

REPUBLIQUE ALGERIENNE DEMOCRATIQUE ET POPULAIRE  
MINISTERE DE L'ENSEIGNEMENT SUPERIEUR ET DE LA RECHERCHE SCIENTIFIQUE  
UNIVERSITE DE JIJEL  
FACULTE DES SCIENCES EXACTES ET DES SCIENCES DE LA NATURE ET DE LA VIE  
DEPARTEMENT DE MATHEMATIQUES  
ECOLE DOCTORALE – POLE DE CONSTANTINE

N° d'ordre : .....

N° de série : .....

## **MEMOIRE**

*Présenté pour l'obtention du diplôme de :*

**MAGISTER**

*Spécialité Mathématiques*

**OPTION**

*Probabilités et Statistiques*

**THEME**

*Sur l'approximation de la Vraisemblance  
d'un Processus Gaussien*

*Par :*

***Sellami Nawel***

Soutenu le :..../..../2010 devant le jury :

Président	: M. Denche	Professeur	<b>Université Constantine</b>
Rapporteur	: M. Dakhmouche	M.C	<b>Université Constantine</b>
Examineur	: M. Boushaba	M.C	<b>Université Constantine</b>
Examineur	: A. Benchettah	Professeur	<b>Université Annaba</b>

## REMERCIEMENTS

*C'est avec un grand plaisir que je réserve ces lignes en signe de gratitude et de reconnaissance à tous ceux qui ont contribué de près ou de loin à l'élaboration de ce travail.*

*Je tiens à exprimer ma plus profonde gratitude à mon directeur de recherche Monsieur: **Meghlaoui Dakhmouche** maître de conférences à l'université de **Constantine**, pour sa fraîcheur d'esprit, pour ses conseils judicieux, pour son infinie patience ainsi que pour les orientations qu'il m'a donnés tout au long de ce travail.*

*Je suis également très honoré que Monsieur **M. Denche** professeur à l'université de **Constantine** soit le président du jury de mon mémoire.*

*Mes remerciements s'adressent également à Monsieur **M. Boushaba** maître de conférences à l'université de **Constantine** et à Monsieur **A. Benchetteh** professeur à l'université de **Annaba** d'avoir accepté de juger ce travail et d'en être les examinateurs.*

*Je tiens aussi à remercier vivement Monsieur **Gherda Mebrouk** pour les conseils qu'il ma donnés.*

*Et pour finir, à tous mes amis en leur espérant bonne continuation dans leurs travaux.*

# Table des matières

<b>Introduction</b>	<b>iii</b>
<b>1 Rappels sur la paramétrisation des processus linéaires <i>ARMA</i></b>	<b>1</b>
1.1 Processus linéaires . . . . .	1
1.2 Processus autorégressif d'ordre $p$ . . . . .	4
1.2.1 Propriétés d'un processus $AR(p)$ . . . . .	5
1.3 Processus moyennes mobiles d'ordre $q$ . . . . .	6
1.3.1 Propriétés d'un processus $MA(q)$ . . . . .	7
1.4 Processus mixtes : autorégressif-moyennes mobiles d'ordres $p$ et $q$ . . . . .	7
1.4.1 Propriétés d'un processus $ARMA(p, q)$ . . . . .	8
<b>2 Approximation de Whittle de la fonction de vraisemblance d'un processus gaussien</b>	<b>10</b>
2.1 Fonction de vraisemblance exacte d'un processus gaussien . . .	10
2.2 Représentation spectrale d'un processus stochastique . . . . .	11
2.2.1 La fonction de densité spectrale . . . . .	11
2.2.2 Transformation de Toeplitz de la densité spectrale . . .	12
2.2.3 Périodogramme d'une série chronologique . . . . .	14
2.2.4 Fonction génératrice des autocovariances . . . . .	14
2.2.5 Calcul explicite de la densité spectrale . . . . .	15
2.3 Approximation de Whittle de la fonction de vraisemblance . .	16
<b>3 Approximation de Box-Jenkins de la fonction de vraisemblance d'un processus gaussien</b>	<b>20</b>
3.1 Forme innovation de la fonction de vraisemblance d'un pro- cessus aléatoire gaussien . . . . .	21

3.2	Approximation de la fonction de vraisemblance d'un processus autorégressif . . . . .	22
3.3	Approximation de la fonction de vraisemblance d'un processus en moyennes mobiles . . . . .	24
3.4	Approximation de la fonction de vraisemblance d'un processus autorégressif-moyennes mobiles . . . . .	27
<b>4</b>	<b>Détermination des valeurs initiales pour la somme des carrés inconditionnelle</b>	<b>31</b>
4.1	Principe de l'aller-retour . . . . .	31
4.2	Détermination des valeurs initiales d'un processus $MA(q)$ par la méthode d'aller-retour . . . . .	34
4.3	Détermination des valeurs initiales d'un processus $ARMA(p, q)$ par la méthode d'aller-retour . . . . .	41
4.3.1	Cas particulier : Calcul des valeurs initiales pour un processus $ARMA(1, 1)$ par la méthode d'aller-retour . . . . .	46

# Introduction

Dans l'étude de la statistique des processus linéaires ([4], [5], [7], [13], [16]), l'un des problèmes les plus fréquemment rencontrés est celui d'exprimer explicitement la fonction de vraisemblance. On sait que même en temps discret le calcul de cette fonction n'est jamais simple. Le problème est encore plus compliqué quand il s'agit de manipuler le rapport de vraisemblance qui joue un rôle essentiel pour la construction et la détermination de tests sur les modèles.

Dans ce travail nous présentons les principaux résultats sur l'étude de la vraisemblance des processus linéaires gaussiens. Ainsi, nous allons nous intéresser à l'étude de quelques types d'approximation de la vraisemblance. La première méthode exposée est due à Whittle ([24]). Cette technique, assez intuitive quoique peu générale, permet le calcul de l'approximation de la vraisemblance, notée  $L_N^W$ , au moyen du périodogramme. En effet, il a été démontré que la forme quadratique dans l'exponentielle est approchée par  $\frac{N}{4\pi} \int \frac{I_N(\omega)}{f(\omega, \theta)} d\omega$ . En d'autres termes, la log-vraisemblance  $L_N^W$  peut s'exprimer :

$$\ln L_N^W = C \int_{-\pi}^{+\pi} \ln f(\omega) d\omega - \frac{N}{4\pi} \int \frac{I_N(\omega)}{f(\omega, \theta)} d\omega$$

Le deuxième type d'approximation consiste à considérer tout processus comme limite de processus autorégressif. Cette idée a été largement popularisée par Box et Jenkins ([5]). Par ailleurs, on constate que la méthode de Box et Jenkins est une application particulière de la méthode de Whittle dans le cas d'un spectre rationnel et de ce fait la forme quadratique de la vraisemblance

est approchée par  $L_N^{BJ}$  définie telle que :

$$L_N^{BJ} = \sum_{j=1}^N \varepsilon_j^2$$

En général, le calcul des innovations nécessite la détermination de quelques valeurs de la série chronologique survenues avant l'instant origine. Alors, une procédure itérative ([11], [13]) permet dans le cas rationnel de calculer cette somme des carrés.

Cependant, sur la log-vraisemblance les deux approximations précédentes se prêtent mal aux calculs asymptotiques. On a donc introduit un troisième type d'approximation résumé dans ce qu'on va appeler la méthode de Toeplitz-Szegö ([3], [13]). Cette troisième méthode est basée sur le principe que la matrice d'autocovariance d'un processus n'est autre que la transformée de Toeplitz de sa densité spectrale. On adopte la représentation spectrale d'un processus aléatoire qui est essentiellement la transformée de Fourier de ce processus. L'avantage de cette forme de paramétrisation est de replacer l'étude des processus aléatoires dans le cadre élégant de l'analyse harmonique.

# Chapitre 1

## Rappels sur la paramétrisation des processus linéaires *ARMA*

Ce chapitre est consacré à l'étude d'une classe particulière de modèles pour les processus stochastiques, les modèles linéaires *ARMA*. Nous allons rappeler quelques définitions principales non détaillées sur les modèles *AR*( $p$ ), *MA*( $q$ ) et *ARMA*( $p, q$ ) qu'on va utiliser dans ce travail. Ces modèles sont linéaires et sous certaines conditions stationnaires, leur linéarité fournira une théorie simple de la prévision.

Pour faciliter les notations, nous introduisons et nous utilisons l'opérateur de retard qui est appelé aussi opérateur backshift, noté  $B$  défini tel que :

$$BX_t = X_{t-1} \quad \text{et} \quad B^m X_t = X_{t-m}$$

L'opérateur inverse ou translation positive est réalisée à l'aide de l'opérateur d'avance qui est appelé aussi opérateur forward, noté  $F$  défini tel que :

$$FX_t = X_{t+1} \quad \text{et} \quad F^m X_t = X_{t+m}$$

**Remarque 1.0.1** *L'opérateur de retard  $B$  et l'opérateur d'avance  $F$  sont inverses l'un de l'autre, i.e.,  $F = B^{-1}$ .*

### 1.1 Processus linéaires

La paramétrisation des modèles stochastiques que l'on va utiliser est basée sur l'idée de Yule ([5]) qui stipule qu'une série chronologique dont les valeurs

CHAPITRE 1. RAPPELS SUR LA PARAMÉTRISATION DES  
PROCESSUS LINÉAIRES ARMA

---

successives sont fortement dépendantes peut être considérée comme le résultat du passage à travers un filtre linéaire d'une suite de variables aléatoires  $\varepsilon_t$  indépendantes et centrées. Ces variables sont des tirages aléatoires d'une distribution donnée qui est en général une loi normale de moyenne nulle et de variance  $\sigma_\varepsilon^2$ . La suite des variables aléatoires  $\varepsilon_t, \varepsilon_{t-1}, \varepsilon_{t-2}, \dots$  est appelée processus bruit blanc et est notée  $\varepsilon_t$ . Cette opération de transformation du processus  $\varepsilon_t$  en un processus  $X_t$  par filtrage linéaire n'est autre que le produit de convolution d'une suite  $\{\psi_j\}_{j \geq 0}$  de poids et de la suite des valeurs  $\{\varepsilon_{t-j}\}_{j \geq 0}$  du processus des innovations. En d'autres termes, la valeur d'un processus à un instant  $t$  est le résultat de la somme pondérée des innovations antérieures à cet instant.

**Définition 1.1.1** *Tout processus linéaire  $X_t$  peut être exprimé sous la forme suivante :*

$$X_t = \varepsilon_t + \sum_{j=1}^{+\infty} \psi_j \varepsilon_{t-j} \quad (1.1)$$

où  $\varepsilon_t$  est un processus d'innovations de moyenne nulle et de variance  $\sigma_\varepsilon^2$ .

D'après (1.1) on a :

$$\begin{aligned} X_t &= \varepsilon_t + \sum_{j=1}^{\infty} \psi_j B^j \varepsilon_t \\ &= \left( 1 + \sum_{j=1}^{\infty} \psi_j B^j \right) \varepsilon_t \end{aligned}$$

On note par  $\psi(B) = \left( 1 + \sum_{j=1}^{\infty} \psi_j B^j \right)$ , alors le processus  $X_t$  s'écrit plus simplement comme suit :

$$X_t = \psi(B) \varepsilon_t \quad (1.2)$$

**Remarque 1.1.1** *L'opérateur  $\psi(B) = \sum_{j=0}^{\infty} \psi_j B^j$  est le filtre linéaire qui transforme  $\varepsilon_t$  en  $X_t$ . L'opérateur  $\psi(B)$  est appelé fonction de transfert. Si la série  $\sum_{j=0}^{\infty} \psi_j B^j$  est convergente, i.e. si  $\sum_{j=0}^{+\infty} \psi_j^2 < \infty$ , alors le filtre  $\psi(B)$  est dit stable et le processus  $X_t$  est dit stationnaire.*

On peut écrire la forme du processus linéaire  $X_t$  dans (1.1) d'une manière équivalente telle que :

$$X_t = \varepsilon_t + \sum_{j=1}^{\infty} \pi_j X_{t-j} \quad (1.3)$$

D'après (1.3) on obtient :

$$X_t = \varepsilon_t + \sum_{j=1}^{\infty} \pi_j B^j X_t$$

D'où

$$\left(1 - \sum_{j=1}^{\infty} \pi_j B^j\right) X_t = \varepsilon_t$$

On note par  $\pi(B) = \left(1 - \sum_{j=1}^{\infty} \pi_j B^j\right)$ , l'expression (1.3) nous donne :

$$\pi(B) X_t = \varepsilon_t \quad (1.4)$$

D'après (1.2) et (1.4) il y a une relation entre  $\psi(B)$  et  $\pi(B)$  qui peut être donnée telle que :

$$\pi(B) \psi(B) = 1$$

La notion de stationnarité occupe un rôle majeur dans la modélisation des séries temporelles, elle est à la base d'une théorie asymptotique générale. On considère généralement deux types de stationnarité : la stationnarité stricte et la stationnarité au second ordre.

**Définition 1.1.2** *Le processus  $X_t$  est dit strictement stationnaire si les vecteurs  $(X_1, X_2, \dots, X_k)^t$  et  $(X_{1+h}, X_{2+h}, \dots, X_{k+h})^t$  ont même loi jointe, pour tout entier  $k$  et tout entier relatif  $h$ .*

On peut également définir la stationnarité au sens faible, i.e. au second ordre, qui peut sembler moins exigeante car elle n'importe des contraintes qu'aux deux premiers moments des variables  $X_t$ . Pour l'étude probabiliste on se limite généralement à requérir la stationnarité au second ordre du processus étudié.

**Définition 1.1.3** *Le processus  $X_t$  est dit stationnaire au second ordre si*

- (i)  $E(X_t^2) < \infty \quad \forall t \in \mathbb{Z}$
- (ii)  $E(X_t) = \mu \quad \forall t \in \mathbb{Z}$
- (iii)  $\text{cov}(X_t, X_{t+h}) = \gamma(h) \quad \forall t, h \in \mathbb{Z}$

*La fonction  $\gamma(\cdot)$  (respectivement  $\rho(\cdot) = \frac{\gamma(\cdot)}{\gamma(0)}$ ) est appelée fonction d'auto-covariance (respectivement d'autocorrélation) de  $X_t$ .*

L'exemple le plus simple du processus stationnaire au second ordre est celui de bruit blanc. Ce processus est particulièrement important car il permet de construire des processus stationnaires plus complexes.

**Définition 1.1.4** *Le processus  $\varepsilon_t$  est appelé bruit blanc faible si les trois propriétés sont respectées :*

- (i)  $E(\varepsilon_t) = 0 \quad \forall t \in \mathbb{Z}$
- (ii)  $E(\varepsilon_t^2) = \sigma_\varepsilon^2$  est constante et strictement positive
- (iii)  $\text{cov}(\varepsilon_t, \varepsilon_s) = 0 \quad \forall s \neq t$

*Les  $\varepsilon_t$  sont indépendantes et identiquement distribuées, pour simplifier on le note par  $\varepsilon_t$  sont i.i.d.*

Le lien entre les deux types de stationnarité sera énoncé dans la lemme suivante :

**Lemme 1.1.1** *Si le processus  $X_t$  est strictement stationnaire et  $E(X_t^2) < \infty$ , alors  $X_t$  est stationnaire au second ordre. La réciproque est fautive en général. Par contre si  $X_t$  est gaussien alors les deux concepts de stationnarité coïncident.*

## 1.2 Processus autorégressif d'ordre $p$

**Définition 1.2.1** *On appelle processus autorégressif d'ordre  $p$ , noté  $AR(p)$ , un processus  $X_t$  vérifiant une relation du type :*

$$P(B)X_t = \varepsilon_t \tag{1.5}$$

## 1.2. PROCESSUS AUTORÉGRESSIF D'ORDRE $P$

où  $\varepsilon_t$  est un processus d'innovations qui satisfait les conditions  $E(\varepsilon_t) = 0$  et  $E(\varepsilon_t^2) = \sigma_\varepsilon^2$  et  $P(B)$  est appelé polynôme caractéristique du degré  $p$  du processus  $X_t$  défini tel que :

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

En outre, l'équation (1.5) peut être écrite sous la forme suivante :

$$X_t = \varepsilon_t + \phi_1 X_{t-1} + \phi_2 X_{t-2} + \dots + \phi_p X_{t-p} \quad (1.6)$$

### 1.2.1 Propriétés d'un processus $AR(p)$

#### Stationnarité

Il est aisé de constater qu'un processus  $AR(p)$  est un processus à filtre linéaire. Par exemple, on peut remplacer dans l'expression (1.6)  $X_{t-1}$  par :

$$X_{t-1} = \varepsilon_{t-1} + \phi_1 X_{t-2} + \phi_2 X_{t-3} + \dots + \phi_p X_{t-p-1}$$

Et ainsi de suite, on remplace  $X_{t-2}, X_{t-3}, \dots$  etc. Finalement,  $X_t$  s'exprimera comme une série en  $\varepsilon_t$ , i.e.

$$\begin{aligned} X_t &= \psi(B) \varepsilon_t = P^{-1}(B) \varepsilon_t \\ \text{avec } \psi(B) &= P^{-1}(B) \end{aligned}$$

Considérons l'équation  $P(B) = 0$  et  $C_i^{-1}$  ses racines, on peut donc écrire  $P(B)$  de la façon suivante :

$$P(B) = (1 - C_1 B)(1 - C_2 B) \dots (1 - C_p B)$$

et

$$\psi(B) = \sum_{i=1}^p \frac{K_i}{(1 - C_i B)} \quad K_i \in \mathbb{R}$$

La condition nécessaire pour que la série  $\psi(B)$  soit convergente est que  $|C_i| < 1$ ,  $i = 1, 2, \dots, p$ . Donc les racines de l'équation  $P(B) = 0$  doivent être à l'extérieur du cercle unité.

#### Inversibilité

Un processus  $AR(p)$  est toujours inversible puisque la série  $\pi(B) = 1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p$  est finie.

### Causalité

Un processus  $AR(p)$  est dit causal lorsqu'il existe une suite  $\{\psi_k\}$  telle que  $\sum_{k \in \mathbb{Z}} |\psi_k| < \infty$  et

$$X_t = \sum_{k=0}^{\infty} \psi_k \varepsilon_{t-k} \quad (1.7)$$

Le résultat suivant donne les conditions nécessaires et suffisantes de stationnarité et de causalité des modèles autorégressifs sur le comportement de l'opérateur  $B$ .

**Proposition 1.2.1** *Le processus autorégressif défini par (1.5) est causal et stationnaire si et seulement si le polynôme en  $z$ ,  $P(z) = 1 - \sum_{i=1}^p \phi_i z^i$ , est tel que :*

$$P(z) \neq 0 \quad \text{pour tout } z \in \mathbb{Z} \quad \text{tel que } |z| \leq 1$$

Les coefficients  $\psi_k$  apparaissant dans la représentation causale (1.7) sont déterminées par :

$$1 - \sum_{j=1}^{\infty} \psi_j z^j = \frac{1}{P(z)}$$

## 1.3 Processus moyennes mobiles d'ordre $q$

**Définition 1.3.1** *On appelle processus moyennes mobiles d'ordre  $q$ , noté  $MA(q)$ , un processus  $X_t$  qu'on peut écrire de la forme suivante :*

$$X_t = Q(B) \varepsilon_t \quad (1.8)$$

où  $\varepsilon_t$  est un processus d'innovations qui satisfait les conditions  $E(\varepsilon_t) = 0$  et  $E(\varepsilon_t^2) = \sigma_\varepsilon^2$ .  $Q(B)$  est appelé polynôme caractéristique du degré  $q$  du processus  $X_t$  défini par :

$$Q(B) = 1 - \theta_1 B - \theta_2 B^2 - \dots - \theta_q B^q$$

En outre, l'équation (1.8) peut être écrite de la façon suivante :

$$X_t = \varepsilon_t - \theta_1 \varepsilon_{t-1} - \dots - \theta_q \varepsilon_{t-q}$$

Il est évident que le processus  $MA(q)$  est un modèle à filtre linéaire.

### 1.3.1 Propriétés d'un processus $MA(q)$

#### Stationnarité

Contrairement aux processus  $AR(p)$ , les processus  $MA(q)$  sont toujours des processus stationnaires puisque la série  $Q(B) = 1 - \theta_1 B - \theta_2 B^2 - \dots - \theta_q B^q$  est finie.

#### Inversibilité

D'après l'équation (1.8), on déduit que  $\varepsilon_t = Q^{-1}(B) X_t$  alors on a  $\pi(B) = Q^{-1}(B)$ .

Considérons l'équation  $Q(B) = 0$  et notons par  $H_i^{-1}$  pour  $i = 1, 2, \dots, q$  ses racines, alors on peut écrire  $Q(B)$  de la manière suivante :

$$Q(B) = (1 - H_1 B)(1 - H_2 B) \dots (1 - H_q B)$$

et

$$\pi(B) = \sum_{i=1}^q \frac{M_i}{(1 - H_i B)} \quad M_i \in \mathbb{R}$$

La condition nécessaire pour que la série  $\pi(B)$  soit convergente est que  $|H_i| < 1$  pour  $i = 1, 2, \dots, q$ . Donc les racines de  $Q(B) = 0$  doivent être à l'extérieur du cercle unité.

#### Causalité

D'après la relation (1.7), nous remarquons immédiatement que tout processus moyennes mobiles est causal par définition.

## 1.4 Processus mixtes : autorégressif-moyennes mobiles d'ordres $p$ et $q$

Pour permettre une meilleure flexibilité dans l'ajustement des séries chronologiques, il est souvent utile de combiner les formes autorégressives et les formes moyennes mobiles.

**Définition 1.4.1** On appelle processus autorégressif-moyennes mobiles d'ordres  $p$  et  $q$ , noté  $ARMA(p, q)$ , un processus  $X_t$  vérifiant une relation du type :

$$P(B) X_t = Q(B) \varepsilon_t \quad (1.9)$$

CHAPITRE 1. RAPPELS SUR LA PARAMÉTRISATION DES  
PROCESSUS LINÉAIRES ARMA

---

où  $\varepsilon_t$  est un processus d'innovations et  $P(B)$  et  $Q(B)$  sont les polynômes caractéristiques des parties autorégressives et moyennes mobiles du processus  $X_t$  des degrés  $p$  et  $q$  respectivement et sont définis par :

$$\begin{cases} P(B) = 1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p \\ Q(B) = 1 - \theta_1 B - \theta_2 B^2 - \dots - \theta_q B^q \end{cases}$$

En outre, l'équation (1.9) peut être écrite sous la forme suivante :

$$X_t - \phi_1 X_{t-1} - \dots - \phi_p X_{t-p} = \varepsilon_t - \theta_1 \varepsilon_{t-1} - \dots - \theta_q \varepsilon_{t-q} \quad (1.10)$$

Par ailleurs, on peut démontrer que les modèles  $ARMA(p, q)$  sont des modèles à filtre linéaire. En effet, on peut remplacer dans l'expression (1.10)  $X_t$  par  $X_{t-1}$  tel que :

$$X_{t-1} - \phi_1 X_{t-2} - \dots - \phi_p X_{t-p-1} = \varepsilon_{t-1} - \theta_1 \varepsilon_{t-2} - \dots - \theta_q \varepsilon_{t-q-1}$$

On procédera de la même manière pour  $X_t, X_{t-1}, X_{t-2}, \dots etc.$  Finalement  $X_t$  peut s'exprimer comme une série en  $\varepsilon_t$ , i.e.

$$X_t = \psi(B) \varepsilon_t = P^{-1}(B) Q(B) \varepsilon_t$$

### 1.4.1 Propriétés d'un processus $ARMA(p, q)$

#### Stationnarité

La condition de stationnarité est que les racines de l'équation  $Q(B) = 0$  doivent être situées à l'extérieur du cercle unité.

#### Inversibilité

Un processus  $ARMA(p, q)$  défini par l'équation (1.9) est dit inversible s'il existe une suite  $\{\pi_j\}$  telle que  $\sum_{j=0}^{+\infty} |\pi_j| < \infty$  et pour tout  $t \in \mathbb{Z}$  :

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

Autrement dit un processus  $ARMA(p, q)$  est inversible si les racines de l'équation  $P(B) = 0$  sont à l'extérieur du cercle unité.

### Causalité

Un processus  $ARMA(p, q)$  défini par l'équation (1.9) est dit causal s'il existe une suite  $\{\psi_j\}$  telle que  $\sum_{j=0}^{+\infty} |\psi_j| < \infty$  et pour tout  $t \in \mathbb{Z}$  :

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

Dans la proposition suivante, on donnera une condition nécessaire et suffisante pour la stationnarité, l'inversibilité et la causalité d'un processus  $ARMA(p, q)$ .

**Proposition 1.4.1** *Soit  $X_t$  un processus  $ARMA(p, q)$  vérifiant l'équation (1.9). Supposons que les polynômes  $P(z) = 1 - \sum_{i=1}^p \phi_i z^i$  et  $Q(z) = 1 - \sum_{i=1}^q \theta_i z^i$  n'ont pas de racines communes.*

•  $X_t$  est stationnaire et causal si et seulement si  $P(z) \neq 0$  pour tout  $z \in \mathbb{Z}$  tel que  $|z| \leq 1$ . Les coefficients  $\psi_j$  sont déterminés par :

$$\sum_{j=0}^{\infty} \psi_j z^j = \frac{Q(z)}{P(z)} \quad \text{si } |z| \leq 1$$

•  $X_t$  est inversible si et seulement si  $Q(z) \neq 0$  pour tout  $z \in \mathbb{Z}$  tel que  $|z| \leq 1$ . Les coefficients  $\pi_j$  sont déterminés par :

$$\sum_{j=0}^{+\infty} \pi_j z^j = \frac{P(z)}{Q(z)} \quad \text{si } |z| \leq 1$$

## Chapitre 2

# Approximation de Whittle de la fonction de vraisemblance d'un processus gaussien

Whittle a proposé une approximation, notée  $L_N^W$ , de la fonction log-vraisemblance  $\ln L_N$  d'une série chronologique issue d'un processus gaussien basé sur le périodogramme et la densité spectrale. Pour cela, nous allons recadrer l'étude de la série chronologique dans le domaine fréquentiel. Cependant le périodogramme est un estimateur non consistant de la densité spectrale d'un processus aléatoire.

### 2.1 Fonction de vraisemblance exacte d'un processus gaussien

L'estimation des paramètres d'un processus linéaire par la méthode du maximum de vraisemblance donne lieu à des équations normales n'ayant pas de solutions explicites. Le problème reste entier quand on utilise la vraisemblance d'une série chronologique, car le calcul de l'inverse de la matrice d'autocovariance ou d'autocorrélation n'est pas simple.

Supposons que l'étape de l'identification se soit soldée par le choix d'un modèle qui convient pour la série chronologique étudiée (i.e. les ordres des parties autorégressives et moyennes mobiles ont été déterminés). Pour l'étape de l'estimation des paramètres, il est indispensable d'exprimer la fonction de vraisemblance du processus avec autant d'exactitude que possible ; puisqu'il

## 2.2. REPRÉSENTATION SPECTRALE D'UN PROCESSUS STOCHASTIQUE

---

est bien établi qu'une bonne approximation de la fonction de vraisemblance fournira de meilleurs résultats pour l'estimation des paramètres du modèle.

**Définition 2.1.1** *Considérons une suite de  $N$  variables aléatoires  $X_1, X_2, \dots, X_N$  de distribution normale centrée. Alors, la distribution jointe de  $X_1, X_2, \dots, X_N$ , notée  $f(x_1, x_2, \dots, x_N)$  est définie telle que :*

$$f(x_1, x_2, \dots, x_N, \boldsymbol{\theta}) = \frac{1}{(2\pi)^{N/2} (\det \Gamma_N)^{\frac{1}{2}}} e^{-\frac{1}{2} X^t \Gamma_N^{-1} X}$$

où  $\Gamma_N$  est la matrice d'autocovariance de  $X = (X_1, X_2, \dots, X_N)^t$  et  $\boldsymbol{\theta}$  le vecteur des paramètres.

**Remarque 2.1.1** *La vraisemblance de  $X_1, X_2, \dots, X_N$ , notée  $L_N(\boldsymbol{\theta})$ , est la distribution jointe  $f(x_1, x_2, \dots, x_N, \boldsymbol{\theta})$  de  $X_1, X_2, \dots, X_N$  considérée comme fonction de  $\boldsymbol{\theta}$  uniquement.*

*Le logarithme de la fonction de vraisemblance, noté  $\ln L_N(\boldsymbol{\theta})$ , s'exprime tel que :*

$$\ln L_N(\boldsymbol{\theta}) = -\frac{1}{2} (N \ln 2\pi + \ln \det \Gamma_N + X^t \Gamma_N^{-1} X)$$

Pour développer la méthode de Whittle, il est utile de rappeler quelques résultats sur l'analyse spectrale des processus aléatoires.

## 2.2 Représentation spectrale d'un processus stochastique

Soit  $X_t$  un processus stochastique centré et stationnaire du second ordre. Il est entièrement caractérisé par sa fonction d'autocovariance ou sa fonction de densité spectrale. C'est sur l'une d'entr'elles que porte l'inférence statistique.

### 2.2.1 La fonction de densité spectrale

L'utilité de l'analyse spectrale des processus stationnaires nous permet de replacer l'étude de la statistique des processus dans le cadre de l'analyse harmonique.

Soit  $\gamma(k)$  la fonction d'autocovariance définie telle que :

$$\gamma(k) = \text{cov}(X_t X_{t+k})$$

La fonction  $\gamma(k)$  est symétrique. De plus, la matrice  $(\gamma(i-j))_{i,j=1,2,\dots,N}$  est définie positive, i.e.

$$\sum_{i,j=1}^N b_i \gamma(i-j) b_j > 0$$

où  $(b_1, b_2, \dots, b_N)^t \in \mathbb{R}^N$ .

La fonction d'autocorrélation  $\rho(k)$  est la fonction d'autocovariance normalisée possédant les mêmes propriétés que la fonction d'autocovariance et est définie telle que :

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

Donc la matrice d'autocorrélation est définie positive.

Supposons que la fonction d'autocovariance  $\gamma(k)$  soit absolument sommable, i.e.

$$\sum_{k=-\infty}^{\infty} |\gamma(k)| < \infty$$

Alors la série  $\sum_{k=-\infty}^{\infty} \gamma(k) \exp(i\omega k)$  est uniformément convergente et admet pour limite une fonction continue  $f(\omega)$ .

**Définition 2.2.1** On appelle densité spectrale la fonction  $f(\omega)$  définie telle que :

$$f(\omega) = \sum_{k=-\infty}^{\infty} \gamma(k) \exp(i\omega k) \quad \forall \omega \in [-\pi, \pi] \quad (2.1)$$

$f(\omega)$  est alors la transformation de Fourier discrète de la fonction d'autocovariance  $\gamma(k)$ .

## 2.2.2 Transformation de Toeplitz de la densité spectrale

D'après Grenander-Szegö ([15]) la fonction d'autocovariance d'un processus  $X_t$  centré et stationnaire peut être exprimée par l'intégrale de Fourier-Stieltjes telle que :

$$\gamma(k) = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(\omega) \exp(i\omega k) d\omega \quad (2.2)$$

2.2. REPRÉSENTATION SPECTRALE D'UN PROCESSUS  
STOCHASTIQUE

---

où  $f$  est la densité spectrale de  $X_t$ .

**Définition 2.2.2** La matrice d'autocovariance  $\Gamma_N = \{\gamma(k-j)\}_{k,j=1,2,\dots,N}$  d'un processus  $X_t$  est appelée transformée de Toeplitz de la densité spectrale  $f$  et est notée  $T_N(f)$ .

La matrice  $T_N(f)$  peut être exprimée sous la forme matricielle telle que :

$$T_N(f) = \Gamma_N = \begin{pmatrix} \gamma_0 & \gamma_1 & \gamma_2 & \cdots & \gamma_{N-1} \\ \gamma_1 & \gamma_0 & \gamma_1 & \cdots & \gamma_{N-2} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ \gamma_{N-1} & \gamma_{N-2} & \gamma_{N-3} & \cdots & \gamma_0 \end{pmatrix}$$

Il est clair que la transformée de Toeplitz  $T_N(f)$  est une troncature de la fonction d'autocovariance du processus  $X_t$ . La matrice de Toeplitz est définie positive et symétrique, alors il existe une matrice orthogonale  $U_N$ , i.e.  $U_N^t = U_N^{-1}$ , telle que :

$$U_N = \{N^{-1/2}e^{i2\pi kjN}\}_{k,j=1,2,\dots,N}$$

et dont les colonnes sont les vecteurs propres de  $T_N(f)$  tel que :

$$T_N(f) = U_N^t L_N U_N \quad (2.3)$$

où  $L_N = \text{diag}(\lambda_1^{(N)}, \lambda_2^{(N)}, \dots, \lambda_N^{(N)})$  est une matrice diagonale et  $\lambda_j^{(N)}$ ,  $j = 1, 2, \dots, N$  sont les valeurs propres de  $T_N(f)$ .

Notons par  $D_N(f)$  le déterminant de la matrice de Toeplitz. Alors, il peut être exprimé tel que :

$$D_N(f) = \det(T_N(f)) = \prod_{k=1}^N \lambda_k^{(N)} \quad k = 1, 2, \dots, N \quad (2.4)$$

Par ailleurs, la transformée de Toeplitz vérifie la propriété suivante :

$$(T_N(f))^s = T_N^s(f) = T_N(f^s) \quad \forall s \in \mathbb{Z}$$

Pour  $s = -1$ ,  $T_N^{-1}(f)$  est la matrice inverse de la matrice de Toeplitz qui se calcule simplement de la façon suivante :

$$T_N^{-1}(f) = T_N(f^{-1}) = \frac{1}{2\pi} \int_{-\pi}^{\pi} \frac{e^{i(k-j)\omega}}{f(\omega)} d\omega \quad k, j = 1, 2, \dots, N \quad (2.5)$$

### 2.2.3 Périodogramme d'une série chronologique

Dans la pratique, la fonction d'autocovariance  $\gamma(k)$  est inconnue. Alors pour estimer la densité spectrale  $f(\omega)$ , on remplace dans la relation (2.1)  $\gamma(k)$  par son estimation  $\hat{\gamma}(k)$ . Ainsi, on obtient :

$$\hat{f}(\omega) = \sum_{k=-\infty}^{+\infty} \hat{\gamma}(k) \exp(ik\omega)$$

Cependant, l'estimation de  $f$  est assez laborieuse à déterminer. Pour cela, nous introduisons la définition suivante :

**Définition 2.2.3** *Le périodogramme d'un processus  $X_t$  pour une fréquence  $\omega_j = \frac{2\pi j}{N}$ ,  $-\pi \leq \omega_j \leq \pi$ , notée  $I_N(\omega_j)$ , est donné par la transformation de Fourier discrète de la variable aléatoire  $X_t$  telle que :*

$$I_N(\omega_j) = \frac{1}{N} \left| \sum_{k=1}^N X_k \exp(ik\omega_j) \right|^2$$

$I_N$  est souvent utilisé comme un estimateur de la densité spectrale  $f$ .

### 2.2.4 Fonction génératrice des autocovariances

D'après ([5]) tout processus aléatoire peut être exprimé sous forme d'un processus moyennes mobiles infinies tel que :

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

Par définition

$$\begin{aligned} \gamma(k) &= E(X_t X_{t+k}) = E\left(\sum_{j=0}^{\infty} \sum_{h=0}^{\infty} \psi_j \psi_h \varepsilon_{t-j} \varepsilon_{t+k-h}\right) \\ &= \sum_{j=0}^{\infty} \sum_{h=0}^{\infty} \psi_j \psi_h E(\varepsilon_{t-j} \varepsilon_{t+k-h}) = \sigma_\varepsilon^2 \sum_{j=0}^{\infty} \psi_j \psi_{j+k} \end{aligned}$$

Une autre manière d'obtenir la fonction d'autocovariance d'un processus linéaire par l'intermédiaire de la fonction génératrice des autocovariances définies dans ([5]) telle que :

$$\gamma(B) = \sum_{k=-\infty}^{\infty} \gamma_k B^k = \sigma_\varepsilon^2 \sum_{k=-\infty}^{\infty} \sum_{j=0}^{\infty} \psi_j \psi_{j+k} B^k$$

## 2.2. REPRÉSENTATION SPECTRALE D'UN PROCESSUS STOCHASTIQUE

---

Notons que  $\psi_l = 0$  pour  $l < 0$ . De plus, en utilisant le changement d'indice  $h = j + k$  dans la relation ci-dessus, on obtient :

$$\gamma(B) = \sigma_\varepsilon^2 \sum_{j=0}^{\infty} \sum_{k=0}^{\infty} \psi_j \psi_{j+k} B^k = \sigma_\varepsilon^2 \sum_{h=0}^{\infty} \psi_h B^h \sum_{j=0}^{\infty} \psi_j B^{-j}$$

D'où

$$\begin{aligned} \gamma(B) &= \sigma_\varepsilon^2 \psi(B) \psi(B^{-1}) \\ &= \sigma_\varepsilon^2 \psi(B) \psi(F) \end{aligned}$$

Finalement, la fonction génératrice des autocovariances prend la forme suivante :

$$\gamma(B) = \sigma_\varepsilon^2 \psi(B) \psi(F) \quad (2.6)$$

### 2.2.5 Calcul explicite de la densité spectrale

En substituant  $B$  par  $e^{-i\omega}$  et  $F = B^{-1}$  par  $e^{i\omega}$  dans la fonction génératrice des autocovariances (2.6) nous obtenons la forme générale de la densité spectrale d'un processus linéaire. La forme du spectre est alors donnée par :

$$f(\omega) = \frac{\sigma_\varepsilon^2}{2\pi} \psi(e^{-i\omega}) \psi(e^{i\omega}) = \frac{\sigma_\varepsilon^2}{2\pi} |\psi(e^{-i\omega})|^2 \quad -\pi \leq \omega \leq \pi \quad (2.7)$$

Pour un processus  $MA(q)$  on a :

$$\psi(B) = Q(B) = 1 - \theta_1 B - \theta_2 B^2 - \dots - \theta_q B^q$$

Alors la densité spectrale d'un processus  $MA(q)$  est obtenue en substituant dans la relation (2.7)

$$f(\omega) = \frac{\sigma_\varepsilon^2}{2\pi} |1 - \theta_1 e^{-i\omega} - \theta_2 e^{-i2\omega} - \dots - \theta_q e^{-iq\omega}|^2 \quad -\pi \leq \omega \leq \pi$$

Pour un processus  $AR(p)$  on a :

$$\psi(B) = P^{-1}(B)$$

où  $P(B) = 1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p$ .

D'après (2.7), la densité spectrale d'un processus  $AR(p)$  est alors :

$$f(\omega) = \frac{1}{2\pi} \frac{\sigma_\varepsilon^2}{|1 - \phi_1 e^{-i\omega} - \phi_2 e^{-i2\omega} - \dots - \phi_p e^{-ip\omega}|^2} \quad -\pi \leq \omega \leq \pi$$

Et pour un processus  $ARMA(p, q)$  on a :

$$\psi(B) = P^{-1}(B)Q(B)$$

D'où, d'après (2.7), le spectre d'un processus  $ARMA(p, q)$  s'exprime tel que :

$$f(\omega) = \frac{\sigma_\varepsilon^2 |Q(e^{-i\omega})|^2}{2\pi |P(e^{-i\omega})|^2}$$

ou bien

$$f(\omega) = \frac{\sigma_\varepsilon^2 |1 - \theta_1 e^{-i\omega} - \theta_2 e^{-i2\omega} - \dots - \theta_q e^{-iq\omega}|^2}{2\pi |1 - \phi_1 e^{-i\omega} - \phi_2 e^{-i2\omega} - \dots - \phi_p e^{-ip\omega}|^2} \quad \pi \leq \omega \leq \pi$$

## 2.3 Approximation de Whittle de la fonction de vraisemblance

Soit  $X = (X_1, X_2, \dots, X_N)' \in \mathbb{R}^N$  une série chronologique, avec  $X \sim N(0, \Gamma_N)$  où  $\Gamma_N$  est la matrice d'autocovariance de dimension  $N \times N$ . Il a été établi dans ([15]) que la matrice d'autocovariance peut être exprimée en fonction de la densité spectrale  $f$  du processus. En effet

$$\Gamma_N = [\gamma(i - j)] \quad i, j = 1, \dots, N$$

avec

$$\gamma(k) = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(\omega, \boldsymbol{\theta}) \exp(i\omega k) d\omega$$

où  $\boldsymbol{\theta} = (\phi_1, \dots, \phi_p, \theta_1, \dots, \theta_q, \sigma_\varepsilon^2)$  est le vecteur des paramètres du modèle. La fonction log-vraisemblance peut alors être exprimée telle que :

$$\ln L_N(\boldsymbol{\theta}) = -\frac{N}{2} \ln(2\pi) - \frac{1}{2} \ln \det \Gamma_N - \frac{1}{2} X^t \Gamma_N^{-1} X \quad (2.8)$$

L'approximation de la log-vraisemblance se fait en deux étapes. On détermine une approximation de  $\ln \det \Gamma_N$ , et ensuite une approximation de la forme quadratique  $\frac{1}{2} X^t \Gamma_N^{-1} X$ .

Pour l'approximation de  $\ln \det \Gamma_N$ , Grenander et Szegö ([15]) ont utilisé le théorème suivant :

2.3. APPROXIMATION DE WHITTLE DE LA FONCTION DE  
VRAISEMBLANCE

---

**Théorème 2.3.1** Soit  $f$  une fonction à valeurs réelles. Notons par  $m$  et  $M$  les limites finies inférieure et supérieure de  $f$ . Si  $F(\lambda)$  est une fonction continue et définie dans l'intervalle  $m \leq \lambda \leq M$ . Alors on a :

$$\lim_{N \rightarrow \infty} \frac{F(\lambda_1^{(N)}) + F(\lambda_2^{(N)}) + \dots + F(\lambda_N^{(N)})}{N} = \frac{1}{2\pi} \int_{-\pi}^{\pi} F[f(\omega)] d\omega \quad (2.9)$$

où  $\lambda_1^{(N)}, \lambda_2^{(N)}, \dots, \lambda_N^{(N)}$  sont les valeurs propres de la matrice de Toeplitz ([15]) que l'on note  $T_N(f)$ .

**Remarque 2.3.1** Une matrice dont les éléments sont les coefficients de Fourier d'une fonction  $f$  est appelée transformée de Toeplitz de la fonction  $f$ . En particulier, la matrice d'autocovariance  $\Gamma_N$  d'un processus stationnaire  $\{X_t\}$  est une matrice de Toeplitz notée  $T_N(f)$ .

**Corollaire 2.3.1** Soit  $m > 0$  et soient  $\lambda_k \geq m > 0, k = 1, 2, \dots, N$ . Notons par  $D_N(f)$  le déterminant de la matrice de Toeplitz de  $f(\omega, \theta)$ , alors :

$$\lim_{N \rightarrow \infty} \frac{1}{N} \ln(D_N(f)) = \frac{1}{2\pi} \int_{-\pi}^{\pi} \ln f(\omega, \theta) d\omega$$

Il est donc clair d'après (2.9) et en posant  $F(\lambda) = \ln(\lambda)$  que :

$$\lim_{N \rightarrow \infty} \frac{1}{N} \ln(D_N(f)) = \lim_{N \rightarrow \infty} \frac{1}{N} \ln \det \Gamma_N = \frac{1}{2\pi} \int_{-\pi}^{\pi} \ln f(\omega, \theta) d\omega \quad (2.10)$$

Pour déterminer l'approximation de la forme quadratique Grenander et Szegö ([15]) ont utilisé l'équation inverse de Toeplitz. ce qui résulte :

$$\begin{aligned} \frac{1}{2} X^t \Gamma_N^{-1} X &\approx \sum_{\ell=1}^N \sum_{j=1}^N X_{\ell} \left\{ \frac{1}{4\pi} \int_{-\pi}^{\pi} f^{-1}(\omega, \theta) \exp[i\omega(\ell - j)] d\omega \right\} X_j \\ &= \frac{1}{4\pi} \int_{-\pi}^{\pi} f^{-1}(\omega, \theta) \sum_{\ell=1}^N \sum_{j=1}^N X_{\ell} X_j \exp[i\omega(\ell - j)] d\omega \\ &= \frac{1}{4\pi} \int_{-\pi}^{\pi} f^{-1}(\omega, \theta) \left| \sum_{j=1}^N X_j \exp(i\omega j) \right|^2 d\omega \end{aligned}$$

D'où

$$\frac{1}{2}X^t\Gamma_N^{-1}X \approx \frac{N}{4\pi} \int_{-\pi}^{\pi} \frac{I_N(\omega)}{f(\omega, \boldsymbol{\theta})} d\omega \quad (2.11)$$

où  $I_N(\omega) = \frac{1}{N} \left| \sum_{j=1}^N X_j \exp(i\omega_j) \right|^2$  est le périodogramme de la série chronologique  $X_1, X_2, \dots, X_N$ , où  $\omega_j = \frac{2\pi j}{N}$  les fréquences aux points  $j = 1, \dots, \frac{N-1}{2}$  et  $f(\omega, \boldsymbol{\theta})$  la densité spectrale du processus.

En remplaçant dans l'équation (2.8) les résultats (2.10) et (2.11), on obtient :

$$\ln L_N^W(\boldsymbol{\theta}) = \frac{-N}{2} \ln(2\pi) - \frac{N}{4\pi} \int_{-\pi}^{\pi} \left[ \ln f(\omega, \boldsymbol{\theta}) + \frac{I_N(\omega)}{f(\omega, \boldsymbol{\theta})} \right] d\omega$$

Finalement, l'estimateur de Whittle est obtenu en minimisant  $\ln L_N^W(\boldsymbol{\theta})$ . D'où

$$\frac{\partial \ln L_N^W(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} = \frac{\partial}{\partial \boldsymbol{\theta}} \left\{ \frac{-N}{4\pi} \int_{-\pi}^{\pi} \left[ \ln f(\omega, \boldsymbol{\theta}) + \frac{I_N(\omega)}{f(\omega, \boldsymbol{\theta})} \right] d\omega \right\}$$

Ceci revient à chercher la solution d'un système de  $M$  équations non linéaires :

$$\begin{aligned} \frac{\partial}{\partial \theta_k} \ln L_N^W(\boldsymbol{\theta})|_{\boldsymbol{\theta}=\hat{\boldsymbol{\theta}}} &= -\frac{N}{4\pi} \frac{\partial}{\partial \theta_k} \int_{-\pi}^{\pi} \ln f(\omega, \boldsymbol{\theta}) d\omega - \frac{N}{4\pi} \frac{\partial}{\partial \theta_k} \int_{-\pi}^{\pi} \frac{I_N(\omega)}{f(\omega, \boldsymbol{\theta})} d\omega = 0 \\ k &= 1, 2, \dots, p + q + 1 = M \end{aligned}$$

En notant  $\boldsymbol{\theta}^* = (1, \phi_1, \dots, \phi_p, \theta_1, \dots, \theta_q)^t$  et en choisissant le paramètre d'échelle  $\sigma_\varepsilon^2$  tel que :

$$f(\omega, \boldsymbol{\theta}) = \sigma_\varepsilon^2 f(\omega, \boldsymbol{\theta}^*)$$

alors, d'après ([18]) on a :

$$\int_{-\pi}^{\pi} \ln f(\omega, \boldsymbol{\theta}^*) d\omega = 0$$

La fonction log-vraisemblance de Whittle  $\ln L_N^W(\boldsymbol{\theta})$  s'écrit plus simplement sous la forme suivante :

$$\ln L_N^W(\boldsymbol{\theta}^*) = -\frac{N}{2} \ln(\sigma_\varepsilon^2) - \frac{N}{4\pi\sigma_\varepsilon^2} \int_{-\pi}^{\pi} \frac{I_N(\omega)}{f(\omega, \boldsymbol{\theta}^*)} d\omega$$

### 2.3. APPROXIMATION DE WHITTLE DE LA FONCTION DE VRAISEMBLANCE

---

L'estimation de la log-vraisemblance exige le calcul des intégrales  $\int_{-\pi}^{\pi} \frac{I_N(\omega)}{f(\omega, \boldsymbol{\theta}^*)} d\omega$ .  
Pour simplifier ces calculs, les intégrales peuvent être remplacées approximativement par les sommes de Riemann telle que :

$$\int_{-\pi}^{\pi} \frac{I_N(\omega)}{f(\omega, \boldsymbol{\theta}^*)} d\omega \approx \frac{2\pi}{N} \sum_{j=1}^N \frac{I_N(\omega_j)}{f(\omega_j, \boldsymbol{\theta}^*)}$$

Donc

$$\ln L_N^W(\boldsymbol{\theta}^*) = -\frac{N}{2} \ln \sigma_\varepsilon^2 - \frac{1}{2\sigma_\varepsilon^2} \sum_{j=1}^N \frac{I_N(\omega_j)}{f(\omega_j, \boldsymbol{\theta}^*)} \quad (2.12)$$

Nous obtenons l'estimateur de Whittle  $\hat{\boldsymbol{\theta}}$  en minimisant la fonction  $L_N^W(\boldsymbol{\theta}^*)$  par rapport à  $\boldsymbol{\theta}^*$  et à  $\sigma_\varepsilon^2$ . Cela revient pour un échantillon fini à minimiser par rapport à  $\boldsymbol{\theta}^*$  la fonction suivante :

$$Q_N(\boldsymbol{\theta}^*) = \sum_{j=1}^N \frac{I_N(\omega_j)}{f(\omega_j, \boldsymbol{\theta}^*)}$$

La résolution de l'équation non linéaire (2.12) se fait en deux étapes. D'abord on résout l'équation suivante pour trouver l'estimation des paramètres  $\boldsymbol{\theta}^*$

$$\frac{\partial}{\partial \boldsymbol{\theta}^*} Q_N(\boldsymbol{\theta}^*) = 0$$

Ensuite l'estimateur de  $\sigma_\varepsilon^2$  est obtenu en minimisant l'équation suivante par rapport à  $\sigma_\varepsilon^2$  :

$$\ln L_N^W(\boldsymbol{\theta}^*) = -\frac{N}{2} \ln \sigma_\varepsilon^2 - \frac{1}{2\sigma_\varepsilon^2} Q_N(\hat{\boldsymbol{\theta}}^*)$$

Alors on obtient :

$$\hat{\sigma}_\varepsilon^2 = \frac{1}{N} Q(\hat{\boldsymbol{\theta}}^*)$$

## Chapitre 3

# Approximation de Box-Jenkins de la fonction de vraisemblance d'un processus gaussien

Il a été établi que la fonction de vraisemblance est dominée par la forme quadratique  $X^t \Gamma_N^{-1} X$ , le problème auquel nous sommes confrontés se situe dans la difficulté de manipuler l'inverse de la matrice d'autocovariance  $\Gamma_N^{-1}$ . Nous allons démontrer que la somme des carrés exacte des innovations  $\sum_{k=-\infty}^N \widehat{\varepsilon}_k^2$  est une bonne approximation de la forme quadratique de la vraisemblance. Box et Jenkins ont proposé une expression de la fonction de vraisemblance à l'aide des espérances conditionnelles des innovations  $[\varepsilon_j] = \widehat{\varepsilon}_j = E(\varepsilon_j / X_1, X_2, \dots, X_N)$ ,  $j \in \mathbb{Z}$  du processus qu'ils ont appelé fonction de vraisemblance inconditionnelle. Celle-ci est proportionnelle à la somme des carrés des innovations de l'échantillon  $\sum_{j=1}^N \widehat{\varepsilon}_j^2$ . Nous allons essayer d'exprimer explicitement la fonction de vraisemblance inconditionnelle d'une série chronologique dans le cas d'un modèle  $AR(p)$ , d'un modèle  $MA(q)$  et d'un modèle  $ARMA(p, q)$ . Dans ce chapitre, nous présentons une synthèse des résultats des travaux de Box-Jenkins ([5]), Dacunha-Castelle et Azencot ([3]), Brockwell et Davies ([7]) et Hamilton ([16]) principalement.

### 3.1 Forme innovation de la fonction de vraisemblance d'un processus aléatoire gaussien

Soit  $X_1, X_2, \dots, X_N$  une série chronologique identifiée par un processus  $ARMA(p, q)$  d'innovation  $\varepsilon$ . Le vecteur  $X = (X_1, X_2, \dots, X_N)^t$  a donc pour densité sur  $\mathbb{R}^N$  :

$$f_N(x_1, x_2, \dots, x_N, \boldsymbol{\theta}) = (2\pi)^{-N/2} (\det \Gamma_N)^{-1/2} \exp\left(-\frac{1}{2} X^t \Gamma_N^{-1} X\right) \quad (3.1)$$

où  $\Gamma_n$  est la matrice d'autocovariance de  $X$ .

D'après ([5]) tout processus aléatoire peut être exprimé sous forme d'un processus moyennes mobiles infinies. Ainsi, il existe une suite  $(\psi_k)_k$  telle que  $\sum_{k \geq 0} |\psi_k| < \infty$  pour tout  $t \in \mathbb{Z}$  et tel que :

$$X_N = \sum_{k \geq 0} \psi_k \varepsilon_{N-k}$$

**Lemme 3.1.1** *Soit  $\widehat{\varepsilon}_k = E(\varepsilon_k / X_1, X_2, \dots, X_N)$  la régression linéaire de  $\varepsilon_k$  sur  $X_1, X_2, \dots, X_N$  alors*

$$X^t \Gamma_N^{-1} X = \frac{1}{\sigma_\varepsilon^2} \sum_{k=-\infty}^N \widehat{\varepsilon}_k^2 \quad (3.2)$$

**Preuve** On pose  $\widehat{\varepsilon}_k = G_k^t X = X^t G_k$  où  $G_k$  est une matrice de dimension  $N \times 1$ . De plus, on a  $X = \sum_{k=-\infty}^N M_k \varepsilon_k$  où  $M_k$  est une matrice de dimension  $N \times 1$  définie à partir de la représentation  $X_k$ . D'une part, pour tout  $k$ ,  $0 \leq k \leq N$ , on a :

$$\sigma_\varepsilon^2 M_k = E(X \varepsilon_k) = E(X \widehat{\varepsilon}_k) = E(X X^t G_k) = \Gamma_N G_k$$

D'autre part

$$\Gamma_N = E(X X^t) = E\left[\left(\sum_{k=-\infty}^N M_k \varepsilon_k\right) \left(\sum_{k=-\infty}^N \varepsilon_k M_k^t\right)\right] = \sigma_\varepsilon^2 \sum_{k=-\infty}^N M_k M_k^t$$

Donc

$$\sigma_\varepsilon^2 \Gamma_N = \sum_{k=-\infty}^N \Gamma_N G_k G_k^t \Gamma_N$$

et

$$\sigma_\varepsilon^2 \Gamma_N^{-1} = \sum_{k=-\infty}^N G_k G_k^t$$

Finalement

$$X^t \Gamma_N^{-1} X = \frac{1}{\sigma_\varepsilon^2} \sum_{k=-\infty}^N X^t G_k G_k^t X = \frac{1}{\sigma_\varepsilon^2} \sum_{k=-\infty}^N \widehat{\varepsilon}_k^2$$

■

Nous admettons sans démonstration, le lemme suivant :

**Lemme 3.1.2** *Soit  $X_1, X_2, \dots, X_N$  une suite d'observations d'un processus  $X_t$  de matrice d'autocovariance  $\Gamma_N$ , et soit  $\varepsilon_t$  le processus d'innovation de  $X_t$ . Alors, on a :*

$$\ln \sigma_\varepsilon^2 = \lim_{N \rightarrow \infty} \frac{1}{N} \ln \det \Gamma_N$$

La log-vraisemblance de  $X$  peut alors être exprimée telle que :

$$\ln L_N(\boldsymbol{\theta}) = -\frac{N}{2} \ln(2\pi\sigma_\varepsilon^2) - \frac{1}{2\sigma_\varepsilon^2} S(\boldsymbol{\theta})$$

où  $S(\boldsymbol{\theta}) = \sum_{k=-\infty}^N \widehat{\varepsilon}_k^2$ .

## 3.2 Approximation de la fonction de vraisemblance d'un processus autorégressif

Supposons que la série chronologique  $X_1, X_2, \dots, X_N$  soit identifiée comme étant un modèle  $AR(p)$  défini tel que :

$$X_t - \phi_1 X_{t-1} - \phi_2 X_{t-2} - \dots - \phi_p X_{t-p} = \varepsilon_t \quad (3.3)$$

où  $\varepsilon_t$  est un bruit blanc de loi normale telle que  $E(\varepsilon_t) = 0$  et  $E(\varepsilon_t^2) = \sigma_\varepsilon^2$ . Soit  $\phi = (\phi_1, \phi_2, \dots, \phi_p, \sigma_\varepsilon^2)^t$  le vecteur des paramètres du modèle. Soit  $\Gamma_N$  la matrice d'autocovariance du vecteur  $X = (X_1, X_2, \dots, X_N)^t$ . La densité de probabilité jointe de  $X$  est alors définie telle que :

$$f(X/\phi, \sigma_\varepsilon) = (2\pi)^{-N/2} |\Gamma_N^{-1}|^{\frac{1}{2}} \exp \left\{ -\frac{1}{2} X^t \Gamma_N^{-1} X \right\} \quad (3.4)$$

### 3.2. APPROXIMATION DE LA FONCTION DE VRAISEMBLANCE D'UN PROCESSUS AUTORÉGRESSIF

---

En notant  $X^{(p)} = (X_1, X_2, \dots, X_p)^t$ , d'après l'axiome des probabilités conditionnelles, on peut écrire la relation (3.4) sous la forme :

$$f(X/\phi, \sigma_\varepsilon) = f(X_{p+1}, \dots, X_N/X^{(p)}, \phi, \sigma_\varepsilon) \times f(X^{(p)}/\phi, \sigma_\varepsilon)$$

où  $f(X^{(p)}/\phi, \sigma_\varepsilon)$  est la densité jointe des  $p$  premières observations de  $X$  que l'on peut exprimer telle que :

$$f(X^{(p)}/\phi, \sigma_\varepsilon) = (2\pi)^{-p/2} |\Gamma_p^{-1}|^{\frac{1}{2}} \exp \left\{ \frac{-1}{2} X^{(p)t} \Gamma_p^{-1} X^{(p)} \right\}$$

où  $\Gamma_p$  est la matrice d'autocovariance de  $X_1, X_2, \dots, X_p$ .

Pour expliciter la densité jointe  $f(X_{p+1}, \dots, X_N/X^{(p)}, \phi, \sigma_\varepsilon)$  des observations restantes de l'échantillon  $(X_{p+1}, \dots, X_N)$ , il est possible d'utiliser la relation linéaire (3.3) entre les  $X_t$  et les  $\varepsilon_t$ . On exprime d'abord la densité jointe en fonction des  $\varepsilon_{p+1}, \dots, \varepsilon_N$  telle que :

$$f(\varepsilon_{p+1}, \dots, \varepsilon_N) = (2\pi\sigma_\varepsilon^2)^{-(N-p)/2} \exp \left\{ \frac{-1}{2\sigma_\varepsilon^2} \sum_{t=p+1}^N \varepsilon_t^2 \right\}$$

Ensuite, en remplaçant dans l'équation ci-dessus les  $\varepsilon_t$  par leur valeur dans la relation linéaire (3.3), on abouti à la densité jointe de  $(X_{p+1}, \dots, X_N)$  sous la forme :

$$f(X_{p+1}, \dots, X_N/X_p, \phi, \sigma_\varepsilon) = (2\pi\sigma_\varepsilon^2)^{-(N-p)/2} \times \exp \left\{ \frac{-1}{2\sigma_\varepsilon^2} \sum_{t=p+1}^N (X_t - \phi_1 X_{t-1} - \dots - \phi_p X_{t-p})^2 \right\}$$

Finalement, la fonction de la vraisemblance  $L_N(X/\phi, \sigma_\varepsilon)$  s'écrit alors telle que :

$$L_N(X/\phi, \sigma_\varepsilon) = (2\pi\sigma_\varepsilon^2)^{-N/2} |\Gamma_p^{-1}|^{\frac{1}{2}} \exp \left\{ \frac{-1}{2\sigma_\varepsilon^2} S(\phi) \right\}$$

où

$$S(\phi) = \sum_{i=1}^p \sum_{j=1}^p m_{ij}^{(p)} X_i X_j + \sum_{t=p+1}^N (X_t - \phi_1 X_{t-1} - \dots - \phi_p X_{t-p})^2$$

D'après ([16]), les éléments  $m_{ij}^{(p)}$  de la matrice inverse  $\Gamma_p^{-1}$  peuvent être calculés par la formule suivante :

$$m^{ij}(p) = \left\{ \sum_{k=0}^{i-1} \phi_k \phi_{k+j-i} - \sum_{k=p+1-j}^{p+i-j} \phi_k \phi_{k+j-i} \right\} \quad 1 \leq i \leq j \leq p$$

où  $\phi_0 \equiv -1$ .

On ne calcule que les valeurs de  $m^{ij}(p)$  pour  $i > j$  car  $\Gamma_p^{-1}$  est symétrique et donc  $m_{ij}^{(p)} = m_{ji}^{(p)}$ .

La fonction log-vraisemblance  $\ln L_N(\theta)$  s'écrit comme suit :

$$\ln L_N(\phi) = \frac{-N}{2} \ln(2\pi\sigma_\varepsilon^2) - \frac{1}{2} \ln |\Gamma_p| - \frac{1}{2\sigma_\varepsilon^2} S(\phi)$$

L'expression de la vraisemblance d'un processus *AR* peut être obtenue assez aisément. Le problème d'estimation des paramètres par la méthode du maximum de vraisemblance est résolu à l'aide des équations de Yule-Walker ([3], [5]).

### 3.3 Approximation de la fonction de vraisemblance d'un processus en moyennes mobiles

Supposons qu'on associe à la série chronologique  $X_1, X_2, \dots, X_N$  un modèle *MA*( $q$ ) défini tel que :

$$X_t = \varepsilon_t - \theta_1\varepsilon_{t-1} - \theta_2\varepsilon_{t-2} - \dots - \theta_q\varepsilon_{t-q} \quad (3.5)$$

où  $\varepsilon_t$  est un bruit blanc de loi  $N(0, \sigma_\varepsilon^2)$ .

Soit  $\theta = (\theta_1, \theta_2, \dots, \theta_q, \sigma_\varepsilon^2)^t$  le vecteur des paramètres du modèle. Alors la densité jointe du vecteur  $X = (X_1, X_2, \dots, X_N)^t$  peut être exprimée telle que :

$$f(X/\theta, \sigma_\varepsilon) = (2\pi)^{-N/2} |\Gamma_N|^{-1/2} \exp \left\{ \frac{-1}{2} X^t \Gamma_N^{-1} X \right\}$$

où  $\Gamma_N$  est la matrice des autocovariances de  $X = (X_1, X_2, \dots, X_N)^t$  pour un processus *MA*( $q$ ).

D'après l'équation (3.2) du lemme 3.1.1, la forme quadratique dans la fonction de vraisemblance est égale à la somme des carrés exacte

$S(\theta) = \sum_{k=-\infty}^N \widehat{\varepsilon}_k^2$ . Par ailleurs, il a été établi dans ([5], [3]) que pour  $N$  assez grand, la fonction de vraisemblance est dominée par le terme  $\frac{1}{2\sigma_\varepsilon^2} S(\theta)$ , et

### 3.3. APPROXIMATION DE LA FONCTION DE VRAISEMBLANCE D'UN PROCESSUS EN MOYENNES MOBILES

---

ainsi les estimateurs des paramètres  $\theta_1, \theta_2, \dots, \theta_q, \sigma_\varepsilon^2$  sont obtenus en maximisant la log-vraisemblance ou ce qui revient au même, en minimisant la somme des carrés  $S(\theta)$ .

De plus, il a été établi dans ([3], [12]) que la somme des carrés exacte  $S(\theta)$  est équivalente à la somme des carrés  $\sum_{t=1-q}^N \widehat{\varepsilon}_t^2$ , appelée somme des carrés inconditionnelle ([5]).

A l'aide de l'équation (3.5) et pour  $t = 1, 2, \dots, N$ , on obtient un système d'équations. En posant  $\varepsilon = (\varepsilon_1, \varepsilon_2, \dots, \varepsilon_N)^t$  et  $\varepsilon^{(0)} = (\varepsilon_{1-q}, \varepsilon_{2-q}, \dots, \varepsilon_0)^t$  le vecteur des  $q$  valeurs initiales, ce système d'équations peut être écrit sous la forme matricielle suivante :

$$X = L_\theta \varepsilon + F \varepsilon^{(0)} \quad (3.6)$$

où  $L_\theta$  est une matrice de dimension  $N \times N$  triangulaire inférieure et  $F$  est une matrice de dimension  $N \times q$ , définies telles que :

$$L_\theta = \begin{pmatrix} 1 & 0 & \dots & \dots & \dots & \dots & \dots & 0 \\ -\theta_1 & 1 & 0 & \dots & \dots & \dots & \dots & 0 \\ -\theta_2 & -\theta_1 & 1 & 0 & \dots & \dots & \dots & 0 \\ \dots & \dots \\ -\theta_q & -\theta_{q-1} & \dots & -\theta_1 & 1 & 0 & \dots & 0 \\ 0 & 0 & \dots & \dots & 0 & 1 & \dots & 0 \\ \dots & 0 \\ 0 & \dots & \dots & \dots & \dots & \dots & 0 & 1 \end{pmatrix}$$

et

$$F = (\Theta^t, \mathbf{0}^t)^t$$

où

$$\Theta = - \begin{pmatrix} \theta_q & \theta_{q-1} & \dots & \theta_1 \\ 0 & \theta_q & \dots & \theta_2 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & \theta_q \end{pmatrix}$$

La distribution jointe du vecteur  $(\varepsilon^t, \varepsilon^{(0)t})$  de taille  $N \times q$  peut s'exprimer telle que :

$$f(\varepsilon^t, \varepsilon^{(0)t} / \sigma_\varepsilon) = (2\pi\sigma_\varepsilon^2)^{-(N+q)/2} \exp \left\{ \frac{-1}{2\sigma_\varepsilon^2} \left( \varepsilon^t \varepsilon + \varepsilon^{(0)t} \varepsilon^{(0)} \right) \right\}$$

*CHAPITRE 3. APPROXIMATION DE BOX-JENKINS DE LA  
FONCTION DE VRAISEMBLANCE D'UN PROCESSUS GAUSSIEN*

---

Considérons la transformation linéaire  $(\varepsilon, \varepsilon^{(0)}) \longrightarrow (X, \varepsilon^{(0)})$  de jacobian unité avec  $\varepsilon = L_{\theta}^{-1} (X - F\varepsilon^{(0)})$ . Alors la distribution jointe de  $(X, \varepsilon^{(0)})$  est :

$$f(X, \varepsilon^{(0)} / \theta, \sigma_{\varepsilon}) = (2\pi\sigma_{\varepsilon}^2)^{-(N+q)/2} \exp \left\{ \frac{-1}{2\sigma_{\varepsilon}^2} S(\theta, \varepsilon^{(0)}) \right\}$$

où

$$S(\theta, \varepsilon^{(0)}) = (X - F\varepsilon^{(0)})^t (L_{\theta}^t)^{-1} L_{\theta}^{-1} (X - F\varepsilon^{(0)}) + \varepsilon^{(0)t} \varepsilon^{(0)} \quad (3.7)$$

On remarque que  $S(\theta, \varepsilon^{(0)})$  dépend du vecteur des valeurs initiales  $\varepsilon^{(0)}$  à déterminer.

Pour cela, notons par  $\widehat{\varepsilon}^{(0)}$  le vecteur qui minimise  $S(\theta, \varepsilon^{(0)})$ . Et en utilisant la méthode des moindres carrés généralisée on trouve :

$$\widehat{\varepsilon}^{(0)} = D^{-1} F^t (L_{\theta}^t)^{-1} L_{\theta}^{-1} X \quad (3.8)$$

où

$$D = I_q + F^t (L_{\theta}^t)^{-1} L_{\theta}^{-1} F \quad (3.9)$$

Par ailleurs, la somme des carrés peut être factorisée telle que ([5]) :

$$S(\theta, \varepsilon^{(0)}) = S(\theta, \widehat{\varepsilon}^{(0)}) + (\varepsilon^{(0)} - \widehat{\varepsilon}^{(0)})^t D (\varepsilon^{(0)} - \widehat{\varepsilon}^{(0)})$$

En utilisant les relations (3.7), (3.8) et (3.9) on obtient :

$$\begin{aligned} S(\theta, \widehat{\varepsilon}^{(0)}) &= (X - F\widehat{\varepsilon}^{(0)})^t (L_{\theta}^t)^{-1} L_{\theta}^{-1} (X - F\widehat{\varepsilon}^{(0)}) \\ &= X^t (L_{\theta}^t)^{-1} \left( I_n - L_{\theta}^{-1} F D^{-1} F^t (L_{\theta}^t)^{-1} \right) L_{\theta}^{-1} X \end{aligned}$$

Il est clair que  $S(\theta, \widehat{\varepsilon}^{(0)})$  est une fonction des observations, alors la distribution jointe de  $X$  et  $\varepsilon^{(0)}$  est donnée par :

$$\begin{aligned} f(X, \varepsilon^{(0)} / \theta, \sigma_{\varepsilon}) &= (2\pi\sigma_{\varepsilon}^2)^{-(N+q)/2} \\ &\exp \left\{ \frac{-1}{2\sigma_{\varepsilon}^2} \left[ S(\theta, \widehat{\varepsilon}^{(0)}) + (\varepsilon^{(0)} - \widehat{\varepsilon}^{(0)})^t D (\varepsilon^{(0)} - \widehat{\varepsilon}^{(0)}) \right] \right\} \end{aligned}$$

### 3.4. APPROXIMATION DE LA FONCTION DE VRAISEMBLANCE D'UN PROCESSUS AUTORÉGRESSIF-MOYENNES MOBILES

---

Par ailleurs, sachant que  $f(X, \varepsilon^{(0)}/\theta, \sigma_\varepsilon) = f(X, /\theta, \sigma_\varepsilon) \times f(\varepsilon^{(0)}/X, \theta, \sigma_\varepsilon)$ , il est clair que l'on a :

$$f(\varepsilon^{(0)}/X, \theta, \sigma_\varepsilon) = (2\pi\sigma_\varepsilon^2)^{-q/2} |D|^{1/2} \exp \left\{ \frac{-1}{2\sigma_\varepsilon^2} \left( \varepsilon^{(0)} - \widehat{\varepsilon}^{(0)} \right)^t D \left( \varepsilon^{(0)} - \widehat{\varepsilon}^{(0)} \right) \right\}$$

et

$$f(X/\theta, \sigma_\varepsilon) = (2\pi\sigma_\varepsilon^2)^{-N/2} |D|^{-1/2} \exp \left\{ \frac{-1}{2\sigma_\varepsilon^2} S \left( \theta, \widehat{\varepsilon}^{(0)} \right) \right\}$$

D'autre part, si on note  $\widehat{\varepsilon}^{(0)} = E[\varepsilon^{(0)}/X, \theta] = [\varepsilon^{(0)}]$  l'espérance conditionnelle de  $\varepsilon^{(0)}$  sachant  $(X, \theta)$ , alors d'après la relation (3.6),  $[\varepsilon] = L_\theta^{-1}(X - F[\varepsilon^{(0)}])$  est l'espérance conditionnelle de  $\varepsilon$  sachant  $(X, \theta)$ , d'où on obtient :

$$S \left( \theta, \widehat{\varepsilon}^{(0)} \right) = [\varepsilon^t] [\varepsilon] + [\varepsilon^{(0)t}] [\varepsilon^{(0)}] = \sum_{t=1-q}^N [\varepsilon_t]^2$$

Finalement la fonction de vraisemblance inconditionnelle est donnée par :

$$L_N(\theta, \sigma_\varepsilon/X) = (2\pi\sigma_\varepsilon^2)^{-N/2} |D|^{-1/2} \exp \left\{ \frac{-1}{2\sigma_\varepsilon^2} \sum_{t=1-q}^N [\varepsilon_t]^2 \right\}$$

Il est évident que pour calculer la somme des carrés inconditionnelle  $S(\theta, \widehat{\varepsilon}^{(0)})$  on a besoin de déterminer  $q$  valeurs initiales  $(\varepsilon_0, \varepsilon_{-1}, \dots, \varepsilon_{1-q})$  en utilisant (3.8) ou bien en utilisant le principe de la "backcast" proposé par Box-Jenkins ([5], [3], [12]).

## 3.4 Approximation de la fonction de vraisemblance d'un processus autorégressif-moyennes mobiles

Soit  $X_1, X_2, \dots, X_N$  une série chronologique identifiée par un processus  $ARMA(p, q)$ . Donc elle vérifie la relation :

$$P(B) X_t = Q(B) \varepsilon_t \tag{3.10}$$

où  $P(B)$  et  $Q(B)$  sont deux polynômes en  $B$  de degrés respectifs  $p$  et  $q$ . La relation ci-dessus peut être écrite sous la forme matricielle comme suit :

$$L_\phi X = L_\theta \varepsilon + F e_*$$

*CHAPITRE 3. APPROXIMATION DE BOX-JENKINS DE LA  
FONCTION DE VRAISEMBLANCE D'UN PROCESSUS GAUSSIEN*

---

$L_\theta$  est la matrice triangulaire inférieure de dimension  $N \times N$  avec 1 sur la diagonale principale,  $-\theta_1$  sur la première subdiagonale,  $-\theta_2$  sur la deuxième subdiagonale,  $\theta_i = 0$  pour  $i > q$  et  $L_\phi$  est une matrice de  $(N \times N)$  qui a la même forme de  $L_\theta$  mais avec  $\phi_i$  à la place de  $\theta_i$ .

$e_*^t = \left( X^{(0)t}, \varepsilon^{(0)t} \right)^t = (X_{1-p}, \dots, X_0, \varepsilon_{1-q}, \dots, \varepsilon_0)^t$  est le vecteur de dimension  $p + q$  des valeurs initiales et  $F$  est la matrice définie par :

$$F = \begin{bmatrix} \tilde{\Phi} & \Theta \\ 0 & 0 \end{bmatrix}$$

Avec

$$\tilde{\Phi} = \begin{bmatrix} \phi_p & \phi_{p-1} & \dots & \dots & \phi_1 \\ 0 & \phi_p & \dots & \dots & \phi_2 \\ \dots & \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & \dots & \phi_p \end{bmatrix}$$

et

$$\Theta = - \begin{bmatrix} \theta_q & \theta_{q-1} & \dots & \dots & \theta_1 \\ 0 & \theta_q & \dots & \dots & \theta_2 \\ \dots & \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & \dots & \theta_q \end{bmatrix}$$

La matrice  $\Omega$  est la matrice d'autocovariance de  $e_*$  qui s'exprime sous la forme suivante :

$$\Omega = \begin{bmatrix} \sigma_\varepsilon^{-2} J_p & C^t \\ C & I_q \end{bmatrix}$$

où  $J_p = E \left[ X^{(0)} X^{(0)t} \right]$  est la matrice de dimension  $p \times p$  de terme général  $\gamma_{|i-j|}^{(0)}$ , et où  $\sigma_\varepsilon^2 C = E \left[ \varepsilon^{(0)} X^{(0)t} \right]$  une matrice ayant pour terme général

$$E [\varepsilon_{i-q}, X_{j-q}] = \begin{cases} \sigma_\varepsilon^2 \psi_{j-i-p-q} & \text{si } j - i - p - q \geq 0 \\ 0 & \text{sinon} \end{cases}$$

Les  $\psi_k$  sont les coefficients de la forme moyennes mobiles infinies d'un  $ARMA(p, q)$ , i.e.

$$X_t = \psi(B) \varepsilon_t = \phi^{-1}(B) \theta(B) \varepsilon_t = \sum_{k=0}^{\infty} \psi_k \varepsilon_{t-k}$$

### 3.4. APPROXIMATION DE LA FONCTION DE VRAISEMBLANCE D'UN PROCESSUS AUTORÉGRESSIF-MOYENNES MOBILES

---

Puisque  $\varepsilon = L_\theta^{-1}(L_\phi X - Fe_*)$  et le vecteur  $e_*$  sont indépendants, alors la distribution jointe de  $X$  et  $e_*$  est

$$f(X, e_*/\phi, \theta, \sigma_\varepsilon) = (2\pi\sigma_\varepsilon^2)^{-(N+p+q)/2} |\Omega|^{-\frac{1}{2}} \exp \left[ \frac{-1}{2\sigma_\varepsilon^2} S(\phi, \theta, e_*) \right]$$

où

$$S(\phi, \theta, e_*) = (L_\phi X - Fe_*)^t (L_\theta^t)^{-1} L_\theta^{-1} (L_\phi X - Fe_*) + e_*^t \Omega^{-1} e_*$$

D'après la méthode des moindres carrés généralisée ([5]), nous pouvons écrire

$$S(\phi, \theta, e_*) = S(\phi, \theta) + (e_* - \hat{e}_*)^t D (e_* - \hat{e}_*)$$

où

$$\begin{aligned} S(\phi, \theta) &= S(\phi, \theta, \hat{e}_*) = \hat{\varepsilon}^t \hat{\varepsilon} + \hat{e}_*^t \Omega^{-1} \hat{e}_* \\ &= \sum_{t=1}^n [\varepsilon_t]^2 + \hat{e}_*^t \Omega^{-1} \hat{e}_* \end{aligned} \quad (3.11)$$

Par ailleurs

$$\hat{e}_* = E(e_*/X^{(0)}, \varepsilon^{(0)}) = [e_*] = D^{-1} F^t (L_\theta^t)^{-1} L_\theta^{-1} L_\phi X \quad (3.12)$$

où  $D = \Omega^{-1} + F^t (L_\theta^t)^{-1} L_\theta^{-1} F$

De plus

$$\hat{\varepsilon} = E(\varepsilon/X^{(0)}, \varepsilon^{(0)}) = [\varepsilon] = L_\theta^{-1} (L_\phi X - F\hat{e}_*).$$

D'autre part, sachant que  $f(X, e_*/\phi, \theta, \sigma_\varepsilon) = f(X/\phi, \theta, \sigma_\varepsilon) \times f(e_*/X, \phi, \theta, \sigma_\varepsilon)$ , il est clair que l'on a :

$$f(X/\phi, \theta, \sigma_\varepsilon) = (2\pi\sigma_\varepsilon^2)^{-N/2} |\Omega|^{-\frac{1}{2}} |D|^{-\frac{1}{2}} \exp \left[ \frac{-1}{2\sigma_\varepsilon^2} S(\phi, \theta) \right]$$

Alors, la fonction de la vraisemblance est définie par :

$$L_N(X/\phi, \theta, \sigma_\varepsilon) = (2\pi\sigma_\varepsilon^2)^{-N/2} |\Omega|^{-\frac{1}{2}} |D|^{-\frac{1}{2}} \exp \left[ \frac{-1}{2\sigma_\varepsilon^2} S(\phi, \theta) \right]$$

Finalement, la log-vraisemblance d'un  $ARMA(p, q)$  peut être exprimée telle que :

$$\ln L_N(X/\phi, \theta, \sigma_\varepsilon) = \frac{-N}{2} \ln(2\pi\sigma_\varepsilon^2) - \frac{1}{2} \ln |\Omega| |D| - \frac{1}{2\sigma_\varepsilon^2} S(\phi, \theta)$$

*CHAPITRE 3. APPROXIMATION DE BOX-JENKINS DE LA  
FONCTION DE VRAISEMBLANCE D'UN PROCESSUS GAUSSIEN*

---

Il est évident que pour calculer la somme des carrés inconditionnelle  $S(\phi, \theta)$  on a besoin de déterminer  $p + q$  valeurs initiales  $(X_0, X_{-1}, \dots, X_{1-p}, \varepsilon_0, \varepsilon_{-1}, \dots, \varepsilon_{1-q})$  en utilisant le principe de la "backcast" proposé par Box-Jenkins ([5], [3], [11], [12]). Par la suite, on se propose de répondre à ce problème dans le cas des modèles en moyennes mobiles et des modèles mixtes.

# Chapitre 4

## Détermination des valeurs initiales pour la somme des carrés inconditionnelle

Pour exprimer la vraisemblance d'une série chronologique centrée et identifiée par un modèle  $ARMA(p, q)$ , nous avons besoin de valeurs initiales  $(X^{(0)}, \varepsilon^{(0)})$ . L'idée de l'aller-retour est basée sur le principe du retournement du temps et de la prévision du passé ([5]). Le principe consiste à générer un mouvement de va et vient entre le futur et le passé de la série chronologique en utilisant les deux expansions duales du processus, i.e. l'expansion backward dans l'espace de Hilbert engendré par  $\{X_N, X_{N-1}, \dots, X_1\}$  et l'expansion forward dans l'espace de Hilbert engendré par  $\{X_1, X_2, \dots, X_N\}$ . En effet, il a été établi dans ([5]) qu'il est possible de représenter un processus aléatoire soit dans le sens croissant du temps, i.e. sur l'espace de Hilbert  $\vec{H}_N$  engendré par  $\{X_1, X_2, \dots, X_N\}$  ou dans le sens contraire du temps, i.e. sur l'espace de Hilbert  $\overleftarrow{H}_N$  engendré par  $\{X_N, X_{N-1}, \dots, X_1\}$ . Cependant, la structure de covariance du phénomène ne change pas.

### 4.1 Principe de l'aller-retour

Soit  $X_1, X_2, \dots, X_N$  une série chronologique identifiée par un processus  $ARMA(p, q)$  et donc elle vérifie la relation

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

*CHAPITRE 4. DÉTERMINATION DES VALEURS INITIALES POUR  
LA SOMME DES CARRÉS INCONDITIONNELLE*

---

Le calcul d'innovations  $\varepsilon_t$  à l'aide de l'équation itérative (4.1), ne peut se faire que si l'on se donne  $p + q$  valeurs initiales.

Il a été démontré plus haut que le logarithme de la fonction de vraisemblance inconditionnelle de  $X_1, X_2, \dots, X_N$  a la forme suivante :

$$\ln L(\phi, \theta, \sigma_\varepsilon^2) = -\frac{N}{2} \ln(2\pi) - \frac{N}{2} \ln \sigma_\varepsilon^2 - \frac{1}{2\sigma_\varepsilon^2} S(\phi, \theta) \quad (4.2)$$

où  $S(\phi, \theta)$  est la somme des carrés exacte définie telle que :

$$S(\phi, \theta) = \sum_{t=-\infty}^N \varepsilon_t^2(\phi, \theta, X_n, X_{n-1}, \dots) \quad (4.3)$$

où  $[\varepsilon_t] = E[\varepsilon_t/X, \phi, \theta]$  représente l'espérance de  $\varepsilon_t$  conditionnelle à  $(X, \phi, \theta)$ . Il a été établi dans ([5], [6]), que pour  $N$  assez grand la relation (4.2) est dominée par le terme  $\frac{S(\phi, \theta)}{2\sigma_\varepsilon^2}$ . Cela veut dire que les contours de la somme des carrés dans l'espace des paramètres  $(\phi, \theta)$  sont quasiment les mêmes contours que ceux de la log-vraisemblance. Il vient en particulier que les estimateurs des moindres carrés des paramètres sont très approximativement les mêmes que ceux obtenus par la méthode du maximum de vraisemblance.

On sait que le modèle (4.1) peut être exprimé sous la forme d'un processus en moyennes mobiles d'ordre infini ([2], [5], [16]). Cependant, pour les processus qui ont un intérêt pratique les poids  $\psi_j$  décroissent de manière géométrique. Alors, on peut supposer que les  $\psi_j$  s'annulent à partir d'un certain indice. Ainsi, on peut approcher la somme des carrés (4.3) avec une certaine précision par la somme des carrés associée à un processus en moyennes mobiles finis d'ordre  $s = \max(p, q + 1)$ . En pratique, les valeurs des innovations passées  $\varepsilon_t$  ( $t < 1$ ) peuvent être calculées récursivement à l'aide de l'équation (4.1) en partant de l'instant  $t = 1 - s$  avant lequel les innovations  $\varepsilon_t$  sont considérées comme négligeables. Donc, la somme des carrés (4.3) peut être exprimée telle que :

$$S(\phi, \theta) = \sum_{t=1-s}^N \varepsilon_t^2(\phi, \theta, X, X^{(0)}, \varepsilon^{(0)})$$

Box-Jenkins ([5]) soutiennent que s'ils remplaçaient l'opérateur  $B$  par son dual  $F$  dans l'expression (3.10) ils obtiendraient le modèle retourné ou modèle dual possédant les mêmes propriétés que le modèle original et dans lequel la valeur  $X_t$  du processus est exprimée en fonction de ses valeurs futures. Cette constatation a permis à ses auteurs de définir la notion de

"backcast" qui leurs permettait de procéder à la prévision des valeurs du processus survenues avant le début des observations. De plus, la structure de covariance du modèle (3.10) ne change pas si on remplace l'opérateur  $B$  par son dual  $F$  ([5]). Ainsi, on obtient l'expansion "backward" du processus  $ARMA(p, q)$  qui s'écrit telle que :

$$P(F) X_t = Q(F) e_t \quad (4.4)$$

où  $e_t$  est un autre processus d'innovations tel que  $E(e_t) = E(\varepsilon_t) = 0$  et  $Var(e_t) = Var(\varepsilon_t)$ . Nous verrons plus loin la nécessité de déterminer certaines valeurs du processus antérieur à la première observation de la série chronologique. Pour le calcul des innovations  $[\varepsilon_t]$  nous prenons l'espérance conditionnelle dans l'équation (3.10) écrite en extension telle que :

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

Sachant que les valeurs des prévisions sont négligeables à partir d'un certain indice  $s = \max(p, q + 1)$ , le calcul se fera selon le principe suivant :

$$[\varepsilon_j] = \begin{cases} 0 & si \quad j > N \\ \varepsilon_j & si \quad j \in [1, N] \end{cases}$$

et

$$[X_j] = \begin{cases} X_j & si \quad j \in [1, N] \\ \widehat{X}_j & si \quad j > N \text{ et } j < 1 \end{cases}$$

En retournant le temps et en utilisant (4.4), on procède au calcul des innovations  $[e_t]$  à l'aide de l'équation :

$$[e_t] = [X_t] - \phi_1 [X_{t+1}] - \dots - \phi_p [X_{t+p}] + \theta_1 [e_{t+1}] + \dots + \theta_q [e_{t+q}]$$

Suivant le principe suivant :

$$[e_j] = \begin{cases} 0 & si \quad j < 1 \\ e_j & si \quad j \geq 1 \end{cases}$$

et

$$[X_j] = \begin{cases} X_j & si \quad j \geq 1 \\ \widehat{X}_j & si \quad j < 1 \end{cases}$$

## 4.2 Détermination des valeurs initiales d'un processus $MA(q)$ par la méthode d'aller-retour

Cette section est inspirée par les travaux de Dakhmouche ([10]). Supposons qu'on associe à la série chronologique centrée  $X_1, X_2, \dots, X_N$  un modèle  $MA(q)$  stationnaire défini tel que :

$$X_t = Q(B)\varepsilon_t \quad (4.5)$$

L'expression(4.5) peut être écrite sous la forme :

$$\varepsilon_t = X_t + \theta_1\varepsilon_{t-1} + \dots + \theta_q\varepsilon_{t-q}$$

où  $\varepsilon_t$  est un processus d'innovations de distribution  $N(0, \sigma_\varepsilon^2)$ .

La représentation duale de (4.5) s'obtient en y remplaçant  $B$  par  $F$  ([2], [5]) telle que :

$$X_t = Q(F)e_t \text{ où } e_t \sim N(0, \sigma_\varepsilon^2)$$

Dans l'espace  $\overrightarrow{H}_N$  les valeurs  $(X_{N+\ell}, \ell > 0)$  du processus sont estimées par prévision d'origine  $N$ . Et dans l'espace  $\overleftarrow{H}_N$  les valeurs  $X_0, X_{-1}, \dots$  seront estimées par prévision d'origine 1 ([3], [4], [5], [12], [19]).

La fonction de prévision d'origine  $N$  pour l'expansion forward du processus est définie telle que :

$$\widehat{X}_{N+\ell} = Q(B)\widehat{\varepsilon}_{N+\ell} \quad (4.6)$$

avec

$$\widehat{X}_{N+\ell} = \begin{cases} \widehat{X}_{N+\ell} & \text{si } \ell > 0 \\ X_{N+\ell} & \text{si } \ell \leq 0 \end{cases} \quad \text{et} \quad \widehat{\varepsilon}_{N+\ell} = \begin{cases} 0 & \text{si } \ell > 0 \\ \varepsilon_{N+\ell} & \text{si } \ell \leq 0 \end{cases}$$

Il est évident que  $\widehat{X}_{N+\ell} = 0$  dès que  $\ell \geq q + 1$ .

Nous allons adopter les notations suivantes :

- Le vecteur des prévisions d'origine  $N$  des  $q$  valeurs futures du processus  $X_t$  obtenues au  $k^{\text{ème}}$  passage de l'algorithme d'aller-retour est noté par :

$$\widehat{X}_{(N,q)}^{(k)} = \left( \widehat{X}_{N+1}^{(k)} \quad \widehat{X}_{N+2}^{(k)} \quad \dots \quad \widehat{X}_{N+q}^{(k)} \right)^t$$

4.2. DÉTERMINATION DES VALEURS INITIALES D'UN PROCESSUS  
 MA(Q) PAR LA MÉTHODE D'ALLER-RETOUR

---

• Le vecteur des  $q$  dernières innovations pour l'expansion forward du processus obtenu au  $k^{\text{ème}}$  passage de l'algorithme est noté tel que :

$$\vec{\eta}^{(k)} = \left( \varepsilon_{N-q+1}^{(k)} \quad \varepsilon_{N-q+2}^{(k)} \quad \dots \quad \varepsilon_N^{(k)} \right)^t$$

• La matrice triangulaire  $\Theta$  d'ordre  $q$  est définie telle que :

$$\Theta = \begin{pmatrix} \theta_q & \theta_{q-1} & \dots & \theta_1 \\ 0 & \theta_q & \dots & \theta_2 \\ \dots & \dots & \dots & \dots \\ 0 & \dots & 0 & \theta_q \end{pmatrix}$$

De la relation (4.6) et pour  $\ell = 1, 2, \dots, q$ , on obtient au  $k^{\text{ème}}$  passage de l'algorithme l'équation ci-après :

$$\widehat{X}_{(N,q)}^{(k)} + \Theta \vec{\eta}^{(k)} = 0 \quad (4.7)$$

Considérons les notations suivantes :

• Le vecteur des  $q$  dernières observations pour l'expansion forward du processus est noté tel que :

$$\vec{X}_{(q)} = (X_{N-q+1} \quad X_{N-q+2} \quad \dots \quad X_N)^t$$

• La matrice diagonale  $G_B$  d'ordre  $q$  sera notée telle que :

$$G_B = \text{diag} \left( G_{N-q}(B) \quad G_{N-q+1}(B) \quad \dots \quad G_{N-1}(B) \right)^t$$

où

$$G_{N-\ell}(B) = \sum_{j=0}^{N-\ell} a_j B^j \quad \ell = 1, 2, \dots, q$$

• La matrice symétrique  $A$  d'ordre  $q$  sera notée telle que :

$$A = \begin{pmatrix} a_{N-q+1} & a_{N-q+2} & \dots & a_N \\ a_{N-q+2} & a_{N-q+3} & \dots & a_{N-1} \\ \dots & \dots & \dots & \dots \\ a_N & a_{N+1} & \dots & a_{N+q-1} \end{pmatrix}$$

où

$$a_m = \sum_{j=1}^q \theta_j a_{m-j} \quad \text{avec } a_0 = 1 \text{ et } a_1 = \theta_1$$

CHAPITRE 4. DÉTERMINATION DES VALEURS INITIALES POUR  
LA SOMME DES CARRÉS INCONDITIONNELLE

---

• Le vecteur des prévisions des  $q$  valeurs du processus  $X_t$  avant l'instant  $t = 1$ , au  $k^{\text{ème}}$  passage de l'algorithme est noté de la façon suivante :

$$\widehat{X}_{(1,q)}^{(k)} = \left( \widehat{X}_0^{(k)} \quad \widehat{X}_{-1}^{(k)} \quad \dots \quad \widehat{X}_{1-q}^{(k)} \right)^t$$

Au  $k^{\text{ème}}$  passage de l'algorithme, le calcul d'innovations à l'aide de l'expansion forward du processus est obtenu à l'aide de l'équation :

$$\overrightarrow{\eta}^{(k)} = G_B \overrightarrow{X}_{(q)} + A \widehat{X}_{(1,q)}^{(k-1)} \quad (4.8)$$

La fonction de prévision d'origine 1 et d'horizon  $\ell$  de l'expansion backward du processus est définie telle que ([5]) :

$$\widehat{X}_{1-\ell} = Q(F) \widehat{e}_{1-\ell} \quad (4.9)$$

où

$$\widehat{X}_{1-\ell} = \begin{cases} \widehat{X}_{1-\ell} & \text{si } \ell > 0 \\ X_{1-\ell} & \text{si } \ell \leq 0 \end{cases} \quad \text{et} \quad \widehat{e}_{1-\ell} = \begin{cases} 0 & \text{si } \ell > 0 \\ e_{1-\ell} & \text{si } \ell \leq 0 \end{cases}$$

Il est évident que  $\widehat{X}_{1-\ell} = 0$  dès que  $\ell \geq q + 1$ .

Considérons les notations suivantes :

• Le vecteur des  $q$  premières innovations pour l'expansion backward du processus sera noté tel que :

$$\overleftarrow{e}^{(k)} = \left( e_q^{(k)} \quad e_{q-1}^{(k)} \quad \dots \quad e_1^{(k)} \right)^t$$

Au  $k^{\text{ème}}$  passage de l'algorithme d'aller-retour et en utilisant la relation (4.9) pour  $\ell = 1, 2, \dots, q$ , on obtient un système d'équation qu'on résume ainsi :

$$\widehat{X}_{(1,q)}^{(k)} + \Theta \overleftarrow{e}^{(k)} = 0 \quad (4.10)$$

En notant :

•  $G_F$  la matrice diagonale d'ordre  $q$  définie telle que :

$$G_F = \text{diag} \left( G_{N-q}(F) \quad G_{N-q+1}(F) \quad \dots \quad G_{N-1}(F) \right)$$

où

$$G_{N-\ell}(F) = \sum_{j=0}^{N-\ell} a_j F^j \quad \ell = 1, 2, \dots, q$$

4.2. DÉTERMINATION DES VALEURS INITIALES D'UN PROCESSUS  
 $MA(Q)$  PAR LA MÉTHODE D'ALLER-RETOUR

---

• Le vecteur des  $q$  dernières observations pour l'expansion backward du processus sera noté tel que :

$$\overleftarrow{X}_{(q)} = (X_q \ X_{q-1} \ \dots \ X_1)^t$$

Alors, au  $k^{\text{ème}}$  passage de l'algorithme d'aller-retour, le calcul d'innovations à l'aide de l'expansion backward du processus est obtenu à l'aide de l'équation suivante :

$$\overleftarrow{e}^{(k)} = G_F \overleftarrow{X}_{(q)} + A \widehat{X}_{(N,q)}^{(k)} \quad (4.11)$$

**Proposition 4.2.1** Soit  $\left\{ \widehat{X}_{(1,q)}^{(k)} \right\}_{k \geq 1}$  la suite des vecteurs des valeurs initiales obtenues aux différents passages de l'algorithme d'aller-retour. Alors, cet algorithme converge si la suite  $\left( \widehat{X}_{(1,q)}^{(k)} \right)_{k \geq 1}$  admet un point fixe.

**Preuve** Des relations (4.7), (4.8), (4.10) et (4.11), on déduit le système d'équations :

$$\begin{cases} \widehat{X}_{(N,q)}^{(k)} + \Theta G_B \overrightarrow{X}_{(q)} + \Theta A \widehat{X}_{(1,q)}^{(k-1)} = 0 \\ \widehat{X}_{(1,q)}^{(k)} + \Theta G_F \overleftarrow{X}_{(q)} + \Theta A \widehat{X}_{(N,q)}^{(k)} = 0 \end{cases}$$

D'où, on obtient :

$$\widehat{X}_{(1,q)}^{(k)} = \Theta A \Theta G_B \overrightarrow{X}_{(q)} - \Theta G_F \overleftarrow{X}_{(q)} + (\Theta A)^2 \widehat{X}_{(1,q)}^{(k-1)} \quad (4.12)$$

Si la matrice est contractante, alors le point fixe est obtenu en passant à la limite dans (4.12) ■

**Proposition 4.2.2** La matrice  $\Theta A$  est une matrice contractante, i.e.  $\|\Theta A\| < 1$ .

**Preuve** Pour simplifier les calculs, nous allons considérer le cas  $q = 2$ .  
On sait que :

$$a_N = \theta_1 a_{N-1} + \theta_2 a_{N-2} \quad (4.13)$$

Alors, il est clair que :

$$\Theta A = \begin{pmatrix} a_{N+1} & a_{N+2} \\ \theta_2 a_N & \theta_2 a_{N+1} \end{pmatrix}$$

*CHAPITRE 4. DÉTERMINATION DES VALEURS INITIALES POUR  
LA SOMME DES CARRÉS INCONDITIONNELLE*

---

D'où on a :

$$\|\Theta A\| = |\theta_2 (a_{N+1}^2 - a_N a_{N+2})|$$

De plus, on remarque que les racines du polynôme caractéristique associé à l'équation de récurrence (4.13) sont les mêmes que celles de l'équation définie telle que :

$$r^2 - \theta_1 r - \theta_2 = 0 \tag{4.14}$$

Notons les deux racines de l'équation (4.14) par  $\rho_1$  et  $\rho_2$ .

On rappelle que le polynôme caractéristique d'un processus en moyennes mobiles d'ordre 2 est  $Q(B) = 1 - \theta_1 B - \theta_2 B^2$ . Il admet toutes ses racines à l'extérieur du cercle unité à cause de la propriété de causalité du processus ([3], [5], [7], [16]). De plus, il est évident que le polynôme  $Q(B)$  est le polynôme dual du polynôme (4.13). Par conséquent, les racines de l'équation (4.14) sont à l'intérieur du cercle unité.

La solution générale de l'équation de récurrence (4.13) est telle que :

$$\begin{aligned} a_N &= \lambda \rho_1^N + \mu \rho_2^N & si \quad \rho_1 \neq \rho_2 \\ a_N &= \delta \rho^N & si \quad \rho_1 = \rho_2 = \rho \end{aligned}$$

Par conséquent :

$$\|\Theta A\| = |\theta_2 (a_{N+1}^2 - a_N a_{N+2})| = |\theta_2 \lambda \mu (\rho_1 \rho_2)^N (\rho_1 - \rho_2)^2|$$

Sachant que les constantes  $\lambda$  et  $\mu$  sont données par les équations  $\lambda + \mu = 1$  et  $\lambda \rho_1 + \mu \rho_2 = \theta_1$ , alors :

$$\begin{aligned} \lambda &= \frac{\rho_1}{\sqrt{\Delta}} \quad et \quad \mu = 1 - \lambda = -\frac{\rho_2}{\sqrt{\Delta}} \\ O\grave{u} \quad \Delta &= \theta_1^2 + 4\theta_2 \end{aligned}$$

Donc,

$$\|\Theta A\| = |\theta_2 (\rho_1 \rho_2)^{N+1}|$$

D'où

$$\|\Theta A\| = |\theta_2| Max^{2N+2} (|\rho_1|, |\rho_2|) < 1$$

■

Finalement, en notant le vecteur limite  $\widehat{X}_{(1,q)}^{(\infty)}$  et par passage à la limite dans l'expression (4.12) nous donne :

$$\widehat{X}_{(1,q)}^{(\infty)} = [I_q - (\Theta A)^2]^{-1} \left\{ \Theta A \Theta G_B \vec{X}_{(q)} - \Theta G_F \overleftarrow{X}_{(q)} \right\}$$

4.2. DÉTERMINATION DES VALEURS INITIALES D'UN PROCESSUS  
 MA(Q) PAR LA MÉTHODE D'ALLER-RETOUR

---

où  $I_q$  est la matrice unité d'ordre  $q$  et  $\widehat{X}_{(1,q)}^{(\infty)} = \left( \widehat{X}_0^{(\infty)} \quad \widehat{X}_{-1}^{(\infty)} \quad \dots \quad \widehat{X}_{1-q}^{(\infty)} \right)^t$ .

Il reste à établir que le vecteur  $\widehat{X}_{(1,q)}^{(\infty)}$  coïncide avec le vecteur  $X^{(0)}$  des valeurs du processus avant l'instant  $t = 1$ , au sens de la topologie associée au produit scalaire dans  $\overrightarrow{H}_N$ .

**Proposition 4.2.3** *Soit  $X = (X_1, X_2, \dots, X_N)^t$  un système de générateurs de l'espace de Hilbert  $\overrightarrow{H}_N$ . Alors*

$$E \left\{ \widehat{X}_{(1,q)}^{(\infty)} X^t \right\} = E \left\{ X^{(0)} X^t \right\} = \Gamma$$

**Preuve** Considérons le cas  $q = 2$  et rappelons les notations adoptées précédemment :

$$\Theta = \begin{pmatrix} \theta_2 & \theta_1 \\ 0 & \theta_2 \end{pmatrix} \quad A = \begin{pmatrix} a_{N-1} & a_N \\ a_N & a_{N+1} \end{pmatrix}$$

$$G_B \overrightarrow{X}_{(2)} = \begin{pmatrix} \sum_{j=0}^{N-2} a_j B^j X_{N-1} \\ \sum_{j=0}^{N-1} a_j B^j X_N \end{pmatrix} \quad G_F \overleftarrow{X}_{(2)} = \begin{pmatrix} \sum_{j=0}^{N-2} a_j F^j X_2 \\ \sum_{j=0}^{N-1} a_j F^j X_1 \end{pmatrix}$$

Calculons :

$$E \left( \widehat{X}_{(1,2)}^{(\infty)} X^t \right) = [I_2 - (\Theta A)^2]^{-1} E \left( \Theta A \Theta G_B \overrightarrow{X}_{(2)} X^t - \Theta G_F \overleftarrow{X}_{(2)} X^t \right)$$

On a :

$$E \left( G_B \overrightarrow{X}_{(2)} X^t \right) = \begin{pmatrix} \sum_{j=0}^{N-2} a_j \gamma_{|N-j-2|} & \sum_{j=0}^{N-2} a_j \gamma_{|N-j-3|} & \dots & \sum_{j=0}^{N-2} a_j \gamma_{j+1} \\ \sum_{j=0}^{N-1} a_j \gamma_{|N-j-1|} & \sum_{j=0}^{N-1} a_j \gamma_{|N-j-2|} & \dots & \sum_{j=0}^{N-1} a_j \gamma_j \end{pmatrix}$$

Par ailleurs, la fonction d'autocovariance  $\gamma_m$  d'un MA(2) vérifie les condi-

CHAPITRE 4. DÉTERMINATION DES VALEURS INITIALES POUR  
LA SOMME DES CARRÉS INCONDITIONNELLE

---

tions suivantes :

$$(C) \left\{ \begin{array}{l} \frac{1}{\sigma_\varepsilon^2} \sum_{j=0}^{N-2} a_j \gamma_{|N-r-j|} = \begin{cases} \theta_1 a_{N-1} + \theta_2^2 a_{N-2} & \text{si } r = 2 \\ \theta_2 a_{N-1} & \text{si } r = 3 \\ 1 & \text{si } r = N-1 \\ -\theta_1 & \text{si } r = N \end{cases} \\ \frac{1}{\sigma_\varepsilon^2} \sum_{j=0}^{N-2} a_j \gamma_{|N-r-j|} = 0 & \text{si } 4 \leq r \leq N-2 \\ \frac{1}{\sigma_\varepsilon^2} \sum_{j=0}^{N-1} a_j \gamma_{|N-r-j|} = \begin{cases} \theta_1 a_N + \theta_2^2 a_{N-1} & \text{si } r = 1 \\ \theta_2 a_N & \text{si } r = 2 \\ 0 & \text{si } r = N \\ -\theta_1 & \text{si } r = N-1 \end{cases} \\ \frac{1}{\sigma_\varepsilon^2} \sum_{j=0}^{N-1} a_j \gamma_{|N-r-j|} = 0 & \text{si } 3 \leq r \leq N-2 \end{array} \right.$$

Donc

$$E \left( G_B \overrightarrow{X}_{(2)} X^t \right) = \sigma_\varepsilon^2 \begin{pmatrix} \theta_2^2 a_{N-2} + \theta_1 a_{N-1} & \theta_2 a_{N-1} & ..0.. & 1 & -\theta_1 \\ \theta_2^2 a_{N-1} + \theta_1 a_N & \theta_2 a_N & ..0.. & 0 & 1 \end{pmatrix} \quad (4.15)$$

En procédant de la même manière que précédemment,  $E \left( G_F \overleftarrow{X}_{(2)} X^t \right)$  peut être exprimée telle que :

$$E \left( G_F \overleftarrow{X}_{(2)} X^t \right) = \begin{pmatrix} \sum_{j=0}^{N-2} a_j \gamma_{j+1} & \sum_{j=0}^{N-2} a_j \gamma_j & \dots & \sum_{j=0}^{N-2} a_j \gamma_{|N-j-2|} \\ \sum_{j=0}^{N-1} a_j \gamma_j & \sum_{j=0}^{N-1} a_j \gamma_{|j-1|} & \dots & \sum_{j=0}^{N-1} a_j \gamma_{|N-j-1|} \end{pmatrix}$$

D'après les conditions (C), l'expression ci-dessus s'écrit alors telle que :

$$E \left( G_F \overleftarrow{X}_{(2)} X^t \right) = \sigma_\varepsilon^2 \begin{pmatrix} -\theta_1 & 1 & ..0.. & \theta_2 a_{N-1} & \theta_2^2 a_{N-2} + \theta_1 a_{N-1} \\ 1 & 0 & ..0.. & \theta_2 a_N & \theta_2^2 a_{N-1} + \theta_1 a_N \end{pmatrix} \quad (4.16)$$

Notons :

$$\Phi = \begin{pmatrix} 1 & -\theta_1 \\ 0 & 1 \end{pmatrix} \quad \text{et} \quad \Sigma = \begin{pmatrix} -\theta_1 & 1 \\ 1 & 0 \end{pmatrix}$$

Il est aisé de remarquer que :

$$-\Theta \Sigma = \frac{1}{\sigma_\varepsilon^2} \Gamma = \begin{pmatrix} -\theta_1 + \theta_1 \theta_2 & -\theta_2 \\ -\theta_2 & 0 \end{pmatrix}$$

4.3. DÉTERMINATION DES VALEURS INITIALES D'UN PROCESSUS  
 ARMA (P, Q) PAR LA MÉTHODE D'ALLER-RETOUR

---

$$-A\Gamma = \sigma_\varepsilon^2 \begin{pmatrix} \theta_2^2 a_{N-2} + \theta_1 a_{N-1} & \theta_2 a_{N-1} \\ \theta_2^2 a_{N-1} + \theta_1 a_N & \theta_2 a_N \end{pmatrix}$$

$$\Theta\Phi = \begin{pmatrix} \theta_2 & \theta_1 - \theta_1\theta_2 \\ 0 & \theta_2 \end{pmatrix}$$

$$A\Theta\Phi = \begin{pmatrix} \theta_2 a_{N-2} & \theta_1 a_{N-1} + \theta_2^2 a_{N-2} \\ \theta_2 a_N & \theta_1 a_N + \theta_2^2 a_{N-1} \end{pmatrix}$$

Ainsi, les expressions (4.15) et (4.16) peuvent être exprimées plus simplement sous la forme :

$$E \left( G_B \overrightarrow{X}_{(2)} X^t \right) = \sigma_\varepsilon^2 \begin{pmatrix} A\Theta\Phi & 0 & \Phi \end{pmatrix}$$

$$E \left( G_F \overleftarrow{X}_{(2)} X^t \right) = \sigma_\varepsilon^2 \begin{pmatrix} \Sigma & 0 & A\Theta\Phi \end{pmatrix}$$

Finalement, on obtient :

$$\begin{aligned} E \left( \widehat{X}_{(1,2)}^{(\infty)} X^t \right) &= \sigma_\varepsilon^2 [I_2 - (\Theta A)^2]^{-1} ((\Theta A)^2 \Theta \Sigma - \Theta \Sigma \quad 0 \quad 0) \\ &= \sigma_\varepsilon^2 [I_2 - (\Theta A)^2]^{-1} [(\Theta A)^2 - I_2] \Theta \Sigma = -\sigma_\varepsilon^2 \Theta \Sigma = \Gamma \end{aligned}$$

■

Ainsi, on montre bien que le vecteur-limite  $\widehat{X}_{(1,q)}^{(\infty)}$  coïncide avec le vecteur des valeurs initiales  $X^{(0)}$ , relativement à la norme associée avec le produit scalaire défini sur l'espace de Hilbert  $L^2$ .

### 4.3 Détermination des valeurs initiales d'un processus ARMA (p, q) par la méthode d'aller-retour

Cette section est inspirée par les travaux de Dakhmouche ([11]). Le calcul de la somme des carrés inconditionnelle pour un processus ARMA(p, q) nécessite la détermination de  $s = \max(p, q + 1)$  valeurs initiales  $X^{(0)} = (X_0, X_{-1}, \dots, X_{1-s})^t$  et  $\varepsilon^{(0)} = (\varepsilon_0, \varepsilon_{-1}, \dots, \varepsilon_{1-s})^t$ . En pratique, les valeurs d'innovations passées  $\varepsilon_t$  ( $t < 1$ ) peuvent être calculées récursivement à l'aide de l'équation (4.1) en partant de l'instant  $t = 1 - s$  avant lequel les innovations  $\varepsilon_t$  sont considérées comme négligeables. Donc, on peut approcher la somme

*CHAPITRE 4. DÉTERMINATION DES VALEURS INITIALES POUR  
LA SOMME DES CARRÉS INCONDITIONNELLE*

---

des carrés exacte (4.3) avec une précision contrôlée, par la somme des carrés inconditionnelle telle que :

$$S(\phi, \theta) = \sum_{t=1-s}^N \varepsilon_t^2(\phi, \theta, X, X^{(0)}, \varepsilon^{(0)})$$

La fonction de prévision d'origine  $N$  à l'horizon  $\ell$  pour l'expansion forward du processus est définie telle que :

$$P(B) \widehat{X}_{N+\ell} = Q(B) \widehat{\varepsilon}_{N+\ell} \quad (4.17)$$

avec

$$\widehat{X}_{N+\ell} = \begin{cases} \widehat{X}_{N+\ell} & \text{si } \ell > 0 \\ X_{N+\ell} & \text{si } \ell \leq 0 \end{cases} \quad \text{et} \quad \widehat{\varepsilon}_{N+\ell} = \begin{cases} 0 & \text{si } \ell > 0 \\ \varepsilon_{N+\ell} & \text{si } \ell \leq 0 \end{cases}$$

Puisque  $P(B)$  a ses racines à l'extérieur du cercle unité,  $\widehat{X}_{N+\ell}$  tend exponentiellement vers zéro lorsque  $\ell \rightarrow \infty$  ([3], [5], ), à partir d'un indice  $s$  les valeurs  $\widehat{X}_{N+\ell}$  sont pratiquement nulles. En général, nous prenons  $s = \max(p, q + 1)$ .

Nous allons adopter les notations suivantes :

- Le vecteur des prévisions d'origine  $N$  des  $s$  valeurs futures pour l'expansion forward du processus obtenu au  $k^{\text{ème}}$  passage de l'algorithme est noté tel que :

$$\widehat{X}_{(N,s)}^{(k)} = \left( \widehat{X}_{N+1}^{(k)} \quad \widehat{X}_{N+2}^{(k)} \quad \dots \quad \widehat{X}_{N+s}^{(k)} \right)^t$$

- Le vecteur des  $s$  dernières observations pour l'expansion forward du processus est noté tel que :

$$\overrightarrow{X}_{(s)} = \left( X_{N-s+1} \quad X_{N-s+2} \quad \dots \quad X_N \right)^t$$

- Le vecteur des  $s$  dernières valeurs d'innovations pour l'expansion forward du processus obtenu au  $k^{\text{ème}}$  passage de l'algorithme est noté tel que :

$$\overrightarrow{\eta}^{(k)} = \left( \varepsilon_{N-s+1}^{(k)} \quad \varepsilon_{N-s+2}^{(k)} \quad \dots \quad \varepsilon_N^{(k)} \right)^t$$

- Soit  $\Phi$ ,  $\widetilde{\Phi}$  et  $\Theta$  les trois matrices triangulaires d'ordre  $s$  définies telle que :

$$\Phi = \begin{pmatrix} 1 & 0 & \dots & \dots & 0 \\ -\phi_1 & 1 & 0 & \dots & 0 \\ -\phi_2 & -\phi_1 & 1 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ -\phi_{s-1} & -\phi_{s-2} & \dots & -\phi_1 & 1 \end{pmatrix}$$

4.3. DÉTERMINATION DES VALEURS INITIALES D'UN PROCESSUS  
ARMA (P, Q) PAR LA MÉTHODE D'ALLER-RETOUR

---

$$\tilde{\Phi} = \begin{pmatrix} \phi_s & \phi_{s-1} & \dots & \dots & \phi_1 \\ 0 & \phi_s & \phi_{s-1} & \dots & \phi_2 \\ 0 & 0 & \phi_s & \dots & \phi_3 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 0 & \phi_s \end{pmatrix}$$

$$\Theta = - \begin{pmatrix} \theta_s & \theta_{s-1} & \dots & \dots & \theta_1 \\ 0 & \theta_s & \theta_{s-1} & \dots & \theta_2 \\ 0 & 0 & \theta_s & \dots & \theta_3 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 0 & \theta_s \end{pmatrix}$$

De la relation (4.17) et pour  $\ell = 1, 2, \dots, \max(p, q + 1) = s$  on obtient au  $k^{\text{ème}}$  passage de l'algorithme l'équation :

$$\Phi \widehat{X}_{(N,s)}^{(k)} - \tilde{\Phi} \overrightarrow{X}_{(s)} + \Theta \overrightarrow{\eta}^{(k)} = 0 \quad (4.18)$$

Considérons les notations suivantes :

- La matrice diagonale  $G_B$  d'ordre  $s$  sera notée telle que :

$$G_B = \text{diag} (G_{N-s}(B) \quad G_{N-s+1}(B) \quad \dots \quad G_{N-1}(B))^t$$

où

$$G_{N-\ell}(B) = \sum_{j=0}^{N-\ell} a_j B^j \quad \ell = 1, 2, \dots, s$$

- La matrice symétrique  $A$  d'ordre  $s$  sera notée telle que :

$$A = \begin{pmatrix} a_{N-s+1} & a_{N-s+2} & \dots & a_N \\ a_{N-s+2} & a_{N-s+3} & \dots & a_{N-1} \\ \dots & \dots & \dots & \dots \\ a_N & a_{N+1} & \dots & a_{N+s-1} \end{pmatrix}$$

où

$$a_m = \sum_{j=1}^s \theta_j a_{m-j} \quad \text{avec } a_0 = 1 \text{ et } a_1 = \theta_1$$

- Le vecteur des prévisions des  $s$  valeurs du processus  $\widehat{X}_t^{(k)}$ ,  $t = 0, -1, -2, \dots, 1-s$ , obtenu au  $k^{\text{ème}}$  passage de l'algorithme est noté de la manière ci-après :

$$\widehat{X}_{(1,s)}^{(k)} = \left( \widehat{X}_0^{(k)} \quad \widehat{X}_{-1}^{(k)} \quad \dots \quad \widehat{X}_{1-s}^{(k)} \right)^t$$

*CHAPITRE 4. DÉTERMINATION DES VALEURS INITIALES POUR  
LA SOMME DES CARRÉS INCONDITIONNELLE*

---

Au  $k^{\text{ème}}$  passage de l'algorithme, le calcul d'innovations à l'aide de l'expansion forward du processus est obtenu à l'aide de l'équation :

$$\vec{\eta}^{(k)} = G_B \vec{X}_{(s)} + A \widehat{X}_{(1,s)}^{(k-1)} \quad (4.19)$$

retournons au passé et en utilisant l'expansion backward du processus. La fonction de prévision d'origine 1 et d'horizon  $\ell$  est définie telle que ([5]) :

$$P(F) \widehat{X}_{1-\ell} = Q(F) \widehat{e}_{1-\ell} \quad (4.20)$$

où

$$\widehat{X}_{1-\ell} = \begin{cases} \widehat{X}_{1-\ell} & \text{si } \ell > 0 \\ X_{1-\ell} & \text{si } \ell \leq 0 \end{cases} \quad \text{et} \quad \widehat{e}_{1-\ell} = \begin{cases} 0 & \text{si } \ell > 0 \\ e_{1-\ell} & \text{si } \ell \leq 0 \end{cases}$$

considérons la symétrie qui existe entre l'expansion forward et l'expansion backward du processus ([5]). Ce qui résulte que  $\widehat{X}_{1-\ell}$  tend exponentiellement vers zéro lorsque  $\ell \rightarrow \infty$ , à partir d'un indice  $s$  les valeurs  $\widehat{X}_{1-\ell}$  sont pratiquement nulles.

Considérons les notations suivantes

- Le vecteur des  $s$  dernières observations pour l'expansion backward du processus sera noté tel que :

$$\overleftarrow{X}_{(s)} = (X_s \quad X_{s-1} \quad \dots \quad X_1)^t$$

- Le vecteur des  $s$  dernières innovations pour l'expansion backward du processus sera noté tel que :

$$\overleftarrow{e}^{(k)} = \left( e_s^{(k)} \quad e_{s-1}^{(k)} \quad \dots \quad e_1^{(k)} \right)^t$$

Au  $k^{\text{ème}}$  passage de l'algorithme d'aller-retour et en utilisant la relation (4.20) pour  $\ell = 1, 2, \dots, q$ , on obtient un système d'équation qu'on résume ainsi :

$$\Phi \widehat{X}_{(1,s)}^{(k)} - \widetilde{\Phi} \overleftarrow{X}_{(s)}^{(k)} + \Theta \overleftarrow{e}^{(k)} = 0 \quad (4.21)$$

En notant :

- $G_F$  la matrice diagonale d'ordre  $s \times s$  définie telle que :

$$G_F = \text{diag} (G_{N-s}(F) \quad G_{N-s+1}(F) \quad \dots \quad G_{N-1}(F))$$

où

$$G_{N-\ell}(F) = \sum_{j=0}^{N-\ell} a_j F^j \quad \ell = 1, 2, \dots, s$$

### 4.3. DÉTERMINATION DES VALEURS INITIALES D'UN PROCESSUS ARMA (P, Q) PAR LA MÉTHODE D'ALLER-RETOUR

---

Alors, au  $k^{\text{ème}}$  passage de l'algorithme d'aller-retour, le calcul des innovations à l'aide de l'expansion backward du processus est obtenu par l'équation suivante :

$$\overleftarrow{e}^{(k)} = G_F \overleftarrow{X}_{(s)} + A \widehat{X}_{(N,s)}^{(k)} \quad (4.22)$$

**Proposition 4.3.1** Soit  $\left\{ \widehat{X}_{(1,s)}^{(k)} \right\}_{k \geq 1}$  la suite des vecteurs des valeurs initiales obtenues aux différents passages de l'algorithme d'aller-retour. Alors, cet algorithme converge si la suite  $\left( \widehat{X}_{(1,s)}^{(k)} \right)_{k \geq 1}$  admet un point fixe.

**Preuve** Des relations (4.18), (4.19), (4.21) et (4.22), on déduit le système d'équations :

$$\begin{cases} \Phi \widehat{X}_{(N,s)}^{(k)} - \widetilde{\Phi} \overrightarrow{X}_{(s)} + \Theta G_B \overrightarrow{X}_{(s)} + \Theta A \widehat{X}_{(1,s)}^{(k-1)} = 0 \\ \Phi \widehat{X}_{(1,s)}^{(k)} - \widetilde{\Phi} \overleftarrow{X}_{(s)} + \Theta G_F \overleftarrow{X}_{(s)} + \Theta A \widehat{X}_{(N,s)}^{(k)} = 0 \end{cases}$$

D'où, on obtient :

$$\begin{aligned} \widehat{X}_{(1,s)}^{(k)} &= \Phi^{-1} \Theta A \Phi^{-1} \left( \Theta G_B - \widetilde{\Phi} \right) \overrightarrow{X}_{(s)} - \\ &\quad \Phi^{-1} \left( \Theta G_F - \widetilde{\Phi} \right) \overleftarrow{X}_{(s)} + [\Phi^{-1} \Theta A]^2 \widehat{X}_{(1,s)}^{(k-1)} \end{aligned} \quad (4.23)$$

Si la matrice est contractante, alors le point fixe est obtenu en passant à la limite dans (4.23). ■

**Proposition 4.3.2** La matrice  $\Phi^{-1} \Theta A$  est une matrice contractante, i.e.  $\|\Phi^{-1} \Theta A\| < 1$ .

**Preuve** Il est évident que le module du déterminant de la matrice  $\Phi$  est égale à 1. De plus, il est clair que le polynôme caractéristique de la matrice  $\Theta A$  est le polynôme dual du polynôme  $Q(B)$  dont les racines sont à l'extérieur du cercle unité à cause de la propriété de l'inversibilité du processus ([3], [6], [12]). Donc, tous les valeurs propres de la matrice  $\Theta A$  sont à l'intérieur d'un cercle unité. Par conséquent on a :

$$\|\Phi^{-1} \Theta A\| \leq \|\Phi\|^{-1} \|\Theta A\| \leq \max_i (|v_i|)^{Ns} \quad (4.24)$$

où  $v_i$ ,  $i = 1, 2, \dots, s$  sont les valeurs propres de la matrice  $\Theta A$ .  
Puisque  $\max_i (|v_i|) < 1$ , l'inégalité (4.24) est inférieure à 1. ■

Alors on obtient :

$$\widehat{X}_{(1,s)}^{(\infty)} = \left\{ I_s - [\Phi^{-1}\Theta A]^2 \right\}^{-1} \Phi^{-1}\Theta A \Phi^{-1} \left( \Theta G_B - \widetilde{\Phi} \right) \overrightarrow{X}_{(s)} - \Phi^{-1} \left( \Theta G_F - \widetilde{\Phi} \right) \overleftarrow{X}_{(s)} \quad (4.25)$$

### 4.3.1 Cas particulier : Calcul des valeurs initiales pour un processus $ARMA(1, 1)$ par la méthode d'aller-retour

La forme générale d'un processus  $ARMA(1, 1)$  peut être exprimée telle que :

$$X_t - \alpha X_{t-1} = \varepsilon_t - \beta \varepsilon_{t-1}$$

Alors, les innovations  $\varepsilon_t$  peuvent être calculées récursivement telles que :

$$\varepsilon_t = X_t - \alpha X_{t-1} + \beta \varepsilon_{t-1} \quad (4.26)$$

Soit  $\widehat{X}_0$  la valeur initiale et soit  $\widehat{X}_{N+1}$  la prévision de la valeur future  $X_{N+1}$ . En procédant de la même manière que dans le paragraphe précédent, on obtient au  $k^{th}$  passage de l'algorithme de "l'aller-retour" les équations suivantes :

$$\widehat{X}_{N+1}^{(k)} = (\alpha - \beta) \sum_{j=0}^{N-1} \beta^j X_{N-j} + \beta^N \left( \alpha \widehat{X}_0^{(k-1)} - \beta \varepsilon_0^{(k-1)} \right) \quad (4.27)$$

quand on utilise l'expansion "forward" et

$$\widehat{X}_0^{(k)} = (\alpha - \beta) \sum_{j=0}^{N-1} \beta^j X_{j+1} + B^N \left( \alpha \widehat{X}_{N+1}^{(k)} - \beta \varepsilon_{N+1}^{(k)} \right) \quad (4.28)$$

quand on utilise l'expansion "backward".

De plus, il a été établi dans ([5], [7], [16]) que les prévisions à l'horizon  $\ell$  en utilisant l'expansion "forward" ou l'expansion "backward", sont obtenues telles que :

$$\widehat{X}_{N+\ell} = \alpha^{\ell-1} \widehat{X}_{N+1}$$

et

$$\widehat{X}_{-\ell} = \alpha^{\ell} \widehat{X}_0$$

### 4.3. DÉTERMINATION DES VALEURS INITIALES D'UN PROCESSUS ARMA (P, Q) PAR LA MÉTHODE D'ALLER-RETOUR

---

Alors, au  $k^{\text{th}}$  passage de l'algorithme de "l'aller-retour" et à l'horizon  $\ell$ , les innovations  $\varepsilon_0^{(k)}$  et  $e_{N+1}^{(k)}$  sont déterminées telles que :

$$\varepsilon_0^{(k,\ell)} = \frac{(1 - \alpha^2) \left(1 - (\alpha\beta)^\ell\right)}{1 - \alpha\beta} \widehat{X}_0^{(k)} \quad (4.29)$$

$$e_{N+1}^{(k)} = \frac{(1 - \alpha^2) \left(1 - (\alpha\beta)^\ell\right)}{1 - \alpha\beta} \widehat{X}_{N+1}^{(k)} \quad (4.30)$$

Faire des prévisions à l'horizon infini à chaque passage de l'algorithme, consiste à passer à la limite sur  $\ell$  dans les équations (4.29) et (4.30). De plus, on sait que  $\alpha\beta < 1$ . Alors, on obtient

$$\varepsilon_0^{(k,\infty)} = \frac{(1 - \alpha^2)}{1 - \alpha\beta} \widehat{X}_0^{(k)} \quad (4.31)$$

$$e_{N+1}^{(k,\infty)} = \frac{(1 - \alpha^2)}{1 - \alpha\beta} \widehat{X}_{N+1}^{(k)} \quad (4.32)$$

Des équations (4.27),(4.28),(4.31),(4.32), et en reprenant les notations utilisées plus haut, on déduit :

$$\begin{aligned} \widehat{X}_{N+1}^{(k)} &= (\alpha - \beta) G_{N-1}(B) X_N + \frac{B^N (\alpha - \beta)}{1 - \alpha\beta} \widehat{X}_0^{(k)} \\ \widehat{X}_0^{(k+1)} &= (\alpha - \beta) G_{N-1}(F) X_1 + \frac{B^N (\alpha - \beta)}{1 - \alpha\beta} \widehat{X}_{N+1}^{(k)} \end{aligned}$$

Où

$$G_{N-\ell}(B) = \sum_{j=0}^{N-\ell} a_j B^j \quad \ell = 1, 2, \dots, q$$

et

$$G_{N-\ell}(F) = \sum_{j=0}^{N-\ell} a_j F^j \quad \ell = 1, 2, \dots, q$$

En notant

$$\mathbf{\Lambda} = \frac{(\alpha - \beta)}{1 - \alpha\beta} B^N$$

On obtient finalement :

$$\widehat{X}_0^{(k+1)} = \mathbf{\Lambda} (\alpha - \beta) G_{N-1}(B) X_N + (\alpha - \beta) G_{N-1}(F) X_1 + \mathbf{\Lambda}^2 \widehat{X}_0^{(k)} \quad (4.33)$$

CHAPITRE 4. DÉTERMINATION DES VALEURS INITIALES POUR  
LA SOMME DES CARRÉS INCONDITIONNELLE

---

**Proposition 4.3.3** *l'équation de récurrence (4.33) admet un point fixe si  $|\Lambda| < 1$ .*

**Preuve** Rappelons qu'un processus  $ARMA(1, 1)$  est stationnaire si  $|\alpha| < 1$  et il est inversible si  $|\beta| < 1$  voir ([5], [7], [16]). De plus, il est clair que  $\left| \frac{(\alpha - \beta)}{1 - \alpha\beta} \right| < M$ . Par conséquent,  $|\Lambda| = \left| \frac{(\alpha - \beta)}{1 - \alpha\beta} B^N \right| < M |B^N| < 1$ . ■

Finalement, le point fixe ou la valeur limite  $\widehat{X}_0^{(\infty)}$  est obtenu tel que :

$$\widehat{X}_0^{(\infty)} = \frac{\Lambda(\alpha - \beta) G_{N-1}(B) X_N + (\alpha - \beta) G_{N-1}(F) X_1}{1 - \Lambda^2}$$

**Proposition 4.3.4** *Soit  $(X_1, X_2, \dots, X_N)$  un système de générateurs de l'espace de Hilbert  $H_N$ . Si le vecteur  $\widehat{X}_0^{(\infty)}$  est la valeur recherchée, alors la condition suivante est satisfaite :*

$$E\left(\widehat{X}_0^{(\infty)} X_k\right) = \begin{cases} \gamma_1 & \text{si } k = 1 \\ \alpha^{k-1} \gamma_1 & \text{si } k \geq 2 \end{cases}$$

**Preuve** Calculons

$$E\left(\widehat{X}_0^{(\infty)} X_1\right) = \frac{(\alpha - \beta)}{1 - \Lambda^2} [\Lambda E(G_{N-1}(B) X_N X_1) + E(G_{N-1}(F) X_1 X_1)]$$

D'une part, on calcule :

$$\begin{aligned} E(G_{N-1}(B) X_N X_1) &= \sum_{j=0}^{N-1} B^j E(X_{N-j} X_1) = \sum_{j=0}^{N-1} B^j \gamma_{N-j-1} \\ &= B^{N-1} \gamma_0 + \frac{\alpha^{N-1} - \beta^{N-1}}{\alpha - \beta} \gamma_1 \end{aligned}$$

Et d'autre part, on a :

$$\begin{aligned} E(G_{N-1}(F) X_1 X_1) &= \sum_{j=0}^{N-1} B^j E(X_{j+1} X_1) = \sum_{j=0}^{N-1} B^j \gamma_j \\ &= \gamma_0 + \beta \gamma_1 + \frac{\alpha \beta^2}{1 - \alpha\beta} \gamma_1 - \frac{\alpha^{N-1} \Lambda}{\alpha - \beta} \gamma_1 \end{aligned}$$

Par conséquent,

$$\Lambda E(G_{N-1}(B) X_N X_1) + E(G_{N-1}(F) X_1 X_1) = \frac{1 - \alpha\beta}{1 - \alpha^2} (1 - \Lambda^2) \sigma_\varepsilon^2$$

4.3. DÉTERMINATION DES VALEURS INITIALES D'UN PROCESSUS  
ARMA (P, Q) PAR LA MÉTHODE D'ALLER-RETOUR

---

Il a été établi dans ([5], [7], [16]) que :

$$\begin{cases} \gamma_0 = \frac{1 + \beta^2 - 2\alpha\beta}{1 - \alpha^2} \sigma_\varepsilon^2 \\ \gamma_1 = \frac{(1 - \alpha\beta)(\alpha - \beta)}{1 - \alpha^2} \sigma_\varepsilon^2 \\ \gamma_k = \alpha\gamma_{k-1} \quad k \geq 2 \end{cases}$$

D'où

$$E\left(\widehat{X}_0^{(\infty)} X_1\right) = \frac{(\alpha - \beta)(1 - \alpha\beta)}{1 - \alpha^2} \sigma_\varepsilon^2 = \gamma_1$$

Considérons maintenant le cas  $k = 2$ . Alors :

$$E\left(\widehat{X}_0^{(\infty)} X_2\right) = \frac{\alpha - \beta}{1 - \Lambda^2} [\Lambda E(G_{N-1}(B) X_N X_2) + E(G_{N-1}(F) X_1 X_2)]$$

Ou bien

$$\frac{1 - \Lambda^2}{\alpha - \beta} E\left(\widehat{X}_0^{(\infty)} X_2\right) = \Lambda E(G_{N-1}(B) X_N X_2) + E(G_{N-1}(F) X_1 X_2)$$

En premier lieu, on a :

$$\begin{aligned} E(G_{N-1}(B) X_N X_2) &= \sum_{j=0}^{N-1} \beta^j E(X_{N-j} X_2) = \sum_{j=0}^{N-1} \beta^j \gamma_{|N-j-2|} \\ &= \beta^{N-1} \gamma_1 + \beta^{N-2} \gamma_0 + \beta^{N-3} \gamma_1 + \sum_{j=0}^{N-4} \beta^j \gamma_{|N-j-2|} \\ &= (\beta^N - \alpha\beta^{N-1} + \beta^{N-2} - \alpha^{N-2}) \frac{\gamma_1}{\beta - \alpha} + \beta^{N-2} \gamma_0 \end{aligned}$$

Et ensuite, on obtient :

$$\begin{aligned} E(G_{N-1}(F) X_1 X_2) &= \sum_{j=0}^{N-1} \beta^j \gamma_{|j-1|} = \gamma_1 + \beta\gamma_0 + \beta^2\gamma_1 + \sum_{j=3}^{N-1} \beta^j \gamma_{|j-1|} \\ &= \gamma_1 + \beta\gamma_0 + \frac{\beta^2\gamma_1}{1 - \alpha\beta} - \frac{\Lambda\alpha^{N-2}}{\alpha - \beta} \gamma_1 \end{aligned}$$

Par conséquent,

$$\frac{1 - \Lambda^2}{\alpha - \beta} E\left(\widehat{X}_0^{(\infty)} X_2\right) = (\beta^N - \alpha\beta^{N-1} + \beta^{N-2}) \frac{\Lambda}{\beta - \alpha} \gamma_1 +$$

CHAPITRE 4. DÉTERMINATION DES VALEURS INITIALES POUR  
LA SOMME DES CARRÉS INCONDITIONNELLE

---

$$\begin{aligned}
& +\Lambda\beta^{N-2}\gamma_0 + \gamma_1 + \beta\gamma_0 + \frac{\beta}{1-\alpha\beta}\gamma_1 \\
& = -\frac{\Lambda^2(1-\alpha\beta)}{\beta^2(\alpha-\beta)^2} \left\{ (1+\beta^2-\alpha\beta)\gamma_1 + (\beta-\alpha)\gamma_0 \right\} \\
& \quad + \gamma_1 + \beta\gamma_0 + \frac{\beta^2\gamma_1}{1-\alpha\beta} \\
& = \frac{\alpha(1-\alpha\beta)}{1-\alpha^2} (1-\Lambda^2)\sigma_\varepsilon^2
\end{aligned}$$

Donc

$$E\left(\widehat{X}_0^{(\infty)} X_2\right) = \alpha\gamma_1$$

Pour terminer montrons, qu'en général, on a :

$$E\left(\widehat{X}_0^{(\infty)} X_h\right) = \alpha^{h-1}\gamma_1 \quad \forall h \geq 2$$

Un premier calcul nous donne :

$$\begin{aligned}
E(G_{N-1}(B) X_N X_h) &= \sum_{j=0}^{N-1} \beta^j \gamma_{N-j-h} = \sum_{j=0}^{N-h-2} \beta^j \gamma_{N-j-h} + \sum_{j=N-h+2}^{N-1} \beta^j \gamma_{j+h-N} \\
& \quad + \beta^{N-h-1}\gamma_1 + \beta^{N-h}\gamma_0 + \beta^{N-h+1}\gamma_1 \\
& = \frac{\alpha\beta^N [1 - (\alpha\beta)^{h-2}]}{(1-\alpha\beta)\beta^{h-2}} \gamma_1 + \frac{\alpha}{\beta-\alpha} (\beta^{N-h-1} - \alpha^{N-h-1}) \gamma_1 \\
& \quad + \beta^{N-h-1}\gamma_1 + \beta^{N-h}\gamma_0 + \beta^{N-h+1}\gamma_1 \\
& = \frac{\alpha\gamma_1\Lambda}{(\alpha-\beta)\beta^{h-2}} - \frac{\alpha^{h-1}}{\alpha-\beta}\Lambda\gamma_1 - \frac{\alpha(1-\alpha\beta)\gamma_1\Lambda}{(\alpha-\beta)^2\beta^{h+1}} + \frac{\alpha^{N-h}}{\alpha-\beta}\gamma_1 + \\
& \quad + \frac{1-\alpha\beta}{(\alpha-\beta)\beta^{h+1}}\Lambda\gamma_1 + \frac{1-\alpha\beta}{(\alpha-\beta)\beta^h}\Lambda\gamma_0 + \frac{1-\alpha\beta}{(\alpha-\beta)\beta^{h-1}}\Lambda\gamma_1 \\
& = \frac{\Lambda(1-\alpha\beta)}{(\alpha-\beta)\beta^{h+1}} \left\{ \frac{\alpha\beta^3}{1-\alpha\beta}\gamma_1 - \frac{\alpha}{\alpha-\beta}\gamma_1 + \beta^2\gamma_1 + \gamma_1 + \beta\gamma_0 \right\} + \\
& \quad + \frac{\alpha^{N-h}}{\alpha-\beta}\gamma_1 - \frac{\alpha^{h-1}}{\alpha-\beta}\Lambda\gamma_1
\end{aligned}$$

Il est évident que :

$$\frac{\alpha\beta^3}{1-\alpha\beta}\gamma_1 - \frac{\alpha}{\alpha-\beta}\gamma_1 + \beta^2\gamma_1 + \gamma_1 + \beta\gamma_0 = 0 \quad (4.34)$$

D'où

$$E(G_{N-1}(B) X_N X_h) = \frac{\alpha^{N-h}}{\alpha-\beta}\gamma_1 - \frac{\alpha^{h-1}}{\alpha-\beta}\gamma_1\Lambda \quad (4.35)$$

4.3. DÉTERMINATION DES VALEURS INITIALES D'UN PROCESSUS  
ARMA (P, Q) PAR LA MÉTHODE D'ALLER-RETOUR

---

De plus

$$\begin{aligned}
E(G_{N-1}(F) X_1 X_h) &= \sum_{j=0}^{N-1} \beta^j \gamma_{j-h-1} = \sum_{j=0}^{h-3} \beta^j \gamma_{h-j-1} + \sum_{j=h+1}^{N-1} \beta^j \gamma_{j-h+1} \\
&\quad + \beta^{h-2} \gamma_1 + \beta^{h-1} \gamma_0 + \beta^h \gamma_1 \\
&= \frac{\alpha^{h-1}}{\alpha - \beta} \gamma_1 - \frac{\alpha \beta^{h-2}}{\alpha - \beta} \gamma_1 + \frac{\alpha \beta^{h+1}}{\alpha - \beta} \gamma_1 - \frac{\alpha^{N-h}}{\alpha - \beta} \Lambda \gamma_1 + \\
&\quad + \beta^{h-2} \gamma_1 + \beta^{h-1} \gamma_0 + \beta^h \gamma_1 \\
&= \beta^{h-2} \left( \frac{\alpha \beta^3}{1 - \alpha \beta} \gamma_1 - \frac{\alpha}{\alpha - \beta} \gamma_1 + \gamma_1 + \beta \gamma_0 + \beta^2 \gamma_1 \right) + \\
&\quad + \frac{\alpha^{h-1}}{\alpha - \beta} \gamma_1 - \frac{\alpha^{N-h}}{\alpha - \beta} \Lambda \gamma_1
\end{aligned}$$

D'après l'équation (4.34), on a :

$$E(G_{N-1}(F) X_1 X_h) = \frac{\alpha^{h-1}}{\alpha - \beta} \gamma_1 - \frac{\alpha^{N-h}}{\alpha - \beta} \Lambda \gamma_1 \quad (4.36)$$

Finalement, en combinant les équations (4.35) et (4.36), on obtient

$$\begin{aligned}
E(X_0^{(\infty)} X_h) &= \frac{\alpha - \beta}{1 - \Lambda^2} [\Lambda E(G_{N-1}(B) X_N X_h) + E(G_{N-1}(F) X_1 X_h)] \\
&= \frac{\alpha - \beta}{1 - \Lambda^2} \left[ \frac{\alpha^{h-1}}{\alpha - \beta} \gamma_1 - \frac{\alpha^{h-1}}{\alpha - \beta} \Lambda^2 \gamma_1 \right] = \alpha^{h-1} \gamma_1
\end{aligned}$$

■

Ainsi, on vient d'établir que la valeur limite  $\widehat{X}_0^{(\infty)}$  coïncide avec la vraie valeur  $X^{(0)}$  du processus antérieure à l'instant  $t = 1$  relativement à la norme associée au produit scalaire défini sur l'espace de Hilbert  $H$ .

# Conclusion

Il est clair que les approximations de la fonction de vraisemblance de Whittle et de Box-Jenkins se prêtent mal aux calculs asymptotiques. Pour cela, il est possible d'introduire un troisième type d'approximation que l'on appellera méthode d'approximation de "Toeplitz-Szegö". Cette méthode est basée sur le principe que la matrice d'autocovariance d'un processus n'est autre que la transformée de Toeplitz de sa densité spectrale. L'essentiel des résultats que l'on propose sont basés sur les travaux de Szegö ([15]) et de Dzaparidge ([13]). Les propriétés de cette technique d'approximation seront très utiles pour résoudre les problèmes de tests dans lesquels interviennent les rapports des vraisemblances de deux processus gaussiens.

Nous allons nous limiter au cas des processus à temps continu et nous nous intéresserons plus particulièrement à l'approximation de la vraisemblance de deux processus gaussiens à temps continu associés à des mesures gaussiennes absolument continues. Alors, si  $h \in L^1(d\lambda)$  est une fonction, on notera  $T_N(h)$  sa transformée de Toeplitz ([15]), i.e. la matrice de taille  $(N, N)$  de terme général :

$$h_{m-n} = \int_{-\pi}^{+\pi} \exp\{i\lambda(m-n)\} h(\lambda) d\lambda \quad 1 < m, n < N$$

où  $h_k$  désigne le  $k^{\text{ème}}$  coefficient de Fourier de  $h$ .

$T_N(h)$  est une matrice hermitienne et on a en particulier :

$$\Gamma_N = E_f(X^N \tau X^N) = T_N(h)$$

La matrice de covariance est donc la transformée de Toeplitz de la densité spectrale ([15], [20]).

Dans la suite, on s'intéresse au cas des fonctions  $h$  de la classe  $\mathcal{C}$  suivante :

1.  $h$  et  $h^{-1} \in L^\infty$
2.  $\|h\|^2 = \sum_{k=-\infty}^{+\infty} k h_k^2 < \infty$

(4.37)

4.3. DÉTERMINATION DES VALEURS INITIALES D'UN PROCESSUS  
ARMA (P, Q) PAR LA MÉTHODE D'ALLER-RETOUR

---

La norme ci-dessus induit des propriétés intéressantes du type algèbre de Banach. Les densités spectrales seront choisies dans la classe  $\mathcal{C}$ . De toute manière pour avoir des théorèmes de type Szegö, des propriétés de ce genre sont indispensables. Sur l'espace de Banach des fonctions  $h$  telles que  $\|h\| < \infty$ , on peut mettre une norme équivalente ainsi définie :

$$\|h\|^2 = \sum_{n=1}^{\infty} |\omega_h(n)|^2$$

où

$$|\omega_h(n)|^2 = \sup_{|a| < \frac{1}{n}} \int |h(\lambda + a) - h(\lambda)|^2 d\lambda$$

On suppose qu'on a deux probabilités gaussiennes  $F$  et  $G$  définies sur  $\mathbb{R}^Z$  et possédant des densités spectrales  $f$  et  $g$  dans la classe  $\mathcal{C}$ . La condition 4.37 implique en particulier que pour tout  $N$ ,  $T_N f$  et  $T_N g$  sont régulières. En notant par  $L_N^S$  l'approximation de Toeplitz-Szegö du rapport de vraisemblance de  $F$  et  $G$ , on peut alors écrire :

$$L_N^S = \frac{1}{2} \{ \Sigma_N + \tau X^N ((T_N f)^{-1} - (T_N g)^{-1}) X^N \}$$

où

$$\Sigma_N = \ln \left( \frac{\det T_N f}{\det T_N g} \right)$$

Si on considère la transformation  $Y_N = T_N^{-\frac{1}{2}}(f) X^N$ , alors on a  $E_F(Y^{\tau} Y) = I_N$  (identité), et en posant

$$A_N = I - T_N^{\frac{1}{2}} f T_N^{-1} g T_N^{\frac{1}{2}} f$$

où l'on note

$$T_N^{\frac{1}{2}} f = (T_N f)^{\frac{1}{2}}$$

On obtient alors :

$$L_N^S = \frac{1}{2} (\Sigma_N + \tau Y^N A_N Y^N)$$

Par ailleurs,  $L_N^S$  peut être écrite telle que :

$$L_N^S = \frac{1}{2} \{ \Sigma_N + \tau X^N (T_N f^{-1} - T_N g^{-1}) X^N \} = \frac{1}{2} (\Sigma_N + \tau Y^N C_N Y^N)$$

*CHAPITRE 4. DÉTERMINATION DES VALEURS INITIALES POUR  
LA SOMME DES CARRÉS INCONDITIONNELLE*

---

où

$$C_N = T_N^{\frac{1}{2}} f (T_N f^{-1} - T_N g^{-1}) T_N^{\frac{1}{2}} f$$

De plus, il est possible de montrer encore que :

$$L_N^W = \frac{1}{2} \left\{ \Sigma_N + N \int I_N(\lambda) \left( \frac{1}{f(\lambda)} - \frac{1}{g(\lambda)} \right) d\lambda \right\}$$

où

$$I_N(\lambda) = \frac{1}{N} \left| \sum_{k=1}^N X_k \exp(ik\lambda) \right|^2$$

# Bibliographie

- [1] Akaiké H. Markovian representation of stochastic process by Canonical Variables, SIAM Journal on Control (1976).
- [2] Anderson O. D. Time Series Analysis : The Box-Jenkins Approach. Butterworths, London (1975)
- [3] Azencot R. et Dacunha-castelle D. Séries d'observations irrégulières, Modélisation et prévision, Masson, Paris (1984).
- [4] Bovas A. and Ledolter J. Statistical Methods for Forecasting. John Wiley and Sons, Inc (1983)
- [5] Box G.E.P. and Jenkins G.M. Time Series Analysis, Forecasting and Control. Holden Day, San Francisco (1976).
- [6] Brockwell P.J. and Davis R.A. Applications of innovation representations in time series analysis, probability and statistics, Essays in Honor of Franklin A. Graybil, J.N. Srivastava 5<sup>th</sup> ed. Elsevier, Amsterdam, 61 – 84 (1988).
- [7] Brockwell P.J. and Davis R.A. Time Series : "Theory and methods". Springer-Verlag, New York, 2<sup>nd</sup> Ed. (1996).
- [8] Broersen P.M.T. and de Waele S. "Empirical Time Series Analysis and Maximum Likelihood Estimation". Proc. 2<sup>nd</sup> IEEE Benelux Signal Processing (SPS 2000), Hilvarenbeek, The Netherlands, (2000).
- [9] Chatfield C. Time series forecasting. Chapman et Hall. CRC (2000).
- [10] Dakhmouche M. Determination of initial values for the Inconditional Likelihood Function of a Moving Average Process. Far East Journal of Theoretical Statistics, 17 (1) (2005), pp. 31 – 42.
- [11] Dakhmouche M. Approximation of the exact sum of squares function of a gaussian ARMA process, Survey in Approximation Theory, (2009) (to appear).

- [12] Dakhmouche M. Statistique des Processus Linéaires : Test et Estimation. Thèse de doctorat (2006).
- [13] Dzaparidze K. Parameter Estimation and Hypothesis Testing in Spectral Analysis of Stationary Time Series, Springer-Verlag, Tokyo (1986).
- [14] George A. F. Seber. A Matrix Handbook for Statisticians. John Wiley et Sons, Inc (2008) .
- [15] Grenander U. et Szegö G. Toeplitz forms and their applications. Univ. of Calif. Press, Berkeley (1968).
- [16] Hamilton J.D. Time Series Analysis. Princeton University press, Princeton (1994).
- [17] Hannan E.J. and Deistler M. the Statistical Theory of Linear Systems, John Wiley, New York (1988).
- [18] Palma W. Long Memory Time Series Theory and Methods. John Wiley and Sons. Inc (2007).
- [19] Pankratz A. Forecasting With Univariate Box-Jenkins Models. John Wiley and Sons. Inc (1983).
- [20] Peter B. Fourier analysis of time series. John Wiley et Sons, Inc (2000).
- [21] Philip G. and Thijsse A. Hankel and Toeplitz matrices and Forms, Algebraic Theory. Birkhäuser Boston (1982).
- [22] Robert M Gray. Toeplitz and Circulant matrices : A review. Stanford University (2001).
- [23] Stoica P. and Moses R. "Introduction to Spectral Analysis", Upper Sadd River, New Jersey, Printice Hall (1997).
- [24] Whittle P. "Estimation and information in stationary time series", Ark. Math, 2, 423 – 434, (1953).

## **Résumé :**

L'objectif de ce travail est de présenter les différentes méthodes pour exprimer la vraisemblance d'un processus aléatoire gaussien facile à utiliser dans le cadre de l'estimation des paramètres d'un modèle et pour développer des tests d'hypothèses sur ce modèle. On sait que même en temps discret les problèmes de calcul de vraisemblance ne sont jamais simples. Ainsi, nous nous sommes intéressés à l'étude de deux méthodes classiques d'approximation de la vraisemblance d'un processus gaussien. La première, appelée méthode de Whittle, utilise la représentation spectrale d'un processus gaussien, historiquement appelée périodogramme. Elle est essentiellement basée sur la transformée de Fourier d'un processus. La seconde, due à Box et Jenkins, est de loin la plus populaire et la plus utilisée. Ces auteurs utilisent la forme innovation de la vraisemblance en remplaçant la forme quadratique dans l'exposant par ce qu'ils appellent la somme des carrés inconditionnelle. On constate que cette somme de carrés n'est pas directement calculable car les innovations dépendent des observations avant l'instant d'origine. Ces valeurs initiales sont calculées par backcast et sont utilisées dans une procédure itérative pour calculer la somme des carrés inconditionnelle. Cependant, une méthode de détermination moins heuristique des valeurs initiales, appelée algorithme d'aller-retour. Finalement, nous démontrons dans le cas d'un ARMA(1,1) la convergence de cet algorithme.

**Mots clés:** Modèles linéaires ARMA - Fonction de vraisemblance - Approximation de vraisemblance - Méthode de Whittle - Algorithme d'aller-retour - Retournement du temps - Matrice d'autocovariance - Matrice de Toeplitz - Périodogramme - Densité spectrale.

## ملخص:

إنّ الهدف من هذا العمل هو تقديم مختلف الطرق لصياغة المتراجحات أو معقولات العمليات التصادفية للزمن لأنماط العشوائية الغوصية بطريقة مبسطة و استخدامها في تقدير وسائط النموذج وتطوير اختبار الفرضيات على هذا النموذج، ونعلم أنه حتى إذا كان المجال الزمني للدراسة متقطع مشكلة حساب هذه المتراجحات معقد و لهذا الغرض نحن مهتمين بدراسة نوعين من التقريب لهذه الدالة: تقريب "ويتل" الذي يستعمل التحليل الطيفي لأنماط العشوائية و هو مبني على تحويل فوريي و يسمى عادة الشكل الدوري و تقريب "بوكس-جنكينس" المبني على مجموع المربعات التي لا يمكن حسابها إلا بعد تحديد قيم للمعلومة التصادفية التي تحققت قبل الحصول على السلسلة الزمنية، وهذه القيم نحصل عليها باستعمال خوارزم الذهاب و الرجوع. في الأخير، سنقوم بدراسة تقارب هذه الخوارزم على النماذج الخطية  $ARMA(1,1)$ .

### كلمات دلالية:

النماذج الخطية  $ARMA$  - دالة المتراجحات- تقريب المتراجحات- طريقة ويتل- خوارزم الذهاب و الرجوع- دوران الوقت- مصفوفة التغيرات- مصفوفة توبليز- الشكل الدوري- الكثافة الطيفية.

## **Abstract:**

The objective of this thesis is to present a summary of principal results which led us to express the function of likelihood of Gaussian linear processes. It is known that even in discrete time the problems of calculation of likelihood are never simple. Thus, we have interested on the study of two methods of likelihood approximation: Approximation of Whittle that utilize the spectral representation of a random process historically called periodogram, it is primarily based on the Fourier's transform of this process and approximation of Box-Jenkins is in so far as the most popular, it is noted that the sum of the squares in the approximation of Box-Jenkins is not directly computable because the innovations depend on the values previous to the time series occurrence. The initial values are calculated by backcast and are used in an iterative procedure to calculate the unconditional sum of the squares. However, a method of determination less heuristic of the initial values, called algorithm of the returns. Finally, we show in the case of one ARMA(1,1) the convergence of this algorithm.

**Key Words:** Linear models ARMA - Likelihood function - Likelihood approximation  
- Method of Whittle - Algorithm of the returns - Reversal of the time - Autocovariance matrix  
- Toeplitz matrix - Periodogram - Spectral density.