Time Series 2

-Quantifying the World-

Lecturer: Darren Homrighausen, PhD

ARIMA

Key processes

Autoregressive Process: An AR(p) process is:

$$Y_t = \phi_1 Y_{t-1} + \ldots + \phi_p Y_{t-p} + \epsilon_t = \sum_{j=1}^{p} \phi_j Y_{t-j} + \epsilon_t$$

(This is a linear combination of the previous p values, plus a random noise term)

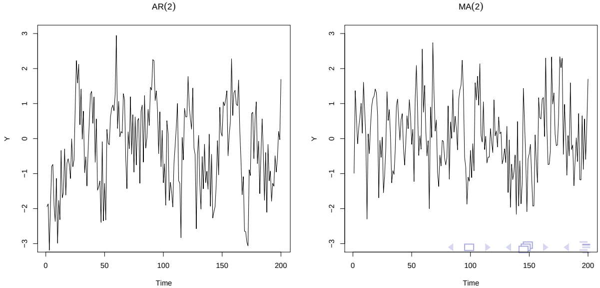
Moving Average Process: An MA(q) process is:

$$Y_t = \epsilon_t + \theta_1 \epsilon_{t-1} + \ldots + \theta_q \epsilon_{t-q} = \epsilon_t + \sum_{j=1}^{q} \theta_j \epsilon_{t-j}$$

(This is a linear combination of the previous q random terms (sometimes referred to as shocks, plus a random noise term)

Examples: R

```
require(astsa)
ylim  = c(-3,3)
parms = c(.35,.35)
plot(arima.sim(list(order=c(length(parms),0,0),
                    ar=parms),n=200),
     ylab="Y", main="AR(2)",ylim=ylim)
plot(arima.sim(list(order=c(0,0,length(parms)),
                    ma=parms),n=200),
     ylab="Y",main="MA(2)",ylim=ylim)
```



Key processes

ARMA process: A time series is an ARMA(p,q) process if it is stationary and a combination of AR and MA processes

It turns out the model is unidentified which makes it possible for a white noise process to look like a ARMA(k,k) process.

$$x_t = \epsilon_t \Leftrightarrow x_t - 0.5x_{t-1} = \epsilon_t - 0.5\epsilon_{t-1} \Leftrightarrow \mathrm{ARMA}(1,1)$$

(Thus, we usually demand a stricter form of the ARMA model known as "causal" and "invertible")

Reminder: Autocovariance Function (ACF)

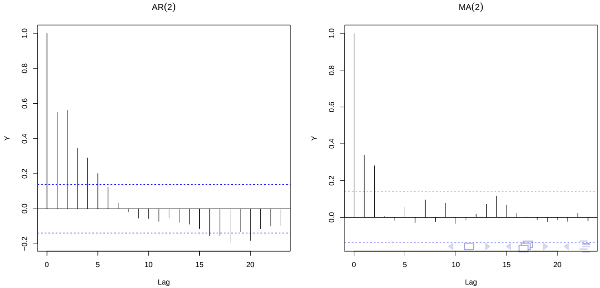
The ACF for a stationary process is:

$$\rho(h) = \frac{\gamma(h)}{\gamma(0)}$$

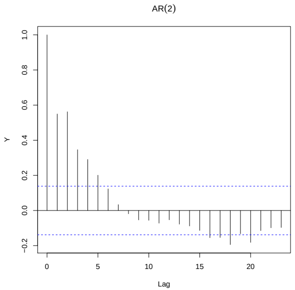
- AR(p): Will decay exponentially
- MA(q): Will have a sharp cut-off after $h > q$
- ARMA(p,q): Will decay exponentially

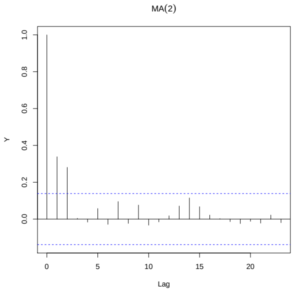
Examples: R

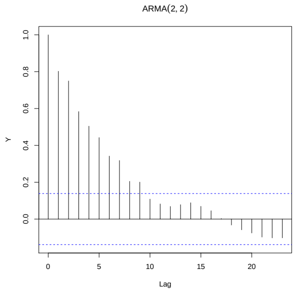
```
require(astsa)
set.seed(10)
acf(arima.sim(list(order=c(length(parms),0,0),
                   ar=parms),n=200),
    ylab="Y", main="AR(2)")
set.seed(1)
acf(arima.sim(list(order=c(0,0,length(parms)),
                   ma=parms),n=200),
    ylab="Y",main="MA(2)")
```



Examples: R







ACF for ARMA models

Take away:

- The ACF gives a lot of information about the order of a MA process
- The ACF behavior is similar for AR and ARMA models
- The ACF gives little information about the order of a AR/ARMA process

To correct for these last two points, we need a different type of covariance function

Partial Autocovariance (PACF)

For three random variables A, B, C, the partial covariance is

- Regressing A onto C to produce $\hat{A}$
- Regressing B onto C to produce $\hat{B}$
- Finding:

$$\text{partialcov}(A, B|C) = \text{cov}(A - \hat{A}, B - \hat{B})$$

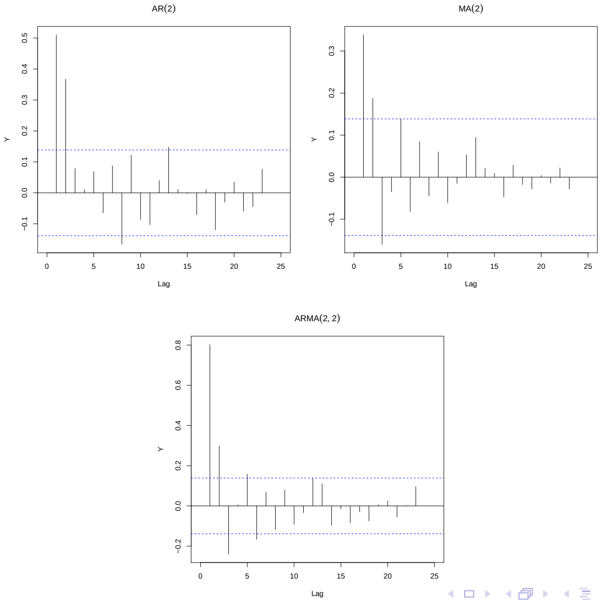
For an AR(p), Y_t is unrelated to Y_{t-p-1} given $\{Y_{t-1}, \ldots, Y_{t-p}\}$

$$Y_t = \phi_1 Y_{t-1} + \ldots + \phi_p Y_{t-p} + \epsilon_t = \sum_{j=1}^{p} \phi_j Y_{t-j} + \epsilon_t$$

$\rightarrow$ Use the partial covariance to determine the presence/order of AR processes

(The partial autocovariance is the covariance between Y_t and Y_s with the linear effect of everything "in the middle" removed)

PACF


AR(2)
MA(2)
ARMA(2, 2)
Y
Lag
0.5
0.4
0.3
0.2
0.1
0.0
−0.1
0.8
0.6
0.4
0.2
−0.2
0
5
10
15
20
25

Summary: Using ACF vs. PACF

	AR(p)	MA(q)	ARMA(p,q)
ACF	Tails off	Cuts off after q lags	Tails off
PACF	Cuts off after p lags	Tails off	Tails off

Differencing nonstationary time series

Integrated ARMA (ARIMA)

Example:

$$Y_t = Y_{t-1} + \epsilon_t$$

If we difference this series, we end up with a white noise process (and hence stationary)

$$Y_t - Y_{t-1} = \epsilon_t$$

It turns out, many time series are of this form:

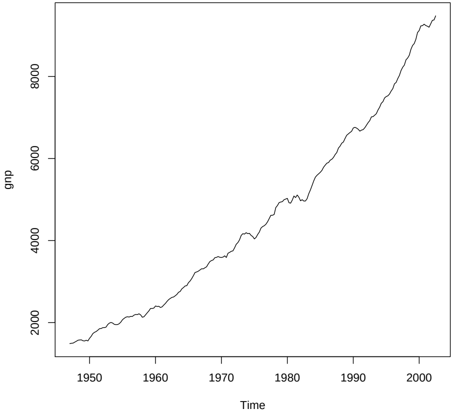
- A nonstationary (trend or seasonal) component
- A zero-mean stationary component

If differencing (Y_t) d-times produces an ARMA(p, q) process, then (Y_t) is an ARIMA(p, d, q)

An algorithm for ARIMA

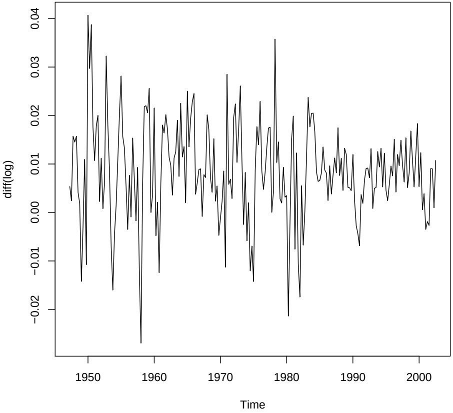
1. Plotting the data
2. Considering transformations
3. Identifying the dependence-based model parameters
   (That is, p, d, and q)
4. Parameter estimation
5. Diagnostics
   - Look at the ACF/QQ plot of the residuals

United States GNP



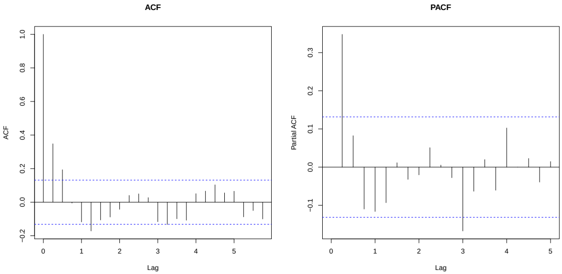
- The growth is probably exponential
- Definitely not stationary

United States GNP



- The growth is probably exponential → log transform
- Definitely not stationary → difference

United States GNP



- Could be an MA(2) process $\rightarrow$ we estimate that log(GNP) follows an ARIMA(0,1,2)
- Could be an AR(1) process $\rightarrow$ we estimate that log(GNP) follows an ARIMA(1,1,0)
- Could be a combination

We can fit all three models

United States GNP

To the R code `Unit4lecture.R`