

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

M/S 10.062

Université 8 Mai 1945 Guelma

Faculté des Mathématiques et de l'Informatique
et des Sciences de la Matière
Département de Mathématiques



Mémoire

Présenté en vue de l'obtention du diplôme de

Master Académique en Mathématiques

Option : **Probabilités et Applications**



Par :

Melle: Ibtissam. Maghmouli

Melle : Sara. Kerdoussi

Intitulé

**Analyse des séries chronologiques
Modèles ARMA**

Dirigé par : A. Yousfi

Devant le jury

PRESIDENT

G. Rebai

Univ-Guelma

RAPPORTEUR

G. Rebai

Univ-Guelma

EXAMINATEUR1

S. Bakhouch

Univ-Guelma

Session Juin 2013

Remerciements

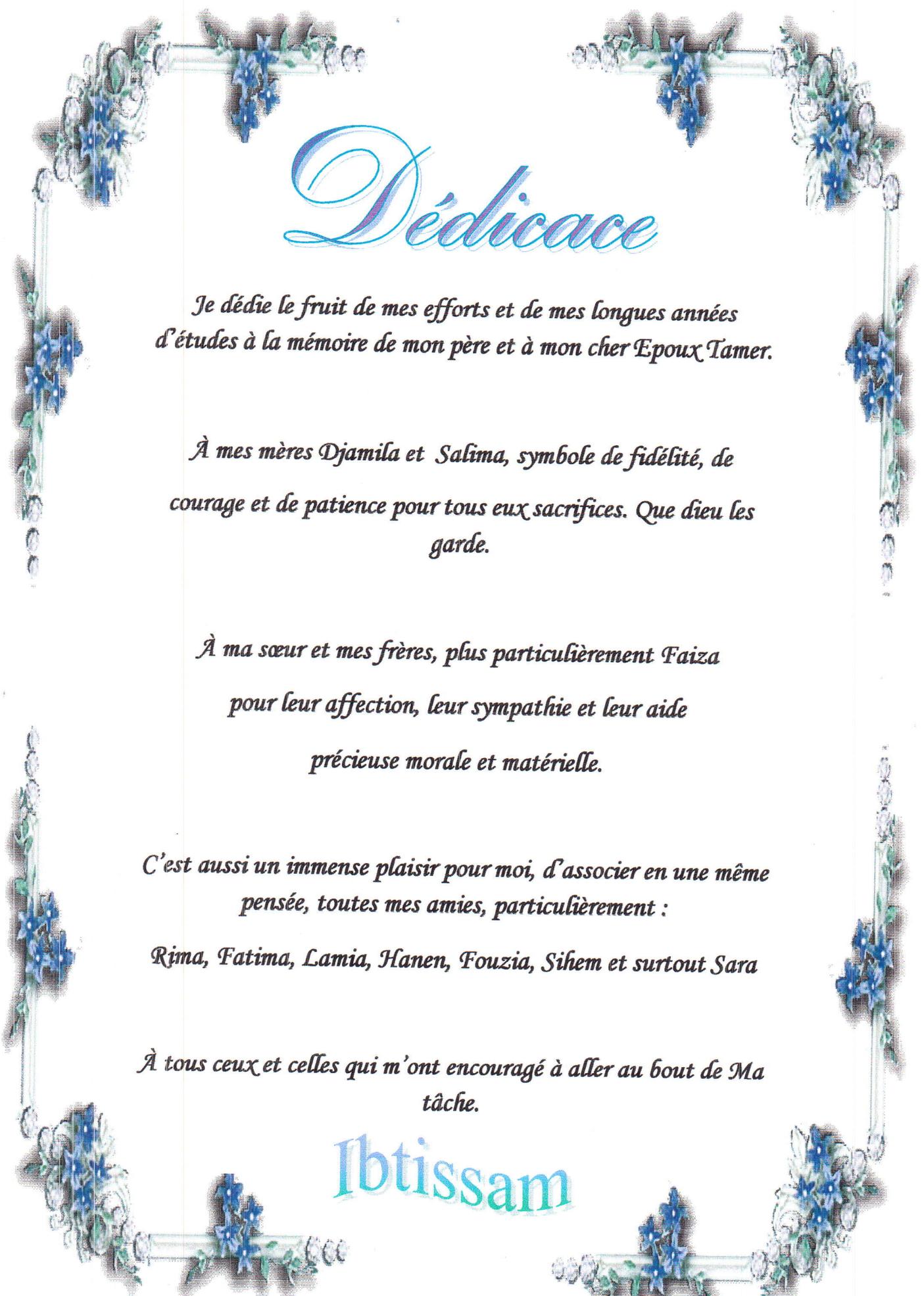
Nous remercions en premier lieu Dieu de nos avoir donné la force et la patience de terminer ce travail, nous souhaitons adresser aussi nos vifs remerciements à certaines personnes qu'ont aidé au cours de ces deux dernières années.

Nous voudrions remercier très chaleureusement notre encadreur « A. Yousfi » qui a été à côtés de nous tout au long de cette étude et qui nous a fait bénéfice de son savoir.

*Nous tenons également à remercier monsieur « Le chef de département »
Pour l'aide précieuse qu'il nous a apporté pendant nos études et la préparation de ce mémoire.*

Nos remerciements les plus sincères s'adressent à tous les enseignants et personnelles du Département de Mathématique.

Sara et Ibtissam
Sara et Ibtissam



Dédicace

Je dédie le fruit de mes efforts et de mes longues années d'études à la mémoire de mon père et à mon cher Epoux Tamer.

À mes mères Djamila et Salima, symbole de fidélité, de courage et de patience pour tous eux sacrifices. Que dieu les garde.

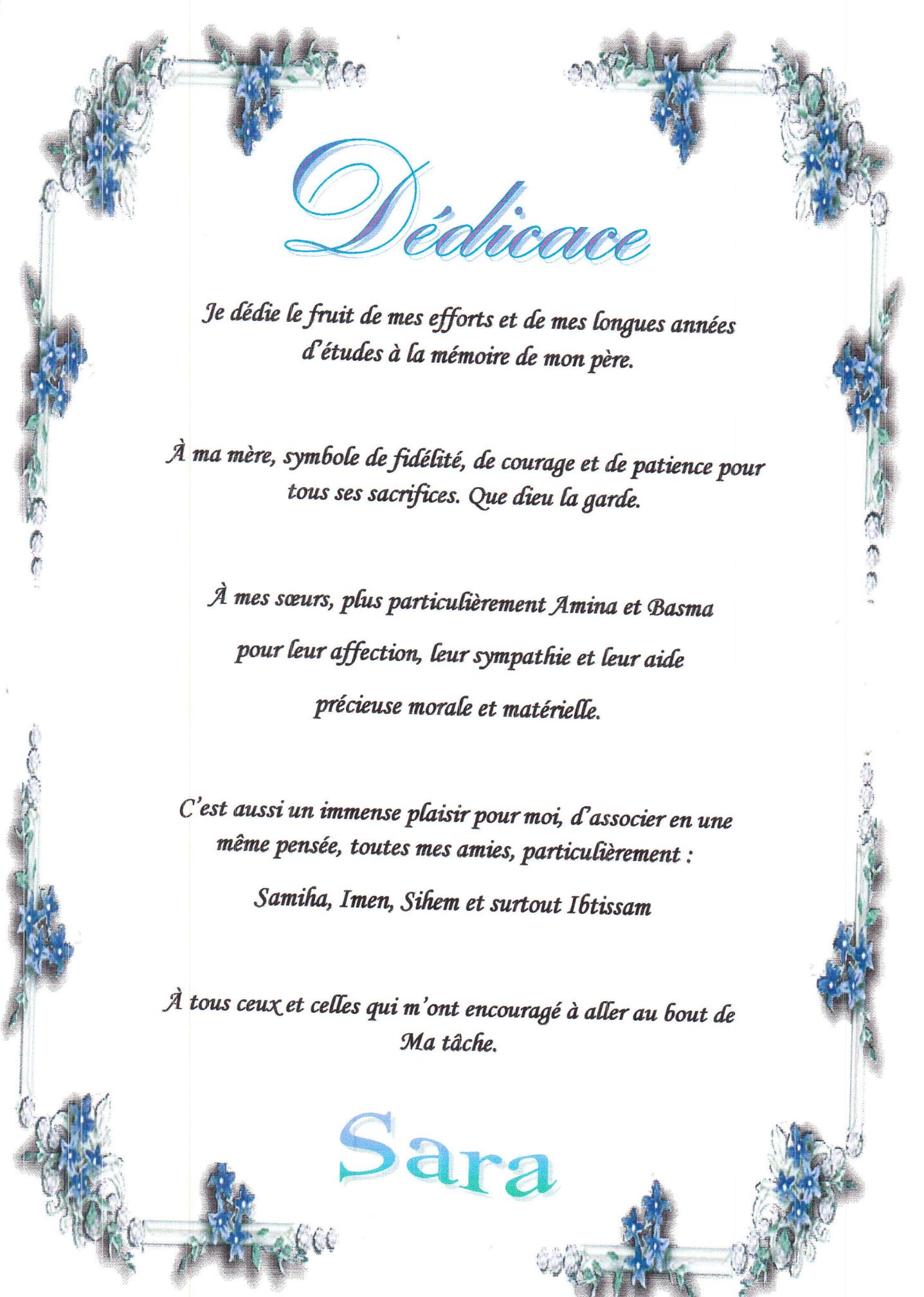
À ma sœur et mes frères, plus particulièrement Faiza pour leur affection, leur sympathie et leur aide précieuse morale et matérielle.

C'est aussi un immense plaisir pour moi, d'associer en une même pensée, toutes mes amies, particulièrement :

Rima, Fatima, Lamia, Hanen, Fouzia, Sihem et surtout Sara

À tous ceux et celles qui m'ont encouragé à aller au bout de Ma tâche.

Ibtissam



Dédicace

*Je dédie le fruit de mes efforts et de mes longues années
d'études à la mémoire de mon père.*

*À ma mère, symbole de fidélité, de courage et de patience pour
tous ses sacrifices. Que dieu la garde.*

*À mes sœurs, plus particulièrement Amina et Basma
pour leur affection, leur sympathie et leur aide
précieuse morale et matérielle.*

*C'est aussi un immense plaisir pour moi, d'associer en une
même pensée, toutes mes amies, particulièrement :*

Samiha, Imen, Sihem et surtout Ibtissam

*À tous ceux et celles qui m'ont encouragé à aller au bout de
Ma tâche.*

Sara

Table des matières

Introduction générale	iii
1 Processus stochastiques	1
1.1 Stationnarité	1
1.1.1 Définition (stationnarité en moyenne et en variance)	1
1.1.2 Définition (stationnarité faible)	2
1.1.3 Définition (stationnarité forte)	2
1.1.4 Définition (bruit blanc)	2
1.2 Fonction d'autocovariance	4
1.2.1 Définition	4
1.2.2 Propriétés	4
1.3 Fonction d'autocorrélation	5
1.3.1 Définition	5
1.3.2 Propriétés	5
1.4 Fonction d'autocorrélation partielle	5
1.5 Causalité et inversibilité	7
1.5.1 Définition	7
1.5.2 Définition	7
1.6 Théorème de décomposition de Wold	7
1.6.1 Théorème	7
2 Modèle AutoRégressif Moyenne Mobile (ARMA)	10
2.1 Les opérateurs	10
2.1.1 Opérateur retard (l'opérateur B)	10
2.1.2 Inversion des polynômes	11
2.1.3 Théorème	12
2.1.4 Opérateur différence (l'opérateur Δ)	13

TABLE DES MATIÈRES

2.2	Processus AutoRégressif (AR)	13
2.2.1	Propriétés statistiques du processus $AR(1)$	14
2.2.2	Propriétés statistiques du processus $AR(p)$	18
2.3	Processus Moyenne Mobile (MA)	20
2.3.1	Propriétés statistiques du processus $MA(1)$	21
2.3.2	Propriétés statistiques du processus $MA(q)$	23
2.4	Modélisation statistique des processus (ARMA)	25
2.4.1	Propriétés statistiques du processus $ARMA(1,1)$: . . .	25
2.4.2	Propriétés statistiques du processus $ARMA(p,q)$: . .	28
3	Méthodologie de Box-Jenkins	30
3.1	Identification des processus ARMA	30
3.1.1	Test de Bartlett	31
3.2	Estimation des paramètres du modèle	32
3.2.1	Méthodes des moments :	32
3.2.2	La méthode de maximum de vraisemblance	33
3.2.3	La méthode des moindres carrées	33
3.3	Validation du processus $ARMA(p,q)$	34
3.3.1	Tests de Student sur les coefficients	34
3.3.2	Test d'autocorrélation des résidus	35
3.3.3	Tests d'homoscédasticité	36
3.3.4	Tests de normalité	37
3.3.5	Critères de choix des modèles	38
3.4	Prévision	39
3.4.1	Transformation de la série	39
3.4.2	Prédicteur pour un processus ARMA	40
4	Conclusion générale	42
A	Annexe	43
A.1	La convergence en moyenne quadratique	43
A.2	Liste des abreviations	43
A.3	Tables de valeurs numériques	44
A.3.1	Lois continues	44
Bibliographie		50

Introduction générale

Parmi l'ensemble des modèles stochastiques, une classe particulière, appelée classe des processus aléatoires stationnaires va permettre de caractériser la structure de corrélation d'une série.

À la différence de la modélisation économétrique, qui consistait à construire un modèle statistique reliant variable endogène et variables exogènes en se basant sur la théorie économique, l'approche en séries chronologiques relie la valeur courante de la variable endogène uniquement à ses valeurs passées ainsi qu'aux valeurs passées et courante d'une perturbation aléatoire. Le modèle statistique issu de cette approche est appelé modèle de série chronologique. Ce type de modèle est différent des modèles économétriques principalement en ce qu'il n'est pas issu d'un cadre de théorie économique.

Une série chronologique ou série temporelle est un ensemble d'observations qui se distinguent par le rôle important que joue l'ordre dans lequel elles ont été recueillies.

L'étude des séries chronologiques, ou séries temporelles, correspond à l'analyse statistique d'observations régulièrement espacées dans le temps. Elles ont été utilisées en :

- Astronomie ('on the periodicity of sunspots', 1906).
- En météorologie ('time-series regression of sealevel on weather', 1968).
- En théorie du signal ('Noise in FM receivers', 1963).
- En biologie ('the autocorrelation curves of schizophrenic brain waves and the power spectrum', 1960)
- En économie ('time-series analysis of imports, exports and other economic variables', 1971)...etc.

De fait, le recours à l'analyse en séries chronologiques peut sembler pertinent lorsque disposant d'un nombre de données suffisamment important l'on souhaite obtenir des prévisions à court terme sans investir en temps et en énergie dans la construction d'un modèle économétrique.

Chapitre 1

Processus stochastiques

L'objectif de la théorie des processus stochastique est l'étude des phénomènes aléatoires dépendant du temps. Soit $(\Omega, \mathcal{F}, \mathbb{P})$ un espace probabilisé, un ensemble T appelé ensemble des temps ($T = \mathbb{R}^+$ ou $[0, t]$).

Un processus stochastique $(X_t, t \in \mathbb{N}$ (ou \mathbb{Z})) est une famille de variables aléatoires de $(\Omega, \mathcal{F}, \mathbb{P})$ dans (E, \mathcal{E}) où (E, \mathcal{E}) est appelé l'espace des états du processus aléatoire.

1.1 Stationnarité

Pour travailler avec des données temporelles, elles doivent conserver une distribution constante dans le temps. C'est le concept de *stationnarité*.

Ainsi, si nos variables passées sont semblables à nos variables futures, on peut utiliser le passé pour tenter de prédire le futur. Un concept de stationnarité généralement utilisé est celui de la stationnarité de second ordre.

1.1.1 Définition (stationnarité en moyenne et en variance)

Considérons une suite de variables aléatoires $(X_t), t \in \mathbb{N}$ ou \mathbb{Z} . On dit que cette suite est stationnaire en moyenne si :

$$\mathbb{E}(X_t) = \mu_t = \mu, \forall t.$$

De même, cette suite sera stationnaire en variance si :

$$\text{Var}(X_t) = \mathbb{E}[(X_t - \mu_t)^2] = \sigma_t^2, \forall t.$$

La définition suivante caractérise les suites qui sont stationnaires en moyenne et dont la structure de covariance reste elle aussi constante.

1.1.2 Définition (stationnarité faible)

Un processus (X_t) , est stationnaire au second ordre (ou faiblement stationnaire) si et seulement si :

- (i) X_t est stationnaire en moyenne : $\mathbb{E}(X_t) = \mu, \forall t$;
- (ii) X_t est de carré intégrable pour tout t : $\mathbb{E}(X_t^2) < \infty$;
- (iii) X_t est stationnaire en covariance :

$$\text{Cov}(X_s, X_{s+t}) = \gamma(h), \forall t, h \in \mathbb{Z}$$

1.1.3 Définition (stationnarité forte)

Soit X un processus aléatoire indexé par $T = \mathbb{N}$ (ou \mathbb{Z}). On dit que X est strictement stationnaire si pour toute famille finie d'instants $(t_1, t_2, \dots, t_n \in T)$ et tout entier s , les lois de probabilité de $(X_{t_1}, X_{t_2}, \dots, X_{t_n})$ et de $(X_{t_1+s}, X_{t_2+s}, \dots, X_{t_n+s})$ sont les même c-à-d :

$$\mathcal{L}(X_{t_1}, X_{t_2}, \dots, X_{t_n}) = \mathcal{L}(X_{t_1+s}, X_{t_2+s}, \dots, X_{t_n+s}), \forall t \in T, \forall s \in \mathbb{N}$$

Exemple de processus stationnaire (bruit blanc) :

1.1.4 Définition (bruit blanc)

Soit $(\varepsilon_t)_{t \in T}$ un processus stochastique, on dit que $(\varepsilon_t)_{t \in T}$ est un processus stochastique ou bruit blanc faible (resp fort) si les trois propriétés sont vérifiées :

1. $\mathbb{E}(\varepsilon_t) = 0, \forall t \in T$
2. $\text{Var}(\varepsilon_t) = \sigma^2, \forall t \in T$
3. $\text{Cov}[\varepsilon_t, \varepsilon_s] = \mathbb{E}(\varepsilon_t, \varepsilon_s) = 0, \forall t \neq s$

La propriété 3 implique que les ε_t sont non corrélés entre eux (resp les ε_t sont iid)

Ce processus (bruit blanc) est un processus stationnaire d'ordre deux telle que toutes les variables sont de même moyenne nulle et de variance σ_ε^2 (constante finie) et non corrélées entre elles.

Exemple de processus non stationnaire (marche aléatoire)

Soit le processus (X_t) défini par :

$$X_t = \begin{cases} \varepsilon_t, & \text{si } t = 1 \\ X_{t-1} + \varepsilon_t, & \text{si } t = 2, 3, \dots \end{cases}$$

où ε_t est le processus bruit blanc, ce processus (marche aléatoire) peut s'écrire sous la forme :

$$X_t = \sum_{j=1}^t \varepsilon_j$$

$$\mathbb{E}(X_t) = 0, \quad \forall t = 1, 2, \dots$$

$$\begin{aligned} \text{Cov}[X_{t+h}, X_t] &= \text{Cov}\left(\sum_{s=1}^{t+h} \varepsilon_s, \sum_{n=1}^t \varepsilon_n\right) \\ &= \sum_{s=1}^{t+h} \sum_{n=1}^t \text{Cov}(\varepsilon_s, \varepsilon_n) \\ &= \sum_{\substack{s \\ t}} \sum_n \sigma_\varepsilon^2 \delta_{sn} \\ &= \sum_{s=1}^t \sigma_\varepsilon^2 = t\sigma_\varepsilon^2 \quad \forall h \end{aligned}$$

où le symbole δ est "le delta de Kronecker" :

$$\delta_{st} = \begin{cases} 1, & \text{si } t = s \\ 0, & \text{si } t \neq s \end{cases}$$

donc une marche aléatoire (Random Walk) n'est pas stationnaire car la suite des covariances dépend de t .

Remarque :

La stationnarité forte est plus exigeante que la stationnarité faible.

1.2 Fonction d'autocovariance

1.2.1 Définition

Soit $(X_t), t \in \mathbb{Z}$ un processus stationnaire. On appelle fonction d'autocovariance (ACV) la fonction γ définie de \mathbb{Z} dans \mathbb{R} par

$$\forall h, t \in \mathbb{Z}, \gamma(h) = \text{Cov}(X_t, X_{t+h}) \quad (1.1)$$

qui est indépendante de t .

1.2.2 Propriétés

La fonction d'autocovariance d'un processus stationnaire vérifie :

- i. $\forall h \in \mathbb{Z}, \gamma(-h) = \gamma(h)$
- ii. $\gamma(0) = \text{Var}(X_t)$
- iii. $|\gamma(h)| \leq \gamma(0), \forall h$

Démonstrations :

- i. $\gamma(-h) = \text{Cov}(X_{t-h}, X_t)$. On pose $s = t - h \Rightarrow t = s + h$, d'où :

$$\gamma(-h) = \text{Cov}(X_s, X_{s+h}) = \gamma(h)$$

- ii. Il suffit de remplacer h par 0 dans (1.1), on a :

$$\gamma(0) = \text{Cov}(X_t, X_t) = \text{Var}(X_t)$$

- iii. Il suffit de vérifier que $\gamma(h) \leq \gamma(0)$ et $\gamma(h) \geq -\gamma(0)$ on aura :

$$\begin{aligned} \mathbb{E}(X_t - X_{t+h})^2 &= \mathbb{E}(X_t^2) - 2\mathbb{E}(X_t X_{t+h}) + \mathbb{E}(X_{t+h}^2) \\ &= 2\gamma(0) - 2\gamma(h) \Rightarrow \end{aligned}$$

$$\gamma(0) = \frac{1}{2}\mathbb{E}(X_t - X_{t+h})^2 + \gamma(h) \Rightarrow \gamma(0) \geq \gamma(h)$$

on montre la deuxième inégalité en utilisant $\mathbb{E}(X_t + X_{t+h})^2$.

1.3 Fonction d'autocorrélation

1.3.1 Définition

On étudie la mémoire d'un processus en calculant sa fonction d'autocorrélation (ACF) de retard h noté $\rho(h)$:

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

qui mesure le lien entre les valeurs du processus à deux dates distante de h . Pour un processus stationnaire, $\rho(h)$ prend une forme plus simple :

$$\rho(h) = \frac{\text{Cov}(X_t, X_{t-h})}{\text{Var}(X_t)} = \frac{\gamma(h)}{\gamma(0)}$$

1.3.2 Propriétés

La fonction d'autocorrélation vérifie :

- i. $\rho(0) = 1$.
- ii. ρ est une fonction symétrique : $\rho(-h) = \rho(h)$; $\forall h \in \mathbb{N}$.
- iii. $|\rho(h)| \leq 1 \quad \forall h$.

1.4 Fonction d'autocorrélation partielle

La Fonction d'autocorrélation partielle (PACF) mesure la liaison entre X_t et X_{t-h} une fois retiré les liens transitant par les variables intermédiaires $X_{t-1}, X_{t-2}, \dots, X_{t-h+1}$, le coefficient d'autocorrélation partielle d'ordre h , noté ϕ_{hh} est le coefficient de corrélation entre :

$$X_t - \mathbb{E}[X_t/X_{t-1}, \dots, X_{t-h+1}] \quad \text{et} \quad X_{t-h} - \mathbb{E}[X_{t-h}/X_{t-1}, \dots, X_{t-h+1}]$$

on a donc :

$$\phi_{hh} = \text{corr} [X_t, X_{t-h}/X_{t-1}, \dots, X_{t-h+1}]$$

C'est donc le coefficient de X_{t-h} dans la régression de X_t sur $X_{t-1}, \dots, X_{t-h+1}$.

CHAPITRE 1. PROCESSUS STOCHASTIQUES

Si X_t est un processus stationnaire centré, la prédiction optimale de X_t est donnée par :

$$\mathbb{E} [X_t / X_{t-1}, \dots, X_{t-h+1}] = a_1 X_{t-1} + \dots + a_h X_{t-h}$$

que l'on peut réécrire matriciellement :

$$\begin{aligned} \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ \vdots \\ a_h \end{bmatrix} &= \begin{bmatrix} \gamma_0 & \gamma_1 & \dots & \gamma_{h-1} \\ \gamma_1 & \gamma_0 & \dots & \gamma_{h-2} \\ \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots \\ \gamma_{h-1} & \gamma_{h-2} & \dots & \gamma_0 \end{bmatrix}^{-1} \begin{bmatrix} \gamma_1 \\ \gamma_2 \\ \vdots \\ \vdots \\ \gamma_h \end{bmatrix} \\ &= \begin{bmatrix} 1 & \rho_1 & \dots & \rho_{h-1} \\ \rho_1 & 1 & \dots & \rho_{h-2} \\ \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots \\ \rho_{h-1} & \rho_{h-2} & \dots & \rho_0 \end{bmatrix}^{-1} \begin{bmatrix} \rho_1 \\ \rho_2 \\ \vdots \\ \vdots \\ \rho_h \end{bmatrix} \end{aligned}$$

Le coefficient d'autocorrélation partielle d'ordre h d'un processus stationnaire est alors a_h et se calcule de la manière suivante :

$$\phi_{hh} = \frac{|R(h)^*|}{|R(h)|}, \text{ avec :}$$

$$R(h) = \begin{bmatrix} 1 & \rho_1 & \dots & \rho_{h-1} \\ \rho_1 & 1 & \dots & \rho_{h-2} \\ \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots \\ \rho_{h-1} & \rho_{h-2} & \dots & \rho_0 \end{bmatrix}$$

et $R(h)^*$ est la matrice $R(h)$ dans laquelle on a remplacé la colonne h par :

$$\begin{bmatrix} \rho_1 \\ \rho_2 \\ \vdots \\ \vdots \\ \rho_h \end{bmatrix}$$

1.5 Causalité et inversibilité

1.5.1 Définition

Un processus X_t s'appelle causal s'il peut être représenté sous la forme :

$$X_t = \sum_{i=0}^{\infty} \Psi_i \varepsilon_{t-i}$$

où ε_t est un bruit blanc et $\sum \Psi_i^2 < \infty$.

1.5.2 Définition

Une représentation causale $X_t = \sum_{i=0}^{\infty} \Psi_i \varepsilon_{t-i}$ d'un processus stationnaire Y_t s'appelle inversible si on peut représenter le bruit blanc par une représentation causale c-à-d :

$$\varepsilon_t = \sum_{i=0}^{\infty} \pi_i X_{t-i} \text{ , où } \sum \pi_i^2 < \infty .$$

1.6 Théorème de décomposition de Wold

Le théorème de Wold (1938) est considéré comme un théorème fondamental dans le domaine des séries temporelles. En vertu de ce théorème, tout processus stochastique X_t stationnaire au sens faible peut être écrit comme une combinaison linéaire (dite encore filtre linéaire).

1.6.1 Théorème

Tout processus stationnaire d'ordre deux ($X_t, t \in \mathbb{Z}$) peut être représenté sous la forme :

$$X_t = \sum_{i=0}^{\infty} \Psi_i \varepsilon_{t-i} + k_t$$

où les paramètres Ψ_i satisfont $\Psi_0 = 1, \Psi_i \in \mathbb{R}, \forall i \in \mathbb{N}^*, \sum_{i=0}^{\infty} \Psi_i^2 < \infty$ et où $\varepsilon_t \sim iid(0, \sigma_\varepsilon^2)$. On dit que la somme des chocs passés correspond à la composante linéaire stochastique de X_t . Le terme k_t désigne la composante linéaire déterministe telle que $Cov(k_t, \varepsilon_{t-i}) = 0, \forall i \in \mathbb{Z}$.

Remarques :

1. L'implication forte de ce théorème est que si l'on connaît les pondérations $\Psi_i, \forall i \in \mathbb{N}$ et si l'on connaît la variance σ_ε^2 du bruit blanc, on est en mesure de proposer une représentation de n'importe quelle processus stationnaire.

2. La représentation de Wold est équivalente à la représentation moyenne mobile infinie $MA(\infty)$.

3. La condition $Cov(k_t, \varepsilon_{t-i}) = 0$, implique que le terme k_t est indépendant des chocs.

4. La représentation de Wold du processus $(X_t, t \in \mathbb{Z})$ suppose que l'on ajoute à la somme pondérée des chocs passés une composante déterministe k_t qui n'est autre que l'espérance du processus, car

$$\mathbb{E}(X_t) = \mathbb{E} \left[\sum_{i=0}^{\infty} \Psi_i \varepsilon_{t-i} \right] + k_t = \sum_{i=0}^{\infty} \Psi_i \mathbb{E}(\varepsilon_{t-i}) + k_t = k_t = m$$

5. Les conditions sur les pondérations :

5.1. Première condition : $\Psi_0 = 1$

Le poids du choc présent ε_t est unitaire, il s'agit ici tout simplement d'une condition de normalisation qui porte sur la détermination de la variance du bruit blanc.

5.2. Deuxième condition : $\Psi_i \in \mathbb{R}, \forall i \in \mathbb{N}^*$

Cette condition est triviale et signifie simplement que les pondérations des bruits blanc strictement passés peuvent être nulle pour certains retards.

5.3. Troisième condition : $\sum_{i=0}^{\infty} \Psi_i^2 < \infty$

Définition

La condition $\sum_{i=0}^{\infty} \Psi_i^2 < \infty$ est dite condition de sommabilité des carrés (ou d'intégrabilité des carrés).

Cette condition assure l'existence du moment d'ordre deux du processus $(X_t, t \in \mathbb{Z})$ sous cette condition on dit alors que X_t converge en moyenne quadratique.

Pour simplifier on suppose que $\mathbb{E}(X_t) = 0$

$$\gamma(h) = \mathbb{E}(X_{t+h} X_t) = \mathbb{E} \left[\sum_{i=0}^{\infty} \Psi_i \varepsilon_{t+h-i} \sum_{j=0}^{\infty} \Psi_j \varepsilon_{t-j} \right]$$

1.6. THÉORÈME DE DÉCOMPOSITION DE WOLD

$$\gamma(h) = \begin{cases} \sum_{i=0}^{\infty} \Psi_i^2 \mathbb{E}(\varepsilon_{t-i})^2 & \text{si } h = 0 \\ \sum_{i=0}^{\infty} \Psi_i \Psi_{i+h} \mathbb{E}(\varepsilon_{t-i}) & \text{si } h \neq 0 \end{cases}$$

$$\gamma(h) = \begin{cases} \sigma_\varepsilon^2 \sum_{i=0}^{\infty} \Psi_i^2 & \text{si } h = 0 \\ \sigma_\varepsilon^2 \sum_{i=0}^{\infty} \Psi_i \Psi_{i+h} & \text{si } h \neq 0 \end{cases}$$

On peut montrer que si $\sum_{i=0}^{\infty} \Psi_i^2 < \infty$ implique $\sum_{i=0}^{\infty} \Psi_i \Psi_{i+h} < \infty$ Dès lors l'existence de l'ensemble des moments d'ordre deux du processus X_t .

Chapitre 2

Modèle AutoRégressif Moyenne Mobile (ARMA)

Le but de ce chapitre est d'introduire la notion de *processus temporels* (séries temporelles) et plus particulièrement la classe des processus ARMA qui sont utiles pour décrire le comportement des séries temporelles univariées, parmi les représentations possibles pour les processus temporels figurent les représentations ARMA pour : *Auto Regressive Moving Average* cette représentation consiste en l'adjonction d'une composante *auto régressive* d'ordre fini (AR) et d'une composante *moyenne mobile* d'ordre fini (MA). Nous allons donc commencer par définir la classe des processus AR, MA et ARMA.

2.1 Les opérateurs

Pour étudier des processus stochastiques et donc des séries chronologiques, on définit des opérateurs retard et de différenciation.

2.1.1 Opérateur retard (l'opérateur B)

L'opérateur B décale le processus d'une unité de temps vers le passé :

$$B(X_t) = X_{t-1}.$$

Propriétés :

1. $B(B(\dots B(X_t)\dots)) = B^h(X_t) = X_{t-h}; \forall h \in \mathbb{Z}$,
En particulier, on a : $B^0(X_t) = X_t$.
2. Si $X_t = c; \forall t \in \mathbb{Z}$ avec $c \in \mathbb{R}$, $B^h(X_t) = B^h(c) = c$.
3. $B^h(B^k(X_t)) = B^{h+k}(X_t) = X_{t-h-k}; \forall (h, k) \in \mathbb{Z}^2$.
4. $B^{-h}(X_t) = X_{t+h}; \forall h \in \mathbb{Z}$.
5. $(B^h + B^k)(X_t) = B^h(X_t) + B^k(X_t)$
 $= X_{t-h} + X_{t-k}; \forall (h, k) \in \mathbb{Z}^2$
6. Si $|a| < 1$, $(1 - aB)^{-1}(X_t) = \frac{X_t}{1 - aB}$
 $= \lim_{h \rightarrow \infty} (1 + aB + a^2B^2 + \dots + a^hB^h)X_t$

2.1.2 Inversion des polynômes

Pour parler de l'*inverse* d'un polynôme de retard, il est commode dans un premier temps de considérer un polynôme particulier qui est le polynôme de retard de degré un défini par :

$$A(B) = 1 - \alpha B$$

Pour $|\alpha| < 1$ ce polynôme possède un inverse, c'est à dire que l'on peut définir :

$$A^{-1}(B) = \frac{1}{1 - \alpha B} = \sum_{i=0}^{\infty} \alpha^i B^i;$$

en utilisant l'expression de la somme d'une progression géométrique. Considérons maintenant le polynôme $A(B)$ de degré p que l'on note :

$$A(B) = 1 - \alpha_1 B - \alpha_2 B^2 - \dots - \alpha_p B^p$$

On définit l'équation caractéristique associée à ce polynôme comme l'expression en z :

$$A(z) = 1 - \alpha_1 z - \alpha_2 z^2 - \dots - \alpha_p z^p = 0.$$

Proposition

L'opérateur $A(B)$ est inversible si et seulement si les racines du polynôme sont de module différent de 1.

On a le théorème suivant :

2.1.3 Théorème

Le polynôme $A(B)$ est inversible si les p racines λ_j de son équation caractéristique associée sont toutes extérieures du cercle unité. Son inverse est donné par

$$A(B)^{-1} = \prod_{j=1}^p \left[\sum_{k=0}^{\infty} \left(\frac{1}{\lambda_j} \right)^k B^k \right] ;$$

la preuve de ce théorème utilise le résultat précédent en se ramenant à une série d'opérations élémentaires où l'on aurait à inverser que des polynômes de degré un.

Preuve :

Factoriser le polynôme $A(z)$ en utilisant les p racines de l'équation caractéristique :

$$A(z) = \prod_{j=1}^p (z - \lambda_j) \alpha_r$$

On peut remarquer que le produit des racines est égal à $1/\alpha_r$ car :

$$A(0) = 1 = \prod_{j=1}^p (-\lambda_j) \alpha_r$$

d'autre part on a la factorisation :

$$(z - \lambda_j) = -\lambda_j \left(1 - \frac{z}{\lambda_j} \right)$$

ce qui permet d'exprimer le polynôme en z sous la forme :

$$A(z) = \prod_{j=1}^p \left(1 - \frac{z}{\lambda_j} \right)$$

2.2. PROCESSUS AUTORÉGRESSIF (AR)

On peut alors ramener le calcul de l'inverse de $A(z)$ au calcul simple suivant, que l'on sait effectuer :

$$A^{-1}(z) = \prod_{j=1}^p \left(1 - \frac{z}{\lambda_j}\right)^{-1}$$

car :

$$\left(1 - \frac{z}{\lambda_j}\right)^{-1} = \sum_{k=0}^{\infty} \left(\frac{1}{\lambda_j}\right)^k z^k$$

Cet inverse existe si les racines λ_j de l'équation caractéristique sont toutes en dehors du cercle unité.

Remarque :

Les λ_j sont appelés les racines du polynôme $A(B)$.

2.1.4 Opérateur différence (l'opérateur Δ)

L'opérateur Δ fait la différence entre le processus et sa version décalée d'une unité de temps. Cet opérateur se construit en utilisant l'opérateur précédent :

$$\begin{aligned}\Delta(X_t) &= X_t - X_{t-1} = X_t - B(X_t) \\ \Leftrightarrow \Delta &= I - B\end{aligned}$$

où I est l'opérateur identité :

$$I(X_t) = X_t.$$

2.2 Processus AutoRégressif (AR)

Un processus autorégressif est un processus dont chaque valeur est décrite comme une combinaison linéaire des valeurs précédentes plus une composante aléatoire qui est le bruit blanc (choc). Le nombre de valeurs précédentes considérées est appelé *ordre* du processus.

Définition

Le processus $(X_t, t \in \mathbb{N} \text{ (ou } \mathbb{Z}))$ satisfait l'équation générale d'un processus AR d'ordre p si :

$$X_t = \delta + \sum_{j=1}^p \phi_j X_{t-j} + \varepsilon_t \quad (2.1)$$

où :

δ : le coefficient d'accroissement.

ϕ_j : les coefficients autorégressifs.

ε_t : le bruit blanc.

2.2.1 Propriétés statistiques du processus $AR(1)$

Définition

On dit que le processus $(X_t, t \in T)$ est de type $AR(1)$ i.e. autorégressif d'ordre 1 si l'on peut écrire :

$$\left\{ \begin{array}{l} X_t = \delta + \phi_1 X_{t-1} + \varepsilon_t \\ \text{où} \\ (\delta, \phi_1) \in \mathbb{R}^2 \\ \mathbb{E}[\varepsilon_t] = 0; V[\varepsilon_t] = \sigma_\varepsilon^2; \text{cov}[\varepsilon_t, \varepsilon_s] = 0 \quad \forall t \neq s \end{array} \right. \quad (2.2)$$

Le couple (δ, ϕ_1) représente les paramètres inconnus du modèle que l'on cherche à estimer.

La première étape dans l'analyse des propriétés statistiques du processus $AR(1)$ consiste à calculer sa moyenne, variance, covariance, fonction d'autocorrélation (ACF) et sa fonction d'autocorrélation partielle (PACF).

Dans un premier temps, on réécrit le processus $(X_t, t \in T)$ en supposant pour simplifier les calculs que $\delta = 0$.

$$X_t = \phi_1 X_{t-1} + \varepsilon_t \quad (2.3)$$

puis on fait des itérations arrières successives :

2.2. PROCESSUS AUTORÉGRESSIF (AR)

$$\begin{aligned}
 X_t &= \phi_1(\phi_1 X_{t-2} + \varepsilon_{t-1}) + \varepsilon_t \\
 X_t &= \varepsilon_t + \phi_1 \varepsilon_{t-1} + \phi_1^2(\phi_1 X_{t-3} + \varepsilon_{t-2}) \\
 X_t &= (\phi_1)^t X_0 + \varepsilon_t + \phi_1 \varepsilon_{t-1} + \dots + \phi_1^{t-1} \varepsilon_1 \\
 X_t &= (\phi_1)^t X_0 + \sum_{i=0}^{t-1} \phi_1^i \varepsilon_{t-i}
 \end{aligned} \tag{2.4}$$

On constate d'après la dernière égalité que les paramètres en ε_t décroissent lorsque i augmente à condition que $|\phi_1| < 1$.

Ainsi, si $|\phi| < 1$, alors le processus $(X_t)_{t \leq T}$ admet une décomposition de Wold avec $\psi_j = \phi_j$ et :

$$\sum_{j=1}^{\infty} \psi_j^2 = \sum_{j=1}^{\infty} \phi_j^2 = \frac{1}{1 - \phi^2} < \infty,$$

d'où, le processus $AR(1)$ est faiblement stationnaire.

Au contraire, le modèle a un comportement explosif dès que $|\phi_1| > 1$, lorsque $\phi_1 = 1$, le modèle s'écrit alors :

$$X_t = X_0 + \sum_{i=0}^{t-1} \varepsilon_{t-i} \tag{2.5}$$

Dans ce cas, on dit que le modèle est non stationnaire de type stochastique dans le sens où un choc exogène de grand ampleur à un instant donné du processus (ε_t grand) va faire diverger X_t de sa trajectoire d'origine sans que par la suite le processus puisse revenir sur sa trajectoire d'origine. On dit que le modèle se comporte comme une marche aléatoire ou encore qu'il possède une racine unitaire. Compte tenu des implications en termes de modélisation d'un tel cas. Par la suite, on supposera toujours que $|\phi_1| < 1$ (cette hypothèse doit être testée sur les données d'observation), i.e. que le modèle est stationnaire.

On peut également écrire le processus $AR(1)$ en utilisant l'opérateur retard B :

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

CHAPITRE 2. MODÈLE AUTORÉGRESSIF MOYENNE MOBILE
(ARMA)

A condition que le polynôme $(1 - \phi_1 B)$ soit inversible ($\phi_1 \neq 1$), on retrouve l'expression précédente :

$$X_t = (1 - \phi_1 B)^{-1} \varepsilon_t = \sum_{i=0}^{t-1} \phi_1^i \varepsilon_{t-i}$$

Le processus $(X_t, t \in T)$ s'écrit donc comme une somme de termes de bruit blanc. On en déduit ses deux premiers moments (on suppose pour simplifier les calculs que $X_0 = 0$) :

• La moyenne s'écrit :

$$\mathbb{E}[X_t] = \mathbb{E} \left[\sum_{i=0}^{t-1} \phi_1^i \varepsilon_{t-i} \right] = 0 ; \quad \text{car } \mathbb{E}[\varepsilon_t] = 0 \quad \forall t.$$

• La variance s'écrit :

$$\begin{aligned} V[X_t] &= \mathbb{E}[(X_t - \mathbb{E}[X_t])^2] = \mathbb{E}[(X_t)^2] \\ V[X_t] &= \mathbb{E} \left[\left(\sum_{i=0}^{t-1} \phi_1^i \varepsilon_{t-i} \right)^2 \right] = \sigma_\varepsilon^2 (1 + \phi_1^2 + \dots + \phi_1^{2t-2}) \end{aligned}$$

car, on a supposé que $\text{cov}[\varepsilon_t, \varepsilon_s] = 0 \quad \forall t \neq s$. D'où l'on déduit :

$$\begin{cases} V[X_t] = \sigma_\varepsilon^2 \frac{1 - \phi_1^{2t}}{1 - \phi_1^2} & \text{si } (\phi_1 \neq 1) \\ V[X_t] = \sigma_\varepsilon^2 t & \text{si } (\phi_1 = 1) \end{cases}$$

• La covariance s'écrit :

$$\begin{aligned} \text{cov}[X_t, X_s] &= \mathbb{E}[(X_t - \mathbb{E}[X_t])(X_s - \mathbb{E}[X_s])] = \mathbb{E}[(X_t)(X_s)] ; \quad \text{car } \mathbb{E}[X_t] = 0 \quad \forall t \\ \text{cov}[X_t, X_s] &= \mathbb{E} \left[\left(\sum_{i=0}^{t-1} \phi_1^i \varepsilon_{t-i} \right) \left(\sum_{j=0}^{s-1} \phi_1^j \varepsilon_{s-j} \right) \right] \end{aligned}$$

Il vient :

$$\begin{cases} \text{cov}[X_t, X_s] = \sigma_\varepsilon^2 \phi_1^s \left(\frac{1 - \phi_1^{2t}}{1 - \phi_1^2} \right) & \text{si } (\phi_1 \neq 1) \\ \text{cov}[X_t, X_s] = \sigma_\varepsilon^2 t & \text{si } (\phi_1 = 1) \end{cases}$$

2.2. PROCESSUS AUTORÉGRESSIF (AR)

En conclusion de l'étude des deux premiers moments du processus $(X_t, t \in T)$, il apparaît que $|\phi_1| < 1$ est une condition indispensable pour que ce processus soit stationnaire au sens large. Si cette condition est vérifiée, on a en effet la variance et la covariance du processus qui sont indépendantes du temps :

$$\begin{cases} \lim_{t \rightarrow +\infty} V[X_t] = \frac{\sigma_\varepsilon^2}{1 - \phi_1^2} = \sigma_X^2 \\ \lim_{t \rightarrow +\infty} cov[X_t, X_s] = \left(\frac{\sigma_\varepsilon^2 \phi_1^s}{1 - \phi_1^2} \right) = \sigma_X^2 \phi_1^s \end{cases}$$

• Le coefficient d'autocorrélation s'écrit :

$$\rho_k = corr[X_t, X_{t-k}] = \frac{cov[X_t, X_{t-k}]}{\sqrt{V[X_t]V[X_{t-k}]}}$$

• La fonction de covariance étant symétrique et sachant de plus que $\rho_0 = 1$, on aura :

$$\rho_k = \frac{\sigma_X^2 \phi_1^k}{\sqrt{\sigma_X^2 \sigma_X^2}} = \phi_1^k$$

• Le coefficient d'autocorrélation partielle s'écrit :

$$\begin{cases} \phi_{11} = \rho_1 = \phi_1 \\ \phi_{kk} = \frac{\rho_k - \sum_{j=1}^{k-1} \phi_{k-1,j} \rho_{k-j}}{1 - \sum_{j=1}^{k-1} \phi_{k-1,j} \rho_j} & k=2,3,\dots \\ \phi_{kj} = \phi_{k-1,j} - \phi_{k-1,k-j} & k=2,3,\dots \quad j=1,\dots,k-1 \end{cases}$$

On voit alors que,

$$\phi_{22} = \frac{\rho_2 - \phi_{1,1}\rho_1}{1 - \phi_{1,1}\rho_1} = \frac{\phi_1^2 - \phi_1\phi_1}{1 - \phi_1\phi_1} = 0$$

On en déduit aisément pour un processus AR(1) :

$$\phi_{kk} = 0 \quad k > 1$$

2.2.2 Propriétés statistiques du processus $AR(p)$

Définition

On dit que le processus $(X_t, t \in T)$ est de type $AR(p)$ i.e, *AutoRégressif* d'ordre p si l'on peut écrire :

$$\left\{ \begin{array}{l} X_t = \delta + \phi_1 X_{t-1} + \phi_2 X_{t-2} + \dots + \phi_p X_{t-p} + \varepsilon_t \\ \text{où} \\ (\delta, \phi_1, \phi_2, \dots, \phi_p) \in \mathbb{R}^{p+1} \\ \mathbb{E}[\varepsilon_t] = 0; V[\varepsilon_t] = \sigma_\varepsilon^2; \text{cov}[\varepsilon_t, \varepsilon_s] = 0 \quad \forall t \neq s \end{array} \right. \quad (2.6)$$

Afin de simplifier les calculs, on suppose de même que pour le processus $AR(1)$, que $X_0 = 0$ et $\delta = 0$. On réécrit le processus $AR(p)$ à l'aide du polynôme retard :

$$X_t - \phi_1 X_{t-1} - \phi_2 X_{t-2} - \dots - \phi_p X_{t-p} = (1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p) X_t = \varepsilon_t$$

Le polynôme retard a donc pour expression :

$$\Phi(B) = (1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p)$$

L'équation caractéristique associée s'écrit :

$$1 - \phi_1 z - \phi_2 z^2 - \dots - \phi_p z^p = 0$$

Le polynôme retard est inversible si les racines de l'équation caractéristique sont toutes de module supérieur à l'unité. Dans ce cas , on peut écrire :

$$X_t = \Phi^{-1}(B)\varepsilon_t \quad (2.7)$$

On en déduit ses deux premiers moments :

• La moyenne s'écrit :

$$\mathbb{E}[X_t] = \mathbb{E}[\Phi^{-1}(B)\varepsilon_t] = 0; \quad \text{car } \mathbb{E}[\varepsilon_t] = 0 \quad \forall t$$

2.2. PROCESSUS AUTORÉGRESSIF (AR)

• Afin de calculer *la variance*, on multiplie chaque terme du processus par X_t :

$$X_t X_t = \phi_1 X_{t-1} X_t + \phi_2 X_{t-2} X_t + \dots + \phi_p X_{t-p} X_t + \varepsilon_t X_t$$

puis on prend l'espérance de chacun des termes de l'égalité :

$$\mathbb{E}[X_t X_t] = \mathbb{E}[\phi_1 X_{t-1} X_t] + \mathbb{E}[\phi_2 X_{t-2} X_t] + \dots + \mathbb{E}[\phi_p X_{t-p} X_t] + \mathbb{E}[\varepsilon_t X_t]$$

Puisque $\mathbb{E}[\varepsilon_t X_t] = \mathbb{E}[\varepsilon_t(\phi_1 X_{t-1} + \phi_2 X_{t-2} + \dots + \phi_p X_{t-p} + \varepsilon_t)] = \sigma_\varepsilon^2$, l'expression s'écrit alors :

$$V[X_t] = \phi_1 \text{cov}[X_{t-1}, X_t] + \phi_2 \text{cov}[X_{t-2}, X_t] + \dots + \phi_p \text{cov}[X_{t-p}, X_t] + \sigma_\varepsilon^2$$

Soit par convention d'écriture :

$$\gamma_0 = \phi_1 \gamma_1 + \phi_2 \gamma_2 + \dots + \phi_p \gamma_p + \sigma_\varepsilon^2$$

• On procède d'une manière similaire pour calculer *la covariance* à l'ordre k , à savoir que l'on multiplie chaque terme de l'égalité (2.6) par X_{t-k} :

$$X_t X_{t-k} = \phi_1 X_{t-1} X_{t-k} + \phi_2 X_{t-2} X_{t-k} + \dots + \phi_p X_{t-p} X_{t-k} + \varepsilon_t X_{t-k}$$

Puis on prend l'espérance de chacun des termes de l'égalité ,

$$\mathbb{E}[X_t X_{t-k}] = \mathbb{E}[\phi_1 X_{t-1} X_{t-k}] + \mathbb{E}[\phi_2 X_{t-2} X_{t-k}] + \dots + \mathbb{E}[\phi_p X_{t-p} X_{t-k}] + \mathbb{E}[\varepsilon_t X_{t-k}]$$

Finalement on obtient :

$$\text{cov}[X_t, X_{t-k}] = \phi_1 \text{cov}[X_{t-1}, X_{t-k}] + \phi_2 \text{cov}[X_{t-2}, X_{t-k}] + \dots + \phi_p \text{cov}[X_{t-p}, X_{t-k}]$$

Soit par convention d'écriture :

$$\gamma_k = \phi_1 \gamma_{k-1} + \phi_2 \gamma_{k-2} + \dots + \phi_p \gamma_{k-p}$$

•L'écriture de l'ACF est alors immédiate, puisqu'il suffit de diviser chaque terme de covariance par la variance de processus :

$$\rho_k = \frac{\gamma_k}{\gamma_0} = \phi_1 \rho_{k-1} + \phi_2 \rho_{k-2} + \dots + \phi_p \rho_{k-p}$$

2.3 Processus Moyenne Mobile (MA)

L'idée de modéliser un processus aléatoire temporel $(X_t, t \in T)$ par un processus de moyenne mobile i.e. par une moyenne pondérée d'erreurs aléatoires présentes et passées vient de l'étude des marchés financiers. En effet, on a constaté que les variations d'une période à l'autre de prix d'actifs (actions, obligations), de taux de change ou de taux d'intérêt se comportaient comme une série de variables aléatoires non corrélées. Il apparait alors que la composante aléatoire reflète les nouvelles informations non anticipées, l'émergence de nouveaux concurrents sur le marché ou bien encore l'annonce de découvertes technologiques.

Définition :

Le processus $(X_t, t \in \mathbb{N} \text{ (ou } \mathbb{Z}))$ satisfait l'équation générale d'un processus MA d'ordre q :

$$X_t = \delta + \sum_{j=1}^q \theta_j \varepsilon_{t-j} + \varepsilon_t \quad (2.8)$$

où

θ_j : Les coefficients moyenne mobile.

ε_{t-j} : Les chocs ou le processus purement aléatoire.

Comme pour le processus AR, les coefficients du processus MA doivent faire l'objet d'une optimisation avec toujours comme indicateur, la minimisation d'erreurs.

Le processus moyenne mobile d'un ordre infini a déjà eu lieu lorsque nous avons présenté le théorème de décomposition de Wold. Il est, avant tout, d'importance théorique dans la pratique, seul un nombre fini de paramètres peut être estimé. Dans la suite, on considère un ordre finie du processus moyenne mobile.

2.3.1 Propriétés statistiques du processus MA (1)

Définition :

On dit que le processus $(X_t, t \in T)$ est de type MA(1) i.e. *Moving Average* (moyenne mobile) d'ordre 1 si l'on peut écrire :

$$\left\{ \begin{array}{l} X_t = \delta + \varepsilon_t + \theta_1 \varepsilon_{t-1} \\ \text{où} \\ (\delta, \theta_1) \in \mathbb{R}^2 \\ \mathbb{E}[\varepsilon_t] = 0; V[\varepsilon_t] = \sigma_\varepsilon^2; \text{cov}[\varepsilon_t, \varepsilon_s] = 0 \quad \forall t \neq s \end{array} \right. \quad (2.9)$$

Ecrire $(X_t, t \in T)$ sous la forme d'un MA(1) signifie que toute l'information apportée par l'erreur à la date $t - 1$ n'a pu être prise en compte dans la v.a X en $t - 1$.

Comme pour le processus AR(1), on caractérise le processus MA(1) à travers ses propriétés statistiques de moyenne, variance, covariance, ACF et PACF.

•La moyenne s'écrit :

$$\begin{aligned} \mathbb{E}[X_t] &= \mathbb{E}[\delta + \varepsilon_t + \theta_1 \varepsilon_{t-1}] \\ &= \delta + \mathbb{E}[\varepsilon_t] + \mathbb{E}[\theta_1 \varepsilon_{t-1}] \\ &= \delta; \text{ car } \mathbb{E}[\varepsilon_t] = 0 \quad \forall t \end{aligned}$$

$(X_t, t \in T)$ est donc stationnaire à l'ordre 1 puisque la moyenne est constante sur toutes les observations.

•La variance s'écrit :

$$\begin{aligned} V[X_t] &= \mathbb{E}[(X_t - \mathbb{E}[X_t])^2] = \mathbb{E}[(X_t - \delta)^2] \\ &= \mathbb{E}[(\varepsilon_t + \theta_1 \varepsilon_{t-1})^2] = \mathbb{E}[(\varepsilon_t^2 + 2\theta_1 \varepsilon_t \varepsilon_{t-1} + \theta_1^2 \varepsilon_{t-1}^2)] \\ &= \mathbb{E}[\varepsilon_t^2] + 2\theta_1 \mathbb{E}[\varepsilon_t \varepsilon_{t-1}] + \theta_1^2 \mathbb{E}[\varepsilon_{t-1}^2] \\ &= (1 + \theta_1^2) \sigma_\varepsilon^2; \text{ car } \text{cov}(\varepsilon_t, \varepsilon_s) = 0 \quad \forall t \neq s \end{aligned}$$

On constate que la variance est indépendante du temps.

•La covariance s'écrit :

$$\begin{aligned} \text{cov}[X_t, X_s] &= \mathbb{E}[(X_t - \mathbb{E}[X_t])(X_s - \mathbb{E}[X_s])] = \mathbb{E}[(X_t - \delta)(X_s - \delta)] \\ &= \mathbb{E}[(\varepsilon_t + \theta_1 \varepsilon_{t-1})(\varepsilon_s + \theta_1 \varepsilon_{s-1})] \\ &= \mathbb{E}[\varepsilon_t \varepsilon_s + \theta_1 \varepsilon_t \varepsilon_{s-1} + \theta_1 \varepsilon_{t-1} \varepsilon_s + \theta_1^2 \varepsilon_{t-1} \varepsilon_{s-1}] \end{aligned}$$

CHAPITRE 2. MODÈLE AUTORÉGRESSIF MOYENNE MOBILE
(ARMA)

Il vient d'après les hypothèses posées sur ε_t :

$$\begin{cases} \text{cov}[X_t, X_{t-1}] = \theta_1 \sigma_\varepsilon^2 \\ \text{cov}[X_t, X_s] = 0 \end{cases} \quad \text{pour tout } t \neq s$$

En conclusion de l'étude des deux premiers moments du processus $(X_t, t \in T)$, il apparaît que le processus $MA(1)$ est stationnaire à l'ordre deux quelques soient les valeurs de θ_1 .

• Le coefficient d'autocorrélation s'écrit :

$$\rho_k = \text{coor}(X_t, X_{t-k}) = \frac{\text{cov}(X_t, X_{t-k})}{\sqrt{V(X_t)V(X_{t-k})}} \quad k = 0, \pm 1, \pm 2, \dots$$

La fonction de covariance étant symétrique et sachant de plus que $\rho_0 = 1$, on aura :

$$\begin{cases} \rho_1 = \frac{\theta_1 \sigma_\varepsilon^2}{\sqrt{(1 + \theta_1^2) \sigma_\varepsilon^2 (1 + \theta_1^2) \sigma_\varepsilon^2}} = \frac{\theta_1}{(1 + \theta_1^2)} \\ \rho_k = 0 \end{cases} \quad \text{pour tout } k > 1$$

• Le coefficient d'autocorrélation partielle s'écrit (algorithme de Durbin) :

$$\begin{cases} \phi_{11} = \rho_1 = \frac{\theta_1}{(1 + \theta_1^2)} \\ \phi_{22} = \frac{\rho_2 - \phi_{11}\rho_1}{1 - \phi_{11}\rho_1} = \frac{-\rho_1^2}{1 - \rho_1^2} = \frac{-\theta_1^2}{1 + \theta_1^2 + \theta_1^4} \end{cases}$$

On voit que $1 + \theta_1^2 + \theta_1^4$ est une suite géométrique de premier terme 1 et de raison θ_1^2 d'où l'on déduit :

$$1 + \theta_1^2 + \theta_1^4 = \frac{1 - \theta_1^6}{1 - \theta_1^2} \implies \phi_{22} = \frac{-\theta_1^2(1 - \theta_1^2)}{1 - \theta_1^6}$$

2.3. PROCESSUS MOYENNE MOBILE (MA)

On procède selon le même raisonnement avec les autres termes :

$$\begin{aligned}\phi_{33} &= \frac{\rho_3 - \phi_{21}\rho_2 - \phi_{22}\rho_1}{1 - \phi_{21}\rho_1 - \phi_{22}\rho_2} = \frac{-\phi_{22}\rho_1}{1 - \phi_{21}\rho_1} \\ &= \frac{\rho_1^3}{(1 - \rho_1^2)(1 - \phi_{21}\rho_1)} = \frac{\rho_1^3}{1 - 2\rho_1^2} = \frac{\theta_1^3(1 - \theta_1^2)}{1 - \theta_1^8}\end{aligned}$$

Il apparaît la généralisation suivante :

$$\phi_{kk} = \frac{-\theta_1^k(1 - \theta_1^2)}{1 - \theta_1^{2(k+1)}}$$

En conclusion, il apparaît que l'on peut caractériser le comportement d'un processus $MA(1)$ à travers ses fonctions ACF et PACF :

- La fonction ACF s'annule pour tout retard $k > 1$.
- La fonction PACF d'éroît vers 0 rapidement si θ_1 est proche de 0 ou lentement si θ_1 est proche de 1. Son allure est de type exponentiel si $\theta_1 > 0$.

Définition

Soit $(X_t, t \in T)$ un processus aléatoire temporel de forme MA .

On dit que ce processus est inversible, si l'on peut le représenter sous forme d'un processus AR .

La condition d'inversibilité est que la racine de l'équation caractéristique $1 + \phi_1 z = 0$ soit en dehors du cercle unité i.e. $|\theta_1| < 1$. Lorsque le processus $MA(1)$ possède une racine unitaire ($|\theta_1| = 1$), alors le processus MA n'est pas inversible.

2.3.2 Propriétés statistiques du processus $MA(q)$

Définition :

On dit que le processus $(X_t, t \in T)$ est de type $MA(q)$ i.e. *Moving Average* (moyenne mobile) d'ordre q si l'on peut écrire :

$$\left\{ \begin{array}{l} X_t = \delta + \varepsilon_t + \theta_1\varepsilon_{t-1} + \theta_2\varepsilon_{t-2} + \dots + \theta_q\varepsilon_{t-q} \\ \text{où} \\ (\delta, \theta_1, \theta_2, \dots, \theta_q) \in \mathbb{R}^{q+1} \\ \mathbb{E}[\varepsilon_t] = 0; V[\varepsilon_t] = \sigma_\varepsilon^2; \text{cov}[\varepsilon_t, \varepsilon_s] = 0 \quad \forall t \neq s \end{array} \right. \quad (2.10)$$

CHAPITRE 2. MODÈLE AUTORÉGRESSIF MOYENNE MOBILE
(ARMA)

Comme pour le processus $MA(1)$, la condition d'inversibilité du $MA(q)$ consiste à écrire que les racines de l'équation caractéristique :

$$1 + \theta_1 z + \dots + \theta_q z^q = 0$$

sont en dehors du cercle unité.

On caractérise les propriétés statistiques du processus $MA(q)$ en étudiant ses deux premiers moments :

• La moyenne s'écrit :

$$\mathbb{E}[X_t] = \mathbb{E}[\delta + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \dots + \theta_q \varepsilon_{t-q}] = \delta; \text{ car } \mathbb{E}[\varepsilon_t] = 0 \quad \forall t$$

$(X_t, t \in T)$ est donc stationnaire à l'ordre 1 puisque la moyenne est constante sur toutes les observations.

• La variance s'écrit :

$$\begin{aligned} V[X_t] &= \mathbb{E}[(X_t - \mathbb{E}[X_t])^2] = \mathbb{E}[(X_t - \delta)^2] \\ &= \mathbb{E}[(\varepsilon_t + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \dots + \theta_q \varepsilon_{t-q})^2] \\ &= \mathbb{E}[\varepsilon_t^2] + \theta_1^2 \mathbb{E}[\varepsilon_{t-1}^2] + \dots + \theta_q^2 \mathbb{E}[\varepsilon_{t-q}^2] \text{ car } \text{cov}[\varepsilon_t, \varepsilon_s] = 0 \quad \forall t \neq s \\ &= (1 + \theta_1^2 + \dots + \theta_q^2) \sigma_\varepsilon^2; \text{ car } \text{cov}[\varepsilon_t, \varepsilon_s] = 0 \quad \forall t \neq s \end{aligned}$$

On constate que la variance est indépendante du temps.

• La covariance s'écrit :

$$\begin{aligned} \text{cov}[X_t, X_s] &= \mathbb{E}[(X_t - \mathbb{E}[X_t])(X_s - \mathbb{E}[X_s])] = \mathbb{E}[(X_t - \delta)(X_s - \delta)] \\ &= \mathbb{E}[(\varepsilon_t + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \dots + \theta_q \varepsilon_{t-q}) \\ &\quad (\varepsilon_s + \theta_1 \varepsilon_{s-1} + \theta_2 \varepsilon_{s-2} + \dots + \theta_q \varepsilon_{s-q})] \\ &= \mathbb{E}[(\varepsilon_t \varepsilon_s + \theta_1 \varepsilon_t \varepsilon_{t-1} + \theta_1 \varepsilon_{t-1} \varepsilon_s + \dots + \theta_q^2 \varepsilon_{t-q} \varepsilon_{s-q})] \end{aligned}$$

Il vient d'après les hypothèses posées sur ε_t :

$$\begin{cases} \text{cov}[X_t, X_s] = (\theta_1 + \theta_1 \phi_{s+1} + \dots + \phi_{q-s} \phi_q) \sigma_\varepsilon^2 & \text{pour } t = s \\ \text{cov}[X_t, X_s] = 0 & \text{pour } t \neq s \end{cases}$$

En conclusion de l'étude des deux premiers moments du processus $(X_t, t \in T)$, il apparaît que le processus $MA(q)$ est stationnaire à l'ordre deux quelques soient les valeurs de ϕ_1 .

2.4. MODÉLISATION STATISTIQUE DES PROCESSUS (ARMA)

• Le coefficient d'autocorrélation s'écrit :

$$\rho_k = \text{corr}(X_t, X_{t-k}) = \frac{\text{cov}(X_t, X_{t-k})}{\sqrt{V(X_t)V(X_{t-k})}} \quad k = 0, \pm 1, \pm 2, \dots$$

La fonction de covariance étant symétrique et sachant de plus que $\rho_0 = 1$, on aura :

$$\begin{cases} \rho_1 = \frac{\phi_1 \sigma_\varepsilon^2}{\sqrt{(1 + \phi_1^2) \sigma_\varepsilon^2 (1 + \phi_1^2) \sigma_\varepsilon^2}} = \frac{\phi_1}{(1 + \phi_1^2)} \\ \rho_k = 0 \end{cases} \quad \text{pour tout } k > 1$$

2.4 Modélisation statistique des processus (ARMA)

Le modèle de régression le plus courant est le modèle *ARMA* qui combine simplement les deux principes *AR* et *MA*, dont l'équation générale peut s'écrire pour un *ARMA*(p, q)

$$X_t = \delta + \sum_{j=1}^p \phi_j X_{t-j} + \sum_{j=1}^q \theta_j \varepsilon_{t-j} \quad (2.11)$$

Dans ce cas on définit p comme l'ordre de processus autorégressif et q l'ordre de processus moyenne mobile. Ce modèle implique que la série traitée soit stationnaire au sens faible. En pratique, on se contentera d'une stationnarité à l'ordre deux localement, cela signifie qu'il est suffisant que la suite soit stationnaire sur un nombre de points au moins équivalents à l'ordre de processus autorégressif ou de processus moyenne mobile, parfois appelé Modèles de Box-Jenkins. Etant donné une série chronologique X_t , le modèle *ARMA* est un outil pour l'arrangement et, peut-être, la prévision des valeurs futures de cette série.

2.4.1 Propriétés statistiques du processus ARMA(1,1) :

Définition

On dit que le processus aléatoire temporel ($X_t, t \in T$) est de type *ARMA*(1, 1) i.e. *Auto-Regressive Moving Average* (auto-régressif moyenne mobile) d'ordre (1,1) si l'on peut écrire :

$$\left\{ \begin{array}{l} X_t = \delta + \phi_1 X_{t-1} + \theta_1 \varepsilon_{t-1} + \varepsilon_t \\ \text{où} \\ (\delta, \phi_1, \theta_1) \in \mathbb{R}^3 \\ \mathbb{E}[\varepsilon_t] = 0 \quad \forall t; \quad V[\varepsilon_t] = \sigma_\varepsilon^2 \quad \forall t; \quad \text{cov}[\varepsilon_t, \varepsilon_s] = 0 \quad \forall t \neq s \end{array} \right. \quad (2.12)$$

Par la suite, on supposera $\delta = 0$. On peut également écrire le processus ARMA à partir des polynômes retards :

$$(1 - \phi_1 B)X_t = (1 - \theta_1 B)\varepsilon_t \iff \Phi(B)X_t = \Theta(B)\varepsilon_t \quad (2.13)$$

On peut noter que le processus est stationnaire et inversible si les racines des polynômes respectivement Φ et Θ sont à l'extérieur du cercle unité (i.e. si $|\phi_1| < 1$ et si $|\theta_1| < 1$). Par la suite, on suppose ces deux conditions vérifiées.

On caractérise les propriétés statistiques du processus ARMA en étudiant ses deux premiers moments :

• La moyenne s'écrit :

$$\mathbb{E}[X_t] = \mathbb{E}[\phi_1 X_{t-1} + \theta_1 \varepsilon_{t-1} + \varepsilon_t] = 0; \text{ car } \mathbb{E}[\varepsilon_t] = 0 \quad \forall t$$

• Pour calculer la variance on procède de la même manière que dans le cas de l'AR(p). On commence par multiplier chaque terme par X_t :

$$X_t X_t = \phi_1 X_{t-1} X_t + \theta_1 \varepsilon_{t-1} X_t + \varepsilon_t X_t$$

puis on prend l'espérance de chacun des termes de l'égalité :

$$\mathbb{E}[X_t X_t] = \mathbb{E}[\phi_1 X_{t-1} X_t] + \mathbb{E}[\theta_1 \varepsilon_{t-1} X_t] + \mathbb{E}[\varepsilon_t X_t]$$

l'expression s'écrit alors :

$$V[X_t] = \phi_1 \text{cov}[X_{t-1}, X_t] + \mathbb{E}[\varepsilon_t(\phi_1 X_{t-1} + \theta_1 \varepsilon_{t-1} + \varepsilon_t)] + \theta_1 \mathbb{E}[\varepsilon_{t-1}(\phi_1 X_{t-1} + \theta_1 \varepsilon_{t-1} + \varepsilon_t)]$$

Soit encore :

$$\gamma_0 = \phi_1 \gamma_1 + \sigma_\varepsilon^2 + \theta_1 \mathbb{E}[\varepsilon_{t-1}(\phi_1(\phi_1 X_{t-2} + \varepsilon_{t-1} + \theta_1 \varepsilon_{t-2}) + \varepsilon_t + \theta_1 \varepsilon_{t-1})]$$

D'où,

$$\gamma_0 = \phi_1 \gamma_1 + \sigma_\varepsilon^2 + \theta_1 \phi_1 \sigma_\varepsilon^2 + \theta_1^2 \sigma_\varepsilon^2$$

2.4. MODÉLISATION STATISTIQUE DES PROCESSUS (ARMA)

Finalement,

$$\gamma_0 = \phi_1 \gamma_1 + [1 + \theta_1(\theta_1 + \phi_1)] \sigma_\varepsilon^2$$

On procède d'une manière similaire pour calculer la covariance à l'ordre 1, à savoir que l'on multiplie chaque terme de l'égalité (2.10) par X_{t-1}

$$X_t X_{t-1} = \phi_1 X_{t-1} X_{t-1} + \varepsilon_t X_{t-1} + \theta_1 \varepsilon_{t-1} X_{t-1}$$

Puis on prend l'espérance de chaque terme de l'égalité,

$$\mathbb{E}[X_t X_{t-1}] = \mathbb{E}[\phi_1 X_{t-1} X_{t-1}] + \mathbb{E}[\varepsilon_t X_{t-1}] + \mathbb{E}[\theta_1 \varepsilon_{t-1} X_{t-1}]$$

Finalement après développement du membre droit de l'égalité, on obtient :

$$\text{cov}[X_t, X_{t-1}] = \phi_1 \gamma_0 + 0 + \mathbb{E}[\theta_1 \varepsilon_{t-1} (\phi_1 X_{t-2} - \varepsilon_{t-1} + \theta_1 \varepsilon_{t-2})]$$

Soit par convention d'écriture :

$$\gamma_1 = \phi_1 \gamma_0 + \theta_1 \sigma_\varepsilon^2$$

• La covariance à l'ordre 2 s'écrit :

$$X_t X_{t-2} = \phi_1 X_{t-1} X_{t-2} + \varepsilon_t X_{t-2} + \theta_1 \varepsilon_{t-1} X_{t-2}$$

D'où en prenant l'espérance de chaque terme de l'égalité :

$$\mathbb{E}[X_t X_{t-2}] = \mathbb{E}[\phi_1 X_{t-1} X_{t-2}] + \mathbb{E}[\varepsilon_t X_{t-2}] + \mathbb{E}[\theta_1 \varepsilon_{t-1} X_{t-2}]$$

Soit :

$$\gamma_2 = \phi_1 \gamma_1$$

• La covariance à l'ordre k s'écrit :

$$\gamma_k = \phi_1 \gamma_{k-1} \quad \text{pour} \quad k > 1$$

• Le coefficient d'autocorrélation s'écrit :

$$\begin{cases} \rho_1 = \frac{\gamma_1}{\gamma_0} = \frac{(1 + \phi_1\theta_1)(\phi_1 + \theta_1)}{1 + 2\phi_1\theta_1 + \theta_1^2} \\ \rho_k = \frac{\gamma_k}{\gamma_0} = \frac{\phi_1\gamma_{k-1}}{\gamma_0} = \phi_1\rho_{k-1} \quad k > 1 \end{cases}$$

2.4.2 Propriétés statistiques du processus ARMA (p,q) :

Définition :

On dit que le processus $(X_t, t \in T)$ est de type $ARMA(p, q)$ i.e. *Autoregressive Moving Average* (auto-régressif moyenne mobile) d'ordre (p,q) si l'on peut écrire :

$$\begin{cases} X_t = \delta + \phi_1 X_{t-1} + \dots + \phi_p X_{t-p} + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \dots + \theta_q \varepsilon_{t-q} \\ \text{où} \\ (\delta, \phi_1, \dots, \phi_p, \theta_1, \dots, \theta_q) \in \mathbb{R}^{p+q+1} \\ \mathbb{E}[\varepsilon_t] = 0 \quad \forall t; \quad V[\varepsilon_t] = \sigma_\varepsilon^2 \quad \forall t; \quad cov[\varepsilon_t, \varepsilon_s] = 0 \quad \forall t \neq s \end{cases} \quad (2.14)$$

Par la suite, on supposera que $\delta = 0$. On peut également écrire le processus $ARMA$ à partir des polynômes retards :

$$\begin{aligned} (1 - \phi_1 B - \dots - \phi_p B^p) X_t &= (1 - \theta_1 B - \dots - \theta_q B^q) \varepsilon_t \quad (2.15) \\ \iff \Phi(B) X_t &= \Theta(B) \varepsilon_t \end{aligned}$$

On peut noter que le processus est stationnaire et inversible si toutes les racines des polynômes respectivement Φ et Θ sont à l'extérieur du cercle unité.

Par la suite, on suppose ces deux conditions vérifiées.

On caractérise les propriétés statistiques du processus $ARMA$ en étudiant ses deux premiers moments :

2.4. MODÉLISATION STATISTIQUE DES PROCESSUS (ARMA)

• La moyenne s'écrit :

$$\mathbb{E}[\phi_1 X_{t-1} + \dots + \phi_p X_{t-p} + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \dots + \theta_q \varepsilon_{t-q}] = 0; \text{ car } \mathbb{E}[\varepsilon_t] = 0 \quad \forall t$$

• La covariance :

Pour l'expression de la covariance à l'ordre k , on commence par multiplier chaque terme par X_{t-k} :

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

Puis on prend l'espérance de chacun des termes de l'égalité et l'on obtient après simplifications

$$\gamma_k = \phi_1 \gamma_{k-1} + \dots + \phi_p \gamma_{k-p} \quad \text{pour } k > p$$

• Le coefficient d'autocorrélation s'écrit :

$$\rho_k = \phi_1 \rho_{k-1} + \dots + \phi_p \rho_{k-p} \quad \text{pour } k > p$$

Chapitre 3

Méthodologie de Box-Jenkins

Ce sont G.BOX et G.JENKINS qui ont contribué dans les années soixante dix à formaliser la théorie et la pratique des modèles de séries temporelles. L'objectif auquel ils se proposent de répondre dans leur ouvrage "Time Series Analysis : Forecasting and Control", Holden Day, 1976 second ed, est de construire un modèle aléatoire de type ARMA permettant de reproduire au mieux les réalisations $(X_t)_1^T$ d'une série temporelle. L'une des difficultés mise en avant par les deux auteurs est que la classe des modèles ARMA est vaste. Aussi, ils proposent de standardiser certaines procédures afin de retenir le meilleur modèle dans des délais les plus courts possibles. Trois étapes sont nécessaires à la construction des modèles : une phase d'identification, une phase d'estimation, une phase de validation. A l'issue de ses trois phases, une fois déterminé le meilleur modèle ARMA, on utilise ce modèle afin de faire des prévisions.

3.1 Identification des processus ARMA

Cette étape est la plus difficile dans le cycle de trois étapes définies par Box-Jenkins. Aussi, on s'attachera à retenir plusieurs modèles candidats qui seront éliminés successivement dans les étapes deux et trois du cycle. Plusieurs critères sont nécessaires pour identifier les modèles :

1. Il faut tout d'abord s'assurer que la série est stationnaire en variance et en moyenne.

- (a) Si la série n'est pas stationnaire en variance, on transforme la série d'origine en appliquant la fonction logarithme, racine carrée ou la transfor-

3.1. IDENTIFICATION DES PROCESSUS ARMA

mation puissance de Box-Cox.

(b) Si la série n'est pas stationnaire en moyenne, il s'agit tout d'abord d'identifier la nature de la non stationnarité avant d'appliquer la transformation adéquate.

i. Si la non stationnarité en moyenne est de type déterministe, alors il faut régresser la série d'origine sur une constante et une tendance. Puis, on modélise à l'aide des modèles ARMA la série des résidus de la régression.

ii. Si la non stationnarité en moyenne est de type stochastique, alors on différencie la série d'origine autant de fois que nécessaire afin d'aboutir à une série stationnaire. Puis on modélise la série différenciée à l'aide de modèle ARMA.

2. Une fois la série stationnarisée, on cherche à identifier les ordres p et q d'un modèle ARMA. On compare pour ce faire les fonctions estimées des ACF et PACF avec les représentations théoriques de ces deux fonctions. On applique le test de Bartlett sur chaque coefficient de corrélation et de corrélation partielle afin de vérifier s'il est ou non significativement différent de 0. Box et Jenkins proposent de retenir $K = \frac{N}{4}$ coefficients de corrélation. On rappelle la procédure de test :

3.1.1 Test de Bartlett

Bartlett [1944] a proposé de tester si les observations $(X_t)_{t=1}^N$ étaient non corrélées à l'ordre k . Soit la stratégie de test suivante :

$$\begin{cases} H_0 : \rho_k = 0 \\ H_1 : \rho_k \neq 0 \end{cases}$$

où ρ_k est le coefficient de corrélation théorique à l'ordre k du processus X_t . Sous H_0 vraie, il montre alors que la statistique de test a pour expression :

$$\frac{\hat{\rho}_k - 0}{\sqrt{V[\hat{\rho}_k]}} \underset{H_0}{\rightsquigarrow} \mathcal{N}(0, 1).$$

Or, $V[\hat{\rho}_k]$ étant inconnue on peut l'approximer par :

$$\hat{V}[\hat{\rho}_k] = \frac{1}{N} \left(1 + 2 \sum_{k=1}^q \rho_k^2 \right)$$

Sous H_0 , un estimateur de la variance est $\hat{V}[\hat{\rho}_k] = \frac{1}{N}$. D'où la statistique de test s'écrit selon la taille de l'échantillon (N) :

$$\begin{cases} \frac{\hat{\rho}_k - 0}{\sqrt{1/N}} \underset{H_0}{\rightsquigarrow} \mathcal{N}(0, 1) & \text{si : } N \geq 30 \\ \frac{\hat{\rho}_k - 0}{\sqrt{1/N}} \underset{H_0}{\rightsquigarrow} t_{(N)} & \text{si : } N < 30 \end{cases}$$

Graphiquement, lorsque l'on trace le corrélogramme de la série résiduelle issue d'un modèle de régression, on doit observer sous H_0 que tous les coefficients de corrélation sont compris dans l'intervalle de confiance $\left[-z_{\frac{\alpha}{2}} \sqrt{\frac{1}{N}}; z_{\frac{\alpha}{2}} \sqrt{\frac{1}{N}}\right]$ où $z_{\frac{\alpha}{2}}$ est le fractile au seuil de α de la loi $N(0, 1)$ ou de la loi de student de degré N ($t_{(N)}$). On en déduit des règles utiles à la caractérisation des processus *AR* et *MA* :

- Pour un processus *AR*(p), les p premiers coefficients estimés de la PACF sont en dehors des bornes de l'intervalle de confiance.
- Pour un processus *MA* (q), les q premiers coefficients estimés de l'ACF sont en dehors des bornes de l'intervalle de confiance.

3.2 Estimation des paramètres du modèle

Cette étape consiste à estimer les paramètres du modèle adéquat retenu en utilisant plusieurs méthodes comme :

- La méthode des moments.
- La méthode de maximum de vraisemblance.
- La méthode des moindres carrées.

3.2.1 Méthodes des moments :

Pour un modèle donné, on sait que les autocorrélations ρ_h et τ_h dépendent des paramètres $\phi_1, \dots, \phi_p, \theta_1, \dots, \theta_q$ selon des équations théoriques connues :

$(\rho_1, \dots, \rho_p, \tau_1, \dots, \tau_q) = F(\phi_1, \dots, \phi_p, \theta_1, \dots, \theta_q)$. Or on sait facilement estimer les autocorrélations ρ_h et τ_h , il suffit donc d'inverser les équations pour estimer les paramètres : $(\hat{\phi}_1, \dots, \hat{\phi}_p, \hat{\theta}_1, \dots, \hat{\theta}_q) = F^{-1}(\hat{\rho}_1, \dots, \hat{\rho}_p, \hat{\tau}_1, \dots, \hat{\tau}_q)$. Il n'est pas nécessaire d'explicitier ici la forme des résultats.

3.2.2 La méthode de maximum de vraisemblance

Si le processus ε_t est gaussien $N(0, \sigma^2)$, alors $\varepsilon_t \rightsquigarrow N(0, \sigma^2 I)$ et $X \rightsquigarrow N(0, \sigma^2 \Omega)$. Ainsi, la vraisemblance (densité du vecteur $X = (X_1, \dots, X_T)'$) est donnée par :

$$\mathcal{L}(X_1, \dots, X_T; \phi, \theta, \sigma^2) = \frac{1}{(2\pi\sigma^2)^{T/2}} \frac{1}{(\det \Omega)^{1/2}} \exp\left(\frac{-1}{2\sigma^2} X' \Omega^{-1} X\right)$$

Le log-vraisemblance est alors donné par :

$$\ln \mathcal{L}(X_1, \dots, X_T; \phi, \theta, \sigma^2) = -\frac{T}{2} \ln(2\pi) - \frac{T}{2} \ln(\sigma^2) - \frac{1}{2} \ln(\det \Omega) - \frac{1}{2\sigma^2} X' \Omega^{-1} X$$

Le problème est que ce log-vraisemblance est difficile à calculer et donc à maximiser à cause de $(\det \Omega)$ et de Ω^{-1} (matrice $T \times T$). De plus, il faut se donner des valeurs préliminaires pour les paramètres, puisque la maximisation du log-vraisemblance utilise des algorithmes de maximisation itératifs.

3.2.3 La méthode des moindres carrés

Soit le modèle général *ARMA* (p, q) : $\Phi(B)X_t - \delta = \Theta(B)\varepsilon_t$, où $\Phi(B)$ et $\Theta(B)$ sont des polynômes de degrés respectifs p et q et de coefficients ϕ et θ . On conçoit que pour une trajectoire donnée du processus, et donc un ensemble de T réalisations X_1, X_2, \dots, X_T , il est possible de calculer pour chaque ensemble de $p+q+1$ valeurs $\tilde{\phi}, \tilde{\theta}, \tilde{\delta}$ une série de résidus empiriques e_1, e_2, \dots, e_T à condition de se donner deux séquences de réalisations non observées. L'une de longueur p relative à la trajectoire du processus X_t antérieurement à la période d'observation : $X_0, X_{-1}, \dots, X_{1-p}$. L'autre, de longueur q , relative à la trajectoire du résidu également antérieure à la période d'observation : $e_0, e_{-1}, \dots, e_{1-q}$. Ceci est en effet nécessaire au calcul de la somme des carrés des résidus $S(\tilde{\phi}, \tilde{\theta}, \tilde{\delta})$ sur la totalité de la période d'observation, avec

$$S(\tilde{\phi}, \tilde{\theta}, \tilde{\delta}) = \sum_{t=1}^T e_t^2;$$

comme on peut s'en apercevoir sur les exemples suivants qui explicitent le calcul de e_1 :

3.3.2 Test d'autocorrélation des résidus

Si les résidus $(\varepsilon_t, t \in \mathbb{Z})$ obéissent à un bruit blanc, il ne doit pas exister d'autocorrélation dans la série. On peut alors utiliser les différents tests suivants :

1. Test de Durbin Watson : test d'autocorrélation d'ordre 1.
 2. Etude de la FAC et de la FACP : on doit vérifier qu'il n'existe aucune autocorrélation ou autocorrélation partielle significativement non nulle pour le processus étudié. Cette étude est prolongée par les tests du "porte-manteau"
 3. Test du "porte-manteau" ou tests d'adéquation globale du modèle.
- Ces tests reposent sur l'idée que la FAC d'un bruit blanc ne doit pas révéler d'autocorrélations non nulles.

En pratique, on utilise deux tests :

Test de Box et Pierce (1970) :

Ce test, encore appelé "porte-manteau", a pour objet de tester le caractère non autocorrélé des résidus. La statistique de ce test est :

$$Q_{BP} = T \sum_{k=1}^K \rho_k^2 \xrightarrow[T \rightarrow \infty]{L} \chi^2(K - p - q)$$

où T est le nombre d'observations et ρ_k est le coefficient d'autocorrélation d'ordre k des résidus estimés.

Les hypothèses du test de Box-Pierce sont les suivantes :

$$\begin{cases} H_0 : \rho_1 = \rho_2 = \dots = \rho_k = 0 \\ H_1 : \text{il existe au moins un } \rho_i \neq 0 \end{cases}$$

Sous l'hypothèse H_0 , la statistique Q_{BP} obéit à un $\chi^2_{(1-\alpha)}(K-p-q)$ degrés de liberté, où K est le nombre de retards choisis pour les autocorrélations.

Pour effectuer ce test il est conseillé de choisir $K = \frac{T}{4}$ (d'après Box-Jenkins).

La règle de décision est la suivante :

- Si $Q_{BP} < \chi^2_{(1-\alpha)}(K-p-q)$ on accepte $H_0 \implies$ les ε_t forment un bruit blanc.

- Si $Q_{BP} > \chi^2_{(1-\alpha)}(K-p-q)$ on rejette $H_0 \implies$ les ε_t ne forment pas un bruit blanc.

Test de Ljung-Box (1978) :

Ce test est à appliquer, de préférence au test de Box-Pierce, lorsque l'échantillon est de petite taille. La distribution de la statistique du test de Ljung-Box est en effet plus proche de celle de Khi-deux en petit échantillon. Ces statistiques, définissent pour un ordre K , correspondant à l'hypothèse nulle

$$H_0 : \rho_k = 0 \quad \forall k \leq K$$

et sont construites de la façon suivante :

$$Q_{LB} = T(T+2) \sum_{k=1}^K \frac{\rho_k^2}{T-k} \xrightarrow[T \rightarrow \infty]{\mathcal{L}} \chi^2(K-p-q)$$

La statistique Q_{LB} suit une loi de Khi-deux à $(K-p-q)$ degrés de liberté.

3.3.3 Tests d'homoscédasticité

Le test le plus utilisé dans le domaine des séries temporelles est le test d'effet ARCH proposé par Engle 1982, car il est très fréquemment employé en économétrie des séries temporelles financières.

Soit $(\varepsilon_t, t \in T)$ le processus d'erreur, on a la stratégie de test suivante :

$$\begin{cases} H_0 : V[\varepsilon_t] = \sigma_\varepsilon^2 \quad \forall t \\ H_1 : V[\varepsilon_t] = h_t = h(\theta_0 + \theta_1 \varepsilon_{t-1}^2 + \dots + \theta_q \varepsilon_{t-q}^2) \end{cases}$$

Sous H_0 vraie, il montre que :

$$\chi_{ARCH}^2 = NR^2 \rightsquigarrow \chi^2(q) \tag{3.1}$$

où R^2 est le coefficient de détermination obtenu dans la régression des résidus carrés issus de la modélisation ARMA sur les résidus carrés retardés jusqu'à l'ordre q .

En pratique, on procède selon les étapes suivantes :

1. On estime le modèle ARMA et l'on en déduit la série des résidus :

$$\Phi(B)X_t = \Theta(B)e_t \implies e_t = \frac{\Phi(B)}{\Theta(B)}X_t$$

2. On régresse la série des résidus carrés sur les résidus carrés retardés.

3.3. VALIDATION DU PROCESSUS ARMA(P,Q)

$$\begin{cases} e_t^2 = \theta_0 + \sum_{j=1}^q \theta_j e_t^2 + \varepsilon_t \\ \text{où} \\ \varepsilon_t \rightsquigarrow \mathcal{N}(0, \sigma_\varepsilon^2) \end{cases}$$

3. On calcule la statistique de test χ_{ARCH}^2 donnée par l'équation (3.1).
4. On calcule en appliquant la règle de décision conventionnelle :
 - (a) Si $\chi_{ARCH}^2 < \chi^2(q)$, alors on ne rejette pas H_0 ; la variance de l'erreur ne fait pas apparaître de forme d'hétéroscédasticité telle que supposée.
 - (b) Si $\chi_{ARCH}^2 > \chi^2(q)$, alors on rejette H_0 ; la variance de l'erreur fait apparaître une forme d'hétéroscédasticité conditionnelle.

3.3.4 Tests de normalité

Pour vérifier si le processus des résidus $(\varepsilon_t, t \in T)$ est un bruit blanc gaussien, le test le plus courant est celui de Jarque et Bera. Ce dernier est fondé sur la notion de Skewness (moment d'ordre 3 et asymétrie) et de Kurtosis (moment d'ordre 4 et queue de distribution). Soit μ_k le moment empirique d'ordre k du processus $(\varepsilon_t, t \in T)$:

$$\mu_k = \frac{1}{T} \sum_{i=1}^t (\hat{\varepsilon}_t - \bar{\varepsilon}_t)^k$$

Les coefficients de la Skewness (S_k) et de la Kurtosis (K_u) est alors définie par :

$$(S_k)^{\frac{1}{2}} = \frac{\mu_3}{\mu_2^{\frac{3}{2}}} \xrightarrow[T \rightarrow \infty]{\mathcal{L}} \mathcal{N}\left(0, \sqrt{\frac{6}{T}}\right)$$

$$K_u = \frac{\mu_4}{\mu_2^2} \xrightarrow[T \rightarrow \infty]{\mathcal{L}} \mathcal{N}\left(3, \sqrt{\frac{24}{T}}\right)$$

On construit alors les statistiques centrées réduites correspondantes à $(S_k)^{\frac{1}{2}}$ et K_u que l'on compare aux seuils d'une loi normale centrée réduite. Ce test est pour examiner si la série est normalement distribuée. La statistique est calculée comme suit :

$$\frac{(S_k)^{\frac{1}{2}}}{\sqrt{\frac{6}{T}}} \xrightarrow[T \rightarrow \infty]{\mathcal{L}} \mathcal{N}(0, 1)$$

$$\frac{K_u - 3}{\sqrt{\frac{24}{T}}} \xrightarrow[T \rightarrow \infty]{\mathcal{L}} \mathcal{N}(0, 1)$$

Si la statistique centrée réduite de $(S_k)^{\frac{1}{2}}$ est inférieure au seuil 1,96 à 5%, on accepte l'hypothèse de symétrie et l'hypothèse de normalité. Si la statistique centrée réduite de K_u est inférieure au seuil 1,96 à 5%, on accepte l'hypothèse de queue de distribution plates et l'hypothèse de normalité.

Le test de Jarque et Bera regroupe ces deux tests en un seul test. On construit la statistique :

$$s = \frac{T}{6} S_k + \frac{T}{24} (K_u - 3)^2 \xrightarrow[T \rightarrow \infty]{\mathcal{L}} \chi^2(2)$$

Donc si $s \geq \chi_{(1-\alpha)}^2(2)$ on rejette l'hypothèse H_0 de normalité des résidus au seuil de α .

3.3.5 Critères de choix des modèles

Après examen des coefficients et des résidus, certains modèles sont écartés pour départager les modèles restants, on fait appel aux critères standards et critères d'information.

Critères standards

L'erreur absolue moyenne (MAE)

$$MAE = \frac{1}{t} \sum |e_t|$$

Racine de l'erreur quadratique moyenne (RMSE)

$$RMSE = \sqrt{\frac{1}{T} \sum e_t^2}$$

Ecart absolu moyen en pourcentage (MAPE)

$$MAPE = 100 \frac{1}{T} \sum \left| \frac{e_t}{X_t} \right|$$

Plus la valeur de ces critères est faible, plus le modèle estimé est proche des observations.

3.4. PRÉVISION

$$X_{t+k} - \hat{X}_t(k) = \sum_{j=0}^{k-1} \pi_j \varepsilon_{t+k-j}$$

Déterminons un intervalle de confiance sur la prévision $\hat{X}_t(k)$, sous l'hypothèse de normalité des résidus ε_t . On montre alors que :

$$\frac{X_{t+k} - \hat{X}_t(k)}{\text{var} [X_{t+k} - \hat{X}_t(k)]^{1/2}} \xrightarrow[T \rightarrow \infty]{\mathcal{L}} \mathcal{N}(0, 1)$$

Or on sait que :

$$\mathbb{E} \left\{ [X_{t+k} - \hat{X}_t(k)]^2 \right\} = \mathbb{E} \left[\left(\sum_{j=0}^{k-1} \pi_j \varepsilon_{t+k-j} \right)^2 \right] = \sum_{j=0}^{k-1} \pi_j^2 \sigma_\varepsilon^2$$

D'où

$$\frac{X_{t+k} - \hat{X}_t(k)}{\sigma_\varepsilon \left[\sum_{j=0}^{k-1} \pi_j^2 \right]^{1/2}} \xrightarrow[T \rightarrow \infty]{\mathcal{L}} \mathcal{N}(0, 1)$$

On peut donc construire un intervalle de confiance sous la forme :

$$IC = \left[\hat{X}_t(k) \pm t^{\alpha/2} \left(\sum_{j=0}^{k-1} \pi_j^2 \right)^{1/2} \hat{\sigma}_\varepsilon \right]$$

Chapitre 4

Conclusion générale

Le but de ce travail était de s'intéresser à la prévision des séries chronologiques en se basant sur la théorie des processus temporelles stationnaires. L'essentiel de ce document était consacré à la représentation des processus linéaires de type ARMA qui sont devenus un outil important pour les prévisions à court terme. En particulier, on insiste sur les différentes étapes de modélisation : identification du processus, estimation du modèle, validation par tests et prévision.

Il ya plusieurs extensions possibles des modèles ARMA :

- En fait nous avons considérés uniquement les séries temporelles stationnaires alors que beaucoup de séries chronologiques économiques sont non stationnaires. Sous l'approche de Box et Jenkins ce problème est traité par les modèles ARIMA (ARMA intégré).

Annexe A

Annexe

A.1 La convergence en moyenne quadratique

Une suite de v.a.r $(X_n)_{n \in \mathbb{N}}$ converge en moyenne quadratique vers une v.a.r X si :

$$\lim_{n \rightarrow \infty} (\mathbb{E}(X_n(\omega) - X(\omega))^2) = 0$$

où : $X_n \xrightarrow{m.q} X$

A.2 Liste des abreviations

AR : Autoregressive ;
MA : Moving Average ;
ARMA : Autoregressive Moving Average ;
ACV : Fonction d'autocovariance ;
PAC : Partial Autocorrelation ;
ACF : Autocorelation Function ;
PACF : Partial Autocorelation Function ;
MAE : Mean Absolute Error ;
RMSE : Root Mean Square Error ;
MAPE : Mean Average Percentage Error ;
AIC : Akaike information critetirion ;
BIC : Bayesien Autocorelation Function ;

A.3 Tables de valeurs numériques

A.3.1 Lois continues

Loi normale ou de Laplace-Gauss :

On dit qu'une variable aléatoire X , à valeurs dans \mathbb{R} , suit une loi normale de paramètres μ et σ si sa densité de probabilité f est définie par :

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}, \quad \forall x \in \mathbb{R}$$

La fonction de répartition de la loi normale est définie à partir de la densité par :

$$F(x) = \int_{-\infty}^x \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(t-\mu)^2}{2\sigma^2}} dt$$

Propriétés :

Si $X \rightsquigarrow \mathcal{N}(\mu, \sigma^2)$ alors $\begin{cases} E(X) = \mu \\ Var(X) = \sigma^2. \end{cases}$

A.3. TABLES DE VALEURS NUMÉRIQUES

Table 1

Fonction de répartition de la loi normale réduite : $\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x \exp\left(-\frac{t^2}{2}\right) dt$

x	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.	0.5	0.50399	0.50798	0.51197	0.51595	0.51994	0.52392	0.52790	0.53188	0.53586
0.1	0.53983	0.54380	0.54776	0.55172	0.55567	0.55962	0.56356	0.56749	0.57142	0.57535
0.2	0.57926	0.58317	0.58706	0.59095	0.59483	0.59871	0.60257	0.60642	0.61026	0.61409
0.3	0.61791	0.62172	0.62552	0.62930	0.63307	0.63683	0.64058	0.64431	0.64803	0.65173
0.4	0.65542	0.65910	0.66276	0.66640	0.67003	0.67364	0.67724	0.68082	0.68439	0.68795
0.5	0.69146	0.69497	0.69847	0.70194	0.70540	0.70884	0.71226	0.71566	0.71904	0.72240
0.6	0.72575	0.72907	0.73237	0.73565	0.73891	0.74215	0.74537	0.74857	0.75175	0.75490
0.7	0.75804	0.76115	0.76424	0.76730	0.77035	0.77337	0.77637	0.77935	0.78230	0.78524
0.8	0.78814	0.79103	0.79389	0.79673	0.79955	0.80234	0.80511	0.80785	0.81057	0.81327
0.9	0.81594	0.81859	0.82121	0.82381	0.82639	0.82894	0.83147	0.83398	0.83646	0.83891
1.	0.84134	0.84375	0.84614	0.84849	0.85083	0.85314	0.85543	0.85769	0.85993	0.86214
1.1	0.86433	0.86650	0.86864	0.87076	0.87286	0.87493	0.87698	0.87900	0.88100	0.88298
1.2	0.88493	0.88686	0.88877	0.89065	0.89251	0.89435	0.89617	0.89796	0.89973	0.90147
1.3	0.90320	0.90490	0.90658	0.90824	0.90988	0.91149	0.91309	0.91466	0.91621	0.91774
1.4	0.91924	0.92073	0.92220	0.92364	0.92507	0.92647	0.92785	0.92922	0.93056	0.93189
1.5	0.93319	0.93448	0.93574	0.93699	0.93822	0.93943	0.94062	0.94179	0.94295	0.94408
1.6	0.94520	0.94630	0.94738	0.94845	0.94950	0.95053	0.95154	0.95254	0.95352	0.95449
1.7	0.95543	0.95637	0.95728	0.95818	0.95907	0.95994	0.96080	0.96164	0.96246	0.96327
1.8	0.96407	0.96485	0.96562	0.96638	0.96712	0.96784	0.96856	0.96926	0.96995	0.97062
1.9	0.97128	0.97193	0.97257	0.97320	0.97381	0.97441	0.97500	0.97558	0.97615	0.97670
2.	0.97725	0.97778	0.97831	0.97882	0.97932	0.97982	0.98030	0.98077	0.98124	0.98169
2.1	0.98214	0.98257	0.98300	0.98341	0.98382	0.98422	0.98461	0.98500	0.98537	0.98574
2.2	0.98610	0.98645	0.98679	0.98713	0.98745	0.98778	0.98809	0.98840	0.98870	0.98899
2.3	0.98928	0.98956	0.98983	0.99010	0.99036	0.99061	0.99086	0.99111	0.99134	0.99158
2.4	0.99180	0.99202	0.99224	0.99245	0.99266	0.99286	0.99305	0.99324	0.99343	0.99361
2.5	0.99379	0.99396	0.99413	0.99430	0.99446	0.99461	0.99477	0.99492	0.99506	0.99520
2.6	0.99534	0.99547	0.99560	0.99573	0.99585	0.99598	0.99609	0.99621	0.99632	0.99643
2.7	0.99653	0.99664	0.99674	0.99683	0.99693	0.99702	0.99711	0.99720	0.99728	0.99736
2.8	0.99744	0.99752	0.99760	0.99767	0.99774	0.99781	0.99788	0.99795	0.99801	0.99807
2.9	0.99813	0.99819	0.99825	0.99831	0.99836	0.99841	0.99846	0.99851	0.99856	0.99861

Tableau1- Loi normale centrée réduite.

Loi du khi-deux χ_n^2 :

Soient X_1, X_2, \dots, X_n n variables aléatoires indépendantes de même loi $\mathcal{N}(0, 1)$, et soit

$$K = X_1^2 + X_2^2 + \dots + X_n^2 = \sum_{i=1}^n X_i^2$$

alors, K suit une loi du Khi-deux à n degrés de liberté ($K \rightsquigarrow \chi_n^2$).

Fonction de densité :

$$l_n(y) = \begin{cases} 0 & , \text{ si } y \leq 0 \\ \frac{1}{2^{\frac{n}{2}} \left(\frac{n}{2} - 1\right)} y^{\frac{n}{2}-1} e^{-\frac{y}{2}}, & \text{ si } y > 0. \end{cases}$$

Fonction de répartition :

$$F(y) = \int_{-\infty}^y l_n(t) dt.$$

Propriétés :

$$\text{Si } K \rightsquigarrow \chi_n^2 \text{ alors } \begin{cases} E(K) = n \\ \text{Var}(K) = 2n. \end{cases}$$

A.3. TABLES DE VALEURS NUMÉRIQUES

Table 2

Fractiles de la loi du khi-deux

$\nu \backslash p$	0.0005	0.001	0.005	0.01	0.025	0.05	0.1	0.5
1	0.000	0.000	0.000	0.000	0.001	0.004	0.016	0.455
2	0.001	0.002	0.010	0.020	0.051	0.103	0.211	1.386
3	0.015	0.024	0.072	0.115	0.216	0.352	0.584	2.366
4	0.064	0.091	0.207	0.297	0.484	0.711	1.064	3.357
5	0.158	0.210	0.412	0.554	0.831	1.145	1.610	4.351
6	0.299	0.381	0.676	0.872	1.237	1.635	2.204	5.348
7	0.485	0.598	0.989	1.239	1.690	2.167	2.833	6.346
8	0.710	0.857	1.344	1.646	2.180	2.733	3.490	7.344
9	0.972	1.152	1.735	2.088	2.700	3.325	4.168	8.343
10	1.265	1.479	2.156	2.558	3.247	3.940	4.865	9.342
11	1.587	1.834	2.603	3.053	3.816	4.575	5.578	10.341
12	1.934	2.214	3.074	3.571	4.404	5.226	6.304	11.340
13	2.305	2.617	3.565	4.107	5.009	5.892	7.042	12.340
14	2.697	3.041	4.075	4.660	5.629	6.571	7.790	13.339
15	3.108	3.483	4.601	5.229	6.262	7.261	8.547	14.339
16	3.536	3.942	5.142	5.812	6.908	7.962	9.312	15.338
17	3.980	4.416	5.697	6.408	7.564	8.672	10.085	16.338
18	4.439	4.905	6.265	7.015	8.231	9.390	10.865	17.338
19	4.912	5.407	6.844	7.633	8.907	10.117	11.651	18.338
20	5.398	5.921	7.434	8.260	9.591	10.851	12.443	19.337
21	5.896	6.447	8.034	8.897	10.283	11.591	13.240	20.337
22	6.404	6.983	8.643	9.542	10.982	12.338	14.041	21.337
23	6.924	7.529	9.260	10.196	11.689	13.091	14.848	22.337
24	7.453	8.085	9.886	10.856	12.401	13.848	15.659	23.337
25	7.991	8.649	10.520	11.524	13.120	14.611	16.473	24.337
26	8.538	9.222	11.160	12.198	13.844	15.379	17.292	25.336
27	9.093	9.803	11.808	12.879	14.573	16.151	18.114	26.336
28	9.656	10.391	12.461	13.565	15.308	16.928	18.939	27.336
29	10.227	10.986	13.121	14.256	16.047	17.708	19.768	28.336
30	10.804	11.588	13.787	14.953	16.791	18.493	20.599	29.336

Le fractile $\chi_{\nu, p}^2$ est tel que $P(\chi_{\nu}^2 < \chi_{\nu, p}^2) = p$.

Tableau 2- Lois du khi-deux.

Bibliographie

- [1] Corinne Perraudin. "Séries Chronologiques, Quelques éléments du cours", Chapitre 1 : "Les modèles ARMA stationnaires ". Pages : 6-12,16-36. Université Paris I, Magistère d'économie-Deuxième année. Année : 2004-2005.
- [2] Dagnelie Pierre. Statistique théorique appliquée (T1). Boeck. 2007.
- [3] Grais Bernard. Méthodes statistiques. Dunod. Paris. 1992.
- [4] Jean-Marie Dufor. "Histoire de l'analyse des séries chronologiques", Version :10-Janvier-2006.
- [5] J.J. Daudin, C. DUBY, S. Robin & P. Trécourt. "Analyse des séries chronologiques", INA-PG, Mathématique. Mai-1996.
- [6] Rainer Von Sachs, Sébastien Van Bellegem. "STAT 2414, Séries chronologiques", Institut de statistique, Université Catholique de Louvain, 4^e édition, Septembre 2005.
- [7] Tassi Philippe. Méthodes statistiques (3^{ème} édition). Economica. Paris. 2004.