République Algérienne Démocratique et

Populaire Ministère de l'Enseignement Supérieur et de la Recherche Scientifique

Université Mentouri de Constantine

FACULTE DES SCIENCES EXACTES

Département de Mathématique

N° d'ordre: ....

Série: ....

Mémoire
Présenté pour l'obtention du diplôme de
MAGISTER
OPTION
Probabilités et statistiques
Par
GHEZAL Ahmed
Intitulé

Analyse statistique des processus bilinéaires à changement de régimes markoviens

Soutenue le 29/01/2012, devant le jury composé de:

Mr.	F. L Rahmani	MC	U.M.C	Président
Mr.	A. Bibi	PR	U.M.C	Rapporteur
Mme.	S. Belaloui	MC	U.M.C	Examinatrice
Mr.	S. Djezzar	PR	U.M.C	Examinateur

Promotion : 2009-2012

ملخص

هذه المذكرة تعالج بعض مشاكل السلاسل الزمنية غير الخطية. القسم الخاص من النماذج غير الخطية الذي تمت دراسته بشكل متسع هو قسم النماذج BL ماركوف المترابط التبديل.

الشروط اللازمة والكافية الاستقرار بصفة دقيقة و من الدرجة الثانية.

الهدف الرئيسي من هذه المذكرة هو التقدير في النماذج BL ماركوف المترابط التبديل باستعمال

طريقة MLE انطلاقا من شروط نحصل على تقديرات ثابتة وتسلك سلوك تقاربيا يتبع القانون الطبيعي.

كلمات المفتاحية

النماذج غير الخطية، المترابط BL، ماركوف التبديل، التمثيل الشعاعي، ضرب كرنكر، الاستقرار،

MLE، الاستقرارية الطافية، السلوك المقارب.

Table des matières

0.1 Remerciements

Je remercie, tout d'abord, Dieu tout puissant, pour avoir guidé mes pas vers un avenir inchaallah prometteur, et le travail, la persévérance et la quête du savoir.

Je remercie vivement Monsieur le professeur ***A. Bibi****, de l'Université Mentouri de Constantine, pour ses critiques constructives et ses encouragements, grâce à son soutien constant et à ses conseils précieux qu'il m'a prodigué.*

Je reste très reconnaissant à Monsieur ***F. L Rahmani****, Maitre de Conférences à l'Université Mentouri de Constantine, pour l'honneur qu'il ma fait en acceptant de présider le jury de ce mémoire.*

Je remercie aussi vivement Madame ***S. Belaloui****, Maitre de Conférences à l'Université Mentouri de Constantine, pour sa participation au jury.*

Je remercie également et profondément Monsieur le professeur ***S. Djezzar*** *de l'Université Mentouri de Constantine, pour sa participation au jury.*

Enfin, je remercie aussi ma famille pour le soutient chaleureux apporté lors de la réalisation de ce travail et hors unique pour remercier tous les enseignants et les professeurs qui étaient des étudiants il ya de la phase préparatoire à l'université, et merci à tous ceux qui ont contribué de près ou de loin à la réalisation de ce travail.

0.2 Dédicaces

Je dédie ce mémoire à ma mère et mon père miséricorde de Dieu, et la sœur Nesrine, et ma mère frères et ma grand-mère et tous famille et mes amis et mes collègues et de tous mes maîtres, en particulier professeur A. Bibi et l'enseignant Saheb Abd al-Majid en lycée 5 Juillet 1961 et tous les travailleurs et enseignants du lycée Chihani Bachir-Teleghma et ses élèves.

Chapitre 1

Introduction

1.1 Motivations

Les séries temporelles sont considérées à tort comme étant une branche exclusive de l'économétrie. Cette dernière est une discipline qui est relativement jeune alors que les séries temporelles ont été utilisées bien avant, par exemple en astronomie (1906) et en météorologie (1968).

L'objet des séries temporelles est l'étude des variables au cours du temps. Même s'ils n'ont pas été à l'origine de cette discipline, ce sont les économètres qui ont assuré les grandes avancées qu'a connues cette discipline (beaucoup de « Prix Nobel » d'économie sont des économètres).

Parmi ses principaux objectifs figurent la détermination de tendances au sein de ces séries ainsi que la stabilité des valeurs (et de leur variation) au cours du temps. Citons par exemple : le volume des ventes hebdomadaires d'un produit, le prix des actions de la banque de clôture du jour, le volume de la production quotidienne de pétrole brut en Algérie, le taux de chômage dans une période connue, ...Ces applications nécessitent beaucoup de recherches et d'expériences. Les techniques standards d'analyse de séries temporelles ont longtemps reposé sur les propriétés fondamentales de linéarité et stationnarité. L'essor considérable qu'a connu l'analyse statistique des séries chronologiques au cours des ces derniers trois décennies est lié essentiellement au développement de l'approche temporelle sous deux hypothèses remarquables :

Stationnarité et linéarité.

L'analyse des séries chronologiques a été considérablement développée depuis la publication de l'ouvrage dev Box et Jenkins [1] (1970) qui a été décisive. En effet, dans l'ouvrage les deux auteurs développent le très populaire modèle $ARMA$ (Auto Regressive Moving Average). Les modèles ($ARMA$) ont ainsi fait l'objet d'un intérêt croissant sur une vaste étendue disciplinaire allant de l'économétrie et la finance à la climatologie ou l'électrotechnique. Cependant, de nombreuses recherches ont démontré que les hypothèses de linéarité n'étaient qu'un pis-aller utopique apportant un confort appréciable dans l'étude probabiliste et statistique du modèle, et cette classe de processus ($ARMA$) jouera le rôle prépondérant dans notre modélisation concrète des processus stationnaires, tandis que la classe encore plus large des processus linéaires caractérisés par leurs flexibilités, facilités d'utilisation est ses utilisation pour prédire les valeurs futures, ils sont facile à interpréter parmi la plupart des autres modèles. Cependant, les méthodes d'analyse et d'inférence statistique sont mieux développés dans cette classe de processus, et de plus elles peuvent être employées comme une analyse préliminaire. Habituellement, elles donnent une représentation parcimonieuse et interprétable.

Comme il a été mentionné par Hallin (1987), dans de nombreux cas l'hypothèse qui se trouve à la base de l'utilisation des méthodes de Box et Jenkins (dans le cas non stationnaire) peut escamoter les problèmes de la non stationnarité plutôt que de les résoudre. Cependant, et à partir des années 70, et dans le but de résoudre certains problèmes liés à la non stationnarité, on trouve un intérêt croissant aux modèles dont les coefficients eux mêmes sont susceptibles d'évoluer avec le temps. On peut distinguer deux

[1] George BOX et Gwilym JENKINS sont deux statisticiens qui ont contribué, dans les années (1970), à populariser la théorie des séries temporelles univariées. Les procédures de modélisation sont présentées dans leur célèbre ouvrage intitulé « Time Series Analysis : Forecasting and control ».

Ils ont proposé une démarche générale de prévision pour les séries chronologiques. Cette démarche est fondée sur la notion de processus ($ARMA$) et elle comprend quatre phases : l'identification a priori, l'estimation du modèle ($ARMA$) identifié, l'identification a posteriori et la prévision.

classes de modèles à coefficients dépendant du temps selon que cette évolution est de nature déterministe ou stochastique

Evolution déterministe Dans cette classe de modèles qui visent à décrire des processus de nature linéaires mais non stationnaires, que l'on conviendra d'appeler évolutifs, où les trajectoires des coefficients s'expriment comme des combinaisons linéaires de fonction du temps, supposées connues a priori et en nombre fini. D'une part et du point de vue technique, l'idée de faire une projection des trajectoires des coefficients sur une base de fonctions était déjà sous-jacente aux travaux de Mendel (1969) et Rao (1970). Dans ces travaux, la représentation évolutive apparait plutôt comme une "astuce" de calcul. Liporace (1975) a été le premier à étudier l'estimation d'un modèle autorégressif suivi par Hall, Oppenheim et Willsky (1983). Grenier (1986) a travaillé sur ce point et a proposé un jeu d'algorithmes rapides adaptés à plusieurs variantes de modèles. D'autre part, Cramer (1961) a étendu le résultat fondamental de Wold au cas non stationnaire (avec variance finie) donc nous parlerons désormais de la décomposition de Wold-Cramer, et les mêmes arguments ; prévision et modélisation, peuvent être obtenus à partir de la décomposition de Wold-Cramer en utilisant des modèles $ARMA$ à coefficients dépendant du temps. Dans cette perspective, nous trouvons Mélard (1985) et Hallin (1986, 1989), et autres (cf. Priestley (1988)), ont enrichi la littérature des séries chronologiques avec de nombreux travaux fondamentaux.

Evolution stochastique C'est la classe des modèles qui a reçu jusqu'a présent le plus d'attention surtout dans les domaines de l'économétrie, l'automatique, le traitement du signal, et des séries chronologiques. Ces modèles (et comme dans le cas linéaire à coefficients constants) visent à décrire principalement des processus de nature stationnaires, mais non linéaires. La littérature sur ces modèles étant assez dispersée. La monographie de Nicolls et Quinn (1982) et la bibliographie qu'elle contient font d'excellentes références dans ce domaine. Plusieurs articles, citons Andel (1976) et Andel (1982) font appel non seulement à l'erreur associée à la spécification du modèle, mais aussi à la nature même des rapports économiques pour justifier l'existence d'une composante non-déterministe dans les coefficients de modèles. Cette hypothèse nous semble aussi plus naturellement adaptée lors de l'analyse des séries chronologiques multiples (cf. Hannan (1970)).

Ainsi, face à une réalisation $\{X_1, X_2, \ldots X_N\}$ d'un processus $(X_t)_{t\in\mathbb{Z}}$, $\mathbb{Z} = \{0, \pm 1, \pm 2, ...\}$, un statisticien, un économètre cherchent à identifier le "bon" modèle. Donc, il est devant deux choix fondamentaux de modèles : modèles linéaires et modèles non linéaires, bien que la frontière entre les concepts de linéarité et de non linéarité soit difficile à concevoir, Rao et Gabr (1984) ont observés que la série décrivant le taux de chômage dans l'Allemagne de l'ouest a un comportement non linéaire et que les prévisions obtenues en utilisant des modèles bilinéaires sont mieux que celles obtenues par des modèles $ARMA$. Par conséquent, et durant ces deux dernières décennies, les modèles non linéaires ont reçus plus d'attention. Une classe particulière de modèles stationnaires et non linéaires qui est introduite dans la littérature de la théorie du contrôle, et qui a trouvé d'autres champs d'applications (cf. Mohler (1988)) est la classe des modèles bilinéaires. Cette classe de modèles, qui peut être regardée comme une extension des modèles $ARMA$, a été suffisamment étudiée par Granger et Andersen (1978a), Pham et Tran (1981), Subba Rao (1981), Subba Rao et Gabr (1984), Guégan (1994), Shu-Ing (1985), Priestley (1988), Liu et Brockwell (1988), Liu (1989) , 1992, et Terdik (2000) et autres. Notons aussi qu'une sous classe particulière de modèles bilinéaires peut être utilisée comme des résidus dans la représentation $ARMA$ (cf. Francq (1999), Francq et Zakoian (2000)). Dans plusieurs situations pratiques, nous sommes devant des données générées par un certain processus non seulement non linéaire mais aussi non stationnaire. Kendall en 1953 a déjà mentionné que *"No economic system yet observed has been stationary over long periods [...]. It seems natural [to consider] the case when the constants [in the model] are themselves slowly moving through time as the economy changes"*. Ceci suggère de considérer des modèles à coefficients dépendant du temps. 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). Il existe cependant des séries qui sont soupçonées d'être à la fois non linéaires et non stationnaires, 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 dépendant du temps nous semble raisonnable.

Motivés par la précédente discussion, nous allons abondonner l'hypothèse de linéarité, bien que ces modèles ont été relativement utilisés, nous allons néanmoins consacrer notre travail dans un cadre probabiliste et statistique à l'étude (bien entendu non exhaustive) d'une classe de processus $(X_t)_{t\in\mathbb{Z}}$ définis sur un espace de probabilité $(\Omega, \Im, P)$ non linéaires. C'est la classe de processus générés par des modèles bilinéaires à changements de régimes markoviens définis par l'équation aux différences stochastiques suivante

$$X_t = \sum_{i=1}^{p} a_i(s_t)X_{t-i} + \sum_{i=1}^{q} b_i(s_t)\varepsilon_{t-i} + \sum_{i=1}^{P}\sum_{j=1}^{Q} c_{ij}(s_t)X_{t-i}\varepsilon_{t-j} + \varepsilon_t \tag{0.1}$$

noté $MS - BL\,(p, q, P, Q)$ où $(a_i(s_t))_{1\leq i\leq p}\,, (b_i(s_t))_{1\leq i\leq q}\,, (c_{ij}(s_t))_{1\leq i\leq P, 1\leq j\leq Q}$ sont des fonctions bornées, déterministes, dépendant éventuellement d'une chaîne de Markov à éspace d'état fini, i.e., $\mathbb{S} = \{1, ..., d\}$ et où $(\varepsilon_t)_{t\in\mathbb{Z}}$ est un processus de bruit blanc fort centré de variance finie et satisfaisant l'hypothèse suivante

$$\left\{\varepsilon_t \text{ et } X_s \text{ sont indépendants pour tout } s < t\right\}.$$

La classe de modèles (0.1) contient trois sous classes

1. La classe de modèles $MS - ARMA\,(p, q)$ peut être obtenue en posant $c_{ij}(.) = 0$ pour tout i, j, donc (0.1) est une extension naturelle des processus $MS - ARMA$.
2. La classe de modèles superdiagonaux obtenue en supposant $c_{ij}(.) = 0$ for $i \leq j$.
3. La classe de modèles sous diagonaux obtenue en supposant $c_{ij}(.) = 0$ for $i > j$.

Du fait de la dépendance non linéaire entre X_t et ε_{t-k}, $k \geq 1$ ceci rend délicat la manipulation des termes de type $X_t\varepsilon_{t-j}, j > 0$. Pour cette raison, seule la classe des modèles super-diagonaux a reçu plus d'attention. Notons que, si dans le modèle général (0.1) les coefficients sont constants et le bruit blanc $(\varepsilon_t)_{t\in\mathbb{Z}}$ est stationnaire, nous trouvons ainsi une littérature abondante. Cette abondance est dûe aux conditions sous lesquelles le modèle devient stationnaire et ergodique. Cependant de nombreux travaux de recherche existent sur le développement des propriétés probabilistes et statistiques, l'identification, les tests et l'estimation des paramètres de certains modèles bilinéaires stationnaires (pour une bibliographie récente voir Terdik (2000)). En revanche, lorsque le modèle est non stationnaire les méthodes classiques sont inapplicables.

Certes l'étude des modèles bilinéaires à coefficients dépendant du temps est loin d'être achevée. De nombreux problèmes restent ouverts. Néanmoins on peut se demander s'il est possible de résoudre, par exemple, le problème de l'identification de certains modèles bilinéaires à coefficients dépendant du temps comme ce fut le cas pour certains modèles de séries chronologiques bilinéaires stationnaires, dans la mesure où la classe des modèles considérés est très riche et très complexe. Par contre la théorie des tests qui jusqu'à présent a été peu étudiée (dans le cas stationnaire) doit permettre d'aboutir assez rapidement à quelques résultats : outre les tests de stationnarité (respectivement de linéarité) pour lesquels quelques procédures ont été proposées sous l'hypothèse de la linéarité (respectivement sous l'hypothèse de stationnarité) on a besoin de tests portant sur le choix de la nature des coefficients, autrement dit le choix des modèles.

1.2 Le contenu de la thèse

La thèse examine certaines propriétés structurelles de modèle bilinéaire à changement de régime Markovien $MS-BL$. Dans cette recherche, des conditions nécessaires et suffisantes assurant l'existence de solutions stationnaires (au sens faible et forte) et de l'inversibilité des solutions sont donnés. En conséquence, nous observons que la stationnarité locale du processus observé n'est ni suffisante ni nécessaire, d'obtenir la stationnarité globale. Existence de moments finis et $\boldsymbol{\beta}-$mélange pour les modèles super-diagonale bilinéaire sont obtenus, et le théorème limite pour les modèle super-diagonaux. D'autre part, la consistance et la normalité asymptotique de l'estimation par MLE pour les modèles $MS-BL$ bilinéaire sont envisagés.

Première partie

Les propriétés probabilistes des processus $MS - BL(p, q, P, Q)$

Chapitre 2

Les modèles bilinéaires à changements de régimes markoviens

2.1 Introduction

De nombreuses séries rencontrées en pratique comme par exemple en sciences biologiques, en finance sont non linéarité,et qui ne peuvent être expliquées en utilisant des modèles $ARMA$, et donc le recours aux modèles non linéaires est inévitable. Cependant les modèles bilinéaire (**BL**) ont reçus récemment une grande attention pour modéliser des phénomènes non gaussiens. Le développement stupéfiant de ces modèles est basé sous l'invariance de ses paramètres, mais cette hypothèse ne peut être appropriée dans plusieurs situations.

Les modèles à changements de régimes markoviens c'est une classe importante de modèles caractérisés par sont dépendance de ses paramètres en une chaîne de Markov caché et à espce d'état fini. Les changements de régime peuvent être lisses ou brutal et ils peuvent se produire fréquemment ou occasionnellement selon une probabilités de transition. L'histoire des modèles à changements de régimes soulève aux œuvres de Quandt (1958) et Goldfeld et Quandt (1973). L'analyse des cycles économiques a été longtemps limitée aux techniques d'analyse linéaire, qui sont incapables de suivre le rythme rapide et accéléré de l'évolution constante de l'économie d'un pays.

Cela a conduit à un intérêt croissant pour les modèles non linéaires capables de traiter les régimes de commutation associés à une série macroéconomique. Beaucoup de séries temporelles économiques parfois présenter des ruptures dramatiques dans leur comportement, associée à des événement tels que les crises financières (Jeann et Masson (2000)) et par conséquent nous avons besoin de modèles non linéaires et stationnaires par exemple les modèles markoviens (**MSM**) qui ont reçus récemment une grande attention et qui sont capables de décrire adéquatement des séries temporelles aux différentes périodes.

Un (**MSM**) à temps discret est un processus aléatoire à deux variables $(X_t, \mathbf{s}_t)_{t\in\mathbb{Z}}$ tels que

(i) $(\mathbf{s}_t)_{t\in\mathbb{Z}}$ est une chaîne de Markov homogène (appelé désormais "régime") à éspace d'état fini,

(ii) La distribution conditionnelle de X_k donne $(X_{t-1}, \mathbf{s}_t)_{t\leq k}$ dépend uniquement de $(X_{t-1}, \mathbf{s}_k)_{t\leq k}$

L'utilisation du **MSM** est devenue de plus en plus populaire dans la modélisation des données économétriques, et les modèles **MSM** s'avère un outil puissant pour l'analyse des cycles économiques et donner lieu à des distributions très souple.

2.2 Mise en équation et représentation Markovienne

Un processus $(X_t)_{t\in\mathbb{Z}}$ est dit bilinéaire à changements de régimes markoviens (noté $\mathbf{MS-BL(p,q,P,Q)}$), si $(X_t)_{t\in\mathbb{Z}}$ est généré par des équations de la forme

$$\forall t \in \mathbb{Z},\ X_t = \sum_{i=1}^{p} a_i(\mathbf{s}_t) X_{t-i} + \sum_{j=0}^{q} b_j(\mathbf{s}_t)\varepsilon_{t-j} + \sum_{i=1}^{P}\sum_{j=1}^{Q} c_{ij}(\mathbf{s}_t) X_{t-i}\varepsilon_{t-j} \tag{2.1}$$

où $(\varepsilon_t)_{t\in\mathbb{Z}}$ est un processus strictement stationnaire, ergodique, les fonctions $a_i(\mathbf{s}_t), b_j(\mathbf{s}_t)$ et $c_{ij}(\mathbf{s}_t)$ dépendent d'une chaîne $(\mathbf{s}_t)_{t\in\mathbb{Z}}$ de Markov homogène à espace d'état fini $\mathbb{S} = \{1; 2; \ldots; k\}$ et indépendante de $(\varepsilon_t)_{t\in\mathbb{Z}}$.

Le modèle (2.1) contient comme cas particuliers plusieurs classes de modèles intéressants ayant étudié dans la littérature, en effet

(i) Les modèles $\mathbf{BL}(p, q, P, Q)$ standard. Ces modèles sont obtenus en supposant que la chaîne $(\mathbf{s}_t)_{t\in\mathbb{Z}}$ a un seul régime (par exemple, Granger et Anderson (c.f.,[23])).

(ii) Les modèles de Markov cachés[1] (**HMM**) (Hidden-Markov Models) En contraste avec (**MSM**), (**HMM**) sont caractérisés par le fait que compte tenu $(\mathbf{s}_t)_{t\in\mathbb{Z}}$, $(X_t)_{t\in\mathbb{Z}}$ est une suite de variables aléatoires indépendantes. Cette classe est obtenue par $X_t = b_0(\mathbf{s}_t)\,\boldsymbol{\varepsilon}_t$ c'est à dire $a_i(.) = c_{ij}(.) = 0$ pour tout i, j et $b_j(.) = 0$ pour $1 \leq j \leq q$ dans (2.1) avec $b_0(\mathbf{s}_t) > 0$ (par exemple, Fancq et Roussignol (c.f.,[15])).

(iii) Les modèles $MS - ARMA$. Ces modèles sont obtenus en posant $c_{ij}(.) = 0$ pour tout i et j dans (2.1) (Par exemple, Francq et Zakoïan (c.f.,[19]), Stelzer (c.f.,[67]), Lee (c.f.,[39]) et Yao et Attali (c.f.,[73])).

(iv) Certaines classe de modèles bilinéaires et **GARCH** périodiques Ces modèles sont obtenus par la mise en $\mathbf{s}_t := \sum_{k=1}^{d} k\mathbb{I}_{\triangle(k)}(t)$ où $\Delta(k)$ représentant l'ensemble des indices correspondant à *kième* régime de l'instant t et $\mathbb{I}_{\triangle}$ désigne la fonction indicatrice de l'ensemble $\triangle$ (par exemple, Bibi et al (c.f.,[7]) et (c.f.,[6]) et les références citées).

(v) Certaines classe de $\mathbf{MS}-\mathbf{GARCH}\,(p,q)$: (par exemple, Abramson et Cohen (c.f.,[1]), Francq et Zakoïan (c.f.,[20]) et Liu (c.f.,[51])),(voir aussi Kristensen (c.f.,[38]) pour la construction de $(\mathbf{GARCH}\,(p,q))$ comme un cas particulier des modèles $\mathbf{BL}\,(p,q,P,Q)$).

Pour la commodité et l'élégance des résultats, nous réécrivons le modèle (2.1) sous la forme vectorielle. Sans perdre de généralité, on suppose que $P = p$

$$\forall t \in \mathbb{Z},\ X_t = \sum_{i=1}^{p} a_i(\mathbf{s}_t)\,X_{t-i} + \sum_{j=0}^{q} b_j(\mathbf{s}_t)\,\varepsilon_{t-j} + \sum_{i=1}^{p}\sum_{j=1}^{Q} c_{ij}(\mathbf{s}_t)\,X_{t-i}\varepsilon_{t-j} \tag{2.2}$$

et on considère les vecteurs $\underline{X}_t := (X_t, X_{t-1}, ..., X_{t-p+1})'$ et $\underline{H} := (1, 0, ..., 0)'$, $\underline{\varepsilon}_t := \left(\sum_{j=0}^{q} b_j(\mathbf{s}_t)\,\varepsilon_{t-j}\right)\underline{H}$,

[1]Un modèle de Markov caché est un modèle statistique dans lequel le système modélisé est supposé être un processus markovien de paramètres inconnus.

Les modèles de Markov cachés sont massivement utilisés notamment en reconnaissance de formes, en intelligence artificielle ou encore en traitement automatique du langage naturel, les (HMM) ont commencé à être appliqués à l'analyse de séquences biologique, en particulier l'ADN.

Les (HMM) ont été décrits pour la première fois dans une série de publication de statistiques par Leonard E. Baum et d'autres auteurs après 1965.

et les matrices

$$A_0(\mathbf{s}_t) := \begin{pmatrix} a_1(\mathbf{s}_t) & \dots & \dots & \dots & a_p(\mathbf{s}_t) \\ 1 & 0 & \dots & \dots & 0 \\ 0 & 1 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & 1 & 0 \end{pmatrix}_{p\times p} \quad \text{et } A_j(\mathbf{s}_t) := \begin{pmatrix} c_{1j}(\mathbf{s}_t) & \dots & \dots & \dots & c_{pj}(\mathbf{s}_t) \\ 0 & 0 & \dots & \dots & 0 \\ \vdots & \ddots & \ddots & 0 & \vdots \\ \vdots & 0 & \ddots & \ddots & \vdots \\ 0 & \dots & \dots & 0 & 0 \end{pmatrix}_{p\times p}$$

donc

$$\begin{cases} \underline{X}_t = A_t \underline{X}_{t-1} + \underline{\varepsilon}_t, t \in \mathbb{Z} \\ X_t = \underline{H}' \underline{X}_t \end{cases} \tag{2.3}$$

où $A_t = A_0(\mathbf{s}_t) + \sum_{j=1}^{Q} A_j(\mathbf{s}_t)\varepsilon_{t-j}$. Nous appelons le modèle (2.3) le modèle vectoriel. Notons que dans le modèle super-diagonale ($\mathbf{MS} - \mathbf{SBL}(p, q, p, Q)$) pour lequel $c_{ij}(.) = 0$ dans (2.1) pour $i < j$, la représentation suivante est considérée

$$\underline{X}_t = A_0(\mathbf{s}_t)\underline{X}_{t-1} + \sum_{j=1}^{Q} A_j(\mathbf{s}_t)\underline{X}_{t-j}\varepsilon_{t-j} + \underline{\varepsilon}_t \tag{2.4}$$

où les première lignes des $A_j(\mathbf{s}_t)$ sont $\left(\underbrace{c_{jj}(\mathbf{s}_t), \dots, c_{pj}(\mathbf{s}_t)}_{p-j+1}, \underbrace{0, \dots, 0}_{j-1}\right)$, $1 \leq j \leq Q$.

La représentation (2.3) n'est pas markovienne ni au sens fort ni faible il s'agit uniquement d'une représentation d'espace d'état. Un représentation markovienne peut être envisagée comme suit : On pose $r = p + q$, et soient les $r-$vecteurs $\underline{Y}_t := (X_t, X_{t-1}, \dots, X_{t-p+1}, \varepsilon_t, \varepsilon_{t-1}, \dots, \varepsilon_{t-q+1})'_{r\times 1}$,

$\underline{\eta}_t := (b_0(\mathbf{s}_t)\varepsilon_t, 0 \dots, 0, \varepsilon_t, 0, \dots, 0)'_{r\times 1}$, $\underline{b}_0(\mathbf{s}_t) := (b_0(\mathbf{s}_t), 0 \dots, 0, 1, 0, \dots, 0)'_{r\times 1}$et $\underline{\eta}_t := \varepsilon_t \underline{b}_0(\mathbf{s}_t)$. Alors on a la représentation suivante $\underline{Y}_t = g_t(\underline{Y}_{t-1}) + \underline{\eta}_t$ où, pour une $t \in \mathbb{Z}$, la fonction $g_t : \mathbb{R}^r \to \mathbb{R}^r$ est définie par $g_t(\underline{Y}_{t-1}) = B_t \underline{Y}_{t-1}$, où B_t est une $r \times r$-matrice aléatoire définie par

$$B_t = \begin{pmatrix} \mathcal{R}_t & \mathcal{T}_t \\ O_{(q,p)} & \mathcal{J} \end{pmatrix}_{r\times r}$$

avec

$$\mathcal{R}_t = \begin{pmatrix} a_1(\mathbf{s}_t) + \sum_{j=1}^{Q} c_{1j}(\mathbf{s}_t)\varepsilon_{t-j} & a_2(\mathbf{s}_t) + \sum_{j=1}^{Q} c_{2j}(\mathbf{s}_t)\varepsilon_{t-j} & \ldots & \ldots & a_p(\mathbf{s}_t) + \sum_{j=1}^{Q} c_{pj}(\mathbf{s}_t)\varepsilon_{t-j} \\ 1 & 0 & \ldots & \ldots & 0 \\ 0 & 1 & 0 & \ldots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \ldots & \ldots & 1 & 0 \end{pmatrix}_{p\times p}$$

$$\mathcal{T}_t = \begin{pmatrix} b_1(\mathbf{s}_t) & \ldots & \ldots & \ldots & b_q(\mathbf{s}_t) \\ 0 & \ldots & \ldots & \ldots & 0 \\ 0 & \ddots & \ddots & \ddots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \ldots & \ldots & \ldots & 0 \end{pmatrix}_{p\times q}, \mathcal{J} = \begin{pmatrix} 0 & 0 & \ldots & \ldots & 0 \\ 1 & 0 & \ldots & \ldots & 0 \\ 0 & 1 & 0 & \ldots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & 0 & \ldots & 1 & 0 \end{pmatrix}_{q\times q}$$

on obtient

$$\begin{cases} \underline{Y}_t = B_t \underline{Y}_{t-1} + \underline{\eta}_t \\ X_t = \underline{H}' \underline{Y}_t \end{cases} \tag{2.5}$$

où $\underline{H}' = (1, 0, \ldots, 0)_{1\times r}$. Généralement la représentation (2.5) n'est pas markovienne mais le processus augmenté $\underline{Z}_t = (\underline{Y}_t', s_t)'$ est un processus markovien.

Afin d'étudier les propriétés probabilistes du modèle (2.2), nous faisons les hypothèses suivantes.

H1. $E\left(\log^+ |\boldsymbol{\varepsilon}_t|\right) < +\infty$ où $\log^+ x = \max\{\log x, 0\}, x > 0$,

H2. nous supposons que $\boldsymbol{\varepsilon}_t$ et $(X_{s-1}, \mathbf{s}_t)_{s\leq t}$ sont indépendants.

H3. La chaîne $(\mathbf{s}_t)_{t\in\mathbb{Z}}$ est supposée stationnaire, irréductible et apériodiques de matrice de transition $\mathbb{P} := (p_{ij}, (i,j) \in \mathbb{S}\times\mathbb{S})$ où $p_{ij} := p_{ij} = P(s_t = j | s_{t-1} = i)$ pour $i, j \in \mathbb{S}$, de distribution stationnaire initiale $\Pi = (\pi(1), ..., \pi(d))'$ avec $\pi(i) = P(s_t = i)$, $i = 1, ..., d$ tq : $\Pi' = \Pi'\mathbb{P}$.

Pour l'analyse statistique, il est souvent demandé que le processus solutions $(X_t)_{t\in\mathbb{Z}}$ pour (2.1) devraient être stationnaires, ergodiques et satisfait $X_t = f(\boldsymbol{\varepsilon}_t, \mathbf{s}_t, \boldsymbol{\varepsilon}_{t-1}, \mathbf{s}_{t-1}, \ldots)$ presque sûrement (p.s) où f est une fonction réelle mesurable de $\mathbb{R}^\infty$, ces solutions sont appelées "causale" et inversible si $\boldsymbol{\varepsilon}_t = f(X_t, s_t, X_{t-1}, s_{t-1}, \ldots)$.

Ces propriétés mentionnées ont été étudiées récemment pour la (**MS − ARMA**) par francq et Zakoïan (c.f.,[19]), Stelzer (c.f.,[67]), Lee (c.f.,[39]) et par Yao et Attali (c.f.,[73]), et pour la (**MS − GARCH**) par Francq et Zakoïan (c.f.,[20]), Liu (c.f.,[51]) et Abramson et Cohen (c.f.,[1]). Toutefois, l'absence de linéarité et de la positivité des coefficients dans notre cas rend les détails plus pénibles.

Dans la suite nous utilisons les notations suivantes : Soit $\|.\|$ désigner une norme sur l'ensemble des matrices $m\times n$ et $m\times 1$, la norme d'une matrice $M := (m_{ij})$ est définie par $\|M\| := \max\limits_i \left\{\sum\limits_{j=1}^{p} |m_{ij}|\right\}$ et pour $\gamma \in\,]0;1]$ on définit une opération matricielle $|M|^\gamma$ comme $|M|^\gamma := (|m_{ij}|^\gamma)$. Il est évident que $|M\underline{x}|^\gamma \leq |M|^\gamma |\underline{x}|^\gamma$ pour tout vecteur $\underline{x}$ approprié, et $\left|\sum\limits_i M_i\right|^\gamma \leq \sum\limits_i |M_i|^\gamma$ et $\left|\prod\limits_i M_i\right|^\gamma \leq \prod\limits_i |M_i|^\gamma$ où la relation matricielle $M \leq N$ désigne la relation élément par élément $m_{ij} \leq n_{ij}$ pour tous les i et j et $N := (n_{ij})$. $\mathbb{L}_r := L_r(\Omega, \Im, P)$ désigne l'espace des vecteurs aléatoires $\underline{X}$ de telle sorte que $E\{\|\underline{X}\|^r\} < +\infty$ muni de la norme habituelle $\|.\|_r$ définie par $\|\underline{X}\|_r = (E\|\underline{X}\|^r)^{\frac{1}{r}}$, $r \geq 1$. $\otimes$ désigne le produit de Kronecker. Si $(M(i))_{i\in I}$ est une séquence de matrices $n \times n$, on notera pour tout entier l et j, $\prod\limits_{i=l}^{j} M(i) = M(l)\, M(l+1) \ldots M(j)$ si $l \leq j$ lorsque les matrices $M(i)$, $i \in I := \{1, \ldots, d\}$ sont des matrices non aléatoires, on notera

$$\mathbb{P}(M) = \begin{pmatrix} p_{11}M(1) & \ldots & p_{1d}M(1) \\ \vdots & \ldots & \vdots \\ p_{d1}M(d) & \ldots & p_{dd}M(d) \end{pmatrix} \quad \Pi(M) = \begin{pmatrix} \pi(1)\, M(1) \\ \vdots \\ \pi(d)\, M(d) \end{pmatrix}$$

le rayon spectral d'une matrice carrée M est à noter $\rho(M)$.

2.3 Stationnarité de $MS - BL$

D'après (2.3) on a par récurrence

$$\underline{X}_t = \left(\prod_{j=0}^{t-1} A_{t-j}\right) \underline{X}_0 + \sum_{k=1}^{t-1} \left(\prod_{j=0}^{k-1} A_{t-j}\right) \underline{\varepsilon}_{t-k} + \underline{\varepsilon}_t.$$

Rappelons tout d'abord quelques définitions.

Définition 2.3.1 *Soit* $(\underline{X}_t(\underline{X}_0))_{t\geq 0}$ *le processus défini récursivement par* (2.3) *en initialisant par un vecteur aléatoire* $\underline{X}_0 \in \mathbb{R}^p$ *borné (en probabilité) indépendant de* $(\mathbf{s}_t, \varepsilon_t)_{t\geq 1}$, *et défini par* $\underline{X}_t(\underline{X}_0) =$

$\sum_{k=0}^{t-1} X_t(k)\,\underline{\varepsilon}_{t-k} + X_t(t)\,\underline{X}_0$ *où* $X_t(k) := \left\{\prod_{j=0}^{k-1} A_{t-j}\right\}$ *avec la convention* $X_t(0) = I_{(p)}$. *Alors, le processus d'état* (2.3) *est dit asymptotiquement stable (resp. en moyenne) si* (2.3) *admet une sotution strictement (resp. du second ordre) stationnaire* $(X_t)_{t\in\mathbb{Z}}$ *tels que* $\|\underline{X}_t - \underline{X}_t(\underline{X}_0)\| \longrightarrow 0$ *en probabilité (resp.*$\|\underline{X}_t - \underline{X}_t(\underline{X}_0)\|_2 \longrightarrow 0$ *) quand* $t \longrightarrow \infty$,*il est dit géométriquement stable (resp. en moyenne) si presque sûrement* $\|\underline{X}_t - \underline{X}_t(\underline{X}_0)\| = O(\rho^t)$ *(resp.*$\|\underline{X}_t - \underline{X}_t(\underline{X}_0)\|_2 = O(\rho^t)$*) où* $\rho \in]0;1[$.

2.3.1 Stationnarité stricte

Comme $(\mathbf{s}_t, \varepsilon_t)_{t\in\mathbb{Z}}$ est un processus stationnaire et ergodique[2], $(\mathbf{A}_t, \underline{\varepsilon}_t)_{t\in\mathbb{Z}}$est aussi un processus stationnaire et ergodique et les deux quantités $E_\pi\left\{\log^+\|\underline{\varepsilon}_1\|\right\}$ et $E_\pi\left\{\log^+\|A_1\|\right\}$ sont finis, donc d'après Brandt (c.f.,[10]) (voir Bougerol et Piccard (c.f.,[9])), l'unique solution causale, borné en probabilité, strictement stationnaire[3] et ergodique de (2.3) est donnée par

$$\underline{X}_t = \sum_{k=1}^{\infty} X_t(k)\,\underline{\varepsilon}_{t-k} + \underline{\varepsilon}_t \tag{2.6}$$

dès que l'exposant de Lyapunov $\gamma_L(A)$ associé à la suite des matrices aléatoires $A = (A_t)_{t\in\mathbb{Z}}$ défini par $\gamma_L(A) := \inf_{t>0} E\left\{\frac{1}{t}\log\|X_t(t)\|\right\} \overset{p.s}{=} \lim_{t\longrightarrow\infty}\left\{\frac{1}{t}\log\|X_t(t)\|\right\}$ est strictement négatif. Bien sûr, $\underline{H}'\underline{X}_t$ constituent l'unique solution de (2.1) ayant les mêmes propriétés que $\underline{X}_t$.

La nécessité de la condition $\gamma_L(A) < 0$ pour l'existence de solution strictement stationnaire de (2.3) peut être établis sous condition de contrôlabilité. Notant que, comme l'exposant de lyapunov est indépendant de la partie moyenne mobile, nous allons supposer, sans perte de généralité que $q = 0$ dans (2.1) (sinon,

[2] Un processus $(X_t)_{t\in\mathbb{Z}}$ tel que $E\{|X_1|\} < \infty$ est dit ergodique p.s. (resp. en m.q.) si

$$\frac{1}{n}(X_1 + \ldots + X_n) \text{ converge p.s. (resp. en m.q.) vers } E\{X_1\} \text{ quand } n \text{ tend vers } +\infty.$$

[3] Un processus $(X_t)_{t\in\mathbb{Z}}$ est dite strictement stationnaire si les distributions de $(X_{t_1}, \ldots, X_{t_k})$ et de $(X_{t_1+h}, \ldots, X_{t_k+h})$ sont les mêmes pour tout entier k et pour tout $t_1, t_2, \ldots, t_k, h \in \mathbb{Z}$.

La stationnarité au sens strict signifie intuitivement que les graphes sur deux périodes de temps de même longueur finie, d'une réalisation d'un processus stochastique, présentent les mêmes caractéristiques statistiques.

une représentation appropriée est nécessaire, voir Bibi et Aknouche (c.f.,[5])) et dans ce cas

$$\underline{X}_t := A_t \underline{X}_{t-1} + \varepsilon_t \underline{H}\left(\mathbf{s}_t\right) \tag{2.7}$$

avec $\underline{H}\left(\mathbf{s}_t\right) := b_0\left(\mathbf{s}_t\right)\underline{H}$. Rappelons ici, que la représentation (2.1) est dite contrôlable (resp. observable) si les matrices $\mathcal{C}_p\left(\mathbf{s}_t\right) := \left[\underline{c}_1\left(\mathbf{s}_t\right) \vdots \underline{c}_2\left(\mathbf{s}_t\right) \vdots \ldots \vdots \underline{c}_p\left(\mathbf{s}_t\right)\right]$ (resp.$\mathcal{O}_p\left(\mathbf{s}_t\right) := \left[\underline{O}_{(1)}\left(\mathbf{s}_t\right) \vdots \underline{O}_{(2)}\left(\mathbf{s}_t\right) \vdots \ldots \vdots \underline{O}_{(p)}\left(\mathbf{s}_t\right)\right]$) ont des rangs égaux à p (p.s) où les matrices $\left(\underline{c}_j\left(\mathbf{s}_t\right)\right)_{1\leq j\leq p}$ (resp. $\left(\underline{O}_j\left(\mathbf{s}_t\right)\right)_{1\leq j\leq p}$) sont définies récursivement par $\underline{c}_1\left(\mathbf{s}_t\right) = \underline{H}\left(\mathbf{s}_t\right)$ et $\underline{c}_j\left(\mathbf{s}_t\right) = \left[A_i\left(\mathbf{s}_t\right)\underline{c}_{j-1}\left(\mathbf{s}_{t-j+1}\right); i \in \{0,\ldots,Q\}\right]$ pour $j \geq 2$. (resp. $\underline{O}_{(1)}\left(\mathbf{s}_t\right) = \underline{H}$ et $\underline{O}_{(j)}\left(\mathbf{s}_t\right) = \left[A_i\left(\mathbf{s}_t\right)\underline{O}_{(j-1)}\left(\mathbf{s}_{t-j+1}\right); i \in \{0,\ldots,Q\}\right]$ pour $j \geq 2$).

Les résultats sur la stationnarité stricte sont résumés dans le théorème suivant

Théorème 2.3.1 *On considère le modèle* (2.1) *et sa représentation d'état* (2.3). *Alors*

$\gamma_L(A) < 0$ *est une condition suffisante pour que* (2.3) *admet une solution unique, strictement stationnaire, causale, ergodique et géométrique stable donnée par la série* (2.6) *qui convergent presque sûrement pour tout* $t \in \mathbb{Z}$.

Si le modèle (2.7) *est contrôlable et admet une solution asymptotiquement stable alors* $\gamma_L(A) < 0$.

Preuve Les premières assertions sont des conséquences immédiates du théorème (1) de Brandt (c.f.,[10])

et théorème (1.1) de Bougerol et Picard (c.f.,[9]).

D'autre part,

$$\begin{aligned}
\lim\ \sup_k \left\|X_t(k)\,\underline{\varepsilon}_{t-k}\right\|^{\frac{1}{k}} &= \exp \limsup_k \frac{1}{k}\log\left\|X_t(k)\,\underline{\varepsilon}_{t-k}\right\| \\
&\leq \exp \limsup_k \frac{1}{k}\log\left\{\left\|X_t(k)\right\|\left\|\underline{\varepsilon}_{t-k}\right\|\right\} \\
&\leq \exp \limsup_k \frac{1}{k}\log\left\|X_t(k)\right\| \\
&\leq \exp\left\{\gamma_L(A)\right\} < 1
\end{aligned}$$

alors

$$\begin{aligned}
\left\|\underline{X}_t - \underline{X}_t\left(\underline{X}_0\right)\right\| &= \left\|\sum_{k=0}^{\infty} X_t\left(k\right)\underline{\varepsilon}_{t-k} - \sum_{k=0}^{t-1} X_t\left(k\right)\underline{\varepsilon}_{t-k} - X_t\left(t\right)\underline{X}_0\right\| \\
&= \left\|\sum_{k=t}^{\infty} X_t\left(k\right)\underline{\varepsilon}_{t-k} - X_t\left(t\right)\underline{X}_0\right\| \\
&\leq K\exp\left\{t\gamma_L\left(A\right)\right\}
\end{aligned}$$

où K est une constante positive. D'où la stabilité géométrique. Pour prouver la seconde assertion, nous observons que s'il existe une solution asymptotiquement stable pour (2.6) alors presque sûrement $\lim\limits_{t\longrightarrow\infty}\left\|X_t\left(t\right)\underline{H}\right\| = 0$, par contrôlabilité on obtient presque sûrement $\lim\limits_{t\longrightarrow\infty}\left\|X_t\left(t\right)\right\| = 0$. Simple modification du lemme (3.4) dans Picard (c.f.,[9]) montre que $\gamma_L\left(A\right) < 0$.

Remarque 2.3.1 *Il n'est pas difficile de voir que la condition de la contrôlabilité de la représentation* (2.7) *implique son irréductibilité dans le sens où il n'existe pas un sous-espace affine $\mathcal{H}$ de $\mathbb{R}^p$ tels que $\left\{A_t\underline{X} + \underline{H}\left(\mathbf{s}_t\right)\varepsilon_t : \underline{X}\in\mathcal{H}\right\}\subset\mathcal{H}$ p.s (c.f.,Kristense (c.f.,[38]) pour plus de détails et d'exemples).*

Exemple 2.3.1 [*non-nécessité de stationnarité locale*] *considérons le $\left(MS - BL\left(1,0,1,1\right)\right)$ modèle $X_t = \left(a_1\left(\mathbf{s}_t\right) + c_{11}\left(\mathbf{s}_t\right)\varepsilon_{t-1}\right)X_{t-1} + b_0\left(\mathbf{s}_t\right)\varepsilon_t$ où $\left(\varepsilon_t\right)_{t\in\mathbb{Z}}$ est un i.i.d séquence. La condition de stationnarité est*

$$\gamma_L\left(A\right) = \inf_{t>0} E\left\{\frac{1}{t}\log\left\|X_t\left(t\right)\right\|\right\} = \sum_{i=1}^{d}\pi\left(i\right)E\left\{\log\left|a_1\left(i\right) + c_{11}\left(i\right)\varepsilon_0\right|\right\} < 0$$

si $b_0\left(i\right)\neq 0$ pour tous les $i\in\mathbb{S}$, alors $\gamma_L\left(A\right) < 0$ est également nécessaire car $rg\left\{\underline{c}_1\left(\mathbf{s}_t\right)\right\} = 1$, autrement, $X_t = 0$ est l'unique solution stationnaire sans aucune contrainte sur $\gamma_L\left(A\right)$.

Cet exemple montre l'importance de la présence de la partie moyenne mobile pour la condition nécessaire. Il montre également que la stationnarité locale n'est pas nécessaire, à savoir l'existence de régimes explosifs c'est à dire $E\left\{\log\left|a_1\left(i\right) + c_{11}\left(i\right)\varepsilon_t\right|\right\} > 0$ ne s'oppose pas à la stationnarité globale.

Exemple 2.3.2 [*non-suffisance de stationnarité locale*] *Considérons le modèle*

$$X_t = \begin{cases} X_{t-1} + c_{11}\left(1\right)X_{t-1}\varepsilon_t + \varepsilon_t & si \;\; \mathbf{s}_t = 1 \\ X_{t-1} + c_{11}\left(2\right)X_{t-1}\varepsilon_t + \varepsilon_t & si \;\; \mathbf{s}_t = 2 \end{cases}$$

où $(\varepsilon_t)_{t\in\mathbb{Z}}$ *est un processes i.i.d. Ensuite, le processus est non stationnaire chaque fois*

$$\gamma_L(A) = \{\pi(1) E\{\log|1 + c_{11}(1)\varepsilon_t|\} + \pi(2) E\{\log|1 + c_{11}(2)\varepsilon_t|\}\} > 0.$$

Mais cette condition est compatible avec la stationnarité locale. En effet, chaque régime admet toujours une solution stationnaire unique donnés par $X_t = -\frac{1}{c_{11}(i)}\mathbb{I}_{\{\mathbf{s}_t=i\}}$ *sans aucune contrainte sur* $\gamma_L^{(i)}(A) = \pi(i) E\{\log|1 + c_{11}(i)\varepsilon_t|\}, i = 1, 2$ *qu'il soit négatif ou non.*

Bien que la condition $\gamma_L(A) < 0$ peut être utilisé comme une condition suffisante pour la stationnarité,

il est peu d'utilité en pratique car cette condition est basée sur le calcul d'un produit infini de matrices aléatoires. Par conséquent, certaines conditions suffisantes simples assurant la négativité de $\gamma_L(A)$ peuvent être envisagées.

Proposition 2.3.1 *Considérons la représentation* (2.1) *avec* $Q = 1$ *et supposons que* $(\varepsilon_t)_{t\in\mathbb{Z}}$ *est une suite i.i.d avec* $E|\varepsilon_0|^\delta < \infty$ *pour certain* $0 < \delta \leq 1$ *et soit* $\Gamma(i) := E\left\{|A_0(\mathbf{s}_t) + A_1(\mathbf{s}_t)\varepsilon_t|^\delta \mathbb{I}_{\{\mathbf{s}_t=i\}}\right\}$. *Alors* $\rho(\mathbb{P}(\Gamma)) < 1$ *où* $\Gamma := (\Gamma(i), 1 \leq i \leq d)$, *implique que* $\gamma_L(A) < 0$ *et donc les premières assertions du théorème* (2.3.1) *sont vraies.*

Preuve Comme l'exposant de lyapunov est indépendant de la norme choisie, alors on peut choisir une norme $\|.\|$ de telle sorte que $\|M\|^\delta \leq \left\||M|^\delta\right\|$ (par exemple,$\|M\| = \sum\limits_{i,j}|m_{ij}|$) par conséquent, par hypothèse $\rho(\mathbb{P}(\Gamma)) < 1$, il existe $0 < \lambda < 1$ de telle sorte que $\lim\sup\limits_n \|\mathbb{P}^n(\Gamma)\|^{\frac{1}{n}} < \lambda$. Posons $\mathbb{I}' := \left(I_{(p)} \vdots \ldots \vdots I_{(p)}\right)$ et $\Pi'(I) := \left(\pi(1) I_{(p)} \vdots \ldots \vdots \pi(d) I_{(p)}\right)$. L'inégalité de Jensen et la sous multiplicativité de l'opérateur $|.|^\delta$ nous avons

$$\begin{aligned}
\delta\gamma_L(A) &= \lim_{t\longrightarrow\infty}\frac{1}{t}E\left\{\log\left\|\prod_{j=0}^{t-1}\left(A_0(\mathbf{s}_{t-j})+A_1(\mathbf{s}_{t-j})\varepsilon_{t-j-1}\right)\right\|^{\delta}\right\}\\
&\leq \lim_{t\longrightarrow\infty}\frac{1}{t}\log E\left\{\left\|\prod_{j=0}^{t-1}\left(A_0(\mathbf{s}_{t-j})+A_1(\mathbf{s}_{t-j})\varepsilon_{t-j-1}\right)\right\|^{\delta}\right\}\\
&\leq \lim_{t\longrightarrow\infty}\frac{1}{t}\log E\left\{\left\|\prod_{j=0}^{t-1}\left|A_0(\mathbf{s}_{t-j})+A_1(\mathbf{s}_{t-j})\varepsilon_{t-j-1}\right|^{\delta}\right\|\right\}\\
&\leq \lim_{t\longrightarrow\infty}\frac{1}{t}\log\left\|E\left\{\prod_{j=0}^{t-1}\Gamma(\mathbf{s}_{t-j})\right\}\right\|\\
&\leq \lim\sup_{t\longrightarrow\infty}\log\left\|\mathbb{I}'\mathbb{P}^{t-1}(\Gamma)\Pi(\Gamma)\right\|^{\frac{1}{t}}<0
\end{aligned}$$

Autres conditions suffisantes plus faibles basés sur les produits fini de matrices aléatoires peut être utilisées. En effet, puisque $\gamma_L(A)\leq\gamma(A)=\log E\{\|X_t(p)\|\}$ alors nous pouvons explorer la condition $E\left\{\|X_t(p)\|^{\delta}\right\}<1$ pour certains $\delta>0$ pour donner le résultat suivant.

Proposition 2.3.2 *Supposons que* $(\varepsilon_t)_{t\in\mathbb{Z}}$ *est une suite i.i.d, avec* $E|\varepsilon_0|^{\delta}<\infty$ *pour certains* $0<\delta\leq 1$. *Si* $\sum_{i=1}^{p}E\left\{\left|a_i(\mathbf{s}_t)+\sum_{j=1}^{Q}c_{ij}(\mathbf{s}_t)\varepsilon_{t-j-1}\right|^{\delta}\right\}<\frac{1}{p}$ *alors* $\gamma(A)<0$ *et donc les résultats de la première assertion du théorème* (2.3.1) *sont vraies.*

Preuve La preuve découle essentiellement des mêmes arguments que dans Liu (c.f.,[47]). Notant que dans le cas où $(\varepsilon_t)_{t\in\mathbb{Z}}$ a une variance finie σ^2 et de moyenne nulle, une simple condition suffisante est $E\left\{\left(\sum_{i=1}^{p}\left|a_i(\mathbf{s}_t)+\sum_{j=1}^{Q}c_{ij}(\mathbf{s}_t)\varepsilon_{t-j-1}\right|\right)^2\right\}<\frac{1}{p}$ tant qu'une condition est légèrement plus forte

$$\sum_{i=1}^{p}\sum_{k=1}^{d}\pi(k)\left\{a_i^2(k)+\sum_{j=1}^{Q}c_{ij}^2(k)\ \sigma^2\right\}<\frac{1}{p^2}.$$

Le critère de Lyapunov $\gamma_L(A) < 0$ semble difficile d'obtenir explicitement lorsque $p > 1$. Cependant, on peut vérifier la négativité de $\gamma_L(A)$ simulation en utilisant l'équation (2.1). Ce fait limite considérablement l'intérêt de ce critère dans les applications statistiques. Dans la section suivante, nous donnons des conditions assurant l'existence des moments pour la solution strictement stationnaire.

2.3.2 Stationnarité au second ordre

Les conditions assurants l'existence de solutions stationnaires aux second ordres [4] pour les modèles $\mathbf{MS-ARMA}(p,q)$ ont été connées par Francq et Zakoïan (c.f.,[19]) (voir aussi les références citées). Résultats généraux pour les modèles ($\mathbf{MS-BL}$) n'ont pas été obtenus, à notre connaissance. Donc, dans ce paragraphe, nous nous intéressons à des conditions garantissant l'existence de solutions stationnaires aux second ordre qui sont aussi causales, strictement stationnaires et ergodiques. Nous donnons les résultats sous forme explicite pour le cas $Q = 2$, mais les mêmes arguments peuvent être utilisés pour $Q > 2$.

$MS - BL(p,q,p,2)$

Le théorème suivant examine les conditions garantissant l'existence des solutions stationnaires au second ordre pour l'équation (2.3) lorsque $Q = 2$.

Théorème 2.3.2 *Considérons l'équation* (2.3) *avec* $Q = 2$, *et supposons que* $(\varepsilon_t)_{t\in\mathbb{Z}}$ *est une suite i.i.d*

[4]Un processus $(X_t)_{t\in\mathbb{Z}}$ est dit stationnaire du second ordre si

$$\begin{aligned} i)\ E\left\{|X_t|^2\right\} &< \infty, t\in\mathbb{Z} \\ ii)\ E(X_t) &= m,\ \forall t\in\mathbb{Z} \text{ (la moyenne ne dépend pas du temps } t) \\ iii)\ \varphi_X(r,s) &= \varphi_X(r+t,s+t), \forall t,r,s\in\mathbb{Z}\ (\varphi_X \text{ est invariante par translation dans le temps).} \end{aligned}$$

Il serait alors convenable de redéfinir une fonction d'auto-covariance d'un processus stationnaire du second ordre comme une fonction d'une seule variable

$$\varphi_X(h) = \varphi_X(h,0) = cov(X_{t+h}, X_t), \forall h,t\in\mathbb{Z}.$$

avec

$$E\left\{\varepsilon_t\right\} = E\left\{\varepsilon_t^3\right\} = 0, E\left\{\varepsilon_t^2\right\} = \sigma^2 \text{ et } K_4 := E\left\{\varepsilon_t^4\right\} < +\infty. \tag{2.8}$$

Posons $\Gamma = (\Gamma(i), 1 \leq i \leq d)$ *où*

$$\Gamma(i) := \begin{pmatrix} A_0^{\otimes 2}(i) + \sigma^2 A_1^{\otimes 2}(i) & A_0(i) \otimes A_2(i) + A_2(i) \otimes A_0(i) & A_2^{\otimes 2}(i) \\ \sigma^2\left(A_0(i) \otimes A_1(i) + A_1(i) \otimes A_0(i)\right) & \sigma^2\left(A_1(i) \otimes A_2(i) + A_2(i) \otimes A_1(i)\right) & O_{(p^2)} \\ \sigma^2 A_0^{\otimes 2}(i) + K_4 A_1^{\otimes 2}(i) & \sigma^2\left(A_0(i) \otimes A_2(i) + A_2(i) \otimes A_0(i)\right) & \sigma^2 A_2^{\otimes 2}(i) \end{pmatrix}$$

alors si

$$\lambda := \rho\left(\mathbb{P}(\Gamma)\right) < 1 \tag{2.9}$$

l'équation (2.3) *a une unique solution stationnaire au second ordre, géométriquement stable et en moyenne quadratique donnée par la série* (2.6) *qui converge presque sûrement et en moyenne quadratique.*

De plus cette solution est strictement stationnaire et ergodique.

Corollaire 2.3.1 *Pour le modèle* $(MS - BL(1, q, 1, 1))$. *La condition* (2.9) *se réduit à* $\rho\left(\mathbb{P}\left(a_1^2 + \sigma^2 c_{11}^2\right)\right) < 1$. *En particulier, lorsque* $d = 2$, *avec* $p_{11} = p_{22} = 1 - p$, $p_{12} = p_{21} = p$, *alors la condition est équivalente aux deux conditions suivantes*

$$\begin{cases} (2p-1)\left(a_1^2(1) + c_{11}^2(1)\right)\left(a_1^2(2) + c_{11}^2(2)\right) + (1-p)\left(a_1^2(1) + c_{11}^2(1) + a_1^2(2) + c_{11}^2(2)\right) < 1 \\ (1-p)\left(a_1^2(1) + c_{11}^2(1) + a_1^2(2) + c_{11}^2(2)\right) < 2 \end{cases}$$

Preuve Dans ce cas, $A_0(\mathbf{s}_t) = a_1(\mathbf{s}_t)$, $A_1(\mathbf{s}_t) = c_{11}(\mathbf{s}_t)$, $A_2(\mathbf{s}_t) = 0$ et

$$\Gamma(i) := \begin{pmatrix} a_1^2(i) + \sigma^2 c_{11}^2(i) & 0 & 0 \\ 2\sigma^2 c_{11}(i)\, a_1(i) & 0 & 0 \\ \sigma^2 a_1^2(i) + K_4 c_{11}^2(i) & 0 & 0 \end{pmatrix}$$

Un calcul simple montre que les valeurs propres non nulles de $\mathbb{P}(\Gamma)$ sont les mêmes que celle de $\mathbb{P}\left(a_1^2 + \sigma^2 c_{11}^2\right)$ d'où le résultat. A noter que cet exemple montre que la stationnarité locale du second ordre n'est ni suffisante ni nécessaire pour la stationnarité globale au second ordre.

Corollaire 2.3.2 *Lorsque* $Q = 1$*, la condition* (2.9) *se réduit à* $\rho\left(\mathbb{P}\left(A_0^{\otimes 2} + \sigma^2 A_1^{\otimes 2}\right)\right) < 1$.

Preuve Dans ce cas, les matrices $\Gamma(i)$ prendront la forme

$$\Gamma(i) := \begin{pmatrix} A_0^{\otimes 2}(i) + \sigma^2 A_1^{\otimes 2}(i) & O_{(p^2)} & O_{(p^2)} \\ \sigma^2\left(A_0(i) \otimes A_1(i) + A_1(i) \otimes A_0(i)\right) & O_{(p^2)} & O_{(p^2)} \\ \sigma^2 A_0^{\otimes 2}(i) + K_4 A_1^{\otimes 2}(i) & O_{(p^2)} & O_{(p^2)} \end{pmatrix}$$

et par conséquent les valeurs propres non nulles de $\mathbb{P}(\Gamma)$ sont les mêmes que celles de $\mathbb{P}\left(A_0^{\otimes 2} + \sigma^2 A_1^{\otimes 2}\right)$. Ainsi $\rho(\mathbb{P}(\Gamma)) = \rho\left(\mathbb{P}\left(A_0^{\otimes 2} + \sigma^2 A_1^{\otimes 2}\right)\right)$.

Corollaire 2.3.3 *Pour la super-diagonale* (2.4) *la condition* (2.9) *se réduit à* $\rho(\mathbb{P}(L)) < 1$*, où* $L := (L(i), 1 \leq i \leq d)$

$$L(i) = \begin{pmatrix} \Gamma_1(i) & \Gamma_2(i) & \dots & \dots & \Gamma_Q(i) \\ I_{(p^2)} & O_{(p^2)} & \dots & \dots & O_{(p^2)} \\ O_{(p^2)} & I_{(p^2)} & O_{(p^2)} & \dots & O_{(p^2)} \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ O_{(p^2)} & \dots & O_{(p^2)} & I_{(p^2)} & O_{(p^2)} \end{pmatrix}$$

avec $\Gamma_1(i) := A_0^{\otimes 2}(i) + \sigma^2 A_1^{\otimes 2}(i)$ *et*

$$\Gamma_j(i) := \sigma^2\left[A_j(i) \otimes \left(\sum_{l=1}^{j-1} A_0^l(i) A_{j-l}(i)\right) + \left(\sum_{l=1}^{j-1} A_0^l(i) A_{j-l}(i)\right) \otimes A_j(i) + A_j^{\otimes 2}(i)\right],\ j = 2, \cdots, Q.$$

Preuve La preuve découle les mêmes arguments utilisés dans le théorème (2.3.2).

Corollaire 2.3.4 [*Le modèle* $MS - ARMA$] *Dans le cas linéaire, lorsque les coefficients* $c_{ij}(.)$ *dans* (2.1) *sont tous nuls, la condition* (2.9) *se réduit à* $\rho\left(\mathbb{P}\left(A_0^{\otimes 2}\right)\right) < 1$.

Preuve Dans ce cas, les matrices $\Gamma(i)$ prennent la forme

$$\Gamma(i) = \begin{pmatrix} A_0^{\otimes 2}(i) & O_{(p^2)} & O_{(p^2)} \\ O_{(p^2)} & O_{(p^2)} & O_{(p^2)} \\ \sigma^2 A_0^{\otimes 2}(i) & O_{(p^2)} & O_{(p^2)} \end{pmatrix}$$

de sorte que le non nuls valeurs propres de $\mathbb{P}(\Gamma)$ sont les mêmes de $\mathbb{P}\left(A_0^{\otimes 2}\right)$. Ainsi $\rho(\mathbb{P}(\Gamma)) = \rho\left(\mathbb{P}\left(A_0^{\otimes 2}\right)\right)$, la preuve découle des arguments standards.

Définissons maintenant les processus de $\mathbb{R}^p$ suivants

$$\underline{S}_n(t) = \begin{cases} \underline{O}_{(p)} & \text{si } \; n < 0 \\ \underline{\varepsilon}_t & \text{si } \; n = 0 \\ A_t \underline{S}_{n-1}(t-1) + \underline{\varepsilon}_t & \text{si } n > 0 \end{cases} \tag{2.10}$$

et posons $\underline{\triangle}_n(t) = \underline{S}_n(t) - \underline{S}_{n-1}(t)$ donc

$$\underline{\triangle}_n(t) = \begin{cases} \underline{O}_{(p)} & \text{si } n < 0 \\ \underline{\varepsilon}_t & \text{si } n = 0 \\ A_t \underline{\triangle}_{n-1}(t-1) & \text{si } n > 0 \end{cases} \tag{2.11}$$

Preuve les récurrences (2.10) et (2.11) permettent de construire les fonctions mesurables $g_n : \mathbb{R}^\infty \longrightarrow \mathbb{R}^p$, $G_n : \mathbb{R}^\infty \longrightarrow \mathbb{R}^p$ telles que $\underline{\triangle}_n(t) = g_n(\varepsilon_{t-1}, \varepsilon_{t-2}, ...)$, $n \geq 1$ et $\underline{S}_n(t) = G_n(\varepsilon_{t-1}, \varepsilon_{t-2}, ...)$, $n \geq 0$. Il est clair $\underline{S}_n(t)$ et $\underline{\triangle}_n(t)$ sont mesurables par rapport à $\Im_t := \sigma(\{(\varepsilon_s, \mathbf{s}_s), s \leq t\})$ et sont donc strictement stationnaire pour tout $n \in \mathbb{Z}$ fixé. D'après la théorie des espaces $\mathbb{L}_p$ $(p > 1)$ le problème d'existence de solution stationnaire au second ordre de (2.3) se réduit à la convergence de $(\underline{S}_n(t))_{n \geq 1}$ vers $\underline{X}_t$ dans $\mathbb{L}_2$. La quantité clé de la convergence est par conséquent $E\left\{\underline{\triangle}_n(t)\underline{\triangle}'_n(t)\right\} \to 0$, ce qui implique que $\underline{S}_n(t)$ est une suite de Cauchy dans $\mathbb{L}_2$. Pour ce faire, nous devons évaluer les vecteurs des moments $\underline{V}_n(i) = E\left\{\underline{\triangle}_n^{\otimes 2}(t)\middle| \mathbf{s}_t = i\right\}$, $\underline{D}_n(i) = E\left\{\underline{\triangle}_n^{\otimes 2}(t)\varepsilon_{t-1}\middle| \mathbf{s}_t = i\right\}$ et $\underline{F}_n(i) = E\left\{\underline{\triangle}_n^{\otimes 2}(t)\varepsilon_{t-1}^2\middle| \mathbf{s}_t = i\right\}$ qui peuvent être évalués comme suit, pour tout $n \geq 3$ nous avons d'après l'indépendance entre ε_t et $\mathbf{s}_t$ et de condition (2.7).

$$\begin{aligned}
\pi(i)\underline{V}_n(i) &= \sum_{j=1}^{d}\left\{\begin{array}{c}\left(A_0^{\otimes 2}(i)+\sigma^2 A_1^{\otimes 2}(i)\right)\underline{V}_{n-1}(j)+\\ \left(A_0(i)\otimes A_2(i)+A_2(i)\otimes A_0(i)\right)\underline{D}_{n-1}(j)+A_2^{\otimes 2}(i)\underline{F}_{n-1}(j)\end{array}\right\}\pi(j)p_{ji}\\
\pi(i)\underline{D}_n(i) &= \sum_{j=1}^{d}\left\{\begin{array}{c}\sigma^2\left(A_0(i)\otimes A_1(i)+A_1(i)\otimes A_0(i)\right)\underline{V}_{n-1}(j)+\\ \sigma^2\left(A_1(i)\otimes A_2(i)+A_2(i)\otimes A_1(i)\right)\underline{D}_{n-1}(j)\end{array}\right\}\pi(j)p_{ji}\\
\pi(i)\underline{F}_n(i) &= \sum_{j=1}^{d}\left\{\begin{array}{c}\left(\sigma^2 A_0^{\otimes 2}(i)+K_4 A_1^{\otimes 2}(i)\right)\underline{V}_{n-1}(j)+\\ \sigma^2\left(A_0(i)\otimes A_2(i)+A_2(i)\otimes A_0(i)\right)\underline{D}_{n-1}(j)+\sigma^2 A_2^{\otimes 2}(i)\underline{F}_{n-1}(j)\end{array}\right\}\pi(j)p_{ji}
\end{aligned}$$

Endéfinissant $\underline{U}_n = (\underline{U}_n(i), 1\leq i\leq d)$ où $\underline{U}_n(i) := \left(\pi(i)\underline{V}'_n(i), \pi(i)\underline{D}'_n(i), \pi(i)\underline{F}'_n(i)\right)'$ nous obtenons l'équation récursive homogène suivante

$$\Pi(\underline{U}_n) = \mathbb{P}(\Gamma)\,\Pi\left(\underline{U}_{n-1}\right)$$

Ceci montre que la condition (2.9) est suffisante pour $(\Pi(\underline{U}_n))_n$ converger vers zéro géométriquement. Plus précisément, il existe une constante positive K telle que

$$\begin{aligned}
E\left\|\underline{S}_n(t)-\underline{S}_{n-1}(t)\right\|^2 &= E\left\|\underline{\triangle}_n(t)\right\|^2 = E\left\{trace\left(\underline{\triangle}_n(t)\,\underline{\triangle}'_n(t)\right)\right\}\leq\left\|\Pi(\underline{U}_n)\right\|\\
&\leq K\lambda^{\frac{n}{2}}
\end{aligned}$$

Cela implique que pour chaque t fixé, $(\underline{S}_n(t))_n$ est une suite de Cauchy, donc elle converge dans $\mathbb{L}_2$ et p.s que $n\longrightarrow\infty$. D'autre part, depuis $\underline{S}_n(t) = \sum_{k=1}^{n} X_t(k)\,\underline{\varepsilon}_{t-k}+\underline{\varepsilon}_t$ alors $\underline{X}_t = \lim_{n\longrightarrow\infty}\underline{S}_n(t)$ satisfait (2.3) pour $Q=2$. Il convient de noter que $(\underline{S}_n(t))_n$ est une solution causale, strictement stationnaire, géométriquement stable en moyenne et ergodique, et donc sa limite $\underline{X}_t$ a les mêmes propriétés.

Discutons maintenant la nécessité de la condition (2.9). supposons que $(\underline{X}_t)_{t\in\mathbb{Z}}$ admet une solution causale, stationnaire au second ordre solution à avec $Q=2$. Alors pour tout $k\geq 1$ nous avons $\underline{X}_t = \underline{X}_t^{(1)}(k)+\underline{X}_t^{(2)}(k)+X_t(k)\,\underline{X}_{t-k-1}$ où $\underline{X}_t^{(1)}(k) := \sum_{l=0}^{k} X_t(l-1)\,\underline{\eta}_{t-l}^{(1)}$, $\underline{X}_t^{(2)}(k) := \sum_{l=0}^{k} X_t(l-1)\,\underline{\eta}_{t-l}^{(2)}$ avec $\underline{\eta}_t^{(i)} = A_{i-1}(\mathbf{s}_{t+1})\,\underline{H}(\mathbf{s}_t)\,\varepsilon_t^i$, $i=1,2$. Les matrices de covariance de $\underline{\eta}_t^{(i)}$, $i=1,2$ sont désignés $\sum^{(i)}$, $i=1,2$ respectivement. Notons que $\underline{X}_t^{(1)}(k)$ et $\underline{X}_t^{(2)}(k)+X_t(k)\,\underline{X}_{t-k-1}$ sont non corrélés avec $E\left\{\underline{X}_t^{(1)}(k)\right\}=0$ et donc pour tous $k \geq 1$, $E\left\{\underline{X}_t^{\prime(1)}(k)\,\underline{X}_t^{(1)}(k)\right\} \leq E\left\{\underline{X}_t^{\prime}\underline{X}_t\right\}$ par conséquent, nous devons avoir $\sum_{l=0}^{\infty}\left\|X_t(l)\,\underline{\eta}_{t-l}^{(1)}\right\|_2^2 < +\infty$ ce qui implique que $\sum_{l=0}^{\infty}\left\|X_t(l)\,\underline{\eta}_{t-l}^{(1)}\underline{\eta}_{t-l}^{\prime(1)}X_t^{\prime}(l)\right\| = \sum_{l=0}^{\infty}\left\|\mathbb{I}^{\prime}\mathbb{P}^l\left(A_0^{\otimes 2}+\sigma^2 A_1^{\otimes 2}\right)\Pi\left(\overset{(1)}{\Sigma}\right)\right\| < +\infty$ où $\mathbb{I}^{\prime} = \left(\mathbb{I}_{(p^2)}\vdots\ldots\vdots\mathbb{I}_{(p^2)}\right)$ est une matrice $p^2 \times dp^2$.

La discussion précédente conduit au résultat suivant.

Théorème 2.3.3 *Une condition nécessaire pour l'existence d'une solution causale stationnaire au second ordre de* (2.7) *est que* $\sum_{l=0}^{\infty}\left\|\mathbb{I}^{\prime}\mathbb{P}^l(\Gamma)\,\Pi\left(\overset{(1)}{\Sigma}\right)\right\| < +\infty$ *en outre, si* (2.7) *est contrôlable et admet une solution asymptotiquement stable en moyen, alors la condition* (2.9) *est vraie.*

2.3.3 Le modèle $MS - BL$ général

Pour un entier $Q > 2$, le même argument utilisé dans le paragraphe précédent peut encore être appliqué. Dans ce cas, nous supposons que $(\varepsilon_t)_{t\in\mathbb{Z}}$ satisfaisant les conditions $E\left\{\varepsilon_t^{2Q}\right\} < +\infty$ et $E\left\{\varepsilon_t^{r}\right\} = 0$ pour tout entier r impair $r < 2Q$. Ainsi les processus $\underline{S}_n(t)$ et $\underline{\triangle}_n(t)$ restent les mêmes que dans le cas $MS - BL(p,q;p,2)$. Cependant, les propriétés de convergence de la séquence $\{\underline{S}_n(t)\}$ peut être étudiée en utilisant les mêmes arguments que dans le cas $Q = 2$ en introduisant les vecteurs suivants

$$\begin{cases} \underline{V}_n(i) = E\left\{\underline{\triangle}_n^{\otimes 2}(t)\middle|\,\mathbf{s}_t = i\right\} & \\ \underline{D}_n^{(j)}(i) = E\left\{\varepsilon_{t-j}\underline{\triangle}_n^{\otimes 2}(t)\middle|\,\mathbf{s}_t = i\right\}, & 1 \leq j \leq Q-1 \\ \underline{F}_n^{(jk)}(i) = E\left\{\varepsilon_{t-j}\varepsilon_{t-k}\underline{\triangle}_n^{\otimes 2}(t)\middle|\,\mathbf{s}_t = i\right\}, & 1 \leq j \leq k \leq Q-1 \end{cases}$$

et les moments d'ordre supérieurs et nous pouvons définir la récurence $\Pi(\underline{U}_n) = \mathbb{P}(\Gamma)\,\Pi\left(\underline{U}_{n-1}\right)$.

2.3.4 Ergodicité géométrique et la propriété $\beta-$mélange

Nous continuons notre inversement ici par l'examinassions des conditions assurants l'ergodicité géométrique et la $\beta-$mélange. Comme il est difficile à manipuler les termes de produits $X_t\varepsilon_{t-k}$, $k > 0$, dans (2.1), nous nous limiterons aux modèles super-diagonaux (2.4) et sans perte de généralité, nous considérons $MS - SBL(p,0,p,Q)$. Bien que ces modèles peuvent être considérés comme un $MS - ARMA$

faible, alors les développements sont assez fastidieux. Soit $r = p(Q+1)$ et définissons les $r-$vecteurs

$\underline{Y}_t := \left(\underline{X}_t', \underline{X}_t'\varepsilon_t, \ldots, \underline{X}_{t-Q+1}'\varepsilon_{t-Q+1}\right)_{r\times 1}'$ et $\underline{\eta}_{\mathbf{s}_t}(\varepsilon_t) := \left(\underline{H}'(\mathbf{s}_t)\varepsilon_t, \underline{H}'(\mathbf{s}_t)\varepsilon_t^2, \underline{O}_{(p(Q-1))}'\right)_{r\times 1}'$ et posons

$$A_{\mathbf{s}_t}(\varepsilon_t) := \begin{pmatrix} A_0(\mathbf{s}_t) & A_1(\mathbf{s}_t) & A_2(\mathbf{s}_t) & \ldots & \ldots & A_Q(\mathbf{s}_t) \\ A_0(\mathbf{s}_t)\varepsilon_t & A_1(\mathbf{s}_t)\varepsilon_t & A_2(\mathbf{s}_t)\varepsilon_t & \ldots & \ldots & A_Q(\mathbf{s}_t)\varepsilon_t \\ O_{(p)} & I_{(p)} & O_{(p)} & \ldots & \ldots & O_{(p)} \\ O_{(p)} & O_{(p)} & I_{(p)} & O_{(p)} & \ldots & O_{(p)} \\ \vdots & \ddots & \ddots & \ddots & \ddots & \vdots \\ O_{(p)} & \ldots & \ldots & O_{(p)} & I_{(p)} & O_{(p)} \end{pmatrix}_{r\times r}$$

alors on a

$$\underline{Y}_t = A_{\mathbf{s}_t}(\varepsilon_t)\underline{Y}_{t-1} + \underline{\eta}_{\mathbf{s}_t}(\varepsilon_t). \tag{2.12}$$

Pour la représentation (2.12) nous avons les résultat suivant

Proposition 2.3.3 *Soit $\widetilde{\gamma}_L(A)$ l'exposant de Lyapunov associé à la suite des matrices aléatoires $A := (A_{\mathbf{s}_t}(\varepsilon_t))_{t\in\mathbb{Z}}$ et supposons que $(\varepsilon_t)_{t\in\mathbb{Z}}$ est une suite i.i.d telle que $E\left\{|\varepsilon_0|^\delta\right\} < +\infty$ pour certains $\delta > 0$. Alors si $\widetilde{\gamma}_L(A) < 0$ il existe $\delta^* \in]0;1[$ de telle sorte que $E\left\{|X_t|^{\delta^*}\right\} < +\infty$.*

Preuve Notans que puisque $\widetilde{\gamma}_L(A) < 0$, alors l'équation (2.12) admette une solution unique, causale, strictement stationnaire, géométriquement stable et ergodique donnée par

$$\underline{Y}_t = \sum_{k\geq 0} Y_t(k)\,\underline{\eta}_{\mathbf{s}_{t-k}}(\varepsilon_{t-k}) \tag{2.13}$$

où $Y_t(k) := \left\{\prod_{i=0}^{k-1} A_{\mathbf{s}_{t-i}}(\varepsilon_{t-i})\right\}$ et la série (2.13) converge absolument p.s. Nous devons d'abord montrer que si $\widetilde{\gamma}_L(A) < 0$ il existe $\delta^* \in]0;1[$ et $t_0 > 0$ tels que $E\left\{\|Y_t(t_0)\|^{\delta^*}\right\} < 1$. En effet, considérons une norme $\|.\|$ de telle sort que $\|M\|^\delta \leq \left\||M|^\delta\right\|$, par la définition de $\widetilde{\gamma}_L(A)$, il existe un entier $t_0 \geq 1$ tel que

$E\left\{\log\left\|Y_t\left(t_0\right)\right\|\right\}<0$. D'autre part d'après la sous-multiplicativité de $|.|^{\delta}$ nous avons

$$\begin{aligned}
E\left\{\left\|Y_t\left(t_0\right)\right\|^{\delta}\right\} &\leq E\left\{\left\|\prod_{i=0}^{t_0-1}\left|A_{\mathbf{s}_{t-i}}\left(\varepsilon_{t-i}\right)\right|^{\delta}\right\|\right\}\\
&= \left\|E\left\{\prod_{i=0}^{t_0-1}\left|A_{\mathbf{s}_{t-i}}\left(\varepsilon_{t-i}\right)\right|^{\delta}\right\}\right\|\\
&= \left\|E\left\{E\left\{\prod_{i=0}^{t_0-1}\left|A_{\mathbf{s}_{t-i}}\left(\varepsilon_{t-i}\right)\right|^{\delta}\left|\mathbf{s}_t,\mathbf{s}_{t-1},\ldots,\mathbf{s}_{t-t_0+1}\right.\right\}\right\}\right\|\\
&= \left\|E\left\{\prod_{i=0}^{t_0-1}\left|A_{\mathbf{s}_{t-i}}\left(\mu\right)\right|^{\delta}\right\}\right\|\\
&= \left\|\mathbb{I}'\mathbb{P}^{t_0-1}\left(A\left(\mu\right)\right)\Pi\left(A\left(\mu\right)\right)\right\|<+\infty
\end{aligned}$$

où $A\left(\mu\right)=\left(\left|A_i\left(\mu\right)\right|^{\delta},1\leq i\leq d\right)$ et $\mu:=E\left\{\left|\varepsilon_0\right|^{\delta}\right\}$. Soit la fonction $f\left(t\right)=E\left\{\left\|Y_t\left(t_0\right)\right\|^{t}\right\}$ comme $f'\left(0\right)=E\left\{\log\left\|Y_t\left(t_0\right)\right\|\right\}<0$ alors $f\left(t\right)$ décroit dans un voisinage de 0 et comme $f\left(0\right)=1$, il s'ensuit qu'il existe $\delta^*\in\left]0,1\right[$ de telle sorte que $E\left\{\left\|Y_t\left(t_0\right)\right\|^{\delta^*}\right\}<1$ et donc $\left\|\underline{Y}_t\right\|\leq\sum_{k\geq1}\left\|Y_t\left(k\right)\underline{\eta}_{\mathbf{s}_{t-k}}\left(\varepsilon_{t-k}\right)\right\|+\left\|\underline{\eta}_{\mathbf{s}_t}\left(\varepsilon_t\right)\right\|$. Comme $\delta^*\in\left]0;1\right[$, nous obtenons $\left\|\underline{Y}_t\right\|^{\delta^*}\leq\sum_{k\geq1}\left\|Y_t\left(k\right)\underline{\eta}_{\mathbf{s}_{t-k}}\left(\varepsilon_{t-k}\right)\right\|^{\delta^*}+\left\|\underline{\eta}_{\mathbf{s}_t}\left(\varepsilon_t\right)\right\|^{\delta^*}$ et par le théorème de la convergence dominée et en utilisant le fait que $\limsup_{k\longrightarrow\infty}\left\|Y_t\left(k\right)\underline{\eta}_{\mathbf{s}_{t-k}}\left(\varepsilon_{t-k}\right)\right\|^{\frac{1}{k}}\leq\exp\left\{\widetilde{\gamma}_L\left(A\right)\right\}<1$ la dernière série converge en $\mathbb{L}_{\delta^*}$.

La proposition suivante établit un résultat similaire à la proposition (2.3.1) pour la représentation (2.12).

Proposition 2.3.4 *Supposons que* $\left(\varepsilon_t\right)_{t\in\mathbb{Z}}$ *est une suite i.i.d avec* $E\left|\varepsilon_0\right|^{2\delta}<+\infty$ *pour certains* $0<\delta\leq1$, *si* $\rho\left(\mathbb{P}\left(A\left(\mu\right)\right)\right)<1$, *alors l'équation* (2.12) *admette une solution unique strictement stationnaire, asymptotiquement stable et ergodique donnée par la série* (2.13). *De plus, pour tout* $t\in\mathbb{Z}$ *; la séquence de* $\left(Y_t\left(k\right)\underline{y}\right)_{k\geq1}$ *converge p.s. vers* $\underline{0}$ *pour tout vecteur* $\underline{y}\in\mathbb{R}^r$.

Preuve La preuve décolle essentiellement de la même façon que dans la proposition (2.3.1).

La représentation (2.12) montre que ε_t est indépendante de $\left\{\left(\underline{Y}'_{s-1},\mathbf{s}_s\right)',s\leq t\right\}$ et donc $\left(\left(\underline{Y}'_t,\mathbf{s}_t\right)'\right)_{t\in\mathbb{Z}}$ est une chaîne homogène de Feller à valeurs dans $\mathbb{R}^r\times\mathbb{S}$. Cependant, une des conditions les plus utilisés dans l'établissement de stationnarité et d'ergodicité géométrique des chaînes $\left(\underline{Z}_t,t\in\mathbb{Z}\right)$ sur $\mathbb{R}^k$ en temps discret, apériodique, $\phi-$irréductibles avec probabilité de transition à l'étape n notée $Q^{(n)}(\underline{z}, A) = Q(\underline{Z}_n \in A \,|\underline{Z}_0 = \underline{z})$, $(A, \underline{z}) \in \mathcal{B}_{\mathbb{R}^k} \times \mathbb{R}^k$ est la condition de dérive (voir Meyn et Tweedie (c.f.,[56]) et les références citées). Cette condition nécessite l'existence d'une fonction $g : \mathbb{R}^k \longrightarrow [1; +\infty[$ un compact $C \subset \mathbb{R}^k$ et une paire de constantes $(b, \alpha) \in [0; +\infty[\times]0; 1[$ telle que $\int_{\mathbb{R}^k} g(\underline{y})\, Q(\underline{z}, d\underline{y}) \leq \alpha g(\underline{z}) + b\mathbb{I}_C(\underline{z})$. Pour établir cette condition, faisons les hypothèses suivantes

A.0 $(\varepsilon_t)_{t\in\mathbb{Z}}$ est une séquence i.i.d des variables aléatoires de distribution marginale absolument continue par rapport à la mesure de Lebesgue λ. Le support de ε_t est définie par sa densité strictement positive contenat un ensemble ouvert et le zéro

A.1 $E\left\{|\varepsilon_t|^{2\delta}\right\} < +\infty$ pour un certain $0 < \delta \leq 1$.

A.2 $\widetilde{\gamma}_L(A) < 0$.

La condition **A.0** assure que $\left((\underline{Y}'_t, \mathbf{s}_t)'\right)_{t\in\mathbb{Z}}$ est une chaîne de Feller $(\lambda \otimes \nu)-$irréductible où ν est la mesure de comptage sur $\mathbb{S}$. Lorsqu'elle est associée aux propriétés de la chaîne (s_t), alors tout ensemble compact est un petit ensemble. Posons $\underline{\widetilde{Y}}_t := (\underline{Y}'_t, \mathbf{s}_t)'$ avec représentation d'espace état

$$\underline{\widetilde{Y}}_t = \widetilde{A}_{\mathbf{s}_t}(\varepsilon_t)\,\underline{\widetilde{Y}}_{t-1} + \widetilde{\underline{\eta}}_{\mathbf{s}_t}(\varepsilon_t) \tag{2.14}$$

où $\widetilde{\underline{\eta}}_{\mathbf{s}_t}(\varepsilon_t) := \left(\underline{\eta}'_{\mathbf{s}_t}(\varepsilon_t), \mathbf{s}_t\right)'$ et $\widetilde{A}_{\mathbf{s}_t}(\varepsilon_t)$ est une matrice appropriée.

Théorème 2.3.4 *Sous les conditions **A.0-A.2** la chaîne $\left(\underline{\widetilde{Y}}_t\right)_{t\in\mathbb{Z}}$ est géométriquement ergodique. En outre, si $\underline{\widetilde{Y}}_t$ est initialisée par sa distribution stationnaire, alors la première composante de $\underline{\widetilde{Y}}_t$ est strictement stationnaire et $\boldsymbol{\beta}-$mélangeante à vitesse exponentielle de décroissance.*

Preuve Soit $\widetilde{Y}_t(k) = \left\{\prod_{l=0}^{k-1} \widetilde{A}_{\mathbf{s}_{t-l}}(\varepsilon_{t-l})\right\}$ et de (2.14), nous avons par récurrence

$$\left\|\underline{\widetilde{Y}}_t\right\|^{\frac{1}{t}} \leq \sum_{k=0}^{t-1} \left\|\widetilde{Y}_t(k)\right\|^{\frac{1}{t}} \left\|\widetilde{\eta}_{\mathbf{s}_{t-k}}(\varepsilon_{t-k})\right\|^{\frac{1}{t}} + \left\|\widetilde{Y}_t(t)\right\|^{\frac{1}{t}} \left\|\underline{\widetilde{Y}}_0\right\|^{\frac{1}{t}}. \tag{2.15}$$

Sous **A.2** nous avons $\frac{1}{t}\log\left\|\widetilde{Y}_t(t)\right\| \longrightarrow \widetilde{\gamma}_L(A) < 0$ et donc $\left\|\widetilde{Y}_t(t)\right\|^{\frac{1}{t}} \longrightarrow \exp^{\widetilde{\gamma}_L(A)} < 1$ d'autre part, les variables aléatoires $\left\|\widetilde{Y}_t(t)\right\|^{\frac{1}{t}}$ sont p.s. bornées par $\sup_t \|A_{\mathbf{s}_t}(\varepsilon_t)\|$. D'après le théorème de convergence dominée, nous avons pour tous les $x \in \mathbb{S}$, $\lim_{t \longrightarrow \infty} E\left\{ \left\| \widetilde{Y}_t(t)\, \mathbb{I}_{\{\mathbf{s}_1 = x\}} \right\|^{\frac{1}{t}} \right\} < 1$ il s'ensuive qu'il existe $p \geq \frac{1}{\delta}$ et une constante positive $K < 1$ de telle sorte que $\alpha_p := \sup_{x \in \mathbb{S}} E\left\{ \left\| \widetilde{Y}_p(p)\, \mathbb{I}_{\{\mathbf{s}_1 = x\}} \right\|^{\frac{1}{p}} \right\} \leq K$. Prenons maintenant l'espérance conditionnelle des deux côtes de l'inégalité (2.15), nous obtenons $E\left\{ \left\| \underline{\widetilde{Y}}_p \right\|^{\frac{1}{p}} \middle| \underline{\widetilde{Y}}_0 = \underline{z} \right\} \leq \alpha_p \left\| \underline{y} \right\|^{\frac{1}{p}} + \beta_p$ pour tous les $\underline{z} = \left(\underline{y}', x \right)' \in \mathbb{R}^r \times \mathbb{S}$ où $\beta_p = \left\{ \sum_{k=0}^{p-1} \left\| \widetilde{Y}_p(k) \right\|^{\frac{1}{p}} \left\| \underline{\widetilde{\eta}}_{\mathbf{s}_{p-k}}(\varepsilon_{p-k}) \right\|^{\frac{1}{p}} | \underline{Z}_0 = \underline{z} \right\}$. Soit V la fonction de Lyapunov définie par $V(\underline{z}) = \|\underline{z}\|^{\frac{1}{p}} + 1$, alors nous avons $E\left\{ V\left(\widetilde{Y}_p \right) \middle| \underline{\widetilde{Y}}_0 = \underline{z} \right\} \leq \alpha_p V(\underline{z}) + \beta_p + 1$ posons $K' = \alpha_p + \frac{1 - \alpha_p}{2}$ et définissons le compact

$$C = \left\{ \underline{z} = \left(\underline{y}', x \right)' \in \mathbb{R}^r \times \mathbb{S} : K'V(\underline{z}) \leq \alpha_p V(\underline{z}) + \beta_p + 1 \right\}.$$

Alors, il est facile de voir que $0 < K' < 1$ et que $\int_{\mathbb{R}^r} V\left(\underline{y}\right) Q\left(\underline{z}, d\underline{y}\right) \leq K'V(\underline{z}) - \alpha_p \mathbb{I}_C(\underline{z})$ ce qui achève la démonstration.

2.4 L'inversibilité

L'inversibilité est une notion très importante dans l'étude des modèles de série chronologiques tant linéaire que non linéaire, car elle permet d'une part de faire des prévisions, d'autre part, d'estimer dans certains cas les paramètres du modèle. Précisons dans quelles circonstances ont été utilisées dans le cadre des modèles bilinéaires. Guégan (c.f.,[26]), Quinn (c.f.,[62]), Pham et Tran (c.f.,[59]). Pour obtenir des conditions d'inversibilité notamment pour les modèles bilinéaires. Guégan et Pam (c.f.,[27]), Subba Rao (c.f.,[64]) ont suggérés des conditions suffisantes d'inversibilité pour une large classes de modèles bilinéaires. Pour fixer les idées, considérons la classe de modèles

2.4.1 MS-BL$(1,0,1,1)$

Soit le modèle

$$X_t = a_1(\mathbf{s}_t) X_{t-1} + \varepsilon_t + c_{11}(\mathbf{s}_t) X_{t-1}\varepsilon_{t-1} \tag{3.1}$$

ou encore

$$X_t = Z_{t-1} + \varepsilon_t \tag{3.2}$$

où $Z_{t-1} = [a_1(\mathbf{s}_t) + c_{11}(\mathbf{s}_t)\varepsilon_{t-1}] X_{t-1}$ alors ε_t peut être calculé de manière récursive $\varepsilon_t = X_t - Z_{t-1}$ et

$$\begin{aligned} Z_{t-1} &= (a_1(\mathbf{s}_t) + c_{11}(\mathbf{s}_t)\varepsilon_{t-1}) X_{t-1} \qquad (3.3) \\ &= (a_1(\mathbf{s}_t) + c_{11}(\mathbf{s}_t) X_{t-1}) X_{t-1} - c_{11}(\mathbf{s}_t) X_{t-1} Z_{t-2}. \end{aligned}$$

Si Z_0 est connu, alors (3.3) peut être utilisé pour calculer ε_t en fonction de $X_1, X_2, \ldots, X_t$. Comme les Z_t ne sont pas observables, on peut prendre Z_0 égale à une valeur arbitraire, par exemple z_0, ce qui donne une valeur initiale du bruit que nous désignons par $\varepsilon(t|\, z_0)$. Cependant, et selon Granger et Andersen, nous disons que le processus X_t est inversible si $\varepsilon(t|\, z_0) - \varepsilon_t$ converge vers 0 indépendamment de z_0 au sens que l'on précisera par la suite (c.f.,[60]). La discussion ci-dessus suppose que nous connaissons les valeurs des paramètres $a_1(\mathbf{s}_t)$, $c_{11}(\mathbf{s}_t)$ du modèle. En pratique, nous ajustons le modèle aux données en supposant que le processus d'observation obéit au modèle (3.1) de paramètre inconnus $\underline{\theta} = (a_1(\mathbf{s}_t), c_{11}(\mathbf{s}_t))$ que l'on cherche à estimer. Nous sommes amenés donc à remplacer $\underline{\theta}$ par un certain estimateur $\tilde{\underline{\theta}} = (\tilde{a}_1(\mathbf{s}_t), \tilde{c}_{11}(\mathbf{s}_t))$. Le bruit est alors calculé par la méthode de récursion (3.3) avec $\tilde{\underline{\theta}}$ au lieu de $\underline{\theta}$, la valeur obtenue sera notée $\varepsilon_{\tilde{\underline{\theta}}}(t|\, z_0)$ pour indiquer sa dépendance en $\tilde{\underline{\theta}}$. Ainsi, nous disons que le modèle (3.1) est inversible par rapport au processus d'observation X_t relativement à $\tilde{\underline{\theta}}$, s'il existe un processus stationnaire $\varepsilon_{\tilde{\underline{\theta}}}(t)$ de telle sorte que $\varepsilon_{\tilde{\underline{\theta}}}(t|\, z_0) - \varepsilon_{\tilde{\underline{\theta}}}(t)$ converge vers 0. Notons que cette définition est similaire à celle de Granger et Andersen (1978b) qui exigent seulement que le second moment de $\varepsilon_{\tilde{\underline{\theta}}}(t|\, z_0) - \varepsilon_t$ tend vers une limite finie.

Théorème 2.4.1 *Le modèle* (3.1) *est fortement inversible en* $\tilde{\underline{\theta}}$ *selon que* $\prod_{i=1}^{d} (\tilde{c}_{11}(i))^{\pi(i)}$ *est strictement inférieur ou strictement supérieur à* $\exp\{-E\log|X_t|\}$ *et* $E\log^{+}|X_t| < +\infty$ *(c.f.,[60]).*

Preuve Considérons l'équation (3.3) avec $\tilde{a}_1(\mathbf{s}_t), \tilde{c}_{11}(\mathbf{s}_t)$ à la place $a_1(\mathbf{s}_t), c_{11}(\mathbf{s}_t)$ et avec une valeur initiale z_0 alors

$$\begin{aligned}
\varepsilon_{\tilde{\underline{\theta}}}(t|\, z_0) &= X_t - [\tilde{a}_1(\mathbf{s}_t) + \tilde{c}_{11}(\mathbf{s}_t) X_{t-1}] X_{t-1} + \tilde{c}_{11}(\mathbf{s}_t) X_{t-1} Z_{t-2} \\
&= X_t - [\tilde{a}_1(\mathbf{s}_t) + \tilde{c}_{11}(\mathbf{s}_t) X_{t-1}] X_{t-1} + \\
&\quad \tilde{c}_{11}(\mathbf{s}_t) X_{t-1} \{[\tilde{a}_1(\mathbf{s}_{t-1}) + \tilde{c}_{11}(\mathbf{s}_{t-1}) X_{t-2}] X_{t-2} - \tilde{c}_{11}(\mathbf{s}_{t-1}) X_{t-2} Z_{t-3}\} \\
&= X_t - [\tilde{a}_1(\mathbf{s}_t) + \tilde{c}_{11}(\mathbf{s}_t) X_{t-1}] X_{t-1} + \tilde{c}_{11}(\mathbf{s}_t) [\tilde{a}_1(\mathbf{s}_{t-1}) + \tilde{c}_{11}(\mathbf{s}_{t-1}) X_{t-2}] X_{t-1} X_{t-2} \\
&\quad - \tilde{c}_{11}(\mathbf{s}_t) \tilde{c}_{11}(\mathbf{s}_{t-1}) X_{t-1} X_{t-2} Z_{t-3}
\end{aligned}$$

par récurrence, on obtient

$$\varepsilon_{\tilde{\underline{\theta}}}(t|\, z_0) = X_t - \sum_{j=1}^{t-1} (-1)^{j-1} \tilde{c}_t(j) \left\{\prod_{k=1}^{j} X_{t-k}\right\} \{\tilde{a}_1(\mathbf{s}_{t-j+1}) + \tilde{c}_{11}(\mathbf{s}_{t-j+1}) X_{t-j}\} - (-1)^{t-1} \tilde{c}_t(t) \left\{\prod_{k=1}^{t-1} X_{t-k}\right\} z_0 \tag{3.4}$$

où $\tilde{c}_t(j) = \prod_{k=1}^{j-1} \{\tilde{c}_{11}(\mathbf{s}_{t-k+1})\}$ avec $\tilde{c}_t(0) = 0$. Le théorème ergodique montre que $\frac{1}{j}\log\left|\prod_{k=1}^{j} X_{t-k}\right| = \frac{1}{j}\sum_{k=1}^{j}\log|X_{t-k}|$ converge p.s. quand $j \longrightarrow \infty$ vers $E\log|X_t|$ donc $\lim_{j\longrightarrow\infty}\left|\prod_{k=1}^{j} X_{t-k}\right|^{\frac{1}{j}} = \exp\{E\log|X_t|\}\, p.s.$

D'autre part si $\prod_{i=1}^{d} |\tilde{c}_{11}(i)|^{\pi(i)} < \exp\{-E \log |X_t|\}$ alors presque surement la série

$$X_t - \sum_{j=1}^{\infty} (-1)^{j-1} \tilde{c}_t(j) \left\{ \prod_{k=1}^{j} X_{t-k} \right\} \{\tilde{a}_1(\mathbf{s}_{t-j+1}) + \tilde{c}_{11}(\mathbf{s}_{t-j+1}) X_{t-j}\}$$

convergente avec une somme égale $\varepsilon_{\underline{\tilde{\theta}}}(t)$ et nous avons $\varepsilon_{\underline{\tilde{\theta}}}(t|\, z_0) - \varepsilon_{\underline{\tilde{\theta}}}(t) \longrightarrow 0$ p.s. comme que $t \longrightarrow \infty$. Si $\prod_{i=1}^{d} |\tilde{c}_{11}(i)|^{\pi(i)} > \exp\{-E \log |X_t|\}$ alors la distribution de $\varepsilon_{\underline{\tilde{\theta}}}(0|\, z_0) - \varepsilon_{\underline{\tilde{\theta}}}(0)$ qui est la même que celle de $X_0 - \sum_{j=1}^{t-1} (-1)^{j-1} \tilde{c}_t(j) \left\{ \prod_{k=1}^{j} X_{-k} \right\} \{\tilde{a}_1(\mathbf{s}_{-j+1}) + \tilde{c}_{11}(\mathbf{s}_{-j+1}) X_{-j}\} - \varepsilon_{\underline{\tilde{\theta}}}(0)$ ne peut converger vers une limite quand $t \longrightarrow \infty$.

Corollaire 2.4.1 *Le processus* (3.1) *est fortement inversible ou non selon que* $\prod_{i=1}^{d} (c_{11}(i))^{\pi(i)}$ *est strictement inférieur ou strictement supérieur à* $\exp\{-E \log |X_t|\}$ *et* $E \log^+ |X_t| < +\infty$.

2.4.2 MS-BL(p, q, P, Q)

Sans perdre de généralité, on suppose que $Q = q$

$$\forall t \in \mathbb{Z},\ X_t = \sum_{i=1}^{p} a_i(\mathbf{s}_t) X_{t-i} + \sum_{j=0}^{q} b_j(\mathbf{s}_t) \varepsilon_{t-j} + \sum_{i=1}^{P} \sum_{j=1}^{q} c_{ij}(\mathbf{s}_t) X_{t-i} \varepsilon_{t-j}, b_0(\mathbf{s}_t) = 1$$

i.e.,

$$\forall t \in \mathbb{Z},\ X_t = \sum_{i=1}^{p} a_i(\mathbf{s}_t) X_{t-i} + \sum_{j=1}^{q} \left[b_j(\mathbf{s}_t) + \sum_{i=1}^{p} c_{ij}(\mathbf{s}_t) X_{t-i} \right] \varepsilon_{t-j} + \varepsilon_t$$

donc le modèle $MS - BL(p, q, P, q)$ admet la représentation suivante

$$\underline{\varepsilon}_t = M_t \underline{\varepsilon}_{t-1} + \underline{K}_t \tag{3.9}$$

où $\underline{\varepsilon}_t := (\varepsilon_t, \varepsilon_{t-1}, \ldots, \varepsilon_{t-q+1})'_{q\times 1}$, $\underline{K}_t := N(\mathbf{s}_t)\,\underline{X}_{t-1}$, $\underline{X}_t := (X_t, X_{t-1}, \ldots, X_{t-p+1})'_{p\times 1}$,et $M_t := M_0(\mathbf{s}_t) + \sum_{j=1}^{P} M_j(\mathbf{s}_t)\,X_{t-j}$ avec $M_0(\mathbf{s}_t)$, $M_j(\mathbf{s}_t)$ et $N(\mathbf{s}_t)$ sont $q\times q$, $q\times q$, $q\times p$ matrices définies par

$$\begin{aligned}
M_0(\mathbf{s}_t) &:= \begin{pmatrix} b_1(\mathbf{s}_t) & \ldots & b_q(\mathbf{s}_t) \\ I_{(q-1)} & & \underline{O}_{(q-1)} \end{pmatrix}_{q\times q} \\
M_j(\mathbf{s}_t) &:= \begin{pmatrix} c_{j1}(\mathbf{s}_t) & \ldots & c_{jq}(\mathbf{s}_t) \\ & O_{(q-1,p)} & \end{pmatrix}_{q\times p}, j = 1, \ldots, P \\
N(\mathbf{s}_t) &:= \begin{pmatrix} a_1(\mathbf{s}_t) & \ldots & a_p(\mathbf{s}_t) \\ & O_{(q-1,p)} & \end{pmatrix}_{q\times p}.
\end{aligned}$$

Alors on a

Proposition 2.4.1 *Le modèle $(MS - BL(p,1,p,1))$ est fortement inversible si*

$$\sum_{i=1}^{d} \pi(i)\, E\left\{\log\left|\sum_{j=1}^{p} c_{j1}(i)\, X_{t-j} + b_1(i)\right|\right\} < 0. \tag{3.10}$$

En particulier, pour $p = 1$ et $b_1(.) = 0$ la condition (3.10) est équivalente à $\prod_{i=1}^{d} |c_{11}(i)|^{\pi(i)} < \exp\{-E\log|X_t|\}$.

Chapitre 3

Théorèmes limites pour les $MS - BL$ super-diagonaux

Nous examinons maintenant les modèles $MS - BL$ super-diagonaux

$$\underline{X}_t = \underline{H}\varepsilon_t + A_0(\mathbf{s}_t)\underline{X}_{t-1} + A_1(\mathbf{s}_t)\underline{X}_{t-1}\varepsilon_{t-1} \tag{4.1}$$

où $(\varepsilon_t)_{t\in\mathbb{Z}}$ est une suite i.i.d. De toute évidence, (4.1) est un cas particulier de (2.2) avec $q = 0$ et $Q = 1$. Dans cette section, nous étudions des conditions appropriées, pour que la moyenne et covariances empiriques satisfaire le théorème central limite $(T.C.L)$ et la loi du logarithme itéré $(L.I.L)$. Soient $\Im_t = \sigma\{\varepsilon_u, \mathbf{s}_u, u \leq t\}$ et T l'opérateur de retard , c-à-d pour $\underline{x} = (..., x_{-1}, x_0, x_1, ...) \in \mathbb{R}^\infty$, $T(\underline{x}) = (..., x_0, x_1, x_2, ...)$ dans cette section, on suppose que $(\varepsilon_t)_{t\in\mathbb{Z}}$ vérifié (2.8) et supposons que

$$\lambda = \rho\left(\mathbb{P}\left(A_0^{\otimes 2} + \sigma^2 A_1^{\otimes 2}\right)\right) < 1 \tag{4.2}$$

Alors il existe une fonction mesurable $g(.)$ de $\mathbb{R}^\infty$ à $\mathbb{R}$ de telle sorte que $X_t = g(\varepsilon_t, s_t, \varepsilon_{t-1}, s_{t-1}, ...)$ $p.s.\forall t \in \mathbb{Z}$ c-à-d $(X_t)_{t\in\mathbb{Z}}$ est causale i.e.,

$$\underline{X}_t = \underline{H}\varepsilon_t + \sum_{k=1}^{\infty}\prod_{j=0}^{k-1}\left[A_0(\mathbf{s}_{t-j+1}) + A_1(\mathbf{s}_{t-j+1})\varepsilon_{t-j}\right]\underline{H}\varepsilon_{t-k} \tag{4.3}$$

où la série converge absolument presque sûrement et dans $\mathbb{L}^2$. Pour le processus $(X_t)_{t\in\mathbb{Z}}$ univarié défini par la première composante du processus $(\underline{X}_t)_{t\in\mathbb{Z}}$ satisfaisant (4.1), alors posons $Y_t = g(\varepsilon_t, s_t, \varepsilon_{t-1}, s_{t-1}, ...) - \mu$

ici $\mu = E(X_0)$ et $g(.)$ est la fonction mesurable déterminée par (4.3) de sorte que $X_t = g(\varepsilon_t, s_t, \varepsilon_{t-1}, s_{t-1}, ...)$. Comme $EY_0 = 0$, $EY_0^2 < \infty$ et Y_0 est $\Im_0-$mesurable. En outre $Y_t = Y_0(T'(\underline{\varepsilon}))$ où $\underline{\varepsilon} = (\ldots, \varepsilon_{-1}, s_t, \varepsilon_0, s_0, \varepsilon_1, s_1 \ldots)$ est strictement stationnaire et ergodique. La première somme partielle qui nous s'intéresse est

$$S_n = \sum_{t=1}^{n} Y_t = \sum_{t=1}^{n} X_t - n\mu.$$

Théorème 3.0.2 *Considérons le modèle bilinéaire (4.1) avec $(X_t)_{\ t\in\mathbb{Z}}$ défini comme étant la première composante du processus $(\underline{X}_t)_{\ t\in\mathbb{Z}}$. Sous les hypothèses (2.8) et (4.2), on a*

$$\frac{S_n}{\sqrt{n}} \underset{n\longrightarrow+\infty}{\overset{d}{\longrightarrow}} \mathcal{N}\left(0, \sigma^{*2}\right) \ p.s. \tag{4.4}$$

où $\sigma^* = \sqrt{\lim\limits_{n\longrightarrow\infty} E\left(\frac{S_n^2}{n}\right)} < \infty$ si en plus $\sigma^* > 0$ puis

$$\limsup_{n\longrightarrow+\infty} \frac{S_n}{\sqrt{2n\ln\ln n}} = \sigma^* \ p.s., \ \liminf_{n\longrightarrow+\infty} \frac{S_n}{\sqrt{2n\ln\ln n}} = -\sigma^* \ p.s. \tag{4.5}$$

Preuve Nous avons pour chaque $t > 0$

$$\begin{aligned} E(\underline{X}_t |\Im_0) &= \sum_{j=0}^{t-2} E\left(\prod_{k=0}^{j} (A_0(\mathbf{s}_{t-k}) + A_1(\mathbf{s}_{t-k})\varepsilon_{t-k-1}) \underline{H}\varepsilon_{t-j-1} |\Im_0\right) \\ &\quad + E\left(\prod_{j=0}^{t-1} (A_0(\mathbf{s}_{t-j}) + A_1(\mathbf{s}_{t-j})\varepsilon_{t-j-1}) \underline{X}_0 |\Im_0\right) \\ &= \sigma^2 \sum_{j=0}^{t-2} \underline{\mathbf{1}}'\mathbb{P}^j(A_0)\Pi(A_1)\underline{H} + \underline{\mathbf{1}}'\mathbb{P}^{t-1}(A_0)\Pi((A_0 + A_1\varepsilon_0)\underline{X}_0). \end{aligned}$$

Pour établir le théorème ci-dessus, nous vérifions seulement les conditions du théorème (5.5) de Hall et Heyde (1980). Soit $y_t = E(Y_t |\Im_0) - E(Y_t |\Im_{-1}) = E(X_t |\Im_0) - E(X_t |\Im_{-1}) = \underline{H}'(E(\underline{X}_t |\Im_0) - E(\underline{X}_t |\Im_{-1}))$

et

$$E(\underline{X}_t |\Im_0) = \begin{cases} \sigma^2 \sum\limits_{j=0}^{t-2} \underline{\mathbf{1}}'\mathbb{P}^j(A_0)\Pi(A_1)\underline{H} + \underline{\mathbf{1}}'\mathbb{P}^{t-1}(A_0)\Pi((A_0 + A_1\varepsilon_0)\underline{X}_0) & \text{si } t > 0 \\ \underline{X}_t & \text{si } t \leq 0 \end{cases} \tag{4.6}$$

et

$$E\left(\underline{X}_t\left|\Im_{-1}\right.\right)=\begin{cases}\sigma^2\sum_{j=0}^{t-1}\underline{\mathbf{1}}'\mathbb{P}^j\left(A_0\right)\Pi\left(A_1\right)\underline{H}+\underline{\mathbf{1}}'\mathbb{P}^t\left(A_0\right)\Pi\left(\left(A_0+A_1\varepsilon_{-1}\right)\underline{X}_{-1}\right) & \text{si } t>0\\ \Pi\left(\left(A_0+A_1\varepsilon_{-1}\right)\underline{X}_{-1}\right) & \text{si } t=0\\ \underline{X}_t & \text{si } t<0\end{cases}$$

donc

$$E\left(\underline{X}_t\left|\Im_0\right.\right)-E\left(\underline{X}_t\left|\Im_{-1}\right.\right)=\begin{cases}\underline{\mathbf{1}}'\mathbb{P}^{t-1}\left(A_0\right)\begin{pmatrix}\Pi\left(\left(A_0+A_1\varepsilon_0\right)\underline{X}_0-\sigma^2A_1\underline{H}\right)-\\ \underline{A}_0'\underline{\mathbf{1}}'\mathbb{P}\Pi\left(\left(A_0+A_1\varepsilon_{-1}\right)\underline{X}_{-1}\right)\end{pmatrix} & \text{si } t>0\\ \underline{X}_0-\Pi\left(\left(A_0+A_1\varepsilon_{-1}\right)\underline{X}_{-1}\right) & \text{si } t=0\\ \underline{0} & \text{si } t<0\end{cases}$$

et $y_t=\underline{H}'\left(\underline{\mathbf{1}}'\mathbb{P}^{t-1}\left(A_0\right)\left(\Pi\left(\left(A_0+A_1\varepsilon_0\right)\underline{X}_0-\sigma^2A_1\underline{H}\right)-\underline{A}_0'\underline{\mathbf{1}}'\mathbb{P}\Pi\left(\left(A_0+A_1\varepsilon_{-1}\right)\underline{X}_{-1}\right)\right)\right)$ pour tout $t>0$. Notant que $\left(\rho\left(A\right)\right)^2=\rho\left(A\otimes A\right)<1$ nous avons par la proposition (2.1) de Liu et Brockwell (1988)

$$Ey_t^2=\underline{H}'E\begin{pmatrix}\underline{\mathbf{1}}'\mathbb{P}^{t-1}\left(A_0\right)\left(\Pi\left(\left(A_0+A_1\varepsilon_0\right)\underline{X}_0-\sigma^2A_1\underline{H}\right)-\underline{A}_0'\underline{\mathbf{1}}'\mathbb{P}\Pi\left(\left(A_0+A_1\varepsilon_{-1}\right)\underline{X}_{-1}\right)\right)\\ \left(\Pi'\left(\left(A_0+A_1\varepsilon_0\right)\underline{X}_0-\sigma^2A_1\underline{H}\right)-\Pi'\left(\left(A_0+A_1\varepsilon_{-1}\right)\underline{X}_{-1}\right)\mathbb{P}'\underline{\mathbf{1}}\underline{A}_0\right)\mathbb{P}'^{t-1}\left(A_0\right)\underline{\mathbf{1}}\end{pmatrix}\underline{H}$$

$$\leq \quad const.\lambda^t \qquad \text{pour tout } t>0.$$

D'qaprès l'inégalité de Cauchy-Schwarz ona $E\left(\sum_{j=m}^{n}y_j\right)^2\leq const.\lambda^m$ pour certains $0<\lambda<1$ et $\forall 0<m<n$. Ceci établit (5.24) de Hall et Heyde (1980). D'autre part

$$\sum_{m=1}^{\infty}\left(\limsup_n E\left(\sum_{j=m}^{n}y_j\right)^2+\limsup_n E\left(\sum_{j=m}^{n}y_{-j}\right)^2\right)<\infty.$$

Enfin, montrons que $E\left(Y_0\left|\Im_\infty\right.\right)=Y_0$ p.s. et $E\left(Y_0\left|\Im_{-\infty}\right.\right)=0$ p.s., où $\Im_\infty=\sigma\left\{\bigcup_t\Im_t\right\}$ et $\Im_{-\infty}=\bigcap_t\Im_t$. La première relation est évidente depuis Y_0 est $\Im_0$ mesurable. Pour montrer la seconde, observons que pour $\underline{Y}_t=\underline{X}_t-\underline{\mu}$ on a pour tout $\forall m>0$

$$\begin{aligned}E\left(\underline{Y}_0\left|\Im_{-m}\right.\right)&=E\left(\underline{X}_0-\underline{\mu}\left|\Im_{-m}\right.\right)=E\left(\underline{X}_0\left|\Im_{-m}\right.\right)-\underline{\mu}\\ &=\sigma^2\sum_{k=1}^{m-1}\underline{\mathbf{1}}'\mathbb{P}^{k-1}\left(A_0\right)\Pi\left(A_1\right)\underline{H}+\underline{\mathbf{1}}'\mathbb{P}^{m-1}\left(A_0\right)\Pi\left(\left(A_0+A_1\varepsilon_{-m}\right)\underline{X}_{-m}\right)-\underline{\mu}\end{aligned}$$

et $\underline{\mu} = \sigma^2 \sum_{k=1}^{\infty} \underline{\mathbf{1}}' \mathbb{P}^{k-1}(A_0)\,\Pi(A_1)\,\underline{H}$ et en appliquant l'inégalité de Cauchy-Schwarz on a

$$E\left(E\left(Y_0 \left|\Im_{-\infty}\right.\right)\right)^2 = E\left(E\left(E\left(Y_0 \left|\Im_{-m}\right.\right)\left|\Im_{-\infty}\right.\right)\right)^2 \leq E\left(E\left(\left(E\left(Y_0 \left|\Im_{-m}\right.\right)\right)^2 \left|\Im_{-\infty}\right.\right)\right) \underset{m\longrightarrow\infty}{\longrightarrow} 0$$

De même, on peut considérer la fonction de covariance empirique $\hat{\gamma}(k) = \frac{1}{n}\sum_{j=k+1}^{n}\left(X_j - \overline{X}\right)\left(X_{j-k} - \overline{X}\right)$, $k \geq 0$ où $\overline{X} = \frac{1}{n}\sum_{t=1}^{n} X_t$ en considérons la seconde somme partielle suivante

$$S_n(k) = \sum_{j=k+1}^{n} X_j X_{j-k} - (n-k)\,\mu(k)\,, k \geq 0$$

où $\mu(k) = E(X_k X_0)$ alors

Théorème 3.0.3 *Soit $(X_t)_{t\in\mathbb{Z}}$ le processus défini par la première composante du processus $(\underline{X}_t)_{t\in\mathbb{Z}}$ défini par* (4.1)*. Supposons que les conditions* (2.8) *et* (4.2) *sont satisfaites. Alors*

$$\frac{S_n(k)}{\sqrt{n}} \underset{n\longrightarrow+\infty}{\overset{d}{\longrightarrow}} \mathcal{N}\left(0, \sigma^2(k)\right)$$

où $\sigma(k) = \sqrt{\lim_{n\longrightarrow+\infty} \frac{E(S_n^2(k))}{n}} < \infty$ si de plus $\sigma(k) > 0$, presque sûrement on a

$$\limsup_{n\longrightarrow+\infty} \frac{S_n(k)}{\sqrt{2n\ln\ln n}} = \sigma(k) \quad et \quad \liminf_{n\longrightarrow+\infty} \frac{S_n(k)}{\sqrt{2n\ln\ln n}} = -\sigma(k)$$

Preuve La preuve est la même comme pour le théorème précédent et en manipulant des espérances conditionnelles.

Deuxième partie

L'inférence statistique asymptotique dans les $MS - BL$

Chapitre 4

L'estimation dans les modèles $MS - BL$

Dans les applications de la statistique, il arrive fréquemment que certaines valeurs soient manquantes dans les jeux de données utilisés par les praticiens. Lorsque cette situation survient, plusieurs méthodes sont mises à la disposition des statisticiens pour remplacer les données manquantes par des valeurs imputées. De plus, lorsque toutes les données sont présentes et que leur distribution est connue, certaines méthodes peuvent être utilisées pour estimer les paramètres de la distribution. Dans ce chapitre, une technique qui permet d'estimer les paramètres lorsque des données sont absentes est étudiée.

4.0.3 La fonction de vraisemblance dans les modèles $MS - BL\,(p, 0, P, Q)$

Soit $b_0(s_t) = 1, m_x = \max\{p, P\}$, $m_\varepsilon = Q$ et supposons qu'on a $\{X_{-m_x+1}, ..., X_0, X_1, ..., X_N\}$ une observation du processus $MS - BL\,(p, 0, P, Q)$ stationnaire au second order $(X_t)_{t\in\mathbb{Z}}$ causal et inversible. Notons $\underline{\theta}:= (\underline{\theta}(k), k = 1, ..., d)$ où $\underline{\theta}(k) = \left(\underline{a}'(k), \underline{c}'(k), \underline{p}'(k), \pi(k), \sigma\,(k)\right)'$ où $\underline{a}'(k) = (a_j(k), 1 \leq j \leq p)$, $\underline{c}'_{i,j}(k) = (c_{i,j}(k), 1 \leq i \leq P, 1 \leq j \leq Q)$, $\underline{p}'(k) = (p_{k,j}, 1 \leq j \leq d)$. Supposons que

H0 la suite $(\varepsilon_t)_t$ est une suite $i.i$ distribuée selon une loi nornmale $\mathcal{N}\left(0, \sigma^2\right)$.

On suppose que les observation $\underline{X}_0$ où $\underline{X}_t = \{(X_{-m_x+1}, \varepsilon_{-m_\varepsilon+1}), ..., (X_t, \varepsilon_t)\}$ connues, alors la vraisemblance conditiontinalle (sachant $\underline{X}_{t-1}$) par rapport à la mesure $d(\lambda \otimes \nu)$ (où λ désigne la mesure de

Lebesgue et ν désigne celle de comptage sur $\mathbb{S}$) est définie pour tout $\underline{\theta}\in\Theta$ par

$$\begin{aligned} L_N(X_1,X_2,...,X_N,\underline{\theta}) &= \prod_{k=1}^{N} L(\ X_k|\underline{X}_{k-1},\underline{\theta}) = L(\ X_N|\underline{X}_{N-1},\underline{\theta}) \prod_{k=1}^{N-1} L(\ X_k|\underline{X}_{k-1},\underline{\theta}) \\ &= \sum_{i=1}^{N} L(\ X_N|\underline{X}_{N-1}, s_N=e_i,\underline{\theta}) P\left(s_N=e_i|\underline{X}_{N-1},\underline{\theta}\right) \prod_{k=1}^{N-1} L(\ X_k|\underline{X}_{k-1},\underline{\theta}) \end{aligned}$$

Si on note

1. $\pi_t(i) = P\left(s_t=e_i|\underline{X}_{t-1},\underline{\theta}\right)$

2. $b_t(i) = L(\ X_t|\underline{X}_{t-1}, s_t=e_i,\underline{\theta})$, la densité conditionnelle de X_t sachant $\underline{X}_{t-1}$ et $s_t=e_i$.

3. $\underline{\pi}'_i = (\pi_t(1),\pi_t(2),\ldots,\pi_t(i))$ et $\underline{b}'_i = (b_t(1),b_t(2),\ldots,b_t(i))$.

Alors on aura

$$L_N(X_1,X_2,...,X_N,\underline{\theta}) = \underline{b}'_N\underline{\pi}_N \prod_{k=1}^{N-1} L(\ X_k|\underline{X}_{k-1},\underline{\theta}) = \prod_{k=1}^{N} \underline{b}'_k\underline{\pi}_k$$

On en déduit une forme pratique de la log-vraisemblance

$$\log L_N(X_1,X_2,...,X_N,\underline{\theta}) = \sum_{k=1}^{N} \log\left\{\underline{b}'_k\underline{\pi}_k\right\}$$

Comme la Jacobienne de la transformation de (X_t) vers (ε_t) est égale à 1, alors

$$\begin{aligned} b_t(i) &= L(\ X_t|\underline{X}_{t-1}, s_t=e_i,\underline{\theta}) = f_{e_i}(X_t|\underline{X}_{t-1},\underline{\theta}) \\ &= \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left\{-\frac{1}{2\sigma^2}\left\{X_t - \sum_{l=1}^{p} a_l\left(e_i\right) X_{t-l} - \sum_{l=1}^{P}\sum_{j=1}^{Q} c_{lj}\left(e_i\right) X_{t-l}\varepsilon_{t-j}(e_i)\right\}^2\right\} \\ &= \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left\{-\frac{1}{2\sigma^2}\varepsilon_t^2\left(e_i\right)\right\} \end{aligned}$$

Notons que la vraisemblance des données complètes (4.2) peut s'écrire sous la forme

$$L_N(X_1,X_2,...,X_N,\underline{\theta}) = \sum_{(e_1,\ldots,e_N)\in\mathbb{S}^N} p_{\underline{\theta}}\left(X_1,X_2,...,X_N|\underline{X}_0\right) \tag{4.2}$$

où

$$p_{\underline{\theta}}\left(X_1,X_2,...,X_N|\underline{X}_0\right) = \pi(e_1)\prod_{t=2}^{N} p_{e_{t-1},e_t}\left\{\prod_{t=1}^{N} f_{e_t}(X_t|\underline{X}_{t-1},\underline{\theta})\right\}$$

ou encore

$$L_N(X_1, X_2, ..., X_N, \underline{\theta}) = \underline{1}' \left\{ \prod_{t=2}^{N} M_{\underline{\theta}}(f(X_t)) \right\} \Pi\left(f(X_1)\right) \tag{4.3}$$

où $M_{\underline{\theta}}(f\left(x_t\right)) = \mathbb{P}\left(f(x_t|\underline{x}_{t-1}, \underline{\theta})\right)$ et $\Pi\left(f(X_1)\right) = \Pi(f(x_1|\underline{x}_0, \theta))$ où $f(x_t|\underline{x}_{t-1}, \underline{\theta}) = \left(f_1(x_t|\underline{x}_{t-1}, \underline{\theta}), ..., f_d(x_t|\underline{x}_{t-1}, \underline{\theta})\right)$.

Il est aussi cependant préférable d'utiliser la formule

$$L_N(X_1, X_2, ..., X_N, \underline{\theta}) = \prod_{t=1}^{N} p_{\underline{\theta}}(X_t|\underline{X}_{t-1}) \tag{4.4}$$

où $p_{\underline{\theta}}(X_t|\underline{X}_{t-1}) = \sum_{s\in\mathbb{S}} p_{\underline{\theta}}\left(X_t|\underline{X}_{t-1}, s_t = s\right) p_{\underline{\theta}}(s_t = s|\underline{X}_{t-1})$. D'autre part, il est facile (pour des techniques de démonstrations) de considérer la vraisemblance conditionnelle suivante

$$\begin{aligned} L_N^*(X_1, X_2, ..., X_N, \underline{\theta}) &= \sum_{(e_1,...,e_N)\in\mathbb{S}^N} \pi(e_1) \left\{ \prod_{t=2}^{N} p_{e_{t-1},e_t} \right\} \left\{ \prod_{t=1}^{N} f_{e_t}(X_t|\underline{X}_{-\infty}, \underline{\theta}) \right\} \\ &= \underline{1}' \left\{ \prod_{t=2}^{N} M_{\underline{\theta}}^*(f(X_t)) \right\} \Pi\left(f(X_1)\right) \end{aligned}$$

dans lesquelles $\underline{X}_{-\infty} = \{(X_{t-s}, \varepsilon_{t-s}), s \leq 1\}$. Notons

1. $p_{\underline{\theta}}(X_t|\underline{X}_{t-1})$ la densité de X_t sachant $\underline{X}_{t-1}$ et $g_{\underline{\theta}}(X_t|\underline{X}_{t-1})$ sa logarithme
2. $p_{\underline{\theta}}^*(X_t|\underline{X}_{-\infty})$ la densité de X_t sachant $\underline{X}_{-\infty}$ et $g_{\underline{\theta}}^*(X_t|\underline{X}_{-\infty})$ sa logarithme

4.0.4 Consistence du MLE

Dans la suite nous supposerons que

H1 Θ est un compact

H2 Pour $\underline{\theta} = \underline{\theta}_0$ la chaîne $(s_t)_t$ est irréductible et apériodique

H3 Pour $\underline{\theta} = \underline{\theta}_0$, le modèle (2.2) admit une solution strictement stationnaire et ergodique

H4 Pour tout $\underline{\theta} \in \Theta$, $0 < \min_k f_k(x_t|\underline{x}_{-\infty}, \underline{\theta}) \leq \max_k f_k(x_t|\underline{x}_{-\infty}, \underline{\theta}) < +\infty$ et il existe un voisinage $\mathcal{V}(\underline{\theta}) = \left\{\underline{\theta}' : \left\|\underline{\theta} - \underline{\theta}'\right\| < \delta\right\}$ de $\underline{\theta}$ tel que $E_{\underline{\theta}_0}\left\{ \sup_{\underline{\theta}'\in\mathcal{V}(\underline{\theta})} \left|g_{\underline{\theta}'}^*(X_t|\underline{X}_{-\infty})\right| \right\} < +\infty$.

H5 Pour tout $\underline{\theta}_1, \underline{\theta}_2 \in \Theta$, si presque sûrement $p_{\underline{\theta}_1}^*(X_t|\underline{X}_{-\infty}) = p_{\underline{\theta}_2}^*(X_t|\underline{X}_{-\infty})$ alors $\underline{\theta}_1 = \underline{\theta}_2$

Commençons par montrer le résultat général suivant

Lemma 4.0.1 *Pour tout $\underline{\theta}\in\Theta$*

$$\begin{aligned}\lim_{N\to\infty}\frac{1}{N}\log L_N^*(X_1,X_2,...,X_N,\underline{\theta}) &= \lim_{N\to\infty}\frac{1}{N}\log L_N(X_1,X_2,...,X_N,\underline{\theta}) \\ &= E_{\underline{\theta}_0}\left\{g_{\underline{\theta}}^*(X_t|\underline{X}_{-\infty})\right\}\end{aligned}$$

Preuve La seconde égalité découle du théorème ergodique car le processus $\left(g_{\underline{\theta}}^*(X_t|\underline{X}_{t-1})\right)_t \geq 1$ est stationnaire est ergodique et en observant que

$$\begin{aligned}\frac{1}{N}\log L_N^*(X_1,X_2,...,X_N,\underline{\theta}) &= \frac{1}{N}\sum_{t=1}^{N}g_{\underline{\theta}}^*(X_t|\underline{X}_{-\infty}) \\ \frac{1}{N}\log L_N(X_1,X_2,...,X_N,\underline{\theta}) &= \frac{1}{N}\sum_{t=1}^{N}g_{\underline{\theta}}(X_t|\underline{X}_{t-1}).\end{aligned}$$

Montrons maintenant que ces deux quantités sont asymptotiquement presque sûrement égales quand $n\to+\infty$. Soit $Z_t(T)=\sup_{l\geq T}\left|g_{\underline{\theta}}(X_t|\underline{X}_{t-l})-g_{\underline{\theta}}^*(X_t|\underline{X}_{-\infty})\right|$, alors le processus $Z_t(T)$ est statitionnaire, erqodique et $E_{\underline{\theta}_0}\{|Z_t(T)|\}<+\infty$ et on a

$$\begin{aligned}\limsup_{N\to\infty}\left|\frac{1}{N}\sum_{t=1}^{N}\left(g_{\theta}^*(X_t|\underline{X}_{-\infty})-g_{\theta}(X_t|\underline{X}_{t-1})\right)\right| &\leq \limsup_{N\to\infty}\frac{1}{N}\sum_{t=1}^{N}\left|g_{\underline{\theta}}^*(X_t|\underline{X}_{-\infty})-g_{\underline{\theta}}(X_t|\underline{X}_{t-1})\right| \\ &\leq \limsup_{N\to\infty}\frac{1}{N}\sum_{t=T+1}^{N}Z_t(T)=E_{\underline{\theta}_0}\{|Z_t(T)|\}\end{aligned}$$

mais lorsque $T\to\infty, Z_t(T)\to 0$. D'où le résultat.

Lemma 4.0.2 *Posons*

$$Q_N(\underline{\theta})=\frac{1}{N}\log\frac{L_N^*(X_1,X_2,...,X_N,\underline{\theta})}{L_N^*(X_1,X_2,...,X_N,\underline{\theta}_0)}$$

pour tout $\underline{\theta}\in\Theta$. Alors sous les condition $H1-H4$, presque sûrement $\lim_{N\to\infty}Q_N(\underline{\theta}) \leq 0$ avec égalité si et seulement si $\underline{\theta}=\underline{\theta}_0$.

Preuve Sous les condition $H1-H4$, presque sûrement $Q_N(\theta)$ est bien définie. D'après le lemme précédent et l'inégalité de Jensen, on a

$$\lim_{N\to\infty}Q_N(\underline{\theta})= E_{\underline{\theta}_0}\left\{\log\frac{p_{\underline{\theta}}^*(X_t|\underline{X}_{-\infty})}{p_{\underline{\theta}_0}^*(X_t|\underline{X}_{-\infty})}\right\}\leq\log E_{\underline{\theta}_0}\left\{\frac{p_{\underline{\theta}}^*(X_t|\underline{X}_{-\infty})}{p_{\underline{\theta}_0}^*(X_t|\underline{X}_{-\infty})}\right\}=0.$$

D'où le résultat.

Notons $\lambda(\theta)$ l'exposent de Lypunov associé à la suite des matrices aléatoires $M_\theta = (M_\theta(\varepsilon_t))_t$, alors

$$\begin{aligned}\lambda(\theta) &= \lim_{N\to\infty} E_{\theta_0}\left\{\frac{1}{N}\log\left\|\prod_{t=2}^{N} M_{\underline{\theta}}(f(X_t))\right\|\right\}\\ &= \inf_{N>0} E_{\theta_0}\left\{\frac{1}{N}\log\left\|\prod_{t=2}^{N} M_{\underline{\theta}}(f(X_t))\right\|\right\}\\ &= \lim_{N\to\infty}\left\{\frac{1}{N}\log\left\|\prod_{t=2}^{N} M_{\underline{\theta}}(f(X_t))\right\|\right\}\end{aligned}$$

Notons que d'après (4.3). On a

$$\begin{aligned}&\min_j \pi(e_j) f_{ej}(X_t|\underline{X}_{-\infty},\theta)\left\|\prod_{t=2}^{N} M_\theta^*(f(X_t))\right\|\\ \leq\ & L_N^*(X_1,X_2,...,X_N,\theta)\\ \leq\ & \max_j \pi(e_j) f_{e_j}(X_t|\underline{X}_{-\infty},\theta)\left\|\prod_{t=2}^{N} M_\theta^*(f(X_t))\right\|\end{aligned}$$

et on obtient

$$\lambda(\theta) = \lim_{N\to\infty} E_{\theta_0}\left\{\frac{1}{N}\log\ L_N(X_1,X_2,...,X_N,\theta)\right\} = \lim_{N\to\infty}\frac{1}{N}\log\ L_N(X_1,X_2,...,X_N,\theta)\ \ p.s.$$

Cette discussion nous amènes à

Lemma 4.0.3 *Sous les hypothèses $H1-H5$, pour tout θ_1 d'un sous compact de Θ et $\theta_0 \neq \theta_1$, il existe un voisinage $\mathcal{V}(\theta_1)$ de θ_1 tel que presque surement*

$$\lim\sup_{N\to\infty}\ \sup_{\theta\in\mathcal{V}(\theta_1)} Q_N(\theta) < 0$$

Preuve Cf. [15].

Théorème 4.0.4 *Pour le modèle $MS-BL(p,0,P,Q)$, sous les hypothèses $H0-H5$, soit $\widehat{\theta}_N$ l'estimateur MLE sur un sous compact Θ^* vérifiant presque sûrement*

$$L_N(X_1,X_2,...,X_N,\widehat{\theta}_N) = \sup_{\theta\in\Theta^*} L_N(X_1,X_2,...,X_N,\theta)$$

alors $\widehat{\theta}_N \to \theta_0$, p.s. quand $N\to\infty$.

Preuve Supposons que $\widehat{\theta}_N$ ne converge pas vers θ_0, *p.s.* quand $N \to \infty$. Cela signifie que pour tout N (suffisamment grand), $\exists\delta > 0$ et $n > N$ tels que $\left\|\widehat{\theta}_n - \theta_0\right\| > \delta$. D'après le lemme 4.0.3 il vient que $L_n(X_1, X_2, ..., X_n, \widehat{\theta}_n) < L_n(X_1, X_2, ..., X_n, \theta_0)$. Cependant d'après la définition du MLE on a

$$L_n(X_1, X_2, ..., X_n, \widehat{\theta}_n) = \sup_{\theta\in\Theta^*} L_n(X_1, X_2, ..., X_n, \theta) \geq L_n(X_1, X_2, ..., X_n, \theta_0)$$

Ce qui contredis les résultats précédents

4.0.5 Normalité asymptotique du MLE

La preuve de la normalité asymptotique reste standard, et donc elle n'est pas surprenante, mais elle nécessite des hypothèses supplémentaires faisons intervenant le vecteur gradient ∇_θ et la matrice hessienne ∇_θ^2.

H6 La fonction $\theta \to g_\theta^*(X_t|\underline{X}_{-\infty})$ est de classe $\mathcal{C}^{(2)}$ sur $\mathcal{V}(\theta_0)$

H7 $\sup\limits_{\theta\in\Theta} \left\|\nabla_\theta g_\theta^*(X_t|\underline{X}_{-\infty})\right\| < +\infty$, $\sup\limits_{\theta\in\Theta} \left\|\nabla_\theta^2 g_\theta^*(X_t|\underline{X}_{-\infty})\right\| < +\infty$,

H8 $E\left\{\sup\limits_{\theta\in\Theta} \left\|\nabla_\theta g_\theta^*(X_t|\underline{X}_{-\infty})\right\|\right\} < +\infty$, $E\left\{\sup\limits_{\theta\in\Theta} \left\|\nabla_\theta^2 g_\theta^*(X_t|\underline{X}_{-\infty})\right\|\right\} < +\infty$

Pour montrer la normalité asymptotique du MLE, nous avons besoin à montrer deux lemme extrêmement importants..

Lemma 4.0.4 *Sous les conditions $H1 - H8$, et pour toute suite $\left(\widehat{\theta}_n\right)_{n\geq 1}$ convergente p.s. vers θ_0, alors en probabilité*

$$\frac{1}{n}\nabla_\theta^2 L_n(\widehat{\theta}_n) \to -\mathcal{I}(\theta_0)$$

où $\mathcal{I}(\theta_0)$ est la matrice d'information de Fisher.

Le second lemme concerne un théorème central limite du vecteur gradient.

Lemma 4.0.5 *Sous les conditions $H1 - H8$, on a*

$$n^{-1/2}\nabla_\theta L_n(\widehat{\theta}_n) \rightsquigarrow \mathcal{N}(0, \mathcal{I}(\theta_0))$$

Ces deux lemmes, nous permet de montrer le théorème suivant

Théorème 4.0.5 *Sous les hypothèses $H1 - H8$ nous avons*

$$\sqrt{n}\left(\widehat{\theta}_n - \theta_0\right) \rightsquigarrow \mathcal{N}\left(0, \mathcal{I}^{-1}\left(\theta_0\right)\right)$$

Preuve Cf [14] and [8]

4.1 Simulation

Nous limitons notre étude expérimentale aux modèles autorégressifs à changement de régimes markoviens. A cet effet nous simulons un ensemble de modèles d'ordres $p \in \{3, 4\}$ de taille $N \in \{500, 1000, 5000\}$ dont l'espace d'état de notre chaîne est de cardinal $s = 2$ et 4.

4.1.1 Modèle (1)

Le premier modèle est défini par

$$X_t = a_1(s_t)X_{t-1} + a_2(s_t)X_{t-2} + a_3(s_t)X_{t-3} + e_t$$

où $(e_t)_{t=1,\ldots N}$ est un processus *i.i.d.* distribuée selon une loi $N(0,1)$, $(s_t)_{t=1,\ldots N}$ est une chaîne de Markov à espace d'état $\{1, 2\}$ de matrice de transition

$$P = \begin{pmatrix} p_{11} & p_{12} \\ p_{21} & p_{22} \end{pmatrix}$$

Noton $\underline{\theta}$ le vecteur des paramètres $(a_i(s_t))_{0 \le i \le 2}$ i,e,

$$\underline{\theta} = \begin{pmatrix} a_1(1) & a_1(2) \\ a_2(1) & a_2(2) \\ a_3(1) & a_3(2) \end{pmatrix}$$

Les trajectoires de longeurs 250 associées à chaque matrice de transition P et à chaque vecteur $\underline{\theta}$ sont raportées dans les figures ci-dessous.

Les résultats de la simulation sont donnés dans le tableau suivant

N	Vraies valeurs	Valeurs estimées
500	$\underline{\theta} = \begin{pmatrix} 05 & -0.2 \\ -0.3 & 0.1 \\ -0.1 & 0.1 \end{pmatrix}$	$\widehat{\underline{\theta}} = \begin{pmatrix} 0.379 & -0.070 \\ -0.238 & 0.203 \\ -0.0917 & 0.044 \end{pmatrix}$
	$P = \begin{pmatrix} 0.9 & 0.1 \\ 0.2 & 0.8 \end{pmatrix}$	$\widehat{P} = \begin{pmatrix} 0.92 & 0.08 \\ 0.16 & 0.84 \end{pmatrix}$
1000	$\underline{\theta} = \begin{pmatrix} 05 & -0.2 \\ -0.3 & 0.1 \\ -0.1 & 0.1 \end{pmatrix}$	$\widehat{\underline{\theta}} = \begin{pmatrix} 0.445 & -0.128 \\ -0.265 & 0.068 \\ -0.130 & 0.138 \end{pmatrix}$
	$P = \begin{pmatrix} 0.9 & 0.1 \\ 0.2 & 0.8 \end{pmatrix}$	$\widehat{P} = \begin{pmatrix} 0.92 & 0.08 \\ 0.24 & 0.76 \end{pmatrix}$
5000	$\underline{\theta} = \begin{pmatrix} 05 & -0.2 \\ -0.3 & 0.1 \\ -0.1 & 0.1 \end{pmatrix}$	$\widehat{\underline{\theta}} = \begin{pmatrix} 0.505 & -0.209 \\ -0.331 & 0.093 \\ -0.006 & 0.097 \end{pmatrix}$
	$P = \begin{pmatrix} 0.9 & 0.1 \\ 0.2 & 0.8 \end{pmatrix}$	$\widehat{P} = \begin{pmatrix} 0.90 & 0.1 \\ 0.19 & 0.81 \end{pmatrix}$

Troisième partie

Appendices

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

4.2 Annexe1

4.2.1 Chaînes de Markov à temps discret

Cette annexe à pour objectifs de rappeler les définitions de base concernant les chaînes de Markov [1] à temps discret.

Définition 4.2.1 *Soit* $(\mathbf{s}_t)_{t\in\mathbb{Z}}$ *une suite de v.a* $(\Omega, \Im, P) \longrightarrow (\mathbb{S}, \mathcal{P}(\mathbb{S}))$ *est dite chaîne de Markov si* $\forall t \in \mathbb{Z}, \forall s_0,\ s_1,\ \ldots,\ s_t \in \mathbb{S}$ *on a* $P(\mathbf{s}_t = s_t \,|\mathbf{s}_{t-1} = s_{t-1}, \ldots, \mathbf{s}_0 = s_0) = P(\mathbf{s}_t = s_t \,|\mathbf{s}_{t-1} = s_{t-1})$ *(c.f.,[11]). Donc chaîne de Markov est un processus aléatoire dont le prochain état ne dépend que de l'état actuel et non sur le passé. L'ensemble* $\mathbb{S}$ *est appelé l'espace des états.*

La probabilité de transition

La probabilité de transition de l'état j à l'état i est notée p_{ij} tq : $p_{ij} = P(\mathbf{s}_t = j \,|\mathbf{s}_{t-1} = i) \forall\, i, j \in \mathbb{S}, t \in \mathbb{Z}$ et $p_{ij}^{(n)} = P(\mathbf{s}_t = j \,|\mathbf{s}_{t-n} = i)$ (c.f.,[52] et [36]) est appelée la probabilité de transition de l'état j à l'état i à n^{ieme} étape.

Homogénéité

On dira de plus que la chaîne de Markov est homogène si $p_{ij}^{(n)} = p_{ij}\ \forall\ n \geq 1$ (c.f.,[29]).

[1] Historique brève Markov : Andrei Andreevich Markov $(1856 - 1992)$ était un professeur à l'Université de Saint-Petersberg de l'ex-Union soviétique en (1886), il devint membre à part entière de l'Académie des Sciences de Saint-Petersberg depuis (1896), a commencé à étudier les chaînes simples en (1906) et a publié de nombreux des recherches sur ce sujet et après une période de la renommée a un chef de file dans ce domaine, a comparu devant la chaîne de Markov terme en (1926).

Matrice de transition

On appelle une matrice de transition tout matrice $\mathbb{P}$ tel que $\mathbb{P} = (p_{ij})_{(i,j \in \mathbb{S})}$.

Exemple 4.2.1 *L'étude statistique du phénomène de la pluie sur une ville donnée a été, convenablement, caractérisée ou décrit par les chaînes de Markov à* 2 *états :*

L'état 1 *représente les jours secs.*

L'état 2 *simule les jours pluvieux.*

Sous les 2 *hypothèses suivantes :*

Si le jour actuel est pluvieux alors la probabilité que le lendemain sera pluvieux est de $P_{22} = 0.8$.

Si le jour actuel est sec, alors la probabilité que le lendemain sera pluvieux est de $P_{12} = 0.4$.

De sorte que l'on peut écrire la matrice de transition comme suit :

$$\mathbb{P} = \begin{pmatrix} 0.6 & 0.4 \\ 0.2 & 0.8 \end{pmatrix}$$

L'analyse de cette matrice de transition donne ce quit suit :

1. $P_{11} = 0.6$ *caractérise l'état que si audjourd'hui est sec, alors le lendemain restera sec avec la probabilité de* 0.6.

2. $P_{21} = 0.2$ *représente l'état que si audjourd'hui est pluvieux, alors le lendemain sera sec avec la probabilité de* 0.2.

Théorème 4.2.1 $\forall\ n \geq 0\ \mathbb{P}^{(n)} = \mathbb{P}^n$ *la matrice* $\mathbb{P}^{(n)}$ *est appelée la matrice de transition en* n *étapes* $\mathbb{P}^{(n)} = (p_{ij}^{(n)})_{(i,j \in \mathbb{S})}$ *(c.f.,[11] et [12]).*

Remarque 4.2.1 *On a évidemment* $p_{ij}^{(1)} = p_{ij}$ *et* $\mathbb{P}^{(1)} = \mathbb{P}$ *(c.f.,[12]).Convention* $\mathbb{P}^{(0)} = I$ *c-à-d* $p_{ij}^{(0)} = \delta_{ij} = \begin{cases} 1 \text{ si } i = j \\ 0 \text{ sinon} \end{cases}$.

Equation de Chapman – Kolmogorov

Soit $\mathbb{P}^{(n)}$, $n \geq 0$ la matrice de transition de n étapes d'une chaîne de Markov alors $\forall\ l,\ m \geq 0$, $\mathbb{P}^{(l+m)} = \mathbb{P}^{(l)}\mathbb{P}^{(m)} = \mathbb{P}^{l}\mathbb{P}^{m} = \mathbb{P}^{l+m}$ (c.f.,[21]) Ce qui s'écrit sous forme développée $p_{ij}^{(l+m)} = \sum\limits_{k\in\mathbb{S}} p_{ik}^{(1)} p_{kj}^{(m)} \forall i, j \in \mathbb{S}$.

Classification des états

Communiquent

Définition 4.2.2 *L'état j est accessible depuis l'état i, s'il existe au moins un chemin de i à j $(i \longrightarrow j)$ (c.f.,[53]).*

Remarque 4.2.2 *Tout état j est accessible depuis lui – même.*

Lemma 4.2.1 *L'état j est accessible depuis l'état i ssi s'il existe $n > 0$ t.q : $p_{ij}^{(n)} > 0$ (c.f.,[68]).*

Définition 4.2.3 *Les états i et j communiquent s'ils sont accessibles l'un à partir de l'autre $i \longrightarrow j$ et $j \longrightarrow i$ (c.f.,[66]).*

Lemma 4.2.2 *Les états i et j communiquent ssi il existe $n > 0$ et $m > 0$ t.q : $p_{ij}^{(n)} > 0$ et $p_{ji}^{(m)} > 0$ (c.f.,[66]).*

Proposition 4.2.1 *Les états i est absorbant ssi $p_{ii} = 1$ (Ou a alors $p_{ij} = 0, \forall\ i \neq j$) (c.f.,[37] et [53]).*

Irréductible

Définition 4.2.4 *Une chaîne de Markov $(\mathbf{s}_t)_{t\in\mathbb{Z}}$ à espace d'état $\mathbb{S}$ et la matrice de transition $\mathbb{P}$ est dite irréductible si pour tous $i,\ j \in \mathbb{S}$ nous avons que $i \longleftrightarrow j$. Sinon la chaîne est dite réductible (c.f.,[29]).*

Proposition 4.2.2 *Une chaîne de Markov est irréductible ssi toutes ses paires d'états communiquent.*

Périodicité La période d de l'état i d'une chaîne de Markov est égale au plus grand diviseur commun de tous les n pour lesquels $p_{ii}^{(n)} > 0$ (c-à-d $d = PGCD(n \geq 1,\, p_{ii}^{(n)} > 0)$).

L'état i est périodique lorsque $d > 1$ et apériodique lorsque $d = 1$ à la convention $d = +\infty$ s'il n'existe pas de $n \geq 1$ avec $p_{ii}^{(n)} > 0$ (c.f.,[11] et [68]).

Lemma 4.2.3 *Si $p_{ii} > 0$ l'état i est apériodique.*

Distribution initiale

La distribution des états d'une chaîne de Markov après n transitions est notée $\pi(n)$ tq : $\pi(n) = P(\mathbf{s}_n = i), \forall\, i \in \mathbb{S}$ (c.f.,[11] et [36]). et la distribution initiale $\pi(0)$ est $\pi(0) = P(\mathbf{s}_0 = i)$.

Théorème 4.2.2 *Soit $\mathbb{P}$ la matrice de transition d'une chaîne de Markov et $\pi(0)$ la distribution initiale, pour tout $n \geq 1$ on a (c.f.,[29] et [40])*

$$\begin{cases} \pi(n) = \pi(0)\,\mathbb{P}^n \\ \pi(n) = \pi(n-1)\,\mathbb{P} \end{cases}.$$

4.3 Annexe2

Dans cette annexe, nous présentons plusieurs résultats utiles de théorie de la matrice, en particulier en ce qui concerne le produit Kronecker des vecteurs et des matrices, ainsi les concepts et les outilles fondamentaux qui s'avèrent nécessaires dans notre étude.

4.3.1 Produit de Kronecker

Produit de Kronecker, également connu comme un produit direct ou d'un produit tensoriel est un concept ayant son origine dans la théorie des groupes et a d'importantes applications en physique des particules. Mais la technique a été appliquée avec succès dans divers domaines de la théorie des matrices, par exemple dans la solution des équations de la matrice qui se posent lors de l'utilisation approche Lyapunov à la théorie de la stabilité. En mathématiques le produit de Kronecker est une opération portant sur les matrices, Il est ainsi dénommé en hommage au mathématicien allemand Leopold Kronecker[2].

Définition 4.3.1 *Soient $A = (a_{ij})$ une matrice de taille $(m \times n)$ et $B = (b_{ij})$ une matrice de taille $(p \times q)$. Leur produit tensoriel est la matrice de taille $(mp \times nq)$, définie par blocs successifs de taille $p \times q$, le bloc d'indice i, j valant $a_{ij}B$ En d'autres termes (c.f.,[71] et [2] et [35])*

$$A \otimes B = \begin{pmatrix} a_{11}B & \dots & a_{1n}B \\ \vdots & \ddots & \vdots \\ a_{m1}B & \dots & a_{mn}B \end{pmatrix}_{(mp \times nq)}$$

Notez que le produit de Kronecker est défini indépendamment de l'ordre des marques concernées, De ce point de vue, c'est un concept plus général que la multiplication de matrices.

[2]Leopold Kronecker (7 décembre 1823 - 29 décembre 1891) est un mathématicien et logicien allemand. Persuadé que l'arithmétique et l'analyse doivent être fondées sur les « nombres entiers », il est célèbre pour la citation suivante « Dieu fit les nombres naturels ; tout le reste est l'œuvre de l'homme. »

Les propriétés remarquables sont

Pour toutes trois matrices A, B, C, alors $(A \otimes B) \otimes C = A \otimes (B \otimes C)$ (associativité).

Pour toutes quatre matrices A, B, C, D, où A et B sont des mêmes ordres, et C, D sont des mêmes ordres et $\lambda \in \mathbb{R}$, alors

$$\begin{aligned} (A+B) \otimes (C+D) &= (A \otimes C) + (A \otimes D) + (B \otimes C) + (B \otimes D) \\ (A \otimes B)(C \otimes D) &= (AC) \otimes (BD) \text{ (Propriétés sur le produit usuel)} \\ & \left\{ \begin{array}{l} (A \otimes B) + \lambda (A \otimes C) = A \otimes (B + \lambda \cdot C) \\ (A \otimes C) + \lambda (B \otimes C) = (A + \lambda \cdot B) \otimes C \end{array} \right. \text{(Bilinéarité)} . \end{aligned}$$

Pour toutes matrices A, B, C, D,et B, C sont des mêmes ordres

$$\left\{ \begin{array}{l} A \otimes (B \pm C) = A \otimes B \pm A \otimes C \\ (B \pm C) \otimes A = B \otimes A \pm C \otimes A \end{array} \right. .$$

Le produit de Kronecker n'est pas commutatif en général $A \otimes B \neq B \otimes A$. On peut en déduire que $(A \otimes B)$ inversible si et seulement si A et B sont inversibles, auquel cas $(A \otimes B)^{-1} = A^{-1} \otimes B^{-1}$ et

donc si $\lambda_1, \lambda_2, \ldots, \lambda_n$ et $\mu_1, \mu_2, \ldots, \mu_n$ sont les valeurs propres de A et B, alors

$\{\lambda_i \cdot \mu_j, i = 1, \ldots, n \text{ et } j = 1, \ldots, m\}$ sont les valeurs propres de $(A \otimes B)$.

$$\begin{aligned} tr\,(A \otimes B) &= tr\,(A)\, tr\,(B) \\ \rho\,(A \otimes B) &= \rho\,(A)\, rg\,(B) \end{aligned}$$

où tr désigne la trace.

$$(A \otimes B)' = A' \otimes B' (Transposition).$$

Si A est $n \times n$ matrice et B est $m \times m$ matrice, alors $\det (A \otimes B) = (\det A)^n (\det B)^m$, où det désigne le déterminant.

$A \otimes B = 0$ si et seulement si $A = 0$ ou $B = 0$.

Pour toute matrice A le Kroneckerian puissance est donnée par $A^{\otimes k} = \underbrace{A \otimes \ldots \otimes A}_{k \text{ fois}}$ donc $A^{\otimes k} B^{\otimes k} = (AB)^{\otimes k}$

$\exp(I_m \otimes A) = I_m \otimes \exp(A)$ et $\exp(A \otimes I_m) = \exp(A) \otimes I_m$.

Si $A_{(m\times n)}, B_{(p\times q)}, C_{(r\times s)}, D_{(n\times k)}, E_{(q\times l)}, F_{(s\times t)}$ $(A \otimes B \otimes C)(D \otimes E \otimes F) = AD \otimes BE \otimes CF$.

Si α est un scalaire, alors $\alpha \otimes A = \alpha A = A\alpha = A \otimes \alpha$.

4.4 Annexe3

4.4.1 Normes matricielles et rayon spectral

Rappelons norme matricielle.

Soient $A, B \in \mathcal{M}_n$, ensemble des matrices réelles de format $n \times n$. Alors, l'application $\|.\| : \mathcal{M}_n \longrightarrow \mathbb{R}$ est une norme matricielle telle que $\|A\| = \max\limits_i \sum\limits_j |a_{ij}|$ et $\|B\| = \max\limits_i \sum\limits_j |b_{ij}|$ et $A = (a_{ij})$, $B = (b_{ij})$ si elle satisfait les propriétés suivantes

1) $\|A\| \geq 0, \forall A \in \mathcal{M}_n$ et $\|A\| = 0 \Leftrightarrow A = O_{(n)}$.

2) $\|\beta A\| = |\beta| \, \|A\|$ et ce pour tout scalaire β et tout A dans $\mathcal{M}_n$.

3) $\|A + B\| \leq \|A\| + \|B\|$ et ce pour tout A et B dans $\mathcal{M}_n$.

4) $\|AB\| \leq \|A\| \, \|B\|$ et ce pour tout A et B dans $\mathcal{M}_n$.

La rayon spectral $\rho(A) = \max\limits_i |\lambda_i|$ où λ_i sont les valeurs propres de A. On a $\rho(A) \leq \|A\|$.

4.4.2 Dérivation Vectorielle et Matricielle

Soient les vecteurs $\underline{X}$,$\underline{Y}$ et la matrice M, on a

1) $\frac{d\underline{Y}'\underline{X}}{d\underline{X}} = \frac{d\underline{X}'\underline{Y}}{d\underline{X}} = \underline{Y}$ où $\underline{X}$ et $\underline{Y}$ de dimensions $n \times 1$.

2)$\frac{d\underline{X}'M\underline{Y}}{dM} = \underline{XY'}$ et $\frac{d\underline{X}'M\underline{Y}}{dM'} = \underline{YX'}$ où $\underline{X}$ et $\underline{Y}$ de dimensions $n \times 1$ et m$\times$1 resp et la matrice M est de dimension $n \times m$

Bibliographie

[1] Abramson, A., Cohen, I. (2007) On the stationarity of Markov-switching GARCH processes. Econometric Theory,23, pp 485-500.

[2] Alexander, G. (1981) Kronecker Products and Matrix Calculus : with Applications, A. Graham/Ellis Horwood Ltd, pp 21-32.

[3] Andel Jiri (1976). Autoregressive series with random parameters. *Math. Operationsforsch. u. Statist. 7 735-741.*

[4] Andel Jiri (1982). On autoregressive models with random parameters. *Commun. Statist. Theory and Meth, 17-30.*

[5] Bibi, A., Aknouche, A. (2010) Stationnarité et β-mélange des processus bilinéaires généraux à changement de régime markovien. C.R. Acad. Sci. Paris, Ser. I., pp 185-188.

[6] Bibi, A., Lescheb, I. (2010) Strong consistency and asymptotic normality of least squares estimators for PGARCH and PARMA-PGARCH models. Statist. Probab. Lett. 80(19), pp 1532-1542.

[7] Bibi, A., et All, (2009) On stationarity and β-mixing of periodic bilinear processes. Statist. Probab. Lett. 79(1), pp 79-87.

[8] Bickel, P., R. Ya'acov and R. Tobias (1998). Asymptotic Normality of the maximum likelihood estimator for general hidden Markov models. The Annals of Statistics. Vol. 26, No. 4, pp. 1614-1635.

[9] Bougerol, P., Picard, N. (1992) Strict stationarity of generalized autoregressive processes. Annals of Probability 20, pp 1714-1730.

[10] Brandt, A. (1986) The stochastic Equation $Y_{n+1} = A_n Y_n + B_n$ with stationary coefficients. Adv. In Appl. Probab. 18, pp 211-220.

[11] Bremand, P. (1999) Markov Chains, Gibbs Fields, Monte Carlo Simulation, and Queues, Springer-Verlag New York, Inc, pp 52-75.

[12] Ching, W. K., Michael. K. Ng. (2006) Markov Chains : Models, Algorithms and Application, Springer Science-Business Media, Inc, pp 1-15.

[13] Cramer, H. (1961). On some classes of non-stationary stochastic processes, in Proceeding of the fourth Berkeley Symposium on Mathematical statistics and Probability. University of California Press, Berkeley and Los Angeles, Vol. 2, 57-78.

[14] Duc, R ; E. Moulines et T. Rydèn (2004). Asymptotic properties of the maximum likelihood estimator in autoregressive models with Markov regime. The Annals of Statistics. Vol. 32, No. 5, pp. 2254-2304.

[15] Francq, C., Roussignol, M. (1997) On white noises driven by hidden Markov chains. J. of Time Ser. Anal. 18, pp 553-578.

[16] Francq, C. (1999). ARMA models with bilinear innovations. Stochastic Models, 15, 29–52.

[17] Francq, C. and Zakoïan, J-M. (2000, a). Multivariate ARMA models with generalized autoregressive linear innovation. Stochastic Analysis and Applications, 18, 231-260.

[18] Francq, C. and Zakoïan, J-M. (2000, b). Estimating weak GARCH representations. Econometric theory, Vol.16, 692-728.

[19] Francq, C., Zakoïan, J. M. (2001) Stationarity of multivariate Markov-switching ARMA models. Journal of Econo-metrics, 102, pp 339-364.

[20] Francq, C., Zakoïan, J. M. (2005) $\mathbb{L}_2$-structures of standard and switching-regime GARCH models. Stoch. Proc. Appl. 115, pp 1557-1582.

[21] George Yin. G., Zhang. Q. (2005) Discrete-Time Markov Chains : Two-Time-Scale Methods and Applications, Springer Science-Business Media, Inc, pp 7-71.

[22] Goldfeld, S. M., Quandt, R. E. (1973) A Markov model for switching regressions. Journal of Econometrics., 3-16.

[23] Granger, C. W. J and Andersen, A. P (1978, a). Introduction to bilinear time series models. Vandenhoeck and Ruprecht.

[24] Granger, C. W. J and Andersen, A. P (1978, b). On the invertibility of time series models. Stoch. Proc. and their appl. 8. 87-92.

[25] Grenier, Y. (1986). Modèle ARMA à coefficients dépendant du temps : estimateurs et applications, Traitement du signal, 3, n4-5. 219-233.

[26] Guégan, D. (1981). Etude d'un modèle non linéaire, le modèle superdiagonal d'ordre 1. C. R. A. S. T. 293. Série I. 95-98.

[27] Guégan, D., Pham Dinh T. (1987, a). Minimalité et inversibilité des modèles bilinéaires à temps discret. C. R. A. S. T. 448. Série I. 159-162.

[28] Guégan, D. (1994). Séries chronologiques non linéaires à temps discret. Economica.

[29] Haggstrom, O. (2002) Finite Markov Chains and Algorithmic Applications, Cambridge University Press, pp 7-39.

[30] Hall, M., Oppenheim, A. V., Willsky, A. (1983). Time-varying parametric modelling of speech. Signal processing, vol 5, n3, 267-285.

[31] Hall, P. and Heyde, C. C. (1980). Martingale limit theory and its application. Academic Press, New York.

[32] Hallin, M. (1986). Non-Stationary q-dependent processes and time-varying moving-average models : Invertibility properties and forecasting problem. Adv. Appl. Prob. 18, 170-210.

[33] Hallin, M. (1989). Modèles non stationnaires : séries univariées et multivariées 157-196. In : Séries chronologiques. Economica.

[34] Hannan, E. J. (1970). Multiple time series. John Wiley & Sons, Inc.

[35] Jan, R.M., Heinz, N. (1988), Matrix Differential Calculus with Applications in Statistics and Econometrics, John Wiley & Sons, pp 31-35.

[36] Kemeny, J. G., Snell, J. L. (1980) Finite Markov Chains, Springer-Verlag New York Berlin Heidelberg Tokyo, pp 24-36.

[37] Kemeny, J. G., Snell, J. L., Knapp, A. W. (1976) Denumerable Markov Chains, Springer-Verlag New York Heidelberg Berlin, pp 79-98.

[38] Kristensen, D. (2009) On stationarity and ergodicity of the bilinear model with applications to the GARCH models. J. Time Ser. Anal. Vol.30(1), pp 125-144.

[39] Lee, O. (2005) Probabilistic properties of a nonlinear ARMA process with Markov-switching. Comm. Statist : Theorie and Methods, 34, pp 193-204.

[40] Levin, D. A., Peres, Y., Wilmer, E. L. Markov Chains and Mixing Times, AMS, pp 2-16

[41] Li, K. (1984). On the autocorrelation structure and identification of some bilinear time series.*J. Time ser. Anal.*, Vol.5, 173-181.

[42] Liporace, L. A. (1975). Linear estimation of non-stationary signals. J. Acoust. Soc. Amer., Vol 58, n6 1288-1295.

[43] Shu-Ing Liu. (1985). Theory of bilinear time series models. Commun. Statist. Theory Meth. 14(10), 2549-2561.

[44] Liu, J. and Brockwell, P. J. (1988). On the general bilinear time series model. J. Appl. Prob 25, 553-564.

[45] Liu, J. (1989, a). A simple condition for the existence of some stationary bilinear time series. *J. Time ser. Anal.*, Vol. 8, 33-39.

[46] Liu, J. (1989, b). On the existence of a general multiple bilinear time series. *J. Time ser. Anal.*, Vol.10, 341-355.

[47] Liu, J.(1990, a). A note on causality and invertibility of a general bilinear time series models. Adv. Appl. Prob. 22, 247-250.

[48] Liu, J. (1992, a). Spectral radius Kronecker products and stationarity. *J. Time ser. Anal.*, Vol. 13, 319-325.

[49] Liu, J. (1992, b). Higher order moments and limit theory of a general time series. Techn. Report British Columbia Univ.

[50] Liu. J. (1992, c) On Stationarity And Asymptotic Inference Of Bilinear Time Series Models, Statistica Sinica, pp 488-494.

[51] Liu, J. C. (2006) Stationarity of a Markov-switching GARCH model. Journal of Financial Econometrics, 4, pp 573-593.

[52] Major , P. (1981) Multiple Wiener-Itô integrals. Lecture Notes in mathematics 849, Springer-Verlag, New York.

[53] Martin, L. P. (1994) Markov Decision Processes, Discrete Stochastic, DynamicProgramming, A JOHN WILEY & SONS, INC., PUBLICATION, pp 587-590.

[54] Martin, C. M. (1999, a). Higher order moments of bilinear subdiagonal models with non-independent shocks. Pub. Inst. Stat. Univ. Paris, XXXXIII, fasc.1, 29-42.

[55] Mélard, G. (1985) Analyse des données chronologiques. Les presses de l'université de Montreal.

[56] Meyn, S. P., Tweedie, R. L. (2009) Markov chains and stochastic stability. Cambridge university press.

[57] Mohler, R. R. (1988). Nonlinear time series and signal processing. Lecture notes in control and information sciences N106. Berlin : Springer Verlag.

[58] Des F. Nicholls and Barry G. Quinn (1982). Random coefficients autoregressive models : An introduction. Springer-Verlag, New York Heidelberg Berlin.

[59] Pham, T. D., Tran, L. T. (1980) Quelques résultats sur les modèles bilinèaires de séries chronologiques. C. R. Acad. Sci. Paris 290, pp 235-238.

[60] Pham, D. T. and Tran, L. T. (1981). On the first order bilinear time series models. *Jour. Appl. Prob.*, 18, 617-627.

[61] Priestley, M. P. (1988). Non-linear and non-stationary time series analysis. Academic Press, London and New York.

[62] Quinn, B. G. (1982, a). Stationarity and invertibility of simple bilinear models. Stoch. Proc. and their Appl. 12. 225-230.

[63] Quinn, B. G. (1982, b). A note on the existence of strictly stationary solutions to bilinear equation. *J. Time ser. Anal.*, Vol.3, 249-252.

[64] Subba Rao, T. (1981). On the theory of bilinear time series models. J. R. S. S. Ser B43, 224-255.

[65] Subba Rao, T. and Gabr, M. M. (1984). An introduction to bispectral analysis and bilinear time series models. *Lecture Notes in statistics* $N^0 24$*. Springer Verlag.*

[66] Suhov, Y., Kelbert, M. (2008) Probability and Statistics by Example : II Markov Chains : a Primer in Random Processes and their Applications, Cambridge University Press, pp 17-18.

[67] Stelzer, R. (2009) On Markov-switching ARMA processes : Stationarity, existence of moments and geometric ergodicity. Econometric theory, 25, pp 43-62.

[68] Stroock, D. W. (2005) An introduction to Markov Processes, Springer-Verlag Berlin Heidelberg, pp 23-52.

[69] Terdik, G. (1985). Transfer functions and conditions for stationarity of bilinear models with Gaussian residuals. Proc. Roy. Soc. London A 400, 315-330.

[70] Terdik, G. (2000). Bilinear stochastic models and related problems of nonlinear time series, A frequency domain approach. Technical report N98/2. Debrecen University, Hungary.

[71] Tonu,. K., Dietrich, V. R. (2005) Advanced Multivariate Statistics With Matrices, Springer, pp 80-90.

[72] Tsay, R. S. (1986). Nonlinearity tests for time series. *Biometrika*, 73, 461-466.

[73] Yao, J. F., Attali, J. G. (2000) On stability of nonlinear AR processes with Markov-switching. Advances in Applied Probability, 32, pp 394-407.

Résumé

Dans cette étude nous examinons quelques problèmes de séries temporelles non linéaires. La classe particulière des modèles bilinéaires à changement de régimes markoviens. Les conditions nécessaires et suffisantes pour la stationnarité au second ordre, au sens stricte.

L'objectif principal de ce travail est l'estimation dans les modèles bilinéaires à changement de régimes markoviens par la méthode **MLE**, à partir des conditions on obtient des estimateurs consistants forts et asymptotiquement normaux.

Mots Cles Les modèles Non Linéaires, bilinéaire **BL**, Markov commutation, Produit de Kronecker, stationnarité, Ergodicité, **MLE**, Consistance forte, Normalité Asymptotique.

Abstract

In this memory we examine some problems of non linear time series. The particular class of bilinear models Markov-switching. The necessary and sufficient conditions for stationarity to second order in the strict sense.

The main objective of this memory is the estimation in the bilinear models Markov-switching by the **MLE**, from the conditions we obtain strong consistent estimators and asymptotically normal.

KeyWords. Nonlinear models, bilinear **BL**, Markov-switching, Kronecker product, Stationarity, Ergodicity, **MLE**, Strong consistency, Asymptotic normality.