

Introduction aux modèles de séries temporelles

1.1 Qu'est-ce qu'une série temporelle ?

Une série temporelle est un ensemble d'observations répétées d'une même variable qui sont collectées à des intervalles de temps réguliers. Dans ce qui suit, on note une série temporelle de la façon suivante :

$$\{X_1, X_2, \dots, X_T\} \text{ ou } \{X_t\}_{t=1}^T$$

X_t est appelée variable aléatoire.

1.1.1 Définitions

Définition 1.1. *Un processus stochastique est une famille $\{X_t, t \in T\}$ de variables aléatoires définies sur un même espace de probabilité.*¹

L'indice t représente le temps (jour, semaine, mois ou année). Le processus est en temps **continu** si T est continu et en temps **discret** si T est discret.

Définition 1.2. *Une série temporelle $\{X_t\}_{t=1}^T$ est la partie de dimension finie d'une réalisation d'un processus stochastique.*

Une série temporelle est, donc toute suite d'observation indexé par le temps t .

1.1.2 Applications et exemples

Une série temporelle peut agir des données :

1. Un espace de probabilité est un triplet $(\Omega, \mathcal{A}, \mathbb{P})$ où (Ω, \mathcal{A}) est un espace mesurable et \mathbb{P} une probabilité sur \mathcal{A}

- macro-économiques telles que le PIB, le taux d'inflation, le niveau des exportations, le taux de chômage, ...

Dans la figure suivante, on trace (à gauche) l'évolution du PIB tunisien au prix de marché en millions de dinars du premier trimestre 2000 au dernier trimestre 2011.

A droite, on donne l'évolution des exportations tunisiennes en valeur.

```
pib <- read.table("http://hamrita.e-monsite.com/medias/files/tunpib.txt", skip = 2)
export <- read.table("http://hamrita.e-monsite.com/medias/files/exportation.txt",
  skip = 2)
pib <- ts(pib, start = 2000, freq = 4)
Export <- ts(export, start = c(1993, 1), freq = 12)
par(mfrow = c(1, 2))
plot(pib, ylab = "PIB", main = "PIB au pm en MD")
plot(Export, ylab = "", main = "Exportations en valeur")
```

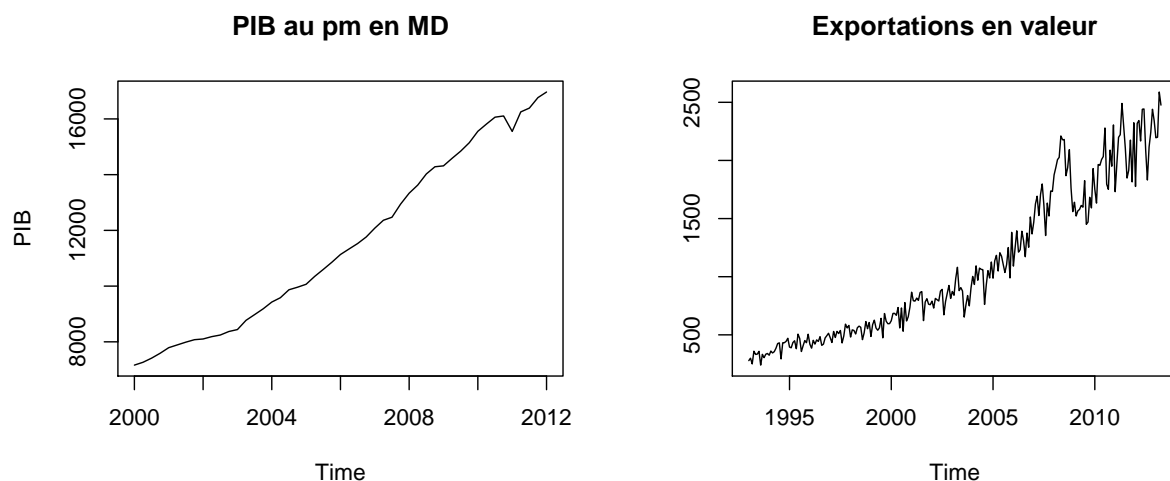


FIGURE 1.1 – Evolution du PIB et Exportations

- micro-économique : les ventes d'une entreprise, nombres d'employés, le revenu des employés, ...
- financières telles que le prix d'un actif financier, le rendement d'un actif, le prix d'une option de vente ou d'achat, ...

```
sp <- read.table("http://hamrita.e-monsite.com/medias/files/sp500.txt")
sp <- ts(sp, start = c(2000, 1), freq = 12)
returns <- diff(log(sp))
par(mfrow = c(1, 2))
plot(sp, ylab = "", xlab = "", main = "Le cours du SP500", col = "red", panel.first = grid())
plot(returns, xlab = "", ylab = "", main = "Le rendement du SP500", col = "red",
  panel.first = grid())
```

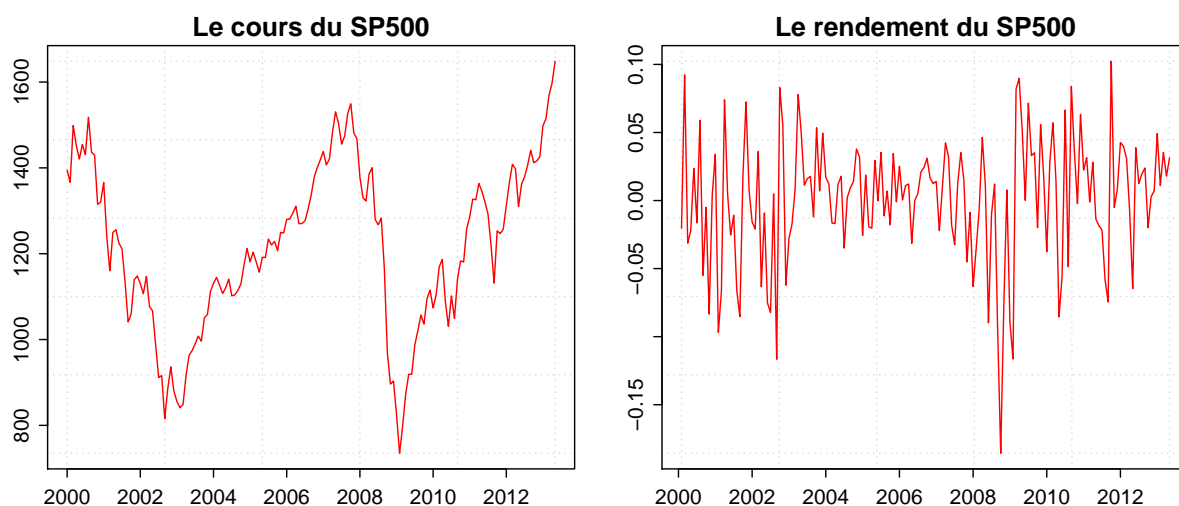


FIGURE 1.2 – Le cours du SP500 et son rendement

– politiques telles que le nombre de votants dans une élection, ...

1.2 Objectifs principaux

On peut poser quelques problèmes à résoudre lorsqu'on observe une série temporelle.

Tout d'abord, la **prévision** qui consiste à prévoir les valeurs futures de la série temporelle en question à partir de ses valeurs observées.

Un autre problème intéressant est la **détection des ruptures** résultants, par exemple, d'un changement d'une politique. La prévision de ces dates de ruptures est bien évidemment très importante.

La détermination de la **causalité** est considérée comme un problème important dans l'analyse d'une série temporelle. Pour qu'un mouvement en provoque un autre, il est nécessaire qu'il le précède.

1.3 Stationnarité, ergodicité et représentation de Wold :

La stationnarité d'une série temporelle en est une propriété fondamentale. Elle indique si les caractéristiques de celui-ci changent avec le temps ou non. Si le processus n'est pas stationnaire alors les informations recueillies dans les données peuvent être mal interprétées. Par exemple, si on observe dans les données deux phases avec des moyennes clairement distinctes, alors les informations telle que la loi marginale ou la structure de corrélation, calculées sur l'ensemble des données n'ont plus un sens. Il faudra dans ce cas séparer les phases et utiliser un modèle différent pour chacune d'entre elles. Formellement, il existe deux types de stationnarité pour un processus $\{X_t\}_t$.

Définition 1.3 (Stationnarité au sens strict). Une série temporelle $\{X_t\}_t$ est dite **strictement** stationnaire si, pour des entiers arbitraires $m_1 < m_2 < \dots < m_n$, la distribution jointe de $X_{t-m_1}, X_{t-m_2}, \dots, X_{t-m_n}$ ne dépend pas du temps t .

En pratique, cette définition s'avère trop restrictive, c'est pourquoi on se limitera à la stationnarité au sens large.

Définition 1.4 (Stationnarité au sens large). Une série temporelle $\{X_t\}_t$ est dite stationnaire au **sens large** (ou faiblement stationnaire ou stationnaire de second ordre) si :

$E(X_t) = \mu$ (constante) $\forall t$ et $E(X_t^2) < \infty$ et ne dépend pas du temps t et

$cov(X_t, X_{t-k}) = \gamma(k)$ ne dépend pas du temps t .

$\gamma(k)$ est dite la fonction d'auto-covariance et $\rho(k) = \gamma(k)/\gamma(0) = corr(X_t, X_{t-k})$ est dite la fonction d'auto-corrélation (ACF). Il est à noter que $\gamma(-k) = \gamma(k)$, $\rho(-k) = \rho(k)$, $\gamma(0) = V(X_t)$ et $\rho(0) = 1$.

Outils graphique pour la stationnarité : On dispose d'une trajectoire d'un processus $\{X_t\}$ et on veut se faire une première idée de la stationnarité de ce processus par l'observation du chrono-gramme (graphique) de la trajectoire. Une condition nécessaire de stationnarité est que la moyenne et la variance de la série soient constantes. Elle implique donc que le graphe de la série en fonction du temps montre un niveau moyen à peu près constant et des fluctuations à peu près de même ampleur autour de la moyenne supposée, quelle que soit la date autour de laquelle on examine la série.

Exemple 1.1 Exemples des processus stationnaires

Bruit blanc faible (Weak White Noise) : Un processus ϵ_t est dit bruit blanc faible si :

- $E(\epsilon_t) = 0$ et $V(\epsilon_t) = \sigma^2 < \infty$ pour tout t .
- $E(\epsilon_t \epsilon_{t-k}) = 0$ pour tout t et $k \neq 0$.

Bruit blanc fort (Strong White Noise) : Un processus ϵ_t est dit bruit blanc fort (ou bruit blanc indépendant) si'il est un bruit blanc faible et les variables aléatoires ϵ_t sont indépendantes.

Bruit blanc gaussien (Gaussian White Noise) : Un processus ϵ_t est dit bruit blanc gaussien si : $\epsilon_t \stackrel{iid}{\sim} N(0, \sigma^2)$

On donne ici le graphique d'un $BB(0, 1)$

```
set.seed(123)
bb <- rnorm(500, 0, 1)
plot.ts(bb, xlab = "", ylab = "", main = "Réalisation d'un BB(0,1)", panel.first = grid(),
        col = "red")
```

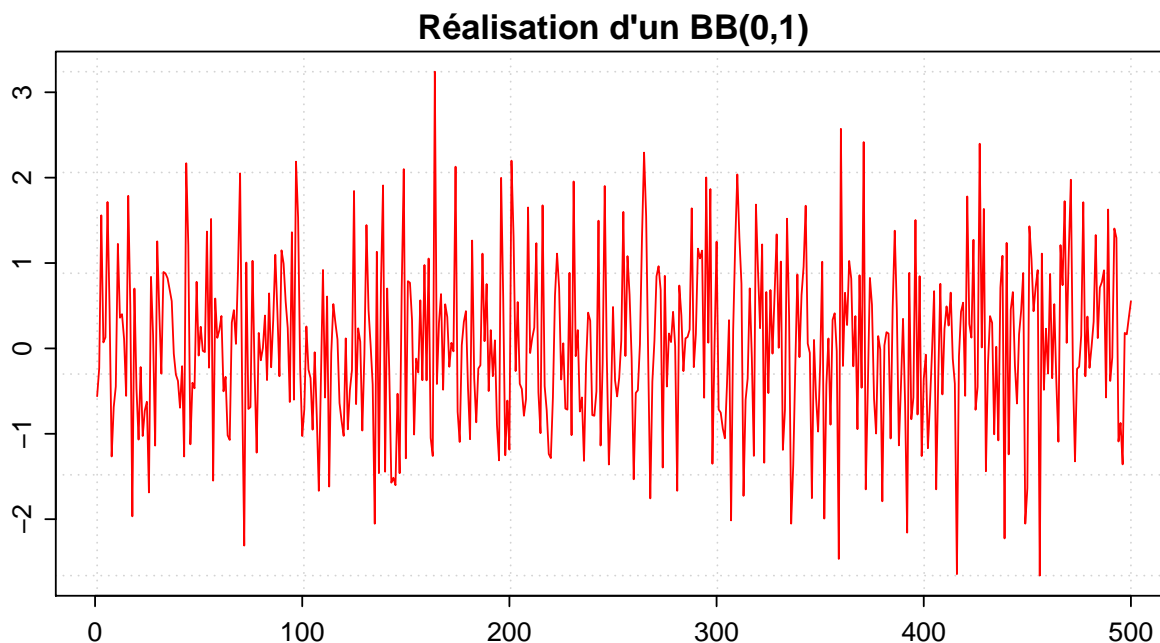


FIGURE 1.3 – Simulation d'un BB(0,1)

Exemple 1.2 Processus non stationnaire

a) Processus marche aléatoire (random walk) : $X_t = X_{t-1} + \epsilon_t$ où ϵ_t est un $BB(0,1)$.

b) Processus avec rupture :

$$X_t = \begin{cases} \alpha + \epsilon_t, & \text{si } t \leq k \\ \alpha + \beta + \epsilon_t, & \text{si } t > k \end{cases}$$

```
X1 <- cumsum(rnorm(500)) # marche aléatoire.
X2 <- c(rnorm(200, 0, 10), 40 + rnorm(300, 0, 10)) # processus avec rupture.
plot.ts(X1, xlab = "", ylab = "", main = "Processus marche aléatoire")
plot.ts(X2, xlab = "", ylab = "", main = "Processus avec rupture")
```

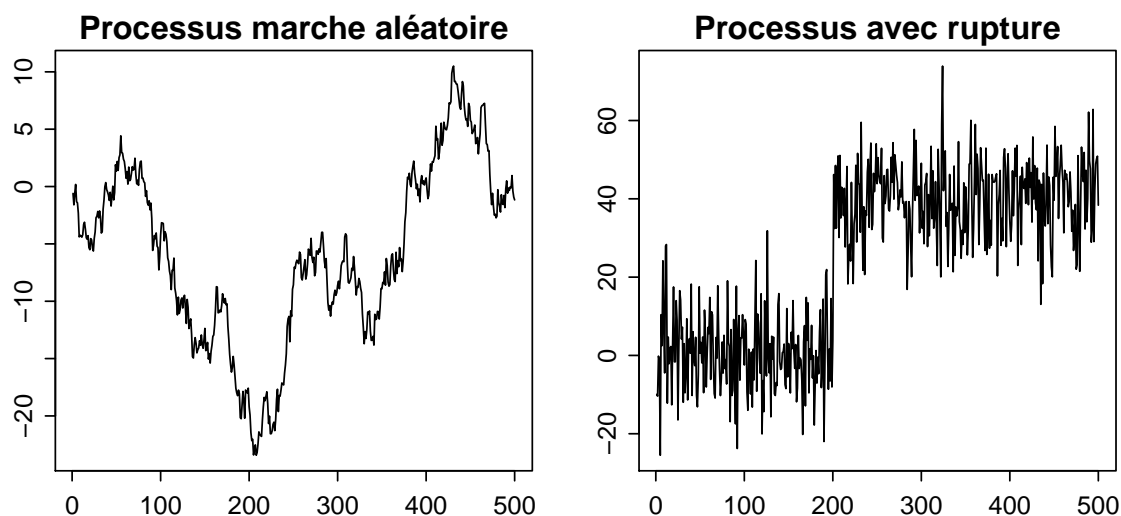


FIGURE 1.4 – Exemples des processus non stationnaires

Définition 1.5 (Ergodicité). Une série temporelle stationnaire est **ergodique** si $\gamma(k) \rightarrow 0$ quand $k \rightarrow \infty$.

Théorème 1.1. Si $\{X_t\}_t$ est strictement stationnaire et ergodique et $E(|X_t|) < \infty$, alors, lorsque $T \rightarrow \infty$:

$$\frac{1}{T} \sum_{t=1}^T X_t \xrightarrow{p} E(X_t)$$

Ce résultat nous permet de déduire des estimateurs consistants des différentes statistiques :

La moyenne empirique : $\hat{\mu} = \bar{X} = \frac{1}{T} \sum_{t=1}^T X_t.$

L'auto-covariance empirique : $\hat{\gamma}(k) = \frac{1}{T} \sum_{t=1}^{T-|k|} (X_t - \bar{X})(X_{t-|k|} - \bar{X}).$

L'auto-corrélation empirique : $\hat{\rho}(k) = \frac{\hat{\gamma}(k)}{\hat{\gamma}(0)}.$

Pour un échantillon de taille T assez large, les corrélations empiriques $\hat{\rho}(1), \dots, \hat{\rho}(k)$ sont, approximativement, *iid* suivant une loi normale de moyenne nulle et de variance T^{-1} , i.e :

$$\hat{\rho}(k) \stackrel{iid}{\sim} N(0, T^{-1})$$

D'où, $t_{\rho(k)} = \sqrt{T}(\hat{\rho}(k) - \rho(k)) \stackrel{\mathcal{L}}{\sim} N(0, 1).$

On déduit aisément la région de confiance pour une hypothèse nulle $\rho(k) = 0$ avec un seuil d'erreur $\alpha = 5\%$:

$$IC = \left[\pm 1.96 \times \sqrt{T} \right]$$

Avec R, on peut utiliser la fonction `acf()` qui représente le corrélogramme empirique d'une série temporelle ainsi que les deux bornes de l'intervalle de confiance des corrélations.

Dans le graphique suivant, on a tracé 40 premières corrélations empiriques d'un échantillon de 200 observations d'une loi $N(0, 1)$.

```
set.seed(123)
X <- rnorm(200) # Simulation de 200 observations d'une loi N(0,1).
acf(X, lag.max = 40, ylim = c(-0.6, 1.02), main = "", col = 2)
```

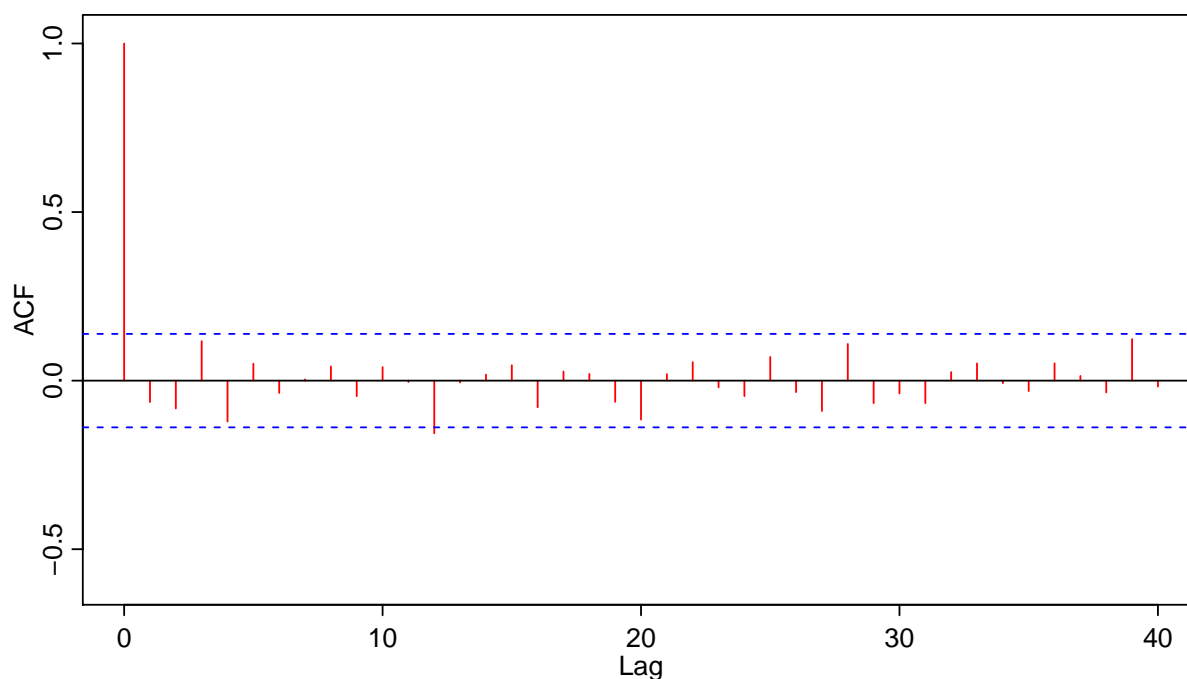


FIGURE 1.5 – Corrélogramme empirique d'un bruit blanc

Il est clair que toutes les corrélations varient entre les bornes $\pm 1.96\sqrt{200}$ sans les dépasser (sauf la première corrélation $\hat{\rho}(0) = 1$).

Théorème 1.2 (représentation de Wold ([Wold, 1938](#))). *Toute série temporelle $\{X_t\}_t$ stationnaire au sens large peut être représentée sous la forme :*

$$X_t - \mu = \sum_{j=0}^{\infty} \psi_j \epsilon_{t-j} \quad (1.3.1)$$

où les paramètres ψ_j satisfont : $\psi_0 = 1$, $\psi_j \in \mathbb{R} \forall j \in \mathbb{N}$ et $\sum_{j=0}^{\infty} \psi_j^2 < \infty$ et avec $\epsilon_t \stackrel{iid}{\sim} N(0, \sigma^2)$.

On dit que la somme des chocs passés correspond à la composante linéaire stochastique de X_t . On peut aisément montrer que l'équation (1.3.1) est très utile pour la détermination des auto-corrélations d'un processus x_t qui admet cette représentation. Depuis cette équation, il en suit :

$$\begin{aligned} E(x_t) &= \mu \\ \gamma(0) &= V(x_t) = E(x_t - \mu)^2 \\ &= E(\epsilon_t + \epsilon_{t-1} + \epsilon_{t-2} + \dots)^2 \\ &= E(\epsilon_t^2) + E(\epsilon_{t-1}^2) + E(\epsilon_{t-2}^2) + \dots \\ &= \sigma^2 + \sigma^2 \psi_1^2 + \sigma^2 \psi_2^2 + \dots \\ &= \sigma^2 \sum_{j=0}^{\infty} \psi_j^2. \end{aligned}$$

en utilisant la condition $E(\epsilon_t \epsilon_{t-k}) = 0$ pour tout $t \neq k$.

$$\begin{aligned} \gamma(k) &= E(x_t - \mu)(x_{t-k} - \mu) \\ &= E(\epsilon_t + \psi_1 \epsilon_{t-1} + \dots + \psi_k \epsilon_{t-k} + \dots)(\epsilon_{t-k} + \psi_1 \epsilon_{t-k-1} + \dots) \\ &= \sigma^2(\psi_k + \psi_1 \psi_{k+1} + \psi_2 \psi_{k+2} + \dots) \\ &= \sigma^2 \sum_{j=0}^{\infty} \psi_j \psi_{j+k}. \end{aligned}$$

Ce qui nous donne :

$$\rho(k) = \frac{\gamma(k)}{\gamma(0)} = \frac{\sum_{j=0}^{\infty} \psi_j \psi_{j+k}}{\sum_{j=0}^{\infty} \psi_j^2}.$$

Si les coefficients ψ_j sont infinis, on peut supposer qu'ils sont absolument sommable, i.e., $\sum |\psi_j| < \infty$. On peut démontrer que cette condition est équivalente à celle de stationnarité du processus x_t , et garantie l'existence et l'indépendance du temps de tous les moments du processus.

1.4 Opérateurs retard, avance et différence

Définition 1.6.

i) On appelle **opérateur retard** l'opérateur L (Lag) ou B (Backward) qui à tout processus X_t , on associe le processus Y_t défini par :

$$\forall t \in \mathbb{Z} Y_t = LX_t = X_{t-1}$$

Avec R , on fabrique des séries retardées à l'aide de la fonction `lag` (si la série est de classe `ts`) ou la fonction `Lag` de la librairie (package) `Hmisc`.

```
set.seed(123)
x <- as.ts(rnorm(6)) # génération de 6 valeurs aléatoires normales de classe ts.
lag1.x <- lag(x, -1) # série retardée d'une seule observation.
lag2.x <- lag(x, -2) # série retardée de deux observations.
(result <- t(cbind(x, lag1.x, lag2.x)))

##           [,1]      [,2]      [,3]      [,4]      [,5]      [,6]      [,7]      [,8]
## x          -0.5605 -0.2302  1.5587  0.07051 0.12929 1.71506      NA      NA
## lag1.x       NA -0.5605 -0.2302  1.55871 0.07051 0.12929 1.7151      NA
## lag2.x       NA      NA -0.5605 -0.23018 1.55871 0.07051 0.1293 1.715

res2 <- ts.intersect(x, lag1.x, lag2.x)
t(res2)

##           [,1]      [,2]      [,3]      [,4]
## x          1.5587  0.07051 0.12929 1.71506
## lag1.x     -0.2302  1.55871 0.07051 0.12929
## lag2.x     -0.5605 -0.23018 1.55871 0.07051
```


ii) De la même façon, on appelle **opérateur avance** l'opérateur F (forward) qui a tout processus X_t , on associe le processus Y_t défini par :

$$\forall t \in \mathbb{Z} Y_t = F X_t = X_{t+1}$$

iii) On appelle **opérateur différence** l'opérateur ∇ qui a tout processus X_t , on associe le processus Y_t défini par :

$$\nabla X_t = Y_t = (1 - L)X_t = X_t - X_{t-1}.$$

Avec **R**, on applique l'opérateur différence à l'aide de la fonction `diff`.

```
set.seed(123)
xt <- as.ts(rnorm(5))
xt_1 <- lag(xt, -1) # série retardée d'une observation
xt_2 <- lag(xt, -2) # série retardée de 2 observations
xt - xt_1 # première différence

## Time Series:
## Start = 2
## End = 5
## Frequency = 1
## [1] 0.33030 1.78889 -1.48820 0.05878

xt_2 - 2 * xt_1 + xt # différence d'ordre 2: (1-L)^2=L^2-2*L+1

## Time Series:
## Start = 3
## End = 5
## Frequency = 1
## [1] 1.459 -3.277 1.547

(deltaX <- diff(xt)) # première différence

## Time Series:
## Start = 2
## End = 5
## Frequency = 1
## [1] 0.33030 1.78889 -1.48820 0.05878

(delta2X <- diff(xt, differences = 2)) # différence d'ordre 2

## Time Series:
## Start = 3
## End = 5
## Frequency = 1
## [1] 1.459 -3.277 1.547
```

Propriétés 1.1.

i) $L^0 X_t = X_t$.

ii) Si on compose n fois l'opérateur L dans lui même, on obtient :

$$\underbrace{L \circ L \circ L \circ \dots \circ L}_n = L^n \text{ et } L^n X_t = X_{t-n}.$$

iii) $L(X_t + Y_t) = L X_t + L Y_t$ (opérateur retard est distributif).

iv) $L(\alpha X_t) = \alpha L X_t$. (α est un réel).

v) $\nabla^k X_t = (1 - L)^k X_t$.

On peut définir un polynôme retard tout polynôme qui s'écrit en fonction de l'opérateur retard. On donne un exemple :

Soit le processus X_t défini par : $X_t = \phi_1 X_{t-1} + \phi_2 X_{t-2} + \epsilon_t$.

Ce processus peut s'écrire comme suit :

$$X_t = \phi_1 L X_t + \phi_2 L^2 X_t + \epsilon_t = (\phi_1 L + \phi_2 L^2) X_t + \epsilon_t$$

D'où $(1 - \phi_1 L - \phi_2 L^2) X_t = \epsilon_t \implies \Phi(L) X_t = \epsilon_t$.

$\Phi(L) = (1 - \phi_1 L - \phi_2 L^2)$ est le polynôme retard associé au processus X_t .

Exercices

Exercice 1.1

Soit ϵ_t un $BB(0,1)$. Étudier la stationnarité des processus suivants :

$$(a) x_t = (-1)^t \epsilon_t \quad (b) y_t = x_t + \epsilon_t$$

Avec R, simuler 500 observations de chacun des deux processus et les représenter graphiquement.

Exercice 1.2

Soit $(\epsilon_t)_t$ un BB gaussien *i.i.d* de moyenne nulle et variance 1, et a, b et c des constantes réelles. Les processus suivants sont-ils stationnaires (au second ordre)? Si oui, donner leur moyenne et leur fonction d'auto-covariance.

$$(a) X_t = a + b\epsilon_t + c\epsilon_{t-1}, \quad (b) Y_t = \epsilon_1 \cos(ct) + \epsilon_2 \sin(ct)$$

$$(c) Z_t = \epsilon_t \cos(ct) + \epsilon_{t-1} \sin(ct), \quad (d) W_t = \epsilon_0 \cos(ct)$$

Simuler 500 observations de chacun de ces processus et les représenter graphiquement et représenter leurs corrélogrammes empiriques.

Exercice 1.3

Soit X_t un processus stationnaire (au sens faible). Montrer par récurrence que $\nabla^k X_t$ est stationnaire. ($\nabla^k = (1-L)^k$ où L est l'opérateur retard).

Exercice 1.4

1 désigne la fonction constante et t la fonction : $t \rightarrow t$. Calculer :

$$1) (1 - 0.75L)1, \quad 2) (1 - 0.75L)t, \quad 3) \nabla^2 t$$

Exercice 1.5

Nous avons simulé les séries suivantes, où ϵ_t est un bruit blanc.

$$X_t = 4 + \epsilon_t; \quad X_t = X_{t-1} + \epsilon_t;$$

$$X_t = \begin{cases} \epsilon_t, & \text{si } t \leq 300; \\ 140 + \epsilon_t, & \text{si } 300 < t \leq 500. \end{cases} ; \quad X_t = 0.5 + 0.7t + \epsilon_t$$

Associer à chaque série sa représentation graphique, en justifiant.

