

REPUBLIQUE ALGERIENNE DEMOCRATIQUE et POPULAIRE.  
Ministère de l'Enseignement Supérieur et de la Recherche Scientifique.

UNIVERSITE MOULOU D MAMMERI, TIZI-OUZOU  
Faculté des Sciences  
Département de Mathématiques

**MEMOIRE DE MASTER**  
en  
**MATHEMATIQUES**

Spécialité  
Processus stochastiques et statistique de décision

Thème  
**Sur les modèles AutoRégressifs Double (DAR)**

Présenté par  
AMARI Kahina

Devant le jury d'examen composé de :

Fazia BEDOUHENE	Professeur	UMMTO	Présidente
Abdelghani HAMAZ	MCB	UMMTO	Rapporteur
Farida ACHEMINE	MCA	UMMTO	Examinatrice

Soutenu le 14/10 / 2017

# Table des matières

<b>Introduction</b>	<b>1</b>
<b>1 Généralités sur les séries chronologiques</b>	<b>3</b>
1.1 Notion de stationnarité . . . . .	4
1.2 Modèle ARMA . . . . .	5
1.2.1 Processus autorégressif d'ordre $p$ AR( $p$ ) . . . . .	7
1.2.2 Processus moyenne mobile d'ordre $q$ MA( $q$ ) . . . . .	8
1.2.3 Processus autorégressif moyenne mobile ARMA( $p, q$ ) . . . . .	9
1.3 Processus Conditionnellement Hétéroscédastiques . . . . .	9
1.3.1 Modèles Autorégressif Conditionnellement Hétéroscédastiques ARCH( $p$ )	10
1.3.2 Propriétés des processus ARCH. . . . .	10
1.3.3 Modèle Autorégressif a Erreur ARCH . . . . .	11
1.3.4 Processus GARCH( $p, q$ ) . . . . .	14
1.3.5 Etude de la Stationnarité . . . . .	15
<b>2 Modèles Autorégressive Doubles</b>	<b>20</b>
2.1 Introduction . . . . .	20
2.2 Stationnarité faible et forte des modèles DAR( $p$ ) . . . . .	20
2.3 Non-stationnarité du modèle AR(1) Double . . . . .	27
2.4 L'Estimature du Maximum de Vraisemblance des paramètres du modèles DAR(1)(EMV) . . . . .	30
2.5 Maximum de vraisemblance du modèle DAR( $P$ ) . . . . .	34
2.6 Etude Par simulation . . . . .	36
2.6.1 Interpretation . . . . .	39
<b>Conclusion</b>	<b>39</b>

<b>Table des matières</b>	<b>2</b>
---------------------------	----------

---

<b>Bibliographie</b>	<b>41</b>
----------------------	-----------

### *Remerciements*

*Je tiens à exprimer toute ma reconnaissance à M.HAMAZ Abdelghani qui a proposé et accepté de diriger ce travail. Je le remercie aussi pour sa grande contribution à l'aboutissement de ce travail, et pour sa grande disponibilité malgré son emploi de temps chargé. Je me dois aussi de le saluer pour la largesse des aides qu'il m'a faites au-delà de sa responsabilité d'encadreur.*

*Je remercie vivement l'ensemble des membres du jury pour l'honneur qu'ils me font en acceptant de juger ce travail. Mes remerciements chaleureux s'adressent également à mes amis au près au de loin... qui m'ont beaucoup aidé dans ce travail.*

# Introduction

L'analyse classique des séries temporelles repose sur la construction de modèles du processus stochastique sous-jacent. Celui-ci est ensuite utilisé d'un point de vue statistique, que ce soit pour analyser la structure "causale" du processus ou pour produire des prévisions "optimales".

Une classe très importante de modèles utilisés pour la prévision des processus stationnaires est celle des AutoRégressifs-Moyenne Ajusté (ARMA). Ceux-ci découlent naturellement d'un résultat fondamental dû à Wold (1938) qui peut s'énoncer comme suit : tout processus "purement non déterministe", stationnaire au second ordre et tel que  $\mathbb{E}(X_t) = 0$ , admet une écriture moyenne mobile infinie. Les modèles linéaires occupent une place centrale dans la théorie et le traitement des séries temporelles. Ils s'appuient sur des concepts tels que la stationnarité, l'innovation dont nous rappelons, dans ce chapitre, les principales propriétés et utilisations. De même que de nombreux processus stationnaires au second ordre ont une espérance conditionnelle non constante, l'hétéroscédasticité conditionnelle est parfaitement compatible avec la stationnarité (au sens strict ou au second ordre). Les modèles introduits dans la littérature économétrique afin de prendre en compte les propriétés particulières des séries financières (variations de prix ou du logarithme des prix, taux d'intérêt etc.) se présentent généralement sous la forme multiplicative  $\epsilon_t = \sigma_t \eta_t$ . Les modèles ARCH (autorégressifs conditionnellement hétéroscédastiques) ont été introduits par Engle (1982) et leur extension GARCH (ARCH généralisés) est due à Bollerslev (1986). Leur caractérisation repose essentiellement sur le concept de variance conditionnelle. Dans ces modèles, celle-ci s'écrit comme une fonction affine des valeurs passées du carré de la série. Cette spécification particulière se révèle très fructueuse car elle permet une étude complète des propriétés des solutions tout en étant assez générale.

Les modèles GARCH sont en effet susceptibles de capter les propriétés caractéristiques de certaines séries. Les processus GARCH étudiés dans ce premier chapitre illustreront abon-

damment ce point. La structure “linéaire” de ces processus est mise en évidence à travers plusieurs représentations que nous allons détailler. Une autre classe de modèles de séries chronologiques, qui constitue un cas particulier des ARMA-GARCH et qui est différente des GARCH est les autorégressifs doubles DAR. Ces derniers ont été introduits dans la littérature des séries temporelles en (1984). Ils ont fait l’objet de plusieurs auteurs dont : Min Chen Dong Li(2013) et Chiqing Ling(2003).

Ce mémoire est organisé comme suit: le premier chapitre sera constitué une brève présentation de certains modèles de séries temporelles. L’accent sera mis sur l’étude de la stationnarité des modèles ARMA et ARCH-GARCH en caractérisant d’une façon précise les conditions nécessaires et/ou suffisantes de la faible et forte stationnarité. Le second chapitre fera l’objet de l’étude des modèles Autorégressifs Doubles(DAR) et l’étude de leurs conditions suffisantes et nécessaires de stationnarité et de non-stationnarité. Ensuite, nous nous intéresserons à l’estimation des paramètres de ce type de modèle par la méthode de maximum de vraisemblance(EMV). Aussi, il sera question de développer les propriétés asymptotiques de cet estimateur. L’étude sera achevée par un travail de simulation dont l’objectif est de quantifier la performance de l’estimateur retenu.

# Chapitre 1

## Généralités sur les séries chronologiques

L'analyse des séries temporelles, ou séries chronologiques, se réfère à la branche de la statistique où les observations sont régulièrement espacées dans le Temps. Une série chronologique, est une série d'observations ordonnées Chronologiquement. Elles se rencontrent naturellement dans une grande variété de domaines. On peut citer: l'économie (taux de chômage, . . .), la finance (cours d'action, taux d'intérêt, . . .) l'écologie (pollution à l'ozone, au  $CO_2$ , . . .), le transport (avec l'exemple célèbre du trafic aérien international), la démographie. . . .etc. Historiquement, c'est dans les XVIe - XVIIe siècles qu'apparaissent les premières séries temporelles (ou chronologiques) dans les écrits scientifiques. Mais le plus ancien graphique connu remonte à l'antiquité (en terme de représentation de l'inclinaison des orbites des planètes en fonction du temps), les Romains savaient déjà que l'année dure 365 et 1/4 jours. Les objectifs d'étude sont multiples. La prévision est sans doute le but le plus fréquent. Il s'agit de prévoir les valeurs futures d'une variable grâce aux valeurs observées dans le présent et le passé de cette même variable, la problématique n'est donc pas la même qu'en régression où l'on cherche à prévoir le niveau d'une variable (la réponse) en fonction du niveau d'autres variables (les prédicteurs). Les séries chronologique ont donc suscité un très vif intérêt, ce qui a eu par conséquence le développement de nombreux modèles: AR, MA, ARMA, ARCH et GARCH, ARMA-GARCH,...

Tout au long de ce mémoire nous noterons  $(\Omega, F, P)$  un espace probabilisé, sur lequel nous considérerons une suite de variables aléatoires réelles  $\{X_t, t \in \mathbb{Z}\}$ . Une telle suite est appelée série temporelle et constitue un exemple de processus stochastique à temps discret.

## 1.1 Notion de stationnarité

La théorie des processus aléatoires est essentiellement basée sur la notion de stationnarité. En particulier, cette notion a rendu facile la modélisation des plusieurs phénomènes réels par l'approche des séries chronologiques. Bon nombre des phénomènes dans la vie réelle montrent une évolution qui ne s'écarte pas trop loin d'un état " d'équilibre statistique ". Ces phénomènes peuvent être assimilés à des processus aléatoires stationnaires dans la mesure où leurs propriétés statistiques et probabilistes restent les mêmes au cours du temps. L'une des grandes questions dans l'étude des séries chronologique, est donc de savoir si celles-ci suivent un processus stationnaire.

Conceptuellement, on distingue la stationnarité forte et la stationnarité faible.

**Définition 1.1 (Stationnarité forte).** Le processus  $\{X_t; t \in \mathbb{Z}\}$  est dit strictement stationnaire, ou fortement stationnaire. Si quel que soit le  $n$ -uplet de temps  $t_1, t_2, \dots, t_n$  tel que  $t_1 < t_2 < \dots < t_n$  Le vecteur  $(X_{t_1+h}, X_{t_2+h}, \dots, X_{t_n+h})'$  a la même loi que le vecteur  $(X_{t_1}, \dots, X_{t_n})'$ .

Cependant, cette notion est rarement vérifiée en pratique, alors que la notion suivante peut sembler moins exigeante car elle n'impose des contraintes qu'aux deux premiers moments des variables  $X_t; t \in \mathbb{Z}$ .

**Définition 1.2 (Stationnarité Faible).** Un processus  $\{X_t; t \in \mathbb{Z}\}$  est dit stationnaire au second ordre, ou stationnaire au sens faible, ou stationnaire d'ordre deux si les trois conditions suivantes sont satisfaites:

$$\begin{aligned} \forall t \in \mathbb{Z} \quad & E(X_t^2) < \infty. \\ \forall t \in \mathbb{Z}, \quad & E(X_t) = m, \quad \text{indépendante de } t. \\ \forall (t, h) \in \mathbb{Z}^2, \quad & \text{Cov}(X_t, X_{t+h}) = E[(X_{t+h} - m)(X_t - m)] = \gamma_X(h), \quad \text{indépendante de } t. \end{aligned}$$

La première condition  $E(X_t^2) < \infty$  garantit tout simplement l'existence (ou la convergence) des moments d'ordre deux. La seconde condition porte sur les moments d'ordre un, et signifie tout simplement que les variables aléatoires  $X_t, t \in \mathbb{Z}$  doivent avoir la même espérance quelle que soit la date  $t$ . Autrement dit, l'espérance du processus  $\{X_t, t \in \mathbb{Z}\}$  doit être indépendante du temps. Enfin, la troisième condition,  $\gamma_X(t)$  indépendant de  $t$ , porte sur les moments d'ordre deux résumés par la fonction d'autocovariance. Cette condition implique que ces moments doivent être indépendants de la date considérée, et ne doivent dépendre uniquement que de l'ordre des retards. Autrement dit, la fonction d'autocovariance du processus  $\{X_t, t \in \mathbb{Z}\}$  doit être indépendante du temps.

En résumé, un processus est stationnaire au second ordre si les moments d'ordre 1 et 2 sont

indépendants du temps, et que la covariance ne dépend que du pas  $h$ . Par conséquent, il convient de noter que la stationnarité implique que la variance  $\gamma(0)$  du processus  $\{X_t, t \in \mathbb{Z}\}$  est constante au cours du temps.

**Remarque 1.1.** • La stationnarité forte implique les deux dernières conditions de la stationnarité faible. Autrement dit, la stationnarité forte implique la stationnarité faible si le moment d'ordre deux est fini. La réciproque n'est généralement pas vraie.

**Contre exemple:** Soit  $\{X_t\}_t$  une suite de variables aléatoires indépendantes telle que

$$\begin{cases} X_t \sim \exp(1), & \text{lorsque } t \text{ est pair} \\ X_t \sim N(1,1), & \text{lorsque } t \text{ est impair} \end{cases}$$

alors  $\{X_t\}_{t \in \mathbb{Z}}$  est stationnaire avec  $\gamma_X(0) = 1$  et  $\gamma_X(h) = 0$  lorsque  $h \neq 0$ . Cependant  $X_1$  et  $X_2$  n'ont pas la même loi, donc  $\{X_t\}$  n'est pas stationnaire.

Cependant, il y'a une classe importante de processus pour laquelle l'assertion: stationnarité faible implique stationnarité stricte est vraie: il s'agit des processus gaussiens.

En effet, ils sont caractérisés par leur espérance et leur covariance.

- La stationnarité faible est bien plus facile à étudier et à vérifier que la stationnarité stricte.
- Un processus stationnaire n'est pas obligatoirement borné ou sympathique.
- Pour un processus du second ordre, la stationnarité stricte implique la stationnarité faible. La réciproque est fausse (elle n'est vraie que pour les processus gaussiens)

## 1.2 Modèle ARMA

Décrire et prédire les comportements des phénomènes, qui évoluent dans le temps conduit à l'introduction de plusieurs modèles probabiliste rencontrés dans la littérature. Parmi ces modèles, les plus célèbres depuis à peu près un demi siècle, sont les modèles autorégressifs moyennes ajusté qui constituent la classe ARMA.

Il est possible de définir cette classe à partir de la fameuse décomposition de Wold (1938). Les modèles AR et MA peuvent être parfait, dans certain cas, mais il peut arriver que l'on ait besoin d'estimer un grand nombre de paramètres pour ajuster le modèle retenu.

Si l'on dispose de peu d'observations, ces estimations auront tendance à être plus précises. De plus, si un modèle contenant  $p$  paramètres convient pour la situation, il n'est pas bon de tenter d'ajuster un modèle qui contiendra plus que  $p$ . Les modèles ARMA consiste à avoir une partie autorégressive et une partie à moyenne mobile. Herman Wold(1938) à

montré que les processus ARMA pouvaient être utilisés pour modéliser n'importe quelle série stationnaire pour autant que les order p et q soient bien choisis.

Box et Jinkins ont travaillé pour développer une méthodologie pour: Edentification, l'estimation et prévision.

Dans cette section, il sera question de présenter rigoureusement les différents modèles issus de la classe ARMA. Pour ce faire, nous allons présenter le bruit blanc qui est à l'origine de ces modèles et les notions d'inversibilité et de causalité ainsi que décomposition de wold(1938).

**Décomposition de Wold** Le théorème de Wold qui sera énoncé ci-dessous était à l'origine de la décomposition des processus stationnaires en général. En particulier, cette décomposition va permettre de justifier théoriquement l'utilisation de la classe ARMA.

**Théorème 1.1.** "Décomposition de Wold", *Tout processus stationnaire  $(Y_t)_{t \in \mathbb{Z}}$  possède une décomposition unique donnée par :*

$$Y_t = C_t + X_t,$$

*telle que : Les deux processus  $C_t$  et  $X_t$  soient orthogonaux. De plus le processus  $C_t$  est purement déterminable (singulier) et  $X_t$  est purement indéterminable (régulier). Ainsi,  $\{X_t\}$  peut être représenté par une combinaison linéaire, du présent et du passé d'un processus bruit blanc, convergente en moyenne quadratique unique de la forme :*

$$X_t = \sum_{j=0}^{\infty} \theta_j \epsilon_{t-j} = \epsilon_t + \theta_1 \epsilon_{t-1} + \theta_2 \epsilon_{t-2} + \dots$$

*avec  $\sum_j \theta_j^2 < \infty, \theta_0 = 1, \{\epsilon_t\}_{t \in \mathbb{T}}$  un processus bruit blanc dit processus d'innovation du processus  $\{X_t\}_{t \in \mathbb{Z}}$ .*

**Remarque 1.2.** ◦ Le processus bruit blanc qui est utilisé dans le théorème précédant représente l'erreur de prévision à l'horizon 1, il sera appelé innovation du processus  $\{X_t\}$ .

◦ Etant donné que la composante  $C_t$  est déterministe, on peut donc l'ajuster facilement par des méthodes type moindres carrés. Par contre l'étude de la composante aléatoire  $\{X_t\}$  sera d'un intérêt particulier.

◦ L'implication forte de ce théorème est que, si on omet la composante déterministe  $C_t$  tout processus stationnaire peut s'écrire comme une somme pondérée des résidus  $\epsilon_t$  d'une part et d'autre part, si on connaît les pondérations  $\theta_j$  et la variance  $\sigma^2$  du bruit blanc, on peut proposer une représentation dite moyenne mobile de tout processus stationnaire.

**Définition 1.3 (Le processus bruit blanc).** Soit  $\{\epsilon_t, t \in \mathbb{Z}\}$  un processus stochastique, on dit que  $\{\epsilon_t, t \in \mathbb{Z}\}$  est un processus stochastique hasard pure ou bruit blanc faible si les trois propriétés suivantes sont vérifiées:

- $\mathbb{E}(\epsilon_t) = m, \forall t \in \mathbb{Z}$
- $v(\epsilon_t) = \sigma^2, \forall t \in \mathbb{Z}$
- $cov(\epsilon_t, \epsilon_{t+h}) = 0, \forall t \in \mathbb{Z}, \forall h \in \mathbb{Z}$ .

**Définition 1.4 (Notion de causalité).** D'une manière générale, un modèle de série chronologique (linéaire ou non) de la forme

$$X_t = g(X_{t-1}, \dots, X_{t-p}, \epsilon_t, \dots, \epsilon_{t-q})$$

avec  $\mathcal{BB}$  un bruit blanc gaussienne, est dit causal si et seulement si, on peut exprimer le processus stochastique  $\{X_t\}$  sous forme de combinaison linéaire (finie ou infinie), convergente au sens de  $L_2$ , du présent et du passé du bruit blanc  $\epsilon_t$ . Autrement dit, s'il existe une suite  $(a_i)$  des réels tels que  $\sum_{i=0}^{\infty} |a_i| < \infty$  et que  $X_t = \sum_{i=0}^{+\infty} a_i \epsilon_{t-i}, t \in \mathbb{Z}$

**Définition 1.5 (Notion D'invisibilité).** D'une manière générale, un processus est dit inversible si on peut l'écrire comme une autorégression infinie.

Autrement dit, s'il existe une suite de nombre  $\Pi_j$  absolument sommable tel que,  $\epsilon_t = \sum_{j=1}^{+\infty} \Pi_j X_{t-j}$ .

### 1.2.1 Processus autorégressif d'ordre p AR(p)

**Définition 1.6.** On appelle processus autorégressif d'ordre p noté AR(p), un processus stationnaire  $\{X_t, t \in \mathbb{Z}\}$  vérifiant l'équation aux différences stochastique suivante :

$$X_t - \phi_1 X_{t-1} - \phi_2 X_{t-2} - \dots - \phi_p X_{t-p} = \epsilon_t$$

où  $\epsilon_t \sim \mathcal{BB}(0, \sigma^2)$  ou  $\mathcal{BB}$  est un Bruit Blanc et les  $\phi_j$  sont des réels. En utilisant l'opérateur retard B, on a l'écriture équivalente suivante:  $\phi(B)X_t = \epsilon_t$  avec  $\phi(B)$  est le polynôme retard de degré p défini par:  $\phi(B) = 1 - \phi_1 B^1 - \phi_2 B^2 - \dots - \phi_p B^p$

**Remarque 1.3.** La définition donnée du processus autorégressif est implicite et il n'est pas évident que l'équation  $\epsilon_t = X_t - \sum_{j=1}^p \phi_j X_{t-j}$  admette une solution stationnaire.

1. Un processus autorégressif est toujours inversible d'après la définition de l'invisibilité d'un processus.

2. On démontre qu'une condition nécessaire et suffisante pour qu'il existe un processus  $\{X_t, t \in \mathbb{Z}\}$  AR(p) stationnaire, solution de l'équation  $\phi(B) = \epsilon_t$  est que toutes les racines du polynôme caractéristique  $\phi(z) = 1 - \phi_1 z - \phi_2 z^2 - \dots - \phi_p z^p$  soient différentes de 1. Pour l'existence d'une solution stationnaire causale il faut que les racines du polynôme soient inférieures à 1.

### 1.2.2 Processus moyenne mobile d'ordre q MA(q)

**Définition 1.7.** Un processus du second ordre  $\{X_t, t \in \mathbb{Z}\}$  est appelé moyenne mobile ou moyenne ajustée d'ordre q notée MA(q) s'il satisfait à l'équation aux différences stochastiques suivante:

$$X_t = \epsilon_t - \sum_{j=1}^q \theta_j \epsilon_{t-j} = \epsilon_t - \theta_1 \epsilon_{t-1} - \theta_2 \epsilon_{t-2} - \dots - \theta_q \epsilon_{t-q}$$

où  $\epsilon_t \sim \mathcal{BB}(0, \sigma^2)$  On a l'écriture équivalente suivante en utilisant l'opérateur retard B:  $X_t = \theta(B)\epsilon_t$  avec  $\theta(B) = 1 - \theta_1 B^1 - \theta_2 B^2 - \dots - \theta_q B^q$  avec  $\theta_j \in \mathbb{R}$

**Remarque 1.4.** 1. La définition d'un MA(q) est, contrairement au cas d'un AR(p) explicite et ne pose pas de problème.

2. Un processus MA(q) est toujours stationnaire (car défini à partir d'un bruit blanc stationnaire).
3. Un processus MA(q) comme le montre sa définition ne s'exprime pas directement en fonction de ses observations passées, mais en fonction du bruit blanc qui n'est pas observé. Il n'est donc pas toujours inversible.

On démontre qu'une condition nécessaire et suffisante pour qu'un processus MA(q) soit inversible est que les racines du polynôme caractéristique  $Z^q - \theta_1 Z^{q-1} - \theta_2 Z^{q-2} - \dots - \theta_q = 0$  soient différentes de 1. Pour des raisons identiques au cas d'un processus AR(p), on supposera les racines strictement supérieures à 1 (c'est à dire  $|Z| > 1$ .)

4. Il est clair que le processus MA(q) est toujours causal par définition de la causalité.

**Fonction d'autocorrelation d'un MA(q)** Soit  $\{X_t; t \in \mathbb{T}\}$  un processus MA(q) tel que:  $X_t = \epsilon_t - \theta_1 \epsilon_{t-1} - \theta_2 \epsilon_{t-2} - \dots - \theta_q \epsilon_{t-q}$  La fonction d'autocovariance du processus MA(q) est donnée par:  $\gamma(h) = (1 + \theta_1^2 + \theta_2^2 + \dots + \theta_q^2)\sigma^2$  si  $|h| < q$

### 1.2.3 Processus autorégressif moyenne mobile ARMA(p, q)

Les modèles ARMA appelés aussi ARMA mixtes constituent une généralisation des modèles AR et MA.

**Définition 1.8.** Un processus du second ordre  $\{X_t; t \in \mathbb{Z}\}$  est dit admettre une forme ARMA(p;q), s'il est solution de l'équation aux différences stochastiques suivante:  $X_t - \sum_j^p \phi_j X_{t-j} = \epsilon_t - \sum \theta_t \epsilon_{t-j}$  ou encore  $\phi(B)X_t = \theta(B)\epsilon_t$  avec  $\epsilon_t \sim \mathcal{BB}(0; \sigma^2)$ .

**Définition 1.9.** La représentation ARMA(p;q) est dite minimale si elle vérifie les propriétés suivantes :

- $\phi_p \neq 0, \theta_q \neq 0$
- Les polynômes  $\phi$  et  $\theta$  ont leurs racines en module strictement supérieur à 1,
- $\phi$  et  $\theta$  n'ont pas de racines communes,
- $\epsilon_t \sim \mathcal{BB}(0; \sigma^2)$  tel que  $\sigma^2 \neq 0$ .

**Remarque 1.5.** Dans la représentation d'un ARMA(p;q), il est facile de voir que :

- si  $q=0$ , alors ARMA(p;0) est un AR(p)
- si  $p=0$ , alors ARMA(0;q) est un MA(q), et donc le processus AR(p) et MA(q) sont bien des ARMA particulières.
- si  $q=0, p=0$ , le modèle théorique ARMA(0;0) est l'unique processus qui admet simultanément une représentation AR pure et une représentation MA pure.

## 1.3 Processus Conditionnellement Hétéroscédastiques

La modélisation des séries financières se révèle être un problème complexe. Cette complexité tient surtout à l'existence de régularité statistique commune à un très grand nombre de séries financières, de ce fait, celles-ci sont difficiles à reproduire à partir des modèles stochastiques. Les processus ARCH (Autoregressive Conditionnel Heteroskedasticity), proposé par Engle en 1982 et leur extensions GARCH (ARCH généralisés) proposé par Bollerslev (1986) représente une bonne approximation pour la modélisation de ces séries. Ces modèles économétriques se présentent sous la forme multiplicative  $\epsilon_t = \sigma_t \eta_t$ , à la variable aléatoire  $\sigma_t$  appelée volatilité (variance conditionnelle) est parfaitement connue à une date t comme fonction affine des valeurs passées du carré de la série  $\epsilon_t^2$ . Cette spécification particulière se révèle très fructueuse, car elle permet une étude complète des propriétés des solutions. L'objectif de cette section, est, en premier temps de mettre en exergue les définitions des modèles ARCH ainsi que ses propriétés, et GARCH. Et en deuxième étape, nous donnons

les conditions de stationnarité au sens fort et au sens faible de tels processus.

### 1.3.1 Modèles Autorégressif Conditionnellement Hétéroscédastiques ARCH(p)

**Définition 1.10.** Un processus  $\{\epsilon_t\}$  satisfait une représentation ARCH(p) si  $\epsilon_t = \eta_t h_t$  avec

$$h_t = \sqrt{\alpha_0 + \sum_{i=1}^p \alpha_i \epsilon_{t-i}^2},$$

où  $\eta_t$  désigne un bruit blanc fort, tel que  $\mathbb{E}[\eta_t] = 0$  et  $v[\eta_t] = \sigma^2$ . Pour ce type de processus, on retrouve les deux propriétés essentielles vues précédemment à savoir, la propriété de différence de martingale (ou bruit blanc faible)  $\mathbb{E}(\epsilon_t/I_{t-s}) = 0$  et la propriété de variance conditionnelle dans le temps puisque :

$$v[\epsilon_t/I_{t-1}] = h_t^2 = \alpha_0 + \sum_{i=1}^p \alpha_i \epsilon_{t-i}^2,$$

de plus

$$\epsilon_t^2 = \alpha_0 + \alpha_1 \epsilon_{t-1}^2 + v_t,$$

$$\mathbb{E}[v_t/I_{t-1}] = 0.$$

### 1.3.2 Propriétés des processus ARCH.

**Propriétés 1.3.1.** On peut noter que pour tout  $s < 1$ :  $\mathbb{E}(\epsilon_t/I_{t-s}) = 0$ . Pour démontrer cela utilisons la propriété des espérances itérées. En effet, on a

$$\begin{aligned} \mathbb{E}[\epsilon_t/I_{t-1}] &= \mathbb{E}[\mathbb{E}[\epsilon_t/I_{t-1}]/I_{t-s}] \\ &= \mathbb{E}[0/I_{t-s}] \\ &= 0 \quad \forall s > 1 \end{aligned}$$

car

$$I_{t-s} \subset I_{t-1}, \forall s > 1.$$

**Propriétés 1.3.2.** La variance conditionnelle du processus  $\{\epsilon_t\}$  satisfaisant une représentation ARCH(1), définie par l'équation  $\epsilon_t = \eta_t h_t$  est non constante dans le temps et vérifie:

$$v[\epsilon_t/I_{t-s}] = \alpha_0 \left[ \frac{1 - \alpha_1^s}{1 - \alpha_1} \right] + \alpha_1^s \epsilon_{t-s}^2, \forall t.$$

C'est la propriété centrale des processus ARCH, le processus  $\{\epsilon_t\}$  a une variance conditionnelle qui dépend du temps.

En effet, On sait que  $\mathbb{E}[\epsilon_t/I_{t-s}] = 0$  dès lors,  $v[\epsilon_t/I_{t-s}] = \mathbb{E}[\epsilon_t^2/I_{t-s}]$ . Considérons le processus  $\epsilon_t^2$  défini par la relation  $\epsilon_t^2 = \alpha_0 + \alpha_1 \epsilon_{t-1}^2 + \nu_t$  où  $\nu$  est un bruit blanc faible. Par itération successive, on a:

**Propriétés 1.3.3.** Les autocovariances conditionnelles du processus  $\{\epsilon_t\}$ , ARCH(1), définies par l'équation  $\epsilon_t = \eta_t h_t$  sont nulles

$$cov[\epsilon_t, \epsilon_{t+k}/I_{t-s}] = 0, \forall k, s \geq 1$$

Le processus est donc un processus sans mémoire conditionnellement à  $I_{t-s}, \forall s > 1$

**Remarque 1.6.** L'absence de corrélations entre les valeurs d'un processus ARCH est une caractéristique très importante de cette famille de modèle, qui les rend utiles pour modéliser certaines séries financières.

Toute, ces propriétés peuvent être généralisées au cas d'un processus ARCH(p).

### 1.3.3 Modèle Autorégressif a Erreur ARCH

On considère dorénavant non plus un processus ARCH pour modéliser directement la série financière, mais les résidus d'un modèle linéaire. Prenant l'exemple d'un modèle linéaire autorégressif avec résidus de type ARCH(p).

On suppose que le résidu admet un représentation autorégressif de type ARCH(p)  $\epsilon_t = \eta_t h_t$  avec  $h_t = \sqrt{\alpha_0 + \sum_{i=1}^p \alpha_i \epsilon_{t-i}^2}$  et où  $\eta_t$  désigne un bruit blanc faible. On a un modèle qui décrit à la fois l'évolution de l'espérance conditionnelle et la variance conditionnelle du processus  $\{X_t\}$  dans le temps. Envisageons le cas le plus simple d'un processus de type AR(1) avec erreur ARCH(1)

$$X_t = \delta + \alpha X_{t-1} + \epsilon_t, \quad |\alpha| < 1,$$

avec

$$\epsilon_t = \eta_t \sqrt{\alpha_0 + \alpha_1 \epsilon_{t-1}^2}$$

Dans ce cas, les résidus satisfont les principales propriétés étudiées précédemment.

1. Le processus  $\{\epsilon_t\}$  est orthogonal aux valeurs passées, quelque soit le retard,

$$\mathbb{E}[\epsilon_t/I_{t-s}] = 0, \forall s \geq 1,$$

et suit un processus ARCH(1). La variance conditionnelle est

$$v[\epsilon_t/I_{t-s}] = \alpha_0 \left[ \frac{1 - \alpha_1^s}{1 - \alpha_1} \right] + \alpha_1^s \epsilon_{t-s}^2$$

$$v[\epsilon_t] = \frac{\alpha_0}{1 - \alpha_1}.$$

2. La propriété d'orthogonalité implique que les corrélations conditionnelles sont nulles

$$Cov[\epsilon_t \epsilon_{t+k}/I_{t-s}] = 0.$$

En effet, il y a donc une absence de corrélation entre les valeurs présentes et futures du processus, quels que soient les retards  $s$  et  $k$ . Mais si la variance conditionnelle de  $\epsilon_t$  n'est pas constante, la variance non conditionnelle est constante. On peut, en outre, en déduire un certain nombre de conclusions quant au processus  $\{X_t\}$  lui-même. On peut montrer tout d'abord, que l'espérance conditionnelle de  $X_t$  vérifie :

$$\mathbb{E}[X_t/I_{t-s}] = \delta + \alpha \mathbb{E}[X_{t-1}/I_{t-s}], \forall s \geq 1,$$

ce qui montre que les prévisions non linéaires de  $X_t$  s'obtiennent comme les prévisions linéaires d'un processus AR (1). Plus généralement

$$X_t = \delta \frac{1 - \alpha^s}{1 - \alpha} + \alpha^s X_{t-s} + \epsilon_t + \alpha \epsilon_{t-1} + \dots + \alpha^{s-1} \epsilon_{t-s+1}.$$

En effet

$$\begin{aligned} X_t &= \delta + \alpha X_{t-1} + \epsilon_t \\ &= \delta + \alpha(\delta + \alpha X_{t-2} + \epsilon_{t-1}) + \epsilon_t \\ &= \delta(1 + \alpha) + \alpha^2 X_{t-2} + \alpha \epsilon_{t-1} + \epsilon_t \\ &= \delta(1 + \alpha + \alpha^2) + \alpha^3 X_{t-3} + \alpha^2 \epsilon_{t-2} + \alpha \epsilon_{t-1} + \epsilon_t \\ &\vdots \\ &= \delta(1 + \alpha + \alpha^2 + \dots + \alpha^{s-1}) + \alpha^s X_{t-s} + \alpha^{s-1} X_{t-s+1} + \dots + \alpha \epsilon_{t-1} + \epsilon_t \\ &= \delta \frac{1 - \alpha^s}{1 - \alpha} + \alpha^s X_{t-s} + \alpha^{s-1} \epsilon_{t-s+1} + \dots + \alpha \epsilon_{t-1} + \epsilon_t \end{aligned}$$

En prenant l'espérance conditionnelle de deux cotés, on obtient

$$\mathbb{E}[X_t/I_{t-s}] = \delta \frac{1 - \alpha^s}{1 - \alpha} + \alpha^s X_{t-s}.$$

De même façon, on peut montrer que la variance conditionnelle de  $X_t$  dépend du temps. En effet, on montre qu'elle dépend du processus  $\{\epsilon_{t-s}^2\}$  de la façon suivante.

**Propriétés 1.3.4.** *La variance conditionnelle du processus AR(1) avec erreur ARCH(1),  $X_t$ , s'écrit*

$$v[X_t/I_{t-s}] = \frac{\delta}{1 - \alpha_1} \left[ \left( \frac{1 - \alpha^{2s}}{1 - \alpha^2} \right) - \alpha_1 \left( \frac{\alpha_1^s - \alpha^{2s}}{\alpha_1 - \alpha^2} \right) \right] + \alpha_1 \left[ \frac{\alpha_1^s - \alpha^{2s}}{\alpha_1 - \alpha^2} \right] \epsilon_{t-s}^2$$

Ainsi, la variance conditionnelle d'une erreur de prévision à l'horizon 1, s'écrit :

$$v[X_t/I_{t-1}] = \delta + \alpha_1 \epsilon_{t-1}^2.$$

Le cas particulier  $p = 0$  correspond à un ARCH(1)

**Définition 1.11.** Un processus  $\{\epsilon_t, t \in \mathbb{Z}\}$  satisfait une représentation ARCH(1) si :

$$\epsilon_t = \eta_t h_t$$

avec

$$h_t = \sqrt{\alpha_0 + \alpha_1 \epsilon_{t-1}^2}$$

où  $\eta_t$  est un bruit blanc faible, tel que  $\mathbb{E}(\eta_t) = 0$  et  $v(\eta_t) = 1$  et  $\alpha_0 > 0, 0 < \alpha_1 < 1$ , la composante  $h_t$  désigne une variable qui, conditionnellement à l'ensemble d'information des valeurs passées de  $\epsilon_t$ , à  $I_{t-1} = \sigma\{\epsilon_{t-1}, \epsilon_{t-2}, \dots, \epsilon_{t-j}, \dots\}$ , est déterministe et positive. Dans le cas, le processus  $\epsilon_t$  est caractérisé par des autocorrélations nulles  $\mathbb{E}(\epsilon_t \epsilon_s) = 0$  pour  $t \neq s$ , ce qui signifie que les  $\epsilon_t$  sont non corrélés dans le temps, en effet,  $\epsilon_t$  et constitue un bruit blanc faible.

**Remarque 1.7.** Un bruit blanc faible implique que les résidus ont une moyenne nulle et ils sont corrélés dans le temps. Ainsi, la variance conditionnelle varie dans le temps, mais  $\{\epsilon_t\}$  est non conditionnellement homoscedastique, c'est-à-dire qu'il y a l'existence d'une variance inconditionnelle finie.

On peut établir des résultats intéressants en considérant le processus autorégressif sur  $\{\epsilon_t^2\}$ , nous pouvons écrire le modèle ARCH sous deux autres formes. prenons un modèle ARCH(1) pour les illustrer.

### 1. Forme d'équilibre:

$$h_t^2 = \sigma_t^2 + \alpha_1(\epsilon_{t-1}^2 - \sigma^2).$$

Sachant que  $\sigma^2 = \frac{\alpha_0}{1-\alpha_1}$ , nous retrouvons la forme habituelle du modèle ARCH(1) ainsi :

$$\begin{aligned} h_t^2 &= \frac{\alpha_0}{1-\alpha_1} + \alpha_1(\epsilon_{t-1}^2 - \frac{\alpha_0}{1-\alpha_1}) \\ &= \frac{\alpha_0}{1-\alpha_1} + \alpha_1\epsilon_{t-1}^2 - \frac{\alpha_0\alpha_1}{1-\alpha_1} \\ &= \alpha_0 + \alpha_1\epsilon_{t-1}^2. \end{aligned}$$

- 2.
- 3.

### 1.3.4 Processus GARCH(p,q)

Nous donnons une première définition d'un processus GARCH fondée sur les deux premières moments  $\epsilon_t$  à son passée.

**Définition 1.12 (Processus GARCH(p,q) faible).** On dit que  $\{\epsilon_t\}$  est un processus GARCH(p,q), si ses deux moments conditionnelles existent et vérifient

★  $\mathbb{E}(\epsilon_t \mid \epsilon_u, u < t) = 0, t \in \mathbb{Z}$

★ Il existe des constantes  $\omega, \alpha_i, i = 1, \dots, q$  et  $\beta_j, j = 1, \dots, p$  telle que

$$\sigma_t^2 = V(\epsilon_t \mid \epsilon_u, u < t) = \omega + \sum_{i=1}^q \alpha_i \epsilon_{t-i}^2 + \sum_{j=1}^p \beta_j \sigma_{t-j}^2, t \in \mathbb{Z}. \tag{1.1}$$

L'équation (1.1) peut être écrite de manière symbolique sous la forme plus compacte

$$\sigma_t^2 = \omega + \alpha(B)\epsilon_t^2 + \beta(B)\sigma_t^2, t \in \mathbb{Z} \tag{1.2}$$

où B est l'opérateur retard  $B^i \epsilon_t^2 = \epsilon_{t-i}^2$  et  $B^i \sigma_t^2 = \sigma_{t-i}^2$ , pour tout entier i,  $\alpha, \beta$  sont des polynômes de degrés q et p:

$$\alpha(B) = \sum_{i=1}^q \alpha_i B^i, \beta(B) = \sum_{j=1}^p \beta_j B^j.$$

Si

$$\beta(B) < 1, \sigma_t^2 = \omega + \sum_{i=1}^q \alpha_i \epsilon_{t-i}^2. \tag{1.3}$$

et le processus est appelé ARCH(q).

L'innovation du processus  $\{\epsilon_t^2\}$  est par définition la variable  $\nu_t = \epsilon_t^2 - \sigma_t^2$ . En remplaçant, dans l'équation(1.1), les variables  $\sigma_{t-j}^2$  par  $\epsilon_{t-j}^2 - \nu_{t-j}$  on obtient la représentation

$$\epsilon_t^2 = \omega + \sum_{i=1}^q (\alpha_i + \beta_i) \epsilon_{t-i}^2 + \nu_t - \sum_{j=1}^p \beta_j \nu_{t-j}, t \in \mathbb{Z}, \tag{1.4}$$

où  $r = \max(p, q)$ , avec la convention  $\alpha_i = 0$  (resp.  $\beta_j = 0$ ) si  $i > q$  (resp.  $j > p$ ). On retrouve ainsi dans cette équation la structure linéaire des modèles ARMA, permettant par exemple un calcul très simple des prévisions linéaires. Sous des hypothèses supplémentaires (impliquant la stationnarité de  $\epsilon_t^2$ ), on peut affirmer que si  $\{\epsilon_t\}$  est un GARCH(p,q),  $\{\epsilon_t^2\}$  est un processus ARMA(r,p). En particulier, le carré d'un processus ARCH(q) admet, s'il est stationnaire, une représentation AR(q). Ces représentations ARMA seront utiles pour l'estimation et l'identification des processus GARCH. Elles seront en revanche de peu d'utilité pour l'étude de la stationnarité du processus  $\{\epsilon_t; t \in \mathbb{Z}\}$  car le bruit  $\{\nu_t\}$  dépend, par construction, du passé de  $\epsilon_t$ .

**Définition 1.13 (Processus GARCH(p,q) fort).** Soit  $\{\eta_t\}$  une suite de variables (iid) de loi  $\eta$ . On dit que  $\epsilon_t$  est un processus GARCH(p,q) au sens fort (relativement à la suite  $\eta_t$ ) s'il vérifie:

$$\begin{cases} \epsilon_t &= \sigma_t \eta_t \\ \sigma_t^2 &= \omega + \sum_{i=1}^q \alpha_i \epsilon_{t-i}^2 + \sum_{j=1}^p \beta_j \sigma_{t-j}^2, \end{cases} \quad (1.5)$$

où les  $\alpha_i, \beta_j$  sont des constantes positives et  $\omega$  est une constante strictement positive. Il est clair qu'un processus GARCH fort tel que  $\sigma_t^2$  est mesurable par rapport à la tribu  $\sigma(\epsilon_u, u < t)$  est un processus GARCH au sens de définition (2.1). La réciproque n'est cependant pas vraie.

Les processus GARCH au sens de la définition (1.1) sont souvent qualifiés de demi-forts depuis l'article de Drost Nijman(1993) consacré à l'agrégation temporelle. En remplaçant  $\epsilon_{t-i}$  par  $\sigma_{t-i} \eta_{t-i}$  dans l'équation (1.1) on obtient

$$\sigma_t^2 = \omega + \sum_{i=1}^q \alpha_i \sigma_{t-i}^2 \eta_{t-i}^2 + \sum_{j=1}^p \beta_j \sigma_{t-j}^2, \quad (1.6)$$

que l'on peut écrire

$$\sigma_t^2 = \omega + \sum_{i=1}^r a_i (\eta_{t-i}) \sigma_{t-i}^2, \quad (1.7)$$

où

$$a_i(z) = \alpha_i z^2 + \beta_i, i = 1, \dots, r.$$

Cette représentation montre que dans le cas d'un GARCH fort, le processus de volatilité vérifie une équation autorégressive, mais avec coefficient aléatoire.

### 1.3.5 Etude de la Stationnarité

Dans cette section, nous allons chercher sous quelles conditions il existe des processus stationnaires au sens strict et au sens faible vérifiant les définitions (1.1) et/ou (1.2). On

s'intéresse plus particulièrement aux solutions non anticipative du modèle (1.2) c'est-à-dire aux processus  $\{\epsilon_t\}$  tel que  $\epsilon_t$  soit une fonction mesurable des variables  $\eta_{t-s}, s \geq 0$ . Nous examinons d'abord le cas du modèle GARCH(1.1) qui peut se traiter avec des techniques élémentaires. On notera,  $x > 0, \log^+ x = \max(\log x, 0)$ .

### Modèle GARCH(1.1)

Dans le cas où  $p=q=1$ , le modèle (1.5) s'écrit:

$$\begin{cases} \epsilon_t &= \sigma_t \eta_t \\ \sigma_t^2 &= \omega + \alpha \epsilon_{t-1}^2 + \beta \sigma_{t-1}^2, \end{cases} \quad (1.8)$$

avec  $\omega \geq 0, \alpha \geq 0, \beta \geq 0$ . On pose  $a(z) = \alpha z^2 + \beta$ .

**Théorème 1.2 (Stationnarité stricte du modèle GARCH(1,1)fort).** *Si :*

$$-\infty \leq \gamma := \mathbb{E}(\log \alpha \eta_t^2 + \beta < 0), \quad (1.9)$$

la série

$$h_t = \left\{ 1 + \sum_{i=1}^{\infty} a(\eta_{t-1} \dots a(\eta_{t-i})) \right\} \omega, \quad (1.10)$$

converge presque sûrement (p.s) et le processus  $\epsilon_t$  défini par  $\epsilon_t = \sqrt{h_t} \eta_t$  est l'unique solution strictement stationnaire du modèle (1.8). Cette solution est non anticipative et ergodique. Si  $\gamma \geq 0$  et  $\omega > 0$ , il n'existe pas de solution strictement stationnaire.

**Remarque 1.8.** • Le coefficient  $\gamma = \mathbb{E}[\log\{a\eta_t\}]$  existe toujours dans  $[-\infty, +\infty[$  car

$$\mathbb{E}[\log^+\{(a\eta_t)\}] \leq \mathbb{E}[a(\eta_t)] = \alpha + \beta.$$

- Dans le cas où  $\omega = 0$  et  $\gamma < 0$ , il est clair d'après (1.10) que la seule solution strictement stationnaire du modèle est  $\epsilon = 0$ . Il n'est donc naturel d'imposer  $\omega > 0$  dans la pratique.
- On voit que la condition (1.9) dépend de la loi du processus  $\eta_t$  et qu'elle n'est pas symétrique en  $\alpha$  et  $\beta$ .
- La condition (1.9) implique  $\beta \leq 1$ . Inversement, si  $\alpha + \beta < 1$ , (1.9) est vérifiée, car par application de l'inégalité de Jensen

$$\mathbb{E}[\log\{a(\eta_t)\}] \leq \log \mathbb{E}\{a(\eta_t)\} = \log(\alpha + \beta) < 0.$$

- Si (1.9) est satisfaite, elle l'est également pour tout couple  $(\alpha_1, \beta_1)$  tel que  $\alpha_1 \leq \alpha$  et  $\{\beta_1 \leq \beta\}$ . En particulier la stationnarité stricte du modèle GARCH implique celle du

modèle ARCH obtenu en supprimant  $\beta$ .

- Dans le cas ARCH(1) ( $\beta = 0$ ), la contrainte de stationnarité strict s'écrit

$$0 < \alpha < \exp\{-\mathbb{E}(\log \eta_t^2)\}. \quad (1.11)$$

Par exemple dans le cas où  $\eta_t \sim \mathcal{N}(0,1)$  la condition est :  $\alpha < 3.56$ .

**Preuve. théorème**

Utilisant de manière itérative la seconde équation du modèle (1.8), on obtient, pour  $N \geq 1$ ,

$$\begin{aligned} \sigma^2 = \omega + a(\eta_{t-1})\sigma_{t-1}^2 &= \omega[1 + \sum_{n=1}^N a(\eta_{t-1})\dots a(\eta_{t-n})] + a(\eta_{t-1})\dots a(\eta_{t-N-1})\sigma_{t-N-1}^2 \\ &:= h_t(N) + a(\eta_{t-1})\dots a(\eta_{t-N-1})\sigma_{t-N-1}^2, \end{aligned}$$

$$\begin{aligned} \sigma^2 &= \omega + a(\eta_{t-1})\sigma_{t-1}^2 \\ &= \omega[1 + \sum_{n=1}^N a(\eta_{t-1})\dots a(\eta_{t-n})] + a(\eta_{t-1})\dots a(\eta_{t-N-1})\sigma_{t-N-1}^2 \\ &:= h_t(N) + a(\eta_{t-1})\dots a(\eta_{t-N-1})\sigma_{t-N-1}^2, \end{aligned}$$

Le processus limite  $h_t = \lim_{t \rightarrow \infty} h_t(N)$  existe dans  $\overline{\mathbb{R}}^+ = [0, +\infty]$  puisque les termes de la somme sont positifs. De plus, en faisant tendre N vers l'infini dans la relation  $h_t(N) = \omega + a(\eta_{t-1})h_{t-1}(N-1)$ , on obtient:

$$h_t = \omega + a(\eta_{t-1})h_{t-1}.$$

Nous allons montrer que  $h_t$  est presque sûrement finie si et seulement si  $\gamma < 0$ . Supposon  $\gamma < 0$ . On utilise la règle de Cauchy pour les séries à termes positifs. On a :

$$[a(\eta_{t-1})\dots a(\eta_{t-n})]^{\frac{1}{n}} = \exp\left[\frac{1}{n} \sum_{i=1}^n \log a(\eta_{t-i})\right] \rightarrow \exp^{\gamma} p.s \quad (1.12)$$

quan  $n \rightarrow \infty$ , par application de la loi forte des grands nombres à la suite iid  $(\log a(\eta_t))$ . La série définie en (1.10) converge alors presque sûrement dans  $\mathbb{R}$ , par application de la règle da Cauchy, et le processus limite,  $(h_t)$ , est à valeurs réelles positives. Par suite, le processus  $\{\epsilon_t\}$  défini par:

$$\epsilon_t = \sqrt{h_t \eta_t} = \left\{ \omega + \sum_{i=1}^{\infty} a(\eta_{t-1})\dots a(\eta_{t-i})\omega \right\}^{\frac{1}{2}} \eta_t \quad (1.13)$$

est strictement stationnaire et ergodique. Il est non anticipatif comme fonction mesurables  $\eta_{t-i}, i \geq 0$ . De plus  $\epsilon_t$  vérifie le modèle (2.8) ona

$$\sigma_t^2 = h_t(N) + a(\eta_{t-1})\dots a(\eta_{t-N-1})\sigma_{t-N-1}^2.$$

par suite

$$\sigma_t^2 - h_t = \{h_{t(N)-h_t} + a(\eta_{t-1}) \dots a(\eta_{t-N-1}) \sigma_{t-N-1}^2\}$$

Le terme entre accolades à droite de l'égalité tend vers 0 p.s. Quand  $N \rightarrow \infty$ . Par ailleurs, puisque la série définissant  $h_t$  converge P.s., On a  $a(\eta_{t-1}) \dots a(\eta_{t-n}) \rightarrow 0$  avec probabilité 1 quand  $n \rightarrow \infty$ . De plus la loi de  $\sigma_{t-N-1}^2$  est indépendante de  $N$  par stationnarité. Par suite  $a(\eta_{t-1}) \dots a(\eta_{t-n}) \sigma_{t-N-1}^2 \rightarrow 0$  en probabilité lorsque  $N \rightarrow \infty$ .

On a montré que  $\sigma_t^2 - h_t \rightarrow 0$  en probabilité quand  $N \rightarrow \infty$ . Ce terme étant indépendant de  $N$  on a nécessairement  $h_t = \sigma_t^2$  pour tout  $t$  p.s.

Si  $\gamma > 0$ , d'après (ref2.13) et la règle de Cauchy,  $\text{quad} \sum_{n=0}^N a(\eta_{-1} \dots a(\eta_{-n})) \rightarrow \infty$ , p.s. lorsque  $N \rightarrow \infty$ . Donc si  $\omega > 0$ ,  $h_t = \infty$ , p.s. D'après (1.11) il est clair que  $\sigma_t^2 = +\infty$ , p.s. Par suite, il n'existe pas de solution finie p.s. de (1.8).

Dans le cas  $\gamma = 0$ , nous procéderons par l'absurd. Supposons qu'il existe une solution strictement stationnaire  $(\epsilon_t, \sigma_t^2)$  de (1.8). Nous avons pour  $n > 0$ ,

$$\sigma_0^2 \geq \omega \left\{ 1 + \sum_{i=1}^n a(\eta_{-1} \dots a(\eta_{-i})) \right\}$$

d'où on déduit que le terme général  $a(\eta_{-1} \dots a(\eta_{-i})) \omega$  converge vers zéro, p.s, quand  $n \rightarrow \infty$ , ou, de manière équivalente, que

$$\sum_{i=1}^n \log a(\eta_i) + \log \omega \rightarrow -\infty, \text{ p.s. quand } n \rightarrow \infty. \quad (1.14)$$

D'après le théorème de Chung-Fuchs nous avons  $\sum_{i=1}^n \log(\eta_t) = +\infty$  avec probabilité 1, ce qui contredit (1.14).

**Théorème 1.3 (Stationnarité au second ordre du GARCH(1,1)).** *Supposons que  $\omega > 0$ .*

*Si  $\alpha + \beta \geq 0$ , il n'existe pas de solution GARCH(1,1) non anticipative et stationnaire au second ordre.*

*Si  $\alpha + \beta < 1$ , le processus  $\{\epsilon_t, t \in \mathbb{Z}\}$  défini par (1.13), est stationnaire au second ordre. Plus précisément,  $\{\epsilon_t\}$ , est un bruit blanc. De plus,*

il n'existe pas d'autre solution stationnaire au second ordre est non anticipative.

# Chapitre 2

## Modèles Autorégressive Doubles

### 2.1 Introduction

Ce chapitre est consacré à l'étude des modèles autorégressifs double d'ordre  $p$  (DAR( $p$ )). Ces modèles ont été introduits dans la littérature des séries chronologiques par Weiss(1982). Ils ont fait l'objet de travaux de plusieurs auteurs tels que Lu (1998), Ling(2007), Zhu et Ting(2013) etc. Ces modèles Autorégressifs doubles sont considérés comme des modèles de régression qui s'appliquent dans des domaines aussi variée des séries chronologique économique et financières, dans lesquels la série est expliquée par ses valeurs passées plutôt que par d'autres variables. Ces modèles prédit un comportement futur en fonction du comportement passé. Comme tout Modèle de séries chronologiques, ces derniers sont utilisés pour établir des prévisions à court et long terme. L'objectif de ce chapitre est de présenter brièvement les modèles DAR( $p$ ) et d'étudier leurs conditions suffisantes et nécessaires de stationnarité forte et faible. Puis, un intérêt particulier sera accordé a l'estimation des paramètres du modèle où l'accent sera mis sur l'Estimateur du Maximum de Vraisemblance (EMV). Bien évidemment, l'étude de la consistance et de la normalité asymptotique sera présentée.

### 2.2 Stationnarité faible et forte des modèles DAR( $p$ )

le processus  $\{Y_t\}$  est dit modèle autorégressif double d'ordre  $p$ , s'il satisfait les équations de récurrences suivantes:

$$y_t = \sum_{i=1}^p \phi_i y_{t-i} + \eta_t \sqrt{\omega + \sum_{i=1}^p \alpha_i y_{t-i}^2}, \quad t \in \mathbb{N} = \{1, 2, \dots\} \quad (2.1)$$

avec  $\omega, \alpha_0 \geq 0$ ,  $t \in \mathbb{N} \equiv \{-p, \dots, 0, 1, 2, \dots\}$ ,  $\{\eta_t\}$  est une suite de variables aléatoires iid, et  $\eta_t : t \geq 1$  et indépendante de  $\{Y_s : s < t\}$ . Nous allons dans ce qui suit reprendre le travail de Shiqing Ling(2007) pour caractériser les conditions nécessaires et suffisantes de l'existence et l'unicité de solutions strictement stationnaire du modèle (2.1). Aussi nous allons intéresser à l'étude de l'estimateur du maximum de vraisemblance des paramètres de ce modèle, et de montrer qu'il est asymptotiquement gaussien.

Soit  $F_t$  la  $\sigma$ -algèbre générée par:

$$F_t = \{\eta_t, \eta_{t-1}, \dots, \eta_1, y_0, y_{-p}, \quad t \in \mathbb{N}\}.$$

La variance conditionnelle de  $\{y_t\}$  est

$$\text{var}\{y_t/F_{t-1}\} = \omega + \alpha_i y_{t-1}^2.$$

Il est important de remarquer que le modèle (2.1) est un cas particulier d'un ARMA-GARCH introduit par Wiess(1986), mais il est différent des modèles GARCH introduit par Engel(1982). Ces modèles ont fait l'objet de travaux de plusieurs auteurs, ces deux dernières désunies, nous citons: Men Chen Dog Li et Shiqing Ling(2013). Pour l'étude de la stationnarité forte du modèle (2.1), introduisons le vecteur  $\xi = (\xi_{1t}, \dots, \xi_{pt})$  de variables aléatoires gaussiennes indépendantes de  $\{x_t\}$  et soit  $A_t$  la matrice aléatoires de dimension  $p * p$ :

$$\begin{pmatrix} \phi_1 + \sqrt{\alpha_1 \xi_{1t}} & \dots & \phi_{p-1} + \sqrt{\alpha_{p-1} \xi_{p-1,t}} \phi_p + \sqrt{\alpha_p \xi_{pt}} \\ & I_{p-1} & o_{(p-1)*1} \end{pmatrix}$$

où  $I_r$  est la matrice identité d'ordre  $r * r$  et  $o_{r*s}$  est la matrice nulle de dimension  $r * s$ . Soit  $\|M\| = \sqrt{\text{tr}(MM')}$ , pour tout vecteur ou matrice  $M$ , Pour caractériser les conditions de stationnarités, nous considérons

Comme  $\xi_{it} \sim N(0,1)$ , avec  $\mathbb{E} \ln^+ \|A_1\| < \infty$ , où,  $\ln^+(X) = \max\{\ln(X), 0\}$ . Par application du théorème ergodique, il s'en suit que

$$\gamma = \lim_{n \rightarrow \infty} \frac{1}{n} \ln \|A_1 \dots A_n\|. \quad (2.2)$$

presque sûrement (p.s). Il est, en générale difficile de calculer  $\gamma$  à partir des  $\phi_i$  et  $\alpha_i$ , mais il est possible d'estimer  $\gamma$  par simulation pour des valeurs assez élevé de  $n$ .

A présent, nous pouvons énoncer le théorème suivant.

**Théorème 2.2.1 (Shiqing Ling, 2007).** *La condition nécessaire et suffisante de l'existence d'une solution strictement stationnaire du modèle (2.1) est  $\gamma < 0$ . Cette solution est unique, ergodique et  $\mathbb{E} |y_t|^u < \infty$  pour  $u > 0$ .*

**Preuve.** soit  $Y_t = (y_t, y_{t-1}, \dots, y_{t-p-1})'$   $B_p$  la classe de borel dans  $\mathbb{R}^p$ , et soit  $\nu_p$  la mesure de lebesgue dans  $(\mathbb{R}^p, B^p)$ .

Soit  $(\mathbb{R}^p, B^p, \nu_p)$  l'espace des états du processus  $\{y_t\}$ . Soit

$$m : \mathbb{R}^p \rightarrow \mathbb{R}.$$

la première coordonné de la projectoire ,  $m(X) = X_1$  pour  $X \in \mathbb{R}^p$ . Soit  $\{y_t\}$  la chaine de markov avec l'espace des états  $(\mathbb{R}^p, B, \nu_p$  avec une probabilité de transition donnée par:

$$p(X, A) = \int_{m(A)} \frac{1}{\sqrt{\lambda' \tilde{X}}} f \left( \frac{z_1 - \lambda'_1 X}{\sqrt{\lambda'_2 \tilde{X}}} dz_1 \right), \quad X \in \mathbb{R}^p, A \in B, \quad (2.3)$$

où  $X = (X_p, \dots, X_1)$ ,  $\tilde{X} = (1, X'_p, \dots, X_1^2)$  et  $f(X) = (2\pi i)^{-0.5} \exp^{-\frac{X^2}{2}}$ . Premièrement, nous montrons que  $\nu_p$  est irréductible de la chaine de markov  $\{y_t\}$  Soit  $\sum_{n=1}^p P^n(X, A) > 0$  pour chaque  $\nu_p(A) > 0$ , quand

$$P^n(X, A) = \int_{\mathbb{R}^p} P^{n-1}(y, A) P(X, dy), \quad X \in \mathbb{R}^p, A \in B^p,$$

Il est facile de voir que la probabilité de transition de la chaine de markov  $\{y_t\}$  a p étapes est:

$$P^p(X, A) = \int_A \prod_{i=1}^p \frac{1}{\sqrt{\lambda'_2 \tilde{X}_i}} \left\{ \left( \frac{z_i - \lambda'_1 X_i}{\sqrt{\lambda'_2 \tilde{X}_i}} \right) dz_1 \dots dz_p, \right. \quad (2.4)$$

où  $X_i = (z_i, \dots, z_1, X_1, \dots, X_{p-i})$  et

$$\tilde{X}_i = (1, z_i^2, \dots, z_1^2, X_1^2, \dots, X_{p-i}^2).$$

D'après la densité positif de transition de kernel dans (2.4), on sais que  $\{y_t\}$  et  $\nu_p$ -irréductible. D'après (2.2), il existe un entier s tel que

$$\mathbb{E}(\ln \| A_1, \dots, A_s \|) < 0.$$

Soit  $\tilde{A}_t = A_t \dots A_{t-s+1}$ , et écrire  $q(u) = \mathbb{E} \| \tilde{A}_t \| ^u$ . Quand  $\eta_t \sim \mathcal{N}(0,1)$ , il est evident de voir que  $q(u)$  est differentiable sur  $[0,2)$  et

$$q'(u) = \mathbb{E}[\| \tilde{A}_t \| ^u \ln \| \tilde{A}_t \|].$$

On note que

$$| \ln(X^\delta) | \leq \max\{X^\delta, X^{-\delta}\} - 1$$

pour toutes  $X > 0$  et  $\delta \in [0, \frac{u}{2}]$ . nous pouvons montrer que

$$\mathbb{E}(\sup_{u \in [0,1]} [\| \tilde{A}_t \|^u \ln \| \tilde{A}_t \|]) < \infty.$$

Par le théorème de convergence dominé,

$$\lim_{u \rightarrow 0} q'(u) = \mathbb{E}(\ln \| \tilde{A}_t \|) < 0.$$

Ainsi, il existe une constante  $\tilde{\delta} \in (0,2)$  tel que  $q(u)$  est strictement décroissant sur  $[0, \tilde{\delta}]$ . Par conséquent, il existe une constante  $u \in (0,1)$  tel que:

$$\mathbb{E}(\| \tilde{A}_t \|^u) < q(0) = 1. \quad (2.5)$$

On utilisant (2.5), nous prouvons ensuite montrer que le s-étapes de la chain de markov  $\{y_{ts}\}$  satisfait la condition dérivé donner dans théorème de Tweedie(1983), il existe un espace des compacts noté par  $K$  et la fonction continue non-négatif  $g(X)$  tel que  $\nu_p(K) > 0, g(X) \leq 1$  sur  $K$ , et:

$$\mathbb{E}(g(y_{st}/y_{(t-1)s} = X) \leq (1 - \epsilon)g(X)), \quad X \in K^c, \quad (2.6)$$

$$\mathbb{E}(g(y_{st}/y_{(t-1)s} = X) \leq M, \quad X \in K, \quad (2.7)$$

pour tout  $\epsilon > 0$ . Notre but est de trouver la fonction de  $g$  tel que (2.6)(2.7) est vérifiées. Il est difficile d'obtenir la fonction  $g$  par une méthode directe. considérons tout d'abord, le modèle RCAR(p) défini par:

$$\tilde{y}_t = A_t \tilde{y}_{t-1} + \tilde{\eta}_t, \quad (2.8)$$

où  $\tilde{\eta}_t = (\sqrt{\omega} \eta_t, 0, \dots, 0)'$  et  $\tilde{y}_t$  est indépendant de  $\{\eta_{t'} : t' < t\}$ . Il est facile de voir que  $\tilde{y}_t$  est une chaîne de markov homogène, avec  $(\mathbb{R}^p, \mathcal{B}^p, \nu_p)$  l'espace d'état, et sa probabilité de transition est

$$P(X, A) = \int_{m(A)} \frac{1}{\sqrt{\lambda'_2 \tilde{X}}} \left( \frac{z_1 - \lambda'_1 X}{\sqrt{\lambda'_2 \tilde{X}}} dz_1 \right), \quad X \in \mathbb{R}^p, A \in \mathcal{B}^p. \quad (2.9)$$

On pose  $g = 1 + \| X \|^u$ , où  $X \in \mathbb{R}^p$  et  $u$  est défini comme dans (2.5.) Pour un  $s$  fixé tel que (2.5) vérifié, nous itérons (2.8) pour obtenir l'expression suivant:

$$\tilde{y}_{ts} = \left( \tilde{\eta}_{ts} + \sum_{j=1}^{s-1} \prod_{r=0}^{j-1} A_{ts-r} \tilde{\eta}_{ts-j} + \tilde{A}_{ts} y_{(t-s)s} \right). \quad (2.10)$$

De (2.10) nous avons:

$$\begin{aligned} \mathbb{E}(g\tilde{y}_{ts}/\tilde{y}_{(t-1)s} = X) &\leq 1 + \mathbb{E}(\|\tilde{\eta}_{ts} + \sum_{j=1}^{s-1} \prod_{r=0}^{j-1} A_{ts-r}\tilde{\eta}_{ts-j}\|^u) + \mathbb{E}\|\tilde{A}_{ts}\|^u\|X\|^u \\ &= \mathbb{E}\|\tilde{A}_{ts}\|^u\|X\|^u + \mathcal{C}, \end{aligned} \quad (2.11)$$

où  $\mathcal{C}$  est une constate. Soit

$$K = \{X : \|X\| \leq L\}$$

et  $L$  est une constate positifs. Il est facile de voire que

$$\mathbb{E}(g(\tilde{y}_{st})/\tilde{y}_{(t-1)s} = X \leq M), \quad X \in K, \quad (2.12)$$

pour une constante  $M$ . Notons que  $\mathbb{E}(\|\tilde{A}_{ts}\|^u) = \mathbb{E}(\|\tilde{A}_t\|^u)$ . Comme  $L$  est largement suffisant, et  $X \in K^c$ , par l'equation (2.5) il existe  $\epsilon > 0$  tel que:

$$\begin{aligned} \mathbb{E}(g(\tilde{y}_{st})/\tilde{y}_{(t-1)s} = X) &= g(X) + (\mathbb{E}\|\tilde{A}_{ts}\|^u - 1)\|X\|^u + \mathcal{C} - 1 \\ &\leq g(X)\{1 - [(1 - \mathbb{E}\|\tilde{A}_{ts}\|^u) + \frac{\mathcal{C}-1}{1+\|X\|^u}]\} \\ &\leq (1 - \epsilon)g(X) \end{aligned} \quad (2.13)$$

De (2.3), (2.9), nous savons que  $\{y_t\}$  et  $\{\tilde{y}_t\}$  qu'ils ont la même densité de transition.

par (2.12)-(2.13), nous savons que (2.6)-(2.7), vérifiées avec la même fonction  $g(X)$  et  $K$ . Pour chaque fonction bornée continue  $G$  dans  $\mathbb{R}^p$ ,  $\mathbb{E}[G(y_{st})/y_{s(t-1)} = X]$  est continu en  $X$ . Ainsi,  $\{y_{st}\}$  la chaine de Feller. En autre, on a  $\{y_{st}\}$  est  $\nu_p$ -irréductible, d'après Feigne et Tweedie (1985), nous savons que  $\{y_{st}\}$  est géométriquement ergodique, qui garantit qu'il existe une unique distribution stationnaire  $\pi$  pour  $\{y_{st}\}$ , et

$$\int \|y_{st}\|^u \leq_1 g(X)\pi(dX) < \infty. \quad (2.14)$$

D'après Tjøstheim (1990),  $y_t$  est géométriquement ergodique. Ainsi,  $y_t$  à une unique distribution  $\pi$ . Soit  $y_0$  la valeur initial de la distribution stationnaire  $\pi$ .  $\{y_t : t \in \mathbb{N}\}$  est l'unique distribution stationnaire du modèle (2.1) et il est géométriquement ergodique. En autre, de (2.14),  $\mathbb{E}\|y_t\|^u < \infty$ . Dans ce qui suit nous démontrons la nécessité. supposons que  $\{y_t : t \in \mathbb{N}\}$  est la solution strictement stationnaire de modèle (2.1), alors,  $\pi$  est la distribution stationnaire de la chaine de markov  $\{y_t\}$ .

puisque  $\{y_t\}$  et  $\{\tilde{y}_t\}$  ont la même densité de transition de Kernel, la chaine de markov  $\{\tilde{y}_t\}$  à la distribution stationnaire  $\pi$ . soit  $y_0$  la valeur initial de la distribution  $\pi$ , et  $\{\tilde{y}_t, t \in \mathbb{N}\}$  est la solution stationnaire du modèle (2.8). De l'equations (2.10) avec  $s=p$ , nous avons:

$$\tilde{y}_{tp} = \mathbf{B}_{tp} + \tilde{A}_{tp}\tilde{y}_{t-1p}, \quad (2.15)$$

où

$$\mathbf{B}_{tp} = \tilde{\eta}_{tp} + \sum_{j=1}^{p-1} \prod_{r=0}^{j-1} A_{tp-r} \tilde{\eta}_{tp-j}.$$

Notez que  $\tilde{A}_t = A_t \dots A_{t-p+1}$ . Puisque  $\xi$  est normal et  $\ln(1+x) \leq x$  pour  $x \geq 0$ , on aura

$$\begin{aligned} \mathbb{E}(\ln^+ \|\tilde{A}_{tp}\|) &= \mathbb{E}(\max\{\ln \|\tilde{A}_{tp}\|, 0\}) \\ &= \mathbb{E}[I\{\|\tilde{A}_{tp}\| > 1\} \ln \|\tilde{A}_{tp}\|] \\ &\leq \mathbb{E}\|\tilde{A}_{tp}\| + 1 \\ &< \infty \end{aligned}$$

De la même manière nous pouvons montrer que :

$$\mathbb{E}(\ln^+ \|\mathbf{B}_{tp}\|) < \infty.$$

D'après l'hypothèse du modèle (2.8),  $\{\tilde{Y}_{tp} : t \in \mathbb{N}\}$  est une solution non-anticipatif à (2.15), puisque  $\{\tilde{y}_{tp} : t \in \mathbb{N}\}$  est strictement stationnaire, il s'ensuit que

$$P(\tilde{y}_{tp} \in A/\tilde{y}_{(t-1)p} = x) = P(\tilde{y}_p \in A/\tilde{y}_0 = x) = P^p(x, A). \quad (2.16)$$

De l'équation (2.4) et (2.16), la densité de  $\{\tilde{y}_{tp}\}$  donne  $\{\tilde{y}_{(t-1)p} = x\}$  est positive. Soit  $H$  tout sous-espace affine invariant de  $\mathbb{R}^p$  sous le modèle (2.15) ( $B_{tp} + \tilde{A}_{tp}x : x \in \subseteq H(P.S.)$ ). Si  $\nu_p(\mathbb{R}^p - H) \neq 0$ , Alors, pour tout  $x \in H$ ,

$$P(B_{tp} + \tilde{A}_{tp}x \in H) = P(\tilde{y}_{tp} \in H/\tilde{y}_{(t-1)p} = x) = P(\tilde{y}_{tp} \in \mathbb{R}^p/\tilde{y}_{(t-1)p} = x) - P(\tilde{y}_{tp} \in \mathbb{R}^p - H/\tilde{y}_{(t-1)p} = x) < 1. \quad (2.17)$$

Il est évident que  $\mathbb{R}^p$  est un sous espace invariant affine. De l'équation (2.17) le sous-espace invariante est unique. Ainsi que le modèle (2.5) est irréductible. D'après Bougerol et Picard(1992), la condition nécessaire pour la solution strictement stationnaire non anticipatif  $\tilde{y}_{tp}$  de modèle (2.15) est que l'exposant de Lyapounov supérieur soit :

$$\tilde{\gamma} = \inf\left\{\frac{1}{n}\mathbb{E}(\ln \|\tilde{A}_p \tilde{A}_{2p} \dots \tilde{A}_{tp}\|), \quad t \geq 1\right\} < 0. \quad (2.18)$$

Notons que  $\tilde{A}_p \tilde{A}_{2p} \dots \tilde{A}_{tp} = A_1 \dots A_{tp}$ . D'après (2.18), il existe un  $s$  sachant que  $\mathbb{E}(\ln \|\tilde{A}_1 \dots \tilde{A}_{sp}\|) < 0$ . On note  $\tilde{A}_t^* = A_t \dots A_{t-sp+1}$ . Puis,

$$\mathbb{E}(\ln \|\tilde{A}_{st}^*\|) = \mathbb{E}(\ln \|\tilde{A}_1 \dots \tilde{A}_{sp}\|) < 0.$$

pour ce qui est de (2.5), il existe  $u \in (0,1)$  sachant que

$$\mathbb{E}(\|\tilde{A}_{sp}^*\|^u) < 1.$$

Soit  $n = msp + r$ , où  $r = 0, \dots, sp - 1$ . Nous avons

$$A_1 \dots A_n = \tilde{A}_{sp}^* \tilde{A}_{2sp}^* \dots \tilde{A}_{msp}^* \tilde{A}_{msp+1}^* \dots \tilde{A}_{msp+r}^*,$$

où  $\tilde{A}_{msp+1}^* \dots \tilde{A}_{msp+r}^* = 1$  quand  $r = 0$ . Puisque  $\{A_i\}$  est une suit des matrices aléatoires (iid), nous avons:

$$\begin{aligned} \mathbb{E} \| A_1 \dots A_n \|^u &\leq \mathbb{E} \| \tilde{A}_{sp}^* \tilde{A}_{2sp}^* \dots \tilde{A}_{msp+1}^* \|^u \mathbb{E} (\| \tilde{A}_{msp+1}^* \dots \tilde{A}_{msp+r}^* \|^u) \\ &\leq [\mathbb{E} \| \tilde{A}_{sp}^* \|^u]^m [\mathbb{E} \| A_1 \|^u]^r \\ &= O([\mathbb{E} \| \tilde{A}_{sp}^* \|^u]^m) \\ &\rightarrow 0. \end{aligned}$$

Quand  $n \rightarrow \infty$ , qui implique que  $m \rightarrow \infty$ . Ainsi, il existe un  $n$  tel que  $\mathbb{E}(\ln \| A_1 \dots A_n \|) \leq u^{-1} \ln \mathbb{E}(\| A_1 \dots A_n \|^u) < 0$ , Par conséquent  $\gamma < 0$ .

Soit  $\rho(A)$  le module de la matrice A ( $\rho(A) = \max | \lambda_i |$ ) où  $\lambda_i$  sont les valeurs propres de A. Nous donnons dans le théorème qui suit la condition d'existence du moment d'ordre 2 de  $y_t$ .

**Théorème 2.2.2 (Ling 2007).** *La condition nécessaire et suffisante d'existence d'une solution strictement stationnaire  $\{y_t, t \in \mathbb{N}\}$  du modèle (2.1) tel que  $\mathbb{E}\{y_t^2\} < \infty$ , est:*

$$\rho \left[ \mathbb{E} \left( A_t \otimes A_t \right) \right] < 1,$$

où  $\otimes$  est le produit Kronecker des matrices. La solution est unique et ergodique.

**Preuve (Ling 2007).** Soit  $\{y_t\}$  et  $\tilde{y}_t$  définies comme précédemment dans la preuve de théorème (2.2). Puisque  $\{y_t\}$  et  $\tilde{y}_t$  ont la même densité de transition, il suffit de montrer que  $\rho[\mathbb{E}(A_t \otimes A_t)] < 1$  est nécessaire et suffisant pour l'existence et l'unicité d'une solution strictement stationnaire  $\tilde{y}_t$  à l'équation (2.8) avec  $\mathbb{E}(\| \tilde{y}_t \|^2) < \infty$ . la suffisance et l'unicité ont été données par Feigin et Tweedie(1985), alors que la nécessité est donnée par Nicholls et Quinn(1982)et Ling(1999).

**Remarque 2.1.** Lorsque  $\rho = 1, \gamma = \mathbb{E}\{\log | \phi_1 + \sqrt{\alpha_1} \xi_t |\}$ . Dans ce cas, la condition du théorème (2.2) se réduit à:

$$\phi_1^2 + \alpha_1 < 1.$$

La figure 1 donne les régions de stationnarité de  $(\phi_1 + \alpha_1)$  telles que  $\phi_1^2 + \alpha_1 < 1$  est  $\mathbb{E}\{\log | \phi_1 + \sqrt{\alpha_1} \xi_t |\} < 0$ .

**Remarque 2.2.** Le modèle(2.1) est un cas particulier du modèle générale:

$$y_t = \{(y_{t-1}, \dots, y_{t-p}) + \eta_t \sqrt{h(y_{t-1}, \dots, y_{t-p})},$$

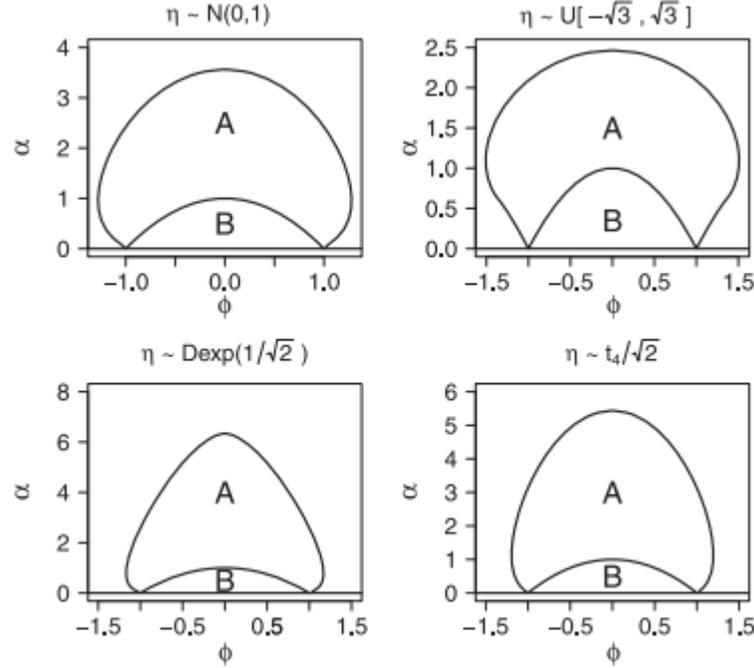


FIG. 2.1 – les régions de stationnarité et de non-stationnarité du modèles (2.1), quand  $\eta$  suit respectivement:  $\mathcal{N}(0,1)$ ,  $U[-\sqrt{3},\sqrt{3}]$ , et la distribution exp-double de paramètre  $\frac{1}{2}$

où  $h(y_{t-1}, \dots, y_{t-p}) > 0$  et  $\eta_t$  est une suite iid de variables aléatoires centré de variance 1, étudié par Ango Nze (1992). En appliquant leurs résultat au modèles (2.1), on obtient la condition suffisante de stationnarité faible qui est:

$$\sum_{i=1}^p |\phi_i| + \mathbb{E} |\eta_t| \sum_{i=1}^p \alpha_i < 1,$$

où  $\mathbb{E}(|\eta_t|) = \sqrt{\frac{2}{\pi}}$  quand  $\eta_t$  est normal, cette condition est aussi suffisante si  $\mathbb{E}\{|(y_t)|\} < \infty$ . La condition suffisante d'existence de module d'ordre 2 ( $\mathbb{E}(y_t^2) < +\infty$ ) est donc:

$$\left( \sum_{i=1}^p |\phi_i| \right)^2 + \sum_{i=1}^p \alpha_i < 1.$$

### 2.3 Non-stationnarité du modèle AR(1) Double

Soit  $\eta$  une variable aléatoire de même distribution que  $\eta_t$ . Borkovec Kluppelberge (2001) BK(2001) à prouvé que la condition suffisante de la stationnarité strict du modèle

(2.1), est:

$$\mathbb{E}(\log |\phi_0 + \eta\sqrt{\alpha_0}|) < 0,$$

et ce, lorsque  $\eta$  symétrique et de densité continue et strictement positive sur  $\mathbb{R}$  avec  $\mathbb{E}(\eta^2) < \infty$ . Ici, l'objectif, est de prouver que sous la condition  $\mathbb{E}[\log |\phi_0 + \eta\sqrt{\alpha_0}|] > 0$ , de modèle (2.1) est explosif pour des innovations symétriques, ce qui entraîne qu'il est non-stationnaire.

A présent, introduisons le processus  $\{X_t\}$  caractérisé par les équations de récurrences suivantes:

$$X_t = |\phi_0 X_{t-1} + \eta_t \sqrt{\gamma_0 + \alpha_0 x_{t-1}^2}|, \quad (2.19)$$

avec les même notation que celle de (2.1), La valeur initiale  $x_0 = |Y_0|$ , p.s presque sûrement, et ce en raison de la symétrie de  $\{X_t\}$ , l'indépendance de  $\eta_t, Y_{t-1}$  dans (2.1) et la structure markovienne de  $\{x_t\}$  et  $\{Y_t\}$ , il est aussi évident que  $\{x_t, t \in \mathbb{N}\} = \{|Y_t|, t \in \mathbb{N}\}$  (voire: Borkovec(2000),BK(2001),et autres).

Soit

$$\omega_t = |\phi_0 x_{t-1} + \eta_t \sqrt{\gamma_0 + \alpha_0 x_{t-1}^2}| - |\phi_0 + \eta_t \sqrt{\alpha_0}| x_{t-1}.$$

D'après (??), il s'ensuit que pour  $t \geq 1$ . on a :

$$x_t = |\phi_0 + \eta_t \sqrt{\alpha_0}| x_{t-1} + \omega_t = \dots$$

$$= \sum_{j=0}^{t-1} \left( \prod_{i=0}^{j-1} |\phi_0 + \eta_{t-i} \sqrt{\alpha_0}| \right) \omega_{t-j} + \left( \prod_{i=0}^{t-1} |\phi_0 + \eta_{t-i} \sqrt{\alpha_0}| \right) x_0$$

avec la convention  $\prod_{i=0}^{-1} = 1$ . Ainsi,

$$\left( \prod_{k=1}^t |\phi_0 + \eta_k \sqrt{\alpha_0}| \right)^{-1} x_t = x_0 + \sum_{j=1}^t \left( \prod_{k=1}^j |\phi_0 + \eta_k \sqrt{\alpha_0}| \right)^{-1} \omega_j.$$

Notons que  $|\omega_t| \leq \sqrt{\gamma_0} |\eta_t|$ . Si  $\mathbb{E} \log |\phi_0 + \eta\sqrt{\alpha_0}| > 0$ , alors nous pouvons montrer que

$$\begin{aligned} \left( \prod_{k=1}^t |\phi_0 + \eta_k \sqrt{\alpha_0}| \right)^{-1} x_t &\rightarrow x_0 + \sum_{j=1}^{\infty} \left( \prod_{k=1}^j |\phi_0 + \eta_k \sqrt{\alpha_0}| \right)^{-1} \omega_j \\ &:= x_0 + W \end{aligned}$$

presque sûrement. quand  $t \rightarrow \infty$ , on note que

$$|W| \leq \sqrt{\gamma_0} \sum_{j=1}^{\infty} \left( \prod_{k=1}^j |\phi_0 + \eta_k \sqrt{\alpha_0}| \right)^{-1} |\eta_j| < \infty$$

presque sûrement.

**Théorème 2.3.1.** *Si  $\{Y_t\}$  est une suite de variables aléatoires symétrique et iid, avec  $\mathbb{E}(\log |\phi_0 + \eta \sqrt{\alpha_0}|) > 0$ , et  $\mathbb{P}(|Y_0| + W = 0) = 0$ , Alors  $\rho^{-1} |Y_t| \rightarrow \infty$  presque sûrement, pour tous*

$$1 < \rho < \exp\{\mathbb{E}(\log |\phi_0 + \eta \sqrt{\alpha_0}|)\},$$

presque sûrement quand  $t \rightarrow \infty$ .

Contrairement à BK(2001) dans lequel  $\eta$  a une densité positive et continue dans  $\mathbb{R}$  avec  $\mathbb{E}\eta^2 < \infty$ , cette condition est faible et en particulier permet de ne pas avoir du moment fractionnelle. Dans ce cas, l'unique condition de  $\eta$  est la symétrie. Cette condition est plutôt faible, et satisfaite par les variables les plus utilisées, par exemple: (Normal, exponentielle double, Student's, et autres...). La figure 1 montre les régions de la stationnarité et de la non-stationnarité du processus DAR(1) quand  $\eta$  suit une loi de distribution exponentielle double (Dexp) avec le paramètre  $\frac{1}{\sqrt{2}}$  et la loi de student standard  $t_4$  respectivement, il est clair que la région de stationnarité de modèle AR(1) est  $|\phi_0| < 1$ .

D'après la figure 1, on peut voir que le modèle 1 peut être stationnaire même si  $|\phi_0| \geq 1$ . Cependant, dans le théorème 1,  $\eta$  est une variable aléatoire discrète, ceci montre la grande différence entre la stationnarité étudiée dans Borkovec(2000), BK(2001) et autres dans la littérature. Dans le théorème 1, on fixe  $Y_0$  ou  $x_0$  comme une valeur initiale sachant que  $\mathbb{P}(|Y_0| + W = 0) = 0$ . Berkes et Al ont utilisé les mêmes conditions. En (2009) montre quand ils ont établi le coefficient aléatoire du modèle autorégressif de la croissance de la non-stationnarité Borkovec (2000) lui aussi a utilisé l'hypothèse que  $Y_0$  largement suffisante quand il a étudié le comportement du modèle DAR(1). D'après l'expression de  $W$ , il semble très difficile d'obtenir la forme fermée de la distribution de  $W$  puisqu'il dépend fortement du processus  $\{x_t\}$  lui-même. Cependant, on peut simuler la distribution ou la densité de  $|Y_0| + W$ .

Figure 3, montre la fonction de densité de  $|Y_0| + W$  quand

$$\sim N(0.1), U[-\sqrt{3}, \sqrt{3}]$$

Dexp( $\frac{1}{2}$ ), student  $t_4$ -distribution, la distribution Rdemacher et la distribution discrète et la (FPDD), définie en (??) respectivement.

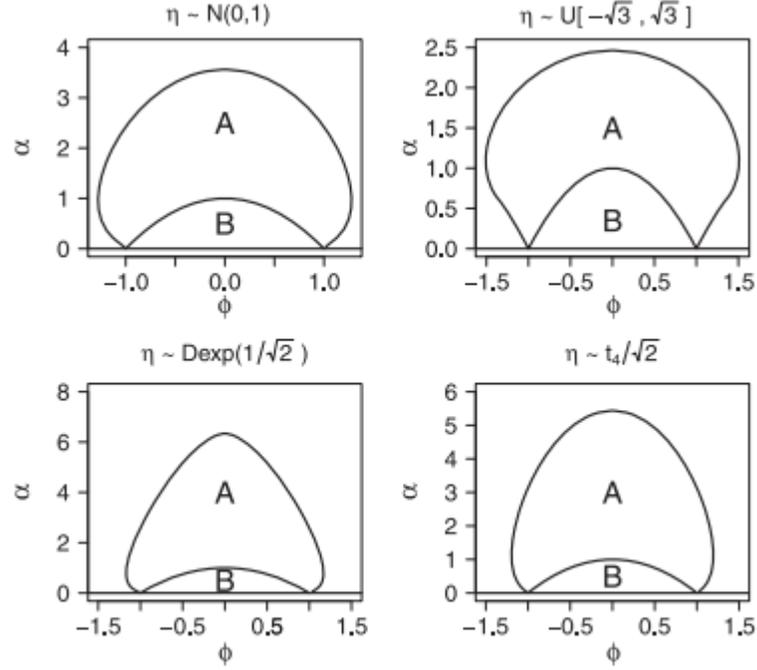


FIG. 2.2 – les régions de stationnarité et de non-stationnarité du modèles (2.1), quand  $\eta$  suit respectivement:  $\mathcal{N}(0,1)$ ,  $U[-\sqrt{3},\sqrt{3}]$ , et la distribution exp-double de paramètre  $\frac{1}{2}$

Dans ce cas on prend les valeurs initiale  $(\phi_0, \alpha_0, \gamma_0) = (2.1.1)$  et la valeur initiale  $Y_0 = 2$ . Premièrement on utilise le 100 term, ainsi il semble que  $\mathbb{P}(|Y_0| + W = 0) = 0$  qui n'a pas de solution. Il sera claire que, pour la cas de  $\mathbb{E} \log |\phi_0 + \eta \sqrt{\alpha_0}| = 0$  nous constate que le modèle DAR(1) reste non-stationnaire et son comportement asymptotique reste compliqué.

## 2.4 L'Estimation du Maximum de Vraisemblance des paramètres du modèles DAR(1)(EMV)

En statistique, l'une des méthodes les plus utilisées pour l'estimation de paramètres est celle du maximum de vraisemblance.

C'est Ronald Fisher qui, en (1912), a publié sa méthode initialement appelée « Le critère absolu ». À l'époque, l'une des méthodes couramment utilisées était celle des moindres carrés.

Soit l'échantillon  $\{y_0, y_1, \dots, Y_n\}$  du modèle 1. on note par  $(\gamma, \Theta)' = (\gamma, \alpha, \theta)'$  les paramètres du modèle DAR(1) et on défine la maximum de vraisemblance (QMC)  $(\hat{\gamma}_n, \hat{\theta}_n)'$  comme

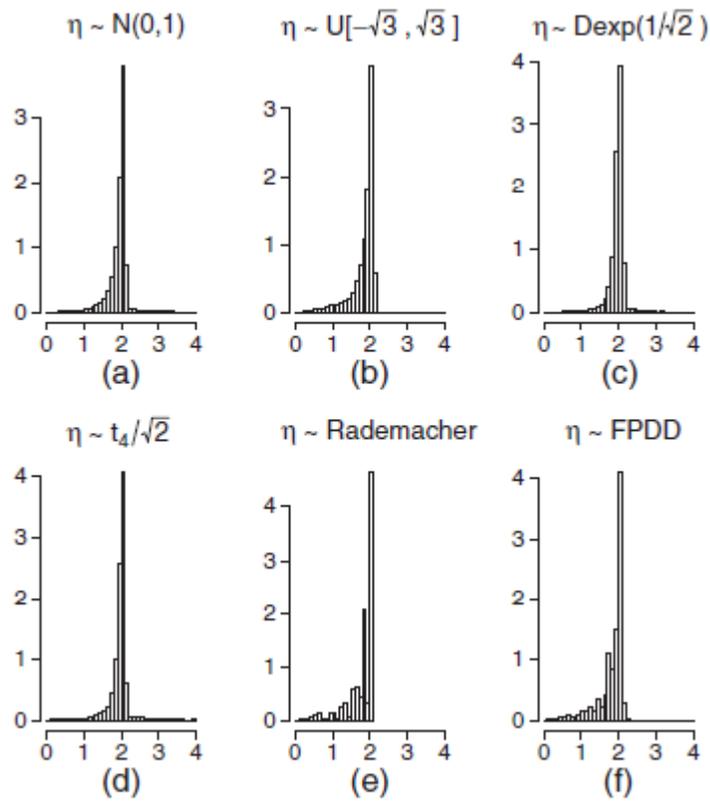


FIG. 2.3 – les regions de stationnarité et de non-stationnarité du modèles (2.1), quand  $\eta$  suit respectivement:  $\mathcal{N}(0,1), U[-\sqrt{3}, \sqrt{3}]$ , et la distribution exp-double de paramètre  $\frac{1}{2}$

une solution mesurable de

$$L_n(\gamma, \theta) = \sum_{t=1}^n l_t(\gamma, \theta),$$

avec

$$\{l_t(\gamma, \theta) = -\frac{1}{2} \left\{ \log(\gamma + \alpha Y_{t-1}^2) + \frac{(Y_t - \theta Y_{t-1})^2}{\gamma + \alpha Y_{t-1}^2} \right\},$$

qui est:

$$\{(\hat{\gamma}_n, \hat{\theta}_n)' = \arg \max_{\Gamma * \Theta} \{L_n(\gamma, \theta)\}$$

où  $\Gamma * \Theta$  le compact défini de  $\mathbb{R}_+ * \mathbb{R} * \mathbb{R}_+$  avec  $\mathbb{R} \equiv (0, +\infty)$  qui contient les valeurs réelles

$$(\gamma_0, \theta_0)' = (\gamma_0, \theta_0, \alpha_0)'.$$

Dans le cas de  $\alpha_0 = 0$ , qui ont étudié les propriétés asymptotiques de  $(\hat{\gamma}_n, \hat{\theta}_n)'$ . Quand  $\{Y_t\}$  est explosif,  $\gamma$  n'aura pas de sens dans la limite de  $L_n(\gamma, \theta)/n$ ; voire le lemme 2. Le même phénomène est observée dans Jensen et Rahbek (2004a) quand ils ont étudié la non-stationnarité du modèle ARCH(1), Ling et Li (2008) quand  $\eta$  suit une lois normal, voire Francq et Zakaoin(2012). Ainsi que, l'estimateur consistant de  $\gamma_0$  n'existe pas, ce point est conclure d'après la preuve du théorème 2 que le log-vraisemblance est complètement plat dans la direction où ona fixé  $(\theta, \alpha)'$  et  $\gamma_0$  varie. plus précisément, pour tout séquence réelle  $\lambda_n$  allant de zéro a  $n \rightarrow \infty$ , on à :

$$\bigwedge_n \left( \begin{array}{c} \frac{\partial L_n(\gamma, \theta_0)}{\partial \gamma} \\ \frac{\partial L_n(\gamma, \theta_0)}{\partial \theta} \end{array} \right) \rightarrow^{\mathcal{D}} \mathcal{N}\left(0, \text{diag}\left\{0, \frac{1}{\alpha_0}, \frac{\mathbb{E}\eta_t^4 - 1}{4\alpha_0^2}\right\}\right),$$

avec  $\bigwedge_n = \text{diag}\{\lambda_n, \frac{1}{\sqrt{n}}, \frac{1}{\sqrt{n}}\}$ . pour plus d'information sur le comportement asymptotique de  $\hat{\gamma}_n$  vous pouvez consulter Francq et Zakoan(2012). L'incohérence de  $\hat{\gamma}_n$  est illustrée par des études de simulation dans la section qui suit. Pour des raison pratiques, nous donnons d'abord des hypothèses sur  $\{\eta_t\}$  et l'espace des paramètre:

1.  $\{\eta_t\}$  est une suite de variables aléatoires iid symétrique avec la moyenne null et de variance 1.
2. Le paramètre réel  $\theta_0 \in \Theta$  satisfait  $\mathbb{E}(|\phi_0 + \eta\sqrt{\alpha_0}|) > 0$  et  $\mathbb{P}(|Y_0| + \omega = 0) = 0$ ,  $\Gamma$  et  $\Theta$  est compact.

Afin d'obtenir des propriétés asymptotiques de  $\hat{\theta}_n$ , le taux de divergence de  $|Y_t|$  à l'infini est nécessaire, ce qui est donner par le théorème 1. Le théorème 2 suivant indique la consistance et la normalité asymptotique de  $\hat{\theta}_n$ .

**Théorème 2.1.** *Si les hypothèses 1 et 2 sont vérifiées, alors, pour  $n \rightarrow \infty$*

1.  $\hat{\theta}_n \rightarrow \theta_0$  en probabilité;

2. If  $K = \mathbb{E}[\eta^4] < \infty$ , et  $\theta_0$  est un point intérieur dans l'ensemble des paramètres  $\Theta$ , alors,

$$\sqrt{n} \left( \hat{\theta}_n - \theta_0 \right) \rightarrow \mathcal{N}(0, \Sigma),$$

$$\text{quand } \Sigma = \text{diag}\{\alpha_0, (K-1)\alpha_0^2\}.$$

En combinant avec les résultats dans Ling(2004) pour le cas stationnaire, la normalité asymptotique de l'estimateur de quasi-maximisation de vraisemblance (QMV) de  $(\phi_0, \alpha_0)$  s'obtient avec le taux de convergence pour n'importe quel

$$\phi_0 \in \mathbb{R}, \alpha_0 \in \mathbb{R}_+$$

avec

$$\mathbb{E} \log |\phi_0 + \eta \sqrt{\alpha_0}| \neq 0.$$

En effet, dans le cas stationnaire, l'estimateur de  $\gamma_0$  est consistant et asymptotiquement normal (Ling(2004)). Cependant, dans le cas explosif, on ne peut pas estimer  $\gamma_0$  en plus, on ne peut pas obtenir l'estimateur consistant, mais la propriété asymptotique de (QMV) de  $(\phi_0, \alpha_0)$  reste telle qu'elle est. Ling et Li(2008) ont présenté un estimateur asymptotiquement sans biais pour  $\gamma_0$ , mais qui n'est pas consistant. Contrairement au résultat des modèles classiques AR(1), le problème de racines unitaires n'existe pas dans le modèle (2.1) (voir Rmks 1-3 dans Ling et Li (2008)). Afin de construire des intervalles de confiance pour  $\theta_0$ , on a besoin d'estimer  $K$ . Enfin, considérons le résidu  $\{\hat{\eta}_t\}$  où

$$\hat{\eta}_t = \frac{Y_t - \hat{\phi}_n Y_{t-1}}{\sqrt{\hat{\gamma}_n + \hat{\alpha}_n Y_{t-1}^2}},$$

et nous définissons l'estimateur de  $K$  qui est

$$\hat{K} = \frac{1}{n} \sum_{t=1}^n \hat{\eta}_t^4.$$

Sous les hypothèses 1 et 2, nous pouvons montrer que: Si  $K < \infty$ , Alors,  $\hat{K}_n = K + o_p(1)$ . C'est-à-dire que  $\hat{K}_n$  est un estimateur consistant de  $K$ .

Avant de donner la démonstration du théorème, le lemme suivant est nécessaire:

**Lemme 2.1.** *Si les hypothèses 1 et 2 sont vérifiées, alors:*

1.

$$\sup_{\theta \in \Theta} \sup_{\gamma \in \Gamma} \left\| \frac{1}{n} \sum_{t=1}^n \frac{\partial \ell_t(\gamma, \theta)}{\partial \theta} \right\| = o_p(1)$$

2.

$$\sup_{\theta \in \Theta} \sup_{\gamma \in \Gamma} \left| \frac{1}{n} \sum_{t=1}^n \frac{\partial^3 \ell_t(\gamma, \theta)}{\partial \theta_i \partial \theta_j \partial \theta_k} \right| = o_p(1)$$

avec

$$\theta_1 = \phi, \theta_2 = \alpha, i, j, k = 1, 2.$$

**Preuve.** Ona l'espace des paramètres  $\Gamma * \Theta$  est compact, donc il existe des constantes  $0 < \underline{\alpha} < \bar{\alpha} < \infty$ ,  $0 < \underline{\gamma} < \bar{\gamma} < \infty$  et  $0 < M < \infty$  sachant que  $|\phi| \leq M, \gamma \in [\underline{\gamma}, \bar{\gamma}]$  et  $\alpha \in [\underline{\alpha}, \bar{\alpha}]$ . D'après l'expression de  $\ell_t(\gamma, \theta)$ , on a

$$\frac{\partial \ell_t(\gamma, \theta)}{\partial \phi} = \frac{Y_{t-1}(Y_t - \phi Y_{t-1})}{\gamma + \alpha Y_{t-1}^2}.$$

d'après le théorème 1 et la lois des grands nombres , on aura :

$$\sup_{\theta \in \Theta} \sup_{\gamma \in \Gamma} \left| \frac{1}{n} \sum_{t=1}^n \frac{\partial \ell_t(\gamma, \theta)}{\partial \phi} \right| < \frac{2M}{\underline{\alpha}} + \frac{1}{\sqrt{\underline{\alpha}}} \sqrt{\frac{\max\{\gamma_0, \alpha_0\}}{\min\{\underline{\gamma}, \underline{\alpha}\}}} \frac{1}{n} \sum_{t=1}^n |\eta_t| = o_p(1).$$

De même, nous pouvons montrer que

$$\sup_{\theta \in \Theta} \sup_{\gamma \in \Gamma} \left| \frac{1}{n} \frac{\partial L_n(\gamma, \theta)}{\partial \alpha} \right| = o_p(1).$$

Avec  $o_p(1)$  converge en probabilité vers 0. De plus, l'hypothèse 1 est vérifiée, reste à verifer la deuxième hypothèses. 2 Considérons  $\partial^3 \ell_t(\gamma, \theta) / \partial \alpha$ . Après les calculs, on a:

$$\frac{\partial^3 \ell_t(\gamma, \theta)}{\partial \alpha^3} = \frac{3Y_{t-1}^6 (Y_t - \phi Y_{t-1})^2}{(\gamma + \alpha y_{t-1}^2)^4} - \frac{Y_{t-1}^6}{(\gamma + \alpha Y_{t-1}^2)^3}.$$

D'après le théorème 1 et les lois des grands nombres, on à :

$$\sup_{\theta \in \Theta} \sup_{\gamma \in \Gamma} \left| \frac{1}{n} \sum_{t=1}^n \frac{\partial^3 \ell_t(\gamma, \theta)}{\partial \alpha^3} \right| \leq \frac{24M^2}{\underline{\alpha}^4} + \frac{6}{\underline{\alpha}^3} \frac{\max\{\gamma_0, \alpha_0\}}{\min\{\underline{\gamma}, \underline{\alpha}\}} \sum_{t=1}^n \eta_t^2 = o_p(1).$$

De même, nous pouvons montrer que les autres dérivées partielles mixtes d'ordre trois de  $\frac{1}{n} \sum_{t=1}^n \ell_t(\gamma, \theta)$  avec  $\theta$  est  $o_p(1)$ . D'où la verifcation de l'hypothèse 2.

## 2.5 Maximum de vraisemblance du modèle DAR(P)

Soit  $\lambda = (\lambda'_1, \lambda'_2)'$ , avec:  $\lambda'_1 = (\phi_1, \dots, \phi_p)'$  et  $\lambda'_2 = (\omega, \alpha_1, \dots, \alpha_p)'$ . Supposons que  $\{y_{-p}, \dots, y_n\}$  sont généré par le modèle (2.1) avec le paramètre  $\lambda_0 = (\lambda'_{10}, \lambda'_{20})'$ . La fonction conditionnelle de log-vraisemblance peut s'écrire comme suit:

$$L_n(\lambda) = \sum_{t=2}^n l_t(\lambda), \text{ avec } l_t = -\frac{1}{2} \ln(\lambda'_2 y_{2t-1}) - \frac{\epsilon_t^2(\lambda)}{2(\lambda'_2 y_{2t-1})} \quad (2.20)$$

où  $\epsilon_t(\lambda) = y_t - \lambda'_1 y_{1t-1}, Y_{1t-1} = (y_t, \dots, y_{t-p+1})'$  et  $Y_{2t} = (1, y_t^2, \dots, y_{t-p+1}^2)'$ . Soit  $\Theta$  l'espace des paramètres. Trouver l'estimateur de maximum du vraisemblance de  $\lambda_0$  noté par  $\tilde{\lambda}_n$  revient à maximiser la fonction  $L_n(\lambda)$  sur  $\Theta$ . Soit  $\bar{\omega}, \underline{\omega}, \underline{\alpha}$  et  $\bar{\alpha}$  des constantes positives.

**Hypothèse 1.** *On suppose que  $\Theta$  est compact, avec  $\underline{\omega} < \omega < \bar{\omega}$  et  $\underline{\alpha} < \alpha < \bar{\alpha}, (i = 1, \dots, p)$ . et  $\lambda_0$  est un point intérieure de  $\Theta$  et  $\gamma < 0$  pour toutes  $\lambda \in \Theta$ .*

Le théorème suivant montre les propriété asymptotiques de l'estimateur du maximum de vraisemblance.

**Théorème 2.2.** *Supposons que la solution  $\{y_t = t \in \mathbb{N}\}$  est strictement stationnaire et ergodique du modèle (2.1). Si l'hypothèse (1) est vérifie, Alors*

$$\sqrt{n}(\hat{\lambda}_n - \lambda_0) \rightarrow_{\varphi} N(0, \Omega^{-1}),$$

quand  $n \rightarrow \infty$ , où  $\rightarrow_{\varphi}$  est la convergence en lois, et

$$\Omega = \text{diag} \mathbb{E} \left( \frac{y_{1t} y'_{1t}}{(\lambda'_{20} y_{2t})} \right), \frac{1}{2} \mathbb{E} \left[ \frac{y_{2t} y'_{2t}}{(\lambda' y_{2t})^2} \right].$$

**Remarque 2.3.** Quand  $\eta_t$  est normal, l'équation (2.20) est toujours vraie. Dans ce cas  $\hat{\lambda}_n$  et l'unique estimateur du maximum du vraisemblance (EMV) de  $\lambda_0$ , et le théorème (2.2) s'applique en remplaçons  $\Omega^{-1}$  par  $\Omega^{-1} \sigma \Omega^{-1}$  si  $\mathbb{E} \eta_t^4 < \infty$  pour tous  $j > 0$ , alors:

$$\Sigma = \mathbb{E} \left( \begin{pmatrix} \frac{y_{1t}}{\sqrt{\lambda'_{20} y_{2t}}} & \frac{y_{2t}}{\lambda'_{20}} \end{pmatrix} J \left( \begin{pmatrix} \frac{y'_{1t}}{\sqrt{\lambda'_{20} y_{2t}}} \\ \frac{y_{2t}}{\lambda'_{20} y_{2t}} \end{pmatrix} \right) \right)$$

et

$$J = \begin{pmatrix} 1 & \mathbb{E}(\eta_t^3) \\ \mathbb{E}(\eta_t^3) & \mathbb{E}(\eta_t^4) - 1 \end{pmatrix}.$$

la preuve de cette remarque et comme la preuve de théorème (2.2) donnée précédemment, à condition de remplacer  $\Omega$  par  $\Sigma$ .

**Remarque 2.4.** La condition  $\underline{\alpha} > 0$  dans l'hypothèse (1) est utiliser car  $\omega + \sum_{i=1}^p \alpha_i y_{t-i}^2$  vérifie la fonction log de vraisemblance, et la matrice d'information, de telle sorte qu'ils sont délimités, sachant que, sans cette condition le théorème (2.2) peut pas être vérifie. L'estimateur de maximum de vraisemblance, est considéré comme un estimateur optimale au sens de Lecam, et son normalité asymptotique implique que l'équation (2.20) de log-maximum de vraisemblance est localement asymptotiquement normal. Cependant, le théorème (2.2) montre que l'EMV de  $\lambda_{10}$  peut être asymptotiquement normal dans le cas où  $\phi(z) = 0$  à l'intérieur de cercle unité, et aussi dans le cas où  $\mathbb{E}(y_t^2) = 0$ . Il est évident

que tout types d'estimateur  $\phi_i$  dans ces cases sont pas asymptotiquement normal dans les modèles classique AR(p) avec des erreurs (iid) (voire Dickey et Fuller(1979), Chan et Tran(1989), Davis, Knight et Liu(1992)).

## 2.6 Etude Par simulation

La méthode de Monte Carlo est une méthode numérique de simulation qui rend les décisions plus faciles, et qui utilise des tirages aléatoires pour réaliser le calcul d'une quantité déterministe. L'objectif de cette partie est de quantifier les performances de l'estimateur de maximum de vraisemblance. Nous utilisons pour cela le langage de programmation R. Rappelons que Ce langage est développé initialement au début des années (1990) par Ross Ihaka et Robert Gentleman, ce langage a pour but de pouvoir traiter et organiser des jeux de données afin de pouvoir y appliquer des tests statistiques, et représenter ces données graphiquement à l'aide d'une grande variété de graphiques disponibles. Ce langage est très utilisé pour diverses raisons:

- Il est entièrement gratuit et sous licence GPL, ce qui signifie que vous pouvez en télécharger les sources et les modifier. C'est notamment grâce à cela qu'une active communauté peut améliorer le langage en permanence.
- Il permet d'assez facilement créer des graphiques paramétrables afin de pouvoir mieux visualiser le résultat de vos analyses.
- Il est multi-plateforme. Vous pourrez en effet l'utiliser sous Windows, Mac OS ou Linux.
- Il permet d'organiser et traiter des volumes importants de données de manière rapide et flexible. Comme il s'agit d'un langage de programmation, vous pouvez (à peu près) en faire tout ce que nous voulons.

A base de simulation de plusieurs échantillons (des réalisations) de taille  $n$ , avec  $n = \{100,200,1001\}$  du modèles DAR(p). Nous allons calculer pour chaque échantillons l'estimateur de maximum de vraisemblance. En faisant varier les paramètres  $(\phi_0, \alpha_0)$  dans l'ensemble

$$E = \{(1,3); (1,4); (2,1); (2,2); (-2,3); (-2,4)\},$$

nous calculons à chaque itération les quantités:

1. EM: La moyenne empirique des estimations des paramètres.
2. ESD: l'écart type empirique des estimations des paramètres.

3. ASD: l'écart type asymptotique des estimations des paramètres.

Quant aux diverses lois des  $\eta_t$  retenues pour effectuer ces simulations sont les suivantes :

1.  $U[-\sqrt{3}, \sqrt{3}]$  est la loi uniforme sur l'intervalle  $[-\sqrt{3}, \sqrt{3}]$  avec  $n = 1001$
2. La distribution exponentiel double ( $\text{Dexp}(\frac{1}{\sqrt{2}})$ ) avec la densité  $f(x) = \frac{1}{\sqrt{2}} \exp(-\sqrt{2} |x|)$ ,  $x \in \mathbb{R}$ .
3. La distribution de student standard  $t_6(\sqrt{\frac{2}{3}}t_6)$  avec la densité  $f(x) = \frac{15}{32}(1+\frac{x^2}{4})^{-\frac{7}{2}}$ ,  $x \in \mathbb{R}$ .
4. La distribution de Radmacher, défini par  $\mathbb{P}(\eta = \pm 1) = \frac{1}{2}$
5. La distribution FPDD.

Les tableaux suivants illustrent les résultat du simulation des estimateurs  $\tilde{\gamma}_n, \tilde{\phi}_n, \tilde{\alpha}_n$  quand  $(\phi_0, \alpha_0)$  prennent leurs valeurs dans X.

$\phi_0$	$\alpha_0$		$\tilde{\gamma}_n$	$\tilde{\phi}_n$	$\tilde{\alpha}_n$	$\tilde{\gamma}_n$	$\tilde{\phi}_n$	$\tilde{\alpha}_n$
1.0	3.0	EM	1.2504	1.0091	2.9554	1.2686	0.9965	2.9792
		ESD	0.9774	0.1812	0.2932	1.0514	0.1251	0.1918
		ASD	-	0.1731	0.2684	-	0.1224	0.1896
1.0	4.0	EM	1.2515	1.0091	3.9476	1.2341	0.9994	3.9832
		ESD	0.8788	0.2023	0.3870	0.9287	0.1465	0.2592
		ASD	-	0.2001	0/3577	-	0.1415	0.2531
2.0	1.0	EM	1.1921	1.9981	0.9844	1.1412	1.9992	0.9912
		ESD	0.7408	0.1028	0.0934	0.7241	0.0724	0.0631
		ASD	-	0.1000	0.0893	-	0.0706	0.0631
2.0	2.0	EM	1.1165	2.0022	1.9748	1.1117	1.9981	1.9902
		ESD	0.6721	0.1501	0.1868	0.6642	0.1053	0.1323
		ASD	-	0.1413	0.1788	-	0.1001	0.1264
-2.0	3.0	EM	1.1644	-2.0136	2.9802	1.1770	-2.0013	2.9876
		ESD	0.9341	0.1731	0.2710	0.8252	0.1182	0.1948
		ASD	-	0.1732	0.2682	-	0.1224	0.1898
-2.0	4.0	EM	1.1948	-2.01423	3.9632	1.1586	-1.9974	3.9736
		ASD	0.9148	0.2123	0.3592	0.8381	0.1488	0.2515
		ASD	-	0.2001	0.3577	-	0.1415	0.2531

Tableau1: La simulation quand  $\eta$  suit une lois  $U[-\sqrt{3}, \sqrt{3}]$  avec n=1001.

$\phi_0$	$\alpha_0$		$\tilde{\gamma}_n$	$\tilde{\phi}_n$	$\tilde{\alpha}_n$	$\tilde{\gamma}_n$	$\tilde{\phi}_n$	$\tilde{\alpha}_n$
1.0	4.0	EM	2.4801	1.0076	3.9617	2.3242	0.9987	3.9692
		ESD	9.1837	0.2153	0.9664	12.573	0.1423	0.8737
		ASD	-	0.2001	0.8943	-	0.1414	0.6324
1.0	6.0	EM	5.3165	0.9911	5.8732	7.2251	1.0041	5.9427
		ESD	67.394	0.2578	1.4067	138.21	0.1764	0.19687
		ASD	-	0.2448	1.3415	-	0.1731	0.9486
2.0	2.0	EM	1453.5	2.0012	2.0151	2.0401	2.0043	1.9762
		ESD	45875	0.1482	0.4604	6.5247	0.01012	0.3068
		ASD	-	0.1415	0.4471	-	0.1001	0.3162
2.0	4.0	EM	3.4803	1.9996	3.9671	2.1051	1.9986	3.9426
		ESD	25.506	0.1995	0.9271	7.2692	0.1426	0.5914
		ASD	-	0.2000	0.8944	-	0.1414	0.6324
-1.0	6.0	EM	4.6524	-1.0095	5.8352	4.4432	-1.0000	5.9190
		ESD	46.197	0.2408	1.3418	28.841	0.1720	0.9289
		ASD	-	0.2448	1.3416	-	0.1732	0.9487
-2.0	4.0	EM	2.8904	-2.0038	3.9575	8.1027	-2.0038	3.9852
		ESD	16.776	0.2013	0.9154	178.65	0.1398	0.6502
		ASD	-	0.2000	0.8944	-	0.1414	0.6325

Tableau2:  $\eta \sim \text{Dexp}(\frac{1}{\sqrt{2}})$ .

$\phi_0$	$\alpha_0$		$\tilde{\gamma}_n$	$\tilde{\phi}_n$	$\tilde{\alpha}_n$	$\tilde{\gamma}_n$	$\tilde{\phi}_n$	$\tilde{\alpha}_n$
1.0	5.0	EM	3.6877	1.0077	4.9581	2.6543	1.0012	4.9886
		ESD	35.719	0.2338	1.2439	14.339	0.1649	0.8823
		ASD	-	0.2235	1.1181	-	0.1582	0.7906
1.0	6.0	EM	3.7139	1.0092	5.9564	4.1232	1.0008	5.9121
		ESD	36.719	0.1791	0.9822	41.095	0.1739	0.9106
		ASD	-	0.2448	1.3415	-	0.1732	0.9487
2.0	2.0	EM	2.7431	2.0048	1.9667	2.7162	2.0029	1.9673
		ESD	16.908	0.1402	0.4351	16.233	0.0982	0.3110
		ASD	-	0.1414	0.4471	-	0.1000	0.3163
2.0	4.0	EM	36.942	2.0001	3.9605	2.6050	2.0013	4.0022
		ESD	1082.2	0.2008	0.8896	13.976	0.1393	0.6325
		ASD	-	0.2000	0.8944	-	0.1415	0.6325
-2.0	7.0	EM	2.8831	-2.0092	6.9522	6.6975	-2.0086	6.9396
		ESD	24.584	0.2677	1.6144	94.661	0.1938	1.1227
		ASD	-	0.2645	1.5651	-	0.1872	1.1067
-2.0	5.0	EM	8.7822	-2.0042	4.9296	6.0256	-1.9935	4.9772
		ESD	212.82	0.2314	1.1188	97.152	0.1578	0.7617
		ASD	-	0.2235	1.1181	-	0.1582	0.7905

Tableau3:  $\eta \rightarrow$  student's standard  $t_6$

$\phi_0$	$\alpha_0$		$\tilde{\gamma}_n$	$\tilde{\phi}_n$	$\tilde{\alpha}_n$	$\tilde{\gamma}_n$	$\tilde{\phi}_n$	$\tilde{\alpha}_n$
1.0	4.0	EM	1.0074	0.9973	3.9544	0.9993	0.9971	3.9808
		ESD	0.0502	0.2141	0.0627	0.0282	0.1383	0.0262
		ASD	-	0.2001	0.0000	-	0.1415	0.000
2.0	1.0	EM	1.0024	2.0063	0.9886	0.9967	1.9992	0.9947
		ESD	0.0503	0.1055	0.0152	0.0358	0.0711	0.0071
		ASD	-	0.1000	0.0000	-	0.0706	0.0000
-1.0	5.0	EM	1.0025	-0.9971	4.9477	1.0015	-1.0087	4.9537
		ESD	0.0366	0.2285	0.0671	0.0251	0.1616	0.0355
		ASD	-	0.2235	0.0000	-	0.1582	0.0000
-2.0	1.0	EM	1.0038	-1.9982	0.9894	0.9982	-2.0005	0.9948
		ESD	0.0502	0.01028	0.0156	0.0357	0.0700	0.0070
		ASD	-	0.01001	0.0000	-	0.0708	0.0000

Tableau4:  $\eta$  suit la distribution Radmacher.

$\phi_0$	$\alpha_0$		$\tilde{\gamma}_n$	$\tilde{\phi}_n$	$\tilde{\alpha}_n$	$\tilde{\gamma}_n$	$\tilde{\phi}_n$	$\tilde{\alpha}_n$
4.0	3.0	EM	1.0843	3.9947	2.9793	1.0621	3.9975	2.9812
		ESD	0.5952	0.1782	0.1836	0.5971	0.1187	0.1294
		ASD	-	0.1731	0.1801	-	0.1224	0.1272
2.0	5.0	EM	1.0691	1.9842	4.9337	1.0501	1.9957	4.9782
		ESD	0.5722	0.2182	0.3141	0.5672	0.1592	0.2192
		ASD	-	0.2235	0.3000	-	0.1582	0.2122
0.0	3.0	EM	1.0567	-0.0002	2.9636	1.0474	-0.0042	2.9814
		ESD	0.5941	0.1779	0.1904	0.5923	0.1211	0.1267
		ASD	-	0.1731	0.1800	-	0.1224	0.1272
-2.0	4.0	EM	1.0556	-1.9942	3.9532	1.0511	-1.9880	3.9803
		ESD	0.5614	0.2071	0.2551	0.5504	0.1451	0.1714
		ASD	-	0.2000	0.2401	-	0.1415	0.1698
-4.0	2.0	EM	1.0314	-3.9937	1.9767	1.0845	-3.9977	1.9904
		ESD	0.5917	0.1434	0.1293	0.5753	0.1013	0.0848
		ASD	-	0.1415	0.1201	-	0.1000	0.0848

Tableau4:  $\eta$  suit la loi de (FPDD).

## 2.6.1 Interpretation

Il résulte des différents tableaux que les biais des estimateurs  $(\phi_0, \alpha_0)$  sont très petits, et les Différences entre les ESD et ASD sont aussi très petites. Ainsi, à la lumière de ces tableaux on constate que plus  $n$  augmente, plus ESD diminue. Pour un échantillon de taille  $n$  fixé, plus  $\alpha$  augmente plus le ESD et le ASD augmentent. Tous les résultats confirment que l'ESD est performant, ce qui confirme le bien fondé de la théorie. Toutefois l'estimateur de maximum de vraisemblance de  $\gamma_0$  a un biais élevé, particulièrement pour  $\eta \rightarrow D \exp(\frac{1}{\sqrt{2}})$ .

# Conclusion

Dans l'introduction générale, il a été expliqué que l'analyse des séries temporelles est bien plus qu'une "simple" estimation de la valeur future d'une série temporelle. Ce mémoire a été consacré à l'étude théorique des séries temporelles. Nous avons traité dans le premier chapitre la classe des processus "ARMA", les modèles GARCH et les ARMA-GARCH, dans le deuxième chapitre, nous avons étudié le modèles autorégressives double. L'étude de la stationnarité a été intensivement relaté au niveau du deux chapitres où nous avons mis en évidence les diverse conditions nécessaires et/ou suffisantes assurant la faible et la forte stationnarité et l'existence et l'unicité de ces solutions stationnaires. Comme perspectives futures, nous envisageons l'extension de l'étude de l'estimateur du maximum de vraisemblance des estimateurs des paramètres des modèles DAR(p) aux cas de données manquantes et aberrantes.

# Annexe

**Définition 2.1 (Hétéroscédasticité).** En statistique, on parle d'hétéroscédasticité lorsque les variances des variables examinées sont différentes. Cette notion est composée du préfixe hétéro ("autre"), et de skedasê ("dissipation"). Une collection de variables aléatoires est hétéroscédastique, s'il y a des sous-populations qui ont des variabilités différentes des autres. La notion d'hétéroscédasticité s'oppose à celle d'homoscédasticité, qui correspond au cas où la variance de l'erreur des variables est constante. Tandis que dans le cas d'homoscédasticité, nous avons  $var(\epsilon_i) = \sigma^2$ ,  $\forall i$  nous avons désormais  $var(\epsilon_i) = \sigma_i^2$ , et  $var(\epsilon_j) = \sigma_j^2$  où  $\sigma_i^2$  est différente de  $\sigma_j^2$  pour  $i \neq j$

**Définition 2.2 (Autorégressivité).** On dira que le processus  $\{X_t, t \in \mathbb{Z}\}$  suit un modèle autorégressif d'ordre  $p$  (AR( $p$ )) s'il existe un bruit blanc centré réduit  $(0, \sigma^2)$ , noté  $\{\epsilon_t\}$ , tel que  $\{\epsilon_t\}$ , soit indépendant de  $\{X_0, \dots, X_{t-1}\}$  et des constantes  $a_1, a_2, \dots, a_p$  et  $\sigma$  tels que pour  $t \in \{1, \dots, n\}$  on a :

$$X_t - \mu = a_1(X_{t-1} - \mu) + a_2(X_{t-2} - \mu) + \dots + a_p(X_{t-p} - \mu) + \sigma\epsilon_t$$

**Définition 2.3 (convergence presque sûrement).** Une suite  $\{X_t\}$  de variables aléatoires converge presque sûrement vers  $X$ , si

$$P(\omega \in \Omega / \lim_{n \rightarrow \infty} X(\omega) = x(\omega)) = 1.$$

Et on note  $X_n \xrightarrow{p.s.} X$ .

**Définition 2.4 (Ergodicité).** Une suite stationnaire  $\{X_t, t \in \mathbb{Z}\}$  est dite ergodique si elle vérifie la loi forte des grands nombres.

# Bibliographie

- [1] Shiqing Ling. A DOUBLE AR(P) MODÈL: STRUCTURE AND ESTIMATION
- [2] Men Chen Dog Li et Shiqing Ling 2013. Non-stationarity eand Quasi-Maximum Likelihood Estimation on a Double Autoregressive Model. Journal Of Time Series Analysis J. Time Ser. Anal. 35: 189-202(2014).
- [3] C.Francq, J.Zakoian, (2010) GARCH MODELS, Ed.Wiley.
- [4] C.Francq, J.Zakoian, (2009) Modèles GARCH et à volatilité stochastique, Université de Lille.
- [5] Christophe Hurlin, (2006-2007), Modèles ARCH - GARCH Applications à la VaR, Université d'Orléans.
- [6] Bose, A. and Mukherjee, K. (2003) Estimating the ARCH parameters by solving linear equations. Journal of Time Series Analysis 24, 127–136.
- [7] Christophe Hurlin (2004) Econométrie pour la Finance, Modèles ARCH - GARCH.
- [8] Francq , C. and Zakoian, J.-M. (2008) Can one really estimate nonstationary GARCH models?
- [9] Francq , C. and Zakoian, J.-M. (2010) GARCH models : Structure, Statistical Inference and Financial Applications, 1st edition. WILEY.
- [10] Jensen, S.T. and Rahbek, A. (2004a) Asymptotic normality of the QMLE estimator of ARCH in the nonstationary case. Econometrica 72, 641–646.
- [11] Zakoian, J.-M. (1992) Modèles ARCH : une revue de la littérature. Journal de la société statistique de Paris, tome 133, n° 1-2(1992), p.40-57.
- [12] Peter J. Brockwell Richard A. Davis, (2006) Time Series: Theory and Methods , Springer Science, Business Media.