

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

UNIVERSITÉ MOHAMED KHIDER, BISKRA

FACULTÉ des SCIENCES EXACTES et des SCIENCES de la NATURE et de la VIE

DÉPARTEMENT DE MATHÉMATIQUES



Mémoire présenté en vue de l'obtention du Diplôme :

MASTER en Mathématiques

Option : **Statistique**

Par

DENDOUGA Samah

Titre :

Séries Temporelles : Théorie et Application

Membres du Comité d'Examen :

Dr. YAHIA Djabrane	UMKB	Président
Dr. BENAMEUR Sana	UMKB	Encadreur
Dr. DJABER Ibtisem	UMKB	Examineur

21 Septembre 2020

DÉDICACE

Je dédie ce modeste travail

À ma très cher Mère et mon très cher Père

À ceux qui m'ont toujours encouragé pour que je réussisse dans mes études

À tous mes frères

À toute ma famille

À tous ceux qui m'ont enseigné

À tous mes compagnons de promotion et à tous ce que ma réussite leur tient à cœur

Que ce travail soit le témoignage de ma gratitude et de mon profond respect.

REMERCIEMENTS

Tout d'abord, je remercie ALLAH Le Tout Puissant qui m'a accordée la volonté et le courage pour réaliser ce mémoire.

Mes plus vifs remerciements et ma profonde gratitude vont à Dr. Benameur Sana qui a accepté de diriger ce travail. Sa grande disponibilité et ses encouragements ont joué un rôle important dans la réalisation de ce mémoire.

Je remercie sincèrement, les membres de jury : Pr : Yahia Djabrane et Dr. Djaber Ibtissem, d'avoir bien voulu accepter de faire partie de la commission d'examineur.

Je remercie également toute l'équipe pédagogique de notre département de mathématiques de l'université de Mohamed Khider et les intervenants professionnels responsables.

Merci à tous et à toutes.

Table des matières

Remerciements	ii
Table des matières	iii
Table des figures	vi
Introduction	1
1 Analyse d'une série temporelle	2
1.1 Décomposition d'une série temporelle	2
1.2 Types de schéma d'une série temporelle	3
1.3 Opérateurs et moyennes mobiles	6
1.4 Calcul de prévision	9
1.4.1 Méthode de Buys-Ballot	9
1.4.2 Lissage exponentiel	10
2 Modèles stationnaires	13
2.1 Propriétés statistiques	13
2.2 Modèles stationnaires	16
2.3 Densité spectrale	19
2.4 Modèles linéaires	20

2.5	Conditions de stationnarité et d'inversibilité d'un modèle sous l'écriture polynomiale	21
3	Modèles ARMA	23
3.1	Modèle autorégressif	23
3.1.1	Autocovariance et autocorrélation d'un $AR(p)$	24
3.1.2	Autocorrélation partielle	27
3.1.3	Modèle autorégressif d'ordre infini	30
3.2	Modèle à moyenne mobile	30
3.2.1	Autocovariance et Autocorrélation d'un $MA(q)$	31
3.2.2	Modèle à moyenne mobile d'ordre infini	33
3.2.3	Théorème de Wold	33
3.3	Modèle ARMA	34
3.3.1	Fonction génératrice d'autocovariance	35
3.4	Modèles ARIMA et SARIMA	36
3.5	Prévision par la méthode de Box-Jenkins	37
3.5.1	Identification	37
3.5.2	Estimation des paramètres	39
3.5.3	Validation	39
3.5.4	Prévision	42
3.6	Application sous \mathbb{R}	46
3.6.1	Analyse d'une série temporelle sous \mathbb{R}	46
3.6.2	Simulation des modèles stationnaire sous \mathbb{R}	50
3.6.3	Simulation des modèles linéaire sous \mathbb{R}	51
3.6.4	Méthode de Box-Jenkins sous \mathbb{R}	53
	Conclusion	60

Bibliographie	61
Annexe : Abréviations et Notations	63

Table des figures

1.1	Représentation graphique d'un schéma additif (série co2)	4
1.2	Représentation graphique d'un schéma multiplicatif (Série AirPassengers)	4
1.3	Représentation graphique d'un schéma mixte	5
1.4	Représentation graphique d'un exemple d'élimination de la saisonnalité S_t par moyenne mobile	8
2.1	Représentation graphique d'un Bruit Blanc de taille 200 ($\mathcal{BB}(0, 2.5)$)	17
2.2	Représentation graphique d'une marche aléatoire	17
3.1	Représentation graphique d'une trajectoire (en haut), autocorrélogramme (au milieu) et autocorrélogramme partiel (en bas) d'un modèle $AR(1)$	28
3.2	Représentation graphique d'une trajectoire, autocorrélogramme et autocor- rélogramme partiel d'un $AR(2)$	29
3.3	Représentation graphique d'une trajectoire (en haut), autocorrélogramme (en bas) d'un modèle $MA(1)$	32
3.4	Organigramme de la méthode de Box & Jenkins	38
3.5	Décomposition de la série co2	48
3.6	Représentation graphique de la série co2 après différentiation	48
3.7	Décomposition de la série $\log(X_t)$	49
3.8	Représentation graphique de la série $\log(X_t)$ après une double différentiation	50
3.9	Autocorrélogramme et autocorrélogramme partiel de la série Y_t	55

3.10 Analyse des résidus	58
3.11 Représentation graphique des valeurs prédictives de la série AirPassenegrs sur 2 ans	59

Introduction

Une série temporelle ou série chronologique est une suite formée d'observations au cours du temps. L'analyse des séries temporelles est un outil couramment utilisé de nos jours pour mieux comprendre et décrire l'évolution d'un phénomène présenté par la série. L'idée est de prendre un échantillon de données et de construire le meilleur modèle qui ajuste ces données. Ce modèle nous permet de tirer certaines conclusions sur la série, et de faire une prédiction du futur de la série à partir de la partie observée.

La théorie des séries temporelles trouve des applications dans de très nombreux domaines : Finance et économétrie (évolution des indices boursiers, des prix, des productions agricoles ou industrielles), Assurance (analyse des sinistres), Astronomie, Médecine, Biologie, ...

L'objectif de ce travail est de regrouper les différentes techniques d'analyse et de modélisation des séries temporelles afin de pouvoir faire des prévisions.

Ce mémoire s'organise en trois chapitres :

Le premier chapitre est consacré à l'analyse d'une série temporelle où on expose des préliminaires, quelques notions et définitions sur les séries chronologiques.

Dans le second chapitre, nous illustrons les propriétés statistiques d'une série temporelle, les modèles stationnaires et la densité spectrale.

Le troisième chapitre porte sur les modèles linéaires : AR , MA , $ARMA$, $ARIMA$ et $SARIMA$. Dans un deuxième volet nous présentons la méthode de prévision des séries temporelles de Box et Jenkins et les différentes étapes spécifient cette méthode, pour arriver finalement à la mise en œuvre des séries temporelles sous R.

Chapitre 1

Analyse d'une série temporelle

Dans ce chapitre, certaines notations et définitions importantes liés à l'analyse des séries chronologiques, dont on aura besoin dans la suite de ce mémoire, seront abordées. Nous commençons tous d'abord par la présentation d'une série temporelle et leur décomposition. Une bonne référence pour la théorie des séries temporelles les livres de Brockwell & Davis (2016) [6] et Anderson (1971) [1].

1.1 Décomposition d'une série temporelle

Définition 1.1.1 (Série temporelle)

Une série temporelle (ou encore une série chronologique), notée X_t , est une suite d'observations repérées dans le temps, elle représente l'évolution d'un phénomène au cours du temps, telle que

$$X_t = \{X_t, t = 1, \dots, T\}, T \in \mathbb{R}, \mathbb{Z}, \mathbb{N}, \dots$$

– Dans la suite le temps t est une partie de l'année (jour, mois, semaine, ...) et $T \in \mathbb{N}$.

On considère que X_t est la résultante de différentes composantes fondamentales :

1. Tendance : On appelle une tendance, une fonction notée Z_t , qui représente l'aspect générale d'une série temporelle X_t à long terme.

Il y a trois types de tendance : $\forall a, b, c, d \in \mathbb{R}$ et $n \in \mathbb{N}$

a) Tendence polynomiale : $Z_t = a + bt + ct^2 + \dots + dt^n$, (linéaire si $n = 1$)

b) Tendence exponentielle : $Z_t = a \times \exp(bt + c)$.

c) Tendence logistique : $Z_t = \frac{c}{a + bt}$.

2. Saisonnalité : Une saisonnalité (ou composante saisonnière) correspond à un phénomène qui se répète à un intervalle de temps régulier (périodique), l'étendue de cet intervalle appelé période, notée p . Elle est donc liée au rythme imposé par les saisons (toutes les 12 périodes pour des données mensuelles, toutes les 4 périodes pour des données trimestrielles, ...).

Très précisément, on appelle composante saisonnière, notée S_t , une fonction périodique de période p (c-à-d $S_t = S_{t \pm p}$). Cette fonction est entièrement déterminée par ses p composantes :

$\gamma_1, \gamma_2, \dots, \gamma_p$. On suppose que

$$\sum_{j=1}^p \gamma_j = 0.$$

En d'autre terme $S_t = \gamma_1 S^1 + \gamma_2 S^2 + \dots + \gamma_p S^p$, telle que :

$$S^j = \begin{cases} 1 & \text{si } j = t, \\ 0 & \text{si } j \neq t. \end{cases}$$

3. Résidus : On appelle résidu (erreur) ou variation irrégulière, notée ε_t , la variable aléatoire à effet de faible intensité et à court durée.

1.2 Types de schéma d'une série temporelle

On modélise souvent une série temporelle par un des modèles suivants :

1. Modèle additif : C'est le modèle classique de décomposition dans le traitement des modèles d'ajustement. La variable X_t s'écrit comme la somme des trois composantes :

$$X_t = Z_t + S_t + \varepsilon_t, \quad t = 1, \dots, T.$$

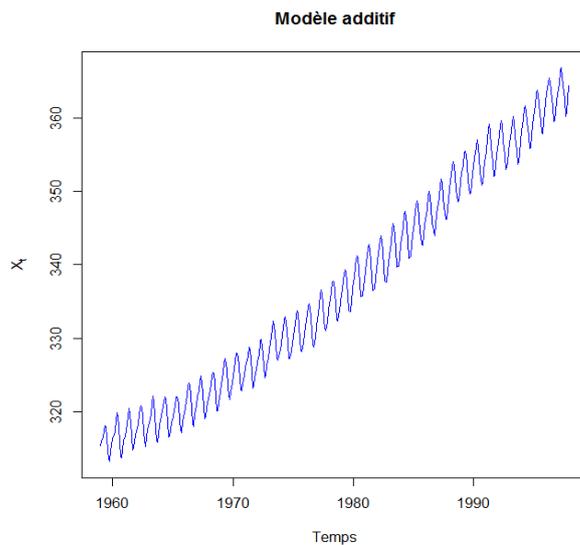


FIG. 1.1 – Représentation graphique d'un schéma additif (série co2)

2. Modèle multiplicatif : Dans ce cas, la série X_t s'écrit comme le produit des composantes :

$$X_t = Z_t \times S_t \times \varepsilon_t, \quad t = 1, \dots, T.$$

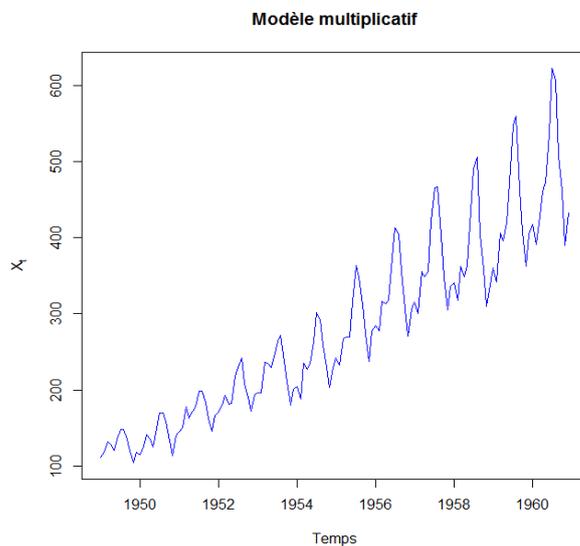


FIG. 1.2 – Représentation graphique d'un schéma multiplicatif (Série AirPassengers)

En pratique, dans le cas d'une série X_t à valeurs positives on se ramène à un modèle additif par passage au logarithme, puisqu'on a :

$$\ln(X_t) = \ln(Z_t \times S_t \times \varepsilon_t) = \ln(Z_t) + \ln(S_t) + \ln(\varepsilon_t)$$

3. Modèle mixte : Il s'agit des différentes combinaisons de modèles additifs et de modèles multiplicatifs, par exemple :

$$X_t = Z_t \times S_t + \varepsilon_t$$

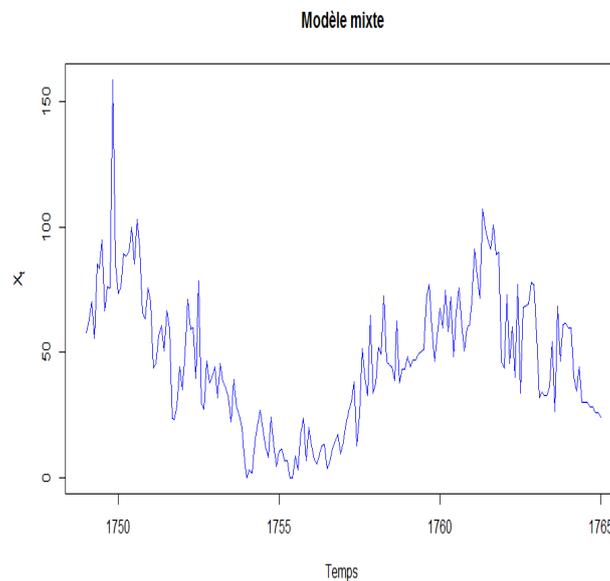


FIG. 1.3 – Représentation graphique d'un schéma mixte

Remarque 1.2.1

Afin de faire la distinction entre les types de modèle, on peut se baser sur une méthode graphique. Sur le graphique de la série, on trace une droite passant par les minima de la courbe et une autre droite passant par les maxima pour chaque saison. Cette méthode s'appelle la méthode de Bande.

Si ces droites sont parallèles, c-à-d l'amplitude de la composante saisonnière et de l'erreur reste constante autour de la tendance au cours du temps, on est en présence d'un modèle additif.

Par contre, si les droites ne sont pas parallèles, c-à-d l'amplitude de la composante saisonnière varie de façon exponentielle dans le temps, on est en présence d'un modèle multiplicatif.

1.3 Opérateurs et moyennes mobiles

Définition 1.3.1 (Opérateur retard et avance)

L'opérateur retard L sur une série temporelle est défini par :

$$LX_t = X_{t-1}, \forall t \in T.$$

On note d'une façon générale :

$$L^k X_t = X_{t-k}, \forall k \in \mathbb{N}^*.$$

L'opérateur avance F sur une série temporelle est défini par :

$$FX_t = L^{-1}X_t = X_{t+1}, \forall t \in T.$$

On note aussi

$$F^j X_t = L^{-j}X_t = X_{t+j}, \forall j \in \mathbb{N}^*.$$

Définition 1.3.2 (Opérateur de différentiation)

L'opérateur de différentiation (à l'ordre 1) Δ d'une série temporelle est déterminé par :

$$\Delta X_t = X_t - X_{t-1} = X_t - LX_t = (1 - L)X_t, \forall t \in T.$$

En générale, la différentiation à l'ordre k de la série X_t est défini par :

$$\Delta^k X_t = (1 - L)^k X_t, \forall t \in T.$$

Remarque 1.3.1

1. *L'opérateur de différentiation d'ordre k permet d'éliminer (ou encore réduire à une constante ou 0) une tendance polynomiales de degré k , pour un modèle additif.*
2. *L'opérateur de différentiation saisonnière, noté Δ_k , donnée par :*

$$\Delta_k = 1 - L^k,$$

permet d'éliminer une saisonnalité de période k , pour un modèle additif.

Définition 1.3.3 (Moyenne mobile)

Une moyenne mobile (ou encore filtre), notée $M(L)$ est un opérateur linéaire de forme générale :

$$M(L) = \sum_{j=-m_1}^{m_2} \theta_j L^j,$$

où $\theta_j \in \mathbb{R}$ pour tout j , $(m_1, m_2) \in \mathbb{N} \times \mathbb{N}$.

L'ordre de la moyenne mobile (ou encore le nombre de termes considérés dans la somme) est l'entier $m_1 + m_2 + 1$.

Propriétés 1.3.1

- 1) *Une moyenne mobile $M(L)$ conserve X_t (ou encore laisse invariante) la série si :*

$$M(L) X_t = X_t.$$
- 2) *Une série X_t est absorbée par une moyenne mobile $M(L)$ si : $M(L) X_t = 0$.*
- 3) *Une moyenne mobile est centré lorsque : $m_1 = m_2 = m$.*
- 4) *Une moyenne mobile centré symétrique si et seulement si : $\theta_{-i} = \theta_i$; $i = 1, \dots, m$.*
- 5) *Une moyenne mobile arithmétique est centré, d'ordre (impaire) $2m + 1$ et telle que :*

$$\theta_i = \frac{1}{2m + 1} ; \forall i = -m, \dots, m.$$

- 6) *Le moyenne mobile est normalisée si : $\sum_{j=-m_1}^{m_2} \theta_j = 1$.*

Propriété 1.3.1 (Elimination de la composante saisonnière)

Si la série X_t possède une composante saisonnière de période p alors l'application d'une moyenne mobile d'ordre p supprime cette saisonnalité. Deux cas se présentent :

– Si p est impaire, on applique la moyenne mobile

$$\begin{aligned} M(L) X_t &= \frac{1}{p} \sum_{i=-q}^q X_{t+i} \text{ où } q \text{ vérifié } p = 2q + 1, \\ &= \frac{1}{2q + 1} [L^q X_t + \dots + LX_t + X_t + L^{-1} X_t + \dots + L^{-q} X_t]. \end{aligned}$$

– Si p est paire, on applique la moyenne mobile

$$\begin{aligned} M(L) X_t &= \frac{1}{p} \left[\sum_{i=-q+1}^{q-1} X_{t+i} + \frac{X_{t-q} + X_{t+q}}{2} \right] \text{ où } q \text{ vérifié } p = 2q \\ &= \frac{1}{2q} \left[\frac{1}{2} L^q X_t + L^{q-1} X_t + \dots + LX_t + X_t + L^{-1} X_t + \dots + L^{-q+1} X_t + \frac{1}{2} L^{-q} X_t \right]. \end{aligned}$$

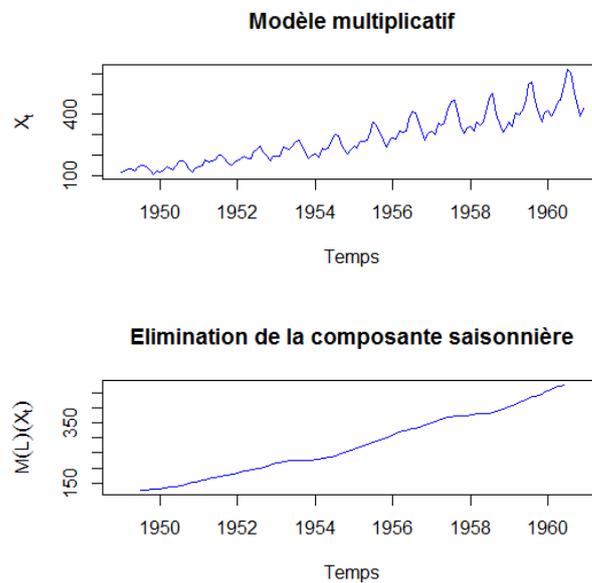


FIG. 1.4 – Représentation graphique d'un exemple d'élimination de la saisonnalité S_t par moyenne mobile

La série représenté dans la figure (1.4) en haut est celle du nombre mensuels de passagers aériens (en milliers), de janvier 1949 à décembre 1960. Cette série est disponible sous logiciel R (AirPassengers). Elle comporte une saisonnalité et une tendance au cours du temps. Le graphique du bas figure la même série après l'application d'une moyenne mobile d'ordre 12 qui élimine la composante saisonnière.

1.4 Calcule de prévision

1.4.1 Méthode de Buys-Ballot

La méthode de Buys-Ballot concerne les séries temporelles suivant un modèle additif et dont la tendance est linéaire. Nous supposons que la modèle est donné par :

$$X_t = a + bt + \gamma_1 S^1 + \gamma_2 S^2 + \dots + \gamma_p S^p + \varepsilon_t.$$

Cette méthode consiste à estimer les paramètres de ce modèle suivant le critère de moindres carrés, et permet ensuite d'effectuer des prévisions.

L'estimateur des paramètres $\theta = (a, b, \gamma_1, \gamma_2, \dots, \gamma_p)^t$ est donc la solution du système

$$\arg \min_{\theta} \sum_{t=1}^T \left[X_t - a - bt - \sum_{j=1}^p \gamma_j S^j \right].$$

Les formules des estimateurs sont donc :

$$\begin{cases} \hat{b} = \frac{12}{mN(N^2 - 1)} \left(\sum_{k=1}^N k \overline{X}_k - \frac{N(N+1)}{2} \overline{\overline{X}} \right), \\ \hat{a} = \overline{\overline{X}} - \hat{b} \frac{Nm + 1}{2}, \\ \hat{\gamma}_j = \overline{X}_j - \overline{\overline{X}} - \hat{b} \left(j - \frac{m+1}{2} \right), \end{cases}$$

avec les notations suivantes :

– N : le nombre d'année.

- m : le nombre de saisons pour chaque année.
- posons : $t = j + m(k - 1)$.
- \bar{X}_k : moyenne de X_t relatives à l'année k ($k = \overline{1, N}$).
- \bar{X}_j : moyenne de X_t relatives au saison j ($j = \overline{1, m}$).
- $\bar{\bar{X}} = \frac{1}{T} \sum_{t=1}^T X_t = \frac{1}{N} \sum_{k=1}^N \bar{X}_k = \frac{1}{m} \sum_{j=1}^m \bar{X}_j$: moyenne de la totalité des observations.

Dans ce cas, la prédiction d'une valeur futur X_{T+h} à l'horizon h est donnée par :

$$\hat{X}_T(h) = \hat{X}_{T+h} = \hat{a} + \hat{b}(T+h) + \sum_{j=1}^m \hat{\gamma}_j S^j.$$

Remarque 1.4.1

- Si les valeurs de X_t sont positives et le modèle est de la forme suivante :

$$X_t = (a + bt) (\gamma_1 S^1 + \gamma_2 S^2 + \dots + \gamma_p S^p) \varepsilon_t,$$

les logarithmes des observations vérifient les conditions précédentes.

- Lorsque la série suit le modèle multiplicatif et que la tendance est exponentielle, en utilise un changement de variable $Y_t = \ln(X_t)$, le modèle devient additif et la méthode de Buys-Ballot s'applique sur la nouvelle série Y_t .

1.4.2 Lissage exponentiel

Les méthodes de lissages exponentiels constituent un outil permettant de réaliser des prévisions à partir de l'observation d'une série temporelle.

Lissage exponentiel simple

Le lissage exponentiel simple permet d'effectuer des prévisions pour des séries temporelles dont la tendance est constante et sans saisonnalité. On suppose que l'on observe la série

temporelle sur les instant $(1, \dots, T)$, telle que

$$X_t = a + \varepsilon_t,$$

Définition 1.4.1

Une prévision de la série X_t à l'horizon h fournie par le lissage exponentiel simple est définie par :

$$\hat{X}_T(h) = \hat{X}_{T+h} = (1 - \alpha) \sum_{j=0}^{T-1} \alpha^j X_{T-j},$$

où $\alpha \in [0, 1]$ est la constante de lissage.

Remarque 1.4.2

- La méthode de lissage exponentiel simple prend en compte tous les observations passées mais en accordant d'importance aux observations les plus éloignées de l'instant T .
- Plus α est petit (ou encore α proche de 0), la prévision est souple, c-à-d fortement influencée par les observations les plus récentes
- Plus α est grand (ou encore α proche de 1), la prévision est rigide, c-à-d l'influence des observations anciennes est plus importante
- Si $\alpha = 0$, la prévision est égale à la dernière valeur observée et si $\alpha = 1$, alors toutes les prévisions sont identiques. En pratique, on prend $\alpha \in]0, 1[$ afin d'exclure ces deux cas.

Remarque 1.4.3

Pour le choix de α , on peut se servir de la remarque (1.4.2) ou de choisir α qui minimise la somme des carrés des erreurs de prévision

$$\arg \min_{\alpha} \sum_{t=1}^{T-h} \left(X_{t+h} - \hat{X}_{t+h} \right)^2.$$

Lissage exponentiel double

Le lissage exponentiel double généralise l'idée du lissage exponentiel simple au cas où la série peut être ajustée par une droite au voisinage de T comme suit

$$X_t = a(t - T) + b + \varepsilon_t.$$

On cherche dans ce cas une prévision à l'horizon h , $\hat{X}_T(h)$ de la forme :

$$\hat{X}_T(h) = \hat{a}_T h + \hat{b}_T,$$

où le couple (\hat{a}_T, \hat{b}_T) minimise la fonction :

$$\sum_{j=0}^{T-1} \alpha^j (X_{T-j} - aj - b)^2.$$

En notant la série lissée s_1 et la série doublement lissée s_2 , définies par :

$$s_1(t) = (1 - \alpha) \sum_{j=0}^{t-1} \alpha^j X_{t-j}, \text{ et } s_2(t) = (1 - \alpha) \sum_{j=0}^{t-1} \alpha^j s_1(t - j),$$

on déduit alors la définition suivante.

Définition 1.4.2

La prévision de la série X_t à l'horizon h , $\hat{X}_T(h)$, fournie par la méthode de lissage exponentiel double est donnée par :

$$\hat{X}_T(h) = \hat{a}_T h + \hat{b}_T,$$

où

$$\begin{cases} \hat{a}_T = \frac{(1 - \alpha)}{\alpha} (s_1(T) - s_2(T)), \\ \hat{b}_T = 2s_1(T) - s_2(T). \end{cases}$$

Chapitre 2

Modèles stationnaires

La modélisation du comportement aléatoire d'une série temporelle X_t devient beaucoup plus simple lorsque certains aspects de sa distribution ne changent pas quand t varie. On parle alors de modèles stationnaires. Dans ce chapitre, nous présentons tout d'abord les propriétés statistiques et la notion de stationnarité d'une série temporelle. Ensuite, nous introduisons la définition de la densité spectrale, ainsi que la propriété de la linéarité, la causalité et l'inversibilité des modèles.

Dans la suite de ce travail, on désigne par un modèle, le processus stochastique¹ qui modélise la série temporelle. Les éléments de cette dernière sont considérés comme des variables aléatoires.

2.1 Propriétés statistiques

Définition 2.1.1 (Moyenne)

La moyenne d'une série temporelle X_t est la fonction donnée par :

$$E(X_t) := \mu_t, \quad \forall t \in T.$$

¹Un processus stochastique est une famille de variable aléatoire X_t ($t \in T$) définie sur un espace probabilisé (Ω, A, P) .

Définition 2.1.2 (Variance)

La variance d'une série temporelle X_t est définie par :

$$\text{var}(X_t) := E(X_t^2) - E(X_t)^2 = \sigma^2.$$

Définition 2.1.3 (Fonction d'autocovariance)

La fonction d'autocovariance d'une série temporelle X_t , notée $\gamma_X(h)$, est définie par :

$$\gamma_X(h) := \text{cov}(X_t, X_{t\pm h}) = E(X_t X_{t\pm h}) - E(X_t)E(X_{t\pm h}), \forall h \in \mathbb{Z}.$$

Proposition 2.1.1

La fonction d'autocovariance vérifie les propriétés suivantes :

- $\text{cov}(X_t, X_s) = \gamma_X(|t - s|)$.
- $\gamma_X(0) = \text{var}(X_t) = \sigma^2 \geq 0$.
- $|\gamma_X(h)| \leq \gamma_X(0)$, pour tout $h \in \mathbb{Z}$, la fonction d'autocovariance est donc bornée.
- $\gamma_X(h) = \gamma_X(-h)$, pour tout $h \in \mathbb{Z}$, autrement dit la fonction d'autocovariance est paire.

Définition 2.1.4 (Fonction d'autocorrélation)

La fonction d'autocorrélation d'une série temporelle, notée $\rho_X(h)$, est définie par :

$$\rho_X(h) := \text{cor}(X_t, X_{t\pm h}) = \frac{\text{cov}(X_t, X_{t\pm h})}{\sqrt{\text{var}(X_t)}\sqrt{\text{var}(X_{t\pm h})}} = \frac{\gamma_X(h)}{\gamma_X(0)}, \forall h \in \mathbb{Z}.$$

Proposition 2.1.2

- $\rho_X(0) = 1$.
- $|\rho_X(h)| < 1$, pour tout $h \in \mathbb{Z}$.
- $\rho_X(h) = \rho_X(-h)$, pour tout $h \in \mathbb{Z}$.

Définition 2.1.5 (Matrice d'autocorrélation)

On appelle matrice d'autocorrélation de (X_t, \dots, X_{t+h-1}) pour $h \in \mathbb{N}^*$, la matrice

$$\mathcal{R}_h = \begin{pmatrix} 1 & \rho_X(1) & \rho_X(2) & \dots & \rho_X(h-1) \\ \rho_X(1) & 1 & \rho_X(1) & \dots & \rho_X(h-2) \\ \rho_X(2) & \rho_X(1) & 1 & \dots & \rho_X(h-3) \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \rho_X(h-1) & \rho_X(h-2) & \rho_X(h-3) & \dots & 1 \end{pmatrix} \quad (2.1)$$

\mathcal{R}_h est de rang h et de forme Toeplitz (symétrique avec égalité des termes diagonaux).

Remarque 2.1.1

Empiriquement pour un échantillon : X_1, X_2, \dots, X_T :

1 La moyenne empirique est donnée par :

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

2 La variance empirique est donnée par :

$$\hat{\gamma}_X(0) := \frac{1}{T-1} \sum_{t=1}^T (X_t - \bar{X})^2.$$

3 La fonction d'autocovariance empirique est donnée par :

$$\hat{\gamma}_X(h) := \frac{1}{T-h-1} \sum_{t=h+1}^T (X_t - \bar{X})(X_{t-h} - \bar{X}).$$

4 La fonction d'autocorrélation empirique est donnée donc par :

$$\hat{\rho}_X(h) := \frac{\hat{\gamma}_X(h)}{\hat{\gamma}_X(0)}.$$

2.2 Modèles stationnaires

Définition 2.2.1 (Stationnarité du second ordre)

Une série temporelle X_t est dite stationnaire du 2nd ordre (ou encore faiblement stationnaire) ou tout simplement stationnaire si :

$$\begin{cases} 1) E(X_t) = \mu, \forall t \in T, \\ 2) \text{var}(X_t) = \sigma^2 < \infty, \forall t \in T, \\ 3) \text{cov}(X_s, X_r) = \text{cov}(X_{s+t}, X_{r+t}), \forall r, s, t \in T. \end{cases}$$

Définition 2.2.2 (Stationnarité stricte)

Une série temporelle X_t est dite strictement (fortement) stationnaire si et seulement si la distribution conjointe de $(X_{t_1}, X_{t_2}, \dots, X_{t_n})^t$ et de $(X_{t_1+h}, X_{t_2+h}, \dots, X_{t_n+h})^t$ sont les mêmes, pour $t_1, t_2, \dots, t_n \in T$ et $h \in \mathbb{Z}$.

– La stationnarité faible est bien plus facile à étudier et à vérifier que la stationnarité stricte.

Exemple 2.2.1 (Bruit Blanc)

Un bruit blanc faible (ou un bruit blanc seulement) est une suite de variables aléatoires centré ($E(\varepsilon_t) = 0$), de variance finie ($\sigma_\varepsilon^2 < \infty$) et de fonction d'autocovariance :

$$\gamma_X(h) = \begin{cases} \sigma_\varepsilon^2 & \text{si } h = 0, \\ 0 & \text{si } h \neq 0. \end{cases}$$

On note $\varepsilon_t \sim BB(0, \sigma_\varepsilon^2)$, il est clair qu'un bruit blanc est stationnaire.

Remarque 2.2.1

On dit que ε_t est un bruit blanc fort s'il est un bruit blanc faible et que les variables aléatoires ε_t sont indépendantes et identiquement distribuées (iid).

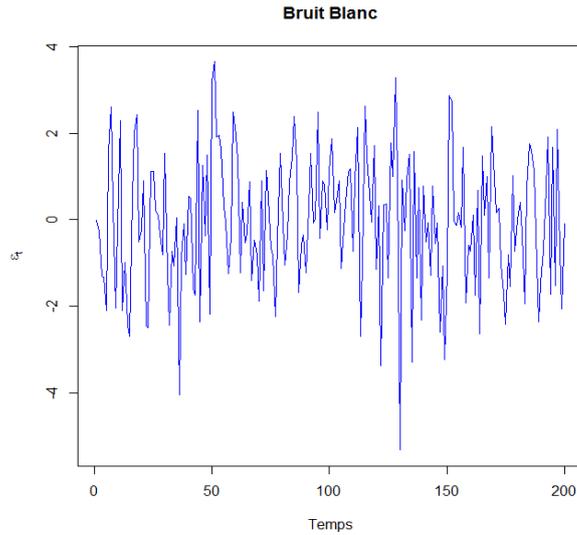


FIG. 2.1 – Représentation graphique d'un Bruit Blanc de taille 200 ($\mathcal{BB}(0, 2.5)$)

Exemple 2.2.2 (Marche aléatoire)

Le modèle définie par :

$$X_t = X_{t-1} + \varepsilon_t, \quad X_0 = 0,$$

est une marche aléatoire, telle que : $E(X_t) = 0$, $var(X_t) = t\sigma_\varepsilon^2$ et $\gamma_X(h) = (t-h)\sigma_\varepsilon^2$. Il est clair qu'une marche aléatoire n'est pas stationnaire.

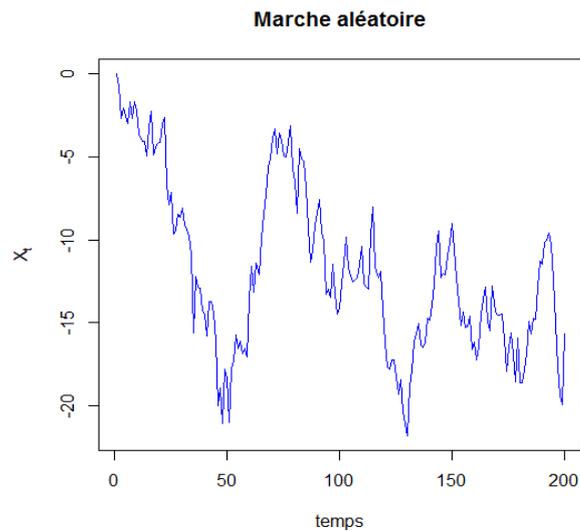


FIG. 2.2 – Représentation graphique d'une marche aléatoire

Proposition 2.2.1

1. Une série fortement stationnaire est toujours stationnaire, la réciproque est généralement fausse.
2. La somme de deux séries stationnaires n'est pas toujours stationnaire.

Exemple 2.2.3

Soit X_t une suite de variables aléatoires indépendantes, telle que :

$$X_t \sim \begin{cases} \exp(1) & \text{si } t \text{ est pair,} \\ \mathcal{N}(1, 1) & \text{si } t \text{ est impair.} \end{cases}$$

Alors X_t est stationnaire avec $\gamma_X(0) = 1$, $\gamma_X(h) = 0$, $\forall h \neq 0$. Cependant X_1 et X_2 n'ont pas la même distribution, donc X_t n'est pas strictement stationnaire.

Exemple 2.2.4

On pose $X_t = \varepsilon_t$ et $Y_t = (-1)^t X_t$ deux modèles stationnaires, mais $Z_t = X_t + Y_t$ n'est pas stationnaire.

Définition 2.2.3 (Stationnarité en différence)

Un modèle X_t est dit stationnaire en différence (DS : difference-stationary) s'il peut s'écrire sous l'une des formes suivantes :

$$(1 - L) X_t = \varepsilon_t.$$

$$(1 - L) X_t = \varepsilon_t + \beta, \quad \beta \in \mathbb{R}^*.$$

Définition 2.2.4 (Stationnarité en tendance)

Un modèle X_t sera dit stationnaire en tendance (TS : trend-stationary) s'il peut se mettre sous la forme :

$$X_t = f(t) + \varepsilon_t,$$

avec $f(t)$ est une fonction polynômiale du temps.

2.3 Densité spectrale

Dans cette section, on va introduire la notion de la densité spectrale qui contient la même information que la fonction d'autocovariance, mais elle a comme avantage les propriétés d'inversibilité et de périodicité.

Définition 2.3.1 (Densité spectrales)

Soit X_t une série temporelle stationnaire, de fonction d'autocovariance $\gamma_X(h)$. On appelle densité spectrale, la fonction $f_X(\lambda)$ définie par :

$$f_X(\lambda) := \frac{1}{2\pi} \sum_{h \in \mathbb{Z}} \gamma_X(h) \exp(i\lambda h), \quad \forall \lambda \in \mathbb{R}. \quad (2.2)$$

C'est donc la transformée de Fourier (discrète) de la fonction $\gamma_X(h)$.

Proposition 2.3.1

La densité spectrale est une fonction paire, continue, périodique de période 2π et positive, de plus on a :

$$f_X(\lambda) = \frac{1}{2\pi} \sum_{h \in \mathbb{Z}} \gamma_X(h) \cos(\lambda h), \quad \forall \lambda \in \mathbb{R}.$$

Proposition 2.3.2

Il est équivalent de connaître la fonction d'autocovariance ou la densité spectrale, d'un modèle stationnaire X_t . On a en effet :

$$\gamma_X(h) = \int_{-\pi}^{\pi} f_X(\lambda) \exp(-i\lambda h) d\lambda = \int_{-\pi}^{\pi} f_X(\lambda) \cos(\lambda h) d\lambda. \quad (2.3)$$

Proposition 2.3.3

1. La densité spectrale d'un bruit blanc ($\varepsilon_t \sim \mathcal{BB}(0, \sigma_\varepsilon^2)$) est constante en λ , telle que

$$f_\varepsilon(\lambda) = \frac{\sigma_\varepsilon^2}{2\pi}.$$

2. Tout modèle stationnaire de densité spectrale constante est un bruit blanc.

Proposition 2.3.4

Si Y_t est un modèle stationnaire de densité spectrale $f_Y(\lambda)$, alors la densité spectrale du modèle X_t , telle que :

$$\Phi(L) X_t = \Theta(L) Y_t,$$

est

$$f_X(\lambda) = \frac{|\Theta(\exp(i\lambda))|^2}{|\Phi(\exp(i\lambda))|^2} f_Y(\lambda).$$

Et si de plus $Y_t \sim \mathcal{BB}(0, \sigma_\varepsilon^2)$, alors :

$$f_X(\lambda) = \frac{|\Theta(\exp(i\lambda))|^2}{|\Phi(\exp(i\lambda))|^2} \frac{\sigma_\varepsilon^2}{2\pi}.$$

2.4 Modèles linéaires

Définition 2.4.1 (Modèle linéaire)

X_t est un modèle linéaire de moyenne c , s'il peut être écrit sous la forme :

$$X_t = c + \sum_{j=-\infty}^{+\infty} \psi_j \varepsilon_{t-j},$$

où $\varepsilon_t \sim \mathcal{BB}(0, \sigma_\varepsilon^2)$ et ψ_j est une suite absolument sommable, c-à-d $\sum_{j=-\infty}^{+\infty} |\psi_j| < +\infty$.

Définition 2.4.2 (Modèle causal)

X_t est un modèle causal s'il s'écrit en fonction de $(\varepsilon_k, k \leq t)$ seulement :

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

Définition 2.4.3 (Modèle inversible)

On dit qu'un modèle X_t est inversible, si ε_t s'écrit en fonction de $(X_k, k \leq t)$ seulement :

$$\varepsilon_t = \sum_{j=0}^{+\infty} \psi_j X_{t-j}.$$

Proposition 2.4.1

Soit Y_t un modèle stationnaire centré

1. Si $\sum_{j=-\infty}^{+\infty} |\phi_j| < \infty$ alors $X_t = \sum_{j=-\infty}^{+\infty} \phi_j Y_{t-j}$ est stationnaire centré de fonction d'autocovariance

$$\gamma_X(h) = \sum_{j=-\infty}^{+\infty} \sum_{k=-\infty}^{+\infty} \phi_j \phi_k \gamma_Y(h+k-j). \tag{2.4}$$

2. Si X_t est linéaire et si $Y_t \sim \mathcal{BB}(0, \sigma_\varepsilon^2)$ alors

$$\gamma_X(h) = \sum_{j=-\infty}^{+\infty} \phi_j \phi_{j+h} \sigma_\varepsilon^2.$$

2.5 Conditions de stationnarité et d'inversibilité d'un modèle sous l'écriture polynomiale

Soit un filtre linéaire d'ordre $(p+1)$

$$\Phi(L) = \sum_{i=0}^p \phi_i L^i = 1 + \phi_1 L + \phi_2 L^2 + \dots + \phi_p L^p, \tag{2.5}$$

où ϕ_i sont des nombres réels avec $\phi_0 = 1$. Le polynôme caractéristique associé à ce filtre est $\Phi(z) = 1 + z + \phi_2 z^2 + \dots + \phi_p z^p$. Ce polynôme possède p racines réelles ou complexes

ψ_j est une suite absolument sommable, c-à-d $\sum_{j=-\infty}^{+\infty} |\psi_j| < +\infty$. $z_j = \frac{1}{\lambda_j}$ ($j = \overline{1, p}$), telles que :

$$\Phi(z) = \prod_{j=1}^p (1 - \lambda_j z).$$

Un opérateur de la forme $1 - \lambda L$ est inversible si et seulement si $|\lambda| \neq 1$. De plus

$$(1 - \lambda L)^{-1} = \begin{cases} \sum_{k=0}^{\infty} (\lambda L)^k & \text{si } |\lambda| < 1, \\ -\sum_{k=0}^{\infty} (\lambda L)^{-k-1} & \text{si } |\lambda| > 1. \end{cases}$$

Lorsque toutes les racines du polynôme $\Phi(z)$ sont de module différent de 1, on a

$$\Phi(L)^{-1} = \prod_{j=1}^p (1 - \lambda_j L)^{-1}.$$

Proposition 2.5.1

Le filtre $\Phi(L)$ est inversible si et seulement si toutes les racines de Φ ont un module différent de 1. Dans ce cas

$$\Phi(L)^{-1} = \sum_{j=-\infty}^{+\infty} \psi_j L^j,$$

où $\sum_{j=-\infty}^{+\infty} |\psi_j| < +\infty$.

Proposition 2.5.2

Si X_t s'écrit sous forme :

$$\Phi(L) X_t = \Theta(L) \varepsilon_t + c,$$

où $\Phi(L)$ et $\Theta(L)$ sont des filtres linéaire de la forme (2.5), alors :

- i) X_t est stationnaire si et seulement si toutes les racines de Φ sont en module strictement supérieur à 1.
- ii) X_t est inversible si et seulement si toutes les racines de Θ sont en module strictement supérieur à 1.

Chapitre 3

Modèles ARMA

Dans ce chapitre, on considère certains modèles linéaires fréquemment utilisés pour une série temporelle. Nous commençons par deux modèles simples : le modèle autorégressif (*AR*) et le modèle à moyenne mobile (*MA*). On continue avec le modèle *ARMA* qui permet de combiner ces deux modèles. Ce chapitre traite aussi la méthode de prévision des séries temporelles, telle que la méthode de Box et Jenkins. Les références les plus importantes dans ce domaine sont les ouvrages de Box et al (2016) [8], Box et Jenkins (1976) [7], Shumway & Stoffer (2017) [17] et Aragon (2011) [2]

3.1 Modèle autorégressif

Définition 3.1.1 (Modèle autorégressif)

X_t est un modèle autorégressif d'ordre p , noté $AR(p)$, si :

$$X_t = c + \sum_{j=1}^p \alpha_j X_{t-j} + \varepsilon_t, \quad (3.1)$$

où $\alpha_j \in \mathbb{R}$, avec $\alpha_p \neq 0$, $\varepsilon_t \sim \mathcal{BB}(0, \sigma_\varepsilon^2)$ et c une terme constante. La relation (3.1) est équivalente à l'écriture suivante :

$$\Phi(L) X_t = c + \varepsilon_t, \quad (3.2)$$

avec

$$\Phi(L) = 1 - \alpha_1 L - \alpha_2 L^2 - \dots - \alpha_p L^p.$$

Remarque 3.1.1

1) Un modèle $AR(p)$ est stationnaire si toutes les racines du polynôme caractéristique $\Phi(z)$ sont de module supérieur strictement à 1. Dans ce cas, X_t est de moyenne $\mu = c / (1 - \alpha_1 - \alpha_2 - \dots - \alpha_p)$.

2) Sous l'hypothèse de stationnarité de X_t , il est possible de se ramener à un modèle centré par une simple translation $Y_t = X_t - \mu$, tel que :

$$E(Y_t) = 0, \quad \gamma_Y(h) = \gamma_X(h) \quad \text{et} \quad \rho_Y(h) = \rho_X(h).$$

3) Un modèle $AR(p)$ est toujours inversible.

3.1.1 Autocovariance et autocorrélation d'un $AR(p)$

On considère le cas d'un modèle $AR(p)$ stationnaire et centré :

$$X_t = \alpha_1 X_{t-1} + \alpha_2 X_{t-2} + \dots + \alpha_p X_{t-p} + \varepsilon_t \tag{3.3}$$

La fonction d'autocovariance est :

$$\gamma_X(h) = E(X_t X_{t-h}), \quad \forall h \geq 0$$

En multipliant la relation (3.3) par X_{t-h} , nous obtenons

$$X_t X_{t-h} = \alpha_1 X_{t-1} X_{t-h} + \alpha_2 X_{t-2} X_{t-h} + \dots + \alpha_p X_{t-p} X_{t-h} + \varepsilon_t X_{t-h}.$$

Alors

$$\gamma_X(h) = \alpha_1 \gamma_X(h-1) + \alpha_2 \gamma_X(h-2) + \dots + \alpha_p \gamma_X(h-p) + E(\varepsilon_t X_{t-h}), \quad \forall h \geq 0.$$

Et comme

$$E(\varepsilon_t X_{t-h}) = \begin{cases} \sigma_\varepsilon^2 & \text{si } h = 0, \\ 0 & \text{si } h \geq 1. \end{cases}$$

On obtient donc

$$\gamma_X(h) = \sum_{j=1}^p \alpha_j \gamma_X(h-j), \quad \forall h \geq 1. \quad (3.4)$$

Pour $h = 0$, on procède de la même façon

$$\begin{aligned} \gamma_X(0) &= E(X_t^2) = E(\alpha_1 X_{t-1} X_t + \alpha_2 X_{t-2} X_t + \dots + \alpha_p X_{t-p} X_t + \varepsilon_t X_t) \\ &= \sum_{j=1}^p \alpha_j \gamma_X(j) + \sigma_\varepsilon^2. \end{aligned} \quad (3.5)$$

On divise les deux membres de (3.5) par $\gamma_X(0)$, on obtient la variance de X_t

$$\gamma_X(0) = \sigma_X^2 = \frac{\sigma_\varepsilon^2}{1 - \sum_{j=1}^p \alpha_j \rho_X(j)}.$$

En divisant (3.4) par $\gamma_X(0)$, nous obtenons la fonction d'autocorrélation

$$\rho_X(h) = \frac{\gamma_X(h)}{\gamma_X(0)} = \sum_{j=1}^p \alpha_j \frac{\gamma_X(h-1)}{\gamma_X(0)} = \sum_{j=1}^p \alpha_j \rho_X(h-j), \quad \forall h \geq 1 \quad (3.6)$$

Les deux équations (3.4) et (3.6) s'appellent le système de **Yule-Walker**.

Estimation des paramètres

Tout d'abord, pour obtenir $\rho_X(1), \rho_X(2), \dots, \rho_X(p)$, il suffit résoudre le système d'équation suivant :

$$\begin{array}{l}
 h = 1 \\
 h = 2 \\
 \vdots \\
 h = p - 1 \\
 h = p
 \end{array}
 \left\{ \begin{array}{l}
 \rho_X(1) = \alpha_1 \rho_X(0) + \alpha_2 \rho_X(-1) + \cdots + \alpha_p \rho_X(1-p) \\
 \rho_X(2) = \alpha_1 \rho_X(1) + \alpha_2 \rho_X(0) + \cdots + \alpha_p \rho_X(2-p) \\
 \vdots \\
 \rho_X(p-1) = \alpha_1 \rho_X(p-2) + \alpha_2 \rho_X(p-3) + \cdots + \alpha_p \rho_X(-1) \\
 \rho_X(p) = \alpha_1 \rho_X(p-1) + \alpha_2 \rho_X(p-2) + \cdots + \alpha_p \rho_X(0)
 \end{array} \right.$$

Et puisque la fonction $\rho_X(h)$ est paire et $\rho_X(0) = 1$, alors :

$$\left\{ \begin{array}{l}
 \rho_X(1) = \alpha_1 + \alpha_2 \rho_X(1) + \cdots + \alpha_p \rho_X(p-1) \\
 \rho_X(2) = \alpha_1 \rho_X(1) + \alpha_2 + \cdots + \alpha_p \rho_X(p-2) \\
 \vdots \\
 \rho_X(p-1) = \alpha_1 \rho_X(p-2) + \alpha_2 \rho_X(p-3) + \cdots + \alpha_p \rho_X(1) \\
 \rho_X(p) = \alpha_1 \rho_X(p-1) + \alpha_2 \rho_X(p-2) + \cdots + \alpha_p
 \end{array} \right.$$

Par conséquent, les équations de Yule-Walker peuvent s'écrire sous la forme matricielle suivante :

$$\rho = \mathcal{R}_p \alpha, \tag{3.7}$$

où $\rho = (\rho_X(1), \rho_X(2), \dots, \rho_X(p))^t$, $\alpha = (\alpha_1, \alpha_2, \dots, \alpha_p)^t$ et \mathcal{R}_p est la matrice des auto-corrélations définie par (2.1), elle est du rang p .

D'après (3.7), on déduit que l'inversion de la matrice \mathcal{R}_p nous permet de calculer les α_j en fonction de $\rho_X(j)$ seulement

$$\alpha = \mathcal{R}_p^{-1} \rho,$$

ce qui nous permet d'obtenir une estimation du vecteur des paramètres α à l'aide de $\hat{\rho}_X(j)$

$$\hat{\alpha} = \hat{\mathcal{R}}_p^{-1} \hat{\rho},$$

3.1.2 Autocorrélation partielle

Les fonctions γ_X et ρ_X n'interviennent que les variables X_t et $X_{t\pm h}$, nous allons introduire ici la fonction d'autocorrélation partielle, notée ϕ_h , faisant intervenir de plus les variables intermédiaires $(X_{t\pm 1}, X_{t\pm 2}, \dots, X_{t\pm(h-1)})$. Pour un modèle $AR(p)$, le calcul des autocorrélations partielles se fait par l'algorithme de Durbin, comme suit :

$$\begin{aligned}\phi_0 &= \rho_X(0) = 1, \\ \phi_1 &= \rho_X(1), \\ \phi_h &= \frac{\det \mathcal{R}_h^*}{\det \mathcal{R}_h}, h \geq 2,\end{aligned}$$

où \mathcal{R}_h^* est la matrice \mathcal{R}_h en remplaçant la dernière colonne par $(\rho_X(1), \rho_X(2), \dots, \rho_X(h))^t$

Proposition 3.1.1

Pour un modèle $AR(p)$ les autocorrélations partielles sont nulles au-delà du rang p : $\phi_h = 0, \forall h > p$. Cette propriété nous permet d'estimer l'ordre p du modèle $AR(p)$.

Exemple 3.1.1 ($X_t \sim AR(1)$)

Considérons un modèle $AR(1)$, avec $|\alpha| < 1$: $X_t = \alpha X_{t-1} + \varepsilon_t$, alors la fonction d'autocovariance est :

$$\begin{aligned}\gamma_X(h) &= E(X_t X_{t-h}) = \alpha E(X_{t-1} X_{t-h}) = \alpha \gamma_X(h-1) = \alpha^2 \gamma_X(h-2) \\ &\vdots \\ &= \alpha^h \gamma_X(0), \forall h \geq 1,\end{aligned}$$

par conséquent la fonction d'autocorrélation est :

$$\rho_X(h) = \alpha \rho_X(h-1) = \dots = \alpha^h, \forall h \geq 1, \quad (3.8)$$

et

$$\gamma_X(0) = \frac{\sigma_\varepsilon^2}{1 - \alpha \rho_X(1)}.$$

Nous représentons dans la figure 3.1 (en haut) les trajectoires de deux exemples d'un modèle $AR(1)$ selon le signe du coefficient α ($\alpha = 0.6$ et $\alpha = -0.6$). Les modèles considérés sont centrés, stationnaires et de taille $n = 500$ avec $\sigma_\varepsilon^2 = 1$. Cette figure illustre le graphique de la fonction d'autocorrélation qui s'appelle autocorrélogramme (au milieu) et celui de la fonction d'autocorrélation partielles qui s'appelle autocorrélogramme partiel (en bas).

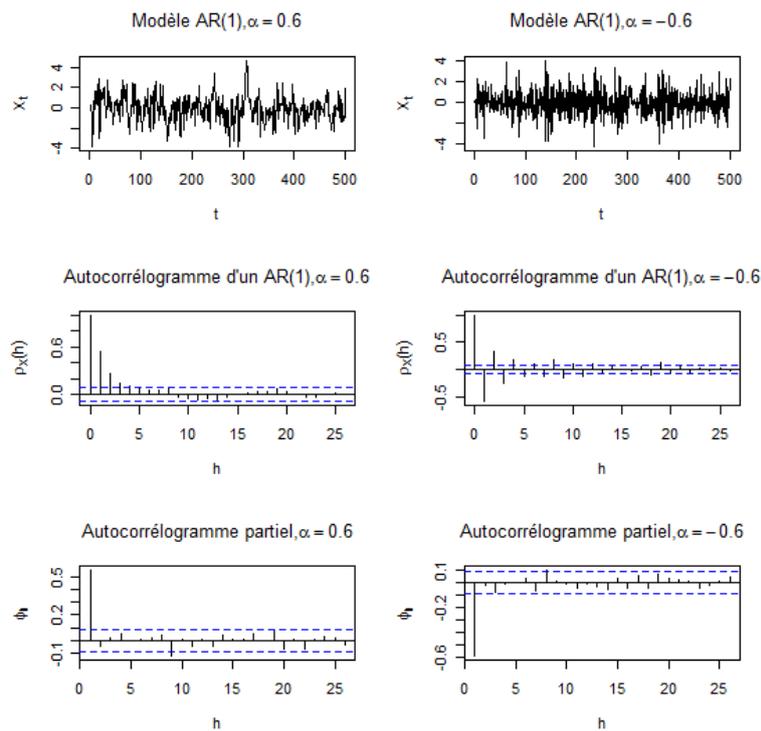


FIG. 3.1 – Représentation graphique d'une trajectoire (en haut), autocorrélogramme (au milieu) et autocorrélogramme partiel (en bas) d'un modèle $AR(1)$

L'autocorrélogramme montre une décroissance exponentielle vers 0, pour un coefficient positif ($\alpha = 0.6$) et une alternance en signe pour un coefficient négatif ($\alpha = -0.6$). Cela confirme le résultat théorique obtenue en (3.8).

A partir les autocorrélogrammes partiels pour un $AR(1)$, on observe un pic significatif pour le premier retard ($h = 1$) : positif si $\alpha = 0.6 > 0$ et négatif si $\alpha = -0.6 < 0$, les autres autocorrélations partielles sont nulles pour des retards $h > 1$.

Exemple 3.1.2 ($X_t \sim AR(2)$)

Considérons maintenant un modèle $AR(2)$, centré et stationnaire : $X_t = \alpha_1 X_{t-1} + \alpha_2 X_{t-2} + \varepsilon_t$. On a d'après le système de Yule-Walker

$$\begin{cases} \rho_X(1) = \alpha_1 + \alpha_2 \rho_X(1), \\ \rho_X(2) = \alpha_1 \rho_X(1) + \alpha_2. \end{cases}$$

Il résulte

$$\rho_X(1) = \frac{\alpha_1}{1 - \alpha_2}, \quad \rho_X(2) = \frac{\alpha_1^2 + \alpha_2^2 + \alpha_2}{1 - \alpha_2} \quad \text{et} \quad \gamma_X(0) = \frac{(1 - \alpha_2) \sigma_\varepsilon^2}{1 - \alpha_2 - \alpha_1^2 - \alpha_2 \alpha_1^2 - \alpha_2^2 - \alpha_2^3},$$

d'où

$$\gamma_X(1) = \frac{\alpha_1 \gamma_X(0)}{1 - \alpha_2} \quad \text{et} \quad \gamma_X(2) = \alpha_1 \gamma_X(1) + \alpha_2 \gamma_X(0).$$

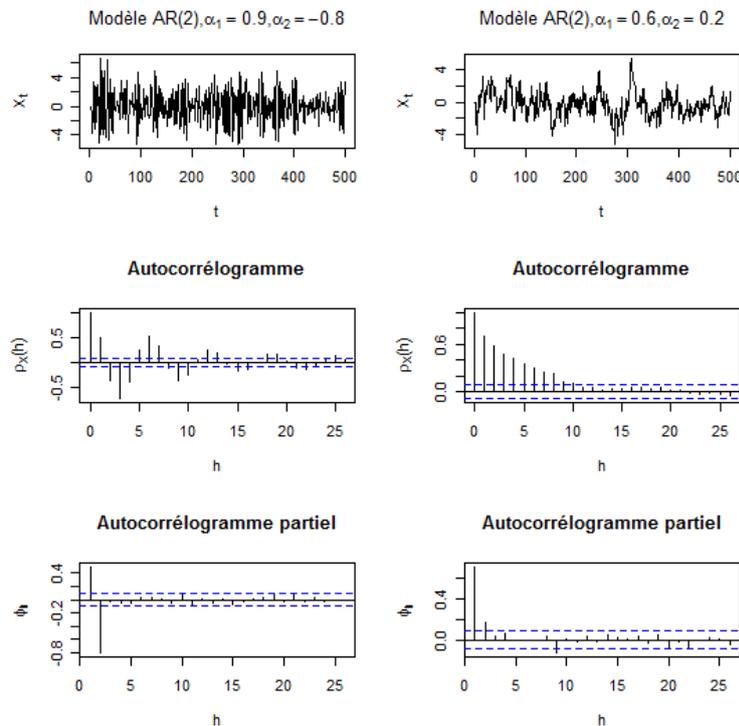


FIG. 3.2 – Représentation graphique d'une trajectoire, autocorrélogramme et autocorrélogramme partiel d'un $AR(2)$

La figure (3.2) représente les trajectoires de deux modèles $AR(2)$ stationnaires, de taille $n = 500$ avec $\sigma_\varepsilon^2 = 1$ (les racines du polynôme Φ associé au premier modèle sont des

nombre complexes, et elles sont des nombre réelles pour le second modèle).

Les autocorrélogrammes (au milieu) montre une décroissance exponentielle si les racines sont réelles, et une forme sinusoïdale si les racines sont complexes. On constate d'après le graphique des autocorrélations partielles (en bas) qu'il y on a deux pics significatifs pour le premier et le second retard ($h = 1, 2$), les autres coefficients sont nuls pour des retard $h > 2$.

3.1.3 Modèle autorégressif d'ordre infini

Définition 3.1.2

On dit que X_t est un modèle $AR(\infty)$, si

$$X_t = \sum_{j=1}^{+\infty} \alpha_j X_{t-j} + \varepsilon_t.$$

3.2 Modèle à moyenne mobile

Définition 3.2.1 (Modèle à moyenne mobile)

X_t est un modèle à moyenne mobile d'ordre q , noté $MA(q)$, si :

$$\begin{aligned} X_t &= \mu + \varepsilon_t + \beta_1 \varepsilon_{t-1} + \beta_2 \varepsilon_{t-2} + \dots + \beta_q \varepsilon_{t-q} \\ &= \mu + \varepsilon_t + \sum_{j=1}^q \beta_j \varepsilon_{t-j}, \quad \beta_j \in \mathbb{R}, \quad \beta_q \neq 0 \quad \text{et} \quad \varepsilon_t \sim \mathcal{BB}(0, \sigma_\varepsilon^2). \end{aligned}$$

On peut noter de façon équivalente

$$X_t = \mu + \Theta(L) \varepsilon_t,$$

où $\Theta(L) = 1 + \beta_1 L + \beta_2 L^2 + \dots + \beta_q L^q$.

Remarque 3.2.1

- 1) Un modèle $MA(q)$ est toujours stationnaire et causal, il est de moyenne μ .
- 2) Il est possible d'inverser le polynôme $\Theta(L)$ dans le cas où ses racines sont de module différent de 1 (mais un modèle $MA(q)$ est inversible si et seulement si ces racines sont de module strictement supérieur à 1).

3.2.1 Autocovariance et Autocorrélation d'un $MA(q)$

Nous considérons dans ce qui suit que le modèle $MA(q)$ est centré.

– La fonction d'autocovariance d'un $MA(q)$ est :

$$\begin{aligned} \gamma_X(h) &= E(X_t X_{t-h}) \\ &= E((\varepsilon_t + \beta_1 \varepsilon_{t-1} + \dots + \beta_q \varepsilon_{t-q})(\varepsilon_{t-h} + \beta_1 \varepsilon_{t-h-1} + \dots + \beta_q \varepsilon_{t-h-q})), \quad h \geq 0 \\ &= \begin{cases} (1 + \beta_1^2 + \beta_2^2 + \dots + \beta_q^2) \sigma_\varepsilon^2 & \text{si } h = 0, \\ (\beta_h + \beta_{h+1} \beta_1 + \dots + \beta_q \beta_{h+q}) \sigma_\varepsilon^2 & \text{si } 1 \leq h \leq q, \\ 0 & \text{si } h > q. \end{cases} \end{aligned}$$

– La fonction autocorrélation :

$$\rho_X(h) = \frac{\gamma_X(h)}{\gamma_X(0)} = \begin{cases} 1 & \text{si } h = 0, \\ \frac{\beta_h + \beta_{h+1} \beta_1 + \dots + \beta_q \beta_{h+q}}{1 + \beta_1^2 + \beta_2^2 + \dots + \beta_q^2} & \text{si } 1 \leq h \leq q, \\ 0 & \text{si } h > q. \end{cases}$$

Remarque 3.2.2 La fonction d'autocorrélation d'un $MA(q)$ s'annule au-delà du rang q ($h > q$).

Exemple 3.2.1 ($X_t \sim MA(1)$)

Considère un modèle $MA(1)$: $X_t = \varepsilon_t + \beta \varepsilon_{t-1}$, $\varepsilon_t \sim BB(0, \sigma_\varepsilon^2)$. Alors $E(X_t) = 0$,

$$\gamma_X(h) = \begin{cases} (1 + \beta^2) \sigma_\varepsilon^2 & \text{si } h = 0, \\ \beta \sigma_\varepsilon^2 & \text{si } h = 1, \\ 0 & \text{si } h > 1, \end{cases}$$

et la fonction d'autocorrélation

$$\rho_X(h) = \frac{\gamma_X(h)}{\gamma_X(0)} = \begin{cases} 1 & \text{si } h = 0, \\ \frac{\beta}{1 + \beta^2} & \text{si } h = 1, \\ 0 & \text{si } h > 1. \end{cases}$$

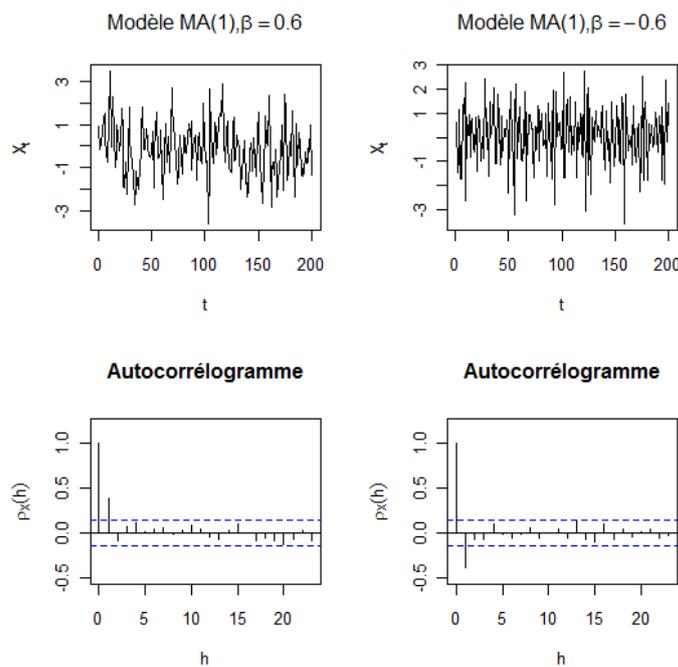


FIG. 3.3 – Représentation graphique d’une trajectoire (en haut), autocorrélogramme (en bas) d’un modèle $MA(1)$

La figure (3.3) montre la trajectoire (en haut) de deux exemples d’un $MA(1)$, de taille $n = 200$ avec $\sigma_\varepsilon^2 = 1$, selon le signe du coefficient β ($\beta = 0.6$ et $\beta = -0.6$). L’autocorrélogramme (en bas) montre un pic significatif pour le premier retard ($h = 1$) : positif si $\beta > 0$ et négatif si $\beta < 0$. Les autres autocorrélations sont nulles pour des retards $h > 1$.

3.2.2 Modèle à moyenne mobile d'ordre infini

Définition 3.2.2

Un modèle X_t est dit un $MA(\infty)$, s'il peut s'écrire sous la forme :

$$X_t = \sum_{j=-\infty}^{+\infty} \beta_j \varepsilon_{t-j}, \quad \forall t \in \mathbb{Z}, \quad \varepsilon_t \sim \mathcal{BB}(0, \sigma_\varepsilon^2).$$

On dit alors que X_t est un modèle linéaire. Si de plus $\beta_j = 0, \forall j < 0$

$$X_t = \sum_{j=0}^{+\infty} \beta_j \varepsilon_{t-j},$$

on dit que X_t est causal.

Proposition 3.2.1 Pour un $MA(\infty)$, on a les propriétés suivantes :

$$1) E(X_t) = 0 \quad 2) \text{var}(X_t) = \sigma_\varepsilon^2 \sum_{j=-\infty}^{+\infty} \beta_j^2. \quad 3) \gamma_X(h) = \sigma_\varepsilon^2 \sum_{j=-\infty}^{+\infty} \beta_j \beta_{j+|h|}$$

3.2.3 Théorème de Wold

Le théorème de Wold est fondamental pour l'analyse de séries temporelles stationnaires.

Théorème 3.2.1 (Wold) [10]

Tout processus X_t stationnaire peut se mettre sous forme d'un modèle $MA(\infty)$:

$$X_t = \kappa_t + \sum_{i=0}^{\infty} \psi_i \varepsilon_{t-i}, \tag{3.9}$$

ψ_i : les paramètres sont des réels, tel que $\Psi_0 = 1$ et $\sum_{i=0}^{\infty} \psi_i^2 < \infty$.

ε_t : Bruit Blanc.

κ_t : est une composante linéaire telle que $\text{cov}(\kappa_t, \varepsilon_{t-i}) = 0, \forall t \in \mathbb{Z}$.

Exemple 3.2.2

Soit un modèle stationnaire :

$$\begin{aligned}
 X_t &= \phi X_{t-1} + \varepsilon_t, \text{ où } |\phi| < 1 \text{ et } \varepsilon_t \sim \mathcal{BB}(0, 1), \\
 &= \phi(\phi X_{t-2} + \varepsilon_{t-1}) + \varepsilon_t \\
 &= \phi^2 X_{t-2} + \phi \varepsilon_{t-1} + \varepsilon_t \\
 &= \vdots \\
 &= \sum_{i=0}^{+\infty} \phi^i \varepsilon_{t-i}
 \end{aligned}$$

Alors X_t a une représentation MA(∞), avec $\psi_i = \phi^i$. En utilisant le résultat sur la somme géométrique infinie, on note que :

$$\sum_{i=0}^{\infty} \psi_i^2 = \sum_{i=0}^{\infty} (\phi^i)^2 = \sum_{i=0}^{\infty} (\phi^2)^i = \frac{1}{1 - \phi^2} < \infty.$$

3.3 Modèle ARMA

La classe des modèles ARMA généralise les modèles AR et les modèles à moyenne mobile MA

Définition 3.3.1 (Modèle Autorégressif et moyenne mobile)

Un modèle est dit autorégressif et moyenne mobile d'ordre p et q , noté ARMA(p, q), si :

$$X_t - \sum_{j=1}^p \alpha_j X_{t-j} = c + \varepsilon_t + \sum_{k=1}^q \beta_k \varepsilon_{t-k}, \tag{3.10}$$

où $\varepsilon_t \sim \mathcal{BB}(0, \sigma_\varepsilon^2)$, $\alpha_p \neq 0$, $\beta_q \neq 0$, c est une constante.

L'équation (3.10) est équivalente à l'écriture :

$$\Phi(L) X_t = c + \Theta(L) \varepsilon_t \tag{3.11}$$

Où : $\Phi(L) = 1 - \alpha_1 L - \alpha_2 L^2 - \dots - \alpha_p L^p$ et $\Theta_q(L) = 1 + \beta_1 L + \beta_2 L^2 + \dots + \beta_q L^q$. Les polynômes Φ et Θ n'ont pas de racines communes. On considère que Φ et Θ ont tous les racines de module strictement supérieur à 1,

Remarque 3.3.1

- 1) Les conditions de stationnarité d'un processus ARMA sont déterminées par les racines du polynôme associé à sa composante AR, alors que son inversibilité dépend des racines du polynôme associé à sa composante MA.
- 2) X_t est un ARMA(p, q) de moyenne μ , si $Y_t = X_t - \mu$ est un ARMA(p, q) centré. Dans ce cas $\mu = c / (1 - \alpha_1 - \alpha_2 - \dots - \alpha_p)$
- 3) Un modèle AR(p) (respectivement MA(q)) est un modèle ARMA($p, 0$) (respectivement ARMA($0, q$)).
- 4) Si $q = 0, p = 0$, le modèle théorique ARMA($0, 0$) est l'unique processus qui admet simultanément une représentation AR pure et une représentation MA pure.

Propriété 3.3.1

Si X_t est un ARMA(p, q) stationnaire et centré, alors :

1. X_t admet la représentation MA(∞) : $X_t = \frac{\Theta(L)}{\Phi(L)} \varepsilon_t = \sum_{i=0}^{\infty} \beta_i \varepsilon_{t-i}$,
2. X_t admet la représentation AR(∞) : $\varepsilon_t = \frac{\Phi(L)}{\Theta(L)} X_t = \sum_{i=0}^{\infty} \psi_i X_{t-i}$.

3.3.1 Fonction génératrice d'autocovariance

Définition 3.3.2 (Fonction génératrice d'autocovariance)

Pour un modèle X_t stationnaire la fonction génératrice d'autocovariance est donnée par :

$$G(z) = \sum_{h=-\infty}^{+\infty} \gamma_X(h) z^h = \sigma_\varepsilon^2 \Psi(z) \Psi(z^{-1}),$$

où

$$\Psi(z) = \frac{\Theta(z)}{\Phi(z)}.$$

Cette fonction permet de calculer les $\gamma_X(h)$, en identifiant les coefficients de z^h .

Exemple 3.3.1 ($X_t \sim ARMA(0, 2)$)

Considérons un modèle ARMA(0, 2) : $X_t = \varepsilon_t + \beta_1\varepsilon_{t-1} + \beta_2\varepsilon_{t-2}$. On a $\Theta(L) = 1 + \beta_1L + \beta_2L^2$, alors

$$\begin{aligned} G(z) &= \sigma_\varepsilon^2 \Psi(z) \Psi(z^{-1}) \\ &= \sigma_\varepsilon^2 \Theta(z) \Theta(z^{-1}) \\ &= \sigma_\varepsilon^2 (1 + \beta_1z + \beta_2z^2) (1 + \beta_1z^{-1} + \beta_2z^{-2}) \\ &= \sigma_\varepsilon^2 [(1 + \beta_1^2 + \beta_2^2) + (\beta_1 + \beta_1\beta_2)(z + z^{-1}) + \beta_2(z^2 + z^{-2})]. \end{aligned}$$

Par conséquent

$$\gamma_X(h) = \begin{cases} (1 + \beta_1^2 + \beta_2^2) \sigma_\varepsilon^2 & \text{si } h = 0, \\ (\beta_1 + \beta_1\beta_2) \sigma_\varepsilon^2 & \text{si } h = 1, \\ \beta_2 \sigma_\varepsilon^2 & \text{si } h = 2, \\ 0 & \text{si } h > 2. \end{cases}$$

3.4 Modèles ARIMA et SARIMA

Les modèles ARIMA et SARIMA sont des modèles non stationnaires qui présentent des tendances polynomiales de degré d (avec $d \geq 1$) et/ou des variations saisonnières de période s .

Définition 3.4.1 (Modèle Autorégressif et moyenne mobile intégré)

Un modèle X_t est dit autorégressif et moyenne mobile intégré d'ordre p , d et q , noté ARIMA(p, q), si le modèle $Y_t = \Delta^d X_t = (1 - L)^d X_t$ est un modèle ARMA(p, q) de moyenne nulle. X_t admet donc la représentation suivant :

$$\Phi(L) \Delta^d X_t = \Theta(L) \varepsilon_t,$$

où

$$\begin{cases} \Phi(L) = 1 - \alpha_1 L - \alpha_2 L^2 - \dots - \alpha_p L^p, & \alpha_p \neq 0, \\ \Theta_q(L) = 1 + \beta_1 L + \beta_2 L^2 + \dots + \beta_q L^q, & \beta_q \neq 0. \end{cases}$$

Définition 3.4.2 (Modèle SARIMA)

X_t est un modèle SARIMA(p, d, q)(P, Q, D) $_s$ si le modèle différentié :

$$Y_t = \Delta^d \Delta_s^D X_t = (1 - L)^d (1 - L^s)^D X_t,$$

est un ARMA(p, q), tel que :

$$\Delta^d \Delta_s^D \Phi(L) \Phi_s(L^s) X_t = \Theta(L) \Theta_s(L^s) \varepsilon_t,$$

Φ_s et Θ_s sont des polynômes de degrés P et Q respectivement

$$\begin{cases} \Phi_s(z) = 1 - \alpha_{s,1} z - \alpha_{s,2} z^2 - \dots - \alpha_{s,P} z^P, & \alpha_{s,P} \neq 0, \\ \Theta_q(L) = 1 + \beta_{s,1} L + \beta_{s,2} L^2 + \dots + \beta_{s,Q} L^Q, & \beta_{s,Q} \neq 0. \end{cases}$$

3.5 Prédiction par la méthode de Box-Jenkins

Box et Jenkins (1976) [7] ont promu une méthodologie consistant à modéliser les séries temporelles univariées au moyen des processus ARMA. Ces processus sont parcimonieux et constituent une bonne approximation de processus plus généraux pourvu que l'on se restreigne au cadre linéaire. Les modèles ARMA donnent souvent de bon résultats en prédiction. La figure (3.4) représente l'organigramme de cette méthodologie.

3.5.1 Identification

La stationnarité de la série X_t est d'abord testée soit graphiquement, ou théoriquement par les tests de racines unitaire (ou de non stationnarité), les plus connues sont : Test de Dickey-Fuller simple, Test de Dickey-Fuller augmenté, Test de Philippe et Perron, ...

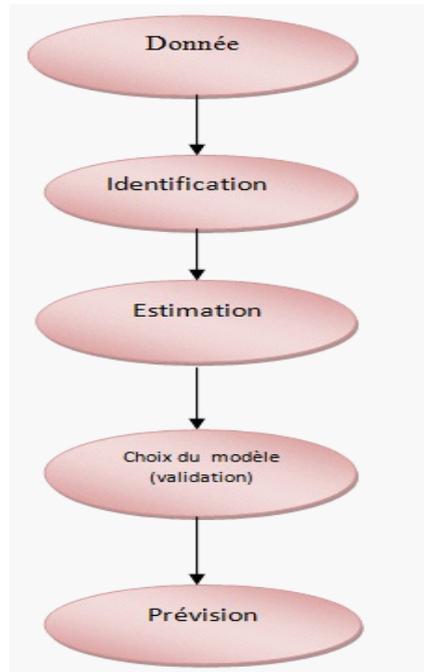


FIG. 3.4 – Organigramme de la méthode de Box & Jenkins

Si la série n'est pas stationnaire, il convient de la transformer par différenciation ΔX_t , $\Delta^2 X_t, \dots$, pour obtenir une série stationnaire.

L'étape d'identification d'un processus *ARMA* (choix entre *AR*, *MA* et *ARMA*, et choix de p et q) de Box et Jenkins est basée sur la comparaison de caractéristiques théoriques des modèles *ARMA* à leurs équivalents empiriques.

- 1) Pour un modèle stationnaire satisfaisant une représentation $AR(p)$, les autocorrélations partielles sont identiquement nulle au-delà de l'ordre p .
- 2) Pour un modèle satisfaisant une représentation $MA(q)$, les autocorrélations s'annulent à partir de q .
- 3) Pour déterminer les ordres p et q d'un modèle *ARMA*. On peut baser sur les autocorrélations et les autocorrélations partielles. Pour obtenir l'ordre de la composante *MA*, il faut identifier l'autocorrélation significative dont l'ordre est le plus élevé ; pour la composante *AR*, il faut identifier l'autocorrélation partielle significative dont l'ordre est le plus élevé.

Remarque 3.5.1

1. Les entiers d et D pour un modèle SARIMA sont choisis de sorte que la série différenciée : $(1 - L)^d(1 - L^s)^D$ soit stationnaire.
2. On utilise les autocorrélations et les autocorrélations partielles pour les valeurs multiples de s pour le choix des ordres P et Q d'un modèle SARIMA.

3.5.2 Estimation des paramètres

Etant donné le modèle X_t admettant une représentation $ARMA(p, q)$, donnée par (3.10). L'estimation des paramètres α_j ($j = \overline{1, p}$), β_k ($k = \overline{1, q}$) et σ_ε^2 du (des) modèle (s) sélectionné (s) peut se réaliser par les méthodes suivantes.

- 1) Méthode du maximum de vraisemblance.
- 2) Dans le cas d'un $AR(p)$, on utilise les équations de Yule-Walker.

3.5.3 Validation

Après l'étape l'identification, si on dispose de plusieurs modèles, il est nécessaire d'effectuer des tests sur les paramètres et sur les résidus.

1. Test de significativité des coefficients

On vérifie tout d'abord que les racines des polynômes AR et MA ne sont pas égales à 1. On teste la significativité des coefficients du modèle $ARMA$ par le test de Student.

2. Test sur le bruit blanc

Pour que les modèles obtenues soient valides, il convient de vérifier que les résidus estimés, notés $\hat{\varepsilon}_t$ suivent un bruit blanc, non autocorrélé et de même variance $\hat{\sigma}_\varepsilon^2$ et qu'ils suivent une loi normale. Si ces hypothèses ne sont pas rejetées, on peut alors mener des tests sur les paramètres.

a) Test sur la moyenne des résidus

Pour vérifier que les résidus estimés sont centrés, comme il se doit, il s'agit de confronter les hypothèses

$$\begin{cases} H_0 : E(\varepsilon_t) = 0, \\ H_1 : E(\varepsilon_t) \neq 0. \end{cases}$$

Une statistique de test est définie par :

$$\zeta = \frac{\bar{\hat{\varepsilon}}}{S_\varepsilon / \sqrt{T}},$$

où $\bar{\hat{\varepsilon}} = \frac{1}{T} \sum_{t=1}^T \varepsilon_t$ et $S_\varepsilon = \frac{1}{T} \sum_{t=1}^T \varepsilon_t^2$. La statistique ζ est approximativement distribuée selon la loi normale centrée réduite, notée $\mathcal{N}(0, 1)$. Ainsi, on rejette H_0 si $|\zeta| > z_{\alpha/2}$, où $z_{\alpha/2}$ est la quantile d'ordre $\alpha/2$ de la loi $\mathcal{N}(0, 1)$.

b) Test de Box-Pierce et Ljung-Box

Le test de Box-Pierce permet d'identifier les processus de bruit blanc. Ce test s'écrit :

$$\begin{cases} H_0 : \rho(1) = \rho(2) = \dots = \rho(h) = 0, \\ H_1 : \text{il existe } i \text{ telle que } \rho(i) \neq 0. \end{cases}$$

Pour effectuer ce test, on utilise la statistique de Box et Pierce (1970) [9], donnée par :

$$Q = T \sum_{k=1}^h \hat{\rho}^2(k),$$

où h est le nombre de retards, T est le nombre d'observations et $\hat{\rho}(k)$ l'autocorrélation empirique d'ordre k . Sous l'hypothèse que les résidus suivent un bruit blanc. Cette statistique, suit une loi du khi-deux à h degré de liberté, notée \mathcal{X}_h^2 . Nous rejetons donc H_0 si $Q > \mathcal{X}_h^2(1 - \alpha)$, où $\mathcal{X}_h^2(1 - \alpha)$ est le quantile d'ordre $1 - \alpha$ de la loi du \mathcal{X}_h^2 .

Ljung et Box (1978) [14] ont améliorés ce test en considérant la statistique

$$Q_{LB} = T(T+2) \sum_{k=1}^h \frac{\hat{\rho}^2(k)}{T-j}.$$

c) Test de normalité

Plusieurs des modèles de séries temporelles supposent que les résidus sont indépendants et distribués selon la loi Normale. Un des tests permettant de vérifier la normalité des résidus est celui de Jarque-Bera. Les hypothèses à confronter sont :

$$\begin{cases} H_0 : \varepsilon_t \sim \mathcal{N}(0, 1), \\ H_1 : \varepsilon_t \not\sim \mathcal{N}(0, 1). \end{cases}$$

La statistique du test est définie par :

$$JB = \frac{T}{6} \beta_1^2 + \frac{T}{24} (\beta_2 - 3)^2,$$

où β_1 et β_2 sont respectivement les coefficients d'asymétrie (Skewness) et d'aplatissement (Kurtosis), donné par :

$$\beta_1 = \frac{1}{T} \sum_{t=1}^T \left(\frac{X_t - \bar{X}}{S} \right)^3 \text{ et } \beta_2 = \frac{1}{T} \sum_{t=1}^T \left(\frac{X_t - \bar{X}}{S} \right)^4,$$

ici, \bar{X} et S sont respectivement la moyenne et l'écart type empiriques. La statistique JB suit asymptotiquement la loi du χ^2_2 . On rejette donc l'hypothèse de normalité des résidus H_0 , si $JB > \chi^2$.

3. Choix du modèle

Une fois quelques modèles choisis, l'erreur de prévision dépend de la variance σ_ε^2 , notre but est donc de choisir le meilleur modèle qui minimise cette erreur par l'intermédiaire de l'un des critères suivants :

Critères standard :

i) L'erreur quadratique moyenne (MSE : Mean Square Error) :

$$MSE = \frac{1}{T} \sum_{t=1}^T \varepsilon_t^2.$$

ii) La racine de l'erreur quadratique moyenne (RMSE : Root Mean Square Error) :

$$RMSE = \sqrt{\frac{1}{T} \sum_{t=1}^T \varepsilon_t^2}.$$

iii) L'erreur absolue moyenne (MAE : Mean absolute Error) :

$$MAE = \frac{1}{T} \sum_{t=1}^T |\varepsilon_t|.$$

Critères d'information :

1. Le critère AIC (Akaike Information Criterion) :

$$AIC(p, q) = \ln \hat{\sigma}_\varepsilon^2 + 2(p + q) / T,$$

2. Le critère BIC (Schwartz Bayesian Information Criterion) :

$$BIC(p, q) = \ln \hat{\sigma}_\varepsilon^2 + (p + q) \ln(T) / T,$$

3. Hannan-Quinn :

$$\phi(p, q) = \ln \hat{\sigma}_\varepsilon^2 + (p + q) c \ln(\ln(T)) / T, \text{ avec } c > 2.$$

3.5.4 Prévision

Dans cette étape étant donnée une série stationnaire X_t , observée entre 1 et T , on cherche à faire de la prévision à horizon h : X_{T+1}, \dots, X_{T+h} . Tous les processus AR , MA et

ARMA seront supposés sous forme canonique, et n'avoir aucune racine unité. Aussi, les polynômes Φ et Θ ont leurs racines de module strictement supérieur à 1.

Prévisions à l'aide d'un modèle $AR(p)$

Soit un modèle $AR(p)$

$$X_t = c + \alpha_1 X_{t-1} + \alpha_2 X_{t-2} + \dots + \alpha_p X_{t-p} + \varepsilon_t, \text{ avec } \varepsilon_t \sim BB(0, \sigma_\varepsilon^2).$$

On appelle prévision optimale à la date $T+h$, la valeur

$$\hat{X}_T(h) = \hat{X}_{T+h} = E[X_{T+h} | \sigma(X)],$$

où $\sigma(X)$ est l'information disponible $f(X_1, \dots, X_T)$. On trouve ensuite

$$\begin{aligned} \hat{X}_{T+1} &= c + \alpha_1 X_T + \alpha_2 X_{T-1} + \dots + \alpha_p X_{T+1-p} \\ \hat{X}_{T+2} &= c + \alpha_1 \hat{X}_{T+1} + \alpha_2 X_T + \dots + \alpha_p X_{T+2-p+2} \\ &\vdots \\ \hat{X}_{T+h} &= c + \alpha_1 \hat{X}_{T+h-1} + \alpha_2 \hat{X}_{T+h-2} + \dots + \alpha_p \hat{X}_{T+h-p} \end{aligned}$$

On a ε_t est indépendant de $\sigma(\varepsilon)$ alors $E(\varepsilon_T | \sigma(\varepsilon)) = 0$. De façon générale, on a :

$$\hat{X}_{T+h} = \begin{cases} c + \alpha_1 \hat{X}_{T+h-1} + \dots + \alpha_{h-1} \hat{X}_{T+1} + \alpha_h X_T + \dots + \alpha_p X_{T+h-p} & \text{si } h \leq p, \\ c + \alpha_1 \hat{X}_{T+h-1} + \alpha_2 \hat{X}_{T+h-2} + \dots + \alpha_p \hat{X}_{T+h-p} & \text{si } h > p. \end{cases}$$

Prévisions à l'aide d'un modèle $MA(q)$

Considérons un modèle $MA(q)$

$$X_t = c + \Theta(L)\varepsilon_t = c + \varepsilon_t + \beta_1 \varepsilon_{t-1} + \dots + \beta_q \varepsilon_{t-q}, \quad \varepsilon_t \sim BB(0, \sigma_\varepsilon^2).$$

La prévision optimale à la date $T + h$ est donnée par :

$$\hat{X}_T(h) = \hat{X}_{T+h} = E(X_{T+h} | \varepsilon_T, \varepsilon_{T-1}, \dots).$$

On trouve alors

$$X_{T+1} = c + \varepsilon_{T+1} + \beta_1 \varepsilon_T + \beta_2 \varepsilon_{T-1} + \dots + \beta_q \varepsilon_{T+1-q},$$

$$\hat{X}_{T+1} = c + 0 + \beta_1 \varepsilon_T + \beta_2 \varepsilon_{T-1} + \dots + \beta_q \varepsilon_{T+1-q}$$

et

$$X_{T+2} = c + \varepsilon_{T+2} + \beta_1 \varepsilon_{T+1} + \beta_2 \varepsilon_T + \dots + \beta_q \varepsilon_{T+2-q},$$

$$\hat{X}_{T+2} = c + 0 + \theta_2 \varepsilon_T + \dots + \theta_q \varepsilon_{T+2-q}.$$

Jusqu'à

$$X_{T+h} = c + \varepsilon_{T+h} + \beta_1 \varepsilon_{T+h-1} + \beta_2 \varepsilon_{T+h-2} + \dots + \beta_q \varepsilon_{T+h-q},$$

$$\hat{X}_{T+h} = c + \hat{\varepsilon}_{T+h} + \beta_1 \hat{\varepsilon}_{T+h-1} + \beta_2 \hat{\varepsilon}_{T+h-2} + \dots + \beta_q \hat{\varepsilon}_{T+h-q},$$

où

$$\hat{\varepsilon}_{T+h} = E[\varepsilon_{T+h} | \sigma(X)] = \begin{cases} 0 & \text{si } h \geq 1, \\ \varepsilon_T & \text{si } h = 0. \end{cases}$$

$$\hat{\varepsilon}_{T+h-1} = E[\varepsilon_{T+h-1} | \sigma(X)] = \begin{cases} 0 & \text{si } h \geq 2, \\ \varepsilon_{T+h-1} & \text{si } h \leq 1. \end{cases}$$

⋮

$$\hat{\varepsilon}_{T+h-q} = E[\varepsilon_{T+h-q} | \sigma(X)] = \begin{cases} 0 & \text{si } h \geq q + 1, \\ \varepsilon_{T+h-q} & \text{si } h \leq q. \end{cases}.$$

Pour X_{T+h} on peut distinguer deux cas

$$\hat{X}_{T+h} = \begin{cases} c + \beta_h \varepsilon_T + \dots + \beta_h \varepsilon_{T+h-q} & \text{si } h \leq q, \\ c & \text{si } h > q. \end{cases}$$

Remarque 3.5.2 Cette forme est exacte mais pratiquement n'est pas utilisable, car le passé du bruit ne peut pas être connu, et en utilisant la représentation $AR(\infty)$ pour calculer les prévisions. On a

$$\Theta(L)^{-1} X_t = c + \varepsilon_t,$$

et

$$\Theta(L)^{-1} = \sum_{k=0}^{\infty} \rho_k L^k, \quad \sum_{k=0}^{\infty} |\rho_k| < \infty.$$

Alors

$$X_t = c - \sum_{k=0}^{\infty} \rho_k X_{t-k} + \varepsilon_t, \quad \text{et}$$

$$X_{T+h} = c - \sum_{k=0}^{\infty} \rho_k X_{T+h-k} + \varepsilon_{T+h}.$$

Pour les prévisions optimales

$$\begin{aligned} \hat{X}_{T+1} &= c - \sum_{k=1}^{+\infty} \rho_k \hat{X}_{T+1-k} \\ \hat{X}_{T+2} &= c - \sum_{k=1}^{+\infty} \rho_k \hat{X}_{T+2-k} \\ &\vdots \\ \hat{X}_{T+h} &= c - \sum_{k=1}^{+\infty} \rho_k \hat{X}_{T+h-k} = c - \sum_{k=1}^{h-1} \rho_k \hat{X}_{T+h-k} - \sum_{k=h}^{+\infty} \rho_k X_{T+h-k} \\ \hat{X}_{T+h} &= c - \sum_{k=1}^{h-1} \rho_k \hat{X}_{T+h-k} - \sum_{k=h}^{T+h-1} \rho_k X_{T+h-k} + \sum_{k=T+h}^{\infty} \rho_k X_{T+h-k}. \end{aligned}$$

Comme en pratique, on n'observe pas les X_t pour $t \leq 0$, on néglige donc le troisième

terme. Ce qui permet d'approcher les prévisions par :

$$\hat{X}_{T+h} = c - \sum_{k=1}^{+\infty} \rho_k \hat{X}_{T+h-k} = c - \sum_{k=1}^{h-1} \rho_k \hat{X}_{T+h-k} - \sum_{k=h}^{T+h-1} \rho_k X_{T+h-k}.$$

Prévisions à l'aide d'un modèle $ARMA(p, q)$

Considérons un modèle $ARMA(p, q)$: $\Phi(L) X_t = c + \Theta(L) \varepsilon_t$, où $\varepsilon_t \sim BB(0, \sigma_\varepsilon^2)$, ou encore

$$X_t = c + \sum_{j=1}^p \alpha_j X_{t-j} + \sum_{k=1}^q \beta_k \varepsilon_{t-k} + \varepsilon_t$$

avec $E(X_t) = \mu = \frac{c}{\Phi(1)} \neq 0$. En posant $Y_t = X_t - m$, on obtient

$$Y_t = \sum_{j=1}^p \alpha_j Y_{t-j} + \sum_{k=1}^q \beta_k \varepsilon_{t-k} + \varepsilon_t,$$

ce qui implique que

$$Y_{T+h} = \sum_{j=1}^p \alpha_j Y_{T+h-j} + \sum_{k=1}^q \beta_k \varepsilon_{T+h-k} + \varepsilon_{T+h}$$

La prévision optimale est donnée par

$$\hat{Y}_{T+h} = \sum_{j=1}^p \alpha_j \hat{Y}_{T+h-j} + \sum_{k=1}^q \beta_k \hat{\varepsilon}_{T+h-k}.$$

3.6 Application sous R

3.6.1 Analyse d'une série temporelle sous R

1. Créer un objet de type série temporelle : `série=ts(data, start, end, frequency)`

- `ts` : utilisée pour la création d'un objet série temporelle.
- `data` : contient le vecteur des données.
- `start` et `end` : mentionne les dates de début et de fin de la série.

- frequency : mentionne le nombre de données par unité de temps.

2. Représenter graphiquement un objet de type série temporelle : `plot.ts(série)`

3. Transformation et décomposition d'une série temporelle :

- La fonction `log` : permet de stabiliser la variance.
- La fonction `diff` permet d'appliquer l'opérateur de différentiation Δ^k ou Δ_s :
`diff(série, lag=s, difference=k)`
- La fonction `decompose` permet d'extraire d'une série temporelle : la composante saisonnière (`decomp$seasonal`), la tendance (`decomp$trend`), et la partie aléatoire (`decomp$random`), dont la syntaxe est donnée par :
`decomp=decompose(série, type=c("additive", "mutliplicative"))`

4. Lissage exponentiel : Les méthodes de lissages exponentiels sont disponibles sous R, grâce à la fonction `HoltWinters`.

- Lissage exponentiel simple : `sliss=HoltWinters(série, alpha= α , beta=F, gamma=F)`
- Lissage exponentiel double : `dliss=HoltWinters(série, alpha= α , beta= β , gamma=F)`

Exemple 3.6.1 (Modèle additif)

On considère la série de la concentration atmosphérique de `co2` à Mauna Loa (Hawai), c'est une série temporelle mensuelle de 468 observations entre 1959 et 1997. Le fichier de donnée contenu dans le logiciel R

```
série=ts(co2,c(1959,1),c(1997,12),f=12)
plot.ts(série,main="Modèle additif",xlab="Temps",ylab=expression(X[t]),col="4")
decomp=decompose(série,type="additive")
plot(decomp,xlab="Temps ",col="4")
```

Après avoir représenté graphiquement cette série (1.1) et sa décomposition (3.5), elle semble contenir une tendance linéaire et une composante saisonnière annuelle.

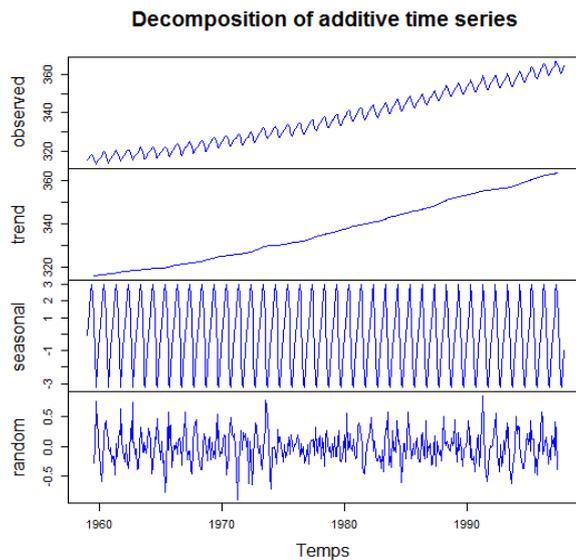


FIG. 3.5 – Décomposition de la série co2

Pour enlever ces deux composantes, on applique le filtre : $\Delta(\Delta_{12})X_t = (1 - L)(1 - L^{12})X_t$ comme suit :

```
D=diff(diff(série,lag=12),difference=1)
plot(D,xlab="Temps ",ylab=expression(Delta(X[t])),col="4")
```

On obtient le graphique de la nouvelle série stationnaire (3.6).

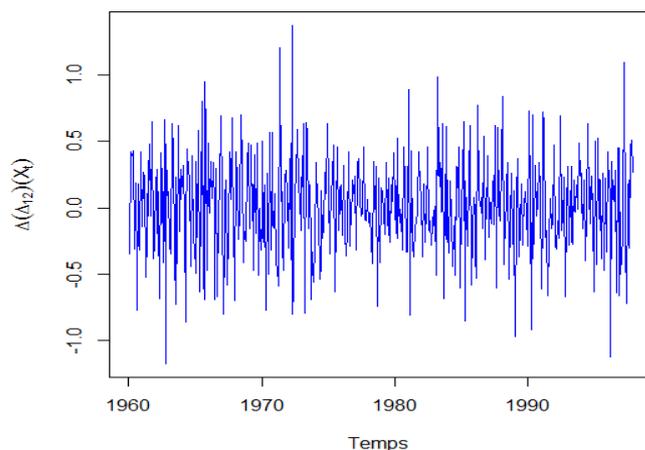


FIG. 3.6 – Représentation graphique de la série co2 après différenciation

Exemple 3.6.2 (Modèle multiplicatif)

On considère la série de nombre mondial mensuel de passagers aériens (en milliers) entre Janvier 1949 et Décembre 1960, pris de Box et Jenkins [?]. C'est une des séries les plus utilisées comme exemple d'application. Les données (AirPassengers) sont disponibles sous logiciel R. A l'aide des commandes suivantes :

```
X=AirPassengers
plot.ts(X,main="Modèle multiplicatif",xlab="Temps",
        ylab=expression(X[t]),col="4")
```

On obtient le graphique de la série (1.2), qui montre une tendance de type exponentiel ainsi qu'une saisonnalité de période 12 (saisonnalité annuelle). Si on effectue une transformation logarithmique de cette série, on se ramène à une série additive, représenté dans la figure 3.7 (en haut), et qui présente une tendance linéaire et conserve une saisonnalité (au milieu).

```
decomp=decompose(log(X),type="additive")
plot(decomp,xlab="Temps ",col="4")
```

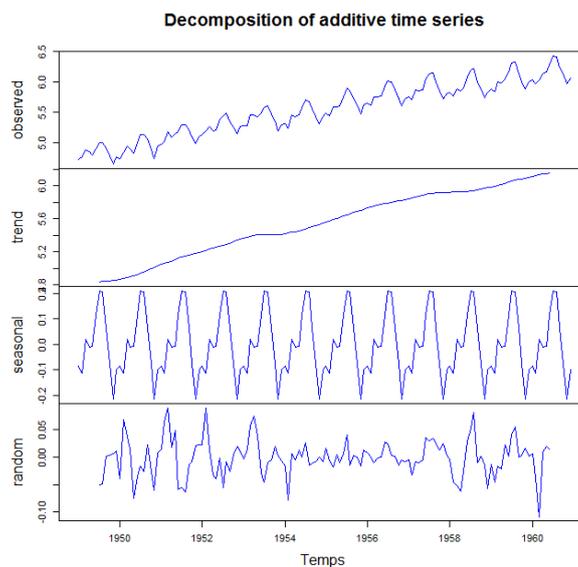


FIG. 3.7 – Décomposition de la série $\log(X_t)$

Pour éviter le problème de non stationnarité de la série, on effectue donc une double différenciation $(1 - L)$ et $(1 - L^{12})$, on obtient la figure (3.8), à l'aide de ce code R.

```
Xdd=diff(diff(log(X),lag=12),difference=1)
plot.ts(Xdd,main="Série stationnaire",
        ylab=expression(Delta(Delta[12])(X[t])),col="4")
```

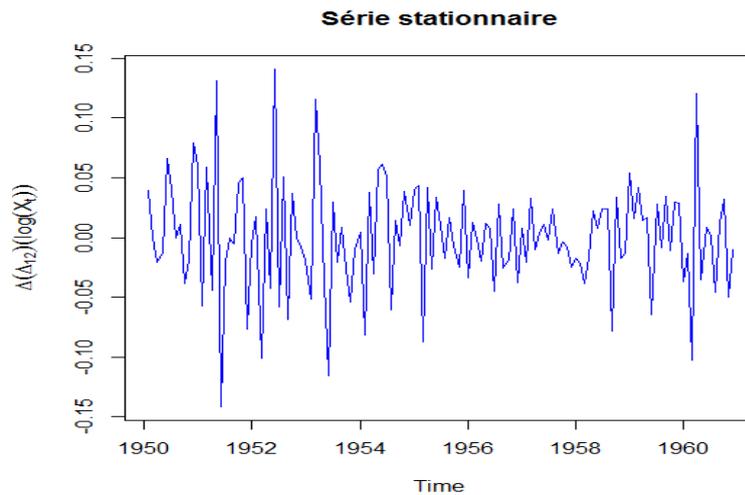


FIG. 3.8 – Représentation graphique de la série $\log(X_t)$ après une double différenciation

3.6.2 Simulation des modèles stationnaire sous R

Exemple 3.6.3 (Modèle stationnaire)

On simule un bruit blanc gaussien de variance $\sigma_\varepsilon^2 = 2.5$, par : (voir la figure 2.1)

```
n=200
bruit=ts(rnorm(n,0,sqrt(2.5)))
plot.ts(bruit,main="Bruit Blanc",col="4",xlab="temps",
        ylab=expression(epsilon[t]))
```

Exemple 3.6.4 (Modèle non stationnaire)

On simule une marche aléatoire comme suit : (voir la figure 2.2)

```
x=rep(0,n)
for(i in 2 :n){x[i]=x[(i-1)]+bruit[i]}
plot.ts(x,main="Marche aléatoire",xlab="temps",ylab=expression(X[t]),col="4")
```

3.6.3 Simulation des modèles linéaire sous R

- **Déclaration d'un modèle $AR(p)$, $MA(q)$ et $ARMA(p, q)$ est respectivement :**

```
ts=arima.sim(list(order=c(p,0,0),ar=x),sd=sqrt( $\sigma_\varepsilon^2$ ),n)
```

```
ts=arima.sim(list(order=c(0,0,q),ma= $\varepsilon$ ),sd=sqrt( $\sigma_\varepsilon^2$ ),n)
```

```
ts.arma=arima.sim(list(order=c(p,0,q),ar=x,ma= $\varepsilon$ ),sd=sqrt( $\sigma_\varepsilon^2$ ),n)
```

- **Test de stationnarité (Test de Dickey-Fuller augmenté) :**

Charger le package tseries : `library(tseries)`

```
adf.test(ts,alternative=c("stationary"),k=0)
```

Si p -value < à 0.05 la série est stationnaire.

Tous les tests sont effectués au seuil 5%.

- **Calcul et représentation des autocovariances , autocorrélations et autocorrelations partielles :**

```
acf(ts,type="covariance")
```

```
acf(ts,type="correlation")
```

```
pacf(ts)
```

- **Estimation des paramètres d'un $AR(p)$**

La fonction `ar.yw(ts,order=p)` nous permet de fournir l'estimation des paramètres d'un modèle $AR(p)$ sous R.

Exemple 3.6.5 (Modèle AR(2))

On génère une série de données de taille $n = 500$ à partir d'un modèle AR(2), tel que :

$$X_t = 0.8X_{t-1} - 0.4X_{t-2} + \varepsilon_t \text{ et } \varepsilon_t \sim \mathcal{BB}(0, 1).$$

```
ts.ar=arima.sim(list(order=c(2,0,0),ar=c(0.9,-0.8)),sd=1,n=500)
par(mfrow=c(3,1))
plot(ts.ar,main=expression(paste("Modèle AR(2)",alpha[1]==0.9,"",
    alpha[2]==-0.8)),ylab=expression(X[t]),xlab="Temps")
acf(ts.ar,main="Autocorrélogramme",xlab="h",ylab=expression(rho[X](h)))
pacf(ts.ar,main="Autocorrélogramme partiel",xlab="h",
    ylab=expression(phi[h]),col="4")
```

Les graphiques de la trajectoire, la fonction d'autocorrélation et la fonction d'autocorrélation partielle de ce modèle sont déjà représentés par la figure (3.2) et commentés dans l'exemple 3.1.2, avec la ligne pointillée bleue indique le niveau en-dessous duquel l'autocorrélation ou l'autocorrélation partielle n'est plus statistiquement significative.

```
library(tseries)
adf.test(ts.ar,alternative=c("stationary"),k=0)
```

D'après le test de stationnarité la p -value = 0.01 < 0.05, alors la série de modèle AR(2) est stationnaire.

Exemple 3.6.6 (Modèle MA(1))

On génère une série de donnée de taille $n = 200$ à partir d'un modèle MA(1), tel que :

$$X_t = \varepsilon_t + 0.6\varepsilon_{t-1} \text{ avec } \varepsilon_t \sim \mathcal{BB}(0, 1).$$

```
ts.ma=arima.sim(list(order=c(0,0,1),ma=0.6),sd=sqrt(4),n=200)
par(mfrow=c(2,1))
ts.plot(ts.ma,type="l",xlab="temps",ylab="MA(1)",col="4")
plot(ts.ma,main=expression(paste("Modèle MA(1)",beta==0.6)),
      ylab=expression(X[t]),xlab="t")
acf(ts.ma,main="Autocorrélogramme",xlab="h",
     ylab=expression(rho[X](h)),ylim=c(-0.5,1.2))
```

Les graphiques de la trajectoire et la fonction d'autocorrélation de ce modèle sont déjà représentés par la figure (3.3) et commentées dans l'exemple 3.2.1.

```
library(tseries)
adf.test(ts.ma,alternative=c("stationary"),k=0)
```

Le test donne un $p - value = 0.01 < 0.05$, alors le modèle est stationnaire et inversible.

3.6.4 Méthode de Box-Jenkins sous R

1. Identification : La fonction `adf.test()` teste la stationnarité de la série. On peut également utiliser le test de Philippe et Perron à travers la fonction `kpss.test()` ; si $p - value < 0.05$ la série est stationnaire. On peut aussi tester la stationnarité graphiquement.
2. La fonction `M=arima(série,order=c(p,0,q))` pour estimer les paramètres d'un modèle $ARMA(p, q)$
3. La fonction `M=arima(série,order=c(p,d,q),list=(order=c(P,D,Q),period=s))` pour estimer les paramètres d'un modèle $SARIMA(p, d, q)(P, D, Q)_s$
4. La fonction `Box.test(série)` teste la bruité de la série ; si $p - value > 0.05$ la série est un bruit blanc.
5. `qqplot`, `qqnorm` et `jarque.bera.test` ce sont des fonctions pour tester la normalité des résidus.
6. La fonction `forecast(M,h)` permet de faire les prévisions à l'horizon h .

Exemple 3.6.7 (Méthode de Box et Jenkins)

On applique la méthode de Box et Jenkins sur la série `AirPassengers`, introduit dans l'exemple 3.6.2. Cet exemple s'inspire fortement de [10], [20], [21], [22] et [13].

Etape 1 : Identification du modèle

L'analyse graphique de cette série (figure 1.2) montre l'existence d'une tendance et une saisonnalité, ce qui traduit un problème de non stationnarité. Une autre façon de tester la stationnarité est d'appliquer le test de Dickey-Fuller augmenté sur notre série en utilisant la fonction `adf.test` du package `tseries` sous R.

```
library(tseries)
adf.test(X,alternative="stationary",k=12)
```

Le test donne un $p - value = 0.7807 > 0.05$, ce qui confirme que la série n'est pas stationnaire.

Maintenant, on travaille sur le logarithme de la série afin de pallier l'accroissement de la saisonnalité (voir la figure 3.7). On passe ainsi d'un modèle multiplicatif à un modèle additif. On effectue donc une double différenciation $(1 - L)$ et $(1 - L^{12})$. Les données transformées semblent stationnaires, le graphe de la nouvelle série oscille davantage autour de 0. Dans ce cas $d = 1$ et $D = 1$, on identifiera donc un modèle *ARMA* sur la série :

$$Y_t = (1 - L)(1 - L^{12}) \log X_t$$

L'autocorrélogramme (*ACF*) et l'autocorrélogramme partiel (*PACF*) de la série Y_t sont illustrés dans la figure (3.9).

```
library(caschrono)
acf2(Xdd,max.lag=50,main=expression(Delta(Delta[12])(log(X[t])))
```

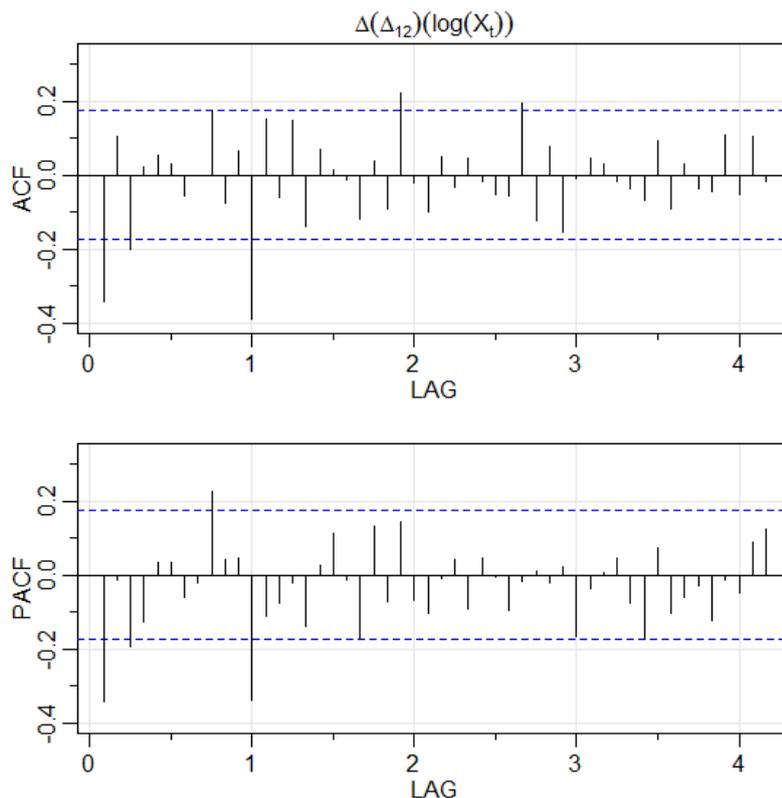


FIG. 3.9 – Autocorrélogramme et autocorrélogramme partiel de la série Y_t

a) **Composante saisonnière** : A l'échelle annuelle, la série présente une forte autocorrélation au premier retard ($1s$ avec $s = 12$). Ainsi, que l'ACF décroît rapidement après le premier retard, alors que le PACF n'est pas très claire du fait du "rebond" au $3^{\text{ème}}$ retard ($3s$), ceci suggère un $SARMA(P, Q)$ avec $Q = 1$ et $P = 0$ dans la saison.

b) **Composante non saisonnière** : En inspectant le même graphique, on se concentre juste sur une saison. Nous constatons que l'ACF et le PACF n'a pas vraiment de décrochage sur une année. Ceci suggère donc pour la composante non saisonnière trois modèle $ARMA(1, 1)$, $ARMA(0, 1)$ et $ARMA(1, 0)$.

Pour résumer, on essaie d'abord un premier modèle $SARIMA(1, 1, 1)(0, 1, 1)_{12}$ sur la série $\log X_t$

Etape 2 : Estimation

On estime les paramètres du modèle à l'aide du code suivant :

```
library(forecast)
modèle1=Arima(log(X),order=c(1,1,1),list(order=c(0,1,1),period=12),
             include.mean=FALSE)
```

La sortie du code, nous donne les estimateurs des paramètres et la valeur des critères AIC et BIC.

```
#Series : log(X)
#ARIMA(1, 1, 1)(0, 1, 1)[12]
#Coefficients :
#   ar1    ma1    sma1
#   0.1960 -0.5784 -0.5643
#s.e 0.2475 0.2132 0.0747
#sigma^2 estimated as 0.001375 : loglikelihood = 244.95
#AIC = -481.9 AICc = -481.58 BIC = -470.4
```

Etape 3 : Validation

1. Test de significativité des coefficients : On applique le test de student, en utilisant la package `caschrono`, comme suit :

```
library(caschrono)
t_stat(modèle1)
```

```
#   ar1    ma1    sma1
# t.stat 0.792074 -2.712696 -7.554402
# p.val 0.428318 0.006674 0.000000
```

La sortie du test nous indique que le coefficient du $AR(1)$ est non significatif au seuil 5% (car la p -value > 0.05). Nous devrions donc essayer de supprimer un paramètre de la partie intra-saisons. Dans ce cas, il semblerait pertinent de prendre le second modèle $SARIMA(0, 1, 1)(0, 1, 1)_{12}$ et le troisième modèle $SARIMA(1, 1, 0)(0, 1, 1)_{12}$. On estime donc les coefficients des deux modèles.

```

modèle2=Arima(log(X),order=c(0,1,1),list(order=c(0,1,1),period=12),
  include.mean=FALSE)

modèle3=Arima(log(X),order=c(1,1,0),list(order=c(0,1,1),period=12),
  include.mean=FALSE)

```

Les résultats obtenus sont les suivants :

<i>#Series : log(X)</i>	<i>#Series : log(X)</i>
<i>#ARIMA(0, 1, 1)(0, 1, 1)[12]</i>	<i>#ARIMA(1, 1, 0)(0, 1, 1)[12]</i>
<i>#Coefficients :</i>	<i>#Coefficients :</i>
<i># ma1 sma1</i>	<i># ar1 sma1</i>
<i># -0.4018 -0.5569</i>	<i># -0.3395 -0.5619</i>
<i>#s.e 0.0896 0.0731</i>	<i>#s.e 0.0822 0.0748</i>
<i>#sigma^2 estimated as 0.001371</i>	<i>#sigma^2 estimated as 0.001391</i>
<i># : loglikelihood = 244.7</i>	<i>v : loglikelihood = 244.7</i>
<i>#AIC = -483.4 AICc = -483.21</i>	<i>#AIC = -481.49 AICc = -481.3</i>
<i>#BIC = -474.77</i>	<i>#BIC = -472.86</i>

On va tester maintenant la signification des coefficients des deux modèles

```

t_stat(modèle2)

t_stat(modèle3)

```

<i># ma1 sma1</i>	<i># ar1 sma1</i>
<i># t.stat -4.482497 -7.618985</i>	<i># t.stat -4.129478 -7.510892</i>
<i># p.val 0.000007 0.000000</i>	<i># p.val 0.000036 0.000000</i>

Les paramètres des modèles sont donc significatifs au seuil 5%.

2. Test sur le bruit blanc :

Le diagnostic des résidus associés au deuxième modèle est illustré dans la figure (3.10).

La série des résidus semble visuellement stationnaire, l'ACF des résidus indique bien un bruit blanc, le qqplot selon la loi normale indique la normalité des résidus et les p-valeurs du test de Ljung-Box, sauf pour une ou deux valeurs aberrantes, indiquent que la partie résiduelle est bien un bruit blanc.

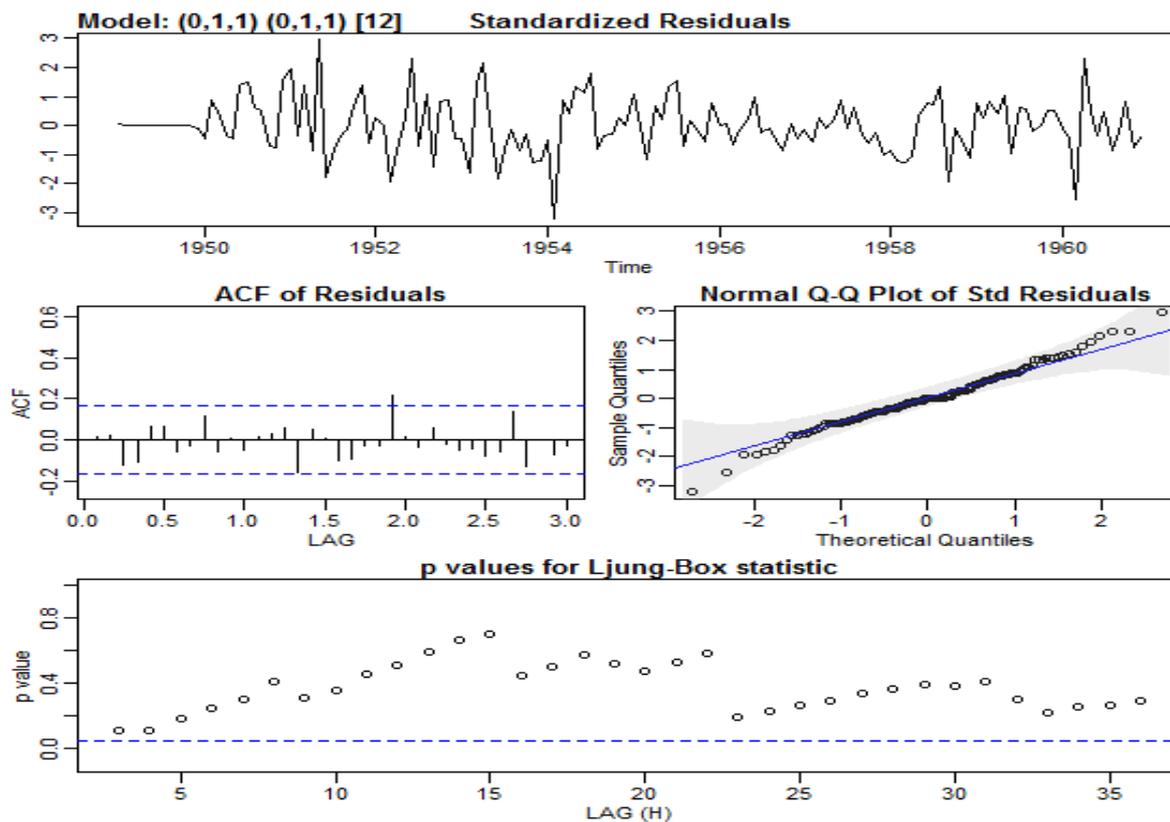


FIG. 3.10 – Analyse des résidus

3. Choix du modèle : Tous les critères d'information AIC et BIC préfèrent le modèle $SARIMA(0, 1, 1)(0, 1, 1)_{12}$. Le meilleur modèle ajusté est donc :

$$(1 - L)(1 - L^{12}) \log(X_t) = (1 - 4.482497L)(1 - 7.618985L^{12}) \varepsilon_t.$$

Étape 4 : Prévision

D'après le modèle retenu on peut prédire des résultats sur 2 ans (de 1961 à 1962). Les résultats sont illustrés dans la figure 3.11.

```

pred_mode2=forecast(modèle2,h=24,level=95)
pred=exp(pred_mode2$mean)
pred_l=ts(exp(pred_mode2$lower),start=c(1961,1),frequency=12)
pred_u=ts(exp(pred_mode2$upper),start=c(1961,1),frequency=12)
ts.plot(X,pred,pred_l,pred_u,xlab="t",ylab="AirPassengers",
        col=c(1,2,4,4),lty=c(1,1,2,2),lwd=c(1,2,2,2),main="Prévision (1961-1962)")
    
```

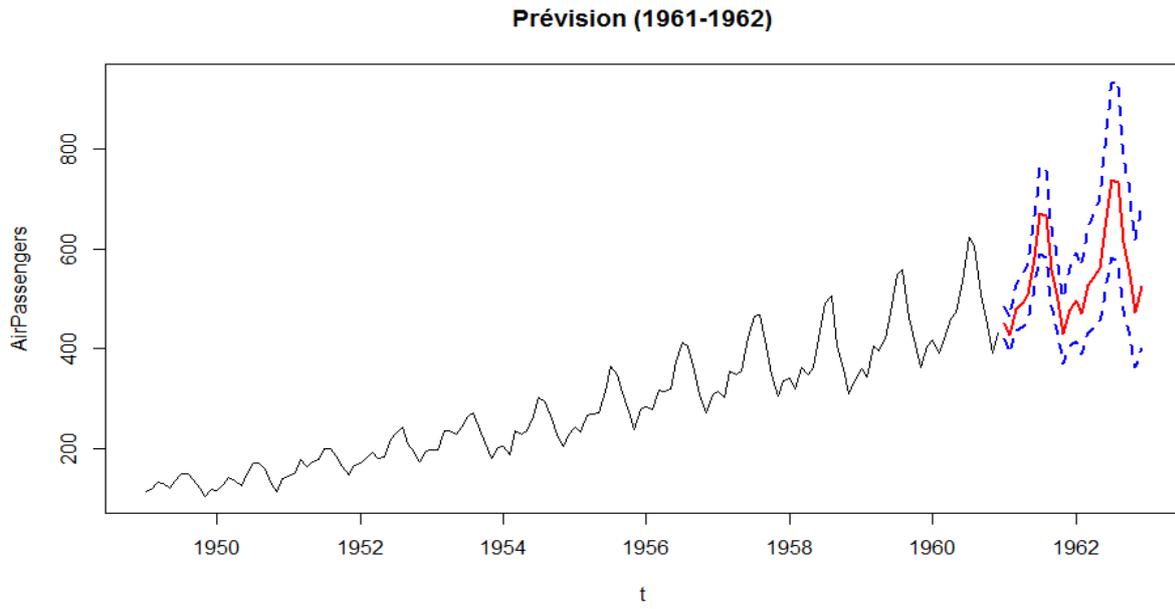


FIG. 3.11 – Représentation graphique des valeurs prédictives de la série AirPassengers sur 2 ans

Conclusion

Au cours de ce mémoire, nous avons regroupés de près quelques méthodes qui permettent de modéliser (c-à-d identifier un modèle adéquat) des séries temporelles. Nous sommes intéressés aux modèles linéaires dont on a synthétisé les propriétés statistiques des modèles *AR*, *MA* et *ARMA*.

Nous nous servons de ces méthodes de modélisation pour prédire le comportement des séries chronologiques. Pour ce faire, nous avons utilisé la méthode de Box et Jenkins.

Nous avons complétés cette étude théorique par une mise en pratique sous R à l'aide des packages *tseries*, *caschrono*, *forecast* et *astsa*. Après avoir appliqué la méthode de Box et Jenkins qui nous a permis de modéliser la série du nombre de passagers aériens (*Airpassnegers*) et d'aboutir au modèle $SARIMA(0, 1, 1)(0, 1, 1)_{12}$, nous avons procédé au calcul des prévisions sur une période de deux ans.

En conclusion, les notions présentées dans ce mémoire ne constituent qu'une petite part des connaissances actuelles sur le sujet, il est extrêmement important de modéliser et de prédire des séries multivariées, des séries non stationnaires, ainsi de présenter des modèles non linéaires (*ARCH*, *GARCH*, ...).

Bibliographie

- [1] Anderson, T. W., (1971). *The Statistical Analysis of Time Series*. Wiley, New York.
- [2] Aragon, Y. (2011). *Séries temporelles avec R. Méthodes et cas*. Springer, Paris.
- [3] Avram, F. (2017). *Séries temporelles, avec R*.
- [4] Avram, F. *Series temporelles, Régression, Interpolation, et Géostatistique*.
- [5] Benameur, S., & Yahia, D. (2020). *Note de cours : Séries Temporelles*. Université Biskra.
- [6] Brockwell, P. J., & Davis, R. A. (2016). *Introduction to time series and forecasting*. springer.
- [7] Box, G.E.P. & Jenkins, G.M. (1976) *Time Series Analysis : Forecasting and Control*. Holden Day, San Francisco.
- [8] Box, G.E.P., Jenkins, G.M., Reinsel, G.C., Ljung, G.M. (2016) *Time Series Analysis : Forecasting and Control. Fifth Edition*,.Wiley.
- [9] Box G.E.P. & Pierce D.A. (1970). *Distribution of Residual Autocorrelations in Autoregressive Moving Average Time Series Models*. *Journal of the American Statistical Association*, vol. 65 .
- [10] Charpentier, A. (2012). *Modèles de prévision-Séries temporelles*. Paris, France.
- [11] Dufour, J. M. (2003). *Histoire de l'analyse des séries chronologiques*.
- [12] Fortier, S. (2013). *Les modèles MA, AR et ARMA multidimensionnels : estimation et causalité*. *CaMUS (Cahiers Mathématiques de l'Université de Sherbrooke)*, 4, 112-136.

- [13] Lefieux, V. Séries temporelles sous R, Exemples de séries temporelles. Université de Sorbonne.
- [14] Ljung, G. & Box, G. (1978). On a measure of lack of fit in time series models. *Biometrika* 65, 297.
- [15] Monbet, V. (2011). Modélisation de séries temporelles. Université de Rennes.
- [16] Perraudin, C. (2004–2005). Chapitre 1 : Les modèles ARMA stationnaires, université Paris I.
- [17] Shumway, R. H., & Stoffer, D. S. (2017). Time series analysis and its applications : with R examples. Fourth Edition, Springer.
- [18] von Sachs, R., & Van Bellegem, S. (2005). Séries chronologiques, *STAT*, 2414.
- [19] Yahia, D. (2005). Séries Temporelles et test d'adéquation pour une modèle GARCH(1, 1). Mémoire de Magister. Université Biskra.
- [20] <https://rpubs.com/nohaelprince/47545>
- [21] <http://www.math.univ-montp2.fr/~ribatet/CentraleNantes/STA2/AirPassengersCaseStudy.Rmd>
- [22] http://rstudio-pubs-static.s3.amazonaws.com/311446_08b00d63cc794e158b1f4763eb70d43a.html
- [23] https://members.loria.fr/FSur/enseignement/modseries/TP3_eleves.R

Annexe : Abréviations et Notations

Les différentes abréviations et notations utilisées tout au long de ce mémoire sont expliquées ci-dessous :

- E : Espérance.
- var : Variance.
- cov : Covariance.
- $corr$: corrélation.
- ACF : Fonction d'autocorrélation simple.
- $PACF$: Fonction d'autocorrélation partielle.
- AIC : Critère d'information d'Akaike.
- BIC : Critère d'information bayésien.
- \mathcal{BB} : Bruit blanc.
- AR : Autorégressif.
- MA : Moyenne mobile.
- $ARMA$: Autorégressif moyenne mobile.
- γ : Fonction d'autocovariance.
- ρ : Fonction d'autocorrélation.
- \mathcal{R}_P : Matrice des corrélations.

X_t : Série temporelle.

S_t : Saisonalité.

Z_t : Tendance.

ε_t : Erreurs.

L : Opérateur de retard.

F : Opérateur d'avance.

Δ : Opérateur de différentiation.

$\stackrel{D}{=}$: Egalité en distribution.

JB : Test de Jarque-Bira.

\mathcal{X} : Khi-deux.

H_0 : Hypothèse nulle.

H_1 : Hypothèse alternative.

Résumé

Nous présentons dans ce mémoire un aperçu sur les séries temporelles; les types, les propriétés statistiques et la stationnarité, dont l'objectif est de regrouper les différentes méthodes permettent d'analyser et modéliser des séries chronologique afin de pouvoir faire des prévisions.

Nous nous intéressons particulièrement à la modélisation et la prévision des séries chronologiques à travers la méthodologie élaborée par Box et Jenkins, au moyen des modèles linéaires à savoir le modèle: AR, MA, ARMA, ARIMA et SARIMA.

On arrive finalement à une mise en pratique d'un exemple d'une série temporelle réelle sous logiciel R.

Mots clés : Série temporelle ; Stationnarité ; Modèles linéaires ; Fonction d'autocorrélation, Fonction d'autocorrélation partielle ; Méthode de Box et Jenkins ; Prévision.

Abstract

We present in this memory an overview on time series; the types, the statistical properties and the stationarity, the objective of which is to gather the different methods allow time series to be analyzed and modeled in order to be able to make forecasts.

We are particularly interested in modeling and forecasting time series through the methodology worked out by Box and Jenkins, by means of linear models namely the AR, MA, ARMA, ARIMA and SARIMA model.

We finally arrive at a practical application of an example of a real time series under the software R.

Keywords: Time series; Stationarity; Linear models; Autocorrelation function, Partial autocorrelation function; Box and Jenkins method; Forecast.

ملخص

نقدم في هاته المذكرة لمحة عامة عن السلاسل الزمنية؛ الأنواع ، الخصائص الإحصائية و الثبات، بحيث الهدف منها هو جمع الطرق المختلفة التي تسمح بتحليل ونمذجة السلاسل الزمنية من أجل التمكن من إجراء التنبؤات.

نحن مهتمون بشكل خاص بنمذجة السلاسل الزمنية والتنبؤ بها من خلال المنهجية التي وضعها بوكس و جنكنس، عن طريق النماذج الخطية وهي: AR، MA، ARMA، ARIMA و SARIMA.

وصلنا أخيراً إلى تطبيق عملي لمثال لسلسلة زمنية حقيقية في إطار البرنامج R.

الكلمات الرئيسية: سلسلة زمنية ؛ الثبات؛ نماذج خطية؛ دالة الارتباط الذاتي؛ دالة الارتباط الذاتي الجزئي؛ طريقة بوكس و جنكنس؛ توقعات.