

République Algérienne Démocratique et Populaire
Ministère de l'Enseignement Supérieur et de la Recherche Scientifique
Université de GHARDIAIA



Faculté des Sciences et de la Technologie
Département de Mathématiques et Informatique
Laboratoire de Mathématiques et Sciences Appliquées

Mémoire

Présenté par :

Kessouri Chahra

pour obtenir

le diplôme de Master

Spécialité : Analyse Fonctionnelle et appliquée

THÈME

Processus ARMA : aspect théorique et simulations

Soutenu le : 02/ 07/ 2019

Devant le jury :

Président :	Mr Latreche Smail	M. A(Univ. Ghardaia)
Promoteur :	Mr Lani Abd Nour	M. A(Univ. Ghardaia)
Examinatrice :	Mme Bahdi Bahia	M. A(Univ. Ghardaia)

Année universitaire 2018/2019



Dédicace

Je dédie ce Modeste travail à ceux qui m'ont soutenu au long de mes études, qui m'ont toujours poussé vers le chemin du savoir, à ma source d'amour et d'affection, les deux êtres les plus chères au monde, Mes Chers parents: Mon Père & ma Mère que dieu les miséricorde , À mon Cher frère : Amine , À mes Soeurs : Souhila et Lamia et Fella, À ma grande famille: Kessouri, À tout mes amies À mes amis de l'université de GHARDAIA. À tous les étudiants de math & info . À tous ceux qui me sont cher(e)s.



❖ K. Chahra

Remerciements

Je tiens à témoigner ma reconnaissance à DIEU tout puissant, de m'avoir donné le courage et la force de mener à terme ce projet. Qui m'a ouvert les portes du savoir.

Je remercie mon encadreur Monsieur *Lani Abd nour* pour le sujet qu'il m'a proposé et pour l'attention et la disponibilité dont il a su faire preuve au long de la préparation de ce mémoire.

Je remercie sincèrement les membres de jury pour avoir accepté de juger ce travail.

Je remercie les enseignants du département des mathématiques et d'informatique qui nous ont encadrés depuis notre première année universitaire.

Merci à tous ceux qui m'ont enseigné primaire, collège, lycée, université et ailleurs, et tous les personnels de ces établissements, sans eux je suis rien. Merci à tous les mathématiciens...

merci à tout ceux qui ont contribué de près ou de loin dans ce travail.

Résumé

Les processus ARMA sont très utilisés en pratique

Ces processus trouvent des applications dans divers domaines comme la finance, la météorologie,

En finance, les processus ARMA modélisent les valeurs des actions boursières. En météorologie, on les utilise pour étudier la température du jour dans une région donnée. Le but de cette modélisation est la prévision : se basant sur les valeurs passées du processus et quand celui-ci est stationnaire on peut estimer ses valeurs dans le futur ce qui est cruciale dans la vie quotidienne, malheureusement la stationnarité n'est pas toujours assurée en pratique dans ce cas on utilise des processus plus généraux comme les modèles ARCH et GARCH.

L'outil informatique comme le logiciel R permet de procéder à des simulations.

Abstract

The ARMA processes are very useful in practice.

These processes are applied in diverse fields for example finance, meteorology,....

In finance, ARMA processes modeling the stock values. In meteorology, we use them to study the temperature of the day in a given region.

The purpose of this modeling is the forecast : based on the past values of the processes and when it is stationary we can estimate the future values which is crucial in daily life, unfortunately the stationary state is not always insured in practice in this case we use more general processes : ARCH and GARCH models.

The computer tool as the software R is allowing to simulation.

Table des matières

Introduction	1
1 Préliminaires	2
1.1 Processus	2
1.2 Espérance	2
1.2.1 Espérance d'une variable aléatoire discrète	2
1.2.2 Espérance d'une variable aléatoire continue	2
1.3 Variance	3
1.4 Covariance	3
1.5 Corrélation	3
1.5.1 Le coefficient d'auto corrélation	3
1.5.2 Fonction d'autocorrélation partielle	4
2 Processus stationnaires	8
2.1 Stationnarité stricte	8
2.2 Stationnarité faible	9
2.3 Remarques diverses sur la stationnarité	10
2.4 Les opérateurs B et F	11
3 Exemples de processus	13
3.1 Bruit blanc	13
3.2 Les processus Linéaires	13
3.3 Les processus AR	15
3.3.1 Le système de Yulle-Walker	16
3.3.2 Etude du processus autorégressif (AR)	17
3.3.3 Simulation d'un échantillons de 1000 données du processus de Markov de l'exemple 1 :	18
3.4 Les processus MA	26
3.4.1 Propriétés d'un MA(q)	26
3.5 Remarque sur le lien entre AR et MA	29
3.6 Les processus mixtes ARMA et ARIMA	29

4	Modèles de Box et Jenkins	31
4.1	Estimation des autocorrélations et identification	31
4.2	Estimation des paramètres et validation du modèle	37
4.3	Calcul de prévisions	39
5	Exemples d'applications	43
5.1	L'estimation des coefficients ARMA sous logiciel R	43
	Conclusion	48
	Bibliographie	49

Table des figures

3.1	Grappe la série X_t	19
3.2	Corrélogramme	19
3.3	Corrélogramme partiel	20
3.4	Grappe la série X_t	22
3.5	Corrélogramme	23
3.6	Corrélogramme partiel	23
3.7	24
3.8	Grappe la série X_t	28
3.9	Corrélogramme	28
3.10	Corrélogramme partiel	28
4.1	Évolution de la production	34
4.2	Illustration d'une tendance linéaire estimée par régression.	35
4.3	La série initiale	36
4.4	La série des différences d'ordre 1.	36
4.5	La fonction r_h pour la série des différences d'ordre 1.	37
4.6	La fonction $\varphi_{h,h}$ pour la série des différences d'ordre 1.	37
5.1	Grappe de la série chronologique	43
5.2	Simulation des coefficients du processus ARMA	44
5.3	Grappe de température	46
5.4	Corrélogramme	46
5.5	Corrélogramme partiel	47

Introduction

Les processus stochastiques en général et les séries chronologiques en particulier occupent une place importante en statistique.

Quand on cherche à modéliser une série temporelle, on a généralement recours à la classe des modèles **ARMA**, qui conviennent pour modéliser beaucoup de situations pratiques. C'est le cas en finance où les processus ARMA sont utilisés pour le traitement des données des actions boursières. Ces processus sont utilisés aussi en météorologie pour modéliser la température du jour.

Dans cette classe de modèle, on distingue :

- les processus **AR**(p) : $X_t = \sum_{i=1}^p \phi_i X_{t-i} + \varepsilon_t$
- les processus **AM**(q) : $X_t = \varepsilon_t - \sum_{i=1}^q \theta_i \varepsilon_{t-i}$
- les processus **ARMA**(p,q) : $X_t - \sum_{i=1}^p \phi_i X_{t-i} = \varepsilon_t - \sum_{i=1}^q \theta_i \varepsilon_{t-i}$

où ε_t est un bruit blanc centré de variance σ^2

Les précurseurs de ces méthodes de prévision sont incontestablement Box et Jenkins qui ont donné une technique de modélisation qui a révolutionné la pratique de la statistique, et son utilisation comme outil d'estimation et de prévision. Leur méthode comporte les étapes suivantes :

- Identification
- Estimation
- Validation
- Prévision

L'utilisation de l'outil informatique comme le logiciel libre **R** ont fait évoluer les méthodes statistiques en permettant aux praticiens de procéder à des simulations.

Chapitre 1

Préliminaires

Dans ce chapitre, on donne quelques résultats de probabilités et statistique utiles pour la suite.

1.1 Processus

Définition 1.1.1. *Une série temporelle représente l'évolution d'une variable aléatoire en fonction du temps*

1.2 Espérance

1.2.1 Espérance d'une variable aléatoire discrète

Définition 1.2.1. *Soit une variable aléatoire réelle discrète X . On pose $X(\Omega) = \{x_i : i \in I\}$ où I est un ensemble dénombrable d'indices. On dit que X admet une espérance si l'on a*

$$\sum |x_i| \mathbb{P}(X = x_i) < \infty.$$

Dans ce cas, l'espérance de X est égale à

$$E[X] := \sum_{i \in I} x_i \mathbb{P}(X = x_i).$$

1.2.2 Espérance d'une variable aléatoire continue

Définition 1.2.2. *Soit une variable aléatoire réelle absolument continue X . Soit f la densité de probabilité de la variable aléatoire réelle X . On dit que X admet une espérance si l'on a*

$$\int_{\mathbb{R}} |x| f(x) dx < \infty.$$

Dans ce cas, l'espérance de X est égale à

$$E[x] := \int_{\mathbb{R}} x f(x) dx.$$

1.3 Variance

Définition 1.3.1. La variance de la série statistique se note σ_x^2 (ou encore $V(x)$) et se définit comme suit :

$$V(x) = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2$$

Elle correspond à la moyenne des carrés des différences entre les observations et leur moyenne \bar{x} .

Dans le cas de n observations, ordonnées dans un tableau statistique (x_i, n_i) , présentant r modalités.

$$V(x) = \frac{1}{n} \sum_{i=1}^n n_i (x_i - \bar{x})^2$$

La variance (ou fluctuation) est la moyenne arithmétique des carrés des écarts à la moyenne.

1.4 Covariance

Définition 1.4.1. La covariance de deux variables aléatoires réelles X et Y ayant chacune une variance, notée $Cov(X, Y)$ ou parfois σ_{XY} , est la valeur :

$$Cov(X, Y) = E[(X - E[X])(Y - E[Y])]$$

La variance de X est donc $var(x) = cov(x, x)$. Pour deux variables discrètes X et Y prenant respectivement leurs valeurs dans deux ensembles finis $\{x_i | 1 \leq i \leq n\}$, et $\{y_j | 1 \leq j \leq m\}$, on a

$$Cov(x, y) = \sum_{i=1}^n \sum_{j=1}^m x_i y_j P(X = x_i \text{ et } Y = y_j) - E[X]E[Y].$$

Tandis que :

$$\sigma_x^2 = \sum_{i=1}^n x_i^2 P(X = x_i) - E[X]^2.$$

et

$$\sigma_y^2 = \sum_{j=1}^m x_j^2 P(X = y_j) - E[Y]^2.$$

1.5 Corrélation

1.5.1 Le coefficient d'auto corrélation

Définition 1.5.1. Soit le processus $(X_t, t \in T)$ on appelle coefficient d'autocorrélation pour les temps t et s la valeur

$$\rho(X_t, X_s) = \frac{Cov(X_t, X_s)}{\sqrt{Var(X_t)Var(X_s)}}$$

Si le processus X_t est faiblement stationnaire $\rho(X_t, X_s)$ est défini pour les temps t et $t+h$ par

$$\rho(h) = \rho(X_t, X_{t+h}) = \frac{\text{Cov}(X_t, X_{t+h})}{\sqrt{\text{Var}(X_t)\text{Var}(X_{t+h})}}$$

et puisque $\text{Var}(X_t) = \text{Var}(X_{t+h}) = \gamma(0)$ alors,

$$\rho(h) = \frac{\gamma(h)}{\gamma(0)}$$

Propriété

Si le processus X_t est faiblement stationnaire, alors

$$\rho(h) = \rho(-h) \text{ pour tout } h \in \mathbb{Z}$$

Soient alors $(k+1)$ valeurs successives du processus, notées X_t, \dots, X_{t+k} on appelle matrice de Toeplitz la matrice des autocorrélations de (X_t, \dots, X_{t+k}) (donc symétrique)

1.5.2 Fonction d'autocorrélation partielle

Elle permet de quantifier les dépendances qui existent entre les observations du processus stationnaire

Pour un processus stationnaire au second ordre

$$\begin{aligned} \rho(h) &= \frac{\text{Cov}(X_t, X_{t-h})}{\sqrt{\text{Var}(X_t)\text{Var}(X_{t-h})}} \\ &= \frac{\text{Cov}(X_t, X_{t-h})}{\text{Var}(X_t)} \end{aligned}$$

$$\rho(0) = 1, \quad \rho(h) = \frac{\gamma(h)}{\gamma(0)}, \quad \rho(h) = \rho(-h)$$

Comme les coefficients $\rho(h)$, $h = 1, 2, \dots$ ne sont pas observables, on les estime

$$\rho(h) = \frac{\sum_{t=h+1}^T (X_t - \bar{X})(X_{t-h} - \bar{X})}{\sum_{t=1}^T (X_t - \bar{X})^2} = r_h \quad (1.1)$$

C'est la fonction d'autocorrélation empirique ; $h=1,2,\dots$

Remarque :

Lorsque le processus est gaussien, Bartlett (1946) à déterminé les variances et les covariances asymptotiques de ces estimateurs.

$$\begin{aligned} 1). \text{ var}(r_h) &\simeq \frac{1}{T} \sum_{j=-\infty}^{+\infty} (\rho_j^2 + \rho_{h+j}\rho_{-h+j} - 4\rho_h\rho_j\rho_{-h+j} + 2\rho_j^2\rho_h^2) \\ &\simeq \frac{1}{T} \sum_{j=-\infty}^{+\infty} \rho_j^2 \\ &= \frac{1}{T} (1 + 2 \sum_{j=1}^{+\infty} \rho_j^2) \end{aligned}$$

$$2). \text{ cov}(r_h, r_{h+k}) \simeq \frac{1}{T} \sum_{j=-\infty}^{+\infty} (\rho_j\rho_{j+k}); h, k \in \mathbb{Z}$$

Comme le processus est supposé stationnaire $h, k \in \mathbb{N}$

* pour T grand, r_h suit asymptotiquement une loi normale

$$\frac{r_h - \rho_h}{\sqrt{\text{var}(r_h)}} \sim N(0, 1)$$

Il existe une représentation matricielle qui donne les relations entre les auto corrélations et les autocorrélations partielles -elle est donnée par les équations de Yule-Walker pour un processus stationnaire du second ordre .

Soit le modèle $X_t = \varphi_{h1}X_{t-1} + \dots + \varphi_{hh}X_{t-h} + \varepsilon_t$ la fonction d'auto covariance d'ordre l .

$$\begin{aligned} \gamma_l &= \text{cov}(X_t, X_{t-l}) \\ &= \text{cov}(\varphi_{h1}X_{t-1} + \dots + \varphi_{hh}X_{t-h} + \varepsilon_t, X_{t-l}) \\ \gamma_l &= \varphi_{h1}\gamma_{l-1} + \varphi_{h2}\gamma_{l-2} + \dots + \varphi_{hh}\gamma_{l-h} \\ &= \text{cov}(\varepsilon_t, X_{t-l}) = 0, \text{ si } l > 0 \end{aligned}$$

En faisant varier l, $l = 1, \dots, h$ on obtient le système d'équations suivant appelé système de Yule-Walker :

$$\begin{aligned}
\gamma_1 &= \varphi_{h1}\gamma_0 + \varphi_{h2}\gamma_1 + \dots + \varphi_{hh}\gamma_{h-1} \\
\gamma_2 &= \varphi_{h1}\gamma_1 + \varphi_{h2}\gamma_0 + \dots + \varphi_{hh}\gamma_{h-2} \\
&\cdot \\
&\cdot \\
&\cdot \\
\gamma_h &= \varphi_{h1}\gamma_{h-1} + \varphi_{h2}\gamma_{h-2} + \dots + \varphi_{hh}\gamma_0
\end{aligned}$$

On encore, en utilisant les auto corrélations; $\rho_0 = 1$

$$\left\{ \begin{array}{l}
\rho_1 = \varphi_{h1}\rho_0 + \varphi_{h2}\rho_1 + \dots + \varphi_{hh}\rho_{h-1} \\
\rho_2 = \varphi_{h1}\rho_1 + \varphi_{h2}\rho_0 + \dots + \varphi_{hh}\rho_{h-2} \\
\cdot \\
\cdot \\
\cdot \\
\rho_h = \varphi_{h1}\rho_{h-1} + \varphi_{h2}\rho_{h-2} + \dots + \varphi_{hh}\rho_0
\end{array} \right.$$

On encore, matériellement :

$$\begin{pmatrix} \rho_1 \\ \cdot \\ \cdot \\ \cdot \\ \cdot \\ \rho_h \end{pmatrix} = \begin{pmatrix} 1 & \rho_1 & \cdot & \cdot & \cdot & \rho_{h-1} \\ \rho_1 & 1 & \rho_1 & \cdot & \cdot & \rho_{h-2} \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \rho_{h-1} & \cdot & \cdot & \cdot & \cdot & 1 \end{pmatrix} \begin{pmatrix} \varphi_{h1} \\ \cdot \\ \cdot \\ \cdot \\ \cdot \\ \varphi_{hh} \end{pmatrix} \quad (1.2)$$

φ_{hh} est auto corrélation partielle d'ordre h

$$\begin{pmatrix} \varphi_{h1} \\ \cdot \\ \cdot \\ \cdot \\ \cdot \\ \varphi_{hh} \end{pmatrix} = \begin{pmatrix} 1 & \rho_1 & \cdot & \cdot & \cdot & \rho_{h-1} \\ \rho_1 & 1 & \rho_1 & \cdot & \cdot & \rho_{h-2} \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \rho_{h-1} & \cdot & \cdot & \cdot & \cdot & 1 \end{pmatrix}^{-1} \begin{pmatrix} \rho_1 \\ \cdot \\ \cdot \\ \cdot \\ \cdot \\ \rho_h \end{pmatrix}$$

$$\varphi_{hh} = \frac{\begin{vmatrix} 1 & \cdot & \cdot & \dots & \rho_{h-2} & \rho_1 \\ \cdot & & & & \cdot & \cdot \\ \cdot & & & & \cdot & \cdot \\ \varphi_{h-1} & & & & \rho_1 & \rho_h \end{vmatrix}}{\begin{vmatrix} 1 & \rho_1 & \cdot & \dots & \rho_{h-1} \\ \cdot & & & & \cdot & \cdot \\ \cdot & & & & \cdot & \cdot \\ \varphi_{h-1} & & & & \rho_1 & 1 \end{vmatrix}}$$

$$\varphi_{11} = \frac{|\rho_1|}{|1|} = \rho_1, \quad \varphi_{22} = \frac{\begin{vmatrix} 1 & \rho_1 \\ \rho_1 & \rho_2 \end{vmatrix}}{\begin{vmatrix} 1 & \rho_1 \\ \rho_1 & 1 \end{vmatrix}} = \frac{\rho_2 - \rho_1^2}{1 - \rho_1^2}, \quad \varphi_{33} = \frac{\begin{vmatrix} 1 & \rho_1 & \rho_1 \\ \rho_1 & 1 & \rho_2 \\ \rho_2 & \rho_1 & \rho_3 \end{vmatrix}}{\begin{vmatrix} 1 & \rho_1 & \rho_2 \\ \rho_1 & 1 & \rho_1 \\ \rho_2 & \rho_1 & 1 \end{vmatrix}},$$

Il existe un algorithme récursif (Algorithme de Durbin (1960)) qui fournit les auto corrélations partielles

$$\varphi_{11} = \rho_1$$

$$\varphi_{h,h} = \frac{\rho_h - \sum_{j=1}^{h-1} \varphi_{h-1,j} \rho_{h-j}}{1 - \sum_{j=1}^{h-1} \varphi_{h-1,j} \rho_j}; \quad (1.3)$$

avec $\varphi_{h,j} = \varphi_{h-1,j} - \varphi_{h,h}\varphi_{h-1,h-j}$ $h = 2, 3, 4, \dots$ $j = 1, 2, 3, \dots, h-1$.

Chapitre 2

Processus stationnaires

Dans l'étude des processus stochastiques, une place particulièrement importante est tenue par les processus dont les lois de probabilité présentent une invariance pour toute translation dans le temps. Cette propriété d'invariance temporelle est couramment utilisée en économétrie, en analyse statistique des séries temporelles à l'aide de processus auto-régressifs moyennes mobiles (ARMA).

2.1 Stationnarité stricte

Définition 2.1.1. *Le processus réel $(X_t, t \in T)$ est dit strictement stationnaire si pour tout n -uplet de temps $t_1 < t_2 < \dots < t_n$ avec $t_i = 1 \dots n$, la suite $(X_{t_1+h}, X_{t_2+h}, \dots, X_{t_n+h})$ a la même loi de probabilité que la suite $(X_{t_1}, X_{t_2}, \dots, X_{t_n})$ c'est à dire*

$$P(X_{t_1+h}, X_{t_2+h}, \dots, X_{t_n+h}) = P(X_{t_1}, X_{t_2}, \dots, X_{t_n}) \quad \forall h \in T$$

où $T \subset \mathbb{Z}$ (X_t processus discrète)

Remarque :

Puisqu'une distribution de probabilité est caractérisée par sa fonction de répartition, la définition précédente est équivalente à pour tous x_1, x_2, \dots, x_n , tous t_1, t_2, \dots, t_n et tout h

$$P(X_{t_1+h} < x_1, X_{t_2+h} < x_2, \dots, X_{t_n+h} < x_n) = P(X_{t_1} < x_1, X_{t_2} < x_2, \dots, X_{t_n} < x_n) \text{ pour tout } h \in T$$

Théorème 2.1.1. *Si $(X_t, t \in T)$ avec $(T = \mathbb{R}, \mathbb{N} \text{ ou } \mathbb{Z})$ est stationnaire, alors on a en particulier*

1. $E(X_t) = m \quad \forall t \in T$
2. $Var(X_t) = \sigma^2 \quad \forall t \in T$
3. $Cov(X_t, X_s) = v(|t - s|) \quad \forall (t, s) \in T^2$

2.2 Stationnarité faible

Définition 2.2.1. *Le processus $(X_t, t \in T)$ est dit faiblement stationnaire ou stationnaire ou second ordre s'il vérifie les trois conditions suivantes :*

1. $E(X_t) = m \quad \forall t \in T$
2. $E(X_t^2) < +\infty \quad \forall t \in T$
3. $Cov(X_t, X_{t-h}) = cov(X_{t+s}, X_{t+s-h}) = v(h) \quad \forall s, t, h \in T$

* la première condition signifie que le processus est stationnaire en moyenne.

* la deuxième condition signifie que la variance du processus existe et est finie.

* la troisième condition signifie que la covariance est invariante dans le temps et quelle ne dépend que du pas h .

$$v(0) = var(X_t) = \text{constante} = \sigma_X^2$$

Définition 2.2.2. $v(h)$ porte le nom de fonction d'auto-covariance du processus

Propriété

1. $v(0) = var(X_t)$
2. $v(h) = v(-h) \quad \forall h \in \mathbb{Z}$
3. $|v(h)| \leq v(0)$

Preure :

$$\begin{aligned}
 1). \quad \nu(0) &= Cov(X_t, X_t) \\
 &= E[(X_t - \mu)(X_t - \mu)] \\
 &= E[(X_t - \mu)^2] \\
 &= var(X_t) \\
 2). \quad \nu(h) &= Cov(X_t, X_{t-h}) \\
 &= Cov(X_{t+h}, X_{t+h-h}) \\
 &= Cov(X_{t+h}, X_t) \\
 &= Cov(X_t, X_{t-(-h)}) \\
 &= \nu(-h) \\
 3). \quad |\nu(h)| &= |Cov(X_t, X_{t-h})| \\
 &= |E[(X_t - E(X_t))(X_{t-h} - E(X_{t-h}))]| \\
 &= |E[(X_t - m)(X_{t-h} - m)]| \\
 &\leq |(E[(X_t - m)^2])^{\frac{1}{2}}(E[(X_{t-h} - m)^2])^{\frac{1}{2}}| \\
 &\leq (\nu(0))^{\frac{1}{2}}(\nu(0))^{\frac{1}{2}} \\
 &\leq \nu(0)
 \end{aligned}$$

2.3 Remarques diverses sur la stationnarité

a- Exemple de processus non stationnaire

Considérons le processus $X_t = at + \varepsilon_t$ où les variables aléatoires, ε_t sont indépendantes et identiquement distribuées (i.i.d) d'espérance nulle et de variance σ^2 . Le processus X_t n'est pas faiblement stationnaire car $E(X_t) = at + b$ pour le rendre stationnaire, considérons le processus différence première :

$$Y_t = X_t - X_{t-h} = a - \varepsilon_t - \varepsilon_{t-h}$$

On a $E(Y_t) = a$, $Var(Y_t) = 2\sigma^2$ pour tout $t \in T$ et d'après les hypothèses sur ε_t

$$Cov(Y_t, Y_{t-1}) = \begin{cases} 2\sigma^2 & \text{si } h = 0 \\ -\sigma^2 & \text{si } h = -1 \text{ ou } h = 1 \\ 0 & \text{si } h \in \mathbb{Z} - \{-1, 1\} \end{cases}$$

Donc le processus Y_t est faiblement stationnaire.

b- Processus stationnaires

Il est fréquent d'assimiler la notion de processus stationnaire à celle de processus maîtrisable, régulier. Il n'en est rien et les exemples suivants sont à retenir.

Exemple 1

Soit le processus X_t à valeurs sur $\{-2, -1, 1, 2\}$ tel que

$$P(X_t = -2) = P(X_t = -1) = P(X_t = 1) = P(X_t = 2) = \frac{1}{4} \text{ pour tout } t$$

On a

$$\begin{aligned} E(X_t) &= -2P(X_t = -2) - 1P(X_t = -1) + 1P(X_t = 1) + 2P(X_t = 2) \\ &= -2 \times \frac{1}{4} - \frac{1}{4} + \frac{1}{4} + 2 \times \frac{1}{4} \\ &= 0 \end{aligned}$$

$$\begin{aligned} Var(X_t) &= (-2)^2P(X_t = -2) + (-1)^2P(X_t = -1) + (1)^2P(X_t = 1) + (2)^2P(X_t = 2) \\ &= 1 + \frac{1}{4} + \frac{1}{4} + 1 \\ &= \frac{10}{4} \end{aligned}$$

Le processus est faiblement stationnaire au deuxième degré mais imprévisible car la meilleure prévision de X_t par une constante est $E(X_t) = 0$ qui n'appartient pas au domaine de X_t .

Exemple 2

On considère le processus X_t qui à la date t prend les valeurs :

$$X_t = \begin{cases} 0 & \text{avec la probabilité } 1 - \frac{1}{t} \\ \sqrt{t} & \text{avec la probabilité } \frac{1}{2t} \\ -\sqrt{t} & \text{avec la probabilité } \frac{1}{2t} \end{cases}$$

$$\begin{aligned} E(X_t) &= 0\left(1 - \frac{1}{t}\right) + \sqrt{t}\frac{1}{2t} - \sqrt{t}\frac{1}{2t} \\ &= 0 \end{aligned}$$

$$\begin{aligned} \text{Var}(X_t) &= E(X_t^2) - (E(X_t))^2 \\ &= E(X_t^2) \\ &= t\frac{1}{2t} + t\frac{1}{2t} \\ &= \frac{2t}{2t} \\ &= 1 \end{aligned}$$

$$\text{Cov}(X_t, X_{t'}) = E(X_t, X_{t'}).$$

$$X_t X_{t'} = \begin{cases} 0 & \text{avec la probabilité } 1 - \frac{1}{tt'} \\ -\sqrt{t}\sqrt{t'} & \text{avec la probabilité } \frac{1}{2tt'} \\ \sqrt{t}\sqrt{t'} & \text{avec la probabilité } \frac{1}{2tt'} \end{cases}$$

d'où $E(X_t X_{t'}) = 0$

X_t est donc stationnaire au sens faible.

2.4 Les opérateurs B et F

Soient $(X_t, t \in \mathbb{Z})$ un processus.

On appelle opérateur retard ou backward, l'opérateur B qui à X_t . associe $BX_t = X_{t-1}$.

On appelle opérateur avance ou forward, l'opérateur F qui à X_t associe $FX_t = X_{t+1}$.

Ces opérateurs peuvent opérer sur autre chose que les séries temporelles : ainsi, si est un polynôme de degré n en la variable réelle X

$$BP_n(x) = P_n(x - 1) \text{ et } FP_n(x) = P_n(x + 1)$$

On peut définir également, à partir des opérateurs de décalage B et F, les opérateurs différence première :

$$\nabla = I - B \text{ et } \Delta = F - I$$

I représente l'opérateur neutre (identité)

$$\begin{aligned}\nabla X_t &= (I - B)X_t = X_t - BX_t = X_t - X_{t-1} \\ \Delta X_t &= (F - I)X_t = FX_t - X_t = X_{t+1} - X_t\end{aligned}$$

Ces opérateurs permettent de générer tous les opérateurs ∇^d , Δ^d , $d \in \mathbb{N}$ dont l'expression s'obtient par la formule du Binôme de Newton.

$$\nabla^d = (I - B)^d = I - C_1^d B + \dots + (-1)^d B^d$$

La puissance d'ordre k de B ou de F n'est autre que l'opérateur de décalage opérant k fois

$$\begin{aligned}B^k X_t &= X_{t-k} \\ F^k X_t &= X_{t+k}\end{aligned}$$

Il est évident que :

$$BF = FB = I \text{ et } B^{-1} = F, F^{-1} = B$$

Enfin, il est possible de donner un sens au développement en séries de la quantité

$$\begin{aligned}\frac{1}{1 - \lambda B} \text{ ou } \frac{1}{1 - \lambda F} \text{ si } |\lambda| < 1 \\ \frac{1}{1 - \lambda B} &= 1 + \lambda B + \lambda^2 B^2 + \dots \\ \frac{1}{1 - \lambda F} &= 1 + \lambda F + \lambda^2 F^2 + \dots\end{aligned}$$

Chapitre 3

Exemples de processus

Il existe dans la théorie probabiliste un grand nombre de processus, certains d'entre eux étant adaptés à une situation bien définie, comme par exemple les processus étudiant l'évolution d'une population.

Parmi les processus importants dans l'analyse statistique des séries temporelles on peut citer : le processus de poisson, les processus de vie et de mort, le mouvement Brownien, les processus autorégressifs et les processus moyenne mobile.

Dans cette partie, on s'intéresse en particulier aux deux derniers types de processus.

3.1 Bruit blanc

Définition 3.1.1. *Un processus stochastique $(\epsilon_t, t \in \mathbb{Z})$ constitue un bruit blanc si :*

$$E(\epsilon_t) = 0 \quad \text{pour chaque } t \text{ et}$$
$$\text{Cov}(\epsilon_t, \epsilon_{t+h}) = E(\epsilon_t \epsilon_{t+h}) = \begin{cases} \sigma^2 & \text{si } h = 0 \\ 0 & \text{si } h \neq 0. \end{cases}$$

Autrement dit, le processus stochastique $(\epsilon_t, t \in \mathbb{Z})$ est constitué d'une suite de variables aléatoires de moyennes nulles, non corrélées et de variances constantes $\sigma^2 < \infty$.

Ces variables, également appelées chocs représentent généralement des erreurs aléatoires.

3.2 Les processus Linéaires

Définition 3.2.1. *Un processus $(X_t, t \in \mathbb{Z})$ est dit linéaire s'il admet la représentation suivante :*

$$X_t = \mu + \sum_{j=0}^{\infty} \psi_j \epsilon_{t-j}.$$

Où $\psi_0 = 1$, μ , ψ_j , $j = 1, 2, \dots$ sont des paramètres réels et $(\epsilon_t, t \in \mathbb{Z})$ est un bruit blanc.

Le membre de droite de l'équation précédente converge en moyenne quadratique. Doob a démontré que cette convergence a lieu si et seulement si

$$\sum_{j=0}^{\infty} \psi_j^2 < \infty.$$

la condition ci-dessus assure l'existence des moments de second ordre ainsi que la stationnarité au sens large du processus linéaire d'où le théorème :

Théorème 3.2.1. *Soit $(X_t, t \in \mathbb{Z})$ un processus linéaire tel que :*

$$\sum_{j=0}^{\infty} \psi_j^2 < \infty, \quad \text{alors pour tout } t \in \mathbb{Z}$$

1. $E(X_t) = \mu$.
2. $\text{cov}(X_t, X_{t+h}) = \sigma^2 \sum_{j=0}^{\infty} \psi_j \psi_{j+h}$, pour tout $t \in \mathbb{Z}$ avec $\psi_0 = 1$.
3. $\text{var}(X_t) = \sigma^2 (1 + \sum_{j=0}^{\infty} \psi_j^2)$.

Démonstration

Puisque $(X_t, t \in \mathbb{Z})$ est un processus linéaire alors :

$$X_t = \mu + \sum_{j=0}^{\infty} \psi_j \epsilon_{t-j}.$$

1. On sait que $\sum_{j=0}^{\infty} \psi_j^2 < \infty$ donc on peut intervertir les symboles E et \sum on obtient :

$$\begin{aligned} E(X_t) &= \mu + E\left(\sum_{j=0}^{\infty} \psi_j \epsilon_{t-j}\right) \\ &= \mu + \sum_{j=0}^{\infty} E(\psi_j \epsilon_{t-j}) \\ &= \mu + \sum_{j=0}^{\infty} \psi_j E(\epsilon_{t-j}) \\ &= \mu. \end{aligned}$$

- 2.

$$\begin{aligned} \text{cov}(X_t, X_{t+h}) &= E\left((X_t - \mu)(X_{t+h} - \mu)\right) \\ &= E\left(\left(\sum_{j=0}^{\infty} \psi_j \epsilon_{t-j}\right)\left(\sum_{i=0}^{\infty} \psi_i \epsilon_{t+h-i}\right)\right) \\ &= \sum_{j=0}^{\infty} \sum_{i=0}^{\infty} \psi_j \psi_i E(\epsilon_{t-j} \epsilon_{t+h-i}) \\ &= \sigma^2 \left(\sum_{j=0}^{\infty} \psi_j \psi_{j+h}\right). \end{aligned}$$

Le résultat vient du fait que :

$$E(\epsilon_{t-j} \epsilon_{t+h-i}) = \begin{cases} \sigma^2 & \text{si } i = j + h \\ 0 & \text{si } i \neq j + h. \end{cases}$$

3. On déduit que :

$$\text{var}(X_t) = \sigma^2 \left(\sum_{j=0}^{\infty} \psi_j^2 \right).$$

3.3 Les processus AR

Dans cette partie, on ne considérera que des processus réels discrets du second ordre.

Définition 3.3.1. Soit un processus $(X_t, t \in \mathbb{Z})$, X_t est dit processus auto-régressif d'ordre p noté $X_t \rightarrow AR(p)$ si :

$$X_t = \sum_{i=1}^p \varphi_i X_{t-i} + \varepsilon_t$$

ε_t est bruit blanc

$$E(\varepsilon_t) = 0, \quad \text{Var}(\varepsilon_t) = \sigma^2, \quad E(\varepsilon_t \varepsilon_{t'}) = 0 \text{ telle que } (t \neq t')$$

Théorème 3.3.1. Soit (X_t) un processus AR d'ordre p . Si le polynôme caractéristique $\phi(B)$ a toutes ses racines à l'extérieur du cercle unité, alors :

$$\gamma_k = \varphi_1 \gamma_{k-1} + \varphi_2 \gamma_{k-2} + \dots + \varphi_p \gamma_{k-p} \quad k > 0.$$

$$\rho_k = \varphi_1 \rho_{k-1} + \varphi_2 \rho_{k-2} + \dots + \varphi_p \rho_{k-p} \quad k > 0. \quad (3.1)$$

$$\gamma_0 = \text{Var}(X_t) = \frac{\sigma^2}{1 - \sum_{j=1}^p \varphi_j \rho_j}.$$

$$\mu = E(X_t) = \frac{c}{1 - \sum_{j=1}^p \varphi_j}$$

Démonstration

(X_t) est un processus autorégressif d'ordre p sa forme générale est :

$$X_t = c + \varphi_1 X_{t-1} + \varphi_2 X_{t-2} + \dots + \varphi_p X_{t-p} + \varepsilon_t$$

où c est une constante réelle, φ_j sont des paramètres réels et ε_t est un bruit blanc.

En posant $E(X_t) = \mu$ pour tout t et en prenant l'espérance des deux membres de l'équation ci-dessus on obtient :

$$\begin{aligned} E(X_t) &= c + \varphi_1 E(X_{t-1}) + \varphi_2 E(X_{t-2}) + \dots + \varphi_p E(X_{t-p}) + E(\varepsilon_t) \\ &= c + \mu(\varphi_1 + \varphi_2 + \dots + \varphi_p) \end{aligned}$$

D'où

$$\mu \left(1 - \sum_{j=1}^p \varphi_j \right) = c$$

On en déduit

$$\mu = \frac{c}{(1 - \sum_{j=1}^p \varphi_j)}.$$

Posons $\tilde{X}_t = X_t - \mu$. En multipliant des deux côtés dans la forme générale de l'AR(p) par \tilde{X}_t et en prenant l'espérance on obtient :

$$\begin{aligned} E(\tilde{X}_t^2) &= \text{Var}(X_t) \\ &= \varphi_1 E(\tilde{X}_{t-1}\tilde{X}_t) + \varphi_2 E(\tilde{X}_{t-2}\tilde{X}_t) + \dots + \varphi_p E(\tilde{X}_{t-p}\tilde{X}_t) + E(\varepsilon_t\tilde{X}_t) \\ &= \varphi_1\gamma_1 + \varphi_2\gamma_2 + \dots + \varphi_p\gamma_p + \sigma^2 \end{aligned}$$

En divisant les deux membres de l'équation précédente par γ_0 on obtient :

$$\gamma_0 = \frac{\sigma^2}{1 - \sum_{j=1}^p \varphi_j \rho_j}.$$

En multipliant dans l'équation de l'autorégressif avec les données centrées les deux membres par \tilde{X}_{t-k} et en prenant l'espérance on a :

$$\begin{aligned} E(\tilde{X}_t\tilde{X}_{t-k}) &= \gamma_k \\ &= \varphi_1 E(\tilde{X}_{t-1}\tilde{X}_{t-k}) + \dots + \varphi_p E(\tilde{X}_{t-p}\tilde{X}_{t-k}) + E(\varepsilon_t\tilde{X}_{t-k}) \\ &= \varphi_1\gamma_{k-1} + \varphi_2\gamma_{k-2} + \dots + \varphi_p\gamma_{k-p}. \end{aligned}$$

Notons que $E(\varepsilon_t\tilde{X}_{t-k}) = 0$ car \tilde{X}_{t-k} dépend uniquement des quantités $\varepsilon_{t-k}, \varepsilon_{t-k-1}, \dots$ et celles-ci n'ont aucune corrélation avec ε_t .

En divisant les deux membres de la dernière égalité par γ_0 on obtient :

$$\rho_k = \varphi_1\rho_{k-1} + \varphi_2\rho_{k-2} + \dots + \varphi_p\rho_{k-p}.$$

3.3.1 Le système de Yulle-Walker

La deuxième équation du théorème 3.3.1 sur l'AR(p) permet de calculer les coefficients ϕ_i en fonction des corrélations ρ_h .

En effet en prenant successivement $h=1, h=2, \dots, h=p$, dans cette équation, on obtient le système d'équation suivant, ou équation de Yulle-Walker :

$$\begin{pmatrix} 1 & \rho_1 & \cdot & \cdot & \cdot & \rho_{h-1} \\ \rho_1 & 1 & \cdot & \cdot & \cdot & \rho_{h-2} \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \rho_{h-1} & \rho_{h-2} & \cdot & \cdot & \cdot & 1 \end{pmatrix} \begin{pmatrix} \phi_1 \\ \phi_2 \\ \cdot \\ \cdot \\ \cdot \\ \phi_p \end{pmatrix} = \begin{pmatrix} \rho_1 \\ \rho_2 \\ \cdot \\ \cdot \\ \cdot \\ \rho_p \end{pmatrix}$$

Corollaire 3.3.1. Soit $(X_t, t \in \mathbb{Z})$ un processus autorégressif d'ordre p stationnaire.

On a :

1. La fonction d'autocorrélation partielle est telle que $\tau_k \neq 0$ pour $k \leq p$, $\tau_p = \varphi_p$ et $\tau_k = 0$ pour $k > p$.
2. Si $\phi(B)$ possède p racines distinctes $\eta_1, \eta_2, \dots, \eta_p$, la fonction d'autocorrélation est alors de la forme générale :

$$\rho_k = d_1 \eta_1^{-k} + \dots + d_p \eta_p^{-k}, \quad k > 0.$$

où d_1, d_2, \dots, d_p sont des constantes déterminées par les valeurs initiales ρ_0, ρ_1, \dots

Remarques

1. En faisant intervenir l'opérateur B on peut écrire :

$$(I - \varphi_1 B - \dots - \varphi_p B^p) X_t = \varepsilon_t$$

soit $\phi(B) X_t = \varepsilon_t$ avec le polynôme : $\phi(B) = I - \varphi_1 B - \dots - \varphi_p B^p$

2. Sous la forme la plus générale, un processus AR(p) peut contenir un terme constant θ_0 :

$$X_t = \theta_0 + \sum_{i=1}^p \varphi_i X_{t-i} + \varepsilon_t$$

3. Supposons $E(X_t) = m$ constante pour tout t , alors m vérifie :

$$m(1 - \sum_{i=1}^p \varphi_i) = 0$$

3.3.2 Etude du processus autorégressif (AR)

a- Inversibilité du processus AR(p)

Soit $\phi(z) = I - \varphi_1 z - \dots - \varphi_p z^p$ le polynôme caractéristique du processus AR(p).

Soient z_1, z_2, \dots, z_p les racines de $\phi(z)$ dans \mathbb{C} et notons z_1, z_2, \dots, z_r les r -racines de module supérieur à 1, $z_{r+1}, z_{r+2}, \dots, z_p$ les $p-r$ racines de modules inférieur à 1 (notons que : $z_1 z_2 \dots z_p = 1$). On peut écrire :

$$\phi(z) = \prod_{i=1}^r (1 - \frac{z}{z_i}) \prod_{j=r+1}^p (1 - \frac{z}{z_j})$$

Par analogie on a :

$$\phi(B) = \prod_{i=1}^r (1 - \frac{B}{z_i}) \prod_{j=r+1}^p (1 - \frac{B}{z_j})$$

Exemple 1 :

Soit X_t le processus de Markov AR(p) défini par :

$$X_t = 0.8X_{t-1} + \varepsilon_t$$

Le polynôme caractéristique de ce processus est :

$$\phi(z) = 1 - 0.8z$$

La solution du polynôme est : $z_1 = 1.25$ elle est de module supérieur à 1.

On peut écrire : $(1 - 0.8B)X_t = \varepsilon_t$ d'où

$$\begin{aligned} X_t &= \frac{1}{1 - 0.8B} \varepsilon_t \\ &= (1 + 0.8B + (0.8)^2 B^2 + \dots) \varepsilon_t \\ &= \varepsilon_t + 0.8\varepsilon_{t-1} + (0.8)^2 \varepsilon_{t-2} + \dots (0.8)^k \varepsilon_{t-k} + \dots \end{aligned}$$

Le processus AR(1) est inversé.

Calculons $E(\varepsilon_t X_{t-k})$ pour $k > 1$

On a :

$$X_t = \varepsilon_{t-k} + 0.8\varepsilon_{t-k-1} + (0.8)^2 \varepsilon_{t-k-2} + \dots + (0.8)^j \varepsilon_{t-k-j} + \dots$$

Donc

$$E(\varepsilon_t X_{t-k}) = E(\varepsilon_t \sum_{j=0}^{\infty} (0.8)^j \varepsilon_{t-k-j})$$

Puisque ε_t est un bruit blanc, donc $E(\varepsilon_t X_{t-k}) = 0$ par conséquent, le bruit est non Corrélé avec les observations passées du processus.

3.3.3 Simulation d'un échantillons de 1000 données du processus de Markov de l'exemple 1 :

```
t=1000
t0=500
phi=c(0.8,0)
p=length(phi)
init=rep(1,2)
innov=rnorm(t+t0)
n=length(innov)
x=init
for(i in (p+1):n)
{
aux=innov[i]+sum(phi*x[(i-1):(i-p)])
x=c(x,aux)
}
x=x[(t0+1):n]
x
plot.ts(x)
acf(x)
pacf(x)
```

Dans les graphiques suivants, nous représentons la série X_t , la fonction d'autocorrélation et d'autocorrélation partielle caractéristiques d'un processus de Markov

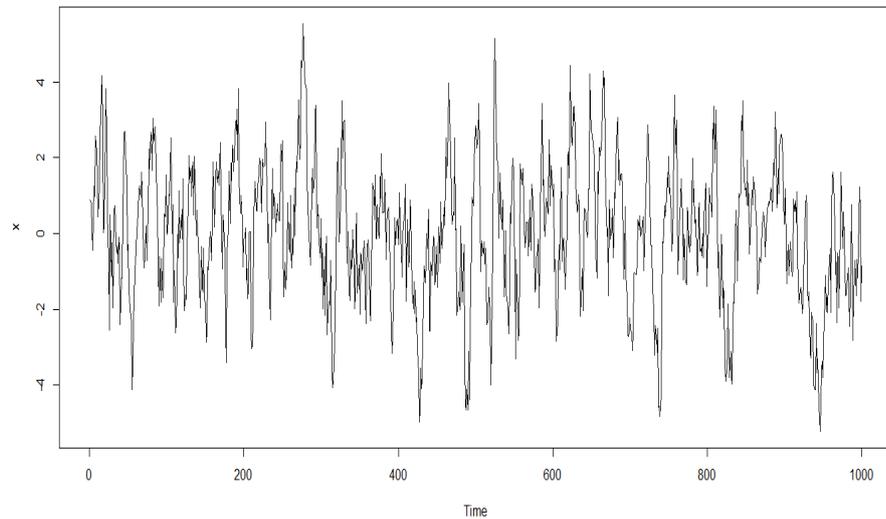


FIGURE 3.1 – Graphe la série X_t

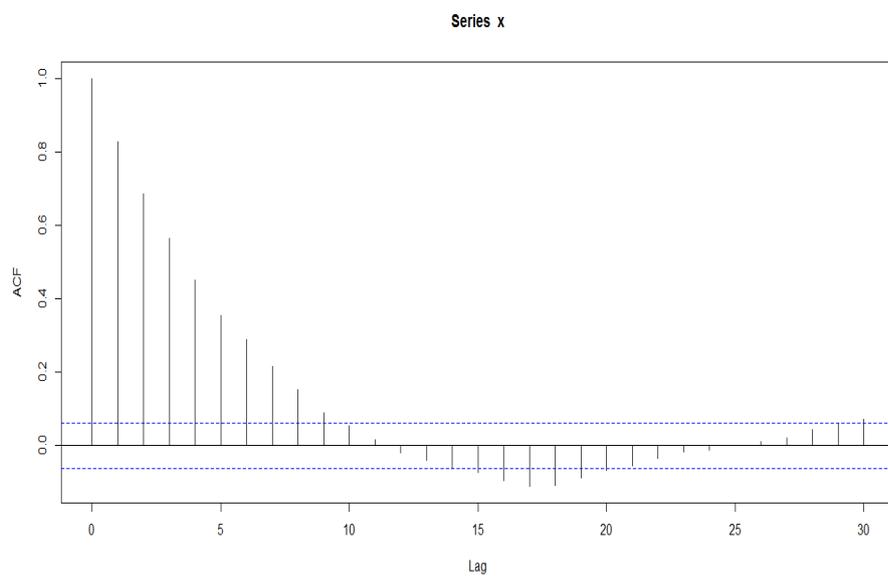


FIGURE 3.2 – Corrélogramme

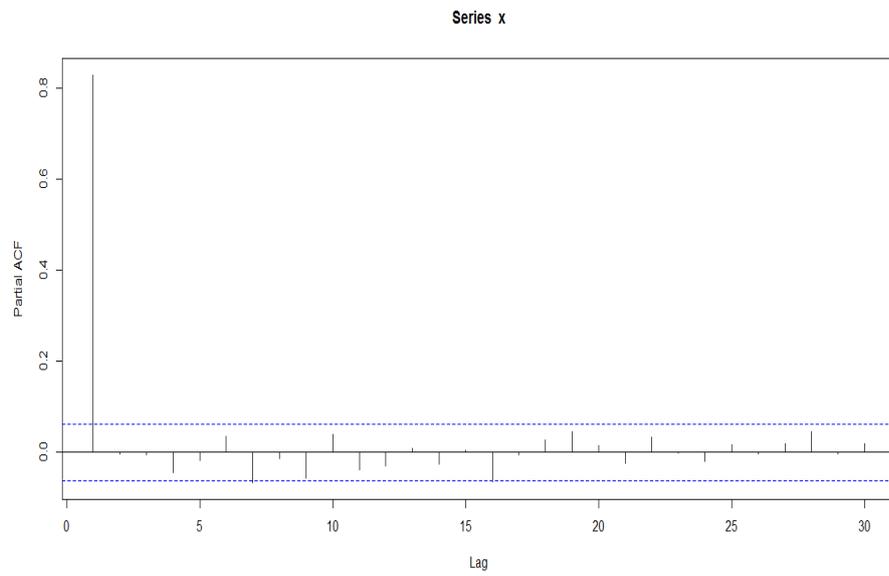


FIGURE 3.3 – Corrélogramme partiel

Exemple 2 :

Soit le processus de Yule AR(2) défini par :

$$X_t = 1.2X_{t-1} + 0.3X_{t-2} + \varepsilon_t$$

En utilisant l'opérateur B on obtient :

$$X_t = 1.2BX_t + 0.3B^2X_t + \varepsilon_t$$

C'est à dire :

$$X_t - 1.2BX_t - 0.3B^2X_t = \varepsilon_t$$

Ou encore :

$$(1 - 1.2B - 0.3B^2)X_t = \varepsilon_t$$

Le polynôme caractéristique du processus de Yule est

$$\phi(z) = 1 - 1.2z - 0.3z^2$$

Les racines de ce polynôme sont $z_1 = 0.7$ et $z_2 = -4.7$ d'où :

$$(1 + 0.7B)(1 - 4.7B)X_t = \varepsilon_t$$

D'autre part :

$$1 + 0.7B = BF + 0.7B = 0.7B(4.7F + 1)$$

D'où :

$$\begin{aligned} (1 - 4.7B)X_t &= \frac{4.7F}{1 + 4.7F}\varepsilon_t \\ &= 4.7F(1 - 4.7F + \dots + (-1)^k(4.7)^k F^k + \dots)\varepsilon_t \\ &= \sum_{k=1}^{\infty} \alpha_k \varepsilon_{t+k} \end{aligned}$$

Avec :

$$\alpha_k = (-1)^{k-1}(4.7)^k$$

Par conséquent :

$$\begin{aligned} X_t &= \frac{1}{1 - 4.7B} \sum_{k=1}^{\infty} \alpha_k \varepsilon_{t+k} \\ &= \sum_{j=0}^{\infty} (4.7)^j B^j \sum_{k=1}^{\infty} \alpha_k \varepsilon_{t+k} \\ &= \sum_{k=1}^{\infty} \alpha_k \left(\sum_{j=0}^{\infty} (4.7)^j B^{t+k-j} \right) \end{aligned}$$

Puisque les racines de polynôme caractéristique pour tous supposons en module dans le processus n'est pas stationnarité dans ce cas on ne peut faire de la prévision.

Exemple 3 :

Soit le processus AR(2) définie par :

$$X_t - 0.6X_{t-1} + 0.2X_{t-2} = \varepsilon_t$$

$$X_t = 0.6X_{t-1} - 0.2X_{t-2} + \varepsilon_t$$

ε_t Sont suivent la loi $N(0,1)$

pour expliquer les différents graphiques, on calcule le polynôme caractéristique en introduisant l'opérateur retard B .

$$X_t - 0.6BX_t + 0.2BX_t = \varepsilon_t$$

$$(1 - 0.6B + 0.2B^2)X_t = \varepsilon_t$$

$$\phi(B) = 1 - 0.6B + 0.2B^2$$

$$\Delta = b^2 - 4ac$$

$$= (-0.6)^2 - 4(1)(0.2)$$

$$= 0.36 - 0.8 = -0.44 = 0.44i^2$$

On peut le simuler en utilisant la commande suivante :

```
x=arima.sim(model = monmodele, n = 100, rand.gen = rnorm)
plot(x)
acf(x)
pacf(x)
```

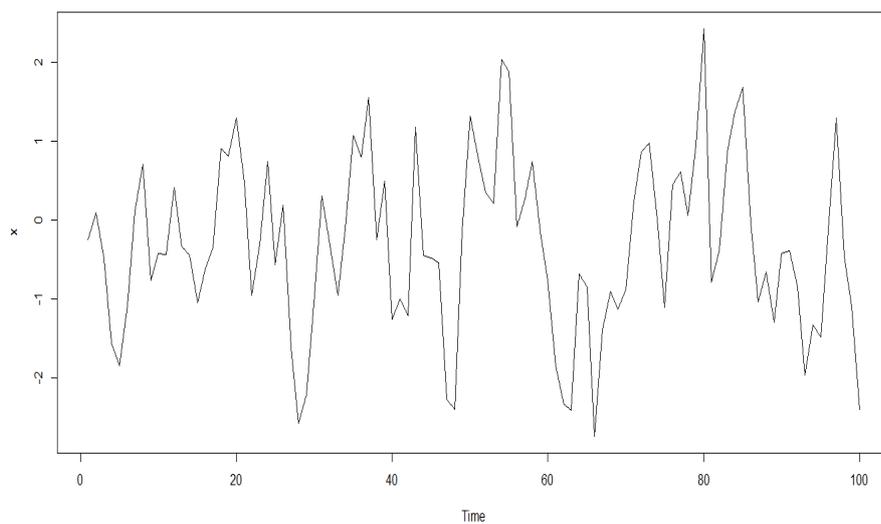


FIGURE 3.4 – Graphe la série X_t

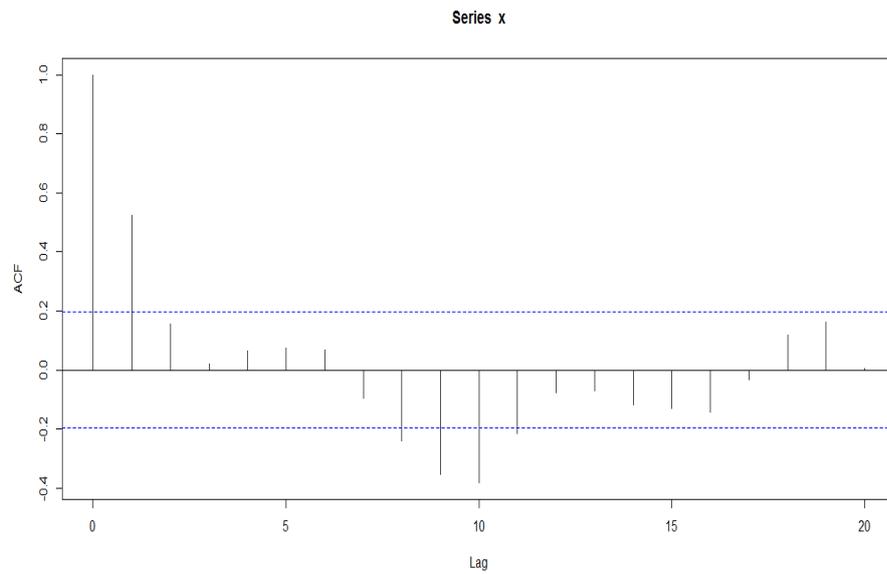


FIGURE 3.5 – Corrélogramme

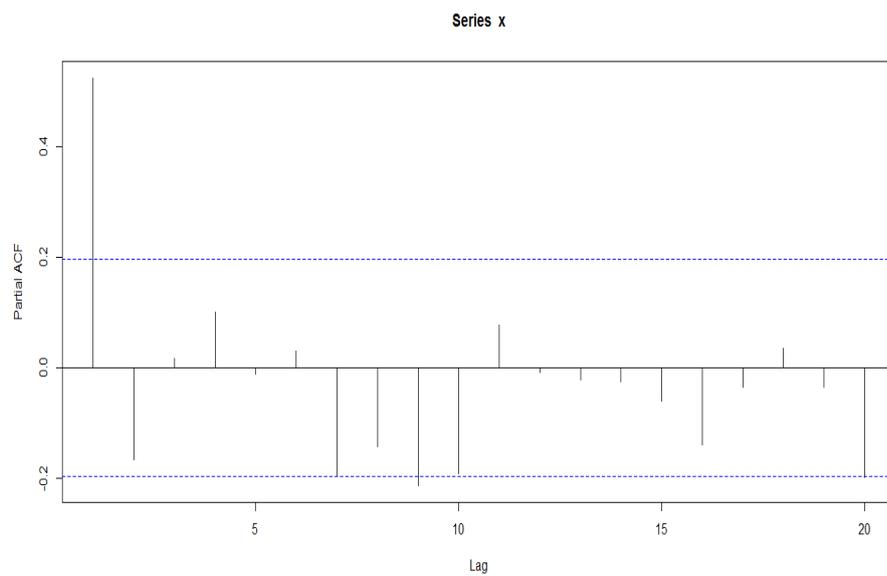


FIGURE 3.6 – Corrélogramme partiel

```
> monmodele = list(ar=c(0.6,-0.2))
> X=arima.sim(model=monmodele,n=100,rand.gen=rnorm)
> X.ord=c(2,0,0)
> X.arima=arima(X,order=X.ord)
> tsdiag(X.arima)

> X.pred=predict(X.arima,10)
```

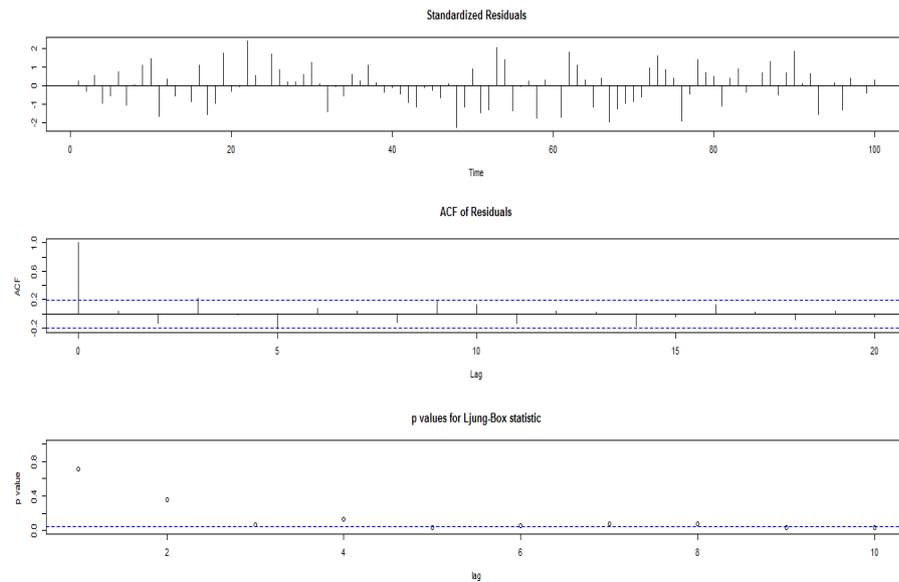


FIGURE 3.7 –

```
> X.pred
```

```
$pred
```

```
Time Series:
```

```
Start = 101
```

```
End = 110
```

```
Frequency = 1
```

```
[1] -0.1403998 -0.1863111 -0.2317215 -0.2557670 -0.2641830 -0.2654610
```

```
[7] -0.2646804 -0.2638441 -0.2633880 -0.2632232
```

```
$se
```

```
Time Series:
```

```
Start = 101
```

```
End = 110
```

```
Frequency = 1
```

```
[1] 1.007969 1.251754 1.297193 1.300947 1.300951 1.301074 1.301148 1.301164
```

```
[9] 1.301165 1.301165
```

Stationnarité

Compte tenu de ce que nous venons de voir quant à la valeur de $E(X_{t-k}\varepsilon_t)$ la stationnarité d'un AR(p) n'est pas toujours établie.

On a :

$$\begin{aligned}
 \text{Var}(X_t) &= E(X_t^2) - (E(X_t))^2 \\
 &= E(X_t^2) \\
 &= E\left(X_t \left(\sum_{\varphi_i}^p X_{t-i} + \varepsilon_t\right)\right) \\
 &= \sum_{i=1}^p E(X_t X_{t-i}) + E(\varepsilon_t X_t) \\
 &= \sum_{i=1}^p \varphi_i \nu(i) + E(\varepsilon_t X_t)
 \end{aligned}$$

Or :

$$E(\varepsilon_t X_t) = \sum_{i=t}^p \varphi_i E(X_t X_{t-i}) + E(\varepsilon_t^2)$$

Premier cas : le processus est inversible, c'est dire toutes les racines du polynôme caractéristique $\phi(z)$ sont supérieures à 1 en module alors :

$$E(X_t X_{t-i}) = 0 \text{ et } E(\varepsilon_t X_t) = \sigma^2$$

D'où :

$$\nu(0) = \text{Var}(X_t) = \frac{\sigma^2}{1 - \sum_{i=1}^p \varphi_i \rho(i)}$$

$\text{Var}(X_t)$ ne dépend pas de t, le processus est faiblement stationnaire.

Deuxième cas : les racine du polynôme caractéristique $\phi(z)$ ne sont pas toutes supérieures à 1 en module

Nous nous contenterons d'énoncer le résultat : on peut toujours supposer que les racines du polynôme $\phi(z)$ sont de module supérieur à 1, en remplaçant le processus dans $\phi(B)x_t = \varepsilon_t$ par le processus $\phi^*(B)X_t = \alpha_t$ où $\phi^*(B)$ s'obtient en remplaçant, dans $\phi(B)$, les racines inférieures à 1 en module par leur inverse, α_t étant un nouveau bruit blanc.

3.4 Les processus MA

On considère un processus $(X_t, t \in \mathbb{Z})$ réel du second ordre.

Définition 3.4.1. On dit que le processus $(X_t, t \in \mathbb{Z})$, est une moyenne mobile d'ordre q noté $X_t \rightarrow MA(q)$ si :

$$\begin{aligned} X_t &= \varepsilon_t - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2} - \dots - \theta_q \varepsilon_{t-q} \\ X_t &= \varepsilon_t - \sum_{i=1}^q \theta_i \varepsilon_{t-i} \quad \forall t \in \mathbb{Z} \end{aligned}$$

$$E(\varepsilon_t) = 0, \text{Var}(\varepsilon_t) = \sigma^2, E(\varepsilon_t \varepsilon_{t'}) = 0 \quad t \neq t'.$$

Remarque :

$$X_t = (I - \theta_1 B - \dots - \theta_q B^q) \varepsilon_t$$

En notant $\Theta(B)$ le polynôme de degré q , qui égale à $(I - \theta_1 B - \dots - \theta_q B^q)$ On a :

$$X_t = \Theta(B) \varepsilon_t$$

3.4.1 Propriétés d'un MA(q)

a)- Espérance

$$\begin{aligned} E(X_t) &= E(\varepsilon_t - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2} - \dots - \theta_q \varepsilon_{t-q}) \\ &= E(\varepsilon_t) - \theta_1 E(\varepsilon_{t-1}) - \dots - \theta_q E(\varepsilon_{t-q}) \\ &= 0 \end{aligned}$$

b)- Variance

$$\text{Var}(X_t) = \sigma^2 (1 + \sum_{i=1}^q \theta_i^2)$$

c)- Fonction d'auto covariance :

Soit à calculer : $E(X_t X_{t-h})$,

Si

$$h > q \quad E(X_t X_{t-h}) = 0$$

Si

$$h < q \quad E(X_t X_{t-h}) = (-\theta_h + \theta_1 \theta_{h+1} + \dots + \theta_q \theta_{q-h}) \sigma^2.$$

Le processus est donc stationnaire (au sens faible) et on en déduit la fonction d'auto corrélation :

$$\begin{aligned}\rho(h) &= 0 && \forall h > q. \\ \rho(h) &= \frac{(-\theta_h + \theta_1\theta_{h-1} + \dots + \theta_q\theta_{q-h})}{1 + \theta_1^2 + \dots + \theta_q^2} && \forall h < q.\end{aligned}$$

Le calcul de $r(h)$ ne présente pas de particularité, et est exécuté selon le système décrit auparavant. Il résulte des calculs précédents qu'un processus moyenne mobile est toujours stationnaire (au sens faible) ce qui n'était pas pour un processus autorégressif quelconque, d'où le théorème suivant :

Théorème 3.4.1. Soit $(X_t, t \in \mathbb{Z})$ un processus moyenne mobile d'ordre q , on a :

$$\begin{aligned}E(X_t) &= 0, \text{ pour tout } t, \\ \gamma_h &= \begin{cases} \sigma^2(1 + \theta_1^2 + \theta_2^2 + \dots + \theta_q^2) & \text{si } h = 0 \\ \sigma^2(-\theta_h + \theta_1\theta_{h+1} + \dots + \theta_q\theta_{q-h}) & \text{si } 1 \leq h \leq q \\ 0 & \text{si } h > q \end{cases} \\ \rho_h &= \begin{cases} 1 & \text{si } h = 0 \\ \frac{(-\theta_h + \theta_1\theta_{h+1} + \dots + \theta_q\theta_{q-h})}{(1 + \theta_1^2 + \theta_2^2 + \dots + \theta_q^2)} & \text{si } 1 \leq h \leq q \\ 0 & \text{si } h > q \end{cases}\end{aligned}$$

Ce théorème est un cas particulier du théorème 3.2.1 énoncé pour les processus linéaires.

Conclusion

Le problème de l'identification d'un processus MA(q) c'est à dire la détermination du paramètre q se fera sur l'examen de la nullité ou non des $\rho(h)$

Exemple 3.4.1. Soit le processus défini par :

$$X_t = \varepsilon_t - \frac{1}{2}\varepsilon_{t-1}$$

C'est un processus moyenne mobile d'ordre $q = 1$

Supposons que l'on veuille générer $n = 100$ réalisation du processus MA(1)

On pourra écrire les lignes de commandes correspondant à l'algorithme on utilise la commande **arim.sim** c'est à dire :

```
X=arima.sim(100,model=list(ma=-1/2))
plot.ts(x)
acf(x)
pacf(x)
```

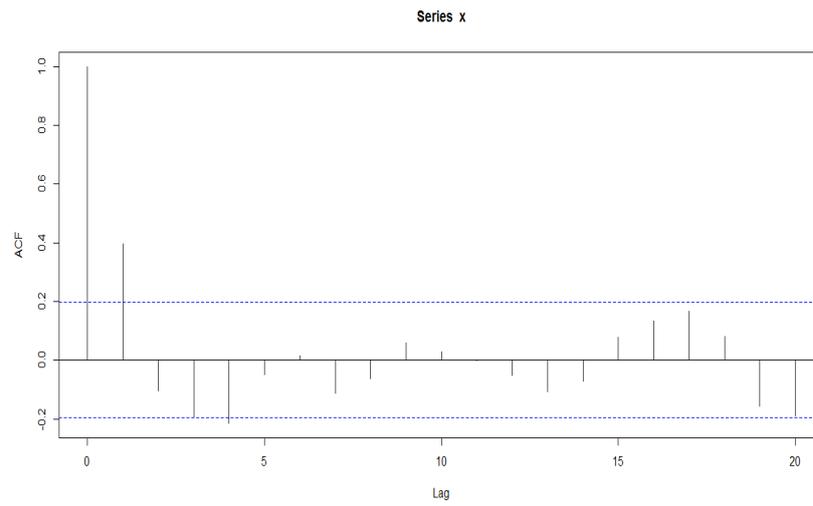
FIGURE 3.8 – Graphe la série X_t 

FIGURE 3.9 – Corrélogramme

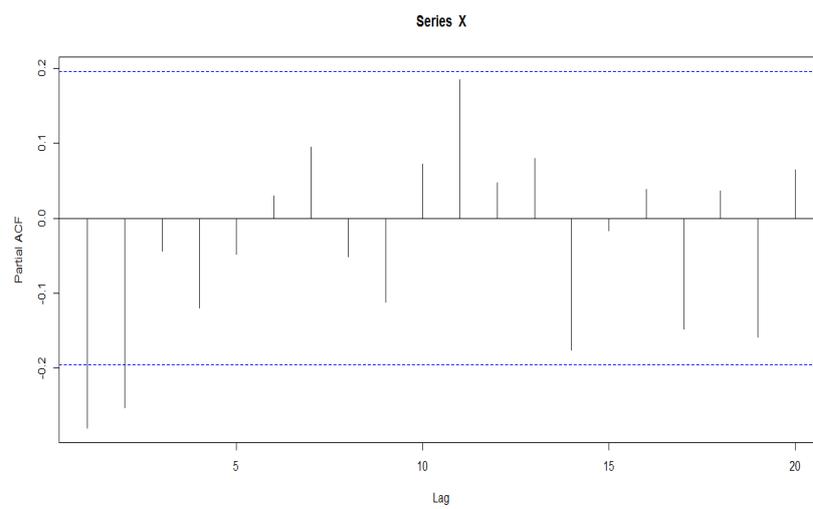


FIGURE 3.10 – Corrélogramme partiel

Au lieu de simuler 1000 observations, on se contente d'un échantillon de taille $n=100$ pour avoir un graphe plus claire.

3.5 Remarque sur le lien entre AR et MA

La démonstration de l'inversibilité d'un $AR(p)$ sous certaines conditions laissent apparaître une équivalence entre $AR(p)$ et $MA(\infty)$.

Considérons, par exemple, un processus $AR(1)$ défini par :

$$X_t = \rho X_{t-1} + \varepsilon_t$$

Par substitution successives, il est clair que :

$$X_t = \varepsilon_t + \rho\varepsilon_{t-1} + \dots + \rho^q\varepsilon_{t-q} + \rho^{q+1}X_{t-q-1}$$

D'où

$$(X_t - \sum_{j=0}^q \rho^j \varepsilon_{t-j})^2 = \rho^{2q+2} X_{t-q-1}^2$$

Donc :

$$E((X_t - \sum_{j=0}^q \rho^j \varepsilon_{t-j})^2) = \rho^{2q+2} E(X_{t-q-1}^2)$$

Supposons que le processus X_t soit faiblement stationnaire, centré de variance σ^2 , alors :

$$E((X_t - \sum_{j=0}^q \rho^j \varepsilon_{t-j})^2) = \rho^{2q+2} \sigma^2$$

Notons $Y_{t,q} = \sum_{j=0}^q \rho^j \varepsilon_{t-j}$

Il vient :

$$E[(X_t - Y_{t,q})^2] \longrightarrow 0 \text{ quand } q \longrightarrow \infty$$

La suite $Y_{t,q}$ converge en moyenne quadratique vers X_t

3.6 Les processus mixtes ARMA et ARIMA

Les processus examinés jusqu'ici étaient soit purement de type autorégressif, soit purement de type moyenne mobile. Certains processus peuvent être composés des deux types. Il s'agit alors de processus mixtes appelés les processus **ARMA**.

Définition 3.6.1. *Un processus $(X_t, t \in \mathbb{Z})$ est dit de type mixte d'ordres p et q , noté **ARMA**(p, q) (où p et q sont deux entiers positifs), s'il satisfait l'équation*

$$X_t = c + \phi_1 X_{t-1} + \dots + \phi_p X_{t-p} - \theta_1 \varepsilon_{t-1} - \dots - \theta_q \varepsilon_{t-q} + \varepsilon_t \quad (3.2)$$

où c est une constante, ϕ_j ,

$j = 1, 2, \dots, p$, sont les paramètres autorégressifs **AR**, θ_j , $j = 1, 2, \dots, q$, sont les paramètres

moyenne mobile **MA**, et $(\epsilon_t, t \in \mathbb{Z})$ est un bruit blanc.

En introduisant l'opérateur retard B , le processus **ARMA** (p, q) s'écrit simplement :

$$\phi(B)X_t = \theta(B)\epsilon_t. \quad (3.3)$$

Dans le cas particulier où $p = q = 1$, le modèle **ARMA** $(1, 1)$ s'écrit :

$$X_t = c + \phi_1 X_{t-1} - \theta_1 \epsilon_{t-1} + \epsilon_t.$$

Les processus étudiés jusqu'à présent sont stationnaires cette condition n'est pas toujours réalisée en pratique. Dans certain cas toutefois, on peut rendre la série stationnaire en introduisant l'opérateur différence. Les processus **ARMA** (p, q) pour lesquels des différences ont été effectuées sont appelés processus **ARIMA** (p, d, q) .

- un processus **ARIMA** $(p, 0, q)$ est un processus **ARMA** (p, q) .
- un processus **ARIMA** $(p, 0, 0)$ est de type autorégressif **AR** d'ordre p .
- un processus **ARIMA** $(0, 0, q)$ est de type moyenne mobile **MA** d'ordre q .

Chapitre 4

Modèles de Box et Jenkins

Durant l'année 1976 George. E.P. Box et Gwilym. M. Jenkins ont proposé une méthode qui a révolutionné la statistique et actuellement très utilisée par divers praticiens dans des domaines comme la finance, la météorologie ...etc.

les méthodes de Box et Jenkins sont des méthodes itératives pour l'analyse des séries chronologiques. Elles utilisent principalement les propriétés des fonctions d'autocorrélation et d'autocorrélation partielle dans les cas suivants :

1. Les données sont recueillies à intervalles de temps réguliers.
2. La série d'observations est assez longue (au moins 50 observations).
3. La série d'observations est stationnaire dans le cas contraire, on la rend stationnaire (calcul des différences, extraction de la tendance,...) avant d'appliquer la méthode.

4.1 Estimation des autocorrélations et identification

Les méthodes de Box et Jenkins supposent que la série chronologique observée est générée par un processus stochastique. L'identification dudit processus est basée sur l'examen de sa fonction d'autocorrélation et de sa fonction d'autocorrélation partielle. L'estimation de ces fonction à partir des données de la série constitue une phase primordiale à l'étape d'identification.

Pour une série chronologique X_1, \dots, X_n donnée, on définit les autocovariances empirique, de pas (ou délai) h , par

$$C_h = \frac{\sum_{t=1}^{n-h} (X_{t+h} - \bar{X})(X_t - \bar{X})}{n}, \quad h = 0, 1, 2, \dots$$

Où \bar{X} désigne la moyenne des observation de la série, $\bar{X} = \frac{\sum_{t=1}^n X_t}{n}$.

On défini les coefficients d'autocorrélation de pas (ou délai) h , notés r_h , par les équations (1.1).

La suite $\rho(h)$ en fonction de h constitue la fonction d'autocorrélation empirique de la série chronologique.

On définit les coefficients d'autocorrélation partielle de pas(ou délai) h , notés φ_{hh} , par les équation(1.3).

La suite φ_{hh} en fonction de h constitue la fonction d'autocorrélation partielle empirique de la série chronologique.

Remarque

les équation (1.3) constituent une approche de résolution par récurrence des équation(1.2), dans laquelle on utilise les r_j pour estimer les ρ_j . Puisque les équations(1.2) découlent de la relation(3.1), les équation(1.3) découlent de l'analogie empirique

$$\rho_j = \varphi_{h,1}\rho_{j-1} + \varphi_{h,2}\rho_{j-2} + \dots + \varphi_{h,h}\rho_{j-h}, \quad j = 1, 2, \dots, h.$$

Dans laquelle seule $\varphi_{h,h}$ est le coefficient d'intérêt. Il existe d'autres méthodes, notamment celle qui consiste à considérer l'ajustement de modèles autorégressif d'ordres successifs $h = 2$, $h = 3$, etc, et à conserver chaque fois le dernier coefficient φ_h comme $\varphi_{h,h}$.

Une analyse des fonctions d'autocorrélation et d'autocorrélation partielle empiriques permet d'identifier des modèles potentiels pour la série étudiée.

L'analyse consiste à essayer de déterminer si l'une des fonctions d'autocorrélation ρ_k et φ_{hh} est nulle à partir d'une valeur h , ou si elle s'amortit lentement lorsque h augmente. Pour ce faire, on considéré généralement les lois asymptotiques suivant lorsque T est grand.

1. Si un processus est tel que $\rho_s = 0$ dès que $s > q$, alors r_h est approximativement de loi normale de moyenne nulle et de variance

$$\text{var}(r_h) \simeq \frac{1}{n} \left(1 + 2 \sum_{j=1}^q \rho_j^2 \right) \quad \text{pour } h > q$$

Remarque : La variance et l'erreur type (s.e) de r_h sont généralement estimées en supposant que le processus est MA(q) d'ordre $q = h - 1$ pour chaque h . Précisément, pour $h = 1, 2, \dots$

$$\widehat{\text{var}}(r_h) = \frac{1}{n} \left(1 + 2 \sum_{j=1}^{h-1} r_j^2 \right), \quad \text{et } s.e.(r_h) = \sqrt{\widehat{\text{var}}(r_h)}$$

Certaines logiciels donnent chaque r_h ainsi que son erreur type $s.e.(r_h)$. Des intervalles de confiance (de niveau $\simeq 95\%$) de la forme $0 \pm 1,96 \times s.e.(r_h)$ sont parfois obtenus.

2. Si un processus est tel que $\phi_{ss} = 0$ dès que $s > p$, alors $\varphi_{h,h}$ est approximativement de loi normale de moyenne nulle et de variance

$$\text{var}(\varphi_{h,h}) \simeq \frac{1}{n}, \quad \text{pour } h > p.$$

Les logiciels donnent chaque $\varphi_{h,h}$ ainsi que son erreur type $s.e.(\varphi_{h,h}) \simeq \frac{1}{\sqrt{n}}$.

Remarque : En pratique, pour $h = 1, 2, \dots$, on utilise les statistique

$$t_{r_h} = \frac{r_h}{s.e.(r_h)} \quad \text{et} \quad t_{\varphi_{h,h}} = \frac{\varphi_{h,h}}{s.e.(\varphi_{h,h})} = \frac{\varphi_{h,h}}{1/n}$$

Pour effectuer approximativement les tests des hypothèses $H_0 : \rho_h = 0$ d'une part et $H_0 : \phi_{hh} = 0$ d'autre part. Ainsi,

- Pour tester $H_0 : \rho_h = 0$ contre $H_1 : \rho_h \neq 0$, la règle est de rejeter H_0 si $|t_{rh}|$ est grand, c'est-à-dire si $|t_{rh}| > z_{\alpha/2}$, ou si p -value $\simeq 2(1 - \Phi(|t_{rh}|))$ est petite.
- Pour tester $H_0 : \phi_{hh} = 0$ contre $H_1 : \phi_{hh} \neq 0$, la règle est de rejeter H_0 si $|t_{\varphi_{h,h}}|$ est grand, c'est-à-dire si $|t_{\varphi_{h,h}}| > z_{\alpha/2}$, ou si p -value $\simeq 2(1 - \Phi(|t_{\varphi_{h,h}}|))$ est petite.

Exemple 4.1.1. Les données du tableau représentent la production d'une usine (en tonnes), de 1975 à 1994.

	<i>Année</i>	<i>Production</i>
1	1975	19.8
2	1976	21.2
3	1977	17.4
4	1978	18.9
5	1979	25.2
6	1980	27.2
7	1981	27.4
8	1982	21.1
9	1983	23.9
10	1984	30.2
11	1985	26.9
12	1986	20.7
13	1987	24.2
14	1988	31.8
15	1989	30.1
16	1990	29.5
17	1991	28.3
18	1992	22.3
19	1993	25.7
20	1994	28.6

Tableau 3.1- Production annuelle d'une usine

Ensuite, on tape les lignes de commande suivantes sous logiciel R :

```
a<-c(19.8,21.2,17.4,18.9,25.2,27.2,27.4,21.1,23.9,30.2,26.9,20.7,24.2,31.2,30.1,
29.5,28.3,22.3,25.7,28.6)
X=ts(c(a), frequency = 1, start = c(1975, 1),end=1994)
X
plot.ts(X)
```

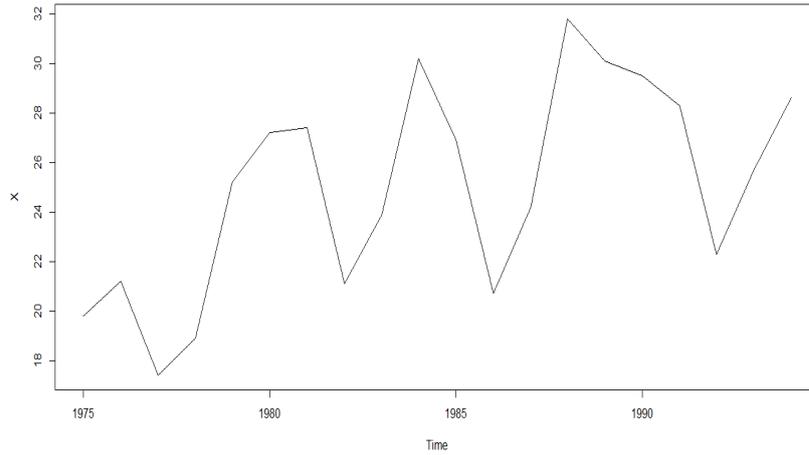


FIGURE 4.1 – Évolution de la production

Exemple 4.1.2. *Considérons les observations de l'exemple 4.1.1. Un examen du graphique du tableau 4.1 montre que la moyenne pour les mois autour de l'année 1977 est significativement différente de celle des mois autour de l'année 1989. La série n'est donc pas stationnaire. On peut effectuer des différences d'ordre 1 et obtenir une nouvelle série. On peut noter que la série des différences d'ordre 1 paraît stationnaire et de moyenne presque nulle.*

Exemple 4.1.3. *(Suite de l'exemple 4.1.1)*

Supposons qu'une tendance linéaire soit plausible pour X_t , la production annuelle (en tonnes) d'une usine de 1975 à 1994.

Année	t	X_t	Année	t	X_t
1	1975	19.8	11	1985	26.9
2	1976	21.2	12	1986	20.7
3	1977	17.4	13	1987	24.2
4	1978	18.9	14	1988	31.8
5	1979	25.2	15	1989	30.1
6	1980	27.2	16	1990	29.5
7	1981	27.4	17	1991	28.3
8	1982	21.1	18	1992	22.3
9	1983	23.9	19	1993	25.7
10	1984	30.2	20	1994	28.6

Puisque $n = 20$, on peut vérifier qu'on a bien :

$$\sum_{t=1}^{20} t = 210; \quad \sum_{t=1}^{20} X_t = 500.4; \quad \sum_{t=1}^{20} tX_t = 5527.3;$$

$$\sum_{t=1}^{20} t^2 = 2870; \quad \sum_{t=1}^{20} X_t^2 = 12857.42.$$

C'est-à-dire :

$$\beta_1 = 0.411; \quad \beta_0 = 20.708. \quad \text{Donc } T_t = 20.71 + 0.41t.$$

β_0, β_1 étant les coefficients de la droite de régression linéaire.

Ensuite, on tape les lignes de commande suivantes sous logiciel R :

```
a<-c(19.8,21.2,17.4,18.9,25.2,27.2,27.4,21.1,23.9,30.2,26.9,20.7,24.2,31.8,
30.1,29.5,28.3,22.3,25.7,28.6)
```

```
X=ts(c(a), frequency = 1, start = c(1, 1),end=20)
```

```
X
```

```
plot.ts(X)
```

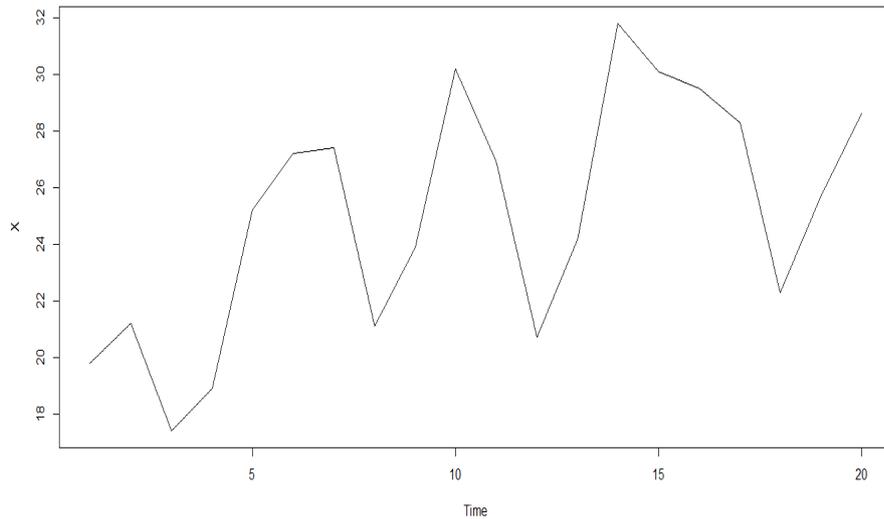


FIGURE 4.2 – Illustration d'une tendance linéaire estimée par régression.

Remarque :

On peut également obtenir une série stationnaire après avoir soustrait la tendance. Dans cet exemple, on pourrait soustraire de la série initiale la tendance linéaire obtenue à l'exemple 4.1.3.

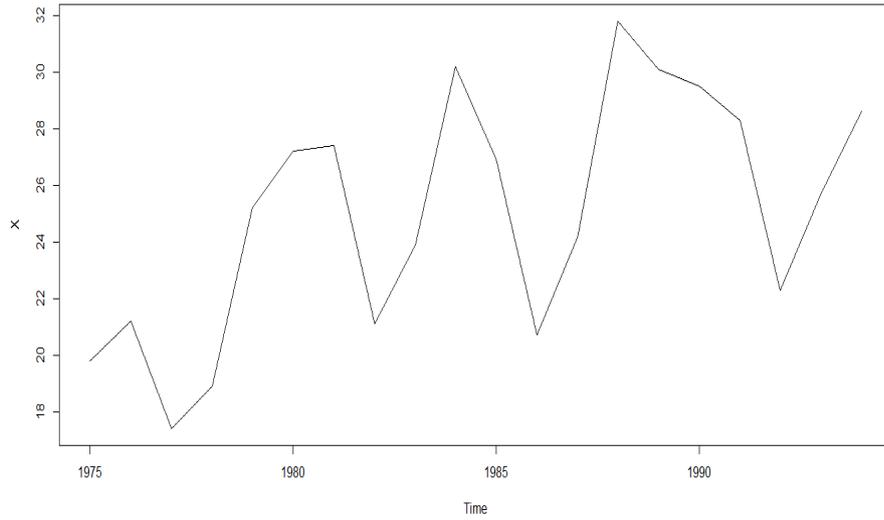


FIGURE 4.3 – La série initiale

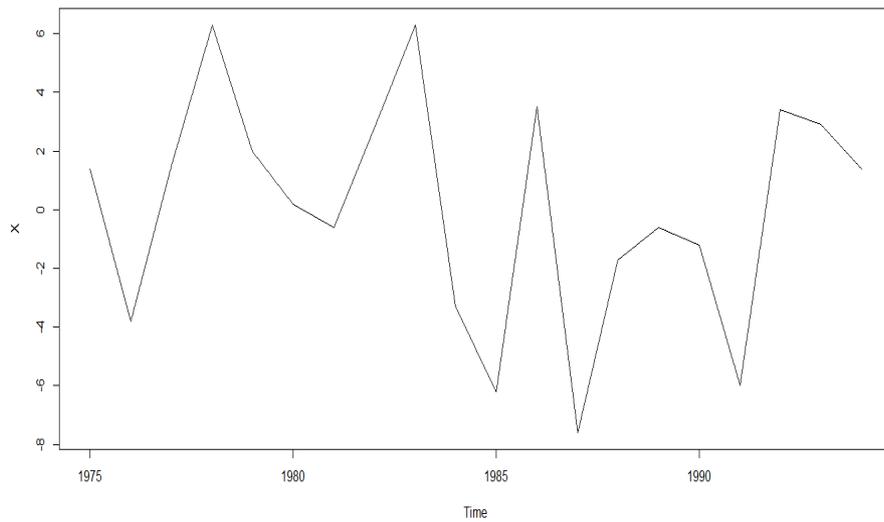


FIGURE 4.4 – La série des différences d'ordre 1.

Exemple 4.1.4. Pour les donnée de l'exemple (4.1.1), le tableau suivante présente les huit premières autocorrélations ε_t et les huit premières autocorrélations partielle $\varphi_{h,h}$ de la série des différence d'ordre 1.

h	1	2	3	4	5	6	7	8
r_h	-0.006	-0.631	-0.098	0.209	0.171	-0.083	-0.148	-0.055
$\varphi_{h,h}$	-0.006	-0.631	-0.180	-0.349	-0.033	-0.232	-0.092	-0.331

Les figures (4.5), (4.6) illustrent le comportement des fonction r_h et $\varphi_{h,h}$ pour la série des différences d'ordre 1.

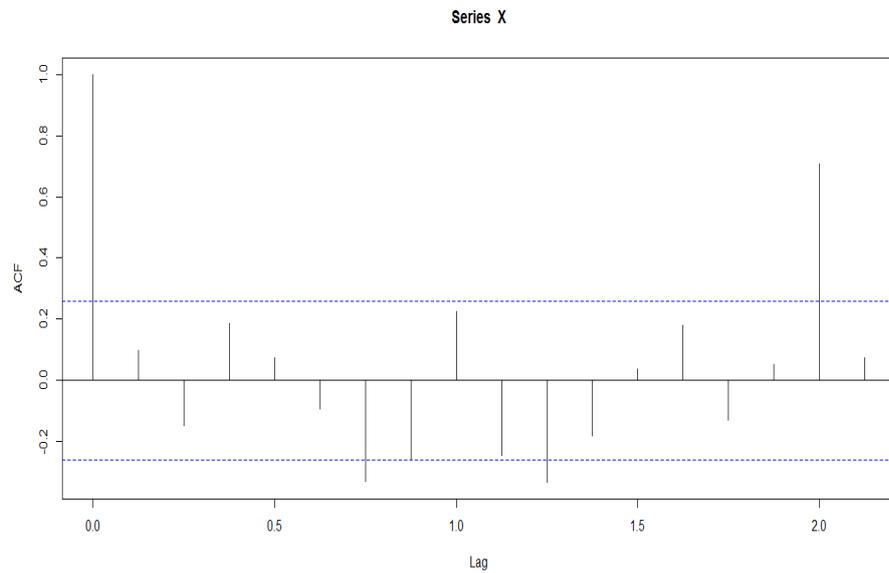


FIGURE 4.5 – La fonction r_h pour la série des différences d'ordre 1.

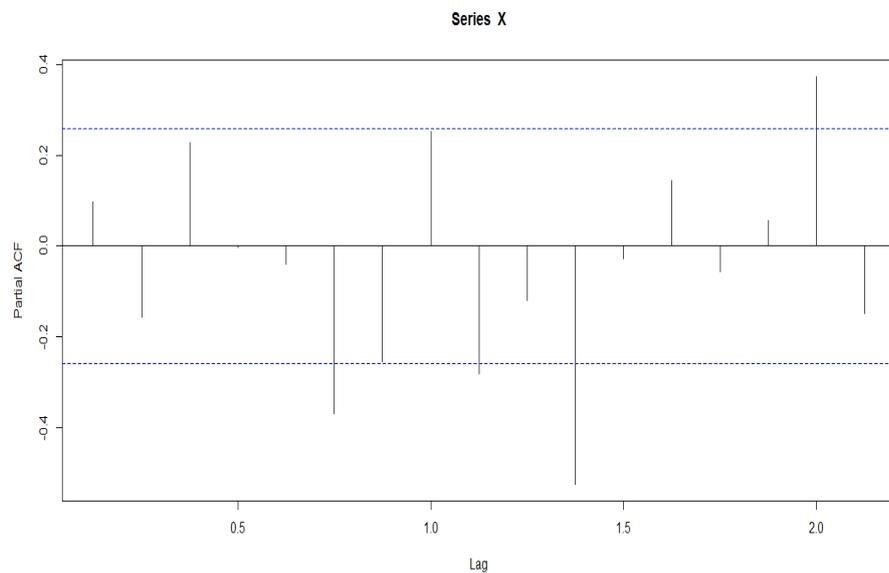


FIGURE 4.6 – La fonction $\varphi_{h,h}$ pour la série des différences d'ordre 1.

4.2 Estimation des paramètres et validation du modèle

Lorsqu'un modèle est identifié, la deuxième étape de la méthode de Box et Jenkins consiste à estimer les paramètres de ce modèle. Pour un modèle correspondant à un processus **ARIMA**(p, d, q) tel que défini par l'équation (3.3), on considère l'hypothèse que les erreurs (ou chocs) ε_t sont

de distribution normale et donc indépendantes. En partant de la fonction de densité conjointe des erreurs, on obtient celle des observations de la série.

La vraisemblance qui en résulte n'a toutefois pas une forme simple, et les estimations du "maximum de vraisemblance" recherchées s'obtiennent généralement par minimisation itérative d'une somme de carrés.

On distingue un certain nombre de méthodes d'approche pour obtenir les estimations recherchées : la méthode de la vraisemblance conditionnelle, la méthode du maximum de vraisemblance, la méthode des prévisions à rebours. Certaines méthodes sont utilisées conjointement : une première est utilisée pour déterminer des valeurs initiales et une deuxième pour effectuer les itérations. Notons que d'un point de vue pratique, l'obtention des estimations par maximum de vraisemblance requiert l'usage d'un logiciel.

Les propriétés des estimateurs du maximum de vraisemblance (e.v.m.) présentées au chapitre 2 sont également valides pour les séries chronologiques.

Précisément, pour un processus ARIMA(p, d, q), si $\varphi_1, \dots, \varphi_p$ et $\hat{\theta}_1, \dots, \hat{\theta}_q$ sont les e.v.m. des paramètres ϕ_1, \dots, ϕ_p et $\theta_1, \dots, \theta_q$ du processus obtenus avec une série chronologique de grande longueur T , on a alors approximativement

$$E(\varphi_j) = \phi_j, \quad j = 1, \dots, p \quad \text{et} \quad E(\hat{\theta}_h) = \theta_h, \quad h = 1, \dots, q;$$

$$\frac{\varphi_j - \phi_j}{\sqrt{\text{Var}(\varphi_j)}} \sim N(0, 1) \quad j = 1, \dots, p \quad \text{et} \quad \frac{\hat{\theta}_h - \theta_h}{\sqrt{\text{Var}(\hat{\theta}_h)}} \sim N(0, 1) \quad h = 1, \dots, q$$

Les variances $\text{Var}(\varphi_j)$ et $\text{Var}(\hat{\theta}_h)$ sont des fonctions des paramètres des φ_j et θ_h . Ces expressions sont les suivantes pour les modèles usuels :

$$AR(1) : \quad \text{Var}(\varphi_1) \simeq T^{-1}(1 - \phi_1^2)$$

$$AR(2) : \quad \text{Var}(\phi_1, \phi_2) \simeq T^{-1} \begin{pmatrix} 1 - \phi_2^2 & -\phi_1(1 + \phi_2) \\ -\phi_1(1 + \phi_2) & 1 - \phi_2^2 \end{pmatrix}$$

$$MA(1) : \quad \text{Var}(\hat{\theta}_1) \simeq T^{-1}(1 - \theta_1^2)$$

$$MA(2) : \quad \text{Var}(\hat{\theta}_1, \hat{\theta}_2) \simeq T^{-1} \begin{pmatrix} 1 - \theta_2^2 & -\theta_1(1 + \theta_2) \\ -\theta_1(1 + \theta_2) & 1 - \theta_2^2 \end{pmatrix}.$$

Les variances sont estimées à des l'aide des estimations des paramètres. Par exemple, pour un processus $AR(1)$, la variance et l'erreur type de φ_1 sont estimées par

$$\widehat{\text{Var}}(\varphi_1) = T^{-1}(1 - \varphi_1^2) \quad \text{et} \quad s.e.(\varphi_1) = \sqrt{\widehat{\text{Var}}(\varphi_1)}.$$

En général, les logiciels fournissent les estimations des paramètres ainsi que leurs erreurs types, c'est-à-dire φ_j et $s.e.(\varphi_j)$, $j = 1, \dots, p$, pour les paramètres autorégressifs du modèle ; $\hat{\theta}_i$ et $s.e.(\hat{\theta}_i)$, $i = 1, \dots, q$, pour les paramètres moyenne mobile du modèle. De plus, lorsque le nombre d'observations, T , de la série est grand, les estimations sont alors approximativement

distribuées selon des lois normales. Précisément, pour $j = 1, \dots, p$; $h = 1, \dots, q$

$$\frac{\varphi_j - \hat{\varphi}_j}{s.e.(\hat{\varphi}_j)} \sim N(0, 1), \quad \text{et} \quad \frac{\hat{\theta}_h - \theta_h}{s.e.(\hat{\theta}_h)} \sim N(0, 1).$$

Notons que les lois T de student sont parfois utilisées, mais que le nombre d'observations (et donc le nombre de degrés de liberté) étant généralement élevé, la loi normale constitue une approximation valide. Les résultats ci-dessus permettent de tester si chaque paramètre du modèle est significatif et donc, si nécessaire, de réduire le nombre de paramètres du modèle ou de changer de modèle. Ainsi,

- Pour tester $H_0 : \phi_j = 0$ contre $H_1 : \phi_j \neq 0$, la règle est de rejeter H_0 si $|t_{\varphi_j}| = \frac{|\varphi_j|}{s.e.(\hat{\varphi}_j)}$ est grand, c'est-à-dire si $|t_{\varphi_j}| > z_{\alpha/2}$, ou si p -value $\simeq 2(1 - \Phi(|t_{\varphi_j}|))$ est petite.
- Pour tester $H_0 : \theta_h = 0$ contre $H_1 : \theta_h \neq 0$, la règle est de rejeter H_0 si $|t_{\hat{\theta}_h}| = \frac{|\hat{\theta}_h|}{s.e.(\hat{\theta}_h)}$ est grand, c'est-à-dire si $|t_{\hat{\theta}_h}| > z_{\alpha/2}$, ou si p -value $\simeq 2(1 - \Phi(|t_{\hat{\theta}_h}|))$ est petite. Une analyse des résidus est également effectuée afin de valider si l'hypothèse de normalité et l'hypothèse d'homoscédasticité des chocs r_t est plausible.

4.3 Calcul de prévisions

Le calcul des prévisions par la méthode de Box et Jenkins peut être effectué en considérant directement la forme (3.2) d'un processus ARMA(p,q) ou une des formes alternatives. Pour simplifier la présentation, on suppose ici que la série X_t , $t = 1, \dots$, est stationnaire et inversible, et que les coefficients $\phi_j, j = 1, \dots, p$, $\theta_i, i = 1, 2, \dots, q$, sont connus. On suppose de plus que la moyenne $\mu = E(X_t) = 0$. Le modèle considéré pour la série est un processus d'équation

$$X_t = \phi_1 X_{t-1} + \dots + \phi_p X_{t-p} - \theta_1 \varepsilon_{t-1} - \dots - \theta_q \varepsilon_{t-q} + \varepsilon_t \quad (4.1)$$

c'est-à-dire $\phi(B)X_t = \theta(B)\varepsilon_t$,

ou $\phi(B)$ vérifie la condition de stationnarité et $\theta(B)$, celle d'inversibilité.

Supposons qu'on dispose des données de la série jusqu'au temps $t = T$ et qu'on aimerait obtenir une prévision d'origine T et d'horizon h (avec $h > 0$) pour X_{t+h} , notée $\hat{X}_t(h)$, de sorte que la moyenne du carré de l'erreur de prévision soit minimale. Autrement dit, on veut que

$$E((\hat{X}_t(h) - X_{t+h})^2)$$

$$\hat{X}_t(h) = E(X_{t+h} | X_t, X_{t-1}, \dots).$$

En se limitant aux prévisions linéaires on peut donc exprimer X_{t+h} selon l'équation (4.1)

$$X_{t+h} = \phi_1 X_{t+h-1} + \dots + \phi_p X_{t+h-p} - \theta_1 \varepsilon_{t+h-1} - \dots - \theta_q \varepsilon_{t+h-q} + \varepsilon_{t+h}.$$

Considérons la notation $E(Y | X_t, X_{t-1}, \dots) = [Y]$ pour l'espérance conditionnelle étant donné le présent et le passé X_t, X_{t-1}, \dots . On peut alors montrer que

$$\hat{X}_t(h) = [X_{t+h}] \quad (4.2)$$

$$= \phi_1[X_{t+h-1}] + \dots + \phi_p[X_{t+h-p}] - \theta_1[\varepsilon_{t+h-1}] - \dots - \theta_q[\varepsilon_{t+h-q}] + [\varepsilon_{t+h}].$$

Remarque Lors du calcul d'une prévision avec l'équation (4.2), on considère que

$$\begin{aligned} [X_{t-j}] &= X_{t-j}, \text{ pour } j = 0, 1, 2, \dots \\ [X_{t+j}] &= \widehat{X}_{t-j}, \text{ pour } j = 0, 1, 2, \dots \\ [\varepsilon_{t-j}] &= \varepsilon_{t-j} = X_{t-j} - \widehat{X}_{t-j-1}(1), \text{ pour } j = 0, 1, 2, \dots \\ [\varepsilon_{t+j}] &= 0, \text{ pour } j = 0, 1, 2, \dots \end{aligned}$$

Précisément, lors du calcul, on applique le principe suivant : les X_{t-j} , $j = 0, 1, 2, \dots$ (données disponibles), sont laissées telles quelles ; les X_{t-j} , $j = 1, 2, \dots$ (données non disponibles), sont remplacées par leurs prévisions $\widehat{X}_t(j)$; les ε_{t-j} , $j = 0, 1, 2, \dots$ (chocs du passé), sont calculés comme $\varepsilon_{t-j} = X_{t-j} - \widehat{X}_{t-j-1}(1)$; finalement les chocs ε_{t+j} , $j = 1, 2, \dots$, sont remplacés par leur moyenne non conditionnelle, c'est-à-dire 0.

Il existe d'autres expressions équivalentes à l'équation (4.2). L'une de celles-ci utilise les coefficients du polynôme $\psi(B)$ qu'on obtient de ceux de $\phi(B)$ et $\theta(B)$ par

$$\phi(B)\psi(B),$$

c'est-à-dire

$$\psi(B) = \phi^{-1}(B)\theta(B). \quad (4.3)$$

En effet, de l'équation (4.1), puisque le processus est inversible et stationnaire, on a donc

$$\phi(B)X_t = \theta(B)\varepsilon_t, \text{ et } X_t = \psi(B)\varepsilon_t,$$

d'où l'équation (4.3).

Pour cette forme de prévision, on peut établir les résultats suivants :

1. La prévision (4.2) peut également s'écrire

$$\begin{aligned} \widehat{X}_t &= [X_{t+h}] = \psi_h\varepsilon_t + \psi_{h+1}\varepsilon_{t-1} + \dots \\ &= \sum_{l=0}^{\infty} \psi_{h+l}\varepsilon_{t-l}, \text{ pour tout } h > 0. \end{aligned}$$

2. L'erreur de la prévision de X_{t+h} par $\widehat{X}_t(h)$ est

$$e_n(h) = X_{t+h} - \widehat{X}_t(h) = \varepsilon_{t+h} + \psi_1\varepsilon_{t+h-1} + \dots + \psi_{h-1}\varepsilon_{t+1}.$$

3. L'erreur $e_n(h)$ est de moyenne nulle puisque les chocs ε_t le sont, et il en résulte que $\widehat{X}_t(h)$ est une estimation (prévision) sans biais pour X_{t+h} .

4. La variance de l'erreur $e_n(h)$ est

$$V(h) = \text{Var}(e_n(h)) = (1 + \psi_1^2 + \dots + \psi_{h-1}^2)\sigma_\varepsilon^2.$$

5. Lorsque l'hypothèse de la normalité des chocs ε_t est valide, on peut alors calculer un intervalle de prévision pour la valeur future X_{t+h} par

$$X_{t+h} \in \widehat{X}_t(h) \pm z_{\alpha/2} \sqrt{(1 + \psi_1^2 + \dots + \psi_{h-1}^2)S_\varepsilon^2}$$

ou S_ε^2 est une estimation de σ_ε^2 (obtenue avec les résidus).

Exemple 4.3.1. *Illustration du calcul de prévisions dans le cas d'un modèle ARIMA(1,0,1). On suppose que les coefficients ϕ_1 et θ_1 sont connus (ou estimés). On a*

$$(1 - \phi_1 B)X_t = (1 - \theta_1 B)\varepsilon_t.$$

On suppose que l'on dispose des données de la série jusqu'au temps $t = n$, X_1, X_2, \dots, X_n . Donc, pour un horizon $h \geq 1$, on peut écrire

$$X_{n+h} = \phi_1 X_{n+h-1} - \theta_1 \varepsilon_{n+h-1} + \varepsilon_{n+h}.$$

On en déduit que $\widehat{X}_n(h) = \phi_1[X_{n+h-1}] - \theta_1[\varepsilon_{n+h-1}]$.

D'où les prévisions successives suivantes :

$$\begin{aligned} \widehat{X}_n(1) &= \phi_1 X_n - \theta_1 \varepsilon_n \\ \widehat{X}_n(2) &= \phi_1 \widehat{X}_n(1) \\ \widehat{X}_n(3) &= \phi_1 \widehat{X}_n(2) = \phi_1^2 \widehat{X}_n(1) \\ &\vdots \\ &\vdots \\ &\vdots \\ \widehat{X}_n(n) &= \phi_1^{j-1} \widehat{X}_n(1), \quad j \geq 2. \end{aligned}$$

Notons que ε_n est nécessaire au calcul de $\widehat{X}_n(1)$ et vaut $\varepsilon_n = X_n - \widehat{X}_{n-1}(1)$.

Exemple 4.3.2. *Illustration du calcul de prévisions dans le cas d'un modèle ARIMA(1,1,0). On suppose que les paramètres ϕ_1 et c sont connus (ou estimés). On a*

$$\begin{aligned} \varepsilon_t &= (1 - \phi_1 B) \nabla \widehat{X}_t \\ \varepsilon_t &= (1 - \phi_1 B)(1 - B)\widehat{X}_t \\ \varepsilon_t &= (1 - (1 + \phi_1)B + B^2)\widehat{X}_t \end{aligned}$$

C'est-à-dire

$$X_t = c + (1 + \phi_1)X_{t-1} - \phi_1 X_{t-2} + \varepsilon_t. \quad (4.4)$$

On suppose que l'on dispose des données de la série jusqu'au temps $t = n$, X_1, X_2, \dots, X_n . De l'équation (4.4), et pour un horizon $h \geq 1$, on peut écrire

$$X_{n+h} = c + (1 + \phi_1)X_{n+h-1} - \phi_1 X_{n+h-2} + \varepsilon_{n+h}.$$

On en déduit que $\widehat{X}_n(h) = c + (1 + \phi_1)[X_{n+h-1}] - \phi_1[X_{n+h-2}] + [\varepsilon_{n+h}]$. D'où les prévisions suivantes :

$$\begin{aligned}\widehat{X}_n(1) &= c + (1 + \phi_1)[X_n] - \phi_1[X_{n-1}] + [\varepsilon_{n+1}] \\ &= c + (1 + \phi_1)X_n - \phi_1X_{n-1}.\end{aligned}$$

$$\begin{aligned}\text{De plus } \widehat{X}_n(2) &= c + (1 + \phi_1)[X_{n+1}] - \phi_1[X_n] + [\varepsilon_{n+2}] \\ &= c + (1 + \phi_1)\widehat{X}_n(1) - \phi_1X_n.\end{aligned}$$

$$\text{et } \widehat{X}_n(3) = c + (1 + \phi_1)\widehat{X}_n(2) - \phi_1\widehat{X}_n(1).$$

Chapitre 5

Exemples d'applications

Les séries chronologiques ont des applications dans beaucoup de domaines par exemple :

- En finance elle modélise les valeurs des actions boursières.
- En croissance économique, évolution climatologique à long terme.
- En météorologie, elle modélise la température du jour.
- Énergie : prévision de la consommation d'électricité

5.1 L'estimation des coefficients ARMA sous logiciel R

Exemple 5.1.1. On considère la processus ARMA(2,2) suivant :

$$X_t - X_{t-1} + \frac{1}{8}X_{t-2} = \varepsilon_t - \varepsilon_{t-1} + \varepsilon_{t-2}$$

Pour avoir le graphe de la série chronologique on tape les instructions suivantes :

```
monmodele=list(ar=c(1,-0.125),ma=c(1,-0.25))
x=arima.sim(model=monmodele,n=100,rand.gen=rnorm)
plot.ts(x)
```

FIGURE 5.1 – Graphe de la série chronologique

Maintenant, on va faire comme si ne connaissait pas les coefficients et donner seulement l'ordre donc ARIMA(2,0,2)

```
x.ord=c(2,0,2)
x.arima=arima(x,order=x.ord)
x
```

```

[1] 3.0757580 2.6785311 2.9384490 2.1234210 0.8348832 0.6405096
[7] 0.2909609 -1.8269692 0.0442448 1.8180343 1.6376503 1.4231154
[13] -0.7720584 -1.0824324 -0.7699031 -1.4404487 -1.8757890 -1.4684131
[19] 0.8707307 1.9848475 2.1078921 1.9926176 -1.2302398 -3.2284884
[25] -3.4238014 -4.7713515 -4.1803750 -3.1916205 -1.2972553 -0.6334049
[31] 0.9083295 2.9280997 3.4677457 3.2546412 0.5952487 -2.8420652
[37] -3.0924407 0.9732007 3.7192281 3.3023188 4.5418373 4.7680536
[43] 4.8659300 8.1473999 8.4009645 6.0297963 4.8867617 3.3688883
[49] 3.0113815 4.7880639 4.7859315 5.0784469 3.1860825 0.5783610
[55] 0.7049111 0.4367237 -0.8306197 1.5003349 3.2688433 2.4387450
[61] 2.1314130 2.6322561 3.4816696 1.0895145 -0.4882363 0.2825068
[67] -0.3948105 0.0145689 2.1279086 2.6674661 1.5260111 0.3041060
[73] 1.1255847 1.5064616 0.5810112 1.9215615 3.8765735 5.1283459
[79] 5.9158909 5.5475497 5.4241682 5.5766761 3.8905231 2.8310171
[85] 3.6105916 2.6177044 0.8916162 1.1463136 0.9985654 -1.4782531
[91] -1.8510180 0.5533920 0.9856236 0.5116335 0.5066352 2.4907445
[97] 5.2390227 1.5473554 -1.3641885 -0.1298123

```

FIGURE 5.2 – Simulation des coefficients du processus ARMA

Les valeurs données dans la figure 5.2. sont très proches de celles données par la figure 5.1. Pour augmenter la précision des estimations on peut prendre n plus grand.

Exemple 5.1.2. *On considère les températures moyennes mensuelles pour la période (2010-2018) dans la région de Ghardaia*

	janvier	fevrier	mars	avril	mai	juin	juillet	Aout	sep	oct	nov	dec
2010	13.6	16.5	18.5	22.5	24.3	31.4	35.1	34.7	28.8	22.4	17.0	14.1
2011	12.6	12.4	15.2	21.8	24.6	29.4	35.1	23.6	31.2	20.9	16.3	12.6
2012	10.8	9.7	16.1	21.0	27.8	34.0	26.4	27.4	29.1	24.7	17.1	12.6
2013	12.7	12.6	18.7	21.5	25.6	30.0	34.6	32.1	29.5	26.9	16.5	10.5
2014	12.3	14.4	15.8	22.7	26.8	30.1	35.5	35.6	31.4	24.4	17.5	11.4
2015	11.0	11.5	16.4	23.0	27.9	30.6	33.4	33.7	29.3	23.7	16.4	12.4
2016	13.8	14.7	17.0	22.7	26.8	31.3	34.2	33.0	29.3	25.2	16.6	12.6
2017	9.80	14.7	18.1	21.3	28.5	31.4	33.9	33.7	28.1	21.8	16.0	11.5
2018	12.8	12.1	17.9	20.8	24.8	30.6	37.9	31.6	29.3	22.0	16.3	13.3

TABLE 5.1 – Moyennes mensuelles des températures

Ensuite, on tape les lignes de commande suivantes sous logiciel R :

```

a<-c(13.6,16.5,18.5,22.5,24.3,31.4,35.1,34.7,28.8,22.4,17.0,14.1)
b<-c(12.6,12.4,15.2,21.8,24.6,29.4,35.1,23.6,31.2,20.9,16.3,12.6)
d<-c(10.8,9.7,16.1,21.0,27.8,34.0,26.4,27.4,29.1,24.7,17.1,12.6)
e<-c(12.7,12.6,18.7,21.5,25.6,30.0,34.6,32.1,29.5,26.9,16.5,10.5)
f<-c(12.3,14.4,15.8,22.7,26.8,30.1,35.5,35.6,31.4,24.4,17.5,11.4)
g<-c(11.0,11.5,16.4,23.0,27.9,30.6,33.4,33.7,29.3,23.7,16.4,12.4)
h<-c(13.8,14.7,17.0,22.7,26.8,31.3,34.2,33.0,29.3,25.2,16.6,12.6)
j<-c(9.80,14.7,18.1,21.3,28.5,31.4,33.9,33.7,28.1,21.8,16.0,11.5)
k<-c(12.8,12.1,17.9,20.8,24.8,30.6,37.9,31.6,29.3,22.0,16.3,13.3)
x=ts(c(a,b,d,e,f,g,h,j,k), frequency = 12, start = c(2010, 1),end=2019)
x
plot.ts(x)
acf(x)
pacf(x)

```

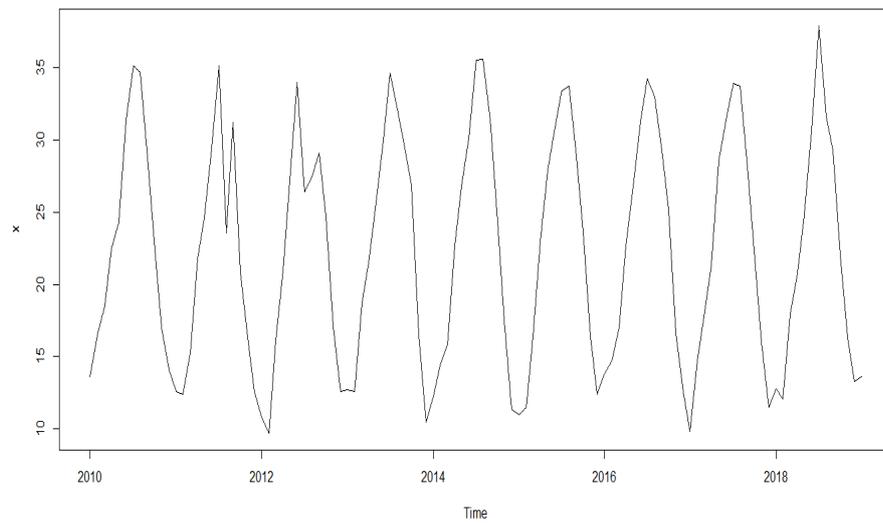


FIGURE 5.3 – Graphe de température

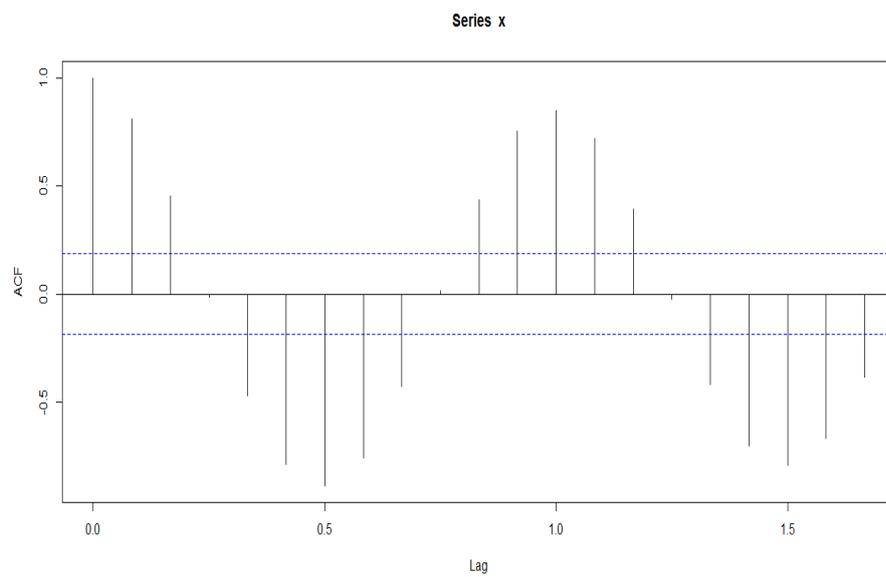


FIGURE 5.4 – Corrélogramme

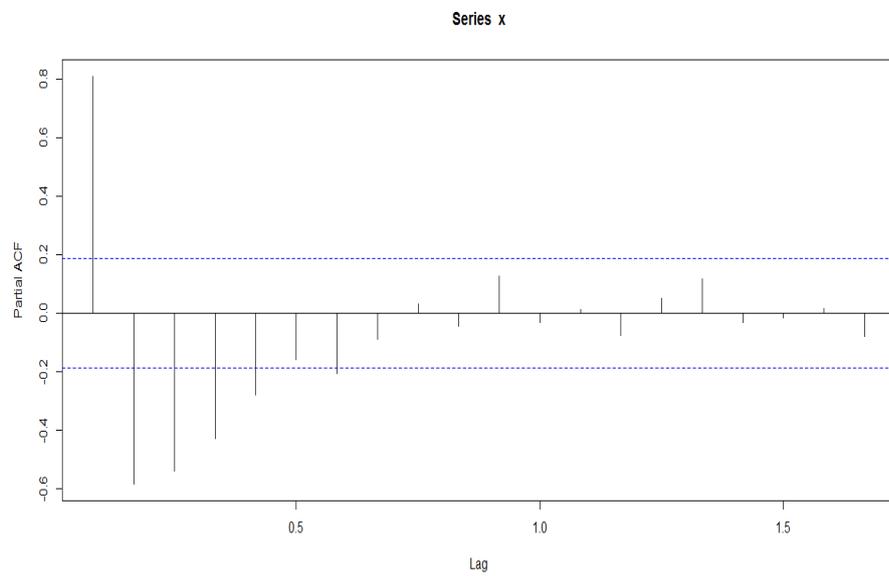


FIGURE 5.5 – Corrélogramme partiel

Conclusion

Dans ce travail on a étudié les processus autorégressifs moyenne mobile (ARMA) qui ont des application dans beaucoup de domaines.

En finance, on utilise ces processus pour modéliser les prix des action boursières.

En météorologie, la température du jour peut être modélisée par un processus autorégressif.

L'utilité de ces modèles réside dans la prévision dans le futur des valeurs prises par le phénomène étudié.

Actuellement les spécialistes arrivent à donner par exemple les prix du pétrole à court et à moyen terme ce qui est très utile pour les pays dépendants des hydrocarbures pour équilibrer leurs balances financières.

Bien entendu les valeurs prévisionnelles du processus ne sont que des estimations qui peuvent changer à cause de paramètres géopolitiques mais qui sont assez précises pour être prises en considération.

C'est le cas aussi en météorologie ; les services de chaque région arrivent à donner les température prévisionnelles plusieurs jours à l'avance et ceci grâce à l' utilisation de la modélisation par les processus stochastiques et la simulation à l'aide des logiciels.

Il est important de mettre l'accent sur l'hypothèse de stationnant des processus étudié qui n'est pas toujours assurée en pratique, l'utilisation des différences premiers permet de stationnant quelques séries.

Dans le cas général on utilise les ARCH ou bien les GARCH.

Box et Jenkins (1970) ont proposée une méthode qu'a révolutionné les processus ARMA qui trouvent beaucoup d'application pratiques.

Bibliographie

- [1] Adjenque Luc, *méthodes statistiques concepts, applications et exercices*, presses internationales polytechnique, canada (2014)
- [2] Box George and Jenkins Gwilym, *time series analysis forecasting and control Revised Edition*. University of wisconsin and University of lancaster
- [3] Box George et Jenkis Gwilym, *series analysis* (1976)
- [4] Lagnoux Agnès. *Séries chronologiques*. Université le Mirail Toulouse.
- [5] Lagnoux Agnès. *Renforcement statistique*. Université le Mirail Toulouse.
- [6] Lecoutre Jean Pierre ,*Statistique et probabilités*, maître de conférences à l'université Panthéon-Assas (Paris II)
- [7] Legaie Sylvia et Lecoutre Jean Pierre. *Méthodes Statistique*. à l'université Panthéon-Assas (Paris II)
- [8] Perradin Corinne. *séries chronologiques*. 2004-2005. Université Paris I
- [9] Yan Girard. *séries chronologiques à une et plusieurs variables*. Université du Québec. mai 2010