MAP 565
Time series analysis : Lecture VII

François Roueff
http://perso.telecom-paristech.fr/~roueff/

Telecom ParisTech – École Polytechnique

January 27, 2016

Outline of the course

- ▷ Stochastic modeling
  - I Random processes.
  - II Spectral representation.
- ▷ Linear models
  - III Linear filtering, innovation process.
  - IV ARMA processes.
  - V Linear forecasting.
- ▷ Statistical inference
  - VI Overview of goals and methods.
  - VII Asymptotic statistics in a dependent context. ←
- ▷ Non-linear models
  - VIII Standard models for financial time series.
  - IX Complements.

Outline of lecture VII

1. Asymptotic statistics for time series
   - Basic definitions
   - Consistency
   - Asymptotic normality
2. An illustrative example with R

1. Asymptotic statistics for time series
   - Basic definitions
   - Consistency
   - Asymptotic normality
2. An illustrative example with R

Convergence of random variables

Let $(W_n)_{n\geq 1}$ and W be real valued random variables defined on the same probability space.

We denote

(i) $W_n \xrightarrow{\text{a.s.}} W$ if W_n converges to W almost surely.

(ii) $W_n \xrightarrow{P} W$ or $W_n = W + o_P(1)$ if W_n converges to W in probability.

(iii) $W_n \Longrightarrow W$ if W_n converges weakly to W.

(iv) $W_n = O_p(1)$ if W_n is bounded in probability.

Recall that

$$(i) \Rightarrow (ii) \Rightarrow (iii) \Rightarrow (iv)$$

and

$$\lim_{n\to\infty} W_n = W \text{ in } L^2 \Rightarrow (ii) \; .$$

Extension to more general spaces

All these definitions can be extended to random variables valued in a general metric space (introducing the notion of tight sequences in place of bounded in probability).

- Finite dimensional space : In $\mathbb{R}^p$, the convergence of a random vector is equivalent to the convergence of its entries, except for the weak convergence : the weak convergence of a vector is equivalent to the weak convergence of all linear combinations (this is called the Cramér-Wold device).
- Infinite dimensional space : In $\mathbb{R}^T$ for an infinite T, the fidi weak convergence is defined by

$$W_n \overset{\text{fidi}}{\Longrightarrow} W \quad \text{if} \quad \left[W_n(t)\right]_{t \in I} \Longrightarrow \left[W(t)\right]_{t \in I} \quad \text{for all finite } I \subset T \; .$$

Consistency and asymptotic normality

Let θ be an unknown parameter of a stationary model on a time series $(X_t)_{t\in\mathbb{Z}}$ and let $\widehat{\theta}_n$ be an estimator based on the n-sample $X_{1:n}$.

Definition : weak consistency

$\widehat{\theta}_n$ is said to be weakly consistent if $\widehat{\theta}_n \xrightarrow{P} \theta$.

Definition : strong consistency

$\widehat{\theta}_n$ is said to be strongly consistent if $\widehat{\theta}_n \xrightarrow{\text{a.s.}} \theta$.

Definition : asymptotic normality

$\widehat{\theta}_n$ is said to be asymptotically normal with asymptotic covariance matrix Σ if

$$\sqrt{n}(\widehat{\theta}_n - \theta) \Longrightarrow \mathcal{N}(0, \Sigma) \ .$$

Classical examples in the i.i.d. setting

Let $(X_t)_{t\in\mathbb{Z}}$ be an i.i.d. sequence with mean μ and variance σ^2 and let $\widehat{\mu}_n$ denote the empirical mean estimator.

Theorem : Law of large numbers (LLN)

If μ is finite, then $\widehat{\mu}_n$ is a strongly consistent estimator of μ,

$$\widehat{\mu}_n \xrightarrow{\text{a.s.}} \mu \;.$$

Theorem : Central limit theorem (CLT)

If σ^2 is finite, then $\widehat{\mu}_n$ is an asymptotically normal estimator of μ with asymptotic variance σ^2.

$$\sqrt{n}(\widehat{\mu}_n - \mu) \Longrightarrow \mathcal{N}(0, \sigma^2) \;.$$

Consistency of the empirical mean

Assumption : weak stationarity

Let (X_t) be a real-valued weakly stationary process with mean μ and autocovariance function γ/spectral density f.

Theorem

Then the empirical mean $\widehat{\mu}_n$ satisfies the following assertions.

- $\widehat{\mu}_n$ is an unbiased estimator of μ ($\mathbb{E}\left[\widehat{\mu}_n\right] = \mu$ for all $n \geq 1$).
- If $\lim_{h\to\infty} \gamma(h) = 0$, then $\lim_{n\to\infty} \mathbb{E}\left[(\widehat{\mu}_n - \mu)^2\right] = 0$ and $\widehat{\mu}_n$ is a weakly consistent estimator of μ.
- If moreover $\gamma \in \ell^1$, then, as $n \to \infty$,

$$\mathrm{Var}\left(\widehat{\mu}_n\right) \leq n^{-1} \|\gamma\|_1 ,$$
$$\mathrm{Var}\left(\widehat{\mu}_n\right) = n^{-1}(2\pi f(0) + o(1)) ,$$

  and $\widehat{\mu}_n$ is a strongly consistent estimator of μ.

Consistency of the empirical autocovariance : assumptions

Assumption : linear process with short memory

X admits the representation $X = \mu + \mathrm{F}_\psi(Z)$ where $\mu \in \mathbb{R}$, $Z \sim \mathrm{WN}(0, \sigma^2)$ is real valued and $(\psi_t)_{t\in\mathbb{Z}} \in \ell^1$ is also real valued.

Then X is weakly stationary with mean μ and autocovariance and spectral density given by

$$\gamma(t) = \sigma^2 \sum_{k\in\mathbb{Z}} \psi_{k+h}\psi_k \quad \text{and} \quad f(\lambda) = \frac{\sigma^2}{2\pi} \left| \sum_{k\in\mathbb{Z}} \psi_k \mathrm{e}^{-\mathrm{i}k\lambda} \right|^2 .$$

Assumption on the noise

Moreover Z satisfies, for a constant $\eta \geq 1$, for all $s \leq t \leq u \leq v$,

$$\mathbb{E}[Z_s\, Z_t\, Z_u\, Z_v] = \begin{cases} \eta\sigma^4 & \text{if } s = t = u = v, \\ \sigma^4 & \text{if } s = t < u = v, \\ 0 & \text{otherwise.} \end{cases}$$

Consistency of the empirical autocovariance : conclusions

Theorem

Then, for all $p, q \in \mathbb{Z}$,

$$\mathbb{E}\left[\widehat{\gamma}_n(p)\right] = \gamma(p) + O(n^{-1}),$$
$$\lim_{n\to\infty} n\,\mathrm{Cov}\left(\widehat{\gamma}_n(p), \widehat{\gamma}_n(q)\right) = V(p,q),$$

where

$$V(p,q) = (\eta - 3)\gamma(p)\gamma(q) + \sum_{u\in\mathbb{Z}} \left[\gamma(u)\gamma(u-p+q) + \gamma(u+q)\gamma(u-p)\right]. \quad (1)$$

Corollary

We have $\widehat{\gamma}_n(p) = \gamma(p) + O_P(n^{-1/2})$ and $\widehat{\gamma}_n(p)$ is a weakly consistent estimator of $\gamma(p)$,

Asymptotic normality : basic assumption

We use the same kind of assumption as for consistency of the empirical autocovariance function.

Assumption : linear process with short memory

X admits the representation $X = \mu + \mathrm{F}_\psi(Z)$ where $\mu \in \mathbb{R}$, $Z \sim \mathrm{WN}(0, \sigma^2)$ is real valued and $(\psi_t)_{t\in\mathbb{Z}} \in \ell^1$ is also real valued.

Recall that it implies that X is weakly stationary with mean μ and autocovariance and spectral density given by

$$\gamma(t) = \sigma^2 \sum_{k\in\mathbb{Z}} \psi_{k+h}\psi_k \quad \text{and} \quad f(\lambda) = \frac{\sigma^2}{2\pi} \left| \sum_{k\in\mathbb{Z}} \psi_k \mathrm{e}^{-\mathrm{i}k\lambda} \right|^2 .$$

Additional assumptions will be made on Z.

Asymptotic normality of the empirical mean

Assumption

Let X be a linear process with short memory as above and suppose that Z satisfies

$$n^{-1/2}\sum_{t=1}^{n} Z_t \Longrightarrow \mathcal{N}(0,\sigma^2)\;.$$

Theorem

Then the empirical mean is asymptotically normal with asymptotic variance $2\pi f(0)$,

$$\sqrt{n}(\widehat{\mu}_n-\mu) \Longrightarrow \mathcal{N}(0,2\pi f(0))\;.$$

Remark: note that the general result follows from the case where $\mu=0$, which we assume in the following.

Proof, key ingredient : approximation lemma

$$
\begin{array}{ccc}
X_n & \overset{?}{\Longrightarrow} \;\Longrightarrow & W \\
\Uparrow\Uparrow & & \Uparrow \\
W_{n,m} & \Longrightarrow & W_m
\end{array}
$$

Lemma

Let $(W_{n,m})_{n,m\geq 1}$ be an array of random variables in $\mathcal{X}$. Suppose that for all $m \geq 1$, $W_{n,m}$ converges weakly to W_m as $n \to \infty$ and that W_m converges weakly to W as $m \to \infty$. Let now $(X_n)_{n\geq 1}$ be random variables in $\mathcal{X}$ such that, for all $\epsilon > 0$,

$$\lim_{m\to\infty} \limsup_{n\to\infty} \mathbb{P}(d(X_n, W_{m,n}) > \epsilon) = 0 \ .$$

Then X_n converges weakly to W as $n \to \infty$.

Proof, Step 1 : FIR approximation

FIR approximation

Consider a linear process $X = \mathrm{F}_{\psi}(Z)$. An approximation of X is obtained by setting

$$X^{(m)} = \mathrm{F}_{\psi^m}(Z) \;,$$

where $\psi^m_k = \psi_k \mathbb{1}_{\{|k| \leq m\}}$.

In order to use the approximation lemma, we must show that

$$\lim_{m \to \infty} \limsup_{n \to \infty} \mathbb{P}(\left|\sqrt{n}\left(\widehat{\mu}_n - \widehat{\mu}_n^{(m)}\right)\right| > \epsilon) = 0 \;,$$

where

$$\widehat{\mu}_n^{(m)} = \frac{1}{n} \sum_{k=1}^{n} X_k^{(m)} \;.$$

Proof, Step 2 : the FIR case

To conclude, it only remains to show that (remember that $\mu = 0$ here), for all $m \geq 1$, as $n \to \infty$,

$$\sqrt{n}\widehat{\mu}_n^{(m)} = \frac{1}{\sqrt{n}} \sum_{k=1}^{n} \sum_{t=-m}^{m} \psi_t Z_{k-t} \Longrightarrow \mathcal{N}(0, 2\pi f^{(m)}(0)) \ , \tag{2}$$

with $f^{(m)}$ denoting the spectral density of $X^{(m)}$, and that, as $m \to \infty$,

$$\mathcal{N}(0, 2\pi f^{(m)}(0)) \Longrightarrow \mathcal{N}(0, 2\pi f(0)) \ . \tag{3}$$

- ▷ Eq. (2) is obtained by using that, for all $t \in \mathbb{Z}$, as $n \to \infty$,

$$\sum_{k=1}^{n} Z_{k-t} = \sum_{k=1}^{n} Z_k + O_P(1) \ .$$

- ▷ Eq. (3) is equivalent to $\lim_{m \to \infty} f^{(m)}(0) = f(0)$.

Asymptotic normality of the empirical autocovariance

Assumption

Let X be a linear process with short memory as above and suppose that Z is i.i.d. with $\mathbb{E}\left[Z_0^4\right] = \eta\sigma^4$.

Theorem

Then the empirical autocovariance is asymptotically normal with asymptotic covariance V,

$$\sqrt{n}\left(\widehat{\gamma}_n - \gamma\right) \overset{\text{fidi}}{\Longrightarrow} \mathcal{N}(0, V) ,$$

where
$$V(p,q) = (\eta - 3)\gamma(p)\gamma(q) + \sum_{u \in \mathbb{Z}} \left[\gamma(u)\gamma(u-p+q) + \gamma(u+q)\gamma(u-p)\right] .$$

Remark: note that we can again take $\mu = 0$ without loss of generality.

Proof, Step 1 : FIR approximation, again

The proof follows the same path but under the stronger assumption used in this therem, Step 1 leads to a stronger consequence.

Definition : m-dependent processes

A process $(W_t)_{t\in\mathbb{Z}}$ is said to be m-dependent if for all $t \in \mathbb{Z}$, $(W_s)_{s\leq t}$ is independent of $(W_s)_{s>t+m}$.

Consequence of FIR approximation

If $Z \sim \mathrm{IID}(0, \sigma^2)$, then $X^{(m)}$ is a $(2m)$-dependent approximation of X.

Proof, Step 2 : independent blocks approximations

An m-dependent sequence can be approximated by a sequence of independent blocks :

$$W_1, \dots, W_k, W_{k+1}, \dots, W_{k+m}, W_{k+m+1}, \dots, W_{2k+m}, \dots\dots,$$

$$\underbrace{W_1, \dots, W_k}, \cancel{W_{k+1}, \dots, W_{k+m}}, \underbrace{W_{k+m+1}, \dots, W_{2k+m}}, \dots\dots,$$

Consequence

Show that the approximation lemma can be applied and use the usual central limit theorem on the sequence of independent blocks.

```
###############################################
#          Examples stat. inference           #
#             ARMA FORECASTING                #
###############################################
plotarimapred <- function(x,myorder=c(0,0,0),ratio=0.9){
  # select a part for arma modeling and then forecast
  dur <- end(x)[1]-start(x)[1]
  y <- window(x,start=start(x)[1],end=start(x)[1]+dur*ratio)
  yf <-  window(x,start=start(x)[1]+dur*ratio,end=end(x)[1])
  # fit an ARMA(1,1) model (with mean)
  est <- arima(y,order=myorder)
  print(est)
  yp <- predict(est,n.ahead=length(x)-floor(length(x)*ratio))
  ts.plot(y,yf,yp$pred,yp$pred+1.96*yp$se,yp$pred-1.96*yp$se,
          col=c(1,2,3,4,4))
}
# SOI Southern Oscillation Index
# pressure differences Tahiti - Darwin (Pacific Ocean)
#(ne. -> El Nino/ pos. -> La nina)
x <- t(read.table('~/data/dataset/climate/soi.tsv'))
x <- x[is.finite(x)]
soi <- ts(x,start=1951,frequency=12)
op <- par(mfrow=c(3,1))
acf(soi)
pacf(soi)
ts.plot(soi)
abline(h=0)
par(op)
plotarimapred(soi,myorder=c(0,0,0),ratio=0.9)
plotarimapred(soi,myorder=c(3,0,0),ratio=0.9)
plotarimapred(soi,myorder=c(2,0,2),ratio=0.9)
graphics.off()
# Steel import time series
steeldat <- paste('~/data/dataset/financial/macro/',
                  'steel-sheets-shipments-usa.csv',sep='')
st <-  read.csv(steeldat,header=TRUE,sep=";")
```

```
# this is monthly data
all <- ts(st$Steel.Sheets.Shipments.for.United.States,
          start=1919,frequency=12)
ts.plot(all)
plotarimapred(all,myorder=c(0,0,0),ratio=0.9)
plotarimapred(all,myorder=c(0,1,0),ratio=0.9)
plotarimapred(all,myorder=c(2,1,5),ratio=0.9)
graphics.off()
# Volatility prediction for sp500 index
spdata <- paste('~/data/dataset/financial/sp500/',
                'sp500-1950--2010.csv',sep='')
sp <- read.csv(spdata,header=TRUE,sep=",")
attach(sp)
plot(as.POSIXct(Date),Open, type='l',xlab='Date',ylab="SP500 index")
# order dates and select post 2009
ind <- order(Date[as.POSIXct(Date) >
                  as.POSIXct("2009-01-01")])
r <- ts(diff(log(Open[ind])))
sd <- Date[ind[2:length(ind)]]
plot(as.POSIXct(sd),r, type='l',xlab='Date',ylab="SP500 index log returns")
# correlations for log returns and squared log returns
nn <- 15
op <- par(mfrow=c(2,1))
acf(r, lag.max=nn, main='Log-returns autocor')
par(op)
# Squarred log-returns
op <- par(mfrow=c(2,1))
acf(r**2,lag.max=nn)
pacf(r**2,lag.max=nn)
par(op)
plotarimapred(r**2,myorder=c(1,0,1))
```