properties of MA estimates

For an MA(1) model

$$Var(\hat{\theta}) \;=\; \frac{1-\theta^2}{n} \qquad (Var \downarrow \text{ as } \theta \to 1).$$

For an MA(2) model

$$Var(\hat{\theta}_1) \;\approx\; Var(\hat{\theta}_2) \;=\; \frac{1-\theta_2^2}{n}.$$

$$Corr(\hat{\theta}_1, \hat{\theta}_2) \;\approx\; -\frac{\theta_1}{1-\theta_2}.$$

If MA(2) model fitted instead of MA(1), $Var(\hat{\theta}_1) \uparrow$.

Properties of ARMA estimates

For an ARMA(1,1) model

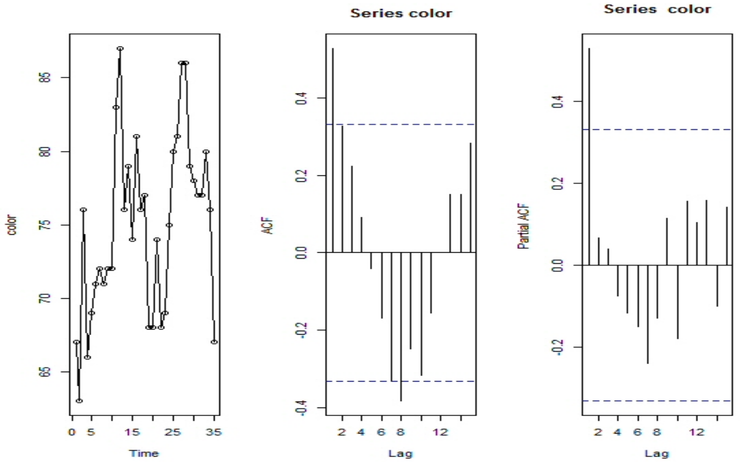
$$\begin{aligned} Var(\hat{\phi}) &= \frac{1-\phi^2}{n}\left(\frac{1-\phi\theta}{\phi-\theta}\right)^2. \\ Var(\hat{\theta}) &= \frac{1-\theta^2}{n}\left(\frac{1-\phi\theta}{\phi-\theta}\right)^2. \\ Corr(\hat{\phi},\hat{\theta}) &\approx \frac{\sqrt{(1-\phi^2)(1-\theta^2)}}{1-\phi\theta}. \end{aligned}$$

If ARMA(1,1) model fitted instead of AR(1) or MA(1), $Var(\hat{\phi})$ and $Var(\hat{\theta}) \uparrow$.

$Var(\hat{\phi})$ and $Var(\hat{\theta})$ also $\uparrow$ if ϕ and θ values are close to one another.

Colour property series Model Fitting

```
data(color) ; n=length(color)
plot(color,type="o");acf(color);pacf(color)
```



Colour property series Model Fitting

```
#Try AR(1) model
a<-acf(color)$acf[1]; a
arima(color,order=c(1,0,0),method="CSS")
arima(color,order=c(1,0,0),method="ML")
```

There is no method of finding MOM estimators with arima().

```
ar(color,order.max=1, AIC=F, method="yw")
ar(color,order.max=1, AIC=F, method="ols")
ar(color,order.max=1, AIC=F, method="mle")
```

$\hat{\phi}$ values for Colour property series

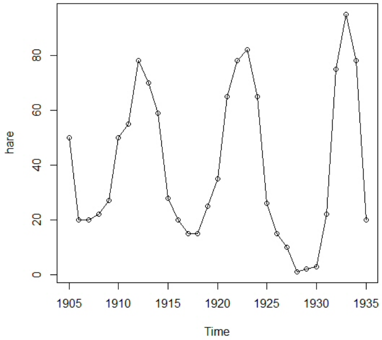
Parameter	Mom	CSS	MLE	n
ϕ	0.528	0.555	0.570	35

$$
\begin{aligned}
Se(\hat{\phi}) &= \sqrt{\frac{1-\phi^2}{n}} \\
&\approx \sqrt{\frac{1-0.57^2}{35}} \\
&\approx 0.14.
\end{aligned}
$$

Thus all the estimates are comparable.

Canadian Hare Abundance Model Fitting

```
data(hare);
plot(hare,type="o")
```



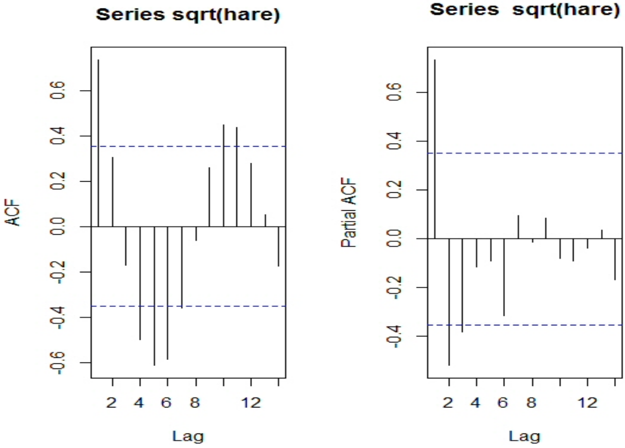
Canadian Hare Abundance Model Fitting

```
ar(diff(hare)) #8
adf.test(hare,k=8)
#Dickey-Fuller = -0.5785,
#Lag order = 8, p-value = 0.9698
#Ho: data not stationary
#High p value, do not reject Ho

haretransform<-BoxCox.ar(hare,
lambda=seq(0.3,0.7,0.01))
haretransform
#mle =0.46, so use square root approx.
adf.test(sqrt(hare))
# p-value < 0.01 i.e., reject Ho,
#Transformed data stationary
```

Canadian Hare Abundance Model Fitting

```
acf(sqrt(hare))
pacf(sqrt(hare))
```



Canadian Hare Abundance Model Fitting

```
# try AR(3 model)
arima(sqrt(hare),order=c(3,0,0))
```

method ="ML" is the default in the arima function. Equivalently

```
ar(sqrt(hare),order.max=3,method="mle")
```

Parameter output values for $\sqrt{Hare}$ series

```
Call:
arima(x = sqrt(hare), order = c(3, 0, 0))

Coefficients:
         ar1      ar2      ar3  intercept
      1.0519  -0.2292  -0.3931     5.6923
s.e.  0.1877   0.2942   0.1915     0.3371

sigma^2 estimated as 1.066:  log likelihood = -46.54,  aic = 101.08
```

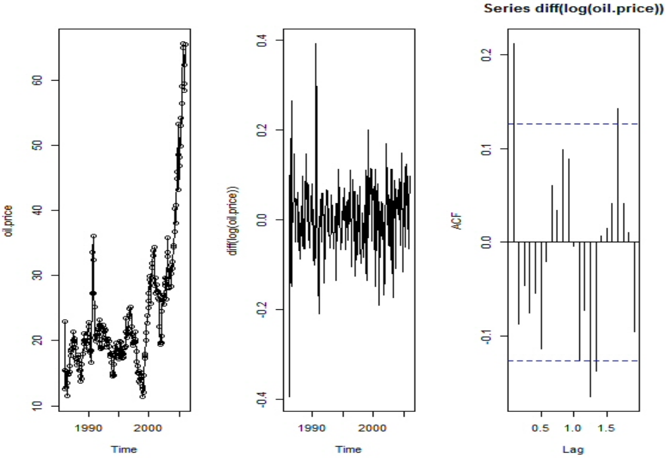
$\hat{\phi}_2$ with its high se is not significantly different from 0.

$$\begin{aligned}(\sqrt{Y_t} - 5.6923) &= 1.0519(\sqrt{Y_{t-1}} - 5.6923) - 0.2292(\sqrt{Y_{t-2}} - 5.6923) \\ &\quad -0.393(\sqrt{Y_{t-3}} - 5.6923) + e_t\end{aligned}$$

$$\sqrt{Y_t} = 3.25 + 1.0519\sqrt{Y_{t-1}} - 0.2292\sqrt{Y_{t-2}} - 0.393\sqrt{Y_{t-3}} + e_t.$$

Oil.Price Model Fitting

```
data(oil.price); plot(oil.price,type="o")
plot(diff(log(oil.price)));acf(diff(log(oil.price))
```



Oil.Price Model Fitting

```
# try MA(1) on difference of log(oil.price)
arima(log(oil.price),order=c(0,1,1),method='CSS')
arima(log(oil.price),order=c(0,1,1),method='ML')
r1=acf(diff(log(oil.price)))$acf[1]; r1# 0.2117
```

MoM estimate:

$$\begin{aligned}\hat{\theta} &= -\frac{1}{2r_1} \pm \sqrt{\frac{1}{4r_1^2} - 1} \\ &= -0.225 \text{ and } -4.502.\end{aligned}$$

Discard the latter because $|\theta| > 1$.

Parameter output values for Oil Price series

Parameter	MoM	CSS	MLE	n
θ	-0.2225	-0.2731	-0.2956	241

The MoM estimate looks quite different from the LS and ML estimates but is actually only one standard error of $\hat{\theta}$ from them, i.e., $se(\hat{\theta}) = \sqrt{\frac{1-\theta^2}{n}} \approx 0.06$, because n is quite large.

Model diagnostics

- Residual analysis
- Residual plots
- Normality of residuals
- Auto-correlation of residuals
- Ljung-Box test

- Over-fitting and parameter redundancy