

N° d'ordre : 03/2007-E/MT

REPUBLIQUE ALGERIENNE DEMOCRATIQUE ET POPULAIRE
MINISTERE DE L'ENSEIGNEMENT SUPERIEUR ET DE LA RECHERCHE
SCIENTIFIQUE
UNIVERSITE DES SCIENCES ET DE LA TECHNOLOGIE HOUARI BOUMEDIENE
FACULTE DE MATHEMATIQUES



Thèse

Présentée pour l'obtention du diplôme de DOCTORAT d'Etat
En : Mathématiques

Spécialité : Probabilités-Statistique
Par
GUERBYENNE Hafida

Sujet

Modèles de Séries Chronologiques Conditionnellement Hétéroscédastiques

Soutenue le 21 Mars 2007, devant le jury composé de Messieurs :

Mohamed	DJEDOUR	Professeur,	U.S.T.H.B.,	Alger	Président
Mohamed	BENTARZI	Professeur,	U.S.T.H.B.,	Alger	Directeur de thèse
Amar	AISSANI	Professeur,	U.S.T.H.B.,	Alger	Examineur
Hocine	FELLAG	Professeur,	U.M.M.T.O.,	Tizi-Ouzou	Examineur
Abdelouahab	BIBI	M.C.,	U.T.M.,	Constantine	Examineur
Mohamed	IBAZIZEN	M.C.,	U.M.M.T.O.,	Tizi-Ouzou	Examineur
Khaled	KHALDI	M.C.,	U.M.B.,	Boumerdes	Examineur

Modèles de Séries Chronologiques Conditionnellement Hétéroscédastiques

Hafida Guerbyenne
Département de Probabilités et Statistique
Faculté de Mathématiques,
Université des Sciences et de la Technologie Houari Boumediene,
U. S. T. H. B.

Table des matières

Introduction	i
0.1 Apport et présentation de la thèse	iii
1 Modèles conditionnellement hétéroscédastiques	1
1.1 Introduction	1
1.2 Modèles ARCH et leurs extensions	2
1.2.1 Modèle ARCH	2
1.3 Stationnarité	5
1.4 Equation de Yule Walker pour le carré d'un processus <i>GARCH</i>	7
1.5 Modèle IGARCH	8
1.6 Modèle $\beta - ARCH$	8
1.7 Etude statistique du modèle <i>ARCH</i>	9
1.7.1 Estimation	9
1.7.2 Cas du modèle de régression avec erreurs <i>ARCH</i>	10
1.7.3 Procédure d'estimation en deux étapes	11
1.7.4 Identification	12
2 Estimation récursive en-ligne des modèles ARCH	13
2.1 Introduction	13
2.2 Modèle et hypothèses	14
2.3 Algorithme <i>RLS</i> en deux étapes (<i>2S-RLS</i>)	15

2.4	Propriétés asymptotiques des estimateurs 2S-RLS	21
2.5	Etude de simulation	25
2.6	Conclusion	27
2.7	Appendice	27
3	Estimation récursive en-ligne des modèles <i>GARCH</i>	28
3.1	Introduction	28
3.2	Pseudo régression linéaire récursive pour des modèles GARCH	30
3.3	Algorithme d'erreur de prédiction récursive en deux étapes	35
3.4	Simulation	39
3.5	Appendice : Preuves des théorèmes	40
3.5.1	Preuve du théorème 3.2.1	40
3.5.2	Preuve du Théorème 3.3.2	59
3.5.3	Quelques propriétés du lemme d'inversion matricelle	62
3.6	Conclusion	63
4	Modèles autorégressifs doubles périodiques	64
4.1	Introduction	64
4.2	Definitions et hypothèses	65
4.2.1	Représentation Markovienne du modèle	67
4.3	Stationnarité périodique stricte	68
4.4	Existence des moments d'ordres supérieurs	75
4.5	Conclusion	77
5	Estimation Adaptative des Modèles AR-périodiques	78
5.1	Introduction	78
5.2	Quelques éléments de la théorie asymptotique des expériences statistiques	81

5.2.1	Normalité locale asymptotique	81
5.2.2	Contiguïté	84
5.2.3	Les lemmes de Le Cam	84
5.3	Normalité Locale Asymptotique des modèles autorégressifs d'ordre p_t , d -périodiques	85
5.3.1	Définitions et formulation.	85
5.4	Résultats préliminaires : le modèle et le modèle multivarié associé	87
5.5	Lemme de Swensen (1985) et les conditions suffisantes de la propriété LAN .	90
5.6	Propriété LAN pour un modèle $AR(p_t)$ évolutif d -périodique	99
5.6.1	Normalité locale asymptotique	99
5.6.2	Distributions asymptotiques locales de $\Delta_f^{(n)}(\underline{\theta})$ et de $\Lambda_{(\underline{\theta} + \nu^{(n)} \underline{\tau}^{(n)})/\underline{\theta}}^{(n)}$ sous $H_f^{(n)}(\underline{\theta}^{(n)})$	100
5.7	Les Familles LAN	100
5.8	Existence et construction d'estimateurs LAM	101
5.8.1	Borne inférieure d'un estimateur dans les modèles LAN	101
5.8.2	Le problème de l'alignement	102
5.8.3	Propriété de linéarité asymptotique de $\Delta_f^{(n)}(\underline{\theta})$ sous $H_f^{(n)}(\underline{\theta}^{(n)})$. . .	102
5.9	Construction d'estimateurs adaptatifs	105
5.10	Résultats numériques	107

Introduction

Depuis les travaux de Wold (1938), l'intérêt pour le développement de modèles de séries chronologiques, pouvant répondre aux besoins de l'utilisateur, et par là, celui de techniques statistiques devant répondre aux problèmes statistiques inévitablement rencontrés est allé crescendo. Les modèles de séries chronologiques linéaires à coefficients constants ont connu une ère de prospérité grâce, en particulier, aux travaux de Box et Jenkins (1970) et leur fameux ouvrage qui les a popularisés avec, en particulier, le fameux triptyque : identification, estimation, validation. Ces modèles, qui supposent une variance des erreurs constante, ont vite montré leurs limites, en particulier, dans la modélisation des séries chronologiques macroéconomiques et financières où la focalisation sur les premiers moments conditionnels- les moments d'ordre supérieurs étant traités comme des paramètres de nuisance- supposés constants par rapport au temps, s'est révélée limitative mais aussi dans les séries chronologiques présentant des dynamiques périodiques qui ne peuvent pas être prises en charge par les modèles *SARIMA* classiques (cf. Bentarzi ((1995)et bibliographie intra), Bentarzi (1998), Bentarzi et Hallin (1994), (1996), (1998)). De plus, l'importance croissante motivée par les considérations sur le risque et sur l'incertitude dans la théorie économique moderne ont nécessité le développement de nouvelles techniques pour les séries chronologiques économétriques permettant à la variance et à la covariance de dépendre du temps. Ainsi est née, sous l'impulsion du génie d'Engle (1982), la classe des modèles *ARCH* (autorégressifs conditionnellement hétéroscédastiques) suggérés afin de saisir les caractéristiques particulières des séries de données d'observations financières. Parallèlement aux succès des modèles de séries chronologiques linéaires standards fondés sur l'utilisation de la moyenne conditionnelle contre l'utilisation de la moyenne inconditionnelle qui prévalait, la clé offerte par les modèles *ARCH* réside dans la distinction faite entre les moments du second ordre conditionnels et inconditionnels (marginaux). Alors que les covariances marginales des variables d'intérêt

peuvent être invariantes par rapport au temps, les variances et les covariances conditionnelles dépendent souvent et de façon non triviale, des états du passé du processus. Comprendre la nature exacte de cette dépendance temporelle est crucialement important car la perte en efficacité, si l'hétéroscédasticité sous jacente est négligée, peut se révéler importante, en particulier dans l'évaluation de prévisions.

De nombreux processus sont mieux modélisés par des modèles non linéaires plutôt que par une représentation linéaire. L'approche par l'hétéroscédasticité conditionnelle peut se targuer de posséder, maintenant, des outils puissants d'analyse et de modélisation fondés sur des bases théoriques solides pour modéliser des séries chronologiques stationnaires présentant une dynamique non linéaire. Depuis l'article fondateur d'Engle (1982), les modèles dits de type *ARCH* ont connu un développement époustoufflant qui est le reflet de leur grande importance non seulement pour modéliser des séries chronologiques financières exhibant une volatilité instantanée, -comme c'est à l'origine-, reflétant la forte dépendance de la variabilité instantanée d'une série chronologique par rapport à son propre passé, mais aussi dans divers autres domaines de recherche tous aussi différents, tels que, par exemple, l'environnement (Bordignon (2000), Fasso (2000), les données sur le climat, (Tol, (1996)), etc...

De plus, les modèles dits de type *ARCH* ne sont pas enclavés dans un monde à parois rigides puisque des ponts sont jetés entre ces modèles et d'autres modèles susceptibles de refléter d'autres dynamiques non linéaires, à savoir, les modèles dits bilinéaires (Granger et Andersen (1978)), (cf. Tong (1990)), ou des modèles autorégressifs à coefficients aléatoires (*RCA* models), (Nicholls et Quinn, (1982), Tsay, (1987), Bera *et al* (1996))

Ceci a, bien entendu, été déterminant dans le développement de ces approches. (les modèles *GARCH* ou Modèles *ARCH* généralisés, Bollerslev, 1986), modèles à mémoire longue (Ding and Granger, 1993 ; Baillie *et al*, 1996) dépendances simultanées (*ARCH* multivarié, Ling and McAleer, 2003). Bien entendu, de nombreuses difficultés sont rencontrées dans la recherche d'extensions consistantes dans le cas non stationnaire (cf, Dahlhauss et Subba Rao, 2006). Une classe importante de modèles non stationnaires est celle des modèles périodiques. Bentarzi 1995, 1998, et bibliographie *intra*, Bentarzi et Hallin 1994, 1996, 1998, Bollerslev and Ghysels, 1996 ; Franses and Paap, 2000 Bibi et Francq, 2003).

Un autre volet non moins important est celui de l'estimation dite *adaptive* qui a eu

pour point de départ les travaux de Stein (1956). Stein a donné une condition suffisante d'adaptativité fondée sur les espaces tangents. Stone (1975) a déterminé un estimateur adaptatif pour la médiane d'une fonction densité symétrique. Pfanzagl (1976) a montré la non existence d'estimateurs adaptatifs pour un quantile dans le cas de densités asymétriques. Bickel (1982), toujours dans le contexte d'une suite de variables aléatoires indépendantes mais aussi pour un modèles de régression a montré que la condition donnée par Stein (1956) est également nécessaire et que, donc, il existe un estimateur adaptatif pour le paramètre pour la classe des lois de probabilité absolument continues et symétriques ; Linton (1993) a, dans le cadre des modèles *ARCH* démontré la propriété de normalité locale asymptotique LAN (*local asymptotic normality*) et déterminé un estimateur adaptatif pour le paramètre vectoriel sous-jacent après avoir reparamétrisé son modèle pour éviter le problème de non identifiabilité du modèle. Dong Wan Shin and Beong Soo So (1999) ont développé des tests de racine unité fondés sur l'estimation adaptative, Ling (2003) s'est intéressé à l'estimation adaptative dans les modèles de mémoire longue. L'estimation adaptative trouve sa source dans la propriété de normalité locale asymptotique d'une suite de logarithmes de rapports de vraisemblance (Le Cam (1960, 1986), Swensen (1985), Le Cam et Yang (1990), Fabian et Hannan (1982), Kreiss (1987 a,b), Bentarzi et Hallin (1996)) qui est fondée sur le concept de contiguïté de deux suites de mesures de probabilité. Le cas d'un modèle autorégressif évolutif périodique dont la variance de l'innovation est périodique et de densité non spécifiée ne semble pas avoir été envisagé et fera l'objet du chapitre cinq

Les modèles de rupture font partie des modèles non linéaires et/ou hétéroscédastiques (Broemeling, 1990 et bibliographie *intra* Guerbyenne et Kezim 1994 etc... dans le cas d'une rupture simple, Stephen (1994), Inclan (1992), Benayache et Guerbyenne (2005) dans le cas de ruptures multiples etc...), mais ils ne seront pas considérés dans ce travail.

0.1 Apport et présentation de la thèse

Chapitre 1 : Modèles conditionnellement hétéroscédastiques

Dans ce chapitre, nous présentons quelques modèles non linéaires *ARCH*, *GARCH*, β -*ARCH*, *ARCH* - *M*. Les conditions de stationnarité stricte et au second ordre sont rappelées

notamment les résultats de Nelson (1990) et Bougerol et Picard (1992) fondés sur le plus grand exposant de Lyapunov.

Les équations de Yule-Walker pour un processus $GARCH(q, p)$ sont rappelées établissant ainsi une analogie claire entre ces processus et les processus $ARMA(p, q)$ classiques.

L'inférence statistique des modèles $ARCH$ a été le centre d'intérêt de nombreux chercheurs où l'on rencontre essentiellement trois préoccupations différentes en ce qui concerne l'estimation des paramètres :

- méthode d'estimation fondée sur la méthode classique des moindres carrés : Engle (1982), Boose et Mukherjee (2003) qui ont utilisé une méthode en deux étapes.
- méthode fondée sur le pseudo-maximum de vraisemblance : Weiss (1984), Francq et Zakoian (2004) etc...
- méthode LAN fournissant des estimateurs adaptatifs au sens de la définition de Stein (1956) : Linton (1993), pour un modèle $ARCH$, Klaassen (1995) pour un $GARCH(1, 1)$ etc...

Toutes ces méthodes ne prennent pas en considération le caractère éventuellement nouvellement en ligne de données d'observations. Dans les chapitres 2 et 3, nous proposons des algorithmes récurrents en ligne qui permettent de répondre à ces objectifs et qui fournissent des estimateurs possédant de bonnes propriétés asymptotiques.

Chapitre 2 : Estimation récurrente en-ligne des modèles ARCH

Les modèles $ARCH$ introduits par Engel (1982) et leurs extensions trouvent des applications importantes dans une diversité de domaines où les séries chronologiques rencontrées présentent une dynamique de haute fréquence, en particulier, dans le monde financier (cours boursiers, taux de change,...), les domaines du contrôle adaptatif et du traitement de signal.

En effet, dans ces genres de problèmes et dans d'autres études intéressantes théoriquement et pratiquement, l'observation des phénomènes sous études donne lieu à des données progressivement disponibles (*en-ligne*) à des intervalles de temps très courts. Ainsi, l'emploi de méthodes classiques d'estimation (maximum de vraisemblance, moindres carrés) (cf.

Weiss, 1986 ; Lumsdaine, 1996 ; Lee et Hansen, 1994 ; Li et *al.* 2002 ; Berkes et *al.* 2003 ; Li et McAleer, 2003 ; Francq et Zakoïan, 2004) qui supportent des observations de taille fixe (*hors-ligne*) peut s'avérer trop lourd. De plus, il est bien connu, de nos jours que les modèles *ARCH* sont de bonnes approximations pour certains modèles de séries chronologiques en temps continu (voir Nelson 1990 ; Nelson et Foster, 1994 ; Drost et Werker, 1996).

Toutes ces raisons, entre autres, fournissent donc une grande motivation pour la recherche de méthodes d'estimation qui prennent en charge l'obtention des estimations de façon récursive sans avoir recours au calcul par bloc dès la présentation d'une nouvelle observation. Ceci fait réduire considérablement le temps d'exécution et l'espace mémoire.

Visant ces objectifs, le présent chapitre est consacré à la mise en œuvre d'un algorithme récurrent pour l'estimation, en deux étapes, des modèles *ARCH*. Cette méthode, inspirée des algorithmes bien connus d'estimation en-ligne (par exemple Kushner et Clark, 1978 ; Ljung et Söderström, 1983 ; Goodwin et Sin, 1984 ; Young, 1984 ; Benveniste et al. , 1990 ; Duflo, 1990 ; Solo et Kong 1995 ; Kushner et Yin, 1997 ; Ljung, 1999 ; Pollock, 2003), fournit des estimateurs asymptotiquement gaussiens et permet de prendre en charge le caractère nouvellement en-ligne des données économétriques. L'algorithme récursif en deux étapes, nommé *2S-RLS* (Two stage Recursive Least Squares)(*2S-RLS*), donné par la proposition **Algorithme 2.3.1**, que nous établissons, étend la méthode des moindres carrés en deux étapes proposée par Bose et Mukherjee (2003) dans le cas d'un échantillon de taille fixée (voir aussi Pantula (1988) pour le cas particulier *ARCH* d'ordre un) au cas d'observations progressives dans le temps. Un autre avantage important de notre algorithme sur celui de Bose et Mukherjee (2003) est que dans ce dernier la variance conditionnelle à l'instant k , dans le modèle de base, a été remplacée pour k par une seule constante (estimée) ; par contre dans le nôtre, cette variance conditionnelle a été estimée récursivement pour chaque k . Ceci a contribué sensiblement à la performance de notre algorithme. La distribution asymptotique des estimateurs obtenus (*2S-RLS*) est donnée par le **Théorème 2.3.1**. De plus, nous avons montré leurs convergences presque sûre vers les vraies valeurs des paramètres.

La performance de notre algorithme (*2S-RLS*) est assistée par une étude de simulation intensive. L'étude a montré d'une part, que le présent algorithme est plus performant que celui de Bose et Mukherjee (2003) (noté ici *B&M*) et, d'autre part, qu'il donne de bons

résultats, au sens de $RMSE$, même pour de petites tailles d'échantillons. Concernant la procédure de contrôle, nous avons appliqué la méthode de Ljung et Söderström (1983) qui teste l'admissibilité de l'estimation courante et la projette en cas de rejet dans la région de positivité. La procédure de projection consiste en une réduction successive du terme de mise à jour en le multipliant par une constante entre 0 et 1. Des extensions de cet algorithme à des modèles plus généraux tels que les modèles $GARCH$ peuvent être développées à l'aide de l'approche pseudo-régression linéaire (Solo, 1979 et Duflo, 1990) que nous abordons dans le prochain chapitre consacré à l'estimation récursive des modèles $GARCH$.

Chapitre 3 : Estimation Récursive des Modèles $GARCH$

Le succès remarquable qu'a connu la classe des modèles $ARCH$, introduite par Engle (1982) en la gestion, par le biais de la variance conditionnelle stochastique, de la volatilité instantanée qui caractérise, en particulier, les séries financières (taux de change, taux d'inflation, indices boursiers,...) a, très rapidement, motivé l'étude théorique et pratique qui a donné, pour la raison du principe de parcimonie et par analogie aux cas des modèles linéaires, lieu à la classe des modèles $GARCH$ (Bollerslev, 1986) et ses diverses extensions ($EGARCH$, $TARCH$, $FIGARCH$, ...). Les méthodes d'estimation existantes dans la littérature, pour de tels modèles sont fondées sur une taille d'échantillon préalablement fixée. Ainsi, les mêmes raisons mentionnées dans le **Chapitre 2**, nous ont fourni la motivation pour le développement de quelques algorithmes récurrents (en-ligne), basés sur le schéma en deux étapes pour l'estimation d'un modèle de cette classe. D'abord, nous établissons, en exploitant l'approche pseudo régression linéaire PLR , un algorithme baptisé $2S - PLR$ (Two stage Pseudo Linear Regression), donné dans le présent document par la proposition **Algorithme 3.2.1** (Aknouche et Guerbyenne, 2006b), qui est une généralisation de l'algorithme ($2S - RLS$) développé au chapitre précédent pour des modèles $ARCH$, au cas de modèles $GARCH$. La convergence presque sûre des estimateurs fournis par cet algorithme est établie, sous quelques conditions, par le **Théorème 3.2.2**, (Aknouche et Guerbyenne, 2006b). Le point faible de l'algorithme $2S - PLR$ réside dans le fait qu'il produit des estimateurs non efficaces bien qu'ils soient convergents presque sûrement. Pour remédier à ce problème,

nous établissons, en adoptant l'approche de l'erreur de prédiction récursive (Recursive Predictor Error Method, *RPEM*), un autre algorithme qui fournit des estimateurs convergeant presque sûrement et asymptotiquement gaussiens. L'algorithme en deux étapes ainsi obtenu baptisé *2S-RML* (*Two-Stage Recursive Maximum Likelihood*) est donnée par la proposition **Algorithme** 3.3.1 (Aknouche et Guerbyenne, 2006b). D'autre part, la normalité asymptotique des estimateurs fournis par cet algorithme est, sous quelques conditions, assurée par le **Théorème** 3.3.2 (Aknouche et Guerbyenne, 2006b). La performance des algorithmes *2S-PLR* et *2S-RML*, a fait l'objet d'une étude de simulation basée sur des données simulées, générées à partir d'un modèle *GARCH*(1, 1) gaussien.

Chapitre 4 : Modèles autorégressifs doubles périodiques

L'objectif principal de ce chapitre est l'étude des propriétés probabilistes de la classe des modèles autorégressifs S -périodiques (d'ordre p constant) à erreurs $ARCH(q)$ périodiques que l'on note $PAR - PARCH_S(p, q)$. Nous nous proposons, d'abord, de donner une nouvelle formulation du modèle $PAR - PARCH_S(p, q)$ dans laquelle l'équation d'innovation est à coefficients aléatoires périodiquement distribués, dans le but de décrire de façon parcimonieuse cette caractéristique de la variance conditionnelle d'être dépendante du temps. Le modèle proposé que nous nommerons *AR périodique double* (*DPAR*) est, d'un point de vue du second ordre, équivalent au modèle $PAR - PARCH$ obtenu dans Franses et Paap (2000), en ce sens que, les deux modèles ont les mêmes espérances et variances conditionnelles. La simplicité du modèle proposé nous permet d'étudier aisément la propriété de stationnarité périodique au sens strict et l'existence des moments d'ordres supérieurs. Le chapitre se compose de cinq sections. Dans la seconde section, nous donnons une autre formulation du modèle $PAR - PARCH_S(p, q)$ sous la forme d'un modèle autorégressif périodique double à coefficients aléatoires moyennant certaines hypothèses. Dans la section suivante nous donnons une représentation markovienne périodique à espace d'état d'un modèle $PAR - PARCH_S(p, q)$. Dans la section quatre, nous utiliserons cette représentation markovienne, à partir de laquelle nous obtiendrons des conditions nécessaires et suffisantes de stationnarité périodique stricte du modèle PAR double. Dans la section cinq sera donnée une condition suffisante pour qu'un modèle $PAR - PARCH_S(p, q)$ possède des moments

d'ordres supérieurs finis. Une conclusion fera l'objet de la section six où seront également soulignés quelques problèmes ouverts

Chapitre 5 : Estimation Adaptative des Modèles AR-périodiques

Le présent chapitre est consacré principalement à la présentation des outils de base notamment la propriété de contiguïté d'une suite de lois de probabilité, concept introduit par Le Cam (1960, 1986) et repris par Roussas dans sa monographie (1972), et le concept de normalité asymptotique locale pour une suite de variables aléatoires appelées, dans ce contexte, suites centrales, qui sont utilisés et qui feront l'objet de la section 2. La section 2 concerne les résultats de Hájek (1972) modifiés par Fabian et Hanan (1982). Dans la section 3, en suivant l'approche de Kreiss (1987 a,b) et, en utilisant la propriété de normalité locale asymptotique, LAN (*local asymptotic normality*) pour un modèle autorégressif périodique — prouvée dans Bentarzi et Hallin (1996) sur la base du lemme de Swensen (1985) — adapté à notre contexte où l'on suppose que la variance de l'innovation est périodique et inconnue et que le modèle sous l'hypothèse nulle est périodique— on détermine des estimateurs adaptatifs pour les paramètres d'un modèle autorégressif périodique. La performance des estimateurs adaptatifs obtenus est comparée à celle des estimateurs classiques des moindres carrés (MCO) via une étude de simulation intensive. La densité de l'innovation est supposé vérifier certaines conditions de régularité. Le critère de comparaison utilisé est le RMSE (Root Mean Square Error).

Chapitre 1

Modèles conditionnellement hétéroscédastiques

1.1 Introduction

Le concept de variance conditionnelle a commencé à jouer un grand rôle au début des années quatre vingt avec l'article fondateur de Engle (1982). Il caractérise les modèles venus élargir la classe des modèles classiques fondés essentiellement sur une structure de dépendance linéaire entre une variable à un instant t et ses valeurs passées et celles d'un bruit blanc et de ses valeurs passées. Les formulations de type *ARMA* sont quasiment centrées sur la structure d'autocovariance des processus. Or, de nombreuses séries, économiques, en particulier, celles des rendements ne diffèrent pratiquement pas des bruits blancs. En revanche, les séries de carrés ou des valeurs absolues sont souvent fortement autocorrélées. Ces deux propriétés ne sont pas incompatibles mais montre que le bruit blanc n'est pas indépendant.

De plus les grandes valeurs (en valeur absolue) des données d'observation tendent à être suivies de grandes valeurs, et les petites valeurs, de petites. On dit que le marché est fortement volatil ou faiblement volatil.

Malgré ce phénomène, le processus peut-être stationnaire et donc, en particulier, homoscedastique (de variance marginale constante). Seulement, puisqu'une forte valeur de la donnée d'observation, au temps $(t - 1)$ tend à augmenter la probabilité d'observer une forte valeur (en valeur absolue) au temps t , la variance de la variable au temps t conditionnellement à ses valeurs passées (appelée volatilité) ne semble pas constante.

L'hétéroscédasticité conditionnelle n'est incompatible ni avec l'homoscédasticité marginale, ni avec la stationnarité.

Il est important de noter que, lorsqu'on considère les distributions de fréquence de séries de rendements, de variations de prix ou du logarithme de ces variations de prix, on remarque qu'elles ne correspondent pas à une distribution gaussienne. Elles sont à queues épaisses (à décroissance plus lente que $\exp\left\{\frac{-x^2}{2}\right\}$) et présentent un pic en zéro : elle sont dites leptokurtiques ; leur coefficient de kurtosis est nettement supérieur à 3. Les modèles hétéroscédastiques se sont révélés particulièrement adaptés à la prise en compte de ces caractéristiques.

1.2 Modèles ARCH et leurs extensions

Une profusion de modèles animés généralement par le souci d'ajuster au mieux une série de données d'observations ont vu le jour depuis l'article fondateur d'Engle (1982).

1.2.1 Modèle ARCH

Un processus $\{u_t, t \in \mathbb{Z}\}$ est dit cadrer avec un modèle $ARCH(q)$ s'il est solution de l'équation stochastique

$$u_t = \sqrt{h_t} \eta_t, t \in \mathbb{Z}, \quad (1.2.1)$$

$$h_t = \alpha_0 + \sum_{i=1}^q \alpha_i u_{t-i}^2, \quad (1.2.2)$$

où $\{\eta_t, t \in \mathbb{Z}\}$ est une suite de variable aléatoire i.i.d centrée réduite et h_t est la variance conditionnelle du processus, appelée volatilité.

Si l'on note $\mathcal{F}_{t-1}$ la tribu engendrée par les valeurs passées de u_s jusqu'au temps $t - 1$, les deux premiers moments conditionnels de u_t découlent simplement de la définition

$$\begin{aligned} E(u_t / \mathcal{F}_{t-1}) &= 0, \\ \text{Var}(u_t / \mathcal{F}_{t-1}) &= h_t. \end{aligned}$$

La première condition reflète le fait que u_t est une différence de martingale. Ainsi $\{u_t, t \in \mathbb{Z}\}$ est un processus centré non corrélé, mais sa variance conditionnelle peut évoluer avec le temps et sa variance marginale peut ou non exister.

Le plus souvent, u_t n'est pas (supposé directement) observable mais apparaît plutôt soit comme l'innovation dans un modèle de régression

$$y_t = X_t' \beta + u_t, \tag{1.2.3}$$

soit comme l'innovation d'un processus $\{y_t, t \in \mathbb{Z}\}$ dans un modèle *ARMA*, écrit sous forme polynomiale

$$A(L)y_t = B(L)u_t. \tag{1.2.4}$$

La classe de modèle ainsi définie est très large et l'hypothèse (1.2.1) très générale. La première famille de modèles repose sur une paramétrisation quadratique de la variance conditionnelle.

Définition 1.2.1 On dit que le processus $\{u_t, t \in \mathbb{Z}\}$, suit un modèle *ARCH*(q) s'il est défini par l'équation (1.2.1) et une paramétrisation affine de valeurs passées du carré du bruit

$$h_t = \alpha_0 + \sum_{i=1}^q \alpha_i u_{t-i}^2 \tag{1.2.5}$$

avec les conditions $\alpha_0 > 0, \alpha_i \geq 0, i = 1, \dots, q$ pour garantir la stricte positivité de la variance conditionnelle. Ce modèle a été introduit par Engle (1982) et étudié dans le cas où η_t est un bruit blanc gaussien, centré de variance unité.

Une généralisation naturelle *GARCH*(p, q) (Bollerslev, 1986) analogue à celle des *AR*(p) par les *ARMA*(p, q), permet de prendre en compte des processus à mémoires longues avec

un nombre restreint de paramètres

$$h_t = \alpha_0 + \sum_{i=1}^q \alpha_i u_{t-i}^2 + \sum_{j=1}^p \beta_j h_{t-j} \quad (1.2.6)$$

avec $\alpha_0 > 0, \alpha_i \geq 0, i = 1, \dots, q$ et $\beta_j \geq 0, j = 1, \dots, p$.

Les modèles *ARCH* se présentent comme une extension possible de bruit qui interviennent dans les différents modèles (modèles de régression, modèles *ARMA*, ...). Il existe d'autres extensions à l'étude des modèles *ARCH* purs.

Le bruit *ARCH* a conduit à un bruit plus sophistiqué :

- Les modèles *ARCH-M* (Engle, Lilien et Robins, 1987), pour lesquels le modèle s'écrit comme suit

$$y_t = X_t' \beta + \delta h_t + u_t \quad (1.2.7)$$

$$u_t = \sqrt{h_t} \eta_t, \quad (1.2.8)$$

$$h_t = \alpha_0 + \sum_{i=1}^q \alpha_i \varepsilon_{t-i}^2 \quad (1.2.9)$$

où $\{\eta_t, t \in \mathbb{Z}\}$ est une suite de variable aléatoire *i.i.d* centrée réduite.

- Les modèles *β-ARCH* (Diebold et Guegan, 1992).

- Une utilisation des modèles de régression ou des modèles *ARMA* avec ces différents bruits, par exemple, les modèles *ARMA* avec bruit *ARCH* (Weiss, 1984).

Supposons que l'on envisage pour le même processus d'être caractérisé par divers schémas récurrents et pour chacun considérons leurs variance et / ou leurs variances conditionnelles :

1) Le processus $\{u_t, t \in \mathbb{Z}\}$ est un processus *AR*(1) défini par :

$$u_t = a u_{t-1} + \varepsilon_t \text{ où } \varepsilon_t \sim \mathcal{N}(0, \sigma^2),$$

alors

$$\text{Var}(u_t) = \frac{\sigma^2}{1-a^2} \text{ et } \text{Var}(u_t/u_{t-1}) = \sigma^2.$$

2) Le processus $\{u_t, t \in \mathbb{Z}\}$ suit un modèle bilinéaire markovien

$$u_t = \beta u_{t-1} \varepsilon_t \text{ où } \varepsilon_t \sim \mathcal{N}(0, \sigma^2), \quad (1.2.10)$$

alors

$$\text{Var}(u_t) = 0 \text{ ou } \infty \text{ et } \text{Var}(u_t/u_{t-1}) = \beta^2 \sigma^2 u_{t-1}^2.$$

3) Le processus $\{u_t, t \in \mathbb{Z}\}$ suit un modèle $ARCH(1)$ défini par (1.2.1), alors

$$\text{Var}(u_t/u_{t-1}) = \alpha_0 + \alpha_1 u_{t-1}^2.$$

A partir de ces exemples, on constate que :

- 1) Dans le cas d'un processus $AR(1)$, la variance et la variance conditionnelle ne dépendent pas du temps.
- 2) Dans le cas du modèle bilinéaire (1.2.10) la variance conditionnelle du processus dépend du passé du processus, mais la variance est soit nulle soit infinie.

Du point de vue statistique, ces modèles constituent une classe spécifique de modèles non linéaires (classe qui contient également les modèles bilinéaires introduits par Granger et Andersen (1978), les modèles autorégressifs à seuils (TAR) introduits par Tong (1978),... etc). Ces modèles ont fait l'objet d'extensions au cas où le processus sous jacent est multivarié :

Engle, Granger et Kraft (1984), Bollerslev, Engle et Woolridge (1988), Engle et Rodrigues (1989), Ling et Deng (1993), Engle et Kroner (1995), Wong et Li (1997), Ling et McAleer (2003) parmi d'autres.

1.3 Stationnarité

Le processus $GARCH(p, q)$ défini par (1.2.1) et (1.2.6) est dit faiblement stationnaire ou stationnaire au second ordre si (Bollerslev, 1986)

$$\sum_{i=1}^q \alpha_i + \sum_{j=1}^p \beta_j < 1. \quad (1.3.1)$$

Une autre représentation de la variance conditionnelle est

$$h_t = \sum_{i=1}^{\tau} A_i u_{t-i} h_{t-i}, \tau = \max(p, q), \quad (1.3.2)$$

où

$$A_i(u_{t-i}) = \alpha_i u_{t-i}^2 + \beta_i, i = 1, \dots, \tau. \quad (1.3.3)$$

Cette représentation permet d'obtenir des propriétés de stationnarité stricte du processus $\{u_t, t \in \mathbb{Z}\}$. Le premier, Nelson (1990) a posé le problème de l'existence d'une solution strictement stationnaire et l'a résolu dans le cas d'un $GARCH(1, 1)$. La condition (nécessaire et suffisante) s'écrit

$$E[\log A(u_t)] = E[\log(\alpha_1 u_{t-1}^2 + \beta_1)] < 0. \quad (1.3.4)$$

Le problème a été résolu pour le cas $GARCH(p, q)$ par Bougerol et Picard (1992). Avant de donner le résultat concernant la stationnarité stricte du modèle $GARCH$, nous définissons le plus grand exposant de Lyapunov associé à une suite de matrices.

Définition 1.3.1 Le plus grand exposant de Lyapunov γ associé à une suite de matrices $\{A_n, n \in \mathbb{N}\}$ est défini par :

$$\gamma = \inf \left\{ E \left(\frac{1}{n+1} \log \|A_0 A_{-1} \cdots A_{-n}\| \right), n \in \mathbb{N} \right\}, \quad (1.3.5)$$

où $\|A\| = \sup \left\{ \frac{\|Ax\|}{\|x\|}, x \in \mathbb{R}^k, x \neq 0_{\mathbb{R}^k} \right\}$.

Proposition 1.3.2 (Bougerol et Picard, 1992) Le processus $GARCH(p, q)$ défini par (1.2.1) et (1.2.6) est strictement stationnaire si et seulement si le plus grand exposant de Lyapunov γ associé à la suite de matrices $\{A_n, n \in \mathbb{N}\}$ définie par :

$$A_n = \begin{pmatrix} \zeta_n & \beta_p & \underline{\alpha} & \alpha_q \\ I_{p-1} & 0 & 0 & 0 \\ \xi_n & 0 & 0 & 0 \\ 0 & 0 & I_{q-2} & 0 \end{pmatrix}, \quad (1.3.6)$$

où $\zeta_n = (\alpha_1 u_n^2 + \beta_1, \beta_2, \dots, \beta_{p-1})$, $\xi_n = (u_n^2, 0, \dots, 0)$ et $\underline{\alpha} = (\alpha_2, \dots, \alpha_{q-1})$ est strictement négatif.

Dans le cas particulier du modèle $GARCH(1, 1)$, la condition précédente devient

$$E [\log (\alpha_1 u_n^2 + \beta_1)] < 0,$$

et ce résultat a été obtenu par Nelson (1990).

1.4 Equation de Yule Walker pour le carré d'un processus $GARCH$

L'étude de la fonction d'autocorrélation et d'autocorrélation partielle du modèle $GARCH(p, q)$ donne des renseignements intéressants quant à l'identification de tels modèles ainsi que des modèles $ARCH$.

On s'intéresse aux fonctions d'autocorrélation et d'autocorrélation partielle du processus carré u_t^2 . Posons

$$\gamma(n) = \gamma(-n) = cov(u_t^2, u_{t-n}^2),$$

alors

$$\gamma(n) = \sum_{i=1}^p \beta_i \gamma(n-i) + \sum_{i=1}^q \alpha_i \gamma(n-i) = \sum_{i=1}^m \varphi_i \gamma(n-i), \quad n \geq p+1 \quad (1.4.1)$$

où $m = \max(p, q)$, $\varphi_i = \alpha_i + \beta_i$, $i = 1, \dots, m$ avec $\alpha_i = 0$ pour $i > q$ et $\beta_i = 0$ pour $i > p$.

Les autocorrélations $\rho(n) = \frac{\gamma(n)}{\gamma(0)}$ vérifient

$$\begin{aligned} \rho(n) &= \rho(-n), \quad \forall n \in \mathbb{N}, \\ &= \sum_{i=1}^m \varphi_i \rho(n-i), \quad n \geq p+1. \end{aligned} \quad (1.4.2)$$

Donc les p autocorrélations de u_t^2 dépendent directement des paramètres $\alpha_1, \dots, \alpha_q$ et $\beta_1, \dots, \beta_p$.

Les autocorrélations partielles d'ordre k

$$\pi_{k,k} = Corr(u_t^2, u_{t-k}^2 / u_{t-1}^2, \dots, u_{t-k+1}^2),$$

sont obtenues en résolvant le système

$$\rho(n) = \sum_{i=1}^k \pi_{k,i} \rho(n-i), \quad n = 1, \dots, k, \quad (1.4.3)$$

en les k inconnues $\pi_{k,i}, i = 1, \dots, k$.

Pour un processus $ARCH(q)$, $\pi_{k,k} \neq 0, k \leq q$ et $\pi_{k,k} = 0, k > q$. Ainsi, ce comportement est semblable à celui d'un processus $AR(q)$. Pour un processus $GARCH(p, q)$ la fonction d'autocorrélation partielle décroît exponentiellement vers zéro. Ces résultats sont conservés quand on utilise les autocorrélations et autocorrélations partielles empiriques (Weiss (1984), McLeod et Li (1983)).

1.5 Modèle IGARCH

Afin de mieux prendre en compte l'information acquise lors de la prédiction à un horizon lointain de la variance conditionnelle, on introduit un modèle "intégré" de la variance, pour avoir un effet "racine unité" sur la variance.

Définition 1.5.1 Un processus $\{u_t, t \in \mathbb{Z}\}$ suit un modèle $IGARCH(p, d, q)$ s'il est défini par les équations (1.2.1) et (1.2.6) et où le polynôme

$$1 - \sum_{i=1}^q \alpha_i z^i + \sum_{j=1}^p \beta_j z^j = 0, \quad (1.5.1)$$

admet d racines unités et $\max(p, q) - d$ racines en dehors du cercle unité.

1.6 Modèle $\beta - ARCH$

C'est une classe de modèles qui permet de comprendre comment se comporte la loi du processus $\{u_t, t \in \mathbb{Z}\}$ en fonction de la loi de $\{\eta_t, t \in \mathbb{Z}\}$.

Définition 1.6.1 On appelle processus $\beta - ARCH$, un processus $\{u_t, t \in \mathbb{Z}\}$ défini par les relations

$$u_t = T(u_{t-1}) + h(u_{t-1}) \eta_t, \quad (1.6.1)$$

où

$$T(x) = ax \text{ et } h(x) = \left[a_0 + (a_1^+ x^+ + a_1^- x^-)^{2\beta} \right]^{1/2}, \quad (1.6.2)$$

où $x^+ = \max(x, 0)$, $x^- = \max(-x, 0)$, a_0, a_1^+, a_1^- et β sont des constantes positives et $\{\eta_t, t \in \mathbb{Z}\}$ est un bruit blanc non supposé gaussien.

Remarque 1.6.2

- 1) Si $a = 0$, alors on a un modèle $\beta - ARCH$ pur. Si $\beta = 1$, on trouve un modèle $AR(1)$ avec erreurs $ARCH(1)$.
- 2) Ce modèle permet de prendre en compte une grande variabilité des données, de modéliser des distributions marginales plus ou moins leptokurtiques en fonction de β .
- 3) Si h n'est pas bornée, il est possible de donner des conditions sur l'existence des moments polynômiaux et exponentiels et des indications sur le comportement asymptotique de la loi de probabilité stationnaire de u_t en fonction de la densité de η_t (voir Diebolt et Guegan, 1994).

1.7 Etude statistique du modèle $ARCH$

1.7.1 Estimation

La pratique de l'inférence statistique sur modèle $ARCH$ consiste à déterminer l'estimateur du (pseudo) maximum de vraisemblance basé sur l'hypothèse de la loi conditionnelle normale. Si $l_t(y; \theta)$ désigne la vraisemblance associée à y_t et conditionnelle au passé, la fonction de vraisemblance de $y_1, \dots, y_t$ conditionnelle à y_0 est $L(y; \theta) = \prod_{i=1}^t l_i(y; \theta)$.

L'estimateur est alors défini comme solution du problème de la recherche du maximum de $L(y; \theta)$. Il est connu que les propriétés de cet estimateur, noté $\hat{\theta}^{(n)}$ et appelé estimateur du pseudo maximum de vraisemblance ($P.M.V$) dépend à la fois de la vraie distribution sous-jacente et de celle utilisée pour calculer la vraisemblance (ici la loi normale). Sous certaines conditions de régularité Gallant (1987), Gouriéroux, Monfort (1989), Gouriéroux, Monfort et Trognon (1984), cet estimateur est convergent, même si la véritable loi sous-jacente n'est pas

conditionnellement normale, c'est à dire ne correspond pas à la loi utilisée pour calculer la vraisemblance. De plus $\sqrt{n} \left(\hat{\theta}^{(n)} - \theta \right)$ est asymptotiquement gaussien de moyenne nulle et de matrice variance covariance $J^{-1} I J^{-1}$ où

$$J = E_0 \left[- \frac{\partial^2 \log(l_t(y; \theta))}{\partial \theta \partial \theta'} \right] \quad (1.7.1)$$

$$I = E_0 \left[\frac{\partial \log(l_t(y; \theta))}{\partial \theta} \frac{\partial \log(l_t(y; \theta))}{\partial \theta'} \right], \quad (1.7.2)$$

et E_0 désignant l'espérance prise par rapport à la vraie loi.

Généralement les deux matrice I et J sont différentes, cependant elles coïncident si la vraie loi est compatible avec la fonction de vraisemblance c'est à dire dans notre cas s'il y a normalité conditionnelle.

1.7.2 Cas du modèle de régression avec erreurs *ARCH*

Considérons le modèle suivant

$$\begin{cases} y_t = \beta' X_t + u_t, \\ u_t = \sigma_t \varepsilon_t, \\ \sigma_t^2 = \alpha_0 + \sum_{i=1}^p \alpha_i u_{t-i}^2, \end{cases} \quad (1.7.3)$$

où ε_t est une suite de variables aléatoires indépendantes de loi normale centrée réduite.

Soit $\theta = (\beta', \alpha')'$ le paramètre global du modèle avec $\alpha = (\alpha_0, \alpha_1, \dots, \alpha_p)'$ le paramètre de la partie *ARCH*.

En appliquant le résultat précédent, nous avons

$$Var \left(\sqrt{n} \begin{pmatrix} \hat{\alpha}_n - \alpha \\ \hat{\beta}_n - \beta \end{pmatrix} \right) = J^{-1} = \begin{pmatrix} J_{\alpha\alpha} & J_{\alpha\beta} \\ J_{\beta\alpha} & J_{\beta\beta} \end{pmatrix}^{-1}, \quad (1.7.4)$$

où

$$J_{\alpha\alpha} = E_0 \left[\frac{1}{\sigma_t^2} X_t' X_t + \frac{2}{\sigma_t^4} \left(\sum_{i=1}^p \alpha_i X_{t-i}' u_{t-i} \right) \left(\sum_{i=1}^p \alpha_i X_{t-i} u_{t-i} \right) \right] \quad (1.7.5)$$

$$J_{\beta\beta} = E_0 \left[\frac{1}{\sigma_t^4} \begin{pmatrix} 1 \\ u_{t-1} \\ \vdots \\ u_{t-p} \end{pmatrix} \begin{pmatrix} 1 & u_{t-1}^2 & \cdots & u_{t-p}^2 \end{pmatrix} \right] \quad (1.7.6)$$

$$J_{\beta\alpha} = -E_0 \left[\frac{1}{\sigma_t^4} \begin{pmatrix} 1 \\ u_{t-1} \\ \vdots \\ u_{t-p} \end{pmatrix} \left(\sum_{i=1}^p \alpha_i X_{t-i} u_{t-i} \right) \right] \quad (1.7.7)$$

on peut montrer que les matrices $J_{\beta\alpha}$ et $J_{\alpha\beta}$ sont nulles.

Donc si ε_t a une représentation ARCH conditionnellement gaussienne, les estimateurs β et α sont asymptotiquement non corrélés, mais corrélés dans le cas général.

1.7.3 Procédure d'estimation en deux étapes

Nous allons montrer le mécanisme de cette procédure sur le modèle défini dans (1.7.3). Par la régression de y_t sur X_t , on obtient l'estimateur des moindres carrés ordinaire de β . On le note $\tilde{\beta}_n$. C'est un estimateur convergent de β . On calcule $\tilde{u}_t = y_t - \tilde{\beta}_n' X_t$ et la régression de $\tilde{u}_t^2$ sur $1, \tilde{u}_{t-1}^2, \dots, \tilde{u}_{t-p}^2$, on obtient l'estimateur des moindres carrés ordinaire de $\alpha = (\alpha_0, \alpha_1, \dots, \alpha_p)'$ noté $\tilde{\alpha}^{(n)}$. Afin d'obtenir un meilleur estimateur on minimise la quantité $\sum_{t=1}^n \frac{(y_t - \beta' X_t)^2}{\tilde{\sigma}_t^2}$ avec $\tilde{\sigma}_t = \tilde{\alpha}_0 + \sum_{i=1}^p \tilde{\alpha}_i \tilde{u}_{t-i}^2$; l'estimateur obtenu noté $\tilde{\tilde{\beta}}$ est appelé l'estimateur des moindres carrés quasi-généralisés (MCQG).

Pour obtenir une seconde estimation de $\alpha_0, \alpha_1, \dots, \alpha_p$, on minimise

$$\sum_{t=1}^n \frac{1}{2\tilde{\sigma}_t^4} \left(\varepsilon_t^2 - \alpha_0 - \sum_{i=1}^p \alpha_i u_{t-i}^2 \right)^2,$$

les estimateurs obtenus sont notés $\tilde{\tilde{\alpha}}_0, \tilde{\tilde{\alpha}}_1, \dots, \tilde{\tilde{\alpha}}_p$.

Propriétés

- 1) L'estimateur des moindres carrés quasi-généralisé ($MQQG$) de β est asymptotiquement moins efficace que l'estimateur du pseudo maximum de vraisemblance (PMV).
- 2) L'estimateur des moindres carrés quasi-généralisé ($MQQG$) de $\alpha = (\alpha_0, \alpha_1, \dots, \alpha_p)'$ est asymptotiquement efficace.
- 3) Les estimateurs des moindres carrés quasi-généralisés ($MQQG$) de β et de $\alpha = (\alpha_0, \alpha_1, \dots, \alpha_p)'$ sont asymptotiquement non corrélés.

Le cas du modèle $ARMA$ avec bruit $GARCH(p, q)$ a été étudié par Weiss (1986).

L'estimation adaptative pour le modèle $ARCH$ a été étudié par Linton (1993) pour le modèle $GARCH(1, 1)$ par Drost et Werker (1996).

1.7.4 Identification

Nous considérons des modèles $ARMA$ avec bruit $ARCH$; on procède de la manière suivante :

- 1— On ajuste un modèle $ARMA$
- 2— On estime les paramètres de la partie $ARMA$ par la méthode des moindres carrés ou par le maximum de vraisemblance.
- 3— A partir du calcul de l'erreur, après identification du modèle $ARMA$, on ajuste un modèle $ARCH$ (régression sur le bruit).
- 4— On estime les paramètres du modèle $ARCH$ par les moindres carrés.

Chapitre 2

Estimation récursive en-ligne des modèles ARCH

2.1 Introduction

Les modèles *ARCH* introduits par Engel (1982) et leurs extensions jouent un rôle prédominant parmi les diverses formulations proposées pour représenter les séries économiques, notamment les séries financières. De nos jours, de telles séries (cours boursiers, taux de change,...) présentent une dynamique de haute fréquence similaire à celle des données rencontrées dans les domaines du contrôle adaptatif et du traitement de signal, dans le sens où elles tendent à être progressivement disponibles (en-ligne) à des intervalles de temps très courts. Ainsi, l'emploi de méthodes classiques d'estimation (maximum de vraisemblance, moindres carrés : Weiss, 1986 ; Li et al. 2002 ; Berkes et al. 2003 ; Ling et McAleer, 2003 ; Francq et Zakoïan, 2004) qui supportent des observations de taille fixe (hors-ligne) peut s'avérer trop lourd.

De plus, il est bien connu, de nos jours que les modèles *ARCH* sont de bonnes approximations pour certains modèles de séries chronologiques en temps continu (voir Nelson 1990 ; Nelson and Foster, 1994 ; Drost and Werker, 1996). Ceci fournit donc une motivation pour la recherche de méthodes d'estimation qui prennent en charge ces aspects. L'objet du présent chapitre est la mise en oeuvre d'un algorithme récurrent pour l'estimation, en deux étapes, d'un modèle *ARCH*. Cette méthode, inspirée des algorithmes bien connus d'estimation en-ligne (par exemple Kushner et Clark, 1978 ; Ljung et Söderström, 1983 ; Goodwin et Sin, 1984 ; Young, 1984 ; Benveniste et al., 1990 ; Duflo, 1990 ; Solo et Kong 1995 ; Kushner et Yin, 1997 ; Ljung, 1999 ; Pollock, 2003), fournit des estimateurs asymptotiquement gaussiens et permet de prendre en charge le caractère nouvellement en-ligne des données économétriques. Il s'agit, en fait, d'une fusion de la méthode des moindres carrés en deux étapes proposée par Bose et Mukherjee (2003) à taille d'échantillon fixée (voir aussi Pantula (1988) pour le cas particulier *ARCH* d'ordre un), avec la célèbre méthode des moindres carrés récursive (*Recursive Least Squares*, *RLS* en anglais) développée pour des modèles *AR* linéaires et permettant de prendre en charge des données en temps réel.

Le reste de ce chapitre est organisé comme suit. La section 2 donne les hypothèses sur lesquelles se base la méthode d'estimation que nous proposons. La section 3 est consacrée à la mise en oeuvre d'un algorithme *RLS* en deux étapes pour l'estimation des paramètres d'un modèle *ARCH*. Dans la section 4 nous nous intéressons à l'obtention de la distribution asymptotique des estimateurs obtenus par cet algorithme. Nous montrons que ces estimateurs sont asymptotiquement gaussiens et que, de plus, l'estimateur de la deuxième étape est asymptotiquement efficace. Enfin, la section 5 montre la performance de l'algorithme proposé à travers une étude de simulation.

2.2 Modèle et hypothèses

Un processus du second ordre $\{\varepsilon_t; t \in \mathbb{Z}\}$ admet une représentation *ARCH* d'ordre q s'il est solution d'une équation aux différences stochastique de la forme

$$\begin{cases} \varepsilon_t = \sqrt{h_t} \zeta_t, \\ h_t = \alpha_0 + \alpha(L) \varepsilon_t^2, \quad t \in \mathbb{Z}, \end{cases} \quad (2.2.1)$$

où $\{\zeta_t, t \in \mathbb{Z}\}$ est une suite de variables aléatoires indépendantes de même loi de probabilité centrée réduite (*i.i.d.*) et de moments d'ordre quatre finis; le polynôme $\alpha(L)$ est donné par : $\alpha(L) = \sum_{i=1}^q \alpha_i L^i$, où L est l'opérateur retard. On suppose que, pour tout $t \in \mathbb{Z}$, ζ_t est indépendante de la tribu $\mathcal{F}_{t-1}$ engendrée par les variables $\{\varepsilon_s, s < t\}$. Ainsi $E(\varepsilon_t / \mathcal{F}_{t-1}) = 0$, i.e. ε_t est une différence de martingales et $h_t = E(\varepsilon_t^2 / \mathcal{F}_{t-1})$ est la variance conditionnelle de ε_t sachant $\mathcal{F}_{t-1}$. Pour la condition de positivité de la variance conditionnelle, les paramètres réels α_i sont tels que $\alpha_0 > 0$, $\alpha_i \geq 0$, $i = 1, \dots, q$. Posons $v_t = \varepsilon_t^2 - h_t = h_t \eta_t$, où $\eta_t = \zeta_t^2 - 1$; nous pouvons donner une représentation équivalente au modèle (2.2.1) en l'écrivant sous une forme de régression linéaire dans laquelle le carré du processus ε_t^2 , $t \in \mathbb{Z}$ est la variable à expliquer.

$$\varepsilon_t^2 = \xi_t' \alpha + h_t(\alpha) \eta_t, \quad (2.2.2)$$

où $\xi_t = (1, \varepsilon_{t-1}^2, \varepsilon_{t-2}^2, \dots, \varepsilon_{t-q}^2)'$, $h_t(\alpha) = \xi_t' \alpha$ et $\alpha = (\alpha_0, \alpha_1, \dots, \alpha_q)'$ de dimension $(q+1) \times 1$ est le vecteur des paramètres du modèle. Considérons les hypothèses suivantes :

H1 : Le processus $\{\varepsilon_t; t \in \mathbb{Z}\}$ est strictement stationnaire et ergodique.

H2 : Les moments $E(\varepsilon_i^2 \varepsilon_j^2)$ existent et sont finis pour tout i, j .

H3 : Les moments $E(\varepsilon_i^4 \varepsilon_j^4 \varepsilon_k^4)$ sont finis pour tout i, j, k .

Pour le cas hors-ligne, Bose et Mukherjee (2003) ont considéré l'hypothèse *H2* pour la normalité asymptotique de l'estimateur de la première étape et l'hypothèse d'existence des moments d'ordre huit pour la normalité de l'estimateur d'intérêt. Dans notre cas, *H3* est exigée pour l'efficacité asymptotique de l'estimateur d'intérêt fourni par la méthode que nous proposons.

2.3 Algorithme *RLS* en deux étapes (*2S-RLS*)

Posons $\varepsilon^t = \{\varepsilon_{1-q}, \dots, \varepsilon_0, \varepsilon_1, \dots, \varepsilon_t\}$ où la taille t de la série chronologique n'est pas a priori fixée et où les valeurs initiales $\{\varepsilon_{1-q}, \dots, \varepsilon_0\}$ sont, par exemple, supposées connues. Bose et Mukherjee (2003) ont proposé une méthode des moindres carrés en deux étapes pour estimer

les paramètres du modèle (2.2.1) en résolvant les deux systèmes d'équations suivants

$$\tilde{\alpha}(N) = \left(\sum_{k=1}^N \xi_k \xi_k' \right)^{-1} \sum_{k=1}^N \varepsilon_k^2 \xi_k, \quad (2.3.1a)$$

$$\hat{\alpha}(N) = \left(\sum_{k=1}^N \frac{1}{h_k^2(\tilde{\alpha}(N))} \xi_k \xi_k' \right)^{-1} \sum_{k=1}^N \frac{\varepsilon_k^2}{h_k^2(\tilde{\alpha}(N))} \xi_k, \quad (2.3.1b)$$

où $\tilde{\alpha}(N)$ est un estimateur préliminaire de α et $\hat{\alpha}(N)$ est l'estimateur d'intérêt, tous deux fondés sur un échantillon d'observations de taille N . Ils ont étudié la distribution asymptotique de $\tilde{\alpha}(N)$ et $\hat{\alpha}(N)$ et ont montré que ces estimateurs sont asymptotiquement gaussiens et que, de plus, sous l'hypothèse d'existence des moments d'ordre huit, l'estimateur d'intérêt $\hat{\alpha}(N)$ est asymptotiquement efficace et sa performance est supérieure à celle de l'estimateur du pseudo-maximum de vraisemblance, même pour de petits échantillons.

Dans ce chapitre nous proposons une alternative récursive à (2.3.1) qui permette de prendre en charge des données en temps réel ε^t et pour laquelle les estimateurs obtenus possèdent les mêmes propriétés asymptotiques que le couple $(\tilde{\alpha}(N), \hat{\alpha}(N))$. Notons que l'estimateur préliminaire $\tilde{\alpha}(N)$ donné dans (2.3.1a) exhibe une forme *OLS* standard et peut être obtenu de façon récursive. En effet, en posant

$$\tilde{R}_t = \sum_{k=1}^t \xi_k \xi_k', \quad (2.3.2a)$$

$$\tilde{\alpha}_t = \tilde{\alpha}(t), \quad (2.3.2b)$$

le système (2.3.1a) peut être résolu de façon récursive en utilisant la méthode, bien connue, des moindres carrés récurrents (*RLS*) que nous rappelons ici (voir également Ljung et Söderström (6, p. 17)).

En utilisant la notation dans (2.3.2) l'équation normale (2.3.1a) peut se réécrire comme suit

$$\tilde{R}_t \tilde{\alpha}_t = \sum_{k=1}^t \varepsilon_k^2 \xi_k, \quad (2.3.3)$$

et l'on a

$$\tilde{R}_t = \tilde{R}_{t-1} + \xi_t \xi_t'. \quad (2.3.4)$$

Dès lors, (2.3.3) devient

$$\begin{aligned}
\tilde{\alpha}_t &= \tilde{R}_t^{-1} \sum_{k=1}^t \varepsilon_k^2 \xi_k, \\
&= \tilde{R}_t^{-1} \left[\sum_{k=1}^{t-1} \varepsilon_k^2 \xi_k + \varepsilon_t^2 \xi_t \right], \\
&= \tilde{R}_t^{-1} \left[\tilde{R}_{t-1} \tilde{\alpha}_{t-1} + \varepsilon_t^2 \xi_t \right], \\
&= \tilde{R}_t^{-1} \left[\left(\tilde{R}_t - \xi_t \xi_t' \right) \tilde{\alpha}_{t-1} + \varepsilon_t^2 \xi_t \right], \\
&= \tilde{\alpha}_{t-1} + \tilde{R}_t^{-1} \xi_t \left(\varepsilon_t^2 - \xi_t' \tilde{\alpha}_{t-1} \right).
\end{aligned} \tag{2.3.5}$$

Ainsi, en regroupant (2.3.4) et (2.3.5), on obtient les équations récurrentes suivantes permettant d'obtenir l'estimateur hors ligne $\tilde{\alpha}(t)$ récursivement

$$\begin{cases} \tilde{\alpha}_t = \tilde{\alpha}_{t-1} + \tilde{R}_t^{-1} \xi_t (\varepsilon_t^2 - \xi_t' \tilde{\alpha}_{t-1}), & t \geq 2, \\ \tilde{R}_t = \tilde{R}_{t-1} + \xi_t \xi_t', \end{cases} \tag{2.3.6}$$

où $\tilde{R}_1 = \xi_1 \xi_1'$ et $\tilde{\alpha}_1 = \tilde{R}_1^{-1} \xi_1 \varepsilon_1^2$ sont des valeurs initiales et l'on suppose que $\tilde{R}_1$ est non singulière.

Seulement, (2.3.1b) ne peut être obtenue de façon récursive comme (2.3.1a) en raison de la présence de $\tilde{\alpha}(N)$ dans les poids $h_k^{-2}(\tilde{\alpha}(N))$, $k = 1, \dots, N$.

En effet, si on démarre de l'équation normale correspondante (2.3.1b), alors en posant

$$\hat{R}_t = \sum_{k=1}^t \frac{1}{h_k^2(\tilde{\alpha}_t)} \xi_k \xi_k', \tag{2.3.7}$$

on trouve

$$\begin{aligned}
\hat{R}_t &= \sum_{k=1}^{t-1} \frac{1}{h_k^2(\tilde{\alpha}_t)} \xi_k \xi_k' + \frac{1}{h_t^2(\tilde{\alpha}_t)} \xi_t \xi_t', \\
&\neq \hat{R}_{t-1} + \frac{1}{h_t^2(\tilde{\alpha}_t)} \xi_t \xi_t',
\end{aligned}$$

car d'après (2.3.7), $\widehat{R}_{t-1}$ serait nécessairement égale à

$$\sum_{k=1}^{t-1} \frac{1}{h_k^2(\widetilde{\alpha}_{t-1})} \xi_k \xi'_k,$$

et non à

$$\sum_{k=1}^{t-1} \frac{1}{h_k^2(\widetilde{\alpha}_t)} \xi_k \xi'_k.$$

C'est pourquoi nous sommes obligés, pour écrire $\widehat{\alpha}_t$ de façon récursive tout en garantissant la même distribution asymptotique de l'estimateur hors-ligne d'intérêt $\widehat{\alpha}(N)$, de définir notre estimateur d'intérêt récursif comme solution de la nouvelle équation normale modifiée

$$\widehat{\alpha}_t = \left(\sum_{k=1}^t \frac{1}{h_k^2(\widetilde{\alpha}_k)} \xi_k \xi'_k \right)^{-1} \sum_{k=1}^t \frac{\varepsilon_k^2}{h_k^2(\widetilde{\alpha}_k)} \xi_k, \quad (2.3.8)$$

dans laquelle les poids $h_k^{-2}(\widetilde{\alpha}(N))$ sont remplacés par $h_k^{-2}(\widetilde{\alpha}_k)$, ($\widetilde{\alpha}_k$ étant l'estimateur récursif d'intérêt du paramètre α au temps k) plutôt que de l'équation normale (2.3.1b). Il est clair que (2.3.1b) et (2.3.8) fournissent des estimateurs différents. Cependant, ils possèdent, comme nous le montrons ci-dessous, les mêmes propriétés asymptotiques. Avec (2.3.8), il est clair qu'en suivant les mêmes étapes que pour (2.3.2)-(2.3.6), on trouve :

$$\begin{cases} \widehat{\alpha}_t = \widehat{\alpha}_{t-1} + \frac{\widehat{R}_t^{-1} \xi_t (\varepsilon_t^2 - \xi'_t \widehat{\alpha}_{t-1})}{(\xi'_t \widetilde{\alpha}_t)^2}, \\ \widehat{R}_t = \widehat{R}_{t-1} + \frac{\xi_t \xi'_t}{(\xi'_t \widetilde{\alpha}_t)^2}, \end{cases}, \quad t \geq 2, \quad (2.3.9)$$

avec valeurs de démarrage

$$\widehat{R}_1 = \frac{1}{h_1^2(\widetilde{\alpha}_1)} \xi_1 \xi'_1 \text{ et } \widehat{\alpha}_1 = \widehat{R}_1^{-1} \xi_1 \frac{\varepsilon_1^2}{h_1^2(\widetilde{\alpha}_1)},$$

pourvu que $\widehat{R}_1$ soit non singulière.

Il est encore possible d'améliorer à nouveau la complexité algorithmique de (2.3.6) et (2.3.9) en évitant d'inverser les matrices $\widetilde{R}_t$ et $\widehat{R}_t$. Ceci peut être atteint en utilisant, dans

les systèmes d'équations (2.3.6) et (2.3.9) le lemme suivant bien connu sous le nom lemme d'inversion matricielle (voir aussi Appendice, Plackett (1950), Solo (1979), Ljung et Söderström (1983)).

En posant $\tilde{P}_t = \tilde{R}_t^{-1}$ et par application du Lemme 2.3.1 sur (2.3.4) avec $A = \tilde{P}_{t-1}^{-1}$, $B = \xi_t$, $C = 1$ et $D = \xi'_t$, on peut trouver des équations récurrentes sur $\tilde{P}_t$ plus simples :

$$\begin{aligned}\tilde{P}_t &= \left(\tilde{P}_{t-1}^{-1} + \xi_t \xi'_t \right)^{-1} \\ &= \tilde{P}_{t-1} - \frac{\tilde{P}_{t-1} \xi_t \xi'_t \tilde{P}_{t-1}}{1 + \xi'_t \tilde{P}_{t-1} \xi_t}.\end{aligned}\tag{2.3.10}$$

De même, posons $\hat{P}_t = \hat{R}_t^{-1}$ et par application du Lemme 2.3.1 sur la seconde équation de (2.3.9) avec cette fois-ci $A = \hat{P}_{t-1}^{-1}$, $B = \xi_t$, $C = (\xi'_t \tilde{\alpha}_t)^{-2}$ et $D = \xi'_t$, on trouve

$$\begin{aligned}\hat{P}_t &= \left(\hat{P}_{t-1}^{-1} + \xi_t (\xi'_t \tilde{\alpha}_t)^{-2} \xi'_t \right)^{-1} \\ &= \hat{P}_{t-1} - \frac{\hat{P}_{t-1} \xi_t \xi'_t \hat{P}_{t-1}}{(\xi'_t \tilde{\alpha}_t)^2 + \xi'_t \hat{P}_{t-1} \xi_t}.\end{aligned}\tag{2.3.11}$$

Ainsi, exploitant (2.3.10) et (2.3.11) nous obtenons l'algorithme suivant que nous nommerons l'algorithme des moindres carrés récursif en deux étapes (*2S-RLS*).

Algorithme 2.3.1 (Aknouche et Guerbyenne, 2006a)

L'algorithme (2S-RLS) pour estimer récursivement les paramètres d'un modèle ARCH est donné par le système d'équations récurrentes suivant :

$$\underbrace{\begin{aligned} \tilde{\alpha}_t &= \tilde{\alpha}_{t-1} + \frac{\tilde{P}_{t-1}\xi_t(\varepsilon_t^2 - \xi_t'\tilde{\alpha}_{t-1})}{1 + \xi_t'\tilde{P}_{t-1}\xi_t}, \tilde{\alpha}_0 = 0, \\ \tilde{P}_t &= \tilde{P}_{t-1} - \frac{\tilde{P}_{t-1}\xi_t\xi_t'\tilde{P}_{t-1}}{1 + \xi_t'\tilde{P}_{t-1}\xi_t}, \tilde{P}_0 = M.I, \end{aligned}}_{\text{première étape}} \rightarrow \underbrace{\begin{aligned} \hat{\alpha}_t &= \hat{\alpha}_{t-1} + \frac{\hat{P}_{t-1}\xi_t(\varepsilon_t^2 - \xi_t'\hat{\alpha}_{t-1})}{(\xi_t'\hat{\alpha}_t)^2 + \xi_t'\hat{P}_{t-1}\xi_t}, \hat{\alpha}_0 = 0, \\ \hat{P}_t &= \hat{P}_{t-1} - \frac{\hat{P}_{t-1}\xi_t\xi_t'\hat{P}_{t-1}}{(\xi_t'\hat{\alpha}_t)^2 + \xi_t'\hat{P}_{t-1}\xi_t}, \hat{P}_0 = M.I, \end{aligned}}_{\text{seconde étape}}$$

où $\tilde{P}_t = \tilde{R}_t^{-1}$, $\hat{P}_t = \hat{R}_t^{-1}$ et M est un nombre réel positif suffisamment grand.

Les valeurs de démarrage de l'algorithme $\tilde{\alpha}_0$, $\tilde{P}_0$, $\hat{\alpha}_0$ et $\hat{P}_0$ ont été fixées, d'une part de façon à éviter le problème de singularité pour les deux matrices $\tilde{R}_1$ et $\hat{R}_1$ et d'autre part pour obtenir des estimateurs qui soient proches des solutions de (2.3.1a) et (2.3.8).

Remarque 2.3.1

L'estimateur d'intérêt $\hat{\alpha}(N)$ de Bose et Mukherjee (2003) est obtenu (par la méthode des moindres carrés) en approximant le modèle ARCH écrit sous la forme de régression linéaire

$$\frac{\varepsilon_k^2}{h_k(\alpha)} = \left(\frac{\xi_k}{h_k(\alpha)} \right)' \alpha + \eta_k, \quad k = 1, \dots, N.$$

par le modèle suivant

$$\frac{\varepsilon_k^2}{h_k(\tilde{\alpha}(N))} = \left(\frac{\xi_k}{h_k(\tilde{\alpha}(N))} \right)' \alpha + \eta_k, \quad k = 1, \dots, N.$$

dans lequel $h_k(\tilde{\alpha}(N))$ approxime la variance conditionnelle $h_k(\alpha)$ à l'instant k .

Alors que le nôtre, $\hat{\alpha}_t$, est obtenu (par la méthode des moindres carrés récursive) à partir du modèle de régression linéaire suivant

$$\frac{\varepsilon_k^2}{h_k(\tilde{\alpha}_k)} = \left(\frac{\xi_k}{h_k(\tilde{\alpha}_k)} \right)' \alpha + \eta_k, \quad k = 1, \dots, t.$$

où la variance conditionnelle $h_k(\alpha)$ est remplacée par $h_k(\tilde{\alpha}_k)$.

2.4 Propriétés asymptotiques des estimateurs 2S-RLS

Dans ce qui suit, nous donnons la distribution asymptotique des estimateurs obtenus par (2.3.13) en nous basant sur les hypothèses $H1$ - $H3$.

Théorème 2.3.1 (Aknouche et Guerbyenne, 2006a)

i) Sous les hypothèses $H1$ et $H2$

$$t^{1/2}(\tilde{\alpha}_t - \alpha^*) \xrightarrow{L} N \left(0, \text{var}(\eta_0^2) \{E(\xi_0 \xi_0')\}^{-1} E \left(\xi_0 \xi_0' (\xi_0' \alpha^*)^2 \right) \{E(\xi_0 \xi_0')\}^{-1} \right), \quad (2.4.1)$$

ii) Si, de plus, $H3$ est vérifiée, alors

$$t^{1/2}(\hat{\alpha}_t - \alpha^*) \xrightarrow{L} N \left(0, \text{var}(\eta_0^2) \left\{ E \left(\xi_0 \xi_0' (\xi_0' \alpha^*)^{-2} \right) \right\}^{-1} \right). \quad (2.4.2)$$

Preuve

Notons que, puisque $\tilde{\alpha}_t = \tilde{\alpha}(t)$, (2.4.1) résulte directement, sous $H1$ et $H2$, de Bose et Mukherjee (2003). Pour montrer (2.4.2), nous pouvons écrire à partir de (2.3.8) et (2.2.2)

$$t^{1/2}(\hat{\alpha}_t - \alpha^*) = \left[t^{-1} \sum_{k=1}^t \frac{\xi_k \xi_k'}{(\xi_k' \tilde{\alpha}_k)^2} \right]^{-1} \left[t^{-1/2} \sum_{k=1}^t \frac{(\xi_k' \alpha^*) \eta_k}{(\xi_k' \tilde{\alpha}_k)^2} \xi_k \right], \quad (2.4.3)$$

où α est remplacé par α^* la vraie valeur du paramètre. Comme dans Bose et Mukherjee (2003), pour établir (2.4.2), il suffit de montrer que

$$t^{-1} \sum_{k=1}^t \left(\frac{1}{(\xi_k' \tilde{\alpha}_k)^2} - \frac{1}{(\xi_k' \alpha^*)^2} \right) \xi_k \xi_k' = o_p(1), \quad (2.4.4a)$$

$$t^{-1/2} \sum_{k=1}^t \left(\frac{1}{(\xi_k' \tilde{\alpha}_{k-1})^2} - \frac{1}{(\xi_k' \alpha^*)^2} \right) (\xi_k' \alpha^*) \eta_k \xi_k = o_p(1), \quad (2.4.4b)$$

pour pouvoir réécrire (2.3.4) comme suit

$$t^{1/2}(\hat{\alpha}_t - \alpha^*) = \left[t^{-1} \sum_{k=1}^t \frac{\xi_k \xi_k'}{(\xi_k' \alpha^*)^2} \right]^{-1} \left[t^{-1/2} \sum_{k=1}^t \frac{(\xi_k' \alpha^*) \eta_k}{(\xi_k' \alpha^*)^2} \xi_k \right] + o_p(1), \quad (2.4.5)$$

et, d'après le théorème central limite des martingales (Hall et Heyde 1980, chapitre 3) qui est facilement applicable à (2.4.5), (Bose et Mukherjee, 2003) nous concluons que (2.4.2) est vraie. Maintenant pour prouver (2.4.4a), en utilisant le théorème de la valeur moyenne (Bose et Mukherjee 2003, formule (15), p. 133), nous avons

$$\frac{1}{t} \sum_{k=1}^t \left(\frac{1}{(\xi'_k \tilde{\alpha}_k)^2} - \frac{1}{(\xi'_k \alpha^*)^2} \right) \xi_k \xi'_k = -\frac{2}{t} \sum_{k=1}^t \frac{(\tilde{\alpha}_k - \alpha^*)' \xi_k}{\chi_k^3} \xi_k \xi'_k = -2S_1, \quad (2.4.6)$$

la constante χ_k est telle que

$$\begin{aligned} 0 &< \frac{1}{\chi_k^3} \leq \frac{1}{(\xi'_k \alpha^*)^3} \left[1 + \frac{\xi'_k \alpha^*}{\xi'_k \tilde{\alpha}_k} \right]^3, \\ &\leq \frac{1}{(\xi'_k \alpha^*)^3} \left[1 + \left\{ 1 + \sum_{j=0}^q \frac{\alpha_j^*}{\tilde{\alpha}_j^{(k)}} \right\} \right]^3 \leq (\alpha_0^*)^{-3} \left[2 + \sum_{j=0}^q \frac{\alpha_j^*}{\tilde{\alpha}_j^{(k)}} \right]^3, \end{aligned} \quad (2.4.7)$$

où dans la première inégalité de (2.4.7), nous avons utilisé la relation (17) dans Bose et Mukherjee (2003, p. 133). La borne supérieure de la dernière inégalité est bornée en probabilité, pour tout k , en raison de (2.4.1). Soit $C_1 > 0$ une telle borne. Alors de (2.4.6) et (2.4.7) nous avons

$$\|S_1\| \leq C_1 t^{-1} \sum_{k=1}^t \|\tilde{\alpha}_k - \alpha^*\| \|\xi_k\|^3,$$

($\|\cdot\|$ désigne la norme Euclidienne).

Puisque à partir de (2.4.1) il existe $C_2 > 0$ tel que

$$k^{1/2} \|\tilde{\alpha}_k - \alpha^*\| \leq C_2,$$

(en probabilité), sous l'hypothèse d'existence des moments d'ordre six, garantie par H3, il s'en suit que

$$\sum_{k=1}^{\infty} k^{-3/2} E(\|\xi_k\|^3) < \infty,$$

et donc (voir par exemple Lukacs (1975, p. 80))

$$\sum_{k=1}^{\infty} k^{-1} \|\tilde{\alpha}_k - \alpha^*\| \|\xi_k\|^3 < \infty, \text{ p.s.} \quad (2.4.8)$$

Ainsi (2.4.4a) est obtenue à partir de (2.4.8) et du lemme de Kronecker.

Pour établir (2.4.4b), nous utilisons à nouveau le théorème de la valeur moyenne qui conduit à

$$t^{-1/2} \sum_{k=1}^t \left(\frac{1}{(\xi'_k \tilde{\alpha}_k)^2} - \frac{1}{(\xi'_k \alpha^*)^2} \right) (\xi'_k \alpha^*) \eta_k \xi_k = -2t^{-1/2} \sum_{k=1}^t \frac{((\tilde{\alpha}_k - \alpha^*)' \xi_k) (\xi'_k \alpha^*) \eta_k}{\chi_k^3} \xi_k,$$

$$\stackrel{def}{=} -2t^{-\delta} S_{t,\delta},$$

pour un certain δ tel que $0 < \delta < 1/2$. Si nous pouvons montrer que

$$E \|S_{t,\delta}\|^2 < \infty,$$

alors d'après l'inégalité de Tchebychev, nous aurons

$$t^{-\delta} \|S_{t,\delta}\| = o_p(1),$$

ce qui équivaut à (2.4.4b).

En utilisant *H1* et l'indépendance de la suite (η_k) , nous pouvons écrire

$$\begin{aligned} E (\|S_t\|^2) &\leq t^{-1+2\delta} \sum_{k=1}^t \frac{E \left\| \frac{((\tilde{\alpha}_k - \alpha^*)' \xi_k) (\xi'_k \alpha^*) \xi_k}{\chi_k^6} \right\|^2}{\chi_k^6} E \eta_k^2, \\ &\leq C_1^2 t^{-1+2\delta} \sum_{k=1}^t E (C_2^2 \|\alpha^*\|^2 \|\xi_k\|^6) k^{-1}, \\ &\leq (C_1^2 C_2^2 \|\alpha^*\|^2 E \|\xi_0\|^6) t^{-1+2\delta} \log t < \infty. \end{aligned}$$

Maintenant (2.4.4b) est établie. ■

Remarque 2.4.1

L'hypothèse *H3* paraît contraignante. Cette hypothèse, très forte d'ailleurs que celles exigées pour la normalité asymptotique d'autres estimateurs tels que l'estimateur du pseudo maximum de vraisemblance (*QMLE*), induit une restriction sur l'espace des paramètres, et donc nuit à la modélisation des processus à queues épaisses, auxquels sont préconisés les modèles *ARCH*. Cependant *H3* simplifie la preuve et les techniques utilisées pour établir la convergence de l'estimateur *QML* sans aucune condition sur les moments (voir par exemple Francq et Zakoïan (2004)) ne nous semblent pas adaptables à notre problème. Néanmoins,

nous pensons que notre résultat peut être un point de départ pour d'éventuelles améliorations, puisqu'il nous semble qu'il n'y a pas dans la littérature de l'estimation en-ligne des méthodes récursives pour estimer les modèles *ARCH*.

En outre, ce problème de nécessité d'existence des moments d'ordres supérieurs est une des faiblesses de la méthode des moindres carrés. Faiblesse rencontrée spécialement lorsqu'elle est appliquée à l'estimation des modèles *ARCH*, contrairement à la méthode du maximum de vraisemblance pour laquelle il existe actuellement des travaux établissant la distribution asymptotique de l'estimateur du pseudo maximum de vraisemblance sans aucune condition sur les moments (voir par exemple, Lumsdaine, 1996 ; Lee et Hansen, 1994 ; Li et al. 2002 ; Berkes et al. 2003 ; Li et McAleer, 2003 ; Francq et Zakoïan, 2004). Le problème réside à notre avis dans le fait que, pour la méthode des moindres carrés, les matrices limites de l'estimateur *OLS* sont donnés explicitement à travers des moments d'ordres élevés, alors que les matrices limites associées à l'estimateur du pseudo maximum de vraisemblance sont exprimées implicitement à travers l'espérance des dérivées de la fonction de vraisemblance (voir Chapitre 1). En outre, Bose et Mukherjee (2003) ont exigé l'existence des moments d'ordre 8 pour l'existence de la matrice limite de leur estimateur d'intérêt $\hat{\alpha}(N)$ fondé sur l'équation

$$\hat{\alpha}(N) = \left(\sum_{k=1}^N \frac{1}{h_k^2(\hat{\alpha}(N))} \xi_k \xi_k' \right)^{-1} \sum_{k=1}^N \frac{\varepsilon_k^2}{h_k^2(\hat{\alpha}(N))} \xi_k,$$

De même, Weiss (1986) a également exigé la même condition pour l'estimateur des moindres carrés.

Si notre estimateur d'intérêt $\hat{\alpha}_t$ défini par

$$\hat{\alpha}_t = \left(\sum_{k=1}^t \frac{1}{h_k^2(\hat{\alpha}_k)} \xi_k \xi_k' \right)^{-1} \sum_{k=1}^t \frac{\varepsilon_k^2}{h_k^2(\hat{\alpha}_k)} \xi_k,$$

exige des moments d'ordre encore plus élevé (d'ordre 12) que l'estimateur d'intérêt hors-ligne $\hat{\alpha}(N)$, c'est parce que, nous-semble-t-il, les poids $h_k^{-2}(\hat{\alpha}_k)$ fonction de k , ($k = 1, \dots, t$) contribuent beaucoup moins rapidement à la convergence de $\hat{\alpha}_t$ vers α^* que ne le font les poids $h_k^2(\hat{\alpha}(N))$, $k = 1, \dots, N$, (fonction du dernier rang N , N tendant vers l'infini) pour $\hat{\alpha}(N)$.

Remarque 2.4.2

Ayant montré la convergence en loi des estimateurs récursifs établis, nous pouvons aller au delà et montrer encore la convergence presque sûre des deux estimateurs vers la vraie valeur du paramètre. La démonstration repose sur la même technique que celle de la convergence en loi, c'est à dire en se basant sur la même équation normale. En effet, pour montrer la convergence presque sûre de l'estimateur de la première étape, nous nous basons sur l'équation

$$(\tilde{\alpha}_t - \alpha^*) = \left[t^{-1} \sum_{k=1}^t \xi_k \xi_k' \right]^{-1} \left[t^{-1} \sum_{k=1}^t \xi_k' \alpha^* \xi_k \eta_k \right],$$

obtenue en combinant (2.3.1a) et (2.2.2). Sous $H1$ et l'hypothèse de l'existence des moments d'ordre 4, nous pourrions utiliser le théorème ergodique qui donne

$$t^{-1} \sum_{k=1}^t \xi_k \xi_k' \xrightarrow[p.s.]{t \rightarrow \infty} E(\xi_0 \xi_0'),$$

et

$$t^{-1} \sum_{k=1}^t \xi_k' \alpha^* \xi_k \eta_k \xrightarrow[p.s.]{t \rightarrow \infty} E(\xi_0' \alpha^* \xi_0 \eta_0) = E(\xi_0' \alpha^* \xi_0) E(\eta_0) = 0,$$

ce qui montre que

$$\tilde{\alpha}_t \xrightarrow[p.s.]{t \rightarrow \infty} \alpha^*.$$

Le même raisonnement peut être appliqué pour établir la convergence presque sûre de l'estimateur d'intérêt $\hat{\alpha}_t$.

2.5 Etude de simulation

Dans cette section nous étudions la performance de notre algorithme ($2S\text{-}RLS$) par le biais de données simulées. Le tableau 2.5.1 compare les résultats obtenus par notre algorithme $2S\text{-}RLS$ et ceux de la méthode des moindres carrés en deux étapes ($2S\text{-}LS$) de Bose et Mukherjee (2003) (noté ici $B\&M$), sur plusieurs séries générées à partir d'un modèle $ARCH(1)$ gaussien avec paramètres $(\alpha_0, \alpha_1) = (0.1, 0.3)$.

Pour toutes les simulations nous considérons 1000 répliques de Monte Carlo. Les valeurs de démarrage ont été fixées à

$$\begin{aligned}\tilde{\alpha}_0 &= \hat{\alpha}_0 = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \\ \tilde{P}_0 &= \hat{P}_0 = 100000 \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.\end{aligned}$$

La moyennes des estimations obtenues par les deux algorithmes et (entre parenthèses) leurs écart-types correspondants sont reportés dans la table suivante.

Data length	$\hat{\alpha}_0^{(2S-RLS)}$	$\hat{\alpha}_1^{(2S-RLS)}$	$\hat{\alpha}_0^{(B\&M)}$	$\hat{\alpha}_1^{(B\&M)}$
40	.0902 (.0112)	.2923 (.0737)	.1066 (.0103)	.2831 (.0608)
60	.0912 (.0092)	.3054 (.0493)	.1039 (.0084)	.2869 (.0465)
80	.0942 (.0085)	.2953 (.0409)	.1037 (.0072)	.2891 (.0399)
100	.0946 (.0047)	.2974 (.0369)	.1024 (.0043)	.2900 (.0300)

Tableau 2.5.1

Nous observons que même pour de petits échantillons, l'estimateur *2S-RLS* donne de bons résultats (sur la base de l'écart type des estimations) et est comparable à son homologue de *B&M*. Pour des séries de taille modérées l'estimateur *2S-RLS* semble plus précis que le *2S-RLS* et ceux grâce à sa capacité de contrôler la positivité des estimations. Pour la procédure de contrôle, nous avons appliqué la méthode de Ljung et Söderström (1983) qui teste l'admissibilité de l'estimation courante et la projette en cas de rejet dans la région de positivité. La procédure de projection consiste en une réduction successive du terme de mise à jour en le multipliant par une constante entre 0 et 1.

2.6 Conclusion

Dans ce chapitre, nous avons proposé un algorithme récursif, nommé *2S-RLS*, pour estimer les paramètres de modèles *ARCH*. Cette méthode, inspirée des algorithmes d'estimation en-ligne, exclusivement utilisés pour des processus linéaires, s'étend au cas de modèles *ARCH* et donne des estimateurs asymptotiquement gaussiens. Bien entendu, la méthode proposée peut aussi être appliquée à des données hors-ligne sur la base desquelles, elle garantit un bon contrôle des estimations afin de les maintenir dans les régions de positivité et de stationnarité. La performance de l'algorithme *2S-RLS* a été montrée dans une étude de simulation intensive sur de nombreux modèles répondant à la formulation *ARCH* et, cela, pour diverses tailles d'échantillons. Ainsi, l'algorithme *2S-RLS* est recommandé aussi bien pour sa simplicité et sa précision que pour son caractère séquentiel qui s'adapte aux données en temps réel telles, par exemple, les modèles de données financières en temps continu pour lesquelles les modèles *ARCH* en sont de bonnes approximations. Il est clair que l'hypothèse d'existence des moments d'ordre douze *H3* paraît contraignante mais cela peut être un point de départ pour d'éventuelles améliorations. Des extensions de cet algorithme à des modèles plus généraux tels que les modèles *GARCH* peuvent être développées à l'aide de l'approche pseudo-régression linéaire (Solo, 1979 et Duflo, 1990) que nous abordons dans le prochain chapitre consacré à l'estimation récursive des modèles *GARCH*.

2.7 Appendice

Lemme 2.5.1 (La formule de Wooldbury)

Soit A , B et C trois matrices de dimensions compatibles de sorte que la somme $A + BCD$ ait un sens ; alors :

$$[A + BCD]^{-1} = A^{-1} - A^{-1}B [DA^{-1}B + C^{-1}]^{-1} DA^{-1}. \quad (2.5.1)$$

■

Chapitre 3

Estimation récursive en-ligne des modèles $GARCH$

3.1 Introduction

Les modèles lineaires à coefficients constants classiques, fondés sur l'hypothèse que la variance des erreurs est constante, ne peuvent pas gérer la volatilité instantanée qui caractérise, en particulier, les séries financières (taux de change, taux d'inflation, indices boursiers,...). La classe des modèles $ARCH$, introduite par Engle (1982) s'est alors imposée comme alternative attrayante et fructueuse. En effet, les modèles $ARCH$ ont vite connu un développement conséquent avec l'apparition des modèles $GARCH$ (Bollerslev, 1986) et leur forme similaire à celle des modèles $ARMA$ classiques, susceptibles d'absorber des dynamiques de moyenne mobile, qui ont propulsé l'adoption de cette nouvelle classe de modèles par la maîtrise des propriétés probabilistes inspirées de celles des modèles $ARMA$ à savoir les propriétés de causalité et d'inversibilité, l'existence d'une solution strictement stationnaire (propriété LAN) et l'existence des moments d'ordre finis. Diverses extensions en ont résulté ($EGARCH$, $TARCH$, $FIGARCH$), sans oublier de mentionner les modèles à coefficients dépendant du temps et en particulier périodiques (Bentarzi et Aknouche (2005), Bentarzi et Aknouche (2006), Aknouche (2006), et références *intra*).

Les méthodes d'estimation existant, pour de tels modèles (Pantula (1988), Lumsdaine (1996), Bose and Mukherjee (2003), Ling and McAleer (2003), Berkes *et al* (2003, 2004), Francq and Zakoian (2004),...) sont fondées sur une taille d'échantillon préalablement fixée ; comme mentionné dans le chapitre 2, ceci ne permet pas de prendre en charge le caractère séquentiel de certaines données ; les données financières constituent un bon exemple puisqu'elles sont progressivement mises à jour et sont donc disponibles en temps réel.

L'objectif de ce chapitre est de développer quelques algorithmes récurrents (en-ligne), basés sur le schémas en deux étapes pour l'estimation d'un modèle *GARCH*. Ces algorithmes, inspirés du domaine du contrôle adaptatif fournissent des estimateurs asymptotiquement Gaussiens et convergeants presque sûrement. De plus, ils paraissent insensibles à l'hypothèse restrictive de normalité conditionnelle. Dans un premier temps, nous généralisons l'algorithme *2S-RLS* développé au chapitre précédent pour des modèles *ARCH*, au cas de modèles *GARCH* et ce, à travers l'approche pseudo régression linéaire (Pseudo Linear regression, *PLR*, voir). L'algorithme proposé nommé *PLR* en deux étapes (*2S-PLR*) parallèle les algorithmes de type *PLR* disponibles pour des modèles linéaires à coefficients constants. Bien qu'il donne des estimateurs convergeants presque sûrement, la distribution asymptotique des estimateurs fourni par cet algorithme n'est cependant pas connue et dont la matrice de covariance est strictement supérieure à la borne de Rao-Cramer (voir Ljung et Söderström 1983 dans le cas linéaire). Ainsi, nous contournons ce problème en proposant un autre algorithme récurrent via l'approche de l'erreur de prédiction récursive (Recursive Predictor Error Method, *RPEM*) qui fourni des estimateurs convergeants presque sûrement et asymptotiquement gaussiens. L'algorithme ainsi proposé qu'on nomme *RPE* en deux étapes (*2S-RPE*) est une adaptation de l'algorithme traditionnel du maximum de vraisemblance récursif (Recursive Maximum Likelihood, *RLM*) aux modèles *GARCH*.

Le reste de chapitre est organisé comme suit. La section 2 généralise l'algorithme *2S-RLS* proposé dans le chapitre précédent, à travers l'approche pseudo régression linéaire, pour estimer des modèles *GARCH*. En se basant sur l'approche de la fonction de Lyapunov stochastique (cf, Solo, 1979 et Moore et Ledwich, 1980), nous montrons la convergence presque sûre des estimateurs fournis. La section 5 développe un algorithme *2S-RPE* basé sur la méthode *RPE*. Nous montrons que les estimateurs obtenus sont fortement convergeants

et asymptotiquement gaussiens. Enfin la performance des algorithmes proposés est montrée via une étude de simulation. Les démonstrations des théorèmes proposés sont données en annexe du chapitre.

3.2 Pseudo régression linéaire récursive pour des modèles GARCH

Dans cette section nous adaptons l'approche pseudo-régression linéaire (voir par exemple, Panuska, 1968 ; Young, 1974 ; Ljung, 1977a ; Solo, 1979 ; Ljung and Söderström, 1983 ; Goodwin and Sin, 1984 ; Duffo, 1990 ; Adams and Goodwin, 1995 ; Ljung, 1999) pour l'estimation récursive des modèles à composantes moyenne mobile au cas de modèles *GARCH*. Il s'agit d'écrire un modèle *GARCH* sous forme de régression linéaire (par rapport au carré du processus $\{\varepsilon_t^2, t \in \mathbb{N}\}$) dans laquelle le régresseur contient des composantes non observables et alors estimer, récursivement, par la méthode des moindres carrés ordinaire, et les paramètres du modèle et les composantes non observables. Cependant si les estimateurs fournis par la méthode des moindres carrés ordinaires sont fortement convergents, ils ne sont pas asymptotiquement efficaces, même sous l'hypothèse de normalité conditionnelle. Cela est dû au fait que les innovations, fonctions de la variance conditionnelle, sont corrélées avec le passé de la variable dépendante ε_t^2 . C'est pourquoi nous faisons appel à la méthode des moindres carrés en deux étapes basée sur les travaux de Pantula (1988) et Bose et Mukherjee (2003).

Considérons le processus $\{\varepsilon_t; t \in \mathbb{Z}\}$ supposé admettre la représentation *GARCH*(p, q) suivante

$$\begin{cases} \varepsilon_t = \sqrt{h_t} \zeta_t, \\ E(\varepsilon_t^2 / \mathcal{F}_{t-1}) = h_t = \alpha_0 + \sum_{i=1}^q \alpha_i \varepsilon_{t-i}^2 + \sum_{j=1}^p \beta_j h_{t-j}, \quad t \in \mathbb{Z}, \end{cases} \quad (3.2.1)$$

où $\{\zeta_t, t \in \mathbb{Z}\}$ est une suite de variables aléatoires indépendantes et identiquement distribuées (*i.i.d.*) de moyenne nulle et de variance unité.

Posons $v_t = \varepsilon_t^2 - h_t$ et récrivons (3.2.1) comme suit

$$\varepsilon_t^2 = \alpha_0 + \sum_{i=1}^q \alpha_i \varepsilon_{t-i}^2 + v_t + \sum_{j=1}^p \beta_j (\varepsilon_{t-j}^2 - v_{t-j}). \quad (3.2.2)$$

Soit θ le $(p + q + 1) \times 1$ -vecteur des paramètres du modèle donné par $\theta = (\alpha_0, \alpha', \beta')'$ où $\alpha = (\alpha_1, \alpha_2, \dots, \alpha_q)'$ et $\beta = (\beta_1, \beta_2, \dots, \beta_p)'$.

Etant donnée une série chronologique $\varepsilon^t = \{\varepsilon_1, \dots, \varepsilon_t\}$ de taille t non fixée à l'avance, nous considérons le critère des moindres carrés pondéré suivant

$$V_t(\theta) = \frac{1}{2} \sum_{k=1}^t \frac{v_k^2(\theta)}{h_k^2(\theta)}, \quad (3.2.3)$$

où l'inclusion des poids $\frac{1}{h_k^2(\theta)}$ nous semble adéquate puisque le carré de l'innovation $v_t^2(\theta^*) = v_t^2 = h_t^2(\zeta^2 - 1)^2$ (θ^* étant la vraie valeur du paramètre θ) est conditionnellement corrélé avec le passé de ε_t^2 . Ainsi, l'algorithme que nous proposons pour l'estimation récursive de θ a pour objectif de minimiser l'expression (3.2.3) avec la plus simple complexité numérique possible.

Pour $t \in \mathbb{Z}$, soit φ_t le $(p + q + 1) \times 1$ "pseudo régresseur" défini par

$$\varphi_t = (1, \varepsilon_{t-1}^2, \dots, \varepsilon_{t-q}^2, \varepsilon_{t-1}^2 - v_{t-1}, \dots, \varepsilon_{t-p}^2 - v_{t-p})', \quad (3.2.4)$$

Alors le modèle *GARCH* (3.2.1) peut se mettre sous la forme pseudo-régression linéaire

$$\varepsilon_t^2 = \varphi_t' \theta + v_t, \quad (3.2.5)$$

dans laquelle le régresseur φ_t contient des innovations non observables. Donc le paramètre θ ne peut être estimé par la méthode des moindres carrés ordinaire à moins que l'on connaisse les composantes non observables. Cependant si l'on suppose pour un moment que les innovations dans (3.2.5) sont connues, le modèle (3.2.5) peut être vu comme une régression linéaire, à erreurs conditionnellement hétéroscédastiques, pour laquelle on peut faire appel à l'algorithme *RLS* en deux étapes (*2S-RLS*) (cf, Chapitre 2). On obtient d'abord dans la première étape, l'estimation préliminaire $\tilde{\theta}_t$ correspondant à la solution des moindres carrés récursive lorsqu'on ignore l'hétéroscédasticité conditionnelle des erreurs. Cette estimation est alors utilisée pour obtenir l'estimateur d'intérêt $\hat{\theta}_t$ tenant compte de l'hétéroscédasticité conditionnelle au moyen des poids appropriés. D'après le chapitre 2, on obtient le système

d'équations récurrentes suivant

$$\left\{ \begin{array}{l} \tilde{v}_t = \varepsilon_t^2 - \tilde{\theta}'_{t-1} \varphi_t, \\ \tilde{\theta}_t = \tilde{\theta}_{t-1} + \tilde{R}_t^{-1} \varphi_t \tilde{v}_t, \\ \tilde{R}_t = \tilde{R}_{t-1} + \varphi_t \varphi'_t, \\ \hat{v}_t = \varepsilon_t^2 - \hat{\theta}'_{t-1} \varphi_t, \\ \hat{\theta}_t = \hat{\theta}_{t-1} + \hat{R}_t^{-1} \varphi_t \frac{\hat{v}_t}{\left(\varphi'_t \tilde{\theta}_t\right)^2}, \\ \hat{R}_t = \hat{R}_{t-1} + \frac{\varphi_t \varphi'_t}{\left(\varphi'_t \tilde{\theta}_t\right)^2}, \end{array} \right. \quad (3.2.6)$$

où $\tilde{R}_t$ est une approximation de la matrice de plan associée au problème de regression (3.2.5) par arpport à $\{\varepsilon_k^2, k = 1, \dots, t\}$ et $\hat{R}_t$ est également la matrice de plan lorsqu'on dévise les deux membres de (3.2.5) par les poids $\frac{1}{h_k(\tilde{\theta}_k)}$.

Jusque là, nous avons supposé que les innovations (v_t) sont connues. Cependant, comme elles sont inconnues en pratique, nous pouvons les estimer en utilisant une approximation de l'expression (3.2.2) dans laquelle la vraie valeur du paramètre θ est remplacée, dans la première étape, par son estimation préliminaire $\tilde{\theta}_t$ obtenue à partir de l'algorithme (3.2.6) et, dans la seconde étape, par l'estimateur d'intérêt $\hat{\theta}_t$. Soit $\tilde{\vartheta}_t$ et $\hat{\vartheta}_t$ l'estimation préliminaire et l'estimation d'intérêt de l'innovation v_t lorsque θ dans (3.2.2) est remplacé respectivement par $\tilde{\theta}_t$ et $\hat{\theta}_t$. Alors $\tilde{\vartheta}_t$ et $\hat{\vartheta}_t$ sont données par

$$\tilde{\vartheta}_t = \varepsilon_t^2 - \tilde{\varphi}'_t \tilde{\theta}_t, \quad (3.2.7a)$$

$$\hat{\vartheta}_t = \varepsilon_t^2 - \hat{\varphi}'_t \hat{\theta}_t, \quad (3.2.7b)$$

où $\tilde{\varphi}_t$ et $\hat{\varphi}_t$ sont des estimations du regressor φ_t , obtenues de (3.2.46) lorsqu'on remplace v_t par $\tilde{\vartheta}_t$ et $\hat{\vartheta}_t$ respectivement. En utilisant ces notations et combinant l'estimation des composantes non observables avec l'estimation des paramètres, nous pouvons énoncé l'algorithme suivant que nous nommons *PLR* en deux étapes (*2S-PLR*) pour l'estimation récursive des paramètres *GARCH*.

Algorithme 3.2.1 (Aknouche et Guerbyenne, 2006b)

L'algorithme pseudo regression linéaire en deux étapes ($2S - PLR$) pour l'estimation d'un modèle GARCH est donné par les équations suivantes

$$\begin{array}{lcl}
 \begin{array}{l}
 \tilde{v}_t = \varepsilon_t^2 - \tilde{\varphi}_t' \tilde{\theta}_{t-1}, \\
 \tilde{\theta}_t = \tilde{\theta}_{t-1} + \frac{\tilde{P}_{t-1} \tilde{\varphi}_t \tilde{v}_t}{1 + \tilde{\varphi}_t' \tilde{P}_{t-1} \tilde{\varphi}_t}, \tilde{\theta}_{-1} = 0, \\
 \tilde{P}_t = \tilde{P}_{t-1} - \frac{\tilde{P}_{t-1} \tilde{\varphi}_t \tilde{\varphi}_t' \tilde{P}_{t-1}}{1 + \tilde{\varphi}_t' \tilde{P}_{t-1} \tilde{\varphi}_t}, \tilde{P}_{-1} = MI
 \end{array}
 & \rightarrow &
 \begin{array}{l}
 \hat{v}_t = \varepsilon_t^2 - \hat{\varphi}_t' \hat{\theta}_{t-1}, \\
 \hat{\theta}_t = \hat{\theta}_{t-1} + \frac{\hat{P}_{t-1} \hat{\varphi}_t \hat{v}_t}{\left(\tilde{\varphi}_t' \tilde{\theta}_{t-1}\right)^2 + \hat{\varphi}_t' \hat{P}_{t-1} \hat{\varphi}_t}, \hat{\theta}_0 = 0, \\
 \hat{P}_t = \hat{P}_{t-1} - \frac{\hat{P}_{t-1} \hat{\varphi}_t \hat{\varphi}_t' \hat{P}_{t-1}}{\left(\tilde{\varphi}_t' \tilde{\theta}_{t-1}\right)^2 + \hat{\varphi}_t' \hat{P}_{t-1} \hat{\varphi}_t}, \hat{P}_0 = MI,
 \end{array}
 \end{array}$$

première étape
seconde étape

(3.2.8)

où $\tilde{P}_t = \tilde{R}_t^{-1}$, $\hat{P}_t = \hat{R}_t^{-1}$ et M est un nombre positive suffisamment grand. ■

Notons que dans les poids $\left(\tilde{\varphi}_t' \tilde{\theta}_{t-1}\right)^2$ nous avons utilisé $\tilde{\theta}_{t-1}$ plutôt que $\tilde{\theta}_t$ et ce pour être dans le cadre de la théorie des martingale que nous utiliserons pour la convergence des estimateurs obtenus.

Pour étudier la propriété de convergence des estimateurs $2S-PLR$, nous pouvons utiliser l'une des approches connues dans la littérature de l'estimation en-ligne, telles que l'approche de l'équation différentielle ordinaire (Ljung 1977a; 1977b), l'approche de Hannan (1976, 1980) ou l'approche de la fonction de Lyapunov stochastique due à Moore et Ledwich (1980) et Solo (1979). Nous avons choisi d'adopter cette dernière approche puisqu'elle assure la convergence des estimateurs sans avoir besoin de l'hypothèse de contrôle de l'estimateur pour le garder dans la région de stabilité (stationnarité, positivité). Considérons les hypothèses suivantes :

$H1$: Le processus GARCH $\{\varepsilon_t\}$ donné par (2.3.1) est stationnaire du second ordre et ergodique.

$H2$: Le processus GARCH $\{\varepsilon_t\}$ donné par (2.3.1) a ses moments d'ordre 4 finis et l'on a

$$\lim_{t \rightarrow \infty} \sup \frac{1}{t} \sum_{k=1}^t \varepsilon_k^4 < \infty, \quad p.s. \tag{3.2.9}$$

$H3$: Il existe $K > 0$ tel que $h_t < K$, $p.s.$

$H4$: Le filtre $1 - \beta(L)$ est stable et l'on a

$$\operatorname{Re} \left[(1 - \beta(e^{i\omega}))^{-1} - \frac{1}{2} \right] > 0, \quad \forall \omega \in]-\pi, \pi[.$$

La condition de positivité de la partie réelle $H4$ est équivalente à la suivante

$H'4$: Le filtre $1 - \beta(L)$ est stable et l'on a

$$|\beta(e^{i\omega})| < 1, \quad \forall \omega \in]-\pi, \pi[.$$

Cette condition est plus facile à vérifier (voir Ljung, 1977 pour d'autres discussions).

Pour le modèle $GARCH(1,1)$ avec filtre $1 - \beta(L)$ stable, la condition 4 est toujours vérifiée, i.e. $H4$ est vérifiée pourvu que le modèle soit stationnaire au second ordre.

Sous les hypothèses précédentes nous pouvons montrer le théorème suivant qui montre la convergence presque sûre des estimateurs.

Théorème 3.2.1 (*Aknouche Guerbyenne, 2006b*)

Les estimateurs $2S-PLR$ $\tilde{\theta}_t$ et $\hat{\theta}_t$ convergent presque sûrement vers θ^* , dans le sens suivant

i)

$$\frac{1}{t} \left(\tilde{\theta}_t - \theta^* \right)' \tilde{R}_t \left(\tilde{\theta}_t - \theta^* \right) \xrightarrow{p.s.} 0, \quad \text{lorsque } t \rightarrow \infty, \quad (3.2.10a)$$

$$\frac{1}{t} \sum_{k=1}^t \left(\tilde{\vartheta}_k - v_k \right)^2 \xrightarrow{p.s.} 0, \quad \text{lorsque } t \rightarrow \infty, \quad (3.2.10b)$$

et

ii)

$$\frac{1}{t} \left(\hat{\theta}_t - \theta^* \right)' \hat{R}_t \left(\hat{\theta}_t - \theta^* \right) \xrightarrow{p.s.} 0, \quad \text{lorsque } t \rightarrow \infty, \quad (3.2.11a)$$

$$\frac{1}{t} \sum_{k=1}^t \left(\hat{\vartheta}_k - v_k \right)^2 \xrightarrow{p.s.} 0, \quad \text{lorsque } t \rightarrow \infty. \quad (3.2.11b)$$

Preuve Voir appendice du chapitre.

Les conclusions du théorème précédent n'impliquent pas directement que les estimateurs sont convergents. Cependant si nous rajoutons l'hypothèse suivante

$H5$: Les matrices limites

$$\tilde{R} = \lim_{t \rightarrow \infty} \frac{1}{t} \sum_{k=1}^t \tilde{\varphi}_k \tilde{\varphi}_k' \quad (3.2.12a)$$

et

$$\hat{R} = \lim_{t \rightarrow \infty} \frac{1}{t} \sum_{k=1}^t (\tilde{\varphi}_k' \theta_{k-1})^{-2} \hat{\varphi}_k \hat{\varphi}_k' \quad (3.2.12b)$$

existent, alors on a le corollaire suivant.

Corollaire 2.3.1

Sous les hypothèses $H1 - H5$ on a

$$\tilde{\theta}_t \xrightarrow{p.s.} \theta^*,$$

$$\hat{\theta}_t \xrightarrow{p.s.} \theta^*,$$

lorsque $t \rightarrow \infty$.

Preuve Nous ométons la preuve qui est tout à fait similaire à celles de Solo (1979) et de Ljung et Söderström (1983). ■

L'hypothèse $H3$ selon laquelle la variance conditionnelle est bornée semble très restrictive puisqu'elle exclut une large gamme de processus, tels que les processus gaussiens. En revanche, $H3$ n'est pas nécessaire pour le théorème et peut être omise. Nous l'avons donnée seulement pour simplifier la preuve.

3.3 Algorithme d'erreur de prédiction récursive en deux étapes

Alors que l'algorithme $2S-PLR$ donne des estimateurs convergents presque sûrement sans avoir besoin de contrôler les estimations, il a deux limites majeures, à savoir : la convergence sous l'hypothèse de positivité de la partie réelle et l'inefficacité asymptotique. En effet, Ljung

and Söderström (1983, p. 231) ont montré sur un contre exemple, sans donner la distribution asymptotique, que l'estimateur PLR n'est pas asymptotiquement efficace. Pour remédier à ces inconvénients nous proposons (voir aussi Ljung 1977a dans le cas de modèles linéaires) de remplacer les pseudo régresseurs $\tilde{\varphi}_t$ et $\hat{\varphi}_t$, dans (3.2.8) par les vecteurs $\tilde{\psi}_t$ et $\hat{\psi}_t$ obtenus à travers un filtre linéaire que nous expliciterons plus bas. Cela conduit à une procédure dont l'essence est similaire à la méthode bien connue du maximum de vraisemblance récursive (Recursive Maximum Likelihood RML , voir Ljung, 1977a).

Avons d'exposer notre algorithme, nous montrons d'abord comment calculer l'erreur de prévision sous forme récursive. La prédiction de ε_t^2 basée sur l'information disponible à l'instant $t - 1$ est donnée par

$$\hat{\varepsilon}_t(\theta) = \varepsilon_t^2 - v_t(\theta), \quad t \in \mathbb{Z}, \quad (3.3.1)$$

où $v_t(\theta)$ est la fonction erreur de prédiction qui peut être évaluée récursivement en termes de ε_{t-i}^2 , $i = 1, \dots, q$ et $v_{t-j}(\theta)$, $j = 1, \dots, p$, à partir de l'équation aux différences stochastique suivante

$$v_t(\theta) = \varepsilon_t^2 - \alpha_0 - \sum_{i=1}^q \alpha_i \varepsilon_{t-i}^2 + \sum_{j=1}^p \beta_j (\varepsilon_{t-j}^2 - v_{t-j}(\theta)), \quad t \in \mathbb{Z}. \quad (3.3.2)$$

Puisque l'effet des valeurs initiales de (3.3.2) n'est pas persistant pour un modèle causal, on peut les fixer arbitrairement sans affecter les prévisions (voir par exemple Francq and Zakoïan, 2004 pour la fixation des valeurs initiales). En combinant (3.3.1) et (3.3.2), $\hat{\varepsilon}_t(\theta)$ peut être écrite sous la forme

$$\hat{\varepsilon}_t(\theta) = \sum_{j=1}^p \beta_j \hat{\varepsilon}_{t-j}(\theta) + \alpha_0 + \sum_{i=1}^q \alpha_i \varepsilon_{t-i}^2, \quad t \geq 1, \quad (3.3.3)$$

où les valeurs de démarrage sont fixées.

Dérivons les deux membres de (3.3.3) par rapport à θ et posons

$$\psi_t(\theta) = \frac{\partial \hat{\varepsilon}_t(\theta)}{\partial \theta'},$$

avec

$$\psi_t(\theta^*) = -\frac{\partial v_t}{\partial \theta'},$$

nous obtenons

$$\psi_t(\theta) - \sum_{s=1}^p \beta_s \psi_{t-s}(\theta) = \varphi_t(\theta), \quad (3.3.4)$$

où

$$\varphi_t(\theta) = (1, \varepsilon_{t-1}^2, \varepsilon_{t-2}^2, \dots, \varepsilon_{t-q}^2, \varepsilon_{t-1}^2 - v_{t-1}(\theta), \dots, \varepsilon_{t-p}^2 - v_{t-p}(\theta))'.$$

Pour évaluer (3.3.4) nous approximations le gradient $\psi_t(\theta)$ par les estimations $\tilde{\psi}_t$ et $\hat{\psi}_t$ obtenues en remplaçant respectivement θ par $\tilde{\theta}_t$ et $\hat{\theta}_t$. Ce qui donne

$$\tilde{\psi}_t - \sum_{s=1}^p \tilde{\beta}_s^{(t-1)} \tilde{\psi}_{t-s} = \tilde{\varphi}_t, \quad (3.3.5a)$$

et

$$\hat{\psi}_t - \sum_{s=1}^p \hat{\beta}_s^{(t-1)} \hat{\psi}_{t-s} = \hat{\varphi}_t. \quad (3.3.5b)$$

En utilisant les notations précédentes nous pouvons construire l'algorithme *RPE*-type en deux étapes suivant (voir Ljung et Söderström 1983, p. 26 pour la formulation).

Algorithm 3.3.1 (Aknouche et Guerbyenne, 2006b)

L'algorithme du *RPE* en deux étapes pour l'estimation d'un modèle GARCH est donné par

$$\underbrace{\left\{ \begin{array}{l} \tilde{v}_t = \varepsilon_t^2 - \tilde{\varphi}_t' \tilde{\theta}_{t-1}, \\ \tilde{\theta}_t = \tilde{\theta}_{t-1} + \frac{\tilde{P}_{t-1} \tilde{\psi}_t \tilde{v}_t}{1 + \tilde{\psi}_t' \tilde{P}_{t-1} \tilde{\psi}_t}, \tilde{\theta}_{-1} = 0, \\ \tilde{P}_t = \tilde{P}_{t-1} - \frac{\tilde{P}_{t-1} \tilde{\psi}_t \tilde{\psi}_t' \tilde{P}_{t-1}}{1 + \tilde{\psi}_t' \tilde{P}_{t-1} \tilde{\psi}_t}, \tilde{P}_{-1} = M.I, \\ \tilde{\psi}_{t+1} = \sum_{k=1}^q \tilde{\beta}_k^{(t)} \tilde{\psi}_{t-k+1} + \tilde{\varphi}_{t+1} \end{array} \right\}}_{\text{première étape}}; \underbrace{\left\{ \begin{array}{l} \hat{v}_t = \varepsilon_t^2 - \hat{\varphi}_t' \hat{\theta}_{t-1}, \\ \hat{\theta}_t = \hat{\theta}_{t-1} + \frac{\hat{P}_{t-1} \hat{\psi}_t \hat{v}_t}{(\tilde{\varphi}_t' \tilde{\theta}_{t-1})^2 + \hat{\psi}_t' \hat{P}_{t-1} \hat{\psi}_t}, \hat{\theta}_0 = 0, \\ \hat{P}_t = \hat{P}_{t-1} - \frac{\hat{P}_{t-1} \hat{\psi}_t \hat{\psi}_t' \hat{P}_{t-1}}{(\tilde{\varphi}_t' \tilde{\theta}_{t-1})^2 + \hat{\psi}_t' \hat{P}_{t-1} \hat{\psi}_t}, \hat{P}_0 = M.I, \\ \hat{\psi}_{t+1} = \sum_{k=1}^q \hat{\beta}_k^{(t)} \hat{\psi}_{t-k+1} + \hat{\varphi}_{t+1}, \end{array} \right\}}_{\text{seconde étape}} \quad (3.3.6)$$

où $\tilde{\varphi}_t$ et $\hat{\varphi}_t$ sont les estimations du regressor φ_t , obtenues de (3.2.46) lorsqu'on remplace v_t par $\tilde{v}_t$ et $\hat{v}_t$ respectivement.

Concernant la convergence des estimateurs fournis par l'algorithme 3.3.1, on note qu'en utilisant $\tilde{\psi}_t$ et $\hat{\psi}_t$ à la place de $\tilde{\varphi}_t$ et $\hat{\varphi}_t$, respectivement, la convergence presque sûre des estimateurs *2S-RPE* est assurée de la même manière que celle dans le théorème 3.3.1 mais en considérant l'hypothèse *H7* (voir ci-dessous) à la place de *H4*. Cependant le présent algorithme doit être contrôlé pour assurer que l'estimation courante ne sorte de la région de stabilité. Considérons les hypothèses suivantes :

H6 : On suppose que l'algorithme (3.3.6) inclut une procédure de projection permettant de garder $\tilde{\theta}_t$ et $\hat{\theta}_t$ dans la région de stabilité.

H7 :

$$\begin{aligned} \operatorname{Re} \left[(1 - \beta(e^{i\omega}))^{-1} (1 - \tilde{\beta}^{(t-1)}(e^{i\omega})) - \frac{1}{2} \right] &> 0 \quad \forall \omega \in]-\pi, \pi[\\ \operatorname{Re} \left[(1 - \beta(e^{i\omega}))^{-1} (1 - \hat{\beta}^{(t-1)}(e^{i\omega})) - \frac{1}{2} \right] &> 0 \quad \forall \omega \in]-\pi, \pi[. \end{aligned}$$

Les dernières hypothèses sont vérifiées une fois que $1 - \beta(L)$ est stable, où

$$\begin{aligned} \tilde{\beta}^{(t)}(L) &= \sum_{j=1}^p \tilde{\beta}_j^{(t)} L^j, \\ \hat{\beta}^{(t)}(L) &= \sum_{j=1}^p \hat{\beta}_j^{(t)} L^j. \end{aligned}$$

En plus de la convergence presque sûre, nous pouvons garantir la normalité asymptotique des estimateurs *2S-RPE* sous l'hypothèse additionnelle suivante

H8 : Il existe $K^* > 0$ tel que $|\varepsilon_t^2| < K^*$ *p.s.*

Alors nous avons le résultat suivant

Théorème 3.3.2 (Akouché et Guerbyenne, 2006b)

Sous les hypothèses *H1* – *H3* et *H5* – *H8*, on a

$$\begin{aligned} t^{1/2}(\tilde{\theta}_t - \theta^*) &\xrightarrow{L} N \left(0, \{E(\psi_0(\theta^*)\psi_0'(\theta^*))\}^{-1} E(\psi_0(\theta^*)\psi_0'(\theta^*)v_0^2) \{E(\psi_0(\theta^*)\psi_0'(\theta^*))\}^{-1} \right) \\ t^{1/2}(\hat{\theta}_t - \theta^*) &\xrightarrow{L} N \left(0, \operatorname{Var}(\eta_0^2) \{E(\psi_0(\theta^*)\psi_0'(\theta^*)h_0^{-2}(\theta^*))\}^{-1} \right). \end{aligned} \quad (3.3.8)$$

où " $\xrightarrow{L}$ " désigne la convergence en loi.

Preuve Voir appendice du chapitre.

3.4 Simulation

Dans le tableau 3.4.1 nous évaluons la performance des algorithmes $2S-PLR$ et $2S-RML$, en se basant sur des données simulées, générées d'un modèle $GARCH(1, 1)$ gaussien avec paramètres

$$(\alpha_0, \alpha_1, \beta_1) = (.3, .05, .8).$$

Nous considérons 1000 répliques de Monte Carlo avec valeurs de démarrage fixées à

$$\begin{aligned} \tilde{\beta}_0 &= \hat{\beta}_0 = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \\ \tilde{P}_0 &= \hat{P}_0 = 100000 \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}. \end{aligned}$$

Data length	$\hat{\alpha}_0^{(PL)}$	$\hat{\alpha}_1^{(PL)}$	$\hat{\beta}_1^{(PL)}$	$\hat{\alpha}_0^{(RM)}$	$\hat{\alpha}_1^{(RM)}$	$\hat{\beta}_1^{(RM)}$
80	.3371 (.0925)	.0622 (.0081)	.8541 (.1028)	.3445 (.1079)	.0653 (.0091)	.8465 (.0912)
300	.3213 (.0649)	.0588 (.0066)	.8341 (.0897)	.3208 (.0592)	.0547 (.0063)	.8265 (.0714)
600	.3105 (.0417)	.0518 (.0047)	.8241 (.0718)	.3086 (.0491)	.0509 (.0043)	.8165 (.0676)

Tableau 3.4.1

Pour des séries chronologiques de petite taille, nous observons que les estimateurs $2S-PLR$ et $2S-RML$ sont plus ou moins acceptables et que le $2S-PLR$ donne de meilleurs résultats. Cependant pour des séries modérées l'algorithme $2S-RML$ est meilleur que son homologue. Notons également que l'algorithme doit être contrôlé pour ne pas avoir des valeurs négatives dans les estimations. Dans le cas GARCH ces valeurs négatives sont plus présentes. Nous avons appliqué la procédure de projection de Ljung and Söderström (1983, p. 367).

3.5 Appendice : Preuves des théorèmes

3.5.1 Preuve du théorème 3.2.1

Preuve de la partie i)

Récrivons la deuxième et la troisième équation du système (3.2.6)

$$\tilde{\theta}_t = \tilde{\theta}_{t-1} + \tilde{R}_t^{-1} \tilde{\varphi}_t \tilde{v}_t \quad (3.5.1a)$$

$$\tilde{R}_t = \tilde{R}_{t-1} + \tilde{\varphi}_t \tilde{\varphi}_t' \quad (3.5.1b)$$

et soit $\bar{\theta}_t = \tilde{\theta}_t - \theta^*$, alors (3.5.1) peut s'écrire comme

$$\bar{\theta}_t = \bar{\theta}_{t-1} + \tilde{R}_t^{-1} \tilde{\varphi}_t \tilde{v}_t \quad (3.5.2a)$$

$$\tilde{R}_t = \tilde{R}_{t-1} + \tilde{\varphi}_t \tilde{\varphi}_t' \quad (3.5.2b)$$

Sachant que via le lemme d'inversion matricielle (cf, Lemme 2.3.1), on a

$$\frac{\tilde{R}_t^{-1} \tilde{\varphi}_t}{1 - \tilde{\varphi}_t' \tilde{R}_t^{-1} \tilde{\varphi}_t} = \tilde{R}_{t-1}^{-1} \tilde{\varphi}_t, \quad (3.5.2c)$$

et que

$$\begin{aligned} \tilde{\vartheta}_t &= \varepsilon_t^2 - \tilde{\varphi}_t' \tilde{\theta}_t, \\ &= \varepsilon_t^2 - \tilde{\varphi}_t' \tilde{\theta}_{t-1} - \tilde{\varphi}_t' \tilde{R}_t^{-1} \tilde{\varphi}_t \tilde{v}_t, \\ &= (1 - \tilde{\varphi}_t' \tilde{R}_t^{-1} \tilde{\varphi}_t) \tilde{v}_t, \end{aligned}$$

alors (3.5.2a) devient

$$\begin{aligned} \bar{\theta}_t &= \bar{\theta}_{t-1} + \frac{\tilde{R}_t^{-1} \tilde{\varphi}_t \tilde{\vartheta}_t}{1 - \tilde{\varphi}_t' \tilde{R}_t^{-1} \tilde{\varphi}_t}, \\ &= \bar{\theta}_{t-1} + \tilde{R}_{t-1}^{-1} \tilde{\varphi}_t \tilde{\vartheta}_t \end{aligned}$$

maintenant (3.5.2) peut s'écrire comme suit

$$\bar{\theta}_t = \bar{\theta}_{t-1} + \tilde{R}_{t-1}^{-1} \tilde{\varphi}_t \tilde{\vartheta}_t, \quad (3.5.3a)$$

$$\tilde{R}_t = \tilde{R}_{t-1} + \tilde{\varphi}_t \tilde{\varphi}_t' \quad (3.5.3b)$$

multiplions (3.5.3a) à gauche par $\tilde{R}_{t-1}$ alors on trouve

$$\tilde{R}_{t-1} \bar{\theta}_t = \tilde{R}_{t-1} \bar{\theta}_{t-1} + \tilde{\varphi}_t \tilde{\vartheta}_t, \quad (3.5.4)$$

Dans le membre gauche de (3.5.4), utilisons (3.5.3b), alors (3.5.4) devient

$$\tilde{R}_t \bar{\theta}_t - \tilde{\varphi}_t \tilde{\varphi}_t' \bar{\theta}_t = \tilde{R}_{t-1} \bar{\theta}_{t-1} + \tilde{\varphi}_t \tilde{\vartheta}_t, \quad (3.5.5)$$

multiplions maintenant (3.5.5), à gauche par $\bar{\theta}_t'$

$$\bar{\theta}_t' \tilde{R}_t \bar{\theta}_t = \bar{\theta}_t' \tilde{R}_{t-1} \bar{\theta}_{t-1} + \bar{\theta}_t' \tilde{\varphi}_t \tilde{\vartheta}_t + (\bar{\theta}_t' \tilde{\varphi}_t)^2, \quad (3.5.6)$$

utilisons (3.5.3a) dans le premier terme du membre droit de (3.5.6), cela donne

$$\bar{\theta}_t' \tilde{R}_t \bar{\theta}_t = \bar{\theta}_{t-1}' \tilde{R}_{t-1} \bar{\theta}_{t-1} + \tilde{\varphi}_t' \bar{\theta}_{t-1} \tilde{\vartheta}_t + \bar{\theta}_t' \tilde{\varphi}_t \tilde{\vartheta}_t + (\bar{\theta}_t' \tilde{\varphi}_t)^2 \quad (3.5.7)$$

En remplaçant la valeur de $\bar{\theta}_{t-1}$ de (3.5.3a) dans le second terme du membre droit de (3.5.7), alors on trouve

$$\begin{aligned} T_t &= T_{t-1} + \tilde{\varphi}_t' \bar{\theta}_t \tilde{\vartheta}_t - \tilde{\varphi}_t' \tilde{R}_{t-1}^{-1} \tilde{\varphi}_t \tilde{\vartheta}_t^2 + \bar{\theta}_t' \tilde{\varphi}_t \tilde{\vartheta}_t + (\bar{\theta}_t' \tilde{\varphi}_t)^2 \\ &= T_{t-1} + 2\tilde{\varphi}_t' \bar{\theta}_t \tilde{\vartheta}_t - \tilde{\varphi}_t' \tilde{R}_{t-1}^{-1} \tilde{\varphi}_t \tilde{\vartheta}_t^2 + (\bar{\theta}_t' \tilde{\varphi}_t)^2 \end{aligned} \quad (3.5.8)$$

où $T_t = \bar{\theta}_t' \tilde{R}_t \bar{\theta}_t$. Notons que (3.5.8) peut s'écrire comme suit

$$\begin{aligned} T_t &= T_{t-1} + 2\tilde{\varphi}_t' \bar{\theta}_t (\tilde{\vartheta}_t + \frac{1}{2} \tilde{\varphi}_t' \bar{\theta}_t) - \tilde{\varphi}_t' \tilde{R}_{t-1}^{-1} \tilde{\varphi}_t \tilde{\vartheta}_t^2 \\ &= T_{t-1} + 2\tilde{\varphi}_t' \bar{\theta}_t (\tilde{\vartheta}_t - v_t + \frac{1}{2} \tilde{\varphi}_t' \bar{\theta}_t) + 2\tilde{\varphi}_t' \bar{\theta}_t v_t - \tilde{\varphi}_t' \tilde{R}_{t-1}^{-1} \tilde{\varphi}_t \tilde{\vartheta}_t^2 \end{aligned}$$

d'où

$$T_t = T_{t-1} - 2a_t b_t + 2\tilde{\varphi}_t' \bar{\theta}_t v_t - \tilde{\varphi}_t' \tilde{R}_{t-1}^{-1} \tilde{\varphi}_t \tilde{\vartheta}_t^2, \quad (3.5.9)$$

où

$$a_t = -\tilde{\varphi}'_t \bar{\theta}_t, \quad (3.5.10a)$$

$$b_t = \tilde{\vartheta}_t - v_t + \frac{1}{2} \tilde{\varphi}'_t \bar{\theta}_t, \quad (3.5.10b)$$

Exploitions (3.5.2a) dans le 3ème terme du membre droit de (3.5.9), ce qui donne

$$T_t = T_{t-1} - 2a_t b_t + 2\tilde{\varphi}'_t \bar{\theta}_{t-1} v_t + 2\tilde{\varphi}'_t \tilde{R}_t^{-1} \tilde{\varphi}_t \tilde{v}_t v_t - \tilde{\varphi}'_t \tilde{R}_{t-1}^{-1} \tilde{\varphi}_t \tilde{\vartheta}_t^2,$$

ou encore

$$T_t = T_{t-1} - 2a_t b_t + 2\tilde{\varphi}'_t \bar{\theta}_{t-1} v_t + 2\tilde{\varphi}'_t \tilde{R}_t^{-1} \tilde{\varphi}_t (\tilde{v}_t - v_t) v_t + 2\tilde{\varphi}'_t \tilde{R}_t^{-1} \tilde{\varphi}_t v_t^2 - \tilde{\varphi}'_t \tilde{R}_{t-1}^{-1} \tilde{\varphi}_t \tilde{\vartheta}_t^2 \quad (3.5.11a)$$

puisque le dernier terme du second membre de (3.5.11a) est positif, alors on en déduit la relation suivante

$$T_t \leq T_{t-1} - 2a_t b_t + 2\tilde{\varphi}'_t \bar{\theta}_{t-1} v_t + 2\tilde{\varphi}'_t \tilde{R}_t^{-1} \tilde{\varphi}_t (\tilde{v}_t - v_t) v_t + 2\tilde{\varphi}'_t \tilde{R}_t^{-1} \tilde{\varphi}_t v_t^2 \quad (3.5.11b)$$

Sachant que les quantités : $\tilde{\varphi}'_t$, $\tilde{R}_t$, $\bar{\theta}_{t-1}$ et $\tilde{v}_t - v_t = \tilde{\varphi}'_t \bar{\theta}_{t-1} - \varphi'_t \theta^*$ sont $\mathcal{F}_{t-1}$ -mesurables, que $E(v_t/\mathcal{F}_{t-1}) = 0$ et enfin que

$$E(v_t^2/\mathcal{F}_{t-1}) = (E(\eta_t^4) - 1) h_t = K.h_t,$$

($K = E(\eta_t^4) - 1 \geq 0$ par Cauchy Schwartz), donc en prenant l'espérance conditionnelle de (3.5.11b) par rapport à $\mathcal{F}_{t-1}$, on trouve

$$\begin{aligned} E(T_t + 2a_t b_t/\mathcal{F}_{t-1}) &\leq T_{t-1} + 2\tilde{\varphi}'_t \bar{\theta}_{t-1} E(v_t/\mathcal{F}_{t-1}) + 2\tilde{\varphi}'_t \tilde{R}_t^{-1} \tilde{\varphi}_t (\tilde{v}_t - v_t) \times \\ &\quad E(v_t/\mathcal{F}_{t-1}) + 2\tilde{\varphi}'_t \tilde{R}_t^{-1} \tilde{\varphi}_t E(v_t^2/\mathcal{F}_{t-1}) \\ &= T_{t-1} + 2K h_t \tilde{\varphi}'_t \tilde{R}_t^{-1} \tilde{\varphi}_t \end{aligned} \quad (3.5.12)$$

Rajoutons aux deux membres de (3.5.12) la quantité $2 \sum_{k=1}^{t-1} a_k b_k$, pour trouver

$$E\left(T_t + 2 \sum_{k=1}^t a_k b_k/\mathcal{F}_{t-1}\right) \leq T_{t-1} + 2 \sum_{k=1}^{t-1} a_k b_k + 2K h_t \tilde{\varphi}'_t \tilde{R}_t^{-1} \tilde{\varphi}_t \quad (3.5.13)$$

Posons $\tilde{T}_t = \frac{1}{t} \left(T_t + 2 \sum_{k=1}^t a_k b_k \right)$ alors on peut réécrire (3.5.13) comme suit

$$E \left(t\tilde{T}_t / \mathcal{F}_{t-1} \right) \leq (t-1)\tilde{T}_{t-1} + 2Kh_t \tilde{\varphi}'_t \tilde{R}_t^{-1} \tilde{\varphi}_t,$$

ou encore

$$E \left(\tilde{T}_t / \mathcal{F}_{t-1} \right) \leq \tilde{T}_{t-1} - \frac{1}{t} \tilde{T}_{t-1} + \frac{2Kh_t}{t} \tilde{\varphi}'_t \tilde{R}_t^{-1} \tilde{\varphi}_t. \quad (3.5.14)$$

On rappelle maintenant le théorème de Robbins et Siegmund (1971) que nous allons appliquer (sous certaines hypothèses) à (3.5.14).

Lemme 3.5.1 (Robbins et Siegmund, 1971)

Soit $\{T_t\}$, $\{\alpha_{t+1}\}$, $\{\beta_{t+1}\}$ trois suites de variables aléatoires non négatives, adaptées à une filtration $\{\mathcal{F}_t\}$, telles que

$$E(T_t / \mathcal{F}_{t-1}) \leq T_{t-1} + \alpha_t - \beta_t,$$

et

$$\sum_{t=1}^{\infty} \alpha_t < \infty,$$

alors

$$T_t \rightarrow T \quad p.s.$$

et

$$\sum_{t=1}^{\infty} \beta_t < \infty.$$

Pour pouvoir appliquer ce théorème sur (3.5.14) avec

$$\alpha_t = \frac{2Kh_t}{t} \tilde{\varphi}'_t \tilde{R}_t^{-1} \tilde{\varphi}_t,$$

et

$$\beta_t = \frac{1}{t} \tilde{T}_{t-1},$$

montrons d'abord que

$$\tilde{T}_t \geq 0, \quad \forall t, \quad (3.5.15a)$$

(la quantité $\frac{2Kh_t}{t}\tilde{\varphi}'_t\tilde{R}_t^{-1}\tilde{\varphi}'_t$ est d'emblée positive) qui est vraie dès que

$$\sum_{k=1}^t a_k b_k \geq 0, \quad (3.5.15b)$$

et que

$$2K \sum_{t=1}^{\infty} \frac{h_t}{t} \tilde{\varphi}'_t \tilde{R}_t^{-1} \tilde{\varphi}_t < \infty, \quad (3.5.15c)$$

Or d'après (3.5.10) (i.e. la définition de a_t et b_t) on a :

$$b_t = \tilde{\vartheta}_t - v_t - 1/2a_t, \quad (3.5.16)$$

Par ailleurs

$$\begin{aligned} \tilde{\vartheta}_t - v_t &= \varepsilon_t^2 - \tilde{\varphi}'_t \tilde{\theta}_t - \varepsilon_t^2 + \varphi'_t \theta^*, \\ &= \varphi'_t \theta^* - \tilde{\varphi}'_t \bar{\theta}_t + \tilde{\varphi}'_t \theta^* - \tilde{\varphi}'_t \theta^*, \\ &= -\tilde{\varphi}'_t \bar{\theta}_t + (\varphi_t - \tilde{\varphi}'_t)' \theta^*, \\ &= -\tilde{\varphi}'_t \bar{\theta}_t + (\xi_t - \bar{\xi}_t)' A \theta^*, \\ &= -\tilde{\varphi}'_t \bar{\theta}_t + \sum_{j=1}^p \beta_j (\tilde{\vartheta}_{t-j} - v_{t-j}), \\ &= -\tilde{\varphi}'_t \bar{\theta}_t + \beta(L) (\tilde{\vartheta}_t - v_t), \end{aligned} \quad (3.5.17)$$

où $\beta(L) = \sum_{j=1}^p \beta_j L^j$. Ainsi, (3.5.17) peut se mettre comme suit

$$(\tilde{\vartheta}_t - v_t) = -(1 - \beta(L))^{-1} \tilde{\varphi}'_t \bar{\theta}_t, \quad (3.5.18a)$$

$$= (1 - \beta(L))^{-1} a_t \quad (3.5.18b)$$

en combinant entre (3.5.16) et (3.5.18), on obtient

$$b_t = ((1 - \beta(L))^{-1} - 1/2) a_t \quad (3.5.19a)$$

$$= H(L) a_t \quad (3.5.19a)$$

$$= \left(\sum_{s=1}^{\infty} H_s L^s \right) a_t \quad (3.5.19c)$$

Enfin, pour montrer (3.5.15b) on a

$$\begin{aligned}
 \sum_{k=1}^t a_k b_k &= \sum_{k=1}^t a_k \left(\sum_{s=0}^{\infty} H_s L^s \right) a_k, \\
 &= \sum_{k=1}^t a_k \sum_{s=0}^k H_{k-s} a_s \\
 &= \frac{1}{2\pi} \int_{-\pi}^{\pi} \left| \sum_{k=1}^t a_k e^{i\omega k} \right|^2 H(e^{i\omega}) d\omega \\
 &= \frac{1}{2\pi} \int_{-\pi}^{\pi} \left| \sum_{k=1}^t a_k e^{i\omega k} \right|^2 \operatorname{Re} (H(e^{i\omega})) d\omega \\
 &\geq \lambda \frac{1}{2\pi} \int_{-\pi}^{\pi} \left| \sum_{k=1}^t a_k e^{i\omega k} \right|^2 d\omega \\
 &= \lambda \sum_{k=1}^t a_k^2
 \end{aligned} \tag{3.5.20}$$

où

$$\lambda = \inf_{-\pi \leq \omega \leq \pi} [\operatorname{Re}(1 - \beta(e^{i\omega}))^{-1} - 1/2]. \tag{3.5.21a}$$

Donc sous l'hypothèse que

$$\operatorname{Re}(1 - \beta(e^{i\omega}))^{-1} - 1/2 \geq 0, \tag{3.5.21b}$$

on a (3.5.15b) et par là même (3.5.15a).

Pour montrer (3.5.15c), utilisons H3, c'est à dire le fait que la variance conditionnelle h_t soit bornée presque sûrement et (3.5.15c) est vérifiée dès que

$$\sum_{t=1}^{\infty} \frac{1}{t} \tilde{\varphi}'_t \tilde{R}_t^{-1} \tilde{\varphi}'_t < \infty. \tag{3.5.22}$$

Par ailleurs, on a

$$\begin{aligned}
 \tilde{\varphi}'_t \tilde{R}_t^{-1} \tilde{\varphi}'_t &= \tilde{\varphi}'_t \left(\tilde{R}_t^{-1} \right)^2 \tilde{R}_t \tilde{\varphi}'_t, \\
 &= \text{tr} \left(\tilde{\varphi}'_t \left(\tilde{R}_t^{-1} \right)^2 \tilde{R}_t \tilde{\varphi}'_t \right), \\
 &= \text{tr} \left(\tilde{\varphi}'_t \left(\tilde{R}_t^{-1} \right)^2 \tilde{\varphi}'_t \right) \text{tr} \left(\tilde{R}_t \right), \\
 &= \tilde{\varphi}'_t \left(\tilde{R}_t^{-1} \right)^2 \tilde{\varphi}'_t \text{tr} \left(\sum_{k=1}^t \tilde{\varphi}_k \tilde{\varphi}'_k \right), \\
 &= \tilde{\varphi}'_t \left(\tilde{R}_t^{-1} \right)^2 \tilde{\varphi}'_t \sum_{k=1}^t \|\tilde{\varphi}_k\|^2,
 \end{aligned} \tag{3.5.23}$$

et comme par le lemme d'inversion matricielle, on a également

$$\frac{1}{1 - \tilde{\varphi}'_t \tilde{R}_t^{-1} \tilde{\varphi}_t} = 1 + \tilde{\varphi}'_t \tilde{R}_{t-1}^{-1} \tilde{\varphi}_t,$$

donc (3.5.2c) devient

$$\tilde{R}_t^{-1} \tilde{\varphi}'_t = \frac{\tilde{R}_{t-1}^{-1} \tilde{\varphi}'_t}{1 + \tilde{\varphi}'_t \tilde{R}_{t-1}^{-1} \tilde{\varphi}_t}.$$

Ainsi nous avons

$$\begin{aligned}
 \sum_{t=1}^{\infty} \frac{1}{t} \tilde{\varphi}'_t \tilde{R}_t^{-1} \tilde{\varphi}'_t &= \sum_{t=1}^{\infty} \tilde{\varphi}'_t \left(\tilde{R}_t^{-1} \right)^2 \tilde{\varphi}'_t \frac{1}{t} \sum_{k=1}^t \|\tilde{\varphi}_k\|^2, \\
 &= \sum_{t=1}^{\infty} \tilde{\varphi}'_t \tilde{R}_t^{-1} \tilde{R}_t^{-1} \tilde{\varphi}'_t \frac{1}{t} \sum_{k=1}^t \|\tilde{\varphi}_k\|^2, \\
 &= \sum_{t=1}^{\infty} \tilde{\varphi}'_t \tilde{R}_t^{-1} \frac{\tilde{R}_{t-1}^{-1} \tilde{\varphi}'_t}{1 + \tilde{\varphi}'_t \tilde{R}_{t-1}^{-1} \tilde{\varphi}_t} \frac{1}{t} \sum_{k=1}^t \|\tilde{\varphi}_k\|^2,
 \end{aligned}$$

$$\begin{aligned}
 &\leq \sum_{t=1}^{\infty} \tilde{\varphi}'_t \tilde{R}_t^{-1} \tilde{R}_{t-1}^{-1} \tilde{\varphi}'_t \frac{1}{t} \sum_{k=1}^t \|\tilde{\varphi}_t\|^2, \\
 &= \sum_{t=1}^{\infty} tr \left(\tilde{\varphi}'_t \tilde{R}_t^{-1} \tilde{R}_{t-1}^{-1} \tilde{\varphi}'_t \right) \frac{1}{t} \sum_{k=1}^t \|\tilde{\varphi}_t\|^2 \\
 &= \sum_{t=1}^{\infty} tr \left(\tilde{\varphi}'_t \tilde{\varphi}'_t \tilde{R}_t^{-1} \tilde{R}_{t-1}^{-1} \right) \frac{1}{t} \sum_{k=1}^t \|\tilde{\varphi}_t\|^2 \\
 &= \sum_{t=1}^{\infty} tr \left(\left(\tilde{R}_t - \tilde{R}_{t-1} \right) \tilde{R}_t^{-1} \tilde{R}_{t-1}^{-1} \right) \frac{1}{t} \sum_{k=1}^t \|\tilde{\varphi}_t\|^2 \\
 &= \sum_{t=1}^{\infty} tr \left(\tilde{R}_{t-1}^{-1} - \tilde{R}_t^{-1} \right) \frac{1}{t} \sum_{k=1}^t \|\tilde{\varphi}_t\|^2
 \end{aligned}$$

Puisque

$$\sum_{t=1}^{\infty} tr \left(\tilde{R}_{t-1}^{-1} - \tilde{R}_t^{-1} \right) = tr(\tilde{R}_0^{-1}) - tr(\tilde{R}_{\infty}^{-1}) < \infty,$$

et que par hypothèse, on a

$$\lim_{t \rightarrow \infty} \sup \frac{1}{t} \sum_{k=1}^t \|\tilde{\varphi}_t\|^2 < \infty, \tag{3.5.24}$$

par conséquent, on obtient

$$\sum_{t=1}^{\infty} \frac{1}{t} \tilde{\varphi}'_t \tilde{R}_t^{-1} \tilde{\varphi}'_t < \infty.$$

Ce qui prouve complètement la première tranche du théorème. Notons que l'hypothèse (3.5.24) peut être remplacée par une autre plus simple

Lemme 3.5.2 Si

$$\lim_{t \rightarrow \infty} \sup \frac{1}{t} \sum_{k=1}^t \varepsilon_k^4 < \infty, \tag{3.5.25}$$

alors

$$\lim_{t \rightarrow \infty} \sup \frac{1}{t} \sum_{k=1}^t \|\tilde{\varphi}_t\|^2 < \infty.$$

Preuve : Il suffit de montrer que (3.5.25) entraîne nécessairement

$$\lim_{t \rightarrow \infty} \sup \frac{1}{t} \sum_{k=1}^t \tilde{\vartheta}_k^2 < \infty. \quad (3.5.26)$$

Pour ce faire, multiplions l'équation (3.5.1a) par $\tilde{R}_t$,

$$\tilde{R}_t \tilde{\theta}_t = \tilde{R}_t \tilde{\theta}_{t-1} + \tilde{\varphi}'_t \tilde{v}_t, \quad (3.5.27)$$

multiplions encore (3.5.27) par $\tilde{\theta}'_t$, ce qui donne

$$\begin{aligned} \tilde{\theta}'_t \tilde{R}_t \tilde{\theta}_t &= \tilde{\theta}'_t \tilde{R}_t \tilde{\theta}_{t-1} + \tilde{\theta}'_t \tilde{\varphi}'_t \tilde{v}_t, \\ &= \left(\tilde{\theta}_{t-1} + \tilde{R}_t^{-1} \tilde{\varphi}'_t \tilde{v}_t \right)' \tilde{R}_t \tilde{\theta}_{t-1} + \tilde{\theta}'_t \tilde{\varphi}'_t \tilde{v}_t, \\ &= \tilde{\theta}'_{t-1} \tilde{R}_t \tilde{\theta}_{t-1} + \tilde{\varphi}'_t \tilde{\theta}_{t-1} \tilde{v}_t + \tilde{\theta}'_t \tilde{\varphi}'_t \tilde{v}_t, \\ &= \tilde{\theta}'_{t-1} \left(\tilde{R}_{t-1} + \tilde{\varphi}'_t \tilde{\varphi}'_t \right) \tilde{\theta}_{t-1} + \tilde{\varphi}'_t \tilde{\theta}_{t-1} \tilde{v}_t + \tilde{\theta}'_t \tilde{\varphi}'_t \tilde{v}_t, \\ &= \tilde{\theta}'_{t-1} \tilde{R}_{t-1} \tilde{\theta}_{t-1} + \left(\tilde{\varphi}'_t \tilde{\theta}_{t-1} \right)^2 + 2 \tilde{\varphi}'_t \tilde{\theta}_{t-1} \tilde{v}_t + \tilde{\varphi}'_t \tilde{R}_t^{-1} \tilde{\varphi}'_t \tilde{v}_t^2, \end{aligned} \quad (3.5.28)$$

Sachant que $\tilde{\varphi}'_t \tilde{\theta}_{t-1} = \varepsilon_t^2 - \tilde{v}_t$, alors (3.5.28) devient

$$\begin{aligned} \tilde{\theta}'_t \tilde{R}_t \tilde{\theta}_t &= \tilde{\theta}'_{t-1} \tilde{R}_{t-1} \tilde{\theta}_{t-1} + \varepsilon_t^4 + \tilde{v}_t^2 - 2 \varepsilon_t^2 \tilde{v}_t + 2 \tilde{\varphi}'_t \tilde{\theta}_{t-1} \tilde{v}_t + \tilde{\varphi}'_t \tilde{R}_t^{-1} \tilde{\varphi}'_t \tilde{v}_t^2 \\ &= \tilde{\theta}'_{t-1} \tilde{R}_{t-1} \tilde{\theta}_{t-1} + \varepsilon_t^4 + \tilde{v}_t^2 - 2(\tilde{\varphi}'_t \tilde{\theta}_{t-1} + \tilde{v}_t) \tilde{v}_t + 2 \tilde{\varphi}'_t \tilde{\theta}_{t-1} \tilde{v}_t \\ &\quad + \tilde{\varphi}'_t \tilde{R}_t^{-1} \tilde{\varphi}'_t \tilde{v}_t^2 \\ &= \tilde{\theta}'_{t-1} \tilde{R}_{t-1} \tilde{\theta}_{t-1} + \varepsilon_t^4 - \tilde{v}_t^2 \left(1 - \tilde{\varphi}'_t \tilde{R}_t^{-1} \tilde{\varphi}'_t \right) \end{aligned} \quad (3.5.29)$$

Par sommation des membres de (3.5.29) pour $k = 1$ à t , on trouve

$$\begin{aligned} \tilde{\theta}'_t \tilde{R}_t \tilde{\theta}_t &= \tilde{\theta}'_0 \tilde{R}_0 \tilde{\theta}_0 + \sum_{k=1}^t \varepsilon_k^4 - \sum_{k=1}^t \tilde{v}_k^2 \left(1 - \tilde{\varphi}'_k \tilde{R}_k^{-1} \tilde{\varphi}'_k \right), \\ &= \tilde{\theta}'_0 \tilde{R}_0 \tilde{\theta}_0 + \sum_{k=1}^t \varepsilon_k^4 - \sum_{k=1}^t \frac{\tilde{\vartheta}_k^2}{1 - \tilde{\varphi}'_k \tilde{R}_k^{-1} \tilde{\varphi}'_k} \end{aligned}$$

D'où

$$\frac{1}{t} \sum_{k=1}^t \frac{\tilde{\vartheta}_k^2}{1 - \tilde{\varphi}'_t \tilde{R}_t^{-1} \tilde{\varphi}'_t} = \frac{1}{t} \sum_{k=1}^t \varepsilon_k^4 + \frac{1}{t} \tilde{\theta}'_0 \tilde{R}_0 \tilde{\theta}_0 - \frac{1}{t} \tilde{\theta}'_t \tilde{R}_t \tilde{\theta}_t. \quad (3.5.30)$$

Ainsi, d'après (3.5.30), il est clair que (3.5.25) entraîne que

$$\lim_{t \rightarrow \infty} \sup \frac{1}{t} \sum_{k=1}^t \frac{\tilde{\vartheta}_k^2}{1 - \tilde{\varphi}'_t \tilde{R}_t^{-1} \tilde{\varphi}'_t} < \infty,$$

et comme $\tilde{\varphi}'_t \tilde{R}_t^{-1} \tilde{\varphi}'_t < 1$ (voir Solo (1979), appendix), on a donc

$$\frac{\tilde{\vartheta}_k^2}{1 - \tilde{\varphi}'_t \tilde{R}_t^{-1} \tilde{\varphi}'_t} \geq \tilde{\vartheta}_k^2,$$

ce qui implique (3.5.26) et par là même que

$$\lim_{t \rightarrow \infty} \sup \frac{1}{t} \sum_{k=1}^t \left[q + \sum_{i=1}^q \varepsilon_{k-i}^4 + \sum_{j=1}^p \left(\varepsilon_{k-j}^2 - \tilde{\vartheta}_{k-j} \right)^2 \right] < \infty, \quad (3.5.31)$$

qui n'est rien d'autre que (3.5.24). D'après ce qui précède, il est possible d'appliquer le théorème de Robbins et Sigmund (1971), donc il existe une variable aléatoire non négative $\tilde{T}$ tel que $\tilde{T}_t \rightarrow \tilde{T}$ *p.s.* quand $t \rightarrow \infty$ et que $\sum_{k=1}^t \frac{1}{t} \tilde{T}_t < \infty$ *p.s.* Cette dernière condition implique

que $\tilde{T} = 0$ *p.s.* ■

Preuve de la partie ii)

Démontrons maintenant de la même manière la seconde tranche du Théorème 3.2.1.

Récrivons la cinquième et la sixième équations de (3.2.6)

$$\hat{\theta}_t = \hat{\theta}_{t-1} + \bar{h}_t^{-2} R_t^{-1} \hat{\varphi}_t \hat{v}_t, \quad (3.5.32a)$$

$$R_t = R_{t-1} + \bar{h}_t^{-2} \hat{\varphi}_t \hat{\varphi}_t' \quad (3.5.32b)$$

où $\bar{h}_t = \tilde{\varphi}'_t \tilde{\theta}_{t-1}$, et posons

$$\bar{\bar{\theta}}_t = \hat{\theta}_t - \theta^*,$$

alors (3.5.32) peut s'écrire comme

$$\bar{\theta}_t = \bar{\theta}_{t-1} + \bar{h}_t^{-2} R_t^{-1} \hat{\varphi}_t \hat{v}_t, \quad (3.5.33a)$$

$$R_t = R_{t-1} + \bar{h}_t^{-2} \hat{\varphi}_t \hat{\varphi}_t' \quad (3.5.33b)$$

Sachant que via le lemme d'inversion matricielle (cf, Lemme 2.3.1), on a

$$\frac{R_t^{-1} \hat{\varphi}_t}{\bar{h}_t^2 - \hat{\varphi}_t' R_t^{-1} \hat{\varphi}_t} = \bar{h}_t^{-2} R_{t-1}^{-1} \hat{\varphi}_t, \quad (3.5.33c)$$

et que

$$\begin{aligned} \hat{\vartheta}_t &= \varepsilon_t^2 - \hat{\varphi}_t' \hat{\theta}_t, \\ &= \varepsilon_t^2 - \hat{\varphi}_t' \hat{\theta}_{t-1} - \bar{h}_t^{-2} \hat{\varphi}_t' R_t^{-1} \hat{\varphi}_t \hat{v}_t, \\ &= \left(1 - \bar{h}_t^{-2} \hat{\varphi}_t' R_t^{-1} \hat{\varphi}_t\right) \hat{v}_t, \end{aligned}$$

alors (3.5.33a) devient

$$\begin{aligned} \bar{\bar{\theta}}_t &= \bar{\bar{\theta}}_{t-1} + \frac{R_t^{-1} \hat{\varphi}_t \hat{\vartheta}_t}{\bar{h}_t^2 - \hat{\varphi}_t' R_t^{-1} \hat{\varphi}_t}, \\ &= \bar{\bar{\theta}}_{t-1} + \bar{h}_t^{-2} R_{t-1}^{-1} \hat{\varphi}_t \hat{\vartheta}_t \end{aligned}$$

maintenant (3.5.33) peut s'écrire comme suit

$$\bar{\bar{\theta}}_t = \bar{\bar{\theta}}_{t-1} + \bar{h}_t^{-2} R_{t-1}^{-1} \hat{\varphi}_t \hat{\vartheta}_t, \quad (3.5.34a)$$

$$R_t = R_{t-1} + \hat{\varphi}_t \hat{\varphi}_t' \quad (3.5.34b)$$

multiplions (3.5.34a) à gauche par R_{t-1} alors on trouve

$$R_{t-1} \bar{\bar{\theta}}_t = R_{t-1} \bar{\bar{\theta}}_{t-1} + \bar{h}_t^{-2} \hat{\varphi}_t \hat{\vartheta}_t. \quad (3.5.35)$$

Dans le membre gauche de (3.5.35), utilisons (3.5.34b), alors (3.5.35) devient

$$R_t \bar{\bar{\theta}}_t - \bar{h}_t^{-2} \hat{\varphi}_t \hat{\varphi}_t' \bar{\bar{\theta}}_t = R_{t-1} \bar{\bar{\theta}}_{t-1} + \bar{h}_t^{-2} \hat{\varphi}_t \hat{\vartheta}_t, \quad (3.5.36)$$

multiplions maintenant (3.5.36), à gauche par $\bar{\theta}'_t$

$$\bar{\theta}'_t R_t \bar{\theta}_t = \bar{\theta}'_t R_{t-1} \bar{\theta}_{t-1} + \bar{h}_t^{-2} \bar{\theta}'_t \hat{\varphi}_t \hat{\vartheta}_t + \bar{h}_t^{-2} (\bar{\theta}'_t \hat{\varphi}_t)^2, \quad (3.5.37)$$

utilisons (3.5.34a) dans le premier terme du membre de droite de (3.5.37), cela donne

$$\bar{\theta}'_t R_t \bar{\theta}_t = \bar{\theta}'_{t-1} R_{t-1} \bar{\theta}_{t-1} + \bar{h}_t^{-2} \hat{\varphi}_t \bar{\theta}_{t-1} \hat{\vartheta}_t + \bar{h}_t^{-2} \bar{\theta}'_{t-1} \hat{\varphi}_t \hat{\vartheta}_t + \bar{h}_t^{-2} (\bar{\theta}'_{t-1} \hat{\varphi}_t)^2 \quad (3.5.38)$$

En remplaçant la valeur de $\bar{\theta}_{t-1}$ de (3.5.34a) dans le second terme du membre de droite de (3.5.38), alors on trouve

$$\begin{aligned} \bar{T}_t &= \bar{T}_{t-1} + \bar{h}_t^{-2} \hat{\varphi}_t \bar{\theta}_t \hat{\vartheta}_t - \bar{h}_t^{-4} \hat{\varphi}_t R_{t-1}^{-1} \hat{\varphi}_t \hat{\vartheta}_t^2 + \bar{h}_t^{-2} \hat{\varphi}_t \bar{\theta}_t \hat{\vartheta}_t + \bar{h}_t^{-2} (\bar{\theta}'_t \hat{\varphi}_t)^2 \\ &= \bar{T}_{t-1} + 2\bar{h}_t^{-2} \hat{\varphi}_t \bar{\theta}_t \hat{\vartheta}_t - \bar{h}_t^{-4} \hat{\varphi}_t R_{t-1}^{-1} \hat{\varphi}_t \hat{\vartheta}_t^2 + \bar{h}_t^{-2} (\bar{\theta}'_t \hat{\varphi}_t)^2 \end{aligned} \quad (3.5.39)$$

où $\bar{T}_t = \bar{\theta}'_t R_t \bar{\theta}_t$. Notons que (3.5.39) peut s'écrire comme suit

$$\begin{aligned} \bar{T}_t &= \bar{T}_{t-1} + 2\bar{h}_t^{-2} \hat{\varphi}_t \bar{\theta}_t (\hat{\vartheta}_t + \frac{1}{2} \hat{\varphi}_t \bar{\theta}_t) - \bar{h}_t^{-4} \hat{\varphi}_t R_{t-1}^{-1} \hat{\varphi}_t \hat{\vartheta}_t^2 \\ &= \bar{T}_{t-1} + 2\bar{h}_t^{-2} \hat{\varphi}_t \bar{\theta}_t (\hat{\vartheta}_t - v_t + \frac{1}{2} \hat{\varphi}_t \bar{\theta}_t) + 2\bar{h}_t^{-2} \hat{\varphi}_t \bar{\theta}_t v_t - \bar{h}_t^{-4} \hat{\varphi}_t R_{t-1}^{-1} \hat{\varphi}_t \hat{\vartheta}_t^2 \end{aligned}$$

d'où

$$\bar{T}_t = \bar{T}_{t-1} - 2\bar{h}_t^{-2} \bar{a}_t \bar{b}_t + 2\bar{h}_t^{-2} \hat{\varphi}_t \bar{\theta}_t v_t - \bar{h}_t^{-4} \hat{\varphi}_t R_{t-1}^{-1} \hat{\varphi}_t \hat{\vartheta}_t^2, \quad (3.5.40)$$

où

$$\bar{a}_t = -\hat{\varphi}_t \bar{\theta}_t, \quad (3.5.41a)$$

$$\bar{b}_t = \hat{\vartheta}_t - v_t + \frac{1}{2} \hat{\varphi}_t \bar{\theta}_t. \quad (3.5.41b)$$

Exploitions (3.5.33a) dans le 3ème terme du membre de droite de (3.5.40), ce qui donne

$$\bar{T}_t = \bar{T}_{t-1} - 2\bar{h}_t^{-2} \bar{a}_t \bar{b}_t + 2\bar{h}_t^{-2} \hat{\varphi}_t \bar{\theta}_{t-1} v_t + 2\bar{h}_t^{-2} \hat{\varphi}_t R_{t-1}^{-1} \hat{\varphi}_t \hat{v}_t v_t - \bar{h}_t^{-4} \hat{\varphi}_t R_{t-1}^{-1} \hat{\varphi}_t \hat{\vartheta}_t^2$$

ou encore

$$\begin{aligned} \bar{T}_t &= \bar{T}_{t-1} - 2\bar{h}_t^{-2} \bar{a}_t \bar{b}_t + 2\bar{h}_t^{-2} \hat{\varphi}_t \bar{\theta}_{t-1} v_t + 2\bar{h}_t^{-2} \hat{\varphi}_t R_{t-1}^{-1} \hat{\varphi}_t (\hat{v}_t - v_t) v_t \\ &\quad + 2\bar{h}_t^{-2} \hat{\varphi}_t R_{t-1}^{-1} \hat{\varphi}_t v_t^2 - \bar{h}_t^{-4} \hat{\varphi}_t R_{t-1}^{-1} \hat{\varphi}_t \hat{\vartheta}_t^2 \end{aligned} \quad (3.5.42a)$$

puisque le dernier terme du second membre de (3.5.42a) est positif, alors on en déduit la relation suivante

$$\bar{T}_t \leq \bar{T}_{t-1} - 2\bar{h}_t^{-2}\bar{a}_t\bar{b}_t + 2\bar{h}_t^{-2}\hat{\varphi}_t\bar{\theta}_{t-1}v_t + 2\bar{h}_t^{-2}\hat{\varphi}_tR_t^{-1}\hat{\varphi}_t(\hat{v}_t - v_t)v_t + 2\bar{h}_t^{-2}\hat{\varphi}_tR_t^{-1}\hat{\varphi}_tv_t^2 \quad (3.5.42b)$$

Sachant que les quantités : $\bar{h}_t$, $\hat{\varphi}_t$, R_t , $\bar{\theta}_{t-1}$ et $\hat{v}_t - v_t = \hat{\varphi}_t\hat{\theta}_{t-1} - \varphi'_t\theta^*$ sont $\mathcal{F}_{t-1}$ -mesurables, que $E(v_t/\mathcal{F}_{t-1}) = 0$ et enfin que

$$E(v_t^2/\mathcal{F}_{t-1}) = (E(\eta_t^4) - 1)h_t = Kh_t,$$

($K = E(\eta_t^4) - 1 \geq 0$), donc en prenant l'espérance conditionnelle de (3.5.42b) par rapport à $\mathcal{F}_{t-1}$, on trouve

$$\begin{aligned} E(\bar{T}_t + 2\bar{h}_t^{-2}\bar{a}_t\bar{b}_t/\mathcal{F}_{t-1}) &\leq \bar{T}_{t-1} + 2\bar{h}_t^{-2}\hat{\varphi}_t\bar{\theta}_{t-1}E(v_t/\mathcal{F}_{t-1}) + \\ &\quad 2\bar{h}_t^{-2}\hat{\varphi}_tR_t^{-1}\hat{\varphi}_t(\hat{v}_t - v_t)E(v_t/\mathcal{F}_{t-1}) + 2\bar{h}_t^{-2}\hat{\varphi}_tR_t^{-1}\hat{\varphi}_tE(v_t^2/\mathcal{F}_{t-1}) \\ &= \bar{T}_{t-1} + 2K\bar{h}_t^{-2}h_t\hat{\varphi}_tR_t^{-1}\hat{\varphi}_t \end{aligned} \quad (3.5.43)$$

Rajoutons aux deux membres de (3.5.43) la quantité $2 \sum_{k=1}^{t-1} \bar{h}_k^{-2}\bar{a}_k\bar{b}_k$, pour trouver

$$E\left(\bar{T}_t + 2 \sum_{k=1}^t \bar{h}_k^{-2}\bar{a}_k\bar{b}_k/\mathcal{F}_{t-1}\right) \leq \bar{T}_{t-1} + 2 \sum_{k=1}^{t-1} \bar{h}_k^{-2}\bar{a}_k\bar{b}_k + 2K\bar{h}_t^{-2}h_t\hat{\varphi}_tR_t^{-1}\hat{\varphi}_t \quad (3.5.44)$$

Posons $\bar{\bar{T}}_t = \frac{1}{t} \left(\bar{T}_t + 2 \sum_{k=1}^t \bar{h}_k^{-2}\bar{a}_k\bar{b}_k \right)$ alors on peut réécrire (3.5.44) comme suit

$$E\left(t\bar{\bar{T}}_t/\mathcal{F}_{t-1}\right) \leq (t-1)\bar{\bar{T}}_{t-1} + 2K\bar{h}_t^{-2}h_t\hat{\varphi}_tR_t^{-1}\hat{\varphi}_t,$$

ou encore

$$E\left(\bar{\bar{T}}_t/\mathcal{F}_{t-1}\right) \leq \bar{\bar{T}}_{t-1} - \frac{1}{t}\bar{\bar{T}}_{t-1} + 2t^{-1}K\bar{h}_t^{-2}h_t\hat{\varphi}_tR_t^{-1}\hat{\varphi}_t, \quad (3.5.45)$$

pour appliquer à nouveau le théorème de Robbins et Siegmund (1971), il suffit de montrer, comme pour la première tranche, que

$$\bar{\bar{T}}_t \geq 0,$$

et que

$$\sum_{t=1}^{\infty} t^{-1} K \bar{h}_t^{-2} h_t \hat{\varphi}_t R_t^{-1} \hat{\varphi}_t < \infty,$$

ou encore que

$$\sum_{k=1}^t \bar{h}_k^{-2} \bar{a}_k \bar{b}_k \geq 0, \quad (3.5.46a)$$

et

$$\sum_{t=1}^{\infty} t^{-1} \bar{h}_t^{-2} \hat{\varphi}_t R_t^{-1} \hat{\varphi}_t < \infty. \quad (3.5.46b)$$

Pour établir (3.5.46a), remarquons que d'après (3.5.41) (i.e. la définition de $\bar{a}_t$ et $\bar{b}_t$) et suivant la même démarche que (3.5.6)-(3.5.18), on peut facilement montrer que

$$\bar{b}_t = ((1 - \beta(L))^{-1} - 1/2) \bar{a}_t, \quad (3.5.47a)$$

$$= H(L) \bar{a}_t \quad (3.5.47b)$$

$$= \left(\sum_{s=1}^{\infty} H_s L^s \right) \bar{a}_t \quad (3.5.47c)$$

Par ailleurs, on a

$$\sum_{k=1}^t \bar{h}_k^{-2} \bar{a}_k \bar{b}_k = \sum_{k=1}^t \bar{\bar{a}}_k \bar{\bar{b}}_k,$$

où $\bar{\bar{a}}_k = \bar{h}_k^{-1} \bar{a}_k$, $\bar{\bar{b}}_k = \bar{h}_k^{-1} \bar{b}_k$ et

$$\bar{\bar{b}}_k = H(L) \bar{\bar{a}}_k. \quad (3.5.48)$$

D'après (3.5.47)-(3.5.48) et de la même façon que pour les équations (3.5.20), on a

$$\begin{aligned} \sum_{k=1}^t \bar{\bar{a}}_k \bar{\bar{b}}_k &\geq \lambda \sum_{k=1}^t \bar{\bar{a}}_k^2, \\ &= \lambda \sum_{k=1}^t \bar{h}_k^{-2} \bar{a}_k^2 \end{aligned}$$

où λ est donné par (3.5.21a), ce qui montre que (3.5.46) est vraie dès que l'hypothèse (3.5.21b) est vérifiée.

Montrons maintenant (3.5.46b). D'abord nous avons

$$\begin{aligned}
 \widehat{\varphi}_t R_t^{-1} \widehat{\varphi}_t &= \widehat{\varphi}_t (R_t^{-1})^2 R_t \widehat{\varphi}_t \\
 &= \text{tr} \left(\widehat{\varphi}_t (R_t^{-1})^2 R_t \widehat{\varphi}_t \right) \\
 &= \text{tr} \left(\widehat{\varphi}_t (R_t^{-1})^2 \widehat{\varphi}_t \right) \text{tr} (R_t) \\
 &= \widehat{\varphi}_t (R_t^{-1})^2 \widehat{\varphi}_t \text{tr} \left(\sum_{k=1}^t \bar{h}_k^{-2} \widehat{\varphi}_k \widehat{\varphi}_k' \right), \\
 &= \widehat{\varphi}_t (R_t^{-1})^2 \widehat{\varphi}_t \sum_{k=1}^t \bar{h}_k^{-2} \text{tr} \left(\widehat{\varphi}_k \widehat{\varphi}_k' \right) \\
 &= \widehat{\varphi}_t (R_t^{-1})^2 \widehat{\varphi}_t \sum_{k=1}^t \bar{h}_k^{-2} \|\varphi_k\|^2
 \end{aligned} \tag{3.5.49}$$

et comme par le lemme d'inversion matricielle, on a

$$\bar{h}_t^{-2} R_t^{-1} \widehat{\varphi}_t = \frac{R_{t-1}^{-1} \widehat{\varphi}_t}{\bar{h}_t^{-2} + \widehat{\varphi}_t R_{t-1}^{-1} \widehat{\varphi}_t}.$$

Ainsi nous avons

$$\begin{aligned}
 \sum_{t=1}^{\infty} t^{-1} \bar{h}_t^{-2} \widehat{\varphi}_t R_t^{-1} \widehat{\varphi}_t &= \sum_{t=1}^{\infty} \bar{h}_t^{-2} \widehat{\varphi}_t (R_t^{-1})^2 \widehat{\varphi}_t \left(t^{-1} \sum_{k=1}^t \bar{h}_k^{-2} \|\widehat{\varphi}_k\|^2 \right), \\
 &= \sum_{t=1}^{\infty} \bar{h}_t^{-2} \widehat{\varphi}_t R_t^{-1} R_t^{-1} \widehat{\varphi}_t \left(t^{-1} \sum_{k=1}^t \bar{h}_k^{-2} \|\widehat{\varphi}_k\|^2 \right), \\
 &= \sum_{t=1}^{\infty} \widehat{\varphi}_t R_t^{-1} \frac{R_{t-1}^{-1} \widehat{\varphi}_t}{\bar{h}_t^{-2} + \widehat{\varphi}_t R_{t-1}^{-1} \widehat{\varphi}_t} \left(t^{-1} \sum_{k=1}^t \bar{h}_k^{-2} \|\widehat{\varphi}_k\|^2 \right),
 \end{aligned}$$

$$\begin{aligned}
 &\leq \sum_{t=1}^{\infty} \bar{h}_t^{-2} \hat{\varphi}_t R_t^{-1} R_{t-1}^{-1} \hat{\varphi}_t \left(t^{-1} \sum_{k=1}^t \bar{h}_k^{-2} \|\hat{\varphi}_k\|^2 \right), \\
 &= \sum_{t=1}^{\infty} \text{tr} \left(\bar{h}_t^{-2} \hat{\varphi}_t R_t^{-1} R_{t-1}^{-1} \hat{\varphi}_t \right) \left(t^{-1} \sum_{k=1}^t \bar{h}_k^{-2} \|\hat{\varphi}_k\|^2 \right), \\
 &= \sum_{t=1}^{\infty} \text{tr} \left(\bar{h}_t^{-2} \hat{\varphi}_t \hat{\varphi}_t R_t^{-1} R_{t-1}^{-1} \right) \left(t^{-1} \sum_{k=1}^t \bar{h}_k^{-2} \|\hat{\varphi}_k\|^2 \right) \\
 &= \sum_{t=1}^{\infty} \text{tr} \left((R_t - R_{t-1}) R_t^{-1} R_{t-1}^{-1} \right) \left(t^{-1} \sum_{k=1}^t \bar{h}_k^{-2} \|\hat{\varphi}_k\|^2 \right) \\
 &= \sum_{t=1}^{\infty} \text{tr} \left(R_{t-1}^{-1} - R_t^{-1} \right) \left(t^{-1} \sum_{k=1}^t \bar{h}_k^{-2} \|\hat{\varphi}_k\|^2 \right)
 \end{aligned}$$

Donc si on montre que

$$\lim_{t \rightarrow \infty} \sup t^{-1} \sum_{k=1}^t \bar{h}_k^{-2} \|\hat{\varphi}_k\|^2 < \infty \text{ en probabilité} \quad (3.5.50)$$

alors puisque,

$$\sum_{t=1}^{\infty} \text{tr} \left(R_{t-1}^{-1} - R_t^{-1} \right) = \text{tr}(R_0^{-1}) - \text{tr}(R_{\infty}^{-1}) < \infty,$$

on en déduit que (3.5.46b) est vraie. Montrons maintenant que l'hypothèse

$$\lim_{t \rightarrow \infty} \sup \frac{1}{t} \sum_{k=1}^t \frac{\varepsilon_k^4}{(\varphi'_k \theta^*)^2} < \infty, \text{ en probabilité,}$$

entraîne (3.5.50). Nous avons d'abord besoin de montrer ceci

Lemme 3.5.3 : Si

$$\lim_{t \rightarrow \infty} \sup \frac{1}{t} \sum_{k=1}^t h_k^{-2} \varepsilon_k^4 < \infty, \quad (3.5.51a)$$

alors

$$\lim_{t \rightarrow \infty} \sup \frac{1}{t} \sum_{k=1}^t \bar{h}_k^{-2} \varepsilon_k^4 < \infty. \quad (3.5.51b)$$

Preuve : Il suffit de montrer que $\frac{1}{t} \sum_{k=1}^t \left(\bar{h}_k^{-2} - h_k^{-2} \right) \varepsilon_k^4 = o_p(1)$, i.e.

$$\frac{1}{t} \sum_{k=1}^t \left((\varphi'_k \theta^*)^{-2} - (\bar{\varphi}'_k \tilde{\theta}_{k-1})^{-2} \right) \varepsilon_k^4 = o_p(1). \quad (3.5.52)$$

En vertu du théorème de la valeur moyenne qui s'énonce comme suit : pour $x, y > 0$, on a

$$x^{-2} - y^{-2} = -2c^{-3}(x - y),$$

où

$$0 < c < x^{-1}(1 + xy^{-1}),$$

on peut réécrire le membre de droite de (3.5.52) comme suit

$$\begin{aligned} S_t &= t^{-1} \sum_{k=1}^t \sum_{k=1}^t \left((\varphi'_k \theta^*)^{-2} - (\bar{\varphi}'_k \tilde{\theta}_{k-1})^{-2} \right) \varepsilon_k^4, \\ &= 2t^{-1} \sum_{k=1}^t c_k^{-3} \left(\bar{\varphi}'_k \tilde{\theta}_{k-1} - \varphi'_k \theta^* \right) \varepsilon_k^4 \end{aligned}$$

où

$$0 < c_k < (\varphi'_k \theta^*)^{-1} \left[1 + \varphi'_k \theta^* \left(\bar{\varphi}'_k \tilde{\theta}_{k-1} \right)^{-1} \right]. \quad (3.5.53)$$

Par ailleurs, remarquons que :

$$\begin{aligned} \left| \bar{\varphi}'_k \tilde{\theta}_{k-1} - \varphi'_k \theta^* \right| &= \left| \bar{\varphi}'_k \tilde{\theta}_{k-1} - \bar{\varphi}'_k \theta^* + \bar{\varphi}'_k \theta^* - \varphi'_k \theta^* \right|, \\ &\leq \left| \bar{\varphi}'_k (\tilde{\theta}_{k-1} - \theta^*) \right| + |(\bar{\varphi}_k - \varphi_k)' \theta^*| \\ &\xrightarrow[k \rightarrow \infty]{} 0 \end{aligned}$$

qui est une conséquence directe de la première tranche du théorème. Cela entraîne d'une part que

$$(\varphi'_k \theta^*)^{-1} \bar{\varphi}'_k \tilde{\theta}_{k-1} < C \quad p.s.$$

ce qui implique que

$$\begin{aligned}
 c_k^3 &< (\varphi'_k \theta^*)^{-3} \left[1 + \varphi'_k \theta^* \left(\bar{\varphi}'_k \tilde{\theta}_{k-1} \right)^{-1} \right]^3, \\
 &< (\alpha_0^*)^{-3} \left[1 + \varphi'_k \theta^* \left(\bar{\varphi}'_k \tilde{\theta}_{k-1} \right)^{-1} \right]^3 \\
 &< C^*
 \end{aligned}$$

et d'autre part d'après le lemme de Toeplitz, on a

$$\begin{aligned}
 |S_t| &\leq 2t^{-1} \sum_{k=1}^t c_k^{-3} \left| \bar{\varphi}'_k \tilde{\theta}_{k-1} - \varphi'_k \theta^* \right| \varepsilon_k^4 \\
 &< 2C^* t^{-1} \sum_{k=1}^t \left| \bar{\varphi}'_k \tilde{\theta}_{k-1} - \varphi'_k \theta^* \right| \varepsilon_k^4 \\
 &\leq 2C^* \left[\left(t^{-1} \sum_{k=1}^t \left| \bar{\varphi}'_k \tilde{\theta}_{k-1} - \varphi'_k \theta^* \right|^2 \right) \left(t^{-1} \sum_{k=1}^t \varepsilon_k^8 \right) \right]^{1/2} \\
 &\xrightarrow[t \rightarrow \infty]{} 0
 \end{aligned}$$

pourvu que $E(\varepsilon_k^4) < \infty$. Ce qui achève la preuve du Lemme 3.5.3. ■

Il reste maintenant à montrer que (3.5.51b) entraîne que

$$\lim_{t \rightarrow \infty} \sup \frac{1}{t} \sum_{k=1}^t \bar{h}_k^{-2} \hat{\vartheta}_k^2 < \infty. \quad (3.5.54)$$

Pour cela, multiplions l'équation (3.5.32a) par R_t ,

$$R_t \hat{\theta}_t = R_t \hat{\theta}_{t-1} + \bar{h}_t^{-2} \hat{\varphi}_t \hat{v}_t, \quad (3.5.55)$$

multiplions encore (3.5.55) par $\widehat{\theta}'_t$, ce qui donne

$$\begin{aligned}
 \widehat{\theta}'_t R_t \widehat{\theta}_t &= \widehat{\theta}'_t R_t \widehat{\theta}_{t-1} + \bar{h}_t^{-2} \widehat{\theta}'_t \widehat{\varphi}_t \widehat{v}_t \\
 &= \left(\widehat{\theta}_{t-1} + \bar{h}_t^{-2} R_t^{-1} \widehat{\varphi}_t \widehat{v}_t \right)' R_t \widehat{\theta}_{t-1} + \bar{h}_t^{-2} \widehat{\theta}'_t \widehat{\varphi}_t \widehat{v}_t \\
 &= \widehat{\theta}'_{t-1} R_t \widehat{\theta}_{t-1} + \bar{h}_t^{-2} \widehat{\varphi}_t \widehat{\theta}_{t-1} \widehat{v}_t + \bar{h}_t^{-2} \widehat{\theta}'_t \widehat{\varphi}_t \widehat{v}_t \\
 &= \widehat{\theta}'_{t-1} \left(R_{t-1} + \bar{h}_t^{-2} \widehat{\varphi}'_t \widehat{\varphi}_t \right) \widehat{\theta}_{t-1} + \bar{h}_t^{-2} \widehat{\varphi}_t \widehat{\theta}_{t-1} \widehat{v}_t + \bar{h}_t^{-2} \widehat{\theta}'_t \widehat{\varphi}_t \widehat{v}_t \\
 &= \widehat{\theta}'_{t-1} R_{t-1} \widehat{\theta}_{t-1} + \bar{h}_t^{-2} \left(\widehat{\varphi}_t \widehat{\theta}_{t-1} \right)^2 + 2 \bar{h}_t^{-2} \widehat{\varphi}_t \widehat{\theta}_{t-1} \widehat{v}_t + \bar{h}_t^{-4} \widehat{\varphi}_t R_t^{-1} \widehat{\varphi}_t \widehat{v}_t^2
 \end{aligned} \tag{3.5.56}$$

Sachant que $\widehat{\varphi}_t \widehat{\theta}_{t-1} = \varepsilon_t^2 - \widehat{v}_t$, alors la dernière équation de (3.5.56) devient

$$\begin{aligned}
 \widehat{\theta}'_t R_t \widehat{\theta}_t &= \widehat{\theta}'_{t-1} R_{t-1} \widehat{\theta}_{t-1} + \bar{h}_t^{-2} \varepsilon_t^4 + \bar{h}_t^{-2} \widehat{v}_t^2 - 2 \bar{h}_t^{-2} \varepsilon_t^2 \widehat{v}_t \\
 &\quad + 2 \bar{h}_t^{-2} \widehat{\varphi}_t \widehat{\theta}_{t-1} \widehat{v}_t + \bar{h}_t^{-4} \widehat{\varphi}_t R_t^{-1} \widehat{\varphi}_t \widehat{v}_t^2 \\
 &= \widehat{\theta}'_{t-1} R_{t-1} \widehat{\theta}_{t-1} + \bar{h}_t^{-2} \varepsilon_t^4 + \bar{h}_t^{-2} \widehat{v}_t^2 - 2 \bar{h}_t^{-2} (\widehat{\varphi}_t \widehat{\theta}_{t-1} + \widehat{v}_t) \widehat{v}_t \\
 &\quad + 2 \bar{h}_t^{-2} \widehat{\varphi}_t \widehat{\theta}_{t-1} \widehat{v}_t + \bar{h}_t^{-4} \widehat{\varphi}_t R_t^{-1} \widehat{\varphi}_t \widehat{v}_t^2 \\
 &= \widehat{\theta}'_{t-1} R_{t-1} \widehat{\theta}_{t-1} + \bar{h}_t^{-2} \varepsilon_t^4 - \bar{h}_t^{-2} \widehat{v}_t^2 \left(1 - \bar{h}_t^{-2} \widehat{\varphi}_t R_t^{-1} \widehat{\varphi}_t \right)
 \end{aligned} \tag{3.5.57}$$

Par sommation des membres de (3.5.57) pour $k = 1$ à t , on trouve

$$\begin{aligned}
 \widehat{\theta}'_t R_t \widehat{\theta}_t &= \widehat{\theta}'_0 R_0 \widehat{\theta}_0 + \sum_{k=1}^t \bar{h}_k^{-2} \varepsilon_k^4 - \sum_{k=1}^t \bar{h}_k^{-2} \widehat{v}_k^2 \left(1 - \bar{h}_k^{-2} \widehat{\varphi}'_k R_k^{-1} \widehat{\varphi}_k \right) \\
 &= \widehat{\theta}'_0 R_0 \widehat{\theta}_0 + \sum_{k=1}^t \bar{h}_k^{-2} \varepsilon_k^4 - \sum_{k=1}^t \frac{\bar{h}_k^{-2} \widehat{v}_k^2}{1 - \bar{h}_k^{-2} \widehat{\varphi}'_k R_k^{-1} \widehat{\varphi}_k}
 \end{aligned}$$

D'où

$$\frac{1}{t} \sum_{k=1}^t \frac{\bar{h}_k^{-2} \widehat{v}_k^2}{1 - \bar{h}_k^{-2} \widehat{\varphi}'_k R_k^{-1} \widehat{\varphi}_k} = \frac{1}{t} \sum_{k=1}^t \bar{h}_k^{-2} \varepsilon_k^4 + \frac{1}{t} \widehat{\theta}'_0 R_0 \widehat{\theta}_0 - \frac{1}{t} \widehat{\theta}'_t R_t \widehat{\theta}_t \tag{3.5.58}$$

Ainsi, d'après (3.5.58), il est clair que (3.5.51b) entraîne que

$$\lim_{t \rightarrow \infty} \sup \frac{1}{t} \sum_{k=1}^t \frac{\bar{h}_k^{-2} \widehat{v}_k^2}{1 - \bar{h}_k^{-2} \widehat{\varphi}'_k R_k^{-1} \widehat{\varphi}_k} < \infty, \text{ en probabilité}$$

et comme $\bar{h}_k^{-2} \hat{\varphi}_k' R_k^{-1} \hat{\varphi}_k < 1$, presque sûrement (voir Solo, 1979, appendix), on a donc

$$\frac{\bar{h}_k^{-2} \hat{\vartheta}_k^2}{1 - \bar{h}_k^{-2} \hat{\varphi}_k' R_k^{-1} \hat{\varphi}_k} \geq \bar{h}_k^{-2} \hat{\vartheta}_k^2,$$

ce qui implique (3.5.54) et par là même que

$$\lim_{t \rightarrow \infty} \sup \frac{1}{t} \sum_{k=1}^t \left[\bar{h}_k^{-2} q + \sum_{i=1}^q \bar{h}_k^{-2} \varepsilon_{k-i}^4 + \sum_{j=1}^p \bar{h}_k^{-2} \left(\varepsilon_{k-j}^2 - \hat{\vartheta}_{k-j} \right)^2 \right] < \infty \text{ en probabilité} \quad (3.5.59)$$

compte tenu de l'hypothèse $H3$ qui entraîne que

$$\lim_{t \rightarrow \infty} \sup \frac{1}{t} \sum_{k=1}^t \bar{h}_k^{-2} < \infty \text{ en probabilité} \quad (3.5.60)$$

Notons que (3.5.59) n'est rien d'autre que (3.5.50).

D'après ce qui précède, il est possible d'appliquer le théorème de Robbins et Siegmund (1971), donc il existe une variable aléatoire positive $\bar{T}$ tel que $\bar{T}_t \rightarrow \bar{T}$ *p.s.* quand $t \rightarrow \infty$ et que $\sum_{k=1}^t \frac{1}{t} \bar{T}_t < \infty$ *p.s.* Cette dernière condition implique que $\bar{T} = 0$ *p.s.* ■

3.5.2 Preuve du Théorème 3.3.2

La preuve des résultats (3.3.7) et (5.3.8) peut être faite à l'aide de techniques similaires à celles utilisées pour d'autres algorithmes donnés dans Solo (1981), Ljung and Söderström (1983, Appendix 4.B) ou Aknouche (2006). Donc nous donnons seulement les grandes lignes de la démonstration de (3.3.8) tant la preuve de (3.5.7) peut se faire de manière similaire. Par récursion descendante, on peut montrer qu'à partir des trois dernières équations dans l'algorithme 3.3.1 on a la relation suivante

$$\hat{R}_t \bar{\theta}_t = \hat{R}_0 \bar{\theta}_0 + \sum_{k=1}^t \hat{\psi}_k \frac{v_k}{\left(\tilde{\varphi}_k' \tilde{\theta}_k \right)^2} + \sum_{k=1}^t \hat{\psi}_k \left(\frac{\hat{\psi}_k' \bar{\theta}_{k-1} + \hat{v}_k - v_k}{\left(\tilde{\varphi}_k' \tilde{\theta}_k \right)^2} \right). \quad (3.5.61)$$

Maintenant nous analysons chaque terme de (3.5.61) que l'on réécrit comme suit

$$t^{1/2}\bar{\theta}_t = \left(\frac{1}{t}\widehat{R}_t\right)^{-1} t^{-1/2} \sum_{k=1}^t \psi_k(\theta^*) \frac{v_k}{\left(\widetilde{\varphi}'_k \widetilde{\theta}_k\right)^2} + t^{-1/2} \left(\frac{1}{t}\widehat{R}_t\right)^{-1} \omega_t, \quad (3.5.62)$$

où

$$\omega_t = \sum_{k=1}^t \left(\widehat{\psi}_k - \psi_k(\theta^*)\right) \frac{v_k}{\left(\widetilde{\varphi}'_k \widetilde{\theta}_k\right)^2} + \widehat{R}_0 \bar{\theta}_0 + \sum_{k=1}^t \widehat{\psi}_k \left(\frac{\widehat{\psi}'_k \bar{\theta}_{k-1} + \widehat{v}_k - v_k}{\left(\widetilde{\varphi}'_k \widetilde{\theta}_k\right)^2} \right). \quad (3.5.63)$$

D'abord, nous avons à montrer que

$$\frac{1}{t}\widehat{R}_t \xrightarrow{p.s.} E \left(\frac{\psi_t(\theta^*) \psi'_t(\theta^*)}{(\varphi'_t \theta^*)^2} \right) \quad \text{lorsque } t \longrightarrow \infty. \quad (3.5.64)$$

Or,

$$\begin{aligned} \frac{1}{t}\widehat{R}_t &= \frac{1}{t} \sum_{k=1}^t \frac{\widehat{\psi}_k \widehat{\psi}'_k}{\left(\widetilde{\varphi}'_k \widetilde{\theta}_k\right)^2}, \\ &= \frac{1}{t} \sum_{k=1}^t \frac{\widehat{\psi}_k \widehat{\psi}'_k}{(\varphi'_k \theta^*)^2} + \frac{1}{t} \sum_{k=1}^t \left(\frac{1}{\left(\widetilde{\varphi}'_k \widetilde{\theta}_k\right)^2} - \frac{1}{(\varphi'_k \theta^*)^2} \right) \widehat{\psi}_k \widehat{\psi}'_k. \end{aligned}$$

En utilisant *H3* et *H5*, il est facile de montrer que le premier terme peut s'écrire comme suit

$$\begin{aligned} \frac{1}{t} \sum_{k=1}^t \frac{\widehat{\psi}_k \widehat{\psi}'_k}{(\varphi'_k \theta^*)^2} &= \frac{1}{t} \sum_{k=1}^t \frac{\psi_k(\theta^*) \psi'_k(\theta^*)}{(\varphi'_k \theta^*)^2} + \frac{1}{t} \sum_{k=1}^t \frac{\left(\widehat{\psi}_k - \psi_k(\theta^*)\right) \widehat{\psi}'_k}{(\varphi'_k \theta^*)^2} \\ &\quad + \frac{1}{t} \sum_{k=1}^t \frac{\left(\widehat{\psi}_k - \psi_k(\theta^*)\right) \psi'_k(\theta^*)}{(\varphi'_k \theta^*)^2}, \\ &= \frac{1}{t} \sum_{k=1}^t \frac{\psi_k(\theta^*) \psi'_k(\theta^*)}{(\varphi'_k \theta^*)^2} + o_p(1). \end{aligned}$$

et par stationnarité et ergodicité du processus *GARCH*, on a

$$\frac{1}{t} \sum_{k=1}^t \frac{\widehat{\psi}_k \widehat{\psi}'_k}{(\varphi'_k \theta^*)^2} \xrightarrow{p.s.} E \left(\frac{\psi_t(\theta^*) \psi'_t(\theta^*)}{(\varphi'_t \theta^*)^2} \right)$$

Donc (3.5.64) est établie pourvu que

$$\frac{1}{t} \sum_{k=1}^t \left(\frac{1}{\left(\tilde{\varphi}'_k \tilde{\theta}_k \right)^2} - \frac{1}{\left(\varphi'_k \theta^* \right)^2} \right) \hat{\psi}_k \hat{\psi}'_k = o_p(1),$$

ce qui peut être obtenu facilement (voir Lemme 3.5.3) en utilisant le théorème de la valeur moyenne, l'hypothèse $H5$ et le lemme de Toeplitz. Considérant le dernier terme dans le second membre de (3.5.62), on peut facilement montrer sous $H3$ et $H5$ et suivant les mêmes étapes que celles dans Ljung and Söderström (1983, Lemmas 4.B1, 4.B4 and 4.B.6), que

$$t^{-1/2} \left(\frac{1}{t} \hat{R}_t \right)^{-1} \omega_t = o_p(1).$$

Finalement, pour obtenir (3.5.60), nous avons besoin de montrer la convergence en loi suivante

$$t^{-1/2} \sum_{k=1}^t \psi_k(\theta^*) \frac{v_k}{\left(\tilde{\varphi}'_k \tilde{\theta}_k \right)^2} \xrightarrow{L} N \left(0, E \left(\frac{\psi_t(\theta^*) \psi'_t(\theta^*)}{\left(\varphi'_t \theta^* \right)^2} \eta_t^2 \right) \right).$$

Or,

$$\begin{aligned} t^{-1/2} \sum_{k=1}^t \psi_k(\theta^*) \frac{v_k}{\left(\tilde{\varphi}'_k \tilde{\theta}_k \right)^2} &= t^{-1/2} \sum_{k=1}^t \psi_k(\theta^*) \frac{v_k}{\left(\varphi'_k \theta^* \right)^2} + \\ &\quad t^{-1/2} \sum_{k=1}^t \left(\frac{1}{\left(\tilde{\varphi}'_k \tilde{\theta}_k \right)^2} - \frac{1}{\left(\varphi'_k \theta^* \right)^2} \right) \psi_k(\theta^*) v_k \end{aligned} \quad (3.5.65)$$

Dans cette dernière relation le dernier terme dans le membre de droite est " $o_p(1)$ " (négligeable en probabilité). En effet, d'après le théorème de la valeur moyenne, nous avons

$$\begin{aligned} t^{-1/2+\delta} \sum_{k=1}^t \left(\frac{1}{\left(\tilde{\varphi}'_k \tilde{\theta}_k \right)^2} - \frac{1}{\left(\varphi'_s \theta^* \right)^2} \right) \psi_k(\theta^*) v_k &= \\ t^{-1/2+\delta} \sum_{k=1}^t \left(\frac{\tilde{\varphi}'_k \tilde{\theta}_k - \varphi'_s \theta^*}{\chi_k^3} \right) \psi_k(\theta^*) v_k &= S_t \end{aligned}$$

pour un certain δ tel que $0 < \delta < 1/2$, où $\frac{1}{\chi_k^3}$ est borné en probabilité. Puisque sous $H1 - H5$, la moyenne $E \|S_t\|^2$ est finie, il s'ensuit d'après l'inégalité de Tchebyshev que

$$t^{-\delta} \|S_t\| = o_p(1),$$

et donc

$$t^{-\delta} S_t = o_p(1).$$

D'où en vertu du théorème central limite des martingales (cf, Hall and Heyde, 1980, Corollary 3.2) qui est facilement applicable et des hypothèses d'ergodicité et de stationnarité du processus *GARCH* on trouve

$$t^{-1/2} \sum_{k=1}^t \psi_k(\theta^*) \frac{v_k}{(\varphi_s' \theta^*)^2} \xrightarrow{L} N \left(0, E \left(\frac{\psi_t(\theta^*) \psi_t'(\theta^*)}{(\varphi_t' \theta^*)^2} \eta_t^2 \right) \right),$$

ce qui achève la preuve.

3.5.3 Quelques propriétés du lemme d'inversion matricelle

Soit R_t une matrice définie par :

$$\begin{aligned} R_t &= \sum_{k=1}^t \lambda_k \xi_k \xi_k' \\ &= R_{t-1} + \lambda_t \xi_t \xi_t', \end{aligned}$$

où $\lambda_t > 0$, alors par application du Lemme 2.3.1, on a

$$\begin{aligned} R_t^{-1} &= (R_{t-1} + \lambda_t \xi_t \xi_t')^{-1} \\ &= R_{t-1}^{-1} - \frac{R_{t-1}^{-1} \xi_t \xi_t' R_{t-1}^{-1}}{\lambda_t^{-1} + \xi_t' R_{t-1}^{-1} \xi_t}, \end{aligned}$$

Ce qui implique que

$$\begin{aligned} \lambda_t R_t^{-1} \xi_t &= \lambda_t R_{t-1}^{-1} \xi_t - \frac{\lambda_t R_{t-1}^{-1} \xi_t \xi_t' R_{t-1}^{-1} \xi_t}{\lambda_t^{-1} + \xi_t' R_{t-1}^{-1} \xi_t}, \\ &= \frac{R_{t-1}^{-1} \xi_t}{\lambda_t^{-1} + \xi_t' R_{t-1}^{-1} \xi_t} \end{aligned}$$

d'où

$$\begin{aligned}\lambda_t \xi_t' R_t^{-1} \xi_t &= \frac{\xi_t' R_{t-1}^{-1} \xi_t}{\lambda_t^{-1} + \xi_t' R_{t-1}^{-1} \xi_t}, \\ &< 1\end{aligned}$$

Nous avons également d'après le Lemme 2.3.1

$$(\lambda_t^{-1} + \xi_t' R_{t-1}^{-1} \xi_t)^{-1} = \lambda_t (1 - \lambda_t \xi_t' R_{t-1}^{-1} \xi_t),$$

et enfin

$$\begin{aligned}R_{t-1}^{-1} &= (R_t - \lambda_t \xi_t \xi_t')^{-1}, \\ &= R_t^{-1} - \frac{R_t^{-1} \xi_t \xi_t' R_t^{-1}}{-\lambda_t^{-1} + \xi_t' R_t^{-1} \xi_t}\end{aligned}$$

d'où

$$\lambda_t R_{t-1}^{-1} \xi_t = \frac{R_t^{-1} \xi_t}{\lambda_t^{-1} - \xi_t' R_t^{-1} \xi_t}.$$

■

3.6 Conclusion

Dans ce chapitre, on a construit des algorithmes récurrents pour estimer des modèles *GARCH* sans avoir recours à l'hypothèse par trop restrictive de normalité conditionnelle. Les méthodes s'inspirent des méthodes en ligne existantes. Il s'agit d'extensions de la méthode des moindres carrés récurrents, de la méthode de pseudo régression linéaire et de celle du maximum de vraisemblance récurrente au cas de modèles *GARCH* et qui donnent sous des conditions très générales des estimateurs consistents. Bien entendu, les méthodes proposées peuvent également être appliquées à des données hors ligne pour lesquelles elles garantissent un bon contrôle des estimateurs afin qu'ils restent dans la région de stabilité et de positivité. Il est possible d'inclure une tendance *ARMAX* avec des erreurs *GARCH*. Ces algorithmes sont adéquats pour des séries chronologiques en temps réel comme par exemple les séries financières en temps continu pour lesquels les modèles *GARCH* en sont de bonnes approximations.

Chapitre 4

Modèles autorégressifs doubles périodiques

4.1 Introduction

L'objectif principal de ce chapitre est l'étude des propriétés probabilistes de la classe des modèles autorégressifs S -périodiques (d'ordre p constant) à erreurs $ARCH(q)$ périodiques que l'on note $PAR - PARCH_S(p, q)$. Nous nous proposons, d'abord, de donner une nouvelle formulation du modèle $PAR - PARCH_S(p, q)$ dans laquelle l'équation d'innovation est à coefficients aléatoires périodiquement distribuées, dans le but de décrire de façon parcimonieuse cette caractéristique de la variance conditionnelle d'être dépendante du temps. Le modèle proposé que nous nommerons *AR double périodique* ($DPAR$) est, d'un point de vue

du second ordre, équivalent au modèle $PAR - PARCH$ obtenu dans Franses et Paap (2000), en ce sens que, les deux modèles ont les mêmes espérances et variances conditionnelles. La simplicité du modèle proposé nous permet d'étudier aisément la propriété de stationnarité périodique au sens strict et l'existence des moments d'ordres supérieurs. Le chapitre se compose de cinq sections. Dans la seconde section, nous donnons une autre formulation du modèle $PAR - PARCH_S(p, q)$ sous la forme d'un modèle autorégressif double périodique à coefficients aléatoires moyennant certaines hypothèses. Dans la section suivante nous donnons une représentation markovienne périodique à espace d'état d'un modèle $PAR - PARCH_S(p, q)$. Dans la section quatre, nous utiliserons cette représentation markovienne, à partir de laquelle nous obtiendrons des conditions nécessaires et suffisantes de stationnarité périodique stricte du modèle PAR double. Dans la section cinq sera donnée une condition suffisante pour qu'un modèle $PAR - PARCH_S(p, q)$ possède des moments d'ordre supérieurs finis. Une conclusion fera l'objet de la section six où seront également soulignés quelques problèmes ouverts.

4.2 Définitions et hypothèses

On dit qu'un processus du second ordre $\{y_t, t \in \mathbb{Z}\}$ possède une représentation autorégressive à erreurs conditionnellement hétéroscédastiques d'ordre p et q et de période S , notée $PAR - PARCH_S(p, q)$, s'il est une solution d'une équation aux différences stochastique de la forme

$$\begin{cases} y_t = \sum_{j=1}^p \phi_{tj} y_{t-j} + \varepsilon_t, \\ \varepsilon_t = \sqrt{h_t} \zeta_t, \\ h_t = E(\varepsilon_t^2 / \mathcal{F}_{t-1}) = \alpha_{t,0} + \sum_{i=1}^q \alpha_{t,i} \varepsilon_{t-i}^2, \quad t \in \mathbb{Z}, \end{cases} \quad (4.2.1)$$

où $\{\zeta_t, t \in \mathbb{Z}\}$ sont des variables aléatoires indépendantes et identiquement distribuées (*i.i.d.*) de moyenne zero et de variance un et où $\mathcal{F}_t$ désigne la σ -algèbre représentant l'information jusqu'au temps t . Le processus $\{\varepsilon_t, t \in \mathbb{Z}\}$ est adapté à la filtration $\{\mathcal{F}_t, t \in \mathbb{Z}\}$. Les paramètres de la partie autorégressive et ceux de la partie hétéroscédastique ϕ_{tj} and $\alpha_{t,i}$ sont périodiques du temps de période S , i.e., $\phi_{t+kS,j} = \phi_{t,j}$, $j = 1, \dots, p$ and $\alpha_{t+kS,i} = \alpha_{t,i}$, $i = 0, \dots, q$, où $\alpha_{t,0} > 0$, $\alpha_{t,i} \geq 0$, $i = 1, \dots, q$, $\forall t \in \mathbb{Z}$.

Sous des hypothèses adéquates données (**A1-A5**), le modèle (4.3.1) peut être représenté de façon équivalente, au sens du second ordre, sous la forme suivante d'un modèle *AR* double périodique dans lequel l'équation d'innovation est à coefficients aléatoires périodiquement corrélés :

$$\begin{cases} y_t = \sum_{j=1}^p \phi_{tj} y_{t-j} + \varepsilon_t & \text{(Equation d'observation)}, \\ \varepsilon_t = e_t + \sum_{i=1}^q \delta_{t,i} \varepsilon_{t-i} & \text{(Equation d'innovation)}, \end{cases} \quad (4.2.2)$$

où les suites de variables aléatoires $\{e_t\}$ $\{y_t\}$ $\{\varepsilon_t\}$ et celles des vecteurs aléatoires $\{\delta_t\}$, où $\delta_t = (\delta_{t1}, \delta_{t2}, \dots, \delta_{tq})'$, sont assujettis à vérifier les hypothèses suivantes :

A1 : $\{e_t\}$ est une suite de variables aléatoires indépendantes de moyenne zéro et de variance S -périodique $E(e_t^2) = \alpha_{t,0}$, où, il est à noter que $\alpha_{t,0}$ est la même constante que celle donnée dans (4.2.1) ;

A2 : $\{\delta_t\}$ est une suite de vecteurs indépendants de moyenne le vecteur nul et de matrice de variance covariance diagonale $\Sigma_t = \text{diag}(\alpha_{t,1}, \alpha_{t,2}, \dots, \alpha_{t,q})$, où $\alpha_{t,j}, j = 1, \dots, q$ sont tels que donnés en (4.2.1) ;

A3 : les suites $\{e_t\}$ et $\{\delta_t\}$ sont mutuellement indépendantes ;

A4 : $\{e_t\}$ et $\{\delta_t\}$ sont S -périodiquement distribués, i.e. la distribution de δ_s est la même que celle de $\delta_{s+S\tau}$ pour tout $s \in \{1, \dots, S\}$, $\tau \in \mathbb{Z}$ et ainsi, pour $\{e_t\}$;

A5 : y_t et ε_t sont $\mathcal{F}_t$ -mesurables.

Dans **A2**, l'hypothèse selon laquelle la matrice de covariance Σ_t est diagonale peut être relaxée à des matrices de covariances plus générales. Cela conduit à une formulation *ARCH* périodique augmentée (notée ici par $P-AARCH$), qui coïncide pour $S = 1$ avec le modèle *ARCH* augmenté (*AARCH*) proposé par Bera *et al.* (1992). Il est utile de noter que sous l'hypothèse de normalité conditionnelle i.e. si $\{e_t\}$ et $\{\delta_t\}$ sont conjointement gaussiennes, et $\{\zeta_t, t \in \mathbb{Z}\}$ est normalement distribuée, alors les modèles (4.2.1) et (4.2.2) sont équivalents aux ordres supérieurs, dans le sens où ils ont les mêmes moments conditionnels d'ordre supérieurs. Cette condition n'est pas requise pour l'obtention des propriétés probabilistes du modèle (4.2.2) et peut donc être omise. La spécification de l'équation d'innovation condition-

nellement hétéroscédastique comme un modèle autorégressif avec coefficients aléatoires a été introduite d'abord par Tsay (1987) pour les modèles *ARMA* conditionnellement hétéroscédastiques (*CHARMA*), généralisé par Wong et Li (1997) au cas multivarié et étudié par Li et al. (2001) pour des modèles *AR* partiellement non stationnaires à erreurs conditionnellement hétéroscédastiques. Une représentation similaire appliquée au modèles *GARCH* a été proposée par Bera et al. (1996). Tous ces travaux se basent sur l'hypothèse de la stationnarité stricte des coefficients aléatoires alors que dans notre cas, les coefficients aléatoires sont supposés périodiquement distribués. Dans le reste de ce chapitre, nous utilisons seulement la seconde représentation qui est utile pour l'obtention des propriétés probabilistes du modèle (4.2.2).

4.2.1 Représentation Markovienne du modèle

Dans cette sous section nous réécrivons le modèle double $PAR_S(p, q)$ donné par (4.2.2) sous forme Markovienne, ce qui permettra d'étudier aisément la périodicité stricte, l'ergodicité périodique et l'existence des moments d'ordres supérieurs. Réécrivons le modèle (4.2.2) comme suit

$$\begin{cases} y_t = \sum_{j=1}^p \phi_{tj} y_{t-j} + \sum_{i=1}^q \delta_{t,i} \varepsilon_{t-i} + e_t, \\ \varepsilon_t = \sum_{i=1}^q \delta_{t,i} \varepsilon_{t-i} + e_t, \end{cases} \quad (4.2.3)$$

et définissons les $(p+q)$ -vecteurs $Y_t = (y_t, \dots, y_{t-p+1}, \varepsilon_t, \dots, \varepsilon_{t-q+1})'$ et

$B_t = (e_t, 0'_{(p-1) \times 1}, e_t, 0'_{(q-1) \times 1})'$, et la matrice A_t de dimensions $(p+q) \times (p+q)$

$$A_t = \begin{pmatrix} \phi_{t,1} & \phi_{t,2} & \dots & \phi_{t,p-1} & \phi_{t,p} & \delta_{t,1} & \delta_{t,2} & \dots & \delta_{t,q-1} & \delta_{t,q} \\ & I_{(p-1) \times (p-1)} & & & 0_{(p-1) \times 1} & & & 0_{(p-1) \times q} & & \\ & & & 0_{q \times p} & & \delta_{t,1} & \delta_{t,2} & \dots & \delta_{t,q-1} & \delta_{t,q} \\ & & & & & I_{(q-1) \times (q-1)} & & & & 0_{(q-1) \times 1} \end{pmatrix},$$

où $0_{m \times 1}$ est le vecteur nul d'ordre m .

Avec de telles notations, le modèle (4.2.3) peut se mettre sous la forme espace d'état Markovienne suivante

$$Y_t = A_t Y_{t-1} + B_t, \quad (4.2.4)$$

dans laquelle la matrice A_t et le vecteur B_t sont des variables aléatoires de moments d'ordre supérieurs S -périodiques en vertu de l'hypothèse **A4**.

Cette représentation Markovienne a été d'abord explorée par Nicholls et Quinn (1982) pour des modèles autorégressifs à coefficients aléatoires (*RCA*), ensuite par Tsay (1987) et Ling (1999) pour représenter les modèles *CHARMA*, Tong (1990) pour les modèles *ARCH*, Ling et Deng (1993) pour les modèles *VAR* à erreurs conditionnellement hétéroscédastiques et par Bougerol et Picard (1992), Chen et An (1998) et Ling et McAleer (2002) pour les modèles *GARCH* à coefficients constants. L'article de Li et al., (2002) contient une littérature intensive sur le sujet.

4.3 Stationnarité périodique stricte

Il est facile de vérifier que l'unique solution, pour l'instant de départ s , de l'équation aux différences stochastique (4.2.4) est donnée par

$$Y_t = B_t + A_t A_{t-1} A_{t-2} \dots A_{s+2} A_{s+1} Y_s + \sum_{j=1}^{t-s-1} (A_t A_{t-1} A_{t-2} \dots A_{t-j+1}) B_{t-j}. \quad (4.3.1)$$

En outre, le théorème suivant donne une condition suffisante et une autre nécessaire pour l'existence d'une solution strictement périodique de (4.3.1) lorsque $s \rightarrow -\infty$.

Théorème 4.3.1 (Aknouche et Guerbyenne, 2005)

Pour que le modèle $D\text{-}PAR_S(p, q)$ (4.2.2) admette une unique solution strictement périodique, $\mathcal{F}_t$ -mesurable de la forme

$$Y_t = B_t + \sum_{j=1}^{\infty} \prod_{i=0}^{j-1} A_{t-i} B_{t-j}, \quad (4.3.2)$$

il suffit que

$$\rho \left(\prod_{s=0}^{S-1} E \left(A_{S-s}^{\otimes 2} \right) \right) < 1, \quad (4.3.3)$$

et il faut que

$$\lim_{j \rightarrow \infty} \left\| \left(\prod_{l=0}^{S-1} E \left(A_{n^{(t)}-l}^{\otimes 2} \right) \right)^j B^{\otimes 2} \right\| = 0, \quad (4.3.4)$$

où $B = (1, 0'_{(p-1) \times 1}, 1, 0'_{(q-1) \times 1})'$.

Preuve

i) Preuve de la condition suffisante. Posons $P_{t,j-1} = \prod_{i=0}^{j-1} A_{t-i}$ et $S_{t,m} = \sum_{j=0}^m P_{t,j-1} B_{t-j}$, où par

convention $\prod_{i=x}^y A_i = 1$ pour $x > y$. Alors on a

$$Y_t = S_{t,m} + R_{t,m}, \quad (4.3.5)$$

où $R_{t,m} = \prod_{i=0}^m A_{t-i} Y_{t-m-1}$. Pour prouver (4.3.3) il suffit de montrer que

$$E(S_{t,m} S'_{t,m}) \text{ est finie lorsque } m \rightarrow \infty,$$

pour tout $t \in \{1, \dots, S\}$. De l'hypothèse d'indépendance des matrices B_t et A_t et celle de A_{t-i} et A_{t-j} pour $j \neq i$ (voir hypothèses **A1-A3**) il vient que

$$\begin{aligned} \text{vec} [E(S_{t,m} S'_{t,m})] &= \text{vec} \left[\sum_{j=0}^m E(P_{t,j-1} B_{t-j} B'_{t-j} P'_{t,j-1}) \right], \\ &= E \left[\sum_{j=0}^m P_{t,j-1}^{\otimes 2} \text{vec}(B_{t-j} B'_{t-j}) \right], \\ &= \left[\sum_{j=0}^m \prod_{i=0}^{j-1} E(A_{t-i}^{\otimes 2}) \text{vec} E(B_{t-j} B'_{t-j}) \right]. \end{aligned}$$

En posant $t = n^{(t)} + S\tau^{(t)}$ et $j = n^{(j)} + S\tau^{(j)}$ tel que $n^{(t)}, n^{(j)} \in \{1, \dots, S\}$, $\tau^{(t)} \in \mathbb{Z}$ et $\tau^{(j)} \in \mathbb{N}$, on peut écrire la limite suivante :

$$\lim_{m \rightarrow \infty} \text{vec} \left[E \left(S_{t,m} S'_{t,m} \right) \right],$$

puisque $E \left(A_t^{\otimes 2} \right)$ et $E \left(\text{vec} \left(B_t B'_t \right) \right)$ sont S -périodiques, comme suit

$$\begin{aligned} \lim_{m \rightarrow \infty} \text{vec} \left[E \left(S_{t,m} S'_{t,m} \right) \right] = \\ \sum_{j=0}^{\infty} \left(\prod_{l=0}^{S-1} E \left(A_{n^{(t)}-l}^{\otimes 2} \right) \right)^{h^{(j)}} \left(\prod_{l=0}^{n^{(j)}-1} E \left(A_{n^{(t)}-l}^{\otimes 2} \right) \right) \text{vec} E \left(B_d B'_d \right), \end{aligned} \quad (4.3.6)$$

où $\tau^{(j)} = \left\lceil \frac{j-1}{S} \right\rceil$ et $d \in \{1, \dots, S\}$ est tel que $n^{(t)} - n^{(j)} = d + Sn$ avec $n \in \mathbb{Z}$.

Puisque la matrice

$$\prod_{l=0}^{n^{(j)}-1} E \left(A_{n^{(t)}-l}^{\otimes 2} \right),$$

est uniformément bornée par

$$\max_{0 \leq i \leq S} \left\| \prod_{l=0}^{S-1} E \left(A_{i-l}^{\otimes 2} \right) \right\|,$$

alors une condition suffisante pour que le modèle (4.2.3) admette une représentation (4.3.6), est que

$$\lim_{j \rightarrow \infty} \left\| \prod_{l=0}^{S-1} E \left(A_{n^{(t)}-l}^{\otimes 2} \right) \right\|^{[j/S]} = 0,$$

ce qui est vérifié si et seulement si le rayon spectral de la matrice

$$\prod_{l=0}^{S-1} E \left(A_{n^{(t)}-l}^{\otimes 2} \right),$$

(à partir de laquelle on obtient par un réarrangement cyclique la matrice $\prod_{s=0}^{S-1} E \left(A_{S-s}^{\otimes 2} \right)$) est inférieur à un en valeur absolue.

ii) Pour montrer la condition nécessaire, nous avons d'abord à montrer que la périodicité stricte du modèle (4.2.3) entraîne $\sum_{j=0}^m (z' \otimes z') \prod_{i=0}^{j-1} E(A_{t-i}^{\otimes 2}) \text{vec} E(B_{t-j} B'_{t-j}) < \infty$ pour tout vecteur z fixé de dimension $(p+q) \times 1$.

Considérons la notation (4.3.5) et posons $V_t = E(Y_t Y'_t)$. Alors puisque

$$E(S_{t,m} R'_{t,m}) = 0,$$

nous avons

$$\begin{aligned} \text{vec}(V_t) &= \text{vec}(E(S_{t,m} S'_{t,m})) + \text{vec}(E(R_{t,m} R'_{t,m})), \\ &= \sum_{j=0}^m \prod_{i=0}^{j-1} E(A_{t-i}^{\otimes 2}) \text{vec} E(B_{t-j} B'_{t-j}) + \prod_{i=0}^m E(A_{t-i}^{\otimes 2}) \text{vec} E(Y'_{t-m-1} Y_{t-m-1}), \\ &= \sum_{j=0}^m \prod_{i=0}^{j-1} E(A_{t-i}^{\otimes 2}) \text{vec} E(B_{t-j} B'_{t-j}) + \prod_{i=0}^m E(A_{t-i}^{\otimes 2}) \text{vec}(V_{t-m-1}). \end{aligned}$$

Par ailleurs, posons

$$\begin{cases} Q_{t-j,0} = E(B_{t-j} B'_{t-j}) & \text{et } Q_{t,j} = E(A_t Q_{t-1,j-1} A'_t) \\ R_{t-j,0} = V_{t-j} & \text{et } R_{t,j} = E(A_t R_{t-1,j-1} A'_t) \end{cases} \quad j = 1, 2, 3, \dots$$

Il est clair que les matrices S -périodiques $Q_{t,j}$ et $R_{t,j}$, $j \in N$, sont semi-définies positives pour tout $t \in \{1, \dots, S\}$ et nous avons

$$V_t = \sum_{j=0}^m Q_{t,j} + R_{t,m+1}; \quad m = 0, 1, 2, \dots$$

Si z est un $(p+q) \times 1$ -vecteur arbitrairement fixé, alors

$$\begin{aligned} \infty &> z' V_t z = \sum_{j=0}^m z' Q_{t,j} z + z' R_{t,m+1} z, \\ &\geq \sum_{j=0}^m z' Q_{t,j} z, \end{aligned}$$

pour tout $m \in N$ et $t \in \{1, \dots, S\}$.

Ainsi $\sum_{j=0}^m z'^{\otimes 2} \text{vec} Q_{t,j}$ converge lorsque $m \rightarrow \infty$ vers une matrice semi-définie positive qui

est exactement donnée par (4.3.7).

Maintenant on sait que (4.3.7) implique que

$$\lim_{j \rightarrow \infty} W_{t,j} = \lim_{j \rightarrow \infty} \prod_{i=0}^{j-1} E(A_{t-i}^{\otimes 2}) \text{vec} E(B_{t-j} B'_{t-j}) = 0, \text{ pour tout } t,$$

et par conséquent

$$\lim_{j \rightarrow \infty} \left\| \prod_{i=0}^{j-1} E(A_{t-i}^{\otimes 2}) \text{vec} E(B B') \right\| = 0.$$

Ainsi toutes les sous suites de la suite $(\|W_{t,j}\|)_j$ convergent vers 0, lorsque $j \rightarrow \infty$, en particulier lorsque j est multiple de S , i.e.

$$\lim_{j \rightarrow \infty} \|W_{t,Sj}\| = 0,$$

pour tout t fixé. Ainsi en adoptant la notation (4.3.6) pour cette dernière équation nous trouvons (4.3.4) comme requis. Ce qui achève la preuve. ■

Notons que la condition (4.3.3) peut être simplifiée en écrivant la matrice A_t comme une matrice triangulaire par bloc, comme suit

$$A_t = \begin{pmatrix} \Phi_t & \Lambda_t \\ \mathbf{0} & \Delta_t \end{pmatrix},$$

où

$$\Phi_t = \begin{pmatrix} \phi_{t,1} & \cdots & \phi_{t,p} \\ I_{(p-1) \times (p-1)} & 0_{(p-1) \times 1} \end{pmatrix},$$

$$\Lambda_t = \begin{pmatrix} \delta_{t,1} \cdots \delta_{t,q} \\ 0_{(p-1) \times q} \end{pmatrix},$$

$$\Delta_t = \begin{pmatrix} \delta_{t,1} & \cdots & \delta_{t,q} \\ I_{(q-1) \times (q-1)} & 0_{(q-1) \times 1} \end{pmatrix}.$$

D'où nous avons le corollaire suivant.

Corollaire 4.3.2 (Aknouche et Guerbyenne, 2005)

Le modèle $D\text{-}PAR_S(p, q)$ (4.3.2) admet une unique solution strictement périodique, $\mathcal{F}_t$ – mesurable de la forme (4.3.6) si

$$\rho \left(\prod_{s=0}^{S-1} \Phi_{S-s} \right) < 1, \quad (4.3.8a)$$

et

$$\rho \left(\prod_{s=0}^{S-1} E \left(\Delta_{S-s}^{\otimes 2} \right) \right) < 1. \quad (4.3.8b)$$

Preuve

D’abord nous avons

$$\begin{aligned} \prod_{s=0}^{S-1} E \left(A_{S-s}^{\otimes 2} \right) &= \prod_{s=0}^{S-1} E \left[\left(\begin{array}{cc} \Phi_{S-s} & \Lambda_{S-s} \\ \mathbf{0} & \Delta_{S-s} \end{array} \right) \otimes \left(\begin{array}{cc} \Phi_{S-s} & \Lambda_{S-s} \\ \mathbf{0} & \Delta_{S-s} \end{array} \right) \right] \\ &= \left(\begin{array}{cc} \prod_{s=0}^{S-1} E \left(\Phi_{S-s} \otimes \left(\begin{array}{cc} \Phi_{S-s} & \Lambda_{S-s} \\ 0 & \Delta_{S-s} \end{array} \right) \right) & K \\ \mathbf{0} & \prod_{s=0}^{S-1} E \left(\Delta_{S-s} \otimes \left(\begin{array}{cc} \Phi_{S-s} & \Lambda_{S-s} \\ 0 & \Delta_{S-s} \end{array} \right) \right) \end{array} \right) \end{aligned}$$

où K est une matrice par bloc dont l’expression n’est pas importante pour nous. Puisque

$$\prod_{s=0}^{S-1} E \left(\Phi_{S-s} \otimes \left(\begin{array}{cc} \Phi_{S-s} & \Lambda_{S-s} \\ 0 & \Delta_{S-s} \end{array} \right) \right) = \prod_{s=0}^{S-1} \Phi_{S-s} \otimes \left(\begin{array}{ccc} \Phi_{S-s} & \mathbf{0} & \\ \mathbf{0} & 0 & \dots & 0 \\ & I_{(q-1) \times (q-1)} & \mathbf{0} & \Delta_{S-s} \end{array} \right),$$

alors par un calcul direct on obtient

$$\det \left[\lambda I_{(p+q)^2} - \prod_{s=0}^{S-1} E \left(\Phi_{S-s} \otimes \left(\begin{array}{cc} \Phi_{S-s} & \Lambda_{S-s} \\ 0 & \Delta_{S-s} \end{array} \right) \right) \right] = \lambda^{pq} \det \left[\lambda I_{p^2} - \prod_{s=0}^{S-1} \Phi_{S-s}^{\otimes 2} \right]$$

De même, on peut aisément montrer que

$$\det \left[\lambda I - \prod_{s=0}^{S-1} E \left(\Delta_{S-s} \otimes \left(\begin{array}{cc} \Phi_{S-s} & \Lambda_{S-s} \\ 0 & \Delta_{S-s} \end{array} \right) \right) \right] = \lambda^{pq} \det \left[\lambda I_{q^2} - \prod_{s=1}^S E \left(\Delta_{S-s}^{\otimes 2} \right) \right].$$

D'où

$$\rho \left(\prod_{s=0}^{S-1} E \left(A_{S-s}^{\otimes 2} \right) \right) < 1,$$

si et seulement si

$$\rho \left(\prod_{s=0}^{S-1} \Phi_{S-s}^{\otimes 2} \right) < 1,$$

et

$$\rho \left(\prod_{s=0}^{S-1} E \left(\Delta_{S-s}^{\otimes 2} \right) \right) < 1.$$

Notons que

$$\rho \left(\prod_{s=0}^{S-1} \Phi_{S-s}^{\otimes 2} \right) = \rho \left(\prod_{s=0}^{S-1} \Phi_{S-s} \right)^{\otimes 2} < 1,$$

(voir par exemple Ling (1999a, Lemma 2.2)) si et seulement si

$$\rho \left(\prod_{s=0}^{S-1} \Phi_{S-s} \right) < 1,$$

comme requis. ■

Remarque 4.3.1

i) Pour le cas particulier d'un modèle *ARCH* périodique (*P-ARCH*) (cf, Bollerslev et Ghysels 1996), obtenu à partir de (4.2.1) lorsque $\phi_{tj} = 0$, $j = 1, \dots, p$ pour tout t , la condition de stationarité périodique stricte est donnée par (4.3.8b). Cette condition se réduit à celle de Bollerslev et Ghysels (1996) pour le modèle particulier $P-ARCH_S(1)$ de période $S = 2$ i.e.

$$\rho \left(\prod_{s=1}^S \alpha_{s1} \right) = \rho \left(\prod_{s=1}^S E \left(\delta_{s1}^2 \right) \right) < 1.$$

ii) Le modèle *AR* périodique (*PAR*) obtenu à partir de (4.2.2) lorsque $\delta_{t,i} = 0$, $i = 1, \dots, q$, est périodiquement stationnaire si et seulement si (4.3.8a) est vérifiée. Cette condition a déjà été donnée par Aknouche (2007).

4.4 Existence des moments d'ordres supérieurs

L'existence des moments d'ordres supérieurs pour un modèle de séries chronologiques est primordiale pour l'inférence statistique. Ce problème a fait l'objet d'une grande attention dans le cadre des modèles conditionnellement hétéroscédastiques (cf., Engle, 1982 ; Milhoj, 1985 ; Chen et An, 1998 ; Ling, 1999 ; Ling et McAleer, 2002, Bibi et Aknouche, 2006) parce que d'une part, les conditions des moments sur les paramètres sont plus restrictives que celles de la stationnarité ou stabilité et, d'autre part, les procédures d'inférence requièrent l'existence des moments d'ordres plus grands pour les modèles de type *ARCH* que pour des modèles linéaires traditionnels. Dans ce qui suit, une condition suffisante pour l'existence des moments d'un processus généré par (4.2.2) est donnée.

Théorème 4.4.1 (Aknouche et Guerbyenne, 2006a)

Supposons que $E(e_t^{2r})$ est fini pour tout $t \in \mathbb{Z}$ et $r > 1$ et que

$$\prod_{s=0}^{S-1} E(A_{S-s}^{\otimes 2r}) < 1. \quad (4.4.1)$$

Alors, le processus généré par (4.2.2) admet une solution (y_t, ε_t) périodique stricte avec

$$E(y_t^{2r}) < \infty,$$

et

$$E(\varepsilon_t^{2r}) < \infty,$$

pour tout t .

Preuve

HOUNA Selon la théorie de la convergence faible (e.g. Billingsley, 1968), il est bien connu qu'une condition suffisante pour l'existence de

$$E(\varphi(X)),$$

est que

$$\liminf_{\tau \rightarrow \infty} E\varphi(X_\tau) \text{ est finie,}$$

où (X_τ) est une suite de variables aléatoires convergeant en loi vers la variable X et $\varphi(\cdot)$ est une fonction mesurable. Nous allons utiliser ce résultat pour prouver le théorème. Posons

$$\begin{cases} Y_t^{(n)} = B_t + \sum_{j=1}^{n-1} \prod_{i=0}^{j-1} A_{t-i} B_{t-j} & \text{pour } n > 0, \\ Y_t^{(0)} = 0. \end{cases}$$

Il est clair que $Y_t^{(n)}$ vérifie (4.2.2) et que $Y_t^{(n)}$ converge *p.s.* vers l'unique solution de (4.2.2) donnée par (4.3.6) lorsque $n \rightarrow \infty$. Soit

$$\pi_n(t) = Y_t^{(n)} - Y_t^{(n-1)} = \begin{cases} \prod_{i=0}^{n-1} A_{t-i} B_{t-n+1} = A_t \pi_{n-1}(t-1) & n \geq 1, \\ B_t & n = 0, \end{cases}$$

Alors en utilisant les propriétés du produit de Kronecker on a

$$\pi_n^{\otimes 2r}(t) = A_t^{\otimes 2r} \pi_{n-1}^{\otimes 2r}(t-1),$$

pour un certain $r > 0$ et

$$\begin{aligned} E(\pi_n^{\otimes 2r}(t)) &= E(A_t^{\otimes 2r}) E(\pi_{n-1}^{\otimes 2r}(t-1)), \\ &= \prod_{i=0}^{n-1} E(A_{t-i}^{\otimes 2r}) E(B_{t-n+1}^{\otimes 2r}). \end{aligned} \tag{4.4.2}$$

La dernière équation est une équation aux différences S -périodique d'ordre 1. Il est bien connu (e.g. Bentarzi et Hallin, 1994, Bentarzi, 1998 ; Bibi et Aknouche, 2006) que, puisque $E(B_{t-n+1}^{\otimes 2r})$ est finie donc une condition suffisante pour que $E(\pi_n^{\otimes 2r}(t))$ ait une limite finie lorsque $n \rightarrow \infty$ est que (4.4.2) soit vérifiée. Puisque $y_t^{(n)}$ converge vers y_t où $y_t^{(n)}$ est la première composante du vecteur $Y_t^{(n)}$ et

$$\lim_{n \rightarrow \infty} E \left(\left(y_t^{(n)} - y_t^{(n-1)} \right)^{\otimes 2r} \right) \text{ est finie,}$$

alors en utilisant la théorie de la convergence faible il s'ensuit que

$$E(y_t^{2r}) < \infty.$$

Le même argument peut être utilisé pour montrer la finitude de $E(\varepsilon_t^{2r})$. Ceci achève la preuve.

4.5 Conclusion

Dans ce chapitre nous avons étudié quelques propriétés probabilistes des modèles PAR à erreurs $P-ARCH$. Pour cela, nous avons introduit une nouvelle classe de modèles que nous avons appelée Modèles autorégressifs doubles périodiques formés de deux modèles autorégressifs périodiques dont le premier est à coefficients constants décrivant l'observation et le second est à coefficients aléatoires pour représenter la partie innovation. En particulier nous avons étudié la périodicité stricte (stationnarité périodique stricte) et l'existence des moments d'ordres supérieurs en nous basant sur une représentation Markovienne dont les coefficients sont plutôt périodiquement distribués. La formulation proposée contient les représentations $ARCH$, $ARCH$ augmentée, $AR - ARCH$, AR périodiques comme cas particuliers.

Chapitre 5

Estimation Adaptative des Modèles AR-périodiques

5.1 Introduction

La classe des modèles autorégressifs moyenne mobile linéaires dépendant du temps périodiques $PARMA_d(p_t, q_t)$ a fait l'objet de nombreux travaux. (cf., Pagano (1978), Troutman (1979), Tiao and Grupe (1980), Cipra (1985), Vecchia (1985), Osborn (1992), Vecchia and Balerini (1991), Anderson and Vecchia (1993), Franses (1996), Bentarzi and Hallin (1994), (1995) and (1998), Bentarzi (1998), Adams and Goodwin (1995) et de nombreux autres auteurs comme une alternative prometteuse à la classe traditionnelle des modèles saisonniers *SARIMA* à coefficients constants popularisés par le fameux ouvrage de Box et Jenkins (1976).

Cet intérêt considérable accordé à la classe des modèles linéaires non stationnaires périodiques - théoriquement justifié grâce au théorème de décomposition de Cramér (Cramér (1961)) qui généralise la décomposition de Wold (1938) des processus stationnaires au second ordre- s'explique par son utilité et son caractère approprié pour modéliser des processus non stationnaire périodiquement corrélés (Gladyshev (1961)) qui sont très souvent rencontrés dans divers champs d'application ,en particulier, en économie (voir, e.g., Osborn and Smith (1989), et autres) en hydrologie et dans les études environnementales (Vecchia et al (1983), Ula and Smadi (1997), Bhuiya (1971) et autres). Comme une conséquence de la reconnaissance que de nombreuses séries de données d'observations exhibent des structures d'autocorrélation périodique -un aspect qui ne peut pas être géré par les modèles saisonniers traditionnels- les problèmes d'identification, d'estimation d'invertibilité et de tests liés aux modèles autorégressifs moyenne mobile périodiques à coefficients périodiques du temps ont été et sont intensivement étudiés par de nombreux auteurs.

Le présent chapitre est consacré principalement, d'une part, à la présentation des outils de base notamment la propriété de contiguïté d'une suite de lois de probabilité, concept introduit par Le Cam (1960) et repris par Roussas dans sa monographie (1972), et le concept de normalité locale asymptotique pour une suite de variables aléatoires appelées, dans ce contexte, suites centrales, qui sont utilisés et qui feront l'objet de la section 1 . La section 2 concerne les résultats de Hájek (1972) modifiés par Fabian et Hannan (1982). Dans la section 3, en suivant l'approche de Kreiss (1987) et, en utilisant la propriété LAN (*local asymptotic normality*) pour un modèle autorégressif périodique — prouvée dans Bentarzi et Hallin (1996) sur la base du lemme de Swensen (1985) — adapté à notre contexte où l'on suppose que la variance de l'innovation est périodique et que le modèle sous l'hypothèse nulle est périodique, on détermine des estimateurs adaptatifs pour les paramètres d'un modèle autorégressif périodique. Stein (1956) a, le premier, émit l'idée de déterminer, pour une suite de variables aléatoires indépendantes de moyenne θ inconnue, un estimateur qui serait asymptotiquement aussi bon lorsque la loi de probabilité qui régit le phénomène aléatoire est connue que lorsqu'elle est inconnue. Cet estimateur est dit *adaptatif*. Stein (1956) adonné une condition suffisante d'adaptativité fondée sur les espaces tangents. Bickel (1982), toujours dans le contexte d'une suite de variables aléatoires indépendantes mais aussi pour un

modèles de régression a montré que la condition donnée par Stein (1956) est également nécessaire et que, donc, il existe un estimateur adaptatif pour le paramètre pour la classe des lois de probabilité absolument continues et symétriques ; Linton (1993) a, dans le cadre des modèles *ARCH* démontré la propriété de normalité asymptotique locale et déterminé un estimateur adaptatif pour le paramètre vectoriel sous-jacent après avoir reparamétrisé son modèle pour éviter le problème de non identifiabilité du modèle. Dong Wan Shin and Beong Soo So (1999) ont développé des tests de racine unité fondés sur l'estimation adaptative, Ling (2003) s'est intéressé à l'estimation adaptative dans les modèles de mémoire longue. Dans la section 2, nous donnons les éléments de la théorie asymptotique des expériences statistiques. Dans la section 3, les notations de base sont présentées ainsi que les définitions relatives à un modèle autorégressif périodique indispensables pour les sections suivantes. Ces outils sont utilisés dans la section 4 pour l'obtention de la propriété de normalité locale asymptotique (en anglais LAN (*local asymptotic normality*) (cf., Le Cam (1960), (1986), Le Cam and Yang (1990) et Swensen (1985) ainsi que la propriété de linéarité (cf., Bentarzi and Hallin (1994)) du modèle périodique sous-jacent. Dans la section 4, on obtient un estimateur localement asymptotiquement minimax (L.A.M.E), au sens de Fabian and Hannan (1982). La section 5 est consacrée à l'obtention des estimateurs adaptatifs (E.A) des paramètres d'un processus autorégressif périodique dont la fonction densité de l'innovation du modèle n'est pas spécifiée, de variance périodique supposée inconnue. La performance des estimateurs adaptatifs obtenus est comparée à celle des estimateurs classiques des moindres carrés (MCO) via une étude de simulation intensive. La densité de l'innovation est supposé vérifier certaines conditions de régularité. Le critère de comparaison utilisé est le RMSE (Root Mean Square Error) est utilisé.

5.2 Quelques éléments de la théorie asymptotique des expériences statistiques

5.2.1 Normalité locale asymptotique

Soit $\underline{Y}^{(n)} = (Y_1^{(n)}, \dots, Y_n^{(n)})$, $n \in \mathbb{N}^*$, une suite d'observations décrite par la suite de modèles ou expériences statistiques $\mathcal{E}^{(n)} = (\mathbb{R}^n, \mathcal{B}_{\mathbb{R}^n}, \mathcal{P}^{(n)})$, où $\mathcal{P}^{(n)} = (P_\theta^{(n)}, \theta \in \Theta)$ est une famille paramétrique de lois de probabilités définies sur $(\mathbb{R}^n, \mathcal{B}_{\mathbb{R}^n})$ et indexées (de façon injective) par le paramètre θ , $\theta \in \Theta$ où on suppose que Θ est un ouvert de $\mathbb{R}^p$. Pour chaque $\theta \in \Theta$ sont définies des suites dites locales de valeurs $\theta^{(n)}$ du paramètre convergeant vers θ lorsque $n \rightarrow \infty$. La forme de ces suites locales se caractérise par une suite $\nu(n)$ de matrices $p \times p$ non singulières telles que $\|\nu(n)\| \rightarrow 0$ lorsque $n \rightarrow \infty$ ($\|\cdot\|$ désignant la norme euclidienne ou une norme équivalente).. Ces suites locales (appelées contre hypothèses locales) sont de la forme $\theta^{(n)} = \theta + \nu(n)\underline{\tau}^{(n)}$, $\underline{\tau}^{(n)} \in \mathbb{R}^p$ et $\sup_n \underline{\tau}^{(n)'} \underline{\tau}^{(n)} < \infty$ et la suite d'expériences $\mathcal{E}^{(n)}(\theta) = (\mathbb{R}^n, \mathcal{B}_{\mathbb{R}^n}, P_{\theta + \nu(n)\underline{\tau}^{(n)}}^{(n)}, \underline{\tau}^{(n)} \in \mathbb{R}^p)$ est appelée suite d'expériences locales (localisées en θ).

Remarque 5.2.1 Θ étant un ouvert, $\theta^{(n)} \in \Theta$ pour n suffisamment grand.

Remarque 5.2.2 Dans les modèles de position ou d'autorégression $\nu(n) = n^{-\frac{1}{2}}I$, où I est la matrice identité d'ordre p .

Pour les suites d'expériences locales $\mathcal{E}^{(n)}$ et $\mathcal{E}^{(n)}(\theta)$, on considère

$$\Lambda_{(\underline{\theta} + \nu(n)\underline{\tau}^{(n)})/\underline{\theta}}^{(n)} = \log \frac{dP_{\theta + \nu(n)\underline{\tau}^{(n)}}^{(n)}}{dP_\theta^{(n)}}. \quad (5.2.1)$$

La log-vraisemblance considérée en (5.2.1) correspond à une décomposition de Lebesgue arbitraire de $P_{\theta + \nu(n)\underline{\tau}^{(n)}}^{(n)}$ par rapport à $P_\theta^{(n)}$.

Les procédures d'inférence paramétriques et non paramétriques asymptotiquement optimales résultent de la normalité locale asymptotique appelée propriété LAN qui est caractérisée par la décomposition et le comportement de $\Lambda_{(\underline{\theta} + \nu(n)\underline{\tau}^{(n)})/\underline{\theta}}^{(n)}$.

Le point de départ des développements de la théorie asymptotique moderne est la mise en évidence par Le Cam (1960) du fait remarquable qu'une structure extrêmement générale (appelée ci-dessous structure de normalité locale asymptotique) suffisante et presque nécessaire à l'étude des performances asymptotique de la quasi totalité des procédures statistiques, peut être caractérisée très simplement à partir du comportement de $\Lambda_{(\underline{\theta} + \nu^{(n)} \underline{\tau}^{(n)})/\underline{\theta}}^{(n)}$.

Définition 5.2.3 (Propriété LAN)

La suite d'expérience $\mathcal{E}^{(n)} = (\mathbb{R}^n, \mathcal{B}_{\mathbb{R}^n}, \mathcal{P}^{(n)})$ est dite localement asymptotiquement normale (LAN) si, pour tout $\theta \in \Theta$, il existe une suite $\Delta^{(n)}(\underline{\theta})$ de vecteurs aléatoires $\sigma(Y^{(n)}, \underline{\theta})$ -mesurables et une matrice semi définie positive $\Gamma(\underline{\theta})$ fonction continue en $\underline{\theta}$ telles que

(i) pour toute suite $\underline{\tau}^{(n)} \in \mathbb{R}^p$ telle que $\sup_n \underline{\tau}^{(n)'} \underline{\tau}^{(n)} < \infty$

$$\Lambda_{(\underline{\theta} + \nu^{(n)} \underline{\tau}^{(n)})/\underline{\theta}}^{(n)} = \underline{\tau}^{(n)'} \Delta^{(n)}(\underline{\theta}) - \frac{1}{2} \underline{\tau}^{(n)'} \Gamma(\underline{\theta}) \underline{\tau}^{(n)} + o_P(1) \text{ sous } P_{\theta}^{(n)} \text{ et pour } n \rightarrow \infty. \quad (5.2.3a)$$

(ii) sous $P_{\theta}^{(n)}$, $\Delta^{(n)}(\underline{\theta})$ est asymptotiquement normale de moyenne nulle et de matrice de variance covariance $\Gamma(\underline{\theta})$. Le vecteur $\Delta^{(n)}(\underline{\theta})$ est appelée suite centrale. (5.2.3b).

Remarque 5.2.4

1) Modèle de position gaussien

La forme de l'approximation quadratique (5.2.3) de la log-vraisemblance est une forme familière, si l'on considère le modèle gaussien

$$G(\underline{\theta}) = \{ \mathcal{N}(\Gamma(\underline{\theta}) \underline{\tau}^{(n)}, \Gamma(\underline{\theta})), \underline{\tau}^{(n)} \in \mathbb{R}^p \}, \quad (5.2.4)$$

appelé expérience gaussienne locale, (localisée en $\underline{\theta}$; $\underline{\tau}^{(n)}$ est appelé paramètre local).

Une suite d'expériences possédant la propriété LAN est donc une suite d'expériences dont les log-vraisemblances $\Lambda_{(\underline{\theta} + \nu^{(n)} \underline{\tau}^{(n)})/\underline{\theta}}^{(n)}$ peuvent être approchées (ponctuellement, en chaque valeur de $\underline{\theta}$) par les log-vraisemblances d'une expérience gaussienne de la forme (5.2.4). Une telle approximation présente une signification statistique c'est à dire décisionnelle. On

montre que l'approximation, dans une famille LAN, des expériences locales par des expériences gaussiennes locales implique que, en particulier, les estimateurs obtenus possédant de bonnes propriétés dans ces expériences gaussiennes locales, possèdent également, sous forme approchée ou asymptotique ces bonnes propriétés dans les expériences initialement considérées.

2) La normalité asymptotique (5.2.3b) de la suite centrale $\Delta^{(n)}(\underline{\theta})$ entraîne celle des log-vraisemblances $\Lambda_{(\underline{\theta} + \nu^{(n)} \underline{\tau}^{(n)})/\underline{\theta}}^{(n)}$: sous $P_{\underline{\theta}}^{(n)}$, lorsque $n \rightarrow \infty$, et pour toute suite $\underline{\tau}^{(n)}$,

$$\frac{\Lambda_{(\underline{\theta} + \nu^{(n)} \underline{\tau}^{(n)})/\underline{\theta}}^{(n)} + \frac{1}{2} \underline{\tau}^{(n)'} \Gamma(\underline{\theta}) \underline{\tau}^{(n)}}{[\underline{\tau}^{(n)'} \Gamma(\underline{\theta}) \underline{\tau}^{(n)}]^{1/2}},$$

est asymptotiquement normale de moyenne zéro et de variance un.

Cette propriété LAN a été établie pour plusieurs modèles, par exemple, les modèles autorégressifs avec tendance linéaire (Swensen (1985)), les modèles *ARMA* multivariés avec tendance linéaire (Hallin et Garel (1995), Hallin et Puri (1994)), les modèles bilinéaires (Benghabrit et Hallin (1992, 1996)), les modèles de régression avec erreurs *ARCH* (Linton (1993)), les modèles *GARCH* (Klaassen (1995)), les modèles de régression à coefficients aléatoires (Akharif et Hallin (1995)), Werker (1996) pour les modèles de séries chronologiques.

La propriété LAN est utilisée pour deux objectifs différents. Le premier objectif est l'estimation adaptative qui permet de construire des estimateurs uniformément optimaux (Bickel (1982), Kreiss (1987), Klaassen (1995), Linton (1993), etc...). Le deuxième objectif est la construction de tests et de tests de rang optimaux, (Hallin et Puri (1987, 1988, 1994), Benghabrit et Hallin (1992, 1996), Assili (1997), Zafri (1998) etc...).

L'une des conséquences de cette propriété LAN est la contiguïté mutuelle des suites d'expériences locales qui nous permet d'avoir la loi sous la contre hypothèse locale.

5.2.2 Contiguïté

Définition 5.2.5

Soient $(\mathcal{X}^{(n)}, \mathcal{A}^{(n)})$, $n \in \mathbb{N}^*$, une suite d'espaces mesurables, $P^{(n)}$ et $Q^{(n)}$ deux suites de mesures de probabilités sur $(\mathcal{X}^{(n)}, \mathcal{A}^{(n)})$. Si pour toute suite $A^{(n)} \in \mathcal{A}^{(n)}$,

$$\lim_{n \rightarrow \infty} P^{(n)}(A^{(n)}) = 0 \Rightarrow \lim_{n \rightarrow \infty} Q^{(n)}(A^{(n)}) = 0,$$

on dit que la suite $Q^{(n)}$ est contiguë à la suite $P^{(n)}$. Si $P^{(n)}$ et $Q^{(n)}$ sont mutuellement contiguës, nous dirons que $P^{(n)}$ et $Q^{(n)}$ sont contiguës.

Remarque 5.2.6

Il n'est pas facile d'établir la contiguïté de deux suites de mesures de probabilité au moyen de la définition, et il est donc utile d'énoncer des critères de contiguïté plus opérationnels.

5.2.3 Les lemmes de Le Cam

Premier lemme de Le Cam

Proposition 5.2.7

Soit $F^{(n)}$ la fonction de répartition sous $P^{(n)}$ de $\frac{dQ^{(n)}}{dP^{(n)}}$:

$$F^{(n)}(x) = P^{(n)}\left(\frac{dQ^{(n)}}{dP^{(n)}} \leq x\right).$$

Si la suite $F^{(n)}$ converge faiblement vers une fonction de répartition F telle que $\int x dF(x) = 1$, alors $P^{(n)}$ et $Q^{(n)}$ sont contiguës.

Corollaire 5.2.8

Si $\Lambda^{(n)} = \frac{dQ^{(n)}}{dP^{(n)}}$ est asymptotiquement normale sous $P^{(n)}$, c'est à dire s'il existe $\mu^{(n)}$ et $\sigma^{(n)}$ tels que sous $P^{(n)}$, lorsque $n \rightarrow \infty$, $\frac{\Lambda^{(n)} - \mu^{(n)}}{\sigma^{(n)}}$ est asymptotiquement normale de moyenne zéro et de variance un, alors $P^{(n)}$ et $Q^{(n)}$ sont mutuellement contiguës si et seulement si

- (i) $\limsup_{n \rightarrow \infty} \sigma^{(n)} < \infty$
- (ii) $\lim_{n \rightarrow \infty} (\mu^{(n)} + \frac{1}{2}\sigma^{(n)}) = 0$.

D'après 2) de la **Remarque 5.2.4**, nous sommes précisément dans la situation décrite dans le corollaire précédent. On a donc le résultat suivant.

Corollaire 5.2.9

Dans une famille LAN, les suites $P_{\theta+\nu(n)\underline{\tau}^{(n)}}^{(n)}$ et $P_{\theta}^{(n)}$ sont mutuellement contiguës, quel que soit θ , et pour toute suite bornée $\underline{\tau}^{(n)}$.

5.3 Normalité Locale Asymptotique des modèles auto-régressifs d'ordre p_t , d -périodiques

5.3.1 Définitions et formulation.

Hypothèse nulle et contre hypothèse locale

Soit $\{y_t^{(n)}, t \in \mathbb{Z}\}$ une suite de réalisations d'un processus.

Posons $t = i + d\tau; i = 1, \dots, d, d \in N^* - \{1\}, \tau \in \mathbb{Z}$.

Nous désignons par $H_f^{(n)}(\underline{\theta})$ la suite d'hypothèses nulles sous lesquelles nous supposons que le processus $\{y_t^{(n)}, t \in \mathbb{Z}\}$ est une solution de l'équation aux différences linéaire stochastique :

$$y_t^{(n)} = \sum_{j=1}^{p_t} \theta_{t,j} y_{t-j}^{(n)} + \varepsilon_t^{(n)}, \quad (5.2.1)$$

$$y_t^{(n)} = \sum_{j=1}^{p_t} \theta_{t,j} y_{t-j}^{(n)} + \varepsilon_t^{(n)} \text{ (modèle } PAR(p_t), d\text{-périodique)}.$$

Les coefficients $\theta_{t,j}; j = 1, \dots, p_t; t \in \mathbb{Z}$ sont dépendants du temps et périodiques de période d , i.e., $\theta_{t+d,j} = \theta_{t,j}; j = 1, \dots, p_t, p_{t+d} = p_t, t \in \mathbb{Z}$. Ainsi, le processus $\{y_t^{(n)}, t \in \mathbb{Z}\}$ est une solution de l'équation aux différences linéaire stochastique à ordres et coefficient périodiques :

$$y_{i+d\tau}^{(n)} = \sum_{j=1}^{p_t} \theta_{i,j} y_{i-j+d\tau}^{(n)} + \varepsilon_{i+d\tau}^{(n)}; i = 1, \dots, d; \tau \in \mathbb{Z}, \quad (5.2.2a)$$

Let

$$\underline{\theta}_i = (\theta_{i,1}, \theta_{i,2}, \dots, \theta_{i,p_i})' ; i = 1, \dots, d \text{ et } \underline{\theta} = (\underline{\theta}'_1, \underline{\theta}'_2, \dots, \underline{\theta}'_d)' \quad \underline{\theta} \text{ non spécifié}$$

le signe $(.)'$ désigne la transposition ; $\{\varepsilon_t^{(n)} ; t \in \mathbb{Z}\}$ est un processus d'innovation, i.e., une suite de variables *i.i.d* de densité f_{σ_t} de moyenne nulle et de variance finie σ_t^2 , non spécifiée et périodique de période d .

$H_f^{(n)}(\underline{\theta}^{(n)})$ désigne une suite d'hypothèses alternatives locales, contiguës aux hypothèses nulles $H_f^{(n)}(\underline{\theta})$, sous lesquelles nous supposons que le processus $y_t^{(n)}$ est une solution de l'équation aux différences linéaire stochastique de la forme

$$y_t^{(n)} = \sum_{j=1}^{p_t} \theta_{t,j}^{(n)} y_{t-j}^{(n)} + \varepsilon_t^{(n)}, \quad (5.2.2b)$$

(Modèle $PAR(p_t)$ d -périodique localement perturbé) ; $\varepsilon_t^{(n)}$ est un processus indépendant, de moyenne nulle, de variance finie σ_t^2 et de densité $f_{\sigma_t}(\cdot)$. Les coefficients $\theta_{t,j}^{(n)} ; j = 1, \dots, p_t ; t \in \mathbb{Z}$, tels que $\theta_{t,p}^{(n)} \neq 0, \forall t \in \mathbb{Z}$ (modèle $AR(p_t)$ dépendant du temps, minimal d'ordre périodique p_t) sont périodiques en t , de période d ($d \geq 2$), i.e., $\theta_{t+d,j}^{(n)} = \theta_{t,j}^{(n)} ; j = 1, \dots, p_t, t \in \mathbb{Z}$

$\forall t \in \mathbb{Z}, i \in \{1, \dots, d\}$ et $\tau \in \mathbb{Z}$, tels que $t = i + d\tau$; alors, le modèle (5.2.2b) s'écrit encore :

$$y_{i+d\tau}^{(n)} = \sum_{j=1}^{p_i} \theta_{i,j}^{(n)} y_{i-j+d\tau}^{(n)} + \varepsilon_{i+d\tau}^{(n)} ; i = 1, \dots, d ; \tau \in \mathbb{Z}, \quad (5.2.3)$$

où les constantes

$$\theta_{i,j}^{(n)} = \theta_{i,j} + \frac{h_{i,j}^{(n)}}{\sqrt{n}}, \quad \underline{\theta}^{(n)} = (\underline{\theta}_1^{(n)'}, \underline{\theta}_2^{(n)'}, \dots, \underline{\theta}_d^{(n)'})' \in \mathbb{R}^p \text{ avec } \underline{\theta}_i^{(n)} = (\theta_{i,1}^{(n)}, \theta_{i,2}^{(n)}, \dots, \theta_{i,p_i}^{(n)})'$$

et $\frac{h_{i,j}^{(n)}}{\sqrt{n}}$ sont les perturbations périodiques locales qui dépendent de la saison i et de j .

Pour des raisons techniques, on suppose que la suite de vecteurs

$$\underline{\mathcal{I}}_{i,\cdot}^{(n)} = (h_{i,1}^{(n)}, h_{i,2}^{(n)}, \dots, h_{i,p_i}^{(n)})' \in \mathbb{R}^{p_i}, \quad (5.2.4)$$

est telle que $\sup_n \sum_{j=1}^{p_i} \left(h_{i,j}^{(n)} \right)^2 < \infty; i = 1, \dots, d \iff \sup_n \left\| \underline{\tau}_{i,\cdot}^{(n)} \right\|^2 < \infty, i = 1, \dots, d$

ainsi

$$\underline{\theta}_i^{(n)} = \underline{\theta}_i + \frac{\underline{\tau}_{i,\cdot}^{(n)}}{\sqrt{n}} \text{ et } \underline{\tau}^{(n)} = \left(\underline{\tau}_{1,\cdot}^{(n)}, \underline{\tau}_{2,\cdot}^{(n)}, \dots, \underline{\tau}_{d,\cdot}^{(n)} \right)' \in \mathbb{R}^p.$$

5.4 Résultats préliminaires : le modèle et le modèle multivarié associé

Dans ce chapitre, le modèle d'intérêt est donné par l'équation aux différences stochastique suivante :

$$y_t^{(n)} - \sum_{j=1}^{p_t} \theta_{t,j} y_{t-j}^{(n)} = \varepsilon_t^{(n)}, t \in \mathbb{Z},$$

où l'on suppose que les paramètres autorégressifs, les ordres du modèle et la variance du bruit blanc sont périodiques du temps, i.e.,

$$\theta_{t+d,j} = \theta_{t,j}, j = 1, \dots, p_t, p_{t+d} = p_t, \sigma_{t+d}^2 = \sigma_t^2, \forall t \in \mathbb{Z}.$$

Gladyshev (1961) a montré le lien existant entre les processus périodiquement corrélés de période d et les processus multivariés stationnaires au second ordre de dimension d .

Soit $X_i(\tau) = y_{i+d\tau}^{(n)}; i = 1, \dots, d, \tau \in \mathbb{Z}$. Alors le modèle (1.1) est équivalent à

$$\Theta(0)X(\tau) + \Theta(1)X(\tau - 1) + \dots + \Theta(p)X(\tau - p) = \eta(\tau),$$

où $\eta(\tau) = (\eta_1(\tau), \eta_2(\tau), \dots, \eta_d(\tau))$, $\eta_i(\tau) = \varepsilon_{i+d\tau}$, $i = 1, \dots, d; \tau \in \mathbb{Z}$, $p = \max_{1 \leq i \leq d} \left[\frac{p_i - i}{d} \right] + 1$, et où

$$(\Theta(0))_{r,s} = \begin{cases} 0, & \text{if } r < s, \\ 1, & \text{if } r = s, \\ \theta_{r,r-s}, & \text{if } r > s, \end{cases} \quad (5.3.1)$$

et

$$(\Theta(k))_{r,s} = \theta_{r,dk+r-s}, \quad r, s = 1, \dots, d$$

Définition 5.4.1 Un processus $\{y_t, t \in \mathbb{Z}\}$ est un processus périodiquement stationnaire au second ordre si c'est un processus autorégressif périodique et, de plus, si le processus vectoriel associé $\{X(\tau), \tau \in \mathbb{Z}\}$ est stationnaire au second ordre.

Proposition 5.4.2 Le processus univarié $PAR_d(p_t)$ est causal si le processus $VAR(p)$ de dimension d associé est causal, i.e., que les racines du polynôme déterminantal $\det(\Theta(0) + \sum_{j=1}^p \Theta(j)z^j)$ doivent être à l'extérieur du cercle unité.

H) On suppose tout le long que le processus $\{y_t^{(n)}, t \in \mathbb{Z}\}$ est périodiquement stationnaire au second ordre. Ainsi, on suppose que l'équation aux différences stochastique périodique (5.2.1) possède une solution strictement stationnaire pour chaque période $i; i = 1, \dots, d$ en raison du comportement asymptotique du rapport de vraisemblance qui est le même pour chaque période fixée $i, i = 1, \dots, d$. (Billingsley (1961))

Aussi, on peut supposer, sans restreindre la généralité que le processus $\{y_{i+d\tau}^{(n)}\}_{\tau \geq 0}$ est strictement stationnaire et ergodique.

Les conditions de régularité

Afin de prouver la propriété de normalité locale asymptotique (LAN (local asymptotic normality)), on impose à la densité f_{σ_t} de vérifier les conditions de régularité suivantes :

CR1 : $\varepsilon_t^{(n)}, t \in \mathbb{Z}$ sont i.i.d de fonction densité f_{σ_t} qui n'est pas nécessairement gaussienne.

$$\int x f_{\sigma_t}(x) dx = 0 \text{ et } \int x^2 f_{\sigma_t}(x) dx = \sigma_t^2 \text{ qui est périodique.}$$

CR2 $f_{\sigma_t}(x) > 0, \forall x \in \mathbb{R}, \forall t \in \mathbb{Z}$ et $f_{\sigma_t}(x) = \frac{1}{\sigma_t} f_1\left(\frac{x}{\sigma_t}\right)$ i.e., σ_t est un paramètre d'échelle.

CR3 $f_{\sigma_t}(\cdot)$ est absolument continue, i.e., $\exists \dot{f}_{\sigma_t}$ telle que pour $\forall a < b \in \mathbb{R}$

$$f_{\sigma_t}(b) - f_{\sigma_t}(a) = \int_a^b \dot{f}_{\sigma_t}(x) d\mu(x),$$

$f_{\sigma_t}(\cdot)$ est à information de Fisher finie, i.e., $I(f_{\sigma_t}) = E(\varphi_{\sigma_t}(X_t))^2 < \infty$ où $\varphi_{\sigma_t}(\cdot) =$

$-\frac{\dot{f}_{\sigma_t}(\cdot)}{f_{\sigma_t}(\cdot)}$ est la fonction score correspondant à la densité $f_{\sigma_t}(\cdot)$. $I(f_{\sigma_t}) = \frac{1}{\sigma_t^2} I(f_1)$ et

$$\varphi_{\sigma_t}(x) = \frac{1}{\sigma_t} \varphi_{f_1}\left(\frac{x}{\sigma_t}\right).$$

D'après les conditions précédentes $f_{\sigma_t}^{\frac{1}{2}}$ est différentiable en moyenne quadratique, i.e.,

$$\lim_{\lambda \rightarrow 0} \lambda^{-2} \int_{-\infty}^{+\infty} \left[f_{\sigma_t}^{\frac{1}{2}}(x + \lambda) - f_{\sigma_t}^{\frac{1}{2}}(x) - \lambda \frac{\dot{f}_{\sigma_t}(x)}{2f_{\sigma_t}^{\frac{1}{2}}(x)} \right]^2 dx = 0,$$

(cf. Hájek et Šidak (1967))

Suite de rapports de vraisemblance

Notons $\underline{y}^{(n)} = (y_1^{(n)}, \dots, y_n^{(n)})$ une réalisation de taille n , n fini, d'un processus $\{y_t^{(n)}; t \in \mathbb{Z}\}$ et soit $\underline{y}_0^{(n)} = (y_{\min(i-p_i)}^{(n)}, \dots, y_{-1}^{(n)}, y_0^{(n)})$, $1 + \max_{1 \leq i \leq d} (p_i - i)$ valeurs initiales pour la suite d'équations aux différences linéaires stochastiques (1.1.2) dont les densités, sous $H_f^{(n)}(\underline{\theta})$ et $H_f^{(n)}(\underline{\theta}^{(n)})$ sont respectivement : $g_0^{(n)}(y_{\min(i-p_i)}^{(n)}, \dots, y_{-1}^{(n)}, y_0^{(n)}; \underline{\theta}, \underline{\sigma})$ et $g_0^{(n)}(y_{\min(i-p_i)}^{(n)}, \dots, y_{-1}^{(n)}, y_0^{(n)}; \underline{\theta}^{(n)}, \underline{\sigma})$, $\underline{\sigma} = (\sigma_1, \dots, \sigma_d)'$.

Supposons de plus :

CR4 : La suite $g_0^{(n)}(y_{\min(i-p_i)}^{(n)}, \dots, y_{-1}^{(n)}, y_0^{(n)}; \underline{\theta}^{(n)}, \underline{\sigma}) - g_0^{(n)}(y_{\min(i-p_i)}^{(n)}, \dots, y_{-1}^{(n)}, y_0^{(n)}; \underline{\theta}, \underline{\sigma})$ converge en probabilité vers 0 lorsque $\underline{\theta}^{(n)} \rightarrow \underline{\theta}$ quand $n \rightarrow \infty$.

En notant, $Z_{i,\tau}^{(n)}(\underline{\theta}_i)$, pour $i = 1, \dots, d$ et $\tau \in \mathbb{Z}$, les résidus calculés sous la suite d'hypothèses nulles $H_f^{(n)}(\underline{\theta})$, qui, dans ce cas coïncident avec le bruit blanc du modèle, et par $Z_{i,\tau}^{(n)}(\underline{\theta}_i^{(n)})$ les résidus calculés sous la suite d'hypothèses alternatives locales $H_f^{(n)}(\underline{\theta}^{(n)})$, on vérifie que : $Z_{i,\tau}^{(n)}(\underline{\theta}_i^{(n)}) = Z_{i,\tau}^{(n)}(\underline{\theta}_i) - \gamma_{i,\tau}^{(n)}$ où $\gamma_{i,\tau}^{(n)} = \frac{1}{\sqrt{n}} \sum_{j=1}^{p_i} h_{i,j}^{(n)} y_{i-j+d\tau}^{(n)}$.

La suite des logarithmes de rapports de vraisemblance pour la suite d'hypothèses nulles

$H_f^{(n)}(\underline{\theta})$ contre la suite d'hypothèses alternatives locales $H_f^{(n)}(\underline{\theta}^{(n)})$, notée $\Lambda_{(\underline{\theta} + \nu^{(n)} \underline{\tau}^{(n)})/\underline{\theta}}^{(n)}$ est localement donnée par :

$$\begin{aligned} \Lambda_{(\underline{\theta} + \nu^{(n)} \underline{\tau}^{(n)})/\underline{\theta}}^{(n)} &= \sum_{i=1}^d \sum_{\tau=0}^{m-1} \log \frac{f_{\sigma_i} \left(Z_{i,\tau}^{(n)}(\underline{\theta}_i) - \gamma_{i,\tau}^{(n)} \right)}{f_{\sigma_i} \left(Z_{i,\tau}^{(n)}(\underline{\theta}_i) \right)} \\ &\quad + \log \frac{g_0^{(n)} \left(y_{\min(i-p_i)}^{(n)}, \dots, y_{-1}^{(n)}, y_0; \underline{\theta}^{(n)}, \underline{\sigma} \right)}{g_0^{(n)} \left(y_{\min(i-p_i)}^{(n)}, \dots, y_{-1}^{(n)}, y_0; \underline{\theta}, \underline{\sigma} \right)}. \end{aligned} \quad (5.3.2)$$

Contrairement à Swensen (1985), Kreiss (1987) et Bentarzi (1995) où sous l'hypothèse nulle le processus est stationnaire, dans ce contexte, le processus est périodiquement corrélé ce qui nécessite l'hypothèse H) donnée ci-dessus.

De là, découle l'approximation locale asymptotique :

$$\Lambda_{(\underline{\theta} + \nu^{(n)} \underline{\tau}^{(n)})/\underline{\theta}}^{(n)} = \sum_{i=1}^d \sum_{\tau=0}^{m-1} \log \frac{f_{\sigma_i} \left(Z_{i,\tau}^{(n)}(\underline{\theta}_i) - \gamma_{i,\tau}^{(n)} \right)}{f_{\sigma_i} \left(Z_{i,\tau}^{(n)}(\underline{\theta}_i) \right)} + o_p(1), \quad (5.3.3)$$

sous $H_f^{(n)}(\underline{\theta})$, où pour des raisons de simplicité de notation, nous avons supposé que la taille n de la série chronologique observée est un multiple de d , i.e., $n = md, m \in \mathbb{N}^*$. Alors, pour tout $t \in \{1, \dots, n\}$, nous pouvons toujours écrire $t = i + d\tau$, où $\tau = 0, 1, \dots, m-1$ et $i = 1, \dots, d$.

5.5 Lemme de Swensen (1985) et les conditions suffisantes de la propriété LAN

Soit pour $i = 1, \dots, d$ et $\tau = 0, \dots, m-1$, les variables aléatoires suivantes :

$$\begin{aligned} X_{i,\tau}^{(n)}(\underline{\theta}_i) &= \frac{f_{\sigma_i}^{\frac{1}{2}} \left(Z_{i,\tau}^{(n)}(\underline{\theta}_i^{(n)}) \right)}{f_{\sigma_i}^{\frac{1}{2}} \left(Z_{i,\tau}^{(n)}(\underline{\theta}_i) \right)} - 1, \\ &= \frac{f_{\sigma_i}^{\frac{1}{2}} \left(Z_{i,\tau}^{(n)}(\underline{\theta}_i) - \gamma_{i,\tau}^{(n)} \right)}{f_{\sigma_i}^{\frac{1}{2}} \left(Z_{i,\tau}^{(n)}(\underline{\theta}_i) \right)} - 1. \end{aligned}$$

Posons

$$\zeta_{i,\tau}^{(n)}(\underline{\theta}_i) = \frac{1}{2} \varphi_{\sigma_i} \left(Z_{i,\tau}^{(n)}(\underline{\theta}_i) \right) \gamma_{i,\tau}^{(n)} \text{ où } \varphi_{\sigma_i}(\cdot) = -\frac{\dot{f}_{\sigma_i}(\cdot)}{f_{\sigma_i}(\cdot)}. \quad (5.4.1)$$

Soit les vecteurs aléatoires $\underline{\delta}^{(n)}(\underline{\theta})$ et $\underline{y}_{i,\tau}^{(n)}; i = 1, \dots, d$ et $\tau \in \mathbb{Z}$ définis comme suit :

$$\underline{\delta}_{i,.,\tau}^{(n)}(\underline{\theta}_i) = \left(\underline{\delta}_{i,1,\tau}^{(n)}(\underline{\theta}_i), \underline{\delta}_{i,2,\tau}^{(n)}(\underline{\theta}_i), \dots, \underline{\delta}_{i,p_i,\tau}^{(n)}(\underline{\theta}_i) \right)' \in \mathbb{R}^{p_i},$$

où : $\underline{\delta}_{i,j,\tau}^{(n)}(\underline{\theta}_i) = \frac{1}{\sqrt{n}} \varphi_{\sigma_i} \left(Z_{i,\tau}^{(n)}(\underline{\theta}_i) \right) y_{i-j+d\tau}^{(n)}$

Soit $\underline{\delta}_{i,.}^{(n)}(\underline{\theta}_i) = \left(\underline{\delta}_{i,1}^{(n)}(\underline{\theta}_i), \dots, \underline{\delta}_{i,p_i}^{(n)}(\underline{\theta}_i) \right)' \in \mathbb{R}^{p_i}$, avec :

$$\begin{aligned} \underline{\delta}_{i,j}^{(n)}(\underline{\theta}_i) &= \frac{1}{\sqrt{n}} \sum_{\tau=0}^{m-1} \varphi_{\sigma_i} \left(Z_{i,\tau}^{(n)}(\underline{\theta}_i) \right) y_{i-j+d\tau}^{(n)} \text{ et } \underline{\delta}^{(n)}(\underline{\theta}), \\ &= \left(\underline{\delta}_{1,.}^{(n)'}(\underline{\theta}_1), \underline{\delta}_{2,.}^{(n)'}(\underline{\theta}_2), \dots, \underline{\delta}_{d,.}^{(n)'}(\underline{\theta}_d) \right) \in \mathbb{R}^{\sum_{i=1}^d p_i}, \\ \underline{y}_{i,\tau}^{(n)} &= \left(y_{i-1+d\tau}^{(n)}, \dots, y_{i-p_i+d\tau}^{(n)} \right)' \end{aligned}$$

Sa matrice de variance covariance notée $\Gamma^i(\underline{\theta}, \underline{\sigma})$ sous $H_f^{(n)}(\underline{\theta})$ de dimension $p_i \times p_i$ est donnée par son élément général :

$$\begin{aligned} (\Gamma^i(\underline{\theta}, \underline{\sigma}))_{k,l} &= \sum_{u=-\infty}^{i+d\tau-\max(k,l)} g(i-k+d\tau, u)(\underline{\theta}) g(i-l+d\tau, u)(\underline{\theta}) \sigma_u^2, \\ &= \sum_{s=1}^{i-\max(k,l)+d|v|} \sum_{\tau'=-\infty}^{\tau-|v|} g(i-k+d\tau, s+d\tau')(\underline{\theta}) \times g(i-l+d\tau, s+d\tau')(\underline{\theta}) \sigma_s^2 \end{aligned}$$

où v est le premier entier tel que

$$1 - d|v| \leq i - \max(k, l) \leq d(1 - |v|), v \in \mathbb{Z}, k, l = 1, \dots, p_i, i = 1, \dots, d.$$

où $g(t, u)$ $t, u \in \mathbb{Z}$ est la fonction de Green unilatérale liée à l'opérateur aux différences associé à l'équation stochastique (5.2.1).

Remarque 5.5.1 : $(\Gamma^i(\underline{\theta}, \underline{\sigma}))_{k,l}$ ne dépend de $t = i + d\tau$ que de par la période i (i.e., elle ne dépend pas de τ) and $\widehat{\Gamma}^i = \frac{1}{m} \sum_{\tau=0}^{m-1} \underline{y}_{i,\tau}^{(n)} \underline{y}_{i,\tau}^{(n)'}$ est un estimateur de $\Gamma^i(\underline{\theta}, \underline{\sigma})$.

Soit $\Gamma(\underline{\theta}, \underline{\sigma})$ la matrice carrée blocs diagonale de dimension $\sum_{i=1}^d p_i \times \sum_{i=1}^d p_i$, donnée par :

$$\Gamma(\underline{\theta}, \underline{\sigma}) = \begin{pmatrix} \frac{\Gamma^1(\underline{\theta}, \underline{\sigma})}{\sigma_1^2} & \cdots & \cdots & 0 \\ 0 & \frac{\Gamma^2(\underline{\theta}, \underline{\sigma})}{\sigma_2^2} & \cdots & 0 \\ \vdots & 0 & \ddots & 0 \\ 0 & 0 & 0 & \frac{\Gamma^d(\underline{\theta}, \underline{\sigma})}{\sigma_d^2} \end{pmatrix}$$

Remarque 5.5.2 : $\zeta_{i,\tau}^{(n)}(\underline{\theta}_i)$ donné par (5.4.1) s'écrit $\zeta_{i,\tau}^{(n)}(\underline{\theta}_i) = \frac{1}{2} \underline{\delta}_{i,\tau}^{(n)'}(\underline{\theta}_i) \underline{\tau}_{i,\tau}^{(n)}$; $i = 1, \dots, d, \tau = 0, \dots, m-1$ où $\underline{\tau}_{i,\tau}^{(n)}$ est donné par (5.2.4) On a encore :

$$\zeta_{i,\tau}^{(n)}(\underline{\theta}_i) = \frac{1}{2\sigma_i} \varphi_{f_1} \left(\frac{Z_{i,\tau}^{(n)}(\underline{\theta}_i)}{\sigma_i} \right) \gamma_{i,\tau}^{(n)}; i = 1, \dots, d, \tau = 0, \dots, m-1$$

Sous les notations précédentes, nous pouvons énoncer la proposition ci-dessous, qui fournit les conditions suffisantes, que nous avons adaptées à notre contexte, pour l'obtention de la propriété LAN du modèle $PAR(p_t)$ évolutif d -périodique.

Proposition 5.5.3 (Conditions de Swensen 1985)

Sous $CR1$, $CR2$ et $CR3$, les conditions suivantes sont satisfaites et impliquent celles de Le Cam :

- (1) $\lim_{m \rightarrow \infty} E \sum_{i=1}^d \sum_{\tau=0}^{m-1} \left(X_{i,\tau}^{(n)}(\underline{\theta}_i) - \zeta_{i,\tau}^{(n)}(\underline{\theta}_i) \right)^2 = 0,$
- (2) $\sup_m E \sum_{i=1}^d \sum_{\tau=0}^{m-1} \left(\zeta_{i,\tau}^{(n)}(\underline{\theta}_i) \right)^2 < \infty,$
- (3) $\max_i \max_{\tau} \left| \zeta_{i,\tau}^{(n)}(\underline{\theta}_i) \right| = o_p(1),$
- (4) $\sum_{i=1}^d \sum_{\tau=0}^{m-1} \left(\zeta_{i,\tau}^{(n)}(\underline{\theta}_i) \right)^2 = \frac{I(f_1)}{4d} \underline{\tau}^{(n)'} \Gamma(\underline{\theta}, \underline{\sigma}) \underline{\tau}^{(n)} + o_p(1),$
- (5) $\sum_{i=1}^d \sum_{\tau=0}^{m-1} E \left[\left(\zeta_{i,\tau}^{(n)}(\underline{\theta}_i) \right)^2 I_{\left[\left| \zeta_{i,\tau}^{(n)}(\underline{\theta}_i) \right| > \frac{1}{2} \right]} / B_{n,i-1+d\tau} \right] = o_p(1),$ où $B_{n,i-1+d\tau}$ est la tribu engendrée par le passé du processus jusqu'au temps $i-1+d\tau$.

$$(6) \quad E \left(\zeta_{i,\tau}^{(n)}(\underline{\theta}_i) / B_{n,i-1+d\tau} \right) = 0_p(1),$$

$$(7) \quad \lim_{m \rightarrow \infty} \sum_{i=1}^d \sum_{\tau=0}^{m-1} E \left((X_{i,\tau}^{(n)}(\underline{\theta}_i))^2 + 2X_{i,\tau}^{(n)}(\underline{\theta}_i) / B_{n,i-1+d\tau} \right) = 0_p(1)$$

Démonstration. Preuve de la condition (1). $\lim_{m \rightarrow \infty} E \left(\sum_{i=1}^d \sum_{\tau=0}^{m-1} \left(X_{i,\tau}^{(n)}(\underline{\theta}_i) - \zeta_{i,\tau}^{(n)}(\underline{\theta}_i) \right)^2 \right).$

En remplaçant $X_{i,\tau}^{(n)}(\underline{\theta}_i)$ et $\zeta_{i,\tau}^{(n)}(\underline{\theta}_i)$ par leurs valeurs respectives dans

$$E \left(\sum_{i=1}^d \sum_{\tau=0}^{m-1} \left(X_{i,\tau}^{(n)}(\underline{\theta}_i) - \zeta_{i,\tau}^{(n)}(\underline{\theta}_i) \right)^2 \right),$$

on obtient :

$$E \sum_{i=1}^d \sum_{\tau=0}^{m-1} \left(f_{\sigma_i}^{\frac{1}{2}} \left(Z_{i,\tau}^{(n)}(\underline{\theta}_i) - \gamma_{i,\tau}^{(n)} \right) - f_{\sigma_i}^{\frac{1}{2}} \left(Z_{i,\tau}^{(n)}(\underline{\theta}_i) \right) - \frac{1}{2}(-\gamma_{i,\tau}^{(n)}) \frac{\dot{f}_{\sigma_i} \left(Z_{i,\tau}^{(n)}(\underline{\theta}_i) \right)}{f_{\sigma_i}^{\frac{1}{2}} \left(Z_{i,\tau}^{(n)}(\underline{\theta}_i) \right)} \right)^2 \\ \times f_{\sigma_i}^{-1} \left(Z_{i,\tau}^{(n)}(\underline{\theta}_i) \right)$$

Pour montrer que cette expression tend vers 0 quand $m \rightarrow \infty$, il suffit de montrer que, pour i fixé, $i = 1, \dots, d$,

$$E \sum_{\tau=0}^{m-1} \left(f_{\sigma_i}^{\frac{1}{2}} \left(Z_{i,\tau}^{(n)}(\underline{\theta}_i) - \gamma_{i,\tau}^{(n)} \right) - f_{\sigma_i}^{\frac{1}{2}} \left(Z_{i,\tau}^{(n)}(\underline{\theta}_i) \right) - \frac{1}{2}(-\gamma_{i,\tau}^{(n)}) \frac{\dot{f}_{\sigma_i} \left(Z_{i,\tau}^{(n)}(\underline{\theta}_i) \right)}{f_{\sigma_i}^{\frac{1}{2}} \left(Z_{i,\tau}^{(n)}(\underline{\theta}_i) \right)} \right)^2 f_{\sigma_i}^{-1} \left(Z_{i,\tau}^{(n)}(\underline{\theta}_i) \right)$$

tend vers 0. Cette quantité s'écrit encore

$$E \left\{ \sum_{\tau=0}^{m-1} \left(\frac{1}{\sqrt{\sigma_i}} f_1^{\frac{1}{2}} \left(\frac{Z_{i,\tau}^{(n)}(\underline{\theta}_i) - \gamma_{i,\tau}^{(n)}}{\sigma_i} \right) - \frac{1}{\sqrt{\sigma_i}} f_1^{\frac{1}{2}} \left(\frac{Z_{i,\tau}^{(n)}(\underline{\theta}_i)}{\sigma_i} \right) - \frac{1}{2}(-\gamma_{i,\tau}^{(n)}) \frac{\frac{1}{\sigma_i^2} \dot{f}_1 \left(\frac{Z_{i,\tau}^{(n)}(\underline{\theta}_i)}{\sigma_i} \right)}{\frac{1}{\sqrt{\sigma_i}} f_1^{\frac{1}{2}} \left(\frac{Z_{i,\tau}^{(n)}(\underline{\theta}_i)}{\sigma_i} \right)} \right) \right. \\ \left. \times \sigma_i f_1^{-1} \left(\frac{Z_{i,\tau}^{(n)}(\underline{\theta}_i)}{\sigma_i} \right) \right\} \\ = E \left\{ \sum_{\tau=0}^{m-1} \left(f_1^{\frac{1}{2}} \left(\frac{Z_{i,\tau}^{(n)}(\underline{\theta}_i) - \gamma_{i,\tau}^{(n)}}{\sigma_i} \right) - f_1^{\frac{1}{2}} \left(\frac{Z_{i,\tau}^{(n)}(\underline{\theta}_i)}{\sigma_i} \right) - \frac{1}{2}(-\gamma_{i,\tau}^{(n)}) \frac{1}{\sigma_i} \frac{\dot{f}_1 \left(\frac{Z_{i,\tau}^{(n)}(\underline{\theta}_i)}{\sigma_i} \right)}{f_1^{\frac{1}{2}} \left(\frac{Z_{i,\tau}^{(n)}(\underline{\theta}_i)}{\sigma_i} \right)} \right) \right\}^2$$

Soit K un réel positif quelconque, alors l'espérance peut-être décomposée comme suit

$$\begin{aligned}
 & \sum_{\tau=0}^{m-1} E \left\{ \left(f_1^{\frac{1}{2}} \left(\frac{Z_{i,\tau}^{(n)}(\underline{\theta}_i) - \gamma_{i,\tau}^{(n)}}{\sigma_i} \right) - f_1^{\frac{1}{2}} \left(\frac{Z_{i,\tau}^{(n)}(\underline{\theta}_i)}{\sigma_i} \right) - \frac{1}{2}(-\gamma_{i,\tau}^{(n)}) \frac{1}{\sigma_i} \frac{\dot{f}_1 \left(\frac{Z_{i,\tau}^{(n)}(\underline{\theta}_i)}{\sigma_i} \right)}{f_1^{\frac{1}{2}} \left(\frac{Z_{i,\tau}^{(n)}(\underline{\theta}_i)}{\sigma_i} \right)} \right. \right. \\
 & \quad \left. \left. \times f_1^{-1} \left(\frac{Z_{i,\tau}^{(n)}(\underline{\theta}_i)}{\sigma_i} \right) \right) \right\} \\
 &= E \left\{ \sum_{\tau=0}^{m-1} \left(f_1^{\frac{1}{2}} \left(\frac{Z_{i,\tau}^{(n)}(\underline{\theta}_i) - \gamma_{i,\tau}^{(n)}}{\sigma_i} \right) - f_1^{\frac{1}{2}} \left(\frac{Z_{i,\tau}^{(n)}(\underline{\theta}_i)}{\sigma_i} \right) - \frac{1}{2}(-\gamma_{i,\tau}^{(n)}) \frac{1}{\sigma_i} \frac{\dot{f}_1 \left(\frac{Z_{i,\tau}^{(n)}(\underline{\theta}_i)}{\sigma_i} \right)}{f_1^{\frac{1}{2}} \left(\frac{Z_{i,\tau}^{(n)}(\underline{\theta}_i)}{\sigma_i} \right)} \right)^2 \right. \\
 &= \sum_{\tau=0}^{m-1} EI_{[|y_{i-j+d\tau}| < K \text{ pour tout } j; j=1, \dots, p_i]} \left(f_1^{\frac{1}{2}} \left(\frac{Z_{i,\tau}^{(n)}(\underline{\theta}_i) - \gamma_{i,\tau}^{(n)}}{\sigma_i} \right) - f_1^{\frac{1}{2}} \left(\frac{Z_{i,\tau}^{(n)}(\underline{\theta}_i)}{\sigma_i} \right) \right. \\
 & \quad \left. - \frac{1}{2}(-\gamma_{i,\tau}^{(n)}) \frac{1}{\sigma_i} \left(\frac{\dot{f}_1 \left(\frac{Z_{i,\tau}^{(n)}(\underline{\theta}_i)}{\sigma_i} \right)}{f_1^{\frac{1}{2}} \left(\frac{Z_{i,\tau}^{(n)}(\underline{\theta}_i)}{\sigma_i} \right)} \right)^2 \times f_1^{-1} \left(\frac{Z_{i,\tau}^{(n)}(\underline{\theta}_i)}{\sigma_i} \right) \right\} \\
 &+ \sum_{\tau=0}^{m-1} E \left\{ I_{[|y_{i-j+d\tau}| \geq K \text{ pour quelque valeurs de } j; j=1, \dots, p_i]} f_1^{\frac{1}{2}} \left(\frac{Z_{i,\tau}^{(n)}(\underline{\theta}_i) - \gamma_{i,\tau}^{(n)}}{\sigma_i} \right) - f_1^{\frac{1}{2}} \left(\frac{Z_{i,\tau}^{(n)}(\underline{\theta}_i)}{\sigma_i} \right) \right. \\
 & \quad \left. - \frac{1}{2}(-\gamma_{i,\tau}^{(n)}) \frac{1}{\sigma_i} \frac{\dot{f}_1 \left(\frac{Z_{i,\tau}^{(n)}(\underline{\theta}_i)}{\sigma_i} \right)}{f_1^{\frac{1}{2}} \left(\frac{Z_{i,\tau}^{(n)}(\underline{\theta}_i)}{\sigma_i} \right)} \right)^2 \times f_1^{-1} \left(\frac{Z_{i,\tau}^{(n)}(\underline{\theta}_i)}{\sigma_i} \right) \right\}
 \end{aligned}$$

Montrons que la première somme, notée $B_{1,m}$ tend vers 0, pour tout K .

$$\begin{aligned}
 B_{1,m} &= \sum_{\tau=0}^{m-1} \int_{\{|y_{i-j+d\tau}| < K; j=1, \dots, p_i\}} \frac{\gamma_{i,\tau}^{(n)2}}{\sigma_i^2} \int_{\{|y_{i-j+d\tau}| < K; j=1, \dots, p_i\}} \left[\frac{f_1^{\frac{1}{2}} \left(\frac{Z_{i,\tau}^{(n)}(\underline{\theta}_i) - \gamma_{i,\tau}^{(n)}}{\sigma_i} \right) - f_1^{\frac{1}{2}} \left(\frac{Z_{i,\tau}^{(n)}(\underline{\theta}_i)}{\sigma_i} \right)}{\frac{-\gamma_{i,\tau}^{(n)}}{\sigma_i}} \right. \\
 & \quad \left. - \frac{1}{2} \frac{\dot{f}_1 \left(\frac{Z_{i,\tau}^{(n)}(\underline{\theta}_i)}{\sigma_i} \right)}{f_1^{\frac{1}{2}} \left(\frac{Z_{i,\tau}^{(n)}(\underline{\theta}_i)}{\sigma_i} \right)} \right]^2 d \frac{Z_{i,\tau}^{(n)}(\underline{\theta}_i)}{\sigma_i} dG_Y
 \end{aligned}$$

où G_Y est la distribution conjointe des variables aléatoires $y_{i-j+d\tau}; j = 1, \dots, p_i$, $B_{1,m}$ peut-être

majorée par

$$B_{1,m} \leq \sum_{\tau=0}^{m-1} \int_{R^{p_i}} \frac{\gamma_{i,\tau}^{(n)2}}{\sigma_i^2} \left[\int_{\{|y_{i-j+d\tau}| < K; j=1, \dots, p_i\}} \left[\frac{f_1^{\frac{1}{2}} \left(\frac{Z_{i,\tau}^{(n)}(\underline{\theta}_i) - \gamma_{i,\tau}^{(n)}}{\sigma_i} \right) - f_1^{\frac{1}{2}} \left(\frac{Z_{i,\tau}^{(n)}(\underline{\theta}_i)}{\sigma_i} \right)}{\frac{-\gamma_{i,\tau}^{(n)}}{\sigma_i}} - \frac{\dot{f}_1 \left(\frac{Z_{i,\tau}^{(n)}(\underline{\theta}_i)}{\sigma_i} \right)}{f_1^{\frac{1}{2}} \left(\frac{Z_{i,\tau}^{(n)}(\underline{\theta}_i)}{\sigma_i} \right)} \right]^2 d \frac{Z_{i,\tau}^{(n)}(\underline{\theta}_i)}{\sigma_i} \right] dG_Y$$

Soit $u_i^{(n)} = \sqrt{n} \left(-\frac{\gamma_{i,\tau}^{(n)}}{\sigma_i} \right); i = 1, \dots, d$ et $\tau = 0, \dots, m-1$. Si on pose $x_i = \frac{Z_{i,\tau}^{(n)}(\underline{\theta}_i)}{\sigma_i}$, alors

l'intégrale entre crochets ci-dessus peut-être majorée comme suit

$$\begin{aligned} & \int_{\{|y_{i-j+d\tau}| < K; j=1, \dots, p_i\}} \left[\left(\frac{f_1^{\frac{1}{2}}(x_i + n^{-\frac{1}{2}}u_i^{(n)}) - f_1^{\frac{1}{2}}(x_i)}{n^{-\frac{1}{2}}u_i^{(n)}} - \frac{\frac{1}{2}\dot{f}_1(x_i)}{f_1^{\frac{1}{2}}(x_i)} \right)^2 dx_i \right] - \frac{\dot{f}_1(x_i)}{f_1^{\frac{1}{2}}(x_i)} \Big]^2 dx_i \\ & \leq \limsup_{\left| u_i^{(n)} \right| \leq \sum_{j=1}^{p_i} \frac{|h_{ij}(n)|}{\sigma_i} . K} \int_R \left[\frac{f_1^{\frac{1}{2}}(x_i + n^{-\frac{1}{2}}u_i^{(n)}) - f_1^{\frac{1}{2}}(x_i)}{n^{-\frac{1}{2}}u_i^{(n)}} - \frac{\frac{1}{2}\dot{f}_1(x_i)}{f_1^{\frac{1}{2}}(x_i)} \right]^2 dx_i = C_n(K) \end{aligned}$$

Puisque $f_1^{\frac{1}{2}}$ est différentiable en moyenne quadratique et $u_i^{(n)}$ est bornée, la quantité ci-dessus tend vers 0 lorsque $n \rightarrow \infty$.

Par conséquent, $B_{1,m} \leq \sum_{\tau=0}^{m-1} E \left(\frac{\gamma_{i,\tau}^{(n)2}}{\sigma_i^2} \right) \times C_n(K)$ et $C_n(K)$ tend vers 0 quand $n \rightarrow \infty$.

Il reste à montrer que $\sum_{\tau=0}^{m-1} E \left(\frac{\gamma_{i,\tau}^{(n)2}}{\sigma_i^2} \right)$ est une quantité finie. En suivant les mêmes étapes

que Bentarzi (1995)

$$\sum_{\tau=0}^{m-1} E \left(\frac{\gamma_{i,\tau}^{(n)2}}{\sigma_i^2} \right) = \frac{1}{n\sigma_i^2} \sum_{\tau=0}^{m-1} \sum_{j_1=1}^{p_i} \sum_{j_2=1}^{p_i} h_{ij_1} h_{ij_2} E(y_{i-j_1+d\tau}^{(n)} y_{i-j_2+d\tau}^{(n)}) = \frac{1}{d\sigma_i^2} \sum_{j_1=1}^{p_i} \sum_{j_2=1}^{p_i} h_{ij_1} h_{ij_2} \gamma_{j_1-j_2}^{i-j_1}$$

indépendant de τ car i -stationnaire. On peut, majorer cette quantité par

$$\max_{j_1, j_2} \gamma_{j_1 - j_2}^{i - j_1} \frac{1}{d\sigma_i^2} C(p_i) \sup_n \sum_{j=1}^{p_i} h_{ij}^{(n)2} < \infty.$$

Montrons que la deuxième somme notée $B_{2,m}$ tend vers 0 pour K suffisamment grand.

$$B_{2,m} = \sum_{\tau=0}^{m-1} E \left\{ I_{[|y_{i-j+d\tau}| \geq K \text{ pour } \forall j; j=1, \dots, p_i]} \left(f_1^{\frac{1}{2}} \left(\frac{Z_{i,\tau}^{(n)}(\underline{\theta}_i) - \gamma_{i,\tau}^{(n)}}{\sigma_i} \right) - f_1^{\frac{1}{2}} \left(\frac{Z_{i,\tau}^{(n)}(\underline{\theta}_i)}{\sigma_i} \right) \right) \right. \\ \left. - \frac{1}{2} (-\gamma_{i,\tau}^{(n)}) \frac{1}{\sigma_i} \left(\frac{f_1 \left(\frac{Z_{i,\tau}^{(n)}(\underline{\theta}_i)}{\sigma_i} \right)}{f_1^{\frac{1}{2}} \left(\frac{Z_{i,\tau}^{(n)}(\underline{\theta}_i)}{\sigma_i} \right)} \right)^2 \times f_1^{-1} \left(\frac{Z_{i,\tau}^{(n)}(\underline{\theta}_i)}{\sigma_i} \right) \right\}$$

Si on pose $x_i = \frac{Z_{i,\tau}^{(n)}(\underline{\theta}_i)}{\sigma_i}$, et $u_i^{(n)} = -\frac{\gamma_{i,\tau}^{(n)}}{\sigma_i}$; $i = 1, \dots, d$ et $\tau = 0, \dots, m-1$, $B_{2,m}$ peut être majorée comme suit

$$B_{2,m} \leq \sum_{\tau=0}^{m-1} \int_{\{|y_{i-j+d\tau}| \geq K, \text{ pour quelques valeurs de } j\}} [f_1^{\frac{1}{2}}(x_i + u_i^{(n)}) - f_1^{\frac{1}{2}}(x_i) - \frac{1}{2}u_i^{(n)}(-\varphi_{f_1}(x_i)f_1^{\frac{1}{2}}(x_i))]^2 dx_i dG_Y(\underline{y})$$

En appliquant la seconde partie du lemme 2 de Swensen(1985), on obtient

$$\int_R [f_1^{\frac{1}{2}}(x_i + u_i^{(n)}) - f_1^{\frac{1}{2}}(x_i) - \frac{1}{2}u_i^{(n)}(-\varphi_{f_1}(x_i)f_1^{\frac{1}{2}}(x_i))]^2 dx_i \leq u_i^{(n)2} I(f_1) = I(f_1) \frac{\gamma_{i,\tau}^{(n)2}}{\sigma_i^2}$$

et

$$B_{2,m} \leq \frac{I(f_1)}{\sigma_i^2} \sum_{\tau=0}^{m-1} E \left(I_{[|y_{i-j+d\tau}| \geq K \text{ pour quelques valeurs de } j]} \gamma_{i,\tau}^{(n)2} \right) \\ \sum_{\tau=0}^{m-1} E \left(I_{[|y_{i-j+d\tau}| \geq K \text{ pour quelques valeurs de } j]} \gamma_{i,\tau}^{(n)2} \right) \\ = \sum_{\tau=0}^{m-1} E \left(I_{[|y_{i-j+d\tau}| \geq K \text{ pour quelques valeurs de } j]} \frac{1}{n} \sum_{j_1=1}^{p_i} \sum_{j_2=1}^{p_i} h_{ij_1} h_{ij_2} y_{i-j_1+d\tau}^{(n)} y_{i-j_2+d\tau}^{(n)} \right), \\ = \frac{1}{n} \sum_{j_1=1}^{p_i} \sum_{j_2=1}^{p_i} h_{ij_1} h_{ij_2} \sum_{\tau=0}^{m-1} E \left(I_{[|y_{i-j+d\tau}| \geq K \text{ pour quelques valeurs de } j]} y_{i-j_1+d\tau}^{(n)} y_{i-j_2+d\tau}^{(n)} \right), \\ \leq \frac{1}{n} c(p_i) \sum_{j=1}^{p_i} h_{ij}^{(n)2} \sum_{\tau=0}^{m-1} E \left(I_{[|y_{i-j+d\tau}| \geq K \text{ pour quelques valeurs de } j]} y_{i-j+d\tau}^{(n)2} \right)$$

Le processus $\{y_{i+d\tau}\}_{\tau \geq 0}$ étant strictement stationnaire et $E(y_{i-j+d\tau}^{(n)2}) = \gamma_0^{i-j} < \infty$ alors $y_{i-j+d\tau}^{(n)2}$ est équi-intégrable et $E(I_{[|y_{i-j+d\tau}| \geq K \text{ pour quelques valeurs de } j]} y_{i-j+d\tau}^{(n)2}) \rightarrow 0$ lorsque $K \rightarrow \infty$ indépendamment de τ .

Preuve de la condition (2)

$$\zeta_{i,\tau}^{(n)}(\underline{\theta}_i) = \frac{1}{2\sigma_i} \varphi_{f_1}\left(\frac{Z_{i,\tau}^{(n)}(\underline{\theta}_i)}{\sigma_i}\right) \frac{1}{\sqrt{n}} \sum_{j=1}^{p_i} h_{i,j}^{(n)} y_{i-j+d\tau}^{(n)},$$

On doit montrer que $\sup_m E \sum_{i=1}^d \sum_{\tau=0}^{m-1} \left(\zeta_{i,\tau}^{(n)}(\underline{\theta}_i) \right)^2 < \infty$

Il suffit de montrer que $\sup_m E \sum_{\tau=0}^{m-1} \left(\zeta_{i,\tau}^{(n)}(\underline{\theta}_i) \right)^2 < \infty$ pour i fixé.

$$\begin{aligned} E \sum_{\tau=0}^{m-1} \left(\zeta_{i,\tau}^{(n)}(\underline{\theta}_i) \right)^2 &= \sum_{\tau=0}^{m-1} E \left(\zeta_{i,\tau}^{(n)}(\underline{\theta}_i) \right)^2 = \frac{1}{4} \sum_{\tau=0}^{m-1} E \left(\varphi_{\sigma_i}^2 \left(Z_{i,\tau}^{(n)}(\underline{\theta}_i) \right) \gamma_{i,\tau}^{(n)2} \right), \\ &= \frac{1}{4} \sum_{\tau=0}^{m-1} \frac{1}{\sigma_i^2} I(f_1) E \left(\gamma_{i,\tau}^{(n)2} \right) \end{aligned}$$

car sous $H_f^{(n)}(\underline{\theta})$ le résidu $Z_{i,\tau}^{(n)}(\underline{\theta}_i)$ est un bruit blanc indépendant du passé du processus, i.e., de $\gamma_{i,\tau}^{(n)}$. De plus $\sum_{\tau=0}^{m-1} E(\gamma_{i,\tau}^{(n)2})$ est finie comme cela a été montré dans la démonstration de la condition (1).

preuve de la condition (3)

$$\max_i \max_{\tau} \left| \zeta_{i,\tau}^{(n)}(\underline{\theta}_i) \right| = \max_i \max_{\tau} \frac{1}{2} \left| \varphi_{\sigma_i}(Z_{i,\tau}^{(n)}(\underline{\theta}_i)) \frac{1}{\sqrt{n}} \sum_{j=1}^{p_i} h_{i,j}^{(n)} y_{i-j+d\tau}^{(n)} \right| = o_p(1)$$

A la manière de Bentarzi (1995), il suffit de montrer que $\max_{\tau} |\zeta_{i,\tau}^{(n)}(\underline{\theta}_i)| = o_p(1)$.

$$\begin{aligned} P(\max_{\tau} |\zeta_{i,\tau}^{(n)}(\underline{\theta}_i)| > \delta) &\leq \frac{1}{\delta^2} \sum_{\tau=0}^{m-1} E \left(\zeta_{i,\tau}^{(n)2}(\underline{\theta}_i) I_{|\zeta_{i,\tau}^{(n)}(\underline{\theta}_i)| > \delta} \right) \\ &= \frac{1}{4n\delta^2} \sum_{\tau=0}^{m-1} E \left((\varphi_{\sigma_i}(Z_{i,\tau}^{(n)}(\underline{\theta}_i))) \sum_{j=1}^{p_i} h_{i,j}^{(n)} y_{i-j+d\tau}^{(n)})^2 I_{\left| \varphi_{\sigma_i}(Z_{i,\tau}^{(n)}(\underline{\theta}_i)) \sum_{j=1}^{p_i} h_{i,j}^{(n)} y_{i-j+d\tau}^{(n)} \right| > 2\delta\sqrt{n}} \right) \\ &= \frac{m}{4n\delta^2} E \left((\varphi_{\sigma_i}(Z_{i,1}^{(n)}(\underline{\theta}_i))) \sum_{j=1}^{p_i} h_{i,j}^{(n)} y_{i-j,1}^{(n)})^2 I_{\left| \varphi_{\sigma_i}(Z_{i,1}^{(n)}(\underline{\theta}_i)) \sum_{j=1}^{p_i} h_{i,j}^{(n)} y_{i-j,1}^{(n)} \right| > 2\delta\sqrt{n}} \right) \end{aligned}$$

Par suite de la stationnarité stricte du processus $\{y_{i+d\tau}\}_{\tau \geq 0}$ cette espérance tend vers 0 lorsque n tend vers ∞ car $\zeta_{i,\tau}^{(n)2}(\underline{\theta}_i)$ est uniformément intégrable. D'où

$$\lim_{m \rightarrow \infty} P(\max_{\tau} |\zeta_{i,\tau}^{(n)}(\underline{\theta}_i)| > \delta) \leq \frac{1}{\delta^2} \lim_{m \rightarrow \infty} \sum_{\tau=0}^{m-1} E \left(\zeta_{i,\tau}^{(n)2}(\underline{\theta}_i) I_{|\zeta_{i,\tau}^{(n)}(\underline{\theta}_i)| > \delta} \right) = 0$$

(3), (5), (6) and (7) sont semblables à celles données dans Swensen (1985), et ont été omises.

Comme pour la condition (4)

$$\sum_{i=1}^d \sum_{\tau=0}^{m-1} \left(\zeta_{i,\tau}^{(n)}(\underline{\theta}_i) \right)^2 = \frac{1}{4d} \sum_{i=1}^d \frac{1}{m\sigma_i^2} \sum_{\tau=0}^{m-1} \varphi_{f_1}^2 \left(\frac{Z_{i,\tau}^{(n)}(\underline{\theta}_i)}{\sigma_i} \right) \sum_{j_1=1}^{p_i} \sum_{j_2=1}^{p_i} h_{i,j_1} h_{i,j_2} y_{i-j_1+d\tau}^{(n)} y_{i-j_2+d\tau}^{(n)}$$

converge en probabilité vers

$$\frac{1}{4d} \sum_{i=1}^d \frac{1}{\sigma_i^2} I(f_1) \underline{\tau}_{i,\cdot}^{(n)'} \Gamma^i(\underline{\theta}, \underline{\sigma}) \underline{\tau}_{i,\cdot}^{(n)} = \frac{I(f_1)}{4d} \underline{\tau}^{(n)'} \Gamma(\underline{\theta}, \underline{\sigma}) \underline{\tau}^{(n)}.$$

■

5.6 Propriété LAN pour un modèle $AR(p_t)$ évolutif d -périodique

Définition 5.6.1 On appelle suite centrale et on note $\Delta_f^{(n)}(\underline{\theta})$ la quantité donnée par $\Delta_f^{(n)}(\underline{\theta}) = \underline{\delta}^{(n)}(\underline{\theta}) \in \mathbb{R}^p$

Les distributions asymptotiques de la suite centrale $\Delta_f^{(n)}(\underline{\theta})$ et de la suite du logarithme du rapport de vraisemblance $\Lambda_{(\underline{\theta} + \nu^{(n)} \underline{\tau}^{(n)})/\underline{\theta}}^{(n)}$, $(\nu^{(n)} = \frac{1}{\sqrt{n}})$ sont obtenues sous la suite d'hypothèses nulles $H_f^{(n)}(\underline{\theta})$. Grâce aux techniques de contiguïté, les distributions asymptotiques de ces variables sont aussi obtenues sous une suite d'hypothèses alternatives locales $H_f^{(n)}(\underline{\theta}^{(n)})$.

5.6.1 Normalité locale asymptotique

La proposition suivante exprime le résultat essentiel concernant la propriété LAN d'un modèle autorégressif à ordres et coefficients périodiques du temps.

Proposition 5.6.2 (Bentarzi, Guerbyenne et Merzougui 2006) Sous les conditions de régularité $CR1$, $CR2$ et $CR3$, nous avons :

(i) LAQ : (Quadratique locale asymptotique)

$$\Lambda_{(\underline{\theta} + \nu^{(n)} \underline{\tau}^{(n)})/\underline{\theta}}^{(n)} - \underline{\tau}^{(n)'} \Delta_f^{(n)}(\underline{\theta}) + \frac{1}{2} \underline{\tau}^{(n)'} \Gamma^\Delta(\underline{\theta}) \underline{\tau}^{(n)} = o_p(1), \text{ sous } H_f^{(n)}(\underline{\theta}) \quad (5.5.1)$$

où la matrice carrée Γ^Δ et la suite centrale $\Delta_f^{(n)}(\underline{\theta})$ sont respectivement données par

$$\Gamma^\Delta(\underline{\theta}) = \frac{I(f_1)}{d} \Gamma(\underline{\theta}, \underline{\sigma}),$$

et

$$\Delta_f^{(n)}(\underline{\theta}) = \left(\delta_{1,1}^{(n)}(\underline{\theta}_1), \delta_{1,2}^{(n)}(\underline{\theta}_1), \dots, \delta_{1,p_1}^{(n)}(\underline{\theta}_1), \delta_{2,1}^{(n)}(\underline{\theta}_2), \dots, \delta_{2,p_2}^{(n)}(\underline{\theta}_2), \dots, \delta_{d,1}^{(n)}(\underline{\theta}_d), \dots, \delta_{d,p_d}^{(n)}(\underline{\theta}_d) \right) \quad (5.5.2)$$

(ii) Normalité locale asymptotique de la suite centrale

$$\Delta_f^{(n)}(\underline{\theta}) \xrightarrow[n \rightarrow \infty]{\mathcal{L}} \mathcal{N}_{\mathbf{p}}(0_{\mathbb{R}^{\mathbf{p}}}, \Gamma^{\Delta}(\underline{\theta})) \text{ sous } H_f^{(n)}(\underline{\theta}). \quad (5.5.3)$$

5.6.2 Distributions asymptotiques locales de $\Delta_f^{(n)}(\underline{\theta})$ et de

$$\Lambda_{(\underline{\theta} + \nu^{(n)} \underline{\tau}^{(n)})/\underline{\theta}}^{(n)} \text{ sous } H_f^{(n)}\left(\underline{\theta}^{(n)}\right)$$

Corollaire 5.6.3 Sous les mêmes conditions et notations de la **Proposition 5.5.2**, nous avons :

$$\begin{aligned} \Lambda_{(\underline{\theta} + \nu^{(n)} \underline{\tau}^{(n)})/\underline{\theta}}^{(n)} - \underline{\tau}^{(n)'} \Delta_f^{(n)}(\underline{\theta}) + \frac{1}{2} \underline{\tau}^{(n)'} \Gamma^{\Delta}(\underline{\theta}) \underline{\tau}^{(n)} &= 0_p(1) \text{ sous } H_f^{(n)}\left(\underline{\theta}^{(n)}\right), \\ \Delta_f^{(n)}(\underline{\theta}) - \Gamma^{\Delta}(\underline{\theta}) \underline{\tau}^{(n)} &\xrightarrow[n \rightarrow \infty]{\mathcal{L}} \mathcal{N}_{\mathbf{p}}(0, \Gamma^{\Delta}(\underline{\theta})) \text{ sous } H_f^{(n)}\left(\underline{\theta}^{(n)}\right) \end{aligned} \quad (5.5.4)$$

5.7 Les Familles LAN

Définition 5.7.1 : Soit $\Theta \subset \mathbb{R}^{\mathbf{p}}$, $\mathcal{A}_n$ une σ -algèbre et $P_{\underline{\theta}}^{(n)}$ une probabilité définie sur $\mathcal{A}_n$ pour tout $\underline{\theta}, n \in \mathbb{N}^*$. (la notation $P_{\underline{\theta}}$ désignera la probabilité définie sur $\bigcup_{n=1}^{\infty} \mathcal{A}_n$, si $\mathcal{A}_n$ désigne

une suite de tribus adaptées au processus d'intérêt). On dit que la condition $LAN(\underline{\theta}, M_n, \gamma_n)$ est satisfaite, pour tout $\underline{\theta}, n \in \mathbb{N}^*$, M_n matrice de dimension $\mathbf{p} \times \mathbf{p}$ définie positive, γ_n vecteur

aléatoire de dimension $\sum_{i=1}^d p_i$, si $\mathcal{L}(\gamma_n / P_{\underline{\theta}}^{(n)}) \Rightarrow N\left(0_{\mathbb{R}^{\mathbf{p}}}, I_{\sum_{i=1}^d p_i \times \sum_{i=1}^d p_i}\right)$ ($\mathbf{p}$ désignant également

$\sum_{i=1}^d p_i$) notée $\mathcal{N}(0, I)$ et, pour toute suite $\{t_n\} \subset \mathbb{R}^{\sum_{i=1}^d p_i}$, $\underline{\theta}^{(n)} = \underline{\theta} + M_n^{-\frac{1}{2}} t_n \in \Theta$ et $g_n \in \frac{dP_{\underline{\theta}^{(n)}}^{(n)}}{dP_{\underline{\theta}}^{(n)}}$

impliquent $\frac{g_n}{\rho_{t_n}(\gamma_n)} \xrightarrow[n \rightarrow \infty]{} 1$ sous $P_{\underline{\theta}}^{(n)}$ où $\rho_t : \mathbb{R}^{\sum_{i=1}^d p_i} \rightarrow \mathbb{R}$, $(\text{voir Fabian and Hannan (1982),$

déf 2, page 461).

Théorème 5.7.2 Dans les modèles $AR_d(p_t)$ et sous les conditions $CR1$, $CR2$ et $CR3$, la

condition $LAN(\underline{\theta}_0, n^{\frac{I(f)}{d}} \Gamma(\underline{\theta}_0), C^{-1}(\underline{\theta}_0) \frac{\sqrt{d}}{\sqrt{I(f)}} \Delta_f^{(n)}(\underline{\theta}_0))$ est satisfaite où $C^{-1}(\underline{\theta}_0)$ est la racine carrée symétrique de $\Gamma(\underline{\theta}_0)$

5.8 Existence et construction d'estimateurs LAM

5.8.1 Borne inférieure d'un estimateur dans les modèles LAN

Soit $l : \mathbb{R}^p \rightarrow \mathbb{R}$ une fonction de perte qui satisfait aux conditions suivantes

- (i) $l \geq 0$,
 - (ii) $l(x) = l(-x), \forall x \in \mathbb{R}^p$,
 - (iii) $\{x/l(x) \leq u\}$ est convexe $\forall u \in \mathbb{R}_+^*$.
- (5.7.1)

Le résultat qui suit donne une borne inférieure du risque encouru en prenant $\{Z_n\}$ comme suite d'estimateurs arbitraires du paramètre $\underline{\theta}$

Théorème 5.8.1 (Hájek 1972) Soit un modèle $AR_d(p_t)$ d -périodique satisfaisant les conditions $CR1$, $CR2$ et $CR3$ et soit $\{Z_n\}$ une suite d'estimateurs quelconques de $\underline{\theta}$, l une fonction convexe définie sur $\mathbb{R}^p$ et $\{Q_n\}$ une suite de matrices appartenant à l'ensemble $\mathcal{M}_{\sum_{i=1}^d p_i}$ des

matrices orthogonales de dimensions $\sum_{i=1}^d p_i \times \sum_{i=1}^d p_i$. Alors :

$$\lim_{K \rightarrow \infty} \liminf_{n \rightarrow \infty} \sup_{\left\| \frac{\sqrt{nI(f)}}{\sqrt{d}} C(\underline{\theta}_0)(\underline{\theta} - \underline{\theta}_0) \right\| \leq K} E_{n, \underline{\theta} l} \left\{ Q_n \frac{\sqrt{n} \sqrt{I(f)}}{\sqrt{d}} C(\underline{\theta}_0)(Z_n - \underline{\theta}) \right\} \geq \int l(x) dN(0, 1)$$
(5.7.2)

Définition 5.8.2 Soit la condition $LAN(\underline{\theta}_0, M_n, \gamma_n)$ vérifiée. Une suite d'estimateurs $\{Z_n\}$ est dite $LAM(\underline{\theta}_0)$ si

$$\lim_{K \rightarrow \infty} \liminf_{n \rightarrow \infty} \sup_{\left\| \frac{\sqrt{nI(f)}}{\sqrt{d}} C(\underline{\theta}_0)(\underline{\theta} - \underline{\theta}_0) \right\| \leq K} E_{n, \underline{\theta} l} \left\{ Q_n \frac{\sqrt{n} \sqrt{I(f)}}{\sqrt{d}} C(\underline{\theta}_0)(Z_n - \underline{\theta}) \right\} = \int l(x) dN(0, 1)$$

Définition 5.8.3 La suite d'estimateurs $\{Z_n\}$ est dite $\underline{\theta}_0$ -régulière, sous la condition $LAN(\underline{\theta}_0, M_n, \gamma_n)$

si $M_n^{-\frac{1}{2}} (Z_n - \underline{\theta}_0) - \gamma_n = o(P_{\underline{\theta}_0})$

Remarque 5.8.4 Dans les modèles $AR_d(p_t)$ d -périodiques, et sous les hypothèses $CR1$, $CR2$ et $CR3$, une suite $\{Z_n\}$ est dite $\underline{\theta}_0$ -régulière si

Lemme 5.8.5 Sous les hypothèses $CR1$, $CR2$ et $CR3$, on a pour toute suite d'estimateurs $\{Z_n\}$, l'implication suivante :

$$\sqrt{n}(Z_n - \underline{\theta}_0) - \frac{d \Gamma^{-1}(\underline{\theta}_0, \underline{\sigma})}{I(f)} \Delta_f^{(n)}(\underline{\theta}_0) = o(P_{\underline{\theta}_0}), \quad (5.7.3)$$

Ce lemme implique que $\{Z_n\}$ est LAM

$\underline{\theta}_0$ restant inconnu, ces résultats fondamentaux peuvent sembler inutiles ! Ceci motive le développement suivant (Le Cam (1960), Kreiss (1987), Hallin (1995), Bentarzi (1995)) :

5.8.2 Le problème de l'alignement

Puisque la valeur de $\underline{\theta}$ sous l'hypothèse $H_f^{(n)}(\underline{\theta})$ est non spécifiée, on peut songer à l'estimer car afin de construire des estimateurs réguliers, l'existence d'estimateurs $\sqrt{n}$ -consistents est essentielle.

On suppose que l'on dispose d'un estimateur initial $\widehat{\underline{\theta}}_n$ $\sqrt{n}$ -consistant ; d'où l'hypothèse $CR4 : \sqrt{n}(\widehat{\underline{\theta}}_n - \underline{\theta}_0) = O(P_{\underline{\theta}_0})$ où encore $\exists c > 0$, tel que $\sqrt{n} \left| \widehat{\underline{\theta}}_n - \underline{\theta}_0 \right| \leq c, \forall n \in \mathbb{N}, \underline{\theta}_0 \in \Theta$

Remarque 5.8.6 : La condition $CR4$ est vérifiée pour n'importe quel estimateur pour lequel le théorème central limite est valide, *i.e.*, pour tous les estimateurs standards.

5.8.3 Propriété de linéarité asymptotique de $\Delta_f^{(n)}(\underline{\theta})$ sous $H_f^{(n)}(\underline{\theta}^{(n)})$

Une conséquence essentielle de la propriété de la normalité asymptotique locale (LAN) établie dans la proposition 5 §5.5.1 est que, pour une densité f_{σ_i} d'un bruit blanc, l'expérience

locale $\varepsilon_{f_\sigma}^{(n)}(\theta) = \left\{ P_{\underline{\theta} + \nu^{(n)} \underline{\tau}^{(n)}, \underline{\tau}^{(n)'}} \in \mathbb{R}^p / \sup_n \left\| \underline{\tau}^{(n)} \right\|^2 < \infty, \underline{\tau}^{(n)} \xrightarrow{n \rightarrow \infty} \underline{\tau} \right\}$ converge faiblement

vers l'expérience shift gaussien $\sum_{i=1}^d p_i$ -dimensionnelle. (Le Cam (1960), (1986)) :

$$\ell_f(\underline{\theta}) = \left\{ \mathcal{N}(\Gamma^\Delta(\underline{\theta}) \underline{\tau}, \Gamma^\Delta(\underline{\theta})), \underline{\tau} \in \mathbb{R}^{\mathbf{p}} \right\},$$

Brièvement, la famille des fonctions de vraisemblance, indexées par le paramètre $\underline{\theta}$ et correspondant à la densité f_σ , notée $L_{\underline{\theta}, f_\sigma}$ restreinte à $\underline{\theta} \in \Theta \subset \mathbb{R}^{\mathbf{p}}$ est asymptotiquement gaussienne au voisinage de $\underline{\theta}$. L'étude des divers problèmes statistiques liés à l'expérience locale $\varepsilon_{f_\sigma}^{(n)}(\theta)$, peut être ainsi, ramenée à celle des problèmes statistiques liés à l'expérience shift gaussien $\ell_f(\underline{\theta})$ bien connue dans la littérature statistique. Les propriétés (exactes) des procédures statistiques établies concernant le problème en question peuvent, alors hériter localement asymptotiquement du problème originel en question. On exploite ce résultat fondamental dû à Le Cam (cf; Le Cam (1986)), Le Cam and Yang (1990)).

Proposition 5.8.6 (Propriété de linéarité)

La suite centrale $\Delta_f^{(n)}(\underline{\theta})$ possède sous $H_f^{(n)}(\underline{\theta})$ et donc sous $H_f^{(n)}(\underline{\theta}^{(n)})$ par contiguïté la propriété suivante, dite propriété de linéarité (asymptotique) :

linéarité sous $H_f^{(n)}(\underline{\theta})$ et donc sous $H_f^{(n)}(\underline{\theta}^{(n)})$

$$\Delta_f^{(n)}(\underline{\theta} + \nu^{(n)} \underline{\tau}^{(n)}) - \Delta_f^{(n)}(\underline{\theta}) = -\Gamma^\Delta(\underline{\theta}) \underline{\tau} + o_{P_{\underline{\theta}}}(1)$$

Théorème 5.8.7 (Existence des estimateurs LAM). Supposons que $\{\bar{\theta}_n\} \in \Theta$ est une suite discrète d'estimateurs, $\sqrt{n}$ -consistents pour $\theta_0 \in \Theta$. Alors $\hat{\underline{\theta}}_n$ définie par

$$\hat{\underline{\theta}}_n = \bar{\theta}_n + \frac{1}{\sqrt{n}} \frac{d \hat{\Gamma}_n(\bar{\theta}_n)^{-1}}{I(f)} \Delta_f^{(n)}(\bar{\theta}_n), \quad (5.7.5)$$

est régulière avec

$$\hat{\Gamma}_n(\bar{\theta}_n) = \begin{pmatrix} \frac{\hat{G}^1(\bar{\theta}_n)}{\hat{\sigma}_1^2} & \dots & \dots & 0 \\ 0 & \frac{\hat{G}^2(\bar{\theta}_n)}{\hat{\sigma}_2^2} & \dots & 0 \\ \vdots & 0 & \ddots & 0 \\ 0 & 0 & 0 & \frac{\hat{G}^d(\bar{\theta}_n)}{\hat{\sigma}_d^2} \end{pmatrix},$$

$$\text{où } \widehat{G}^i(\bar{\theta}_n) = \frac{1}{m} \sum_{\tau=0}^{m-1} \underline{y}_{i,\tau}^{(n)} \underline{y}_{i,\tau}^{(n)'} \text{ et } \widehat{\sigma}_i^2 = \frac{1}{m} \sum_{\tau=0}^{m-1} \left(Z_{i,\tau}^{(n)}(\bar{\theta}_i) \right)^2$$

Pour des raisons techniques, nous nous sommes restreints à des suites discrètes d'estimateurs $\{\bar{\theta}_n\}$ telles que, en supposant que $\widehat{\theta}_n$ vérifie *CR4*, $\bar{\theta}_n$ est donnée par un des points de l'ensemble $\{\underline{\theta} : \underline{\theta} = n^{-\frac{1}{2}}(i_1, \dots, i_{\sum_{i=1}^d p_i}), i_j \in \mathbb{Z}\}$ le plus proche de $\widehat{\theta}_n$. Bien sûr, $\{\bar{\theta}_n\}$ satisfait la propriété plus générale de discrétisation.

Définition 5.8.8 Une suite $\{\bar{\theta}_n\}$ d'estimateurs est dite discrète s'il existe $K \in \mathbb{N}^*$ tel que, indépendamment de $n \in \mathbb{N}^*$, $\bar{\theta}_n$ prend au plus K valeurs différentes dans l'ensemble :

$$Q_n = \{\underline{\theta} \in \mathbb{R}^p : \sqrt{n} |\underline{\theta} - \underline{\theta}_0| \leq c\}, c > 0 \text{ fixé.}$$

Le grand avantage des estimateurs discrets, est le résultat suivant dû à Le Cam (1960), et utilisé entre autres, par Bickel (1982) et Kreiss (1987).

Lemme 5.8.9 Soit $\{S_n(\underline{\theta}), n \in N\}$ une suite de variables aléatoires qui dépendent de $\underline{\theta}$. Si pour toute suite d'estimateurs $\{\underline{\theta}_n\} \subset \Theta$ satisfaisant

$$\sqrt{n}(\underline{\theta}_n - \underline{\theta}_0) \text{ est borné par une constante } c > 0, \text{ on a } S_n(\underline{\theta}_n) = 0_{P_{\underline{\theta}_0}^{(n)}}(1),$$

alors on a aussi $S_n(\bar{\theta}_n) = 0_{P_{\underline{\theta}_0}^{(n)}}(1)$ pour toute suite d'estimateurs localement discrète $\bar{\theta}_n$

Preuve

$$\begin{aligned} \forall \varepsilon > 0, \forall M > 0, P_{\underline{\theta}_0}^{(n)}(\|S_n(\bar{\theta}_n)\| \geq \varepsilon) &= P_{\underline{\theta}_0}^{(n)}(\|S_n(\bar{\theta}_n)\| \geq \varepsilon \text{ et } \sqrt{n}(\bar{\theta}_n - \underline{\theta}_0) \leq \\ M) &+ P_{\underline{\theta}_0}^{(n)}(\|S_n(\bar{\theta}_n)\| \geq \varepsilon \text{ et } \sqrt{n}(\bar{\theta}_n - \underline{\theta}_0) > M) \leq P_{\underline{\theta}_0}^{(n)}(\|S_n(\bar{\theta}_n)\| \geq \varepsilon \text{ et } \sqrt{n}(\bar{\theta}_n - \underline{\theta}_0) \leq \\ M) &+ P_{\underline{\theta}_0}^{(n)}(\sqrt{n}(\bar{\theta}_n - \underline{\theta}_0) > M) \end{aligned}$$

Pour $M = M(\varepsilon), \exists N_1(\varepsilon) \in \mathbb{N}, \forall n \geq N_1(\varepsilon), P_{\underline{\theta}_0}^{(n)}(\sqrt{n}(\bar{\theta}_n - \underline{\theta}_0) > M) < \frac{\varepsilon}{2}$, puisque $\bar{\theta}_n$ est $\sqrt{n}$ -convergent.

L'événement $[\|S_n(\bar{\theta}_n)\| \geq \varepsilon \text{ et } \sqrt{n}(\bar{\theta}_n - \underline{\theta}_0) \leq M]$ est constitué d'un nombre fini de points borné par K . Toute suite de valeurs de S_n en ces K points est une suite de type $S_n(\underline{\theta}_n)$ satisfaisant aux hypothèses du lemme. Soit alors $\underline{\theta}_n$ la suite construite de façon que pour tout n $\|S_n(\underline{\theta}_n)\| = \max \|S_n(\bar{\theta}_n)\|$ (ce maximum est calculé sur un nombre fini de points).

Cette suite satisfait encore aux hypothèses du lemme et donc $S_n(\bar{\theta}_n) = 0_{P_{\bar{\theta}_0}^{(n)}}(1)$. Il existe donc un $N_2(\varepsilon) \in \mathbb{N}$, tel que $\forall n \geq N_2(\varepsilon)$, $P_{\bar{\theta}_0}^{(n)}(\|S_n(\bar{\theta}_n)\| \geq \varepsilon \text{ et } \sqrt{n}(\bar{\theta}_n - \bar{\theta}_0) \leq M) < \frac{\varepsilon}{2}$. Donc $\forall \varepsilon > 0, \exists M(\varepsilon)$ et $N(\varepsilon) = \max(N_1(\varepsilon), N_2(\varepsilon))$ tels que $P_{\bar{\theta}_0}^{(n)}(\|S_n(\bar{\theta}_n)\| \geq \varepsilon \text{ et } \sqrt{n}(\bar{\theta}_n - \bar{\theta}_0) \leq M) + P_{\bar{\theta}_0}^{(n)}(\sqrt{n}(\bar{\theta}_n - \bar{\theta}_0) > M) < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon$.

5.9 Construction d'estimateurs adaptatifs

Les estimateurs obtenus ci-dessus dépendent de la loi sous-jacente du bruit blanc. Dans cette section, et en suivant étroitement la méthode utilisée par Kreiss (1987), on apporte une réponse à la question s'il est possible de construire des estimateurs qui soient indépendants de la loi sous-jacente du bruit blanc mais qui sont *LAM*, simultanément pour plusieurs types de lois de probabilité des erreurs. De tels estimateurs, s'ils existent, sont dits **adaptatifs** pour la classe spécifiée de densités

Pour donner une réponse à la question ci-dessus, on recherche des estimateurs pour la fonction score et pour la quantité d'information de Fisher qui sont inconnues. Ensuite, on définit l'estimateur adaptatif semblable à $\hat{\theta}_n$. D'abord, on restreint notre attention à des estimateurs de φ .

On introduit les notations additionnelles :

$$(i) \quad g(x) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{x^2}{2}\right), x \in \mathbb{R},$$

$$(ii) \quad f(x) = \int_{-\infty}^{+\infty} g(x-y)f_1(y)dy.$$

Notons $z_k, k = 1, \dots, n$ les résidus centrés réduits qui sous $H_f^{(n)}(\underline{\theta})$ coïncident avec le bruit blanc réduit.

$$(iii) \quad \hat{f}_j(x, \theta) = \frac{1}{2n} \sum_{\substack{k=1 \\ k \neq j}}^n \{g(x+z_k, \eta) + g(x-z_k, \eta)\}; j = 1, \dots, n$$

Soit $\hat{q}_j(x, \theta)$, l'estimateur suivant de φ

$$\hat{q}_j(x, \theta) = \begin{cases} -\frac{1}{2} \frac{\hat{f}_j(x, \theta)}{\hat{f}_j(x, \theta)} & \text{si } \begin{cases} \hat{f}_j(x, \theta) \geq d_n, |x| \leq g_n \\ \hat{f}_j(x, \theta) \leq c_n \hat{f}_j(x, \theta) \end{cases} \\ 0 & \text{sinon} \end{cases}$$

avec $c_n \rightarrow \infty, g_n \rightarrow \infty, \eta(n) \rightarrow 0, d_n \rightarrow 0$.

$$\varphi_{\sigma_t}(x) = \frac{1}{\sigma_t} \varphi_{f_1}(x) \text{ et } \widehat{\widehat{q}}_j(x, \theta) = \frac{1}{\widehat{\sigma}_j} \widehat{q}_j(x, \theta)$$

Soit

$$\widetilde{\delta}_{i,\cdot}(\underline{\theta}_i) = \frac{1}{\sqrt{n}} \sum_{\tau=0}^{m-1} \widehat{\widehat{q}}_j(x, \underline{\theta}_i) \underline{y}_{i,\tau},$$

et

$$\widetilde{\Delta}_m(\underline{\theta}) = (\widetilde{\delta}_{1,\cdot}(\underline{\theta}_1), \widetilde{\delta}_{2,\cdot}(\underline{\theta}_2), \dots, \widetilde{\delta}_{d,\cdot}(\underline{\theta}_d))',$$

la version estimée de $\Delta_f^{(n)}(\underline{\theta})$.

Lemme 5.9.1 (Estimation de la quantité d'information de Fisher).

L'estimateur de la quantité d'information de Fisher donné par $\widehat{I}_n = \frac{4}{n} \sum_{j=0}^n \widehat{\widehat{q}}_{n,j}^2(e_j(\bar{\theta}_n); \bar{\theta}_n)$ est

consistant où $\widehat{\widehat{q}}_{n,j}(e_j(\bar{\theta}_n); \bar{\theta}_n)$ sont des estimateurs de $\varphi(e_j(\theta), j = 1, \dots, n$ et $\frac{4}{n} \sum_{j=0}^n \varphi^2(e_j(\theta_n)) \xrightarrow{P_{\theta_n}}$

$I(f_1)$, lorsque $n \rightarrow \infty$.

Théorème 5.9.2 (Bentarzi, Guerbyenne et Merzougui (2006)) (Existence des estimateurs adaptatifs)

L'estimateur $\widetilde{\underline{\theta}}_n = \bar{\theta}_n + \frac{1}{\sqrt{n}} \frac{d \widehat{\Gamma}_n(\bar{\theta}_n)^{-1}}{\widehat{I}} \widetilde{\Delta}_m(\bar{\theta}_n)$ où $\widehat{I}$ est un estimateur de I et est défini à partir de $\widehat{f}_1$ qui est un estimateur de f_1 lorsque la méthode du noyau est utilisée. Il satisfait la condition de régularité :

$$\sqrt{n} \left(\widetilde{\underline{\theta}}_n - \underline{\theta}_0 \right) - \frac{d \Gamma^{-1}(\underline{\theta}_0, \underline{\sigma})}{I(f)} \Delta_f^{(n)}(\underline{\theta}_0) = o(P_{\underline{\theta}_0}^{(n)}),$$

i.e., $\widetilde{\underline{\theta}}_n$ est LAM pour toute densité f satisfaisant aux hypothèses

(i) f est symétrique autour de l'origine

(ii) $\int x^4 f(x) dx < \infty$

et on a : $L(\sqrt{n} (\widetilde{\underline{\theta}}_n - \underline{\theta}_0) / P_{n, \underline{\theta}_0}) \Rightarrow N \left(0_{\mathbb{R}^p}, \frac{d \Gamma^{-1}(\underline{\theta}_0)}{I(f)} \right)$ pour toute densité satisfaisant (i)

et (ii)

La démonstration est identique à celle de Kreiss (1987).

5.10 Résultats numériques

Des simulations intensives ont été exécutées pour étudier le comportement de l'estimateur adaptatif AE obtenu ci-dessus. Nous avons simulé un processus $PAR_2(1)$

$$\begin{cases} y_{1+2\tau} = \theta_{1,1}y_{2\tau} + \varepsilon_{1+2\tau}, \\ y_{2+2\tau} = \theta_{2,1}y_{1+2\tau} + \varepsilon_{2+2\tau} \end{cases}$$

pour les quatre densités des erreurs telles que données dans Kreiss (1987) mais f_3 et f_4 ont été modifiées afin que la densité du bruit blanc soit de variance 1 ($CR2$).

$$\begin{aligned} f_1(x) &= \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{x^2}{2}\right), \\ f_2(x) &= \frac{0.05}{\sqrt{50\pi}} \exp\left(-\frac{x^2}{50}\right) + \frac{0.95}{\sqrt{2\pi}} \exp\left(-\frac{x^2}{2}\right), \\ f_3(x) &= \frac{0.5\sqrt{10}}{\sqrt{2\pi}} \exp\left(-\frac{(\sqrt{10}x - 3)^2}{2}\right) + \frac{0.5\sqrt{10}}{\sqrt{2\pi}} \exp\left(-\frac{(\sqrt{10}x - 3)^2}{2}\right), \\ f_4(x) &= \frac{1}{\sqrt{2}} \exp(-\sqrt{2}|x|). \end{aligned}$$

f_1 et f_4 sont fréquemment rencontrées dans la pratique, f_2 et f_3 sont des mélanges de lois normales alors que f_2 est communément utilisée en présence de valeurs aberrantes ou simplement extrêmes; f_3 a été choisie par Kreiss pour montrer l'efficacité de la procédure adaptative.

On utilise toujours l'estimateur des moindres carrés (MCO) comme estimateur initial $\bar{\theta}_n$ et on compare son comportement avec celui de l'estimateur adaptatif (EA) : $\tilde{\theta}_n$ relativement à un premier critère qui est la longueur de l'intervalle de confiance (CI).

intervalle de confiance à 90% pour $\sqrt{n}(\bar{\theta}_n - \theta)$ et $\sqrt{n}(\tilde{\theta}_n - \theta)$,
pour $\theta = (0.8, 0.9)$, $\sigma^2 = (1, 0.8)$ et $n = 100$

densité	$\theta_{1,1}$	$\theta_{2,1}$
$f_1 MCO$	[-1.4882, 1.0802]	[-1.3345, 1.0306]
EA	[-1.5639, 1.1174]	[-1.3524, 1.2479]
$f_2 MCO$	[-1.6715, 1.0937]	[-1.9485, 1.0313]
EA	[-1.5822, 0.8721]	[-1.3286, 0.8621]
$f_3 MCO$	[-1.8184, 1.0557]	[-1.5748, 0.8213]
EA	[-0.8406, 0.7249]	[-0.5357, 0.8823]
$f_4 MCO$	[-2.1268, 0.7822]	[-1.4040, 0.8662]
EA	[-1.9113, 0.7856]	[-1.0482, 0.7694]

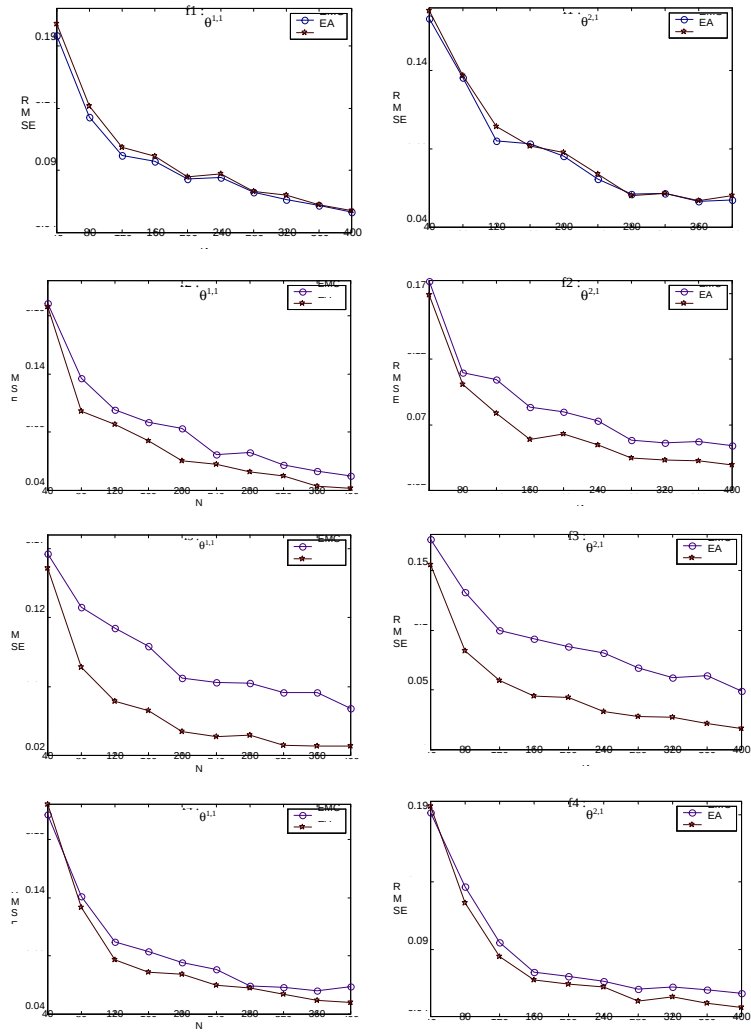
Le second critère est le *RMSE* (Root Mean Square Error ou racine carrée de l'erreur quadratique moyenne) définie par : $RMSE = \sqrt{(Biais)^2 + Variance}$.

Les valeurs du *RMSE* sont calculées pour diverses valeurs de la taille n en commençant à partir de 40 jusqu'à 400, pour $\theta = (0.5, 0.7)'$ et $\sigma^2 = (1, 0.8)'$, Les résultats sont représentés par les graphes suivants.

Les simulations (CI et graphes) montrent que lorsque la densité du bruit blanc normalisé est f_2 et f_3 l'estimateur adaptatif est toujours meilleur principalement pour f_3 où la différence est très grande en faveur de *EA*, bien que l'estimateur MCO soit utilisé comme estimateur initial.

Dans les situations standard f_1, f_4 les deux estimateurs ont presque la même performance principalement pour f_1 les estimateurs *EA* et *MCO* sont presque les mêmes contrairement à f_4 où l'estimateur adaptatif *EA* semble être légèrement meilleur.

Cependant, l'estimateur adaptatif est relativement coûteux.



Bibliographie

- [1] Adams, G. J. and Goodwin, G. C. (1995). Parameter estimation for periodic *ARMA* Models. *Journal of Time Series Analysis*, **16**, 127-145.
- [2] Aknouche, A., (2006). Etude théorique et algorithmique de la modélisation des séries chronologiques périodiques. Thèse de Doctorat. Faculté de Mathématiques, U.S.T.H.B. Alger.
- [3] Aknouche, A (2007) Causality conditions and autocovariance calculation in *PVAR* models. *Journal of Statistical Computation and Simulation*. A paraître.
- [4] Aknouche A. and Guerbyenne, H. (2005). On some probabilistic properties of double periodic *AR* models. Submitted to *Statistics and Probability Letters*.
- [5] Aknouche, A. and Guerbyenne, H. (2006a). Algorithme RLS en deux étapes pour l'estimation d'un modèle *ARCH*. Comptes rendus de l'académie des sciences- Mathématiques. I **343** 535–540
- [6] Aknouche, A. and Guerbyenne, H. (2006b). Recursive estimation of *GARCH* models. *Communications in Statistics- Simulation and Computation*, **35**, 925-938.
- [7] Anderson, PL and Vecchia, A. V. (1993). Asymptotic results for periodic autoregressive moving-average processes. *Journal of Time Series Analysis*. **14**, 1-18.
- [8] Baillie, R. T., Bollerslev, T. and Mikkelsen, H. O. (1996). Fractionally integrated generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics* **74**, 3-30.
- [9] Bentarzi, M. (1995). "Modèles de séries chronologiques à coefficients périodiques". Thèse de Doctorat es Sciences. Institut de Mathématiques, U.S.T.H.B, Alger, Algérie.
- [10] Bentarzi, M., (1998). Model-building problem of periodically m-variate moving average processes. *Journal of Multivariate Analysis*, **66**, 1-21.

- [11] Bentarzi, M. and Aknouche A. (2005). Calculation of the Fisher information matrix for periodic *ARMA* models. *Communications in Statistics-Theory and Methods*, **34**, 891-903.
- [12] Bentarzi, M. and Aknouche A. (2006). An on-line estimation algorithm for periodic autoregressive models. *Communications in Statistics-Theory and Methods*, **35**, 1495-1512.
- [13] Bentarzi, M, Guerbyenne, H and Merzougui, M (2006). Adaptive estimation in periodic *AR*-models (to be submitted)
- [14] Bentarzi, M. and Hallin, M. (1994). On the Invertibility of Periodic Moving Average Models. *Journal of Time Series Analysis*, **15**, 263-268.
- [15] Bentarzi, M. and Hallin, M. (1996). Locally optimal tests against periodic autoregression. *Econometric Theory*, **12**, 88-112.
- [16] Bentarzi, M. and Hallin, M. (1998). Spectral factorization of periodically correlated *MA*(1) processes. *Journal of Applied Probability*, **35**, 46-54.
- [17] Benveniste, A., Métivier, M. and Priouret, P. (1990). *Adaptive algorithms and stochastic approximation*. Springer-Verlag, Berlin.
- [18] Bera, A. K., Higgins, M. L. and Lee, S. (1992). Interaction between autocorrelation and conditional conditional heteroskedasticity : a random coefficient approach. *Journal of Business and Economic Statistics*, **10**, 133-142.
- [19] Bera, A. K., Higgins, M. L. and Lee, S. (1996). Random coefficients formulation of conditional heteroskedasticity and augmented *ARCH* models. *Sankhya*, **58**, 199-220.
- [20] Berkes, I., and Horvath, L. (2004). The efficiency of the estimators of the parameters in *GARCH* processes. *Annals of statistics*, **32**, 633-655.
- [21] Berkes, I., Horvath, L. and Kokoskza, P. (2003). *GARCH* processes : Structure and estimation. *Bernoulli*, **9**, 201-227.
- [22] Bhuiya, R. K. (1971). Stochastic analysis of periodic hydrologic processes. *Journal of Hydraulic Division of the American Society of Civil Engineering*, **97**, 949-962.
- [23] Bibi, A. et Aknouche, A (2006). On the probabilistic properties of periodic *GARCH* processes. Preprint.

- [24] Bibi, A. and Francq, C. (2003). Consistent and asymptotically normal estimators for time-dependent linear models. *Annals of the Institute of Statistical Mathematics*, **55**, 41-68.
- [25] Bickel, P.J, (1982). On adaptive Estimation. *Ann.Statist.***10** 647-671.
- [26] Billingsley, P. (1968). *Convergence of probability measure*. Wiley, New York.
- [27] Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics*, **31**, 307-327.
- [28] Bollerslev, T. R., Engle and Wooldridge, J. (1988). A Capital Asset Pricing Model with Time-Varying Covariances, *Journal of Political Economy* **96**, 116-131.
- [29] Bollerslev, T. and Ghysels, E. (1996). Periodic Autoregressive Conditional Heteroskedasticity. *Journal of Business and Economic Statistics*, **14**, 139-152.
- [30] Bordignon, S. and Lisi, F. (2000). Nonlinear analysis and prediction of river flow time series. *Environmetrics*, **11**, 463-477.
- [31] Bose, A. and Mukherjee, K. (2003). Estimating the *ARCH* parameters by solving linear equations. *Journal of Time Series Analysis*, **24**, 127-136.
- [32] Bougerol, P. and Picard, N. (1992a). Stationarity of *GARCH* processes and some nonnegative time series. *Journal of Econometrics*, **52**, 115-127.
- [33] Bougerol, P. and Picard, N. (1992b). Strict stationarity of generalised autoregressive processes. *Annals of Probability*, **20**, 1714-1730.
- [34] Box, G. E. P. and Jenkins, G. M. (1976). *Time Series Analysis, Forecasting and Control*, 2nd edi., Holden-Day San Francisco, CA.
- [35] Brockwell, P. J. and Davis, R. A. (1991). *Time Series : Theory and methods*. Springer-Verlag. New York.
- [36] Chen, M., An, H. Z. (1998). A note on the stationarity and the existence of moments of the *GARCH* model. *Statistica Sinica*, **8**, 505-510.
- [37] Čipra, T. (1985). Periodic moving average processes. *Aplikace Matematiky*, **30**, 218-229.
- [38] Cramér, H. (1961). On some classes of nonstationary stochastic processes. *Proceeding of the 4th Berkley Symposium on Mathematical Statistics and Probability*, **2**, 57-78.

- [39] Dahlhaus, R. and Subba Rao, S. (2006). Statistical inference for time-varying *ARCH* processes. To appear in *The Annals of statistics*, Vol. **34**.
- [40] Ding, Z., Granger, C.W.J. and Engle, R. (1993) A long memory property of stock market returns and a new model. *J. Empirical Finance* **1**, 83-106.
- [41] Drost, F. C. and Werker, B. J. M. (1996). Closing the *GARCH* gap : Continuous time *GARCH* modeling. *Journal of Econometrics*, **74**, 31-57.
- [42] Duflo, M. (1990). *Méthodes récursives aléatoires*. Techniques stochastiques. Massons, Paris.
- [43] Durbin, J. (1959). Efficient estimation of parameters in moving average models. *Biometrika*, **46**, 306-316.
- [44] Engle, R. F. (1982). Autoregressive Conditional Heteroskedasticity with estimates of variance of U.K. Inflation. *Econometrica*, **50**, 987-1008.
- [45] Engle, R. F., C. W. J. Granger and D. Kraft. (1984). Combining competing forecasts of inflation using a bivariate *ARCH* model, *Journal of Economic Dynamics and Control*, **8**, 151-165.
- [46] Engle, R. and Kroner, K. (1995). Multivariate Simultaneous Generalised *ARCH*. *Econometric Theory*, **11**, 122-150
- [47] Engle, R., Lilien, D. and Robins, R. (1987). Estimating Time Varying Risk Premia in the Term Structure. The *ARCH* Model, *Econometrica*, **55**, 391-407.
- [48] Engle, C. and Rodrigues, A. P. (1989). Tests of International CAPM with Time-Varying Covariances. *Journal of Applied Econometrics*, **4**, 119-138.
- [49] Fabian, V. and Hannan, J. (1982). On estimation and adaptive estimation for locally asymptotically normal families. *Z. Wahrsh. verw. gebiete* **59** 459-478.
- [50] Fasso, A. (2000). Residual Autocorrelation Distribution in the Validation Data Set. *Journal of Time Series Analysis*, Vol. **21**, 143-153.
- [51] Francq, C. and Zakoïan, J. M. (2004). Maximum likelihood estimation of pure *GARCH* and *ARMA-GARCH* processes. *Bernoulli*, **10**, 605-637.
- [52] Franses, P. H. (1996). *Periodicity and Stochastic Trends in Economic Times Series*. Oxford University Press.

- [53] Franses, P. and Paap, R., (2000). Modeling Day-of-the-Week Seasonality in the S&P 500 Index. *Applied Financial Economics*, **10**, 483-488.
- [54] Gallant RA. (1987). *Nonlinear statistical models*, New York, John Wiley and Sons.
- [55] Gladyshev, E. G. (1961). Periodically correlated random sequences. Soviet. Math., **2**, 385-88.
- [56] Goodwin, G. C. and Sin, K. S. (1984). Adaptive Filtering : Prediction and Control. Englewood Cliffs, NJ : Prentice Hall.
- [57] Gouriéroux, C. and Monfort (1989). A General Framework for Testing Null Hypothesis in a Mixed Form, *Econometric Theory*, **5**, 63-82.
- [58] Gouriéroux, C., Monfort and Trognon (1984a)..Pseudo Maximum Likelihood Methods : Theory. *Econometrica*, **52**, 680-700.
- [59] Gouriéroux, C., Monfort and Trognon (1984a). Pseudo Maximum Likelihood Methods : Application to Poisson Models. *Econometrica*, **52**, 701-721.
- [60] Granger, C. W. J. and Andersen, A. P. (1978). Introduction to Bilinear Time Series Models. Vandenhoeck & Ruprect, Gottingen.
- [61] Hall, P. and Heyde, C. C. (1980). *Martingale Limit Theory and its Applications*. Academic Press, New York.
- [62] Hallin, M. (1984). Spectral factorization of nonstationary moving average processes. *The Annals of Statistics*, **12**, 172-192.
- [63] Hallin, M. (1986). Nonstationary q-dependent processes and time-varying moving average models : invertibility properties and the forecasting problem. *Adv. Appl. Probab.*, **18**, 170-210.
- [64] Hannan, E. J. (1976). The convergence of some recursions. *The Annals of Statistics*, **4**, 1258-1270.
- [65] Hannan, E. J. (1980). Recursive estimation based on *ARMA* models. *The Annals of Statistics*, **8**, 762-777.
- [66] Kreiss, J.P. (1987) On Adaptive Estimation in Stationary Arma Processes. *The Annals of Statistics*. Vol 15, N°1, 112-133.

- [67] Kushner, H. J. and Clark, D. S. (1978). *Stochastic approximation methods for constrained and unconstrained systems*, *Applied Mathematics Science*. Springer Verlag, Berlin.
- [68] Kushner, H. J. and Yin, G. G. (1997). *Stochastic approximation algorithms and applications*. Springer, New York.
- [69] Le Cam, L. (1960) Locally Asymptotically Normal Families of Distributions. Univ California Publ. Statist. **3** 37-98.
- [70] Li, W. K., Ling, S. and McAleer, M. (2002). Recent theoretical results for time series models with *GARCH* errors. *Journal of Economic Surveys*, **16**, 245-269.
- [71] Li, W. K., Ling, S. and Wong, H. (2001). Estimation for partially nonstationary multivariate autoregressive models with conditional heteroscedasticity. *Biometrika*, **8**, 1135-1152.
- [72] Ling, S. (1999a). On probability properties of a double threshold *ARMA* conditional heteroskedasticity model. *Journal of Applied Probability*, **36**, 688-705.
- [73] Ling, S. (1999b). On the stationarity and the existence of moments of conditional heteroskedastic *ARMA* models. *Statistica Sinica*, **9**, 1119-1130.
- [74] Ling, S. and Deng, W. C. (1993). Parametric estimate of multivariate autoregressive models with conditional heterocovariance matrix errors. *Acta Math. Appl. Sinica*, **16**, 517-533.
- [75] Ling, S., McAleer, M. (2002). Necessary and sufficient moment conditions for the *GARCH*(p, q) and asymmetric power *GARCH*(p, q) models. *Econometric Theory*, **18**, 722-729.
- [76] Ling, S. and McAleer, M. (2003). Asymptotic theory for a new vector *ARMA-GARCH* model. *Econometric Theory*, **19**, 280-310.
- [77] Linton, O. (1993). Adaptive Estimation in *ARCH* Models. *Econometric Theory* **9**, 539– 569.
- [78] Ljung, L. (1977a). On positive real transfer functions and the convergence of some recursive schemes. *IEEE Trans. Autom. Control*. **22** 539-551.
- [79] Ljung, L. (1977b). Analysis of recursive stochastic algorithms. *IEEE Trans. Autom. Control*. **22**, 551-575.

- [80] Ljung, L. (1999). *System identification : theory for the user*. 2nd ed., PTR Prentice Hall, Upper Saddle River, N.J.
- [81] Ljung, L. and Söderström, T. (1983). *Theory and practice of recursive identification*. MIT Press, Cambridge, Massachusetts.
- [82] Lukacs, E. (1975). *Stochastic convergence*. 2nd ed. Academic.
- [83] Lumsdaine, R. L. (1996). Consistency and asymptotic normality of the quasi-maximum likelihood estimator in $IGARCH(1, 1)$ and covariance stationary $GARCH(1, 1)$ models. *Econometrica*, **64**, 575-596.
- [84] McLeod, A. I. and Li, W. K. (1983). Diagnostic checking ARMA time series models using squared-residuals autocorrelations. *Journal of Time Series Analysis*, **4**, 269-273.
- [85] Milhoj, A. (1985). The moment structure of $ARCH$ processes. *Scandinavian Journal of Statistics*, **12**, 281-292.
- [86] Moore, J. B. and Ledwich, G. (1980). Multivariable adaptive parameter and state estimators with convergence analysis. *Journal of the Australian Mathematical Society*, **21**, 176-197.
- [87] Nelson, D. B. (1990). $ARCH$ Models as Diffusion Approximations. *Journal of Econometrics*, **45**, 7-38.
- [88] Nelson, D. B., and Foster, D. P. (1994). Asymptotic filtering theory for univariate $ARCH$ models. *Econometrica*, **62**, 1-41.
- [89] Nicholls, D. F. and Quinn, B. G. (1982). *Random coefficient autoregressive model : An introduction*. Springer Verlag, New York.
- [90] Osborn, D. R. (1992). The implication of periodically varying coefficients for seasonal time-series processes. *Journal of Econometrics*, **48**, 373-384.
- [91] Osborn, D. R. and Smith, J. P. (1989). The performance of periodic autoregressive models in forecasting seasonal U. K. consumption. *Journal of Business and Economic Statistics*, **7**, 117-127.
- [92] Pagano, M. (1978). On periodic and multiple autoregression. *The Annals of Statistics*, **6**, 1310-1317.

- [93] Panuska, V. (1968). A stochastic approximation method for identification of linear systems using adaptive filtering. *In Joint Automatic Control Conference*, 1014-1021. Ann. Arbor. Mich.
- [94] Pantula, S. G. (1988). Estimation of autoregressive models with *ARCH* errors. *Sankhya*, **50**, 119-138.
- [95] Plackett, R. L. (1950). Some theorems in least squares. *Biometrika*, **32**, 149.
- [96] Pollock, D. S. G. (2003). Recursive estimation in econometrics. *Computational Statistics and Data Analysis*, **44**, 37-75.
- [97] Robbins, H., and Siegmund, D. (1971). A convergence theorem for nonnegative almost supermartingales and some applications. *In Optimizing Meth. Statist., Proc. Sympos (p.233-257)*. Ohio State University.
- [98] Solo, V. (1979). The convergence of AML. *IEEE Trans. Autom. Control*. **24**, 958-963.
- [99] Solo, V. (1980). Some aspects of recursive parameter estimation. *International Journal of control*, **32**, 395-410.
- [100] Solo, V. (1981). The second order properties of a time series recursion. *The Annals of Statistics*, **9**, 307-317.
- [101] Solo, V. and Kong, X. (1995). *Adaptive Signal Processing Algorithms*. Prentice-Hall, Englewood Cliffs, NJ.
- [102] Swensen, A. R. (1985). The Asymptotic Distribution of the likelihood ratio for autoregressive time series with a regression trend. *Journal of Multivariate Analysis*. **16**, 54-70.
- [103] Tiao, G. C. and Grupe, M. R. (1980). Hidden periodic autoregressive-moving average models in time series data. *Biometrika*, **67**, 365-373.
- [104] Tol, R. J. (1996). Autoregressive conditional heteroscedasticity in daily temperature measurements, *Environmetrics*, **7**, 67-76.
- [105] Tong, H. (1990). *Nonlinear Time Series : a Dynamical System Approach*. Oxford University Press, Oxford.
- [106] Tong, H. (1978). *On a threshold model*, in C. H. Chen (ed), *Pattern Recognition and Signal Processing*, Sijthoff and Noordhoff, The Netherlands.

- [107] Troutman, B. M. (1979). Some results in periodic autoregression. *Biometrika*, **66**, 219-227.
- [108] Tsay R. S. (1987). Conditional heteroskedastic time series models. *Journal of the American Statistical Association*, **7**, 590-604.
- [109] Vecchia, A. V. (1985a). Maximum likelihood estimation for periodic autoregressive Moving average models. *Technometrics*, **27**, 375-384.
- [110] Vecchia, A. V. (1985b). Periodic autoregressive-moving average (*PARMA*) modeling with application to water resources. *Water Resources Bulletin*, **21**, 721-730.
- [111] Vecchia, A. V. and Ballerini, R. (1991). Testing for periodic autocorrelations in seasonal time series data. *Biometrika*, **78**, 53-63.
- [112] Vecchia, A. V., Obeysekera, J. T. and Salas, J. D. (1983). Aggregation and estimation for low-order periodic ARMA models. *Water Resources Research*, **19**, 1297-1306.
- [113] Weiss, A. A. (1984). ARMA models with ARCH errors. *Journal of Time Series Analysis*, **3**, 129-143.
- [114] Weiss, A. A. (1986). Asymptotic theory for *ARCH* models : estimation and testing. *Econometric Theory*, **2**, 107-131.
- [115] Whittle, P. (1953). Estimation and information in time series. *Ark. Math. Fys. Astr.*, **2**, 433-434.
- [116] Wold, H. (1938). *A study in the stationary time series*. Almqvist and Wiksell, Stockholm (2nd edition : 1954).
- [117] Wong, H. and Li, W. K., (1997). On multivariate conditional heteroskedasticity model. *Biometrika*, **4**, 111-123.
- [118] Young, P. C. (1974). Recursive approaches to time series analysis. *Bull. Inst. Math. Appl.*, **10**, 209-224.
- [119] Young, P. C. (1984). *Recursive Estimation and Time-Series Analysis, An Introduction*. Springer-Verlag.