Stationary models

MA, AR and ARMA

Matthieu Stigler

November 14, 2008

Version 1.1

Lectures list

Outline

Recall: auto-covariance

Definition (autocovariance)

$\text{Cov}(X_t, X_{t-k}) \equiv \gamma_k(t) \equiv \text{E}\left[(X_t - \mu)(X_{t-k} - \mu)\right]$

Definition (Autocorrelation)

$\text{Corr}(X_t, X_{t-k}) \equiv \rho_k(t) \equiv \frac{\text{Cov}(X_t, X_{t-k})}{\text{Var}(X_t)}$

Proposition

$\text{Corr}(X_t, X_{t-0}) = \text{Var}(X_t)$

$\text{Corr}(X_t, X_{t-j}) = \phi^j$ depend on the lage: plot its values at each lag.

Recall: stationarity

The stationarity is an essential property to define a time series process:

Definition

A process is said to be **covariance-stationary**, or **weakly stationary**, if its first and second moments are time invariant.

$$\begin{aligned} &E(Y_t) = E[Y_{t-1}] = \mu && \forall\, t \\ &\mathrm{Var}(Y_t) = \gamma_0 < \infty && \forall\, t \\ &\mathrm{Cov}(Y_t, Y_{t-k}) = \gamma_k && \forall\, t, \forall\, k \end{aligned}$$

Recall: The AR(1)

The AR(1): $Y_t = c + \varphi Y_{t-1} + \varepsilon_t \qquad \varepsilon_t \sim iid(0, \sigma^2)$
with $|\varphi| < 1$, it can be can be written as:

$$Y_t = \frac{c}{1-\varphi} + \sum_{i=0}^{t-1} \varphi^i \varepsilon_{t-i}$$

Its 'moments' do not depend on the time: :

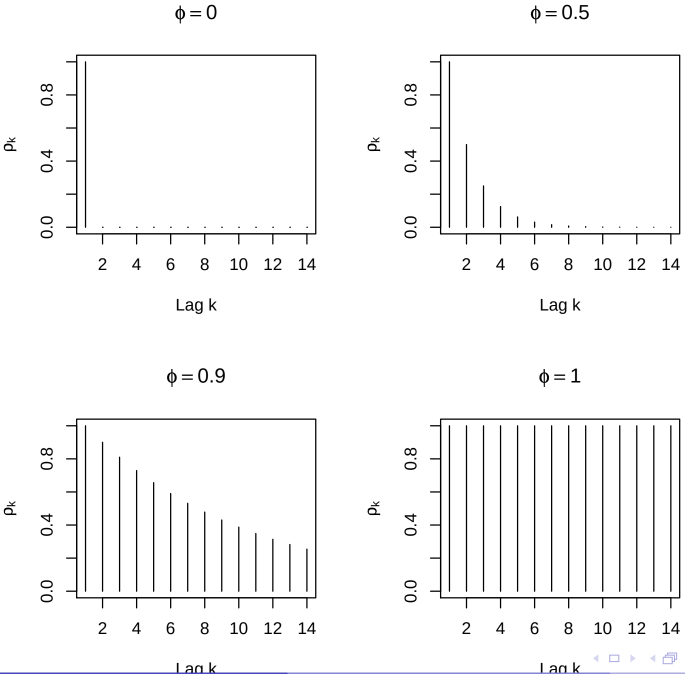
- $\text{E}(X_t) = \frac{c}{1-\varphi}$
- $\text{Var}(X_t) = \frac{\sigma^2}{1-\varphi^2}$
- $\text{Cov}(X_t, X_{t-j}) = \frac{\varphi^j}{1-\varphi^2}\sigma^2$
- $\text{Corr}(X_t, X_{t-j}) = \phi^j$

Outline

Outline

Autocorrelation function

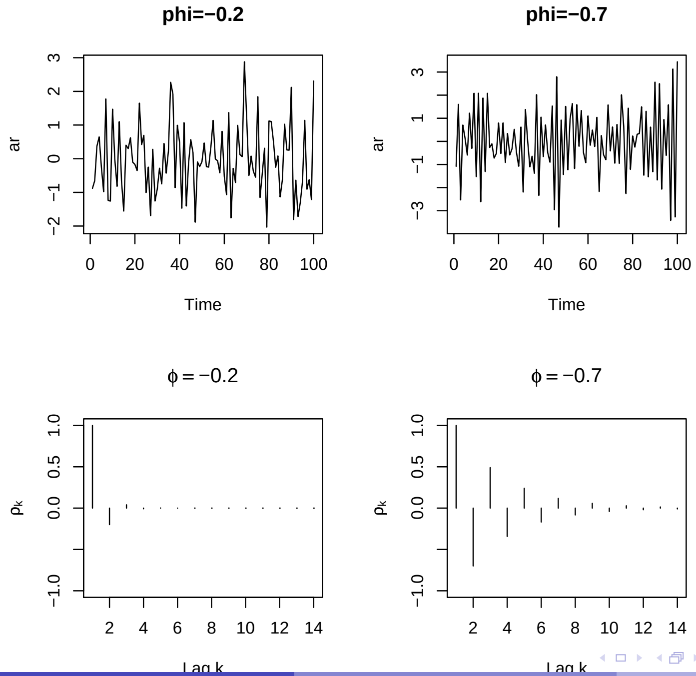
A usefull plot to understand the dynamic of a process is the autocorrelation function:
Plot the autocorrelation value for different lags.



φ = 0
φ = 0.5
φ = 0.9
φ = 1
ρk
0.0
0.4
0.8
2
4
6
8
10
12
14
Lag k

AR(1) with $-1 < \phi < 0$

in the AR(1): $Y_t = c + \varphi Y_{t-1} + \varepsilon_t \qquad \varepsilon_t \sim iid(0, \sigma^2)$
with $-1 < \phi < 0$
we have negative autocorrelation.


phi=−0.2
phi=−0.7
ar
Time
ϕ = −0.2
ϕ = −0.7
ρk
Lag k

Definition (AR(p))

$y_t = c + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \ldots + \phi_p y_{t-p} + \varepsilon_t$

- Expectation?
- Variance?
- Auto-covariance?
- Stationary conditions?

Lag operator

Definition (Backshift /Lag operator)

$LX_t = X_{t-1}$

Proposition

See that: $L^2 X_t = X_{t-2}$

Proposition (Generalisation)

$L^k X_t = X_{t-k}$

Lag polynomial

We can thus rewrite:

Example (AR(2))

$X_t = c + \varphi_1 X_{t-1} + \varphi_2 X_{t-2} + \varepsilon_t$

$(1 - \varphi_1 L - \varphi_2 L^2) X_t = c + \varepsilon_t$

Definition (lag polynomial)

We call lag polynomial: $\Phi(L) = (1 - \varphi_1 L - \varphi_2 L^2 - \ldots - \phi_p L^p)$

So we write compactly:

Example (AR(2))

$\Phi(L) X_t = c + \varepsilon_t$

Outline

1. Last Lecture
2. AR(p) models
   - Autocorrelation of AR(1)
   - Stationarity Conditions
   - Estimation
3. MA models
   - ARMA(p,q)
   - The Box-Jenkins approach
4. Forecasting

Definition (Characteristic polynomial)

$(1 - \varphi_1 z - \varphi_2 z^2 - \ldots - \phi_p z^p)$

Stability condition:

Proposition

The AR(p) process is stable if the roots of the lag polynomial lie outside the unit circle.

Example (AR(1))

The AR(1): $X_t = \varphi X_{t-1} + \varepsilon_t$
can be written as: $(1 - \varphi L)X_t = \varepsilon_t$
Solving it gives: $1 - \varphi x = 0 \Rightarrow x = \frac{1}{\varphi}$
And finally: $|\frac{1}{\varphi}| > 1 \Rightarrow |\varphi| < 1$

Proof.

1. Write an AR(p) as AR(1)
2. Show conditions for the augmented AR(1)
3. Transpose the result to the AR(p)

□

Proof.

The AR(p):

$$y_t = \phi_1 y_{t-1} + \phi_2 y_{t-2} + \ldots + \phi_p y_{t-p} + \varepsilon_t$$

can be recast as the AR(1) model:
$\xi_t = F\xi_{t-1} + \varepsilon_t$

$$\begin{bmatrix} y_t \\ y_{t-1} \\ y_{t-2} \\ \vdots \\ y_{t-p+1} \end{bmatrix} = \begin{bmatrix} \phi_1 & \phi_2 & \phi_3 & \ldots & \phi_{p-1} & \phi_p \\ 1 & 0 & 0 & \ldots & 0 & 0 \\ 0 & 1 & 0 & \ldots & 0 & 0 \\ \vdots & \vdots & \vdots & \ldots & \vdots & \vdots \\ 0 & 0 & 0 & \ldots & 1 & 0 \end{bmatrix} \begin{bmatrix} y_{t-1} \\ y_{t-2} \\ y_{t-3} \\ \vdots \\ y_{t-p} \end{bmatrix} + \begin{bmatrix} \varepsilon_t \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

$$\begin{cases} y_t & = c + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \ldots + \phi_p y_{t-p} + \varepsilon_t \\ y_{t-1} & = y_{t-1} \\ \ldots & \\ y_{t-p+1} & = y_{t-p+1} \end{cases}$$

Proof.

Starting from the augmented AR(1) notation:

$$\xi_t = F\xi_{t-1} + \varepsilon_t$$

Similarly as in the simple case, we can write the AR model recursively:

$$\xi_t = F^t\xi_0 + \varepsilon_t + F\varepsilon_{t-1} + F^2\varepsilon_{t-2} + \ldots + F^{t-1}\varepsilon_1 + F^t\varepsilon_0$$

Remember the eigenvalue decomposition: $F = T\Lambda T^{-1}$
and the propriety that: $F^j = T\Lambda^j T^{-1}$
with

$$\Lambda^j = \begin{bmatrix} \lambda_1^j & 0 & \ldots & 0 \\ 0 & \lambda_2^j & \ldots & 0 \\ \vdots & \vdots & \ldots & \vdots \\ 0 & 0 & \ldots & \lambda_3^j \end{bmatrix}$$

So the AR(1) model is stable if $|\lambda_i| < 1 \quad \forall i$ □

Proof.

So the condition on F is that all λ from $|F - \lambda I| = 0$ are < 1.
One can show that the eigenvalues of F are:

Proposition

$\lambda^p - \phi_1 \lambda^{p-1} - \phi_2 \lambda^{p-2} - \ldots - \phi_{p-1} \lambda - \phi_p = 0$

But the λ are the reciprocal of the values z that solve the characteristic polynomial of the AR(p):
$(1 - \varphi_1 z - \varphi_2 z^2 - \ldots - \phi_p z^p) = 0$
So the roots of the polynomial should be > 1, or, with complex values, outside the unit circle. □

Stationarity conditions

The conditions of roots outside the unit circle lead to:

- AR(1): $|\phi| < 1$
- AR(2):
  - $\phi_1 + \phi_2 < 1$
  - $\phi_1 - \phi_2 < 1$
  - $|\phi_2| < 1$

Example

Consider the AR(2) model:

$$Y_t = 0.8Y_{t-1} + 0.09Y_{t-2} + \varepsilon_t$$

Its AR(1) representation is:

$$\begin{bmatrix} y_t \\ y_{t-1} \end{bmatrix} = \begin{bmatrix} 0.8 & 0.09 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} y_{t-1} \\ y_{t-2} \end{bmatrix} + \begin{bmatrix} \varepsilon_t \\ 0 \end{bmatrix}$$

Hence its eigenvalues are taken from:

$$\begin{vmatrix} 0.8-\lambda & 0.09 \\ 1 & 0-\lambda \end{vmatrix} = \lambda^2 - 0.8\lambda - 0.09 = 0$$

And the eigenvalues are smaller than one:

```
> Re(polyroot(c(-0.09, -0.8, 1)))

[1] -0.1  0.9
```

Example

$$Y_t = 0.8Y_{t-1} + 0.09Y_{t-2} + \varepsilon_t$$

Its lag polynomial representation is: $(1 - 0.8L - 0.09L^2)X_t = \varepsilon_t$
Its characteristic polynomial is hence: $(1 - 0.8x - 0.09x^2) = 0$
whose solutions lie outside the unit circle:

```
> Re(polyroot(c(1, -0.8, -0.09)))
```

```
[1]   1.111111 -10.000000
```

And it is the inverse of the previous solutions:

```
> all.equal(sort(1/Re(polyroot(c(1, -0.8, -0.09)))), Re(polyroot(c(-0.09,
+     -0.8, 1))))
```

```
[1] TRUE
```

Unit root and integration order

Definition

A process is said to be integrated of order d if it becomes stationary after being differenced d times.

Proposition

An AR(p) process with k unit roots (or eigenvalues) is integrated of order k.

Example

Take the random walk: $X_t = X_{t-1} + \varepsilon_t$
Its polynomial is (1-L), and the roots is $1 - x = 0 \Rightarrow x = 1$
The eigenvalue of the trivial AR(1) is $1 - \lambda = 0 \Rightarrow \lambda = 1$

So the random walk is integrated of order 1 (or difference stationary).

Integrated process

Take an AR(p):

$$y_t = \phi_1 y_{t-1} + \phi_2 y_{t-2} + \ldots + \phi_p y_{t-p} + \varepsilon_t$$

With the lag polynomial:

$$\Phi(L) X_t = \varepsilon_t$$

If one of its p (not necessarily distinct) eigenvalues is equal to 1, it can be rewritten:

$$(1 - L)\Phi'(L) X_t = \varepsilon_t$$

Equivalently:

$$\Phi'(L) \Delta X_t = \varepsilon_t$$

The AR(p) in detail

Moments of a **stationary** AR(p)

- $E(X_t) = \frac{c}{1-\varphi_1-\varphi_2-\ldots-\varphi_p}$
- $\mathrm{Var}(X_t) = \varphi_1\gamma_1 + \varphi_2\gamma_2 + \ldots + \varphi_p\gamma_p + \sigma^2$
- $\mathrm{Cov}(X_t, X_{t-j}) = \varphi_1\gamma_{j-1} + \varphi_2\gamma_{j-2} + \ldots + \varphi_p\gamma_{j-p}$

Note that $\gamma_j \equiv \mathrm{Cov}(X_t, X_{t-j})$ so we can rewrite both last equations as:

$$\begin{cases} \gamma_0 & = \varphi_1\gamma_1 + \varphi_2\gamma_2 + \ldots + \varphi_p\gamma_p + \sigma^2 \\ \gamma_j & = \varphi_1\gamma_{j-1} + \varphi_2\gamma_{j-2} + \ldots + \varphi_p\gamma_{j-p} \end{cases}$$

They are known under the name of **Yule-Walker** equations.

Yule-Walker equations

Dividing by γ_0 gives:

$$\begin{cases} \rho_0 & = \varphi_1 \rho_1 + \varphi_2 \rho_2 + \ldots + \varphi_p \rho_p + \sigma^2 \\ \rho_j & = \varphi_1 \rho_{j-1} + \varphi_2 \rho_{j-2} + \ldots + \varphi_p \rho_{j-p} \end{cases}$$

Example (AR(1))

We saw that:

- $\text{Var}(X_t) = \frac{\sigma^2}{1-\varphi^2}$
- $\text{Cov}(X_t, X_{t-j}) = \frac{\varphi^j}{1-\varphi^2}\sigma^2$
- $\text{Corr}(X_t, X_{t-j}) = \phi^j$

And we have effectively: $\rho_1 = \phi \rho_0 = \phi$ and $\rho_2 = \phi \rho_1 = \phi^2$

Utility:

- Determination of autocorrelation function
- Estimation

Outline

To estimate a AR(p) model from a sample of T we take t=T-p

- M**ethods of moments**: estimate sample moments (the γ_i), and find parameters (the ϕ) correspondingly
- **Unconditional ML:** assume $y_p, \dots, y_1 \sim \mathcal{N}(0, \sigma^2)$. Need numerical optimisation methods.
- **Conditional Maximum likelihood** (=OLS): estimate $f(y_T, t_{T-1}, \dots, y_{p+1} | y_p, \dots, y_1; \theta)$ and assume $\varepsilon_t \sim \mathcal{N}(0, \sigma^2)$ and that $y_p, \dots, y_1$ are given

What if errors are not normally distributed? *Quasi-maximum likelihood estimator*, is still consistent (in this case) but standard erros need to be corrected.

Outline

Moving average models

Two significations!

- regression model
- Smoothing technique!

MA(1)

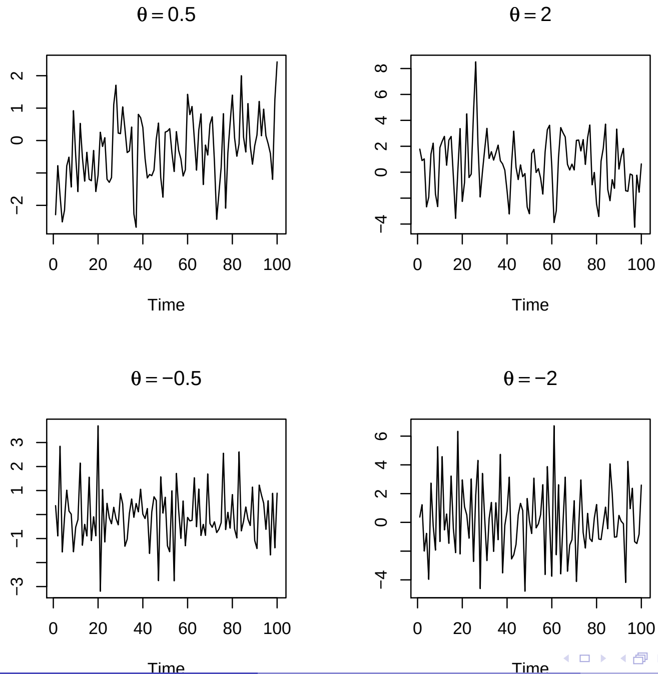
Definition (MA(1))

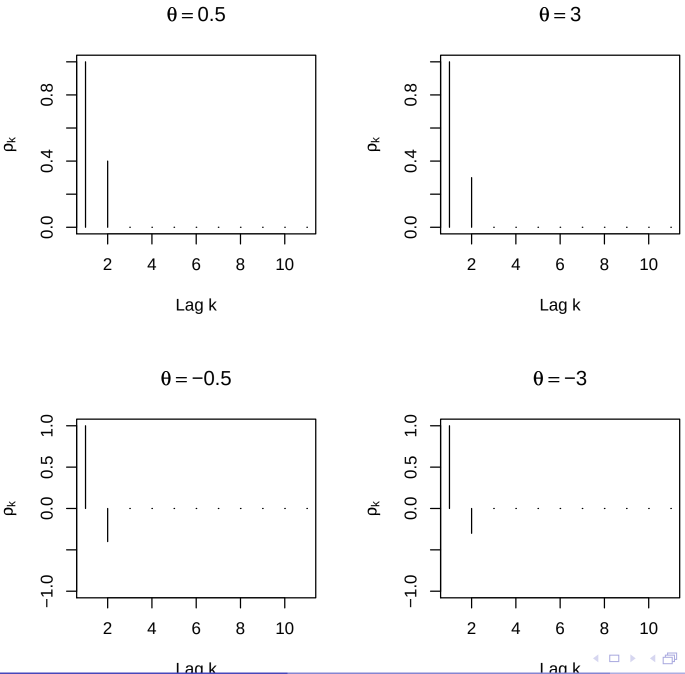
$Y_t = c + \varepsilon_t + \theta\epsilon_{t-1}$

- $\mathrm{E}(Y_t) = c$
- $\mathrm{Var}(Y_t) = (1+\theta^2)\sigma^2$
- $\mathrm{Cov}(X_t, X_{t-j}) = \begin{cases} \theta\sigma^2 & \text{if } j = 1 \\ 0 & \text{if } j > 1 \end{cases}$
- $\mathrm{Corr}(X_t, X_{t-j}) = \begin{cases} \frac{\theta}{(1+\theta^2)} & \text{if } j = 1 \\ 0 & \text{if } j > 1 \end{cases}$

Proposition

A MA(1) is stationnary **for every** θ


θ = 0.5
θ = 2
θ = −0.5
θ = −2
Time
Time
Time
Time


θ = 0.5
θ = 3
θ = −0.5
θ = −3
ρk
Lag k
0.0
0.4
0.8
0.5
1.0
−1.0
2
4
6
8
10

MA(q)

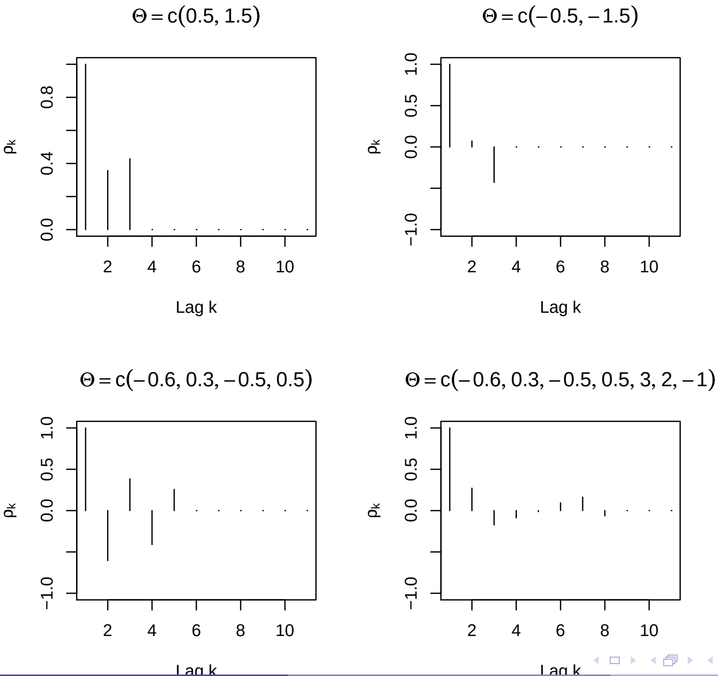
The MA(q) is given by:

$$Y_t = c + \varepsilon_t + \theta_1 \epsilon_{t-1} + \theta_2 \epsilon_{t-2} + \ldots + \theta_1 \epsilon_{t-q}$$

- $\mathrm{E}(Y_t) = c$
- $\mathrm{Var}(Y_t) = (1 + \theta_1^2 + \theta_2^2 + \ldots + \theta_q^2)\sigma^2$
- $\mathrm{Cov}(X_t, X_{t-j}) =$
  $\begin{cases} \sigma^2(\theta_j + \theta_{j+1}\theta_1 + \theta_{j+2}\theta_2 + \ldots + \theta_q\theta_{q-1}) & \text{if } j = 1 \\ 0 & \text{if } j > 1 \end{cases}$
- $\mathrm{Corr}(X_t, X_{t-j}) = \begin{cases} \frac{\theta}{(1+\theta^2)} & \text{if } j = 1 \\ 0 & \text{if } j > 1 \end{cases}$

Proposition

A MA(q) is stationary for every sequence $\{\theta_1, \theta_2, \ldots, \theta_q\}$


Θ = c(0.5, 1.5)
Θ = c(−0.5, −1.5)
Θ = c(−0.6, 0.3, −0.5, 0.5)
Θ = c(−0.6, 0.3, −0.5, 0.5, 3, 2, −1)
ρk
Lag k
0.0
0.4
0.8
−1.0
0.5
1.0
2
4
6
8
10

The MA(∞)

Take now the MA(∞):

$$Y_t = \varepsilon_t + \theta_1 \epsilon_{t-1} + \theta_2 \epsilon_{t-2} + \ldots + \theta_\infty \epsilon_\infty = \sum_{j=0}^{\infty} \theta_j \varepsilon_{t-j}$$

Definition (Absolute summability)

A sequence is *absolute summable* if $\sum_{i=0}^{\infty} |\alpha_i| < 0$

Proposition

The MA(∞) is stationary if the coefficients are absolute summable.

Back to AR(p)

Recall:

Proposition

If the characteristic polynomial of a AR(p) has roots =1, it is not stationary.

See that:
$(1 - \phi_1 y_{t-1} - \phi_2 y_{t-2} - \ldots - \phi_p y_{t-p}) y_t =$
$(1 - \alpha_1 L)(1 - \alpha_2 L) \ldots (1 - \alpha_p L) y_t = \varepsilon_t$
It has a MA(∞) representation if: $\alpha_1 \neq 1$:
$y_t = \frac{1}{(1-\alpha_1 L)(1-\alpha_2 L)\ldots(1-\alpha_p L)} \varepsilon_t$
Furthermore, if the α_i (the eigenvalues of the augmented AR(1)) are smaller than 1, we can write it:

$$y_t = \sum_{i=0}^{\infty} \beta_i \varepsilon_t$$

Estimation of a MA(1)

We do not observe neither ε_t nor ε_{t-1}
But if we know ε_0, we know $\varepsilon_1 = Y_t - \theta\varepsilon_0$
So obtain them recursively and minimize the conditional SSR:
$S(\theta) = \sum_{t=1}^{T}(y_t - \varepsilon_{t-1})^2$
This recquires numerical optimization and works only if $|\theta| < 1$.

Outline

ARMA models

The ARMA model is a composite of AR and MA:

Definition (ARMA(p,q))

$X_t = c + \phi_1 X_{t-1} + \phi_2 X_{t-2} + \ldots + \phi_p X_{t-p} + \varepsilon_{t-1} + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \ldots + \theta \varepsilon_{t-q}$

It can be rewritten properly as:

$$\Phi(L) Y_t = c + \Theta(L) \varepsilon_t$$

Theorem

The ARMA(p,q) model is stationary provided the roots of the $\Phi(L)$ polynomial lie outside the unit circle.

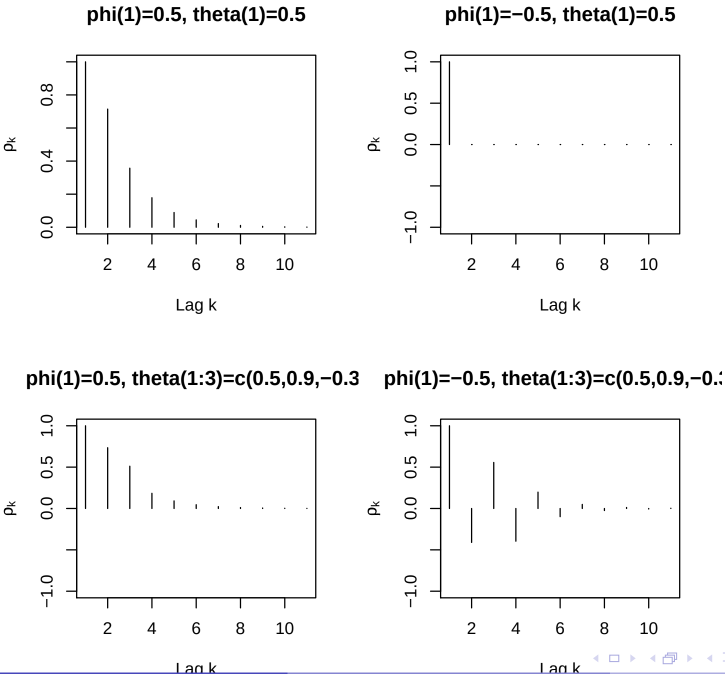
So only the AR part is involved!

Autocorrelation function of a ARMA(p,q)

Proposition

After q lags, the autocorrelation function follows the pattern of the AR component.

Remember: this is then given by the Yule-Walker equations.


phi(1)=0.5, theta(1)=0.5
phi(1)=−0.5, theta(1)=0.5
phi(1)=0.5, theta(1:3)=c(0.5,0.9,−0.3
phi(1)=−0.5, theta(1:3)=c(0.5,0.9,−0.
ρk
Lag k
0.0
0.4
0.8
−1.0
0.5
1.0
2
4
6
8
10

ARIMA(p,d,q)

Now we add a parameter d representing the order of integration (so the I in ARIMA)

Definition (ARIMA(p,d,q))

ARIMA(p,d,q): $\Phi(L)\Delta^d Y_t = \Theta(L)\varepsilon_t$

Example (Special cases)

- White noise: ARIMA(0,0,0) $X_t = \varepsilon_t$
- Random walk : ARIMA(0,1,0): $\Delta X_t = \varepsilon_t \Rightarrow X_t = X_{t-1} + \varepsilon_t$

Estimation and inference

The MLE estimator has to be found numerically.
Provided the errors are normaly distributed, the estimator has the usual asymptotical properties:

- Consistent
- Asymptotically efficients
- Normally distributed

If we take into account that the variance had to be estimated, one can rather use the T distribution in small samples.

Outline

The Box-Jenkins approach

1. Transform data to achieve stationarity
2. Identify the model, i.e. the parameters of ARMA(p,d,q)
3. Estimation
4. Diagnostic analysis: test residuals

Step 1

Transformations:

- Log
- Square root
- Differenciation
- Box-Cox transformation: $Y_t^{(\lambda)} \begin{cases} \frac{Y_t^\lambda - 1}{\lambda} & \text{for } \lambda \neq 0 \\ \log(Y - t) & \text{for } \lambda = 0 \end{cases}$

Is log legitimate?

- Process is: $y_t e^{\delta t}$ Then $z_t = \log(y_t) = \delta t$ and remove trend
- Process is $y_t = y_{t-1} + \varepsilon_t$ Then (by $\log(1 + x) \cong x$)
  $\Delta \log(y_t) = \frac{y_t - y_{t-1}}{y_t}$

Step 2

Identification of p,q (d should now be 0 after convenient transformation)
Principle of parsimony: prefer small models Recall that

- incoporating variables increases fit (R^2) but reduces the degrees of freedom and hence precision of estimation and tests.
- A AR(1) has a MA(∞) representation
- If the MA(q) and AR(p) polynomials have a common root, the ARMA(p,q) is similar to ARMA(p-1,q-1).
- Usual techniques recquire that the MA polynomial has roots outside the unit circle (i.e. is invertible)

Step 2: identification

How can we determine the parameters p,q?

- Look at ACF and PACF with confidence interval
- Use information criteria
  - Akaike Criterion (AIC)
  - Schwarz criterion (BIC)

Definition (IC)

$AIC(p) = n \log \hat{\sigma}^2 + 2p$
$BIC(p) = n \log \hat{\sigma}^2 + p \log n$

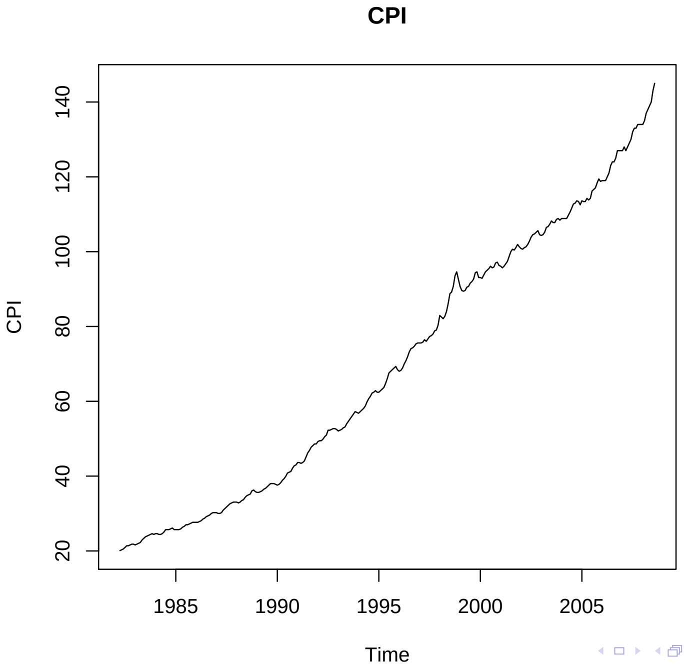
Step 3: estimation

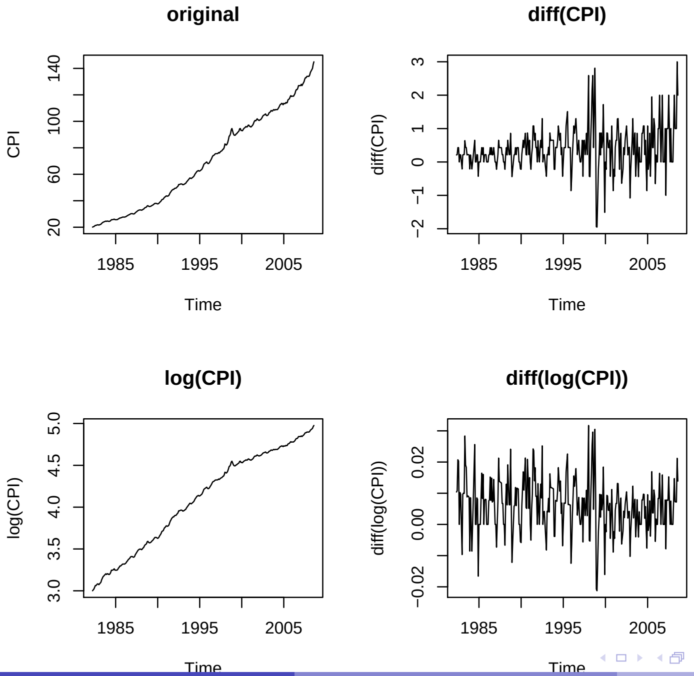
Estimate the model...
R function: arima() argument: order=c(p,d,q)

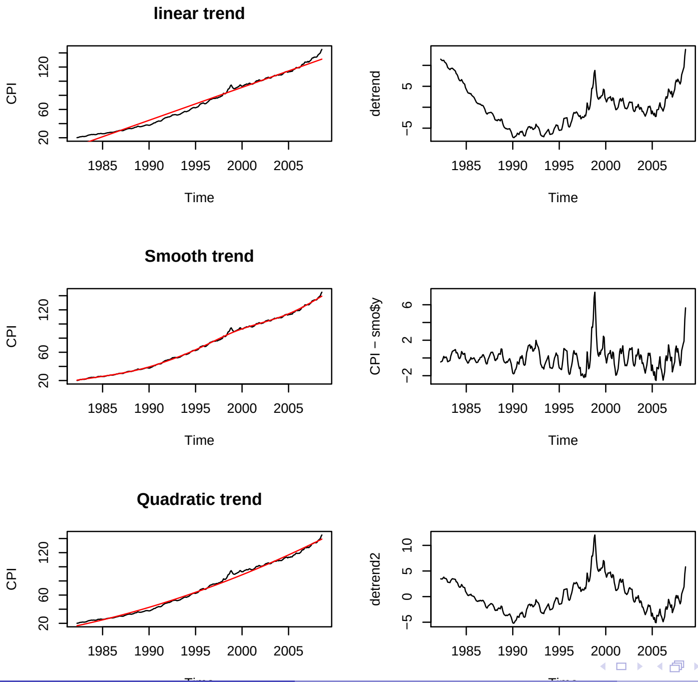
Step 4: diagnostic checks

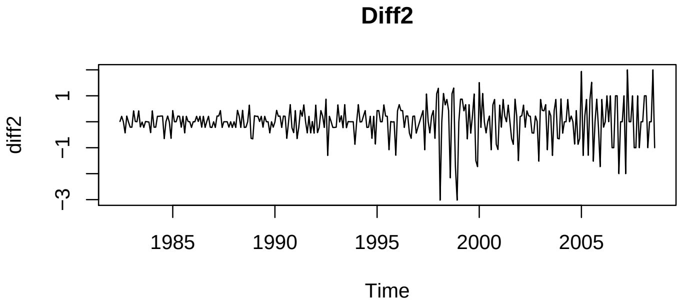
Test if the residuals are white noise:

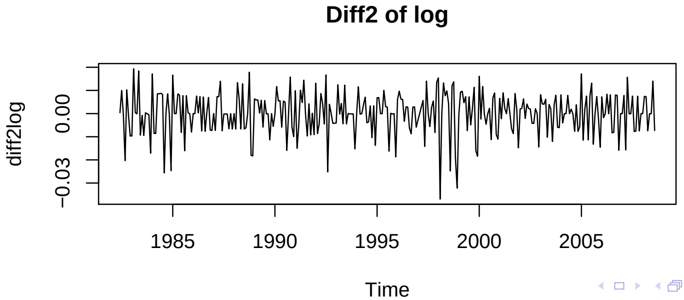
1. Autocorrelation
2. Heteroscedasticity
3. Normality

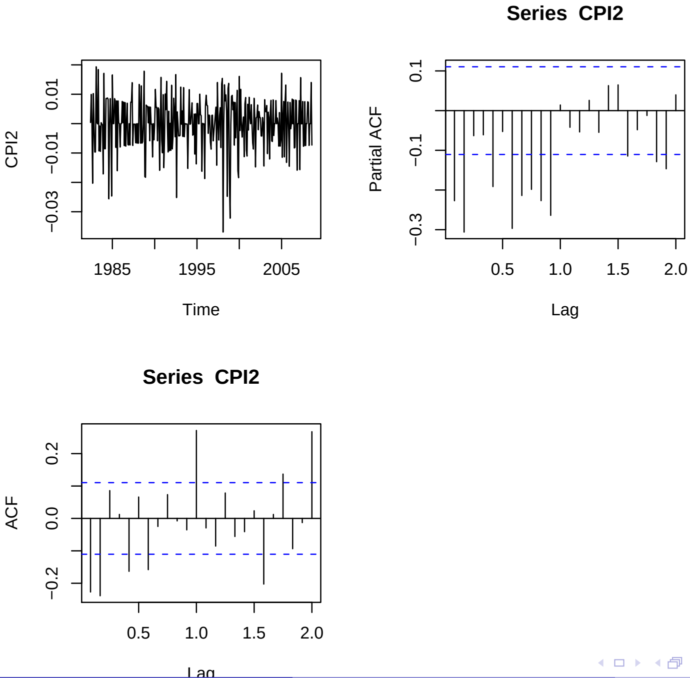

CPI
CPI
20
40
60
80
100
120
140
1985
1990
1995
2000
2005
Time


original
CPI
20
60
100
140
1985
1995
2005
Time
diff(CPI)
diff(CPI)
−2
−1
0
1
2
3
1985
1995
2005
Time
log(CPI)
log(CPI)
3.0
3.5
4.0
4.5
5.0
1985
1995
2005
Time
diff(log(CPI))
diff(log(CPI))
−0.02
0.00
0.02
1985
1995
2005
Time

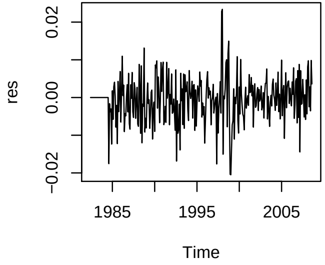

linear trend
CPI
20
60
120
1985
1990
1995
2000
2005
Time
detrend
−5
5
Smooth trend
CPI − smo$y
−2
2
6
Quadratic trend
detrend2
0
10

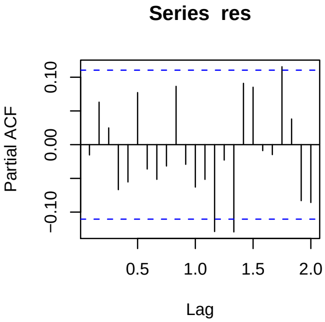

Diff2
diff2
1
−1
−3
1985
1990
1995
2000
2005
Time

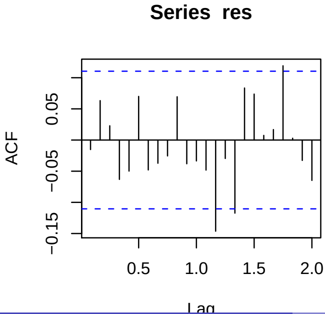

Diff2 of log
diff2log
0.00
−0.03
1985
1990
1995
2000
2005
Time


Series CPI2
CPI2
0.01
−0.01
−0.03
1985
1995
2005
Time
Partial ACF
0.1
−0.1
−0.3
0.5
1.0
1.5
2.0
Lag
Series CPI2
ACF
0.2
0.0
−0.2
0.5
1.0
1.5
2.0
Lag

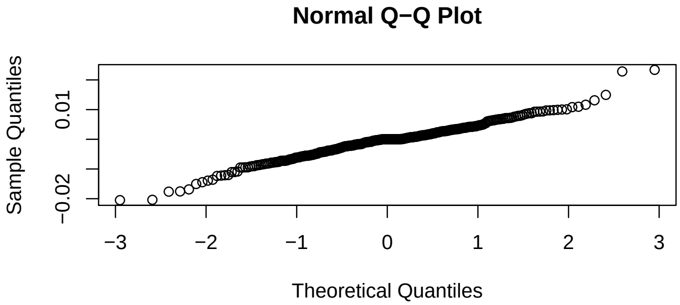
```
> library(forecast)

This is forecast 1.17

> fit <- auto.arima(CPI2, start.p = 1, start.q = 1)
> fit

Series: CPI2
ARIMA(2,0,1)(2,0,2)[12] with zero mean

Coefficients:
         ar1      ar2      ma1    sar1    sar2     sma1     sma2
      0.2953  -0.2658  -0.9011  0.6021  0.3516  -0.5400  -0.2850
s.e.  0.0630   0.0578   0.0304  0.1067  0.1051   0.1286   0.1212

sigma^2 estimated as 4.031e-05:  log likelihood = 1146.75
AIC = -2277.46   AICc = -2276.99   BIC = -2247.44

> res <- residuals(fit)
```

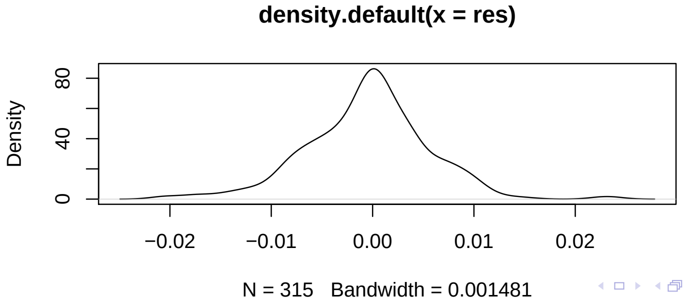

res
0.02
0.00
−0.02
1985
1995
2005
Time


Series res
Partial ACF
0.10
0.00
−0.10
0.5
1.0
1.5
2.0
Lag


Series res
ACF
0.05
−0.05
−0.15
0.5
1.0
1.5
2.0

```
> Box.test(res)

        Box-Pierce test

data:  res
X-squared = 0.0736, df = 1, p-value = 0.7862
```


Normal Q−Q Plot
Sample Quantiles
−0.02
0.01
−3
−2
−1
0
1
2
3
Theoretical Quantiles


density.default(x = res)
Density
0
40
80
−0.02
−0.01
0.00
0.01
0.02
N = 315   Bandwidth = 0.001481

Outline

Notation (Forecast)

$\hat{y}_{t+j} \equiv \mathrm{E}_t(y_{t+j}) = \mathrm{E}(y_{t+j}|y_t, y_{t-1}, \ldots, \varepsilon_t, \varepsilon_{t-1}, \ldots)$ *is the conditional expectation of* y_{t+j} *given the information available at* t.

Definition (J-step-ahead forecast error)

$e_t(j) \equiv y_{t+j} - \hat{y}_{t+j}$

Definition (Mean square prediction error)

$MSPE \equiv \frac{1}{H} \sum_{i=1}^{H} e_i^2$

R implementation

To run this file you will need:

- R Package forecast
- R Package TSA
- Data file AjaySeries2.csv put it in a folder called Datasets in the same level than your.Rnw file
- (Optional) File Sweave.sty which change output style: result is in blue, R commands are smaller. Also in same folder as .Rnw file.