Time Series Analysis

Christian FRANCQ Jean-Michel ZAKOIAN

CREST

Chapter 1: Introduction

Objective: **study linear univariate time series models**

Time series analysis is central to a wide range of applications

- business cycle measurement,
- financial risk management,
- policy analysis,
- forecasting...

Some topics of interest are particular to time series analysis:

- stationarity/non stationarity,
- trends and cycles
- seasonality, periodicity,
- predictability,
- structural changes,
- linearity/non linearity

Classical textbooks

- Brockwell and Davis (1991) Time Series: Theory and Methods. Springer Verlag.
- Brockwell and Davis (2002) Introduction to Time Series and Forecasting. Springer Verlag.
- Gouriéroux et Monfort (1995) Séries temporelles et modèles dynamiques. Economica.
- Hamilton (1994) Times Series Analysis. Princeton University Press.
- Box and Jenkins (1970) Time Series Analysis: Forecasting and Control. Holden-Day.

Resources **freely available** on the web

- Alex Aue's 'Time Series Analysis' Link
- Bruce Hansen's 'Advanced Time Series and Forecasting' Link
- Frank Diebold's 'Time-Series Econometrics: a concise course' Link
- John Cochrane's 'Time Series for Macroeconomics and Finance' Link
- ...

Plan of the lecture

Chapter 1: Introduction
Chapter 2: ARMA models
Chapter 3: Using ARMA and SARIMA models
Chapter 4: Unit root tests

Outline

1. Introduction
2. Basic time series models
3. Estimating the 1st and 2nd order moments

1 Introduction
- Definition and examples
- Stationary models
- White noise, Theoretical Predictions

2 Basic time series models

3 Estimating the 1st and 2nd order moments

Time Series

Any series of observations ordered along time (or any other single dimension) may be thought of as **a time series**.

Many economic and financial variables are observed over time:

- prices, stock returns,
- sales, stocks,
- GDP,
- interest rates and foreign exchange rates...

In addition to being interested in

the interrelationships among such variables,

we are also concerned with

relationships among the current and past values of one or more of them,

that is, relationship over time.

At the theoretical level

Modern time series analysis is related to the **theory of stochastic processes**.

$\mathscr{T}$: finite set of dates

A time series is a collection of random variables

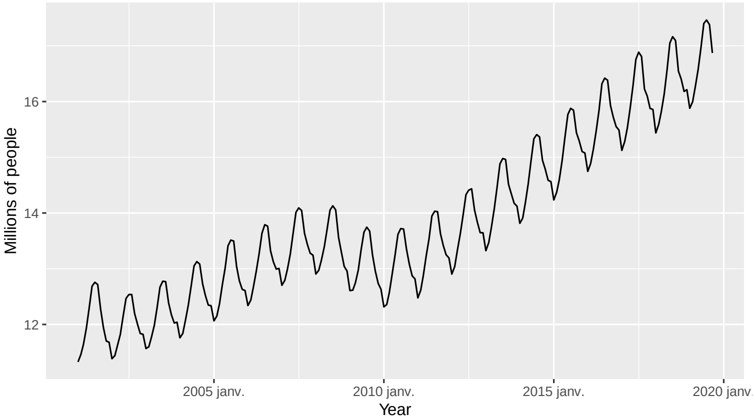
$$X = (X_t)_{t \in \mathscr{T}}$$

$X_t \in \mathbb{R}^d$: $d = 1$ univariate time series
$d > 1$ multivariate time series

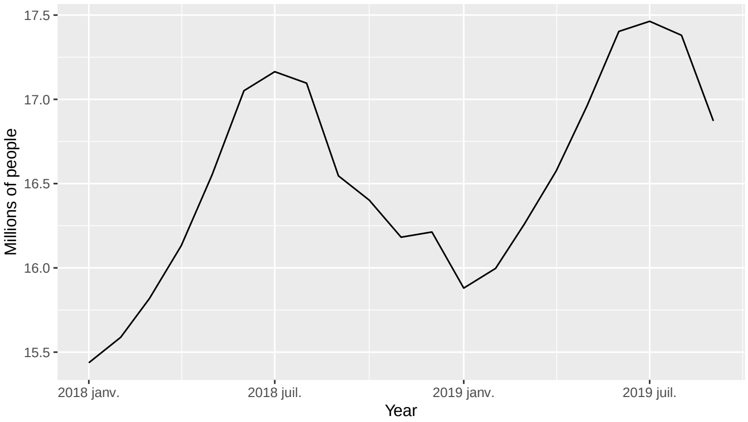
Examples:

- Monthly leisure and hospitality jobs
- Black and white pepper price (bivariate time series)
- Daily returns of the CAC index

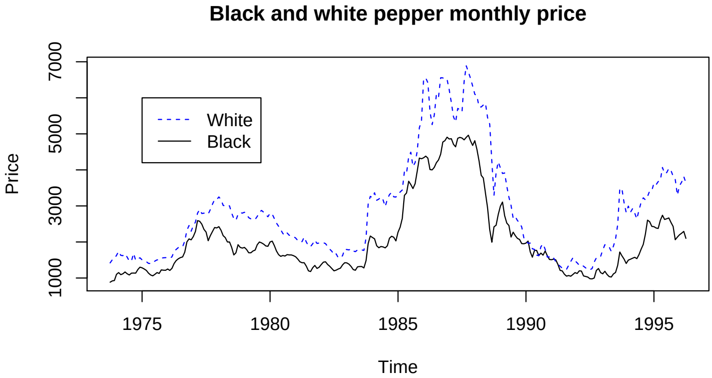
Monthly US leisure and hospitality jobs from January 2000 to September 2019



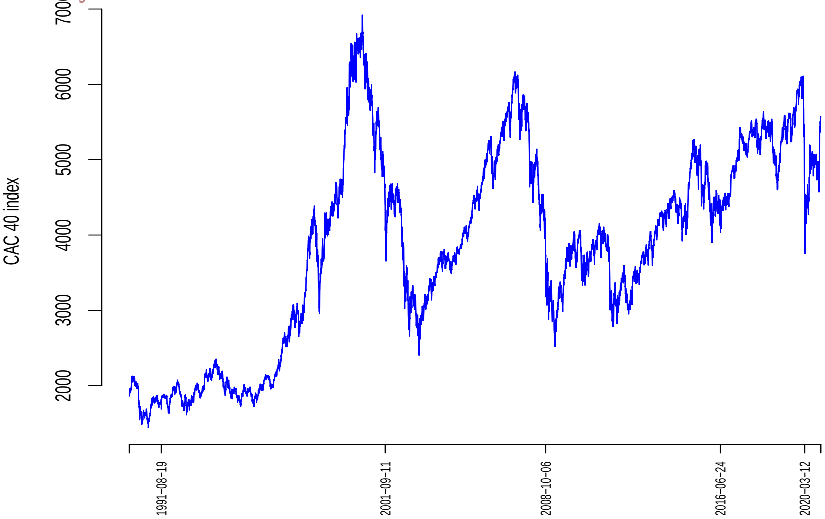
US leisure and hospitality jobs (zoom on the last values)



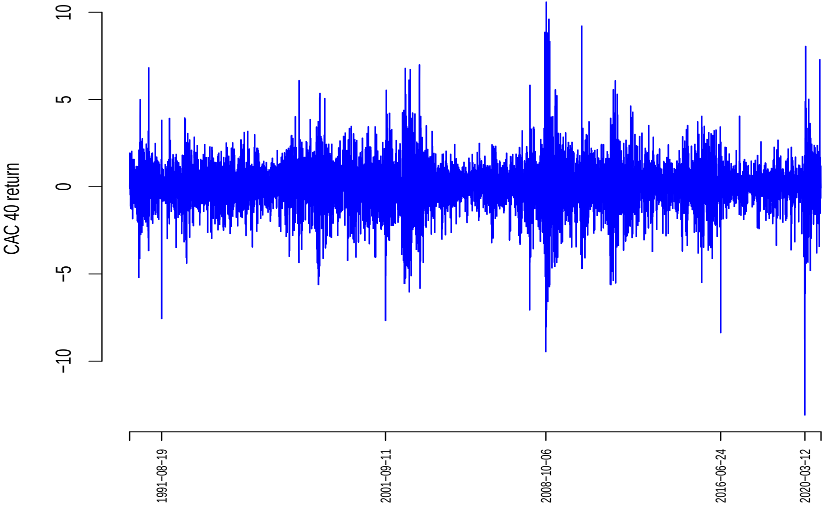
Bivariate series of black and white pepper prices



CAC daily index from 1992-01-03 to 2020-11-26



Log-returns $\epsilon_t = \log(p_t/p_{t-1}) \approx \frac{p_t - p_{t-1}}{p_{t-1}}$



Stochastic process point of view

Assume that each observation x_t $(t = 1, \ldots, n)$ is the realisation of some random variable X_t.

Definition

A **time series model** is an assumption on the joint distribution of the sequence of variables (or stochastic process) (X_t).

⚠The term "time series" is used for both the (partial) realisation $(x_t)_{1 \leq t \leq n}$ and the sequence of variables $(X_t)_{t \in \mathbb{Z}}$.

1 Introduction
- Definition and examples
- Stationary models
- White noise, Theoretical Predictions

2 Basic time series models

3 Estimating the 1st and 2nd order moments

Strict stationarity

A time series (X_t) is called stationary if its probabilistic properties are the same as those of the series (X_{t+h}), for any integer h.

Definition

(X_t) is strictly stationary if

$$(X_1, X_2, \ldots, X_k) \text{ has the same distribution as } (X_{1+h}, X_{2+h}, \ldots, X_{k+h})$$

for any h and any $k \geq 1$.

This concept can be difficult to manipulate.

Second-order stationarity

Definition

Let (X_t) such that $EX_t^2 < \infty$. The mean function of (X_t) is

$$\mu_X(t) = E(X_t)$$

The autocovariance function of (X_t) is

$$\gamma_X(r,s) = \mathrm{Cov}(X_r, X_s)$$

Definition

(X_t) is (second-order) stationary if

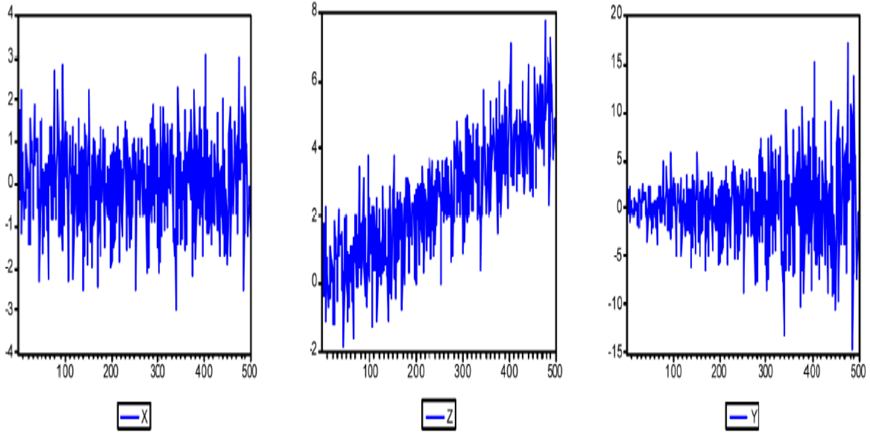
(i) $\mu_X(t)$ is independent of t, and

(ii) $\gamma_X(t, t+h)$ is independent of t, for any h.

Illustration on simulated series

(see the previous graphs for real examples)

(X_t) is stationary, (Y_t) and (Z_t) are not



Autocovariance and autocorrelation functions

Definition (case $d=1$)

Let (X_t) a univariate stationary time series.
The autocovariance function of (X_t) is

$$\gamma_X(h) = \mathrm{Cov}(X_t, X_{t+h}), \quad h = 0, \pm 1, \pm 2, \ldots$$

The autocorrelation function is

$$\rho_X(h) = \frac{\gamma_X(h)}{\gamma_X(0)} = \mathrm{Cor}(X_t, X_{t+h}), \quad h = 0, \pm 1, \pm 2, \ldots$$

Remark: even functions

$$\gamma_X(h) = \gamma_X(-h), \quad \rho_X(h) = \rho_X(-h)$$

Links between strict and 2nd-order stationarity

- **(X_t) strictly stationary $+EX_t^2 < \infty$**

  $\Rightarrow$ (X_t) 2nd-order stationary

- **(X_t) Gaussian* and 2nd-order stationary**

  $\Rightarrow$ (X_t) strictly stationary

*i.e. any linear combination of the X_t's is Gaussian

Linear transformation of a stationary time series

Proposition

Let Z a stationary time series, $(a_i)_{i\in Z}$ a sequence of coefficients such that $\sum_{i=-\infty}^{\infty}|a_i| < \infty$. Then

$$X_t = \sum_{i=-\infty}^{\infty} a_i Z_{t-i}$$

defines a new stationary time series.

Remark: In particular, any finite linear combination of the variables Z_{t-i} is always stationary.

More details

Autocovariance of a linear transformation

Proposition

Let (Z_t) a stationary time series with mean $E(Z_t) = 0$ and autocovariance function γ_Z. Then

$$X_t = \sum_{i=-\infty}^{\infty} a_i Z_{t-i}, \quad \sum_{i=-\infty}^{\infty} |a_i| < \infty$$

is stationary with mean 0, and autocovariance function

$$\gamma_X(h) = \sum_{i=-\infty}^{\infty} \sum_{j=-\infty}^{\infty} a_i a_j \gamma_Z(h+i-j).$$

1. Introduction
   - Definition and examples
   - Stationary models
   - White noise, Theoretical Predictions
2. Basic time series models
3. Estimating the 1st and 2nd order moments

White noise

Definition

A (weak) white noise is a sequence (ϵ_t) of uncorrelated variables with zero mean and constant variance:

$$E(\epsilon_t) = 0, \quad \mathrm{Var}(\epsilon_t) = \sigma^2, \qquad \mathrm{Cov}(\epsilon_t, \epsilon_s) = 0, \; t \neq s$$

Notation: $(\epsilon_t) \sim \; WN\,(0, \sigma^2)$

Autocovariance function: when $d = 1$,

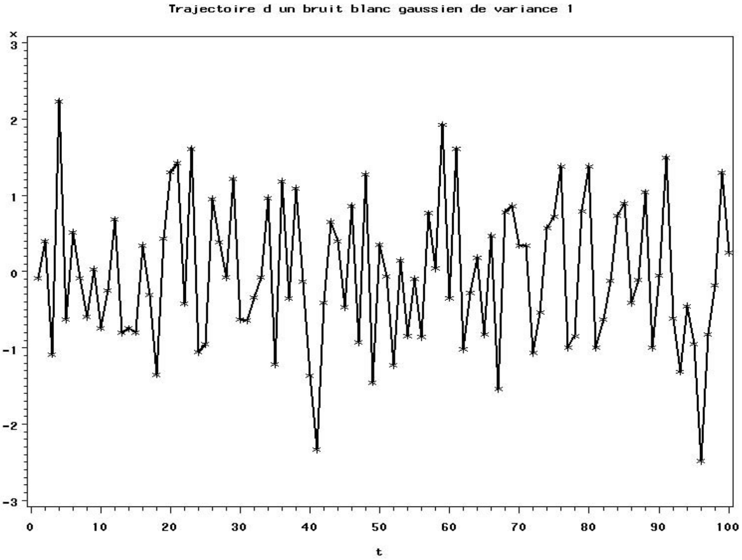
$$\gamma_\epsilon(h) = \begin{cases} \sigma^2, & h = 0 \\ 0, & h \neq 0 \end{cases}$$

Strong white noise

Definition

A strong WN is a sequence (ϵ_t) of **independent and identically distributed (iid)** variables, with zero mean and finite variance σ^2.

Notation: $(\epsilon_t)\ iid\,(0, \sigma^2)$


Trajectoire d un bruit blanc gaussien de variance 1
x
3
2
1
0
-1
-2
-3
0
10
20
30
40
50
60
70
80
90
100
t

Links between the different types of WN

weak WN $\supset$ strong $\supset$ Gaussian WN

Most time series models can be written under the form

$$X_t = \varphi(X_{t-1}, X_{t-2}, \dots) + \epsilon_t$$

In this course, the focus is on linear functions φ.

Conditional expectation

For $X, Y \in L^2$, the conditional expectation of X given Y is the random variable function of Y which minimizes (in Z)

$$E(X-Z)^2$$

and it is denoted $E(X|Y)$.

It can be interpreted as the best approximation of X as a function of Y.

Note that $E(X|Y)$ is also the mean of the conditional distribution of X given Y.

Optimal prediction of X_t

Suppose $EX_t^2 < \infty$

Conditional expectation =

the best approximation of X_t as a function of the past.

$E(X_t \mid X_u, u < t)$ is the **random variable** which

- is a function of the past variables $X_{t-1}, X_{t-2}, \ldots$
- minimizes (in Z) the expectation $E(X_t - Z)^2$

We have the **characterization**

$$X_t = E(X_t \mid X_u, u < t) + \epsilon_t,$$

where $E(\epsilon_t \mid X_u, u < t) = 0.$

Optimal linear prediction of X_t

Suppose $EX_t^2 < \infty$ $(d = 1)$

Conditional linear expectation =

the best approximation of X_t as a **linear** function of the past.

$EL(X_t \mid X_u, u < t)$ is the **random variable** which

- is a **linear** function of the past variables $X_{t-1}, X_{t-2}, \ldots$
- minimizes (in Z) the expectation $E(X_t - Z)^2$

We have the **characterization**

$$X_t = EL(X_t \mid X_u, u < t) + \epsilon_t,$$

where $E(\epsilon_t) = 0$ and $\mathrm{Cov}(\epsilon_t, X_u) = 0$ for $u < 0$.

1 Introduction

2 Basic time series models
- Examples of stationary processes
- Examples of non stationary processes

3 Estimating the 1st and 2nd order moments

Moving Average of order 1: MA(1)

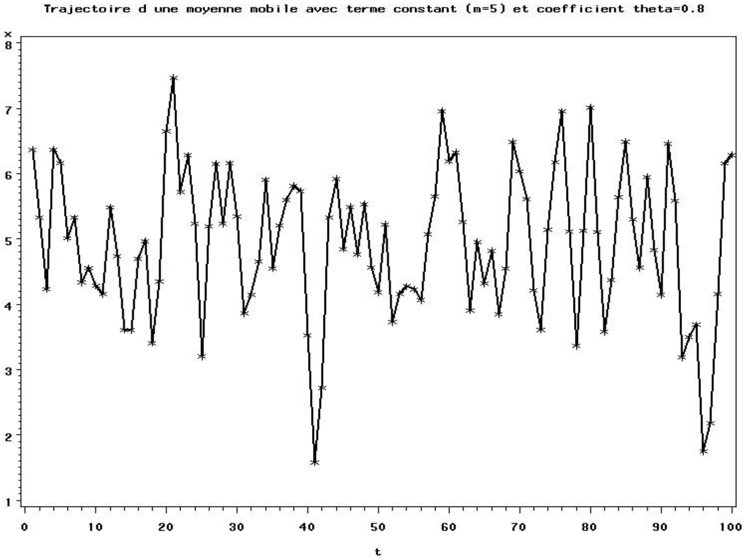
$$X_t = m + \epsilon_t + \theta \epsilon_{t-1}, \quad (\epsilon_t) \sim WN(0, \sigma^2), \quad m, \theta \in \mathbb{R}.$$

$$E(X_t) = m, \qquad V(X_t) = \sigma^2 (1 + \theta^2)$$

$$\gamma_X(h) = \begin{cases} \sigma^2(1+\theta^2), & h = 0 \\ \sigma^2\theta, & h = \pm 1 \\ 0, & |h| > 1. \end{cases}$$

$\Rightarrow (X_t)$ is stationary

$$\rho_X(h) = \begin{cases} 1, & h = 0 \\ \theta/(1+\theta^2), & h = \pm 1 \\ 0, & |h| > 1. \end{cases}$$


Trajectoire d une moyenne mobile avec terme constant (m=5) et coefficient theta=0.8
x
8
7
6
5
4
3
2
1
0
10
20
30
40
50
60
70
80
90
100
t

First-order autoregressive: AR(1)

Let $(\epsilon_t) \sim WN(0, \sigma^2)$ et $\phi \in \mathbb{R}$.

Is there a (stationary) process (X_t) satisfying the autoregressive equation

$$X_t = \phi X_{t-1} + \epsilon_t?$$

Starting from an initial value X_0, it is easy to construct a process - generally non stationary - satisfying the AR(1) equation for all $t \geq 1$.

Conditions are needed to ensure the existence of solutions for all $t \in \mathbb{Z}$.

Causal stationary solution of $X_t = \phi X_{t-1} + \epsilon_t$

If $|\phi| < 1$ then $X_t(N) := \sum_{i=0}^{N} \phi^i \epsilon_{t-i}$ satisfies

$$X_t(N) = \phi X_{t-1}(N-1) + \epsilon_t.$$

Moreover $X_t = \lim_{N\to\infty} X_t(N)$ exists with probability 1 because

$$E \sum_{i=0}^{\infty} |\phi|^i |\epsilon_{t-i}| = E|\epsilon_t| \frac{1}{1-|\phi|} < \infty.$$

Hence, with probability 1, the series

$$X_t = \sum_{i=0}^{\infty} \phi^i \epsilon_{t-i}$$

is well defined, and it is a stationary solution to the AR(1) model.

Stationary anticipative solution of $X_t = \phi X_{t-1} + \epsilon_t$

If $|\phi| > 1$ then

$$X_t = -\frac{1}{\phi}\epsilon_{t+1} + \frac{1}{\phi}X_{t+1} = -\sum_{i=1}^{\infty}\frac{1}{\phi^i}\epsilon_{t+i}$$

is called the anticipative (or non causal) solution of the AR(1) model.

This solution is 2nd-order stationary (and also strictly if the WN is strong).

No stationary solution

If $|\phi| = 1$ et $\sigma^2 > 0$ then, there exists **no stationary solution**

Indeed, it (X_t) was a stationary solution,

$$\mathrm{Var}(X_t - \phi^N X_{t-N}) = 2\,\{\mathrm{Var}(X_t) \pm \mathrm{Cov}(X_t, X_{t-N})\}$$

would be bounded. But

$$X_t - \phi^N X_{t-N} = \sum_{i=0}^{N-1} \phi^i \epsilon_{t-i}$$

$$\Rightarrow \quad \mathrm{Var}(X_t - \phi^N X_{t-N}) = \sum_{i=0}^{N-1} \phi^{2i} \mathrm{Var}\,\epsilon_{t-i} = N\sigma^2 \to \infty,$$

which leads to a contradiction.

Autocorrelations of the causal AR(1)

Let (X_t) be the solution of

$$X_t = \phi X_{t-1} + \epsilon_t, \quad (\epsilon_t) \sim WN(0, \sigma^2), \; |\phi| < 1$$

where ϵ_t is non correlated with the X_{t-i}, $i > 0$.

$$E(X_t) = 0, \qquad \gamma_X(h) = \phi\gamma_X(h-1) + \mathrm{Cov}(\epsilon_t, X_{t-h}).$$

Thus

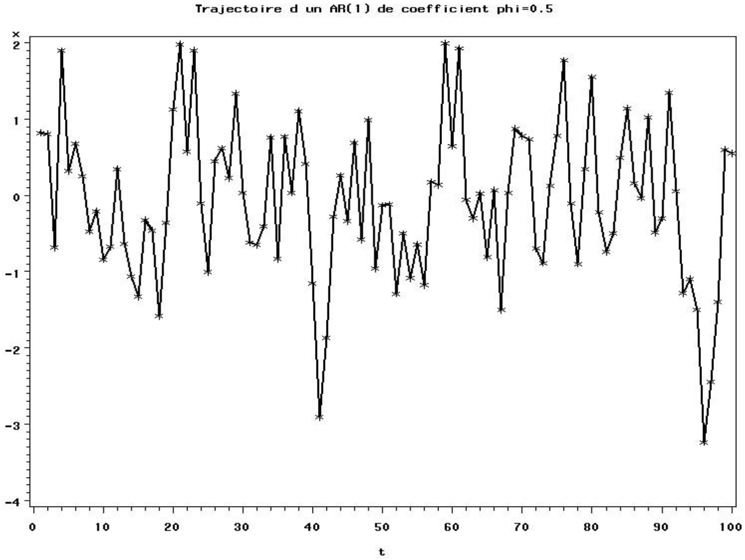
$$\gamma_X(h) = \phi\gamma_X(h-1) = \phi^h\gamma_X(0), \quad h > 0$$

and, using $\gamma_X(h) = \gamma_X(-h)$,

$$\gamma_X(0) = \phi\gamma_X(-1) + \sigma^2 = \phi^2\gamma_X(0) + \sigma^2 = \frac{\sigma^2}{1-\phi^2}$$

Autocorrelation function:

$$\rho_X(h) = \phi^{|h|}$$


Trajectoire d un AR(1) de coefficient phi=0.5
x
2
1
0
-1
-2
-3
-4
0
10
20
30
40
50
60
70
80
90
100
t

ARMA models

MA(1) and AR(1) are particular ARMA(p,q) models:

$$\begin{cases} X_t - \phi_1 X_{t-1} - \cdots - \phi_p X_{t-p} = \epsilon_t - \psi_1 \epsilon_{t-1} - \cdots - \psi_q \epsilon_{t-q} \\ (\epsilon_t) \text{ weak white noise} \end{cases}$$

which will be studied in the next chapter.

1 Introduction

2 Basic time series models
- Examples of stationary processes
- Examples of non stationary processes

3 Estimating the 1st and 2nd order moments

Models with trends

Many time series sample paths display a trend:

$$X_t = m_t + Y_t$$

where m_t is a non constant function of time, called trend, and (Y_t) is a centred time series.

$$E(X_t) = m_t \Rightarrow (X_t) \quad \text{is not stationary}$$

Examples:

- linear trend: $m_t = a_0 + a_1 t,\ a_1 \neq 0$
- quadratic trend: $m_t = a_0 + a_1 t + a_2 t^2,\ a_2 \neq 0$

The coefficients a_0, a_1, a_2 can be estimated by least-squares:

$$\min_{a_0, a_1, a_2} \sum_{t=1}^{n} (x_t - m_t)^2 \quad \Longrightarrow \quad \hat{a}_0, \hat{a}_1, \hat{a}_2$$

Models with tend and seasonality

Many series also display seasonal features, generally linked to the seasons or the economic activity (Ex: decrease in consumption of certain goods in August etc)..

It is natural to complete the model with trend as

$$X_t = m_t + s_t + Y_t$$

where s_t is a periodic function of time, called seasonality:

$$s_1, \dots, s_d \text{ and } s_t = s_{t-d} \text{ for } t > d.$$

Random walk

Let $(\epsilon_t) \sim WN(0, \sigma^2)$. The random walk is defined by

$$X_0 = 0, \quad X_t = X_{t-1} + \epsilon_t, \quad t > 0.$$

Thus

$$X_t = \epsilon_t + \cdots + \epsilon_1, \quad t > 0.$$

Thus $EX_t = 0$ but

$$EX_t^2 = E\epsilon_t^2 + \cdots + E\epsilon_1^2 = t\sigma^2.$$

Therefore the random walk is not stationary.

Random walk with trend

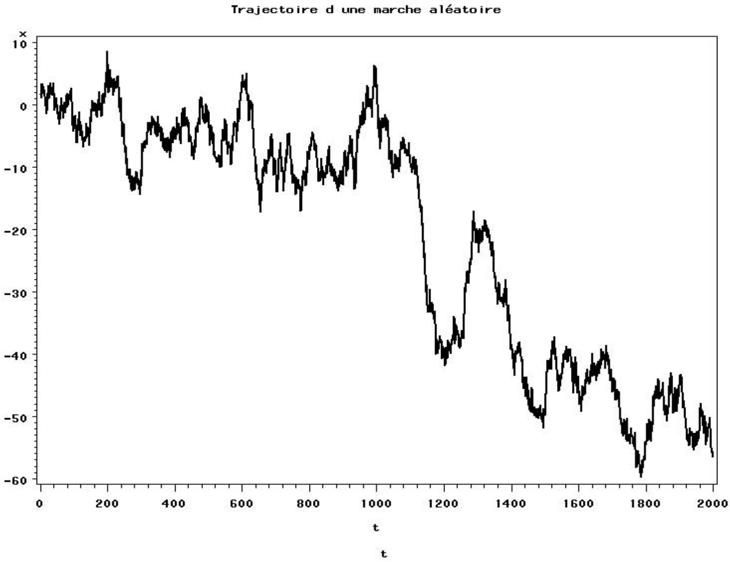
Let $(\epsilon_t) \sim WN(0, \sigma^2)$. The random walk with trend is defined by

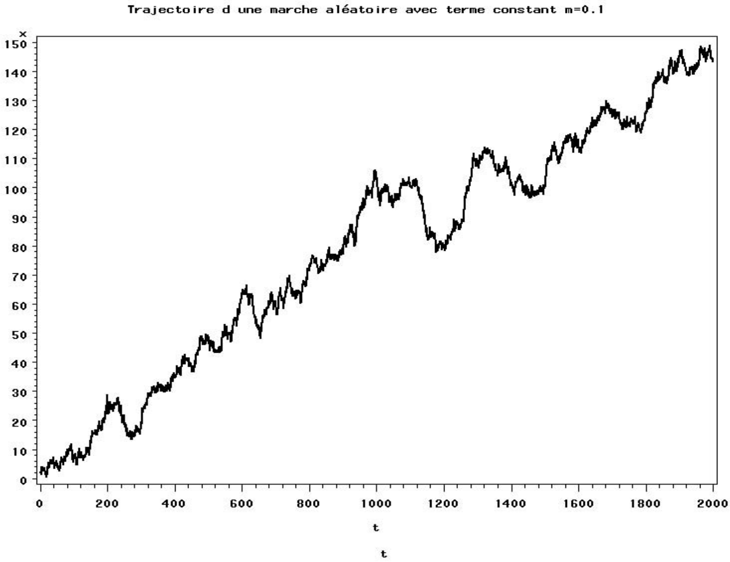
$$X_0 = a, \quad X_t = X_{t-1} + b + \epsilon_t, \quad t > 0.$$

Hence

$$X_t = \epsilon_t + \cdots + \epsilon_1 + bt + a, \quad t > 0.$$

The component $a + bt$ is the (linear) deterministic trend and $\epsilon_1 + \cdots + \epsilon_t$ is a stochastic trend.


Trajectoire d une marche aléatoire
x
10
0
-10
-20
-30
-40
-50
-60
0
200
400
600
800
1000
1200
1400
1600
1800
2000
t
t


Trajectoire d une marche aléatoire avec terme constant m=0.1
x
150
140
130
120
110
100
90
80
70
60
50
40
30
20
10
0
0
200
400
600
800
1000
1200
1400
1600
1800
2000
t
t

The Box and Jenkins (1976) approach

Non stationary time series can **often** be made stationary by applying repetitively the difference operator

$$\begin{aligned} \Delta X_t &= X_t - X_{t-1}, \\ \Delta^2 X_t &= \Delta X_t - \Delta X_{t-1} = X_t - 2X_{t-1} + X_{t-2}, \qquad \text{etc.} \end{aligned}$$

or the seasonal difference operator

$$\Delta_s X_t = X_t - X_{t-s}.$$

Examples:

$$\begin{aligned} X_t &= a_0 + a_1 t + Y_t &&\Longrightarrow \quad \Delta X_t = a_1 + Y_t - Y_{t-1} \\ X_t &= a_0 + a_1 t + a_2 t^2 + Y_t &&\Longrightarrow \quad \Delta^2 X_t = 2a_2 + \Delta^2 Y_t \\ X_t &= a_0 + a_1 t + s_t + Y_t &&\Longrightarrow \quad \Delta_s X_t = a_1 s + Y_t - Y_{t-s} \end{aligned}$$

1 Introduction

2 Basic time series models

3 Estimating the 1st and 2nd order moments
- Ergodicity
- Empirical mean
- Empirical autocorrelations

Estimation of moments

The theoretical moments can be estimated from the observations $X_1, \ldots, X_n$. A natural estimator for the expectation EX_1 is the empirical mean

$$\overline{X}_n = \frac{1}{n} \sum_{t=1}^{n} X_t.$$

As the observations are not iid, the usual Law of Large Numbers (LLN) does not apply. Instead, we will rely on the concept of ergodicity.

A stationary sequence is called ergodic if it satisfies the strong LLN.

Ergodic stationary process

A stationary sequence is ergodic if it satisfies the strong LLN (even if the usual conditions are not satisfyed).

Definition

A strictly stationary process $(Z_t)_{t\in\mathbb{Z}}$, valued in $\mathbb{R}^d$, is called ergodic if for any integer k, and any Borel set B of $\mathbb{R}^{dk}$,

$$n^{-1}\sum_{t=1}^{n} I_B(Z_t, Z_{t+1}, \ldots, Z_{t+k-1}) \to P\{(Z_1, \ldots, Z_{1+k}) \in B\}\ a.s.$$

Ergodic Theorem

Any fixed transformation of a stationary ergodic sequence is also stationary and ergodic.

Theorem

If $(Z_t)_{t\in\mathbb{Z}}$ is a strictly stationary and ergodic sequence and if $(Y_t)_{t\in\mathbb{Z}}$ is defined by

$$Y_t = f(Z_t, Z_{t+1}, \ldots; Z_{t-1}, Z_{t-2}, \ldots),$$

then $(Y_t)_{t\in\mathbb{Z}}$ is strictly stationary and ergodic.

Examples and counter-examples of ergodic processes

- A strong WN (ϵ_t) is stationary and ergodic.
- A MA(q)

$$X_t = \sum_{i=0}^{q} c_i \epsilon_{t-i}$$

  or the *causal* solution of an AR(1)

$$X_t = aX_{t-1} + \epsilon_t, \qquad |a| < 1,$$

  are stationary and ergodic.
- The process defined by

$$X_t = X, \quad \forall t,$$

  where X is a non-degenerate r.v. is stationary but is not ergodic.

1 Introduction

2 Basic time series models

3 Estimating the 1st and 2nd order moments
- Ergodicity
- Empirical mean
- Empirical autocorrelations

Weak ergodicity for the mean

If (X_t) is a univariate, 2nd-order stationary process, whose autocovariance function satisfies $\gamma(h) \to 0$ when $h \to \infty$, then

$$\overline{X}_n := \frac{1}{n}\sum_{t=1}^{n} X_t \quad \xrightarrow{L^2} \quad EX_1 \quad \text{as } n \to \infty.$$

Proof: Using the Cesàro lemma

$$E\left(\overline{X}_n - EX_1\right)^2 = \frac{1}{n^2}\sum_{t,s=1}^{n} \mathrm{Cov}(X_t, X_s)$$

$$= \frac{1}{n}\sum_{|h|<n}\left(1 - \frac{|h|}{n}\right)\gamma(h)$$

$$\leq \frac{1}{n}\sum_{|h|<n}\left|\gamma(h)\right| \to 0.$$

1 Introduction

2 Basic time series models

3 Estimating the 1st and 2nd order moments
- Ergodicity
- Empirical mean
- Empirical autocorrelations

Empirical autocorrelations

Definition (univariate case)

The empirical autocovariance function is, for $|h| < n$,

$$\hat{\gamma}(h) = \frac{1}{n} \sum_{t=|h|+1}^{n} (X_t - \overline{X}_n)(X_{t-|h|} - \overline{X}_n), \quad \overline{X}_n = \frac{1}{n} \sum_{t=1}^{n} X_t.$$

The empirical autocorrelation function is

$$\hat{\rho}(h) = \frac{\hat{\gamma}(h)}{\hat{\gamma}(0)}, \quad |h| < n.$$

Empirical autocorrelations of a WN

If the observations are the realizations of a strong WN,

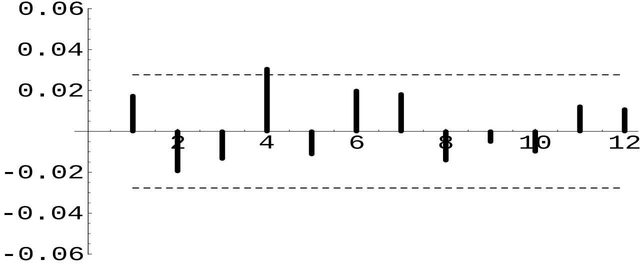
- $\overline{X}_n \to 0, \quad \hat{\gamma}(0) \to \sigma^2$ (LLN) and, if $h \neq 0, \quad \hat{\gamma}(h) \to 0$ a.s.
- $\sqrt{n}\,\overline{X}_n \xrightarrow{\mathcal{L}} \mathcal{N}(0, \sigma^2)$,

  for $h \neq 0$, $\sqrt{n}\hat{\gamma}(h) \xrightarrow{\mathcal{L}} \mathcal{N}(0, \sigma^4)$

  for $h \neq 0$, $\sqrt{n}\hat{\rho}(h) \xrightarrow{\mathcal{L}} \mathcal{N}(0, 1)$ ("almost" usual CLT + Slutsky).

Empirical autocorrelogram of a WN

For a strong WN, $\hat{\rho}(h)$ belongs to the interval $\pm 1.96/\sqrt{n}$ in dotted lines with a probability $\approx$ 95%.



Empirical autocorrelations of a strong WN, for $n = 5000$.

Empirical autocorrelations of a stationary process

For a stationary (strict and 2nd-order) and ergodic process,

$$\overline{X}_n \to EX_1, \qquad \hat{\gamma}(0) \to \mathsf{V}(X_1)$$

and, if $h \neq 0$,

$$\hat{\gamma}(h) \to \gamma(h) \text{ a.s. (ergodic theorem).}$$

For ARMA-type processes (see chapter 7 in Brockwell Davis)

$$\sqrt{n}\{\hat{\rho}(h) - \rho(h)\} \xrightarrow{\mathcal{L}} \mathcal{N}(0, \sigma_h^2)$$

where σ_h^2 only depends on $\rho(\cdot)$ (Bartlett's formula).
End of Chapter 1

Linear transformation of a stationary series

Proposition

Let (Z_t) a stationary (2nd order) time series, $(a_i)_{i \in Z}$ a sequence of coefficients such that e $\sum_{i=-\infty}^{\infty} |a_i| < \infty$. Then the sum

$$X_t = \sum_{i=-\infty}^{\infty} a_i Z_{t-i}$$

converges in mean-square and almost surely. Moreover, (X_t) is stationary.

Proof of the convergence

- By Fubini,

$$E\sum_{i=-\infty}^{\infty}|a_i Z_{t-i}|<\infty,$$

  hence the series is finite with probability 1.

- For $0<p<q$,

$$\left\|\sum_{p\le|i|\le q}a_i Z_{t-i}\right\|_2\le\sum_{p\le|i|\le q}|a_i|\,\|Z_{t-i}\|_2\to 0$$

  when $p,q\to\infty$, thus the series converges in mean square by the Cauchy criterion.

Proof of stationarity

The inversions of E and $\sum$, by Lebesgue and Fubini, give

$$EX_t = a(B)EZ_t = a(1)EZ_1 = \sum_{i=-\infty}^{\infty} a_i EZ_1$$

and

$$E(X_t - EX_1)(X_{t-h} - EX_1) = \sum_{i=-\infty}^{\infty} \sum_{j=-\infty}^{\infty} a_i a_j E(Z_{t-i} - EZ_1)(Z_{t-j-h} - EZ_1).$$

Return