NDA: Time Series Analysis - Part 2

Lionel Tabourier

LIP6 – CNRS and Sobonne University

first_name.last_name@lip6.fr

January 2021

Outline

Outline

1. What to do with the residuals?
2. Autoregressive models
3. Moving Average models
4. ARMA models

How to analyze a time series? (3)

Third step:

- fit the residuals

If the residuals is an IID time series, nothing else to model...

Simple method to test if a time series is IID

General idea: suppose IID random variables
What should we observe? Is it the case?

Simple method to test if a time series is IID

General idea: suppose IID random variables
What should we observe? Is it the case?

Sample ACF criterion

Because of the Central Limit Theorem:

- suppose x_t IID with mean 0 and variance 1 (white noise)
- if n large enough, $\hat{\rho}_x(h)$ is approx. distributed as $\mathcal{N}(0, \frac{1}{\sqrt{n}})$

In practice, consider the 95% confidence interval:
How many values fall out of $\left[\frac{-1.96}{\sqrt{n}}, \frac{+1.96}{\sqrt{n}}\right]$? By how much?

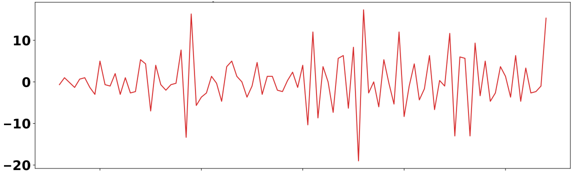
Simple method to test if a time series is IID

General idea: suppose IID random variables
What should we observe? Is it the case?

Sample ACF criterion

Because of the Central Limit Theorem:

- suppose x_t IID with mean 0 and variance 1 (white noise)
- if n large enough, $\hat{\rho}_x(h)$ is approx. distributed as $\mathcal{N}(0, \frac{1}{\sqrt{n}})$



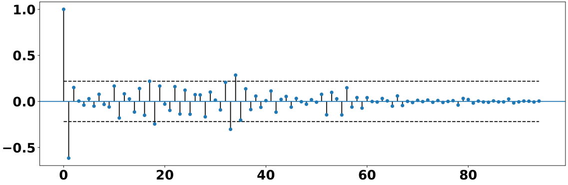
Simple method to test if a time series is IID

General idea: suppose IID random variables
What should we observe? Is it the case?

Sample ACF criterion

Because of the Central Limit Theorem:

- suppose x_t IID with mean 0 and variance 1 (white noise)
- if n large enough, $\hat{\rho}_x(h)$ is approx. distributed as $\mathcal{N}(0, \frac{1}{\sqrt{n}})$



not IID noise

How to analyze a time series? (3)

Third step:

- fit the residuals

If the residuals is an IID time series, nothing else to model...
Otherwise, we use ARMA models

Outline

1. What to do with the residuals?
2. Autoregressive models
3. Moving Average models
4. ARMA models

AR(1) process

What is autoregression?

auto means self $\Rightarrow$ regression from itself

The most basic AR model: 1^{st} order regression or AR(1)

$\{X_t\}$ is a series satisfying:

$$X_t = \phi X_{t-1} + W_t \ , \quad |\phi| < 1$$

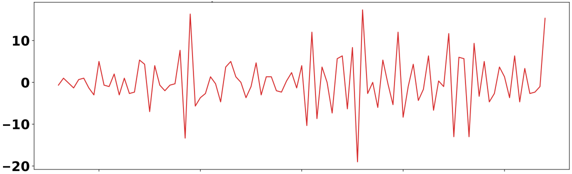
where W_t is a white noise (mean 0, variance σ^2)

Reminder: stationary

- $\mathbb{E}[X_t] = 0$
- $\gamma_X(h) = \phi^{|h|}\gamma_X(0) = \phi^{|h|}\frac{\sigma^2}{1-\phi^2}$

AR(1) process

Consider the following time series r_t:



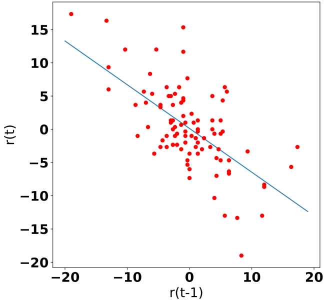
Suppose AR(1) model for a time series r_t, how to compute ϕ?

- plot r_t as a function of r_{t-1} *(lag–1 plot)*
- linear fit, slope is an estimation of ϕ

AR(1) process

Suppose AR(1) model for a time series r_t, how to compute ϕ?

- plot r_t as a function of r_{t-1} *(lag–1 plot)*
- linear fit, slope is an estimation of ϕ

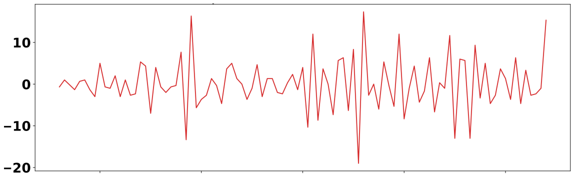


on the example $\phi \simeq -0.659$

AR(1) process

Now, we compute the residuals of the TS – AR(1) model:

$$r_t - \phi r_{t-1}$$

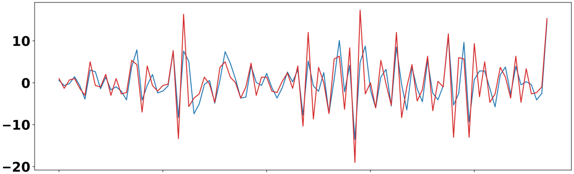


Is it closer to an IID noise?

AR(1) process

Now, we compute the residuals of the TS – AR(1) model:

$$r_t - \phi r_{t-1}$$



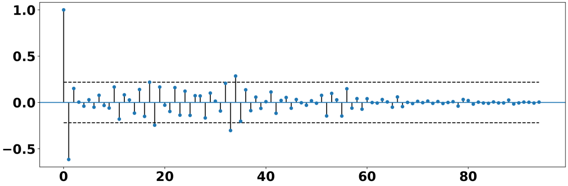
Is it closer to an IID noise?

AR(1) process

Now, we compute the residuals of the TS – AR(1) model:

$$r_t - \phi r_{t-1}$$

Is it closer to an IID noise?



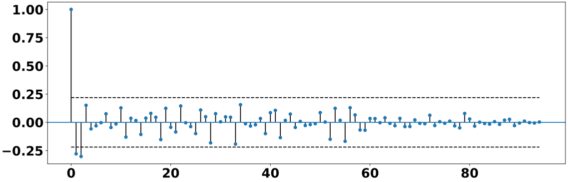
before AR(1): not IID

AR(1) process

Now, we compute the residuals of the TS – AR(1) model:

$$r_t - \phi r_{t-1}$$

Is it closer to an IID noise?



after AR(1): closer to IID $\longrightarrow$ other tests?

Generalized AR models

AR(1) model

W_t is white noise
signal = noise and (weighted) influence of the signal at $t - 1$

- $X_t = \phi X_{t-1} + W_t$
- $X_t = \phi_1 X_{t-1} + \phi_2 X_{t-2} + \ldots + \phi_p X_{t-p} + W_t$

Generalized AR models

AR(p) model

W_t is white noise
signal = noise and influence of the signal at p previous steps

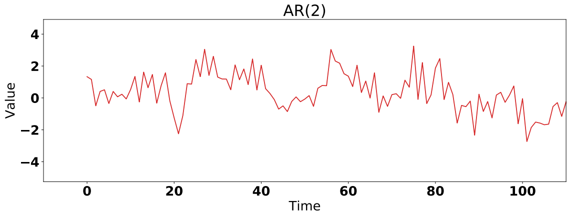
- $X_t = \phi X_{t-1} + W_t$
- $X_t = \phi_1 X_{t-1} + \phi_2 X_{t-2} + \ldots + \phi_p X_{t-p} + W_t$

Generalized AR models

$$X_t - \phi_1 X_{t-1} - \phi_2 X_{t-2} - \ldots - \phi_p X_{t-p} = W_t$$

Some characteristics

- Stationary process? see characteristic polynomial
  $P(x) = 1 - \phi_1 x - \phi_2 x^2 - \ldots - \phi_p x^p \rightarrow$ if no unit root AR(p) stationary
- typical shape: smooth decay, no cut-off

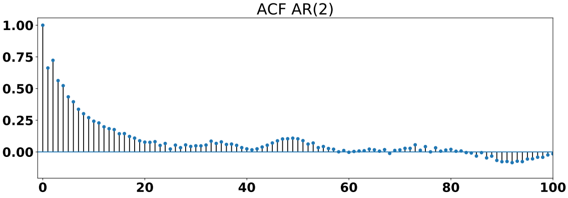


Generalized AR models

$$X_t - \phi_1 X_{t-1} - \phi_2 X_{t-2} - \ldots - \phi_p X_{t-p} = W_t$$

Some characteristics

- Stationary process? see characteristic polynomial
  $P(x) = 1 - \phi_1 x - \phi_2 x^2 - \ldots - \phi_p x^p \rightarrow$ if no unit root AR(p) stationary
- typical shape: smooth decay, no cut-off



How to find AR(p) ACF coefficients?

$$X_t = \phi_1 X_{t-1} + \phi_2 X_{t-2} + \ldots + W_t$$

Yule-Walker equations

We have seen that for AR(1):

$$\gamma(h) = \phi^{|h|}\gamma(0) = \phi^{|h|}\frac{\sigma^2}{1-\phi^2}$$

What about the general case?

How to find AR(p) ACF coefficients?

$$X_t = \phi_1 X_{t-1} + \phi_2 X_{t-2} + \ldots + W_t$$

Yule-Walker equations

We have seen that for AR(1):

$$\gamma(h) = \phi^{|h|}\gamma(0) = \phi^{|h|}\frac{\sigma^2}{1-\phi^2}$$

What about the general case?

- $\gamma(h) = \phi_1\gamma(h-1) + \phi_2\gamma(h-2) + \ldots$
- $\rho(h) = \phi_1\rho(h-1) + \phi_2\rho(h-2) + \ldots$

these are called the Yule-Walker equations

How to find AR(p) ACF coefficients?

$$X_t = \phi_1 X_{t-1} + \phi_2 X_{t-2} + \ldots + W_t$$

Yule-Walker equations

We have seen that for AR(1):

$$\gamma(h) = \phi^{|h|}\gamma(0) = \phi^{|h|}\frac{\sigma^2}{1-\phi^2}$$

What about the general case?

- $\gamma(h) = \phi_1\gamma(h-1) + \phi_2\gamma(h-2) + \ldots$
- $\rho(h) = \phi_1\rho(h-1) + \phi_2\rho(h-2) + \ldots$

these are called the Yule-Walker equations

How to find AR(p) ACF coefficients?

$$X_t = \phi_1 X_{t-1} + \phi_2 X_{t-2} + \ldots + W_t$$

Yule-Walker equations

We have seen that for AR(1):

$$\gamma(h) = \phi^{|h|}\gamma(0) = \phi^{|h|}\frac{\sigma^2}{1-\phi^2}$$

What about the general case?

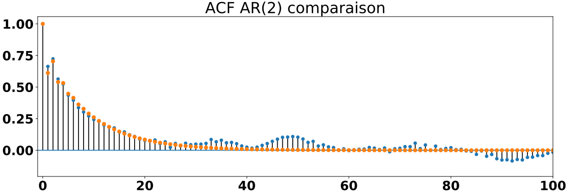
- $\gamma(h) = \phi_1\gamma(h-1) + \phi_2\gamma(h-2) + \ldots$
- $\rho(h) = \phi_1\rho(h-1) + \phi_2\rho(h-2) + \ldots$

these are called the Yule-Walker equations

How to find AR(p) ACF coefficients?

Illustration on a practical case

$$X_t = \frac{1}{3}X_{t-1} + \frac{1}{2}X_{t-2} + W_t$$

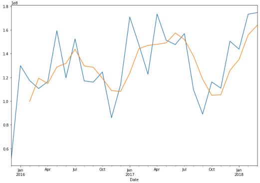


Sample ACF (blue) vs Yule-Walker coefficients (orange)

Outline

MA(q) processes

Moving average to smooth a signal:



Original signal $f(x)$ smoothed over a window size w, typically:

ex1: $g_1(x) = \frac{f(x)+f(x-1)+f(x+1)}{3}$ (no weight, two-sided, $w = 3$)

ex2: $g_2(x) = \frac{f(x)+f(x-1)+f(x-2)}{3}$ (no weight, one-sided, $w = 3$)

ex3: $g_3(x) = \frac{f(x)+a.f(x-1)+b.f(x-2)+c.f(x-3)}{(1+a+b+c)}$ (weighted, one-sided, $w = 4$)

MA(q) processes

MA(1) model

W_t is white noise
signal = weighted average of noise at t and of noise at $t-1$

- $X_t = \beta_0 W_t + \beta_1 W_{t-1}$
- $X_t = \beta_0 W_t + \beta_1 W_{t-1} + \ldots + \beta_q W_{t-q}$

MA(q) processes

MA(q) model

W_t is white noise
signal = weighted average of noise at t and q previous steps

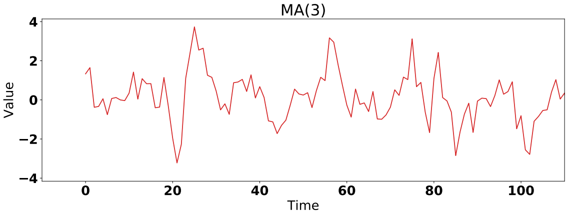
- $X_t = \beta_0 W_t + \beta_1 W_{t-1}$
- $X_t = \beta_0 W_t + \beta_1 W_{t-1} + \ldots + \beta_q W_{t-q}$

MA(q) processes

$$X_t = \beta_0 W_t + \beta_1 W_{t-1} + \ldots + \beta_q W_{t-q}$$

Some characteristics

- Stationary process
  mean and autocovariance at lag h do not depend on time
  Ex: prove if $h \leq q$, $Cov(X_t, X_{t+h}) = \sigma^2 \sum_{i=0}^{q-h} \beta_i \beta_{i+h}$
- ACF cuts off at lag q

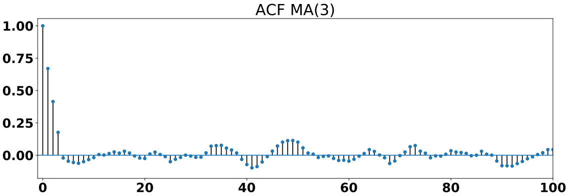


MA(q) processes

$$X_t = \beta_0 W_t + \beta_1 W_{t-1} + \ldots + \beta_q W_{t-q}$$

Some characteristics

- Stationary process
  mean and autocovariance at lag h do not depend on time
  Ex: prove if $h \leq q$, $Cov(X_t, X_{t+h}) = \sigma^2 \sum_{i=0}^{q-h} \beta_i \beta_{i+h}$
- ACF cuts off at lag q



Connection between MA and AR

From MA(1) to AR

Considering the MA(1) process:

$$
\begin{aligned}
X_t &= W_t + \beta W_{t-1} \\
\Rightarrow W_t &= X_t - \beta W_{t-1} \\
\Rightarrow W_t &= X_t - \beta(X_{t-1} - \beta W_{t-2}) \\
\Rightarrow W_t &= X_t - \beta X_{t-1} + \beta^2 X_{t-2} - \beta^3 X_{t-3} + \dots \\
\Rightarrow X_t &= W_t + \beta X_{t-1} - \beta^2 X_{t-2} + \beta^3 X_{t-3} - \dots
\end{aligned}
$$

In other words, MA(1) is an AR(∞) process

More generally, MA(q) can (often) be seen as AR(∞) process

Connection between MA and AR

From MA(1) to AR

Considering the MA(1) process:

$$
\begin{aligned}
X_t &= W_t + \beta W_{t-1} \\
\Rightarrow W_t &= X_t - \beta W_{t-1} \\
\Rightarrow W_t &= X_t - \beta(X_{t-1} - \beta W_{t-2}) \\
\Rightarrow W_t &= X_t - \beta X_{t-1} + \beta^2 X_{t-2} - \beta^3 X_{t-3} + \dots \\
\Rightarrow X_t &= W_t + \beta X_{t-1} - \beta^2 X_{t-2} + \beta^3 X_{t-3} - \dots
\end{aligned}
$$

In other words, MA(1) is an AR(∞) process

More generally, MA(q) can (often) be seen as AR(∞) process

Connection between MA and AR

From AR(1) to MA

$$
\begin{aligned}
X_t &= W_t + \phi X_{t-1} \\
\Rightarrow X_t &= W_t + \phi(W_{t-1} + \phi X_{t-2}) \\
\Rightarrow X_t &= \dots \\
\Rightarrow X_t &= W_t + \phi W_{t-1} + \phi^2 W_{t-2} + \phi^3 W_{t-3} + \dots
\end{aligned}
$$

In other words, AR(1) is an MA(∞) process

More generally, AR(p) can (often) be seen as MA(∞) process

Connection between MA and AR

From AR(1) to MA

$$\begin{aligned} X_t &= W_t + \phi X_{t-1} \\ \Rightarrow X_t &= W_t + \phi(W_{t-1} + \phi X_{t-2}) \\ \Rightarrow X_t &= \dots \\ \Rightarrow X_t &= W_t + \phi W_{t-1} + \phi^2 W_{t-2} + \phi^3 W_{t-3} + \dots \end{aligned}$$

In other words, AR(1) is an MA(∞) process

More generally, AR(p) can (often) be seen as MA(∞) process

Outline

ARMA models

ARMA(p,q) model

ARMA(p,q) model is a combination of AR(p) and MA(q) model:

$$X_t = \phi_1 X_{t-1} + \ldots + \phi_p X_{t-p} + W_t + \beta_1 W_{t-1} + \ldots + \beta_q W_{t-q}$$

As for AR(p) and MA(q) parameters can be found from the ACF

In practice

- fit the residuals with several (low) values of p and q
- select what is the *best* model
- $\Rightarrow$ complete model:

  trend + seasonality + ARMA(p,q) residuals

ARMA models

ARMA(p,q) model

ARMA(p,q) model is a combination of AR(p) and MA(q) model:

$$X_t = \phi_1 X_{t-1} + \ldots + \phi_p X_{t-p} + W_t + \beta_1 W_{t-1} + \ldots + \beta_q W_{t-q}$$

As for AR(p) and MA(q) parameters can be found from the ACF

In practice

- fit the residuals with several (low) values of p and q
- select what is the *best* model
- $\Rightarrow$ complete model:

  trend + seasonality + ARMA(p,q) residuals

A few questions that we have left aside ...

- How to choose p and q of an ARMA process?
  $\rightarrow$ *useful to define partial ACF*
- How to find the coefficients of an ARMA process?
  $\rightarrow$ *transform it in AR(∞) (or MA(∞)) process*
- How do we select the best model?
  $\rightarrow$ *complexity criteria and the overfitting problem (AIC, BIC)*
- How to deal more systematically with the trend?
  $\rightarrow$ *differentiation method*

A few questions that we have left aside . . .

- How to choose p and q of an ARMA process?
  $\rightarrow$ *useful to define partial ACF*
- How to find the coefficients of an ARMA process?
  $\rightarrow$ *transform it in AR(∞) (or MA(∞)) process*
- How do we select the best model?
  $\rightarrow$ *complexity criteria and the overfitting problem (AIC, BIC)*
- How to deal more systematically with the trend?
  $\rightarrow$ *differentiation method*

A few questions that we have left aside . . .

- How to choose p and q of an ARMA process?
  $\rightarrow$ *useful to define partial ACF*
- How to find the coefficients of an ARMA process?
  $\rightarrow$ *transform it in AR(∞) (or MA(∞)) process*
- How do we select the best model?
  $\rightarrow$ *complexity criteria and the overfitting problem (AIC, BIC)*
- How to deal more systematically with the trend?
  $\rightarrow$ *differentiation method*

A few questions that we have left aside ...

- How to choose p and q of an ARMA process?
  $\rightarrow$ *useful to define partial ACF*
- How to find the coefficients of an ARMA process?
  $\rightarrow$ *transform it in AR(∞) (or MA(∞)) process*
- How do we select the best model?
  $\rightarrow$ *complexity criteria and the overfitting problem (AIC, BIC)*
- How to deal more systematically with the trend?
  $\rightarrow$ *differentiation method*

Studying time series in python

Among several options, `pandas` library

A few useful functions:

- Load data as dataframe:
  `read_csv` from `pandas` library
- Moving average:
  `rolling` from `pandas` library
- Fitting:
  `curve_fit` in `scipy.optimize` library
- Autocorrelation function:
  `plot_ACF` in `statsmodels` library
- ARMA model fit:
  `ARMA.fit` in `statsmodels` library