

# Économétrie des séries temporelles

Thomas Chuffart

[thomas.chuffart@univ-fcomte.fr](mailto:thomas.chuffart@univ-fcomte.fr)

# Informations générales

- 15h CM + 3h TD
- des heures de forecasts entre des séances
- Évaluation : Devoir sur table + projet
- Pré-requis : Économétrie en coupe transversale

# Introduction

## Definition

Une série temporelle est une suite de variables stochastiques  $\{y_1, y_2, \dots, y_t\}$  indicé par  $t = 1, \dots, T$ , le temps. Cette série représente un vecteur de variables observées de façon régulières.

## Remarques :

- Il existe une dépendance temporelle entre les observations : la distribution conditionnelle de  $y_t$  dépend des valeurs passées.
- $\{y_t\}_{t=-\infty}^{+\infty} = \{y_{-\infty}, \dots, 0, y_{+\infty}\}$
- Si l'on connaît le PGD, il est possible d'obtenir  $I$  collections de réalisations de  $\{y_t\}_{t=-\infty}^{+\infty}$

# Introduction

- L'économètre dispose d'une seule collection de réalisations de  $T$  v.a
- Il faut choisir (ou se restreindre) suivant la disponibilité des données, à la meilleure périodicité possible.
- On cherche à identifier le meilleur processus stochastique. On pourra alors :
  - effectuer des prévisions
  - réaliser des simulations à des fins statistiques

# CPI and GDP: get data from FRED Database

```
library(tidyverse)
library(data.table)
library(viridis)
library(tseries)
library(quantmod)

my.names <- data.table(var=c("CPILFESL","GDPC1"), name=c("CPI","GDP"))
df= getSymbols('CPILFESL',src='FRED', auto.assign=F)
df = data.frame(date=time(df), coredata(df))
df.gdp_us =getSymbols('GDPC1',src='FRED', auto.assign=F)
df.gdp_us = data.frame(date=time(df.gdp_us), coredata(df.gdp_us))

df<-merge(df,df.gdp_us,by="date")
dt<-data.table(df)
dt %>% gather(var,value,-date) %>% data.table() -> dt2
dt2<-merge(dt2,my.names,by="var")

# Base 100 pour la date t = 1

dt2=dt2[,id:=1:.N, by=var]
dt2=dt2[,var0:=100*value/sum(ifelse(id==1,value,0)),by=var]
```

# CPI et GDP Figure

```
ggplot(data=dt2,aes(x=date,y=var0,color=name,linetype=name))+geom_line(size=1.1)+  
  scale_y_log10(breaks=c(100,200,400,800))+theme_minimal()+  
  theme(plot.caption=element_text(hjust=0),legend.position="top")+guides(linetype=F)+  
  scale_color_viridis(name="Variable",discrete=T,end=0.8)
```

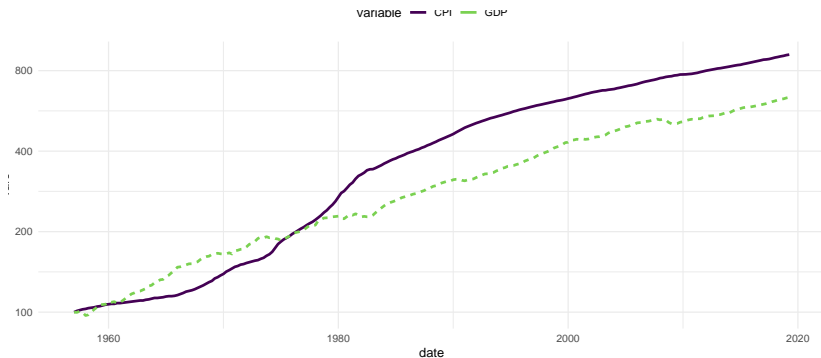


Figure 1: US CPI and GDP

# CPI expectation change et GDP change: get data from FRED

```
rm(list = ls())
my.names <- data.table(var=c("MICH", "A191RL1Q225SBEA"),
                        name=c("Inflation_Expectation", "GDP Change"))
df= getSymbols('MICH',src='FRED', auto.assign=F)
df = data.frame(date=time(df), coredata(df))
df.gdp_us =getSymbols('A191RL1Q225SBEA',src='FRED', auto.assign=F)
df.gdp_us = data.frame(date=time(df.gdp_us), coredata(df.gdp_us))

df<-merge(df,df.gdp_us,by="date")
dt<-data.table(df)
dt %>% gather(var,value,-date) %>% data.table() -> dt2
dt2<-merge(dt2,my.names,by="var")
```

# Michigan inflation expectation et GDP change Figure

```
ggplot(data=dt2,aes(x=date,y=value,color=name,linetype=name))+  
  geom_line(size=1.1)+  
  theme_minimal()+theme(plot.caption=element_text(hjust=0),  
    legend.position="top")+  
  guides(linetype=F)+  
  scale_color_viridis(name="Variable",discrete=T,end=0.8)
```

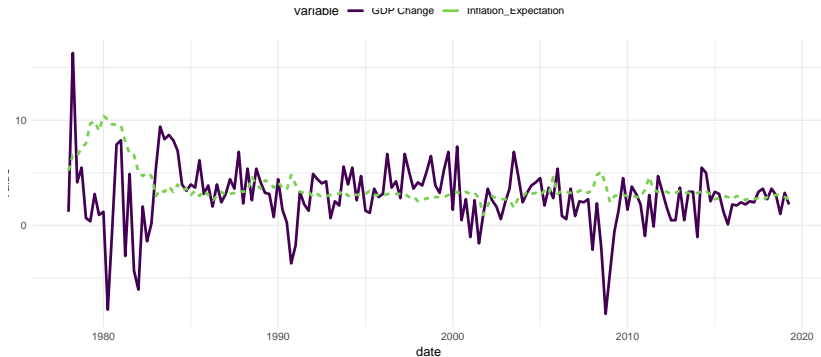


Figure 2: US CPI and GDP change



# Introduction

## Outline

- 1 Stationnarité
- 2 Non-sationnarité
- 3 Dynamiques non-linéaires

# Séries temporelles stationnaires

## Introduction

### Definition

Un processus stochastique, noté  $\{y_t(\omega), \omega \in \Omega, t \in \mathbb{Z}\}$  est une séquence ordonnée de variables aléatoires définies sur un espace probabilisé  $(\Omega, \mathbb{F}, \mathbb{P})$  avec :

- $\Omega$  l'ensemble des possibles
- $\mathbb{F}$  une tribu ou un  $\sigma$ -algèbre représentant les évènements
- $\mathbb{P}$  une mesure de probabilité

Ce processus sera noté  $y_t$ .

# Exchange rate China/US in level and % change: get data

```
rm(list = ls())
# set dates
date.start <- "2006-01-01"
date.end <- "2015-12-31"
my.names <- data.table(var=c("DEXCHUS"),
                        name=c("Exchange rates"))
df= getSymbols('DEXCHUS',src='FRED', auto.assign=F)
df <- df[!(is.na(df$DEXCHUS)),]
df <- df[paste(date.start,date.end,sep="/")]
df.exrate_change <- log(df$DEXCHUS) - log(lag(df$DEXCHUS))
df.exrate_change <- data.frame(date=time(df.exrate_change), coredata(df.exrate_change))

df = data.frame(date=time(df), coredata(df))
dt<-data.table(df)
dt %>% gather(var,value,-date) %>% data.table() -> dt2
dt2<-merge(dt2,my.names,by="var")
```

# Exchange rate China/US in level

```
ggplot(data=dt2,aes(x=date,y=value,color=name,linetype=name))+geom_line(size=1.1)+  
theme_minimal()+theme(plot.caption=element_text(hjust=0),legend.position="top")+  
guides(linetype=F)+scale_color_viridis(name="Variable",discrete=T,end=0.8)
```

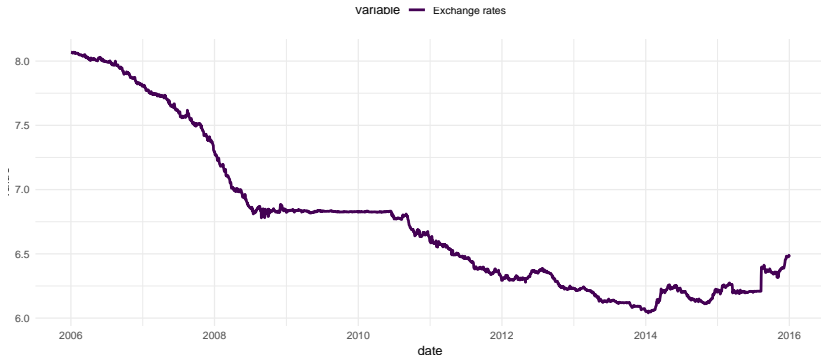


Figure 3: China exchange rates

# Exchange rate China/US % change

```
plot(df.exrate_change, type="l")
```

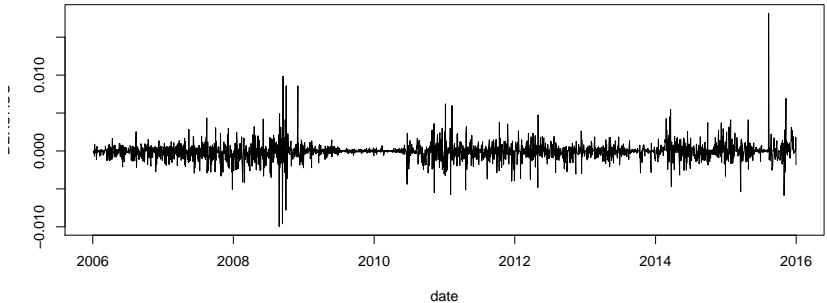


Figure 4: China exchange rates %

# Séries temporelles stationnaires

## Outline

- **Définition de la stationnarité**
- Non-stationnarité
- Rappels
- Spurious regressions
- Stationnariser un processus

# Séries temporelles stationnaires

## Définitions de la stationnarité

### Definition

Les moments inconditionnels ordinaires et centrés de  $y_t$  peuvent s'exprimer comme l'espérance de  $h(y_t)$ , une fonction continue de  $y_t$  :

$$E[h(y_t)] = \int_{t=1}^T h(y_t) f(y_t) dy_t \quad (1)$$

avec  $f(y_t)$  la fonction de densité inconditionnelle de  $y$ .

- Pour l'espérance  $h(y_t) = y_t$
- Pour la variance  $h(y_t) = (y_t - \mathbb{E}[y_t])^2$

# Séries temporelles stationnaires

## Définitions de la stationnarité

### Definition

La fonction d'autocovariance de  $y_t$  s'obtient à partir de la densité jointe de  $(y_t, y_{t+1}, \dots, y_{t+h})$  :

$$\begin{aligned} \text{Cov}(y_t, y_{t+h}) &= \mathbb{E}[(y_t - \mu_t)(y_{t+h} - \mu_{t+h})] \\ &= \int \dots \int (y_t - \mu_t)(y_{t+h} - \mu_{t+h}) \\ &\quad f(y_t, \dots, y_{t+h}) dy_t \dots dy_{t+h} \end{aligned}$$



# Séries temporelles stationnaires

## Définitions de la stationnarité

### Definition (Stationnarité Faible)

Soit  $y_t$  une séquence de v.a,  $y_t$  est (faiblement) stationnaire au second ordre si

- $\forall t \in \mathbb{Z}, \mathbb{E}[y_t] = \mathbb{E}[y_{t-1}] = \dots = \mu < \infty$
- $\forall t, h \in \mathbb{Z}, \text{Cov}(y_t, y_{t+h}) = \gamma(h) < \infty$

avec pour  $h = 0, \gamma(0) = \sigma^2$

# Séries temporelles stationnaires

## Définitions de la stationnarité

### Definition (Stationnarité Forte)

Soit  $y_t$  une séquence de v.a,  $y_t$  est strictement stationnaire si la distribution jointe de  $y_t$  et  $y_{t+h}$  dépend uniquement de  $h$  :

$$f(y_t, \dots, y_{t+h}) = f(y_\tau, \dots, y_{\tau+h}) \quad (2)$$

avec  $t \neq \tau$ . La distribution jointe du processus  $y_t$  doit donc être invariante par translation dans le temps.

# Séries temporelles stationnaires

## Définitions de la stationnarité

### Definition

Soit  $y_t$  un processus stationnaire au second ordre, il est ergodique si

$$\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{h=1}^n \gamma(h) = 0 \quad (3)$$

- La mémoire du processus est finie
- l'ergodicité décrit une forme faible d'indépendance asymptotique

# Séries temporelles stationnaires

## Définitions de la stationnarité

### Theorem (Décomposition de Wold)

*Soit  $\{y_t\}$  un processus stationnaire au second ordre. On peut montrer que  $y_t$  peut toujours se décomposer en une somme pondérée des innovations de  $y_t$  et une composante déterministe  $\mu_t$  :*

$$y_t = \mu_t + \sum_{i=0}^{\infty} \alpha_i \varepsilon_{t-i}, \varepsilon_t \sim IID(0, \sigma_\varepsilon < \infty) \quad (4)$$

avec  $\sum_{i=0}^{\infty} \alpha_i^2 < \infty$

# Séries temporelles stationnaires

## Définitions de la stationnarité

$$\begin{aligned}\gamma(h) &= E [(y_t - \mu_t) (y_{t+h} - \mu_t)] \\ &= E \left[ \sum_{i=0}^{\infty} \alpha_i \varepsilon_{t-i} \sum_{j=0}^{\infty} \alpha_{j+h} \varepsilon_{t-j+h} \right] \\ &= \sum_{i,j=0}^{\infty} \alpha_i \alpha_{j+h} \gamma_{\varepsilon}(i - j + h) \\ &= \sigma_{\varepsilon}^2 \sum_{i,j=0}^{\infty} \alpha_i \alpha_{j+h}\end{aligned}$$

# Séries temporelles stationnaires

## Outline

- 1 Processus AR
- 2 Processus MA
- 3 Processus ARMA

# Séries temporelles stationnaires

## Processus AR

### Definition (Processus AR)

Le processus stationnaire  $(y_t, t \in \mathbb{Z})$  satisfait une représentation AR d'ordre  $p$ , notée  $AR(p)$ , si et seulement si :

$$y_t = c + \sum_{i=1}^p \phi_i y_{t-i} + \varepsilon_t \quad (5)$$

# Séries temporelles stationnaires

## Processus AR

### Definition (Opérateur retard)

L'opérateur retard est défini par  $L : y_t \rightarrow L(y_t) = Ly_t = y_{t-1}$ . Le polynôme retard est défini par  $P(L) = (1 - \phi_1 L - \phi_2 L^2 - \dots - \phi_p L^p)$ .

### Definition (Processus AR)

Le processus (5) peut s'écrire

$$\Phi(L)y_t = c + \varepsilon_t \quad (6)$$

avec  $c \in \mathbb{R}$ ,  $\Phi(L) = \sum_{i=0}^p \phi_i L^i$  et  $\phi_0 = 1$  et  $\varepsilon_t \sim iid(0, \sigma_\varepsilon^2)$



# Séries temporelles stationnaires

## Processus AR

Soit  $y_t = c + \phi_1 y_{t-1} + \varepsilon_t$ , calculez

- $\mathbb{E}[y_t] = ?$
- $\text{Var}(y_t) = ?$
- $\text{Cov}(y_t, y_{t+h}) = ?$

# Séries temporelles stationnaires

## Processus AR

Quelques simulations, pour  $T = 100$ ,  $\varepsilon_t \sim N(0, 1)$

- simuler un processus AR(1) avec  $c = 0$  et  $\phi_1 = 0.1$
- simuler un processus AR(1) avec  $c = 0$  et  $\phi_1 = -0.1$
- simuler un processus AR(1) avec  $c = 0$  et  $\phi_1 = 0.9$
- simuler un processus AR(1) avec  $c = 0$  et  $\phi_1 = -0.9$
- simuler un processus AR(1) avec  $c = 0$  et  $\phi_1 = 1.01$

## Processus AR $\mu = 0$ $\phi_1 = 0.2$

```
T <- 100
mu <- 0
phi1 <- 0.2
e <- rnorm(T)
y <- rep(NA,T)
y[1] = 0
for (t in 2:T){
  y[t] <- mu + phi1*y[t-1] + e[t]
}
plot(y,type='l',
     xlab="Time")
```

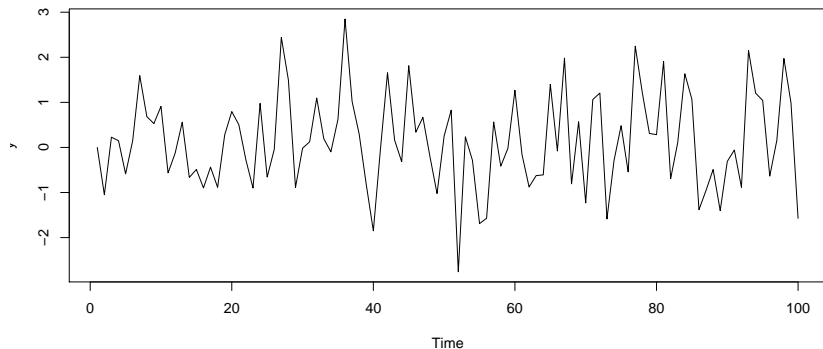


Figure 5: AR(1) with  $\phi_1 = 0.2$

## Processus AR $\mu = 0$ AR(1) $\phi_1 = -0.2$

```
T <- 100
mu <- 0
phi1 <- -0.2
e <- rnorm(T)
y <- rep(NA,T)
y[1] = 0
for (t in 2:T){
  y[t] <- mu + phi1*y[t-1] + e[t]
}
plot(y,type='l')
```

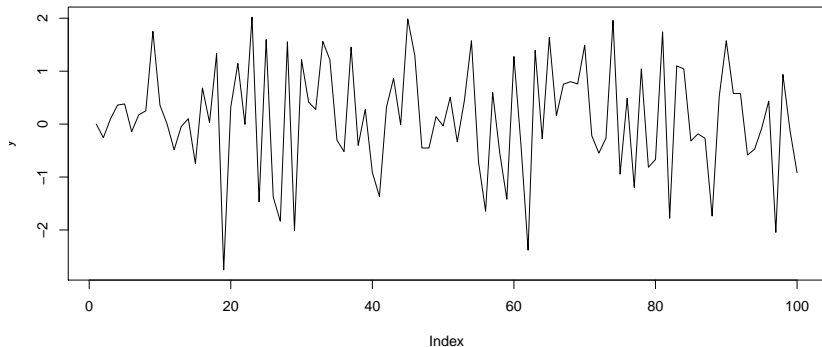


Figure 6: AR(1) with  $\phi_1 = -0.2$

## Processus AR $\mu = 0$ $\phi_1 = 0.9$

```
T <- 100
mu <- 0
phi1 <- 0.9
e <- rnorm(T)
y <- rep(NA, T)
y[1] = 0
for (t in 2:T){
  y[t] <- mu + phi1*y[t-1] + e[t]
}
plot(y, type='l')
```

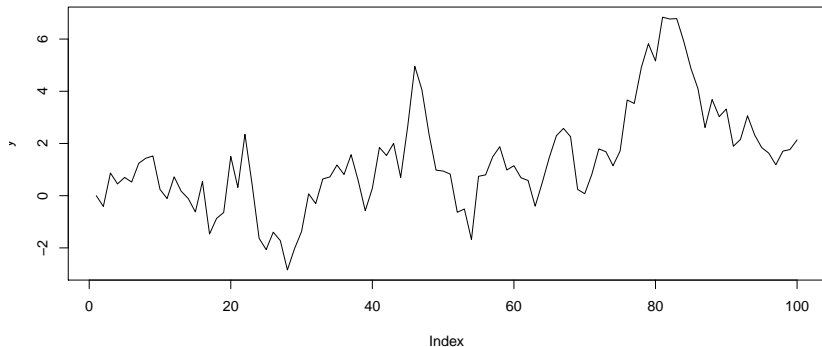


Figure 7: AR(1) with  $\phi_1 = 0.9$

## Processus AR $\mu = 0$ $\phi_1 = -0.9$

```
T <- 100
mu <- 0
phi1 <- -0.9
e <- rnorm(T)
y <- rep(NA,T)
y[1] = 0
for (t in 2:T){
  y[t] <- mu + phi1*y[t-1] + e[t]
}
plot(y,type='l')
```

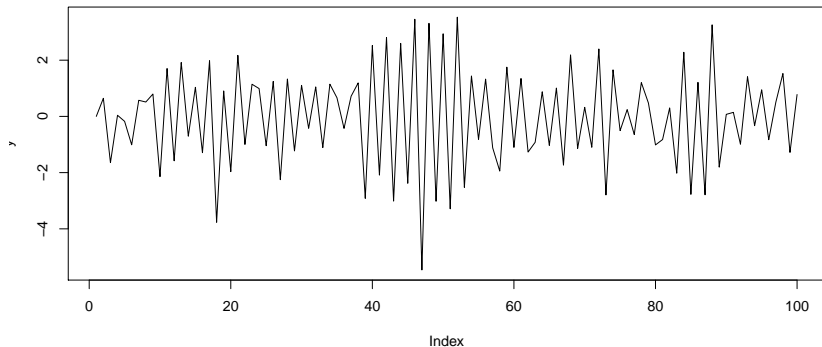


Figure 8: AR(1) with  $\phi_1 = -0.9$

## Processus AR, $\mu = 0$ , $\phi_1 = 1.01$

```
T <- 100
mu <- 0
phi1 <- 1.2
e <- rnorm(T)
y <- rep(NA,T)
y[1] = 0
for (t in 2:T){
  y[t] <- mu + phi1*y[t-1] + e[t]
}
plot(y,type='l')
```

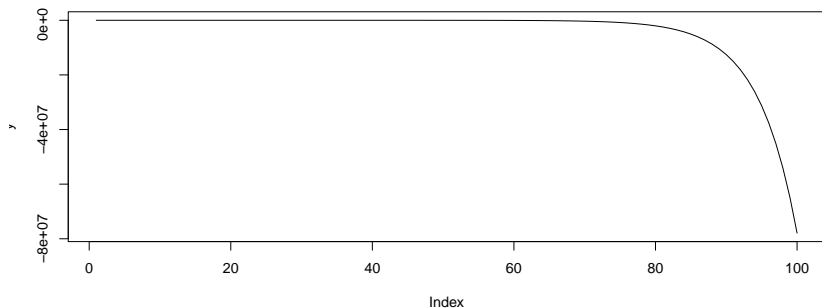


Figure 9: AR(1) with  $\phi_1 = 1.01$

# Séries temporelles stationnaires

## Processus AR

### Propriétés du modèle AR(p) :

- Fonction d'auto-covariance :

$$y_t = \phi_1 y_{t-1} + \cdots + \phi_p y_{t-p} + \varepsilon_t$$

- Fonction d'auto-corrélation :

$$\rho(h) = \text{Corr}(y_t, y_{t+h}) = \frac{\text{Cov}(y_t, y_{t+h})}{\text{Var}(y_t)}$$

$$\Leftrightarrow \rho(h) = \phi_1 \rho(h-1) + \cdots + \phi_p \rho(h-p)$$



# Séries temporelles stationnaires

## Processus AR

### Propriétés du modèle AR(p) :

- Fonction d'auto-covariance :

$$y_t = \phi_1 y_{t-1} + \cdots + \phi_p y_{t-p} + \varepsilon_t$$
$$\Leftrightarrow y_t^2 = \phi_1 y_{t-1} y_t + \cdots + \phi_p y_{t-p} y_t + \varepsilon_t y_t \quad (\times \mathbf{y_t})$$

- Fonction d'auto-corrélation :

$$\rho(h) = \text{Corr}(y_t, y_{t+h}) = \frac{\text{Cov}(y_t, y_{t+h})}{\text{Var}(y_t)}$$

$$\Leftrightarrow \rho(h) = \phi_1 \rho(h-1) + \cdots + \phi_p \rho(h-p)$$

# Séries temporelles stationnaires

## Processus AR

### Propriétés du modèle AR(p) :

- Fonction d'auto-covariance :

$$\begin{aligned}y_t &= \phi_1 y_{t-1} + \cdots + \phi_p y_{t-p} + \varepsilon_t \\ \Leftrightarrow y_t^2 &= \phi_1 y_{t-1} y_t + \cdots + \phi_p y_{t-p} y_t + \varepsilon_t y_t \quad (\times \mathbf{y_t}) \\ \Leftrightarrow \gamma(0) &= \phi_1 \gamma(1) + \cdots + \phi_p \gamma(p) + \sigma_\varepsilon^2\end{aligned}$$

- Fonction d'auto-corrélation :

$$\rho(h) = \text{Corr}(y_t, y_{t+h}) = \frac{\text{Cov}(y_t, y_{t+h})}{\text{Var}(y_t)}$$

$$\Leftrightarrow \rho(h) = \phi_1 \rho(h-1) + \cdots + \phi_p \rho(h-p)$$

# Séries temporelles stationnaires

## Processus AR

### Propriétés du modèle AR(p) :

- Fonction d'auto-covariance :

$$y_t = \phi_1 y_{t-1} + \cdots + \phi_p y_{t-p} + \varepsilon_t$$

$$\Leftrightarrow y_t^2 = \phi_1 y_{t-1} y_t + \cdots + \phi_p y_{t-p} y_t + \varepsilon_t y_t \quad (\times \mathbf{y_t})$$

$$\Leftrightarrow \gamma(0) = \phi_1 \gamma(1) + \cdots + \phi_p \gamma(p) + \sigma_\varepsilon^2$$

$$\Leftrightarrow \gamma(h) = \phi_1 \gamma(h-1) + \cdots + \phi_p \gamma(h-p)$$

- Fonction d'auto-corrélation :

$$\rho(h) = \text{Corr}(y_t, y_{t+h}) = \frac{\text{Cov}(y_t, y_{t+h})}{\text{Var}(y_t)}$$

$$\Leftrightarrow \rho(h) = \phi_1 \rho(h-1) + \cdots + \phi_p \rho(h-p)$$

# Séries temporelles stationnaires

## Processus AR

Équations de Yule-Walker :

$$\begin{pmatrix} 1 & \rho(1) & \dots & \rho(p-1) \\ \rho(1) & 1 & \rho(1) & \rho(p-2) \\ \vdots & & \ddots & \vdots \\ \rho(p-1) & \dots & \dots & 1 \end{pmatrix} \begin{pmatrix} \phi_1 \\ \vdots \\ \phi_p \end{pmatrix} = \begin{pmatrix} \rho(1) \\ \vdots \\ \rho(p) \end{pmatrix}$$

# Séries temporelles stationnaires

## Processus AR

### Proposition

*Si  $y_t \sim AR(p)$  stationnaire alors les  $|\rho(h)|$  et les  $\gamma(h)$  décroissent vers 0 exponentiellement*

# Séries temporelles stationnaires

## Processus AR

### Auto-corrélations partielle

#### Definition (Auto-corrélation partielle)

La corrélation entre  $y_t$  et  $y_{t-2}$  survient compte tenu de la dépendance avec  $y_{t-1}$ . L'auto-corrélation partielle vise à retirer la dépendance avec la variable intermédiaire  $y_{t-1}$ .

#### Proposition

Si  $(y_t)_{t \in \mathbb{Z}} \sim AR(p)$  tel que  $\Phi(L)y_t = \varepsilon_t$ , alors

$$r(h) \begin{cases} = 0 & \text{si } h > p \\ \neq 0 & \text{sinon.} \end{cases}$$

# Séries temporelles stationnaires

## Processus AR

**Démonstration :**  $r(h)$  est le coefficient de  $y_{t-h}$  dans  $E(y_t | y_{t-1} \dots y_{t-h})$  :

$$y_t = \phi_1 y_{t-1} + \dots + \phi_p y_{t-p} + \varepsilon_t$$

$$E(y_t | y_{t-1}, \dots, y_{t-p}) = \phi_1 y_{t-1} + \dots + \phi_p y_{t-p}$$

# ACF AR(1)

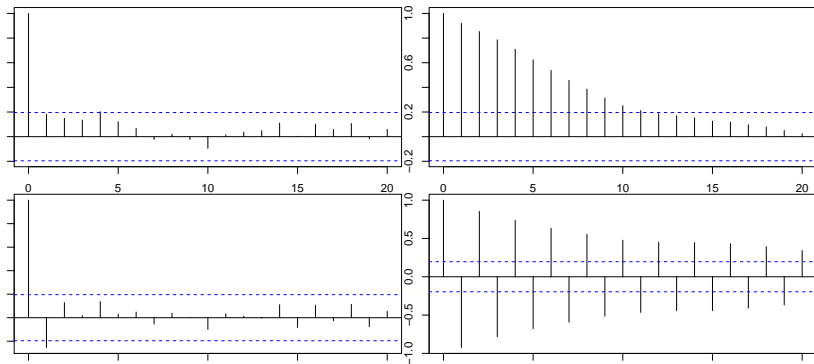


Figure 10: ACF AR(1)



# PACF AR(1)

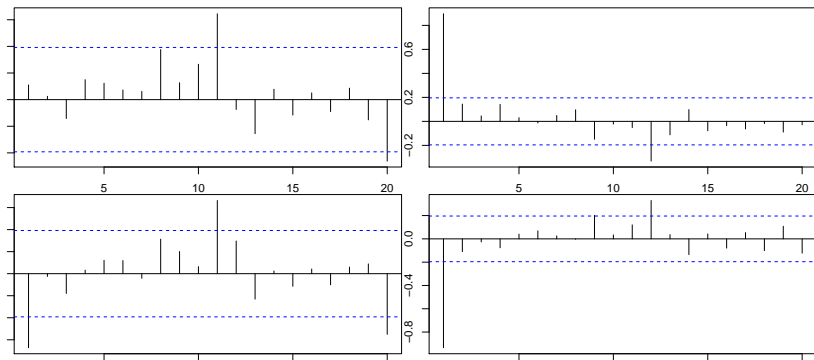


Figure 11: PACF AR(1)

# Séries temporelles stationnaires

## Processus AR

Quels sont les conditions de stationnarité? **Petit rappel**

### Definition

$A(L)$  est inversible  $\Leftrightarrow \exists B(L)$  tel que  $A(L) \circ B(L) = Id$ .

# Séries temporelles stationnaires

## Processus AR

### Exemple (AR(1))

Par déf, le processus est inversible car il est tourné vers le passé

$$y_t = \phi_1 y_{t-1} + \varepsilon_t$$

$$\Leftrightarrow y_t = \phi_1 L y_t + \varepsilon_t$$

$$\Leftrightarrow (1 - \phi_1 L) y_t = \varepsilon_t$$

$$\Leftrightarrow y_t = (1 - \phi_1 L)^{-1} \varepsilon_t$$

# Séries temporelles stationnaires

## Processus AR

- La racine du polynôme  $1 - \phi_1 L$  est  $z_1 = \frac{1}{\phi_1}$  car  $1 - \phi_1 \times \frac{1}{\phi_1} = 0$ .
- $|\phi_1| = 1$  la variance du processus dépend de  $t$ ,
- $|\phi_1| < 1 \Leftrightarrow |z_1| > 1$  et  $(1 - \phi_1 L)^{-1} = \sum_{i=0}^{\infty} \phi_1^i L^i$ , on a une relation causale,
- $|\phi_1| > 1 \Leftrightarrow |z_1| < 1$  et  $(1 - \phi_1 L)^{-1} = -\sum_{i=1}^{\infty} \phi_1^{-i} L^{-i}$ , la relation est non causale.

# Séries temporelles stationnaires

## Processus MA

### Definition (MA(1) process)

Un processus MA d'ordre 1 est défini par une combinaison linéaire des innovations :

$$y_t = \varepsilon_t - \theta_1 \varepsilon_{t-1} \quad (7)$$

ou

$$y_t = (1 - \theta_1 L) \varepsilon_t \quad (8)$$

### Definition (MA(q) process)

Un processus MA d'ordre q est défini par :

$$y_t = \varepsilon_t - \sum_{i=1}^q \theta_i \varepsilon_{t-i} \quad (9)$$

ou

$$y_t = \Theta(L) \varepsilon_t = (1 - \theta_1 L - \theta_2 L^2 - \dots - \theta_q L^q) \varepsilon_t \quad (10)$$

# PACF AR(2)

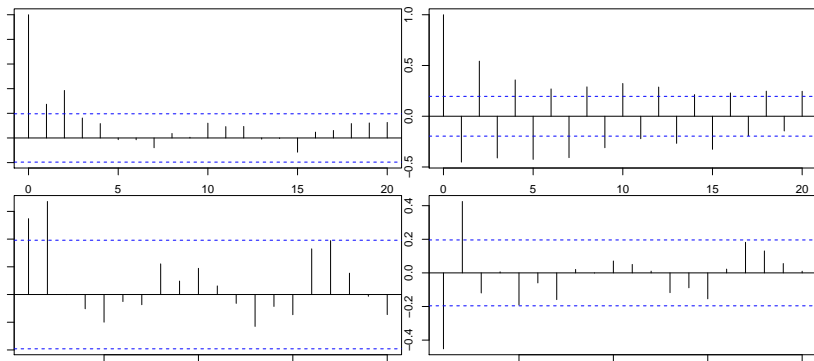


Figure 12: ACF et PACF AR(2)

# Séries temporelles stationnaires

## Processus MA

La variance d'un MA(1) :

- La variance du processus est

$$\mathbb{E}(y_t^2) = \mathbb{E}(\varepsilon_t^2) + \theta_1^2 \mathbb{E}(\varepsilon_{t-1}^2) - 2\theta_1 \mathbb{E}(\varepsilon_t \varepsilon_{t-1})$$

- Soit  $\sigma_y^2 = \sigma^2 (1 + \theta_1^2)$

# Séries temporelles stationnaires

## Processus MA

### ACF et PACF d'un MA(1)

$$\begin{aligned}\gamma_1 &= \mathbb{E}(\varepsilon_t y_{t-1}) - \theta_1 \mathbb{E}(\varepsilon_{t-1} y_{t-1}) \\ &= 0 - \theta_1 \mathbb{E}(\varepsilon_{t-1} (\varepsilon_{t-1} - \theta_1 \varepsilon_{t-2})) \\ &= -\theta_1 \sigma^2\end{aligned}$$

$$\begin{aligned}\gamma_2 &= \mathbb{E}(\varepsilon_t y_{t-2}) - \theta_1 \mathbb{E}(\varepsilon_{t-1} y_{t-2}) \\ &= 0\end{aligned}$$



# Séries temporelles stationnaires

## Processus MA

- $\rho_1 = \frac{\theta_1}{1+\theta_1^2}$
- $\rho_j = 0, \quad j > 1$
- Par contre, la PACF est non nulle et décroît exponentiellement.
- On peut voir ici la dualité entre les modèles MA et les modèles AR.

# Séries temporelles Stationnaires

## Processus MA

### ACF et PACF

$$\gamma_0 = (1 + \theta_1^2 + \dots + \theta_q^2) \sigma^2$$

$$\gamma_k = (-\theta_k + \theta_1\theta_{k+1} + \dots + \theta_{q-1}\theta_q^2) \sigma^2 \quad k = 1, \dots, q$$

$$\gamma_k = 0 \quad k > q$$

$$\rho_k = \frac{\sum_{i=0}^q \theta_i \theta_{k+i}}{\sum_{i=0}^q \theta_i^2} \quad k = 1, \dots, q$$

$$\rho_k = 0 \quad k > q$$

avec  $\theta_0 = -1$ .

# ACF MA(1)

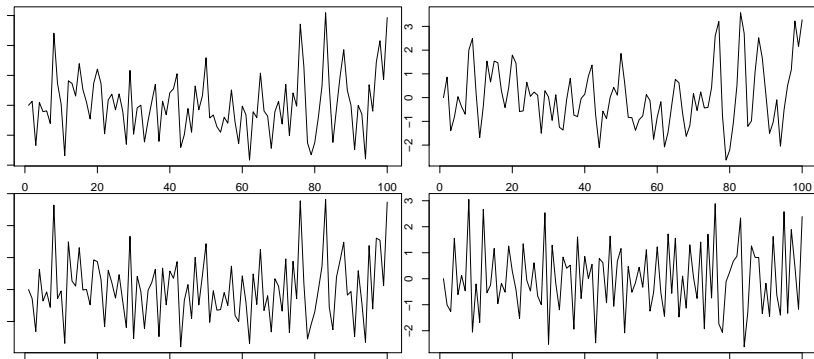


Figure 13: MA(1),  $\theta_{11} = 0.2$ ,  $\theta_{12} = 0.9$ ,  $\theta_{13} = -0.2$ ,  $\theta_{14} = -0.9$

# ACF MA(1)

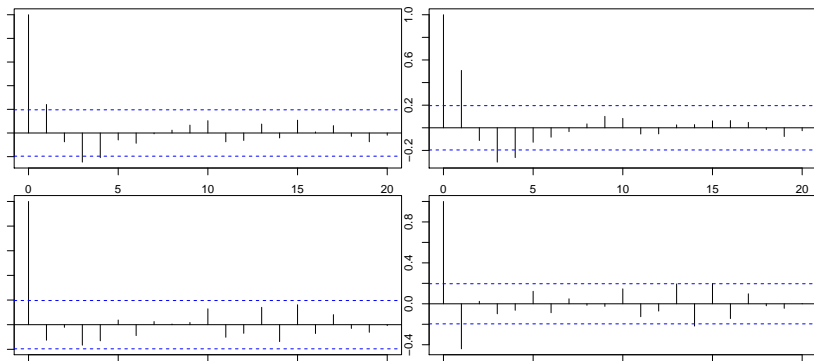


Figure 14: ACF MA(1)

# Séries temporelles stationnaires

## Processus MA

- Par définition ce processus est stationnaire et à une représentation causale.
- Soit  $z_i = \frac{1}{\lambda_i}$  les racines du polynôme  $\Theta(L)$ , alors ce polynôme peut se factoriser :

$$\Theta(L) = (1 - \lambda_1 L)(1 - \lambda_2 L) \dots (1 - \lambda_q L) \quad (11)$$

- si  $z_i \neq 1$  alors

$$\begin{aligned} \Theta(L)^{-1} &= (1 - \lambda_1 L)^{-1}(1 - \lambda_2 L)^{-1} \dots (1 - \lambda_q L)^{-1} \\ &= \frac{k_1}{1 - \lambda_1 L} + \frac{k_2}{1 - \lambda_2 L} + \dots + \frac{k_q}{1 - \lambda_q L} \end{aligned}$$

- Et donc  $\Theta(L)^{-1}y_t = \sum_{-\infty}^{\infty} \pi_i y_{t-i} = \varepsilon_t$

# Séries temporelles stationnaires

## Processus MA

Si les racines de  $\Theta(L) = 0$  sont toutes de module supérieur à 1, on peut montrer que  $\pi_i = 0, \forall i < 0$  On dit que le processus est inversible.

# Séries temporelles stationnaires

## Processus ARMA

### Definition

Un processus ARMA combine ces propriétés et nous autorise une représentation en forme réduite de ces deux dynamiques.

### Example (ARMA(1,1))

$$y_t = \phi_1 y_{t-1} + \varepsilon_t + \theta_1 \varepsilon_{t-1}$$
$$(1 - \phi_1 L)y_t = (1 - \theta_1 L)\varepsilon_t$$

# Séries temporelles stationnaires

## Processus ARMA

- Condition de stationnarité :  $|\phi_1| < 1$
- Condition d'inversibilité :  $|\theta_1| < 1$
- Il n'y a pas de racines communes dans les polynômes AR et MA



# Séries temporelles stationnaires

## Processus ARMA

### ACF

$$\gamma_k = \phi_1 \gamma_{k-1} + \mathbb{E}(\varepsilon_t y_{t-k}) - \theta_1 \mathbb{E}(\varepsilon_{t-1} y_{t-1}) \quad (12)$$

Pour  $k > 1$  les espérances sont nulles :

$$\gamma_k = \phi_1 \gamma_{k-1}, \quad k > 1 \quad (13)$$

Pour  $k = 0$ ,  $\mathbb{E}(\varepsilon_t y_t) = \sigma^2$  et  $\mathbb{E}(\varepsilon_{t-1} y_t) = \sigma^2(\phi_1 - \theta_1)$

# Séries temporelles stationnaires

## Processus ARMA

ACF :

$$\gamma_0 = \phi_1 \gamma_1 + \sigma^2 - \theta_1 \sigma^2 (\phi_1 - \theta_1)$$

$$\gamma_1 = \phi_1 \gamma_0 - \theta_1 \sigma^2$$

Soit :

$$\gamma_0 = \sigma^2 \frac{1 - 2\phi_1\theta_1 + \phi_1^2}{1 - \phi_1^2} \quad (14)$$

et

$$\rho_1 = \frac{(\phi_1 - \theta_1)(1 - \phi_1\theta_1)}{1 - 2\phi_1\theta_1 + \theta_1^2}$$

$$\rho_k = \phi_1 \rho_{k-1}$$

# Séries temporelles stationnaires

## Processus ARMA

**PACF** : On réécrit le modèle ARMA(1,1) comme un  $AR(\infty)$  :

$$(1 - \theta_1 L)^{-1}(1 - \phi_1 L)y_t = \varepsilon_t$$

$$y_t = (\phi_1 - \theta_1)y_{t-1} + \theta_1(\phi_1 - \theta_1)y_{t-2} + \theta_1^2(\phi_1 - \theta_1)y_{t-3} + \cdots + \varepsilon_t$$

# La méthode Box-Jenkins

La procédure de modélisation de Box et Jenkins (1976) est composée de 5 étapes :

- Stationnarisation
- **Identification** (Choix du modèle stationnaire)
- Estimation
- Validation et Test
- Prévisions

# Séries temporelles stationnaires

## Identification processus ARMA

### Comment choisir l'ordre des retards ?

- Certains sont basés sur la divergence Kullback-Leibler : comment une distribution de probabilité diverge d'une autre distribution de probabilité.
- Il propose une estimation de la perte d'information lorsqu'on utilise le modèle considéré pour représenter le processus qui génère les données.
- Il servent donc à comparer deux modèles.

# Séries temporelles stationnaires

## Identification processus ARMA

- AIC :  $2k - 2\log(L)$  ou  $k$  est le nombre de paramètres à estimer et  $L$  est le maximum de la fonction de vraisemblance.
- BIC :  $\log(T)k - 2\log(L)$  avec  $T$  la taille de l'échantillon.
- L'AIC pénalise le nombre de paramètres moins fortement que le BIC.
- Il en existe d'autres bien entendu !

# Séries temporelles stationnaires

## Identification processus ARMA

Algorithme de choix de modèle :

- Pour  $q$  allant de 1 à  $n$ ,
- Pour  $p$  allant de 1 à  $n$ ,
- Estimer  $\text{ARMA}(p,q)$
- Calculer AIC, BIC
- Fin
- Fin
- Calculer min AIC BIC

# Séries temporelles stationnaires

## Identification processus ARMA

- En pratique, on vérifie la significativité des coefficients
- Il se peut que les critères d'information choisissent  $p = q = 2$  mais que le paramètre  $\theta_1$  soit non significatifs.
- On peut opérer un test de Student classique.



# La méthode Box-Jenkins

La procédure de modélisation de Box et Jenkins (1976) est composée de 5 étapes :

- Stationnarisation
- Identification (Choix du modèle stationnaire)
- **Estimation**
- Validation et Test
- Prévisions

# Séries temporelles stationnaires

## Estimation processus ARMA

**Estimation** On estime généralement les processus ARMA via  
**Maximum de Vraisemblance**

- c'est une méthode **numérique** via un algorithme de maximisation
- Elle nécessite une hypothèse sur la **loi** des résidus

# Séries temporelles stationnaires

## Processus ARMA

La fonction de vraisemblance est une **fonction de probabilité** conditionnelle

- Elle prend comme arguments les paramètres du modèle pour les valeurs  $y_i$  observées ;
- Elle mesure une adéquation entre la distribution observée sur un échantillon aléatoire et une loi de probabilité supposée décrire une réalité sur la population ;
- On la note généralement  $L(y_1, y_2, \dots, y_t; \theta)$ .

# Séries temporelles stationnaires

## Estimation processus ARMA

### Definition

Notons  $f_\theta$  la densité de probabilité de la loi  $P_\theta$ . On appelle vraisemblance associée à  $P_\theta$ , la fonction qui à un n-uplet  $(y_1, \dots, y_t)$  et à une valeur de  $\theta$  du paramètre associe la quantité :

$$L(y_1, \dots, y_t, \theta) = \prod_{i=1}^T f_\theta(y_i).$$

L'estimateur de vraisemblance  $\hat{\theta}$  est la valeur de  $\theta$  qui maximise cette fonction.

# Séries temporelles stationnaires

## Estimation processus ARMA

Exemple : le lancer de pièce

- On lance un pièce 10 fois de suite et on obtient l'échantillon 0,1,0,1,0,1,1,1,0,0.

- Le lancer de pièce suit probablement une loi de Bernoulli :

$$P(X = x) = \begin{cases} p & \text{si } x = 1 \\ 1 - p & \text{si } x = 0 \\ 0 & \text{sinon} \end{cases} \quad \text{ou}$$

$$P(X = x) = p^x(1 - p)^{(1-x)}.$$

- La vraisemblance s'écrit :  $L(X_1, \dots, X_{10}, p) = (p^0(1 - p)^1) \times p^1(1 - p)^0 \times \dots \times p^0(1 - p)^1 = p^5(1 - p)^5$
- L'estimateur du Maximum de vraisemblance est la valeur de  $p$  qui maximise cette fonction.

# Séries temporelles stationnaires

## Estimation processus ARMA

### Definition

La vraisemblance d'une variable distribuée selon une loi normale est définie par

$$\prod_{i=1}^T f_{\theta}(y_i) = \left(\frac{1}{\sigma\sqrt{2\pi}}\right)^T e^{-\frac{1}{2} \sum_{i=1}^T \left(\frac{y_i - \mu}{\sigma}\right)^2}$$

### Definition

La log-vraisemblance d'une variable distribuée selon une loi normale est définie par

$$\log L(y_1, \dots, Y_T; \theta) = 0 - \frac{T}{2} T \log(\sigma^2) - \frac{T}{2} T \log(2\pi) - \frac{1}{2} \sum_{i=1}^T \frac{y_i - \mu}{\sigma^2}$$

# Séries temporelles stationnaires

## Estimation processus ARMA

- Supposons les erreurs du processus  
 $AR(1)$  :  $y_t = c + \phi_1 y_{t-1} + \varepsilon_t$  avec  $\varepsilon_t \sim N(0, \sigma_\varepsilon^2)$ ,
- Les variables aléatoires  $y_1, \dots, y_T$  suivent une loi normale  $\mathcal{N}(\frac{c}{1-\phi_1}, \frac{\varepsilon^2}{(1-\phi_1)^2})$  si  $y_t$  est stationnaire.
- Il suffit alors de remplacer dans la fonction de log-vraisemblance et de maximiser.

# Séries temporelles stationnaires

## Estimation processus ARMA

- Soit le modèle  $\Phi(L)y_t = \Theta(L)\varepsilon_t$  stationnaire.
- On suppose que la population des résidus  $\varepsilon_t$  peut-être décrite par un processus Gaussien  $N(0, \sigma_\varepsilon^2)$ .
- La vraisemblance associée au vecteur  $Y = (y_1, \dots, y_t)$  s'écrit :

$$\left( \frac{1}{\sqrt{(2\pi\sigma_\varepsilon^2)}} \right)^T \times \frac{1}{\det(\Omega)^{\frac{1}{2}}} \exp \left\{ -\frac{1}{2\sigma_\varepsilon^2} Y' \Omega^{-1} Y \right\}$$



# La méthode Box-Jenkins

La procédure de modélisation de Box et Jenkins (1976) est composée de 5 étapes :

- Stationnarisation
- Identification (Choix du modèle stationnaire)
- Estimation
- **Validation et Test**
- Prévisions

# Séries temporelles stationnaires

## Validation processus ARMA

### Comment vérifier que la série est bien modélisée ?

- Graphe d'un ACF, si il a le comportement d'un bruit blanc : ok !
- il faut aussi le vérifier à l'aide de tests statistiques sur la série filtrée.
- $\hat{\varepsilon}_t = y_t - \hat{\phi}_1 y_{t-1}$

# Séries temporelles stationnaires

## Validation processus ARMA

### Test de normalité

- Le test de Jarque-Bera :  
 $H_0$  : Les données suivent une loi normale
- C'est un test joint sur moments 3 et 4 d'une loi normale (Skewness et Kurtosis)
- Le Skewness représente l'asymétrie. Il est donc sensé être nul.
- Le Kurtosis représente l'aplatissement de la loi

# Séries temporelles stationnaires

## Validation processus ARMA

### Definition

La statistique du test de **Jarque-Bera** est donnée par :

$$JB = \frac{T - k}{6} \left( S^2 + \frac{(K - 3)^2}{4} \right)$$

- T le nombre d'observation ;
- k le nombre de variables explicatives ;
- S le Skewness ;
- K le kurtosis.
- Cette statistique de test suit une loi du  $\chi^2$  à 2 degrés de liberté.

# Séries temporelles stationnaires

## Validation processus ARMA

### Test de non-autocorrélation

- Test de Box-Pierce et **Ljung-Box**
- $H_0$  : Non autocorrelation
- $H_1$  : Autocorrelation.
- L'idée : On estime la fonction d'autocorrelation.
- $H_0 : \rho(1) = \rho(2) = \dots = \rho(p) = 0$

# Séries temporelles stationnaires

## Validation processus ARMA

### Definition

La statistique du test Ljung-Box est donnée par :

$$Q(m) = T(T+2) \sum_{h=1}^m \frac{\hat{\rho}(h)^2}{T-h} \quad (15)$$

- $\hat{\rho}_h = \frac{\sum_{t=h+1}^T (\varepsilon_t - \bar{\varepsilon})(\varepsilon_{t-h} - \bar{\varepsilon})}{(\varepsilon_t - \bar{\varepsilon})^2}$
- On rejette  $H_0$  si  $Q(m) > q_\alpha$  avec  $q_\alpha$  le  $100(1 - \alpha)$  quantile d'une  $\chi^2$  avec  $m - p - q$  degrés de liberté

# Séries temporelles stationnaires

## Validation processus ARMA

### Remarque

- *Le carré d'un bruit blanc Gaussien est aussi un bruit blanc Gaussien.*
- *Les résidus élevés au carré doivent donc aussi être non-corrélés.*
- *Économiquement, si on rejette l'hypothèse nulle, il faudrait modéliser la variance de la série d'une manière différente.*

# Séries temporelles stationnaires

## La méthode Box-Jenkins

La procédure de modélisation de Box et Jenkins (1976) est composée de 5 étapes :

- Stationnarisation
- Identification (Choix du modèle stationnaire)
- Estimation
- Validation et Test
- Prévisions



# Séries temporelles non stationnaires

## Introduction

Comment modéliser une série **non-stationnaire** ?

- Il faut la **stationnariser**. On peut ensuite l'étudier de façon **univariée** via ses propriétés **stochastiques**.
- L'ajout de **covariates** peut par contre poser des problèmes (**Cointégration**).
- **Cointegration** : deux séries ont une tendance de long-terme commune.
- Avant de régresser deux séries stationnaires, ils faut vérifier cette tendance commune.
- On ne s'y intéresse pas dans ce cours malheureusement.

# Séries temporelles non-stationnaires

## Outline

- **Non-stationnarité**
- Spurious regressions
- Rappels
- Stationnariser un processus

# Séries temporelles non-stationnaires

## Non-stationnarité

### Definition

Si un processus ne respecte pas les conditions de stationnarité et l'hypothèse d'ergodicité, alors il est non stationnaire

Il existe donc de nombreux processus non stationnaires

- La tendance déterministe linéaire :

$$y_t = \mu + \delta t + \varepsilon_t$$

- La tendance déterministe non-linéaire :

$$y_t = \mu + \delta(t) + \varepsilon_t \quad (16)$$

# Séries temporelles non-stationnaires

## Non-stationnarité

### Definition

$(y_t, t \in \mathbb{Z})$  est un processus **TS** si il peut s'écrire  $y_t = \delta(t) + u_t$ .

- $y_t$  est composé d'une partie **déterministe** et d'une partie **stochastique stationnaire**
- Exemple :  $y_t = b_0 + b_1 t + \varepsilon_t$ ,  $y_t - b_0 - b_1 t = \varepsilon_t$  est stationnaire
- Un choc sur la partie stochastique de ce genre de processus est dit **non persistant** car la tendance du modèle est **déterministe**.
- Économiquement, la trajectoire de long terme de la série est **insensible** aux aléas conjoncturels.

# Séries temporelles non-stationnaires

## Non-stationnarité

### Theorem

*Pour **stationnariser** un processus  $TS$ , il convient de retirer la composante déterministe  $\delta(t)$  en régressant la série  $y_t$  sur des puissances de  $t$ .*

- Exemple :  $y_t = b_0 + b_1 t + z_t + \varepsilon_t$ . On enlève la trend :  
 $y_t - \hat{b}_0 - \hat{b}_1 t = z_t + \varepsilon_t$ .

# Séries temporelles non-stationnaires

## Non-stationnarité

### Proposition

*L'influence d'un choc  $\varepsilon_t$  à une date  $T$  sur  $y_t$  défini par  $y_t = f(t) + z_t$  avec  $z_t$  stationnaire et  $E[z_t] = 0$ , est transitoire. Après le choc  $\varepsilon_T$ , la séquence des  $y_t$  converge ainsi vers sa valeur de long terme  $f(t)$ .*

# Séries temporelles non-stationnaires

## Non-stationnarité

### Definition

Une marche aléatoire est un processus stochastique non stationnaire respectant

$$y_t = y_{t-1} + \varepsilon_t, \quad \varepsilon_t \sim IID \quad (17)$$

- $\mathbb{E}[y_t] = \mathbb{E}[y_{t-1} + \varepsilon_t] = \dots = \mathbb{E}\left[y_0 + \sum_{i=0}^{t-1} \varepsilon_{t-i}\right]$
- $\text{Cov}(y_t, y_{t+h}) = (t+h) \sigma_\varepsilon^2$
- $y_t$  est une martingale de tendance stochastique  $\sum_i \varepsilon_{t-i}$

# Séries temporelles non-stationnaires

## Non-stationnarité

### Definition

Un processus est dit DS, Differency Stationary, si la non-stationnarité est causée par une source stochastique

- $(y_t, t \in \mathbb{Z})$  est un processus DS d'ordre  $d$  si le processus **filtré** défini par  $(1 - L)^d y_t$  est stationnaire. On dit que  $y_t$  est intégré d'ordre  $d$ .
- On peut définir une classe de processus stochastiques qui ne satisfont pas les conditions de la stationnarité, mais dont la différence à l'ordre  $d$  satisfait les propriétés de la stationnarité.



# Séries temporelles non-stationnaires

## Non-stationnarité

- Soit la marche aléatoire,  $y_t = y_{t-1} + \varepsilon_t$ . Ce processus est non-stationnaire.
- Soit le processus différencié,  $\Delta y_t = y_t - y_{t-1} = \varepsilon_t$ . Ce processus est stationnaire.

# Séries temporelles non-stationnaires

## Non-stationnarité

Les processus possédant une tendance déterministe ou stochastique sont globalement non-stationnaires

- le processus viole les conditions de stationnarité
- les paramètres du processus sont invariants
- la non-stationnarité existe pour toute évolution du processus

# Séries temporelles non-stationnaires

## Non-stationnarité

Les processus dont les paramètres évoluent dans le temps sont possiblement localement non-stationnaires

- un modèle à changement de régime peut être stationnaire dans un régime non-stationnaire dans autre régime (localement non-stationnaire)
- globalement stationnaire ou non-stationnaire, ce sont des processus non-linéaires

# Séries temporelles non-stationnaires

## Outline

- Définition de la stationnarité
- Non-stationnarité
- **Spurious regressions**
- Rappels
- Stationnariser un processus

# Séries temporelles non-stationnaires

## Spurious regression

Soit deux marches aléatoires  $y_t = y_{t-1} + \varepsilon_t$  et  $y_t = y_{t-1} + u_t$  :

- $x_0 = y_0 = 0$
- $\varepsilon_t \sim IID(0, \sigma_\varepsilon^2)$  et  $u_t \sim IID(0, \sigma_u^2)$ .
- $E[\varepsilon_t u_s] = 0$

Que se passe-t-il si l'on régresse  $y_t$  sur  $y_t$  sans avoir diagnostiquer la non stationnarité ?

$$y_t = \alpha + \beta y_t + v_t$$

# Séries temporelles non-stationnaires

## Outline

- Définition de la stationnarité
- Non-stationnarité
- Spurious regressions
- **Rappels**
- Stationnariser un processus

# Séries temporelles non-stationnaires

## Rappels

Soit  $y_n$  une fonction de  $n$  variables aléatoires  $y_n = f(Y_1, \dots, Y_n)$ .

- On étudie le comportement de  $y_n$  quand  $n \rightarrow \infty$ .
- La fonction  $f(\cdot)$  est souvent un estimateur. Cela peut aussi être une statistique de test.
- Cette étude repose sur des notions de convergence :
  - convergence presque sûre
  - convergence en probabilité
  - convergence en moyenne quadratique
  - convergence en loi

# Séries temporelles non-stationnaires

## Rappels

### Definition

$y_n$  converge presque sûrement vers une constante  $c$  si,

$$Pr\left(\lim_{n \rightarrow \infty} y_n = c\right) = 1$$

Explications :  $y_n$  tend vers une valeur constante de manière certaine, sa distribution asymptotique est une masse ponctuelle



# Séries temporelles non-stationnaires

## Rappels

### Definition (Convergence en probabilité)

$y_n$  converge en probabilité vers une constante  $c$ , si pour toute valeur de  $\varepsilon > 0$

$$\lim_{n \rightarrow \infty} Pr(|y_n - c| > \varepsilon) = 0 \quad (18)$$

### Definition (Convergence en moyenne quadratique)

$y_n$  converge en moyenne quadratique vers une constante  $c$ , si  $\mathbb{E}[|y_n|^2] < \infty$  et si pour tout  $\gamma$  :

$$\mathbb{E}[|y_n - c|^2] < \gamma \quad (19)$$

# Séries temporelles non-stationnaires

## Rappels

### Definition (Convergence en Loi)

Soit  $F_n(\cdot)$  la fonction de répartition de  $y_n$ .  $y_n$  converge en loi vers une variable aléatoire  $Y$  définie sur un support  $Y(\Omega)$  et ayant pour fonction de répartition  $F(\cdot)$  si,

$$\lim_{n \rightarrow \infty} F_n(z) = F(z), \forall z \in Y(\Omega)$$

# Séries temporelles non-stationnaires

## Rappels

### Theorem (Loi faible des grands nombres)

*Pour une séquence de variables aléatoires IID,  $y_t = y_1, \dots, y_n$ , la moyenne empirique de ces variables converge en probabilité vers l'espérance de  $y_t$*

$$\frac{1}{n} \sum_{i=1}^n y_i \xrightarrow[p]{} \mathbb{E}[y_t] = \bar{y}$$

# Séries temporelles non-stationnaires

## Rappels

### Theorem (Loi forte des grands nombres)

*Pour une séquence de variables aléatoires IID,  $y_t = y_1, \dots, y_n$ , la moyenne empirique de ces variables converge p.s vers l'espérance de  $y_t$*

$$\frac{1}{n} \sum_{i=1}^n y_i \xrightarrow{a.s} \mathbb{E}[y_t] = \bar{y}$$

*si  $\mathbb{E}[|y_t|] < \infty$ .*

# Séries temporelles non-stationnaires

## Rappels

### Theorem (Théorème central limite )

*Soit une séquence IID,  $y_t = y_1, \dots, y_n$  d'espérance  $\mathbb{E}[y_t] = m$  et de variance finie  $V(y_t) = \sigma^2$ . D'après le théorème central limite de Lindeberg-Levy,*

$$Z_n = \sqrt{n}(\bar{y} - m) \xrightarrow{d} \mathcal{N}(0, \sigma^2)$$

# Séries temporelles non-stationnaires

## Rappels

Soit le modèle linéaire :

$$y_t = X_t\beta + \varepsilon_t, \varepsilon_t \sim IID(0, \sigma_\varepsilon^2) \quad (20)$$

L'estimateur OLS est alors donné par

$$\frac{1}{n} \left( \sum_{t=1}^T x_t^2 \right)^{-1} \frac{1}{n} \sum_{t=1}^T x_t y_t = \beta + \frac{n^{-1} \sum_{t=1}^T x_t \varepsilon_t}{n^{-1} \sum_{t=1}^T x_t^2}$$

# Séries temporelles non-stationnaires

## Rappels

Est-ce que  $\hat{\beta}$  est convergent ?

$$\text{plim}_{n \rightarrow \infty} \hat{\beta} = \text{plim} \beta + \frac{\text{plim } n^{-1} \sum_{t=1}^T x_t \varepsilon_t}{\text{plim } n^{-1} \sum_{t=1}^T x_t^2}$$

- On a  $\mathbb{E}[x_t \varepsilon_t] = 0$  et  $V([x_t \varepsilon_t]) = \sigma_x^2 \sigma_\varepsilon^2$
- D'après la loi forte de grand nombre,  $\sum \frac{\sigma_x^2 \sigma_\varepsilon^2}{t^2} < \infty$
- Donc  $\frac{\text{plim } n^{-1} \sum_{t=1}^T x_t \varepsilon_t}{\text{plim } n^{-1} \sum_{t=1}^T x_t^2} = \frac{0}{\text{plim } n^{-1} \sum_{t=1}^T x_t^2} = 0$
- Notez que faille la condition  $\mathbb{E}[x_t^4] < \infty$  est nécessaire

# Séries temporelles non-stationnaires

## Rappels

- La consistance donne une distribution dégénérée :  $\hat{\beta} \xrightarrow{p} \beta$
- on peut montrer que  $\sqrt{n} \left( \hat{\beta} - \beta \right) \xrightarrow{p} \mathcal{N}(0, \sigma_{\varepsilon}^2)$



# Séries temporelles non-stationnaires

## Rappels

Que se passe-t-il dans le cas d'un AR(1) ?  $y_t = \rho y_{t-1} + \varepsilon_t$

- $\hat{\rho} = \rho + \frac{\sum_{t=1}^T y_{t-1} \varepsilon_t}{\sum_{t=1}^T y_{t-1}^2}$
- $\sigma_y^2 = \mathbb{E} \left[ \sum_{t=1}^T y_t^2 \right] = \mathbb{E} \left[ \sum_{t=1}^T \rho^2 y_{t-1}^2 + \varepsilon_t^2 \right] = \rho^2 \sigma_y^2 + \sigma_\varepsilon^2$
- Donc  $\sigma_y^2 = \frac{\sigma_\varepsilon^2}{1-\rho^2}$
- On a donc  $\sqrt{n}(\hat{\rho} - \rho) \xrightarrow{\rho} \mathcal{N}(0, 1 - \rho^2)$
- Comment tester  $\rho = 1$  ?

# Séries temporelles non-stationnaires

## Outline

- Définition de la stationnarité
- Non-stationnarité
- Spurious regressions
- Rappels
- **Stationnariser un processus**

# Séries temporelles non-stationnaires

## Stationnariser un processus

- Supposons que  $y_t = b_0 + b_1 t + \varepsilon_t$  et qu'on lui applique un filtre  $(1 - L)^d$
- $y_t - y_{t-1} = (1 - L)y_t = \Delta y_t$
- $\Delta y_t = b_0 + b_1 t + \varepsilon_t - b_0 - b_1(t - 1) - \varepsilon_{t-1} = b_1 + \varepsilon_t - \varepsilon_{t-1}$
- Le processus est bien stationnaire mais on a introduit de l'auto-corrélation dans les résidus...

# Séries temporelles non-stationnaires

## Stationnariser un processus

- Supposons que  $y_t = y_{t-1} + \varepsilon_t$  et qu'on le stationnarise à l'aide d'une régression sur une tendance
- On obtient un processus qui n'a aucun sens. D'autant plus que les paramètres estimés de la régression sont parfois significatifs

# Séries temporelles non-stationnaires

## Stationnariser un processus

- Test de **Dickey Fuller** de racine unitaire dont l'**hypothèse nulle** est la **non stationnarité** d'un processus AR(1) :  
$$y_t = \rho y_{t-1} + \varepsilon_t.$$
- Le test DF revient à tester les hypothèses  $H_0 : \rho = 1$  vs  $H_1 : \rho < 1$ .
- Il suffirait alors d'appliquer un test de Student mais sous l'hypothèse de non stationnarité, l'estimateur des MCO n'a pas une distribution asymptotique standard.
- Il faudra calculer les seuils de significativité.

# Séries temporelles non-stationnaires

## Stationnariser un processus

### Remarque

*La distribution **asymptotique**, sous  $H_0$ , de la statistique de Student  $t_{\hat{\rho}=1}$  du test de Dickey Fuller n'est pas **standard**.  
L'utilisation, à tort, des seuils standard associés à une distribution normale peut conduire à un mauvais diagnostic quant à la non stationnarité de la série étudiée. Ce type d'erreur conduit à **rejeter trop souvent** l'hypothèse de non stationnarité.*

# Séries temporelles non-stationnaires

## Stationnariser un processus

### Proposition

*Sous l'hypothèse  $H_0$  de non stationnarité, la distribution asymptotique de la statistique de Student  $t_{\hat{\rho}=1}$  diffère suivant le modèle utilisé :*

1  $y_t = \rho y_{t-1} + \varepsilon_t \Leftrightarrow \Delta y_t = \phi y_{t-1} + \varepsilon_t$

2  $y_t = b_0 + \rho y_{t-1} + \varepsilon_t \Leftrightarrow \Delta y_t = b_0 + \phi y_{t-1} + \varepsilon_t$

3  $y_t = b_0 + b_1 t + \rho y_{t-1} + \varepsilon_t \Leftrightarrow \Delta y_t = b_0 + b_1 t + \phi y_{t-1} + \varepsilon_t$

avec  $\phi = 1 - \rho$ .

On ne sait pas ce que l'on doit inclure dans le modèle testé. On propose généralement **une stratégie de tests de Dickey Fuller**, et non pas un seul test unique.

# Les processus linéaires non stationnaires

## Stationnariser un processus

La stratégie de test : on va du plus général au plus spécifique.

1 Modèle 1 :  $(1 - \phi L)(y_t - \alpha - \beta t) = \varepsilon_t$

2 Modèle 2 :  $(1 - \phi L)(y_t - \alpha) = \varepsilon_t$

3 Modèle 3 :  $(1 - \phi L)y_t = \varepsilon_t$



# Les processus linéaires non stationnaires

## Stationnariser un processus

- Stage 1, Modèle 1  $(1 - \phi L)(y_t - \alpha - \beta t) = \varepsilon_t$
- On test  $H_0 : \phi = 0$  vs  $H_1 : \phi < 0$
- Si on accepte  $H_0 : \phi = 0$  alors le processus peut contenir une racine unitaire
- Si on rejette  $H_0$ , le processus peut ne pas contenir de racine unitaire

# Les processus linéaires non stationnaires

## Stationnariser un processus

- Stage 2, Modèle 1  $(1 - \phi L)(y_t - \alpha - \beta t) = \varepsilon_t$
- Il faut vérifier que le modèle (1) est le bon
- Si on a accepté  $H_0 : \phi = 0$ , on test  $H_0 : \phi = \beta = 0$  sinon on test  $H_0 : \beta = 0$ 
  - 1er cas : Valeur critique simulée. Si on accepte  $H_0$ , le modèle (1) n'est pas le bon. Sinon  $I(1) + T + C$
  - 2ème cas : Valeur critique classique. Si on accepte  $H_0$ , le modèle (1) n'est pas le bon. Sinon  $I(0) + T + C$

# Les processus linéaires non stationnaires

## Stationnariser un processus

- Stage 3, Modèle 2  $(1 - \phi L)(y_t - \alpha) = \varepsilon_t$
- On test  $H_0 : \phi = 0$  vs  $H_1 : \phi < 0$
- Si on accepte  $H_0 : \phi = 0$  alors le processus peut contenir une racine unitaire et/ou
- Si on rejette  $H_0$ , le processus peut ne pas contenir de racine unitaire

# Les processus linéaires non stationnaires

## Stationnariser un processus

- Stage 4, Modèle 2  $(1 - \phi L)(y_t - \alpha) = \varepsilon_t$
- Il faut vérifier que le modèle (2) est le bon
- Si on a accepté  $H_0 : \phi = 0$ , on test  $H_0 : \phi = \alpha = 0$  sinon on test  $H_0 : \alpha = 0$ 
  - 1er cas : Valeur critique simulée. Si on accepte  $H_0$ , le modèle (1) n'est pas le bon. Sinon  $I(1) + C$
  - 2ème cas : Valeur critique classique. Si on accepte  $H_0$ , le modèle (1) n'est pas le bon. Sinon  $I(0) + C$

# Les processus linéaires non stationnaires

## Stationnariser un processus

- Stage 5, Modèle 3  $(1 - \phi L) y_t = \varepsilon_t$
- On test  $H_0 : \phi = 0$  vs  $H_1 : \phi < 0$
- Si on accepte  $H_0 : \phi = 0$  alors le processus est  $I(1)$
- Si on rejette  $H_0$ , le processus est  $I(0)$

# Dynamiques non-linéaires

## Introduction

- **Faits stylisés**
- Définition
- Les modèles nonlinéaires en moyenne
- Les modèles nonlinéaires en variance

# Dynamiques non-linéaires

## Introduction

- Pourquoi s'intéresser aux modèles non-linéaires ?
- Quelles sont les limites des modèles linéaires ?
- Quels instruments de diagnostic (visuels et statistiques) peuvent indiquer une incompatibilité entre les données et un modèle linéaire ?
- Quelles caractéristiques des données ne peuvent pas être modélisées par les méthodes linéaires ?

# Dynamiques non-linéaires

## Introduction

- **Objectifs** : révéler la distribution probabiliste qui décrit le processus sous-jacent dans le but de comprendre et d'interpréter les mécanismes générateurs de données, de **prévoir** les événements futurs et de contrôler la survenue des événements futurs à travers une **intervention**.
- Sur la base des **objectifs** fixés ex-ante et des **faits stylisés** identifiés sur les données observées, le modèle probabiliste le plus approprié est choisi



# Dynamiques non-linéaires

## Introduction

### Remarques

La caractérisation complète de la dynamique du processus stochastique est impossible si on ne recourt pas à des hypothèses simplificatrices.

- Si  $\Omega$  est fini,  $t \in 1, 2, \dots, T$  et qu'on se limite aux deux seuls premiers moments du processus stochastique, il faudrait estimer  $E(y_1), E(y_2), \dots, E(y_t)$ , soit  $T$  éléments, ainsi que les  $T(T+1)/2$  éléments de la matrice de variance-covariance résumant la dépendance temporelle.
- Hypothèse simplificatrice standard : stationnarité (au sens faible) du processus stochastique

# Dynamiques non-linéaires

## Introduction

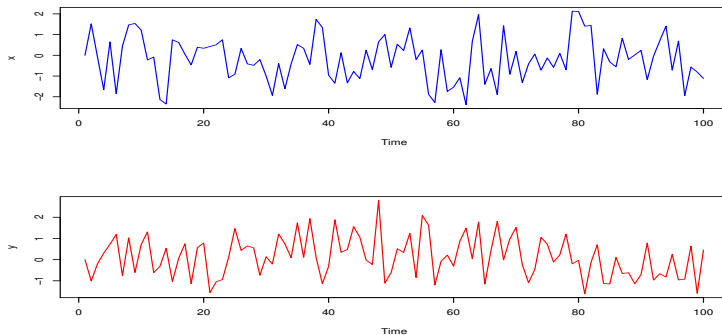


FIGURE – Deux processus, un linéaire, l'autre non

# Dynamiques non-linéaires

## Introduction

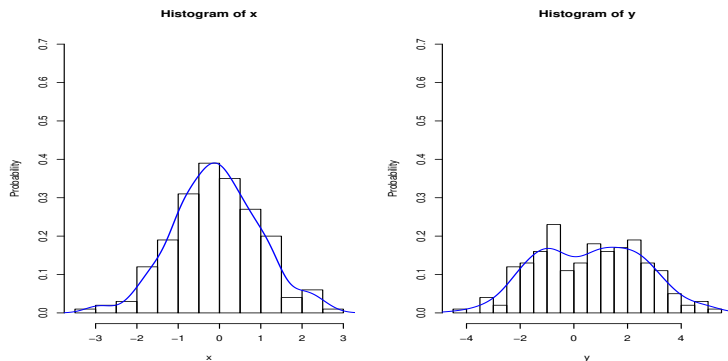
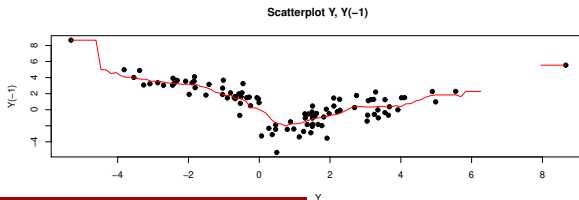
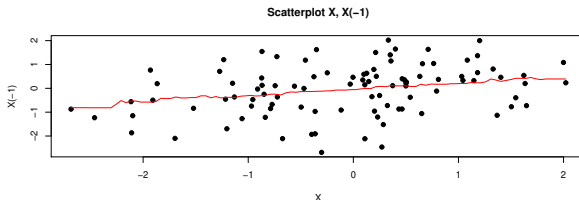


FIGURE – Deux processus, un linéaire, l'autre non

# Dynamiques non-linéaires

## Introduction

- Analyse de la relation entre  $y_t$  et  $y_{t-1}$  par une regression non paramétrique



# Dynamiques non-linéaires

## Introduction

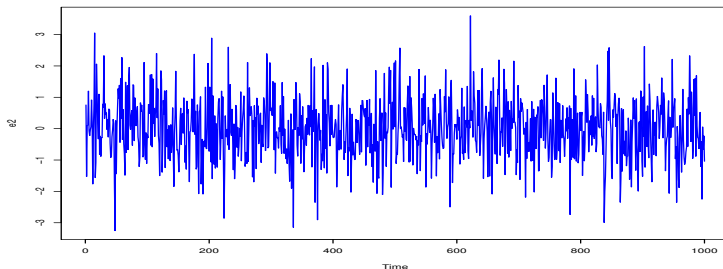


FIGURE – Bruit blanc Gaussien

- $Z_t \sim IID(0, \sigma^2)$  est un processus time-reversible

# Dynamiques non-linéaires

## Introduction

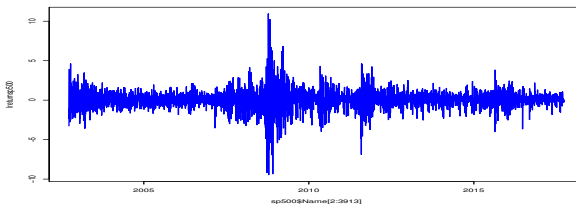


FIGURE – Log-returns du SP500

- Est-ce que ce processus est time-reversible ?

# Dynamiques non-linéaires

## Introduction

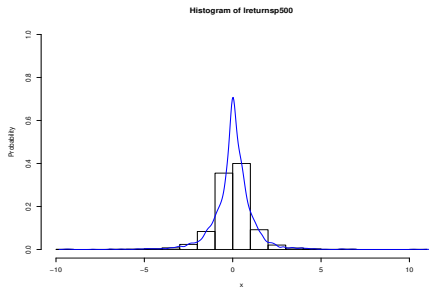


FIGURE – Histogramme et densité des rendements du SP500

# Dynamiques non-linéaires

## Introduction

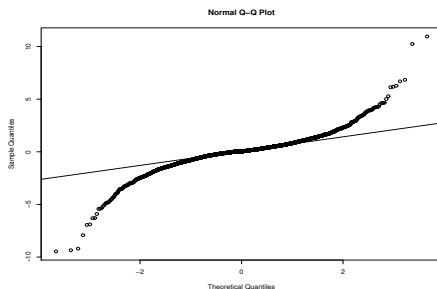


FIGURE – QQplot du SP500 vs Loi normale



# Dynamiques non-linéaires

## Définition

**Definition :** Tout processus qui n'est pas linéaire est un processus stochastique nonlinéaire. L'équation dynamique du processus doit être nonlinéaire.

- *Time-changing* Variance
- Assymétrie (business cycle)
- Effet de seuil
- Break, changements structurels

# Dynamiques non-linéaires

## Rappels sur les processus linéaires

### Définition

Dans ce cours, un processus stochastique  $\{y_t\}$ ,  $t \in 1, \dots, T$  est un processus linéaire s'il est linéaire dans les paramètres

- $y_t = \alpha + \beta x_t + u_t$  est linéaire ;
- $y_t = \alpha + \beta x_t^2 + u_t$  est linéaire ;
- $y_t = \alpha + \beta^2 x_t + u_t$  est nonlinéaire.

Un processus ARMA(p,q) est linéaire et peut s'écrire à l'aide de l'opérateur retard  $L$  :

$$\Phi(L)y_t = \Theta(L)u_t$$

# Dynamiques non-linéaires

## Définition

- Modèle linéaire = modèle de base, fondé sur la normalité
- Vaste classe de modèles non-linéaires, adaptés aux propriétés des séries économiques
- On distingue notamment la non-linéarité en moyenne de la non-linéarité en variance

# Dynamiques non-linéaires

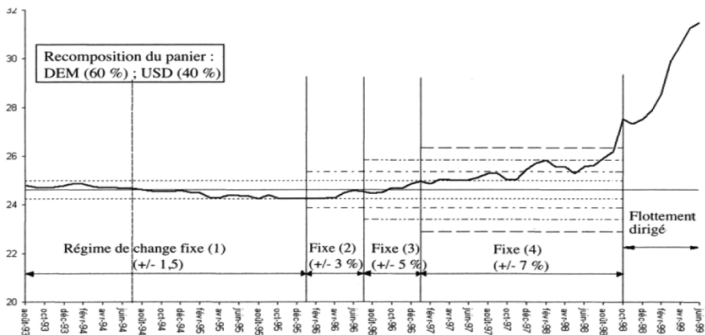
## Nonlinéaires en moyenne

- Les séries économiques et financières exhibent bien souvent des comportements non-linéaires
- La non-linéarité peut prendre différentes formes...
  - déterministe (e.g. saisonnalité intra-hebdomadaire des marchés)
  - stochastique (e.g. régime de change / crise économique)
- et dépendre de variables
  - latentes (e.g. crise économique : de nombreux facteurs inobservés)
  - observées (e.g. régime de change : politique monétaire)

# Dynamiques non-linéaires

Nonlinéaires en moyenne

Taux de change de la couronne slovaque  
par rapport au panier de devises constitué du DEM et de l'USD



# Dynamiques non-linéaires

## Nonlinéaires en moyenne

- On suppose que La variable  $y_t$  possède une dynamique non-linéaire qui dépend d'une variable observée  $z_t$
- La variable  $z_t$  étant observée, elle est considérée non-stochastique
- $z_t$  est néanmoins une réalisation d'une variable aléatoire
- c'est pourquoi il ne s'agit pas d'une non-linéarité déterministe

# Dynamiques non-linéaires

## Nonlinéaires en moyenne

- Tong (1990) introduit les modèles autorégressifs à seuil  $\Rightarrow$  **Threshold Autoregressiv (TAR)**
- Le **régime** est déterminé par une valeur seuil notée  $c$
- Il s'écrit :

$$y_t = \begin{cases} \phi_{0,1} + \phi_{1,1}y_{t-1} + \varepsilon_t, & \text{si } z_t < c \\ \phi_{0,2} + \phi_{1,2}y_{t-1} + \varepsilon_t, & \text{si } z_t > c \end{cases}$$

- Spécificité du modèle : la transition d'un régime à l'autre est abrupte.

# Dynamiques non-linéaires

## Nonlinéaires en moyenne

- Une spécification possible du TAR consiste à choisir  $z_t = y_{t-k}$ ,  $k > 0$
- On parle alors de modèle modèles auto-excité autorégressif à seuil **Self-Exciting Threshold AutoRegressive (SETAR)**
- Écriture alternative à l'aide de la fonction indicatrice

$$y_t = (\phi_{0,1} + \phi_{1,1}y_{t-1})(1 - \mathbb{1}(z_t > c)) + \dots \\ (\phi_{0,2} + \phi_{1,2}y_{t-1})\mathbb{1}(z_t > c) + \varepsilon_t$$



# Dynamiques non-linéaires

## Nonlinéaires en moyenne

- La transition d'un régime à l'autre dans les modèles TAR et SETAR est abrupte
- On peut modéliser la transition entre deux régimes via une fonction lisse/souple en remplaçant la fonction indicatrice par une fonction continue
- **Smooth Transition AutoRegressive (STAR)**

$$y_t = (\phi_{0,1} + \phi_{1,1}y_{t-1})(1 - G(z_t; \gamma, c)) + \dots \\ (\phi_{0,2} + \phi_{1,2}y_{t-1})G(z_t; \gamma, c) + \varepsilon_t$$

# Dynamiques non-linéaires

## Nonlinéaires en moyenne

- Généralement, on utilise la fonction logistique :

$$G(z_t; \gamma, c) = \frac{1}{1 + \exp(\gamma(z_t - c))}$$

- ou exponentielle :

$$G(z_t; \gamma, c) = 1 - \exp(-\gamma(z_t - c)^2)$$

- **Smooth Transition AutoRegressive (STAR)**

$$y_t = (\phi_{0,1} + \phi_{1,1}y_{t-1})(1 - G(z_t; \gamma, c)) + \dots \\ (\phi_{0,2} + \phi_{1,2}y_{t-1})G(z_t; \gamma, c) + \varepsilon_t$$

# Dynamiques non-linéaires

## Nonlinéaires en moyenne

- Avec La fonction logistique, les deux régimes correspondent à des valeurs faibles et élevées de la variable de transition  $z_t$  relativement à  $c$
- La fonction exponentielle permet de considerer des regimes associés avec des valeurs absolues faibles et importantes de la variable de transition

# Dynamiques non-linéaires

## Nonlinéaires en moyenne

Il est possible d'augmenter le nombre de régime des modèles à seuil de deux manières

**1** en augmentant le nombre de seuils :

$$y_t = \begin{cases} \phi_{0,1} + \phi_{1,1}y_{t-1} + \varepsilon_t, & \text{si } z_t < c_1 \\ \phi_{0,2} + \phi_{1,2}y_{t-1} + \varepsilon_t, & \text{si } c_1 < z_t < c_2 \\ \phi_{0,3} + \phi_{1,3}y_{t-1} + \varepsilon_t, & \text{si } c_2 < z_t \end{cases}$$

**2** en augmentant le nombre de variables interagissant avec le seuil :

$$y_t = \begin{cases} \phi_{0,1} + \phi_{1,1}y_{t-1} + \varepsilon_t, & \text{si } z_{1,t} < c_1 ; z_{2,t-1} < c_1 \\ \phi_{0,2} + \phi_{1,2}y_{t-1} + \varepsilon_t, & \text{si } z_{1,t} < c_1 ; z_{2,t-1} > c_1 \\ \phi_{0,3} + \phi_{1,3}y_{t-1} + \varepsilon_t, & \text{si } z_{1,t} > c_1 ; z_{2,t-1} < c_1 \\ \phi_{0,4} + \phi_{1,4}y_{t-1} + \varepsilon_t, & \text{si } z_{1,t} > c_1 ; z_{2,t-1} > c_1 \end{cases}$$

# Dynamiques non-linéaires

## Nonlinéaires en moyenne

$$H_0 : \phi_{0,1} = \phi_{0,2} \text{ et } \phi_{1,1} = \phi_{1,2}$$

- Le seuil  $c$  c'est pas identifiable sous  $H_0$ , c'est ce qu'on appelle un paramètre de nuisance. La distribution de la statistique de test est non-standard et les valeurs critiques sont obtenues par simulation.

$$t = n \frac{\hat{\sigma}_c^2 - \hat{\sigma}_{nc}^2}{\hat{\sigma}_c^2}$$

# Dynamiques non-linéaires

## Nonlinéaires en moyenne

$H_0 : \phi_{0,1} = \phi_{0,2} \text{ et } \phi_{1,1} = \phi_{1,2}$

- Sous  $H_0$ , on est en présence de paramètres de nuisance (non-identifiés)  $(c, \gamma)$
- Il faut reformuler l'hypothèse nulle  $H_0 : \gamma = 0$ .
- Dans ce cas, le paramètre de seuil  $c$  et les paramètres autoregréssifs  $\phi_{1,1}$  et  $\phi_{1,2}$  ne sont pas identifiés ;
- Les statistiques de test ont des distributions non-standards et les p-values sont calculables uniquement par simulation.

# Dynamiques non-linéaires

## Nonlinéaires en moyenne

$$H_0 : \gamma = 0$$

- **Solution** : Remplacer la fonction de transition par un développement de Taylor approprié et tester la linéarité par une statistique classique du Multiplicateur de Lagrange (LM) qui suit asymptotiquement une  $\chi^2$ . (Luukkonen, Saikkonen, Teräsvirta, 1988)

Avantages :

- 1 Estimer uniquement le modèle sous  $H_0$
- 2 Utiliser la théorie asymptotique standard pour calculer les valeurs critiques

# Dynamiques non-linéaires

## Nonlinéaires en moyenne

Les changements de régime dans les modèles à seuil peuvent être catégorisé de *déterministe*, les MS-AR proposent des modèles à changement de régimes **stochastiques**. La probabilité de changer de régime est

- égale à 1 dans les modèles à seuil
- **strictement** inférieure à 1 dans les modèles MS-AR



# Dynamiques non-linéaires

## Nonlinéaires en moyenne

- Les phénomènes cycliques sont un sujet d'étude central : Récessions/expansions, Bull/Bear
- Ces dynamiques cycliques irrégulières échappent aux modèles linéaires. Les cycles sont persistants et de durées diverses. Ajouter des retards dans un modèle autorégressif ne suffit pas !
- Ces dynamiques impliquent des distributions multimodales : l'hypothèse Gaussienne ne tient pas.

# Dynamiques non-linéaires

## Nonlinéaires en moyenne

### Definition

On dit qu'un variable aléatoire  $\{\Delta_t\}$  suit une chaîne de Markov d'ordre  $k$  si le futur de la variable ne dépend que des  $m$  valeurs passées.

$$\begin{aligned} P(\Delta_{t+1} = j | \Delta_t, \Delta_{t-1}, \dots, \Delta_{t-m}, \dots, \Delta_0) \\ = P(\Delta_{t+1} = j | \Delta_t, \Delta_{t-1}, \dots, \Delta_{t-m}) \end{aligned}$$

- $j$  représente les états de la nature, il est supposé dénombrable.
- Les modèles de MS-AR suppose que  $m = 0$ .

# Dynamiques non-linéaires

## Nonlinéaires en moyenne

- Une chaîne de Markov peut-être représenté par une matrice de transition de taille  $k \times k$ .
- $p_{ij} = P(\Delta_{t+1} = j | \Delta_t = i)$

$$P = \begin{pmatrix} p_{11} & \dots & p_{1k} \\ \vdots & \ddots & \vdots \\ p_{k1} & \dots & p_{kk} \end{pmatrix}$$

- Possible de la représenter graphiquement.

# Dynamiques non-linéaires

## Nonlinéaires en moyenne

Pour faire de l'inférence sur une chaîne de Markov il faut plusieurs hypothèses :

- 1 homogène :  $P(\Delta_t = i | \Delta_{t-1} = j) = P(\Delta_{t-l} = i | \Delta_{t-l-1} = j)$
- 2 irréductible : tout état est accessible depuis n'importe quel autre état à toute date  $t$ .
- 3 ergodique : apériodique et récurrente positive. Elle assure que la distribution de  $\Delta_t$  converge vers une distribution invariante quelque soit l'état initial.

# Dynamiques non-linéaires

## Nonlinéaires en moyenne

Pour faire de l'inférence sur une chaîne de Markov il faut plusieurs hypothèses :

- Grâce à toutes ces hypothèses on peut calculer les probabilités non conditionnelles :

$$\pi_1 = \frac{1 - p_{22}}{2 - p_{11} - p_{22}}$$

$$\pi_2 = 1 - \pi_1$$

# Dynamiques non-linéaires

## Nonlinéaires en moyenne

Partons du modèle le plus simple possible :

$$y_t = \mu_{\Delta_t} + \varepsilon_t, \varepsilon_t \sim B(0, \sigma^2)$$

avec  $\Delta_t = 1, 2$ .

- $\mu$  dépend de la chaîne de Markov :  $\theta = (\mu_1, \mu_2, \sigma_\varepsilon)'$
- $y_t = \begin{cases} \mu_1 + \varepsilon_t, & \text{si } \Delta_t = 1 \\ \mu_2 + \varepsilon_t, & \text{si } \Delta_t = 2 \end{cases}$

Cette classe de modèle dit Markov-Switching (MS) a été popularisée par Hamilton (1989).  $y_t$  est donc un mélange Markovien de 2 distributions normales.

# Dynamiques non-linéaires

## Nonlinéaires en moyenne

Ce modèle peut-être complexifié :

$$y_t = \mu_{\Delta_t} + \varepsilon_t, \varepsilon_t \sim B(0, \sigma_{\Delta_t}^2)$$

- $\mu$  et  $\sigma$  dépendent de la chaîne de Markov :

$$\theta = (\mu_1, \mu_2, \sigma_1, \sigma_2)'$$

Amusez-vous à simuler ces processus et voir le comportement de la densité estimée. Il faut aussi savoir qu'il y a un lien entre les modèles MS et les modèles ARMA.

# Dynamiques non-linéaires

## Nonlinéaires en moyenne

Ce modèle peut-être **encore** complexifié :

$$y_t = \mu_{\Delta_t} + \phi_{\Delta_t} y_{t-1} + \varepsilon_t, \varepsilon_t \sim N(0, \sigma_{\Delta_t}^2)$$

- $\mu$ ,  $\phi$  et  $\sigma$  dépendent de la chaîne de Markov :

$$\theta = (\mu_1, \mu_2, \phi_1, \phi_2, \sigma_1, \sigma_2)'$$



# Dynamiques non-linéaires

## Nonlinéaires en moyenne

Plusieurs tests peuvent être intéressant :

- tests de structure : la structure nonlinéaire du modèle est-elle pertinente ?
- tests de diagnostic variable omise ?

Si on souhaite tester la linéarité du modèle, l'hypothèse nulle est :

- $\mu_1 = \mu_2, \phi_1 = \phi_2, \sigma_1 = \sigma_2 \Rightarrow$  la matrice  $P$  n'est pas identifiée.
- ou  $p_{11} = 1$  et  $p_{21} = 0 \Rightarrow$  les paramètres du régime 2 ne sont identifiés.

Carrasco *et al* (2014) proposent un test très lourd à mettre en place.

# Dynamiques non-linéaires

## Nonlinéaires en variance

### Le modèle GARCH(p,q)

Soit  $\{y_t\}$  un processus stochastique tel que :

$$y_t = E[y_t | \mathcal{F}_{t-1}] + \varepsilon_t, \quad \text{et} \quad \text{Var}[y_t | \mathcal{F}_{t-1}] = h_t,$$

pour  $t = 1, \dots, T$  ;  $E[y_t | \mathcal{F}_{t-1}]$  et  $\text{Var}[y_t | \mathcal{F}_{t-1}]$  sont l'espérance et la variance conditionnelles de  $y_t$ .

# Dynamiques non-linéaires

## Nonlinéaires en variance

- La moyenne conditionnelle peut être très générale et doit être spécifiée indépendamment

$$E[y_t | \mathcal{F}_{t-1}] = m(x_t, \lambda),$$

- $m$  doit être continue et deux fois différentiable par rapport à  $\lambda$ .
- Les erreurs sont définies par

$$\varepsilon_t = h_t^{1/2} \eta_t, \quad \eta_t \sim \text{IID}(0, 1)$$

$$h_t = \alpha_0 + \sum_{j=1}^q \alpha_j \varepsilon_{t-j}^2 + \sum_{j=1}^p \gamma_j h_{t-j}.$$

# Dynamiques non-linéaires

## Nonlinéaires en variance

Soit le processus  $y_t = \phi_0 + \phi_1 y_{t-1} + \varepsilon_t$ ,  $\varepsilon_t \sim N(0, \sigma_\varepsilon)$

- Calculez l'espérance et la variance non conditionnelles et conditionnelles de  $y_t$
- Soit  $\varepsilon_t = h_t^{1/2} \eta_t$  avec  $h_t = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \gamma_1 h_{t-1}$ , répondez de nouveau à la question précédente.
- Proposez des conditions de stationnarité.

# Introduction aux modèles multivariés

## VAR Models

L'origine des modèle VAR

*Macroeconomics and Reality, Sims, Econometrica (1980)*

- Estimer des gros modèles Macro. Chaque variable est endogène
- Extension naturelle des modèles AR sous une forme multivariée
- Permet de décrire le comportement dynamique de variables économiques et financières

# Introduction aux modèles multivariés

## VAR Models

$$y_t = \beta_{10} + \phi_{11}y_{t-1} + \phi_{12}z_{t-1} + \varepsilon_{yt}$$

$$z_t = \beta_{20} + \phi_{21}z_{t-1} + \phi_{22}y_{t-1} + \varepsilon_{zt}$$

- $y$  et  $z$  sont endogènes, les erreurs sont des bruits blancs
- $\varepsilon_{yt}$  affecte  $z$  de deux façons.
- Il y a 10 paramètres à estimer.
- $\Sigma\varepsilon = \begin{pmatrix} \sigma_y^2 & \sigma_y\sigma_z \\ \sigma_z\sigma_y & \sigma_z^2 \end{pmatrix}$

# Introduction aux modèles multivariés

## VAR Models

Cette forme est appelé forme-réduite (reduced form) du modèle VAR

- Valeurs courantes exprimées en fonction des valeurs retardées
- Utile pour la prévision
- Utile pour résumer les propriétés des données
- et utile pour quantifier la vitesse à laquelle les variables reviennent à leur équilibre après un choc.

# Introduction aux modèles multivariés

## VAR Models

Structural form

$$y_t = \beta_{10} - \beta_{12}z_t + \phi_{11}y_{t-1} + \phi_{12}z_{t-1} + \varepsilon_{yt}$$

$$z_t = \beta_{20} - \beta_{22}y_t + \phi_{21}z_{t-1} + \phi_{22}y_{t-1} + \varepsilon_{zt}$$

$$\begin{pmatrix} 1 & \beta_{12} \\ \beta_{22} & 1 \end{pmatrix} \begin{pmatrix} y_t \\ z_t \end{pmatrix} = \begin{pmatrix} \beta_{10} \\ \beta_{20} \end{pmatrix} + \begin{pmatrix} \phi_{11} & \phi_{12} \\ \phi_{21} & \phi_{22} \end{pmatrix} \begin{pmatrix} y_{t-1} \\ z_{t-1} \end{pmatrix} + \begin{pmatrix} \varepsilon_{yt} \\ \varepsilon_{zt} \end{pmatrix}$$

ou

$$BY_t = \Gamma_1 + \Gamma_2 Y_{t-1} + \varepsilon_t$$



# Introduction aux modèles multivariés

## VAR Models

### Definition

Un modèle VAR peut s'écrire sous la forme réduite suivante :

$$Y_t = A_1 + A_2 Y_{t-1} + A_3 \varepsilon_t$$

avec  $A_1 = B^{-1}\Gamma_1$ ,  $A_2 = B^{-1}\Gamma_2$  et  $A_3 = B^{-1}$ .

# Introduction aux modèles multivariés

## VAR Models

### Remarque

- 1 *On utilise l'AIC pour choisir l'ordre des retards*
- 2 *Le système doit être stable (stationnaire). Il est stable si les racines de la matrice  $A_2$  sont plus petites que 1 en VA.*
- 3 *Si non, il faut utiliser un Vector Error Correction Models*

# Introduction aux modèles multivariés

## VAR Models

### Impulse Response Functions (IRF)

Objectif : La réaction du système à un choc

$$Y_t = A_1 Y_{t-1} + \varepsilon_t$$

Si le système est stationnaire en covariance :

$$Y_t = \Psi(L)a_t = \psi_1 a_{t-1} + \psi_2 a_{t-2} + \dots$$

$$Y_{t+s} = \Psi(L)a_{t+s} = \psi_1 a_{t+s-1} + \psi_2 a_{t+s-2} + \dots \quad \frac{\partial Y_{t+s}}{\partial a_t} = \psi_s$$

# Introduction aux modèles multivariés

## VAR Models

### Remarque

- 1 *Est-ce possible de retrouver les paramètres du modèle VAR initial ?*
- 2 *Sims propose d'identifier le modèle via une décomposition de Cholesky.*
- 3 *On pose  $\beta_{12} = 0$ .*

# Introduction aux modèles multivariés

## VAR Models

### Exemple

$$\begin{pmatrix} y_t \\ \pi_t \\ r_t \end{pmatrix} = \begin{pmatrix} c_1 \\ c_2 \\ c_3 \end{pmatrix} + \begin{pmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \\ b_{31} & b_{32} & b_{33} \end{pmatrix} \begin{pmatrix} y_{t-1} \\ \pi_{t-1} \\ r_{t-1} \end{pmatrix} + \begin{pmatrix} u_{yt} \\ u_{\pi t} \\ u_{rt} \end{pmatrix}$$

$$\begin{pmatrix} y_t \\ \pi_t \\ r_t \end{pmatrix} = \begin{pmatrix} c_1 \\ c_2 \\ c_3 \end{pmatrix} + \begin{pmatrix} \tilde{b}_{11} & \tilde{b}_{12} & \tilde{b}_{13} \\ \tilde{b}_{21} & \tilde{b}_{22} & \tilde{b}_{23} \\ \tilde{b}_{31} & \tilde{b}_{32} & \tilde{b}_{33} \end{pmatrix} \begin{pmatrix} y_{t-1} \\ \pi_{t-1} \\ r_{t-1} \end{pmatrix} + \begin{pmatrix} \tilde{a}_{11} & 0 & 0 \\ \tilde{a}_{21} & \tilde{a}_{22} & 0 \\ \tilde{a}_{31} & \tilde{a}_{32} & \tilde{a}_{33} \end{pmatrix} \begin{pmatrix} \varepsilon_{yt} \\ \varepsilon_{\pi t} \\ \varepsilon_{rt} \end{pmatrix}$$

# Introduction aux modèles multivariés

## VAR Models

### Definition

Décomposition de Cholesky :  $\Sigma_u = P'P$ ,  $P'$  lower triangular

On est donc dans ce cas là !

# Introduction aux modèles multivariés

## VAR Models

$$y_t = c_1 + \cdots + \tilde{a}_{11}\varepsilon_{yt}$$

$$\pi_t = c_2 + \cdots + \tilde{a}_{21}\varepsilon_{yt} + \tilde{a}_{22}\varepsilon_{\pi t}$$

$$r_t = c_3 + \cdots + \tilde{a}_{31}\varepsilon_{yt} + \tilde{a}_{32}\varepsilon_{\pi t} + \tilde{a}_{33}\varepsilon_{rt}$$

# Introduction aux modèles multivariés

## VAR Models