

REPUBLIQUE ALGERIENNE DEMOCRATIQUE ET POPULAIRE

Ministère de l'enseignement Supérieur et de la Recherche Scientifique

**Université Mentouri de Constantine
Département de mathématiques**

N° d'Ordre:
N° de Série:

Pour l'obtention du diplôme de Magister

**Option
Mathématiques appliquées :
Statistique Des Processus Aléatoires**

Estimation Dans Les Processus Bilinéaires
" Théorie et Applications "

**Présenté par
Wafa MILI**

Devant le jury composé de:

Présidente	Mlle. BOUGHABA Soria	M. C	Université Mentouri
Rapporteur	Mr. BIBI Abdelouhab	M. C	Université Mentouri
Examineur	Mr. RAHMANI Fouad Lazhar	M. C	Université Mentouri
Examineur	Mr. BOUSSEBOUA Moussek	M. C	Université Mentouri

Soutenu le: 14- 03- 2007

Remerciements

D'abord, je tiens tout particulièrement à remercier mon directeur de thèse, Monsieur Abdelouhab Bibi qui à suivi mon travail durant tout l'année universitaire, et le support qu'il m'a apporté, il m'a donné le goût de la recherche. De plus, ses conseils scientifiques et ses encouragements ont sans aucun doute été déterminants dans l'aboutissement de cette thèse.

Je souhaite ensuite remercier également les membres du jury de ma soutenance.

Enfin, mes amis et à ceux ou celles qui sont très chers, témoins de mes joies, de mes fatigues et de mes hauts et bas, qui m'ont soutenu durant toute l'année.

Dédicace

Je dédie la présente thèse à mes parents en témoignage de leur soutien et de leurs encouragements tout au long de mes études. A mes chers frères: Saïd, Fateh. A mes chères sœurs: Malika, Fahima, Nadia, Nadjet.

Table des matières

1 Les modèles de séries temporelles non linéaires	6
1.1 Introduction	6
1.1.1 Exemples	6
1.2 Les grandes classes de modèles non linéaires	7
1.3 Différentes approches pour l'étude des modèles non linéaires	11
1.3.1 Les Chaos de Wiener	12
1.3.2 L'approche directe	12
1.3.3 La représentation markovienne	12
2 Étude probabiliste des modèles bilinéaires	14
2.1 Introduction	14
2.2 Théorèmes d'existence de solutions strictement stationnaires	15
2.3 Théorèmes d'existence de solutions stationnaires au second ordre	17
2.3.1 Le modèle $BL(p, q, p, 2)$	18
2.3.2 Le modèle $BL(p, q, p, Q)$	23
2.4 Inversibilité des modèles bilinéaires	25
2.5 Structure de covariance	27
2.5.1 Expression des moments pour le processus $BL(p, 0, p, 1)$	28
2.5.2 Expression des moments pour les processus $BL(0, 0, 2, 1)$ superdiagonaux	32
2.5.3 Expression des moments pour le processus $BL(1, 1, 1, 1)$	33
3 L'inférence statistique dans les modèles bilinéaires	35
3.1 Un aperçu sur la méthode des moments généralisés (GMM)	35
3.1.1 Exemple introductif	36
3.2 Description de la méthode des moments généralisés	37

3.2.1	Matrice de poids optimal	39
3.3	Comportement asymptotique des estimateurs	40
3.3.1	Consistance de l'estimateur à GMM	41
3.3.2	Normalité asymptotique de l'estimateur GMM	43
4	Simulation et expérimentation	47
4.1	Introduction	47
4.2	Description des modèles	47
4.2.1	Bruit blanc non linéaire	47
4.2.2	Modèle diagonal purement bilinéaire	53
4.2.3	Le modèle diagonal	57
4.2.4	Résultats et commentaires	60
4.3	Conclusion	61
5	Appendices	62
5.1	Dérivation vectorielle et matricielle	62
5.1.1	Dérivée de Ax	62
5.2	Dérivées de $z'Ax$	64
5.3	Produit de Kronecker	65
5.3.1	Propriétés	65

Introduction

Les techniques standards d'analyse de séries temporelles ont longtemps reposent sur les propriétés fondamentales de **linéarité** et de **stationnarité**. Eu égard à leurs coefficients constants et à leur dynamique linéaire, les modèles autorégressifs moyennes mobiles (*ARMA*) inhérents à Box et Jenkins [8] ont ainsi fait l'objet d'un intérêt croissant sur une vaste étendue disciplinaire allant de l'économétrie et la finance à la climatologie ou l'électrotechnique. Cependant, de nombreuses recherches ont démontré que les hypothèses de **linéarité** et de **stationnarité** n'étaient qu'un pis-aller utopique apportant un confort appréciable dans l'étude probabiliste et statistique du modèle (Guégan [13]). Le recours à des modèles plus souples est alors vivement apparu comme une nécessité dans l'analyse de séries temporelles.

Ainsi, la structure non linéaire de nombreuses séries temporelles et l'insuffisance pragmatique des modèles *ARMA* ont souligné dans un premier temps la nécessité d'introduire d'autre classes de modèles qui répondent aux défaillances portées par la classe de modèles linéaires. Par ailleurs, de nombreuses séries chronologiques possèdent des caractéristiques non-linéaires (Subba Rao et Gabr [43]; Gabr et Subba Rao [10]), ce qui exclut toute éventualité de modélisation *ARMA*. Pour pallier cette absence de linéarité, la classe des modèles bilinéaires, introduits dans la théorie du contrôle par Mohler [30], a reçu une attention exacerbée dans la littérature probabiliste et statistique dès le début des années 80 (Granger et Andersen [11]; Subba Rao [42]; Priestley [41]; Liu et Brockwell [27]; Liu [26]; Terdik [45]; et les références citées par la suite). Comme illustration, considérons la série macroéconomique mensuelle (étudiée par Subba Rao et Gabr [43]) du nombre de demandeurs d'emploi en Allemagne de l'ouest depuis Janvier 1948 à Décembre 1980 donc 396 observations (voir figure 0-1).

Cette série exhibe clairement un caractère non-linéaire associé à une évolution "périodique annuelle". Il se pose alors naturellement, les problèmes de la modélisation et de l'identification pour les séries temporelles présentant une structure non-linéaire.

La diversité des méthodes de modélisation et d'identification et notamment l'estimation, est non seulement fonction d'une évolution historique du problème, mais aussi du fait qu'il n'existe pas nécessairement une seule "bonne" façon de procéder à une telle étude. Cependant et à l'égard des méthodes d'estimation dans les processus bilinéaires nous pouvons les regrouper en deux classes distinctes

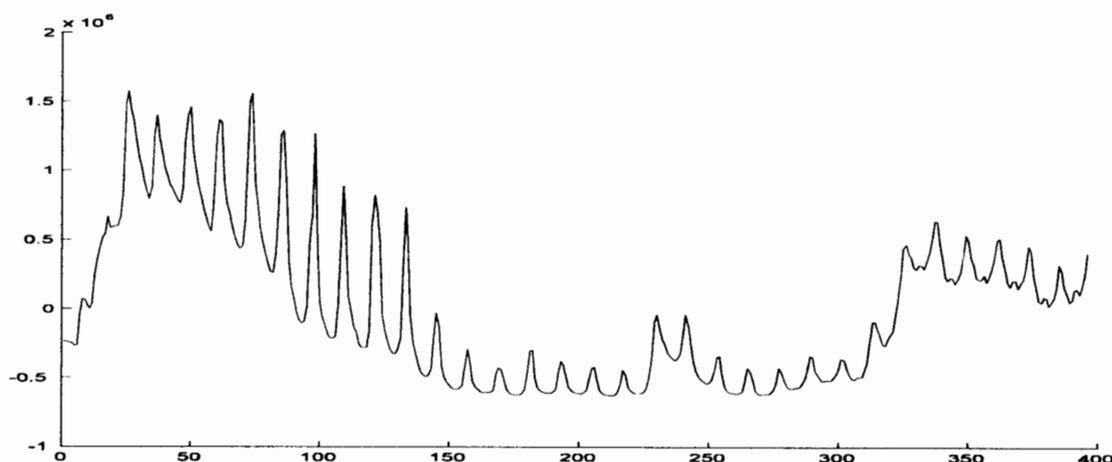


FIG. 0-1 – Série de nombre de demandeurs d'emploi en Allemagne de l'ouest depuis Janvier 1948 à Décembre 1980

1. Classe des méthodes à information complète : C'est la classe des méthodes de type Maximum de vraisemblances, Quasi maximum de vraisemblance. Nous citons par exemples Gabr [10], Guégan [12] et [13], Kim et Basawa [24], Pham Dinh [39], Priestley [41], Subba Rao [42], Subba Rao et Gabr [43] et autres.
2. Classe des méthodes à information limitée où contrairement à la classe précédente : On se contente d'estimer une équation faisant intervenir uniquement quelques informations (cf. Berlinet et Francq [4], Kim et Billard [24] et Subba Rao et Silva [44]). Nous citons les méthodes de moments et plus spécifiquement les **GMM** ou encore **MAD** (Minimum distance estimation) introduites par Hansen [21].

Notons que les propriétés probabilistes et statistiques que possèdent les **GMM** leurs ont donné un grand succès et une vaste application dans les domaines de l'économétrie, finance, économie et dans l'analyse des séries temporelles (cf. [34] pour les applications de la méthode **GMM**). L'objectif majeur de notre travail est l'estimation dans des modèles bilinéaires par la méthode **GMM** et sous certaines conditions nous établissons la consistance et la normalité asymptotique des estimateurs.

Organisation du document :

Notre thèse se décompose de quatre chapitres : Dans le chapitre 1, nous donnons une vue globale sur la classe des modèles non linéaires les plus utilisés dans la littérature, dans cette classes

en s'intéresse aux modèles bilinéaires. Dans le deuxième chapitre nous faisons une étude probabiliste des modèles bilinéaires pour établir leurs propriétés. On construit la représentation vectorielle des modèles bilinéaires ; cette représentation permet, ensuite de donner des conditions nécessaires et suffisantes d'existence et d'unicité de processus bilinéaires stationnaire au second ordre. L'existence des moments d'ordre supérieur est également abordée. Dans le troisième chapitre on rappelle en détail les théorèmes déjà existants sur la méthode **GMM** que nous avons choisie pour l'estimation des paramètres des modèles bilinéaires, nous étudions également ses propriétés statistiques, son efficacité et son comportement asymptotique. Dans le dernier chapitre nous présentons une étude expérimentale exhaustive basée sur des données simulées afin de valider la mise en évidence les résultats théoriques.

Chapitre 1

Les modèles de séries temporelles non linéaires

1.1 Introduction

Depuis, le livre de Box & Jenkins [8], et les travaux de Akaike [1] pour la classification structurelle, par exemple, les techniques de modélisation de séries temporelles ont parfaitement établi leurs maturité et connaissent un grand succès dans de nombreux domaines. Cependant, comme nous l'avons souligné précédemment, leurs champs d'application reste limité aux modèles linéaires. Nous citons à titre d'exemples :

1.1.1 Exemples

1. Lors de l'utilisation des méthodes de type Box Jenkins pour traiter et identifier des données météorologiques, les chercheurs ont constaté que ces données ne présentaient pas l'aspect classique des séries linéaires (voir figure 1-1 ci-après). Ce qui a poussé les chercheurs à étendre la classe des modèles linéaires ou à trouver d'autres modèles qui pourraient eu servir dans la modélisation de ces processus non linéaires.
2. Les phénomènes réversibles peuvent se modéliser par des modèles linéaires car ces derniers gardent leurs structures et leurs propriétés en retournant le sens du temps, mais certaines observations donnent des trajectoire de type irréversible par exemple les " montées " sont plus abruptes que les " descentes ". Ces phénomènes ne peuvent se modéliser par des modèles du type ARMA.

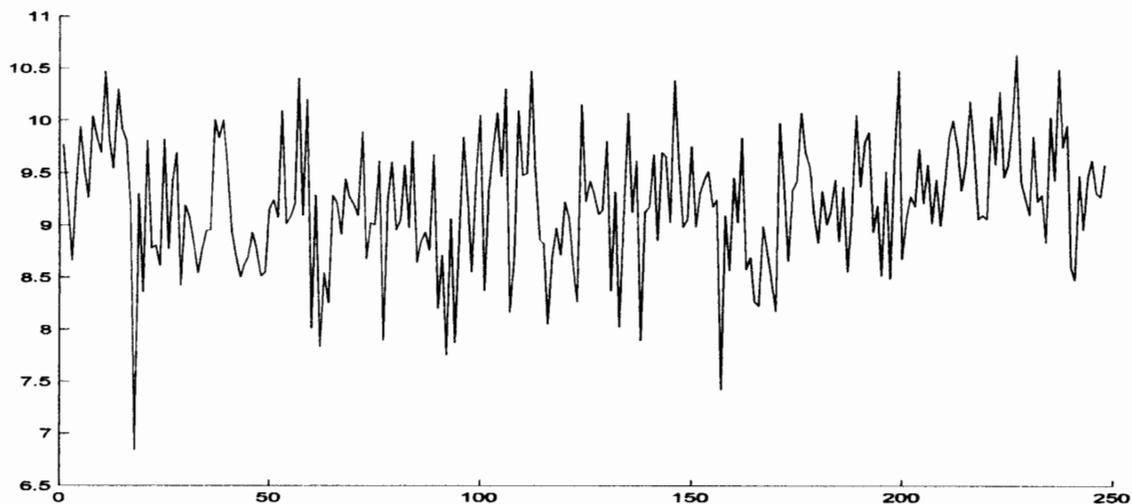


FIG. 1-1 - La température moyenne annuelle en Angleterre centrale entre 1723-1970

3. Lorsque les innovations linéaires d'un processus stationnaire $(X_t)_{t \in \mathbb{Z}}$ ne sont pas indépendantes, ce qui est le cas dans de nombreuses situations pratiques, on sait que la modélisation de $(X_t)_{t \in \mathbb{Z}}$ par une équation *ARMA* ne fournit généralement pas les meilleures prévisions. Pour répondre à ce problème la solution habituellement retenue consiste à ne pas tenir compte de la modélisation *ARMA* et à rechercher une modélisation non linéaire adéquate. La solution qui est envisagée par Francq et Berline [4] est de compléter la représentation *ARMA* par une représentation non linéaire du bruit qui intervient dans l'équation *ARMA*.

C'est pour cela qu'on a tenté de développer des classes de modèles dits "non linéaires" qui pourraient rendre compte de ces phénomènes, mais la construction d'un modèle non linéaire comporte à l'écarter quelques précautions qui ne doivent pas conduire à des solutions explosives et qui permettent de faire des prévisions fiables, très proches de la réalité, en prenant les disparités de la linéarité.

1.2 Les grandes classes de modèles non linéaires

Pour la création des modèles non linéaires, on distingue deux points de vue, le premier va consister à ne considérer que des modèles dont la structure sous-jacente ne sera plus linéaire, l'autre est celui qui prend en compte des modèles évolutifs (ce sont des modèles non stationnaires pour lesquels la moyenne et les covariances dépendent du temps, cf. [6]), le cas des problèmes issus de l'économie, de l'hydrologie, de la météorologie. Dans la suite, nous donnons les principales définitions

de modèles non linéaires les plus utilisés en pratique, avec les principales références permettant d'en approfondir l'étude (voir Guégan [13] pour un inventaire exhaustif des modèles non linéaires).

Les modèles autorégressifs généraux

Le processus $(X_t)_{t \in \mathbb{Z}}$ est dit autorégressif général d'ordre k s'il est généré par l'équation suivante

$$X_t = f(X_{t-1}, X_{t-2}, \dots, X_{t-k}, \varepsilon_t), t \in \mathbb{Z} \quad (1.1)$$

où f est une fonction non linéaire de \mathbb{R}^{k+1} à valeurs dans \mathbb{R} et où $(\varepsilon_t)_{t \in \mathbb{Z}}$ est un bruit blanc¹ et ε_t est indépendante de X_k pour tout $k < t$.

Remarque 1 Il est commode de récrire (1.1) sous la forme vectorielle, en effet, soient les vecteurs $\underline{X}_t = (X_t, X_{t-1}, \dots, X_{t-k+1})'$ et $\underline{\eta}_t = (\varepsilon_t, 0 \dots, 0)'$ de \mathbb{R}^k ; alors l'équation (1.1) peut être réécrite sous la forme

$$\underline{X}_t = \varphi(\underline{X}_{t-1}, \underline{\eta}_t) \quad (1.2)$$

où φ est une fonction de \mathbb{R}^{2k} à valeurs dans \mathbb{R}^k définie par $X_t^{(1)} = f(X_{t-1}^{(1)} \dots X_{t-1}^{(k)}, \eta_t^{(1)})$, $X_t^{(2)} = X_{t-1}^{(1)} \dots X_{t-1}^{(k-1)} = X_{t-1}^{(k-1)}$.

Les modèles autorégressifs exponentiels (EXPAR)

Un processus $(X_t)_{t \in \mathbb{Z}}$ est dit autorégressif exponentiel d'ordre k (EXPAR(k)) si pour tout $t \in \mathbb{Z}$

$$X_t = \sum_{j=1}^k (\alpha_j + \beta_j \exp(-\delta X_{t-1}^2)) X_{t-j} + \varepsilon_t, \delta > 0 \quad (1.3)$$

où $(\varepsilon_t)_{t \in \mathbb{Z}}$ est un bruit blanc. Ces modèles ont été introduits par Haggan et Ozaki [19], leurs propriétés probabilistes ont été étudiées par Ozaki [35]. Ces modèles trouvent leurs applications dans différents champs d'application.

¹Une suite $(\varepsilon_t)_{t \in \mathbb{Z}}$ de variables aléatoires de moyenne 0 et de variance σ^2 est appelé bruit blanc fort (respectivement faible) si elle est indépendante et identiquement distribuée (*i.i.d.*) (respectivement $Cov(\varepsilon_t, \varepsilon_{t+h}) = 0 \forall t, h \in \mathbb{Z}$ et $h \neq 0$)

Les modèles autorégressifs fractionnaires (*FAR*)

Un processus $(X_t)_{t \in \mathbb{Z}}$ est dit autorégressif fractionnaire d'ordre 1, si pour tout $t \in \mathbb{Z}$

$$X_t = \frac{a_0 + \sum_{j=1}^p a_j X_{t-1}^j}{b_0 + \sum_{j=1}^q b_j X_{t-1}^j} + \varepsilon_t \quad (1.4)$$

où $(\varepsilon_t)_{t \in \mathbb{Z}}$ est un bruit blanc, $0 \leq p \leq q + 1 < \infty$, $a_p \neq 0$, $b_q \neq 0$. Ces modèles ont été proposés originairement par Jones [22]. Avec une étude expérimentale pour la prédiction non linéaire en météorologie.

Les modèles autorégressifs produits (*PAR*)

Les modèles autorégressifs produits ont pour expression

$$X_t = \xi_t X_{t-1}^\alpha \quad (1.5)$$

où, $\alpha \in [0, 1]$, et $\xi_t = \exp(\varepsilon_t)$, $(\varepsilon_t)_{t \in \mathbb{Z}}$ étant un bruit blanc. Ces modèles ont été introduits par McKenzie [29]. Actuellement, il n'existe pas d'applications connues de ces modèles.

Les modèles autorégressifs à coefficients aléatoires (*RCAR*)

Les modèles autorégressifs à coefficients aléatoires d'ordre k (*RCAR*(k)) ont pour expression :

$$X_t = \sum_{i=1}^k (B_i + B_i(t)) X_{t-i} + \varepsilon_t \quad (1.6)$$

où

1. $(\varepsilon_t)_{t \in \mathbb{Z}}$ est un bruit blanc.
2. Les B_i , $i = 1, \dots, k$, sont des constantes.
3. $(B(t))_{t \in \mathbb{Z}}$ est indépendante de $(\varepsilon_t)_{t \in \mathbb{Z}}$ où $B(t) = (B_k(t), \dots, B_1(t))$, $(B(t))$ est une suite indépendante de $1 \times k$ vecteurs de moyenne 0 et $E[B^T(t) B(t)] = \Sigma$.

Ces modèles ont été introduits par Andel [3], Nicholls et Quinn [33]².

²Notons que le modèle 1.6 admet une représentation vectorielle similaire à la représentation (1.2)

Les Modèles bilinéaires (BL)

Le processus $(X_t)_{t \in \mathbb{Z}}$ défini sur un certain espace de probabilité $(\Omega, \mathfrak{F}, P)$ est dit bilinéaire d'ordre p, q, P, Q et noté $BL(p, q, P, Q)$ s'il satisfait l'équation suivante :

$$X_t = \sum_{i=1}^p a_i X_{t-i} + \sum_{i=1}^q c_i \varepsilon_{t-i} + \sum_{j=1}^P \sum_{k=1}^Q b_{jk} X_{t-j} \varepsilon_{t-k} + \varepsilon_t \quad (1.7)$$

où $(\varepsilon_t)_{t \in \mathbb{Z}}$ est un processus de bruit blanc défini sur le même espace $(\Omega, \mathfrak{F}, P)$, $(a_i)_{1 \leq i \leq p}$, $(c_i)_{1 \leq i \leq q}$ et $(b_{jk})_{1 \leq j \leq P, 1 \leq k \leq Q}$ sont des suites de constantes réelles.

Ces modèles ont été présentés par Granger et Andersen [16] et les principaux travaux sont ceux de Subba Rao et Gabr [43]. Ces modèles ont été utilisés dans la modélisation des phénomènes complexes et notamment dans la chimie, la météorologie, la physique et l'astronomie.

Remarque 2 De (1.7) on retrouve les trois classes de modèles bilinéaires :

1. Les modèles diagonaux sont ceux pour lesquels $b_{ij} = 0, \forall i \neq j$
2. Les modèles superdiagonaux sont ceux pour lesquels $b_{ij} = 0, \forall i \leq j$
3. Les modèles sousdiagonaux ceux pour lequel $b_{ij} = 0, \forall i > j$.

Les modèles autorégressifs conditionnellement hétéroscédastiques (ARCH)

Un processus $(X_t)_{t \in \mathbb{Z}}$ est dit autorégressif conditionnellement hétéroscédastique d'ordre p ($ARCH(p)$) si les équations suivantes sont vérifiées

$$X_t = \varepsilon_t h_t^{\frac{1}{2}} \text{ et } h_t = \alpha_0 + \sum_{i=1}^p \alpha_i X_{t-i}^2 \quad (1.8)$$

avec $\alpha_0 > 0, \alpha_i \geq 0$ pour $i = 1, \dots, p$ et $(\varepsilon_t)_{t \in \mathbb{Z}}$ est un processus de bruit blanc gaussien centré de variance 1. Cependant le processus $(X_t)_{t \in \mathbb{Z}}$ apparaît comme un bruit blanc gaussien, multiplié à chaque instant t par une variable aléatoire dont le carré dépend (de manière linéaire) des p valeurs passées du carré du processus. On peut écrire le modèle (1.8) sous la forme $X_t | X_{t-1} \rightsquigarrow N(0, h_t)$ et $h_t = \alpha_0 + \sum_{i=1}^p \alpha_i X_{t-i}^2$ ce qui montre que la distribution conditionnelle de X_t est normale centrée et de variance h_t .

Remarque 3 Une troisième formulation fait apparaître le processus ARCH comme une différence

de martingale, puisque $E\{X_t|\mathfrak{S}_{t-1}\} = 0$ mais hétéroscédastique puisque $Var\{X_t|\mathfrak{S}_{t-1}\} = \alpha_0 + \sum_{i=1}^p \alpha_i X_{t-i}^2$ où $\mathfrak{S}_t := \sigma\{X_s, s \leq t\}$.

Remarque 4 Il est évident que le processus $(X_t^2)_{t \in \mathbb{Z}}$ est un processus bilinéaire si $(X_t)_{t \in \mathbb{Z}}$ est un processus ARCH. Inversement, il existe des processus bilinéaires qui se comporte comme des processus ARCH. En effet le processus bilinéaire stationnaire au second ordre $(X_t^2)_{t \in \mathbb{Z}}$ généré par $X_t = bX_{t-2}\varepsilon_{t-1} + \varepsilon_t$ satisfait

$$E\{X_t|\mathfrak{S}_{t-1}\} = 0 \text{ et } Var\{X_t|\mathfrak{S}_{t-1}\} = \sigma^2 + \sigma^2 b^2 X_{t-2}^2.$$

Les modèles ARCH ont été proposés par Engle [9]. Les propriétés probabilistes et statistiques ont été étudiées par plusieurs auteurs (voir en particulier Engle [9] et Hamilton [20]). Ses applications en économétrie, finances et en économie sont très développées ([15]).

Les modèles ARCH généralisés (GARCH)

Un processus $(X_t)_{t \in \mathbb{Z}}$ est dit un processus GARCH s'il est généré par

$$\begin{cases} X_t = \varepsilon_t h_t^{\frac{1}{2}} \\ h_t = \alpha_0 + \sum_{i=1}^p \alpha_i X_{t-i}^2 + \sum_{j=1}^q \beta_j h_{t-j} \end{cases} \quad (1.9)$$

où $(\varepsilon_t)_{t \in \mathbb{Z}}$ est un processus de bruit blanc, et où avec $\alpha_0 > 0, \alpha_i \geq 0$ pour $i = 1, \dots, p, \beta_j \geq 0, j = 1, \dots, q$. Cette classe de modèles est due à Bollerslev [7]. Lorsque le processus $(\varepsilon_t)_{t \in \mathbb{Z}}$ est gaussien, les formulations décrites pour les processus ARCH sont manifestement valables pour les processus GARCH. Les modèles GARCH sont en effet susceptibles de capter les propriétés caractéristiques des modèles linéaires ARMA.

1.3 Différentes approches pour l'étude des modèles non linéaires

Vu la difficulté de la structure profonde des modèles non linéaires, on rencontre plusieurs problèmes concernant leurs définitions car si on s'intéresse par exemple, par les outils utilisés classiquement en théorie des modèles linéaires ou l'outil hilbertien on trouve qu'ils sont totalement

insuffisants pour les étudier. Pour cela différentes approches ont été utilisées certaines ont conduit à des résultats qui ne concernent que des cas particuliers, par contre d'autres ont permis d'obtenir des résultats théoriques plus généraux. Citons, dans ce qui suit, les approches fréquemment utilisées en analyses des séries temporelles.

1.3.1 Les Chaos de Wiener

Les chaos de Wiener forment un cadre plus général pour l'étude des modèles non linéaire. En effet, si $(\varepsilon_t)_{t \in \mathbb{Z}}$ est un bruit blanc gaussien, soit $F(\varepsilon) = \sigma\{\varepsilon_t, t \in \mathbb{Z}\}$, l'approche par les chaos de Wiener est basée sur l'identification de deux espaces $\mathbb{L}_2(F(\varepsilon))$ et $\bigoplus_{n \geq 0}^{\perp} \mathcal{H}^{\odot n}(\varepsilon)$ (chaos de Wiener) la somme directe orthogonale des sous espaces $\mathcal{H}^{\odot n}(\varepsilon)$ fermés engendrés par les combinaisons linéaires des produits symétriques $\varepsilon_{t_1} \odot \varepsilon_{t_2} \odot \dots \odot \varepsilon_{t_n}$ où $t_1, \dots, t_n \in \mathbb{Z}$. Mais l'algèbre sous jacente aux chaos est très compliquée par rapport à l'algèbre des polynômes d'observations qui ont des propriétés moins intéressantes et on est alors conduit à la représentation de Volterra à partir de laquelle on peut construire des tests permettant d'identifier ces modèles ou de les différentier composante par composante.

1.3.2 L'approche directe

L'étude des modèles non linéaires a été restreinte tout d'abord à l'existence de tels modèles stationnaires. Certains auteurs se sont intéressés aux problèmes de l'inversibilité³, de l'ergodicité⁴, de la structure de covariances et les moments d'ordre supérieur. Subba Rao et Gabr [43] et Terdik [45] ont d'une part présenté une analyse spectrale des modèles bilinéaires pour des cas particulier et, d'autre part, ont montré que les propriétés du second ordre ne suffisent pas pour identifier les modèles bilinéaires.

1.3.3 La représentation markovienne

A partir de la représentation Markovienne dans la théorie des modèles linéaires (qui joue un rôle important dans la théorie de filtrage), Pham [36] a établi cette représentation pour les modèles bilinéaires superdiagonaux et en augmentant suffisamment l'espace d'état, il a établi pour des modèles

³Un processus $(X_t)_{t \in \mathbb{Z}}$ est dit inversible si son innovation ε_t est $\mathfrak{F}_t^{(X)}$ mesurable où $\mathfrak{F}_t^{(X)} := \sigma(X_s, s \leq t)$

⁴Un exposé complet sur le concept d'ergodicité peut être trouver dans Breiman, L. (1968). Probability, Addison-Wesley publishing company, Massachusetts.

bilinéaires généraux. Ces représentations Markoviennes ont permis d'avoir l'existence d'une solution stationnaire unique, la minimalité de la représentation, l'ergodicité, l'inversibilité et les calculs des moments. Cette représentation peut être établie pour d'autres classes de processus non linéaires à savoir les processus *GARCH* et les processus *RCAR*.

Chapitre 2

Étude probabiliste des modèles bilinéaires

2.1 Introduction

Dans ce chapitre nous passons en revue les premières études sur la structure probabiliste des modèles bilinéaires d'ordre (p, q, P, Q) , noté $BL(p, q, P, Q)$ défini par

$$X_t = \sum_{i=1}^p a_i X_{t-i} + \sum_{i=1}^q c_i \varepsilon_{t-i} + \sum_{j=1}^P \sum_{k=1}^Q b_{jk} X_{t-j} \varepsilon_{t-k} + \varepsilon_t \quad (2.1)$$

où $(\varepsilon_t)_{t \in \mathbb{Z}}$ est un bruit blanc gaussien, $(a_i)_{1 \leq i \leq p}$, $(c_i)_{1 \leq i \leq q}$ et $(b_{jk})_{1 \leq j \leq P, 1 \leq k \leq Q}$ sont des suites de constantes réelles. Les principaux thèmes qui ont retenu l'attention des chercheurs sont l'existence et l'unicité d'une solution stationnaire (au second ordre ¹ ou strict²), l'inversibilité, l'ergodicité, la structure de covariance et l'existence des moments d'ordre supérieur. A cause de la complexité structurale du modèle, peu de résultats sont généraux. On peut distinguer trois sous classes de modèles bilinéaires : Les modèles bilinéaires diagonaux pour lesquels : $b_{ij} = 0$ pour tout $i \neq j$, les modèles bilinéaires superdiagonaux pour lesquels : $b_{ij} = 0$ si $j < i$, les modèles bilinéaires sousdiagonaux pour lesquels : $b_{ij} = 0$ si $j > i$. La plupart des résultats, dont on dispose actuellement, ne concernent, en fait, qu'une ou deux de ces sous classes. Par exemple, si l'existence d'une solution

¹Un processus $(X_t)_{t \in \mathbb{Z}}$ est dit stationnaire au second-ordre si $\forall t, h \in \mathbb{Z} : E\{X_t^2\} < +\infty, E\{X_t\} = m$ et $Cov(X_{t+h}X_t) = \gamma(h)$

²Un processus $(X_t)_{t \in \mathbb{Z}}$ est dit strictement stationnaire si les vecteurs $(X_1, \dots, X_k)'$ et $(X_{1+h}, \dots, X_{k+h})'$ ont même loi jointe pour tout entier k et tout entier relatif h

stationnaire unique est un problème résolu pour certains modèles superdiagonaux, les résultats sont parcellaires pour l'ensemble des modèles sousdiagonaux et diagonaux. Les principaux résultats présentés sur ce sujet sont dus aux chercheurs suivants : Pham Dinh et Tran [39] ont fait une étude complète du modèle diagonal d'ordre 1, Subba Rao, Bhaskara Rao, Subba Rao et Walker [5], Akamanam, Bhaskara Rao et Subramanyam [2].

2.2 Théorèmes d'existence de solutions strictement stationnaires

Sans perdre en généralité, on suppose que $p = P$ (sinon des zéros peuvent être ajoutés), soient les vecteurs : $\underline{X}_t := (X_t, X_{t-1}, \dots, X_{t-p+1})' \in \mathbb{R}^p$, $\underline{\eta}_t := \left(\varepsilon_t + \sum_{j=1}^q c_j \varepsilon_{t-j}, 0, \dots, 0 \right)' \in \mathbb{R}^p$, $H := (1, 0, \dots, 0)' \in \mathbb{R}^p$ et soient les matrices

$$A = \begin{pmatrix} a_1 & a_2 & \dots & \dots & a_p \\ 1 & 0 & \dots & \dots & 0 \\ 0 & 1 & 0 & \dots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \dots & 0 & 1 & 0 \end{pmatrix}_{p \times p} \quad B_j = \begin{pmatrix} b_{1j} & \dots & \dots & b_{pj} \\ 0 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots \\ 0 & \dots & 0 & 0 \end{pmatrix}_{p \times p}$$

alors avec les notations précédentes, nous avons la représentation markovienne suivante

$$\begin{cases} \underline{X}_t = A(t)\underline{X}_{t-1} + \underline{\eta}_t \\ A(t) := A + \sum_{j=1}^Q B_j \varepsilon_{t-j} \end{cases} \quad (2.2)$$

et $X_t = H' \underline{X}_t$. Il est bien clair que le processus $(A(t), \eta_t)_{t \in \mathbb{Z}}$ est strictement stationnaire et ergodique, de plus on a $E \left\{ \ln^+ \sum_{i=1}^P \left| a_i + \sum_{j=1}^Q b_{ij} \varepsilon_{t-j} \right| \right\}^3 < \infty$. Cependant l'exposant de Lyapunov noté δ_L

³ $\log^+(x) = \max(0, \log x)$, $\forall x > 0$

associé à la suite des matrices aléatoires défini par

$$\delta_L = \inf_{n>0} \left\{ \frac{1}{n} E \left\{ \ln \left\| \prod_{j=0}^{n-1} A(t-j) \right\| \right\} \right\}^4 \quad (2.3)$$

existe et on a, presque sûrement $\delta_L = \lim_{n \rightarrow \infty} \left\{ \frac{1}{n} \left\{ \ln \left\| \prod_{j=0}^{n-1} A(t-j) \right\| \right\} \right\}$. D'où

Théorème 5 Soit δ_L défini par (2.3). Supposons que l'exposant de Lyapunov est strictement négatif, alors $\forall t \in \mathbb{Z}$ la série

$$\underline{X}_t = \sum_{l=1}^{\infty} \left\{ \prod_{j=0}^{l-1} A(t-j) \right\} \underline{\eta}_{t-l} + \underline{\eta}_t \quad (2.4)$$

converge presque sûrement et le processus $(X_t)_{t \in \mathbb{Z}}$ (défini comme étant la première composante de $(\underline{X}_t)_{t \in \mathbb{Z}}$) constitue l'unique solution strictement stationnaire et ergodique de (1.7).

Exemple 6 Considérons le modèle

$$X_t = a_1 X_{t-1} + \varepsilon_t + b_{11} X_{t-1} \varepsilon_{t-1} \quad (2.5)$$

où $(\varepsilon_t)_{t \in \mathbb{Z}}$ est un bruit blanc gaussien de moyenne 0 et variance σ^2 . Alors

$$\delta_L := \lim_{n \rightarrow \infty} \left\{ \frac{1}{2n} \sum_{j=0}^{n-1} \ln (a_1 + b_{11} \varepsilon_{t-j-1})^2 \right\}$$

converge vers $\frac{1}{2} E \{ \ln (a_1 + b_{11} \varepsilon_t)^2 \}$. D'après la loi forte des grands nombres et par l'inégalité de Jensen, on a

Théorème 7 (Pham et Tran) [39] Si $a_1^2 + b_{11}^2 \sigma^2 < 1$ alors, il existe un unique processus $(X_t)_{t \in \mathbb{Z}}$ strictement stationnaire et ergodique satisfaisant (2.5) donné par

$$X_t = \varepsilon_t + \sum_{j=1}^{\infty} \left[\prod_{k=1}^j \{a_1 + b_{11} \varepsilon_{t-k}\} \right] \varepsilon_{t-j} \quad (2.6)$$

où la série précédente converge presque sûrement.

⁴ $\|\cdot\|$ est une norme matricielle

Notons que le théorème (5) donne une condition suffisante d'existence et d'unicité de solution strictement stationnaire basée sur l'exposant de Lyapunov. Néanmoins, cette condition n'a pas beaucoup d'intérêt pratique car elle dépend de calculs de produit infini de matrices aléatoires, ce qui rend délicat et même parfois impossible d'effectuer ces calculs lorsque p est "assez grand". D'autre part, d'après la définition de δ_L , on a $\delta_L \leq E \{\ln \|A(t) A(t-1) \cdots A(t-p)\|\}$ donc la condition

$$\delta(p) := E \{\ln \|A(t) A(t-1) \cdots A(t-p)\|\} < 0 \quad (2.7)$$

pourrait être employée comme étant une condition suffisante pour l'existence de solution strictement stationnaire et on a

Théorème 8 *Si la condition (2.7) est vérifiée, alors le processus $(\underline{X}_t)_{t \in \mathbb{Z}}$ défini par (2.2) admet une unique solution strictement stationnaire donnée par (2.4) dont la première composante constitue l'unique solution strictement stationnaire et ergodique de (1.7).*

Exemple 9 *Considérons le modèle bilinéaire diagonal*

$$X_t = \sum_{i=1}^2 (a_i + b_{ii}\varepsilon_{t-i}) X_{t-i} + \varepsilon_t \quad (2.8)$$

et soit la norme $\|A\| = \max_i \left\{ \sum_{j=1}^p |a_{ij}| \right\}$, nous avons

$$\|A(t) A(t-1)\| = \max \{ |a_1 + b_{11}\varepsilon_{t-2}| + |a_2 + b_{22}\varepsilon_{t-3}|, |a_1 + b_{11}\varepsilon_{t-1}| |a_1 + b_{11}\varepsilon_{t-2}| + |a_2 + b_{22}\varepsilon_{t-2}| + |a_1 + b_{11}\varepsilon_{t-1}| |a_2 + b_{22}\varepsilon_{t-3}| \}$$

par conséquent $E \{\ln \|A(t) A(t-1)\|\} < 0$ si $(A+B)(A+2) < 1$ où $A := E \{|a_1 + b_{11}\varepsilon_t|\}$ et $B := E \{|a_2 + b_{22}\varepsilon_t|\}$.

2.3 Théorèmes d'existence de solutions stationnaires au second ordre

On cherche à résoudre le problème suivant : Etant donné un bruit blanc gaussien centré, de variance finie $\sigma^2 < \infty$, et les matrices $A, B_j, j = 1, \dots, Q$, existe-t-il un processus stationnaire au second-ordre unique $(\underline{X}_t)_{t \in \mathbb{Z}}$ satisfaisant (2.2) ?

2.3.1 Le modèle $BL(p, q, p, 2)$

Le théorème suivant donne une condition suffisante pour que l'équation (2.2) avec $Q = 2$, i.e.,

$$\underline{X}_t = A\underline{X}_{t-1} + (B_1\varepsilon_{t-1} + B_2\varepsilon_{t-2})\underline{X}_{t-1} + \underline{\eta}_t \quad (2.9)$$

admette une solution stationnaire au second ordre.

Théorème 10 (*Liu et Brockwell*) [28] *Soit le processus $(\underline{X}_t)_{t \in \mathbb{Z}}$ défini par (2.9). Posons :*

$$\Gamma := \begin{pmatrix} A \otimes A + \sigma^2 B_1 \otimes B_1 & A \otimes B_2 + B_2 \otimes A & B_2 \otimes B_2 \\ \sigma^2 (A \otimes B_1 + B_1 \otimes A) & \sigma^2 (B_2 \otimes B_1 + B_1 \otimes B_2) & 0 \\ \sigma^2 (A \otimes A + 3\sigma^2 B_1 \otimes B_1) & \sigma^2 (A \otimes B_2 + B_2 \otimes A) & \sigma^2 (B_2 \otimes B_2) \end{pmatrix}$$

Alors si

$$\lambda := \rho(\Gamma) < 1 \quad (2.10)$$

l'équation (2.9) admet une unique solution causal⁵ stationnaire au second ordre donnée par

$$\underline{X}_t = \underline{\eta}_t + \sum_{n=1}^{\infty} \left\{ \prod_{j=1}^n (A + B_1\varepsilon_{t-j} + B_2\varepsilon_{t-j-1}) \right\} \underline{\eta}_{t-n}$$

dont la première composante X_t est l'unique solution stationnaire au second ordre, causal et ergodique de (2.1). La solution est aussi strictement stationnaire.

Avant de donner les grandes lignes de la démonstration, nous énonçons quelques corollaires

Corollaire 11 *Pour le modèle (2.5), la condition (2.10) se réduit à la condition donnée par le théorème (7).*

Preuve. Dans ce cas on a $A = a_1$ et $B_1 = b_{11}$ et $B_2 = 0$, donc

$$\Gamma := \begin{pmatrix} a_1^2 + \sigma^2 b_{11}^2 & 0 & 0 \\ 2\sigma^2 a_1 b_{11} & 0 & 0 \\ \sigma^2 a_1^2 + 3\sigma^4 b_{11}^2 & 0 & 0 \end{pmatrix}$$

⁵Le processus bilinéaire $BL(p, qP, Q)$ est dit causal s'il existe une fonction mesurable $f : \mathbb{R}^\infty \rightarrow \mathbb{R}$ tel que $X_t = f(\varepsilon_t, \varepsilon_{t-1}, \dots)$ presque sûrement pour tout $t \in \mathbb{Z}$.

et un simple calcul montre que les valeurs propres de la matrice Γ sont $\{0, 0, a_1^2 + \sigma^2 b_{11}^2\}$. D'où le résultat. ■

Corollaire 12 *Pour le modèle $BL(p, q, p, 1)$, i.e.,*

$$\underline{X}_t = A\underline{X}_{t-1} + B_1\underline{X}_{t-1}\varepsilon_{t-1} + \underline{\eta}_t \quad (2.11)$$

la condition (2.10) se réduit à la condition $\rho(A \otimes A + \sigma^2 B_1 \otimes B_1) < 1$.

Preuve. Dans ce cas, la matrice Γ prend la forme

$$\Gamma := \begin{pmatrix} A \otimes A + \sigma^2 B_1 \otimes B_1 & 0 & 0 \\ \sigma^2 (A \otimes B_1 + B_1 \otimes A) & 0 & 0 \\ \sigma^2 (A \otimes A + 3\sigma^2 B_1 \otimes B_1) & 0 & 0 \end{pmatrix}$$

et par conséquent, les valeurs propres non nulles de Γ sont les mêmes que celles de $A \otimes A + \sigma^2 B_1 \otimes B_1$. Ainsi $\rho(\Gamma) = \rho(A \otimes A + \sigma^2 B_1 \otimes B_1)$. ■

Corollaire 13 *Dans les modèles $ARMA(p, q)$, la condition (2.10) se réduit à la condition $\rho(A) < 1$.*

Preuve. Dans ce cas, les coefficients b_{ij} sont nuls et la matrice Γ prend la forme

$$\Gamma := \begin{pmatrix} A \otimes A & 0 & 0 \\ 0 & 0 & 0 \\ \sigma^2 A \otimes A & 0 & 0 \end{pmatrix}$$

et par conséquent, les valeurs propres non nulles de Γ sont celles de $A \otimes A$. Ainsi $\rho(\Gamma) = \rho(A \otimes A) = \rho(A)$. ■

Preuve. [Du Théorème 10] Pour montrer le théorème 10, on introduit les deux processus vectoriels suivants :

$$\underline{S}_n(t) = \begin{cases} 0 & \text{si } n < 0 \\ \underline{\eta}_t & \text{si } n = 0 \\ \underline{\eta}_t + (A + B_1\varepsilon_{t-1} + B_2\varepsilon_{t-2})\underline{S}_{n-1}(t-1) & \text{si } n \geq 1 \end{cases} \quad (2.12)$$

et soit $\underline{\Delta}_n(t) = \underline{S}_n(t) - \underline{S}_{n-1}(t)$. Alors d'après la définition de $\Delta_n(t)$ nous avons

$$\underline{\Delta}_n(t) = \begin{cases} 0 & \text{si } n < 0 \\ \underline{\eta}_t & \text{si } n = 0 \\ (A + B_1\varepsilon_{t-1} + B_2\varepsilon_{t-2})\underline{\Delta}_{n-1}(t-1) & \text{si } n \geq 1 \end{cases} \quad (2.13)$$

On peut construire, à partir de ces deux processus, les fonctions mesurables $g_n : \mathbb{R}^\infty \rightarrow \mathbb{R}^p$ et $G_n : \mathbb{R}^\infty \rightarrow \mathbb{R}^p$ telles que pour tout $t \in \mathbb{Z}$, $\underline{\Delta}_n(t) = g_n(\varepsilon_{t-1}, \varepsilon_{t-2}, \dots)$, $n \geq 1$ et $\underline{S}_n(t) = G_n(\varepsilon_t, \varepsilon_{t-1}, \dots)$, $n \geq 0$. Alors pour n fixé et $n \geq 1$ à partir de deux fonctions $g_n(\cdot)$ et $G_n(\cdot)$ on déduit que le processus $(\underline{\Delta}_n(t), \underline{S}_n(t))_{t \in \mathbb{Z}}$ est strictement stationnaire pour $n \in \mathbb{Z}$. Calculons maintenant les matrices des moments :

$$\begin{cases} V_n = E \left(\underline{\Delta}_n(t) \underline{\Delta}_n'(t) \right) \\ D_n = E \left(\varepsilon_{t-1} \underline{\Delta}_n(t) \underline{\Delta}_n'(t) \right) \\ F_n = E \left(\varepsilon_{t-1}^2 \underline{\Delta}_n(t) \underline{\Delta}_n'(t) \right) \end{cases} \quad (2.14)$$

Pour $n \geq 3$, on a :

$$\begin{aligned} V_n &= E \left[(A + B_1\varepsilon_{t-1} + B_2\varepsilon_{t-2}) \underline{\Delta}_{n-1}(t-1) \underline{\Delta}_{n-1}'(t-1) (A' + B_1'\varepsilon_{t-1} + B_2'\varepsilon_{t-2}) \right] \\ &= AV_{n-1}A' + AD_{n-1}B_2' + B_2D_{n-1}A' + \sigma^2 B_1V_{n-1}B_1' + B_2F_{n-1}B_2' \\ D_n &= E \left[\varepsilon_{t-1} (A + B_1\varepsilon_{t-1} + B_2\varepsilon_{t-2}) \underline{\Delta}_{n-1}(t-1) \underline{\Delta}_{n-1}'(t-1) (A' + B_1'\varepsilon_{t-1} + B_2'\varepsilon_{t-2}) \right] \\ &= \sigma^2 (AV_{n-1}B_1' + B_1V_{n-1}A') + \sigma^2 (B_2D_{n-1}B_1' + B_1D_{n-1}B_2') \\ F_n &= E \left[\varepsilon_{t-1}^2 (A + B_1\varepsilon_{t-1} + B_2\varepsilon_{t-2}) \underline{\Delta}_{n-1}(t-1) \underline{\Delta}_{n-1}'(t-1) (A' + B_1'\varepsilon_{t-1} + B_2'\varepsilon_{t-2}) \right] \\ &= \sigma^2 AV_{n-1}A' + 3\sigma^4 B_1V_{n-1}B_1' + \sigma^2 (AD_{n-1}B_2' + B_2D_{n-1}A') + \sigma^2 B_2F_{n-1}B_2' \end{aligned}$$

Ces expressions ont une forme plus agréable si on introduit la notation $\vec{M} = \text{Vec}(M)$, où M est une matrice. On peut alors écrire :

$$\begin{aligned} \vec{V}_n &= (A \otimes A + \sigma^2 B_1 \otimes B_1) \vec{V}_{n-1} + (A \otimes B_2 + B_2 \otimes A) \vec{D}_{n-1} + (B_2 \otimes B_2) \vec{F}_{n-1} \\ \vec{D}_n &= \sigma^2 (A \otimes B_1 + B_1 \otimes A) \vec{V}_{n-1} + \sigma^2 (B_2 \otimes B_1 + B_1 \otimes B_2) \vec{D}_{n-1} \\ \vec{F}_n &= (\sigma^2 A \otimes A + 3\sigma^4 B_1 \otimes B_1) \vec{V}_{n-1} + \sigma^2 (A \otimes B_2 + B_2 \otimes A) \vec{D}_{n-1} + \sigma^2 (B_2 \otimes B_2) \vec{F}_{n-1} \end{aligned}$$

ou de manière équivalente $\underline{u}_n = \Gamma \underline{u}_{n-1} = \Gamma^2 \underline{u}_{n-2} = \dots = \Gamma^{n-3} u_3$ où

$$\underline{u}_n = \begin{bmatrix} \vec{V}_n \\ \vec{D}_n \\ \vec{F}_n \end{bmatrix}, n \geq 3.$$

1. Nous allons montrer que le processus $\underline{S}_n(t)$ converge en moyenne quadratique pour chaque t et que les limites : $\underline{X}_t := \lim_{n \rightarrow \infty} \underline{S}_n(t)$ constituent une solution strictement stationnaire de (2.9). D'après (2.10), on a l'inégalité suivante $|\underline{u}_n| \leq c\lambda^{\frac{n}{2}}$ où c est une certaine constante. On déduit que :

$$E \|\underline{S}_n(t) - \underline{S}_{n-1}(t)\|^2 = E \|\underline{\Delta}_n(t)\|^2 = \text{trace}(V_n) \leq |\underline{u}_n| \leq c\lambda^{\frac{n}{2}}$$

et donc pour chaque t fixé, $\underline{S}_n(t)$ converge à la fois dans \mathbb{L}^2 et en probabilité 1 quand $n \rightarrow \infty$. Il existe donc une fonction mesurable $f_t : R^\infty \rightarrow R^p$, telle que

$$\underline{X}_t = \lim_{n \rightarrow \infty} \underline{S}_n(t) = f_t(\varepsilon_t, \varepsilon_{t-1}, \dots) \quad (2.15)$$

Pour montrer que \underline{X}_t est strictement stationnaire, on considère une fonction h bornée continue sur R^{ps} telle que :

$$\begin{aligned} E \{h(\underline{X}_{t+1}, \dots, \underline{X}_{t+s})\} &= \lim_{n \rightarrow \infty} E \{h(\underline{S}_n(t+1), \dots, \underline{S}_n(t+s))\} \\ &= \lim_{n \rightarrow \infty} E \{h(\underline{S}_n(t+1+p), \dots, \underline{S}_n(t+s+p))\} \\ &= E \{h(\underline{X}_{t+1+p}, \dots, \underline{X}_{t+s+p})\} \end{aligned}$$

à partir du moment où $\underline{S}_n(t)$ est strictement stationnaire pour chaque n fixé. Il est alors facile de voir, en prenant les limites dans (2.12), que \underline{X}_t satisfait l'équation aux différences (2.9) et donc que la première composante $X_t, t \in \mathbb{Z}$ satisfait (2.1) avec $Q = 2$. Pour établir l'ergodicité de X_t , nous devons montrer qu'il existe une fonction $f : R^\infty \rightarrow R$, telle que, presque sûrement, $X_t = f(\varepsilon_t, \varepsilon_{t-1}, \dots)$ pour tout $t \in \mathbb{Z}$. Avec la suite de fonction f_t définies

par (2.15) on observe que

$$\begin{aligned}
E \|\underline{S}_n(t) - f_0(\varepsilon_t, \varepsilon_{t-1}, \dots)\|^2 &= E \|G_n(\varepsilon_t, \varepsilon_{t-1}, \dots) - f_0(\varepsilon_t, \varepsilon_{t-1}, \dots)\|^2 \\
&= E \|G_n(\varepsilon_0, \varepsilon_{-1}, \dots) - f_0(\varepsilon_0, \varepsilon_{-1}, \dots)\|^2 \\
&= E \|\underline{S}_n(0) - f_0(\varepsilon_0, \varepsilon_{-1}, \dots)\|^2 \\
&\rightarrow 0, \text{ quand } n \rightarrow \infty.
\end{aligned}$$

Ainsi, presque sûrement, pour tout $t \in \mathbb{Z}$, $f_t(\varepsilon_t, \varepsilon_{t-1}, \dots) = f_0(\varepsilon_t, \varepsilon_{t-1}, \dots)$. Maintenant, si on définit $f = f_0^{(1)}$, et $f_0 := (f_0^{(1)}, \dots, f_0^{(p)})'$, on peut écrire : $\underline{X}_t = f(\varepsilon_t, \varepsilon_{t-1}, \dots)$ pour presque tout $t \in \mathbb{Z}$, ce qui montre que \underline{X}_t est ergodique. La stationnarité de X_t provient de la stricte stationnarité de \underline{X}_t et du fait que, pour chaque t , \underline{X}_t est la limite dans \mathbb{L}^2 de $\underline{S}_n(t)$.

2. Soit \underline{Y}_t une autre solution (2.9) telle que

$$\underline{Y}_t = g(\varepsilon_t, \varepsilon_{t-1}, \dots) \quad (2.16)$$

où g est une fonction mesurable de \mathbb{R}^∞ dans \mathbb{R}^p . Posons $\underline{U}_t = \underline{X}_t - \underline{Y}_t$ où \underline{X}_t est la solution de (2.9) trouvée dans (1). Alors le processus $(\underline{U}_t)_{t \in \mathbb{Z}}$ est strictement stationnaire et satisfait les équations suivantes

$$\underline{U}_t = (A + B_1\varepsilon_{t-1} + B_2\varepsilon_{t-2})\underline{U}_{t-1}, t \in \mathbb{Z}$$

On pose $V = E(\underline{U}_t \underline{U}_t')$, $D = (\varepsilon_{t-1} \underline{U}_t \underline{U}_t')$, $F = E(\varepsilon_{t-1}^2 \underline{U}_t \underline{U}_t')$ et $\underline{u} = [\vec{V}, \vec{D}, \vec{F}]'$, on obtient alors, en utilisant les mêmes arguments que précédemment $\underline{u} = \Gamma \underline{u} = \dots = \Gamma^t \underline{u}$, $t \geq 0$. Comme $\lambda = \rho(\Gamma) < 1$, on en déduit que, presque sûrement, $\underline{u} = 0$ et donc que \underline{X}_t est la solution unique de (2.9) de la forme (2.16). Pour montrer que \underline{X}_t est l'unique solution stationnaire de (2.9), on définit le processus

$$\underline{S}_n^*(t) = \begin{cases} 0 & \text{si } n < 0 \\ \underline{\eta}_t^* = (\varepsilon_{t-2}, \dots, \varepsilon_{t-p+1})' & \text{si } n = 0 \\ \underline{\eta}_t + (A + B_1\varepsilon_{t-1} + B_2\varepsilon_{t-2})\underline{S}_{n-1}^*(t-1) & \text{si } n \geq 1 \text{ si } n > 0 \end{cases} \quad (2.17)$$

et on procède de la même façon qu'en (1) pour montrer que chaque t , $(\underline{S}_n^*(t), n \geq 0)$ converge à la fois, presque sûrement et dans L^2 , vers un vecteur aléatoire \underline{Y}_t de la forme (2.16) satisfaisant (2.9). A partir de l'unicité établit pour le processus $(\underline{X}_t)_{t \in \mathbb{Z}}$, on en déduit que, presque sûrement, $\underline{Y}_t = \underline{X}_t$ pour tout $t \in \mathbb{Z}$. Il est ensuite facile de voir, par simple itération, à partir

de (2.17) que

$$S_n(t) = \underline{\eta}_t + \sum_{r=1}^n \prod_{j=1}^r (A + B_1 \varepsilon_{t-j} + B_2 \varepsilon_{t-j-1}) \underline{\eta}_{t-r}$$

et donc que

$$\underline{X}_t = \underline{\eta}_t + \sum_{r=1}^{\infty} \prod_{j=1}^r (A + B_1 \varepsilon_{t-j} + B_2 \varepsilon_{t-j-1}) \underline{\eta}_{t-r}$$

où la somme peut être interprétée à la fois comme une limite presque sûre ou dans L^2 .

■

2.3.2 Le modèle $BL(p, q, p, Q)$

Théorème 14 Soit le processus $(\underline{X}_t)_{t \in \mathbf{Z}}$ défini par (2.9). Posons :

$$\Gamma_1 = A \otimes A + \sigma^2 (B_1 \otimes B_1)$$

et pour $j = 2, \dots, Q$

$$\Gamma_j = \sigma^2 (B_j \otimes (A^{j-1} B_1 + A^{j-2} B_2 + \dots + AB_{j-1}) + (A^{j-1} B_1 + A^{j-2} B_2 + \dots + AB_{j-1}) \otimes B_j + B_j \otimes B_j)$$

Si toutes les valeurs propres de la matrice

$$\Gamma = \begin{pmatrix} \Gamma_1 & \Gamma_2 & \dots & \Gamma_{Q-1} & \Gamma_Q \\ I_{(p^2)} & 0 & \dots & 0 & 0 \\ 0 & I_{(p^2)} & 0 & \dots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \dots & 0 & I_{(p^2)} & 0 \end{pmatrix}$$

ont un module plus petit que 1, alors il existe un processus strictement stationnaire $(\underline{X}_t)_{t \in \mathbf{Z}}$ solution de de l'équation (2.2). De plus, le processus solution est unique, ergodique est donné par (2.4).

Preuve. Pour montrer le théorème 14, on utilise les mêmes arguments que précédemment, mais les suite $(\underline{S}_n(t))_{t \in \mathbb{Z}}$ définies dans (2.12) sont remplacées par

$$\underline{S}_n(t) = \begin{cases} 0 & \text{si } n < 0 \\ \eta_t & \text{si } n = 0 \\ \eta_t + \Phi(t) \underline{S}_{n-1}(t-1) & \text{si } n \geq 1 \end{cases}$$

où $\Phi(t) := A + B_1 \varepsilon_{t-1} + \dots + B_Q \varepsilon_{t-Q}$ et le processus $(\underline{\Delta}_n(t))_{t \in \mathbb{Z}}$ devient

$$\underline{\Delta}_n(t) = \begin{cases} 0 & \text{si } n < 0 \\ \eta_t & \text{si } n = 0 \\ \Phi(t) \underline{\Delta}_{n-1}(t-1) & \text{si } n \geq 1 \end{cases}$$

Dans ce cas, les propriétés de la convergence de $(\underline{S}_n(t))_{t \in \mathbb{Z}}$ peuvent être étudiées en utilisant les mêmes arguments que dans la sous-section précédente. Donc, on introduit les matrices suivantes :

$$\begin{cases} \underline{V}_n = E \left\{ \underline{\Delta}_n(t) \underline{\Delta}_n'(t) \right\}, \\ \underline{D}_n^{(j)} = E \left\{ \varepsilon_{t-j}(t) \underline{\Delta}_n(t) \underline{\Delta}_n'(t) \right\}, j = 1, \dots, Q-1, \\ \underline{F}_n^{(jk)} = E \left\{ \varepsilon_{t-j} \varepsilon_{t-k} \underline{\Delta}_n(t) \underline{\Delta}_n'(t) \right\}, 1 \leq j \leq k < Q. \end{cases}$$

■

Dans le cas de certains modèles simples, on dispose de résultats plus forts que ceux que l'on peut obtenir du résultat général. Voici quelques exemples

Exemple 15 *Considérons le processus $(X_t)_{t \in \mathbb{Z}}$ bilinéaire superdiagonal $BL(1, 0, p, 1)$*

$$X_t = \varepsilon_t + \sum_{i=2}^p b_{i1} X_{t-i} \varepsilon_{t-1}, \quad t \in \mathbb{Z} \quad (2.18)$$

où $(\varepsilon_t)_{t \in \mathbb{Z}}$ est un bruit blanc gaussien de variance σ^2 . Soient les vecteurs $\underline{X}_t = (X_t, X_{t-1}, \dots, X_{t-p+1})'$, $H = (1, 0, \dots, 0)'$ de \mathbb{R}^p et les matrices carrées d'ordre p

$$A = \begin{pmatrix} 0 & 0 & \dots & 0 \\ 1 & 0 & \dots & 0 \\ 0 & \ddots & \dots & \vdots \\ \vdots & 0 & 1 & 0 \end{pmatrix}, \quad B = \begin{pmatrix} b_{21} & \dots & b_{p1} & 0 \\ 0 & 0 & \dots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \ddots & \ddots & 0 \end{pmatrix}.$$

Alors on a

$$\begin{cases} X_t = H' \underline{X}_t \\ \underline{X}_t = H \varepsilon_t + A \underline{X}_{t-1} + B \underline{X}_{t-2} \varepsilon_{t-1} \end{cases}$$

et la condition $\sigma^2 (b_{21}^2 + \dots + b_{p1}^2) < 1$ est nécessaire et suffisante pour l'existence et l'unicité d'un processus stationnaire au second ordre satisfaisant (2.18).

Exemple 16 [Pham et Tran] Soit le modèle bilinéaire diagonal $BL(1, 0, 0, 1)$

$$X_t = a_1 X_{t-1} + b_{11} X_{t-1} \varepsilon_{t-1} + \varepsilon_t, t \in \mathbb{Z}$$

où $(\varepsilon_t)_{t \in \mathbb{Z}}$ est un bruit blanc gaussien de variance σ^2 . Alors, une condition nécessaire et suffisante d'existence et d'unicité d'un processus strictement stationnaire de carré intégrable $(X_t)_{t \in \mathbb{Z}}$ vérifiant l'équation précédente est $a_1^2 + b_{11}^2 \sigma^2 < 1$.

2.4 Inversibilité des modèles bilinéaires

L'inversibilité est une notion très importante dans l'étude des modèles tant linéaire que non linéaires car elle permet d'une part, de faire des prévisions et d'autre part, d'estimer dans certains cas, les paramètres des modèles. Dans la présente section on s'intéresse à la classe de modèles

$$X_t = \sum_{i=1}^p a_i X_{t-i} + \left(c + \sum_{j=1}^p b_j X_{t-j} \right) \varepsilon_{t-1} + \varepsilon_t \quad (2.19)$$

où $(\varepsilon_t)_{t \in \mathbb{Z}}$ est un bruit blanc gaussien. Le modèle (2.19) peut s'écrire sous la forme vectorielle

$$\underline{X}_t = A \underline{X}_{t-1} + B_1 \underline{X}_{t-1} \varepsilon_{t-1} + c H \varepsilon_{t-1} + H \varepsilon_t$$

et $X_t = H' \underline{X}_t$. Notons que la généralisation des résultats aux modèles bilinéaires généraux, nécessite simplement une écriture matricielle.

Théorème 17 Soit $(\varepsilon_t)_{t \in \mathbb{Z}}$ un bruit blanc gaussien de variance σ^2 . Alors sous les conditions du corollaire (12) le modèle (2.19) est inversible si et seulement si

$$E \left\{ \log \left| c + \sum_{j=1}^P b_j X_{t-j} \right| \right\} < 0. \quad (2.20)$$

Preuve. D'après (2.19) nous avons pour tout $n > 1$

$$\begin{aligned} \varepsilon_t &= \left(X_t - \sum_{i=1}^p a_i X_{t-i} \right) + \sum_{k=1}^{n-1} (-1)^k \prod_{l=0}^{k-1} \left(c + \sum_{j=1}^P b_j X_{t-l-j} \right) \left(X_{t-k} - \sum_{i=1}^p a_i X_{t-k-i} \right) \\ &\quad + (-1)^n \prod_{l=0}^{n-1} \left(c + \sum_{j=1}^P b_j X_{t-l-j} \right) \varepsilon_{t-n} \end{aligned}$$

Pour montrer que $\mathfrak{S}_t^{(\varepsilon)} \subset \mathfrak{S}_t^{(X)}$, il suffit de montrer que, presque sûrement, $\lim_{n \rightarrow \infty} \prod_{l=1}^n \left(c + \sum_{j=1}^P b_j X_{t-l-j} \right) \varepsilon_{t-n} = 0$ et comme $(\varepsilon_t)_{t \in \mathbb{Z}}$ est un processus *i.i.d.*, il suffit alors de montrer que

$$\lim_{n \rightarrow \infty} \prod_{l=1}^n \left(c + \sum_{j=1}^P b_j X_{t-l-j} \right) = 0$$

Observons tous d'abord que

1. Sous les hypothèses du théorème, le processus X_t est ergodique
2. $\forall t \in \mathbb{Z}$, alors $\left(c + \sum_{j=1}^P b_j X_{t-j} \right), \left(c + \sum_{j=1}^P b_j X_{t-1-j} \right), \left(c + \sum_{j=1}^P b_j X_{t-3-j} \right), \dots$ est ergodique
3. $\forall t \in \mathbb{Z}$, alors $\log \left| \left(c + \sum_{j=1}^P b_j X_{t-j} \right) \right|, \log \left| c + \sum_{j=1}^P b_j X_{t-1-j} \right|, \log \left| c + \sum_{j=1}^P b_j X_{t-3-j} \right|, \dots$ est ergodique.
4. Posons $P_n(t) := \prod_{l=0}^{n-1} \left(c + \sum_{j=1}^P b_j X_{t-l-j} \right)$, alors

$$\frac{1}{n} \log |P_n(t)| := \frac{1}{n} \log \left| \prod_{l=0}^{n-1} \left(c + \sum_{j=1}^P b_j X_{t-l-j} \right) \right| = \frac{1}{n} \sum_{l=0}^{n-1} \log \left| c + \sum_{j=1}^P b_j X_{t-l-j} \right|$$

Alors le théorème ergodique montre que

$$\lim_{n \rightarrow \infty} \frac{1}{n} \log \left| \prod_{l=0}^{n-1} \left(c + \sum_{j=1}^P b_j X_{t-l-j} \right) \right| = E \log \left| c + \sum_{j=1}^P b_j X_{t-1-j} \right| < 0$$

D'où, presque sûrement, $\lim_{n \rightarrow \infty} \prod_{l=0}^{n-1} \left(c + \sum_{j=1}^P b_j X_{t-l-j} \right) = 0$.

■

Remarque 18 La présence de la partie autorégressive $\sum_{i=1}^p a_i X_{t-j}$ n'a aucun impact sur l'inversibilité du modèle. D'autre part la condition (2.20) n'est rien d'autre que l'exposant de Lyapunov associé à l'équation

$$\varepsilon_t = - \left(c + \sum_{j=1}^p b_j X_{t-j} \right) \varepsilon_{t-1} + \left(X_t - \sum_{i=1}^p a_i X_{t-j} \right).$$

Cependant, la généralisation de ce résultat aux processus bilinéaires généraux est immédiate.

Remarque 19 Une condition suffisante pour l'inversibilité de modèle bilinéaire (2.19) pour chaque $t \in \mathbb{Z}$ est

$$c^2 + 2c\mu \sum_{j=1}^P b_j + \sum_{k=1}^P \sum_{j=1}^P b_k b_j E \{ X_{t-k} X_{t-j} \} < 1 \quad (2.21)$$

où $\mu = E \{ X_t \}$.

Exemple 20 Soit le processus $(X_t)_{t \in \mathbb{Z}}$ généré par $X_t = \varepsilon_t + \sum_{i=2}^p b_i X_{t-i} \varepsilon_{t-1}$, $t \in \mathbb{Z}$. Alors la condition nécessaire et suffisamment, pour $(X_t)_{t \in \mathbb{Z}}$ soit stationnaire au second ordre est que $\sigma^2 \sum_{i=2}^p b_i^2 < 1$.

D'autre part pour ce modèle on a $\mu = 0$ et

$$E \{ X_{t-k} X_{t-j} \} = \begin{cases} \frac{\sigma^2}{1 - \sigma^2 \sum_{i=2}^p b_i^2}, & \text{si } k = j \\ 0 & \text{sinon} \end{cases}$$

Donc d'après (2.21) une condition suffisante d'inversibilité est $2\sigma^2 \sum_{i=2}^p b_i^2 < 1$.

2.5 Structure de covariance

La méthode des moments consiste à exploiter des équations "récurives" des moments ou des cumulants et remplacer les valeurs théoriques par leurs estimations. Les propriétés des seconds ordres pour certains modèles bilinéaires sont semblables à ceux d'un modèle linéaire. Plus précisément, la fonction covariance du processus $BL(p, q, p, q)$ est la même que celle qu'un processus $ARMA$ ayant la même partie autorégressive. Par conséquence, nous ne pouvons pas distinguer le modèle bilinéaire

d'un linéaire en analysant la fonction de covariance, il est alors nécessaire d'étudier quelques cumulants d'ordre supérieur. Le calcul des moments d'ordre supérieur est en principe possible. Néanmoins les calculs sont très fastidieux.

2.5.1 Expression des moments pour le processus $BL(p, 0, p, 1)$

Considérons le modèle suivant

$$X_t = \sum_{i=1}^p a_i X_{t-i} + \sum_{i=1}^p b_i X_{t-i} \varepsilon_{t-1} + \varepsilon_t \quad (2.22)$$

où $(\varepsilon_t)_{t \in \mathbb{Z}}$ est un bruit blanc gaussien de variance σ^2 . Supposons que

$$\rho(A \otimes A + \sigma^2 B_1 \otimes B_1) < 1. \quad (2.23)$$

Proposition 21 *Considérons le processus bilinéaire $BL(p, 0, p, 1)$ satisfaisant (2.23) et notons*

$$\begin{aligned} \mu & : = E\{X_t\} \\ c(h) & : = E\{(X_t - \mu)(X_{t-h} - \mu)\} \\ c(k, h) & : = E\{(X_t - \mu)(X_{t-k} - \mu)(X_{t-h} - \mu)\}. \end{aligned}$$

Alors nous avons :

1. $\mu = \frac{\sigma^2 b_1}{1 - \sum_{i=1}^p a_i}$
2. $c(h) = \begin{cases} \sum_{i=1}^p a_i c(h-i), & \text{si } h > 1 \\ \sum_{i=1}^p a_i c(h-i) + \sigma^2 \sum_{i=1}^p \sum_{j=1}^p b_i b_j c(i-j) + \sigma^2 \mu \sum_{i=1}^p b_i \left\{ a_1 + \mu \sum_{i=1}^p b_i \right\} + b_1^2 \sigma^4 + \sigma^2, & \text{si } h = 0 \\ \sum_{i=1}^p a_i c(h-i) + \mu \sigma^2 \sum_{i=1}^p b_i, & \text{si } h = 1 \end{cases}$
3. $c(1, h) = \sum_{i=1}^p a_i c(i-1, h-1) + \sigma^2 \sum_{i=1}^p b_i c(h-i), \text{ si } h \geq 2.$

Preuve. Sous la condition (2.23), le processus est stationnaire au second ordre et on a aussi $\rho(A \otimes A) < 1$ ceci implique que $\sum_{i=1}^p a_i < 1$. Comme $E\{X_{t-i} \varepsilon_t\} = 0$ pour $i > 0$, alors on a d'une part

$$\mu = \left(\sum_{i=1}^p a_i \right) \mu + E\{b_1 X_{t-1} \varepsilon_{t-1}\} \quad (2.24)$$

et comme $E\{X_t \varepsilon_t\} = \sigma^2$ on a le résultat. D'autre part, il est relativement simple de voir que $X_t - \mu$ satisfait l'équation suivante :

$$X_t - \mu = \sum_{i=1}^p a_i (X_{t-i} - \mu) + \sum_{i=1}^p b_i (X_{t-i} - \mu) \varepsilon_{t-1} + \varepsilon_t + \mu \left(\sum_{i=1}^p b_i \right) \varepsilon_{t-1} - b_1 \sigma^2$$

Donc pour tout $h \geq 0$

$$c(h) = E \left\{ \underbrace{\sum_{i=1}^p a_i (X_{t-i} - \mu) (X_{t-h} - \mu)}_1 + \underbrace{\sum_{i=1}^p b_i (X_{t-i} - \mu) (X_{t-h} - \mu) \varepsilon_{t-1}}_2 \right. \\ \left. + \underbrace{(X_{t-h} - \mu) \varepsilon_t}_3 + \underbrace{\mu \left(\sum_{i=1}^p b_i \right) \varepsilon_{t-1} (X_{t-h} - \mu)}_4 - \underbrace{b_1 \sigma^2 (X_{t-h} - \mu)}_5 \right\}$$

d'où

$$E\{1\} = \sum_{i=1}^p a_i c(h-i), E\{5\} = 0, E\{3\} = \begin{cases} \sigma^2, h=0 \\ 0, h>0 \end{cases}, E\{4\} = \begin{cases} \mu \left(\sum_{i=1}^p b_i \right) \sigma^2, h=1 \\ 0, h>1 \end{cases}$$

Pour évaluer l'expression $E\{(X_{t-l} - \mu)(X_{t-h} - \mu)\varepsilon_t\}$ remarquons que :

- Si $l > 0$ et $h > 0$, alors $E\{(X_{t-l} - \mu)(X_{t-h} - \mu)\varepsilon_t\} = 0$
- Si $l = 0$ et $h > 0$ on a $E\{(X_t - \mu)(X_{t-h} - \mu)\varepsilon_t\} = 0, h > 0$
- Si $l = h = 0$, alors

$$E\{(X_t - \mu)^2 \varepsilon_t\} = E \left\{ \left(\sum_{i=1}^p a_i (X_{t-i} - \mu) + \sum_{i=1}^p b_i (X_{t-i} - \mu) \varepsilon_{t-1} + \varepsilon_t + \mu \left(\sum_{i=1}^p b_i \right) \varepsilon_{t-1} - b_1 \sigma^2 \right)^2 \varepsilon_t \right\} \\ = 0$$

Donc, en conclusion $E\{(X_{t-l} - \mu)(X_{t-h} - \mu)\varepsilon_t\} = 0, \forall l, h \geq 0$ et le résultat s'achève.

Pour calculer $c(0)$, nous procédons comme suit :

$$c(0) = E \left\{ \begin{array}{l} \underbrace{\sum_{i=1}^p \sum_{j=1}^p a_i a_j (X_{t-i} - \mu) (X_{t-j} - \mu)}_1 + \underbrace{\sum_{i=1}^p \sum_{j=1}^p b_i b_j (X_{t-i} - \mu) (X_{t-j} - \mu) \varepsilon_{t-1}^2}_2 + \underbrace{\varepsilon_t^2}_3 \\ + \underbrace{\mu^2 \left(\sum_{i=1}^p b_i \right)^2 \varepsilon_{t-1}^2}_4 + \underbrace{b_1^2 \sigma^4}_5 + \underbrace{2 \sum_{i=1}^p \sum_{j=1}^p a_i b_j (X_{t-i} - \mu) (X_{t-j} - \mu) \varepsilon_{t-1}}_6 \\ + \underbrace{2 \sum_{i=1}^p a_i (X_{t-i} - \mu) \varepsilon_t}_7 + \underbrace{2 \sum_{i=1}^p a_i (X_{t-i} - \mu) \mu \left(\sum_{l=1}^p b_l \right) \varepsilon_{t-1}}_8 - \underbrace{2 \sum_{i=1}^p a_i (X_{t-i} - \mu) b_1 \sigma^2}_9 \\ + \underbrace{2 \sum_{i=1}^p b_i (X_{t-i} - \mu) \varepsilon_{t-1} \varepsilon_t}_{10} + \underbrace{2 \sum_{i=1}^p b_i (X_{t-i} - \mu) \varepsilon_{t-1}^2 \mu \left(\sum_{l=1}^p b_l \right)}_{11} \\ - \underbrace{2b_1 \sigma^2 \sum_{i=1}^p b_i (X_{t-i} - \mu) \varepsilon_{t-1}}_{12} + \underbrace{2\varepsilon_t \left(\sum_{i=1}^p b_i \right) \mu \varepsilon_{t-1}}_{13} - \underbrace{2b_1 \sigma^2 \varepsilon_t}_{14} - \underbrace{2b_1 \sigma^2 \mu \left(\sum_{i=1}^p b_i \right) \varepsilon_{t-1}}_{15} \end{array} \right\}$$

Une vérification directe montre que les espérances des quantités $\{6\}, \{7\}, \{9\}, \{10\}, \{13\}, \{14\}$ et $\{15\}$ sont nulles, $E\{8\} = 2\mu \left(\sum_{i=1}^p b_i \right) a_1 \sigma^2$, $E\{12\} = -2b_1^2 \sigma^4$. Aussi

$$E\{(X_{t-l} - \mu)(X_{t-h} - \mu)\varepsilon_t^2\} = \begin{cases} c(h-l)\sigma^2, & \text{si } h > 0 \text{ ou } l > 0 \\ c(0)\sigma^2 + 2\sigma^4 & \text{si } h = l = 0 \end{cases}$$

ce qui implique que $E\{2\} = \sigma^2 \sum_{i=1}^p \sum_{j=1}^p b_i b_j c(i-j) + 2b_1^2 \sigma^4$. Finalement, on peut montrer aussi que $E\{(X_t - \mu)\varepsilon_t^2\} = 0$, alors $E\{11\} = 0$ et le résultat s'achève pour la variance.

L'expression de $c(1, h)$ peut être trouvée d'une façon similaire

$$c(1, h) = E \left\{ \begin{array}{l} \underbrace{\sum_{i=1}^p a_i (X_{t-i} - \mu) (X_{t-1} - \mu) (X_{t-h} - \mu)}_1 \\ + \underbrace{\sum_{i=1}^p b_i (X_{t-i} - \mu) (X_{t-1} - \mu) (X_{t-h} - \mu) \varepsilon_{t-1}}_2 + \underbrace{(X_{t-1} - \mu) (X_{t-h} - \mu) \varepsilon_t}_3 \\ + \mu \left(\underbrace{\sum_{i=1}^p b_i}_{4} \right) (X_{t-1} - \mu) (X_{t-h} - \mu) \varepsilon_{t-1} - \underbrace{b_1 \sigma^2 (X_{t-1} - \mu) (X_{t-h} - \mu)}_5 \end{array} \right\}$$

Il vient que les espérances des quantités {3} et {4} sont nulles pour tout $h \geq 1$, et

$$E \{ (X_{t-i} - \mu) (X_{t-1} - \mu) (X_{t-h} - \mu) \varepsilon_{t-1} \} = \sigma^2 c(h-i), i, h > 1$$

alors

$$E \{ (X_{t-1} - \mu)^2 (X_{t-h} - \mu) \varepsilon_{t-1} \} = E \left\{ \begin{array}{l} \underbrace{2 \sum_{l=1}^p a_l (X_{t-l} - \mu) (X_{t-h} - \mu) \varepsilon_{t-1}^2}_1 \\ + 2 \underbrace{\sum_{l=1}^p b_l (X_{t-l-1} - \mu) (X_{t-h} - \mu) \varepsilon_{t-2} \varepsilon_{t-1}^2}_2 \\ + 2\mu \underbrace{\left(\sum_{l=1}^p b_l \right) \varepsilon_{t-2} \varepsilon_{t-1}^2 (X_{t-h} - \mu)}_3 \end{array} \right\}$$

on a

$$E \{2\} = 0, h \geq 2, \quad E \{3\} = \begin{cases} 0, h \geq 3 \\ 2\mu \left(\sum_{i=1}^p b_i \right) \sigma^4, h = 2 \end{cases}$$

et

$$E \{1\} = 2\sigma^2 \sum_{i=1}^p a_i c(h-l-1) = \begin{cases} 2\sigma^2 c(h-1), h > 2 \\ 2\sigma^2 \left(c(h-1) - \mu \left(\sum_{i=1}^p b_i \right) \sigma^2 \right), h = 2 \end{cases}$$

D'où le résultat. ■

2.5.2 Expression des moments pour les processus $BL(0, 0, 2, 1)$ super-diagonaux

Considérons le modèle

$$X_t = \varepsilon_t + bX_{t-2}\varepsilon_{t-1} \quad (2.25)$$

où $(\varepsilon_t)_{t \in \mathbb{Z}}$ est un bruit blanc gaussien de moyenne nulle et de variance σ^2 . Notons, comme précédemment, pour tout $h, k, l \geq 0$

$$\begin{aligned} \mu & : = E\{X_t\} \\ c(s) & : = E\{X_t X_{t-s}\} \\ c(k, l) & : = E\{X_t X_{t-k} X_{t-l}\} \\ c(h, k, l) & : = E\{X_t X_{t-h} X_{t-k} X_{t-l}\}. \end{aligned}$$

En développant X_t^m selon le binôme de Newton, alors on obtient les résultats suivants

$$\begin{aligned} \sigma^2 b^2 < 1 & \iff (X_t)_{t \in \mathbb{Z}} \text{ strictement stationnaire, ergodique et causal} \\ \sigma^4 b^4 < 1/3 & \iff E\{X_t^4\} < +\infty \text{ (Convergence forte) } (X_t)_{t \in \mathbb{Z}} \text{ est inversible} \\ \sigma^6 b^6 < 1/15 & \iff E\{X_t^6\} < +\infty \\ \sigma^8 b^8 < 1/105 & \iff E\{X_t^8\} < +\infty \end{aligned}$$

Le corollaire suivant est une conséquence immédiate de la proposition 21

Corollaire 22 *Soit le processus $(X_t)_{t \in \mathbb{Z}}$ défini par l'équation (2.25) où $(\varepsilon_t)_{t \in \mathbb{Z}}$ est un bruit blanc gaussien centrées, de variance $\sigma^2 < \infty$, alors $(X_t)_{t \in \mathbb{Z}}$ est un processus centré. Si $b^2 \sigma^2 < 1$, $X_t \in \mathbb{L}^2$ et si $b^4 \sigma^4 < \frac{1}{3}$, alors X_t a ses moments d'ordre 4 finis. Sous ces conditions on a les résultats suivants. En posant $u = b^2 \sigma^2$:*

1. $c(0) = \frac{\sigma^2}{1-u}$, $c(k) = 0 \quad \forall k > 0$
2. $c(1, 2) = b\sigma^2 c(0)$
3. $c(k, l) = 0$, $\forall k \neq 1, \forall l \neq 2$
4. $c(0, 0, 0) = \frac{3\sigma^4(1+u)}{(1-u)(1-3u^2)}$

5. $c(0, 1, 1) = \frac{\sigma^4(1+2u)}{(1-u)^2}$
6. $c(0, 2, 2) = \frac{\sigma^4(1+3u)}{(1-u)(1-3u^2)}$
7. $c(0, 2k-1, 2k-1) = \frac{\sigma^4(1+2u^k)}{(1-u)^2}$
8. $c(0, 2k, 2k) = \frac{\sigma^4}{(1-u)^2} + \frac{2\sigma^4 u^{2k}}{(1-u)^2(1-3u^2)}$
9. $c(1, 3, 3) = \frac{2u\sigma^4}{1-u}$
10. $c(0, 1, 4) = \frac{2u}{1-u}$
11. $c(0, 2k+1, 2k+4) = \frac{2\sigma^4 u^{k+2}}{1-u}$
12. $c(0, 2k+1, 2k+4) = \frac{2\sigma^4 u^{k+2}}{1-u}$
13. $c(1, 3, 4) = \frac{u\sigma^4}{1-u}$
14. $c(h, k, l) = 0$ pour toutes autres valeurs de h, k, l .

2.5.3 Expression des moments pour le processus $BL(1, 1, 1, 1)$

Nous considérons le modèle bilinéaire

$$X_t = aX_{t-1} + bX_{t-1}\varepsilon_{t-1} + c\varepsilon_{t-1} + \varepsilon_t$$

où $(\varepsilon_t)_{t \in \mathbb{Z}}$ est un processus de bruit blanc gaussien centré de variance σ^2 . Alors on a, sous les hypothèses de stationnarité au second ordre $\mu = \frac{b\sigma^2}{1-a}$. Les moments d'ordres deux de X_t sont

$$\begin{aligned} c(0) &= \sigma^2 \frac{(2b^2\sigma^2 + c^2 + 2ac + 1 + 2b(2a+c)\mu)}{1 - a^2 - b^2\sigma^2} \\ c(1) &= ac(0) + 2b\sigma^2\mu + c\sigma^2 \\ c(k) &= ac(k-1) + b\sigma^2\mu, \quad k \geq 2 \end{aligned}$$

L'expression de $c(k_1, k_2) = E\{X_t X_{t-k_1} X_{t-k_1-k_2}\}$ s'écrit

$$c(0, 0) = \frac{b_1 c(0) + c_1 \mu + d_1}{1 - a^3 - 3ab^2\sigma^2}$$

où $b_1 = 3b\sigma^2(3b^2\sigma^2 + 3a^2 + 2ac)$, $c_1 = 3\sigma^2(6b^2\sigma^2(a+c) + 2a^2c + 1)$ et $d_1 = 3\sigma^4b(2b^2\sigma^2 + 4ac + 3c^2)$.
D'autre part

$$c(0, 1) = a_2c(0, 0) + b_2c(0) + c_2\mu + d_2$$

où $a_2 = a^2 + b^2\sigma^2$, $b_2 = 2b\sigma^2(3a + c)$, $c_2 = \sigma^2(b^2\sigma^2 + c^2 + 4ac + 1)$ et $d_2 = 4bc\sigma^4$ et pour tout $k > 1$

$$c(0, k) = a_3c(0, k-1) + b_3c(k-1) + c_3\mu$$

avec $a_3 = a_2$, $b_3 = 2b\sigma^2(2a + c)$ et $c_3 = \sigma^2(2b^2\sigma^2 + c^2 + 2ac + 1)$. Pour finir

$$c(1, 0) = ac(0, 0) + 3b\sigma^2c(0) + 2c\sigma^2\mu$$

$$c(1, k) = ac(0, k) + 2b\sigma^2c(k-1) + c\sigma^2\mu$$

pour tout $k > 0$, et

$$c(k_1, k_2) = ac(k_1 - 1, k_2) + b\sigma^2c(k_2)$$

pour tout $k_1 > 1$ et $k_2 \geq 0$.

Chapitre 3

L'inférence statistique dans les modèles bilinéaires

3.1 Un aperçu sur la méthode des moments généralisés (*GMM*)

La méthode des moments généralisés (*GMM*) est une méthode d'estimation qui a plusieurs applications dans l'économétrie et dans les séries temporelles. Cette méthode qui peut être considérée comme une méthode à information limitée, occupe une place centrale chez les économètres. Ceci est dû aux deux motivations suivantes :

1. Un des nombreux avantages de *GMM* est que c'est une méthode englobante permettant de retrouver comme cas particuliers un grand nombre d'estimateurs usuels parmi lesquels

- Les moindres carrés ordinaires (OLS)
- Le maximum de vraisemblance (MLE)
- Les moindres carrés non linéaires
- Les *M*-estimateurs

2. *GMM* est une alternative fondée sur des hypothèses minimales.

L'objectif de la présente section, est l'estimation dans les modèles bilinéaires en utilisant la méthode *GMM*. Nous renvoyons le lecteur intéressé au chapitre 8 du livre de Hamilton [20] et les références citées.

3.1.1 Exemple introductif

On considère une variable aléatoire X distribuée selon une distribution de Student à ν degrés de liberté, de densité

$$f_X(x, \nu) = \frac{\Gamma\left(\frac{\nu+1}{2}\right)}{(\pi\nu)^{\frac{1}{2}} \Gamma\left(\frac{\nu}{2}\right)} \left[1 + \left(\frac{x^2}{\nu}\right)\right]^{-\left(\frac{\nu+1}{2}\right)}$$

où $\Gamma(\cdot)$ désigne la fonction Gamma¹. Supposons que l'on dispose d'un échantillon de n réalisations $\{x_1, \dots, x_n\}$ et que l'on désire à partir de cet échantillon estimer le nombre de degrés de liberté ν . La première approche que l'on pourrait qualifier d'approche à information complète consiste en une estimation par maximum de vraisemblance (*MLE*). La log-vraisemblance de l'échantillon s'écrit dans ce cas $L(\nu) = \sum_{t=1}^n \log f_{X_t}(x_t, \nu)$ et l'estimateur $\hat{\nu}_n$ est alors défini par : $\hat{\nu}_n = \underset{\nu > 0}{\text{Arg max}} L(\nu)$.

Une méthode alternative consiste au lieu d'exploiter l'information complète de la fonction de densité $f_{X_t}(x_t, \nu)$, équivalente à la fonction génératrice des moments, de n'exploiter qu'un nombre restreint de moments. On sait en effet que la connaissance de la densité $f_{X_t}(x_t, \nu)$ est équivalente à la connaissance de la fonction génératrice $G(h) = E\{X^h\} = \int_{R_+} x^h f_X(x, \nu) dx, h > 0$. Mais plutôt que d'utiliser l'ensemble des moments $\{G(1), \dots, G(h)\}$ pour estimer ν on peut se contenter d'utiliser uniquement un sous-ensemble de moments. Supposons que $\nu > 2$, alors on sait en particulier que les deux premiers moments (non centrés) sont tels que : $\mu_1 = E(X) = 0, \mu_2 = E(X^2) = \frac{\nu}{\nu-2}$. Dans ce cas précis, si l'on connaît la valeur de $E(X_t^2)$ alors on a $\nu = \frac{2E(X_t^2)}{E(X_t^2) - 1}$. Soit $\hat{\mu}_{2,n}$ le moment

empirique (non centré) d'ordre deux $\hat{\mu}_{2,n} = \frac{1}{n} \sum_{t=1}^n x_t^2$. Cet estimateur converge en probabilité vers μ_2 .

On peut en déduire que si n est très grand alors : $\frac{\nu}{\nu-2} \simeq \hat{\mu}_{2,n}$ et que par conséquent on peut déduire

l'estimateur $\hat{\nu}_n = \frac{2\hat{\mu}_{2,n}}{\hat{\mu}_{2,n} - 1}$ de ν qui converge vers ν . Cet estimateur existe dès lors que $\hat{\mu}_{2,n} > 1$.

Cet estimateur $\hat{\nu}_n$ est qualifié d'estimateur de la méthode des moments classique. De façon plus générale, si l'on considère un vecteur de paramètre $\theta \in \mathbb{R}^k$ caractérisant la densité $f_X(x, \theta)$ d'une variable X et si l'on suppose que k moments distincts $E(X^i) = \mu_i(\theta), i = i_1, i_2, \dots, i_k$ dépendant de θ , alors l'estimateur $\hat{\theta}_n$ de la méthode des moments classique est obtenu par la résolution d'un système à k équations et k inconnues : $\mu_i(\hat{\theta}_n) = \hat{\mu}_{i,n}$ où $\hat{\mu}_{i,n}$ désigne l'estimateur du moment empirique $\hat{\mu}_{i,n} = \frac{1}{n} \sum_{t=1}^n x_t^{i_i}, i = i_1, i_2, \dots, i_k$.

¹La fonction Gamma est définie pour tout $r > 0$ par $\Gamma(r) = \int_0^{\infty} x^{r-1} \exp\{-x\} dx$, si $r \in \mathbb{N}^*$, $\Gamma(r) := (r-1)!$

Dans cet exemple, on estime le paramètre v en utilisant un seul moment empirique, on aurait pu utiliser à la place de ce moment n'importe quel autre moment de X dépendant lui aussi du paramètre v . Par exemple, si $v > 4$, on sait que $\mu_4(v) = E(X_t^4) = \frac{3v^2}{(v-2)(v-4)}$. Cependant un

estimateur de v peut être déduit en résolvant l'équation $\hat{\mu}_{4,n} = \frac{3\hat{v}_n^2}{(\hat{v}_n-2)(\hat{v}_n-4)}$ où $\hat{\mu}_{4,n} = \frac{1}{n} \sum_{t=1}^n x_t^4$.

Une autre possibilité consiste à déterminer un estimateur \hat{v}_n qui permettent d'obtenir des valeurs des moments d'ordre deux et quatre aussi près que possible des moments empiriques $\hat{\mu}_{2,n}$ et $\hat{\mu}_{4,n}$. Autrement dit on cherche une valeur unique de v permettant de résoudre le système suivant :

$$\begin{cases} \hat{\mu}_{2,n} - \frac{v}{v-2} = 0 \\ \hat{\mu}_{4,n} - \frac{3v^2}{(v-2)(v-4)} = 0 \end{cases}$$

C'est pourquoi on cherche à déterminer l'estimateur \hat{v}_n qui minimise une **fonction critère** de la forme

$$Q_n(\hat{v}_n, x_1, \dots, x_n) = h_n' \underset{(1,1)}{W_n} h_n \underset{(1,2)(2,2)(2,1)}{h_n}$$

où la suite des vecteurs h_n sont est définies par

$$h_n = \begin{pmatrix} \hat{\mu}_{2,n} - \frac{v}{v-2} \\ \hat{\mu}_{4,n} - \frac{3v^2}{(v-2)(v-4)} \end{pmatrix}$$

et où les matrice W_n sont des matrices poids symétriques définies positives qui reflètent l'importance attribuée à chacun des deux moments que l'on désire reproduire. Un tel estimateur est appelé "**Estimateur des Moments Généralisés**"

3.2 Description de la méthode des moments généralisés

Soit $(\underline{X}_t)_{1 \leq t \leq N}$ une réalisation de longueur N d'un processus strictement stationnaire et ergodique de \mathbb{R}^m , qui est paramétrisé par un vecteur de paramètres $\theta \in \Theta \subset \mathbb{R}^p$, $h(\theta, \underline{X}_t)$ une fonction de $\mathbb{R}^p \times \mathbb{R}^m$ à valeurs dans \mathbb{R}^r , c'est un vecteur de variables aléatoires de \mathbb{R}^r . Soit θ_0 la vraie valeur de θ . On appelle conditions d'orthogonalité les r conditions définies par le système d'équations

$$E_{\theta_0} \{h(\theta, \underline{X}_t)\} = 0$$

A la variable aléatoire $h(\theta, \underline{X}_t)$, on lui associe les moments empiriques correspondants tel que

$$h_N(\theta) := \frac{1}{N} \sum_{t=1}^N h(\theta, \underline{X}_t)$$

donc $h_N(\theta)$ est une fonction définie de \mathbb{R}^p à valeurs dans \mathbb{R}^r . L'idée de base des *GMM* consiste à déterminer la valeur de θ telle que les r moments empiriques $h_N(\theta)$ soient aussi proches que possible de *ZERO*. Ainsi, on peut définir l'estimateur *GMM* de la façon suivante

Définition 23 L'estimateur *GMM* $\hat{\theta}_N$ consiste à minimiser la forme quadratique :

$$\begin{cases} \hat{\theta}_N := \underset{\theta \in \Theta}{\text{Arg min}} Q_N(\theta) \\ Q_N(\theta) := h'_N(\theta) W_N h_N(\theta) \end{cases} \quad (4.1)$$

où W_N désigne une suite de matrices $r \times r$ de poids symétriques définies positives qui peuvent être fonction de \underline{X}_t et de θ .

Notons que dans la plus part des cas, le programme de minimisation (4.1) ne peut être achevé que numériquement. L'intuition est très simple, on sait que $\forall \theta$, d'après le théorème ergodique $h_N(\theta)$ converge presque sûrement vers $E_{\theta_0} \{h(\theta, \underline{X}_t)\}$. Supposons que $E_{\theta_0} \{h(\theta, \underline{X}_t)\}$ soit continue en θ et que θ_0 soit la seule telle que $E_{\theta_0} \{h(\theta_0, \underline{X}_t)\} = 0$, alors sous les conditions de stationnarité, continuité et des conditions sur les moments, l'estimateur $\hat{\theta}_N$ est un estimateur convergent.

Exemple 24 Considérons le modèle bilinéaire diagonal $BL(1, 0, 1, 1)$

$$X_t = (b + be_{t-1}) X_{t-1} + e_t \quad (4.2)$$

où $(e_t)_{t \in \mathbb{Z}}$ est un bruit blanc gaussien de variance 1. Sous la condition de stationnarité ($2b^2 < 1$) on a $E\{X_t\} = \frac{b}{1-b}$. Soit $h(b, X_t) = X_t - \frac{b}{1-b}$, telle que $E_{b_0} \{h(b, X_t)\} = 0$. A cette condition d'orthogonalité en lui associée la fonction $h_N(b) := \bar{X}_N - \frac{b}{1-b}$. En posant $W_N = 1$, on retrouve le programme de minimisation

$$Q_N(b) := \left(\bar{X}_N - \frac{b}{1-b} \right)^2$$

La plus petite valeur admissible de $Q_N(b)$ est 0 obtenue pour $\bar{X}_N - \frac{\hat{b}}{1-\hat{b}}$ soit $\hat{b}_N = \frac{\bar{X}_N}{1 - \bar{X}_N}$. On retrouve donc l'estimateur de la méthode classique des moments.

Généralement, on distingue deux cas selon les valeurs de p et de r .

1. Lorsque il existe autant de conditions d'orthogonalité que de paramètres ($p = r$), on dit que le système (4.1) est juste identifié et l'estimateur *GMM* se ramène au vecteur $\hat{\theta}_N$ de dimension r qui permet de résoudre les r équations $h_N(\hat{\theta}_N) = 0$. Dans ce cas, il s'agit juste de résoudre un système d'équations éventuellement non linéaire à r équations et r inconnues. Le choix de la matrice W_N est totalement arbitraire, ce qui explique qu'elle ne soit pas spécifiée.
2. Lorsque il existe plus de conditions d'orthogonalité que de paramètres ($r > p$) on dit que le système est sur-identifié. L'estimateur *GMM* dépend alors du choix de la matrice poids W_N . Dans ce cas se pose un problème crucial du choix de la matrice de poids optimale.

3.2.1 Matrice de poids optimal

Comment choisit on la matrice de poids W_N afin d'obtenir un estimateur convergent efficace du vecteur θ . Supposons que le processus $(\underline{X}_t)_{t \in \mathbb{Z}}$ est stationnaire au second ordre, posons $\Omega(\theta) := E \{h(\theta, \underline{X}_t) h'(\theta, \underline{X}_t)\}$, alors la valeur optimale de la matrice poids W_N dans la fonction $Q_N(\theta)$ est donnée par $W_N = \Omega_N^{-1}(\theta)$ où $\Omega_N(\theta) := \frac{1}{N} \sum_{t=1}^N h(\tilde{\theta}_N, \underline{X}_t) h'(\tilde{\theta}_N, \underline{X}_t)$ où $\tilde{\theta}_N$ désigne une estimation préliminaire du vecteur θ . Cependant l'estimateur *GMM* correspondant est appelé le meilleur estimateur. Notons que la "meilleur" façon d'obtenir un estimateur préliminaire consistant $\tilde{\theta}_N$ de θ est de minimiser le programme (4.1) avec $W_N = I_r$.

Soit $G(\theta)$ la matrice Jacobienne i.e.,

$$G(\theta) := -E \left\{ \frac{\partial h(\theta, \underline{X}_t)}{\partial \theta} \right\} = -E \left(\begin{array}{c} \frac{\partial h(\theta, \underline{X}_t)}{\partial \theta_1} \\ \vdots \\ \frac{\partial h(\theta, \underline{X}_t)}{\partial \theta_p} \end{array} \right)_{p \times r}$$

et $G_N(\theta) := -\frac{1}{N} \sum_{t=1}^N \frac{\partial h(\theta, \underline{X}_t)}{\partial \theta}$, alors $G_N(\tilde{\theta}_N)$ converge en probabilité vers $G(\theta_0)$. Désormais Ω_N et G_N désigne $\Omega_N(\tilde{\theta}_N)$ et $G_N(\tilde{\theta}_N)$ respectivement et Ω et G désigne $\Omega(\theta_0)$ et $G(\theta_0)$. Sous des hypothèses de régularité que nous donnons ultérieurement, on montre que l'estimateur *GMM* pour lequel W_N converge en probabilité vers une matrice W symétrique définie positive, est asymptotiquement normale de fonction de covariance asymptotique $(G'WG)^{-1} G'W\Omega WG(G'WG)^{-1}$, et que l'estimateur *GMM* pour lequel W_N converge en probabilité vers $\Omega^{-1}(\theta)$ est asymptotiquement normale de fonction de covariance asymptotique $(G'WG)^{-1}$. Cette propriété suggère que le dernier estimateur est "meilleur" ceci est justifié par le lemme suivant

Lemme 25 La matrice $(G'WG)^{-1} G'W\Omega WG (G'WG)^{-1} - (G'WG)^{-1}$ est semi définie positive.

Preuve. Considérons la matrice $M := I - \Omega^{-1/2}G (G'\Omega^{-1}G)^{-1} G'\Omega^{-1/2}$, alors on a $M^2 = M$, donc la matrice M est idempotente et par conséquent, elle est semi définie positive, il s'en suit que la matrice

$$\left(\Omega^{1/2}WG (G'WG)^{-1}\right)' M\Omega^{1/2}WG (G'WG)^{-1}$$

est aussi demi définie positive. Or on vérifie facilement que

$$\left(\Omega^{1/2}WG (G'WG)^{-1}\right)' M\Omega^{1/2}WG (G'WG)^{-1} = (G'WG)^{-1} G'W\Omega WG (G'WG)^{-1} - (G'WG)^{-1}$$

d'où le résultat. ■

3.3 Comportement asymptotique des estimateurs

Dans cette section, nous étudions les propriétés asymptotiques ; c-à-d la consistance et la normalité asymptotique des estimateurs **GMM** pour le processus $BL(p, 0, p, 1)$, i.e.,

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

où $(\varepsilon_t)_{t \in \mathbb{Z}}$ est un bruit blanc gaussien de variance σ^2 considérée comme un paramètre de nuisance. Notre approche est basée sur les relations 2 et 3 de la proposition (21). A cet effet, nous définissons les vecteurs : $\underline{C}_t \in \mathbb{R}^r$ ($r \geq 2p + 1$) dont la $h^{\text{ième}}$ composante est $(X_t - \mu)(X_{t-1} - \mu) \dots (X_{t-h} - \mu)$, $\underline{a}' := (a_1, \dots, a_p)$, $\underline{b}' := (b_1, \dots, b_p)$ et les matrices

$$\begin{aligned} (A_t)_{h,j} &: = (X_t - \mu)(X_{t+1-j} - \mu) \dots (X_{t+1-h} - \mu), \quad h = 2, \dots, r, j = 1, \dots, p \\ (B_t)_{h,j} &: = (X_t - \mu)(X_{t+j-h} - \mu), \quad h = 2, \dots, r, j = 1, \dots, p \end{aligned}$$

Avec ces notations les relations 3 de la proposition (21) satisfaisant les équations

$$\begin{cases} E\{h(\theta, \underline{X}_t)\} = 0 \\ h(\theta, \underline{X}_t) := \underline{C}_t - D_t' \theta \\ D_t := \begin{pmatrix} A_t' \\ B_t' \end{pmatrix}, \quad \theta := \begin{pmatrix} \underline{a} \\ \underline{b} \end{pmatrix} \end{cases} \quad (4.3)$$

ce qui permis de définir r conditions d'orthogonalités, soit $h_N(\theta) := \frac{1}{N} \sum_{t=1}^N h(\theta, \underline{X}_t)$.

3.3.1 Consistance de l'estimateur à GMM

Pour obtenir la consistance de l'estimateur GMM nous avons besoin du théorème ergodique, pour s'appliquer nous faisons donc la prétention suivante :

$$h_N(\theta) := \left(\frac{1}{n} \sum_{t=1}^n h(\theta, \underline{X}_t) \right) \xrightarrow[n \rightarrow \infty]{P} h(\theta) := E \{h(\theta, \underline{X}_t)\}$$

Cependant, si W_N converge en probabilité vers W , alors $Q_N(\theta) := h'_N(\theta) W_N h_N(\theta)$ converge en probabilité vers $Q(\theta) := h'(\theta) W h(\theta)$. Malheureusement, ceci n'implique pas $\text{Arg min}_{\theta \in \Theta} Q_N(\theta)$ converge en probabilité vers $\text{Arg min}_{\theta \in \Theta} Q(\theta)$ car la fonction Arg min n'est pas un corollaire du théorème des fonctions continues. Une condition suffisante pour que le minimum de la limite $Q(\theta)$ soit la limite du minimum $\hat{\theta}_N$ est que la convergence en probabilité de $Q_N(\theta)$ vers $Q(\theta)$ soit uniforme sur un compact Θ au sens que $\sup_{\theta \in \Theta} |Q_N(\theta) - Q(\theta)|$ converge en probabilité vers 0 (cf. [32]). Donc, il faut imposer des conditions sur $Q_N(\theta)$ assurant l'uniformité de la convergence en probabilité. Considérons les hypothèses suivantes

- A1. Le processus $(\underline{X}_t)_{t \in \mathbb{Z}}$ est strictement stationnaire, ergodique et admettant des moments d'ordre quatre finis
- A2. L'espace des paramètres Θ est compact
- A3. Pour tout $\theta \in \Theta$, $h(\theta, \cdot)$ est une fonction mesurable
- A4. Pour tout $t \in \mathbb{Z}$, $h(\cdot, \underline{X}_t)$ est continue sur Θ
- A5. $|h_i(\theta, \underline{X}_t)| \leq b(\underline{X}_t)$ pour tout $\theta \in \Theta$ et $i = 1, \dots, r$, où $b(\cdot)$ est une fonction non négative telle que, $E_{\theta_0} \{b(\underline{X}_t)\} < \infty$
- A6. $W_N \xrightarrow{P} W$, où W est une matrice $r \times r$ définie positive
- A7. θ_0 est une solution unique de $E_{\theta_0} \{h(\theta, \underline{X}_t)\} = 0$ si et seulement si $\theta = \theta_0$.

Théorème 26 (Consistance de GMM) *Sous les hypothèses A1 – A7, l'estimateur GMM $\hat{\theta}_N$ défini par (4.1) est faiblement consistant, i.e., $\hat{\theta}_N$ converge en probabilité vers θ_0 .*

Preuve. Il suffit de montrer que $\sup_{\theta \in \Theta} |Q_N(\theta) - Q(\theta)|$ converge vers 0 en probabilité. On a

$$\begin{aligned} Q_N(\theta) - Q(\theta) &= h'_N(\theta) W_N h_N(\theta) - h'(\theta) W h(\theta) \\ &= (h_N(\theta) - h(\theta))' W_N (h_N(\theta) - h(\theta)) h'(\theta) (W_N + W'_N) (h_N(\theta) - h(\theta)) + \\ &\quad h'(\theta) (W_N - W) h(\theta) \end{aligned}$$

D'après l'inégalité triangulaire on a

$$\begin{aligned} |Q_N(\theta) - Q(\theta)| &\leq \left| (h_N(\theta) - h(\theta))' W_N (h_N(\theta) - h(\theta)) \right| + \\ &\quad \left| h'(\theta) (W_N + W'_N) (h_N(\theta) - h(\theta)) \right| + \left| h'(\theta) (W_N - W) h(\theta) \right| \\ &\leq \|h_N(\theta) - h(\theta)\|^2 \|W_N\| + 2 \|h(\theta)\| \|h_N(\theta) - h(\theta)\| \|W_N\| + \\ &\quad \|h(\theta)\|^2 \|W_N - W\| \end{aligned}$$

Donc

$$\begin{aligned} \sup_{\theta \in \Theta} |Q_N(\theta) - Q(\theta)| &\leq \sup_{\theta \in \Theta} \|h_N(\theta) - h(\theta)\|^2 \|W_N\| + 2 \sup_{\theta \in \Theta} \|h(\theta)\| \|h_N(\theta) - h(\theta)\| \|W_N\| + \\ &\quad \sup_{\theta \in \Theta} \|h(\theta)\|^2 \|W_N - W\| \end{aligned}$$

Nous allons montrer que les quantités

$$\sup_{\theta \in \Theta} \|h_N(\theta) - h(\theta)\|^2 \|W_N\|, \sup_{\theta \in \Theta} \|h(\theta)\| \|h_N(\theta) - h(\theta)\| \|W_N\|, \sup_{\theta \in \Theta} \|h(\theta)\|^2 \|W_N - W\|$$

convergent en probabilité vers 0 lorsque $N \rightarrow \infty$. Premièrement, Comme $\|W_N\|$ converge en probabilité vers $\|W\|$, et d'après la condition A7. et le fait que $h(\theta, \underline{X}_t)$ est dominé uniformément par une fonction intégrable, alors, $p \lim_{N \rightarrow \infty} \sup_{\theta \in \Theta} \|h_N(\theta) - h(\theta)\| = 0$. Comme $h(\theta)$ est une fonction continue sur un compact Θ , alors on obtient $\sup_{\theta \in \Theta} \|h(\theta)\|$ est borné, donc $p \lim_{N \rightarrow \infty} \sup_{\theta \in \Theta} |Q_N(\theta) - Q(\theta)| = 0$ uniformément sur Θ . De la condition A7., $Q(\theta_0) = \inf_{\theta \in \Theta} Q(\theta)$ et $Q_N(\hat{\theta}_N) = \inf_{\theta \in \Theta} Q_N(\theta)$, donc

$$\begin{aligned} 0 &\leq Q(\hat{\theta}_N) - Q(\theta_0) = \left(Q(\hat{\theta}_N) - Q_N(\hat{\theta}_N) \right) + \left(Q_N(\hat{\theta}_N) - Q(\theta_0) \right) \\ &\leq \left(Q(\hat{\theta}_N) - Q_N(\hat{\theta}_N) \right) + (Q_N(\theta_0) - Q(\theta_0)) \\ &\leq 2 \sup_{\theta \in \Theta} |Q_N(\theta) - Q(\theta)|. \end{aligned}$$

Comme $\sup_{\theta \in \Theta} |Q_N(\theta) - Q(\theta)|$ converge vers 0 en probabilité, alors que $(Q(\hat{\theta}_N) - Q(\theta_0))$ converge vers 0 en probabilité. D'autre part, comme la fonction Q est continue sur le compact Θ et que $Q(\theta_0) = 0$ et $Q(\theta) = 0 \iff E_{\theta_0} \{h(\theta, \underline{X}_t)\} = 0 \iff \theta = \theta_0$, c'est-à-dire, pour chaque $\epsilon > 0$, il existe $\delta > 0$ tel que $\|\hat{\theta}_N - \theta_0\| > \delta$ implique $|Q(\hat{\theta}_N) - Q(\theta_0)| > \epsilon$, donc

$$P\left(\|\hat{\theta}_N - \theta_0\| > \delta\right) \leq P\left(|Q(\hat{\theta}_N) - Q(\theta_0)| > \epsilon\right)$$

de façon que

$$\lim_{N \rightarrow \infty} P\left(|Q(\hat{\theta}_N) - Q(\theta_0)| > \epsilon\right) = 0$$

et par conséquent $\lim_{N \rightarrow \infty} P\left(\|\hat{\theta}_N - \theta_0\| > \delta\right) = 0$. D'où $p \lim_{N \rightarrow \infty} \hat{\theta}_N = \theta_0$. ■

3.3.2 Normalité asymptotique de l'estimateur GMM

Pour montrer la normalité asymptotique de l'estimateur GMM, des hypothèses supplémentaires sont nécessaires

A8. La vraie valeur $\theta_0 \in \overset{\circ}{\Theta}$ (l'intérieur de Θ)

A9. $h(\theta, \underline{X}_t)$ est deux fois continûment différentiable sur un voisinage \mathcal{V} de θ_0

A10 $\left\| \frac{\partial h(\theta, \underline{X}_t)}{\partial \theta} \right\| \leq f(\underline{X}_t)$ pour tout $\theta \in \mathcal{V}$ et $E_{\theta_0} \{f(\underline{X}_t)\} < \infty$

A11 La matrice $G'WG$ est non singulier où $G = -E_{\theta_0} \left\{ \frac{\partial h(\theta_0, \underline{X}_t)}{\partial \theta} \right\}$

A12 $\sqrt{N}h_N(\theta_0) := \frac{1}{\sqrt{N}} \sum_{t=1}^N h(\theta_0, \underline{X}_t) \xrightarrow{\mathcal{L}} (0, \Omega)$ où $\Omega := E \left\{ h(\theta_0, \underline{X}_t) h(\theta_0, \underline{X}_t)' \right\}$.

Théorème 27 (Normalité asymptotique de GMM) Sous les hypothèses A1 – A12, on a

$$\sqrt{N}(\hat{\theta}_N - \theta_0) \rightsquigarrow N\left(0, V^{as}(\hat{\theta}_N)\right)$$

où

$$\begin{aligned}
V^{as}(\hat{\theta}_N) &= (G'WG)^{-1} G'W\Omega W'G (G'W'G)^{-1} \\
\Omega &= E \left\{ h(\theta_0, \underline{X}_t) h(\theta_0, \underline{X}_t)' \right\} \\
\Omega_N &= \frac{1}{N} \sum_{t=1}^N h(\hat{\theta}_N, \underline{X}_t) h'(\hat{\theta}_N, \underline{X}_t) \xrightarrow{P} \Omega \\
G_N &= -\frac{1}{N} \sum_{t=1}^N \frac{\partial h(\hat{\theta}_N, \underline{X}_t)}{\partial \theta} \xrightarrow{P} G = -E \left\{ \frac{\partial h(\theta_0, \underline{X}_t)}{\partial \theta} \right\} \\
V_N^{as}(\hat{\theta}_N) &= (G'_N W_N G_N)^{-1} G'_N W_N \Omega_N W_N G_N (G'_N W_N G_N)^{-1}.
\end{aligned}$$

Preuve. Le critère à minimiser est

$$Q_N(\theta) = h'_N(\theta) W_N h_N(\theta) = m(h_N(\theta))$$

avec $m(h) = h' W_N h$. Donc

$$\left(\frac{\partial Q_N(\theta)}{\partial \theta'} \right)_{(1,p)} = \left(\frac{\partial m(h_N(\theta))}{\partial h'_N(\theta)} \right)_{(1,r)} \left(\frac{\partial h_N(\theta)}{\partial \theta'} \right)_{(r,p)}$$

où

$$\begin{aligned}
\frac{\partial h_N(\theta)}{\partial \theta'} &= \begin{pmatrix} \frac{\partial h_{N,1}(\theta)}{\partial \theta_1} & \dots & \frac{\partial h_{N,1}(\theta)}{\partial \theta_p} \\ \vdots & \vdots & \vdots \\ \frac{\partial h_{N,r}(\theta)}{\partial \theta_1} & \dots & \frac{\partial h_{N,r}(\theta)}{\partial \theta_p} \end{pmatrix}_{(r,p)} \\
\left(\frac{\partial m(h_N(\theta))}{\partial h'_N(\theta)} \right)_{(1,r)} &= 2h'_N(\theta) W_N
\end{aligned}$$

Donc la condition de minimisation du premier ordre donne

$$\left(\frac{\partial Q_N(\hat{\theta}_N)}{\partial \theta'} \right) = 0 \iff 2 \frac{\partial h'_N(\hat{\theta}_N)}{\partial \theta} W_N h_N(\hat{\theta}_N) = 0$$

Un développement limité $h_N(\hat{\theta}_N)$ au voisinage de θ_0 donne :

$$\sqrt{N}h_N(\hat{\theta}_N) = \sqrt{N}h_N(\theta_0) + \frac{\partial h_N(\theta_N^*)}{\partial \theta'} \sqrt{N}(\hat{\theta}_N - \theta_0) \quad (4.4)$$

où $\theta_N^* = \lambda\theta_0 + (1-\lambda)\hat{\theta}_N$, θ_N^* est un point intermédiaire entre $\hat{\theta}_N$ et θ_0 qui converge aussi en probabilité vers θ_0 lorsque $N \rightarrow \infty$. En multipliant (4.4) par $\frac{\partial h'_N(\hat{\theta}_N)}{\partial \theta} W_N$ on obtient

$$\frac{\partial h'_N(\hat{\theta}_N)}{\partial \theta} W_N \sqrt{N}h_N(\theta_0) + \frac{\partial h'_N(\hat{\theta}_N)}{\partial \theta} W_N \frac{\partial h_N(\theta_N^*)}{\partial \theta'} \sqrt{N}(\hat{\theta}_N - \theta_0) = 0$$

Alors d'après les conditions A10 et A11, $\frac{\partial h_N(\hat{\theta}_N)}{\partial \theta}$ converge uniformément en probabilité vers $E\left\{\frac{\partial h(\theta_0, \underline{X}_t)}{\partial \theta}\right\} = G$, d'où

$$\begin{aligned} \frac{\partial h_N(\hat{\theta}_N)'}{\partial \theta} W_N &\xrightarrow{P} G'W \\ \frac{\partial h_N(\hat{\theta}_N)'}{\partial \theta} W_N \frac{\partial h_N(\theta_N^*)}{\partial \theta'} &\xrightarrow{P} G'WG \end{aligned}$$

A'après hypothèse A12 et le théorème de Slutsky on déduit

$$\sqrt{N}(\hat{\theta}_N - \theta_0) \overset{\mathcal{L}}{\rightsquigarrow} N(0, V^{as}(\hat{\theta}_N))$$

où

$$V^{as}(\hat{\theta}_N) = (G'WG)^{-1} G'W\Omega WG (G'WG)^{-1}$$

■

Remarque 28 *Il est clair que les conditions d'orthogonalités (4.3) sont similaires aux celles d'un modèle de régression linéaires. Cependant, si $r = 2p$ la solution du système (4.1) est donnée par*

$$\begin{aligned}\hat{\theta}_N &= \left(\frac{1}{N} \sum_{t=1}^N D_t D_t' \right)^{-1} \frac{1}{N} \sum_{t=1}^N D_t C_t \\ &= \left(\begin{array}{cc} \frac{1}{N} \sum_{t=1}^N A_t' A_t & \frac{1}{N} \sum_{t=1}^N A_t' B_t \\ \frac{1}{N} \sum_{t=1}^N B_t' A_t & \frac{1}{N} \sum_{t=1}^N B_t' B_t \end{array} \right)^{-1} \frac{1}{N} \sum_{t=1}^N \begin{pmatrix} A_t' \\ B_t' \end{pmatrix} C_t\end{aligned}$$

pourvu que l'inverse existe, mais le cas où $r > 2p$ est le plus intéressant.

Chapitre 4

Simulation et expérimentation

4.1 Introduction

Dans cette partie, nous avons réalisé 500 simulations de trois modèles différents :

$$X_t = bX_{t-2}e_{t-1} + e_t \text{ [bruit blanc non linéaire]} \quad (4.1)$$

$$X_t = bX_{t-1}e_{t-1} + e_t \text{ [Modèle diagonal purement bilinéaire]} \quad (4.2)$$

$$X_t = aX_{t-1} + bX_{t-1}e_{t-1} + e_t \text{ [Modèle diagonal]} \quad (4.3)$$

où $(e_t)_{t \in \mathbb{Z}}$ est un bruit blanc Gaussien $\mathcal{N}(1, 0)$. Chacune des 500 trajectoires simulées est de longueur $N \in \{500, 1000, 2000, 3000, 4000, 5000, 10000\}$. Pour chaque simulation, les paramètres ont été estimés par la méthode GMM. Cette étude expérimental a pour but :

4.2 Description des modèles

4.2.1 Bruit blanc non linéaire

D'après l'étude envisagée dans le chapitre 2, on vérifie aisément que

$$\begin{aligned} b^2 < 1 &\iff (X_t)_{t \in \mathbb{Z}} \text{ est stationnaire au second ordre et ergodique} \\ 3b^4 < 1 &\iff E\{X_t^4\} < +\infty \\ 15b^6 < 1 &\iff E\{X_t^6\} < +\infty \\ 105b^8 < 1 &\iff E\{X_t^8\} < +\infty. \end{aligned}$$

D'après la propriété (22) le processus se comporte comme un processus de bruit blanc faible, du point de vue de sa fonction de covariance. La condition d'orthogonalité est cependant

$$h(b, X_t) = \frac{1}{N} \sum_{t=1}^N \{X_t X_{t-1} X_{t-2} - b X_t^2\}.$$

La Figure 1, montre une réalisation de longueur $N = 250$ avec $b = -0.75$.

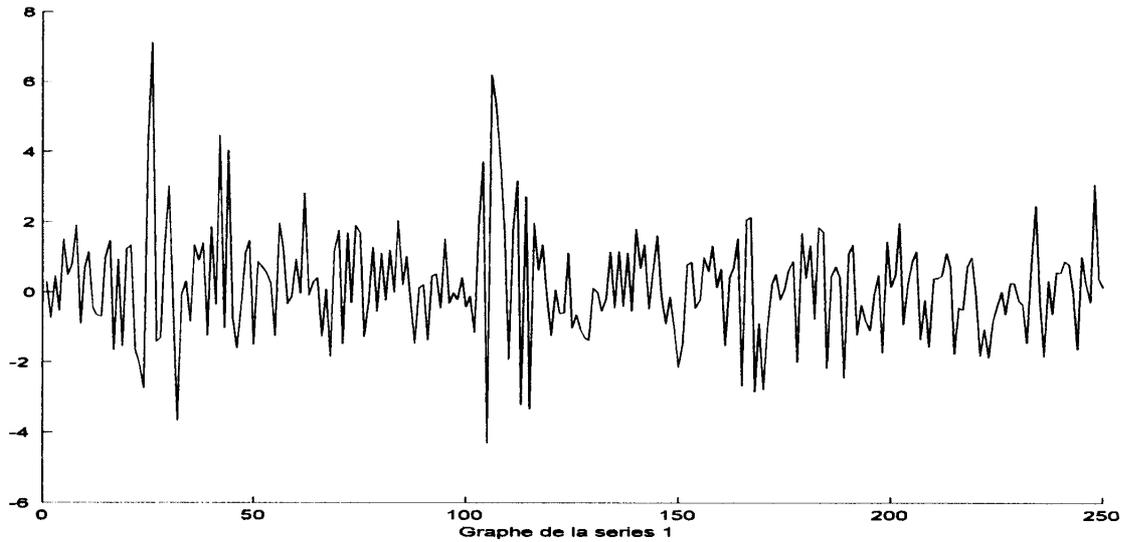


FIG. 4-1 – Trajectoire de longueur $N = 250$

La figure 2 montre les covariances de la série des 250 dernières valeurs qui ressemblent assez à celles d'un bruit blanc, comme on pouvait s'y attendre.

Les tableaux *Table(1.1) – Table(1.8)* contient les résultats de la simulation. Les résultats sont acceptables et même sont très bonnes. Notons que les résultats des tables (1.5) – (1.8) semblent un peu loin des vraies valeurs, ceci n'est pas surprenant car les valeurs attribuées excluaient le processus d'être dans \mathbb{L}_4 . Les coefficients d'aplatissements et d'asymétries témoignent cette confirmation.

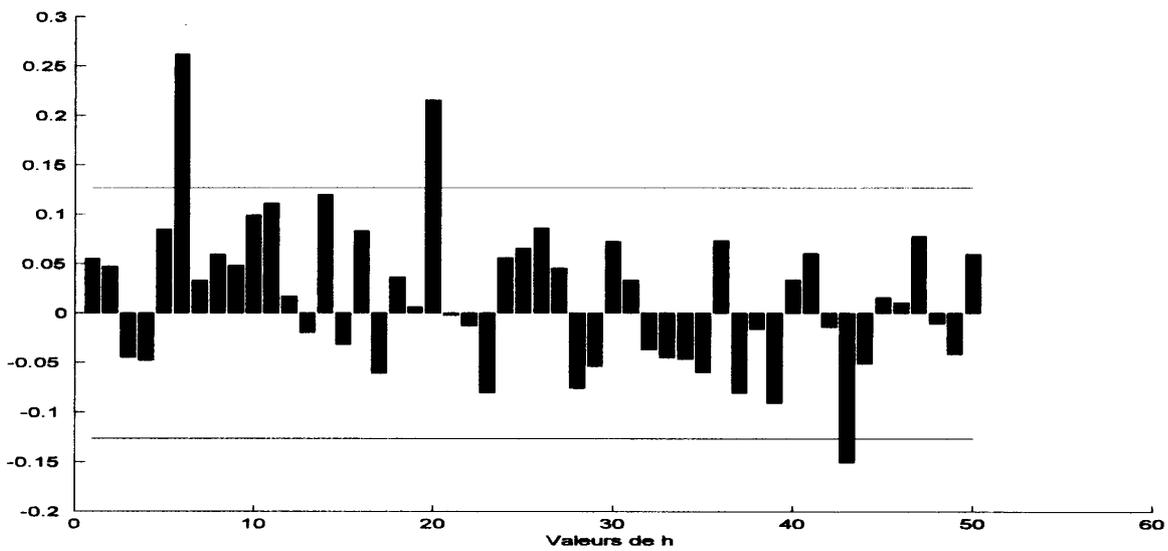


FIG. 4-2 – Fonction de covariance

N	Vraie valeur	Valeurs Minimales	Valeurs Maximales	Moyennes	Ecart Types	RMSE ¹	Kurtosis ²	Skewness ³
500	0.15	0.031900	0.288787	0.149173	0.049104	0.049062	2.807609	0.244835
1000		0.048747	0.246548	0.148889	0.036178	0.036159	2.801506	0.042743
2000		0.080121	0.219104	0.149332	0.024500	0.024485	2.779312	0.020240
3000		0.088119	0.208133	0.149444	0.020647	0.020634	3.052231	0.061942
4000		0.100410	0.200548	0.149505	0.018327	0.018315	2.880937	0.026482
5000		0.102532	0.201294	0.149373	0.016327	0.016322	3.264846	0.098760
10000		0.119312	0.192487	0.150086	0.011056	0.011046	3.309269	0.269949

Table(1.1)

¹Roots mean squares errors

²Appelé aussi coefficient d'aplatissement, il est défini par $E \left\{ (X - E \{X\})^4 \right\} / \sigma^4$

³Appelé coefficient d'asymétrie et défini par $E \left\{ (X - E \{X\})^3 \right\} / \sigma^3$

N	Vraie valeur	Valeurs Minimales	Valeurs Maximales	Moyennes	Ecart Types	RMSE	Kurtosis	Skewness
500	-0.15	-0.305688	-0.025751	-0.148475	0.051215	0.051186	2.845172	-0.372834
1000		-0.303789	-0.057482	-0.148831	0.036546	0.036529	3.331641	-0.463185
2000		-0.229341	-0.076459	-0.149624	0.025336	0.025313	2.938518	-0.070072
3000		-0.213311	-0.082773	-0.149807	0.021494	0.021473	2.905320	-0.001650
4000		-0.196895	-0.098245	-0.149701	0.018649	0.018632	2.775805	0.034401
5000		-0.194619	-0.103920	-0.149820	0.016587	0.016572	2.943636	-0.008635
10000		-0.180578	-0.113844	-0.149206	0.011435	0.011451	2.906027	0.174013

Table(1.2)

N	Vraie valeur	Valeurs Minimales	Valeurs Maximales	Moyennes	Ecart Types	RMSE	Kurtosis	Skewness
500	0.5	0.219261	0.923256	0.489463	0.101820	0.102263	4.080008	0.766658
1000		0.332384	0.804838	0.494724	0.077944	0.078044	3.580456	0.547183
2000		0.367583	0.813362	0.498178	0.058301	0.058271	5.156342	0.832100
3000		0.384510	0.679591	0.497660	0.046779	0.046791	3.697651	0.421988
4000		0.393611	0.653002	0.498466	0.042084	0.042070	3.349215	0.308335
5000		0.393048	0.631151	0.497506	0.037173	0.037220	3.374309	0.293362
10000		0.433656	0.597462	0.498015	0.025613	0.025664	3.516798	0.383249

Table(1.3)

N	Vraie valeur	Valeurs Minimales	Valeurs Maximales	Moyennes	Ecart Types	RMSE	Kurtosis	Skewness
500	-0.5	-0.900296	-0.262786	-0.490271	0.106521	0.106858	3.912241	-0.74657
1000		-1.308013	-0.311569	-0.495105	0.086061	0.086114	18.51711	-2.086263
2000		-0.954441	-0.346084	-0.499180	0.065331	0.065271	9.833419	-1.391495
3000		-0.831365	-0.367128	-0.498998	0.065331	0.065271	6.484617	-0.956646
4000		-0.741971	-0.377650	-0.498602	0.045761	0.045737	4.844893	-0.785019
5000		-0.689227	-0.395798	-0.499025	0.040715	0.040686	3.908310	-0.569466
10000		-0.609799	-0.423131	-0.499077	0.029203	0.029189	3.404437	-0.302449

Table(1.4)

N	Vraie valeur	Valeurs Minimales	Valeurs Maximales	Moyennes	Ecart Types	RMSE	Kurtosis	Skewness
500	0.75	0.220542	1.715995	0.706540	0.192792	0.197442	5.822484	1.188900
1000		-0.699759	1.749542	0.722016	0.169472	0.171600	15.331121	-0.167511
2000		-0.284929	2.271193	0.739770	0.152219	0.152411	29.783371	2.420172
3000		-0.020070	1.854511	0.743421	0.129723	0.129760	25.554329	2.583854
4000		0.173976	1.661458	0.744856	0.113607	0.113610	19.546043	2.302956
5000		0.261361	1.474363	0.742481	0.101877	0.102052	15.109303	1.941718
10000		0.466103	1.453641	0.744285	0.078222	0.078352	18.650134	2.452496

Table(1.5)

N	Vraie valeur	Valeurs Minimales	Valeurs Maximales	Moyennes	Ecart Types	RMSE	Kurtosis	Skewness
500	-0.75	-2.235340	6.044031	-0.699873	0.366391	0.369441	231.551501	12.174715
1000		-2.631004	3.924927	-0.727175	0.289654	0.290263	141.451174	6.986744
2000		-2.125006	2.119108	-0.736249	0.197016	0.197299	96.733733	4.744968
3000		-1.865243	1.269343	-0.738331	0.153849	0.154137	67.314081	2.970782
4000		-1.711214	0.822787	-0.740977	0.131335	0.131513	49.403695	1.860495
5000		-1.849989	0.598203	-0.744255	0.125250	0.125257	44.981579	-0.020051
10000		-1.514648	-0.067956	-0.747043	0.095284	0.095235	23.970152	-2.083577

Table(1.6)

N	Vraie valeur	Valeurs Minimales	Valeurs Maximales	Moyennes	Ecart Types	RMSE	Kurtosis	Skewness
500	0.95	-4.692332	2.780635	0.804063	0.481369	0.502544	38.634431	-2.537785
1000		-8.563654	3.955384	0.802084	0.671132	0.686583	99.773044	-7.624522
2000		-8.208191	3.982638	0.851055	0.647059	0.653940	96.680718	-7.034364
3000		-7.844210	3.617019	0.873034	0.601312	0.605621	103.198538	-7.002708
4000		-7.187570	3.248651	0.884280	0.553440	0.556779	106.137297	-7.265519
5000		-6.934808	2.997401	0.886235	0.527483	0.530800	109.976778	-7.398343
10000		-5.625394	3.086419	0.890745	0.445854	0.449332	100.580979	-6.280355

Table(1.7)

N	Vraie valeur	Valeurs Minimales	Valeurs Maximales	Moyennes	Ecart Types	RMSE	Kurtosis	Skewness
500	-0.95	-4.125002	30.804056	-0.744654	1.493991	1.506557	399.503078	18.860470
1000		-3.220532	27.233578	-0.803142	1.327824	1.334600	399.285230	18.831010
2000		-7.436513	22.283159	-0.822309	1.136035	1.142059	346.023084	16.442526
3000		-6.798837	16.011460	-0.848956	0.874073	0.879025	282.611121	13.662640
4000		-6.020338	13.667254	-0.863447	0.768031	0.772129	261.772056	12.843510
5000		-8.724003	12.176391	-0.882127	0.779298	0.781471	181.532503	6.909672
10000		-4.443427	5.694977	-0.882226	0.518713	0.522607	71.444693	3.923038

Table(1.8)

4.2.2 Modèle diagonal purement bilinéaire

Pour ce modèle, la condition d'orthogonalité est

$$h(a, X_t) = \frac{1}{N} \sum_{t=1}^N \{X_t X_{t-1} - 2bX_t\}.$$

La Figure 2, montre une réalisation de longueur $N = 250$ avec $b = 0.5$. Les résultats de la simulation sont portés dans les tables Table(2,1)-Table(2,8).

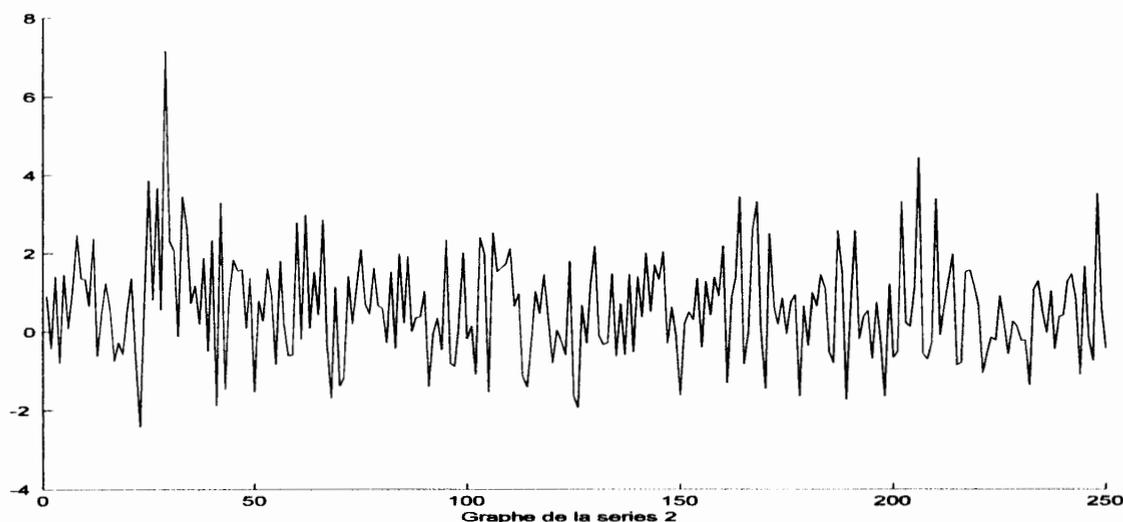


FIG. 4-3 – Trajectoire de longueur $N = 250$

N	Vraie valeur	Valeurs Minimales	Valeurs Maximales	Moyennes	Ecart Types	RMSE	Kurtosis	Skewness
500	0.15	0.019404	0.291040	0.148049	0.047540	0.047533	2.727920	-0.075323
1000		0.051051	0.259974	0.148727	0.032931	0.032923	3.464073	0.186353
2000		0.064720	0.224105	0.150678	0.024168	0.024153	3.278971	-0.028068
3000		0.075690	0.205114	0.149058	0.019011	0.019015	3.481697	-0.028508
4000		0.080193	0.200859	0.149250	0.016616	0.016616	3.456664	-0.063338
5000		0.090535	0.201585	0.148768	0.015043	0.015078	3.124189	-0.134763
10000		0.110284	0.183082	0.149393	0.010682	0.010689	3.214715	-0.073221

Table(2.1)

N	Vraie valeur	Valeurs Minimales	Valeurs Maximales	Moyennes	Ecart Types	RMSE	Kurtosis	Skewness
500	-0.15	-0.286354	-0.029620	-0.148880	0.047263	0.047229	2.772096	-0.225042
1000		-0.257838	-0.047347	-0.149525	0.033137	0.033107	3.249245	0.007953
2000		-0.226045	-0.077887	-0.148768	0.023873	0.023881	3.270108	-0.054724
3000		-0.213273	-0.091418	-0.150230	0.019203	0.019185	3.050214	-0.072958
4000		-0.214914	-0.097298	-0.150254	0.016408	0.016394	3.150119	-0.078940
5000		-0.207627	-0.102630	-0.151020	0.014878	0.014898	2.997076	-0.054623
10000		-0.187344	-0.117242	-0.150514	0.010613	0.010615	3.013678	-0.050146

Table(2.2)

N	Vraie valeur	Valeurs Minimales	Valeurs Maximales	Moyennes	Ecart Types	RMSE	Kurtosis	Skewness
500	0.5	0.325076	0.721047	0.490370	0.065838	0.066473	2.841483	0.081845
1000		0.374871	0.655839	0.493884	0.045691	0.046053	3.225149	0.251789
2000		0.374664	0.590643	0.498862	0.034518	0.034502	3.038555	-0.015935
3000		0.388433	0.580674	0.496958	0.027752	0.027891	3.602223	-0.051344
4000		0.403680	0.573117	0.497844	0.024304	0.024375	3.392521	-0.019671
5000		0.417669	0.565704	0.497525	0.021872	0.021990	2.963240	-0.126136
10000		0.442435	0.549892	0.499051	0.015788	0.015801	3.164754	-0.061503

Table(2.3)

N	Vraie valeur	Valeurs Minimales	Valeurs Maximales	Moyennes	Ecart Types	RMSE	Kurtosis	Skewness
500	-0.5	-0.703884	-0.303777	-0.490503	0.066365	0.066975	2.939967	-0.369187
1000		-0.663101	-0.378942	-0.495158	0.046889	0.047092	3.083858	-0.200438
2000		-0.595368	-0.404397	-0.496175	0.034241	0.034420	3.082200	0.011281
3000		-0.578803	-0.429217	-0.498225	0.027744	0.027773	2.771156	-0.150159
4000		-0.581376	-0.440830	-0.498594	0.023470	0.023489	2.899324	-0.153298
5000		-0.576455	-0.446816	-0.500167	0.021325	0.021304	2.824927	-0.076964
10000		-0.545640	-0.451180	-0.499962	0.015149	0.015134	3.019212	0.021575

Table(2.4)

N	Vraie valeur	Valeurs Minimales	Valeurs Maximales	Moyennes	Ecart Types	RMSE	Kurtosis	Skewness
500	0.75	0.501091	1.051068	0.733765	0.097074	0.098326	3.153271	0.251873
1000		0.558494	0.963842	0.739668	0.066899	0.067626	3.083489	0.311427
2000		0.573453	0.877254	0.748271	0.050912	0.050890	2.964043	0.018779
3000		0.593142	0.896048	0.745795	0.041442	0.041614	3.698242	-0.038385
4000		0.624255	0.897312	0.747382	0.035639	0.035700	3.784829	0.116145
5000		0.636233	0.861542	0.746867	0.032258	0.032378	3.217428	-0.018940
10000		0.666742	0.819469	0.749055	0.023760	0.023755	3.204793	-0.029975

Table(2.5)

N	Vraie valeur	Valeurs Minimales	Valeurs Maximales	Moyennes	Ecart Types	RMSE	Kurtosis	Skewness
500	-0.75	-1.080845	-0.121216	-0.734000	0.099254	0.100437	5.726203	0.041604
1000		-0.978492	-0.472197	-0.743908	0.070586	0.070778	3.338009	-0.032066
2000		-0.908768	-0.607437	-0.745099	0.051306	0.051488	3.184085	-0.059897
3000		-0.872316	-0.642462	-0.747368	0.041323	0.041365	3.128593	-0.177613
4000		-0.876202	-0.640968	-0.748022	0.035295	0.035315	3.342206	-0.220310
5000		-0.870001	-0.655992	-0.750530	0.032276	0.032248	3.147495	-0.089741
10000		-0.812595	-0.688146	-0.750068	0.022349	0.022327	3.025408	-0.070319

Table(2.6)

N	Vraie valeur	Valeurs Minimales	Valeurs Maximales	Moyennes	Ecart Types	RMSE	Kurtosis	Skewness
500	0.95	0.183741	1.620585	0.928038	0.162693	0.164007	4.747512	0.387680
1000		0.576656	1.365568	0.936134	0.119089	0.119775	3.669951	0.534550
2000		0.009301	1.236831	0.947210	0.098808	0.098749	18.545760	-1.529520
3000		0.385941	1.253819	0.942537	0.077637	0.077917	8.452895	-0.645284
4000		0.624381	1.251657	0.946296	0.064700	0.064742	5.708781	-0.181055
5000		0.687222	1.178602	0.945094	0.058172	0.058321	4.079300	-0.154565
10000		0.664959	1.129267	0.947744	0.046077	0.046086	6.226058	-0.355996

Table(2.7)

N	Vraie valeur	Valeurs Minimales	Valeurs Maximales	Moyennes	Ecart Types	RMSE	Kurtosis	Skewness
500	-0.95	-1.581648	3.320257	-0.929004	0.248292	0.248930	173.02589	9.833162
1000		-1.318315	1.143668	-0.941730	0.156342	0.156405	66.488824	4.969173
2000		-1.377555	-0.153515	-0.943077	0.100570	0.100707	11.202344	0.703617
3000		-1.298943	-0.496187	-0.946968	0.082332	0.082305	5.586584	0.121990
4000		-1.197517	-0.629820	-0.947983	0.069561	0.069520	4.514014	0.002550
5000		-1.144008	-0.665997	-0.952424	0.063590	0.063572	4.101583	0.179145
10000		-1.078347	-0.641128	-0.949673	0.044718	0.044675	7.076049	0.563250

Table(2.8)

4.2.3 Le modèle diagonal

C'est le modèle le plus utilisé dans la littérature. Pour ce modèle, les conditions d'orthogonalités sont

$$\begin{cases} h_1(\theta, X_t) = \frac{1}{N} \sum_{t=1}^N \{X_t X_{t-1} - aX_t^2 - 2bX_t\} \\ h_2(\theta, X_t) = \frac{1}{N} \sum_{t=1}^N \{X_t X_{t-2} - aX_t X_{t-1} - bX_t\} \\ \theta = (a, b)' \end{cases}$$

La Figure 3, montre une réalisation de longueur $N = 250$ avec $a = 0.15$ et $b = -0.75$. Les résultats de la simulation sont portés dans les tables Table(3.1)-Table(3.8).

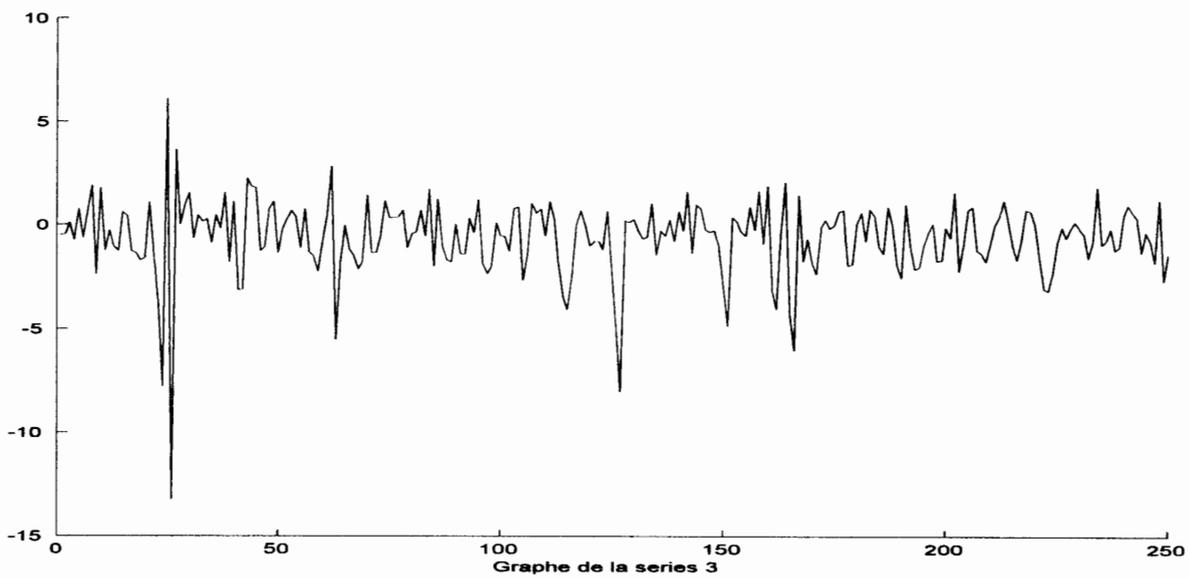


FIG. 4-4 - Trajectoire de longueur $N = 250$

Longueur de la série	Vraie valeur	Valeurs Minimales	Valeurs Maximales	Moyennes	Ecart Types	RMSE	Kurtosis	Skewness
500	0.15 0.75	-0.280356	0.610364	0.124806	0.133321	0.135550	3.621538	0.219500
		0.442179	1.274806	0.762794	0.125944	0.126467	3.644951	0.426792
1000		-0.301598	0.475666	0.129845	0.103305	0.105151	3.870316	0.287840
		0.440714	1.109327	0.760560	0.096676	0.097155	4.234717	0.428626
2000		-0.088257	0.523648	0.138250	0.083076	0.083821	4.493996	0.593054
		0.501206	0.996632	0.760027	0.078476	0.079036	3.437744	0.051766
3000		-0.113320	0.461952	0.139496	0.074706	0.075367	4.080021	0.364528
		0.532529	1.077295	0.755799	0.067259	0.067442	4.306034	0.191833
4000		-0.159367	0.477077	0.141388	0.067987	0.068463	5.081354	0.370292
		0.505433	1.101315	0.755694	0.062178	0.062376	5.322105	0.140791
5000		-0.134634	0.464862	0.145106	0.066311	0.066425	5.357929	0.477892
		0.448398	1.047248	0.751714	0.060108	0.060072	5.642156	-0.244512
10000		-0.065855	0.417432	0.146611	0.049152	0.049219	6.014939	0.615339
		0.540007	0.957436	0.751942	0.042900	0.042901	5.821213	-0.301455

Table(3.1)

Longueur de la série	Vraie valeur	Valeurs Minimales	Valeurs Maximales	Moyennes	Ecart Types	RMSE	Kurtosis	Skewness
500	-0.5 0.5	-0.716872	-0.285037	-0.496668	0.062999	0.063024	3.158355	0.131867
		0.327420	0.755831	0.498831	0.067079	0.067022	3.431147	0.253883
1000		-0.658478	-0.321715	-0.499507	0.044728	0.044686	4.190466	0.105798
		0.361158	0.669226	0.499340	0.045982	0.045941	3.597700	0.263704
2000		-0.603669	-0.382987	-0.499375	0.034509	0.034480	3.050679	0.056403
		0.392279	0.618403	0.502024	0.034856	0.034880	3.070196	0.061387
3000		-0.586663	-0.410703	-0.499843	0.028432	0.028404	2.867477	-0.032919
		0.394899	0.590968	0.499788	0.027491	0.027464	3.449780	-0.038511
4000		-0.571795	-0.413098	-0.500370	0.023868	0.023847	2.990839	-0.091853
		0.412889	0.567185	0.499990	0.023578	0.023554	3.378292	-0.042300
5000		-0.558401	-0.439915	-0.500137	0.021729	0.021707	2.647781	-0.124873
		0.424549	0.561121	0.499126	0.021369	0.021365	3.222685	-0.160357
10000		-0.548603	-0.464606	-0.500278	0.016025	0.016011	2.789314	-0.255236
		0.445194	0.546447	0.499958	0.015815	0.015799	3.137818	-0.064153

Table(3.2)

Longueur de la série	Vraie valeur	Valeurs Minimales	Valeurs Maximales	Moyennes	Ecart Types	RMSE	Kurtosis	Skewness
500	0.75 0.00	0.629245	0.826688	0.737967	0.031439	0.033634	3.289990	-0.155867
		-0.180365	0.134635	-0.000545	0.049079	0.049033	3.106759	-0.192470
1000		0.673073	0.806494	0.743730	0.021155	0.022044	3.134682	-0.123199
		-0.110440	0.116113	-0.000426	0.032902	0.032872	3.610027	0.072852
2000		0.699126	0.786435	0.746725	0.015182	0.015517	2.990898	-0.133780
		-0.088477	0.069073	0.000976	0.023432	0.023429	3.419461	-0.083653
3000		0.706755	0.778988	0.747483	0.012843	0.013075	2.864544	-0.190520
		-0.073768	0.065801	-0.000518	0.018776	0.018765	3.590703	-0.067363
4000		0.710601	0.777112	0.748456	0.010813	0.010912	3.119740	-0.240908
		-0.069025	0.054736	-0.000499	0.016045	0.016036	3.519360	-0.086408
5000		0.722007	0.775702	0.749182	0.009555	0.009581	2.913900	-0.197533
		-0.057735	0.047371	-0.001124	0.014372	0.014402	3.120914	-0.102849
10000		0.729611	0.768878	0.749611	0.006824	0.006828	2.863784	-0.083952
		-0.038647	0.030859	-0.000529	0.010198	0.010202	3.136756	-0.073060

Table(3.3)

Longueur de la série	Vraie valeur	Valeurs Minimales	Valeurs Maximales	Moyennes	Ecart Types	RMSE	Kurtosis	Skewness
500	0.0 0.75	-0.388035	0.452084	-0.011516	0.124333	0.124741	3.670259	0.242829
		0.470925	1.212116	0.752150	0.112481	0.112389	3.527874	0.435907
1000		-0.467454	0.325534	-0.011849	0.096151	0.096782	4.208023	0.106580
		0.406310	1.092587	0.753817	0.085290	0.085290	4.638914	0.495209
2000		-0.281217	0.275087	-0.005886	0.075713	0.075866	3.795223	0.246764
		0.551620	0.989773	0.755163	0.067334	0.067464	3.380707	0.251585
3000		-0.192133	0.259759	-0.004214	0.066279	0.066347	3.668566	0.256865
		0.539277	0.945827	0.750680	0.055370	0.055319	3.723908	0.015808
4000		-0.234395	0.370454	-0.002828	0.061408	0.061412	6.159933	0.507493
		0.501969	0.971392	0.750944	0.051497	0.051454	4.669185	-0.136061
5000	-0.205124	0.370828	-0.000368	0.060329	0.060270	7.975752	0.860592	
	0.428680	0.921025	0.748233	0.049773	0.049755	7.168902	-0.761933	
10000	-0.152193	0.265725	-0.000030	0.044465	0.044420	7.260229	0.818798	
	0.532854	0.867497	0.749474	0.035196	0.035164	6.501994	-0.524686	

Table(3.4)

Longueur de la série	Vraie valeur	Valeurs Minimales	Valeurs Maximales	Moyennes	Ecart Types	RMSE	Kurtosis	Skewness
500	0.5 0.5	0.135475	0.812384	0.460266	0.099673	0.107208	3.373289	0.089166
		0.256224	0.959223	0.532299	0.109907	0.114449	3.429262	0.467537
1000		0.297314	0.745745	0.476357	0.072823	0.076495	3.272145	0.408832
		0.279567	0.859522	0.519067	0.078809	0.081006	3.870656	0.429135
2000		0.330151	0.790349	0.484935	0.059501	0.061321	4.616713	0.679014
		0.254152	0.693904	0.514625	0.060081	0.061777	3.812750	0.044288
3000		0.353542	0.744307	0.487286	0.052535	0.054000	4.311386	0.622702
		0.291141	0.691969	0.509594	0.050580	0.051432	4.085622	0.046140
4000		0.386812	0.714056	0.490398	0.045236	0.046200	4.484002	0.761184
		0.318661	0.647999	0.507414	0.044075	0.044651	3.821023	-0.130219
5000	0.403681	0.695367	0.493102	0.042512	0.043026	4.360376	0.755605	
	0.315546	0.634482	0.504468	0.041372	0.041571	4.233039	-0.347886	
10000	0.415357	0.622324	0.495196	0.031205	0.031542	3.908589	0.603870	
	0.379540	0.602622	0.503669	0.030192	0.030384	3.733004	-0.231874	

Table(3.5)

Longueur de la série	Vraie valeur	Valeurs Minimales	Valeurs Maximales	Moyennes	Ecart Types	RMSE	Kurtosis	Skewness
500	0.85 0.25	0.819534	0.897112	0.848110	0.043737	0.048538	3.387797	0.395224
		0.174590	0.306430	0.251874	0.095445	0.095949	3.484526	-0.155083
1000		0.639457	0.944596	0.828896	0.034017	0.036259	4.431460	-0.548467
		0.112163	0.698118	0.279412	0.065924	0.066404	6.699684	1.479991
2000		0.751599	0.934100	0.837930	0.025384	0.026216	3.242350	0.159856
		0.133204	0.474173	0.268487	0.049451	0.049626	3.827481	0.617374
3000		0.770775	0.921668	0.841270	0.020856	0.021399	3.223193	0.081421
		0.137708	0.418321	0.260950	0.038883	0.038887	3.662867	0.397878
4000		0.780354	0.910931	0.843452	0.017982	0.018430	3.285482	0.092001
		0.160354	0.386329	0.258171	0.034465	0.034534	3.545601	0.345082
5000	0.785082	0.906498	0.844866	0.016124	0.016493	3.260443	0.086493	
	0.160102	0.387710	0.255832	0.032494	0.032584	3.688417	0.316083	
10000	0.810853	0.912675	0.847254	0.011828	0.011920	3.425694	0.354317	
	0.148182	0.342852	0.253451	0.023618	0.023621	3.454872	0.067836	

Table(3.6)

Longueur de la série	Vraie valeur	Valeurs Minimales	Valeurs Maximales	Moyennes	Ecart Types	RMSE	Kurtosis	Skewness
500	-0.85 -0.25	-0.926826	-0.726155	-0.843211	0.030462	0.031179	3.093559	0.173518
		-0.423609	-0.110847	-0.249528	0.053818	0.053767	2.748383	-0.150898
1000		-0.939119	-0.783742	-0.845811	0.022596	0.022959	3.439305	-0.278369
		-0.378838	-0.142100	-0.249452	0.037200	0.037167	3.210443	-0.102258
2000		-0.917171	-0.803802	-0.848640	0.016713	0.016751	3.197360	-0.218565
		-0.340314	-0.172626	-0.248628	0.025900	0.025910	3.214374	0.007484
3000		-0.908410	-0.812917	-0.848887	0.014260	0.014289	3.216195	-0.292459
		-0.333329	-0.191713	-0.249944	0.021010	0.020989	3.242328	-0.197969
4000		-0.897587	-0.815675	-0.849340	0.012055	0.012061	3.287594	-0.262100
		-0.316995	-0.204029	-0.249917	0.017877	0.017860	3.021428	-0.151252
5000		-0.893944	-0.821723	-0.849568	0.010755	0.010753	3.499754	-0.243074
		-0.299954	-0.207318	-0.250840	0.016318	0.016324	2.761626	-0.106717
10000		-0.880090	-0.827324	-0.850004	0.007735	0.007727	3.880197	-0.395836
		-0.288811	-0.216736	-0.250403	0.011550	0.011546	2.955317	-0.055416

Table(3.7)

Longueur de la série	Vraie valeur	Valeurs Minimales	Valeurs Maximales	Moyennes	Ecart Types	RMSE	Kurtosis	Skewness
500	0.95 0.15	0.856927	0.980044	0.929466	0.023181	0.030951	3.045565	-0.448584
		0.011640	0.543042	0.186017	0.079899	0.087569	4.774466	0.899701
1000		0.872224	0.979457	0.938385	0.017465	0.020960	3.311946	-0.503084
		0.044901	0.370809	0.170563	0.050365	0.054354	4.006478	0.761661
2000		0.904782	0.980699	0.943616	0.012596	0.014110	3.200910	0.001586
		0.069548	0.324729	0.164126	0.035692	0.038352	3.832042	0.444831
3000		0.913561	0.979544	0.945255	0.010745	0.011736	3.208728	0.191342
		0.073859	0.255588	0.158873	0.028626	0.029943	3.342032	0.126185
4000		0.919039	0.977928	0.946215	0.009788	0.010486	3.107607	0.194709
		0.080487	0.241204	0.157101	0.025159	0.026118	3.526358	0.061226
5000		0.925242	0.976702	0.947059	0.008963	0.009424	3.116452	0.260826
		0.075814	0.243018	0.155031	0.023974	0.024473	3.262629	-0.151701
10000		0.930713	0.976197	0.948491	0.007019	0.007172	3.146503	0.274143
		0.086733	0.200299	0.152873	0.017929	0.018140	3.203265	-0.096636

Table(3.8)

4.2.4 Résultats et commentaires

D'après les résultats empiriques qui sont reportés dans les tableaux (1.1) – (3.8), on observe les principales conclusions suivantes :

L'augmentation de la taille d'échantillon joue un rôle essentiel pour l'obtention des meilleurs résultats. Cela est prouvé quand la moyenne des estimateurs $\hat{\theta}_N$ calculée sur les 500 simulations est plus proche de la vraie valeur, plus la taille d'échantillon est élevée, ce qui améliore les résultats en termes de réduction de biais. D'autres parts on observe que plus la taille d'échantillon est élevée, plus l'écart type des estimations sont proche de zéro, par conséquent les estimateurs sont plus précis. Donc, la théorie asymptotique présentée dans le théorème (26) est confirmée.

On observe aussi que lorsque la taille d'échantillon augmente, la moyenne de coefficients d'aplatissement se rapproche de la valeur 3, et le coefficient d'asymétrie est plus proche de zéro. Cette observation confirme la normalité asymptotique de l'estimateur $\hat{\theta}_N$ obtenue dans le théorème (27).

4.3 Conclusion

Nous nous intéressons ici à l'estimation des paramètres des modèles bilinéaires. Sous certaines conditions, nous établissons la convergence et la normalité asymptotiques des estimateurs GMM de la distribution empirique vers la distribution théorique, avec une matrice de covariance connue. Les résultats empiriques obtenus dans les tableaux (1.1) – (3.8) confirment la théorie asymptotique présentée dans le chapitre (3).

La raison principale du choix de cette méthode d'estimation provient du fait que

1. Beaucoup d'estimateurs peuvent être vus comme cas spéciaux de GMM.
2. GMM est une alternative fondée sur des hypothèses minimales.
3. L'estimation de GMM est souvent possible où une analyse de la probabilité est extrêmement difficile.

Chapitre 5

Appendices

Dans ces appendices on donne les définition, les concepts et les outils fondamentaux qui s'avèrent nécessaires dans notre étude.

5.1 Dérivation vectorielle et matricielle

5.1.1 Dérivée de Ax

Soient $a \in \mathbb{R}^n$, $x \in \mathbb{R}^n$ (tous les vecteurs sont des vecteurs colonnes). Alors

$$\begin{aligned}\frac{\partial (a'x)}{\partial x} &= \begin{pmatrix} \frac{\partial y}{\partial x_1} \\ \vdots \\ \frac{\partial y}{\partial x_n} \end{pmatrix} = \begin{pmatrix} \frac{a'x}{\partial x_1} \\ \vdots \\ \frac{a'x}{\partial x_n} \end{pmatrix} = \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} = a. \\ \frac{\partial (a'x)}{\partial x'} &= \left(\frac{\partial y}{\partial x_1} \dots \frac{\partial y}{\partial x_n} \right) = a'.\end{aligned}$$

Soit A une matrice de $m \times n$,

$$A = \begin{pmatrix} a'_1 \\ \vdots \\ a'_m \end{pmatrix},$$

où $a_j \in \mathbb{R}^n$ pour $j = 1, \dots, m$.

$$\frac{\partial (Ax)}{\partial x'} = \begin{pmatrix} \frac{\partial (a'_1 x)}{\partial x'} \\ \vdots \\ \frac{\partial (a'_m x)}{\partial x'} \end{pmatrix} = \begin{pmatrix} a'_1 \\ \vdots \\ a'_m \end{pmatrix} = A.$$

De la même façon,

$$\frac{\partial (x' A')}{\partial x} = \left(\frac{\partial (x' a_1)}{\partial x} \dots \frac{\partial (x' a_m)}{\partial x} \right) = (a_1 \dots a_m) = A'.$$

Soient $\alpha \in \mathbb{R}^r$, et $x = x(\alpha)$. Alors

$$\frac{\partial x}{\partial \alpha'} = \begin{pmatrix} \frac{\partial x_1}{\partial \alpha'} \\ \vdots \\ \frac{\partial x_n}{\partial \alpha'} \end{pmatrix},$$

c'est une matrice $n \times r$. Soit $x = x(\alpha)$. Alors

$$\frac{\partial (a' x)}{\partial \alpha} = \begin{pmatrix} \frac{\partial (a_1 x_1 + \dots + a_n x_n)}{\partial \alpha_1} \\ \vdots \\ \frac{\partial (a_1 x_1 + \dots + a_n x_n)}{\partial \alpha_r} \end{pmatrix} = \begin{pmatrix} a_1 \frac{\partial x_1}{\partial \alpha_1} + \dots + a_n \frac{\partial x_n}{\partial \alpha_1} \\ \vdots \\ a_1 \frac{\partial x_1}{\partial \alpha_r} + \dots + a_n \frac{\partial x_n}{\partial \alpha_r} \end{pmatrix} = \frac{\partial x'}{\partial \alpha} a,$$

un vecteur de r , et

$$\frac{\partial (a' x)}{\partial \alpha'} = a' \frac{\partial x}{\partial \alpha'}.$$

Ensuite,

$$\frac{\partial (Ax)}{\partial \alpha'} = \begin{pmatrix} \frac{\partial (a'_1 x)}{\partial \alpha'} \\ \vdots \\ \frac{\partial (a'_m x)}{\partial \alpha'} \end{pmatrix} = \begin{pmatrix} a'_1 \frac{\partial x}{\partial \alpha'} \\ \vdots \\ a'_m \frac{\partial x}{\partial \alpha'} \end{pmatrix} = A \frac{\partial x}{\partial \alpha'}.$$

Notons que

$$\frac{\partial x}{\partial x'} = I_n.$$

5.2 Dérivées de $z'Ax$

Soient $x \in \mathbb{R}^n$, $z \in \mathbb{R}^m$, et A est une matrice $m \times n$. Posons $c = A'z \in \mathbb{R}^n$. Alors,

$$\frac{\partial (z'Ax)}{\partial x} = \frac{\partial (c'x)}{\partial x} = A'z.$$

Ensuite,

$$\frac{\partial (z'Ax)}{\partial z} = Ax.$$

Soit $x = x(\alpha)$, et $z = z(\alpha)$. Alors

$$\begin{aligned} \frac{\partial (z'Ax)}{\partial \alpha} &= \frac{\partial z'}{\partial \alpha} \frac{\partial (z'Ax)}{\partial z} + \frac{\partial x'}{\partial \alpha} \frac{\partial (z'Ax)}{\partial x} \\ &= \frac{\partial z'}{\partial \alpha} Ax + \frac{\partial x'}{\partial \alpha} A'z. \end{aligned}$$

Ensuite, si $m = n$,

$$\frac{\partial (x'Ax)}{\partial x} = Ax + A'x = (A + A')x.$$

Si A est symétrique,

$$\frac{\partial (x'Ax)}{\partial x} = 2Ax.$$

Si $x = x(\alpha)$, A est symétrique et indépendant de α ,

$$\frac{\partial (x'Ax)}{\partial \alpha} = 2 \frac{\partial x'}{\partial \alpha} Ax.$$

La dérivée de $z'Ax$ en ce qui concerne A est donné par

$$\begin{aligned}
\frac{\partial (z'Ax)}{\partial A} &= \begin{pmatrix} \frac{\partial (z'Ax)}{\partial A_{11}} & \cdots & \frac{\partial (z'Ax)}{\partial A_{1n}} \\ \vdots & \cdots & \vdots \\ \frac{\partial (z'Ax)}{\partial A_{m1}} & \cdots & \frac{\partial (z'Ax)}{\partial A_{mn}} \end{pmatrix} = \begin{pmatrix} \frac{\partial (\sum_{i=1}^m \sum_{j=1}^n z_i A_{ij} x_j)}{\partial A_{11}} & \cdots & \frac{\partial (\sum_{i=1}^m \sum_{j=1}^n z_i A_{ij} x_j)}{\partial A_{1n}} \\ \vdots & \cdots & \vdots \\ \frac{\partial (\sum_{i=1}^m \sum_{j=1}^n z_i A_{ij} x_j)}{\partial A_{m1}} & \cdots & \frac{\partial (\sum_{i=1}^m \sum_{j=1}^n z_i A_{ij} x_j)}{\partial A_{mn}} \end{pmatrix} \\
&= \begin{pmatrix} z_1 x_1 & \cdots & z_1 x_n \\ \vdots & \cdots & \vdots \\ z_m x_1 & \cdots & z_m x_n \end{pmatrix} \\
&= z x'.
\end{aligned}$$

5.3 Produit de Kronecker

Définition 29 Soit $A = (a_{ij})$ est une matrice $(m \times n)$ et B une matrice $(p \times q)$, $A \otimes B$ désigne la matrice $(mp \times nq)$ admettant les éléments blocs $a_{ij}B$.

$$A \otimes B = \begin{bmatrix} a_{11}B & \cdots & a_{1n}B \\ \vdots & \vdots & \vdots \\ a_{m1}B & \cdots & a_{mn}B \end{bmatrix}$$

5.3.1 Propriétés

Les propriétés remarquables sont

1. Pour tous trois matrices A, B, C , alors $(A \otimes B) \otimes C = A \otimes (B \otimes C)$
2. Pour tous quatre matrices A, B, C et D , où A et B sont des même ordres, et C, D sont des mêmes ordres, alors

$$(A + B) \otimes (C + D) = (A \otimes C) + (A \otimes D) + (B \otimes C) + (B \otimes D).$$

3. $(A \otimes B)(C \otimes D) = (AC) \otimes (BD)$

Soit $A = (a_{ij})$ une matrice carrée d'ordre $n \times n$ avec les valeurs $\lambda_1, \lambda_2, \dots, \lambda_n$. Nous dénotons le maximum de $|\lambda_1|, |\lambda_2|, \dots, |\lambda_n|$ par $\rho(A)$, $\rho(A)$ est appelé le rayon spectral de la matrice A . Aussi, il existe un constant positif K tel que pour tout nombre entier positif m , nous avons

$$|(A^m)_{ij}| \leq K (\rho(A))^m$$

pour chaque i et j , I_n la matrice de l'identité d'ordre $n \times n$ dans le quelle chaque entrée diagonale est égale à unité et chaque entrée non diagonale est égale à zéro.

Définition 30 Soit $A = (a_1, \dots, a_n)$ est une matrice $(m \times n)$ avec $(m \times 1)$ colonnes a_i , le **Vect** opérateur transforme A dans un $(mn \times 1)$ vecteur en empilant les colonnes, c'est,

$$\text{Vec}(A) = \begin{bmatrix} a_1 \\ \vdots \\ a_n \end{bmatrix}$$

Remarque : Soit A, B, C des matrices avec dimensions appropriées.

- (a) $\text{Vec}(A + B) = \text{Vec}(A) + \text{Vec}(B)$.
- (b) $\text{Vec}(ABC) = (C' \otimes A) \text{Vec}(B)$.
- (c) $\text{Vec}(AB) = (I \otimes A) \text{Vec}(B) = (B' \otimes I) \text{Vec}(A)$.
- (d) $\text{Vec}(ABC) = (I \otimes AB) \text{Vec}(C) = (C' B' \otimes I) \text{Vec}(A)$.

Bibliographie

- [1] Akaike, H. (1974). Markovian representation of stochastic processes and its application to the analysis of autoregressive moving average processes. *Ann. Stat. Math.* 26. 363-387.
- [2] Akamanam, S. I., Rao, B. M. and Subramanyam, K. On the ergodicity of bilinear time series models. *J. Time Series Analysis*, 7, (1986) 157-163.
- [3] Andel, J. (1976). Autoregressive series with random parameters. *Math. Op. Stat.*, 7, 735-41.
- [4] Berlinet A., Francq C., Stationnarité et identification d'un processus bilinéaire strictement superdiagonal, *Statistique et Analyse des données* 15 (1990) 1-24.
- [5] Bhaskara Rao, M., Subba Rao, T. and Walker, A. M. (1983). On the existence of some bilinear time series models. *J. Time ser. Anal.*, Vol. 4, 95-110.
- [6] Bibi A. (2003). On the covariance structure of the time-varying bilinear models, *Stochastic Anal. Appl.*, 21, 25-60.
- [7] Bollerslev, T. (1986). Generalized autoregressive conditional heterpskedasticity. *Journal of Econometrics*, 31, 307-327.
- [8] Box G. E. P., Jenkins G. M., *Time Series Analysis : Forecasting and Control*, Prentice Hall : New Jersey, (1976).
- [9] Engle, R.F. (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of U.K. inflation. *Econometrica*, 50, 987-1008.
- [10] Gabr M. M., Subba Rao T., The estimation and prediction of subset bilinear time series models with applications, *Journal of Time Series Analysis* 2 (1981) 155-171.
- [11] Granger C. W. J., Andersen A. P., *Introduction to Bilinear Time Series Models*, Vandenhoeck and Ruprecht : Göttingen, (1978).
- [12] Guégan D., Etude d'un modèle non linéaire, le modèle superdiagonal d'ordre 1, *Comptes Rendus de l'Académie des Sciences de Paris, Série I*, 293 (1981) 95-98.

- [13] Guégan D., Séries chronologiques non linéaires à temps discret, *Economica* : Paris, (1994).
- [14] Gabr, M.M, and Subba Rao, T. (1984). On the identification of bilinear systems from operating records. *Int. J. Control, Vol.40, 121-128*.
- [15] Gouriéroux, C. (1997). ARCH models and financial applications. Springer-Verlag, New York.
- [16] [23] Granger, C. W. J and Andersen, A. P (1978, b). On the invertibility of bilinear time series models. *Stoch. Proc. and their appl. 8. 87-92*.
- [17] Guégan. T. D. Pham.(1987). Minimalité et inversibilité des modèles bilinéaires à temps discret. *C. R. A. S. 488. 159-162*.
- [18] Guégan, D. (1988). Modèles bilinéaires et polynômiaux de séries chronologiques : Etude probabilistes et analyse statistiques. Thèse de Doctorat, Université de Grenoble, France.
- [19] Haggan , V. and Ozaki, T. (1981). Modelling non linéaire random vibrations using an amplitude-dependent autoregressive time series models. *Biometrika*, 68, 189-196.
- [20] Hamilton, J.D. (1994). Time Series Analysis, Princeton University Press.
- [21] Hansen, L.P. (1982). "Large Sample Properties of Generalized Method of Moments Estimators," *Econometrica*, 50, 1029-1054.
- [22] Jones, R.H. (1965). An experiment in non linear prediction. *J. Appl. Meteo.* 4, 701-705.
- [23] Kim W. K., Billard L., Asymptotic properties for the first-order bilinear time series model, *Communications in Statistics. Theory and Methods* 19 (1990) 1171–1183.
- [24] Kim W. K., Basawa I. V., Estimation of the first order diagonal bilinear time series model, *Journal of Time Series Analysis* 11 (1990) 215–227.
- [25] Liu J., On the existence of a general multiple bilinear time series, *J. of Time Series Analysis* 10 (1989) 341–355.
- [26] Liu J., A note on causality and invertibility of a general bilinear time series models, *Ad. Applied Probability* 22 (1990) 247–250.
- [27] Liu J., Brockwell P. J., On the general bilinear time series model, *J. of Applied Probability* 25 (1988) 553–564.
- [28] Liu, J. and Brockwell, P. J. (1988). On the general bilinear time series model. *J. Appl. Prob* 25, 553-564.
- [29] McKenzie, E. (1985). Some simple models for discrete variate time series. In *Time series analysis in water resource*, (ed. K.W. Hipel) pp.645-650.

- [30] Mohler R. R., *Bilinear Control Processes*, Academic Press : New York, (1973).
- [31] Neudecker H., Some theorems on matrix differentiation with special reference to Kronecker matrix products, *Journal of the American Statistical Association* 64 (1969) 953–963.
- [32] Newey, W.K. and McFadden D. (1994). Large sample estimation and hypothesis testing. In *Handbook of Econometrics*, Vol. IV, Edited by R.F. Engle and D.L. McFadden. pp 2111-2245
- [33] Nicholls, D.F. and Quinn, B.G. (1982). Random coefficient autoregressive models : An introduction, lecture notes in statistics. Vol.11, Springer, N. Y.
- [34] Ogaki, M. (1992). “Generalized Method of Moments : Econometric Applications,” in G. Maddala, C. Rao, and H. Vinod (eds.), *Handbook of Statistics, Volume 11 : Econometrics, North-Holland, Amsterdam*.
- [35] Ozaki, T. (1982). The statistical analysis of perturbed limit cycle processes using nonlinear time series models. *J. Time Ser. Anal.* , 3, 29-41.
- [36] Pham, D. T. (1985). Bilinear Markovian representation and bilinear models. *Stoch. Processes and their applications*. 20, 295-306.
- [37] Tjøstheim. (1986). Estimation in non linear time series models, *Stoch. Proc. and their Appl.* 21251-273
- [38] Tong, H. and Lim, K. S (1980). threshold autoregression : limit cycles and cyclical data. *J. R. Statist. Soc. B.* 42. 245-292.
- [39] Pham D. T., Tran L.T.,(1981) On the first order bilinear time series models, *Journal of Applied Probability* 18, 617–627.
- [40] Pham, Dinh T., (1985). Bilinear Markovian representation and bilinear models. *Stoch. Processes and their appl.* 20, 295-306.
- [41] Priestley M. P., *Non-Linear and Non-Stationary Time Series Analysis*, Academic Press : London and New York, (1988).
- [42] Subba Rao T., On the theory of bilinear time series models, *Journal of the Royal Statistical Society, Series B*, 43 (1981) 244–255.
- [43] Subba Rao T., Gabr M. M., An introduction to bispectral analysis and bilinear time series models. *Lecture Notes in Statistics* No. 24, Springer Verlag, New York (1984).
- [44] Subba Rao T., Silva M. E., Identification of bilinear time series models $BL(p, 0, p, 1)$, *Statistica Sinica* 2 (1992) 464–478.

- [45] Terdik G., *Bilinear Stochastic Models and Related Problems of Nonlinear Time Series : a Frequency Domain Approach*, Lecture Notes in Statistics 124, Springer Verlag : New York, (1999).

ملخص

هذه المذكرة تعالج بعض مشاكل السلاسل الزمنية غير الخطية. القسم الخاص من النماذج غير الخطية الذي تمت دراسته بشكل متسع هو قسم النماذج الثنائية الخطية. الشروط اللازمة و الكافية للاستقرار بصفة دقيقة و من الدرجة الثانية، الانقلابية و وجود العزوم من مراتب عليا تم التطرق اليها. الهدف الرئيسي من هذه المذكرة هو التقدير في النماذج الثنائية الخطية باستعمال طريقة العزوم المعممة. انطلاقا من شروط التعامد نحصل على تقديرات ثابتة وتسلك سلوكا تقاربيا يتبع القانون الطبيعي. هذه الطريقة تعطي نتائج جد مرضية تم تبويبها في التطبيق.

بعض الكلمات المفتاحية

النماذج غير الخطية، السلاسل الزمنية الثنائية الخطية، التمثيل الشعاعي للثنائي الخطية، حاصل ضرب كرنكر، القطر الطيفي، مستقر، سببية، طاقة، الانقلابية، العزوم من مراتب عليا، طريقة العزوم المعممة، السلوك المقارب.

Résumé

Ce mémoire examine quelques problèmes de séries temporelles non linéaires. La classe particulière des modèles non linéaires qui a été extensivement discutée dans la littérature est la classe des modèles bilinéaires. Les conditions nécessaires et suffisantes pour la stationnarité au sens stricte et au second ordre, l'inversibilité et l'existence des moments d'ordre supérieur sont abordées. L'objectif principal de ce travail est l'estimation dans les modèles bilinéaires par la méthode des moments généralisés, à partir des conditions d'orthogonalité on obtient des estimateurs consistants et asymptotiquement normaux. Cette méthode donne des résultats très satisfaisants qui sont montrés dans la simulation.

Quelques mots clés

Les Modèles Non Linéaires, Bilineaires Series Temporelles, La représentation Vectorielle Bilineaire, Produit Kronecker, Rayon Spectral, Stationnarité, Causalité, Ergodicité, Inversibilité, Moments D'ordre Supérieur, Estimateur GMM, Consistance, Normalité Asymptotique.