Time Series
2. Stationary Processes. ARMA Models

Josep A. Sanchez-Espigares

Department of Statistics and Operations Research
Universitat Polit?cnica de Catalunya - BarcelonaTECH
Barcelona, Spain

josep.a.sanchez@upc.edu

Outline

- AR(p) models
- MA(q) models
- ARMA(p,q) models

Gaussian process $\{ X_t \}$, $t = 1, \cdots, n$

$$X = (X_t) \sim \text{Multivariate Normal/Gaussian distribution}$$

$$E(X) = \mu = (\mu_1, \cdots, \mu_n)'$$

Covariance matrix (nxn) is

$$V(X) = \Gamma = \{\gamma(t_i, t_j) | t, j = 1, \cdots, n\}$$

And the multivariate Normal density function can be written as

$$f(x) = (2\pi)^{-n/2} |\Gamma|^{-1/2} exp\{-1/2(x-\mu)^{'} \Gamma^{-1} (x-\mu)\}$$

where $|\Gamma| \equiv$ determinant

General Stochastic model for a time series:

$$X_t = G(X_{t-1}, X_{t-2}, ..., Z_{t-1}, Z_{t-2}, ...) + Z_t \quad Z_t \sim WN(0, \sigma_z)$$

The most easy mathematical function G(.) is the linear combination of the components:

Linear Stochastic model for a time series:

$$X_t = \phi_1 X_{t-1} + \phi_2 X_{t-2} + ... + \theta_1 Z_{t-1} + \theta_2 Z_{t-2} + ... + Z_t \quad Z_t \sim WN(0, \sigma_Z)$$

$$\phi_1, \phi_2, ..., \theta_1, \theta_2, ... \in \mathbf{R}$$

The variable at time t is a linear combination of past observations and disturbances plus a new disturbance independent form the past

AR(p) models

A **pth-order autoregressive mode**l, or AR(p), takes the form:

$$X_t = \phi_1 X_{t-1} + \phi_2 X_{t-2} + \cdots + \phi_p X_{t-p} + Z_t$$

$$(1 - \phi_1 B - \phi_2 B^2 - \cdots + \phi_p B^p) X_t = Z_t$$

{where}

X_t is stationary and $\phi_1, \cdots, \phi_p$ are the parameters (constants)

Z_t is a Gaussian white noise ($E(Z_t) = 0$, $V(Z_t) = \sigma_Z^2$)

B is the Backshift operator: $BX_t = X_{t-1}$

p is the lag of the farthest observation included

An **AR(p) model** is a regression model with lagged values of the dependent variable in the independent variable positions, hence the name **Auto-Regressive model**.

AR(p) models

- If $\mu \neq 0$, then

$$(X_t - \mu) = \phi_1(X_{t-1} - \mu) + \cdots + \phi_p(X_{t-p} - \mu) + Z_t$$

or

$$X_t = \alpha + \phi_1 X_{t-1} + \cdots + \phi_p X_{t-p} + z_t$$

where the constant term is:

$$\alpha = \mu(1 - \phi_1 - \cdots - \phi_p)$$

Or a more concise expression:

$$\phi_p(B) X_t = Z_t$$

where $\phi(B) = (1 - \phi_1 B - \cdots - \phi_p B^p)$ is the characteristic autoregressive polynomial of order p.

AR(p) models

AR(1) process, considering $\mu = 0$:

$$(1 - \phi B)X_t = Z_t, \quad Z_t \sim N(0, \sigma_Z^2).$$

Or equivalently:

$$X_t = \phi X_{t-1} + Z_t, \quad Z_t \sim N(0, \sigma_Z^2).$$

Autocovariance function derivation:

$$\gamma(0) = E[X_t^2] = E[(\phi X_{t-1} + Z_t)^2] = \phi^2 \gamma(0) + \sigma_Z^2 \Rightarrow \gamma(0) = \frac{\sigma_Z^2}{1 - \phi^2}$$

$$\gamma(1) = E[X_t X_{t-1}] = E[(\phi X_{t-1} + z_t)X_{t-1}] = \phi \gamma(0)$$

$$\vdots$$

$$\gamma(h) = E[X_t X_{t-h}] = E[(\phi X_{t-1} + Z_t)X_{t-h}] = \phi \gamma(h-1) = \phi^h \gamma(0)$$

Reminder: The noise at time t is independent of the past:
$E[Z_t X_s] = E[Z_t Z_s] = 0 \quad s < t$.

AR(1) process: $(1-\phi B)X_t = Z_t$

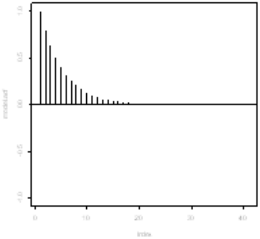
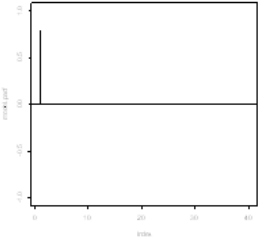
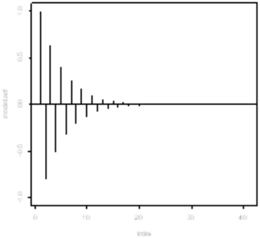
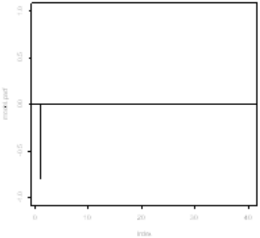
Autocorrelation function:

$$\rho(1) = \frac{\gamma(1)}{\gamma(0)} = \phi$$
$$\vdots$$
$$\rho(h) = \frac{\gamma(h)}{\gamma(0)} = \phi^h$$

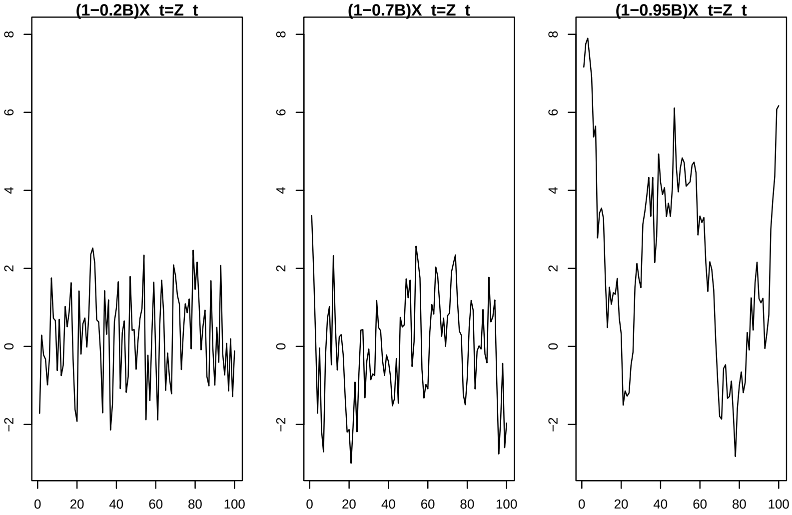
Recursion: $\rho(h) = \phi\rho(h-1)$
Partial Autocorrelation function:

$$\phi_{1,1} = \phi$$
$$\vdots$$
$$\phi_{h,h} = 0 \quad h > 1$$

AR(1) process: $(1 - \phi B)X_t = Z_t$

AR(1)	$\phi>0$		
	$\phi<0$		

AR(1) models



AR(p) models

AR(2) process: $(1 - \phi_1 B - \phi_2 B^2)X_t = Z_t$

Considering $\mu = 0$, then

$$X_t = \phi_1 X_{t-1} + \phi_2 X_{t-1} + Z_t \quad Z_t \sim N(0, \sigma_Z^2)$$

Autocovariance function:

$$\begin{gathered}
\gamma(0) = E[X_t^2] = (\phi_1^2 + \phi_2^2)\gamma(0) + 2\phi_1\phi_2\gamma(1) + \sigma_Z^2 \\
\gamma(1) = E[X_t X_{t-1}] = \phi_1\gamma(0) + \phi_2\gamma(1) \\
\vdots \\
\gamma(h) = E[X_t X_{t-h}] = \phi_1\gamma(h-1) + \phi_2\gamma(h-2) \quad h > 1
\end{gathered}$$

MA(q) models

A **qth-order moving average model**, or MA(q), takes the form:

$$X_t = z_t + \theta_1 Z_{t-1} + \theta_2 Z_{t-2} + \cdots + \theta_q Z_{t-q}$$

In other words,

$$X_t = \theta_q(B) Z_t$$

where

$$\theta(B) = 1 + \theta_1 B + \theta_2 B^2 + \cdots + \theta_q B^q$$

is the moving average operator or characteristic polynomial.

An **MA(q) model** is a regression model with the dependent variable, X_t, depending on previous values of the errors rather than on the variable itself.

MA(q) models

MA(1) process, considering $\mu = 0$:

$$X_t = (1 + \theta B)Z_t, \quad Z_t \sim N(0, \sigma_Z^2)$$

Or equivalently:

$$X_t = Z_t + \theta Z_{t-1}, \quad Z_t \sim N(0, \sigma_Z^2)$$

Autocovariance function derivation:

$$\gamma(0) = E[X_t^2] = E[(Z_t + \theta Z_{t-1})^2] = (1 + \theta^2)\sigma_Z^2$$

$$\gamma(1) = E[X_t X_{t-1}] = E[(Z_t + \theta Z_{t-1})(Z_{t-1} + \theta Z_{t-2})] = \theta\sigma_Z^2$$

$$\vdots$$

$$\gamma(h) = E[X_t X_{t-h}] = 0 \quad h > 1$$

MA(q) models

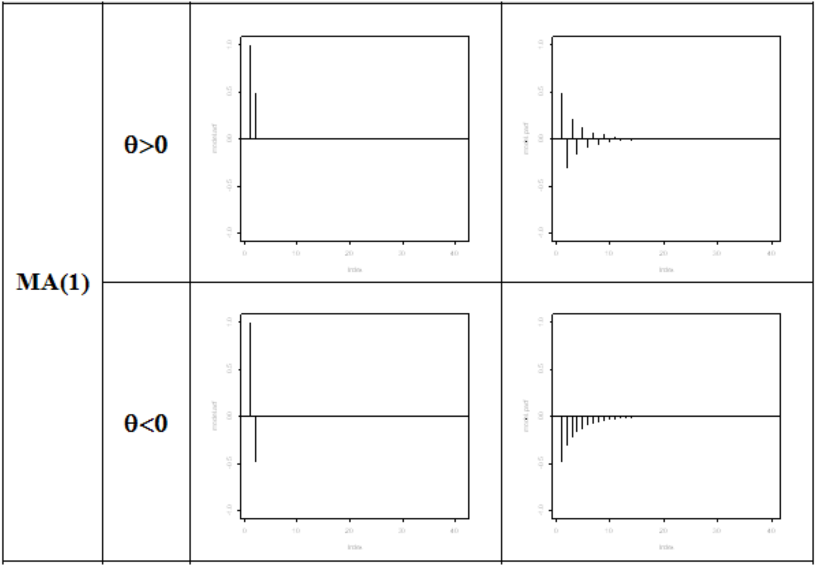
MA(1) process: $X_t = (1 + \theta B) Z_t$.

Autocorrelation function:

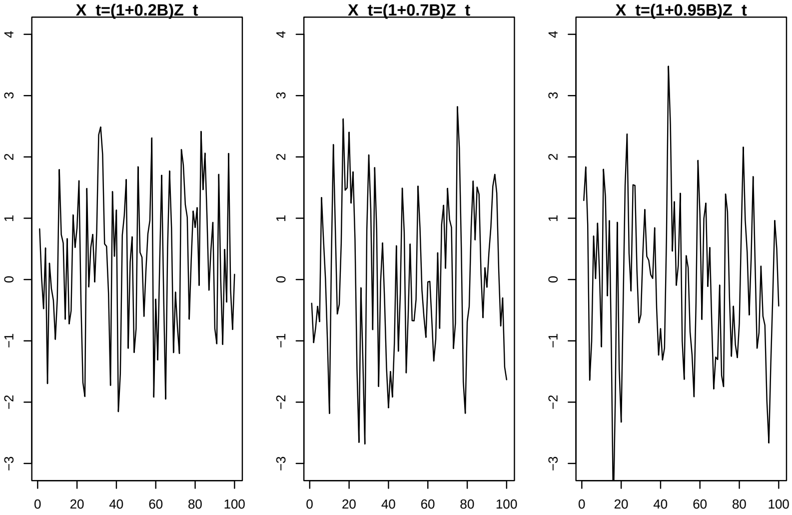
$$\rho(1) = \frac{\gamma(1)}{\gamma(0)} = \frac{\theta}{1+\theta^2}$$
$$\vdots$$
$$\rho(h) = \frac{\gamma(h)}{\gamma(0)} = 0 \quad h > 1$$

MA(q) models

MA(1): $x_t = (1 + \theta B) z_t$



MA(1) models



A times series $\{X_t; t = 0, \pm 1, \pm 2, \cdots\}$ is an **AutoRegressive Moving Average model**, ARMA (p,q), if it is stationary and

$$X_t = \phi_1 X_{t-1} + \cdots + \phi_p X_{t-p} + Z_t + \theta_1 Z_{t-1} + \cdots + \theta_q Z_{t-q}$$

with $\phi_p \neq 0$, $\theta_q \neq 0$ and $\sigma_z^2 > 0$

The parameters p and q are called the autoregressive and the moving average orders, respectively.

ARMA(p,q) models

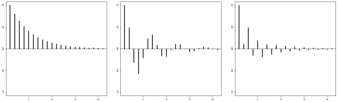
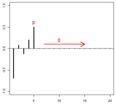
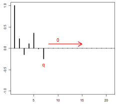
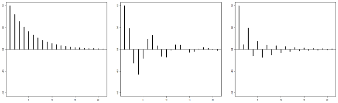
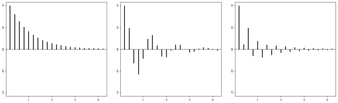
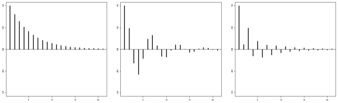
If X_t has a nonzero mean μ and $\alpha = \mu(1 - \phi_1 - \cdots - \phi_p)$, the model del will be:

$$X_t = \alpha + \phi_1 X_{t-1} + \cdots + \phi_p X_{t-p} + Z_t + \theta_1 Z_{t-1} + \cdots + \theta_q Z_{t-q}$$

When $q = 0$, the model is an AR(p).
When $p = 0$, the model is a MA(q)

ARMA(p,q) model can be written in concise form as,

$$\phi(B)_p X_t = \theta_q(B) Z_t$$

MODEL	ACF	PACF
AR(p)	Decreasing patterns (∞ lags not null) 	$p =$ Last lag not null 
MA(q)	$q =$ Last lag not null 	Decreasing patterns (∞ lags not null) 
ARMA(p,q)	Decreasing patterns (∞ lags not null) 	Decreasing patterns (∞ lags not null) 

Expression of an $ARMA(p, q)$ model, using characteristic polynomials:

$$\phi(B)x_t = \theta(B)z_t$$

Under certain conditions, ARMA(p, q) models can be expressed as an AR(∞) or MA(∞)model.

Expression as an AR(∞):

$$\frac{\phi(B)}{\theta(B)}x_t = \pi(B)x_t = z_t$$

Expression as an MA(∞):

$$x_t = \frac{\theta(B)}{\phi(B)}z_t = \psi(B)z_t$$

Expression of an $ARMA(p,q)$ model, using characteristic polynomials:

$$(1-\phi_1 B - ... - \phi_p B^p)x_t = (1+\theta_1 B + ... + \theta_q B^q)z_t$$

Under certain conditions, $ARMA(p,q)$ models can be expressed as a pure AR(∞) or MA(∞) process.

Expression as an AR(∞):

$$\frac{1-\phi_1 B - ... - \phi_p B^p}{1+\theta_1 B + ... + \theta_q B^q}x_t = (1-\pi_1 B - \pi_2 B^2 - ...)x_t = z_t$$

Expression as an MA(∞):

$$x_t = \frac{1+\theta_1 B + ... + \theta_q B^q}{1-\phi_1 B - ... - \phi_p B^p}z_t = (1+\psi_1 B + \psi_2 B^2 + ...)z_t$$

ARMA(p,q) models

Under certain conditions of stationarity, the ARMA(p, q) models can be expressed as an AR(∞) or MA(∞)model:

$$(1 - \phi_1 B - ... - \phi_p B^p)X_t = (1 + \theta_1 B + ... + \theta_q B^q)Z_t$$

Expression as an AR(∞):

$$\frac{1 - \phi_1 B - ... - \phi_p B^p}{1 + \theta_1 B + ... + \theta_q B^q} X_t = (1 - \pi_1 B - \pi_2 B^2 - ...)X_t = Z_t$$

Expression as an MA(∞):

$$X_t = \frac{1 + \theta_1 B + ... + \theta_q B^q}{1 - \phi_1 B - ... - \phi_p B^p} z_t = (1 + \psi_1 B + \psi_2 B^2 + ...)Z_t$$

$AR(\infty)$: π-weights

- The π weights come from the power expansion of the rational function in B: $\pi(B) = \frac{\phi_p(B)}{\theta_q(B)}$
- The expansion will converge $(\sum_{i=0}^{\infty} \pi_i^2 < \infty) \Leftrightarrow$ The module of all roots in $\theta_q(B)$ are greater than one
- This means that all complex roots of the characteristic polynomial of the MA part lie outside the unit circle.
- This condition implies that the model is **invertible**

Example: $MA(1) \quad X_t = (1+\theta B)Z_t$

Expression as an $AR(\infty)$:

$$\frac{1}{1+\theta B}X_t = Z_t$$

$$(1-\theta B+\theta^2 B^2 - ... + (-1)^k \theta^k B^k + ...)X_t = Z_t$$

In this case: $\pi_k = (-1)^k \theta^k$
The model is **invertible** $\Leftrightarrow |B| = |-\frac{1}{\theta}| > 1 \Leftrightarrow |\theta| < 1$

$MA(\infty)$: ψ-weights

- The ψ weights come from the power expansion of the rational function in B: $\psi(B) = \frac{\theta_q(B)}{\phi_p(B)}$
- The expansion will converge $(\sum_{i=0}^{\infty} \psi_i^2 < \infty) \Leftrightarrow$ The module of all roots in $\phi_p(B)$ are greater than one
- This means that all complex roots of the characteristic polynomial of the AR part lie outside the unit circle.
- This condition implies that the model is **causal** (stationary)

$MA(\infty)$: ψ-weights

Example: $AR(1) \quad (1-\phi B)X_t = Z_t$

Expression as an $MA(\infty)$:

$$X_t = \frac{1}{1-\phi B} Z_t$$

$$X_t = (1 + \phi B + \phi^2 B^2 - ... + \phi^k B^k + ...)Z_t$$

In this case: $\psi_k = \phi^k$

The model is **causal** $\Leftrightarrow |B| = |\frac{1}{\phi}| > 1 \Leftrightarrow |\phi| < 1$

- All $AR(p)$ models are **invertible**:

$$\pi_k = \phi_k \quad k = 1 \ldots p \Rightarrow \sum_{i=0}^{\infty} \pi_i^2 = \sum_{i=0}^{p} \pi_i^2 < \infty$$

- All $MA(q)$ models are **causal**:

$$\psi_k = \theta_k \quad k = 1 \ldots q \Rightarrow \sum_{i=0}^{\infty} \psi_i^2 = \sum_{i=0}^{q} \psi_i^2 < \infty$$

- An $ARMA(p, q)$ model will be...:
  - Invertible $\Leftrightarrow$ All roots of $\theta_q(B)$ have module greater than one
  - Causal $\Leftrightarrow$ All roots of $\phi_p(B)$ have module greater than one

For an **causal** and **invertible** $ARMA(p, q)$ model:

- Expression as an $AR(\infty)$ can be truncated when the π_k weight is very small. This is useful to calculate the point prediction by using past observations and the π-weights

- Expression as an $MA(\infty)$ can be truncated when the ψ_k weight is very small.
This is useful to calculate the variance of the prediction by using the ψ-weights

Example: $ARMA(1,2) \quad X_t = 0.8X_{t-1} + Z_t - 3Z_{t-1} + 2Z_{t-2}$

$$(1 - 0.8B)X_t = (1 - 3B + 2B^2)Z_t$$

Roots of the AR-polynomial: $1 - 0.8B = 0 \Rightarrow B = 1/0.8 = 1.25 > 1$

```
polyroot(c(1,-0.8))
```

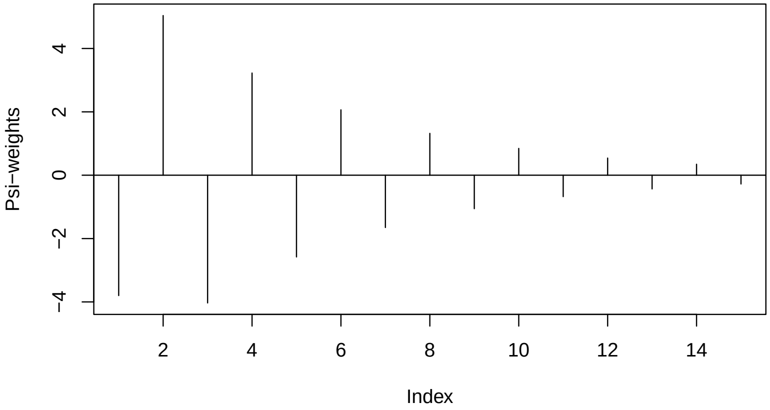
```
## [1] 1.25+0i
```

```
Mod(polyroot(c(1,-0.8)))
```

```
## [1] 1.25
```

Stationarity and Invertibility

The model $X_t = 0.8X_{t-1} + z_t - 3Z_{t-1} + 2Z_{t-2}$ is **causal**



Example: $ARMA(1,2) \quad x_t = 0.8x_{t-1} + z_t - 3z_{t-1} + 2z_{t-2}$

$$(1-0.8B)x_t = (1-3B+2B^2)z_t$$

Roots of the MA-polynomial: $1-3B+2B^2 = 0 \Rightarrow B = \{0.5, 1\}$

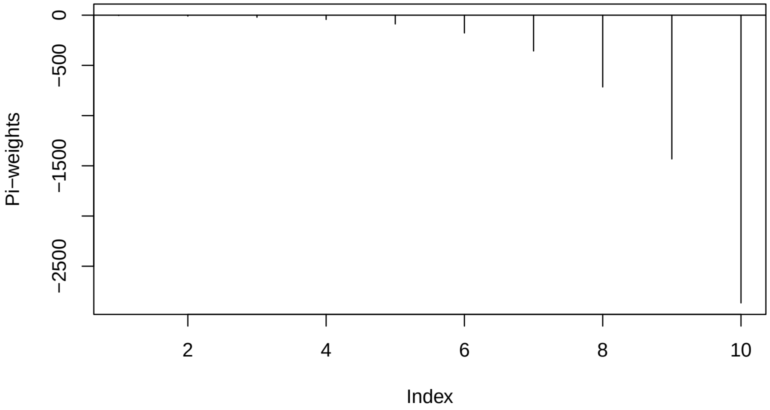
```
polyroot(c(1,-3,2))
```

```
## [1] 0.5+0i 1.0-0i
```

```
Mod(polyroot(c(1,-3,2)))
```

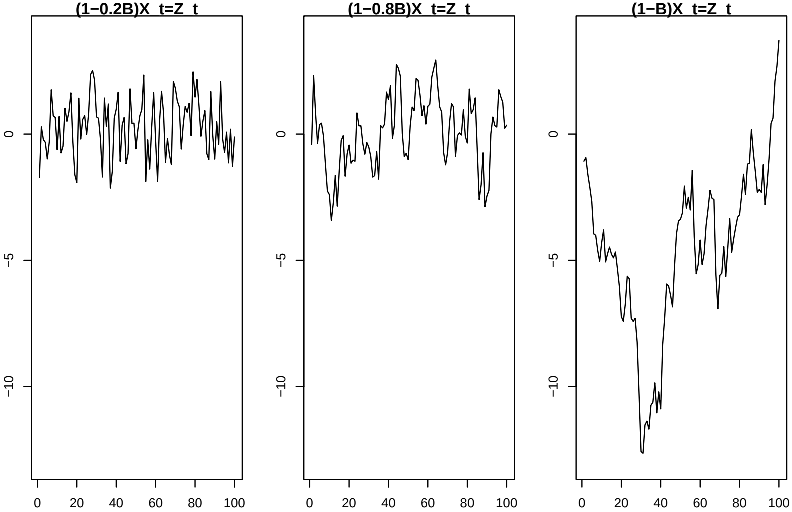
```
## [1] 0.5 1.0
```

The model $X_t = 0.8X_{t-1} + z_t - 3Z_{t-1} + 2Z_{t-2}$ is not **invertible**



Let's consider several $AR(1)$ models

$$(1-0.2B)X_t = Z_t \quad (1-0.8B)X_t = Z_t \quad (1-B)X_t = Z_t$$



Non-Stationary models with unit roots:Random Walk

Random Walk:

$$X_t = X_{t-1} + Z_t \qquad (1-B)X_t = Z_t \qquad Z_t \sim WN(0, \sigma_Z^2)$$

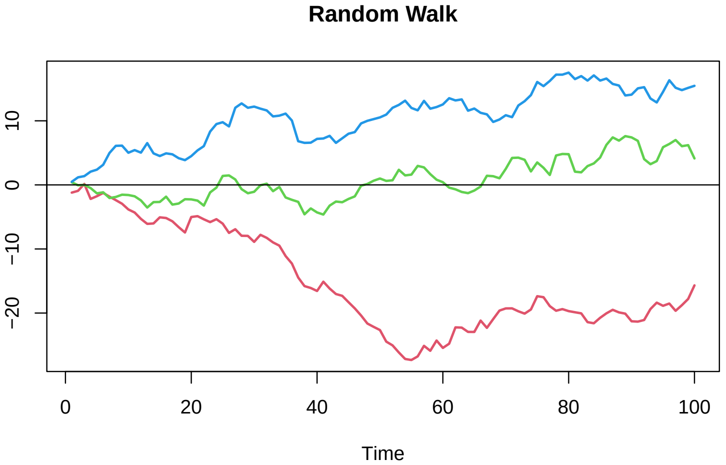
$$X_t = Z_t + Z_{t-1} + ... + Z_1$$

Non stationary process:

$$V(X_t) = t\sigma_Z^2$$

$$\gamma(X_t, X_{t+k}) = t\sigma_Z^2$$

Random Walk:



Non-Stationary models with unit roots

Non-stationary $ARMA(p,q)$ with 1 unit root:

$$\phi_p(B)X_t = \theta_q(B)Z_t$$

$$\phi_p(B) = 1 - \phi_1 B - ... - \phi_p B^p = (1 - \phi_1' B - ... - \phi_{p-1}' B)(1 - B)$$

x_t is a process with 1 unit root: Integrated of order 1 ($x_t \sim I(1)$)

This means that $W_t = (1 - B)X_t$ is an stationary $ARMA(p-1,q)$ process

x_t is an $ARIMA(p-1,1,q)$

Non-stationary $ARMA(p,q)$ with d unit roots:

$$\phi_p(B)X_t = \theta_q(B)Z_t$$

$$\phi_p(B) = 1 - \phi_1 B - ... - \phi_p B^p = (1 - \phi_1' B - ... - \phi_{p-d}' B)(1 - B)^d$$

x_t is a process with d unit roots: Integrated of order d ($X_t \sim I(d)$)

This means that $W_t = (1-B)^d X_t$ is an stationary $ARMA(p-d,q)$ process

X_t is an $ARIMA(p-d,d,q)$

Non-Stationary models

Non-stationary $ARIMA(p, d, q)$ with d unit roots:

$$\phi_p(B)(1-B)^d X_t = \theta_q(B)Z_t$$
$$(1-\phi_1 B - ... - \phi_p B^p)(1-B)^d X_t = (1+\theta_1 B + ... + \theta_q B^q)Z_t$$

In the Identification step:

- d = number of regular differences to reach stationarity
- p,q = analysis of ACF and PCF of the transformed series

Example: ARIMA(0,1,1)

$ARMA(1,1)$ with $\phi = 1 \Rightarrow$ Non-stationary and invertible (if $|\theta| < 1$)

$$(1 - B)X_t = (1 + \theta B)Z_t$$

The π-weights show an exponential decay:

$$X_t = (\theta + 1)X_{t-1} - \theta(\theta + 1)X_{t-2} + \theta^2(\theta + 1)X_{t-3} + ... + Z_t$$

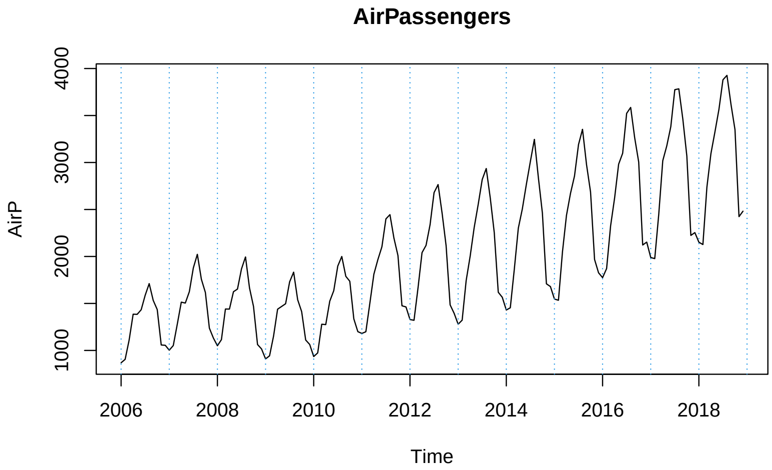
This model is the basis for the **EWMA** filter (Exponential Weighted Moving Average)

AirP: Number of monthly passengers (in thousands) of international air flights at El Prat between January 2006 and December 2018

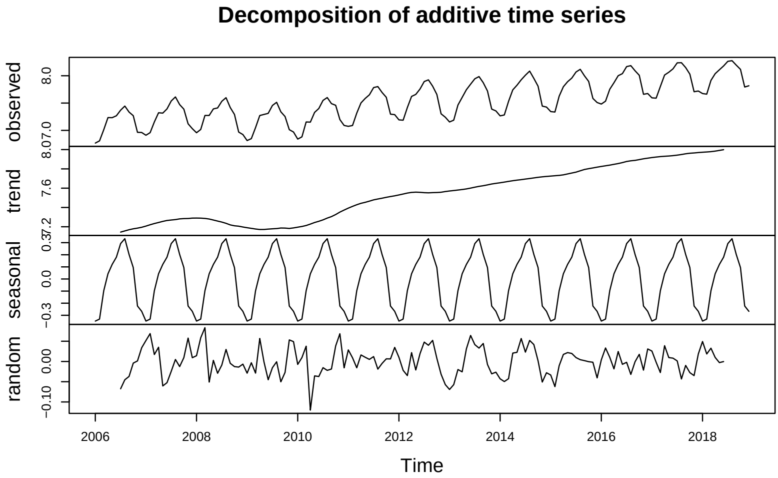
Source: Ministry of Public Works of Spain (http://www.fomento.gob.es/BE/?nivel=2&orden=03000000)

```
##       Jan  Feb  Mar  Apr  May  Jun  Jul  Aug  Sep  Oct  Nov  Dec
## 2006  869  905 1110 1386 1384 1433 1587 1711 1533 1435 1058 1055
## 2007 1003 1052 1279 1514 1504 1625 1878 2023 1758 1617 1237 1132
## 2008 1050 1115 1441 1440 1625 1653 1867 1995 1665 1468 1063 1016
## 2009  910  943 1154 1439 1468 1497 1730 1833 1540 1413 1111 1064
## 2010  935  974 1279 1275 1524 1636 1898 2000 1789 1736 1336 1200
## 2011 1181 1199 1504 1812 1967 2103 2400 2445 2195 2010 1475 1463
## 2012 1329 1321 1665 2041 2118 2337 2680 2765 2462 2115 1485 1395
## 2013 1279 1322 1744 2005 2308 2554 2819 2936 2625 2255 1619 1566
## 2014 1430 1454 1865 2301 2506 2767 3011 3246 2838 2461 1709 1680
## 2015 1547 1534 2049 2437 2671 2856 3191 3353 2977 2686 1968 1826
## 2016 1775 1871 2327 2623 2982 3098 3523 3586 3263 3005 2122 2153
## 2017 1989 1977 2448 3019 3177 3381 3775 3784 3462 3068 2224 2254
## 2018 2152 2127 2737 3094 3325 3566 3880 3927 3619 3349 2424 2482
```

X_t = Increasing Variance + Linear Trend + Seasonal component + stationary process

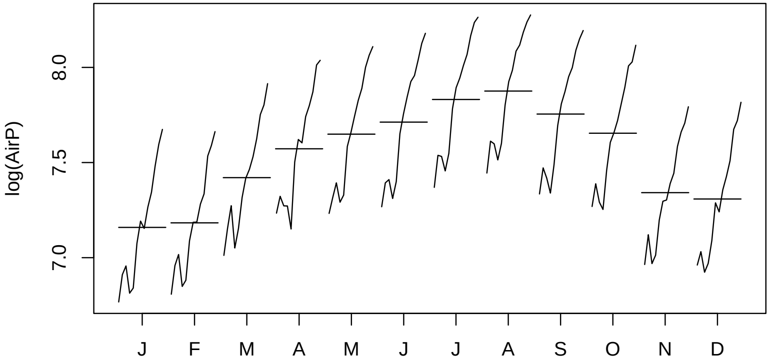


$log(X_t)$ = Linear Trend + Seasonal component + stationary process

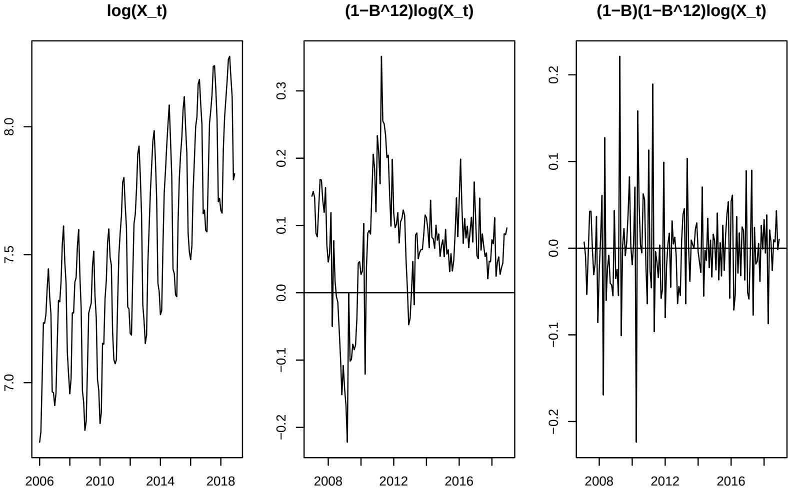


Seasonal Model: AirPassengers

$log(X_t)$ = Linear Trend + Seasonal component + stationary process



Seasonal Model: AirPassengers



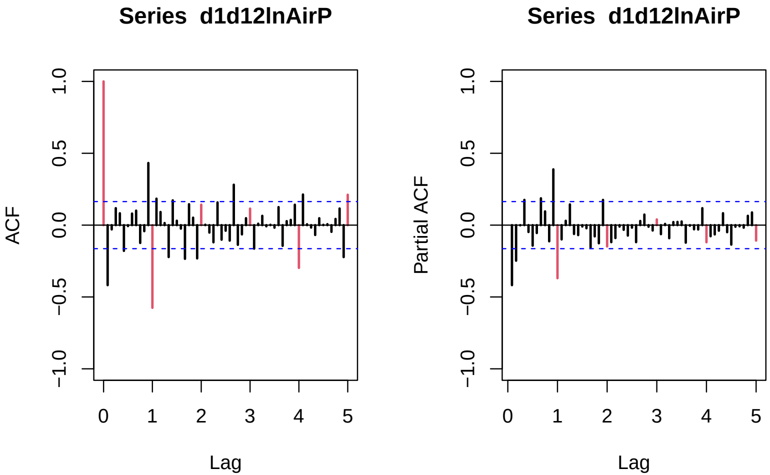
Seasonal Model: AirPassengers

Transformation into an stationary series:

$$W_t = (1 - B)(1 - B^{12})log(X_t)$$

- Logarithm scale: to stabilize the variance
- Seasonal difference: remove the seasonal pattern (and perhaps a global linear trend: $(1 - B^{12}) = (1 - B)(1 + B + ... + B^{11})$)
- Regular difference: to reach a constant mean

Seasonal Model: AirPassengers



Non-stationary Seasonal $ARIMA(p, d, q)(P, D, Q)_s$ (or SARIMA):

$$\phi_p(B)\Phi_P(B^s)(1-B)^d(1-B^s)^D X_t = \theta_q(B)\Theta_Q(B^s)Z_t$$

$$(1-\phi_1 B - \ldots - \phi_p B^p)(1-\Phi_1 B^s - \ldots - \Phi_P B^s P)(1-B)^d(1-B^s)^D X_t$$
$$= (1+\theta_1 B + \ldots + \theta_q B^q)(1+\Theta_1 B^s + \ldots + \Theta_Q B^s Q)Z_t$$

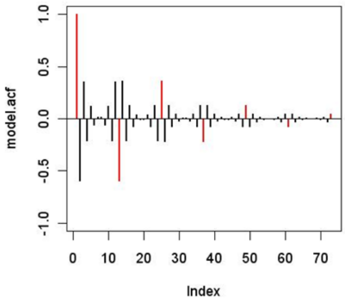
In the Identification step:

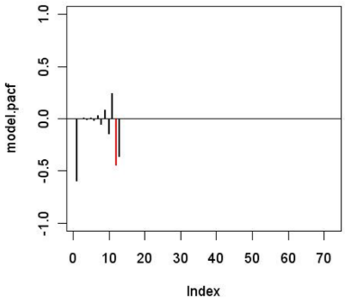
- d = number of regular differences to reach stationarity
- D = number of seasonal differences to reach stationarity (usually 0 or 1)
- p,q = analysis of ACF and PCF (only first lags)
- P,Q = analysis of ACF and PCF (only lags multiple of s)

Example 1: The following Theoretical ACF and PACF belong to models $ARMA(p,q)(P,Q)_{12}$ of this form:

$$(1-\phi B)(1-\Phi B^{12})W_t = (1+\theta B)(1+\Theta B^{12})Z_t$$

and $\phi, \Phi, \theta, \Theta \in \{-0.6, 0, 0.6\}$



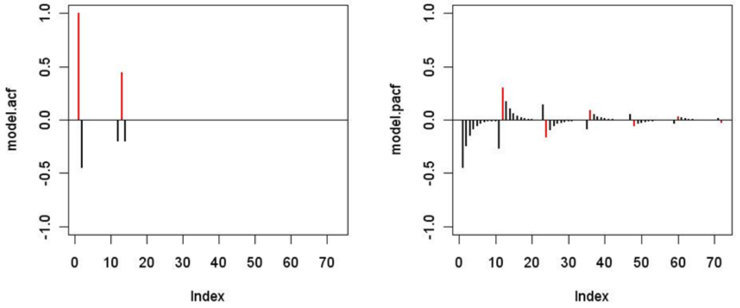


Seasonal ARIMA models

Example 2: The following Theoretical ACF and PACF belong to models $ARMA(p,q)(P,Q)_{12}$ of this form:

$$(1-\phi B)(1-\Phi B^{12})W_t = (1+\theta B)(1+\Theta B^{12})Z_t$$

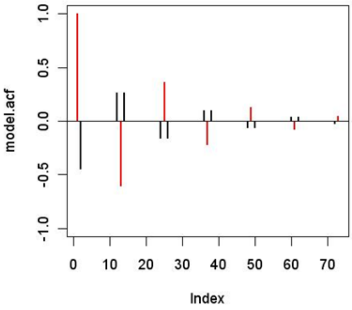
and $\phi, \Phi, \theta, \Theta \in \{-0.6, 0, 0.6\}$



Example 3: The following Theoretical ACF and PACF belong to models $ARMA(p,q)(P,Q)_{12}$ of this form:

$$(1-\phi B)(1-\Phi B^{12})W_t = (1+\theta B)(1+\Theta B^{12})Z_t$$

and $\phi, \Phi, \theta, \Theta \in \{-0.6, 0, 0.6\}$



Seasonal ARIMA models

Example 1: $(1 + 0.6B)(1 + 0.6B^{12})W_t = Z_t$

Example 2: $W_t = (1 - 0.6B)(1 + 0.6B^{12})Z_t$

Example 3: $(1 + 0.6B^{12})W_t = (1 - 0.6B)Z_t$