

N° d'ordre : 08/2011- M/MT

République Algérienne Démocratique et Populaire
Ministère de l'Enseignement Supérieur et de la Recherche Scientifique
Université des Sciences et de la Technologie Houari Boumediene
Faculté de Mathématiques



MEMOIRE

Présenté pour l'obtention du diplôme de **Magister**

En : MATHEMATIQUES

Spécialité : **Recherche Opérationnelle : Méthodes Stochastiques**

Par

Sara BENDJEDDOU

Thème

**Etude de Quelques Modèles Bilinéaires
Périodiques**

Soutenu, publiquement, le 11/10/2011, devant le Jury composé de :

M ^{me} GUERBYENNE	Hafida	MCA,	U.S.T.H.B :	Présidente
M ^r BENTARZI	Mohamed	Professeur,	U.S.T.H.B :	Directeur de mémoire
M ^r BERKOUNE	Youcef	MCA,	U.M.M, Tizi Ouzou	Examineur
M ^r HAMDI	Fayçal	MCB,	U.S.T.H.B :	Invité

Table des matières

Introduction Générale	ii
1 Rappels des concepts fondamentaux	1
1.1 Introduction	1
1.2 Processus stochastiques	3
1.2.1 Introduction	3
1.2.2 Processus du second ordre	4
1.2.3 Processus bruit blanc : faible et fort	4
1.2.4 Stationnarité (stabilité)	5
1.2.5 Causalité	7
1.2.6 Régularité et singularité	7
1.2.7 Inversibilité	10
1.3 Processus de Markov	11
1.3.1 Définition et classification des processus de Markov	11
1.3.2 Propriété de Markov à espace d'états fini ou dénombrable	12
1.3.3 Propriété de Markov à espace d'états général	16
1.4 Ergodicité	23
1.4.1 Introduction	23
1.4.2 Théorème ergodique et ergodicité pour des processus stationnaires	24

1.4.3	Ergodicité pour des chaînes de Markov à espace d'états général, topologique et normé	26
1.4.4	Ergodicité géométrique	28
1.5	Représentation markovienne des modèles linéaires	29
1.5.1	Meilleure prévision linéaire	29
1.5.2	Représentation markovienne	33
1.5.3	Les équations aux différences stochastiques	35
2	Les Modèles Bilinéaires	38
2.1	Introduction	38
2.2	Etude probabiliste des modèles bilinéaires via leurs représentations markoviennes	42
2.2.1	Construction des représentations markoviennes	43
2.2.2	Condition d'existence d'une représentation markovienne	47
2.2.3	Inversibilité d'une représentation markovienne	49
2.2.4	Minimalité d'une représentation markovienne	52
2.2.5	Ergodicité géométrique et régularité absolue	55
2.2.6	Moments et cumulants	57
2.3	Etude Probabiliste des Modèles Bilinéaires à partir de l'approche directe . .	60
2.3.1	Représentation vectorielle	60
2.3.2	Existence d'une solution stationnaire stricte, causale et ergodique . .	60
2.3.3	Existence d'une solution stationnaire au second ordre	62
2.3.4	Inversibilité	64
2.3.5	Le modèle $BL(p, 0, p, 1)$	65
2.4	Illustrations	69
2.4.1	Le modèle $BL(0, 0, 2, 1)$	69

2.4.2	Le modèle $BL(1, 0, 1, 1)$	75
2.5	Comparaison entre les Modèles Bilinéaires et d'autres modèles de séries chronologiques	79
2.5.1	Comparaison selon la structure probabiliste	79
2.5.2	Comparaison selon la structure de covariance	81
2.6	Conclusion	85
3	Les modèles bilinéaires périodiques	86
3.1	Introduction	86
3.2	Processus périodiquement corrélés et les modèles bilinéaires périodiques . . .	87
3.2.1	Processus périodiquement corrélés	87
3.2.2	Processus strictement périodique	88
3.2.3	Processus périodiquement ergodique	89
3.2.4	Les modèles bilinéaires périodiques	90
3.3	Le modèle $PBL(0, 0, P, 1)_S$	90
3.3.1	Ecriture vectorielle	91
3.3.2	Stationnarité strictement périodique	94
3.3.3	Stationnarité périodique du second ordre	96
3.3.4	Structure d'autocovariance	100
3.3.5	Les moments d'ordre supérieurs	103
3.4	Illustrations	104
3.4.1	Le modèle $PBL(0, 0, 2, 1)_S$	104
3.4.2	Le modèle $PBL(1, 0, 1, 1)_S$	115
4	Conclusion et perspectives	121
5	Annexe	123

Alà barakati Ellah

Introduction Générale

D'un point de vue algébrique, la linéarité d'un modèle ou d'un système se traduit par les équations où n'interviennent que les sommes ; la non linéarité se traduit en introduisant dans les équations des multiplications entre les variables.

Dans ce sens, les premiers travaux de N. Wiener (1958) traduisent une classe très générale de modèles non linéaires : les développements en série de Volterra. Si X_t représente l'entrée du système et Y_t la sortie, de tels modèles ont pour expression

$$Y_t = \sum_{i=0}^{\infty} a_i X_{t-i} + \sum_{i=0}^{\infty} \sum_{j=0}^{\infty} a_{ij} X_{t-i} X_{t-j} + \sum_{i=0}^{\infty} \sum_{j=0}^{\infty} \sum_{k=0}^{\infty} a_{ijk} X_{t-i} X_{t-j} X_{t-k} \dots$$

Une classe de modèles, qui est en quelque sorte un cas particulier des développements de Volterra, a été alors considérée par les automaticiens : il s'agit des modèles bilinéaires qui ont la particularité d'être linéaires par rapport à chacune des variables (entrée et sortie) quand on considère l'autre constante. L'autre particularité de ces modèles est qu'ils se présentent comme une extension naturelle des modèles linéaires avec un nombre fini de paramètres. Dans ce cas les applications sont essentiellement orientées en automatique, en théorie du contrôle et en économie (étude des populations dynamiques en biologie, établissement des lois en chimie, étude de l'accroissement des économies nationales...) , on peut citer les travaux de Mohler (1970), Ruberti, Isidori, d'Alessandro (1972), Bruni, Dippilo et Koch (1974), Beghelli et Guidorzi (1978), Fliesset Normant-Cyrot (1980), Desay (1986), etc.

Puis, c'est à partir de la démarche des automaticiens, que C.W.J. Granger et A.P. Anderson (1978a) ont introduit la classe des modèles bilinéaires analogue à celle des automaticiens où l'entrée déterministe a été remplacée par une suite de variables aléatoires indépendantes équidistribuées non observées, supposées généralement de lois gaussiennes.

Dans ce cas, ces modèles ont été utilisés dans différents domaines, en particulier en économie (C.W.J.Granger et A.P.Anderson, 1978a), et en dynamiques des populations (T.Suba Rao et M.M.Gabr, 1984). En outre, en faisant une étude empirique des trajectoires, on peut penser à utiliser ces modèles pour modéliser des phénomènes physiques explosifs (étude de la vitesse du vent et les éruptions volcaniques ; c.f.Guégan 1988), et ce grâce à leur comportement qui se caractérise par de fortes explosions suivies de longues plages de calme.

A la suite de leur travail , beaucoup de chercheurs se sont intéressés à cette présentation stochastique et ont étudié les propriétés probabilistes des cas particuliers des modèles bilinéaires à cause de leur complexité structurale. On peut citer les travaux de Subba Rao (1981), Pham et Tran (1981) , Pham (1981) , Tong (1981) , Guégan (1981, 1983, 1984) , Quinn (1982) et Hannan (1982) .

L'année (1985) représente un tournant essentiel dans l'étude des modèles bilinéaires et une nouvelle orientation se dessine avec la représentation markovienne. En effet, Pham fut le pionnier de la représentation markovienne des modèles superdiagonaux. Guégan (1987) a montré que cette représentation s'étend difficilement au cas des modèles sous diagonaux mais elle a introduit, en revanche, la forme stochastique des modèles polynomiaux en entrées affines en état définis par Sontag (1979) .

En (1986), Pham a établi la représentation markovienne de l'ensemble des modèles bilinéaires. Ces représentations markoviennes permettent souvent d'aborder de façon efficace et élégante, un grand nombre de problèmes : existence d'une solution stationnaire, minimalité de la représentation, ergodicité, inversibilité et calcul des moments. On peut citer les travaux de Pham (1985, 1986) , Guégan (1986, 1987) , Guégan et Pham (1987, 1989) .

Ensuite, Liu et Brockwell (1988) et Liu (1989, 1990) ont utilisé une représentation vectorielle pour les modèles bilinéaires et ont réussi à établir la stationnarité, la causalité, l'ergodicité et l'inversibilité du modèle bilinéaire générale $BL(p, q, P, Q)$ sous des conditions simples et sans supposer nécessairement l'existence de la variance du bruit $(\varepsilon_t, t \in \mathbb{Z})$.

Ce panorama de travaux nous a permis d'observer qu'il existe plusieurs voies de recherches sur lesquelles on peut effectuer l'analyse des modèles bilinéaires dont : les représentations vectorielles, les représentations markoviennes, et en particulier les fameuses équations

tions aux différences stochastiques ayant une représentation markovienne linéaire de la forme $X_n = A_n X_{n-1} + B_n, \forall n \in \mathbb{Z}$ où (A_t, B_t) est une suite de matrices carrées aléatoires iid. En effet, cette variété nous a permis de cerner plusieurs modes de recherche permettant de découvrir des cheminements cautionnant éventuellement toutes nos questions d'une spécificité qui caractérise chacune de ces voies.

Par ailleurs, dans plusieurs situations pratiques, nous sommes devant des données générées par un certain processus non seulement non linéaire mais aussi non stationnaire, par exemple, en théorie de l'économie, la plupart des indices d'un stock d'une marchandise sont des différences de martingales (donc non nécessairement un processus i.i.d), pour de telles séries les techniques des modèles linéaires habituelles sont inapplicables. Lorsque l'économie change, il est difficile de justifier l'utilisation du même modèle (non linéaire) sur une longue période. Donc, le recours aux modèles non linéaires à coefficients dépendants du temps nous semble raisonnable.

Autres exemples de coefficients dépendant du temps qui ont un intérêt particulier sont les coefficients périodiques (pour des données saisonnières), ou les coefficients de rupture à un instant connu t_0 (pour une série qui subit un changement en un instant t_0). C'est pour cela, dans le cadre de ce travail, notre objectif est ciblé sur l'étude de quelques propriétés probabilistes des modèles bilinéaires à travers la propriété de la périodicité et ce après les avoir examiné sans la présence de cet effet. De plus, cette observation nous a permis d'organiser notre travail de recherche comme suit :

Organisation du mémoire :

Ce mémoire est composé de trois chapitres :

Le chapitre 1 se place dans une perspective assez générale où nous introduisons les outils et les concepts nécessaires pour l'analyse des modèles bilinéaires.

Le chapitre 2 est consacré essentiellement à l'étude probabiliste des modèles bilinéaires via deux approches différentes dont on a relevé les résultats de la première auprès des travaux de Pham (1985, 1986), Guégan (1987), Guégan et Pham (1987, 1992) et autres, quant à la deuxième on s'est référé aux travaux de Liu, J.(1989, 1990), Liu, J. et Brockwell, P.J.(1988).

Plus précisément, ce chapitre est composé en outre de la section de l'introduction de quatre sections dont :

- La première section présente l'étude probabiliste des modèles bilinéaires généraux via leur représentation markovienne.

- - La deuxième section traite, dans un premier temps, l'étude probabiliste de ces modèles en s'appuyant sur leur représentation vectorielle.

Parmi les propriétés probabilistes auxquelles on s'intéresse au cours de ces deux sections figure essentiellement : la construction, les conditions d'existence, l'inversibilité, la minimalité, l'ergodicité géométrique et régularité absolue, les moments et les cumulants des représentations markoviennes ainsi la stationnarité stricte, la stationnarité au second ordre, l'inversibilité, l'existence des moments d'ordre supérieur des représentations vectorielles de ces modèles.

Ensuite, afin de manipuler certains outils théoriques (stationnarité du 1^{er} et du 2^{ième} ordre asymptotique, ergodicité, inversibilité, structure d'autocovariance,...), nous étudions une classe particulière des modèles bilinéaires superdiagonaux ($BL(p, 0, p, 1)$).

- - - La troisième section met le point sur le modèle purement bilinéaire et strictement superdiagonal d'ordre un ($BL(0, 0, 2, 1)$) ainsi le modèle bilinéaire diagonal d'ordre un ($BL(1, 0, 1, 1)$) qui sont respectivement des cas particuliers de deux classes différentes connues par la classe des modèles purement bilinéaires et strictement superdiagonaux et la classe des modèles bilinéaires diagonaux. Alors, cette section est une illustration ou une application des résultats donnés auparavant pour les modèles bilinéaires généraux ou le modèle ($BL(p, 0, p, 1)$).

- - - - La quatrième section repose sur une vision comparative afin de voir la position des modèles bilinéaires par rapport à d'autres modèles de séries chronologiques linéaires tels que les modèles *ARMA* et non linéaires tels que les modèle *GARCH*.

Le chapitre 3 traite l'aspect périodique des modèles bilinéaires tout en développant quelques propriétés probabilistes de certaines classes de ces modèles non stationnaires à coefficients périodiques, dans le temps ($PBL(0, 0, P, 1)$, $PBL(p, 0, p, 1)$, $PBL(0, 0, 2, 1)$ et $PBL(1, 0, 1, 1)$). Il est composé, en outre de la section de l'introduction, de trois sections dont :

- Au cours de la première section, nous passons en revue les définitions de base telles que le concept de la corrélation périodique, la périodicité stricte ainsi que l'ergodicité périodique

qui seront très utiles par la suite.

- - La deuxième section se concentre sur l'étude probabiliste du modèle $PBL(0, 0, P, 1)$ (et parallèlement le modèle $PBL(p, 0, p, 1)$). Elle est composée de 05 phases réparties selon l'enchaînement suivant :

· La première phase répond à l'étape préalable qui précède cette étude et qui permet d'ouvrir un champs de dispersion de sophistication ou encore elle permet de tracer deux axes principaux sur lesquels on peut effectuer notre étude probabiliste. Ces deux axes sont, en effet, les deux approches connues dans la littérature de l'analyse des séries temporelles à coefficients périodiques par :

*L'approche de Gladyshev (1961).

**L'approche de Bentarzi et Hallin (1995).

Alors, dans un premier temps on explique le principe de chacune de ces approches, puis on montre que lorsqu'on s'appuie sur l'une ou l'autre on obtient toujours une équation aux différences stochastique vectorielle d'ordre un mais expliquée par des modèles différents, i.e, la première approche nous conduira à un modèle autorégressif à coefficient aléatoire (RCA) tandis que la deuxième nous ramènera à un modèle autorégressif à coefficient aléatoire périodique ($RCPAR$). Donc, sur le fait que la deuxième approche conserve la propriété de la périodicité, notre intention sera mise par la suite sur celle ci tout en développant quelques propriétés probabilistes de certains modèles bilinéaires périodiques à travers cette voie de recherche.

· La deuxième et la troisième phases établissent respectivement les conditions suffisantes d'existence des solutions strictement stationnaires périodiques et au second ordre périodique qui restent vrai pour l'ensemble de tous les modèles bilinéaires périodiques ayant une représentation de la forme $X_t = A_t X_{t-1} + B_t$ où (A_t, B_t) est une suite de matrices carrées aléatoires non nécessairement iid.

· La quatrième phase s'occupe du calcul des autocovariances du modèle $PBL(0, 0, P, 1)$ via deux procédures différentes :

*Approche basée sur l'équation aux différences.

**Approche basée sur la représentation espace d'état.

et cela tout en remettant en cause les qualités de chacune d'entre elle, i.e, en faisant ressortir la lacune rencontrée dans ce cadre de travail, qui reste un maillon manquant nécessitant une

attention particulière pour les recherches à venir.

- - - Enfin, la dernière section vient pour illustrer tous les résultats présentés tout au long de ce chapitre à l'aide de deux cas particuliers des modèles bilinéaires périodiques ; à savoir le modèle purement bilinéaire et strictement superdiagonal d'ordre un périodique ($PBL(0, 0, 2, 1)$) et diagonal d'ordre un périodique ($PBL(1, 0, 1, 1)$).

Chapitre 1

Rappels des concepts fondamentaux

1.1 Introduction

L'étude des modèles de séries temporelles comporte généralement une partie cruciale liée aux conditions d'existence de solutions stationnaires. Il n'est pas toujours simple d'obtenir des conditions suffisantes (portant sur les paramètres, la loi des erreurs,...) pour qu'un modèle admette des solutions stationnaires (au sens strict ou au second ordre).

Par exemple, pour les modèles ARMA (autorégressif moyenne mobile), les conditions portent sur les racines des polynômes retard (autorégressif et moyenne mobile) mais, la majorité des modèles de séries chronologiques acceptent des représentations markoviennes permettant d'étudier la propriété de stationnarité de ces modèles à travers leurs représentations. Ces représentations peuvent être linéaires de la forme $X_n = A_n X_{n-1} + B_n$, où les matrices A_n et B_n sont aléatoires et constituent une suite (A_n, B_n) indépendante et identiquement distribuée, ou non linéaires de la forme $X_n = F(X_{n-1}, \varepsilon_n)$, où (ε_n) est une suite de vecteurs aléatoires i.i.d et F est une application mesurable non linéaire.

Une approche du problème de la stationnarité, pour le premier type de représentations (linéaires), est fondée sur les travaux de Brandt (1986) et Bougerol et Picard (1992*a, b*) tandis que pour le deuxième type de représentations (non linéaires), une telle approche est fondée sur la théorie des chaînes de Markov à espace d'états continu, exposée dans le livre de Meyn et Tweedie (1996).

Dans le présent chapitre, en guise d'introduction, nous énonçons certaines notions et certains résultats qui s'avèrent nécessaires à la compréhension des chapitres suivants.

Dans la première et la deuxième section, nous déclinons quelques outils proportionnels aux processus stochastiques et aux processus de Markov respectivement en prenant comme particularité celui de l'ergodicité, qui fait l'objet d'étude de la troisième section de ce chapitre.

Dans la dernière section, nous définissons, en utilisant le théorème de projection, certains concepts qui permettent d'établir la représentation markovienne des modèles linéaires et/ou bilinéaires; à savoir le concept d'espérance conditionnelle, meilleure prévision en moyenne quadratique et meilleure prévision linéaire. Ensuite, nous présentons certains résultats relatifs à l'étude probabiliste non seulement des modèles linéaires ou bilinéaires mais de la quasi totalité des modèles de séries chronologiques, et cela via les fameuses équations aux différences stochastiques ayant une représentation markovienne linéaire.

1.2 Processus stochastiques

1.2.1 Introduction

L'analyse d'une série chronologique consiste à trouver un modèle mathématique adéquat décrivant le mécanisme ayant donné lieu à cette série temporelle dont elle a pour but la description des principales propriétés du processus générateur de cette série. Formellement, cette série ce n'est autre que la trajectoire de ce processus comme le montre la définition suivante :

Définition 1.2.1

Un processus stochastique est une application de S "l'espace des réalisations" dans un espace de fonctions de variables réelles (temps)

$$X_t : S \rightarrow F$$

$$s \mapsto \{X_t(s), t \in T\}$$

C'est-à-dire pour chaque point s de l'espace des échantillons S , on fait correspondre une fonction du temps $X_t(s)$ appelée "la trajectoire du processus", où les observations successives forment l'histoire du processus.

Une autre définition équivalente, qui vise à présenter les phénomènes aléatoires évoluant dans le temps, est donnée par

Définition 1.2.2

*Un processus stochastique est une collection de variables aléatoires indexées $X_t = \{X_t, t \in T\}$, où :
 · T est le domaine d'évolution défini sur un espace de probabilité (Ω, \mathcal{A}, P) et à valeurs dans un espace d'état E .*

· L'espace d'état est l'espace sur lequel toutes les variables aléatoires prennent leurs valeurs, le plus souvent $(\mathbb{R}, \mathcal{B}_{\mathbb{R}})$ où $\mathcal{B}_{\mathbb{R}}$ est la tribu de Borel.

· X_t est une variable aléatoire pour chaque t , où t peut appartenir à un domaine d'évolution discret ($T = \mathbb{Z}, T = \mathbb{N}$) ou continu ($T = \mathbb{R}$), ce qui définit un processus à temps discret et un processus à temps continu respectivement.

Par ailleurs, lorsque chacune des variables X_t vérifie $E(X_t) < \infty$, la loi du processus est partiellement résumée par l'espérance des différentes variables et par leurs covariances. Si de plus $E(X_t^2) < \infty$, le processus devient un processus du second ordre, défini explicitement

dans ce qui suit.

1.2.2 Processus du second ordre

Un processus aléatoire $\{X_t, t \in \mathbb{Z}\}$ est dit du second ordre (ou encore de carré intégrable) si et seulement si ses deux premiers moments existent et sont finis, i.e,

$$E(X_t) < \infty \text{ et } E|X_t|^2 < \infty, \forall t \in \mathbb{Z}.$$

Ainsi, la fonction d'autocovariance d'un processus stochastique du second ordre est définie par :

$$\gamma(t, s) = Cov(X_t, X_s) = E((X_t - \mu_t)(X_s - \mu_s)), \forall t, s \in \mathbb{Z},$$

où μ_t est la moyenne du processus à l'instant t .

Par ailleurs, le processus du second ordre le plus simple, le plus utilisable en analyse des séries chronologiques est le processus bruit blanc. Ces processus ont donné un mouvement important et remarquable à l'analyse des séries chronologiques stationnaires (depuis 1938 grâce au théorème introduit par *Wold*, concernant l'existence et l'unicité d'une décomposition linéaire d'un processus du second ordre faiblement stationnaire) et non stationnaires (depuis 1961 lorsque *Cramer* a généralisé le résultat du *Wold* au cas non stationnaires, appelé "décomposition de *Wold-Cramer*"). De plus, grâce à cette décomposition que la classe des modèles linéaires à coefficients évolutifs dans le temps a été introduite.

1.2.3 Processus bruit blanc : faible et fort

Un bruit blanc $\{\varepsilon_t\}$ est une suite de variables aléatoires représentant une série de référence grâce à la transformation spécifique qui lui caractérise, car trouver les transformations qui décrivent une série donnée comme une transformation d'un bruit blanc est un enjeu de base dans l'analyse des séries temporelles. De plus, du fait que le bruit blanc peut ramener l'étude à des variables aléatoires non corrélées et même souvent indépendantes, on peut distinguer deux types de bruit dont :

▷ Le processus $(\varepsilon_t, t \in \mathbb{Z})$ est un bruit blanc fort si ce processus est iid et $E(\varepsilon_t) = 0, \forall t$. De plus, si $\varepsilon_t \sim N(0, h), \forall t$, on parle de bruit blanc gaussien (standard si $h = 1$).

▷ Le processus $(\varepsilon_t, t \in \mathbb{Z})$ est un bruit blanc faible si les ε_t sont identiquement distribués et

non corrélés où la propriété de la non corrélation veut dire que les fonctions d'autocovariance et d'autocorrélation sont respectivement données par :

$$\cdot \gamma(h) = \begin{cases} \sigma^2, & \text{si } h = 0, \\ 0, & \text{ailleurs.} \end{cases}$$

$$\cdot \rho_h = \begin{cases} 1, & \text{si } h = 0, \\ 0, & \text{ailleurs.} \end{cases}$$

Dans le cadre de ce mémoire, on utilise la notion du bruit blanc fort lorsque l'on parle de bruit blanc.

Remarque 1.2.1

Il semble important de rappeler que deux variables non corrélées ne sont pas toujours indépendantes :

$X \perp\!\!\!\perp Y \implies Cov(X, Y) = 0$ mais la réciproque est fausse.

Cependant, on obtient l'équivalence dans le cas de variables aléatoires gaussiennes c'est-à-dire : $X \perp\!\!\!\perp Y \Leftrightarrow Cov(X, Y) = 0$.

Comme on a mentionné dans l'introduction, la loi du processus est partiellement résumée par l'espérance des différentes variables et par leurs covariances, lorsque chacune des variables X_t vérifie $E(X_t) < \infty$. En effet, ces moments dépendent en général du temps, ce qui est gênant, quand à partir de l'observation de réalisations du processus on veut tirer de l'information sur la loi sous-jacente de ce dernier. Donc, pour pouvoir obtenir une accumulation d'information on est amené à considérer des processus dits stationnaires.

1.2.4 Stationnarité (stabilité)

Intuitivement, un processus stochastique $(X_t, t \in \mathbb{Z})$ est stationnaire si les propriétés probabilistes du processus ne changent pas au cours du temps. Pourtant on distingue deux types de stationnarité, la stationnarité au second ordre (ou la stationnarité faible ou encore la stationnarité au sens large) et la stationnarité stricte (ou la stationnarité forte) .

Définition 1.2.3

Un processus $(X_t, t \in \mathbb{Z})$ est dit stationnaire au second ordre, si $(X_t, t \in \mathbb{Z})$ est un processus du second ordre dont la moyenne et la fonction d'autocovariance sont indépendantes du temps.

Ce type de stationnarité est aussi une propriété d'invariance ou de stabilité des deux pre-

miers moments par translation dans le temps, i.e : $E(X_t) = E(X_s)$ et $V(X_t) = V(X_s)$ pour $t \neq s$. Mais on ne dit rien sur les moments d'ordre supérieur (autrement dit que ce type de stationnarité n'empêche pas une variation des moments d'ordre plus élevés ce qui explique l'existence de l'asymétrie de la loi ou l'épaisseur des queues fonctions du temps), ce qui fait que cette définition, très commode par ailleurs, est sans doute trop floue.

Au lieu de considérer simplement les deux premiers moments d'un processus, on peut décider de s'intéresser à la distribution complète des observations. On peut alors accéder à une définition plus stricte de la stationnarité.

Définition 1.2.4

Un processus stochastique est dit strictement stationnaire si la distribution jointe de X_t et X_{t+h} ne dépend pas de t , mais seulement de h . Autrement dit pour tout k appartenant à \mathbb{N} , toute famille d'instantes t_1, t_2, \dots, t_k appartenant à \mathbb{Z} et tout $n \in \mathbb{N}$ les lois jointes de $(X_{t_1}, X_{t_2}, \dots, X_{t_k})$ et de $(X_{t_1+n}, X_{t_2+n}, \dots, X_{t_k+n})$ sont les mêmes.

Ici, c'est la distribution jointe qui est invariante par translation. Cette propriété est plus forte que la précédente car un processus stationnaire au second ordre peut posséder des moments d'ordre supérieur qui ne sont pas invariants par translation. La notion de stationnarité stricte et de stationnarité au second ordre se confondent par contre pour les processus Gaussiens car ceux-ci sont entièrement résumés par leurs deux premiers moments. Par contre dès que l'on sort du cadre Gaussien comme dans les modèles GARCH, on n'a plus de coïncidence entre les deux notions. Il est cependant courant de se contenter de la stationnarité au second ordre, car elle est plus facile à décrire.

Remarque 1.2.2

1) La notion de stabilité peut être défini de la manière suivante :

Définition 1.2.5 (stabilité)

Etant donné un processus $(X_t, t \in \mathbb{Z})$, celui ci est stable si pour une suite de variables aléatoires réelles de $L^2(\Omega, \mathcal{A}, \mathcal{P})$, il existe deux constantes m, M positives vérifiant

$$E|X_t| \leq m, \text{Var}(X_t) \leq M, \forall t \in \mathbb{Z}.$$

Ainsi, pour la classe des processus stationnaires et ergodiques, on dispose du théorème suivant (Breiman, 1968) :

Théorème 1.2.1 (stabilité par translation mesurable)

Soit $(X_t, t \in \mathbb{Z})$ un processus stationnaire et ergodique et f une fonction mesurable définie de $\mathbb{R}^{\mathbb{Z}}$ dans \mathbb{R} , alors le processus $(Z_t, t \in \mathbb{Z})$ défini par $Z_t = f(\dots, X_{t-1}, X_t, X_{t+1}, \dots)$, $t \in \mathbb{Z}$ est un processus stationnaire et ergodique.

2) Si l'on considère la dépendance temporelle, la stationnarité du second ordre suppose uniquement une stabilité de la corrélation (moment d'ordre 2) : $cov(X_t, X_{t+h}) = cov(X_s, X_{s+h})$. La stationnarité au sens fort est beaucoup plus forte que cette définition sur le moment d'ordre 2, puisqu'elle suppose une stabilité de toutes les lois jointes : en particulier, cette condition implique l'égalité en loi des couples (X_t, X_{t+h}) et (X_s, X_{s+h}) .

3) Il est très important de rappeler que, si X_1 et X_2 sont deux variables aléatoires de même loi et de même pour Y_1 et Y_2 , telles que $cov(X_1, Y_1) = cov(X_2, Y_2)$, alors on n'a pas égalité des lois jointes : $\mathcal{L}(X_1, Y_1) \neq \mathcal{L}(X_2, Y_2)$. De plus, si X et Y suivent des lois normales, la loi du couple n'est pas nécessairement de loi gaussienne.

1.2.5 Causalité

Intuitivement, un processus stochastique est dit causal si on peut exprimer les X_t uniquement en fonction des $\varepsilon_t, \varepsilon_{t-1}, \dots$

Plus précisément on a la définition suivante

Définition 1.2.6

Un processus stochastique $(X_t, t \in \mathbb{Z})$ est dit causal s'il existe une fonction mesurable $f : \mathbb{R}^{\infty} \rightarrow \mathbb{R}$ telle que $X_t = f(\varepsilon_t, \varepsilon_{t-1}, \dots)$ presque sûrement, pour tout $t \in \mathbb{Z}$.

1.2.6 Régularité et singularité

Dans cette partie, on note par

· $\mathcal{F}(X)$ la σ -algèbre engendrée par $(X_t)_{t \in \mathbb{Z}}$,

· $\mathcal{F}_t(X)$ la σ -algèbre engendrée par $\{X_s, s \leq t\}$,

· $\mathcal{F}_{-\infty}(X) = \bigcap_{t \in \mathbb{Z}} \mathcal{F}_t(X)$, $\mathcal{F}_{+\infty}(X) = \bigvee_{t \in \mathbb{Z}} \mathcal{F}_t(X)$,

· $\mathcal{M}(X)$ l'espace vectoriel fermé engendré par $(X_t)_{t \in \mathbb{Z}}$. Tout élément de $\mathcal{M}(X)$ est alors, ou

bien combinaison linéaire d'un certain nombre fini des X_t avec des coefficients constants ou non, ou bien limite en moyenne quadratique d'une telle combinaison.

· $\mathcal{M}_t(X)$ sous espace fermé linéaire de $\mathcal{M}(X)$ engendré par $\{X_s, s \leq t\}$, c'est un sous espace de Hilbert de $\mathbb{L}^2(\Omega, \mathcal{A}, \mathcal{P})$.

$$\cdot \mathcal{M}_{-\infty}(X) = \bigcap_{t \in \mathbb{Z}} \mathcal{M}_t(X), \mathcal{M}_{+\infty}(X) = \bigvee_{t \in \mathbb{Z}} \mathcal{M}_t(X),$$

· $H_n(x)$ désigne le polynôme de Hermite de degré n défini par

$$H_n(x) = (-1)^n \exp\left\{\frac{x^2}{2}\right\} \frac{d^n}{dx^n} \exp\left\{-\frac{x^2}{2}\right\}$$

Avec ces notations, la définition de ces deux concepts est donnée par

Définition 1.2.7

Etant donné un processus $(X_t)_{t \in \mathbb{Z}}$ stable, alors

1. *Le processus $(X_t)_{t \in \mathbb{Z}}$ est dit régulier linéaire si $\mathcal{M}_{-\infty}(X) = \{0\}$, et dit régulier non linéaire, si $\mathcal{F}_{-\infty}(X) = \{\emptyset, \Omega\}$.*

2. *Le processus $(X_t)_{t \in \mathbb{Z}}$ est dit singulier linéaire si $\mathcal{M}_{-\infty}(X) = \mathcal{M}_{+\infty}(X) = \mathcal{M}(X)$, et dit singulier non linéaire, si $\mathcal{F}_{-\infty}(X) = \bigvee_{t \in \mathbb{Z}} \mathcal{F}_t(X)$.*

Les relations entre ces deux notions, nous les donnons dans la proposition suivante

Proposition 1.2.2

1. *Tout processus $(X_t)_{t \in \mathbb{Z}}$ stable, régulier non linéaire est régulier linéaire.*
2. *Tout processus $(X_t)_{t \in \mathbb{Z}}$ stable, singulier linéaire est singulier non linéaire.*

Remarque 1.2.3

Cette proposition signifie clairement que la régularité non linéaire est plus forte que la régularité linéaire, et que la singularité linéaire est plus forte que la singularité non linéaire, mais l'inverse n'est pas vrai (cf. Guégan (1983)). Dans le cas linéaire, on dispose de critères, de conditions nécessaires et suffisantes permettant de dire qu'un processus est régulier, citons néanmoins la condition de Kolmogorov-Szego où la condition porte sur la continuité de la densité spectrale évolutive, ou qu'il possède une décomposition de Wold-Cramer dans la base de ses innovations. Dans le cas non linéaire, on peut se demander si des résultats analogues peuvent être obtenus. En général, il n'est pas possible dans le cas non linéaire (cf. Guégan (1981)) de disposer de critères permettant de dire qu'un processus est régulier. On doit donc travailler cas par cas.

Si, par exemple, on considère le processus $(X_t)_{t \in \mathbb{Z}}$ défini par : $X_t = b_{21}X_{t-2}\varepsilon_{t-1} + \varepsilon_t$, où $(\varepsilon_t)_{t \in \mathbb{Z}}$ est un processus de bruit blanc fort (i.e une suite de variables aléatoires centrées et indépendantes), avec $\text{var} \{\varepsilon_t\} = \sigma^2$, on suppose de plus que la condition d'inversibilité de ce modèle est vérifiée (i.e $b_{21}^2\sigma^2 < 1$), on peut vérifier que le processus $(X_t)_{t \in \mathbb{Z}}$ est un processus de bruit blanc faible, il est donc régulier linéaire, il est aussi régulier non linéaire car pour tout $t \in \mathbb{Z}$ la variable aléatoire X_t peut s'exprimer comme combinaison linéaire des polynômes de Hermite en ε_t . Ceci implique que $\mathcal{F}_{-\infty}(X) \subset \mathcal{F}_{-\infty}(\varepsilon) = \{\emptyset, \Omega\}$.

1.2.7 Inversibilité

Le concept d'inversibilité joue un rôle fondamental dans l'analyse des séries temporelles et dans les applications statistiques telles que : prévisions, estimation des paramètres, évaluation de la fonction de vraisemblance, certains tests...

Etant donné un processus stochastique $X_t, t \in \mathbb{Z}$, celui-ci est inversible s'il est possible d'estimer la suite des variables aléatoires inobservables $\{\varepsilon_t\}$, ce qui revient à écrire ε_t en fonction du passé et du présent de $\{X_s, s \leq t\}$, telle que :

$$\hat{\varepsilon}_n - \varepsilon_n \rightarrow 0 \text{ en un certain sens, quand } n \rightarrow \infty.$$

En outre, selon les types de convergence, on peut tirer trois formes d'inversibilité ; à savoir l'inversibilité au sens faible par la convergence en probabilité, l'inversibilité au sens fort par la convergence presque sûre (autrement dit que si ε_t est $\mathcal{F}_t(X)$ -mesurable pour tout $t \in \mathbb{Z}$ alors le processus $(X_t, t \in \mathbb{Z})$ est fortement inversible) et l'inversibilité au sens de Granger et Anderson par la convergence en moyenne quadratique. Plus précisément, la suite $\hat{\varepsilon}_n$ est une fonction de X_1, \dots, X_t uniquement pour les deux premières définitions (inversibilité faible et inversibilité forte) quant à la dernière la suite $\hat{\varepsilon}_n$ ne dépend pas uniquement de X_1, \dots, X_t comme le montre la définition suivante :

Définition 1.2.8 (Granger et Anderson, 1978)

Soit X_t un paramètre de série chronologique discret satisfaisant le modèle

$$X_t = f(X_{t-j}, \varepsilon_{t-j}, j = 1, 2, \dots, p) + \varepsilon_t, \quad (1.2.1)$$

où $\{\varepsilon_t\}_{t \in \mathbb{Z}}$ est une suite de variables aléatoires indépendantes inobservables.

En commençant par quelques valeurs initiales

$$\hat{\varepsilon}_0 = \hat{\varepsilon}_{-1} = \dots = \hat{\varepsilon}_{-p} = 0, \text{ et } x_0 = x_{-1} = \dots = x_{-p} = 0,$$

on obtient une suite d'estimateur $\hat{\varepsilon}_t$ défini par l'équation aux différences suivante

$$\hat{\varepsilon}_t = X_t - f(X_{t-j}, \hat{\varepsilon}_{t-j}, j = 1, 2, \dots, p),$$

où $\hat{\varepsilon}_{t-j} = 0$ pour $t < j$. Ainsi, ce modèle est inversible si $\lim_{t \rightarrow \infty} E(\hat{\varepsilon}_t - \varepsilon_t)^2 \rightarrow 0$.

1.3 Processus de Markov

Outre la représentation de Wold-cramer, les processus de Markov se considèrent comme une classe de processus stochastiques d'une importance majeure qui a donnée une nouvelle orientation dans les recherches dès son invention due aux Andreï A. Markov.

Cette partie a comme but de présenter quelques résultats basiques sur la théorie des processus de Markov ; à savoir les chaînes de Markov à espace d'états fini (faisant leur première apparition dans un article publié en 1906), les chaînes de Markov à temps continu et les chaînes de Markov avec un nombre infini d'états (généralisations du premier concept par Kolmogorov en 1931 et 1936 respectivement).

1.3.1 Définition et classification des processus de Markov

Un processus stochastique X_t est dit processus de Markov si pour tout sous-ensemble fini d'instants $\{t_i, i = 1, \dots, M\}$ tels que $t_1 < t_2 < \dots < t_M$, la distribution conditionnelle de X_{t_M} connaissant $X_{t_1}, X_{t_2}, \dots, X_{t_{M-1}}$ dépend seulement de $X_{t_{M-1}}$:

$$P \{X_{t_M} = x_M / X_{t_1} = x_1, \dots, X_{t_{M-1}} = x_{M-1}\} = P \{X_{t_M} = x_M / X_{t_{M-1}} = x_{M-1}\}.$$

Ce qui signifie que connaissant le "présent" du processus, le "futur" est indépendant du "passé". Cette propriété est connue sous le nom **propriété de markov**. L'ensemble \mathcal{X} des valeurs possibles du processus X_t est appelé **l'espace d'états** de ce processus.

Suivant la nature de l'espace d'états ainsi que de l'espace des temps sur lequel est défini le paramètre "t", les processus de Markov sont divisés en quatre classes, figure 1.3.1

Nature du paramètre t	Espace d'états	
	Discret	Continu
Discret	Chaîne de M à temps discret	Processus de M à temps discret
Continu	Chaîne de M à temps continu	Processus de M à temps continu

Fig. 1.3.1 Classification des processus de Markov

Remarque 1.3.1

La définition d'un processus de Markov peut être généralisée en définissant les processus de Markov d'ordre "r" en considérant que l'état d'un processus X_t à un instant donné t_M ne dépend que des "r" états précédents :

$$P \{X_{t_M} = x_M / X_{t_1} = x_1, \dots, X_{t_{M-1}} = x_{M-1}\} = P \{X_{t_M} = x_M / X_{t_{M-1}} = x_{M-1}, \dots, X_{t_{M-r}} = x_{M-r}\}.$$

Ici la propriété de Markov correspond à l'idée que l'on souhaite résumer l'information contenue dans les variables passées du processus par un nombre fini de variables (les variables d'états). Dans le cas le plus simple, on souhaite que les variables d'états soient des valeurs retardées du processus : toute l'information est contenue dans les r valeurs les plus récentes. Par conséquent, la définition déjà donnée d'un processus de Markov concerne les processus de Markov d'ordre 1. Cette généralisation est plaisante et présente souvent plus de réalisme que la seule définition des processus de Markov d'ordre 1. Néanmoins, il est important de noter que le nombre de paramètres nécessaires pour la description de tels processus croît exponentiellement avec la mémoire " r " du processus. Ceci explique pourquoi on est souvent amené, en pratique, à se limiter aux processus de Markov d'ordre 1.

Afin de saisir intellectuellement les règles suivant lesquelles on formule et on raisonne dans le cadre de la théorie des chaînes de Markov, on s'intéresse d'abord à développer des résultats en s'appuyant sur un espace d'états dénombrable et puis on donne l'analogie pour des processus à espace d'états continu sachant qu'un espace d'états est dit :

- ▷ Dénombrable s'il est discret avec un nombre d'éléments fini ou infini dénombrable,
- ▷ Général s'il est muni d'une tribu (σ -algèbre) générée d'une manière dénombrable, notée désormais \mathcal{B} .

1.3.2 Propriété de Markov à espace d'états fini ou dénombrable

Une chaîne de Markov à temps discret est une séquence X_1, X_2, X_3, \dots de variables aléatoires. L'ensemble de leurs valeurs possibles est appelé l'espace d'états, la valeur X_t étant l'état du processus à l'instant t . Dans toute cette partie, nous travaillons avec un espace d'états discret "fini ou infini dénombrable" donné dans la plupart des cas par :

$$1^\circ. E = \{1, 2, \dots, K\},$$

$$2^\circ. E = \{0, 1, 2, \dots, K\},$$

$$3^\circ. E = \mathbb{N} = \{0, 1, 2, \dots\},$$

$$4^\circ. E = \mathbb{Z} = \{\dots, -1, 0, 1, \dots\}.$$

Définition 1.3.1 (chaîne de Markov)

Un processus stochastique $\{X_t, t \in \mathbb{N}\}$ défini sur un espace de probabilité (Ω, P) à valeurs dans un espace d'états E (fini ou infini dénombrable) est dit chaîne de Markov homogène s'il vérifie les propriétés suivantes :

1) Pour tout $i, j, i_0, \dots, i_{t-1} \in E$:

$$P(X_{t+1} = j / X_t = i, X_{t-1} = i_{t-1}, \dots, X_0 = i_0) = P(X_{t+1} = j / X_t = i) = P_{ij}(t).$$

2) L'évolution de cette chaîne ne dépend pas de l'instant t , mais seulement des états concernés, i.e, $P_{ij}(t) = P_{ij} \forall t$.

Remarque 1.3.2

1°. La condition 1) s'appelle la propriété de Markov.

2°. La condition 2) s'appelle propriété d'homogénéité dans le temps.

3°. Les probabilités P_{ij} s'appellent les probabilités de transition de l'état i vers l'état j , ainsi ces probabilités permettent de construire la matrice de Markov donnée par la définition suivante :

Dfinition 1.3.2 (matrice de transition)

On appelle matrice des probabilités de transition la matrice carrée $P = (P_{ij})_{i,j \in E}$ telle que :

$$\cdot \forall i, j \in E : 0 < P_{ij} < 1,$$

$$\cdot \sum_{j \in E} P_{ij} = 1 \forall i \in E.$$

A noter que la distribution de X_0 joue un rôle très important, notamment pour simuler les chaînes de Markov, d'où on a la définition suivante :

Définition 1.3.3 (Distribution initiale d'une chaîne de Markov)

Soit $\{X_t, t \in \mathbb{N}\}$ une chaîne de markov sur E , la distribution de probabilité $(p_i, i \in E)$ définie par $P(X_0 = i) = p_i$, s'appelle la loi initiale.

Cette distribution ainsi que la matrice des probabilités de transition contribuent à déterminer la structure de probabilité de la chaîne $\{X_t, t \in \mathbb{N}\}$ caractérisée par ses distributions fini-dimensionnelles données par le corollaire suivant :

Corrolaire 1.3.1

Les distributions fini-dimensionnelles de la chaîne de Markov $\{X_t, t \in \mathbb{N}\}$ admettant la loi initiale $P_{i_0} = P(X_0 = i_0)$ sont données par :

$$\begin{aligned} P(X_n = i_n, X_{n-1} = i_{n-1}, \dots, X_0 = i_0) \\ = P(X_n = i_n / X_{n-1} = i_{n-1}) \times \dots \times P(X_1 = i_1 / X_0 = i_0) \times P(X_0 = i_0), n \in \mathbb{N}, \forall i_0, i_1, \dots, i_n \in E. \end{aligned}$$

Probabilités de transition d'ordre k (Equation de Chapman Kolmogorov)

La probabilité de transition d'ordre k , i.e, après k étapes, est donnée par

$$\begin{aligned} P_{ij}^{(0)} &= \begin{cases} 1, & \text{si } i = j, \\ 0, & \text{sinon} \end{cases} \\ P_{ij}^{(1)} &= P_{ij}, \\ P_{ij}^{(k)} &= P(X_{t+k} = j / X_t = i). \end{aligned}$$

Pour calculer les probabilités de transition d'ordre k , on calcule d'abord la probabilité de transition d'ordre 2, $P_{ij}^{(2)}$. En utilisant la loi de probabilité totale, on voit que

$$\begin{aligned} P_{ij}^{(2)} &= P(X_{t+2} = j / X_t = i) = P((X_{t+2} = j) \cap \Omega / X_t = i) \\ &= P\left((X_{t+2} = j) \cap \bigcup_{k \in E} (X_{t+1} = k) / X_t = i\right) \\ &= P\left(\bigcup_{k \in E} (X_{t+2} = j) \cap (X_{t+1} = k) / X_t = i\right) \\ &\stackrel{\text{les événements sont disjoints}}{=} \sum_{k \in E} P(X_{t+2} = j / X_{t+1} = k) P(X_{t+1} = k / X_t = i) \\ &= \sum_{k \in E} P_{kj} P_{ik} = P_{ik} P_{kj} \end{aligned}$$

Par conséquent, les équations de Chapman Kolmogorov sont données par la relation suivante :

$$P_{ij}^{n+m} = \sum_{k \in E} P_{ik}^{(m)} P_{kj}^{(n)}.$$

Distribution stationnaire et distribution limite

Après avoir déterminé les distributions conditionnelles, on pense généralement à étudier l'existence de ses distributions limites autrement dit on cherche à vérifier la convergence d'une distribution limite vers une distribution stationnaire. En outre, cette limite dépend de certaines propriétés structurelles de la chaîne (irréductibilité, apériodicité, récurrence et positivité) comme le montrent les théorèmes suivants :

Théorème 1.3.2 (ergodicité ou régularité)

Soit $\{X_t, t \in \mathbb{N}\}$ une chaîne de Markov défini sur un espace de probabilité (Ω, \mathcal{A}, P) à valeurs dans un espace d'états E discret. Si cette chaîne est :

- irréductible,
- apériodique,
- récurrente positif.

Alors, elle est ergodique. Dans ce cas, l'unique distribution stationnaire est confondue avec la distribution limite.

Théorème 1.3.3 (stabilité)

Pour toute chaîne de Markov ergodique, on a $\lim_{n \rightarrow \infty} P_{ij}^{(n)}$ existe et ne dépend pas de l'état initial i dans lequel se trouve la chaîne, i.e, elle dépend seulement de l'état d'arrivée j . Ainsi, on note : $\lim_{n \rightarrow \infty} P_{ij}^{(n)} = \pi_j$.

A noter que le vecteur des probabilités (π_0, π_1, \dots) est appelé loi stationnaire. Pour déterminer cette loi, on commence d'abord par écrire

$$P_{ij}^{(n+1)} = \sum_{k \geq 0} P_{ik}^{(n)} P_{kj}$$

et en utilisant le fait que

$$\lim_{n \rightarrow \infty} P_{ij}^{(n)} = \lim_{n \rightarrow \infty} P_{ij}^{(n+1)} = \pi_j.$$

Il vient que

$$\pi_j = \sum_{k \geq 0} \pi_k P_{kj}, \text{ avec : } \sum_{k \geq 0} \pi_k = 1.$$

Ces deux dernières équations peuvent s'écrire sous la forme d'un produit matriciel liant la loi stationnaire $\pi = (\pi_0, \pi_1, \dots)$ à la matrice de transition $P = (P_{ij})$:

$$\begin{cases} \pi = \pi P \\ \text{et} \\ \sum_{k \geq 0} \pi_k = 1 \end{cases}$$

donc, la résolution du système d'équations ci-dessus fournit la loi stationnaire $\pi = (\pi_0, \pi_1, \dots)$.

1.3.3 Propriété de Markov à espace d'états général

Avant, la distribution du processus était donnée en fonction des probabilités conditionnelles expliquant le passage du processus d'un état à un autre, connues par probabilités de transition, mais dorénavant ces dernières vont être remplacées par des mesures de probabilités que le processus se trouve dans un ensemble quelconque appartenant à l'espace \mathcal{X} muni d'une tribu \mathcal{B} , connues par noyau de transition. Si de plus cette tribu est la tribu de Borel, l'espace d'états est un espace topologique et si elle est générée d'une manière dénombrable, il s'agit d'un espace général. En effet, la relation entre les deux, outre le fait que l'espace topologique est une spécification d'un espace général, est que le cas général nous fournisse des résultats pour des ensembles plutôt que pour des états individuels (lorsqu'il y a une topologie).

Définition 1.3.4 (noyau de transition)

On appelle noyau de probabilité de transition ou fonction de transition de Markov toute famille $\mathcal{P} = \{P(x, B), x \in \mathcal{X}, B \in \mathcal{B}\}$ telle que :

1. *Pour tout $B \in \mathcal{B}$, $P(\cdot, B)$ est une fonction mesurable non négative sur \mathcal{X} ,*
2. *Pour tout $x \in \mathcal{X}$, $P(x, \cdot)$ est une mesure de probabilité sur \mathcal{B} .*

Les $P(x, B)$ sont appelées probabilités de transition (de x vers B).

Par analogie au cas où l'espace d'états est dénombrable, i.e en remplaçant la matrice de transition par le noyau de transition, la structure de probabilité de la chaîne est toujours caractérisée par sa loi initiale et son noyau de transition; alors étant donnée une mesure de probabilité initiale μ sur \mathcal{B} , les distributions fini dimensionnelles sont données par le théorème suivant :

Théorème 1.3.4 (Meyn et Tweedie, 1993)

Pour toute mesure initiale μ sur \mathcal{B} et pour n'importe quel noyau de transition \mathcal{P} , il existe un processus stochastique $X = \{X_0, X_1, \dots\}$ mesurable défini sur Ω l'espace $X^{\mathbb{N}}$, $\left(\prod_{i=0}^{\infty} X_i\right)$, muni de la tribu \mathcal{F} , le produit $B^{\otimes \mathbb{N}}$, et une mesure de probabilité \mathbb{P}_μ sur \mathcal{F} telle que $\mathbb{P}_\mu(B)$ est la probabilité de l'évènement $(X \in B)$, où $B \in \mathcal{F}$, donnée pour tout entier n et tout $(n + 1)$ -uplet (B_0, \dots, B_n) de parties de B par

$$\mathbb{P}_\mu(X_0 \in B_0, \dots, X_n \in B_n) = \int_{x_0 \in B_0} \dots \int_{x_{n-1} \in B_{n-1}} \mu(dx_0) P(x_0, dx_1) \dots P(x_{n-1}, B_n) \quad (1.3.1)$$

A l'aide de ces instruments, on peut énoncer la définition suivante

Définition 1.3.5 (chaîne de Markov homogène)

Le processus stochastique X défini sur (Ω, \mathcal{F}) est appelé chaîne de Markov homogène de noyau de transition \mathcal{P} et de mesure initiale μ si les lois fini-dimensionnelles satisfont (1.3.1) pour tout n .

Concernant le passage des lois fini-dimensionnelles aux probabilités d'évènements, on dispose du résultat établi par Chapman-Kolmogorov imposant des conditions sur l'espace d'états $(\mathcal{X}, \mathcal{B})$ qui doit appartenir à l'un des cas suivants :

- ▷ $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$, qui est général,
- ▷ un espace dénombrable muni de la tribu de toutes ses parties, qui est dénombrable,
- ▷ un espace métrique complet possédant une base dénombrable d'ouverts muni de sa tribu borélienne, qui est topologique.

Noyau de transition d'ordre supérieur "Equation de Chapman Kolmogorov"

Soit n l'ordre supérieur et supposant qu'à l'instant initial la chaîne est à l'état x et elle va atteindre un autre état B en n étapes, alors cette chaîne doit nécessairement passer par un autre état y à la $m^{i\grave{e}me}$ étape (autrement dit qu'elle sera à l'état y dans l'étape m , telle que $m \leq n$). Alors passer de x à B en n étapes est équivalent de passer du y à B en $n - m$ étapes, ce qui présente le principe général de l'équation de Chapman Kolmogorov. Le déroulement de cette dernière est comme suit :

En prenant comme mesure initiale, la mesure de Dirac définie par :

$$P^0(x, B) = \delta_x(B), \delta_x(B) = \begin{cases} 1 & \text{si } x \in B, \\ 0 & \text{sinon} \end{cases},$$

Alors, le noyau de transition relatif à l'ordre n , noté par P^n , en effectuant le produit de ce noyau $(n - 1)$ fois avec lui même, est donné par :

$$P^n(x, B) = \int_{y \in \mathcal{X}} P(x, dy) P^{n-1}(y, B), \forall x \in \mathcal{X}, \forall B \in \mathcal{B}, \forall n \geq 1.$$

Par conséquent, l'équation de Chapman Kolmogorov peut être écrite par une simple généralisation de cette dernière et elle est donnée par

$$P^n(x, B) = \int_{y \in \mathcal{X}} P^m(x, dy) P^{n-m}(y, B), \forall x \in \mathcal{X}, \forall B \in \mathcal{B}, \forall 0 \leq m \leq n \text{ et } \forall n \geq 1.$$

Cette dernière signifie également que les valeurs prises antérieurement en y n'ont plus d'importance pour ce qui est postérieur à l'étape m ce qui explique la propriété de l'indépendance entre futur et passé conditionnellement au présent exprimée par Meyn et Tweedie en 1993 par la propriété suivante

Propriété 1.3.6

Si X est une chaîne de Markov sur (Ω, \mathcal{F}) , avec une mesure initiale μ , et $h : \Omega \rightarrow \mathbb{R}$ une fonction mesurable bornée, alors

$$\mathbb{E}(h(X_{n+1}, X_{n+2}, \dots) / X_0, \dots, X_n; X_n = x) = \mathbb{E}(h(X_1, X_2, \dots) / X_0 = x).$$

Classification des ensembles

Définissons pour tout ensemble $B \in \mathcal{B}$:

* Le temps de séjour dans B (après la date initiale) / le temps moyen de séjour :

$$\eta_B = \sum_{n=1}^{\infty} 1_{X_n \in B}, \text{ où } 1 \text{ désigne la variable indicatrice d'un évènement } A;$$

$$U(x, B) = E(\eta_B / X_0 = x) = \sum_{n=1}^{\infty} P^n(x, B), \text{ où la famille } \{U(x, B)\} \text{ a des propriétés simi-}$$

laires à celles d'un noyau de transition mais les $U(x, \cdot)$ ne sont pas des probabilités.

* Le temps de premier retour à B (hormis l'état initial) :

$$\tau_B = \min \{n \geq 1 : X_n \in B\}.$$

* La probabilité de premier retour à B partant de x :

$$L(x, B) = \mathbb{P}(\tau_B < \infty / X_0 = x).$$

Alors, la classification repose sur la famille $\{U(x, B)\}$, $U(x, B)$ représente le temps de séjour moyen dans B partant de x , dont l'ensemble B est dit :

▷ récurrent si $U(x, B) = +\infty, \forall x \in B$;

▷ uniformément transitoire si $\exists M < \infty$ tq $U(x, B) \leq M, \forall x \in B$;

▷ transitoire s'il existe une famille dénombrable (B_n) d'ensembles uniformément transitoires telle que $B = \bigcup_n B_n$.

Avec cette définition, un ensemble récurrent peut également être transitoire. Par conséquent, l'ensemble \mathcal{X} est toujours récurrent et s'il est dénombrable, on peut évidemment l'écrire comme la réunion de tous les singletons uniformément transitoires.

Remarque 1.3.3

Pour tout $B \in \mathcal{B}$, η_B et τ_B sont clairement des fonctions mesurables de Ω à $\mathbb{Z}_+ \cup \{\infty\}$.

Une autre définition d'importance majeur et qui nous fournisse la plupart des propriétés par la suite est celle d'un ensemble petit, introduite par Nummelin et Tuominen, 1982, p.188.

Définition 1.3.7 (ensemble petit)

un ensemble A est dit petit si A est relativement compact et $\varphi(A) > 0$.

Ainsi, une définition équivalente à cette dernière a été proposée par Meyn et Tweedie, 1993, est donnée par :

Définition 1.3.8

Un ensemble C est dit un ensemble petit s'il existe $m > 0$, et une mesure non triviale ν_m sur \mathcal{B} , tel que $\forall x \in C, B \in \mathcal{B}$,

$$\mathbf{P}^m(x, B) \geq \nu_m(B). \tag{1.3.2}$$

Lorsque (1.3.2) est vérifiée, on dit que C est ν_m small.

Formes de Stabilité

* Etant donnée une chaîne de Markov $X = (X_t, t \in \mathbb{N})$, la forme la moins restrictive de stabilité qu'on peut exiger est que la chaîne ne consiste pas, en réalité, de deux chaînes : la collection des ensembles qu'on puisse atteindre à partir de différents points initiaux n'est pas différente. Ce qui nous conduit d'abord à introduire la notion de l'irréductibilité.

Définition 1.3.9 (φ – irréductibilité)

Soit φ une mesure non identiquement nulle définie sur \mathcal{B} . On dit que φ est une mesure d'irréductibilité si $\varphi(B) > 0 \Rightarrow L(x, B) > 0, \forall x \in \mathcal{X}, \forall B \in \mathcal{B}$.

On dit alors que la chaîne $X = (X_t, t \in \mathbb{N})$ est φ – irréductible.

Maintenant, si cette chaîne est φ – irréductible, la mesure d'irréductibilité est générée comme

une mesure d'irréductibilité maximale qui se considère comme son extension du fait qu'elle définit le rang de la chaîne plus convenablement qu'une autre mesure d'irréductibilité arbitraire qu'on pouvait construire initialement.

Définition 1.3.10 (irréductibilité maximale)

Si X est φ – irréductible pour une certaine mesure φ , alors il existe une mesure de probabilité ψ sur \mathcal{B} telle que :

1. X est ψ – irréductible;
2. Pour toute autre mesure φ' sur \mathcal{B} , X est φ' – irréductible si et seulement si φ' est absolument continue par rapport à ψ ($\psi > \varphi'$);

Si $\psi(B) = 0$, alors $\psi(\bar{B}) = 0$, où $\bar{B} = \{x : L(x, B) > 0\}$.

La mesure ψ est alors appelée mesure d'irréductibilité maximale.

Nous noterons pour une chaîne d'irréductibilité maximale ψ :

$$B^+ = \{B \in \mathcal{B} : \psi(B) > 0\}.$$

Définition 1.3.11 (périodicité)

On dit que le noyau de transition \mathcal{P} , φ – irréductible, est périodique s'il existe un entier $d \geq 2$ et une suite $\{E_1, \dots, E_d\}$ de d événements non vides t.q pour $i = 0, \dots, d-1$ et $x \in E_i$, $P(x, E_j) = 1$ pour $j = i + 1 \pmod{d}$. Sinon, on dit que P est apériodique.

** Le deuxième niveau de stabilité est l'exigence, non seulement qu'il devrait être une possibilité d'atteindre les mêmes états à partir de points de départ différents, mais que atteindre tels ensembles d'états devrait être finalement garanti. Ceci nous conduit à introduire les concepts suivants :

Définition 1.3.12 (chaînes récurrentes/ transitoires)

Soit X une chaîne φ – irréductible. Elle est dite :

· Récurrente $\Leftrightarrow U(x, B) = +\infty, \forall x \in \mathcal{X}, \forall B \in \mathcal{B}^+$,

· Transitoire \Leftrightarrow l'espace d'états est un ensemble transitoire, i.e : $\exists (B_j)_j, \mathcal{X} = \bigcup_j B_j, U(x, B_j) \leq$

$M_j < +\infty, \forall x \in \mathcal{X}$.

Théorème 1.3.5 (Meyn et Tweedie, 1993)

Si X est une chaîne φ – irréductible, elle est soit récurrente soit transitoire.

Un autre concept de récurrence un peu plus fort fournissant des résultats également plus forts est donné par celui du Harris-récurrence.

Définition 1.3.13 (Harris-récurrence)

Une chaîne X , φ – irréductible, est dite Harris-récurrente si chaque ensemble de \mathcal{B}^+ est Harris récurrent tel que :

$$Q(x, B) = P(\eta_B = +\infty / X_0 = x) = 1, x \in B.$$

De cette définition, il est clair que si un ensemble est Harris récurrent, il est évidemment récurrent car dans le cas de récurrence, le nombre de visite d'un ensemble est espéré être infini quant au deuxième cas le nombre de visite de cet ensemble est presque sûrement infini expliqué autrement par l'existence d'un renforcement du cas récurrent au cas Harris récurrent.

On peut de manière équivalente définir la Harris-récurrence pour chaque ensemble de \mathcal{B}^+ ou bien en fonction des probabilités de premier retour

$$L(x, B) = P(\tau_B < \infty / X_0 = x) = 1, x \in B.$$

ou bien en fonction des ensembles petits comme suit :

Théorème 1.3.6 (Meyn et Tweedie, 1993)

Si X est une chaîne φ – irréductible. S'il existe un ensemble petit C tel que

$$\forall x \in \mathcal{X}, L(x, C) = 1 \text{ alors } (X_n, n \in \mathbb{N}) \text{ est harris-récurrente.}$$

***Au cours d'une variation persistante, la forme maximale de stabilité qu'on peut exiger est que la distribution de X_n ne change pas lorsque n prend différentes valeurs. Ceci est expliqué, par la propriété de Markov, par l'invariance, sous une translation dans le temps, des distributions fini-dimensionnelles de cette chaîne. Une telle considération nous conduit à introduire le concept des mesures invariantes et sous invariantes.

Définition 1.3.14 (mesure invariante / sous invariante)

Une mesure non nulle σ – finie π sur B est dite invariante (resp. sous-invariante) si

$$\pi(B) = \int P(x, B) \pi(dx) \text{ (resp. } \pi(B) \geq \int P(x, B) \pi(dx)), \forall B \in \mathcal{B}.$$

De plus,

· pour une chaîne de Markov φ – irréductible admettant une mesure de probabilité invariante π est dite "positive" et dans le cas contraire elle est dite "nulle", et

· pour une chaîne récurrente, une mesure invariante peut ou non exister. Ces dernières ont une importance majeure pour produire des processus stationnaires en considérant uniquement une forme de stationnarité dès la première étape :

Etant donnée une mesure de probabilité initiale invariante μ , i.e :

$$\mu(B) = \int_{\mathcal{X}} \mu(dw) P(w, B) = P_\mu(X_1 \in B),$$

ce qui devient par des remplacements successifs :

$$\begin{aligned} \mu(B) &= \int_{\mathcal{X}} \left(\int_{\mathcal{X}} P(x, dw) \mu(dx) \right) P(x, B) = \int_{\mathcal{X}} \mu(dx) \int_{\mathcal{X}} P(x, dw) P(x, B) \\ &= \int_{\mathcal{X}} \mu(dx) P^2(x, B) = P_\mu(X_2 \in B) \\ &\vdots \\ &= \int_{\mathcal{X}} \mu(dx) P^n(x, B) = P_\mu(X_n \in B). \end{aligned}$$

De la propriété de Markov, il est clair que la chaîne est strictement stationnaire. Inversement, elle est strictement stationnaire si et seulement si sa loi de probabilité initiale est invariante.

Une autre raison de l'importance des mesures de probabilité invariantes est liée au comportement ergodique (ou de long terme) de la chaîne. Supposons en effet qu'il existe une mesure de probabilité limite π , définie par : $\forall B \in \mathcal{B}, P(X_n \in B) \rightarrow \pi(B)$ quand $n \rightarrow \infty$. Alors

$$\begin{aligned} \pi(B) &= \lim_{n \rightarrow +\infty} \int \mu(dx) P^n(x, B) \\ &= \lim_{n \rightarrow +\infty} \int \mu(dx) \int P^{n-1}(x, dy) P(y, B) \\ &= \int P(y, B) \lim_{n \rightarrow +\infty} \int P^{n-1}(x, dy) \mu(dx) \\ &= \int P(y, B) \pi(dy) \end{aligned}$$

Donc π est invariante. En particulier s'il existe une unique mesure de probabilité invariante, la probabilité limite sera indépendante de μ . On a le résultat suivant

Théorème 1.3.7 (Meyn et Tweedie, 1993)

Si X est récurrente, elle admet une unique mesure sous-invariante π (à une constante multiplicative près) et cette mesure est invariante. De plus π est équivalente à toute mesure d'irréductibilité maximale ψ .

1.4 Ergodicité

1.4.1 Introduction

L'hypothèse ergodique est une hypothèse fondamentale de la physique statistique. Elle s'appliquait aux systèmes composés d'un très grand nombre de particules, et affirmait qu'à l'équilibre, la valeur moyenne d'une grandeur calculée de manière statistique est égale à la moyenne d'un très grand nombre de mesures prises dans le temps. Elle est donc fondamentale pour un bon rapprochement entre la théorie et l'expérience, et elle intervient également pour le traitement du signal, où elle consiste à admettre que l'évolution d'un signal au cours du temps apporte la même information qu'un ensemble de réalisations. Un exemple concret qui facilite la compréhension est le suivant : si, en premier lieu, on prend un dé et on le jette N fois en l'air, et en second lieu, on prend N dés identiques et on les jette en l'air tous en même temps ; dont le but est de vouloir calculer la moyenne de la face qui sort : l'hypothèse ergodique permet de dire que les deux expériences fournissent le même résultat, c'est-à-dire le calcul de la moyenne au fur et à mesure sur les N dés ; et cela en supposant que les dés soient suffisamment séparés les uns des autres lors du jet, pour qu'ils ne s'interchoquent pas. Cette propriété fondamentale ne peut être achevée que si la liaison entre les moyennes statistiques et les moyennes temporelles existe, donc on doit appeler une autre hypothèse connue par l'hypothèse d'ergodicité.

La notion d'ergodicité est une notion fondamentale dans l'étude des séries chronologiques de variables aléatoires dépendantes afin de pouvoir utiliser des théorèmes du type de la loi des grands nombres. En particulier, elle est importante dans l'étude des processus stationnaires, des chaînes de Markov et des mélanges. Du fait que les chaînes de Markov et les systèmes mélangeants présentent deux exemples de systèmes dynamiques mesurés, il en sera judicieux d'éclaircir ces notions.

Un système dynamique est un quadruplet $(\Omega, \mathcal{F}, \mu, \theta)$ où $(\Omega, \mathcal{F}, \mu)$ est un espace mesuré et θ est une application mesurable de (Ω, \mathcal{F}) qui préserve la mesure μ , c'est-à-dire que : $\forall A \in \mathcal{F} : \mu(\theta^{-1}(A)) = \mu(A)$. On dit aussi que la mesure μ est invariante par θ ainsi s'il est le cas, on a presque sûrement pour tout n l'invariance de μ par θ^n . Alors, ce système est ergodique si la tribu \mathcal{I} , la famille des événements invariants, est triviale sous μ , c'est-à-dire

que : $\forall A \in \mathcal{I}, \mu(A) \in \{0, 1\}$. Ce système est dit mélangeant si pour tous évènements A et B appartenant à \mathcal{F} , on a : $\mu(A \cap \theta^{-n}(B)) \rightarrow \mu(A)\mu(B)$, et si oui, il est forcément ergodique car l'invariance d'un ensemble A peut être exprimé aussi, pour tout n , par $1_A = 1_A \circ \theta^n$ presque sûrement, donc $1_A 1_{\theta^{-n}(A)} = 1_A^2 = 1_A$, presque sûrement, d'où $\mu(A \cap \theta^{-n}(A)) = \mu(A)$. En faisant tendre n vers $+\infty$, on obtient $\mu(A) = \mu^2(A)$, d'où $\mu(A) \in \{0, 1\}$, ce qui explique le fait que tout système mélangeant est ergodique.

1.4.2 Théorème ergodique et ergodicité pour des processus stationnaires

La notion de l'ergodicité pour les processus stationnaires exprime le fait que le processus peut prendre n'importe quelle valeur dans l'espace des états indépendamment de sa valeur initiale c'est pour cela qu'elle est expliquée autrement par la propriété d'irréductibilité. Du fait que la propriété d'ergodicité lie les moyennes statistiques (effectuées sur l'espace des réalisations sous-jacent à la définition des variables aléatoires qui constituent le processus) et les moyennes temporelles (effectuées sur les fonctions du temps qui sont les réalisations du processus), cette propriété est très importante pour l'établissement du théorème ergodique.

Intuitivement, un processus aléatoire est ergodique si ses moments peuvent être obtenus comme des moyennes à partir d'une seule de ses réalisations. Ceci doit être vrai en particulier pour les moments d'ordre 1 et 2 :

$$\frac{1}{T} \int_T x(t, w) dt \xrightarrow{T \rightarrow \infty} m_x(t)$$

$$\frac{1}{T} \int_T x(t, w) x(t + \tau, w) dt \xrightarrow{T \rightarrow \infty} R_x(t, t + \tau)$$

Si l'on examine la première équation, on se rend compte que le membre gauche ne dépend pas du temps, et donc pour que cette équation puisse être vérifiée, il faut que le processus ait une moyenne constante. De la même façon, pour que la deuxième équation soit possible le processus doit être stationnaire au sens large, on aura alors

$$\frac{1}{T} \int_T x(t, w) dt \xrightarrow{T \rightarrow \infty} m_x$$

$$\frac{1}{T} \int_T x(t, w) x(t + \tau, w) dt \xrightarrow{T \rightarrow \infty} R_x(\tau)$$

On peut donc affirmer que pour qu'un processus soit ergodique, il doit nécessairement être

stationnaire (au sens large) : ergodicité \Rightarrow stationnarité.

Formellement, ce théorème est énoncé par :

Théorème 1.4.1 (*Karlin et Taylor, 1975*)

Soit $(X_t, t \in \mathbb{Z})$ un processus strictement stationnaire et ergodique avec une moyenne finie μ . Alors

$$P \left(\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{k=1}^n X_k = \mu \right) = 1$$

Remarquons que la notion d'ergodicité, qui implique à son tour le théorème ergodique, est intimement liée à la propriété de convergence c'est-à-dire on peut distinguer deux types d'ergodicité ; à savoir l'ergodicité forte par la convergence presque sûre et l'ergodicité faible par la convergence en probabilité ou en moyenne quadratique.

Alors, si le processus est strictement stationnaire, la notion d'ergodicité forte est donnée par :

Définition 1.4.1 (*Ergodicité forte*)

Un processus strictement stationnaire est dit ergodique si $P((\dots, X_{-1}, X_0, X_1, \dots) \in A) = 0$ ou 1 pour chaque ensemble A \mathcal{T} -invariant.

Où, \mathcal{T} est une transformation $\mathcal{T} : \mathbb{R}^{\mathbb{Z}} \rightarrow \mathbb{R}^{\mathbb{Z}}$ et le sous ensemble $A \subset \mathbb{R}^{\mathbb{Z}}$ est dit \mathcal{T} -invariant si $\mathcal{T}A = A$.

De même, si le processus est stationnaire au second ordre, on a la définition suivante

Définition 1.4.2 (*Ergodicité faible*)

Soit $(X_t, t \in \mathbb{Z})$ un processus stationnaire au sens faible, tel que $E(X_t) = \mu < \infty$ et $\text{var}(X_t) = \sigma^2 < \infty$ pour tout t . Soit $\bar{X}_t = \frac{1}{n} \sum_{t=1}^n X_t$ la moyenne temporelle. Si \bar{X}_t converge en probabilité vers μ quand $n \rightarrow \infty$, alors (X_t) est ergodique pour la moyenne.

1.4.3 Ergodicité pour des chaînes de Markov à espace d'états général, topologique et normé

Dans ce qui suit, on donne des conditions suffisantes assurant l'ergodicité des chaînes de Markov à espace d'états général ainsi deux autres spécifications : espace topologique et espace normé. Etant donnée une chaîne de Markov X , φ – irréductible, prenant ses valeurs sur un espace arbitraire $(\mathcal{X}, \mathcal{F})$ avec des probabilités de transition homogènes $P^n(x, A)$. Soit μ une mesure sous-invariante de X (pour mieux comprendre, μ devient π lorsque la chaîne est ergodique), sur $\mathcal{F}_\mu = \{A \in \mathcal{F} : 0 < \mu(A) < \infty\}$.

► Conditions sur un espace d'états général

Théorème 1.4.2 (Tweedie, 1975)

Une condition suffisante pour que $(X_t, t \in \mathbb{N})$ soit ergodique est l'existence d'un ensemble $K \in \mathcal{F}_\mu$ et une fonction mesurable non négative g sur \mathcal{X} tels que

$$\int_{\mathcal{X}} P(x, dy) g(y) \leq g(x) - 1, x \notin K,$$

et pour un entier $B > 0$,

$$\int_{\mathcal{X}} P(x, dy) g(y) = \lambda(x) \leq B < \infty, x \in K.$$

Si de plus la chaîne est apériodique, on dispose du résultat suivant :

Théorème 1.4.3 (Meyn et Tweedie, 1996)

Si $(X_t, t \in \mathbb{N})$ est une chaîne de Markov Harris récurrente, positive de mesure invariante π et apériodique, alors pour toute mesure initiale μ

$$\left\| \int \mu(dx) P^n(x, \cdot) - \pi \right\| \rightarrow 0, n \rightarrow \infty$$

► Conditions sur un espace d'états topologique

Dans la plupart des applications des chaînes de Markov, l'espace d'états n'est pas complètement général, mais muni avec une topologie \mathcal{I} . Intuitivement, cette dernière est une structure englobant les propriétés de l'invariance d'un ensemble de fonctions, après une déformation continue ce qui justifie qu'on doit imposer une condition de continuité sur les fonctions de transition de la chaîne. Mathématiquement, c'est l'ensemble de tous les sous-ensembles ouverts de \mathcal{X} , c'est-à-dire c'est une collection des ensembles contenant :

- (i). Des unions arbitraires des membres de \mathcal{I} ,
- (ii). Des intersections finies des membres de \mathcal{I} ,

·(iii). L'ensemble complet \mathcal{X} et l'ensemble vide ϕ .

Les membres contenant un point x sont appelés "les proximités de x ", et les compléments des ensembles ouverts sont appelés "fermés". L'imposition de cette continuité assure, en effet, que les ensembles compacts ont de mesures finies dont la compacité présente l'analogie de la finitude d'un espace dénombrable (espace d'états fini).

On dit qu'une loi de transition $\{P(x, \cdot)\}$ est fortement continue si, pour tout $A \in \mathcal{F}$, $P(x, \cdot)$ est une fonction continue en x , autrement dit qu'elle est fortement continue si $\int P(x, dy) g(y)$ est une fonction continue bornée en x lorsque g est une fonction continue bornée sur \mathcal{X} . Cette forme de continuité est exprimée en appelant la chaîne $\{X_t\}$ par chaîne de Feller.

Alors, lorsque \mathcal{X} est l'ensemble englobant les ensembles compacts d'une mesure positive φ , les conditions de l'ergodicité sont données par le théorème suivant :

Théorème 1.4.3 (Tweedie, 1975)

Soit $\{X_t\}$ une chaîne de Markov φ – irréductible sur un espace topologique $(\mathcal{X}, \mathcal{I})$. Si $\{P(x, \cdot)\}$ est fortement continue, une condition suffisante pour que $\{X_t\}$ soit ergodique est l'existence d'un ensemble $K \in \mathcal{X}$ et une fonction mesurable non négative g sur \mathcal{X} tels que

$$(i) \int_{\mathcal{X}} P(x, dy) g(y) \leq g(x) - 1, x \notin K;$$

$$(ii) \int_{\mathcal{X}} P(x, dy) g(y) = \lambda(x) \leq B < \infty, x \in K, \text{ pour un ensemble } B > 0.$$

► **Conditions sur un espace normé**

Supposant maintenant que l'espace \mathcal{X} est muni d'une norme, noté par $\|\cdot\|$, \mathcal{I} est une topologie générée par la norme et \mathcal{F} est la tribu de Borel sur $(\mathcal{X}, \mathcal{I})$. Dans ce cas, sous les conditions du théorème précédent, la fonction g peut être soit $g(x) = \|x\|$ ou $g(x) = \|x\|/c$ pour une certaine constante c . En posant $\gamma_x = E\{\|X_{n+1}\| - \|X_n\| / X_n = x\}$, on a le résultat suivant :

Théorème 1.4.4 (Tweedie, 1975)

Soit $(X_t, t \in \mathbb{N})$ une chaîne de Markov φ – irréductible avec une probabilité de transition fortement continue, une condition suffisante d'ergodicité pour X_t est qu'il existe un compact K sur \mathcal{F} avec $\varphi(K) > 0$, et une constante $c > 0$ telle que :

$$\gamma_x \leq -c, \forall x \notin K,$$

$$\text{et pour un ensemble } B > 0, \gamma_x \leq B < \infty, \forall x \in K.$$

1.4.4 Ergodicité géométrique

Comme pour l'ergodicité, l'ergodicité géométrique est une notion importante dans le cas de variables dépendantes mais elle est de plus une notion propre aux processus markoviens dont elle englobe la condition la plus forte que l'on puisse espérer obtenir pour ce type de processus car elle englobe différentes propriétés comme Harris récurrente, ergodicité et mélange. La définition de cette propriété, liée aux chaînes de Markov φ -irréductibles et faiblement de Feller (i.e, pour chaque fonction continue bornée g sur E , la fonction de x définie par $E(g(X_t)/X_{t-1} = x)$ est aussi continue), est bien donnée par le théorème suivant :

Théorème 1.4.5 (Tweedie, 1983)

Supposons que $X_t, t \geq 0$ soit une **chaîne de Feller**, qu'il existe une mesure φ et un ensemble compact A avec $\varphi(A) > 0$ tel que :

- (i) $X_t, t \geq 0$ soit φ -irréductible,
- (ii) il existe une fonction positive continue $g : E \rightarrow R$ satisfaisant $g(x) \geq 1 \forall x \in A$, et pour $\delta > 0 : E(g(X_t)/X_{t-1} = x) \leq (1 - \delta)g(x), x \in A^c$.

Alors $X_t, t \geq 0$ est géométriquement ergodique.

Et si de plus cette chaîne est récurrente positive et apériodique, on dispose du résultat suivant

Propriété 1.4.3 (Meyn et Tweedie, 1993)

On suppose que pour une chaîne $(X_t, t \in \mathbb{N})$ récurrente positive, apériodique et ψ -irréductible, il existe un ensemble petit $C \in \mathcal{B}^+$ et les constantes $M_C < \infty, \rho_C < 1$ tels que

$$|P^n(x, C) - \pi(C)| \leq M_C \rho_C^n, \text{ pour tout } n \text{ et tout } x \in \mathcal{X}.$$

Alors, il existe $\rho < 1$ et $M_s < \infty$ pour tout $x \in \mathcal{X}$ tels que

$$\|P^n(x, \cdot) - \pi\| \leq M_s \rho^n, n \geq 1, \tag{1.4.1}$$

où π est la mesure de probabilité invariante.

On dit que la chaîne $(X_t, t \in \mathbb{N})$ est géométriquement ergodique si (1.4.1) est vérifiée.

Remarque 1.4.1

Certains processus ne vérifient pas la propriété markovienne, autrement dit que ces processus ne sont pas des processus de Markov donc le recours à une représentation markovienne semble important pour vérifier la propriété de l'ergodicité géométrique de ces processus. Dans le cas des processus bilinéaires, la classe des processus bilinéaires superdiagonaux ne sont

pas des processus de Markov car : le futur et le passé sont conditionnellement dépendants, étant donné le présent ; et si on prend le processus bilinéaire superdiagonal le plus simple défini par le modèle suivant : $X_t = b_{21}X_{t-2}\varepsilon_{t-1} + \varepsilon_t$, il devient clair que ce processus est non markovien donc il n'est pas géométriquement ergodique.

1.5 Représentation markovienne des modèles linéaires

Du fait que la représentation markovienne bilinéaire est une généralisation de la représentation markovienne des modèles linéaires en remplaçant les matrices de constantes de cette dernière par des matrices polynomiales en bruits blancs, cette section est une étape préliminaire à la compréhension de la représentation markovienne bilinéaire, et cela en introduisant quelques concepts de base liés à la représentation markovienne linéaire.

1.5.1 Meilleure prévision linéaire

Dans cette section, nous énonçons quelques définitions et théorèmes qui proviennent essentiellement de Brockwell et Davis (1991). Ici, ν dénote un espace vectoriel sur le corps \mathcal{K} (le corps des réels \mathbb{R} ou des complexes \mathbb{C}) et \mathcal{H} indique l'espace de Hilbert.

Définition 1.5.1

Un sous espace vectoriel \mathcal{M} d'un espace \mathcal{H} est un **sous-espace vectoriel fermé** de \mathcal{H} si \mathcal{M} contient tous ses points limites (c-à-d. si $x_n \in \mathcal{M}$ et $\|x_n - x\| \rightarrow 0$ ceci implique que $x \in \mathcal{M}$).

Théorème 1.5.1

Soit \mathcal{M} un sous-espace vectoriel fermé non-vide de \mathcal{H} . Alors pour tout $x \in \mathcal{H}$, il existe un unique $\hat{x} \in \mathcal{M}$ tel que

$$\|x - \hat{x}\| = \inf_{y \in \mathcal{M}} \|x - y\|.$$

L'élément \hat{x} est appelé la projection (orthogonale) de x sur \mathcal{M} , qu'on note $P_{\mathcal{M}}x$; aussi, nous interprétons $\hat{x} = P_{\mathcal{M}}x$ comme la meilleure prévision de x dans \mathcal{M} .

Théorème 1.5.2

Soit \mathcal{M} un sous - espace vectoriel fermé non-vide de \mathcal{H} ; alors $P_{\mathcal{M}}x$ vérifie l'équation

$$\langle x - P_{\mathcal{M}}x, y \rangle = 0 \text{ pour tout } y \in \mathcal{M}.$$

Dorénavant, cette équation sera appelée l'équation de prévision. Cette équation s'avère très utile puisque c'est elle qui détermine l'expression de la projection $P_{\mathcal{M}}x$.

Définition 1.5.2

Soit un espace de probabilité (Ω, F, \mathbb{P}) , nous disons que X appartient à l'espace

$$">\mathbb{L}^2 = \mathbb{L}^2(\Omega, F, \mathbb{P})"$$
 si $E[X^2] = \int_{\Omega} X(w)^2 \mathbb{P}(dw) < \infty$.

\mathbb{L}^2 est un espace vectoriel dont la norme est induite par le produit scalaire

$$\langle X, Y \rangle = E[XY], \quad \|X\| = (E[X^2])^{\frac{1}{2}}.$$

\mathbb{L}^2 est complet (si on identifie des variables presque sûrement égales) et par conséquent \mathbb{L}^2 est un espace de Hilbert.

Définition 1.5.3

Si \mathcal{M} est un sous - espace vectoriel fermé de \mathbb{L}^2 et $X \in \mathbb{L}^2$, alors la meilleure prévision en moyenne quadratique de X dans \mathcal{M} est l'unique élément $\hat{X} \in \mathcal{M}$ qui vérifie :

$$\|X - \hat{X}\|^2 = \inf_{W \in \mathcal{M}} \|X - W\|^2 = \inf_{W \in \mathcal{M}} E|X - W|^2.$$

Le théorème de projection identifie donc immédiatement \hat{X} , l'unique meilleure prévision en moyenne quadratique de X dans \mathcal{M} , comme la projection $P_{\mathcal{M}}X$. Dans \mathbb{L}^2 , $P_{\mathcal{M}}X$ vérifie l'équation de prévision : $E[WP_{\mathcal{M}}X] = E[WX]$ pour tout $W \in \mathcal{M}$.

Il existe bien entendu d'autres définitions possibles pour la "meilleure" prévision, mais pour les processus avec des moments d'ordre deux finis, la définition de "meilleure prévision en moyenne quadratique" nous permet d'avoir une théorie sur les prévisions qui est simple, élégante et très utile dans la pratique. Lorsque nous sommes plus spécifiques concernant le sous - espace fermé \mathcal{M} (i.e, en précisant la forme qui peut prendre les v.a de cet espace), la définition (1.5.3) nous amène aux concepts d'espérance conditionnelle et de meilleure prévision linéaire.

Définition 1.5.4

Soit Z_1, \dots, Z_n des variables aléatoires appartenant à \mathbb{L}^2 , $\mathcal{M}(Z_1, \dots, Z_n)$ le sous espace vectoriel fermé de \mathbb{L}^2 composé de toutes les variables de \mathbb{L}^2 de la forme $\phi(Z_1, \dots, Z_n)$ pour une fonction de Borel quelconque $\phi : \mathbb{R}^n \rightarrow \mathbb{R}$, et $X \in \mathbb{L}^2$.

Soit $E_{\mathcal{M}(Z_1, \dots, Z_n)}[X]$ l'espérance conditionnelle de X dans $\mathcal{M}(Z_1, \dots, Z_n)$, c-à-d

$$E_{\mathcal{M}(Z_1, \dots, Z_n)}[X] = E[X/\mathcal{M}(Z_1, \dots, Z_n)] = E[X/Z_1, \dots, Z_n].$$

La meilleure prévision en moyenne quadratique de X dans $\mathcal{M}(Z_1, \dots, Z_n)$ est l'espérance conditionnelle $E_{\mathcal{M}(Z_1, \dots, Z_n)}[X]$.

Selon le théorème de projection et la définition de meilleure prévision en moyenne quadratique, la projection de X sur $\mathcal{M}(Z_1, \dots, Z_n)$, $P_{\mathcal{M}(Z_1, \dots, Z_n)}(X)$, correspond à $E_{\mathcal{M}(Z_1, \dots, Z_n)}[X]$.

L'expression de $E_{\mathcal{M}(Z_1, \dots, Z_n)}[X]$ vérifie bien l'équation de prévision :

$$E\left[w E_{\mathcal{M}(Z_1, \dots, Z_n)}[X]\right] = E[w X], \text{ pour tout } w \in \mathcal{M}(Z_1, \dots, Z_n).$$

Définition 1.5.5

Soit Y_1, \dots, Y_n des variables aléatoires sur \mathbb{L}^2 . "L'enveloppe linéaire" des variables $1, Y_1, \dots, Y_n$ est le sous - espace vectoriel fermé de \mathbb{L}^2 composé de l'ensemble de toutes les combinaisons linéaires $V = \lambda_0 + \lambda_1 Y_1 + \dots + \lambda_n Y_n$, ou $\lambda_0, \dots, \lambda_n \in \mathcal{K}$.

Nous notons cette enveloppe par $s\bar{p}\{1, Y_1, \dots, Y_n\}$.

Définition 1.5.6

La meilleure prévision linéaire de X en fonction de $1, Z_1, \dots, Z_n$ est la meilleure prévision en moyenne quadratique de X dans $s\bar{p}\{1, Z_1, \dots, Z_n\}$, $P_{s\bar{p}\{1, Z_1, \dots, Z_n\}}(X)$.

Remarque 1.5.1

Lorsque X est de moyenne nulle alors : $P_{s\bar{p}\{1, Z_1, \dots, Z_n\}}(X) = P_{s\bar{p}\{Z_1, \dots, Z_n\}}(X)$

Avec la meilleure prévision linéaire nous ne pouvons jamais avoir une erreur en moyenne quadratique inférieure à celle obtenue avec "L'espérance conditionnelle, $E_{\mathcal{M}(Z_1, \dots, Z_n)}[X]$,

c - à - d nous n'avons jamais $\|X - P_{s\bar{p}\{1, Z_1, \dots, Z_n\}}(X)\|^2 < \|X - P_{\mathcal{M}\{Z_1, \dots, Z_n\}}(X)\|^2$, ceci est évidemment dû au fait que $s\bar{p}\{1, Z_1, \dots, Z_n\} \subseteq \mathcal{M}(Z_1, \dots, Z_n)$.

Néanmoins, la meilleure prévision linéaire joue un rôle d'une très grande importance pour

les raisons énumérées ci dessous :

(a) Elle est plus facile à calculer que $E_{\mathcal{M}(Z_1, \dots, Z_n)} [X]$, puisque la détermination des projections sur $\mathcal{M}(Z_1, \dots, Z_n)$ est habituellement très difficile, étant donné la nature complexe de l'équation de prévision pour le sous-espace $\mathcal{M}(Z_1, \dots, Z_n)$.

(b) Elle dépend seulement des moments d'ordre un et deux, $E[x]$, $E[z_i]$, $E[z_i z_j]$ et $E[x z_j]$ de la distribution conjointe de (X, Z_1, \dots, Z_n) .

(c) Si (X, Z_1, \dots, Z_n) est une distribution normale multivariée alors :

$$P_{s\bar{p}\{1, Z_1, \dots, Z_n\}}(X) = E_{\mathcal{M}\{Z_1, \dots, Z_n\}} [X].$$

Nous comprenons donc mieux pourquoi la meilleure prévision linéaire est celle qui se rencontre le plus fréquemment. C'est d'ailleurs cette prévision qui nous permettra d'établir la représentation markovienne. Toute fois, c'est à partir de l'espérance conditionnelle que nous obtiendrons la représentation markovienne bilinéaire.

1.5.2 Représentation markovienne

La représentation markovienne (ou représentation de l'espace d'états) est considérée comme une approche très générale dans la modélisation des séries chronologiques univariées et multivariées. Cette représentation joue un très grand rôle dans l'étude des séries chronologiques, tout spécialement en ce qui a trait aux filtres de Kalman. Ce concept a d'ailleurs été originalement développé par des ingénieurs en contrôle (voir Kalman [1960], Kalman et Bucy [1961] et Kalman, Falb et Arbib [1969]). Akaike semble être le premier à avoir appliqué le concept directement à l'analyse des processus ARMA.

L'état d'un système est généralement interprété comme l'ensemble minimum d'informations provenant du présent et du passé qui fait en sorte que le comportement futur du système peut être complètement déterminé par la connaissance de l'état présent et de "l'input" futur. La représentation de l'espace d'états est donc basée sur la propriété markovienne qui implique que, connaissant l'état présent d'un système, le futur de ce dernier est indépendant du passé. Cela explique pourquoi la représentation de l'espace d'états est également appelée la représentation markovienne.

Le fait d'avoir une représentation markovienne pour les processus stochastiques constitue un atout important, cela nous permet d'avoir une **meilleure** connaissance de la structure probabiliste de ces modèles.

Dans cette section, nous traitons uniquement la représentation markovienne des processus linéaires. Plus loin, dans le chapitre 2, la représentation markovienne bilinéaire sera étudiée. Nous citons maintenant quelques définitions tirées de Wei (1990) et de Akaike (1974b).

Définition 1.5.7

Soit $X_t \in \mathbb{R}^m$ l'out put d'un système linéaire homogène (dans le temps). La représentation de l'espace des états (ou représentation markovienne) de X_t est de la forme :

$$\begin{cases} Z_{t+1} = AZ_t + BW_{t+1}, \\ X_t = CZ_t, \end{cases}$$

où t représente le temps, Z_t est un vecteur d'états du système au temps t de dimension k , W_{t+1} est un vecteur de dimension n de l'input du système au temps t et X_t est un vecteur de dimension m représentant l'out put du système au temps t . A , B et C sont respectivement des matrices constantes de dimensions $k \times k$, $k \times n$ et $m \times k$.

Notation : Dans cette section, nous notons par \hat{Y}_{t+j/H_t} la meilleure prévision linéaire de Y_{t+j} , ($j = 0, 1, \dots$), dans l'enveloppe linéaire $\{W_t, W_{t-1}, \dots, W_{t-p}\}$. La variable W_{t-p} est la dernière donnée du passé prise en considération.

Définition 1.5.8

Lorsque l'input W_t et l'output X_t sont tous deux stochastiques, la représentation markovienne de $\{X_t\}$ est alors de la forme :

$$\begin{cases} Z_{t+1} = AZ_t + BH_{t+1}, \\ X_t = CZ_t + U_t, \end{cases} \quad (1.5.1)$$

où H_{t+1} est un vecteur de dimension n qui correspond à l'innovation de $\{W_t\}$ au temps $t+1$, $c - \hat{a} - d$ $H_{t+1} = W_{t+1} - \hat{W}_{t+1/H_t}$. La variable U_t est un vecteur de dimension m indépendant de Z_t (pour que la représentation soit markovienne) et des inputs présents et passés $W_t, W_{t-1}, \dots, W_{t-p}$.

Théorème 1.5.3 (Akaike, 1974 b)

Considérons deux processus, $\{X_t\}$ (linéaire) et $\{W_t\}$, appartenant à l'espace \mathbb{L}^2 , qui sont de moyennes nulles, stationnaires, conjointement gaussiens et mutuellement corrélés avec covariance $R_j = E(X_{t+j}W_t')$ pour tout t . Les variables X_t et W_t sont respectivement des vecteurs de dimensions m et n .

Si nous avons pour hypothèse :

H_1 : L'espace de prévision $\mathcal{P}_t = \overline{\text{span}} \left\{ \hat{X}_{t+j/H_t}, (j = 0, 1, \dots) \right\}$ est de dimension finie, alors $\{X_t\}$ admet une représentation markovienne de la forme (1.5.1).

Théorème 1.5.4 (Akaike, 1974 a)

Soit $\{X_t\} \in \mathbb{L}^2$ un processus stationnaire de moyenne nulle et de dimension m . La représentation ARMA de $\{X_t\}$ et la représentation markovienne suivante sont équivalentes :

$$\begin{cases} Z_{t+1} = AZ_t + BH_{t+1}, \\ X_t = CZ_t, \end{cases}$$

où $U_t = 0$ lorsque $X_t = W_t$ pour tout t .

1.5.3 Les équations aux différences stochastiques

Il est très connu que "les équations aux différences stochastiques" posent leur intervention dans certaines classes de modèles souvent utilisés dans la littérature de l'analyse des séries chronologiques. Autrement dit, la quasi totalité des modèles de séries chronologiques peuvent être exprimés sous la forme de cette équation qui est du type suivant

$$X_{n+1} = A_{n+1}X_n + B_{n+1}, n \in \mathbb{Z}, \quad (1.5.2)$$

où X_n et B_n sont des vecteurs aléatoires définis en \mathbb{R}^d , les matrices A_n sont de dimension $d \times d$, et le processus $\{(A_n, B_n), n \in \mathbb{Z}\}$ est strictement stationnaire et ergodique. En outre, les propriétés de stationnarité de ces modèles sont directement liées à celles des solutions de cette équation. Par conséquent, il est suffisant de restreindre l'étude sur cette dernière.

Comme notre principal objectif porte sur l'existence des conditions nécessaires et suffisantes d'une solution stationnaire unique de cette équation, Brandt (1986) a établi le résultat suivant

Théorème 1.5.5 (Brandt, 1986)

Soit $\{(A_n, B_n), n \in \mathbb{Z}\}$ un processus strictement stationnaire ergodique tel que $E(\log^+ \|A_0\|)$ et $E(\log^+ \|B_0\|)$ sont toutes les deux finies. Supposons que le plus grand exposant de Lyapounov γ défini par $\gamma = \inf \{E(\frac{1}{n+1} \log \|A_0 A_{-1} \dots A_{-n}\|), n \in \mathbb{N}\}$ est strictement négatif. Alors, pour tout $n \in \mathbb{Z}$, la série

$$X_n = \sum_{k=0}^{+\infty} A_n A_{n-1} \dots A_{n-k+1} B_{n-k},$$

converge presque sûrement, et le processus $\{X_n, n \in \mathbb{Z}\}$ est l'unique solution strictement stationnaire de (1.5.2).

A noter que lorsqu'on se place dans le cas où (A_n, B_n) est i.i.d, l'équation (1.5.2) vérifiera un modèle autorégressif généralisé comme le montre la définition suivante :

Définition 1.5.9

Un modèle autorégressif généralisé est un modèle de la forme

$$X_{n+1} = A_{n+1}X_n + B_{n+1}, n \in \mathbb{Z}, \quad (1.5.3)$$

où $\{(A_n, B_n), n \in \mathbb{Z}\}$ est une suite de variables aléatoires i.i.d définies sur (Ω, \mathcal{A}, P) avec des valeurs en \mathbb{R}^d . La solution de cette équation est une séquence de \mathbb{R}^d variables aléatoires pour laquelle (1.5.3) est vérifiée.

Dans ce qui suit, on s'est intéressé aux solutions stationnaires non anticipatives, dans le sens de la définition suivante et dont les propriétés de causalité et de mesurabilité ne sont pas supposées. En effet, il en sera ainsi une conséquence du premier théorème principal qui est actuellement vrai sous la condition d'irréductibilité.

Définition 1.5.10

Une solution strictement stationnaire non anticipative de (1.5.3) est un processus strictement stationnaire $\{X_n, n \in \mathbb{Z}\}$ qui est une solution de (1.5.3) tel que, pour tout $p \in \mathbb{Z}$, X_p est indépendant des variables aléatoires $\{(A_n, B_n), n > p\}$.

En dehors de cette hypothèse, le théorème inverse de Brandt n'est pas vrai ; par exemple, si B_n est identiquement nulle, alors $X_n = 0$ est une solution stationnaire pour n'importe quelle séquence $\{A_n\}$. Afin d'éviter cette situation dégénérée, on introduit la condition d'irréductibilité suivante. La proposition (1.5.8) et le corollaire (1.5.9) vont nettement éclaircir le rôle de cette hypothèse.

Définition 1.5.11

Un sous espace affine H de \mathbb{R}^d est dit invariant sous le modèle (1.5.3) si $\{A_0x + B_0; x \in H\}$ est inclus dans H presque sûrement. Le modèle (1.5.3) est irréductible si \mathbb{R}^d est l'unique sous espace affine invariant.

Alors le premier résultat principal n'exige aucune condition d'intégrabilité. Il est donné par le théorème suivant :

Théorème 1.5.6 (Bougerol et Picard, 1992)

Considérons le modèle (1.5.3) avec des coefficients i.i.d. Supposons que ce modèle est irréductible et qu'il a une solution strictement stationnaire non anticipative $\{X_n, n \in \mathbb{Z}\}$. Alors il ressort

(i) $A_0 A_{-1} \dots A_{-k}$ converge vers zéro presque sûrement lorsque $k \rightarrow +\infty$.

(ii) Pour tout entier n , $X_n = \sum_{k=0}^{+\infty} A_n A_{n-1} \dots A_{n-k+1} B_{n-k}$, où la série converge presque sûrement.

(iii) Cette solution est l'unique solution strictement stationnaire de (1.5.3).

Le deuxième résultat principal est donné par le théorème inverse de Brandt suivant

Théorème 1.5.7 (Bougerol et Picard, 1992)

Supposons que le modèle autorégressif généralisé (1.5.3) avec des coefficients *i.i.d* est irréductible et que $E(\log^+ \|A_0\|)$ et $E(\log^+ \|B_0\|)$ sont toutes les deux finies. Alors (1.5.3) admet une solution strictement stationnaire non anticipative si et seulement si l'exposant de Lyapounov γ est strictement négatif.

La prochaine proposition peut être la première étape dans l'étude de situations non irréductibles.

Proposition 1.5.8 (Bougerol et Picard, 1992)

Supposons que le modèle (1.5.3) a une solution strictement stationnaire non anticipative $\{X_n, n \in \mathbb{Z}\}$. Soit H le sous espace affine minimal de \mathbb{R}^d tel que $P(X_0 \in H) = 1$. Alors H est invariant sous le modèle et n'importe quel sous espace invariant de H porte une solution strictement stationnaire non anticipative.

Une conséquence immédiate du théorème 1.5.5, théorème 1.5.7 et la proposition 1.5.8 est le corollaire suivant :

Corollaire 1.5.9 (Bougerol et Picard, 1992)

Considérons le modèle (1.5.3) avec $E(\log^+ \|A_0\|)$ et $E(\log^+ \|B_0\|)$ sont finies. Supposons qu'il existe une solution strictement stationnaire non anticipative qui n'est pas porté par l'hyperplan affine. Alors les conditions suivantes sont équivalentes

(i) L'exposant de Lyapounov est strictement négatif.

(ii) Le modèle est irréductible.

(iii) Il existe une unique solution stationnaire.

Chapitre 2

Les Modèles Bilinéaires

2.1 Introduction

Récemment, plusieurs travaux qui traitent les modèles non linéaires de séries chronologiques ont vu le jour, démontrant ainsi la limitation du champ d'application des modèles linéaires de séries chronologiques et le gain qu'on peut gagner en s'intéressant aux modèles non linéaires. En effet, dans plusieurs disciplines, l'existence de la non linéarité entre variables est reconnue.

Parmi les modèles non linéaires, on peut citer les modèles bilinéaires (*BL*), les modèles autorégressifs à seuil (*TAR*), les modèles autorégressifs exponentiels (*EXPAR*), les modèles *ARCH*, etc.

Dans ce chapitre on étudie le modèle bilinéaire d'ordre (p, q, P, Q) à valeurs réelles, noté $BL(p, q, P, Q)$ défini par la relation de récurrence suivante :

$$X_t = \sum_{i=1}^p a_i X_{t-i} + \sum_{j=1}^q c_j \varepsilon_{t-j} + \varepsilon_t + \sum_{j=1}^P \sum_{k=1}^Q b_{jk} X_{t-j} \varepsilon_{t-k}, \quad (2.1.1)$$

où X_t est un processus stochastique défini sur un certain espace de probabilité (Ω, A, P) , pour tout $t \in \mathbb{Z}$ et où les coefficients $(a_i, 1 \leq i \leq p)$, $(c_j, 1 \leq j \leq q)$, $(b_{jk}, 1 \leq j \leq P, 1 \leq k \leq Q)$ sont des constantes et où $(\varepsilon_t, t \in \mathbb{Z})$ est une suite de variables aléatoires indépendantes équidistribuées, gaussiennes, centrées, de variance finie σ^2 , définies aussi sur l'espace de probabilité (Ω, A, P) .

On peut distinguer trois classes de modèles bilinéaires :

* Les modèles bilinéaires diagonaux :

$$X_t = \sum_{i=1}^p a_i X_{t-i} + \sum_{j=1}^q c_j \varepsilon_{t-j} + \varepsilon_t + \sum_{j=1}^P b_j X_{t-j} \varepsilon_{t-j}.$$

* Les modèles bilinéaires superdiagonaux :

$$X_t = \sum_{i=1}^p a_i X_{t-i} + \sum_{j=1}^q c_j \varepsilon_{t-j} + \varepsilon_t + \sum_{j=1}^P \sum_{\substack{k=1 \\ j \geq k}}^Q b_{jk} X_{t-j} \varepsilon_{t-k}.$$

* Les modèles bilinéaires sousdiagonaux :

$$X_t = \sum_{i=1}^p a_i X_{t-i} + \sum_{j=1}^q c_j \varepsilon_{t-j} + \varepsilon_t + \sum_{j=1}^P \sum_{\substack{k=1 \\ j < k}}^Q b_{jk} X_{t-j} \varepsilon_{t-k}.$$

Certains résultats ne concernent que des classes particulières de modèles que nous préciserons. Par ailleurs, il est très visible que (2.1.1) est une extension naturelle non linéaire des modèles ARMA obtenue en ajoutant la forme bilinéaire $\{X_{t-i}, \varepsilon_{t-j}\}$.

Une approche naturelle pour commencer l'étude d'un modèle est d'en faire une description approfondie afin d'avoir une première idée de son comportement. Pour ce faire, on a choisi 05 séries caractérisant deux cas particuliers des modèles bilinéaires suivants :

▷ Le modèle $BL(1, 0, 1, 1)$:

$$(S1) : X_t = 0.4X_{t-1} + 0.4X_{t-1}\varepsilon_{t-1} + \varepsilon_t,$$

$$(S2) : X_t = 0.7X_{t-1} + 0.4X_{t-1}\varepsilon_{t-1} + \varepsilon_t,$$

$$(S3) : X_t = 0.7X_{t-1} + 0.7X_{t-1}\varepsilon_{t-1} + \varepsilon_t.$$

▷ Le modèle $BL(0, 0, 2, 1)$:

$$(S4) : X_t = 0.2X_{t-2}\varepsilon_{t-1} + \varepsilon_t,$$

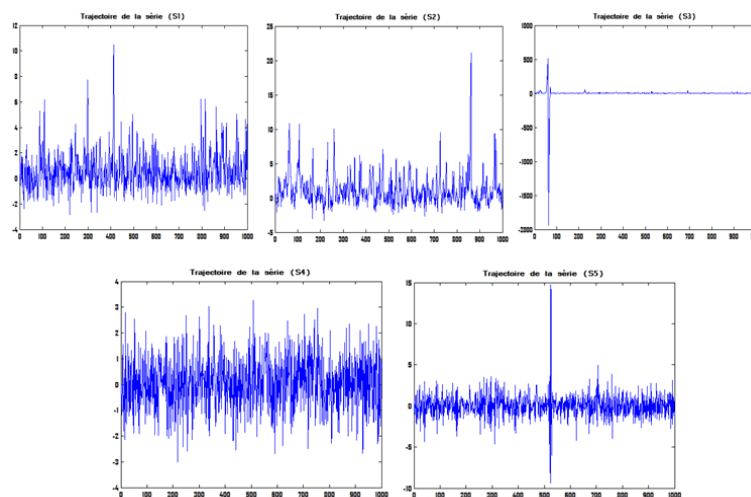
$$(S5) : X_t = 0.7X_{t-2}\varepsilon_{t-1} + \varepsilon_t,$$

Les caractéristiques marginales des cinq séries (moyenne, écart-type, quantiles) sont données dans le tableau suivant :

	Min	Max	Moyenne	Ecart-type	Q1%	Q5%	Q10%	Q25%
(S1)	-2.8394	10.4606	0.6283	1.4607	-1.9492	-1.3217	-0.9461	-0.2899
(S2)	-3.2908	21.0776	1.2004	2.2339	-2.0108	-1.3177	-0.9122	-0.1678
(S3)	$1.9456 \times e^3$	514.8527	-3.8766	104.5780	-8.4512	-1.8942	-1.0507	-0.1748
(S4)	-3.0151	3.2657	0.0343	1.0200	-2.2941	-1.6056	-1.2650	-0.6694
(S5)	-9.3847	14.7106	-0.0154	1.4699	-3.4247	-2.0384	-1.4938	-0.8530

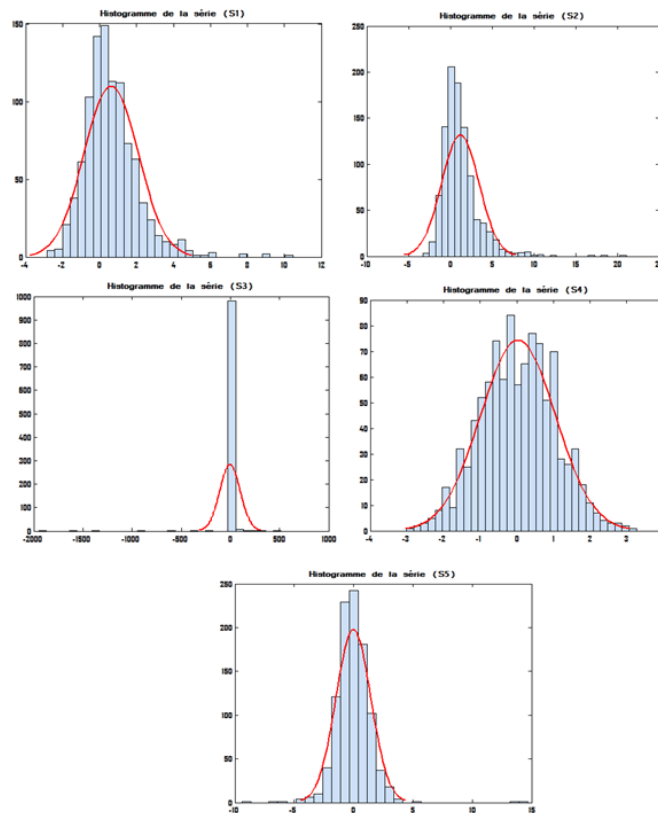
	Q50%	Q75%	Q90%	Q95%	Q99%	interquartile
(S1)	0.4087	1.3289	2.3084	3.1576	5.1566	1.6188
(S2)	0.8111	1.9664	3.7816	5.0846	9.1575	2.1342
(S3)	0.9432	3.2820	8.0520	12.6681	55.3898	3.4568
(S4)	0.0206	0.7459	1.3659	1.6871	2.4080	1.4153
(S5)	-0.0518	0.8008	1.6020	2.1030	3.2368	1.6538

On peut constater que pour la série (S3) il existe un grand pourcentage de valeurs extrêmes par rapport à la série (S2). Cela se traduit par un écart-type très important par rapport à l'interquartile, et par une augmentation de la moyenne par rapport à la médiane. Ainsi, ces deux dernières séries définissent un modèle bilinéaire diagonal d'ordre un, ce qui montre le fait que les séries du modèle purement bilinéaire et strictement superdiagonal d'ordre un ne portent pas de valeurs extrêmes. Afin d'illustrer quelques caractéristiques de non linéarité qui porte ces modèles, on présente dans les figures ci-dessus des réalisations présentant la trajectoire de ces séries où dans chaque cas les $\{\varepsilon_t\}$ sont des variables aléatoires indépendantes suivant la loi $N(0, 1)$ et chaque réalisation est formée de 1000 informations :



Remarquons que pour les deux types de modèles, l'amplitude de l'explosion devient importante de plus en plus que le coefficient du terme bilinéaire est élevé. Ainsi, la série (S3) est une série explosive dont on a déjà remarqué la forte proportion de valeurs extrêmes.

Par ailleurs, on peut examiner la distribution de la loi de probabilité de ces modèles à travers leurs histogrammes correspondants :



On peut voir que les queues de ces séries ne sont pas fines ou encore plus épaisses en les comparant avec celles de la loi normale. Ceci est expliqué théoriquement par le fait que la valeur du Kurtosis des modèles bilinéaires est supérieure à celle de la loi normale (qui est égale à 3).

L'étude des modèles bilinéaires à travers leurs principales propriétés probabilistes laisse préconiser l'existence et l'unicité d'une solution stationnaire (stricte et au second ordre), l'inversibilité, l'ergodicité, la fonction de covariance, l'existence des moments d'ordre supérieur et des cumulants. Elle sera établie selon deux approches dont les preuves de certains théorèmes sont renvoyées en annexe pour ne pas alourdir la lecture :

-L'approche par représentation d'état qui permet de généraliser de façon élégante des résultats obtenus par Akaike (1974a) pour les processus ARMA.

-L'approche directe qui s'appuie sur une représentation vectorielle des modèles bilinéaires.

2.2 Etude probabiliste des modèles bilinéaires via leurs représentations markoviennes

Tous les résultats de cette étude ont été établis sur la base que les processus bilinéaires représentent une généralisation des processus ARMA. Toutefois, en nous inspirant des résultats connus dans l'étude des séries chronologiques linéaires, qui proviennent principalement d'Akaike (1974a, 1974b), que nous expliquons les étapes de construction de quelques résultats relatifs aux séries chronologiques non linéaires tels que les représentations markoviennes. En outre, pour les modèles bilinéaires, il existe deux représentations markoviennes ; à savoir la représentation markovienne bilinéaire pour les processus bilinéaires diagonaux et super-diagonaux ainsi la représentation affine en l'état polynomial en entrées pour les processus bilinéaires généraux.

La suite de cette partie est consacrée à l'étude d'un certain nombre de propriétés probabilistes des processus bilinéaires à partir de leurs représentations.

2.2.1 Construction des représentations markoviennes

Dans toute cette partie, le processus $\{X_t, t \in \mathbb{Z}\}$ est un processus centré strictement stationnaire ayant ses moments d'ordre deux finis.

Représentation bilinéaire markovienne

On note :

* \mathbf{H}_t = l'espace des variables aléatoires mesurables de carré intégrable.

* F_t = la σ -algèbre générée par les $X_s, s \leq t$.

* $\hat{Y}(u \setminus \mathbf{H}_t)$ = la valeur prévu conditionnelle de la variable (ou le vecteur) Y_u donnée par \mathbf{H}_t .

On définit :

* l'espace des prédicteurs non linéaires \mathcal{P}_t comme le sous-espace fermé engendré par les prédicteurs : $\hat{X}_{t+k \setminus \mathbf{H}_t} = E(X_{t+k} \setminus F_t), k = 1, 2, \dots$, c'est-à-dire : $\mathcal{P}_t = s\bar{p} \left\{ \hat{X}_{t+k \setminus \mathbf{H}_t}, k = 1, 2, \dots \right\}$

*L'espace d'innovation \mathcal{I}_t comme le sous espace fermé engendré par les variables aléatoires $\hat{X}_{t+k \setminus \mathbf{H}_t} - \hat{X}_{t+k \setminus \mathbf{H}_{t-1}}, k = 1, 2, \dots$, (ou bien c'est l'espace qui correspond à l'enveloppe linéaire de ces variables), c'est-à-dire : $\mathcal{I}_t = s\bar{p} \left\{ \hat{X}_{t+k \setminus \mathbf{H}_t} - \hat{X}_{t+k \setminus \mathbf{H}_{t-1}}, k = 1, 2, \dots \right\}$.

Par analogie au cas linéaire, on introduit les hypothèses suivantes :

* (H_0) : L'espace des prédicteurs est de dimension finie.

* (H_1) : Les éléments de \mathcal{I}_t peuvent s'exprimer sous la forme : $\varepsilon_t Y_{t-1} + c\varepsilon_t + d(\varepsilon_t^2 - \sigma^2)$,

où :

- ▶ $\varepsilon_t = X_t - \hat{X}_{t \setminus \mathbf{H}_{t-1}}$
- ▶ $Y(t-1) \in \mathbf{H}_{t-1}$
- ▶ c, d sont des constantes
- ▶ $\sigma^2 = E(\varepsilon_t^2 / F_{t-1})$ est supposée constante.

* (H_2) : Les variables aléatoires Y_t de l'hypothèse H_1 appartiennent au sous-espace vectoriel $\mathcal{P}_t(m)$ engendré par \mathcal{P}_t et par les variables aléatoires X_t, \dots, X_{t-m+1} , pour un certain $m \geq 0$.
 Sous ces hypothèses, la représentation markovienne bilinéaire des modèles diagonaux et superdiagonaux s'obtient naturellement, et elle est donnée par la proposition suivante :

Proposition 2.2.1

Soit Z_t un vecteur ayant pour composantes les éléments d'une base de $\mathcal{P}_t(m)$. En utilisant (H_0) , (H_1) et (H_2) , on obtient

$$\begin{cases} X_t = H Z_{t-1} + \varepsilon_t, \\ Z_t = A Z_{t-1} + B Z_{t-1} \varepsilon_t + C \varepsilon_t + D(\varepsilon_t^2 - \sigma^2), \end{cases} \quad (2.2.1)$$

où ε_t est une suite de variables aléatoires indépendantes identiquement distribuées avec :

$E(\varepsilon_t | Z_{t-1}, \dots) = 0$, $E(\varepsilon_t^2 | Z_{t-1}, \dots) = \sigma^2$, et où A , B , sont des matrices carrées, $(n \times n)$, C , D , sont des matrices colonnes, $(n \times 1)$ et H est une matrice ligne $(1 \times n)$.

En effet :

Soit (Z_1, \dots, Z_n) une base de l'espace de prévision \mathcal{P}_t et Z_t le vecteur des composantes $Z_{i,t}$.

L'hypothèse H_0 implique que les variables \hat{X}_{t/\mathbf{H}_t} et $\hat{Z}_{t/\mathbf{H}_{t-1}}$ sont obtenues par des transformations de Z_{t-1} , c'est à dire :

$$\begin{cases} \hat{X}_{t/H_t} = H Z_{t-1}, \\ \hat{Z}_{t/H_{t-1}} = A Z_{t-1}, \end{cases}$$

pour certaines matrices lignes et carrées, H et A (qui ne dépendent pas de t par la stationnarité de $\{X_t\}$).

Toutefois, contrairement au cas linéaire, cette hypothèse n'est pas suffisante pour déterminer une représentation markovienne adéquate de X_t . Il en est ainsi puisque, dans le cas non linéaire, $X_t - \hat{X}_{t/\mathbf{H}_{t-1}}$ et $Z_t - \hat{Z}_{t/\mathbf{H}_{t-1}}$ ne sont pas toujours des multiples de ε_t .

Nous savons que les composantes du vecteur $Z_t - \hat{Z}_{t/\mathbf{H}_{t-1}}$ appartiennent à l'espace \mathcal{I}_t engendré par $\hat{X}_{s/\mathbf{H}_t} - \hat{X}_{s/\mathbf{H}_{t-1}}$, $s > t$.

De plus $X_t - \hat{X}_{t/\mathbf{H}_{t-1}}$ et les éléments de \mathcal{I}_t ne sont pas corrélés avec les éléments de l'espace \mathbf{H}_{t-1} , c'est -à -dire :

$$E\left(\left(\hat{X}_{s/\mathbf{H}_t} - \hat{X}_{s/\mathbf{H}_{t-1}}\right) V\right) = 0, \forall V \in H_{t-1}, s > t.$$

Donc, il s'avère nécessaire d'avoir l'hypothèse H_1 qui assure l'orthogonalité entre les éléments de \mathcal{I}_t et \mathbf{H}_t .

Si nous permettons à Y_t de varier librement dans \mathbf{H}_t , l'espace \mathcal{I}_t ne peut être déterminé par un ensemble fini de paramètres, et nous ne pouvons alors espérer avoir une représentation

avec un nombre fini de paramètres.

L'hypothèse H_2 vient restreindre le domaine de Y_t .

Comme nous l'avons vu, l'hypothèse H_0 implique que :

$$\begin{cases} \hat{X}_{t/\mathbf{H}_{t-1}} = HZ_{t-1}, \\ \hat{Z}_{t/\mathbf{H}_{t-1}} = AZ_{t-1}. \end{cases}$$

Et si, en vertu de l'hypothèse H_2 , nous posons $Y_{t-1} = Z_{t-1}$ dans l'hypothèse H_1 , il s'ensuit alors que $\{X_t\}$ a une représentation de la forme (2.2.1).

Remarque 2.2.1

1°/ la représentation (2.2.1) est une représentation bilinéaire markovienne dont :

- $(\varepsilon_t, t \in \mathbb{Z})$ est le bruit,
- $Z_t = (Z_{1,t}, \dots, Z_{n,t})'$ et ses composantes sont appelées vecteur d'état et variables d'état, respectivement.
- $(Z_t, t \in \mathbb{Z})$ est markoviennes si ε_t est indépendant de $Z_s, s < t$.
- La présence du terme σ^2 a pour but de centrer le modèle.

2°/ les matrices H, A, B, C, D qui vérifient l'équation (2.2.1) se diffèrent, selon que les modèles bilinéaires soient superdiagonaux ou diagonaux, comme suit :

· Les modèles bilinéaires superdiagonaux admettent la représentation (2.2.1), *Pham (1985)*, avec :

$$A_{(n \times n)} = \begin{pmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \vdots & & & \ddots & \vdots \\ & & & & 1 \\ 0 & & & & a_1 & 1 \\ 0 & & & & a_2 & \\ \vdots & & & & \vdots & \\ a_n & a_{n-1} & \dots & a_{n-m} & \dots & 0 \end{pmatrix}, B_{(n \times n)} = \begin{pmatrix} 0 & \dots & 0 & 0 & 0 \\ \vdots & & & \vdots & \\ b_{1m} & \dots & b_{10} & 0 & \dots & 0 \\ \vdots & & & & \vdots & \\ b_{n-m,m} & \dots & b_{n-m,0} & 0 & \dots & 0 \end{pmatrix},$$

$$C_{(n \times 1)} = (0, 0, \dots, 1, a_1 + c_1, \dots, a_{n-m} + c_{n-m})', D_{(n \times 1)} = (0, 0, \dots, b_{10}, \dots, b_{n-m,0})', H_{(1 \times n)} = (1, \dots, 0).$$

· Les modèles bilinéaires diagonaux admettent la représentation (2.2.1), *Guégan (1988)*, avec :

$$A_{(r \times r)} = \begin{pmatrix} a_1 & 1 & 0 & \cdots & 0 \\ a_2 & 0 & 1 & \cdots & 0 \\ \vdots & & \ddots & \ddots & \vdots \\ a_{r-1} & 0 & \cdots & 0 & 1 \\ a_r & 0 & 0 & \cdots & 0 \end{pmatrix}, B_{(r \times r)} = \begin{pmatrix} b_1 & 0 & \cdots & 0 \\ \vdots & & & \vdots \\ b_r & 0 & \cdots & 0 \end{pmatrix},$$

$$C_{(r \times 1)} = (a_1 + c_1, \dots, a_r + c_r)', D_{(r \times 1)} = (b_1, \dots, b_r)', H_{(1 \times r)} = (1, \dots, 0), a_j = 0 \text{ pour } j > p, b_j = 0 \text{ pour } j > P, c_j = 0 \text{ pour } j > q \text{ et } r = \max(p, q, P).$$

3° / En ce qui concerne les modèles bilinéaires sous diagonaux, Guégan (1988) a donné une représentation markovienne pour ces modèles, qu'elle a nommé représentation l -markovienne. Pour plus de détail on peut se référer à Guégan (1988, *chapitre 5*, p.165).

Représentation affine en l'état des modèles bilinéaires

En posant

$$\begin{cases} A_t = A + B\varepsilon_t, \\ H_t = C\varepsilon_t + D\varepsilon_t^2, \end{cases}$$

la deuxième équation de (2.2.1) devient : $Z_t = A_t Z_{t-1} + H_t$.

Cette dernière, d'après Pham (1986), reste vraie dans le cas général où A_t, H_t sont une matrice et un vecteur polynomial en ε_t de degrés $P' + 1$ et $P' + 2$ respectivement.

D'où :

Une représentation markovienne affine en l'état des processus bilinéaires définis par (1.1.1), stationnaires et appartenant à \mathbb{L}^2 , a été introduit par Pham (1986), et elle est de la forme :

$$\begin{cases} X_t = B(\varepsilon_t) Z_{t-1} + C(\varepsilon_t), & (2.2.2a) \\ Z_t = A(\varepsilon_{t-1}) Z_{t-1} + H(\varepsilon_{t-1}), & (2.2.2b) \end{cases}$$

et cela en utilisant les mêmes arguments que Pham (1985) rien que l'hypothèse H_1 devient :

* H_1 : les éléments de \mathcal{I}_t peuvent s'exprimer sous la forme :

$A(\varepsilon_{t-1}) Y_{t-1} + H(\varepsilon_{t-1})$, où :

► $\varepsilon_t = X_t - \hat{X}_{t/H_{t-1}}$,

► $Y_{t-1} \in H_{t-1}$,

► $A(\varepsilon_t)$ et $H(\varepsilon_t)$ sont respectivement des matrices carrée ($n \times n$) et ligne $(1, n)$ polynomiales en ε_t de degré fini .

► $C(\varepsilon_t)$ est un polynôme en ε_t de degré fini. On supposera en fait que $E(C(\varepsilon_t)) = 0$, mais cette hypothèse peut être levée.

► Z_0 est donné et est dépendant de $A(\varepsilon_t), H(\varepsilon_t)$ pour tout $t = 1, 2, \dots$

2.2.2 Condition d'existence d'une représentation markovienne

Nous allons maintenant donner les conditions d'existence des représentations markoviennes (2.2.1) et (2.2.2).

* Dans le cas de la représentation markovienne (2.2.1) on a le résultat suivant :

Théorème 2.2.2 (Pham, 1985)

*On suppose que les ε_t sont indépendants avec des moments d'ordre 4 finis. Une condition **nécessaire et suffisante** pour l'existence d'une solution stationnaire de (2.2.1) est que l'une des deux conditions suivantes soit vérifiée :*

(i) *L'équation matricielle $Q = AQ\acute{A} + BQ\acute{B}\sigma^2 + \Sigma$ où Σ représente la matrice de covariance de $C\varepsilon_t + D\varepsilon_t^2$ admet une solution positive en Q .*

(ii) *La suite Q_k définie par :*

$$\begin{cases} Q_0 = 0, \\ Q_{k+1} = AQ_k\acute{A} + BQ_k\acute{B}\sigma^2 + \Sigma, \end{cases}$$

est convergente.

Alors sous les conditions ci-dessus, une solution stationnaire à (2.2.1) est donnée par :

$$\begin{cases} X_t = HZ_{t-1} + \varepsilon_t, \\ Z_t = \sum_{j=0}^{\infty} \prod_{k=0}^{j-1} (A + B\varepsilon_{t-k}) (C\varepsilon_{t-j} + D(\varepsilon_{t-j}^2 - \sigma^2)), \end{cases} \quad (2.2.3)$$

avec les séries ci-dessus convergent en moyenne quadratique et presque sûrement. La solution est unique si l'équation $Q = AQ\acute{A} + BQ\acute{B}\sigma^2 + \Sigma$ n'admet pas de solution positive non triviale.

* Dans le cas des processus définis par la relation (2.2.2), on dispose d'une solution d'existence similaire. Pour cela on pose les notations suivantes : $E(M(\varepsilon_t) \otimes N(\varepsilon_t)) = M\tilde{\otimes}N$, $E(M(\varepsilon_t) N(\varepsilon_t)) = MN$, $E(M(\varepsilon_t)) = \tilde{M}$, pour tout t , où $M\tilde{\otimes}N$, $M\tilde{N}$, \tilde{M} sont des fonctions de σ , pour simplifier on omettra σ , \otimes représente le produit de Kronecker des matrices.

Théorème 2.2.3 (Guégan, 1987)

Soit le processus $X_t, t \in \mathbb{Z}$ défini en (2.2.2) et supposons que $E(H(\varepsilon_t)) = 0$, alors une condition **nécessaire et suffisante** pour l'existence d'une solution stationnaire de (2.2.2) est que l'une des deux conditions suivantes soit vérifiée :

- (i) L'équation matricielle $\text{vec}W = A \otimes A \text{vec}W + H \tilde{\otimes} H$ admet une solution positive en W .
 (ii) La suite W_k définie par

$$\begin{cases} W_0 = W, \\ W_k = E \left(A(\varepsilon_t) W_{k-1} \dot{A}(\varepsilon_t) \right) + \Delta_1, \text{ où} \\ \Delta_1 = E \left(A(\varepsilon_t) Z_t \dot{H}(\varepsilon_t) + H(\varepsilon_t) \dot{Z}(t) \dot{A}(\varepsilon_t) + H(\varepsilon_t) \dot{H}(\varepsilon_t) \right), \end{cases} \text{ est convergente.}$$

Sous les conditions précédentes, une solution stationnaire de (2.2.2) est donnée par :

$$\begin{cases} X_t = B(\varepsilon_t) Z_{t-1} + C(\varepsilon_t), \\ Z_t = \sum_{j \geq 0} P_{t,j-1} H(\varepsilon_{t-j-1}), \text{ où} \\ P_{t,r} = \prod_{k=1}^r A(\varepsilon_{t-k}), \quad r > 0. \end{cases} \quad (2.2.4)$$

La solution est unique si l'équation $W = \left(A(\varepsilon_t) W \dot{A}(\varepsilon_t) \right)$ n'admet pas de solution non triviale.

Notons que le théorème (2.2.3) concerne les modèles centrés et il peut être étendu au cas des modèles non centrés. En effet, on a le résultat suivant :

Théorème 2.2.4 (Guégan, Pham, 1987)

Une condition **nécessaire et suffisante** d'existence d'une solution stationnaire de (2.2.2), c'est à dire l'existence de deux processus stationnaires au second ordre satisfaisant (2.2.2) est que :

- (a) L'équation $\mu = E(A(\varepsilon_t))\mu + E(H(\varepsilon_t))$ admette une solution en μ , et que l'une des conditions suivantes soit vraie :

- (b) L'équation matricielle $\text{vec}W = A \tilde{\otimes} A \text{vec}W + \Delta$, où

$$\Delta = E \left(A(\varepsilon_t) \mu \dot{H}(\varepsilon_t) + H(\varepsilon_t) \dot{\mu} \dot{A}(\varepsilon_t) + H(\varepsilon_t) \dot{H}(\varepsilon_t) \right) \text{ admet une solution } W \geq \mu \dot{\mu};$$

- (c) La suite W_k définie par :

$$\begin{cases} W_0 = \mu \dot{\mu}, \\ \text{vec}W_k = A \tilde{\otimes} A \text{vec}W_k + \Delta. \end{cases} \text{ est convergente.}$$

Sous ces conditions, une solution stationnaire est donnée par :

$$\begin{cases} X_t = B(\varepsilon_t) Z_{t-1} + \varepsilon_t, \\ Z_t = \sum_{j \geq 0} \prod_{k=1}^{j-1} A(\varepsilon_{t-k}) H(\varepsilon_{t-j-1}). \end{cases}$$

La solution est unique si $I - E(A(\varepsilon_t))$ est inversible et si l'équation $W = (A(\varepsilon_t) W A(\varepsilon_t))$ n'admet pas de solution non triviale.

Pour la preuve, on montre tout d'abord que la stationnarité implique (a) et (b), puis que sous (a), (b) implique (c), enfin que (a) et (c) impliquent la stationnarité. On peut toujours se ramener au cas $\mu = 0$ en travaillant sur le processus $\tilde{Z}_t = Z_t - \mu$. La démonstration du théorème (2.2.4) se fait alors à partir du processus \tilde{Z}_t défini par :

$\tilde{Z}_t = A(\varepsilon_t) \tilde{Z}_{t-1} + \zeta(\varepsilon_t)$ où $\zeta(\varepsilon_t) = A(\varepsilon_t) - E(A(\varepsilon_t))\mu + H(\varepsilon_t) - E(H(\varepsilon_t))$ qui est un processus centré.

2.2.3 Inversibilité d'une représentation markovienne

Nous établissons l'inversibilité des modèles bilinéaires admettant la représentation markovienne (2.2.1).

Pour cela nous utilisons la définition faible de l'inversibilité et nous donnons un algorithme permettant de construire la suite $\hat{\varepsilon}_t$.

Dans le cas des modèles que nous considérons, une construction naturelle de la suite $\hat{\varepsilon}_t$ est la suivante : étant donnée une valeur arbitraire z_0 de Z_0 , on définit $\varepsilon_{t/z_0}, Z_{t/z_0}$ récursivement à partir des équations du modèle (2.2.1), on obtient :

$$\begin{cases} \varepsilon_{t/z_0} = X_t - H Z_{t-1/z_0}, \\ Z_{t/z_0} = (A + B \varepsilon_{t/z_0}) Z_{t-1/z_0} + C \varepsilon_{t/z_0} + D \varepsilon_{t/z_0}^2 - D\sigma^2, \end{cases} \quad (2.2.5)$$

avec $Z_{0/z_0} = z_0$.

Dans la suite, on utilisera :

> la norme opérateur définie par $\|K\|_\varphi = \sup_{\varphi(x) \leq 1} \varphi(Kx)$, où K est une matrice carrée, et φ une norme vectorielle.

> La condition (H_1) : " $B = D H$ " qui est en particulier vérifiée pour les modèles diagonaux.

On a alors le théorème suivant pour le processus (2.2.1) :

Théorème 2.2.5 (Guégan, Pham, 1987)

S'il existe une norme vectorielle φ telle que : $E \left(\log \|K\|_\varphi \right) < 0$,

où $K_t = A - CH - BX_t$, alors $\hat{\varepsilon}_{t/z_0} - \varepsilon_t \rightarrow 0$ presque sûrement quand $t \rightarrow \infty \forall z_0$.

Preuve :

On définit la suite $\hat{\varepsilon}_{t/z_0}$ par $\hat{\varepsilon}_{t/z_0} = X_t - H \hat{Z}_{t-1/z_0}$ dont \hat{Z}_{t-1/z_0} étant défini par l'équation (2.2.5) où l'on remplace ε_{t/z_0} par $\hat{\varepsilon}_{t/z_0}$.

Alors :

$$\begin{aligned}
 \hat{Z}_{t/z_0} &= (A + B \hat{\varepsilon}_{t/z_0}) \hat{Z}_{t-1/z_0} + C \hat{\varepsilon}_{t/z_0} + D \hat{\varepsilon}_{t/z_0}^2 - D \sigma^2, \\
 &= \left(A + B X_t - B H \hat{Z}_{t-1/z_0} \right) \hat{Z}_{t-1/z_0} + C X_t - CH \hat{Z}_{t-1/z_0} + D \left(X_t - H \hat{Z}_{t-1/z_0} \right)^2 \\
 &\quad - D \sigma^2, \\
 &= (A - CH) \hat{Z}_{t-1/z_0} + B X_t \hat{Z}_{t-1/z_0} - B H \hat{Z}_{t-1/z_0}^2 + C X_t + D X_t \\
 &\quad - 2 D H X_t \hat{Z}_{t-1/z_0} + D H^2 \hat{Z}_{t-1/z_0}^2 - D \sigma^2, \\
 &= (A - CH) \hat{Z}_{t-1/z_0} + C X_t + D X_t^2 + B X_t \hat{Z}_{t-1/z_0} - D H^2 \hat{Z}_{t-1/z_0}^2 \\
 &\quad - 2 B X_t \hat{Z}_{t-1/z_0} + D H^2 \hat{Z}_{t-1/z_0}^2 - D \sigma^2, \\
 &= (A - CH - B X_t) \hat{Z}_{t-1/z_0} + C X_t + D X_t^2 - D \sigma^2.
 \end{aligned}$$

$$\hat{Z}_{t/z_0} = K_t \hat{Z}_{t-1/z_0} + C X_t + D X_t^2 - D \sigma^2,$$

$$\text{où } K_t = A - CH - B X_t.$$

Il suffit alors de montrer la convergence vers 0 de $\hat{Z}_t - Z_t$.

On a :

$$\begin{aligned}
 \hat{Z}_t - Z_t &= K_t \left(\hat{Z}_{t-1} - Z_{t-1} \right) \\
 &= K_t \left(K_{t-1} \hat{Z}_{t-2} - K_{t-1} Z_{t-2} \right) \\
 &\quad \vdots \\
 &= \Pi_{j=0}^{t-1} K_{t-j} \left(\hat{Z}_0 - Z_0 \right)
 \end{aligned}$$

En posant $k = t - j$, on obtient :

$$\hat{Z}_t - Z_t = \Pi_{k=1}^t K_k \left(\hat{Z}_0 - Z_0 \right)$$

Or :

$$\begin{aligned} \frac{1}{t} \log \|\Pi_{k=1}^t K_k\|_\varphi &\leq \frac{1}{t} \log \Pi_{k=1}^t \|K_k\|_\varphi \\ &\leq \frac{1}{t} \sum_{k=1}^t \log \|K_k\|_\varphi \end{aligned}$$

Ce qui converge presque sûrement par le théorème ergodique vers :

$$E \left(\log \|K_k\|_\varphi \right) = \alpha < 0.$$

Donc pour tout α' , $\alpha < \alpha' < 0$, il existe un nombre T assez grand tel que :

$$\frac{1}{t} \log \|\Pi_{k=1}^t K_k\|_\varphi \leq \alpha < \alpha' < 0$$

$$\Rightarrow \log \|\Pi_{k=1}^t K_k\|_\varphi \leq \alpha' t$$

$$\Rightarrow \|\Pi_{k=1}^t K_k\|_\varphi \leq e^{\alpha' t}, t > T,$$

D'où :

$$\begin{aligned} \lim_{t \rightarrow \infty} \hat{Z}_t - Z_t &= \lim_{t \rightarrow \infty} \left\{ C \|\Pi_{k=1}^t K_k\|_\varphi \right\} \\ &\leq \lim_{t \rightarrow \infty} e^{\alpha' t} = 0 \quad \text{car } \alpha' < 0, \end{aligned}$$

et où $C = \hat{Z}_0 - Z_0$, ce qui termine la démonstration . ■

2.2.4 Minimalité d'une représentation markovienne

Dans la construction d'une représentation markovienne on a intérêt à faire en sorte que la dimension de l'espace d'état soit la plus petite possible pour la mémorisation et les calculs. Ce problème est lié à la notion de minimalité de la représentation.

Suivant Kalman, Falb et Arbib (1969), ou Faurre, Clerget et Germain (1979), nous disposons de la définition suivante :

Définition 2.2.6

Etant donnée une représentation interne d'un système dynamique, celle-ci est minimale s'il n'existe pas d'autre (représentant le même système) pour laquelle l'espace d'état soit de dimension inférieure.

Dans le cas de la représentation markovienne d'un modèle bilinéaire on dispose de critères permettant d'assurer une propriété un peu plus faible que la minimalité définie ci-dessus, toutefois il n'y a pas de lien explicite de la minimalité avec l'unicité de la représentation markovienne.

Définition 2.2.7

On dit qu'une représentation est "quasi-minimale" s'il n'existe pas d'autre représentation définie à partir du même bruit et dont le vecteur d'état soit la transformation linéaire du vecteur d'état d'origine avec une dimension plus petite.

Notons que si on se restreint aux représentations inversibles alors cette notion est la même que la notion de minimalité donnée dans la définition (2.2.6).

** On dispose alors, dans le cas de la représentation (2.2.1), du théorème suivant :

Théorème 2.2.8 (Pham, 1985)

On suppose que la représentation (2.2.1) est régulière (c'est à dire que le processus $X_t, t \in Z$ est donné par (2.2.3)). Soit :

• n : la dimension de l'espace d'état,

• Q_k la suite définie par :
$$\begin{cases} Q_0 = 0, \\ Q_{k+1} = A Q_k A' + B Q_k B' \sigma^2 + \Sigma, \end{cases}$$

• P_k la suite définie par :
$$\begin{cases} P_0 = 0, \\ P_{k+1} = A' P_k A + B' P_k B \sigma^2 + H' H, \end{cases}$$

• C_k la matrice formée par les colonnes de : $A^{n_1} B^{m_1} \dots A^{n_r} B^{m_r} \Sigma$,

• O_k la matrice formée par les vecteurs lignes : $H A^{n_1} B^{m_1} \dots A^{n_r} B^{m_r}$,

où Σ est la matrice de covariance de $C\varepsilon_t + D\varepsilon_t^2$, et où $n_1, m_1, \dots, n_r, m_r$ sont des entiers non

négatifs tels que $\sum_{i=1}^r (n_i + m_i) \leq k$.

Alors la représentation (2.2.1) est quasiminimale si et seulement si les conditions (c) et (O) sont vérifiées, où (c) est l'une des conditions suivantes :

(c₁) : Q_k est non singulière pour tout $k \leq n$,

(c₂) : C_k est de rang n pour tout $k \leq n$,

et (O) est l'une des conditions suivantes :

(O₁) : P_k est non singulière pour tout $k \leq n$,

(O₂) : O_k est de rang n pour tout $k \leq n$.

** Dans le cas de la représentation (2.2.2), on a un résultat du même type pour la minimalité de la représentation, que celui énoncé dans le théorème (2.2.8) :

Théorème 2.2.9 (Guégan, Pham, 1987)

Soient $X_t, t \in \mathbb{Z}$ le processus défini par les équations (2.2.2), p la dimension de l'espace d'état, Δ et W_k définies comme dans le théorème (2.2.2), P_k la suite définie par :

$$P_{k+1} = E \left(A(\varepsilon_t) P_k \dot{A}(\varepsilon_t) \right) + D, \text{ où } D = E \left(\dot{B}(\varepsilon_t) B(\varepsilon_t) \right).$$

Soit enfin la décomposition de $A(\varepsilon_t)$ suivant les polynômes d'Hermite h_j :

$$A(\varepsilon_t) = \sum_{j=0}^r A_j h_j(\varepsilon_t). \text{ Alors la représentation (2.2.2) est quasiminimale si et seulement}$$

si l'une des conditions équivalentes de (a) et l'une des conditions équivalentes de (b) est réalisée :

(a) (i) W_k est non singulière pour $k \leq p$.

(ii) La matrice C_k formée par les colonnes des matrices $\Delta, A_{i_1} \dots A_{i_n} \Delta, 1 \leq n \leq k$ est de rang p , pour un $k \leq p$.

(b) (i) P_k est non singulière pour $k \leq p$.

(ii) La matrice L_k formée par les lignes des matrices $D, DA_{i_1} \dots A_{i_n}, 1 \leq n \leq k$ est de rang p , pour un $k \leq p$.

Notons que la décomposition de $A(\cdot)$ suivant les polynômes d'Hermite est utilisée pour faciliter la démonstration, mais les résultats restent vrais pour toute décomposition de $A(\cdot)$ suivant une autre base de polynômes.

2.2.5 Ergodicité géométrique et régularité absolue

Bien que l'ergodicité géométrique soit définie uniquement pour des processus markoviens, la régularité absolue est définie pour des processus généraux. En effet :

- Si notre processus d'origine est stationnaire, on parle dans ce cas de la régularité absolue de la représentation markovienne de ce processus,
- Si notre processus d'origine est markovien, on parle directement de l'ergodicité géométrique.

Dans notre cas, on dispose d'une représentation markovienne d'un processus bilinéaire, donc on parle de la régularité absolue. Ainsi, ce processus est relativement lié à un autre processus qui est markovien donc l'ergodicité géométrique de ce dernier implique la régularité absolue du premier processus mais avec une vitesse de convergence géométrique. Cela nous permet de conclure que ces deux notions (ergodicité géométrique et régularité absolue avec une vitesse de convergence géométrique) se coïncident pour les processus markoviens, **c.f (Pham, 1986)**. Etant donnée la représentation markovienne des modèles bilinéaires diagonaux et superdiagonaux donnée par

$$\begin{cases} X_t = H Z_{t-1} + \varepsilon_t, \\ Z_t = A Z_{t-1} + B Z_{t-1} \varepsilon_t + C \varepsilon_t + D(\varepsilon_t^2 - \sigma^2), \end{cases}$$

où Z_t est un **processus vectoriel stationnaire** avec Z_t indépendant de $\varepsilon_s, s > t$. Ainsi, la notion de la régularité absolue du processus Z_t est définie comme suit :

Définition 2.2.8 (Pham, 1986)

Soit $(Z_t, t \in \mathbb{Z})$ un processus vectoriel stationnaire, et posant $\Delta_n = \sup_{h \in S_n} E(h/Z_s, s < 1) -$

$E(h)$, où S_n représente l'ensemble des variables aléatoires mesurables qui correspondent à la tribu engendrée par $Z_t, t \geq n$, et bornées par 1 en valeur absolue. Ce processus est dit absolument régulier si $E(\Delta_n)$ tend vers 0 quand n tend vers l'infini.

* Afin d'atteindre la régularité absolue du processus X_t , on montre d'abord que le processus Z_t est markovien :

En posant $A_t = A + B\varepsilon_t, \zeta_t = C\varepsilon_t + D\varepsilon_t^2$, la deuxième égalité de cette dernière peut être écrit comme

$$Z_t = A_t Z_{t-1} + \zeta_t. \tag{2.2.6}$$

Cette relation est la même que celle définie pour un modèle *RCA*, à l'exception que la matrice aléatoire A_t n'est pas indépendante de ζ_t comme il est exigé dans ce modèle. On doit donc appelé "processus *RCA* généralisé", le processus Z_t défini en (2.2.6) avec (A_t, ζ_t) est un couple d'un vecteur et matrice aléatoires conjointement indépendants et identiquement distribués et indépendant de Z_s pour $s < t$. Un tel processus est clairement markovien avec une mesure de probabilité de transition $P(z, \cdot)$ égale à la distribution de $A_t z + \zeta_t$.

** Lorsque Z_t est markovien, il devient clair que Δ_n ne dépend que de Z_0 , qui est donnée par $\|P^n(Z_0, \cdot) - \pi(\cdot)\|$ où $P^n(z, \cdot)$ est la distribution conditionnelle de $Z(t+n)$ étant donnée $Z(t) = z$ et π est la distribution de $Z(t)$. D'où, l'ergodicité de la chaîne de Markov est équivalente à la régularité absolue, avec une vitesse de convergence géométrique, du processus stationnaire avec une probabilité initiale invariante.

*** Comme X_t est un processus subordonné au processus de Markov Z_t dans le sens où la distribution conditionnelle de (X_t, Z_t) étant donnée (X_s, Z_s) , $s < t$ dépend uniquement de Z_{t-1} . Alors, par plusieurs conditionnement sur (X_s, Z_s) , $s < n+k$, $k \in \mathbb{N}$, il peut être vue que Δ_n (qui correspond au processus X_s) est bornée par l'espérance conditionnelle, étant donnée X_s , $s < 1$, de Δ_n (qui correspond au processus Z_t). Ainsi, la régularité absolue du processus X_t s'ensuit à partir de celle du processus Z_t , avec une vitesse de convergence identique ou plus rapide.

2.2.6 Moments et cumulants

Fonction de covariance

Grâce à la représentation markovienne, le calcul des moments peut se faire plus facilement que quand utilise l'approche directe. Considérons les processus bilinéaires qui admettent la représentation (2.2.1) donnée par :

$$\begin{cases} X_t = HZ_{t-1} + \varepsilon_t, \\ Z_t = AZ_{t-1} + BZ_{t-1}\varepsilon_t + D(\varepsilon_t^2 - \sigma^2) + C\varepsilon_t. \end{cases}$$

Fonction de covariance de Z_t :

$$\begin{aligned} Q &= E(Z_t Z_t'), \\ &= E(AZ_{t-1}Z_{t-1}'A') + E(BZ_{t-1}Z_{t-1}'B'\varepsilon_t^2) + E(CC'\varepsilon_t^2) + E(CD'\varepsilon_t^3) \\ &\quad + E(DC'\varepsilon_t^3) + DD'E\left((\varepsilon_t^2 - \sigma^2)(\varepsilon_t^2 - \sigma^2)'\right), \\ &= AQA' + BQB'\sigma^2 + CC'\sigma^2 + E(\varepsilon_t^3)(CD' + DC') + DD'E\left((\varepsilon_t^2 - \sigma^2)(\varepsilon_t^2 - \sigma^2)'\right). \end{aligned}$$

D'où :

$$\begin{aligned} Q &= AQA' + BQB'\sigma^2 + \Sigma, \text{ et où} \\ \Sigma &= CC'\sigma^2 + DD'Var(\varepsilon_t^2) + E(\varepsilon_t^3)(CD' + DC'). \end{aligned}$$

Fonction de covariance de X_t :

$$\begin{aligned} R_0 &= E(X_t X_t') = E\left((HZ_{t-1} + \varepsilon_t)(HZ_{t-1} + \varepsilon_t)'\right), \\ &= E(HZ_{t-1}Z_{t-1}'H' + 2HZ_{t-1}\varepsilon_t + \varepsilon_t^2), \\ &= HQH' + \sigma^2. \end{aligned}$$

$R_t = E(X_t X_{t+k})$. On a :

$$\begin{aligned} E(X_t X_{t+1}) &= E(X_t(HZ_t + \varepsilon_{t+1})) = HE(X_t Z_t) \\ E(X_t X_{t+2}) &= E(X_t(HZ_{t+1} + \varepsilon_{t+2})) = HAE(X_t Z_t) \\ E(X_t X_{t+3}) &= E(X_t(HZ_{t+2} + \varepsilon_{t+3})) \\ &= HAE(X_t Z_{t+1}) = HAE(X_t AZ_t) \\ &= HA^2E(X_t Z_t) \end{aligned}$$

⋮

$$E(X_t X_{t+k}) = HA^{k-1}E(X_t Z_t)$$

D'où :

$$\begin{aligned}
 R_t &= E(X_t X_{t+k}) = HA^{k-1}E(X_t Z_t) \\
 &= HA^{k-1}E((AZ_{t-1} + BZ_{t-1} \varepsilon_t + C \varepsilon_t + D (\varepsilon_t^2 - \sigma^2))(HZ_{t-1} + \varepsilon_t)) \\
 &= HA^{k-1}E(AZ_{t-1}^2 H' + BZ_{t-1}^2 H' \varepsilon_t + C H \varepsilon_t Z_{t-1} \\
 &\quad + D H (\varepsilon_t^2 - \sigma^2) Z_{t-1}) + HA^{k-1}E(A Z_{t-1} \varepsilon_t + B Z_{t-1} \varepsilon_t^2 + C \varepsilon_t^2 + D \varepsilon_t (\varepsilon_t^2 - \sigma^2)) \\
 R_t &= HA^k Q H' + HA^{k-1} C \sigma^2 + HA^{k-1} E(\varepsilon_t^3), \quad \forall k \geq 1
 \end{aligned}$$

Moments d'ordre supérieurs

Le calcul des moments d'ordre supérieur est en principe possible.

Néanmoins les calculs sont très fastidieux

On considère la représentation (2.2.1) des modèles bilinéaires, et on suppose Z_t scalaire, on a la représentation suivante :

$$\begin{cases} X_t = Z_{t-1} + \varepsilon_t, \\ Z_t = a_t Z_{t-1} + \eta_t, \end{cases}$$

où : $a_t = A + B \varepsilon_t$, $\eta_t = C \varepsilon_t + D \varepsilon_t^2 + F$ et $F = -D \sigma^2$, alors le moment d'ordre k , μ_k de Z_t , est donné par :

$$\begin{aligned}
 \mu_k &= E(Z_t^k) = E(a_t Z_{t-1} + \eta_t)^k, \\
 &= \sum_{j=0}^k C_k^j E\left((a_t Z_{t-1})^{k-j} (\eta_t)^j\right), \\
 &= \sum_{j=0}^k C_k^j E(Z_{t-1})^{k-j} E\left(a_t^{k-j} (\eta_t)^j\right), \\
 \mu_k &= \sum_{j=0}^k C_k^j \mu_{k-j} w_{k,j}, \quad \text{où } w_{k,j} = E\left(a_t^{k-j} (\eta_t)^j\right),
 \end{aligned}$$

et il peut être calculé récursivement à partir de $\mu_0 = 1$ pour vu que $w_{j,0} \neq 1$, $j = 1, \dots, k$.

Remarque 2.2.2

1. Notons que certains résultats se généralisent au processus $X_t, t \in \mathbb{Z}$ défini par l'équation :

$$X_t = \varepsilon_t + \nu \varepsilon_{t-i} X_{t-j} \text{ et l'on a en particulier : } c(i, j) = \frac{\nu \sigma^4}{1 - \nu}, \text{ et } \forall k \neq i, \forall l \neq j \quad c(k, l) = 0.$$

2. Le calcul des moments d'ordre trois a été généralisé par Berline et Francq (1990) pour les modèles superdiagonaux suivants : $X_t = \varepsilon_t + \sum_{k=1}^{\varphi} \sum_{j=k+1}^P b_{jk} X_{t-j} \varepsilon_{t-k}$

3. Pour les modèles $BL(1, 0, 1, 1)$, Kim et Billard (1989) donnent l'expression des moments d'ordre trois.

4. Si nous considérons le processus défini par l'équation (2.3.1) "superdiagonal d'ordre 1", un calcul rapide montre que l'existence des moments d'ordre k : $E[X_t^k]$ n'est assu-

rée que si $b_{21}^k E[\varepsilon_t^k] < 1$. Ainsi on peut voir que la distribution du processus X_t défini par (2.3.7), "diagonal d'ordre 1", a des moments d'ordre supérieur qui n'existeront pas toujours.

5. Les remarques précédentes permettent de mettre en évidence une propriété fondamentale des modèles bilinéaires qui est la non existence de tous les moments, ce qui, en particulier, fait la difficulté de l'étude de ces modèles.

Cumulants :

L'étude des moments d'ordre supérieur est fondamentale pour une connaissance approfondie des modèles bilinéaires, puisque ceux-ci ne sont pas identifiables seulement par la fonction de covariance.

L'étude des cumulants est aussi d'une très grande richesse pour identifier ces modèles. On dispose pour les cumulants d'ordre trois et quatre d'une relation de récurrence type Yule - Walker, pour les modèles diagonaux et superdiagonaux. Afin d'établir cette relation dans le cas des cumulants d'ordre trois, on considère la représentation suivante des modèles bilinéaires diagonaux et superdiagonaux. Pour cela on réécrit (1, 1, 1) sous la forme suivante :

$$X_t = \sum_{i=1}^P a_i X_{t-i} + \varepsilon_t + \sum_{j=1}^{\max(q,Q)} (b_j + \sum_{k=j}^P b_{jk} X_{t-k}) \varepsilon_{t-j}. \quad (2.2.7)$$

Théorème 2.2.10 (Guégan, Pham, 1992)

Soit le processus $X_t, t \in \mathbb{Z}$ défini par (2.2.7). On note $c(r, s)$ le cumulants d'ordre trois défini par : $c(r, s) = cum(X_t, X_{t-r}, X_{t+r})$.

Alors :

$$c(r, s) = \sum_{j=1}^P a_j c(s, r - j) \quad \text{pour tout } r > s + \max(q, Q). \quad (2.2.8)$$

Remarque 2.2.3

1. La relation (2.2.8) n'est établie que pour des modèles bilinéaires diagonaux et superdiagonaux, car la démonstration repose sur l'expression (2.2.7) qui n'est vraie que pour ces modèles.

2. Le résultat (2.2.8) généralise celui établi par Sesay et Subba Rao (1988, 1991) respectivement pour les modèles $BL(1, 0, 1, 1)$ (voir aussi, Jou et Shaman, (1988)) , et les modèles $BL(p, 0, p, 1)$. On a aussi l'expression des cumulants d'ordre quatre pour ces modèles.

2.3 Etude Probabiliste des Modèles Bilinéaires à partir de l'approche directe

2.3.1 Représentation vectorielle

Cette approche fait appel à une version sous forme vectorielle du modèle (2.1.1). Pour cela, nous nous servons des notations suivantes :

$\mathbf{X}_t = (X_t, X_{t-1}, \dots, X_{t-r+1})'$, $\tilde{\varepsilon}_t = (\varepsilon_t, \varepsilon_{t-1}, \dots, \varepsilon_{t-q})'$, avec $r = \max(p, P)$, A est la matrice $r \times r$,

$$A = \begin{pmatrix} a_1 & a_2 & \cdots & a_p & \cdots & \cdots & a_r \\ 1 & 0 & \cdots & 0 & & \cdots & 0 \\ \vdots & & & \vdots & & & \vdots \\ 0 & 0 & \cdots & 0 & \cdots & 1 & 0 \end{pmatrix},$$

B_j , des matrices $r \times r$, $j = 1, \dots, Q$

$$B_j = \begin{pmatrix} b_{1j} & b_{2j} & \cdots & b_{rj} \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & 0 & 0 \end{pmatrix},$$

et C une matrice $r \times (q+1)$:

$$C = \begin{pmatrix} c_0 & c_1 & \cdots & c_q \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & 0 \end{pmatrix}, c_0 = 1.$$

Alors (2.1.1) se réécrit :

$$\mathbf{X}_t = A\mathbf{X}_{t-1} + C\tilde{\varepsilon}_t + \sum_{j=1}^Q B_j \mathbf{X}_{t-1} \varepsilon_{t-j}. \quad (2.3.1)$$

2.3.2 Existence d'une solution stationnaire stricte, causale et ergodique

Une forme équivalente à (2.3.1) est la suivante :

$$\mathbf{X}_t = A(t) \mathbf{X}_{t-1} + C\tilde{\varepsilon}_t, \quad (2.3.2)$$

$$\text{où } A(t) = \begin{pmatrix} a_1 + \sum_{j=1}^Q b_{1j}\varepsilon_{t-j} & a_2 + \sum_{j=1}^Q b_{2j}\varepsilon_{t-j} & \cdots & \cdots & a_r + \sum_{j=1}^Q b_{rj}\varepsilon_{t-j} \\ 1 & 0 & \cdots & 0 & 0 \\ \vdots & & & & \vdots \\ 0 & 0 & \cdots & 1 & 0 \end{pmatrix}.$$

Ainsi, étudier la solution stricte de (2.1.1) revient à étudier l'existence d'une solution strictement stationnaire de l'équation aux différences (2.3.2) dont l'outil clé est l'exposant de Lyapunov. Définons alors, pour chaque matrice $A = (a_{ij})$, d'ordre $r \times r$, la norme de A par

$$\|A\| := \max_i \left\{ \sum_{j=1}^r |a_{ij}| \right\} \text{ et l'opérateur matriciel } |\cdot|^\delta, \text{ pour chaque } \delta > 0, \text{ par } |A|^\delta := \left[|a_{ij}|^\delta \right].$$

Il est facile de voir que, pour l'opérateur matriciel $|\cdot|^\delta$, où $\delta \in [0, 1]$, et la suite des matrices A_n d'ordre $r \times r$,

$$\left| \sum_n A_n \right|^\delta \leq \sum_n |A_n|^\delta, \quad \left| \prod_n A_n \right|^\delta \leq \prod_n |A_n|^\delta,$$

où la relation matricielle $A \leq B$ signifie que $a_{ij} \leq b_{ij}$ pour tout i et j .

Il peut être montré que si $E \left\{ \ln^+ \left(\sum_{i=1}^r \left| a_i + \sum_{j=1}^Q b_{ij}\varepsilon_{t-j} \right| \right) \right\} < \infty$, l'exposant de Lyapunov

donné par :

$$\gamma := \limsup_{n \rightarrow \infty} n^{-1} E \{ \ln \|A(t) A(t-1) \cdots A(t-n)\| \}$$

existe toujours et de plus,

$$\gamma \leq E \left\{ \ln \left\| \prod_{j=0}^{r-1} A(t-j) \right\| \right\}, \text{ p.s, où pour chaque } x \in \mathbb{R}^1, \ln^+ x := \max \{0, \ln x\}. \text{ Cepen-}$$

dant, la condition $\gamma < 0$ peut être utilisée comme **une condition suffisante** de stationnarité; et comme cette condition demande une limite du produit d'un nombre important de matrices aléatoires, elle est modérément utilisée pour la vérification pratique de la stationnarité. Cela conduit à utiliser une condition un peu plus forte, elle est donnée par le théorème suivant :

Théorème 2.3.1 (Liu, 1990)

Supposons que $E \{\ln^+ |\varepsilon_1|\} < \infty$. Si

$$E \left\{ \ln \left\| \prod_{j=0}^{r-1} A(t-j) \right\| \right\} < 0, \quad (2.3.3)$$

alors il existe une unique solution strictement stationnaire, causale et ergodique de (2.3.2) qui peut être formulée comme la première composante de la somme presque sûre de la série infinie,

$$C\tilde{\varepsilon}_t + \sum_{n=1}^{\infty} \left[\prod_{j=0}^{n-1} A(t-j) \right] C\tilde{\varepsilon}_{t-n},$$

où la matrice C , d'ordre $r \times (q+1)$, ainsi que le vecteur $\tilde{\varepsilon}_t$ sont définis précédemment.

2.3.3 Existence d'une solution stationnaire au second ordre

A noter que la stationnarité du modèle (2.1.1) découle de l'existence d'une solution stationnaire au second ordre, qui est de plus strictement stationnaire, de l'équation (2.3.1). La condition **suffisante** d'existence d'une telle solution est donnée par le prochain théorème qui s'appuie sur les hypothèses suivantes,

$$E\varepsilon_t = 0, E\varepsilon_t^2 = \sigma^2 < \infty, E\varepsilon_t^3 = 0, E\varepsilon_t^4 = \gamma^4 < \infty. \quad (2.3.4)$$

Théorème 2.3.2 (Liu et Brockwell, 1988)

Soit le processus $\{\mathbf{X}_t, t \in \mathbb{Z}\}$ défini par (2.3.1) – (2.3.4), où les matrices A, B_j et C ont été définies précédemment. On suppose de plus que $E \varepsilon_t^{2q} < \infty$ et $E \varepsilon_t^s = 0$ pour tout entier impair $s < q$. Posons :

$$\Gamma = \begin{pmatrix} A \otimes A + \sigma^2 B_1 \otimes B_1 & A \otimes B_2 + B_2 \otimes A & B_2 \otimes B_2 \\ \sigma^2 (A \otimes B_1 + B_1 \otimes A) & \sigma^2 (B_2 \otimes B_1 + B_1 \otimes B_2) & 0_{(p^2)} \\ \sigma^2 (A \otimes A + \gamma^4 B_1 \otimes B_1) & \sigma^2 (A \otimes B_2 + B_2 \otimes A) & \sigma^2 (B_2 \otimes B_2) \end{pmatrix}. \quad (2.3.5)$$

Alors si

$$\lambda = \rho(\Gamma) < 1, \quad (2.3.6)$$

(i) l'équation (2.3.1) a une solution strictement stationnaire de carré intégrable \mathbf{X}_t dont la première composante X_t est une solution stationnaire et ergodique de (2.1.1).

(ii) la solution unique de (2.1.1) est la première composante de

$$\mathbf{X}_t = C \tilde{\varepsilon}_t + \sum_{n=1}^{\infty} \prod_{j=1}^n \Phi(t-j) C \tilde{\varepsilon}_{t-n}, \quad (2.3.7)$$

où $\Phi(t) = A + B_1 \varepsilon_{t-1} + \dots + B_Q \varepsilon_{t-Q}$, $t = 0, \mp 1, \dots$

Remarque 2.3.1

A noter que la condition d'existence d'une solution strictement stationnaire du modèle bilinéaire générale donnée par le théorème (2.3.2) exige la finitude de la variance du bruit quand à la condition donnée par le théorème (2.3.1) n'est pas restreinte c'est-à-dire que la variance du bruit est quelconque (finie ou infinie).

2.3.4 Inversibilité

Considérons le modèle général $BL(p, q, P, Q)$ défini par (2.1.1), et soit les matrices $A(t)$, d'ordre $(r \times r)$, associées avec $r = \max(q, Q)$

$$A(t) = \begin{pmatrix} c_1 + \sum_{j=1}^r b_{j1} X_{t-j} & c_2 + \sum_{j=1}^r b_{j2} X_{t-j} & \cdots & c_r + \sum_{j=1}^r b_{jr} X_{t-j} \\ 1 & 0 & 0 & 0 \\ \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 1 & 0 \end{pmatrix}$$

À partir de la définition de l'inversibilité faible, Liu a obtenu des conditions d'inversibilité pour les processus bilinéaires généraux, qui est en effet identique à celle de l'existence d'une solution stationnaire stricte. Cette condition est donnée dans la proposition suivante :

Proposition 2.3.4 (Liu, (1990 b))

Supposons qu'un processus $(\mathbf{X}_t, t \in \mathbb{Z})$ vérifie les conditions du théorème (2.3.2) et la condition $E(\log |\mathbf{X}_1|) < \infty$. Alors si,

$$E \left(\log \left\| \prod_{j=1}^r A(t-j) \right\| \right) < 0, \quad (2.3.12)$$

le processus vérifiant (2.1.1) est inversible.

Remarque 2.3.2

Pour quelques cas particuliers, on dispose des résultats suivants :

* Granger et Andersen (1978 b) établissent l'inversibilité pour le modèle bilinéaire diagonal pur d'ordre un, $BL(0, 0, 1, 1)$, défini par :

$$X_t = b_{11} X_{t-1} \varepsilon_{t-1} + \varepsilon_t, \quad (2.3.13)$$

La condition suffisante est :

$$2b_{11}^2 \sigma^2 < 1, \quad (2.3.14)$$

* Guégan (1981) donne une condition suffisante d'inversibilité du modèle superdiagonal à un terme $BL(0, 0, 2, 1)$, défini par (2.3.6).

La condition suffisante est :

$$2b_{21}^2 \sigma^2 < 1. \quad (2.3.15)$$

2.3.5 Le modèle $BL(p, 0, p, 1)$

Pour ce modèle, on donne les expressions de covariances, les conditions de stationnarité asymptotique ainsi que les conditions d'inversibilité, et cela après l'avoir présenté sous forme vectorielle. Son expression est telle que

$$X_t + \sum_{i=1}^p a_i X_{t-i} = \varepsilon_t + \sum_{j=1}^p b_{j1} X_{t-j} \varepsilon_{t-1}. \quad (2.3.16)$$

*** Représentation vectorielle :**

En posant :

$$A = \begin{pmatrix} -a_1 & -a_2 & \cdots & -a_p \\ 1 & 0 & \cdots & 0 \\ 0 & 0 & & 0 \end{pmatrix}, B = \begin{pmatrix} b_{11} & b_{21} & \cdots & b_{p1} \\ 0 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \end{pmatrix},$$

$$D' = (1, 0, \dots, 0), H' = (1, 0, \dots, 0) \text{ et } \mathbf{X}_t = (X_t, X_{t-1}, \dots, X_{t-p+1}),$$

la représentation vectorielle, dénotée par $VBL(p)$, du modèle $BL(p, 0, p, 1)$ est comme suit

$$X_t = AX_{t-1} + BX_{t-1}\varepsilon_{t-1} + D\varepsilon_t, \quad (2.3.17a)$$

$$\mathbf{X}_t = H'X_t, \quad (2.3.17b)$$

qui est un cas particulier de la représentation défini en (2.3.1), lorsque $Q = 1$ et $r = p$.

*** Expressions de covariances et conditions de stationnarité :**

Théorème 2.3.5 (Subba Rao, 1981)

Une condition **suffisante** pour que le processus $X_t, t \in \mathbb{Z}$ défini par l'équation (2.3.16), où ε_t est une suite de variables aléatoires indépendantes équidistribuées gaussiennes, centrées, soit stationnaire au second ordre, est donnée par :

$$\rho(A \otimes A + \sigma^2 B_1 \otimes B_1) < 1. \quad (2.3.18)$$

Sous cette condition, si $\mu = E(X_t)$, $V = E(X_t X_t')$ et $c(s) = E((X_{t-s} - \mu)(X_t - \mu)')$,

on a les relations suivantes :

(i) $\mu = (I - A)^{-1} B C,$

(ii) $V = A V A' + B_1 V B_1' + \Delta_1,$

$$\begin{aligned} \text{où } \Delta_1 &= A S B'_1 + B_1 S A' + 2 B_1 C C' B'_1 + C C', \\ \text{avec } S &= A \mu C' + B_1 C C' + C \mu' A' + C C' B'_1. \end{aligned}$$

$$(iii) c(0) = A c(0) A' + B_1 c(0) B'_1 + \Delta_2,$$

$$c(1) = A c(0) + \Delta_3,$$

$$c(s) = A c(s-1) = A^{s-1} c(1), \quad s = 2, 3, \dots \quad (2.3.19)$$

$$\text{où } \Delta_2 = B_1 \mu \mu' B'_1 + A \mu \mu' A' + A S B'_1,$$

$$\text{et } \Delta_3 = A \mu \dot{\mu} + B_1 S - \mu \dot{\mu}.$$

Remarque 2.3.3

1/ Les résultats donnés au théorème (2.3.5) généralisent et confirment ceux obtenus par Granger et Anderson (1978a) pour des cas particuliers.

2/ Dans le cas où $\sigma_e^2 \neq 1$, la condition (2.3.24) devient :

$$\rho(A \otimes A + B \otimes B \sigma_e^2) < 1.$$

3/ Dans le cas où $p \neq P$ et $p = 0$, le modèle $BL(p, 0, p, 1)$ devient $BL(0, 0, P, 1)$ et la condition (2.3.18) se réduit facilement pour ce modèle à :

$$\rho(\sigma^2 B_1 \otimes B_1) < 1. \quad (2.3.25)$$

Ainsi, les expressions de covariances sont données dans la proposition suivante :

Proposition 2.3.6

Sous la condition (2.3.25), si $\mu = E(X_t)$, $V = E(X_t X_t')$ et $c(s) = E((X_{t-s} - \mu)(X_t - \mu)')$,

on a les relations suivantes :

$$(i) \mu = B C,$$

$$(ii) V = B_1 V B'_1 + \Delta_1, \text{ où}$$

$$\text{où } \Delta_1 = -B_1 C_0 C'_0 B'_1 + C_0 C'_0,$$

$$(iii) c(0) = B_1 c(0) B'_1 + \Delta_2,$$

$$c(1) = B_1 S - \mu \dot{\mu},$$

$$c(s) = 0, \quad s = 2, 3,$$

$$\text{où } \Delta_2 = \Delta_1 + B_1 \mu \mu' B'_1 - \mu \mu', \quad (2.3.26)$$

4/ En (1989a), Liu a introduit un théorème d'existence d'une solution strictement stationnaire de ce modèle sous une condition plus simple à vérifier que la condition (2.3.18) :

Théorème 2.3.7 (Liu, 1989)

Considérons l'équation (2.3.17) avec une suite de variables aléatoires $(\varepsilon_t, t \in \mathbb{Z})$ indépendantes identiquement distribuées telle que $E|\varepsilon_t|^\gamma < \infty$ pour un certain $\gamma \in]0, 2]$.

Soit $\Gamma = E(|A + \varepsilon_1 B|^{\gamma/2})$. Alors, si

$$\lambda = \rho(\Gamma) < 1, \quad (2.3.27)$$

(i) L'équation (2.3.17) admet une et une solution strictement stationnaire $\{\mathbf{X}_t, t \in \mathbb{Z}\}$ dont la première composante $\{X_t, t \in \mathbb{Z}\}$ est une solution strictement stationnaire unique, causale et ergodique de (2.3.16).

(ii) L'unique solution strictement stationnaire de (2.3.16) est donnée par la première composante de

$$X_t = C\varepsilon_t + \sum_{i=1}^{\infty} \prod_{j=1}^i (A + \varepsilon_{t-l-j+1}B) C\varepsilon_{t-i}, \quad (2.3.28)$$

où la série qui figure dans le second nombre de (2.3.28) converge absolument presque sûrement.

5/ la condition (2.3.27) fait appel au calcul des valeurs propres d'une matrice d'ordre $p \times p$, alors que la condition (2.3.18) fait appel au calcul des valeurs propres d'une matrice d'ordre $p^2 \times p^2$. Donc, on obtient ici une condition plus simple.

*** Inversibilité :**

En utilisant la définition de Granger et Anderson, décrite précédemment, on obtient une condition suffisante pour l'inversibilité du modèle $VBL(p)$. Considérons le modèle défini par (2.3.17) et soit $\hat{\varepsilon}_t$ l'estimateur de ε_t satisfaisant l'équation aux différences

$$\mathbf{X}_t = H'AX_{t-1} + H'BX_{t-1}\hat{\varepsilon}_{t-1} + H'D\hat{\varepsilon}_t, \quad (2.3.29)$$

du (2.3.17) et (2.3.29), l'erreur de l'estimation vérifie

$$H'D\xi_1(t) = -H'BX_{t-1}\xi_1(t-1),$$

où $\xi_1(t) = \varepsilon_t - \hat{\varepsilon}_t$. Alors par des remplacements successifs de cette dernière, on obtient :

$$\xi_1(t) = \prod_{j=1}^t \xi(t-j) \xi_1(0) \stackrel{s=t-j, \forall j=1, \dots, t}{=} (\xi(s))^t \xi_1(0) \text{ avec } \xi(s) = -\frac{H'BX_s}{H'D}.$$

D'où, en prenant l'espérance et en utilisant le théorème ergodique on obtient

$$\lim_{t \rightarrow \infty} E(\xi_1^2(t)) \leq \lim_{t \rightarrow \infty} E(\xi^2(s))^t \xi_1^2(0), \text{ avec } \xi_1^2(0) \neq 0,$$

dont le terme de droite de cette inégalité tend vers zéro lorsque $t \rightarrow \infty$ si $E(\xi^2(s)) < 1$, ce qui implique que

$$H'BE(X_s X_s')B'H < (H'D)^2. \quad (2.3.30)$$

Etant donné A, B, H et D , on peut évaluer explicitement $V = E(X_s X_s')$.

La condition (2.3.30) est une condition suffisante de l'inversibilité du modèle $VBL(p)$.

2.4 Illustrations

Dans un cadre d'application des résultats présentés avant pour les modèles bilinéaires généraux ainsi pour le modèle bilinéaire $BL(p, 0, p, 1)$ (via l'approche directe), cette partie est consacrée à l'étude probabiliste de quelques cas particuliers des modèles bilinéaires ; à savoir le modèle purement bilinéaire et strictement superdiagonal d'ordre un et le modèle diagonal d'ordre un .

2.4.1 Le modèle $BL(0, 0, 2, 1)$

* Définition 2.4.1

Un modèle purement bilinéaire et strictement superdiagonal d'ordre un noté par $BL(0, 0, 2, 1)$ est une équation aux différences stochastique non linéaire de la forme :

$$X_t = b_{21}X_{t-2}\varepsilon_{t-1} + \varepsilon_t, \quad (2.4.1)$$

où X_t est un processus stochastique défini sur un certain espace de probabilité (Ω, \mathcal{A}, P) , pour tout $t \in \mathbb{Z}$ et où $\{\varepsilon_t, t \in \mathbb{Z}\}$ est une suite de variables aléatoires indépendantes, équidistribuées, centrées, de variance finie σ^2 , définies aussi sur le même espace de probabilité (Ω, \mathcal{A}, P) , telle que ε_t est indépendant de $X_k, k < t$.

* Représentation vectorielle

En posant $\mathbf{X}_t = (X_t, X_{t-1})'$, $\tilde{\varepsilon}_t = \varepsilon_t$ et $A(t) = \begin{pmatrix} 0 & b_{21}\varepsilon_{t-1} \\ 1 & 0 \end{pmatrix}$,

le modèle (2.4.1) se réécrit facilement sous la forme suivante

$$\mathbf{X}_t = A(t) \mathbf{X}_{t-1} + \tilde{\varepsilon}_t.$$

* Existence d'une solution stationnaire stricte, causale et ergodique

la condition suffisante d'existence d'une solution strictement stationnaire, causale et ergodique donnée précédemment pour les modèles bilinéaires généraux par (2.3.3) se réduit pour ce cas particulier à

$$E \ln \|A(t) A(t-1)\| < 0, \quad (2.4.2)$$

dont

$$A(t) A(t-1) = \begin{pmatrix} b_{21}\varepsilon_{t-1} & 0 \\ 0 & b_{21}\varepsilon_{t-2} \end{pmatrix} .\text{et,}$$

$$\|A(t)A(t-1)\| = \max\{|b_{21}\varepsilon_{t-1}|, |b_{21}\varepsilon_{t-2}|\}$$

D'où, la condition (2.4.2) est vérifiée, si pour un certain entier $\delta > 0$,

$$E\left\{[|b_{21}\varepsilon_{t-1}| + |b_{21}\varepsilon_{t-2}|]^\delta\right\} < 1.$$

Et comme ε_t a une variance finie, alors cette condition se réduit à :

$$2|b_{21}|^2\sigma^2 < 1. \tag{2.4.3}$$

*** Existence d'une solution stationnaire au second ordre**

Trouver une condition nécessaire et suffisante d'existence d'une solution non anticipative strictement stationnaire de carré intégrable unique causale de (2.4.1) revient à trouver la condition sous laquelle la solution de ce modèle converge en moyenne quadratique. En effet, par des remplacements successifs en (2.4.1) on obtient :

$$X_t = b_{21}^m X_{t-2m} \left(\prod_{k=1}^m \varepsilon_{t-2k+1} \right) + \sum_{j=1}^{m-1} b_{21}^j \varepsilon_{t-2j} \prod_{k=1}^j \varepsilon_{t-2k+1} + \varepsilon_t.$$

De cette dernière :

$$\begin{aligned} & \lim_{m \rightarrow \infty} \text{var} \left[X_t - \sum_{j=1}^{m-1} b_{21}^j \varepsilon_{t-2j} \prod_{k=1}^j \varepsilon_{t-2k+1} + \varepsilon_t \right] \\ &= \lim_{m \rightarrow \infty} \text{var} \left[b_{21}^m X_{t-2m} \left(\prod_{k=1}^m \varepsilon_{t-2k+1} \right) \right] \\ &= \lim_{m \rightarrow \infty} [(b_{21}^2 \sigma^2)^m \text{var}(X_{t-2m})] \end{aligned}$$

Il est très facile de remarquer que la limite de cette dernière se découle à partir de la $\lim_{m \rightarrow \infty} (b_{21}^2 \sigma^2)^m$ car le processus $(X_t, t \in \mathbb{Z})$ est un processus de carré intégrable, donc $\text{var}(X_{t-2m})$ existe toujours et elle est finie. D'où, pour que

$$\lim_{m \rightarrow \infty} \text{var} \left[X_t - \sum_{j=1}^{m-1} b_{21}^j \varepsilon_{t-2j} \prod_{k=1}^j \varepsilon_{t-2k+1} + \varepsilon_t \right] \rightarrow 0 \text{ lorsque } m \rightarrow \infty, \text{ il suffit que}$$

$$b_{21}^2 \sigma^2 < 1. \tag{2.4.4}$$

En outre, cette condition a été donnée par Guégan (1981) et elle est de plus un cas particulier du résultat trouvé par Liu (1990) pour les modèles bilinéaires généraux étudiés dans le paragraphe 3.

En effet, on a pour ce modèle :

$$p = q = 0, P = 2, Q = 1, \text{ d'où } r = \max(0, 2) = 2,$$

$$A = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, B_1 = \begin{pmatrix} 0 & b_{21} \\ 0 & 0 \end{pmatrix}, C = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \Gamma = \begin{pmatrix} A \otimes A + \sigma^2 B_1 \otimes B_1 & 0 & 0 \\ \sigma^2 (A \otimes B_1 + B_1 \otimes A) & 0 & 0 \\ \sigma^2 (A \otimes A + \gamma^4 B_1 \otimes B_1) & 0 & 0 \end{pmatrix}, \text{ et}$$

$$\rho(\Gamma) = |\sigma b_{21}|.$$

Il en découle que $\rho(\Gamma) < 1$ ssi $b_{21}^2 \sigma^2 < 1$.

Remarque 2.4.1

Lorsque la convergence vers zéro est obtenue, ce modèle a une solution strictement stationnaire causale unique donnée par

$$X_t = \sum_{j \geq 1} b_{21}^j \varepsilon_{t-2j} \prod_{k=1}^j \varepsilon_{t-2k+1} + \varepsilon_t.$$

* Inversibilité

En suivant la définition de l'inversibilité au sens de Granger et Anderson, on obtient une condition suffisante de l'inversibilité pour ce modèle :

Considérons le modèle défini par (2.4.1) et soit $\hat{\varepsilon}_t$ l'estimateur de ε_t satisfaisant l'équation aux différences

$$X_t = b_{21} X_{t-2} \hat{\varepsilon}_{t-1} + \hat{\varepsilon}_t, \tag{2.4.5}$$

Du (2.4.1) et (2.4.5), l'erreur de l'estimation vérifie :

$$\xi_t = -b_{21} X_{t-2} \xi_{t-1},$$

où $\xi_t = \varepsilon_t - \hat{\varepsilon}_t$. Alors le carré de j remplacements successifs de cette dernière, par la propriété de l'ergodicité, donne :

$$\xi_t^2 = \prod_{j=1}^t b_{21}^{2j} X_{t-j-1}^2 \xi_0^2 \stackrel{s=t-j-1, \forall j=\overline{1,t}}{=} (b_{21}^2 X_s^2)^t \xi_0^2. \text{ D'où,}$$

$$\lim_{t \rightarrow \infty} E(\xi_t^2) = \lim_{t \rightarrow \infty} E(b_{21}^2 X_s^2)^t \xi_0^2, \text{ avec } \xi_0^2 \neq 0,$$

dont le terme de droite de cette inégalité tend vers zéro, p.s, lorsque $t \rightarrow \infty$ si $E(b_{21}^2 X_s^2) < 1$,

ce qui implique que $b_{21}^2 E(X_s^2) < 1$, qui se réduit facilement, lorsque $E(X_s^2) = \frac{\sigma^2}{1 - b_{21}^2 \sigma^2}$, à

la condition (2.3.15) trouvée par Guégan (1981).

Remarque 2.4.2

Pham et Tran (1981), Guégan et Pham (1987) ont utilisé une autre définition d'inversibilité, qu'ils ont nommé inversibilité au sens forte. Nous allons expliciter cette définition pour ce cas particulier des modèles bilinéaires :

Ce modèle est défini par : $X_t = b_{21}X_{t-2}\varepsilon_{t-1} + \varepsilon_t$, on pose :

$$Z_t = b_{21}X_{t-1}\varepsilon_t, \quad (2.4.6)$$

alors X_t peut s'écrire : $X_t = Z_{t-1} + \varepsilon_t$.

Jusqu'à présent on supposait connaître les valeurs des paramètres, or en fait, si on note $\theta = (b_{21})$ le paramètre, celui ci doit être estimé quand on ajuste le modèle aux données.

Considérons une valeur quelconque $\tilde{\theta}$ de θ . Le bruit dans l'équation réursive (2.4.6) est alors calculé en fonction de $\tilde{\theta}$, on le note $\varepsilon_{t/z_0}^{\tilde{\theta}}$. On se donne alors une valeur arbitraire z_0 de Z_0 et on définit récursivement $\varepsilon_{t/z_0}^{\tilde{\theta}}$ et Z_{t/z_0} à partir des équations (2.4.6) avec $Z_{0/z_0} = z_0$.

Les équations récursives de Z_t sont telles que :

$$\begin{aligned} Z_t &= \tilde{b}_{21}X_{t-1}(X_t - Z_{t-1}), \\ &= \tilde{b}_{21}X_{t-1}X_t - \tilde{b}_{21}X_{t-1}(\tilde{b}_{21}X_{t-2}X_{t-1} - \tilde{b}_{21}X_{t-2}Z_{t-2}), \\ &= \tilde{b}_{21}X_{t-1}X_t + (-\tilde{b}_{21})\tilde{b}_{21}(X_{t-1}X_{t-2})X_{t-1} + (-\tilde{b}_{21})^2X_{t-1}X_{t-2}(\tilde{b}_{21}X_{t-3}X_{t-2} \\ &\quad - \tilde{b}_{21}X_{t-3}Z_{t-3}). \end{aligned}$$

De proche en proche on obtient pour Z_t l'expression suivante :

$$Z_t = \tilde{b}_{21}X_{t-1}X_t + \sum_{i=1}^{t-1} (-\tilde{b}_{21})^i (\tilde{b}_{21}X_{t-i}X_{t-i-1}) \left\{ \prod_{k=1}^i X_{t-k} \right\} + (-\tilde{b}_{21})^t \left\{ \prod_{k=1}^t X_{t-k} \right\} z_0.$$

D'où l'équation de $\varepsilon_{t/z_0}^{\tilde{\theta}}$ est donnée par :

$$\begin{aligned} \varepsilon_{t/z_0}^{\tilde{\theta}} &= \\ &X_t - X_{t-2}X_{t-1} - \sum_{i=1}^{t-2} \left[(-\tilde{b}_{21})^i \tilde{b}_{21}X_{t-i-2}X_{t-i-1} \prod_{j=1}^i X_{t-j-1} \right] - (-\tilde{b}_{21})^{t-1} \left[\prod_{j=1}^{t-1} X_{t-j-1} \right] z_0. \end{aligned}$$

Une forme forte de l'inversibilité est de montrer, lorsque la condition (2.3.15) est vérifiée, qu'il existe un processus stationnaire $\varepsilon_t^{\tilde{\theta}}$ tel que $\varepsilon_{t/z_0}^{\tilde{\theta}} - \varepsilon_t^{\tilde{\theta}} \rightarrow 0$ en probabilité quand $t \rightarrow \infty, \forall z_0$. En effet, il est très facile de voir que, sous la condition $-\tilde{b}_{21} \log EX_j < 1$,

$\left(-\tilde{b}_{21}\right)^{t-1} \left[\prod_{j=1}^{t-1} X_{t-j-1} \right] z_0$ tend vers zéro lorsque $t \rightarrow \infty$ donc il devient clair de poser

$$\varepsilon_t^{\tilde{\theta}} = X_t - X_{t-2}X_{t-1} - \sum_{i \geq 1} \left[\left(-\tilde{b}_{21}\right)^j \tilde{b}_{21} X_{t-i-2} X_{t-i-1} \prod_{j=1}^i X_{t-j-1} \right]$$

pour assurer la convergence vers zéro de $\varepsilon_{t/z_0}^{\tilde{\theta}} - \varepsilon_t^{\tilde{\theta}}$.

*** Structure d'autocovariance :**

En posant $\mu = E(X_t)$, $V_t = E(X_t^2)$ et $c(h) = E((X_t - \mu)(X_{t-h} - \mu))$, la moyenne ainsi que l'autocovariance sont respectivement données par :

► La moyenne :

Le processus X_t est de moyenne nulle étant donné que

$$\mu = b_{21}E(X_{t-2}\varepsilon_{t-1}) + E(\varepsilon_t) = 0,$$

car ε_{t-1} est indépendant du passé du processus, et en particulier de X_{t-2} , et $E(\varepsilon_t) = 0$ pour tout t .

► L'autocovariance de retard h vaut :

$$E(X_t X_{t-h}) = E(\varepsilon_t \varepsilon_{t-h} + b_{21}^2 X_{t-2} \varepsilon_{t-1} X_{t-h-2} \varepsilon_{t-h-1} + b_{21} \varepsilon_t X_{t-h-2} \varepsilon_{t-h-1} + b_{21} \varepsilon_{t-h} X_{t-2} \varepsilon_{t-1}).$$

· Si $h > 1$, nous avons :

$$E(X_t X_{t-h}) = E(\varepsilon_t) E(\varepsilon_{t-h}) + b_{21}^2 E(\varepsilon_{t-1}) E(X_{t-2} X_{t-h-2} \varepsilon_{t-h-1}) + b_{21} E(\varepsilon_t) E(X_{t-h-2} \varepsilon_{t-h-1}) + b_{21} E(X_{t-2} \varepsilon_{t-h}) E(\varepsilon_{t-1}) = 0.$$

· Si $h = 1$, le regroupement des facteurs est différent pour le dernier terme mais conduit au même résultat :

$$E(X_t X_{t-1}) = E(\varepsilon_t) E(\varepsilon_{t-1}) + b_{21}^2 E(\varepsilon_{t-1}) E(X_{t-2} X_{t-3} \varepsilon_{t-2}) + b_{21} E(\varepsilon_t) E(X_{t-3} \varepsilon_{t-2}) + b_{21} E(X_{t-2}) E(\varepsilon_{t-1}^2) = 0.$$

· En remplaçant h par 0, nous trouvons une relation de récurrence pour la variance marginale $E(X_t^2)$:

$$E(X_t^2) = E(\varepsilon_t^2) + b_{21}^2 E(X_{t-2}^2) E(\varepsilon_{t-1}^2) + 2b_{21} E(\varepsilon_t) E(X_{t-2}) E(\varepsilon_{t-1}),$$

et comme le dernier terme s'annule,

$$V_t = \sigma^2 + b_{21}^2 \sigma^2 V_{t-2}.$$

Il existe une solution stationnaire du second ordre pour ce modèle défini par l'équation (2.4.1)

à condition que $b_{21}^2 \sigma^2 < 1$. La variance marginale vaut alors

$$V = \frac{\sigma^2}{1 - b_{21}^2 \sigma^2}.$$

En conclusion, la structure d'autocovariance est résumée comme suit :

$$c(h) = \begin{cases} \frac{\sigma^2}{1 - b_{21}^2 \sigma^2} \text{ sous la condition } b_{21}^2 \sigma^2 < 1, \text{ si } h = 0, \\ 0, \forall h \geq 1. \end{cases} \quad (2.4.7)$$

Remarque 2.4.3 :

A noter que ce processus se comporte comme un bruit blanc, du point de vue de sa fonction d'autocorrélation.

2.4.2 Le modèle $BL(1, 0, 1, 1)$

* Définition 2.4.3

Un modèle bilinéaire diagonal d'ordre un noté par $BL(1, 0, 1, 1)$ est une équation aux différences stochastique non linéaire de la forme :

$$X_t = a_1 X_{t-1} + b_{11} X_{t-1} \varepsilon_{t-1} + \varepsilon_t, \quad (2.4.8)$$

où X_t est un processus stochastique défini sur un certain espace de probabilité (Ω, \mathcal{A}, P) , pour tout $t \in \mathbb{Z}$ et où $\{\varepsilon_t, t \in \mathbb{Z}\}$ est une suite de variables aléatoires indépendantes, équidistribuées, centrées, de variance finie σ^2 , définies aussi sur le même espace de probabilité (Ω, \mathcal{A}, P) , telle que ε_t est indépendant de $X_k, k < t$.

* Représentation vectorielle

À la première vue, on peut dire que l'équation (2.4.8) s'écrit facilement sous la forme vectorielle $X_t = A(t) X_{t-1} + \varepsilon_t$, où $A(t) = a_1 + b_{11} \varepsilon_{t-1}$. Par ailleurs, une représentation markovienne de ce modèle, analogue à celle donnée par Pham (1985), s'obtient facilement de la manière suivante :

$$\begin{aligned} X_t &= a_1 X_{t-1} + b_{11} X_{t-1} \varepsilon_{t-1} + \varepsilon_t, \\ &= (a_1 + b_{11} \varepsilon_{t-1}) X_{t-1} + \varepsilon_t, \\ &= Z_{t-1} + \varepsilon_t. \end{aligned}$$

où

$$\begin{aligned} Z_t &= (a_1 + b_{11} \varepsilon_t) X_t, \\ &= (a_1 + b_{11} \varepsilon_t) (Z_{t-1} + \varepsilon_t), \\ &= (a_1 + b_{11} \varepsilon_t) Z_{t-1} + (a_1 + b_{11} \varepsilon_t) \varepsilon_t. \end{aligned}$$

D'où, une représentation markovienne de ce modèle est donnée par :

$$\begin{cases} X_t = Z_{t-1} + \varepsilon_t, \\ (a_1 + b_{11} \varepsilon_t)^{-1} Z_t = Z_{t-1} + \varepsilon_t. \end{cases}$$

Dans toute la suite on s'intéressera à l'étude probabiliste de ce modèle à partir de la représentation vectorielle.

*** Existence d'une solution stationnaire stricte, causale et ergodique**

Il est très facile de voir que la condition (2.3.3), donnée précédemment pour les modèles bilinéaires généraux dont $A(t)$ est une matrice d'ordre $r \times r$, se réduit pour ce cas particulier de modèle, où $A(t)$ est un scalaire, à

$$E \ln |a_1 + b_{11}\varepsilon_{t-1}| < 0. \quad (2.4.9)$$

Remarque 2.4.4

Si on s'appuie sur la condition donnée par Liu (1989) pour le modèle $BL(p, 0, p, 1)$, $\rho \left(E \left(|A + \varepsilon_1 B|^{\gamma/2} \right) \right) < 1$, on retrouve facilement la condition (2.4.9) lorsque $A(t)$ est un scalaire et $\gamma = 2$ (car le bruit a une variance finie).

*** Existence d'une solution stationnaire au second ordre**

Ici on a $p = 1, q = 0, P = 1$ et $Q = 1$, d'où $r = \max(1, 1) = 1$,

$$A = a_1, B_1 = b_{11}, D = 1, \Gamma = \begin{pmatrix} a_1^2 + \sigma^2 b_{11}^2 & 0 & 0 \\ 2\sigma^2 a_1 b_{11} & 0 & 0 \\ \sigma^2 a_1^2 + \gamma^4 b_{11}^2 & 0 & 0 \end{pmatrix}.$$

Il en découle que $\rho(\Gamma) < 1$ ssi

$$a_1^2 + \sigma^2 b_{11}^2 < 1, \quad (2.4.10)$$

ce qui est bien la même condition donnée par Pham et Tran (1981).

*** Inversibilité**

Avant, on a montré en utilisant la définition de l'inversibilité au sens de Granger et Anderson que $H'BE(X_s X_s')B'H < (H'D)^2$ est une condition suffisante de l'inversibilité pour le modèle $BL(p, 0, p, 1)$. Donc, lorsque $B = b_{11}$ et $D = 1$, cette condition se réduit pour ce modèle, $BL(1, 0, 1, 1)$, à :

$$b_{11}^2 E(X_s^2) < 1. \quad (2.4.11)$$

Remarque 2.4.5

1°/ Pham et Tran (1981) ont donné une condition forte de l'inversibilité pour ce modèle et elle est donnée par le corollaire suivant :

Corollaire 2.4.6 (Pham et Tran, (1981))

Selon que $|b_{11}| < e^{-E(\log|X_t|)}$ ou $|b_{11}| > e^{-E(\log|X_t|)}$ le modèle (2.4.7) est fortement inversible ou non. Dans le cas où il est inversible i.e ε_t est $F_t(X) -$ mesurable où $\{X_t\}$ est une solution arbitraire de (2.4.7), on a

$$\varepsilon_t = X_t - \sum_{j=1}^{\infty} (-b_{11})^{j-1} (a_1 + b_{11}X_{t-j}) \prod_{k=1}^j X_{t-k}, t \in \mathbb{Z}.$$

2° / La condition donnée par ce corollaire est très forte et on peut la simplifier comme suit :

On a :

$$|b_{11}| > e^{-E(\log|X_t|)} \Leftrightarrow \log |b_{11}| > -E(\log |X_t|).$$

En utilisant les propriétés suivantes :

$$\cdot \log |X_t| = \frac{1}{2} \log X_t^2, \log |b_{11}| = \frac{1}{2} \log b_{11}^2.$$

$$\cdot \text{L'inégalité de Jensen : } E \log (|X_t|) < \log E (|X_t|) \Leftrightarrow \frac{1}{2} E (\log X_t^2) < \frac{1}{2} \log E (X_t^2) \Rightarrow -\frac{1}{2} E (\log X_t^2) > -\frac{1}{2} \log E (X_t^2),$$

on obtient :

$$\log b_{11}^2 > -E \log X_t^2$$

$$\Leftrightarrow \log b_{11}^2 > -\log E (X_t^2)$$

$$\Leftrightarrow \log b_{11}^2 > \log \frac{1}{E (X_t^2)}$$

$$\Leftrightarrow b_{11}^2 E (X_t^2) > 1.$$

Par conséquent, une condition suffisante pour l'inversibilité du modèle défini par (2.4.7) est $b_{11}^2 E (X_t^2) \leq 1$, qui est la même condition trouvée précédemment en (2.4.11) à travers la définition de Granger et Anderson.

* **Expression de covariance :**

Comme ce modèle est un cas particulier du modèle $BL(p, 0, p, 1)$, lorsque $p = 1$, qui a été déjà étudié précédemment, alors on omettra les moindres détails et on ne donnera que les principaux résultats dont :

▷ Les expressions de la moyenne et de la variance asymptotique sont respectivement données par :

La moyenne asymptotique :

On a : $\mu = (I - A)^{-1} BC$

Dans ce cas : $A=a, B=b, C=1$.

$$\text{D'où : } E(X_t) = \frac{b}{1-a}$$

La variance asymptotique :

De même, en remplaçant A par a , B par b , C par 1 et μ par $\frac{b}{1-a}$, on trouve :

$$S = \frac{2ab}{1-a} + 2b$$

$$\Delta_1 = \frac{4a^2b^2}{1-a} + 4ab^2 + 2b^2 + 1$$

$$V = E(X_t^2) = \frac{1+2b^2}{1-a^2-b^2} + \frac{4ab^2}{(1-a^2-b^2)(1-a)}$$

D'où

$$c(0) = V - \mu\mu' = \frac{1+2b^2}{1-a^2-b^2} + \frac{-3a^2b^2 + 4ab^2 + b^4 - b^2}{(1-a)^2(1-a^2-b^2)}$$

▷ L'expression de la covariance est résumée dans le système d'équations linéaires suivants :

$$\cdot c(1) = a_{11} c(0) + \Delta_3, \text{ où } \Delta_3 = (a_{11} - 1)\mu^2 + b_{11}S,$$

$$\cdot c(s) = a_{11} c(s-1) = a_{11}^{s-1} c(1), \forall s \geq 2.$$

2.5 Comparaison entre les Modèles Bilinéaires et d'autres modèles de séries chronologiques

Essentiellement, il existe deux approches pour étudier ou comparer les séries chronologiques, à savoir l'approche dans le domaine du temps dite aussi l'approche temporelle ou l'approche dans le domaine des fréquences dite aussi l'approche spectrale.

Alors, sous le titre de comparaison entre les modèles bilinéaires et d'autres modèles de séries chronologiques, on donne importance dans ce qui suit à la première approche et cela via la structure probabiliste ainsi que la structure de covariance.

2.5.1 Comparaison selon la structure probabiliste

▷ Outre le fait que les modèles bilinéaires sont une généralisation des modèles autorégressif moyenne mobile (*ARMA*), le modèle autorégressif (*AR*) à erreur autorégressif conditionnellement hétéroscédastique (*ARCH*) défini par

$$X_t = aX_{t-1} + \{b_1 1_{\{X_{t-1} < c\}} + b_2 1_{\{X_{t-1} \geq c\}}\} X_{t-1} \varepsilon_t + d\varepsilon_t; t \in \mathbb{N}^+,$$

où $\{\varepsilon_t, t \in \mathbb{N}^+\}$ est une suite de variables aléatoires indépendantes identiquement distribuées absolument continues, a, b_1, b_2, c et d sont des nombres réelles,

peut être vu comme un modèle de mélange entre les modèles bilinéaires et les modèles autorégressif non linéaire à seuil introduit par Tong (1978) dont il présente deux régimes différents expliqués par deux sous modèles bilinéaires avec un point de variation spécifié par le paramètre c .

▷ Une classe de modèles non linéaires qui peut être écrite sous la forme d'un modèle bilinéaire est celle des modèles *GARCH* (p, q) qui peut être formulée, à son tour, à un modèle autorégressif à coefficient aléatoire (*RCA*). On considère le processus $\{x_t\}$ défini par :

$$x_t = \sigma_t z_t, \tag{2.5.1}$$

où $\{z_t\}$ est une suite de variables aléatoires i.i.d, et le processus σ_t satisfait $\sigma_t = \Lambda(y_t)$, pour certain processus sous-jacent y_t vérifiant soit

$$y_t = a_0 + \sum_{i=1}^p a_i y_{t-i} + \sum_{i=1}^q b_i \nu(z_{t-i}),$$

ou

$$y_t = a_0 + \sum_{i=1}^p a_i y_{t-i} + \sum_{i=1}^q c_i \nu(z_{t-i}) y_{t-i},$$

pour des fonctions quelconques $\Lambda : \mathbb{R} \rightarrow \mathbb{R}$ et $\nu : \mathbb{R} \rightarrow \mathbb{R}$. Définons $\varepsilon_t = \nu(z_{t-1})$, ces deux classes de modèles *GARCH* prennent les formes des modèles *ARMA* et *AR - BIL* respectivement dont la combinaison de ces deux classes donne la classe des processus de volatilité avec y_t vérifie un modèle bilinéaire sous diagonal.

Par ailleurs, il existe plusieurs spécifications des modèles *GARCH* où dans chaque cas le processus de volatilité appartient soit à la première classe soit à la deuxième, y compris :

*Asymmetric Power GARCH (Ding, Granger and Engle, 1993 ; Guegan and Diebolt, 1994 ; Hentschel, 1995),

*NGARCH(p, q) (Engle and Ng, 1993),

*LGARCH(p, q) (Robinson, 1991 ; Giraitis et al, 2004),

*Log-GARCH(p, q) (Geweke, 1986 ; Pantula, 1986),

*EGARCH(p, q) (Nelson, 1991).

Ainsi, parmi ceux cités précédemment, les 03 types de modèles qui peuvent être reformulés à la classe des modèles *AR - BIL*, deuxième classe, sont les suivants :

Asymmetric Power GARCH :

Ce modèle vérifie l'équation aux différences stochastique suivante

$$\sigma_t^\mu = w + \sum_{i=1}^p \beta_i \sigma_{t-i}^\mu + \sum_{i=1}^q \alpha_i [|z_{t-i} - b|^\mu 1\{z_{t-i} < b\} + c |z_{t-i} - b|^\mu 1\{z_{t-i} \geq b\}] \sigma_{t-i}^\mu,$$

où $1\{\cdot\}$ dénote la fonction indicatrice. Ici, les paramètres se restreignent à $w > 0$, $\alpha_i \geq 0$, $i = 1, \dots, q$, $\beta_i \geq 0$, $i = 1, \dots, p$, $\mu, c \geq 0$ et $b \in \mathbb{R}$ qui est une forte spécification générale qui contient de nombreux modèles, comme cas spéciaux, proposés dans la littérature. Exemples des modèles inclus sont : les modèles GARCH linéaires, Power GARCH, Asymmetric / GJR GARCH, TS-GARCH/ T-GARCH.

Par conséquent, on retombe facilement sur la deuxième classe en posant : $y_t = \sigma_t^\mu$, $a_0 = w$, $a_i = \beta_i$, $c_i = \alpha_i$.

NGARCH(p, q) :

Ce modèle est donné par :

$$\sigma_t^2 = w + \sum_{i=1}^p \beta_i \sigma_{t-i}^2 + \sum_{i=1}^q \alpha_i (z_{t-i} - b)^2 \sigma_{t-i}^2,$$

où $w > 0$, $\alpha_i \geq 0$, $i = 1, \dots, q$, $\beta_i \geq 0$, $i = 1, \dots, p$, et $b \in \mathbb{R}$.

En posant : $y_t = \sigma_t^2$, $a_0 = w$, $a_i = \beta_i$, $c_i = \alpha_i$, et $\nu(z) = (z - b)^2 \geq 0$, on voit clairement que ce modèle appartient à la deuxième classe citée précédemment.

LGARCH (p, q) :

Ce modèle n'impose pas la positivité de σ , et il est défini comme suit :

$$\sigma_t = w + \sum_{i=1}^p \beta_i \sigma_{t-i} + \sum_{i=1}^p \alpha_i z_{t-i} \sigma_{t-i},$$

De plus, il se considère comme une version restreinte des modèles $ARCH(\infty)$ présenté en Robinson (1991). Sa structure probabiliste peut être reformulée à un modèle $AR - BIL$ en posant : $y_t = \sigma_t$, $a_0 = w$, $a_i = \beta_i$, $c_i = \alpha_i$, et $\nu(z) = z$.

2.5.2 Comparaison selon la structure de covariance

Il est très connu que, d'après (Pham, 1985), la structure de covariance d'un processus admettant une représentation bilinéaire markovienne (2.4.1), c'est-à-dire celle des modèles bilinéaires diagonaux et superdiagonaux, a la même forme que celle d'un processus $ARMA$. Prenons à titre d'exemple le modèle suivant :

* **Le modèle $BL(p, 0, p, 1)$ et son équivalent linéaire $ARMA(p, 1)$:**

Prenons $\gamma(s) = cov(x(t+s), x(t))$, i.e $\gamma(s) = A.\gamma(s-1)$ d'après (2.2.19).

Trouvons alors les équations de Yule - Walker du modèle $BL(p, 0, p, 1)$ défini par :

$$X_t = -a_1 X_{t-1} - \dots - a_p X_{t-p} + b_{11} X(t-1) \varepsilon(t-1) + \dots + b_{p1} X(t-p) \varepsilon(t-1).$$

En multipliant les deux membres de cette dernière par X_{t-h} , $h = 0, 1, 2, \dots, p$, et en prenant l'espérance, on obtient :

$$\gamma(0) = -a_1 \gamma(1) - a_2 \gamma(2) - \dots - a_p \gamma(p),$$

$$\gamma(1) = -a_1 \gamma(0) - a_2 \gamma(1) - \dots - a_p \gamma(p-1),$$

et pour $h=2, \dots, p$, on a, le système d'équations :

$$\begin{cases} \gamma(2) = -a_1 \gamma(1) - a_2 \gamma(0) - \dots - a_p \gamma(p-2), \\ \gamma(3) = -a_1 \gamma(2) - a_2 \gamma(1) - a_3 \gamma(0) - \dots - a_p \gamma(p-3), \\ \vdots \\ \gamma(p) = -a_1 \gamma(p-1) - a_2 \gamma(p-2) - \dots - a_p \gamma(0). \end{cases}$$

Ces équations sont les mêmes que les équations de Yule - Walker pour le modèle $ARMA(p, 1)$. Donc, on peut dire que les deux modèles " $BL(p, 0, p, 1)$ et $ARMA(p, 1)$ " ont la même structure de la fonction de covariance. Cette vérité fondamentale, c'est à dire la similarité célèbre dans la structure de covariance entre les processus bilinéaires et les processus $ARMA$, conduit souvent à une misclassification d'un processus à un autre. Néanmoins, à l'intérieur de cette concordance dans la structure de la fonction de covariance, il existe, pour certains modèles, des différences cachées entre ces deux processus. Prenons à titre d'exemple le modèle suivant :

***Le modèle $PDB(1)$ et son équivalent linéaire le modèle non centré $MA(1)$:**

Un processus $\{X_t, t \in Z\}$ est un modèle bilinéaire purement diagonal d'ordre un, dénoté par $PDB(1)$, s'il vérifié l'équation aux différences suivante :

$$X_t = \theta_1 X_{t-1} \varepsilon_{t-1} + \varepsilon_t, \quad (2.5.2)$$

où $\{\varepsilon_t, t \in Z\}$ est une suite de variables aléatoires indépendantes identiquement distribuées avec $E(\varepsilon_t) = 0$ et $E(\varepsilon_t^2) = \sigma_1^2 < \infty$. L'équivalent linéaire du modèle (2.5.2) est le processus moyenne mobile non centré, dénoté par $MA(1)$, et est donné par :

$$X_t = \beta_0 + \beta_1 u_{t-1} + u_t, \quad (2.5.3)$$

où $u_t, t \in Z$ est tel que $E(u_t) = 0$ et $E(u_t^2) = \sigma_2^2 < \infty$.

Etant donné le modèle (2.5.2), les résultats suivants ont été établis avec $\lambda_1 = \sigma_1 \theta_1$ et $|\lambda_1| < 1$:

$$E(X_t) = \sigma_1^2 \theta_1 = \sigma_1 \lambda_1, \quad (2.5.4)$$

$$\rho_k = \begin{cases} 1, & k = 0, \\ \frac{\sigma_1^2 \theta_1^2 (1 - \sigma_1^2 \theta_1^2)}{1 + \sigma_1^2 \theta_1^2 + \sigma_1^4 \theta_1^4} = \frac{\lambda_1^2 (1 - \lambda_1^2)}{1 + \lambda_1^2 + \lambda_1^4}, & k = 1, \\ 0, & k \geq 2. \end{cases} \quad (2.5.5)$$

De plus, pour le modèle (2.5.3), les résultats suivants ont été établi facilement par Box et al, 1994 :

$$E(X_t) = \beta_0, \quad (2.5.6)$$

$$\rho_k = \begin{cases} 1, & k = 0, \\ \frac{\beta_1}{1 + \beta_1^2}, & k = 1, \\ 0, & k \geq 2, \end{cases} \quad (2.5.7)$$

avec $|\beta_1| < 1$ pour l'inversibilité.

Afin d'obtenir la valeur maximale de ρ_1 pour le modèle (2.5.2), $PDB(1)$, on différencie (2.5.5) qui correspond à λ_1 , on égale le résultat à zéro et on résoud l'équation résultante de λ_1 pour obtenir :

$$1 - 2\lambda_1^2 - 2\lambda_1^4 = 0. \tag{2.5.8}$$

La résolution de (2.5.8) donne $\lambda_1 = \sqrt{0.3660} = \pm 0.6051$. De même, la valeur maximale de ρ_1 pour le modèle (2.5.3), $MA(1)$, est à $\beta_1 = \pm 1$. Le tableau n°1 montre les valeurs calculées de ρ_1 pour les processus $PDB(1)$ et $MA(1)$, et cela en prenant les valeurs de β_1 et λ_1 respectivement. De plus, la figure 1 montre le graphe de ρ_1 à partir des valeurs admissibles de λ_1 pour le processus $PDB(1)$ alors que la figure 2 montre le graphe de ρ_1 à partir des valeurs admissibles de β_1 pour le processus $MA(1)$. En raison d'espace, le tableau 1 est un tableau abrégé (le tableau entier a les valeurs de $\lambda_1 = [-0.99, -0.98, \dots, 0.98, 0.99]$ et $\beta_1 = [-0.99, -0.98, \dots, 0.98, 0.99]$). Le tracé des graphes 1 et 2 a été basé sur le tableau complet qui est disponible en Ohakwe (2008).

PDB(1)		MA(1)	
λ_1	ρ_1	β_1	ρ_1
-0.9	0.0624	-0.9	-0.4972
-0.8	0.1124	-0.8	-0.4878
-0.7	0.1444	-0.7	-0.4698
-0.6	0.1547	-0.6	-0.4412
-0.5	0.1429	-0.5	-0.4000
-0.4	0.1134	-0.4	-0.3448
-0.3	0.0746	-0.3	-0.2752
-0.2	0.0369	-0.2	-0.1923
-0.1	0.0098	-0.1	-0.0990
0.0	0.0000	0.0	0.0000
0.1	0.0098	0.1	0.0990
0.2	0.0369	0.2	0.1923
0.3	0.0746	0.3	0.2752
0.4	0.1134	0.4	0.3448
0.5	0.1429	0.5	0.4000
0.6	0.1547	0.6	0.4412
0.7	0.1444	0.7	0.4698
0.8	0.1124	0.8	0.4878
0.9	0.0624	0.9	0.4972

Table 1 :

Autocorrélation d'ordre 1 pour les modèles PDB(1) et MA(1)

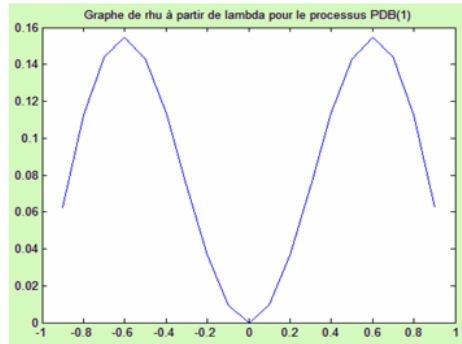


Figure 1

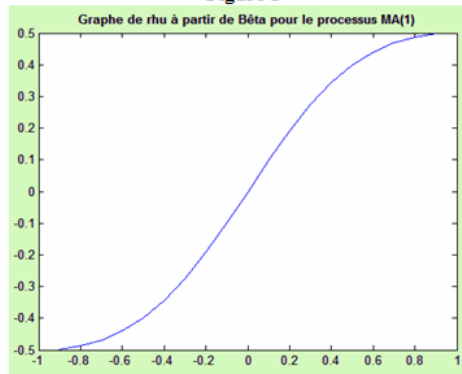


Figure 2

Du tableau 1, figure 1 et figure 2, les dissimilarités suivantes sont évidentes :

1. Intervalles pour ρ_1 : Pour le processus $PDB(1)$, $0 < \rho_1 < 0.16$ tandis que pour le processus $MA(1)$, $|\rho_1| < 0.5$. Cela signifie que le processus $MA(1)$ peut être comparé avec le processus $PDB(1)$ uniquement dans le cas où $0 < \rho_1 < 0.16$.

2. Graphes de ρ_1 (en fonction de λ_1 ou β_1) : Pour le processus $PDB(1)$, le graphe de ρ_1 auprès λ_1 est un graphe ondulatoire avec trois points de retournement $(-0.6051, 0.1547)$, (0.0) , et $(0.6051, 0.1547)$ alors que pour le processus $MA(1)$, le graphe de ρ_1 auprès β_1 ressemble à un arctangent.

Remarque 2.5.1

A noter que le carré du processus $PDB(1)$ s'identifie comme un processus autorégressif moyenne mobile d'ordre $p = 1$ et $q = 1$ alors que le carré du processus non centré $MA(1)$ s'identifie toujours comme un processus $MA(1)$.

2.6 Conclusion

A travers ce chapitre, on a étudié via deux approches différentes des modèles bilinéaires à coefficients constants qui ont la particularité :

- d'être une extension naturelle des modèles linéaires (ARMA) avec un nombre fini de paramètres,
- d'être linéaires par rapport à chacune des variables (entrée et sortie) quand on considère l'autre constante,
- d'avoir une structure ARMA des autocovariances,
- d'avoir un comportement un peu particulier caractérisé par de fortes explosions suivies de longues plages de calme.

Ainsi, cette étude mène à conclure qu'à l'aide de la représentation vectorielle utilisée par *Liu et Brockwell* (1988), puis par *Liu* (1989 – 1990), on dispose de résultats plus généraux et plus simples à vérifier que ceux établis par *Pham* (1985, 1986) et *Guégan* (1987 – 1988) à travers la représentation markovienne. En effet, dans les travaux de ces derniers on ne considère que le cas où le bruit admet une variance finie, alors que *Liu* donne une condition de stationnarité et d'inversibilité pour les modèles bilinéaires généraux sans imposer de condition sur la variance du bruit.

De l'autre côté, le calcul des moments ainsi des moments d'ordre supérieurs est fastidieux et très long en utilisant cette approche tandis que le calcul peut se faire très rapidement et plus facilement en utilisant la représentation markovienne.

Chapitre 3

Les modèles bilinéaires périodiques

3.1 Introduction

Une classe particulière de modèles bilinéaires à coefficients dépendants du temps, rarement étudiés à notre connaissance dans la littérature de l'analyse des séries temporelles non stationnaires, hormis dans Subba Rao (1997), Bibi (2003) et Bibi & Oyet (2004) qui visent essentiellement à décrire des propriétés de nature probabiliste ou statistique de certaines sous-classes de modèles, est la classe des modèles où les coefficients $(a_{i,t})_{0 \leq i \leq p}$, $(b_{j,t})_{0 \leq j \leq q}$, $(c_{ij,t})_{1 \leq i \leq p, 1 \leq j \leq q}$ ainsi que la variance du bruit blanc sont des fonctions périodiques, du temps, de période S ($S \geq 2$). De plus, la classe des modèles bilinéaires périodiques appartient aux processus du second ordre périodiquement corrélés c'est-à-dire que les fonctions d'autocovariances sont périodiques, au sens de Gladyshev (1961), du temps de période S .

Ce chapitre a comme but de présenter quelques propriétés probabilistes des modèles bilinéaires périodiques après être passé en revue les définitions de base qui seront très utiles par la suite.

3.2 Processus périodiquement corrélés et les modèles bilinéaires périodiques

Cette section a comme but de présenter quelques résultats basic qui ont prouvé leurs utilité et opportunité pour la modélisation des séries chronologiques non linéaires périodiques ; à savoir le concept de la corrélation périodique, la propriété de la périodicité stricte ainsi que l'ergodicité périodique ensuite on donnera la structure générale du modèle bilinéaire périodique. Toutefois, l'existence des moments d'ordre un et deux des processus stochastiques permet de trouver quelques propriétés statistiques ; ainsi, la périodicité dans ces deux moments peut être suffisante pour expliquer la propriété de périodicité du processus. Cela conduit au concept de la corrélation périodique.

3.2.1 Processus périodiquement corrélés

Soit $\{X_t, t \in \mathbb{Z}\}$ un processus, du second ordre, de moyenne

$$\mu_t = E(X_t), t \in \mathbb{Z}$$

et de fonction d'autocovariance

$$\gamma(t, s) = E\{(X_t - \mu_t)(X_s - \mu_s)\}; \forall t, s \in \mathbb{Z}.$$

Si notre intérêt ne porte pas sur la non-stationnarité de la moyenne, on peut, sans perte de généralité, supposer que le processus $\{X_t, t \in \mathbb{Z}\}$ est un processus centré, i.e. $\mu_t = 0, \forall t \in \mathbb{Z}$.

Une classe particulière de processus non-stationnaires est celle des processus dont les fonctions d'autocovariance sont des fonctions périodiques du temps. La définition du processus périodiquement corrélé a été introduite, dans la littérature de l'étude des processus aléatoires, par Gladyshev (1961). Elle est donnée par :

Définition 3.2.1

Un processus du second ordre $\{X_t, t \in \mathbb{Z}\}$ est dit périodiquement corrélé ou faiblement périodique s'il existe un entier positif S tel que la moyenne et la fonction d'autocovariance sont S -périodique dans le temps, i.e

- i) $\mu_{t+S\tau} = \mu_t$,
- ii) $\gamma(t + S\tau, s + S\tau) = \gamma(t, s), \forall t, h, \tau \in \mathbb{Z}$,

L'entier positif S permettant de satisfaire ces deux conditions est appelé la période du pro-

cessus périodiquement corrélé. Dans toute la suite, nous supposons que cette période est le plus petit entier positif satisfaisant les deux conditions citées précédemment.

Un simple exemple d'un processus périodiquement corrélé est celui du bruit blanc périodique, qui est une suite de variables aléatoires non corrélées de moyenne zéro et de variance périodique de période S , c'est-à-dire que $\sigma_{s+S\tau}^2 = \sigma_s^2$ pour tout $s = 1, \dots, S, \tau \in \mathbb{Z}$.

3.2.2 Processus strictement périodique

Définition 3.2.2

Un processus du second ordre $\{X_t, t \in \mathbb{Z}\}$ est dit strictement périodique de période $S \in \mathbb{N}^*$ (\mathbb{N}^* indique l'ensemble des entiers positifs) si ses distributions, de dimension finie, sont invariantes sous un multiple shift de S , i.e que la distribution conjointe du vecteur aléatoire $(X_{t_1}, X_{t_2}, \dots, X_{t_n})'$ est la même que celle de $(X_{t_1+S_h}, X_{t_2+S_h}, \dots, X_{t_n+S_h})'$, pour tout entier $n \in \mathbb{N}^*$, $h \in \mathbb{Z}$ et $(t_1, t_2, \dots, t_n) \in \mathbb{Z}^n$.

Remarque 3.2.2

1 -Il est très clair que le concept de la périodicité stricte généralise celui de la stationnarité stricte. En effet, de la définition **3.2.2** il s'ensuit qu'un processus strictement périodique de période $S = 1$ est strictement stationnaire.

2 -Une sous classe, utile, particulière des processus strictement périodiques est la classe où les variables aléatoires sont indépendantes et périodiquement distribuées (notée désormais **i.p.d**).

Définition 3.2.3

Une suite aléatoire $\{\varepsilon_t, t \in \mathbb{Z}\}$ est dite indépendante et périodiquement distribuée de période S , dénotée par (**i.p.d**), si :

i) $\{\varepsilon_t, t \in \mathbb{Z}\}$ est une suite de variables aléatoires indépendantes,

ii) La distribution de probabilité de ε_t est la même que celle de $\varepsilon_{t+S_h} \forall t, h \in \mathbb{Z}$.

Evidemment, une suite (*i.p.d*) de période S est un processus strictement S -périodique. De plus, pour $S = 1$ on retombe sur le cas d'une suite indépendante identiquement distribuée (*iid*).

Comme dans la stationnarité stricte, il est impossible de vérifier pratiquement la périodicité stricte dans les processus stochastiques à l'exception de quelques cas particuliers.

3.2.3 Processus périodiquement ergodique

A noter que la définition d'un processus périodiquement ergodique est intimement liée à la propriété de l'invariance d'un ensemble qui est défini comme suit :

Définition 3.2.4

Soit T une transformation shift et écrivons T^S pour la $S^{\text{ième}}$ puissance de T :

$T^S = T \circ T \circ \dots \circ T$, S fois. Donc,

un ensemble C est appelé S -invariant si et seulement si $T^S X \in C$ pour toutes les séquences doublement infinie de nombres réels, appartenant à C , $X = \{\dots, X_{-1}, X_0, X_1, \dots\} \in \mathbb{R}^{\mathbb{Z}}$. A l'aide de cet instrument, on peut alors citer étroitement la définition de l'ergodicité périodique comme le montre la définition suivante :

Définition 3.2.5

Un processus s.p.s $\{X_t, t \in \mathbb{Z}\}$ est dit périodiquement ergodique si :

$P(\{\dots, X_{-1}, X_0, X_1, \dots\} \in C) = 0$ ou 1 , pour tous les ensembles de Borel S -invariant C de $\mathbb{R}^{\mathbb{Z}}$.

Remarque 3.2.3

Comme dans le cas de la stationnarité stricte (voir par exemple la proposition 6.3.1 donnée par Breiman (1968) ou le théorème 3.6.4 par Billingsley (1995)), la stationnarité strictement périodique et l'ergodicité périodique sont conservées sous certaines transformations. Par exemple si $\{X_t, t \in \mathbb{Z}\}$ est **s.p.s** et périodiquement ergodique, et si $\{Z_t, t \in \mathbb{Z}\}$ est donné par :

$$Z_t = f_t(\dots, X_{t-1}, X_t, X_{t+1}, \dots),$$

où f_t est une fonction de $\mathbb{R}^{\mathbb{Z}}$ à \mathbb{R} qui est mesurable, périodique en t de période S ($f_t = f_{t+nS} \forall n$ et t), alors $\{Z_t, t \in \mathbb{Z}\}$ est aussi **s.p.s** et périodiquement ergodique.

3.2.4 Les modèles bilinéaires périodiques

Définition 3.2.6

On dit qu'un processus périodiquement corrélé $\{X_t, t \in \mathbb{Z}\}$, $\mathbb{Z} = \{0, \pm 1, \pm 2, \dots\}$, du second ordre, défini sur un espace probabilisé (Ω, F, P) est un modèle bilinéaire général périodique, du temps, de période S s'il est solution de l'équation stochastique non linéaire suivante :

$$X_t = a_{t,0} + \sum_{i=1}^p a_{t,i} X_{t-i} + \sum_{j=1}^q c_{t,j} \varepsilon_{t-j} + \sum_{i=1}^P \sum_{j=1}^Q b_{t,ij} X_{t-i} \varepsilon_{t-j} + \varepsilon_t, t \in \mathbb{Z}, \quad (3.2.1)$$

où $(\varepsilon_t)_{t \in \mathbb{Z}}$ est une suite de variables aléatoires indépendantes et périodiquement distribuées (i.p.d), de moyennes nulles et de variances σ_t^2 , définie sur le même espace de probabilité (Ω, F, P) ; et où les coefficients $(a_{t,i})_{0 \leq i \leq p}$, $(c_{t,j})_{0 \leq j \leq q}$, $(b_{t,ij})_{1 \leq i \leq P, 1 \leq j \leq Q}$ ainsi que la variance du bruit blanc sont des fonctions périodiques, du temps, de période S , i.e., $(a_{t+kS,i} = a_{t,i})$, $(c_{t+kS,j} = c_{t,j})$, $(b_{t+kS,ij} = b_{t,ij})$ et $\sigma_{t+kS}^2 = \sigma_t^2 \forall t, k \in \mathbb{Z}$.

3.3 Le modèle PBL(0, 0, P, 1)_S

Définition 3.3.1

On dit qu'un processus périodiquement corrélé $\{X_t, t \in \mathbb{Z}\}$, $\mathbb{Z} = \{0, \pm 1, \pm 2, \dots\}$, du second ordre, défini sur un espace probabilisé (Ω, F, P) , est purement bilinéaire et strictement superdiagonal périodique, du temps, de période S , noté désormais par PBL(0, 0, P, 1)_S, s'il est solution de l'équation stochastique non linéaire suivante :

$$X_t = \sum_{i=2}^P b_{t,i1} X_{t-i} \varepsilon_{t-1} + \varepsilon_t, t \in \mathbb{Z}, \quad (3.3.1)$$

dont $(\varepsilon_t)_{t \in \mathbb{Z}}$ est une suite de variables aléatoires indépendantes et périodiquement distribuées (i.p.d), de moyennes nulles et de variances σ_t^2 , définie sur le même espace de probabilité (Ω, F, P) , telle que ε_l est indépendant de $X_k, k < l$; et où les paramètres $b_{t,i1}, i = 1, \dots, P$ ainsi que σ_t^2 sont périodiques de période S , i.e., $(b_{t+kS,i1} = b_{t,i1})$ et $\sigma_{t+kS}^2 = \sigma_t^2, \forall t, k \in \mathbb{Z}$.

3.3.1 Ecriture vectorielle

Afin d'étudier quelques propriétés probabilistes de ce modèle, on doit répondre à l'étape préliminaire qui consiste à reformuler le modèle sous forme d'une équation aux différences stochastique vectorielle d'ordre un qui est, en effet, une représentation markovienne linéaire de la forme $X_n = A_n X_{n-1} + B_n, \forall n \in \mathbb{Z}$, où (A_n, B_n) est une suite de matrices aléatoires indépendantes ou non selon le cas. Pour cela, on dispose de deux approches différentes ; à savoir celle basée sur la période, introduite par Gladyshev en (1961) sous le nom de "Period Span lumping", et la deuxième approche basée sur l'ordre, introduite par Bentarzi et Hallin en (1994) nommée "Order Span Lumping". Donc, dans un premier temps on montre que la première approche nous conduira à un modèle **RCA** "modèle autorégressif à coefficient aléatoire" tandis que la seconde nous ramènera à un modèle **RCPAR** "modèle autorégressif à coefficient aléatoire périodique". Cette dernière conservant la propriété de la périodicité qui représente l'une de ses particularités, soulèvera notre intérêt par la suite. Le principe de chaque approche est bien expliqué dans ce qui suit :

Period Span Lumping :

Posant $X_\tau(s) = X_{S\tau+s}$ et $\varepsilon_\tau(s) = \varepsilon_{S\tau+s}$, lorsque $t = S\tau + s$, l'équation (3.3.1) devient :

$$X_\tau(s) = \sum_{i=2}^P b_i(s) X_\tau(s-i) \varepsilon_\tau(s-1) + \varepsilon_\tau(s). \quad (3.3.2)$$

En (3.3.2), la notation $X_\tau(s)$ fait référence à la variable X durant la $s^{\text{ième}}$ saison (régime) tel que $s = \{1, 2, \dots, S\}$ du $\tau^{\text{ième}}$ cycle.

Par commodité, $X_\tau(i) = X_{\tau-1}(S+i), \varepsilon_\tau(i) = \varepsilon_{\tau-1}(S+i)$ si $i \leq 0$; à noter aussi que, lorsque $S > 1$, le processus est globalement non stationnaire, mais il est stationnaire à l'intérieur de chaque période.

La version "espace d'état" est alors obtenue par la définition des vecteurs suivants

$$\underline{X}_\tau(s) = (X_\tau(s), \dots, X_\tau(s-P+1))', \quad H = (1, 0, \dots, 0)', \quad \underline{\varepsilon}_\tau(s) = (\varepsilon_\tau(s)) H,$$

$$\text{et la matrice } A_\tau(s) = \begin{pmatrix} 0 & b_2(s) \varepsilon_\tau(s-1) & \dots & b_P(s) \varepsilon_\tau(s-1) \\ I_{P-1} & & \dots & 0_{P-1 \times 1} \end{pmatrix}_{P \times P},$$

Avec ces notations, on réécrit le modèle (3.3.2) sous la forme

$$\underline{X}_\tau(s) = A_\tau(s) \underline{X}_\tau(s-1) + \underline{\varepsilon}_\tau(s), \quad (3.3.3)$$

où $X_\tau(s) = H' \underline{X}_\tau(s)$.

Ainsi, la relation en (3.3.3) est la même que l'équation définie pour un processus autorégressif périodique multivarié. Maintenant, il est très connu que dans les modèles de série chronologique avec des coefficients périodiques, il est possible d'empiler les saisons en un processus multivarié.

Plus précisément $(\underline{X}_\tau)_{\tau \in \mathbb{Z}}$ où $\underline{X}_\tau = (\underline{X}'_\tau(1), \underline{X}'_\tau(2), \dots, \underline{X}'_\tau(S))'$ est un processus **RCA**, i.e.,

$$\underline{X}_\tau = A_\tau \underline{X}_{\tau-1} + B_\tau, \quad (3.3.4)$$

où $(A_\tau, B_\tau)_{\tau \in \mathbb{Z}}$ sont des paires de matrice et vecteur aléatoires strictement stationnaires et ergodiques définis respectivement par bloc comme :

$$A_\tau = \begin{pmatrix} 0_{(p)} & \cdots & 0_{(p)} & A_\tau(1) \\ 0_{(p)} & \cdots & 0_{(p)} & A_\tau(2) A_\tau(1) \\ \vdots & & & \\ 0_{(p)} & \cdots & 0_{(p)} & \prod_{s=0}^{S-1} A_\tau(S-s) \end{pmatrix}_{SP \times SP}, \quad B_\tau = \begin{pmatrix} \varepsilon_\tau(1) \\ A_\tau(2) \varepsilon_\tau(1) + \varepsilon_\tau(2) \\ \vdots \\ \sum_{k=1}^S \left\{ \prod_{s=0}^{S-k-1} A_\tau(S-s) \right\} \varepsilon_\tau(k) \end{pmatrix}_{SP \times 1}$$

où comme d'habitude les produits vides sont posés égaux à $I_{(P)}$.

En effet, le résultat découle facilement en remarquant que

$$\underline{X}_\tau(0) = \begin{pmatrix} X_\tau(0) \\ X_\tau(-1) \\ \vdots \\ X_\tau(-p+1) \end{pmatrix} = \begin{pmatrix} X_{\tau-1}(S) \\ X_{\tau-1}(S-1) \\ \vdots \\ X_{\tau-1}(S-p+1) \end{pmatrix} = \underline{X}_{\tau-1}(S).$$

Order Span Lumping :

Cette approche nécessite la définition des processus P -varié suivant

$$\underline{X}_t = (X_{Pt}, X_{Pt-1}, \dots, X_{Pt-P+1})'; \quad \underline{\varepsilon}_t = (\varepsilon_{Pt}, \varepsilon_{Pt-1}, \dots, \varepsilon_{Pt-P+1})', \quad t \in \mathbb{Z}$$

et les matrices périodiques suivantes

$$A_{t,0} = \begin{pmatrix} 1 & 0 & -\beta_{Pt,2} & -\beta_{Pt,3} & \cdots & -\beta_{Pt,P-2} & -\beta_{Pt,P-1} \\ 0 & 1 & 0 & -\beta_{Pt-1,2} & \cdots & -\beta_{Pt-1,P-3} & -\beta_{Pt-1,P-2} \\ 0 & 0 & 1 & 0 & \cdots & -\beta_{Pt-2,P-4} & -\beta_{Pt-2,P-3} \\ \vdots & \vdots & \vdots & \ddots & & & \vdots \\ & & & \ddots & 0 & & -\beta_{Pt-P+3,2} \\ 0 & \dots & & & 1 & & 0 \\ & & & & 0 & & 1 \end{pmatrix},$$

et

$$A_{t,1} = \begin{pmatrix} \beta_{Pt,P} & 0 & & \dots & & 0 \\ \beta_{Pt-1,P-1} & \beta_{Pt-1,P} & 0 & & & \\ \beta_{Pt-2,P-2} & \beta_{Pt-2,P-1} & \beta_{Pt-2,P} & & & \vdots \\ \vdots & \vdots & \vdots & \ddots & & \\ \beta_{Pt-P+2,2} & \beta_{Pt-P+2,3} & \beta_{Pt-P+2,4} & & \beta_{Pt-P+2,P} & 0 \\ 0 & \beta_{Pt-P+1,2} & \beta_{Pt-P+1,3} & \cdots & \beta_{Pt-P+1,P-1} & \beta_{Pt-P+1,P} \end{pmatrix},$$

où $\beta_{t,i} = b_{t,i}\varepsilon_{t-1}$, dont les termes généraux sont donnés par :

$$(A_{t,0})_{i,j} = \begin{cases} 1 & \text{si } i = j \\ 0 & \text{si } j = i + 1 \\ 0 & \text{si } j < i \\ -\beta_{Pt-i,j-i} & \text{si } j \geq i + 2 \end{cases};$$

$$(A_{t,1})_{i,j} = \begin{cases} \beta_{Pt-i,P} & \text{si } i = j \\ 0 & \text{si } j > i \\ \beta_{Pt-i,P+j-i} & \text{si } j < i \end{cases}, \text{ avec } \beta_{Pt-P+1,1} = 0.$$

Avec les notations précédentes, le modèle (3.3.1) peut être réécrit sous la forme

$$A_{t,0}\underline{X}_t = A_{t,1}\underline{X}_{t-1} + \underline{\varepsilon}_t, \quad (3.3.5)$$

et comme $A_{t,0}$ est une matrice triangulaire supérieure inversible, la représentation (3.3.5) s'écrit également sous la forme d'un processus **RCPAR**

$$\underline{X}_t = A_t \underline{X}_{t-1} + B_t, \quad (3.3.6)$$

où A_t et B_t sont respectivement une matrice, de dimension $P \times P$, et un vecteur colonne de dimension $P \times 1$, donnés par

$$A_t = A_{t,0}^{-1}A_{t,1}, \text{ et } B_t = A_{t,0}^{-1}\underline{\varepsilon}_t, \quad t \in \mathbb{Z}.$$

Remarque 3.3.1

* Dans le cas du modèle $PBL(p, o, p, 1)_S$, on a :

$$X_t = \sum_{i=1}^p a_{t,i} X_{t-i} + \sum_{i=1}^p b_{t,i1} X_{t-i} \varepsilon_{t-1} + \varepsilon_t = \sum_{i=1}^p a_{t,i} X_{t-i} + \sum_{i=1}^p \beta_{t,i1} X_{t-i} + \varepsilon_t$$

$$X_t = \sum_{i=1}^p (a_{t,i} + \beta_{t,i1}) X_{t-i} + \varepsilon_t = \sum_{i=1}^p k_{t,i} X_{t-i} + \varepsilon_t, \text{ qui admet une représentation vectorielle}$$

identique à celle du modèle $PBL(0, 0, P, 1)$ lorsque $P = p$ et $k_{t,i} = a_{t,i} + \beta_{t,i1}$ rien que les termes généraux des matrices $A_{t,0}$ et $A_{t,1}$ sont respectivement donnés par :

$$(A_{t,0})_{i,j} = \begin{cases} 1 & \text{si } j = i \\ 0 & \text{si } j < i \\ -k_{pt-i,j-i} & \text{si } j > i \end{cases} ;$$

$$(A_{t,1})_{i,j} = \begin{cases} k_{pt-i,p} & \text{si } j = i \\ 0 & \text{si } j > i \\ k_{pt-i,p+j-i} & \text{si } j < i. \end{cases}$$

ou encore plus précisément,

$$A_{t,0} = \begin{pmatrix} 1 & -k_{Pt,1} & -k_{Pt,2} & \dots & -k_{Pt,P-2} & -k_{Pt,P-1} \\ 0 & 1 & -k_{Pt-1,1} & \dots & -k_{Pt-1,P-3} & -k_{Pt-1,P-2} \\ 0 & 0 & 1 & -k_{Pt-2,1} & \dots & -k_{Pt-2,P-4} & -k_{Pt-2,P-3} \\ \vdots & \vdots & \vdots & \ddots & & \vdots & \\ & & & \ddots & -k_{Pt-P+3,1} & -k_{Pt-P+3,2} & \\ & & & & 1 & -k_{Pt-P+2,1} & \\ 0 & \dots & & & 0 & 1 & \end{pmatrix},$$

et

$$A_{t,1} = \begin{pmatrix} k_{Pt,P} & 0 & \dots & 0 \\ k_{Pt-1,P-1} & k_{Pt-1,P} & 0 & \vdots \\ k_{Pt-2,P-2} & k_{Pt-2,P-1} & k_{Pt-2,P} & \\ \vdots & \vdots & \vdots & \ddots \\ k_{Pt-P+2,2} & k_{Pt-P+2,3} & k_{Pt-P+2,4} & k_{Pt-P+2,P} & 0 \\ k_{Pt-P+1,1} & k_{Pt-P+1,2} & k_{Pt-P+1,3} & \dots & k_{Pt-P+1,P-1} & k_{Pt-P+1,P} \end{pmatrix},$$

3.3.2 Stationnarité strictement périodique

Cette partie étudie l'existence d'une solution strictement périodique stationnaire (notée **s.p.s**), X_t , au modèle (3.3.1). Notons qu'on est intéressé uniquement par des solutions non anticipatives (e.g. Nicholls et Quinn (1982); Bougerol et Picard (1992a, b) de ce modèle. Ainsi, étudier la stationnarité périodique de la solution de (3.3.1) se traduit immédiatement à étudier la solution de l'équation aux différences stochastique (3.3.6), qui a des coefficients (A_t, B_t) **i.p.d** au lieu des coefficients **i.i.d** (cf. Bougerol et Picard (1992a, b)). Comme pour la stationnarité stricte, l'outil clé dans l'étude de la stationnarité strictement périodique est "le plus grand exposant de Lyapunov" associé avec des matrices aléatoires i.p.d.

Soit $\|\cdot\|$ une norme arbitraire en \mathcal{M}^P , l'espace des matrices carrées réelles de dimension p . Alors, une extension naturelle de la définition standard de l'exposant de Lyapunov au cas

périodique donne :

$$\gamma^S(A) = \inf_{n \in \mathbb{N}^*} \frac{1}{n} E \{ \log \|A_{nS} A_{nS-1} \dots A_1\| \}, \quad (3.3.7)$$

et cela lorsque $\sum_{s=1}^S E(\log^+ \|A_s\|) < \infty$, où $\forall x > 0 \log^+(x) = \max(\log(x), 0)$.

A noter que $\gamma^S(\cdot)$ hérite des propriétés de l'exposant de Lyapunov standard (dédit du (3.3.7), lorsque $S = 1$). Par exemple, la propriété $\gamma^S(A) \leq \sum_{s=1}^S E(\log \|A_s\|)$ est vérifiée avec l'égalité dans le cas scalaire $P = 1$. De plus, si les coefficients $\{(A_t, B_t), t \in \mathbb{Z}\}$ sont périodiquement stationnaires et périodiquement ergodiques (parallèlement th de Brandt (1986)), on peut énoncer le résultat suivant :

Proposition 3.3.1 :

Une condition suffisante pour que le modèle (3.3.1) possède une solution strictement périodique stationnaire est que

$$\gamma^S(A) < 0, \quad (3.3.8)$$

Sous cette condition, l'unique solution s.p.s, non anticipative, et périodiquement ergodique de (3.3.1) est la première composante de la somme presque sûre de la série infinie,

$$\underline{X}_t = B_t + \sum_{n=1}^{\infty} \prod_{i=1}^n A_{t-i} B_{t-n}, t \in \mathbb{Z}. \quad (3.3.9)$$

Preuve :

La démonstration de cette proposition s'appuie sur les 03 fréquentes étapes :

1/ Un simple argument, en utilisant le critère de Cauchy, montre que la série (3.3.9) converge absolument et donc presque sûrement, i.e :

$$\sqrt[n]{\underline{X}_t} = \sqrt[n]{\prod_{i=0}^{n-1} A_{t-i} B_{t-n}} = e^{\frac{1}{n} \log \prod_{i=0}^{n-1} A_{t-i} B_{t-n}} = e^{\frac{1}{n} \sum_{i=0}^{n-1} \log A_{t-i} + \frac{1}{n} \log B_{t-n}}.$$

Vu que :

- $\frac{1}{n} \sum_{i=0}^{n-1} \log A_{t-i}$ converge p.s par le théorème ergodique vers $e^{E \log A_0}$,
- $\frac{1}{n} \log B_{t-n}$ tend vers zéro lorsque n tend vers l'infini, car $E |B_{t-n}| < \infty$.

Alors :

$\sqrt[n]{X_t} \xrightarrow{n \rightarrow \infty} 0 < 1$ puisque $E \log A_0 < 0$. Ceci signifie que la série (3.3.9) converge absolument et donc presque sûrement.

2/ Il est très clair que le processus donné par (3.3.9) est une solution de (3.3.6), i.e :

$$\begin{aligned} \underline{X}_t &= B_t + \sum_{n=1}^{\infty} A_t A_{t-1} \dots A_{t-n+1} B_{t-n} \\ &= B_t + A_t B_{t-1} + A_t A_{t-1} B_{t-2} + \dots + A_t A_{t-1} A_{t-2} B_{t-3} + \dots \\ &= B_t + A_t (B_{t-1} + A_{t-1} B_{t-2} + A_{t-1} A_{t-2} B_{t-3} + \dots) \\ &= B_t + A_t \underline{X}_{t-1}. \end{aligned}$$

3/ Cette solution est clairement s.p.s et périodiquement ergodique puisque c'est une fonction mesurable d'une suite i.p.d $\{(A_t, B_t), t \in \mathbb{Z}\}$ qui est aussi s.p.s et périodiquement ergodique ce qui termine la démonstration. ■

3.3.3 Stationnarité périodique du second ordre

Une condition suffisante pour l'existence d'une solution s.p.s à (3.3.1) satisfaisant $E(X_t^2) < \infty$ est bien donnée par la proposition suivante :

Proposition 3.3.2

Une condition suffisante pour que le modèle (3.3.1) possède une solution $\{X_t, t \in \mathbb{Z}\}$ s.p.s telle que $E(X_t^2) < \infty, \forall 1 \leq t \leq S$ est que

$$\rho \left(\prod_{s=0}^{S-1} E(A_{S-s}^{\otimes 2}) \right) < 1, \quad (3.3.10)$$

de plus, la solution est unique et elle est donnée par (3.3.9) qui converge en moyenne quadratique.

Preuve :

1/ **Trouvons la condition sous laquelle** $\underline{X}_t - B_t - \sum_{j \geq 1} \left(\prod_{k=0}^{j-1} A_{t-k} \right) B_{t-j}$ **converge en**

moyenne quadratique vers zéro :

Rappelons qu'une série X_n converge en moyenne quadratique vers X , si :

$$\lim_{m \rightarrow \infty} E[(X_m - X)^2] = 0.$$

Si on pose $P_{t,n} = \prod_{i=0}^{n-1} A_{t-i}$, $S_{t,m} = \sum_{n=1}^{m-1} P_{t,n} B_{t-n}$ et $R_{t,m} = \prod_{n=0}^{m-1} A_{t-n} X_{t-m}$ où les produits vides sont posés égaux à l'identité, alors par des remplacements successifs en (3.3.6), on obtient pour un certain entier positive m ,

$$X_t = R_{t,m} + S_{t,m} + B_t \quad (*)$$

Alors, du (*), il revient à trouver la condition sous laquelle $R_{t,m}$ converge en moyenne quadratique vers zéro ; i.e :

$$\begin{aligned} & \lim_{m \rightarrow \infty} E \left[\underline{X}_t - B_t - \sum_{n=1}^{m-1} \left(\prod_{k=0}^{n-1} A_{t-k} \right) B_{t-n} \right]^2 \\ &= \lim_{m \rightarrow \infty} E [R_{t,m} R'_{t,m}] \\ &= \lim_{m \rightarrow \infty} E \left[\left(\prod_{n=0}^{m-1} A_{t-n} \right) \underline{X}_{t-m} \underline{X}'_{t-m} \left(\prod_{n=0}^{m-1} A_{t-n} \right)' \right] \\ &= \lim_{m \rightarrow \infty} \text{vec} E \left[\left(\prod_{n=0}^{m-1} A_{t-n} \right) \underline{X}_{t-m} \underline{X}'_{t-m} \left(\prod_{n=0}^{m-1} A_{t-n} \right)' \right] \\ &= \lim_{m \rightarrow \infty} E \left[\prod_{n=0}^{m-1} (A_{t-n}^{\otimes 2}) \text{vec} (\underline{X}_{t-m} \underline{X}'_{t-m}) \right] \end{aligned}$$

Remarquons que la condition suffisante d'existence d'une solution s.p.s de (3.3.1) existe même si la suite des matrices A_t et B_t sont dépendantes. Autrement dit que cette condition n'exige pas l'indépendance entre ces deux suites qui est par analogie entre A_t et \underline{X}_{t-m} . En outre, puisque la matrice A_t n'est pas donnée au cas explicite pour le modèle $PBL(0, 0, P, 1)_S$ à cause de la difficulté du calcul de la matrice inverse $A_{t,0}$, alors, afin d'illustrer cette vérité, on pose $P = 2$. Donc, on a d'une part :

$$\underline{X}_{t-m} \underline{X}'_{t-m} = \begin{pmatrix} X_{2(t-m)}^2 & X_{2(t-m)} X_{2(t-m)-1} \\ X_{2(t-m)-1} X_{2(t-m)} & X_{2(t-m)-1}^2 \end{pmatrix}, \quad A_t = \begin{pmatrix} b_{2t} \varepsilon_{2t-1} & 0 \\ 0 & b_{2t-1} \varepsilon_{2t-2} \end{pmatrix}.$$

Et d'autre part lorsque les deux dernières matrices composant le $\prod_{n=0}^{m-1} (A_{t-n}^{\otimes 2})$ sont comme suit

$$\begin{aligned} A_{t-m+1} &= \begin{pmatrix} b_{2(t-m)+2} \varepsilon_{2(t-m)+1} & 0 \\ 0 & b_{2(t-m)+1} \varepsilon_{2(t-m)} \end{pmatrix}, \\ A_{t-m+2} &= \begin{pmatrix} b_{2(t-m)+4} \varepsilon_{2(t-m)+3} & 0 \\ 0 & b_{2(t-m)+3} \varepsilon_{2(t-m)+2} \end{pmatrix}, \end{aligned}$$

alors il est très facile de remarquer que A_{t-m+1} est la seule matrice dépendante de \underline{X}_{t-m} mais indépendante du produit des matrices qui restent. Cela nous permet d'écrire :

$$\begin{aligned} & \lim_{m \rightarrow \infty} E \left[\prod_{n=0}^{m-1} (A_{t-n}^{\otimes 2}) \text{vec} (\underline{X}_{t-m} \underline{X}'_{t-m}) \right] \\ &= \lim_{m \rightarrow \infty} E \left[\prod_{n=0}^{m-2} (A_{t-n}^{\otimes 2}) \right] E [A_{t-m+1}^{\otimes 2} \text{vec} (\underline{X}_{t-m} \underline{X}'_{t-m})]. \end{aligned}$$

En utilisant la propriété de la périodicité, i.e; lorsque $t = s^{(t)} + S\tau^{(t)}$, $m = s^{(m)} + S\tau^{(m)}$ tel que $s^{(t)}, s^{(m)} \in \{0, \dots, S-1\}$, $\tau^{(t)} \in \mathbb{Z}$ et $\tau^{(m)} \in \mathbb{N}$, on a :

$$\left(\prod_{n=0}^{m-2} A_{t-n} \right) = \left(\prod_{l=0}^{S-1} A_{s^{(t)}-l} \right)^{\tau^{(m)}} \left(\prod_{l=0}^{s^{(m)}-2} A_{s^{(t)}-l} \right) \text{ où } \tau^{(m)} = \left[\frac{m}{S} \right],$$

d'où,

$$\begin{aligned} & \lim_{m \rightarrow \infty} E [R_{t,m} R'_{t,m}] \\ &= \lim_{\tau^{(m)} \rightarrow \infty} \left(\prod_{l=0}^{S-1} E (A_{s^{(t)}-l}^{\otimes 2}) \right)^{\tau^{(m)}} \left(\prod_{l=0}^{s^{(m)}-2} E (A_{s^{(t)}-l}^{\otimes 2}) \right) E (A_{s^{(t)}-s^{(m)}+1}^{\otimes 2} \text{vec} (\underline{X}_{s^{(t)}-s^{(m)}} \underline{X}'_{s^{(t)}-s^{(m)}})) \end{aligned}$$

Il est très facile de remarquer que la limite de cette dernière découle de la $\lim_{\tau^{(m)} \rightarrow \infty} \left(\prod_{l=0}^{S-1} E (A_{s^{(t)}-l}^{\otimes 2}) \right)^{\tau^{(m)}}$

car :

* Le processus $\underline{X}_t, t \in \mathbb{Z}$ est un processus de carré intégrable, donc la quantité :

$$E \left(A_{s^{(t)}-s^{(m)}+1}^{\otimes 2} \text{vec} (\underline{X}_{s^{(t)}-s^{(m)}} \underline{X}'_{s^{(t)}-s^{(m)}}) \right) = \begin{pmatrix} b_{2(s^{(t)}-s^{(m)})+2}^2 \sigma_{2(s^{(t)}-s^{(m)})+2}^2 E \left(X_{2(s^{(t)}-s^{(m)})}^2 \right) \\ 0 \\ 0 \\ b_{2(s^{(t)}-s^{(m)})+1}^2 \sigma_{2(s^{(t)}-s^{(m)})}^2 E \left(X_{2(s^{(t)}-s^{(m)})-1}^2 \right) \end{pmatrix}$$

existe toujours et elle est finie,

** $\prod_{l=0}^{s^{(m)}-2} E (A_{s^{(t)}-l}^{\otimes 2})$ est uniformément borné par $\max_{0 \leq s^{(t)} \leq S-1} \left(\prod_{l=0}^{s^{(m)}-2} E (A_{s^{(t)}-l}^{\otimes 2}) \right)$ ce qui im-

plique que sa limite existe et elle est finie.

D'où, pour que $\lim_{m \rightarrow \infty} E \left[\underline{X}_t - B_t - \sum_{n=1}^{m-1} \left(\prod_{k=0}^{n-1} A_{t-k} \right) B_{t-n} \right]^2 \rightarrow 0$ lorsque $m \rightarrow \infty$, il suffit

que $\rho \left(\prod_{l=0}^{S-1} E (A_{s^{(t)}-l}^{\otimes 2}) \right) < 1$.

Par la propriété circulaire du rayon spectral, il est très facile de remarquer que :

$$\rho \left(\prod_{l=0}^{S-1} E \left(A_{s^{(t)}-l}^{\otimes 2} \right) \right) = \rho \left(\prod_{s=0}^{S-1} E \left(A_{S-s}^{\otimes 2} \right) \right), \forall s^{(t)} \text{ ce qui détermine la condition (3.3.10).}$$

2/ Lorsque la convergence vers zéro est atteinte, l'unique solution strictement stationnaire de carré intégrable est bien donnée par :

$$\underline{X}_t = \sum_{n \geq 1} \left(\prod_{k=0}^{n-1} A_{t-k} \right) B_{t-n} + B_t,$$

qui est clairement une solution de (3.3.6). De plus, puisque c'est une fonction mesurable d'une suite **i.p.d** $\{(A_t, B_t), t \in \mathbb{Z}\}$ alors la solution est aussi *s.p.s* et périodiquement ergodique.

3.3.4 Structure d'autocovariance

Cette section s'occupe du calcul des autocovariances théoriques du processus P – *varié* $PBL(0, 0, P, 1)_S$. Posant $t = s + S\tau$, où $\tau = \lfloor \frac{t-1}{S} \rfloor$, $s = 1, \dots, S$, $\tau \in \mathbb{Z}$, la représentation (3.3.6) s'écrit sous la forme :

$$\underline{X}_{s+S\tau} = A_s \underline{X}_{s+S\tau-1} + B_{s+S\tau}, s = 1, \dots, S, \tau \in \mathbb{Z},$$

$$\text{où } E(B_{s+S\tau} B'_{s-h+S\tau}) = \delta_{h0} \Delta_s, \Delta_s = \begin{pmatrix} \Lambda_0 & 0 & \cdots & 0 \\ 0 & \Lambda_1 & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \cdots & 0 & \Lambda_{p-1} \end{pmatrix},$$

$$\text{et } \Lambda_i = E(\varepsilon_{P(s+S\tau)-i} \varepsilon'_{P(s+S\tau)-i}), i = 0, \dots, p-1.$$

Soit $\gamma_h^{(t)} = E(X_t X'_{t-h})$ et $\Gamma_h^{(t)} = E(\underline{X}_t \underline{X}'_{t-h})$ les fonctions d'autocovariances à l'instant t et retard h de $\{X_t, t \in \mathbb{Z}\}$ et $\{\underline{X}_t, t \in \mathbb{Z}\}$ respectivement.

Alors, $\gamma_h^{(t)}$ et $\Gamma_h^{(t)}$ sont S – *périodique* en t , i.e, $\gamma_h^{(s+S\tau)} = \gamma_h^{(s)}$ et $\Gamma_h^{(s+S\tau)} = \Gamma_h^{(s)}$, $s = 1, \dots, S$, et on a

$$\Gamma_h^{(s)} = \begin{pmatrix} \gamma_{Ph}^{(Ps)} & \gamma_{Ph+1}^{(Ps)} & \cdots & \gamma_{Ph+P-1}^{(Ps)} \\ \gamma_{Ph-1}^{(Ps-1)} & \gamma_{Ph}^{(Ps-1)} & \cdots & \gamma_{Ph+P-2}^{(Ps-1)} \\ \vdots & & \ddots & \\ \gamma_{Ph-P+1}^{(Ps-P+1)} & \gamma_{Ph-P+2}^{(Ps-P+1)} & \cdots & \gamma_{Ph}^{(Ps-P+1)} \end{pmatrix},$$

$$\text{avec la propriété } \gamma_{-Ph}^{(Ps)} = \left(\gamma_{Ph}^{(Ps+h)} \right)'$$

Il est possible d'obtenir les autocovariances du modèle $PBL(0, 0, P, 1)_S$ via deux procédures différentes :

* La première procédure est une simple méthode basée sur l'approche de l'équation aux différences, qui nécessite la résolution d'un système par bloc, obtenu à partir des équations de Yule-Walker périodiques. En effet, ces équations peuvent être obtenues en multipliant les deux côtés de l'équation (3.3.1) par X_{t-h} , $h = 0, \dots, P$, et en prenant l'espérance, on obtient S systèmes linéaires des équations (d'ordre $P+1$) de la forme :

$$\gamma_h^{(s)} = \begin{cases} \sigma_{s-1}^2 \underline{b}_s \Gamma_s \underline{b}'_s + \Lambda_0, & h = 0 \\ 0, & h \in \mathbb{N}^* \end{cases}, s = 1, \dots, S. \quad (3.3.11)$$

où $\underline{b}_s = (b_{s,21}, b_{s,31}, \dots, b_{s,P1})$, $(\Gamma_s)_{lk} = \gamma_{l-k}^{(s-k)}$, $l, k = 2, \dots, P$. Alors, le calcul des autocovariances initiales exige la résolution du système précédent, qui est d'ordre $S(P+1)$, associé à toutes les saisons. A noter que cette méthode est très coûteuse dans les calculs notamment lorsqu'il s'agit des modèles avec une période importante comme dans le cas où les données sont mensuelles (*i.e* $S = 12$), c'est pour cela que la prochaine méthode vient pour faciliter l'opération des calculs et pour restreindre le temps de cette dernière.

** La deuxième procédure est basée sur la représentation espace d'état des modèles $PBL(0, 0, P, 1)_S$ donnée par l'équation (3.3.6), utilisée récemment par *Aknouche* (2007) pour calculer les autocovariances de démarrage des modèles $PVAR$ purs et où les autocovariances de chacune des saisons peuvent être calculées séparément, ce qui permet d'alléger le poids du calcul surtout pour les modèles de grandes périodes. En effet, en multipliant les deux côtés de l'équation (3.3.6) par \underline{X}_{t-h} , $h = 0, \dots, P$, et en prenant l'espérance, on obtient :

$$\begin{cases} \Gamma_0^{(s)} = E \left(A_s \underline{X}_{s+S\tau-1} \underline{X}'_{s+S\tau} \right) + E \left(B_{s+S\tau} \underline{X}'_{s+S\tau} \right), & (3.3.12.a) \\ \Gamma_h^{(s)} = 0, \quad \forall h \in \mathbb{N}^*, & (3.3.12.b) \end{cases}$$

Et à cause de la non disponibilité de la matrice A_s pour le modèle $PBL(0, 0, P, 1)_S$, on ne peut donner l'expression d'autocovariance que pour des cas particuliers des modèles purement bilinéaires et strictement superdiagonaux comme on va le montrer, en détail, ultérieurement dans le cas où $P = 2$.

Remarque 3.3.2

Concernant le modèle bilinéaire superdiagonal périodique, $PBL(p, 0, p, 1)_S$, le calcul est fastidieux par rapport au modèle purement bilinéaire et strictement superdiagonal périodique $PBL(0, 0, P, 1)_S$ dont :

Le calcul de la moyenne et de la variance donne :

▷ $\mu_s = \underline{a}_s \underline{\mu} + b_{s,1} \sigma_{s-1}^2$, $s = 1, \dots, S$, où \underline{a}_s et $\underline{\mu}$ sont respectivement des vecteurs lignes et colonnes donnés par $\underline{a}_s = (a_{s,1}, \dots, a_{s,p})_{1 \times p}$, $\underline{\mu} = (\mu_{s-1}, \dots, \mu_{s-p})'_{p \times 1}$.

$$\begin{aligned} \triangleright \gamma_0^{(s)} &= (a_{s,1}, \dots, a_{s,p}) \Psi_s^{(p,p)} (a_{s,2}, \dots, a_{s,p})' \\ &+ \sigma_{s-1}^2 (b_{s,21}, \dots, b_{s,p1}) K_s^{(p-1,p-1)} (b_{s,21}, \dots, b_{s,p1})' \\ &+ 4a_{s,1} b_{s,11} \sigma_{s-1}^2 \left[(a_{s-1,1}, \dots, a_{s-1,p}) (\mu_{s-2}, \dots, \mu_{s-p-1})' + b_{s-1,11} \sigma_{s-2}^2 \right] \\ &+ 2\sigma_{s-1}^2 [a_{s,1} (b_{s,21}, \dots, b_{s,p1}) + b_{s,11} (a_{s,2}, \dots, a_{s,p})] (\mu_{s-2}, \dots, \mu_{s-p})' \\ &+ 2b_{s,11} \sigma_{s-1}^2 (a_{s-1,1}, \dots, a_{s-1,p}) \Theta_s^{(p,p-1)} (b_{s,21}, \dots, b_{s,p1})' \end{aligned}$$

$$\begin{aligned}
 & +2b_{s,11}b_{s-1,11}\sigma_{s-1}^2\sigma_{s-2}^2(\mu_{s-3}, \dots, \mu_{s-p-2})(b_{s,31}, \dots, b_{s,p1})' \\
 & +2b_{s,11}b_{s,21}\sigma_{s-1}^2\sigma_{s-2}^2\left\{(b_{s-1,21}, \dots, b_{s-1,p1})(\mu_{s-3}, \dots, \mu_{s-p-1})'\right. \\
 & \left.+2b_{s-1,11}\left[(a_{s-2,1}, \dots, a_{s-2,p})(\mu_{s-3}, \dots, \mu_{s-p-2}) + b_{s-2,11}\sigma_{s-3}^2\right]\right\} + \sigma_s^2 \\
 & + \left(1 - \prod_{j=1}^S (b_{s-j,11}^2\sigma_{s-j}^2)\right)^{-1} \times \\
 & \left[\sum_{j=1}^m \prod_{l=1}^j (b_{s-l+1,11}^2\sigma_{s-l}^2) \sigma_{s-j-1,1}^2 \left\{ 4a_{s-j,1}b_{s-j,11} \left(\sum_{k=1}^p a_{s-j-1,k}\mu_{s-j-1-k} + b_{s-j-1,11}\sigma_{s-j-2}^2 \right) \right. \right. \\
 & + 2 \sum_{k=2}^p (a_{s-j,1}b_{s-j,k1} + b_{s-j,11}a_{s-j,k}) \mu_{s-j-k} \\
 & + 2b_{s-j,11} \left((a_{s-j-1,1}, \dots, a_{s-j-1,p}) \Gamma_{s-j+1}^{(p,p-1)} (b_{s-j,21}, \dots, b_{s-j,p1})' \times \right. \\
 & \left. + 2b_{s-j,21}b_{s-j-1,11}\sigma_{s-j-2}^2 \left[\sum_{k=1}^p a_{s-j-2,k}\mu_{s-j-2-k} + b_{s-j-2,11}\sigma_{s-j-3}^2 \right] \right. \\
 & \left. + b_{s-j,21}\sigma_{s-j-2}^2 \left(\sum_{k=2}^p b_{s-j-1,k1}\mu_{s-j-1-k} \right) + 2b_{s-j-1,11}\sigma_{s-j-2}^2 \left(\sum_{k=3}^p b_{s-j,k1}\mu_{s-j-k} \right) \right) \\
 & \left. + \sum_{k=2}^p b_{s-j,k1}^2 \Gamma_{s-j+1}^{(p-1,p-1)} \right\} \\
 & \left. + \sum_{j=1}^m \prod_{l=1}^j (b_{s-l+1,11}^2\sigma_{s-l}^2) \sum_{k=1}^p a_{s-j,k}^2 \Phi_{s-j+1}^{(p,p)} + b_{s,11}^2 \sum_{j=1}^m \prod_{l=1}^{j-1} (b_{s-l,11}^2\sigma_{s-l}^2) E(\varepsilon_{s-j}^4) \right] ,
 \end{aligned}$$

où les matrices $\Psi_s, K_s, \Theta_s, \Gamma_s, \Phi_s$ sont comme suit :

$$\Psi_s = \begin{pmatrix} \gamma_0^{(s-1)} & \cdots & \gamma_{1-p}^{(s-p)} \\ \gamma_1^{(s-1)} & \cdots & \gamma_{2-p}^{(s-p)} \\ \vdots & \ddots & \vdots \\ \gamma_{p-1}^{(s-1)} & \cdots & \gamma_0^{(s-p)} \end{pmatrix}_{p \times p}, \quad K_s = \begin{pmatrix} \gamma_0^{(s-2)} & \cdots & \gamma_{2-p}^{(s-p)} \\ \gamma_1^{(s-2)} & \cdots & \gamma_{3-p}^{(s-p)} \\ \vdots & \ddots & \vdots \\ \gamma_{p-2}^{(s-2)} & \cdots & \gamma_0^{(s-p)} \end{pmatrix}_{p-1 \times p-1},$$

$$\Theta_s = \begin{pmatrix} \gamma_0^{(s-2)} & \cdots & \gamma_{2-p}^{(s-p)} \\ \gamma_1^{(s-2)} & \cdots & \gamma_{3-p}^{(s-p)} \\ \vdots & \ddots & \vdots \\ \gamma_{p-1}^{(s-2)} & \cdots & \gamma_1^{(s-p)} \end{pmatrix}_{p \times p-1}, \quad \Gamma_s = \begin{pmatrix} \gamma_0^{(s-3)} & \cdots & \gamma_{2-p}^{(s-p-1)} \\ \gamma_1^{(s-3)} & \cdots & \gamma_{3-p}^{(s-p-1)} \\ \vdots & \ddots & \vdots \\ \gamma_{p-1}^{(s-3)} & \cdots & \gamma_1^{(s-p-1)} \end{pmatrix}_{p \times p-1}$$

$$\text{et } \Phi_s = \begin{pmatrix} \gamma_0^{(s-2)} & \cdots & \gamma_{1-p}^{(s-p-1)} \\ \gamma_1^{(s-2)} & \cdots & \gamma_{2-p}^{(s-p-1)} \\ \vdots & \ddots & \vdots \\ \gamma_{p-1}^{(s-2)} & \cdots & \gamma_0^{(s-p-1)} \end{pmatrix}_{p \times p}.$$

Ainsi, les autocovariances de ce modèle peuvent être obtenues en résolvant le système suivant :

$$\begin{aligned}
 * \gamma_1^{(s)} &= (a_{s,1}, a_{s,2}, \dots, a_{s,p}) \left(\gamma_0^{(s-1)}, \gamma_1^{(s-1)}, \dots, \gamma_p^{(s-1)} \right)' + 2b_{s,11} \sigma_{s-1}^2 [(a_{s-1,1}, a_{s-1,2}, \dots, a_{s-1,p}) \\
 &\quad (\mu_{s-2}, \mu_{s-3}, \dots, \mu_{s-p-1})' + b_{s-1,11} \sigma_{s-2}^2] + \sigma_{s-1}^2 (b_{s,21}, \dots, b_{s,p1}) (\mu_{s-2}, \dots, \mu_{s-p})' \\
 * \gamma_2^{(s)} &= (a_{s,1}, a_{s,2}, \dots, a_{s,p}) \left(\gamma_{-1}^{(s-2)}, \gamma_0^{(s-2)}, \dots, \gamma_{p-2}^{(s-2)} \right)' + b_{s,11} \sigma_{s-1}^2 \mu_{s-2} \\
 &\vdots \\
 * \gamma_p^{(s)} &= (a_{s,1}, a_{s,2}, \dots, a_{s,p}) \left(\gamma_{1-p}^{(s-p)}, \gamma_{2-p}^{(s-p)}, \dots, \gamma_0^{(s-p)} \right)' + b_{s,11} \sigma_{s-1}^2 \mu_{s-p} \\
 &\vdots \\
 * \gamma_j^{(s)} &= (a_{s,1}, a_{s,2}, \dots, a_{s,p}) \left(\gamma_{1-j}^{(s-j)}, \gamma_{2-j}^{(s-j)}, \dots, \gamma_{p-j}^{(s-j)} \right)' + b_{s,11} \sigma_{s-1}^2 \mu_{s-j}, j \geq 2 \\
 \gamma_j^{(s)} &= (a_{s,1}, a_{s,2}, \dots, a_{s,p}) \left(\gamma_{1-j}^{(s-j)}, \gamma_{2-j}^{(s-j)}, \dots, \gamma_{p-j}^{(s-j)} \right)' + b_{s,11} \sigma_{s-1}^2 (a_{s-j,1}, a_{s-j,2}, \dots, a_{s-j,p}) \times \\
 &\quad (\mu_{s-j-1}, \mu_{s-j-2}, \dots, \mu_{s-j-p})', j \geq 2
 \end{aligned}$$

3.3.5 Les moments d'ordre supérieurs

Comme pour les modèles bilinéaires classiques, un simple argument en utilisant le binôme de Newton nous permet d'écrire l'expression des moments d'ordre r pour les modèles bilinéaires superdiagonaux périodiques, d'ordre P quelconque, sous les formes suivantes :

· Pour le modèle $PBL(0, 0, P, 1)$:

$$E(X_t^r) = \begin{cases} E(\varepsilon_t)^r + \sum_{j=1}^{r/2} C_r^{2j} E\left(\sum_{i=2}^P b_{t,i1} X_{t-i}\right)^{2j} E(\varepsilon_{t-1})^{2j} E(\varepsilon_t)^{r-2j} & \text{r pair} \\ 0 & \text{r impair} \end{cases} \quad (3.3.13)$$

· Pour le modèle $PBL(p, 0, p, 1)$:

$$\begin{aligned}
 E(X_t^r) &= E(\varepsilon_t)^r + \sum_{j=1}^r C_r^j E \left[\left(\sum_{i=1}^p X_{t-i} b_{t,i1} \varepsilon_{t-1} \right)^j + \sum_{l=1}^j C_j^l \left(\sum_{i=1}^p X_{t-i} a_{t,i} \right)^l \times \right. \\
 &\quad \left. \left(\sum_{i=1}^p X_{t-i} b_{t,i1} \varepsilon_{t-1} \right)^{j-l} \right] E(\varepsilon_t)^{r-j}. \quad (3.3.14)
 \end{aligned}$$

3.4 Illustrations

Cette partie a comme but de faire une petite application des résultats présentés dans la section précédente pour le modèle **RCPAR** général, et cela sur des cas particuliers des modèles bilinéaires périodiques ; à savoir le modèle purement bilinéaire et strictement superdiagonal d'ordre un périodique et le modèle bilinéaire diagonal d'ordre un périodique .

3.4.1 Le modèle $\text{PBL}(0, 0, 2, 1)_S$

Définition 3.4.1

On dit qu'un processus périodiquement corrélé $\{X_t, t \in \mathbb{Z}\}$, $\mathbb{Z} = \{0, \pm 1, \pm 2, \dots\}$, du second ordre, défini sur un espace probabilisé (Ω, F, P) , est purement bilinéaire et strictement superdiagonal **d'ordre un** périodique, du temps, de période S , noté désormais par $\text{PBL}(0, 0, 2, 1)_S$, s'il est solution de l'équation stochastique non linéaire suivante :

$$X_t = b_{t,21}X_{t-2}\varepsilon_{t-1} + \varepsilon_t, \quad t \in \mathbb{Z}, \quad (3.4.1)$$

dont $(\varepsilon_t)_{t \in \mathbb{Z}}$ est une suite de variables aléatoires indépendantes et périodiquement distribuées (i.p.d), de moyenne nulle et de variance σ_t^2 , définie sur le même espace de probabilité (Ω, F, P) , telle que ε_l est indépendant de $X_k, k < l$; et où le paramètre $b_{t,21}$ ainsi que σ_t^2 sont périodiques de période S , i.e, $b_{t+kS,21} = b_{t,21}$ et $\sigma_{t+kS}^2 = \sigma_t^2, \forall t, k \in \mathbb{Z}$.

Écriture Vectorielle

Comme il a été vu précédemment, l'écriture du modèle sous forme d'une équation aux différences stochastique d'ordre un en utilisant la deuxième approche nécessite la définition des processus 2 – varié suivant :

$$\underline{X}_t = (X_{2t}, X_{2t-1})'; \quad \underline{\varepsilon}_t = (\varepsilon_{2t}, \varepsilon_{2t-1})', \quad t \in \mathbb{Z},$$

et les matrices périodiques suivantes

$$A_{t,0} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \text{ et } A_{t,1} = \begin{pmatrix} b_{2t,2}\varepsilon_{2t-1} & 0 \\ 0 & b_{2t-1,2}\varepsilon_{2t-2} \end{pmatrix},$$

Avec les notations précédentes, le modèle (3.4.1) peut être réécrit sous la forme

$$A_{t,0}\underline{X}_t = A_{t,1}\underline{X}_{t-1} + \underline{\varepsilon}_t, \quad (3.4.2)$$

et comme $A_{t,0}$ est une matrice inversible, la représentation (3.4.2) s'écrit également sous la forme d'un processus **RCPAR** "autorégressif à coefficient aléatoire périodique" suivant

$$\underline{X}_t = A_t \underline{X}_{t-1} + B_t, \quad (3.4.3)$$

où A_t et B_t sont respectivement une matrice, de dimension 2×2 , et un vecteur colonne de dimension 2×1 , donnés par :

$$A_t = A_{t,0}^{-1} A_{t,1} = A_{t,1}, \text{ et } B_t = A_{t,0}^{-1} \underline{\varepsilon}_t = \underline{\varepsilon}_t, t \in \mathbb{Z}.$$

Existence d'une solution stationnaire stricte

En suivant le raisonnement de Liu (1990), i.e lorsque les propriétés suivantes sont vérifiées

$$* \|A\| := \max_i \left\{ \sum_{j=1}^p |a_{ij}| \right\},$$

$$** \gamma := E \left\{ \ln \left\| \prod_{j=0}^{p-1} A(t-j) \right\| \right\},$$

la condition (2.3.3), en exploitant la propriété de la périodicité, se réduit dans ce cas particulier comme suit :

$$\begin{aligned} E \log \left\| \prod_{s=1}^S A_{nS-s} \right\| &= E \log \left\| \prod_{s=1}^S A_{S-s} \right\| \\ &= E \log \left\| \prod_{s=1}^S A_{-s} \right\| = E \log \left\| \prod_{s=1}^S A_s \right\|, \end{aligned} \quad (3.4.4)$$

et lorsque le produit de ces matrices est égaux à

$$\left(\prod_{s=1}^S A_s \right) = \begin{pmatrix} \prod_{s=0}^{S-1} (b_{2s,21} \varepsilon_{2s-1}) & 0 \\ 0 & \prod_{s=0}^{S-1} (b_{2s-1,21} \varepsilon_{2s-2}) \end{pmatrix},$$

alors la norme de ce produit est bien donnée par

$$\left\| \prod_{s=1}^S A_s \right\| = \max \left\{ \left| \prod_{s=0}^{S-1} (b_{2s,21} \varepsilon_{2s-1}) \right|, \left| \prod_{s=0}^{S-1} (b_{2s-1,21} \varepsilon_{2s-2}) \right| \right\}.$$

Enfin, il est très facile de voir que l'exposant de Lyapunov est strictement négatif si pour un

$$\text{entier } \delta > 0 : E \left\{ \left| \prod_{s=0}^{S-1} (b_{2s,21} \varepsilon_{2s-1}) \right| + \left| \prod_{s=0}^{S-1} (b_{2s-1,21} \varepsilon_{2s-2}) \right| \right\}^\delta < 1.$$

Et comme ε_t a une variance finie, alors cette dernière se réduit à

$$\left\{ \left| \prod_{s=1}^S b_{2s,21}^2 \sigma_{2s-1}^2 \right| + \left| \prod_{s=1}^S b_{2s-1,21}^2 \sigma_{2s-2}^2 \right| \right\} < 1,$$

qui est la même condition (2.4.3) trouvée par **Liu** (1990) lorsque $S = 1$.

Existence d'une solution stationnaire au second ordre

Trouver la condition de la stationnarité du second ordre périodique pour le modèle (3.4.3) revient à trouver le rayon spectral du produit des espérances de la quantité suivante :

$$A_s^{\otimes 2} = \begin{pmatrix} b_{2s,21}^2 \varepsilon_{2s-1}^2 & 0 & 0 & 0 \\ 0 & b_{2s,21} b_{2s-1,21} \varepsilon_{2s-1} \varepsilon_{2s-2} & 0 & 0 \\ 0 & 0 & b_{2s-1,21} b_{2s,21} \varepsilon_{2s-2} \varepsilon_{2s-1} & 0 \\ 0 & 0 & 0 & b_{2s-1,21}^2 \varepsilon_{2s-2}^2 \end{pmatrix}$$

i.e $\rho \left(\prod_{s=1}^S E(A_s^{\otimes 2}) \right)$ qui est clairement inférieur à 1 lorsque :

$$\max \left\{ \prod_{s=1}^S b_{2s,21}^2 \sigma_{2s-1}^2, \prod_{s=1}^S b_{2s-1,21}^2 \sigma_{2s-2}^2 \right\} < 1.$$

C'est la même condition (2.4.4) trouvée par Guégan (1981) dans le cas classique ($S = 1$).

Structure d'autocovariance

Dans ce qui suit, on donne les autocovariances de ce modèle en suivant les deux approches citées précédemment :

1) Approche basée sur l'équation aux différences stochastique (3.4.1)

Du (3.3.11), le calcul des autocovariances de ce modèle se fait uniquement pour $h = 0$; elles sont données par l'équation suivante :

$$\gamma_0^{(t)} = b_{t,21}^2 \gamma_0^{(t-2)} \sigma_{t-1}^2 + \sigma_t^2, t \in \mathbb{Z}$$

ce qui devient après k remplacements successifs :

$$\gamma_0^{(t)} = \prod_{j=1}^k (b_{t-2j+2,21}^2 \sigma_{t-2j+1}^2) \gamma_0^{(t-2k)} + \sum_{j=1}^k \prod_{l=1}^{j-1} (b_{t-2l+2,21}^2 \sigma_{t-2l+1}^2) \sigma_{t-2j+2}^2$$

En utilisant la propriété de la périodicité, i.e $t = s + S\tau$, $s = \overline{1, S}$, $\tau \in \mathbb{Z}$, et lorsque $k = S$,

l'expression de la variance est donnée par

$$\gamma_0^{(s)} = \left(1 - \prod_{j=1}^S b_{s-2j+2,21}^2 \sigma_{s-2j+1}^2 \right)^{-1} \left(\sum_{j=1}^S \prod_{l=1}^{j-1} (b_{s-2l+2,21}^2 \sigma_{s-2l+1}^2) \sigma_{s-2j+2}^2 \right)$$

où on retrouve le même résultat trouvé précédemment, (2.4.7), dans le cas classique lorsque $S = 1$.

Ainsi, pour $S = 4$, on a :

$s = 1$	$\gamma_0^{(1)} = E(X_{1+4\tau} X'_{1+4\tau}),$
$s = 2$	$\gamma_0^{(2)} = E(X_{2+4\tau} X'_{2+4\tau}),$
$s = 3$	$\gamma_0^{(3)} = E(X_{3+4\tau} X'_{3+4\tau}),$
$s = 4$	$\gamma_0^{(4)} = E(X_{4+4\tau} X'_{4+4\tau}),$

qui sont transformées sous les équations de Yule Walker périodiques suivantes :

$$\begin{cases} * \gamma_0^{(1)} = b_{1,21}^2 \gamma_0^{(3)} \sigma_4^2 + \sigma_1^2, \\ * \gamma_0^{(2)} = b_{2,21}^2 \gamma_0^{(4)} \sigma_1^2 + \sigma_2^2, \\ * \gamma_0^{(3)} = b_{3,21}^2 \gamma_0^{(1)} \sigma_2^2 + \sigma_3^2, \\ * \gamma_0^{(4)} = b_{4,21}^2 \gamma_0^{(2)} \sigma_3^2 + \sigma_4^2. \end{cases} \quad (3.4.5)$$

La résolution de ces équations nous donne les autocovariances suivantes :

$$\begin{cases} * \gamma_0^{(1)} = \frac{b_{1,21}^2 \sigma_4^2 \sigma_3^2 + \sigma_1^2}{1 - b_{1,21}^2 b_{3,21}^2 \sigma_4^2 \sigma_2^2}, \\ * \gamma_0^{(2)} = \frac{b_{2,21}^2 \sigma_1^2 \sigma_4^2 + \sigma_2^2}{1 - b_{2,21}^2 b_{4,21}^2 \sigma_1^2 \sigma_3^2}, \\ * \gamma_0^{(3)} = \frac{b_{3,21}^2 \sigma_1^2 \sigma_2^2 + \sigma_3^2}{1 - b_{1,21}^2 b_{3,21}^2 \sigma_4^2 \sigma_2^2}, \\ * \gamma_0^{(4)} = \frac{b_{4,21}^2 \sigma_2^2 \sigma_3^2 + \sigma_4^2}{1 - b_{2,21}^2 b_{4,21}^2 \sigma_1^2 \sigma_3^2}. \end{cases} \quad (3.4.6)$$

2) Approche basée sur la représentation d'état (3.4.3)

Lorsque $P = 2$, on a :

$$\cdot \Gamma_h^{(s)} = \begin{pmatrix} \gamma_{2h}^{(2s)} & \gamma_{2h+1}^{(2s)} \\ \gamma_{2h-1}^{(2s-1)} & \gamma_{2h}^{(2s-1)} \end{pmatrix}, \quad (3.4.7)$$

$$A_t = \begin{pmatrix} b_{2t,21}\varepsilon_{2t-1} & 0 \\ 0 & b_{2t-1,21}\varepsilon_{2t-2} \end{pmatrix} \text{ et } \Delta_t = \begin{pmatrix} \Lambda_0 & 0 \\ 0 & \Lambda_1 \end{pmatrix}.$$

Alors, en multipliant les deux côtés de (3.4.3) par \underline{X}'_{t-h} , $h = 0, \dots, P$, et en prenant l'espérance, on obtient le même système (3.3.12) dont :

$$E \left(A_s \underline{X}_{s+S\tau-1} \underline{X}'_{s+S\tau} \right) = \begin{pmatrix} b_{2s,21}^2 \gamma_0^{(2s-2)} \sigma_{2s-1}^2 & 0 \\ 0 & b_{2s-1,21}^2 \gamma_0^{(2s-3)} \sigma_{2s-2}^2 \end{pmatrix},$$

$$E \left(B_{s+S\tau} \underline{X}'_{s+S\tau} \right) = \begin{pmatrix} \sigma_{2s}^2 & 0 \\ 0 & \sigma_{2s-1}^2 \end{pmatrix}.$$

Alors, la résolution de (3.3.12.a), dans ce cas, nécessite la résolution de l'équation aux différences linéaire du premier ordre suivante :

$$\Gamma_0^{(s)} = \phi_s \Gamma_0^{(s-1)} + \Delta_s, \text{ où}$$

$$\phi_s = \begin{pmatrix} b_{2s,21}^2 \sigma_{2s-1}^2 & 0 \\ 0 & b_{2s-1,21}^2 \sigma_{2s-2}^2 \end{pmatrix} \text{ et } \Delta_s = \begin{pmatrix} \sigma_{2s}^2 & 0 \\ 0 & \sigma_{2s-1}^2 \end{pmatrix}.$$

Et comme cette dernière équation est matricielle, on utilise l'opérateur "vec" et la propriété $\text{vec}(AB) = (I \otimes A) \text{vec}(B)$ pour obtenir l'équation vectorielle suivante :

$$\text{vec} \left(\Gamma_0^{(s)} \right) = (I \otimes \phi_s) \text{vec} \left(\Gamma_0^{(s-1)} \right) + \text{vec}(\Delta_s),$$

ce qui devient, après r ($r \geq 1$) remplacements successifs :

$$\text{vec} \left(\Gamma_0^{(s)} \right) = \prod_{j=1}^r (I \otimes \phi_{s-j+1}) \text{vec} \left(\Gamma_0^{(s-r)} \right) + \text{vec}(\Delta_s) + \sum_{j=1}^{r-1} \prod_{k=1}^j (I \otimes \phi_{s-k+1}) \text{vec}(\Delta_{s-j}). \quad (3.4.8)$$

En posant $r = S$, la propriété de la périodicité de $\Gamma_0^{(s)}$ augmente la performance de (3.4.8) à :

$$\text{vec} \left(\Gamma_0^{(s)} \right) = \left(I_{P^2 \times P^2} - \prod_{j=1}^S (I_{P \times P} \otimes \phi_{s-j+1}) \right)^{-1} \left(\text{vec}(\Delta_s) + \sum_{j=1}^{S-1} \prod_{k=1}^j (I \otimes \phi_{s-k+1}) \text{vec}(\Delta_{s-j}) \right), \quad (3.4.9)$$

et malgré cet avantage, il existe deux inconvénients au niveau de la résolution de l'équation (3.4.9) :

1°/ Le calcul d'une matrice inverse, d'ordre $P^2 \times P^2$, pour chaque saison s , $s = 1, \dots, S$.

2°/ Lorsque $P > S$, la symétrie de $\text{vec} \left(\Gamma_0^{(s)} \right)$ contient des entrées identiques qui conduisent à une redondance indésirable dans le calcul.

En outre, pour pallier ces deux inconvénients, on procède comme suit :

1°/ En posant $r = s$, (3.4.8) se réécrit sous la forme équivalente suivante :

$$vec\left(\Gamma_0^{(s)}\right) = \prod_{j=1}^s (I \otimes \phi_{s-j+1}) vec\left(\Gamma_0^{(S)}\right) + vec(\Delta_s) + \sum_{j=1}^{s-1} \prod_{k=1}^j (I \otimes \phi_{s-k+1}) vec(\Delta_{s-j}). \quad (3.4.8.a)$$

De plus, pour $s = S$, (3.4.8.a) devient :

$$\left(I_{P^2 \times P^2} - \prod_{j=1}^S (I \otimes \phi_{S-j+1}) \right) vec\left(\Gamma_0^{(S)}\right) = vec(\Delta_S) + \sum_{j=1}^{S-1} \prod_{k=1}^j (I \otimes \phi_{S-k+1}) vec(\Delta_{S-j}). \quad (3.4.8.b)$$

2°/ L'amélioration des calculs se fait selon les deux cas suivants :

i) Pour le cas où $S > P$, qui est beaucoup rencontré dans le cas des modèles périodiques, on sait que la redondance en $\Gamma_0^{(S)}$ est causé uniquement par la propriété de la symétrie. Cependant, on peut éviter cet inconvénient en utilisant l'opérateur *vech* au lieu de *vec* dont le lien entre ces deux opérateurs est donné par $vec(A) = D_k vech(A)$, où A est une matrice symétrique d'ordre k et D_k est la matrice de duplication usuelle d'ordre $(k^2 \times \frac{1}{2}k(k+1))$ (eg, Magnus et Neudecker, 1989). Ainsi, l'équation (3.4.8.b) peut se réécrire sous la forme équivalente

$$Bvech\left(\Gamma_0^{(S)}\right) = C, \quad (3.4.10)$$

où la matrice B est le premier bloc de dimension $\frac{P(P+1)}{2}$ de la matrice

$$\left(I_{P^2 \times P^2} - \prod_{j=1}^S (I_{P \times P} \otimes \phi_{S-j+1}) \right) D_{P^2 \times \frac{P(P+1)}{2}}$$

et C est le vecteur contenant les $\frac{P(P+1)}{2}$ premiers éléments du vecteur du côté droit de (3.4.8.b). Par conséquent, nous n'avons besoins que de résoudre un système linéaire d'ordre $\frac{P(P+1)}{2}$.

ii) Pour le cas où $S \leq P$, il est possible d'améliorer la complexité de calcul de (3.4.8.b) en superposant tous les éléments distincts de $\Gamma_0^{(S)}$ et en utilisant ensuite une matrice de permutation appropriée au lieu de la matrice de duplication D_k pour formuler un système similaire à (3.4.10).

Afin de manipuler la résolution de l'équation (3.4.10), on donne les illustrations suivantes, pour $P = 2$ et $S = 4$, qui confirment l'efficacité de cette approche :

▷ La matrice B est le premier bloc de dimension 3 de la matrice $(I_{4 \times 4} - P) D_{4 \times 3}$, où :

$$\cdot (I_{4 \times 4} - P) = \left(I_{4 \times 4} - \prod_{j=1}^4 (I_{2 \times 2} \otimes \phi_{4-j+1}) \right) = \begin{pmatrix} 1 - P_{11} & 0 & 0 & 0 \\ 0 & 1 - P_{22} & 0 & 0 \\ 0 & 0 & 1 - P_{33} & 0 \\ 0 & 0 & 0 & 1 - P_{44} \end{pmatrix},$$

$$\cdot D_{4 \times 3} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \text{ avec}$$

$$P_{11} = b_{8,21}^2 b_{6,21}^2 b_{4,21}^2 b_{2,21}^2 \sigma_7^2 \sigma_5^2 \sigma_3^2 \sigma_1^2,$$

$$P_{22} = b_{7,21}^2 b_{5,21}^2 b_{3,21}^2 b_{1,21}^2 \sigma_6^2 \sigma_4^2 \sigma_2^2 \sigma_0^2,$$

$$P_{33} = b_{8,21}^2 b_{6,21}^2 b_{4,21}^2 b_{2,21}^2 \sigma_7^2 \sigma_5^2 \sigma_3^2 \sigma_1^2,$$

$$P_{44} = b_{7,21}^2 b_{5,21}^2 b_{3,21}^2 b_{1,21}^2 \sigma_6^2 \sigma_4^2 \sigma_2^2 \sigma_0^2.$$

D'où,

$$(I_{4 \times 4} - P) D_{4 \times 3} =$$

$$\begin{pmatrix} 1 - b_{8,21}^2 b_{6,21}^2 b_{4,21}^2 b_{2,21}^2 \sigma_7^2 \sigma_5^2 \sigma_3^2 \sigma_1^2 & 0 & 0 \\ 0 & 1 - b_{7,21}^2 b_{5,21}^2 b_{3,21}^2 b_{1,21}^2 \sigma_6^2 \sigma_4^2 \sigma_2^2 \sigma_0^2 & 0 \\ 0 & 1 - b_{8,21}^2 b_{6,21}^2 b_{4,21}^2 b_{2,21}^2 \sigma_7^2 \sigma_5^2 \sigma_3^2 \sigma_1^2 & 0 \\ 0 & 0 & 1 - b_{7,21}^2 b_{5,21}^2 b_{3,21}^2 b_{1,21}^2 \sigma_6^2 \sigma_4^2 \sigma_2^2 \sigma_0^2 \end{pmatrix},$$

et la matrice B est bien donnée par :

$$B = \begin{pmatrix} 1 - b_{8,21}^2 b_{6,21}^2 b_{4,21}^2 b_{2,21}^2 \sigma_7^2 \sigma_5^2 \sigma_3^2 \sigma_1^2 & 0 & 0 \\ 0 & 1 - b_{7,21}^2 b_{5,21}^2 b_{3,21}^2 b_{1,21}^2 \sigma_6^2 \sigma_4^2 \sigma_2^2 \sigma_0^2 & 0 \\ 0 & 1 - b_{8,21}^2 b_{6,21}^2 b_{4,21}^2 b_{2,21}^2 \sigma_7^2 \sigma_5^2 \sigma_3^2 \sigma_1^2 & 0 \end{pmatrix} = \begin{pmatrix} b_{11} & 0 & 0 \\ 0 & b_{22} & 0 \\ 0 & b_{32} & 0 \end{pmatrix}.$$

▷ C est le vecteur contenant les 03 premiers éléments de :

$$\text{vec}(\Delta_4) + \sum_{j=1}^3 \prod_{k=1}^j (I \otimes \phi_{S-k+1}) \text{vec}(\Delta_{4-j}) = \text{vec}(\Delta_4) + S_1 + S_2 + S_3, \text{ où :}$$

$$\cdot \Delta_1 = \begin{pmatrix} \sigma_2^2 & 0 \\ 0 & \sigma_1^2 \end{pmatrix}, \Delta_2 = \begin{pmatrix} \sigma_4^2 & 0 \\ 0 & \sigma_3^2 \end{pmatrix}, \Delta_3 = \begin{pmatrix} \sigma_6^2 & 0 \\ 0 & \sigma_5^2 \end{pmatrix}, \Delta_4 = \begin{pmatrix} \sigma_8^2 & 0 \\ 0 & \sigma_7^2 \end{pmatrix},$$

$$\cdot S_1 = (I \otimes \phi_4) \text{vec}(\Delta_3) = \begin{pmatrix} b_{8,21}^2 \sigma_7^2 \sigma_6^2 \\ 0 \\ 0 \\ b_{7,21}^2 \sigma_6^2 \sigma_5^2 \end{pmatrix},$$

$$\cdot S_2 = (I \otimes \phi_4) (I \otimes \phi_3) \text{vec}(\Delta_2) = \begin{pmatrix} b_{8,21}^2 b_{6,21}^2 \sigma_7^2 \sigma_5^2 \sigma_4^2 \\ 0 \\ 0 \\ b_{7,21}^2 b_{5,21}^2 \sigma_6^2 \sigma_4^2 \sigma_3^2 \end{pmatrix},$$

$$\cdot S_3 = (I \otimes \phi_4) (I \otimes \phi_3) (I \otimes \phi_2) \text{vec}(\Delta_1) = \begin{pmatrix} b_{8,21}^2 b_{6,21}^2 b_{4,21}^2 \sigma_7^2 \sigma_5^2 \sigma_3^2 \sigma_2^2 \\ 0 \\ 0 \\ b_{7,21}^2 b_{5,21}^2 b_{3,21}^2 \sigma_6^2 \sigma_4^2 \sigma_2^2 \sigma_1^2 \end{pmatrix}.$$

D'où,

$$\begin{aligned} \text{vec}(\Delta_4) + \sum_{j=1}^3 \prod_{k=1}^j (I \otimes \phi_{S-k+1}) \text{vec}(\Delta_{4-j}) = \\ \begin{pmatrix} \sigma_8^2 + b_{8,21}^2 \sigma_7^2 \sigma_6^2 + b_{8,21}^2 b_{6,21}^2 \sigma_7^2 \sigma_5^2 \sigma_4^2 + b_{8,21}^2 b_{6,21}^2 b_{4,21}^2 \sigma_7^2 \sigma_5^2 \sigma_3^2 \sigma_2^2 \\ 0 \\ 0 \\ \sigma_7^2 + b_{7,21}^2 \sigma_6^2 \sigma_5^2 + b_{7,21}^2 b_{5,21}^2 \sigma_6^2 \sigma_4^2 \sigma_3^2 + b_{7,21}^2 b_{5,21}^2 b_{3,21}^2 \sigma_6^2 \sigma_4^2 \sigma_2^2 \sigma_1^2 \end{pmatrix}, \end{aligned}$$

et le vecteur C est bien donné par :

$$C = \begin{pmatrix} \sigma_8^2 + b_{8,21}^2 \sigma_7^2 \sigma_6^2 + b_{8,21}^2 b_{6,21}^2 \sigma_7^2 \sigma_5^2 \sigma_4^2 + b_{8,21}^2 b_{6,21}^2 b_{4,21}^2 \sigma_7^2 \sigma_5^2 \sigma_3^2 \sigma_2^2 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} c_{11} \\ c_{12} \\ c_{13} \end{pmatrix}.$$

Enfin, sous les résultats précédents, l'équation (3.4.10) se transforme comme suit :

$$\begin{pmatrix} b_{11} & 0 & 0 \\ 0 & b_{22} & 0 \\ 0 & b_{32} & 0 \end{pmatrix} \begin{pmatrix} \gamma_0^{(8)} \\ \gamma_1^{(8)} \\ \gamma_0^{(7)} \end{pmatrix} = \begin{pmatrix} c_{11} \\ c_{12} \\ c_{13} \end{pmatrix},$$

et la résolution de ce dernier système donne facilement :

$$\begin{aligned}
 \cdot\gamma_1^{(8)} &= \frac{c_{12}}{b_{22}} = \frac{c_{13}}{b_{32}} = 0, \\
 \cdot\gamma_0^{(8)} &= \frac{c_{11}}{b_{11}} = \frac{\sigma_8^2 + b_{8,21}^2 \sigma_7^2 \sigma_6^2 + b_{8,21}^2 b_{6,21}^2 \sigma_7^2 \sigma_5^2 \sigma_4^2 + b_{8,21}^2 b_{6,21}^2 b_{4,21}^2 \sigma_7^2 \sigma_5^2 \sigma_3^2 \sigma_2^2}{1 - b_{8,21}^2 b_{6,21}^2 b_{4,21}^2 b_{2,21}^2 \sigma_7^2 \sigma_5^2 \sigma_3^2 \sigma_1^2} \\
 &= \frac{\sigma_4^2 + b_{4,21}^2 \sigma_3^2 \sigma_2^2 + b_{4,21}^2 b_{2,21}^2 \sigma_3^2 \sigma_1^2 \sigma_4^2 + b_{4,21}^4 b_{2,21}^2 \sigma_3^4 \sigma_1^2 \sigma_2^2}{1 - b_{4,21}^4 b_{2,21}^4 \sigma_3^4 \sigma_1^4} \\
 &= \frac{(\sigma_4^2 + b_{4,21}^2 \sigma_3^2 \sigma_2^2) + b_{4,21}^2 b_{2,21}^2 \sigma_3^2 \sigma_1^2 (\sigma_4^2 + b_{4,21}^2 \sigma_3^2 \sigma_2^2)}{(1 - b_{4,21}^2 b_{2,21}^2 \sigma_3^2 \sigma_1^2) (1 + b_{4,21}^2 b_{2,21}^2 \sigma_3^2 \sigma_1^2)} \\
 &= \frac{\sigma_4^2 + b_{4,21}^2 \sigma_3^2 \sigma_2^2}{1 - b_{4,21}^2 b_{2,21}^2 \sigma_3^2 \sigma_1^2} = \gamma_0^{(4)}.
 \end{aligned}$$

De même, en suivant les mêmes démarches précédentes sur l'équation (3.4.8.a) on retrouve $\gamma_0^{(1)}, \gamma_0^{(2)}, \gamma_0^{(3)}$, trouvées précédemment en (3.4.6).

Remarque 3.4.1

A noter que les équations de Yule-Wolker périodiques trouvées précédemment en (3.4.5) par la première approche peuvent être obtenues aussi à travers la représentation d'état (3.4.3).

En effet, on a :

▷ D'une part, du (3.3.12.a) :

$$\Gamma_0^{(s)} = \begin{pmatrix} \sigma_{2s}^2 + b_{2s,21}^2 \gamma_0^{(2s-2)} \sigma_{2s-1}^2 & 0 \\ 0 & \sigma_{2s-1}^2 + b_{2s-1,21}^2 \gamma_0^{(2s-3)} \sigma_{2s-2}^2 \end{pmatrix},$$

▷ d'autre part, du (3.4.7) :

$$\Gamma_0^{(s)} = \begin{pmatrix} \gamma_0^{(2s)} & \gamma_1^{(2s)} \\ \gamma_{-1}^{(2s-1)} & \gamma_0^{(2s-1)} \end{pmatrix}.$$

D'où, par identification, on obtient facilement :

$$\begin{cases} \gamma_0^{(2s)} = \sigma_{2s}^2 + b_{2s,21}^2 \gamma_0^{(2s-2)} \sigma_{2s-1}^2, \\ \gamma_0^{(2s-1)} = \sigma_{2s-1}^2 + b_{2s-1,21}^2 \gamma_0^{(2s-3)} \sigma_{2s-2}^2, \end{cases} \quad s = \overline{1, 4}$$

Enfin, d'après ces équations et grâce à la propriété de la périodicité, on est besoin uniquement de poser $s = 1$ et $s = 2$ pour obtenir le même système périodique trouvé en (3.4.5).

Les moments d'ordre supérieurs

En remplaçant P par 2, l'expression (3.3.13) devient :

$$E(X_t)^r = E(\varepsilon_t)^r + \sum_{j=1}^{r/2} C_r^{2j} b_{t,21}^{2j} E(X_{t-2})^{2j} E(\varepsilon_{t-1})^{2j} E(\varepsilon_t)^{r-2j}, \quad (r \text{ pair})$$

et en posant $r/2 = j_0$, cette dernière se réécrit comme suit :

$$\begin{aligned} E(X_t)^r &= E(\varepsilon_t)^r + \sum_{j_1=1}^{j_0} C_{2j_0}^{2j_1} b_{t,21}^{2j_1} E(X_{t-2})^{2j_1} E(\varepsilon_{t-1})^{2j_1} E(\varepsilon_t)^{2j_0-2j_1}, \quad \text{avec } j_0 > j_1 \\ &= E(\varepsilon_t)^r + \sum_{j_1=1}^{j_0-1} C_{2j_0}^{2j_1} b_{t,21}^{2j_1} E(X_{t-2})^{2j_1} E(\varepsilon_{t-1})^{2j_1} E(\varepsilon_t)^{2j_0-2j_1} + b_{t,21}^r E(X_{t-2})^r E(\varepsilon_{t-1})^r. \end{aligned}$$

Alors, par des remplacements successifs sur cette dernière on obtient à l'étape m

$$\begin{aligned} E(X_t)^r &= \prod_{i=1}^m b_{t-2i+2,21}^r E(\varepsilon_{t-2i+1})^r E(X_{t-2m})^r + \sum_{l=0}^{m-1} \left(\prod_{i=1}^l b_{t-2i+2,21}^r E(\varepsilon_{t-2i+1})^r \right) \times \\ &\left\{ E(\varepsilon_{t-2l})^r + \sum_{j_1=1}^{j_0-1} C_{2j_0}^{2j_1} b_{t-2l,21}^{2j_1} E(X_{t-2l-2})^{2j_1} E(\varepsilon_{t-2l})^{2j_0-2j_1} E(\varepsilon_{t-2l-1})^{2j_1} \right\} \end{aligned}$$

En utilisant la propriété de la périodicité, i.e $t = s + S\tau$, $s = \overline{1, S}$, $\tau \in \mathbb{Z}$, et lorsque $m = S$,

l'expression du moment d'ordre r pour ce modèle est donnée par

$$\begin{aligned} E(X_s)^r &= \left(1 - \prod_{i=1}^S b_{s-2i+2,21}^r E(\varepsilon_{s-2i+1})^r \right)^{-1} \left[\sum_{l=0}^{S-1} \left(\prod_{i=1}^l b_{s-2i+2,21}^r E(\varepsilon_{s-2i+1})^r \right) \times \right. \\ &\left. \left\{ E(\varepsilon_{s-2l})^r + \sum_{j_1=1}^{j_0-1} C_{2j_0}^{2j_1} b_{s-2l,21}^{2j_1} E(X_{s-2l-2})^{2j_1} E(\varepsilon_{s-2l})^{2j_0-2j_1} E(\varepsilon_{s-2l-1})^{2j_1} \right\} \right], \quad (3.4.11) \end{aligned}$$

avec $j_0 = r/2 > j_1$ et $\prod_{i=1}^0 (\cdot) = 1$.

Remarque 3.4.2

Grâce à l'expression (3.4.11), le problème du calcul des moments pour le modèle superdiagonale d'ordre un peut être résolu facilement en utilisant la formule suivante :

$$E(X_t)^r = (1 - b_{t,21}^r E(\varepsilon_{t-1})^r)^{-1} \times \left\{ E(\varepsilon_t)^r + \sum_{j_1=1}^{j_0-1} C_{2j_0}^{2j_1} b_{t,2}^{2j_1} E(X_{t-2})^{2j_1} E(\varepsilon_t)^{2j_0-2j_1} E(\varepsilon_{t-1})^{2j_1} \right\},$$

avec $j_0 > j_1$, $\prod_{i=1}^0 (\cdot) = 1$ et $\sum_{i=1}^0 (\cdot) = 0$.

Par exemple, en posant $r = 4$, l'expression du moment d'ordre quatre est donnée par :

$$\begin{aligned}
 E(X_t)^4 &= (1 - b_{21}^4 E(\varepsilon_{t-1})^4)^{-1} \times \{E(\varepsilon_t)^4 + C_4^2 b_{i,2}^2 E(X_{t-2})^2 E(\varepsilon_t)^2 E(\varepsilon_{t-1})^2\}, \\
 &= (1 - b_{21}^4 E(\varepsilon_{t-1})^4)^{-1} \times (E(\varepsilon_t)^4 + 6b_{i,2}^2 E(X_{t-2})^2 \sigma^4).
 \end{aligned}$$

Et lorsque le bruit blanc est gaussien, $N(0, 1) : E(\varepsilon_t)^4 = 3\sigma^4 = 3, V = \frac{1}{1 - b_{21}^2}$ et l'expres-

sion de cette dernière devient :

$$\begin{aligned}
 E(X_t)^4 &= \frac{1}{1 - 3b_{21}^4} \times \left(3 + \frac{6b_{21}^2}{1 - b_{21}^2}\right) \\
 &= \frac{3 + 3b_{21}^2}{(1 - 3b_{21}^4)(1 - b_{21}^2)} = \frac{3(1 + b_{21}^2)}{(1 - 3b_{21}^4)(1 - b_{21}^2)}.
 \end{aligned}$$

Ainsi, grâce à ces résultats, l'expression du Kertosis pour ce modèle est telle que

$$\begin{aligned}
 K &= \frac{E(X_t)^4}{(E(X_t)^2)^2} \\
 &= \frac{3(1 + b_{21}^2)}{(1 - 3b_{21}^4)(1 - b_{21}^2)} \times (1 - b_{21}^2)^2 \\
 &= \frac{3(1 + b_{21}^2)(1 - b_{21}^2)}{(1 - 3b_{21}^4)} = \frac{3(1 - b_{21}^4)}{(1 - 3b_{21}^4)} \\
 &= \frac{3(1 - 3b_{21}^4 + 2b_{21}^4)}{(1 - 3b_{21}^4)} = 3 + \frac{6b_{21}^4}{(1 - 3b_{21}^4)} > 3.
 \end{aligned}$$

Alors, cette dernière indique et confirme le fait que la distribution du modèle bilinéaire est plus épaisse que celle de la loi normale, comme on l'a déjà vu précédemment (*intro.ch.2*)

3.4.2 Le modèle $PBL(1, 0, 1, 1)_S$

Définition 3.4.1

On dit qu'un processus périodiquement corrélé $\{X_t, t \in \mathbb{Z}\}$, $\mathbb{Z} = \{0, \pm 1, \pm 2, \dots\}$, du second ordre, défini sur un espace probabilisé (Ω, F, P) , est diagonal **d'ordre un** périodique, du temps, de période S , noté désormais par $PBL(1, 0, 1, 1)_S$, s'il est solution de l'équation stochastique non linéaire suivante :

$$X_t = a_{t,1}X_{t-1} + b_{t,11}X_{t-1}\varepsilon_{t-1} + \varepsilon_t, \quad t \in \mathbb{Z}, \quad (3.4.12)$$

dont $(\varepsilon_t)_{t \in \mathbb{Z}}$ est une suite de variables aléatoires indépendantes et périodiquement distribuées (i.p.d), de moyenne nulle et de variance σ_t^2 , définie sur le même espace de probabilité (Ω, F, P) ; et où les paramètres $a_{t,1}, b_{t,11}$ ainsi que σ_t^2 sont périodiques de période S , i.e, $(a_{t+kS,1} = a_{t,1}, b_{t+kS,11} = b_{t,11})$ et $\sigma_{t+kS}^2 = \sigma_t^2, \forall t, k \in \mathbb{Z}$.

Écriture Vectorielle

A première vue, on peut écrire le modèle (3.4.12) comme suit :

$$X_t = (a_{t,1} + b_{t,11}\varepsilon_{t-1}) X_{t-1} + \varepsilon_t, \quad t \in \mathbb{Z}$$

Donc, on peut dire que c'est un modèle *RCPAR* mais dans le cas scalaire puisque :

$$\begin{cases} A_t = a_{t,1} + b_{t,11}\varepsilon_{t-1}, \\ et \\ B_t = \varepsilon_t. \end{cases}$$

Existence d'une solution stationnaire stricte

Du (3.4.4), l'exposant de Lyapunov est strictement négatif si

$$\sum_{s=1}^S E \log |a_{s,1} + b_{s,11}\varepsilon_{s-1}| < 0, \quad (3.4.13)$$

qui est la même condition (2.4.8) trouvée par **Liu** (1990) lorsque $S = 1$.

Existence d'une solution stationnaire au second ordre

De même, la condition de stationnarité du second ordre périodique est vérifiée lorsque

$$\prod_{s=1}^S (a_{s,1}^2 + b_{s,11}^2 \sigma_{s-1}^2) < 1, \quad (3.4.14)$$

qui est la même condition trouvée par Bibi et Ho (2006) dans le cas périodique et on retrouve aussi la condition (2.4.9) donnée par Pham et Tran (1981) dans le cas classique.

Structure d'autocovariance

Contrairement aux modèles bilinéaires superdiagonaux, les modèles bilinéaires diagonaux ne sont pas centrés donc le calcul de la moyenne de ce modèle, diagonal d'ordre un, est donné par :

$$E(X_t) = a_{t,1}E(X_{t-1}) + b_{t,11}E(X_{t-1}\varepsilon_{t-1}).$$

En posant $\mu_t = E(X_t)$, on obtient l'équation aux différences linéaire du premier ordre suivante :

$$\mu_t = a_{t,1} \mu_{t-1} + b_{t,11}\sigma_{t-1}^2.$$

Par des remplacements successifs de cette dernière, on obtient :

$$\mu_t = \left(\prod_{i=0}^{m-1} a_{t-i,1} \right) \mu_{t-m} + \sum_{j=1}^m \prod_{i=0}^{j-2} a_{t-i,1} b_{t-j+1,11} \sigma_{t-j}^2. \quad (3.4.15)$$

En utilisant la propriété de la périodicité, i.e, lorsque $t = s+S\tau$, tel que $s \in \{0, \dots, S-1\}$, $\tau \in \mathbb{Z}$, on a :

$$\mu_s = \left(\prod_{i=1}^m a_{s-i,1} \right) \mu_{s-m} + \sum_{j=1}^m \prod_{i=1}^{j-1} a_{s-i,1} b_{s-j+1,11} \sigma_{s-j}^2. \quad (3.4.16)$$

En posant $m = S$, une représentation équivalente à (3.4.16) est donnée par

$$\mu_s = \left(\prod_{s=1}^S a_{s,1} \right) \mu_s + \sum_{j=1}^S \prod_{i=1}^{j-1} a_{s-i,1} b_{s-j+1,11} \sigma_{s-j}^2, \text{ car } \prod_{i=1}^S a_{s-i,1} = \prod_{s=1}^S a_{s,1} \text{ par la propriété de la}$$

périodicité des coefficients $a_{s,1}$.

D'où, la moyenne du modèle $PBL(1, 0, 1, 1)_S$ est donnée par :

$$\mu_s = \left(1 - \prod_{s=1}^S a_{s,1} \right)^{-1} B_s, \prod_{s=1}^S |a_{s,1}| < 1, \quad (3.4.16)$$

$$\text{avec } B_s = \sum_{j=1}^S \prod_{i=1}^{j-1} a_{s-i,1} b_{s-j+1,11} \sigma_{s-j}^2.$$

Maintenant, le calcul de la variance est bien détaillé dans ce qui suit :

$$\bullet \gamma_0^{(t)} = a_{t,1}E(X_t X_{t-1}) + b_{t,11}E(X_t X_{t-1} \varepsilon_{t-1}) + E(X_t \varepsilon_t), \text{ or}$$

$$\left\{ \begin{array}{l} \blacktriangleright E(X_t X_{t-1}) = a_{t,1} \gamma_0^{(t-1)} + b_{t,11}E(X_{t-1}^2 \varepsilon_{t-1}), \\ \blacktriangleright E(X_t X_{t-1} \varepsilon_{t-1}) = a_{t,1}E(X_{t-1}^2 \varepsilon_{t-1}) + b_{t,11}E(X_{t-1}^2 \varepsilon_{t-1}^2), \\ \blacktriangleright E(X_t \varepsilon_t) = \sigma_t^2, \end{array} \right.$$

ce qui donne

$$\bullet \gamma_0^{(t)} = a_{t,1}^2 \gamma_0^{(t-1)} + 2a_{t,1}b_{t,11}E(X_{t-1}^2 \varepsilon_{t-1}) + b_{t,11}^2 E(X_{t-1}^2 \varepsilon_{t-1}^2) + \sigma_t^2, \text{ où}$$

$$\triangleright E(X_{t-1}^2 \varepsilon_{t-1}) = a_{t-1,1}E(X_{t-2} X_{t-1} \varepsilon_{t-1}) + b_{t-1,11}E(X_{t-1} \varepsilon_{t-1} X_{t-2} \varepsilon_{t-2}) + E(X_{t-1} \varepsilon_{t-1}^2),$$

$$\text{avec } \left\{ \begin{array}{l} E(X_{t-2} X_{t-1} \varepsilon_{t-1}) = \sigma_{t-1}^2 \mu_{t-2}, \\ E(X_{t-1} \varepsilon_{t-1} X_{t-2} \varepsilon_{t-2}) = \sigma_{t-1}^2 \sigma_{t-2}^2 \\ E(X_{t-1} \varepsilon_{t-1}^2) = \sigma_{t-1}^2 (a_{t-1,1} \mu_{t-2} + b_{t-1,11} \sigma_{t-2}^2) + E(\varepsilon_{t-1}^3). \end{array} \right.$$

D'où

$$** E(X_{t-1}^2 \varepsilon_{t-1}) = 2\sigma_{t-1}^2 (a_{t-1,1} \mu_{t-2} + b_{t-1,11} \sigma_{t-2}^2) + E(\varepsilon_{t-1}^3),$$

$$\triangleright b_{t,11}^2 E(X_{t-1}^2 \varepsilon_{t-1}^2) = b_{t,11}^2 a_{t-1,1} E(X_{t-2} X_{t-1} \varepsilon_{t-1}^2) + b_{t,11}^2 b_{t-1,11} E(X_{t-2} \varepsilon_{t-2} X_{t-1} \varepsilon_{t-1}^2) + b_{t,11}^2 E(X_{t-1} \varepsilon_{t-1}^3), \text{ dont}$$

$$\left\{ \begin{array}{l} E(X_{t-2} X_{t-1} \varepsilon_{t-1}^2) = a_{t-1,1} \sigma_{t-1}^2 \gamma_0^{(t-2)} + b_{t-1,11} \sigma_{t-1}^2 E(X_{t-2}^2 \varepsilon_{t-2}) + \mu_{t-2} E(\varepsilon_{t-1}^3), \\ \quad = a_{t-1,1} \sigma_{t-1}^2 \gamma_0^{(t-2)} + 2b_{t-1,11} \sigma_{t-1}^2 \sigma_{t-2}^2 (a_{t-2,1} \mu_{t-3} + b_{t-2,11} \sigma_{t-3}^2) + b_{t-1,11} \sigma_{t-1}^2 E(\varepsilon_{t-2}^3) + \mu_{t-2} E(\varepsilon_{t-1}^3). \\ E(X_{t-2} \varepsilon_{t-2} X_{t-1} \varepsilon_{t-1}^2) = a_{t-1,1} \sigma_{t-1}^2 E(X_{t-2}^2 \varepsilon_{t-2}) + b_{t-1,11} \sigma_{t-1}^2 E(X_{t-2}^2 \varepsilon_{t-2}^2) + \sigma_{t-2}^2 E(\varepsilon_{t-1}^3), \\ \quad = 2a_{t-1,1} \sigma_{t-1}^2 \sigma_{t-2}^2 (a_{t-2,1} \mu_{t-3} + b_{t-2,11} \sigma_{t-3}^2) + a_{t-1,1} \sigma_{t-1}^2 E(\varepsilon_{t-2}^3) + b_{t-1,11} \sigma_{t-1}^2 E(X_{t-2}^2 \varepsilon_{t-2}^2) + \sigma_{t-2}^2 E(\varepsilon_{t-1}^3). \\ E(X_{t-1} \varepsilon_{t-1}^3) = a_{t-1,1} \mu_{t-2} E(\varepsilon_{t-1}^3) + b_{t-1,11} \sigma_{t-2}^2 E(\varepsilon_{t-1}^3) + E(\varepsilon_{t-1}^4). \end{array} \right.$$

D'où

$$\begin{aligned}
 & ** b_{t,11}^2 E(X_{t-1}^2 \varepsilon_{t-1}^2) = (b_{t,11}^2 \sigma_{t-1}^2) a_{t-1,1}^2 \gamma_0^{(t-2)} \\
 & \quad + 4 (b_{t,11}^2 \sigma_{t-1}^2) a_{t-1,1} b_{t-1,11} \sigma_{t-2}^2 (a_{t-2,1} \mu_{t-3} + b_{t-2,11} \sigma_{t-3}^2) + b_{t,11}^2 E(\varepsilon_{t-1}^4) + A \\
 & \quad + b_{t,11}^2 (b_{t-1,11}^2 \sigma_{t-1}^2) E(X_{t-2}^2 \varepsilon_{t-2}^2), \\
 & \quad \vdots \\
 & = (b_{t,11}^2 \sigma_{t-1}^2) a_{t-1,1}^2 \gamma_0^{(t-2)} + (b_{t,11}^2 \sigma_{t-1}^2) (b_{t-1,11}^2 \sigma_{t-2}^2) a_{t-2,1}^2 \gamma_0^{(t-3)} + \\
 & (b_{t,11}^2 \sigma_{t-1}^2) (b_{t-1,11}^2 \sigma_{t-2}^2) (b_{t-2,11}^2 \sigma_{t-3}^2) a_{t-3,1}^2 \gamma_0^{(t-4)} + \\
 & + 4 (b_{t,11}^2 \sigma_{t-1}^2) a_{t-1,1} b_{t-1,11} \sigma_{t-2}^2 (a_{t-2,1} \mu_{t-3} + \\
 & b_{t-2,11} \sigma_{t-3}^2) + 4 (b_{t,11}^2 \sigma_{t-1}^2) (b_{t-1,11}^2 \sigma_{t-2}^2) a_{t-2,1} b_{t-2,11} \sigma_{t-3}^2 (a_{t-3,1} \mu_{t-4} + b_{t-3,11} \sigma_{t-4}^2) \\
 & + 4 (b_{t,11}^2 \sigma_{t-1}^2) (b_{t-1,11}^2 \sigma_{t-2}^2) (b_{t-2,11}^2 \sigma_{t-3}^2) a_{t-3,1} b_{t-3,11} \sigma_{t-4}^2 \times (a_{t-4,1} \mu_{t-5} + b_{t-4,11} \sigma_{t-5}^2) \\
 & + b_{t,11}^2 E(\varepsilon_{t-1}^4) + b_{t,11}^2 (b_{t-1,11}^2 \sigma_{t-1}^2) E(\varepsilon_{t-2}^4) + b_{t,11}^2 (b_{t-1,11}^2 \sigma_{t-1}^2) (b_{t-2,11}^2 \sigma_{t-2}^2) E(\varepsilon_{t-3}^4) \\
 & + A + B + C \\
 & + b_{t,11}^2 (b_{t-1,11}^2 \sigma_{t-1}^2) (b_{t-2,11}^2 \sigma_{t-2}^2) (b_{t-3,11}^2 \sigma_{t-3}^2) E(X_{t-4}^2 \varepsilon_{t-4}^2).
 \end{aligned}$$

où,

$$\begin{aligned}
 A & = 2 (b_{t,11}^2) a_{t-1,1} \mu_{t-2} E(\varepsilon_{t-1}^3) + 2b_{t,11}^2 (b_{t-1,11} \sigma_{t-2}^2) E(\varepsilon_{t-1}^3) + \\
 & \quad 2 (b_{t,11}^2 \sigma_{t-1}^2) a_{t-1,1} b_{t-1,11} E(\varepsilon_{t-2}^3), \\
 B & = 2 (b_{t,11}^2 \sigma_{t-1}^2) (b_{t-1,11}^2 \sigma_{t-2}^2) a_{t-2,1} b_{t-2,11} E(\varepsilon_{t-3}^3) + 2b_{t,11}^2 (b_{t-1,11}^2 \sigma_{t-1}^2) a_{t-2,1} \mu_{t-3} E(\varepsilon_{t-2}^3) + \\
 & \quad 2b_{t,11}^2 (b_{t-1,11}^2 \sigma_{t-1}^2) b_{t-2,11} \sigma_{t-3}^2 E(\varepsilon_{t-2}^3), \\
 C & = 2 (b_{t,11}^2 \sigma_{t-1}^2) (b_{t-1,11}^2 \sigma_{t-2}^2) (b_{t-2,11}^2 \sigma_{t-3}^2) a_{t-3,1} b_{t-3,11} E(\varepsilon_{t-4}^3) + \\
 & \quad 2b_{t,11}^2 (b_{t-1,11}^2 \sigma_{t-1}^2) (b_{t-2,11}^2 \sigma_{t-2}^2) a_{t-3,1} \mu_{t-4} E(\varepsilon_{t-3}^3) + \\
 & \quad 2b_{t,11}^2 (b_{t-1,11}^2 \sigma_{t-1}^2) (b_{t-2,11}^2 \sigma_{t-2}^2) b_{t-3,11} \sigma_{t-4}^2 E(\varepsilon_{t-3}^3).
 \end{aligned}$$

sont égales à zéro car le bruit est gaussien.

En posant $D_{t-1} = E(X_{t-1}^2 \varepsilon_{t-1}^2)$, l'expression de cette espérance à la $(m+1)^{\text{ème}}$ étape est donnée par

$$\begin{aligned}
 b_{t,11}^2 D_{t-1} & = b_{t,11}^2 \prod_{j=1}^m (b_{t-j,11}^2 \sigma_{t-j}^2) D_{t-m-1} + \sum_{j=1}^m \left[\prod_{l=1}^j (b_{t-l+1,11}^2 \sigma_{t-l}^2) \left\{ a_{t-j,1}^2 \gamma_0^{(t-j-1)} + \right. \right. \\
 & \quad \left. \left. 4 (a_{t-j,1} b_{t-j,11} \sigma_{t-j-1}^2) (a_{t-j-1,1} \mu_{t-j-2} + b_{t-j-1,11} \sigma_{t-j-2}^2) \right\} + b_{t,11}^2 \prod_{l=1}^{j-1} (b_{t-l,11}^2 \sigma_{t-l}^2) E(\varepsilon_{t-j}^4) \right],
 \end{aligned}$$

En utilisant la propriété de la périodicité, i.e, lorsque $t = s + S\tau$, tel que $s \in \{0, \dots, S-1\}$, $\tau \in \mathbb{Z}$, on a :

$$b_{s,11}^2 D_{s-1} = b_{s,11}^2 \prod_{j=1}^m (b_{s-j,11}^2 \sigma_{s-j}^2) D_{s-m-1} + \sum_{j=1}^m \left[\prod_{l=1}^j (b_{s-l+1,11}^2 \sigma_{s-l}^2) \left\{ a_{s-j,1}^2 \gamma_0^{(s-j-1)} + 4 (a_{s-j,1} b_{s-j,11} \sigma_{s-j-1}^2) (a_{s-j-1,1} \mu_{s-j-2} + b_{s-j-1,11} \sigma_{s-j-2}^2) \right\} + b_{s,11}^2 \prod_{l=1}^{j-1} (b_{s-l,11}^2 \sigma_{s-l}^2) E(\varepsilon_{s-j}^4) \right].$$

En posant $m = S$, une représentation équivalente à cette dernière est donnée par

$$b_{s,11}^2 D_{s-1} = \left(1 - \prod_{j=1}^S (b_{s-j,11}^2 \sigma_{s-j}^2) \right)^{-1} \left\{ \sum_{j=1}^S \left[\prod_{l=1}^j (b_{s-l+1,11}^2 \sigma_{s-l}^2) \left\{ a_{s-j,1}^2 \gamma_0^{(s-j-1)} + 4 (a_{s-j,1} b_{s-j,11} \sigma_{s-j-1}^2) (a_{s-j-1,1} \mu_{s-j-2} + b_{s-j-1,11} \sigma_{s-j-2}^2) \right\} + b_{s,11}^2 \prod_{l=1}^{j-1} (b_{s-l,11}^2 \sigma_{s-l}^2) E(\varepsilon_{s-j}^4) \right] \right\}.$$

Par conséquent, la variance de ce modèle devient

$$\begin{aligned} \gamma_0^{(s)} &= a_{s,1}^2 \gamma_0^{(s-1)} + 4a_{s,1} b_{s,11} \sigma_{s-1}^2 (a_{s-1,1} \mu_{s-2} + b_{s-1,11} \sigma_{s-2}^2) + \sigma_s^2 \\ &+ \left(1 - \prod_{j=1}^S (b_{s-j,11}^2 \sigma_{s-j}^2) \right)^{-1} \left\{ \sum_{j=1}^S \left[\prod_{l=1}^j (b_{s-l+1,11}^2 \sigma_{s-l}^2) \left\{ a_{s-j,1}^2 \gamma_0^{(s-j-1)} + 4 (a_{s-j,1} b_{s-j,11} \sigma_{s-j-1}^2) (a_{s-j-1,1} \mu_{s-j-2} + b_{s-j-1,11} \sigma_{s-j-2}^2) \right\} + b_{s,11}^2 \prod_{l=1}^{j-1} (b_{s-l,11}^2 \sigma_{s-l}^2) E(\varepsilon_{s-j}^4) \right] \right\}. \end{aligned}$$

Remarquons que lorsque $S = 1$, l'expression de cette variance devient

$$\begin{aligned} \gamma_0^{(t)} &= a_1^2 \gamma_0^{(t)} + 4a_1 b_{11} \sigma^2 (a_1 \mu + b_{11} \sigma^2) + \sigma^2 \\ &+ (1 - b_{11}^2 \sigma^2)^{-1} \left\{ \left[(b_{11}^2 \sigma^2) \left\{ a_1^2 \gamma_0^{(t)} + 4 (a_1 b_{11} \sigma^2) (a_1 \mu + b_{11} \sigma^2) \right\} + b_{11}^2 E(\varepsilon_t^4) \right] \right\} \\ &= (1 - a_1^2 - b_{11}^2 \sigma^2)^{-1} [4a_1 b_{11} \sigma^2 (a_1 \mu + b_{11} \sigma^2) + \sigma^2 (1 - b_{11}^2 \sigma^2) + b_{11}^2 E(\varepsilon_t^4)] \end{aligned}$$

qui est, lorsque $E(\varepsilon_t^4) = 3\sigma^4$, le même résultat trouvé par Pham et Tran, 1981.

Par ailleurs, le calcul des autocovariances est bien expliqué dans ce qui suit :

$$\begin{aligned} \bullet \gamma_1^{(t)} &= a_{t,1} E(X_{t-1}^2) + b_{t,11} E(X_{t-1}^2 \varepsilon_{t-1}), \\ &= a_{t,1} \gamma_0^{(t-1)} + 2b_{t,11} \sigma_{t-1}^2 (a_{t-1,1} \mu_{t-2} + b_{t-1,11} \sigma_{t-2}^2), \\ \bullet \gamma_2^{(t)} &= a_{t,1} E(X_{t-1} X_{t-2}) + b_{t,11} E(X_{t-1} X_{t-2} \varepsilon_{t-1}), \\ &= a_t \gamma_1^{(t-1)} + b_{t,11} \sigma_{t-1}^2 \mu_{t-2}, \\ \bullet \gamma_3^{(t)} &= a_{t,1} E(X_{t-1} X_{t-3}) + b_{t,11} E(X_{t-1} X_{t-3} \varepsilon_{t-1}), \\ &= a_{t,1} \gamma_2^{(t-1)} + b_{t,11} \sigma_{t-1}^2 \mu_{t-3}, \\ &\vdots \\ \bullet \gamma_j^{(t)} &= a_{t,1} \gamma_{j-1}^{(t-1)} + b_{t,11} \sigma_{t-1}^2 \mu_{t-j}, j \geq 2. \end{aligned}$$

Enfin, les autocovariances peuvent être obtenues en résolvant le système suivant :

$$\begin{cases} \gamma_0^{(s)} = a_{s,1}^2 \gamma_0^{(s-1)} + 4a_{s,1} b_{s,11} \sigma_{s-1}^2 (a_{s-1,1} \mu_{s-2} + b_{s-1,11} \sigma_{s-2}^2) + \sigma_s^2 + \Delta_s, \\ \gamma_1^{(s)} = a_{s,1} \gamma_0^{(s-1)} + \Lambda_s, \\ \gamma_j^{(s)} = a_{s,1} \gamma_{j-1}^{(s-1)} + b_{s,11} \sigma_{s-1}^2 \mu_{s-j}, j \geq 2. \end{cases}, \text{ avec}$$

$$\Delta_s = b_{s,11}^2 D_{s-1}, \text{ et}$$

$$\Lambda_s = 2b_{s,11} \sigma_{s-1}^2 (a_{s-1,1} \mu_{s-2} + b_{s-1,11} \sigma_{s-2}^2).$$

Les moments d'ordre supérieurs

De même que précédemment, en remplaçant p par 1 dans (3.3.14), on obtient

$$E(X_t^r) = E(\varepsilon_t)^r + \sum_{j=1}^r C_r^j E \left[(X_{t-1} b_{t,11} \varepsilon_{t-1})^j + \sum_{l=1}^j C_j^l (X_{t-1} a_{t,1})^l \times (X_{t-1} b_{t,11} \varepsilon_{t-1})^{j-l} \right] E(\varepsilon_t)^{r-j},$$

et en posant $r = j_0$, l'expression de cette dernière devient

$$E(X_t^r) = E(\varepsilon_t)^r + \sum_{j_1=1}^{j_0-1} C_{j_0}^{j_1} \left[b_{t,11}^{j_1} E(X_{t-1} \varepsilon_{t-1})^{j_1} + \sum_{j_2=1}^{j_1} C_{j_1}^{j_2} a_{t,1}^{j_2} b_{t,11}^{j_1-j_2} \times E \left((X_{t-1})^{j_1} (\varepsilon_{t-1})^{j_1-j_2} \right) \right] E(\varepsilon_t)^{j_0-j_1}.$$

Avec m remplacements successifs, on obtient

$$\begin{aligned} E(X_t^r) &= E(\varepsilon_t)^r + \prod_{i=1}^m a_{t-i+1}^r E(X_{t-m}^r) + \sum_{l=1}^{m-1} \left(\prod_{i=1}^l a_{t-i,1}^r \right) E(\varepsilon_{t-l}^r) \\ &+ \sum_{l=0}^{m-1} \left(\prod_{i=1}^l a_{t-i+1,1}^r \right) \left\{ \sum_{j_2=0}^{r-1} C_r^{j_2} a_{t-l,1}^{j_2} b_{t-l,11}^{r-j_2} \times E \left((X_{t-l-1})^r (\varepsilon_{t-l-1})^{r-j_2} \right) \right. \\ &\left. + \sum_{j_1=1}^{j_0-1} C_{j_0}^{j_1} \left[\sum_{j_2=0}^{j_1} C_{j_1}^{j_2} a_{t-l,1}^{j_2} b_{t-l,11}^{j_1-j_2} \times E \left((X_{t-l-1})^{j_1} (\varepsilon_{t-l-1})^{j_1-j_2} \right) \right] E(\varepsilon_{t-l})^{j_0-j_1} \right\}. \end{aligned}$$

En utilisant la propriété de la périodicité, i.e $t = s + S\tau$, $s = \overline{1, S}$, $\tau \in \mathbb{Z}$, et lorsque $m = S$,

l'expression du moment d'ordre r pour ce modèle est donnée par

$$\begin{aligned} E(X_s^r) &= \left(1 - \prod_{i=1}^S a_{s-i+1}^r \right)^{-1} \left[\sum_{l=0}^{S-1} \left(\prod_{i=1}^l a_{s-i+1,1}^r \right) \left\{ E(\varepsilon_{s-l}^r) + \sum_{j_2=0}^{r-1} C_r^{j_2} a_{s-l,1}^{j_2} b_{s-l,11}^{r-j_2} \right. \right. \\ &\times E \left((X_{s-l-1})^r (\varepsilon_{s-l-1})^{r-j_2} \right) + \sum_{j_1=1}^{j_0-1} C_{j_0}^{j_1} \left(\sum_{j_2=0}^{j_1} C_{j_1}^{j_2} a_{s-l,1}^{j_2} b_{s-l,11}^{j_1-j_2} \times E \left((X_{s-l-1})^{j_1} (\varepsilon_{s-l-1})^{j_1-j_2} \right) \right) \\ &\left. \left. \times E(\varepsilon_{s-l})^{j_0-j_1} \right\} \right], \text{ avec } j_0 = r > j_1 > j_2 \text{ et } \prod_{i=1}^0 (\cdot) = 1. \end{aligned}$$

Chapitre 4

Conclusion et perspectives

Dans le cadre de ce travail, on a essayé de toucher le bord de quelques propriétés probabilistes des modèles bilinéaires non stationnaires à coefficients périodiques, dans le temps, ou encore des modèles bilinéaires périodiques et ce à travers le fond des propriétés probabilistes des modèles bilinéaires stationnaires. Et comme l'objectif de ces modèles n'a pas atteint son apogée jusqu'à ce jour, autrement dit que leur fond n'est pas encore visible, alors il est évident de dire que la propriété de la périodicité reste un point ouvert nécessitant une attention particulière dans les recherches à venir.

Plus précisément, on a donné une condition suffisante d'existence d'une solution strictement stationnaire et au second ordre pour l'ensemble des modèles bilinéaires périodiques ayant une représentation markovienne linéaire identique à une équation aux différences stochastique de la forme $X_n = A_n X_{n-1} + B_n, \forall n \in \mathbb{Z}$, où (A_n, B_n) est une suite de matrices aléatoires périodiques non nécessairement indépendantes. Ensuite, pour la classe des modèles bilinéaires $PBL(0, 0, P, 1)$, on a constaté que ces modèles représentent un bruit blanc faible et ce à travers leur structure d'autocorrélation.

Vu qu'à notre connaissance il existe deux méthodes de résolution de la structure d'autocovariance "approche basée sur l'équation aux différences" et "approche basée sur la représentation espace d'états" des modèles bilinéaires périodiques dont la première est facile à manipuler quant à la deuxième, malgré son efficacité par rapport à la première notamment lorsqu'il s'agit d'une période importante, exige la présence de la matrice A_n , nécessitant à son tour le calcul d'une matrice inverse d'ordre $p \times p$. A cet effet, la résolution du système n'a été appliquée que pour $p = 2$ puisque la matrice A_n n'est pas disponible dans le cas où

l'ordre p est quelconque. D'ailleurs, il est très important de noter que le manque de cette matrice est une clé qui contribue à une nouvelle orientation dans les recherches à venir, i.e, elle offre une forme plus explicite de l'exposant de Lyapunov γ^S ce qui conduit à son tour à donner une condition d'existence d'une solution strictement stationnaire périodique plus simple à celle donnée dans ce cadre de travail.

Outre le modèle purement bilinéaire et strictement superdiagonal périodique $PBL(0, 0, P, 1)$, on a étudié aussi le modèle bilinéaire superdiagonal périodique $PBL(p, 0, p, 1)$ où on a rencontré, par rapport au premier, une difficulté dans le calcul des autocovariances à cause de la diagonalité du terme bilinéaire qui le caractérise. Ensuite, on a illustré les résultats trouvés avant à travers deux cas particuliers, en posant $P=2$ dans le premier modèle et $P=1$ dans le deuxième. C'est-à-dire que, on a abordé deux cas particuliers de deux classes de séries temporelles non linéaires périodiques telles que la classe des modèles purement bilinéaire et strictement superdiagonaux périodiques et la classe des modèles diagonaux périodiques or il est intéressant aussi d'explorer les modèles sous diagonaux périodiques et de les comparer avec les deux classes citées auparavant.

Ainsi, on a considéré dans ce travail uniquement l'aspect temporel alors qu'il y a un intérêt croissant de ces modèles à travers l'approche fréquentielle (dans le cas périodique).

Chapitre 5

Annexe

Dans ce qui suit, on présente un complémentaire au chapitre deux sous format de preuves de quelques théorèmes afin d'approfondir la compréhension et d'élargir la connaissance :

Preuve du théorème (2.2.2)

Il est facile de voir que (ii) implique (i) à partir du moment où la limite Q de Q_k , si elle existe, satisfait l'équation de (i). Inversement, supposons que cette équation admette une solution positive Q . Alors, Q_k étant définie comme en (ii), on a :

$Q_0 \leq Q$ et donc $Q_1 = AQ_0A' + BQ_0B' \sigma^2 + \Sigma \leq AQ_0A' + BQ_0B' \sigma^2 + \Sigma = Q$ et en répétant cet argument, on a $Q_k \leq Q$ pour tout k . Le même argument, en commençant avec $Q_0 \leq Q_1$ donne $Q_k \leq Q_{k+1}$. Ainsi la suite Q_k converge. Or, Q_k est la matrice de covariance de :

$$Z_{(t,k)} = \sum_{i=0}^k (A + B\varepsilon_t) \dots (A + B\varepsilon_{t-i+1}) (C\varepsilon_{t-k} + D(\varepsilon_{t-k}^2 - \sigma^2))$$
 et comme $C\varepsilon_s + D(\varepsilon_s^2 - \sigma^2)$

est de moyenne zéro et est indépendant de $\varepsilon_{s+1}, \varepsilon_{s+2}, \dots$, alors pour tout vecteur x , la suite $xZ_{(t,k)}$, $k = 1, 2, \dots$ est une martingale dont le moment d'ordre deux est borné par xQx et donc, converge en moyenne quadratique et presque sûrement vers la limite $xZ_t x$. Il est facile de vérifier que Z_t est la solution de (2.2.1). Inversement, si (2.2.1) admet une solution stationnaire, alors (i) est vérifiée à partir du moment où la matrice de covariance Q de Z_t satisfait la relation $Q = AQ_0A' + BQ_0B' \sigma^2 + \Sigma$. Pour montrer l'unicité de la solution (2.2.3), on peut observer que toute solution de (2.2.1) a un vecteur d'état qui diffère de Z_t d'une variable aléatoire U_t satisfaisant $U_t = (A + B\varepsilon_t)U_{t-1}$ et donc la matrice de covariance Q , solution de l'équation $Q = AQ_0A' + BQ_0B' \sigma^2$, doit tendre vers zéro par hypothèse. ■

Preuve du théorème (2.2.8)

Tout d'abord, on montre l'équivalence de (c_1) et (c_2) et de (O_1) et (O_2) . Pour cela, on observe que Q_k est la somme de toutes les matrices de la forme : $A^{n_1}(\sigma B)^{m_1} \dots A^{n_r}(\sigma B)^{m_r} \Sigma (\sigma B')^{m_r} A'^{m_r} \dots (\sigma B')^{m_1} A'^{m_1}$ avec $n_1 \geq 1, n_i \geq 0, 1 < i \leq r, m_r \geq 1$ et $\sum (n_i + m_i) \leq k$.

Puisque, $x'Tx = 0$ implique que $x'T = 0$ si T est une matrice positive, la singularité de Q_k est équivalente à l'existence d'un vecteur orthogonal à toutes les colonnes de $A^{n_1}(\sigma B)^{m_1} \dots A^{n_r}(\sigma B)^{m_r}$ où n_i et m_i sont définies comme ci-dessus. Ce qui montre l'équivalence de (c_1) et (c_2) . De même pour (O_1) et (O_2) .

On montre maintenant que si la représentation (2.2.1) n'est pas quasi minimale, alors (c_1) ou (O_2) ne sont pas vérifiées. Si (c_1) n'est pas vérifiée, il n'ya rien à démontrer. On suppose que (c_1) est vérifiée ; ce qui implique que la matrice de covariance W de Z_t est non singulière. Par hypothèse, il existe une deuxième représentation :

$$\begin{cases} X_t = \tilde{H}\tilde{Z}_{t-1} + \varepsilon_t, \\ \tilde{Z}_t = \tilde{A}\tilde{Z}_{t-1} + \tilde{B}\tilde{Z}_{t-1}\varepsilon_t + \tilde{C}\varepsilon_t + \tilde{D}\varepsilon_t^2 - \sigma^2, \end{cases}$$

avec $\begin{cases} \tilde{Z}_t = SZ_t = \tilde{A}SZ_{t-1} + \tilde{B}SZ_{t-1}\varepsilon_t + \tilde{C}\varepsilon_t + \tilde{D}\varepsilon_t^2 - \sigma^2, \\ X_t = \tilde{H}SZ_{t-1} + \varepsilon_t. \end{cases}$

Par comparaison avec (2.2.1), on obtient $\tilde{H}S = H, \tilde{A}S = SA, \tilde{B}S = SB, \tilde{C} = SC, \tilde{D} = SD$. Comme S a moins de lignes que de colonnes, il existe un vecteur $x \neq 0$ tel que $SX = 0$. Par suite $Hx = \tilde{H}Sx = 0, HAx = \tilde{H}\tilde{A}Sx = 0, \dots$ D'où (O_2) n'est pas vérifiée.

Inversement, si (c_1) ou (O_2) ne sont pas vérifiées, alors la représentation n'est pas quasi minimale. Si (c_1) n'est pas vérifiée, alors en notant $Ker(Q_k), k = 1, \dots, n$, le noyau de Q_k , on a : $Ker(Q_k) \neq 0$ et comme $Ker(Q_k) \supset Ker(Q_{k+1})$, on a : $Ker(Q_{p+1}) = Ker(Q_p)$, pour $p < n$. Donc, $\forall x \in Ker(Q_{p+1}) : x'Q_{p+1}x = 0$ c'est à dire $x'AQ_pA'x + x'BQ_pB'x + x'\Sigma x$ ainsi que $x \in Ker(\Sigma)$ et Ax et Bx appartient à $Ker(Q_p) = Ker(Q_{p+1})$. En répétant cet argument, $Ker(Q_k) = Ker(Q_p)$ pour $k > p$. Par suite Q est singulière. Ainsi on peut construire une représentation équivalente en conservant uniquement les composantes linéairement indépendantes de Z_t , et donc la représentation (2.2.1) n'est pas quasi minimale. De l'autre côté, si (O_2) n'est pas vérifiée, ensuite de la même manière que précédemment, P_k n'est pas singulière pour tout k et donc les vecteurs lignes $HA^{n_1}B^{m_1} \dots A^{n_r}B^{m_r}$ peuvent s'exprimer comme des combinaisons linéaires de $p < n$ parmi eux. Soit S la matrice formée par ces

vecteurs, alors $H = \tilde{H}S, SA = \tilde{A}S, SB = \tilde{B}S$ et d'où on a une deuxième représentation avec le vecteur d'état SZ_t . ■

Preuve du théorème (2.3.2)

*Pour établir ce résultat, nous commençons par donner la démonstration pour $Q = 2$ et dans ce cas (2.2.1) devient :

$$\mathbf{X}_t = C\varepsilon_t + (A + B_1\varepsilon_{t-1} + B_2\varepsilon_{t-2}) \mathbf{X}_{t-1}. \quad (2.3.8)$$

Ainsi, avant de procéder à cette preuve, on introduit d'abord une suite de variables de dimension p qui sera utilisée par la suite pour générer une solution strictement stationnaire de (2.3.8), Soit :

$$S_n(t) = \begin{cases} 0, & \text{si } n < 0, \\ C\tilde{\varepsilon}_t, & \text{si } n = 0, \\ C\tilde{\varepsilon}_t + (A + B_1\varepsilon_{t-1} + B_2\varepsilon_{t-2}) S_{n-1}(t-1), & \text{si } n > 0, \end{cases} \quad (2.3.9)$$

et $\Delta_n(t) = S_n(t) - S_{n-1}(t)$ qui vérifié l'équation $\Delta_n(t) = (A + B_1\varepsilon_{t-1} + B_2\varepsilon_{t-2}) \Delta_{n-1}(t-1)$, pour $n \geq 3$.

De plus, par un simple argument en utilisant un critère d'induction, on peut conclure que ces deux suites sont strictement stationnaires. Alors, en utilisant la théorie \mathbb{L}^p , ($p > 1$), le problème de l'existence d'une solution stationnaire au second ordre se réduit maintenant à la convergence de $(S_n(t))_{n \geq 0}$ vers \mathbf{X}_t en \mathbb{L}^2 dont la quantité d'intérêt pour atteindre la convergence est $E \{ \Delta_n(t) \Delta'_n(t) \}$ qui est une quantité matricielle, et comme il est plus approprié d'utiliser les vecteurs des moments on a alors :

$$\begin{aligned} \tilde{V}_n &= E \{ \Delta_n^{\otimes 2}(t) \}, \\ &= (A \otimes A + \sigma^2 B_1 \otimes B_1) \tilde{V}_{n-1} + (A \otimes B_2 + B_2 \otimes A) E \{ \Delta_{n-1}^{\otimes 2}(t) \varepsilon_{t-1} \} + (B_2 \otimes B_2) \cdot \\ &\quad E \{ \Delta_{n-1}^{\otimes 2}(t) \varepsilon_{t-1}^2 \}. \end{aligned}$$

En posant $\tilde{D}_n = E \{ \Delta_n^{\otimes 2}(t) \varepsilon_{t-1} \}$, $\tilde{F}_n = E \{ \Delta_n^{\otimes 2}(t) \varepsilon_{t-1}^2 \}$, on obtient :

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

Ce système se réduit facilement à une seule équation aux différences, en posant $V_n = (\tilde{V}_n, \tilde{D}_n, \tilde{F}_n)$, $n \geq 3$ comme suit :

$$V_n = \Gamma V_{n-1} = \dots = \Gamma^{n-3} V_{n-3}, \quad (2.3.10)$$

où Γ est une matrice définie comme (2.3.5). Donc, la condition nécessaire et suffisante d'existence d'une solution stable de cette équation est que $\rho(\Gamma) < 1$.

Un autre instrument qui précède cette preuve et qui semble intéressant à la compréhension de celle ci se trouve dans la proposition suivante :

Proposition 2.3.3 :

Soit $|V_n|$ la somme en valeur absolue de toutes les composantes de V_n . Alors, si $\lambda := \rho(\Gamma) < 1$, on a l'inégalité $|V_n| \leq c\lambda^{n/2}$, pour tout $n \in \{3, 4, \dots\}$ et pour une certaine constante c .

Preuve :

Si $\Gamma = T^{-1}\Lambda T$ est la décomposition du Jordan de Γ , alors il est clair que pour $n > 2d = 6p^2$ (d est la dimension de V_n) les éléments de Λ^n sont bornés en valeur absolue par $\binom{n}{d} \lambda^{n-d}$,

et d'où par $K\lambda^{n/2}$ pour une certaine constante K . Par conséquent, on obtient :

$$|V_n| = |\Gamma^{n-2}V_2| = |T^{-1}\Lambda^{n-2}TV_2| \leq c\lambda^{n/2}, \text{ pour une certaine constante } c. \square$$

Enfin, on peut entrer dans les trois fameuses étapes de démonstration du théorème (2.3.2) dont il s'agit de montrer que :

· (i) l'équation (2.3.1) a une solution strictement stationnaire de carré intégrable \mathbf{X}_t :

A partir de la proposition (2.3.1), il s'ensuit que :

$$E \|S_n(t) - S_{n-1}(t)\|^2 = E \|\Delta_n(t)\|^2 = \text{trace}(V_n) \leq |u_n| \leq c\lambda^{n/2},$$

et donc pour chaque t fixé, $S_n(t)$ est une suite de Cauchy en \mathbb{L}^2 et ainsi elle converge à la fois en \mathbb{L}^2 et presque sûrement. D'où ses limites $\mathbf{X}_t, t = 0, \pm 1, \dots$ sont des solutions strictement stationnaires de carrés intégrables de (2.3.1), et donc la première composante $X_t, t = 0, \pm 1, \dots$ satisfait (2.1.1).

· (ii) la première composante X_t est une solution stationnaire, causale et ergodique de (2.1.1) :

► La stationnarité de X_t se découle à partir de la stationnarité stricte de \mathbf{X}_t et du fait que pour chaque t , \mathbf{X}_t est la limite en \mathbb{L}^2 de $S_n(t)$.

► Etablir la causalité et l'ergodicité de X_t revient à montrer qu'il existe une fonction $f : \mathbb{R}^\infty \rightarrow \mathbb{R}$, tel que : $X_t = f(\varepsilon_t, \varepsilon_{t-1}, \dots)$, p.s pour tout $t \in \{0, \pm 1, \dots\}$, On a :

d'une part, $\begin{cases} S_n(t) \xrightarrow{mq} \mathbf{X}_t = f_t(\varepsilon_t, \varepsilon_{t-1}, \dots), \\ S_n(0) \xrightarrow{mq} \mathbf{X}_0 = f_0(\varepsilon_0, \varepsilon_{-1}, \dots), \text{ pour } t = 0, \end{cases}$

et d'autre part, puisque $S_n(t)$ est définie comme (2.3.9), alors il existe une fonction mesurable $G_n : \mathbb{R}^\infty \rightarrow \mathbb{R}^p$ telle que pour **tout** $t \in \{0, \pm 1, \dots\}$,

$$\begin{aligned} S_n(t) &= G_n(\varepsilon_t, \varepsilon_{t-1}, \dots), n = 0, 1, 2, \dots \\ &= G_n(\varepsilon_0, \varepsilon_{-1}, \dots). \end{aligned}$$

Alors :

$$\begin{cases} S_n(t) \xrightarrow{mq} f_0(\varepsilon_t, \varepsilon_{t-1}, \dots), \\ f_t = f_0 = (f_0^{(1)}, f_0^{(2)}, \dots). \end{cases}$$

En posant $f = f_0^{(1)}$, on peut écrire : $X_t = f(\varepsilon_t, \varepsilon_{t-1}, \dots)$, p.s pour tout $t \in \{0, \pm 1, \dots\}$ ce qui signifie que X_t est causal et ergodique.

·(iii) Les solutions \mathbf{X}_t et X_t sont uniques :

► Soit \mathbf{Y}_t une solution quelconque de (2.3.1) telle que $\mathbf{Y}_t = g(\varepsilon_t, \varepsilon_{t-1}, \dots)$, (2.3.11)

où g est une fonction mesurable de \mathbb{R}^∞ dans \mathbb{R}^p et soit : $\mathbf{U}_t = \mathbf{X}_t - \mathbf{Y}_t$ où \mathbf{X}_t est la solution de (2.3.1) trouvée en (i). Alors \mathbf{U}_t est strictement stationnaire et satisfait les équations suivantes : $U_t = (A + B_1\varepsilon_{t-1} + B_2\varepsilon_{t-2})U_{t-1}, t \in 0, \pm 1, \dots$

En utilisant les mêmes arguments que précédemment, (2.3.10) devient :

$u = \Gamma u = \dots = \Gamma^t u, \forall t \geq 0$, et comme $\lambda = \rho(\Gamma) < 1$, on en déduit que $u = 0$ (en probabilité 1), et donc que \mathbf{X}_t est la solution unique de (2.3.1) de la forme (2.3.11).

► Pour montrer maintenant que X_t est l'unique solution stationnaire de (2.3.1), on définit :

$$S_n^*(t) = \begin{cases} 0, & \text{si } n < 0, \\ \tilde{\varepsilon}_t^* = (\varepsilon_{t-2}, \dots, \varepsilon_{t-p+1})', & \text{si } n = 0, \\ C\tilde{\varepsilon}_t + (A + B_1\varepsilon_{t-1} + B_2\varepsilon_{t-2})S_{n-1}^*(t-1), & \text{si } n > 0, \end{cases}$$

et on procède de la même façon qu'en (i) pour montrer que pour chaque $t, (S_n^*(t), n \geq 0)$ converge à la fois presque sûrement et dans \mathbb{L}^2 vers un vecteur aléatoire \mathbf{Y}_t de la forme (2.3.11) satisfaisant (2.3.1). A partir de l'unicité établie pour le processus \mathbf{X}_t , on en déduit que $\mathbf{Y}_t = \mathbf{X}_t$ pour tout $t \in (0, \pm 1, \dots)$ en proba 1.

► Il est ensuite facile de voir, par simple itération à partir de (2.3.9) que :

$$S_n(t) = C \tilde{\varepsilon}_t + \sum_{m=1}^n \prod_{j=1}^m (A + B_1\varepsilon_{t-j} + B_2\varepsilon_{t-j-1}) C \tilde{\varepsilon}_{t-m},$$

alors $\mathbf{X}_t = \lim_{n \rightarrow \infty} S_n(t)$ satisfait (2.3.7), et où la somme infinie peut être interprétée comme une limite presque sûre ou en \mathbb{L}^2 .

**Enfin, pour démontrer le théorème dans le cas général, avec Q quelconque, on utilise les mêmes arguments que précédemment, mais les suites $S_n(t)$ définies en (2.3.9) sont remplacées par :

$$S_n(t) = \begin{cases} 0, & \text{si } n < 0, \\ C\tilde{\varepsilon}_t, & \text{si } n = 0, \\ C\tilde{\varepsilon}_t + \Phi(t) S_{n-1}(t-1), & \text{si } n > 0, \end{cases}$$

où $\Phi(t) = A + B_1\varepsilon_{t-1} + \dots + B_Q \varepsilon_{t-Q}$, $t = 0, \mp 1, \dots$ ■

Preuve du théorème (2.3.5)

Dans ce qui suit, on obtient successivement les conditions de stationnarité asymptotique du 1^{er} et 2^{ième} ordre de la série \mathbf{X}_t vérifiant le modèle (2.3.17). On a :

$$E(\mathbf{X}_t) = H' E(X_t)$$

$$\text{cov}(\mathbf{X}_t \mathbf{X}_{t+s}) = H' E [((X_t - E(X_t))(X_{t+s} - E(X_{t+s})))'] H$$

Pour toute la suite on suppose que les variables aléatoires $\{\varepsilon_t\}$ sont indépendantes et chacune d'elles suit la loi $N(0,1)$.

Posant : $\mu_t = E(X_t)$, $V_t = E(X_t X_t')$, $S_t = E(X_t X_t' \varepsilon_t)$ et $W_t = E(X_t X_t' \varepsilon_t^2)$.

Condition de stationnarité du 1^{er} ordre :

Prenons l'espérance sur les deux cotés de (2.3.17) et notons $E(X_t \varepsilon_t) = C$, on obtient :

$$\begin{aligned} \mu_t &= A\mu_{t-1} + BC, \\ \Leftrightarrow \mu_{t+1} &= A\mu_t + BC, \\ &= A^2\mu_{t-1} + ABC + BC, \\ &= A^3\mu_{t-2} + A^2BC + ABC + BC, \\ &\vdots \end{aligned}$$

$$\mu_{t+1} = A^t \Gamma(1) + \left(\sum_{j=0}^{t-1} A^j \right) BC,$$

Dans le cas où $B = 0$ et $\mu_1 = 0$ alors $\mu_t = 0, \forall t \geq 1$, et de là dans ce cas aucune condition sur la matrice A est nécessaire pour la stationnarité de 1^{er} ordre.

Autrement, on procède comme suit :

Soit :

· $\rho(A)$ = le rayon spectral de la matrice A : $\rho(A) = \max \{|\lambda_i(A)|\}$ où $\lambda_i(A)$ est la $i^{\text{ème}}$ valeur propre de A .

· $\|A\|$ est une norme qq d'après (Wilkinson,1965).

Alors, du fait que $\rho(A) \leq \|A\|$, une condition suffisante pour que :

$$\lim_{t \rightarrow \infty} \left(A^t \mu(1) + \left(\sum_{j=0}^{t-1} A^j \right) BC \right) \text{ soit finie est : } \rho(A) < 1.$$

Sous cette condition, la valeur de l'espérance μ est donnée par :

$$\mu = A^t \mu + \left(\sum_{j=0}^{t-1} A^j \right) BC,$$

ce qui devient $\mu = (I - A)^{-1} BC$, lorsque $t = 1$.

Conditions de stationnarité du 2^{ème} ordre :

A partir du modèle (2.3.17), on a :

$$\begin{cases} E \{X_t \varepsilon_{t+1}\} = 0, \\ E \{X_t \varepsilon_t \varepsilon_{t+1}\} = 0. \end{cases}$$

En outre, trouver la condition de stationnarité du second ordre revient à trouver la condition sous laquelle $E(X_t X_t')$ soit finie :

Du (2.3.17a) :

$$\begin{aligned} X_t X_t' &= A X_{t-1} X_t' + B X_{t-1} X_t' \varepsilon_{t-1} + C X_t' \varepsilon_t, \text{ or} \\ X_{t-1} X_t' &= X_{t-1} X_{t-1}' A' + X_{t-1} X_{t-1}' \varepsilon_{t-1}' B' + X_{t-1} \varepsilon_t' C'. \end{aligned}$$

Donc :

$$\begin{aligned} X_t X_t' &= A X_{t-1} X_{t-1}' A' + A X_{t-1} X_{t-1}' \varepsilon_{t-1}' B' + A X_{t-1} \varepsilon_t' C' + B X_{t-1} X_{t-1}' \varepsilon_{t-1} A' + B X_{t-1} \\ & X_{t-1}' \varepsilon_{t-1} \varepsilon_{t-1}' B' + B X_{t-1} \varepsilon_{t-1} \varepsilon_t' C' + C \varepsilon_t X_{t-1}' A' + C \varepsilon_t \varepsilon_{t-1}' X_{t-1}' B' + C \varepsilon_t \varepsilon_t' C'. \end{aligned}$$

En appliquant l'espérance, on trouve :

$$V_t = A V_{t-1} A' + A S_{t-1} B' + B S_{t-1} A' + B W_{t-1} B' + C C', \quad (2.3.20)$$

où S_t et W_t sont calculée comme suit :

$$\begin{aligned} \cdot X_t X_t' \varepsilon_t &= A X_{t-1} X_{t-1}' A' \varepsilon_t + A X_{t-1} X_{t-1}' \varepsilon_{t-1}' \varepsilon_t B' + A X_{t-1} \varepsilon_t' \varepsilon_t C' \\ & + B X_{t-1} \varepsilon_{t-1} X_t' \varepsilon_t + C \varepsilon_t^2 X_{t-1}' A' + C \varepsilon_t^2 \varepsilon_{t-1}' X_{t-1}' B' + C \varepsilon_t^3 C' \\ \cdot X_t X_t' \varepsilon_t^2 &= A X_{t-1} X_{t-1}' \varepsilon_t^2 A' + A X_{t-1} X_{t-1}' \varepsilon_{t-1}' \varepsilon_t^2 B' + A X_{t-1} \varepsilon_t' \varepsilon_t^2 C' \\ & + B X_{t-1} \varepsilon_{t-1} X_t' \varepsilon_t^2 + C \varepsilon_t^3 X_{t-1}' A' + C \varepsilon_t^3 \varepsilon_{t-1}' X_{t-1}' B' + C \varepsilon_t^4 C' \end{aligned}$$

En appliquant l'espérance, on trouve :

$$\begin{aligned}
S_t &= A\mu_{t-1}C' + BCC' + C\mu'_{t-1}A' + CC'B' \\
W_t &= AV_{t-1}A' + AS_{t-1}B' + E(*) + 3CC', \text{ où} \\
* &= BX_{t-1}\varepsilon_{t-1}X'_t\varepsilon_t^2 \Rightarrow E(*) = BS_{t-1}A' + BW_{t-1}B', \text{ d'où on a bien :} \\
W_t &= AV_{t-1}A' + AS_{t-1}B' + BS_{t-1}A' + BW_{t-1}B' + 3CC' \\
&= V_t + 2CC'
\end{aligned} \tag{2.3.21}$$

Au cours de cette dérivation, on a pris en considération le fait que les $\{\varepsilon_t\}$ sont des variables aléatoires gaussiennes avec : $E\{\varepsilon_t\} = 0$, $E\{\varepsilon_t^2\} = 1$ et $E\{\varepsilon_t^4\} = 3$.

De (2.3.20) et (2.3.21) on a :

$$\begin{aligned}
V_t &= AV_{t-1}A' + AS_{t-1}B' + BS_{t-1}A' + B(V_{t-1} + 2CC')B' + CC' \\
&= AV_{t-1}A' + BV_{t-1}B' + AS_{t-1}B' + BS_{t-1}A' + 2BCC'B' + CC'
\end{aligned} \tag{2.3.22}$$

On suppose maintenant que le processus $\{x_t\}$ est stationnaire au premier ordre, on a alors :

$$\mu_t = \mu, S_t = S \text{ où } : S = A\mu C' + BCC' + C\mu' A' + CC'B'$$

Avec les notations précédentes, l'expression (2.3.22) se réécrit comme suit :

$$\begin{aligned}
V_t &= AV_{t-1}A' + BV_{t-1}B' + \Delta_1, \\
\text{où } : \Delta_1 &= ASB' + BSA' + 2BCC'B' + CC'.
\end{aligned} \tag{2.3.23}$$

Afin de trouver la condition sous laquelle, V_t tend vers V lorsque $t \rightarrow \infty$ où V ne dépend pas de t , on utilise l'opérateur vec , pour rendre l'équation matricielle précédente une équation vectorielle, qui est définie (d'après Neudecker, 1969) pour trois matrices carrées, chacune d'ordre $p \times p$, comme suit :

$$\text{Vec}(DEF) = (F' \otimes D) \text{Vec}(E), \text{Vec}(D E) = (I \otimes D) \text{Vec}(E)$$

En utilisant les notations ci - dessus, (2.3.23) devient :

$$\text{Vec}\{V_t\} = (A \otimes A + B \otimes B) \text{Vec}\{V_{t-1}\} + \text{Vec} \Delta_1$$

qui est une équation aux différences du 1^{er} ordre dans $\text{Vec}\{V_t\}$. Ainsi, pour que la solution de cette dernière converge il est suffisant que :

$$\rho(A \otimes A + B \otimes B) < 1, \tag{2.3.24}$$

qui est une condition suffisante de stationnarité asymptotique du 2^{ième} ordre de la série x_t générée à partir de (2.3.17). Supposant maintenant que la condition (2.3.24) est satisfait, on obtient alors l'expression de la variance de $\{X_t\}$:

$$V = A V A' + B V B' + \Delta_1.$$

Afin d'obtenir l'expression de la covariance, on a :

$$\begin{aligned} \cdot E(X_{t+1}X_t') &= E(A X_t X_t' + B X_t X_t' \varepsilon_t + C \varepsilon_{t+1} X_t') = AV + BS, \\ \cdot E(X_{t+2}X_t') &= AE(X_{t+1}X_t') + BE(X_{t+1}X_t' \varepsilon_{t+1}) + CE(\varepsilon_{t+2}X_t'), \\ &= AE(X_{t+1}X_t') + BC\mu', \\ \cdot E(X_{t+3}X_t') &= A^2E(X_{t+1}X_t') + ABE(X_{t+1}X_t' \varepsilon_{t+1}) + A.CE(\varepsilon_{t+2}X_t') + BAE(X_{t+1}\varepsilon_{t+2} \\ &\quad X_t') + B^2.E(X_{t+1}X_t' \varepsilon_{t+1}\varepsilon_{t+2}) + BCE(\varepsilon_{t+2}^2 X_t') + CE(\varepsilon_{t+3}X_t'), \\ &= A^2E(X_{t+1}X_t') + ABC\mu' + BC\mu'. \end{aligned}$$

$$\text{D'où : } E(X_{t+s}X_t') = \begin{cases} AV + BS, & s = 1, \\ A^{s-1}E(X_{t+1}X_t') + (\sum_{j=0}^{s-2} A^j B C) \mu', & \forall s > 1. \end{cases}$$

Soit $c(s) = E((X_{t+s} - \mu)(X_t - \mu)')$. Alors :

$$\begin{aligned} \cdot c(0) &= E((X_t - \mu)(X_t - \mu)') = V_t - \mu\mu', \\ &= AV_{t-1}A' + BV_{t-1}B' + ASB' + BSA' + 2BCC'B' + CC' - \mu\mu', \\ &= A(c(0) + \mu\mu')A' + B(c(0) + \mu\mu')B' + ASB' + BSA' + 2BCC'B' + CC' - \mu\mu', \end{aligned}$$

$$c(0) = Ac(0)A' + Bc(0)B' + \Delta_2,$$

$$\text{où : } \Delta_2 = A\mu\mu'A' + B\mu\mu'B' + ASB' + BSA' + 2BCC'B' + CC' - \mu\mu'.$$

$$\begin{aligned} \cdot c(1) &= E((X_{t+1} - \mu)(X_t - \mu)'), \\ &= E(X_{t+1}X_t' - \mu'X_{t+1} - \mu X_t' + \mu\mu'), \\ &= E(X_{t+1}X_t') - \mu\mu', \\ &= AV + BS - \mu\mu', \\ &= A(c(0) + \mu\mu') + BS - \mu\mu', \end{aligned}$$

$$c(1) = Ac(0) + \Delta_3, \text{ où : } \Delta_3 = A\mu\mu' + BS - \mu\mu'$$

$$\begin{aligned} \cdot c(2) &= E((X_{t+2} - \mu)(X_t - \mu)'), \\ &= E((AX_{t+1} + BX_{t+1} \varepsilon_{t+1} + C \varepsilon_{t+2} - \mu).(X_t - \mu)'), \\ &= AE(X_{t+1}(X_t - \mu)') + BE(X_{t+1} \varepsilon_{t+1} X_t') - B\mu'E(X_{t+1} \varepsilon_{t+1}) - \mu E(X_t - \mu)', \\ &= AE(X_{t+1}X_t') - A\mu\mu' + BC\mu' - B\mu'C, \\ &= A(c(1) + \mu\mu') - A\mu\mu', \\ &= Ac(1) + A\mu\mu' - A\mu\mu', \end{aligned}$$

$$c(2) = Ac(1),$$

⋮

$$\cdot c(s) = Ac(s-1) = A^{s-1}c(1). \quad s = (2, 3, \dots). \blacksquare$$

Bibliographie

- [1] Abdullah, A.S and Taylan, A.U (1997) . Periodic stationarity conditions for periodic autoregressive moving average processes as eigenvalue problems. *Water resources research*, vol **33**, n°08, 1929-1934.
- [2] Akaïke, H.(1974a) . Markovian representation of stochastic processes and its application to the analysis of autoregressive moving average processes. *Ann. Inst. Stat. Math*, **26**, 363-387.
- [3] Akaïke, H.(1974b) . Stochastic theory of minimal realizations. *IEEE. Transactions on Automatic Control*, **19**, 667-674.
- [4] Aknouche, A., (2006). Etude théorique et algorithmique de la modélisation des séries chronologiques périodiques. Thèse de Doctorat. Faculté de Mathématiques, U.S.T.H.B. Alger.
- [5] Aknouche, A.and Guerbyenne, H. (2009). Periodic stationarity of random coefficient periodic autoregressions. *Statistics and Probability Letters*, **79**, 990-996.
- [6] Aknouche, A.and Nadia, R. (2010).On an independent and identically distributed mixture bilinear time-series model. *Journal of Time Series Analysis*, **31**, 113–131.
- [7] Benghabrit, Y, (1992). Inférence asymptotique paramétrique et non paramétrique pour les modèles bilinéaires. Thèse de Doctorat en Sciences. Institut de mathématiques, U.L.B, Bruxelles, Algérie.
- [8] Bentarzi, M. and Hallin, M.(1994). On the Invertibility of Periodic Moving Average Models. *Journal of Time Series Analysis*. **Vol 15**, n°3, 263-268.
- [9] Bentarzi, M., (1995). Modèles de séries chronologiques à coefficients périodiques. Thèse de Doctorat en Sciences. Institut de mathématiques, U.S.T.H.B, Alger, Algérie.

- [10] Berline, A. and Francq, C (1990). Stationnarité et identification d'un processus purement bilinéaire et strictement superdiagonal. *Statistique et Analyse des Données*, **15**, 1–24.
- [11] Bibi, A., (2002). Quelques contributions dans l'analyse des modèles bilinéaires à coefficients dépendant du temps. Thèse de Doctorat en Sciences. Institut de Mathématiques, Université Mentouri, Constantine, Algérie.
- [12] Bibi A⁽¹⁾, Oyet AJ⁽²⁾ (2002) A note on the properties of some time varying bilinear models. *Statistics & Probability Letters*, **58**, 399–411.
- ⁽¹⁾Département de Mathématique, Université Mentouri de Constantine, Algeria.
- ⁽²⁾Department of Mathematics and Statistics, Memorial University of Newfoundland, Canada.
- [13] Bibi, A. (2003). On the covariance structure of the time-varying bilinear models, *Stochastic Anal. Appl*, **21** (1), 25–60.
- [14] Bibi, A.⁽¹⁾ and Moon-Ho Ringo Ho⁽²⁾ (2004). Properties of some bilinear models with periodic regime switching. *Statistics & Probability Letters*, **69**, 221–231.
- [15] Bibi A, Oyet AJ (2004). Estimation of some bilinear time series models with time-varying coefficients. *Stoch Anal Appl*, **22**, 355–376.
- [16] Bibi, A. (2004). On the stability and causality of general time-dependant bilinear models. *C. R. Acad. Sci, Ser. I*, **338**, 245–248
- [17] Bibi, A. (2005). A note on the stability and causality of general time-dependent bilinear models. *Statistics & Probability Letters*, **73**, 131–138
- [18] Bibi, A. (2005). Stationnarité et inférence asymptotique de modèles bilinéaires périodiques. *C. R. Acad. Sci. Paris, Ser. I*, **341**, 679–682.
- [19] Bibi, A. and Gautier, A. (2005). Stationnarité et inférence asymptotique de modèles bilinéaires périodiques. *C. R. Acad. Sci. Paris, Ser. I*, **341**, 679–682
- [20] Bibi, A. and Gautier, A., (2006). Propriétés dans \mathbb{L}^2 et estimation des processus purement bilinéaires et strictement superdiagonaux à coefficients périodiques. *Canad. J. Statist*, **34**(1), 131–148.

- [21] Bibi, A.⁽¹⁾ and Moon-Ho Ringo Ho⁽²⁾.(2006). Estimation of Periodic Bilinear Time Series Models. *Communications in Statistics-Theory and Methods*, **35**, 1745-1756.
- ⁽¹⁾ Département de Mathématique, Université Mentouri de Constantine, Algeria.
- ⁽²⁾ Department of Psychology, McGill University, Montreal, Quebec, Canada.
- [22] Bibi, A. and Lessak, R. (2009). On stationarity and β -mixing of periodic bilinear processes. *Statistics and Probability Letters*, **79**, 79-87.
- [23] Bibi, A and Aknouche, A. (2010). Stationnarité et β -mélange des processus bilinéaires superdiagonaux à changement de régime Markovien, *Comptes Rendus de l'Académie des Sciences*. A paraître.
- [24] Bibi, A and Aknouche, A. (2010). Yule-Walker type estimators in periodic bilinear models : strong consistency and asymptotic normality. *Stat Methods Appl*.
- [25] Billingsley, P., (1995). Probability and Measure, 3rd ed. Wiley, New York.
- [26] Bougerol, P. and Picard, N. (1992). Strict stationarity of generalized autoregressive processes. *Annals of probability*, **20**, 1714-1730.
- [27] Bouzaachane, K., (2006). Modèles de séries chronologiques bilinéaires : Estimations, Algorithmes et Applications. Thèse de Doctorat en Sciences Appliquées. Institut de mathématiques, Université Mohammed V-Agdam. Ecole Mohammadia d'ingénieurs.
- [28] Box, G. E. P. and Jenkins, G. M. (1976). *Time Series Analysis, Forecasting and Control*, 2nd ed., Holden-Day San Francisco, CA.
- [29] Brandt, A., (1986). The stochastic equation $Y_{n+1} = A_n Y_n + B_n$ with stationary coefficients. *Advanced in Applied Probability*, **18**, 211-220.
- [30] Breiman, L., (1968). Probability. Addison Wesley, California.
- [31] Brockwell P. J. and Davis R. A. (1991). *Time Series : Theory and Methods*, Springer-Verlag, New York, .
- [32] Diane, B., (1997). Modélisation stochastique des caisses de retraite. Thèse de Philosophie Doctor (Ph. D.) en Mathématiques Appliquées. Faculté des études supérieures, université de Montréal. Alger, Constantine.
- [33] Ding, Z., Granger, C.W.J. and Engle, R.F. (1993). A long memory property of stock market returns and a new model. *J. Emp. Finance*, **1**, 83-106.

- [34] Engle, R.F. and V.Ng (1993) . Measuring and Testing tne Impact of News on Volatility. *Journal of Finance*, **48**, 1749-1778.
- [35] Faurre, P, Clerget, M, Germain, F (1979) . Opérateurs rationnels. Dunod.
- [36] Geweke, J. (1986) . Modelling the persistence of Conditional Variances : *Comment. Econometric Reviews*, **5**, 57-61.
- [37] Giraitis, L.R. Leipus, P.M. Robinson and D. Surgaïlis (2004) . LARCH, Leverage and Long Memory. *Journal of Financial Econometrics*, **2**, 177-210.
- [38] Gladyshev, E.G., (1961), Periodically correlated random sequences. *Soviet Mathematics*, **2**, 385–388.
- [39] Granger, C.W.J and Anderson, A.P (1978*b*) . On the invertibility of time series models. *Stoch. Proc and their application*, **8**, 87-92.
- [40] Guégan, D. (1981) . Etude d'un modèle non linéaire, le modèle superdiagonal d'ordre 1. *C. R. Acad., Sc*, t.**293**, série I- 95.
- [41] Guégan, D. (1983) . Cadre d'étude pour des modèles non linéaires. *C. R. Acad. Sc*, t.**296**, série I- 167.
- [42] Guégan, D. (1987) . Different representations for bilinear models. *J.Time Ser. Anal* , **8**, 38-58.
- [43] Guégan, D. and Diebolt, J. (1994). Probabilistic properties of the ■-ARCH model. *Statist.Sinica*, **4**, 71-87.
- [44] Guégan, D.and Pham, D.T.(1987) . Minimalité et inversibilité des modèles bilinéaires à temps discrets. *C. R. Acad. Sc*, t.**304**, Série I, n°6.
- [45] Guégan, D. (1988) . Modèles bilinéaires et polynomiaux de séries chronologiques : Etude probabiliste et analyse statistique. Thèse de Doctorat en Sciences. Institut de mathématiques, Université Joseph Fourier-Grenoble I.
- [46] Guégan, D.and Pham, D.T.(1989) .A Note on the Estimation of the Parameters of the Diagonal Bilinear Models by the Least Squares Method, *Scand. Journ of Stat. Th and Appl*, **16**, 129-136.
- [47] Guégan, D.and Pham, D.T.(1992) .Power of the Score Test against bilinear time series models. *Statistica Sinica* **2**, 157-170.

- [48] Guégan, D.(1994). *Séries chronologiques non linéaires à temps discret*, Economica.
- [49] Guerbyenne, H. (2007) :Modèles de séries chronologiques conditionnellement hétéroscédastiques.Thèse de Doctorat. Faculté de Mathématiques, U.S.T.H.B. Alger.
- [50] Gyorgy, T.(1990). Second-Order Properties for Multiple-Bilinear Models. *Journal of multivariate analysis*, **35**, 295-307.
- [51] Hamdi, F., (2008). Modèles Espaces d'Etats et Processus Périodiquement Corrélés. Thèse de Doctorat. Faculté de Mathématiques, U.S.T.H.B. Alger.
- [52] Hannan, E.J.(1982). A note on bilinear time series models. *Stoch Proc and their Applications*, **12**, 221-224.
- [53] Hentschel, L.(1995). All in the family : Nesting Symmetric and Asymmetric GARCH Models. *Journal of Financial Econometrics*, **39**, 71-104.
- [54] Jou, J.J, Shaman, P , (1988). The third order cumulant function and bispectral density of the bilinear time series model $BL(p, 0, p, 1)$, *Tech. Repport. Philadelphia*.
- [55] Kalman, R.E (1960). A new approach to linear filtering and predicting problems, *Trans. ASME, J. Basic Eng*, **82D**, 35–45.
- [56] Kalman, R.E, Falb, P.L et Arbib, M.A (1969). *Topics in Mathematical System Theory*, McGraw-Hill, New York.
- [57] Kalman, R.E et Bucy, R.S (1961). New results in linear filtering and prediction problems, *Trans. ASME, J. Basic Eng*, **83D**, 95–108.
- [58] Kâ Diongue*, A., Guégan, D** .et Wolff, R.C***. Exact Maximum Likelihood estimation for the BL-GARCH model under elliptical distributed innovations.
* Université Gaston Berger, UFR SAT, BP 234, Saint-Louis, Senegal.
** Paris School of Economics, MSE-CES, University Paris I Panthéon-Sorbonne, France.
*** School of Mathematical Sciences, Queensland University of Technology, Australia.
- [59] Kim, W.K. Billard, L (1989). Third order moment structure and asymptotic properties for the first order bilinear time series models, *Tech. Repport. Georgia*.
- [60] Kristensen, D. (2009) On stationarity and ergodicity of the bilinear model with applications to GARCH models. *Journal of Time Series Analysis*, **30**, 125–144.

- [61] Liu, J.et Brockwell, P.J.(1988). On the general bilinear time series model. *J. Appl. Prob.* **25**, 553-564.
- [62] Liu, J.(1989). A simple condition for the existence of some stationnary bilinear time series.*J. Time Ser. Anal.* **10**, 33 - 39
- [63] Liu, J.(1990). A note on causality and invertibility of a general bilinear time series model. *Adv. Appl. Prob.* **22**, 247-250
- [64] Magnus, J.R. and Neudecker, H., (1989), *Matrix Differential Calculus* (NewYork : JohnWiley).
- [65] Meyn, P.S.et Tweedie, R.L (1993) . *Markov chains and stochastic stability*. Springer-Verlag. New York.
- [66] Nelson, D. (1991) . Conditional Heteroskedasticity in Asset Returns : *A New Approach*. *Econometrica*, **59**, 347-370.
- [67] Nicholls, D. F. and Quinn, B. G. (1982). *Random Coefficient Autoregressive Model : An Introduction*. New York : Springer Verlag.
- [68] Ohakwe, J. , (2008). Tables and moments of the purely diagonal bilinear time series model. Ph. D Thesis, Abia State University, Uturu, Nigeria.
- [69] Ohakwe, Johnson⁽¹⁾.et Iheanyi S.Iwueze⁽²⁾. Characterizations of the moments of the purely diagonal bilinear time series model of order one.
- ⁽¹⁾ Department of Statistics, Faculty of Biological and Physical Sciences, Abia State University, P.M.B.2000, Uturu, Abia State, Nigeria.
- ⁽²⁾ Department of Physical Sciences, College of Natural and Applied Sciences, Renaissance University, P.M.B 01183, Ugbawka, Enugu State, Nigeria.
- [70] Olivier, G. (2010) . Modèles probabilistes sur réseaux. Cours spécialisé, 2ième année Master. University de Nancy 1.
- [71] Pantula, S.G.(1986) . Modelling the Persistence of Conditional Variances : *Comment*. *Econometric Reviews*, **5**, 71-73.
- [72] Pham, D.T et Lanh T.Tran.(1980) . Quelques résultats sur les modèles bilinéaires de séries chronologiques. *C.R.Acad. Sc*, **t.290**, série A-335.

- [73] Pham, D.T et Lanh T.Tran.(1981) . On the first order bilinear time series model. *J. Appl. Prob*, **18**, 617-627.
- [74] Pham, D.T.(1985) Bilinear Markovien representation and bilinear models. *Stoch. Process. Appl*, **19**, 297-303.
- [75] Pham, D.T et Lanh T.Tran (1985) . Some mixing properties of time series models. *Stoch. Process. Appl*, **20**, 295-306.
- [76] Pham, D.T.(1986) . The mixing property of Bilinear and Generalised Random Coefficient Autoregressive Models. *Stoch. Process. Appl*, **23**, 291-300.
- [77] Priestly, M.B.(1988) . *Non linear and Non stationary Time Series Analysis*. Time series. Analysis, 519.5.
- [78] Quinn, B.G.(1982) . Stationarity and invertibility of simple bilinear models. *Stochastic Processes and their Applications*, **12**, 225-230.
- [79] Robinson, P.M. (1991) . Testing for Strong Serial Correlation and Dynamic Conditional Heteroskedasticity in Multiple Regression. *Journal of Econometrics*, **47**, 67-84.
- [80] Saidi, Y (2003) . Etude probabiliste et statistique de modèles conditionnellement hétéroscédastiques non linéaires. Thèse de Doctorat en Mathématiques Appliquées aux Sciences Economiques, Université Lille 3.
- [81] Seasay, S. A. O. and Subba Rao, T. (1988) : Yule - Walker type difference equations for higher - order moments and cumulants for bilinear time series models. *J. Time Ser. Anal.* **9**, 385 - 401
- [82] Seasay, S. A. O. and Subba Rao, T. (1991) : Difference equations for higher order moments and cumulants for the bilinear time series model $BL(p, 0, p, 1)$. *J. Time Ser. Anal.* **12**, 159 - 176.
- [83] Seymour,L (1973).*Probability : Cours et problemes. Serie Schaum, temple university.*
- [84] Subba Rao, T.S. et Gabr, M.M. (1984). *An introduction to bispectral analysis and bilinear time series analysis*, Lecture Notes in Statistics, New York, Springer, n°24.
- [85] Subba Rao, T.S (1981) . On the theory of bilinear time series models. *J.R. Statist. Soc. B*, **43**, n°2, pp.244-255.

- [86] Subba Rao, T.S (1997) . Statistical analysis of nonlinear and nongaussian time series models. In Stochastic Differential and Difference Equations; *Cscizar, I., Michaeltzky, G., Eds* ; Birkhäuser Boston, Inc : Basel, 285-298.
- [87] Tiao, G.C.and Grupe, M.R.(1980). Hidden periodic autoregressive-moving average models in time series data, *Biometrika*, **67** , pp. 365–373.
- [88] Tweedie, R.L.(1975) . Sufficient conditions for ergodicity and recurrence of Markov chains on general state space. *Stoch. Process. Appl*, **03**, 385-403.
- [89] Tweedie, R.L.(1983) . The existence of moments for stationary Markov chains. *J. Appl Prob*, **20**, 191-196.
- [90] Wei, W.W.S (1990) . *Time Series Analysis : Univariate and Multivariate Methods*. Addison-Wesley, New York.